

Licença



Este trabalho está licenciado sob uma licença [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/). Fonte: <https://livros.unb.br/index.php/portal/catalog/book/668>. Acesso em: 24 out. 2025.

Referência

MENESES, Paulo Roberto; ALMEIDA, Tati de; SANO, Edson Eyji. (ed.). **Processamento de imagens de sensoriamento remoto**. Brasília: Editora Universidade de Brasília, 2025. E-book (420 p.). Disponível em: <https://livros.unb.br/index.php/portal/catalog/book/668>. Acesso em: 24 out. 2025.

PROCESSAMENTO DE IMAGENS DE **SENSORIAMENTO REMOTO**

*Paulo Roberto Meneses
Tati de Almeida
Edson Eyji Sano*
Editores



EDITORA



UnB

Embrapa

**Ministério da Educação
Universidade de Brasília
Ministério da Agricultura e Pecuária
Embrapa Cerrados**

PROCESSAMENTO DE IMAGENS DE
**SENSORIAMENTO
REMOTO**

*Paulo Roberto Meneses
Tati de Almeida
Edson Eyji Sano*
Editores

*Brasília, DF
2025*



Editora Universidade de Brasília

Centro de Vivência, Bloco A – 2ª etapa, 1º andar
Campus Darcy Ribeiro, Asa Norte, Brasília/DF
CEP: 70910-900
(61) 3107-3700
www.editora.unb.br
contatoeditora@unb.br

Reitora

Rozana Reigota Naves

Vice-Reitor

Márcio Muniz

Diretora da Editora UnB

Marina Dourado Lustosa Cunha

Conselho editorial

Presidente

Marina Dourado Lustosa Cunha

Membros

Maria do Socorro Neri de Sousa

Andrey Rosenthal Schlee

Biagio D'Angelo

Carina Luísa Ochi Flexor

Guilherme Sales Soares de Azevedo Melo

Josivânia Silva Farias

Maria Carmem Aires Gomes

Maria Filomena Pinto da Costa Coelho

Nilce Santos de Melo

Rosineide Magalhães de Sousa

Embrapa Cerrados

BR 020, Km 18, Rod. Brasília-Fortaleza
Caixa Postal 08223
73310-970, Planaltina, DF
www.embrapa.br/cerrados
www.embrapa.br/fale-conosco/sac

Comitê Local de Publicações

Presidente

Eduardo Alano Vieira

Secretária-executiva

Lidiamar Barbosa de Albuquerque

Membros

Alessandra de Jesus Boari, Alessandra Silva Gelape

Faleiro, Angelo Aparecido Barbosa Sussel,

Fábio Gelape Faleiro, Giuliano Marchi, Jussara Flores

de Oliveira Arbues, Karina Pulrolnik, Maria Emilia

Borges Alves e Natália Bortoleto Athayde Maciel

Edição executiva

Jussara Flores de Oliveira Arbues

Revisão de texto

Margit Bergener L. Guimarães

Normalização bibliográfica

Marilaine Schaum Pelufe (CRB-1/2045)

Projeto gráfico, diagramação e capa

Wellington Cavalcanti

1ª edição

Publicação digital (2025): PDF

Dados Internacionais de Catalogação na Publicação (CIP)

(Biblioteca Central da Universidade de Brasília – BCE/UNB)

P963 Processamento de imagens de sensoriamento remoto [recurso eletrônico] / Paulo Roberto Meneses, Tati de Almeida, Edson Eyji Sano, editores. – Brasília : Editora Universidade de Brasília, 2025.
420 p.

Inclui bibliografia.

Formato PDF.

ISBN 978-65-5846-227-9.

1. Sensoriamento remoto. 2. Processamento de imagens – Técnicas digitais. I. Meneses, Paulo Roberto (ed.). II. Almeida, Tati de (ed.). III. Sano, Edson Eyji (ed.).

CDU 528.8

Heloiza dos Santos – CRB 1/1913

Autores

Edson Eyji Sano

Geólogo, doutor em Ciência do Solo, pesquisador da Embrapa Cerrados, Planaltina, DF

Giuliano Sant'anna Marotta

Engenheiro Agrimensor, doutor em Geociências Aplicadas, professor da Universidade de Brasília, Brasília, DF

Gustavo Macedo de Mello Baptista

Geógrafo, doutor em Geologia, professor associado da Universidade de Brasília, Brasília, DF

Paulo Roberto Meneses

Geólogo, doutor em Geologia, professor titular aposentado da Universidade de Brasília, Brasília, DF

Rejane Ennes Cicerelli

Engenheira Cartógrafa, doutora em Ciências Cartográficas, professora adjunta da Universidade de Brasília, Brasília, DF

Tati de Almeida

Geóloga, doutora em Geologia, professora associada da Universidade de Brasília, Brasília, DF



Nota dos Editores

A observação sistemática do planeta Terra por sensores imageadores orbitais completou 50 anos em 2022! Esses valiosos dados têm permitido acompanhar as mudanças na superfície terrestre e monitorar o meio ambiente em escalas global e temporal. Nas últimas décadas, o sucesso dessa tecnologia fez com que a utilização transcendesse o mundo acadêmico, tornando-se popular até mesmo nos celulares utilizados em todo o mundo.

A multidisciplinaridade do tema é latente e não há barreiras para as possibilidades de utilização dessa tecnologia. No entanto, no caminho entre a geração de informações e a sua utilização pelo público em geral, está a necessidade de tratamento dos dados coletados por especialistas para que esses dados sejam traduzidos em informações. Hoje, os programas para processamento de imagens de sensoriamento remoto não exigem habilidade em computação por parte dos especialistas. Porém, o entendimento da formulação matemática ou estatística dos algoritmos envolvidos em vários métodos de processamento de imagens continua importante para que possamos extrair melhor todo o conteúdo de informação contida nas imagens, o que nem sempre é facilmente percebido pelo usuário.

O livro “Processamento de Imagens de Sensoriamento Remoto” apresenta técnicas utilizadas pelos analistas das mais diferentes áreas de conhecimento. Trata-se de um livro interdisciplinar que, através de linguagem simples e direta, aborda conceitos e aplicações da utilização de imagens para a extração da informação, transformando o dado de sensoriamento

remoto em informação. A linguagem do livro é direcionada ao ensino do sensoriamento remoto para o público em geral, aos estudantes de graduação e àqueles que estão se iniciando no sensoriamento remoto, qualquer que seja a área de formação. Por isso, os autores se preocuparam em demonstrar como o operador de uma técnica de processamento transforma a imagem para realçar a sua visualização na tela do computador ou como se ordena e se classifica os dados para construir um mapa digital temático, com o mínimo necessário da compreensão das funções matemáticas e estatísticas.

Apesar de o livro enfatizar o processamento de imagens, os quatro primeiros capítulos do livro apresentam uma síntese dos conceitos básicos de sensoriamento remoto e de sensores imageadores. Os autores entendem que, para a extração da informação, é necessário entender minimamente sobre conceitos dos processos físicos e fotoquímicos e a forma de obtenção dos dados por sensores imageadores para o sucesso do processo.

Esse *e-book*, disponibilizado em formato pdf de publicação e de forma gratuita, pretende possibilitar aos estudantes o livre acesso para o acompanhamento de suas aulas nas disciplinas de processamento de imagens de sensoriamento remoto. Esperamos que este livro possa inspirar o interesse e o entusiasmo desses estudantes! Aproveitem!

Os Editores

Prefácio

O lançamento de um livro técnico-científico em língua portuguesa é sempre um acontecimento digno de destaque, notadamente em um país com uma crônica carência de literatura dessa natureza. Esse destaque torna-se ainda mais relevante quando se trata de um livro de temática transdisciplinar e de amplo espectro de aplicações, como é o caso desta edição do “Processamento de Imagens de Sensoriamento Remoto”.

É impossível imaginar hoje qualquer atividade de monitoramento/mapeamento de processos naturais ou antrópicos em nosso planeta que não seja fortemente apoiado em informações extraídas de imagens de sensoriamento remoto obtidas por plataformas aéreas e orbitais. O mesmo se passa com relação à exploração espacial, na qual o sensoriamento remoto é a fonte primordial de informações essenciais ao avanço do conhecimento sobre todos os corpos do Sistema Solar.

O processamento digital é a forma pela qual é extraída e decodificada a extraordinária quantidade de informações presentes nas imagens da Terra e dos demais corpos planetários registradas pelos sensores remotos. Assim, é de fundamental importância conhecer e dominar o amplo conjunto de técnicas empregadas na desafiadora tarefa de tornar essas informações passíveis de uso nas múltiplas aplicações do sensoriamento remoto.

É essa a proposta desta obra, que detalha os principais conceitos, princípios e métodos de extração de informações das imagens de sensoriamen-

to remoto de vários tipos. Com uma abordagem de elevado rigor teórico, mas com uma linguagem de fácil entendimento, o livro aborda os princípios do sensoriamento remoto e as principais técnicas de processamento digital, desde seus fundamentos teóricos, passando por aspectos práticos e finalizando com a apresentação de exemplos didáticos de suas aplicações.

Uma obra com essas características só pode se concretizar quando organizada e escrita por pesquisadores com extensa experiência e vivência nos temas em foco. É o que ocorre com este livro, elaborado por profissionais de amplo conhecimento e reconhecida competência em sensoriamento remoto e em processamento digital de imagens.

Disponibilizada na forma de e-book de acesso aberto, o que amplia de maneira significativa o seu alcance, este livro será de fundamental importância no treinamento e aperfeiçoamento das próximas gerações de especialistas em sensoriamento remoto, tanto no Brasil como também em outros países da comunidade internacional de língua portuguesa.

Por todas essas qualidades, é louvável a iniciativa de lançamento desta obra por esse distinto grupo de pesquisadores da Universidade de Brasília e da Embrapa Cerrados.

Álvaro Penteado Crósta
Professor titular aposentado
Universidade Estadual de Campinas (Unicamp)

Sumário

CAPÍTULO 1 – Princípios de Sensoriamento Remoto	14
1.1. Breve História do Sensoriamento Remoto.....	15
1.2. Definição de Sensoriamento Remoto	22
1.3. Natureza da Radiação Eletromagnética	23
1.4. Fontes de Radiação Eletromagnética	30
1.5. Espectro Eletromagnético	35
1.6. Síntese das Características mais Importantes dos Intervalos Espectrais..	39
CAPÍTULO 2 – Modelos de Interação da Radiação Eletromagnética	42
2.1. Interação da Radiação Eletromagnética: Fotografias x Imagens	43
2.2. Processos de Interação da Luz Solar.....	46
2.3. Processos de Interação da Radiação de Micro-ondas.....	60
CAPÍTULO 3 – Sistemas Sensores Imageadores	68
3.1. Sensores Multiespectrais	70
3.2. Síntese dos Sistemas que Compõem Sensores Ópticos e Termiais.....	72
3.3. Mecanismos de Imageamento Multiespectral	73
3.4. Radiância e Reflectância.....	78
3.5. Sensores Termiais	82
3.6. Radar de Abertura Sintética	94
3.7. Referências	110

CAPÍTULO 4 – Resoluções de Imagens 112

4.1. A Importância das Resoluções das Imagens.....	113
4.2. Resolução Espacial.....	115
4.3. Resolução Espectral.....	120
4.4. Resolução Radiométrica.....	124
4.5. Resolução Temporal.....	125
4.6. Resolução Espacial de Imagens de Radar.....	128

CAPÍTULO 5 – Modelos de Dados e de Processamento..... 134

5.1. Notação dos Dados.....	135
5.2. Arquivos Raster.....	138
5.3. Formatos de Salvamento.....	139
5.4. Modelo de Processamento.....	139
5.5. Formato de Arquivos Rasterizados.....	143
5.6. Referência.....	144

CAPÍTULO 6 – Correção Atmosférica de Imagens..... 146

6.1. O Efeito da Atmosfera sobre a Radiação Eletromagnética (REM).....	147
6.2. Modelos de Transferência Radiativa.....	158
6.3. Referências.....	164

CAPÍTULO 7 – Correção Geométrica de Imagens 166

7.1. Distorções Geométricas.....	167
7.2. Georreferenciamento de Imagens.....	182
7.3. Referências.....	199

CAPÍTULO 8 – Ampliação de Contraste 200

8.1. Representação Histograma.....	201
8.2. Expansão por Saturação.....	206
8.3. Expansão por Equalização.....	211
8.4. Expansão Gaussiana.....	216
8.5. Função Raiz Quadrada.....	219

CAPÍTULO 9 – Sistemas de Cores.....	222
9.1. Cor.....	223
9.2. Fatiamento de Cores.....	227
9.3. Espaço de Cores RGB.....	228
9.4. Espaço de Cor IHS.....	236
9.5. Referências.....	244
 CAPÍTULO 10 – Fusão de Imagens.....	 246
10.1. IHS.....	247
10.2. Modelos Algébricos.....	250
10.3. Modelos no Domínio Espectral.....	251
10.4. Referências.....	254
 CAPÍTULO 11 – Aritmética de Bandas.....	 256
11.1. Operações Aritméticas.....	257
11.2. Índice de Vegetação.....	272
11.3. Índice de Áreas Construídas.....	279
11.4. Índice de Água.....	281
11.5. Índice de Queimadas.....	284
11.6. Índice de Minerais e de Rochas.....	287
11.7. Referências.....	290
 CAPÍTULO 12 – Rotação Espectral.....	 292
12.1. Transformação por Componentes Principais.....	293
12.2. Descorrelação de Bandas.....	311
12.3. Referências.....	316
 CAPÍTULO 13 – Filtragem Espacial.....	 318
13.1. Modelos de Filtragem.....	319
13.2. Filtros Passa-Alta e Passa-Baixa.....	325
13.3. Filtros de Mediana.....	331

13.4. Filtros de Bordas Direcionais.....	332
13.5. Operadores Laplacianos.....	337
13.6. Filtros Gaussianos.....	338
13.7. Filtros Morfológicos.....	339
13.8. Filtros de Textura.....	341
13.9. Filtragens no Domínio da Frequência.....	345
13.10. Referências.....	353

CAPÍTULO 14 – Classificação de Imagens Pixel-a-Pixel 354

14.1. Modelo Básico de Classificação.....	355
14.2. Fluxo do Processo de Classificação e a Noção de Similaridade.....	362
14.3. Classificação Não Supervisionada.....	366
14.4. Classificação Supervisionada.....	371
14.5. Validação da Classificação.....	386
14.6. Árvore de Decisão.....	388
14.7. Redes Neurais Artificiais.....	390
14.8. Referências.....	394

CAPÍTULO 15 – Processamento de Dados Hiperespectrais 396

15.1. O Conceito de Espectrometria de Imageamento.....	397
15.2. Correção Atmosférica dos Dados.....	399
15.3. Bibliotecas Espectrais.....	404
15.4. Identificação de <i>Endmembers</i>	408
15.5. Medição de Intensidade das Feições Espectrais.....	412
15.6. Referências.....	418

CAPÍTULO 1

Princípios de Sensoriamento Remoto

Paulo Roberto Meneses



Para diferentes propósitos, sensoriamento remoto pode ser definido como uma técnica de medição de propriedades espectrais dos objetos da superfície da Terra a partir de dados coletados na forma de fotos e imagens por aviões e satélites. Não há, dessa forma, um contato direto com os objetos de nosso interesse. O mecanismo que possibilita esse processo é totalmente dependente das propriedades físicas da radiação eletromagnética emitida por fontes naturais como o Sol e a Terra, ou por fontes artificiais, além dos mecanismos de detecção da radiação pelos sensores. O primeiro passo para entender os processos de aquisição de imagens por satélites de sensoriamento remoto é decifrar as informações contidas nas imagens, auxiliando-se, para isso, de técnicas de processamento de imagens e do conhecimento das leis físicas mais básicas que descrevem a natureza da radiação eletromagnética.

1.1. Breve História do Sensoriamento Remoto

O marco inicial do sensoriamento remoto se deu quando as primeiras fotos aéreas foram obtidas a partir de 1908, poucos anos após a invenção do avião. Foi durante a Primeira Guerra Mundial que o valor das fotografias aéreas foi reconhecido para se fazer mapas sobre áreas com pouca informação, para fins de táticas militares. Ainda hoje é uma tecnologia viável e uma importante fonte de dados por causa de sua alta resolução espacial e flexibilidade de cobertura. No Brasil, o maior projeto aerofotográfico foi realizado na metade da década de 1960, com o objetivo de iniciar um programa de mapeamento de recursos naturais em escala nacional. Esse levantamento aerofotogramétrico foi executado pela Força Aérea Americana (missão USAF AST-10) a partir de um acordo entre o Brasil e os Estados Unidos, com abrangência de cobertura de grande parte do território brasileiro. O projeto teve duração de cinco anos (1965–1969) e obteve fotos aéreas estereoscópicas na escala de 1:60.000 que se constituíram em im-

portante insumo básico para a elaboração de cartas topográficas e mapeamentos geológicos, geomorfológicos e florestais.

Outro projeto de sensoriamento de grande impacto foi o Projeto Radam (Projeto Radar na Amazônia, denominado de Projeto RadamBrasil após 1975), que operou entre 1970 e 1985 e que consistiu na cobertura de todo o território brasileiro com imagens aéreas de radar, reproduzidas por meio óptico em mosaicos analógicos na escala de 1:250.000. Com base na interpretação dessas imagens analógicas, foi realizado um amplo estudo integrado do meio físico e biótico que incluiu relatórios analíticos e mapas temáticos sobre geologia, geomorfologia, pedologia, vegetação, uso potencial da terra e capacidade de uso dos recursos naturais renováveis.

A primeira notícia que se tem do uso do termo sensoriamento remoto data do início dos anos de 1960, cunhado por Evelyn L. Pruit e colaboradores, momento em que as primeiras experiências com sensores imageadores digitais começavam a ser testadas em aviões, na expectativa de serem embarcados em satélites artificiais. Foi com essa mudança de plataforma aérea para plataforma orbital que a era moderna do sensoriamento remoto teve início. Dessa forma, o início do sensoriamento remoto liga-se ao espetacular desenvolvimento da área espacial nesses anos, ficando conhecida como a década da corrida espacial. Foi nesse período que se viu o mais rápido desenvolvimento de foguetes lançadores de satélites, possibilitando colocar, no espaço, satélites artificiais para várias finalidades. Os satélites meteorológicos foram os pioneiros. Por meio deles é que o sensoriamento remoto por satélite deu seus primeiros passos. O primeiro satélite meteorológico foi o *Television InfraRed Observation Satellite* (TIROS-1), lançado em 1960 pelos Estados Unidos, que comprovou as enormes vantagens do uso de satélites para observações climáticas em uma escala de tempo reduzido. Em algumas das imagens do TIROS-1, os analistas conseguiram identificar, em áreas com baixa cobertura de nuvens, algumas feições da superfície da Terra que, pela primeira vez, demonstraram ser possíveis de serem vistas do espaço. Com a melhoria dos sensores embarcados nos sa-

télices subsequentes da série TIROS, puderam-se realizar estudos mais detalhados de diversas feições terrestres, tais como de áreas com cobertura de neve, gelos polares, águas oceânicas, cadeias de montanhas e grandes rios, confirmando, dessa forma, que a Terra poderia ser estudada do espaço por meio de imagens.

A década de 1960 foi também a década em que o homem pôde ir ao espaço em satélites tripulados e de lá observar a Terra e tomar as primeiras fotos da superfície terrestre pelas escotilhas da nave, usando câmeras fotográficas manuais. Foram essas experiências, realizadas durante os programas espaciais tripulados das missões Mercury, Gemini e Apollo, que definitivamente asseguraram o futuro do sensoriamento remoto. A missão GT-4 do programa Gemini denominado *Gemini-Titan for a four-day mission* foi o primeiro programa fotográfico espacial com o objetivo específico de identificar feições terrestres, principalmente as de origem geológica. A visão sinóptica na escala de 1:350.000 que essas fotos mostraram foi um fato novo para os intérpretes de fotos aéreas. Suas interpretações levaram à descoberta de novas feições geológicas que não constavam nos mapas geológicos em escalas maiores que 1:250.000.

A partir do sucesso da missão GT-4, foram incluídos, no programa Apollo, experimentos controlados envolvendo a aquisição de fotografias multiespectrais a partir de um conjunto de quatro câmeras Hasselblad, munidas cada uma de filmes pancromáticos com filtros vermelho, verde e infravermelho (Figura 1.1). Foi dessa maneira que se obteve a primeira foto multiespectral da Terra. O satélite americano Corona e o satélite soviético Zenit, lançados no final dos anos 1960, foram os últimos a realizarem missões espaciais fotográficas.

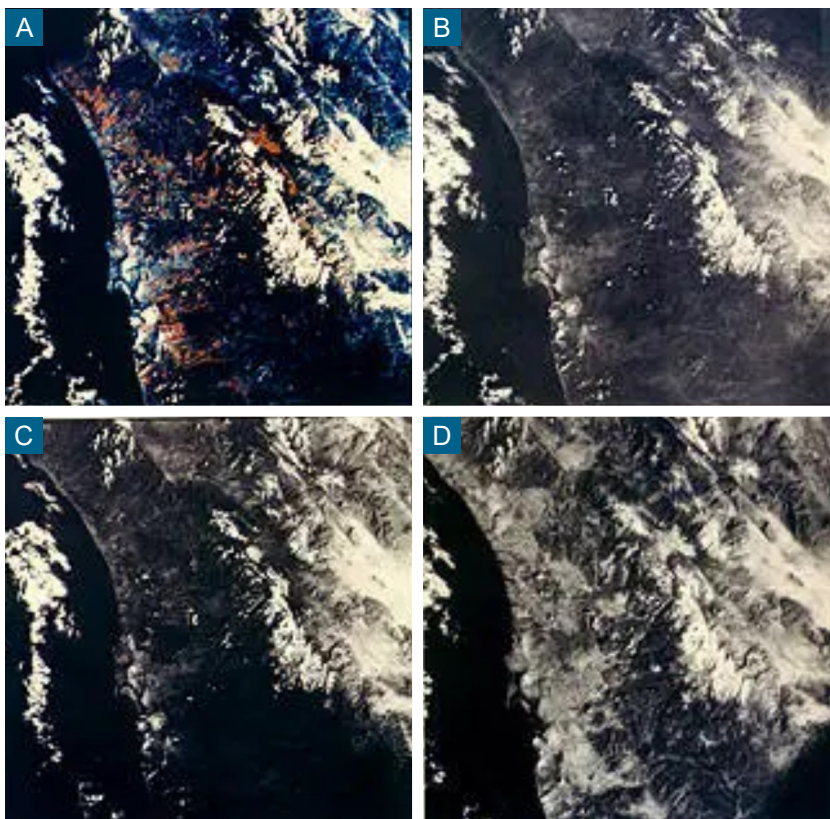


Figura 1.1. Fotos multiespectrais em falsa-cor (A) obtidas pela missão Apollo a partir de filmes pancromáticos filtrados nas cores verde (B), vermelha (C) e infravermelho (D).

Fonte: <https://www.nasa.gov/feature/50-years-ago-apollo-9-completes-its-mission>.

Os excelentes resultados desses programas espaciais impulsionaram, da metade da década de 1960 ao final da mesma década, as pesquisas para a construção de sensores ópticos eletromecânicos, uma nova tecnologia com capacidade de visualizar a Terra a partir do espaço remoto. Tais sensores receberam o nome de sensores imageadores em razão do processo de cobertura do terreno ser realizado na forma de varredura linear

do terreno, ao invés de um mecanismo de tomada instantânea de área em quadro como é feito com câmeras fotográficas. Por isso, os produtos gerados por esses novos sensores receberam a denominação de imagens, a fim de diferenciá-los, pelo seu processo de aquisição, das tradicionais fotos aéreas. Outra importante diferença dos sensores imageadores é a capacidade de obterem imagens simultâneas em várias faixas do espectro eletromagnético.

A despeito das diferenças entre as imagens de satélite e as fotos aéreas, principalmente devido aos mecanismos de aquisição, pode-se afirmar que ambas são produtos que se parecem por possuírem a mesma finalidade, que é a de retratar os objetos ou fenômenos da superfície terrestre. As vantagens de uma sobre a outra são várias, mas apenas para justificar o que impulsionou o uso dos sensores imageadores por satélites, pode-se citar a sua capacidade de imagear, em curto espaço de tempo, toda a superfície do planeta de uma maneira sistemática, dado que um satélite fica continuamente orbitando a Terra. Essa forma de cobertura repetitiva, por meio da obtenção de imagens periódicas de qualquer área do planeta, propiciou a detecção e o monitoramento das mudanças que acontecem na superfície terrestre. Essa é a principal razão pela qual as imagens de satélites passaram a ser a mais eficiente ferramenta para uso nas aplicações que envolvem análises ambientais dos diversos ecossistemas terrestres. A perfeita combinação dessas duas tecnologias, satélites artificiais e sensores imageadores, talvez tenha sido um dos maiores benefícios até hoje propiciados pelo desenvolvimento tecnológico a serviço do levantamento dos recursos naturais terrestres.

No Ocidente, um marco importante que deu início à era do sensoriamento remoto moderno foi o lançamento, pelos Estados Unidos, do primeiro satélite de sensoriamento remoto, o *Earth Resources Technology Satellite-1* (ERTS-1), posteriormente renomeado para Landsat 1, colocado em órbita em 1972 a 919 km de altitude. O sensor imageador multiespectral desse primeiro satélite de sensoriamento remoto possibilitava a obtenção simul-

tânea de quatro imagens nas faixas do espectro eletromagnético do visível e do infravermelho próximo, portanto, além do que era possível com o uso de filmes fotográficos. A largura da área imageada na Terra por cada órbita do satélite era de 185 km, o que permitia a visão em uma única imagem de grandes áreas do terreno, até então impossível de ser obtida com as câmeras fotográficas aéreas. A cada 18 dias, uma nova imagem com resolução espacial de 76 m era obtida de qualquer área do planeta.

Do programa norte-americano Landsat, seguiram-se outros programas europeus, canadenses, asiáticos e sul-americanos, cada um procurando atender, cada vez mais, as necessidades específicas, quer seja em função das características geográficas das diversas regiões terrestres, quer seja em função das aplicações. Atualmente, há em circulação no espaço, dezenas de sensores orbitais imageando a superfície da Terra em todas as faixas espectrais possíveis e com os mais variados parâmetros de resoluções espacial e temporal. De um lado, há os sensores hiperespectrais com centenas de bandas espectrais (centenas de imagens) e, do outro, há os sensores de alta resolução espacial, capazes de detectar objetos na superfície da Terra menores que, por exemplo, 50 cm. Para atravessar a opacidade das nuvens, muito frequente em regiões como a Amazônia, há os sensores de radar com suas fontes artificiais de radiação eletromagnética (REM) que fazem das nuvens um objeto transparente.

A rapidez com que o sensoriamento remoto se desenvolveu em poucas décadas deve-se à revolução que as observações da Terra, em uma escala global e periódica, constituíram-se em meios rápidos para monitorar os fenômenos dinâmicos e as mudanças das feições terrestres. Trata-se de uma tecnologia empregada em escala mundial e que talvez reúna a maior diversidade de pesquisadores e usuários em torno de uma tecnologia de aplicação. O sensoriamento remoto é hoje a única ferramenta disponível à humanidade para integrar os mosaicos de informações espaciais, espectrais e temporais para o monitoramento das mudanças que podem ser com-

preendidas em um sentido global. Algumas das importantes aplicações do sensoriamento remoto por satélites são:

- Monitoramento ambiental urbano, rural e oceanográfico;
- Detecção e monitoramento de mudanças em biomas em escala global;
- Mapeamentos temáticos;
- Previsão de safras;
- Detecção de desastres naturais e artificiais;
- Detecção de incêndios florestais, degradação e desmatamentos;
- Reconhecimento e serviços militares de defesa e vigilância;
- Cadastramentos multifinalitários; e
- Cartografia de precisão.

No Brasil, o Instituto Nacional de Pesquisas Espaciais (INPE) teve uma ação decisiva na consolidação do sensoriamento remoto como uma tecnologia de uso em escala nacional, projetando o País como o pioneiro no Hemisfério Sul a dominar essa tecnologia. Mesmo antes do lançamento do primeiro satélite de sensoriamento remoto em 1972, o INPE se destacava como uma das primeiras instituições na área espacial a investir em pesquisas de sensoriamento remoto. Um exemplo desse investimento foi a Missão 96 que, com a colaboração da Agência Espacial Norte-Americana (NASA), executou um levantamento experimental aerotransportado com diversos tipos de sensores imageadores na região do Quadrilátero Ferrífero em 1969, em Minas Gerais. Com a criação, no início dos anos de 1970, do programa de pós-graduação em sensoriamento remoto em nível de mestrado, o INPE formou os primeiros especialistas em sensoriamento remoto. Esses pesquisadores contribuíram decisivamente, com suas dissertações, para o desenvolvimento de metodologias de interpretação de imagens e de processamento digital e na disseminação do uso de imagens de satélite em todo o território nacional.

1.2. Definição de Sensoriamento Remoto

Historicamente, reconhece-se que o termo sensoriamento remoto foi criado para designar o desenvolvimento dessa nova tecnologia de instrumentos capaz de obter imagens da superfície terrestre a distâncias remotas. Por isso, a definição mais conhecida ou clássica de sensoriamento remoto é: “uma técnica de obtenção de imagens dos objetos da superfície terrestre sem que haja um contato físico de qualquer espécie entre o sensor e o objeto”. A condição principal imposta por essa definição clássica, que é o sensor estar a uma distância remota do objeto, estabelece a base para definir o sensoriamento remoto em uma concepção um pouco mais científica, regida segundo os seguintes preceitos:

- Exigência: ausência de matéria no espaço entre o objeto e o sensor;
- Consequência: a informação do objeto é possível de ser transportada no espaço vazio; e
- Processo: o elo de comunicação entre o objeto e o sensor é a REM, a única forma de energia capaz de se deslocar no espaço.

Com base nesses preceitos, uma definição mais científica que se pode dar ao sensoriamento remoto é:

“Ciência que visa o desenvolvimento da obtenção de imagens da superfície terrestre por meio de medição quantitativa das respostas das interações da REM com os materiais terrestres”.

Essa definição de sensoriamento remoto é explícita em afirmar que o objeto imageado é registrado pelo sensor por meio de medições da REM, seja da luz solar refletida da superfície de qualquer objeto, do calor radiométrico emitido pela Terra, ou da onda emitida por uma antena de radar. Nenhum outro tipo de sensor que obtenha imagens que não seja pela detecção da REM deve ser classificado como de sensoriamento remoto. A confusão mais comum é feita com os sensores geofísicos que também podem registrar informações da superfície da Terra a partir de aeronaves, mas

que se utilizam principalmente de medições de campos de força. Exemplos são os magnetômetros que geram uma imagem do campo magnético da superfície terrestre.

Por outro lado, existe a concepção errônea de considerar como sensoriamento remoto apenas as imagens obtidas de satélites, visto que o termo foi criado quando a era espacial teve início. As fotografias aéreas que, por mais de um século, vêm sendo usadas como uma forma de observar a Terra, são, obrigatoriamente, uma classe de sensores remotos. O filme fotográfico foi o primeiro material construído pelo homem capaz de registrar a REM e de transformá-la em imagem do objeto fotografado. A partir da definição de sensoriamento remoto, fica óbvio que só se pode entender o processo de aquisição e de análise das imagens se conhecermos primeiro o que é a REM.

1.3. Natureza da Radiação Eletromagnética

A explicação do que é a REM começa com o entendimento da dualidade do comportamento da sua natureza: onda e energia. Isso quer dizer que a REM que se propaga pelo espaço vazio, como a luz solar, é, ao mesmo tempo, uma forma de onda e uma forma de energia. Esse conceito de dualidade é extremamente importante para o sensoriamento remoto, pois sempre que alguém estiver analisando qualquer tipo de imagem de sensoriamento remoto, a coexistência da REM na forma de onda e na forma de energia deverá ser considerada para que se possa explicar tudo o que se observa nas imagens com referência às características dos objetos. A dualidade do comportamento da REM, onda e energia, é formulada pelos modelos denominados de ondulatório (onda) e corpuscular (energia), que matematicamente são expressos pelas Equações 1.1 e 1.2.

$$\lambda = \frac{c}{\nu} \quad (\text{modelo ondulatorio}) \quad (1.1)$$

$$E = \frac{hc}{\lambda} \quad (\text{modelo corpuscular}) \quad (1.2)$$

em que

λ = comprimento da onda;

c = velocidade da luz no vácuo;

ν = frequência;

E = energia; e

h = constante de Planck ($6,624 \times 10^{-34} \text{ J s}^{-1}$)

1.3.1. Modelo ondulatorio

Segundo o modelo ondulatorio, a REM pode ser explicada como uma forma de onda senoidal e harmônica. De acordo com as formulações de Maxwell, uma partícula carregada eletricamente gera um campo elétrico (E) em torno de si e o movimento dessa partícula gera, por sua vez, um campo magnético (M). Ambos os campos, elétrico e magnético, vibram ortogonalmente entre si e possuem as mesmas amplitudes, isto é, alcançam os seus máximos ao mesmo tempo. As variações nos campos são causadas pelas vibrações da partícula. Quando essa partícula é acelerada, as perturbações entre os dois campos se propagam repetitivamente no vácuo em uma direção ortogonal à direção dos campos elétrico e magnético, como é mostrado na Figura 1.2.

Essas perturbações dos campos elétrico e magnético são chamadas de ondas eletromagnéticas que se propagam no espaço à velocidade da luz. Como têm a forma de uma onda, elas podem ser medidas pelo seu comprimento e sua frequência. O comprimento de onda da REM depende

de quanto tempo a partícula é acelerada, enquanto a frequência depende da frequência de vibração da partícula. Assim, define-se uma onda eletromagnética como a oscilação dos campos elétrico e magnético segundo um padrão harmônico de ondas, isto é, ondas espaçadas repetitivamente no tempo. Esses campos dinâmicos sempre ocorrem juntos como padrões inseparáveis, de modo que nem o campo elétrico puro e nem o campo magnético puro de ondas irradiadas ocorrerão separadamente um do outro.

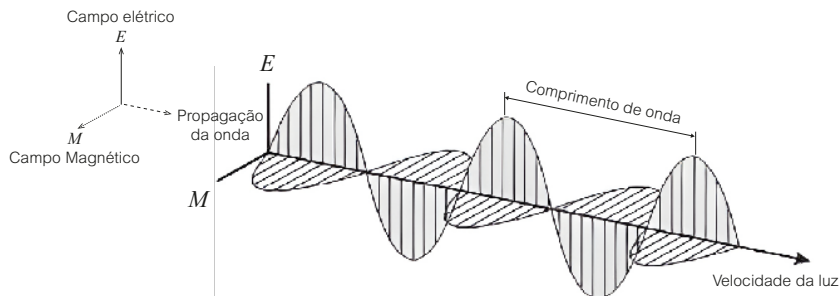


Figura 1.2. Direção de propagação da radiação eletromagnética na forma de uma onda, em função das oscilações ortogonais dos campos magnético (M) e elétrico (E).

As ondas eletromagnéticas propagam-se no vácuo na velocidade da luz ($c = 299.292,46 \text{ km s}^{-1}$ ou aproximadamente $300.000 \text{ km s}^{-1}$). Se a distância entre dois picos sucessivos de ondas for medida, como na Figura 1.2, determina-se o comprimento ou o tamanho da onda que é simbolizado pela letra grega λ e expresso no sistema de unidades métricas. As ondas podem ter comprimentos da ordem de bilionésimo de metro (raios cósmicos) até dimensões de quilômetros (ondas de rádio). Se medirmos o número de picos de ondas que passa por um ponto fixo de referência no espaço, pode-se determinar a sua frequência ν . Como a velocidade de propagação da onda é constante, as ondas de comprimentos menores têm frequências maiores, ou seja, um maior número de ondas passa pelo ponto de referência em um dado intervalo de tempo. Inversamente, as ondas de comprimentos maiores têm frequências menores porque um número menor de ondas irá passar.

Por isso, como mostra a Equação 1.1, o comprimento de uma onda eletromagnética é inversamente proporcional à sua frequência.

Por costume de uso, em sensoriamento remoto, sempre nos referimos à REM pelo seu comprimento de onda e não pela sua frequência. Por exemplo, o comprimento de onda da REM equivalente à luz solar azul é de $0,455 \mu\text{m}$ ($1 \mu\text{m}$ é igual a 10^{-6} m) e o da luz solar vermelha é de $0,617 \mu\text{m}$. Basta substituir esses valores na Equação 1.1 para ter as suas correspondentes frequências. Portanto, se considerarmos a REM apenas de acordo com o modelo ondulatório, as propriedades que os objetos retratam nas imagens são explicadas tendo-se em conta o tamanho ou o comprimento da onda e sua relação com o tamanho dos objetos. Um bom exemplo prático para explicar esse comportamento é a interação da REM com as nuvens. As radiações eletromagnéticas das ondas do visível e do infravermelho da luz solar, com comprimentos de onda de $0,4 \mu\text{m}$ a $2,5 \mu\text{m}$, respectivamente, possuem tamanhos menores que o tamanho médio das moléculas de vapor de água que compõem as nuvens, que é da ordem de $30 \mu\text{m}$. Devido a essa enorme diferença entre o tamanho das ondas do visível e do infravermelho com o tamanho das moléculas de vapor de água das nuvens, a REM da luz solar, quando incide nas nuvens, é barrada por essas moléculas e é totalmente refletida de volta ao espaço, sem que consiga atravessar a nuvem. Então, o sensor registrará apenas a reflectância da nuvem, formando uma imagem da nuvem e não dos objetos que estão sobre a superfície terrestre. Um dos sensores mais conhecidos que operam nos comprimentos de onda do visível e do infravermelho próximo é a câmera fotográfica. É de conhecimento de todos que, se no dia do levantamento aéreo houver presença de nuvens, elas estarão presentes na foto. O mesmo fenômeno ocorrerá nas imagens de satélites com esses mesmos comprimentos de onda.

E o que acontece se o comprimento da onda for maior que o diâmetro das partículas gasosas das nuvens, como é o caso das ondas eletromagnéticas de comprimentos das micro-ondas, com tamanhos da ordem de 1 mm a 100 cm ? Para entender isso, devemos considerar que uma nuvem nada

mais é do que um amontoado disperso de partículas de vapor de água, ou seja, um gás. Nesse caso, apenas uma mínima fração da dimensão da onda de tamanho centimétrico é barrada pelo correspondente tamanho da partícula, o resto da onda passa por ela. Portanto, a nuvem não é capaz de barrar fisicamente a passagem da onda. Assim, a onda atravessa a nuvem e incide nos objetos da superfície terrestre, sendo por eles refletida de volta para o sensor, que registra as imagens desses objetos. Pode-se dizer então que as nuvens são transparentes aos comprimentos de onda relativamente grandes das micro-ondas. Os sensores que trabalham com as micro-ondas são denominados de radar. Em imagens de radar, mesmo que o céu esteja totalmente coberto por nuvens, a imagem formada é uma imagem limpa, aparecendo nela toda a superfície do terreno. Daí decorre a grande utilidade dos sensores de radar nas áreas com frequentes coberturas de nuvens, como é o caso da região amazônica brasileira.

1.3.2. Modelo corpuscular

As formulações de Maxwell sobre o comportamento da REM como um movimento de ondas suaves dos campos elétrico e magnético falham em explicar certos fenômenos importantes quando a radiação interage com a matéria. Esse comportamento é mais evidente para as ondas de pequenos comprimentos da luz e foi criado com base em experimentos com a REM da luz do ultravioleta ao infravermelho próximo.

Desde os primeiros experimentos de geração de ondas eletromagnéticas, observou-se que a radiação é emitida em forma de pulsos que carregam certa quantidade de energia. Em 1901, Planck observou que a REM transferia energia de um corpo para outro em quantidades fixas e discretas, à qual denominou de fótons. Fóton é a forma física de um quantum, a partícula básica estudada em mecânica quântica (ramo da física que se dedica com o muito pequeno, nos níveis atômicos ou subatômicos), e que pode ser descrita como a partícula mensageira da energia da onda eletromagnética. Os fótons são como corpúsculos de energia sem matéria. Assim, uma onda

eletromagnética exibe propriedades de partículas, como as partículas (elétrons) exibem propriedades de onda.

O conceito de energia discreta é muito importante e revolucionou a Física do século XX, que até então considerava que qualquer forma de energia era contínua. O termo discreto significa que a intensidade de energia transportada pela REM em um dado comprimento de onda é sempre a mesma, independente do momento em que é medida. Para cada comprimento de onda, há uma quantidade específica de energia. Segundo o modelo corpuscular, pode-se definir que a REM é uma forma dinâmica de energia que se manifesta somente por suas interações com a matéria. A quantidade de energia E de um quantum é diretamente proporcional à frequência da radiação (Equação 1.3).

$$E = h\nu \quad (1.3)$$

em que

E = energia;

h = constante de Planck ($6,624 \times 10^{-34} \text{ J s}^{-1}$); e

ν = frequência.

Substituindo-se ν por $\frac{c}{\lambda}$, tem-se que (Equação 1.4):

$$E = \frac{hc}{\lambda} \quad (1.4)$$

Da Equação 1.4, verifica-se que um *quantum* ou fóton de energia E é inversamente proporcional ao comprimento de onda. Os comprimentos de ondas menores transportam, conseqüentemente, maiores quantidades de energia. Um fóton de luz visível com comprimento de onda de $0,6 \mu\text{m}$ tem

uma energia de $3,3 \times 10^{-19}$ J, e um fóton de uma onda de rádio, com comprimento de 2 m, tem uma energia muito menor, de $9,9 \times 10^{-26}$ J. A unidade usada em física atômica e óptica quântica é o elétron-Volt, que é igual a $1,59 \times 10^{-19}$ J.

De acordo com o modelo corpuscular, a energia da onda, ao atingir a superfície dos materiais, interage com ela por meio de trocas de energia entre a energia discreta que carrega e as energias, também discretas, contidas nos átomos e moléculas da matéria. As energias contidas nos átomos e moléculas estão distribuídas por níveis de energia que são determinados pelos orbitais dos elétrons em torno do núcleo e pelas forças de vibrações das ligações moleculares. Para haver uma troca ou interação da energia da radiação da onda com a energia da matéria, é preciso que a energia da REM incidente seja em parte absorvida pelos elétrons ou moléculas, o que promove uma mudança do elétron de um orbital para outro de maior energia ou um aumento da intensidade das vibrações moleculares (Figura 1.3).

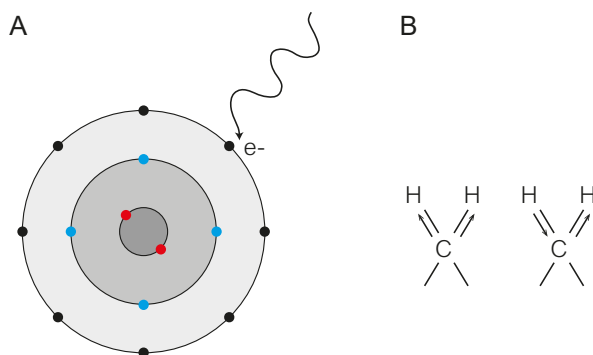


Figura 1.3. A absorção da radiação eletromagnética pelos materiais desloca o elétron para um orbital de maior energia (A) ou aumenta a frequência dos movimentos vibracionais da molécula (B).

Aparte da energia da REM que não é absorvida é refletida para o sensor. Nesse processo, as imagens medem a radiação refletida, mas, indiretamente, podemos avaliar o que foi absorvido pelos materiais. Isso é importante

porque são essas quantidades de energia eletromagnética absorvidas pelos materiais que definem algumas propriedades sobre a composição dos alvos terrestres, as quais podem ser inferidas com as imagens de sensoriamento remoto. Porém, nos grandes comprimentos de onda, como nas micro-ondas, a energia contida é muito pequena para ser absorvida pela matéria e promover trocas de energia. Assim, a onda refletida independe da composição dos materiais.

1.4. Fontes de Radiação Eletromagnética

Todos os objetos do mundo real que estejam a uma temperatura acima de zero absoluto ($-273,15\text{ }^{\circ}\text{C}$ ou zero graus Kelvin) apresentam uma movimentação aleatória das partículas dos átomos e das moléculas. Quando essas partículas se colidem, elas mudam o seu estado de energia e emitem REM. Quanto maior é a temperatura da fonte, maior é a potência irradiante de energia.

As imagens de sensoriamento remoto não dependem apenas das qualidades técnicas dos sensores, mas também da qualidade e intensidade irradiante das fontes de REM. No nosso sistema planetário, o Sol é a fonte mais potente de REM devido à alta temperatura de superfície, próxima a $6.000\text{ }^{\circ}\text{C}$. A Terra também é uma fonte de REM, embora bem menos potente que o Sol, pois a sua temperatura média é de apenas $27\text{ }^{\circ}\text{C}$. Essas duas fontes naturais de REM são as mais usadas em sensoriamento remoto, mas como veremos adiante, elas não conseguem emitir radiação em todos os comprimentos de onda. Por isso, o homem construiu fontes artificiais capazes de emitirem altas intensidades em comprimentos de onda além da capacidade do Sol ou da Terra, como na região das micro-ondas.

O cálculo da intensidade da energia que uma fonte de REM emite foi modelado por Planck, que idealizou uma fonte padrão de REM chamada de corpo negro. Um corpo negro é um modelo físico teórico de um perfeito absorvedor e emissor de energia eletromagnética. Planck utilizou esse modelo de corpo negro para calcular a quantidade máxima de intensidade radiante

que uma fonte emite em um dado comprimento de onda, em função de sua temperatura de superfície. A partir desse modelo, em 1901, Planck formulou a seguinte lei, conhecida como Lei de Planck, que é derivada da mecânica quântica (Equação 1.5).

$$M_{\lambda} = \frac{2\pi hc^2}{\lambda^5 \left[\exp\left(\frac{ch}{\lambda KT}\right) - 1 \right]} h\nu \quad (1.5)$$

em que

M_{λ} = energia radiante espectral emitida em $\text{W m}^{-2} \mu\text{m}^{-1}$;

h = constante de Planck ($6,6256 \times 10^{-34} \text{ J s}^{-1}$);

c = velocidade da luz ($2,997925 \times 10^8 \text{ m s}^{-1}$);

K = constante de Boltzman ($1,38054 \times 10^{-23} \text{ W s K}^{-1}$); e

T = temperatura absoluta (em K) da fonte.

Substituindo as constantes por seus valores, a forma usual da Equação 1.5 é apresentada na Equação 1.6:

$$M_{\lambda} = \frac{3,74151 \times 10^8}{\lambda^5 \left[\exp\left(\frac{1,43879 \times 10^4}{\lambda T}\right) - 1 \right]} \quad (1.6)$$

Vemos, na Equação 1.6, que a energia radiante espectral (M_{λ}) emitida por uma fonte a uma determinada temperatura (T) é calculada para cada comprimento de onda. Resolvendo matematicamente essa equação, chega-se às curvas de distribuição de energia emitida por diferentes fontes (Figura 1.4). Para cada temperatura da fonte, há um pico máximo de emis-

são dentro do intervalo de comprimento de onda que a fonte emite. Quanto mais alta a temperatura da fonte, maior é o intervalo de onda que a fonte emite e mais alto é o pico de emissão de energia eletromagnética. Observe as diferenças desses valores para as fontes com temperaturas típicas do Sol e da Terra. Por lei, qualquer material real não pode emitir, termalmente, a uma taxa que exceda à do corpo negro.

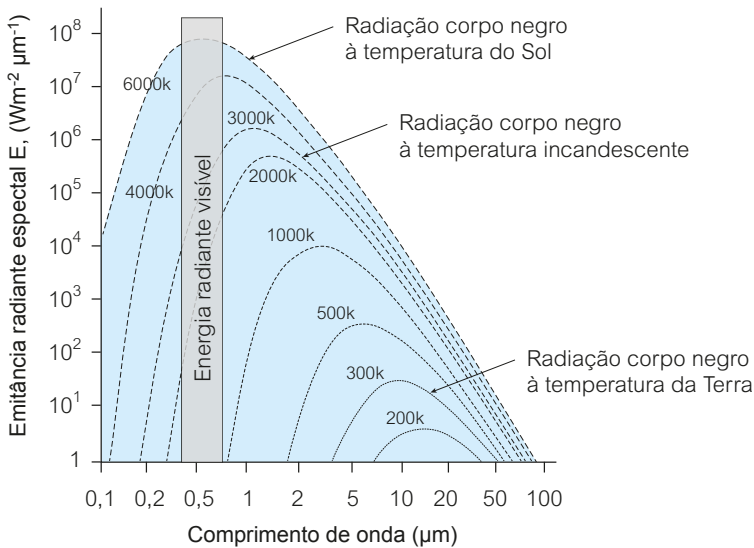


Figura 1.4. Valores de intensidades da energia radiante emitida em diferentes comprimentos de onda (0,1 μm –100 μm) por fontes de radiação com diferentes temperaturas, segundo o modelo de corpo negro de Planck.

Das várias curvas de emitância radiante espectral da Figura 1.4, a que mostra a distribuição da energia a 6.000 K é a que mais se aproxima da curva de emitância espectral do Sol. Nessa figura, o pico máximo de emitância situa-se no comprimento de onda de 0,48 μm , ou seja, na região espectral da luz visível. Para a curva de emitância radiante da fonte a 300 K, que é comparável à temperatura média da superfície da Terra, o seu pico máximo de energia emitida está situado no comprimento de onda de 9,6 μm . Essa

relação entre a temperatura da fonte e seu pico de emissão radiante em comprimento de onda é descrita pela lei de Wien (Equação 1.7).

$$\lambda_{max} = \frac{K}{T} \quad (K = 2.898 \text{ } \mu\text{K}) \quad (1.7)$$

Por exemplo, para o Sol, o comprimento de onda da radiação emitida máxima é de (Equação 1.8):

$$\lambda_{max} = \frac{2.898 \text{ } \mu\text{K}}{6.000 \text{ K}} = 0,483 \text{ } \mu\text{m} \quad (1.8)$$

As fontes artificiais de REM de comprimentos de onda das micro-ondas, isto é, dos sensores de radar, têm suas intensidades de emissão determinadas por componentes eletrônicos. Por isso, elas ficam à parte do modelo corpuscular de Planck que é calculado com base na temperatura das fontes.

1.4.1. Fontes naturais de radiação eletromagnética: o Sol e a Terra

Como o Sol e a Terra não se comportam como um corpo negro, eles irradiam energia a uma taxa menor que um corpo negro à mesma temperatura. O Sol, uma massa gasosa incandescente composta de átomos de hidrogênio (71%) e hélio (26%), é a principal fonte de REM para o sensoriamento remoto. O seu núcleo, por estar submetido à enorme pressão exercida por sua massa, atinge a altíssima temperatura de cerca de 15 milhões de graus Kelvin, por isso, o núcleo solar é considerado um reator estelar. Essa altíssima temperatura transforma o hidrogênio em hélio por meio de fusão de núcleos de hidrogênio em núcleos de hélio. São como reações nucleares que geram energia equivalente à energia liberada por explosões atômicas

(20 milhões de graus). Os gases de hélio em ebulição se propagam para a fotosfera ou corona solar, de onde irradia os raios de luz eletromagnética. Na corona solar, a temperatura torna-se muito mais baixa que na região do núcleo, situando-se em torno de 6.000 K. A corona solar emite uma enorme quantidade de radiação, em torno de 72 milhões de Watts, em uma esfera com 650.000 km de raio. Ao se deslocar no espaço à velocidade da luz, essa onda de energia reparte-se em esferas concêntricas de raio cada vez maiores. Ao chegar à órbita terrestre, a 149,5 milhões de km do centro do Sol, ela cai para 1.367 W m^{-2} , o que equivale a 160 t de luz por minuto atingindo a Terra e, por fim, a quantidade total que atravessa a atmosfera e incide na superfície terrestre é da ordem de 187 W m^{-2} . O intervalo e comprimento de onda da luz solar situam-se, aproximadamente entre a luz ultravioleta (0,1 μm –0,38 μm) e o infravermelho de ondas curtas (1,3 μm –3,0 μm).

A Terra, por ter uma temperatura interna muito mais baixa que o núcleo solar, transmite, por radioatividade, calor para a superfície apenas o suficiente para colocar a temperatura superficial em torno de 27 °C na média. É, portanto, uma fonte de REM muito menos intensa que o Sol e de intervalo espectral mais limitado, o qual é denominado de termal. É por causa da sua baixa temperatura de superfície que não é uma fonte emissora de luz visível.

Semelhante ao que se viu na Figura 1.4 para o corpo negro, na Figura 1.5A, é mostrada, a título de comparação, a configuração do fluxo de energia radiante do Sol e da Terra. Nessa figura, a intensidade da energia radiante da fonte solar decresce abruptamente abaixo dos comprimentos de onda do ultravioleta e atinge valor mínimo próximo ao comprimento de onda de 3,0 μm . Devido a isso, somente o intervalo espectral da REM de 0,45 μm a 2,5 μm é útil ao sensoriamento remoto. Nesse intervalo espectral, o pico máximo de intensidade de radiação encontra-se na faixa de 0,45 μm a 0,76 μm , conhecida como a região do visível. A Terra tem uma intensidade de fluxo de energia bem abaixo do fluxo solar, tendo o seu máximo de emitância radiante em torno de 9,6 μm e um intervalo espectral útil ao sensoriamento remoto entre 8,0 μm a 14,0 μm .

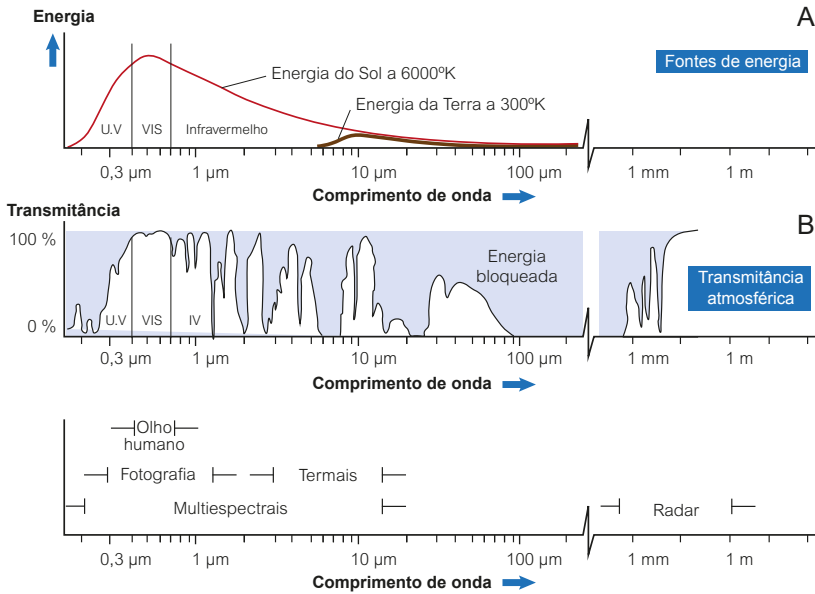


Figura 1.5. (A) Distribuição da intensidade da energia emitida pelas fontes naturais de radiação eletromagnética, medidas no topo da atmosfera. Barras indicam a região de comprimento de onda de máxima emitância; (B) Transmitância da radiação eletromagnética através da atmosfera. As áreas em branco correspondem às regiões espectrais de transparências da atmosfera e as áreas em azul correspondem às regiões espectrais nas quais os gases atmosféricos absorvem a radiação eletromagnética.

1.5. Espectro Eletromagnético

Denomina-se espectro eletromagnético as regiões espectrais da REM conhecidas pela humanidade. O espectro eletromagnético conhecido se estende dos comprimentos de onda dos raios cósmicos aos comprimentos de corrente alternada emitidos pelas redes de alta tensão. A faixa de comprimento de onda da REM é praticamente ilimitada. Há estrelas que emitem radiações com máximo de irradiância em comprimentos de onda do raio-X e outras, como o Sol, que emitem seu máximo de REM no visível.

A REM deve ser vista como um espectro contínuo. Porém, o espectro eletromagnético foi arbitrariamente dividido pelo homem em intervalos de comprimentos de onda com base nos mecanismos físicos geradores da energia eletromagnética e nos mecanismos físicos de sua detecção. A nomenclatura de cada um dos intervalos foi feita em função do uso que o homem encontrou para as suas aplicações. Para os usuários do sensoriamento remoto, é essencial saber os valores dos comprimentos de onda desses intervalos, as denominações que recebem, e quais comprimentos de ondas são possíveis de serem detectados em cada tipo de sensor. Isso porque, quando estivermos com uma imagem de sensoriamento remoto em mãos, a primeira pergunta que se faz é: qual o comprimento de onda dessa imagem? Usando uma terminologia mais correta, perguntar-se-ia: qual é a banda espectral dessa imagem?

Um dos intervalos ou faixas de comprimentos de onda mais familiares é a faixa de luz visível da radiação solar, por ser aquela que o olho humano é capaz de detectar. O intervalo espectral da luz visível foi decomposto pela primeira vez em 1766 por Isaac Newton, atravessando a luz branca por um prisma de vidro (dispersão) e emergindo do lado oposto do prisma em raios de luz coloridos (Figura 1.6). Cada raio colorido tem o seu comprimento de onda específico. A inclinação de cada raio, ao emergir da outra face do prisma, é devido à relação entre o comprimento da onda e o índice de refração do prisma (vidro). Essa experiência se repete toda vez que vemos a formação de um arco-íris no céu, provocado pela dispersão da luz pelas gotas de água em um dia chuvoso. Foi a partir dessa experiência que o homem propôs a divisão do espectro eletromagnético. A Tabela 1.1 apresenta a divisão que é a mais aceita em sensoriamento remoto. Não são limites rigorosamente precisos, sendo comum encontrarmos na literatura pequenas diferenças de limites dos intervalos espectrais. As faixas espectrais em negrito são as mais usadas em sensoriamento remoto. Na Tabela 1.2, é mostrado um detalhamento dos limites de comprimentos de onda do espectro de cores do visível.

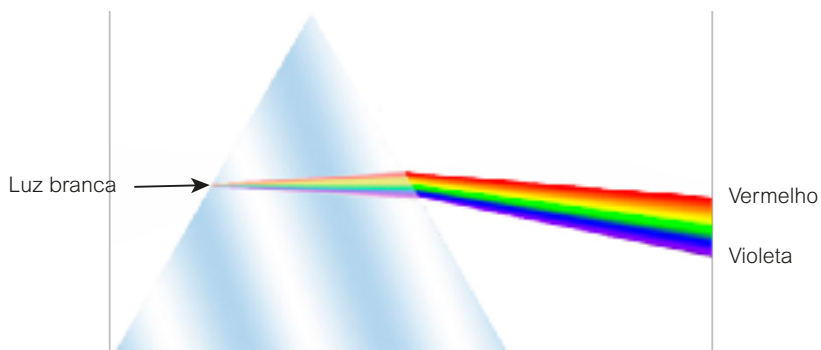


Figura 1.6. Dispersão da luz branca em seis cores espectrais ao atravessar o prisma de vidro.

Tabela 1.1. Divisão dos comprimentos de ondas do espectro eletromagnético. Os intervalos espectrais em negrito correspondem às faixas mais utilizadas em sensoriamento remoto.

Intervalo espectral	Comprimento de onda
Raios cósmicos	0,01 Å°
Raios gama	0,01 Å°–0,1 Å°
Raios X	0,1 Å–10 Å°
Ultravioleta	100 nm–0,38 µm
Visível	0,38 µm–0,76 µm
Infravermelho próximo	0,76 µm–1,2 µm
Infravermelho de ondas curtas	1,2 µm–3,0 µm
Infravermelho médio	3,0 µm–5,0 µm
Infravermelho termal	5,0 µm–1 mm
Micro-ondas	1 mm–100 cm
Rádio	1 m–10 km
Áudio	10 km–100 km
Corrente alternada	> 100 km

Å° = 10^{-10} m; nm = 10^{-9} m; µm = 10^{-6} m.

Tabela 1.2. Divisões dos comprimentos de ondas das cores da região do visível.

Cor da luz refletida	Comprimento de onda (μm)
Violeta	0,380–0,455
Azul	0,455–0,482
Azul esverdeado	0,482–0,487
Azul - verde	0,487–0,493
Verde azulado	0,493–0,498
Verde	0,498–0,530
Verde amarelado	0,530–0,559
Amarelo - verde	0,559–0,571
Amarelo esverdeado	0,571–0,576
Amarelo	0,576–0,580
Laranja amarelado	0,580–0,587
Laranja	0,587–0,597
Laranja avermelhado	0,597–0,617
Vermelho	0,617–0,760

Cada um dos comprimentos de onda da REM interage com os objetos terrestres com intensidades diferentes, resultando em diferentes intensidades de reflectância. Seria como dizer que cada objeto tem a sua assinatura espectral de reflectância. Por isso, um dos parâmetros mais importantes para definir as características de um sensor é a escolha dos comprimentos de onda das imagens que o sensor irá adquirir. As imagens não são definidas em um comprimento de onda específico, isto é, abrangem uma pequena largura espectral para cada imagem, à qual nos referimos como sendo bandas espectrais.

1.6. Síntese das Características mais Importantes dos Intervalos Espectrais

Visível (0,45 μm –0,76 μm)

Faixa do espectro solar com a mais alta intensidade de fluxo radiante e onde há a maior janela atmosférica, bastante transparente, deixando passar uma grande quantidade de radiação. Por isso, é muito usada em sensoriamento remoto. É a faixa responsável pela interação com os pigmentos da vegetação e com os minerais, solos e rochas, dando origem às suas cores. O problema dessa faixa espectral é o alto espalhamento da radiação solar pelos gases atmosféricos que pode reduzir o contraste da refletância dos alvos terrestres. É chamada de visível porque o olho humano é sensível a essa região espectral.

Infravermelho próximo (0,76 μm –1,2 μm)

Faixa do espectro solar onde a atmosfera também é bastante transparente, mas ocorrem algumas bandas de absorções, impedindo que todo o intervalo possa ser continuamente utilizado para sensoriamento remoto. É o intervalo onde ocorrem importantes interações da REM com os níveis de energia eletrônica dos átomos, gerando feições espectrais diagnósticas que permitem identificar a natureza de vários tipos de rochas, principalmente as com metais de transição (Fe, Ni, Cr e Mn) na sua constituição e de solos com composição ferrífera.

Infravermelho de ondas curtas (1,2 μm –3,0 μm)

Faixa espectral geológica, porque é nessa faixa que os vários minerais de alteração hidrotermal apresentam as suas feições de absorção diagnósticas. Também é a faixa onde os comprimentos de onda em torno de 1,4 μm e 1,9 μm são totalmente absorvidos pelas moléculas de vapor de água da atmosfera, proibindo o uso do sensoriamento remoto e, por consequência, a determinação da presença de água molecular nas estruturas dos materiais terrestres. Importante também para a identificação de argilominerais nos solos.

Infravermelho médio (3,0 μm –5,0 μm)

Faixa onde o Sol e a Terra não emitem quantidades suficientes de energia que possam ser detectadas pelos sensores. Somente alvos com elevadas temperaturas, como vulcões e incêndios, podem ser detectados, pois agem como fontes próprias de emissão de radiação. É uma faixa espectral pouco usada no sensoriamento remoto, à exceção de sensores meteorológicos ou atmosféricos.

Infravermelho termal (5,0 μm – 1,0 mm)

Faixa termal devida à radiação emitida pelos objetos terrestres em função das suas temperaturas de superfície. A melhor janela atmosférica nessa faixa espectral para imageamento orbital é o intervalo de 8,0 μm a 14,0 μm , porque, acima de 30 km, a camada de ozônio absorve toda a radiação além de 14 μm emitida pela Terra. Usada para detecções de anomalias de calor no terreno.

Micro-ondas (1 mm – 100 cm)

Faixa de uso de sensores ativos (radar) que utilizam fontes artificiais para a geração da REM. Por causa do tamanho do comprimento de onda, o radar tem a habilidade de operar em condições atmosféricas adversas, sob condições de cobertura de nuvens ou de chuvas, podendo operar tanto de dia como à noite. Importante para a região amazônica, que apresenta cobertura persistente de nuvens.

CAPÍTULO 2

Modelos de Interação da Radiação Eletromagnética

Paulo Roberto Meneses



O entendimento das propriedades espectrais dos objetos registrados nas imagens de sensoriamento remoto passa pelo entendimento dos processos de interação da radiação eletromagnética com os objetos. Os processos de interação são distintos para cada uma das três regiões espectrais: óptica, termal e micro-ondas. Na faixa óptica da luz solar, a radiação recebida pelo sensor corresponde à porção refletida pela superfície do terreno, subtraída da porção da radiação que foi absorvida pelos materiais. Na região das micro-ondas, não há absorção da radiação pelos objetos e, assim, a reflectância é ditada pelo modelo corpuscular. Na região do termal, recorremos a outro fenômeno radiométrico que é a emitância da radiação em função da temperatura dos objetos, sendo esse assunto tratado no Capítulo 3, no qual são descritos os sensores imageadores.

2.1. Interação da Radiação Eletromagnética: Fotografias x Imagens

Em sensoriamento remoto, devemos creditar que o conhecimento sobre as propriedades dos objetos da superfície terrestre iniciou-se a partir de fotos aéreas obtidas por câmeras fotográficas munidas com filmes sensíveis à luz do visível. A partir dos anos de 1930, surgiram os primeiros métodos de fotointerpretação, a maioria dirigida a aplicações para mapeamentos urbanos, geológicos, geomorfológicos, de solos e de vegetação. Invariavelmente, todos esses métodos de interpretação tiveram seus conceitos fundamentados na análise das formas ou texturas dos objetos, e não de suas reflectâncias, em razão da existência de pelo menos três parâmetros que são inerentes aos sensores fotográficos, conforme se segue:

- 1) Os filmes fotográficos, que correspondem aos detectores da radiação eletromagnética (REM), registram a reflectância dos objetos em

uma única imagem ou banda espectral que abrange quase toda a região da luz visível, do azul ao vermelho ($0,45\ \mu\text{m}$ – $0,76\ \mu\text{m}$). Uma banda espectral com essa largura diminui muito a capacidade do sensor de registrar as diferenças de reflectância que os objetos possam exibir no visível. Consequentemente, as fotos aéreas apresentam pouca variação de níveis de cinza, ou seja, pouca variação de reflectância. Por isso, os sensores fotográficos são classificados como de baixa resolução espectral.

- 2) O segundo parâmetro consiste no alto poder de resolução espacial das fotos, o que permite a identificação de objetos de tamanhos menores que 1 m. Isso é possível porque cada segmento do filme é composto de dezenas de milhões de microscópicos cristais de haleto de prata que atuam como detectores da REM. São sensores de alta resolução espacial.
- 3) O terceiro parâmetro deve-se ao tipo de geometria de recobrimento aéreo do terreno, que é feito por tomadas de fotos que se sobrepõem em 60% para possibilitar uma visão tridimensional do terreno com o recurso de um simples instrumento óptico, o estereoscópio.

Como consequência da conjunção desses três fatores, os atributos de forma dos objetos são destacados, enquanto a capacidade de análise dos níveis de cinza de uma foto aérea torna-se secundária por não trazer informações suficientes para uma análise do comportamento espectral da reflectância dos objetos. A Figura 2.1 elucida essa questão a partir dos exemplos do comportamento da reflectância da água e da vegetação. Devemos entender que, no filme fotográfico, a medida da reflectância de qualquer objeto é a média da reflectância dos comprimentos de onda de todo o intervalo da luz visível nas faixas espectrais do azul, verde e vermelho. Como mostram os gráficos da Figura 2.1, a água e a vegetação têm reflectâncias bem distintas nas faixas espectrais do azul, verde e vermelho. Se tivéssemos uma imagem em cada um desses comprimentos de onda, poderíamos diferenciar esses dois objetos pelas suas reflectâncias. No entanto, consi-

derando-se que a medida da reflectância pelos filmes fotográficos engloba toda a faixa do visível, isso resulta em médias de reflectâncias quase iguais às da água e da vegetação.

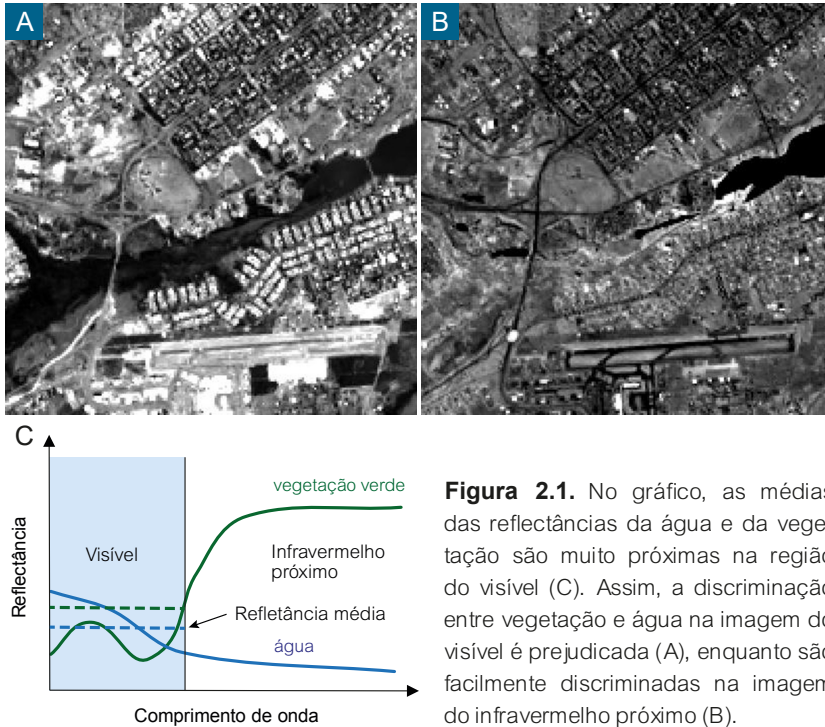


Figura 2.1. No gráfico, as médias das reflectâncias da água e da vegetação são muito próximas na região do visível (C). Assim, a discriminação entre vegetação e água na imagem do visível é prejudicada (A), enquanto são facilmente discriminadas na imagem do infravermelho próximo (B).

Nas fotos aéreas, os níveis de cinza da água e da vegetação são muito próximos, levando-nos a concluir que a análise da reflectância não é um bom parâmetro de fotointerpretação. Um exemplo real dessa situação é mostrado na Figura 2.1 que corresponde a uma parte do Lago Paranoá de Brasília, rodeado por mata de galeria. A foto aérea dessa região (A) mostra que o que parece ser o corpo de água do lago (tom de cinza escuro), de fato, é água mais a mata de galeria. Porém, a imagem do infravermelho (B) mostra que o corpo de água é de fato bastante reduzido. Nessas situações,

os sensores multiespectrais, pelo fato de terem a capacidade de obter imagens simultâneas em vários comprimentos de onda, mostram vantagens em relação às fotos aéreas.

2.2. Processos de Interação da Luz Solar

Ao longo da história do sensoriamento remoto, a concepção de sensores multiespectrais tem valorizado mais os sensores que operam na faixa óptica do espectro eletromagnético. Há diversas razões para essa afirmação, sendo uma delas a mais importante. No intervalo espectral da luz solar, a quantidade de energia do fluxo radiante é muito intensa, suficiente para interagir, ao nível microscópico, com os constituintes da matéria, o que possibilita que as imagens multiespectrais possam diferenciar ou distinguir um material de outro pelas diferenças de suas reflectâncias. Um exemplo é mostrado na Figura 2.2, onde as duas imagens obtidas em diferentes comprimentos de onda indicam que, para identificar os pivôs centrais de irrigação em uma área cultivada, basta identificar os objetos com formas circulares, independente de qual seja a banda espectral da imagem. Adicionalmente, para identificar os pivôs que possuem cultura agrícola em desenvolvimento, basta identificar os pivôs com tons de cinza escuro na imagem do visível (baixa reflectância) e tons de cinza claro na imagem do infravermelho próximo (alta reflectância). Note, na figura, que apenas um dos pivôs, indicado pela seta, não obedece a essa regra, portanto, esse pivô não possuía nenhuma cultura agrícola em pleno desenvolvimento no momento da aquisição da imagem.

É exatamente essa condição de que diferentes tipos de materiais podem se distinguidos por suas diferenças de reflectância que leva a uma das principais motivações do sensoriamento remoto moderno. Com as imagens multiespectrais, é possível acrescentar, à análise textural dos alvos no terreno, a análise espectral da reflectância, adicionando informações sobre a composição dos materiais. Podemos então resumir que, enquanto as fotos aéreas fornecem informações sobre as propriedades da forma dos objetos,

as imagens multiespectrais oferecem, adicionalmente, informações sobre a composição química e física dos objetos.

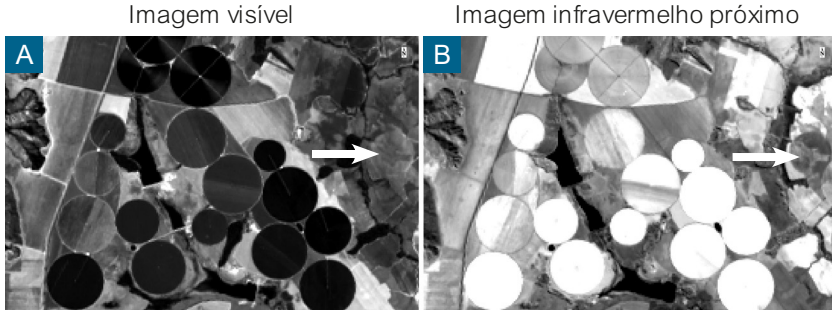


Figura 2.2. As diferenças de comprimentos de onda das duas imagens mostram que os pivôs com cobertura vegetal aparecem com tons de cinza escuro na imagem do visível (A) e correspondentes tons claros na imagem do infravermelho próximo (B). A seta indica o único pivô que foge a essa regra.

Dentro desse contexto, para que uma imagem multiespectral possa ser interpretada, é fundamental conhecer os mecanismos de interação da REM com a superfície dos materiais terrestres. É por meio do entendimento de como se processam esses mecanismos de interação que se pode determinar como e com qual intensidade cada objeto ou material reflete a onda eletromagnética. Sendo assim, o intérprete deverá estar habilitado em saber identificar qual banda espectral contém a informação que se deseja. Os sensores multiespectrais normalmente possuem configurações de bandas espectrais selecionadas para interesses em múltiplas aplicações.

Para melhor compreender a interação da REM com os materiais terrestres, é preciso inicialmente saber que a luz, ao incidir sobre a superfície do objeto, poderá sofrer três possíveis processos de interações (Figura 2.3): refletância (ρ_λ), transmitância (τ_λ) e absortância (α_λ). Como, pela lei de conservação de energia, nenhuma quantidade de energia é perdida, tem-se que $\rho_\lambda + \tau_\lambda + \alpha_\lambda = 1$.

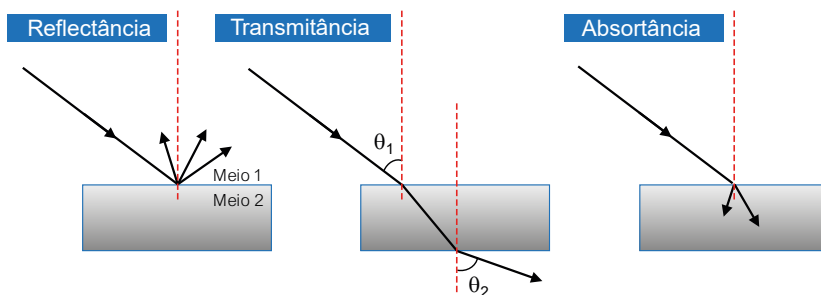


Figura 2.3. Processos de interação da radiação eletromagnética com os materiais.

As quantidades de reflectância, absortância e transmitância dependem dos valores dos coeficientes ópticos dos materiais, que geralmente são propriedades difíceis de serem medidas, dada a heterogeneidade dos materiais. Por sua vez, como o sensor posiciona-se acima do terreno, não há como medir diretamente o componente da transmitância da luz que ocorre em objetos transparentes como a água. Portanto, em sensoriamento remoto, costuma-se simplificar essas quantidades, analisando-se as propriedades dos materiais em função apenas das relações de absortância e reflectância. Um material com alta taxa de absorção reflete pouco e terá, na imagem, um tom de cinza escuro. Para fins práticos, classificamos os materiais em função de suas absortâncias/reflectâncias em dois tipos:

- Opacos: quando contêm, em suas constituições, materiais opacos como minerais de magnetita ou compostos como a matéria orgânica que absorvem bastante a REM em quase todos os comprimentos de onda; e
- Transopacos: quando os materiais exibem baixa reflectância em uma dada região espectral e alta reflectância em outra região espectral, como é o caso da maioria dos objetos terrestres.

Além das propriedades que controlam as quantidades de energia radiante absorvida e refletida pelos objetos, há fatores externos que influenciam as medições da reflectância pelos sensores, tornando complexas a

análise e a interpretação das imagens. Um pixel de tamanho grande corresponde muitas vezes a um pixel contendo mistura de diversos objetos, isto é, com um padrão de reflectância que é difícil de ser decodificado. Além disso, os gases presentes na atmosfera espalham a REM, adicionando reflectância à reflectância propriamente dita do pixel, alterando, dessa forma, a reflectância inerente aos materiais. Como vimos no Capítulo 1, o efeito atmosférico tende a diminuir o contraste espectral entre os alvos.

Por esses motivos, a teoria física e experimental sobre a interação da REM com os materiais e suas implicações na análise e interpretação de imagens é assunto especializado em sensoriamento remoto. Recomenda-se, aos leitores, procurarem por essas publicações especializadas, como o livro de Meneses, Almeida e Baptista, publicado em 2019 sob o título *Reflectância dos Materiais Terrestres: Análise e Interpretação*. No presente contexto, cabe abordar esse assunto apenas em um nível introdutório, mas o suficiente para se entender por que o sensoriamento remoto oferece uma gama variada de imagens em diferentes comprimentos de onda.

O conceito visto no Capítulo 1 sobre a natureza da REM, uma onda carregando corpúsculos de energia, é fundamental para entendermos como a REM proveniente de uma fonte interage com os objetos da superfície terrestre e é registrada pelos sensores para a formação das imagens. Em outras palavras, para analisar os dados que uma imagem armazena, precisamos primeiramente entender quais relações existem entre as propriedades do objeto, o tamanho de uma onda e a quantidade de energia que a onda eletromagnética carrega. Portanto, para melhor compreender esse processo, recorreremos ao princípio da dualidade da REM, analisando separadamente as interações com relação aos tamanhos das ondas e das quantidades de energias da REM. Isso está de acordo com o seguinte princípio básico bastante consagrado dos métodos de fotointerpretação: as propriedades dos objetos contidos nas fotos e nas imagens são decodificadas pela análise de dois atributos: a forma ou textura do objeto e a sua cor. Por isso, em sensoriamento remoto, costuma-se relacionar a forma ou a textura que os ob-

jetos exibem na imagem com o tamanho da onda da REM (lembrando que $\lambda = c/\nu$), enquanto relacionamos a energia que a onda carrega ($E = hc/\lambda$) com as cores dos objetos, que são respostas diretas da composição químico-física dos materiais. Denominamos esses dois processos de interações macroscópica e microscópica da REM com os objetos.

2.2.1. Interação macroscópica

O processo de interação macroscópica tenta explicar, apenas pelo tamanho da onda, como objetos de diferentes tamanhos podem ser reconhecidos e identificados nas imagens pelas intensidades da reflectância da onda. Na prática, entende-se que o tamanho de um objeto não é uma medida da dimensão de sua área, mas de uma propriedade que é descrita nos processos de fotointerpretação como textura de superfície do objeto. A textura representa a rugosidade topográfica da superfície e pode ser estimada pelas variações da altura da superfície do material. É fácil imaginar isso comparando a visão que temos da textura bastante rugosa de dossel de uma floresta com a visão da textura lisa e uniforme de um gramado. A textura se mostra muito eficiente para explicar as diferenças observadas nas imagens entre as diversas formas de relevo, os tipos de dosséis da vegetação, a superfície lisa de um corpo de água sem movimento ou com ondas revoltas, ou de um solo gradeado ou arado.

Em sensoriamento remoto, as texturas das superfícies dos objetos são genericamente classificadas em dois tipos: especulares e difusoras. Usa-se o critério de Rayleigh h para determinar se uma superfície é especular ou difusora, o que é dependente da relação do comprimento de onda e do ângulo de incidência θ da radiação (Equação 2.1):

$$h_s \leq \lambda / (8 \cos \theta) \text{ ou } h \geq \lambda / (8 \cos \theta) \quad (2.1)$$

em que

h_s = altura média da superfície;

θ = ângulo de incidência dos raios de luz; e

λ = comprimento de onda.

Conforme mostrado na Figura 2.4A, quando $h_s \geq \lambda/(8\cos\theta)$, a superfície do objeto é classificada como difusora, com os raios de luz refletindo-se em todas as direções (reflectância hemisférica). Nesse caso, a textura da superfície do objeto é dita ser rugosa. Na Figura 2.4b, quando $h_s \leq \lambda/(8\cos\theta)$, a imagem do objeto é classificada como uma superfície especular, com os raios de luz incidente refletindo-se em direção oposta e com ângulos próximos ao ângulo de incidência da radiação incidente. Nesse caso, diz-se que a textura da superfície do objeto é lisa. Em alguns casos, pode-se ter uma mistura de reflectância especular e difusora. Para todos os efeitos, costuma-se classificar uma superfície como sendo difusora quando 25% ou mais da radiação é refletida difusamente. Na natureza, a maioria dos objetos tem um comportamento difuso em relação aos comprimentos de onda da luz solar.

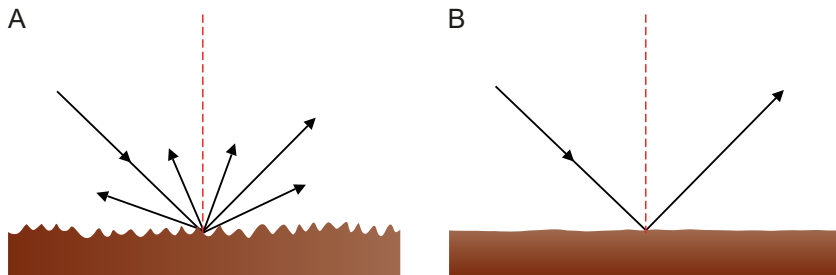


Figura 2.4. (A) Reflectância difusa de uma superfície rugosa com os raios refletidos espalhados em todas as direções; e (B) Reflectância especular de uma superfície lisa com o raio incidente refletido no sentido oposto com o mesmo ângulo.

Simplificando, podemos dizer que o que define se uma superfície é rugosa ou lisa é o comprimento de onda incidente. Consideremos essa situa-

ção. Um solo gradeado, cuja textura medida pelas mínimas variações da altura da sua superfície seja de poucos centímetros, será considerado rugoso em uma imagem de comprimento de onda do visível ($0,40\ \mu\text{m}$ – $0,76\ \mu\text{m}$), pois o tamanho das ondas do visível é milhões de vezes menor que o tamanho das variações de altura h da superfície do solo. No entanto, esse mesmo solo gradeado poderá ser considerado liso, ou seja, especular, em uma imagem adquirida na faixa das micro-ondas, visto que as micro-ondas têm comprimentos de onda de até 100 cm. Dessa forma, as áreas de solos gradeados serão vistas em uma imagem do visível com feições bem diferentes das que são vistas em uma imagem das micro-ondas. O usuário de sensoriamento remoto que não souber das implicações desse processo de interação macroscópica julgará, analisando as duas imagens, que são solos com texturas diferentes, imaginando, talvez, que as imagens tenham sido tomadas em datas diferentes.

De acordo com a Equação 2.1, relativa ao critério de Rayleigh, podemos deduzir que quase todas as superfícies dos objetos terrestres são consideradas de textura rugosa em imagens obtidas em comprimentos de onda micrométricos, do visível ao infravermelho ($0,4\ \mu\text{m}$ – $2,5\ \mu\text{m}$). A possibilidade de se reconhecer, nas imagens desse intervalo espectral, algum grau de diferença de rugosidade de alguns materiais está quase sempre condicionada aos efeitos de iluminação dos raios solares. Em geral, a regra estabelece que, quanto maior é a rugosidade de superfície de um material, menor será a reflectância.

Superfícies muito rugosas tendem a criar uma maior área de superfície sombreada do que as superfícies mais lisas. Considerando-se que as áreas sombreadas apresentam reflectância nula em todas as bandas de um sensor multiespectral da faixa óptica, alvos mais rugosos de um mesmo material terão uma reflectância mais baixa. Um bom exemplo para demonstrar essa regra é mostrado pelas variações de dossel da vegetação. Na composição colorida da Figura 2.5, as áreas de mata de galeria, que possuem estrutura de dossel muito rugosa, comparativamente com as áreas de ve-

getação de cultivos agrícolas, com dossel mais liso, são representadas pela cor verde mais escura. Medidas de reflectâncias de folhas de diferentes espécies não mostram quase nenhuma diferença de padrão de reflectância e, assim, somente as variações de estrutura de rugosidade de dossel explicam as diferenças das intensidades das cores verdes vistas na Figura 2.5. Todos os demais materiais deveriam seguir essa regra, desde que fosse possível comparar um material de mesma composição (rocha ou solo), mas com rugosidades diferentes. Medidas espectrorradiométricas de amostras de solos em laboratório mostram que a granulometria é um parâmetro importante na reflectância, repetindo a regra de que o aumento da granulação do solo diminui a reflectância.



Figura 2.5. Diferenças de reflectância nas faixas espectrais do infravermelho de ondas curtas (SWIR1), infravermelho próximo (NIR) e vermelho (Red) entre as vegetações de mata de galeria e de cultura agrícola sob sistema de irrigação por pivô central devido às diferenças nas rugosidades desses dois tipos de dossel, demonstrando o resultado da interação macroscópica da radiação eletromagnética.

Outro critério para dimensionar a interação macroscópica é estabelecido pela relação entre a direção azimutal da iluminação solar com as di-

reções das estruturas de relevo. Ao longo de um ano, se compararmos as direções azimutais de iluminação entre as estações do inverno e verão para os Hemisférios Norte ou Sul, vemos uma significativa mudança no ângulo de iluminação em relação à linha do equador. Como mostra a Figura 2.6, essa mudança no ângulo de iluminação altera o sombreado do relevo de modo favorável ou desfavorável para o realce do relevo. Quando os raios solares iluminam ortogonalmente as faces frontais do relevo, criando sombreado nas faces opostas, gera-se o par “face iluminada e face sombreada”, condição ideal para o olho humano criar uma visão tridimensional do relevo. Quando os raios solares iluminam igualmente as faces do relevo em um ângulo paralelo às direções do relevo, sem criar, nessa condição, sombreado, a visão tridimensional do relevo fica comprometida. Essa forma de interação macroscópica acontece em todas as bandas e é essencial nos estudos de estruturas geológicas realçadas pelas direções de relevo.



Figura 2.6. Processo de interação macroscópica condicionado pela direção azimutal de iluminação em relação às direções do relevo. No inverno (25 de setembro de 2015) (A), a direção de iluminação é aproximadamente paralela ao relevo, implicando em menor realce de relevo na imagem. No verão (28 de janeiro de 2021) (B), a direção de iluminação é aproximadamente frontal ao relevo, realçando os lineamentos geológicos de forma mais acentuada.

2.2.2. Interação microscópica

O processo de interação microscópica da REM com os objetos é um assunto bem mais complexo e requer dados auxiliares sobre o comportamento espectral da reflectância dos materiais para que se possa analisar e interpretar as propriedades dos objetos vistos nas imagens. Essas propriedades dizem respeito à composição dos materiais. O modelo de interação microscópica fundamenta-se no entendimento da REM como uma forma de energia, expressa pela equação $E = hc/\lambda$. Nesse processo de interação, ocorrem trocas de energia da REM com as energias dos átomos e moléculas dos materiais. Em sensoriamento remoto, o modelo corpuscular é mais adequado para os comprimentos de onda do visível ao infravermelho de ondas curtas porque, como previsto pela equação acima, ondas de pequenos comprimentos possuem maiores energias, com capacidade suficiente para promover saltos quânticos dos elétrons dos átomos ou dos orbitais moleculares. Assim, materiais de diferentes composições, ou seja, com diferentes constituições de elementos químicos de átomos e moléculas, absorverão quantidades diferentes da energia da onda eletromagnética. Consequentemente, as suas reflectâncias serão diferentes, resultando em imagens com diferentes tons de cinza.

O modelo quântico é a chave para explicar as interações no nível microscópico de trocas energéticas da REM com os materiais, pois assim se comportam, com muita semelhança, as energias contidas na REM e nos átomos e moléculas dos materiais. Em outras palavras, cada comprimento de onda carrega uma quantidade de energia específica e única e cada átomo ou molécula tem seus níveis ou estados de energia específicos. Pela teoria quântica, o átomo e a molécula podem existir somente em certos estados ou níveis de energia permitidos. Assim, o átomo ou a molécula somente absorverá uma quantidade específica de energia eletromagnética (número de fótons) se essa quantidade for suficiente para mudar, no átomo ou na molécula, o seu estado de energia de fundamental para outro excitado. Diz-se que ocorreu uma transição entre os estados de energia.

Na Figura 2.7, é mostrado um esquema bastante simplificado para exemplificar uma interação microscópica envolvendo trocas de energia, segundo o modelo corpuscular da REM, para dois tipos de rochas com composições químicas diferentes. Os diagramas hipotéticos dos níveis de energia dos átomos de ferro presentes no xisto e das moléculas CO_3 presentes no calcário indicam, pelas setas, que os saltos quânticos para níveis de energias superiores são diferentes. Isso indica que as quantidades de energia da onda de um determinado comprimento que foram absorvidas pelo átomo de ferro e pela molécula de CO_3 são diferentes. Pelo diagrama da figura, vemos que os átomos de ferro do xisto absorvem uma maior quantidade de energia da REM (maior salto quântico) do que as moléculas de CO_3 do calcário. Por isso, as áreas no terreno que expõem o xisto aparecerão com um tom de cinza mais escuro nas imagens do que as áreas de calcário naquele comprimento de onda. Essa dedução é simples de se ver pelo diagrama. Pela imagem, só poderíamos chegar a essa conclusão se tivéssemos em mãos os dados referentes às curvas de reflectância das duas rochas.

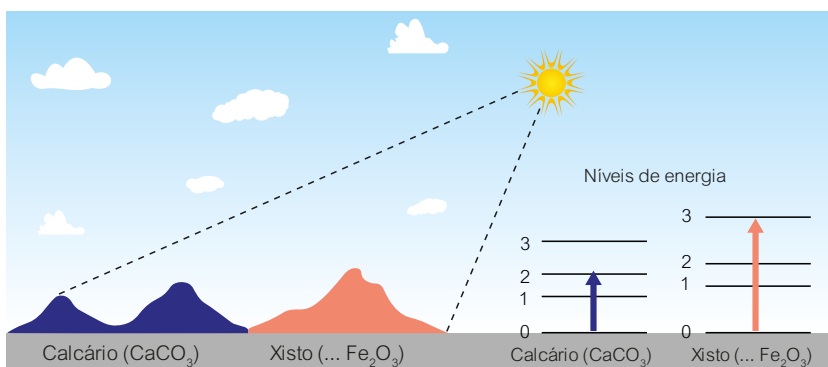


Figura 2.7. Diagrama de energia mostrando os saltos quânticos de energia devido à absorção da radiação eletromagnética pelos átomos de Fe contidos na rocha xisto e pelas moléculas de CO_3 contidas na rocha calcária.

A quantidade de energia que é absorvida pelos materiais tem uma grande importância para o sensoriamento remoto, porque ela se mostrará como

uma assinatura espectral da composição daquele material. São chamadas de bandas de absorção. A maneira prática que se tem para identificar em quais comprimentos de onda ocorrem as bandas de absorção de um material é por meio de medições dos espectros de reflectância dos materiais feitos em laboratório ou no campo. Para obter essas medidas, usam-se os espectrorradiômetros que são instrumentos que plotam, na forma de gráficos, as intensidades da energia refletida. Os espectrorradiômetros têm uma função equivalente aos sensores imageadores. Os sensores imageadores medem a radiância refletida por comprimento de onda (bandas de um sensor) e configuram a imagem dos objetos em níveis de cinza, onde os valores de cinza claro representam alvos com alta reflectância e os valores de cinza escuros representam alvos com baixa reflectância. Os espectrorradiômetros fazem a mesma coisa, apenas não geram uma imagem de saída.

As medidas de reflectância obtidas pelos espectrorradiômetros são mostradas em gráficos que relacionam a porcentagem da reflectância com o comprimento de onda, como os que estão na Figura 2.8 para os materiais rocha, solo e vegetação. As estreitas depressões indicadas pelas setas identificam os comprimentos de onda em que as absorções da REM pelos átomos e moléculas possibilitam identificar alguns dos constituintes presentes no material. Ao especialista em sensoriamento remoto, cabe determinar quais elementos e moléculas presentes em cada um desses materiais são os responsáveis por essas bandas de absorção. No espectro da vegetação, a absorção em $0,68\ \mu\text{m}$ é devida às moléculas de clorofila, enquanto as duas absorções centradas em $1,4\ \mu\text{m}$ e $1,9\ \mu\text{m}$ são devidas às moléculas de água. No espectro da rocha carbonática (calcário), a absorção em $2,27\ \mu\text{m}$ indica a presença do íon CO_3 na calcita que é o principal mineral formador dessa rocha. No espectro do solo, a absorção em $0,95\ \mu\text{m}$ é devida à presença de óxido de ferro. Sendo assim, um sensor será ideal para diagnosticar e identificar esses três tipos de materiais quando as bandas do sensor estiverem posicionadas coincidentemente com as bandas de absorção desses materiais.

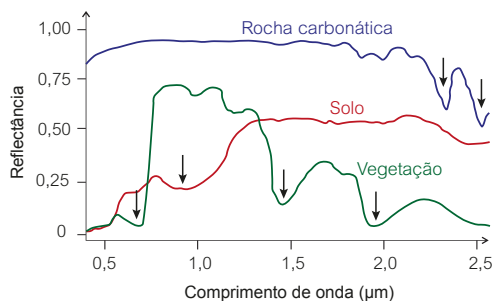


Figura 2.8. Exemplos de espectro de vegetação sadia, rocha carbonática e solo com óxido de ferro. As setas indicam presença de bandas de absorção nos três espectros ilustrados.

Os espectros de reflectância são fundamentais para guiar o intérprete na seleção das bandas do sensor para a interpretação das imagens. Um jeito prático de se fazer isso é mostrado na Figura 2.9. Com base em alguns conhecimentos prévios da área, o intérprete plota os espectros de reflectância dos alvos existentes na área, sobrepondo as posições das bandas espectrais das imagens do sensor que estiver usando. Com isso, ele poderá determinar quais bandas do sensor têm a capacidade de discriminar os alvos, que é na prática, o que interessa em sensoriamento remoto.

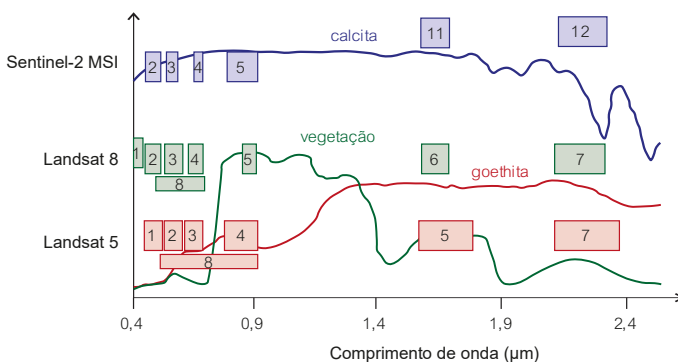


Figura 2.9. Espectros de reflectância e as posições das bandas dos sensores Sentinel-2, Landsat 5 e Landsat 8 para avaliação das bandas que melhor diferenciam espectralmente os três tipos de materiais.

2.2.3. Manifestação conjunta das interações macroscópica e microscópica

Devido à indissociabilidade de onda e energia, somos obrigados a afirmar que as imagens dos objetos sempre são, em maior ou menor proporção, respostas conjuntas em função do tamanho da onda e da intensidade de sua energia. Há, particularmente, uma situação didática que demonstra a atuação conjunta dessas duas formas de interação. Podemos ver isso na Figura 2.10 que mostra um recorte de imagens tomadas de uma área com densa cobertura vegetal da região amazônica. As duas imagens são do sensor *Thematic Mapper* (TM) do satélite Landsat 5, uma da região do infravermelho próximo de $0,76\ \mu\text{m}$ a $0,90\ \mu\text{m}$ (banda 4) e a outra da região do visível de $0,63\ \mu\text{m}$ a $0,69\ \mu\text{m}$ (banda 3). A primeira impressão que se tem é de que, na imagem do infravermelho próximo, consegue-se ver o relevo abaixo das árvores. Na realidade, ambas as imagens refletem a reflectância apenas do topo das árvores, pois, nesses comprimentos de onda, a REM não consegue penetrar a vegetação. Na imagem superior, vemos que a topografia dos dosséis é uma reprodução bem aproximada da topografia do relevo subjacente, pois a radiação solar incidente ilumina as faces frontais do relevo e sombreia as faces opostas.

Combinando as condições de iluminação do relevo com as informações que mostram a curva de reflectância da vegetação, podemos ver que, na imagem do visível, as faces opostas sombreadas exibem níveis de cinzas escuros. As faces frontais que refletem a REM também exibem níveis de cinzas escuros devido à alta absorvância da vegetação no visível. Essa falta de contraste dos níveis de cinza não favorece, ao olho humano, uma visão tridimensional do relevo, isto é, a topografia do dossel. Na imagem do infravermelho próximo, as faces sombreadas continuam a exibir níveis de cinzas escuros, mas devido à alta reflectância da vegetação nesse comprimento de onda, as faces frontais exibem níveis de cinzas claros, criando uma condição de alto contraste entre as faces, o que favorece a percepção visual humana em 3D. Nessa banda do infravermelho próximo, as intera-

ções macroscópica e microscópica se combinam favoravelmente, fornecendo assim, uma percepção do relevo que na realidade está encoberto pela vegetação.

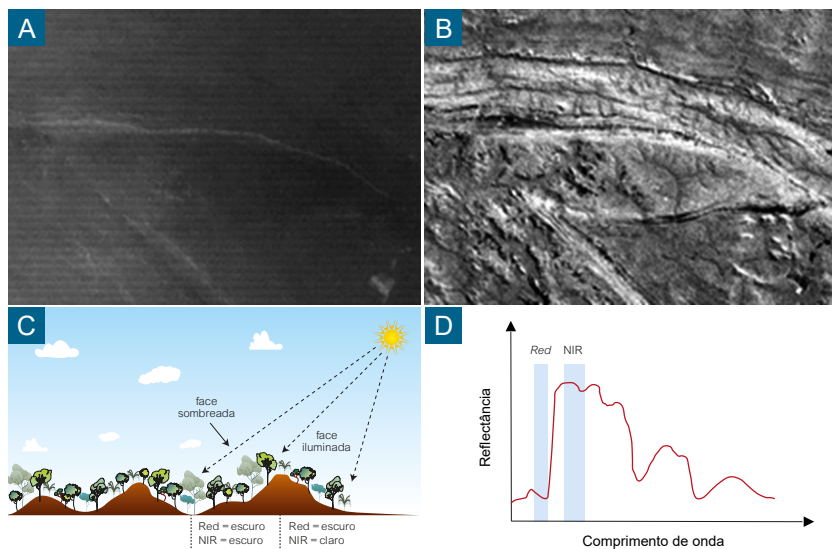


Figura 2.10. Visão do relevo de uma área densamente coberta por vegetação da região amazônica nas bandas espectrais do vermelho (A) e infravermelho próximo (B) do satélite Landsat 5. As figuras (C) e (D) mostram a associação da interação macroscópica e microscópica da radiação eletromagnética no topo do dossel.

2.3. Processos de Interação da Radiação de Micro-ondas

Nos grandes comprimentos das micro-ondas em que operam os sensores de radar, a quantidade de energia contida nessas ondas é muito pequena e insuficiente para interagir ao nível dos átomos e moléculas dos materiais e assim possibilitar a obtenção de informações sobre a composição do material. Consequentemente, as informações extraídas das imagens de radar são quase que exclusivamente dependentes de uma relação entre o

tamanho da onda e as variações texturais e estruturais das superfícies do terreno, podendo-se inferir muito pouco sobre a composição dos materiais. A interação é, portanto, exclusivamente macroscópica. O processo de interação macroscópica da radiação das micro-ondas é mais complexo do que nos pequenos comprimentos de onda da luz solar porque depende muito dos parâmetros do sensor e da geometria de imageamento (ver Capítulo 4). Os principais parâmetros são: comprimento da onda; direção de visada; ângulo de incidência local do pulso no terreno; e constante dielétrica, a única propriedade que depende da composição do objeto.

2.3.1. Interação em relação ao comprimento de onda

O comprimento de onda é importante para se determinar a rugosidade da superfície. Em uma imagem de radar, uma superfície rugosa é um refletor difuso que retroespalha o sinal em todas as direções, que é a condição mais comum para os alvos terrestres. Para isso acontecer, o tamanho da onda deve ser menor que as variações de rugosidade da superfície. A textura da superfície é classificada como lisa quando não há retorno do sinal retroespalhado em direção ao sensor. Nesse caso, a superfície tem um comportamento especular e ocorre quando o tamanho da onda é muitas vezes maior que as variações de rugosidade da superfície. Os espelhos de água calmos são exemplos típicos de superfície lisa.

A Figura 2.11 mostra essa dependência das superfícies dos alvos serem classificadas de rugosa ou lisa em uma imagem de radar com comprimento de onda de 23,5 cm. Alvos como a superfície calma da água do lago, rodovias pavimentadas, pista de pouso e solos expostos exibem reflexão especular da onda de 23,5 cm. Como não há retorno do sinal para a antena de radar, esses alvos aparecem na imagem em tons de cinza escuro. Os demais alvos aparecem na imagem com tons de cinza médio ou claro por causa do retroespalhamento difuso que ocorre quando a onda é menor que a rugosidade dos alvos. Um caso particular de radiação difusa é a refle-

xão de canto. Alvos como edificações urbanas, orientadas ortogonalmente à radiação incidente, refletem com forte intensidade a radiação incidente de volta para a antena, aparecendo com tons de cinza muito claros. Esses alvos são conhecidos como refletores de canto e o fenômeno é conhecido como *double bouncing*.

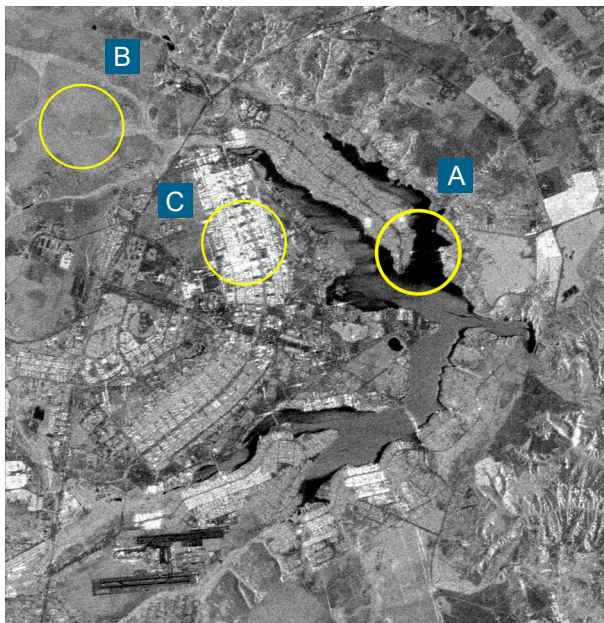


Figura 2.11. Imagem de radar na banda L de comprimento de onda de 23,5 cm do satélite japonês ALOS PALSAR, mostrando alvos com reflexões especulares em tons de cinza escuros (A), alvos com reflexão difusa em tons de cinza variados (B) e alvos com reflexão de canto em tons de cinza muito claro (C).

Pode-se medir a rugosidade de uma superfície em termos de altura h das pequenas variações da altitude da superfície (Figura 2.12). A medida h possui escala centimétrica, o que significa que as variações topográficas do relevo sempre são superfícies difusoras. As rugosidades das superfícies podem ser classificadas como lisa (reflexão especular), intermediária

(reflexão difusa média) e rugosa (reflexão difusa forte). Para isso, usam-se cálculos que levam em consideração o comprimento de onda e o ângulo de iluminação. Na Figura 2.12, o comprimento de onda é de 5,6 cm e o ângulo de iluminação é de 40° .

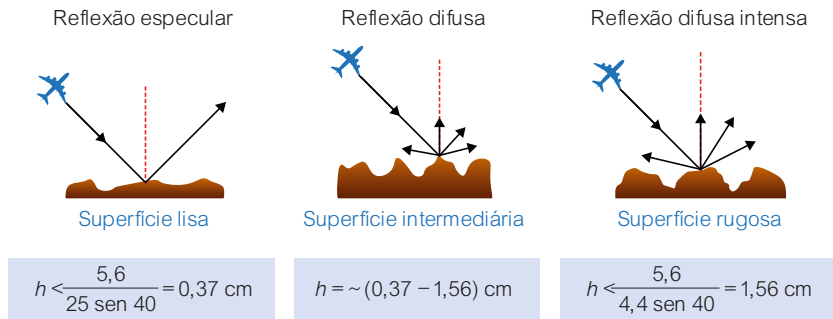


Figura 2.12. Critérios para classificação de superfícies da Terra em rugosidades lisa, intermediária e rugosa.

2.3.2. Interação em relação à direção de iluminação

A relação entre a direção de iluminação da onda e as orientações dos alvos no terreno é um parâmetro importante para o realce das feições do terreno em imagens de qualquer comprimento de onda, porém, trata-se de um fator crítico nas imagens das micro-ondas, pois as interações são unicamente macroscópicas. A ortogonalidade entre a direção da frente de onda e o alvo determina que uma seção maior ou total da frente de onda interaja com o alvo, causando um forte retorno da onda ao sensor, exibindo tons de cinza mais claros nas imagens. Os alvos fora da ortogonalidade têm uma interação com uma seção menor da frente da onda e, portanto, um retorno de sinal mais fraco, exibindo tons de cinza mais escuros nas imagens. O resultado desse processo é o realce diferencial das feições do terreno, sendo mais intenso para feições com direções ortogonais à frente de onda.

2.3.3. Interação em relação ao ângulo de incidência local

O terceiro parâmetro que controla a interação macroscópica é o ângulo de incidência local, formado pela inclinação do pulso de onda e a linha perpendicular à superfície do terreno em cada ponto do terreno onde a onda incide. Em um relevo inclinado, o ângulo de incidência local descreve a relação entre o pulso de radar e a declividade superficial das encostas de relevo. Serão inúmeros os ângulos de incidência, não havendo condições e nem a necessidade do intérprete saber o valor de cada ângulo de incidência local. Um desenho dessa relação é apresentado na Figura 2.13. Os diferentes ângulos de incidência local que são encontrados ao longo da encosta resultarão em diferentes intensidades de retorno do sinal para o sensor, apresentando-se com diferentes tons de cinza nas imagens, dando, ao intérprete, a noção da forma textural da encosta. Quanto menor for o ângulo de incidência local, mais intenso será o retorno do sinal.

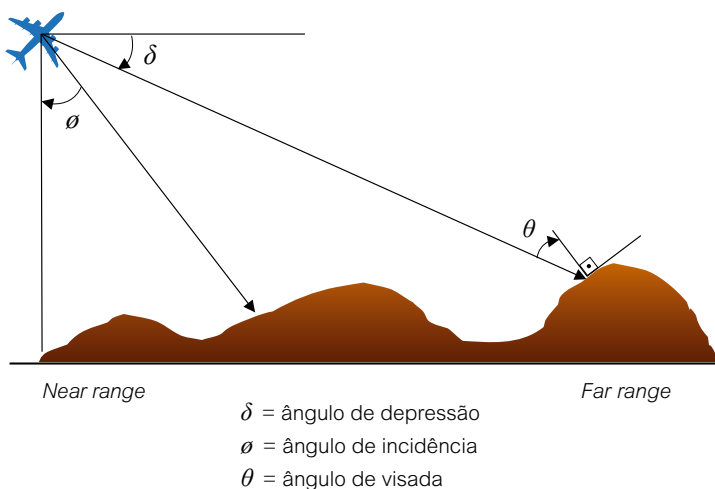


Figura 2.13. Relação entre inclinação do pulso da onda e a normal à superfície do terreno onde a onda incide, como forma de medir o ângulo de incidência local (θ).

2.3.4. Constante dielétrica

Há uma condição única para que os alvos tenham um retorno forte de sinal, independente das condições ideais de tamanho da onda, da rugosidade da superfície ou do ângulo de incidência. Essa condição é quando um alvo tem uma alta constante dielétrica. Constante dielétrica corresponde à capacidade dos materiais em formar dipolos na presença de um campo elétrico externo e de armazenar e conduzir energia elétrica. Os alvos naturais, para terem altas constantes dielétricas, precisam conter água livre na sua estrutura porque a água tem uma alta constante dielétrica, próxima de 80. Rochas e solos, em suas condições naturais (secos), possuem uma baixa constante dielétrica, entre 3 a 8.

Quanto maior for a constante dielétrica do material, maior é a interação do campo elétrico do pulso de micro-ondas com a molécula de água e mais intenso é o retorno do sinal. Há, dessa forma, uma relação direta de proporcionalidade entre a constante dielétrica e a intensidade do pulso de micro-ondas refletido. Por isso, as vegetações fotossinteticamente ativas que armazenam água em suas estruturas foliares internas apresentam tons de cinza muito claros nas imagens de radar (Figura 2.14). Solos úmidos respondem da mesma maneira, o que torna a imagem das micro-ondas uma excelente opção para a separação de solos úmidos e secos. Entretanto, estimar umidade de solo por meio de imagens de radar, que pode ser importante em diversas aplicações, como em modelagens hidrológicas, não é uma tarefa simples por causa das interferências das condições de rugosidade do terreno e da cobertura vegetal.

Essa propriedade da constante dielétrica explica a extraordinária imagem de radar adquirida pela missão *Shuttle Imaging Radar* (SIR-A) em novembro de 1981. A parte da cena colorida na Figura 2.15 corresponde a um recorte de imagem Landsat de uma parte do deserto do Arizona nos Estados Unidos. Pelo fato de as areias secas terem uma baixa constante dielétrica, as ondas de radar conseguiram penetrar o solo até cerca de 3,5 m de profundidade, o que é mostrado pela imagem de radar na faixa central da figu-

ra em tons de cinza. O que se vê são paleocanais subsuperficiais que estão soterrados pela areia e, portanto, invisíveis na imagem Landsat.

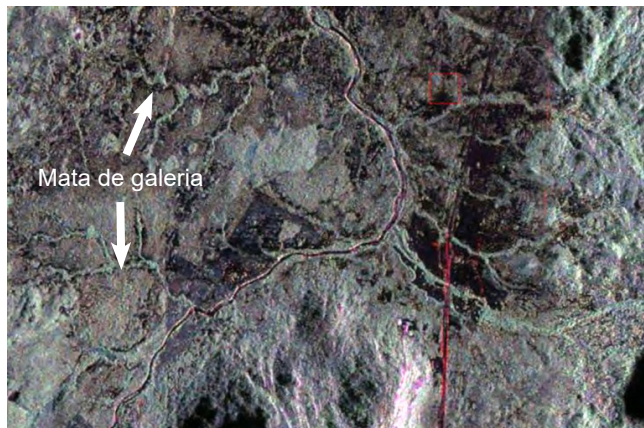


Figura 2.14. Imagem de radar aerotransportada pela aeronave R99B da Força Aérea Brasileira (FAB) mostrando as matas de galerias em tons de cinza claros devido à alta constante dielétrica da vegetação.

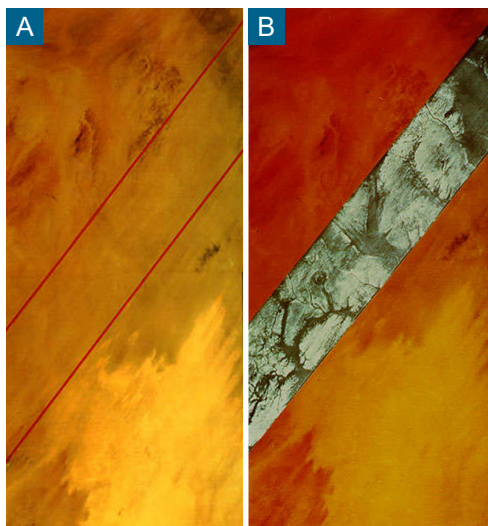


Figura 2.15. Composição colorida RGB de imagens do satélite Landsat da região desértica de Flagstaff, Arizona, EUA (A). Em (B), foi sobreposta uma imagem de radar para realçar paleocanais de drenagem cobertos por uma camada de sedimentos de areia seca.

Fonte: <http://airandspace.si.edu/exhibitions/looking-at-earth/online/orbital-vistas/landsat.cfm>.

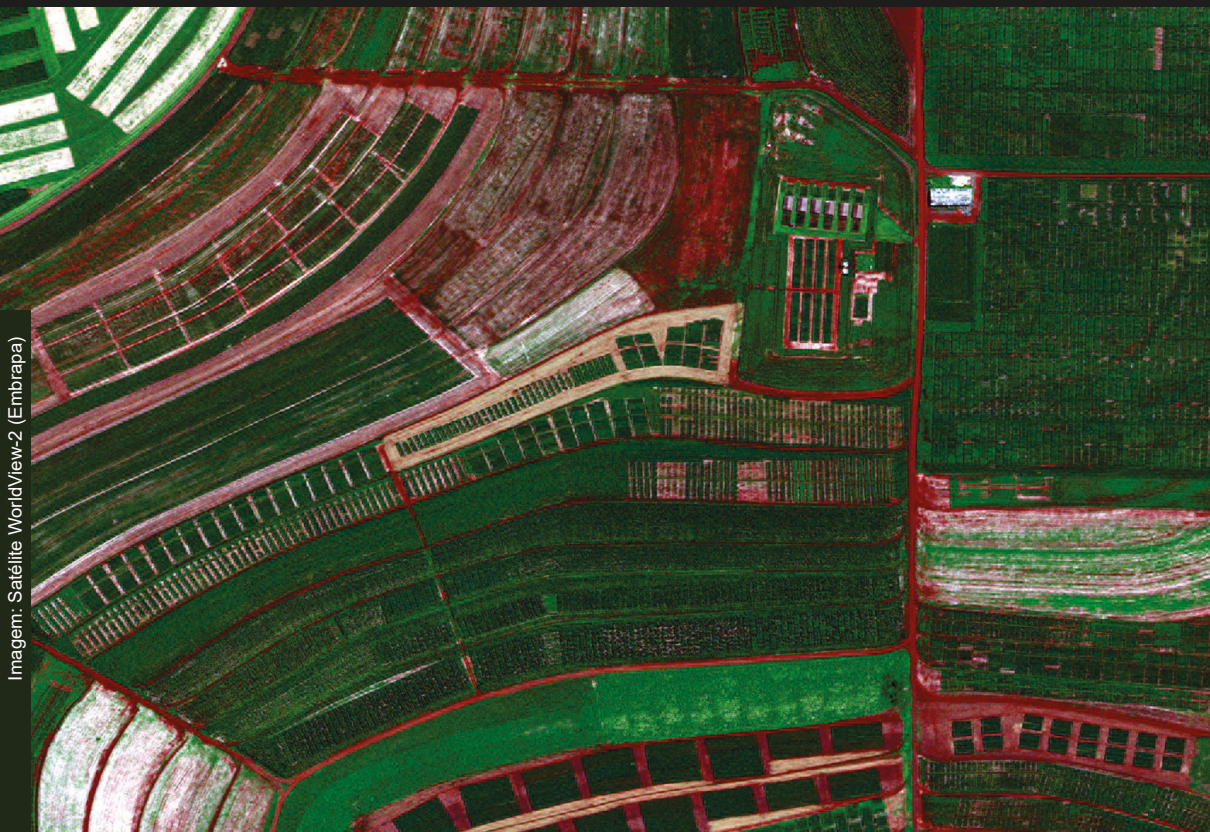
CAPÍTULO 3

Sistemas Sensores Imageadores

Paulo Roberto Meneses

Edson Eyji Sano

Gustavo Macedo de Mello Baptista



O sensoriamento remoto tem sua base nos princípios que levaram ao desenvolvimento das câmeras fotográficas aéreas. Apesar do enorme benefício que as fotografias aéreas proporcionaram para a humanidade, com o tempo, foi-se percebendo que o uso de aeronaves restringia a cobertura de grandes áreas e, sobretudo, a tomada de dados de forma sistemática para fins de estudos e levantamentos dos recursos terrestres. Com o desenvolvimento tecnológico de construção de satélites artificiais de sensoriamento remoto no início da década de 1970, sensores imageadores multiespectrais deram início a uma nova maneira de obter dados da superfície terrestre de forma sistemática e global. Detectores eletrônicos de radiação eletromagnética estenderam a capacidade dos sensores imageadores além do poder dos filmes fotográficos, possibilitando a tomada de imagens desde a região espectral da luz visível até à de micro-ondas. Essa nova era de sistemas sensores remotos foi inaugurada em 1972 com o sensor *Multispectral Scanner System* (MSS) do satélite Landsat 1, que começou a fornecer imagens em múltiplas bandas espectrais do visível ao infravermelho. No início da década de 1990, as agências espaciais do Japão, Canadá e Europa lançaram satélites com sensores de abertura sintética que operam na faixa espectral das micro-ondas. A partir dos anos 2000, para atender às necessidades dos diferentes usuários de sensoriamento remoto, muitos sistemas sensores foram desenvolvidos, oferecendo uma ampla opção de resoluções espaciais, espectrais e temporais, com imagens capazes de detectar objetos menores que 1 m. Neste capítulo, são descritos os mecanismos de imageamento dos sensores multiespectrais ópticos e termais e do sensor de radar da região das micro-ondas.

3.1. Sensores Multiespectrais

Definem-se sensores imageadores multiespectrais como sendo os instrumentos capazes de obter múltiplas imagens simultâneas da superfície terrestre, desde o comprimento de onda da luz visível azul até a região do infravermelho termal. Os sensores das faixas espectrais óptica e do infravermelho termal operam praticamente da mesma forma. O mais comum é ter os dois sensores operando simultaneamente na mesma plataforma, apenas em módulos instrumentais distintos. Com o tempo, acostumou-se a utilizar o termo sensor multiespectral para os sensores que operam na faixa óptica da luz solar, de 0,45 μm a 2,5 μm , reservando o nome de sensor termal exclusivamente para essa faixa espectral do infravermelho termal, muito embora os sensores termais sejam também multiespectrais. Ambos são sensores passivos, pois se utilizam de fontes naturais de radiação eletromagnética (REM), o Sol e a Terra.

Pode-se justificar a grande eficiência dos sensores imageadores multiespectrais pelos seguintes motivos: os dados são obtidos em formato digital; podem ser operados de plataformas espaciais por meio de satélites, permitindo a tomada de imagens de forma repetitiva ao longo de vários anos; e as imagens podem ser obtidas em um amplo intervalo de comprimento de onda, com dezenas a centenas de bandas espectrais. Combinadas, essas características projetaram esse modelo de sensoriamento remoto como uma das melhores e mais bem-sucedidas inovações para a tomada de dados da superfície terrestre. As imagens multiespectrais, quando obtidas por sensores imageadores a bordo de satélites, possibilitam um imageamento global da Terra em curto período de tempo, incomparáveis para uma monitoração precisa dos eventos e mudanças ambientais. Por sua vez, a versatilidade de se obter imagens em diferentes comprimentos de onda ampliou sua utilização para muito além das aplicações temáticas das Ciências da Terra. Além disso, o formato digital trouxe o computador como um meio rápido de visualizar as imagens e processá-las para análises qualitativa e quan-

titativa, resultando no surgimento de uma nova tecnologia de tratamento de dados, denominada de processamento digital de imagens.

Deixando de lado o entendimento da complexa engenharia dos diferentes componentes instrumentais que integram um sensor imageador multiespectral e termal, pode-se dizer que o coração de um sensor imageador é composto pelos detectores da REM e pela eletrônica de sinal que convertem a REM refletida ou emitida em um número digital. É devido à maior efetividade e versatilidade que os detectores têm para registrar a REM em diferentes intervalos de comprimento de onda, que é conferido, aos sensores imageadores, a capacidade de gerar, simultaneamente, um grande número de imagens em diferentes comprimentos de onda.

Um imageador é um instrumento que, a cada instante, capta a energia eletromagnética da radiação de pequenas amostras de área da superfície terrestre (pixel), convertendo-a, por meio dos detectores, em um sinal elétrico cuja intensidade é correspondente à intensidade da radiação de cada pixel. O sinal elétrico, amplificado e filtrado, é amostrado e quantizado em valores digitais dentro de um valor escalar dimensionado em bits. A conversão ou a relação entre o nível do sinal analógico/digital (conversor A/D) é feito usualmente por um quantizador linear que converte o sinal para o valor digital inteiro mais próximo. Com o uso de programas de processamento de imagens, os valores digitais da imagem são convertidos em valores de brilho (tons de cinza) quando a imagem é exposta em um monitor de vídeo. O valor digital zero converte-se na tela em um nível de brilho escuro e o valor digital máximo em um nível de brilho claro, estabelecendo-se, entre esses dois valores extremos, uma escala linear de brilhos de tons cinza.

Os sensores ópticos e termais são constituídos, basicamente, de três componentes instrumentais: um óptico, um de detecção da REM e um eletrônico de registro de dados, os quais são agrupados em uma única unidade aeroespacial. A configuração ou o arranjo desses componentes está representado na Figura 3.1 e suas funções resumidas no quadro abaixo.

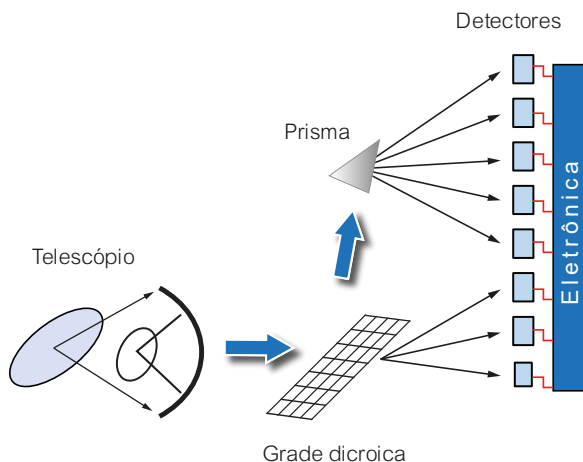


Figura 3.1. Componentes básicos de um sistema sensor multiespectral.

3.2. Síntese dos Sistemas que Compõem Sensores Ópticos e Termais

Sistema óptico

É constituído de espelhos e lentes que possuem a função de focar uma área na superfície terrestre e medir a REM por ela radiada. Os imageadores utilizam um telescópio portátil do tipo Ritchey Chretch, que permite a observação de pequenas áreas da superfície terrestre a uma altitude de centenas ou milhares de quilômetros. Os telescópicos possuem essa capacidade, pois operam com distâncias focais de alguns metros e aberturas ópticas com ângulos de miliradianos.

Sistema de dispersão e detecção

Grades dicróicas e prismas são utilizados como filtros para dividir a REM em vários intervalos de comprimento de onda ou bandas espectrais que são direcionados para incidirem sobre um conjunto de detectores que geram sinais elétricos proporcionais às intensidades das radiações neles incidentes.

Sistema eletrônico

Constitui-se de componentes eletrônicos que ampliam os sinais elétricos de saída dos detectores e os transformam em valores digitais. Os sinais digitais podem ser transmitidos em tempo real para estações terrestres de rastreamento de satélites onde são gravados e, posteriormente, processados em laboratórios de produção de imagens.

3.3. Mecanismos de Imageamento Multiespectral

O tipo de tecnologia dos sensores multiespectrais da faixa óptica e termal pode ser resumido em três sistemas de imageamento que operam com poucas diferenças entre si:

- Sensores multiespectrais que utilizam detectores discretos e espelho de varredura.
- Sensores multiespectrais que utilizam arranjos lineares de detectores.
- Sensores hiperspectrais que utilizam arranjos lineares e de área.

Schowengerdt (2007) viu o uso das imagens geradas por esses sistemas sensores de dois modos. O primeiro modo pode ser chamado de centrado na imagem (*image centered*) e inclui os dois primeiros tipos de sensores multiespectrais. O interesse principal está na construção de uma imagem na qual o intérprete pode analisar as relações espaciais entre os objetos no terreno, o que confere uma similaridade do processo de análise das imagens com os métodos de fotointerpretação das fotos aéreas. O objetivo principal é a criação de mapas por meio de interpretação visual dos dados, requerendo habilidade do intérprete para localizar e relacionar espacialmente os objetos. Cada intérprete tem que ter suas habilidades específicas, caso o interesse seja na criação de mapas geológicos, geomorfológicos, de cobertura vegetal, de solos, de infraestruturas urbanas e rurais, ou de propósitos ambientais e mesmo militares. A gama de aplicações é ampla.

O segundo modo pode ser chamado de centrado nos dados (*data centered*), que inclui os sensores hiperespectrais. Nesse caso, o analista está interessado no próprio dado e não em suas relações espaciais. É uma análise feita do pixel, a qual é baseada nas feições de absorção para estimar a composição dos materiais dentro de cada pixel. Mapas não são produzidos e a análise de centenas de bandas é conduzida por algoritmos que processem os dados de forma automática, com pouca interferência do analista. Dele é requerido um sólido conhecimento em comportamento espectral dos materiais que compõem o terreno.

3.3.1. Sensores multiespectrais que utilizam detectores discretos e espelho de varredura

Também conhecidos como sensores de varredura mecânica, são os de tecnologia mais antiga e que não são mais utilizados. São denominados de sensores tipo *whiskbroom*. Os sensores dos sete primeiros satélites do programa Landsat são os exemplos mais conhecidos de sensores de varredura mecânica. São sensores que usam um espelho externo para varredura do terreno e um conjunto de detectores discretos para cada banda. Nesses sensores, cada conjunto de 16 detectores de cada banda varrem 16 linhas simultâneas no terreno para compensar a velocidade de deslocamento da plataforma com a velocidade de rotação do espelho.

A principal característica desse tipo de sensor é o fato de o imageamento do terreno ser feito pixel a pixel ao longo de uma linha de várias dezenas ou centenas de quilômetros de extensão, no sentido transversal ao deslocamento do satélite. Daí serem também denominados de sensores *cross-track* (Figura 3.2A). O espelho oscilante de varredura do terreno está alinhado a um telescópio que direciona o fluxo radiante refletido e emitido dos objetos ao longo de uma linha de varredura para os componentes de dispersão da REM e a seguir para: 1) o plano focal primário do visível e do infravermelho próximo, onde estão posicionados os conjuntos de detectores de silício;

2) o plano focal resfriado do infravermelho de ondas curtas onde estão posicionados os conjuntos de detectores de antimônio de índio ou outro material; e 3) o plano focal do infravermelho termal com conjuntos de detectores de telureto de mercúrio e cádmio, ou outro material.

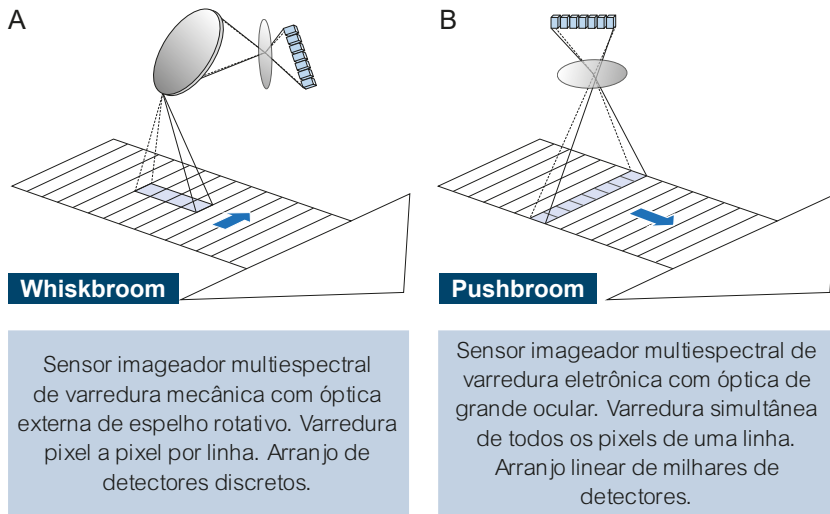


Figura 3.2. Configuração esquemática do mecanismo de imageamento de um sensor de varredura eletrônica, em que todos os pixels de uma linha são imageados simultaneamente.

Nesse processo, o sensor obtém, por meio de dois sistemas separados de detecção, imagens da luz solar refletida e imagens da radiação termal emitida pela Terra. A designação varredura mecânica é devido à existência, no alinhamento de abertura do telescópio, de um espelho plano de berílio inclinado em 45° e ligado a um eixo rotor que imprime uma altíssima rotação ou oscilação no sentido transversal ao deslocamento da plataforma. É esse movimento rotativo ou oscilante do espelho que varre o terreno em linhas perpendiculares à direção de voo. O movimento oscilatório do espelho permite que o sensor escaneador meça a radiação refletida ou emitida de um lado a outro da faixa do terreno, ao longo da trajetória da órbita da

plataforma. O ângulo de oscilação do espelho, de cerca de 5° a 10° , determina a largura da faixa imageada no terreno. Esse ângulo de oscilação é denominado de campo de visada (FOV). Conforme a plataforma avança, linhas sucessivas são varridas até compor uma imagem. Esse mecanismo é similar às imagens produzidas pelo processo de varredura de linhas utilizado pelos escaneadores de mesa.

3.3.2. Sensores multiespectrais que utilizam arranjos lineares

Esses sensores diferem dos sensores de varredura mecânica por não possuírem um espelho rotativo e por terem um conjunto de detectores que opera de modo diferente dos detectores discretos. Os sensores multiespectrais são de tecnologia mais moderna que os sensores de varredura mecânica. A varredura é do tipo eletrônico e são denominados de sensores tipo *pushbroom*. Como exemplos, têm-se os sensores dos satélites Landsat 8, Landsat 9, Sentinel-2, WorldView e *China-Brazil Earth Resources Satellite* (CBERS), entre outros. Os sensores de varredura eletrônica possuem arranjos lineares de milhares de detectores que usam diodos ou dispositivos de cargas acopladas (CCD) que nada mais são do que um chip de metal semicondutor montado em matrizes lineares ou bidimensionais.

A óptica é composta por uma grande objetiva com espelho plano que dispensa o uso de espelho rotativo. De um modo semelhante ao mecanismo dos sensores de varredura mecânica, os sensores de varredura eletrônica registram a imagem ao longo de uma faixa no terreno, por meio de varreduras de linhas sucessivas orientadas ortogonalmente à direção de trajetória da plataforma. A diferença entre esses dois tipos de varredura está na maneira como cada linha é varrida. Por meio da óptica de grande angular, cada linha é imageada no terreno em um único instante, utilizando-se do arranjo de milhares de detectores acoplados a um circuito eletrônico, o CCD. A configuração de um sensor de varredura eletrônica é apresentada na Figura 3.2B. Cada detector registra a energia de um único elemento de área

do terreno (pixel) ao longo de uma linha de varredura. Portanto, a amostragem do terreno não é feita pixel a pixel, mas todos os pixels de uma linha são imageados simultaneamente. À medida que a plataforma se desloca ao longo de sua trajetória de voo ou órbita, linhas sucessivas são imageadas no terreno pelo arranjo dos detectores. São assim, também denominados de sensores *along-track*.

Como os detectores CCD têm tamanhos minúsculos, um único arranjo pode conter dezenas de milhares de detectores individuais. Por exemplo, para adquirir uma imagem com faixa de imageamento no terreno de 60 km de extensão e tamanho de pixel de 10 m, o arranjo linear de CCD será composto de 6.000 elementos para cada banda. Os arranjos dos detectores são normalmente localizados no plano focal do sensor, de modo que todas as linhas varridas são vistas simultaneamente por todos os arranjos das várias bandas. Nesse sistema de varredura linha a linha, e não pixel a pixel, tem-se a vantagem de maior tempo de permanência para a medida da radiância que deixa cada pixel, gerando uma imagem com melhor relação sinal/ruído. A desvantagem é a enorme taxa de dados que é coletada a cada instante, o que é compensado por uma menor largura da faixa de terreno imageada por órbita. Por ser uma tecnologia mais avançada que os sensores mecânicos, exatamente porque dispensa o movimento de varredura oscilatório de um espelho, a tecnologia de sensores *pushbroom* de arranjo linear tem sido, atualmente, a melhor opção.

3.3.3. Sensores multiespectrais que utilizam arranjos lineares e de área

Os sistemas sensores que utilizam arranjos lineares e de área são mais conhecidos como sensoriamento hiperespectral. O nome advém do fato de adquirirem imagens com centenas de bandas espectrais relativamente estreitas (≤ 10 nm), na faixa espectral do visível ao infravermelho de ondas curtas. São também conhecidos pela denominação de espectrometria de imageamento pelo fato de cada pixel ser amostrado em centenas de com-

primentos de onda, podendo fornecer um espectro de reflectância de alta resolução do pixel que se assemelha aos espectros de reflectância obtidos em laboratórios por espectrorradiômetros. O espectro de reflectância pode ser usado para identificar, no pixel, a classe de material e sua composição físico-química.

O sistema de imageamento dos sensores hiperespectrais pode ser do tipo *whiskbroom* de arranjo linear ou do tipo *pushbroom* baseado no uso de CCD em arranjo de área, que é uma matriz bidimensional de milhares de elementos detectores. Pelo fato de poderem gerar espectros de reflectância, acuradas calibrações radiométricas são exigidas para esses sensores. A análise dos dados (*data centered*) demanda o uso de sofisticados algoritmos de processamento por causa da necessidade de converter os dados de radiância para valores de reflectância a fim de remover os efeitos da atmosfera. Para a análise dos espectros dos pixels, existem algoritmos que os comparam com espectros de bibliotecas espectrais de rochas, minerais, solos e vegetação, dentre outros alvos, que são obtidos em campo ou laboratório, os quais ajudam a determinar o tipo de material constituinte do pixel.

3.4. Radiância e Reflectância

A forma como os sensores multiespectrais medem a REM é praticamente a mesma, mas, dependendo da faixa espectral que o sensor opera, há leis específicas que descrevem e calculam as relações entre a energia eletromagnética que incide no alvo e dele se direciona ao sensor para construir a imagem do objeto. São chamadas de leis radiométricas e são fundamentais para explicar o significado físico dos valores da REM medidos pelo sensor. Para os sensores multiespectrais da faixa óptica, os conceitos expressos por essas leis são conhecidos pelos termos radiométricos irradiância, radiância e reflectância. Na Figura 3.3, que mostra um modelo básico de sensoriamento multiespectral, os números identificam cada um desses componentes radiométricos em uma ordem sequencial segundo a trajetória da REM da fonte até o sensor. O subscrito λ indica que o valor

calculado para a grandeza radiométrica que o contém é dependente do comprimento de onda.

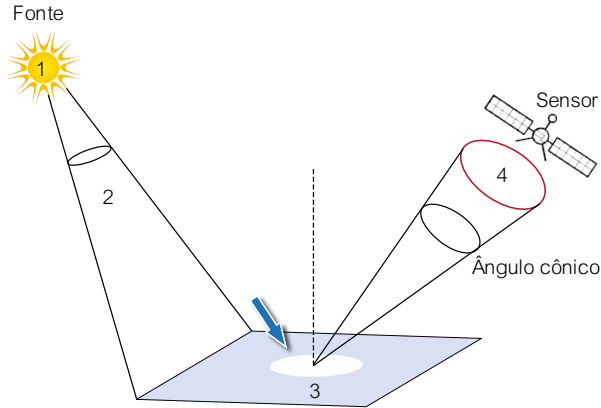


Figura 3.3. Componentes radiométricos da trajetória da radiação eletromagnética (REM) da fonte ao sensor. (1) = fonte de REM; (2) = camada atmosférica; (3) = alvo; (4) = sistema sensor.

- 1) Energia radiante (Q): energia que se propaga da fonte na forma de ondas eletromagnéticas, medida em Joule (J).
- 2) Fluxo radiante ($\Phi = \delta Q / \delta t$): taxa em que a energia radiante é transferida de um ponto ou superfície para outra superfície, medida em Watts por um determinado período de tempo. Também é a taxa de tempo que o fluxo ilumina a superfície do terreno imageada pelo sensor.
- 3) Irradiância ($E_\lambda = \Phi / \Delta A$): fluxo radiante solar incidente na superfície do terreno por área de superfície, medido em Watts por metro quadrado ($W m^{-2}$). No entanto, se efetivamente quiser medir a quantidade de irradiância incidente por área, deve-se integrar as variáveis da atmosfera pela qual a radiação passa. Segundo Jensen (2009), essa integração é uma função da irradiância solar espectral no topo da atmosfera ($E_{o\lambda}$), multiplicada pela transmitância atmosférica ($T_{\theta\theta}$)

em um certo ângulo zenital solar (θ_0), mais a contribuição solar difusa (E_{di}) (Equação 3.1). A unidade de medida é dada por $\text{W m}^{-2} \mu\text{m}^{-1}$.

$$E_{g\lambda} = \int_{\lambda_1}^{\lambda_2} (E_{o\lambda} T_{\theta_0} + E_{di}) d\lambda \quad (3.1)$$

- 4) Radiância ($L_\lambda = \Phi / \Delta\Omega \Delta A \cos\theta$) é a medida feita pelo sensor da densidade de fluxo radiante que deixa um elemento de área da superfície do terreno e que se propaga em uma direção definida por um cone elementar de um ângulo sólido contendo aquela direção. A unidade de medida da radiância é dada por $\text{W m}^{-2} \text{sr}^{-1} \mu\text{m}^{-1}$. Assumindo-se que a superfície da Terra é um refletor difuso perfeito, a quantidade total de radiância que deixa a superfície de um alvo em direção ao sensor é, de acordo com Jensen (2009), dada pela Equação 3.2.

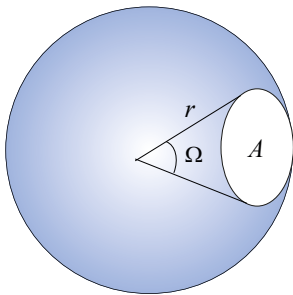
$$L_T = \frac{1}{\pi} \int_{\lambda_1}^{\lambda_2} \rho_\lambda T_{\theta_v} (E_{o\lambda} \cos\theta_0 + E_{di}) d\lambda \quad (3.2)$$

em que

ρ = refletância média da superfície do alvo; o fator π considera o espaço hemisférico.

O conceito de ângulo sólido é importante em sensoriamento remoto. Ele descreve a convergência ou divergência angular, em três dimensões, do fluxo para uma superfície ou de uma superfície. Ele é convenientemente definido em termos da geometria de uma esfera, como mostrada na Figura 3.4. O ângulo sólido subtendido no centro de uma esfera é definido como a razão da superfície de área da esfera pelo quadrado do raio da esfera. Como, em sensoriamento remoto, o ângulo sólido possui valor muito pequeno, isto é, de alguns poucos miliradianos, por conveniência, usa-se a área plana da base do cone, de diâmetro D . O sensor mede a intensidade ra-

diante que deixa a superfície de área do terreno limitada pelo ângulo sólido Ω . Esterradiano (*sr*), ou esferorradiano, é a unidade de medida de ângulo sólido, sendo que 1 esterradiano equivale a 57 graus plano. Além do ângulo sólido, a medida da radiância depende também da posição que o sensor se encontra no espaço, medida pelo cosseno do ângulo em relação a normal à superfície do terreno. Sendo assim, a imagem formada no sensor é uma imagem da radiância ao invés de imagem da reflectância, como se fala comumente. Os valores de radiância medidos pelo sensor são convertidos em valores digitais, gerando a imagem digital.



$$\Omega = A/r^2$$

em que

Ω = Ângulo sólido em esferorradiano;

A = Área da esfera delimitada pelo cone Ω ; e

r = raio da esfera

Figura 3.4. Conceito de ângulo sólido.

Há algo muito mais importante que deve ser considerado. Sabe-se que a densidade de fluxo incidente na superfície do terreno, isto é, a irradiância E , varia no tempo (hora, dia, mês), pois a distância e a posição do Sol em relação à Terra mudam continuamente, mesmo porque o Sol não é uma fonte totalmente uniforme. Além disso, a atmosfera é muito dinâmica, ou seja, a cada momento deixa passar quantidades diferentes de fluxo radiante devido às absorções dos gases. O que isso acarreta? A radiância que será medida no sensor para construir a imagem dos objetos poderá apresentar-se diferente de um dia para outro, mesmo que os alvos não tenham sofrido nenhuma mudança. Essa influência na variação da irradiância poderia ser eliminada se, no exato momento em que o sensor mede a radiância L_λ que deixa o alvo, medisse também a quantidade de fluxo radiante incidente no alvo, ou seja, a irradiância E_λ . A relação de medida da irradiância com a

radiância estabelece um conceito muito importante em sensoriamento remoto, que é a reflectância.

Reflectância é a razão entre a quantidade de energia radiante que deixa uma unidade de área no terreno (radiância) pela quantidade de energia incidente naquela área (irradiância), medida no mesmo instante de tempo. Como irradiância e radiância são densidades de fluxo, o valor dessa razão torna-se adimensional, sendo, portanto, expresso em porcentagem. Qualquer alvo no terreno poderá refletir uma porcentagem maior que 0% e menor que 100% da radiação nele incidente. Usa-se a letra grega ρ para se referir à reflectância (Equação 3.3).

$$\rho_{\lambda} (\%) = \frac{L_{\lambda}}{E_{\lambda}} \quad (3.3)$$

Em processamento de imagens, há algoritmos que convertem uma imagem digital em radiância para uma imagem de reflectância. As complexas formulações matemáticas que possibilitam fazer essa conversão de valores estão embutidas nos algoritmos de processamento de imagens para correção atmosférica. Para a maioria das aplicações de sensoriamento remoto e das técnicas de processamento de imagens, usam-se imagens no formato digital, aceitando-se que os valores digitais da radiância são representações bem próximas da reflectância. Razoáveis análises da reflectância espectral dos alvos podem ser feitas com imagens multiespectrais no formato digital. As imagens dos sensores hiperespectrais, que exigem, em suas análises, algoritmos de classificação espectral, devem ser convertidas para imagem de reflectância.

3.5. Sensores Termiais

Todo e qualquer objeto da superfície terrestre emite, em função de sua temperatura, REM no comprimento de onda do infravermelho termal. Isso

ocorre porque, no material, as moléculas se encontram em um estado de movimento vibracional. Quando essas moléculas se colidem entre si, elas mudam seu estado de energia e emitem REM. Os objetos terrestres são, portanto, fontes de REM, de forma que se pode medir a radiação termal emitida por eles com um sensor. A imagem termal é uma imagem de emitância da radiação. Por isso, ela é diferente da imagem de reflectância em termos de propriedades dos objetos que o sensor termal registra. A radiação termal medida em função da temperatura de superfície dos objetos é detectada pelo sensor em duas janelas atmosféricas: $3,0\text{ }\mu\text{m}$ – $5,0\text{ }\mu\text{m}$ e $8,0\text{ }\mu\text{m}$ – $14\text{ }\mu\text{m}$. A primeira janela é usada para fins meteorológicos, enquanto a segunda janela é a de sensoriamento remoto. O mecanismo de imageamento no termal é igual ao dos sensores multiespectrais da faixa óptica, podendo ser do tipo *whiskbroom* ou do tipo *pushbroom*. Por isso, muitos dos satélites de sensoriamento remoto adquirem, simultaneamente, imagens termais e imagens ópticas da mesma área. Assim, a aquisição de dados termais acaba sendo realizada em função do horário de obtenção das imagens ópticas.

Água, rocha, solo, vegetação e outros materiais não naturais têm a capacidade de internamente conduzir calor para a superfície (condutividade térmica) e armazenar calor (capacidade térmica). Essas propriedades se manifestam pela temperatura, dita temperatura cinética. Muitos dos usos das imagens termais têm interesses em detectar as temperaturas dos objetos como um evento anômalo na paisagem. São os casos, por exemplo, da detecção de plumas quentes em ambientes aquáticos, ilhas de calor de centros urbanos, diferenças entre solos úmidos e solos secos, incêndios florestais, derrame de lavas e atividades vulcânicas. Há também o uso das imagens termais com a finalidade de mapeamentos temáticos. Nesse caso, o objetivo é a espacialização da informação para criar mapas, seja de vegetação, rochas, solos ou água. Diferenças nas estruturas de dossel florestal, biomassa e condições fisiológicas, identificação de rochas com forte presença de minerais quartzosos e de litologias em função das fortes variações de suas propriedades térmicas, corpos de água que podem ter variações de temperatura pela presença de poluentes, sedimentos em sus-

pensão ou macrófitas, são facilmente identificados e até quantificados nas imagens termais.

Do mesmo modo que os sensores multiespectrais, os fótons do fluxo radiante do infravermelho termal emitidos pelos objetos são direcionados para os detectores termais. Esses detectores são feitos dos elementos Ge:Hg e Hg:Cd:Te, que são sensíveis aos comprimentos de onda de 8,0 μm a 14 μm . Esses detectores precisam ser resfriados a baixas temperaturas (-196 °C) usando hélio líquido ou nitrogênio líquido para assegurar que a energia detectada vem do terreno e não da temperatura dos equipamentos que compõem o sensor. Os detectores termais respondem à radiação neles incidentes por meio de um aumento de temperatura do material que são construídos. O aumento de temperatura provoca mudanças na resistência ou voltagem dos detectores, produzindo um sinal correspondente a essa mudança. A desvantagem é que eles têm uma lenta resposta de medida, resultando em uma demora no tempo de retorno para a temperatura base. Por isso, a taxa de amostragem no terreno é menor se comparada com a taxa de amostragem dos fótons detectores, e a sua resolução espacial também é menor. Nos sensores dos satélites Landsat 8, em que a imagem termal é obtida simultaneamente com as imagens do módulo multiespectral com resolução espacial de 30 m, a resolução espacial da imagem termal é de 90 m.

A faixa espectral termal de 8 μm a 14 μm usada no sensoriamento remoto é dividida em duas janelas devido a uma forte absorção atmosférica em 9,5 μm . Um dos exemplos de sensor termal multiespectral é o sensor *Advanced Spaceborne Thermal Emission and Reflection Radiometer* (ASTER) embarcado na plataforma Terra e que opera em cinco bandas espectrais:

- Banda 10: 8,1 μm –8,475 μm ;
- Banda 11: 8,475 μm –8,825 μm ;
- Banda 12: 8,925 μm –9,275 μm ;
- Banda 13: 10,25 μm –10,95 μm ; e
- Banda 14: 10,95 μm –11,65 μm .

Na Figura 3.5, é mostrada uma imagem obtida na faixa espectral do infravermelho, ao lado de outra imagem da mesma área obtida na faixa espectral óptica e tomadas simultaneamente em um horário diurno. Observe que dificilmente um especialista poderia prontamente reconhecer qual é a imagem termal. Isso porque qualquer imagem de qualquer comprimento de onda corresponde ao registro, em tons de cinza, das diferentes intensidades de REM que chegam ao sensor. Porém, a interpretação do significado dos tons de cinza de uma imagem do visível e de uma imagem do termal é baseada em parâmetros de análises diferentes. Enquanto as variações de níveis de cinza na imagem do visível correspondem às diferenças de reflectâncias entre os alvos, na imagem do termal, as variações de cinza correspondem às relações de temperaturas/emissividades dos alvos.

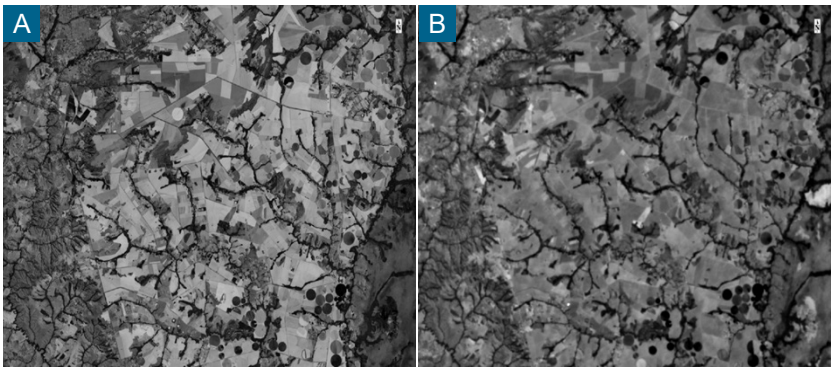


Figura 3.5. Imagens do satélite Landsat 8 OLI nas faixas espectrais do vermelho ($0,64\ \mu\text{m}$ – $0,67\ \mu\text{m}$) (A) e do termal ($10,6\ \mu\text{m}$ – $11,19\ \mu\text{m}$) (B).

A compreensão dos dados da imagem termal depende basicamente da compreensão do significado da relação física da temperatura cinética ou real do objeto com a temperatura medida pelo sensor por meio da REM, chamada de temperatura radiométrica. Os conceitos e leis radiométricas que definem os processos de interação da radiação termal com os objetos são abordados a seguir.

3.5.1. Leis da radiação termal

A percepção mais direta ou a primeira que temos com a radiação termal está relacionada com a emissão de raios solares. Como vimos no Capítulo 1, a REM deve ser entendida como uma onda que carrega diversas partículas de energia. Essas partículas ou fótons, ao serem absorvidas pelo material, aumentam o estado energético, levando o material a um estado que chamamos de excitado. Porém, o material não pode permanecer continuamente excitado e precisa retornar ao seu nível energético original. Para tal, é necessário que ele se “livre” dos fótons absorvidos, que são emitidos sob a forma de ondas longas, perceptíveis na faixa do infravermelho termal.

Os seres humanos não têm sensores ópticos sensíveis à percepção da REM do termal, pois nossa percepção visual está restrita à faixa da luz visível. No entanto, por meio do tato, podemos perceber a radiação calorífica ou termal que é emitida pelos objetos. Um exemplo disso é que se colocarmos dois indivíduos expostos ao Sol, sendo um com uma camisa branca e outro com uma camisa preta, não temos como avaliar a diferença de emissão de fótons dos dois tecidos, mas ao aproximarmos nossas mãos, sem a necessidade de tocá-los, podemos sentir que a camisa preta emite mais calor. Isso porque, ao absorver mais radiação que o tecido branco, o tecido de cor preta teve seu estado energético aumentado e, para retornar à normalidade, houve a necessidade de liberar maior quantidade de fótons absorvidos.

Quando pensamos em sensoriamento remoto termal, temos de pensar no comportamento térmico dos alvos. Para tal, é necessária a compreensão das leis físicas que regem esse comportamento. O primeiro aspecto que devemos pensar é o conceito de corpo negro como fonte de REM. A ideia de corpo negro foi desenvolvida por Kirchoff e designa o corpo que transforma em calor toda a radiação que sobre ele incide. Mais tarde, Planck, aluno de Kirchoff, desenvolveu a lei para explicar como funciona a radiação emitida por um corpo negro em todos os comprimentos de onda (Equação 1.6, Capítulo 1). Como na natureza não há nenhum alvo que seja 100% absorvedor, a solução encontrada para simulá-lo foi utilizar uma esfera oca

com o interior metálico refletor e com um pequeno orifício que permite a entrada da REM. Ao entrar pelo orifício, a radiação tem uma baixa probabilidade de sair, e como fica sendo espalhada pelo interior da esfera, tende a atingir o equilíbrio térmico. Cria-se, assim, a simulação do corpo negro. A quantidade de energia que qualquer corpo emite é, dentre outras coisas, função da temperatura do corpo. Essa propriedade pode ser expressa pela Lei de Stefan-Boltzmann (Equação 3.4).

$$M = \sigma T^4 \quad (3.4)$$

em que

M = emitância radiante da superfície do material (W m^{-2});

σ = constante de Stefan-Boltzmann ($5,6697 \times 10^{-8} \text{ W m}^2 \text{ K}^{-4}$); e

T = temperatura absoluta medida em graus Kelvin.

Por essa equação, notamos que há uma relação direta entre a emitância radiante e a temperatura. Essa relação não é linear, ou seja, um pequeno aumento da temperatura representa um grande aumento na emissão de radiação. Da mesma maneira que a energia total emitida por um corpo varia com a temperatura, a distribuição espectral da energia emitida também varia. O comprimento de onda no qual a emissão de energia é máxima pode ser determinado pela Lei de Wien, também conhecida como Lei do Deslocamento, expressa pela Equação 3.5.

$$\lambda_m = \frac{A}{T} \quad (3.5)$$

em que

λ_m = comprimento de onda de máxima emissão (μm);

$A = 2.898 \mu\text{m K}$; e

T = temperatura absoluta, medida em graus Kelvin.

De acordo com a Figura 3.6, se tomarmos a temperatura média do planeta Terra como de 300 K, um corpo negro a essa temperatura emitirá seu máximo de energia radiante em torno de 9,6 μm , enquanto a sua faixa como fonte de REM será de 8 μm a 35 μm . Por isso, podemos dizer que, se na faixa do espectro óptico refletido de 0,4 μm a 2,5 μm , o Sol é a principal fonte de REM, no espectro emitido, de 8 μm a 14 μm , a Terra é a fonte de radiação (além de 14 μm , a radiação termal é absorvida pela atmosfera). Resumindo, o Sol é a fonte de REM de ondas curtas e a Terra, a fonte de REM de ondas longas.

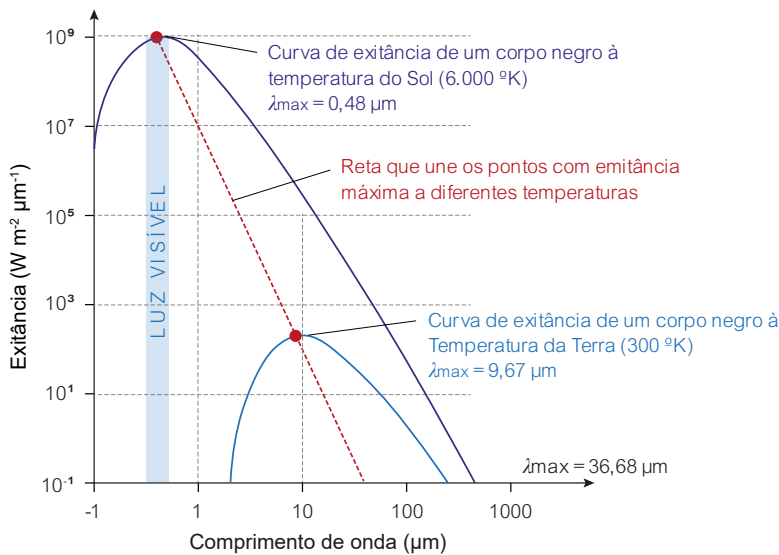


Figura 3.6. Deslocamento em direção aos menores comprimentos de onda do pico máximo de energia emitida, à medida que a temperatura do corpo aumenta.

Fonte: adaptado de Jensen (2009).

3.5.2. Emissividade

Os objetos terrestres como rochas, solo, vegetação ou água, não podem ser considerados um corpo negro radiador puro, que é apenas um modelo físico teórico. Se comparados com um corpo negro que é um perfeito ir-

radiador, veremos que esses objetos terrestres emitem apenas uma parte da energia que seria emitida por um corpo negro que estivesse à mesma temperatura. Essa relação de emissão de energia entre o corpo negro e os objetos terrestres define um importante conceito radiométrico no termal que é chamado de emissividade (ε). Emissividade representa a relação entre a radiância real emitida por um corpo do mundo real (M_r) e a do corpo negro (M_b) à mesma temperatura (Equação 3.6).

$$\varepsilon = \frac{M_r}{M_b} \quad (3.6)$$

A emissividade de um objeto é função de uma série de fatores, tais como cor, composição química, rugosidade da superfície, umidade e tantos outros. A emissividade depende da absorção pretérita de fótons e, para tal, é fundamental lembrar-se dos processos de absorção. Kirchoff observou que, na porção do infravermelho, a emissividade espectral de um objeto geralmente é igual à sua absorptância espectral. Por isso, podemos dizer que bons absorvedores são bons emissores. Se analisarmos as imagens dos objetos de coloração mais escura, ou seja, de baixa reflectância, deduzimos que eles são mais absorvedores. No termal, notaremos que eles são mais emissores, ou seja, apresentam maior emissividade que os alvos de coloração clara, com maior reflectância.

Vaughan et al. (2003) deduziram que se a emissividade geralmente é igual à sua absorptância e se os objetos têm transmitância zero à radiação termal, então a intensidade radiante é dada por (Equação 3.7):

$$I = \rho_\lambda + \varepsilon_\lambda \quad (3.7)$$

Essa equação demonstra que a energia que emana dos objetos tem uma relação entre a reflectância (ρ_λ) e a emissividade (ε_λ). Se a reflec-

tância é alta, a emissividade é baixa. Bons exemplos para ilustrarem essa propriedade é um corpo de água e um teto metálico. Como a água absorve quase toda a energia da luz solar e reflete muito pouco, a sua emissividade é muito alta, próxima a 1. Por outro lado, um teto metálico de um galpão tem alta reflectância e sua emissividade é muito baixa quando comparada à do corpo de água, como se pode ver na Tabela 3.1. Quando pensamos em materiais geológicos, ao analisarmos, por exemplo, espectros na faixa óptica refletida dos materiais que contém caulinita ($\text{Al}_2\text{SiO}_5(\text{OH})_4$) ou gibbsita ($\text{Al}_2\text{O}_3 \cdot 3\text{H}_2\text{O}$), notaremos que os materiais gibbsíticos apresentam bandas de absorção mais intensas, conseqüentemente, apresentam maior emissividade. Portanto, as variações de emissividade dos materiais terrestres, podem, nas imagens termais, revelar diferenças de composições dos materiais tal como nas imagens da região espectral óptica.

A umidade do material tende a reduzir o albedo. Um solo úmido é mais escuro que um solo seco e, com isso, maior é a sua capacidade de absorção da energia e de reemissão posterior de energia. Observe na Tabela 3.1 que um solo seco tem uma emissividade de 0,92. Quando o solo é umedecido, a emissividade aumenta para 0,95. Partículas de solo úmido têm uma alta emissividade, semelhante à da água. Superfícies com maior rugosidade têm uma maior área superficial e possuem maior potencial de absorção, o que as torna maiores emissores.

Como se pode observar na Figura 3.7, a água destilada é o material terrestre que apresenta comportamento termal mais similar ao do corpo negro na mesma temperatura, pois possui uma emissividade bastante próxima a 1, entre 0,98 e 0,99. Note que a curva de emissividade da água é bem próxima e similar à curva de emissividade do corpo negro. A maioria dos outros materiais apresenta um comportamento seletivo, tal como o quartzo mostrado nessa mesma figura.

Tal como as curvas de reflectância dos materiais, as curvas de emitância como as mostradas na Figura 3.7 servem para indicar as diferenças espectrais da emissividade dos diferentes materiais. Dessa forma, quando

analisamos uma imagem termal, não necessariamente nos interessa extrair informações de temperatura dos materiais da imagem. A informação da emissividade dos materiais seria a forma de identificá-los. Para fazer isso, é preciso resolver o problema da separação temperatura/emissividade, de tal forma que os efeitos da temperatura sejam subtraídos para que reste apenas a informação da emissividade dos alvos para ser analisada. Vários trabalhos foram realizados nesse sentido, descrevendo como a radiação termal emanada de superfícies rochosas pode ser usada em conjunto com a informação da emissividade derivada para fins de mapeamento geológico (Vaughan et al., 2003; Jacob et al., 2004).

Tabela 3.1. Emissividades médias de alguns materiais terrestres entre 8 μm a 14 μm de comprimento de onda.

Material	Emissividade
Água destilada	0,99
Água	0,92–0,98
Gelo seco	0,97–0,98
Vegetação verde	0,96–0,99
Vegetação seca	0,88–0,94
Gramma	0,97
Solo úmido	0,95–0,98
Solo seco	0,92–0,94
Neve	0,83–0,85
Areia	0,90
Asfalto	0,95
Basalto	0,95
Granito	0,86
Dunito	0,78
Metal polido	0,16–0,21
Alumínio polido	0,55
Ouro polido	0,02–0,03
Aço inoxidável	0,16

Fonte: Sabins (1997); Lillesand et al. (2008).

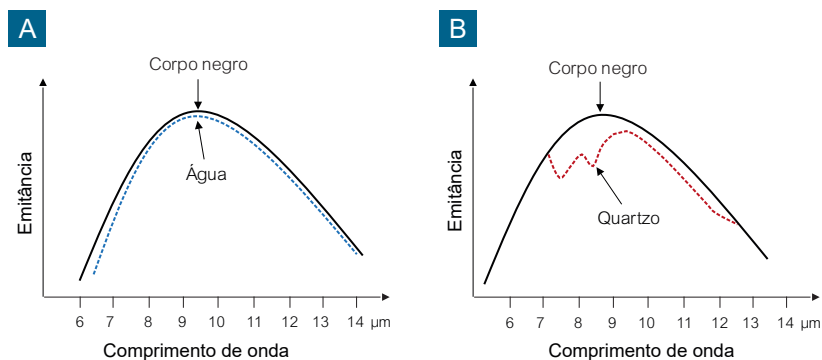


Figura 3.7. Comparação da emitância radiante espectral da água (A) e do quartzo (B) com a do corpo negro à mesma temperatura de superfície.

Fonte: Lillesand et al. (2008).

3.5.3. Temperatura cinética versus temperatura radiométrica

Devido à propriedade de emissividade dos corpos naturais, os sensores do infravermelho termal registram a intensidade da radiação equivalente à temperatura radiante aparente do objeto, ao invés da temperatura cinética verdadeira (T_{kin}), medida que se pode fazer com um termômetro. A REM que sai de um objeto é chamada de fluxo radiante e é medida em Watts. A concentração dessa quantidade de fluxo radiante que sai do objeto é chamada de temperatura radiométrica (T_{rad}) ou temperatura aparente. Portanto, quando um sensor termal mede essa energia emitida, ele não está medindo a temperatura cinética do objeto, como se faz com o termômetro, mas a temperatura radiométrica.

Em algumas aplicações envolvendo imagens termais, o conhecimento da temperatura do objeto pode ser importante. Alguns exemplos são o mapeamento dos padrões de nuvens e os movimentos de frentes frias, o monitoramento das temperaturas da superfície do mar, a detecção de ilhas urbanas de calor, focos de queimada ou mesmo derrames de lavas.

Nesses casos, é necessário determinar a relação entre a temperatura cinética com a temperatura radiométrica. Para fazer isso, utilizamos a Lei de Stefan-Boltzmann, vista na Equação 3.8, que é válida para objetos que são emissores perfeitos (corpos negros). Para compensar a diferença das fontes corpos negros com as fontes naturais, basta acrescentar a ela o valor de emissividade correspondente a cada material. Assim a equação de Stefan-Boltzmann deve ser reescrita como (Equação 3.8):

$$M_r = \varepsilon \sigma T^4 \quad (3.8)$$

Combinando as Equações 3.6 e 3.8, tem-se (Equação 3.9):

$$M_r = M_b \varepsilon \quad (3.9)$$

E assim, tem-se (Equação 3.10):

$$\sigma T_{rad}^4 = \varepsilon \sigma T^4 \quad (3.10)$$

de onde se obtém a relação entre a medida da temperatura radiante T_{rad} de um objeto na imagem com a sua temperatura cinética T_{kin} no terreno (Equação 3.11).

$$T_{rad} = \varepsilon^{1/4} T_{kin} \quad (3.11)$$

Como a emissividade ε varia entre 0 a 1, pode-se deduzir, da Equação 3.11, que a temperatura radiométrica medida pelo sensor sempre será inferior à temperatura cinética real do objeto.

3.6. Radar de Abertura Sintética

Os sensores de radar, um acrônimo de *radio detection and ranging* (detecção de ondas de rádio e posição), são do tipo ativo, isto é, geram a sua própria energia eletromagnética por meio de uma fonte artificial. Nesse caso, o sensor é, ao mesmo tempo, fonte e receptor da REM. Isso tem fortes implicações no sistema de imageamento do terreno. A região das micro-ondas de 3,2 cm a 100 cm, cujos comprimentos de onda chegam a ser da ordem de 200.000 vezes maiores que os comprimentos de onda da luz visível, tem, como principal vantagem, não ser barrada ou absorvida pelas micrométricas partículas ou gases da atmosfera. Essa condição permite obter imagens mesmo quando a cobertura de nuvens é total. Além disso, o imageamento também pode ser feito à noite. Um sensor de radar opera em uma única banda, sendo o único sensor que não tem características multiespectrais.

O entendimento das imagens de radar não é um procedimento tão intuitivo como o que se tem com imagens multiespectrais ópticas, as quais muito se assemelham com as imagens processadas pelo olho humano. Para interpretar uma imagem de radar, é preciso conhecer os vários fatores condicionados à visada lateral que implicam na obtenção da imagem, os quais não são de claro entendimento. Devido a isso, o processamento das imagens de radar requer um módulo distinto dos processamentos das imagens multiespectrais, a começar pela necessidade de primeiro ter que reamostrar os pixels distorcidos devido à visada lateral oblíqua para pixels de tamanho fixos.

Os primeiros usos de imagens de radar começaram na década de 1950 com fins militares. Nas décadas de 1960 e 1970, já eram extensivamente obtidas com aviões, mas somente nos anos de 1990 começaram a ser obtidas por satélites. Praticamente diferem em tudo dos sensores multiespectrais, a começar pela denominação das bandas espectrais. Elas são identificadas por letras, aparentemente sem nenhum significado, fruto de uma tentativa nos anos de 1940 de manter segredo militar das faixas espectrais

de operação (Figura 3.8). Quanto maior o comprimento de onda, maior é a penetração da onda nos alvos.

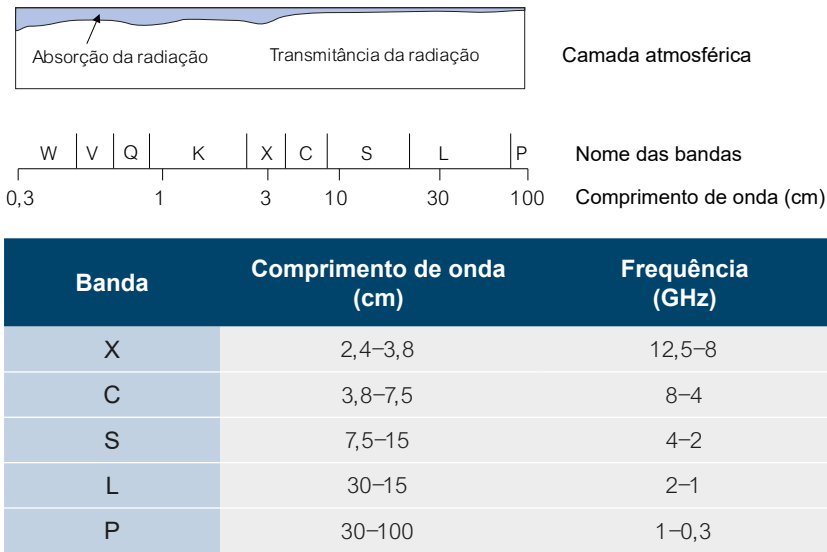


Figura 3.8. O gráfico mostra a transparência quase total das várias bandas de radar e no quadro os respectivos intervalos de comprimento de onda e as correspondentes frequências.

Considerando que, nos grandes comprimentos de onda das micro-on-das, a interação é do tipo macroscópica, a escolha do comprimento de onda do radar deve ser compatível com o tamanho das feições do alvo que se deseja identificar. Por exemplo, a banda de comprimento de onda X é melhor para identificar as variações texturais de solos, enquanto, para o mapeamento geológico, cujas feições do terreno são de grandes dimensões, é mais apropriado usar a banda L. No entanto, se quisermos obter uma maior penetração da onda na vegetação de uma floresta, uma banda de maior comprimento, como a banda P, maior do que o tamanho médio das folhas, é mais indicada.

3.6.1. O pulso de radar e a visada lateral

Diferente da iluminação solar, que consiste no envio de radiações paralelas e contínuas de luz (fótons) sobre uma superfície, o radar envia, por meio de uma antena, séries descontínuas de pulsos de fótons que se espalham sobre o terreno como um feixe angular na forma de um lóbulo. A transmissão é de um pulso eletromagnético de alta potência, de curto período de tempo e de feixe muito estreito. Tem o nome de pulso porque é enviado em intervalos regulares de tempo. Durante a propagação pelo espaço, o feixe se alarga em forma de cone ou de lóbulo até atingir o alvo que está sendo iluminado, sendo então espalhado (refletido em todas as direções). Feixes da energia espalhada retornam para a antena que, neste momento, torna-se receptora de sinais. São chamados de retroespalhamento. O lóbulo no terreno, em relação à direção de propagação da plataforma que transporta o sistema de radar, é amplo na dimensão transversal (dezenas ou centenas de quilômetros) e estreito na dimensão longitudinal da linha de vôo (Figura 3.9).

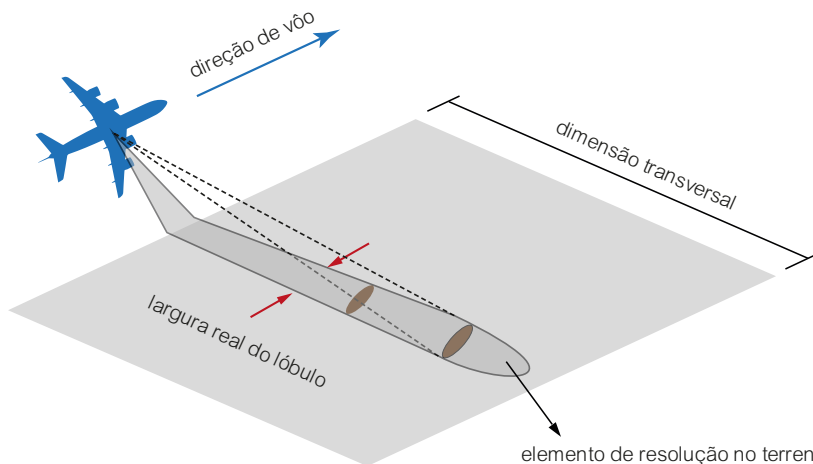


Figura 3.9. No imageamento lateral do terreno, um pulso de radiação eletromagnética emitido por uma antena de radar de visada lateral ilumina uma área na forma de um lóbulo.

O mecanismo se resume em enviar, em um intervalo de tempo programado, sucessivos pulsos de onda eletromagnética na direção do objeto, à medida que a plataforma se desloca, varrendo linhas sucessivas no terreno. Como a velocidade de propagação do pulso é conhecida (velocidade da luz, $3 \times 10^8 \text{ m s}^{-1}$), medindo-se o tempo de envio do pulso até o seu retorno à antena, o sistema calcula a distância em que o objeto está da antena e então registra na imagem a posição que ocupa no terreno (posição do pixel). Um lóbulo de iluminação é equivalente a uma linha de imageamento dos sistemas sensores multiespectrais. Para cada pulso registrado, a sua intensidade é medida. Para emitir esses pulsos de REM, os radares de sensoriamento remoto utilizam uma pequena antena retangular que é fixada na lateral da aeronave ou do satélite e que é apontada lateralmente em relação à direção da trajetória. Por isso são chamados de radares de visada lateral e a imagem é de visada oblíqua. A largura do feixe da antena determina a largura do lóbulo no terreno e o tamanho da antena é um parâmetro que implica no cálculo da resolução espacial. Quanto maior o tamanho da antena, maior a resolução espacial. Devido à impossibilidade de colocar grandes antenas nas plataformas, a engenharia construiu antenas sintéticas que simulam o tamanho efetivo de uma grande antena. São, por isso, conhecidos como radares de abertura sintética (*synthetic aperture radar-SAR*).

A necessidade de uma visada lateral é demonstrada na Figura 3.10, que mostra a incoerência da visada vertical para detectar as posições e distâncias dos alvos em relação ao sensor. Essa é outra diferença básica de imageamento por radar em comparação às imagens multiespectrais de visada vertical. Nessa figura, o alvo A3 e os pares de alvos (A1, A5) e (A2, A4) estão em posições diferentes no terreno. Quando o pulso vertical de onda do radar chega ao terreno, primeiro ele encontra o alvo A3 e faz o registro de sua posição na imagem, como se ele estivesse em uma posição do terreno mais próxima da aeronave. A seguir, o pulso de radar atinge simultaneamente os alvos A2 e A4, registrando os pulsos refletidos com sobreposição de tempo e, portanto, de distância. Para o sistema, os alvos A2 e A4 estão na mesma posição do terreno, sobrepostos. O mesmo se repetirá em um terceiro

momento de tempo para os alvos A1 e A5. Esses pares de alvos são assim registrados na imagem com posições distorcidas, tal como mostra o lado direito da Figura 3.10. Essa incoerência é conhecida como ambiguidade e por isso imagens de radar não podem ser tomadas na vertical.

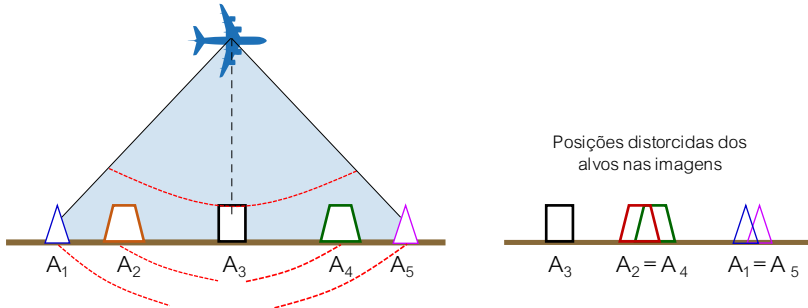


Figura 3.10. Ilustração mostrando o motivo pelo qual não existem sensores de radar com visada vertical.

3.6.2. Medida do tempo de transmissão do pulso

A configuração do envio de pulsos na direção de visada lateral, como é feito por todos os radares imageadores, é mostrada na Figura 3.11. A antena transmite os pulsos de micro-ondas em intervalos de microssegundos (cerca de 1.500 pulsos por segundo) que vão se propagar no lóbulo de iluminação como frentes da onda em sucessivos incrementos de tempo (1 a 8). No tempo 6, o pulso transmitido alcança a casa, e a frente de onda de retorno (retroespalhada) tem início no tempo 7. No tempo 12, o sinal de retorno da casa alcança a antena e é registrado com esse tempo no gráfico de resposta da antena. No tempo 8, a frente de onda transmitida é refletida da árvore e o seu “eco” alcança a antena no tempo 16. Medindo-se eletronicamente o tempo de retorno do sinal, são determinadas as distâncias ou as posições (range) entre o transmissor e os dois objetos refletoras. Como a casa retroespalha o pulso de onda de radar com intensidade maior do que a árvore, a intensidade da resposta do sinal registrado é mais forte. É o mesmo que dizer que a reflectância da casa é maior que a da árvore.

Apenas que o valor dessa reflectância é medida de maneira diferente aos dos sensores multiespectrais.

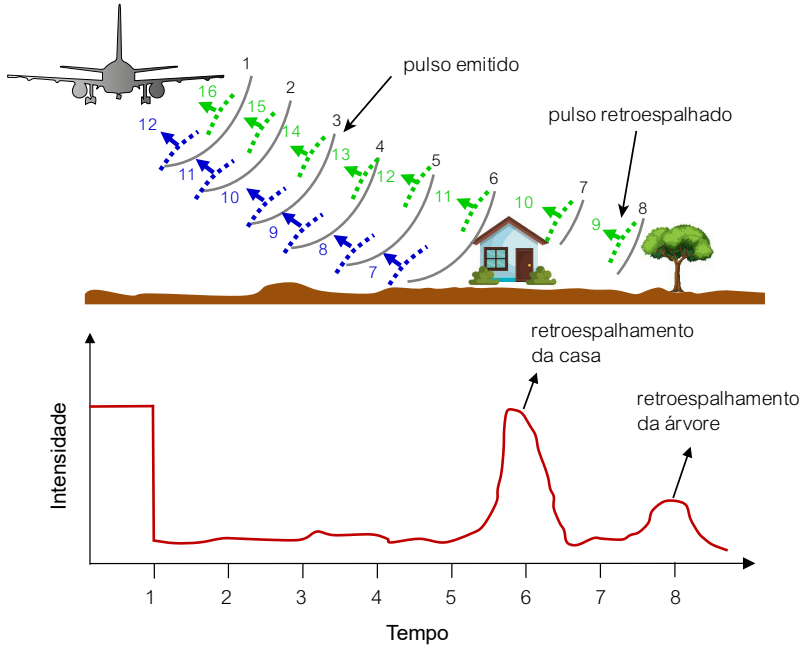


Figura 3.11. Modelo de propagação de pulsos de onda por um radar de visada lateral e o retorno dos pulsos para a antena, após serem retroespalhados pelos alvos.

Fonte: adaptado de Lillesand et al. (2008).

Em todos os momentos, são medidos os tempos de transmissão e de retorno dos pulsos. Por sabermos que a REM, independente do seu comprimento de onda, se propaga à velocidade da luz, podemos calcular a distância do objeto medida no plano inclinado (da antena ao alvo) usando a Equação 3.12.

$$SR = \frac{Ct}{2} \quad (3.12)$$

em que

SR = distância no alcance inclinado (*slant range*);

C = velocidade da luz; e

t = tempo entre o pulso transmitido e recebido.

O tempo é medido pela fase de onda que tem a sua referência em um ponto de partida relativo a um tempo arbitrário no momento em que o pulso é transmitido. Como uma onda é um fenômeno cíclico, a fase é frequentemente medida em graus, onde 90° representa $1/4$ de um ciclo e 360° um ciclo completo. Quando o pulso é transmitido, registra-se o grau da fase no tempo $t = 0$, como mostra a Figura 3.12. A fase do sinal de retorno é comparada com a fase do sinal transmitido e o tempo de retorno do sinal é calculado, determinando-se assim, a posição do objeto no terreno.

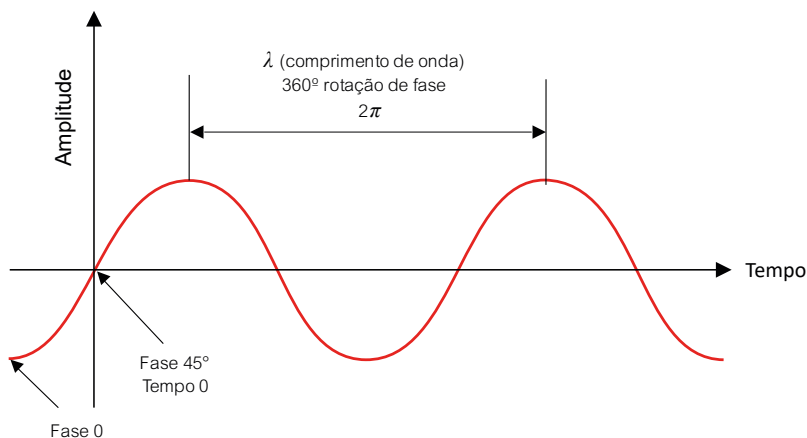


Figura 3.12. Medida do tempo pela fase da onda quando o pulso de radar é transmitido.

A antena funciona duplamente, como uma fonte transmissora da REM e como receptora da radiação de retorno. Uma chave liga a antena para emitir o pulso de alta potência, marca o tempo, e logo a seguir a desliga para registrar o sinal de retorno do pulso que foi enviado e o seu correspondente

tempo de retorno. O sinal de retorno do terreno é bem mais fraco que o enviado e deve ser amplificado significativamente. No sensor de radar a bordo do satélite SeaSat, o pulso transmitido tinha potência de 50 W, enquanto a potência efetiva recebida pela antena de um objeto tendo uma seção de área transversal ao pulso de 10 m^2 (pixel) era cerca de 10^{-17} W .

3.6.3. Polarização da onda

As ondas corpusculares de fontes de radiação como o Sol são naturalmente despolarizadas. No radar, quando um pulso de energia é emitido pela antena, o vetor do seu campo elétrico é controlado por um cristal polarizador com eixo de transmissão situado na horizontal ou na vertical para fazer vibrar a onda transmitida em uma direção horizontal (H) ou vertical (V). Essa onda é chamada de polarizada. A maioria dos pulsos é plano-polarizada, o que quer dizer que se o pulso é emitido na direção H, o retorno do pulso também é registrado na direção H. Diz-se, então, que a polarização é HH. Da mesma maneira, tem-se a polarização plano-paralela VV. A polarização HH é mais aconselhada para as superfícies terrestres enquanto a polarização VV é mais indicada para as superfícies (ondas) do mar. A Figura 3.13 mostra a configuração de polarização de pulsos horizontal e vertical.

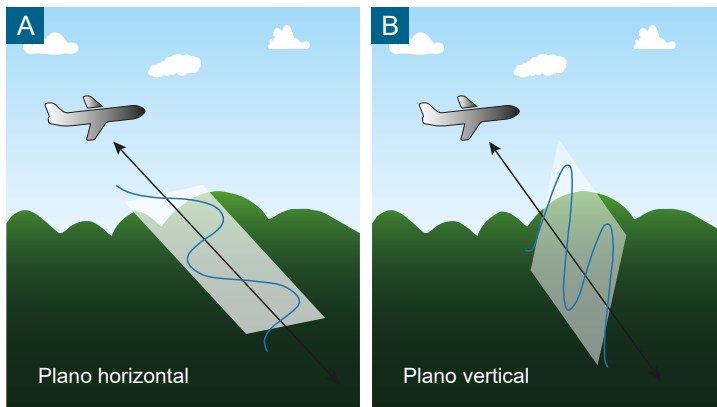


Figura 3.13. Polarização do sinal de radar enviado pela antena nos planos horizontal (A) e vertical (B).

Quando o pulso atinge o alvo, ocorre a despolarização, de modo que o sinal enviado pode retornar à antena com uma intensidade maior em uma direção diferente da direção de polarização emitida. Por exemplo, um sensor que emite um pulso na direção H pode registrar o sinal de retorno na direção polarizada V. Nesse caso, tem-se o que se denomina de polarização cruzada HV. Se o sinal polarizado é emitido em V e o sinal de retorno é registrado pela antena em H, tem-se a polarização cruzada VH.

Atualmente, muitos sensores de radar operam com as combinações de polarizações paralelas e cruzadas, possibilitando a aquisição de imagens em quatro modos de polarização: HH, HV, VV e VH. Isso é equivalente a quatro bandas, mas não podem ser assim denominadas, porque as quatro imagens polarizadas estão todas no mesmo comprimento de onda. A vantagem é que as imagens multipolarizadas aumentam as chances de discriminação dos materiais em função das variações de rugosidade da superfície. A vegetação, em particular, tende a mostrar diferentes respostas nas polarizações HH, HV (ou VH) e VV, por causa da despolarização da REM pelos múltiplos retroespalhamentos dos galhos e folhas, como se vê nas imagens da Figura 3.14.

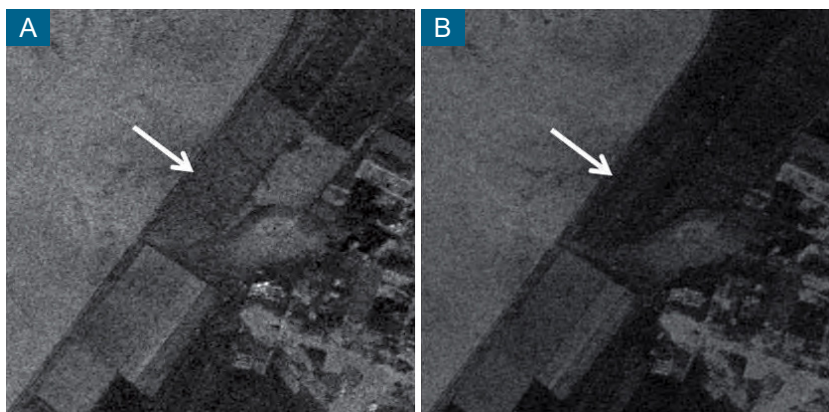


Figura 3.14. Diferenças das tonalidades de cinza dos alvos de uma área agrícola nas imagens de radar com duas polarizações diferentes, HH (A) e HV (B), obtidas pelo satélite japonês ALOS PALSAR.

3.6.4. Radar de abertura sintética

A Figura 3.15 mostra que um pulso de radar configura um lóbulos no terreno com largura estreita no início do pulso e larga no final. Nesse caso, diz-se que é um radar de abertura real. Não se tem uma linha de largura uniforme como nos sensores ópticos e termais. Isso implica que a resolução espacial das imagens de radar varia de valor ao longo do lóbulos. De fato, devido à configuração do lóbulos no terreno, há duas resoluções espaciais, uma medida na direção de iluminação do pulso e outra na direção transversal ao lóbulos. São ditas resolução em alcance (*range*) e resolução azimutal ou resolução longitudinal e transversal ao lóbulos. Os cálculos dessas resoluções dependem da duração de tempo que a antena emite os pulsos de energia, da largura angular do feixe do pulso enviado pela antena e do tamanho da antena. Para solucionar o problema da medida da resolução espacial do radar de abertura real, foram concebidos os radares de abertura sintética.

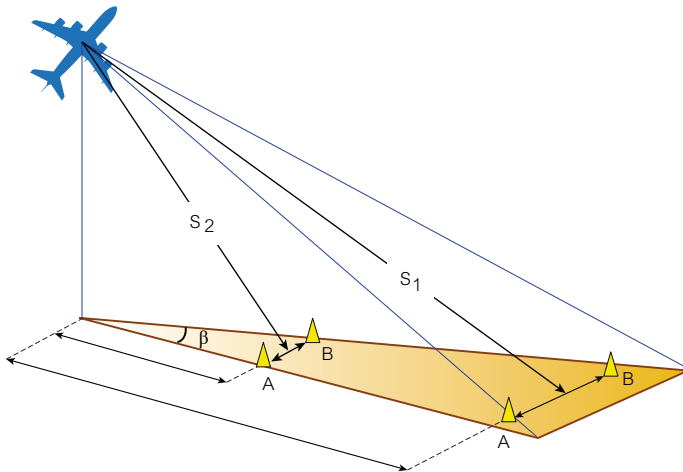


Figura 3.15. Em um radar de abertura real, a medida da resolução espacial varia ao longo da distância longitudinal do lóbulos e da largura transversal do lóbulos.

Os radares de abertura sintética, que são a tecnologia hoje utilizada para todos os sistemas de radares transportados em aviões ou satélites,

empregam uma antena de tamanho físico pequeno que sintetiza uma antena de tamanho de centenas de metros. Isso é conseguido por modificações nos registros dos dados e por técnicas de processamento de sinais. Em detalhe, essa operação é bastante complicada. É suficiente mostrar que se usa o movimento da plataforma ao longo da linha de voo para transformar uma antena pequena em uma matriz de antenas que são unidas matematicamente para detectar o alvo (Figura 3.16). A antena real de tamanho pequeno é mostrada em diversas posições sucessivas ao longo da linha de voo, fazendo com que o alvo seja visto várias vezes (*looks*), como está representado à direita da Figura 3.16. Um objeto posicionado em alcance próximo (*near range*) é visto, proporcionalmente, por menores números de elementos de antena. Quando na posição de alcance distante (*far range*), é visto por maiores números de elementos de antena. Essas posições sucessivas são tratadas matematicamente e eletronicamente como se elas fossem simplesmente elementos de uma única antena longa, resultando em uma resolução azimuthal constante ao longo do lóbulo do feixe de radar, independente da distância dos alvos à antena.

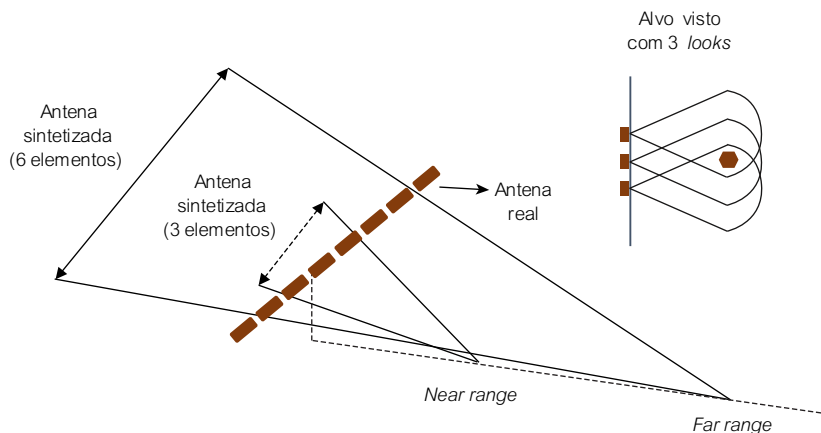


Figura 3.16. Múltiplos elementos de uma antena (*looks*) que sintetizam uma grande antena. Em *near range*, três elementos de antena e, em *far range*, seis elementos.

Fonte: adaptado de Lillesand et al. (2008).

3.6.5. Distorções geométricas da imagem de radar

A visada lateral das imagens de radar é como uma visada oblíqua. A imagem é originalmente construída pelo registro dos sinais medidos no plano inclinado, o que causa uma deformação dos alvos em todas as imagens de radar, deslocando, comprimindo ou achatando as formas dos objetos que se situam mais próximos à antena. Custosos processamentos corrigem esse tipo de distorção, transformando a imagem do plano inclinado (*slant range*) para o plano horizontal no terreno. Outros dois tipos de distorção, de caráter mais complexo, também ocorrem devido às medidas das distâncias dos objetos ao sensor serem feitas no plano inclinado. A inversão de relevo (*layover*) e o encurtamento de rampa (*foreshortening*) são esses dois tipos de distorções. Para entendê-los, basta observar a relação da inclinação da frente de onda com a inclinação das encostas de relevo voltadas em direção ao radar, como são apresentadas na Figura 3.17. Nesse exemplo, os morros são de relevos simétricos. Considerando-se que os radares utilizam medições do tempo de retorno do sinal, a duração de tempo que o feixe de radar leva para iluminar a frente de uma encosta, de sua base ao topo, é o fator que determina a sua posição e a sua dimensão na imagem.

No *layover*, a frente de onda atinge primeiro o topo do morro para depois atingir a base. Portanto, o sinal de retorno do topo do morro chega primeiro ao sensor. Como a imagem de radar determina a posição do alvo em função do tempo de ida e de retorno do sinal e sabendo que imagem é uma superfície em duas dimensões, o topo do morro aparece na imagem antes da base. Observe, na Figura 3.17, que a inversão de relevo ocorre nos morros que estão mais próximos do sensor, ou seja, com ângulos de depressão maiores. Por isso, a inversão de relevo é mais intensa em imagens de radar de satélites do que de avião. O efeito dessa distorção é mostrado pela imagem na Figura 3.18, onde o relevo aparece na forma de “V” com alto retorno do sinal (tons claros), mascarando por completo a forma real do relevo. Quando a frente de onda incide normal à face do relevo, não ocorre distorção.

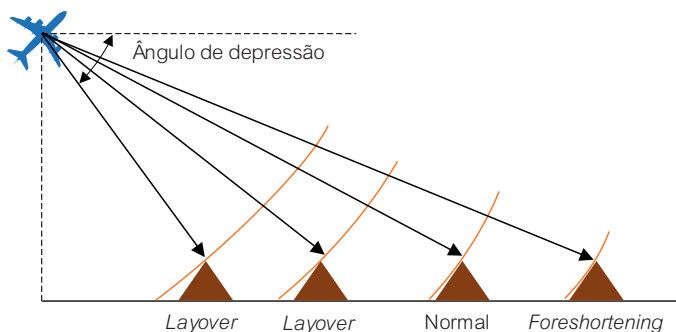


Figura 3.17. Ilustração das distorções geométricas em imagens de radar conhecidas como *layover* (inversão de relevo) e *foreshortening* (encurtamento de rampa). A primeira posição que a frente de onda toca o ponto no terreno é função da inclinação da frente de onda com a inclinação das faces de relevo. À exceção da posição normal, as demais indicam distorções na forma dos objetos que serão vistos na imagem de radar.

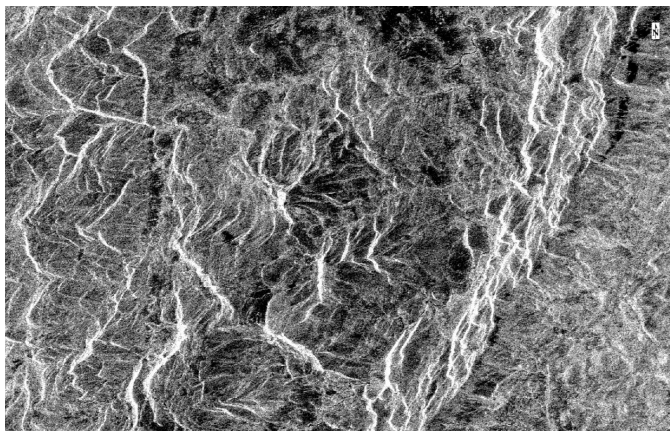


Figura 3.18. Formas em V de relevo como se fossem artefatos que surgem nas imagens de radar pelo efeito de inversão de relevo (ALOS PALSAR, Chapada dos Veadeiros, GO).

O encurtamento de rampa se expressa como uma compressão ou encurtamento da encosta frontal, relativo à encosta dorsal do relevo. Esse efeito ocorre quando o ângulo de incidência local (ângulo de incidência da

onda com o terreno) é menor que o ângulo de iluminação, mas maior que zero. Isso faz com que a face de encosta voltada para a antena apareça na imagem em tons bem claros, indicando forte potência do sinal de retorno, e não por causa de qualquer outra característica do relevo. Com a diminuição do ângulo de incidência local, a geometria faz com que a extensão da vertente (topo para base) diminua, aumentando assim, a intensidade do encurtamento da rampa.

Na Figura 3.19, as encostas frontais aparecem como faixas estreitas e alongadas, com tonalidades claras, enquanto as encostas dorsais aparecem com tonalidades escuras. Essas faixas alongadas e claras correspondem ao encurtamento de rampa. Se, por um lado, a visada inclinada traz esses inconvenientes, por outro lado, ela é a razão principal das imagens de radar serem consideradas como imagens muito eficientes para a análise das texturas de relevo. O raio de onda inclinado ilumina as faces frontais das encostas de relevo e geram sombras nas faces opostas. Essa condição de iluminação gera pares iluminado/sombreado (claro/escuro) que são responsáveis por criarem uma percepção visual tridimensional do relevo.

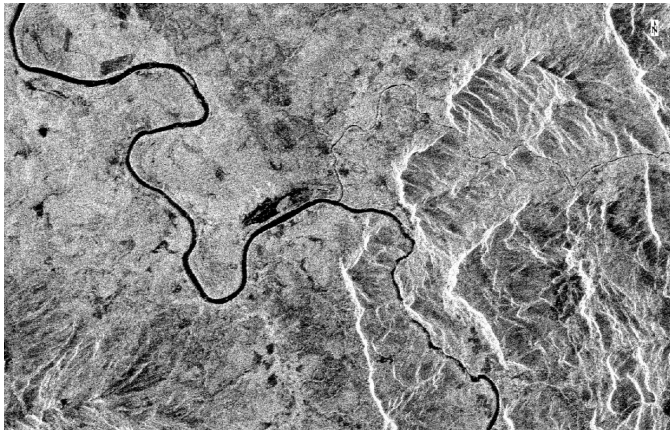


Figura 3.19. Feições de encurtamento de rampa (faixas estreitas, alongadas e claras) em uma imagem de radar do satélite japonês ALOS PALSAR da região de Chapada dos Veadeiros, GO.

3.6.6. Ruído *speckle*

A qualidade radiométrica das imagens de radar é afetada por um parâmetro instrumental que diz respeito à coerência de onda. Luz coerente é aquela formada por ondas de mesma frequência, fase e direção. Por isso, a luz solar é incoerente. No radar, a onda emitida é altamente coerente. Devido a essa propriedade, um tipo de ruído denominado *speckle* está sempre associado aos sensores de radar. O *speckle* é um ruído multiplicativo que é proporcional à intensidade do sinal recebido. Em uma célula de resolução no terreno, pode existir um número muito grande de alvos difusores, aleatoriamente distribuídos, de tal forma que esses alvos podem interferir uns aos outros, construtivamente e destrutivamente. A interferência destrutiva causa o cancelamento do sinal refletido e a interferência construtiva a sua soma, fazendo aparecer variações súbitas na intensidade da imagem que é o ruído *speckle* (Figura 3.20). O efeito visual na imagem é um padrão granulado do tipo sal e pimenta que degrada a qualidade visual da imagem para fins de interpretação (Figura 3.21). O *speckle* pode ser reduzido com múltiplas visadas (*looks*), porém, o aumento de *looks* causa degradação da resolução espacial.

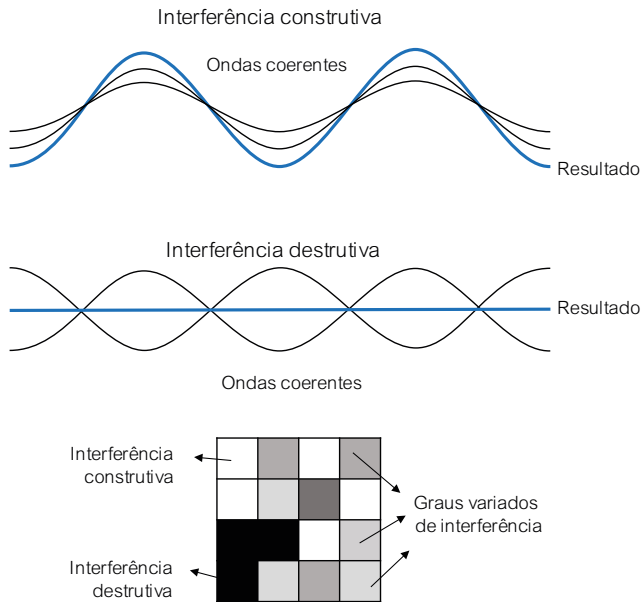


Figura 3.20. Interferência construtiva e destrutiva da onda coerente de radar após sofrer retroespalhamento.



Figura 3.21. Presença do ruído *speckle* (aspecto de sal e pimenta em toda a imagem) em uma cena do satélite canadense Radarsat-1 da região de Lagoa dos Patos, RS.

3.7. Referências

JACOB, F.; PETITCOLIN, F.; SCHMUGGE, T.; VERMOTE, E.; FRENCH, A.; OGAWA, K. Comparison of land surface emissivity and radiometric temperature derived from MODIS and ASTER sensors. **Remote Sensing of Environment**, v. 90, n. 2, p. 137-152, 2004.

JENSEN, J. R. **Sensoriamento remoto do ambiente: uma perspectiva em recursos terrestres**. EPIPHANIO, J. C. N.; FORMAGGIO, A. R.; SANTOS, A. R.; RUDORFF, B. F. T.; ALMEIDA, C. M.; GALVÃO, L. S. (Trad.). São José dos Campos: Parentese, 2009, 598 p.

SABINS, F. F. **Remote sensing**. Principles and interpretation. Nova York: Freeman, 3a ed., 1997, 494 p.

SCHOWENGERDT, R. A. **Remote sensing. Models and methods for image processing**. 3a ed., Amsterdam: Elsevier, 2007, 515 p.

VAUGHAN, R. G.; CALVIN, W. M.; TARANIK, J. V. SEBASS hyperspectral thermal infrared data: Surface emissivity measurement and mineral mapping. **Remote Sensing of Environment**, v. 85, n. 1, p. 48-63, 2003.

CAPÍTULO 4

Resoluções de Imagens

Paulo Roberto Meneses

Tati de Almeida



Atualmente, o sensoriamento remoto é constituído por uma razoável constelação de satélites que oferecem imagens para atender às necessidades de uma ampla demanda de usuários, o que é realizado por meio de medidas de resoluções que obtêm as propriedades dos objetos no terreno. Para os usuários que necessitam de uma observação detalhada do tamanho e das formas dos objetos, há os sensores que detectam áreas unitárias do terreno de tamanho menores que 1 m, sendo muito úteis para levantamentos cadastrais multifinalitários, urbanos e cartografia digital. Os interessados no acompanhamento da evolução e de mudanças da paisagem podem recorrer aos que, em um pequeno intervalo de dias, possibilitam uma alta taxa de revisita à área. E para os que se interessam em determinar a composição ou a constituição dos materiais rocha, solo, água ou vegetação, a procura é pelos sensores com um grande número de bandas espectrais. Por isso, o melhor modo para avaliar as potencialidades de um sensor é pelo dimensionamento de suas resoluções.

4.1. A Importância das Resoluções das Imagens

Nas aplicações de sensoriamento remoto mais comuns, que são as que envolvem mapeamento e análises geotemáticas, duas questões estão sempre presentes na escolha do sensor: qual é a melhor resolução da imagem para se identificar e resolver os objetos de interesse? E qual é a melhor escala para representar os objetos ou fenômenos geográficos? Nem sempre há respostas simples para essas duas indagações. O que mais prontamente pode-se responder é que as imagens dos sensores que recobrem grandes áreas, com largura de faixa de imageamento no terreno de centenas de quilômetros, são associadas com resoluções espaciais pequenas (20 – 30 m), e as imagens dos sensores que recobrem pequenas áreas, com largura de faixa de imageamento de dezenas de quilômetros, são associadas

com resoluções espaciais grandes (≤ 1 m). Consequentemente, as imagens de pequena resolução espacial, por não mostrarem os detalhes dos alvos, servirão para estudos em escalas cartográficas mais regionais, enquanto as imagens com grande resolução espacial se prestarão para estudos locais de detalhe. Essa relação entre a largura de área de imageamento e a resolução espacial acontece para equilibrar o volume de dados que é gerado pelos sensores.

Na prática, a detecção ou a identificação de um objeto nas imagens de sensoriamento remoto não é determinada somente pela resolução espacial, isto é, pelo tamanho do objeto, visto que qualquer objeto, para ser melhor investigado, é necessário conhecermos outras propriedades como a sua composição, cor, texturas, associações, ou até mesmo índices biológicos. Os sensores imageadores cumprem esta tarefa por meio da medida simultânea de quatro diferentes tipos de resoluções. Cada sensor tem o seu conjunto próprio e único de valores de resoluções e a mudança de apenas um de seus valores implicará na definição de outro tipo de sensor. Essas quatro medidas de resoluções são conhecidas como:

- Resolução espacial, a qual, pelo tamanho do campo de visada do sensor, é capaz de medir as dimensões dos objetos e suas relações espaciais;
- Resolução espectral, pela qual identificamos a cor e composição dos objetos, dependendo do número de bandas espectrais que o sensor possui;
- Resolução radiométrica, que define com que precisão a reflectância do alvo é medida, a fim de se ter uma imagem bastante nítida do objeto; e
- Resolução temporal, que define, em um intervalo de tempo, quantas vezes o alvo pode ser observado, a fim de se detectar mudanças ou acompanhar sua evolução.

Essas quatro medidas de resoluções são válidas para os sensores multiespectrais da faixa óptica e termal. Elas atuam em conjunto em um processo interativo nem sempre facilmente percebido quando se analisa a imagem. Para os sensores de radar, por serem ativos e terem visada lateral, as medidas de resoluções espaciais diferem completamente das medidas de resoluções dos sensores multiespectrais e são tratadas à parte por último.

4.2. Resolução Espacial

O tamanho individual do elemento de área imageada no terreno, chamado de pixel, representa, em qualquer tipo de sensor, um parâmetro importante da imagem: a resolução espacial. A resolução espacial determina o tamanho do menor objeto que pode ser identificado em uma imagem. Matematicamente, a resolução espacial é definida pelo espaçamento dos pixels no terreno, o que é nada mais que uma taxa de amostragem realizada pela projeção de um simples detector de tamanho w sobre o terreno, em função do valor do ângulo do campo de visão instantâneo (IFOV) da óptica do sensor. Na denominação em inglês, dá-se o nome desse espaçamento de pixel no terreno de *ground project sample interval* (GSI). A fórmula matemática da Equação 4.1 para calcular o GSI considera os seguintes parâmetros: espaçamento entre detectores discretos que é igual ao tamanho W do detector e que pode ser mencionado como espaçamento interdetectores; altura H que o sensor está no espaço; e distância focal (f) da óptica do sensor (Equação 4.1).

$$GSI = W \frac{H}{f} \quad (4.1)$$

É prática comum denominar a taxa de amostragem GSI como sendo igual ao campo de visada instantânea projetado no terreno (*ground-project instantaneous field of view* - GIFOV). Um pequeno IFOV é desejável quando há a necessidade de se registrar altos detalhes espaciais. Dessa forma,

quanto menor é o tamanho do IFOV, maior é a capacidade de se detectar, nas imagens, os menores elementos que integram a cena. A Figura 4.1 ilustra a configuração geométrica para o cálculo da resolução espacial segundo esses parâmetros.

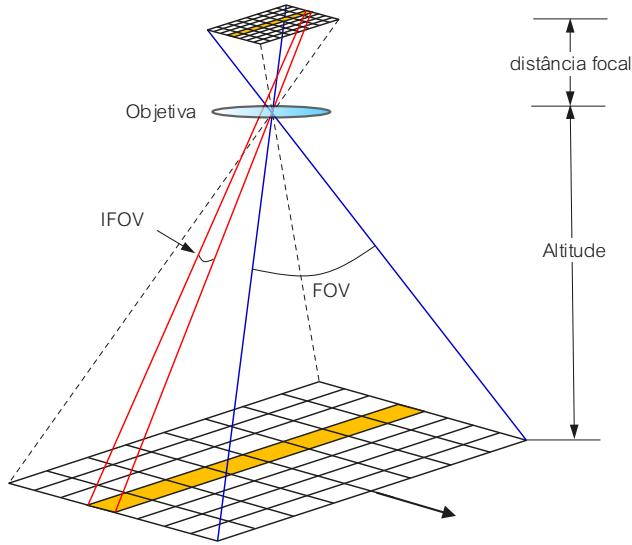


Figura 4.1. Medida geométrica da resolução espacial a partir do valor do ângulo *instantaneous field of view* (IFOV) e da altura H do sensor.

Nos sensores orbitais, a óptica é a de um telescópio portátil com distância focal de 2 a 3 m, que configura o valor do ângulo sólido β do IFOV em centésimos de miliradianos. Também podemos usar a medida de β para calcular a resolução espacial (Equação 4.2).

$$GSI = H\beta \quad (4.2)$$

Usando a Equação 4.2, para um sensor orbital colocado à altura de 700 km e com um IFOV igual a 0,043 miliradianos, o tamanho do elemento de resolução no terreno será (Equação 4.3):

$$D = 700.000 \times (0,043 \times 10^{-3}) = 30,1 \text{ m} \quad (4.3)$$

A título de ilustrar o efeito da resolução espacial na identificação dos objetos, a Figura 4.2 mostra imagens de três sensores ópticos com diferentes resoluções espaciais. Fica evidente que se pode estabelecer uma relação de comparação entre a resolução espacial e a escala de visualização da imagem.

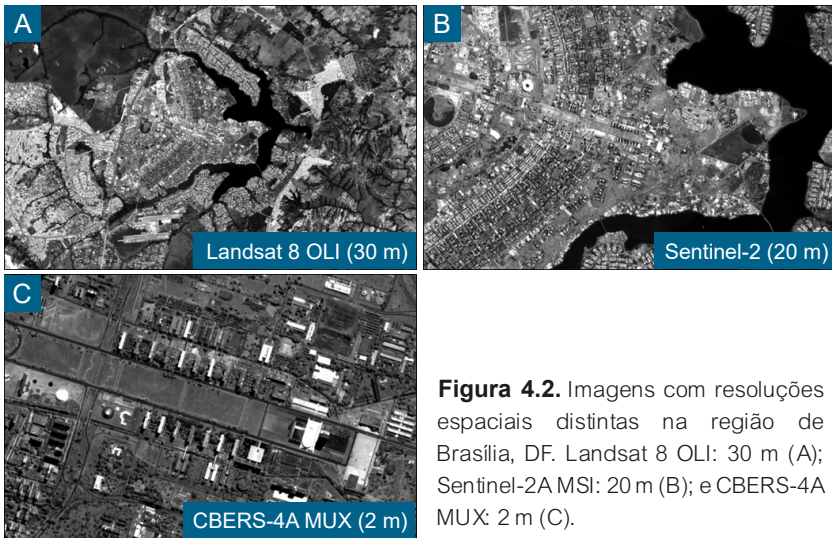


Figura 4.2. Imagens com resoluções espaciais distintas na região de Brasília, DF. Landsat 8 OLI: 30 m (A); Sentinel-2A MSI: 20 m (B); e CBERS-4A MUX: 2 m (C).

As equações anteriores nos indicam que um objeto somente pode ser resolvido (detectado) quando o seu tamanho é, no mínimo, igual ou maior do que o tamanho do elemento de resolução no terreno, ou seja, da resolução espacial. Portanto, se um objeto tem 20 m x 20 m de tamanho, a resolução espacial da imagem deveria ser, no mínimo, de 20 m para que pudesse ser identificado na imagem. Entretanto, a experiência mostra que, de fato, para um objeto ser resolvido na imagem, a resolução espacial nominal deve ser de pelo menos a metade do tamanho do objeto medido na sua menor

dimensão. Mesmo assim, o objeto ainda tem que apresentar um bom contraste de reflectância com os alvos que lhe são vizinhos, vistos na dimensão espacial do pixel. Isso porque o pixel representa uma média de duas dimensões: o espaço e o comprimento de onda definem as características dos dados. A terceira dimensão é o tempo que o sensor detém sobre o pixel para medir a reflectância, o que é muito pequeno para ser considerado. É como dizer que a informação que um pixel integra é uma sobreposição de ambas as dimensões, espacial e espectral.

A Figura 4.3 exemplifica esse conceito de resolução espacial segundo as relações dessas duas dimensões e vizinhanças dos pixels. Em (A), a resolução espacial possibilita detectar individualmente cada uma das duas árvores porque elas têm a mesma dimensão do pixel, mas na imagem, elas não são resolvidas porque, na dimensão espectral, os dois pixels têm o mesmo valor de cinza, fundindo-se como se fosse um único objeto. Em (B), as duas árvores são resolvidas na imagem por estarem separadas por um pixel de reflectância diferente.

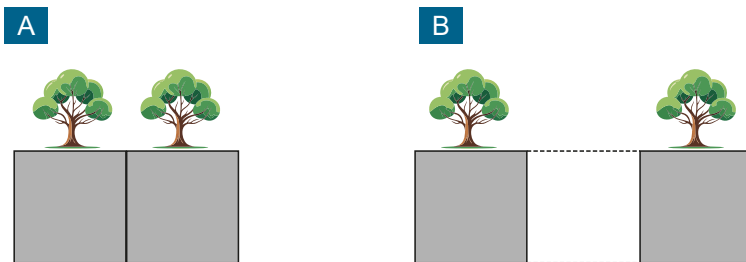


Figura 4.3. Em (A), situação em que a resolução espacial da imagem medida nas dimensões de tamanho e comprimento de onda não permite resolver os objetos e, em (B), situação que permite resolver os objetos.

Outra influência na resolução espacial é a localização relativa dos pixels e dos alvos. Há inúmeras possibilidades dessa relação acontecer. Por exemplo, um alvo de alta reflectância que ocupe 50% de um pixel pode ter metade dele localizado em um pixel e a outra metade no pixel vizinho

(Figura 4.4A), ou estar repartido em mais do que dois pixels, como se vê na Figura 4.4B. Considerando que o *background* da área seja de baixa reflectância, na primeira situação, o alvo pode ser detectado nos dois pixels. Na segunda situação em que o alvo ocupa 30%, 15%, 5% e 10% em quatro pixels adjacentes, o alvo pode ser detectado em três dos quatro pixels. Nessa mesma situação, mas considerando que as reflectâncias do alvo e do *background* não sejam muito diferentes, a possibilidade de detecção do alvo muda completamente. Em resumo, é imprevisível a detecção de um alvo mesmo que o poder de resolução espacial da imagem seja, por definição, capaz de detectá-lo.

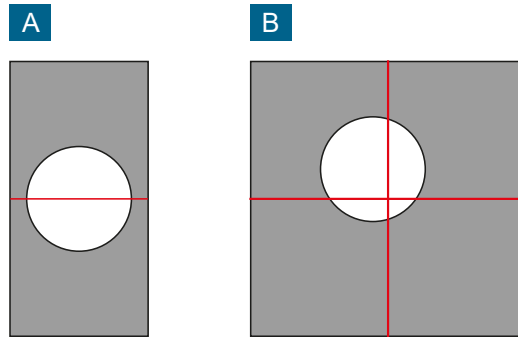


Figura 4.4. Efeito da localização relativa do pixel e do alvo no poder de resolução espacial da imagem.

Finalmente, devemos considerar que, para a maioria das aplicações de sensoriamento remoto, o poder de resolução espacial de um sensor envolve uma análise da relação do grau de autocorrelação da organização espacial dos objetos no terreno. Como regra geral, em terrenos naturais não urbanos, os alvos geralmente apresentam uma alta correlação espacial, mostrando pouca variabilidade ao longo de uma extensa área, não exigindo, para as suas identificações, altas resoluções espaciais. Já para uma área com alta variabilidade de tipos de objetos, como uma área urbana, a exigência seria para um sensor com resolução espacial de poucos metros, para que as casas, ruas, estacionamentos, possam ser resolvidos.

4.3. Resolução Espectral

Muitos dos satélites de sensoriamento não têm resolução espacial capaz de identificar objetos individuais pelo seu tamanho. Isso não é necessariamente um problema, porque o tamanho de um alvo não é a única propriedade para se identificar um objeto, na concepção do sensoriamento remoto multibanda. Por exemplo, várias árvores de um bosque podem caber dentro de um único pixel de 30 m x 30 m. Reconhecer que esse pixel é de vegetação e que pode ser diferenciado do pixel vizinho, que pode ser um solo exposto, é possível, se considerarmos as diferenças de reflectâncias da vegetação e do solo. Nesse caso, é a resolução espectral da imagem, isto é, o comprimento de onda, que soluciona o problema. Mas, para isso, é necessário o conhecimento auxiliar dos padrões de reflectância, no exemplo, da vegetação e do solo e, assim, para todos os objetos ou materiais. Para o sensoriamento remoto, a obtenção simultânea de imagens em múltiplas bandas espectrais é, sem dúvida, a propriedade mais importante dos sensores imageadores. O termo resolução espectral envolve pelo menos três parâmetros de medida:

- 1) número de banda que o sensor possui;
- 2) largura da banda; e
- 3) posição que a banda está situada no espectro eletromagnético.

Comparativamente, um sensor tem melhor resolução espectral se ele possuir maior número de bandas situadas em diferentes regiões espectrais e com larguras estreitas. Quanto maior for o número de bandas, maior é a possibilidade de identificar, nas imagens, pelas variações de reflectância, os diferentes tipos de materiais ou objetos presentes na paisagem. Quanto à largura da banda, ela deve ser dimensionada em concordância com as larguras médias das bandas de absorção dos materiais. Essas bandas de absorções podem ser identificadas nos espectros de reflectância de cada material. As bandas de absorções são identificadores da composição do objeto e normalmente elas têm a largura da ordem de 10 nm a 20 nm.

Sensores com bandas muito largas, além de 20 nm, são menos capazes de diferenciar objetos da mesma classe em função das variações de suas composições (por exemplo, espécies de vegetação).

A Figura 4.5 ilustra esses dois critérios de resolução espectral, número de bandas e largura das bandas, tendo, como exemplo, as bandas do sensor ASTER. Nove entre as 14 bandas desse sensor foram selecionadas para discriminar minerais de alteração hidrotermal. Os espectros de reflectância dos minerais de alteração hidrotermal plotados nessa figura mostram que a faixa espectral do infravermelho entre 2,0 μm a 2,5 μm é onde ocorre a maioria das bandas de absorção que permitem distinguir um mineral do outro. Das nove bandas, cinco foram posicionadas nessa região espectral. Como as bandas de absorção dos minerais são estreitas, as bandas do sensor ASTER são suficientemente estreitas para as imagens diferenciarem e identificarem esses minerais.

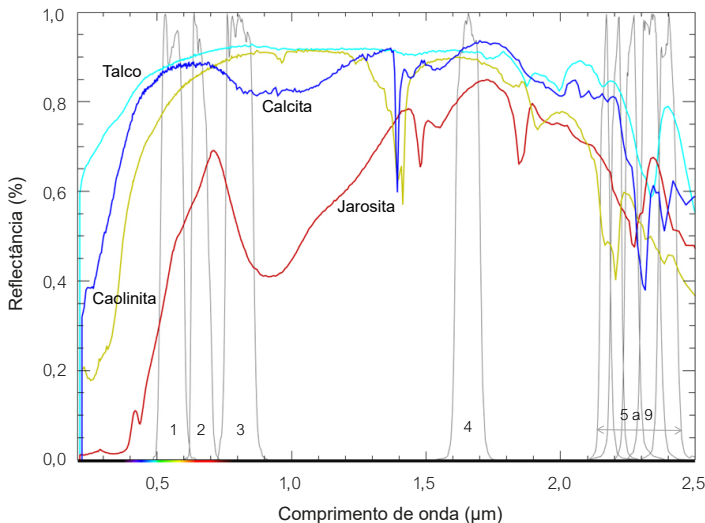


Figura 4.5. Espectros de reflectância de minerais calcita, caolinita, jarosita e talco, sobrepostos às nove bandas do sensor ASTER. A largura da banda do sensor deve ser equivalente à largura da banda de absorção do mineral para que este possa ser identificado na imagem.

Outro exemplo que serve para ilustrar o efeito da resolução espectral na detecção ou identificação de um objeto é mostrado na Figura 4.6, cujas imagens correspondem a uma área da floresta amazônica. Trata-se de imagens do satélite Landsat com resolução espacial de 30 m. Na imagem da esquerda da Figura 4.6, obtida na faixa espectral do vermelho ($0,636\ \mu\text{m}$ – $0,673\ \mu\text{m}$) pelo sensor *Operational Land Imager* (OLI) do satélite Landsat 8, um grande rio de centenas de metros de largura não é facilmente identificado. Isto porque as baixas reflectâncias da água e da vegetação são praticamente iguais neste comprimento de onda, exibindo um baixo contraste espectral (de reflectância) entre os dois tipos de alvos. Para a região amazônica, seria um erro selecionar essa banda espectral para elaborar mapas de drenagem. Já na imagem da direita da Figura 4.6, obtida na faixa espectral do infravermelho próximo ($0,851\ \mu\text{m}$ – $0,879\ \mu\text{m}$), os rios são facilmente identificados porque, enquanto a água possui uma baixa reflectância, a vegetação da floresta tem uma alta reflectância, estabelecendo uma razão de alto contraste entre os dois alvos.

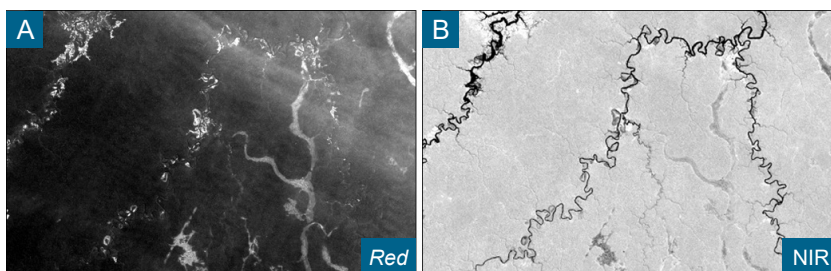


Figura 4.6. Imagens da região amazônica nas bandas do vermelho (A) e do infravermelho próximo (B) do satélite Landsat 8 OLI exemplificando o efeito da resolução espectral no contraste da água do rio com a vegetação.

Valendo-se do poder da resolução espectral é que os sensores multiespectrais aparentam ter a sua resolução espacial melhorada. É bem conhecido dos usuários de sensoriamento remoto que é possível detectar objetos que são consideravelmente menores que a resolução espacial, se

o contraste entre objetos vizinhos é suficientemente alto. O exemplo mais marcante é a identificação de estradas em imagens de satélites. Na Figura 4.7, o exemplo é novamente de uma imagem com resolução de 20 m, onde há estradas sem pavimentação com larguras inferiores a 10 m. Devido ao tamanho do pixel ser de 20 m, ele integra as áreas da estrada e da vegetação marginal, constituindo um pixel de mistura com um valor qualquer de reflectância. Tal valor não será nem o valor puro da reflectância da estrada e nem o da vegetação. Mas como na banda do visível a estrada sem pavimentação possui uma alta reflectância, muito superior à reflectância da vegetação, o valor médio da reflectância do pixel de mistura se aproximará mais do alto valor da reflectância da estrada do que do baixo valor de reflectância da vegetação. Nesse caso, os pixels que contêm a estrada e a vegetação terão um nível de cinza mais claro que os pixels puros escuros da vegetação.

Tal efeito de contraste apresenta-se como uma aparente melhoria na resolução espacial da imagem. Para confirmar esse efeito da resolução espectral, na resolução espacial, observe que, quando a banda espectral selecionada for de comprimento de onda do infravermelho próximo, a estrada é dificilmente identificada porque, neste caso, as reflectâncias da estrada e da vegetação são muito próximas, portanto, sem um bom contraste espectral.

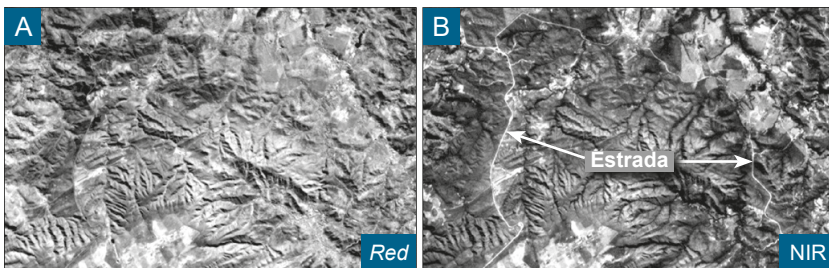


Figura 4.7. Efeito da resolução espectral na identificação de alvos (estradas de terra) em áreas de relativa densa vegetação, com larguras inferiores à resolução espacial de 20 m da imagem do sensor CBERS-4A MUX obtidas nas bandas espectrais do vermelho (A) e do infravermelho próximo (B).

Infelizmente, não é fácil aumentar, indefinidamente, qualquer um dos três parâmetros que medem a resolução espectral do sensor: largura das bandas, posição das bandas ao longo do espectro eletromagnético e número de bandas. Um sensor com centenas de bandas espectrais que permite a obtenção de uma amostragem detalhada do comportamento espectral da reflectância dos objetos aumenta significativamente a taxa de gravação ou de transmissão de dados do satélite para a Terra. Isso implica que apenas pequenos segmentos de área no terreno podem ser imageados na prática. De certa forma, é por essas restrições que se vê atualmente, nos exemplos dos sensores orbitais em operação, certo comprometimento nas medidas das resoluções espectral e espacial. Normalmente os sensores com alta resolução espacial, para compensarem o grande volume de dados que tem de ser transmitido para a Terra, possuem poucas bandas, limitadas às faixas espectrais do visível e do infravermelho próximo. São assim considerados sensores de alta resolução espacial e baixa resolução espectral. Ao contrário, um sensor de baixa resolução espacial pode ser compensado por um maior número de bandas.

4.4. Resolução Radiométrica

A medição dos detectores da intensidade de radiância de cada pixel é denominada de resolução radiométrica. Quanto maior for a capacidade do detector para medir as diferenças de intensidades dos níveis de radiância, maior será a sua resolução radiométrica. A resolução radiométrica define o número mínimo e máximo de níveis de radiância que o detector pode registrar.

Como vimos antes, em função da intensidade da radiação de entrada nos detectores do sensor, o sinal de saída é convertido eletronicamente em valores numéricos digitais discretos. A esses valores, dá-se o nome de quantização, a qual corresponde à medida da resolução radiométrica. Em termos práticos, a quantização do sinal é medida pelo intervalo de números de valores digitais usados para expressar os valores de radiância medi-

dos pelo detector. A quantização é normalmente expressa em termos de números de dígitos binários (bits). Quanto maior é a quantização, melhor é a qualidade visual da imagem, como mostra a Figura 4.8. Nessa figura, observa-se que a imagem de 16 bits ($2^{16} = 65.536$) tem detalhes visuais melhores que as imagens de menores resoluções radiométricas, por exemplo, de 6, 4 ou 2 bits.

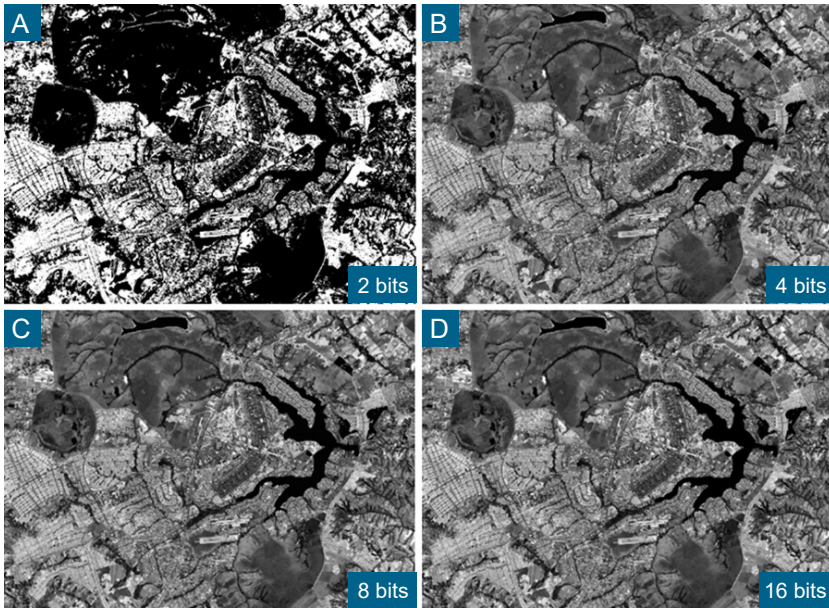


Figura 4.8. Imagens da área urbana do Distrito Federal adquiridas em 2, 4, 6 e 8 bits (A–D) mostrando que, quanto maior for o nível de quantização em bits, melhor é a resolução radiométrica.

4.5. Resolução Temporal

Resolução temporal refere-se à frequência com que o sensor revisita uma área ao longo de sua vida útil. Isso só é possível porque os satélites de sensoriamento remoto executam uma órbita heliossíncrona, que é um caso particular de uma órbita semipolar. O plano de órbita é sempre fixo e

ortogonal ao sentido de rotação da Terra. Assim, o satélite passa sobre um mesmo ponto da superfície da Terra na mesma hora. As órbitas possuem uma inclinação em relação ao equador de 97° a 98° a uma altitude nominal próxima de 550 km a 900 km. O tempo de viagem para completar uma órbita é de aproximadamente 90 minutos. Em um intervalo de tempo de 24 horas, aproximadamente 14 órbitas se completam. Considerando-se que os sensores a bordo dos satélites conseguem imagear somente uma faixa da Terra com algumas dezenas ou centenas de quilômetros de largura, as 14 órbitas imageadas em um dia ficam distanciadas entre si de milhares de quilômetros devido às velocidades relativas entre a órbita do satélite e a rotação da Terra. Como a rotação da Terra é no sentido de oeste para leste, a cada novo dia a posição da órbita progride na direção oeste.

Esse processo de cobertura pode ser exemplificado considerando-se as características de imageamento do satélite Landsat (Figura 4.9). Cada órbita do Landsat cobre uma faixa no terreno de 185 km de largura. As órbitas tomadas no mesmo dia se distanciam entre si de 2.875 km, sendo necessários 16 dias para concluir o recobrimento total do globo. O princípio é o mesmo para qualquer outro satélite. O que irá variar é a resolução temporal do imageamento, pois cada sensor imageia uma determinada largura de faixa no terreno em virtude de sua altitude e parâmetros de visada.

A resolução temporal é fundamental para acompanhar ou detectar as mudanças que ocorrem na Terra, principalmente para alvos dinâmicos, como são os casos do ciclo fenológico de culturas, desmatamentos e desastres ambientais, tendo forte impacto na monitoração ambiental. Na Figura 4.10, é mostrada uma imagem contendo uma área agrícola com vários pivôs de irrigação. Percebe-se que, em um intervalo de 40 dias, ocorreram acentuadas alterações no parcelamento dos cultivos e no número de pivôs. Atualmente, pela facilidade com que se pode reposicionar a visada de um sensor, de nadir para *off-nadir* (fora da vertical), muitos satélites oferecem a possibilidade de revistas à área em períodos de tempo de dois a três dias.

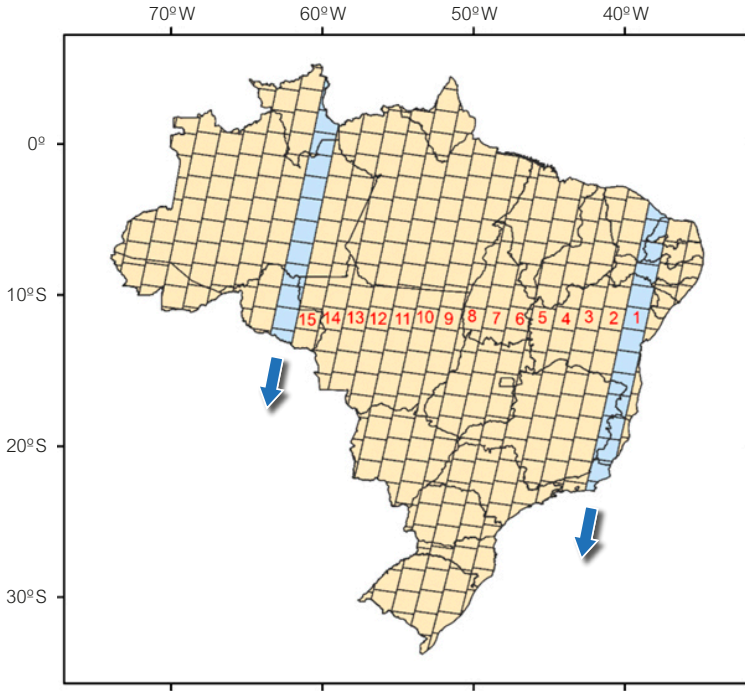


Figura 4.9. Padrão de recobrimento de órbitas completadas em um único dia pelo satélite Landsat em destaque e no período da resolução temporal de 16 dias.

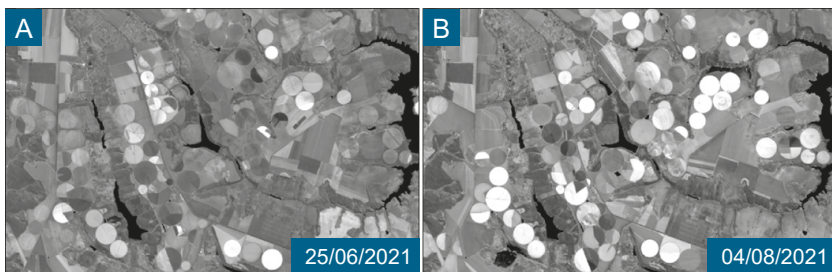


Figura 4.10. Efeito da resolução temporal na detecção de mudanças em áreas de cultivo sob sistema de irrigação por pivô central em função das diferenças de menos de dois meses nas datas das imagens. A imagem (A) foi obtida em 25 de junho de 2021 enquanto a imagem (B) foi adquirida em 04 de agosto de 2021.

4.6. Resolução Espacial de Imagens de Radar

As medidas de resoluções espaciais dos sensores de radar são bem mais complexas que as dos sensores multiespectrais e termais, em razão da linha de imageamento no terreno ter a forma de um lóbulo cuja largura aumenta ao longo de sua dimensão. O elemento ou a célula de resolução espacial de um radar de visada lateral é controlado por dois parâmetros independentes: o comprimento do pulso, determinado pela duração de tempo que a antena emite pulsos de energia, e a largura angular do feixe do pulso enviado pela antena.

Como são parâmetros independentes, cada um deles determina uma resolução espacial, havendo duas medidas de resoluções espaciais: uma na direção de alcance ou de iluminação de propagação do pulso e outra na direção transversal ao lóbulo, que é a direção azimutal da linha de voo. Elas são denominadas respectivamente de resolução em alcance (*range*) e resolução azimutal. Podem também ser ditas como resoluções longitudinal e transversal em relação ao lóbulo de iluminação. Na resolução em alcance, calcula-se o comprimento do elemento de resolução e na resolução azimutal calcula-se a largura.

4.6.1. Resolução em alcance (*range*)

A resolução em alcance é medida ao longo da direção de iluminação do pulso de onda. O valor calculado para a resolução espacial em alcance é diretamente proporcional ao comprimento do pulso τ (determinado pela duração de transmissão do pulso) e à velocidade da luz (c) e inversamente proporcional ao cosseno do ângulo de depressão (γ). A Equação 4.4 determina a resolução em alcance no plano do terreno.

$$R_r = \frac{\tau \times c}{2 \cos \gamma} \quad (4.4)$$

em que

R_r = resolução em alcance;

τ = tempo de duração do pulso;

c = velocidade da luz; e

γ = ângulo de depressão.

Na Figura 4.11, os alvos 1 e 2, situados na posição *near range* e os alvos 3 e 4, situados na posição *far range*, estão igualmente distanciados entre si, no terreno. Eles são iluminados pelo pulso de onda com ângulos de depressão de 50° e de 30° , respectivamente. Considerando-se um tempo de duração do pulso $\tau = 0,1 \times 10^{-6}$ s, a resolução em alcance, calculada de acordo com a Equação 4.4, será igual a 17,3 m em *far range* e de 23,3 m em *near range*. Supondo que, no terreno, os pares de alvos estejam separados de uma distância de 20 m, somente os alvos 3 e 4 em *far range* serão resolvidos, isto é, vistos como dois alvos separados. Os alvos localizados na posição *near range*, apesar de estarem mais próximos da plataforma, não serão resolvidos, formando uma imagem superposta.

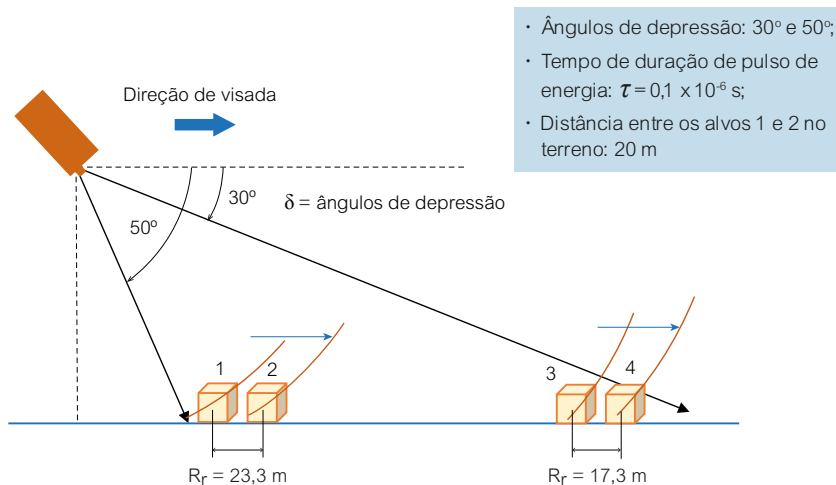


Figura 4.11. Variação da resolução em alcance em função dos ângulos de depressão da antena de radar.

Fonte: adaptado de Lillesand et al. (2008).

4.6.2. Resolução em azimute (azimutal)

A resolução azimutal é medida na direção paralela à direção da linha de voo (direção azimutal). Há dois modos de se calcular a resolução azimutal: pelo tamanho da antena ou pela medida angular do feixe de onda transmitido pela antena. Como se pode ver na Figura 4.12, a largura do lóbulo é estreita em *near range* e mais larga em *far range*. Nesse caso, diz-se que o radar é de abertura real. Quanto mais estreita a largura do lóbulo, melhor é a resolução azimutal. A largura do lóbulo é diretamente proporcional ao comprimento de onda (λ) do pulso transmitido, isto é, quanto menor o comprimento de onda, mais estreito é o lóbulo. Inversamente, os maiores comprimentos de onda resultam em lóbulos mais largos. Também se sabe que a largura do lóbulo é proporcional ao comprimento da antena (L). Quanto maior for o tamanho da antena, menor é a largura do lóbulo e, consequentemente, melhor a resolução azimutal. A Equação 4.5 relaciona o comprimento de onda (λ) e o tamanho da antena (L) para o cálculo da resolução azimutal (R_{AZ}).

$$R_{AZ} = \frac{S \times \lambda}{L} \quad (4.5)$$

em que

S = distância da antena ao ponto no terreno no alcance inclinado; e

L = tamanho da antena.

Usando a Equação 4.5 para calcular a resolução dos objetos situados em *near range* e *far range* em uma situação como a apresentada na Figura 4.12 (as distâncias AB são iguais), podemos facilmente deduzir que a resolução azimutal em *near range* é melhor do que em *far range*, em razão das diferenças de distâncias dos alvos à antena medidas no plano inclinado S . As duas torres na posição em *near range* são individualizadas, ou resolvidas, enquanto as duas torres na posição em *far range* não são resolvidas.

Dessa forma, para melhorar a resolução azimuthal, pode-se diminuir a distância entre o radar e o alvo diminuindo a altura de voo ou aumentando o tamanho da antena, o que, nesse caso, é um grande problema.

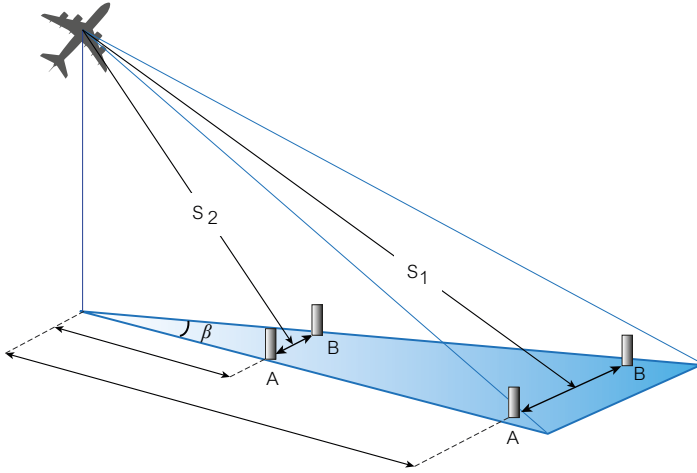


Figura 4.12. Dependência da resolução azimuthal em função das distâncias S medidas no plano inclinado, ou da abertura angular β do lóbulos de iluminação do feixe de radar.

Quanto à medida da resolução azimuthal (R_{AZ}) que é função do ângulo de abertura do feixe de radar, também é facilmente percebido que, quanto menor for a abertura angular β , mais estreito é o feixe de radar e melhor a resolução (Figura 4.12). O cálculo é feito pela Equação 4.6, onde G_r é a distância no terreno.

$$R_{AZ} = G_r \times \beta \quad (4.6)$$

De acordo com essa última equação, as torres nos pontos A e B seriam resolvidas em G_{r1} (distância mais próxima da antena), mas não em G_{r2} , ou seja, à distância G_{r1} elas resultam em sinais separados, ao passo que para a distância G_{r2} elas seriam vistas simultaneamente e sobrepostas.

Por outro lado, a abertura angular β do feixe é diretamente proporcional ao comprimento de onda (λ) do pulso transmitido e inversamente proporcional ao tamanho da antena (L), tal como demonstrado na Equação 4.7.

$$\beta = \frac{\lambda}{L} \quad (4.7)$$

As relações entre as Equações 4.5, 4.6 e 4.7 indicam que, para se ter uma maior resolução azimutal, é preciso utilizar antenas bastante longas, com dezenas ou centenas de metros, o que é inviável de se alojar em aviões ou satélites. A solução foi encontrar uma maneira de simular grandes antenas em antenas pequenas, capazes de serem transportadas por qualquer tipo de plataforma. Disso, resultou a mudança da tecnologia dos radares imageadores de abertura real (RAR) para os radares de abertura sintética (SAR).

CAPÍTULO 5

Modelos de Dados e de Processamento

Paulo Roberto Meneses

Tati de Almeida



Em sensoriamento remoto, a medição do processo físico da radiação eletromagnética está implícita nos dados que geram a imagem. O modelo de dados das imagens e suas notações permitem estabelecer uma conectividade entre o sentido físico do dado e as configurações dos algoritmos de processamento. A energia da radiação eletromagnética refletida, emitida ou retroespalhada pelos objetos é convertida em um valor digital inteiro por meio de um conversor analógico/digital e codificada em uma unidade denominada pixel. A forma digital do dado facilita o uso da computação para processar as imagens, com o objetivo principal de representar porções bem definidas do espaço terrestre, utilizando-se de processamentos matemáticos, estatísticos e probabilísticos. Os limites do espaço são configurados pelos pixels, os quais são organizados na disposição de linhas e colunas. Embora estejamos falando de valores numéricos (imagem digital), os limites de qualquer porção da imagem não são matemáticos. Não são limites como o que vemos em uma carta topográfica, os quais são, de fato, limites matemáticos geralmente definidos por meridianos e paralelos. Então, percebe-se que há uma clara distinção entre a representação dos dados de uma imagem com a representação dos dados do espaço terrestre nas cartas topográficas. Qualquer que seja a imagem, ela não possui um sistema de referência de projeção cartográfica. Daí a razão de nos referirmos à localização dos objetos em uma imagem por meio de coordenadas espaciais dos pixels dispostos em linhas e colunas.

5.1. Notação dos Dados

Conforme descrito nos capítulos anteriores, a imagem de um sensor é obtida pela varredura do terreno em linhas sequenciais em que cada linha é amostrada por uma área unitária (pixel). O sensor registra os pixels no formato de uma grade retangular ou quadrada, e o valor do pixel, correspondente aos valores de radiância, emitância ou de retroespalhamento dos

alvos, é convertido no sensor para um valor digital. O valor digital do pixel representa a média dos valores de radiância dos diferentes alvos que estão contidos no pixel. Os pixels podem ser puros (um só objeto) ou misturas de dois ou mais objetos. Em terra, as estações de rastreamento gravam o sinal digital transmitido pelo satélite por telemetria e a imagem é gravada no formato *raster* em laboratórios de produção de imagens. Dados *raster* são formados por uma matriz de pixels (também chamados de células) em que cada célula possui valor único que representa a posição e a medida física do dado. Trata-se de uma abordagem diferente do formato vetorial que utiliza a geometria (pontos, linhas e polígonos) para representar o mundo real.

Os algoritmos de processamento veem esse formato como uma matriz de dados composta por milhares de linhas e colunas, na qual o valor de cada célula da matriz é definido pelas coordenadas que o pixel nela ocupa e pelo seu valor digital. Genericamente, pode-se definir qualquer imagem como sendo uma função bidimensional representada por $f(x,y)$ (Gonzales e Woods, 2017). As coordenadas matriciais x,y e os valores de amplitude f são quantidades discretas, isto é, números inteiros. Nas imagens de sensoriamento remoto, x,y corresponde às coordenadas espaciais do pixel e o valor de f em qualquer ponto x,y é a medida da energia radiante.

Um pixel posicionado na linha i e na coluna j terá a notação ND_{ij} ou ND_p . O primeiro pixel é numerado de (1,1) e posicionado no lado esquerdo superior da matriz e o último (n,m) no canto direito inferior. Essa notação é natural para os programas de computador por causa do armazenamento sequencial dos dados, diferindo da notação cartesiana (x,y) contínua, com origem dos dados em (0,0) e as coordenadas x,y rotacionadas em 90° (Figura 5.1). Essa convenção é utilizada porque a maioria das imagens é obtida no sentido de órbita descendente do satélite, de norte para sul. Nesse caso, o topo da imagem indica o sentido para norte. Quando uma imagem é obtida no sentido ascendente da órbita, para visualizá-la no monitor e orientá-la para norte, faz-se uma rotação na imagem. O tamanho de uma imagem é uma expressão do tipo: linha x coluna (byte) x número de bandas.

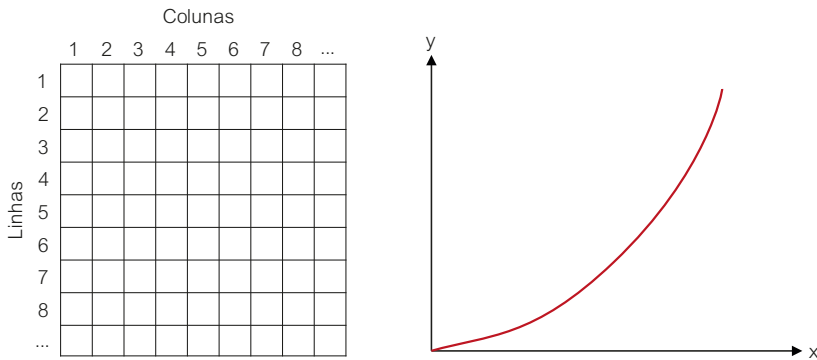


Figura 5.1. Notação dos dados da imagem na forma de uma matriz com a locação dos pixels definidos por linha, coluna, comparada com a notação cartesiana para dados contínuos.

Na forma digital, o valor da função $f(x,y)$ é representado por um número digital quantizado em bits que assume somente valores inteiros positivos, como é mostrado na Figura 5.2. Uma imagem $f(x,y)$, para ser processada e exposta na tela do computador, tem de ser digitalizada tanto espacialmente como em amplitude de níveis de brilho (ou cinza). A digitalização de coordenadas x,y refere-se à amostragem de imagem, enquanto a digitalização de amplitude refere-se à quantização de nível de brilho. Ao olharmos para a Figura 5.2, podemos perguntar: por que não executar o processamento no modo fracionário (*floating point*) que pode representar qualquer valor? O problema é o espaço em disco. Quanto maior o intervalo de valores dos dados, mais espaço em disco é preciso. Por exemplo, um tipo de dado em *byte* usa somente um único byte para representar o valor de cada pixel. Tipos de dados em *integer* usam dois bytes para cada pixel, enquanto o tipo em *floating point* usa quatro bytes. Assim, *floating point* consome duas vezes mais espaço em disco do que o tipo *integer*.

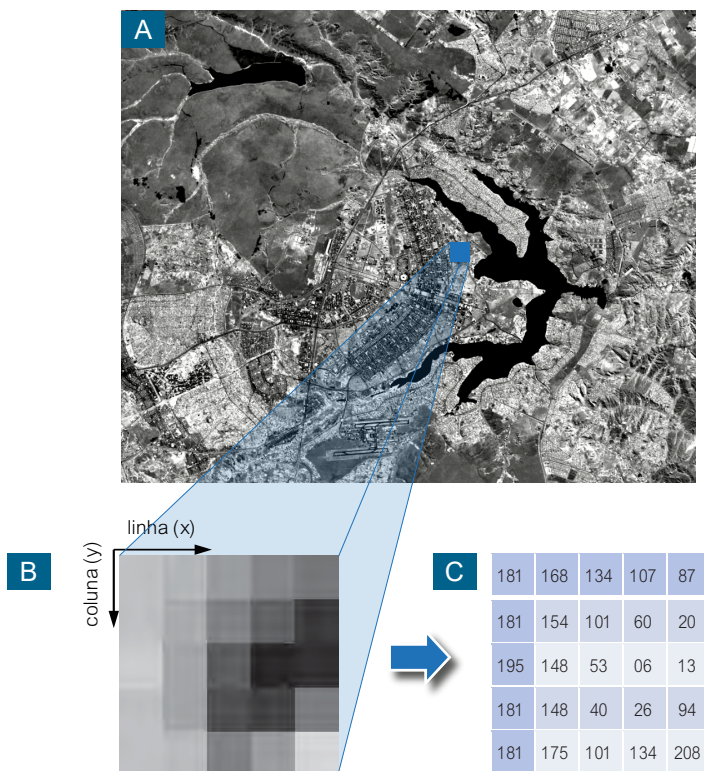


Figura 5.2. Imagem digital (A) com área em zoom de um subconjunto de pixels em níveis de cinza (B) e os correspondentes valores digitais (C).

5.2. Arquivos Raster

Todas as imagens de sensoriamento remoto são no formato raster, que é um formato que contém a descrição de cada pixel da imagem digital de forma sequencial contínua, sendo assim gravada como uma matriz.

Grande parte dos programas comerciais de processamento de imagens usa o formato de dados *raster* com dois arquivos: um arquivo binário (base dois) que contém os dados da imagem e um pequeno arquivo texto que descreve algumas informações sobre os dados da imagem (*header*).

Nos seus formatos nativos, os programas aceitam dados (imagens) em *byte*, *signed*, *unsigned integer*, *long integer*, *floating point*, *integer*, *64-bit*, *complex* e *double complex*, além de outros. O mais comum é armazenar os dados no disco rígido em um fluxo binário de *bytes*, seguindo um dos três tipos básicos de formatos: bandas sequenciais (BSQ), bandas intervaladas por linhas (BIL) e bandas intervaladas por pixels (BIP).

5.3. Formatos de Salvamento

O salvamento de imagens é um procedimento corriqueiro durante uma seção de trabalho de processamento de imagens. A forma de salvar um dado é um procedimento controlado pelo software em uso. O importante é conhecer em que tipos de formatos as imagens podem ser salvas, quando é preciso transferi-las para outros aplicativos.

A maioria dos programas de processamentos de imagens de sensoriamento remoto assegura o salvamento das imagens nos formatos nativos de outros programas. Além desses formatos nativos, há opções de salvamento nos formatos comuns de visualização de imagens. A escolha do formato deve garantir que se preserve, o máximo possível, a imagem a ser salva em termos de qualidade e quantidade de cores, contraste, tamanho máximo do arquivo, além de outros parâmetros. Os formatos mais usados são dos tipos JPEG, BMP, GEOTIFF e PNG.

5.4. Modelo de Processamento

O modelo de processamento de imagens de sensoriamento remoto é um modelo ditado para imagens multiespectrais, cujos algoritmos levam em conta, principalmente, a origem dos dados segundo os intervalos espectrais que foram obtidos: ópticos, termais e das micro-ondas. Isso significa que o processamento de imagens de sensoriamento remoto é orientado para cada tipo de imagem e para cada tipo de problema. Há um conjunto de algoritmos que são próprios para imagens hiperespectrais, outros para

imagens com altíssimas resoluções espaciais e outros para imagens de radar. Os modelos dos sensores e os parâmetros que definem as características básicas da imagem, como as resoluções e o nível de quantização das imagens são determinantes nisso. No entanto, independente de qual seja o tipo de imagem, todo o conjunto de processamento tem, como objetivo principal, operar transformações das imagens originais em imagens com melhores qualidades espectrais e espaciais e que sejam mais apropriadas para uma determinada aplicação ou para a extração de dados pelo analista. Os algoritmos de processamento executam operações matemáticas e estatísticas dos dados, que podem ser agrupados em:

- Correções e calibrações;
- Transformações espectrais;
- Transformações espaciais;
- Fusões; e
- Classificações.

Os algoritmos são extremamente especializados, cada um idealizado com uma função específica e são independentes entre si. Isso significa que o analista não tem, rigidamente, que seguir um receituário de processamento. Além disso, o resultado que o analista obtém da escolha do processamento para a solução do seu problema, não necessariamente atesta que pode ser reproduzível para outra área, pois o mundo real não se replica por cópias. Muito embora os algoritmos tratem matematicamente a imagem como um dado diretamente relacionado ao processo físico que a criou, a representação e o processamento matemático da imagem não é 100% da representação do mundo real. Por isso, a tomada de decisão sempre caberá ao usuário e não ao computador. Para isso, os programas de processamento de imagens são bastante amigáveis com o usuário e de fácil operação.

Como foi dito, não há um fluxo de processamento predefinido que o usuário seja obrigado a seguir. Por exemplo, o usuário que deseja gerar um mapa temático usando algoritmos de classificação supervisionada ne-

cessitará de informações precisas da localização dos alvos que ele deseja investigar e, por isso, as correções das distorções geométricas da imagem são imprescindíveis. Isso porque um mapa é um documento de registro cartográfico dos objetos que representa. Outro usuário que queira extrair informações espectrais dos alvos para identificar a natureza de suas composições estará mais preocupado em calibrar previamente os dados.

A Figura 5.3 deve, portanto, ser vista como um quadro sucinto que mostra as principais técnicas de processamento que a maioria dos aplicativos comerciais dispõe e não como um fluxo de processamento a ser seguido. Não se sugere que a imagem digital deva primeiro ser submetida a uma etapa de pré-processamento que visa correções dos dados, para depois ser processada por técnicas de transformação ou de classificação. Não se sugere, ainda, que as técnicas de transformação estão à parte das técnicas de classificação, porque uma imagem que será classificada pode ter sido antecipadamente submetida a uma técnica de realce espectral.

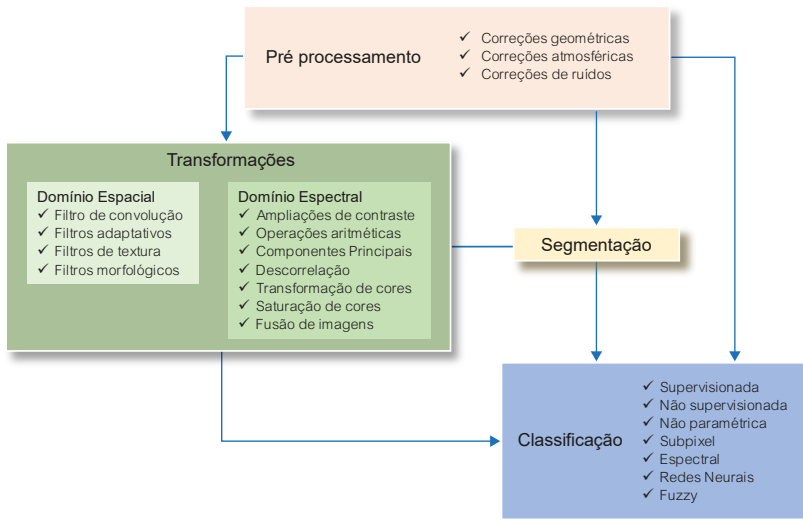


Figura 5.3. Principais técnicas de processamento de imagens segundo um modelo geral de ordenação dos tipos de processamento.

As técnicas de pré-processamento são, essencialmente, funções operacionais para remover ou corrigir erros e distorções que são introduzidos nas imagens pelos sistemas sensores devido a erros instrumentais (ruídos espúrios), às interferências da atmosfera (erros radiométricos) e à geometria de imageamento (distorções geométricas). As técnicas de transformação nos domínios espectral e espacial listadas têm a função básica de gerar uma imagem de alto contraste visual para facilitar o analista na interpretação da cena, em um processo que objetiva otimizar ao máximo a extração das informações e a análise, em termos de eficiência, tempo e custo. Os dados são transformados em função dos valores digitais de cada pixel individual (domínio espectral) ou em função das relações espaciais de conjuntos de pixel vizinhos (domínio espacial). Os resultados são imagens com novos arranjos de valores digitais.

As técnicas de classificação temática tratam a imagem por meio de complexos modelos estatísticos e probabilísticos com o intuito de gerar um mapa temático com uma participação bastante reduzida do intérprete no processo do mapeamento. No conjunto, essas técnicas de processamento constituem-se em um grande esforço para tornar mais acessível, por meio de uma forma interativa homem-máquina, a extração de informações que, no mínimo, seriam difíceis de serem obtidas com o uso de métodos visuais clássicos de interpretação. Desde o início do sensoriamento remoto orbital, muitos pesquisadores perceberam essa dificuldade porque muitas classes de alvos naturais podem exibir baixos contrastes espectrais de reflectância.

Somente com cuidadosos processamentos computacionais pode-se extrair a informação no limite mais ideal de análise. Além do mais, se considerarmos o volume de dados que integra uma cena correspondente a um conjunto de multibandas, haverá centenas de milhões de dados registrados, não havendo como imaginar uma análise sem os recursos de processamento. Os capítulos que seguem no transcorrer do livro descrevem as técnicas de processamento seguindo, aproximadamente, o fluxograma da Figura 5.3.

5.5. Formato de Arquivos Rasterizados

JPEG (*Joint Photographic Experts Group*)

É um tipo de formato de arquivo utilizado para armazenamento de imagens com esquema de 24 bits. Isso significa que esse formato aceita 16,8 milhões de cores. Devido à limitação da visão humana em distinguir, no máximo, entre três a cinco mil cores, são subtraídas algumas informações das cores, preservando somente o que o olho humano pode enxergar. Esse processo é conhecido como compressão, mas não há perda da qualidade da imagem, com a vantagem de gerar um arquivo não muito grande.

PNG (*Portable Network Graphics*)

É uma versão do JPEG com esquema de cores de 24 bits, isto é, com 16,8 milhões de cores. A vantagem sobre o formato JPEG é que a compressão dos dados é mais eficiente, não causando perda de qualidade a cada salvamento, o que possibilita maior fidelidade à imagem original.

BMP (*Bitmap*)

Desenvolvido para aplicações disponíveis no sistema operacional Windows, trata-se de um formato muito simples, com mínima possibilidade de erro na interpretação de arquivos. Como não há compressão de dados, a qualidade da imagem é a melhor possível, em compensação, o tamanho dos arquivos costuma ser grande. Utiliza o esquema de 24 bits de cor para imagens com 16 milhões de cores simultâneas, 16 bits para imagens com 65.536 cores e 8 bits para imagens com 256 cores.

TIFF (*Tag Image File Format*)

Foi desenvolvido como uma tentativa de se criar um padrão para imagens geradas por equipamentos digitais, por exemplo, escaneadores. É um formato capaz de armazenar imagens em preto e branco, em escalas de tons de cinza, ou em paletas de cores com 24 ou 32 bits. Ele tem a vantagem de ser reconhecido por qualquer tipo de programa de processamento de imagens. O formato TIFF é baseado em entidades lógicas (campos TIFF) compostas por etiquetas (tags) que descrevem atributos e parâmetros de dados das imagens.

ASCII (*American Standard Code for Information Interchange*)

Muitas vezes pode ser solicitado, ao usuário, o salvamento de arquivos que não são imagens, mas de dados obtidos durante o processamento e que podem ser armazenados na forma de um arquivo texto. O formato ASCII é usado para essa função. É um tipo de formato binário de codificação de caracteres (letra, número ou símbolo) que são representados por oito bits. O sistema totaliza 128 caracteres. São exemplos de arquivos ASCII, gráficos, coordenadas x,y ou x,y,z de dados, cabeçalhos de imagens, anotações e outros.

5.6. Referência

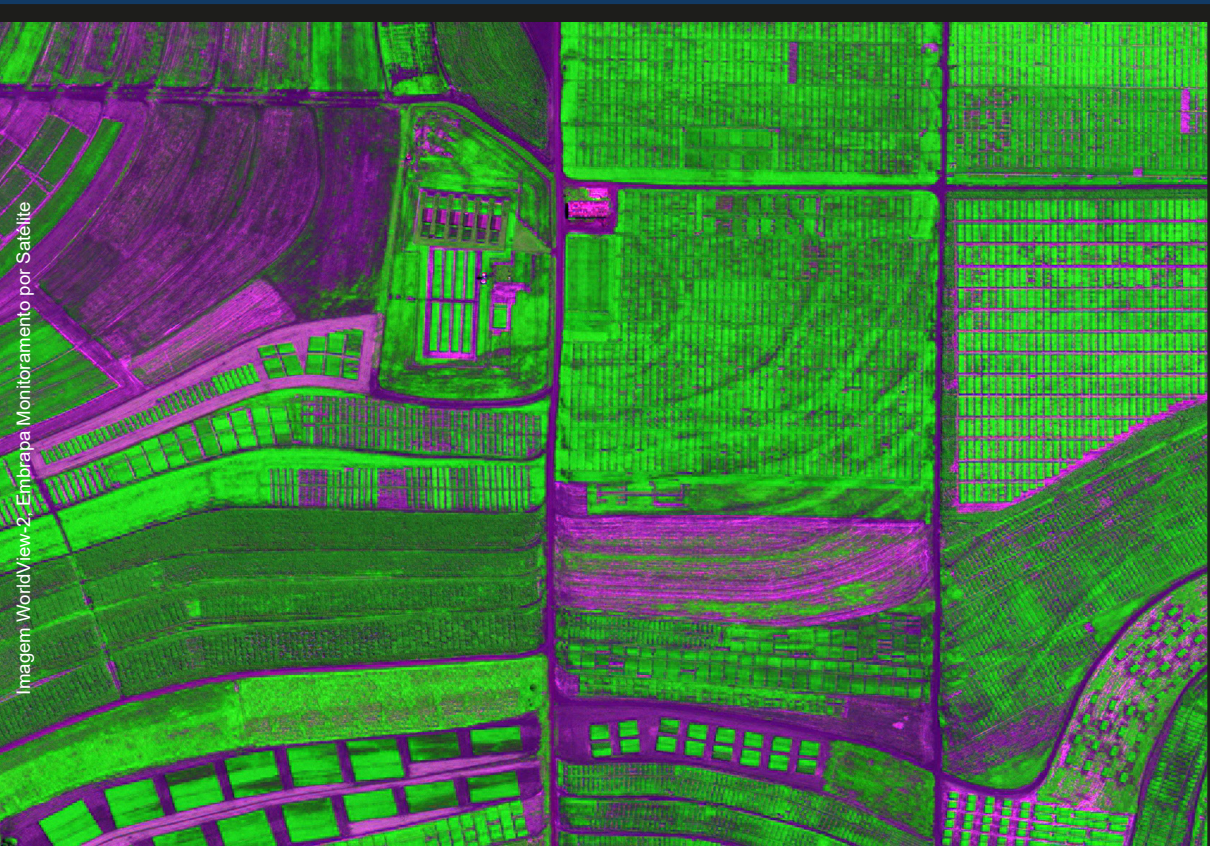
GONZALEZ, R. C.; WOODS, R. E. **Digital image processing**. Nova York: Pearson, 4a ed., 2017, 1.024 p.

CAPÍTULO 6

Correção Atmosférica de Imagens

Rejane Ennes Cicerelli

Tati de Almeida



Imagens da superfície da Terra adquiridas por plataformas orbitais e aéreas sofrem efeitos da atmosfera. Esses efeitos incluem a atenuação da luz refletida e a perda de contraste devido ao espalhamento da luz solar pelos aerossóis e moléculas presentes na atmosfera. As aplicações de sensoriamento remoto requerem que esses efeitos atmosféricos sejam removidos das imagens para recuperar a reflectância espectral inerente aos materiais presentes na superfície terrestre. O processo de remoção dos efeitos atmosféricos é conhecido como correção ou compensação atmosférica. Existem diversos métodos e algoritmos de correção atmosférica como aqueles baseados nos princípios da modelagem física de transporte de radiação e nas abordagens empíricas que se baseiam em reflectâncias conhecidas na imagem. Apesar do número considerável de algoritmos de correção atmosférica disponíveis, pouco se sabe sobre seus méritos. Esse capítulo é dedicado ao entendimento dos efeitos atmosféricos na imagem e à apresentação das características dos diferentes métodos de correção disponíveis.

6.1. O Efeito da Atmosfera sobre a Radiação Eletromagnética (REM)

As reflectâncias dos alvos da superfície terrestre registradas pelos sensores orbitais são influenciadas pelas interferências atmosféricas que são produzidas pela absorção e espalhamento de gases e aerossóis e pelos efeitos de adjacência causados pelas variações de alvos da superfície terrestre. A interferência atmosférica altera a intensidade do fluxo radiante, a distribuição espectral e a direção dos raios incidentes, tanto na sua trajetória descendente entre o Sol e a Terra quanto na sua trajetória ascendente da radiação refletida e emitida da superfície terrestre para o sensor. Esses fatores modificam a magnitude da resposta e suas características espectrais e seus efeitos variam substancialmente com o comprimento de onda da radiação incidente.

Para se ter uma ideia, se não houvesse atmosfera, o céu seria preto e com um disco brilhante (o Sol), do qual receberíamos radiação direta. Na média, 47% do fluxo de radiação que incide na superfície terrestre são absorvidos pelos materiais da superfície terrestre, 37% são refletidos pela Terra (incluindo nuvens e atmosfera) e 17% são absorvidos pela atmosfera. A parte da radiação eletromagnética (REM) que interage diretamente com a atmosfera é questão crucial para o sensoriamento remoto.

A absorção pela atmosfera é o efeito mais prejudicial ao sensoriamento remoto. Como pode ser observado na Figura 6.1, em vários intervalos de comprimentos de onda, a atmosfera mostra-se parcial ou totalmente opaca às passagens da radiação solar e da radiação emitida pela Terra, em razão da absorção pelos gases nela presentes, em particular, moléculas de O_2 , CO_2 , O_3 e H_2O . O conhecimento sobre o tipo de molécula predominante e a sua espessura óptica é fundamental para a precisão dos processos de correção atmosférica. Os tipos de partículas com composições, formas, dimensões e propriedades ópticas são diferentes e variáveis, podendo ser de origem natural, como a poeira oriunda de desertos e o sal presente como sólidos em suspensão nos oceanos, ou de origem antrópica, como a poluição industrial, poluição urbana e queima de biomassa vegetal.

Como consequência, a radiação solar pode ser impedida de atingir a superfície terrestre ou, no mínimo, a sua intensidade é atenuada, o mesmo acontecendo com a radiação emitida pela Terra. Dessa forma, o sensor colocado no espaço ficará impedido de obter imagens da superfície terrestre em alguns comprimentos de onda. Esses intervalos de comprimentos de onda são chamados de bandas de absorção da atmosfera e são restritivos para o uso do sensoriamento remoto. As demais regiões onde a atmosfera não absorve total ou intensamente a radiação solar são chamadas de janelas atmosféricas, as únicas em que é possível usar o sensoriamento remoto.

Observe na Figura 6.1 que a região de maior absorção, portanto, sem uso no sensoriamento remoto, é no intervalo espectral termal de $14\ \mu m$ a $1.000\ \mu m$ devido à total absorção da radiação pelo vapor de água atmosférica. Em contrapartida, na região das micro-ondas, a atmosfera é quase

100% transparente. No intervalo do visível ao infravermelho, que é a região espectral mais usada em sensoriamento remoto, a atmosfera também mostra a sua influência danosa para o uso do sensoriamento remoto. Por exemplo, nos comprimentos de onda de 1,4 μm e 1,9 μm , 100% da radiação solar são absorvidas pelas moléculas de vapor de água, impedindo totalmente o uso de sensoriamento remoto nesses comprimentos de ondas. Alguns sensores, principalmente os hiperespectrais, são capazes de registrar informações da atmosfera nessa região com o intuito de melhor modelar a influência da atmosfera.

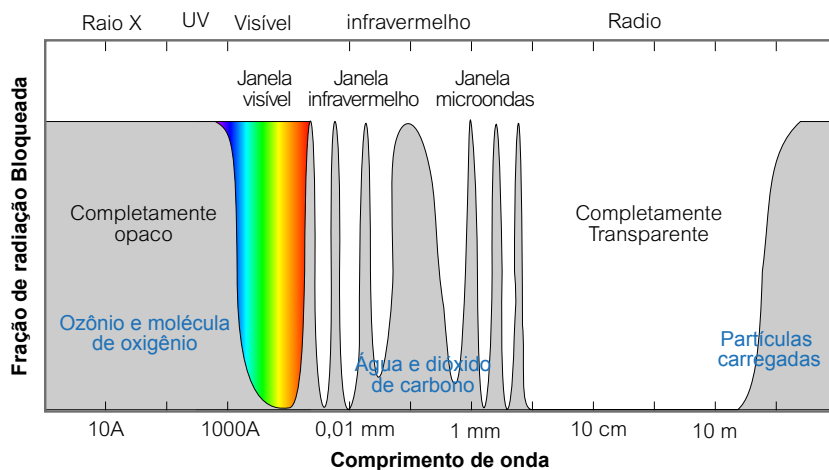


Figura 6.1. Transmitância da radiação eletromagnética através da atmosfera. As áreas em cinza correspondem às regiões espectrais onde os gases atmosféricos absorvem a radiação.

Na Figura 6.2, são mostradas as atenuações que a irradiância solar, medida no topo da atmosfera (linha laranja), sofre quando é medida na superfície terrestre (linha azul), ou seja, depois de atravessar a atmosfera. São vários os comprimentos de onda em que a atmosfera absorve total ou parcialmente a radiação solar. Essas absorções são o motivo dos sensores mostrarem descontinuidades nos comprimentos de onda entre bandas contíguas.

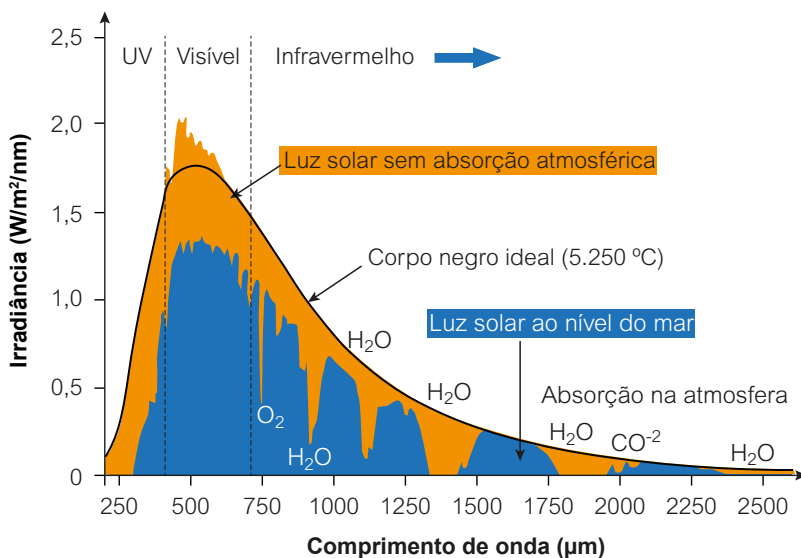


Figura 6.2. Curvas da irradiância solar de um corpo negro no topo da atmosfera e na superfície terrestre. Nessa última curva, é mostrada uma série de absorções associadas à presença de gases na atmosfera.

Considerando um sistema de sensoriamento remoto passivo que tem o Sol como fonte de energia e que opera nas regiões espectrais do visível e do infravermelho refletivo, a trajetória da energia está relacionada com os elementos da REM apresentados na Figura 6.3. Essa figura mostra o efeito que a atmosfera provoca no valor de brilho medido de um único pixel. Na ausência de atmosfera, o sinal medido pelo sensor seria uma função exclusiva do nível de energia solar que incide no pixel e das propriedades de reflectância do próprio pixel. No entanto, a presença da atmosfera modifica esse sinal hipotético significativamente. Primeiro, tem-se a irradiância do céu (E_d) no qual, além da radiação proveniente diretamente da fonte, um alvo (pixel) pode receber parte da radiação espalhada pela atmosfera, além daquela proveniente dos pixels vizinhos. Destaca-se ainda a radiação de trajetória (L_p), que é a radiação que atinge o sensor oriunda do espalhamento atmosférico e proveniente dos pixels adjacentes.

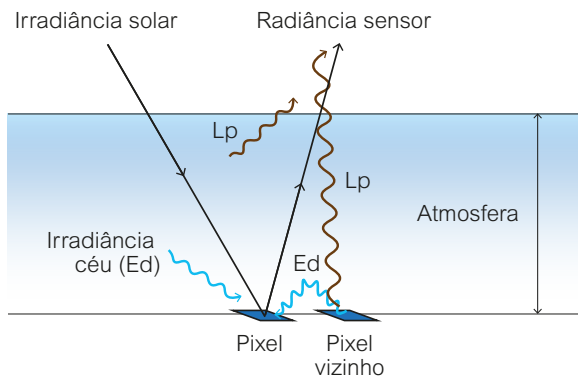


Figura 6.3. Efeito da atmosfera na trajetória da energia solar que atinge a superfície terrestre e, em seguida, chega até o sensor.

Essas radiações provenientes da atmosfera são acrescidas à radiação proveniente do próprio pixel. Nesse valor registrado no pixel, é importante introduzir algumas definições de grandezas radiométricas que possibilitam simplificar explicações e permitem que as equações de correção atmosférica sejam devidamente formuladas. Imagine o Sol como uma fonte de energia que emite a uma determinada taxa de energia (em Joules por segundo ou Watts). Essa energia radiante se propaga isotropicamente no espaço, de modo que, a uma determinada distância, a energia emitida pelo Sol pode ser medida em W m^{-2} . Essa densidade de potência que incide na superfície terrestre é chamada de irradiância (E_λ), uma propriedade que pode ser usada para descrever a força de qualquer emissor de energia eletromagnética, também dada em W m^{-2} . Há ainda a densidade de fluxo radiante que emerge da superfície por unidade de área da superfície, denominada exitância (M_λ) radiante.

A partir do momento em que a irradiância chega ao alvo e é capturada em uma determinada direção, tem-se a densidade de energia espalhada e que é definida por sua densidade por ângulo sólido. Essa quantidade de energia é chamada de radiância (L_λ) e tem unidade de medida em Watts por metro quadrado por esteradiano ($\text{W m}^{-2} \text{sr}^{-1}$). A irradiância é depen-

dente da variação natural da fonte de radiação específica, o que significa que a radiância é diretamente proporcional a E_λ . Essa característica inviabiliza qualquer caracterização espectral das feições da superfície terrestre. Portanto, para exclusão da influência da variação da fonte de iluminação, necessita-se definir a variável reflectância (ρ_λ) que é considerada como a razão adimensional entre a radiância (L_λ) e o fluxo radiante incidente (E_λ). Até então, estamos apenas tratando da intensidade de fluxo incidente sobre um objeto, desprezando a interferência da atmosfera.

Para entender como a atmosfera atua na informação do pixel, será apresentada uma sequência de cálculos até a obtenção da informação de reflectância de superfície, ou seja, sem a influência da atmosfera, considerando o comprimento de onda (λ) e o ângulo zenital solar θ . Suponha que, na ausência de atmosfera, a irradiância espectral solar na Terra seja E_λ . Se o ângulo zenital solar, medido da normal à superfície, for θ , como mostrado na Figura 6.3, então a irradiância espectral (densidade de potência espectral) na superfície da Terra será $E_\lambda \cos \theta$. Isso resulta na irradiância disponível entre os comprimentos de onda λ_1 e λ_2 (Equação 6.1):

$$E_{sup} = \int_{\lambda_1}^{\lambda_2} (E_\lambda \times \cos \theta \times d\lambda) \text{ (Wm}^{-2}\text{)} \quad (6.1)$$

Em sensoriamento remoto, a largura da banda espectral ($\Delta\lambda = \lambda_2 - \lambda_1$) é estreita o suficiente para que se possa assumir que (Equação 6.2):

$$E_{sup} = E_{\Delta\lambda} \times \cos \theta \Delta\lambda \times d\lambda \text{ (Wm}^{-2}\text{)} \quad (6.2)$$

Em que $E_{\Delta\lambda}$ é a irradiância espectral média na banda $\Delta\lambda$. Para uma superfície difusa com reflectância ρ , representando a proporção da energia incidente que é refletida, a energia espalhada no topo do hemisfério e disponível para ser medida é (Equação 6.3):

$$L = E_{\Delta\lambda} \times \cos \theta \times \frac{\rho}{\pi} \partial\lambda \text{ (W m}^2 \text{ sr}^{-1}\text{)} \quad (6.3)$$

Em que π representa o hemisfério superior do ângulo sólido. L permite determinar o fluxo detectado pelo sensor e, a partir disso, o número digital registrado pelo sensor que é diretamente relacionado com a radiância da cena. A partir do valor digital, por exemplo, 2^{11} bits = 2.048 números digitais, pode-se medir a radiância de um pixel (Equação 6.4).

$$LL = CK + L_{min} \text{ (W m}^2 \text{ sr}^{-1}\text{)} \quad (6.4)$$

Em que $K = (L_{max} - L_{min})/C_{max}$; e L_{max} e L_{min} são as radiâncias máximas e mínimas medidas pelo sensor. Esses valores são usualmente disponibilizados pelas empresas fabricantes dos sensores. A Equação 6.3 relata o caso ideal, sem a existência da atmosfera. Quando a atmosfera está presente, vários mecanismos que modificam essa equação devem ser considerados. A absorção atmosférica é um fenômeno termodinâmico que resulta na perda efetiva de energia para os constituintes atmosféricos. Contudo, mecanismos de espalhamento e absorção da radiação pelas partículas atmosféricas modificam o sinal medido pelo sensor. Assim, a simples relação acima necessita compor elementos que representem fenômenos de espalhamento e absorção ocasionados pela atmosfera.

O espalhamento de difusão ou dispersão corresponde à mudança aleatória da direção de propagação da radiação solar incidente devido à sua interação elástica com os componentes atmosféricos. Teoricamente, o espalhamento pode ser dividido em três categorias, dependendo das relações entre o comprimento de onda da radiação e o tamanho das partículas atmosféricas que causam o espalhamento. Os três tipos são espalhamento Rayleigh, espalhamento Mie e espalhamento não-seletivo.

O espalhamento Rayleigh é o mais comum. Ele ocorre quando o diâmetro das partículas é menor do que o comprimento da onda. O espalhamento Rayleigh relaciona-se com as moléculas de gases presentes na atmosfera e, por isso, é também denominado de espalhamento molecular. O cálculo do coeficiente de espalhamento Rayleigh depende do número de partículas por centímetro cúbico, do comprimento da radiação, dos índices refrativos das partículas e do meio, mas de forma mais simplificada, pode-se dizer que ele é proporcional ao inverso da quarta potência do comprimento de onda (Equação 6.5).

$$R = \frac{I}{\lambda^4} \quad (6.5)$$

De acordo com a lei de Rayleigh, a influência mais prejudicial do espalhamento atmosférico ocorre nas imagens obtidas nos menores comprimentos de onda do visível. Desde que o coeficiente de espalhamento é proporcional ao inverso da quarta potência do comprimento de onda, os menores comprimentos de onda da luz azul são espalhados mais fortemente que a luz vermelha. A luz visível azul é espalhada cerca de seis vezes mais que a luz visível vermelha ($\lambda = 0,617 \mu\text{m}$), enquanto o efeito do espalhamento nas imagens do infravermelho ($\lambda > 0,76 \mu\text{m}$) é bem menor, quase imperceptível. O fato de a luz azul espalhar-se mais que os comprimentos de onda mais longos do visível é por ela ser absorvida pelas moléculas e reemitida em todas as direções, modificando, portanto, a direção da radiação.

O espalhamento Rayleigh é a causa de o céu aparecer azul. O vermelho do pôr-do-sol é também causado pelo espalhamento Rayleigh. Tão logo o Sol se aproxima do horizonte, seus raios seguem uma trajetória maior através da atmosfera, espalhando os menores comprimentos de onda e deixando apenas as radiações de maiores comprimentos, laranja e vermelho, alcançarem nossos olhos. Por causa desse forte espalhamento dos comprimentos de onda azul, o uso de imagens multiespectrais na porção do es-

pectro eletromagnético do visível requer correções atmosféricas. Por essa mesma razão é que, nas câmeras fotográficas aéreas, filtros amarelos são usados para cortar a radiação azul.

O espalhamento Mie ocorre quando a onda eletromagnética é de tamanho comparável ao tamanho das partículas atmosféricas. Nesse caso, o coeficiente de espalhamento é um coeficiente de área definido como a razão da frente de onda afetada pela partícula e a seção de área transversal da própria partícula. Em analogia ao espalhamento de Rayleigh, o espalhamento Mie tem a seguinte relação com o comprimento de onda (Equação 6.6):

$$M = \frac{I}{\lambda^{0,7 - 2,0}} \quad (6.6)$$

Como no espalhamento Mie o tamanho das partículas atmosféricas é consideravelmente maior que o considerado no espalhamento Rayleigh, esse tipo de espalhamento ocorrerá apenas com os maiores comprimentos de onda do visível. Em sensoriamento remoto, o espalhamento Mie se manifesta também por uma perda de contraste entre os alvos da imagem, o que acontece apenas sob atmosfera com partículas esféricas de vapor de água e poeiras, isto é, quando a atmosfera está mais densa. Nas imagens de sensoriamento remoto, o principal efeito dos espalhamentos Rayleigh e Mie é o de aumentar a reflectância real dos objetos, ocasionando uma diminuição do contraste entre os diferentes objetos presentes nas imagens. O exemplo dessa diminuição do contraste é mostrado na Figura 6.4. Nesta figura, a cena original possui dois objetos, um com brilho 8 para o alvo mais claro e outro com brilho 4 para o alvo mais escuro, considerando-se uma situação com ausência total de espalhamento atmosférico. O contraste entre esses dois alvos pode ser medido como uma razão do brilho dos dois alvos: 8/4 que é igual a 2. Quando esta mesma cena é colocada sob influência de uma atmosfera, onde são adicionadas quatro unidades de brilho devido

ao espalhamento, o alvo com brilho 8 aumenta para brilho 12 e o alvo com brilho 4 aumenta para 8. A razão de contraste dos alvos que era de 2 diminui para 1,5, reduzindo assim, a possibilidade de diferenciar os dois alvos nas imagens.

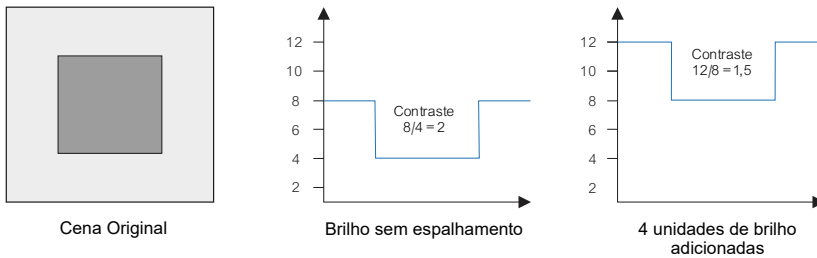


Figura 6.4. Razão de contraste entre os dois alvos de uma cena sem atmosfera e a perda de contraste entre eles quando a cena está sob uma atmosfera.

O mecanismo final de espalhamento é o não-seletivo. Ele ocorre quando o tamanho das partículas é muito maior que o comprimento de onda da faixa espectral que vai do visível ao termal ($0,4 \mu\text{m}$ a $14 \mu\text{m}$). Daí a razão de ser denominado de não-seletivo em relação ao tamanho da onda. Ele tem a soma da contribuição dos três processos envolvidos na interação da radiação com as partículas. O espalhamento não-seletivo ocorre em atmosferas muito densas, com neblinas e nuvens e resulta em severas atenuações da radiação refletida pelos alvos ao sensor. É o responsável pela cor branca das nuvens, que espalham a radiação para cima com alta intensidade. A Tabela 6.1 resume as características dos três tipos de espalhamento.

Tendo definido esses efeitos, estamos agora em condições de determinar como a radiancia medida pelo sensor é afetada pela presença da atmosfera. Primeiro, a irradiância total na superfície da Terra agora se torna (Equação 6.7):

$$E_G = E_{A\lambda} \times T_\theta \cos\theta \Delta\lambda + E_D \quad (6.7)$$

Tabela 6.1. Tipos de espalhamento atmosférico.

Tipo de espalhamento	Dependência com comprimento de onda	Tamanho das partículas	Tipo de partículas	Efeitos na redução do contraste
Rayleigh	λ^{-4}	$\ll \lambda$	Moléculas de gás	Visível de comprimento de onda curto
Mie	$\lambda^{0.7}$ a λ^{-2}	0,1 a 10λ	Gotas de água, aerossóis	Visível de comprimento de onda longo
Não-seletivo	λ^0	$>10 \lambda$	Nuvens, neblinas, névoas	Afeta igualmente todos os comprimentos de onda

Em que, por simplificação, assume-se que a irradiância difusa do céu não é uma função do comprimento de onda na banda de onda de interesse. A radiância, portanto, devido a esta irradiância global do pixel, torna-se (Equação 6.8):

$$L_T = \frac{\rho}{\pi} \{E_{\Delta\lambda} T_{\theta} \cos \theta \Delta\lambda + E_D\} \quad (6.8)$$

Acima da atmosfera, a radiância total disponível para o sensor torna-se (Equação 6.9):

$$L_T = \frac{\rho T_{\theta}}{\pi} \{E_{\Delta\lambda} T_{\theta} \cos \theta \Delta\lambda + E_D\} + E_p \quad (6.9)$$

Essa quantidade, portanto, é a que deve ser usada na Equação 6.8 para relacionar o valor de nível digital com a radiância medida.

6.2. Modelos de Transferência Radiativa

A atmosfera afeta a radiância medida em qualquer ponto da imagem de duas maneiras aparentemente contraditórias. Primeiro, ela atua como um refletor, adicionando uma radiância extra ao sinal que é detectado pelo sensor. Segundo, ela atua com um absorvedor, atenuando a intensidade de energia que ilumina o alvo na superfície. Quando a atenuação é acima de um valor, a transparência da atmosfera torna-se opaca em determinados comprimentos de onda, impossibilitando a obtenção de imagens. Portanto, quando temos uma imagem, nela devemos apenas verificar a degradação que pode ter sido causada pelo espalhamento atmosférico, avaliando quanto de radiação foi adicionada ao pixel e que sabemos não ser proveniente do alvo. Na maioria das vezes, os efeitos do espalhamento atmosférico são sempre o de diminuir o contraste entre os alvos presentes na imagem. Assim, antes de realizar qualquer tipo de análise usando imagens de satélite, especialmente quando capturadas na porção do visível ao infravermelho próximo (NIR) do espectro eletromagnético, é essencial remover ou, pelo menos, reduzir as perturbações atmosféricas introduzidas durante a aquisição da imagem.

A remoção de perturbações atmosféricas, portanto, fornece o equivalente à reflectância de objetos na superfície da Terra ou reflectância da superfície abaixo da atmosfera. A obtenção de reflectância de superfície é um grande desafio para os usuários. Assim, vários métodos foram introduzidos que se enquadram principalmente em duas categorias, os baseados em informações intrínsecas das imagens (métodos relativos) e os baseados em princípios físicos (Tabela 6.2). Os métodos baseados em princípios físicos são os mais complicados. Contudo, são os mais precisos e fazem uso de informações auxiliares sobre profundidade óptica do aerossol (atenuação do feixe de radiação pelos aerossóis ao se propagar pela atmosfera), tipos de aerossóis, vapor de água e concentração de O_3 , além de informações geométricas (ângulo zenital e azimuth solar e do sensor). Nesse caso, para melhor estimar as condições atmosféricas durante o tempo de aquisição

da imagem, tais informações auxiliares precisam ser coletadas simultaneamente por meio de várias fontes de dados (*in situ* e/ou satélite).

Tabela 6.2. Aplicativos desenvolvidos para remoção dos efeitos atmosféricos.

Métodos	Aplicativos	Citação	Observação
Relativo	<i>Dark subtraction</i>	Chavez Jr. (1988)	
	<i>Internal average relative reflectance</i> (IARR)	Kruse et al. (1985)	
	<i>Flat field</i>	Roberts et al. (1986)	
	Método de linha empírica (ELM)	Smith e Milton (1999)	
Absoluto (físico)	<i>Atmospheric Correction Now</i> (ACORN)	Miller (2002)	MODTRAN
	<i>Fast Line-of-sight Atmospheric Analysis of Hypercubes</i> (FLAASH)	Matthew et al. (2000)	MODTRAN
	OPERA	Sterckx et al. (2015)	MODTRAN-5
	<i>Framework for Operational Radiometric Correction for Environmental Monitoring</i> (FORCE)	Frantz et al. (2016)	5S
	<i>Python Interface for the 6S Radiative Transfer Model</i> (Py6S)	Wilson (2013)	6S
	<i>Landsat Ecosystem Disturbance Adaptive Processing System</i> (LEDAPS)	Masek et al. (2006)	6S
	<i>Landsat Surface Reflectance Code</i> (LaSRC)	Vermote et al. (2016)	6S

Os métodos baseados em informações intrínsecas das imagens (métodos relativos) são os mais simples, pois utilizam números digitais existentes em pixels que são representativos de feições específicas e precisam apenas do arquivo de metadados da imagem para a estimativa. Eles incluem a subtração do objeto escuro (*dark object subtraction* - DOS) (Chavez Jr., 1988), método de linha empírica (ELM) (Smith e Milton, 1999), *Internal*

Average Relative Reflectance (IARR) (Kruse et al., 1985), *Flat Field* (Roberts et al., 1986), entre outros.

Quanto aos métodos físicos, destacam-se o *Moderate Resolution Atmospheric Transmission* (MODTRAN) (Berk et al., 2005), o *Simulation of the Satellite Signal in the Solar Spectrum* (5S) (Tanré et al., 1990) e a versão vetorial do *Second Simulation of the Satellite Signal in the Solar Spectrum* (6S) (Kotchenova et al., 2006) (Tabela 6.2). Desconsiderando-se algumas constantes, os aplicativos em média utilizam o mesmo modelo para radiância espectral de um pixel (L^*), fundamentado na superfície plana Lambertiana ou equivalente (Equação 6.10).

$$L^* = Ap / (1 - \rho_e S) + B\rho_e / (1 - \rho_e S) + L_a^* \quad (6.10)$$

em que

ρ = reflectância do pixel na superfície;

ρ_e = reflectância média da superfície sobre o pixel e sobre a região vizinha;

S = albedo global da atmosfera; e

L_a^* = radiância espalhada pela atmosfera.

A e B são coeficientes que dependem das condições atmosféricas e da geometria, mas não da superfície. Cada uma dessas variáveis depende do comprimento de onda, mas o índice foi omitido para simplificar a equação. O primeiro termo da equação corresponde à radiância que é refletida pela superfície em direção ao sensor. O segundo termo corresponde à radiância da superfície que é espalhada pela atmosfera e que chega até o sensor, incluindo os efeitos espaciais de mistura ou de adjacência.

6.2.1. Métodos relativos

Os métodos relativos são correções que não dependem de parâmetros atmosféricos ou de dados da superfície e sim das informações da imagem

para determinar a contribuição da atmosfera em cada banda espectral de forma indireta. Uma de suas limitações consiste na exigência de obtenção de informações sobre alvos específicos nas imagens analisadas. Na técnica de *Dark Subtraction*, é assumida que cada banda da imagem deveria conter alguns pixels com valores próximos ou iguais a zero, cujas origens mais comuns seriam as áreas de sombras de relevo. No entanto, devido aos efeitos do espalhamento atmosférico, é adicionado um valor de brilho a todos os pontos da imagem, por isso as áreas sombreadas deixam de exibir pixels com valores zero ou bem próximos de zero. Consequentemente, se os histogramas de cada banda fossem analisados, os pixels com valores digitais mais baixos não seriam zero.

No ELM, é realizada uma calibração radiométrica a partir da implementação da equação da linha empírica (Smith e Milton, 1999). A técnica realiza uma calibração por regressão linear, no qual pode ser aplicada caso sejam conhecidas as medidas espectrais de campo de alvos específicos e suas respectivas medidas na imagem. Assim, por meio da obtenção dos coeficientes do modelo linear e aplicação do modelo à imagem, subentende-se que os efeitos da variação da irradiância solar e da interferência atmosférica foram removidos.

O método IARR calcula o valor espectral médio da uma cena que é então dividido pelo valor espectral de qualquer pixel na cena. Essa abordagem é aplicável principalmente para imagens adquiridas em áreas áridas e sem vegetação (Gao et al., 2006). Já o método *Flat Field* assume que existe uma área na cena que possui reflectância homogênea e espectralmente uniforme, ou seja, o espectro tem pouca variação em função do comprimento de onda (por exemplo, areia de praia). O espectro médio dessa área é entendido basicamente como efeitos atmosféricos e da irradiância solar e então usado para a derivação de espectros de reflectância relativa de outros pixels na cena. Tanto a abordagem IARR quanto a abordagem de *Flat Field* não precisam de nenhuma medição de campo de espectros de reflectância de alvos de superfície.

6.2.2. Métodos físicos

Nos métodos físicos, é necessário o conhecimento das propriedades ópticas da atmosfera bem como do processo de interação da radiação com a atmosfera e com a superfície. Estes métodos são fundamentados na teoria da transferência radiativa (Chandrasekhar, 1960) que simulam a propagação da REM pela atmosfera. O código 5S é um exemplo de modelo que posteriormente evoluiu para o 6S (Vermote et al., 1997). Esses métodos foram propostos para modelar o sinal do sensor baseando-se nas seguintes características:

- a) Modelos estatísticos baseados nas bandas de absorções atmosféricas provocados pela absorção causada por CO_2 , O_3 , N_2O , O_2 , CH_4 e H_2O ; e
- b) Modelagem de várias funções atmosféricas para um tratamento completo do processo de espalhamento molecular por aerossóis, bem como a reflectância bidirecional de superfícies não uniformes, quando necessário.

Os dados de entrada previstos para a realização dos cálculos são a visibilidade e profundidade óptica de aerossol, a data e o horário de aquisição da imagem, a altitude orbital e a altitude topográfica média da área contida na imagem.

O modelo MODTRAN, derivado do *Low Resolution Transmission Model* (LOWTRAN), considera a maior resolução espectral de 2 cm^{-1} , modela as assinaturas de absorção da atmosfera como função da temperatura e pressão e calcula os parâmetros de modelos de banda para 12 moléculas. O MODTRAN vem sofrendo constantes atualizações para otimizar o seu desempenho. O FLAASH, oriundo do MODTRAN 4, considera a modelagem da radiância registrada pelo sensor como uma função da reflectância do pixel e do seu entorno, do albedo hemisférico descendente, da radiância de trajetória (a qual inclui os efeitos dos aerossóis e de Rayleigh), transmitância e da geometria de aquisição. O código MODTRAN 4 necessita, a priori, da definição do modelo de aerossóis, especificação da localização geográfica, elevação, altitude do satélite, data do voo e horário de passagem do

sistema, além de outros parâmetros como dados de concentração de CO_2 . Como saída, o FLAASH fornece:

- a) Concentração de vapor de água da coluna derivada de dados e relativo a mapas de elevação de superfície (por exemplo, de coluna de oxigênio);
- b) Quantidades de aerossol (valores de visibilidade) derivadas de objetos na cena com reflectâncias conhecidas; e
- c) Imagens atmosféricamente corrigidas, ou seja, reflectância de superfície com a H_2O e outras bandas de absorção e espalhamento de aerossol removidos para comprimentos de onda não térmicos (infravermelho de ondas curtas a ultravioleta), incluindo um efeito de “adjacência” da imagem.

O FLAASH oferece flexibilidade adicional, quando comparado aos outros programas de correção atmosférica, pois permite cálculos personalizados de transferência radiativa para uma ampla gama de condições, incluindo visualização *off-nadir* e todos os modelos de aerossol padrão MODTRAN. O ACORN, também oriundo do MODTRAN 4, usa tabelas de consulta calculadas com o código de transferência radiativa MODTRAN 4 para modelar a absorção de gás atmosférico, bem como os efeitos de dispersão molecular e de aerossol, convertendo as medições de radiância do sensor calibrado em reflectância de superfície aparente. Os gases são limitados pela elevação e pela geometria de observação. O vapor de água é estimado a partir dos dados pixel a pixel usando as bandas de absorção de vapor de água em $0,94\ \mu\text{m}$ e/ou $1.150\ \mu\text{m}$. Uma tabela de pesquisa para uma faixa de densidades de vapor da coluna de água é gerada usando MODTRAN 4 e, em seguida, ajustada em um sentido de mínimos quadrados contra os dados do espectrômetro de imagem. O ACORN fornece correção básica com aprimoramentos e alguns controles sobre as opções MODTRAN e recursos adicionais de correção multiespectral. O FLAASH fornece correções básicas com correções aprimoradas para efeitos de adjacência e é mais flexível para correção de dados hiperespectrais à luz das opções disponíveis e baseadas em MODTRAN.

6.3. Referências

- BERK, A.; ANDERSON, G. P.; ACHARYA, P. K.; BERNSTEIN, L. S.; MURATOV, L.; LEE, J.; FOX, M.; ADLER-GOLDEN, ST.; CHETWYND, J.; HOKE, M.; LOCKWOOD, R.; GARDNER, J.; COOLEY, T.; BOREL, C.; LEWIS, P. **MODTRAN 5: a reformulated atmospheric band model with auxiliary species and practical multiple scattering options: update**. In: PROCEEDINGS. SPIE 5806, Orlando, FL, EUA, 2005, p. 662-667.
- CHANDRASEKHAR, S. **Radiative transfer**. Nova York: Dover, 1960, 393 p.
- CHAVEZ JR., P. S. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. **Remote Sensing of Environment**, v. 24, n. 3, p. 459-479, 1988.
- CHUVIECO, E.; MARTÍN, M. P.; PALACIOS, A. Assessment of different spectral indices in the red-near-infrared spectral domain for burned land discrimination. **International Journal of Remote Sensing**, v. 23, n. 23, p. 5103-5110, 2002.
- FRANTZ, D.; RÖDER, A.; STELLMES, M.; HILL, J. An operational radiometric Landsat preprocessing framework for large-area time series applications. **IEEE Transactions on Geoscience and Remote Sensing**, v. 54, n. 7, p. 3928-3943, 2016.
- GAO, B. C.; DAVIS, C.; GOETZ, A. A review of atmospheric correction techniques for hyperspectral remote sensing of land surfaces and ocean color. In: IEEE INTERNATIONAL SYMPOSIUM ON GEOSCIENCE AND REMOTE SENSING (IGARSS 2006), Denver, CO, EUA. **Proceedings...** IEEE, 2006, p. 1979-1981.
- KOTCHENOVA, S. Y.; VERMOTE, E. F.; MATARRESE, R.; KLEMM, F. L. Validation of a vector version of the 6S radiative transfer code for atmospheric correction of satellite data. Part I: Path radiance. **Applied Optics**, v. 45, n. 26, p. 6762-6774, 2006.
- KRUSE, F. A. Comparison of Atrem, Acorn, and Flaash atmospheric corrections using low-altitude Aviris data of Boulder, CO. In: 13th JPL AIRBORNE GEOSCIENCE WORKSHOP, Pasadena, CA, EUA. **Anais...** Pasadena: NASA/JPL, 2004, 12 p.
- KRUSE, F. A.; RAINES, G. L.; WATSON, K. Analytical techniques for extraction geologic information from multichannel airborne spectroradiometer and airborne imaging spectrometer data. In: INTERNATIONAL SYMPOSIUM ON REMOTE SENSING OF ENVIRONMENT, São Francisco, CA, EUA. **Anais...** Ann Arbor: ERIM, v. 4, p. 309-324, 1985.
- MASEK, J. G.; VERMOTE, E. F.; SALEOUS, N. E.; WOLFE, R.; HALL, F. G.; HUEMMERICH, K. F.; GAO, F.; KUTLER, J.; LIM, T. K. A Landsat surface reflectance dataset for North America, 1990-2000. **IEEE Geoscience and Remote Sensing Letters**, v. 3, n. 1, p. 68-72, 2006.
- MATTHEW, M.; ADLER-GOLDEN, S.; BERK, A.; RICHTSMEIER, S.; LEVINE, R.; BERNSTEIN, L.; ACHARYA, P.; ANDERSON, G.; FELDE, G.; HOKE, M. Status of atmospheric correction using a MODTRAN4-based algorithm. In: AEROSENSE 2000, **Proceedings...** SPIE 4049, Orlando, FL, EUA, 2000, p. 199-207.

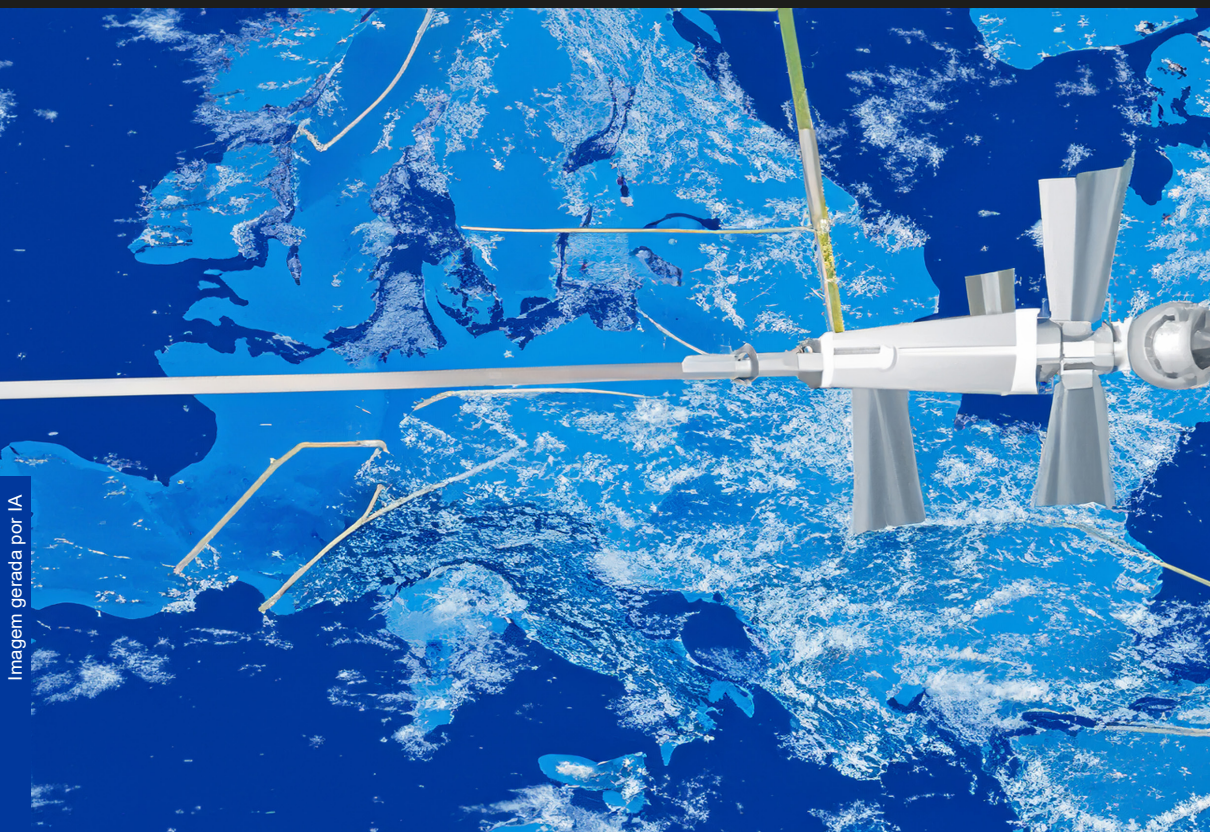
- MILLER, C. J. Performance assessment of Acorn atmospheric correction algorithm. In: AEROSENSE 2002, **Proceedings...** SPIE 4725, Orlando, FL, EUA, 2002, p. 438-449.
- ROBERTS, D. A.; YAMAGUCHI, Y.; LYON, R. J. P. Comparison of various techniques for calibration of AIS data. In: SECOND AIRBORNE IMAGING SPECTROMETER DATA ANALYSIS WORKSHOP. **Proceedings...** Pasadena: NASA/JPL, p. 21-30, 1986.
- SMITH, G. M.; MILTON, E. J. The use of the empirical line method to calibrate remotely sensed data to reflectance. **International Journal of Remote Sensing**, v. 20, n. 13, p. 2653-2662, 1999.
- STERCKX, S.; KNAEPS, E.; ADRIAENSEN, S.; REUSEN, I.; KEUKELAERE, L.; HUNTER, P.; GIARDINO, C.; ODERMATT, D. Opera: an atmospheric correction for land and water. In: Sentinel-3 for Science Workshop. **Proceedings...** ESA SP-734, Venice, Itália, 2015, p. 36.
- TANRÉ, D.; DEROO, C.; DUHAUT, M.; HERMAN, M.; MORCLETTE, J.; PERBOS, J.; DESCHAMPS, P. Y. Technical note description of a computer code to simulate the satellite signal in the solar spectrum: The 5S code. **International Journal of Remote Sensing**, v. 11, n. 4, p. 659-668, 1990.
- VERMOTE, E. F.; TANRE, D.; DEUZE, J. L.; HERMAN, M.; MORCLETTE, J. J. Second simulation of the satellite signal in the solar spectrum, 6S: An overview. **IEEE Transactions on Geoscience and Remote Sensing**, v. 35, n. 3, p. 675-686, 1997.
- WILSON, R.T. Py6S: a Python interface to the 6S radiative transfer model. **Computers & Geosciences**, v. 51, n. 2, p. 166-171, 2013.

CAPÍTULO 7

Correção Geométrica de Imagens

Rejane Ennes Cicerelli

Giuliano Sant'Anna Marotta



O processo de correção geométrica é uma etapa crucial em aplicações envolvendo sensoriamento remoto. Apesar de todo o procedimento empregado e do bom desempenho do controle de qualidade por parte dos órgãos responsáveis pela aquisição de dados, em alguns casos, as imagens corrigidas ainda apresentam erros de posicionamento. O processo de georreferenciamento de imagens minimiza os erros de posicionamento remanescentes, normalmente associados a distorções aleatórias e sistemáticas residuais (desconhecidas) que devem ser corrigidos pelo usuário. Esse processo é frequentemente conduzido manualmente, o que é uma tarefa trabalhosa, já que os métodos automáticos de georreferenciamento de imagens ainda estão longe de serem amplamente aplicados. Nesse capítulo, serão apresentadas as fontes de distorções geométricas em imagens referentes ao ponto visado, à plataforma e ao sensor, bem como as técnicas de georreferenciamento de imagens mais utilizadas. Também serão abordadas as formas de avaliação do processo de correção geométrica, além da tradicional inspeção visual.

7.1. Distorções Geométricas

A correção geométrica de imagens de sensoriamento remoto é fundamental para a integração, gerenciamento e análise de dados de várias fontes para muitas aplicações. A qualidade geométrica das imagens de sensoriamento remoto está diretamente relacionada com a presença de distorções que podem ser sistemáticas ou aleatórias. As distorções sistemáticas têm, como principais causas, tanto as imperfeições na lente, no sensor e nas mudanças de velocidade e atitude da plataforma, quanto os efeitos da curvatura e rotação da Terra. Como a correção dessas distorções depende do conhecimento das fontes de erro e de uma série de parâmetros relacionados ao sensor e à plataforma de aquisição da imagem, o procedimento fica a cargo dos laboratórios de produção de imagens.

As distorções aleatórias são ocasionadas principalmente pela visada central ou cônica dos sensores e pela alteração na posição dos sensores. Para corrigi-las, é necessário um conhecimento empírico que é obtido por meio de coleta de pontos de controle bem identificados na superfície da Terra, juntamente com os seus respectivos pontos homólogos na imagem a ser corrigida. Esse procedimento é denominado de georreferenciamento de imagens e consiste no uso de transformações geométricas simples para estabelecer uma relação entre as coordenadas de imagem e as coordenadas terrestres dispostas em um dado sistema de projeção cartográfica.

Para avaliar os motivos da existência das distorções geométricas, primeiramente é necessário considerar que uma imagem é formada de linhas sequenciais de pixels. Uma imagem é composta por N linhas e M pixels. As linhas são dispostas sucessivamente uma abaixo da outra. Como o campo de visada instantânea (IFOV) e a altitude da plataforma determinam a resolução no terreno, a qual possui dimensões de um quadrado, por exemplo, 20 m x 20 m, a formatação da imagem corresponde ao arranjo dos pixels em uma grade regular, tal como mostrado na Figura 7.1. Os pontos de intersecções da grade são as posições dos pixels e o espaçamento entre os pontos da grade é igual ao IFOV do sensor.

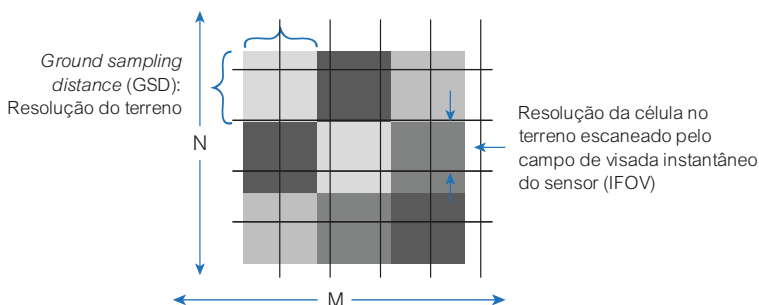


Figura 7.1. Grade usada para formatar os pixels da imagem digital.

As distorções geométricas referem-se aos desajustes na grade de pixels. Em uma imagem digital, isso deve ser entendido como a mudança de

posição do pixel na grade. Como o pixel possui um valor de brilho ou de nível de cinza, se ele mudou de posição, então, pode-se dizer que aquele pixel na grade teve uma mudança no valor de brilho. As distorções sistemáticas e aleatórias nos dados de sensoriamento remoto podem ser classificadas quanto à sua origem, podendo ser em relação ao objeto visado, à plataforma ou ao sensor.

7.1.1. Distorções geométricas referentes ao objeto visado

As distorções oriundas do objeto visado, no caso, a Terra, são devidas à rotação, à esfericidade e ao relevo. O efeito provocado pela rotação da Terra deve-se ao movimento relativo entre o sensor e a Terra. Durante o movimento do sensor, a Terra está no seu processo natural de rotação, provocando descontinuidades entre as linhas consecutivas de imageamento (Figura 7.2). Nessa situação, um ponto imageado fica deslocado para oeste devido ao movimento de rotação da Terra que é de oeste para leste no momento da aquisição da janela de imageamento. Caso as linhas imageadas estejam arranjadas em forma de grade, pode-se perceber que elas estarão deslocadas para leste, em relação ao terreno que elas representam. Para uma correção da posição relativa do pixel no terreno, é necessário realizar uma compensação. Esse deslocamento para oeste depende da velocidade relativa do satélite, da velocidade de rotação da Terra e do comprimento da faixa de imageamento. O deslocamento das linhas é da ordem de uma dezena de quilômetros para satélites como Landsat ou *China-Brazil Earth Resources Satellite* (CBERS). As imagens disponibilizadas para os usuários já vêm processadas com a correção necessária. Em sensores aerotransportados, esse tipo de distorção é pouco relevante.

A distorção provocada pela esfericidade da Terra é ocasionada pelo fato de a Terra não ser um plano e de o sensor imagear em uma perspectiva cilindro-cônica. O efeito dessa distorção é maior ao longo das linhas ou áreas imageadas, acentuando-se com a visada lateral (Figura 7.3).

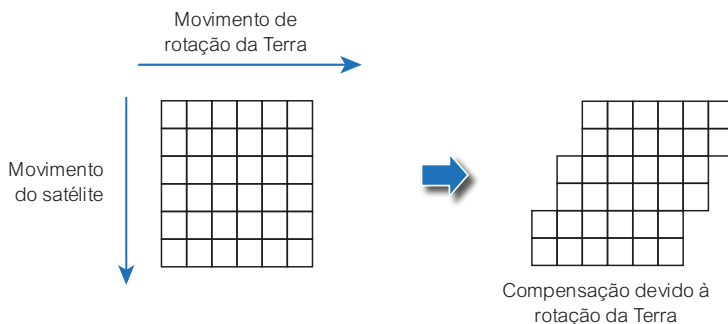


Figura 7.2. Efeito da rotação da Terra no tempo de aquisição da imagem.

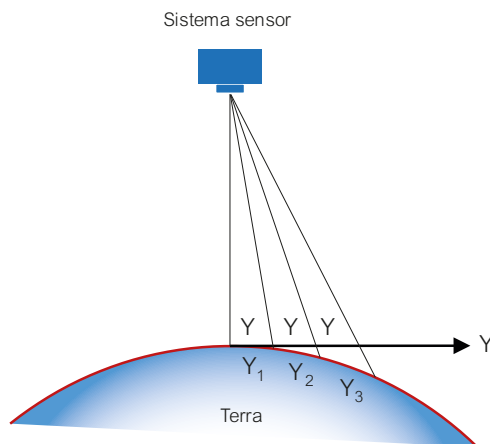


Figura 7.3. Efeito da esfericidade da Terra.

Para os sensores que possuem um pequeno ângulo de visada, isto é, FOV em torno de $\pm 7,5^\circ$, como são os casos dos sensores *High Resolution Visible* (HRV) do *Satellite Pour l'Observation de la Terre* (SPOT), *Operational Land Imager* (OLI) do Landsat 8, ou a câmera multiespectral (MUX) do CBERS-4A, com uma largura da faixa de imageamento no terreno entre uma e duas centenas de quilômetros, a distorção não é relevante. Para ângulos FOV da ordem de $\pm 40^\circ$ a $\pm 50^\circ$, a distorção da imagem passa a ser conside-

rável, pois a largura da faixa de imageamento pode atingir mais de 2.000 km. Nesse caso, o tamanho do pixel ao longo da linha de varredura pode ser de três a cinco vezes maior que o tamanho do pixel a nadir. É o caso do sensor *Moderate Resolution Imaging Spectroradiometer* (MODIS) das plataformas Terra (lançado em 1999) e Aqua (lançado em 2002) ou do satélite *National Oceanic and Atmospheric Administration* (NOAA).

As fontes de distorções devidas ao grande campo de visada do sensor são bem conhecidas dos fotointérpretes. Elas introduzem dois tipos de distorções nas imagens: o aumento da área dos pixels situados no extremo da imagem e o deslocamento das posições dos pixels ao longo da linha de varredura. Ambas são inerentes a todos os tipos de sensores, em função de a visada ser do tipo central ou cônica, semelhante aos sensores fotográficos. Portanto, nesse caso, a correção da imagem trata, prioritariamente, da remoção de erros introduzidos pelos campos de visada do sensor.

A Figura 7.4 ilustra a geometria que os ângulos β (IFOV) e θ (FOV) ocasionam nesses dois tipos de distorções. No caso de aumento da área do pixel, vê-se que os pixels, ao longo da linha de varredura, são amostrados a incrementos regulares, definidos pelo ângulo de visada instantâneo β , na extensão definida pelo movimento do ângulo de visada θ . Como resultado, a área efetiva do pixel na extremidade (T) é maior do que a nadir (N). Dessa forma, há uma pequena variação de escala entre as áreas do terreno a nadir (N) e *off-nadir* (T). Se a dimensão do pixel a nadir é p , então a sua dimensão na direção de varredura estabelecida pelo ângulo θ será $p\theta = \beta h \sec^2\theta$. Para pequenos valores de θ , esse efeito é negligenciável. Por exemplo, no satélite Landsat 7, o maior valor de θ é $7,50^\circ$ de modo que o pixel (T) na extremidade da linha apresenta um pequeno incremento de tamanho ($p\theta = 1,02$) do pixel a nadir. Quando os pixels são arranjados para compor a grade de pixel da imagem digital, eles são transcritos com o mesmo tamanho, ao passo que a área equivalente do terreno não é, mas isso não é tão relevante nesse caso porque o incremento no tamanho do pixel é pequeno.

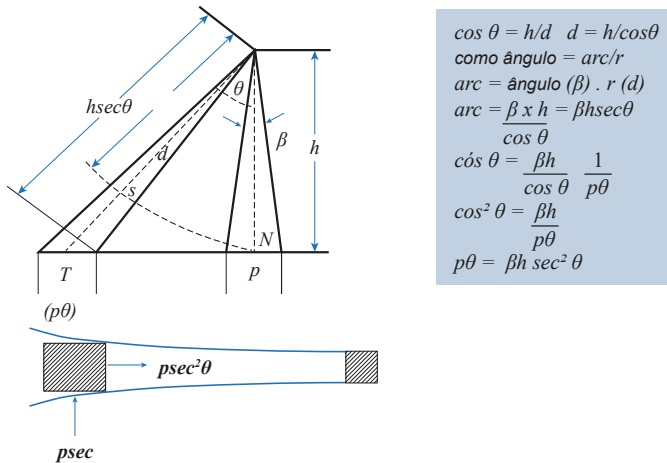


Figura 7.4. Distorção geométrica da imagem devido à visada cônica, causando deslocamento e aumento da área do pixel.

Na Figura 7.4, o arco centrado no sensor e que passa pelo ponto N posiciona S a um ângulo de visada θ normal à superfície e à mesma altura h que o ponto N . Nesse caso, os pixels em N e em S , medidos na visada inclinada, seriam do mesmo tamanho, sem qualquer distorção. No entanto, como a imagem é medida no terreno, o ponto S é a projeção inclinada na posição T . Se projetar S verticalmente ao terreno, poder-se-á, então, medir a sua distância em relação ao ponto T . Essa medida configura o deslocamento que os pixels progressivamente sofrem ao longo da linha de imageamento. Para medir esse deslocamento, basta calcular o valor do arco SN da Figura 7.4, conforme demonstrado por Richards (2013). O grau de distorção transversal à linha é dado por $SN/TN = \theta/\tan \theta$. No caso do satélite Landsat, para $(\theta/\tan)_{\max} = 0,99$, os pixels nas extremidades da linha, isto é, distantes 92,5 km a leste ou a oeste da posição a nadir, estarão deslocados cerca de 314 m da posição real no terreno. Trata-se de uma distorção geométrica significativa.

O efeito do deslocamento do pixel é a deformação visual da imagem. Como os pixels são registrados com uma separação angular igual ao IFOV, no ponto N a nadir em que $\cos \theta$ é igual a 1, os centros dos pixels são es-

paçados em p . A um ângulo θ em relação à posição T , os centros de pixels serão espaçados em $p \sec 2\theta$. Dessa forma, colocando-se os pixels sobre uma grade uniforme, haverá um maior número de pixels por área próxima a nadir do que nas áreas extremas da linha de imageamento. Dessa forma, a imagem sofrerá uma compressão transversal à linha de varredura, com os alvos deformando-se, como ilustra a Figura 7.5.

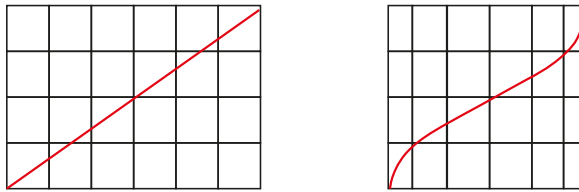


Figura 7.5. Efeito de distorção da imagem por compressão da linha de varredura devido ao *Instantaneous Field of View* (IFOV) e à taxa angular de varredura constante. Fonte: Richards (2013).

O efeito causado pelo relevo ou pela diferença de altura entre objetos que compõem a superfície terrestre leva a uma distorção devida à diferença de altitude. Essa distorção é diretamente proporcional às diferenças de nível entre os pontos. Esse efeito é minimizado para grandes altitudes de voo, na qual a imagem tende a adotar as características de uma projeção ortogonal, como ilustrado na Figura 7.6.

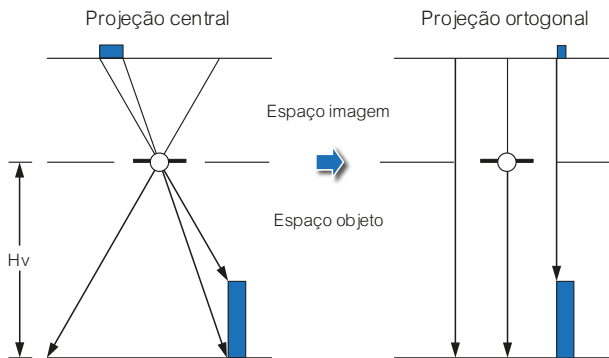


Figura 7.6. Distorção devido ao relevo.

A principal característica dessa distorção é que os objetos detectados no centro da imagem, assumindo imagem registrada a nadir, são isentos de deslocamento. À medida que se afastam do centro da imagem, as distorções vão se agravando. Essa distorção pode ser representada pela Equação 7.1:

$$\delta_R = \frac{r \times h_o}{H_v} \quad (7.1)$$

em que

δ_R = deslocamento devido ao relevo;

r = distância entre o objeto imageado e o centro da imagem;

h_o = altura do objeto; e

H_v = altitude do sensor.

Supondo um sensor com altitude de aproximadamente 450 km, distância focal de 8.800 mm e dimensão do pixel no terreno de 0,6 m, o tamanho do pixel na imagem será de 0,1173 mm. Supondo que um objeto com 100 m de altura seja detectado pelo sensor e representado na imagem a uma distância de 13.000 pixels do centro da imagem, isto é, próximo à borda da imagem, e assumindo uma inclinação do eixo focal do sensor em relação ao nadir igual a zero, haverá um deslocamento devido ao relevo de 0,034 mm, ou seja, cerca de 3 pixels.

7.1.2. Distorções geométricas referentes à plataforma

As principais distorções da imagem durante o imageamento, inerentes à plataforma, são devidas à altitude, velocidade e atitude. Segundo Richards (2013), variações na elevação ou na altitude da plataforma proporcionam variações de escala e, por consequência, mudanças no tamanho do pixel, como apresentado na Figura 7.7.

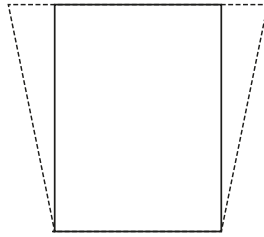


Figura 7.7. Distorção devido à variação da altitude da plataforma.

Variações de velocidade da plataforma alteram o intervalo de tempo (Δt_s) de passagem do sensor e, por consequência, o tamanho do pixel na direção da trajetória. Portanto, esse efeito proporciona mudança de escala em função da direção de trajetória do sensor. Em se tratando de sensor de imageamento em linha, se a velocidade da plataforma for menor que a nominal, haverá superposição positiva no imageamento com parte da superfície da Terra, sendo essa imageada duas vezes. Caso contrário, haverá superposição negativa, o que criará faixas na Terra não imageadas, como ilustrado na Figura 7.8.

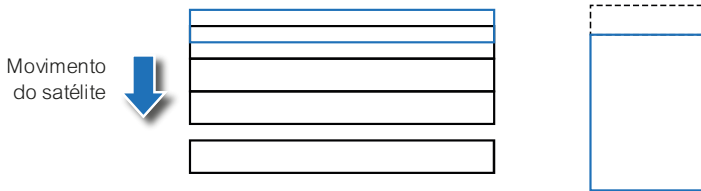


Figura 7.8. Distorções devido à variação da velocidade.

As distorções relativas à atitude da plataforma são representadas por três ângulos denominados de guinada (κ), arfagem (φ) e rolamento (ω), que inferem rotações e deslocamentos longitudinal e transversal de varredura. O ângulo ω é responsável por ligeira distorção na varredura, devido a não ortogonalidade da visada. Já a variação do ω ao longo da cena provoca o não alinhamento das varreduras consecutivas. Essa distorção afeta a imagem apenas no sentido da varredura, como ilustrada na Figura 7.9.

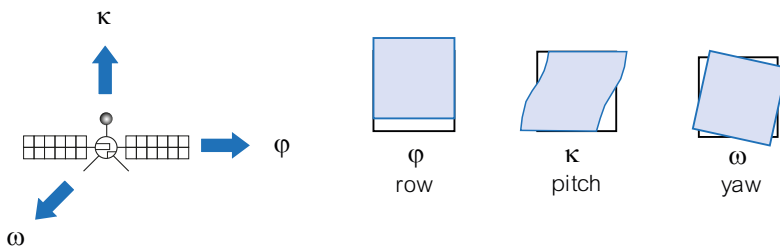


Figura 7.9. Distorções devido à altitude do sensor.

De forma semelhante ao rolamento (ω), porém, na direção transversal à varredura, o ângulo φ é responsável por pequenas distorções na imagem, devido a não ortogonalidade da visada. A variação desse ângulo ao longo da cena provoca superposições, positivas ou negativas, entre varreduras consecutivas. A guinada é responsável por distorções representadas por um não alinhamento das varreduras devido à rotação sofrida pelo sensor. A variação do κ ao longo da cena pode provocar superposição positiva ou negativa entre duas varreduras consecutivas, resultando em um efeito semelhante a um leque.

7.1.3. Distorções geométricas referentes ao sensor

As distorções referentes ao sensor são caracterizadas segundo seu tipo, sendo os de arranjo linear e o de quadro os mais comuns para aquisição de imagens de sensoriamento remoto. Nos sensores de arranjo linear, problemas em seu desenvolvimento proporcionam erros que podem estar associados à mudança de dimensão do fotodetector (pixel), translação ou rotação dos segmentos dos fotodetectores no plano focal em relação à posição nominal e flexão em arco ao longo da linha de varredura.

Na correção decorrente da mudança na dimensão do pixel, inicialmente é definido o sistema de referência da linha de detectores, como ilustrado na Figura 7.10, cuja origem se encontra no centro de cada linha e do detector

central. O eixo y_s se encontra ao longo da linha de varredura e o eixo x_s é perpendicular ao eixo y_s e ao longo da trajetória da plataforma. Como y_s é a referência do eixo na direção da linha e a imagem sofre distorção radial em cada seguimento ao longo da linha de varredura, pode-se dizer que as distorções devido ao tamanho do pixel se encontram ao longo do eixo y_s .

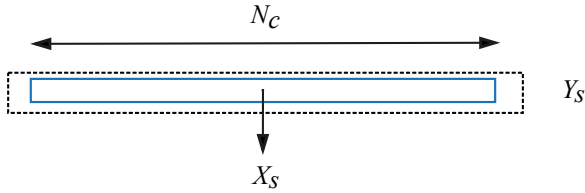


Figura 7.10. Efeito do tamanho do pixel na direção da linha de varredura.

Considerando-se um pixel com dimensão (p_x, p_y) , uma variação na dimensão do pixel ao longo da linha de detectores dp_y provocará uma alteração na coordenada y dada por (Equação 7.2):

$$y_p = y \frac{dp_y}{p_y} \quad (7.2)$$

em que

y = posição do pixel na linha de varredura.

Como uma variação no tamanho do pixel tem, como efeito, uma alteração na escala da imagem, conclui-se que o erro dy_p é fortemente correlacionado com o fator de escala na direção y_s , com variação da distância focal e com a distorção radial do sistema de lentes. Na direção x_s , a correção é dada por (Equação 7.3):

$$dx_p = dp_x \quad (7.3)$$

Os erros devidos às translações na linha de fotodetectores nas direções x e y podem ser modelados inserindo-se as constantes dx_c e dy_c , respectivamente, no modelo.

A correção devida à rotação (θ) dos segmentos de fotodetectores no plano focal em relação à posição nominal (Figura 7.11) é necessária se o sensor não estiver bem alinhado com a direção da trajetória, produzindo assim, erros angulares nas direções x_s e y_s (dx_θ dy_θ) em cada seguimento.

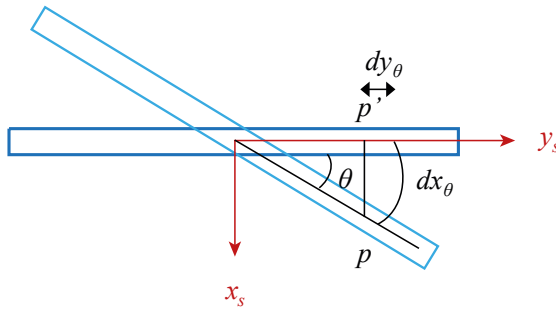


Figura 7.11. Efeito da mudança ou rotação dos segmentos CCD no plano focal em relação à posição nominal.

Essa distorção pode ser corrigida compensando o ângulo formado entre a linha de varredura e a posição nominal, por meio das seguintes expressões (Equações 7.4 e 7.5):

$$dy_\theta = y - y \times \cos\theta = y \times (1 - \cos\theta) \quad (7.4)$$

$$dx_\theta = y \times \sin\theta \quad (7.5)$$

A distorção devida à flexão em arco no plano focal, ao longo da linha de varredura, se dá quando um seguimento de fotodetectores sofre deformação em arco, como mostrado na Figura 7.12. O tamanho do arco é descrito

pelo ângulo central (δ) referenciado ao sistema adotado para cada linha de varredura.

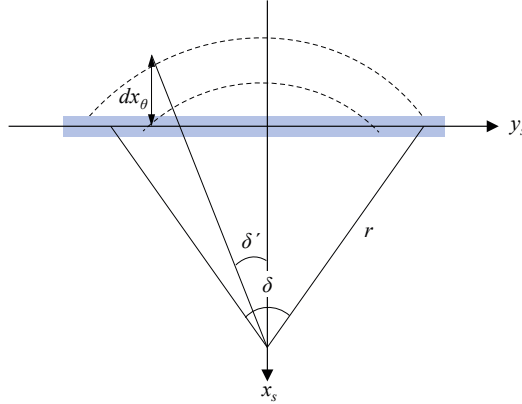


Figura 7.12. Linha de flexão no plano focal. Fonte: Poli (2005).

O raio r desse arco pode ser calculado em função do comprimento da linha de fotodetectores em números de pixels (Np), do tamanho de cada pixel na direção y (p_y) e do ângulo δ por (Equação 7.6):

$$r = \frac{Np \times p_y}{2 \times \sin \frac{\delta}{2}} \quad (7.6)$$

O efeito dessa curvatura tende a ser mais evidente com o aumento do ângulo nadiral. Segundo Poli (2005), esse efeito pode causar erros significantes na direção x , indicado por dx_δ . No centro da linha ($y_s = 0$), esse erro é máximo e dado por (Equação 7.7):

$$dx_\delta = - \left(r - r \times \cos \frac{\delta}{2} \right) \quad (7.7)$$

Nas bordas, esse erro é nulo. Para qualquer outro valor de y_s , ele pode ser modelado inserindo um ângulo δ' (Equação 7.8):

$$0 \leq \delta' \leq \frac{\delta}{2} \quad (7.8)$$

Definido como (Equação 7.9):

$$\delta' \approx \frac{y}{r} \quad (7.9)$$

Utilizando δ' no cálculo da deformação, tem-se que (Equação 7.10):

$$dx_\delta = - \left(r \times \cos \delta' - r \times \cos \frac{\delta}{2} \right) = - r \times \left(\cos \delta' - \cos \frac{\delta}{2} \right) \quad (7.10)$$

Se y_s é negativo, deve-se multiplicar os valores de dx_δ por -1.

Outra fonte de distorção que se torna necessário modelar e corrigir são as distorções das lentes. Essas distorções são mais evidentes em sensores de grande abertura angular, como no caso de sensores aerotransportados em baixa altitude. Segundo Wolf et al. (2014), em fotogrametria, as distorções das lentes produzem imagens com posições deslocadas de suas localizações ideais. Modelos matemáticos são usados para modelar essas distorções que podem ser decompostas em duas componentes, a distorção radial simétrica e a distorção descentrada.

A distorção radial simétrica pode ser encarada como uma parcela não desejável da refração sofrida pela luz ao atravessar uma lente. Essa distorção é inerente ao processo de fabricação das lentes. Já a distorção descentrada deve-se à impossibilidade do fabricante em alinhar perfeitamente os eixos óticos das lentes. Em fotogrametria orbital, a distorção descentrada

da lente, descrita pelas constantes p_1 e p_2 , é modelada por (Equações 7.11 e 7.12):

$$dx_d = p_1 \times (r^2 + 2 \times \bar{x}_p^2) + 2 \times p_1 \times \bar{x}_p \times \bar{y}_p \quad (7.11)$$

$$dy_d = p_2 \times (r^2 + 2 \times \bar{y}_p^2) + 2 \times p_1 \times \bar{x}_p \times \bar{y}_p \quad (7.12)$$

A distorção radial simétrica ocorre em função de coeficientes angulares k_1 , k_2 e k_3 e pode ser modelada como (Equações 7.13 a 7.15):

$$dx_r = (k_1 \times r^2 + k_2 \times r^4 + k_3 \times r^6) \times \bar{x}_p \quad (7.13)$$

$$dy_r = (k_1 \times r^2 + k_2 \times r^4 + k_3 \times r^6) \times \bar{y}_p \quad (7.14)$$

em que

$$r^2 = \bar{x}_p^2 + \bar{y}_p^2 \quad (7.15)$$

Segundo Poli (2005), além das distorções simétrica e descentrada, outros possíveis erros que podem ocorrer em sistemas ópticos de sensores orbitais são o deslocamento do ponto principal com coordenadas (x_p, y_p) , que é modelado com incremento de constantes Δx_p e Δy_p nas direções x_s e y_s , e uma variação Δf na distância focal f , cujo efeito pode ser modelado como (Equações 7.16 e 7.17):

$$r^2 = \bar{x}_p^2 + \bar{y}_p^2 \quad (7.16)$$

$$dx_f = -\frac{\Delta f}{f} \times \bar{x}_p \quad (7.17)$$

em que

$$\bar{x}_p = x - x_p \quad ; \text{ e}$$

$$\bar{y}_p = y - y_p$$

Com relação à variação de escala (s), com ênfase em sensores de arranjo linear na direção y_s , a equação pode ser descrita da seguinte forma (Equação 7.18):

$$dy_s = s_y \times \bar{y}_p \quad (7.18)$$

7.2. Georreferenciamento de Imagens

Imagens corretamente georreferenciadas e/ou registradas são necessárias para várias operações que envolvem dados geoespaciais, tais como fusão de imagens multitemporais ou de multissensores, sobreposição com conjunto de dados espaciais e mapas analógicos ou digitais existentes, detecção de mudanças, atualização de mapeamentos, ou integração de dados realizados em sistemas de informação geográfica. Georreferenciar uma imagem é tornar suas coordenadas terrestres conhecidas em um dado sistema de coordenadas e de referência geodésica, podendo ou não ser adotado um sistema de projeção. O registro de uma imagem corresponde ao processo de superposição de uma mesma cena que aparece em duas ou mais imagens de bandas diferentes ou de datas diferentes, tal que os pontos correspondentes nessas imagens coincidam espacialmente.

Atualmente, a geometria bem caracterizada dos sistemas sensores, combinadas ao sistema de referência do sensor de estrelas *Star Tracker Sensor*, ao sistema inercial *Inertial Measurement Units*, à determinação de posição a bordo do sensor de alta precisão com Sistema de Navegação

Global por Satélite (GNSS) e ao uso de modelos digitais de elevação (MDE) precisos, proporcionam precisão de alguns pixels na geolocalização, dependendo do sistema de satélite/sensor. Apesar de todo o método empregado e do bom desempenho do controle de qualidade dos órgãos responsáveis pela aquisição das cenas, em alguns casos, as imagens corrigidas ainda apresentam erros de posicionamento.

Esses erros remanescentes são devidos principalmente ao conhecimento insuficiente da altitude do satélite, de ângulos de montagem influenciados termicamente entre o sensor óptico e a unidade de medição de atitude, fatores decorrentes da distorção panorâmica, da refração atmosférica, de variações de relevo e da não linearidade na varredura do IFOV de um sensor. O processo de georreferenciamento das imagens minimiza o erro de posicionamento remanescente, normalmente associados às distorções não sistemáticas (aleatórias) e sistemáticas residuais (desconhecidas), e deve ser executado pelo usuário. Por isso, as informações de pontos de controle de campo são necessárias para ajustar os parâmetros do modelo do sensor, a fim de alcançar alta qualidade posicional.

Os pontos de controle de campo (GCP) podem ser obtidos manualmente usando mapas de referência, imagens ortorretificadas ou por medições em campo por meio do posicionamento pelo GNSS. A definição dos GCPs é realizada a partir da seleção de objetos identificáveis na imagem bruta (imagem a ser corrigida) e normalmente correspondem a feições retas como arestas, contornos, interseções de linha, cantos, cruzamento de estradas, pontes e calçadas. Em todos os casos, a medição de GCPs é uma tarefa morosa e onerosa, especialmente no contexto do processamento de dados em massa de centenas de imagens. O processo é ainda mais complexo quando as imagens são predominantemente cobertas por feições naturais, como, por exemplo, florestas densas, desertos, corpos de água amplos, entre outras.

No processo de correção, numerosos GCPs bem distribuídos são localizados tanto em termos de suas coordenadas da imagem u e v na imagem

bruta como em termos de suas coordenadas no terreno x e y . O processo nada mais é do que a transformação geométrica da imagem que consiste em associar dois sistemas de coordenadas de imagem (u, v) e terrestre (x, y), no qual um deles passa a ser referência para a transformação. Quando o processo for concluído, a imagem estará então georreferenciada, com todos os pixels podendo ser lidos por coordenadas referenciadas a um sistema terrestre, tal como se localiza um objeto ou uma feição em uma carta topográfica.

O processo de correção, ou georreferenciamento, é realizado em quatro etapas: (i) identificação dos pontos de controle; (ii) correspondência entre as feições da imagem a ser referenciada e as coordenadas de referência; (iii) aplicação do modelo de transformação; e (iv) cálculo do novo valor de brilho para o pixel na imagem de saída corrigida, por meio de reamostragem ou interpolação de pixels da imagem original. A primeira e a segunda etapa, como já citado, preveem a seleção dos GCPs, a obtenção das coordenadas no sistema de referência terrestre e a identificação/detecção acurada na imagem bruta. Caso o modo tela seja aplicado, espera-se a detecção dos GCPs na imagem de referência e na imagem bruta.

7.2.1. Aplicação do modelo de estimação

A terceira etapa do processo de correção geométrica consiste na definição da função de mapeamento que irá relacionar as coordenadas de imagem (linha e coluna) e as coordenadas de referência. A partir disso, são definidos dois sistemas de coordenadas cartesianas, um que define a localização dos pontos no mapa (x, y), e outro sistema para a localização dos pixels na imagem (u, v linha e coluna) na imagem a ser corrigida, conforme ilustrado no esquema da Figura 7.13. Se essas funções de mapeamento são conhecidas nas suas formas, pode-se então localizar um ponto na imagem conhecendo-se a sua posição no mapa.

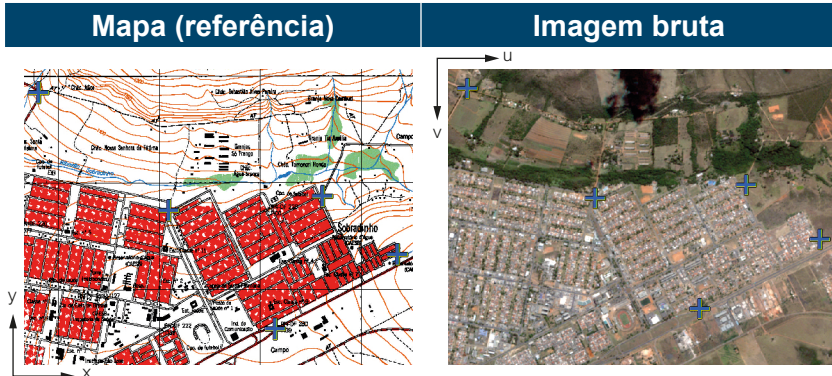


Figura 7.13. Sistemas de coordenadas mapa-imagem para localização dos pontos.

Com isso, constrói-se uma versão corrigida da imagem original. Considera-se então que esses dois sistemas de coordenadas podem ser relacionados por um par de funções de mapeamento espacial, f e g , tal que (Equações 7.19 e 7.20):

$$u = f(x, y) \quad (7.19)$$

$$v = g(x, y) \quad (7.20)$$

- **Modelo rigoroso**

Uma das funções de mapeamento muito conhecidas em fotogrametria é denominada de modelo rigoroso de georreferenciamento e se baseia na equação de colinearidade. Tal procedimento relaciona a localização do pixel medido no plano focal com a localização terrestre de seu objeto correspondente (GCPs). Esse método é conhecido como rigoroso pela necessidade de se conhecer os parâmetros internos do sensor para a reconstrução da geometria interior do sensor no momento da aquisição da imagem, para

o processo de orientação interior, além da orientação angular (atitude) do sensor e sua relação com o terreno para o processo de orientação exterior. Trata-se de uma técnica muito utilizada na aerofotogrametria, porém, pouco aplicada para georreferenciamento de imagens orbitais, pois os parâmetros internos e precisos do sensor normalmente não são disponibilizados aos usuários, inviabilizando a utilização das equações de colinearidade.

- ***Funções polinomiais racionais***

Outro procedimento amplamente utilizado para pré-processamento de imagens de satélite são as chamadas funções polinomiais racionais. Nessa técnica, as orientações exterior e interior não são consideradas, mas sim as correções associadas à translação, rotação e escala, realizadas de forma direta na imagem. Assim, as funções de mapeamento u e v são escolhidas como sendo, na realidade, simples polinômios de primeiro, segundo e terceiro graus para desenvolver o processo. As transformações polinomiais fazem o vínculo entre as coordenadas de imagem e as coordenadas no sistema de referência terrestre. Usando um polinômio de segundo grau, que tem duas variáveis, temos as seguintes expressões que relacionam coordenadas da imagem e coordenadas da carta (referência) (Equações 7.21 e 7.22):

$$u = a_0 + a_1 x + a_2 y + a_3 xy + a_4 x^2 + a_5 y^2 \quad (7.21)$$

$$v = b_0 + b_1 x + b_2 y + b_3 xy + b_4 x^2 + b_5 y^2 \quad (7.22)$$

Para que esse sistema de equações possa ser construído, é preciso conhecer os valores dos coeficientes a_i e b_i . Esses coeficientes são estimados identificando-se as coordenadas de feições no mapa de referência que também podem ser identificadas nas imagens ou vice-versa. As coordena-

das de imagem (linha, coluna) são obtidas quando o usuário clica sobre a feição na imagem e os valores correspondentes das coordenadas de referência são obtidos por meio de produtos cartográficos ou de levantamentos geodésicos em campo.

Os pontos de controle devem representar feições discretas, tais como cruzamento de estradas ou de pistas de aeroportos, confluências de rios consolidados, feições proeminentes de linha de costa e interseções de linhas ou limites construídos pelo homem. Uma vez determinados os n pontos de controle e selecionada a transformação polinomial, um sistema de $2n$ equações é montado para resolver os parâmetros das funções, cuja quantidade varia em relação à ordem do polinômio adotado. Em uma equação do primeiro grau, ou transformação afim no plano, os coeficientes a_i e b_i , modelam duas translações, duas rotações e dois fatores de escala, totalizando seis parâmetros.

A escolha da ordem da função polinomial depende da intensidade da distorção da imagem. Normalmente, transformações de 1ª e 2ª ordem são usadas nas imagens de sensoriamento remoto. Quanto maior a ordem dos polinômios, mais próxima a curva de ajuste deve passar desses pontos. Contudo, a utilização de polinômios de grau muito elevado pode provocar algumas distorções na imagem gerada. Por isso, às vezes, o ajuste linear é uma adaptação aceitável. Porém, os polinômios de maior ordem apresentam maior acurácia na vizinhança dos pontos de controle. Essa característica pode conduzir a erros significativos e, conseqüentemente, a distorções naquelas regiões que estão fora dos limites definidos pelo conjunto de pontos de controle.

Richard (2013) exemplificou claramente que nem sempre uma ordem de transformação maior é mais adequada para a correção de uma imagem. A Figura 7.14 mostra os ajustes dos pontos de controle às curvas de diferentes ordens polinomiais. Aparentemente, a curva de 3ª ordem seria a melhor escolha, porque é a que passa mais próxima aos pontos de controle. No entanto, se mais pontos fossem plotados, é presumível que a tendência dos

pontos seguiria mais próxima à reta de 1ª ordem, enquanto a curva de 3ª ordem se desviaria dessa tendência. Transformações de 3ª ordem devem ser usadas em imagens com severas distorções, como as fotos aéreas e imagens de alta resolução espacial, tomadas *off-nadir*. O processo de corrigir distorções não lineares é também conhecido como *rubber sheeting*.

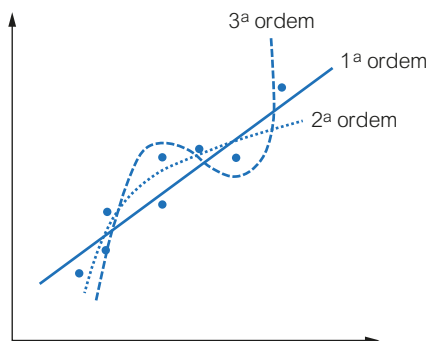


Figura 7.14. Escolha da ordem polinomial para melhor ajuste.

Fonte: Richards (2013).

É exigido um número mínimo de pontos de controle, conforme seja a ordem polinomial, mas, na prática, aumentar o número de pontos de controle aumenta a confiança no resultado, sem a necessidade de aumentar o grau do polinômio. Deve-se também ter em mente que a distribuição dos pontos de controle é muito importante, desde que as transformações polinomiais tendem a se comportar, adequadamente, apenas na região onde se encontram os pontos de controle. O ideal para uma correção homogênea da imagem é que os pontos recubram todos os cantos e a área central da imagem.

• Correspondência

A correspondência de imagens é uma técnica que vem ganhando espaço no sensoriamento remoto e desempenha um papel importante na automação do processo. A técnica se baseia na identificação de pontos ho-

mólogos entre a imagem de referência e a imagem a ser referenciada para identificar candidatos a GCP. O processo de correspondência usa uma pirâmide de múltipla resolução para lidar com grandes diferenças entre a imagem referência e a imagem a ser referenciada. A busca por correspondências começa nas imagens com resoluções mais baixas, onde as diferenças, devido à topografia, iluminação, perspectiva e afins, são menores. Os parâmetros estimados no final dessa etapa são então aplicados como valores iniciais ao nível seguinte.

Primeiro, os possíveis pontos de correspondência são selecionados na imagem de referência aplicando-se um operador de interesse. Com base nesses pontos, janelas padrão de n pixels x n pixels são selecionadas na imagem de referência e localizadas com uma precisão de cerca de um pixel na imagem bruta usando o máximo dos coeficientes de correlação normalizados, calculados por varredura na área de pesquisa. A área de pesquisa é determinada pela estimativa de transformações afins locais com base em pontos homólogos.

Uma limitação envolvida no uso da técnica é que naturalmente espera-se uma imagem de referência e que as posições aproximadas de GCPs devem ser conhecidas a priori em ambas as imagens. Além disso, as janelas em torno delas devem, de preferência, conter características bem contrastantes, a fim de permitir uma boa estimativa dos parâmetros. Esse método passou a ser empregado principalmente na aquisição de imagens por aeronaves remotamente tripuladas (RPAs) em virtude da sobreposição longitudinal e lateral entre as imagens bem como a aquisição simultânea de coordenadas por GNSS.

• **Ortorretificação**

A ortorretificação de imagens é o processo de correção da imagem, pixel a pixel, das distorções causadas pelo relevo, fazendo com que a imagem seja representada em perspectiva ortogonal. A ortorretificação é uma correção geométrica que resulta em bons resultados, em razão de in-

serir a função de mapeamento juntamente com os GCPs, parâmetros de orientação interior do sensor/satélite bem como de um MDE que garantirá a correção da imagem com relação à superfície topográfica. Tal operação é necessária em imagens de sensores remotos, especialmente quando as imagens representam superfícies bastante irregulares e com declividades acentuadas. Esse processo de correção é realizado principalmente pelo fato de que as imagens de satélites são obtidas por meio de perspectiva central, ou seja, são representadas em projeções cônicas (Figura 7.15).

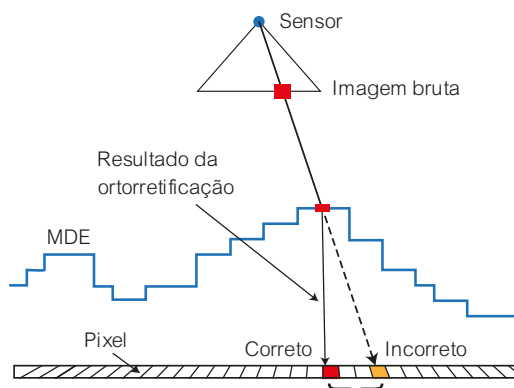


Figura 7.15. Geometria das distorções causadas pelo relevo na imagem.

Fonte: PCI Geomatics (2013).

Para a aplicação da técnica, é importante o conhecimento dos coeficientes polinomiais racionais denominados de *Rational Polynomial Coefficients* ou *Rational Function Coefficients*. Esses coeficientes refletem a realidade física da geometria da visada e corrigem as distorções inerentes à plataforma, ao sensor e à Terra. Assim, pressupõem-se que sejam conhecidos os parâmetros físicos relacionados ao sensor e à plataforma do satélite nos instantes de aquisição das imagens que nem sempre são disponibilizados aos usuários. Em vez disso, os provedores fornecem os coeficientes para a orientação e emprego de um modelo rigoroso alternativo.

7.2.2. Avaliação dos erros de transformação

Difícilmente consegue-se derivar coeficientes que não produzem erros. Isso em virtude da diversidade de distorções, conforme apresentado nos itens anteriores, e pelo fato de o usuário raramente identificar um ponto de controle em uma imagem de forma unívoca, pois muitas das feições não são visualmente discretas. Além disso, nem sempre se consegue uma boa distribuição de pontos de controle que contemple toda a área da imagem. Essas limitações introduzem erros nos coeficientes para a transformação polinomial, levando a uma imprecisão no georreferenciamento da imagem. Essa imprecisão pode ser representada ao imaginarmos a sobreposição da grade de pixel da imagem à grade de pixel do mapa de referência. O que ocorreria seria algo parecido com a Figura 7.16. À medida que se aumenta a quantidade de pontos de controle, melhor ajustada ficará a imagem em relação à imagem de referência.

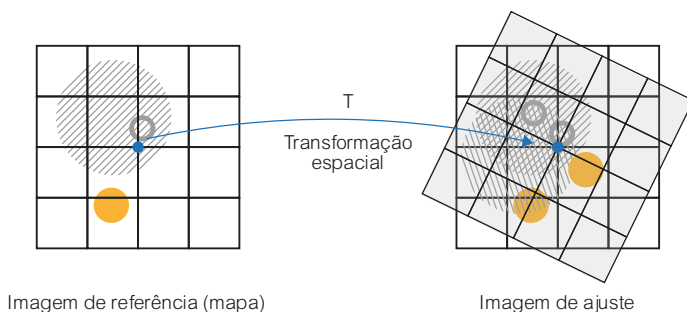


Figura 7.16. Erro na sobreposição da grade da imagem à grade do mapa devido à imprecisão dos pontos de controle.

Matematicamente, a precisão do georreferenciamento é geralmente feita calculando-se o erro médio quadrático (RMS) ou a diferença média quadrática entre as coordenadas de referência dos pontos de controle e as coordenadas obtidas na imagem corrigida após ajuste dos parâmetros. A Figura 7.17 mostra o ajuste dos pontos de controle a uma função polinomial de segunda ordem.

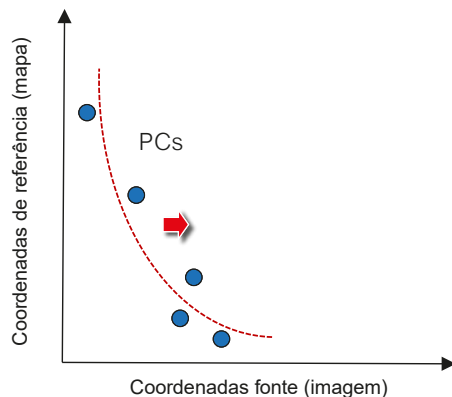


Figura 7.17. Pontos de controle ajustados à ordem de transformação polinomial.

Aqui, é correto dizer que o erro pode ser avaliado individualmente em cada ponto de controle utilizado por meio do resíduo (r_i) (Equação 7.23), ou globalmente por meio do RMS (Equação 7.24).

$$r_i = \sqrt{(x_i - x_r)^2 + (y_i - y_r)^2} \quad (7.23)$$

em que

x_i e y_i = coordenadas transformadas; e

x_r e y_r = coordenadas de referência.

$$RMS = \sqrt{\frac{1}{n} \sum_{i=1}^n (r_i^2)} \quad (7.24)$$

A Figura 7.18 mostra a diferença, nas direções X e Y, entre as coordenadas definidas para um ponto de controle (PC fonte) e as coordenadas ajustadas da imagem após o georreferenciamento (PC saída).

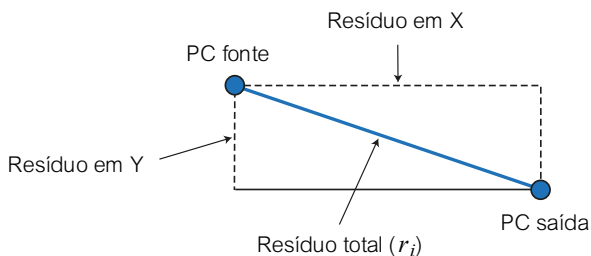


Figura 7.18. Diferença entre as coordenadas definidas para um ponto de controle e definidas na imagem após o georreferenciamento.

O valor de RMS pode ser expresso em números de pixels ou em unidade de medida adotado pelo sistema de coordenadas terrestres utilizado. Por exemplo, um RMS igual a 2 significa que o pixel transformado está distante dois pixels do pixel de referência. Se os pontos de controle mostram valores altos de RMS na direção X ou na direção Y, deve-se adicionar mais pontos naquela direção. Os valores de tolerância do RMS que normalmente são aceitos para os pontos de controle são:

- Imagens com pixel de 30 m: $\text{RMS} \leq 2 \text{ pixels} = 60 \text{ m}$;
- Mapas na escala 1:26.000: precisão próxima de 10 m; e
- Uso do receptor de navegação GNSS: precisão de aproximadamente 15 m.

Para o caso polinomial, é sempre recomendável começar com uma transformação de 1ª ordem. Pode-se computar repetidamente a transformação até conseguir um RMS aceitável. Após cada cálculo da transformação e do RMS, existem quatro opções para melhoramento dos resultados:

- 1) Excluir o ponto de controle com o maior RMS, assumindo que este é o ponto de controle de menor precisão. Outra matriz de transformação pode, então, ser computada dos pontos de controles restantes. Entretanto, se esse é o único ponto de controle em uma determinada região da imagem, isto pode causar um erro maior;
- 2) Tolerar um valor maior de RMS;

- 3) Selecionar somente os pontos que você tem maior confiança; e
- 4) Aumentar a ordem da transformação, criando alterações geométricas mais complexas na imagem.

Uma matriz de transformação pode ser então computada, acomodando os pontos de controles com um erro menor. Importante ressaltar que, caso o interesse for na avaliação da qualidade geométrica de acordo com padrões de exatidão cartográficos (PECs), é importante considerar GPCs diferentes dos adotados para a determinação da função de transformação. Atualmente são adotados PECs para mapeamento digital que segue metodologia e limiares para adoção de escalas em função do RMS alcançado, calculado por meio de GPCs diferentes dos utilizados no modelo.

7.2.3. Reamostragem

A quarta etapa do processo de correção geométrica refere-se ao cálculo do novo valor de brilho para o pixel na imagem de saída corrigida, por meio de reamostragem ou interpolação de pixels da imagem original. A reamostragem consiste em encontrar pontos na imagem correspondentes para cada localização na grade de pixel previamente definida sobre o mapa. Isso é feito automaticamente sem interferência do usuário. O espaçamento da grade deve ser escolhido de acordo com o tamanho do pixel da imagem resultante. Na suposição de que os pontos locados na imagem correspondem exatamente aos centros de pixels, esses pixels, com seus valores de brilho, são então simplesmente transferidos da imagem original para a apropriada localização na grade, para construir a imagem corrigida.

Mas como se espera, os centros da grade de pixels do mapa de referência não são transferidos para a exata localização do centro de pixels na imagem, isto é, a grade de pixel do mapa raramente coincide com a grade da imagem. É preciso então encontrar uma solução, a qual se resume em encontrar quais são os valores de brilho dos pixels a serem escolhidos para substituir os valores de brilho na nova grade. Essa técnica é o que se denomina de reamostragem ou interpolação dos valores de brilho da imagem.

Pode-se definir reamostragem como o processo de extrapolar os valores digitais para os pixels da imagem corrigida, a partir dos valores digitais dos pixels fontes. Isso quer dizer que os níveis de cinza que comporão a imagem corrigida residem na imagem de entrada. Segundo D'Alge (2001), baseada fortemente na ideia de otimização computacional, essa é uma fase que pode ser denominada de mapeamento inverso, que permite que se retorne à imagem original, para que se definam os níveis de cinza que comporão a imagem corrigida. A reamostragem é a última etapa do processo de georreferenciamento.

A Figura 7.19 exemplifica esse processo de mapeamento inverso. Nessa figura, a grade de pixel da imagem original está representada em linhas contínuas e a grade de pixel da imagem corrigida em linhas tracejadas. O pixel destacado na imagem original (coluna 430, linha 289) influencia radiometricamente pelo menos os quatro pixels vizinhos da imagem retificada (colunas 427 e 428 e linhas 288 e 289). O problema da reamostragem se resume em determinar qual desses quatro pixels vizinhos deverá assumir o valor de brilho do pixel em destaque, da imagem original.

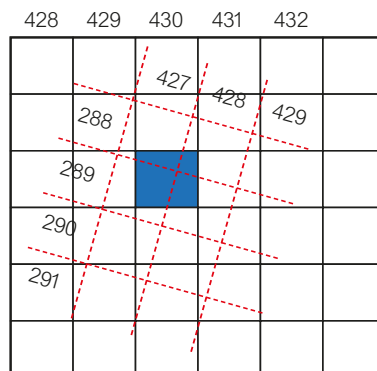


Figura 7.19. Reamostragem do valor de brilho da imagem corrigida a partir do brilho da imagem fonte.

Fonte: Brito e Coelho Filho (2002).

No processo de reamostragem dos pixels, vários métodos de interpolação foram desenvolvidos, incluindo-se a interpolação pelo método do vizinho mais próximo, a interpolação bilinear e a interpolação por convolução cúbica. A interpolação com reamostragem pelo vizinho mais próximo, também conhecida como interpolação de ordem zero, apenas transfere, para a imagem reamostrada, o valor de brilho do pixel da imagem original que estiver mais próximo. No exemplo da Figura 7.19, esse pixel será o da coluna 427 e linha 289. Esse método introduz descontinuidades geométricas (erros) na imagem reamostrada da ordem de 0,5 pixel, que é irrelevante para a visualização da imagem (Schowengerdt, 2007). Sua principal vantagem é a rapidez de processamento e de não requerer nenhum cálculo para derivar o pixel de saída. Além disso, essa reamostragem não altera os valores radiométricos da imagem original, necessários para aplicações que envolvem o conhecimento dos princípios físicos, químicos e biológicos dos alvos.

A interpolação com reamostragem bilinear (primeira ordem) toma, por base, os valores de brilho dos quatro vizinhos mais próximos ao centro do novo pixel (imagem corrigida). Duas interpolações lineares são executadas ao longo das linhas para encontrar o valor a ser interpolado. Utilizando-se uma janela de dimensão 2×2 , é calculada a média ponderada das distâncias dos centros dos quatro pixels ao centro do pixel de saída para se determinar a intensidade do valor de brilho do pixel de saída. O efeito é uma suavização da superfície, acompanhada de uma alteração dos valores de brilho da imagem original, o que pode ser problemático para subsequentes análises numéricas. Sua vantagem em relação ao método do vizinho mais próximo é a melhor precisão geométrica e o desaparecimento de descontinuidades (Novo, 2010).

A interpolação por convolução cúbica utiliza os 16 pixels ao redor. A interpolação é realizada ajustando polinômios cúbicos a cada coluna para depois interpolar um novo polinômio cúbico a esses resultados. Isso incorre em menor erro de interpolação. A reamostragem cúbica produz um aumento de brilho nas áreas com bordas bruscas, o que pode ser interessante para fins de fotointerpretação da imagem. A técnica é consideravelmente

mais complicada do que as anteriores, mas tende a produzir imagens com uma aparência bem mais "natural", embora alguma perda de feições de alta frequência acabe por ocorrer, uma vez que qualquer interpolador é, na verdade, um filtro passa-baixa. Os dados são alterados mais drasticamente do que pelos outros métodos, podendo não ser recomendado se a imagem for usada para classificação ou para análises numéricas que dependem da precisão dos valores radiométricos originais. Dos três métodos, o da interpolação por convolução cúbica é o mais preciso para gerar uma imagem geometricamente corrigida e deve ser utilizado quando o trabalho envolver o uso de imagens de diferentes resoluções espaciais.

A Figura 7.20 mostra a variação visual das diferentes técnicas de reamostragem. É importante visualizar como o processo de suavização da cena é crescente à medida que aumentamos o número de pixels envolvidos na aplicação da técnica.

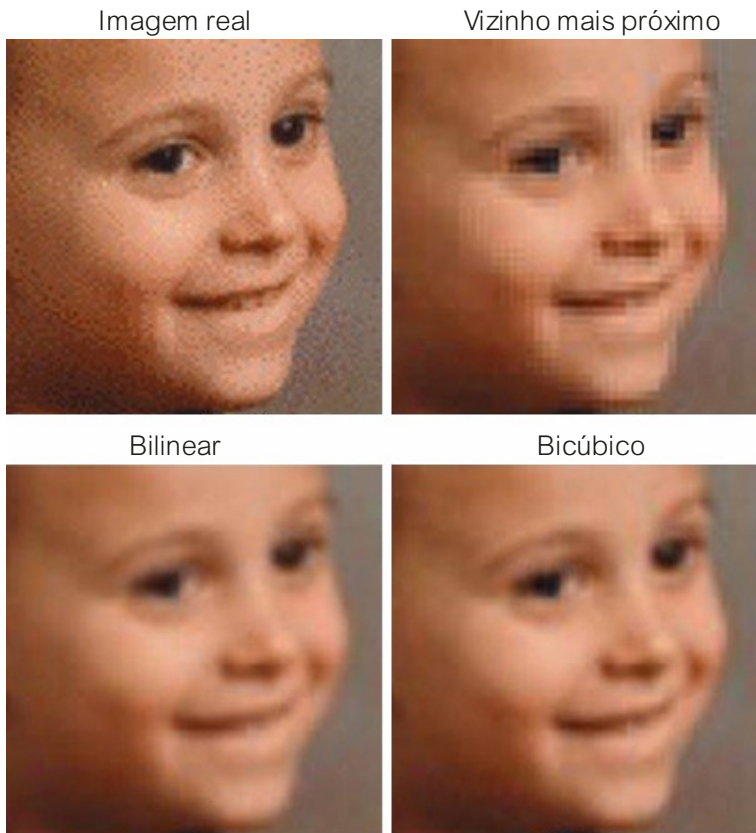


Figura 7.20. Diferenças visuais entre as técnicas de reamostragem.

Fonte: Hasan e Hoque (2017).

7.3. Referências

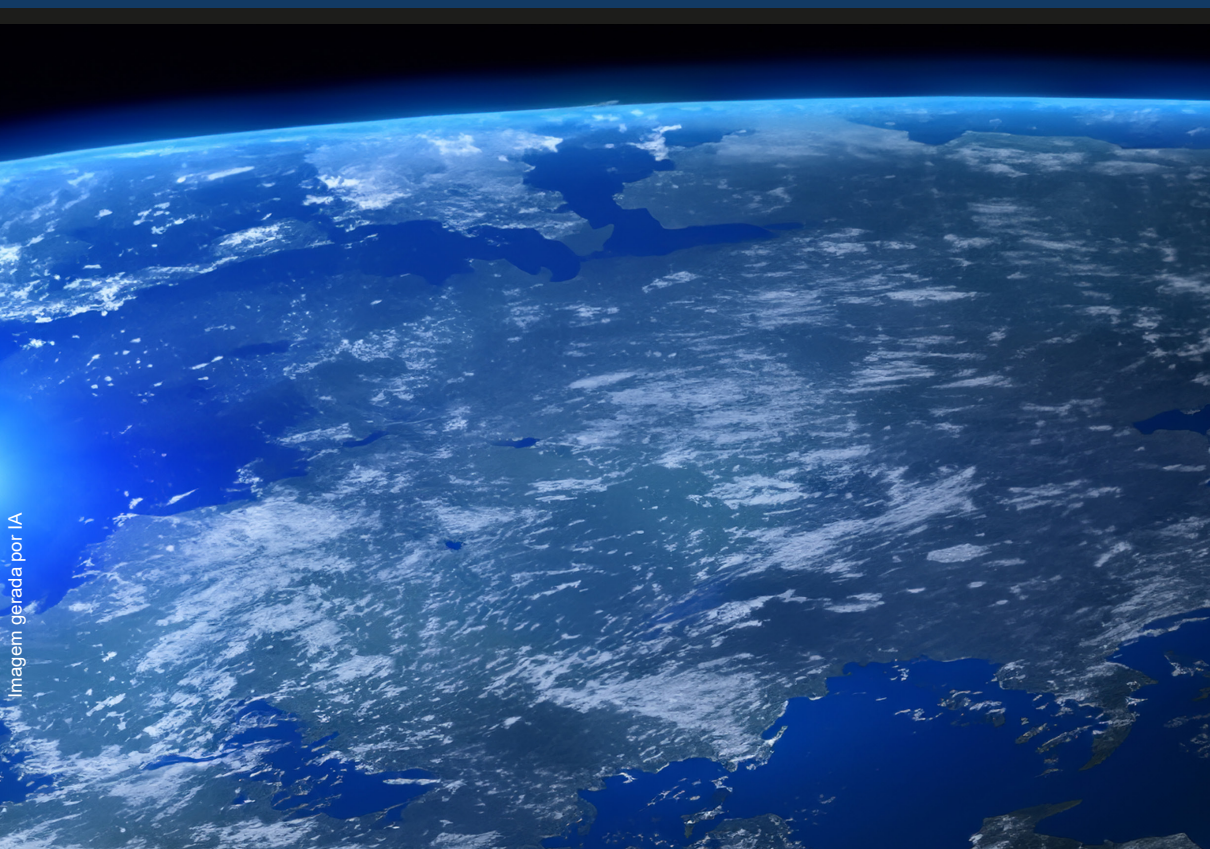
- BRITO, J. L. N. E. S.; COELHO FILHO, L. C. T. Fotogrametria digital. Rio de Janeiro: IME, 1ª ed., 2002.
- D'ALGE, J. C. L. **Cartografia para Geoprocessamento**. São José dos Campos: INPE, 2001, 33 p. (INPE-8079-PRE/3894).
- HASAN, M. S.; HAQUE, S. T. Single image super-resolution using back-propagation neural networks. In: INTERNATIONAL CONFERENCE OF COMPUTER AND INFORMATION TECHNOLOGY (ICCIT), 20. Dhaka, Bangladesh. **Proceedings**. IEEE, 2017, 5 p.
- NOVO, E. M. L. M. **Sensoriamento remoto**: princípios e aplicações, São Paulo: Blucher, 4ª ed., 2010.
- PCI GEOMATICS. **PCI Geomatics** Inc. Toronto, ON, Canada, 2013.
- POLI, D. **Modelling of spaceborne linear array sensors** (tese de doutorado em Ciências Técnicas). Instituto Federal de Tecnologia de Zurique, Zurique, Suíça, 2005, 205 p.
- RICHARDS, J. A. **Remote sensing digital image analysis**. An Introduction. 5a ed., Berlim: Springer, 2013, 494 p.
- WOLF, P. R.; DEWITT, B. A.; WILKINSON, B. E. **Elements of Photogrammetry with Applications in GIS**. Nova York: McGraw-Hill, 4a ed., 2014.

CAPÍTULO 8

Ampliação de Contraste

Paulo Roberto Meneses

Tati de Almeida



A mais simples das técnicas de transformação espectral de imagens é a ampliação de contraste baseada na distribuição histográfica dos dados, cujo único objetivo é melhorar a interpretabilidade visual da imagem, ampliando o contraste dos níveis de cinza das classes de alvos presentes na cena. A simplicidade das operações histográficas está no fato de que elas não promovem uma transformação total da imagem, mas apenas um reescalonamento dos níveis de cinza dentro do intervalo de quantização da imagem. Apesar de sua simplicidade, é a técnica mais usada pelos usuários para iniciar o primeiro contato com as diferentes bandas de um sensor e avaliar o potencial de informação de cada banda. A necessidade de ampliar o contraste da imagem deriva do fato de os sensores multiespectrais terem uma cobertura global, adquirindo imagens com um amplo leque de paisagens, desde cenas com baixos valores de radiância até com muita alta radiância. Isso requer um intervalo de valores digitais (nível de quantização) de no mínimo 8 bits por pixel para acomodar essa variação de radiância. O realce da imagem é efetuado no domínio espectral, de tal forma que cada pixel da imagem modifica o seu valor independentemente do valor do pixel vizinho. São operações efetuadas pixel a pixel e que podem ser ditas do tipo realce radiométrico. Os algoritmos dessa técnica embutem o conceito de um espaço espectral multidimensional.

8.1. Representação Histográfica

As operações de contraste por meio de expansão de histograma são feitas banda a banda. A escolha do melhor método depende das características de cada área, pois elas ditam a forma que o histograma assume. Há várias opções de técnicas de aumento de contraste, algumas baseadas em funções matemáticas e outras em funções estatísticas. Os histogramas são usados para representar uma grande quantidade de dados numéricos e para analisar as informações de forma mais fácil do que por tabelas. O histograma de uma imagem é uma representação estatística da distribui-

ção dos pixels da imagem em termos de número total de pixel de cada valor digital. Os dados são dispostos na forma de um gráfico XY em que os valores no eixo X representam os números digitais, enquanto os valores em Y representam o número total de pixels de cada número digital, isto é, a frequência. Cada número digital deve ser entendido como uma classe de alvo de um valor de radiância.

Na Figura 8.1, é mostrada a representação típica de histograma de uma imagem que frequentemente tem a aparência de uma curva normal. O histograma contém apenas informação radiométrica da imagem, sem nenhuma informação espacial. Um intérprete experiente, ao localizar um ponto na curva histogrâmica, é capaz de estimar a classe do alvo que ele representa e ter uma ideia da área que ocupa na imagem. No entanto, apenas pelo histograma, não saberá dizer onde o alvo se encontra e nem se ocorre na forma de uma única área ou se encontra fragmentado em pequenas áreas na imagem.

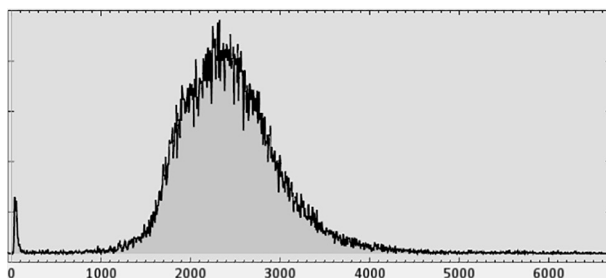


Figura 8.1. Histograma da imagem de uma banda com os valores de brilho ou níveis de cinza no eixo X definidos pela resolução radiométrica do sensor e no eixo Y o número de pixels por classe de brilho.

A análise de um histograma envolve basicamente três parâmetros, tendência central, variabilidade e forma. A tendência central da distribuição de frequência de números digitais ou de brilhos da imagem é controlada pelo intervalo digital de quantização da imagem (por exemplo, 8 bits = 256 valores; 16 bits = 65.536 valores). Duas maneiras de representar a distribuição

central típica são pelo valor mais frequente da variável, ou seja, a moda, que corresponde ao pico mais alto da curva, e pelo valor da média da variável amostrada. A média está relacionada com uma grandeza que indica algo sobre o nível de reflectância dos alvos presentes na cena. Média e moda se igualam somente se o histograma for análogo a uma distribuição normal gaussiana, o que raramente se vê nas imagens de sensoriamento remoto.

Para descrever adequadamente a distribuição de frequências de uma variável quantitativa, é necessário analisar também o quanto esses valores variam (variabilidade), ou seja, quão dispersos eles estão em relação à média. Estatisticamente, a variabilidade é a medida da variância, a qual dá uma ideia do contraste da imagem. Se a curva histográfica está muito concentrada em um pequeno intervalo de valores digitais, a imagem mostra pouca variabilidade de brilho, isto é, pouca variação de radiância. Caso esteja bem dispersa, ocupando um intervalo grande de valores digitais, significa que os alvos presentes na cena têm radiâncias bem diferentes e a medida da variância é alta. Essas informações são automaticamente computadas durante o processamento da imagem e podem ser acessadas pelo usuário por meio de relatórios estatísticos dos dados das imagens.

A forma do histograma indica como os dados na imagem estão distribuídos em torno da média ou da moda. A forma geral é próxima de um sino, com um único pico, mas não necessariamente com as abas da esquerda e da direita simétricas. O mais comum é uma das extremidades da curva estender-se em direção a um dos limites inferior ou superior do intervalo de números digitais. Não se deve associar essa forma de histograma parecida com a de um sino com a estatística de uma distribuição gaussiana denominada função densidade de probabilidade que é definida para variáveis aleatórias contínuas. As imagens são raramente tratadas como exemplos de um processo de medidas de variáveis aleatórias. Por isso, o histograma é mais apropriado para descrever os dados discretos das imagens de sensoriamento remoto.

Os exemplos de histogramas mostrados na Figura 8.2 indicam que esses três parâmetros, moda, média e variância, classificam os padrões de

histogramas das bandas nos tipos unimodal, bimodal, simétrico e assimétrico, os quais podem ser abertos ou fechados. A forma simétrica é mostrada no caso A em que o histograma apresenta média baixa de valores digitais e distribuição fechada. Isso significa que os alvos na cena possuem, simultaneamente, baixa radiância e baixo contraste de níveis de cinza. A aparência visual da imagem é escura, sendo mais difícil de distinguir as diferentes classes de alvos. No caso B, a média dos valores digitais é mais alta, isto é, visualmente, a imagem é mais clara, com uma variância moderada, portanto, com um contraste espectral moderado de alvos. No caso C, o histograma tem uma forma bimodal (dois picos), sugerindo a presença de um conjunto de pixels com valores baixos de radiância, correspondendo a alvos com tons de cinza mais escuros, e de outro conjunto de pixels com tons de cinza mais claros. O caso D é de um histograma assimétrico que estende a curva em direção aos números digitais mais altos devido à presença de vários alvos com alta radiância. No caso E, a curva histográfica é simétrica, mas as baixas reflectâncias das águas dos rios configuram um pico com muitos pixels em torno do valor digital igual a zero.

Em termos de processamento de imagens, a importância de conhecer essas características do histograma é que elas ajudam o analista a selecionar o tipo mais apropriado de técnica de processamento de ampliação histográfica para obter uma imagem com melhor contraste. Como é frequente na cena apresentar bandas com diferentes formas histográficas, pode haver a necessidade de se aplicar uma técnica distinta de ampliação histográfica para cada banda.

Há duas classes de algoritmos de ampliação de histograma que são as mais empregadas: lineares, que podem ser por saturação e por partes, e as não lineares, que podem ser gaussiana, equalização, raiz quadrada, logarítmica e exponencial. As técnicas lineares conseguem manter a radiometria da imagem, enquanto as não lineares distorcem a radiometria. Existe uma máxima em processamento de imagens de que o bom desempenho dos resultados obtidos em uma área não necessariamente se repete em outras áreas.

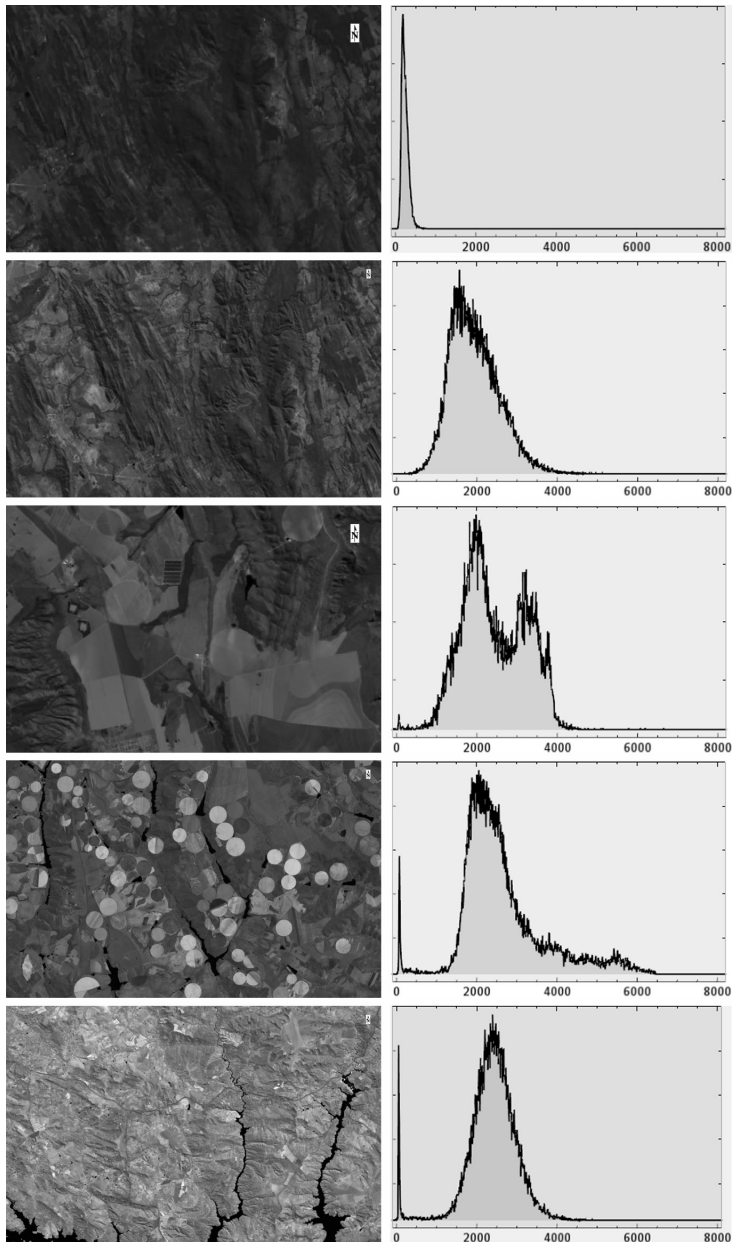


Figura 8.2. Exemplos de formas de histogramas.

8.2. Expansão por Saturação

Os sensores são instrumentos concebidos para registrar, em qualquer faixa espectral, uma ampla variação de intensidade de radiância refletida, emitida ou retroespalhada. Acontece que, na natureza, frequentemente uma boa parte dos alvos contidos em uma cena não apresentam altos contrastes de reflectância.

Na prática, isso mostra que, quando a imagem original é exposta no monitor de vídeo, o analista tem dificuldade para discriminar visualmente as classes de níveis de cinza. Contribui também para isso, a baixa capacidade de percepção visual do olho humano para discriminar tons de cinza. Como pode ser visto nos exemplos da Figura 8.2, geralmente apenas uma pequena proporção do intervalo máximo de valores de brilho de uma imagem (nível de quantização) é ocupada. A necessidade de se ter um intervalo digital grande é para que os alvos possam ser detectados sem levar à saturação dos detectores. O intervalo amplo de valores digitais permite que exista uma boa margem de espaço para expandir o histograma e, conseqüentemente, aumentar o contraste espectral dos alvos presentes na imagem.

A maneira de aumentar o contraste da imagem por saturação é reescalonar a amplitude de cada pixel por meio de uma regra básica de expansão do histograma. A imagem modificada ou realçada é assumida de ser restrita ao mesmo número de pixels da imagem original, apenas expandindo seus valores de brilho para o intervalo total de quantização da imagem. Por isso, o que se realiza é uma saturação em preto e em branco, de modo que o nível de brilho de saída é o nível que mais se aproxima do exato mapeamento de um nível de brilho de entrada. Esse efeito resulta em uma escala de níveis de brilho de saída muito mais diferenciada. Fica claro que, ao se efetuar a expansão linear na imagem, a mesma intensidade de expansão é aplicada igualmente para os pixels que tenham o mesmo valor de brilho, o que significa ser impossível de aplicar, em uma imagem, diferentes transformações histográficas para as áreas com os mesmos valores de brilho. Se acontecer de o realce produzir valores de brilho fora do intervalo definido

pela função de transformação, esses valores serão saturados ao nível máximo ou ao nível mínimo permitido. Por isso que é dito ser realce ou expansão linear por saturação.

A função de mapeamento linear por saturação é uma função de 1º grau que destina novos valores de brilho (y) para cada pixel da imagem original (x). Na função matemática de primeiro grau, o coeficiente a determina a inclinação da reta e exerce, efetivamente, o efeito de ampliação do histograma original, enquanto o coeficiente b define o *offset*, ou seja, o deslocamento do histograma em relação à origem. Para $a > 1$ e $b = 0$, tem-se somente uma ampliação de contraste na imagem de saída; para $b > 0$, acrescenta-se, à ampliação de contraste, um deslocamento dos níveis de cinza a fim de centralizar o histograma próximo à média do intervalo de valores digitais da imagem. A Figura 8.3 mostra que o valor de brilho de saída (V_s) pela transformação linear é calculado por simples relações trigonométricas, a partir dos valores do intervalo de valor digital da imagem de entrada (Equação 8.1).

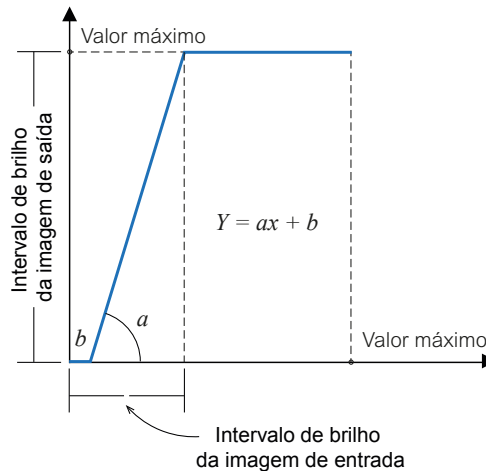


Figura 8.3. Representação trigonométrica da expansão linear de contraste por saturação dos valores de brilho.

$$y = ax + b \quad (8.1)$$

Computacionalmente, essa última equação pode ser expressa por (Equação 8.2):

$$V_s = \frac{255}{Max - Min} (VB_e - Min) \quad (8.2)$$

em que

V_s = valor do brilho de saída;

VB_e = valor do brilho de entrada;

Max = valor de brilho máximo (255 para imagens com 8 bits); e

Min = valor de brilho mínimo.

É importante considerar que o realce linear deve ser aplicado nas imagens que exibem histogramas com formas mais simétricas e unimodais. Para aplicar esse tipo de realce, o programa adquire o histograma de cada banda e determina os valores inferiores e superiores dos níveis de cinza ou brilho, que são fixados como valores de corte. Como o processo é iterativo, o analista pode, ao seu critério, redefinir os valores de corte a fim de controlar o realce desejado. A função de transformação linear é aplicada para cada banda, fazendo com que o valor de corte mínimo seja fixado automaticamente ao valor de brilho zero e o valor de corte máximo ao valor máximo digital. Qualquer valor dentro do intervalo é linearmente reescalonado entre VD_{min} e VD_{max} . Portanto, os pixels vizinhos aos valores min-max são mapeados para preto e branco, respectivamente. A técnica consiste, dessa forma, em um realce de contraste por saturação. Esse é um bom procedimento para manter a base de dados normalizada à escala total do intervalo dinâmico de brilho. Usualmente, o analista, ao observar o histograma, determina livremente quais são os valores digitais mínimo e máximo do dado

original, pois ele pode decidir que, abaixo ou acima de certo valor digital, não há informação de interesse.

Como exemplo do método, a Figura 8.4 representa o histograma de uma suposta banda de 8 bits, com os valores mínimos e máximos dos pixels da imagem de entrada fixados em 5 e 25, respectivamente.

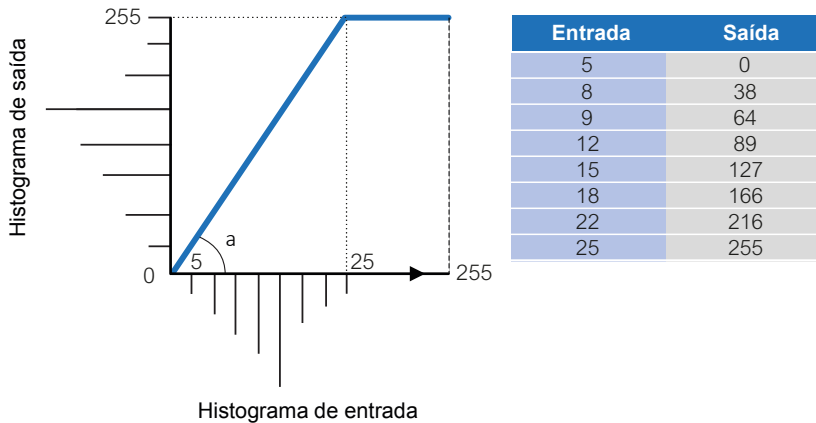


Figura 8.4. Exemplo numérico e gráfico da ampliação dos valores de brilho de uma imagem pelo método de expansão linear.

O problema pode ser assim formulado:

- Valores de brilho < 5 : fixados em 0 ($b = 0$);
- Valores de brilho > 25 : fixados em 255; e
- Valores de brilho entre 5 e 25: expandidos entre 0 e 255.

Portanto, segundo a função computacional de transferência de 1º grau,

- Para $f(x,y) < 5$, $y = 0$;
- Para $f(x,y) > 25$, $y = 255$; e
- Para $f(x,y) \in [5, 25]$, y é reescalado proporcionalmente segundo a Equação 8.3.

$$V_s = \frac{255}{255 - 5} (VB_e - 5) \quad (8.3)$$

Computacionalmente, os cálculos operacionais podem ser numericamente representados em uma tabela denominada de *Look-Up Table* (LUT), que relaciona os valores de brilho da imagem original aos correspondentes valores de brilho da imagem realçada. A LUT é feita mediante uma rotina computacional usando a equação de transformação selecionada para fazer o realce da imagem. A Figura 8.5 mostra o resultado obtido de uma imagem realçada por ampliação linear por saturação e seu respectivo histograma. As duas barras no histograma marcam, por opção do usuário, os valores de corte mínimo e máximo no espaço no histograma, onde todos os pixels com valores digitais compreendidos entre eles serão expandidos para o intervalo total dos valores digitais da imagem de saída.

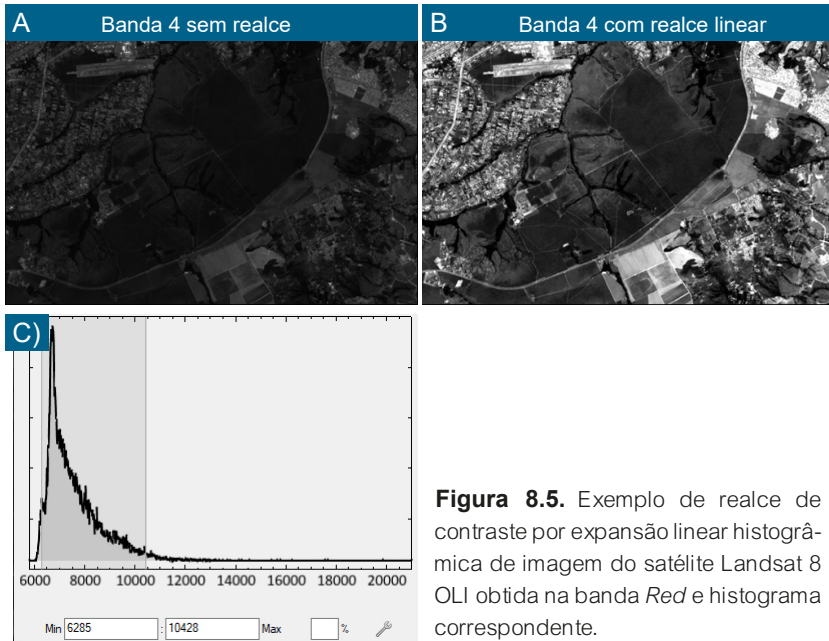


Figura 8.5. Exemplo de realce de contraste por expansão linear histográfica de imagem do satélite Landsat 8 OLI obtida na banda *Red* e histograma correspondente.

A grande vantagem da função linear, como se observa nas imagens da Figura 8.5, é a manutenção na imagem realçada das relações originais de reflectância dos alvos. Não há alteração radiométrica da imagem. Por isso, a opção por ampliação linear por saturação é a mais preferida das opções para aumentar o contraste espectral das imagens da maioria dos sensores multiespectrais, principalmente quando há interesse em estabelecer relações de análises comparativas da imagem com a reflectância dos alvos medida em campo ou laboratório.

8.3. Expansão por Equalização

Quando o histograma da imagem é assimétrico, é impossível, usando uma função linear, controlar simultaneamente o realce dos níveis de cinza com maior população de pixels e a quantidade de saturação nos extremos do histograma com menores populações. Para esse tipo de histograma, será necessário usar uma transformação não linear para executar uma redistribuição dos valores de brilho, tal que os incrementos de ampliação do brilho da imagem sejam desigualmente distribuídos entre VD_{min} e VD_{max} . Realces não lineares distorcem a radiometria da imagem e, portanto, são úteis apenas para melhorar a interpretação visual da imagem.

Equalização refere-se ao fato de o histograma da imagem realçada ser aproximadamente igual em densidade (número de pixels/níveis de cinza), ou seja, a curva histogrâmica tem uma altura aproximadamente igual, dando a aparência de um histograma com uma forma quase uniforme. Ao invés de ser uma modificação matemática presa à forma do histograma original, como condiciona a expansão linear, a equalização estabelece uma forma de histograma de saída previamente concebida. Imagina-se que se o histograma tem a mesma altura, então ele mostrará uma imagem onde há uma boa representação dos detalhes de todos os valores de brilho. Pode ser dito que a equalização de histograma é realizada com o objetivo de normalizar a escala de brilho, trazendo resultados muito satisfatórios de realce de imagens.

Na prática, verificamos que o resultado normalmente não exibe um histograma perfeitamente uniforme, tal como a técnica foi proposta. A equalização só conduz a histogramas efetivamente uniformes quando a gama de brilhos da imagem for contínua e se o número de pixels da imagem fosse infinito. Como consequência disso, a equalização tende a reduzir o contraste dos níveis de brilho muito claros e dos níveis de brilho muito escuros, enquanto expande os níveis medianos que se transformarão em valores de brilho com maiores contrastes. A hipótese de haver uma vantagem dessa técnica em relação à forma de contraste linear está no fato de que o estímulo de percepção visual do olho humano não é linear.

Richard (2013) fizeram a demonstração matemática da função de mapeamento dos valores de brilho para calcular o contraste por equalização, que pode ser resumida pela Equação 8.4. Por essa equação, o aumento de contraste por equalização tem, por base, a função de distribuição cumulativa da imagem. É exigido que o intervalo dos valores de brilho de Y seja de 0 a $(L - 1)$, em que L é o número de classes do alvo (Equação 8.4).

$$Y = \frac{L - 1}{N} C(x) \quad (8.4)$$

em que

Y = valor do brilho de saída;

N = número total de pixels da imagem de entrada;

$C(x)$ = histograma cumulativo; e

Os histogramas apresentados na Figura 8.6 correspondem à base de dados da Tabela 8.1. São dados de uma imagem hipotética com um total de 75 pixels, distribuídos em 16 classes de valores de brilho, apenas para simplificar os cálculos do exemplo que é apresentado. Os histogramas das Figuras 8.6a, 8.6b e 8.6c são, respectivamente, da imagem original, dos valores cumulativos e da imagem realçada de saída. Na Tabela 8.1, a pri-

meira coluna (i) são os valores das 16 classes de brilho da imagem original com 75 pixels. Computado o valor cumulativo (iii) dos números de pixels da imagem de entrada (ii), primeiramente, calcula-se o valor do novo fator de escala (iv) que, pela Equação 8.4, corresponde à divisão $(L - 1)/N = 15/75 = 0,2$. Esses valores são aproximados para valores de brilho inteiros e discretos (v).

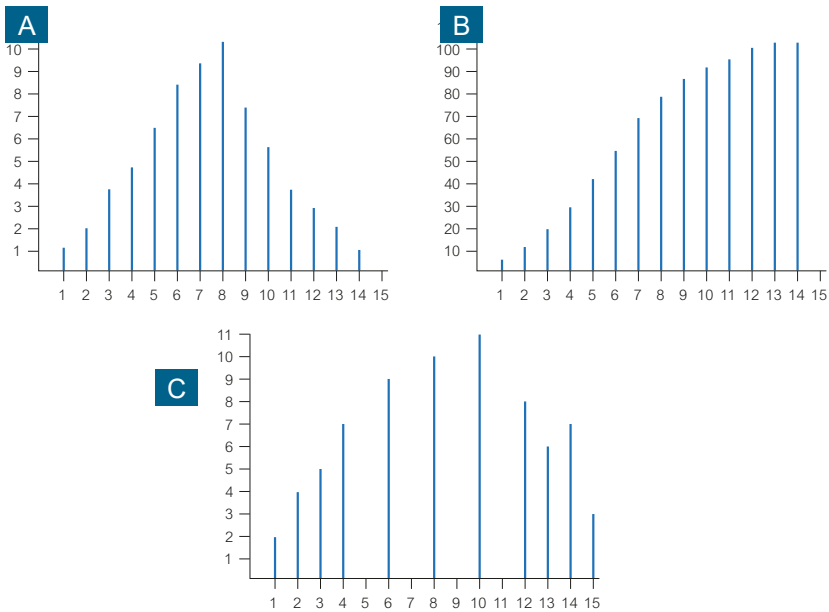


Figura 8.6. Histogramas de uma imagem original hipotética (A), dos valores acumulativos (B) e da imagem realçada por expansão da equalização (C).

Tabela 8.1. Dados do processo acumulativo de pixels para cálculo da equalização do histograma.

Brilho de entrada	Pixels por classe	Número de pixel cumulativo	Novo valor da escala	Aproximação do valor de brilho	Número de pixel da imagem realçada
0	1	1	0,2	0	2
1	1	2	0,4	0	2
2	2	4	0,8	1	4
3	4	8	1,6	2	5
4	5	13	2,6	3	7
5	7	20	4,0	4	0
6	9	29	5,8	6	9
7	10	39	7,8	8	0
8	11	50	10,0	10	10
9	8	58	11,6	12	0
10	6	64	12,8	13	11
11	4	68	13,6	14	0
12	3	71	14,2	14	8
13	2	74	14,8	15	6
14	1	75	15,0	15	7
15	0	75	15,0	15	3

Por último, para calcular a nova distribuição de número de pixels da imagem de saída, ou imagem realçada, relacionam-se os valores de brilho da coluna (v) com os números de pixels da coluna (ii). Aos valores de brilho que não há na coluna (v), o número de pixels é computado como sendo zero. Observe, no histograma de saída da Figura 8.6c, que o maior contraste ocorre para as classes que ocupam a posição mais central do histograma. Portanto, o efeito da equalização aumenta o contraste no centro do histograma que é onde há a mais densa população de pixels, enquanto reduz o contraste nas extremidades. Embora o pressuposto da equalização seja gerar uma forma histogrâmica com as classes (barras) com igual dimensão

de número de pixel, isso não é totalmente conseguido, como se observa no histograma equalizado da Figura 8.6c.

A equalização tem uma aplicação prática além do simples propósito de aumentar o contraste da imagem. Nas aplicações de sensoriamento remoto, é bastante comum haver a necessidade de se compor mosaicos com duas ou mais imagens. Nesse processo, as imagens frequentemente têm, entre si, uma distribuição de brilho variada, ou por serem de datas diferentes ou por apresentarem variações das condições atmosféricas de uma imagem para outra. Nesse caso, para se ter um mosaico com aparência de brilho mais homogênea, é preciso realizar uma transformação nas imagens a fim de reduzir as variações de brilho do mosaico. Essa tarefa se resume em equalizar os brilhos das imagens por um processo igual ao apresentado, com a diferença de que o usuário deverá decidir pela escolha de uma imagem de referência, em relação à qual as outras imagens serão equalizadas. Os programas de processamento de imagens executam essa operação de uma forma bastante automatizada, sem que o usuário tenha a necessidade de analisar os histogramas.

A função de equalização não deve ser aplicada em imagens que tenham uma porção da área extensa e homogênea de brilho constante, como um grande corpo de água, que cria um pico isolado no histograma. Isso porque a equalização tem, como princípio, criar um histograma uniforme na altura, agrupando classes de alvos menores. Um pico isolado de uma única classe de alvo interfere no histograma cumulativo, tendo, como consequência, a perda de contraste, o que é o oposto do pretendido.

Na Figura 8.7, é mostrada uma imagem da banda do infravermelho de ondas curtas (SWIR1) do satélite Landsat 8 *Operational Land Imager* (OLI), original e realçada por equalização. Observe que, na imagem realçada, é mais notável a textura de relevo, enquanto nas áreas de solo exposto, os valores dos pixels, originalmente mais claros, ficam saturados em branco.

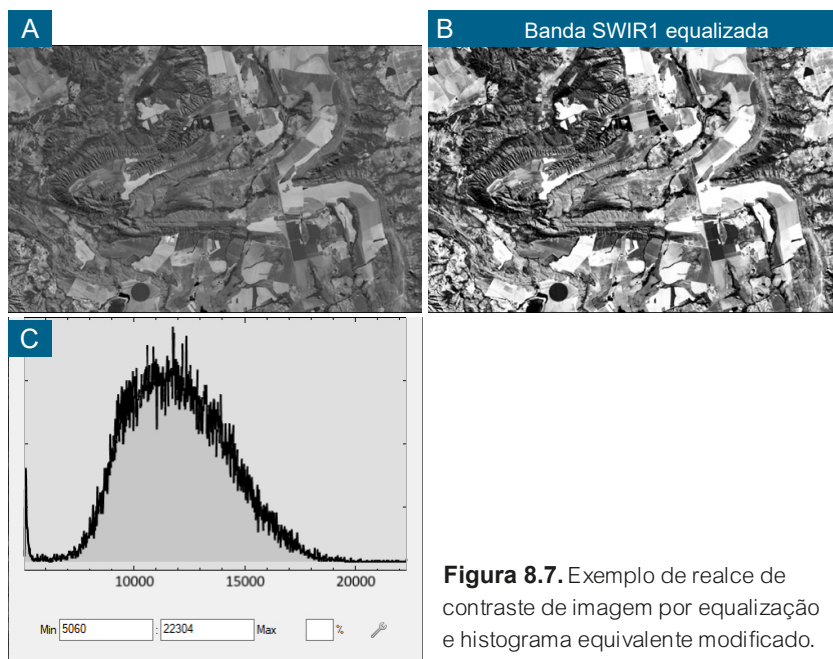


Figura 8.7. Exemplo de realce de contraste de imagem por equalização e histograma equivalente modificado.

8.4. Expansão Gaussiana

A proposta de expansão de histograma baseada em suposições gaussianas fundamenta-se nas formas dos histogramas que frequentemente são observadas nas imagens das diferentes bandas espectrais de sensoriamento remoto. Como temos observado nos exemplos anteriores, há uma tendência de os histogramas apresentarem formas parecidas à de uma distribuição normal gaussiana ou log normal. Portanto, seria lógico imaginar que, para algumas finalidades, poder-se-ia ajustar o histograma de uma imagem a uma forma predefinida de um histograma com forma gaussiana.

Tal qual o processo de equalização, que também força o ajuste do histograma da imagem original a uma forma histogrâmica predefinida, o processo passa por um estágio de usar o histograma cumulativo da imagem original para se obter novos valores de distribuição de pixels da imagem de saída. Para isso, são exigidos dois parâmetros para realizar um processa-

mento estatístico de segunda ordem: a média e o desvio-padrão. Segundo Schowengerdt (2007), ambos os parâmetros fornecem, ao algoritmo, meios convenientes para explicitamente controlar a média dos resultados, ao mesmo tempo em que varia o contraste controlando o desvio-padrão. O aumento do desvio-padrão permite expandir a distribuição dos dados com aumento de contraste. A modificação gaussiana é um processo de normalização particularmente útil para igualar as imagens em termos de média e desvio-padrão, o que ajuda a obter composições coloridas com bom equilíbrio de cores.

A função gaussiana de densidade de probabilidade normal e sua forma analítica são mostradas na Figura 8.8. Os dois parâmetros que definem a distribuição são a média μ e o desvio-padrão σ (Equação 8.5). A área total sob a curva é unitária e a área dentro do intervalo $(+2\sigma$ e $-2\sigma)$, centrado na média, é de 0,95, e a área dentro do intervalo $(+1\sigma$ e $-1\sigma)$ é de 0,68.

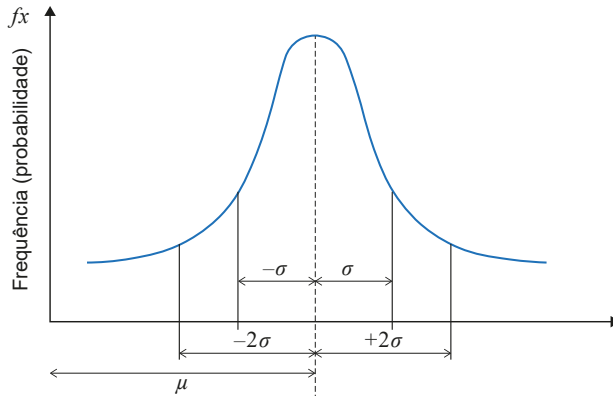


Figura 8.8. Função de densidade de probabilidade normal ou curva gaussiana.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (8.5)$$

em que

μ = média;

σ = desvio-padrão;

$\pi = 3,14159$; e

$e = 2,71828$

Na Figura 8.9, é mostrado um exemplo de uma imagem realçada por ampliação histográfica gaussiana. Como a curva de uma distribuição normal varia de $-\infty$ a $+\infty$, no exemplo, um intervalo de ± 3 desvios-padrões foi usado. Um maior ou menor número de desvios-padrões pode ser especificado pelo usuário. O algoritmo define a priori qual é a probabilidade de cada classe de brilho da imagem original ser mapeada para novos valores de brilho, dentro de uma distribuição normal. Note que, no histograma transformado gaussiano, os extremos, isto é, as áreas muito escuras ou muito claras, foram também realçados.

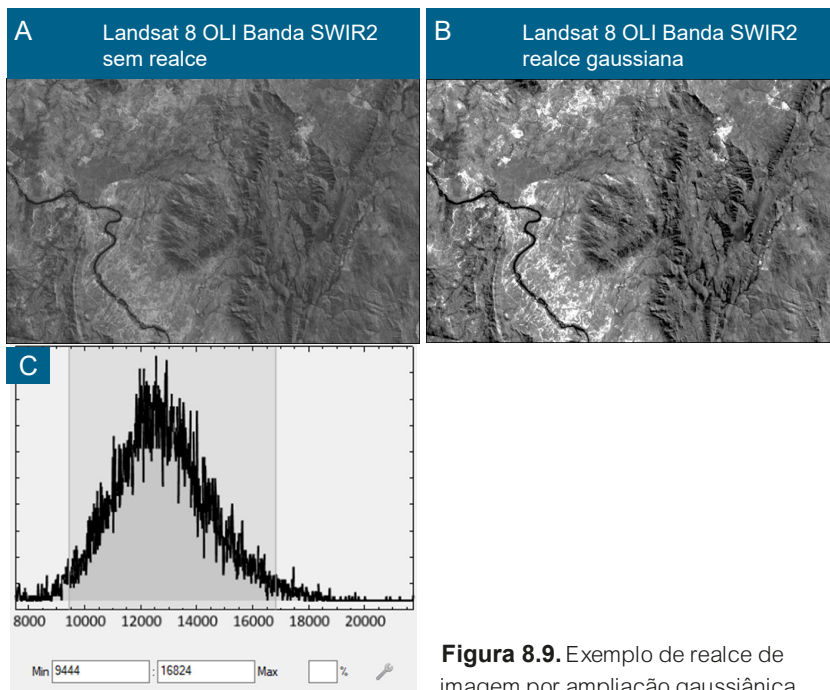


Figura 8.9. Exemplo de realce de imagem por ampliação gaussiana.

8.5. Função Raiz Quadrada

Quando algumas imagens apresentam uma distribuição de dados concentrada na porção de valores de brilho mais escuros do que nas porções mais claras, gerando um histograma de forma assimétrica para a esquerda, é desejável forçar uma maior intensidade de contraste dessa porção mais escura da imagem, e menor da porção mais clara. A função de transformação não linear adequada a esse tipo de histograma é a raiz quadrada, que tem a seguinte formulação matemática (Equação 8.6):

$$g(l, p) = G \times \sqrt{g(l, p)} \quad (8.6)$$

em que

$$G = \frac{ND_{max}}{\sqrt{ND_{max}}} \text{ (fator de ajuste); e}$$

$g(l, p)$ = valor original do pixel.

Nesse caso, propicia-se um realce maior das áreas de intensidade média e clara da imagem (Figura 8.10). A inclinação da curva de transformação será mais acentuada no final do que no início do histograma. Observe que, nas imagens correspondentes aos histogramas de entrada e de saída, as áreas mais escuras na imagem original modificaram-se para valores de níveis de cinza mais claros, visualmente mais discerníveis, aumentando a interpretabilidade da imagem. Equivalentes às funções raiz quadrada e quadrática são as funções logarítmica e exponencial, com a diferença de que promovem o realce de um intervalo menor de níveis de brilho.

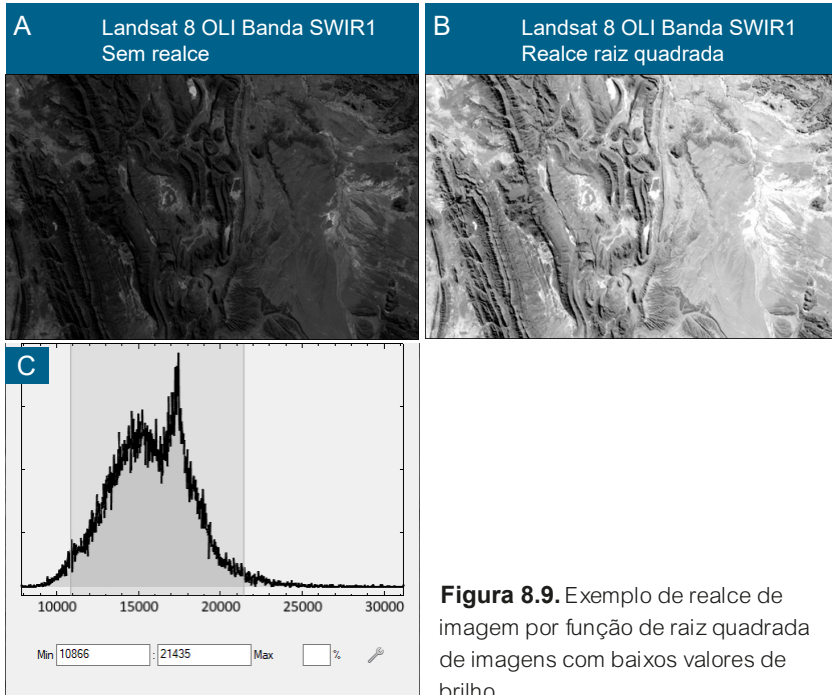


Figura 8.9. Exemplo de realce de imagem por função de raiz quadrada de imagens com baixos valores de brilho.

CAPÍTULO 9

Sistemas de Cores

Paulo Roberto Meneses

Tati de Almeida



Qualquer imagem, quando exposta em bandas individuais no monitor, é uma imagem em preto e branco, pela simples razão de que a onda eletromagnética não tem cor. Uma onda só é medida pelo seu comprimento e sua energia. Cor é um atributo que adicionamos às imagens por meio de técnicas de processamento, visando aproximá-las à capacidade que o olho humano possui em enxergar cores. A cor pode ser usada para simplesmente expor uma imagem no monitor ou pode ser manipulada por técnicas de processamento a fim de realçar os dados que o analista deseja extrair da imagem. Fazemos isso porque nós, humanos, temos a capacidade de distinguir milhares de tons de cores, enquanto nos restringimos a perceber não mais do que duas dezenas de tons de cinza. Neste capítulo, apresentaremos algumas técnicas heurísticas simples para manipular numericamente a cor, utilizando-se de figuras geométricas como cubos para entender os modelos teóricos de cor cujas teorias são complexas demais para serem abordadas aqui

9.1. Cor

O termo cor é usado para se referir a um aspecto da experiência visual de percepção. Para descrevê-la, usamos os atributos tom ou cor dominante, brilho ou luminosidade e saturação para distinguir uma cor opaca de uma cor intensa. Há também fatores físicos que devem ser considerados e que afetam esses atributos, como a iluminação da cena, as propriedades das superfícies dos objetos e até a atitude do observador. Isso torna difícil a obtenção de medidas quantitativas precisas da magnitude de percepção da cor. Matematicamente, para dimensionar uma cor, é preciso numerar os vetores que são necessários para caracterizar diretamente a percepção de cor. Essa formulação matemática precisa estar de acordo com as teorias que tentam explicar como o olho humano enxerga as cores, as quais dependem de conceitos ópticos, físicos e psicofísicos para explicar a resposta interna da visão aos estímulos externos de cores. Desde a primeira teoria

de cor, proposta por James Maxwell no século XIX, a preocupação tem sido como padronizar as cores.

Experimentos realizados a partir de misturas de cores mostraram que é necessária uma combinação aditiva de três estímulos independentes de diferentes energias espectrais para produzir a percepção de todas as cores pelo olho humano. Essa mistura aditiva de três cores primárias é possível de ser representada graficamente, numericamente ou em termos de notação vetorial. A vantagem de modelar quantidade de qualquer cor por meio de uma forma geométrica está na possibilidade de medi-la por meio de coordenadas ou valores numéricos.

Em 1931, a *Commission Internationale de l'Eclairage* (CIE) propôs um padrão de cores que foi adotado internacionalmente. Trata-se de um padrão de cores de origem luminosa, isto é, a luz. O CIE adotou o sistema de especificações triestímulos como padrão internacional para colorimetria, desenvolvendo fórmulas para calcular os valores triestímulos do olho humano e as coordenadas de cromaticidade por meio de padrões de luzes das cores primárias azul (B), verde (G) e vermelha (R) e um grupo de observadores. Com isso, criou-se o diagrama de cromaticidade, visto na Figura 9.1. Foram estabelecidos os valores de comprimentos de onda de 435,8 nm, 546,1 nm e 700 nm para representar espectralmente as três cores primárias azul, verde e vermelha, respectivamente. O diagrama de cromaticidade mostra os limites de todas as cores visíveis que representam as combinações de cores monocromáticas do espectro.

A medição de cores pelo sistema CIE se baseia na possibilidade de matizar qualquer cor como uma combinação das três cores primárias R, G e B, conhecida como sistema de cor RGB. A soma dessas cores primárias resulta no branco. A cor preta corresponde à ausência total de luz. Após o desenvolvimento do sistema CIE, triângulos de cores foram usados como diagramas de cromaticidade, sendo que as coordenadas lineares do triângulo representam os valores de cromaticidade. Como a soma dos três valores de cromaticidade tem um valor fixo, é suficiente representar geometricamente apenas dois dos três valores por meio de coordenadas cartesianas X, Y. Os

vértices do triângulo de cromaticidade representam as cores primárias RGB (coordenadas X, Y e Z nos vértices (1,0), (0,1) e (0,0), respectivamente), enquanto os lados do triângulo representam as cores secundárias ciano, amarelo e magenta, formadas pela mistura das cores primárias. RGB é o sistema de cor usado na captação e reprodução de cores em monitores de computadores, televisões e máquinas fotográficas digitais.

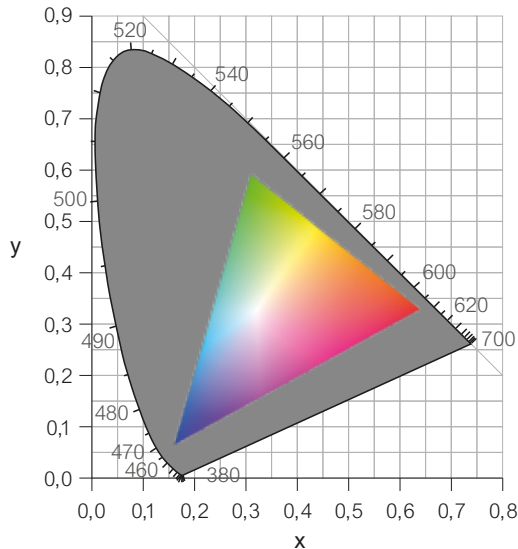


Figura 9.1. Diagrama de cromaticidade mostrando os limites de todas as cores visíveis que representam as combinações de cores monocromáticas do espectro.

Essa teoria triestímulo que modela numericamente a percepção visual de cores pelo sistema RGB é que permitiu a fundamentação das formulações computacionais para usar a cor como forma de expor as imagens multiespectrais de sensoriamento remoto. Como será visto neste capítulo, os algoritmos de processamento de cores usualmente utilizados em processamento de imagens de sensoriamento remoto são numericamente modelados por vetores que compõem os três eixos ortogonais de um cubo representando as cores tricromáticas RGB. Também será mostrada outra

opção que é usar o espaço hexacone para representar a cor por meio dos atributos de brilho (intensidade), matiz (*hue*) e saturação (S). Isso significa a aplicação de dois modelos de cores para construir composições coloridas com imagens de sensoriamento remoto. A representação desses dois modelos RGB e IHS pelos espaços do cubo e do hexacone foi a forma encontrada para uma simplificação computacional confiável e ajustada aos equipamentos de reprodução RGB dos monitores.

Qual é a importância da cor para a visualização das imagens multiespectrais de sensoriamento remoto? Se considerarmos que fisiologicamente a visão humana tem uma baixa percepção para discriminar níveis de cinza, essa seria a resposta a essa questão. Na prática, os intérpretes de fotos aéreas preto e branco quase sempre desprezam a análise dos tons de cinza e recorrem à visão estereoscópica 3D para interpretar as fotos. Teríamos a mesma dificuldade se tentássemos analisar as imagens em preto e branco de uma determinada banda de um sensor. Mesmo que os métodos de ampliação de contraste de imagens, vistos em capítulo anterior, ajudem a melhorar essa deficiência, dificilmente poder-se-ia garantir que todas as informações contidas nas bandas fossem reconhecidas pelo fotointérprete. Por essa razão, quando se modifica o contraste das imagens originais, os resultados poderão ser mais bem avaliados quando as bandas realçadas são combinadas para formar composições coloridas. As composições coloridas procuram reproduzir padrões de cores tricromáticas RGB, combinando três bandas com as cores primárias azul, verde e vermelha.

A intenção de usar a cor pode ser apenas para expor visualmente as imagens para facilitar algumas operações simples como a seleção de amostras de áreas para uma classificação supervisionada. Mas o uso mais importante é propiciar uma alta qualidade de extração de informações para fins de mapeamentos temáticos, cartográficos ou de monitoramento. Veremos que não são apenas as imagens realçadas por modificação de histogramas que são usadas para compor uma imagem colorida. Há diversas outras técnicas como decorrelação de bandas, fusão de imagens com mais de três bandas ou fusão de imagens com diferentes resoluções espaciais.

9.2. Fatiamento de Cores

O procedimento mais simples para associar cores às bandas de uma imagem é por meio do fatiamento dos valores de brilho. Alguns pesquisadores consideram esse processo como uma forma de realce de contraste, muito embora não seja realizada nenhuma operação matemática ou estatística para processar a imagem. O fatiamento consiste em dividir o intervalo digital mínimo-máximo de uma determinada banda em subintervalos digitais especificados pelo analista. O analista deve recorrer ao histograma de valores digitais da banda para definir o número de fatiamento que deve aplicar na imagem. Os pixels que estiverem dentro de um intervalo digital são convertidos para um único valor e visualmente mapeados a uma cor escolhida. Se o analista decidir dividir a imagem em dez fatias, a imagem de saída terá dez variações de cores que são selecionadas previamente de uma paleta de cores. As cores dessa banda são denominadas de pseudo-cores para diferenciar essa imagem colorida de uma composição colorida padrão RGB. Trata-se de uma forma simples de colorir uma imagem composta por tons de cinza. Porém, ao agrupar pixels em um intervalo de brilho, o número total de valores discretos de brilho é reduzido, acarretando em perda de alguns detalhes da imagem.

O valor dessa técnica é transformar, em cores, as sutis variações de níveis de cinza que não são visualmente percebidas pelo analista e que podem representar pequenas diferenças espectrais de um alvo aparentemente homogêneo. O resultado é bastante parecido com uma classificação do tipo paralelepípedo, assunto a ser visto mais adiante. A Figura 9.2 mostra a aplicação de fatiamento de uma banda do infravermelho próximo de uma área de depósito de sal de bórax (borato de sódio) e halita, encontrado em proporções variadas nos Andes argentino. Dificilmente dois intérpretes chegariam a um mesmo resultado de mapeamento baseado apenas na interpretação visual das variações dos níveis de cinza que essa imagem apresenta. Bons resultados também podem ser obtidos para o mapeamento de

plumas de dispersão de sedimentos em corpos d'água, plumas de poluição ou de macrófitas.

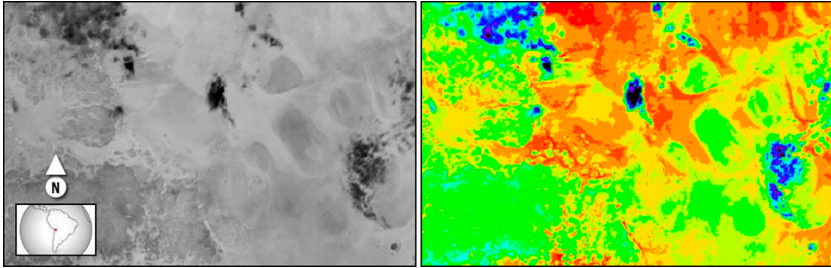


Figura 9.2. Banda do infravermelho próximo de um salar andino e pseudocores obtidas pelo fatiamento dos níveis de cinza da imagem.

9.3. Espaço de Cores RGB

O sistema de cor RGB é o mais usado entre os modelos de cores, especialmente para imagens de oito bits. De acordo com a teoria de mistura de cores, a luz visível, quando passa através da pupila, é focada na retina pelas lentes oculares. Na retina, há um grande número de células fotoreceptoras altamente especializadas, conhecidas como cones e bastonetes que estão conectados ao nervo óptico por uma rede de fibras nervosas, ao longo das quais a energia luminosa detectada pelo olho vai ao cérebro, formando a imagem visual. Há cerca de 5 milhões de cones e 120 milhões de bastonetes. Os bastonetes respondem à luz de baixos níveis de iluminação (intensidade), adaptando-se à sub-obscuridade. Os cones são responsáveis pela visão das cores e seus detalhes. Realizam a distinção de matiz (*hue*) e a percepção de pureza (saturação).

Imagina-se que a cor, ou mais apropriadamente *hue*, esteja associada com a visão diurna dos cones, pois há três tipos de cones, cada um sendo responsável pela criação de cada uma das três cores primárias, azul, verde e vermelha. O mais interessante dessa teoria triestímulo, além de sua simplicidade, está no fato de que ela dá a ideia de que as cores podem

ser formadas pela soma do vermelho, azul e verde em várias combinações de intensidades. Dessa forma, se imaginarmos que cada cor possui uma intensidade bem definida, ao se tomar uma fração da intensidade da cor e somá-la à outra, cria-se uma cor (C) que pode ser expressa por meio da conhecida equação de adição de cores (Equação 9.1).

$$C = x_1B + x_2G + x_3R \quad (9.1)$$

No algoritmo do modelo de espaço de cor das imagens de sensoria-mento remoto, as intensidades x_1 , x_2 e x_3 são determinadas pelos valores digitais de brilho das bandas. Em obediência à teoria triestímulo, o modelo concebido baseia-se na adição de um tripleto de bandas, às quais são as-sociadas livremente aos canais das cores azul, verde e vermelha do moni-tor. O processo de expor, no monitor, as imagens de três bandas, com seus respectivos valores digitais armazenados e organizados matricialmente, é mostrado na Figura 9.3.

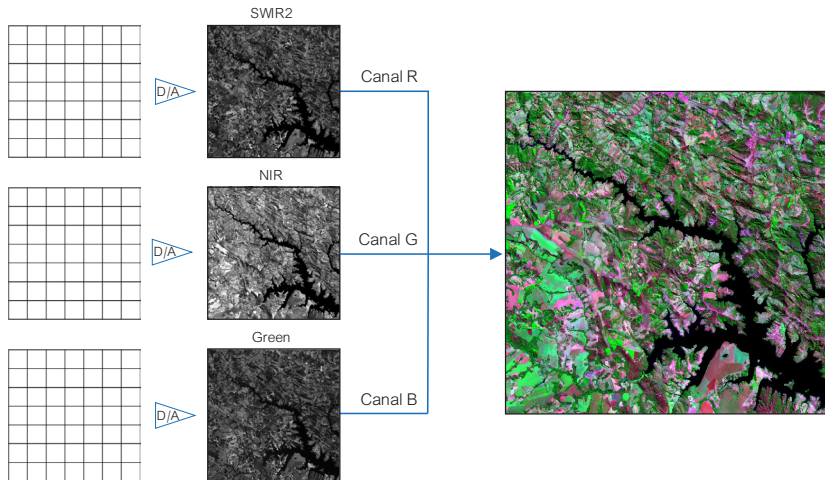


Figura 9.3. Conversão de uma imagem digital para analógica no processo de formação de uma composição colorida padrão RGB.

Para serem expostas nos três canais de cores vermelha, verde e azul do monitor, as imagens das três bandas são processadas por *Look-Up Tables* (LUT), que é essencialmente um arquivo digital que, com suas fórmulas matemáticas, consiste em converter os números digitais (nd) inteiros da imagem para valores analógicos discretos de níveis de cinza (NC), isto é, um conversor analógico/digital, conforme mostra a Equação 9.2.

$$NC = LUT(nd)_{0,255} \quad (9.2)$$

Para compor a imagem colorida na tela do monitor, os correspondentes pixels de cada uma das bandas são endereçados a cada um dos três canais de cores RGB do monitor. Assim, cada um dos pixels da tela RGB do monitor contém a sobreposição de três pixels das três bandas e a soma das cores RGB dos pixels cria a imagem colorida. As telas atuais conseguem gerar 24 bits de informação de cor por pixel, sendo 8 bits para cada canal RGB ($3 \times 8 = 24$). Esses bits são referentes aos índices de cores de profundidade, um termo da computação gráfica em que 8 bits (2^8) é igual a 256 cores. Se cada canal do monitor consegue gerar uma variação de 256 cores, três canais totalizam 16,78 milhões de variações de cores com brilho, tom e saturação. Isso explica o limite de valores que os canais RGB dos monitores utilizam, que vai de 0 a 255. Todas as imagens de sensoriamento remoto, quando expostas na tela do monitor, obedecem a essa condição, independente da resolução radiométrica da imagem ser de 8, 16 ou 32 bits.

O sistema de cores RGB tem, como qualidade destacada, a liberdade dada ao analista para explorar as possíveis combinações de três bandas por ele selecionadas, procurando obter a imagem colorida de melhor contraste. Embora o apelo visual das cores seja dependente da qualidade de contraste de cada banda, ele é controlado pela escolha adequada das bandas em função do comportamento espectral dos alvos que o usuário deseja analisar. Isso requer experiência e conhecimento por parte do intérprete sobre o comportamento espectral dos alvos. Não basta apenas ter habilidade em processamento de imagem.

A cor é um recurso imperativo no uso de imagens multiespectrais da faixa óptica do espectro eletromagnético, porém, composições coloridas podem ser obtidas com qualquer outro tipo de imagem, tais como as imagens termais multibandas do sensor *Advanced Spaceborne Thermal Emission and Reflection Radiometer* (ASTER) ou as imagens polarimétricas de radar do satélite canadense RADARSAT.

A regra universal de adição das três cores primárias é mostrada na Figura 9.4. A combinação da projeção dos focos de luzes azul, verde e vermelha resulta no branco, e a combinação pareada das cores azul + verde resulta na cor ciano. Da mesma maneira, o par de cores azul + vermelha resulta na cor magenta e o par de cores verde + vermelha resulta na cor amarela. Milhões de cores poderão ser obtidas variando as intensidades de luz de qualquer uma das três cores primárias.

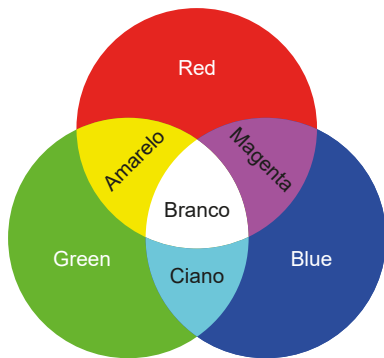


Figura 9.4. Combinação aditiva das cores RGB, gerando as cores secundárias ciano, magenta e amarelo.

Em termos instrumentais, os monitores coloridos são constituídos de materiais fotossensíveis às três cores primárias que se iluminam como um pequeno ponto de luz colorido (pixel da tela). A iluminação da tela é produzida por correntes de elétrons cujas intensidades são determinadas pelos valores de brilho dos pixels da imagem. Quando uma imagem é carregada na tela, um pixel da imagem ocupará um correspondente pixel do monitor. Dessa forma, cada pixel da tela registrará a combinação das três cores,

com intensidades correspondentes aos valores de 8 bits dos pixels das três bandas combinadas. Suponha um pixel de uma imagem com os seguintes valores:

- Banda 1: pixel com valor digital 255 e endereçado ao canal de vídeo R;
- Banda 2: pixel com valor digital 255 e endereçado ao canal de vídeo G; e
- Banda 3: pixel com valor digital 0 e endereçado ao canal de vídeo B.

O valor digital 255 acenderá, na tela do monitor, altos valores de brilho de cores vermelha e verde para as bandas 1 e 2, respectivamente. O valor digital 0 não acenderá nenhuma intensidade de azul no monitor. A soma das intensidades de brilho das cores R intenso + G intenso + B nulo, resulta em um pixel na tela de cor amarela (Y). A Figura 9.5 contém alguns exemplos de cores resultantes das combinações dos valores digitais com as cores básicas RGB. Vemos que a cor dominante é controlada pelo pixel ou pixels que possuem maior valor digital, ou seja, maior radiância. Como a *hardware* de monitores são padronizados para o sistema RGB, para qualquer outra representação de cor como a do sistema IHS, a imagem, após o seu processamento, tem de ser transformada obrigatoriamente para o sistema RGB para que possa ser exposta em monitores coloridos.

Cores	Valores digitais
	192 80 77
	84 141 212
	0 255 0
	255 0 255
	102 204 255
	215 246 20
	255 0 102
	128 128 128
	11 131 62

Figura 9.5. Exemplo de notações de cores RGB correspondentes aos valores digitais de três bandas.

O modelo computacional de mistura das cores RGB é representado pelo cubo de cor, como mostra a Figura 9.6. O modelo do cubo é intuitivamente

relacionado com o modelo triestímulo da visão humana e com o mecanismo de cores dos monitores. Os valores digitais dos pixels de cada banda irão constituir, no sistema de coordenadas RGB, os eixos ortogonais do cubo de cor. O intervalo de valor de brilho da imagem define o tamanho dos três eixos do cubo de cores RGB que, para imagens quantizadas em 8 bits, é de 256 (0–255). Nesse contexto, cada cor é identificada por uma tripla coordenada de números inteiros $0 \leq R \leq 255$, $0 \leq G \leq 255$ e $0 \leq B \leq 255$, representada por um vetor que estará de algum modo dentro dos limites do cubo de cor. Note, na Figura 9.6, que as cores aditivas primárias e secundárias estão nos vértices do cubo, enquanto o eixo diagonal com origem no vértice do centro do cubo e em direção ao vértice oposto às cores, contém os tons de cinza. Sobre as arestas do cubo, localizam-se as cores saturadas nesse espaço. Cor saturada não significa cor pura, como as cores definidas em 1931 pela CIE, onde cores puras são somente aquelas localizadas no contorno do diagrama de tricromacidade. Dessa forma, o uso do cubo de cor não é capaz de reproduzir cores puras, apenas cores saturadas.

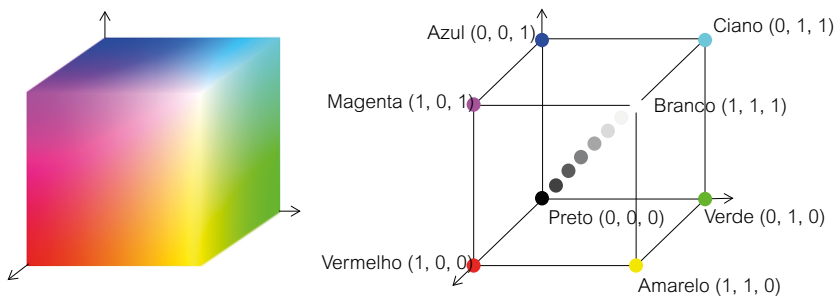


Figura 9.6. Representação do espaço de cubo de cores RGB para geração de imagens coloridas.

Esse modelo computacional de mistura de cores RGB embute um problema. Para que a imagem colorida apresente um alto contraste de cores, cada banda deve ser previamente realçada aplicando uma ampliação de histograma de acordo com os seus valores mínimo e máximo de números

digitais, o que é realizado automaticamente quando a imagem colorida é exposta na tela do monitor. Isso realmente aumenta o contraste das cores da imagem colorida RGB, mas o balanço das cores pode mudar de forma imprevisível porque qualquer uma das bandas pode conter pixels de valores anômalos, aos quais a referida ampliação de contraste é muito sensível.

A melhor combinação de cores para um tripleto de bandas é uma decisão a ser tomada pelo analista. Isso pode parecer simples, mas se considerarmos o número de possíveis combinações coloridas, como no cálculo abaixo, com seis bandas permutadas com três cores, a tarefa é quase impossível.

$${}_nP_r = \frac{n!}{(n-r)!} \quad \text{com} \quad {}_6P_3 = \frac{6!}{(6-3)!} = 120 \text{ combinações possíveis}$$

Para resolver esse problema, dois passos podem ser seguidos. O primeiro passo consiste em selecionar três bandas que resultem na melhor imagem colorida, isto é, com melhor contraste de cores entre os alvos presentes na cena. É preciso que o analista tenha um mínimo de informação de quais alvos existem na área e aqueles que são do seu interesse. A seguir, ele deve plotar, em um gráfico, as curvas de reflectâncias dos alvos que deseja discriminar na imagem e nele indicar as posições das bandas. As melhores bandas são aquelas situadas nas regiões de maior diferença de reflectância entre os alvos do seu interesse, como é mostrado na Figura 9.7 com as bandas do sensor *Enhanced Thematic Mapper Plus* (ETM+) do satélite Landsat 7. Nesse exemplo, a melhor composição colorida é obtida com as bandas 4, 5 e 7. Pelo fato de os comprimentos de onda dessas bandas não serem do visível, a composição colorida por elas formada é chamada de falsa-cor. Para se ter uma composição colorida que se aproxime da visão humana, deve ser usado um tripleto de bandas com comprimentos de onda situados exclusivamente no visível. No exemplo mostrado na Figura 9.7, essa composição seria com as bandas 1, 2 e 3 do satélite Landsat 7.

Uma vez selecionado o tripleto de bandas, o segundo passo é decidir qual cor RGB atribuir às três bandas selecionadas. Experimentos mostram

que o número de cones sensíveis à cor azul é muito menor do que às cores vermelha e verde, por isso a sensibilidade da visão humana à cor azul é mais baixa do que às cores verde e vermelha, como mostra a Figura 9.8. Em função disso, para montar uma composição colorida, é aconselhável associar a cor verde ou a cor vermelha às bandas de maiores variâncias, portanto, de maior contraste, destinando a cor azul à banda de menor variância. Para saber a variância da banda, basta ver o histograma de valores digitais. Pode-se constatar isso nas composições coloridas exibidas na Figura 9.9. São imagens obtidas pela combinação das bandas R, infravermelho próximo (NIR) e infravermelho de ondas curtas (SWIR1) do sensor *Operational Land Imager* (OLI) do satélite Landsat 8. Dessas três bandas, a banda SWIR1 é a que apresenta a maior variância. Observe que, quando as cores verde ou vermelha são destinadas a essa banda, as composições apresentam maior impacto visual do que quando se associa a cor azul a essa banda, muito embora todas as três imagens coloridas contenham o mesmo conteúdo de informação.

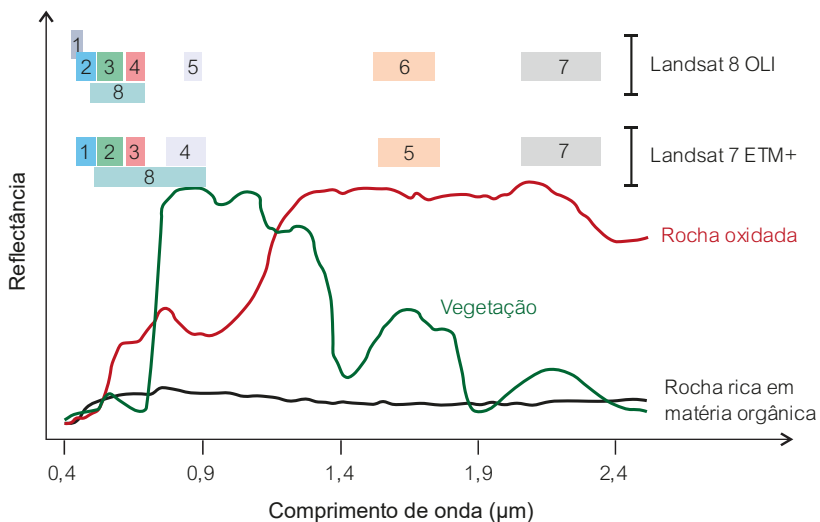


Figura 9.7. Espectros de reflectância de quatro materiais e as posições das bandas do sensor ETM+ como meio prático para selecionar três bandas que melhor possam identificar esses materiais.

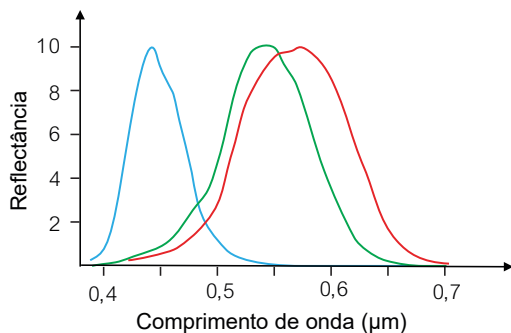


Figura 9.8. Diferenças de sensibilidade do olho humano às cores azul, verde e vermelha.

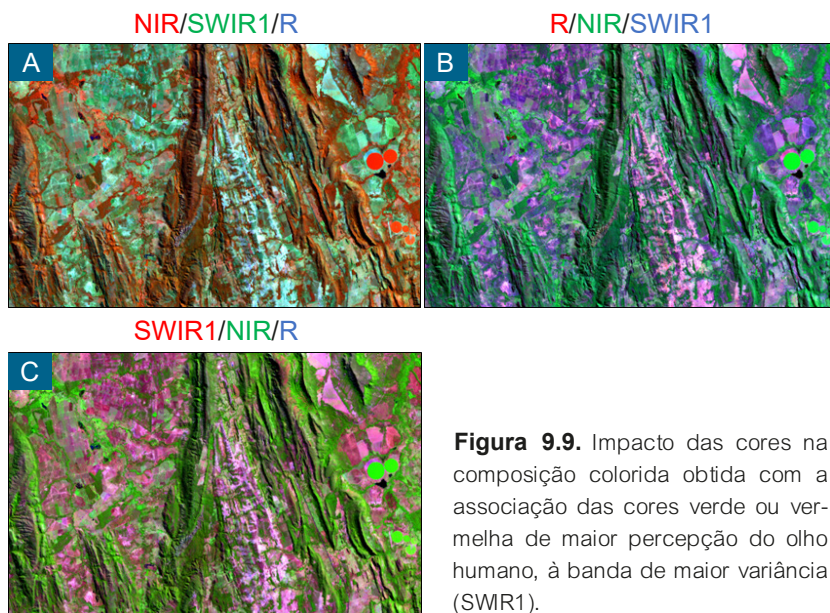


Figura 9.9. Impacto das cores na composição colorida obtida com a associação das cores verde ou vermelha de maior percepção do olho humano, à banda de maior variância (SWIR1).

9.4. Espaço de Cor IHS

No espaço de cores RGB, os eixos do cubo representam a quantização da radiância referente aos comprimentos de onda de um tripleto de bandas, resultando em imagens coloridas que satisfazem a maioria das necessidades do intérprete. Contudo, quando vamos descrever a cor de um objeto,

naturalmente nos referimos, por exemplo, se aquela cor vermelha é brilhante, clara ou escura, ou mesmo se é uma cor pura. Essa é a forma normal de se descrever a cor por meio da percepção fisiológica da visão. Os termos equivalentes em computação gráfica são intensidade (I), *hue* (H) e saturação (S), que são grandezas quantitativamente especificáveis e independentes, sob o ponto de vista de percepção visual. Diz-se, então, que I, H e S são atributos da cor fortemente percebidos pela visão. Dependendo dos valores digitais das bandas e das características do terreno da área imageada, esses atributos podem apresentar valores bem diferentes. A vantagem do espaço de cor IHS sobre o espaço de cor RGB está no fato de que é possível manipular cada componente I, H e S separadamente, obtendo-se, dessa forma, maior controle sobre o realce de cor. Isso é feito por meio de operações lineares e não lineares sobre os componentes I, H e S (Dutra e Meneses, 1986, 1987).

O sistema IHS não se trata de uma teoria de cor, mas de outra forma de modelar a representação espacial desses três atributos para a mesma cor obtida pelo espaço de cubo de cores. Ou seja, é uma transformação numérica do espaço de cor RGB para o espaço IHS. Define-se intensidade como a quantidade de brilho de uma cena, o que é uma medida do albedo refletido das partes iluminadas e sombreadas de todos os alvos presentes na imagem. É equivalente a ver uma cena colorida com uma visão em preto e branco. Assim, a intensidade é independente dos atributos saturação e *hue*, que são apropriadamente, os atributos da cor. Saturação é o grau de diluição de uma cor pura pela luz branca, por isso, é também chamada de pureza da cor, enquanto *hue* ou matiz define a cor espectral dominante. Nessa transformação do espaço de cor, a imagem *hue* derivada de três bandas é livre de sombra, pois os valores dos pixels na imagem *hue* são independentes de iluminação e estão relacionados apenas com as assinaturas de reflectância das três bandas.

Para se usar IHS como uma opção de processamento de cor, devem-se seguir os passos indicados na Figura 9.10. Com as três bandas selecio-

nadas, executa-se a transformação do espaço cúbico RGB para o espaço IHS, obtendo-se separadamente os componentes de I, H e S da imagem RGB. Nesse processo, não há um deslocamento numérico da cor, uma característica que geralmente não é verdadeira porque os componentes R, G, e B no cubo de cor, são sujeitos, como comentado anteriormente, por um aumento de contraste mínimo e máximo. Mudanças de contrastes em H ou S devem ser feitas por transformações apropriadas, como forma de se ter maior controle sobre os realces das cores da imagem. Com a aplicação da transformação IHS inversa, volta-se ao espaço de cores RGB para que se possa observar a imagem nos monitores coloridos. Além disso, é possível aplicar filtros passa-alta nessas composições de maneira a ressaltar detalhes espaciais.

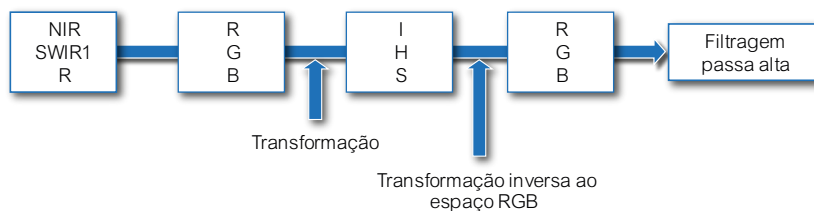


Figura 9.10. Etapas do processo de transformação do espaço de cores RGB para o espaço IHS.

Um exemplo da transformação para o espaço de cores IHS é mostrado na Figura 9.11, a partir do tripleto de bandas R, NIR, SWIR1 do sensor OLI. Os histogramas são apresentados para demonstrar o primeiro passo do processo de transformação, que consiste em extrair, do conjunto das três bandas, os três componentes independentes que caracterizam a cor. Note que os intervalos digitais das bandas são convertidos para o intervalo 0 – 1 para os componentes I e S, enquanto o componente H é expressa no intervalo de 0° a 360°. Antes dos componentes serem revertidos para o espaço RGB, o analista pode ajustar cada componente por métodos de realce de contraste. A diferença desses ajustes com os ajustes feitos no espaço RGB, é que, no espaço IHS, a mudança de cores é uniforme.

Na Figura 9.11, são mostradas, para efeito de comparação, a composição colorida no espaço de cor IHS e a composição colorida no espaço de cor RGB. Observe que a imagem IHS retrata, de forma bem mais contrastante e com finos detalhes, as variações de cores que são mais difíceis de serem percebidas na imagem RGB. Há uma pequena perda de textura do relevo porque, no processo de transformação, as sombras são computadas como valores de brilho.

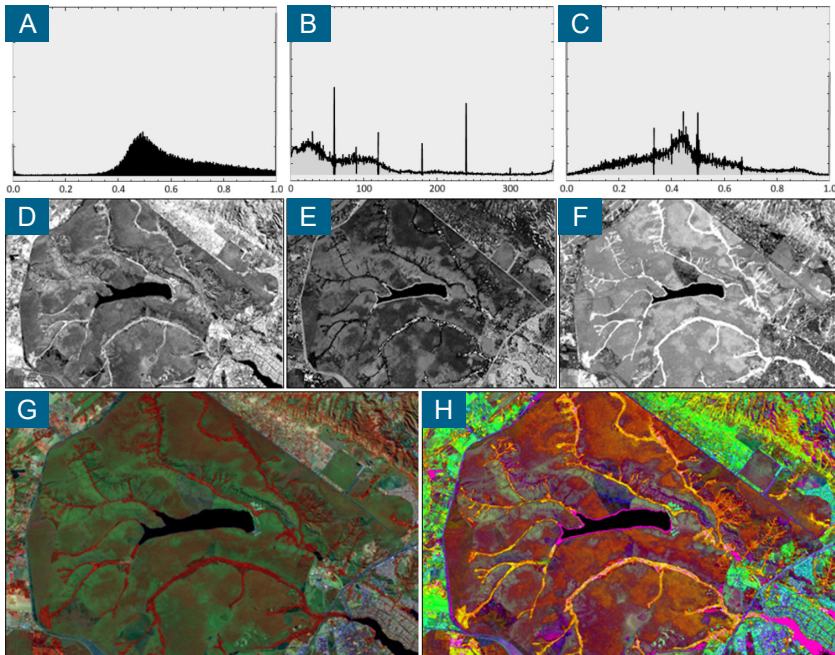


Figura 9.11. Histogramas (A, B e C) e as correspondentes imagens componentes IHS (D, E e F, respectivamente) das bandas do satélite Landsat 8 OLI nas faixas espectrais do infravermelho próximo (NIR), infravermelho de ondas curtas (SWIR1) e vermelho (R). Em (G) e (H), as composições coloridas RGB e IHS, respectivamente.

Para descrever o modelo geométrico do espaço de cor IHS, usa-se normalmente um hexacone que é representativo da maioria das versões de algoritmos usados em processamento de imagens equivalentes ao IHS,

como o HSV (*value*), e o HLS (*lightness*). *Value* e *lightness* são atributos com equivalência ao atributo intensidade I. Outra conhecida transformação é o HSV (carta de Munsell) que se baseia na carta de cores de solos, mas que utiliza a geometria de um cilindro. Variações dessas representações são possíveis e praticamente não se nota diferenças significativas de resultados de processamento entre esses algoritmos.

Schowengert (2007), utilizando a geometria de um hexacone, descreveu como se chega às medidas das coordenadas dos atributos IHS. Como o IHS é uma transformação do sistema RGB, a concepção da geometria do hexacone deve ser subtraída da geometria do cubo de cor. Imagine a projeção de cada subhexágono de cor em um plano perpendicular à linha do vetor preto ao branco, como visto na Figura 9.12. Movendo-se o vértice do vetor preto até o vértice branco, a projeção do plano resulta em uma série de hexágonos de diferentes dimensões. Essa série de hexágonos define o hexacone. O vetor preto ao branco define a intensidade de cada hexágono, enquanto os componentes H e S se localizam nas arestas do hexágono. No local onde o hexacone se reduz ao valor mínimo, a intensidade é zero, isto é, o hexacone é um ponto preto. À medida que se aumenta a intensidade, os tamanhos dos hexágonos vão aumentando, e o ponto preto caminha até atingir o seu máximo, o branco, como se fosse uma linha de níveis de cinza. A dimensão de cada seção transversal do hexacone determina a quantidade de matizes presentes (*hue*). Para um ponto P interno à projeção de um hexágono e com uma dada intensidade, os outros dois componentes da cor, *hue* e saturação, são definidos, geometricamente, na superfície do hexágono. O *hue* é determinado pela posição da matiz em torno do hexágono e é calculado pela rotação do ângulo azimutal de 0° a 360°. O vermelho está em 0°, o amarelo em 60°, o verde em 120°, o ciano em 180°, o azul em 240° e a magenta em 300°.

A saturação é determinada pela distância do ponto P ao centro do hexágono, por isso define a quantidade de branco que está associada à cor. Pontos mais distantes do centro representam matizes mais puros do que

aqueles mais próximos ao centro. Saturação e intensidade possuem valores no intervalo de 0 a 1. Para se efetuar o realce dos atributos IHS de uma forma independente, antes de voltar ao espaço RGB (transformação inversa), a componente I pode ser linearmente ampliada para realçar as áreas de baixo albedo, H é rotacionada pela escolha de um ângulo apropriado para destacar uma cor, e os valores de S podem ser linearmente reescalados a fim de aumentar a saturação e tornar as cores mais vivas.

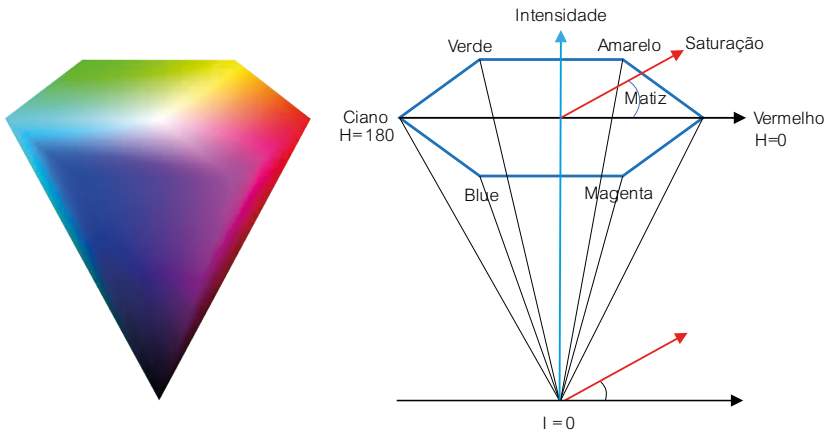


Figura 9.12. Modelo de representação dos atributos IHS da cor por meio da geometria do hexacone.

Fonte: adaptado de Schowengerdt (2007).

Como as imagens são armazenadas utilizando o espaço de cores RGB, torna-se necessária a utilização de equações para a transformação IHS. Transformações lineares podem ser usadas para definir os componentes, tornando os algoritmos de transformação de intensidade, matiz e saturação bastante simples e eficientes. Uma transformação bastante conhecida é a que foi proposta por Pratt (2014), que tem a seguinte formulação, também baseada na geometria do hexacone (Equação 9.3):

$$L \begin{bmatrix} 1 \\ V_1 \\ V_2 \end{bmatrix} \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/\sqrt{6} & 1/\sqrt{6} & 2/\sqrt{6} \\ 1/\sqrt{6} & 2/\sqrt{6} & 0 \end{bmatrix} = \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (9.3)$$

Após essa primeira transformação, os valores de matiz (H), saturação (S) e intensidade ($I = V$) são obtidos por (Equações 9.4 a 9.6):

$$LH = \tan^{-1} \frac{V_1}{V_2} \quad (9.4)$$

$$S = \sqrt{V_1^2 + V_2^2} \quad (9.5)$$

$$V = 1 \quad (9.6)$$

A transformação inversa para o espaço RGB, para que a imagem possa ser visualizada nos monitores coloridos que são de padrão RGB, é dada por (Equações 9.7 a 9.9):

$$V_1 = S \cos H \quad (9.7)$$

$$V_2 = S \sin H \quad (9.8)$$

$$I = V \quad (9.9)$$

Conforme mencionado anteriormente, outras variações de transformações existem tais como o HLS (*hue*, *lightness*, saturação) ou HSV (*hue*, saturação, *value*). No modelo HLS, o espaço de cor é definido por um hexacone duplo (Figura 9.13). Nesse modelo, diferentemente do que ocorre no HSV, os valores de L e S nos quais se obtém o mais forte nível de luminosidade e saturação são em $L = 0,5$ e $S = 1$, enquanto no HSV, para se ter o mesmo efeito, teríamos $V = 1$ e $S = 1$ (Foley et al., 1990).

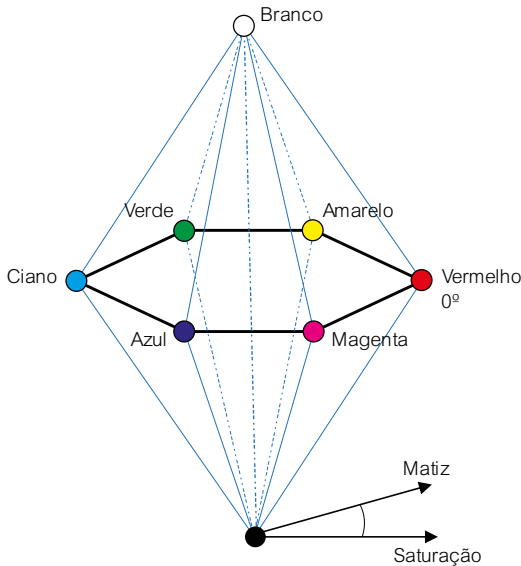


Figura 9.13. Modelo de cor HSL por meio da geometria do hexacone duplo.

Fonte: adaptado de Foley et al. (1990).

Independentemente de quais sejam as equações de transformação, Souto (2000) observou que um problema inerente aos componentes da cor é a instabilidade de *hue* sob condições variadas de saturação e/ou intensidade. Ao se observar as imagens dos componentes, esse autor comprovou que é percebida uma acentuada instabilidade de *hue* (cores) quando a cena real do terreno exibe baixos valores de saturação. À medida que o valor de saturação aumenta, a instabilidade de matiz diminui até chegar a uma situação de estabilidade. Para valores baixos de intensidade, o componente *hue* é mais sensível (ou instável) do que as variações de valores dos componentes no espaço RGB.

9.5. Referências

- DUTRA, L. V.; MENESES, P. R. Aplicação da transformação IHS para realce de cores em imagens Landsat. In: SIMPÓSIO LATINO-AMERICANO DE SENSORIAMENTO REMOTO, Gramado, RS. **Anais...** São José dos Campos: INPE, p. 675-681, 1986.
- DUTRA, L. V.; MENESES, P. R. Realce de cores em imagens de sensoriamento remoto utilizando rotação de matiz no espaço IHS. In: SIMPÓSIO BRASILEIRO DE TELECOMUNICAÇÕES, 5., Campinas, SP. **Anais...** Rio de Janeiro: SBrT, p. 209-212, 1987.
- FOLEY, J. D.; VAN DAM, A.; FEINER, S. K.; HUGHES, J. F. **Computer Graphics**. Principles and Practice. Reading: Addison-Wesley, 2a ed., 1990.
- PRATT, W. K. **Introduction to digital imaging processing**. Boca Raton, CRC Press, 2014.
- SOUTO, R. P. **Segmentação de imagens multiespectrais utilizando-se o atributo matiz** (dissertação de mestrado em Sensoriamento Remoto). São José dos Campos, SP, INPE, 2000, 171 p.

CAPÍTULO 10

Fusão de Imagens

Paulo Roberto Meneses



Na busca de um equilíbrio entre a pequena largura espectral das bandas e a razão sinal/ruído, os sensores multiespectrais tipicamente apresentam campos de visadas instantâneos maiores, ou seja, baixas resoluções espaciais para permitir a coleta de um grande número de fótons refletidos do pixel e manter uma boa relação sinal/ruído. A partir da metade da década de 1980 e baseada na experiência com sensores fotográficos que obtêm fotografias pancromáticas (uma única banda com ampla largura espectral) com alta resolução espacial, a maioria dos sensores orbitais passaram a incluir, no módulo multiespectral, uma banda pancromática com alta resolução espacial. O objetivo de combinar a banda pancromática com as bandas multiespectrais é permitir a produção de imagens multiespectrais com alta resolução espacial. Os algoritmos de processamento dessa operação são conhecidos como fusão de imagens e as primeiras tentativas utilizaram a transformação intensidade-hue-saturação (IHS) para fundir ou fusionar imagens de baixa resolução espacial com uma imagem de alta resolução espacial. É possível não só fusionar imagens do mesmo sensor, mas também imagens de sensores diferentes. A condição imposta é que as imagens precisam estar geometricamente registradas antes da fusão. É aconselhável ainda que também possuam a mesma resolução radiométrica. Neste capítulo, veremos as diferentes opções de algoritmos disponíveis para fusão de imagens.

10.1. IHS

A alternativa para descrever as cores de uma imagem decompondo-as nos componentes de intensidade (I), *hue* (matiz) e saturação (S), estende-se além do seu uso para obtenção de composições coloridas. A transformação IHS tem sido mais usada como uma técnica poderosa para combinar imagens de diferentes resoluções espaciais. A representação no espaço de cores IHS apresenta algumas propriedades, entre as quais se destaca o ordenamento das resoluções espaciais dos componentes I, H e S, na forma

como o olho humano é capaz de percebê-las. A componente I é percebida por sua máxima resolução, seguida de H e S, devido ao fato de que a retina do olho humano é composta de maior número de bastonetes que são responsáveis pela percepção de luminosidade ou brilho do que de cones que são responsáveis pela percepção da matiz e da saturação.

Imagine uma foto aérea pancromática tomada com um filme preto e branco na faixa espectral do visível de um terreno cheio de objetos coloridos aos nossos olhos. As cores mais vivas da paisagem corresponderiam a tons de cinza claros na fotografia pancromática preto e branco, enquanto as cores escuras corresponderiam a tons de cinza escuros. A foto aérea em preto e branco seria, portanto, uma representação da intensidade. Para compensar a falta dos outros dois atributos da cor, elas têm, como principal característica, a alta resolução espacial, de alguns poucos centímetros, o que possibilita a identificação de objetos pela análise de suas formas e não pelas propriedades de matiz e saturação. Elas possuem, assim, uma baixa resolução espectral, mas uma alta resolução espacial. Então, o ideal seria ter um sensor que pudesse registrar com a máxima eficiência os três componentes das cores com alta resolução espacial. Por essa razão, os sensores multiespectrais passaram a incluir, no módulo multiespectral de baixa resolução espacial, uma banda pancromática de maior resolução espacial (banda PAN).

Para que se possa compreender a operação de fusão de forma efetiva, mostramos, na Figura 10.1, os passos para realizar a fusão de imagens pela técnica IHS. O primeiro passo é selecionar três bandas do módulo multiespectral que contenham as melhores informações espectrais de interesse e realizar a transformação para o espaço IHS. Os componentes H e S que guardam a informação espectral das cores são preservados e o componente I é descartado e substituído pela imagem pancromática de alta resolução espacial. O passo seguinte consiste em reamostrar os componentes H e S para a resolução espacial da banda pancromática, sintetizando uma imagem colorida com resolução espacial da banda pancromática e com as cores do módulo multiespectral. Opera-se, então, a transformação inversa para o espaço de cor RGB. Apesar de ser uma técnica bastante utilizada, a

fusão pelo método IHS possui a limitação de permitir apenas a fusão de três bandas com a imagem de alta resolução espacial.

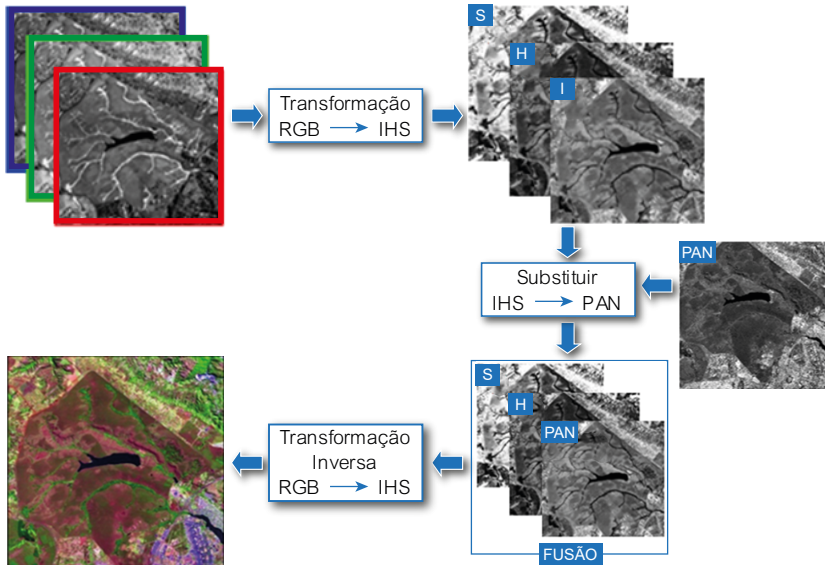


Figura 10.1. Etapas do processo de fusão de imagens pela técnica IHS.

As outras propostas de técnicas de fusão são baseadas em três opções de modelos no domínio espacial, domínio espectral e algébricos. Os modelos no domínio espacial utilizam as transformadas de wavelet e a técnica de filtragem passa-alta para retirar a informação espacial de alta frequência da imagem de alta resolução espacial e combiná-la com a imagem multiespectral. Os três principais modelos no domínio espectral são a técnica IHS, principais componentes e a de Gram-Schmidt. Já nos modelos algébricos incluem-se as técnicas de Brovey e a multiplicativa que utilizam operações aritméticas pixel a pixel. Muitos trabalhos têm sido publicados com essas técnicas para diversas aplicações, podendo-se citar os estudos conduzidos por Carper et al. (1990) e Francisco e Braga (2021). Uma breve descrição dessas técnicas é apresentada a seguir. Maiores detalhes podem ser obtidos em Vrabel (1996) e Welch e Ehlers (1987).

10.2. Modelos Algébricos

10.2.1. Modelo de Brovey

O modelo algébrico de Brovey utiliza uma combinação matemática de soma e multiplicação entre as imagens de diferentes resoluções espaciais. Cada banda da imagem multiespectral (MSI) é multiplicada pela imagem de alta resolução espacial (*PAN*) e dividida pelo somatório das bandas multiespectrais (Equação 10.1). Essa operação normaliza a cor da imagem fusio-nada, contribuindo para diminuir a instabilidade da matiz sob condições de baixa saturação e/ou intensidade. Somente três bandas multiespectrais são usadas. A fusão Brovey possui a mesma limitação da fusão IHS, pois tra-balha com o sistema de cores RGB. No entanto, tem-se mostrado superior ao IHS quando as bandas multiespectrais são da região do infravermelho.

$$CN = \frac{(MS_i + 1) \times (PAN + 1) \times 3}{\sum_i MS_i \times 3} - 1 \quad (10.1)$$

10.2.2. Normalização espectral de cores

A técnica de normalização espectral de cores é uma extensão do algorit-mo Brovey, mas sem restrição ao número de bandas e à resolução radiomé-trica da imagem sintética de saída. Entretanto, há uma única restrição para a execução dessa fusão: o intervalo espectral das bandas de entrada deve ser o mesmo da banda de alta resolução espacial (*PAN*). O processo de fusão começa com o agrupamento das bandas de entrada em segmentos espectrais cujos intervalos espectrais são definidos pelo valor central do comprimento de onda de cada banda, o que é obtido do arquivo cabeçalho (txt) da imagem. Em seguida, é realizada uma operação aritmética seme-lhante à da fusão Brovey. Cada banda de entrada é multiplicada pela banda

de alta resolução espacial e normalizada pela divisão da soma de bandas multiespectrais, na forma da Equação 10.2.

$$CN = \frac{(MS_i \times PAN \text{ (banda segmento)})}{\sum \text{bandas de entrada segmentos} + \text{números bandas segmento}} \quad (10.2)$$

Podemos considerar que o método de fusão mais eficiente é aquele que preserva as cores da imagem original e que consegue transferir a informação espacial de alta frequência da imagem de alta resolução espacial.

10.3. Modelos no Domínio Espectral

10.3.1. Componentes principais

A fusão por componentes principais é a forma de não limitar o número de bandas que podem ser fusionadas, como ocorre com as técnicas Brovey e IHS. Inicia-se com a transformação das bandas multiespectrais em um mesmo número de componentes não correlacionados. O primeiro componente principal (CP1) é substituído pela banda pancromática, que antes passa por um processo de equalização de média e variância para que fique mais parecida possível com a CP1.

10.3.2. Gram-Schmidt e NN-Diffuse

O procedimento de Gram-Schmidt, assim como o de principais componentes, trata-se de uma operação sobre vetores, com o objetivo de torná-los ortogonais. A fusão inicia-se simulando a existência de uma banda pancromática a partir das bandas multiespectrais de baixa resolução espacial. A banda pancromática simulada deve estar dentro do intervalo espectral das bandas de baixa resolução. Na sequência, uma transformação de Gram-Schmidt é aplicada à banda pancromática simulada e às bandas multiespectrais. A banda pancromática simulada é considerada como a primeira

banda Gram-Schmidt que é trocada pela banda pancromática de alta resolução. Finalmente, uma transformação inversa é aplicada para formar a imagem sintética de saída. O método Gram-Schmidt é mais preciso do que o método de principais componentes porque ele usa a função de resposta do sensor para estimar como a banda pancromática deve aparecer.

Exemplos de fusão de imagens pelos métodos acima descritos são apresentados na Figura 10.2. As imagens são do sensor ETM+ do satélite Landsat 7, com seis bandas no módulo multiespectral ($0,45\text{ }\mu\text{m}$ – $2,35\text{ }\mu\text{m}$) de resolução espacial de 30 metros e uma banda pancromática ($0,55\text{ }\mu\text{m}$ – $0,95\text{ }\mu\text{m}$) com resolução espacial de 15 metros.

Sun et al. (2014) propuseram um novo método de pan-sharpening baseado em modelo de mistura, denominado de *Nearest-Neighbour Diffuse Pan-Sharpening* (*NNDiffuse Pan-Sharpening*). Esse método assume que cada pixel da banda pancromática é uma mistura linear ponderada dos superpixels vizinhos das bandas espectrais de baixa resolução espacial. Os pesos são controlados por um modelo de difusão inferido a partir da imagem pancromática que relaciona a similaridade do pixel de interesse com os superpixels vizinhos, isto é, com um subconjunto de pixels conectados que são espectralmente similares. A natureza da operação do espectro por pixel é diferente dos algoritmos existentes, que dependem principalmente de processamento banda a banda. Por causa do seu alto desempenho, essa técnica tem-se tornado cada vez mais popular por causa do seu desempenho superior em relação aos outros métodos baseados em pan-sharpening.

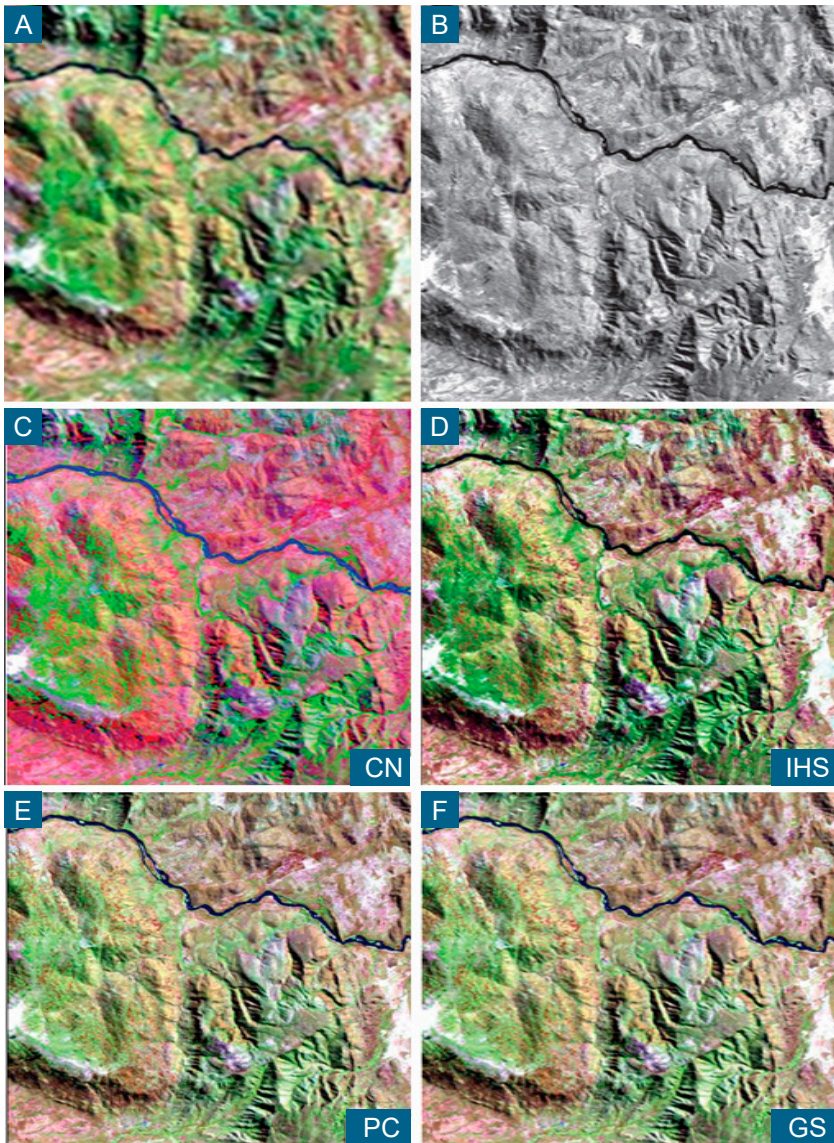


Figura 10.2. Resultados da fusão da imagem multiespectral obtida pelo sensor Landsat 7 ETM+ de resolução espacial de 30 m (A) com a imagem pancromática de resolução de 15 m (B), pelos métodos baseados no modelo algébrico (CN) (C), domínio espectral (IHS) (D), componentes principais (PC) (E) e Gram-Schmidt (GS) (F).

10.4. Referências

CARPER, W.; LILLESAND, T. M.; KIEFER, R. W. The use of Intensity-Hue-Saturation transformation for merging spot panchromatic and multispectral image data.

Photogrammetric Engineering and Remote Sensing, v. 56, n. 4, p. 459-467, 1990.

FRANCISCO, R. F.; BRAGA, A. C. O. Fusão de imagens multissensor para a identificação de feições erosivas no município de Bauru/SP. **Revista Brasileira de Geomática**, v. 9, n. 1, p. 21-35, 2021.

SUN, W.; CHEN, B.; MESSINGER, D. W. Nearest-neighbor diffusion-based pan-sharpening algorithm for spectral images. **Optical Engineering**, v. 53, n. 1, 013107, 2014.

VRABEL, J. Multispectral imagery band sharpening study. **Photogrammetric Engineering & Remote Sensing**, v. 62, n. 9, p. 1075-1083, 1996.

WELCH, R.; EHLERS, M. Merging multiresolution SPOT HRV and Landsat TM data.

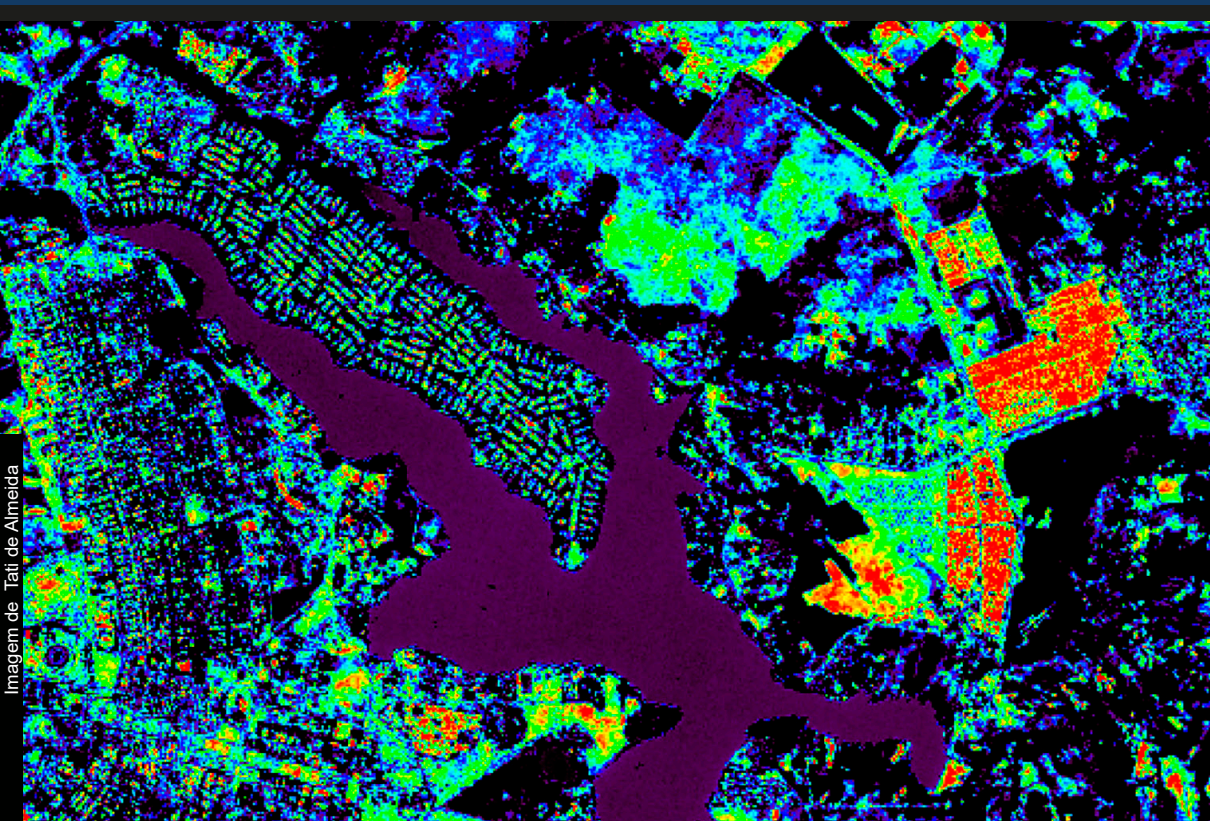
Photogrammetric Engineering & Remote Sensing, v. 53, n. 3, p. 301-303. 1987.

CAPÍTULO 11

Aritmética de Bandas

Paulo Roberto Meneses

Tati de Almeida



A aritmética de bandas é uma das formulações algorítmicas mais simples de processamento de imagens e que produz resultados expressivos. A facilidade para executar as operações aritméticas de soma, subtração, multiplicação e divisão é uma notável demonstração das vantagens do uso de processamento de imagens multiespectrais. Trata-se de um processo bastante rápido para realçar toda a cena de uma imagem ou alvos específicos de interesse do analista e, também, para ajustar e suavizar imagens muito ruidosas. A aritmética de bandas é bastante utilizada para detectar mudanças (*change detection*) no meio ambiente a partir de imagens adquiridas em datas diferentes. As operações aritméticas realizam um processamento de transformação de dados a partir da combinação de bandas, gerando, como resultado, uma nova imagem completamente distinta das imagens de entrada. Uma nova matriz de dados é gerada. As imagens envolvidas nas operações aritméticas devem ter as mesmas dimensões espaciais em termos de número de pixel, número de linhas e resolução espacial.

11.1. Operações Aritméticas

As transformações aritméticas que envolvem soma, subtração, multiplicação e divisão são executadas pixel a pixel por meio de uma regra matemática predefinida envolvendo, normalmente, no mínimo duas bandas do mesmo sensor (Figura 11.1). A soma, subtração e multiplicação são entendidas como transformações lineares, enquanto a divisão de bandas corresponde a uma transformação não linear e é a que mais pode ser interpretada em termos de aspectos físicos do imageamento do terreno pelos sensores.

As operações mais utilizadas em sensoriamento remoto são a divisão e a subtração de bandas, com aplicação reduzida para as operações de soma e multiplicação. Como o resultado é uma única imagem de saída, e que não deve ser chamada de banda, as operações aritméticas são uma forma de recurso para reduzir o número de imagens contidas em um arquivo de

dados, o que é uma vantagem em termos de armazenamento em disco. Em termos gerais, a soma e a multiplicação servem para realçar similaridades espectrais entre imagens de diferentes bandas ou datas, enquanto a subtração e a divisão são utilizadas para realçar diferenças espectrais dos alvos.

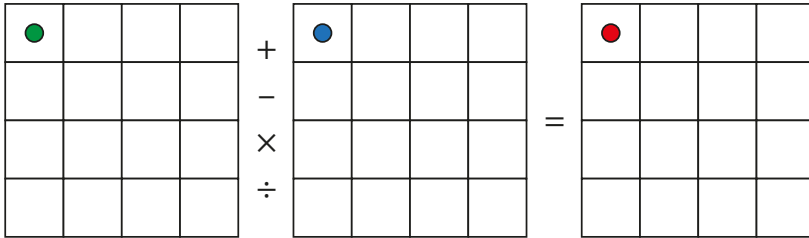


Figura 11.1. Operações aritméticas são executadas pixel-a-pixel entre duas bandas, gerando uma nova imagem.

11.1.1. Soma

Com um par de imagens ou n imagens, a operação soma é definida como a soma de bandas espectrais ou de várias imagens obtidas em datas distintas. Como a soma pode resultar em valores de pixels que caem fora do intervalo de número digital da imagem, uma compressão de dados se faz necessária, o que normalmente pode ser feita dividindo-se o resultado da soma pelo número de bandas somadas. Como a nova imagem (b_{soma}) é uma média aritmética dos pixels de todas as bandas somadas, são somados tanto os pixels bons como os pixels ruidosos, o que tende a minimizar os ruídos que podem estar presentes nas imagens originais. A expressão matemática genérica para a soma consiste em uma média de valores inteiros (Equação 11.1).

$$b_{soma}(l, p) = int \frac{b_i(l, p) \dots + \dots b_n(l, p)}{n} \quad (11.1)$$

em que

l = pixel na linha l ;

p = pixel na coluna p ;

n = número de bandas; e

int = inteiro.

O resultado da soma e da multiplicação de dois números inteiros também é um valor inteiro, enquanto a divisão de dois números inteiros não necessariamente resulta em um número inteiro. Outra opção matemática da soma, tendo-se a mesma situação anterior em que os pixels caem fora do intervalo digital da imagem, é reescalonar os dados de saída para os valores extremos mais altos ou mais baixos dos dados originais. Por exemplo, os valores de saída, para serem expostos no monitor de vídeo de 8 bits, somente assumem valores dentro do intervalo de 0 a 255.

Segundo Crósta (1999), a soma é utilizada para realçar similaridades entre imagens de diferentes bandas ou datas, resultando em um aumento de contraste entre as classes de alvos presentes nas imagens originais. Um exemplo numérico da soma dos valores digitais de duas bandas é mostrado na Figura 11.2. Observe, na imagem da soma, que três conjuntos (classes) de níveis de cinza são destacados ou realçados.

Por livre iniciativa do analista, pode-se também combinar a adição de três pares de bandas com as respectivas cores RGB a fim de se obter uma composição colorida. Os resultados são meras tentativas de sucesso, visto que não há um controle ou regra espectral que possa orientar o analista a alcançar o resultado desejado.

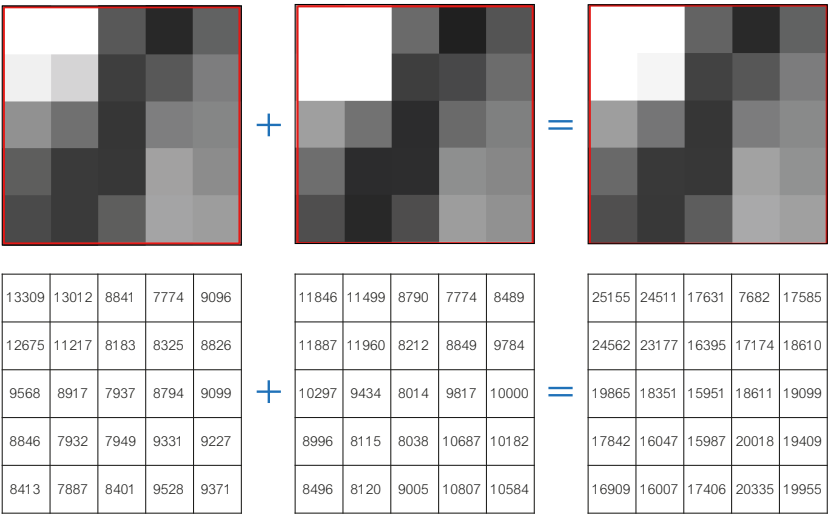


Figura 11.2. Soma de duas bandas com similaridades espectrais dos alvos.

11.1.2. Multiplicação

A multiplicação de bandas é talvez a menos usada das operações aritméticas em sensoriamento remoto. Uma das principais dificuldades é a faixa extrema de valores que pode ser gerada. A expressão matemática abaixo é normalmente usada na multiplicação de duas bandas e tem a finalidade de realçar os dados que são semelhantes em ambas as bandas (Equação 11.2).

$$b(l, p) = \text{int} \frac{b_1(l, p) \times b_n(l, p)}{k - 1} \quad (11.2)$$

Florenzano et al. (2001) encontraram uma forma de usar a multiplicação de bandas como recurso para obter uma imagem que agrega o realce da informação textural de relevo com a informação espectral. Para isso, multiplicaram a banda do infravermelho próximo (NIR) do sensor *Operational Land Imager* (OLI) do satélite Landsat 8 que, em áreas com cobertura ve-

getal verde, consegue ter uma boa informação de relevo, pelas bandas 2 (azul), 5 (infravermelho próximo) e 7 (infravermelho médio), combinando os resultados dessa multiplicação em uma composição colorida. Comparada com uma composição colorida simples RGB das bandas 2, 5 e 7, dá para perceber que o resultado é uma imagem com maior valor para interpretação visual das formas de relevo, enquanto se vê uma leve suavização das cores que ajuda a retirar, em benefício do intérprete, o efeito distrativo de cores saturadas de outros alvos (Figura 11.3).

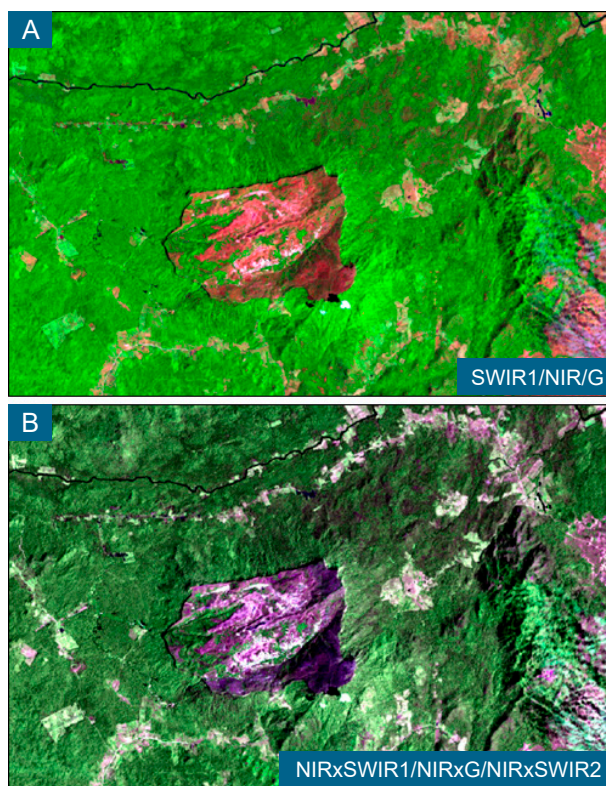


Figura 11.3. Imagens Landsat 8 OLI da região da serra do Tepequém/RO. Em (A), composição SWIR1/NIR/G e em (B), composição da multiplicação $(NIR \times SWIR1)/(NIR \times G)/(NIR \times SWIR2)$. SWIR1 = infravermelho de ondas curtas 1; NIR = infravermelho próximo; G = verde; e SWIR2 = infravermelho de ondas curtas 2.

11.1.3. Subtração

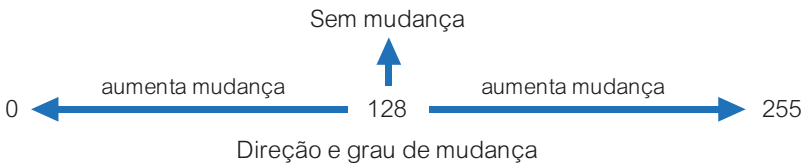
A subtração de bandas é uma operação de grande recurso prático para identificar, espectralmente, feições que tenham mudado entre duas imagens adquiridas sobre a mesma área em datas diferentes. Ela pode ser feita entre duas bandas do mesmo comprimento de onda e de datas diferentes, ou entre dois conjuntos de imagens também de datas diferentes. Os arquivos têm que ter o mesmo número de bandas, a mesma resolução espacial e serem corregistrados ou georreferenciados. Essa operação é conhecida como *change detection* e tem grande uso na detecção de impactos ambientais e na avaliação de desmatamentos florestais.

Nesse tipo de aplicação, é importante considerar alguns fatores que podem causar diferenças entre as imagens de datas diferentes, sem que haja mudanças dos alvos. Esses fatores podem estar relacionados com: diferenças nos modos de imageamento dos sensores, por exemplo, *Enhanced Thematic Mapper Plus* (ETM+) e OLI, que podem ter a mesma largura em termos de faixas espectrais, mas com os centros de comprimentos de onda um pouco diferentes, ou terem diferenças de funções de respostas espectrais; imagens tomadas com ângulos de iluminação (inclinação e azimute solar) diferentes, como por exemplo, uma imagem no verão e a outra no inverno, especialmente se as imagens forem adquiridas em áreas do terreno com forte controle estrutural; diferenças nas condições atmosféricas entre imagens de datas diferentes, as quais causam intensidades diferentes de espalhamento da radiação. O corregistro entre as bandas das duas imagens deve estar assegurado, por isso, a subtração entre imagens do mesmo sensor torna-se mais confiável.

O método mais simples de subtração é o de duas bandas do mesmo comprimento de onda de duas imagens de datas diferentes, o que, neste caso, pode resultar, no extremo, em diferenças negativas máximas de valores digitais, contando que sempre há, nas imagens, alguns pixels com $ND_{min} = 0$. Para compensar os valores negativos dos pixels, pois não existe radiância negativa, é preciso adicionar um valor ao resultado da subtração

para ajustar o valor digital mínimo para zero. De acordo com essa formulação matemática (Equação 11.3), os pixels com valores iguais à mediana do intervalo digital das imagens correspondem às áreas sem mudanças, enquanto as áreas que sofreram mudanças ao longo das duas datas terão pixels com valores tendendo para zero (alvos mais claros na data 1 do que na data 2) ou para ND_{max} (alvos mais escuros na data 1 do que na data 2). Neste caso, pode-se dizer que é diagnosticada uma direção de mudança.

$$g_{dif} = \frac{(g_1 - g_2) + 255}{2} \quad (11.3)$$



Os métodos atuais de detecção de mudanças oferecem uma ampla variedade de opções para identificar, descrever e quantificar diferenças de feições entre conjuntos de imagens da mesma área de datas diferentes. Os seguintes métodos abaixo descritos de detecção de mudanças relativas são empregados no programa ENVI 6.5. Mudanças relativas significam que alguma coisa mudou, mas não especificam o que mudou. É um método mais rápido.

Vemos isso no exemplo da Figura 11.4 de duas imagens tomadas em 2001 e 2006. Subtraindo a banda 3 do satélite Landsat 7 ETM+ de 2006 da imagem de 2001, os pixels na direção ao zero representam mudanças das classes de alvos que eram mais claras (maior reflectância) em 2001 do que eram em 2006, e os pixels na direção a 255 são das classes de alvos que eram mais escuros em 2001 do que eram em 2006.

(imagem ano 200 – imagem 2006) = imagem direfença

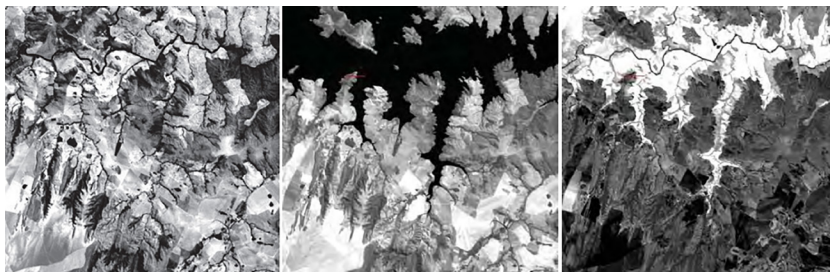


Figura 11.4. Detecção de mudanças por meio de subtração de imagens tomadas em diferentes datas. Na imagem diferença, os tons de cinza mais escuros e mais claros indicam as áreas que sofreram as maiores mudanças.

- ***Transformação***

Técnicas de processamento de transformação de imagens como a análise de componentes principais, fração de ruído mínimo (MNF) e análise de componentes independentes são computadas nas imagens de entrada de datas diferentes, a fim de se extrair as feições correlacionadas a mudanças.

- ***Subtração***

Às imagens de entrada, são aplicadas funções de processamento de divisão (razão) de bandas, tais como o índice de vegetação por diferença normalizada (NDVI) e outras divisões propostas pelo analista. Os resultados dessas divisões são subtraídos das imagens de entrada para se criar as diferenças de imagens.

- ***Visada Múltipla de Duas Cores***

Uma banda de uma imagem da data 1 é exposta no canal vermelho do monitor e a mesma banda da data 2 é exposta nos canais azul e verde. Objetos que são mais claros na imagem da data 1, quando comparados com a imagem da data 2, aparecem em vermelho. Objetos que são mais escuros na imagem da data 1, quando comparados com a imagem da data

2, aparecem em ciano. Essas cores podem ser usadas para indicarem potenciais áreas de mudanças.

- *Ângulo Espectral*

Esse método computa o ângulo espectral entre os espectros de pixels correspondentes das imagens das datas 1 e 2. As diferenças dos ângulos espectrais medem as diferenças entre os espectros e, conseqüentemente, das áreas que sofreram mudanças.

11.1.4. Divisão de bandas

No processo de extração de informações de imagens de sensores multiespectrais, em muitas situações, há o interesse em discriminar variações sutis de reflectância de alvos específicos, como um tipo de litologia, solo, água, vegetação ou outros tipos de alvos de interesse do analista. Nesse caso, deve-se buscar por meios que possam destacar exclusivamente as propriedades daquele tipo de alvo, em detrimento a todos os outros presentes na cena, sem interesse. A divisão ou razão de bandas é a operação aritmética usada para este fim e, por isso, se tornou em uma das mais úteis e abrangentes técnicas de transformação de imagens. Ela se baseia em parâmetros dos aspectos físicos do sensor e no comportamento da reflectância dos alvos, como veremos logo adiante.

Como mostra a Equação 11.4, a divisão de imagens consiste na divisão pixel-a-pixel de uma banda pelo pixel de posição correspondente de outra banda do mesmo sensor, tal como representado na Figura 11.5. A divisão trabalha no realce das assinaturas espectrais dos alvos e, para isso, é necessário que o intérprete selecione corretamente as bandas do sensor. A seleção das bandas é decidida com base no conhecimento que o intérprete tem acerca do padrão de reflectância do alvo de interesse. Daí decorre ser uma técnica controlada pelo intérprete, com resultados que se têm mostrado importantes para mapeamentos litológicos, de solos, vegetação e corpos de água.

$$b_{div}(l, p) = \frac{b_1(l, p)}{b_2(l, p)} \quad (11.4)$$

Para realçar uma determinada classe de alvo, como um tipo de solo ou de rocha, pela técnica de divisão de bandas, as áreas desses alvos precisam expor, na imagem, padrões homogêneos e bem definidos de reflectância. Porém, em áreas de relevos acidentados, os sombreamentos alteram os valores de reflectância dos alvos, e o pressuposto de um padrão de reflectância diagnóstico daquele alvo é perdido. Nesse caso, a recuperação da reflectância do alvo sombreado implica na supressão da sombra. Há, dessa forma, um compromisso mútuo entre a possibilidade de se realçar alvos específicos e a supressão de sombras em imagens multiespectrais.

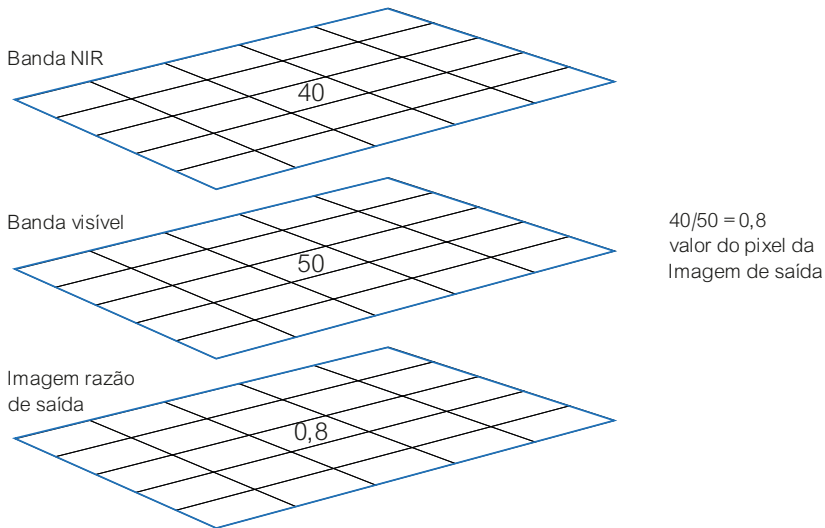


Figura 11.5. Exemplo de processo de divisão de duas bandas para um pixel com valores digitais 40 na banda NIR e 50 na banda visível.

Essa situação é demonstrada na Figura 11.6 para uma forma hipotética de relevo. Nessa figura, vemos, pela geometria de iluminação do relevo, que

o ângulo θ , formado pela direção do raio solar incidente com a normal à superfície, é diferente entre a face frontal da encosta que é fortemente iluminada e a face oposta da encosta que é sombreada. Isso tem uma implicação matemática. Como o valor do cosseno do ângulo θ é parte da equação que mede os valores quantitativos da reflectância de uma superfície, mesmo que as duas encostas fossem constituídas do mesmo material, a reflectância da face oposta seria diferente da face frontal devido ao sombreamento, não devendo ser interpretada como uma variação de reflectância do alvo.

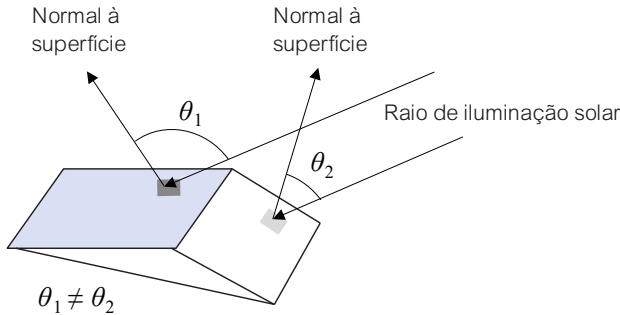


Figura 11.6. Influência da variação do ângulo de iluminação solar (θ) na reflectância de superfície do terreno em virtude das inclinações das vertentes frontal e oposta.

A demonstração matemática da divisão de bandas para supressão do sombreamento é mostrada a seguir. A equação matemática que descreve a radiância medida pelo sensor para cada banda, integra, além do cosseno do ângulo de incidência (θ), as variáveis relacionadas à irradiância da fonte (E), trajetória atmosférica (T), reflectância do alvo (ρ), e parâmetros físicos do sensor que podem ser expressos na forma da Equação 11.5.

$$L_{\lambda} = \frac{1}{\pi} \int E_{\lambda} R_{\lambda} \{ T_{\beta(\lambda)} T_{z(\lambda)} \rho_{(\lambda)} \cos \theta + \rho' \theta \} \quad (11.5)$$

em que

E_{λ} = irradiância solar;

R_λ = função da resposta espectral do sensor;

$T_{\beta(\lambda)}$ = transmitância da atmosfera na trajetória descendente;

$T_{z(\lambda)}$ = transmitância da atmosfera na direção zenital refletida da superfície;

$\rho_{(\lambda)}$ = reflectância da superfície; e

$\rho'\theta$ = reflectância da atmosfera, indicando a dependência da elevação solar.

Reescrevendo a Equação 11.5 de uma forma mais simples, tem-se (Equação 11.6):

$$L = E_\lambda G_\lambda T_\lambda \rho_\lambda \theta + N_\lambda \quad (11.6)$$

em que

E_λ = irradiância solar;

G_λ = fator instrumental;

T_λ = transmitância da atmosfera;

$\rho_{(\lambda)}$ = reflectância da superfície;

θ = ângulos de iluminação solar do terreno; e

N_λ = radiância aditiva da atmosfera.

A divisão de duas bandas baseada na Equação 11.6 é dada por (Equação 11.7):

$$\frac{L_{\lambda 1}}{L_{\lambda 2}} = \frac{E_{\lambda 1} G_{\lambda 1} T_{\lambda 1} \rho_{\lambda 1} \theta + N_{\lambda 1}}{E_{\lambda 2} G_{\lambda 2} T_{\lambda 2} \rho_{\lambda 2} \theta + N_{\lambda 2}} \quad (11.7)$$

Dos parâmetros envolvidos na Equação 11.7, o único termo que possui o mesmo valor nas duas bandas é o ângulo θ de direção de iluminação solar (não confundir com o ângulo θ de incidência na superfície). Nessa equação, a divisão anula o termo θ e, dessa forma, a reflectância das duas faces dei-

xa de ter a influência das variações de iluminação com a normal à superfície das encostas de relevo. O efeito na imagem é a eliminação da textura de relevo (na realidade, uma redução), deixando a superfície topográfica plana, sem sombreamento. Sem o efeito de sombreamento, a reflectância da superfície do alvo torna-se uniforme e configura representar sua assinatura espectral.

Segundo Drury (1987), o efeito do sombreamento não é completamente eliminado por duas razões: primeiro, os terrenos naturais não são uma superfície Lambertiana, ou seja, as reflectâncias das superfícies variam em função do ângulo entre as superfícies e a iluminação incidente, e em função das formas das estruturas, tal como o dossel de uma vegetação; segundo, os efeitos atmosféricos modificam os valores de razões de diferentes encostas, de uma maneira imprevisível. Devido a isso, é aconselhável corrigir a atmosfera, transformando a imagem digital para imagem de reflectância antes de se realizar a divisão.

Quando o sombreamento é suprimido da divisão de bandas após a sua execução, automaticamente o realce do alvo desejado é alcançado, razão pela qual a divisão de bandas é uma técnica de realce de imagens de grande poder. Um exemplo prático da aplicação da divisão de bandas para realçar alvos de interesse do intérprete é ilustrado na Figura 11.7. Imagine quatro diferentes tipos de alvos que tenham padrões de reflectâncias como no esquema dessa figura. Os quatro alvos, quando vistos na banda 1 ou na banda 2, não mostram diferenças significativas de reflectância entre si, à exceção do alvo D. Sendo assim, os alvos A, B e C dificilmente podem ser discriminados entre si, seja na banda 1 ou na banda 2. Como mostrado na tabela da figura, os valores digitais dos pixels equivalentes aos valores de reflectância dos alvos A, B, C e D estão muito próximos entre si. Mas, com base na Equação 11.4, quando a banda 2 é dividida pela banda 1, os valores digitais da imagem de saída tornam-se números fracionários e pequenos que ocupam somente uma estreita faixa do intervalo dinâmico dos valores digitais da imagem. Isso requer que uma operação adicional

seja processada na imagem resultante da divisão, unicamente para expor visualmente a imagem divisão na tela do computador de forma realçada. Vamos usar a Figura 11.7 para explicar melhor.

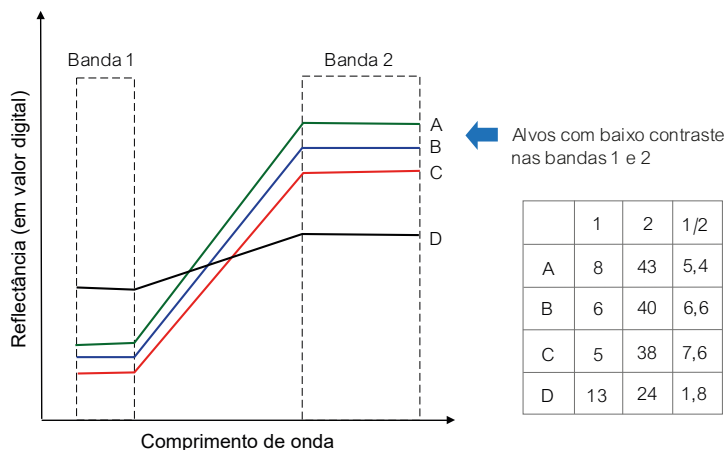


Figura 11.7. Gradiente de diferença de reflectância dos alvos como condição imposta para a divisão da banda 2 pela banda 1.

No exemplo da Figura 11.7, há uma condição estabelecida que é a regra básica para se obter o realce de alvos pela operação de divisão de bandas. Essa regra diz que o alvo deve ter uma boa diferença de reflectância entre as bandas 1 e 2, o que na Figura 11.7 é medida pela inclinação da reta (coeficiente a da reta) entre os valores de reflectância da banda 1 com a banda 2. Observe que os quatros alvos obedecem a essa regra e que quanto maior for a diferença de reflectância entre os alvos das duas bandas, maior é a inclinação da reta. Sabendo-se que a equação da reta tem a forma $y = ax + b$ em que a é o coeficiente de inclinação da reta e x são os valores da variável, computacionalmente, os valores digitais da imagem divisão que correspondem ao x da equação da reta são multiplicados pelo valor do coeficiente da reta, o que faz com que se expanda os valores digitais resultados da divisão para todo o intervalo dinâmico da imagem. Ao se adicionar o valor escalar (coeficiente b na equação da reta) aos números da

divisão, desloca-se o histograma da imagem divisão para o centro do intervalo digital dinâmico. Por esse processo, o resultado da razão é a obtenção de um grande realce dos alvos. Os algoritmos de divisão calculam automaticamente os valores dos coeficientes a e b . Numericamente, os valores da divisão mostrados na tabela inserida na Figura 11.7 passariam a ter os valores listados na Tabela 11.1.

Tabela 11.1. Comparativo entre divisões de bandas com e sem realce utilizando equação linear.

Alvos	Divisão 2/1	$a = 15 \quad b = 7$
A	5,4	88
B	6,6	106
C	7,6	121
D	1,8	34

Observe, nessa tabela, que os valores digitais da divisão, ao serem multiplicados pelo valor escalar 15 do coeficiente a e somados por 7 pelo coeficiente b , apresentam diferenças numéricas significativas, ou seja, visualmente, na tela do computador, seria notado um grande realce entre os alvos. Na prática, para encontrarmos as bandas que satisfaçam essa condição, o intérprete deve primeiro identificar em qual banda do sensor o alvo tem a maior absorção e, a seguir, a banda em que o alvo apresenta a mais alta reflectância. Com isso, o analista encontrará as duas bandas com a maior declividade entre os valores de reflectância. Embora a divisão seja definida em termos de reflectância dos alvos, ela é largamente aplicada em imagens com valores de números digitais não calibrados. É regra dividir a banda de maior reflectância pela banda de menor reflectância para se ter uma imagem de saída em tons de cinza mais claros.

É bastante vantajoso o uso das imagens razão para se obter composições coloridas a partir da combinação de três imagens distintas e resultantes da divisão monocromática. Essas composições têm a vantagem de combinar informações destacadas de várias bandas, facilitando a inter-

pretação das variações espectrais sutis dos alvos, além de ser uma forma de reduzir a dimensão de dados para análise. Três divisões envolvem seis bandas.

A técnica de divisão de bandas deu origem a uma série de algoritmos de processamentos, denominados de índices radiométricos, tais como o índice de vegetação, índice de minerais, índice de água e índice de queimadas. O analista pode criar livremente o índice de seu interesse, desde que tenha boas referências do comportamento espectral do alvo. Todos os índices seguem o mesmo princípio que foi mostrado na Figura 11.7, como veremos nas seções seguintes.

11.2. Índice de Vegetação

O exemplo mais conhecido e didático que ilustra os princípios da divisão de bandas é o índice de vegetação. Na Figura 11.8, está plotada a curva de reflectância da vegetação para demonstrar, na prática, como um índice deve ser criado. Os dois outros espectros de reflectância são de dois alvos sem cobertura vegetal, água e solo exposto, que foram colocados na figura para demonstrar as diferenças de suas divisões com a da vegetação. A principal banda de absorção da vegetação está centrada em $0,65\ \mu\text{m}$, região visível da luz vermelha, e que nas imagens do sensor OLI do Landsat 8 equivale à banda 4, situada entre $0,636\ \mu\text{m}$ e $0,673\ \mu\text{m}$, chamada de banda R do vermelho. A região de maior reflectância da vegetação é no NIR, onde se situa a banda 5 do sensor OLI, nos comprimentos de onda de $0,851\ \mu\text{m}$ a $0,879\ \mu\text{m}$.

Entre essas duas bandas, temos a maior declividade de reflectância da vegetação. Fazendo a divisão da banda 5 pela banda 4, os pixels situados nas áreas com vegetação resultarão em valores bem maiores do que 1. Entretanto, os pixels da água e do solo exposto resultarão em valores próximos a 1, porque, para ambos, não há praticamente nenhuma diferença de declividade em suas reflectâncias entre as duas bandas. Essa é outra regra das imagens divisão: alvos com diferentes radiâncias absolutas, mas

com declividades de curvas espectrais de reflectância similares aparecem com valores de pixels aproximadamente iguais e baixos. Aplicando-se o escalonamento aos valores digitais da razão, tal como vimos no exemplo da Figura 11.7, os pixels da vegetação se transformam em altos valores digitais, aparecendo na imagem divisão em tons muito claros, enquanto os outros dois alvos se transformam em valores digitais praticamente iguais e mais baixos do que os da vegetação, pois possuem valores de razão iguais a 1 (Figura 11.9). Eles surgem na imagem como alvos escuros, indicando serem áreas sem vegetação. Com isso, foi criado um índice para destacar as áreas de vegetação das áreas sem cobertura de vegetação.

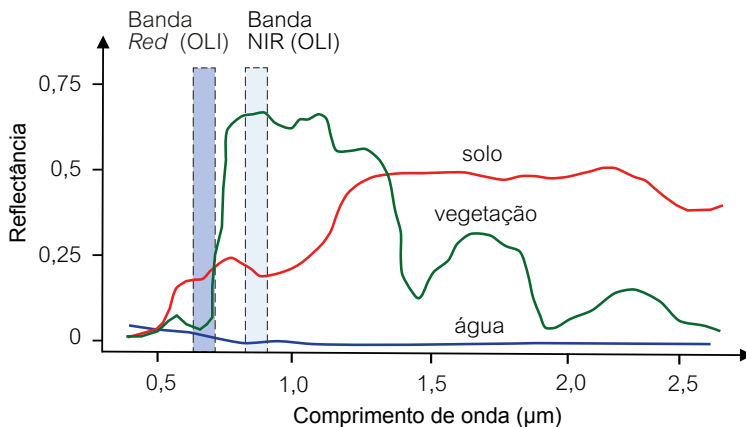


Figura 11.8. Seleção de bandas para definição do índice de vegetação a partir das curvas espectrais de reflectância.

Essa divisão acima exposta é referida como *Ratio Vegetation Index* (RVI) e foi proposto por Jordan (1969), o primeiro a ser estabelecido. Ele consiste na divisão da banda situada no NIR, região que a vegetação mostra a mais intensa reflectância, pela banda situada no vermelho, onde a vegetação possui a mais alta absorção da luz solar visível devido à presença, em sua constituição, do pigmento verde de clorofila, indicativo de uma vegetação

sadia, tal como mostrado na Figura 11.9. A forma matemática é bem simples, como mostra a Equação 11.8.

$$RVI = \frac{NIR}{RED} \quad (11.8)$$

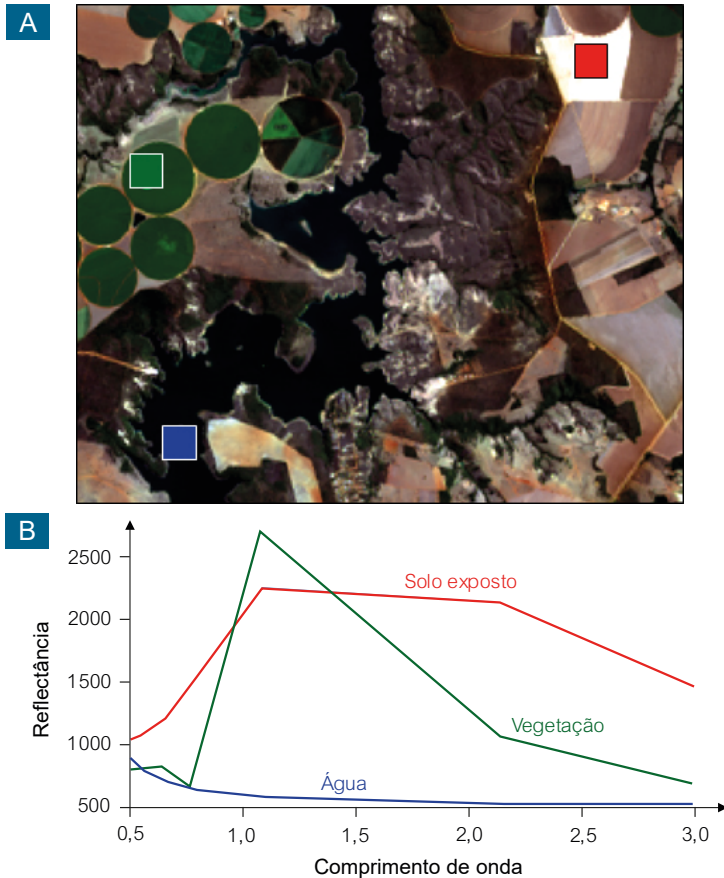


Figura 11.9. Recorte de imagem do satélite Landsat 8 OLI identificando áreas cobertas com solo exposto, vegetação verde e massa de água (A) e suas respectivas curvas espectrais nas bandas do verde, vermelho, infravermelho próximo e infravermelho de ondas curvas.

Os valores absolutos de RVI computados são sempre maiores que zero e assumem valores positivos altos em áreas de densa cobertura vegetal. O escalonamento, como se comentou, é necessário para converter os valores da divisão para números inteiros dentro do intervalo numérico da imagem (por exemplo, 0 – 255), a fim de expor visualmente a imagem ao intérprete. O resultado é uma imagem monocromática, com a gradação de tons de cinza indicando a variação da biomassa ou o índice de área foliar. Quanto mais claro o tom de cinza, maior a densidade de vegetação. Tons de cinzas médios indicam pouca vegetação e tons de cinza totalmente escuros indicam ausência de vegetação, o que só pode ocorrer nas áreas de corpos de água, solos expostos e áreas cobertas por neve ou nuvens, dentre outros.

Um índice similar, um pouco mais complexo, é o NDVI, proposto por Rouse Jr. et al. (1974), que envolve a diferença e a soma entre as duas bandas do infravermelho próximo e do vermelho, segundo a razão (Equação 11.9):

$$NDVI = \frac{NIR - RED}{NIR + RED} \quad (11.9)$$

Os valores de NDVI computados variam de -1 a +1, em que valores positivos são de áreas vegetadas e os valores zero ou negativos são de áreas não vegetadas. A vantagem do NDVI sobre o RVI é que ele tende a ser linearmente mais proporcional à biomassa. O NDVI também é mais apropriado quando se pretende fazer comparações ao longo do tempo de uma mesma área, pois é esperado que seja menos influenciado pelas variações das condições atmosféricas. Trata-se de um índice preferido ao RVI e é amplamente utilizado em escalas continental ou global, até de forma comercial, com satélites como o *Satellite pour l'Observation de la Terre* (SPOT). Inúmeros trabalhos relacionaram NDVI com diversos aspectos da vegetação, tais como índice de área foliar, porcentagem de cobertura do solo e

radiação fotossinteticamente ativa, que são usados em vários modelos para estudos de fotossíntese e sequestro de carbono.

Apesar da grande aceitação do NDVI, Huete e Jackson (1987) constataram que esse índice é um bom indicador de biomassa de vegetação se o terreno tiver uma cobertura esparsa de vegetação, como em regiões semiáridas ou áridas. Esses autores propuseram então um novo índice denominado de índice de vegetação ajustado para solo (SAVI), que tem um melhor desempenho para áreas com baixa cobertura vegetal, ou seja, com presença natural de solos expostos. Sua formulação é dada por (Equação 11.10):

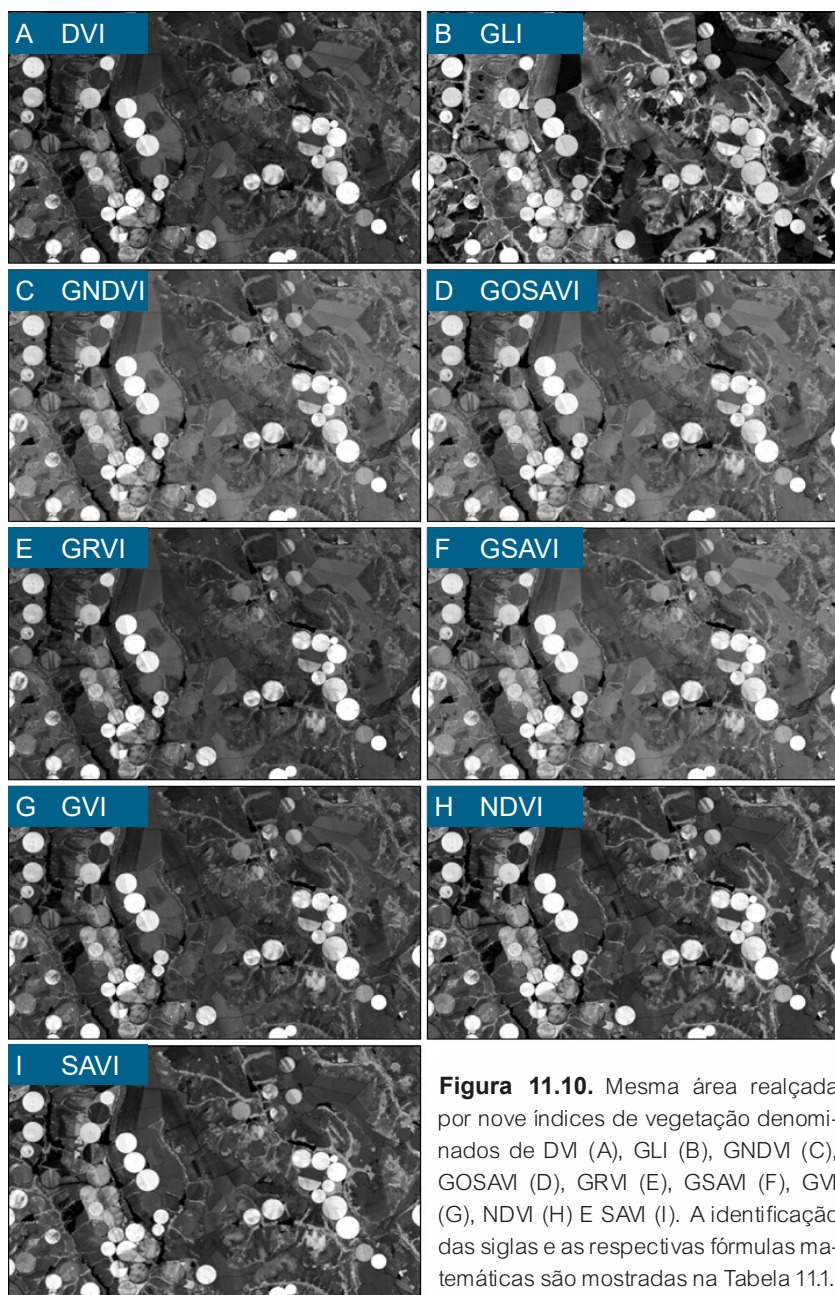
$$SAVI = \left[\frac{\rho N - NIR - \rho RED}{\rho NIR + \rho RED + L} \right] (1 + L) \quad (11.10)$$

O parâmetro L é uma constante empiricamente determinada para minimizar a sensibilidade do índice de vegetação às variações de reflectâncias dos solos. Como a primeira parte da equação SAVI é idêntica ao índice NDVI, se L for igual a zero, o SAVI é igual ao NDVI. Para áreas com cobertura vegetal intermediária, L é tipicamente em torno de 0,5. O fator $(1 + L)$ assegura que o intervalo de valores de SAVI seja o mesmo que o do NDVI, isso é, entre $[-1, +1]$.

A Tabela 11.2 reúne os índices de vegetação mais conhecidos, enquanto na Figura 11.10, é mostrada, comparativamente, os resultados de vários índices de vegetação. Extensões das aplicações dos índices de vegetação são utilizadas para avaliar a senescência da vegetação por meio do *Plant Senescence Reflectance Index* (PSRI) e *Cellulose Absorption Index* (CAI), que servem para investigar a disponibilidade de biomassa susceptível a queimadas.

Tabela 11.2. Índices de vegetação para os sensores TM, ETM+ e OLI usando operações aritméticas.

Nome do índice	Sigla	Formulação matemática	Citação
Difference Vegetation Index	DVI	$DVI = NIR - Red$	Tucker (1979)
Green Leaf Index	GLI	$GLI = \frac{(Green - Red) + (Green + Red)}{(2 \times Green) + Red + Blue}$	Louhaichi et al. (2001)
Green Normalized Difference Vegetation Index	GNDVI	$GNDVI = \frac{NIR - Green}{NIR + Green}$	Gitelson e Merzlyak (1998)
Green Optimized Soil Adjusted Vegetation Index	GOSAVI	$GOSAVI = \frac{NIR - Green}{NIR + Green + 0,16}$	Sripada et al. (2006)
Green Ratio Vegetation Index	GRVI	$GRVI = \frac{NIR}{Green}$	Sripada et al. (2006)
Green Soil Adjusted Vegetation Index	GSAMI	$GSAMI = 1,5 \times \frac{NIR - Green}{NIR + Green + 0,5}$	Sripada et al. (2006)
Green Vegetation Index	GVI	$GVI = (-0,2848 \times Blue) - (0,2435 \times Green) - (0,5436 \times Red) + (0,7243 \times NIR) + (0,0840 \times SWIR1) - (0,18 \times SWIR2)$	Kauth e Thomas (1976)
Normalized Difference Vegetation Index	NDVI	$NDVI = \frac{NIR - Red}{NIR + Red}$	Rouse et al. (1974)
Soil Adjusted Vegetation Index	SAVI	$SAVI = 1,5 \times \frac{NIR - Red}{NIR + Red + 0,5}$	Huete (1988)



Com essa mesma concepção do índice de vegetação, é possível modelar ainda índices para água, minerais, solos, construções e queimadas. A dificuldade associada aos índices voltados para realçar minerais e os solos, é que eles apresentam uma grande diversidade de composição química e, conseqüentemente, uma grande variedade de padrões de reflectância. Nesse sentido, um índice deve ser idealizado para cada tipo de mineral ou de solo. Conseqüentemente, a exigência de o sensor possuir muitas bandas para se ter uma possibilidade efetiva de criar esses índices torna-se maior. Uma rápida apresentação sobre índices desses materiais é mostrada nos itens a seguir.

11.3. Índice de Áreas Construídas

O monitoramento da distribuição espacial e do crescimento de áreas urbanas é de interesse para aplicações de planejamento da ocupação do solo, que tem finalidades para previsões não apenas demográficas, mas inúmeras outras, como escoamento de águas pluviais ou de bacias hidrográficas, controles ambientais ou transporte e tráfico de veículos. A complexidade do arranjo espacial das classes de objetos que compõem áreas urbanas e das suas assinaturas espectrais são somente resolvidas com imagens de alta resolução espacial e modelos de classificação de imagens orientadas a objetos. Porém, para se realizar diagnósticos rápidos que informem sobre a dimensão da área do espaço urbano construído e estimativas de taxa de expansão, foram criados índices de áreas construídas utilizando-se de imagens multiespectrais de resolução espacial mais pobre, como as dos sensores ETM+ ou OLI dos satélites Landsat 7 e Landsat 8, ou equivalentes.

Zha et al. (2003) propuseram o índice *Normalized Difference Built-up Index* (NDBI) para identificar áreas urbanas construídas com base no incremento da resposta espectral de áreas construídas, entre as bandas do NIR e dos comprimentos de ondas curtas (SWIR). O NDBI possui a seguinte formulação (Equação 11.11).

$$NDBI = B_u - NDVI \quad (11.11)$$

em que

$$B_u = \frac{NIR - SWIR}{NIR + SWIR}$$

Por meio dessas equações, são obtidos valores entre +1 e -1 em que valores positivos referem-se às áreas construídas, enquanto todas as outras classes de cobertura terrestres apresentam valores de 0 a -1. O desempenho desse índice pode ser negativamente afetado pela presença de outros tipos de solos sem vegetação, como solos preparados para cultivos, áreas de empréstimos de materiais para pavimentação, minas ou loteamentos ainda não consolidados. O maior desafio na detecção de área urbana construída utilizando o NDBI consiste em modelar áreas de menores dimensões de ocupação. O NDBI possui alta precisão para áreas de moderada e intensa ocupação. Um exemplo da aplicação do NDBI é mostrado na Figura 11.11.

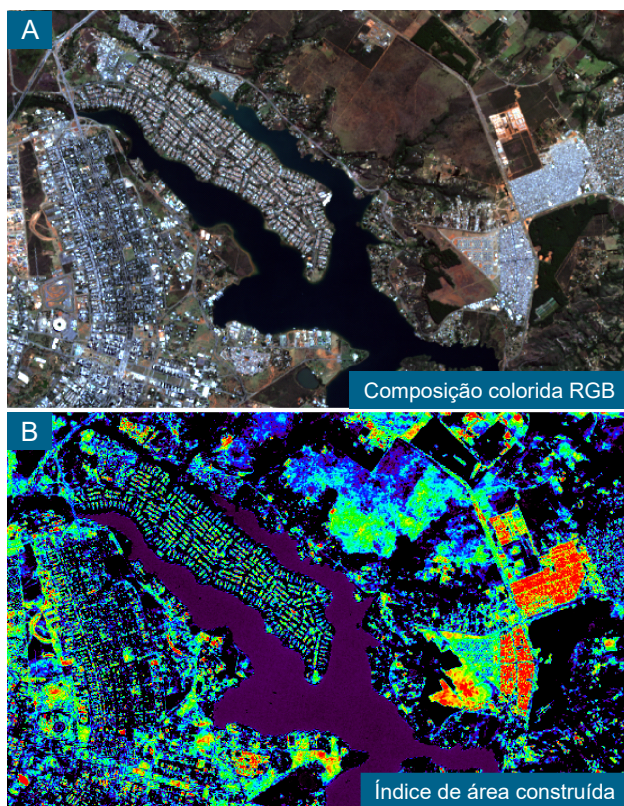


Figura 11.11. Composição colorida RGB real e imagem do índice de área construída (NDBI) em vermelho.

11.4. Índice de Água

Seguindo os mesmos preceitos estipulados para o índice de vegetação, é também possível, com base na curva de reflectância da água, criar um índice de água por diferença normalizada (NDWI). Esse índice requer o uso de uma banda situada na região do comprimento de onda próxima da cor verde ($\pm 0,490 \mu\text{m}$ – $0,580 \mu\text{m}$) e de uma banda situada no comprimento de onda do infravermelho próximo ($\pm 0,700 \mu\text{m}$ – $0,900 \mu\text{m}$), conforme a Equação 11.12, proposta por Gao (1996).

$$B_u = \frac{NIR - SWIR}{NIR + SWIR} \quad (11.12)$$

Na Figura 11.12, são mostradas as curvas de reflectância de corpos de água com diferentes concentrações de partículas de sedimentos em suspensão. Segundo Novo (2010), medidas em campo e experimentos em laboratório comprovam que, com o aumento da concentração de sólidos totais em suspensão na água, o material particulado tem, como principal efeito, o aumento do coeficiente de espalhamento da reflectância da água. A análise das curvas da Figura 11.13 mostra que as curvas de reflectâncias têm uma reflectância ascendente até 0,600 μm (luz verde), que é mais intensa quanto maior é a concentração de sedimentos, e uma queda de reflectância a partir de 0,700 μm que se acentua com a diminuição da concentração de sedimentos. Essa condição configura valores de declividades diferentes para cada nível de concentração de sedimentos, possibilitando, com a aplicação do índice de água (Equação 11.13), um mapeamento das plumas de sedimentos e distinguir água turva de água limpa. Portanto, pela equação que configura esse índice da água, obtém-se um forte realce dos corpos de água, se for usada a banda da luz verde de alta reflectância da água, obtendo, com a banda do infravermelho, a minimização da baixa reflectância da água nesse comprimento de onda, o que também resulta em alto contraste espectral da água com as altas reflectâncias no infravermelho próximo da vegetação e dos solos. Como resultado, a água tem valores digitais positivos, enquanto solo e vegetação têm valor zero ou valores negativos.

Tomando-se imagens do sensor OLI, a aplicação do índice da água pela Equação 11.13 consiste na divisão das bandas do verde (0,533 μm –0,590 μm) pela banda NIR (0,851–0,879 μm). A Figura 11.13, que mostra o encontro das águas do rio Negro com o rio Solimões, próximo à cidade de Manaus, confirma a validade desse índice para diferenciar plumas de sedimentos em suspensão. Nessa imagem divisão, as variações de cores identificam as variações de concentrações de sedimentos em suspensão

nos canais dos dois rios. A cor vermelha indica maior concentração de sedimentos.

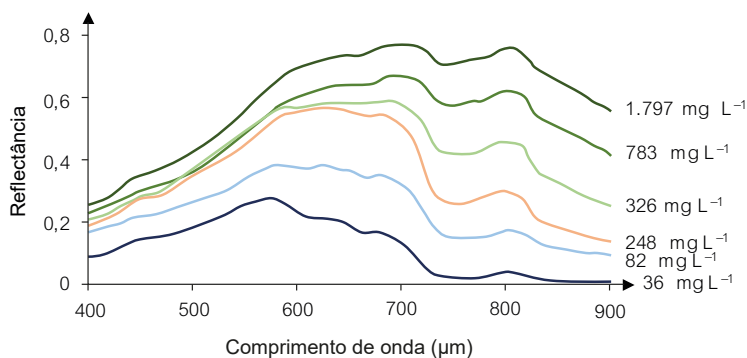


Figura 11.12. Variações na reflectância da água devido às diferenças nas concentrações de partículas sólidas em suspensão.

Fonte: adaptado de Novo (2010).

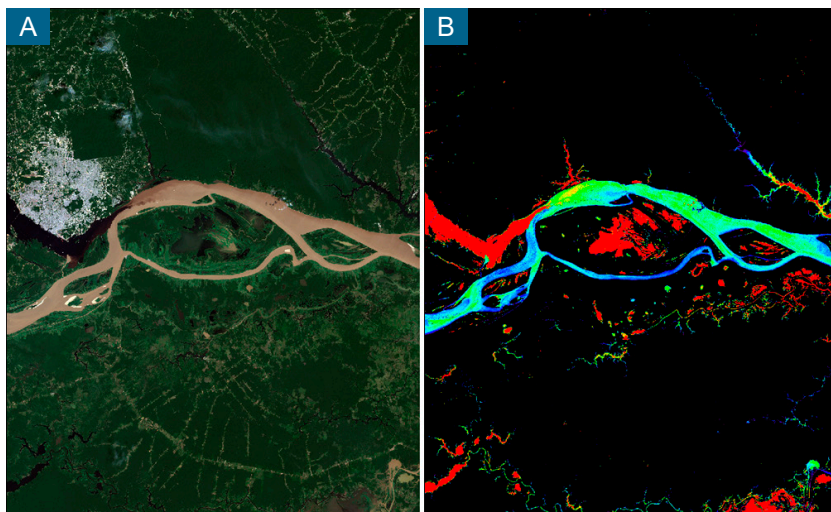


Figura 11.13. Composição colorida das bandas do azul, verde e vermelha do satélite Landsat 8 OLI de 07 de fevereiro de 2018 da região de encontro das águas dos rios Negro e Solimões, em Manaus/AM (A). Em (B), imagem índice de água por diferença normalizada com destaque para as variações nas concentrações de sedimento nas águas dos diferentes rios.

Como a reflectância da água gradativamente diminui em direção aos maiores comprimentos de onda, McFeeters (1996) propôs um novo índice denominado de índice de água por diferença normalizada e modificada (MNDWI) para aumentar o desempenho do índice da água para regiões de águas rasas (Equação 11.14). O cálculo do MNDWI apresenta dois resultados principais: a água tem valores positivos mais altos em relação ao índice NDWI, pelo fato de a reflectância da água ser menor nos comprimentos de onda curtas de 1,50 μm –1,75 μm , região onde se situa a banda SWIR1 do satélite Landsat 8 OLI, em comparação com a banda do infravermelho próximo; e solos e vegetação têm, nesse comprimento de onda, os seus mais altos patamares de reflectâncias e seus valores na imagem divisão são negativos.

11.5. Índice de Queimadas

Para detectar áreas que foram queimadas em imagens pós-incêndio, Key e Benson (1999) propuseram o *Normalized Burn Ratio* (NBR) (Equação 11.13), que consiste em uma razão normalizada entre as bandas do infravermelho próximo (0,76 μm –1,20 μm) e as bandas do infravermelho de ondas curtas (1,20 μm –2,50 μm), mais especificamente, as bandas 5 e 7 (SWIR1) do satélite Landsat 8 OLI. A Figura 11.14 mostra as curvas espectrais de uma queimada e a declividade da reflectância que ocorre entre o pico de reflectância de uma vegetação verde com a da área de queimada.

$$NBR = \frac{NIR - SWIR}{NIR + SWIR} \quad (11.13)$$

Para determinar a severidade de queimadas, utiliza-se o *Differenced Normalized Burn Ratio* (dNBR) que envolve o uso de dois índices NBR antes e depois do evento de fogo (Roy et al., 2006), expresso pela Equação 11.14.

$$dNBR = NBR_{pré-fogo} - NBR_{pós-fogo} \quad (11.14)$$

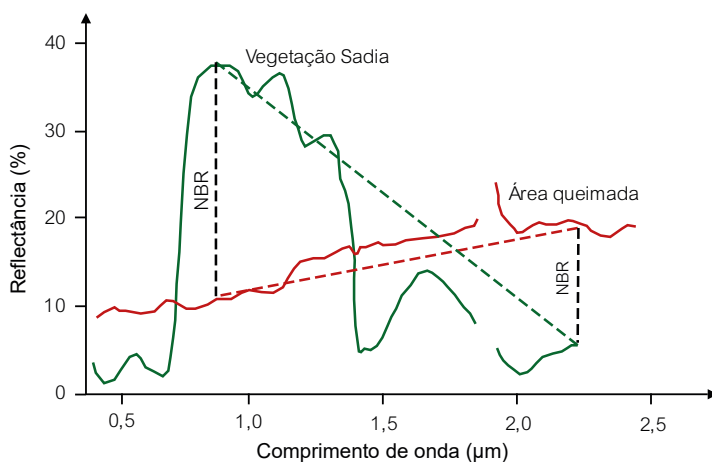


Figura 11.14. Comportamento espectral da vegetação sadia e de uma área queimada a partir de dados do sensor AVIRIS.

Fonte: modificado de Pessoa et al. (2015).

Chuvieco et al. (2002) também propuseram um novo índice denominado de *Burned Area Index* (BAI) que tem a seguinte formulação (Equação 11.15):

$$dBAI = \frac{1}{(0,1 - Red)^2 + (0,06 - NIR)^2} \quad (11.15)$$

Desde que os índices de queimada são designados para enfatizar sinais de matéria vegetal queimada (carvão), eles são altamente dependentes da permanência de carvão após os focos de fogo. A Figura 11.15 ilustra o resultado da aplicação desses índices de queimadas em uma área do cerrado.

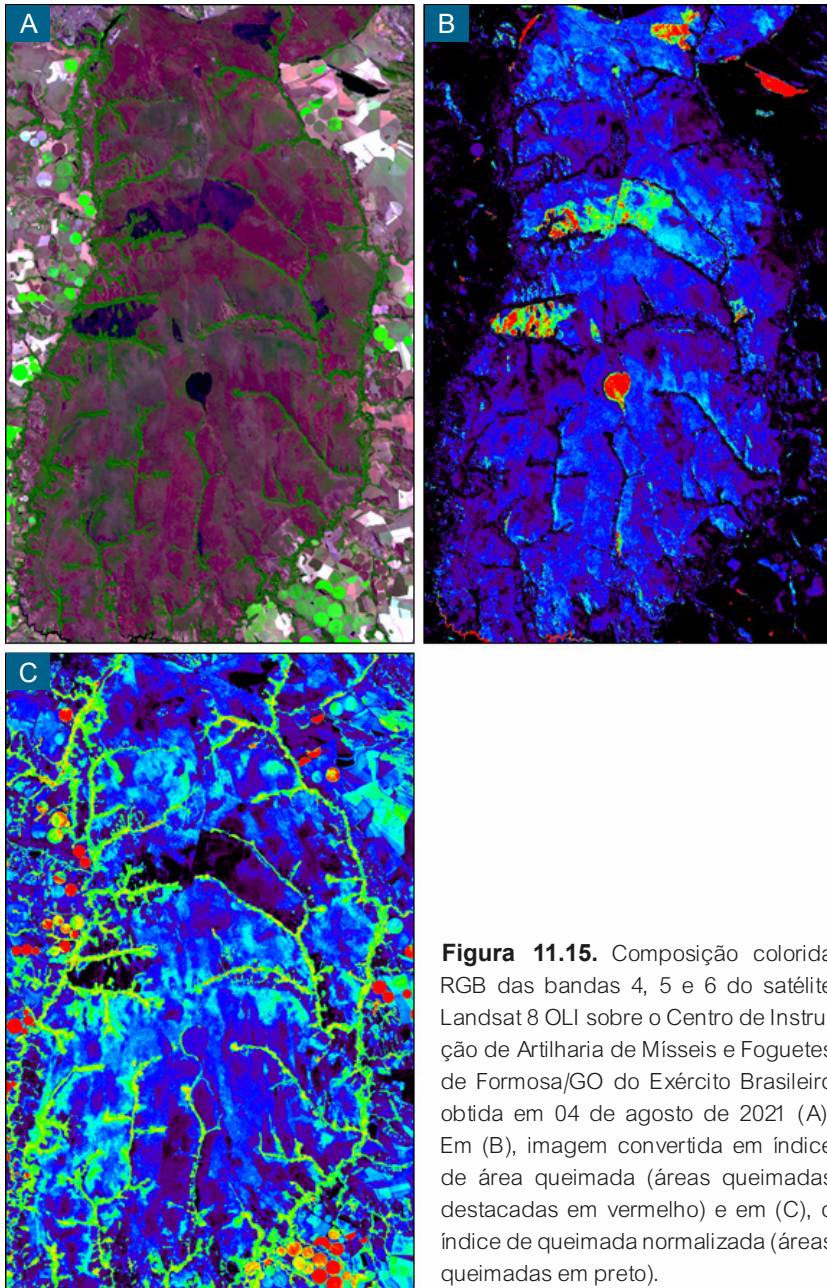


Figura 11.15. Composição colorida RGB das bandas 4, 5 e 6 do satélite Landsat 8 OLI sobre o Centro de Instrução de Artilharia de Mísseis e Foguetes de Formosa/GO do Exército Brasileiro obtida em 04 de agosto de 2021 (A). Em (B), imagem convertida em índice de área queimada (áreas queimadas destacadas em vermelho) e em (C), o índice de queimada normalizada (áreas queimadas em preto).

11.6. Índice de Minerais e de Rochas

Nas aplicações geológicas, o uso da técnica divisão de bandas também pode se constituir em uma excelente opção para prospecção de alvos geológicos com interesses de pesquisa mineral. Devido ao fato de os minerais de rochas serem os materiais que apresentam a maior diversidade de bandas de absorção por causa de sua enorme variedade composicional, teoricamente, seria possível criar índices para vários tipos de minerais. Entretanto, para isso, o sensor multiespectral precisaria satisfazer pelo menos dois requisitos. Primeiro, possuir bandas espectrais posicionadas em torno das principais bandas de absorção que caracterizam a composição dos materiais geológicos, o que implicaria em um sensor com muitas bandas espectrais. Infelizmente, os atuais sensores multiespectrais em operação têm poucas bandas que estão situadas nos comprimentos de onda favoráveis à geologia. Segundo, como as bandas de absorção de minerais normalmente têm larguras muito estreitas, as bandas dos sensores multiespectrais precisariam ter larguras espectrais em torno de 10 nm, o que é bem mais estreito que a largura das bandas dos atuais sensores multiespectrais.

O sensor multiespectral OLI do satélite Landsat 8 serve para ilustrar o referido problema. A banda SWIR2, no infravermelho de ondas curtas (2,11 μm –2,29 μm) foi incluída para ser uma banda geológica para a identificação de áreas com minerais de alteração hidrotermal, pois esses minerais tipicamente apresentam as suas bandas de absorção na região espectral do infravermelho de ondas curtas, como são exemplos os espectros de alguns minerais mostrados na Figura 11.16. Porém, os resultados não se mostraram muito eficientes para este propósito, uma vez que a largura da banda SWIR-2 (banda 8) é muito ampla (180 nm), enquanto, nesse intervalo espectral, as bandas de absorção dos minerais de alteração hidrotermal possuem de 10 a 20 nm de largura, o que torna difícil o uso dessa banda para identificar, com segurança, se há ou não, a presença de halos de alteração hidrotermal em uma determinada área de estudo.

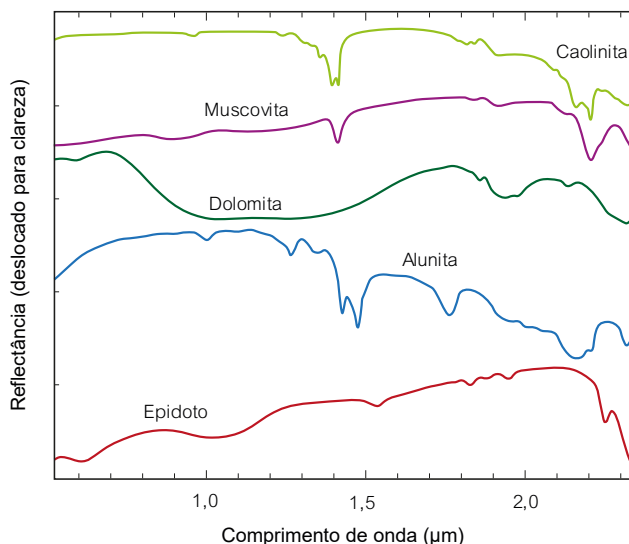


Figura 11.16. Exemplos de espectros de minerais de alteração hidrotermal com intensas bandas de absorção posicionadas na região espectral do infravermelho (2,0 μm a 2,5 μm).

Outra opção seria o uso do sensor ASTER, projetado para aplicações geológicas, que possui seis bandas na região espectral da banda do SWIR2 do satélite Landsat 8 OLI. Com o sensor ASTER, razões de bandas para identificação de halos de alteração hidrotermal podem ser feitas com a banda 4 (1,6 μm –1,7 μm), tipicamente onde os materiais geológicos não alterados exibem altas reflectâncias, e uma das cinco bandas existentes para a identificação de minerais de alteração hidrotermal: bandas 5 (2,145 μm –2,185 μm), 6 (2,185 μm –2,225 μm), 7 (2,235 μm –2,285 μm), 8 (2,295 μm –2,365 μm) e 9 (2,365 μm –2,430 μm). A Tabela 11.3 mostra uma série de índices de minerais, específicos para as bandas do sensor ASTER, que também inclui índices com as bandas do termal deste sensor. Se as imagens ASTER fossem utilizadas para obter uma discriminação de um halo hidrotermal que contivessem os minerais da Figura 11.16 e outros, as razões de bandas da Tabela 11.3 poderiam ser propostas (incluem as bandas 10 a 14 do módulo termal).

Tabela 11.3. Índices minerais para o sensor ASTER usando razões de bandas.

Feições	Bandas ou razões	Referências
Ferro		
Ferro férrico	2/1	Rowan; CSIRO
Ferro ferroso	5/3 + 1/2	Rowan
Laterita	4/5	Bierwith
Gossan	4/2	Volesky
Silicatos férricos (biotita, clorita e anfibólio)	5/4	CSIRO
Óxidos de ferro	4/3	CSIRO
Carbonatos e minerais máficos		
Carbonato, clorita, epidoto	(7 + 9)/8	Rowan
Epidoto, clorita, anfibólio	(6 + 9)/(7 + 8)	CSIRO
Anfibólio, MgOH	(6 + 9)/8	Hewson
Anfibólio	6/8	Bierwith
Dolomita	(6 + 8)/7	Rowan; USGS
Carbonato	13/14	Bierwith, Nimoyima, CSIRO
Silicatos		
Sericita, muscovita, illita, esmectita	(5 + 7)/6	Hewson (CSIRO)
Alunita, caolinita, pirofilita	(4 + 6)/5	Rowan (USGS)
Fengita	5/6	Hewson
Muscovita	7/6	Hewson
Caolinita	7/5	Hewson
Argila	(5 + 7)/62	Bierwith
Alteração	4/5	Volesky
Rocha hospedeira	5/6	Volesky
Sílica		
Rochas ricas em quartzo	14/12	Rowan
Sílica	(11x11)/(10x12)	Bierwith
Minerais de baixo grau de metamorfismo (garnierita, clinopiroxênio, epidoto e clorita)	12/13	Bierwith, CSIRO
SiO ₂	13/12	Palomera

Continua...

Tabela 11.3. Continuação.

Feições	Bandas ou razões	Referências
SiO ₂	12/13	Nimoyima
Rochas silicáticas	(11 11)/(10 12)	Nimoyima
Sílica	11/10	CSIRO
Sílica	11/12	CSIRO
Sílica	13/10	CSIRO

11.7. Referências

CRÓSTA, A. P. **Processamento digital de imagens de sensoriamento remoto.**

Campinas: UNICAMP, 1999, 164 p.

CRÓSTA, A. P.; MOORE, J. M. Enhancement of Landsat Thematic Mapper imagery for residual soil mapping in SW Minas Gerais State, Brazil: A prospecting case history in greenstone belt terrain. In: 7th THEMATIC CONFERENCE ON REMOTE SENSING FOR EXPLORATION GEOLOGY, Calgary, Canada. **Proceedings...** Ann Arbor: ERIM, 1989, p. 1173-1187.

DRURY, S. A. **Image Interpretation in Geology.** Berlim, Springer, 7a ed., 1987, 243 p.

FLORENZANO, T. C.; TAVARES JR., S. S.; LORENA, R. B.; MELO, D. H. C. T. B. Multiplicação e adição de imagens Landsat no realce de feições da paisagem. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 10, Foz do Iguaçu, PR. **Anais...** São José dos Campos: INPE, p. 1257-1263, 2001.

GAO, B. C. NDWI: a normalized difference water index for remote sensing of vegetation liquid water from space. **Remote Sensing of Environment**, v. 58, n. 3, p. 257-266, 1996.

GITELSON, A. A.; MERZLYAK, M. N. Remote sensing of chlorophyll concentration in higher plant leaves. **Advances in Space Research**, v. 22, n. 5, p. 689-692, 1998.

HUETE, A. A soil-adjusted vegetation index (SAVI). **Remote Sensing of Environment**, v. 25, p. 295-309, 1988.

HUETE, A. R.; JACKSON, R. D. Suitability of spectral indices for evaluating vegetation characteristics on arid environments. **Remote Sensing of Environment**, v. 23, n. 2, p. 213-232, 1987.

JORDAN, C.F. Derivation of leafarea index from quality of light on the forest floor. **Ecology**, v. 50, p. 663-666, 1969.

KAUTH, R. J.; THOMAS, G. S. The Tasseled Cap – a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. In: LARS SYMPOSIA, **Anais...** West Lafayette, Universidade de Purdue, p. 41-51, 1976.

KEY, C. H.; BENSON, N. Measuring and remote sensing of burn severity: The CBI and NBR. In: JOINT FIRE SCIENCE CONFERENCE AND WORKSHOP, Boise, Universidade de Idaho, EUA. **Anais...** 1999.

LOUHAICHI, M.; BORMAN, M.M.; JOHNSON, D.E. Spatially located platform and aerial photography for documentation of grazing impacts on wheat. **Geocarto International**, v. 16, n. 1, p. 65-70, 2001.

MCFEETERS, S. K. The use of the normalized difference water index (NDWI) in the delineation of open water features. **Remote Sensing Letters**, v. 17, n. 7, p. 1425-1432, 1996.

PESSOA, O. A. A.; BAPTISTA, G. M. M.; ALMEIDA, T.; MENESES, P. R. Monitoramento da refletância espectral após extinção do fogo e uma proposta para a detecção de áreas queimadas em vegetação campestre de Cerrado. **Revista Brasileira de Cartografia**, v. 67, p. 1131-1143, 2015.

ROUSE JR., J. W.; HAAS, R. H.; SCHELL, J. A.; DEERING, D. W. Monitoring vegetation systems in the great plains with ERTS. In: Third ERTS SYMPOSIUM, Greenbelt, MD, USA. **Anais...** Greenbelt: NASA/GSFC, p. 309-317, 1974.

ROY, D. P.; BOSCHETTI, L.; TRIGG, S. N. Remote sensing of fire severity: assessing the performance of the normalized burn ratio. **IEEE Geoscience and Remote Sensing Letters**, v. 3, n. 1, p. 112-116, 2006.

SRIPADA, R. P.; HEINIGER, R. W.; WHITE, J. G.; MEIJER, A. D. Aerial color infrared photography for determining early in-season nitrogen requirements in corn. **Agronomy Journal**, v. 98, p. 968-977, 2006.

TUCKER, C. Red and photographic infrared linear combinations for monitoring vegetation. **Remote Sensing of Environment**, v. 8, p. 127-150, 1979.

ZHA, Y.; GAO, J.; NI, S. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. **International Journal of Remote Sensing**, v. 24, n. 3, p. 583-594, 2003.

CAPÍTULO 12

Rotação Espectral

Paulo Roberto Meneses

Tati de Almeida



A maioria das aplicações de sensoriamento remoto exige o uso de sensores com um número razoável de bandas, visto que parece ser mais importante descobrir as propriedades dos objetos do que simplesmente detectar as suas presenças. Contudo, essa exigência de um maior número de bandas tem um compromisso crítico. Ao visualizarmos as diferentes bandas dos sensores multiespectrais, principalmente bandas espectralmente contíguas, vemos uma alta semelhança entre elas, o que, em termos estatísticos, significa alta correlação. Essa alta correlação de bandas tem, como causa, a semelhança da reflectância entre os objetos presentes na cena. Destaca-se a combinação de três fatores: i) alvos como vegetação, água, solos e rochas possuem reflectâncias muito próximas em alguns comprimentos de onda do visível e do infravermelho; ii) o efeito espectral do sombreamento topográfico é o mesmo em todas as bandas, aparecendo de forma dominante em áreas de relevo acidentado e com baixo ângulo de iluminação solar, aumentando, com isso, a correlação entre as bandas; e iii) a grande largura das faixas espectrais das bandas reduz as diferenças de reflectância entre os alvos. Essa alta correlação dos dados prejudica a análise ou a interpretação das imagens e minimiza as chances de discriminação de classes de alvos. Em processamento de imagens, as técnicas que podem minimizar a redundância entre os pixels de diferentes bandas, reduzindo a alta correlação, executam transformações do espaço espectral dos dados por meio de rotações dos eixos de coordenadas.

12.1. Transformação por Componentes Principais

No Capítulo 9, discutimos o modelo de cores aplicado às imagens para criação de composições coloridas, associando-se as cores primárias RGB a três bandas espectrais. A regra básica para obter uma imagem com alto

contraste de cores é ter, em um sensor, um alto contraste espectral entre as bandas selecionadas. Entretanto, quando a correlação entre as bandas é alta, as possíveis opções de combinações de bandas para conseguir boas composições coloridas são reduzidas. Sendo assim, é lógico pensar que, por causa dessa redundância, a análise de todas as bandas não parece ser necessária. No entanto, quando encontramos uma alta correlação entre bandas, ela nunca atinge 100%, havendo sempre um mínimo de informação não redundante que pode ser relevante preservá-la, pois pode conter informações importantes para o analista, mas que são difíceis de serem observadas nas imagens.

Uma técnica poderosa de processamento de imagens utilizada para eliminar a alta correlação entre as bandas de um sensor e com perda mínima de informação é a transformação por componentes principais (TCP), que tem seus princípios na estatística para análise de um grande número de variáveis. A TCP é largamente usada na Biologia e na Geologia como um processo de redução da dimensão das variáveis a serem analisadas. Aplicadas às imagens de sensoriamento remoto, a grande vantagem dessa técnica é a remoção da correlação existente entre as bandas multiespectrais por meio de uma compressão simultânea da maioria das informações totais presentes em todas as bandas, isto é, concentração da variância total dos dados. É o mesmo que dizer que, de um conjunto de 10 bandas, cerca de 90% a 95% de toda a informação contida nessas bandas são redistribuídas e concentradas em três novas imagens denominadas de componentes principais, sem que haja correlação entre elas. Os outros 5% a 10% de informações remanescentes são distribuídas nos outros sete componentes, que igualmente não mais se correlacionam.

As imagens dos novos componentes principais são relacionadas aos valores de brilho do conjunto das imagens originais via uma transformação linear, os quais possuem dois efeitos práticos: (i) maior realce das feições dos objetos que não eram anteriormente discerníveis; e (ii) concentração de 90% a 95% das informações de todas as bandas originais nas três primei-

ras componentes principais. Essas três componentes podem ser utilizadas para gerar uma composição colorida RGB, reduzindo assim a dimensão de dados a serem analisados. Porém, nessa transformação, as imagens componentes perdem as relações físicas da reflectância dos alvos, o que torna o processo interpretativo um tanto trabalhoso.

Neste capítulo, abordaremos os princípios da transformação de imagens multiespectrais para o novo conjunto de componentes, primeiramente, com a técnica TCP que permite operações com qualquer número de bandas. Uma derivação dessa técnica é apresentada a seguir para um conjunto de apenas três bandas, visando, nesse caso, a obtenção de composições coloridas RGB com imagens componentes descorrelacionadas, mas com a diferença de que não há perda das relações físicas da reflectância das bandas envolvidas. Os dois processos envolvem uma rotação dos eixos dos dados originais, razão pela qual essas técnicas são denominadas de rotação espectral.

O modelo de formato das imagens multiespectrais possibilita representar os valores de brilho dos pixels por suas posições dentro de um espaço vetorial. O número de eixos ou dimensões correspondem ao número de bandas espectrais, à semelhança do espaço cúbico do modelo de cores RGB. No espaço de múltiplas bandas, a medida estatística é multivariável e a medida das variáveis, isto é, dos números digitais (ND) de cada banda se torna uma medida de vetores NDp com K componentes, como mostra a Figura 12.1.

No contexto da TCP, o objetivo da representação do espaço vetorial n -dimensional é determinar o nível de correlação que pode haver entre n bandas, calculando-se o grau de informação ou variância que os dados apresentam para determinar os coeficientes que transformam as imagens em novos componentes. A TCP é derivada da matriz de covariância entre as bandas. Para facilitar o entendimento matemático e geométrico, a discussão que se segue limitar-se-á ao espaço de duas dimensões, pelo fato de que

dimensões maiores só podem ser representadas algebricamente. Mais informações sobre a teoria da TCP podem ser obtidas no livro *Processamento Digital de Imagens de Sensoriamento Remoto*, editado por Crósta (1999).

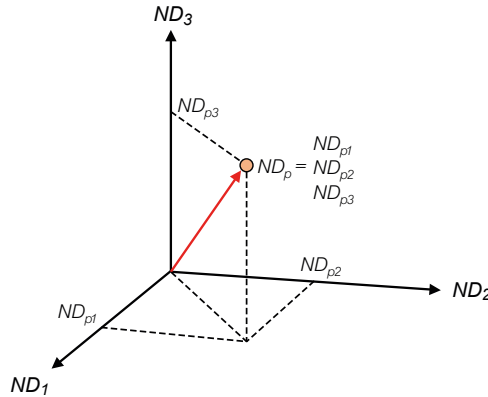


Figura 12.1. Visualização do pixel ND_p de uma imagem de três bandas como um vetor tridimensional.

Fonte: Schowengerdt (2007).

Considere um espaço de duas variáveis correspondentes às bandas x e y de um sensor imageador com um grande número de pixels plotados nesse espaço, como ilustrado na Figura 12.2A. Cada pixel pode ser descrito por seu vetor x_i , enquanto o vetor média m define a média das posições dos pixels nesse espaço. Se os valores dos pixels das duas bandas estiverem muito correlacionados, as medidas simultâneas de x e y formam uma linha aproximadamente reta (eixo AB na Figura 12.2B). Se as medidas não estiverem correlacionadas, os pixels são delimitados pelo espaço definido na forma de uma elipse, com dimensões significativamente diferentes nos seus dois eixos AB e CD, conforme mostrado na Figura 12.2C. Neste caso, haverá uma direção predominante (AB) de variância com um eixo maior em relação ao segundo eixo traçado ortogonalmente (CD), com menor variância. Se esse eixo CD contiver apenas uma pequena proporção do total da

variância dos dados, ele pode ser ignorado sem acarretar muita perda de informação. Isso significaria uma diminuição na dimensão dos dados de dois para um. Dessa maneira, ao representar a distribuição dos pixels na forma de uma elipse, percebe-se que os eixos AB e CD revelam, mais apropriadamente, a estrutura interna que está presente nos dados das imagens, em relação aos eixos x e y separados.

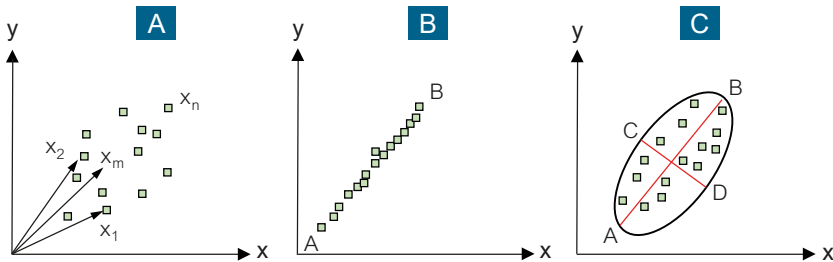


Figura 12.2. Espaço bidimensional da representação dos dados mostrando: A) os vetores x_i dos pixels e o vetor média m ; B) a alta correlação dos dados com a variância em um único eixo AB; e C) a correlação positiva dos dados com os eixos AB e CD medindo a variância dos dados na forma de uma elipse.

Segundo Mather (1987), o exemplo da Figura 12.2 mostra que há uma distinção importante entre o número de variáveis (bandas) e a dimensionalidade do conjunto de dados. Na Figura 12.2B, o número de variáveis é dois, enquanto o da dimensionalidade é um. Na Figura 12.2C, o número de variáveis é dois, mas a dimensionalidade estatística dos dados é efetivamente um, que é o eixo de maior variância, o qual define a orientação da elipse no espaço. O uso de um único eixo AB substitui os eixos separados x e y , com a vantagem de reduzir a dimensionalidade dos dados. Adicionalmente, a informação transmitida pelo eixo AB é maior do que a informação transmitida pelos dois eixos separados. Normalmente, as imagens multiespectrais têm uma dimensionalidade que é menor do que o número de bandas espectrais. A TCP atua nesse sentido de reduzir o conjunto de dados, preservando a informação existente. Com isso, consegue-se minimizar o esforço

de análise de um grande número de variáveis, no caso de sensoriamento remoto, de um grande número de bandas.

Estatisticamente, a forma da elipse como meio de descrever o espalhamento dos pixels no espaço n -dimensional é definida pela matriz de covariância calculada para as n bandas espectrais. Por definição, a matriz de covariância é uma matriz simétrica que mede como duas ou mais variáveis variam conjuntamente. Seus valores são sempre positivos. A covariância é, por vezes, chamada de medida de dependência linear entre duas ou mais variáveis aleatórias. Calculando-se o vetor média e a matriz de covariância, determinam-se a localização e a forma do espalhamento dos pontos no espaço n -dimensional, além de conhecer os valores numéricos da correlação existente entre o conjunto de bandas analisadas.

A matriz de covariância é um dos mais importantes recursos matemáticos usados no processamento de imagens multiespectrais. Os valores da matriz enfatizam propriedades do conjunto das bandas analisadas. No entanto, há uma controvérsia em usar a matriz de covariância ao invés da matriz de coeficientes de correlação para transformações de bandas. Se a matriz de covariância for usada para definir a forma da elipse que contém os pixels em uma direção particular, as medidas de cada variável devem ser comparáveis. Isso não é o que de fato acontece quando se comparam os valores digitais entre bandas diferentes. Por exemplo, um número digital de valor 52 em uma dada banda representa um valor de radiância que fisicamente pode não ser o mesmo valor de radiância para o número 52 em outra banda. Isto porque os detectores de cada banda possuem valores específicos de ganho e *offset* para efeitos de calibração. Sendo assim, são variáveis que não são totalmente comparáveis. Pode-se verificar isso na Figura 12.3, que mostra um subconjunto de pixels de duas bandas em diferentes comprimentos de onda. Na imagem à esquerda, o pixel situado na segunda coluna e na segunda linha possui valor digital igual a 52. Esse mesmo valor digital é encontrado na imagem à direita na segunda coluna e quarta linha. Percebe-se que existe diferença nos tons de cinza nas duas imagens para

o pixel de mesmo valor digital igual a 52. Nessa situação, a matriz de coeficientes de correlação é mais apropriada para medir o grau de correlação entre bandas espectrais. Os coeficientes de correlação são, simplesmente, a covariância medida para variáveis padronizadas e são encontrados nos elementos fora da diagonal da matriz, o que é mostrado adiante.

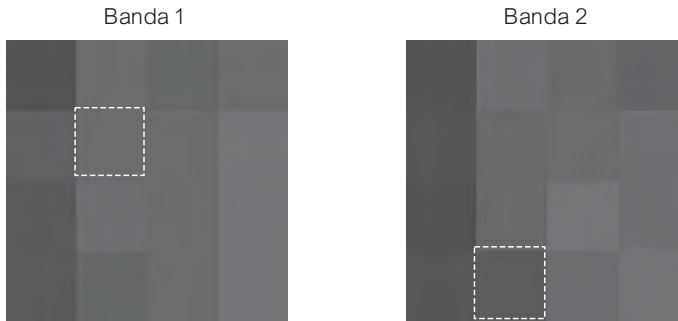


Figura 12.3. Segmento de pixels de duas imagens de diferentes comprimentos de onda mostrando a diferença de nível de cinza de um pixel com valor digital igual a 52 devido às diferenças de calibração dos detectores das duas bandas.

A experiência nos mostra que qualquer sensor multiespectral possui algumas bandas com alta correlação, de modo que os eixos das funções de distribuição de densidade de probabilidade (FDP) dos NDs não são estatisticamente ortogonais, como se vê na Figura 12.4A entre duas bandas (eixos ND_1 e ND_2). No contexto multivariado, a forma da elipse que encerra o espalhamento dos dados no espaço p -dimensional é definida pela matriz de covariância computada das n bandas espectrais. A TCP envolve uma rotação linear dos eixos ou dos dados em um espaço hipotético de atributos do sistema de coordenadas da FDP, produzindo novas variáveis conhecidas como componentes principais, ou eixos principais, que são combinações lineares das variáveis originais (bandas originais).

A rotação das coordenadas dos vetores NDp (Figura 12.4B) corresponde a um esforço para ortogonalizar os novos eixos CP_1 e CP_2 , fazendo-os

coincidirem ao máximo com as direções de distribuição dos dados, ou suas variâncias, como mostra a orientação da elipse na Figura 12.4B. Na TCP, isto é, para um novo sistema de eixos ou componentes ortogonais rotacionados, a distribuição dos pixels apresenta, no primeiro componente (CP_1) ou no primeiro eixo principal (maior eixo da elipse na Figura 12.4C), a maior variância, a qual fica orientada na direção do primeiro componente CP_1 . Na direção do eixo do segundo componente principal CP_2 , a variância é menor (direção do segundo eixo da elipse). Desde que se assume que a variância é uma medida do conteúdo de informação da imagem, o primeiro componente ou o primeiro eixo agrega uma maior quantidade de informação espectral, produzindo a imagem CP_1 com maior quantidade de informação, enquanto a imagem CP_2 reúne pouca quantidade de informação. Em termos percentuais, seria equivalente dizer que o CP_1 contém 95% dos dados das imagens originais, enquanto o CP_2 contém apenas 5%. Nesse novo espaço de eixos rotacionados, a imagem do primeiro componente não tem correlação com a imagem do segundo componente. Situações similares podem ser imaginadas em espaços de n dimensões, ou n bandas, sabendo-se, contudo, que todo o processo de transformação da imagem é realizado em uma concepção puramente matemática.

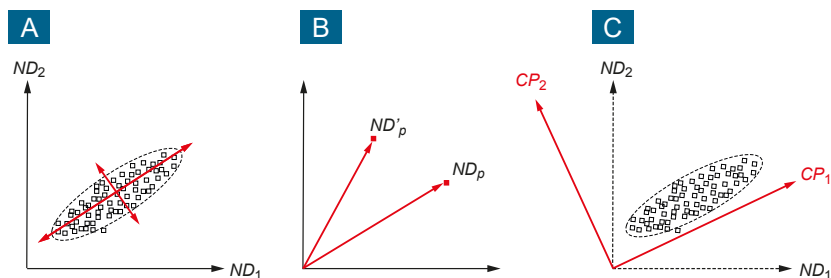


Figura 12.4. Transformação de duas imagens com alta correlação de pixels (A) por rotação espectral dos vetores ND_p (B), para novos eixos ortogonalizados CP descorrelacionados, com as variâncias orientadas aos eixos CP_1 e CP_2 .

A transformação mostrada na Figura 12.4 é uma transformação f do vetor ND da imagem para o novo vetor rotacionado ND' . Uma transformação linear tem a seguinte expressão (Equação 12.1) (Schowengerdt, 2007):

$$ND' = f(ND) \quad (12.1)$$

Se f é uma transformação linear, tem-se que (Equações 12.2 e 12.3):

$$ND' = W \times ND + B \quad (12.2)$$

$$ND' = \begin{bmatrix} w_{11} & \dots & w_{1k} \\ \dots & \dots & \dots \\ w_{k1} & \dots & w_{kk} \end{bmatrix} \times \begin{bmatrix} ND_1 \\ \dots \\ ND_k \end{bmatrix} + \begin{bmatrix} B_1 \\ \dots \\ B_k \end{bmatrix} \quad (12.3)$$

em que

W = matriz de pesos aplicada às bandas originais; e

B = vetor que move a origem do espaço dos ND.

Na TCP, a matriz W descreve operações como rotação e mudança de escala e o vetor adicional B descreve um deslocamento espacial (Equação 12.4).

$$CP = W_{CP} \times ND \quad (12.4)$$

Em síntese, a TCP tem, como objetivo principal, eliminar a correlação entre as bandas, utilizando a matriz de covariância ou a matriz de coeficientes de correlação das imagens para determinar um conjunto de coeficientes chamados de autovalores e autovetores. Os autovalores representam o comprimento dos eixos dos componentes principais da imagem CP, que são medidos em unidades de variância, enquanto os autovetores deter-

minam as direções dos eixos dos componentes principais rotacionados. Eliminar a alta correlação das imagens tem a vantagem de reduzir o volume de dados a ser analisado e de redistribuir a informação espectral das imagens originais entre os novos componentes principais.

O cálculo da TCP é extenso e relativamente complexo. No livro *Statistics and Data Analysis in Geology*, editado por Davis (2002), o leitor encontrará a base do desenvolvimento estatístico, com exemplos de aplicações envolvendo variáveis geológicas. Um exemplo prático é mostrado a seguir com as sete bandas do sensor *Operational Land Imager* (OLI) do satélite Landsat 8 que opera na faixa óptica do espectro eletromagnético. Inicialmente, são computadas as matrizes de coeficiente de correlação ou de covariância das sete bandas, conforme mostradas na Tabela 12.1, mas cabe, ao usuário, decidir se todas as bandas serão utilizadas para os cálculos dos componentes principais. As matrizes são simétricas em relação à diagonal principal, por isso, é desnecessário repetir os mesmos valores acima da diagonal. Os coeficientes de correlação variam de -1 a +1. Os altos valores positivos indicam que se tem uma correlação positiva alta entre as bandas e a correlação diminui quando o coeficiente se aproxima de zero. O sinal negativo indica que há correlação inversa entre as bandas, ou seja, quando um pixel tem um alto valor digital em uma banda, na outra banda ele tem um baixo valor. Valores negativos altos significam alta correlação negativa. Coeficientes negativos normalmente ocorrem em áreas com presença de cobertura vegetal verde na imagem selecionada. Na banda 5 (infravermelho próximo) do sensor OLI, essas áreas apresentam altos valores de brilho, enquanto nas outras cinco bandas, os valores são mais baixos. A Tabela 12.2 corresponde à matriz de covariância cujos valores medem quanto as bandas variam conjuntamente. É também chamada de matriz de variância-covariância porque a diagonal é a medida da variância de cada banda. Note, nesse exemplo, que a banda 5 é a banda que tem a maior variância (5438108,2), ou seja, é a que contém a maior quantidade de informação e o maior contraste espectral.

Tabela 12.1. Matriz de coeficientes de correlação computada de sete bandas do sensor OLI do satélite Landsat 8.

Correlação	Banda 1	Banda 2	Banda 3	Banda 4	Banda 5	Banda 6	Banda 7
Banda 1	1,00000						
Banda 2	0,98539	1,00000					
Banda 3	0,96009	0,98099	1,00000				
Banda 4	0,96098	0,98145	0,97807	1,00000			
Banda 5	0,44782	0,48049	0,56009	0,44821	1,00000		
Banda 6	0,88814	0,91858	0,91677	0,92701	0,55778	1,00000	
Banda 7	0,92947	0,94824	0,93140	0,95868	0,44464	0,98050	1,00000

Tabela 12.2. Matriz de covariância computada de sete bandas do sensor OLI do satélite Landsat 8.

Covariância	Banda 1	Banda 2	Banda 3	Banda 4	Banda 5	Banda 6	Banda 7
Banda 1	75547,0						
Banda 2	103465,2	145931,4					
Banda 3	188416,3	267568,5	509789,8				
Banda 4	325635,4	462222,2	860936,7	1519882,1			
Banda 5	190090,0	283466,6	617586,7	853366,52	2384986,9		
Banda 6	569269,1	818311,1	1526445,0	2665106,5	2008772,5	5438108,2	
Banda 7	431508,1	611841,7	1123254,5	1996284,2	1159852,5	3862045,1	2852898,6

O primeiro passo na TCP consiste em um cálculo algébrico linear que altera a matriz de coeficientes ou de covariância (o analista deve decidir qual usar), determinando um conjunto de quantidades denominado de autovalores (*eigenvalues*). Esse é um processo bastante complexo que não entraremos em discussão. Os autovalores são medidos em termos de unidade de variância na diagonal da matriz da Tabela 12.3. Como a variância é uma das medidas da quantidade de informação presente na imagem, os autovalores representam o tamanho dos eixos do “elipsoide” de sete eixos, mas não a direção do elipsoide. Por convenção, os autovalores são arranjados em ordem decrescente de variância. A variância total das sete imagens componentes principais mantém a variância total das sete imagens originais, apenas foram redistribuídas de tal forma que a maior variância está localizada no primeiro componente principal (11001406), com valores sucessivamente menores para os componentes de ordem maior (CP7 = 920). Fora da diagonal, os valores nulos indicam que não há qualquer correlação entre os componentes, que é esse o objetivo principal da transformação por componentes principais, o de descorrelacionar as imagens. O número de imagens componentes principais sempre será igual ao número de bandas originais.

Tabela 12.3. Matriz de autovalores (*eigenvalues*) calculados pela transformação dos coeficientes de correlação.

	CP1	CP2	CP3	CP4	CP5	CP6	CP7
CP1	11001406						
CP2	0	1669716					
CP3	0	0	216127				
CP4	0	0	0	23547			
CP5	0	0	0	0	10808		
CP6	0	0	0	0	0	4617	
CP7	0	0	0	0	0	0	0920

Para facilitar a percepção do impacto dessa transformação, pode-se apresentar os autovalores, isto é, as variâncias, em porcentagens. Para cal-

cular a porcentagem de variância de cada imagem principal componente, basta dividir a variância total dos sete componentes pela variância de cada componente. Seus valores são mostrados na Tabela 12.4.

Tabela 12.4. Autovalores em porcentagens.

	Componentes principais						
	CP1	CP2	CP3	CP4	CP5	CP6	CP7
Porcentagem da variância (%)	85,103	12,916	1,6718	0,182	0,0836	0,0357	0,0071
Porcentagem acumulada da variância (%)	85,1031	98,0195	99,691	99,8735	99,9571	99,9928	100

Observando-se a porcentagem da variância acumulada da Tabela 12.4, verifica-se que os três primeiros componentes somam 99,69%, o que se pode dizer que é praticamente a variância total do conjunto das sete bandas originais. O restante das variâncias dos CP4, CP5, CP6 e CP7, que é menor que 1%, é muito pouco significativo em termos de conteúdo de informação. Normalmente, considera-se que, a partir do quarto componente, as imagens contêm apenas pixels ruidosos, podendo ser desprezadas. Em outras palavras, os primeiros componentes principais separam a variância não aleatória (conteúdo de informação) da variância aleatória (ruído) que estão nos últimos CPs, enquanto, simultaneamente, qualquer redundância entre as bandas é eliminada (correlação = 0). Como resultado prático, os três primeiros componentes, que contêm 99,691% da variância total original, podem ser combinados em uma única composição colorida RGB. Por essa razão, a TCP é também uma forma de reduzir a dimensão dos dados originais, diminuindo a obrigação de se analisar um grande número de variáveis.

Mas antes que o processo da TCP possa ser concluído, uma segunda operação é processada. Associado à cada autovalor, existe um conjunto de coordenadas que define as direções dos eixos dos sete componentes

principais (“elipsoide” de sete eixos). São chamados de autovetores (*eigenvectors*). A Tabela 12.5 apresenta os autovetores computados.

Tabela 12.5. Matriz de autovetores derivados da matriz de coeficientes de correlação. Eig = *eigenvector*.

Eig.	Banda 1	Banda 2	Banda 3	Banda 4	Banda 5	Banda 6	Banda 7
1	0,07605	0,10889	0,20469	0,35308	0,28459	0,69764	0,49808
2	0,03224	0,03850	0,01857	0,16629	-0,9454	0,10494	0,25432
3	0,16659	0,21258	0,43916	0,66996	0,07412	-0,5281	-0,0299
4	0,15950	0,04701	-0,1777	-0,2746	0,12075	-0,4554	0,80205
5	-0,3738	-0,3733	-0,6293	0,55215	0,07046	-0,0861	0,08635
6	0,59908	0,50100	-0,5760	0,12867	0,01220	0,08707	-0,1844
7	0,66429	-0,7408	0,07342	0,04414	-0,0056	0,02983	-0,0395

Os coeficientes da matriz de autovetores são interpretados como os cossenos das direções da rotação dos novos eixos em relação aos eixos originais, apontando na direção de qual é a contribuição ou peso das bandas originais na formação de cada componente principal. De uma maneira bem simples, pode-se dizer que a banda original que mais contribui para a formação de um componente principal é a que tem, na matriz de autovetores, o valor absoluto maior. Na Tabela 12.5, o primeiro componente tem uma contribuição muito maior da banda 6. Como consequência, a imagem do CP1 será muito semelhante, na aparência, com a banda 6 original. A diferença com a imagem original da banda 6 é que a CP1 é interpretada como sendo, aproximadamente, uma imagem albedo, a qual descreve o brilho médio da cena. A CP1 é quase uma soma das médias das imagens, enquanto os componentes restantes, de algum modo, parecem ser diferenças entre pares das imagens originais. No CP2, o valor mais extremo é da banda 5. Segundo Richards (2013), autovetores negativos não são um problema porque o espaço dos CPs pode ser arbitrariamente deslocado, tornando os componentes com valores positivos, sem que ocorra mudanças nas propriedades da TCP.

Em termos numéricos, pode-se entender que os autovetores são definidos como uma combinação aditiva e linear computada para cada pixel de uma banda original para criar o novo valor de pixel das imagens de componentes principais. Tomando-se, como exemplo, os autovetores do primeiro componente da matriz da Tabela 12.5, a seguinte operação é feita para se calcular o novo valor de cada pixel da imagem componente principal CP1 (Equação 12.5). A mesma operação é repetida para os demais componentes.

$$CP1 = 0,07605x_{ij1} + 0,10889x_{ij2} + 0,20469x_{ij3} + 0,35308x_{ij4} + 0,28459x_{ij5} + 0,69764x_{ij6} + 0,49808x_{ij7} \quad (12.5)$$

em que

$X_{ij1...7}$ = valor do pixel posicionado na linha i e coluna j de cada uma das bandas originais 1 a 7.

A Figura 12.5 mostra as imagens resultantes da TCP, cujos dados são os das matrizes antes apresentadas. Visualmente, é significativa a diminuição gradativa do contraste ou da quantidade de informação da CP1 à CP7, com os últimos componentes parecendo guardar somente os sinais de ruído dos dados, sem valor de informação. Como as imagens componentes principais têm correlação zero, qualquer pixel lido em um componente tem um correspondente diferente do valor digital nos outros componentes. Os três primeiros componentes principais, que reúnem 99,691% da informação total, podem ser combinados em uma composição colorida RGB, obtendo-se uma imagem de alto contraste de cores, mais contrastantes do que as obtidas com as bandas originais. Isso ocorre devido à ausência de correlação entre os componentes. Porém, há uma grande dificuldade para se identificar as naturezas das classes dos alvos, pois, sendo cada uma das imagens componentes formadas pela soma das contribuições de informações de todas as bandas originais, as relações espectrais dos alvos com as imagens são perdidas.

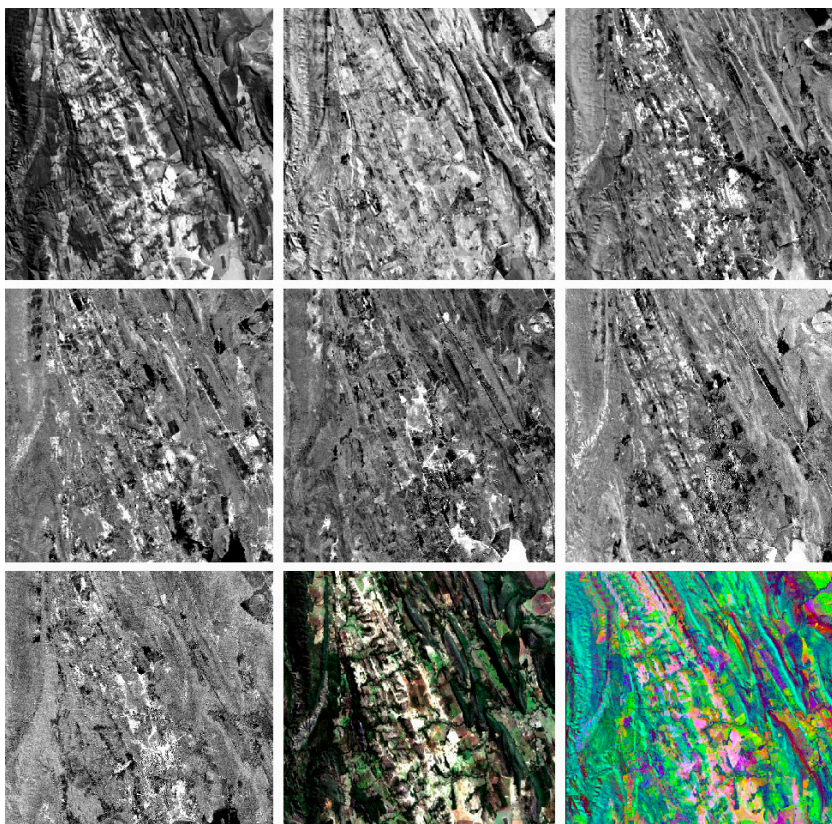


Figura 12.5. Imagens componentes principais de sete bandas OLI correspondentes aos valores da matriz de autovalores em porcentagens da Tabela 12.4. Composição colorida real (R/G/B) e da componentes principais 1/2/3.

Crósta e Moore (1989) desenvolveram uma solução para essa dificuldade de identificação da natureza das classes nas imagens componentes principais, por meio da aplicação de uma técnica de análise denominada *Feature Oriented Principal Component*, também denominada de técnica Crósta. Essa técnica possibilita identificar as classes de alvos por meio da análise da matriz de autovetores, reconhecendo qual componente contém a informação diretamente relacionada à assinatura espectral do alvo.

A primeira etapa dessa técnica é a avaliação do espectro do alvo a ser pesquisado, com o objetivo de identificar duas bandas que contenham respostas de feições espectrais marcantes, uma com alta absorção e outra com alta reflectância. Depois de selecionadas essas duas bandas, são escolhidas outras duas bandas nas quais o alvo tenha reflectância média. Nessas quatro bandas, é aplicada a TCP. A seguir, são analisados os valores e sinais da matriz de autovetores, considerando-se que a imagem componente principal que contém a informação individualizada do alvo é aquela que apresenta os maiores e menores valores absolutos nas bandas que têm as feições espectrais do alvo. Esse componente conterá a informação individualizada do alvo. Vamos verificar isso por um exemplo prático de uma área granítica com alteração hidrotermal moscovítica. Na Figura 12.6, o espectro de reflectância do mineral muscovita apresenta uma forte feição de absorção na banda 7 e uma alta reflectância na banda 5 dos sensores *Thematic Mapper* (TM) e ETM+. A aplicação da técnica Crósta para mapeamento desse mineral utiliza a rotação por componentes principais das bandas 1, 4, 5 e 7. Na análise da matriz de autovetores (Tabela 12.6), deve-se identificar o componente principal que possui sinais opostos e variações de moderado a alto para as bandas 5 e 7 (CP4), sendo esse o componente principal que identifica as áreas que concentram a presença de muscovita nas rochas.

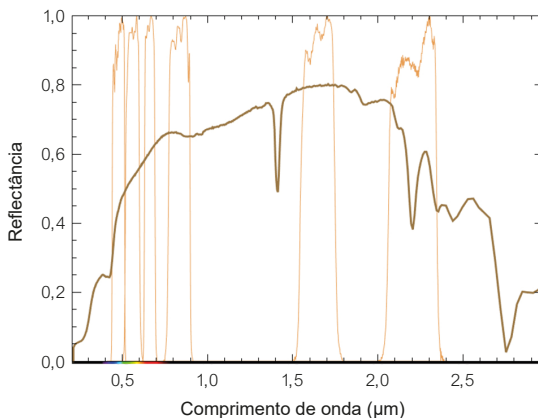


Figura 12.6. Espectro da muscovita sobreposto ao intervalo das bandas 1, 2, 3, 4, 5 e 7 do sensor ETM+ do satélite Landsat 7.

Tabela 12.6. Autovetores obtidos por componentes principais envolvendo as bandas 1, 4, 5 e 7 do satélite Landsat 7 ETM+.

	B1	B4	B5	B7
CP1	0,199966	0,250449	0,780750	0,536394
CP2	0,191043	0,887538	0,050195	0,416244
CP3	0,875854	0,264005	0,387476	0,114209
CP4	0,395470	0,282583	0,487619	0,725244

O resultado pode ser visto na Figura 12.7, relativa ao granito da Serra de Pedra Branca da província estanífera de Goiás, localizada no município de Nova Roma (Figura 12.7A). Na matriz de autovetores da Tabela 12.6, observa-se que os valores que apresentam a informação relativa à muscovita é o CP4 (maiores sinais de valores opostos). A imagem CP4 é mostrada nessa figura, onde se destaca uma pequena área de forma oval que possui forte brilho (indicada pela seta) e que coincide com a faixa de *greissen* muscovítico mineralizado em estanho neste granito, segundo o estudo de Almeida Filho (1983).

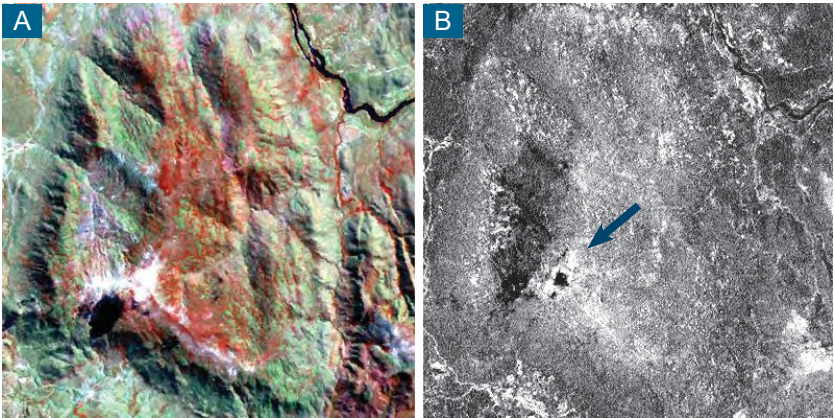


Figura 12.7. Composição colorida da área do granito Serra da Pedra Branca (GO) e a imagem do componente principal segundo o uso da técnica Crôsta (*feature-oriented principal component*) com as bandas 1, 4, 5 e 7 do satélite Landsat 7 ETM+, destacando a área de *greissen* moscovítico, indicada pela seta.

12.2. Descorrelação de Bandas

A descorrelação de bandas, também conhecida como *decorrelation stretching*, corresponde a outra técnica que também zera a correlação entre as bandas de um sensor, com princípios semelhantes à TCP, e é bastante usada para a obtenção de composições coloridas a partir de um conjunto de apenas três bandas. Essa técnica objetiva aumentar o contraste das cores de imagens com alta correlação entre si, tendo em vista que as composições coloridas RGB obtidas por meio de técnicas de ajustes histográficos somente exibem um alto contraste de cores se as imagens possuírem baixas correlações espectrais. Quando as bandas possuem alta correlação, produzem-se composições coloridas RGB em que algumas cores altamente saturadas não são realçadas. Cores saturadas primárias ou secundárias são cores com alto brilho e com alto contraste. A ausência de cores saturadas na imagem comprova que as modificações de contraste por ampliações histográficas não conseguem promover uma redistribuição equilibrada das cores entre bandas que tenham alto grau de correlação.

Graficamente, esse efeito pode ser mostrado no espaço cúbico de cores RGB (Figura 12.8A), onde a representação histográfica tridimensional de três bandas é uma distribuição de forma elíptica, alongada segundo o eixo acromático. O tamanho da elipse é dimensionado pelos seus três eixos, ao longo dos quais são representados os espalhamentos dos pixels de cada uma das três bandas originais, como apresentado anteriormente na Figura 12.4. O eixo maior da elipse e a sua direção representam a variação de intensidade de brilho da imagem, a qual é a maior fonte de correlação interbanda. Depois de aplicar uma ampliação de contraste a cada uma das três bandas, a elipse torna-se mais alongada e ocupa uma porção maior do espaço cúbico de cores (Figura 12.8B). Contudo, a mesma proporção básica dos eixos da elipse continua mantida, indicando que a correlação entre as bandas realçadas continua permanecendo. A principal consequência da manutenção dessa correlação é que grande parte do espaço volumétrico tridimensional do espaço de cores do cubo não é preenchido.

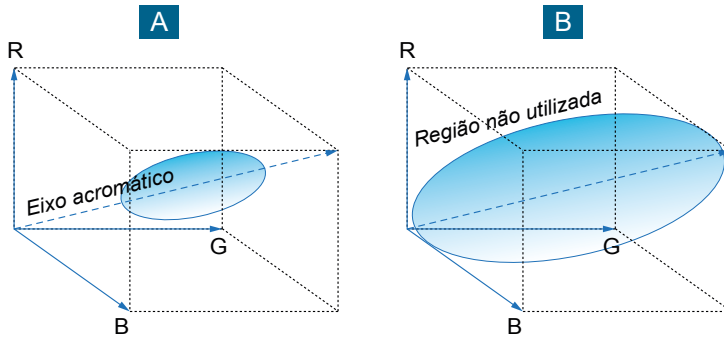


Figura 12.8. No espaço de cores RGB, três bandas geram um histograma em três dimensões que assume a forma de uma elipse alongada (A). Um realce linear amplia a elipse, mas a correlação existente entre as bandas permanece, conforme ilustrado em (B).

Para ocupar todo o espaço de cores do cubo representado na Figura 12.8, e assim se ter uma contribuição total ou saturada de todas as cores representadas pelos três eixos RGB do cubo, o elipsoide deve-se tornar uma esfera. Por definição, estatisticamente, o círculo representa o espaço de distribuição entre variáveis não correlacionadas, ou seja, correlação zero (Figura 12.9). Visto no espaço bidimensional de duas bandas, qualquer ponto posicionado na linha do círculo tem um valor no eixo x que é diferente do valor no eixo y.

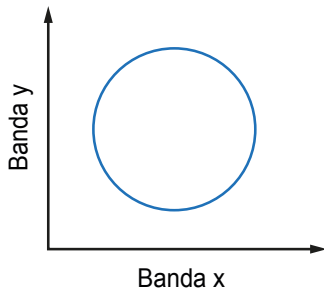


Figura 12.9. O círculo representa o espaço de uma distribuição entre duas variáveis não correlacionadas.

Para se obter a transformação por decorrelação e gerar uma composição colorida de alto contraste, é necessário realizar a transformação de

três bandas em que os valores da imagem de saída, ou novos componentes, dependem dos valores conjuntos das três imagens de entrada e não dos valores individuais de cada banda, como ocorre com a ampliação de contraste por manipulação histográfica. Esse tipo de transformação tem o pressuposto de obter uma distribuição esfericamente simétrica que resulta em três novas imagens descorrelacionadas. Recordando o que se viu na seção anterior, tal transformação segue os mesmos procedimentos da TCP.

Os novos componentes descorrelacionados passam a ter uma distribuição que ocupa quase todo o espaço cúbico de cores. Dessa forma, as feições de cores são maximizadas sem que haja alteração considerável das cores originais da imagem. Isso porque os componentes não perdem a relação dos valores digitais transformados com as propriedades físicas originais da radiação. O resultado é a obtenção de uma composição colorida com uma intensidade de contraste de cores muito maior do que a da composição colorida primária. A vantagem é que a interpretação da imagem é bem menos complexa do que a de componentes principais e pode ser feita com base nas propriedades do comportamento de reflectância dos alvos, como se faz em uma composição RGB normal. A limitação dessa transformação é que a sua aplicação é reduzida a três bandas.

Os procedimentos da transformação por descorrelação de bandas baseiam-se no princípio da transformação *whitening* ou transformação esférica. Supondo-se que X é um vetor de valores de brilho da imagem de entrada com uma matriz de covariância Σ e média zero, tem-se que (Equação 12.6):

$$Y = W \times X \quad (12.6)$$

Por essa equação, a matriz *whitening* $W = \Sigma^{-1/2}$, Σ denotando a matriz de covariância, produz um vetor Y *whitening* com covariância diagonal. O sinal negativo indica a transposição da matriz de covariância (Equação 12.7).

$$Y = (\Sigma\sigma^{-1/2}) \times X \quad (12.7)$$

Essa matriz também corresponde a uma transformação linear como a de componentes principais, que transforma um vetor de variáveis aleatórias com uma matriz de covariância conhecida em um conjunto de novas variáveis cuja covariância é a matriz de identidade, significando que elas são não correlacionáveis e tem a mesma variância, conforme indicado pela matriz abaixo. Na matriz identidade, λ são os valores dos coeficientes dos autovalores (medidos em unidade de variância) e os valores nulos indicam a não correlação dos dados, tal como na TCP (Equações 12.8 e 12.9).

$$|W| = \begin{bmatrix} \lambda_1 & \lambda_1 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \quad (12.8)$$

$$|W| = \begin{bmatrix} \frac{1}{\sqrt{\lambda_1}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{\lambda_2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{\lambda_3}} \end{bmatrix} \quad (12.9)$$

A transformação *whitening* tem esse nome porque ela muda o vetor de dados de entrada em vetores de 'ruído branco'. *White noise vector* refere-se a um sinal tendo igual intensidade em diferentes frequências. Em processamento de imagens digitais, *white noises* possuem a premissa de serem variáveis aleatórias independentes com distribuição de probabilidade uniforme (esfera).

A Figura 12.10A mostra a representação gráfica em duas dimensões da transformação das bandas X e Y que possuem alta correlação (ver orientação da elipse) para os novos eixos ortogonais DEC1 e DEC2 que representam as novas imagens descorrelacionadas e com igual variância (círculo).

Os dois eixos do círculo estão orientados paralelamente aos eixos DEC1 e DEC2. A rotação dos eixos do círculo ao redor do seu centro (Figura 12.10B) não muda o formato da distribuição, ou seja, não muda a correlação entre DEC1 e DEC2. O passo seguinte é a aplicação de um realce individual dos componentes por uma equalização de variância que assegura que os dados de cada componente tenham a mesma variância e ocupem, em cada eixo, o volume máximo esférico no espaço de cores.

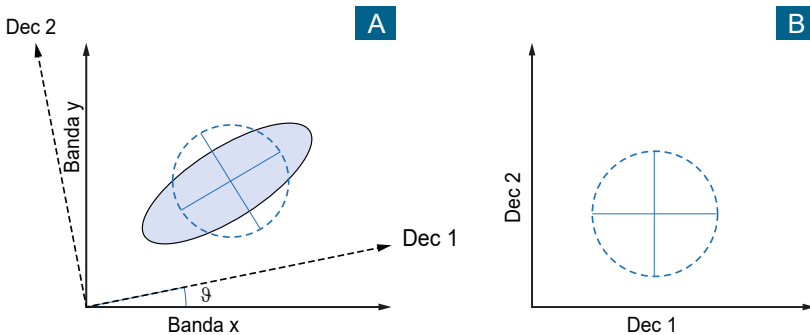


Figura 12.10. Transformação *whitening* da distribuição elíptica dos dados originais com correlação positiva (A) para uma distribuição circular com correlação zero (B) rotacionado por θ .

A principal vantagem do processo de decorrelação de bandas está na etapa em que os dados das imagens transformadas retornam ao espaço original de cores. Nesse caso, o retorno com rotação é mais vantajoso do que no método de componentes principais, pois o método de decorrelação permite que as relações de cores básicas das bandas da imagem original sejam preservadas. Isso facilita a identificação, por parte do intérprete, de alvos baseada nas cores das imagens componentes sem perda de informações relacionadas com as radiancias dos alvos. Para retornar ao espaço de cores originais, efetua-se o inverso da transformação por componentes principais (TPC⁻¹). Desde que a matriz de transformação dos autovalores é ortogonal, sua inversa é igual à sua matriz transposta. Lembre-se que uma matriz transposta é uma matriz em que os elementos das linhas (i, j) são

trocados pelos elementos das colunas (j, i), isto é, a matriz é rotacionada ao redor da diagonal principal.

A eficiência desse processo de realce de cores das imagens pode ser vista no exemplo da Figura 12.11, que compara uma composição colorida RGB das bandas 5, 4, e 3 do satélite Landsat 7 ETM+, realçadas independentemente pela técnica de expansão linear de contraste, com a composição colorida das mesmas bandas descorrelacionadas. Nota-se que as cores na imagem descorrelacionada são mais claras (saturadas), daí o nome da transformação se denominar *whitening*. O lago na imagem da composição colorida normal é de um azul marinho escuro, que se confunde com um tom de cinza preto, induzindo o intérprete a imaginar que o nível de reflectância da água é zero. Na imagem por descorrelação, a cor é realçada para um azul mais claro, mostrando de forma mais correta que a água não tem reflectância zero nos comprimentos de onda das três bandas.

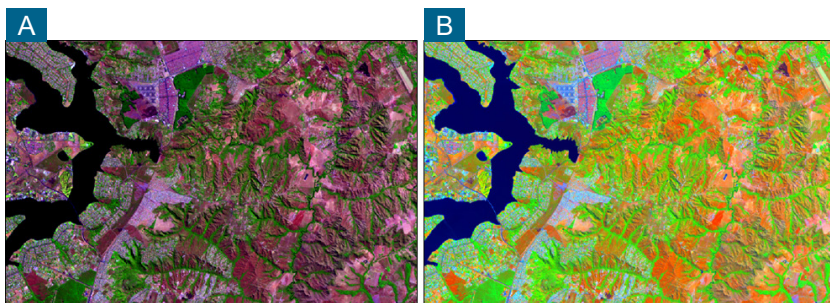


Figura 12.11. Comparação entre as composições coloridas RGB das bandas 4, 5 e 6 obtidas com realce linear de cores (A) e com realce de cores precedida de descorrelação de bandas (B) de uma imagem sobre a porção leste do Distrito Federal pelo satélite Landsat 8 OLI de 04 de agosto de 2021.

12.3. Referências

ALMEIDA FILHO, R. **Sensoriamento remoto orbital aplicado à prospecção mineral nas províncias estaníferas de Goiás e Rondônia: uma contribuição metodológica** (tese de doutorado em Geologia). São Paulo: USP, 1983, 171 p.

DAVIS, C. J. **Statistics and Data Analysis in Geology**. Houston: John Wiley, 3a ed., 2002, 656 p.

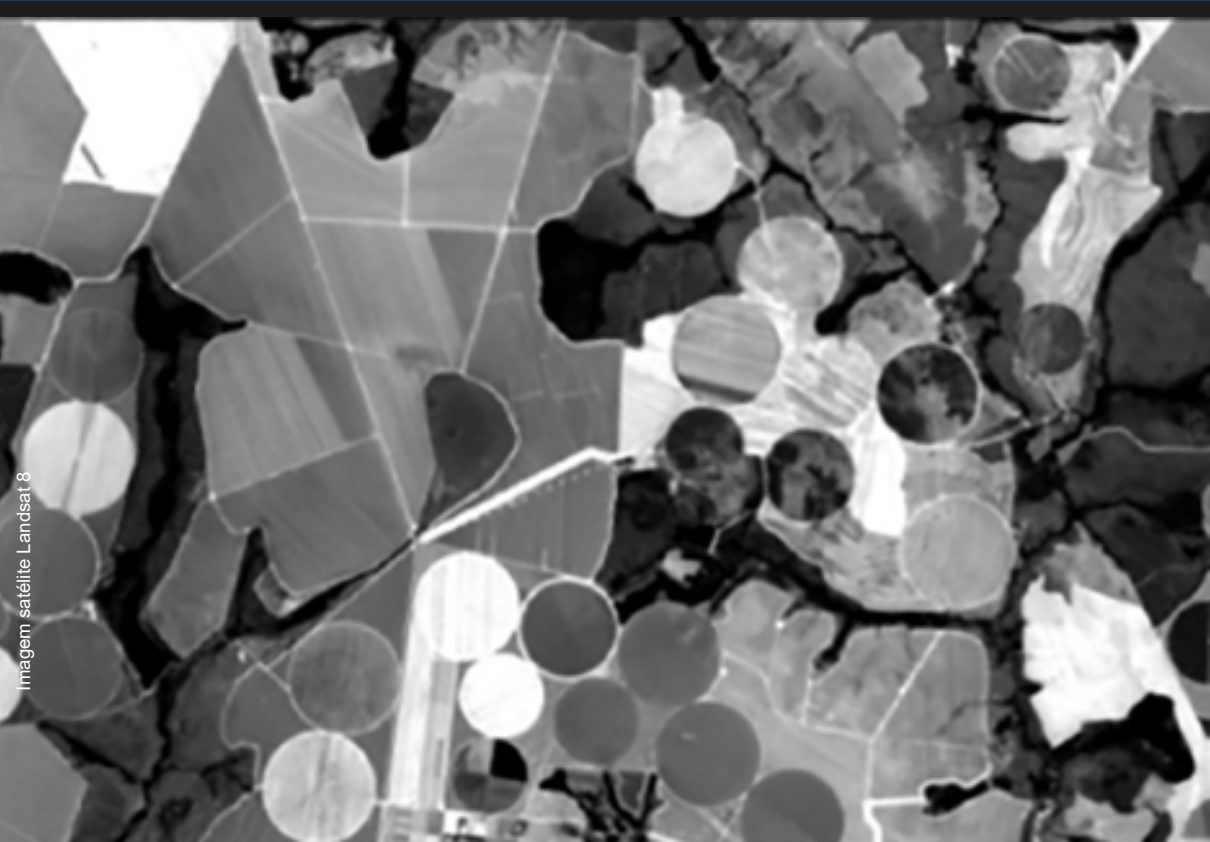
MATHER, P. M. **Computer Processing of Remotely Sensed Images**: an Introduction. Chichester: John Wiley, 1987, 360 p.

CAPÍTULO 13

Filtragem Espacial

Paulo Roberto Meneses

Edson Eyji Sano



Os valores digitais de uma imagem não devem ser vistos apenas como valores relacionados exclusivamente à reflectância dos alvos. Eles devem ser vistos ainda como valores relacionados com duas dimensões, a espectral e a espacial. Em muitos casos, estamos mais interessados em extrair informações das imagens com base na organização espacial dos valores digitais do que propriamente em identificar as propriedades espectrais dos pixels individuais. A filtragem é uma transformação pixel a pixel que não depende apenas do nível de cinza de um determinado pixel, mas também dos valores dos níveis de cinza dos pixels vizinhos. Por isso, a filtragem é considerada como uma técnica de transformação de imagem no domínio espacial, muito embora podemos filtrar a imagem decompondo suas frequências. A filtragem tem o objetivo de realçar o contraste das imagens e extrair informações das estruturas locais de detalhe ou de suprimir os ruídos que contaminam os pixels e que degradam as informações contidas nas imagens.

13.1. Modelos de Filtragem

Em qualquer imagem, sempre é observada uma forte interdependência da vizinhança espacial dos valores dos pixels, pois os alvos na natureza tendem a ser homogêneos dentro de certos espaços. Essa interdependência pode ser utilizada favoravelmente para realçar, de forma seletiva, detalhes espaciais das imagens como as bruscas mudanças ou bordas entre áreas aparentemente homogêneas que, na realidade, configuram-se como limites. Também observamos presença de feições que marcam, por assim dizer, o rompimento dessa interdependência e que não parecem estar associadas a limites ou bordas de áreas. Neste caso, essas feições são identificadas como ruídos.

As técnicas de processamento que atuam no sentido de realçar as bordas ou limites e de reduzir a presença de ruídos são denominadas de filtragens. Basicamente, há duas formas de filtrar uma imagem: filtros que

atuam no domínio do espaço e filtros que atuam no domínio da frequência. A maioria dos filtros promovem transformações no domínio espacial, caracterizando-se por definir operações de filtragem a partir da informação local das imagens, dentro de um pequeno espaço em torno de um pixel. Por isso, são chamados de filtros espaciais. Os filtros que atuam no domínio da frequência, onde o operador é a transformada de Fourier, usa o conteúdo do espaço total da imagem que é decomposta em diferentes escalas ou frequências.

Muitos dos filtros no domínio do espaço que são aplicados nas imagens de sensoriamento remoto usam, como operação matemática, a convolução para suavizar, aumentar o contraste, realçar bordas e texturas e remover ruídos das imagens de entrada. Há pelo menos três tipos de filtros: morfológico, textural e adaptativo. Os filtros morfológicos processam as imagens com base na forma do objeto, usando uma morfologia matemática de elementos estruturantes para filtrá-las. Os filtros texturais extraem informações das imagens sobre os padrões de tons de cinza que definem os objetos. Os filtros adaptativos reduzem o ruído pela suavização das imagens, preservando o realce dos limites ou as bordas dos alvos. No domínio da frequência, as imagens são transformadas do domínio espacial para o domínio das frequências para passar ou eliminar as baixas ou altas frequências das imagens. O propósito é o mesmo dos filtros espaciais para suavizar ou realçar as imagens e eliminar ruídos.

Iniciaremos a descrição das técnicas de filtrações com os filtros que atuam no domínio do espaço, apresentando as diferentes configurações dos modelos de filtros espaciais mais utilizados em sensoriamento remoto. Neste caso, uma imagem deve ser considerada como tendo uma distribuição de padrões de valores de pixels que se repetem com uma determinada taxa. Essa taxa de repetição de pixels pode ocorrer em qualquer direção e representa as variações espaciais de brilho ou de níveis de cinza.

Quando a taxa de repetição é interrompida ou sofre uma mudança, o filtro identificará o local desse acontecimento como sendo um limite. Podemos

ver isso na Figura 13.1 na qual é mostrado o perfil de variação de brilho dos pixels medidos ao longo de uma linha da imagem que atravessa áreas urbanas, lagos e reflorestamento. Note, no perfil dos valores de brilho da imagem, que há áreas com fortes variações de taxa de repetição de brilho e outras áreas com baixíssima variação de taxa de repetição. Os limites físicos entre elas são facilmente identificáveis na leitura do perfil e na imagem. O problema maior é identificar limites que consigam separar distintas taxas de brilho onde há fortes variações de brilho, com as do início e final do perfil, que correspondem às áreas urbanas. A filtragem tem a tarefa de identificar esses limites a fim de realçar o padrão de espacialização dos alvos ou suavizar a alta taxa de repetição.

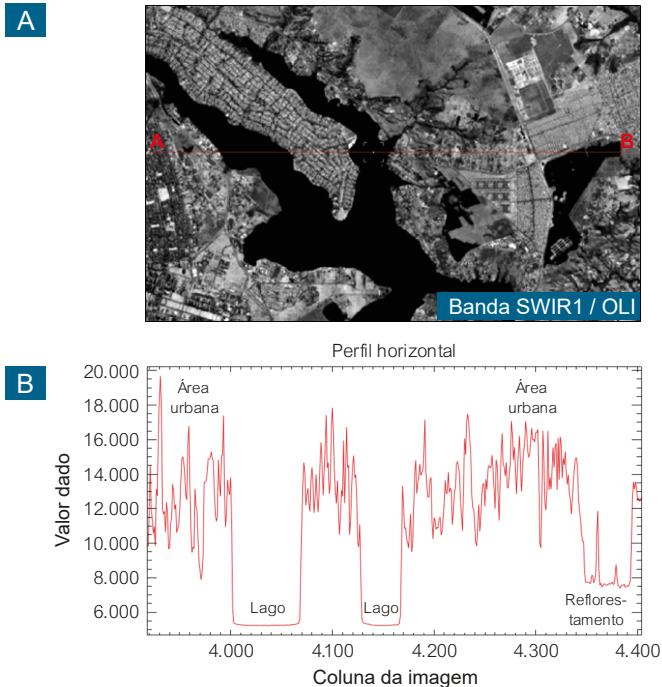


Figura 13.1. Amostragem linear de valores digitais (transecto AB) em uma imagem da banda 6 do satélite Landsat 8 OLI obtida na porção central do Distrito Federal (A). Em (B), é mostrada a variação da intensidade de brilho dos alvos localizados ao longo do transecto.

Esse conceito de variação espacial de brilho é aplicado aos métodos de filtragem espacial que vê as imagens como sendo formadas por componentes espaciais em diferentes escalas. Quando falamos em taxa de repetição de valores, estatisticamente estamos falando de moda, o que é bem entendido pela variação de brilho com a mesma intensidade de locais espacialmente diferentes. No entanto, quando falamos em taxa, intuitivamente não deixamos de pensar em frequência. Por isso, a taxa de variação espacial de brilho e a frequência espacial são muitas vezes empregadas como sinônimos.

Lillesand et al. (2008) referiram-se à "frequência" no domínio espacial como sendo a rugosidade das variações tonais que ocorrem em uma imagem. Áreas com alta frequência espacial são "tonalmente rugosas" e mostram as partes das imagens em que ocorrem rápidas mudanças dos valores de brilho, como se vê nas áreas urbanas da Figura 13.1. Elas expressam os detalhes da imagem em uma escala local. As partes da imagem com baixa frequência espacial, por exemplo, a área do lago na Figura 13.1, são tonalmente lisas, com pouca variação de brilho, ou seja, expressam a homogeneidade do alvo, caracterizando distintamente seu comportamento espectral. Nesse sentido, a filtragem espacial realça ou suaviza localmente parte dessas repetições de padrão de brilho, mesmo quando os padrões de variações de brilho são sutis. Quando se deseja enfatizar as altas frequências espaciais de variações de brilho, o filtro espacial é genericamente chamado de passa-alta. Quando se deseja enfatizar as baixas frequências, o filtro espacial é chamado de passa-baixa. Considerando que qualquer imagem é composta de uma estrutura espacial aleatória de regiões com baixas e altas frequências espaciais, matematicamente, pode-se escrever essa relação como (Equação 13.1):

$$f(x, y) = PB(x, y) + PA(x, y) \quad (13.1)$$

em que

PB = passa-baixa; e

PA = passa-alta.

A decomposição das imagens em uma soma de componentes de baixas e altas frequências de brilho é a base para a filtragem espacial. Os filtros espaciais operam diretamente sobre os pixels da imagem original e utilizam um operador de convolução discreta, por isso, são chamados de filtros de convolução. Convolução é um operador matemático que vai operar dois elementos distintos, a imagem e o filtro. Trata-se simplesmente de uma operação de somatório do produto entre duas funções ao longo da região em que elas se sobrepõem, em razão do deslocamento existente entre elas. A função de processamento de convolução discreta entre a imagem e a janela móvel do filtro pode ser escrita na seguinte forma (Equação 13.2):

$$g(i, j) = T[f(x, y)] \quad (13.2)$$

em que

$g(i, j)$ = imagem processada;

$f(x, y)$ = imagem de entrada; e

T = operador de transformação que atua sobre um conjunto de pixels vizinhos da imagem.

Matematicamente, a imagem $g(i, j)$ de saída é dada pela Equação 13.2 (Schowengerdt, 2007):

$$g(i, j) = \sum_{m=0}^{N_x-1} \sum_{n=0}^{N_y-1} f_{mn} W_{i-m, j-n} \quad (13.3)$$

em que

W = função de resposta da janela, ambas com ambas com tamanho N_x, N_y ; e

f = imagem de entrada.

A operação de convolução discreta para filtrar as imagens usa uma janela de pixel com pesos fixados a ela, que são números positivos e/ou negativos, de acordo com o que se deseja filtrar das imagens. A janela é, de fato, uma matriz, convencionalmente chamada de máscara. Durante a operação de convolução, a janela do filtro se movimenta sobre a imagem e vai multiplicando o pixel pelo peso do filtro e somando os valores sobrepostos, colocando o valor calculado na imagem de saída, na mesma posição do pixel central da janela. O processo é repetido varrendo-se toda a imagem pixel a pixel como no esquema da Figura 13.2. Trata-se de um processo que opera a cada momento em um espaço local da imagem. Quando o último pixel é processado, a imagem original é modificada radiometricamente, resultando em realces, suavizações, contraste de bordas e texturas ou remoção de ruídos, de acordo com a configuração da janela do filtro e dos valores dos pesos que foram estabelecidos. A flexibilidade de se usar praticamente qualquer peso na configuração de um filtro espacial faz da convolução um instrumento de processamento de imagem de muitas aplicabilidades.

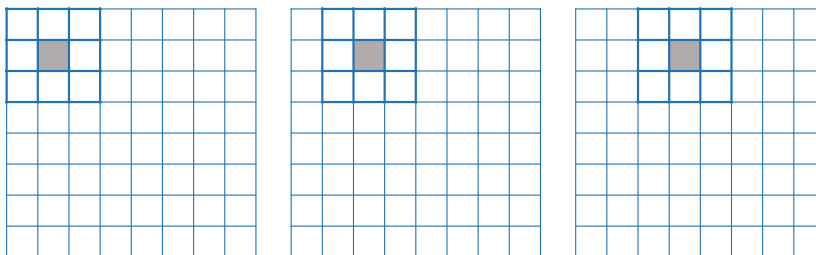


Figura 13.2. Movimento da janela de tamanho 3 x 3 para filtrar imagens. O pixel cinza é o pixel central que terá o seu valor modificado, conforme os pesos dos pixels do filtro.

Na Tabela 13.1, são listadas as principais classes e tipos de filtros. Os filtros no domínio espacial mais utilizados no processamento de imagens multiespectrais usam o processo de convolução e se baseiam em três alternativas: soma dos pesos; estatística específica; e noção de derivada. Os

filtros adaptativos são bastante usados para supressão de ruídos em imagens de radar.

Tabela 13.1. Filtros de uso mais comum em sensoriamento remoto.

Classe	Tipo	Aplicações
Domínio espacial (Operador de convolução)	Passa-alta	Realce de imagem
	Passa-baixa	Supressão de ruído
	Gaussiano passa-alta	Realce de variações locais
	Gaussiano passa-baixa	Suavização de imagens
	Direcional	Realce de bordas nas direções horizontal, vertical e diagonal
	Mediana	Remoção de ruído
	Gradiente	Realce de bordas direcionais
Domínio espacial (Operador sem o uso de convolução)	Laplaciano	Realce de bordas não-direcionais
	Morfológico	Preenchimento ou erosão de contornos
	Textural	Realce de variações espaciais de tons de cinza da imagem
Domínio de frequência	Adaptativo	Supressão de ruído com preservação de detalhes locais
		Realce em relação à proporção das informações associadas com as frequências

13.2. Filtros Passa-Alta e Passa-Baixa

De acordo com a Equação 13.1, a soma dos filtros passa-alta e passa-baixa é igual à imagem original, isto é, é igual à soma das taxas altas e baixas de repetição espacial de brilho da imagem que passaremos a referir como alta ou baixa frequência espacial. Ambos os filtros usam máscaras simples do tipo caixa (*box filters*), as quais têm peso uniforme no filtro passa-baixa e pesos complementares no filtro passa-alta. Somente podem ter máscaras de dimensões ímpares, por exemplo, 3 x 3 ou 5 x 5. Exemplos de máscaras de filtros passa-alta e passa-baixa são mostrados na Tabela 13.2.

O filtro passa-alta tem um valor central alto e positivo e valores negativos circundantes. A soma dos pesos é igual a zero, o que remove a média local e produz uma imagem que mede o desvio do sinal de entrada em relação à média local. Quando a máscara é aplicada em regiões homogêneas das imagens, o resultado será zero ou um valor digital muito baixo, enquanto nas regiões não homogêneas, o resultado é um realce, o que é consistente com o princípio da filtragem passa-alta. A aplicação de um filtro passa-alta com esses valores de peso pode resultar em números digitais negativos. Isso é automaticamente corrigido pelos algoritmos que aplicam, após a convolução, uma normalização dos valores digitais para o intervalo de quantização das imagens.

Tabela 13.2. Exemplos de configurações das janelas dos filtros passa-alta e passa-baixa.

Tamanho	Filtro passa-alta	Filtro passa-baixa																																																		
3 x 3	$\frac{1}{9}$ <table><tr><td>-1</td><td>-1</td><td>-1</td></tr><tr><td>-1</td><td>+8</td><td>-1</td></tr><tr><td>-1</td><td>-1</td><td>-1</td></tr></table>	-1	-1	-1	-1	+8	-1	-1	-1	-1	$\frac{1}{9}$ <table><tr><td>+1</td><td>+1</td><td>+1</td></tr><tr><td>+1</td><td>+1</td><td>+1</td></tr><tr><td>+1</td><td>+1</td><td>+1</td></tr></table>	+1	+1	+1	+1	+1	+1	+1	+1	+1																																
-1	-1	-1																																																		
-1	+8	-1																																																		
-1	-1	-1																																																		
+1	+1	+1																																																		
+1	+1	+1																																																		
+1	+1	+1																																																		
5 x 5	$\frac{1}{25}$ <table><tr><td>-1</td><td>-1</td><td>-1</td><td>-1</td><td>-1</td></tr><tr><td>-1</td><td>-1</td><td>-1</td><td>-1</td><td>-1</td></tr><tr><td>-1</td><td>-1</td><td>+24</td><td>-1</td><td>-1</td></tr><tr><td>-1</td><td>-1</td><td>-1</td><td>-1</td><td>-1</td></tr><tr><td>-1</td><td>-1</td><td>-1</td><td>-1</td><td>-1</td></tr></table>	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	+24	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	$\frac{1}{25}$ <table><tr><td>+1</td><td>+1</td><td>+1</td><td>+1</td><td>+1</td></tr><tr><td>+1</td><td>+1</td><td>+1</td><td>+1</td><td>+1</td></tr><tr><td>+1</td><td>+1</td><td>+1</td><td>+1</td><td>+1</td></tr><tr><td>+1</td><td>+1</td><td>+1</td><td>+1</td><td>+1</td></tr><tr><td>+1</td><td>+1</td><td>+1</td><td>+1</td><td>+1</td></tr></table>	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1
-1	-1	-1	-1	-1																																																
-1	-1	-1	-1	-1																																																
-1	-1	+24	-1	-1																																																
-1	-1	-1	-1	-1																																																
-1	-1	-1	-1	-1																																																
+1	+1	+1	+1	+1																																																
+1	+1	+1	+1	+1																																																
+1	+1	+1	+1	+1																																																
+1	+1	+1	+1	+1																																																
+1	+1	+1	+1	+1																																																

Filtros passa-alta são usados para realçar detalhes locais das imagens, o que implica em remover os componentes de baixa frequência, enquanto retém os de alta frequência (variações locais). O objetivo é tornar mais nítidas as transições tonais entre diferentes alvos, conhecidas como bordas

ou limites. O lado negativo é que realçam os ruídos existentes nas imagens. Às vezes, os filtros passa-alta maiores que 100 pixels x 100 pixels são usados para seletivamente eliminar alguns efeitos de baixa frequência, como sombreamentos. A Figura 13.3 mostra o efeito da aplicação do filtro passa-alta em um segmento de imagem do satélite Landsat 8 onde predominam áreas agrícolas irrigadas sob sistema de pivô-central. As regiões de baixa frequência são praticamente convertidas em um mesmo tom de cinza para que sejam destacadas apenas as bordas. As imagens filtradas perdem o valor para interpretar a natureza das classes de alvos.

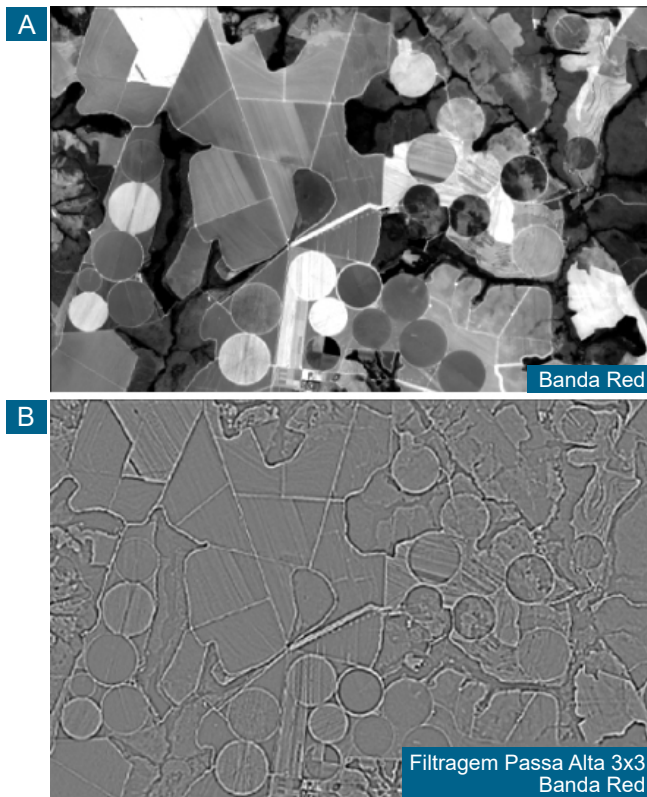


Figura 13.3. Resultado de filtragem passa-alta com janela de 3 x 3 (B) na imagem da banda do visível do satélite Landsat 8 (A) com realce dos limites e bordas entre as classes de alvos.

Quando o objetivo é fazer uma interpretação visual temática da imagem filtrada, é desejável reforçar os componentes de alta frequência que representam os detalhes da imagem, sem eliminar totalmente os componentes de baixa frequência. Nesse caso, utiliza-se um filtro passa-alta que é multiplicado por um fator de amplificação 15 (Equação 13.4):

$$PA(x, y, K) = f(x, y + K \times PA(x, y)) \quad K \geq 0 \quad (13.4)$$

Exemplos de filtros passa-alta reforçados (*high boost filters*) são mostrados na Tabela 13.3. A soma de seus pesos é igual a 1, o que significa que a imagem de saída tem a mesma média de valores digitais da imagem de entrada, preservando a baixa frequência, enquanto a quantidade de realce obtido é proporcional a K . A escolha do valor K depende do grau de realce que é desejado. No exemplo mostrado na Figura 13.4, foi aplicado um fator de amplificação $K = 2$.

Tabela 13.3. Exemplos de filtro passa-alta amplificado, com diferentes valores de K .

$K = 1$				$K = 2$				$K = 3$			
1/9				1/9				1/9			
-1	-1	-1		-2	-2	-2		-3	-3	-3	
-1	+17	-1		-2	+25	-2		-3	+33	-3	
-1	-1	-1		-2	-2	-2		-3	-3	-3	

Os filtros passa-baixa devem conter os mesmos pesos com valores iguais a +1 para cada elemento da máscara. Multiplicando-se os pesos da máscara pelos pixels da imagem e dividindo o resultado pelo número total de pixels da máscara, o valor do pixel central da imagem é substituído pela média dos valores dos pixels vizinhos sob a janela. A saída é a média dos pixels contidos na vizinhança da máscara e, por isso, o filtro passa-baixa é considerado um filtro de média. O efeito do filtro passa-baixa de média é

suavizar a variabilidade dos níveis de cinza, diminuindo um pouco o contraste da imagem. Neste caso, preserva-se os componentes de baixa frequência à custa de reduzir os de alta frequência, atenuando bordas e limites e os detalhes finos que correspondem às transições abruptas de brilho entre as classes de alvos. Isso, de certa forma, consiste em reduzir os detalhes irrelevantes das imagens, isto é, regiões menores que o tamanho da janela.

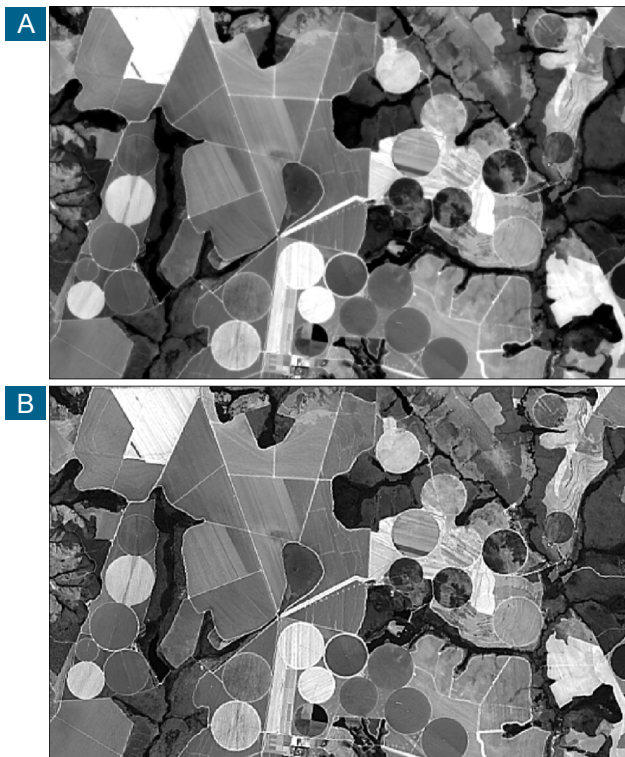


Figura 13.4. Banda 4 do satélite Landsat 8 OLI adquirido em 04 de agosto de 2021 (A), realçada com aplicação de um filtro passa-alta com fator de amplificação $K=2$ para se obter uma imagem com melhor capacidade de interpretação visual (B).

Quanto maior for o tamanho da máscara, maior será a suavização. Os pixels com valores maiores que os seus vizinhos são reduzidos de modo que os detalhes locais são perdidos e, por isso, o produto é uma imagem

de aparência um pouco desfocada. Pela razão dos filtros passa-baixa reduzirem as altas frequências de brilho da imagem, eles são usados para remover ruídos e, conseqüentemente, uniformizam os valores de brilho da imagem. Como o tamanho do filtro está positivamente relacionado à suavização e à resolução espacial das imagens, filtros maiores do que 9×9 não são usados com frequência. É uma técnica que facilita a interpretação da imagem quando o interesse do intérprete é identificar os padrões gerais dos níveis de cinza e não os seus detalhes, sem se incomodar com a perda de contraste de bordas. A Figura 13.5 mostra o resultado da filtragem passa-baixa. O resultado obtido é de uma imagem com aparência um pouco desfocada, pois os detalhes (altas frequências) são atenuados.

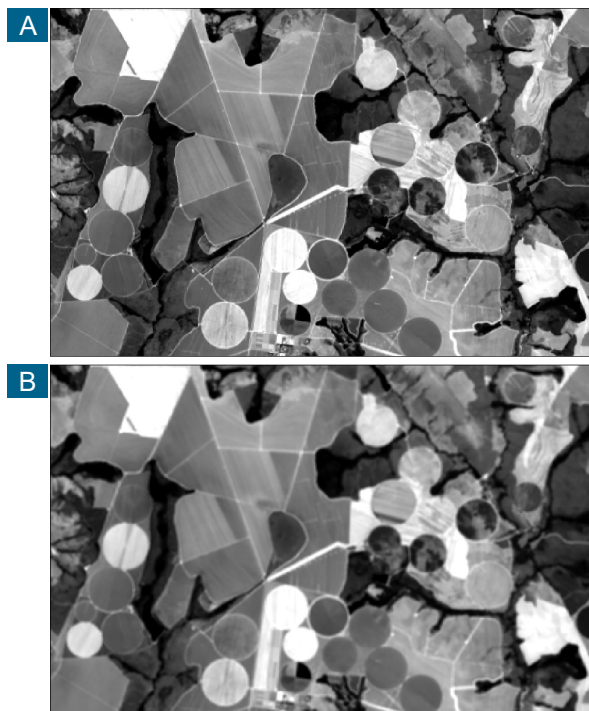


Figura 13.5. Banda 4 do satélite Landsat 8 OLI adquirido em 04 de agosto de 2021 (A), realçada com aplicação de um filtro passa-baixa, o qual suaviza a imagem e reduz a percepção dos limites e bordas entre os alvos.

13.3. Filtros de Mediana

Uma alternativa para minimizar a aparência borrada das imagens pelo filtro de média é usar o filtro de mediana. O valor mediano de um conjunto de pixels na vizinhança do pixel central é de tal forma que a metade dos valores é menor que a mediana e a outra metade é maior. Tomando-se como exemplo um conjunto de pixels de dimensão 3×3 , com os valores [13, 9, 12, 32, 13, 14, 8, 10, 9], para encontrar o valor mediano, deve-se primeiro ranquear esses valores em ordem ascendente de magnitude. No exemplo, a ordem seria [8, 9, 9, 10, 12, 13, 13, 14, 32] com o valor mediano 12 que substituirá o valor do pixel central da imagem. A média é 13,13 que, arredondada para 13, é maior do que seis dos valores exemplificados, indicando ter a influência do valor 32 que pode representar um pixel ruidoso na imagem. Esse valor extremo é eliminado pelo filtro da mediana e as bordas são mais preservadas do que com o filtro de média. Há, assim, duas razões para a escolha da mediana e não da média: (i) a mediana do conjunto de pixels é sempre igual a um dos valores presentes no conjunto de pixels; e (ii) a mediana é menos sensível a pixels de valores extremos. O filtro de mediana é bastante eficiente em imagens de radar para suprimir um tipo muito comum de ruído que essas imagens apresentam, chamado de *speckle* (aspecto de sal e pimenta nas imagens de radar). Um exemplo da aplicação do filtro da mediana é mostrado na Figura 13.6.

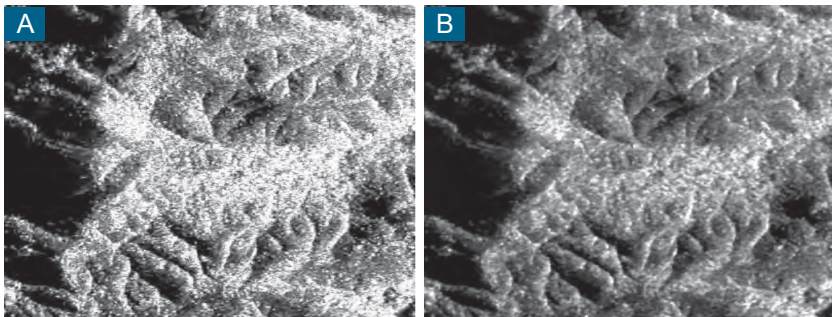


Figura 13.6. Imagem adquirida por um sensor de radar de abertura sintética com ruído *speckle* (A) e com atenuação desse ruído por meio do uso de um filtro de mediana (B).

13.4. Filtros de Bordas Direcionais

Além do simples realce dos componentes de altas e baixas frequências, um clássico problema em processamento de imagem é a detecção de mudanças repentinas nos valores de brilho de um pixel para outro, sem que isso seja um ruído. Esse tipo de mudança indica um limite, isto é, uma borda entre dois objetos espectralmente distintos. Uma borda é simplesmente uma fronteira entre duas regiões com relativa diferença de níveis de cinza. As bordas podem ter variadas direções e a possibilidade de detectar essas direções e realçá-las pode ser algo muito desejado. Esse tipo de filtragem indica a direção preferencial em que será realçado os limites das bordas. Exemplos notáveis de bordas ocorrem em muitas das feições construídas pelo homem. Na Geologia, os lineamentos estruturais que normalmente ocorrem em várias direções são exemplos de feições naturais.

Há diversas configurações bastante simples de filtros direcionais que possibilitam realçar as direções horizontal, vertical, diagonal ou isotrópica. Esses filtros produzem imagens cujos valores digitais dos pixels são proporcionais às diferenças entre os números digitais de pixels vizinhos, em uma dada direção. As bordas que são evidenciadas podem ser ou adicionadas de volta à imagem original para aumentar o contraste nas vizinhanças da borda ou usar pixels saturados pretos ou claros sobre elas para realçá-las. Os dois tipos de filtros de detecção de bordas mais utilizados são os filtros lineares de detecção de borda e os que calculam a primeira derivada espacial.

13.4.1. Filtros lineares

Tipicamente, os filtros lineares correspondem a filtros passa-alta, os quais são lineares, combinados com uma limiarização de valores de brilho. Exemplos de máscaras dos filtros lineares de borda, sugeridos por Richards (2013), são apresentados na Tabela 13.4.

Note que a soma dos elementos do filtro é zero. O resultado é que as áreas com valores de pixels homogêneos são zeradas nas imagens de saída, enquanto as áreas com pixels de brilho variado são apresentadas com bordas brilhantes. Um exemplo da aplicação do filtro direcional linear na direção horizontal é mostrado na Figura 13.7, em uma área com forte presença de lineamentos geológicos. Visualmente, o efeito na imagem filtrada é um forte realce, destacando o relevo associado aos lineamentos na direção horizontal (direção E-W). Na imagem filtrada, as áreas onde foram detectadas as direções horizontais são mais claras, enquanto as áreas ausentes de direção horizontal mostram tons de cinza escuros. Isso pode ser considerado como uma forma muito prática de identificar os domínios geológicos estruturais existentes em uma área, ou as direções de relevo geomorfológico.

Tabela 13.4. Exemplos de filtros lineares de borda.

Vertical	Horizontal	Diagonal	Diagonal																																				
<table><tr><td>-1</td><td>0</td><td>+1</td></tr><tr><td>-1</td><td>0</td><td>+1</td></tr><tr><td>-1</td><td>0</td><td>+1</td></tr></table>	-1	0	+1	-1	0	+1	-1	0	+1	<table><tr><td>-1</td><td>-1</td><td>-1</td></tr><tr><td>0</td><td>0</td><td>0</td></tr><tr><td>+1</td><td>+1</td><td>+1</td></tr></table>	-1	-1	-1	0	0	0	+1	+1	+1	<table><tr><td>0</td><td>+1</td><td>+1</td></tr><tr><td>-1</td><td>0</td><td>+1</td></tr><tr><td>-1</td><td>1</td><td>0</td></tr></table>	0	+1	+1	-1	0	+1	-1	1	0	<table><tr><td>+1</td><td>+1</td><td>0</td></tr><tr><td>+1</td><td>0</td><td>-1</td></tr><tr><td>+1</td><td>-1</td><td>-1</td></tr></table>	+1	+1	0	+1	0	-1	+1	-1	-1
-1	0	+1																																					
-1	0	+1																																					
-1	0	+1																																					
-1	-1	-1																																					
0	0	0																																					
+1	+1	+1																																					
0	+1	+1																																					
-1	0	+1																																					
-1	1	0																																					
+1	+1	0																																					
+1	0	-1																																					
+1	-1	-1																																					

Fonte: Richards (2013).

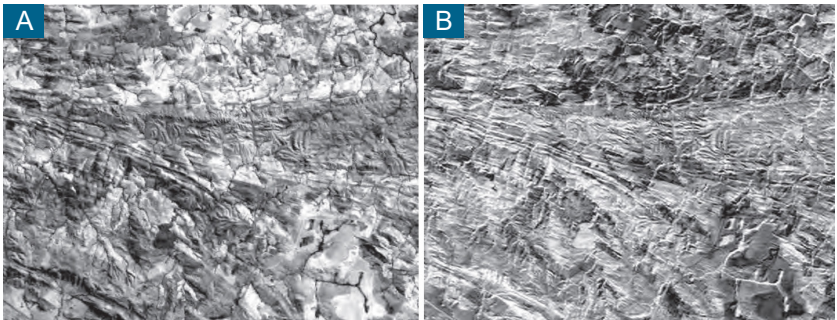


Figura 13.7. Imagem de satélite sobre uma área com elevada ocorrência de lineamentos geológicos (A), realçada com o uso de um filtro linear de direção horizontal (B).

13.4.2. Filtros de primeira derivada

Os filtros de borda não lineares que usam o conceito matemático da primeira derivada, seletivamente, também realçam os componentes de direções específicas que as imagens possam ter, calculando o gradiente direcional dos pixels. Esses filtros são chamados de filtros gradientes. A derivada de uma função contínua é a taxa de mudança da função em um determinado ponto. Sendo a imagem uma função de valores de brilho, a taxa de mudança de brilho dos pixels em um espaço é a primeira derivada. Para um pixel de coordenada (x, y) , as derivadas em x e y são as medidas de quanto rapidamente os valores de brilho mudam na direção x e na direção y . Combinando os resultados por meio de um cálculo de vetor a cada pixel, a magnitude do gradiente local das imagens é obtida para qualquer direção, como demonstrado por Schowengerdt (2007) na Figura 13.8. Nessa figura, a magnitude do gradiente local é dada pela dimensão do vetor $|\Delta|$, que é a derivada total em relação a (x, y) e que é calculada como a raiz quadrada da soma das duas derivadas individuais Δx e Δy . A direção do gradiente local é dada pelo ângulo φ entre o vetor e os eixos x e y . As respectivas equações são (Equações 13.5 a 13.8).

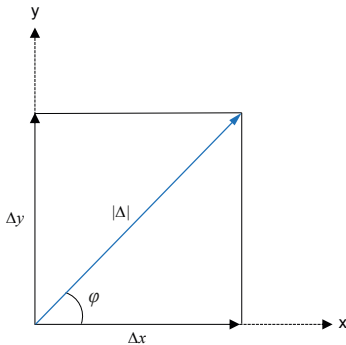


Figura 13.8. Geometria de vetor para cálculo da imagem gradiente.

Fonte: Schowengerdt (2007).

$$\Delta_I = \frac{\partial}{\partial_x} f(x, y) \quad (13.5)$$

$$\Delta_2 = \frac{\partial}{\partial y} f(x, y) \quad (13.6)$$

$$|\Delta| = \sqrt{\Delta_1^2 + \Delta_2^2} \quad (13.7)$$

$$\text{ângulo gradiente} \quad \varphi = \tan \frac{\Delta_y}{\Delta_x} \quad (13.8)$$

Todos esses valores são calculados para cada pixel da imagem. Nas áreas das imagens que são homogêneas, as derivadas nas direções x e y e a derivada total são pequenas, enquanto nas áreas em que ocorrem mudanças abruptas de brilho, as derivadas são altas. Nesse sentido, a primeira derivada, ou a gradiente da imagem, detecta as regiões de alta frequência. Devido ao cálculo do vetor magnitude, os filtros de derivada são não lineares. Com base nessa formulação, as configurações de filtros de derivada se fazem pelo produto escalar da magnitude do vetor gradiente do pixel da imagem com os pesos particulares dos vetores. Os filtros mais utilizados são conhecidos como gradientes de Roberts, Sobel e Prewitt, os quais são mostrados na Tabela 13.5.

Tabela 13.5. Configurações de filtros de borda não lineares (filtros de gradiente).

Filtros	Roberts	Sobel	Prewitt																						
Componente horizontal	<table><tr><td>0</td><td>+1</td></tr><tr><td>-1</td><td>0</td></tr></table>	0	+1	-1	0	<table><tr><td>+1</td><td>+2</td><td>+1</td></tr><tr><td>0</td><td>0</td><td>0</td></tr><tr><td>-1</td><td>-2</td><td>-1</td></tr></table>	+1	+2	+1	0	0	0	-1	-2	-1	<table><tr><td>+1</td><td>+1</td><td>+1</td></tr><tr><td>+1</td><td>-2</td><td>+1</td></tr><tr><td>-1</td><td>-1</td><td>-1</td></tr></table>	+1	+1	+1	+1	-2	+1	-1	-1	-1
0	+1																								
-1	0																								
+1	+2	+1																							
0	0	0																							
-1	-2	-1																							
+1	+1	+1																							
+1	-2	+1																							
-1	-1	-1																							
Componente vertical	<table><tr><td>-1</td><td>0</td></tr><tr><td>0</td><td>-1</td></tr></table>	-1	0	0	-1	<table><tr><td>-1</td><td>0</td><td>+1</td></tr><tr><td>-2</td><td>0</td><td>+2</td></tr><tr><td>-1</td><td>0</td><td>+1</td></tr></table>	-1	0	+1	-2	0	+2	-1	0	+1	<table><tr><td>+1</td><td>+1</td><td>+1</td></tr><tr><td>-1</td><td>-2</td><td>+1</td></tr><tr><td>-1</td><td>+1</td><td>+1</td></tr></table>	+1	+1	+1	-1	-2	+1	-1	+1	+1
-1	0																								
0	-1																								
-1	0	+1																							
-2	0	+2																							
-1	0	+1																							
+1	+1	+1																							
-1	-2	+1																							
-1	+1	+1																							

Fonte: Richards (2013).

O filtro gradiente 2 x 2 de Roberts, por ter uma configuração quadrada, não há simetria porque não tem um pixel central, e, por isso, o deslocamento da máscara é de 1/2 pixel. As máscaras 3 x 3 deslocam-se pixel a pixel, e a detecção de bordas é mais larga do que a produzida pelos filtros 2 x 2. Isto porque o processo de detecção de bordas direcionais computa um gradiente local e porque as bordas detectadas são saturadas com valores de cinza (como uma binarização). É necessária a escolha de um valor de limiar acima do qual as bordas são aceitas. Onde o gradiente exceder o limiar, o valor do pixel é fixado no valor máximo de número digital, de acordo, isso para imagens disponibilizadas com a resolução radiométrica da imagem (8, 12 ou 16 bits, por exemplo). A escolha de um limiar baixo implica em um grande número de bordas largas, enquanto um limiar alto resulta em segmentos de bordas finas. Em geral, os programas comerciais de processamento de imagens embutem um limiar prefixado, sem dar opção ao usuário de definir um novo limiar com base na sua experiência.

As máscaras dos filtros direcionais de Roberts e Sobel não podem ser editadas, o que é possível com o filtro de Prewitt, para direções que o usuário desejar (Tabela 13.6). Exemplos da aplicação de filtros gradientes são mostrados na Figura 13.9. Note que foi detectado um grande número de bordas em todas as direções. A imagem filtrada está adicionada com um *background* de 30% da imagem original. Os resultados das aplicações de filtros direcionais exemplificados nas Figuras 13.7 e 13.9 mostram que os filtros direcionais lineares são mais eficientes na detecção de lineamentos geológicos, enquanto os filtros de primeira derivada aplicam-se melhor nas áreas urbanas e agrícolas.

Tabela 13.6. Exemplos de filtros lineares de borda.

Componente Noroeste			Componente Nordeste		
+1	+1	+1	-1	-1	+1
-1	+2	+1	-1	2	+1
-1	-1	+1	+1	0	+1

Fonte: Richards (2013).

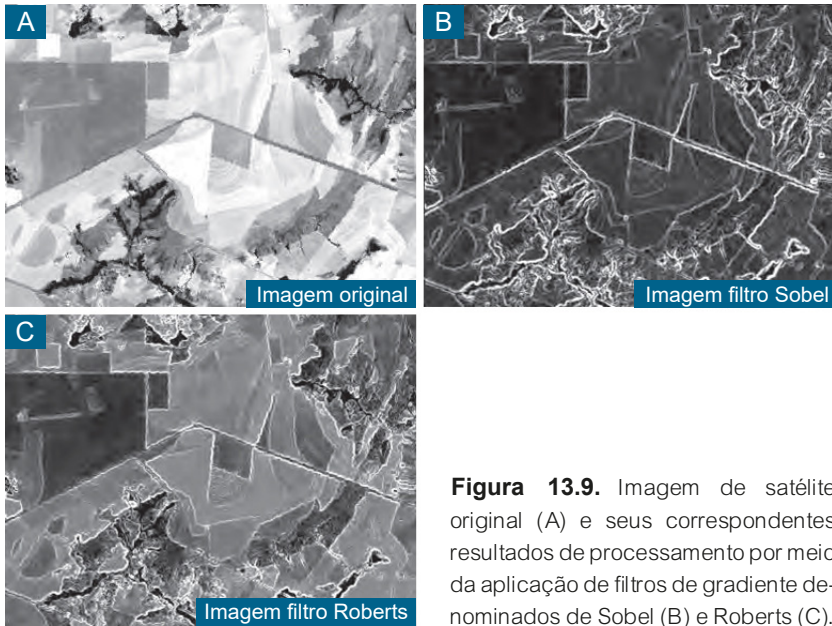


Figura 13.9. Imagem de satélite original (A) e seus correspondentes resultados de processamento por meio da aplicação de filtros de gradiente denominados de Sobel (B) e Roberts (C).

13.5. Operadores Laplacianos

Os operadores laplacianos são filtros que utilizam a segunda derivada para realçar bordas, sem considerar a direção, comportando-se como um filtro passa-alta. Esses operadores geralmente utilizam um alto valor central na máscara, cercado por valores negativos nas direções vertical e horizontal. Como eles podem enfatizar ruídos existentes nas imagens, recomenda-se passar o filtro Laplaciano após o filtro passa-baixa que elimina os ruídos.

O uso da segunda derivada para máximos e mínimos possibilita atribuir valor zero nas áreas onde o gradiente (primeira derivada) é constante, e valores positivos e negativos onde encontrar o topo e a base do declive dos valores de brilho da imagem, respectivamente. A Figura 13.10 exibe a máscara discreta de ordem 3 de um filtro Laplaciano e o resultado de sua aplicação. O somatório dos elementos da máscara é nulo. Todos os filtros Laplacianos devem ter tamanho ímpar.



Figura 13.10. Imagem de satélite original (A) realçada e a correspondente imagem realçada por meio do filtro Laplaciano (B) e a máscara correspondente do filtro.

13.6. Filtros Gaussianos

O filtro gaussiano tem a função de suavizar a imagem quase da mesma forma que o filtro de média passa-baixa, mas um pouco mais suave, preservando os contornos mais do que o da média. O filtro gaussiano usa a função Gaussiana para obter valores da máscara. Normalmente é usada a forma gaussiana 2D, circularmente simétrica (direções x , y), ou seja, o

grau de suavização será o mesmo em todas as direções (filtro isotrópico). O operador gaussiano 2D é dado pela Equação 13.9.

$$g(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2 + y^2}{2\sigma^2}} \quad (13.9)$$

O resultado será tão mais suave quanto maior for o desvio-padrão da gaussiana. Desvios-padrões maiores também fazem com que a máscara deva ser maior, ou seja, quanto maior o desvio-padrão, maior a largura do filtro. Se subtrairmos a versão do filtro gaussiano (passa-baixa) da imagem, obtêm-se imagens realçadas dos altos componentes de frequência, o que poderia ser chamado de filtro gaussiano passa-alta.

13.7. Filtros Morfológicos

Os filtros morfológicos têm como objetivo extrair objetos baseados em alguma geometria predefinida. Eles atuam nas propriedades geométricas dos sinais (níveis de cinza), utilizando dois conceitos básicos, o fechamento e a abertura. Podemos combinar essas duas propriedades com as seguintes sequências: abertura seguida de fechamento, que se denomina dilatação e, fechamento seguido de abertura, que se denomina erosão. Esses filtros podem possuir algoritmos para imagens binárias e em tons de cinza. A forma da máscara é chamada de elemento estruturante e a operação dos componentes pesos morfológicos é completamente diferente da operação de convolução. As formas da máscara podem ser quadradas, retangulares, diagonais, ou qualquer forma desejada. A escolha da forma é importante porque dela depende a mudança induzida na imagem.

O filtro morfológico de erosão provoca efeitos de redução das partes claras da imagem, gerando imagens mais escuras. O filtro morfológico de dilatação age de forma contrária, provocando efeitos de expansão das partes claras da imagem. Eles possuem aplicações na remoção do ruído e na

segmentação da imagem para classificação. Alguns exemplos de pesos estruturantes são mostrados na Tabela 13.7. Os valores dos elementos estruturantes são sempre zero ou 1.

Tabela 13.7. Exemplos de máscaras de filtros morfológicos.

<table><tr><td>0</td><td>1</td><td>0</td></tr><tr><td>1</td><td>1</td><td>1</td></tr><tr><td>0</td><td>1</td><td>0</td></tr></table>	0	1	0	1	1	1	0	1	0	<table><tr><td>1</td><td>0</td><td>1</td></tr><tr><td>0</td><td>1</td><td>0</td></tr><tr><td>1</td><td>0</td><td>1</td></tr></table>	1	0	1	0	1	0	1	0	1	<table><tr><td>0</td><td>0</td><td>0</td></tr><tr><td>1</td><td>1</td><td>1</td></tr><tr><td>0</td><td>0</td><td>0</td></tr></table>	0	0	0	1	1	1	0	0	0
0	1	0																											
1	1	1																											
0	1	0																											
1	0	1																											
0	1	0																											
1	0	1																											
0	0	0																											
1	1	1																											
0	0	0																											
<table><tr><td>0</td><td>1</td><td>0</td></tr><tr><td>0</td><td>1</td><td>0</td></tr><tr><td>0</td><td>1</td><td>0</td></tr></table>	0	1	0	0	1	0	0	1	0	<table><tr><td>1</td><td>0</td><td>0</td></tr><tr><td>0</td><td>1</td><td>0</td></tr><tr><td>0</td><td>0</td><td>1</td></tr></table>	1	0	0	0	1	0	0	0	1	<table><tr><td>0</td><td>0</td><td>1</td></tr><tr><td>0</td><td>1</td><td>0</td></tr><tr><td>1</td><td>0</td><td>0</td></tr></table>	0	0	1	0	1	0	1	0	0
0	1	0																											
0	1	0																											
0	1	0																											
1	0	0																											
0	1	0																											
0	0	1																											
0	0	1																											
0	1	0																											
1	0	0																											

Fonte: Richards (2013).

As Figuras 13.11 e 13.12 ilustram os procedimentos para filtrar uma imagem com os filtros morfológicos. O procedimento apresentado utiliza a primeira forma de máscara da Tabela 13.7 e o mesmo procedimento se repete para quaisquer uma das outras formas. A Figura 13.12C mostra os resultados da filtragem por erosão e dilatação. Na dilatação, o filtro, quando se move sobre a imagem, realiza a união de todos os pontos da imagem quando o elemento estruturante Bx intercepta X.

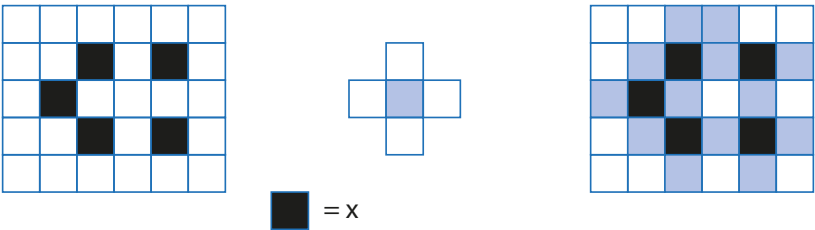


Figura 13.11. Processo de filtragem da imagem X com o filtro de dilatação.

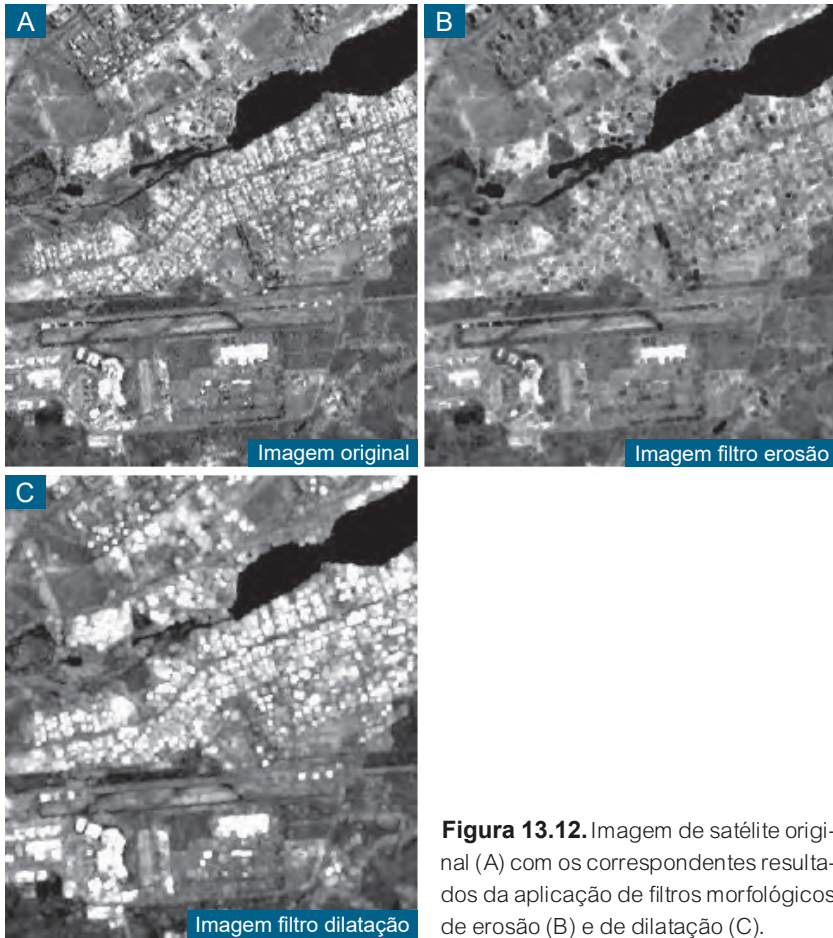


Figura 13.12. Imagem de satélite original (A) com os correspondentes resultados da aplicação de filtros morfológicos de erosão (B) e de dilatação (C).

13.8. Filtros de Textura

No contexto de imagens espectrais de sensoriamento remoto, textura é uma propriedade importante na percepção das características da rugosidade dos objetos e é expresso pela distribuição espacial das variações de tonalidade de cinza dos objetos. Nas imagens, a textura se manifesta pela variação quase periódica de valores digitais (tons) que tem origem na reflectância dos objetos e nos sombreamentos devidos à topografia do rele-

vo. O esforço é tentar reconhecer variações de padrões de formas por meio de uma análise das variações de padrões de tons de cinza. Não há uma definição geral para textura tonal e tal dificuldade é refletida pelo grande número de métodos de análise de texturas encontrados na literatura (Reed e Dubuf, 1993). Haralick et al. (1973) definiram textura a partir de medidas de uniformidade, densidade, aspereza, regularidade e intensidade, dentre outras características da imagem.

O uso da matriz de co-ocorrência está entre os métodos mais utilizados para análise de texturas. Matrizes de co-ocorrência de níveis de cinza são funções de distribuição de probabilidades conjuntas de ocorrência dos atributos de pares de pixels separados por uma determinada distância em uma direção fixa sobre uma imagem em tons de cinza (Sabino, 2004). A partir dessa matriz, podem ser calculados vários parâmetros que representam a distribuição dos valores das probabilidades de co-ocorrências de pixels. É um método baseado na abordagem de propriedades estatísticas de dados espaciais que há tempo é aplicado na exploração mineral, recursos naturais e Medicina, de onde se originou, e que mais recentemente foi estendido para aplicações em imagens de sensoriamento remoto.

A matriz de co-ocorrência contém uma grande quantidade de informações acerca da variação espacial das imagens, mas o maior problema é como extrair as informações que possam descrever as características texturais das imagens. Para isso, a abordagem estatística para a análise da textura utiliza um conjunto de medidas estatísticas locais para representar uma imagem texturizada. Essas medidas estatísticas de textura, extraídas da matriz de co-ocorrência, incluem média, variância, homogeneidade, contraste, segundo momento, correlação e outras. Cada uma dessas medidas é uma tentativa de calcular feições texturais das imagens baseadas nas propriedades da matriz de co-ocorrência. Como um método de filtragem no domínio do espaço, o seu princípio é o mesmo dos filtros apresentados, ou seja, ele descreve as imagens ou uma região de interesse nas imagens em termos das relações entre os valores dos pixels com os valores dos pixels vizinhos. A diferença é que não se busca gerar imagens realçadas para o

analista visualizar a informação desejada. A textura é um método de processamento para ser usado no processamento de segmentação e classificação de imagens (Schwartz e Pedrini, 2006).

Medidas de co-ocorrência usam uma matriz de dependência espacial de tons de cinza entre pixels vizinhos para calcular o valor da textura. Dessa forma, é uma matriz de frequência relativa cujos valores dos pixels ocorrem em duas janelas de processamento vizinhas, separadas por uma distância e uma direção específica (ENVI, 2006). Dita de outra maneira, a matriz de co-ocorrência contém uma tabulação da frequência relativa $P(i, j, h, \theta)$, considerando dois pixels vizinhos (pares) separados pela distância h na orientação θ , um pixel com tom de cinza i e o outro com tom de cinza j . O pixel i é chamado de pixel de referência e o pixel j , de pixel vizinho. A escolha do pixel vizinho pode ser em qualquer direção θ da linha, da coluna, ou das diagonais. A vizinhança pode ser de distância h de 1, 2, 3 pixels, ou qualquer outro valor. Haverá tantas matrizes de co-ocorrência quanto forem os valores de h e θ .

O exemplo mostrado na Figura 13.13 ilustra uma matriz de co-ocorrência utilizando-se um deslocamento de valor $X = 1$ na direção $Y = 0$ (horizontal), para uma máscara 3×3 . A tabulação mostra o número de ocorrências da relação entre o pixel e seus vizinhos especificados. Na primeira célula da matriz, temos a informação de quantos pixels de valor 3 têm como vizinho outro pixel de valor 3 nas imagens de tons de cinza. Na segunda célula da primeira linha, o valor corresponde ao número de pixels de valor 3 que tem, como vizinho, um pixel com valor 4, e assim sucessivamente são calculados os números de pares de pixels de distância h entre os pares de pontos localizados em i e j .

Para executar uma operação de filtragem de textura, deve-se primeiro escolher uma banda e selecionar a região de interesse. É, então, calculada a matriz de co-ocorrência da região selecionada, computando-se uma, ou simultaneamente, várias medidas estatísticas para serem usadas como medidas de textura (por exemplo, entropia, contraste, homogeneidade e mé-

dia). Os programas de processamento de imagem oferecem várias opções de medidas para caracterizar a textura da imagem.

Janela base			Janela de deslocamento			Matriz de Co-ocorrência				
							3	4	5	6
4	3	5	3	5	6	3	0	0	2	1
3	5	6	5	6	3	4	2	0	0	0
6	4	3	4	3	6	5	0	0	0	2
						6	1	1	0	0

Figura 13.13. Valores de níveis de cinza tabulados da matriz de co-ocorrência a partir dos valores dos pixels da janela base com a janela de deslocamento de distância igual a 1 na direção horizontal.

Fonte: ENVI (2006).

Uma vez calculados os valores da matriz de co-ocorrência para as medidas estatísticas selecionadas, eles podem ser comparados por meio de valores numéricos representados em gráficos para se verificar como cada medida distingue as texturas de diferentes classes de alvos. Um exemplo prático é mostrado na Figura 13.14, a partir de três segmentos de áreas extraídos de uma imagem na banda do visível e que se sabe possuírem padrões texturais distintos. Os valores estatísticos são mostrados para contraste e dissimilaridade, apenas para ilustrar o processo.

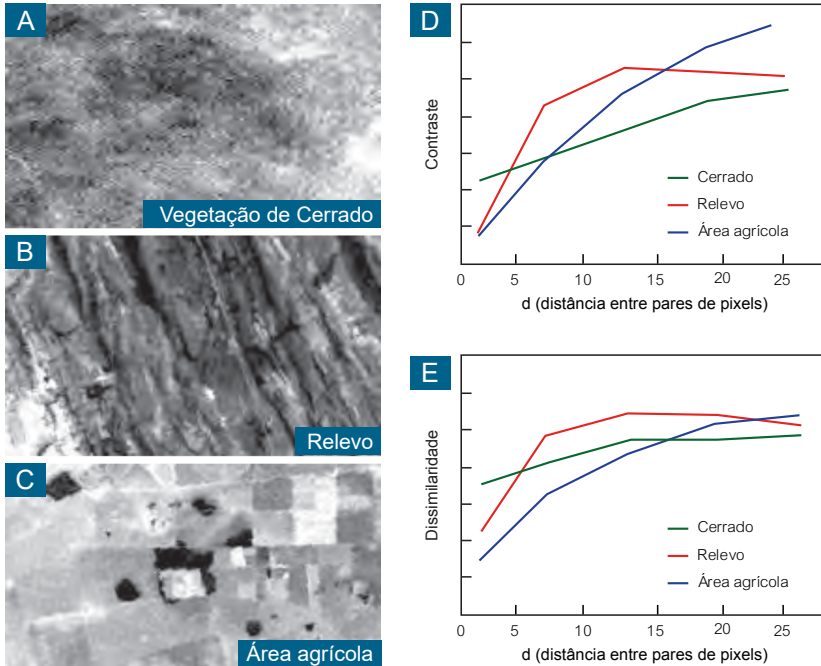


Figura 13.14. Valores das feições texturais de contraste e dissimilaridade obtidos da matriz de co-ocorrência para as classes de vegetação (A), relevo (B) e áreas de cultivo agrícolas (C) dos segmentos de imagens mostrados acima dos gráficos (D e E).

13.9. Filtragens no Domínio da Frequência

Nas seções anteriores, vimos que as imagens são representadas essencialmente no domínio do espaço e que a imagem pode ser decomposta em apenas dois componentes espaciais distintos: alta frequência espacial e baixa frequência espacial. Acontece que as imagens não são só compostas de altas e baixas frequências espaciais. Para analisar uma imagem em todos os seus componentes de frequência, é preciso usar uma transformação da imagem no domínio espacial para o domínio da frequência. Um gráfico no domínio da frequência mostra quanto do sinal reside em cada faixa de frequência. Se plotarmos, em um gráfico, as variações das frequências de

brilhos que a imagem pode conter, vamos obter uma curva complexa do tipo mostrado na Figura 13.15. Essa curva complexa, isto é, não periódica, pode ser decomposta em vários componentes de seno e cosseno que nos dá uma ideia da distribuição espacial de frequência das imagens, pois os senoides variam com uma frequência. Para decompor as imagens em seus componentes de seno e cosseno, como visto na Figura 13.15, usa-se a transformada de Fourier que tem a seguinte forma geral (Equação 13.10).

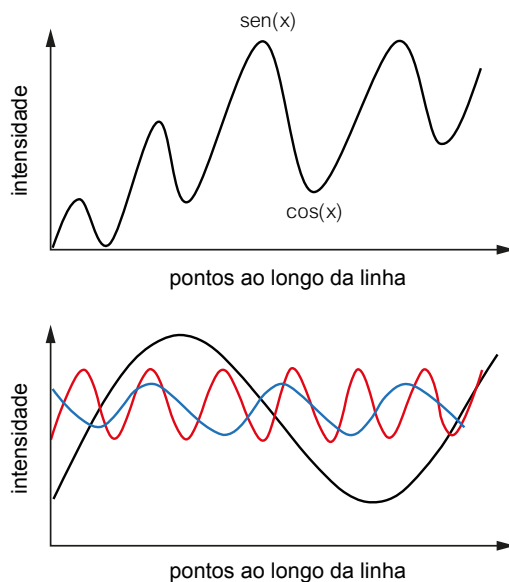


Figura 13.15. Decomposição das frequências da imagem em seus componentes seno e cosseno.

Fonte: adaptado de Crósta (1999).

$$f(x) = a_0 + a_1 \sin x + a_2 \sin 2x + \dots + b_1 \cos x + b_2 \cos 2x + \dots \quad (13.10)$$

No domínio de Fourier, cada ponto da imagem representa uma frequência única que está contida no domínio espacial da imagem. A transformada

de Fourier é mais uma das opções para filtrar as imagens, mas seu uso estende-se a várias aplicações para reconstrução e compressão de imagens, tomografia e detecção de vozes. Em processamento de imagem, a filtragem no domínio da frequência consiste em transformar a representação das imagens no espaço para o domínio da frequência, aplicando, por exemplo, a transformada de Fourier e realizando, em seguida, a filtragem desejada e depois calculando-se a transformada inversa de Fourier para obter os resultados das imagens filtradas. A realização da filtragem dos dados no domínio da frequência possibilita a alteração dos dados originais em novas informações de saída com a aplicação de filtros como o passa-baixa ou o passa-alta, permitindo a remoção de ruídos dos sinais, a suavização dos dados, o aumento do contraste ou até o realce dos detalhes de uma imagem, por exemplo, como as bordas. Matematicamente, seu entendimento é bem mais complexo do que filtrar uma imagem no domínio do espaço, mas a transformada de Fourier torna mais fácil processar ou examinar certas frequências que influenciam na estrutura geométrica das imagens. Como exemplo, pode-se citar a presença de ruídos periódicos que frequentemente são encontrados em quase todos os tipos de imagens de sensoriamento multiespectral.

Para o processamento de imagem digital, a qual é uma função bidimensional discreta, utiliza-se da transformada de Fourier discreta, que é uma versão da transformada de Fourier. A transformada discreta não contém todas as frequências das imagens, mas apenas um conjunto de frequência que é suficiente para descrever completamente o domínio espacial das imagens. Conforme demonstrado na Equação 13.10, González e Woods (2017) definiram a filtragem no domínio da frequência como a multiplicação de uma função filtro $H(u, v)$ pela função $F(u, v)$ que é a transformada de Fourier (F) da imagem que se deseja filtrar, por um processo de convolução. O teorema de convolução estabelece que a convolução de uma máscara na imagem equivale à multiplicação da transformada de Fourier pela transformada da máscara (filtragem no espectro). A utilização da transformada de Fourier em imagens facilita alguns processos de definição de filtros, reco-

nhecimento de texturas e compressão de imagens, sendo realizada usualmente em três passos:

- 1) Cada pixel da imagem $f(x,y)$ é transformado do domínio espacial para o da frequência $F(u,v)$ usando a transformada de Fourier;
- 2) Filtros são aplicados na imagem por meio de técnicas de processamento de imagens; e
- 3) Em seguida, realiza-se o processo inverso, no qual a imagem modificada no domínio da frequência é transformada para o domínio espacial, usando a transformada inversa de Fourier e gerando uma nova imagem processada.

A equação básica da filtragem é dada por (Equação 13.11):

$$g(x, y) = F [H(u, v) F(u, v)] \quad (13.11)$$

em que

F = transformada inversa de Fourier;

$H(u, v)$ = transformada discreta de Fourier; e

$F(u, v)$ = função de transferência de filtro.

Uma questão importante na transformada de Fourier é a localização das frequências da imagem após a transformação. As baixas frequências localizam-se no centro das imagens e representam as mudanças mais suaves na intensidade das imagens, enquanto as frequências mais altas localizam-se mais afastadas das imagens e representam as intensidades mais altas, como bordas de objetos. Os valores de brilho médio das imagens, isto é, o componente de frequência zero, são mostrados nos centros das imagens transformadas. Diversas opções de filtro podem ser utilizadas para remover um componente particular de frequência das imagens. A Figura 13.16 exemplifica o processo em três passos da filtragem transformada discreta de Fourier.

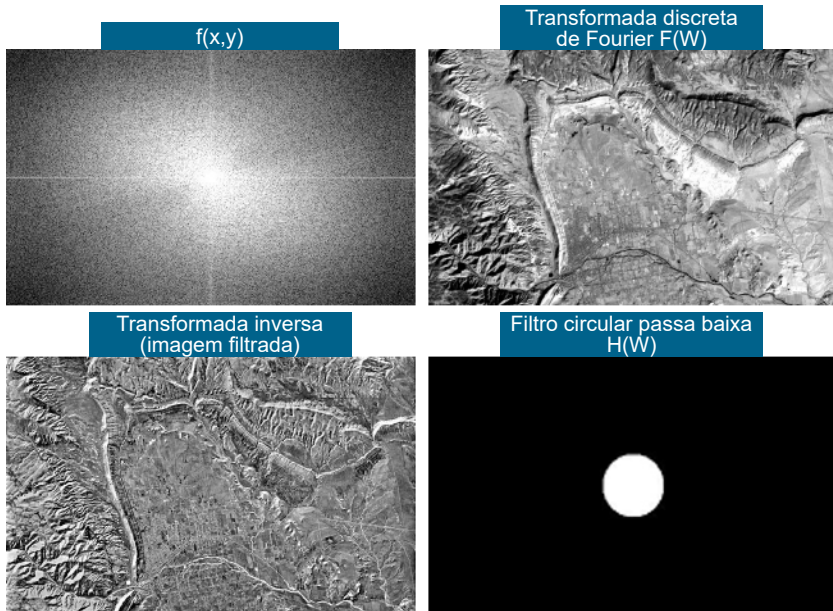


Figura 13.16. Em sentido horário, as imagens ilustram o resultado do processo de filtragem no domínio da frequência pela transformada de Fourier.

13.9.1. Amostragem

Existe uma relação que descreve o compromisso entre o domínio do espaço (tempo) e o domínio da frequência, que é chamada relação de Nyquist (Equação 13.12). Essa relação deve ser respeitada em função de ocorrerem efeitos denominados *aliasing*.

$$d_t < \frac{1}{2} f \quad (13.12)$$

em que

d_t = intervalo de amostragem no tempo; e

f = frequência no intervalo de amostragem.

Esse limite de amostragem permite que a relação de tempo e espaço possa ser realizada sem perda de informação do sinal original. Vamos realizar um exemplo com a imagem original que estamos utilizando de tamanho de 400 pixels x 400 pixels amostrados para 1.200 pixels x 1.200 pixels. Calculando a transformada de Fourier das duas imagens, temos os resultados mostrados na Figura 13.17. A relação entre as operações nos dois domínios, tempo e frequência, isto é, a convolução no domínio do tempo, corresponde à multiplicação no domínio da frequência (Equações 13.13 e 13.14).

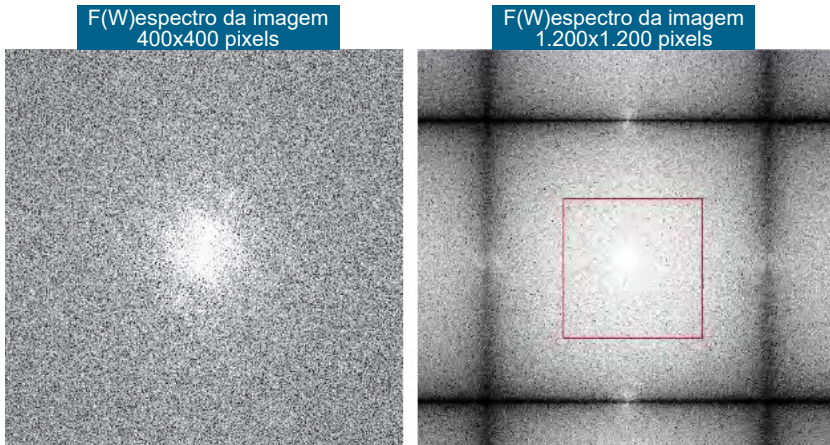


Figura 13.17. Espectros da transformada discreta de Fourier da imagem amostrada de 400 pixels para 1.200 pixels.

$$f(t) \times h(t) \quad (13.13)$$

$$F(W) \times H(W) \quad (13.14)$$

em que

$F(W)$ = espectro do sinal; e

$H(W)$ = espectro do filtro.

Alguns tipos de filtros que são usados na filtragem por transformada discreta de Fourier são apresentados a seguir, com exemplos de imagens. Filtros circulares são construídos calculando-se as distâncias dos pixels em relação ao ponto central do espectro (frequência zero). À medida que a distância aumenta, a frequência de corte também aumenta.

- Filtro banda-passante circular externo: corta as frequências entre os dois círculos e deixa passar as demais frequências. As frequências entre os círculos são de frequências baixas e muito altas. Usa-se esse filtro para obter as feições de maiores detalhes e as regionais (Figura 13.18).
- Filtro passa-alta circular: corta as frequências externas ao círculo (região cinza) e deixa passar as frequências internas ao círculo (região clara). É entendido que a região clara da transformada de Fourier está associada às altas frequências e a região cinza às baixas frequências. É necessário que o analista defina alguns parâmetros, tais como o raio em pixels do círculo. Utiliza-se esse tipo de filtro para obter as feições de detalhe na imagem (passa-alta).
- Filtro passa-baixa circular: corta as frequências internas ao círculo (região clara) e deixa passar as frequências externas ao círculo (região cinza). Nesse caso, utiliza-se esse filtro para obter as feições regionais da imagem (passa-baixa).
- Filtro banda-passante circular interno: corta as frequências internas ao círculo menor e externas ao círculo maior e deixa passar as frequências internas ao círculo maior e externa ao menor (região clara). Pode-se concluir que o filtro banda-passante deixa passar as frequências na região entre os círculos e corta as demais frequências. Utiliza-se esse filtro para obter as feições de detalhes limitados.

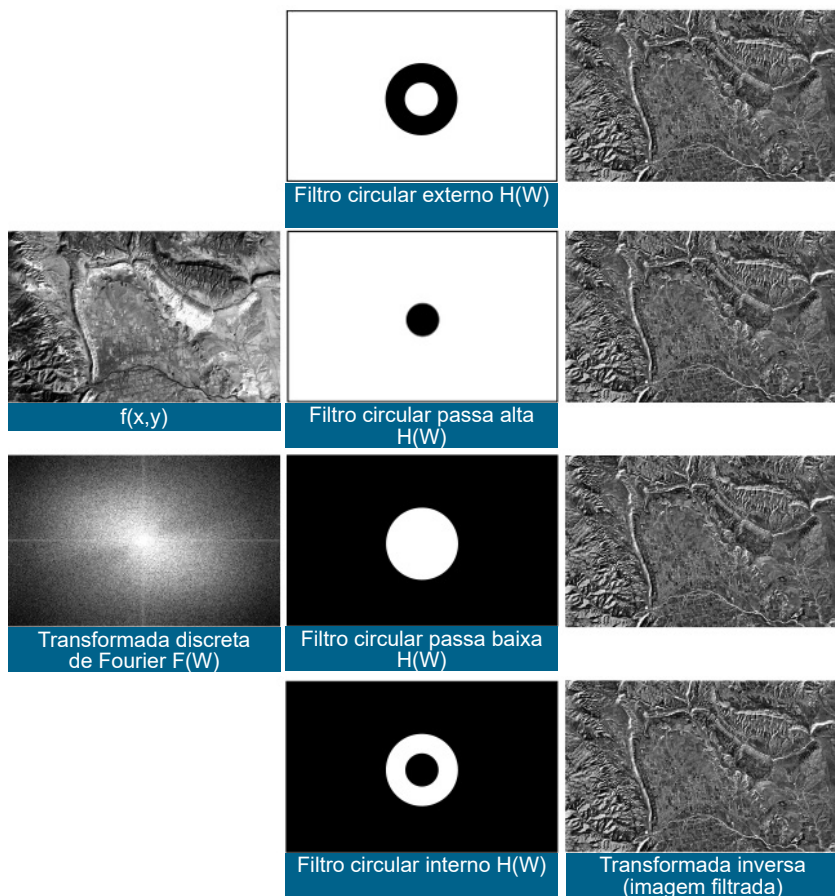


Figura 13.18. Filtragem com a transformada discreta de Fourier com uso de diferentes filtros circulares.

13.10. Referências

ENVI. ENVI User's Guide. 2006.

HARALICK, R. M.; SHANMUGAM, K. S.; DINSTEN, I. Textural features for image classification. **IEEE Transactions on Systems Man and Cybernetics**, v. 3, n. 6, p. 610-621, 1973.

REED, T. R.; DUBUF, J. M. H. A review of recent texture segmentation and feature extraction techniques. **CVGIP: Image Understanding**, v. 57, n. 3, p. 355-372, 1993.

SABINO, N. O.; USHIZIMA, D. M.; COSTA, L. F.; RIZZATI, E. G.; ZAGO, M. A. A texture approach to leukocyte recognition. **Real-Time Imaging**, v. 10, n. 4, p. 205-216, 2004.

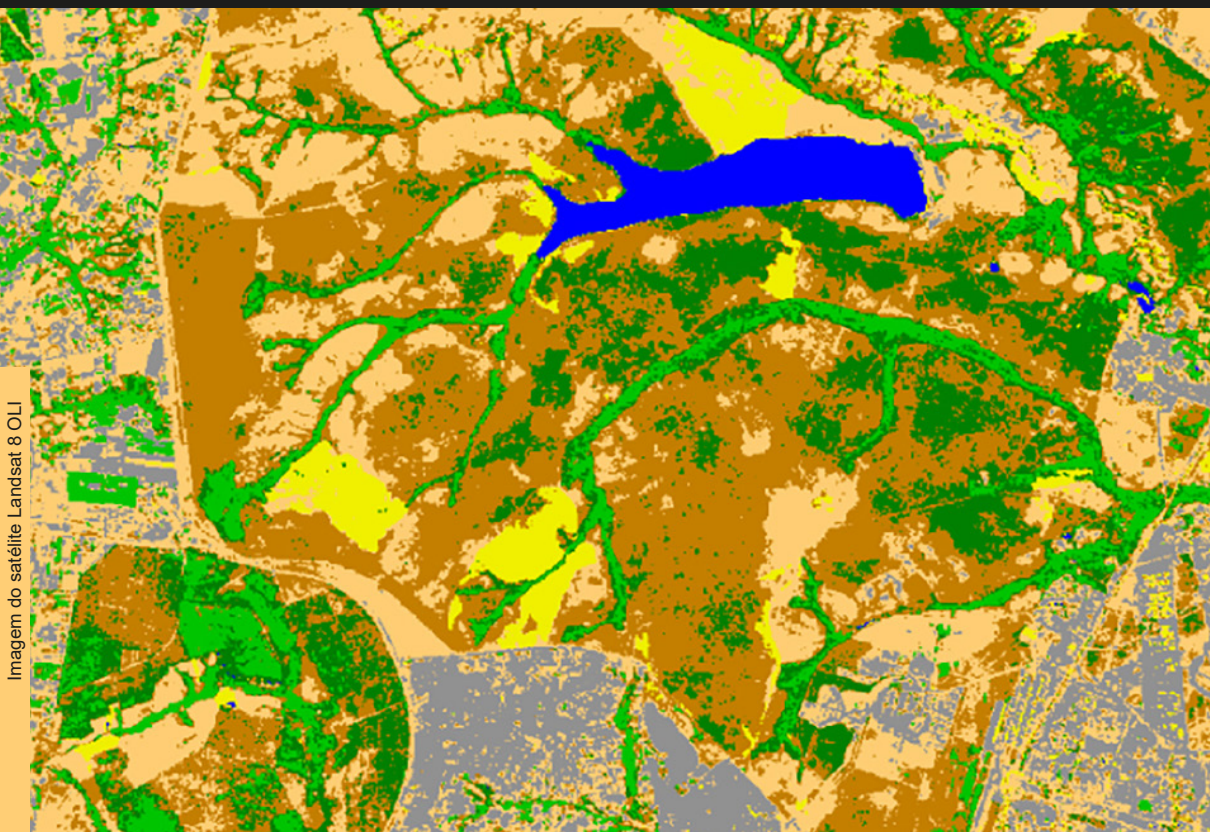
SCHWARTZ, W. R.; PEDRINI, H. Textured image segmentation based on spatial dependence using a Markov random field model. **International Conference on Image Processing**, Atlanta, GA, EUA. IEEE, 2006, p. 2449-2452.

CAPÍTULO 14

Classificação de Imagens Pixel-a-Pixel

Paulo Roberto Meneses

Edson Eyji Sano



Classificação é um processo automático de extração de informações de imagens de sensoriamento remoto para produzir mapas temáticos. Os algoritmos de classificação baseiam-se, fundamentalmente, em princípios estatísticos e determinísticos de reconhecimento de padrões. Na classificação pixel a pixel, apenas a informação espectral de cada pixel é usada para agrupar pixels em classes de objetos de interesse do mapeamento pretendido. Ela não leva em consideração a correlação entre pixels espacialmente próximos, assumindo que os pixels são independentes entre si, mesmo entre os pixels adjacentes. O resultado final da classificação de imagens é a transformação dos dados numéricos da imagem em dados descritivos, rotulados e espacialmente delimitados em áreas, que são armazenados e codificados como um simples arquivo de banda (mapa). Muitas das técnicas de classificação são interativas, homem-máquina, que decide pela seleção de amostras de pixels para treinar o classificador a reconhecer as classes dos alvos desejados. Dois parâmetros são importantes na classificação: a resolução espacial da imagem e a similaridade entre pixels ou grupos de pixels que compõem as classes relativas aos objetos no terreno.

14.1. Modelo Básico de Classificação

Na atualidade, as imagens de sensoriamento remoto constituem o meio mais usado para se produzir mapeamentos temáticos de recursos terrestres em diferentes níveis de escala. Para extrair informações das imagens por meio de métodos manuais, é necessário estabelecer metodologias de análises com regras claras e lógicas para que possam ser replicadas por qualquer analista. Essas regras estabelecem critérios para análise dos elementos que compõem a paisagem, baseados nas propriedades de cor, tonalidade, textura, estrutura, sombras e homologia. Isso é o que ditam os métodos de interpretação, herdados da experiência com a fotointerpretação. Precede a essa etapa que os fotointérpretes denominam de fotoanálise, o

domínio do conhecimento que o analista deve ter sobre todas as características técnicas dos sistemas sensores, sem o qual a fotoanálise não pode ser conduzida em sua plenitude.

A personalidade é outro fator que interfere no processo de extração da informação e no resultado. Uma pessoa que obedecer às regras do método chegará, por exemplo, a um mapa de vegetação que será mais detalhado do que o de outra pessoa que tende a ser mais generalista. Acrescente-se a isso, o demorado tempo que o intérprete gastará para processar e interpretar as várias imagens de um sensor. A classificação de imagens tem a vantagem de automatizar o processo de extração de informações das imagens, eliminar a subjetividade da interpretação humana e reduzir o esforço do analista. O resultado final da classificação é uma imagem digital que corresponde a um mapa de pixels classificados, representando, por meio de polígonos, os padrões homogêneos das classes de alvos (mapa digital temático).

A classificação de imagens multiespectrais é o processo de associar os pixels da imagem a um número finito de classes que representam os objetos do mundo real, com base nos seus valores digitais. Se um pixel satisfaz um determinado conjunto de critérios, normalmente avaliados por medidas estatísticas, o pixel é atribuído à classe de objetos que melhor satisfaz os critérios estabelecidos. Por exemplo, se os valores digitais de um pixel nas bandas dos comprimentos de onda do azul, verde, vermelho, infravermelho próximo e infravermelho de ondas curtas são, respectivamente, 15, 25, 14, 115 e 23, eles poderiam ser rotulados à classe de vegetação, pois esses valores digitais são condizentes com o padrão de comportamento de reflectância da vegetação nessas bandas. Se usarmos esses valores de pixel como amostras de treinamento, um classificador de imagens mapeará todas as áreas na imagem que possuem vegetação, apresentando, ao analista, um mapa digital de vegetação em poucos minutos.

O resultado final de um processo de classificação é uma imagem digital que constitui um mapa de pixels classificados que podem ser rotulados

como classes de objetos. Esse é o processo conhecido como classificação de imagens. A imagem classificada é, assim, uma forma de um mapa digital temático. Está implícito que somente as classes que são espectralmente distinguíveis na imagem poderão ser mapeadas no processo da classificação da imagem. O analista atua no processo de classificação apenas na etapa de seleção de amostras de pixels para treinar o classificador a reconhecer as classes de objetos desejadas. Os conjuntos de valores de pixels que identificam as classes de objetos denominam-se de padrões espectrais. Portanto, a classificação automática de imagens é nada mais do que um reconhecimento de padrões.

Reconhecimento de padrões é a ciência que procura encontrar padrões significativos nos dados da imagem que possam ser extraídos por operadores automatizados. O reconhecimento de padrões é realizado pelo cérebro humano que rotula certas texturas e cores em classes de alvos ou objetos. Esse é um processo realizado por treinamento neuro-cerebral que é desenvolvido ao longo da vida. Em um computador, o reconhecimento de padrões em imagens pode ser feito de modo mais científico, a partir de medições estatísticas derivadas dos valores espectrais dos pixels, usando uma regra de decisão para encontrar os vetores de separabilidade das classes.

Mather (1987) resumiram o problema de alocar pixels individuais às classes mais prováveis em dois modos. Se sabemos previamente que existem diferentes tipos de alvos em uma área com padrões espectralmente separáveis nas imagens, é possível estimar propriedades estatísticas que descrevem e medem cada um desses padrões. Então, cada pixel de uma imagem pode ser comparado a cada um dos padrões e ser rotulado como pertencente ao padrão com a similaridade mais alta. O outro modo é quando não há conhecimento dos tipos de alvos (padrões) que podem estar presentes na imagem. Nesse caso, são definidos arbitrariamente alguns padrões básicos e o método usado aloca e realoca os pixels da imagem a esses padrões. Ao final de cada tentativa, quando todos os pixels tiverem sido alocados, os padrões definidos podem ser alterados de acordo com

a natureza dos pixels que lhe foram associados e, se necessário, alguns desses padrões básicos podem ser descartados ou mesmo divididos. Dá-se a esses dois modos de alocar pixels às suas classes mais prováveis, os nomes de classificação supervisionada e classificação não supervisionada, respectivamente.

Aparentemente, o processo de classificar os alvos presentes na imagem é uma questão simples. Mas há uma série de fatores que podem interferir nessa tarefa. Por exemplo, muitos dos alvos na natureza podem não ser espectralmente distintos nas bandas existentes do sensor, a atmosfera pode reduzir os contrastes espectrais ou mesmo impedir a obtenção de imagens em comprimentos de onda que sofrem fortes absorções, a topografia interfere na reflectância, além de adicionar sombras, ruídos frequentemente estão presentes na imagem, e a resolução espacial torna os pixels uma mistura de alvos. Alguns desses efeitos podem ser modelados pelos algoritmos, mas nem sempre é possível modelá-los completamente. Mesmo que um modelamento completo fosse possível, a maioria dos algoritmos de classificação baseia-se nos valores digitais dos pixels, o que sabemos não ser suficiente para se obter uma perfeita classificação.

Quando o analista interpreta uma imagem, ele é capaz de separar uma vegetação vigorosa de mata de galeria de uma vegetação também vigorosa de um reflorestamento, mesmo que espectralmente ambas sejam muito similares. Elas serão confundidas por um classificador, mas o analista utiliza-se de propriedades de textura, forma e estrutura para diferenciá-las. Modelar textura e estrutura de forma satisfatória é um desafio que não foi solucionado ainda pelos classificadores para todos os tipos de imagens. A classificação de imagens deve, portanto, ser vista como um processo estatístico e probabilístico que tenta aproximar ao máximo os mapas digitais à realidade. O resultado da classificação deve ser avaliado com base no desempenho do classificador proposto e validado por critérios numéricos para estimar a sua precisão e acurácia. Sobre esse assunto, uma discussão é feita no final deste capítulo.

Os métodos de classificação podem ser divididos nas seguintes categorias de classificadores:

- 1) **Árvore de decisão:** o classificador executa a classificação em múltiplos estágios, usando uma série de decisões binárias para alocar os pixels nas classes temáticas. Cada decisão divide os pixels em duas classes baseadas em uma expressão. O analista pode dividir cada nova classe em mais duas classes baseadas em nova expressão, e assim sucessivamente. O resultado da decisão são as classes temáticas.
- 2) **Classificação supervisionada ou não supervisionada:** a principal diferença entre essas duas categorias está na presença ou não de uma fase de treinamento com a participação do analista. Na classificação supervisionada, o analista tem a tarefa de elencar as classes e selecionar amostras de pixels representativas das classes. Na classificação não supervisionada, o algoritmo não utiliza, a priori, nenhum conhecimento prévio sobre as possíveis classes existentes na imagem, isto é, não há interferência do analista.
- 3) **Classificação por região:** é um classificador de imagens que utiliza um algoritmo de segmentação, fazendo com que o contexto em que o pixel se encontra inserido seja considerado na classificação. Esse classificador leva em consideração as características do meio (regiões) ao invés de considerar apenas os valores espectrais dos pixels. É um classificador muito usado para imagens de alta resolução espacial.
- 4) **Classificação espectral:** é um classificador de imagens hiperespectrais que busca por feições espectrais de absorção dos pixels e sua posterior comparação com espectros de reflectância conhecidos, denominados de *endmembers*.

Qualquer algoritmo de classificação de imagens envolve um processo estatístico multivariado em que a medida da variável número digital (ND)

torna-se uma medida do vetor ND com K componentes (bandas). Em outras palavras, o pixel é representado por um vetor em um espaço de K dimensões (vetor K-D), como exemplificado na Figura 14.1 para um vetor com três dimensões.

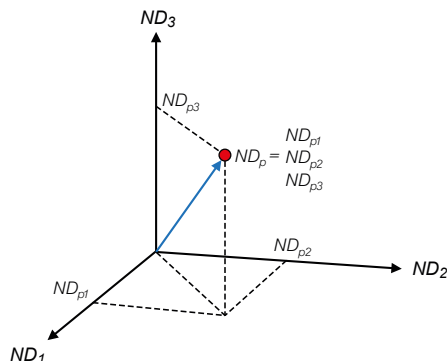


Figura 14.1. Visualização do pixel ND_p de uma imagem com três bandas K-D, como um vetor KD em três dimensões.

A noção de que um pixel é representado por um vetor no espaço de K dimensões em que normalmente, para qualquer sensor multiespectral, é maior que duas bandas, cria a noção de espaço de atributos que é um conceito importante para entender como funciona a classificação de imagens. Se plotarmos duas bandas, uma em relação à outra, podemos formar um espaço de atributos que contém as características espectrais dos pixels de treinamento para as classes de objetos que desejamos classificar. Um modo simples de visualizar esse espaço de atributo é por meio de um diagrama de dispersão.

A Figura 14.2 mostra um diagrama de dispersão de um pequeno segmento de uma imagem sobre o Parque Nacional de Brasília, em Brasília/DF, obtida nas bandas do vermelho e do infravermelho próximo. Nessa imagem, são encontradas a presença de vegetação de cerrado, mata de galeria, solo exposto e um reservatório. Olhando para o diagrama de dispersão da Figura 14.2, vemos que o espalhamento dos pontos está representado

por eixos ortogonais, na forma como se representa o espaço euclidiano. Usando a ideia do espaço euclidiano, podemos imaginar que cada uma dessas três classes irá definir três agrupamentos de pixels com um ponto central (média). O algoritmo de classificação, por meio de uma simples medida de distância euclidiana, calcula a distância de cada pixel da imagem ao centro dos agrupamentos de cada uma das três classes. Os pixels são alocados a cada classe por uma regra de decisão que é a menor distância do pixel ao centro da classe. Um algoritmo de classificação realiza automaticamente essa operação, com a vantagem de trabalhar em um espaço n -dimensional (múltiplas bandas). Veremos adiante que outras medidas de distância, chamadas de distâncias estatísticas, também podem ser empregadas na classificação de imagens.

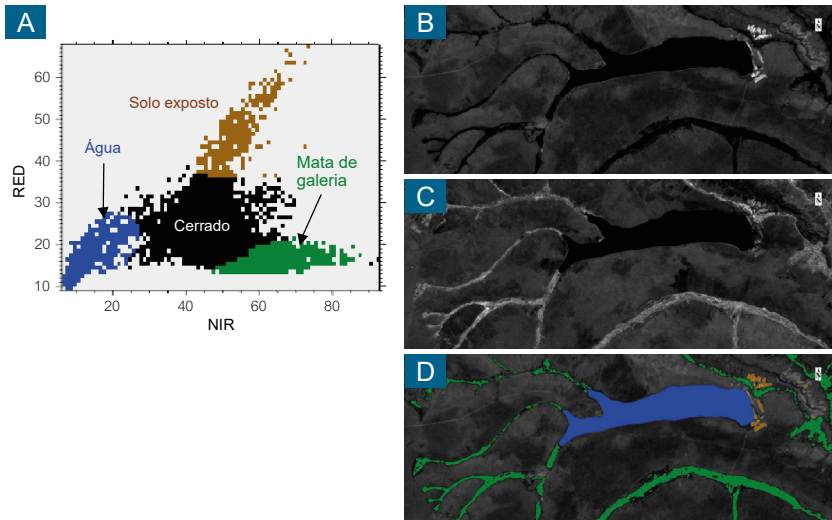


Figura 14.2. Em (A), diagrama de dispersão entre os valores digitais das bandas do vermelho (RED) e do infravermelho próximo (NIR) do satélite Landsat 8 OLI do Parque Nacional de Brasília, DF, próximo ao reservatório de Santa Maria, como um meio de reconhecimento de agrupamento de classes espectrais de alvos. Em (B) e (C), recorte de cenas do mesmo satélite nas bandas do RED e NIR, respectivamente. Em (D), o resultado da classificação de imagens mostrando as áreas cobertas com reservatório, mata de galeria e solo exposto.

14.2. Fluxo do Processo de Classificação e a Noção de Similaridade

Já foi comentado de que a precisão da classificação é afetada por alguns fatores externos que causam confusões entre as assinaturas espectrais das classes, como o sombreamento de relevo dos alvos, o espalhamento atmosférico, a mistura de objetos dentro de um único pixel e a alta correlação entre bandas. As transformações de imagens que foram vistas nos capítulos anteriores podem constituir-se em uma etapa de pré-classificação para corrigir esses efeitos que causam essas confusões espectrais, melhorando o desempenho da classificação.

O fluxo do processo de classificação depende de a classificação ser não supervisionada ou supervisionada. Na classificação não supervisionada, há apenas uma etapa em que o classificador cria diretamente o espaço de atributos com dimensão K-D. Na classificação supervisionada, o processo de classificação envolve três etapas conforme mostra a Figura 14.3.

- 1) Extração dos atributos: transformações espectrais ou espaciais das imagens originais para uma imagem processada pelas técnicas de filtragem, divisão de bandas e componentes principais, dentre outras. Essa etapa é opcional, podendo-se usar diretamente as imagens originais;
- 2) Treinamento do classificador para reconhecer as classes de alvos predefinidas, por meio da seleção de pixels de amostragem das classes, com a supervisão do analista; e
- 3) Rotulação, isto é, aplicação de limites de decisão do espaço de atributos de toda a imagem para classificar todos os pixels, a fim de constituir uma legenda do mapa digital.

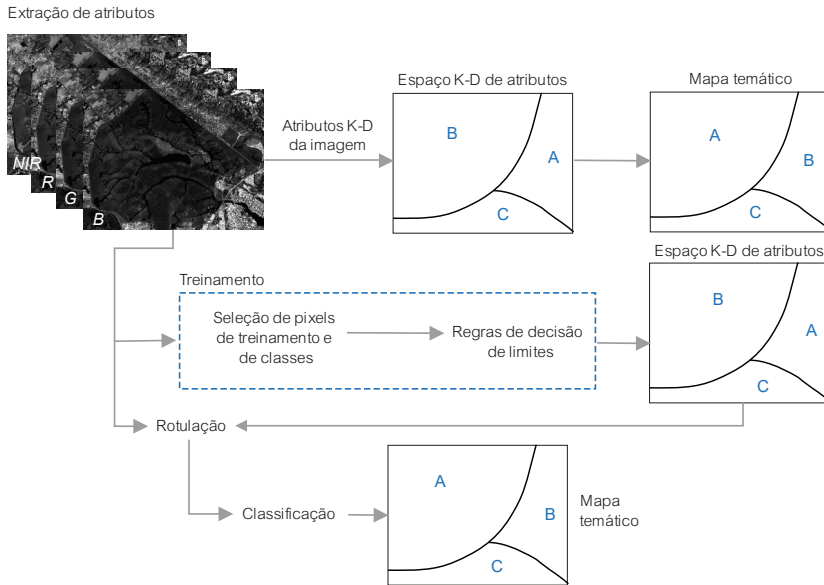


Figura 14.3. Fluxo das etapas de classificação de imagens. Fonte: adaptado de Schowengerdt (2007).

Na classificação supervisionada, o resultado da classificação é fortemente dependente da etapa de treinamento. A seleção de pixels de treinamento para representar os dados da imagem em conjuntos de classes espectrais e a identificação dessas classes como tipos de cobertura do terreno constitui-se em reconhecimento de padrões na imagem multiespectral. Padrões, em terminologia matemática, são ditos vetores de pixel que contêm os valores dos pixels arranjados na forma de coluna. Trata-se de uma simples convenção matemática, como se vê na Figura 14.4.



Figura 14.4. Padrões de classes de treinamento representados por pixels vetores em colunas.

Estatisticamente, é esperado que os vetores de pixels usados como amostras de treinamento para a classificação da imagem formam grupos em um espaço multiespectral correspondente aos tipos de classes desejadas. No exemplo da Figura 14.5, as reflectâncias dos alvos mostradas nas bandas do vermelho e infravermelho próximo configuram três agrupamentos com alta separabilidade das classes. A posição dos agrupamentos é uma relação entre as reflectâncias dos alvos nas duas bandas. A medida euclidiana de distâncias dos centros dos agrupamentos é grande, significando que os limites entre as classes são bem definidos. Na prática, os agrupamentos podem não ser tão simples e bem separados. Usualmente, nas imagens são encontrados diversos grupos. Por exemplo, a classe de vegetação poderá formar diferentes agrupamentos de acordo com os tipos de vegetação existentes na área: densa cobertura de mata de galeria; cerrado ralo, que mistura arbustos com gramíneas; pastos naturais; e áreas verdes encharcadas (veredas). Por isso, apenas uma medida do valor da média dos pixels (centro de classe) não é adequada como critério de separabilidade das classes.

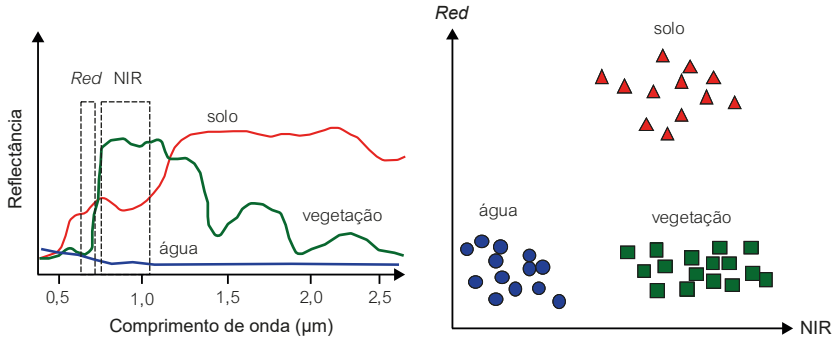


Figura 14.5. As reflectâncias nas bandas do vermelho (*Red*) e do infravermelho próximo (*NIR*) das três classes de alvos definem agrupamentos com alta separabilidade que resultará em um bom desempenho do classificador.

A questão sobre a separabilidade nada mais é do que uma questão de similaridade entre pixels ou conjunto de pixels que é fundamental nos algoritmos de classificação. Para definir a separabilidade de classes, os algoritmos estabelecem regras de decisão de limites. A questão principal é: quão bem as assinaturas de classes dos dados na imagem correspondem a classes físicas no terreno? No terreno, uma determinada classe pode ser constituída por uma associação de diversos objetos que variam em diferentes proporções no espaço ocupado pela classe. Por exemplo, uma área urbana é composta de diversos tipos de objetos, como vegetação, ruas, edifícios e terrenos vazios, resultando em assinaturas espectrais heterogêneas que se sobrepõem e que podem ser confundidas com outras classes de objetos. Consequentemente, a habilidade do classificador em desempenhar uma acurada classificação da imagem multiespectral é determinada pela extensão da sobreposição de assinaturas de classes (sobreposição de limites). Em situações como essa, é conveniente aplicar limiares para o processo de decisão dos limites de separabilidade. Pixels que possuem um valor de probabilidade para as classes abaixo do valor especificado pelo limiar não serão classificados.

Essa questão de separabilidade de classes é analisada na classificação supervisionada de imagens por três métodos distintos:

- 1) Classificadores baseados em distâncias estatísticas dos centros de classes, com o problema de que a distância aumenta sem se relacionar diretamente com a maior separabilidade das classes;
- 2) Um ótimo compromisso pode ser alcançado com um classificador do tipo máxima verossimilhança que minimiza o erro total na classificação se a nossa estimativa da distribuição de probabilidade for correta; e
- 3) Um classificador *fuzzy* aceita o fato de as assinaturas de classes se sobreporem, expressando-se como verossimilhança de membros em cada classe.

Em síntese, o desafio do classificador é tentar resolver as complexas relações entre as medidas físicas pixel a pixel das imagens multiespectrais com as classes reais no terreno.

14.3. Classificação Não Supervisionada

São classificadores que requerem pouca ou nenhuma participação do analista no processo de classificação da imagem. Há duas situações em que isso pode acontecer. Uma delas é quando não se tem conhecimentos suficientes acerca do número e natureza das classes de alvos que possam estar presentes em uma área. A outra é quando desejamos fazer apenas uma classificação exploratória da imagem para, rapidamente e sem grande esforço, saber as possíveis classes de alvos que podem existir na imagem. Não há, portanto, possibilidade de se estimar os centros das classes usando pixels de treinamento representativos de cada classe. As classes espectrais geradas na classificação não supervisionada podem ou não coincidir com as classes de interesse do analista. Trata-se, portanto, de uma operação exploratória para verificar o que é estatisticamente separável ou não, quais são as classes mais evidentes e suas relações com que se deseja discriminar. É da responsabilidade do analista, depois da classificação, as-

sociar uma rotulação às classes resultantes. O resultado da classificação é útil somente se as classes puderem ser interpretadas pelo analista.

O treinamento não supervisionado é quase que totalmente automatizado pelo computador. Ele apenas permite, ao analista, especificar alguns parâmetros que o classificador irá usar para descobrir padrões que são inerentes aos dados. Esses padrões não necessariamente correspondem às características reais da cena em termos de classes de alvos que reconhecemos. Eles são simplesmente agrupamentos de pixels com características espectrais similares. Mas, em alguns casos, pode ser mais importante identificar grupos de pixels com características espectrais similares do que separar os pixels em classes reconhecidas.

Há dois principais algoritmos de classificação não supervisionada de amplo uso em sensoriamento remoto: Isodata e K-médias. Ambos baseiam-se no agrupamento de pixels por suas similaridades, usando técnicas de distância mínima. Os parâmetros mais utilizados para definir os agrupamentos são a média, variância e covariância, pelos quais a proximidade espectral dos pixels no espaço multidimensional é definida. Os agrupamentos constituem, portanto, classes espectrais. Por isso, a classificação não supervisionada é também chamada de *clustering*, devido à técnica usada. Um algoritmo baseado em *clustering* frequentemente usa todos os pixels da imagem de entrada para iniciar a análise.

14.3.1. Isodata

Isodata significa *Iterative Self-Organizing Data Analysis Technique*. Ele é iterativo, uma vez que realiza repetidamente uma classificação inteira, produzindo uma imagem temática de saída após múltiplos cálculos estatísticos. *Self-organizing* refere-se ao modo como ela localiza os agrupamentos com mínima informação fornecida pelo usuário.

O classificador Isodata fundamenta-se no princípio de que todos os agrupamentos encontrados na imagem devem ser compactos, com todos os pixels agrupados em volta da média, ocupando uma região esférica no

espaço. Um valor de desvio-padrão deve ser tomado como medida dessa compacidade (Mather, 1987). Caso o desvio-padrão do agrupamento seja maior do que o valor estabelecido, o agrupamento que representa aquela classe de alvo é dividido. A outra condição é que os agrupamentos serão bem separados se seus centros são separados por uma distância mínima maior do que o valor de entrada que foi especificado. Se essa distância entre os centros de agrupamento for menor do que o valor de entrada, então os agrupamentos são fundidos. Divisão e fusão de agrupamento é um processo típico do Isodata, portanto, não há fixação de um número de agrupamento (classes).

O processo de classificação realiza-se em três etapas (Figura 14.6): (1) começa com um número aleatório de médias de agrupamentos arbitrários uniformemente distribuídos no espaço de dados, os quais representam uma hipotética possibilidade de centros de agrupamentos; (2) então, as distâncias mínimas são calculadas iterativamente para cada pixel da imagem e o pixel é destinado ao agrupamento mais próximo usando a técnica de distância mínima (menor distância ao centro dos agrupamentos); e (3) no fim dessa primeira sequência de alocação, a média dos valores de todos os pixels alocados a cada agrupamento arbitrário é computada para cada uma das bandas, obtendo-se as coordenadas dos centros de cada grupo. Usando esses novos centros, uma nova iteração é realizada e a cada nova iteração é recalculada as médias. Os pixels são então reclassificados em relação às novas médias.

As divisões ou fusões de grupos que irão ocorrer durante o processo de classificação dependem dos parâmetros dos padrões de entrada definidos pelo analista, que é o único momento em que ocorre uma participação interativa do analista com o processo de classificação. Alguns desses parâmetros, a exemplo do que é solicitado pelos programas de processamento de imagens, são:

- 1) Número máximo de agrupamentos a serem considerados. Uma vez que cada agrupamento é a base de uma classe, esse número é o

número máximo de classes a serem formadas. O método Isodata começa determinando N médias de agrupamentos arbitrários. Alguns agrupamentos com poucos pixels podem ser eliminados;

- 2) Limiar de convergência, que é a porcentagem máxima de pixels cujos valores de classe irão permanecer sem mudança entre iterações;
- 3) Um número mínimo de pixels em cada classe e, caso houver um número menor de pixel na classe, ela é desconsiderada e os pixels vão para a classe mais próxima;
- 4) Número de iterações; e
- 5) Se o desvio-padrão de uma classe é menor do que o limiar, então a classe é dividida em duas classes. Em alguns casos, alguns pixels podem não ser classificados, se eles não concordam com os parâmetros selecionados. A classificação será finalizada quando o número máximo de iterações for alcançado.

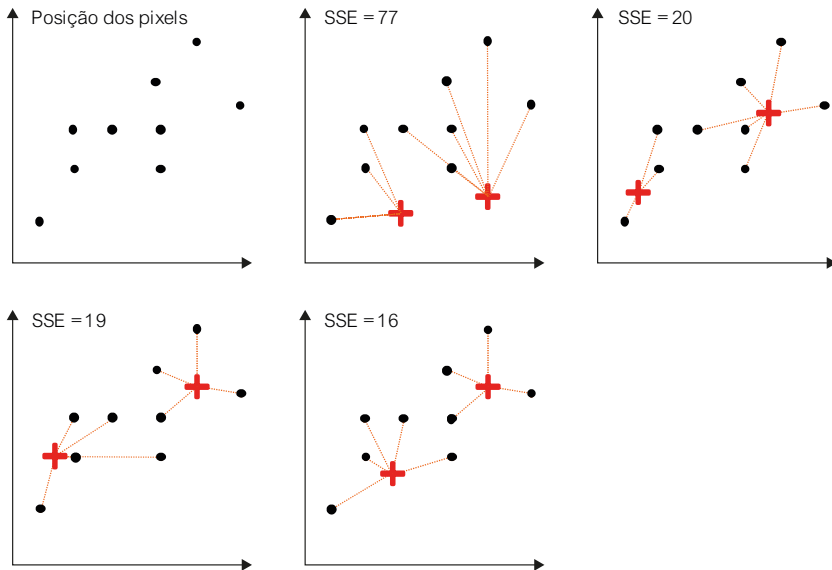


Figura 14.6. Agrupamento de pixels de dois centros de classes e os valores das medidas de similaridade soma do erro médio quadrático.

Essa classificação tem grande sucesso em descobrir os agrupamentos espectrais inerentes aos dados. Não importa onde estão os agrupamentos iniciais, desde que um número suficiente de iterações seja permitido. As desvantagens do método são: o processo pode se repetir diversas vezes, consumindo tempo; o método não leva em consideração a homogeneidade espacial dos pixels; e se o número predefinido de agrupamentos for alto, eles terão de ser fundidos, o que é um gasto desnecessário de processamento. Exemplificando: para 30 agrupamentos e 50.000 pixels, a cada iteração, os cálculos para determinar cada distância do pixel ao centro dos agrupamentos são determinados. Se forem 100 iterações, então teremos $30 \times 50.000 \times 100$, o que requer 150 milhões de cálculos por banda. Por isso, os algoritmos Isodata limitam um número máximo de agrupamentos que o usuário deve definir inicialmente.

14.3.2. K-médias

Nos classificadores não supervisionado K-médias, assume-se que o número de agrupamentos é conhecido a priori. O K-médias fixa um número de agrupamentos, enquanto o Isodata não fixa nenhum número. O objetivo do K-médias é minimizar a variabilidade interna do agrupamento, que é dimensionada pela função soma do erro médio quadrático entre cada pixel e o centro ao qual é alocado. A soma da distância do erro médio ao quadrado (*sum of squared errors* SSE) tem a seguinte fórmula (Equação 14.1):

$$SSE_{distância} = \sum xy [x - C(x)]^2 \quad (14.1)$$

em que

$C(x)$ = média do agrupamento ao qual o pixel x é alocado.

A função SSE corresponde a um critério de avaliação da qualidade do agrupamento que não se encontra no Isodata. No restante, o K-médias se-

que os mesmos processos que foram apresentados para o Isodata, mas como se fixa um número arbitrário de agrupamentos, não ocorre a divisão e fusão de grupos. A Figura 14.6 ilustra o critério de medição de similaridade para dois agrupamentos. Um resultado comparativo de classificação de imagens pelos métodos Isodata e K-médias é apresentado na Figura 14.7.

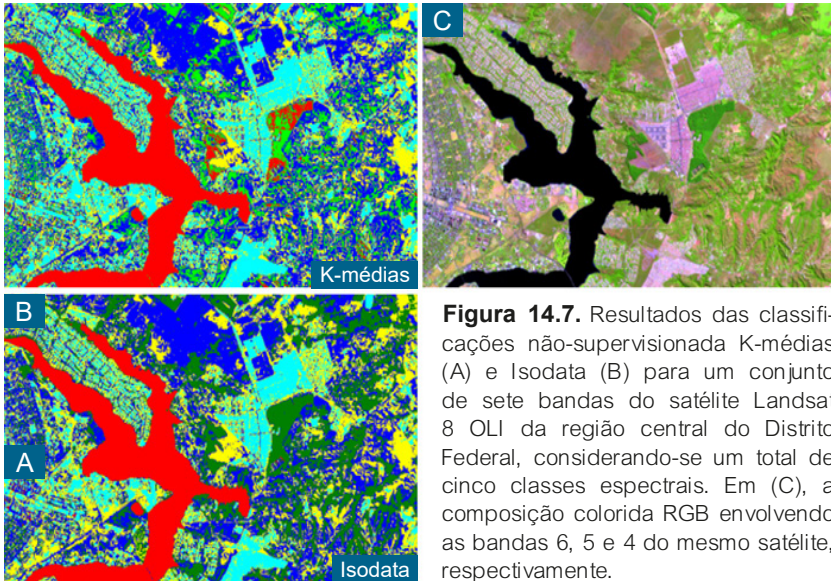


Figura 14.7. Resultados das classificações não-supervisionada K-médias (A) e Isodata (B) para um conjunto de sete bandas do satélite Landsat 8 OLI da região central do Distrito Federal, considerando-se um total de cinco classes espectrais. Em (C), a composição colorida RGB envolvendo as bandas 6, 5 e 4 do mesmo satélite, respectivamente.

14.4. Classificação Supervisionada

Ao contrário da classificação não-supervisionada, a classificação supervisionada requer conhecimentos prévios das classes de alvos, a fim de classificar a imagem nas classes de interesse pré-fixadas pelo analista. O algoritmo necessita ser treinado para poder distinguir as classes uma das outras. O treinamento supervisionado é controlado de perto pelo analista. O analista seleciona, visualmente, pequenas áreas (polígonos) de amostras de pixels (pixels vetores) contendo poucas centenas de pixels que sejam espectralmente bem representativos das classes de interesse previamente

definidas pelo usuário e por ele identificadas nas imagens. Pode-se recorrer a outras fontes de ajuda, tais como dados coletados no campo ou mapas. Portanto, é necessário o conhecimento dos dados e das classes de objetos que existem na área, antes de se iniciar o processo de classificação.

Não há nenhuma restrição quanto ao número de classes selecionadas, apenas que o analista faça uma seleção de classes bem distintas porque, caso contrário, no final da classificação, ocorrerá excesso de sobreposição entre as classes. Se o treinamento tiver uma boa precisão, as classes resultantes representam as categorias dos dados que o analista identificou originalmente. Para assegurar uma boa classificação, as imagens devem ser corrigidas dos efeitos atmosféricos, se severos, e da presença de ruídos. As características estatísticas das classes que são estimadas das amostras de treinamento dependem do método de classificação utilizado. Há vários métodos de classificação supervisionada: estatística não paramétrica (paralelepípedo, distância mínima, distância de Mahalanobis, classificação neural) e paramétrica (máxima verossimilhança).

14.4.1. Método do paralelepípedo

A classificação pelo método do paralelepípedo é o mais simples de todos e tem esse nome porque, em um espaço de dados bidimensional (duas bandas), os valores dos pixels de cada classe de treinamento são estimados por um limite superior e um limite inferior de valores digitais que identificam as dimensões das bordas do paralelepípedo. Para estimar esses limites, tomam-se os valores terminais da curva histogrâmica das amostras de treinamento em cada banda, como mostra a Figura 14.8. Os dois histogramas dessa figura representam o padrão de brilho de uma classe de treinamento visto em duas bandas (eixos x_1 e x_2). Na classificação, os pixels cujos valores caem entre os limites inferior e superior para todas as n bandas que estão sendo classificadas, são designadas àquela classe. Por meio desse processo simples, os padrões dos paralelepípedos são construídos para o

espaço de n dimensão. Na Figura 14.8, eles estão representados para um espaço bidimensional.

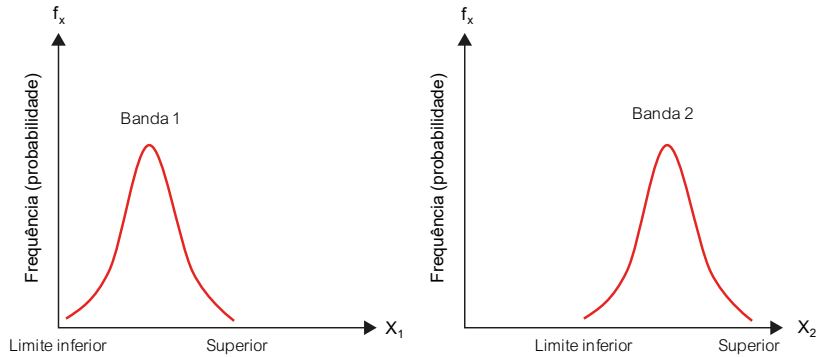


Figura 14.8. Histogramas da classe de treinamento tomada de duas bandas com os limites inferior e superior que estimam os valores de pixels da classe.

Essa simples regra de decisão pode ser estendida, acrescentando-se um limiar baseado no desvio-padrão. Nesse caso, as dimensões da classificação pelo método do paralelepípedo são estimadas pelo limiar do desvio-padrão da média de cada classe selecionada. Da mesma forma, se o valor do pixel cair acima do limiar inferior e abaixo do limiar superior para as n bandas que estão sendo classificadas, ele é alocado àquela classe, como representado na Figura 14.9 para a classe 1. Pode-se optar por um único valor de desvio-padrão para todas as classes ou múltiplos desvios-padrões. Fato comum na classificação de imagens é ocorrer pixels que são candidatos às várias classes ou que não se encaixam em nenhuma classe. No primeiro caso, como os pixels são pixels classificados, o método do paralelepípedo usa a decisão de realocá-los à primeira ou à última das classes. A segunda situação é bem mais comum de acontecer, porque são muitos "espaços vazios" entre os paralelepípedos. Na verdade, nesses espaços, estão os pixels que não caíram dentro dos paralelepípedos e que, na classificação, serão rotulados como não classificados.

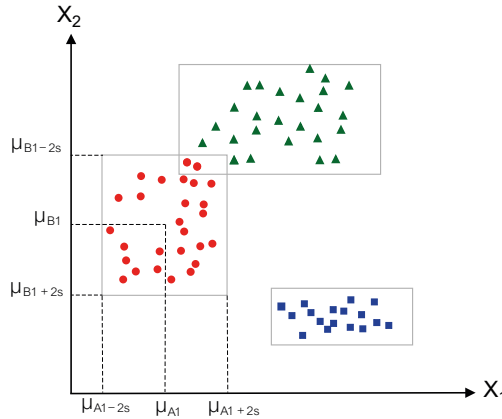


Figura 14.9. Representação das dimensões dos paralelepípedos das classes selecionadas levando em conta um limiar de desvio-padrão, em que μ = média e s = desvio-padrão.

O método do paralelepípedo tem a vantagem de ser rápido e simples, uma vez que os valores dos pixels são comparados aos limites que permanecem constantes para cada banda. Como é um método que não depende de distribuições normais, uma de suas desvantagens é que os paralelepípedos "possuem cantos" e, sendo assim, um pixel pode ser classificado mesmo estando distante espectralmente da média da classe, como mostra a Figura 14.10. Então, as classes poderão conter pixel "espectralmente impuros". A elipse na figura representa a forma da distribuição dos pixels das bandas X_1 e X_2 , forma que é definida pelo coeficiente de correlação entre as bandas.

Para a classificação, é aconselhável o uso de um maior número de bandas porque isso ajuda a aumentar a separabilidade das áreas de treinamento das classes. As amostras de treinamento devem ser homogêneas, mas também devem incluir a variabilidade que possa estar presente na classe, porque, na natureza, os alvos normalmente se compõem de associações de objetos. Também se recomenda adquirir mais de uma amostra de treinamento por classe. Na classificação pelo método do paralelepípedo, não há

necessidade de as amostras de treinamento terem uma grande quantidade de pixels, como é exigido na classificação por máxima verossimilhança. Regra geral, de 50 a 200 pixels são suficientes.

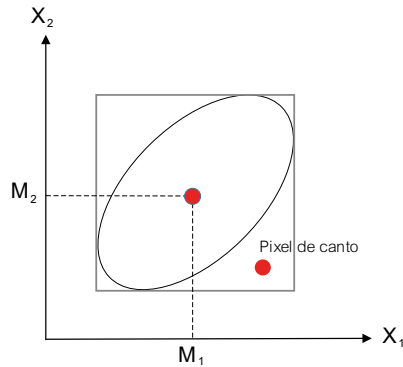


Figura 14.10. Situação em que um pixel de canto que se encontra distante da média da classe é classificado à classe no método de classificação pelo paralelepípedo.

Um resultado prático da classificação pelo método do paralelepípedo é mostrado na Figura 14.11. Uma composição colorida é recomendada para selecionar as amostras de treinamento. Foram selecionadas cinco classes a partir da análise de sete bandas da faixa óptica do sensor OLI do satélite Landsat 8. Embora o resultado possa ser considerado bom, algumas falhas podem ser notadas. Por exemplo, vários trechos da mata de galeria (cor amarela) foram classificados como reflorestamento (cor marrom), enquanto o lago (cor azul), apesar de ser aparentemente homogêneo, não foi completamente classificado. Pixels não classificados são mostrados em cor preta. Uma parte desses pixels correspondem a estradas que não foram incluídas nas classes selecionadas devido à dificuldade de amostragem.

Após a classificação, é aconselhável que o analista compare a imagem original com a imagem classificada para analisar as áreas que porventura não foram classificadas. Nesse caso, podemos refazer a classificação adquirindo mais amostras de cada classe, até que toda imagem esteja clas-

sificada, ou então fazer uma edição matricial, na qual os pixels que não foram classificados são selecionados para serem editados para a classe correspondente. Os programas normalmente incluem ferramentas de pós-classificação com várias opções para ajustar os resultados.

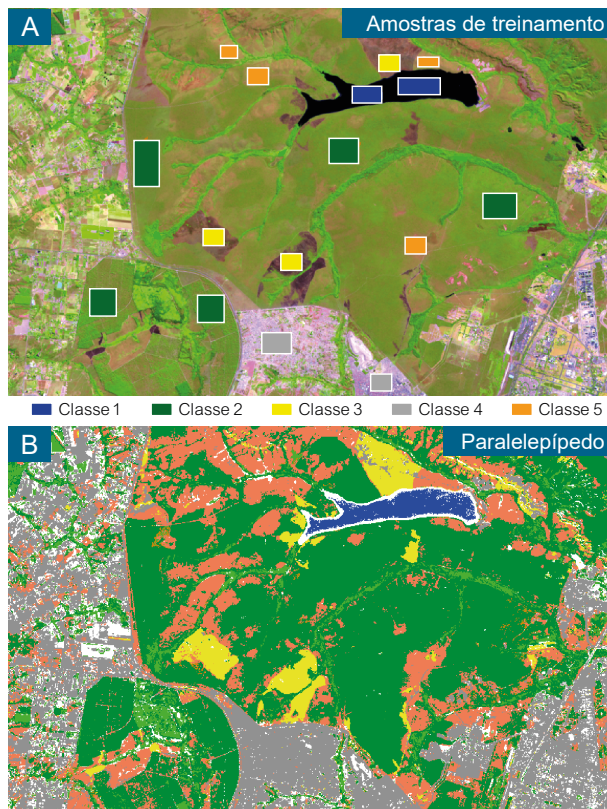


Figura 14.11. Seleção de amostras de treinamento para classificação supervisionada, envolvendo sete bandas do satélite Landsat 8 OLI, cinco classes espectrais e um limiar de desvio-padrão igual a 3,0 (A). Em (B), o resultado da classificação supervisionada pelo método do paralelepípedo. A classe mostrada na cor branca corresponde a áreas que o classificador não identificou nenhuma das cinco classes selecionadas pelo usuário.

14.4.2. Distância mínima

O método de distância mínima usa o vetor média de cada classe e calcula a distância Euclidiana de cada pixel desconhecido ao vetor média de cada classe. Todos os pixels são classificados para a classe de treinamento mais próxima, ao menos que o usuário especifique um limiar de distância. Nesse caso, alguns pixels poderão não ser classificados se eles não estiverem acima desse limiar. Medidas de distâncias como a Euclidiana apresentam o problema de não levarem em conta a sobreposição de distribuições de classes devido às suas variâncias e, assim, não são consideradas boas medidas de separabilidade. Cada pixel será incorporado a um agrupamento por meio da análise da medida de similaridade de distância Euclidiana, que é dada pela Equação 14.2.

$$D(x, n) = \sqrt{(x_i - m_i)^2} \quad (14.2)$$

em que

x_i = pixel candidato;

m_i = média das classes; e

n = número de bandas.

A Figura 14.12 ilustra a medida da distância Euclidiana entre o pixel candidato e as médias de três classes. O classificador compara a distância Euclidiana de cada pixel à média de cada classe. O pixel candidato é designado à classe com média mais próxima, isto é, à classe que apresenta a menor distância Euclidiana.

A vantagem do classificador baseado em distância mínima é de que todos os pixels encontrarão uma média à qual estarão espectralmente mais próximos, e assim, não existirão pixels não classificados. É a regra de decisão mais rápida depois do método do paralelepípedo. As desvantagens são: i) pixels que deveriam permanecer não classificados, isto é, eles não estão espectralmente próximos a nenhuma amostra dentro de limites ra-

zoáveis, serão classificados; e ii) o método não considera a variabilidade espacial da classe. Por exemplo, áreas urbanas têm uma alta variabilidade espacial, incluindo muitos pixels que tendem a estar distantes da média da assinatura da classe. Usando a regra de decisão por distância mínima, esses pixels urbanos podem ser classificados em outras classes. Por outro lado, uma classe com menor variabilidade, como um corpo de água, tende a classificar mais pixels que o esperado para a classe, pois os pixels que pertencem à classe estão normalmente mais próximos espectralmente da sua média do que aqueles das outras classes às suas médias.

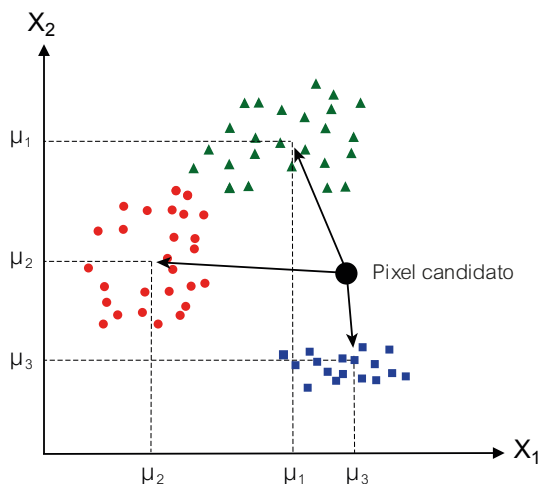


Figura 14.12. Alocação de um pixel candidato às classes pela medida da distância euclidiana.

Na Figura 14.13, é apresentado o resultado do classificador por distância mínima, usando a mesma imagem e as mesmas áreas de treinamento da Figura 14.11, para permitir a comparação do desempenho dos dois classificadores. Notam-se pequenas diferenças nas áreas classificadas, porque o classificador pelo método da distância mínima aglutina mais os polígonos das classes, reduzindo a aparência excessivamente segmentada do mapa digital obtido pela classificação pelo método do paralelepípedo.

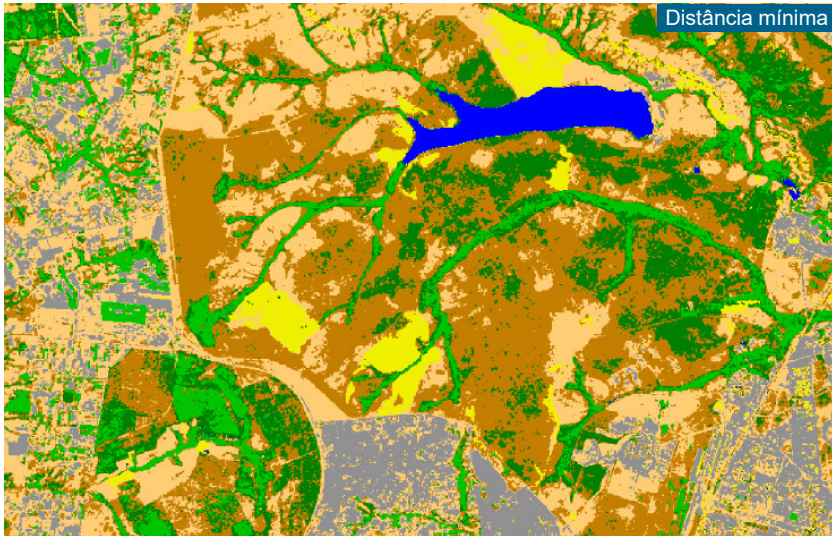


Figura 14.13. Resultado da classificação pelo método de mínima distância envolvendo seis bandas espectrais do satélite Landsat 8 OLI para as cinco classes selecionadas a partir de amostras de treinamento indicadas na composição colorida RGB da Figura 14.11.

14.4.3. Máxima verossimilhança

A classificação por máxima verossimilhança (MaxVer) é a classificação paramétrica mais comum, isto é, as estatísticas calculadas presumem que os dados das imagens são provenientes de um tipo de distribuição de probabilidade, fazendo inferências sobre os parâmetros da distribuição. Se um grande número de pixels de treinamento é disponível para cada classe, pode-se calcular o histograma da classe e usá-lo como uma aproximação da função de densidade de probabilidade contínua. Essa função de densidade de probabilidade condicional da classe, $p(f|\omega_i)$, descreve a probabilidade de um pixel, tendo um vetor de atributo f dado que o pixel está na classe ω_i . É um classificador eficiente porque as classes de treinamento são utilizadas para estimar a forma da distribuição dos pixels contidos em cada classe no espaço de n bandas, como também a localização do centro de cada

classe. Essa abordagem é diferente do que foi mostrado anteriormente em que se fundamentava no princípio de compacidade de que os pontos são distribuídos à volta do centro da média de um modo esférico. Se a distribuição dos pontos não for esférica, como em um elipsoide, é esperado que a probabilidade de um ponto pertencer ao agrupamento irá depender não somente da distância ao centro do agrupamento, mas também das dimensões e direções dos eixos do elipsoide.

Quanto a isso, vimos, no Capítulo 12, como representar a distribuição dos pixels delimitando-os no espaço de uma elipse. A orientação e as dimensões relativas dos eixos da elipse que vão definir qual será a forma, dependem da matriz de covariância. Vamos considerar, no espaço bidimensional, que as distribuições dos pixels de duas classes de treinamento assumem as formas dos elipsoides mostradas na Figura 14.14. A classe 1 tem uma orientação inclinada para a direita, significando uma orientação de covariância positiva alta entre as duas bandas. Já a classe 2 tem uma inclinação para a esquerda, indicando que a orientação da covariância é negativa. Elas têm distribuições de probabilidades distintas, e essas distribuições representam a probabilidade do pixel A pertencer à classe 1 e do pixel B pertencer à classe 2. Isso significa que a alocação dos pixels às suas classes depende da posição do pixel em relação às distribuições. No entanto, note que se fosse usado apenas a medida da distância aos centros das classes, como nos algoritmos anteriores, depararíamos com um problema. Na Figura 14.14, os centros das duas classes são indicados nos elipsoides, assim como os dois pixels A e B, de pertencerem, respectivamente, às classes 1 e 2. Apenas pela medida de distância dos pontos A e B aos centros das duas classes, o pixel B deveria pertencer à classe 1 porque a sua distância ao centro da classe 1 é menor, comparado com a distância ao centro da classe 2.

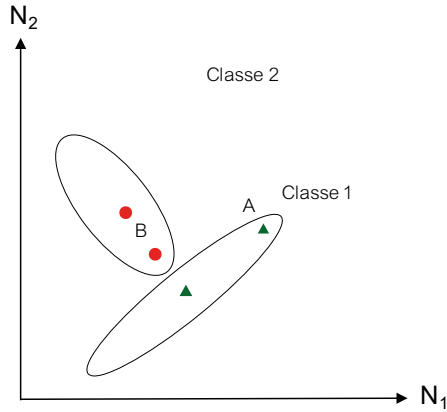


Figura 14.14. As distribuições de probabilidades representadas pelas elipses e calculadas pela covariância mostram que o ponto A pertence à classe 1 a despeito de sua distância ao centro da classe 2 ser menor.

Na classificação por máxima verossimilhança, cada pixel é destinado à classe que tem a probabilidade mais alta, isto é, a máxima verossimilhança. Isso significa que, para um pixel na posição x do espaço multiespectral, um conjunto de probabilidades são computadas e que dão as possibilidades relativas (verossimilhança) do pixel pertencer a cada uma das classes disponíveis. As áreas de treinamento são usadas para estimar a distribuição de probabilidade que descreve a chance de encontrar um pixel da classe ω_i na posição x , o que se representa por $p(x/\omega_i)$. Haverá tantas $p(x/\omega_i)$ quantas forem as classes de cobertura do terreno. Todos os pixels serão classificados, a menos que seja definido um limiar. Portanto, a $p(\omega_i/x)$ desejada e a $p(x/\omega_i)$ disponível, estimadas dos dados de treinamento, são relacionados pelo teorema de Bayes (Equação 14.3).

$$p(x) = \frac{p(\omega_i)p(x/\omega_i)}{P_x} \quad (14.3)$$

em que

$p(\omega_i/x)$ = probabilidade de a classe ocorrer na imagem; e

$p(x/\omega_i)$ = a probabilidade de x pertencer à classe ω_i .

Um pixel será destinado a uma classe se a sua probabilidade for maior do que a das outras classes. Quando ocorrerem situações ruins de classificação devido à sobreposição de classes espectrais, ou quando há conhecimento da existência de outras classes, mas sem áreas suficientes de treinamento para estimar com confiança os parâmetros de suas distribuições, aplica-se um limiar para estabelecer um limite para o processo de decisão de separabilidade. O significado de limiar é mostrado na Figura 14.15. As duas curvas representam a distribuição de probabilidade de duas classes onde há sobreposição, indicando que alguns pixels possuem probabilidades iguais de pertencer às duas classes. Nessa condição, o resultado da classificação é baixo.

O limiar é um valor real positivo que indica a porcentagem de pixels da distribuição de probabilidade de uma classe que será classificada como pertencente a essa classe. Abaixo do limiar, os pixels não são considerados como pertencentes a uma classe, ou seja, pode-se fixar uma verossimilhança máxima a partir da qual o ponto não será considerado como pertencente à classe analisada. Esses pixels normalmente estão situados nas extremidades da curva de distribuição de probabilidade. O limiar fornece, ao usuário, a possibilidade de variar o rigor da classificação. Quanto menor o limiar, mais rigorosa será a classificação. Um ótimo compromisso pode ser alcançado com o classificador de máxima verossimilhança se a estimativa da distribuição de probabilidade estiver correta. Para isso, é preciso selecionar cuidadosamente as áreas de treinamento. A desvantagem que pode apresentar é se existir uma grande dispersão dos pixels em uma classe de treinamento. Nesse caso, a matriz de covariância irá mostrar valores altos.

A Figura 14.16 mostra o resultado da classificação por máxima verossimilhança para cinco classes selecionadas e seis bandas do sensor OLI. O uso desse classificador exige que as áreas de amostragem das classes de treinamento tenham uma grande quantidade de pixels, da ordem de centenas. Por ser um classificador baseado na probabilidade, é recomendável que se tenha números de pixels aproximadamente iguais por classe de amostragem e que sejam tomadas mais de uma área de amostragem por classe.

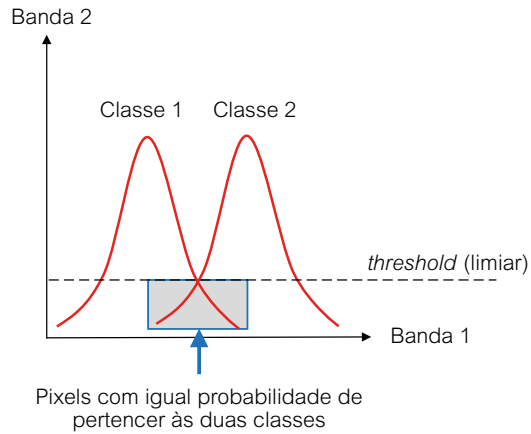


Figura 14.15. Condição em que pixels serão pobremente classificados quando se situam nas extremidades da curva de distribuição de probabilidade. O limiar remove o erro questionável de classificação.

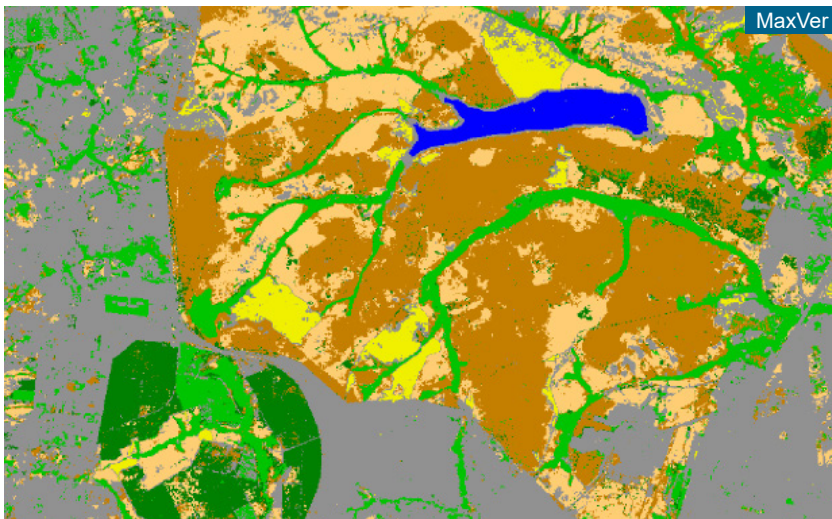


Figura 14.16. Resultado da classificação pelo método da máxima verossimilhança (MaxVer) envolvendo seis bandas espectrais do satélite Landsat 8 OLI para as cinco classes selecionadas a partir de amostras de treinamento indicadas na composição colorida RGB da Figura 14.11.

14.4.4. Distância Mahalanobis

Na Figura 14.14, apresentada na discussão do classificador por máxima verossimilhança, observamos que se a distribuição não for claramente esférica, por exemplo, um elipsoide, devemos esperar que a probabilidade de um ponto teste pertencer ao conjunto vai depender não apenas da distância ao centro do elipsoide, mas também da direção. Nessa figura, vimos que onde o elipsoide tem um eixo menor, o ponto está mais próximo do seu centro, no entanto, na direção do eixo maior, o ponto está mais distante do centro. Colocando isso em uma formulação matemática, o elipsoide que melhor representa a distribuição de probabilidade do conjunto pode ser estimado construindo a matriz de covariância das amostras de cada classe. Usando a interpretação probabilística, isso é equivalente à seleção da classe que apresenta a máxima verossimilhança.

A distância de Mahalanobis é simplesmente a medida da distância do pixel na posição x do espaço multidimensional ao centro da classe, dividida pelo comprimento do elipsoide na direção de x . Isso tem a propriedade de minimizar a distância do ponto ao centro de média. É, dessa forma, um classificador de distância sensível à direção que usa probabilidade para cada classe, o que o faz similar ao classificador por máxima verossimilhança, mas assume que toda classe tem covariâncias iguais. Isso lhe dá a vantagem de ser um classificador mais rápido do que o de máxima verossimilhança e superior ao método de distância mínima, visto anteriormente.

Para se usar a distância Mahalanobis para classificar um pixel a uma das n classes, calcula-se inicialmente a matriz de covariância com base nas amostras de treinamento das n classes. O pixel será destinado à classe na qual a distância de Mahalanobis seja a menor de todas. Usando a interpretação probabilística, isso equivale à classe que apresenta a máxima verossimilhança. A medida da distância de Mahalanobis considera que um pixel tem uma distância menor ao centro de média que um outro pixel candidato, se ele se distribuir ao longo do eixo da classe com maior variabilidade. Ele deve, então, ser classificado a essa classe. As medidas de Mahalanobis

computadas definem um espaço multidimensional cujos limites determinam o intervalo de variação tido como aceitável para que pixels candidatos possam ser classificados como pertencentes a uma das classes existentes. Outra vantagem de usar a medida de Mahalanobis é que as distâncias são calculadas em unidades de desvio-padrão a partir da média do grupo, o que faz com que a elipse englobante formada à volta do agrupamento defina a zona de um desvio-padrão. Isto permite, ao analista, atribuir uma probabilidade estatística a essa medida.

A medida de Mahalanobis é usada para identificar pixels que se distanciam bastante dos pixels da distribuição. Muitas vezes, quando as áreas de amostragens são selecionadas das classes de treinamento, alguns pixels podem contaminar a homogeneidade das classes, levando a erros nos cálculos da média e da matriz de covariância. Técnicas de regressão podem ser utilizadas para determinar se um pixel é atípico ou não no espaço multivariado, usando medidas da distância ao quadrado que classifica os pixels na base da menor das distâncias sensíveis à direção. A Figura 14.17 apresenta o resultado da classificação pela distância Mahalanobis.

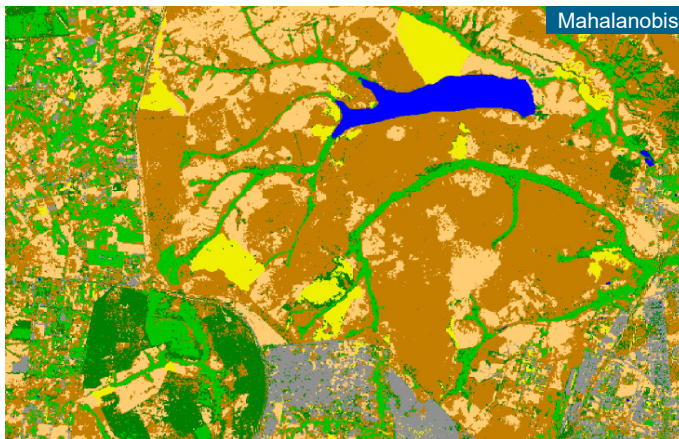


Figura 14.17. Resultado da classificação pela distância de Mahalanobis envolvendo seis bandas espectrais do satélite Landsat 8 OLI para as cinco classes selecionadas a partir de amostras de treinamento indicadas na composição colorida RGB da Figura 14.11.

14.5. Validação da Classificação

Uma das últimas etapas da classificação digital de imagens é a validação dos resultados da classificação. O que se busca nessa etapa é avaliar a acurácia da classificação. Um resultado com 100% de acurácia significa que todos os pixels da imagem foram classificados de forma correta, segundo um conjunto de dados que compõe a verdade terrestre. Um resultado com 50% de acurácia significa que, em teoria, metade dos pixels da imagem foi classificada corretamente. A acurácia depende de uma série de fatores como a complexidade do terreno, as resoluções espaciais e espectrais do sistema sensor, o próprio algoritmo de classificação utilizado, a seleção de classes utilizada no processo de classificação e o conjunto de dados que representa a verdade terrestre. Por exemplo, considere os resultados da classificação de imagens de satélite de um determinado município. Se o número de classes for de duas classes, por exemplo, terra e água, a acurácia tende a ser maior do que se composta por três classes de vegetação natural, por exemplo, formações campestre, savânica e florestal ou de três classes de uso da terra, por exemplo, culturas agrícolas, pastagens cultivadas e reflorestamento.

A acurácia é normalmente expressa em termos de índices que são calculados a partir de matrizes de erros que expressam a concordância entre a imagem classificada e o conjunto de amostras de referência. A matriz de erros compara, classe por classe, a relação entre os dados de verdade terrestre (dados de referência) e os correspondentes resultados da classificação. O número de linhas e o número de colunas dessa matriz devem ser iguais ao número de classes espectrais do estudo. Erros de omissão (exclusão) e de comissão (inclusão) de cada classe são calculados a partir dessa matriz.

Na Tabela 14.1, é mostrada uma típica matriz de erros, apresentada por Sano et al. (2010) para validar o mapa de cobertura vegetal natural e antrópica do bioma Cerrado por meio de 315 verificações de campo. As colunas da matriz referem-se à verdade terrestre, enquanto as linhas representam a interpretação de imagens. Valores em negrito da diagonal principal da

matriz correspondem ao número de pixels que foram corretamente classificados. Valores fora da diagonal correspondem a erros de omissão (EO) e comissão (EC) e que podem ser expressos na forma de percentagens.

Tabela 14.1. Erros de omissão e comissão referentes ao mapeamento de cobertura vegetal natural e antrópica do bioma Cerrado.

Interpretação de imagens	Verdade terrestre									
	Classes	1	2	3	4	5	6	Total	EO (%)	EC (%)
	1	1	0	1	0	0	0	2	75,0	50,0
	2	1	57	20	3	0	24	105	1,7	45,7
	3	1	1	33	0	0	1	36	43,1	8,3
	4	0	0	1	71	21	5	98	10,1	27,6
	5	1	0	0	5	53	1	60	29,3	11,7
	6	0	0	3	0	1	10	14	75,6	28,6
	Total	4	58	58	79	75	41	315		

Outras medidas descritivas podem ser obtidas a partir dessa matriz de confusão. Uma delas é a exatidão global, que é computada dividindo-se o número total de pontos corretamente classificados (soma dos elementos ao longo da diagonal principal = 225 pelo número total de pontos de validação = 315). Dividindo-se 225/315, tem-se uma exatidão global de 71,43%, em uma escala que varia de 0 a 100%.

O índice Kappa (K) é outro método bastante utilizado na avaliação dos resultados de classificação. Esse índice varia de 0 (concordância nula) a 1 (concordância perfeita) (Equação 14.4).

$$K = \frac{\theta_1 - \theta_2}{1 - \theta_1} \quad (14.4)$$

em que

$$\theta_1 = \frac{\sum_{i=1}^c \frac{x_{ii}}{n}}{n} \quad \theta_2 = \frac{\sum_{i=1}^c \frac{x_{i1}}{n^2}}$$

De acordo com os dados da Tabela 14.1, tem-se que $\theta_1 = 0,71$ e $\theta_2 = 0,21$, o que resulta em um índice Kappa de 0,64, considerada como sendo muito boa, de acordo com a qualidade da classificação associada aos valores da estatística Kappa que foi proposta por Landis e Koch (1977): 0–0,2 = ruim; 0,2–0,4 = razoável; 0,4–0,6 = boa; 0,6–0,8 = muito boa; e 0,8–1,0 = excelente.

Mais recentemente, Ma e Redmond (1995) propuseram o uso do coeficiente Tau (τ) para avaliar a acurácia da classificação. Os autores citaram as seguintes vantagens do uso desse coeficiente em relação à estatística Kappa: conceitualmente, é mais simples de entender; é mais simples de ser calculado; e considera o número de categorias ou classes na análise. Uma comparação relativamente detalhada do desempenho dos três métodos aqui mencionados (exatidão global, índice Kappa e índice τ) pode ser encontrada no estudo desenvolvido por Brites et al. (1996).

O tamanho (número) de amostras a serem adquiridas nesse processo de validação é outra questão delicada. Como regra geral, Congalton (1991) e Lillesand et al. (2008) sugeriram que um mínimo de 50 amostras de cada classe deveria ser considerado na construção da matriz de erros. Fitzpatrick-Lins (1981) definiu, com base na teoria de probabilidade binomial, um número mínimo de 203 observações para um percentual de exatidão de 85% e um erro admissível de 5%. No entanto, essa regra nem sempre pode ser seguida na prática, por causa das dificuldades de acesso à área de estudo e do tempo e custos envolvidos na aquisição de números grandes de pontos de validação.

14.6. Árvore de Decisão

A árvore de decisão é um tipo de algoritmo de aprendizagem de máquina supervisionada que se baseia na ideia de divisão de dados em classes homogêneas. Trata-se de um método não paramétrico que é capaz de classificar imagens contendo diferentes distribuições estatísticas. A árvore de decisão não requer conhecimentos estatísticos para a sua interpretação e aceita dados tanto categóricos como numéricos. Ela executa uma clas-

sificação em múltiplos estágios usando uma série de decisões binárias 0 e 1 para alocar o resultado dos pixels às classes temáticas. O resultado 0 é enviado para o braço “Não” e o resultado 1 para o braço “Sim” da árvore de decisão. Cada decisão divide os pixels em duas classes baseadas em uma expressão, por exemplo, $NDVI > 0$. Pode-se dividir cada nova classe em duas outras classes com base em outra expressão. Cada divisão representa um nó da árvore.

O ponto de partida de uma árvore de classificação é chamado de nó raiz, que consiste em todo conjunto de aprendizado e posiciona-se no topo da árvore. Os nós de decisão determinam o caminho dentro da árvore e essas decisões são tomadas a partir de uma amostra teste. Por fim, tem-se ainda os nós folha onde as classes são atribuídas. Podem ser usados dados de diferentes fontes e tipos. Por exemplo, pode-se usar dados de imagens multiespectrais em conjunto com modelos digitais de elevação para encontrar pixels com baixa cobertura vegetal e alta declividade, editar ou suprimir as decisões interativamente e salvar a árvore para aplicá-la para outros conjuntos de dados. O objetivo da árvore de decisão é encontrar o atributo que gera a melhor divisão dos dados, realizando testes sequenciais de fácil compreensão, cuja semântica é praticamente intuitiva. As expressões usadas são de várias categorias, conforme exemplificada na Tabela 14.2.

Tabela 14.2. Expressões usadas em árvores de decisão.

Categorias	Funções
Aritmética básica	Soma, subtração, multiplicação, divisão
Trigonométricas	Seno, cosseno, tangente, arcos, hipérboles
Operadores relacionais	AND, OR, NOR, XOR, LT, LE, EQ, NE, GE, GT
Outras funções	Expoente, raiz quadrada, logaritmo natural

Um nó é um subconjunto do conjunto de atributos e pode ser terminal ou não terminal. Um nó não terminal é um nó que se divide em nós filhos. Tal divisão é determinada por uma condição sobre o valor de um único atributo e que vai dividir os exemplos, de acordo com a condição, em outros nós.

Um nó que não se divide é chamado de nó terminal e a ele é atribuída uma classe.

14.7. Redes Neurais Artificiais

O grupo de classificadores denominado de redes neurais artificiais corresponde a uma técnica computacional que matematicamente procura simular o funcionamento de um neurônio humano. Ele possui uma estrutura complexa, interligada por elementos de processamento mais simples (neurônios) que possuem a capacidade de realizar cálculos paralelos de processamento de dados e representação de conhecimentos. As redes neurais artificiais são uma abstração da rede neural biológica que serve de modelo para aprendizado e resolução de problemas complexos.

Os neurônios humanos são formados por três elementos básicos: dendritos, que são um conjunto de terminais de entrada; corpo central, chamado de soma; e axônios, que são longos terminais de saída. A rede artificial simula essa estrutura de um neurônio por meio de um arranjo de neurônios em camadas interconectadas. Essa topologia é constituída por um conjunto de neurônios denominados “nós” computacionais, como ilustrado na Figura 14.18, que funciona como uma rede progressiva (rede *feedforward*), em que a saída de um neurônio se conecta com outro neurônio da próxima camada, no sentido da esquerda para a direita. A rede possui uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída. A complexidade dessas redes em múltiplas camadas está relacionada com a quantidade de camadas ocultas que houver e com a quantidade de neurônios que essas camadas possuem.

A camada de entrada é encarregada de receber os dados iniciais que correspondem, por exemplo, às bandas de um satélite. Cada sinal, valor ou padrão de entrada é multiplicado por um peso que corresponde a um número real que indica a sua influência na unidade de saída. O peso tem a mesma função das sinapses de um neurônio. A entrada são os atributos usados para a classificação. Em um caso mais simples, são os vetores

espectrais dos pixels de treinamento, uma banda por nó. Trata-se de uma camada de interface sem função computacional.

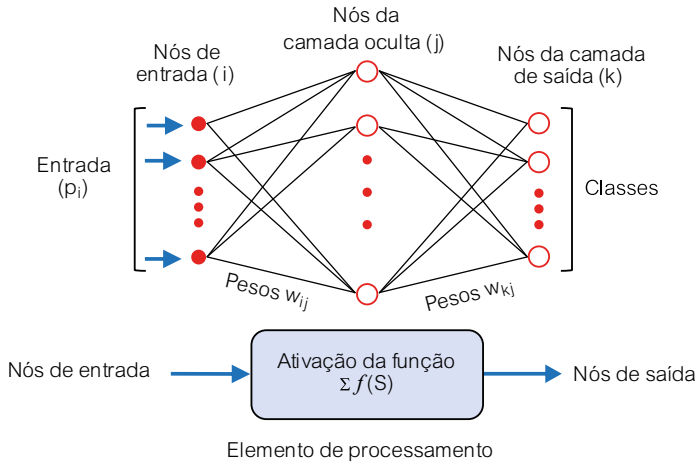


Figura 14.18. Estrutura de uma rede neural artificial de três camadas e elemento de processamento. Fonte: Schowengerdt (2007).

A camada intermediária ou camada oculta de neurônios é onde se realiza a maior parte do processamento. Essa camada contém os elementos de processamento de cada nó. Nela é feita a soma ponderada dos sinais de entrada e onde o algoritmo de transformação é ativado, podendo ser considerada a extratora de características. Se o nível de atividade exceder certo limiar, a unidade produz uma determinada resposta de saída. A camada de saída também contém elementos de processamento e sua função é apresentar o resultado da classificação (classes).

Como podemos ver na Figura 14.18, dentro de cada nó de processamento da camada oculta, tem-se o somatório e a transformação, ou seja, a função de ativação. No fluxo de processamento da rede neural, na camada oculta, cada neurônio artificial (nó j) recebe todos os valores de entradas $p(i)$ que são multiplicados pelos pesos das sinapses w_{ij} . O resultado desse somatório passa por uma função de ativação ou transferência, cujo objetivo

é limitar a amplitude de saída do neurônio, ou seja, o valor obtido no somatório é normalizado dentro de um intervalo fechado, podendo ser interpretado também como a probabilidade do resultado. No final, produz-se um valor de saída h_j (*hidden*) pela seguinte operação (Schowengardt, 2007):

$$\begin{array}{l} \text{camada oculta} \\ \text{(hidden layer)} \end{array} \quad S_j = \sum_i w_{ji} p_i \quad \text{e} \quad h_j = f(S_j)$$

Cada valor produzido nessa camada oculta pela operação acima é direcionado para cada nó k da camada de saída, onde a saída O_k (output) é calculada:

$$\begin{array}{l} \text{camada de saída} \\ \text{(output layer)} \end{array} \quad S_k = \sum_j j h_j \quad \text{e} \quad O_k = f(S_k)$$

Existem diversas funções de ativação, podendo-se destacar as funções linear, hiperbólica, gaussiana, logística sigmoidal e tangencial. A logística sigmoidal, que tem a forma de um S , é a mais usada (Equação 14.5) (Figura 14.19).

$$f(S) = \frac{1}{1 + e^{-s}} \quad (14.5)$$

Conforme visto nas classificações supervisionadas, um conjunto de amostras conhecidas é apresentado à rede neural. Inicialmente, os valores dos pesos que multiplicam os valores de entrada das amostras de treinamento são aleatórios e vão sendo ajustados iterativamente e uma saída é calculada. Após o treinamento de cada amostra, ocorre uma comparação entre a saída da rede neural e o vetor de saída desejado. Por meio dessa comparação, é feita a computação dos erros. Os pesos são reajustados em função do erro cometido e a rede é novamente submetida ao teste, calcu-

lando-se novas saídas para as amostras de treinamento. Se o erro ainda for significativo, a rede é novamente reavaliada e assim sucessivamente. No final do treinamento, obtém-se um conjunto de pesos e, com ele, a rede produz resultados bem próximos aos valores esperados. Portanto, a capacidade de discriminação de uma rede neural artificial é controlada pelos seus pesos.

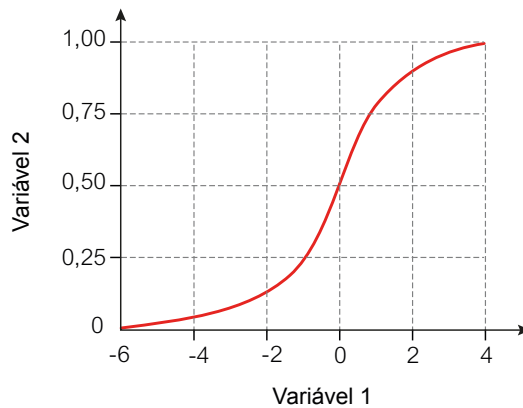


Figura 14.19. Função logarítmica sigmoideal.

A propriedade mais importante das redes neurais é a habilidade de aprendizado e, com isso, melhorar o seu desempenho. O aprendizado ocorre quando a rede neural atinge uma solução generalizada para uma classe de problemas. A rede neural artificial de treinamento supervisionado, denominada *Feed Forward BackPropagation* (FFBP), é uma das mais usadas no processamento de imagens de sensoriamento remoto. O treinamento de redes por meio desse aprendizado por *backpropagation* (retropropagação), caracteriza-se por realizar a correção dos pesos da camada de saída para a camada de entrada da rede. Com relação à correção dos pesos, existem dois parâmetros importantes: *momentum*, que mede a resistência que os neurônios apresentam ao trocar o sentido de ajuste de um peso durante o processo de treinamento da rede neural; e o *training rate*, que controla a rapidez com que os pesos são ajustados durante o processo de treinamento

da rede neural artificial. Uma taxa mais alta acelera o treinamento, mas também aumenta os riscos de oscilação ou não convergência dos resultados de treinamento. Os valores devem ficar entre 0,1 e 1,0. Existem dois tipos de aprendizado: aprendizado supervisionado, quando se utiliza um agente externo que indica a resposta desejada para o padrão de entrada para a rede; e o aprendizado não supervisionado (auto-organização), quando não existe um agente externo indicando a resposta desejada para os padrões de entrada.

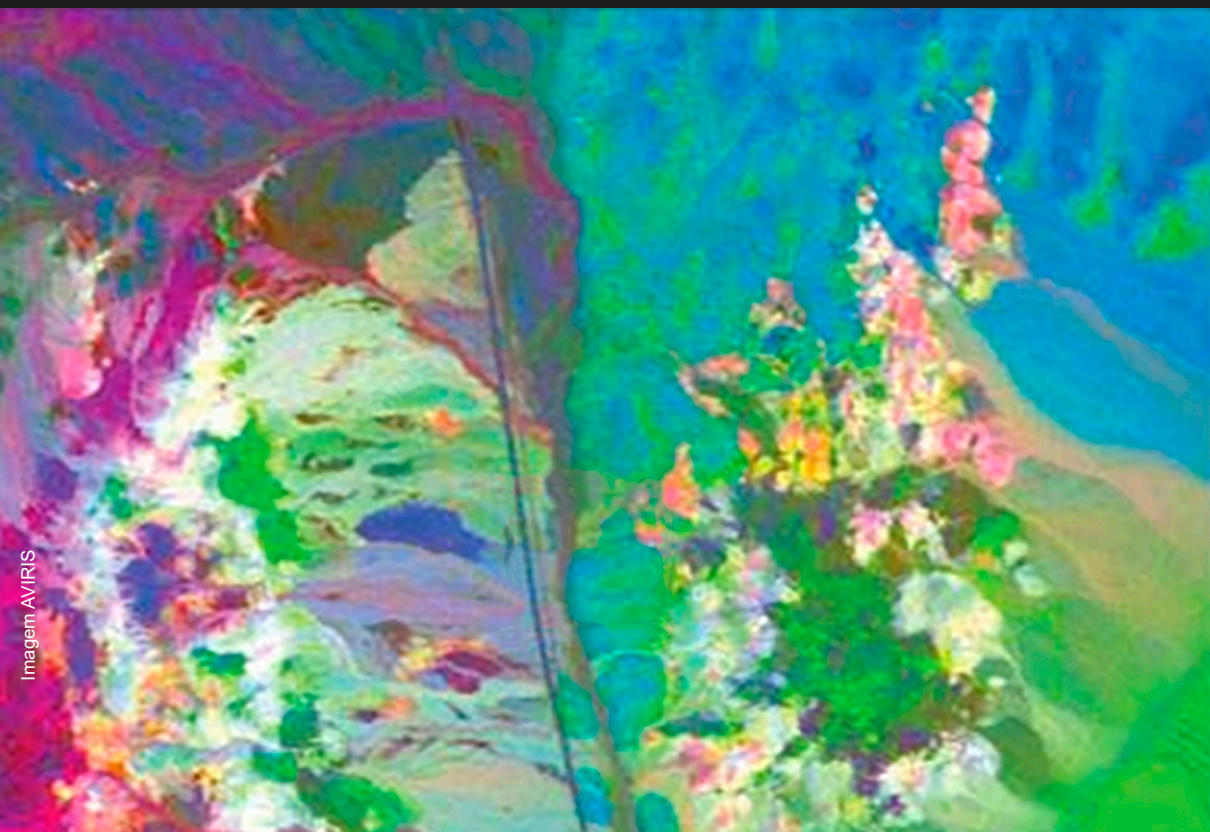
14.8. Referências

- BRITES, R. S.; SOARES, V. P.; RIBEIRO, C. A. A. S. Comparação de desempenho entre três índices de exatidão aplicados a classificações de imagens orbitais. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 8, Salvador, BA. **Anais...** São José dos Campos: INPE, 1996, p. 813-821.
- CONGALTON, R. G. A review of assessing the accuracy of classifications of remotely sensed data. **Remote Sensing of Environment**, v. 37, n. 1, p. 35-46, 1991.
- LILLESAND, T. M.; KIEFER, R. W.; CHIPMAN, J. W. **Remote Sensing and Image Interpretation**. Hoboken: John Wiley & Sons, 6a ed., 2008, 756 p.
- FITZPATRICK-LINS, K. Comparison of sampling procedures and data analysis for a land-use and land-cover map. **Photogrammetric Engineering and Remote Sensing**, v. 47, n. 3, p. 343-351, 1981.
- LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical data. **Biometrics**, v. 33, n. 1, p. 159-174, 1977.
- MA, Z.; REDMOND, R. L. Tau coefficients for accuracy assessment of classification of remote sensing data. **Photogrammetric Engineering and Remote Sensing**, v. 61, n. 4, p. 453-439, 1995.
- SANO, E.E.; ROSA, R.; BRITO, J.L.S.; FERREIRA, L.G. Land cover mapping of the tropical savanna region in Brazil. **Environmental Monitoring and Assessment**, v. 166, p. 113-124, 2010.

CAPÍTULO 15

Processamento de Dados Hiperespectrais

Gustavo Macedo de Mello Baptista



O final do século passado foi marcado por uma ruptura de um dos paradigmas mais impactantes em sensoriamento remoto: os dados hiperespectrais. Esses dados mexeram com a compreensão e os procedimentos usuais de processamento de imagens, pois as imagens saltaram da dimensão de algumas poucas bandas para a dimensão de centenas de bandas contínuas no espectro da reflectância. Como cada pixel da imagem tem o registro da radiância do alvo medida por centenas de bandas, um detalhado espectro de reflectância pode ser obtido de cada pixel. Isso consiste em um processamento estritamente no domínio espectral da imagem, com a ressalva de que reconhecimentos de padrões de pixels (classes) ou vizinhanças (correlações espaciais) não são considerados. Essa é a mudança de paradigma. Com as imagens hiperespectrais, deixamos de lado a "macro" identificação dos alvos para entrarmos na quantificação de observações de caráter mais microscópicos da matéria, apresentados nos espectros de cada pixel.

15.1. O Conceito de Espectrometria de Imageamento

Pode-se dizer que as resoluções dos sensores multiespectrais são capazes de detectar os objetos muito mais pelos detalhes de seus arranjos espaciais do que pelas suas propriedades espectrais e, por isso, são sensores para mapeamento. Nesse sentido, o uso de imagens multiespectrais baseia-se na abordagem tradicional que se pode denominar de centrada na imagem. Naturalmente, o processo de análise do dado tem similaridade com as tradicionais e antigas fotografias aéreas. Entretanto, devido à ampla largura espectral das bandas, falta, aos sensores multiespectrais, a capacidade de identificar os objetos pela natureza de suas composições químicas e biofísicas, que, de fato, é a propriedade mais importante para se discriminar um objeto de outro. Os sensores hiperespectrais vieram para cumprir essa função, criando oportunidade para o desenvolvimento de no-

vas técnicas de extração de informação dos dados. São técnicas baseadas em uma abordagem que podemos chamar de centrada no dado.

O conceito de sensoriamento remoto hiperespectral ou espectroscopia de imageamento, como é assim chamado, é de que as centenas de bandas espectrais passam a ter o valor de um espectrômetro de imageamento, que é similar a um espectrorradiômetro de laboratório e tem a habilidade de fornecer um espectro de reflectância de alta resolução para cada elemento da cena (pixel) na imagem. O espectro do pixel, analisado individualmente, é capaz de fornecer informações da composição dos materiais. A Figura 15.1 ilustra esse conceito de espectroscopia de imageamento que surgiu com a aquisição das primeiras imagens hiperespectrais pelo sensor *Airborne Visible/Infrared Imaging Spectrometer* (AVIRIS), um sensor experimental aerotransportado com 224 bandas que teve uma de suas missões em território brasileiro em 1995.

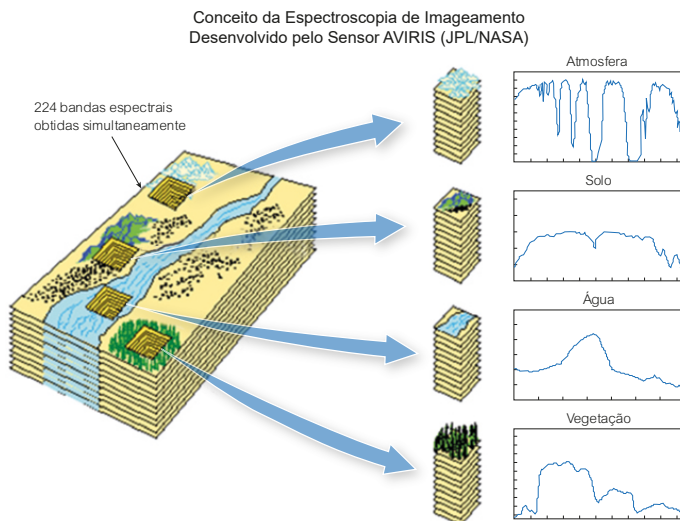


Figura 15.1. Múltiplas imagens se superpõem possibilitando a obtenção do espectro de reflectância do alvo contido no pixel. Os eixos x e y do cubo representam dados espaciais (pixels) e o eixo z representa os dados espectrais de 224 bandas no intervalo espectral de $0,4 \mu\text{m}$ a $2,5 \mu\text{m}$.

Fonte: adaptado de Baptista (2019).

O imageamento hiperespectral é realizado com a tomada de centenas de bandas contíguas e com larguras bastante estreitas, da ordem de 10 nm ou menos, equivalentes às larguras das bandas de absorção dos materiais terrestres. Essa equivalência é que possibilita identificar, nos pixels, as absorções que caracterizam as composições química, biofísica e mineralógica dos constituintes da matéria. A imagem hiperespectral não é interpretada pelo analista nos padrões de interpretação usados para as imagens multiespectrais. A análise dos dados hiperespectrais exige o uso de algoritmos de processamento muito sofisticados e que foram desenvolvidos particularmente para essas imagens, muito embora possam eles ser aplicados em imagens multiespectrais, pelo fato de a espectrometria de imageamento ser do tipo *pushbroom*, semelhante à usada pelos sensores multiespectrais.

Basicamente, o processamento dessas imagens hiperespectrais conduz à obtenção de um espectro de reflectância de cada pixel e a interpretação desse espectro exige um conhecimento profundo do comportamento espectral dos materiais pelo analista. Para isso, recomenda-se a consulta ao livro “Reflectância dos Materiais Terrestres: Análise e Interpretação”, editado pelos professores Paulo Meneses, Tati de Almeida e Gustavo Baptista da Universidade de Brasília (UnB).

15.2. Correção Atmosférica dos Dados

Um dos aspectos mais controversos do processamento de imagens hiperespectrais refere-se ao pré-processamento para correção dos efeitos atmosféricos, visando a transformação ou conversão dos dados de radiância para reflectância de superfície. Os algoritmos que fazem essas operações baseiam-se, normalmente, em modelos de transferência radiativa ou de radiação solar. Nesses modelos, busca-se reduzir os efeitos que o espalhamento e a absorção dos gases atmosféricos causam nos dados para que se possa obter um espectro puro de reflectância do pixel.

De forma bem simplificada, o espalhamento é um desvio de trajetória que elementos atmosféricos causam na radiação eletromagnética (REM),

tanto no sentido descendente (irradiância) como no sentido ascendente (radiância). Esses elementos podem ser gases ou aerossóis, estes entendidos como partículas não gasosas presentes na atmosfera. As absorções que ocorrem na faixa do espectro óptico refletido (0,4 μm a 2,5 μm) reduzem acentuadamente a intensidade da REM incidente e acabam obliterando as feições de absorção dos objetos que queremos identificar ou mesmo quantificar. Os detalhes desse assunto vimos no Capítulo 6, com figuras que mostram as regiões espectrais de absorção causadas pelos principais gases atmosféricos na faixa do espectro óptico refletido, aos quais chamamos de gases de efeito estufa. A quantidade de radiação que chega ao topo da atmosfera é maior que a quantidade medida no nível do mar e essa diferença se dá basicamente pela absorção de gases como O_3 , CO_2 , vapor de água e outros. O espalhamento e absorção ocorrem concomitantemente, mas a maioria dos algoritmos de correção dos efeitos atmosféricos trata os processos de forma separada, integrando-os posteriormente.

Outro problema são as informações que devem ser inseridas nos algoritmos para rodá-los, principalmente quanto à visibilidade da atmosfera no momento da tomada da imagem. A visibilidade, grosso modo, é entendida como o quanto se pode ver através da atmosfera. Para isso, alguns fatores são importantes para sua determinação. Eles incluem as propriedades ópticas da atmosfera, a quantidade de luz, sua distribuição, além dos aspectos inerentes ao objeto a ser observado, bem como as características dos sistemas sensores. Porém, normalmente esses dados não estão tão facilmente disponíveis para o analista. Quando se vai fazer a correção atmosférica da imagem, acaba-se utilizando valores *default* que o *software* sugere para visibilidade.

Outro aspecto salientado por B. C. Gao e colaboradores, autores do algoritmo de correção atmosférica ATREM, é que, por causa da variabilidade da concentração do vapor de água em função da hora de obtenção da imagem e da altitude do sensor, não é possível remover completamente os seus efeitos no procedimento de transformação dos dados de radiância para reflectância, utilizando os modelos de correção atmosférica atualmente em uso. Porém, criou-se um dogma de que não se pode utilizar dados hi-

perespectrais sem antes corrigir os efeitos atmosféricos. Estudos recentes apontam que problemas com os algoritmos de correção atmosférica podem mascarar resultados, principalmente quando as feições de absorção de água em 1,4 μm e 1,9 μm não são bem removidas. Alguns estudos comparativos na aplicação de índices espectrais para a determinação de relações mineralógicas em solos, tanto em dados corrigidos como em dados brutos de radiância, não apresentaram diferenças significativas quando comparados com os dados de reflectância obtidos analiticamente em laboratório.

Os algoritmos mais usuais disponíveis para correção atmosférica são os métodos de Green, *Atmospheric Removal Program* (ATREM), *Fast Line-of-sight Atmospheric Analysis of Hypercubes* (FLAASH) e ACORN. O método de Green, desenvolvido por Robert Green, investigador responsável pelo sistema AVIRIS, opera no sistema operacional Unix e o seu tempo de processamento é muito longo. O ATREM foi desenvolvido pela Universidade de Colorado e é bem mais rápido em termos de processamento e opera no sistema operacional Windows. Os algoritmos FLAASH e ACORN são de processamentos rápidos e estão disponíveis comercialmente. Praticamente todos eles se baseiam em ajustes do tipo mínimos quadrados associados a um modelo de transferência radiativa como apresenta a equação abaixo (Equação 15.1):

$$Lo(\lambda) = Lsun_{\lambda} \times T_{\lambda} \times R_{\lambda} \times \cos \theta + Lpath_{\lambda} \quad (15.1)$$

em que

$Lo(\lambda)$ = radiância observada no nível do sensor;

$Lsun_{\lambda}$ = radiância solar acima da atmosfera;

$T(\lambda)$ = transmitância total atmosférica;

R_{λ} = reflectância de superfície;

$\cos \theta$ = ângulo de incidência; e

$Lpath_{\lambda}$ = radiância de trajetória espalhada.

Na Figura 15.2, são apresentados os resultados da aplicação desses diferentes métodos de correção atmosférica a partir do espectro de reflectância extraído de um pixel da cena AVIRIS do município de São João D'Aliação/GO, obtida no âmbito da missão Smoke, Clouds, and Radiation-Brazil (SCAR-B) de 1995. O espectro A representa o método de correção atmosférica Green, o B representa o ATREM e o C, o FLAASH.

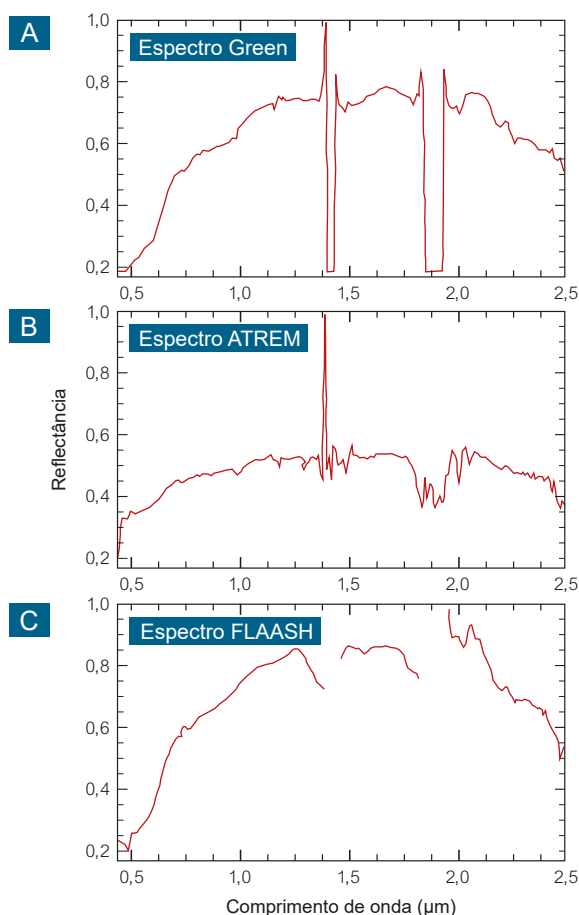


Figura 15.2. Espectros de reflectância de um pixel de solo extraído de uma imagem AVIRIS após correção atmosférica pelos métodos de Green (A), ATREM (B) e FLAASH (C).

Os espectros obtidos pelo método de Green e FLAASH removem as bandas de absorção da água localizadas em 1,4 μm e 1,9 μm , enquanto o ATREM, na tentativa de removê-las, acaba criando um ruído nesses dois locais. As feições de absorção da caulinita (2,05 μm) e a da gibbsita (2,265 μm) aparecem nos três espectros, porém, no método de Green, elas se destacam, seguido do FLAASH, enquanto o ATREM apresenta um espectro mais ruidoso, dificultando a identificação direta da gibbsita. As absorções do Fe_2O_3 no visível e infravermelho próximo aparecem nos três espectros, mas a pior visualização é no espectro obtido pelo método ATREM, seguido do Green, enquanto a melhor visualização aparece no FLAASH. Essas constatações verificadas nesses espectros não devem ser adotadas como regra geral.

A maioria dos trabalhos utiliza a correção atmosférica como uma forma de transformar os dados brutos de radiância para dados de reflectância de superfície, a fim de facilitar a identificação das feições de absorção nos espectros. Porém, existem alguns procedimentos de pré-processamento que permitem essa transformação sem alterar os efeitos de espalhamento e absorção, mantendo, assim, a integridade dos dados. Vamos tratar rapidamente de três métodos disponíveis para esse tipo de calibração visando a transformação dos dados em reflectância aparente: *Flat Field*; *Internal Average Relative Reflectance* (IARR); e *Empirical Line*. Maiores detalhes sobre essas transformações podem ser encontradas no Capítulo 6.

O *Flat Field* consiste em um método simples de normalização dos dados de radiância para reflectância aparente, usando amostras de uma área plana sem grandes variações de reflectância, por exemplo, areias de uma praia, ou seja, a mais homogênea possível ou a mais uniforme espectralmente. O espectro médio de radiância dessa área é entendido como sendo composto basicamente de efeitos atmosféricos e da irradiância solar e, para a sua correção, divide-se o espectro de cada pixel pelo espectro médio. O resultado dessa divisão é a reflectância aparente e que pode ser comparada aos *endmembers* de bibliotecas espectrais.

O IARR é um método de normalização muito semelhante ao *Flat Field*. Parte do pressuposto de que há ausência na cena de uma área espectralmente uniforme e brilhante e, portanto, ele toma o espectro médio de toda a cena e o adota como referência, dividindo todos os pixels da cena por ele. Ele tem apresentado bons resultados em regiões secas e com pouca vegetação, o que o torna mais problemático para uso em regiões intertropicais.

Já o *Empirical Line* é mais complexo, sendo utilizado para coincidir os dados da imagem com os *endmembers* coletados no campo. Portanto, ele requer um conhecimento da região imageada e que dois ou mais alvos sejam identificados na imagem e seus espectros obtidos *in situ*. A partir de regressão linear dos espectros de campo sobre os da imagem, determinam-se os ganhos e *offsets* e, a partir de uma transformação linear, os dados de radiância são convertidos para reflectância aparente, banda a banda, pixel a pixel.

15.3. Bibliotecas Espectrais

No processamento de dados hiperespectrais, o principal objetivo é a busca das feições espectrais de absorção dos objetos a serem investigados e a sua comparação com espectros de reflectância conhecidos, que passam a ser chamados de *endmembers*, termo que é atribuído às referências que serão investigadas nas imagens. Por isso, no processamento de imagens hiperespectrais, o analista deve estar familiarizado com os conceitos e mecanismos de manipulação de bibliotecas espectrais, que são arquivos contendo espectros de reflectância de minerais de solos e de rochas, de vegetação, água e outros, obtidos em campo ou em laboratório. Por exemplo, se procuramos nas imagens pela existência no solo ou rochas de pixels de caulinita, devemos primeiro investigar quais são as feições de absorção diagnóstica deste mineral. Na caulinita ($\text{Al}_2\text{Si}_2\text{O}_5(\text{OH})_4$), a feição de absorção diagnóstica ocorre em $2,2\ \mu\text{m}$ devido à ligação molecular Al-OH ($\nu_1 + \nu_3$), que é caracterizada por uma absorção dupla. Como se vê na Figura 15.3, a absorção em $1,4\ \mu\text{m}$ da hidroxila $2\nu_1$ acaba se misturando

à absorção da água em 1,9 μm e, portanto, não deve ser investigada. Esse espectro é considerado um *endmember*, pois é a referência que deverá ser utilizada para identificar nas imagens os pixels que apresentam em seus espectros a absorção diagnóstica da caulinita.

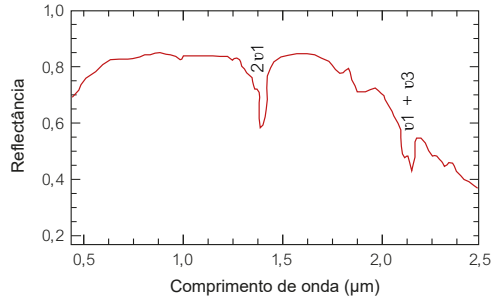


Figura 15.3. Espectro de reflectância da caulinita e com indicação de feição de absorção diagnóstica em 2,2 μm (CM9 da biblioteca do USGS).

Fonte: Clark et al. (1990).

As Figuras 15.4 a 15.6 mostram as simulações de cinco espectros de solos reamostrados para as imagens Landsat e Humap. Todos os cinco espectros contêm as duas feições de absorção típicas da água, em 1,4 μm (1) e 1,9 μm (2), com maior ou menor intensidade. Em 2,2 μm (3), há a mesma dupla feição de absorção que vimos na Figura 15.3 e que indica presença da caulinita. Em 2,265 μm (4), encontramos outra feição típica de solos tropicais decorrente da presença de gibbsita ($\text{Al}_2\text{O}_3 \cdot 3\text{H}_2\text{O}$), e que é identificada nos espectros 1, 2 e 3. As ombreiras que aparecem por volta de 0,5 μm até 1,0 μm são decorrentes da presença de sesquióxidos de ferro, no caso, hematita e goethita. Esses espectros apresentam 419 pontos de medidas, o que, em um sensor hiperespectral, corresponderia a 419 bandas.

Visando simular ou reamostrar esses espectros para as respostas de ganho dos detectores do sensor ETM+ do satélite Landsat 7, obtiveram-se os espectros mostrados na Figura 15.5 que, como se vê, obliteram as feições de absorções destacadas na Figura 15.4. Isso decorre do fato de que as seis bandas do Landsat 7 são largas e não são contínuas ao longo do

intervalo espectral do visível ao infravermelho de ondas curtas. O impedimento da atmosfera de permitir que qualquer sensor possa ter bandas em torno de $1,4\ \mu\text{m}$ e $1,9\ \mu\text{m}$ elimina, nos espectros reamostrados, as feições de absorções decorrentes da presença de água (1 e 2). Outro problema verificado é que, como as absorções da caulinita e da gibbsita ocorrem ambas na banda 7 do sensor ETM+ ($2,08\ \mu\text{m}$ – $2,35\ \mu\text{m}$), suas identificações tornam-se impossíveis. Os espectros na Figura 15.5 apresentam seis pontos de medidas de reflectância correspondentes às seis bandas do sensor ETM+.

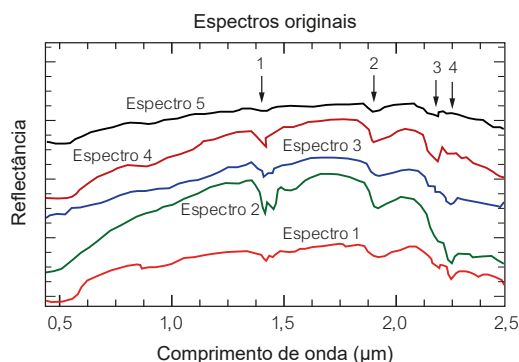


Figura 15.4. Espectros de solos obtidos com o espectrorradiômetro GER-IRIS da Embrapa Cerrados.

Cortesia: José da Silva Madeira Netto.

Convém ressaltar que outros sistemas além dos multispectrais podem ser simulados, como são os casos dos sistemas hiperespectrais. Isso traz facilidades, pois nem sempre temos, de imediato, as imagens de nossa área de estudo e podemos, assim, antecipar o trabalho realizando coletas de amostras em campo. Ao obter os espectros das amostras coletadas, eles podem ser reamostrados para o sistema sensor que irá futuramente imagear a área, possibilitando, dessa maneira, prever como o sensor irá “ver” os alvos que foram amostrados. Um exemplo disso é apresentado na Figura 15.6, em que se simulou o sensor aerotransportado australiano Hymap para os cinco espectros originais apresentados na Figura 15.5. Esses espectros

apresentam 419 bandas no seu formato original, porém, o Hymap possui 125 bandas, o que gera espectros com menor precisão quanto aos pontos de inflexão das feições de absorção. Entretanto, vemos que elas estão presentes, como as da caulinita (1), da gibbsita (2), e as ombreiras decorrentes das transições eletrônicas dos sesquióxidos (3 e 4). Além disso, por apresentar espectros contínuos, as feições decorrentes da presença de água centradas em 1,4 μm e 1,9 μm são detectadas.

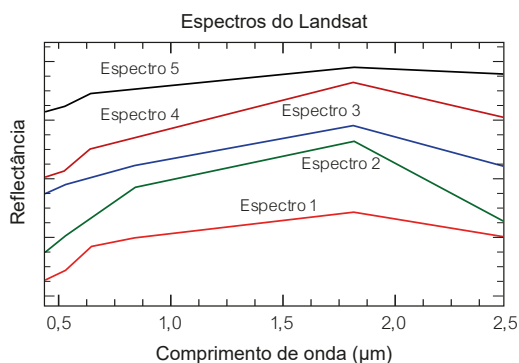


Figura 15.5. Espectros dos solos da Figura 15.4 reamostrados para as bandas do sensor ETM+ do Landsat 7.

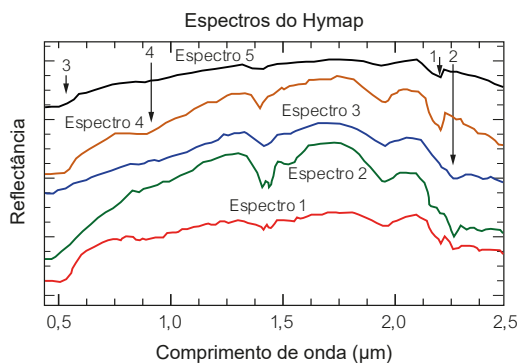


Figura 15.6. Espectros de solos da Figura 15.5 reamostrados para o sensor aerotransportado Hymap que opera com 125 bandas.

15.4. Identificação de *Endmembers*

Como mencionado no início do item anterior, os dados hiperespectrais permitem a identificação de espectros de referência, os *endmembers*. Para tal, é necessário reduzir a dimensão dos dados, ou seja, do número de bandas, pois com centenas de bandas sempre encontraremos alta correção espectral entre muitas das bandas do sensor. O algoritmo de processamento que realiza essa operação é chamado de *Minimum Noise Fraction* (MNF). O MNF é um procedimento baseado na Análise de Componentes Principais (ACP) que, além de reduzir a dimensionalidade dos dados, permite a remoção do ruído e a redução da demanda computacional, por meio de duas transformações. A primeira baseia-se na matriz de covariância do ruído e, por meio de descorrelação, é reescalonado. A segunda transformação ocorre após a remoção do ruído e baseia-se nas componentes principais dos dados. A dimensionalidade dos dados é avaliada a partir do gráfico de autovalores (*eigenvalues*) (Figura 15.7). O exemplo utilizado é baseado em uma cena AVIRIS com 50 bandas a partir de 2,2 μm , da região de Cuprite, estado de Nevada, Estados Unidos.

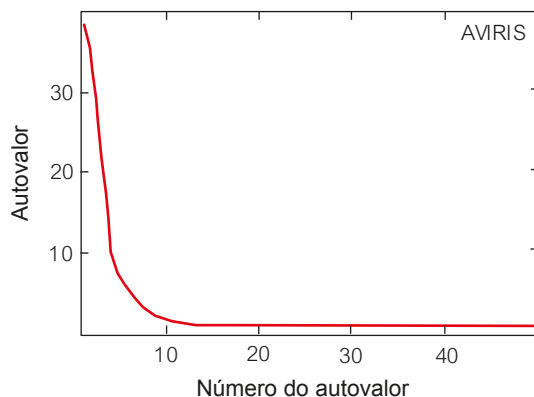


Figura 15.7. Gráfico de autovalores obtidos a partir da transformação MNF de 50 bandas do sensor AVIRIS.

No processamento por PCA, o número de bandas de entrada é igual ao número de MNF geradas. Como, nesse caso, foram 50 bandas, temos 50 MNFs. Note, na Figura 15.8, que, depois da MNF 10, os autovalores baixam significativamente. A partir de então, considera-se haver basicamente imagens ruído. Podemos trabalhar com as 10 imagens MNFs reduzindo em 1/5 a dimensão desses dados, o que agiliza o processamento. A Figura 15.8 exibe uma composição colorida com as três primeiras MNFs da cena de Cuprite, apresentando a diversidade mineralógica que esta área apresenta geologicamente.



Figura 15.8. Composição colorida com as três primeiras imagens MNFs geradas de 50 bandas do AVIRIS da região de Cuprite, Nevada, EUA.

De posse dessas 10 MNFs, pode-se realizar o procedimento inverso, ou seja, de MNF para as bandas em separado, removendo-se, assim, o ruído dos dados. Porém, para a determinação dos *endmembers*, não há a necessidade de realizar esse procedimento inverso. Basta apenas utilizar as 10 imagens MNFs para encontrar os pixels puros existentes na cena, o que é

feito pelo algoritmo de processamento de dados hiperespectrais denominada *Pixel Purity Index* (PPI).

O processamento feito pelo algoritmo PPI baseia-se em um processo iterativo de geometria convexa e ocorre banda a banda. Cada vez que um pixel aparece no extremo do diagrama de dispersão n -dimensional, ele é marcado. A imagem resultante apresenta os pixels brancos que foram marcados e o valor corresponde em quantas vezes ele foi considerado como extremo. O restante da cena apresenta-se com cor preta que representa os pixels não puros ou pixels misturados, como é mostrado na Figura 15.9.

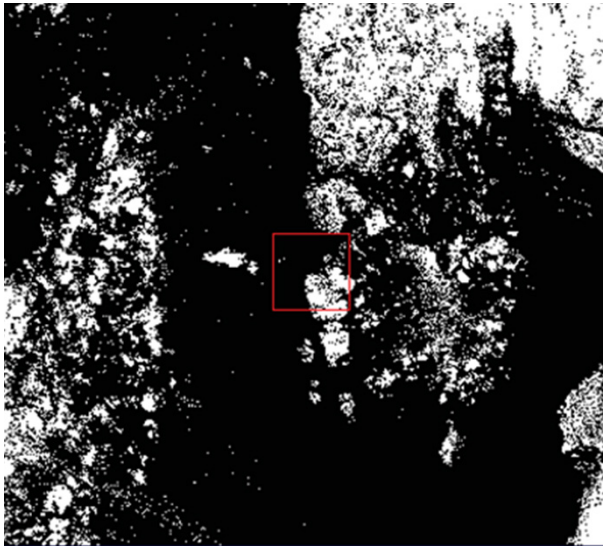


Figura 15.9. Imagem PPI mostrando, em branco, os pixels puros, após 10.000 iterações.

Ao selecionarmos todos os pixels diferentes de zero, estaremos selecionando os chamados pixels puros identificados na imagem de Cuprite e se os exportarmos para o visualizador n -dimensional, poderemos identificar os *clusters* ou as nuvens de pixels que representam os agrupamentos de pixels espectralmente similares (Figura 15.10). Nessa figura, o escaterograma

refere-se às imagens de reflectância do sensor AVIRIS de Cuprite, ou seja, as mesmas que foram utilizadas para determinar os MNFs.

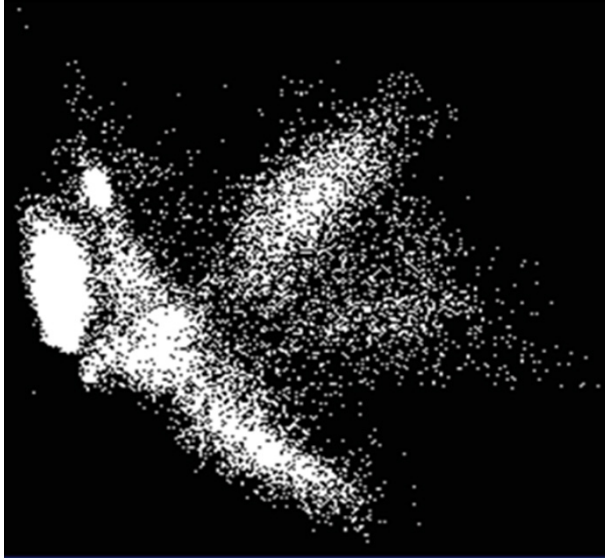


Figura 15.10. Diagrama de dispersão n-dimensional das bandas de reflectância apresentando apenas o comportamento dos pixels puros.

Como esse espaço apresenta mais de duas bandas, ele pode ser rotacionado, facilitando a identificação e demarcação dos agrupamentos (*clusters*). A Figura 15.11 mostra a seleção de vários *clusters*, um para cada cor, e a posterior exportação dos mesmos sobre uma composição colorida de bandas originais de reflectância, a fim de visualizar as diferenças minerais existentes na área mostradas pelos pixels puros.

Se obtivermos os espectros dos pixels que foram exportados, ou seja, dos pixels ditos puros, estaremos extraindo os *endmembers* a partir da imagem, conforme pode ser observado no espectro da Figura 15.12, na qual foi selecionado o mineral caulinita, cuja localização coincide com os mapeamentos de campo mostrado na imagem composição colorida pelo cursor.

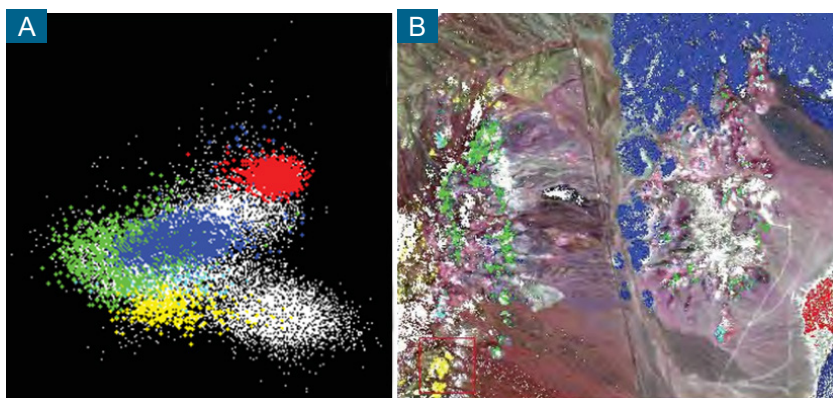


Figura 15.11. Em (A), diagrama de dispersão n -dimensional com os clusters selecionados e, em (B), os *clusters* exportados para as imagens de reflectância, mostrando os diversos minerais existentes em Cuprite.

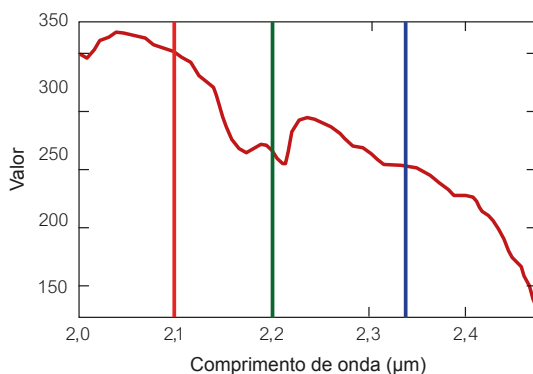


Figura 15.12. Imagem da área de Cuprite com a presença do mineral caulinita e o espectro do *endmember* de caulinita selecionado a partir da análise de PPI.

15.5. Medição de Intensidade das Feições Espectrais

A maioria dos alvos imageados guarda uma relação linear entre a intensidade da absorção e o conteúdo da matéria que lhe deu origem. Por isso, é possível estimar os conteúdos dos materiais por meio das medições das

intensidades das absorções. Um dos procedimentos mais simples para fazer isso é por meio de uma operação denominada de remoção do contínuo espectral. Essa remoção consiste em uma normalização dos espectros de reflectância a partir de uma linha de base comum que deve tangenciar ao máximo a curva espectral. A Figura 15.13 apresenta o conceito de remoção do contínuo espectral.

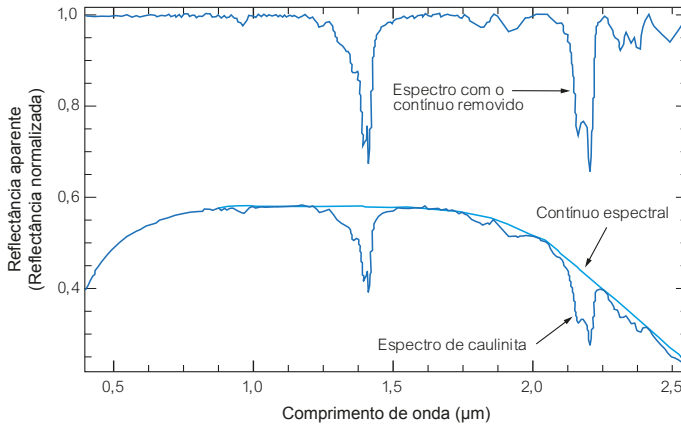


Figura 15.13. Conceito de remoção do contínuo espectral usando, como exemplo, a espectro da caulinita.

Para determinar a intensidade da feição de absorção, pode-se adotar o procedimento descrito pela equação abaixo (Equação 15.1). O resultado desse procedimento é ilustrado na Figura 15.14, que mostra uma feição de absorção da caulinita.

$$D = 1 - R' \quad (15.1)$$

em que

D = profundidade da banda de absorção; e

R' = valor de reflectância correspondente ao ponto de absorção no espectro com o contínuo removido.

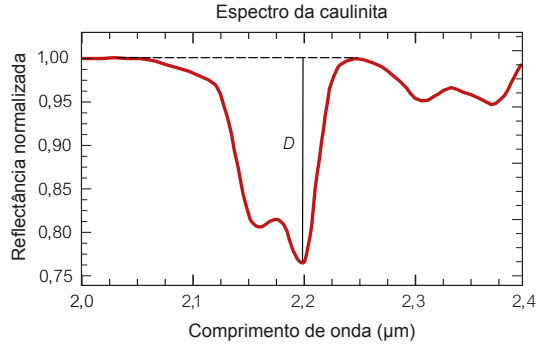


Figura 15.14. Conceito de profundidade de absorção (*band depth*) a partir de espectro com remoção do contínuo.

No caso específico desse espectro, a intensidade da absorção (D) seria de 0,23422, lembrando que esse dado é adimensional, pois os espectros foram normalizados de 0 a 1. Pode-se supor que a feição de absorção apresenta 23,4% de intensidade se comparada com uma feição que apresentasse intensidade variando de 0 a 1. Porém, quando a banda de absorção é dupla, como é o caso da caulinita, o procedimento considera a feição de maior intensidade. No entanto, ao se utilizar álgebra de bandas, é possível contemplar as duas feições, por meio de sua média. No caso da caulinita, o procedimento para estimar a intensidade de sua feição pode ser obtido da seguinte forma (Equação 15.2):

$$D = \left[\left(\frac{b_1 + b_2}{2} \right) - \left(\frac{b_3 + b_4}{2} \right) \right] \quad (15.2)$$

em que

D = profundidade média da feição de absorção da caulinita;

b_1 = banda referente ao início da feição localizada a 2,05 μm ;

b_2 = banda referente ao final da feição localizada a 2,25 μm ;

b_3 = banda referente à primeira absorção centrada a 2,16 μm ; e

b_4 = banda referente à segunda absorção centrada a 2,20 μm .

A Figura 15.15 ilustra a lógica de determinação da profundidade média de feições que apresentam mais de um ponto de inflexão, como a da caulinita.

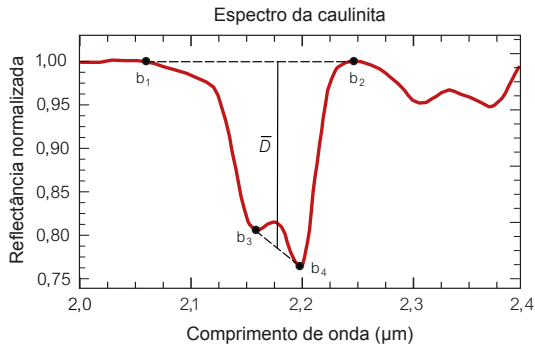


Figura 15.15. Conceito de profundidade média de feição espectral dupla de acordo com a Equação 15.3.

Existe uma maneira mais elaborada para determinar a intensidade da feição espectral usando um algoritmo chamado *Spectral Feature Fitting* (SFF) (Clark et al., 1990). Esse algoritmo realiza uma comparação dos *endmembers* com cada pixel da cena, por meio do ajuste por mínimos quadrados. Para tal, a imagem tem de ser submetida ao procedimento de remoção do contínuo espectral. O SFF gera dois resultados, sendo o primeiro chamado de imagem *scale* que deve ser entendida como a medida da intensidade da feição em cada pixel. Quanto maior o valor, maior é a intensidade da feição do *endmember* no pixel. O segundo produto do SFF é o erro médio quadrático do ajuste por mínimos quadrados, ou imagem RMS. A divisão do *scale* pelo RMS gera um terceiro produto chamado de *fit image*, que pode ser entendido como a probabilidade de o pixel conter o *endmember*.

Para exemplificar esse processamento, utilizaremos a imagem de Cuprite, Nevada. A Figura 15.16 apresenta a imagem *scale* de caulinita e seu respectivo RMS. Repare na área destacada pelo cursor vermelho, a mesma área que foi selecionada quando explicamos como selecionar *endmembers* a partir das imagens. Como é uma área que apresenta uma quantidade alta de caulinita, na imagem *scale* (a da esquerda) ela aparece com um tom de cinza mais claro, o que indica elevado valor de intensidade de absorção por este mineral. Já na imagem RMS ele não se destaca, pois, o RMS do ajuste de seus pixels é baixo.

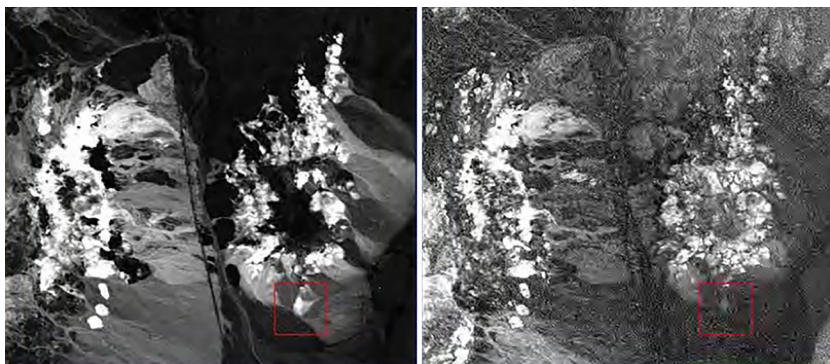


Figura 15.16. Resultados do algoritmo SFF que identifica a intensidade da absorção do mineral caulinita. À esquerda, imagem *scale* e, à direita, imagem RMS.

Outra possibilidade de investigação de intensidade de feições se dá pelos processos de decomposição espectral linear. O algoritmo mais usual para tal fim é o *Linear Spectral Unmixing* que determina a abundância relativa dos materiais (ou *endmembers*) na composição do pixel mistura. Existe uma restrição quanto ao seu uso, que não é problema para sensores hiperspectrais, pois não se pode adotar um número de *endmembers* maior que o número de bandas do sensor. Como resultado do *Linear Spectral Unmixing*, tem-se, para cada *endmember*, uma imagem de abundância e um referente ao RMS da decomposição linear.

Para testar o algoritmo, foram escolhidos os minerais illita e calcita para a área de Cuprite. A área investigada com o algoritmo SFF, apontada na Figura 15.17 pela seta, aparece em um tom mais azulado, indicando que, na decomposição linear espectral, os pixels avaliados não apresentam nem as feições da illita e nem da calcita. O resultado obtido mostra a imagem de abundância de illita em vermelho, calcita em verde e o RMS em azul, sobre uma composição colorida. Portanto, as áreas mais vermelhas apresentam maior abundância de illita e as em verde de calcita. Isso pode ser verificado nas imagens isoladas. A Figura 15.18 apresenta a imagem *Unmixing* de illita e de calcita, respectivamente. Nos locais mais claros é onde há uma maior abundância de feições espectrais correspondentes a esses dois minerais.

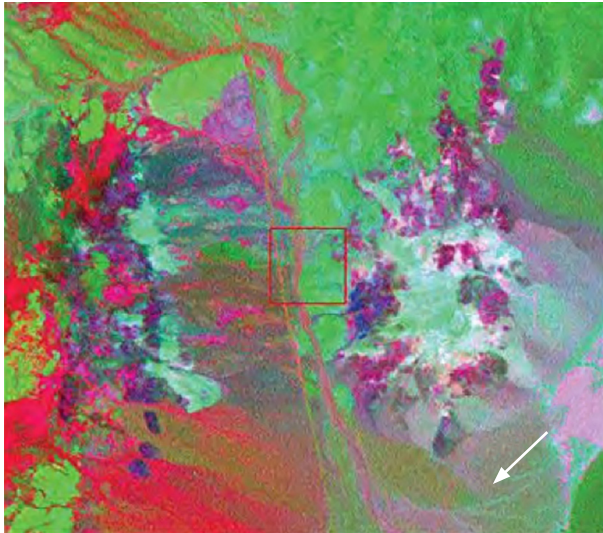


Figura 15.17. Composição colorida com a imagem linear *spectrum unmixing* para os *endmembers* illita em vermelho, calcita em verde e RMS em azul.

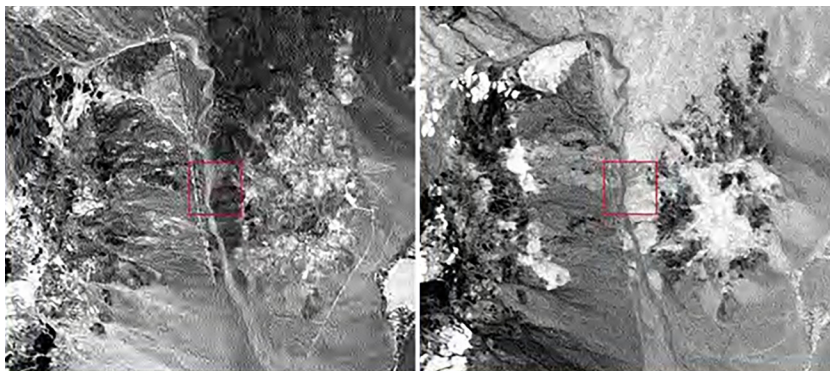


Figura 15.18. Imagens *unmixing* para os endmembers illita à esquerda e calcita à direita. Os pontos mais claros indicam maior abundância desses minerais.

Os procedimentos que acabamos de ver demonstram que a análise dos espectros é o aspecto mais importante a ser trabalhado no processamento de dados hiperespectrais. Normalmente, vamos buscar, nos espectros dos pixels da cena, as feições que os *endmembers* apresentam a fim de identificá-los espacialmente e, posteriormente, quantificá-los. Existem outros algoritmos que tem esse mesmo objetivo de identificar as feições espectrais e a tendência é cada vez mais surgirem novos processamentos, pois já há pesquisas suficientes em várias áreas de aplicação que comprovam o alto poder de mapeamento dos pixels com imagens de sensores hiperespectrais.

15.6. Referências

BAPTISTA, GM.M. **Sensoriamento remoto hiperespectral**. Rio de Janeiro: Interciência, 2^a ed., 2019, 204 p.

CLARK, R. N.; GALLAGHER, A. J.; SWAYZE, G. A. Material absorption band depth mapping of imaging spectrometer data using a complete band shape leastsquares fit with library reference spectra. In: ANNUAL JPL AIRBORNE GEOSCIENCE WORKSHOP, 2, Pasadena, CA. **Summaries**. JPL-Publ. 90-54, 1990, p. 176-186.

