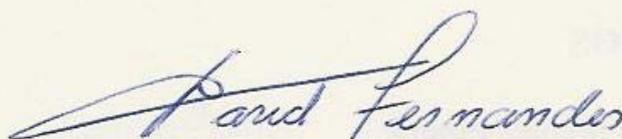


Tese apresentada à Pró-Reitoria de Pós-Graduação e Pesquisa do Instituto Tecnológico de Aeronáutica, como parte dos requisitos para obtenção do título de Doutor em Ciências no Curso de Pós-Graduação em Engenharia Eletrônica e Computação na Área de Telecomunicações.

MARCO ANTONIO PIZARRO

**Triagem de Amostras e Múltiplas Medidas de
Similaridade Espectral para Seleção de Membros de Referência
em Imagens Hiperespectrais**

Tese aprovada em sua versão final pelos abaixo assinados:



Prof. Dr. David Fernandes

Orientador

Prof. Dr. Homero Santiago Maciel
Pró-Reitor de Pós-Graduação e Pesquisa

Campo Montenegro
São José dos Campos, SP – Brasil

2007

Dados Internacionais de Catalogação-na-Publicação (CIP)
Divisão Biblioteca Central do ITA / CTA

Pizarro, Marco Antonio

Triagem de Amostras e Múltiplas Medidas de Similaridade Espectral para Seleção de Membros de Referência em Imagens Hiperespectrais / Marco Antonio Pizarro.

São José dos Campos, 2007.

185f.

Tese de Doutorado – Curso de Pós-Graduação em Engenharia Eletrônica e Computação. Área de Telecomunicações – Instituto Tecnológico de Aeronáutica, 2007. Orientador: D. C. David Fernandes

1. Dados Hiperespectrais. 2. Entropia. 3. Membros de Referência. 4. Sensoriamento Remoto. 5. Espectroscopia de Imageamento. 6. Espectrômetro Imageador Aerotransportado no Visível e Infravermelho (AVIRIS). 7. Mistura Espectral. 8. Processamento de Imagens. 9. Classificação de Imagens.

Comando-Geral de Tecnologia Aeroespacial. Instituto Tecnológico de Aeronáutica. II.Título

REFERÊNCIA BIBLIOGRÁFICA

PIZARRO, Marco Antonio. Triagem de Amostras e Múltiplas Medidas de Similaridade Espectral para Seleção de Membros de Referência em Imagens Hiperespectrais. 2007. 185f. Tese de Mestrado – Instituto Tecnológico de Aeronáutica, São José dos Campos.

CESSÃO DE DIREITOS

NOME DO AUTOR: Marco Antonio Pizarro.

TÍTULO DO TRABALHO: Triagem de Amostras e Múltiplas Medidas de Similaridade Espectral para Seleção de Membros de Referência em Imagens Hiperespectrais.

TIPO DO TRABALHO/ANO: Tese de Doutorado/ 2007.

É concedida ao Instituto Tecnológico de Aeronáutica permissão para reproduzir cópias desta tese e para emprestar ou vender cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte desta tese pode ser reproduzida sem a sua autorização (do autor).

Marco Antonio Pizarro

Rua Curaçao, 354 – Cidade Vista Verde

12223-750 – São José dos Campos – SP

**Triagem de Amostras e Múltiplas Medidas de Similaridade
Espectral para Seleção de Membros de Referência em Imagens
Hiperespectrais**

MARCO ANTONIO PIZARRO

Composição da Banca Examinadora:

Prof. Waldecir João Perrella, D.C.	Presidente	(ITA)
Prof. David Fernandes, D. C.	Orientador	(ITA)
Prof. Marcelo da Silva Pinho, Dr. Sc.		(ITA)
Prof. Lênio Soares Galvão, D. C.	Membro Externo	(INPE)
Vitor Francisco de Araújo Härtel, PhD.	Membro Externo	(UFRGS)

Suplentes:

Prof. Fernando Toshinori Sakane, PhD.		(ITA)
Prof. Álvaro Penteadó Crósta, PhD.	Membro Externo	(UNICAMP)

ITA

A Lúgia, esposa e amiga, pelo inestimável apoio, incentivo e muita paciência; aos meus filhos Raíssa e Pedro Arthur pelos felizes dias, imprescindíveis, ao longo destes anos dedico este trabalho.

AGRADECIMENTOS

Ao Prof. Dr. David Fernandes pela orientação, incentivo, bem como pelas valiosas discussões durante a elaboração do trabalho.

Ao Instituto Nacional de Pesquisas Espaciais (INPE) pela oportunidade de realizar o doutorado no ITA.

Ao Eng. MsC. Amauri Silva Montes coordenador geral da Engenharia e Tecnologia Espacial (ETE/INPE) e ao Eng. Dr. Mário Luiz Selingardi, chefe da Divisão de Eletrônica Aeroespacial (DEA/INPE), pelo incentivo e apoio na realização deste trabalho.

À Divisão de Sensoriamento Remoto (DSR) do INPE, pelas imagens do sensor AVIRIS/JPL utilizadas neste estudo.

À bibliotecária Silvia Marcelino, funcionária do Serviço de Informação e Documentação (SID) do INPE, pela contínua ajuda ao longo da elaboração deste trabalho.

Ao Projeto CAPTAER-FINEP do ITA e ao Termo Aditivo 1 e 3 da Cooperação CTA – CCSIVAM (Comissão para Coordenação do Projeto do Sistema de Vigilância da Amazônia) que envolvia como parceira a ATECH (Fundação Aplicações de Tecnologias Críticas) pelo suporte em equipamentos e softwares.

A todas as pessoas que contribuíram de forma direta ou indireta para a realização deste trabalho.

RESUMO

O objetivo do presente trabalho foi desenvolver um processo semi-automático para identificação de Membros de Referência (MR) em imagens hiperespectrais. O processo faz inicialmente a triagem das amostras espectrais gerando um conjunto de amostras candidatas a MR. A triagem consiste em: determinar uma vizinhança espacialmente uniforme em torno da amostra espectral, rejeitar amostras cujos espectros na vizinhança uniforme não são homogêneos, rejeitar amostras que não formam um contexto uniforme e homogêneo com outras amostras e rejeitar amostras redundantes.

As amostras espectrais candidatas a MR, consideradas uniformes, homogêneas, pertencentes a um contexto e pouco redundantes, são submetidas a um processo de busca dos MR, que utiliza o conceito de entropia, proposto neste trabalho, como medida de similaridade espectral, em conjunto com a Distância Euclidiana e a medida de Coerência entre os espectros. Propõe-se ainda dois limitantes superiores para o número máximo de MR presentes em uma imagem hiperespectral, utilizando-se o conceito de amostras espectrais bem configuradas e um valor de entropia mínima permissível, entre amostras.

Foi implementado com o aplicativo computacional IDL® um processador que realiza todas as funções propostas. Os métodos apresentados são testados, com êxito, em imagens hiperespectrais do sensor aerotransportado AVIRIS (*Airborne Visible Infrared Imaging Spectrometer*) do JPL.

ABSTRACT

This work proposes a new semi-automatic method to find the EndMembers (EM) in hiperespectral images. Initially the methodology does the discrimination of spectral samples in order to create a set of EM candidates samples. The discrimination is conducted by the selection of a uniform spatial neighborhood; surround the samples, with homogeneous spectra. The discrimination processes reject also samples that don't make a uniform and homogeneous cluster with others uniform and homogeneous samples and finely reject redundant samples.

Uniforms and homogeneous samples that belongs to a spatial context and are non redundant are used as input in the developed processor for EM selection. This processor searches for the EM using the entropy concept, proposed in this work, which can be also used jointly with Euclidian Distance and Coherence measurement between spectra. Two upper bounds for the EM numbers, based in the samples well-configured set concept and in the minimum admissible entropy value, for a set of samples is also proposed.

The developed processor was implemented in IDL[®] software. The proposed method is successfully applied in JPL-AVIRIS (Airborne Visible/InfraRed Imaging Spectrometer) images.

SUMÁRIO

Pág.

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE SÍMBOLOS

LISTA DE SIGLAS E ABREVIATURAS

1 Introdução	21
1.1 Membros de Referência em Imagens Hiperespectrais	21
1.2 Objetivos do Trabalho	22
1.3 Organização do trabalho	23
2 Espectroscopia de Imageamento	25
2.1 Conceito do sensoriamento remoto hiperespectral	25
2.2 Resposta espectral da cobertura vegetal, do solo e da água	26
2.2.1 Reflectância da cobertura vegetal com e sem atividade fotossintética.....	26
2.2.2 Reflectância do Solo	29
2.2.3 Reflectância de alvos contendo sombra e água	34
2.3 A questão da informação espacial e espectral	34
2.4 Conceito de <i>pixel</i> misturado	37
2.5 Membros de referência (MR) e modelos de mistura espectral (MME).....	39
3 Sensor AVIRIS e Área Teste	48
3.1 Sensor AVIRIS	48
3.2 Área teste	49
4 Amostras Candidatas a MR	59
4.1 Triagem das amostras associadas às classes de alvos.....	59
4.1.1 Triagem manual com especialista.....	60
4.1.2 Triagem aleatória.....	61
4.1.3 Triagem discriminatória	62
4.2 Uniformidade espacial da vizinhança de uma amostra teste	63
4.3 Homogeneidade espectral da vizinhança de uma amostra teste	67
4.4 Contexto espacial da amostra teste	71
4.5 Redundância entre amostras espectrais	72
4.6 Limiares e fatores no processo de triagem discriminatória	77
5 Condicionamento Espectral.....	79
5.1.1 Condicionamento pela derivada discreta.....	79
5.1.2 Condicionamento pela decomposição com a TWD	80
6 Busca dos MR pela Máxima Entropia.....	85
6.1 Medida de similaridade espectral pela entropia	85
6.2 Amostras espectrais bem configuradas.....	88
6.3 Determinação dos MR pela máxima entropia	90
6.4 Limitante superior do número de MR	91
6.5 Determinação dos limiares de configuração.....	92

6.5.1 Limiar de configuração associado à <i>DE</i>	93
6.5.2 Limiar de configuração para <i>CE</i>	94
6.5.3 Limiar de configuração para entropia.....	95
7 Busca dos MR com Múltiplas Medidas de Similaridade Espectral	96
7.1 Distância Euclidiana como medida de similaridade espectral.....	96
7.2 Coerência espectral como medida de similaridade espectral	98
7.3 Utilização de múltiplas medidas.....	99
7.4 Estrutura geral do processo de busca dos MR.....	103
8 Teste de Desempenho e Análise de Resultados	104
8.1 Teste da triagem discriminatória	104
8.1.1 Amostra d'água contaminada	104
8.1.2 Amostras selecionadas manualmente e aleatoriamente.....	107
8.1.3 Aplicação da triagem discriminatória na subcena4	111
8.2 Busca dos MR pela entropia.....	115
8.2.1 Teste do método de busca dos MR pela máxima entropia na subcena 4.....	115
8.2.2 Aplicação do método de busca pela máxima entropia na subcena 5.....	121
8.3 Condicionamento Espectral.....	124
8.4 Triagem discriminatória e busca dos MR para a subcena 7	135
8.5 Comparação com o algoritmo PPI.....	141
9 Considerações Finais	146
9.1 Sumário do estudo realizado	146
9.2 Contribuições para a área de espectroscopia de imageamento.....	146
9.3 Avaliação do trabalho realizado	148
9.4 Sugestões para continuidade do trabalho realizado.....	149
APÊNDICE A – Estrutura e Arquivos do Processador para Triagem e Busca de MR	164
A.1 - Estrutura básica do processador para busca dos MR	164
A.2 - Arquivo de parâmetros de entrada.....	168
A.3 - Arquivo de dados das amostras escolhidas pela triagem manual.....	170
A.4 - Relatório de saída do programa.....	172
A.5 - Figuras e imagens de saída do programa.....	185

LISTA DE FIGURAS

FIGURA 2.1 – Representação esquemática do sensoriamento hiperespectral.....	26
FIGURA 2.2 - Espectros de reflectância para diferentes conteúdos de umidade em folhas....	28
FIGURA 2.3 – Reflectância espectral de materiais secos obtidos do pinheiro.....	28
FIGURA 2.4 – Espectros de reflectância da caulinita, nontronita e illita.....	33
FIGURA 2.5 – Vizinhança 3 x 3 da amostra espectral (i, j).....	35
FIGURA 2.6 – Conceito de <i>pixel</i> puro e <i>pixel</i> misturado em imagens hiperespectrais.....	38
FIGURA 2.7 - Mistura microscópica entre materiais numa região inferior às dimensões da célula de resolução do sensor.....	39
FIGURA 2.8 – Mistura espectral a nível macroscópico: (a) superposição de vários materiais em um <i>pixel</i> e (b) MLME representativo da combinação de espectros puros da vegetação seca, vegetação verde e solo.....	40
FIGURA 2.9 – Tipos de mistura: (a) modelo linear associado a reflexões simples e (b) modelo não linear associado a múltiplas reflexões. Fonte: adaptação a partir de Keshava e Mustard (2002, p. 45).....	42
FIGURA 2.10 – Diagrama de espalhamento entre duas dimensões dos dados espectrais ilustrando a interpretação do modelo de mistura baseada nos MR.....	44
FIGURA 2.11 – Diagrama de espalhamento bidimensional contendo MR falsos para eliminar o aparecimento dos pontos fora da curva (<i>outliers</i>).....	45
FIGURA 3.1 – Relação sinal-ruído do sensor AVIRIS para os anos de 1987, 1994 e 1997....	49
FIGURA 3.2 - Espectro da vegetação verde extraído de um <i>pixel</i> da imagem original, corrigida para os efeitos de absorção e espalhamento atmosféricos, com 224 bandas originais sem interpolação e 209 bandas com interpolação.....	51
FIGURA 3.3 – Composição colorida RGB para a subcena 4 obtida com as bandas 22 (≈ 667 nm), 10 (≈ 549 nm) e 3 (≈ 481 nm), respectivamente, do sensor AVIRIS.....	52
FIGURA 3.4 - Espectros médios representativos das classes e subclasses pertencentes a subcena 4.....	53
FIGURA 3.5 – Composição colorida RGB para a subcena 5 obtida com as bandas 22 (≈ 667 nm), 10 (≈ 549 nm) e 3 (≈ 481 nm), respectivamente, do sensor AVIRIS.....	54
FIGURA 3.6 - Espectros médios das amostras das classes e subclasses pertencentes a subcena 5.....	55
FIGURA 3.7 – Composição colorida RGB para a subcena 7 obtida com as bandas 22 (≈ 667 nm), 10 (≈ 549 nm) e 3 (≈ 481 nm), respectivamente, do sensor AVIRIS.....	56
FIGURA 3.8 – Espectros médios das amostras pertencentes às classes e subclasses indicadas na subcena 7.....	57
FIGURA 4.1 – Triagem de amostras na subcena 4 pela abordagem aleatória e manual.....	62
FIGURA 4.2 – Tipos de triagem.....	63
FIGURA 5.1 - Estrutura de aplicação da TWD no espectro $\vec{I}(p, q)$ com $(p, q) \in V_k$	81
FIGURA 5.2: Condicionamento dos espectros das amostras $vv1$ e vs	82
FIGURA 5.3: Condicionamento dos espectros das amostras $RQ1$ e $LV1$	83
FIGURA 5.4: Condicionamento do espectro da amostra a (água).....	83
FIGURA 7.1 – Seleção dos MR.....	103
FIGURA 8.1 – Teste da uniformidade espacial. (a): corpo da água pertencente a subcena 7 onde estão indicados os limites da vizinhança 5×5 <i>pixels</i> em torno de uma amostra espectral relativa a água. (b) espectros contidos na vizinhança da amostra espectral água.....	

(c) 19 espectros escolhidos como pertencentes a uma vizinhança uniforme. (d) espectros rejeitados por serem considerados contaminados.....	105
FIGURA 8.2 – Teste da uniformidade espacial relativo à vizinhança de três amostras de água contaminadas por vegetação verde.....	107
FIGURA 8.3 – Triagem discriminatória de amostras na subcena 4 escolhidas inicialmente pela abordagem aleatória e manual.....	110
FIGURA 8.4 – Triagem da subcena 4 quanto a uniformidade, homogeneidade e contexto.....	113
FIGURA 8.5 – Triagem da subcena 4 pelo critério da redundância aplicada às amostras selecionadas pelo contexto espacial.....	114
FIGURA 8.6 – Aglomerados resultante do contexto espacial.....	116
FIGURA 8.7 - Imagem da subcena 4, banda 192 (2.271 nm), com as amostras espectrais utilizadas nos testes de busca de MR.....	117
FIGURA 8.8 - Espectros de amostras candidatas a MR da subcena 4.....	117
FIGURA 8.9 - Imagens pós-classificadas da subcena 4 com $R = 5$ e 6 . Em cores as classes com <i>pixels</i> tendo valores de abundância superiores a 50% da fração total e em branco as áreas não-classificadas (intensa mistura espectral).....	120
FIGURA 8.10 - Amostras candidatas a MR na subcena 5.....	121
FIGURA 8.11 - Espectros dos MR escolhidos para a subcena 5 em correspondência com (<i>vv</i> , <i>vs</i> , <i>a</i> , <i>LV</i> , <i>RQ</i>).....	123
FIGURA 8.13 – Imagens pós-classificadas resultante da aplicação do MLME para $R = 4$, onde foram realçados os <i>pixels</i> com valores de abundância superiores 50 % da fração total. As áreas não-classificadas e, portanto, de intensa mistura espectral das componentes da cena são mostradas em branco.....	130
FIGURA 8.14 – Imagens pós-classificadas resultante da aplicação do MLME para $R = 5$, onde foram realçados os <i>pixels</i> com valores de abundância superiores 50 % da fração total. As áreas não-classificadas e, portanto, de intensa mistura espectral das componentes da cena são mostradas em branco.....	132
FIGURA 8.15 – Imagens pós-classificadas relativas às medidas <i>DE</i> e <i>CE</i> resultante da aplicação do MLME para $R = 6$, onde foram realçados os <i>pixels</i> com valores de abundância superiores 50 % da fração total. As áreas não-classificadas e, portanto, de intensa mistura espectral das componentes da cena são mostradas em branco.....	134
FIGURA 8.17 – Aplicação da triagem discriminatória na subcena 7.....	136
FIGURA 8.18 – Posição das 83 amostras candidatas a MR e das sete amostras selecionadas como MR para subcena7 (banda 192 - (2.271 nm).....	137
FIGURA 8.19 – Espectros médios dos MR da subcena 7 selecionados para $2 \leq R \leq 6$	139
FIGURA 8.20 – Imagens pós-classificadas resultante da aplicação do MLME para $R = 5$ e 6 , onde foram realçados os <i>pixels</i> com valores de abundância superiores 50 % da fração total. As áreas não-classificadas e, portanto, de intensa mistura espectral das componentes da cena são mostradas em branco.....	140
FIGURA 8.21 – Composição colorida R (PPI), G (549 nm), B (PPI) para a subcena 4, identificando os <i>pixels</i> puros, em magenta, pela técnica PPI.....	142
FIGURA 8.22 – Espectros selecionados automaticamente pelo PPI para a subcena 4 com $N_{MNF} = 20$, $L = 10.000$ e $t = 2,5$	144
FIGURA A.1 - Amostras candidatas e MR escolhidos para a subcena 7.....	185

LISTA DE TABELAS

2.1- Bandas de absorção de alguns minerais do grupo dos óxidos de ferro. Fonte: adaptação a partir de Hunt; Salisbury e Lenhoff (1973, p.99); Clark, King e Klejwa (1990, p. 12.653-12.680); Epiphonio et al. (1992, p. 35-41).	31
2.2 Bandas de absorção de alguns minerais do grupo das argilas. Fonte: adaptação a partir de Hunt; Salisbury e Lenhoff (1973, p.99); Clark, King e Klejwa (1990, p. 12.653-12.680); Epiphonio et al. (1992, p. 35-41).	32
4.1 - Triagem manual realizada por especialista.	61
8.1 - MR selecionados pela entropia na subcena 4.....	119
8.2 - MR selecionados na subcena 5.....	122
8.3 - Triagem discriminatória de amostras selecionadas na subcena 4.	125
8.4 -MR selecionados na subcena 4.....	127
8.5 - MR selecionados com para a subcena 7.....	138

LISTA DE SÍMBOLOS

Latinos	
$CE(.,.)$	- coerência espectral entre dois espectros
CE_k	- Coerência espectral entre o espectro de referência \vec{I}_{ref} e o espectro médio \vec{I}_{mk} , $CE_k = CE(\vec{I}_{mk}, \vec{I}_{ref})$.
$CE(e_1, e_2, \dots, e_R)$	- Coerência espectral relativa ao conjunto de amostras (e_1, e_2, \dots, e_R) , $e_1, e_2, \dots, e_R \in \{1, 2, \dots, \mathbb{K}\}$.
$CE_{[1]}, CE_{[2]}, \dots, CE_{[K_C]}$	- conjunto de medidas de coerências ordenadas em ordem crescente
$DE(.,.)$	- distância euclidiana entre dois espectros.
DE_k	- distância euclidiana entre o espectro de referência \vec{I}_{ref} e o espectro médio \vec{I}_{mk} , $DE_k = DE(\vec{I}_{mk}, \vec{I}_{ref})$.
$DE(e_1, e_2, \dots, e_R)$	- distância euclidiana relativa ao conjunto de amostras (e_1, e_2, \dots, e_R) , $e_1, e_2, \dots, e_R \in \{1, 2, \dots, \mathbb{K}\}$.
$DE_{(1)}, DE_{(2)}, \dots, DE_{(K_C)}$	- conjunto de medidas de distâncias ordenadas em ordem crescente.
\vec{e}_r	- vetor espectro referente ao r-ésimo MR, onde $1 \leq r \leq R$.
E	- matriz de dimensão $(R \times N)$ contendo o conjunto de espectros \vec{e}_r .
$H(e_1, e_2, \dots, e_R)$	- entropia relativa ao conjunto de amostras (e_1, e_2, \dots, e_R) , $e_1, e_2, \dots, e_R \in \{1, 2, \dots, \mathbb{K}\}$.
H_{min}	- mínima entropia aceitável.
$I(i, j, n)$	- n-ésima componente do vetor $\vec{I}(i, j)$.
$\vec{I}(i, j)$	- $\vec{I}(i, j) = [I(i, j, 1) I(i, j, 2) I(i, j, 3) \dots I(i, j, N)]^T$
$\vec{I}(i_k, j_k)$	- vetor espectro da amostra espectral k referente ao pixel (i_k, j_k) . Também pode ser representado como: $\vec{I}(i_k, j_k) = \vec{I}_k$.
$\vec{I}'(p, q)$	- espectro com adição de zeros para ter 256 elementos.
$\vec{I}_d(p, q)$	- derivada do espectro $\vec{I}(p, q)$, $\vec{I}_d(p, q) = \dot{\vec{I}}(p, q)$.
$\vec{I}_w(i_k, j_k)$	- vetor espectro processado pela TWD.

$\dot{\bar{I}}(p, q)$	- derivada do espectro $\bar{I}(p, q)$, $\bar{I}_d(p, q) = \dot{\bar{I}}(p, q)$.
$\tilde{\bar{I}}_m(i_k, j_k)$	- espectro médio normalizado.
$\tilde{\bar{I}}_m(i_k, j_k, n)$	- valor do espectro médio normalizado referente a n-ésima banda.
$\bar{\bar{I}}_m(i_k, j_k)$	- Média do vetor espectro $\tilde{\bar{I}}_m(i_k, j_k)$.
$\bar{\bar{I}}_k$	- Vetor espectro de referência no teste da uniformidade espacial.
$\tilde{\bar{I}}_m(i_r, j_r, n)$	- n-ésima componente do vetor espectro médio com a média espectral subtraída: $\tilde{\bar{I}}_m(i_r, j_r, n) = \bar{I}_m(i_r, j_r, n) - \bar{\bar{I}}_m(i_r, j_r)$
$\bar{\bar{I}}_{ref}^{CE}$	- espectro de referência no teste da redundância pela CE.
$\bar{\bar{\bar{I}}}_{ref}^{CE}$	Espectro de referência médio no teste da redundância pela CE.
$\bar{\bar{I}}_{ref}^{DE}$	espectro de referência no teste da redundância pela DE.
$\bar{\bar{I}}_{mk}$	- espectro médio calculado com os espectros pertencentes à vizinhança $V(i_k, j_k)$.
$\bar{\bar{I}}_{mk_c}$	espectro médio calculado com os espectros pertencentes à vizinhança $V(i_{k_c}, j_{k_c})$.
$\bar{\bar{I}}_m(i_k, j_k)$	vetor espectro médio, calculado em uma vizinhança da amostra espectral $\bar{I}(i_k, j_k)$.
$\bar{\bar{I}}_{mr}(i_k, j_k)$	- vetor espectro médio da semi-vizinhança $V_r(i_k, j_k)$, $r=0, 1$.
$\bar{\bar{I}}_{md}(i_k, j_k)$	- vetor espectro médio calculado através dos espectros derivados.
$\bar{\bar{I}}_{mw}(i_k, j_k)$	- vetor espectro médio calculado através dos espectros processados pela TWD.
I_k	matriz de espectros, com dimensão $(M+1) \times (M+1) \times N$, - contendo os espectros na vizinhança $(M+1) \times (M+1)$ do pixel (i_k, j_k) .
$I \times J$	- dimensão das imagens utilizadas nos testes.

K	- número inicial de amostras espectrais candidatas a MR.
\mathbb{k}	- número de amostras espectrais que passaram pelo teste da triagem discriminatória.
K_U	- número de amostras espectrais após o teste da uniformidade espacial.
K_H	- número de amostras espectrais após o teste da homogeneidade espectral.
K_C	- número de amostras espectrais após o teste do contexto espacial.
K_R	- número de amostras espectrais após o teste da redundância.
$L(e_1, e_2, \dots, e_R)$	- matriz covariância relativa a um conjunto de amostras (e_1, e_2, \dots, e_R) , $e_1, e_2, \dots, e_R \in \{1, 2, \dots, \mathbb{k}\}$.
$L_{p,q}$	- elemento da matriz $L(e_1, e_2, \dots, e_R)$.
$M(i_k, j_k)$	- elementos ordenados em ordem crescente da média espectral de cada vetor espectro $\vec{I}(\cdot, \cdot)$ na vizinhança $(M+1) \times (M+1)$.
$(M+1) \times (M+1)$	- dimensões iniciais da vizinhança de uma amostra espectral.
$(M_C+1) \times (M_C+1)$	- dimensões da vizinhança relativa ao teste do contexto espacial.
M_k	- conjunto de valores médios $m(p, q)$ para: $i_k - M/2 \leq p \leq i_k + M/2$, $j_k - M/2 \leq q \leq j_k + M/2$.
$M_{(k)}$	- conjunto M_k ordenado em ordem crescente.
$m(p, q)$	- média do espectro $\vec{I}(p, q)$
N	- número total de bandas do conjunto de espectros interpolados.
p_r	- autovalor λ_r normalizado.
$Q_h(i_k, j_k)$	- indicador de homogeneidade espectral.
R	- número de Membros de Referência (MR).
\mathfrak{R}_\cap	- conjunto de amostras espectrais menos redundantes selecionadas pela distância euclidiana e pela coerência

\mathfrak{R}_\cup	- conjunto de amostras espectrais menos redundantes selecionadas pela distância euclidiana ou pela coerência
\mathfrak{R}_{DE}	- conjunto de amostras espectrais menos redundantes selecionadas pela distância euclidiana.
\mathfrak{R}_{CE}	- conjunto de amostras espectrais menos redundantes selecionadas pela coerência espectral..
\mathfrak{R}_\forall	- conjunto de amostras espectrais obtidos por \mathfrak{R}_\cap , \mathfrak{R}_\cup , \mathfrak{R}_{DE} ou \mathfrak{R}_{CE} .
\mathfrak{R}_1	- limitante superior relativo ao número máximo de MR, relativo a amostras bem configuradas.
\mathfrak{R}_2	- limitante superior relativo ao número máximo de MR, relativo a mínima entropia aceitável.
$\vec{S}_r(i_{k_u}, j_{k_u})$	- vetor variância da semi-vizinhança $V_r(i_k, j_k)$, $r=0, 1$.
$t(i_k, j_k, n)$	- v.a. com distribuição t-Student, relativa a n-ésima banda do pixel (i_k, j_k) , no processo de verificação da igualdade de duas médias do teste da homogeneidade espectral.
U_{bc}^R	- conjunto de amostras espectrais bem configuradas.
U_{DE}	- conjunto formado pelas distâncias de pares de amostras.
$U_{(DE)}$	- conjunto U_{DE} ordenado em ordem crescente.
U_{CE}	- conjunto formado pelas coerências de pares de amostras.
$U_{(CE)}$	- conjunto U_{CE} ordenado em ordem decrescente.
U_H	- conjunto formado pelas entropias de pares de amostras.
$U_{(H)}$	- conjunto U_H ordenado em ordem crescente.
V	- Matriz contendo, como colunas, os autovetores da matriz $L(e_1, e_2, \dots, e_R)$.

V_k	- conjunto de <i>pixels</i> pertencentes a $V(i_k, j_k)$.
\vec{V}_r	- Autovetor da matriz $L(e_1, e_2, \dots, e_R)$.
$V(i_k, j_k)$	- vizinhança espacialmente uniforme do pixel (i_k, j_k) , formada por vetores espectros.
$\#V(i_k, j_k)$	- cardinalidade de $V(i_k, j_k)$.
$V_C(i_k, j_k)$	- vizinhança para o teste do contexto do pixel (i_k, j_k) .
$\#V_C(i_k, j_k)$	cardinalidade de $V_C(i_k, j_k)$.
$V_0(i_k, j_k)$ e	- Semi-vizinhança do pixel (i_k, j_k)
$V_1(i_k, j_k)$	- Semi-vizinhança do pixel (i_k, j_k) , $V_0(i_{k_u}, j_{k_u}) \cap V_1(i_{k_u}, j_{k_u}) = \emptyset$, $V_0(i_k, j_k) \cup V_1(i_k, j_k) = V(i_k, j_k)$
$W_s^{par \text{ ou } imp}(p, q)$	- coeficientes pares ou ímpares da TWD para o nível s de decomposição.
\vec{w}	- vetor erro coluna com dimensão N .

Gregos

α	- grau de significância α no teste estatístico de igualdade de médias no teste de homogeneidade espectral
α_u	- fator de uniformidade espacial, $0,5 < \alpha_u \leq 1$, sendo $\alpha_u (M+1)^2$ o limiar de uniformidade espacial.
α_c	- fator de contexto, $0 \leq \alpha_c \leq 1$, sendo $\alpha_c (M_c+1)^2$ o limiar do teste de contexto espacial.
α_{DE}	- fator de configuração relativo a distância euclidiana, $0 \leq \alpha_{DE} \leq 1$.
α_{CE}	- fator de configuração relativo a coerência espectral, $0 \leq \alpha_{CE} \leq 1$.
α_H	- fator de configuração relativo a entropia, $0 \leq \alpha_H \leq 1$.
$\Delta_{DE}(e_1, \dots, e_R)$	- distância $DE(e_1, \dots, e_R)$ modificada.

- $\Delta_{CE}(e_1, \dots, e_R)$ - coerência $CE(e_1, \dots, e_R)$ modificada.
- $\Delta_H(e_1, \dots, e_R)$ - entropia $H(e_1, \dots, e_R)$ modificada.
- $\Delta DE_{(k_c-1)}$ - Razão de variação da distancia.
- $\Delta CE_{[k_c]}$ - razão de variação da coerência.
- $\bar{\epsilon}_r$ - vetor erro do MLME.
- λ - comprimento de onda em nm.
- λ_r - autovalor real e não nulo.
- $\bar{\Phi}$ - vetor coluna, de dimensão R , formado pelas frações abundâncias.
- ϕ_r - fração abundância, escalar associado ao r -ésimo MR
- η_{CE} - limiar de configuração para a distância euclidiana.
- η_{CE} - limiar de configuração para a coerência espectral.
- η_{CE} - limiar de configuração para a entropia.
- Ω_k - conjunto de espectros da matriz I_k , ou seja, conjunto de espectros na vizinhança $(M+1) \times (M+1)$ do pixel (i_k, j_k) .
- Ω_U - conjunto de espectros que passaram pela triagem realizada pelo teste da uniformidade espacial.
- Ω_H - conjunto de espectros que passaram pelo teste da homogeneidade espectral.
- Ω_C - conjunto de espectros que passaram pelo teste do contexto espacial.
- Ω_R - conjunto de espectros que passaram pelo teste da redundância.
- ψ_e - limiar de seleção espacial, $0 \leq \psi_e \leq 1$.
- ψ_h - limiar de homogeneidade espectral, $0,5 < \psi_h \leq 1$
- ψ_{rde} - limiar de redundância da distância euclidiana, $0 \leq \psi_{rde} \leq 1$
- ψ_{rce} - limiar de redundância da coerência espectral, $0 \leq \psi_{rce} \leq 1$

LISTA DE SIGLAS E ABREVIATURAS

<i>a</i>	- classe água
AESNA	- Amostragem Estratificada Sistemática não Alinhada
AMEE	- <i>Automated Morphological Endmember Extraction</i>
ASTER	- <i>Advanced Spaceborne Thermal Emission and Reflection Radiometer</i>
AVIRIS	- <i>Airborne Visible Infrared Imaging Spectrometer</i>
CE	- Coerência Espectral
CLS	- <i>Constrained Least Squares</i>
Coifz	- Coiflet-z, <i>wavelet</i> mãe Coiflet de ordem <i>z</i>
Dbz	- Daubechies-z, <i>wavelet</i> mãe Daubechies de ordem <i>z</i>
DE	- Distância Euclidiana
EOS	- <i>Earth Observing System</i>
FCLSU	- <i>Fully Constrained Linear Spectral Unmixing</i>
G-SO	- <i>Gram-Schmidt Orthogonalization</i>
<i>H</i>	- entropia
HE	- Homogeneidade Espectral
IDL	- <i>Interactive Data Language</i>
IFOV	- <i>Instantaneous Field of View</i>
IVM	- InfraVermelho Médio
IVP	- InfraVermelho Próximo
LSE	- Least Squares Estimation
MD	- Medida de Divergência
MEA	- Mapeamento Espectral Angular
MESMA	- <i>Multiple Endmember Spectral Mixture Analysis</i>
MEST	- <i>Manual Endmember Selection Tool</i>
MLME	- Modelo Linear de Mistura Espectral

MME	- Modelo de Mistura Espectral
MNLME	- Modelo Não Linear de Mistura Espectral
MODIS	- <i>Moderate Resolution Imaging Spectroradiometer</i>
MR	- membro(s) de referência (" <i>endmember</i> ")
ORASIS	- <i>Optical Real-Time Adaptive Spectral Identification</i>
PPI	- <i>Pixel Purity Index</i>
ENVI	- <i>Environment for Visualizing Images</i>
SAM	- <i>Spectral Angle Mapper</i>
RQ	- classe NEOSSOLO QUARTZARÊNICO
UE	- Uniformidade Espacial
LV	- classe LATOSSOLO VERMELHO
SNR	- <i>Signal to Noise Ratio</i>
NV	- classe NITOSSOLO VERMELHO
SVD	- <i>Singular Value Decomposition</i>
TCP	- Transformação por Componentes Principais
TWD	- Transformada <i>Wavelet</i> Discreta
V	- Vermelho
VIS	- VISível
vs	- classe vegetação seca
vv	- classe vegetação verde
WLS	- Weighted Least Squares

1 Introdução

Apresenta-se neste capítulo o conceito de Membros de Referência (MR) em imagens hiperespectrais e os principais métodos de identificação dos MR de uma cena. Descreve-se o objetivo do trabalho realizado e como ele está organizado.

1.1 Membros de Referência em Imagens Hiperespectrais

A necessidade de descobrir e classificar materiais na superfície da Terra é um tema comum em muitas aplicações de espectroscopia de imageamento. O atendimento a essas necessidades já faz parte da concepção das novas gerações de sensores remotos, tais como o ASTER (*Advanced Spaceborne Thermal Emission and Reflection Radiometer*), o MODIS (*Moderate Resolution Imaging Spectroradiometer*) e o sensor Hyperion, que constitui parte do programa de tecnologias revolucionárias proposto pela NASA para este terceiro milênio do programa EOS (*Earth Observing System*). Esses e outros sensores de alta resolução espectral discriminam mais classes de alvos na superfície e, conseqüentemente, pode-se obter um melhor entendimento sobre a natureza dos materiais que cobrem a superfície da Terra.

Para que se possa melhor discriminar as classes de alvos presentes na superfície há necessidade de se determinar os espectros de elementos menos contaminados na cena (membros de referência ou *endmembers*). De acordo com a definição dada por Schowengerdt (1997) idealmente o MR (membro de referência) representa o espectro puro para uma classe de alvo. O conjunto de MR escolhido deve ser capaz de explicar e abranger a variabilidade espectral inerente a uma dada cena (KESHAVA; MUSTARD, 2002). A seleção adequada dos tipos de MR, bem como o seu número, permitem uma melhor acurácia na estimação das imagens-fração abundâncias no modelo de mistura (ROBERTS et al., 1998).

Com o sensoriamento remoto hiperespectral é possível detectar uma grande variedade de alvos presentes na superfície. Entretanto, a baixa resolução espacial de uma imagem pode comprometer extração dos MR e a detecção e a classificação correta dos alvos. A baixa resolução agrava a mistura dos materiais no *pixel*, fazendo com que este se torne mais contaminado ainda. A suposição de que um *pixel* pertence a uma classe possivelmente nunca poderá ser garantida independente da resolução espacial, pois a mistura estará presente em menor ou maior grau no *pixel* da cena imageada.

Por exemplo, o espectrômetro imageador AVIRIS (*Airborne Visible InfraRed Imaging Spectrometer*) quando adquire imagens a 20 km de altura o *pixel* tem 20 x 20 m de resolução espacial. Em termos práticos a área compreendida pelo *pixel* está coberta inevitavelmente por mais de um tipo de material resultando em uma imagem com *pixels* tipicamente misturados.

Um problema de considerável interesse na comunidade de imageamento hiperespectral nos recentes anos, visto a sua relevância, tem sido o desenvolvimento de algoritmos matemáticos de identificação dos MR em cenas hiperespectrais (KESHAVA; MUSTARD, 2002).

Os principais métodos de busca dos MR são: a) Forma interativa (supervisionada): o procedimento mais simples consiste em localizar, manualmente, os *pixels* na imagem na qual se conhece a priori que a abundância de um determinado MR é alta (DALTON et al., 2001); b) Outro procedimento iterativo é o chamado de Método Manual para Seleção de MR, denominado MEST (*Manual Endmember Selection Tool*) (BATESON; CURTIS, 1996), e o SSEE (*Spatial-Spectral Endmember Extraction*), os quais, segundo os autores, podem ser totalmente automatizados (ROGGE et al., 2007); c) Método da Análise da Mistura Espectral com Múltiplos MR, MESMA (*Multiple Endmember Spectral Mixture Analysis*) (ROBERTS et al., 1998); d) Dentre os métodos automáticos (não-supervisionados): citam-se o N-FINDR (WINTER, 1999); a Análise do Erro Interativo, IEA (*Iterative Error Analysis*) (STAENZ; SZEREDI; SCHWARZ, 1998; NEVILLE et al., 1999); a Extração Morfológica Automática do MR, AMEE (*Automated Morphological Endmember Extraction*) (PLAZA, et al., 2002); o Sistema de Identificação Espectral Óptico Adaptativo em Tempo Real, ORASIS (*Optical Real-Time Adaptive Spectral Identification*) (BOWLES et al., 1995; PALMADESSO et al., 1995) e o método Assistente de Classificação Hiperespectral (Spectral Mapping Wizard) que é um assistente de classificação automática para imagens hiperespectrais, na qual aplica-se a metodologia Fração por Ruído Mínimo (*Minimum Noise Fraction*) e a metodologia do Índice de Pureza de *Pixel* (*Pixel Purity Index*) (ENVI, 2006).

1.2 Objetivos do Trabalho

Neste trabalho propõe-se uma nova metodologia para seleção de MR em imagens hiperespectrais. O método proposto é composto por três fases.

Na primeira fase é feita a triagem das amostras presentes na imagem hiperespectral de modo a identificar quais amostras farão parte de um conjunto de amostras candidatas a MR. Para essa triagem de amostras na imagem consideram-se três abordagens:

- a) **manual:** realizada por um especialista que escolhe empiricamente, baseado na sua experiência e habilidade, quais devem ser as amostras da imagem hiperespectral que potencialmente podem se tornar MR;
- b) **aleatória:** definindo-se a priori um número de amostras e escolhendo-se aleatoriamente essas amostras, com a restrição de que elas fiquem espalhadas uniformemente pela cena considerada e
- c) **discriminatória:** fazendo-se com que as amostras escolhidas possuam uma vizinhança com uniformidade espacial, homogeneidade espectral, um dado contexto espacial e que sejam pouco redundantes entre si. Os critérios e definições de uniformidade, homogeneidade e contexto da vizinhança de uma amostra e da redundância entre amostras são propostas feitas neste trabalho (PIZARRO; FERNANDES, 2007a).

Uma vez identificado na imagem hiperespectral o conjunto de amostras candidatas a MR, manualmente, aleatoriamente ou discriminatoriamente, passa-se para a segunda fase. Nesta fase escolhe-se, dentre as amostras candidatas, o conjunto de MR para a imagem hiperespectral utilizando-se múltiplas medidas de similaridades espectrais e o conceito de amostras bem configuradas. Estas medidas são a Distância Euclidiana (*DE*) e a Coerência Espectral (*CE*) entre dois espectros e a Entropia entre um conjunto de espectros. A medida de similaridade pela entropia e o conceito de amostras bem configuradas também são propostas deste trabalho (PIZARRO; FERNANDES, 2007b).

Na terceira fase estabelece-se um limitante superior para o número máximo de MR para uma dada cena, utilizando-se o conceito de amostras bem configuradas e de entropia mínima para um conjunto de amostras candidatas a MR.

1.3 Organização do trabalho

Este trabalho está dividido em nove capítulos, correspondendo esta introdução ao primeiro capítulo. No Capítulo 2 descreve-se o comportamento espectral de alvos de interesse e discute-se o conceito de espectroscopia de imageamento. No Capítulo 3 descreve-se a área teste e sua imagem hiperespectral obtida pelo sensor AVIRIS.

No Capítulo 4 apresenta-se o método de triagem de amostras espectrais, proposto neste estudo, enfatizando-se os conceitos de uniformidade espacial, homogeneidade espectral, contexto espacial e redundância. Com a triagem é possível o descarte de espectros com falhas na sua aquisição e a rejeição de amostras com alta mistura de classes. O resultado da triagem é um conjunto de amostras candidatas a MR.

O Capítulo 5 apresenta o processo de condicionamento dos espectros, baseado na derivada e na Transformada Wavelet Discreta (TWD). O condicionamento pela TWD é uma proposta deste trabalho. O condicionamento dos espectros é uma opção que para algumas aplicações pode ser conveniente (TSAI e PHILPOT, 2002; DEMETRIADES-SHAH et al., 1990; ZHANG, RIVARD; SANCHEZ-AZOFEIFA, 2004; DEBBA et al. 2006).

No Capítulo 6 define-se o conceito de amostras espectrais bem configuradas e propõe-se o método de busca de MR utilizando-se a medida de similaridade espectral, entre um conjunto de espectros, pelo cálculo da entropia. O conceito de entropia utilizado é uma proposta deste trabalho. Propõe-se ainda neste capítulo dois limitantes superiores para o número máximo de MR relativos a uma imagem, um destes limitantes é estabelecido em função do conceito de amostras bem configuradas e o outro em função da entropia mínima aceitável para um conjunto de amostras.

No Capítulo 7 faz-se a proposta de busca dos MR utilizando-se em conjunto a DE a CE e a entropia. O processo de múltiplas medidas procura explorar as particularidades de cada medida utilizada, de modo que haja complementaridade de propriedades que levem a uma melhor escolha dos MR representativos de uma dada cena.

No Capítulo 8 são apresentados os testes de desempenho e as análises dos resultados referentes aos processos propostos neste trabalho. Neste capítulo comparam-se os resultados obtidos com os obtidos pelo o algoritmo PPI (*Pixel Purity Index*).

No Capítulo 9, comentários finais, descreve-se o trabalho realizado e enfatiza-se as novas propostas e os resultados obtidos. São feitas ainda considerações para a continuidade do estudo realizado.

No Apêndice A apresenta-se a estrutura do processador, implementado com o aplicativo IDL[®], que faz a triagem das amostras espectrais e identifica os MR em imagens hiperespectrais.

2 Espectroscopia de Imageamento

Nesse capítulo apresenta-se o conceito do sensoriamento remoto hiperespectral. Faz-se a análise geral de alguns tipos de alvos e de suas respostas espectrais. Aborda-se a questão das informações espacial-espectral e discute-se o conceito de mistura espectral. Faz-se a conceituação de Membros de Referência (MR) e da sua utilização no Modelo de Mistura Espectral (MME) e em especial no Modelo Linear de Mistura Espectral (MLME).

2.1 Conceito do sensoriamento remoto hiperespectral

O conceito do sensoriamento remoto hiperespectral está ilustrado na Fig. 2.1. De uma mesma área várias imagens (centenas) podem ser obtidas simultaneamente, o que possibilita a extração de um espectro de reflectância, praticamente contínuo, para cada *pixel* da cena. Com isso fica mais fácil identificar as bandas de absorção mineral estreitas, que podem indicar a presença de certos materiais na cena. Essa boa reconstrução das bandas de absorção com os dados hiperespectrais não aparecem, via de regra, nos dados dos sensores multiespectrais, tais como aqueles oriundos dos satélites Landsat, CBERS ou SPOT, devido à baixa resolução espectral destes sensores imageadores.

Os dados espectrais portam informações correspondentes à composição da superfície e à composição da atmosfera. Constata-se, devido a esse fato, que através da espectroscopia de imageamento, mesmo com sensores hiperespectrais, é praticamente impossível a reconstrução das feições de absorção em torno de 1.400 nm e 1.900 nm, devido à alta concentração de vapor de água presente na atmosfera.

Com o advento do sensoriamento remoto hiperespectral abriu-se a possibilidade de se obter espectros na faixa de 400 a 2.500 nm do espectro óptico, o que era possível somente por espectrorradiômetros de laboratório/campo (ELVIDGE, 1988a; LACAPRA et al., 1996; MARTIN; ABER, 1997). Pelo processamento computacional dos dados hiperespectrais é possível produzir uma imagem de qualquer uma das bandas espectrais ou o espectro correspondente a qualquer um dos *pixels* da imagem (PORTER; ENMARK, 1987).

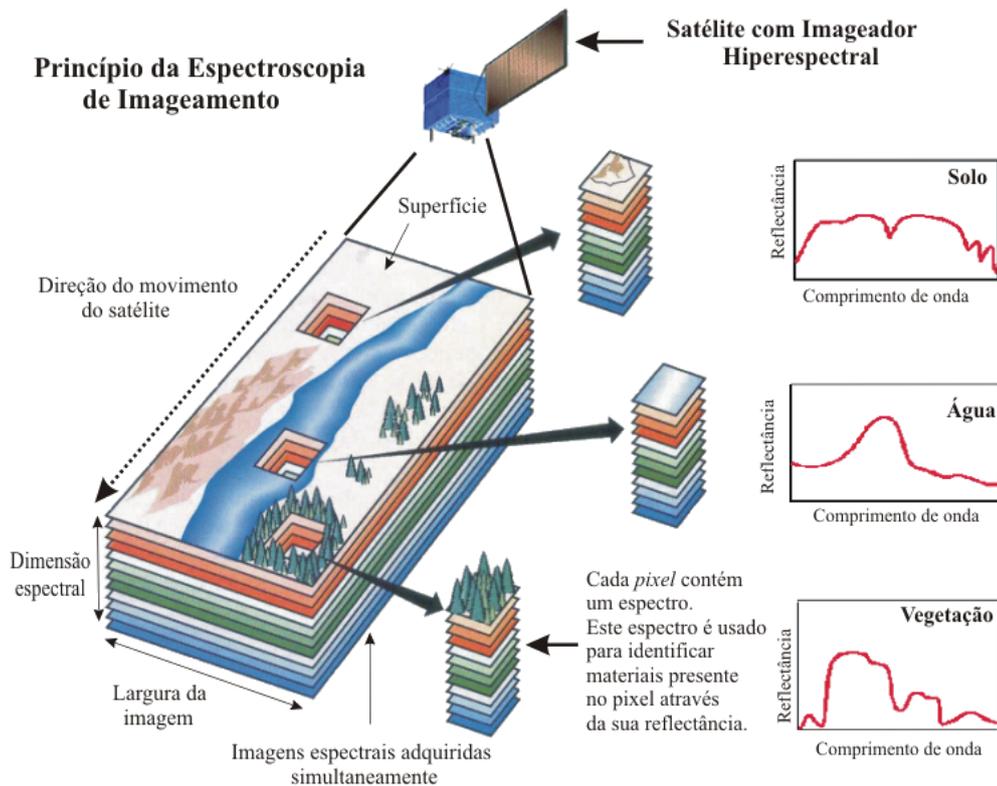


FIGURA 2.1 – Representação esquemática do sensoriamento hiperespectral.
 Fonte: adaptação a partir Shaw e Manolakis (2002, p. 12).

2.2 Resposta espectral da cobertura vegetal, do solo e da água

Cada objeto apresenta um comportamento frente à radiação nele incidente, que depende de diversos fatores, muitos dos quais não se podem controlar. “As frações de energia transmitida, refletida e absorvida pelos objetos variam com suas propriedades físico-químicas, estrutura e posição em relação à fonte de radiação” (NOVO, 1992). Mesmo assim, dentro de determinadas condições, é possível estabelecer a característica espectral de alguns alvos.

2.2.1 Reflectância da cobertura vegetal com e sem atividade fotossintética

A grande variação das propriedades ópticas deve-se às diferenças no conteúdo de água e às variações específicas nas concentrações de lignina e celulose entre as espécies vegetais. Geralmente os vegetais com pouca atividade fotossintética apresentam bandas de absorção por água em menor quantidade e intensidade do que a matéria verde (vegetal com atividade fotossintética), mas exibem feições de absorção de lignina-celulose e de outros componentes orgânicos (ASNER1998).

No estudo da resposta espectral de folhas verdes, constatou-se que no intervalo de comprimentos de onda do espectro óptico de 400 a 700 nm, faixa do Visível (VIS), a reflectância é baixa, da ordem de 10%, com um suave aumento da resposta na região do verde (550 nm). Na faixa do infravermelho próximo (IVP) de 700 a 1300 nm, ocorre outro aumento na reflectância para valores próximos a 50%. Daí em diante, até 2500 nm (infravermelho médio – IVM, 1000 a 2.500 nm), há um gradual decréscimo, aparecendo algumas feições de absorção pelo conteúdo de água líquida. No VIS a queda da reflectância está associada à absorção pelos pigmentos foliares, principalmente pela clorofila. Na região do azul, a absorção ocorre nas proximidades do comprimento de onda de 445 nm e está associada à presença dos pigmentos xantofila, carotenos, clorofilas α e β . Na região do vermelho (V), apenas a clorofila atua, absorvendo energia em torno de 645 nm. Enquanto isso, o aumento de reflectância no IVP é devido à estrutura interna da folha (tamanho e formato das células e à quantidade de espaços intercelulares). No IVM predominam bandas de forte absorção em 1400 e 1900 nm, devido ao conteúdo de água líquida das folhas, e feições de menor intensidade na região entre 700 e 1300 nm (GATES et al., 1965; KNIPLING, 1970), como as observadas em 980 e 1.190 nm (GAO; HEIDEBRECHT; GOETZ, 1993).

A presença de vegetação verde ou de resíduos não fotossinteticamente ativos pode ser facilmente alterada pelo tempo (variação do clima e/ou sazonalidade) e pela lavoura (preparação do terreno para plantio). As mudanças nestas condições afetam a resposta espectral dos solos. A Fig. 2.2 mostra as diferenças espectrais obtidas de um grupo de folhas de milho com diferentes conteúdos de umidade, desde a vegetação fotossinteticamente ativa (vegetação verde) com 66 a 100% de umidade até vegetação não-fotossinteticamente ativa (vegetação seca) com 0 a 40% de umidade.

Na Fig. 2.2, percebe-se muito bem as diferenças no comportamento devido ao estado fisiológico nas faixas do VIS e IVP e no começo do IVM, especialmente pelo forte decaimento do conteúdo da clorofila (remoção da absorção na borda do vermelho).

A Fig. 2.3 mostra espectros de laboratório, coletados com o espectrofotômetro UV-5240 (Beckman), para materiais extraídos do pinheiro. Nestes espectros, as feições de absorção provocadas pela lignina e celulose dominam entre 1.000 e 2.500 nm. As feições provocadas pela lignina são mais bem definidas nas curvas da folha escura e da casca cinza (ELVIDGE, 1990).

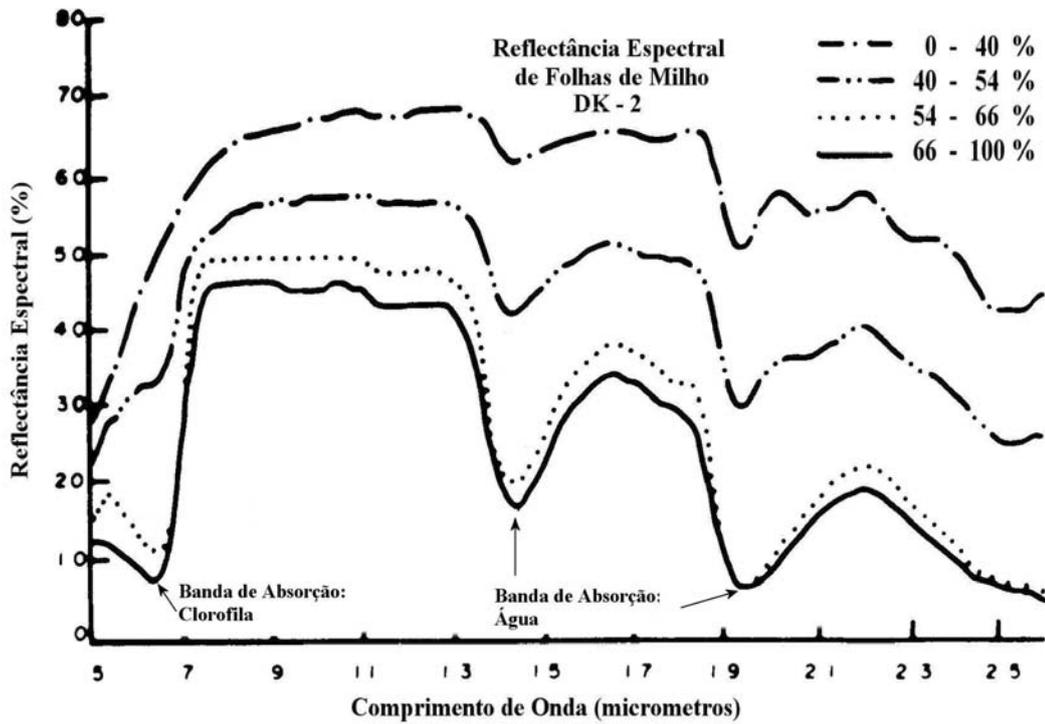


FIGURA 2.2 - Espectros de reflectância para diferentes conteúdos de umidade em folhas.
 Fonte: adaptação a partir de Kumar, 1972, p. 13).

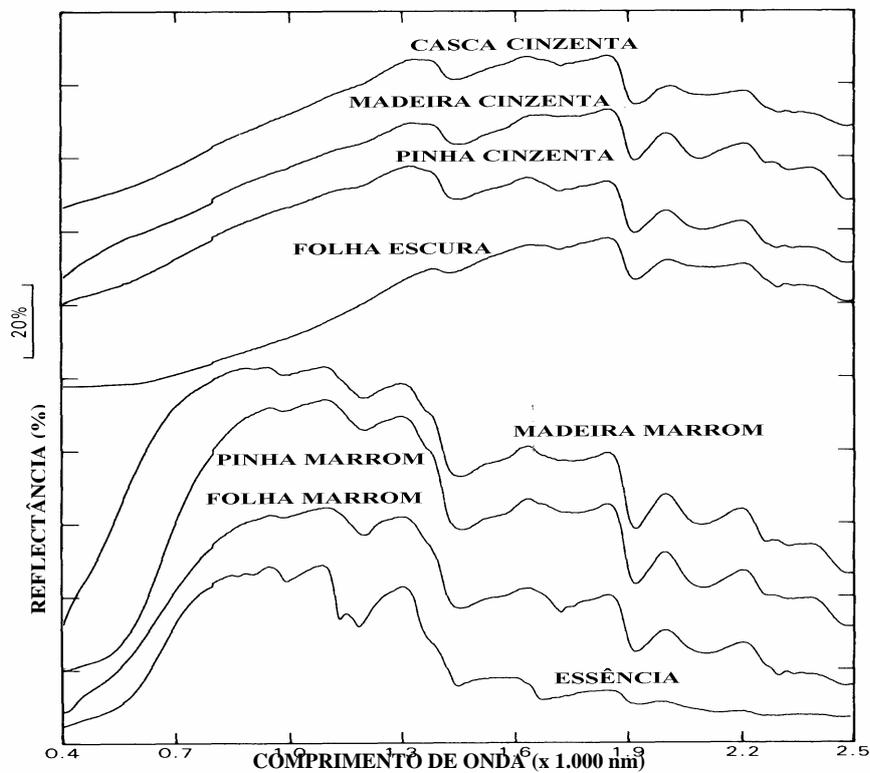


FIGURA 2.3 – Reflectância espectral de materiais secos obtidos do pinheiro.
 Fonte: adaptação a partir de Elvidge (1990, p. 1789).

Em relação à vegetação não-fotossinteticamente ativa ou seca (*vs*), à medida que a vegetação verde (*vv*) entra em processo de senescência, com a desidratação e a perda de pigmentos, a sua curva espectral é gradativamente modificada. A forte absorção no VIS é reduzida, com a perda de pigmentos, passando a apresentar maior reflectância. Nessa região a curva irá aparecer mais suave e crescente, assemelhando-se ao perfil de resposta do solo. No IVP, observa-se uma pequena redução na reflectância, motivada pela diminuição de espaços intercelulares, resultante do processo de desidratação. Finalmente, no IVM, ocorre a diminuição da intensidade das feições de absorção causadas pela água líquida da matéria viva, também resultante da desidratação. Os valores de reflectância dependem das características particulares de cada espécie vegetal, além do estágio fenológico, aspectos sanitários e de condições adversas do clima (como secas e geadas). Pode-se afirmar, contudo, que em geral a vegetação verde sadia apresenta um comportamento muito próximo do padrão de resposta de suas folhas, o que não significa dizer que um dossel terá o mesmo comportamento espectral do tipo de folha que o compõe, quando estudada individualmente. Tratando-se do estudo de um dossel, ou seja, da cobertura vegetal aglomerada, como é encontrado no campo, o número de variáveis aumenta, tornando o estudo mais complexo.

2.2.2 Reflectância do Solo

Quanto ao solo, observa-se um comportamento menos variável que a vegetação, em relação aos máximos e mínimos valores de reflectância. Alguns fatores que afetam a sua reflectância são: o conteúdo de umidade; a textura (proporção de areia, silte e argila); a rugosidade da superfície; a concentração de óxidos de ferro e o conteúdo de matéria orgânica (decomposta ou não) (LILLESAND; KIEFER, 1994). Pode-se acrescentar, também, a mineralogia e a estrutura do solo, a presença de grandes fragmentos, as incrustações, as sombras e as práticas culturais (HUETE, 1989). A textura, a mineralogia, a presença de óxidos de ferro e o conteúdo de matéria orgânica são considerados como propriedades intrínsecas, enquanto os demais fatores são tratados como condições da superfície. De forma geral, a presença de umidade diminui a reflectância dos solos, sendo este efeito maior na faixa do IVM (LILLESAND; KIEFER, 1994).

Na faixa do IVP há a incidência de apenas duas bandas de fraca absorção, em 960 e 1200 nm (IRONS; WEISMILLER; PETERSEN, 1989). Solos argilosos também apresentam bandas de absorção devidas à hidroxila. Textura e umidade estão fortemente relacionadas. Solos de textura grosseira são normalmente bem drenados resultando em um baixo conteúdo

de umidade e uma reflectância mais alta; solos mal drenados, de textura fina, geralmente têm reflectância mais baixa. Para muitos solos, quando úmidos, a reflectância diminui pela metade (JACKSON; SLATER; PINTER JR., 1983).

A presença de matéria orgânica tem uma forte influência nos solos, geralmente diminuindo a reflectância em toda a região do VIS e do IVP. Se presente com um conteúdo maior do que 2% a matéria orgânica pode mascarar as características de absorção de outros constituintes do solo. Quando em conteúdo superior a 5% a curva de reflectância dos solos freqüentemente apresenta um perfil côncavo entre 500 e 1300 nm e, se acima de 20% (solos orgânicos), os solos têm um comportamento espectral variável, dependendo do grau de decomposição do material orgânico (IRONS; WEISMILLER; PETERSEN, 1989).

A rugosidade, a umidade, a presença de grandes fragmentos (como os resíduos vegetais), a estrutura do solo e as incrustações, ou seja as condições da superfície, são facilmente e freqüentemente alteradas pelo clima e pelos tratamentos culturais, acarretando assim um comportamento espectral muito variável do solo.

A Tabela 2.1 indica as principais feições de absorção que ocorrem nos espectros dos minerais mais importantes do grupo dos óxidos de ferro e os agentes responsáveis por tais feições.

As feições de absorção associadas com os óxidos de ferro são provocadas por transições eletrônicas, onde os elétrons em torno de um átomo passam a adotar uma nova configuração eletrônica, em função de uma grande quantidade de energia presente. Estas feições são observadas predominantemente na região do ultravioleta e do visível até aproximadamente 1.200 nm.

Segundo Drury (1993), o ferro é o mais abundante entre os metais de transição, que incluem cobre, níquel, cromo, cobalto, manganês, vanádio e titânio, e por isso seus efeitos sobre os espectros de reflectância, na região do visível, são os mais marcantes. A transição espectral mais freqüentemente observada ocorre entre o ferro e o oxigênio (migração de elétrons do ferro para o oxigênio). Ela é caracterizada por um pronunciado decréscimo da reflectância na região do visível, na direção do comprimento de onda do azul (HUNT, 1980). Stoner et al. (1980) observaram que a banda de absorção em 700 e 900 nm é resultante da presença de ferro férrico. Hunt, Salisbury e Lenhoff (1971) observaram que a maioria das

feições eletrônicas bem delineadas de óxidos de ferro em minerais e rochas pode ser atribuída a transições eletrônicas nos cátions de ferro.

TABELA 2.1- Bandas de absorção de alguns minerais do grupo dos óxidos de ferro. Fonte: adaptação a partir de Hunt; Salisbury e Lenhoff (1973, p.99); Clark, King e Klejwa (1990, p. 12.653-12.680); Epiphanyo et al. (1992, p. 35-41).

Minerais	Principais Feições Espectrais (nm)	Agente Provocador de Absorção
Hematita (Fe ₂ O ₃)	885, 650, 530 e 445	Transições eletrônicas das formas trivalentes de ferro (Fe ³⁺)
Goetita (FeO(OH))	917, 650, 480 e 435	Transições eletrônicas das formas trivalentes (Fe ³⁺)
Magnetita (Fe ₃ O ₄)	Mineral tipicamente opaco, praticamente sem eventos ou feições espectrais no visível, no infravermelho próximo e no infravermelho médio, causados por bandas de condução dos óxidos de ferro e de titânio que se estendem através do espectro.	Transferências de cargas Fe ²⁺ => Fe ³⁺ ; Fe ³⁺ => O; Fe ²⁺ => O e as transições intra -ions Fe ²⁺ ; Fe ³⁺
Ilmenita (FeTiO ₃)	Mesma situação anterior	Transferência de cargas Fe ²⁺ => O; Ti ⁺⁴ => O, Fe ²⁺ => Ti ⁺⁴ e as transições intra-iônicas Fe ²⁺

Stoner e Baumgardner (1980) analisaram amostras de solos sesquioxídicos, provenientes de regiões intensamente intemperizadas como as dos trópicos, e concluíram que os seus espectros de reflectância são dominados por feições devidas do ferro férrico. Os minerais hematita, limonita e goetita apresentavam bandas de absorção bem definidas próximo de 900 nm. Quando ocorreu magnetita notaram-se baixas reflectâncias no VIS e ausência de feições de absorção no IVP e IVM, em função de bandas de condução dos óxidos de ferro e de titânio, que se estenderam através do espectro. Bigham et al. (1978) observaram que o tipo e quantidade relativa de óxidos de ferro imprimem a cor vermelha e amarela dos solos ricos em argilas sesquioxídicas. Observou-se que predominantemente solos amarelos ricos em goetita adsorviam mais fosfatos por unidade de peso que solos similares vermelhos ricos em hematita.

Nas relações entre reflectância espectral e composição dos solos, Karmanov (1970) notou que a intensidade de reflectância dos hidróxidos de ferro, contendo pouca água e tendo cor marrom-avermelhada-escura, aumentou mais fortemente no intervalo de 554 a 596 nm.

Por outro lado, a intensidade de reflectância dos óxidos de ferro hidratados aumentou mais fortemente no intervalo de 500 a 540 nm.

Entre os componentes importantes dos solos estão também os minerais da fração argila, sendo que os principais estão listados na Tabela 2.2.

TABELA 2.2 Bandas de absorção de alguns minerais do grupo das argilas. Fonte: adaptação a partir de Hunt; Salisbury e Lenhoff (1973, p.99); Clark, King e Klejwa (1990, p. 12.653-12.680); Epiphany et al. (1992, p. 35-41).

Minerais	Principais Feições Espectrais	Agente Provocador de Absorção
Caulinita ($\text{Al}_2\text{SiO}_5(\text{OH})_4$)	2.700, 2.730, 2.755, 2.200 e 1.400	Transições vibracionais dos grupos das hidroxilas (OH) da rede cristalina
Gibbsita ($\text{Al}_2\text{O}_3 \cdot 3\text{H}_2\text{O}$)	2.975, 2.960, 2.917, 2.842, 2.765, 2.260; 1.900, 1.400 e 1.550	Mesma situação anterior
Montmorilonita Fórmula geral: ($1/2\text{Ca}, \text{Na}$) _{0,7} (Al, Mg, Fe) ₄ (Si, Al) ₈ O ₂ (OH) ₄ .nH ₂ O	2.200, 1.900 e 1.450	Mesma situação anterior

Considerando que muitos solos se diferenciam pela composição da fração argila, e que esta tem estruturas características, com conseqüente influência no comportamento espectral dos solos, muitos trabalhos vêm sendo desenvolvidos para estudar a influência espectral deste constituinte.

Como se pode observar pela Tabela 2.2, todos os minerais apresentam como agente provocador as transições vibracionais das hidroxilas, que resultam da interação entre a energia de radiação eletromagnética e as moléculas, ou das oscilações dos átomos e, também, de seu centro de equilíbrio. As evidências para tais transições podem ser observadas no IVP e IVM. Deste modo, transições vibracionais necessitam de uma quantidade menor de energia que a necessária nas transições eletrônicas, embora as feições espectrais sejam muitas vezes mais intensas do que as observadas nas eletrônicas. Dentre os grupos moleculares, são poucos os que originam feições de absorção vibracionais no infravermelho. Dentre eles, a hidroxila, Al-OH e Mg-OH, a água, os carbonatos e alguns sulfatos são os principais responsáveis por quase todas as feições espectrais dos materiais geológicos no intervalo de 1.200-3.000 nm. Em particular, para os Latossolos, os eventos espectrais ligados às hidroxilas são

particularmente importantes, pois os radicais OH- participam na rede cristalina da caulinita e da gibbsita, dois componentes maiores destes solos.

Mathews; Cunningham e Petersen (1973) relataram que o tipo de mineral de argila influencia a forma e a intensidade da curva de reflectância em toda a faixa estudada de 500 a 2.600 nm. Eles obtiveram diferentes curvas espectrais para a caulinita, nontronita e ilita, conforme ilustrado na Fig. 2.4. Segundo estes autores, a baixa reflectância em 1.400 e 1.900 nm refere-se à forte influência de água adsorvida.

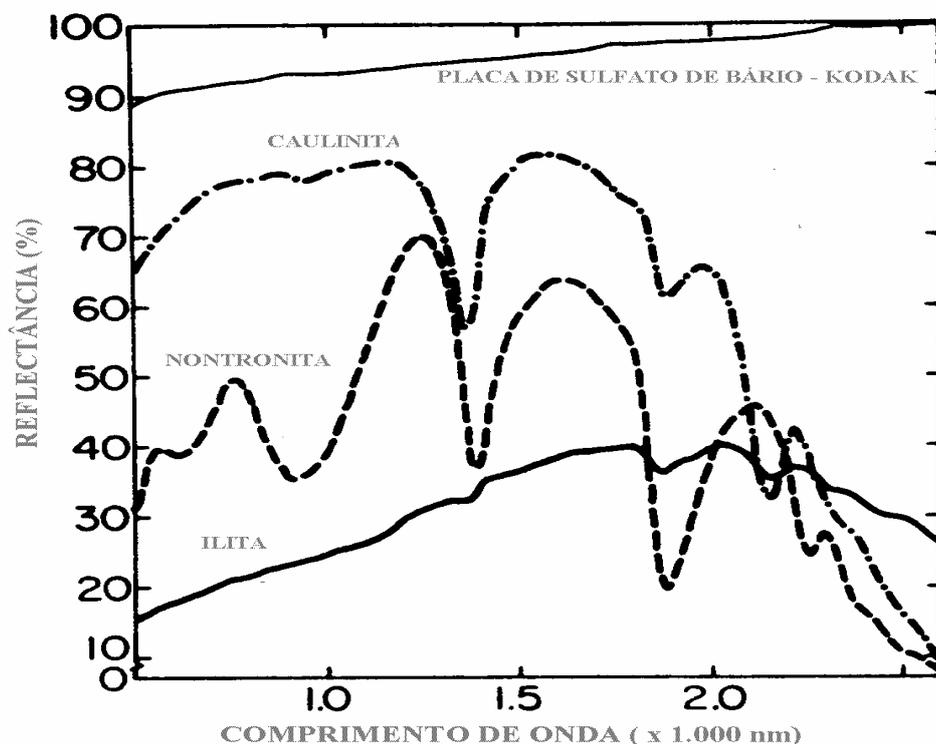


FIGURA 2.4 – Espectros de reflectância da caulinita, nontronita e ilita.

Fonte: adaptação a partir de Mathews; Cunningham e Petersen (1973, p. 422).

O mesmo resultado foi observado por Lindberg e Snyder (1972), que observaram uma redução na reflectância da montmorilonita e nontronita com conseqüentes bandas de absorção, principalmente como resultado de água adsorvida nesses minerais. Hunt e Salisbury (1970) observaram bandas moleculares de água muito fortes em 1.400 e 1.900 nm causadas pela água retida típica das montmorilonitas.

Stoner e Baumgardner (1980) mostraram que nos espectros dos argilo-minerais dos grupos montmorilonita e caulinita, as principais feições de absorção devem-se à presença de

água. A ocorrência de fortes feições de absorção em 1.400 e 2.200 nm, devidas à hidroxila, são típicas da reflectância da caulinita.

2.2.3 Reflectância de alvos contendo sombra e água

A sombra representa a ausência de informação, devido à absorção total ou à falta de incidência de radiação. Em geral, mesmo regiões cujo fluxo direto de iluminação é impedido de alcançar o alvo não apresentam reflectância nula, devido à incidência de uma pequena porção de radiação resultante da componente difusa. Além desse aspecto, a intensidade do sombreamento também depende da opacidade do material que o provoca, podendo variar com o comprimento de onda, como no caso das folhas, que apresentam maior opacidade na região do V do que na região do IVP.

De um modo geral a água líquida apresenta uma baixa reflectância, menor do que 10% na faixa compreendida entre 380 e 700 nm e máxima absorção acima de 700 nm. A assinatura espectral de um corpo d'água resulta dos processos de absorção e espalhamento da radiação solar no seu interior, portanto a radiância detectada por um sensor pode fornecer informações sobre suas características físico-químicas e biológicas (STEFFEN; MORAES; GAMA, 1996).

2.3 A questão da informação espacial e espectral

A cada *pixel* da imagem hiperespectral está associado um espectro. O contexto espacial associado ao espectro do *pixel* pode ser levado em consideração através dos espectros das amostras na sua vizinhança. A título de exemplo na Fig. 2.5 mostra-se uma vizinhança 3×3 do *pixel* (*i, j*) de uma imagem. Os *pixels* são representados pelo vetor espectro coluna $\vec{I}(i, j)$:

$$\vec{I}(i, j) = [I(i, j, 1) I(i, j, 2) I(i, j, 3) \dots I(i, j, N)]^T \quad (2.1)$$

onde T denota o operador transposto e N é o número de bandas.

A imagem considerada é composta de uma coleção de *pixels* (*i, j*) onde *i* e *j* são as coordenadas do *pixel* na imagem com $i = 1, 2, \dots, I$ e $j = 1, 2, \dots, J$, sendo I e J as dimensões da imagem. A vizinhança é uma importante característica da técnica do processamento de

imagens no domínio espacial e que aparece normalmente denominada na literatura como núcleo, *kernel* (KATKOVNIK; SHMULEVICH, 2002).

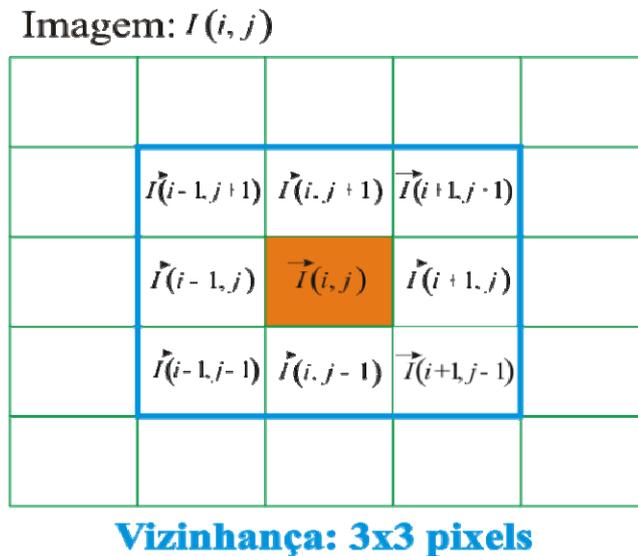


FIGURA 2.5 – Vizinhança 3 x 3 da amostra espectral (i, j) .

Inúmeros métodos de classificação em imagens de sensoriamento remoto baseiam-se na determinação das classes presentes na imagem através da classificação *pixel a pixel*. Esse tipo de classificação não leva em consideração o contexto presente no conjunto dos *pixels* vizinhos (JONG; HORNSTRA; MAAS, 2001). Por outro lado, técnicas que levam em consideração as características espaciais nas análises melhoram a precisão da classificação, embora isto acarrete o aumento do tempo computacional devido ao aumento da complexidade numérica e no volume de dados a serem analisados (LANDGREBE, 1980).

A limitação e a dificuldade do modelamento do contexto espacial fazem também com que as técnicas espectrais ainda sejam as mais utilizadas (JIA; RICHARDS, 1993). Além disso, as técnicas baseadas exclusivamente no domínio espacial possuem limitações na caracterização precisa dos alvos da cena devido à alta resolução espacial requerida dos sensores, resolução esta ainda não disponível nos imageadores atuais devido ao alto custo de construção (LANDGREBE, 2002). Na atualidade os imageadores hiperespectrais tem a sua resolução espacial degradada em favor da melhora da resolução espectral e por decorrência deste fato as imagens são mais bem adaptadas para aplicações na quais as informações espectrais são mais úteis que as informações espaciais (MANOLAKIS; MARDEN; SHAW, 2003).

Os autores Madhok e Landgrebe (1999), Pinzón, Ustin e Pierce (1998) e Jimenez et al. (2005) têm argumentado que a principal limitação das técnicas exclusivamente espectrais, p. ex., a *Spectral Angle Mapper* (SAM) (BOARDMAN; KRUSE; GREEN, 1995) reside no fato de que o processo não leva em consideração a correlação espacial presente na imagem. De fato, segundo Madhok e Landgrebe (1999), as informações espaciais e espectrais complementam-se mutuamente de modo a compensarem suas deficiências particulares.

Existem vários trabalhos (MADHOK; LANDGREBE, 1999; PINZÓN; USTIN; PIERCE, 1998; GREEN; BOARDMAN, 2000; BIELSKI et al., 2002) mostrando a necessidade de se construir algoritmos capazes de combinar simultaneamente as informações espacial e espectral para analisar imagens hiperespectrais, de forma que nenhuma delas seja favorecida ao longo do processo. Pois, se um dos processamentos for realizado antes do outro, ou seja, de forma separada, pode-se obter resultados errados na classificação (PLAZA, 2002).

Em particular neste último trabalho onde é proposto e avaliado o algoritmo AMEE, método não-supervisionado, na determinação dos MR por operações morfológicas, foram obtidos resultados iguais ou melhores do que com outros algoritmos de extração de MR. Mas, ele tem um significativo aumento computacional quando o tamanho da vizinhança aumenta e o algoritmo é capaz de selecionar unicamente um *pixel* na vizinhança analisada (PLAZA, 2002; PLAZA, 2004a).

A questão espacial relaciona-se também com a caracterização da homogeneidade na vizinhança das amostras espectrais. Se os *pixels* em uma vizinhança exibem variabilidade espectral muito intensa, a caracterização espectral média da amostra, calculada na vizinhança, deixa de representar um alvo homogêneo. Por outro lado, a homogeneidade da vizinhança pode ser utilizada de modo que o espectro médio, calculado na vizinhança homogênea, represente bem essa condição espacial de homogeneidade espacial.

Na metodologia proposta neste trabalho os espectros são representados pelo espectro médio de uma vizinhança, da amostra considerada, que seja uniforme espacialmente e homogênea espectralmente. Os conceitos de uniformidade espacial e homogeneidade espectral são definidos no Capítulo 4. O uso conveniente da vizinhança ao invés de um simples *pixel*, no seu interior, também se mostra interessante no contexto de se ter uma maior representatividade da classe do alvo extenso, representada pelos *pixels* na vizinhança, além de se poder ter uma melhora na relação sinal-ruído (KRUSE, 1998).

A dimensão e a forma da vizinhança é um fator determinante no desempenho do processamento da imagem. É necessário analisar as características espaciais das formas e objetos contidos na imagem para que não haja degradação na sua resolução espacial, quando se utiliza o conceito de vizinhança para filtragens e extração de informações. Caso não se disponha de informações prévias sobre as feições de interesse na imagem, bem como as características espaciais dos objetos de interesse, a utilização de um tamanho único de vizinhança pode degradar o desempenho do processamento dificultando a caracterização da totalidade dos objetos presentes na cena.

Entre as técnicas mais utilizadas para solucionar as limitações anteriormente descritas tem-se o uso de vizinhanças com tamanho variável (BURT; ADELSON, 1983). Na literatura esta opção é englobada dentro do conjunto de técnicas denominadas decomposição em multiescala na imagem. Há múltiplas formas de realizar o processamento multiescala (GAUCH; PIZER, 1993; STRICKLAND; HAHN, 1997), mas a mais popular entre elas é certamente a que utiliza vizinhanças de tamanho variável (BISTER; CORNELIS; ROSENFELD, 1990).

Na proposta deste trabalho, que utiliza o espectro médio calculado sobre uma vizinhança homogênea, a vizinhança tem a sua forma variável, ou adaptativa, pois é função das características espacial e espectral da vizinhança do *pixel* considerado. A vizinhança variável proposta parte de uma vizinhança inicial pré definida de dimensões $(M + 1) \times (M + 1)$, com M par, e nesta vizinhança identifica os *pixels* que formam um conjunto homogêneo nas dimensões espacial e espectral. Portanto, a vizinhança final é variável e é uma função do contexto espacial e espectral dos *pixels* considerados na vizinhança pré-definida.

2.4 Conceito de *pixel* misturado

A existência de misturas ao nível de *subpixel*, ou seja, misturas de diferentes materiais dentro da dimensão de uma célula de resolução, é um fenômeno habitual em imagens hiperespectrais, como aparece ilustrado na Fig. 2.6, onde se mostra a diferença entre *pixels* puros e *pixels* misturados.

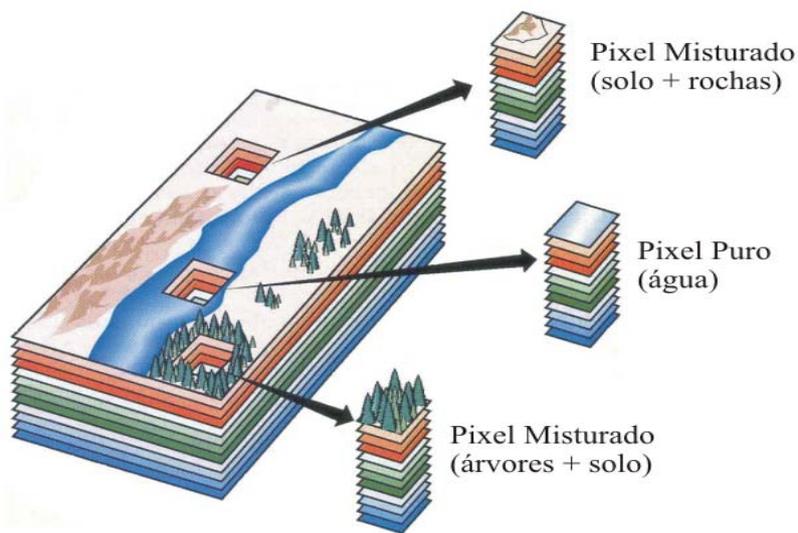


FIGURA 2.6 – Conceito de *pixel* puro e *pixel* misturado em imagens hiperespectrais.
 Fonte: adaptação a partir de Shaw e Manolakis (2002, p. 12).

O *pixel* misturado pode ser definido como aquele que contém diferentes tipos de coberturas (HSIEH; LEE; CHEN, 2001; KUSTAS; NORMAN, 2000; TU et al., 1999; PIZARRO; GALVÃO; EIPHANIO, 2001). Devido a grande variabilidade dos materiais presentes em uma cena pode-se afirmar que quase a totalidade dos *pixels* na imagem hiperespectral são misturados.

A mistura de classes ocorre em dois níveis:

- a) nível microscópico, onde ocorrem continuamente a mistura de diferentes materiais em níveis espaciais bem inferiores a dimensão da célula de resolução do sensor (KESHAVA; MUSTARD, 2002; JOHNSON, 1983). A Fig. 2.7 mostra esta situação que independe da resolução espacial do sensor. O fenômeno da mistura espectral é característico do mundo real e ocorre independente da escala considerada (VERHOEYE; WULF, 2002), e
- b) nível macroscópico, onde na dimensão do *pixel*, ou seja, dentro da célula de resolução do sensor, ocorrem misturas de diferentes classes de alvos, notadamente nas fronteiras entre as diferentes classes de alvos presentes na superfície. Assim, o espectro resultante adquirido pelo sensor será correspondente ao espectro misturado pelos diferentes materiais (componentes) presentes em uma célula de resolução (CHEN, 1999). A Fig. 2.8a e b ilustra a situação descrita.

Os espectros utilizados na Fig. 2.8b foram extraídos da imagem AVIRIS que será utilizada neste trabalho. Nesta figura observa-se como os espectros se diferenciam entre si pela forma da reflectância.

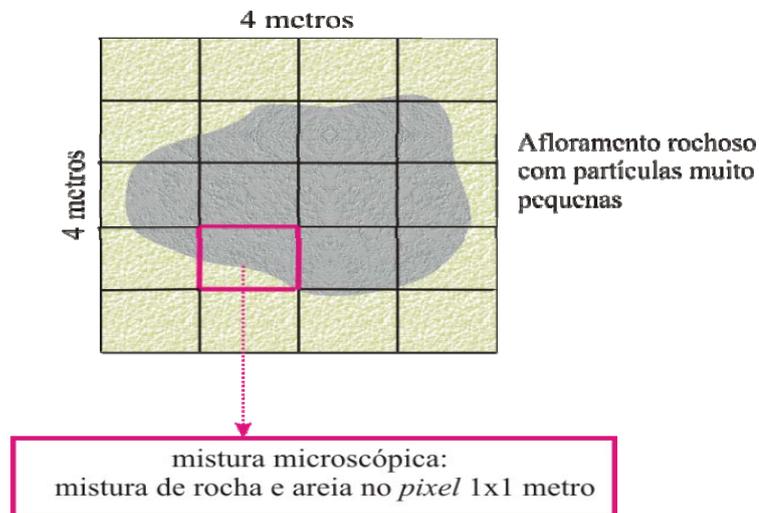


FIGURA 2.7 - Mistura microscópica entre materiais numa região inferior às dimensões da célula de resolução do sensor.

O conceito de mistura deve sempre ser considerado nas técnicas das análises espectrais. Na literatura científica há vários modelos que simulam as propriedades da mistura dos componentes de objetos pertencentes ao mundo real (BOREL; GERSTL, 1994, ICHOKU; KARNIELI, 1996).

2.5 Membros de referência (MR) e modelos de mistura espectral (MME)

Os MR são os espectros que podem representar a variedade espectral presente em uma cena. Segundo Schowengerdt (1997, p. 455) os *pixels* menos contaminados podem ser considerados bons MR. Os MR estão normalmente associados às componentes macroscópicas claramente identificadas na imagem, como água, solo, vegetação, minerais, etc.

Essa representatividade atribuída aos MR implica na possibilidade de se descrever os demais espectros dos *pixels* da imagem hiperespectral uma função dos MR escolhidos. Um dos principais objetivos deste trabalho é precisamente desenvolver uma metodologia supervisionada para a adequada seleção dos MR.

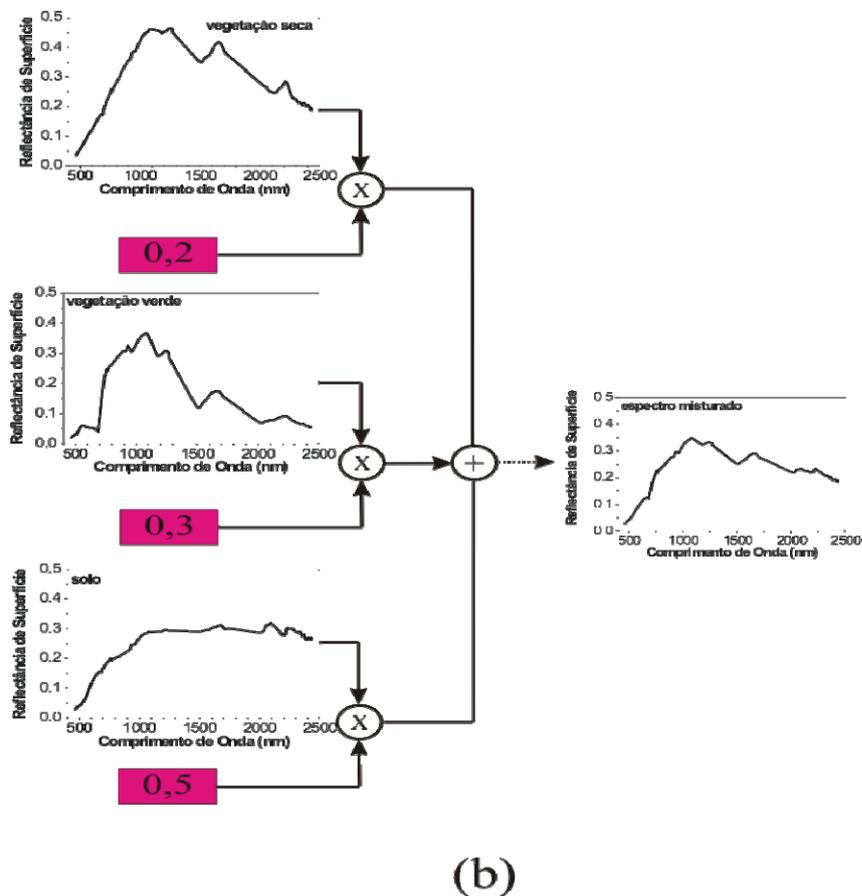
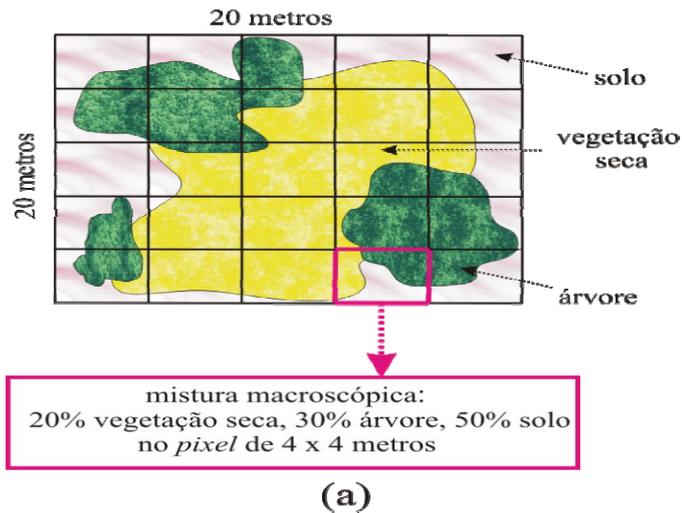


FIGURA 2.8 – Mistura espectral a nível macroscópico: (a) superposição de vários materiais em um *pixel* e (b) MLME representativo da combinação de espectros puros da vegetação seca, vegetação verde e solo.

Os MME que utilizam os MR para representarem os espectros dos *pixels* de uma cena podem ser agrupados em duas categorias básicas. Na primeira categoria está o MLME na qual se assume que a reflectância em cada *pixel* pode ser determinada pela combinação linear da resposta espectral dos MR selecionados dando origem às frações abundâncias (contribuição

de cada MR na constituição do espectro decomposto) (ADAMS; SMITH; GILLESPIE, 1993). O MLME, associado às reflexões simples, é ilustrado na Fig. 2.9a. Na segunda categoria estão os MNLME (Modelos Não Lineares de Mistura Espectral), ilustrado na Fig. 2.9b, onde a interação entre os MR se dá na forma não linear, devido às múltiplas reflexões associadas aos efeitos da dispersão múltipla da luz pelos diferentes materiais (GARCÍA-HARO; SOMMER, 2002; ZHANG et al., 1998).

O modelo mais utilizado é o MLME por proporcionar resultados adequados em várias aplicações (COLLINS; ROBERTS; BOREL, 2001; ROBERTS et al., 1998), como também pela sua simplicidade (BATESON; ASNER; WESSMAN, 2000). Por outro lado, o MNLME tem sido usado em certas aplicações específicas principalmente nos estudos das propriedades de coberturas vegetais (ZARCO-TEJADA, et al., 2001).

Ambos modelos constituem uma simplificação da complexa realidade conforme a análise realizada por Schowengerdt (1997). Uma dessas complexidades é o caso em que há transmissão da radiação através de um material (como nas folhas), seguida pela reflexão em outro material; outra situação complexa está relacionada às múltiplas reflexões que podem ocorrer dentro de um único objeto, ou entre vários objetos.

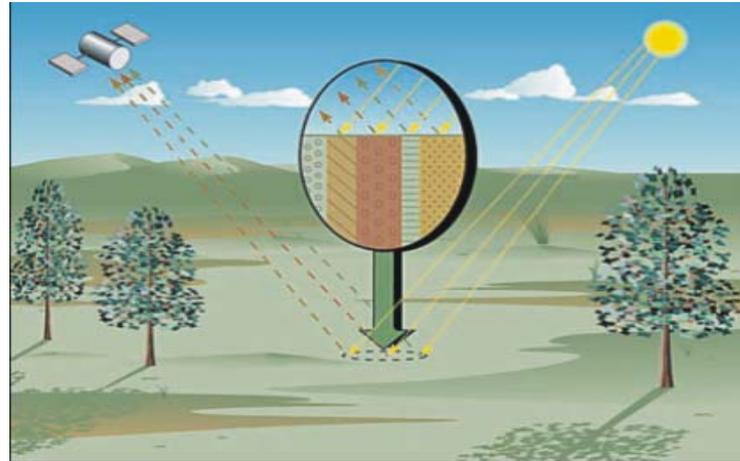
Não há uma resposta clara na literatura científica de qual tipo mistura (linear ou não linear) representa melhor os espectros de maior interesse, pois a escolha adequada depende de uma série de fatores e condições da cena analisada (KESHAVA; MUSTARD, 2002). De uma maneira geral o MLME é o mais utilizado na maioria das aplicações (BATESON; ASNER; WESSMAN, 2000), já que há grande complexidade associada aos modelos não lineares.

O conceito do MLME parte da premissa de que a radiação resultante de uma cena é formada pela mistura de contribuições dos MR, misturados linearmente em proporções variáveis dentro de cada *pixel* da cena. Para descrever o MLME seja:

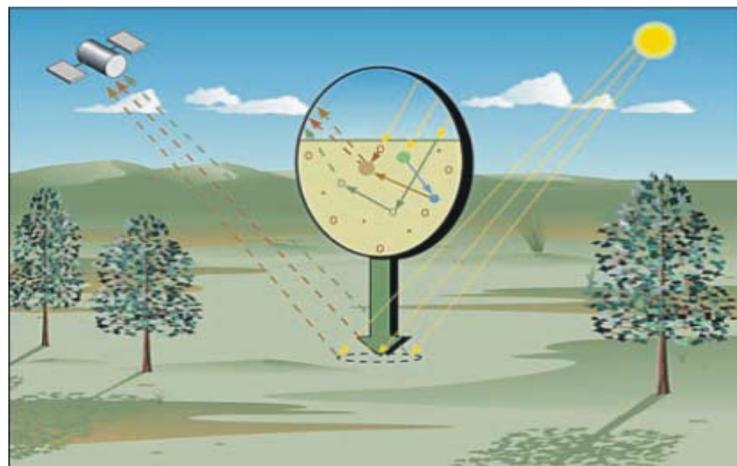
$$\vec{I}(i, j) = [I(i, j, 1)I(i, j, 2)I(i, j, 3)\dots I(i, j, N)]^T \quad (2.2)$$

o vetor espectro referente ao *pixel* (i, j) , $i = 1, 2, \dots, I$ e $j = 1, 2, \dots, J$, sendo I e J as dimensões da imagem e N é o número de bandas em cada *pixel*. Por razões de simplicidade, quando não houver problemas de confusão de notação, o vetor espectro $\vec{I}(i, j)$ também será representado como:

$$\vec{I}(i, j) = \vec{I} = [I(1), I(2), \dots, I(N)]^T \quad (2.3)$$



(a)



(b)

FIGURA 2.9 – Tipos de mistura: (a) modelo linear associado a reflexões simples e (b) modelo não linear associado a múltiplas reflexões. Fonte: adaptação a partir de Keshava e Mustard (2002, p. 45).

O vetor espectro $\vec{I}(i, j)$ pode ser modelado pela combinação linear dos vetores espectros MR, \vec{e}_r , $r = 1, \dots, R$ pela expressão:

$$\vec{I}(i, j) = \sum_{r=1}^R \phi_r \vec{e}_r + \vec{\epsilon} \quad (2.4)$$

onde R é o número total de MR associados à cena, ϕ_r é um escalar que representa a fração abundância do MR \vec{e}_r , no *pixel* (i, j) e \vec{E} é o vetor erro que deve ser minimizado, segundo o critério do mínimo erro quadrático médio, pela escolha adequada dos valores de ϕ_r .

A equação (2.4) pode ser escrita na forma matricial como:

$$\vec{I} = E\vec{\Phi} + \vec{w} \quad (2.5)$$

onde \vec{I} é um vetor coluna de dimensão N , \vec{w} é um vetor erro, coluna, de dimensão N , E é uma matriz de dimensão $N \times R$ formada pelas componentes dos vetores MR, \vec{e}_r , $r = 1, \dots, R$:

$$E = \left[\begin{array}{c} \left[\begin{array}{c} \vec{e}_1 \\ \vec{e}_2 \\ \dots \\ \vec{e}_R \end{array} \right] \end{array} \right] \quad (2.6)$$

e $\vec{\Phi}$ é um vetor coluna, de dimensões R , formado pelas frações abundâncias:

$$\vec{\Phi} = [\phi_1 \quad \phi_2 \quad \dots \quad \phi_R]^T \quad (2.7)$$

A fração abundância representa a contribuição de cada MR nos *pixels* da imagem. O vetor fração abundância $\vec{\Phi}$, da equação (2.7), relativo ao *pixel* (i, j) estimado para todos os *pixels* da imagem, $i = 1, 2, \dots, I$ e $j = 1, 2, \dots, J$, gera um conjunto de mapas de abundâncias ou imagens-fração que estão associados a contribuição de cada um dos MR.

Há vários trabalhos (TOMPKINS et al., 1997; ELMORE et al., 2000; LI; USTIN; LAY, 2005) analisando a decomposição do espectro de um *pixel* misturado em frações abundância dos espectros puros ou MR. As frações abundância são um indicativo do nível de contribuição associada a cada um dos MR (BOARDMAN; KRUSE; GREEN, 1995). Para encontrar o conjunto dos valores das proporções ou frações abundâncias é necessário resolver um sistema de equações onde o número de bandas da imagem deve ser maior que o número de MR, suposição integralmente satisfeita no caso de imagens hiperespectrais (KESHAVA; MUSTARD, 2002).

Os MR podem ser identificados de forma gráfica através do diagrama de dispersão entre duas bandas com pouca correlação, como é o caso das duas primeiras imagens obtidas

da transformação por componentes principais (TCP) (BATESON; CURTISS, 1996). A Fig. 2.10 ilustra essa metodologia de escolha dos MR, onde as coordenadas dos MR são representadas por e_1 , e_2 e e_3 . Observa-se nessa figura que todos os pontos da imagem estão no interior do triângulo formado pelos três pontos mais extremos, escolhidos como MR: $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$. Os vetores relativos aos pontos no interior do triângulo têm um novo sistema de coordenadas com origem no centróide da nuvem dos pontos. Assim qualquer que seja o ponto pertencente à imagem pode ser expresso pela combinação linear dos pontos mais extremos da nuvem os quais são escolhidos como os melhores candidatos a MR (BOARDMAN; KRUSE; GREEN, 1995). A situação mostrada na Fig. 2.10, em que os pontos internos do triângulo podem ser expressos pelos MR é denominada por Boardman (1993) de geometria convexa.

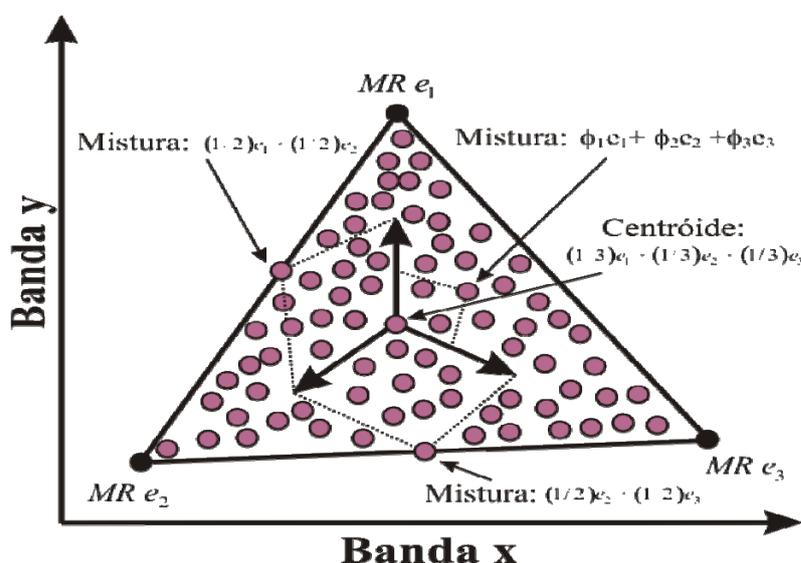


FIGURA 2.10 – Diagrama de espalhamento entre duas dimensões dos dados espectrais ilustrando a interpretação do modelo de mistura baseada nos MR.

Fonte: adaptação a partir de Keshava e Mustard (2002, p. 45).

Segundo Keshava e Mustard (2002) a aproximação dada pela geometria convexa para explicar o sentido físico dos MR tem algumas limitações:

- a) a variabilidade espectral do mesmo tipo de alvo na cena, diversos *pixels* adjacentes pertencentes à mesma classe, é explicada pelas condições de iluminação variável, efeitos de sombra e da geometria de aquisição (incidência e reflexão da luz). Essa variabilidade espectral para o mesmo material pode levar a escolha de MR pertencentes à mesma classe. Em particular a variável angular relacionada com a geometria de aquisição é muito importante nos estudos da variabilidade espectral da vegetação (BERGER et al., 2000), e

b) a presença de pontos fora do triângulo (*outliers*) (HSIEH; LANDGREBE, 1998), mostrada na Fig. 2.11, devido à variabilidade dos espectros associados aos materiais, podem acarretar coeficientes negativos na estimação das frações abundâncias do modelo de mistura. TOMPKINS et al., (1997) propuseram a inclusão de MR falsos para solucionar este problema. Pela configuração dos pontos mostrados na Fig. 2.11 é necessário gerar MR artificiais para que o polígono convexo contenha todos os pontos da imagem. Outra forma de contornar esta situação seria incorporar as frações abundâncias negativas no MLME, mas isto não seria conceitualmente interessante.

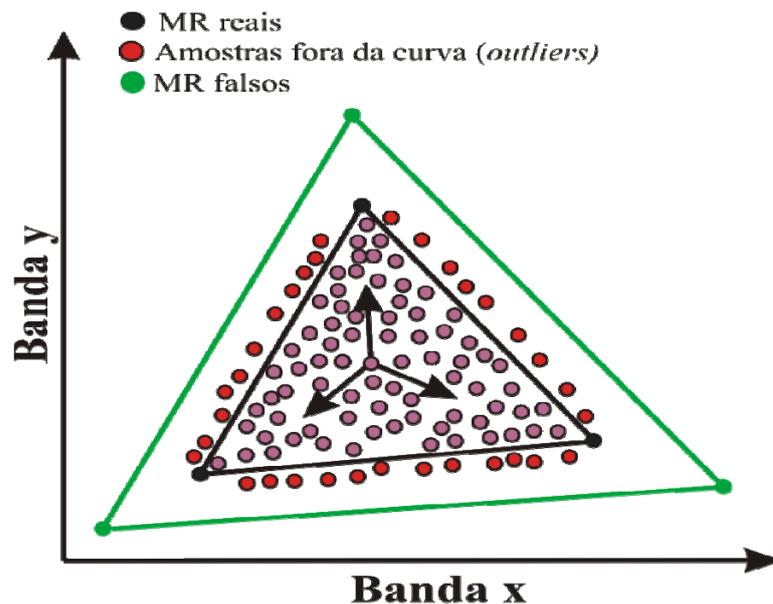


FIGURA 2.11 – Diagrama de espalhamento bidimensional contendo MR falsos para eliminar o aparecimento dos pontos fora da curva (*outliers*).
 Fonte: adaptação a partir de Tompkins et al. (1997, p. 475).

A extensão da determinação dos MR utilizando-se mais de duas bandas, descorrelacionadas, é realizada por Ifarraguerri e Chang (1999) através do conceito de hiperpolígono convexo.

Staez et al. (1998) descartam a utilização de MR falsos por não serem encontrados em nenhuma classe de alvo pertencente ao mundo real. Por outro lado, a sua utilização é defendida por Tompkins et al., (1997), já que explica e torna coerente em termos físico as frações abundâncias finais.

Plaza et al. (2004a) afirmam que os inconvenientes descritos anteriormente decorrente dos pontos fora da curva (*outliers*) podem ser minimizados com a incorporação da informação espacial no processo de seleção dos MR.

A estimação das frações abundâncias ϕ_r , da equação (2.4), relativas ao MLME aparece na literatura científica como técnicas de separação linear das contribuições espectrais dos MR (*Linear Spectral Unmixing*). Para isso duas restrições são desejáveis para a estimativa das frações abundâncias:

a) coeficientes não negativos e menores ou iguais a unidade: $0 \leq \phi_r \leq 1$:

A abundância de um determinado material no *pixel* deve estar sempre entre 0 (abundância nula) e 1 (abundância total). Coeficientes negativos podem ser originários de uma inadequada seleção dos MR ou de um número impróprio de MR, entre outros fatores. Por outro lado, pode ser conveniente aceitar pequenos valores negativos quando o algoritmo de estimação das frações exibe bons resultados globais (GROSS; SCHOTT, 1998). Outros autores (KESHAVA; MUSTARD, 2002) descrevem que valores negativos ou maiores do que 1 não são necessariamente indicativo de erro de aplicação do método de determinação dos MR;

b) a soma unitária de todas as frações abundâncias: $\sum_{r=1}^R \phi_r = 1$:

Esta restrição pode introduzir algumas inconsistências numéricas durante a execução do processo de estimação, principalmente quando faltam ou sobram MR. Sem a restrição a somatória das frações abundâncias no *pixel* pode ser diferente da unidade.

As técnicas de estimação das frações abundâncias dos MR que incorporam as restrições anteriormente descritas denominam-se técnicas totalmente restritivas de separação linear das contribuições espectrais, FCLSU (*Fully Constrained Linear Spectral Unmixing*) dos MR.

Heinz e Chang (2001) discutem métodos de estimativa das frações abundâncias e concluem que a segunda restrição é mais dificilmente realizável e só pode ser alcançada de modo aproximado por métodos numéricos. Devido a estas dificuldades utilizam-se ocasionalmente, para aplicações particulares, métodos não restritivos que produzem soluções que não refletem necessariamente os verdadeiros valores das proporções ou frações

abundâncias dos materiais. Neste caso eles podem ser usados unicamente com o propósito de detecção de materiais, discriminação e classificação, mas não para quantificação dos materiais.

Segundo Plaza et al. (2004b) o método FCLSU melhora a solução numérica relativa ao processo de inversão de matrizes no processo de estimativa das frações abundâncias.

Existem outros métodos para a estimativa das frações abundâncias, entre eles incluem-se: estimação dos mínimos quadrados, LSE (*Least Squares Estimation*) (SMITH et al., 1990); decomposição de valores primários, SVD (*Singular Value Decomposition*) (BOARDMAN, 1989); ortogonalização Gram-Schmidt, G-SO (*Gram-Schmidt Orthogonalization*) (ADAMS; SMITH; GILLESPIE, 1993); mínimos quadrados restritivo, CLS (*Constrained Least Squares*) e mínimos quadrados ponderado, WLS (*Weighted Least Squares*) (SHIMABUKURO; SMITH, 1991). O aplicativo ENVI (2006) (*Environment for Visualizing Images*) muito utilizado nas análises deste estudo utiliza o SVD.

3 Sensor AVIRIS e Área Teste

Neste capítulo descrevem-se as características do sensor *Airborne Visible InfraRed Imaging Spectrometer* (AVIRIS) e da área teste cujas imagens AVIRIS foram utilizadas no desenvolvimento e teste da metodologia proposta neste trabalho.

3.1 Sensor AVIRIS

Imagens do espectrorradiômetro AVIRIS, que se tornou operacional em 1989, foram utilizadas nas avaliações da metodologia proposta neste trabalho. Esse sensor hiperespectral é o resultado de um projeto proposto à NASA em 1983 por pesquisadores do JPL. A sua concepção, na época, visou preparar a comunidade científica para a utilização de imagens geradas por sensores de alta resolução espectral, bem como incorporá-las aos futuros programas orbitais de sensoriamento remoto. O AVIRIS pode ser utilizado para as mais variadas aplicações, incluindo a possibilidade de simulação de bandas para definição das faixas espectrais de operação de sensores de programas espaciais em desenvolvimento.

O sensor AVIRIS é composto, essencialmente, por um grupo de quatro espectrômetros que observam a superfície da terra através de um sistema de varredura mecânica, enquanto se desloca sobre o terreno a bordo de uma aeronave ER-2. A cada momento, os espectrômetros recebem o fluxo de radiação de um ponto no solo com um *Instantaneous Field of View* (IFOV) de 1 mrad, a uma altitude nominal de 20 km, equivalente a uma área nominal de 20 x 20 m², a qual define a resolução do sensor na superfície. Este mesmo *pixel*, correspondente a área de 20 x 20 m² na superfície, é detectado simultaneamente em 224 bandas espectrais, com largura de banda de aproximadamente 10 nm. A imagem é construída a partir do movimento angular do sistema de varredura, o qual define uma linha de imageamento de 614 *pixels* de largura (aproximadamente 11km) perpendicularmente à direção de vôo e pelo movimento da aeronave que define a extensão da imagem (variando de 10 a 100 km), considerando a altura de vôo padrão de 20 km e o campo de visada de 30 graus. Os dados coletados são armazenados em uma unidade de fita magnética para posterior análise. O conjunto de dados registrados forma um cubo, no qual dois eixos representam as dimensões espaciais e o terceiro eixo representa a dimensão espectral.

Os trabalhos de Elvidge e Mouat (1989) e Elvidge (1990) mostraram que é possível detectar na superfície coberturas de vegetações menores do que 10%, do total dos materiais

presentes em um *pixel*, utilizando a banda do vermelho de dados hiperespectrais adquiridos pelo sensor AVIRIS. Elvidge (1988b, 1990) destacou que, através dos dados AVIRIS, é possível detectar a presença das feições de absorção provocadas por outros constituintes químicos da vegetação não-fotossinteticamente ativa, tais como a lignina e celulose, que estão associadas ao dossel ou à palha vegetal (em inglês *litter*) presente sobre o solo

Ao longo destes últimos anos diversas melhorias têm sido incorporadas aos subsistemas do sensor AVIRIS (VANE et al., 1993) proporcionando o aumento da relação sinal-ruído, $SNR(\lambda)$, de cada banda espectral. A Fig. 3.1 mostra como evoluiu essa relação sinal-ruído ao longo dos anos.

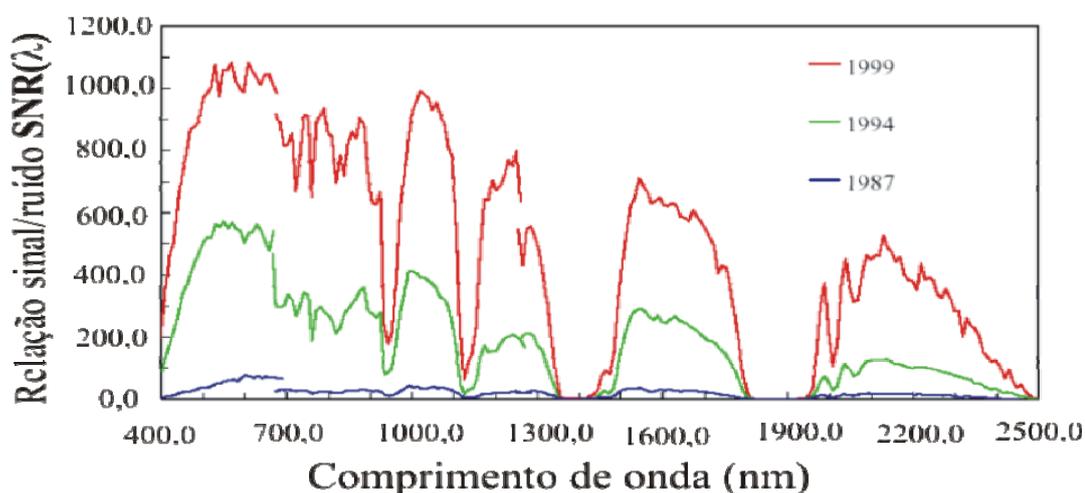


FIGURA 3.1 – Relação sinal-ruído do sensor AVIRIS para os anos de 1987, 1994 e 1997.
Fonte: Adaptada de Green e Pavri (2000).

3.2 Área teste

Para o desenvolvimento, ilustração de conceitos e teste do processo de seleção de MR, proposto neste trabalho, utilizou-se imagens do sensor AVIRIS de uma área teste situa-se próximo à cidade de Campo Grande, MS, e nas coordenadas 19°55'S e 53°30'W. Essa área é dominada pela presença de NEOSSOLO QUARTZARÊNICO (*RQ*), LATOSSOLO VERMELHO (*LV*) e NITOSSOLOS VERMELHOS (*NV*), que representa uma transição química de baixo para altos teores de ferro, óxidos de titânio e fração argila (NOVAES et al., 1983). Também é caracterizada por atividades agrícolas como pasto e alguns talhões de colheitas sazonais (por exemplo, milho e soja). A cobertura de vegetação natural está composta de pequenos bosques de cerrado. A chuva anual e a temperatura média são da

ordem de 1.300 mm e 24°C, respectivamente, e a geologia regional é dominada pela presença de diferentes tipos de arenitos e basaltos (PIZARRO, 1999).

Os dados AVIRIS foram adquiridos em 20 de agosto de 1995, praticamente no ápice da estação regional seca, em 224 faixas estreitas (≈ 10 nm de largura), posicionadas na faixa espectral óptica de aproximadamente $383 \leq \lambda \leq 2.510$, λ em nm, e com resolução espacial nominal de 20 x 20 m. O voo do sensor AVIRIS pela área selecionada produziu 7 subcenas adjacentes, sendo que somente as subcenas 4, 5 e 7 foram selecionadas para o desenvolvimento desse trabalho, não só por possuírem melhores áreas testes, mas também por terem sido objeto de um trabalho de campo (PIZARRO, 1999). As subcenas selecionadas são formadas por uma área retangular de 12 por 10 km (614 x 512 *pixels*), aproximadamente. Os dados, originalmente fornecidos em valores de radiância, foram convertidos para valores de reflectância de superfície, por um método de transferência radiativa baseado no MODTRAN (GREEN, 1991). Os detalhes sobre a correção atmosférica aplicada nas subcenas podem ser encontrados em Pizarro (1999).

O conjunto de imagens hiperespectrais, com os valores de reflectância, é representado por $I_o(i, j, \lambda)$, onde (i, j) denota a localização do *pixel* na imagem e λ comprimento de onda da banda espectral. Do conjunto de imagens testes foram excluídas as bandas iniciais de 1 a 8 ($383 \leq \lambda \leq 451$), por ausência de dados e por apresentarem valores negativos, e as bandas finais de 218 a 224 ($2.450 \leq \lambda \leq 2.510$) por estarem extremamente ruidosas. Suprimiu-se ainda as bandas entre 33 a 35 ($663 \leq \lambda \leq 687$) devido a erro de calibração do sensor AVIRIS, as bandas 78 a 85 ($1.095 \leq \lambda \leq 1.162$), 95 a 123 ($1.258 \leq \lambda \leq 1.512$) e 145 a 175 ($1.731 \leq \lambda \leq 2.021$), pois o MODTRAN introduziu correções excessivas, transformando as bandas de absorção causadas pelas respectivas influências do vapor d'água e do CO₂ em picos de reflectância. Para as bandas suprimidas de 33 a 35, 78 a 85, 95 a 123 e 145 a 175 foram estimados novos valores de reflectância utilizando-se a função *spline* cúbica e considerando-se que os valores discretos dos espectros estavam igualmente espaçados e correspondiam todos a uma mesma largura de banda.

O novo conjunto de imagens foi representado por $I(i, j, n)$, onde n ($1 \leq n \leq N$) denota o número da banda, ou cada uma das imagens. O índice n foi utilizado em substituição a λ por simplicidade de notação. Deste modo, os valores correspondentes ao comprimento de onda foram substituídos pelos respectivos valores do número da banda. A Fig. 3.2 ilustra duas

curvas espectrais para um mesmo alvo com e sem a interpolação. No caso particular todos os conjuntos de imagens hiperespectrais utilizados no trabalho tem $N = 209$.

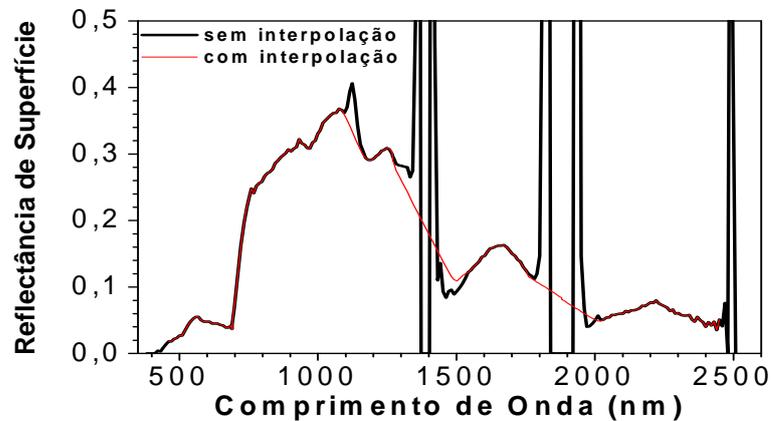


FIGURA 3.2 - Espectro da vegetação verde extraído de um *pixel* da imagem original, corrigida para os efeitos de absorção e espalhamento atmosféricos, com 224 bandas originais sem interpolação e 209 bandas com interpolação.

A Fig. 3.3 mostra uma imagem colorida da área de estudo obtida com as bandas 22 (≈ 667 nm), 10 (≈ 549 nm) e 3 (≈ 481 nm) do AVIRIS em vermelho, verde e azul, respectivamente, a partir da imagem reflectância de superfície da subcena 4. Em geral, as áreas de vegetação verde (*vv*) (p. ex., culturas desenvolvidas) aparecem em tonalidades verdes nesta composição colorida. A vegetação com pouca atividade fotossintética (*vs*) ocorre em tonalidades amareladas, enquanto áreas com predomínio de solos são expressas por cores mais avermelhadas. Obviamente, a diferença temporal entre a aquisição dos dados AVIRIS (Agosto de 1995) e as atividades de campo (Agosto de 1997) limitou uma análise mais profunda do comportamento espectral dos alvos na cena.

Para se ter uma melhor noção das classes e subclasses constituintes da subcena 4 apresentam-se na Fig. 3.3 as posições de amostras candidatas a MR, escolhidas por um especialista, referentes a essas classes e subclasses. As classes e subclasses referentes a esta cena são:

- classe vegetação verde (*vv*) composta por subclasses que representam diferentes estágios evolutivos e/ou diferentes espécies: *vv1*, *vv2*, *vv3*, *vv4* e *vv5*;
- classe vegetação com pouca atividade fotossintética ou vegetação seca (*vs*);
- classe Latossolo Vermelho (*LV*) composta pelas subclasses: *LV1*; *LV2* e *LV3*;

- classe Neossolo Quartzarênico (*RQ*) composta pelas subclasses: *RQ1*, *RQ2* e *RQ3*;
- classe água (*a*) formada por corpos d'água presentes ao longo da mata galeria.
Adicionalmente, escolheram-se duas classes especiais:
- classe mistura composta pelas subclasses: *m1*, *m2*, *m3* e *m4*. Todas as subclasses são provenientes de locais de intensa mistura por estarem situadas em locais heterogêneos e
- classe anômala composta das subclasses: *anoma1*, *anoma2*, *anoma3* e *anoma4*. Todas as amostras pertencentes à classe anômala são de locais na qual pelo menos um *pixel* na vizinhança 5×5 do *pixel* considerado, tem falha na aquisição em algumas bandas do espectro imageado.

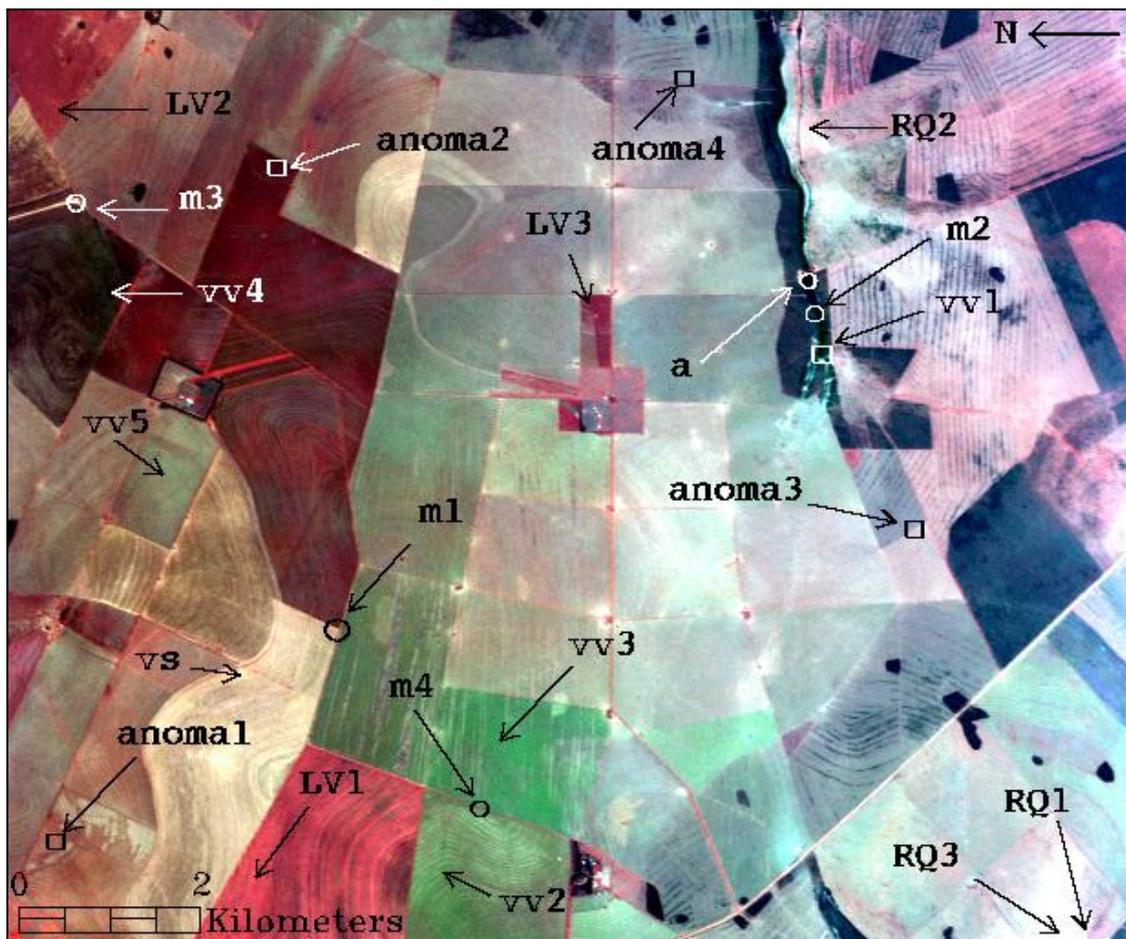


FIGURA 3.3 – Composição colorida RGB para a subcena 4 obtida com as bandas 22 (≈ 667 nm), 10 (≈ 549 nm) e 3 (≈ 481 nm), respectivamente, do sensor AVIRIS.

As amostras pertencentes às classes especiais, mistura e anômala, não são boas amostras candidatas a MR por conterem altas mistura de classes, ou falhas de aquisição, entretanto elas foram escolhidas para se avaliar, posteriormente, o desempenho do processo de escolha dos MR.

A Figura 3.4 mostra os espectros médios, calculados em uma vizinhança 5×5 dos *pixels* representativos das diversas classes atribuídas a subcena 4. Na Fig. 3.4d mostra-se os 25 espectros, na vizinhança 5×5 da amostra *m1*, referente a uma classe com misturas. A variabilidade espectral evidencia a mistura. Na Fig. 3.4e mostra-se também os 25 espectros de uma vizinhança homogênea, pois os espectros têm pouca variabilidade, mas fica evidente a falha de aquisição espectral em algumas bandas presente em um dos 25 espectros.

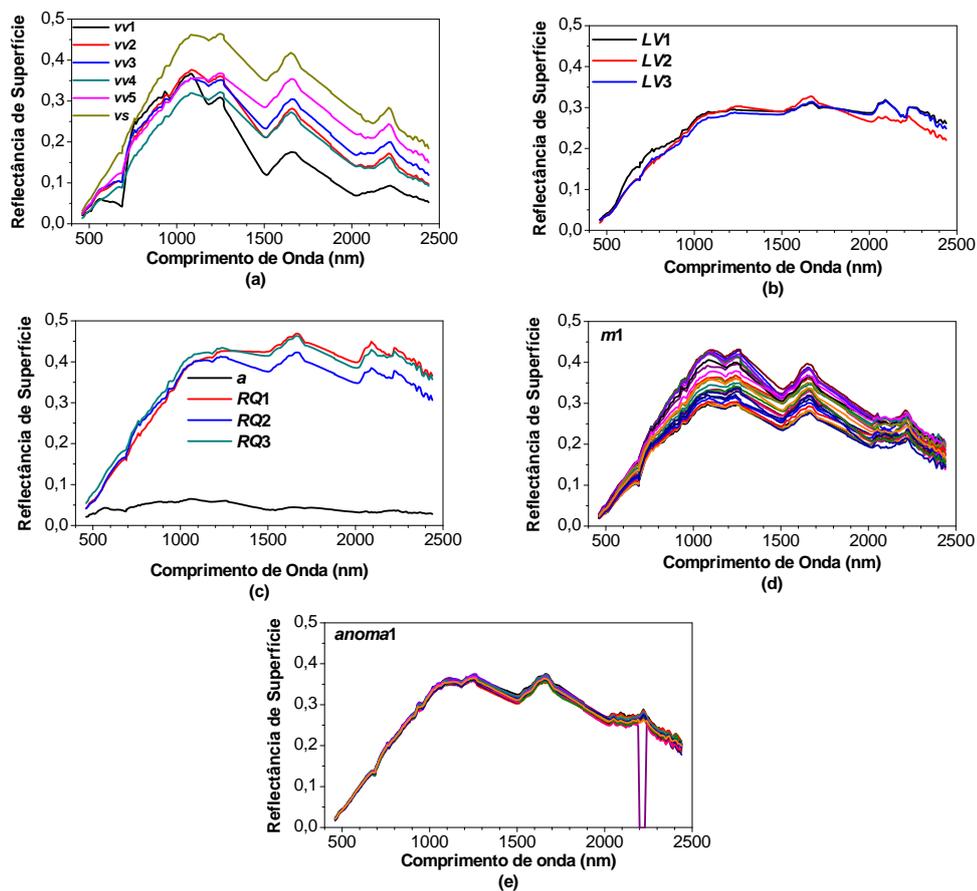


FIGURA 3.4 - Espectros médios representativos das classes e subclasses pertencentes a subcena 4.

Adicionalmente a subcena 4 também foram utilizadas as subcenas 5 e 7 da área teste. Estas subcenas possuem as mesmas características básicas, relativas aos tipos de solos e

vegetação, encontradas na subcena 4. A Fig. 3.5 mostra uma imagem da subcena 5, onde são evidenciadas as posições de amostras candidatas a MR, escolhidas por um especialista, pertencentes às classes e/ou subclasses dessa cena. Nesta subcena as classes e subclasses são:

- classe vegetação verde (*vv*) é composta por subclasses em diferentes estágios evolutivos e/ou diferentes espécies: *vv1*, *vv2*, *vv3*, *vv4* e *vv5*;
- classe vegetação com pouca atividade fotossintética ou seca (*vs*);
- classe Latossolo Vermelho (*LV*) composta pelas subclasses: *LV1*; *LV2*; *LV3*; *LV4* e *LV5*;
- classe Neossolo Quartzarênico (*RQ*) é composta pelas subclasses: *RQ1* e *RQ2*;
- classe água (*a*).

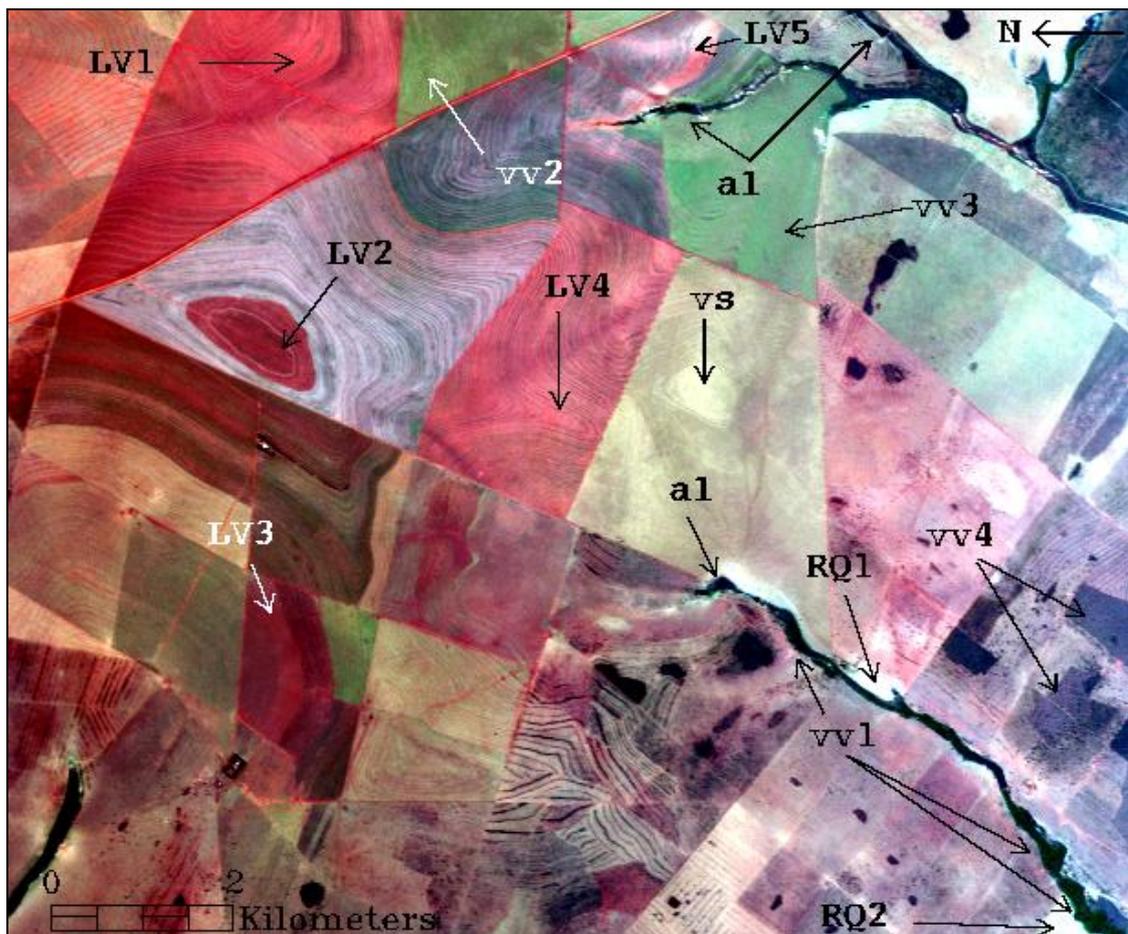


FIGURA 3.5 – Composição colorida RGB para a subcena 5 obtida com as bandas 22 (≈ 667 nm), 10 (≈ 549 nm) e 3 (≈ 481 nm), respectivamente, do sensor AVIRIS.

A Fig. 3.6 mostra os espectros médios, calculados em uma vizinhança 5×5 dos *pixels* representativos das diversas classes atribuídas a subcena 5.

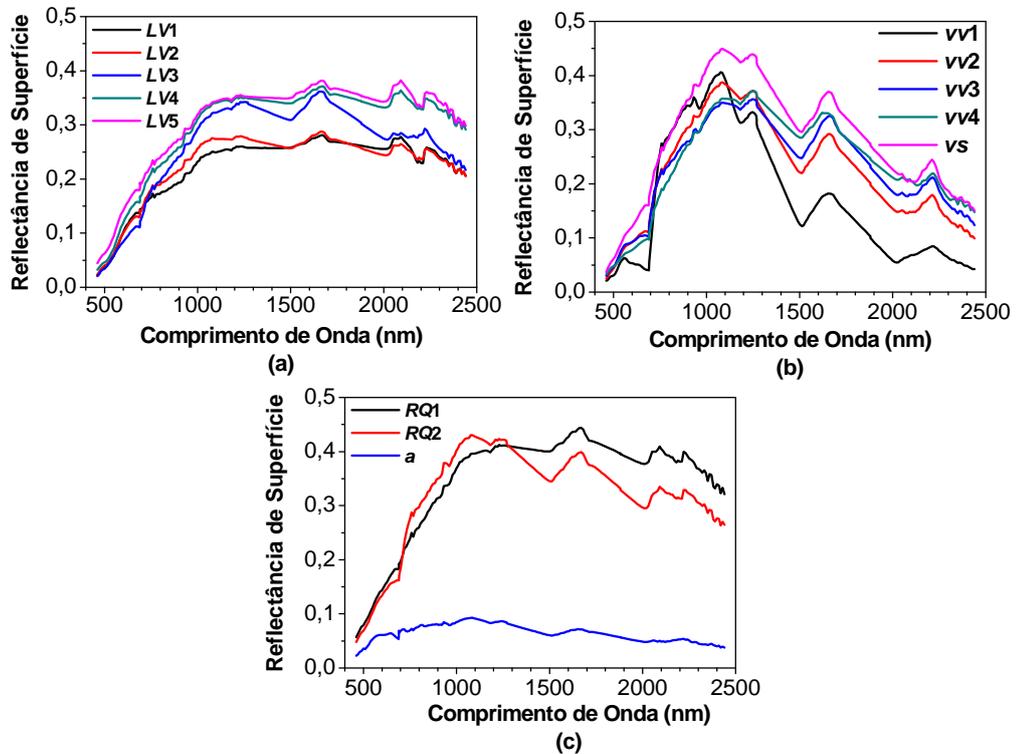


FIGURA 3.6 - Espectros médios das amostras das classes e subclasses pertencentes a subcena 5.

Na Fig. 3.7 mostras-se a subcena 7 e as posições de amostras candidatas a MR, escolhidas por um especialista, pertencentes às classes e/ou subclasses dessa cena. Nesta subcena as classes e subclasses são:

- classe vegetação verde (*vv*);
- classe vegetação com pouca atividade fotossintética (*vs*);
- classe Latossolo Vermelho (*LV*);
- classe Nitossolo Vermelho (*NV*);
- classe Neossolo Quartzarênico (*RQ*) e

- classe água (*a*).

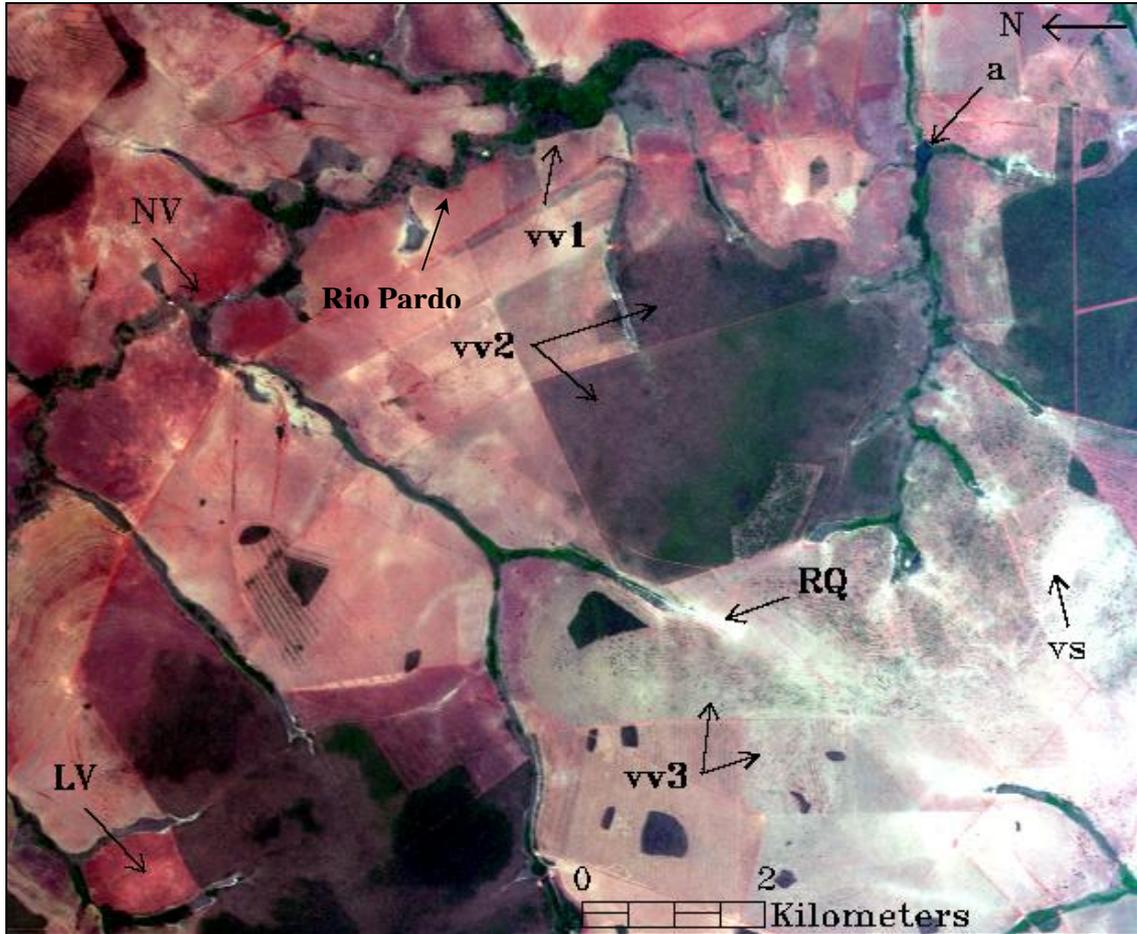


FIGURA 3.7 – Composição colorida RGB para a subcena 7 obtida com as bandas 22 (≈ 667 nm), 10 (≈ 549 nm) e 3 (≈ 481 nm), respectivamente, do sensor AVIRIS.

A Fig. 3.8 mostra os espectros médios, calculados em uma vizinhança 5×5 dos *pixels* representativos das diversas classes atribuídas a subcena 7.

Os espectros médios das amostras da classe vegetação verde (*vv*) e vegetação seca (*vs*), presentes nas Figs. 3.4a, 3.6b e 3.8a, de uma maneira geral ilustram a variabilidade espectral associada a maior ou menor quantidade de umidade presente nas folhas. Nestas figuras *vv1* é a mais vigorosa ou sadia (mais úmida por ser proveniente de uma mata galeria) e *vs* é uma vegetação não-fotossinteticamente ativa ou seca, pois a vegetação verde entrou em processo de senescência.

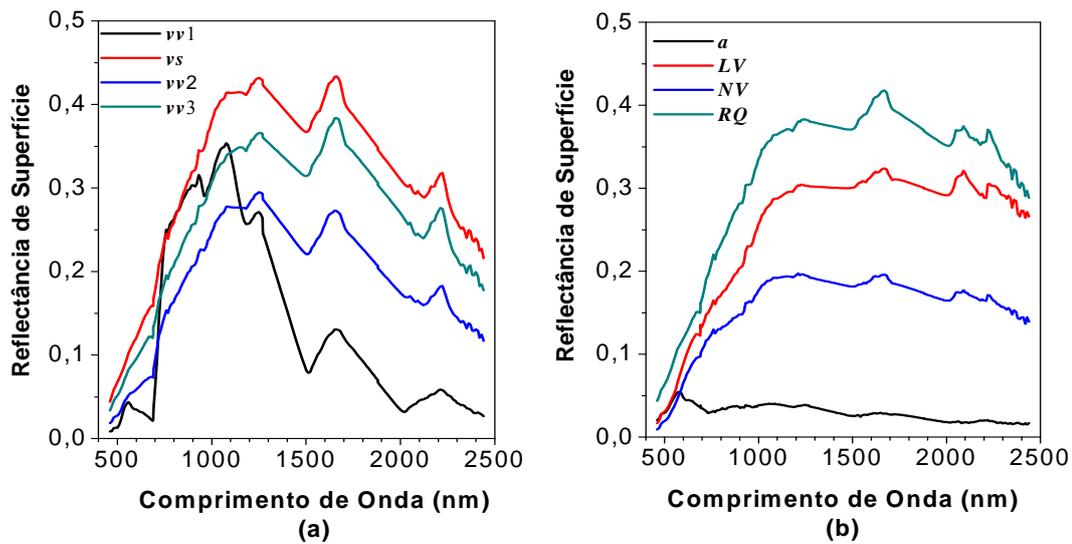


FIGURA 3.8 – Espectros médios das amostras pertencentes às classes e subclasses indicadas na subcena 7.

A variabilidade espectral presente nas subclasses de vegetação verde (*vv1*, *vv2*, *vv3*, etc) pode também, além da umidade, estar associada a uma mesma espécie em diferentes estágios evolutivos da vegetação.

No caso da *vs* com a desidratação e a perda de pigmentos, a sua curva espectral é gradativamente modificada. A forte absorção no VIS, ≈ 711 nm, é reduzida, com a perda de pigmentos, passando a apresentar maior reflectância. Nessa região a curva será mais suave e crescente, assemelhando-se ao perfil de resposta do solo. Na região do IVM, entre ≈ 1.860 a ≈ 2.440 nm, começam a surgir as feições de absorção provocadas pela lignina e celulose.

Quanto aos espectros médios das amostras de solos *NV*, *LV* e *RQ* ilustrados nas Figs. 3.8b refletem as variações de reflectância dos solos, onde o solo *NV* apresenta menores valores de reflectância, ao longo de toda a faixa do espectro óptico imageado em relação aos outros dois solos. O solo *NV* por sua vez está presente somente na subcena 7 ao longo do Rio Pardo (GALVÃO; PIZARRO; EPIPHANIO, 2001) e as suas típicas feições de absorção são provenientes dos materiais listados na Tabela 2.1, principalmente pela hematita (PIZARRO, GALVÃO; EPIPHANIO, 2001).

Assim, *NV* e *LV* possuem menor reflectância em relação ao *RQ* por apresentarem uma maior quantidade de substâncias opacas ou minerais fortemente absorvedores de radiação (p. ex., matéria orgânica e magnetita) na sua composição.

Solos mais claros (p. ex., *RQ*), normalmente mais pobres em substâncias opacas, tendem a serem mais sensíveis aos efeitos espectrais relativos à matéria orgânica e argila, que podem provocar mudanças significativas na forma dos espectros, conforme destacaram Galvão, Vitorello e Formaggio (1997). No caso particular a amostra que melhor representa esta classe de solo é a *RQ1* onde o seu espectro médio é ilustrado na Fig. 3.4c pertencente a subcena 4, devido às típicas feições de absorção (Tabela 2.2) serem mais bem reconstituídas. Quanto ao solo com reflectância intermediária, *LV*, os espectros médios das amostras *LV1* e *LV3* (Fig. 3.5b) da subcena 4 e *LV5* (Fig. 3.6a) da subcena 5 por também terem melhor reconstituição das típicas feições de absorção (Tabela 2.2), principalmente pelo mineral caulinita (PIZARRO, GALVÃO; EPIPHANIO, 2001).

As reconstituições das bandas de absorção descritas acima dependerá da mistura na vizinhança ao *pixel* analisado resultantes da presença de *vv* ou *vs*, conforme discutido por Galvão, Vitorello e Almeida Filho (1999) e da presença de minerais opacos (ferro total e TiO_2) (PIZARRO; EPIPHANIO; GALVAO, 2001).

No caso dos espectros mostrados na Fig. 3.4a, 3.6b e 3.8a há uma pobre definição na reconstituição das bandas de absorção devido à lignina e celulose o que já foi descrito em Galvão, Pizarro e Epiphanio (2001) quando comparados com as ilustradas nas Figs. 2.2 e 2.3. De acordo com Elvidge (1988b, 1990), feições de lignina e celulose são mais bem reconstituídas quando a vegetação contém pouca água líquida no seu interior.

4 Amostras Candidatas a MR

Neste capítulo apresenta-se a primeira fase da metodologia desenvolvida para a seleção dos MR em imagens hiperespectrais. Nesta fase é descrito o método de triagem das amostras da imagem hiperespectral para se compor um conjunto de amostras candidatas a MR. Enfatiza-se a **triagem discriminatória** das amostras espectrais, que compõe a imagem hiperespectral, baseada nos princípios de **uniformidade espacial** e **homogeneidade espectral** da vizinhança da amostra de interesse, no princípio do **contexto espacial** e no princípio da **não-redundância** entre as amostras. A descrição destes princípios é do escopo deste capítulo.

4.1 Triagem das amostras associadas às classes de alvos

A triagem das amostras pertencentes à imagem hiperespectral que irão compor o conjunto de amostras candidatas a MR deve ser feita de modo que todas as classes e subclasses presentes na imagem tenham pelo menos uma representante. Plaza e Chang (2006) alertam sobre o impacto da escolha destas amostras candidatas na seleção final dos MR.

Grande parte da pesquisa de métodos para a classificação de imagens tem seu foco nos algoritmos matemáticos para agrupar *pixels*, com os mesmos tipos de feições, que não levam em consideração a informação semântica entre as classes. Uma cena imageada está diretamente relacionada com conceitos humanos e não só àqueles relacionados ao sinal físico detectado e medido pelos sensores imageadores, esse fato pode tornar subjetiva a definição das classes presentes na imagem (WILKSON, 2005).

Essa subjetividade leva a uma dificuldade na escolha das amostras de treinamento ou representativas de cada classe de interesse. Muitas classes de alvos, presentes na cobertura da cena analisada, não têm um único espectro de reflectância que as caracterizem. Por exemplo, o espectro de reflectância da vegetação muda continuamente com o seu desenvolvimento devido às estações do ano e há o caso em que pode ocorrer variação no desempenho evolutivo de uma espécie de vegetação em diferentes áreas para uma mesma cena e com isso provocar diferentes escolhas para os MR (HOFFBECK; LANDGREBE, 1996; GLENN et al., 2005). Esta última situação pode ser observada na Fig. 2.2 onde se apresentam diferentes estágios evolutivos devido à variação de clorofila e/ou água para o milho. Outra situação é a mistura que pode ocorrer em qualquer proporção entre o solo e a vegetação. Desta forma, em teoria, pode ocorrer um incontável número de classes (SMITH, et al. 1990).

Os dados hiperespectrais têm ótimo potencial para a discriminação de classes de alvos. Em contrapartida os algoritmos de análises podem ser limitados na precisão do seu desempenho em se tratando da modelagem do conjunto de dados e das classes desejadas, pois os dados hiperespectrais não são simples de serem trabalhados, há a necessidade de se conhecer muito bem o espaço de feições, que não é intuitivo, deste tipo de dado que tem uma alta dimensionalidade devido o grande número de bandas (LANDGREBE, 2005).

Nesta primeira fase, relativa a escolha das amostras representativas das classes e subclasses da imagem hiperespectral, que irão compor um conjunto de amostras candidatas a MR, tem-se três abordagens: a manual, a aleatória e a discriminatória, que serão tratadas nos itens a seguir.

4.1.1 Triagem manual com especialista

Esta triagem é realizada por um especialista que escolhe empiricamente, baseado na sua experiência e habilidade, quais devem ser as amostras da imagem hiperespectral que potencialmente podem se tornar MR.

A abordagem manual tem a vantagem de poder contar com a habilidade do especialista em escolher amostras bem representativas das diversas classes e subclasses presentes na cena imageada. Por outro lado, a complexidade dos dados hiperespectrais não é totalmente explorada pelo especialista quando somente algumas bandas são utilizadas no processo de seleção de amostras candidatas a MR ou quando somente são analisados espectros fora do contexto espacial da imagem.

Um exemplo de amostras de classes e subclasses obtidas pela triagem manual de um especialista, relativas a subcena 4, é mostrada na Tabela 4.1, onde as classes e subclasses de solos estão representadas com siglas com letras maiúsculas e as demais classes e subclasses com siglas com letras minúsculas. Algumas dessas amostras escolhidas e os seus espectros estão representados nas Figs. 3.3 e 3.4, respectivamente.

No caso mostrado na Tabela 4.1 o especialista agrupou as amostras escolhidas em cinco classes (*vv*, *vs*, *a*, *LV*, *RQ*), com algumas das classes possuindo subclasses. Essa escolha de amostras espectrais candidatas a MR foi feita pelo especialista consultando mapas com fotointerpretação da cena (Fig. 3.3) e pela análise dos espectros de *pixels* na imagem. Para fins de testes foram escolhidas ainda quatro amostras consideradas amostras com misturas de

classes e quatro amostras com falhas de aquisição nos seus espectros, tendo-se assim um total de $K = 76$ amostras espectrais, que posteriormente, no Capítulo 7, serão utilizadas para o teste da metodologia proposta.

TABELA 4.1 - Triagem manual realizada por especialista.

Classe		Subclasse	Número de amostras
vegetação verde:	vv	vv1(a), ..., vv1(f)	6
		vv2(a), ..., vv2(e)	5
		vv3(a), ..., vv3(e)	5
		vv4(a), ..., vv4(e)	5
		vv5(a), ..., vv5(e)	5
vegetação seca:	vs	vs(a), ..., vs(f)	6
água:	a	a(a), a(b), a(c)	3
LATOSSOLO VERMELHO:	LV	LV1(a), ..., LV1(f)	6
		LV2(a), ..., LV2(e)	5
		LV3(a), ..., LV3(e)	5
NEOSSOLO QUARTZARÊNICO:	RQ	RQ1(a), ..., RQ1(g)	7
		RQ2(a), ..., RQ2(e)	5
		RQ3(a), ..., RQ3(e)	5
misturas:	m	m1, ..., m4	4
amostras com falhas no espectro: <i>anoma</i>		<i>anoma1, ..., anoma4</i>	4

4.1.2 Triagem aleatória

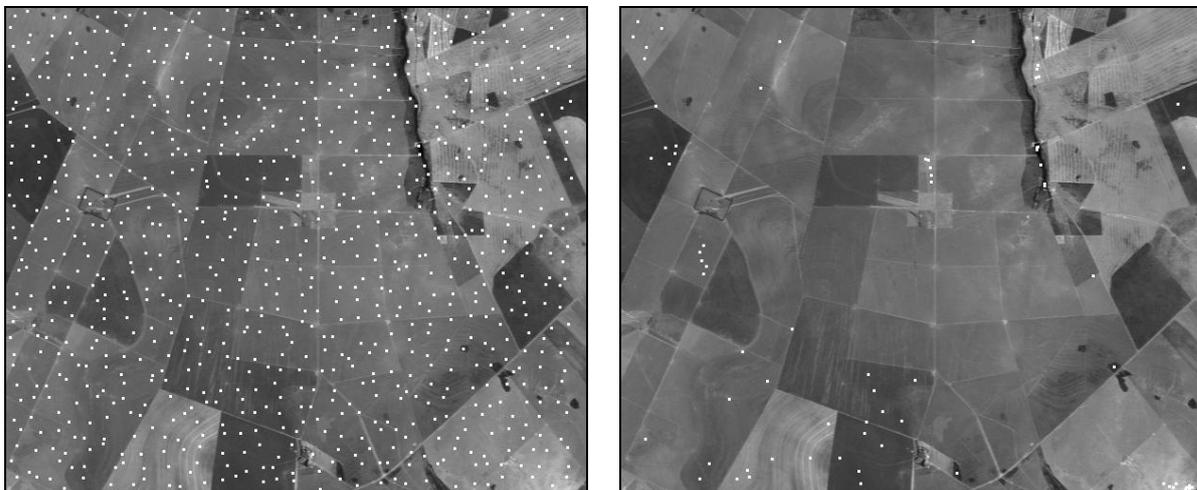
Definido a priori um número de amostras, em geral grande, o processo escolhe automaticamente e aleatoriamente essas amostras na imagem considerada, com a restrição de que as amostras escolhidas fiquem espalhadas uniformemente pela cena considerada.

Um grande problema da amostragem aleatória é que ela, mesmo com a restrição de uniformidade no espalhamento das amostras pela imagem, não garante que todas as classes e subclasses terão sempre amostras selecionadas. Este problema é minimizado selecionando-se muitas amostras no processo de triagem. Entretanto, um número excessivo de amostras candidatas a MR pode aumentar muito o custo computacional do processo de busca dos melhores MR, que será tratado nos Capítulos 5 e 6.

Um aspecto positivo da amostragem aleatória é que as amostras selecionadas na imagem passam a ser candidatas a MR sem serem submetidas a nenhum processo de manipulação que pode, de algum modo e em algumas situações, descaracterizar essas amostras.

A Fig. 4.1a mostra a banda 192 (2.271 nm) da subcena 4 e as posições de $K = 900$ amostras espectrais escolhidas aleatoriamente, mas com a restrição de uniformidade em

termos de concentração de amostras na imagem, realizada pelo método da Amostragem Estratificada Sistemática não Alinhada (AESNA) (CONGALTON, 1988).



(a) Triagem aleatória (AESNA)

(b) Escolha manual (especialista)

FIGURA 4.1 – Triagem de amostras na subcena 4 pela abordagem aleatória e manual.

Para efeitos de comparação a Fig. 4.1b mostra também a posição das 76 amostras candidatas a MR, da Tabela 4.1, escolhidas pelo especialista no processo de triagem manual.

4.1.3 Triagem discriminatória

Neste tipo de abordagem são escolhidas amostras que possuem uma vizinhança espacialmente uniforme e espectralmente homogênea, com contexto espacial e que não sejam redundantes entre si. As definições de uniformidade espacial, de homogeneidade espectral, de contexto espacial e de redundância entre amostras (PIZARRO; FERNANDES, 2007a) serão descritos nas próximas seções, bem como serão apresentados os métodos de testes baseados nestes princípios. A Fig. 4.2 apresenta o fluxograma dos processos de triagem.

A triagem discriminatória apresenta a flexibilidade de poder ser aplicada a todos os *pixels* da imagem hiperespectral ou às amostras selecionadas pela triagem manual ou pela triagem aleatória.

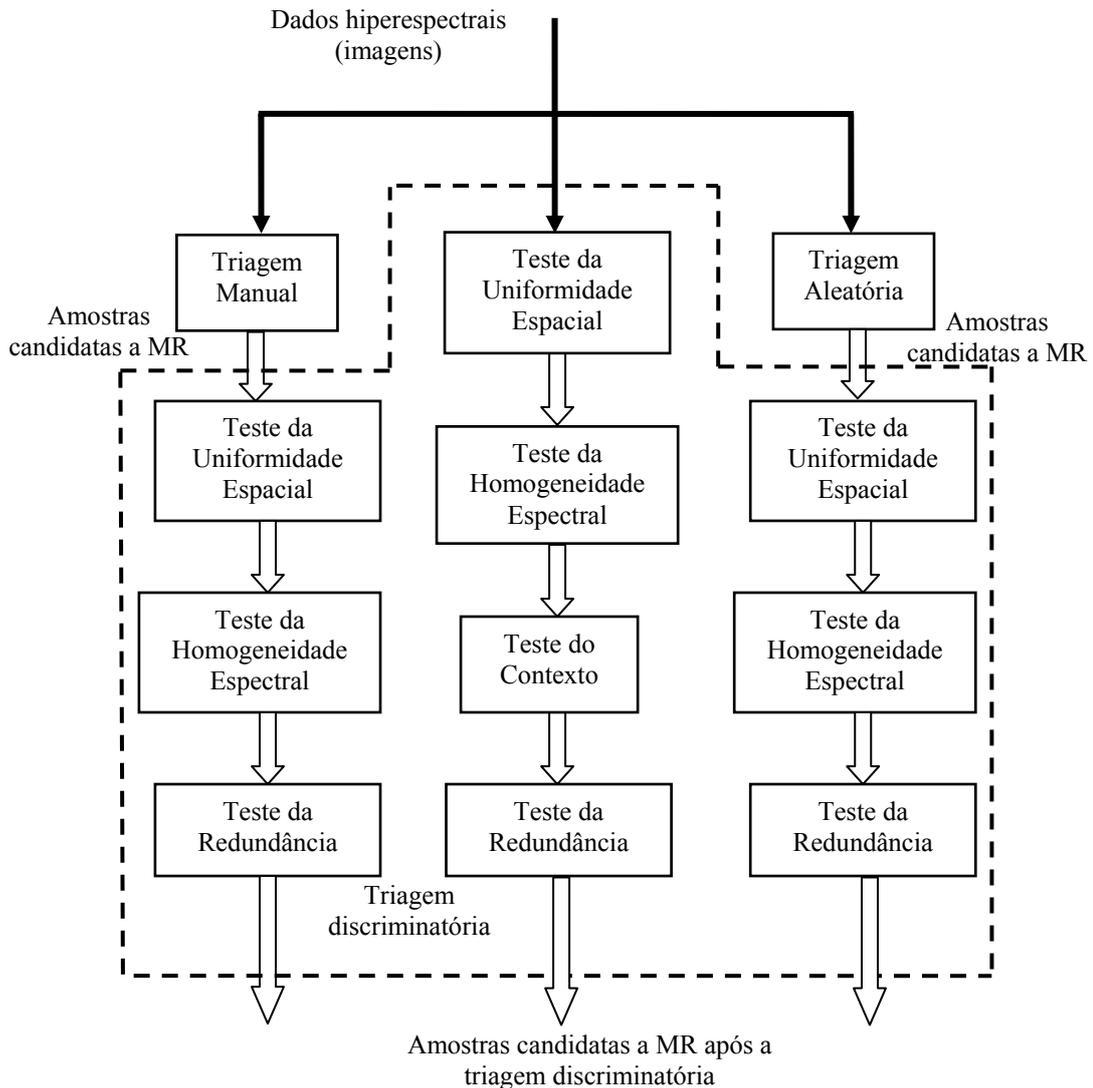


FIGURA 4.2 – Tipos de triagem.

4.2 Uniformidade espacial da vizinhança de uma amostra teste

O teste da uniformidade espacial tem por objetivo verificar a uniformidade das amostras pertencentes à vizinhança $(M + 1) \times (M + 1)$, pré-definida, da amostra espectral sob teste. Este teste verifica quantas amostras, de um total de $(M + 1)^2$ amostras, podem ser consideradas pertencentes a uma mesma classe. Se este número for maior que um limiar de uniformidade espacial dado por $\alpha_u (M + 1)^2$, onde $0,5 < \alpha_u \leq 1$ é o fator de uniformidade espacial, a amostra é considerada uma amostra com vizinhança uniforme. Neste caso uma

nova vizinhança, composta somente pelos espectros considerados de uma mesma classe, é estabelecida.

Consideremos uma amostra k , $1 \leq k \leq K$, onde K é o total de amostras, que está sendo submetida ao processo de triagem, relativa ao *pixel* (i_k, j_k) , e a sua vizinhança $(M+1) \times (M+1)$, com M par, em torno do *pixel* (i_k, j_k) . Os $(M+1)^2$ espectros, de dimensão N , dessa vizinhança, agrupados na matriz I_k , são representados por:

$$I_k = \begin{bmatrix} \bar{I}(i_k - M/2, j_k - M/2) & \dots & \bar{I}(i_k - M/2, j_k + M/2) \\ \dots & \dots & \dots \\ \dots & \bar{I}(i_k, j_k) & \dots \\ \dots & \dots & \dots \\ \bar{I}(i_k + M/2, j_k - M/2) & \dots & \bar{I}(i_k + M/2, j_k + M/2) \end{bmatrix}_{(M+1) \times (M+1) \times N} \quad (4.1)$$

Seja Ω_k o conjunto de espectros da matriz I_k :

$$\Omega_k = \left\{ \begin{array}{l} \bar{I}(p, q) = [I(p, q, 1), I(p, q, 2), \dots, I(p, q, N)]^T \\ | \quad i_k - M/2 \leq p \leq i_k + M/2, \quad j_k - M/2 \leq q \leq j_k + M/2 \end{array} \right\} \quad (4.2)$$

A detecção dos espectros que contaminam a vizinhança da amostra (i_k, j_k) , devido a uma possível mistura de classes, é feita em duas etapas. Na primeira etapa, calcula-se a média espectral de todos os espectros da vizinhança $(M+1) \times (M+1)$ obtendo-se o conjunto:

$$M_k = \{m(p, q), \quad i_k - M/2 \leq p \leq i_k + M/2, \quad j_k - M/2 \leq q \leq j_k + M/2\} \quad (4.3)$$

onde $m(p, q)$ é a média do espectro $\bar{I}(p, q)$:

$$m(p, q) = \frac{1}{N} \sum_{n=1}^N I(p, q, n) \quad (4.4)$$

O conjunto M_k com os seus elementos ordenados, em ordem crescente, será representado por:

$$M_{(k)} = \left\{ M_{(k)}(m), \quad m = 1, 2, \dots, (M+1)^2 \mid M_{(k)}(p) \geq M_{(k)}(q) \text{ para } p \geq q \right\} \quad (4.5)$$

Escolhe-se então como espectro de referência o espectro $\vec{I}(p, q) \in \Omega_k$ cujo valor espectral médio $m(p, q)$ corresponde a mediana do conjunto $M_{(k)}$, ou seja, $m(p, q) = M_{(k)} \left(\left\lfloor (M+1)^2 / 2 \right\rfloor \right)$, onde $\lfloor \cdot \rfloor$ representa o maior inteiro menor ou igual a $(M+1)^2 / 2$.

Portanto, o espectro de referência da vizinhança do *pixel* (i_k, j_k) é dado por:

$$\vec{I}_k = \vec{I}(p, q) \in \Omega_k \mid m(p, q) = M_{(k)} \left(\left\lfloor (M+1)^2 / 2 \right\rfloor \right) \quad (4.6)$$

A escolha do espectro de referência como sendo aquele cuja média espectral corresponde a mediana de $M_{(k)}$ possibilita, em uma primeira aproximação, a rejeição dos espectros que tendem ser muito contrastantes da maioria dos espectros presentes na vizinhança $(M+1) \times (M+1)$ da amostra considerada.

Em uma segunda etapa calcula-se a Coerência Espectral (CE) (CARVALHO JR.; MENESES, 2000), que representa a covariância normalizada, entre o espectro de referência \vec{I}_k e todos os espectros da vizinhança $(M+1) \times (M+1)$ pela expressão:

$$CE(\vec{I}_k, \vec{I}(p, q)) = \frac{\langle \vec{I}_k - \bar{\vec{I}}_k, \vec{I}(p, q) - \bar{\vec{I}}(p, q) \rangle}{\| \vec{I}_k - \bar{\vec{I}}_k \| \cdot \| \vec{I}(p, q) - \bar{\vec{I}}(p, q) \|} \quad (4.7)$$

com $i_k - M/2 \leq p \leq i_k + M/2, \quad j_k - M/2 \leq q \leq j_k + M/2$

onde $\langle \cdot, \cdot \rangle$ representa o produto escalar entre dois vetores, $\| \cdot \|$ representa a norma Euclidiana de um vetor e os vetores $\bar{\vec{I}}_k(p, q)$ e $\bar{\vec{I}}(p, q)$ representam vetores colunas com todos elementos iguais a $M_{(k)} \left(\left\lfloor (M+1)^2 / 2 \right\rfloor \right)$ e $m(p, q)$, respectivamente:

$$\bar{\vec{I}}_k(p, q) = \left[M_{(k)} \left(\left\lfloor (M+1)^2 / 2 \right\rfloor \right), M_{(k)} \left(\left\lfloor (M+1)^2 / 2 \right\rfloor \right), \dots, M_{(k)} \left(\left\lfloor (M+1)^2 / 2 \right\rfloor \right) \right]^T \quad (4.8)$$

$$\vec{\bar{I}}(p, q) = [m(p, q), m(p, q), \dots, m(p, q)]^T \quad (4.9)$$

A escolha de um espectro de referência possibilita uma comparação relativa entre os demais espectros, visto que todos são comparados, através da coerência, com um único padrão.

A seguir rejeita-se os espectros $\vec{\bar{I}}(p, q)$ cuja CE com o espectro de referência for inferior a um certo limiar $0 \leq \psi_e \leq 1$, que será chamado de limiar de seleção espacial. Se $\psi_e = 0$ nenhuma amostra é rejeitada e se $\psi_e = 1$ somente as amostras colineares ao espectro de referência não serão rejeitadas.

Portanto, a nova vizinhança do *pixel* (i_k, j_k) será formada pelo conjunto de vetores espectros:

$$V(i_k, j_k) = \left\{ \begin{array}{l} \vec{\bar{I}}(p, q), \text{ tal que } CE(\vec{\bar{I}}_k, \vec{\bar{I}}(p, q)) \geq \psi_e, \\ \text{com } i_k - M/2 \leq p \leq i_k + M/2, \quad j_k - M/2 \leq q \leq j_k + M/2 \end{array} \right\} \quad (4.10)$$

Uma amostra espectral k , $1 \leq k \leq K$, é considerada candidata a MR se a cardinalidade $\#V(i_k, j_k)$, do conjunto $V(i_k, j_k)$, for igual ou superior ao limiar de uniformidade espacial dado por $\alpha_u \times (M+1)^2$, onde $0,5 < \alpha_u \leq 1$ é o fator de uniformidade espacial.

Seja Ω_U o conjunto de amostras espectrais que passaram pela triagem realizada pelo teste da uniformidade espacial. Então

$$\Omega_U = \left\{ \vec{\bar{I}}(i_k, j_k) \mid \#V(i_k, j_k) \geq \alpha_u \times (M+1)^2, k = 1, 2, \dots, K \right\} \quad (4.11)$$

O fator de uniformidade espacial maior que 50% equivale a se considerar a vizinhança definida por $V(i_k, j_k)$ uniforme somente se esta vizinhança contiver a maioria dos espectros da vizinhança inicial $(M+1) \times (M+1)$.

Após a triagem realizada pelo teste da uniformidade espacial das amostras $1 \leq k \leq K$ obtém-se um novo conjunto com $K_U \leq K$ amostras. Estas amostras são renumeradas de tal

modo que cada amostra k_u , $1 \leq k_u \leq K_U$, tem uma nova vizinhança uniforme dada por $V(i_{k_u}, j_{k_u})$, com cardinalidade

$$\#V(i_{k_u}, j_{k_u}) \geq \alpha_u \times (M+1)^2 \quad (4.12)$$

Deste modo o conjunto Ω_U pode ser reescrito como:

$$\Omega_U = \left\{ \bar{I}(i_{k_u}, j_{k_u}), k_u = 1, 2, \dots, K_U \mid \#V(i_{k_u}, j_{k_u}) \geq \alpha_u \times (M+1)^2 \right\} \quad (4.13)$$

Na seção 8.1, item 8.1.1, mostra-se um exemplo do teste da uniformidade espacial aplicada a três amostras de água (*a*) contaminada com mata galeria (*vv1*). Neste teste a contaminação da água pela vegetação verde é eliminada pela escolha de uma nova vizinhança menos contaminada.

4.3 Homogeneidade espectral da vizinhança de uma amostra teste

O teste de homogeneidade espectral tem como objetivo avaliar a similaridade entre os espectros dos *pixels* dentro da vizinhança $V(i_{k_u}, j_{k_u})$, $1 \leq k_u \leq K_U$, definida como uniforme. Neste teste considera-se que quanto mais similares, banda a banda, são os espectros, das amostras na vizinhança uniforme, mais homogênea espectralmente será a amostra e em consequência pode-se afirmar que os espectros tendem a pertencer à mesma classe e/ou subclasse.

Para o teste da homogeneidade divide-se o conjunto de amostras da vizinhança $V(i_{k_u}, j_{k_u})$ em dois grupos de amostras escolhidas aleatoriamente. Sejam estes grupos de amostras representados por $V_0(i_{k_u}, j_{k_u})$ e $V_1(i_{k_u}, j_{k_u})$ de modo que:

- a) $V_0(i_{k_u}, j_{k_u}) \cap V_1(i_{k_u}, j_{k_u}) = \emptyset$;
- b) $|\#V_0(i_{k_u}, j_{k_u}) - \#V_1(i_{k_u}, j_{k_u})| \leq 1$ e
- c) $\#V_0(i_{k_u}, j_{k_u}) + \#V_1(i_{k_u}, j_{k_u}) = \#V(i_{k_u}, j_{k_u}) \geq \alpha_u (M+1)^2$.

Para cada conjunto de amostras espectrais $V_0(i_{k_u}, j_{k_u})$ e $V_1(i_{k_u}, j_{k_u})$ calcula-se o vetor espectro médio e o vetor variância espectral, dados respectivamente por:

$$\bar{I}_{mr}(i_{k_u}, j_{k_u}) = \frac{1}{\#V_r(i_{k_u}, j_{k_u})} \sum_{\bar{I}(p,q) \in V_r(i_{k_u}, j_{k_u})} \bar{I}(p, q), \quad r = 0, 1 \quad (4.14)$$

e

$$\bar{S}_r(i_{k_u}, j_{k_u}) = \left[S_r(i_{k_u}, j_{k_u}, 1), S_r(i_{k_u}, j_{k_u}, 2), \dots, S_r(i_{k_u}, j_{k_u}, N) \right] \quad (4.15)$$

com

$$S_r(i_{k_u}, j_{k_u}, n) = \frac{1}{\#V_r(i_{k_u}, j_{k_u}) - 1} \sum_{\bar{I}(p,q) \in V_r(i_{k_u}, j_{k_u})} [I(p, q, n) - I_{mr}(i_{k_u}, j_{k_u}, n)]^2 \quad (4.16)$$

com $r = 0, 1$ e $n = 0, 1, \dots, N$ e ainda $I(p, q, n)$ e $I_{mr}(i_{k_u}, j_{k_u}, n)$ representando a n -ésima coordenada dos vetores $\bar{I}(p, q)$ e $\bar{I}_{mr}(i_{k_u}, j_{k_u})$, respectivamente.

A cardinalidade dos conjuntos $V_0(i_{k_u}, j_{k_u})$ e $V_1(i_{k_u}, j_{k_u})$ depende do número de amostras da vizinhança inicial $(M+1) \times (M+1)$ e do resultado do teste da uniformidade espacial. Para que se tenha um número mínimo razoável de amostras nos conjuntos $V_0(i_{k_u}, j_{k_u})$ e $V_1(i_{k_u}, j_{k_u})$ a vizinhança inicial $(M+1) \times (M+1)$ não pode ser muito pequena, entretanto uma vizinhança muito grande pode acarretar uma grande mistura entre classes. Neste trabalho utilizou-se na maioria dos testes apresentados no Capítulo 8 a vizinhança inicial $(M+1) \times (M+1) = 5 \times 5$ pixels.

Faz-se para cada banda espectral um teste estatístico para se verificar a igualdade das médias entre os conjuntos $V_0(i_{k_u}, j_{k_u})$ e $V_1(i_{k_u}, j_{k_u})$. Para cada banda espectral $1 \leq n \leq N$ o teste é definido por:

$$t(i_{k_u}, j_{k_u}, n) = \frac{I_{m0}(i_{k_u}, j_{k_u}, n) - I_{m1}(i_{k_u}, j_{k_u}, n)}{\sqrt{\left(\frac{S_0(i_{k_u}, j_{k_u}, n)}{\#V_0(i_{k_u}, j_{k_u})} + \frac{S_1(i_{k_u}, j_{k_u}, n)}{\#V_1(i_{k_u}, j_{k_u})} \right)}} \quad (4.17)$$

onde os valores $I_{m0}(i_{k_u}, j_{k_u}, n)$, $I_{m1}(i_{k_u}, j_{k_u}, n)$, $S_0(i_{k_u}, j_{k_u}, n)$ e $S_1(i_{k_u}, j_{k_u}, n)$ são as n -ésimas componentes dos vetores $\vec{I}_{m0}(i_{k_u}, j_{k_u})$, $\vec{I}_{m1}(i_{k_u}, j_{k_u})$, $\vec{S}_0(i_{k_u}, j_{k_u})$ e $\vec{S}_1(i_{k_u}, j_{k_u})$, respectivamente.

Admitindo a hipótese de normalidade dos espectros presentes nas vizinhanças testadas a variável aleatória $t(i_{k_u}, j_{k_u}, n)$ tem uma distribuição t-Student (SHANMUGAN; BREIPOHL, 1988, p. 494 e 520) com $g = \#V_0(i_{k_u}, j_{k_u}) + \#V_1(i_{k_u}, j_{k_u}) - 2$ graus de liberdade. Neste caso aceita-se a igualdade das médias com um grau de significância α se $|t(i_{k_u}, j_{k_u}, n)| \leq t_{g; \alpha/2}$. Por exemplo, se $\alpha = 0,1$ tem-se 90% de confiança no teste da igualdade das médias.

A aleatoriedade na escolha das amostras de ambos os grupos assegura maior independência entre os mesmos, facilitando a aplicação do teste estatístico $t(i_k, j_k, n)$.

Seja o indicador de igualdade das médias definido por:

$$I_h(i_{k_u}, j_{k_u}, n) = \begin{cases} 1, & \text{se } |t(i_{k_u}, j_{k_u}, n)| \leq t_{g; \alpha/2} \\ 0, & \text{se } |t(i_{k_u}, j_{k_u}, n)| > t_{g; \alpha/2} \end{cases} \quad (4.18)$$

e o indicador de homogeneidade definido por:

$$Q_h(i_{k_u}, j_{k_u}) = \frac{1}{N} \sum_{n=1}^N I_h(i_{k_u}, j_{k_u}, n) \quad (4.19)$$

Considera-se que o conjunto de amostras na vizinhança $V(i_{k_u}, j_{k_u})$ do *pixel* (i_{k_u}, j_{k_u}) é homogêneo se:

$$Q_h(i_{k_u}, j_{k_u}) \geq \psi_h \quad (4.20)$$

onde $0,5 < \psi_h \leq 1$ é o limiar de homogeneidade. Por exemplo, se o limiar de homogeneidade for $\psi_h = 0,90$, então 90% das faixas espectrais dos espectros médios relativos às duas partições $V_0(i_{k_u}, j_{k_u})$ e $V_1(i_{k_u}, j_{k_u})$ devem ser considerados iguais, com um grau de significância α , para que a vizinhança do *pixel* (i_{k_u}, j_{k_u}) seja considerada espectralmente homogênea.

Seja Ω_H o conjunto de amostras espectrais que passaram pela triagem realizada pelo teste da uniformidade espacial e homogeneidade espectral. Então

$$\Omega_H = \{ \vec{I}(i_{k_u}, j_{k_u}) \mid Q_h(i_{k_u}, j_{k_u}) \geq \psi_h, k_u = 1, 2, \dots, K_U \} \quad (4.21)$$

Após a triagem realizada pelo teste da homogeneidade espectral das amostras $1 \leq k_u \leq K_U \leq K$ obtém-se um novo conjunto com $K_H \leq K_U \leq K$ amostras. Estas amostras são reenumeradas de tal modo que cada amostra k_h , $1 \leq k_h \leq K_H$, tem a vizinhança uniforme e homogênea dada por $V(i_{k_h}, j_{k_h})$.

Deste modo o conjunto Ω_H pode ser reescrito como:

$$\Omega_H = \{ \vec{I}(i_{k_h}, j_{k_h}), k_h = 1, 2, \dots, K_H \mid Q_h(i_{k_h}, j_{k_h}) \geq \psi_h \} \quad (4.22)$$

Ainda a respeito da uniformidade espacial e da homogeneidade espectral, vale comentar que além da mistura devido aos diferentes tipos de materiais associados as diferentes classes e/ou subclasses de alvos presentes na vizinhança, há outros fatores que contribuem para a não uniformidade espacial e para a não homogeneidade espectral. Entre outras causas podem ser citadas as interferências provenientes do ruído térmico nos sistemas eletrônicos dos sensores (CLARK; ROBERTS; CLARK, 2005), falhas de operação dos detectores (HAN et al., 2002; STEIN et al., 2002) e a atenuação atmosférica (WINTER, 2004). Essas interferências afetam as bandas espectrais de forma diferente, fazendo com que algumas bandas espectrais tornem-se menos confiáveis. As variabilidades espectrais impostas por essas interferências podem fazer com que vizinhanças aparentemente homogêneas possam apresentar uma dispersão espectral acentuada (AIAZZI et al., 2001).

Um exemplo de teste de homogeneidade espectral relativo a três amostras d'água com vizinhança espacialmente uniforme é apresentado na seção 8.1, item 8.1.1.

4.4 Contexto espacial da amostra teste

Pode-se ainda impor mais uma restrição às vizinhanças para que de fato possam ser consideradas uniformes e homogêneas. Essa restrição será estabelecida pelo teste de contexto espacial.

Uma amostra espectral $\vec{I}(i_{k_h}, j_{k_h})$ com vizinhança espacialmente uniforme e espectralmente homogênea será considerada uma candidata a MR, somente se $\alpha_c (M_c + 1)^2$ amostras em uma nova vizinhança $(M_c + 1) \times (M_c + 1)$, com M_c par, da amostra considerada no *pixel* (i_{k_h}, j_{k_h}) , também forem espacialmente uniformes e espectralmente homogêneas. O fator $0 \leq \alpha_c \leq 1$ será chamado de fator de contexto.

O contexto espacial procura aglomerados de amostras candidatas a MR que sejam espacialmente uniformes e espectralmente homogêneos. A extensão destes aglomerados é regulada pelo fator de contexto α_c . Quanto mais próximo da unidade for este fator menores extensões terão estes aglomerados e se $\alpha_c = 0$ a imagem tenderá formar um único aglomerado, podendo ficar sem efeito os testes de uniformidade e homogeneidade.

Seja Ω_C o conjunto de amostras que passaram pela triagem realizada pelos testes da uniformidade espacial, homogeneidade espectral e contexto. Então:

$$\Omega_C = \left\{ \vec{I}(i_{k_h}, j_{k_h}) \mid \#V_C(i_{k_h}, j_{k_h}) \geq \alpha_c (M_c + 1)^2, k_h = 1, 2, \dots, K_H \right\} \quad (4.23)$$

onde $\#V_C(i_{k_h}, j_{k_h})$ é a cardinalidade da vizinhança $V_C(i_{k_h}, j_{k_h})$ definida como:

$$V_C(i_{k_h}, j_{k_h}) = \left\{ \vec{I}(p, q) \mid \vec{I}(p, q) \in \Omega_H \right. \\ \left. \text{com } i_{k_h} - M_c/2 \leq p \leq i_{k_h} + M_c/2, j_{k_h} - M_c/2 \leq q \leq j_{k_h} + M_c/2 \right\} \quad (4.24)$$

Após a triagem realizada pelo teste de contexto espacial das amostras $1 \leq k_h \leq K_H \leq K_U \leq K$ obtém-se um novo conjunto com $K_C \leq K_H \leq K_U \leq K$ amostras. Estas

amostras são renumeradas de tal modo que cada amostra k_c , $1 \leq k_c \leq K_C$, tem uma vizinhança uniforme e homogênea dada por $V(i_{k_c}, j_{k_c})$ e cada uma destas k_c amostras tem uma vizinhança $V_C(i_{k_h}, j_{k_h})$ formada por um mínimo de $\alpha_c (M_c + 1)^2$ amostras que também tem uma vizinhança $V(i_{k_c}, j_{k_c})$ espacialmente uniforme e espectralmente homogênea.

Deste modo o conjunto Ω_C pode ser reescrito como:

$$\Omega_C = \left\{ \vec{I}(i_{k_c}, j_{k_c}), k_c = 1, 2, \dots, K_C \mid \#V_C(i_{k_c}, j_{k_c}) \geq \alpha_c (M_c + 1)^2 \right\} \quad (4.25)$$

Um exemplo do teste de contexto espacial, para a subcena 4, na criação de aglomerados é apresentado na seção 8.1, item 8.1.3.

O teste do contexto espacial é aplicável a todos os *pixels* da imagem hiperespectral. Este teste não é aplicável para o conjunto de amostras, candidatas a MR, que tenham sido escolhidas pela triagem manual ou aleatória, pois neste caso a descontinuidade das amostras pode impedir a formação de um contexto espacial.

4.5 Redundância entre amostras espectrais

Uma vez garantido no processo de triagem que as amostras tenham uma vizinhança espacialmente uniforme, espectralmente homogênea e tenham passado pelo teste do contexto, faz-se a eliminação de amostras consideradas redundantes.

O teste da redundância é baseada no cálculo das medidas de similaridades Distância Euclidiana (*DE*) (DUDA; HART; STORK, 2001, p. 37) e *CE* entre um espectro de referência e os espectros médios \vec{I}_{mk_c} , de todas as amostras $1 \leq k_c \leq K_C$ que passaram pelos testes de uniformidade espacial, homogeneidade espectral e contexto espacial. Para a *CE* foi escolhido como espectro de referência \vec{I}_{ref}^{CE} o espectro médio relativo a todos os espectros das amostras $1 \leq k_c \leq K_C$. Para a *DE* o espectro de referência escolhido é $\vec{I}_{ref}^{DE} = -\vec{I}_{ref}^{CE}$, que por ser afastado de todos os demais espectros evita ambigüidades na medida da *DE*.

O espectro médio da amostra k_c é dado por:

$$\bar{I}_{mk_c} = \frac{1}{\#V(i_{k_c}, j_{k_c})} \sum_{\bar{I}(p,q) \in V(i_{k_c}, j_{k_c})} \bar{I}(p, q) \quad (4.26)$$

A *DE* e a *CE* são então calculadas, para $1 \leq k_c \leq K_C$, respectivamente por:

$$DE_{k_c} = DE(\bar{I}_{mk_c}, \bar{I}_{ref}^{DE}) = \|\bar{I}_{mk_c} - \bar{I}_{ref}^{DE}\| = \sqrt{\langle \bar{I}_{mk_c}, \bar{I}_{ref}^{DE} \rangle} = \sqrt{\sum_{n=1}^N (I_{mk_c}(n) - I_{ref}^{DE}(n))^2} \quad (4.27)$$

onde,

$$\bar{I}_{mk_c} = [I_{mk_c}(1), I_{mk_c}(2), \dots, I_{mk_c}(N)] \quad (4.28)$$

e

$$\bar{I}_{ref}^{DE} = [I_{ref}^{DE}(1), I_{ref}^{DE}(2), \dots, I_{ref}^{DE}(N)] \quad (4.29)$$

$$CE_{k_c} = CE(\bar{I}_{mk_c}, \bar{I}_{ref}^{CE}) = \frac{\langle \bar{I}_{mk_c} - \bar{\bar{I}}_{mk_c}, \bar{I}_{ref}^{CE} - \bar{\bar{I}}_{ref}^{CE} \rangle}{\|\bar{I}_{mk_c} - \bar{\bar{I}}_{mk_c}\| \cdot \|\bar{I}_{ref}^{CE} - \bar{\bar{I}}_{ref}^{CE}\|} \quad (4.30)$$

onde,

$$\bar{I}_{ref}^{CE} = [I_{ref}^{CE}(1), I_{ref}^{CE}(2), \dots, I_{ref}^{CE}(N)] \quad (4.31)$$

com

$$\bar{\bar{I}}_{ref}^{CE} = \frac{1}{N} \left[\sum_{n=1}^N I_{ref}^{CE}(n), \sum_{n=1}^N I_{ref}^{CE}(n), \dots, \sum_{n=1}^N I_{ref}^{CE}(n) \right]^T \quad (4.32)$$

e

$$\bar{\bar{I}}_{mk_c} = \frac{1}{N} \left[\sum_{n=1}^N I_{mk_c}(n), \sum_{n=1}^N I_{mk_c}(n), \dots, \sum_{n=1}^N I_{mk_c}(n) \right]^T \quad (4.33)$$

onde $DE_{k_c} \geq 0$, $DE(\vec{I}_{ref}, \vec{I}_{ref}) = 0$, $0 \leq CE_{k_c} \leq 1$ e $CE(\vec{I}_{ref}, \vec{I}_{ref}) = 1$.

Novamente a escolha de um espectro de referência possibilita a comparação relativa entre os espectros, visto que todos são comparados, através da DE e CE , com um único padrão, um para cada tipo de medida.

Ordenando-se as medidas DE_{k_c} , $1 \leq k_c \leq K_C$, em ordem crescente tem-se a seqüência: $DE_{(1)}, DE_{(2)}, \dots, DE_{(K_C)}$ e as medidas CE_{k_c} , $1 \leq k_c \leq K_C$, em ordem decrescente tem-se a seqüência: $CE_{[1]}, CE_{[2]}, \dots, CE_{[K_C]}$.

Para se estimar a proximidade das amostras entre si, baseado no cálculo das medidas de similaridades DE e CE em relação a um espectro de referência, calcula-se, respectivamente, a diferença relativa entre as medidas consecutivas DE e CE das amostras:

$$\Delta DE_{(k_c-1)} = \frac{DE_{(k_c)} - DE_{(k_c-1)}}{DE_{(k_c)}} \quad \text{para } k_c = 2, 3, \dots, K_C \quad (4.34)$$

$$\Delta CE_{[k_c]} = \frac{CE_{[k_c-1]} - CE_{[k_c]}}{CE_{[k_c-1]}} \quad \text{para } k_c = 2, 3, \dots, K_C \quad (4.35)$$

Define-se o limiar de redundância $0 \leq \psi_{rde} \leq 1$ para a medida DE e o limiar de redundância $0 \leq \psi_{rce} \leq 1$ para a medida CE de modo que são consideradas as amostras menos redundantes aquelas que satisfazem os seguinte critérios: $\Delta DE_{(k_c-1)} \geq \psi_{rde}$ e $\Delta CE_{[k_c]} \geq \psi_{rce}$ com $2 \leq k_c \leq K_C$.

Portanto, o conjunto de amostras menos redundantes que satisfaz os dois limiares é dado por:

$$\mathfrak{R}_\cap = \left\{ p \mid p = (K_C) \wedge \left[\begin{array}{l} p = (q-1) \text{ se } \Delta DE_{(q-1)} \geq \psi_{rde}, \\ \text{com } q = 2, 3, \dots, K_C \end{array} \right] \right\} \quad (4.36)$$

$$\cap \left\{ p \mid p = [1] \wedge \left[\begin{array}{l} p = [q] \text{ se } \Delta CE_{[q]} \geq \psi_{rce}, \\ \text{com } q = 2, 3, \dots, K_C \end{array} \right] \right\}$$

Assim sendo o conjunto \mathfrak{R}_\cap é formado pelo conjunto de amostras que entre si sejam distantes, pela medida DE , e pouco correlacionados, pela medida CE . Os limiares $0 \leq \psi_{rde} \leq 1$ e $0 \leq \psi_{rce} \leq 1$ estabelecem o quão distante e o quão coerente devem ser as amostras consideradas não redundantes.

Pode-se ainda utilizar-se a união dos critérios da DE e CE para seleção das amostras menos redundantes:

$$\mathfrak{R}_\cup = \left\{ p \mid p = (K_C) \wedge \left[\begin{array}{l} p = (q-1) \text{ se } \Delta DE_{(q-1)} \geq \psi_{rde}, \\ \text{com } q = 2, 3, \dots, K_C \end{array} \right] \right\} \cup \left\{ p \mid p = [1] \wedge \left[\begin{array}{l} p = [q] \text{ se } \Delta CE_{[q]} \geq \psi_{rce}, \\ \text{com } q = 2, 3, \dots, K_C \end{array} \right] \right\} \quad (4.37)$$

ou ainda as medidas DE ou CE isoladamente:

$$\mathfrak{R}_{DE} = \left\{ p \mid p = (K_C) \wedge \left[\begin{array}{l} p = (q-1) \text{ se } \Delta DE_{(q-1)} \geq \psi_{rde}, \\ \text{com } q = 2, 3, \dots, K_C \end{array} \right] \right\} \quad (4.38)$$

$$\mathfrak{R}_{CE} = \left\{ p \mid p = [1] \wedge \left[\begin{array}{l} p = [q] \text{ se } \Delta CE_{[q]} \geq \psi_{rce}, \\ \text{com } q = 2, 3, \dots, K_C \end{array} \right] \right\} \quad (4.39)$$

Se $\psi_{rde} = 0$ então \mathfrak{R}_{DE} é formado por todas as amostras $1 \leq k_c \leq K_C$. Se $\psi_{rde} = 1$ então \mathfrak{R}_{DE} é formado no máximo por duas amostras: $\mathfrak{R}_{DE} = \{(K_C), (K_C - 1)\}$.

Se $\psi_{rce} = 0$ então \mathfrak{R}_{CE} é formado por todas as amostras $1 \leq k_c \leq K_C$. Se $\psi_{rce} = 1$ então \mathfrak{R}_{CE} é formado no máximo por duas amostras: $\mathfrak{R}_{CE} = \{[1], [2]\}$.

A DE e CE avaliam informações distintas contidas nos espectros. A DE tende a diferenciar os espectros pelos valores de reflectância (brilho), enquanto a CE tende a diferenciar os espectros pela forma (declive e alicive), assim estas medidas avaliam propriedades diferentes nos espectros. Deste modo o conjunto \mathfrak{R}_\cap é o menos indicado para a identificação das amostras menos redundantes, pois somente as amostras que atendem simultaneamente os dois critérios estarão presentes nesse conjunto. Como exemplo consideremos as amostras espectrais da Tabela 4.1: se em \mathfrak{R}_{DE} estiver presente somente a

amostra $a(b)$, como a amostra de água menos redundante, e em \mathfrak{R}_{CE} estiver presente somente a amostra $a(c)$, como a amostra de água menos redundante, então em \mathfrak{R}_\cap não estará presente nenhuma amostra da classe água, apesar da classe água estar presente nos testes relativos a DE e a CE , o que é um péssimo resultado.

Seja Ω_R o conjunto de amostras que passaram pela triagem realizada pelos testes da uniformidade espacial, homogeneidade espectral, contexto e pelo teste da redundância. Então:

$$\Omega_R = \left\{ \vec{I}(i_{k_c}, j_{k_c}) \mid k_c \in \mathfrak{R}_\nabla, k_c = 1, 2, \dots, K_C \right\} \quad (4.40)$$

onde: $\mathfrak{R}_\nabla = \mathfrak{R}_\cup \vee \mathfrak{R}_\nabla = \mathfrak{R}_\cap \vee \mathfrak{R}_\nabla = \mathfrak{R}_{DE} \vee \mathfrak{R}_\nabla = \mathfrak{R}_{CE}$.

Após a triagem realizada pelo teste da redundância das amostras $1 \leq k_c \leq K_C \leq K_H \leq K_U \leq K$ obtém-se um novo conjunto com $K_R \leq K_C \leq K_H \leq K_U \leq K$ amostras. Estas amostras são renumeradas de tal modo que cada amostra k_r , $1 \leq k_r \leq K_R$, tem uma vizinhança uniforme e homogênea dada por $V(i_{k_r}, j_{k_r})$ e formam um conjunto menos redundante.

Deste modo o conjunto Ω_R pode ser reescrito como:

$$\Omega_R = \left\{ \vec{I}(i_{k_r}, j_{k_r}), k_r = 1, 2, \dots, K_R \mid k_r \in \mathfrak{R}_\nabla \right\} \quad (4.41)$$

Um exemplo do teste da redundância aplicado a três amostras d'água com vizinhança espacialmente uniforme é apresentado na seção 8.1, item 8.1.1.

O teste da redundância também pode ser aplicado às amostras a MR escolhidas pela triagem manual ou aleatória, sem que estas amostras tenham sido submetidas ao teste do contexto espacial.

Na seção 8.1 analisa-se o desempenho do processo de triagem discriminatória, relativo à preservação de boas amostras (amostras com vizinhança espacialmente uniforme, espectralmente homogênea e pertencentes a um contexto) e a rejeição de amostras ruins, bem como a sua capacidade de minimizar o conjunto resultante de amostras candidatas a MR, pelo teste de redundância, conservando amostras de todas as classes e subclasses presentes na cena.

No item 8.1.1 analisa-se o resultado da aplicação do teste da uniformidade espacial, homogeneidade espectral e redundância em amostras de água contaminadas com vegetação verde. No item 8.1.2 analisa-se o resultado da aplicação da metodologia de triagem discriminatória, baseada na uniformidade, homogeneidade e redundância nas $K = 900$ amostras escolhidas aleatoriamente e nas $K = 76$ amostras escolhidas manualmente pelo especialista e mostradas na Fig. 4.1. No item 8.1.3 analisa-se a aplicação da triagem discriminatória, incluindo o teste do contexto espacial, em todos os *pixels* da subcena 4.

O resultado da triagem das amostras pela abordagem discriminatória, quer seja ela aplicada a todas as amostras da imagem ou as amostras previamente escolhidas pelo operador ou aleatoriamente, terá como resultado um conjunto de amostras candidatas a MR. Esse conjunto de amostras candidatas é utilizado pelos métodos de busca de MR descritos nos Capítulos 6 e 7.

4.6 Limiares e fatores no processo de triagem discriminatória

O processo de triagem discriminatória é realizado através dos testes de uniformidade espacial, homogeneidade espectral, contexto espacial e redundância.

Os parâmetros, limiares e fatores que regulam a atuação destes testes são resumidos a seguir:

- M : dimensões da vizinhança espacial inicial $(M + 1) \times (M + 1)$ *pixels* de uma amostra espectral;
- α_u : fator de uniformidade espacial, $0,5 < \alpha_u \leq 1$;
- $\alpha_u (M + 1)^2$: limiar de uniformidade espacial;
- ψ_e : limiar de seleção espacial, $0 \leq \psi_e \leq 1$;
- ψ_h : limiar de homogeneidade espectral, $0,5 < \psi_h \leq 1$;
- α : nível de significância do teste estatístico para verificação da igualdade de médias espectrais, banda a banda, entre as vizinhanças $V_0(\cdot, \cdot)$ e $V_1(\cdot, \cdot)$;

- M_c : dimensões da vizinhança espacial $(M_c + 1) \times (M_c + 1)$ *pixels* para o teste do contexto;
- α_c fator de contexto espacial, $0 \leq \alpha_c \leq 1$;
- $\alpha_c (M_c + 1)^2$: limiar do teste de contexto espacial;
- ψ_{rde} : limiar de redundância para a medida *DE*, $0 \leq \psi_{rde} \leq 1$;
- ψ_{rce} : limiar de redundância para a medida *CE*, $0 \leq \psi_{rce} \leq 1$;
- \mathfrak{R}_∇ : conjunto de amostras menos redundantes baseado em:
 - na intersecção dos testes de redundância realizados com a *DE* e *CE* ($\mathfrak{R}_\nabla = \mathfrak{R}_\cap$);
 - na união dos testes de redundância realizados com a *DE* e *CE* ($\mathfrak{R}_\nabla = \mathfrak{R}_\cup$);
 - no teste de redundância realizado com a *DE* ($\mathfrak{R}_\nabla = \mathfrak{R}_{DE}$) e
 - no teste de redundância realizado com a *CE* ($\mathfrak{R}_\nabla = \mathfrak{R}_{CE}$).

No Capítulo 8, onde são mostrados e discutidos os resultados do processo de triagem, alguns limiares e fatores, quando aplicável, são apresentados em forma de porcentagem.

5 Condicionamento Espectral

No Capítulo 4 apresentou-se o método de triagem de amostras da imagem hiperespectral para se compor um conjunto de amostras candidatas a MR. Neste capítulo apresenta-se o condicionamento dos espectros das amostras candidatas a MR. O condicionamento dos espectros consiste na aplicação de uma transformação linear nos espectros médios, calculados em uma vizinhança espacialmente uniforme e espectralmente homogênea, da amostra candidata a MR, com o objetivo de enfatizar-se nuances desejáveis na forma espectral. A transformação linear normalmente utilizada no processamento de dados hiperespectrais é a derivada. Ela atua no espectro original como filtro passa-altas, ressaltando as transições. Além da derivada propõe-se neste trabalho a decomposição pela Transformada Wavelet Discreta (TWD). A decomposição pela TWD proposta atua no espectro original como filtro passa-baixas, filtro suavizador, seguido pelo filtro passa-altas. Na seção 8.3 analisa-se o efeito dos métodos de condicionamentos apresentados nesta seção.

5.1.1 Condicionamento pela derivada discreta

A derivada como função de condicionamento do espectro consegue enfatizar as transições nos espectros provocadas pelas absorções oriundas dos materiais presentes na superfície, elimina o nível médio nos espectros e minimiza os efeitos de fundo na composição do sinal, devido ao solo, quando se analisam áreas com vegetação (TSAI e PHILPOT, 2002; DEMETRIADES-SHAH et al., 1990; ZHANG, RIVARD; SANCHEZ-AZOFEIFA, 2004; DEBBA et al. 2006).

Consideremos uma amostra k , com $1 \leq k \leq \mathbb{k}$, de um grupo de amostras candidatas a MR. A amostra k refere-se ao *pixel* (i_k, j_k) da imagem hiperespectral. Consideremos os *pixels* na vizinhança $V(i_k, j_k)$ da amostra k , vizinhança esta definida na seção 4.2 pela expressão (4.10).

Seja o conjunto de *pixels* V_k o conjunto dos *pixels* que pertencem a vizinhança $V(i_k, j_k)$:

$$V_k = \{(p, q) \mid \vec{I}(p, q) \in V(i_k, j_k)\} \quad (5.1)$$

Os espectros resultantes da derivada discreta dos espectros pertencentes a vizinhança $V(i_k, j_k)$ serão representados por

$$\vec{I}_d(p, q) = \dot{I}(p, q) \text{ com } (p, q) \in V_k \quad (5.2)$$

Calcula-se então o vetor espectro derivado médio da amostra k utilizando-se todos os espectros derivados pertencentes a vizinhança $V(i_k, j_k)$, obtendo-se um espectro de dimensão $(N \times 1)$ dado por:

$$\vec{I}_{md}(i_k, j_k) = \frac{1}{\#V(i_k, j_k)} \sum_{(p, q) \in V_k} \vec{I}_d(p, q) \quad 1 \leq k \leq \mathbb{k} \quad (5.3)$$

A média aplicada em cada amostra k , além de ser utilizada com os propósitos definidos no início desta seção, possibilita a atenuação do ruído presente principalmente nos espectros com baixos valores de reflectância que com o uso da derivada tendem a ser realçados.

5.1.2 Condicionamento pela decomposição com a TWD

Para a aplicação mais eficiente da TWD deve-se inicialmente realizar uma adição de zeros (*zero-padding*) nos espectros. O adição de zeros se faz necessário devido a dimensão do número de bandas do espectro não ser normalmente divisível por 2^s , onde s é o nível de decomposição desejado para os coeficientes *wavelets* obtidos pela TWD (PERCIVAL; WALDEN, 2000, p.308).

Assim adiciona-se zeros a cada espectro da vizinhança V_k da amostra espectral k de tal forma que todos os vetores fiquem com uma dimensão $T > N$ tal que T seja divisível por potências de 2. O vetor espectro da área teste utilizada neste trabalho que era $N = 209$ bandas passa a ter a dimensão $T = 256$ bandas. O novo espectro, com $T-N$ zeros adicionados no fim do vetor $\vec{I}(p, q) \in V(i_k, j_k)$, será representado por:

$$\vec{I}'(p, q) = [I'(p, q, 1), I'(p, q, 2), \dots, I'(p, q, N), I'(p, q, N+1), \dots, I'(p, q, 256)]^T \quad (5.4)$$

com $I'(p, q, n) = 0$, para $N < n \leq T$

Aplica-se então a TWD, com dois níveis de decomposição, ao espectro $\vec{I}^i(p, q)$ com $(p, q) \in V_k$ segundo a estrutura contendo os coeficientes de aproximações ($a^*(.)$) e de detalhes ($d^*(.)$) mostrada na Fig. 5.1.

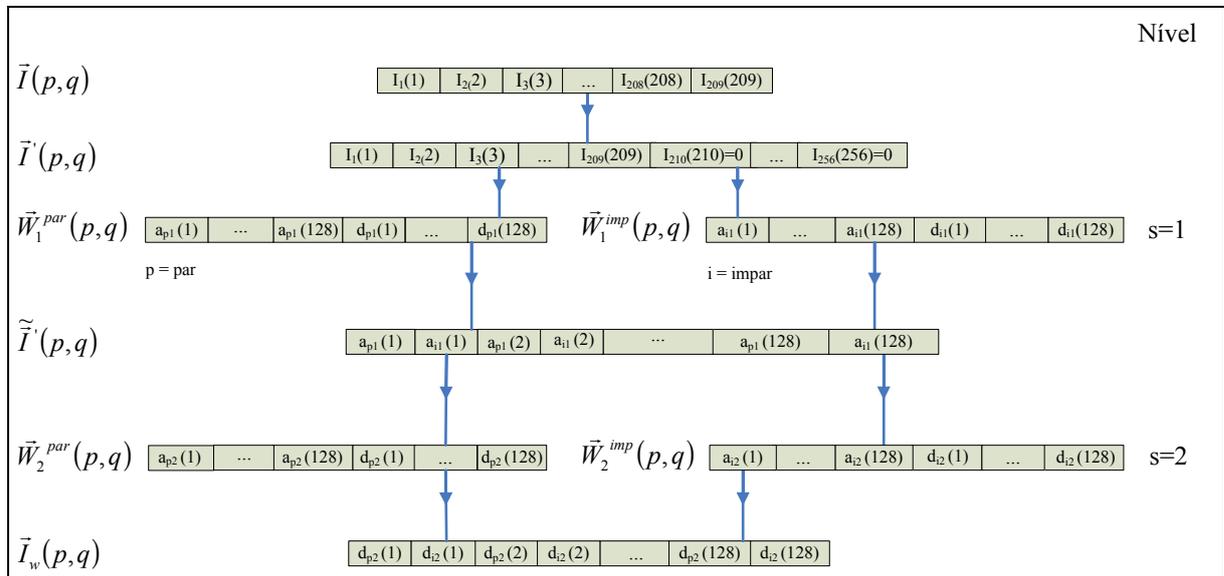


FIGURA 5.1 - Estrutura de aplicação da TWD no espectro $\vec{I}^i(p, q)$ com $(p, q) \in V_k$.

A aplicação da TWD ao espectro $\vec{I}^i(p, q)$, com $(p, q) \in V_k$, segundo a estrutura da Fig. 5.1, é feita de tal modo que os coeficientes do primeiro nível de decomposição ($s = 1$) $W_1^{par}(p, q)$ e $W_1^{imp}(p, q)$ da TWD são aplicados respectivamente a $I^i(p, q, n)$ com n par e n ímpar. A junção, intercalada, adequada dos coeficientes de $W_1^{par}(p, q)$ e $W_1^{imp}(p, q)$ possibilitam a obtenção da TWD sem a dizimação, por 2, de $\vec{I}^i(p, q)$.

A partir da $W_1^{par}(p, q)$ e $W_1^{imp}(p, q)$ são extraídas as aproximações de cada decomposição e composto um novo vetor espectro $\tilde{\vec{I}}^i(p, q)$ com 256 bandas. Aplica-se novamente a TWD e obtêm-se novos coeficientes da decomposição para $s = 2$ (segundo nível de decomposição) $W_2^{par}(p, q)$ e $W_2^{imp}(p, q)$, e o novo vetor espectro $\vec{I}_w(p, q)$ com 209 bandas, sem os zeros inicialmente adicionados, é formado, desta vez a partir dos detalhes deste segundo nível da decomposição.

O espectro $\vec{I}_w(p, q)$ pode ser interpretado como um sinal de diferenças (coeficientes detalhes obtidos no segundo nível da decomposição) do espectro original suavizado (coeficientes aproximações obtidos no primeiro nível da decomposição).

O vetor espectro médio da amostra k utilizando-se os coeficientes da TWD é obtido da expressão:

$$\vec{I}_{mw}(i_k, j_k) = \frac{1}{\#V(i_k, j_k)} \sum_{(p, q) \in V_k} \vec{I}_w(p, q) \quad 1 \leq k \leq \mathbb{k} \quad (5.5)$$

A título de exemplo, as Figs. 5.2 a 5.4 mostram os espectros médios originais e condicionados de cinco classes e subclasses dominantes na subcena 4 ($vv1$, vs , a , $LV1$ e $RQ1$). No condicionamento foram utilizadas além da derivada as TWD com as wavelets Haar e Daubechies de segunda ordem (Db2). O condicionamento foi utilizado no conjunto de amostras que tiveram a sua vizinhança original, de 5×5 pixels, verificada pelo teste da uniformidade espacial. Nas Figs. 5.2 a 5.4 encontram-se ao lado da identificação da amostra no gráfico do espectro original, sem condicionamento, as quantidades de pixels da vizinhança espacialmente uniforme utilizada no cálculo dos espectros médios.

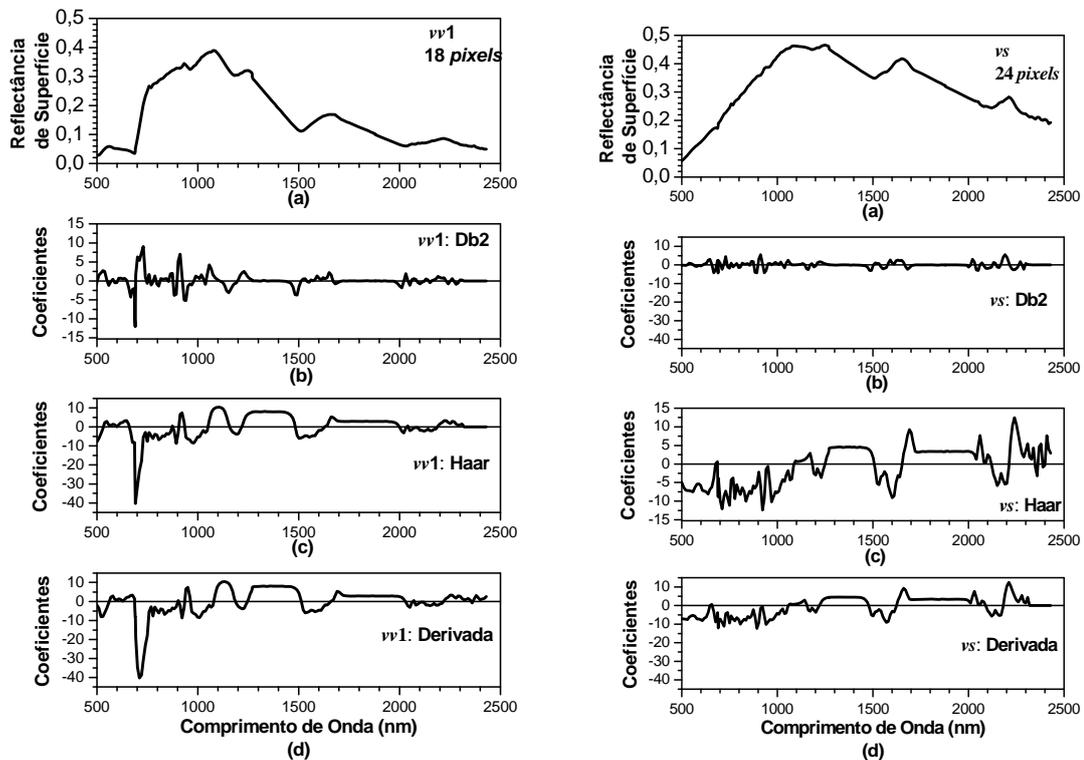


FIGURA 5.2: Condicionamento dos espectros das amostras $vv1$ e vs .

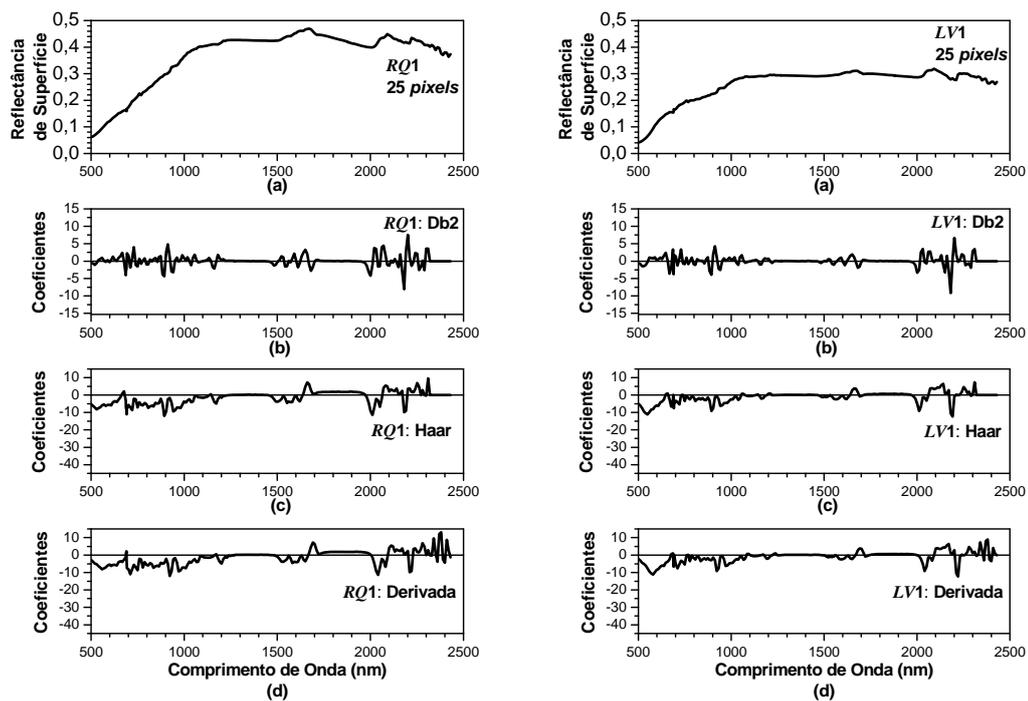


FIGURA 5.3: Condicionamento dos espectros das amostras *RQ1* e *LV1*.

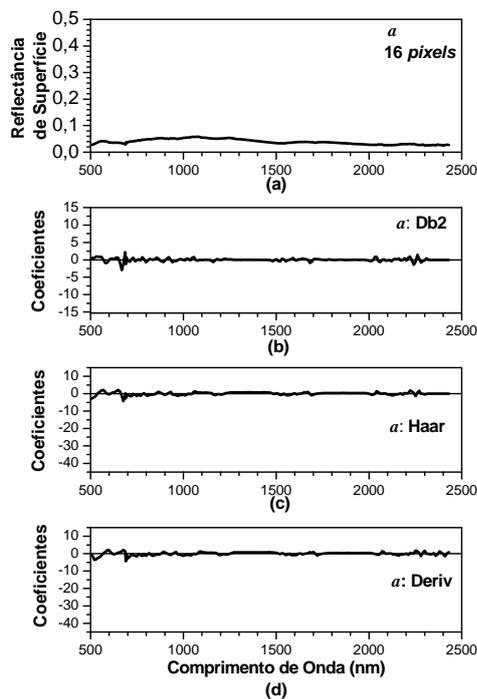


FIGURA 5.4: Condicionamento do espectro da amostra *a* (água).

Nota-se uma grande semelhança entre os espectros condicionados pela derivada e pela TWD-Haar. Nota-se ainda que a TWD-Db2 tem características de derivada segunda, pois os

patamares não nulos presentes nos condicionamentos feitos pela derivada e pela TWD-Haar são anulados no condicionamento feito pela TWD-Db2.

Na seção 8.3 analisa-se os efeitos dos métodos de condicionamento na seleção dos MR, onde se inclui também a TWD com a Coiflet.

6 Busca dos MR pela Máxima Entropia

Este capítulo descreve a proposta de determinação dos MR utilizando-se a medida de similaridade entre espectros através do cálculo da entropia. Com ela é possível medir a similaridade espectral entre várias amostras espectrais ao mesmo tempo. Isto possibilita a determinação do conjunto ótimo de R espectros, dentre um conjunto de \mathbb{k} amostras candidatas a MR. O conjunto ótimo de R espectros será aquele que maximiza a entropia calculada para todos os subconjuntos de R espectros do conjunto de \mathbb{k} amostras. Ressalta-se que diferentemente do que acontece com várias medidas de similaridade espectral, entre elas a DE e a CE , que podem ser utilizadas somente com pares de amostras, a entropia mede diretamente a similaridade entre um conjunto $R \geq 2$ de amostras espectrais. Propõe-se ainda neste capítulo dois métodos para o cálculo do limitante superior do número de MR. Um deles baseia-se na medida de similaridade espectral que utiliza a entropia e o outro no conceito de amostras espectrais bem configuradas.

6.1 Medida de similaridade espectral pela entropia

Considera-se para cada amostra k candidata a MR, com $1 \leq k \leq \mathbb{k}$, o espectro médio $\bar{I}_m(i_k, j_k)$ obtido em uma das seguintes situações:

a) sem condicionamento das amostras candidatas a MR:

$$\bar{I}_m(i_k, j_k) = \frac{1}{\#V(i_k, j_k)} \sum_{(p,q) \in V_k} \bar{I}(p, q) \quad (6.1)$$

b) condicionamento pela derivada:

$$\bar{I}_m(i_k, j_k) = \bar{I}_{md}(i_k, j_k) \quad (6.2)$$

c) condicionamento pela TWD:

$$\bar{I}_m(i_k, j_k) = \bar{I}_{mw}(i_k, j_k) \quad (6.3)$$

Define-se então o espectro médio normalizado como:

$$\tilde{\bar{I}}_m(i_k, j_k) = [\tilde{\bar{I}}_m(i_k, j_k, 1) \ \tilde{\bar{I}}_m(i_k, j_k, 2) \ \dots \ \tilde{\bar{I}}_m(i_k, j_k, N)]^T \quad (6.4)$$

onde as suas componentes normalizadas são descritas por:

$$\tilde{I}_m(i_k, j_k, n) = \frac{I_m(i_k, j_k, n) - \bar{I}_m(i_k, j_k)}{\sqrt{\sum_{u=1}^N \left[\left(I_m(i_k, j_k, u) - \bar{I}_m(i_k, j_k) \right)^2 \right]}} \quad (6.5)$$

com a média do espectro $\bar{I}_m(i_k, j_k)$ dada por

$$\bar{I}_m(i_k, j_k) = \frac{1}{N} \sum_{u=1}^N I_m(i_k, j_k, u) \quad (6.6)$$

e sendo $I_m(i_k, j_k, n)$ a n-ésima componente do vetor espectro médio $\vec{I}_m(i_k, j_k)$:

$$\vec{I}_m(i_k, j_k) = [I_m(i_k, j_k, 1) \ I_m(i_k, j_k, 2) \ \dots \ I_m(i_k, j_k, N)]^T \quad (6.7)$$

A matriz covariância relativa a um conjunto qualquer (e_1, e_2, \dots, e_R) , $2 \leq R \leq \mathbb{k}$, de espectros médios normalizados, diferentes entre si, identificados por $e_1, e_2, \dots, e_R \in \{1, 2, \dots, \mathbb{k}\}$, $e_1 \neq e_2 \neq \dots \neq e_R$, é definida como:

$$L(e_1, e_2, \dots, e_R) = \begin{bmatrix} L_{e_1, e_1} & L_{e_1, e_2} & \dots & L_{e_1, e_R} \\ L_{e_2, e_1} & L_{e_2, e_2} & \dots & L_{e_2, e_R} \\ \dots & \dots & \dots & \dots \\ L_{e_R, e_1} & L_{e_R, e_2} & \dots & L_{e_R, e_R} \end{bmatrix}_{R \times R} \quad (6.8)$$

onde: $L_{p,q}$, $p, q \in \{e_1, e_2, \dots, e_R\}$, é a correlação cruzada entre $\vec{I}_m(i_p, j_p)$ e $\vec{I}_m(i_q, j_q)$ definida por:

$$L_{p,q} = \frac{1}{N} \sum_{n=1}^N \left[\tilde{I}_m(i_p, j_p, n) \tilde{I}_m(i_q, j_q, n) \right] \quad (6.9)$$

sendo:

$$L_{p,q} = L_{q,p} \quad (6.10)$$

Por construção $L(e_1, e_2, \dots, e_R)$ é uma matriz simétrica, ou seja $L(e_1, e_2, \dots, e_R) = L^T(e_1, e_2, \dots, e_R)$, positiva definida e portanto, os seus autovalores calculados através da equação característica $|L(e_1, e_2, \dots, e_R) - \lambda I| = 0$, sendo I a matriz identidade, são números reais não negativos (STRANG e BORRE, 1997). Sejam estes autovalores $\lambda_r \geq 0$, $r = 1, 2, \dots, R$ dispostos em ordem decrescente, tal que: $\lambda_r \leq \lambda_{r-1}$, $r = 2, \dots, R$. Os R autovetores \vec{V}_r , $r = 1, 2, \dots, R$, de dimensões $R \times 1$, associados a cada autovalor λ_r são calculados pela expressão $L(e_1, e_2, \dots, e_R)\vec{V}_r = \lambda_r \vec{V}_r$.

Define-se então a matriz V contendo na sua r -ésima coluna o autovetor \vec{V}_r , $r = 1, 2, \dots, R$. Sendo $L(e_1, e_2, \dots, e_R)$ uma matriz simétrica os autovetores \vec{V}_r , correspondentes a diferentes autovalores, são ortogonais entre si, ou seja $V^T V = I$ e em consequência $V^{-1} = V^T$ (STRANG e BORRE, 1997). Para a situação em que um ou mais autovalores são nulos os autovetores correspondentes a eles estão no *Espaço Nulo* da matriz $L(e_1, e_2, \dots, e_R)$. Isso implica que o *Posto* da Matriz $L(e_1, e_2, \dots, e_R)$ (número máximo de linhas/colunas independentes) é menor que R (BIRKHOFF e MACLANE, 1977).

Uma relação conveniente que pode avaliar a separabilidade de $R \geq 2$ amostras candidatas a MR $e_1, e_2, \dots, e_R \in \{1, 2, \dots, \mathbb{k}\}$, $e_1 \neq e_2 \neq \dots \neq e_R$, é a entropia $H(e_1, e_2, \dots, e_R)$ definida em função dos autovalores normalizados:

$$H(e_1, e_2, \dots, e_R) = - \sum_{r=1}^R p_r \log_R p_r \quad (6.11)$$

onde p_r é o autovalor λ_r normalizado:

$$p_r = \frac{\lambda_r}{\sum_{u=1}^R \lambda_u}, \quad r = 1, 2, \dots, R \quad (6.12)$$

Devido a normalização tem-se que $\sum_{r=1}^R p_r = 1$. Como $\lambda_r \geq 0$ então $p_r \geq 0$ e em virtude da ordenação decrescente dos autovalores tem-se que $p_r \leq p_{r-1}$, $r = 2, \dots, R$.

A entropia $H(e_1, e_2, \dots, e_R)$ é tal que: $0 \leq H(e_1, e_2, \dots, e_R) \leq 1$. Em particular se $H(e_1, e_2, \dots, e_R) = 1$ então todos autovalores são iguais, ou seja, $p_r = 1/R$, $r = 1, 2, \dots, R$, indicando que os $R \geq 2$ espectros que geraram a matriz L_R estão bem alocados no espaço de dimensão R . Se $H(e_1, e_2, \dots, e_R) = 0$ então um dos autovalores normalizados é unitário e os demais são nulos, indicando que os espectros utilizados para a geração da matriz L_R são todos iguais, ou podem ser alocados em uma única dimensão.

6.2 Amostras espectrais bem configuradas

Uma má configuração de amostras (e_1, e_2, \dots, e_R) , $R \geq 2$, $e_1, e_2, \dots, e_R \in \{1, 2, \dots, \mathbb{k}\}$ e $e_1 \neq e_2 \neq \dots \neq e_R$, é aquela em que um ou mais autovalores de $L(e_1, e_2, \dots, e_R)$ são nulos ou muito próximos de zero. Um caso extremo é aquele em que $H(e_1, e_2, \dots, e_R) = 0$. Uma má configuração nas amostras espectrais (e_1, e_2, \dots, e_R) também pode ser caracterizada quando duas amostras quaisquer do conjunto $\{e_1, e_2, \dots, e_R\}$ tem uma entropia muito baixa.

Define-se, segundo o critério da entropia, que um conjunto de $R \geq 2$ espectros médios, calculados por (6.1), ou seja, sem o uso do condicionamento, relativos ao conjunto de amostras (e_1, e_2, \dots, e_R) , $e_1 \neq e_2 \neq \dots \neq e_R$, é bem configurado se:

$$H(p, q) \geq \eta_H \quad \forall p, q \in \{e_1, e_2, \dots, e_R\} \quad (6.13)$$

onde η_H é o limiar de configuração, segundo a medida de similaridade calculada pela entropia.

Ainda, para testar a boa configuração de $R \geq 2$ espectros médios, candidatos a MR, faz-se dois testes adicionais. Um deles baseado na Distância Euclidiana (DE) entre dois vetores espectros e outro baseado na Coerência Espectral (CE), ou coeficiente de correlação normalizado, entre dois vetores espectros.

Para o mesmo conjunto de $R \geq 2$ amostras estabelecido anteriormente define-se, segundo o critério da DE , que este conjunto é bem configurado se:

$$DE(p, q) \geq \eta_{DE} \quad \forall p, q \in \{e_1, e_2, \dots, e_R\} \quad (6.14)$$

onde η_{DE} é o limiar de configuração segundo a DE e

$$DE(p, q) = \sqrt{\sum_{n=1}^N \left[\left(I_m(i_p, j_p, n) - I_m(i_q, j_q, n) \right)^2 \right]} \quad (6.15)$$

onde $\bar{I}_m(i_r, j_r)$ com $r = p, q$ é calculado por (6.1), sem o uso do condicionamento dos espectros. Da expressão (6.15) tem-se que: $DE(p, q) \geq 0$ e $DE(p, p) = 0$.

Um exemplo de conjunto de amostras mal configuradas, segundo o critério da DE , são duas amostras muito próximas entre si, mas muito distantes das demais amostras.

Ainda para o mesmo conjunto de $R \geq 2$ amostras define-se, segundo o critério da CE , que este conjunto é bem configurado se:

$$CE(p, q) \leq \eta_{CE} \quad \forall p, q \in \{e_1, e_2, \dots, e_R\} \quad (6.16)$$

onde η_{CE} é o limiar de configuração segundo a CE e

$$CE(p, q) = \frac{\left| \sum_{n=1}^N \left[\bar{I}_m(i_p, j_p, n) \bar{I}_m(i_q, j_q, n) \right] \right|}{\sqrt{\sum_{n=1}^N \left(\bar{I}_m(i_p, j_p, n) \right)^2 \sum_{n=1}^N \left(\bar{I}_m(i_q, j_q, n) \right)^2}} \quad (6.17)$$

onde

$$\bar{I}_m(i_r, j_r, n) = I_m(i_r, j_r, n) - \bar{\bar{I}}_m(i_r, j_r) \quad r = p, q \quad (6.18)$$

novamente $\bar{I}_m(i_r, j_r)$ é calculado por (6.1), sem o uso do condicionamento dos espectros e $\bar{\bar{I}}_m(i_r, j_r)$ é definido em (6.7). Da expressão (6.17) tem-se que: $0 \leq CE(p, q) \leq 1$ e $CE(p, p) = 1$.

Um exemplo de conjunto de amostras mal configuradas, segundo o critério da CE , são duas amostras muito coerentes entre si, mas pouco coerentes com todas as demais.

Define-se o conjunto U_{bc}^R de amostras espectrais bem configuradas (*bc*) como aquele formado por todos os subconjuntos de R amostras (e_1, e_2, \dots, e_R) que satisfazem pelo menos um dos três critérios de configuração dados por (6.13), (6.14) e (6.16):

$$U_{bc}^R = \left\{ \begin{array}{l} (e_1, e_2, \dots, e_R) \text{ com } e_1, e_2, \dots, e_R \in \{1, 2, \dots, \mathbb{k}\}, \\ e_1 \neq e_2 \neq \dots \neq e_R \text{ tal que, para } \forall p, q \in \{e_1, e_2, \dots, e_R\} \\ \text{tem-se } H(p, q) \geq \eta_H \vee DE(p, q) \geq \eta_{DE} \vee CE(p, q) \leq \eta_{CE} \end{array} \right\} \quad (6.19)$$

A estimativa dos limiares de configuração $(\eta_{DE}, \eta_{CE}$ e $\eta_H)$ será abordada na seção 4.6.

Análises realizadas nas imagens da subcena 4, 5 e 7 mostraram que o teste de boa configuração, entre um conjunto de amostras espectrais, sem o uso do condicionamento espectral, definido no Capítulo 5, deram melhores resultados do que o teste de configuração realizado com os espectros condicionados. No conjunto de imagens testes utilizadas o uso do condicionamento espectral mostrou-se mais eficiente nas tarefas de busca dos MR, onde se procura o conjunto de espectros mais diferentes entre si, do que na tarefa de rejeição de amostras espectrais muito semelhantes, operação esta realizada pelo teste da boa configuração de um conjunto de amostras.

6.3 Determinação dos MR pela máxima entropia

Define-se para um conjunto de amostras espectrais (e_1, e_2, \dots, e_R) , $R \geq 2$, tal que $e_1, e_2, \dots, e_R \in \{1, 2, \dots, \mathbb{k}\}$ e $e_1 \neq e_2 \neq \dots \neq e_R$, a função entropia modificada:

$$\Delta_H(e_1, e_2, \dots, e_R) = \begin{cases} H(e_1, e_2, \dots, e_R), & \text{se } (e_1, e_2, \dots, e_R) \in U_{bc}^R \\ 0 & , \text{se } (e_1, e_2, \dots, e_R) \notin U_{bc}^R \end{cases} \quad (6.20)$$

Escolhe-se como MR os R espectros médios $(e_{h_1}, e_{h_2}, \dots, e_{h_R})$, dentre as \mathbb{k} amostras candidatas, que fazem com que a entropia modificada seja máxima:

$$(e_{h_1}, e_{h_2}, \dots, e_{h_R}) = \arg \left[\max_{(e_1, e_2, \dots, e_R)} (\Delta_H(e_1, e_2, \dots, e_R) \neq 0) \right] \quad (6.21)$$

Segundo este critério os $R \geq 2$ espectros médios escolhidos como MR são dados por:

$$\vec{I}_m(i_{e_{h_r}}, j_{e_{h_r}}) \quad r = 1, 2, \dots, R \quad (6.22)$$

Nota-se que só podem ser escolhidos como MR os $R \geq 2$ espectros médios, referentes as amostras $(e_{h_1}, e_{h_2}, \dots, e_{h_R})$, se estes espectros forem bem configurados.

Na seção 8.2 são realizados testes para a análise do desempenho do método proposto utilizando-se amostras espectrais controladas da subcena 4. Nesta mesma seção a subcena 5 também é submetida aos testes e análise do desempenho do método.

6.4 Limitante superior do número de MR

Dentre um dado conjunto de \mathbb{k} espectros escolhidos como candidatos a MR qual é o número $R \geq 2$ de MR?

Há na literatura algumas tentativas de solução deste difícil problema (GROSS e SCHOTT, 1996; PIZARRO e FERNANDES, 2005a; BIOUCAS-DIAS, 2005; CHANG e DU, 2004). Neste trabalho, propõem-se dois limitantes superiores para o número de MR presentes na cena. Um deles baseia-se no conceito de amostras bem configuradas e depende dos limiares η_{DE} , η_{CE} e η_H , definidos na seção 6.2. O outro baseia-se em um limitante inferior para a entropia calculada para o conjunto de $R \geq 2$ espectros escolhidos como MR.

O cálculo do primeiro limitante considera para o número de MR igual a $2, 3, \dots, \mathfrak{R}_1 \leq \mathbb{k}$ as seguintes escolhas possíveis, bem configuradas, de MR:

$$(e_1^2, e_2^2), (e_1^3, e_2^3, e_3^3), \dots, (e_1^{\mathfrak{R}_1}, e_2^{\mathfrak{R}_1}, \dots, e_{\mathfrak{R}_1}^{\mathfrak{R}_1}),$$

com $e_1^r, e_2^r, \dots, e_r^r \in \{1, 2, \dots, \mathbb{k}\}$ e $r = 2, \dots, \mathfrak{R}_1$ onde não foi possível achar um conjunto de $\mathfrak{R}_1 + 1$ MR que fossem bem configuradas, ou seja:

$$(e_1^R, e_2^R, \dots, e_R^R) \in U_{bc}^R \quad \text{para} \quad 2 \leq R \leq \mathfrak{R}_1 \quad \text{e} \quad U_{bc}^{\mathfrak{R}_1+1} = \emptyset.$$

Com estas condições estabelece-se que o limitante superior do número de MR é o valor \mathfrak{R}_1 , ou seja,

$$2 \leq R \leq \mathfrak{R}_1 \quad (6.23)$$

O cálculo do segundo limitante estabelece um valor mínimo, H_{min} , para o valor da Entropia do conjunto de MR escolhido.

Consideremos que para o número de MR igual a $2, 3, \dots, \mathfrak{R}_2 \leq \mathbb{k}$ tem-se as seguintes escolhas de MR:

$$(e_1^2, e_2^2), (e_1^3, e_2^3, e_3^3), \dots, (e_1^{\mathfrak{R}_2}, e_2^{\mathfrak{R}_2}, \dots, e_{\mathfrak{R}_2}^{\mathfrak{R}_2}),$$

com $e_1^r, e_2^r, \dots, e_r^r \in \{1, 2, \dots, \mathbb{k}\}$, $r = 2, \dots, \mathfrak{R}_2$, $(e_1^R, e_2^R, \dots, e_R^R) \in U_{bc}^R$ para $2 \leq R \leq \mathfrak{R}_2$ e ainda:

$$\Delta_{H_R}(e_1, e_2, \dots, e_R) \geq H_{min} \quad \text{para} \quad 2 \leq R \leq \mathfrak{R}_2 \quad (6.24)$$

sendo que:

$$\Delta_{H_R}(e_1, e_2, \dots, e_{\mathfrak{R}_2+1}) < H_{min} \quad \text{para} \quad (e_1^{\mathfrak{R}_2+1}, e_2^{\mathfrak{R}_2+1}, \dots, e_{\mathfrak{R}_2+1}^{\mathfrak{R}_2+1}) \in U_{bc}^{\mathfrak{R}_2+1} \quad (6.25)$$

Para estas condições estabelece-se que o limitante superior do número de MR é o valor \mathfrak{R}_2 , ou seja:

$$2 \leq R \leq \mathfrak{R}_2 \quad (6.26)$$

Como o segundo limiar exige também que os conjuntos de amostras sejam bem configurados tem-se que:

$$\mathfrak{R}_2 \leq \mathfrak{R}_1 \quad (6.27)$$

Na seção 8.2, além do teste da busca dos MR pela medida de similaridade espectral pela entropia, avalia-se também o número máximo de MR para as subcenas 4 e 5.

6.5 Determinação dos limiares de configuração

Na seção 6.2 foi estabelecido o conceito de amostras espectrais bem configuradas em função de três limiares de configuração: η_{DE} , η_{CE} e η_H . Os valores desses limiares controlam os subconjuntos de $R \geq 2$ amostras, ditas bem configuradas, que podem ser formados com as

\mathbb{k} amostras disponíveis. O subconjunto de $R \geq 2$ amostras que apresentar a maior entropia, como estabelecido na seção 6.3, determina os MR.

O uso de limiares de sensibilidade no processo de busca dos MR, diferentes dos propostos neste trabalho, também são usados nos seguintes procedimentos e medidas: ORASIS (*Optical Real-Time Adaptive Spectral Identification*) (BOWLES et al, 1995); AMEE (*Automated Morphological Endmember Extraction*) (PLAZA et al., 2002) e VCA (*Vertex Component Analysis*) (NASCIMENTO; DIAS, 2005). Nascimento (1997) salienta que não há um valor de limiar fixo ideal, pois os valores dos limiares dependem do tipo da imagem, do uso do solo, do período em que os dados foram adquiridos e dos objetivos da pesquisa. Geralmente a escolha dos valores dos limiares é feita após vários testes e análise dos resultados, embora haja técnicas que procuram escolher de modo automático os limiares (SAHOO et al., 1988).

Para a determinação dos limiares de configuração η_{DE} , η_{CE} e η_H utilizam-se neste trabalho os fatores de configuração α_{DE} , α_{CE} e α_H , onde:

$$0 \leq \alpha_{DE} \leq 1 \quad (6.28)$$

$$0 \leq \alpha_{CE} \leq 1 \quad (6.29)$$

$$0 \leq \alpha_H \leq 1 \quad (6.30)$$

Nos próximos itens mostra-se como os fatores de configuração α_{DE} , α_{CE} e α_H se relacionam com os limiares de configuração η_{DE} , η_{CE} e η_H .

6.5.1 Limiar de configuração associado à DE

A partir dos espectros médios $\bar{I}_m(i_k, j_k)$ calcula-se a DE entre todos os pares de espectros das \mathbb{k} amostras candidatas a MR e forma-se o conjunto U_{DE} contendo estas distâncias:

$$U_{DE} = \{DE(p, q) \mid p \neq q \wedge p, q = 1, 2, \dots, \mathbb{k}\} \quad (6.31)$$

Ordenando-se em ordem crescente os $(\mathbb{k}^2 - \mathbb{k})/2$ elementos do conjunto U_{DE} obtém-se o conjunto:

$$U_{(DE)} = \left\{ \begin{array}{l} DE_{(u)} \mid DE_{(u)} \in U_{DE}, \text{ para } u = 1, 2, \dots, (\mathbb{k}^2 - \mathbb{k})/2 \wedge \\ DE_{(p)} \geq DE_{(q)}, \text{ para } p \geq q \end{array} \right\} \quad (6.32)$$

Adota-se então como limiar de configuração o valor:

$$\eta_{DE} = DE_{\left(\lceil \alpha_{DE} \cdot (\mathbb{k}^2 - \mathbb{k})/2 \rceil\right)} \quad (6.33)$$

onde $\lceil \cdot \rceil$ representa o maior valor inteiro maior ou igual a $\alpha_{DE} \cdot (\mathbb{k}^2 - \mathbb{k})/2$.

Se $\alpha_{DE} = 0,25$ o limiar de configuração η_{DE} rejeita o quartil inferior da seqüência no conjunto $U_{(DE)}$. Se $\alpha_{DE} = 0$ o limiar fica sem efeito, pois todas as medidas $DE_{(u)}$ são maiores ou iguais a $\eta_{DE} = DE_{(0)}$.

6.5.2 Limiar de configuração para CE

Da mesma forma que a DE a partir $\vec{I}_m(i_k, j_k)$ calcula-se a CE entre todos os pares de espectros das \mathbb{k} amostras candidatas a MR e forma-se o conjunto U_{CE} contendo estas coerências:

$$U_{CE} = \{CE(p, q) \mid p \neq q \wedge p, q = 1, 2, \dots, \mathbb{k}\} \quad (6.34)$$

Ordenando-se em ordem decrescente os $(\mathbb{k}^2 - \mathbb{k})/2$ elementos do conjunto U_{CE} obtém-se o conjunto:

$$U_{[CE]} = \left\{ \begin{array}{l} CE_{[u]} \mid CE_{[u]} \in U_{CE}, \text{ para } u = 1, 2, \dots, (\mathbb{k}^2 - \mathbb{k})/2 \wedge \\ CE_{[p]} \leq CE_{[q]}, \text{ para } p \geq q \end{array} \right\} \quad (6.35)$$

Adota-se então como limiar de configuração o valor:

$$\eta_{CE} = CE_{\left[\left[\alpha_{CE} \cdot (\mathbb{k}^2 - \mathbb{k}) / 2\right]\right]} \quad (6.36)$$

Se $\alpha_{CE} = 0,25$ o limiar de configuração η_{CE} rejeita o quartil inferior da seqüência no conjunto $U_{[CE]}$. Se $\alpha_{CE} = 0$ o limiar fica sem efeito, pois todas as medidas $CE_{[u]}$ são menores ou iguais a $\eta_{CE} = CE_{[0]}$.

6.5.3 Limiar de configuração para entropia

Calcula-se a entropia, utilizando-se (6.11) entre todos os pares de espectros das \mathbb{k} amostras candidatas a MR e forma-se o conjunto U_H contendo estas entropias:

$$U_H = \{H(p, q) \mid p \neq q \wedge p, q = 1, 2, \dots, \mathbb{k}\} \quad (6.37)$$

Ordenando-se em ordem crescente os $(\mathbb{k}^2 - \mathbb{k}) / 2$ elementos do conjunto U_H obtém-se o conjunto:

$$U_{(H)} = \left\{ \begin{array}{l} H_{(u)} \mid H_{(u)} \in U_H, \text{ para } u = 1, 2, \dots, (\mathbb{k}^2 - \mathbb{k}) / 2 \wedge \\ H_{(p)} \geq H_{(q)}, \text{ para } p \geq q \end{array} \right\} \quad (6.38)$$

adota-se então como limiar de configuração o valor:

$$\eta_H = H_{\left(\left[\alpha_H \cdot (\mathbb{k}^2 - \mathbb{k}) / 2\right]\right)} \quad (6.39)$$

Se $\alpha_H = 0,25$ o limiar de configuração η_H rejeita o quartil inferior da seqüência no conjunto $U_{(H)}$. Se $\alpha_H = 0$ o limiar fica sem efeito, pois todas as medidas $H_{(u)}$ são maiores ou iguais a $\eta_H = H_{(0)}$.

7 Busca dos MR com Múltiplas Medidas de Similaridade Espectral

No Capítulo 6 foi proposta uma técnica para a determinação dos MR com base no cálculo da entropia como medida de similaridade espectral de um conjunto de várias amostras bem configuradas. Neste capítulo descreve-se um método de escolha de MR em imagens hiperespectrais utilizando-se a combinação de três medidas de similaridade espectral. As medidas de similaridade utilizadas, além da entropia, são a Distância Euclidiana (DE) e a Coerência Espectral (CE). O processo de múltiplas medidas procura explorar as particularidades de cada medida utilizada, de modo que haja complementaridade de propriedades que levem a uma melhor escolha dos MR representativos de uma dada cena. A necessidade de múltiplas medidas já tinha sido observada em Pizarro e Fernandes (2005a e b) que empregaram somente a CE para a busca de MR. As medidas DE e CE foram escolhidas por serem medidas extensivamente empregadas em processamento de sinais, detecção, classificação e reconhecimento de padrões e por apresentarem, de um modo geral, um bom desempenho (ROBILA; GERSHMAN, 2005).

Diferentemente da entropia, a DE e a CE são medidas realizadas entre pares de amostras espectrais. A DE e a CE de um conjunto de amostras espectrais serão consideradas como a DE média e a CE média, respectivamente. Como este processo de média pode trazer resultados inconvenientes, devido à má configuração do conjunto de amostras, utiliza-se também o conceito de amostras bem configuradas, proposto no Capítulo 6, para se descartar conjuntos de amostras inconvenientes.

7.1 Distância Euclidiana como medida de similaridade espectral

Dentre um conjunto de \mathbb{k} amostras candidatas a MR seja um conjunto qualquer de $R \geq 2$ amostras (e_1, e_2, \dots, e_R) tal que $e_1, e_2, \dots, e_R \in \{1, 2, \dots, \mathbb{k}\}$ e $e_1 \neq e_2 \neq \dots \neq e_R$. Define-se o conjunto $U_{DE}(e_1, e_2, \dots, e_R)$ formado por todas as DE entre as diferentes combinações de duas amostras, sem repetição:

$$U_{DE}(e_1, e_2, \dots, e_R) = \{DE(p, q) \mid p \neq q \wedge p, q \in \{e_1, e_2, \dots, e_R\}\} \quad (7.1)$$

onde a DE entre duas amostras, $DE(p, q)$, é definida como:

$$DE(p, q) = \sqrt{\sum_{n=1}^N \left[\left(I_m(i_p, j_p, n) - I_m(i_q, j_q, n) \right)^2 \right]} \quad (7.2)$$

sendo $I_m(i_r, j_r, n)$ a n -ésima componente do vetor $\vec{I}_m(i_r, j_r)$, com $r = p, q$. Este vetor é calculado por (6.1), (6.2) ou (6.3), dependendo do tipo de condicionamento desejado.

O conjunto $U_{DE}(e_1, e_2, \dots, e_R)$ tem $\binom{R}{2} = (R^2 - R)/2$ elementos e a sua média será representada por:

$$DE(e_1, e_2, \dots, e_R) = \frac{2}{R(R-1)} \sum_{p=1}^{R-1} \sum_{q=p+1}^R DE(e_p, e_q) \quad (7.3)$$

Como $DE(e_p, e_q) \geq 0$, então $DE(e_1, e_2, \dots, e_R) \geq 0$.

O conjunto de amostras bem configuradas, proposto na Seção 6.3, é definido por:

$$U_{bc}^R = \left\{ \begin{array}{l} (e_1, e_2, \dots, e_R) \text{ com } e_1, e_2, \dots, e_R \in \{1, 2, \dots, \mathbb{k}\}, e_1 \neq e_2 \neq \dots \neq e_R \\ \text{tal que, para } \forall p, q \in \{e_1, e_2, \dots, e_R\} \text{ tem-se} \\ H(p, q) \geq \eta_H \vee DE(p, q) \geq \eta_{DE} \vee CE(p, q) \leq \eta_{CE} \end{array} \right\} \quad (7.4)$$

Para um conjunto de amostras bem configuradas (e_1, e_2, \dots, e_R) define-se a DE média modificada como:

$$\Delta_{DE}(e_1, e_2, \dots, e_R) = \begin{cases} DE(e_1, e_2, \dots, e_R), & \text{se } (e_1, e_2, \dots, e_R) \in U_{bc}^R \\ 0, & \text{se } (e_1, e_2, \dots, e_R) \notin U_{bc}^R \end{cases} \quad (7.5)$$

Escolhe-se então como MR os R espectros médios $(e_{d_1}, e_{d_2}, \dots, e_{d_R})$, dentre as \mathbb{k} amostras candidatas, que fazem com que a DE média modificada seja máxima:

$$(e_{d_1}, e_{d_2}, \dots, e_{d_R}) = \arg \left[\max_{(e_1, e_2, \dots, e_R)} \left(\Delta_{DE}(e_1, e_2, \dots, e_R) \neq 0 \right) \right] \quad (7.6)$$

7.2 Coerência espectral como medida de similaridade espectral

Dentre um conjunto de \mathbb{k} amostras candidatas a MR seja um conjunto qualquer de $R \geq 2$ amostras (e_1, e_2, \dots, e_R) tal que $e_1, e_2, \dots, e_R \in \{1, 2, \dots, \mathbb{k}\}$ e $e_1 \neq e_2 \neq \dots \neq e_R$. Defina-se o conjunto $U_{CE}(e_1, e_2, \dots, e_R)$ formado por todas as CE entre as diferentes combinações de duas amostras, sem repetição:

$$U_{CE}(e_1, e_2, \dots, e_R) = \{CE(p, q) \mid p \neq q \wedge p, q \in \{e_1, e_2, \dots, e_R\}\} \quad (7.7)$$

onde a CE entre duas amostras, $CE(p, q)$, é definida como:

$$CE(p, q) = \frac{\left| \sum_{n=1}^N [\tilde{I}_m(i_p, j_p, n) \tilde{I}_m(i_q, j_q, n)] \right|}{\sqrt{\sum_{n=1}^N (\tilde{I}_m(i_p, j_p, n))^2 \sum_{n=1}^N (\tilde{I}_m(i_q, j_q, n))^2}} \quad (7.8)$$

sendo $\tilde{I}_m(i_r, j_r, n)$ para $r = p, q$ definido em (6.18).

O conjunto $U_{CE}(e_1, e_2, \dots, e_R)$ tem $\binom{R}{2} = (R^2 - R)/2$ elementos e a sua média será representada por:

$$CE(e_1, e_2, \dots, e_R) = \frac{2}{R(R-1)} \sum_{p=1}^{R-1} \sum_{q=p+1}^R CE(e_p, e_q) \quad (7.9)$$

Como $CE(e_p, e_q) \geq 0$, então $CE(e_1, e_2, \dots, e_R) \geq 0$.

Utilizando-se novamente o conceito de amostras bem configuradas, proposto na Seção 6.3, define-se que a CE média modificada entre um conjunto (e_1, e_2, \dots, e_R) de amostras é dado por:

$$\Delta_{CE}(e_1, e_2, \dots, e_R) = \begin{cases} CE(e_1, e_2, \dots, e_R), & \text{se } (e_1, e_2, \dots, e_R) \in U_{bc}^R \\ 0, & \text{se } (e_1, e_2, \dots, e_R) \notin U_{bc}^R \end{cases} \quad (7.10)$$

Escolhe-se como MR os R espectros médios $(e_{c_1}, e_{c_2}, \dots, e_{c_R})$, dentre as \mathbb{k} amostras candidatas, que fazem com que a CE média modificada seja mínima:

$$(e_{c_1}, e_{c_2}, \dots, e_{c_R}) = \arg \left[\min_{(e_1, e_2, \dots, e_R)} (\Delta_{CE}(e_1, e_2, \dots, e_R) \neq 0) \right] \quad (7.11)$$

7.3 Utilização de múltiplas medidas

Analisa-se nesta seção a utilização conjunta das medidas de similaridades espectral segundo a entropia modificada, proposta na Seção 6.3:

$$(e_{h_1}, e_{h_2}, \dots, e_{h_R}) = \arg \left[\max_{(e_1, e_2, \dots, e_R)} (\Delta_H(e_1, e_2, \dots, e_R) \neq 0) \right] \quad (7.12)$$

e segundo a DE média modificada expressa por (7.5) e à CE média modificada dada por (7.10).

Uma possível escolha para a medida conjunta é definir como MR final as amostras $(e_1^f, e_2^f, \dots, e_R^f)$ que aparecem com maior frequência no conjunto formado pelas soluções dadas pela DE , CE e entropia, soluções estas representadas por:

$$\{e_{d_1}, e_{d_2}, \dots, e_{d_R}, e_{c_1}, e_{c_2}, \dots, e_{c_R}, e_{h_1}, e_{h_2}, \dots, e_{h_R}\}.$$

Uma outra possível abordagem é procurar o conjunto de R amostras $(e_1^3, e_2^3, \dots, e_R^3)$ que satisfaçam simultaneamente os três critérios de medidas, baseados na DE , CE e entropia, ou seja:

$$(e_1^3, e_2^3, \dots, e_R^3) = \max_{(e_1, e_2, \dots, e_R)} (\Delta_{DE}(e_1, e_2, \dots, e_R) \neq 0) \quad (7.13)$$

e

$$(e_1^3, e_2^3, \dots, e_R^3) = \min_{(e_1, e_2, \dots, e_R)} (\Delta_{CE}(e_1, e_2, \dots, e_R) \neq 1) \quad (7.14)$$

e

$$(e_1^3, e_2^3, \dots, e_R^3) = \max_{(e_1, e_2, \dots, e_R)} (\Delta_H(e_1, e_2, \dots, e_R) \neq 0) \quad (7.15)$$

Em uma terceira abordagem procura-se seqüencialmente o conjunto de R amostras $(e_1^2, e_2^2, \dots, e_R^2)$, bem configuradas, que satisfaçam simultaneamente dois dos três critérios de medidas, baseados na DE , CE e entropia.

Nessa abordagem:

a) escolhe-se um primeiro conjunto de R amostras candidatas $(e_1^{2a}, e_2^{2a}, \dots, e_R^{2a})$, faz-se $(e_1^2, e_2^2, \dots, e_R^2) = (e_1^{2a}, e_2^{2a}, \dots, e_R^{2a})$ e estabelece-se os limiares $(\Delta_c, \Delta_d, \Delta_h)$ como:

$$\Delta_d = \Delta_{DE}(e_1^{2a}, e_2^{2a}, \dots, e_R^{2a}) \quad (7.16)$$

$$\Delta_c = \Delta_{CE}(e_1^{2a}, e_2^{2a}, \dots, e_R^{2a}) \quad (7.17)$$

$$\Delta_h = \Delta_H(e_1^{2a}, e_2^{2a}, \dots, e_R^{2a}) \quad (7.18)$$

b) tomando-se um novo conjunto de R amostras candidatas $(e_1^{2b}, e_2^{2b}, \dots, e_R^{2b})$, faz-se $(e_1^2, e_2^2, \dots, e_R^2) = (e_1^{2b}, e_2^{2b}, \dots, e_R^{2b})$ se e somente se:

$$\begin{aligned} \Delta_{CE}(e_1^{2b}, e_2^{2b}, \dots, e_R^{2b}) < \Delta_c \quad \wedge \quad \Delta_{DE}(e_1^{2b}, e_2^{2b}, \dots, e_R^{2b}) > \Delta_d \\ \wedge \quad \Delta_{HE}(e_1^{2b}, e_2^{2b}, \dots, e_R^{2b}) > \Delta_h \end{aligned} \quad (7.19)$$

ou

$$\Delta_{CE}(e_1^{2b}, e_2^{2b}, \dots, e_R^{2b}) < \Delta_c \quad \wedge \quad \Delta_{DE}(e_1^{2b}, e_2^{2b}, \dots, e_R^{2b}) > \Delta_d \quad (7.20)$$

ou

$$\Delta_{CE}(e_1^{2b}, e_2^{2b}, \dots, e_R^{2b}) < \Delta_c \quad \wedge \quad \Delta_{HE}(e_1^{2b}, e_2^{2b}, \dots, e_R^{2b}) > \Delta_h \quad (7.21)$$

ou

$$\Delta_{DE}(e_1^{2b}, e_2^{2b}, \dots, e_R^{2b}) > \Delta_d \wedge \Delta_{HE}(e_1^{2b}, e_2^{2b}, \dots, e_R^{2b}) > \Delta_h \quad (7.22)$$

c) caso $(e_1^2, e_2^2, \dots, e_R^2) = (e_1^{2b}, e_2^{2b}, \dots, e_R^{2b})$ faz-se $(e_1^{2a}, e_2^{2a}, \dots, e_R^{2a}) = (e_1^{2b}, e_2^{2b}, \dots, e_R^{2b})$ e altera-se os limiares $(\Delta_c, \Delta_d, \Delta_h)$ do seguinte modo:

c1) Se (7.19) é satisfeita então:

$$\Delta_d = \Delta_{DE}(e_1^{2b}, e_2^{2b}, \dots, e_R^{2b}) \quad (7.23)$$

$$\Delta_c = \Delta_{CE}(e_1^{2b}, e_2^{2b}, \dots, e_R^{2b}) \quad (7.24)$$

$$\Delta_h = \Delta_H(e_1^{2b}, e_2^{2b}, \dots, e_R^{2b}). \quad (7.25)$$

c2) Se (7.20) é satisfeita então:

Δ_d é dado por (7.23) e

Δ_c é dado por (7.24).

c3) Se (7.21) é satisfeita então:

Δ_c é dado por (7.24) e

Δ_h é dado por (7.25).

c4) Se (7.22) é satisfeita então:

Δ_d é dado por (7.23) e

Δ_h é dado por (7.25).

d) caso $(e_1^2, e_2^2, \dots, e_R^2) \neq (e_1^{2b}, e_2^{2b}, \dots, e_R^{2b})$ (ou seja, a condição do item anterior não é satisfeita) descarta-se $(e_1^{2b}, e_2^{2b}, \dots, e_R^{2b})$, continua-se com $(e_1^2, e_2^2, \dots, e_R^2) = (e_1^{2a}, e_2^{2a}, \dots, e_R^{2a})$ e escolhe-se um novo conjunto de amostra ao qual se atribui a nomenclatura $(e_1^{2b}, e_2^{2b}, \dots, e_R^{2b})$ e

e) repete-se os passos b), c) e d) até que todo conjunto das R amostras tiverem sido testadas.

Este tipo de abordagem depende não só do conjunto de inicialização com \mathbb{k} amostras como também depende da ordem da disposição da seqüência das \mathbb{k} amostras que será testada.

Na Seção 8.4 são mostrados os testes do processo de busca dos MR utilizando-se múltiplas medidas, em especial $(e^3_1, e^3_2, \dots, e^3_R)$, e também $(e_{d_1}, e_{d_2}, \dots, e_{d_R})$, $(e_{c_1}, e_{c_2}, \dots, e_{c_R})$ e $(e_{h_1}, e_{h_2}, \dots, e_{h_R})$.

7.4 Estrutura geral do processo de busca dos MR

A Fig. 7.1 mostra um diagrama de blocos do processo de busca dos MR a partir das amostras candidatas a MR, selecionadas pelo processo de triagem.

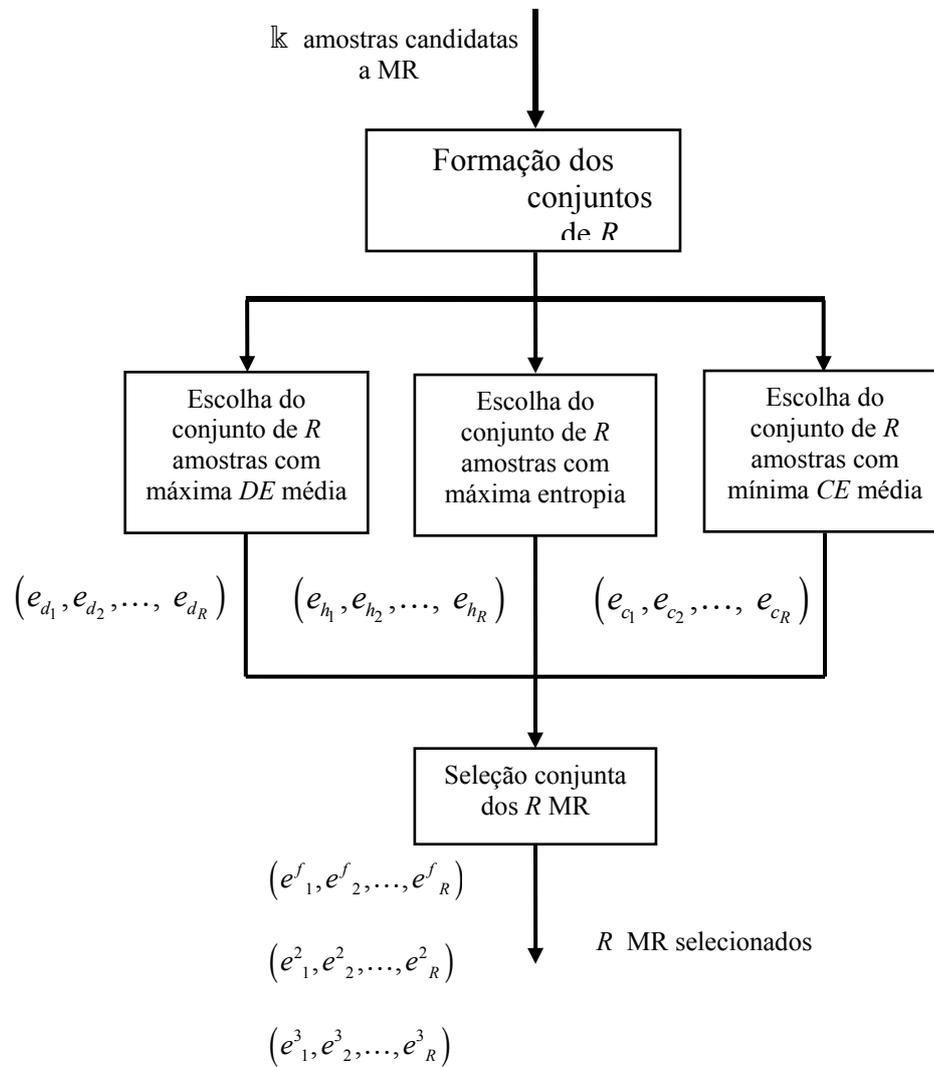


FIGURA 7.1 – Seleção dos MR.

8 Teste de Desempenho e Análise de Resultados

Neste capítulo apresentam-se os testes, as aplicações e as análises relativas à metodologia proposta para a triagem de amostras, para o condicionamento dos espectros e para a seleção dos MR, utilizando-se as imagens do sensor AVIRIS descritas no Capítulo 3. Em algumas situações utiliza-se como medida de desempenho o valor médio da imagem-erro obtida pelo MLME. Essa imagem-erro foi obtida, em todos os casos, utilizando-se os espectros médios originais das imagens hiperespectrias, escolhidos como MR, sem nenhum tipo de condicionamento, como aqueles citados no Capítulo 5. O uso do condicionamento espectral foi utilizado somente no processo de busca dos MR, não sendo utilizado nem no processo de triagem discriminatória e nem no MLME.

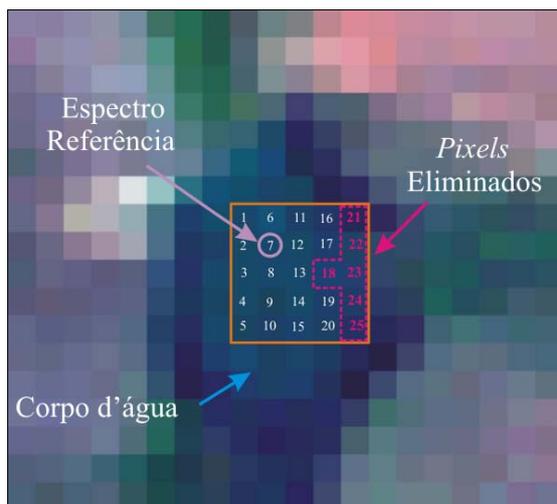
8.1 Teste da triagem discriminatória

Nesta seção apresentam-se e analisam-se alguns resultados do processo de triagem discriminatória descrito no Capítulo 4.

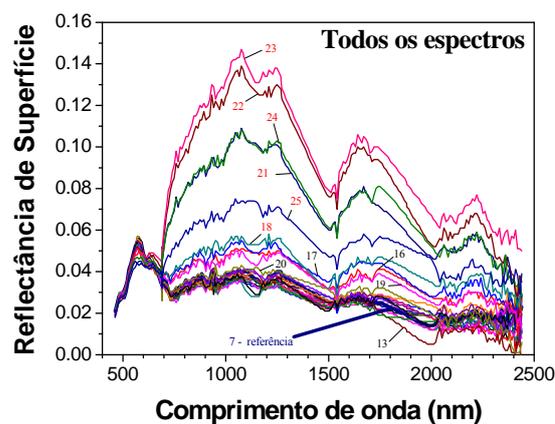
8.1.1 Amostra d'água contaminada

Realizou-se inicialmente o teste da uniformidade espacial considerando-se o limiar de seleção espacial $\psi_e = 78\%$ e uma vizinhança $(M + 1) \times (M + 1) = 5 \times 5$ pixels em torno de uma amostra de água contaminada por vegetação verde (vv1: mata galeria) presente na subcena 7, da imagem do sensor AVIRIS da área teste. A Fig. 8.1a mostra uma pequena porção, 20 x 18 pixels da imagem subcena 7, onde está demarcada a vizinhança com 5 x 5 pixels centralizada na amostra d'água observada. Nesta figura o pixel indicado pelo número sete teve o seu espectro médio utilizado como referência no cálculo da correlação entre os espectros, conforme descrição realizada na seção 4.2.

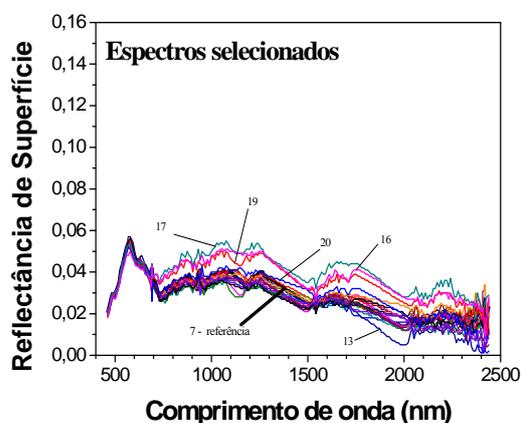
Na Fig. 8.1a os pixels em vermelho (18, 21, 22, 23 e 24) são dos pixels rejeitados devido a mistura, presente na margem do corpo d'água, do qual a amostra teste foi considerada. Observa-se ainda que o pixel 7, que teve o seu espectro escolhido com referência, pela metodologia apresentada na seção 4.2, está afastado dos pixels misturados presentes em uma borda da vizinhança.



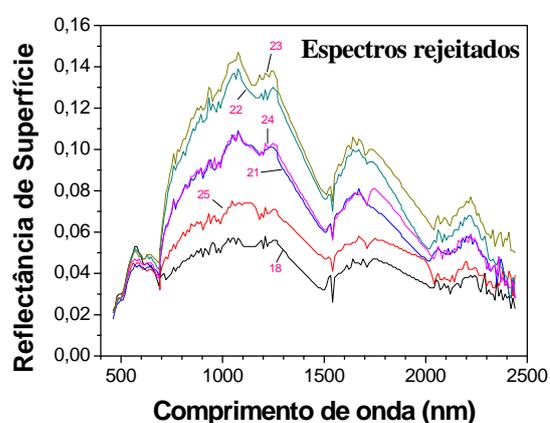
(a)



(b)



(c)



(d)

FIGURA 8.1 – Teste da uniformidade espacial. (a): corpo da água pertencente a subcena 7 onde estão indicados os limites da vizinhança 5×5 pixels em torno de uma amostra espectral relativa a água. (b) espectros contidos na vizinhança da amostra espectral água. (c) 19 espectros escolhidos como pertencentes a uma vizinhança uniforme. (d) espectros rejeitados por serem considerados contaminados.

Na Fig. 8.1b estão apresentados os 25 espectros da vizinhança 5×5 pixels em torno de amostra de água considerada. A Fig. 8.1c mostra os 19 espectros selecionados, 76% de 25 possíveis, pelo processo de verificação da uniformidade espacial. Na Fig. 8.1d estão os espectros das amostras que foram rejeitadas no teste da uniformidade espacial. Se for considerado o fator de uniformidade espacial $\alpha_u = 76\%$ esta amostra, com 19 espectros selecionados, será considerada uma amostra com vizinhança uniforme.

Para esta amostra de água com mistura o teste de homogeneidade (seção 4.3), aplicado na vizinhança $V(i_k, j_k)$ de 19 amostras, consideradas uniformes espacialmente, produziu um indicador de homogeneidade $Q_h(i_k, j_k) = 99\%$.

Considerou-se ainda três amostras d'água, identificadas por $a(a)$, $a(b)$ e $a(c)$, contaminadas com vegetação verde, conforme ilustração da Fig. 8.2a. A amostra $a(c)$ é a amostra na qual foi feito o teste de uniformidade espacial representado pela Fig 8.1.

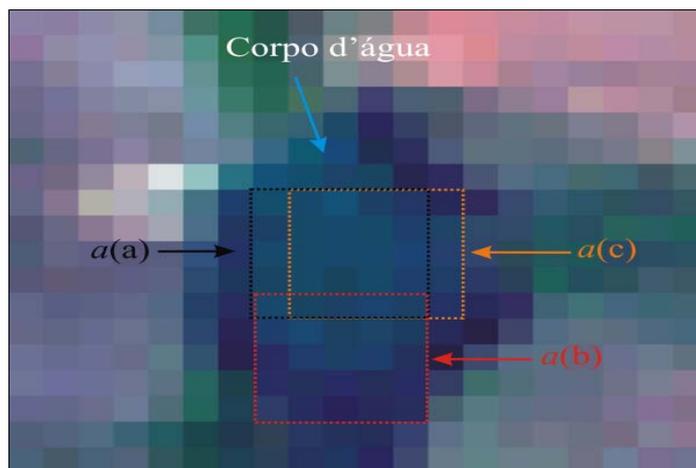
No teste com as três amostras de água considerou-se uma vizinhança $(M+1) \times (M+1) = 5 \times 5$ pixels e os seguintes parâmetros: fator de uniformidade espacial $\alpha_u = 90\%$, limiar de seleção espacial $\psi_e = 78\%$, fator de homogeneidade $\alpha_h = 90\%$, nível de significância $\alpha = 10\%$ no teste de igualdade das médias, limiar de redundância para a medida DE $\psi_{rde} = 5\%$ e limiar de redundância para a medida CE $\psi_{rce} = 5\%$.

Aplicando-se os testes de uniformidade espacial, homogeneidade espectral e redundância, considerando-se o conjunto intersecção \mathfrak{R}_\cup (Eq. 4.37), obtiveram-se os seguintes valores do indicador de homogeneidade $Q_h(i_k, j_k)$:

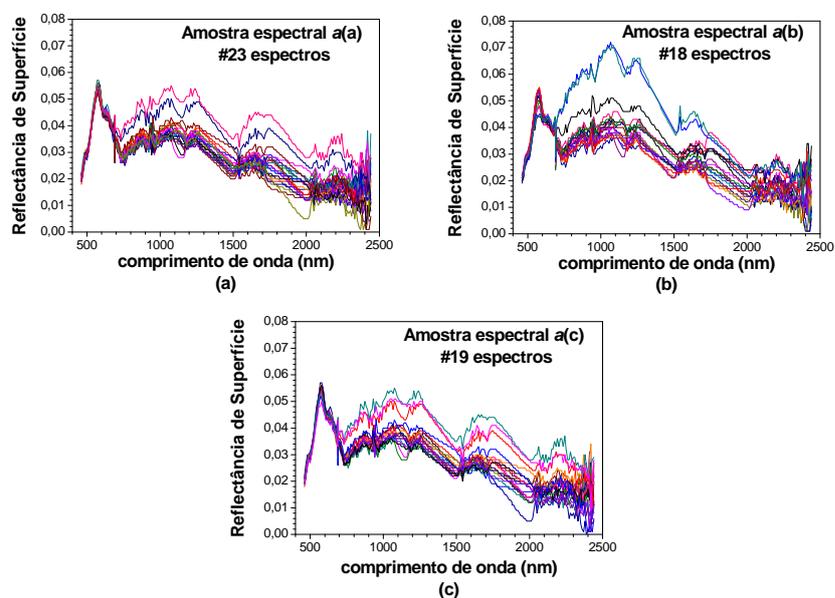
- a) Amostra $a(a)$: nova vizinhança uniforme composta por 23 pixels considerada homogênea pois $Q_h(i_k, j_k) = 97,6\% > \psi_h$;
- b) Amostra $a(b)$: nova vizinhança uniforme composta por 18 pixels considerada homogênea pois $Q_h(i_k, j_k) = 100\% > \psi_h$ e considerada redundante com $a(a)$ e
- c) Amostra $a(c)$: nova vizinhança uniforme composta por 19 pixels considerada homogênea pois $Q_h(i_k, j_k) = 100\% > \psi_h$ e considerada redundante com $a(a)$.

Deste modo as amostras $a(b)$ e $a(c)$ foram rejeitadas por serem consideradas redundantes com a amostra $a(a)$.

Observa-se na Fig. 8.2a que, de fato, existe uma grande superposição entre as vizinhanças iniciais 5×5 das amostras $a(a)$ e $a(c)$. Observa-se também nesta figura que a amostra $a(a)$ é a amostra mais centrada no corpo d'água e, portanto, é a amostra com vizinhança menos contaminada. Isto concorda com o fato de apenas dois espectros da sua vizinhança 5×5 terem sido rejeitados no teste da uniformidade espacial.



(a) vizinhanças de três amostras de água.



(b) espectros resultantes após o teste da uniformidade espacial.

FIGURA 8.2 – Teste da uniformidade espacial relativo à vizinhança de três amostras de água contaminadas por vegetação verde.

8.1.2 Amostras selecionadas manualmente e aleatoriamente

Aplicou-se a metodologia de triagem discriminatória, descrita no Capítulo 4, em 76 amostras escolhidas manualmente por um especialista, na subcena 4 das imagens AVIRIS, mostradas na Fig. 8.3a. Esta escolha foi descrita no item 4.1.1 e resumida na Tabela 4.1. Estas amostras foram divididas, pelo especialista, nas seguintes classes e subclasse:

- $vv1$ (6 amostras), $vv2$ (5 amostras), $vv3$ (5 amostras), $vv4$ (5 amostras), $vv5$ (5 amostras);
- vs (6 amostras);
- a (3 amostras);
- $LV1$ (6 amostras), $LV2$ (5 amostras), $LV3$ (5 amostras);
- $RQ1$ (7 amostras), $RQ2$ (5 amostras), $RQ3$ (5 amostras);
- misturas de classe (4 amostras), e
- espectros com falhas em algumas das suas bandas (4 amostras).

As amostras com misturas de classe e com falhas nas bandas foram adicionadas com o propósito de se testar o desempenho do processo de triagem.

Considerando-se uma vizinhança inicial $(M + 1) \times (M + 1) = 5 \times 5$ pixels, fator de uniformidade espacial $\alpha_u = 60\%$, limiar de seleção espacial $\psi_e = 78\%$, fator de homogeneidade $\alpha_h = 90\%$, nível de significância $\alpha = 10\%$ no teste de igualdade das médias, limiar de redundância para a medida DE $\psi_{rde} = 5\%$, limiar de redundância para a medida CE $\psi_{rce} = 5\%$ e o conjunto união \mathfrak{R}_\cup obteve-se os seguintes resultados no processo de triagem proposto:

- Amostras iniciais: $K = 76$;
- Amostras rejeitadas pelo teste da uniformidade espacial: $4 \Rightarrow KU = 72$;
- Amostras rejeitadas pelo teste da homogeneidade espectral: $24 \Rightarrow KH = 48$;
- Amostras rejeitadas pelo teste da redundância: $20 \Rightarrow KR = 28$.

Das 28 amostras resultantes estavam presentes as seguintes classes e subclasses:

- $vv1$ (2 amostras), $vv2$ (1 amostra), $vv3$ (3 amostra), $vv4$ (2 amostra), $vv5$ (3 amostras);

- *vs* (4 amostras);
- *a* (3 amostras);
- *LV1* (1 amostras), *LV2* (2 amostras), *LV3* (2 amostras);
- *RQ1* (2 amostras), *RQ2* (1 amostra), *RQ3* (1 amostra) e
- uma amostra com mistura de classes.

Observa-se que todas as classes e subclasses, escolhidas pelo especialista, foram mantidas no conjunto final de amostras. Nenhuma amostra com falhas no espectro foi mantida e apenas uma amostra com mistura de classes foi mantida no grupo de amostras candidatas a MR. As amostras candidatas a MR escolhidas são mostradas na Fig. 8.3b.

Constatou-se ainda que a amostra da classe mistura que permaneceu teve a sua vizinhança inicial modificada pelo teste da uniformidade espacial de modo que a nova vizinhança eliminou a mistura da classe menos preponderante na vizinhança.

Aplicou-se ainda a metodologia de triagem discriminatória nas 900 amostras escolhidas aleatoriamente, pela metodologia AESNA, mostradas na Fig. 8.3c. Considerando-se uma vizinhança inicial $(M+1) \times (M+1) = (5 \times 5)$, fator de uniformidade espacial $\alpha_u = 60\%$, limiar de seleção espacial $\psi_e = 78\%$, fator de homogeneidade $\alpha_h = 90\%$, nível de significância $\alpha = 10\%$ no teste de igualdade das médias, limiar de redundância para a medida *DE* $\psi_{rde} = 5\%$, limiar de redundância para a medida *CE* $\psi_{rce} = 5\%$ e o conjunto união \mathfrak{R}_\cup obteve-se os seguintes resultados no processo de triagem proposto:

- Amostras iniciais: $K = 900$;
- Amostras rejeitadas pelo teste da uniformidade espacial: $27 \Rightarrow KU = 873$;
- Amostras rejeitadas pelo teste da homogeneidade espectral: $270 \Rightarrow KH = 603$;
- Amostras rejeitadas pelo teste da redundância: $496 \Rightarrow KR = 107$.

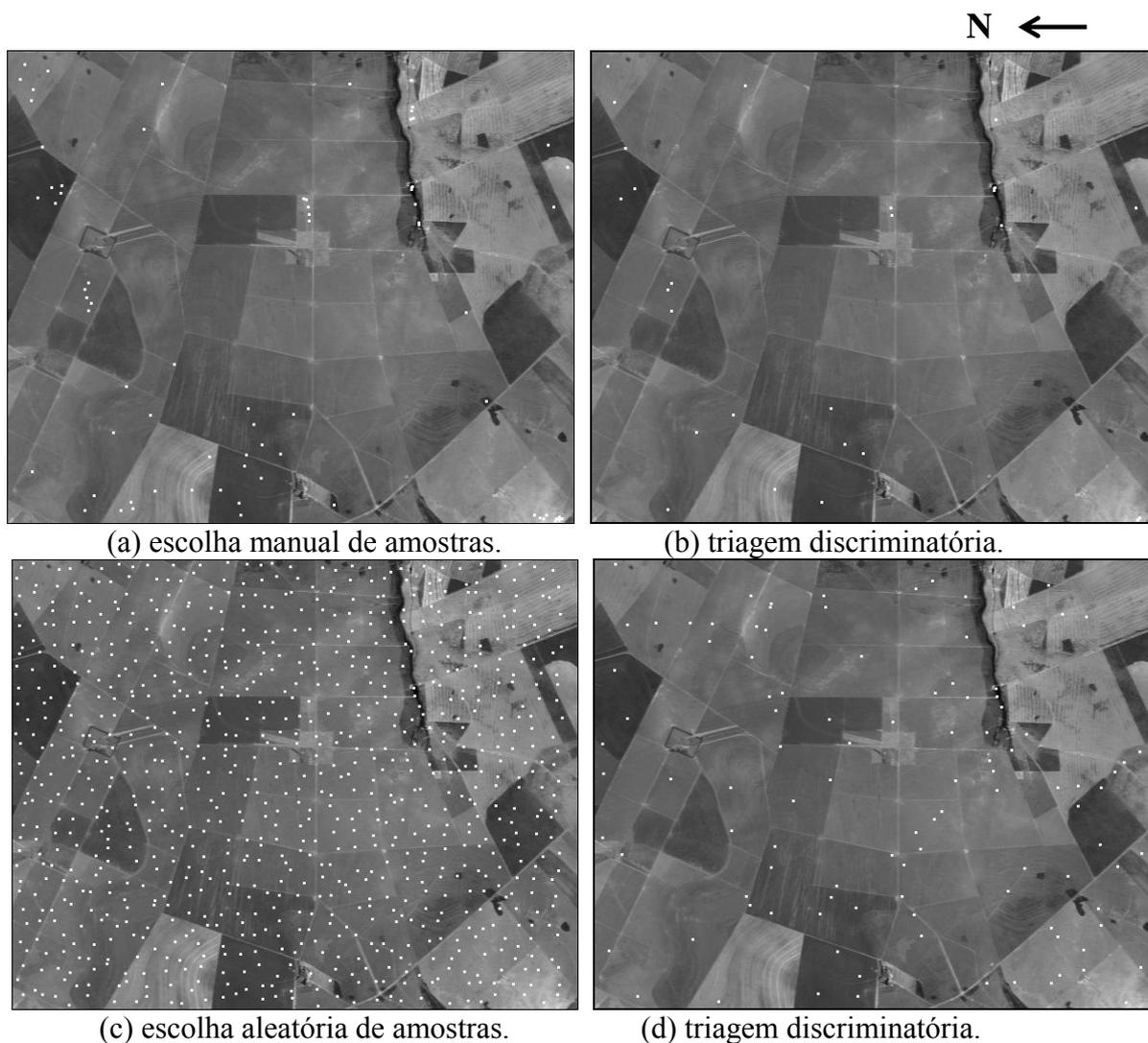


FIGURA 8.3 – Triagem discriminatória de amostras na subcena 4 escolhidas inicialmente pela abordagem aleatória e manual.

Uma análise qualitativa dos resultados do processo de triagem mostra que dentre as 107 amostras candidatas a MR ficaram presentes as seguintes classes e subclasses:

- *vv1* (pelo menos 2 amostras), *vv2* (pelo menos 4 amostras), *vv3* (pelo menos 3 amostras), *vv4* (pelo menos 3 amostras), *vv5* (pelo menos 1 amostra);
- *vs* (pelo menos 2 amostra);
- *a* (não presente);
- *LV1* (pelo menos 1 amostra), *LV2* (pelo menos 1 amostra), *LV3* (pelo menos 1 amostra) e
- *RQ1* (não presente), *RQ2* (não presente), *RQ3* (pelo menos 1 amostra);

Observa-se que somente a classe água não tem nenhum representante. Um dos motivos da ausência da água é a sua diminuta participação na cena relativa a subcena 4, que tem um pequeno corpo de água (riacho) envolto pela mata galeria (vv1). As amostras candidatas a MR escolhidas são mostradas na Fig. 8.3d.

8.1.3 Aplicação da triagem discriminatória na subcena4

Aplicou-se o teste da triagem discriminatória em todos os *pixels* da subcena 4, utilizando-se neste processo uma vizinhança inicial $(M+1) \times (M+1) = 5 \times 5$ *pixels*, fator de uniformidade espacial $\alpha_u = 60\%$, limiar de seleção espacial $\psi_e = 78\%$, fator de homogeneidade $\alpha_h = 90\%$ e nível de significância $\alpha = 10\%$ no teste de igualdade das médias. Para o teste do contexto considerou-se uma vizinhança $(M_c+1) \times (M_c+1) = 5 \times 5$ *pixels* e um fator de contexto $\alpha_c = 80\%$. Considerou-se ainda o limiar de redundância para a medida *DE* como $\psi_{rde} = 71,26 \times 10^{-7}$, o limiar de redundância para a medida *CE* como $\psi_{rce} = 14,36 \times 10^{-5}$.

A Fig. 8.4a mostra a banda 192 (2.271 nm) da subcena 4 e a Fig. 8.4b mostra a imagem da cardinalidade $\#V(i_k, j_k)$ da vizinhança $V(i_k, j_k)$ das amostras, normalizada pelo número de amostras presentes na vizinhança original, ou seja, mostra a grandeza $0 \leq \#V(i_k, j_k)/(M+1)^2 \leq 1$, onde em uma escala de cinza o nulo corresponde ao preto e o valor unitário ao branco. Quanto maior é o valor do parâmetro $\#V(i_k, j_k)/(M+1)^2$ mais uniforme é a vizinhança. Observa-se que as bordas das classes, por serem regiões de mistura, possuem um valor baixo de $\#V(i_k, j_k)/(M+1)^2$, como era esperado. A cardinalidade normalizada pode desta maneira ser utilizada como um detector de bordas em imagens hiperespectrais.

A Fig 8.4c mostra a imagem do indicador de homogeneidade $0 \leq Q_h(i_k, j_k) \leq 1$, após o teste da uniformidade espacial na imagem, novamente em uma escala de cinza o nulo corresponde ao preto e o valor unitário ao branco. Quanto maior o indicador de homogeneidade mais homogênea é a vizinhança. Nota-se que devido a redefinição da vizinhança, que ficou somente com as amostras consideradas espacialmente uniformes, selecionadas pelo teste da uniformidade espacial, algumas bordas entre classes foram consideradas espectralmente homogêneas, aparecendo claras na imagem da Fig. 8.4c.

A Fig. 8.4d mostra a imagem original, onde em preto estão as amostras rejeitadas pelos testes da uniformidade espacial e homogeneidade espectral. Na Fig. 8.4e mostram-se os aglomerados, em branco, devido o teste do contexto espacial. Estes aglomerados, em branco, são os *pixels* da imagem que passaram pelos testes de uniformidade, homogeneidade e de contexto. A Fig 8.4f mostra estes mesmos aglomerados em tons de amarelo, essa figura foi formada pela composição *Red-Green-Blue* das imagens dos aglomerados formados pelo contexto espacial (*Red*), imagem $\#V(i_k, j_k)/(M+1)^2$ (*Green*) e imagem banda 192 (*Blue*), respectivamente. Estes aglomerados são mostrados ainda na Fig. 8.5b em outra composição de cores RGB: imagens dos aglomerados formados (*Red* e *Blue*) e imagem da banda 192 (*Green*).

No teste de uniformidade espacial foram selecionados 95,77% dos *pixels* da imagem original com 612 x 512 *pixels*. Os testes conjuntos de uniformidade espacial e homogeneidade espectral selecionaram 71,77% dos *pixels* originais. Incluindo-se o teste do contexto espacial foram selecionados 33,75% dos *pixels* originais, que equivalem aos *pixels* em branco na Fig. 8.4e.

Realizando-se o teste da redundância nos *pixels* dos aglomerados, gerados pelo contexto espacial, obteve-se os seguintes resultados: 2,80% dos *pixels* originais foram selecionados pela redundância baseada somente na medida da *DE*, conjunto \mathfrak{R}_{DE} de amostras; 2,51% dos *pixels* originais foram selecionados pela redundância baseada somente na medida da *CE*, conjunto \mathfrak{R}_{CE} de amostras; 5,09% dos *pixels* originais foram selecionados pela redundância baseada na união das amostras obtidas pelas medidas realizadas pela *DE* e *CE*, conjunto \mathfrak{R}_{\cup} de amostras; e 0,22% dos *pixels* originais foram selecionados pela redundância baseada na intersecção das amostras obtidas pelas medidas realizadas pela *DE* e *CE*, conjunto \mathfrak{R}_{\cap} de amostras. Estes resultados são mostrados na Fig. 8.5.



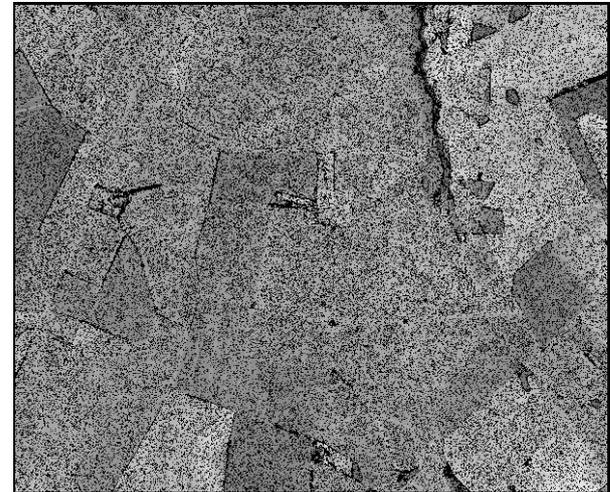
(a) Banda 192 (2.271nm) da sub cena 4.



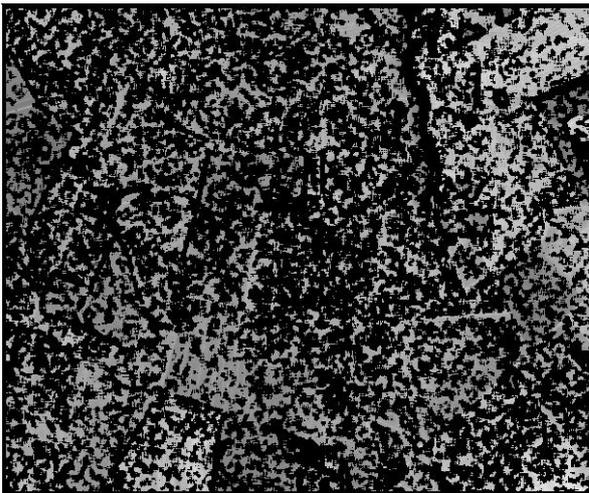
(b) Parâmetro $\#V(i_k, j_k)/(M+1)^2$.



(c) Parâmetro $0 \leq Q_h(i_k, j_k) \leq 1$.



(d) em preto *pixels* rejeitados pela uniformidade e homogeneidade.



(e) em branco os aglomerados aceitos pelo contexto espacial.

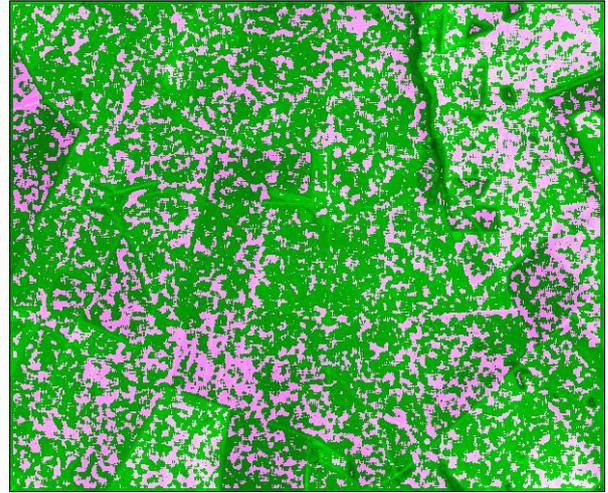


(f) em tons amarelo *pixels* aceitos no contexto espacial.

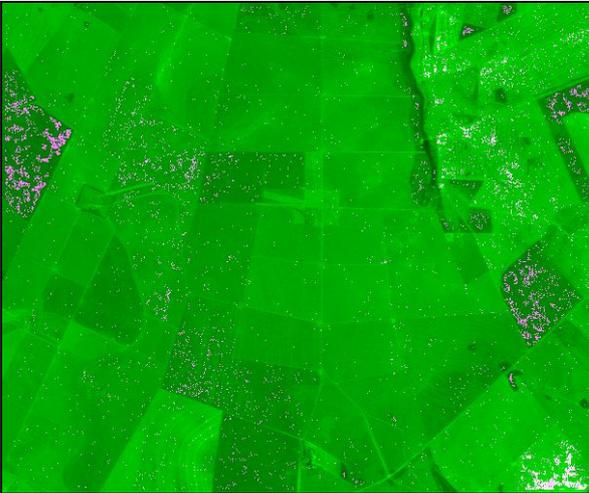
FIGURA 8.4 – Triagem da subcena 4 quanto a uniformidade, homogeneidade e contexto.



(a) Banda 192 (2271 nm) da sub cena 4.



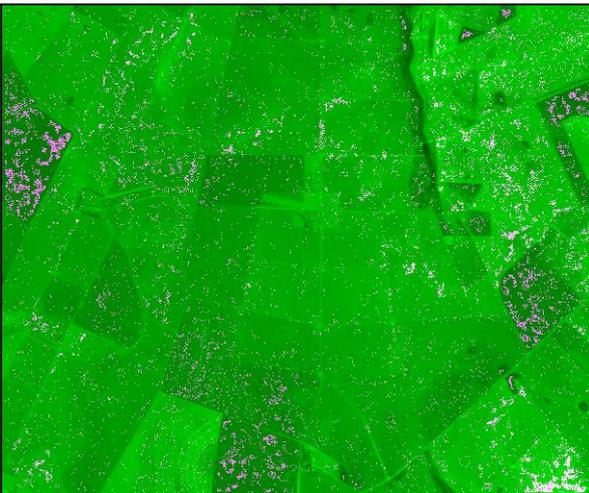
(b) em tons róseo *pixels* aceitos pelo contexto.



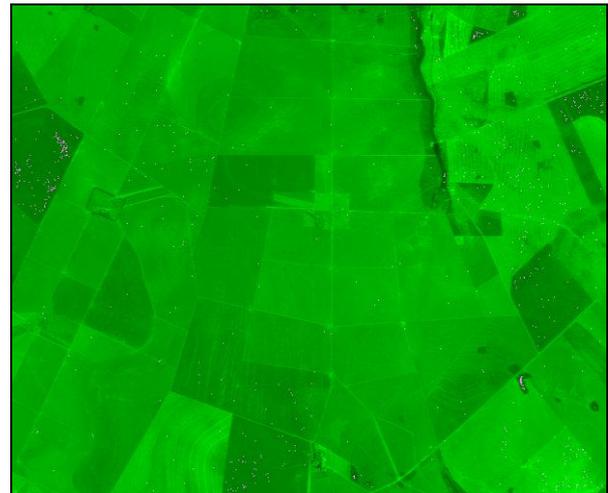
(c) menos redundantes pertencentes a \mathfrak{R}_{DE} .



(d) menos redundantes pertencentes a \mathfrak{R}_{CE} .



(e) menos redundantes pertencentes a \mathfrak{R}_U .



(f) menos redundantes pertencentes a \mathfrak{R}_O .

FIGURA 8.5 – Triagem da subcena 4 pelo critério da redundância aplicada às amostras selecionadas pelo contexto espacial.

A Figura 8.6 mostra na cor branca a estrutura dos aglomerados de *pixels* selecionados pelo contexto, após o teste da uniformidade espacial e da homogeneidade espectral. Nestas figuras foram utilizados diferentes fatores de contexto α_c para uma vizinhança $(M_c + 1) \times (M_c + 1) = 5 \times 5$ *pixels*. Como descrito na seção 4.4 quanto menor for este fator maior serão os aglomerados.

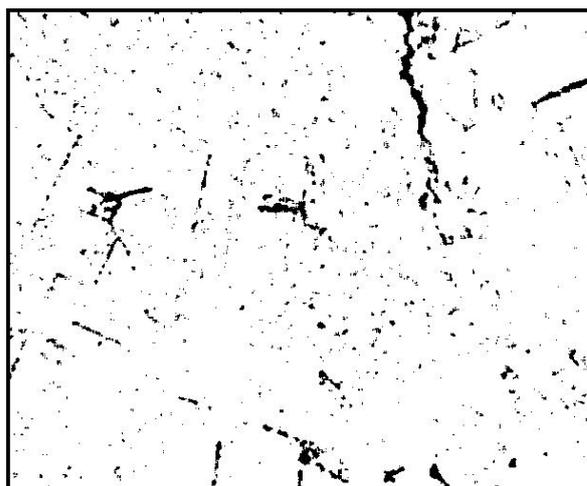
8.2 Busca dos MR pela entropia

Apresenta-se nesta seção uma avaliação da busca dos MR pelo método da máxima entropia proposto no Capítulo 6. Faz-se uma análise do método de busca utilizando-se um conjunto controlado de amostras espectrais candidatas a MR, escolhidas por um especialista na subcena 4 e aplica-se também o método de busca às amostras candidatas a MR escolhidas pela metodologia AESNA na subcena 5.

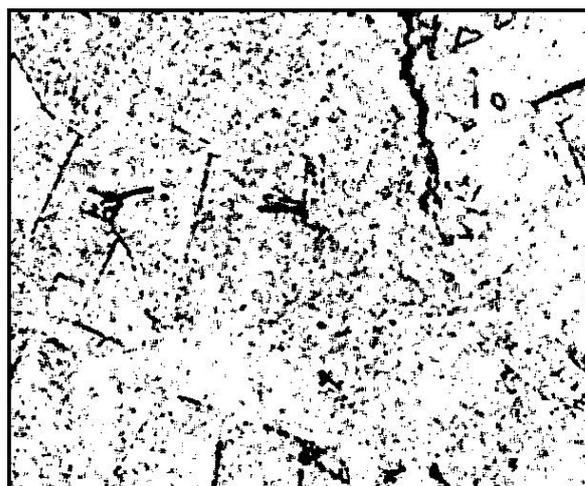
8.2.1 Teste do método de busca dos MR pela máxima entropia na subcena 4

Para o teste da seleção dos MR pelo método da entropia, proposto no Capítulo 6, foram utilizados espectros conhecidos, extraídos de dados hiperespectrais AVIRIS da subcena 4.

Na Fig. 8.7 mostra-se uma das imagens, banda 192 (2.271 nm) da subcena 4 obtida pelo sensor AVIRIS, com a localização, aproximada, dos *pixels* de 14 amostras espectrais ($k = 14$) candidatas a MR e que poderão representar classes e/ou subclasses de alvos. Cinco amostras foram escolhidas pelo especialista com o conhecimento a priori da área, essas amostras representam as classes: vegetação verde (*vv*), vegetação com pouca atividade fotossintética (*vs*), Latossolo Vermelho (*LV*), Neossolo Quartzarênico (*RQ*) e água (*a*). Adicionalmente a estas cinco amostras foram escolhidas pela metodologia AESNA mais nove amostras espectrais na cena que foram rotuladas como: *amo1*, *amo2*, ..., *amo9*.



(a) aglomerados para $\alpha_c = 50\%$.



(b) aglomerados para $\alpha_c = 60\%$.



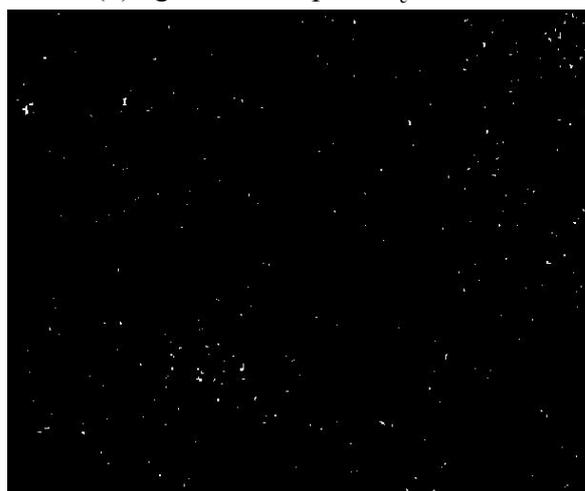
(c) aglomerados para $\alpha_c = 70\%$.



(d) aglomerados para $\alpha_c = 80\%$.

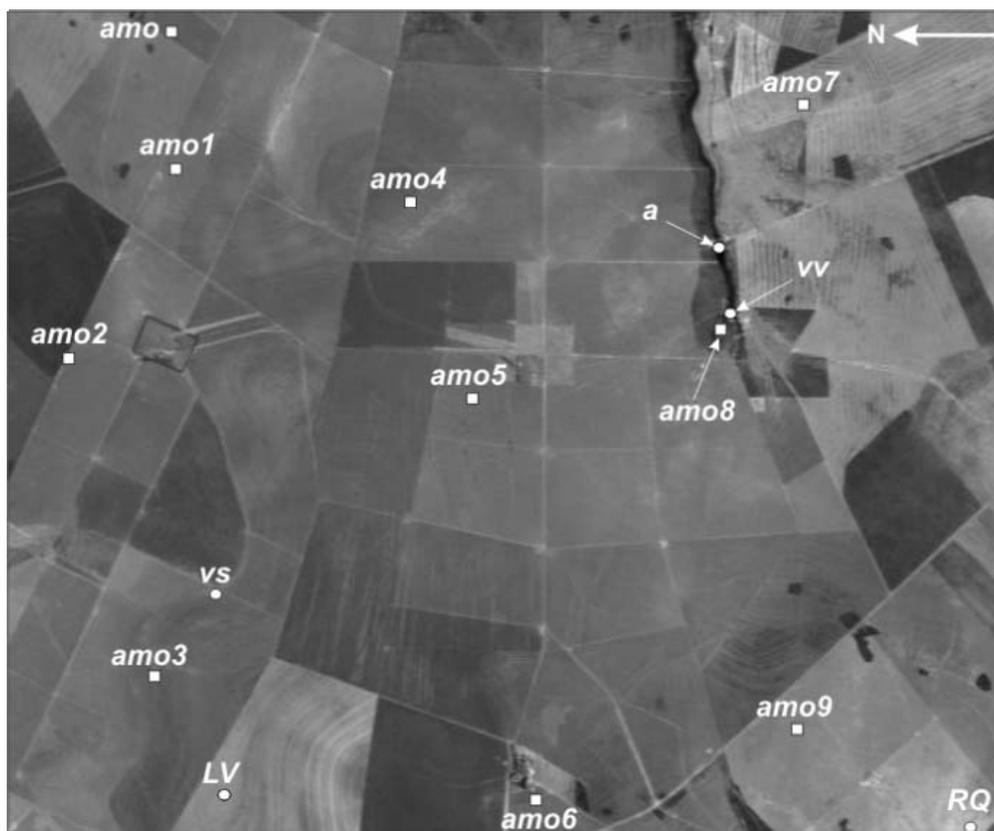


(e) aglomerados para $\alpha_c = 90\%$.



(f) aglomerados para $\alpha_c = 100\%$.

FIGURA 8.6 – Aglomerados resultante do contexto espacial.



- amostras escolhidas aleatoriamente
- amostras escolhidas por especialista

FIGURA 8.7 - Imagem da subcena 4, banda 192 (2.271 nm), com as amostras espectrais utilizadas nos testes de busca de MR.

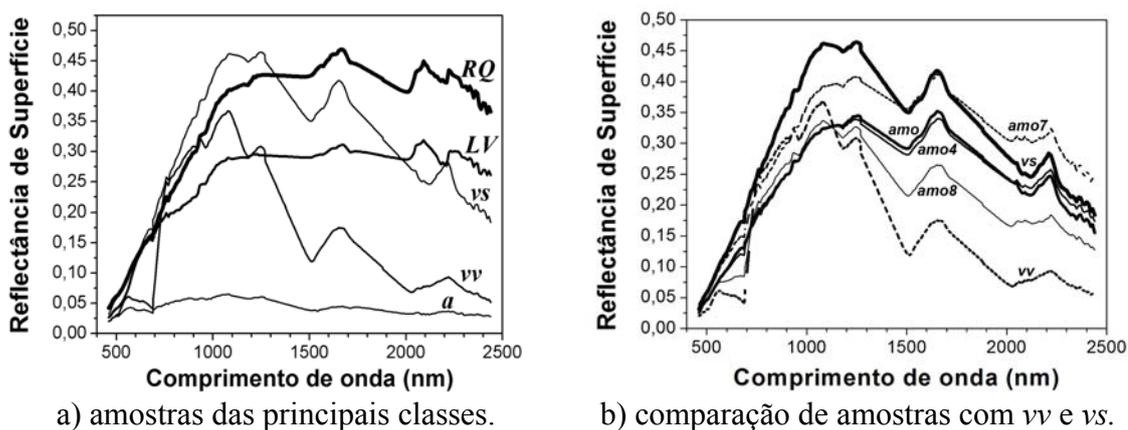


FIGURA 8.8 - Espectros de amostras candidatas a MR da subcena 4.

A Figs. 8.8a e 8.8b mostram os vetores espectros médios $\vec{I}_m(i_k, j_k)$, calculado por (6.1), relativo a uma vizinhança com uniformidade espacial e homogeneidade espectral, para cada classe de alvo escolhido pelo especialista. No cálculo do espectro médio considerou-se

uma vizinhança de $(M+1) \times (M+1) = 5 \times 5 \text{ pixels}$. A Fig. 8.8b mostra outros espectros relevantes para a análise de resultados que será realizada.

Salienta-se que as cinco classes (*vv*, *vs*, *a*, *LV*, *RQ*) foram consideradas pelo especialista como bem representativas da variabilidade espectral na subcena 4 e, assim, os seus espectros configuram uma boa escolha de MR.

Para a busca dos MR utilizando-se a máxima entropia utilizou-se derivada como modo de condicionamento dos espectros das amostras candidatas a MR e calculou-se a média dos espectros condicionados, conforme descrito na seção 5.1, considerando-se uma vizinhança $(M+1) \times (M+1) = 5 \times 5 \text{ pixels}$ da amostra que será testada, sem a utilização da triagem discriminatória relativa aos testes da uniformidade espacial, homogeneidade espectral, contexto e redundância. Foram escolhidos os fatores de configuração para as medidas *DE*, *CE* e entropia: $\alpha_{DE} = \alpha_H = \alpha_{CE} = 25\%$, que corresponderam aos seguintes limiares de configuração: $\eta_{DE} = 547,87$, $\eta_{CE} = 0,98$ e $\eta_H = 0,052$, conforme disposto na seção 6.5.

Obteve-se pelo método da entropia, expresso por (6.21) os MR listados na Tabela 8.1. Nessa Tabela foi suprimido o prefixo *amo* das amostras da AESNA.

Na Tabela 8.1 nota-se que com o aumento de *R* as amostras escolhidas anteriormente como MR vão se mantendo. Para $R = 9$ todas as combinações das $\mathbb{k} = 14$ amostras, agrupadas em conjuntos de 9 amostras, foram consideradas mal configuradas e, deste modo, não houve escolha de nenhum MR. Logo, o limitante superior para o número de MR da cena é $\mathfrak{R}_1 = 8$, conforme descrito na seção 6.4. Para $(2 \leq R \leq 5)$ obteve-se como resposta todos os espectros considerados bons MR pelo especialista.

Para $R = 6$ foi escolhida como MR a amostra rotulada como *amo4*. Analisando-se, o espectro médio dessa amostra (Fig. 6.8b) conclui-se que ela pode ser considerada uma outra espécie de vegetação seca, distinta de *vs*. Para $R = 7$, além de *amo4*, foi escolhida *amo8*. Por inspeção (Fig. 8.8b) o espectro da *amo8* pode ser considerada um tipo de vegetação verde distinto de *vv*. Para $R = 8$, além da *amo4* e *amo8*, foi selecionada a amostra *amo7*, que pode ser também considerada um outro tipo de vegetação seca distinto de *vs* e de *amo4*.

TABELA 8.1 - MR selecionados pela entropia na subcena 4.

MR (R):	Amostras selecionadas como MR:	Entropia	Imagem-erro	
			média	desvio padrão
2	<i>a, LV</i>	0,91	33,75	8,86
3	<i>a, LV, vv,</i>	0,80	22,33	5,77
4	<i>a, LV, vv, vs</i>	0,71	6,93	1,98
5	<i>a, LV, vv, vs, RQ</i>	0,63	5,13	1,47
6	<i>a, LV, vv, vs, RQ, 4</i>	0,54	3,12	0,90
$0,47 < H_{\min} \Rightarrow \mathfrak{R}_2 = 6$				
7	<i>a, LV, vv, vs, RQ, 4, 8</i>	0,47	2,93	0,81
8	<i>a, LV, vv, vs, RQ, 4, 8, 7</i>	0,42	2,74	0,75
9	Nenhuma amostra foi selecionada $\Rightarrow \mathfrak{R}_1 = 8$			

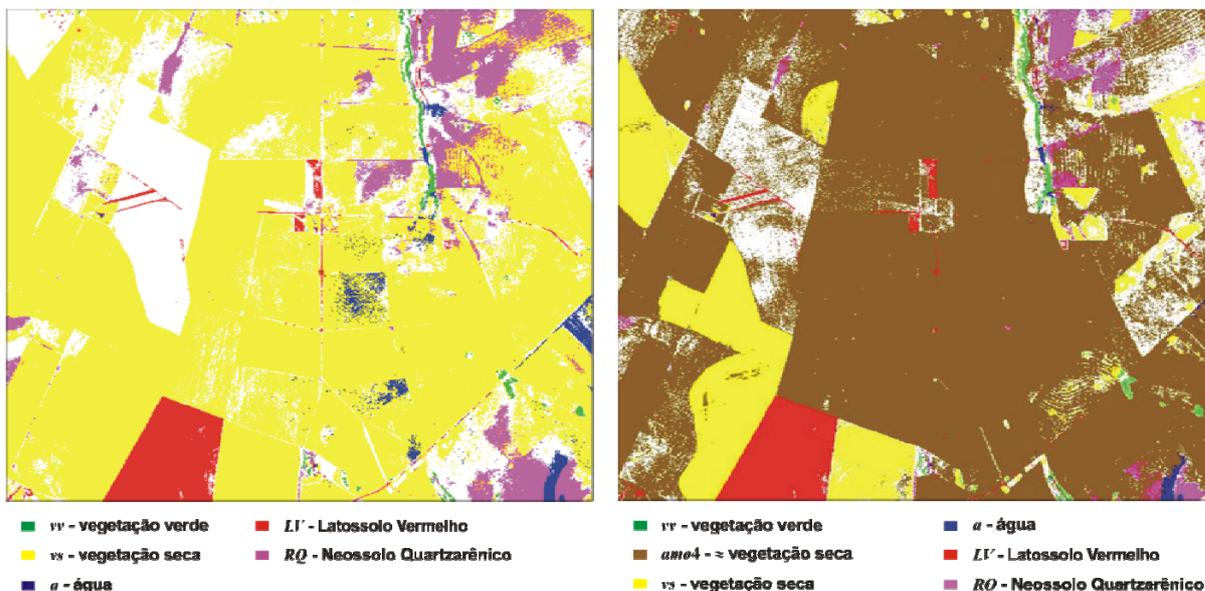
Quanto à escolha final do número de MR, a ser utilizado em um processo de classificação, dependerá da finalidade da classificação. Se houver interesse na classificação de estágios intermediários entre vegetação verde e seca então $R = 6, 7$ ou 8 podem ser recomendados.

A Tabela 8.1 mostra também os valores da entropia, da média e do desvio padrão da imagem-erro, calculados pelo MLME, utilizando-se os MR selecionados. Nota-se que para $R = 4$ a seleção de *vs* como MR fez o erro diminuir consideravelmente, isto porque *vs* é uma classe com alta presença espacial na cena.

De acordo com a seção 6.4, se for estabelecido que o valor mínimo da Entropia é $H_{\min} = 0,50$, então o limitante superior para o número de MR é $\mathfrak{R}_2 = 6$.

As Figs. 8.9a e 8.9b mostram as imagens pós-classificadas da subcena 4, para $R = 5$ e 6 , utilizando-se os MR dispostos na Tabela 8.1. Na imagem pós-classificada são atribuídos aos *pixels* da imagem as classes dos MR que tem mais de 50% na composição espectral do *pixel* segundo o MLME.

Na Fig. 8.9a percebe-se a predominância da classe vegetação seca (correspondente a amostra *vs*). Como *amo4* é um tipo diferente de vegetação seca observa-se na Fig. 8.9b a subdivisão da classe vegetação seca em duas subclasses, uma correspondente a amostra *vs* e a outra a amostra *amo4*. Nota-se ainda que para $R = 6$ o erro na classificação da água (mostrada em azul) foi menor que para o caso de $R = 5$. De modo geral o resultado da classificação com $R = 6$ foi considerada pelo especialista como o mais satisfatório.



a) subcena 4 classificada com cinco MR.

b) subcena 4 classificada com seis MR.

FIGURA 8.9 - Imagens pós-classificadas da subcena 4 com $R = 5$ e 6 . Em cores as classes com *pixels* tendo valores de abundância superiores a 50% da fração total e em branco as áreas não-classificadas (intensa mistura espectral).

Para complementar o teste realizado escolheu-se ainda um novo conjunto de 49 amostras pela AESNA, além das amostras (vv , vs , a , LV , RQ), selecionadas pelo especialista, totalizando assim um conjunto de $k = 54$ amostras candidatas a MR. Novamente não foram aplicados os testes de uniformidade espacial, homogeneidade espectral, contexto e redundância. Para os limiares de configuração escolheu-se: $\alpha_{DE} = 32\%$ e $\alpha_{CE} = \alpha_H = 30\%$.

Utilizando-se novamente o condicionamento pela derivada, conforme descrito na seção 5.1, com $(M + 1) \times (M + 1) = 7 \times 7$ *pixels*, obteve-se para $R = 5$ o seguinte conjunto de MR: (vv , amo , a , LV , RQ). Analisando-se, na Fig. 8.8b, o espectro da amostra rotulada por amo verifica-se que ela representa um tipo de vegetação seca, muito semelhante a $amo4$. Portanto, com um conjunto maior de amostras candidatas o método escolheu amo no lugar de vs .

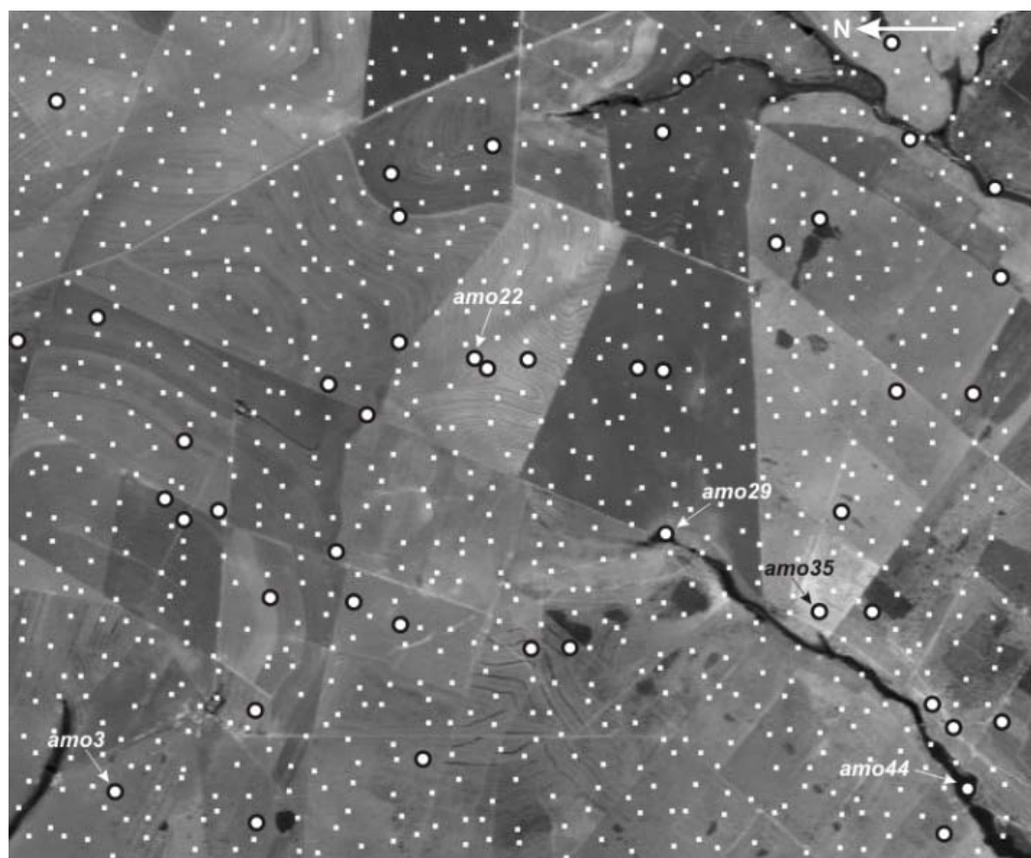
Aplicando-se o MLME para o conjunto (vv , vs , a , LV , RQ) e (vv , amo , a , LV , RQ) gerou-se as imagens-erro para os dois conjuntos de amostras escolhidos como MR. Para o primeiro conjunto obteve-se um erro com média de 5,13 e desvio padrão de 1,47. Para o segundo conjunto de MR, no qual está presente amo no lugar de vs , obteve-se o erro com média de 3,78 e desvio padrão de 1,52. Essas medidas de erro mostram que para a subcena 4

da imagem AVIRIS a escolha da amostra *amo* para MR foi mais acertada que a amostra *vs*, escolhida a priori pelo especialista, tendo-se como critério a média da imagem-erro.

Nos testes realizados observou-se que espectros provenientes d'água por conter baixa relação sinal-ruído foram os mais problemáticos para fins de seleção, pois a reflectância tende aos limites da relação sinal-ruído do sensor.

8.2.2 Aplicação do método de busca pela máxima entropia na subcena 5

O método de busca dos MR pelo método da máxima entropia foi utilizado para a determinação dos MR da subcena 5 dos dados AVIRIS. Escolheu-se pela AESNA $30 \times 30 = 900$ amostras espectrais candidatas a MR. A Fig. 8.10 mostra a posição dessas amostras candidatas a MR na subcena 5.



- 900 amostras escolhidas aleatoriamente.
- 48 amostras selecionadas após os teste de homogeneidade e redundância.

FIGURA 8.10 - Amostras candidatas a MR na subcena 5.

Para se reduzir o número de amostras candidatas a MR utilizou-se a triagem das amostras baseado no teste da homogeneidade espectral (limiar de homogeneidade $\psi_h = 90\%$ e

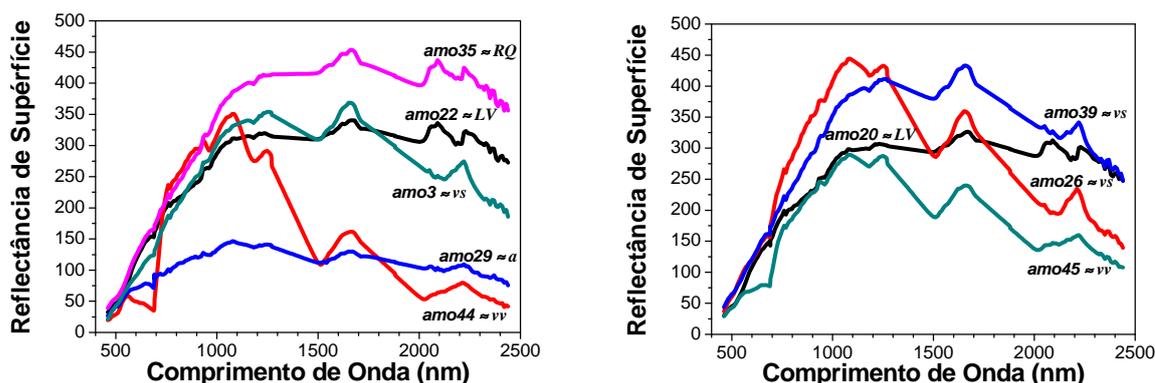
nível de significância $\alpha=10\%$ no teste de igualdade das médias) e teste da redundância (limiares de redundâncias para as medidas DE e CE : $\psi_{rde} = 3\%$ e $\psi_{rce} = 14\%$ e o conjunto intersecção \mathfrak{R}_U para as amostras selecionadas). Foram escolhidos $\alpha_{DE} = \alpha_{CE} = \alpha_H = 25\%$ como fatores de configuração para as medidas DE , CE e entropia. Estes fatores corresponderam aos seguintes limiares de configuração: $\eta_{DE} = 406,09$, $\eta_{CE} = 0,99$, e $\eta_H = 0,045$, conforme disposto na seção 6.5.

Das 900 amostras iniciais o teste de homogeneidade escolheu 727 amostras e destas o teste de redundância escolheu 48 amostras candidatas a MR ($\mathbb{k} = 48$). As amostras candidatas rotuladas como ($amo1, amo2, \dots, amo48$) estão em destaque na Fig. 8.10.

Aplicando-se o condicionamento pela derivada, conforme descrito na seção 5.1 com $(M+1) \times (M+1) = 5 \times 5$ pixels, obteve-se com o cálculo da máxima entropia os MR listados na Tabela 8.2, que também mostra a equivalência das amostras escolhidas com as classes básicas dominantes na subcena 5 (vv, vs, a, LV, RQ). Nessa Tabela foi suprimido o prefixo amo das amostras. Para diminuir o tempo de processamento na escolha dos MR para $R \geq 6$ admitiu-se como MR, escolhidos a priori, os MR escolhidos para $R = 5$. Os espectros dos MR relativos às amostras constantes na Tabela 8.2 são mostrados na Fig. 8.11.

TABELA 8.2 - MR selecionados na subsena 5.

MR (R):	Amostras selecionadas como MR:	Entropia	Imagem-erro	
			média	desvio padrão
2	22, 44 $\approx (LV, vv)$	0,83	18,76	6,84
3	20, 29, 44 $\approx (LV, a, vv)$	0,77	18,36	6,86
4	3, 22, 29, 44 $\approx (vs, LV, a, vv)$	0,65	6,05	2,00
5	3, 22, 29, 35, 44 $\approx (vs, LV, a, RQ, vv)$	0,55	4,35	1,26
6	3, 22, 26, 29, 35, 44 $\approx (vs, LV, vs, a, RQ, vv)$	0,48	3,65	1,08
7	3, 22, 29, 35, 39, 44, 45 $\approx (vs, LV, a, RQ, vs, vv, vv)$	0,42	3,02	0,94
8	3, 22, 26, 29, 35, 39, 44, 45 $\approx (vs, LV, vs, a, RQ, vs, vv, vv)$	0,41	2,70	0,76



a) espectros selecionados para $R = 5$.

b) espectros selecionados para $R \geq 6$.

FIGURA 8.11 - Espectros dos MR escolhidos para a subcena 5 em correspondência com (vv, vs, a, LV, RQ).

Na Fig. 8.11 associa-se as amostras escolhidas (*amo3*, *amo22*, *amo26*, *amo29*, *amo35*, *amo39*, *amo44*, *amo45*) com as classes básicas (vv, vs, a, LV, RQ).

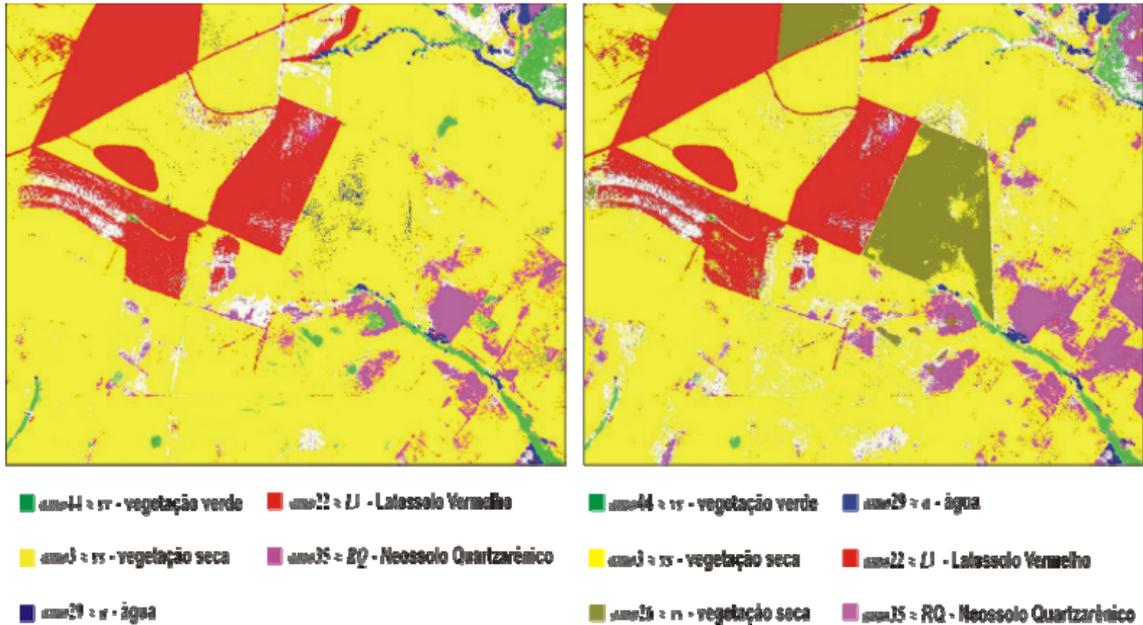
Para diminuir o tempo de processamento na escolha dos MR para $R \geq 6$ admitiu-se como MR, escolhidos a priori, os MR escolhidos para $R = 5$.

Na Tabela 8.2 nota-se que com o aumento de R as classes escolhidas matem a coerência com a escolha anterior. Observa-se ainda que para $R=6$ *amo3* e *amo26* representam dois tipos de amostras próximas a classe vegetação seca. Para $R=7$ foram selecionadas as amostras *amo39* e *amo45* com espectros próximos a vs e a vv, respectivamente.

A Tabela 8.2 mostra também os valores da entropia, da média e do desvio padrão da imagem-erro, calculada pelo MLME utilizando-se os MR listados. Nota-se que a introdução de *amo3* como MR, para $R=4$, o erro diminui consideravelmente, isto porque a *amo3* é um tipo de vegetação seca (vs) que é predominante na área imageada. No caso de $R=6$, 7 e 8 os valores obtidos pela entropia ficaram inferior ao valor mínimo estabelecido ($H_{\min} = 0,50$), assim, o limitante superior de MR, pela entropia mínima, é $\mathfrak{R}_2 = 5$. As imagens pós-classificadas da subcena 5 para $R=5$ e 6 encontram-se na Fig. 8.12.

Observa-se na Fig. 8.12 que, embora, a amostra *amo26* pertença a mesma classe vs da amostra *amo3*, ela contém informações espectrais diferenciadas em relação a esta última.

Deste modo a classificação com $R=6$ foi considerada qualitativamente melhor que a classificação com $R=5$. Este fato contradiz o limitante $\mathfrak{R}_2 = 5$ determinado pela entropia mínima escolhida como $H_{\min} = 0,50$. Qual seria então o melhor número de MR? A resposta vai depender da análise, por parte do especialista, levando em conta as suas necessidades de mapeamento, por ex.: uma maior discriminação de alguns tipos de vegetação seca na imagem.



a) subcena 5 classificada com cinco MR. b) subcena 5 classificada com seis MR.

FIGURA 8.12 - Imagens pós-classificadas da subcena 5 com $R = 5$ e 6 . Em cores as classes com *pixels* tendo valores de abundância superiores a 50% da fração total e em branco as áreas não-classificadas (intensa mistura espectral).

8.3 Condicionamento Espectral

Nesta seção são apresentados e analisados alguns resultados relativo ao processo de condicionamento dos espectros proposto no Capítulo 5.

Para o teste do condicionamento dos espectros foram utilizadas as 76 amostras, $K = 76$, da subcena 4 obtidas pela triagem manual descritas no item 8.1.3 e que aparecem na Fig. 8.3a. As amostras foram divididas em 15 classes/subclasses e submetidas a triagem discriminatória, descrita no Capítulo 4, através dos testes de uniformidade espacial, homogeneidade espectral e redundância, com os seguintes parâmetros: vizinhança inicial $(M + 1) \times (M + 1) = 5 \times 5$ *pixels*, fator de uniformidade espacial $\alpha_u = 60\%$, limiar de seleção

espacial $\psi_e = 78\%$, fator de homogeneidade $\alpha_h = 90\%$, nível de significância $\alpha = 10\%$ para teste de igualdade das médias, limiar de redundância para a medida $DE \psi_{rde} = 3\%$, limiar de redundância para a medida $CE \psi_{rce} = 5\%$ e o conjunto união \mathfrak{R}_U para as amostras selecionadas.

A Tabela 8.3 mostra a disposição das amostras iniciais, em classes/subclasses, e o conjunto de amostras selecionadas em cada uma das fases dos testes da triagem discriminatória.

TABELA 8.3 - Triagem discriminatória de amostras selecionadas na subcena 4.

Classes	Classes e Subclasses	Amostras iniciais	Amostras após o teste da Uniformidade Espacial	Amostras após o teste da teste da Homog. Espectral	Amostras após o teste da Redundância
<i>vv</i>	<i>vv1(a), ..., vv1(f)</i>	6	6	3	<i>vv1(a), vv1(f)</i>
	<i>vv2(a), ..., vv2(e)</i>	5	5	2	<i>vv2(a)</i>
	<i>vv3(a), ..., vv3(e)</i>	5	5	3	<i>vv3(a), vv3(c), vv3(d)</i>
	<i>vv4(a), ..., vv4(e)</i>	5	5	3	<i>vv4(a), vv4(b)</i>
	<i>vv5(a), ..., vv5(e)</i>	5	5	5	<i>vv5(a), vv5(b), vv5(e)</i>
<i>vs</i>	<i>vs(a), ..., vs(f)</i>	6	6	6	<i>vs(b), vs(c), vs(e), vs(f)</i>
<i>a</i>	<i>a(a), a(b), a(c)</i>	3	3	3	<i>a(a), a(b), a(c)</i>
<i>LV</i>	<i>LV1(a), ..., LV1(f)</i>	6	6	4	<i>LV1(e)</i>
	<i>LV2(a), ..., LV2(e)</i>	5	5	4	<i>LV2(a), LV2(e)</i>
	<i>LV3(a), ..., LV3(e)</i>	5	5	5	<i>LV3(b), LV3(c), LV3(e)</i>
<i>RQ</i>	<i>RQ1(a), ..., RQ1(g)</i>	7	7	5	<i>RQ1(a), RQ1(b)</i>
	<i>RQ2(a), ..., RQ2(e)</i>	5	4	1	<i>RQ2(e)</i>
	<i>RQ3(a), ..., RQ3(e)</i>	5	5	3	<i>RQ3(c)</i>
<i>m</i>	<i>m1, ..., m4</i>	4	1	1	<i>m3</i>
<i>anoma</i>	<i>anoma1, ..., anoma4</i>	4	4	nenhuma	nenhuma
Total de amostras		$K = 76$	$K_U = 72$	$K_H = 48$	$\mathbb{k} = K_R = 29$

Após os testes o conjunto inicial $K = 76$ amostras foi reduzido para $\mathbb{k} = K_R = 29$ amostras. Observa-se que as amostras espectrais com espectros com anomalias na sua vizinhança, identificadas com o prefixo *anoma* (falhas de aquisição em algumas bandas espectrais), foram todas rejeitadas e que apenas uma das quatro amostras com mistura (identificadas por $m1$, $m2$, $m3$ e $m4$) permaneceu no conjunto final de amostras. Nota-se também que todas as classes/subclasses significativas estão presentes no conjunto final de amostras candidatas a MR.

Os condicionamentos dos espectros médios de cada amostra espectral candidata a MR, resultantes do processo de triagem, foram feitos utilizando-se a derivada e as TWD relativas as wavelets Haar, Daubechies de ordem 2 (Db2) e Coiflet de ordem 1 e 2 (Coif1 e Coif2). As

diferenças na ordem afetam a regularidade e a simetria das *wavelets* mães, que por sua vez afetam os resultados dos coeficientes da TWD que influenciam o processo de discriminação de alvos, por ex.: entre as classes de cobertura e do solo (KOGER et al., 2003). No Capítulo 5 as Figs. 5.2 a 5.4 mostram os espectros médios originais e condicionados pela derivada e pelas TWD Haar e Db2, das principais classes e subclasses: *vv1*, *vs*, *a*, *LV1* e *RQ1*.

Os principais resultados relativos à busca dos MR, para $R = 4, 5$ e 6 MR, utilizando-se os espectros condicionados são apresentados na Tabela 8.4. Na busca dos MR considerou-se isoladamente as medidas *DE*, *CE* e entropia, e também o processo de múltiplas medidas $(e^3_1, e^3_2, \dots, e^3_R)$ conforme estabelecido na seção 7.3.

A Tabela 8.4 contém também o erro médio e o desvio padrão da imagem-erro calculada pelo MLME, para cada conjunto de MR escolhido. Salienta-se ainda em negrito e em cores as situações em que se obteve os menores erros médios. Os MR escolhidos nestas situações serão objeto de uma análise posterior mais detalhada.

Na elaboração dos dados da Tabela 8.4, para a seleção dos MR, foram utilizados os seguintes fatores de configuração, definidos na seção 6.5: $\alpha_{DE} = 30\%$, $\alpha_{CE} = 40\%$ e $\alpha_H = 40\%$. Os resultados obtidos com o condicionamento feito pela Coif2 não foram satisfatórios e por este motivo não são apresentados na tabela.

Destacam-se em amarelo, na Tabela 8.4, para $R = 5$ e 6 a escolha inadequada das amostras *RQ1(a)* e *RQ3(c)*, pois a subclasse *RQ3* é muito similar a *RQ1* conforme mostrado na Fig. 3.4c. Destacam-se ainda a escolha inconveniente de duas amostras de vegetação seca como MR.

Com exceção dos erros médios da ordem de dezenas, os demais erros médios da imagem-erro, mostrados na Tabela 8.4, não são significativamente diferentes. Isso implica, nestes casos, que o valor do erro médio não pode ser utilizado com um bom discriminador do desempenho dos MR escolhidos.

Na Tabela 8.4 para $R = 4$ sem condicionamento observam-se, com exceção da busca dos MR pela *DE*, grandes valores para os erros médios e desvios padrões associados às imagens-erro. A busca de MR pela *DE* acarretou um menor erro devido ao fato de ter sido

escolhida uma amostra de vegetação seca, $vs(b)$, como MR e está classe de alvo ser a mais abundante na cena.

TABELA 8.4 –MR selecionados na subcena 4.

MR	Tipo de condicionamento	Medidas de similaridade	MR selecionados	Imagem-Erro	
				média	desvio padrão
4	Sem condicionamento	DE	$vv1(a), vs(b), a(a), RQ1(a)$	5,25	1,43
		CE	$vv1(a), a(a), LV3(b), RQ1(a)$	15,90	4,79
		entropia	$vv1(a), a(a), LV3(e), RQ1(a)$	16,03	4,85
		conjunta	$vv1(a), a(a), LV1(e), RQ1(a)$	16,47	5,22
	Derivada	DE	$vv1(a), vs(b), a(a), RQ1(a)$	5,25	1,43
		CE	$vs(b), a(a), LV3(e), RQ1(a)$	6,44	3,03
		entropia	$vv1(a), vs(b), a(a), LV3(e)$	5,27	1,35
		conjunta	$vv1(a), vs(b), a(a), LV3(e)$	5,27	1,35
	Haar	DE	$vv1(a), vs(b), a(a), RQ1(a)$	5,25	1,43
		CE	$vs(b), a(a), LV3(e), RQ1(a)$	6,44	3,03
		entropia	$vv1(a), vs(b), a(a), LV3(e)$	5,27	1,35
		conjunta	$vv1(a), vs(b), a(a), LV3(e)$	5,27	1,35
	Db2 e Coif1	DE	$vv1(a), vs(b), a(a), RQ1(a)$	5,25	1,43
		CE	$vs(e), a(a), LV3(b), RQ1(a)$	6,90	2,47
		entropia	$vv1(a), vs(e), a(a), LV3(b)$	5,96	1,61
		conjunta	$vv1(a), vs(b), a(a), RQ1(a)$	5,25	1,43
5	Sem condicionamento	DE	$vv1(a), vs(b), a(a), RQ1(a), RQ3(c)$	4,73	1,21
		CE	$vv1(a), a(a), LV3(b), RQ1(a), RQ3(c)$	14,25	4,38
		entropia	$vv1(a), vs(b), a(a), LV3(e), RQ1(a)$	4,65	1,23
		conjunta	$vv1(a), vs(b), a(a), LV3(b), RQ1(a)$	4,63	1,23
	derivada	DE	$vv1(a), vs(b), a(a), LV3(e), RQ1(a)$	4,65	1,23
		CE	$vv1(a), vs(b), a(a), LV3(e), RQ1(a)$	4,65	1,23
		entropia	$vv1(a), vs(b), a(a), LV3(e), RQ1(a)$	4,65	1,23
		conjunta	$vv1(a), vs(b), a(a), LV3(e), RQ1(a)$	4,65	1,23
	Haar	DE	$vv1(a), vs(b), a(a), LV3(e), RQ1(a)$	4,65	1,23
		CE	$vv1(a), vs(b), a(a), LV3(e), RQ1(a)$	4,65	1,23
		entropia	$vv1(a), vs(b), a(a), LV3(e), RQ1(a)$	4,65	1,23
		conjunta	$vv1(a), vs(b), a(a), LV3(e), RQ1(a)$	4,65	1,23
	Coif1	DE	$vv1(a), vs(b), a(a), LV3(e), RQ1(a)$	4,65	1,23
		CE	$vv1(a), vs(e), a(a), LV3(b), RQ1(a)$	5,49	1,62
		entropia	$vv1(a), vs(b), a(a), LV3(e), RQ1(a)$	4,65	1,23
		conjunta	$vv1(a), vs(b), a(a), LV3(b), RQ1(a)$	4,63	1,23
Db2	DE	$vv1(a), vs(b), a(a), LV3(e), RQ1(a)$	4,65	1,23	
	CE	$vs(b), vs(e), a(a), LV3(b), RQ1(a)$	4,68	1,49	
	entropia	$vv1(a), vs(b), a(a), LV3(e), RQ1(a)$	4,65	1,23	
	conjunta	$vv1(a), vs(b), a(a), LV3(b), RQ1(a)$	4,63	1,23	
6	Sem condicionamento	DE	$vv1(a), vv2(a), vs(b), a(a), RQ1(a), RQ3(c)$	4,42	1,24
		CE	$vv1(a), vv3(d), a(a), LV3(b), RQ1(a), RQ3(c)$	5,29	1,77
		entropia	$vv1(a), vv1(f), a(a), LV3(e), RQ1(a), RQ3(c)$	4,41	1,17
		conjunta	$vv1(a), vv1(f), vs(b), a(a), LV3(e), RQ1(a)$	4,07	1,10
	derivada	DE	$vv1(a), vv3(d), vs(b), a(a), LV3(e), RQ1(a)$	3,66	0,96
		CE	$vv1(a), vs(b), vs(e), a(a), LV3(e), RQ1(a)$	4,03	1,10
		entropia	$vv1(a), vv3(d), vs(b), a(a), LV3(e), RQ1(a)$	3,66	0,96
		conjunta	$vv1(a), vv5(a), vs(b), a(a), LV3(e), RQ1(a)$	3,09	0,89
	Haar	DE	$vv1(a), vv3(d), vs(b), a(a), LV3(e), RQ1(a)$	3,66	0,96
		CE	$vv1(a), vs(b), vs(e), a(a), LV3(e), RQ1(a)$	4,03	1,10
		entropia	$vv1(a), vv3(d), vs(b), a(a), LV3(e), RQ1(a)$	3,66	0,96
		conjunta	$vv1(a), vv5(a), vs(b), a(a), LV3(e), RQ1(a)$	3,09	0,89
	Coif1	DE	$vv1(a), vv3(d), vs(b), a(a), LV3(e), RQ1(a)$	3,66	0,96
		CE	$vv1(a), vs(b), vs(e), a(a), LV3(b), RQ1(a)$	4,02	1,10
		entropia	$vv1(a), vv3(d), vs(b), a(a), LV3(e), RQ1(a)$	3,66	0,96
		conjunta	$vv1(a), vv1(f), vs(b), a(a), LV3(b), RQ1(a)$	4,01	1,08
Db2	DE	$vv1(a), vv3(d), vs(b), a(a), LV3(e), RQ1(a)$	3,66	0,96	
	CE	$vv1(a), vs(b), vs(e), a(a), LV3(b), RQ1(a)$	4,02	1,10	
	entropia	$vv1(a), vv3(d), vs(b), a(a), LV3(e), RQ1(a)$	3,66	0,96	
	conjunta	$vv1(a), vs(b), vs(e), a(a), LV3(e), RQ1(a)$	4,03	1,10	

As demais medidas de similaridade escolheram no lugar da vegetação seca uma amostra do Latossolo Vermelho (*LV*). Para $R = 4$ com condicionamento as demais medidas também escolheram uma amostra da classe *vs* como MR, diminuindo assim o valor do erro médio da imagem-erro.

Na Tabela 8.4 para $R = 4$ sem condicionamento observam-se, com exceção da busca dos MR pela *DE*, grandes valores para os erros médios e desvios padrões associados às imagens-erro. A busca de MR pela *DE* acarretou um menor erro devido ao fato de ter sido escolhida uma amostra de vegetação seca, *vs(b)*, como MR e esta classe de alvo ser a mais abundante na cena. As demais medidas de similaridade escolheram no lugar da vegetação seca uma amostra do Latossolo Vermelho (*LV*). Para $R = 4$ com condicionamento as demais medidas também escolheram uma amostra da classe *vs* como MR, diminuindo assim o valor do erro médio da imagem-erro.

A classe *vs* está presente em todas as escolhas de MR, com exceção na escolha feita pela *CE* para $R = 5$, sem condicionamento e na escolha feita pela *CE* e entropia para $R = 6$, também sem condicionamento. Para este último caso (medidas *CE* e entropia, sem condicionamento) o erro médio da imagem-erro foi pequeno, pois as amostras das subclasses rotuladas como *vv3(d)* e *vv1(f)* substituíram convenientemente a classe *vs*.

Para $R = 5$ e 6 os valores dos erros para os espectros médios, com e sem condicionamento, são praticamente da mesma ordem de grandeza, isso ocorre, entre outros fatores, devido a presença de uma amostra da classe dominante *vs* como MR. Esse grande efeito da diminuição do erro, devido a presença da classe *vs*, atrapalha de certo modo a análise do desempenho do processo de busca dos MR através do valor médio da imagem-erro.

Para uma melhor análise dos resultados, apresentados na Tabela 8.4, mostram-se nas Figs. 8.13 a 8.16 algumas imagens pós-classificadas resultantes da utilização de conjuntos de MR no MLME.

A Fig. 8.13 mostra algumas imagens pós-classificadas, para $R = 4$, utilizando-se alguns conjuntos de MR selecionados na Tabela 8.4:

- A Fig. 8.13a mostra a imagem pós-classificada com os MR: *vv1(a)*, *vs(b)*, *a(a)* e *RQ1(a)*, escolhidos pela *DE* com e sem condicionamento dos espectros e pela

medida conjunta com condicionamento feito pela Db2 e Coif1. Nota-se que houve um grande erro na classificação da água (em azul), que se encontra somente ao longo da mata galeria $vv1$ (verde claro), como mostrado na Fig. 8.13c;

- A Fig. 8.13b mostra a imagem pós-classificada com os MR: $vs(b)$, $a(a)$, $LV3(e)$ e $RQ1(a)$ escolhidos pela CE , com os espectros condicionados pela derivada e Haar. Observa-se um menor erro na localização da classe água, mesmo considerando o erro médio superior ao do caso anterior;
- A Fig. 8.13c mostra a imagem pós-classificada com os MR: $vv1(a)$, $vs(b)$, $a(a)$ e $LV3(e)$ escolhidos pela entropia e medida conjunta, com os espectros condicionados pela derivada e Haar. Com essa escolha de MR tem-se uma melhor pós-classificação da água. Neste caso o erro médio é praticamente igual ao erro médio do primeiro caso (Fig. 8.13a), e
- A Fig. 8.13d mostra a imagem pós-classificada com os MR: $vs(e)$, $a(a)$, $LV3(b)$, $RQ1(a)$ escolhidas pela CE , com os espectros condicionados pela Db2 e Coif1, que produziram o pior resultado da classificação e têm o maior erro médio da imagem-erro.

De um modo geral os melhores resultados foram obtidos com os MR selecionados com a CE , entropia e medida conjunta, com o uso dos espectros condicionados pela derivada e pela TWD Haar (Figs. 8.13b e 8.13c).

Para $R=4$ observa-se ainda que todas as escolhas de MR, com e sem condicionamento, apresentam amostras pertencentes as classes: vv , vs , a , LV e RQ , presentes a subcena 4. Nota-se também que a grande proximidade dos valores dos erros médios (por ex. : $5,25 \pm 1,43$ e $5,27 \pm 1,35$) impede que se avalie o desempenho do classificador tomando-se por base somente os valores desses erros, as diferenças entre os erros não são estatisticamente significativas.

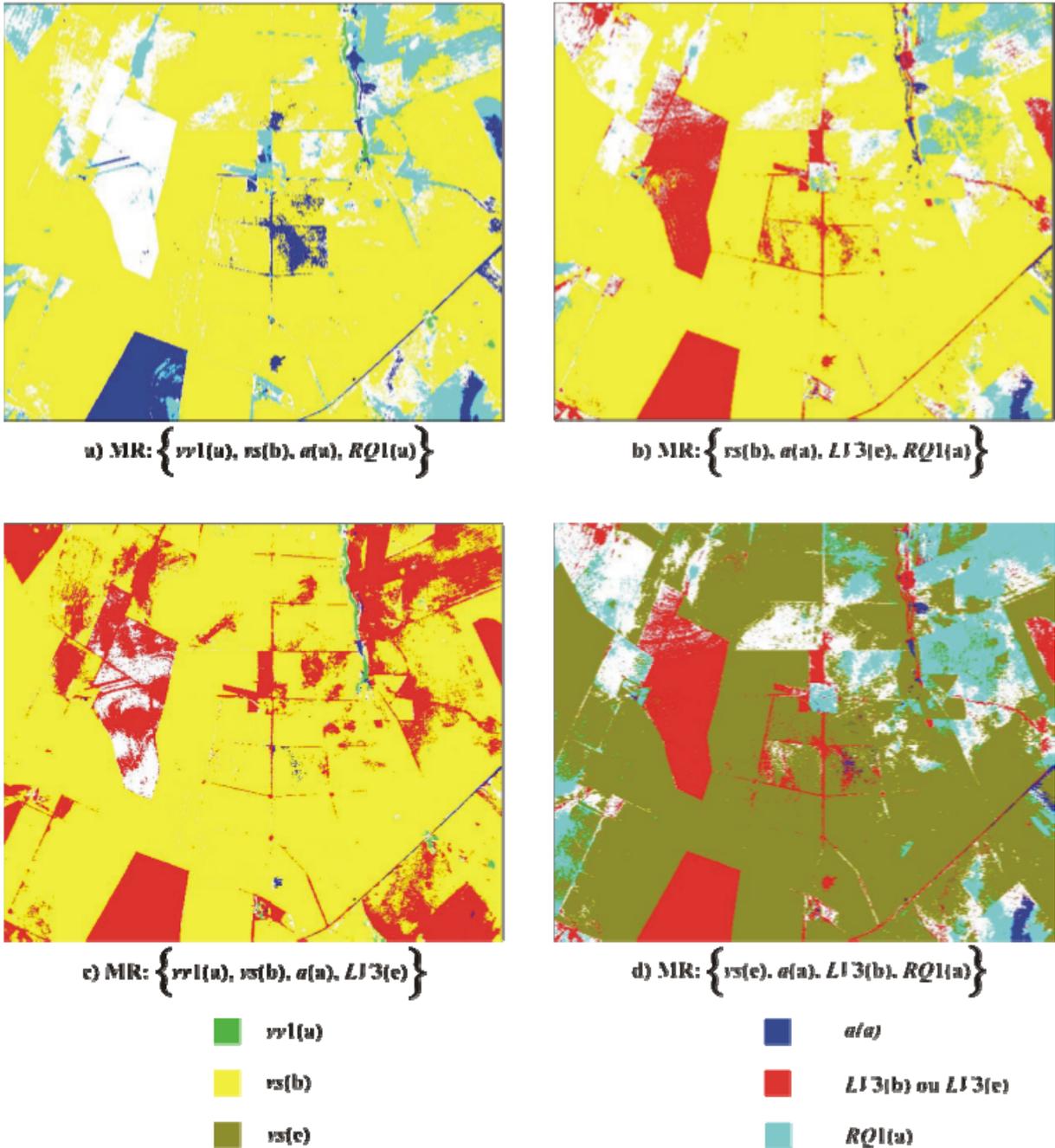


FIGURA 8.13 – Imagens pós-classificadas resultante da aplicação do MLME para $R = 4$, onde foram realçados os *pixels* com valores de abundância superiores 50 % da fração total. As áreas não-classificadas e, portanto, de intensa mistura espectral das componentes da cena são mostradas em branco.

Na Fig. 8.14 mostra algumas das imagens pós-classificadas, para $R = 5$, utilizando-se alguns conjuntos de MR selecionados na Tabela 8.4:

- Na Fig. 8.14a mostra a imagem pós-classificada com os MR: $v1(a)$, $v2(b)$, $a(a)$, $L13(c)$ e $RQ1(a)$, escolhidos pela medida da entropia usando os espectros sem e

com condicionamento e pela *DE* com os espectros condicionados. Estes mesmos MR foram escolhidos pela *CE* e medida conjunta com os espectros condicionados pela derivada e TWD Haar. Nesta escolha dos MR estão presentes as principais classes da imagem (*vv*, *vs*, *a*, *LV* e *RQ*);

- A Fig. 8.14b mostra a imagem pós-classifica com os MR: *vv1(a)*, *vs(b)*, *a(a)*, *LV3(b)* e *RQ1(a)*, escolhidos pela medida conjunta sem condicionamento e com condicionado dos espectros pela *Coif1* e *Db2*. Nas amostras selecionadas como MR estão presentes as principais classes de alvos na cena (*vv*, *vs*, *a*, *LV* e *RQ*). A imagem na Fig. 8.14b é praticamente idêntica a imagem da Fig. 8.1a;
- A Fig. 8.14c mostra a imagem pós-classificada com os MR: *vv1(a)*, *vs(e)*, *a(a)*, *LV3(b)* e *RQ1(a)*, escolhidos pela *CE* com os espectros condicionados pela *Coif1*. O erro médio da imagem-erro é o maior dentre todos, com a exceção do erro devido a *CE* sem condicionamento. O erro na classificação da água é maior que nos casos anteriores (Fig. 8.14a, b e d), e
- A Fig. 8.14d mostra a imagem pós-classificada com os MR: *vs(b)*, *vs(e)*, *a(a)*, *LV3(b)* e *RQ1(a)*, escolhidos pela medida *CE* com os espectros condicionados pela *Db2*. Verifica-se para esta configuração que a amostra *vv1(a)*, presente em todas as outras seleções, foi substituída por uma das amostras da classe *vs*, a *vs(e)*, ou seja, ocorre a seleção de amostras redundantes. Como consequência houve a eliminação do mapeamento da mata galeria presente em todas as outras imagens (Fig. 8.14a, b e c).

Novamente a grande proximidade dos valores dos erros médios da imagem-erro impede que se avalie o desempenho do classificador tomando-se por base somente os valores desses erros, as diferenças entre os erros não são estatisticamente significativas, a única exceção é o erro devido a *CE* sem condicionamento dos espectros, cujo valor do erro médio foi muito acima dos demais.

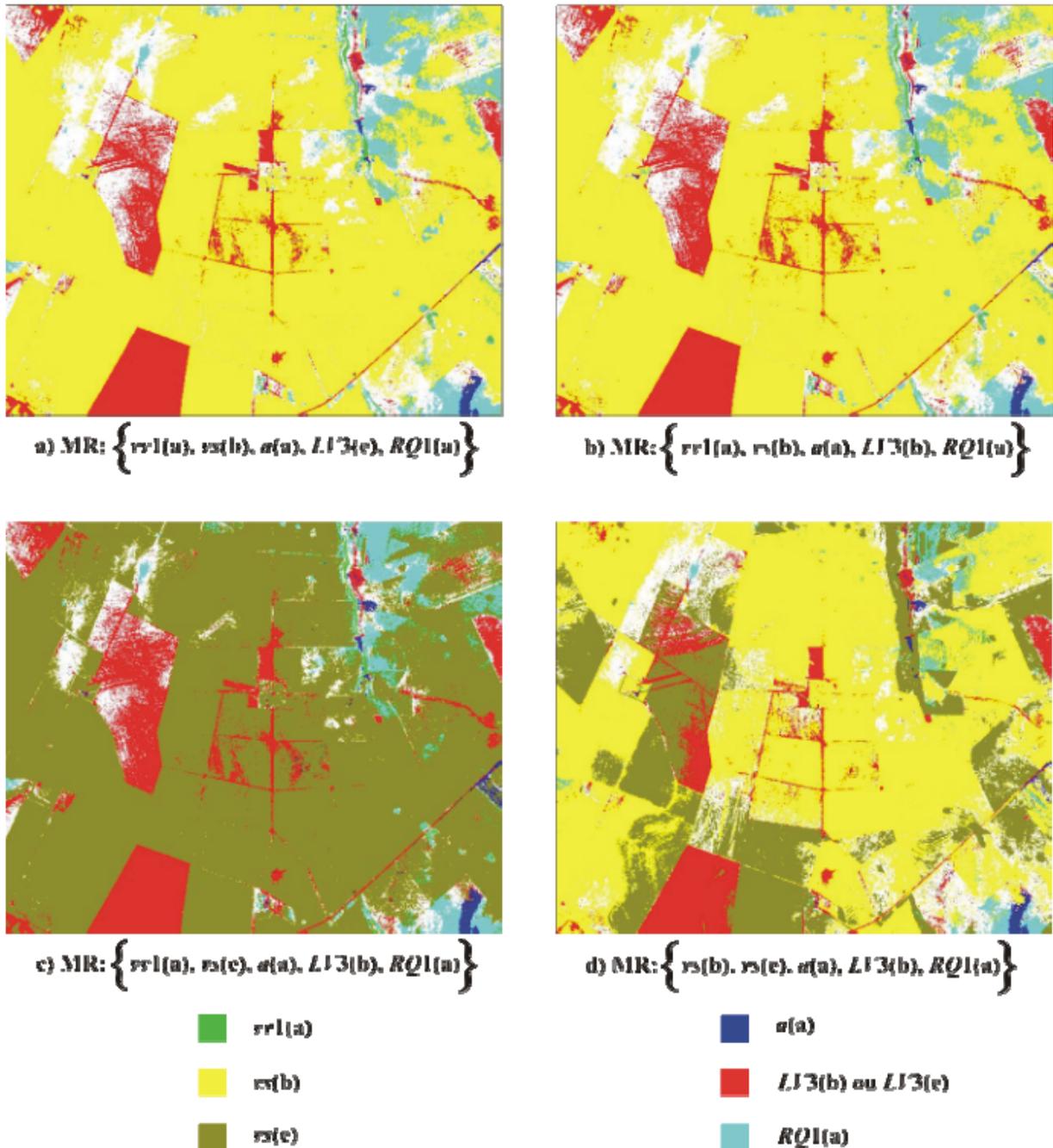


FIGURA 8.14 – Imagens pós-classificadas resultante da aplicação do MLME para $R = 5$, onde foram realçados os *pixels* com valores de abundância superiores 50 % da fração total. As áreas não-classificadas e, portanto, de intensa mistura espectral das componentes da cena são mostradas em branco.

De um modo geral os melhores resultados, para $R = 5$, foram obtidos na busca dos MR pela medida conjunta e pela entropia, ambas medidas utilizadas com e sem condicionamento, ou ainda por todas as medidas com os espectros condicionados pela derivada e pela TWD Haar.

A Fig. 8.15 mostra algumas das imagens pós-classificadas, para $R = 6$, utilizando-se alguns conjuntos de MR selecionados na Tabela 8.4, nestas classificações estão sempre presentes as principais classes (vv , vs , a , LV e RQ), incluindo-se algumas subclasses relativa à vegetação verde ou relativa à vegetação seca:

- A Fig. 8.15a mostra a imagem pós-classificada com os MR: $vv1(a)$, $vv3(d)$, $vs(b)$, $a(a)$, $LV3(e)$ e $RQ1(a)$, escolhidos pela DE e entropia, com os espectros condicionados. Nessa figura aparecem áreas pós-classificadas como água, onde não há ocorrência dessa classe;
- A Fig. 8.15b mostra a imagem pós-classificada com os MR: $vv1(a)$, $vv5(a)$, $vs(b)$, $a(a)$, $LV3(e)$ e $RQ1(a)$, escolhidos pela medida conjunta, com os espectros condicionados pela derivada e Haar. Nota-se que novamente há ocorrência de erros na classificação da classe água, mas neste caso os erros são menores que os da classificação anterior.
- A Fig. 8.15c mostra a imagem pós-classificada com os MR: $vv1(a)$, $vs(b)$, $vs(e)$, $a(a)$, $LV3(e)$ e $RQ1(a)$, selecionados pela CE , com os espectros condicionados pela derivada e pela Haar, e selecionados pela medida conjunta, com os espectros condicionados pela Db2. Nota-se que há duas amostras espectrais para a classe vs , ou seja, amostras redundantes, resultando em uma má escolha, e
- A Fig. 8.15d mostra a imagem pós-classificada com os MR: $vv1(a)$, $vv1(f)$, $vs(b)$, $a(a)$, $LV3(e)$ e $RQ1(a)$, escolhidos pela medida conjunta, com os espectros condicionados pela Coif1. Novamente há duas amostras para a mesma subclasse $vv1$ e que representa uma má escolha.

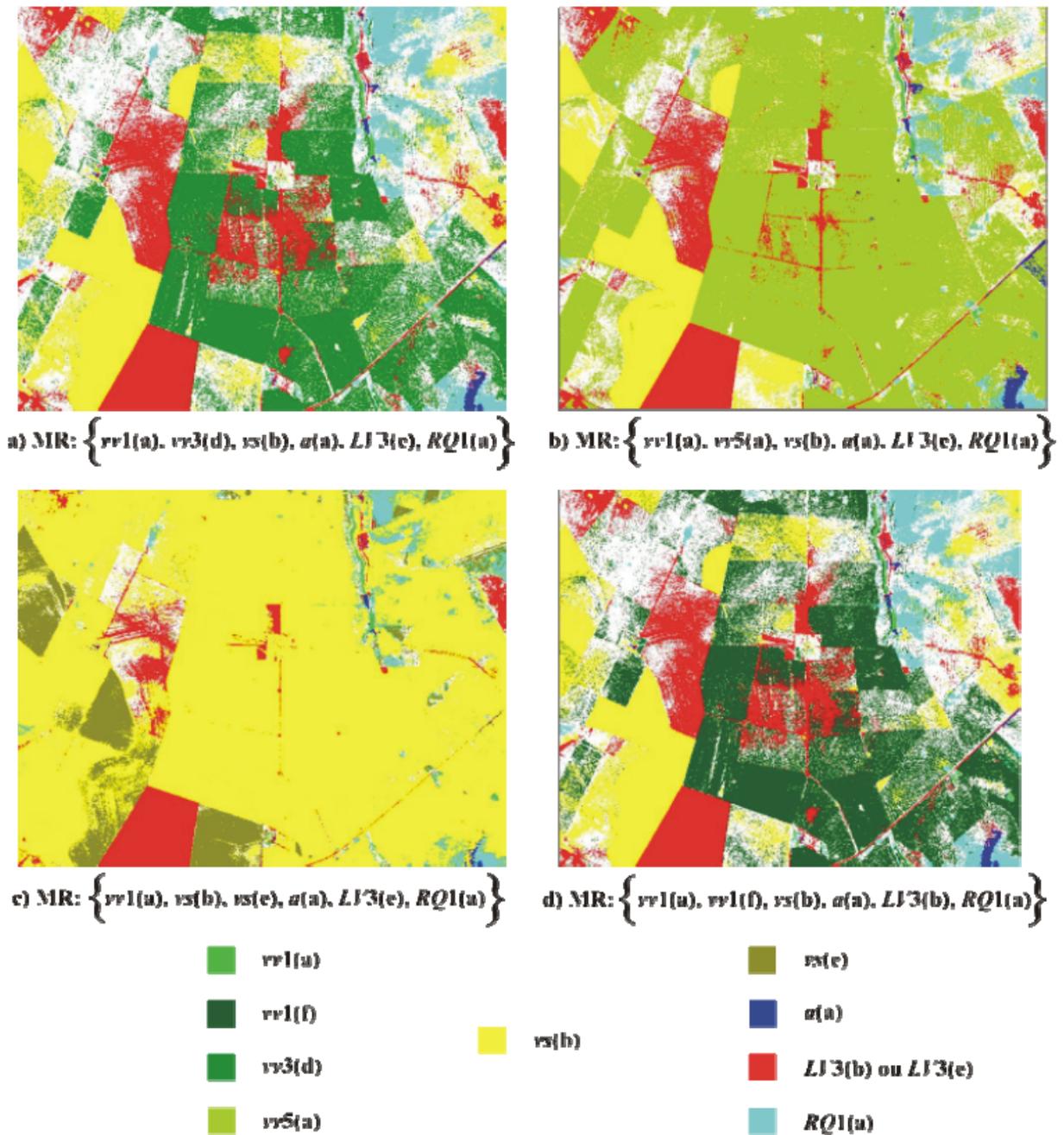


FIGURA 8.15 – Imagens pós-classificada relativas às medidas *DE* e *CE* resultante da aplicação do MLME para $R = 6$, onde foram realçados os *pixels* com valores de abundância superiores 50 % da fração total. As áreas não-classificadas e, portanto, de intensa mistura espectral das componentes da cena são mostradas em branco.

Novamente a grande proximidade dos valores dos erros médios da imagem-erro impede que se avalie o desempenho do classificador tomando-se por base somente os valores desses erros. Da análise qualitativa das imagens pós-classificadas, para $R = 6$, tem-se que os melhores resultados, de uma maneira geral, foram os apresentados pelas *DE*, entropia, com os espectros condicionados.

Da Tabela 8.4 e da análise qualitativa das imagens pós-classificadas conclui-se que o método de busca de MR que produziu os melhores resultados, para a subcena 4 estudada, foi aquele que utilizou a entropia e a medida conjunta, com os espectros condicionados pela derivada e pela TWD Haar.

Nota-se ainda que a medida que o valor de R aumenta as medidas sem condicionamento tendem a selecionar amostras redundantes. Desta forma, para a subcena estudada, convém utilizar o condicionamento dos espectros, principalmente utilizando-se a derivada e a TWD Haar.

Quanto aos erros na classificação da classe água, verificou-se que esse erro depende dos MR escolhidos que não pertencem a classe água. As hipóteses prováveis para explicar a confusão no mapeamento da classe água seriam: a) o MLME não é suficientemente preciso na estimativa das frações abundâncias, ou b) há a presença de sombras na subcena analisada.

8.4 Triagem discriminatória e busca dos MR para a subcena 7

Nesta seção são apresentados alguns resultados referentes ao processo de triagem discriminatória, aplicada a toda a imagem relativa a subcena 7 e a busca dos MR.

Para o teste da triagem discriminatória foram escolhidos os seguintes parâmetros: vizinhança inicial $(M + 1) \times (M + 1) = 5 \times 5 \text{ pixels}$, fator de uniformidade espacial $\alpha_u = 60\%$, limiar de seleção espacial $\psi_e = 78\%$, fator de homogeneidade $\alpha_h = 90\%$, nível de significância $\alpha = 10\%$ para teste de igualdade das médias, vizinhança para teste do contexto $(M_c + 1) \times (M_c + 1) = 5 \times 5 \text{ pixels}$, fator de contexto $\alpha_c = 80\%$, limiar de redundância para a medida DE $\psi_{rde} = 2,1\%$, limiar de redundância para a medida CE $\psi_{rce} = 3,8\%$, e o conjunto união \mathfrak{R}_U para as amostras selecionadas.

Após os testes de uniformidade espacial, homogeneidade espectral, contexto e redundância, com os parâmetros definidos anteriormente, obteve-se $K = 3.048$ amostras espectrais. A Fig. 8.17a mostra a subcena 7 (banda 192 - 2.271 nm) e Fig. 8.17b ilustra a mascara gerada pelo teste do contexto aplicada à imagem. As regiões em preto são provenientes das posições dos *pixels* que foram rejeitados pelos testes da uniformidade espacial, homogeneidade espectral e contexto. Na Fig. 8.17c mostra a posição das $K = 3.048$ amostras menos redundantes selecionadas pelo teste da redundância.

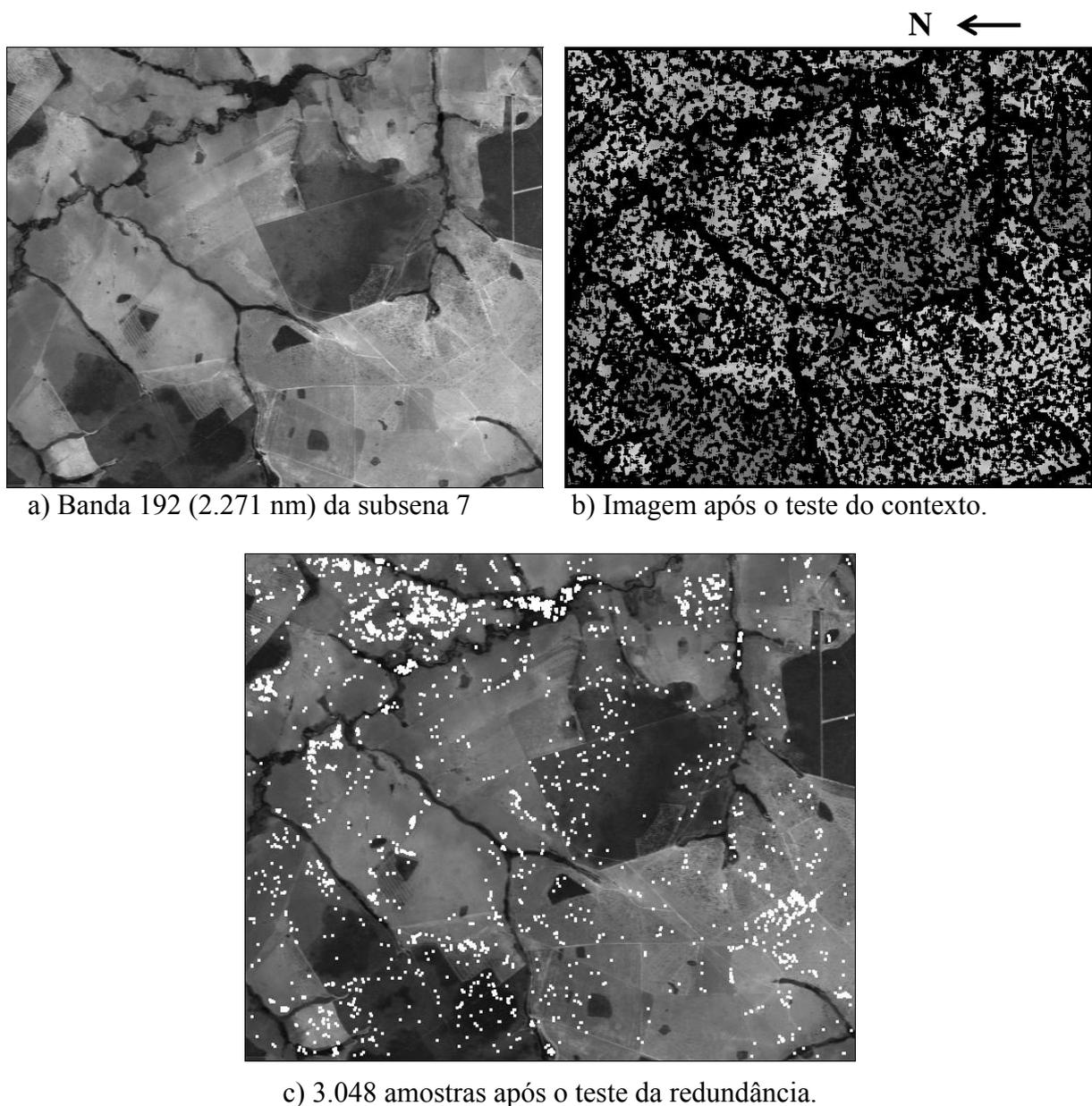


FIGURA 8.17 – Aplicação da triagem discriminatória na subcena 7.

Para redução do tempo computacional no processo de busca dos MR, dentre as $K = 3.048$ amostras, aplicou-se novamente o teste da redundância a este conjunto de amostras. Para este teste utilizou-se $\psi_{rde} = \psi_{rce} = 11\%$ e novamente o conjunto união das amostras, \mathfrak{R}_{\cup} , reduzindo-se para $k = 83$ o número de amostras candidatas a MR.

Estas $k = 83$ amostras estão representadas na Fig. 8.18. Nesta mesma figura destacam-se também as sete amostras que serão escolhidas como MR pelo procedimento de busca dos MR que será descrito em seguida. Na Fig. 8.18 observa-se a boa distribuição das 83 amostras ao longo da cena imageada, com exceção de amostras de solo *RQ* (Neossolo Quartzârenico) todas as demais classes e subclasses de alvos têm pelo menos uma amostra no

conjunto. A classe RQ , presente na subcena 7 e representada nas Figs. 3.7 e 3.8, tem uma cobertura vegetal que dificulta a sua detecção, entretanto esta mesma classe foi detectada e escolhida como MR na subcena 4 e subcena 5, nos Itens 8.2.1 e 8.2.2, respectivamente.

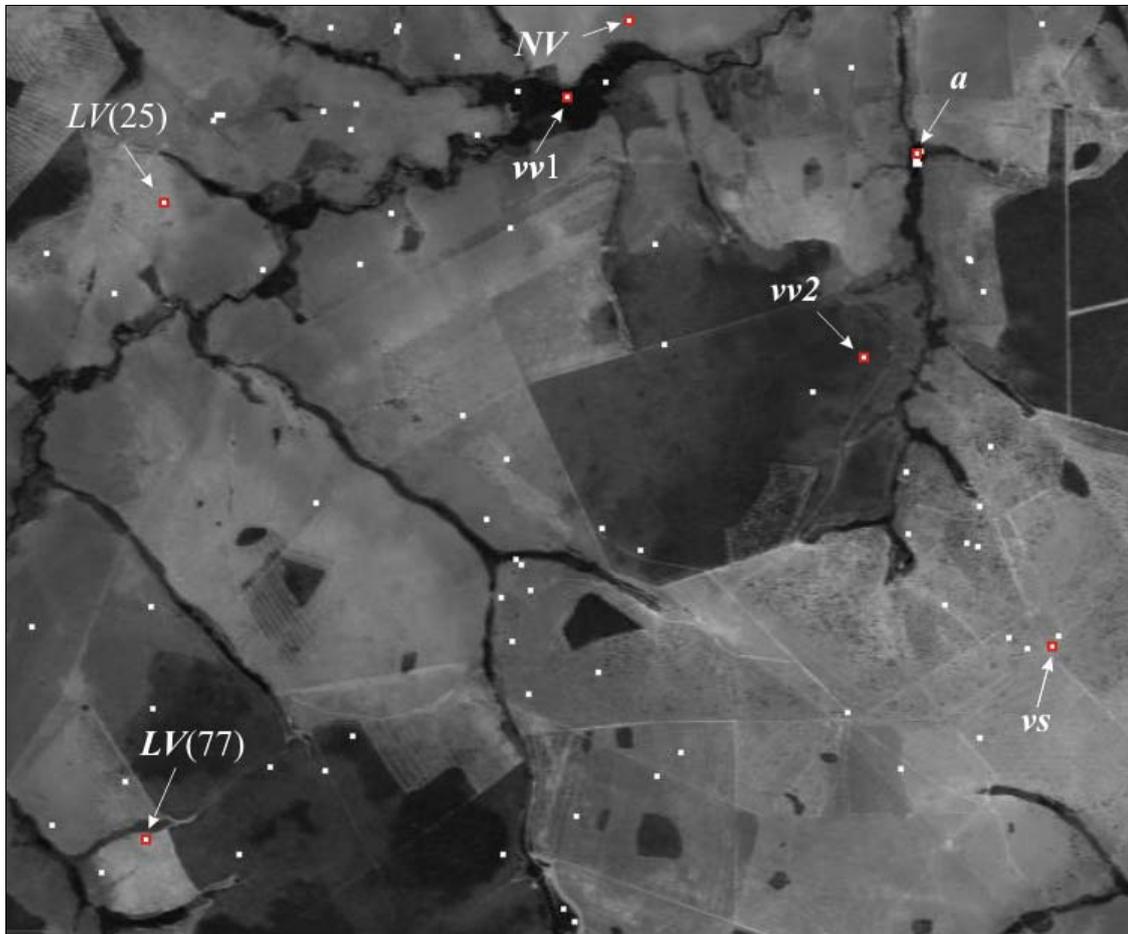


FIGURA 8.18 – Posição das 83 amostras candidatas a MR e das sete amostras selecionadas como MR para subcena7 (banda 192 - (2.271 nm)).

Para a seleção dos MR utilizou-se os seguintes fatores de configuração: $\alpha_{DE} = 50\%$, $\alpha_{CE} = 77\%$ e $\alpha_H = 80\%$. O resultado do processo de seleção dos MR é apresentado na Tabela 8.5, utilizando-se isoladamente as medidas de similaridades espectrais DE , CE e entropia e o processo de múltiplas medidas $(e^3_1, e^3_2, \dots, e^3_R)$, conforme estabelecido na seção 7.3. Nesta tabela são mostrados também os erros médios e desvios padrões associados a imagem-erro obtida do MLME, e os valores da entropia relativos ao conjunto de MR selecionados.

TABELA 8.5 - MR selecionados com para a subcena 7.

<i>R</i>	Medidas de similaridades	Amostras espectrais selecionadas	Entropia	Imagem-Erro	
				média	desvio padrão
2	<i>DE, CE</i> , entropia e conjunta	<i>vv1, a</i>	0,99	88,68	31,035
3	<i>DE</i>	<i>vv1, vs, a</i>	0,87	9,37	4,04
	<i>CE</i> , entropia e conjunta	<i>vv1, a, LV(77)</i>	0,91	15,49	5,32
4	<i>DE</i> e entropia	<i>vv1, vs, a, LV(77)</i>	0,78	5,72	2,54
	<i>CE</i> e conjunta	<i>vv1, a, NV, LV(25)</i>	0,73	7,79	3,20
5	<i>DE</i> e entropia	<i>vv1, vs, a, NV, LV(25)</i>	0,61	4,54	2,00
	<i>CE</i> e conjunta	<i>vv1, vv2, a, NV, LV(25)</i>	0,61	4,73	2,12
6	<i>DE, CE</i> , entropia e conjunta	<i>vv1, vv2, vs, a, NV, LV(25)</i>	0,54	3,01	0,86
7	<i>DE, CE</i> , entropia e conjunta	Todos os conjuntos de amostras foram considerados mal configurados $\Rightarrow \mathfrak{R}_1 = 6$	-----	-----	-----

No processo de triagem e busca dos MR foi utilizado um Computador Pessoal com CPU (núcleo simples e 32 bits) com 2.250 MHz, com três pentes de memória SDRAM DDR com 512 MBytes cada e uma unidade de disco rígido com interface IDE ATA - 7.200 rpm. Para abreviar ao longo tempo de processamento no processo de busca dos MR para $R \geq 5$ utilizou-se os MR identificados por *vv1, a, NV* e *LV(25)*, relativos a medida conjunta para $R = 4$, como MR permanentes nas escolhas.

Pela Tabela 8.5, para $R = 2, 3$ e 4 observa-se que:

- Todas as medidas mantiveram as escolhas anteriores, acrescentando a estas uma nova classe ou subclasse;
- As medidas *CE* e conjunta apresentaram sempre os mesmos resultados;
- A inclusão da classe *vs* como MR diminui muito o valor médio da imagem-erro. Isto ocorre devido a grande presença dessa classe na imagem utilizada e
- As medidas *CE* e conjunta substituíram no processo de busca dos MR a amostra 25, relativa à classe *LV*, pela amostra 77, desta mesma classe.
- Para $R = 6$ todas as medidas coincidiram;

- A partir de $R=7$ todas as combinações das 83 amostras, agrupadas em conjuntos de sete amostras tendo entre as amostras $\{vv1, a, NV, LV(25)\}$, foram consideradas mal configuradas, ou seja de (6.19) tem-se que $U_{bc}^{R=7} = \emptyset$, deste modo, não houve escolha de nenhum MR. Logo, o limitante superior para o número de MR da cena é $\mathfrak{R}_1 = 6$.

As amostras escolhidas como MR estão indicadas Fig. 8.18 e os seus espectros médios calculados em uma vizinhança com uniformidade espacial estão dispostos na Fig. 8.19.

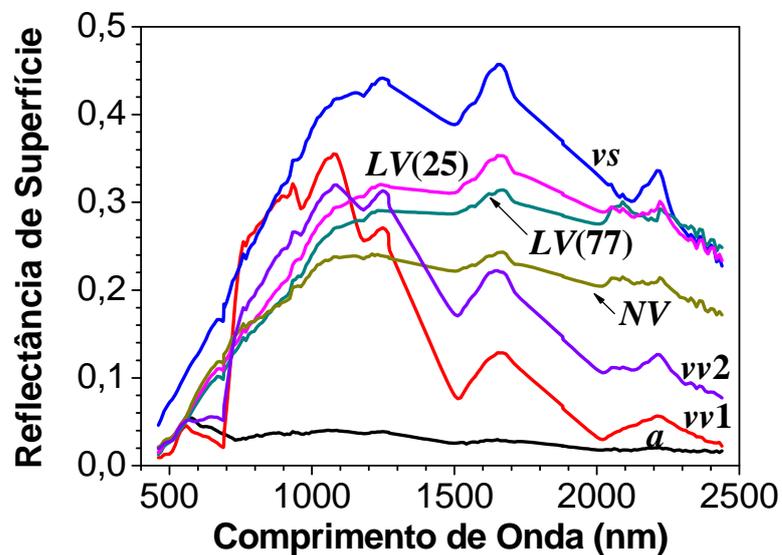


FIGURA 8.19 – Espectros médios dos MR da subcena 7 selecionados para $2 \leq R \leq 6$.

As Figs. 8.20a e 8.20b mostram, respectivamente, as imagens pós-classificadas para $R=5$ com os MR selecionados pela *CE* e conjunta $\{vv1, vv2, a, NV, LV(25)\}$ e com os MR selecionados pela *DE* e entropia $\{vv1, vs, a, NV, LV(25)\}$. Na classificação mostrada na Fig. 8.20a há indicação de *pixels* pós-classificados com a classe *a* (água) em locais com solo ou vegetação e a indicação da classe solo *LV* em áreas com *vs* (vegetação seca). Estes erros não são tão intensos na classificação mostrada na Fig. 8.20b. Portanto, para $R=5$ os MR escolhidos pela *DE* e pela entropia não só apresentam o menor erro médio da imagem-erro, como também possibilitam a melhor classificação da cena. Constata-se, porém, que os valores do erro médio para as duas escolhas citadas ($4,73 \pm 2,12$ e $4,54 \pm 2,00$) não são

significativamente diferentes, fazendo com que seus valores não possam ser utilizados para a discriminação do melhor desempenho na escolha dos MR.

A Fig 8.20c mostra a classificação para $R = 6$ com os MR dados pelo conjunto $\{vv1, vv2, vs, a, NV, LV(25)\}$, escolhido por unanimidade por todas as medidas. Uma análise qualitativa dos resultados mostra que a classificação realizada com os seis MR foi a melhor de todas. A classificação mostrada pela Fig. 8.20c também apresenta uma grande coerência com a classificação representada na Fig. 8.20b. O erro médio da imagem-erro, neste caso, foi menor que os anteriores.

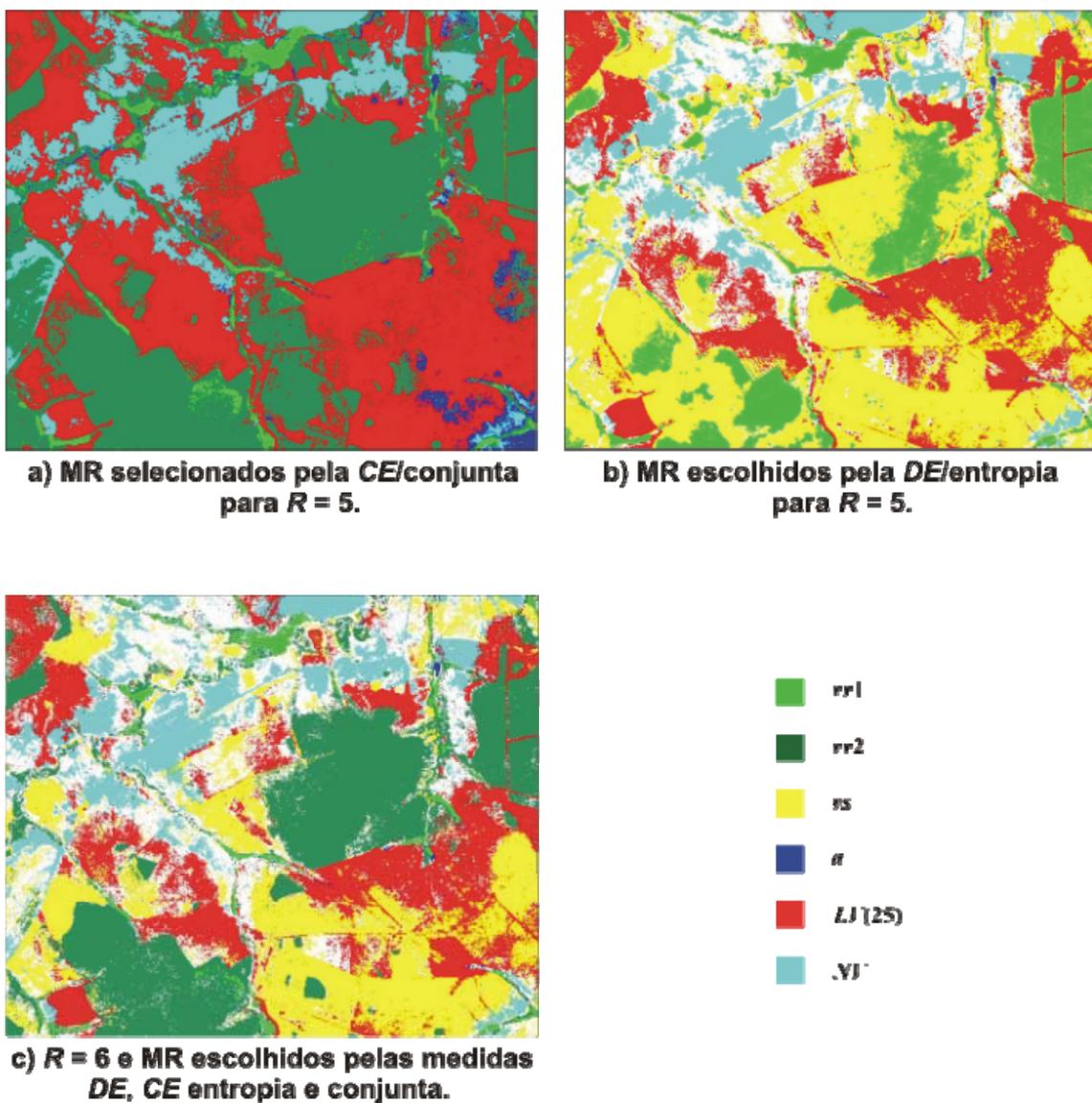


FIGURA 8.20 – Imagens pós-classificadas resultante da aplicação do MLME para $R = 5$ e 6 , onde foram realçados os *pixels* com valores de abundância superiores 50 % da fração total. As áreas não-classificadas e, portanto, de intensa mistura espectral das componentes da cena são mostradas em branco.

Nota-se ainda na Tabela 8.5 que os valores da entropia relativos às amostras espectrais do conjuntos de MR escolhidos são superiores a 0,5, que foi o valor mínimo da entropia escolhido anteriormente para a definição do limitante superior \mathfrak{R}_2 do número de MR.

8.5 Comparação com o algoritmo PPI

O PPI (*Pixel Purity Index*) é um algoritmo que tem sido muito utilizado pela comunidade de sensoriamento remoto devido a sua versatilidade e a sua disponibilidade no aplicativo computacional ENVI® (*Environment for Visualizing Images da RSI*).

Este método tem o objetivo de localizar na imagem hiperespectral os *pixels* espectralmente mais puros, identificando os *pixels* extremos em um espaço de atributos, que é gerado pela transformação MNF (*Minimum Noise Fraction*) (BOARDMAN et al., 1995; CHANG; PLAZA, 2006; ENVI, 2006; GREEN et al., 1988; MOREIRA, R.C., 2000; THEILER et al., 2000). Os *pixels* selecionados pelo PPI podem ser visualizados com o uso da ferramenta denominada *N-dimensional (N-Dimensional Visualizer)*, que é muito similar à ferramenta utilizada pela técnica MEST (BATESON; CURTISS, 1996) que utiliza os diagramas de dispersão dos primeiros autovetores obtidos através da aplicação da transformação MNF nos dados originais.

Foi utilizado o aplicativo ENVI com a opção automática, ou seja, não há qualquer interação intermediária do usuário no processo para a localização dos *pixels* puros, na qual incluem-se as fases MNF, PPI e *N-dimensional*. Utilizou-se como parâmetros: $L = 10.000$ interações, $t = 2,5$ (limite espacial de tolerância, que regula o número de *pixels* puros escolhidos) e número de imagens, ou dimensões, da transformação MNF como sendo $N_{MNF} = 10, 20$ e 30 imagens. Constatou-se que a partir da décima primeira banda MNF até a centésima banda, houve uma grande diminuição na relação sinal-ruído (listras ou má definição dos contornos entre classes de alvos) nas imagens MNF. Essa diminuição da relação sinal-ruído com o aumento do número das bandas foi estudado por Kruse e Huntington (1996) e consta também no manual do aplicativo ENVI (2006).

Para o teste do PPI foi utilizada a subcena 4 (Fig. 3.3) onde se destacam as seguintes classes e subclasses, listadas na seção 3.2: a) classe vegetação verde (*vv*) composta por subclasses: *vv1, vv2, vv3, vv4* e *vv5*; b) classe vegetação seca (*vs*); c) classe Latossolo

Vermelho (*LV*) composta pelas subclasses: *LV1*; *LV2* e *LV3*; d) classe Neossolo Quartzarênico (*RQ*) composta pelas subclasses: *RQ1*, *RQ2* e *RQ3* e e) classe água (*a*).

A Fig. 8.21 realça, na composição R (PPI), G (banda 549 nm) e B (PPI), os *pixels* tidos como puros pela técnica PPI na subcena 4, para $N_{MNF} = 20$. Nessa imagem evidencia-se 5.910 *pixels* puros nas cores magenta e branca obtidos pelo PPI. As indicações de classe e subclasses indicadas nesta figura são as mesmas presentes na Fig. 3.3.

Para $N_{MNF} = 10$ o método PPI selecionou entre os *pixels* considerados puros amostras das classes/subclasses:

vv1, *vv3*, *vv4*, *vs*, *a*, *LV1*; *LV3*, *RQ1*, e *RQ3*.

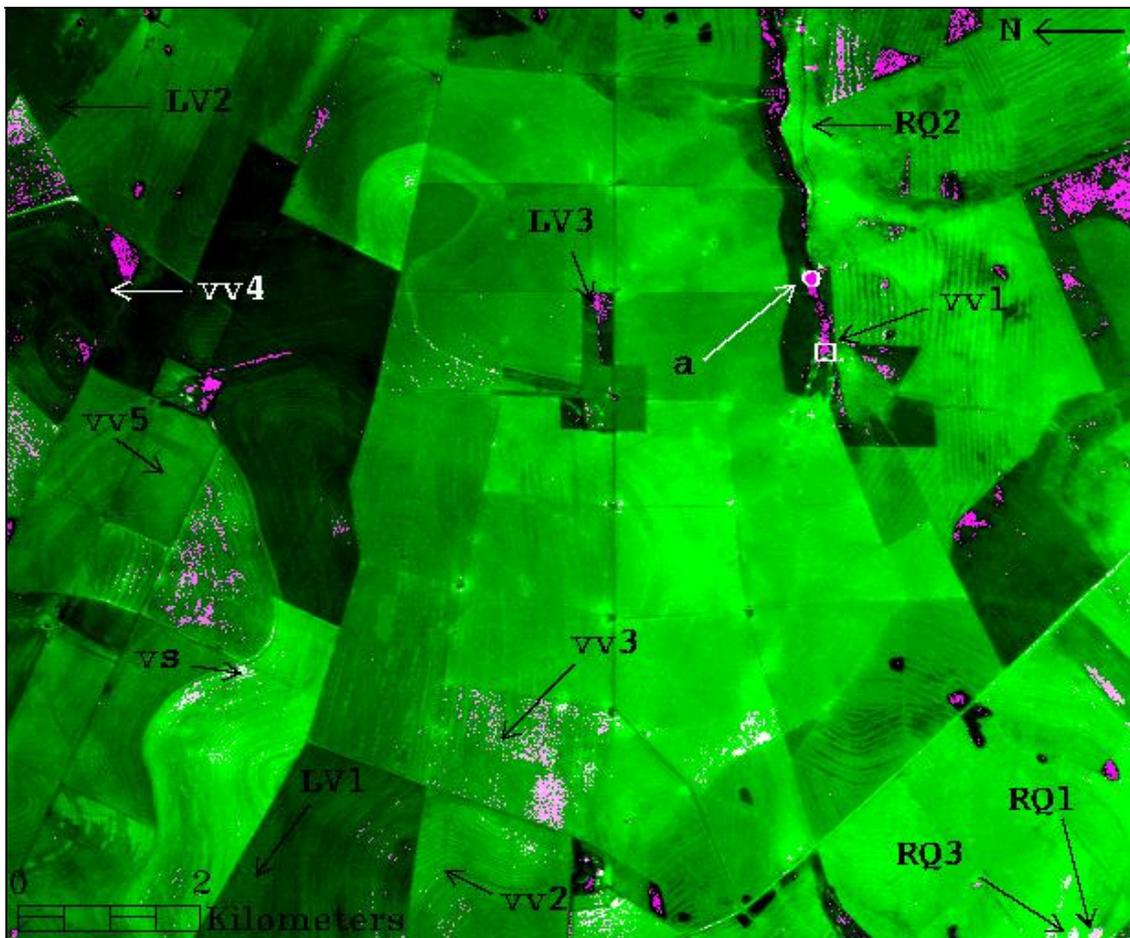


FIGURA 8.21 – Composição colorida R (PPI), G (549 nm), B (PPI) para a subcena 4, identificando os *pixels* puros, em magenta, pela técnica PPI.

Para $N_{MNF} = 20$, mostrado na Fig. 8.21, o método PPI selecionou entre os *pixels* considerados puros amostras das classes/subclasses:

$vv1, vv2, vv3, vv4, vs, a, LV1, LV3, RQ1$ e $RQ3$.

Para $N_{MNF} = 30$ o método PPI selecionou entre os *pixels* considerados puros amostras das classes/subclasses:

$vv1, vv3, vv4, vs, a, LV1, RQ1$ e $RQ3$.

Uma análise mais detalhada mostrou que os melhores resultados foram obtidos para $N_{MNF} = 20$. Nota-se também que em todos os casos ($N_{MNF} = 10, 20$ e 30) foram escolhidas amostras de todas as principais classes (vv, vs, a, LV e RQ). Constatou-se ainda que na seleção automática dos espectros puros realizada pelo PPI, para $N_{MNF} = 10, 20$ e 30 , há um número elevado de espectros que apresentam falhas na aquisição (classes anômalas). Os três testes foram realizados de modo a se minimizar a seleção de espectros com falhas e com isso melhorar a seleção de espectros representativos das principais classes e subclasses de alvos.

No método da triagem aleatória, apresentada na item 8.1.2, para a subcena 4, partindo de 900 amostras escolhidas aleatoriamente, pela metodologia AESNA, obteve um conjunto de 107 amostras candidatas a MR, das quais constavam as classes/subclasses:

$vv1, vv2, vv3, vv4, vv5, vs, LV1, LV2, LV3$ e $RQ3$.

Observa-se nessa escolha a ausência da classe água (a) e a presença de todas as demais classes.

No teste da triagem discriminatória, apresentada no item 8.1.3, também para a subcena 4, utilizando-se o conjunto de amostras menos redundantes \mathfrak{R}_U obteve-se 5,09% do total de 612×512 *pixels*, perfazendo um total de 15.949 *pixels* ou amostras candidatas a MR. Estas amostras estão representadas na Fig. 8.5e. Dentre as amostras candidatas constatou-se a presença das seguintes classes/subclasses:

$vv1, vv2, vv3, vv4, vv5, vs, LV1, LV2, LV3, RQ1, RQ2$ e $RQ3$.

Com a exceção da classe água (a) todas as demais amostras de interesse estão presentes neste ultimo resultado.

Na triagem discriminatória os testes de uniformidade espacial e homogeneidade espectral, ao contrário do PPI, tendem a descartar amostras com falhas de aquisição (amostras

anômalas), além de identificar vizinhanças mais uniformes e homogêneas em torno das amostras candidatas a MR.

A Fig. 8.22 mostra um conjunto de 19 espectros selecionados automaticamente pelo PPI para $N_{MNF} = 20$, tidos como os mais puros dentre todos os *pixels* selecionados. Nota-se que muitos dos espectros selecionados possuem falhas na aquisição em algumas bandas. Dentre estes espectros encontra-se somente as subclasses de interesse *vv1* e *RQ3*.

Para $N_{MNF} = 10$ dentre os espectros selecionados automaticamente pelo PPI, considerados mais puros, encontrava-se presente somente à classe *a*. Para $N_{MNF} = 30$ dentre os espectros selecionados automaticamente pelo PPI, considerados mais puros, encontravam-se presentes somente as classes *a* e *RQ3*.

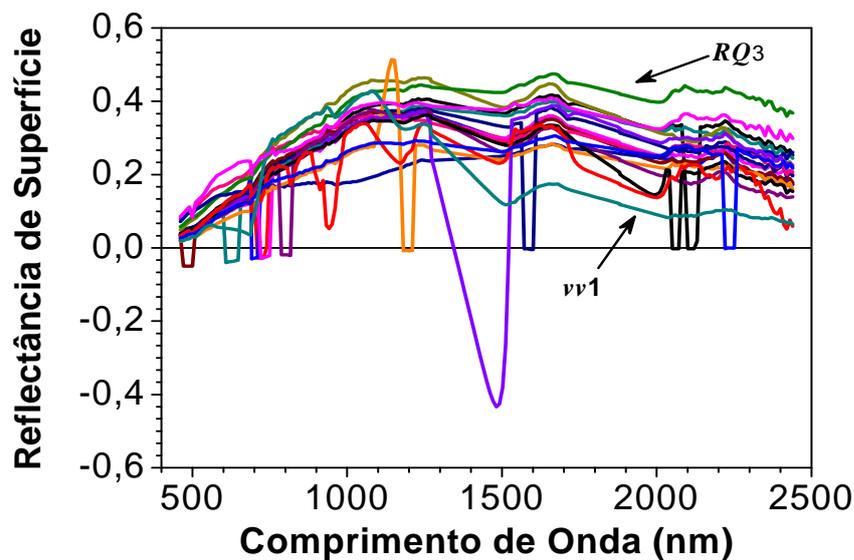


FIGURA 8.22 – Espectros selecionados automaticamente pelo PPI para a subcena 4 com $N_{MNF} = 20$, $L = 10.000$ e $t = 2,5$.

Conclui-se que, para a subcena 4, o PPI, com os parâmetros utilizados, selecionou amostras das principais classes e subclasses presentes na cena. O método da triagem discriminatória, utilizado para a subcena 4 com parâmetros constantes nos itens 8.1.2 e 8.1.3, com exceção da classe *a*, também selecionou amostras das classes e subclasses presentes na cena. Dentre as amostras selecionadas pelo PPI encontram-se muitas amostras com falhas de aquisição (amostras anômalas) estas mesmas amostras são pouco encontradas dentre as amostras selecionadas pela triagem discriminatória.

As posições dos *pixels* mais puros, fornecidas pelo PPI, como aquelas mostradas na Fig. 8.21, podem ainda ser utilizadas por um operador no ENVI para que se procure, manualmente, bons espectros candidatos a MR. Este procedimento que envolve a utilização de um mapa com a posição dos *pixels* mais puros pode também ser realizado utilizando-se as máscaras calculadas pelo teste de triagem discriminatória. Exemplos destes mapas, para a subsena 4, encontram-se nas Figs. 8.4e, 8.5c-f e 8.6.

Observa-se também que a técnica PPI pode ser utilizada como um processo de triagem de amostras candidatas a MR. Estas amostras escolhidas pelo PPI podem então ser submetidas aos processos de busca dos MR propostos nos Capítulos 6 e 7. Chang e Plaza (2006a) também sugerem a utilização do PPI para a seleção do conjunto inicial de amostras espectrais que serão submetidas a um processo de busca dos MR. Além disso, as amostras provenientes do PPI podem ser submetidas a um processo de triagem discriminatória antes de serem utilizadas nos processos de busca dos MR.

9 Considerações Finais

Apresenta-se neste capítulo um resumo do estudo realizado, destacam-se as principais contribuições propostas na área de espectroscopia de imageamento com dados hiperespectrais e sugerem-se temas para a continuação do estudo realizado.

9.1 Sumário do estudo realizado

Fez-se a descrição das principais características dos espectros ópticos presentes em imagens hiperespectrais, em particular, descreveu-se as características dos dados hiperespectrais do sensor AVIRIS utilizados neste estudo. Discutiu-se o conceito de MR e a sua utilização no MLME.

Propôs-se o método de triagem discriminatória de amostras espectrais com a finalidade de se selecionar amostras candidatas a MR. Esse método utiliza os conceitos de uniformidade espacial, homogeneidade espectral, contexto espacial e redundância, definidos neste estudo.

Fez-se a conceituação de um conjunto de amostras espectrais bem configuradas e definiu-se a medida de similaridade espectral baseada no cálculo da entropia e na medida de similaridade com múltiplas medidas. Propôs-se ainda uma nova forma de condicionamento dos espectros, das amostras utilizados pela metodologia de busca de MR, através da TWD.

Realizaram-se testes de desempenho da metodologia proposta utilizando-se três subcenas de imagens hiperespectrais adquiridas pelo sensor AVIRIS e comparou-se a metodologia proposta com o método PPI na busca de MR da cena imageada.

9.2 Contribuições para a área de espectroscopia de imageamento

As principais contribuições deste trabalho foram:

- a) Seleção de amostras espectrais pelo método da triagem discriminatória: Este processo possibilita a identificação, na imagem hiperespectral, de amostras espectrais uniformes espacialmente e homogêneas espectralmente. Possibilita ainda identificar aglomerados espaciais de amostras que também possuam as características de uniformidade e de homogeneidade, estabelecendo assim um contexto espacial entre as amostras candidatas a

MR. Pode ainda fazer uma redução do número de amostras, uniformes espacialmente e homogêneas espectralmente, rejeitando amostras redundantes;

b) Janela espacial adaptativa de processamento: A utilização do teste de uniformidade espacial possibilita, a partir de uma vizinhança inicial $(M + 1) \times (M + 1)$, o estabelecimento de uma vizinhança espacial particular para cada amostra espectral. Dentre os testes realizados, e não descritos em detalhes neste estudo, está o uso da vizinhança inicial $(M + 1) \times (M + 1) = 5 \times 5 \text{ pixels}$. Esta vizinhança é a menor possível visto que a vizinhança $(M + 1) \times (M + 1) = 3 \times 3 \text{ pixels}$ se mostrou muito pequena por conter apenas nove amostras. Este número diminuto de amostras, na maioria das situações, faz que a eliminação de poucas amostras, detectadas como contaminação na vizinhança, levem a rejeição da amostra como candidata a MR.

c) Detector de bordas para imagens hiperespectrais: Como salientado na Seção 8.1, Item 8.1.3, o teste da uniformidade espacial pode ainda ser utilizado como um detector de bordas;

d) Geração de mascaras com aglomerados de amostras com uniformidade espacial, homogeneidade espectral através do teste do contexto espacial: Estas mascaras podem ser utilizadas no ENVI[®], juntamente com o conjunto de imagens hiperespectrais, por um operador no processo de avaliação de classes/materiais presentes na imagem de uma cena;

e) Condicionamento espectral pela TWD: O espectro resultante deste condicionamento são os detalhes do segundo nível de decomposição, obtidos das aproximações do primeiro nível de aproximação. Isto equivale a aplicar o filtro passa baixas no espectro original seguido do processamento equivalente a uma derivação. Como observado na Seção 8.3 este condicionamento pode potencializar a discriminação dos espectros das amostras candidatas a MR, facilitando assim a busca dos melhores MR para a imagem hiperespectral;

f) Identificação de conjuntos de amostras espectrais bem configuradas: Como existe muita similaridade entre os espectros provenientes da maioria dos alvos/materiais presentes nas cenas, fica extremamente difícil a obtenção dos MR que tenham um significado válido. A validade do significado dos MR escolhidos está relacionada com a capacidade do grupo de MR representar a variabilidade espectral, presente na cena, através de espectros escolhidos como MR que tenham sentido físico e não apenas matemático. Essa validade, ou

capacidade de representação útil, pode ser verificada analisando-se, por exemplo, as imagens-fração. Conjuntos de MR bem configurados geram imagens-fração mais coerentes com as verdades terrestres e também coerentes com os elementos básicos presentes na própria cena;

g) Medida de similaridade espectral pela entropia: Muitos métodos de medida da similaridade espectral podem ser utilizados somente entre dois espectros, como exemplo tem-se a *DE* e a *CE*. O método proposto baseado na entropia, conforme definido no Capítulo 6, possibilita a medida da similaridade espectral entre um conjunto de espectros. A maximização dessa medida possibilita ainda a identificação do conjunto de MR referentes à cena;

h) Limitante do número de MR presentes na cena: Aplicando-se o conceito de amostras bem configuradas e estabelecendo-se um limite inferior para a entropia de um conjunto de amostras espectrais estabeleceu-se dois limitantes superiores para o número máximo de MR presentes na imagem hiperespectral;

i) Busca de MR com múltiplas medidas de similaridade: Utilizando-se as medidas *DE*, *CE* e entropia foi estabelecida a metodologia de busca dos MR que utiliza estas três medidas. O processo de múltiplas medidas procura explorar as particularidades de cada medida utilizada, de modo que haja uma complementaridade de propriedades que levem a uma melhor escolha dos MR representativos de uma dada cena e

j) Implementação do processo de triagem e busca dos MR em IDL[®]: Foi implementado um processador que realiza todas as etapas propostas e que pode ser empregado com relativa facilidade pelo usuário.

9.3 Avaliação do trabalho realizado

Todos os processos propostos neste trabalho, triagem de amostras, similaridade espectral pela entropia, busca de MR com múltiplas medidas e estimativa do número máximo de MR, baseiam-se em métodos de cálculos matemáticos/estatísticos. Entretanto, a seqüência de cálculo e a estrutura do processamento dos dados hiperespectrais, para se determinar um conjunto de MR, não está baseada em uma estrutura ótima global de cálculo, determinada a priori por algum método de otimização. A estrutura utilizada pelo processador implementado é formada de blocos de processamento que localmente procuram otimizar o seu desempenho, não garantindo uma otimização global de todo o processo.

A estrutura globalmente não é ótima, mas a estrutura concebida de certa maneira contribuiu para uma melhor compreensão das informações contidas nos complexos dados hiperespectrais. Essa compreensão poderá resultar em uma melhor utilização dos dados hiperespectrais e também na elaboração de uma teoria de processamento de imagens mais ampla, que leve a uma estrutura global otimizada de processamento.

Os processos desenvolvidos também podem ser considerados semi-automáticos ou semi-supervisionados, pois apesar de poderem processar quase que automaticamente as imagens hiperespectrais, na procura dos MR, exige que se forneça uma série de parâmetros (fatores e limiares) para que possa realizar os seus cálculos e as suas tarefas. Destes parâmetros dependerá o desempenho do processador e por esse motivo devem ser ajustados, em torno dos valores *default*, em função do tipo de imagem hiperespectral que se irá processar, em função do tipo de aplicação que se deseja realizar e em função do tipo de informação que se deseja obter dos dados hiperespectrais.

Em resumo o processo proposto e o processador implementado exigem perícia e intervenção humana. O processador e os métodos de cálculo utilizados visam auxiliar a tomada de decisão do operador humano, conhecedor do assunto e não substituí-lo.

9.4 Sugestões para continuidade do trabalho realizado

Os métodos propostos e o processador implementado não constituem uma obra acabada. São as proposições finais de um estudo realizado, que será continuado e aperfeiçoado. Dentre algumas tarefas que poderão dar continuidade ao trabalho realizado propõe-se:

- Testar o processador proposto com outras imagens AVIRIS e imagens de outros sensores imageadores hiperespectrais, avaliando o seu desempenho em relação ao ruído;
- Avaliar melhor o efeito do condicionamento espectral na fase de busca dos MR. Avaliar o uso do condicionamento espectral nos espectros submetidos ao processo de triagem discriminatória. Avaliar o uso dos espectros condicionados, escolhidos como MR, no MLME;

- Otimizar a interface do processador implementado em IDL[®] com usuários e integrá-lo ao aplicativo ENVI[®];
- Desenvolver um detector de bordas utilizando o conceito de uniformidade espacial;
- Trabalhar em métodos de estimação e ajustes automáticos dos parâmetros de processamento (fatores e limiares). Neste caso o emprego de imagens hiperespectrais simuladas podem auxiliar no teste dos estimadores.
- Procurar otimizar a estrutura de utilização conjunta das medidas *DE*, *CE* e entropia para a determinação dos MR;
- Incorporar novas medidas para utilização conjunta com a *DE*, *CE* e entropia na busca dos MR de uma cena e
- Otimizar a estrutura geral de processamento.

Referências Bibliográficas

ADAMS, J.B.; SMITH, M.O.; GILLESPIE, A.R. Imaging spectroscopy: interpretation based on spectral mixture analysis. In: PIETERS, C. M.; ENGLERT, P. **Remote Geochemical Analysis: elemental and mineralogical composition**. New York: Cambridge University Press, 1993. cap. 7, p.145-166

AIAZZI, B. et al. Information-theoretic assessment of sampled hyperspectral imagers. **IEEE Transactions on Geoscience and Remote Sensing**, v.39, n.7, p. 1447-1458, Jul. 2001.

ASNER, G.P. Biophysical and biochemical sources of variability in canopy reflectance. **Remote Sensing of Environment**, v. 64, n. 3, p. 234-253, Jun. 1998.

BATESON, C. A.; ASNER, G. P.; WESSMAN, C. A. Endmember bundles: A new approach to incorporating endmember. **IEEE Transactions on Geoscience and Remote Sensing**, v. 38, n. 2, p. 1083-1094, Mach 2000.

BATESON, A.; CURTISS, B. A method for manual endmember selection and spectral unmixing. **Remote Sensing of Environment**, v. 55, n. 1, p. 229-243, Jan. 1996.

BERGER, M. et al. The digital airborne imaging spectrometer experiment-DAISEX '99. In: INTERNATIONAL GEOSCIENCE AND REMOTE SENSING SYMPOSIUM, 2000, Honolulu. **Proceedings...** Honolulu: IEEE, 2000. v. 7, p. 3039-3041.

BIELSKI, C.M. et al. Spectral space: a new concept for information extraction from imaging spectrometer data. **International Journal of Remote Sensing**, v. 23, n. 10, p. 2005-2022, May 2002.

BIGHAM, J.M. et al. Iron oxide mineralogy of well-drained ultisols and oxisols: II. influence on color, surface area, and phosphate retention. **Soil Science Society of America Journal**, v. 42, n. , p. 825-830, 1978.

BIOUCAS-DIAS, J.M.; NASCIMENTO, J.M.P. Estimation of signal subspace on hyperspectral Data. In: BRUZZONE, L. (Ed.) **Image and Signal Processing for Remote Sensing XI**. Bellingham: SPIE, 2005. p. 191-198 (SPIE Proceedings, v. 5982).

BIRKHOFF, G.; MACLANE, S. **Álgebra moderna básica**. Rio de Janeiro: Guanabara Dois, 1977. 484 p.

BISTER, M.; CORNELIS, J.; ROSENFELD, A. A critical view of pyramid segmentation algorithms. **Pattern Recognition Letters**, v. 11, n. 9, p. 605-617, Sep.1990.

BOARDMAN, J.W; KRUSE, F.A.; GREEN, R.O. Mapping target signatures via partial unmixing of AVIRIS data. In: ANNUAL JPL AIRBORNE GEOSCIENCE WORKSHOP , 5., Pasadena, 1995. **Proceedings...** Pasadena: NASA, 1995. v. 1, p. 23-26. (JPL Publication 95-1).

BOARDMAN, J.W. Automating spectral unmixing of AVIRIS data using convex geometry concepts. In: ANNUAL JPL AIRBORNE GEOSCIENCE WORKSHOP, Pasadena, 1993. **Proceedings...** Pasadena: NASA, 1993. v. 1, p. 11-14. (JPL Publication 93-26).

BOARDMAN, J. W. Inversion of imaging spectrometry data using singular value decomposition. In: INTERNATIONAL GEOSCIENCE AND REMOTE SENSING SYMPOSIUM, 1989, Vancouver. **Proceedings...** Piscataway: IEEE, 1989. v. 4, p. 2069-2072.

BOREL, C.C.; GERSTL, S.A.W. Nonlinear spectral mixing models for vegetative and soil surfaces. **Remote Sensing of Environment**, v. 47, n. 3, p. 403-416, March 1994.

BOWLES, J. et al. Use of filter vectors in hyperspectral data analysis. In: STROJNIK, M.; ANDRESEN, B. F. **Infrared Spaceborne Remote Sensing III**. Bellingham: SPIE, 1995. p. 148-157 (SPIE Proceedings, v. 2553).

BURT, P.J.; ADELSON, A.E. The Laplacian pyramid as a compact image code. **IEEE Transactions on Communications**, v. COM-31, n. 4, p. 532-540, April 1983.

CARVALHO JR., O.A. DE; MENESES, P.R. Spectral correlation mapper (SCM): an improvement on the spectral angle mapper (SAM). In: JPL AIRBORNE EARTH SCIENCE WORKSHOP, 9., 2000, Pasadena. **Proceedings...** Pasadena: JPL, 2000. p. 65-74 (JPL Publication 00-18).

CHANG, C.-I; PLAZA, A. A fast iterative algorithm for implementation of Pixel Purity Index. **IEEE Geoscience and Remote Sensing Letters**, v. 3, n.1, p. 63-67, Jan. 2006.

CHANG, C.-I; DU, Q. Estimation of number of spectrally distinct signal sources in hyperspectral imagery. **IEEE Transactions on Geoscience and Remote Sensing**, v. 42, n. 3, p. 608-619, March 2004.

CHEN, J.M. Spatial scaling of a remotely sensed surface parameter by contexture. **Remote Sensing of Environment**, v. 69, n. 1, p. 30-42, July 1999.

CLARK, M.L.; ROBERTS, D.A.; CLARK, D.B. Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. **Remote Sensing of Environment**, v. 96, n. 3/4, p. 375-398, June 2005.

CLARK, R.N.; KING, T.V.V.; KLEJWA, M. High spectral resolution reflectance spectroscopy of minerals. **Journal of Geophysical Research**, v. 95, n. B8, p. 12653-12680, 1990.

COLLINS, E.F.; ROBERTS, D.A.; BOREL, C.C. Spectral mixture analysis of simulated thermal infrared spectrometry data: an initial temperature estimate bounded TESSMA search approach. **IEEE Transactions on Geoscience and Remote Sensing**, v.39, n.7, p.1435-1446, July 2001.

CONGALTON, R.G. A comparison of sampling schemes used in generating error matrices for assessing the accuracy of maps generated from remotely sensed data. **Photogrammetric Engineering and Remote Sensing**, v. 54, n. 5, p. 593-600, May 1988.

DALTON, J.B. et al. Spectral classification of similar materials using the tetracorder algorithm: the calcite-epidote-chlorite problem. In: AIRBORNE VISIBLE/INFRARED IMAGING SPECTROMETER (AVIRIS) WORKSHOP, 11., 2001, Pasadena, **Proceedings...** Pasadena: NASA, 2001.

DEBBA, P. et al. Abundance estimation of spectrally similar minerals by using derivative spectra in simulated annealing. **IEEE Transactions on Geoscience and Remote Sensing**, v. 44, n. 12, p. 3649-3658, Dec. 2006.

DEMETRIADES-SHAH, TANVIR H.; STEVEN, M.D.; CLARK, J.A. High resolution derivative spectra in remote sensing. **Remote Sensing of Environment**, v. 33, n. 1, p. 55-64, July 1990.

DRURY, S.A. **Image interpretation in geology**. London: Chapman & Hall, 1993. 283p.

DU, Y. et al. New hyperspectral discrimination measure for spectral characterization . **Optical Engineering**, v. 43, n. 8, p. 1777-1786, Aug. 2004.

DUDA, R.O.; HART, P.E.; STORK, D.G. **Pattern classification**. New York: A Willey-Interscience, 2001. XXII+654p.

ELMORE, A.J. et al. Quantifying vegetation change in semiarid environments: precision and accuracy of spectral mixture analysis and the normalized difference vegetation index. **Remote Sensing of Environment**, v. 73, n. 1, p. 87-102, July 2000.

ELVIDGE, C.D. Visible and near infrared reflectance characteristics of dry plants materials. **International Journal of Remote Sensing**, v. 11, n. 10, p. 1775-1759, 1990.

ELVIDGE, C.D.; MOUAT, D.A. Analysis of green vegetation detection limits in 1988 AVIRIS data. In: THEMATIC CONFERENCE ON REMOTE SENSING FOR EXPLORATION GEOLOGY, 7., 1989, Calgary. **Proceedings...** Ann Arbor: ERIM, 1989. p. 791-800.

ELVIDGE, C.D. Vegetation reflectance features in AVIRIS data. In: THEMATIC CONFERENCE ON REMOTE SENSING FOR EXPLORATION GEOLOGY, 6., 1988, Houston. **Proceedings...** Houston: ERIM, 1988a. p. 169-182.

ELVIDGE, C.D. Examination of the spectral features of vegetation in 1987 AVIRIS data. In: AIRBORNE VISIBLE/INFRARED IMAGING SPECTROMETER (AVIRIS) PERFORMANCE EVALUATION WORKSHOP, 1., 1988, Pasadena. **Proceedings...** Pasadena: NASA, 1988b. p. 97-101

ITT VISUAL INFORMATION SOLUTIONS. **ENVI**: tutorial data. Boulder: RSI. 2006. 2 CD-ROM.

EPIPHANIO, J.C.N. et al. **Comportamento espectral de solos do Estado de São Paulo**. São José dos Campos: INPE, junho 1992. 132p. (INPE-5424-PRP/172).

GALVÃO, S.L.; PIZARRO, M.A., EPIPHANIO, J.C.N. Variations in reflectance of tropical soils: spectral-chemical composition relationships from AVIRIS data. **Remote Sensing of Environment**, v. 75, n. 2, p. 245-255, Feb. 2001.

GALVÃO, L.S.; VITORELLO, I.; ALMEIDA FILHO, R. Effects of band positioning and bandwidth on NDVI measurements of tropical savannas. **Remote Sensing of Environment**. v. 67, n. 2, Feb. 1999.

GALVÃO, L.S.; VITORELLO, Í.; FORMAGGIO, A.R. Relationships of spectral reflectance and color among surface and subsurface horizons of tropical soil profiles. **Remote Sensing of Environment**, v. 61, n. 1, p. 24-33, July 1997.

GAO, B.-C.; HEIDEBRECHT, K. B.; GOETZ, A. F. H. Derivation of scaled surface reflectances from AVIRIS data. **Remote Sensing of Environment**, v. 44, n. 2-3, p. 165-178, May-June 1993.

GARCÍA-HARO, F. J.; SOMMER, S. A fast canopy reflectance model to simulate realistic remote sensing scenarios. **Remote Sensing of Environment**, v. 81, n. 2-3, p. 205-227, August 2002.

GATES, D.M. et al. Spectral properties of plants. **Applied Optics**, v. 4, n. 1, p. 11-20, Jan. 1965.

GAUCH, J.M.; PIZER, S.M. Multiresolution analysis of ridges and valleys in grey-scale images. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 15, n. 6, p. 635-646, June 1993.

GLENN, N.F. et al. Hyperspectral data processing for repeat detection of small infestations of leafy spurge. **Remote Sensing of Environment**, v. 95, n. 3, p. 399-412, April 2005.

GREEN, R.O.; BOARDMAN, J.W. Exploration of the relationship between information content and signal-to-noise ratio and spatial resolution in AVIRIS spectral data. In: JPL AIRBORNE EARTH SCIENCE WORKSHOP, 9., 2000, Pasadena. **Proceedings...** Pasadena: JPL, 2000.

GREEN, R.O.; PAVRI, B. AVIRIS in-flight calibration experiment, sensitivity analysis, and intraflight stability. In: AIRBORNE VISIBLE/INFRARED IMAGING SPECTROMETER (AVIRIS) WORKSHOP, 2000, Pasadena. **Proceedings...** Pasadena: NASA/JPL, 2000.

GREEN, R.O.; BOARDMAN, J.W. Exploration of the relationship between information content and signal-to-noise ratio and spatial resolution in AVIRIS spectral data. In: JPL AIRBORNE EARTH SCIENCE WORKSHOP, 9., 2000, Pasadena. **Proceedings...** Pasadena: JPL, 2000.

GREEN, R.O. Retrieval of reflectance from AVIRIS-measured radiance using a radiative transfer code. In: AIRBORNE VISIBLE/INFRARED IMAGING SPECTROMETER (AVIRIS) WORKSHOP, 3., 1991, Pasadena. **Proceedings...** Pasadena: NASA/JPL, 1991. p.200-210. (JPL Publication 91-28).

GREEN, A. A. et al. A transformation for ordering multispectral data in terms of image. **IEEE Transactions on Geoscience And Remote Sensing**, v. 26, n. 1, p. 65-74, Jan. 1988.

GROSS, H.N.; SCHOTT, J.R. Application of spectral mixture analysis and image fusion techniques for image sharpening. **Remote Sensing of Environment**, v.63, n. 2, p. 85-94, Feb. 1998.

GROSS, H.N.; SCHOTT, J.R. Evaluating an image fusion algorithm with synthetic image generation tools. In: IVERSON, A. Evan (Ed) **Algorithms for multispectral and hyperspectral imagery**. Bellingham: SPIE, 1996. p. 136-147 (SPIE Proceedings, v. 2758).

HAN, T. et al. Detection and correction of abnormal pixels in hyperion images. In: INTERNATIONAL GEOSCIENCE AND REMOTE SENSING SYMPOSIUM, 24., 2002, Ontario. **Proceeding...** Piscataway: IEEE, 2002. [CD-ROM]

HARSANYI, J.C.; CHANG, C.-I. Hyperspectral image dimensionality reduction and pixel classification: an orthogonal subspace projection approach. In: CONFERENCE ON INFORMATION SCIENCES AND SYSTEMS. 1993, Baltimore. **Proceedings...** Baltimore: The Johns Hopkins University, 1993. p. 401-406.

HEINZ, D. C.; CHANG, C.-I. Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery. **IEEE Transactions on Geoscience and Remote Sensing**, v. 39, n.3, p. 529-545, March 2001.

HOFFBECK, J.P.; LANDGREBE, D.A. Classification of remote sensing images having spectral resolution. **Remote Sensing of Environment**, v. 57, n. 3, p. 119-126, Sep. 1996.

HSIEH, P.-F.; LEE, L. C.; CHEN, N.-Y. Effect of spatial resolution on classification errors of pure and mixed pixels in remote sensing. **IEEE Transactions on Geoscience and Remote Sensing**, v. 39, n.12, Dec. 2001.

HSIEH, P.-F.; LANDGREBE, D. **Classification of high dimensional data**. V+121f. 1998. Tese (Doutorado) - Purdue University West Lafayette. (TR-ECE 98-4). Disponível em: <<http://dynamo.ecn.purdue.edu/~landgreb/Hsieh.TRECE4.pdf>> Acesso em: 05 abr 2006.

HUETE, A.R. Soil influences in remotely sensed vegetation-canopy spectra. In: Asrar, G.(ed.). **Theory and applications of optical remote sensing**. New York: John Wiley, 1989. Cap. 4, p. 107-141.

HUNT, G.R. Eletromagnetic radiation: the communication link in remote sensing. In: SIEGAL, D.S.; GILLESPIE, A.R (eds.). **Remote sensing in geology**. New York: John Wiley, 1980. Cap. 2, p. 5-45.

HUNT, G.R.; SALISBURY, J.W.; LENHOFF, C.J. Visible and near-infrared spectra of minerals and rocks: VI. Additional silicates. **Modern Geology**, v. 2, p. 85-106, 1973.

HUNT, G. R.; SALISBURY, J.W; LENHOFF, C.J. Visible and near-infrared spectra of minerals and rocks: III. Oxides and hydroxides. **Modern Geology**, v. 2, p. 195-205, 1971.

HUNT, G.R.; SALISBURY, J.W. Visible and near-infrared spectra of minerals and rocks: I. Silicate minerals. **Modern Geology**, v. 1, p. 283-300, 1970.

ICHOKU, C.; KARNIELI, A. A review of mixture modeling techniques for sub-pixel land cover estimation. **Remote Sensing Reviews**, v. 13, p. 161-186, 1996.

IFARRAGUERRI, A.; CHANG, C. I. Multispectral and hyperspectral image analysis with convex cones. **IEEE Transactions on Geoscience and Remote Sensing**, v. 37, n. 2, p. 756-770, March 1999.

IRONS, J. R.; WEISMILLER, R. A.; PETERSEN, G. W. Soil reflectance. In: ASRAR, G. **Theory and applications of optical remote sensing**. New York: John Wiley, 1989. Cap. 3, p. 66-106.

JACKSON, R. D.; SLATER, P. N.; PINTER JR., P. J. Discrimination of growth and water stress in wheat by various vegetation indices through clear and turbid atmospheres. **Remote Sensing of Environment**, v. 13, n. 3, p. 187-208, 1983.

JIA, X.; RICHARDS, J.A. Binary coding of imaging spectrometer data for fast spectral matching and classification. **Remote Sensing of Environment**, v. 43, n. 1, p. 47-53, Jan. 1993.

JIMENEZ, L. O. et al. Integration of spatial and spectral information by means of unsupervised extraction and classification for homogenous objects applied to multispectral and hyperspectral data. **IEEE Transactions on Geoscience and Remote Sensing**, v. 43, n. 4, p. 844-851, Apr. 2005.

JOHNSON, P.E. A semiempirical method for analysis of the reflectance spectra of binary mineral mixtures. **Journal of Geophysical Research**, v. 88, n. B4, p. 3557-3561, April 1983.

JONG, S. M. DE; HORNSTRA, T.; MAAS, H.-G. An integrated spatial and spectral approach to the Classification of Mediterranean land cover types: the SSC method. **International Journal of Applied Earth Observation and Geoinformation**, v. 3, n. 2, p. 176-183, Jan. 2001.

KARMANOV, I.I. Study of soils from the spectral composition of reflected radiation. **Soviet Soil Science**, n. 4, p. 34-37, 1970.

KATKOVNIK, V. SHMULEVICH, I. Kernel density estimation with adaptive varying window size. **Pattern Recognition Letters**, v. 23, n. 14, p. 1641-1648, Dec. 2002.

KESHAVA, N.; MUSTARD, J.F. Spectral unmixing. **IEEE Transactions on Geoscience and Remote Sensing**, v. 19, n.1, p. 44-57, Jan. 2002.

KNIPLING, E. B. Physical and physiological basis for the reflectance of visible and near-infrared radiation from vegetation. **Remote Sensing of Environment**, v. 1, n. 3, p. 155-159, 1970.

KOGER, C.H. et al. Wavelet Analysis of hyperspectral reflectance data for detecting Pitted Morningglory (*Ipomoea lacunosa*) in Soybean (*Glacine max*). **Remote Sensing of Environment**, v. 86, n. 1, p. 108-119, June 2003.

KRUSE, F. A. Spectral identification of image endmembers determined from AVIRIS data. In: JPL AIRBORNE EARTH SCIENCE WORKSHOP, 7., 1998, Pasadena. **Proceedings...** Pasadena: NASA, 1998.

KRUSE, F.A.; HUNTINGTON, J.F. The 1995 AVIRIS geology group shoot. In: ANNUAL JPL AIRBORNE EARTH SCIENCE WORKSHOP, 6., 1996, Pasadena. **Proceedings...** Pasadena: NASA/JPL, 1996. p. 155-164. (JPL Publication 96-1).

KUMAR, R. **Radiation from plants: reflection and emission: a review**. Purdue: Purdue University, 1972. (ES 72-02-02).

KUSTAS, W.P.; NORMAN, J.M. Evaluating the effects of subpixel heterogeneity on pixel average fluxes. **Remote Sensing of Environment**, v. 74, p. 327-342, 2000.

LACAPRA, V.C. et al. Remote sensing of foliar chemistry of inundated rice with imaging spectrometry. **Remote Sensing of Environment**, v. 55, n. 1, p. 309-336, Jan. 1996.

LANDGREBE, D. A. Multispectral land sensing: where from, where to?. **IEEE Transactions on Geoscience and Remote Sensing**, v. 43, n. 3, p. 414-421, March 2005.

LANDGREBE, D. Hyperspectral image data analysis. **IEEE Signal Processing Magazine**, v. 19, n. 1, p. 17-28, Jan. 2002.

LANDGREBE, D. A. The development of a spectral-spatial classifier for earth observational data. **Pattern Recognition**, v. 12, n. 1, p. 165-175, 1980.

LI, L.; USTIN, S.L.; LAY, M. Application of AVIRIS data in detection of oil-induced vegetation stress and cover change at Jornada, New México. **Remote Sensing of Environment**, v. 94, p. 1-16, 2005.

LILLESAND, T.M.; KIEFER, R.W. **Remote Sensing and Image Interpretation**. New York: John Wiley, 1994. 750 p.

LINDBERG, J.D.; SNYDER, D.G. Diffuse reflectance spectra of several clay minerals. **American Mineralogist**, v. 57, p. 485-493, March 1972.

MADHOK, V.; LANDGREBE, D. **Spectral-spatial analysis of remote sensing data: an image model and a procedural design**. 169f. 1999. Tese (Doutorado) Purdue University, West Lafayette. (TR-ECE 99-10). Disponível em: <http://dynamo.ecn.purdue.edu/~landgreb/Saldju_TR.pdf> Acesso em: 05 abr. 2006.

MANOLAKIS, D.; MARDEN, D.; SHAW, G.A. Hyperspectral image processing for automatic target detection applications. **Lincoln Laboratory Journal**, v. 14, n. 1, p. 79-116, 2003.

MARTIN, M.E.; ABER, J.D. High spectral resolution remote sensing of forest canopy lignin, nitrogen, and ecosystem processes. **Ecological Applications**, v. 6, n. 2, p. 431-443, 1997.

MATHEWS, H.L.; CUNNINGHAM, R.L.; PETERSEN, G.W. Spectral reflectance of selected Pennsylvania Soils. **Soil Science Society of America Proceedings**, v. 37, p. 421-424, 1973.

MOREIRA, R.C. **Influencia do posicionamento e da largura de bandas de sensores remotos e dos efeitos atmosfericos na determinacao de indices de vegetacao**. 179 f. 2000. Dissertação (Mestrado em Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos. Disponível em: <<http://mtc-m12.sid.inpe.br/rep-/sid.inpe.br/deise/2000/11.06.10.01>>. Acesso em: 17 jul. 2007.

NASCIMENTO, J. M. P.; DIAS, J. M. B. Vertex component analysis: a fast algorithm to unmix hyperspectral data. **IEEE Transactions on Geoscience and Remote Sensing**, v. 43, n. 4, p. 898-910, April 2005.

NASCIMENTO, P.S.R. **Avaliação de técnicas de segmentação e classificação por regiões em imagens Landsat-TM visando o mapeamento de unidades de paisagem na Amazônia**. São José dos Campos. 120f. 1997. Dissertação (Mestrado em Sensoriamento Remoto) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos.

NEVILLE, R.A. et al. Automatic endmember extraction from hyperspectral data for mineral exploration. In: Fourth International Airborne Remote Sensing Conference and Exhibition / 21st Canadian Symposium on Remote Sensing. **Proceedings**. Ottawa, Ontario, Canada, 21-24 June 1999. <ftp://ftp.nrcan.gc.ca/q294f0f669fn3536419s.pdf>> Acesso em: 10 maio 2006.

NOVAES, A.S.S. et al. **Pedology**. Radam Brasil Project, v. 31, Section SE.22, p. 413-576, 1983. (Technical Report).

NOVO, E.M.L. **Sensoriamento remoto: princípios e aplicações**. São Paulo: Edgar Blücher, 1992. 308 p.

PALMADESSO, P. et al. Use of filter vectors and fast convex set methods in hyperspectral analysis. In: INTERNATIONAL SYMPOSIUM ON SPECTRAL SENSING RESEARCH, 1995, Melbourne. **Proceedings...** Melbourne: ISSSR, 1995.

PERCIVAL, D.B.; WALDEN, A.T. **Wavelet methods for time series analysis**. Cambridge: Cambridge University Press, 2000. XXVI+594p.

PINZÓN, J.E.; USTIN, S.L.; PIERCE, J.F. Robust feature extraction for hyperspectral imagery using both spatial and spectral redundancies. In: JPL AIRBORNE EARTH SCIENCE WORKSHOP, 7., 1998, Pasadena. **Proceedings...** Pasadena: NASA, 1998.

PIZARRO, M.A.; FERNANDES, D. Seleção de amostras espectrais candidatas a membro de referência em imagens hiperespectrais. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 13., 2007a, Florianópolis. **Anais...** São José dos Campos: INPE, 2007. p. 6505-6512. CD-ROM.

PIZARRO, M.A.; FERNANDES, D. Determinação de membros de referência em dados hiperespectrais pelo método da máxima entropia. **Revista Brasileira de Cartografia**, 2007b. enviado em 19/03/2007.

PIZARRO, M.A.; FERNANDES, D. Avaliação de métodos de estimativa de coerência espectral para seleção de membros de referência em imagens hiperespectrais. **Revista Brasileira de Cartografia**, v. 57, n. 2, p. 113-120, ago. 2005a.

PIZARRO, M.A.; FERNANDES, D. Medida de coerência espectral na seleção de membros de referência em dados hiperespectrais, 12., 2005b, Goiânia. **Anais...** São José dos Campos: INPE, 2005a. p. 423-430. 1 CD-ROM.

PIZARRO, M.A.; GALVÃO, L.S.; EPIPHANIO, J.C.N. Caracterização espectral de solos tropicais e identificação mineral através de sensoriamento remoto hiperespectral. **Pesquisa Agropecuária Brasileira**, v. 36, n. 10, p. 1277-1286, out. 2001.

PIZARRO, M.A. **Sensoriamento remoto hiperespectral para a caracterização e identificação mineral em solos tropicais**. 1999 185 f. Dissertação (Mestrado em Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos. (INPE-7249-TDI/693). Disponível em: <<http://mtc-m12.sid.inpe.br/rep-sid.inpe.br/deise/1999/10.22.17.45>>. Acesso em: 18 nov 2004.

PLAZA, A.; CHANG. C.-I. Impact of initialization on design of endmember extraction algorithms. **IEEE Transactions on Geoscience and Remote Sensing**, v. 44, n. 11, p. 3397-3407, Nov. 2006.

PLAZA, A. et al. A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data. **IEEE Transactions on Geoscience and Remote Sensing**, v. 42, n. 3, p. 650-663, Mar. 2004a.

PLAZA, A. et al. A new approach to mixed pixel classification of hyperspectral imagery based on extended morphological profiles. **Pattern Recognition**, v. 37, n. 6, p. 1097-1116, 2004b.

PLAZA, A. et al. Spatial/spectral endmember Extraction by multidimensional morphological operations. **IEEE Transactions on Geoscience and Remote Sensing**, v. 40, n. 9, p. 2025-2041, Sept. 2002.

PORTER, W.M.; ENMARK, H.T. A system overview of the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS). In: VANE, G. (Ed) **Imaging Spectroscopy II**. Bellingham: SPIE, 1987. p.22-31 (SPIE Proceedings, v. 834).

ROBERTS, D.A. et al. Mapping chaparral in the Santa Monica mountains using multiple endmember spectral mixture models. **Remote Sensing of Environment**, v. 65, n. 3, p. 267-279, Sept. 1998.

ROBILA, S.A.; GERSHMAN, A. Spectral matching accuracy in processing hyperspectral data. In: INTERNATIONAL SYMPOSIUM ON SIGNALS, CIRCUITS AND SYSTEMS, 2005, Iasi. **Proceedings...** Iasi: IEEE, 2005. p. 163-166, v. 1 .

ROGGE, D.M.; RIVARD, B.; ZHANG, J.; SANCHEZ, A.; HARRIS, J.; FENG, J. Integration of spatial-spectral information for the improved extraction of endmembers. **Remote Sensing of Environment**, v. 110, n. 3, p. 287-303, Oct. 2007.

SAHOO, P.K.; SOLTANI, S.; WONG, A.K.C. A survey of thresholding techniques **Computer Vision Graphics and Image Processing**, v. 41, n.2, p. 233-260, 1988.

SCHOWENGERDT, R.A. **Remote Sensing: models and methods for image processing**. San Diego: Academic Press, 1997. XLV+522 p.

SHANMUGAN, K.S.; BREIPOHL, A.M. **Random signals: detection estimation and data analysis**. New York: John Wiley & Sons, 1988. XXVI+664p.

SHAW, G.; MANOLAKIS, D. Signal processing for hyperspectral image exploitation. **IEEE Signal Processing Magazine**, v. 19, n. 1, p. 12-16, Jan.. 2002.

SHIMABUKURO, Y. E.; SMITH, J. A. The least-squares mixing models to generate fraction images derived from remote sensing multispectral data. **IEEE Transactions on Geoscience and Remote Sensing**, v. 29, n. 1, p. 16-20, Jan. 1991.

SMITH, M.O.; USTIN, S.L.; ADAMS, J.B.; GILLESPIE, A.R. Vegetation in deserts: I. a Regional Measure of abundance from multispectral images. **Remote Sensing of Environment**, v. 31, p. 1-26, Jan. 1990.

STAENZ, K.; SZEREDI, T.; SCHWARZ, J. ISDAS – A system for processing/analyzing hyperspectral data. **Canadian Journal of Remote Sensing**, v. 24, p. 99-113, 1998.

STEFFEN, C.A.; MORAES, E.C.; GAMA, F.F. Radiometria óptica espectral. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 8., 1996, Salvador. **Anais...** 1996. Salvador: INPE, 1996. Disponível em: <<http://iris.sid.inpe.br:1908/rep-sid.inpe.br/iris@1908/2005/05.25.14.28>>. Acesso em: 09 fev. 2006.

STEIN, D.W. Anomaly detection from hyperspectral imagery. **IEEE Signal Processing Magazine**, v. 19, n.1, p. 58-69, Jan. 2002.

STONER, E.R. et al. **Atlas of soil reflectance properties**. West Lafayette: Purdue University, 1980. 75 p. (Research Bulletin 962).

STONER, E.R.; BAUMGARDNER, M.F. **Physicochemical, site, and bidirectional reflectance factor characteristics of uniformly moist soils**. West Lafayette: Purdue University, 1980. (LARS Technical Report 111679).

STRANG, G.; BORRE, K. **Linear algebra, geodesy, and GPS**. Wellesley: Wellesley-Cambridge Press, 1997. XVI+624 p.

STRICKLAND, R.; HAHN, H.I. Wavelet transform methods for object detection and recovery. **IEEE Transactions on Image Processing**, v. 6, n. 5, p. 724-735, May 1997.

THEILER, J. et al. Using blocks of skewers for faster computation of Pixel Purity Index In: DESCOUR, M. R.; SHEN, S. S. **Imaging Spectrometry VI**. Bellingham: SPIE, 2000, p. 61-71 4132

TSAI, F.; PHILPOT, W.D. A derivative-aided hyperspectral image analysis system for land-cover classification. **IEEE Transactions on Geoscience and Remote Sensing**, v. 40, n. 2, p. 416-425, Feb.2002.

TOMPKINS, S. et al. Optimization of endmembers for spectral mixture analysis. **Remote Sensing of Environment**, v. 59, p. 472-489, 1997.

TU, TE-M. et al. An oblique subspace projection approach for mixed pixel classification in hyperspectral images. **Pattern Recognition**, v.32, n. 8; p. 1399-1408, 1999.

VANE, G.; GREEN, R.O; CHRIEN, T.G; ENMARK, H.T.; HANSEN, E.G.; PORTER, W.M. The Airborne Visible/Infrared Imaging Spectrometer (AVIRIS). **Remote Sensing of Environment**, v. 44, n. 2/3, p. 127-143, May/June 1993.

VERHOEYE, J.; DE WULF, R. Land cover mapping at sub-pixel scales using linear optimization techniques. **Remote Sensing of Environment**, v. 79, n. 1, p. 96-104, Jan. 2002.

WILKINSON, G.G. Results and implications of a study of fifteen years of satellite image classification experiments. **IEEE Transactions on Geoscience and Remote Sensing**, v. 43, n. 3, p. 433-440, Mar. 2005.

WINTER, M.E. A proof of the N-FINDR algorithm for the automated detection of end-members in a hyperspectral image. In: SHEN, S. S.; LEWIS, P. E. (Ed) **Algorithms and technologies for multispectral, hyperspectral and ultra spectral imagery X**. Bellingham: SPIE, 2004. p. 31-41 (SPIE Proceedings, v. 5425).

WINTER, M.E. N-FINDR: an algorithm for fast autonomous spectral end member determination in hyperspectral data. In: DESCOUR, M. R.; SHEN, S. S. (Ed) **Imaging Spectrometry V**. Bellingham: SPIE, 1999. p. 266-275 (SPIE Proceedings, v. 3753).

ZARCO-TEJADA, P. J. et al. Scaling-up and model inversion methods with narrowband optical indices for chlorophyll content estimation in closed forest canopies with hyperspectral data. **IEEE Transactions on Geoscience and Remote Sensing**, v. 39, n. 7, p. 1491-1507, Jul 2001.

ZHANG, J.; RIVARD, B.; SANCHEZ-AZOFEIFA, A. Derivative spectral unmixing of hyperspectral data applied to mixtures of lichen and rock. **IEEE Transactions on Geoscience and Remote Sensing**, v. 42, n. 9, p. 1934-1940, Sep. 2004.

ZHANG, L. et al. Study of the spectral mixture model of soil and vegetation in PoYang lake area, China. **International Journal of Remote Sensing**, v. 19, n. 11, p. 2077-2084, July 1998.

APÊNDICE A – Estrutura e Arquivos do Processador para Triagem e Busca de MR

Apresenta-se a estrutura do programa de cálculo, implementado com o aplicativo computacional IDL[®], que além da triagem e condicionamento dos espectros realiza a busca dos MR utilizando múltiplas medidas. Mostram-se exemplos dos arquivos de entrada e de saída relativos ao programa de busca dos MR implementado em IDL[®]. O exemplo refere-se à busca de cinco MR dentre 32 amostras da subcena 7, selecionadas por um especialista no processo de triagem manual. As amostras foram submetidas ao processo de triagem discriminatória (testes de uniformidade espacial, homogeneidade espectral e redundância) e tiveram os seus espectros condicionados pela derivada.

A.1 - Estrutura básica do processador para busca dos MR

Foi desenvolvido com o aplicativo IDL[®] (Iterative Data Language) da RSI[®] (Research System Inc.) um processador para a triagem de amostras e busca dos MR em imagens hiperespectrais. Este processador tem como entradas:

- um conjunto de imagens hiperespectrais;
- parâmetros, fatores e limiares de configuração que definem etapas de processamento e
- dados e comandos para execução das tarefas.

As imagens hiperespectrais utilizadas como entrada devem estar com os valores de reflectância de superfície e os parâmetros, fatores e limiares de configuração são aqueles definidos neste capítulo e nos Capítulos 4, 5 e 6.

a) parâmetros, fatores e limiares na etapa de triagem das amostras:

- M : dimensões da vizinhança espacial inicial $(M + 1) \times (M + 1)$ de uma amostra espectral;
- α_u : fator de uniformidade espacial, $0,5 < \alpha_u \leq 1$;

- $\alpha_u (M + 1)^2$: limiar de uniformidade espacial;
- ψ_e : limiar de seleção espacial, $0 \leq \psi_e \leq 1$;
- ψ_h : limiar de homogeneidade espectral, $0,5 < \psi_h \leq 1$;
- α : nível de significância do teste estatístico para verificação da igualdade de médias espectrais, banda a banda, entre as vizinhanças $V_0(.,.)$ e $V_1(.,.)$;
- M_c : dimensões da vizinhança espacial $(M_c + 1) \times (M_c + 1)$ para o teste do contexto;
- α_c fator de contexto espacial, $0 \leq \alpha_c \leq 1$;
- $\alpha_c (M_c + 1)^2$: limiar do teste de contexto espacial;
- ψ_{rde} : limiar de redundância para a medida DE , $0 \leq \psi_{rde} \leq 1$;
- ψ_{rce} : limiar de redundância para a medida CE , $0 \leq \psi_{rce} \leq 1$ e
- \mathfrak{R}_∇ : conjunto de amostras menos redundantes baseado:
 - na intersecção dos testes de redundância realizados com a DE e CE ($\mathfrak{R}_\nabla = \mathfrak{R}_\cap$);
 - na união dos testes de redundância realizados com a DE e CE ($\mathfrak{R}_\nabla = \mathfrak{R}_\cup$);
 - no teste de redundância realizado com a DE ($\mathfrak{R}_\nabla = \mathfrak{R}_{DE}$) e
 - no teste de redundância realizado com a CE ($\mathfrak{R}_\nabla = \mathfrak{R}_{CE}$).

b) fatores e limiares na etapa de busca dos MR:

- R : número de MR a serem procurados, $R \geq 2$;
- η_H : é o limiar de configuração das amostras candidatas a MR segundo a entropia;

- η_{DE} : limiar de configuração das amostras candidatas a MR segundo a *DE*;
- η_{CE} : limiar de configuração das amostras candidatas a MR segundo a *CE*;
- α_H : fator de configuração das amostras candidatas a MR segundo a entropia, $0 \leq \alpha_H \leq 1$;
- α_{DE} : fator de configuração das amostras candidatas a MR segundo a *DE*, $0 \leq \alpha_{DE} \leq 1$ e
- α_{CE} : fator de configuração das amostras candidatas a MR segundo a *CE*, $0 \leq \alpha_{CE} \leq 1$.

c) parâmetros e limiares na etapa de estimação do número de MR:

- \mathfrak{R}_1 : limitante superior do número de MR baseado no número máximo de MR bem configurados, $R \leq \mathfrak{R}_1$;
- \mathfrak{R}_2 : limitante superior do número de MR baseado em um valor de entropia mínima H_{\min} , $R \leq \mathfrak{R}_2$ e
- H_{\min} : limiar para o valor da entropia, $\Delta_H(e_1, e_2, \dots, e_{\mathfrak{R}_2}) \geq H_{\min}$;

Os dados e comandos para a execução das tarefas são:

- tipo de triagem (manual, aleatória ou discriminatória). No caso de triagem manual devem ser informadas as posições das amostras selecionadas pelo operador em forma de arquivo. O processador desenvolvido pode ainda oferecer ao especialista, que irá escolher as amostras, um aplicativo que possibilita a escolha das amostras candidatas a MR clicando com o *mouse* sobre as amostras em uma composição colorida RGB . Para o caso do AVIRIS foi escolhida para a imagem colorida a composição R (Banda 22, ~677 nm), G (Banda 10, ~549 nm) e B (Banda 3, ~481 nm). Na Triagem manual o processador desenvolvido permite ainda que sejam atribuídas diversas amostras associadas a cada uma das classes e subclasses escolhidas pelo especialista. Neste caso, na procura dos MR

é feita com a restrição de que não haja mais de uma amostra de uma classe ou subclasse no conjunto dos MR escolhidos. Esta funcionalidade amplia o conceito de amostras bem configuradas, considerando como mal configurado o conjunto de amostras que tenha mais de uma amostra de cada classe ou subclasse previamente definida. No caso da amostragem aleatória deve ser especificado o número de elementos da quadricula, horizontal e vertical, para que seja realizada a AESNA. Por exemplo, se forem escolhidos (30 x 20) elementos da quadricula serão geradas $30 \times 20 = 600$ amostras candidatas a MR uniformemente espalhadas pela imagem. Há a opção para que além da escolha manual do operador seja feita também a escolha de amostras pela AESNA (triagem mista). No caso de triagem discriminatória é feita automaticamente a seleção das amostras candidatas a MR dentre todas as amostras espectrais da cena;

- etapas da triagem das amostras ou pré-processamento dos dados (uniformidade espacial, homogeneidade espectral, redundância e contexto espacial). Pode-se optar por um ou mais tipos de etapas de triagem. No caso da redundância deve-se escolher o conjunto de amostras menos redundantes: \mathfrak{R}_{DE} ; \mathfrak{R}_{CE} ; \mathfrak{R}_{\cup} ou \mathfrak{R}_{\cap} e
- modo de Busca dos MR (critério da *DE*, *CE*, entropia ou medidas conjuntas: baseadas na $(e^f_1, e^f_2, \dots, e^f_R)$, $(e^2_1, e^2_2, \dots, e^2_R)$, ou $(e^3_1, e^3_2, \dots, e^3_R)$).

As saídas do processador são:

- relatório na forma de um arquivo de texto;
- arquivos de imagens e figuras e
- arquivo com os espectros escolhidos como MR.

No relatório constam dados e parâmetros fornecidos e calculados pelo processador. No arquivo de imagens são fornecidos os espectros escolhidos como MR e a posição, em uma imagem em tons de cinza (Banda 192, ~ 2.271 nm), das amostras escolhidas como MR.

A.2 - Arquivo de parâmetros de entrada

Mostra-se a seguir um exemplo do arquivo de dados de entrada contendo o endereço de diretórios, nomes de arquivos, parâmetros, fatores, limiares e comandos para a execução das tarefas.

Escolheu-se a subcena 7, triagem manual com 32 amostras, testes discriminatórios envolvendo a uniformidade espacial, homogeneidade espectral e redundância com o conjunto intersecção \mathfrak{R}_\cap e os parâmetros: vizinhança $(M+1) \times (M+1) = (5 \times 5)$, fator de uniformidade espacial $\alpha_u = 60\%$, limiar de seleção espacial $\psi_e = 78\%$, fator de homogeneidade $\alpha_h = 90\%$, nível de significância $\alpha = 10\%$ no teste de igualdade das médias, limiar de redundância para a medida DE $\psi_{rde} = 20\%$, limiar de redundância para a medida CE $\psi_{rce} = 20\%$, fator de configuração das amostras candidatas a MR segundo a DE $\alpha_{DE} = 32\%$, fator de configuração das amostras candidatas a MR segundo a CE $\alpha_{CE} = 30\%$ e fator de configuração das amostras candidatas a MR segundo a entropia $\alpha_H = 30\%$.

Arquivo de dados de entrada nomeado como:

F:\pizarro\arquivos\hiper_in\sub4.txt.

```
*****
* PARAMETER FILE FOR HIPERSPECTRAL PROCESSOR V3.0 / July-2007 *
*
* ITA/IEE/IEET *
* Marco Antonio Pizarro (ITA-INPE) e David Fernandes (ITA) *
* 22-Jun-2007 16:48:20.00 *
*****

A. DEFINITION OF DIRECTORIES

1. INPUT HIPERSPECTRAL DATA DIRECTORY [\...\...\]: f:\pizarro\arquivos\entradas\
2. HIPERSPECTRAL DATA NAME [\...\...\]: cg_7_25_0.3n_resize.rfl.dat_interpol_209.dat
3. INPUT DIRECTORY [\...\...\]: f:\pizarro\arquivos\hiper_in\
4. OUTPUT DATA DIRECTORY [\...\...\]: f:\pizarro\arquivos\hiper_out\
5. ROOTNAME [Ex.: su4]: sub7
```

B. DEFINITION OF PARAMETERS

1. NUMBER OF IMAGES [209...?]: 209
2. IMAGE DIMENSION - NUMBER OF SAMPES [612...?]: 614
3. IMAGE DIMENSION - NUMBER OF LINES [512...?]: 512
4. ENDMEMBER SEARCH WINDOW [3, 5, 7...?]: 5

C. DEFINITION OF SELECTION SAMPLES TASKS

1. SELECTION TYPE [manual, aesna, mixed, discrimination?]: manual
2. MANUAL:
 - 2.1. MANUAL SELECTION TYPE [defined, assisted?]: defined
 - 2.2. INPUT MANUAL DATA FILE DIRECTORY [\\...\\...\\]: f:\pizarro\arquivos\hiper_in\
 - 2.3. MANUAL DATA FILE NAME [\\...\\...\\]: sub7_arquivo_36_amostras.txt
3. AESNA:
 - 3.1. ELEMENTS IN THE IMAGE - SAMPLES DIMENSION [2, 3,...?]: 5
 - 3.2. ELEMENTS IN THE IMAGE - LINES DIMENSION [2, 3,...?]: 5
4. MIXED (MANUAL AND AESNA):
 - 4.1. ELEMENTS IN THE IMAGE - SAMPLES DIMENSION [2, 3,...?]: 5
 - 4.2. ELEMENTS IN THE IMAGE - LINES DIMENSION [2, 3,...?]: 5
 - 4.3. INPUT MANUAL DATA FILE DIRECTORY [\\...\\...\\]: f:\pizarro\arquivos\hiper_in\
 - 4.4. MANUAL DATA FILE NAME [\\...\\...\\]: sub4_arquivo_21_amostras.txt
5. DISCRIMINATION:
 - 5.1. DISCRIMINATION TYPE [manual, aesna, mixed, image?]: manual
 - 5.2. MANUAL OR MIXED: INPUT MANUAL DATA FILE DIRECTORY [\\...\\...\\]:
f:\pizarro\arquivos\hiper_in\
 - 5.3. MANUAL OR MIXED: MANUAL DATA FILE NAME [\\...\\...\\]:
sub4_arquivo_21_amostras.txt
 - 5.4. AESNA OR MIXED: ELEMENTS IN THE IMAGE - SAMPLES DIMENSION [2, 3, ...?]: 5
 - 5.5. AESNA OR MIXED: ELEMENTS IN THE IMAGE - LINES DIMENSION [2, 3, ...?]: 5

D. SAMPLES SELECTION PARAMETERS

1. SPACIAL UNIFORMITY [yes, no?]: yes
 - 1.1. SPACIAL UNIFORMITY THRESHOLD [0-1?]: 0.780000
 - 1.2. SPACIAL UNIFORMITY FACTOR [0-1?]: 0.600000
2. SPECTRAL HOMOGENEITY [yes, no?]: yes
 - 2.1. HOMOGENEITY FACTOR [0-1?]: 0.900000
 - 2.2. SIGNIFICANCE LEVEL [0.001-0.1?]: 0.100000
3. REDUNDANCY [yes, no?]: yes
 - 3.1. EUCLIDEAN DISTANCE THRESHOLD [0-1?]: 0.200000
 - 3.2. SPECTRAL COERENGE THRESHOLD [0-1?]: 0.200000
 - 3.3. REDUNDANCY MODE [de, ce, union, inter?]: inter
4. SPACIAL CONTEXT [yes, no?]: yes
 - 4.1. CONTEXT WINDOW [3, 5, 7,...?]: 5
 - 4.2. CONTEXT FACTOR [1 - 1?]: 0.800000

E. SPECTRAL PREPARING

1. PREPARE TYPE [none, derivate, DWT?]: derivate
2. DWT: TYPE [Haar, Db2-10]: Haar

F. MEMBER OF REFERENCE (MR) SEARCH

1. EUCLIDEAN DISTANCE FACTOR [0-1?]: 0.320000
2. SPECTRAL COERENGE FACTOR [0-1?]: 0.300000
3. ENTROPY FACTOR [0-1?]: 0.300000
4. SEARCH TYPE [Euclidean, coherence, vote, two, three?]: three

A.3 - Arquivo de dados das amostras escolhidas pela triagem manual

Mostra-se a seguir o arquivo de dados relativo a 32 amostras escolhidas por um especialista, no processo de triagem manual, referenciado no arquivo de parâmetros de entrada da seção anterior. Neste arquivo, após a sexta linha, há uma seqüência de linhas com quatro elementos formados por três números e uma palavra alfanumérica. Os dois primeiros números correspondem às coordenadas das amostras na imagem, o terceiro número refere-se

ao agrupamento das amostras (amostras de um mesmo grupo tem o mesmo valor) e o quarto elemento é a informação referente ao nome da amostra dada pelo especialista.

No caso apresentado o processo de busca dos MR é feita com a restrição de que duas ou mais amostras pertencentes ao mesmo grupo (terceiro número igual) não podem ser escolhidas como MR.

Arquivo de dados nomeado como:

F:\pizarro\arquivos\entradas\ sub7_arquivo_72_amostras.txt.

cena: F:\pizarro\arquivos\entradas\cg_7_25_0.3n_resize.rfl.dat_interpol_209.dat

subcena: sub7

linhas: 512

amostras: 614

dimensao: 209

candidatas: 32

308	57	0	vv1(a)
316	54	0	vv1(b)
290	51	0	vv1(c)
305	50	0	vv1(d)
297	211	4	vv2(a)
347	39	4	vv2(b)
317	203	4	vv2(c)
481	213	4	vv2(d)
417	455	8	vv3(a)
466	385	8	vv3(b)
427	450	8	vv3(c)
421	450	8	vv3(d)
570	303	12	vs(a)
575	301	12	vs(b)
432	347	12	vs(c)
434	355	12	vs(d)

497	81	16	a(a)
497	85	16	a(b)
498	81	16	a(c)
388	342	19	saq(a)
377	335	19	saq(b)
387	343	19	saq(c)
394	344	19	saq(d)
81	476	23	sle(a)
87	475	23	sle(b)
68	487	23	sle(c)
81	461	23	sle(d)
116	153	27	ste(a)
107	159	27	ste(b)
109	140	27	ste(c)
110	140	27	ste(d)
348	15	27	ste(e)

A.4 - Relatório de saída do programa

Mostra-se a seguir o arquivo relativo ao relatório do programa de busca de MR referente ao arquivo de dados de entrada, constante na Seção A.1, e ao arquivo de amostras da Seção A.2. No final deste relatório estão os MR escolhidos.

No exemplo, das 32 amostras iniciais o processo de triagem selecionou nove amostras e o programa de busca dos MR selecionou as cinco melhores amostras, segundo os critérios apresentados nos Capítulos 6 e 7. Os critérios de busca de MR baseados na DE, CE, entropia e medida conjunta apresentaram os mesmos resultados. Foram escolhidos como MR as amostras: $vv2(b)$, $vs(d)$, $a(a)$, $sle(d)$ e $ste(d)$. Na nomenclatura utilizada neste trabalho os solos sle e ste correspondem respectivamente a LV e NV .

Relatório do Programa de Busca de MR:

*versão 3.0 ---- data 18/06/2007

Seleção dos membros de referência em imagens hiperespectrais

Marco Antonio Pizarro (INPE - ITA) David Fernandes (ITA)

Amostras candidatas a MR:

- 0.... 308.x. 57... 0...vv1(a)
- 1.... 316.x. 54... 0...vv1(b)
- 2.... 290.x. 51... 0...vv1(c)
- 3.... 305.x. 50... 0...vv1(d)
- 4.... 297.x. 211... 4...vv2(a)
- 5.... 347.x. 39... 4...vv2(b)
- 6.... 317.x. 203... 4...vv2(c)
- 7.... 481.x. 213... 4...vv2(d)
- 8.... 417.x. 455... 8...vv3(a)
- 9.... 466.x. 385... 8...vv3(b)
- 10.... 427.x. 450... 8...vv3(c)
- 11.... 421.x. 450... 8...vv3(d)
- 12.... 570.x. 303... 12...vs(a)
- 13.... 575.x. 301... 12...vs(b)
- 14.... 432.x. 347... 12...vs(c)
- 15.... 434.x. 355... 12...vs(d)
- 16.... 497.x. 81... 16...a(a)
- 17.... 497.x. 85... 16...a(b)
- 18.... 498.x. 81... 16...a(c)
- 19.... 388.x. 342... 19...saq(a)
- 20.... 377.x. 335... 19...saq(b)
- 21.... 387.x. 343... 19...saq(c)
- 22.... 394.x. 344... 19...saq(d)
- 23.... 81.x. 476... 23...sle(a)
- 24.... 87.x. 475... 23...sle(b)

25.... 68.x. 487... 23...sle(c)

26.... 81.x. 461... 23...sle(d)

27.... 116.x. 153... 27...ste(a)

28.... 107.x. 159... 27...ste(b)

29.... 109.x. 140... 27...ste(c)

30.... 110.x. 140... 27...ste(d)

31.... 348.x. 15... 27...ste(e)

Numero de amostras candidatas a MR = 32

Uniformidade Espacial:

Espectros de cada amostra selecionados pela filtragem espacial:

0 vv1(a) 25

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

1 vv1(b) 25

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

2 vv1(c) 25

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

3 vv1(d) 25

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

4 vv2(a) 25

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

5 vv2(b) 24

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23

6 vv2(c) 25

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

7 vv2(d) 22

0 1 2 3 4 5 6 7 8 9 11 12 13 14 16 17 18 19 21 22 23 24

8 vv3(a) 25

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

9 vv3(b) 25

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

10 vv3(c) 25

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

11 vv3(d) 25

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

12 vs(a) 25

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

13 vs(b) 25

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

14 vs(c) 25

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

15 vs(d) 25

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

16 a(a) 23

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 21 23 24

17 a(b) 18

0 1 2 5 6 7 8 10 11 12 13 14 15 16 17 18 20 21

18 a(c) 19

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 18 19

19 saq(a) 17

0 1 4 5 6 10 11 12 15 16 17 18 20 21 22 23 24

20 saq(b) 20

0 1 3 4 5 6 7 8 9 10 11 12 15 16 17 20 21 22 23 24

21 saq(c) 23

0 1 2 3 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

22 saq(d) 25

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

23 sle(a) 25

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

24 sle(b) 25

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

25 sle(c) 25

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

26 sle(d) 25

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

27 ste(a) 24

0 1 2 3 4 5 6 7 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

28 ste(b) 25

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

29 ste(c) 25

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

30 ste(d) 25

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

31 ste(e) 25

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

Uniformidade Espacial:

seqüência de amostras selecionadas pela filtragem espacial (melhor para a pior):

ste(e) ste(d) ste(c) ste(b) vs(c) sle(d) sle(c) sle(b) sle(a) saq(d) vv1(b)

vv1(c) vv1(d) vv2(a) vv2(c) vv1(a) vs(d) vv3(a) vs(b) vs(a) vv3(d) vv3(c) vv3(b)

ste(a) vv2(b) a(a) saq(c) vv2(d) saq(b) a(c) a(b) saq(a)

amostras selecionadas pela uniformidade espacial 32 de 32

vv1(a) vv1(b) vv1(c) vv1(d) vv2(a) vv2(b) vv2(c) vv2(d) vv3(a) vv3(b) vv3(c)

vv3(d) vs(a) vs(b) vs(c) vs(d) a(a) a(b) a(c) saq(a) saq(b) saq(c) saq(d)

sle(a) sle(b) sle(c) sle(d) ste(a) ste(b) ste(c) ste(d) ste(e)

Teste da homogeneidade:

Estatística t-student das amostras:

0.928230 0.918660 0.952153 0.942584 0.947368 0.956938

0.956938 1.00000 0.933014 1.00000 0.995215 1.00000

0.995215 1.00000 0.961722 0.990431 0.976077 1.00000

0.990431 0.942584 0.885167 0.746412 0.995215 0.727273

1.00000 0.947368 0.995215 1.00000 0.985646 0.880383

0.995215 0.956938

seqüência de amostras quanto a homogeneidade:

sle(b) vv3(b) vv3(d) vv2(d) a(b) ste(a) vs(b) vv3(c) sle(d) saq(d) ste(d) vs(a)

vs(d) a(c) ste(b) a(a) vs(c) ste(e) vv2(c) vv2(b) vv1(c) sle(c) vv2(a) saq(a)

vv1(d) vv3(a) vv1(a) vv1(b) saq(b) ste(c) saq(c) sle(a)

amostras mais homogêneas:

amostras selecionadas 28 de 32

vv1(a) vv1(b) vv1(c) vv1(d) vv2(a) vv2(b) vv2(c) vv2(d) vv3(a) vv3(b) vv3(c)

vv3(d) vs(a) vs(b) vs(c) vs(d) a(a) a(b) a(c) saq(a) saq(d) sle(b) sle(c)

sle(d) ste(a) ste(b) ste(d) ste(e)

Teste da Redundância,

diferença das distâncias:

15.3162 7.59439 7.22809 10.1304 47.6263 24.0374

8.91099 10.7238 10.9831 3.22050 2.81520 2.02432

0.733440 13.5640 15.6647 3.37343 10.4521 12.4616

9.36978 9.32137 8.04913 7.75711 4.20771 0.135730

35.9953 1.12104 0.506448

vv2(b) vv2(a) vv2(c) vv2(d) ste(e) ste(d) ste(b) ste(a) sle(d) sle(c) vv3(a) sle(b) vv3(d) vv3(c) vv3(b)

vs(c) vs(d) vv1(c) vv1(a) vv1(b) vv1(d) vs(b) vs(a)

saq(d) saq(a) a(b) a(c) a(a)

amostras eliminadas: vv1(a) vv1(b) vv1(d) vv2(a) vv2(c) vv3(a) vv3(d) vs(a) vs(b) vs(c) a(b) a(c)

saq(d) sle(b) sle(c) ste(b)

amostras resultantes: x x vv1(c) x x vv2(b) x vv2(d) x vv3(b) vv3(c) x x x x vs(d) a(a) x x saq(a) saq(b)

saq(c) x sle(a) x x sle(d) ste(a) x ste(c) ste(d) ste(e)

diferença das coerências;

42.7271 21.9036 16.0882 6.24891 13.3048 9.55246

27.9327 9.63176 8.26768 0.515242 10.4137 51.1911

0.0665051 11.4110 10.0970 11.2147 20.1867 38.6117

19.7575 2.06356 59.6698 8.29614 2.66195 2.11987

56.7637 14.5243 4.30758

vv3(b) vs(a) vv2(b) vv2(a) vs(b) vv2(c) vs(d) vs(c) vv3(d) vv3(c) vv2(d) vv3(a) ste(a) ste(e) ste(b)

ste(d) saq(d) saq(a) sle(c) sle(b) sle(d) vv1(c) vv1(a) vv1(b) vv1(d) a(b) a(c) a(a)

amostras eliminadas: vv1(a) vv1(b) vv1(c) vv2(a) vv2(c) vv3(c) vv3(d) vs(c) a(c) sle(b) ste(a)

amostras resultantes: x x x vv1(d) x vv2(b) x vv2(d) vv3(a) vv3(b) x x vs(a) vs(b) x vs(d) a(a) a(b) x

saq(a) saq(b) saq(c) saq(d) sle(a) x sle(c) sle(d) x ste(b) ste(c) ste(d) ste(e)

seqüência de -- 12 -- amostras menos redundantes segundo o critério da distância:

critério de aceitação: variação maior que - 10.0000% __ na distância

vv1(c) vv2(b) vv2(d) vv3(b) vv3(c) vs(d) a(a) saq(a) sle(d) ste(a) ste(d) ste(e)

seqüência de -- 17 -- amostras menos redundantes segundo o critério da coerência:

critério de aceitação: variação maior que - 10.0000% __ na coerência

vv1(d) vv2(b) vv2(d) vv3(a) vv3(b) vs(a) vs(b) vs(d) a(a) a(b) saq(a) saq(d)

sle(c) sle(d) ste(b) ste(d) ste(e)

seqüência de -- 20 -- amostras menos redundantes segundo o critério união dis e coer:

critério de aceitação: variação maior que - 10.0000% __ na distância

critério de aceitação: variação maior que - 10.0000% __ na coerência

vv1(c) vv1(d) vv2(b) vv2(d) vv3(a) vv3(b) vv3(c) vs(a) vs(b) vs(d) a(a) a(b)

saq(a) saq(d) sle(c) sle(d) ste(a) ste(b) ste(d) ste(e)

seqüência de -- 9 -- amostras menos redundantes segundo o critério intersecção dis e coer:

critério de aceitação: variação maior que - 10.0000% __ na distância

critério de aceitação: variação menor que - 10.0000% __ na coerência

vv2(b) vv2(d) vv3(b) vs(d) a(a) saq(a) sle(d) ste(d) ste(e)

Resumo:

número inicial de amostras candidatas a MR = 32

número de amostras homogêneas candidatas a MR = 28

novo número de amostras candidatas a MR (critério: inter) = 9

Novas amostras candidatas a MR (critério: inter)

0.... 347.x. 39... 0...vv2(b)

1.... 481.x. 213... 0...vv2(d)

2.... 466.x. 385... 2...vv3(b)

3.... 434.x. 355... 3...vs(d)

4.... 497.x. 81... 4...a(a)

5.... 388.x. 342... 5...saq(a)

6.... 81.x. 461... 6...sle(d)

7.... 110.x. 140... 7...ste(d)

8.... 348.x. 15... 7...ste(e)

Novo número de amostras candidatas a MR (critério: inter) = 9

Entropia duas a duas:

0.00	0.01	0.03	0.10	1.00	0.35	0.51	0.30	0.27
0.01	0.00	0.05	0.13	1.00	0.39	0.55	0.35	0.32
0.03	0.05	0.00	0.03	1.00	0.25	0.41	0.22	0.19
0.10	0.13	0.03	0.00	0.98	0.16	0.32	0.17	0.14
1.00	1.00	1.00	0.98	0.00	0.88	0.79	0.92	0.93
0.35	0.39	0.25	0.16	0.88	0.00	0.07	0.09	0.07
0.51	0.55	0.41	0.32	0.79	0.07	0.00	0.17	0.18
0.30	0.35	0.22	0.17	0.92	0.09	0.17	0.00	0.01
0.27	0.32	0.19	0.14	0.93	0.07	0.18	0.01	0.00

Coerência duas a duas:

1.00	1.00	0.99	0.97	0.03	0.87	0.77	0.89	0.91
1.00	1.00	0.99	0.97	0.07	0.85	0.75	0.87	0.89
0.99	0.99	1.00	0.99	0.06	0.92	0.83	0.93	0.94
0.97	0.97	0.99	1.00	0.15	0.95	0.88	0.95	0.96
0.03	0.07	0.06	0.15	1.00	0.41	0.53	0.32	0.32
0.87	0.85	0.92	0.95	0.41	1.00	0.98	0.98	0.98
0.77	0.75	0.83	0.88	0.53	0.98	1.00	0.95	0.95
0.89	0.87	0.93	0.95	0.32	0.98	0.95	1.00	1.00

0.91 0.89 0.94 0.96 0.32 0.98 0.95 1.00 1.00

Distância/100 duas a duas:

0.00 0.68 12.91 15.17 25.83 19.23 10.79 5.96 4.55

0.68 0.00 12.84 15.12 25.98 19.31 11.08 6.44 5.09

12.91 12.84 0.00 2.63 38.60 7.78 8.96 17.19 13.48

15.17 15.12 2.63 0.00 40.69 5.46 9.49 19.09 15.31

25.83 25.98 38.60 40.69 0.00 44.01 33.78 21.94 25.76

19.23 19.31 7.78 5.46 44.01 0.00 10.84 22.15 18.34

10.79 11.08 8.96 9.49 33.78 10.84 0.00 12.17 8.70

5.96 6.44 17.19 19.09 21.94 22.15 12.17 0.00 3.86

4.55 5.09 13.48 15.31 25.76 18.34 8.70 3.86 0.00

limiar da medida da distância Euclidiana calculado = 778.02196

limiar da medida da coerência calculado = 0.95246360

limiar da medida da entropia calculada = 0.086873733

Dados sobre a uniformidade espacial das amostras selecionadas:

0...vv2(b) 24

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23

1...vv2(d) 22

0 1 2 3 4 5 6 7 8 9 11 12 13 14 16 17 18 19 21 22 23 24

2...vv3(b) 25

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

3...vs(d) 25

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

4...a(a) 23

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 21 23 24

5...saq(a) 17

0 1 4 5 6 10 11 12 15 16 17 18 20 21 22 23 24

6...sle(d) 25

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

7...ste(d) 25

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

8...ste(e) 25

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

Busca dos MR:

Fator da DE: 0.320000

Fator da CE: 0.300000

Fator da entropia: 0.300000

limiar da medida da DE = 778.02196

limiar da medida da CE = 0.95246360

limiar da medida da entropia = 0.086873733

1) primeiro método de escolha - medidas conjuntas d-c-h

endmembers : 0 3 4 6 7

endmembers : vv2(b) vs(d) a(a) sle(d) ste(d)

dist_Euclid = 18.299616 coer media = 56.2607 entropia = 57.7253

medida = 3

2) segundo método de escolha - moda das 3 medidas

endmembers (d): 0 3 4 6 7

vv2(b) vs(d) a(a) sle(d) ste(d)

endmembers (c): 0 3 4 6 7

vv2(b) vs(d) a(a) sle(d) ste(d)

endmembers (h): 0 3 4 6 7

vv2(b) vs(d) a(a) sle(d) ste(d)

endmember: 0 3...vez(es)indicadores: d c h

endmember: 3 3...vez(es)indicadores: c d h

endmember: 4 3...vez(es)indicadores: c d h

endmember: 6 3...vez(es)indicadores: c d h

endmember: 7 3...vez(es)indicadores: d h c

end members escolhidos 3 vezes : 0 7 6 4 3

end members possíveis : 0 7 6 4 3

end members escolhidos : 0 7 6 4 3

3) terceiro método de escolha

eliminando redundâncias de classes

endmembers (d): 0 3 4 6 7

vv2(b) vs(d) a(a) sle(d) ste(d)

endmembers (c): 0 3 4 6 7

vv2(b) vs(d) a(a) sle(d) ste(d)

endmembers (h): 0 3 4 6 7

vv2(b) vs(d) a(a) sle(d) ste(d)

endmember: 0 3...vez(es)indicadores: d c h

endmember: 3 3...vez(es)indicadores: c d h

endmember: 4 3...vez(es)indicadores: c d h

endmember: 6 3...vez(es)indicadores: c d h

endmember: 7 3...vez(es)indicadores: d h c

end members escolhidos 3 vezes : 0 7 6 4 3

end members possíveis : 0 7 6 4 3

end members escolhidos : 0 7 6 4 3

4) quarto método de escolha

a) critério da medida conjunta

antepenúltima escolha= 0 0 0 0 0

vv2(b) vv2(b) vv2(b) vv2(b) vv2(b)

dist_Euclid = 0.000000

coerência = 0.000000

entropia = 0.000000

medida = 0

penúltima escolha= 0 0 0 0 0

vv2(b) vv2(b) vv2(b) vv2(b) vv2(b)

dist_Euclid = 0.000000

coerência = 0.000000

entropia = 0.000000

medida = 0

ultima escolha= 0 3 4 6 7

vv2(b) vs(d) a(a) sle(d) ste(d)

dist_Euclid = 18.299616

coerência = 56.2607

entropia = 57.7253

medida = 3

b) critério da distância

antepenúltima escolha= 0 0 0 0 0

vv2(b) vv2(b) vv2(b) vv2(b) vv2(b)

dist_Euclid = 0.000000

penúltima escolha= 0 0 0 0 0

vv2(b) vv2(b) vv2(b) vv2(b) vv2(b)

dist_Euclid = 0.000000

ultima escolha= 0 3 4 6 7

vv2(b) vs(d) a(a) sle(d) ste(d)

dist_Euclid = 18.299616

c) critério da coerência

antepenúltima escolha= 0 0 0 0 0

vv2(b) vv2(b) vv2(b) vv2(b) vv2(b)

coerência média = 0.000000

penúltima escolha= 0 0 0 0 0

vv2(b) vv2(b) vv2(b) vv2(b) vv2(b)

coerência média = 1000.00

ultima escolha= 0 3 4 6 7

vv2(b) vs(d) a(a) sle(d) ste(d)

coerência média = 56.2607

d) critério da entropia

antepenúltima escolha= 0 0 0 0 0

vv2(b) vv2(b) vv2(b) vv2(b) vv2(b)

Entropia = 0.000000

penúltima escolha= 0 0 0 0 0

vv2(b) vv2(b) vv2(b) vv2(b) vv2(b)

Entropia = 0.000000

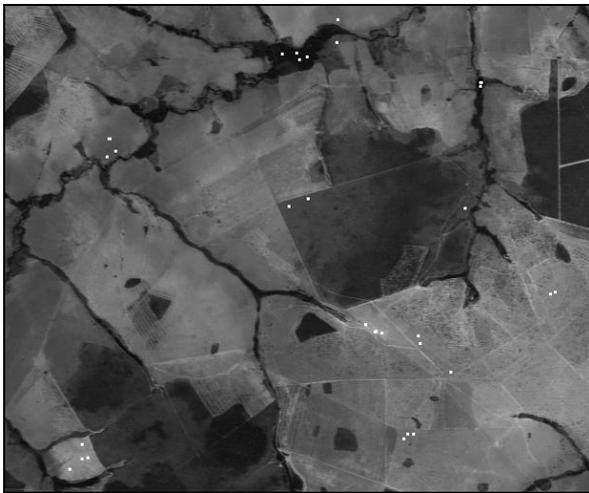
última escolha= 0 3 4 6 7

vv2(b) vs(d) a(a) sle(d) ste(d)

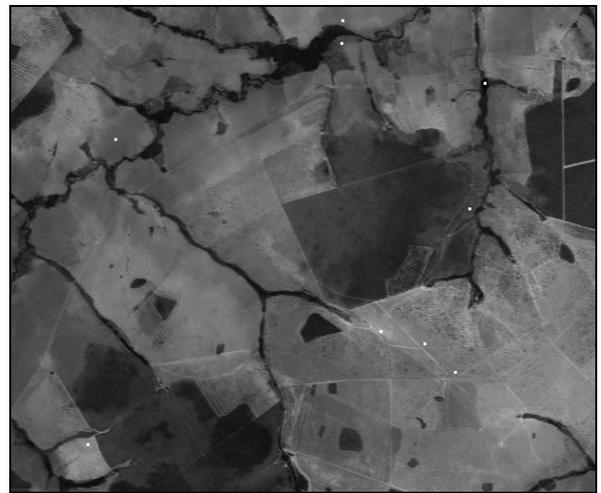
A.5 - Figuras e imagens de saída do programa

Mostra-se a seguir as figuras e imagens geradas pelo programa de triagem e busca dos MR. No exemplo todos os métodos de escolha dos MR (*DE*, *CE*, entropia e medida conjunta) produziram o mesmo resultado: *vv2(b)*, *vs(d)*, *a(a)*, *sle(d)* e *ste(d)*.

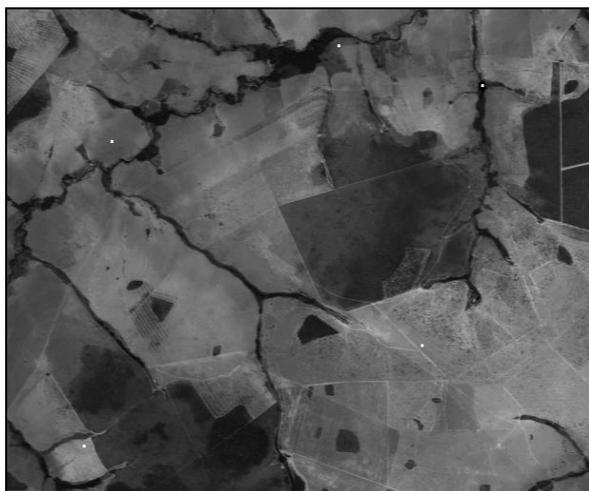
Na Fig. A.1a ilustra-se a imagem da subcena 7 e as posições das amostras iniciais, escolhidas pelo especialista, as posições das nove amostras resultantes do processo de triagem discriminatória (Fig. A.1b) e as posições das amostras escolhidas como MR (Fig. A.1c). Mostram-se ainda os espectros dos MR escolhidos (Fig. A.1d).



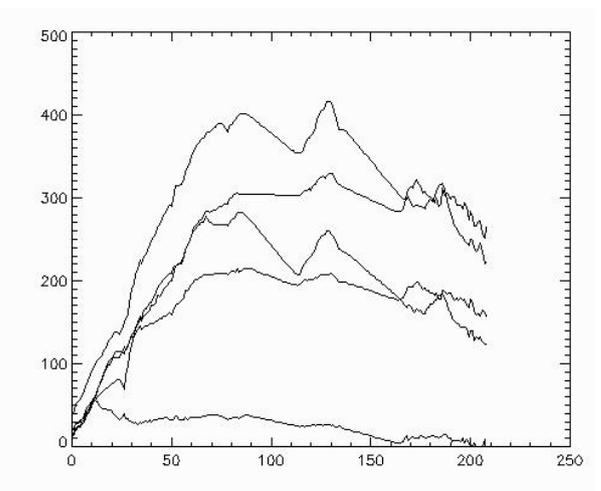
a) 32 amostras iniciais.



b) 9 amostras selecionadas como candidatas a MR.



c) 5 MR escolhidos



d) espectros dos MR escolhidos.

FIGURA A.1 - Amostras candidatas e MR escolhidos para a subcena 7.

FOLHA DE REGISTRO DO DOCUMENTO

1. CLASSIFICAÇÃO/TIPO TD	2. DATA 24 de outubro de 2007	3. DOCUMENTO N° CTA/ITA-IEE/TD-002/2007	4. N° DE PÁGINAS 185
5. TÍTULO E SUBTÍTULO: Triagem de Amostras e Múltiplas Medidas de Similaridade Espectral para Seleção de Membros de Referência em Imagens Hiperespectrais.			
6. AUTOR(ES): Marco Antonio Pizarro			
7. INSTITUIÇÃO(ÕES)/ÓRGÃO(S) INTERNO(S)/DIVISÃO(ÕES): Instituto Tecnológico da Aeronáutica. Divisão de Engenharia Eletrônica – ITA/IEE.			
8. PALAVRAS-CHAVE SUGERIDAS PELO AUTOR: Dados Hiperespectrais; Entropia; Membros de Referência; Sensoriamento Remoto; Espectroscopia de Imageamento; Espectrômetro Imageador Aerotransportado no Visível e Infravermelho (AVIRIS); Mistura Espectral; Processamento de Imagens; Classificação de Imagens.			
9. PALAVRAS-CHAVE RESULTANTES DE INDEXAÇÃO: Sensoriamento remoto; Análise de espectro; Entropia; Processamento de imagens; Classificação de imagens; Técnicas de formação de imagens; Espectroscopia no infravermelho; Telecomunicações; Engenharia eletrônica			
10. APRESENTAÇÃO: ITA, São José dos Campos, 2007. 185 páginas		X Nacional	Internacional
11. RESUMO: O objetivo do presente trabalho foi desenvolver um processo semi-automático para identificação de Membros de Referência (MR) em imagens hiperespectrais. O processo faz inicialmente a triagem das amostras espectrais gerando um conjunto de amostras candidatas a MR. A triagem consiste em: determinar uma vizinhança espacialmente uniforme em torno da amostra espectral, rejeitar amostras cujos espectros na vizinhança uniforme não são homogêneos, rejeitar amostras que não formam um contexto uniforme e homogêneo com outras amostras e rejeitar amostras redundantes. As amostras espectrais candidatas a MR, consideradas uniformes, homogêneas, pertencentes a um contexto e pouco redundantes, são submetidas a um processo de busca dos MR, que utiliza o conceito de entropia, proposto neste trabalho, como medida de similaridade espectral, em conjunto com a Distância Euclidiana e a medida de Coerência entre os espectros. Propõe-se ainda dois limitantes superiores para o número máximo de MR presentes em uma imagem hiperespectral, utilizando-se o conceito de amostras espectrais bem configuradas e um valor de entropia mínima permissível, entre amostras. Foi implementado com o aplicativo computacional IDL® um processador que realiza todas as funções propostas. Os métodos apresentados são testados, com êxito, em imagens hiperespectrais do sensor aerotransportado AVIRIS (<i>Airborne Visible Infrared Imaging Spectrometer</i>) do JPL.			
12. GRAU DE SIGILO: (X) OSTENSIVO () RESERVADO () CONFIDENCIAL () SECRETO			