

UNIVERSIDADE ESTADUAL DE CAMPINAS  
FACULDADE DE ENGENHARIA ELÉTRICA E DE COMPUTAÇÃO  
DEPARTAMENTO DE ENGENHARIA DE COMPUTAÇÃO E AUTOMAÇÃO INDUSTRIAL

# Métodos de Classificação Não-supervisionada de Imagens de Sensoriamento Remoto usando Mapas Auto-organizáveis de Kohonen

**Autor:** Márcio Leandro Gonçalves

**Orientador:** Prof. Dr. Márcio Luiz de Andrade Netto

**Co-orientador:** Prof. Dr. José Alfredo Ferreira Costa

**Tese de Doutorado** apresentada à Faculdade de Engenharia Elétrica e de Computação como parte dos requisitos para obtenção do título de Doutor em Engenharia Elétrica. Área de concentração: **Engenharia de Computação.**

Banca Examinadora

Clésio Luiz Tozzi, Dr. .... DCA/FEEC/UNICAMP  
Fernando José Von Zuben, Dr. .... DCA/FEEC/UNICAMP  
Jurandir Zullo Jr., Dr. .... CEPAGRI/UNICAMP  
Márcio Luiz de Andrade Netto, Dr. .... DCA/FEEC/UNICAMP  
Sarajane Marques Peres, Dra. .... EACH/USP

Campinas, SP

Março/2009

FICHA CATALOGRÁFICA ELABORADA PELA  
BIBLIOTECA DA ÁREA DE ENGENHARIA E ARQUITETURA- BAE - UNICAMP

G586m Gonçalves, Márcio Leandro  
Métodos de classificação não-supervisionada de  
imagens de sensoriamento remoto usando mapas auto-  
organizáveis de Kohonen / Márcio Leandro Gonçalves. –  
Campinas, SP:[s.n.], 2009.

Orientadores: Márcio Luiz de Andrade Netto, José  
Alfredo Ferreira Costa.

Tese (doutorado) - Universidade Estadual de  
Campinas, Faculdade de Engenharia Elétrica e de  
Computação.

1. Redes neurais (Computação). 2. Inteligência  
Artificial. 3. Processamento imagens - Técnicas digitais.  
4. Sensoriamento Remoto. 5. Mapas auto-organizáveis.  
I. Andrade Netto, Márcio Luiz de. II. Costa, José  
Alfredo Ferreira. III. Universidade Estadual de  
Campinas. Faculdade de Engenharia Elétrica e de  
Computação. IV. Título

Título em Inglês: Unsupervised methods of classifying remotely sensed  
images using Kohonen self-organizing maps  
Palavras-chave em Inglês: Neural networks, Artificial intelligence, Digital  
Image processing, Remote Sensing, Kohonen maps  
Área de concentração: Engenharia de Computação  
Titulação: Doutor em Engenharia Elétrica  
Banca Examinadora: Jurandir Zullo Júnior, Sarajane Marques Peres,  
Clésio Luiz Tozzi, Fernando José Von Zuben  
Data da defesa: 11/03/2009  
Programa de Pós Graduação: Engenharia Elétrica

## COMISSÃO JULGADORA - TESE DE DOUTORADO

**Candidato:** Márcio Leandro Gonçalves

**Data da Defesa:** 11 de março de 2009

**Título da Tese:** "Métodos de Classificação Não-supervisionada de Imagens de Sensoriamento Remoto Usando Mapas Auto-organizáveis de Kohonen"

Prof. Dr. Márcio Luiz de Andrade Netto (Presidente):  \_\_\_\_\_

Profa. Dra. Sarajane Marques Peres:  \_\_\_\_\_

Prof. Dr. Jurandir Zullo Júnior:  \_\_\_\_\_

Prof. Dr. Clésio Luis Tozzi: \_\_\_\_\_

Prof. Dr. Fernando José Von Zuben:  \_\_\_\_\_

# Resumo

Esta tese propõe novas metodologias de classificação não-supervisionada de imagens de sensoriamento remoto que particularmente exploram as características e propriedades do Mapa Auto-organizável de Kohonen (SOM - *Self-Organizing Map*). O ponto chave dos métodos de classificação propostos é realizar a análise de agrupamentos das imagens através do mapeamento produzido pelo SOM, ao invés de trabalhar diretamente com os padrões originais das cenas. Tal estratégia reduz significativamente a complexidade da análise dos dados, tornando possível a utilização de técnicas normalmente consideradas computacionalmente inviáveis para o processamento de imagens de sensoriamento remoto, como métodos de agrupamentos hierárquicos e índices de validação de agrupamentos. Diferentemente de outras abordagens, nas quais o SOM é utilizado como ferramenta de auxílio visual para a detecção de agrupamentos, nos métodos de classificação propostos, mecanismos para analisar de maneira automática o arranjo de neurônios de um SOM treinado são aplicados e aprimorados com o objetivo de encontrar as melhores partições para os conjuntos de dados das imagens. Baseando-se nas propriedades estatísticas do SOM, modificações nos cálculos de índices de validação agrupamentos são propostas com o objetivo de reduzir o custo computacional do processo de classificação das imagens. Técnicas de análise de textura em imagens são aplicadas para avaliar e filtrar amostras de treinamento e/ou protótipos do SOM que correspondem a regiões de transição entre classes de cobertura terrestre. Informações espaciais a respeito dos protótipos do SOM, além das informações de distância multiespectral, também são aplicadas em critérios de fusão de agrupamentos procurando facilitar a discriminação de classes de cobertura terrestre que apresentam alto grau de similaridade espectral. Resultados experimentais mostram que os métodos de classificação propostos apresentam vantagens significativas em relação às técnicas de classificação não-supervisionada mais frequentemente utilizadas na área de sensoriamento remoto.

**Palavras-chave:** Redes Neurais Artificiais, Mapas Auto-Organizáveis de Kohonen, Processamento Digital de Imagens, Sensoriamento Remoto, Reconhecimento de Padrões.



# Abstract

This thesis proposes new methods of unsupervised classification for remotely sensed images which particularly exploit the characteristics and properties of the Kohonen Self-Organizing Map (SOM). The key point is to execute the clustering process through a set of prototypes of SOM instead of analyzing directly the original patterns of the image. This strategy significantly reduces the complexity of data analysis, making it possible to use techniques that have not usually been considered computationally viable for processing remotely sensed images, such as hierarchical clustering methods and cluster validation indices. Unlike other approaches in which SOM is used as a visual tool for detection of clusters, the proposed classification methods automatically analyze the neurons grid of a trained SOM in order to find better partitions for data sets of images. Based on the statistical properties of the SOM, clustering validation indices calculated in a modified manner are proposed with the aim of reducing the computational cost of the classification process of images. Image texture analysis techniques are applied to evaluate and filter training samples and/or prototypes of the SOM that correspond to transition regions between land cover classes. Spatial information about the prototypes of the SOM, in addition to multispectral distance information, are also incorporated in criteria for merging clusters with aim to facilitate the discrimination of land cover classes which have high spectral similarity. Experimental results show that the proposed classification methods present significant advantages when compared to unsupervised classification techniques frequently used in remote sensing.

**Keywords:** Artificial Neural Networks, Self-Organizing Maps, Digital Image Processing, Remote Sensing, Pattern Recognition.

*“Dedicação é a capacidade de se entregar à realização de um objetivo. Não conheço ninguém que tenha progredido na carreira sem trabalhar pelo menos doze horas por dia nos primeiros anos. Não conheço ninguém que conseguiu realizar seu sonho sem sacrificar sábados e domingos pelo menos uma centena de vezes. Da mesma forma, se você quiser construir uma relação amigável com seus filhos, terá de se dedicar a isso, superar o cansaço, arrumar tempo para ficar com eles, deixar de lado o orgulho e o comodismo. Se quiser um casamento gratificante, terá de investir tempo, energia e sentimentos nesse objetivo. O sucesso é construído a noite! Durante o dia você faz o que todos fazem. Mas, para conseguir um resultado diferente da maioria, você tem de ser especial. Se fizer igual a todo mundo, obterá os mesmos resultados. Não se compare à maioria, pois, infelizmente, ela não é modelo de sucesso. Se você quiser atingir uma meta especial, terá de estudar no horário em que os outros estão tomando chope com batatas fritas. Terá de planejar, enquanto os outros permanecem à frente da televisão. Terá de trabalhar, enquanto os outros tomam sol à beira da piscina. A realização de um sonho depende da dedicação. Há muita gente que espera que o sonho se realize por magia. Mas toda magia é ilusão. E ilusão não tira ninguém do lugar onde está.*

*Ilusão é combustível de perdedores.”*

**Roberto Shinyashiki**

*Dedico este trabalho à minha esposa, Renata, à minha filha, Bruna, e ao meu filho, Leandro, que está chegando.*

# Agradecimentos

Agradeço primeiramente ao meu orientador e amigo prof. Márcio Luiz de Andrade Netto, que acreditou em mim e abriu (mais de uma vez) as portas da FEEC/Unicamp para que eu pudesse alcançar importantes conquistas acadêmicas. O prof. Márcio foi um companheiro de percurso e de discussões sempre proveitosas, dentro e fora do contexto desse trabalho, contemplando-me cordialmente, inúmeras vezes, com a sua atenção, conhecimento, paciência e amizade. Sem a sua ajuda, certamente os meus sonhos e desejos profissionais não seriam alcançados. Devo a este admirável ser humano grande parte de meu crescimento como pessoa e como profissional. Muitíssimo obrigado, xará!

Ao meu co-orientador prof. José Alfredo F. Costa, pelas valiosas discussões, sugestões e atenção sempre gentilmente concedidas durante o desenvolvimento desse trabalho por meio de incontáveis e-mails.

Ao prof. Jurandir Zullo Jr. pela atenção e apoio. A transmissão de seus conhecimentos referentes às áreas de processamento de imagens e sensoriamento remoto, desde a época do Mestrado, foram fundamentais para o progresso desse trabalho.

Aos professores Fernando José Von Zuben e Clésio Luiz Tozzi pelas críticas e sugestões dadas durante o meu exame de qualificação. Certamente suas contribuições foram de grande valia para o enriquecimento dessa tese.

Aos diversos professores/pesquisadores que foram muito cordiais, tirando dúvidas, discutindo assuntos de interesse ou enviando contribuições. Entre eles, posso citar: Mateus Batistella (Embrapa), Camilo D. Rennó (INPE), Sarajane Marques Peres (EACH/USP), Sanghamitra Bandyopadhyay (Indian Statistical Institute), Kadim Tasdemir (Rice University), Erzsébet Merényi (Rice University), Costas Varotsos (University of Athens), Murat Sezgin (Istanbul Technical University) e Y. Hong (University of California).

À Divisão de Geração de Imagens (DGI) do INPE e à ENGESAT pela concessão das imagens que foram utilizadas neste trabalho e, também, pelo suporte técnico.

À minha esposa Renata e à minha filha Bruna, por serem minha força e fonte de inspiração diária e por terem compreendido a minha ausência em tantas ocasiões em que precisei me dedicar integralmente ao trabalho.

Aos meus pais e irmãos, que apesar da distância, sempre foram fontes de carinho, apoio e incentivo.

À Lourdes, pela atenção, carinho e também pela ajuda nos momentos em que não pude ser um pai presente.

Ao Ricardo, à Lidiane e ao Rafa pela paciência e compreensão por tantas visitas e passeios pensados, mas nunca cumpridos.

Ao professor Iran Calixto Abrão pela amizade, parceria e companheirismo ao longo de muitos anos dentro e fora da PUC Minas.

Aos professores do Curso de Ciência da Computação da PUC Minas em Poços de Caldas, em especial aos professores Faria, Corrêa, Udo e Boca.

À PUC Minas que através do seu Programa Permanente de Capacitação Docente concedeu-me importantes horas de dedicação para realizar esse trabalho.

Ao funcionário Edson da CPG/FEEC pelo atendimento sempre feito de forma atenciosa e prestativa.

A Deus, por ter colocado todas as pessoas citadas aqui em minha vida, por ter me concedido força, discernimento, motivação e saúde para desenvolver essa tese e, principalmente, por ter me feito sentir a Sua presença nos momentos mais difíceis e solitários durante todos esses anos de trabalho. Obrigado mais uma vez, SENHOR!

# Sumário

<b>Lista de Figuras</b>	<b>xv</b>
<b>Lista de Tabelas</b>	<b>xxiii</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Contexto e motivação . . . . .	1
1.2 Objetivos . . . . .	12
1.3 Organização do texto . . . . .	13
<b>2 Processamento Digital de Imagens de Sensoriamento Remoto</b>	<b>15</b>
2.1 Introdução . . . . .	15
2.2 Fundamentos do Sensoriamento Remoto . . . . .	16
2.2.1 Radiação solar e espectro eletromagnético . . . . .	17
2.2.2 Interação da radiação com o ambiente terrestre . . . . .	20
2.2.3 Sistemas sensores . . . . .	22
2.2.4 Níveis de aquisição de dados . . . . .	25
2.2.5 Satélites . . . . .	25
2.3 Características das imagens de sensoriamento remoto . . . . .	30
2.4 Operações do processamento digital de imagens de sensoriamento remoto . . . . .	35
2.5 Técnicas de classificação . . . . .	38
2.5.1 Classificação supervisionada . . . . .	40
2.5.1.1 Classificador estatístico - Máxima Verossimilhança . . . . .	44
2.5.1.2 Classificador neural - <i>Perceptrons</i> multicamadas . . . . .	45

2.5.2	Classificação não-supervisionada . . . . .	46
2.5.2.1	Algoritmo K-médias . . . . .	48
2.5.3	Classificação híbrida . . . . .	51
2.5.4	Incorporação de informações não-espectrais . . . . .	51
2.5.4.1	Informações externas . . . . .	52
2.5.4.2	Textura . . . . .	52
2.6	Mecanismos de validação da acuidade do mapeamento . . . . .	53
<b>3</b>	<b>Mapa Auto-organizável de Kohonen</b>	<b>55</b>
3.1	Introdução . . . . .	55
3.2	Motivação neurobiológica . . . . .	56
3.3	Estrutura básica do SOM . . . . .	58
3.4	Algoritmo de treinamento . . . . .	60
3.4.1	Considerações sobre os parâmetros de treinamento . . . . .	61
3.5	Propriedades do SOM . . . . .	67
3.6	Métodos de visualização . . . . .	68
3.6.1	Matriz de distâncias unificada - <i>U-matrix</i> . . . . .	69
3.7	Métodos de segmentação do SOM . . . . .	71
3.7.1	Segmentação do SOM baseada em técnicas de morfologia matemática . . . . .	75
3.7.1.1	Exemplo de uso . . . . .	79
3.8	Literatura e modelos derivados do SOM . . . . .	91
<b>4</b>	<b>Classificação baseada em SOMs segmentados por morfologia matemática</b>	<b>93</b>
4.1	Introdução . . . . .	93
4.2	Metodologia de classificação de imagens de sensoriamento remoto usando SOMs segmentados por morfologia matemática . . . . .	94
4.2.1	Amostragem da imagem . . . . .	95
4.2.2	Treinamento do SOM . . . . .	104
4.2.2.1	Mapa de Características da Imagem . . . . .	107
4.2.3	Segmentação do SOM . . . . .	107

---

4.2.4	Classificação da imagem . . . . .	109
4.3	Resultados Experimentais . . . . .	110
4.3.1	Imagem teste . . . . .	110
4.3.2	Aplicação da metodologia proposta . . . . .	111
4.3.3	Análise dos resultados . . . . .	127
4.3.3.1	Comparação com o algoritmo K-médias . . . . .	129
4.4	Considerações finais . . . . .	132
<b>5</b>	<b>Classificação baseada em SOMs segmentados por morfologia matemática e em índices de validação de agrupamentos</b>	<b>135</b>
5.1	Introdução . . . . .	135
5.2	Análise de agrupamentos usando SOMs segmentados por morfologia matemática e índices de validação . . . . .	137
5.3	Análise, seleção e modificação dos cálculos de índices de validação . . . . .	140
5.3.1	Levantamento bibliográfico e índices pré-selecionados . . . . .	142
5.3.2	Modificação dos cálculos de índices de validação . . . . .	148
5.3.3	Testes experimentais . . . . .	150
5.4	Metodologia de classificação de imagens usando SOMs segmentados por morfologia matemática e índices de validação de agrupamentos . . . . .	159
5.4.1	Resultados Experimentais . . . . .	160
5.4.1.1	Imagem teste 1 . . . . .	160
5.4.1.2	Imagem teste 2 . . . . .	168
5.5	Geração de bordas mais definidas na <i>U-matrix</i> . . . . .	178
5.6	Abordagem paralela . . . . .	192
5.7	Considerações finais . . . . .	193
<b>6</b>	<b>Classificação baseada em SOMs segmentados hierarquicamente com informações espectrais e espaciais</b>	<b>197</b>
6.1	Introdução . . . . .	197

6.2	Metodologia de classificação de imagens usando SOMs segmentados hierarquicamente com informações espectrais e espaciais . . . . .	199
6.2.1	Filtragem de protótipos . . . . .	201
6.2.2	Análise hierárquica . . . . .	203
6.2.3	Avaliação dos agrupamentos . . . . .	206
6.2.4	Classificação da imagem . . . . .	207
6.2.5	Reclassificação . . . . .	207
6.3	Resultados Experimentais . . . . .	207
6.3.1	Comparação com o algoritmo K-médias padrão e <i>janelado</i> . . . . .	216
6.3.2	Comparação com os métodos anteriores . . . . .	219
6.3.3	Exemplos de reuso . . . . .	223
6.4	Considerações finais . . . . .	227
<b>7</b>	<b>Conclusões</b>	<b>231</b>
7.1	Considerações finais sobre os métodos propostos . . . . .	231
7.2	Contribuições . . . . .	236
7.3	Trabalhos futuros . . . . .	240
7.4	Publicações . . . . .	243
	<b>Referências bibliográficas</b>	<b>245</b>
<b>A</b>	<b>Mídia com informações adicionais</b>	<b>267</b>
A.1	Exemplo de uso do método de análise de agrupamentos proposto no capítulo 5	
A.2	Descrição dos índices de validação de agrupamentos pré-selecionados	
A.3	Versões modificadas dos índices de validação de agrupamentos	
A.4	Imagens correspondentes aos conjuntos de dados <i>SR1</i> e <i>SR2</i>	
A.5	Tabelas e gráficos dos experimentos realizados com índices de validação	
A.6	Árvore de mapas da imagem teste 2	
A.7	Histograma de vencimentos do SOM da imagem teste 2	
A.8	Dendrogramas do SOM da imagem IKONOS	
A.9	Artigos publicados	



# Lista de Figuras

2.1	<i>Esquema representativo dos quatro elementos fundamentais das técnicas de sensoriamento remoto (adaptada de Castro Jr. (2001)). . . . .</i>	17
2.2	<i>Espectro eletromagnético. . . . .</i>	19
2.3	<i>Causas das atenuações da radiação solar na atmosfera terrestre (adaptado de Moreira (2005)). . . . .</i>	21
2.4	<i>Curvas de reflectância espectral de diferentes classes de uso do solo (adaptado de Richards (1995)). . . . .</i>	22
2.5	<i>Componentes de um sistema sensor (adaptado de Moreira (2005)). . . . .</i>	23
2.6	<i>Imagens com diferentes resoluções espaciais: (a) Imagem do satélite IKONOS-II com resolução espacial de 1m. (b) Imagem do satélite CBERS com resolução espacial de 20m (Fonte: INPE (2007)). . . . .</i>	31
2.7	<i>Imagens das bandas espectrais 3, 4 e 5 do sensor TM do satélite LANDSAT. . .</i>	33
2.8	<i>Simulação de uma imagem do satélite CBERS: (a) com dois níveis de cinza; (b) com quatro níveis de cinza. . . . .</i>	34
2.9	<i>Ilustração de um espaço multiespectral bidimensional mostrando sua relação com as reflectâncias espectrais dos tipos de cobertura terrestre. . . . .</i>	40
2.10	<i>Representação de classes de informação por conjuntos de classes espectrais. . . .</i>	41
3.1	<i>Representação das várias regiões corticais especializadas no cérebro humano. . .</i>	57
3.2	<i>Estrutura básica do SOM. . . . .</i>	59
3.3	<i>Duas configurações de arranjo de neurônios do SOM e níveis de vizinhança ao redor do neurônio vencedor. . . . .</i>	60

3.4	<i>Representação das distâncias utilizadas no cálculo da U-matrix em um arranjo de neurônios retangular. Os círculos representam os neurônios e os cubos representam as distâncias.</i>	69
3.5	<i>Ilustração do processo de geração da árvore dinâmica de mapas pelo TS-SL-SOM (adaptado de Costa (1999)).</i>	78
3.6	<i>Conjunto de dados sintético.</i>	79
3.7	<i>Dados e neurônios do SOM após o treinamento.</i>	80
3.8	<i>Neurônios do SOM com tamanho proporcional à sua atividade.</i>	80
3.9	<i>Histograma de vencedores.</i>	80
3.10	<i>U-matrix na forma planar.</i>	81
3.11	<i>U-matrix na forma de superfície.</i>	81
3.12	<i>Configuração dos neurônios do SOM após a eliminação do efeito dos neurônios inativos.</i>	81
3.13	<i>U-matrix na forma de superfície após a eliminação do efeito dos neurônios inativos.</i>	81
3.14	<i>Gráfico do valor de limiar (<math>k</math>) da U-matrix versus o número de regiões conectadas (agrupamentos).</i>	82
3.15	<i>Imagem de marcadores para a U-matrix correspondente ao limiar <math>k = 81</math>.</i>	83
3.16	<i>Partição encontrada pelo algoritmo watershed.</i>	83
3.17	<i>U-matrix rotulada após a segmentação pela watershed.</i>	83
3.18	<i>SOM rotulado.</i>	83
3.19	<i>Neurônios do submapa 1 após o treinamento e eliminação dos neurônios inativos.</i>	84
3.20	<i>Histograma de vencedores do submapa 1.</i>	84
3.21	<i>U-matrix do submapa 1 na forma planar.</i>	84
3.22	<i>U-matrix do submapa 1 na forma de superfície.</i>	84
3.23	<i>Número de regiões conectadas versus limiar da U-matrix - submapa 1.</i>	85
3.24	<i>Imagem de marcadores correspondente ao valor de limiar <math>k = 34</math> - submapa 1.</i>	85
3.25	<i>Partição da U-matrix encontrada pelo algoritmo watershed - submapa 1.</i>	85
3.26	<i>U-matrix segmentada - submapa 1.</i>	86
3.27	<i>Submapa 1 rotulado.</i>	86

3.28	<i>Conjuntos de dados 1 e 2 classificados.</i>	86
3.29	<i>Neurônios do submapa 2 após o treinamento e eliminação dos neurônios inativos.</i>	87
3.30	<i>Histograma de vencedores do submapa 2.</i>	87
3.31	<i>U-matrix do submapa 2 na forma planar.</i>	87
3.32	<i>U-matrix do submapa 2 na forma de superfície.</i>	87
3.33	<i>Número de regiões conectadas versus limiar da U-matrix - submapa 2.</i>	88
3.34	<i>Imagem de marcadores correspondente ao valor de limiar <math>k = 41</math> - submapa 2.</i>	88
3.35	<i>Partição da U-matrix encontrada pelo algoritmo watershed - submapa 2.</i>	88
3.36	<i>U-matrix segmentada - submapa 2.</i>	89
3.37	<i>Submapa 2 rotulado.</i>	89
3.38	<i>Conjuntos de dados 3 e 4 classificados.</i>	89
3.39	<i>Árvore de mapas gerada durante a aplicação do método proposto por Costa e Andrade Netto (1999, 2001) sobre o conjunto de dados da figura 3.6.</i>	90
4.1	<i>Ilustração da metodologia de classificação proposta usando SOMs segmentados por morfologia matemática.</i>	95
4.2	<i>Imagem construída com quatro classes de cobertura terrestre capturadas pelo satélite IKONOS (com resolução espacial de 4 metros).</i>	97
4.3	<i>U-matrix 3D do SOM treinado com janelas de pixels.</i>	97
4.4	<i>Imagem dos protótipos do SOM treinado com janelas de pixels.</i>	97
4.5	<i>U-matrix 3D do SOM treinado com pixels isolados.</i>	98
4.6	<i>Imagem dos protótipos do SOM treinado com pixels isolados.</i>	98
4.7	<i>Imagem da figura 4.2 classificada através dos protótipos do SOM treinado com janelas de pixels.</i>	99
4.8	<i>Imagem da figura 4.2 classificada através dos protótipos do SOM treinado com pixels isolados.</i>	99
4.9	<i>Processo de amostragem da imagem de entrada.</i>	101

- 4.10 *Efeito do espaçamento entre janelas amostrais no treinamento do SOM: (a) Imagem Landsat/TM. (b) Imagem de um SOM treinado com janelas amostrais coletadas em intervalos regulares de 10 pixels. Os protótipos correspondentes à classe água aparecem (circulados) no canto superior esquerdo do mapa. (c) Imagem de um SOM treinado com janelas amostrais coletadas em intervalos regulares de 20 pixels. Apenas um protótipo no mapa (circulado no canto superior esquerdo) corresponde à classe água. (d) Imagem de um SOM treinado com janelas amostrais coletadas em intervalos regulares de 30 pixels. O mapa não apresenta protótipos associados à classe água. . . . . 102*
- 4.11 *Efeito do tamanho do SOM no mapeamento dos padrões da imagem LANDSAT da figura 4.10: (a) SOM 12x12: protótipos correspondentes à classe água aparecem (circulados) no canto superior esquerdo do mapa. (b) SOM 9x9: o mapa também apresenta protótipos (circulados no canto superior esquerdo) correspondentes à classe água. (c) SOM 6x6: a classe água não aparece de maneira representativa no mapa. . . . . 106*
- 4.12 *Ilustração do processo de construção do Mapa de Características da Imagem (MCI). 108*
- 4.13 *Composição colorida das bandas 3(R), 4(G) e 5(B) da imagem teste. . . . . 110*
- 4.14 *Janelas amostrais de tamanho 5x5 coletadas a partir da imagem original. . . . . 111*
- 4.15 *Imagem do MCI após o treinamento do SOM. . . . . 112*
- 4.16 *U-matrix na forma de superfície do SOM treinado. . . . . 113*
- 4.17 *Gráfico do valor de limiar (k) da U-matrix versus o número de regiões conectadas (agrupamentos). . . . . 114*
- 4.18 *Imagem de marcadores correspondente ao valor de limiar  $k = 149$ . . . . . 114*
- 4.19 *Partição da U-matrix encontrada pelo algoritmo watershed. . . . . 114*
- 4.20 *Sobreposição das linhas de watershed sobre a U-matrix apresentada na figura 4.16. 115*
- 4.21 *U-matrix rotulada após a segmentação pela watershed. . . . . 115*
- 4.22 *SOM rotulado. . . . . 115*
- 4.23 *Imagem do MCI do submapa 1 após o treinamento. . . . . 117*
- 4.24 *U-matrix 3D do submapa 1 treinado. . . . . 117*

4.25	<i>Número de regiões conectadas versus limiar da U-matrix - submapa 1.</i>	118
4.26	<i>Imagem de marcadores correspondente ao valor de limiar <math>k = 19</math> - submapa 1.</i>	118
4.27	<i>Partição da U-matrix encontrada pelo algoritmo watershed - submapa 1.</i>	118
4.28	<i>Sobreposição das linhas de watershed sobre a U-matrix apresentada na figura 4.24.</i>	119
4.29	<i>U-matrix segmentada - submapa 1.</i>	119
4.30	<i>Submapa 1 rotulado.</i>	119
4.31	<i>Número de regiões conectadas versus limiar da U-matrix - submapa 1.1.</i>	120
4.32	<i>Número de regiões conectadas versus limiar da U-matrix - submapa 1.2.</i>	121
4.33	<i>Imagem do MCI do submapa 2 após o treinamento.</i>	121
4.34	<i>U-matrix 3D do submapa 1 treinado.</i>	122
4.35	<i>Número de regiões conectadas versus limiar da U-matrix - submapa 2.</i>	123
4.36	<i>Imagem de marcadores correspondente ao valor de limiar <math>k = 78</math> - submapa 2.</i>	123
4.37	<i>Partição da U-matrix encontrada pelo algoritmo watershed - submapa 2.</i>	123
4.38	<i>Sobreposição das linhas de watershed sobre a U-matrix apresentada na figura 4.34.</i>	124
4.39	<i>U-matrix segmentada - submapa 2.</i>	124
4.40	<i>Submapa 2 rotulado.</i>	124
4.41	<i>Número de regiões conectadas versus limiar da U-matrix - submapa 2.1.</i>	125
4.42	<i>Número de regiões conectadas versus limiar da U-matrix - submapa 2.2.</i>	125
4.43	<i>Árvore de mapas gerada pelo método de segmentação do SOM proposto por Costa e Andrade Netto (1999, 2001) aplicado na análise de agrupamentos dos dados da imagem teste. Para cada mapa da árvore são apresentados o tamanho, número padrões e tempo de treinamento em segundos (s).</i>	126
4.44	<i>Imagem teste classificada pela metodologia proposta.</i>	127
4.45	<i>Imagem teste da figura 4.13 classificada pelo algoritmo K-médias considerando <math>k = 4</math>.</i>	130
5.1	<i>Ilustração da estratégia de análise de agrupamentos proposta.</i>	138
5.2	<i>Conjunto de dados sintéticos utilizados nos experimentos.</i>	153
5.3	<i>Ilustração do método de classificação de imagens proposto usando SOMs segmentados por morfologia matemática e índices de validação.</i>	160

5.4	<i>Gráfico do valor de limiar (<math>k</math>) da U-matrix versus o número de regiões conectadas destacando os valores de <math>k</math> que pertencem a <math>S_k</math>.</i>	161
5.5	<i>Diferentes partições para o SOM da imagem teste 1.</i>	162
5.6	<i>Imagem teste 1 classificada pela metodologia de classificação proposta.</i>	164
5.7	<i>Gráficos dos valores das versões originais e modificadas dos três índices de validação utilizados para avaliar as partições do conjunto de dados da imagem teste 1.</i>	167
5.8	<i>Composição colorida das bandas 3(R), 4(G) e 2(B) da imagem teste 2.</i>	168
5.9	<i>Imagem do MCI da imagem teste 2 após o treinamento do SOM.</i>	169
5.10	<i>U-matrix na forma de superfície do SOM da imagem teste 2.</i>	170
5.11	<i>Número de regiões conectadas versus limiar da U-matrix da imagem teste 2.</i>	171
5.12	<i>Diferentes partições do SOM da imagem teste 2.</i>	172
5.13	<i>Imagem teste 2 classificada pela metodologia de classificação proposta.</i>	173
5.14	<i>Gráficos dos valores das versões originais e modificadas dos três índices de validação utilizados para avaliar as partições do conjunto de dados da imagem teste 2.</i>	176
5.15	<i>Imagem teste 2 classificada pelo algoritmo K-médias considerando <math>k = 5</math>.</i>	177
5.16	<i>Segmento de imagem do satélite IKONOS.</i>	180
5.17	<i>Janelas amostrais coletadas do segmento de imagem do satélite IKONOS. As amostras heterogêneas estão marcadas em vermelho.</i>	181
5.18	<i>Projeção do conjunto de amostras usando as duas componentes principais. Os pontos em vermelho correspondem às amostras heterogêneas e os pontos em azul correspondem às amostras homogêneas.</i>	182
5.19	<i>MCI do SOM treinado com as janelas amostrais apresentadas na figura 5.17. Os protótipos heterogêneos estão marcados em vermelho.</i>	182
5.20	<i>Projeção do conjunto de protótipos do SOM usando as duas componentes principais. Os pontos em vermelho são os protótipos heterogêneos.</i>	183
5.21	<i>U-matrix na forma de superfície do SOM treinado com todas as janelas mostrais coletadas da imagem da figura 5.16.</i>	184

5.22	<i>MCI do SOM treinado sem as amostras heterogêneas.</i>	184
5.23	<i>U-matrix na forma de superfície do SOM treinado sem as amostras heterogêneas.</i>	185
5.24	<i>U-matrix resultante do SOM treinado sem as amostras heterogêneas da imagem da figura 4.13.</i>	188
5.25	<i>U-matrix resultante do SOM treinado sem as amostras heterogêneas da imagem da figura 4.13.</i>	189
5.26	<i>Número de regiões conectadas versus limiar da U-matrix da imagem teste 1 derivada do SOM treinado sem amostras heterogêneas.</i>	190
5.27	<i>Número de regiões conectadas versus limiar da U-matrix da imagem teste 2 derivada do SOM treinado sem amostras heterogêneas.</i>	191
5.28	<i>Ilustração do modelo de implementação paralela sugerido para a estratégia de análise de agrupamentos empregada na metodologia de classificação proposta.</i>	193
6.1	<i>Ilustração do método de classificação de imagens proposto usando SOMs segmentados hierarquicamente.</i>	200
6.2	<i>Neurônios adjacentes na grade do SOM.</i>	204
6.3	<i>Composição colorida (3R 2G 1B + PAN) da imagem IKONOS (provida pela Engesat/Brasil, ©Space Imaging).</i>	208
6.4	<i>MCI da imagem da figura 6.3 após o treinamento do SOM.</i>	209
6.5	<i>Histograma de vencimentos do SOM treinado com os dados da imagem da figura 6.3. Os protótipos marcados em vermelho são aqueles que apresentaram atividade nula (inativos).</i>	210
6.6	<i>Valores de energia (ENE) de cada um dos protótipos do SOM. Os protótipos marcados em vermelho são aqueles cujos valores de energia não satisfizeram a relação dada pela equação 5.4.</i>	211
6.7	<i>Valores do índice <math>\mathcal{V}_{CD_{bw}}^{SOM}</math> para as configurações de agrupamentos dos níveis 97 ao 114 do dendrograma.</i>	212
6.8	<i>SOM da imagem da figura 6.3 classificado.</i>	212
6.9	<i>Imagem da figura 6.3 classificada pelo método proposto.</i>	214
6.10	<i>Imagem da figura 6.3 classificada pelo algoritmo K-médias padrão.</i>	217

---

6.11	<i>Imagem da figura 6.3 classificada pelo algoritmo K-médias "janelado". . . . .</i>	218
6.12	<i>Imagem da figura 6.3 classificada pelo método proposto no capítulo 5. . . . .</i>	220
6.13	<i>U-matrix na forma de superfície do SOM da figura 6.4. . . . .</i>	221
6.14	<i>Imagem de uma área da superfície terrestre geograficamente próxima a da figura 6.3.</i>	224
6.15	<i>Imagem da figura 6.14 classificada usando os protótipos do SOM rotulado da figura 6.8. . . . .</i>	225
6.16	<i>Imagem de uma área da superfície terrestre geograficamente próxima a da figura 4.13.</i>	226
6.17	<i>Imagem da figura 6.16 classificada usando os protótipos do SOM rotulado da figura 4.15. . . . .</i>	227



# Lista de Tabelas

2.1	Principais características dos sensores da série LANDSAT (EMBRAPA, 2004). . .	27
2.2	Principais características dos sensores dos satélites do programa CBERS (EMBRAPA, 2004). . . . .	29
2.3	Principais características das bandas espectrais do satélite IKONOS II (EMBRAPA, 2004). . . . .	30
2.4	Classificação da resolução espacial das imagens de sensoriamento remoto. . . . .	32
2.5	Qualidade da classificação associada aos valores do índice de concordância Kappa (adaptada de Landis e Koch (1977)). . . . .	54
4.1	Matriz de confusão das classes (em porcentagem) entre a imagem classificada pelo método proposto e a imagem referência (classificada de modo supervisionado). As linhas da matriz mostram os erros de comissão (porcentagem de pixels que pertencem a outra classe, mas que foi classificado como a classe de interesse) e as colunas mostram os erros de omissão (porcentagem de pixels que pertencem a classe da referência/verdade, mas que foram classificados incorretamente) . . .	128
4.2	Tempo consumido pelas etapas do método de classificação proposto. . . . .	129
5.1	Descrição dos conjuntos de dados utilizados nos testes. . . . .	152
5.2	Tamanhos dos SOMs utilizados e o número de partições obtidas na aplicação da estratégia de análise de agrupamentos proposta na seção 5.2 em cada conjunto de dados. . . . .	154
5.3	Tempos de processamento (em segundos) consumidos pelas versões originais dos índices na avaliação de todas as partições de cada conjunto de dados. . . . .	154

- 5.4 Tempos de processamento (em segundos) consumidos pelas versões modificadas dos índices na avaliação de todas as partições de cada conjunto de dados e média dos *speedups*. . . . . 155
- 5.5 Número de agrupamentos e indicação de acerto (ou não) da melhor partição apontada por cada índice de validação para cada conjunto de dados. O símbolo "O" denota que o índice de validação acertou na indicação da melhor partição e o símbolo "X" denota que o índice errou. A última coluna da tabela mostra a proporção de acertos de cada índice considerando todos os conjuntos de dados. Os números e indicações marcados em negrito na tabela referem-se aos casos em que os comportamentos das duas versões dos índices foram diferentes. . . . . 157
- 5.6 Valores dos índices  $\mathcal{V}_{CDbw}^{SOM}$ ,  $\mathcal{V}_{DBKim}^{SOM}$  e  $\mathcal{V}_{PBM}^{SOM}$  para as 5 partições do SOM da imagem teste 1. . . . . 163
- 5.7 Tempo consumido pelas etapas da metodologia proposta na classificação da imagem teste 1. . . . . 165
- 5.8 Tempos de processamento (em segundos) gastos pelas versões originais e modificadas dos índices de validação na avaliação das cinco partições do SOM da imagem teste 1. Na última linha da tabela, são mostrados os *speedups* de tempo das versões modificadas dos índices em relação as suas versões originais. . . . . 166
- 5.9 Matriz de confusão das classes para a imagem classificada pela metodologia de classificação proposta. As linhas da matriz mostram os erros de comissão (porcentagem de pixels que pertencem a outra classe, mas que foi classificado como a classe de interesse) e as colunas mostram os erros de omissão (porcentagem de pixels que pertencem a classe da referência/verdade, mas que foram classificados incorretamente). . . . . 167
- 5.10 Valores dos índices  $\mathcal{V}_{CDbw}^{SOM}$ ,  $\mathcal{V}_{DBKim}^{SOM}$  e  $\mathcal{V}_{PBM}^{SOM}$  para as 12 partições do SOM da imagem teste 2. . . . . 172
- 5.11 Tempo consumido pelas etapas da metodologia proposta na classificação da imagem teste 2. . . . . 174

---

5.12	Tempos de processamento (em segundos) gastos pelas versões originais e modificadas dos índices de validação na avaliação das 12 partições do SOM da imagem teste 2. . . . .	175
5.13	Matriz de confusão das classes (em porcentagem) entre a imagem classificada pelo método proposto e a imagem referência (resultante da classificação supervisionada). As linhas da matriz mostram os erros de comissão e as colunas mostram os erros de omissão. . . . .	177
6.1	Matriz de confusão das classes para a imagem classificada pela metodologia de classificação proposta. As linhas da matriz mostram os erros de comissão (porcentagem de pixels que pertencem a outra classe, mas que foi classificado como a classe de interesse) e as colunas mostram os erros de omissão (porcentagem de pixels que pertencem a classe da referência/verdade, mas que foram classificados incorretamente). . . . .	215
6.2	Tempo consumido pelas etapas do método de classificação proposto. . . . .	215

# Capítulo 1

## Introdução

### 1.1 Contexto e motivação

Desde os anos 70, período em que foram lançados os primeiros satélites de observação da Terra, o uso de imagens de sensoriamento remoto tem adquirido uma importância crescente em todo o mundo. A ilustração das previsões climáticas nos meios de comunicação por meio de imagens de satélites meteorológicos, a exibição na mídia de áreas de conflitos internacionais e de catástrofes naturais, e a popularização de sistemas de pesquisa e de visualização de mapas (como *Google Earth* e *Google Maps*) mostram claramente que as imagens de satélite vêm ocupando cada vez mais espaço no nosso cotidiano.

Consolidando-se como uma das principais fontes de dados para sistemas de informações geográficas, as imagens adquiridas por sensores remotos orbitais têm sido utilizadas em diversos estudos e levantamentos sobre o ambiente terrestre, ampliando a capacidade do homem em obter informações sobre os recursos naturais e o meio ambiente. O monitoramento ambiental, a avaliação de desflorestamento, a análise de cobertura vegetal, o suporte à previsão de safras, a análise de uso do solo e a identificação de áreas inundáveis são algumas de suas principais aplicações.

Conforme ressaltado em Crósta (1992), as imagens de sensoriamento remoto passaram a representar uma das poucas formas viáveis de monitoramento ambiental em escalas locais e globais. Isso se justifica em razão das vantagens proporcionadas pelos sistemas orbitais de

aquisição de dados, dentre as quais, destacam-se: a possibilidade de obtenção de informações de áreas inóspitas e de difícil acesso, a facilidade do recobrimento de grandes áreas, a cobertura repetitiva com mesma hora local, a grande quantidade de dados pontuais sobre uma mesma área, a transferência de dados Satélite/Terra em tempo real e a capacidade de gerar produtos em diferentes faixas espectrais. A visão sinóptica e os aspectos multiespectral e multitemporal dos sensores orbitais possibilitam a aquisição de informações precisas para a avaliação de diversas variações temáticas da superfície terrestre (Novo, 1989; Pachêco e Ribas, 1998).

Especialmente nesse momento, em que a comunidade mundial começa a encarar com maior seriedade a necessidade de acompanhar as mudanças globais que vêm ocorrendo na superfície do planeta, os sistemas de imageamento em nível orbital firmam-se como ferramentas estratégicas para o futuro. O conhecimento atualizado da distribuição e da área ocupada pela agricultura, vegetação natural, áreas urbanas e edificadas, bem como informações sobre as proporções de suas mudanças, tornam-se cada vez mais necessárias. Compromissos internacionais (como o Protocolo de Kyoto), tomadas de decisões globais sustentáveis, o gerenciamento efetivo de conflitos relacionados à exploração de recursos terrestres, e planos de proteção do meio ambiente, somente são possíveis quando se utiliza uma base de dados que é reconhecida por todos os participantes e que seja capaz de representar a situação atual. Isso tem levado a comunidade mundial a intensificar o uso de informações extraídas a partir de imagens de satélite, reconhecendo e estendendo dessa forma a importância das mesmas nos campos da política, economia e ciência (Pachêco, 2001; Landgrebe, 2005; Kux e Blaschke, 2007).

Conseqüentemente, um número cada vez maior de países e consórcios internacionais tem investido em programas espaciais de lançamento e operação de satélites. O Brasil, em particular, através de uma parceria inédita com a China no setor técnico-científico espacial, ingressou no seleto grupo de países detentores da tecnologia de sensoriamento remoto orbital. Através do conjunto de satélites do programa CBERS (*China-Brazil Earth Resources Satellite*), o Brasil, a partir de 1999, obteve uma poderosa ferramenta para monitorar seu imenso território com satélites próprios de sensoriamento remoto, buscando consolidar uma importante autonomia neste segmento (Epiphanyo, 2005). A base de dados do programa CBERS possui mais de 15.000 usuários de mais de 1.500 instituições cadastradas, distribuindo mais de 300.000 imagens à razão

aproximada de 250 por dia (INPE, 2007).

O aumento expressivo do número de satélites de observação da Terra e o melhoramento contínuo dos sistemas sensores a bordo dos mesmos, que na busca por atender demandas mais sofisticadas dos dados gerados passam a ter resoluções espaciais, espectrais e temporais sem precedentes, certamente aumentam o número de imagens de satélite que são coletadas e arquivadas diariamente. Além disso, o aumento progressivo das capacidades de armazenamento e processamento dos computadores se apresenta como outro componente que colabora significativamente para um armazenamento cada vez maior desses tipos de dados. Esses fatores, em conjunto, têm proporcionado um crescimento substancial no volume de imagens de sensoriamento remoto disponíveis em todo o mundo. A exploração integral dessa enorme quantidade de imagens geradas por satélites cada vez mais numerosos, modernos e sofisticados tornou-se um grande desafio para cientistas e profissionais. Alguns estudos estimaram que somente 5 a 10% de todas as imagens de sensoriamento remoto coletadas estão sendo efetivamente interpretadas (Ehlers, 2007).

É diante desse cenário que nota-se claramente a necessidade de construir e desenvolver novas técnicas de processamento das informações contidas nas imagens de sensores remotos. Atualmente, existem poucos pacotes de software com métodos capazes de processar apropriadamente as altas resoluções e os enormes volumes de dados (Kux e Blaschke, 2007).

Dentre as diferentes técnicas de processamento existentes, particularmente, a classificação de imagens se coloca como um procedimento imprescindível na interpretação quantitativa do gigantesco volume de imagens de satélite disponíveis atualmente. O objetivo geral dos procedimentos de classificação é categorizar cada pixel de uma imagem em uma das várias classes de cobertura terrestre (ou temas), produzindo, essencialmente, um mapa temático da imagem original.

Segundo Richards (2005), novas metodologias de classificação devem ser projetadas de modo que possibilitem uma identificação automática e eficiente do grande volume de dados disponíveis nas imagens e, ao mesmo tempo, tornem o processo de mapeamento da superfície terrestre menos subjetivo e com maior potencial de reuso em situações subseqüentes. De acordo ainda com Blaschke *et al.* (2007), a pesquisa e desenvolvimento na área de sensoriamento remoto

estão diante do desafio de tornar acessível o volume crescente de imagens de satélite a uma comunidade de usuários, cuja tarefa não é o processamento de imagens, mas o planejamento e o gerenciamento. A possibilidade de objetividade, permitindo uma grande independência do técnico-intérprete, deve ser buscada. Isso exige menor especialização na área de classificação de imagens e, portanto, novos métodos de classificação devem permitir também que não especialistas em sensoriamento remoto obtenham bons resultados, em tempo relativamente curto.

Na literatura, há um grande número de algoritmos propostos para a classificação de imagens de sensoriamento remoto. Diferentes técnicas derivadas das mais diversas áreas de pesquisa estão sendo utilizadas como base para o desenvolvimento desses algoritmos, por exemplo, estatística (Dean e Smith, 2003), redes neurais artificiais (Gonçalves, 1997), árvores de classificação (Hansen *et al.*, 1996), lógica *fuzzy* (Wang, 1990), algoritmos genéticos (Pal *et al.*, 2001) e máquinas de vetores-suporte (Mitra *et al.*, 2004).

Embora a literatura apresente uma abundância de abordagens, alguns estudos confirmam que a pesquisa em sensoriamento remoto voltada particularmente para as técnicas de classificação de imagens ainda necessita de avanços mais significativos. Em Wilkinson (2005), resultados e implicações de um estudo de quinze anos sobre experimentos de classificação de imagens de satélite mostram que, embora um número razoável de experiências isoladas demonstrem as vantagens de uma técnica de classificação em relação a outra, o cenário global revela que nenhuma tendência ascendente e significativa em termos de resultados pode ser identificada nas centenas de trabalhos examinados no período de tempo considerado. Segundo o autor, uma possível conclusão sobre este cenário é a de que os níveis de melhoramentos efetuados sobre as técnicas de classificação têm sido muito pequenos para se produzir algum efeito apreciável nesta área de pesquisa. Além disso, conforme ressaltado em Liu *et al.* (2002) e em Campbell (2002), a literatura tem mostrado que nenhum classificador de imagens é "perfeito" ou apresenta um desempenho ótimo para todos os tipos de imagens de sensoriamento remoto. De fato, algoritmos de classificação diferentes produzem resultados diferentes, ainda que aplicados sobre os mesmos conjuntos de dados (Gonçalves *et al.*, 1998b,a; Pal e Mather, 2002; Pasquariello *et al.*, 2002). Em vista disso, e também dos recentes avanços em tecnologias de sensores remotos, esforços devem continuar sendo feitos para se obter maior robustez e melhores desempenhos

por parte dos algoritmos de classificação.

Do ponto de vista de paradigmas de aprendizado, em sensoriamento remoto tradicional, as técnicas de classificação de imagens são divididas em três categorias principais: supervisionada, não-supervisionada e híbrida (Lillesand e Kiefer, 2000). Apesar do número significativo de diferentes métodos de classificação, a maioria deles realiza a classificação das imagens usando o paradigma de aprendizado supervisionado. Neste tipo de classificação ocorre a interação do usuário ou de um analista de imagens que seleciona por meio da imagem original um conjunto de amostras de treinamento para cada uma das classes de interesse. Classificadores estatísticos convencionais, como o método de Máxima Verossimilhança, e redes neurais de *Perceptrons* multicamadas estão entre os métodos supervisionados mais populares (Gonçalves *et al.*, 1998c; Frizzelle e Moody, 2001; Dean e Smith, 2003; Erbek *et al.*, 2004). No entanto, na literatura mais recente, há um crescimento perceptível na sofisticação de métodos de classificação supervisionada, incluindo o uso de algumas técnicas que não eram freqüentemente empregadas pela comunidade de sensoriamento remoto. Em Lawrence *et al.* (2004), por exemplo, técnicas de *boosting* aplicadas no refinamento de árvores de classificação foram exploradas usando uma grande diversidade de padrões espectrais, apresentando uma maior exatidão nos resultados do que métodos de árvores de classificação tradicionais. Em Pal e Mather (2005), uma máquina de vetores-suporte foi utilizada para realizar a classificação supervisionada de imagens de satélite e os resultados mostram que o algoritmo proposto obteve um nível superior de precisão do que classificadores neurais e de máxima verossimilhança. Métodos de classificação supervisionada orientados a objetos, que realizam a análise das imagens integrando valores espectrais, forma e textura, têm sido propostos como alternativas aos métodos supervisionados de classificação pixel a pixel, principalmente para imagens que apresentam alta resolução espacial (Yoon *et al.*, 2004; Kux e Blaschke, 2007).

Considerando que classificadores diferentes possuem capacidades complementares, diversas propostas de métodos combinados que procuram explorar as vantagens de diferentes algoritmos supervisionados têm demonstrado ser uma estratégia poderosa no aumento da acurácia das classificações (Liu *et al.*, 2002; Zhong *et al.*, 2005; Stathakis e Vasilakos, 2006). Notavelmente, *ensembles* de redes neurais supervisionadas também têm sido propostas como uma abordagem



eficiente para o desenvolvimento de sistemas de classificação de imagens. Resultados experimentais registrados na literatura mostram que *ensembles* de redes neurais têm apresentado desempenhos superiores na classificação de dados de sensoriamento remoto em comparação com redes neurais individuais (Giacinto *et al.*, 2000; Espinosa *et al.*, 2004; Van Coillie *et al.*, 2004).

Embora a abordagem supervisionada de classificação seja a mais comumente utilizada, é importante observar que sua performance depende fortemente da qualidade do conjunto de amostras de treinamento. Cada amostra deve ser precisamente escolhida de modo que seja o mais representativa possível para a classe de interesse; caso contrário, os resultados da classificação poderão conter erros ou imprecisões (Swain e Davis, 1978). Como mencionado em Mitra *et al.* (2004), o principal problema encontrado em sistemas de classificação supervisionada é a dificuldade de se ter em algumas imagens uma disponibilidade clara de amostras de treinamento "puras". Geralmente essa dificuldade é amenizada somente quando existem informações sobre verdade terrestre (de referência) da cena, as quais muitas vezes exigem custosos trabalhos de campo para serem adquiridas. Além disso, nesse tipo de classificação, os dados de treinamento são definidos primariamente com base em categorias de informação, e somente secundariamente com base em propriedades espectrais. Tal estratégia pode impor uma estrutura de classificação aos dados da imagem que não necessariamente condiz com a sua estrutura real (Campbell, 2002). Diferenças em conjuntos de dados de treinamento podem produzir diferenças significativas nos resultados de classificação por métodos supervisionados, principalmente no caso de imagens com alta resolução espacial (Chen e Stow, 2002). Por essas razões, algumas estratégias diferenciadas para selecionar e filtrar amostras de treinamento em imagens de sensoriamento remoto têm sido registradas na literatura (Arai, 1992; Foody, 1999; Foody e Mathur, 2004). Em particular, métodos de aprendizagem ativa têm sido utilizados extensivamente para reduzir requisitos de dados em tarefas de classificação (Cohn *et al.*, 1994; Bachmann, 2003; Mitra *et al.*, 2004).

Alternativamente, métodos de classificação híbrida, que combinam os paradigmas de aprendizado supervisionado e não-supervisionado, são também utilizados para facilitar a identificação de classes e a aquisição de amostras de treinamento (Lillesand e Kiefer, 2000). A maioria dos métodos híbridos de classificação efetua uma estratificação inicial da imagem através de téc-

nicas de análise de agrupamentos, associam os agrupamentos descobertos às classes definidas pelo usuário e, então, aplicam um método supervisionado para classificar todos os pixels da imagem. O resultado obtido com a aplicação inicial de um método não-supervisionado forma a base para a seleção de amostras de treinamento que serão utilizadas pelo método supervisionado subsequente. Além disso, o uso precedente de um método não-supervisionado pode ajudar o analista a identificar classes que aparecem com uma densidade muito baixa na imagem. Essas classes provavelmente não seriam reconhecidas em um processo de classificação exclusivamente supervisionado, especialmente no caso de imagens que possuem uma variação muito complexa de padrões espectrais e não apresentam nenhuma informação sobre verdade terrestre (Gonçalves *et al.*, 1996; Kelly *et al.*, 2004; Lo e Choi, 2004).

Diferentemente da classificação supervisionada, métodos não-supervisionados de classificação de imagens de sensoriamento remoto requerem condições mínimas sobre o conjunto de dados de treinamento e estão baseados no princípio de que o algoritmo computacional é capaz de identificar por si só as classes da imagem. Esta abordagem de classificação desempenha um papel especialmente importante quando se tem pouca informação *a priori* sobre os dados da imagem e, por esta razão, se apresenta como uma opção interessante para analistas sem um conhecimento amplo de campo ou para aqueles usuários que procuram minimizar interações e/ou possíveis influências no processo de classificação (Duda e Canty, 2002; Wang e Zhou, 2002; Kelly *et al.*, 2004). Uma vez que não exigem o elevado grau de pressuposições e de decisões detalhadas demandado pelos métodos de classificação supervisionada, as técnicas de classificação não-supervisionadas podem ser consideradas como ferramentas essenciais para a exploração em grande escala do gigantesco volume de imagens de satélite que estão sendo adquiridas e armazenadas atualmente (Gonçalves *et al.*, 2008b).

A classificação não-supervisionada de imagens de sensoriamento remoto é frequentemente realizada através de métodos de análise de agrupamentos (ou de *clustering*). Esses métodos examinam os pixels desconhecidos de uma imagem e os incorporam dentro de um conjunto de classes definidas através dos agrupamentos naturais de seus níveis de cinza. A premissa básica é a de que pixels que pertencem à mesma classe de cobertura terrestre possuam atributos iguais ou similares, enquanto que pixels pertencentes a classes diferentes estejam em sua maioria

relativamente distantes uns dos outros no espaço de atributos. A análise de agrupamentos constitui-se como um método prático para organizar grandes volumes de dados permitindo que a busca por informação seja feita de maneira mais eficiente. Entretanto, embora haja uma grande quantidade de diferentes métodos de análise de agrupamentos na área de reconhecimento de padrões (Xu e Wunsch II, 2005), somente um número limitado deles pode ser utilizado efetivamente em aplicações de sensoriamento remoto. Como destacado por Tran *et al.* (2005), diversos problemas são encontrados quando se realiza a análise de agrupamentos de imagens de sensoriamento remoto. No entanto, os problemas relacionados com o tamanho das imagens e a dimensão dos seus espaços de atributos são aqueles que normalmente inviabilizam a aplicação de um método de agrupamentos.

Os métodos de análise de agrupamentos mais freqüentemente aplicados em dados de sensoriamento remoto são os denominados particionais. Métodos particionais são aqueles que produzem uma partição dos dados em  $K$  agrupamentos, geralmente otimizando uma função objetivo. A maioria dos softwares ou sistemas computacionais voltados para o processamento digital de imagens de sensoriamento remoto contém especificamente dois métodos de análise de agrupamentos particionais para efetuar classificações não-supervisionadas: o algoritmo K-médias, e uma variação dele, o ISODATA (Iterative Self-Organising Data Analysis Technique) (Ball e Hall, 1967).

Apesar de serem comumente utilizados, os algoritmos K-médias e ISODATA possuem limitações importantes. Suas funções objetivo partem do pressuposto de que o número de classes,  $K$ , é conhecido *a priori*. Na hipótese de que um valor inadequado  $K'$  tenha sido definido, o método irá impor, através do uso de técnicas de otimização,  $K'$  agrupamentos aos dados. O usuário também necessita especificar manualmente vários parâmetros para controlar a análise de agrupamentos, por exemplo, os centróides iniciais de cada grupo de dados, o número máximo de iterações, limiares para efetuar a divisão, fusão ou exclusão de grupos. Os algoritmos K-médias e ISODATA são muito sensíveis a esses parâmetros, podendo gerar diferentes partições quando várias simulações são realizadas para um mesmo conjunto de dados. Diante disso, os valores ótimos desses parâmetros são muitas vezes determinados experimentalmente por tentativa e erro. Essas necessidades certamente aumentam o nível de interação do usuário

com o algoritmo computacional, e conseqüentemente aumentam o grau de subjetividade do processo de categorização da imagem. Outras limitações não menos importantes de algoritmos particionais, como o K-médias e o ISODATA, são o alto custo computacional quando o volume de dados a ser analisado é muito grande (em cada iteração todos os pixels da imagem são comparados com todos os centróides) e a existência de suposições sobre as formas dos agrupamentos. Geralmente apenas um único protótipo (centróide) é usado para representar um agrupamento, assim esses métodos tornam-se adequados somente para analisar conjuntos de dados com formatos hipersféricos (Shah *et al.*, 2004; Gonçalves *et al.*, 2005b).

Outra maneira possível, mas não usual, de efetuar a classificação não-supervisionada de imagens de sensoriamento remoto é através de métodos de agrupamentos hierárquicos. Diferentemente dos métodos particionais, métodos hierárquicos não exigem que o usuário especifique previamente o número de agrupamentos e outros parâmetros adicionais e, portanto, são menos afetados por condições iniciais (Richards, 1995). Outra vantagem significativa desses métodos é a possibilidade de visualizar os resultados da classificação por meio de um dendrograma que ilustra de forma hierárquica o grau de similaridade entre os agrupamentos que são formados pelas fusões (ou divisões) em cada estágio sucessivo da análise. Entretanto, métodos hierárquicos apresentam algumas características que dificultam a sua aplicação na classificação de imagens de sensoriamento remoto: (a) em geral requerem espaço de memória da ordem de  $O(N^2)$ , onde  $N$  é o número de registros do conjunto de dados; (b) dendrogramas são impraticáveis para grandes conjuntos de dados e, portanto, os resultados podem ser difíceis de interpretar; (c) para determinar qual a linha de "corte" do dendrograma (número de agrupamentos) há a necessidade de se aplicar algum critério de decisão (Tran *et al.*, 2005; Xu e Wunsch II, 2005).

Conseqüentemente, módulos de agrupamentos hierárquicos são raros em softwares ou sistemas computacionais de sensoriamento remoto e, por esta razão, são normalmente utilizados de uma maneira desacoplada. Na literatura, não há muita pesquisa ou trabalhos que empregam métodos hierárquicos na classificação não-supervisionada de imagens de sensoriamento remoto e, em geral, versões comuns desses métodos são aplicadas na prática quando o objetivo principal é basicamente comparar desempenhos com outros métodos não-supervisionados de classificação

(Duda e Canty, 2002; Wilson *et al.*, 2002). Entretanto, um crescimento relativamente recente de abordagens que procuram viabilizar o uso de métodos hierárquicos em aplicações de sensoriamento remoto pode ser notado. Como exemplo, ao invés de aplicar métodos hierárquicos aglomerativos diretamente sobre uma base pixel a pixel, alguns trabalhos têm empregado esses métodos somente depois da aplicação de alguma outra técnica de agrupamentos sobre os dados originais da imagem. Em Tran *et al.* (2003), o algoritmo K-médias foi utilizado para aglomerar os dados da imagem antes da aplicação de um método hierárquico, e o ISODATA foi usado com a mesma finalidade em Marçal e Castro (2005) e em Marçal e Borges (2005). Embora essa estratégia seja efetiva para evitar a aplicação direta (e custosa) de métodos hierárquicos em grandes conjuntos de dados, uma possível desvantagem da mesma é que os resultados das metodologias de classificação apresentadas nesses trabalhos podem ser influenciados pelas limitações dos algoritmos K-médias e ISODATA.

Embora não sejam encontradas com facilidade em pacotes de processamento digital de imagens de sensoriamento remoto, algumas propostas de métodos não-supervisionados de classificação que não se baseiam apenas em técnicas convencionais de análise de agrupamentos podem ser encontradas na literatura. Em Huang (2002), por exemplo, um método hierárquico divisivo baseado em conceitos de hiperplanos e árvores binárias foi proposto como uma alternativa para o ISODATA. Embora a acurácia global do algoritmo hierárquico proposto tenha se mostrado pouco superior, em todos os testes experimentais apresentados o seu tempo de processamento foi bem inferior ao do ISODATA. Shah *et al.* (2004) propuseram um algoritmo não-supervisionado usando análise de componentes independentes que emprega estatísticas de ordem superior e, contrariamente às técnicas convencionais, o algoritmo proposto modela as distribuições das classes através de densidades não-gaussianas. Em Baraldi *et al.* (2000) e em Tso e Olsen (2005), técnicas de classificação não-supervisionada baseadas em campos aleatórios de Markov com uso combinado de informações espectrais e espaciais, também foram investigadas. No entanto, os algoritmos apresentam uma alta complexidade computacional e consomem muito tempo de processamento. Adicionalmente, pesquisadores também têm projetado abordagens não-supervisionadas para classificar imagens de sensoriamento remoto baseadas em métodos de inteligência computacional, como redes neurais, sistemas nebulosos e computação evolutiva

(Ji, 2003; Bandyopadhyay, 2005; Liu *et al.*, 2005; Xiang *et al.*, 2005).

Em particular, alguns estudos têm investigado o uso de Mapas Auto-organizáveis de Kohonen (*SOM - Self-Organizing Map*) em tarefas de análise de agrupamentos e de classificação de imagens de sensoriamento remoto. Propriedades importantes como a aproximação do espaço de entrada, ordenação topológica e casamento de densidade, aliadas à simplicidade do modelo e à facilidade de implementação do seu algoritmo de aprendizagem, fazem do SOM uma ferramenta promissora para este tipo de aplicação. Ji (2000) mostrou que o SOM é capaz de alcançar desempenhos de classificação superiores quando comparado com o método de Máxima Verossimilhança, e concluiu que a rede neural é uma alternativa viável e útil para a classificação de imagens de sensoriamento remoto. Vilmann *et al.* (2003) estudaram o efeito da aplicação de extensões do SOM na análise de agrupamentos de imagens de satélite multispectrais e hiperspectrais. Os autores mostraram bons resultados de análise de agrupamentos e de classificações através do SOM, e concluíram que algumas considerações sobre suas propriedades e sobre a interpretação do mapa treinado podem ser ainda investigadas. Aplicando o SOM para monitorar queimadas de florestas a partir de imagens multiespectrais de regiões próximas à cidade de Berlim, Schaale e Furrer (1995) relatam que as conexões laterais entre os neurônios do SOM revelam características interessantes da superfície terrestre a partir do espaço topológico da rede. Em Hadjimitsis *et al.* (2003), um método de classificação não-supervisionada usando o SOM foi proposto para realizar o mapeamento de mudanças temporais de cobertura terrestre. Os resultados experimentais mostram que a aplicação do SOM na classificação de imagens de satélite é vantajosa e apresenta um espaço considerável para estudos e pesquisas futuras.

Embora alguns trabalhos já publicados coloquem o SOM como uma ferramenta eficaz para realizar a análise de agrupamentos de imagens de sensoriamento remoto, é importante observar que a detecção de fronteiras ou limites de agrupamentos através do espaço de saída da rede não é uma tarefa trivial. O SOM necessita de ferramentas adicionais para que uma interpretação adequada do seu mapa de saída seja realizada (Costa, 1999; Vesanto e Alhoniemi, 2000). Além disso, conforme ressaltado em Ji (2000), o processo de segmentação do SOM treinado pode afetar a representação das classes pelas quais a imagem original será classificada e comprometer os resultados da classificação. Nesse sentido, a maioria dos métodos de classificação de imagens

baseados no SOM realiza a detecção de agrupamentos com a ajuda de um analista de imagens que segmenta manualmente o mapa de saída da rede usando algum conhecimento prévio sobre os dados e/ou ferramentas alternativas para facilitar a visualização dos agrupamentos através do mapa de neurônios (Merkl e Rauber, 1997; Vesanto, 1999). No entanto, conforme relatado em Mather (2004), o que se espera em abordagens de classificação não-supervisionada, cujo objetivo básico é descobrir a estrutura inerente dos dados, é que a detecção de agrupamentos da imagem seja realizada de modo que a intervenção do usuário seja a menor possível.

## 1.2 Objetivos

O objetivo principal deste trabalho é explorar as características e propriedades do SOM na construção de métodos de classificação não-supervisionada de imagens de sensoriamento remoto. O ponto chave dos métodos de classificação propostos é realizar o processo de análise de agrupamentos das imagens através de um conjunto de protótipos do SOM, ao invés de trabalhar diretamente com os padrões originais das cenas.

Para atender tal objetivo, os seguintes procedimentos foram adotados:

- Avaliar a aplicabilidade de métodos de segmentação automática do SOM, particularmente daqueles baseados em técnicas de morfologia matemática e em métodos de agrupamentos hierárquicos. Mecanismos para analisar e interpretar os neurônios de um SOM treinado são aplicados e aprimorados com o objetivo de particionar automaticamente o conjunto de dados das imagens. Diferentemente de outras aplicações nas quais o SOM é utilizado como ferramenta de auxílio visual para indicar tendências de agrupamentos, nos métodos de classificação propostos neste trabalho, a detecção de agrupamentos através dos neurônios de um SOM treinado é realizada sem a necessidade de participação do usuário ou de um analista de imagens;
- Propor a utilização de índices de validação de agrupamentos na determinação automática do número de agrupamentos, bem como da melhor partição para o conjunto de dados das imagens. Diversos índices de validação são analisados e aplicados para avaliar a qualidade de diferentes configurações de agrupamentos de dados obtidas através do SOM.

Baseando-se ainda nas propriedades estatísticas do SOM, modificações nos cálculos dos índices de validação considerados também são propostas com o objetivo de reduzir o custo computacional do processo de classificação das imagens;

- Aplicar técnicas de análise de textura em imagens para avaliar e filtrar amostras de treinamento e/ou protótipos do SOM que correspondem a regiões de transição entre classes de cobertura terrestre. Tais protótipos, juntamente com os pixels da imagem associados a eles, são tratados de maneira especial procurando garantir um melhor desempenho na classificação das imagens;
- Incorporar informações espaciais a respeito dos protótipos do SOM em critérios de fusão de agrupamentos, além das informações de distância multiespectral. Medidas de compactidade e de fronteira espaciais são utilizadas para facilitar a discriminação de classes de cobertura terrestre que apresentam alto grau de similaridade espectral;
- Comparar o desempenho dos métodos de classificação propostos com métodos convencionais de classificação não-supervisionada de imagens de sensoriamento.

### 1.3 Organização do texto

O presente documento está dividido em sete capítulos. Este primeiro capítulo apresenta o contexto dentro do qual o trabalho foi desenvolvido, seus aspectos relevantes e objetivos.

O capítulo 2 apresenta os fundamentos teóricos do sensoriamento remoto, descrevendo particularmente os mecanismos e sistemas envolvidos no processo de aquisição de dados em nível orbital. O capítulo 2 também apresenta as características e as principais técnicas de processamento digital de imagens de sensores remotos, onde os paradigmas e métodos tradicionais de classificação de imagens são abordados de maneira especial.

O capítulo 3 descreve as características e propriedades do SOM, além de apresentar seu algoritmo de aprendizagem e considerações sobre seus parâmetros de treinamento. O mesmo capítulo apresenta, ainda, uma breve revisão sobre métodos que buscam segmentar de maneira automática os neurônios de um SOM treinado, descrevendo em particular o método de segmentação do SOM proposto por Costa e Andrade Netto (1999, 2001), o qual serviu de base para



os métodos de classificação apresentados nos capítulos 4 e 5 .

Os capítulos 4, 5 e 6 apresentam os métodos de classificação de imagens de sensoriamento remoto propostos neste trabalho. O capítulo 4 apresenta um metodologia de classificação que emprega técnicas de morfologia matemática para realizar o particionamento do SOM. Considerações sobre o processo de amostragem das imagens e explicações sobre os parâmetros adotados para o treinamento do SOM também são relatadas. O capítulo 4 mostra, ainda, a aplicação do método de classificação proposto e compara o seu desempenho com o do algoritmo K-médias.

O capítulo 5 apresenta um aprimoramento da metodologia de classificação apresentada no capítulo 4, onde índices de validação de agrupamentos são aplicados para avaliar a qualidade de diferentes partições dos conjuntos de dados das imagens. Uma estratégia para simplificar o cálculo dos índices baseada na abstração de dados produzido pelo SOM também é apresentada, incluindo uma revisão sobre diferentes índices de validação propostos na literatura e uma análise experimental comparativa sobre um conjunto de índices pré-selecionados. O capítulo 5 apresenta, também, um método de filtragem de amostras baseado em medidas de textura para acentuar as bordas da *U-matrix*.

Diferentemente dos métodos apresentados nos capítulos 4 e 5, o capítulo 6 deste trabalho apresenta um método de classificação de imagens que emprega um método de agrupamento hierárquico com restrições de conectividade, e que incorpora informações espaciais, além de informações de distância, para segmentar os protótipos do SOM. Comparações do método de classificação proposto com o algoritmo K-médias padrão, e com uma versão "janelada" do mesmo, são apresentadas.

O capítulo 7 apresenta algumas considerações finais sobre os métodos de classificação propostos, uma análise das principais contribuições deste trabalho e perspectivas de trabalhos futuros.

Finalmente, esta tese apresenta, ainda, uma mídia (CD) com informações adicionais sobre os desenvolvimentos e testes experimentais realizados neste trabalho.

## Capítulo 2

# Processamento Digital de Imagens de Sensoriamento Remoto

### 2.1 Introdução

A função principal do processamento digital de imagens de sensoriamento remoto é a de prover técnicas e ferramentas que facilitem a identificação e extração de informações contidas nas imagens sobre fenômenos ou alvos da superfície terrestre. Para isso, sistemas dedicados de computação são utilizados para realizar atividades de análise e tratamento de imagens brutas coletadas pelos sensores remotos. Esses sistemas computacionais basicamente apresentam técnicas que tentam simular a capacidade do sistema visual humano em reconhecer padrões espectrais existentes em uma cena da superfície terrestre (representada pelo conjunto de pixels da imagem). O resultado desse processo é a produção de outras imagens, contendo informações específicas, extraídas e/ou realçadas.

A grande vantagem proporcionada pelo uso de técnicas computacionais no reconhecimento de padrões em imagens é a capacidade de processar o grande volume de informações presentes numa cena. Embora o sistema visual humano possua uma habilidade notável de identificar padrões, dificilmente ele é capaz de analisar rapidamente a enorme quantidade de informações detalhadas existentes em uma imagem. Diferentes tipos de ruídos e distorções, ocasionados pelos processos de aquisição, transmissão e visualização de imagens coletadas por sensores re-

mentos, também colaboram para limitar ainda mais a capacidade do olho humano (Crósta, 1992). As técnicas do processamento digital de imagens de sensoriamento visam, portanto, eliminar essas barreiras, promovendo a extração de informações das imagens através da automatização da análise dos dados contidos nas mesmas, substituindo assim procedimentos de interpretação visual.

Um aspecto importante, que deve ser considerado quando se pretende estudar e desenvolver técnicas para o processamento digital de imagens de sensoriamento remoto, é o fato dessas imagens serem adquiridas em diferentes faixas do espectro eletromagnético, o que as confere uma característica multiespectral. Apresentando essa característica, as imagens de sensoriamento remoto geralmente se diferenciam de outras imagens digitais. Em vista disso, conforme ressalta Mather (2004), para que se possa entender as propriedades de uma imagem de sensoriamento remoto e interpretar adequadamente os padrões de cobertura terrestre existentes na mesma é importante compreender os mecanismos envolvidos no seu processo de aquisição.

Nesse sentido, este capítulo tem como objetivo apresentar os fundamentos do sensoriamento remoto, descrevendo basicamente os processos interativos da radiação eletromagnética com a superfície terrestre, os diferentes níveis de aquisição de dados e, em particular, os mecanismos e sistemas envolvidos no processo de aquisição de imagens em nível orbital. Posteriormente, as características das imagens de sensoriamento remoto e suas principais técnicas de processamento são apresentadas, onde os paradigmas e métodos tradicionais de classificação são abordados de maneira especial.

## 2.2 Fundamentos do Sensoriamento Remoto

Segundo Novo (1989), o Sensoriamento Remoto pode ser definido como a utilização conjunta de modernos sensores, equipamentos para processamento e transmissão de dados, aeronaves, espaçonaves, etc., com o objetivo de estudar o ambiente terrestre através do registro e da análise das interações entre a radiação eletromagnética e as substâncias componentes do planeta em suas mais diversas manifestações.

De uma maneira geral, as diferentes técnicas de sensoriamento remoto são basicamente

constituídas por quatro elementos fundamentais, os quais podem ser representados através do esquema apresentado na figura 2.1.

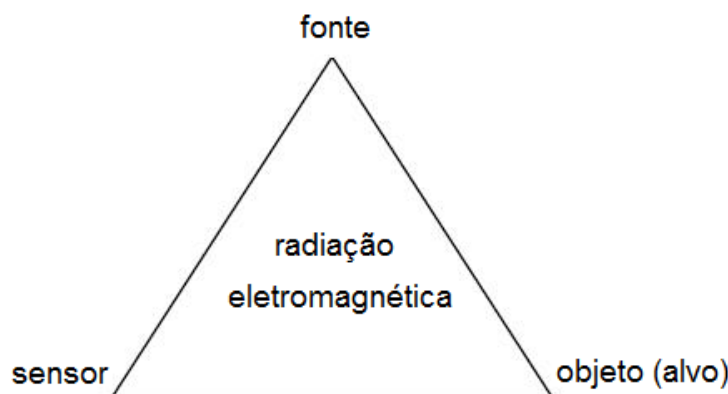


Fig. 2.1: *Esquema representativo dos quatro elementos fundamentais das técnicas de sensoriamento remoto (adaptada de Castro Jr. (2001)).*

No centro do triângulo encontra-se a radiação eletromagnética, que é o elemento de ligação entre todos os seguintes elementos (que se encontram nos vértices): a *fonte* de energia eletromagnética (geralmente o Sol, mas também pode ser a Terra para os sensores passivos termais e antenas de microondas para os sistemas radares); o *sensor*, que é o instrumento que coleta e registra a radiação eletromagnética refletida ou emitida pelo *objeto*, o qual também é denominado *alvo*, e que representa o elemento do qual se pretende extrair informação.

Através desse esquema fica claro que a base sobre a qual se assenta a aquisição de informações em sensoriamento remoto é a interação entre radiação eletromagnética e matéria. O Sol é a principal fonte de energia para todo o sistema solar e, devido à sua elevada temperatura, gera uma grande quantidade de energia que é irradiada para todo o espaço. Para o sensoriamento remoto, a radiação solar é a base de todos os princípios em que se fundamenta essa tecnologia.

### 2.2.1 Radiação solar e espectro eletromagnético

De acordo com a literatura, a radiação solar apresenta uma natureza dual. Ora se comporta como corpúsculos, ora como onda eletromagnética. No sensoriamento remoto tradicional, a

teoria é fundamentada quase toda no conceito ondulatório, principalmente quando se trata de aplicações (Cracknell e Hayes, 1991).

Sob uma perspectiva ondulatória, a radiação se propaga na forma de ondas formadas pela oscilação dos campos elétrico e magnético, caracterizando-se dessa maneira pelo comprimento da onda ou pela frequência de oscilação. O comprimento é definido como a distância que separa dois pontos semelhantes da onda (duas cristas ou cavados consecutivos). O comprimento de onda é normalmente expresso em centímetros ou em micra ( $1 \text{ micra} = 10^{-4} \text{ cm}$ ). A frequência corresponde ao número de vezes que a onda passa por um ponto de referência em determinado intervalo de tempo, ou seja, ao número de oscilações por unidade de tempo. A frequência é geralmente expressa em ciclos por segundo, ou Hertz (Hz).

Existem diversas denominações para as radiações eletromagnéticas, que variam de acordo com a frequência e o comprimento de onda. De acordo com Moreira (2005), essas diferentes radiações receberam denominações de natureza histórica ou decorrentes dos processos utilizados na sua produção, sendo classificadas conforme seguem:

- **Radiação gama:** emitida por materiais radioativos e pelo Sol. Possui altas frequências e, por isso, é muito penetrante (alta energia). Na prática, tem aplicações na medicina (radioterapia) e em processos industriais, principalmente na conservação de alimentos.
- **Raios X:** radiações cujas frequências de onda estão acima daquelas da radiação ultravioleta, ou seja, possuem comprimentos de onda menores. Esta denominação foi dada por seu descobridor, o físico alemão Wilhelm Röntgen, em 1895, por não conhecer suas trajetórias. Os raios X surgem do interior da eletrosfera do átomo, por rearranjos eletrônicos. São muito usados em radiografias e em estudos de estruturas cristalinas de sólidos. Os raios X provenientes do Sol são absorvidos pelos gases na alta atmosfera.
- **Radiação ultravioleta (UV):** conjunto de radiações compreendidas na faixa espectral de  $0,01$  a  $0,40 \mu\text{m}$ . Estas radiações são muito produzidas durante as reações nucleares no Sol. Entretanto, ao atingir o topo da atmosfera terrestre, são quase totalmente absorvidas pelo gás ozônio. O espectro do UV é dividido em três bandas: UV próximo ( $0,32$  a  $0,40 \mu\text{m}$ ), UV distante ( $0,28$  a  $0,32 \mu\text{m}$ ) e UV máximo ( $0,1$  a  $0,28 \mu\text{m}$ ).

- **Radiação visível (luz):** conjunto de radiações eletromagnéticas compreendidas entre 0,39 e 0,70  $\mu\text{m}$ . As radiações contidas nesta faixa de comprimento de onda, ao incidirem no sistema visual humano, são capazes de provocar uma sensação de cor no cérebro.
- **Radiação infravermelha (IV):** conjunto de radiações eletromagnéticas cujos comprimentos de onda variam de 0,7 a 1.000  $\mu\text{m}$ . Situam-se no espectro eletromagnético entre a luz vermelha e as microondas. Esta radiação é dividida em três faixas espectrais: próximo (0,7 a 1,1  $\mu\text{m}$ ), médio (1,1 a 3,0  $\mu\text{m}$ ) e distante (3,0 a 1.000  $\mu\text{m}$ ). Às vezes recebem a denominação de radiação térmica.
- **Microondas:** radiações eletromagnéticas que se estendem pela região do espectro de 1.000  $\mu\text{m}$  até cerca de 1 m. São mais comumente referenciadas em Hertz.
- **Ondas de rádio:** conjunto de radiações com frequências menores que 300 MHz (comprimento de onda maior que 1 m). Estas ondas são utilizadas principalmente em telecomunicação e radiodifusão.

O conjunto de todas estas radiações, desde os raios gama até as ondas de rádio, forma o espectro eletromagnético, que nada mais é do que a ordenação destas radiações em função do comprimento de onda e da frequência (figura 2.2).

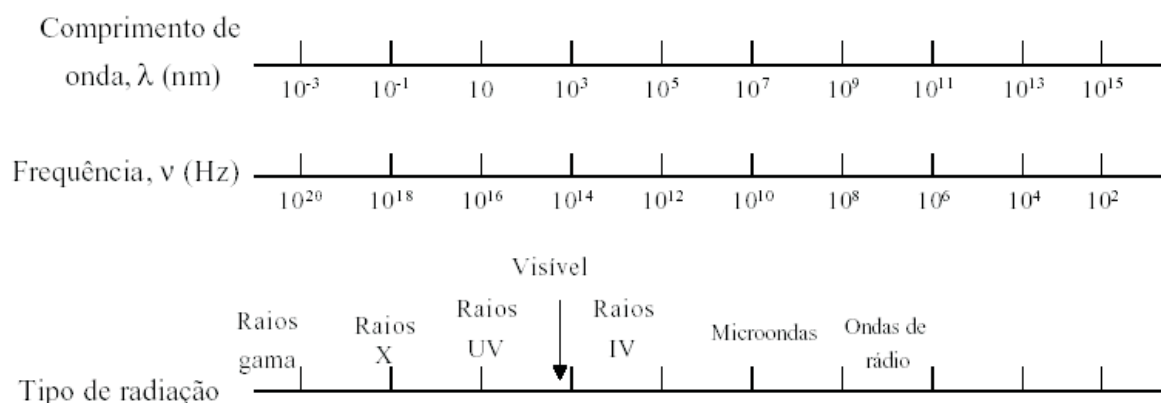


Fig. 2.2: *Espectro eletromagnético.*

Conforme descrito adiante, as imagens de sensoriamento remoto são adquiridas em diferentes porções do espectro eletromagnético, desde o ultravioleta até o infravermelho distante.

### 2.2.2 Interação da radiação com o ambiente terrestre

Independentemente de sua fonte, toda radiação detectada por sensores remotos percorre algum caminho através da atmosfera, podendo sofrer atenuações associadas aos processos de *reflexão*, *espalhamento* e *absorção* causadas pelos constituintes atmosféricos, por partículas dispersas ou nuvens. Dependendo do grau e do tipo da atenuação, a radiação global que chega à superfície terrestre pode ser fortemente reduzida.

Devido à grande variabilidade de gases que compõem a atmosfera terrestre em diferentes zonas e camadas, a radiação solar sofre vários processos de refração. Além da refração, alguns gases constituintes, como o ozônio e  $CO_2$ , absorvem a radiação solar em comprimentos de ondas específicos. Essa absorção pode ser total, como é o caso da radiação ultravioleta (absorvida pelo ozônio), ou parcial, como ocorre em quase todo o espectro a partir de  $0,3 \mu m$ .

Existem também outros fatores atenuantes, como partículas dispersas e presença de nuvens, que provocam o espalhamento da radiação. O espalhamento é um processo físico que ocorre em função do desvio das ondas eletromagnéticas por partículas existentes nas suas trajetórias. Esse desvio pode ser tanto da energia incidente quanto da refletida.

Do ponto de vista do sensoriamento remoto, as nuvens são a maior barreira à radiação solar. Compostas por gotículas e/ou cristais de gelo, as nuvens podem ser consideradas como um aerossol concentrado no espaço, impedindo a passagem da radiação solar na faixa refletiva do espectro eletromagnético, onde operam a maioria dos sistemas sensores.

Segundo Moreira (2005), a radiação solar que volta para o espaço corresponde a 37% da incidente, sendo 26% refletida pelas nuvens e 11% pela dispersão das partículas que se encontram na atmosfera. Os gases e vapor de água são responsáveis pela absorção de 16% da radiação. Dessa forma, somando a radiação que volta para o espaço com aquela que foi absorvida pelos gases e vapor d'água, tem-se um total de 53% de perda da radiação global incidente. Assim, do total que chega ao topo da atmosfera somente 47% atinge a superfície terrestre (figura 2.3). A radiação global é composta de raios solares diretos (19%) e raios difusos (26%).

Ao alcançar uma região qualquer da superfície terrestre, a radiação eletromagnética incidente interage de acordo com as características do sinal e do objeto da superfície em questão. Como resultado desta interação, uma parte da energia incidente é refletida, outra porção é

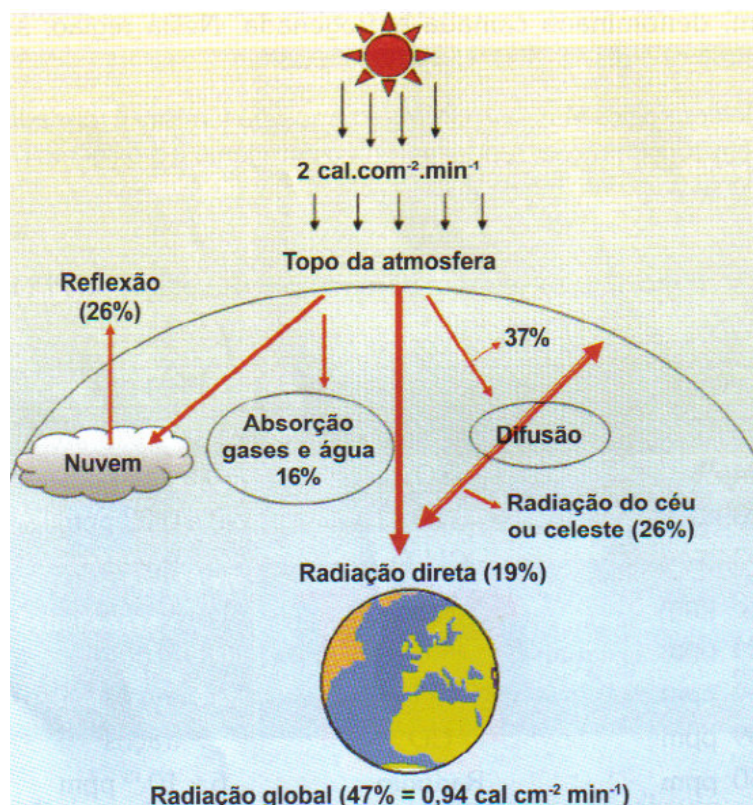


Fig. 2.3: *Causas das atenuações da radiação solar na atmosfera terrestre (adaptado de Moreira (2005)).*

absorvida, e, ainda, uma outra fração é transmitida. A capacidade de um objeto absorver, refletir e transmitir a radiação eletromagnética incidente é denominada, respectivamente, de absorvância, refletância e transmitância.

Os objetos interagem de maneira diferenciada espectralmente com a energia eletromagnética incidente, uma vez que apresentam diferentes propriedades físico-químicas e biológicas. Estas diferentes interações possibilitam a distinção e o reconhecimento dos diversos objetos terrestres sensoriados remotamente, pois são reconhecidos devido a variação da porcentagem de energia refletida em cada comprimento de onda (ou faixa do espectro eletromagnético).

O comportamento espectral de um objeto pode ser definido como sendo o conjunto dos valores sucessivos da refletância do objeto ao longo do espectro eletromagnético, também conhecido como a *assinatura espectral do objeto*. A assinatura espectral do objeto define as feições deste, sendo que a forma, a intensidade e a localização de cada faixa de absorção é que caracteriza o



objeto (Moraes, 2006).

O gráfico que apresenta a variação da refletância espectral de um objeto em função do comprimento de onda é chamado *curva de refletância espectral*. A figura 2.4 apresenta as curvas de refletância espectral médias para as classes vegetação, solo exposto e água.

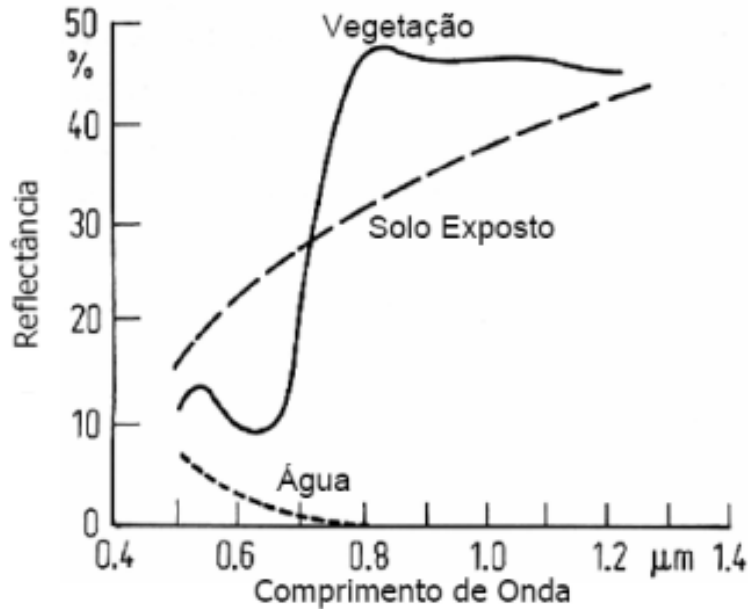


Fig. 2.4: *Curvas de refletância espectral de diferentes classes de uso do solo (adaptado de Richards (1995)).*

Maiores detalhes sobre a interação da radiação solar com diferentes tipos de alvos (ou materiais) da superfície terrestre podem ser obtidos em: Novo (1989); Cracknell e Hayes (1991); Richards (1995); Campbell (2002); Moreira (2005).

### 2.2.3 Sistemas sensores

Os sensores remotos são dispositivos capazes de detectar a radiação eletromagnética proveniente de um objeto em determinadas faixas do espectro eletromagnético, transformá-las em um sinal elétrico e registrar esse sinal de tal forma que ele possa ser armazenado ou transmitido em tempo real, para posteriormente ser convertido num produto passível de interpretação, seja na forma de imagem, gráficos ou tabelas.

Conforme mostra a figura 2.5, os sistemas sensores são basicamente formados por uma parte

óptica (coletor), constituída por lentes e espelhos, que tem o objetivo de captar e direcionar a energia proveniente dos alvos para os detectores. A energia registrada é processada obtendo-se o produto que contém a informação disponível ao usuário.

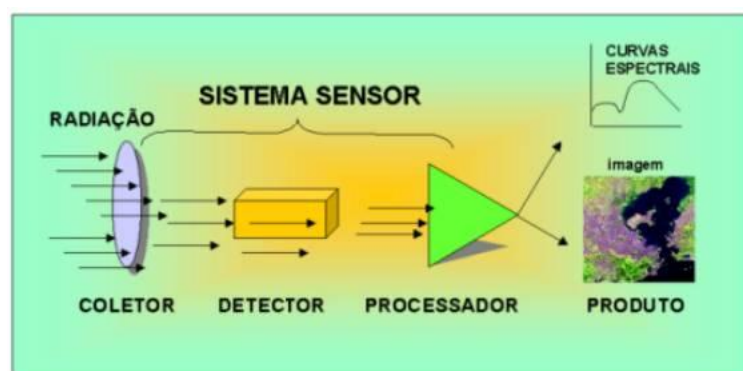


Fig. 2.5: Componentes de um sistema sensor (adaptado de Moreira (2005)).

Os sistemas sensores podem ser classificados em *não-imageadores* e *imageadores*.

Os sistemas não-imageadores, também denominados radiômetros ou espectroradiômetros, apresentam o resultado em forma de dígitos ou gráficos. Esses sensores integram a energia no tempo, no espaço e em comprimento de onda, para produzir uma curva espectral que caracteriza a potência eletromagnética emitida, refletida ou transmitida pelo alvo. São essenciais para a aquisição de informações minuciosas sobre o comportamento espectral dos objetos da superfície terrestre.

Os sistemas imageadores fornecem como produto uma imagem da área observada, como por exemplo, os sistemas de quadro (*framing systems*), que adquirem a imagem da cena em sua totalidade num mesmo instante, e os sistemas de varredura (*scanning systems*), que obtêm a imagem da cena pela aquisição seqüencial de "imagens elementares do terreno" ou "elementos de resolução", também chamados de pixels. As características dos sistemas sensores imageadores (não-fotográficos) são expressas por quatro tipos de resolução, descritas como seguem (Moraes, 2006):

- **Resolução espacial:** representa a capacidade do sensor em distinguir objetos. Ela indica o tamanho do menor elemento da superfície individualizado pelo sensor. A resolução espacial depende principalmente do detector, da altura do posicionamento do sensor em

relação ao objeto. Para um dado nível de posicionamento do sensor, quanto menor for a resolução geométrica deste, maior será o grau de distinção entre objetos próximos. Por exemplo, o sistema sensor CCD a bordo do satélite CBERS possui uma resolução espacial de 20 metros.

- **Resolução espectral:** refere-se à faixa espectral em que opera o sensor. Portanto, ela define o intervalo espectral no qual são realizadas as medidas, e conseqüentemente a composição espectral do fluxo de energia que atinge o detector. Quanto maior for o número de medidas num determinado intervalo de comprimento de onda melhor será a resolução espectral da coleta. Por exemplo, o satélite Landsat 5 possui os sensores TM (*Thematic Mapper*) e MSS (*Multispectral Scanning System*). O sensor TM apresenta algumas bandas (faixas) espectrais mais estreitas do que o sensor MSS, portanto, nestas bandas o TM apresenta melhor resolução espectral do que o MSS.
- **Resolução radiométrica:** define a eficiência do sistema sensor em detectar pequenos sinais, ou seja, refere-se à maior ou menor capacidade do sistema sensor em detectar e registrar diferenças na energia refletida e/ou emitida pelos elementos que compõe a cena, tais como rochas, solos, vegetações e água. Por exemplo, o sistema sensor TM do Landsat 5 distingue até 256 tons distintos de sinais representando-os em 256 níveis de cinza.
- **Resolução temporal:** está relacionada com a repetitividade com que o sistema sensor pode adquirir informações referentes ao objeto. Por exemplo, o sensor CCD do satélite CBERS possui uma resolução temporal de 26 dias, ou seja, a cada 26 dias o satélite CBERS passa sobre um mesmo ponto geográfico da Terra.

Os sistemas sensores podem ser classificados ainda quanto à fonte de radiação. São chamados de *ativos* aqueles sistemas sensores que produzem sua própria radiação, e de *passivos*, aqueles sensores que dependem de uma fonte de radiação externa para que possam operar.

### 2.2.4 Níveis de aquisição de dados

A altitude do sensor em relação ao alvo define o que se convencionou chamar em sensoriamento remoto de *nível de aquisição de dados*.

De acordo com Novo (1989), existem basicamente três níveis de aquisição de dados em sensoriamento remoto:

- Nível de *laboratório/campo*: trabalha-se com porções reduzidas da matéria onde seu comportamento espectral é estudado quase sem interferência de fatores ambientais. Neste nível, os sistemas sensores podem ser instalados em mastros, colocados em barcos, fixados em bóias ou dentro de laboratórios.
- Nível *suborbital*: a energia registrada não se refere a um determinado objeto, mas a um arranjo de objetos da cena (alguns objetos podem ser individualmente detectados por sua configuração). Neste caso, geralmente utilizam-se aeronaves como plataformas de coletas de dados.
- Nível *orbital*: em cada elemento de resolução no terreno, a energia registrada é a integração da resposta de diferentes objetos. Neste nível, empregam-se satélites não tripulados e balões.

No presente trabalho são analisados apenas dados adquiridos por sensores imageadores instalados a bordo de satélites artificiais, ou seja, dados adquiridos em nível orbital.

### 2.2.5 Satélites

Embora outros sistemas façam parte do sensoriamento remoto, como os radiômetros de campo e de laboratório, sensores fotográficos e outros sistemas que operam em aeronaves, são os satélites que, cada vez mais, se tornam os instrumentos cotidianos dos profissionais de sensoriamento remoto. Por isso, pode-se afirmar que há uma estreita associação entre sensoriamento remoto e satélites artificiais.

Os satélites são veículos colocados em órbita da Terra e que promovem continuamente a aquisição de dados relacionados às capacidades de reflexão e emissão de radiação por parte dos

objetos. Por estarem a grandes altitudes, tipicamente entre 600 e 1.000 km, os satélites têm a capacidade de abranger em seu campo de visada uma grande porção de superfície terrestre. Ao mesmo tempo, como orbitam ao redor da Terra, promovem uma cobertura que se repete ao longo do tempo, permitindo o acompanhamento da evolução das propriedades de reflexão ou emissão dos objetos e fenômenos (Epiphanyo, 2006).

A era espacial de satélites artificiais de observação da Terra teve início no ano de 1972, quando os americanos colocaram em órbita o satélite denominado Earth-1, renomeado de Landsat-1. Depois desse, vários outros satélites foram e têm sido colocados em órbita, por muitos países, inclusive o Brasil (Moreira, 2005).

A seguir são apresentadas as características de alguns dos principais satélites em operação, particularmente, dos mais utilizados no Brasil (incluindo aqueles a partir dos quais foram obtidas as imagens utilizadas nos experimentos realizados neste trabalho).

## LANDSAT

A série de satélites do programa LANDSAT foi iniciada no final da década de 60, a partir de um projeto desenvolvido pela Agência Espacial Americana dedicado exclusivamente à observação de recursos naturais da superfície terrestre. O primeiro satélite da série começou a operar em 1972 e a última atualização ocorreu em 1999 com o lançamento do LANDSAT-7.

Atualmente, o único satélite em operação é o LANDSAT-5, que leva a bordo o sensor TM, o qual contribuiu para o mapeamento temático da superfície terrestre. O LANDSAT-7 iniciou suas atividades em 1999 e encerrou em 2003, utilizando o sensor ETM+ (Enhanced Thematic Mapper Plus). Este instrumento foi capaz de ampliar as possibilidades de uso dos produtos LANDSAT, uma vez que manteve a alta resolução espectral e conseguiu ampliar a resolução espacial da banda 6 (Infravermelho Termal) para 60 metros, além de inserir a banda pancromática e permitir a geração de composições coloridas com 15 metros de resolução (EMBRAPA, 2004).

A antena de recepção do Instituto Nacional de Pesquisas Espaciais (INPE) localizada em Cuiabá, capta, desde os anos 70, imagens LANDSAT de todo o território nacional.

A tabela 2.1 mostra as principais características dos diferentes sensores remotos a bordo

dos satélites da série LANDSAT.

Sensor	Bandas espectrais	Resolução espectral	Resolução espacial	Resolução temporal	Faixa imageada
MSS	4	0,5 - 0,6 $\mu\text{m}$	80m	18 dias	185km
	5	0,6 - 0,7 $\mu\text{m}$			
	6	0,7 - 0,8 $\mu\text{m}$			
	7	0,8 - 1,1 $\mu\text{m}$			
	8	10,4 - 12,6 $\mu\text{m}$	120m		
TM	1	0,45 - 0,52 $\mu\text{m}$	30m	16 dias	185km
	2	0,50 - 0,60 $\mu\text{m}$			
	3	0,63 - 0,69 $\mu\text{m}$			
	4	0,76 - 0,90 $\mu\text{m}$			
	5	1,55 - 1,75 $\mu\text{m}$			
	6	10,4 - 12,5 $\mu\text{m}$	120m		
	7	2,08 - 2,35 $\mu\text{m}$	30m		
ETM+	1	0,45 - 0,52 $\mu\text{m}$	30m	16 dias	185km
	2	0,50 - 0,60 $\mu\text{m}$			
	3	0,63 - 0,69 $\mu\text{m}$			
	4	0,76 - 0,90 $\mu\text{m}$			
	5	1,55 - 1,75 $\mu\text{m}$			
	6	10,4 - 12,5 $\mu\text{m}$	60m		
	7	2,08 - 2,35 $\mu\text{m}$	30m		
	8	0,50 - 0,90 $\mu\text{m}$	15m		

Tab. 2.1: Principais características dos sensores da série LANDSAT (EMBRAPA, 2004).

## SPOT

A série SPOT (Satellite pour l'Observation de la Terre) foi iniciada com o satélite franco-europeu SPOT 1 em 1986, sob a responsabilidade do Centre National d'Etudes Spatiales - CNES da França. Hoje a plataforma do SPOT está em órbita com três satélites (2, 4 e 5), o que possibilita acesso a qualquer ponto da Terra em menos de 24 horas, com revisitas em intervalos de 3 a 5 dias.

Os satélites da família SPOT operam com sensores ópticos, em bandas do visível, infravermelho próximo e infravermelho médio. Com o lançamento do SPOT 5, ocorrido em maio de 2002, a missão inaugurou a possibilidade de aquisição de imagens orbitais tridimensionais

graças a sua capacidade de visada lateral de até 27 graus (estereoscopia cilíndrica) e também conseguiu melhorar a resolução espacial do canal pancromático para 5 metros (Mather, 2004).

## NOAA

A família de satélites NOAA (National Oceanic Atmospheric Administration), administrada pelo National Environmental Satellite and Information Service (NESDIS), foi iniciada em 1960 com os satélites da série TIROS.

O objetivo principal do sensor AVHRR a bordo dos satélites da série NOAA era fornecer estimativas de densidade de nuvens e medidas de temperaturas da superfície do mar. Entretanto, em 1981, foi publicado o primeiro trabalho sobre a importância dos dados do AVHRR no uso do solo (Grays e McCrary, 1981). A partir desse trabalho, os dados do AVHRR passaram a ser utilizados também para fins ambientais.

A largura da faixa de imageamento do sensor AVHRR é de 2.700km e a sua resolução radiométrica é de 10bits (1024 níveis de cinza). As suas bandas espectrais compreendem as faixas do vermelho, infravermelho próximo, médio e termal (Moreira, 2005).

## CBERS

O programa CBERS nasceu de uma parceria entre Brasil e China no setor técnico-científico espacial, inserindo o Brasil no seleto grupo de países detentores da tecnologia de sensoriamento remoto orbital. O programa CBERS contemplou num primeiro momento apenas dois satélites de sensoriamento remoto, CBERS-1 e 2. O sucesso tanto do lançamento quanto do perfeito funcionamento do CBERS-1 e CBERS-2 estimulou os dois governos a decidirem expandir o acordo e incluir outros dois satélites da mesma categoria, os satélites CBERS-2B e os CBERS-3 e 4, como uma segunda etapa da parceria sino-brasileira (INPE, 2007).

A tabela 2.2 descreve as principais características dos sensores remotos a bordo dos satélites do programa CBERS.

Sensor	Bandas espectrais	Resolução espectral	Resolução espacial	Resolução temporal	Faixa imageada
CCD	Pan	0,51 - 0,73 $\mu\text{m}$	20m	26/3 dias	113km
	Azul	0,45 - 0,52 $\mu\text{m}$			
	Verde	0,52 - 0,59 $\mu\text{m}$			
	Vermelho	0,63 - 0,69 $\mu\text{m}$			
	Infravermelho próximo	0,77 - 0,89 $\mu\text{m}$			
IRMSS	Pan	0,50 - 1,10 $\mu\text{m}$	80m	26 dias	120km
	Infravermelho médio	1,55 - 1,75 $\mu\text{m}$			
	Infravermelho médio	2,08 - 2,35 $\mu\text{m}$	160m		
	Infravermelho termal	10,40 - 12,50 $\mu\text{m}$			
WFI	Vermelho	0,63 - 0,69 $\mu\text{m}$	260m	5 dias	890km
	Infravermelho próximo	0,77 - 0,89 $\mu\text{m}$			

Tab. 2.2: Principais características dos sensores dos satélites do programa CBERS (EMBRAPA, 2004).

## IKONOS

O satélite IKONOS II, operado pela SPACE IMAGING, foi lançado em 24 de setembro de 1999, sendo o primeiro satélite comercial de alta resolução com capacidade de imageamento de 1 metro. Seu sistema sensor capta imagens no modo pancromático e multiespectral. Adicionalmente, há a possibilidade de combinação de imagens preto e branco com dados multiespectrais para a geração de imagens coloridas, facilitando assim a interpretação visual e substituindo em grande parte dos casos, o uso de fotografias aéreas (EMBRAPA, 2004). Para se conseguir alta resolução espacial, as bandas espectrais dos sensores no visível são largas dentro do espectro da luz, o que permite maior penetração na atmosfera e maior poder de discriminação dos alvos terrestres, principalmente da cobertura vegetal, áreas sombreadas e de corpos d'água.

As principais características das bandas espectrais do IKONOS II são mostradas na tabela 2.3.

Mais informações sobre as características técnicas de outros satélites podem ser encontradas em EMBRAPA (2004) e em Moreira (2005).



Sensor	Bandas espectrais	Resolução espectral	Resolução espacial	Resolução temporal	Faixa imageada
Pancromático	Pan	0,45 - 0,90 $\mu\text{m}$	1m	2,9 dias	13km
Multiespectral	Azul	0,45 - 0,52 $\mu\text{m}$	4m	1,5 dia	
	Verde	0,52 - 0,60 $\mu\text{m}$			
	Vermelho	0,63 - 0,69 $\mu\text{m}$			
	Infravermelho próximo	0,76 - 0,90 $\mu\text{m}$			

Tab. 2.3: Principais características das bandas espectrais do satélite IKONOS II (EMBRAPA, 2004).

## 2.3 Características das imagens de sensoriamento remoto

As imagens de sensoriamento remoto, por sua natureza digital ou discreta, são constituídas por uma matriz bidimensional de elementos chamados pixels. Cada pixel da imagem tem sua localização definida em um sistema de coordenadas espaciais, do tipo linha e coluna, e possui um atributo numérico que indica o nível de cinza desse pixel, que varia do preto (nível mínimo) ao branco (nível máximo). O nível de cinza de um pixel corresponde sempre à média da intensidade da energia eletromagnética refletida ou emitida pelos diferentes materiais presentes em uma área da superfície terrestre. Para um mesmo sensor remoto, cada pixel representa sempre uma área projetada com as mesmas dimensões (Crósta, 1992).

Para entender os fundamentos do processamento digital das imagens de sensoriamento remoto é necessário compreender que alguns parâmetros que caracterizam essas imagens estão diretamente relacionados com as características técnicas dos sensores de imageamento, principalmente no que se refere à resolução dos mesmos. Conforme descrito anteriormente na seção 2.2.3, os sistemas sensores apresentam quatro domínios de resolução. Particularmente, os domínios de resolução espacial, espectral e radiométrico são abordados novamente aqui, porém no contexto das imagens.

A resolução **espacial** das imagens corresponde ao tamanho da área da superfície terrestre que é representada pelos seus pixels. Quanto maior a área projetada pelo pixel, maior é a quantidade de materiais da superfície terrestre representados pelo mesmo e, portanto, menor é a resolução espacial da imagem. Por outro lado, quanto menor a área da superfície terrestre

coberta pelo pixel, menor é a quantidade de materiais representados por ele, logo, maior é a resolução espacial da imagem. Isso significa que a resolução espacial das imagens de satélite está relacionada com o nível de detalhe com que podemos observar os objetos da superfície terrestre. A figura 2.6 (a) mostra uma imagem da cidade de Brasília-DF adquirida pelo satélite IKONOS-II com resolução espacial de 1 m. Através dela pode-se observar nitidamente feições locais como o traçado das ruas e até mesmo árvores e casas. Já através da imagem da figura 2.6 (b), obtida pelo satélite CBERS com resolução espacial de 20 m, que também mostra a cidade de Brasília, pode-se observar apenas feições regionais como, por exemplo, a distribuição do Plano Piloto e do Lago Paranoá.

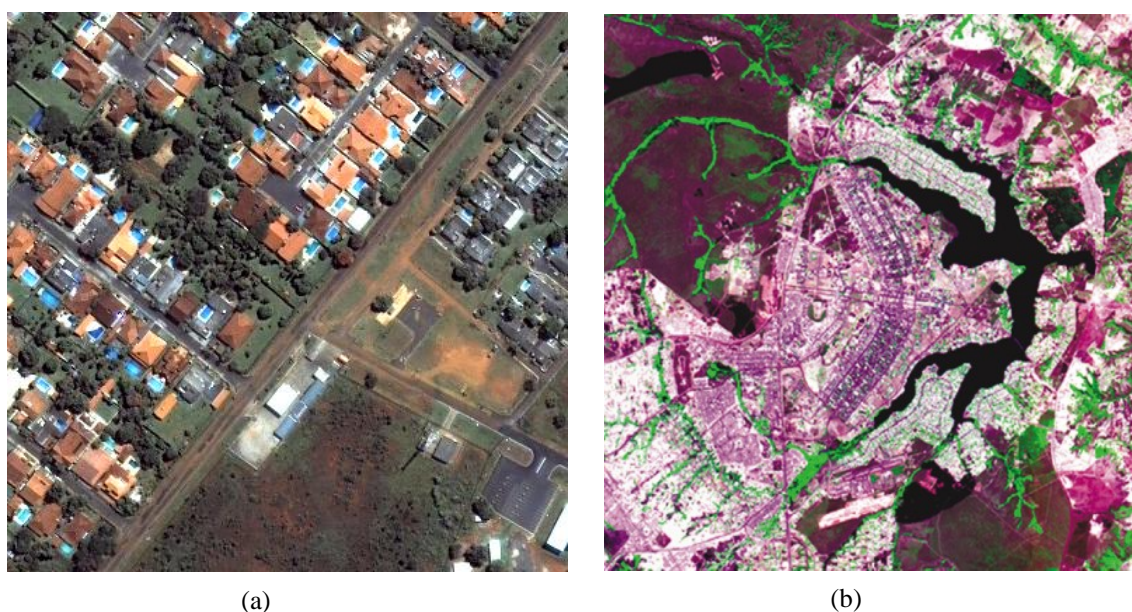


Fig. 2.6: *Imagens com diferentes resoluções espaciais: (a) Imagem do satélite IKONOS-II com resolução espacial de 1m. (b) Imagem do satélite CBERS com resolução espacial de 20m (Fonte: INPE (2007)).*

A tabela 2.4 mostra uma classificação da resolução espacial das imagens de sensoriamento remoto em três categorias: alta, média e baixa, apresentando também as principais aplicações e os principais satélites associados a cada uma delas.

A resolução **espectral** é um conceito inerente às imagens *multiespectrais* de sensoriamento remoto. Ela é definida pelo número de bandas espectrais de um sistema sensor e pela largura do intervalo de comprimento de onda coberto por cada banda. A característica multiespectral

Resolução espacial		Características e aplicações	Principais Satélites
Baixa	50 - 1000 m	- Cartografia de estrutura geológicas regionais; - Cartografia de grandes bacias hidrográficas e e extensas áreas florestais e agrícolas;	NOAA ERS
Média	10 - 15 m	- Identificação dos quarteirões de uma cidade, edifícios e ruas principais; - Detalhamento de áreas florestais; - Identificação de minerações e áreas agrícolas;	CBERS LANDSAT SPOT ASTER RADARSAT
	20 - 30 m	- Identificação de regiões urbanas, aeroportos, rodovias principais e ferrovias; - Identificação de grandes áreas florestais e agrícolas, bacias hidrográficas e caracterização da cobertura do solo; - Identificação de lineamentos geológicos;	
Alta	0,7 - 5 m	- Permite diferenciar tipos de edificações; - Geração de cartografia à escala humana (discrimina árvores, carros, ônibus, monumentos e arruamento intra-urbano); - Identificação de áreas agrícolas menores;	IKONOS Quickbird Eros

Tab. 2.4: Classificação da resolução espacial das imagens de sensoriamento remoto.

que geralmente diferencia algumas imagens de sensoriamento remoto de outras imagens digitais refere-se, portanto, às imagens de uma mesma cena, adquiridas em diferentes bandas espectrais, definidas por intervalos de comprimento de onda específicos. Isso viabiliza a identificação de diferentes padrões de cobertura terrestre, através de uma análise multiespectral realizada com imagens de diferentes regiões do espectro. Dependendo das bandas utilizadas, os alvos da superfície terrestre serão representados por um ou outro nível de cinza nas imagens. A figura 2.7 mostra, por exemplo, imagens adquiridas através das bandas espectrais 3, 4 e 5 do sensor TM do satélite LANDSAT, onde pode-se observar diferenças nas intensidades dos níveis de cinza nas três imagens. A cena mostra a cidade de Brasília-DF.

Embora a característica multiespectral das imagens de sensoriamento remoto seja vantajosa para a discriminação dos padrões de cobertura terrestre, é importante ressaltar que alguns sensores de imageamento apresentam alta correlação entre as informações contidas em algumas de suas bandas espectrais. Isso acontece, por exemplo, nas bandas 4 e 5 (região do visível) e nas bandas 6 e 7 (infravermelho próximo) do sensor MSS do satélite LANDSAT. Portanto,

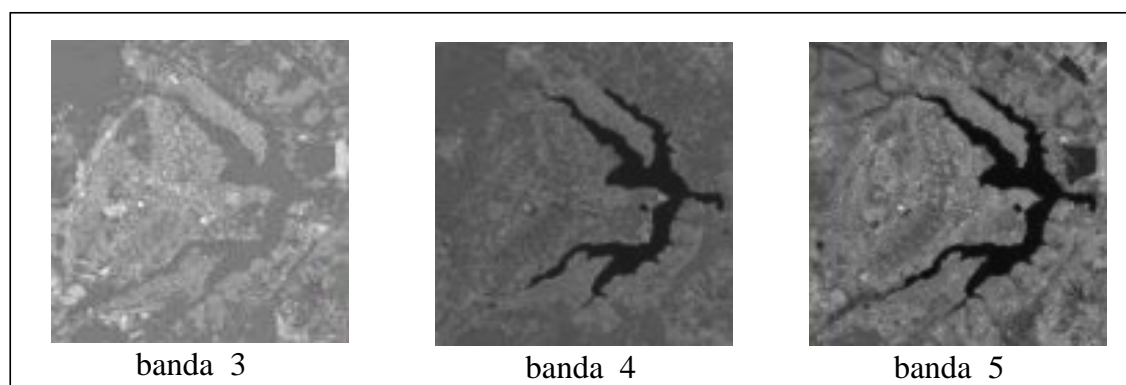


Fig. 2.7: *Imagens das bandas espectrais 3, 4 e 5 do sensor TM do satélite LANDSAT.*

em casos como esses, em que a imagem de uma banda espectral pode ser inferida a partir de outra (dada a alta correlação existente entre elas), fica evidente que o usuário ou o analista de imagens pode utilizar um número menor de bandas espectrais ou aplicar alguma técnica de processamento para reduzir a redundância de informações (Moreira, 2005).

No que se refere à resolução espectral, vale observar ainda que recentes avanços na tecnologia de sensoriamento remoto vêm disponibilizando para a comunidade científica internacional uma nova geração de sensores, capazes de fornecer imagens com uma resolução espectral muito alta. O Hyperion (*Hyperspectral Imager*) é um exemplo desse tipo de sensor. Colocado a bordo do satélite EO1, ele obtém dados com amplitude de 10 nm, em 220 bandas espectrais. Essa tecnologia proporciona a obtenção de imagens em um grande número de bandas estreitas, amostrando o espectro eletromagnético de maneira praticamente contínua. A alta resolução espectral torna a informação de cada elemento da imagem (pixel) próxima à obtida por meio de medições realizadas em nível de laboratório e/ou campo. Essas imagens, adquiridas por sensores com altíssima resolução espectral, são chamadas de *hiperespectrais*. As suas principais aplicações são o mapeamento de minerais, a identificação de feições espectrais de manchas de solos, de afloramentos, do comportamento biofísico da vegetação, além de permitirem a identificação dos principais gases de efeito estufa,  $H_2O$  e  $CO_2$  (Carvalho Jr. *et al.*, 2002; Rodrigues *et al.*, 2007).

A resolução **radiométrica** de uma imagem de sensoriamento remoto é dada pelo número de níveis digitais, representando níveis de cinza usados para expressar os dados coletados pelo

sensor. Quanto maior o número de níveis, maior é a resolução radiométrica. O número de níveis é normalmente expresso em termos do número de dígitos binários (*bits*) necessários para armazenar, em forma digital, o valor de brilho máximo dos pixels da imagem. A figura 2.8 mostra duas imagens de uma mesma cena com resoluções radiométricas diferentes (uma com dois níveis de cinza e outra com quatro).

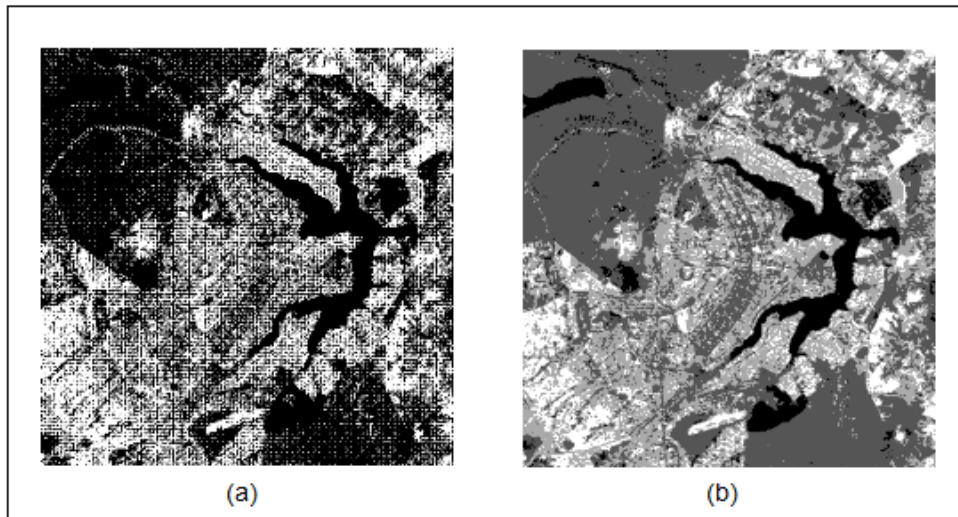


Fig. 2.8: *Simulação de uma imagem do satélite CBERS: (a) com dois níveis de cinza; (b) com quatro níveis de cinza.*

Os sistemas sensores mais comuns, dos satélites LANDSAT e CBERS, possuem resolução radiométrica de 8 bits, o que compreende 256 níveis de cinza. O sistema visual humano não é muito sensível às variações em intensidade, de tal modo que dificilmente são percebidas mais do que 30 diferentes tons de cinza numa imagem; o computador, por sua vez, consegue diferenciar qualquer quantidade de níveis, razão pela qual torna-se importante ter imagens de alta resolução radiométrica (Crósta, 1992).

Diante do contexto de caracterização das imagens de sensores remotos apresentado nesta seção, vale destacar que os métodos de classificação de imagens propostos neste trabalho são indicados para imagens multiespectrais de sensoriamento remoto de média e alta resoluções espaciais, contendo quaisquer tipos de padrões de cobertura terrestre, exceto padrões de área urbana para o caso de imagens de alta resolução espacial.

## 2.4 Operações do processamento digital de imagens de sensoriamento remoto

De acordo com Lillesand e Kiefer (2000), as possíveis formas de processamento de imagens digitais de sensoriamento remoto são inúmeras. Entretanto, elas podem ser virtualmente categorizadas dentro de quatro tipos de operações básicas: pré-processamento, realce, transformações e classificação. Uma breve descrição de cada um desses tipos de operações é dada a seguir.

### Pré-processamento

As operações de pré-processamento têm como objetivo corrigir imagens distorcidas e degradadas visando uma representação mais fiel da cena original. Elas tipicamente envolvem o processamento inicial das imagens brutas procurando corrigir distorções geométricas, calibrar radiometricamente os dados e eliminar ruídos. Por isso, precedem qualquer tipo de manipulação adicional da imagem para extrair alguma informação específica. As principais operações de pré-processamento são descritas como seguem (Novo, 1989):

- *Correção geométrica*: procura eliminar erros que geralmente ocorrem nas imagens em decorrência de fatores como: a variação da altitude do satélite, variação da velocidade da plataforma, deslocamentos sofridos pelo sistema de eixos do sensor e movimento de rotação da Terra. A função das técnicas de correção geométrica é reorganizar os pixels da imagem em relação a determinado sistema de projeção cartográfica.
- *Correção radiométrica*: tem como função minimizar diferenças entre os níveis de cinza registrados por uma matriz de detectores. As principais correções radiométricas são o "stripping" (padrão sucessivo de linhas horizontais que aparecem na imagem devido, por exemplo, à diferença ou desajuste de calibração dos detectores) e o "dropedlines" ou linhas com ausência de informação (padrão horizontal anômalo na imagem que ocorre pela perda de informações quando da gravação ou transmissão defeituosa ou ainda no processamento posterior dos dados).

- *Correção atmosférica*: consiste em reduzir o efeito da interferência da atmosfera sobre os valores de nível de cinza registrados em uma dada cena. Conforme mencionado na seção 2.2.2, a atmosfera afeta o sinal registrado pelos detectores a bordo dos satélites. Os métodos mais sofisticados de correção atmosférica baseiam-se na teoria de transferência radiativa e usam programas para o cálculo de transmitância atmosférica.

## Realce de imagens

Procedimentos de realce basicamente visam melhorar a qualidade visual das imagens. O objetivo é criar uma “nova” imagem a partir da imagem original, aumentando a quantidade de informação que pode ser visualmente interpretada. Não existem regras simples para se produzir a “melhor” imagem para uma aplicação particular. Frequentemente, diversos realces são feitos sobre a mesma imagem original para que se possa alcançar uma visualização satisfatória para a aplicação. As principais técnicas de realce de imagens são resumidas conforme seguem (Brito e Rosa, 2003; Mather, 2004):

- *Ampliação de contraste*: geralmente os níveis de cinza originais de uma cena, obtidos por um sistema sensor qualquer, não ocupam todo o intervalo possível. Esta técnica de realce de contraste (*contrast stretch*), através de uma transformação matemática, amplia o intervalo original para toda a escala de níveis de cinza ou números digitais disponíveis. Embora a transformação mais comum seja a linear, pode-se implementar qualquer outro tipo de transformação, dependendo do histograma original e do alvo ou feição de interesse.
- *Composição colorida*: a utilização de composições coloridas é fundamentada pelo fato de que o olho humano é capaz de discriminar mais facilmente matiz de cores do que tons de cinza. Para cada banda, associa-se uma cor primária (azul, verde ou vermelha) ou, ainda, as suas complementares (amarela, magenta ou cian), de modo que para cada alvo diferente da cena associa-se uma cor ou uma combinação de cores diferentes. Uma das únicas restrições a este método é que ele permite a utilização simultânea de no máximo três bandas. A reconstituição das cores na imagem advém do processo aditivo de formação das cores primárias (azul, verde e vermelho). A imagem resultante é costumeiramente

denominada de *composição colorida*.

- *Divisão de bandas*: esta técnica consiste na divisão do valor digital dos pixels de uma banda pelos correspondentes valores de uma outra banda. Ao se efetuar uma razão entre bandas, os quocientes variam em um intervalo que compreende valores reais contínuos. Esses valores devem atribuir à imagem resultante uma maior variância possível dos níveis de cinza (números digitais) sem saturá-la, e a média deve estar próxima da média do intervalo máximo dos valores digitais da imagem. Essa técnica possui a vantagem de atenuar os efeitos multiplicativos relativos à topografia, ângulo de elevação e azimute solar, além de enfatizar a separação dos alvos com comportamento de gradiente diferente nas curvas de reflectância.
- *Filtragem espacial*: busca enfatizar feições de interesse do usuário, enquanto suprime outras indesejáveis, como por exemplo, o ruído. Os dois principais tipos de filtros são o passa-alta e passa-baixa. Denomina-se passa-alta quando ocorre diminuição dos componentes de baixa frequência e aumento dos de alta frequência, ocorrendo um realce das bordas e detalhes da imagem. Os filtros do tipo passa-baixa tendem a aumentar os componentes de baixa frequência e diminuir os de alta frequência, ocorrendo perda de detalhes e redução do contraste da imagem, porém atenua a influência de processos ruidosos provocados, por exemplo, por defeitos do sensor e erros na transformação matemática.

### Transformações

Conforme descrito em Moreira (2005), para muitas das aplicações de imagens de sensoriamento, o uso de dados espectrais transformados é, às vezes, mais conveniente do que trabalhar com os originais. Dentre os métodos existentes para realizar a transformação de dados de imagens de satélite, um dos mais utilizados é o de *rotação espectral*.

O objetivo das técnicas de rotação espectral é realizar transformações lineares nas imagens, envolvendo processamentos que combinam as informações espectrais coletadas nas várias bandas do sensor, para gerar novas imagens contendo informações espectrais mais contrastadas do que nas bandas originais. Essas transformações são realizadas pixel a pixel e não dependem da



localização espacial dos dados. Para realizar a rotação espectral utiliza-se geralmente a técnica de *Análise de Componentes Principais*.

O uso da técnica de Análise de Componentes Principais visa reduzir a dimensionalidade dos dados, isto é, condensar as informações espectrais dos alvos, contidas em várias bandas do espectro eletromagnético, num número reduzido de bandas transformadas, sem perda de informações relevantes. Este procedimento é interessante para reduzir informações redundantes decorrentes de bandas espectrais altamente correlacionadas (conforme mencionado na seção 2.3), e também para sistemas de classificação, que geralmente admitem, para a análise espectral de um alvo, um número máximo de três ou quatro bandas espectrais.

## Classificação

Os procedimentos de classificação visam substituir a análise ou a interpretação visual da imagem por técnicas quantitativas que identificam automaticamente os padrões de cobertura terrestre existente em uma cena. As técnicas de classificação normalmente envolvem a análise de imagens multiespectrais e a aplicação de regras de decisão para categorizar todos os pixels da imagem dentro de uma das várias classes de cobertura terrestre. Quando esse tipo de operação é efetuado para todos os pixels de uma imagem, o resultado é um *mapa temático*, mostrando a distribuição geográfica de um tema, tal como vegetação, solo ou água. Pode-se dizer então que uma imagem de sensoriamento remoto classificada é uma forma de mapa digital temático.

Considerando que, dentre as diferentes operações do processamento digital de imagens de sensoriamento remoto, a classificação de imagens é o foco de estudo principal deste trabalho, nas próximas seções são apresentados maiores detalhes sobre essa técnica de processamento.

## 2.5 Técnicas de classificação

As técnicas de classificação de imagens de sensoriamento remoto normalmente envolvem a análise de imagens multiespectrais. Segundo Richards (1995), a melhor maneira de representar os dados dessas imagens para compreender e formular métodos de classificação, é fixá-los em um espaço de padrões ou espaço multiespectral com dimensão igual ao número de bandas espectrais

da imagem. Nesse espaço, um pixel da imagem é fixado como um ponto cujas coordenadas são dadas pelo seu nível de cinza em cada banda espectral.

A figura 2.9 ilustra um espaço multiespectral composto apenas por duas bandas espectrais de uma imagem. Supondo que as bandas tenham sido projetadas de forma a proverem uma boa discriminação, é esperado que haja a formação de agrupamentos de pixels no espaço que sejam correspondentes aos tipos de cobertura terrestre presentes na imagem. Esses agrupamentos podem apresentar diferentes tamanhos e formas, uma vez que qualquer material da superfície terrestre sempre apresenta variações nas intensidades dos pixels que os compõe, devido a fatores como mistura de outros materiais, diferentes condições de iluminação, presença de ruído e outros. Os agrupamentos de pixels formados no espaço multiespectral são chamados de **classes de informação**, uma vez que eles são as classes reais de cobertura terrestre presentes na imagem.

Na prática, os agrupamentos das classes de informação podem não ser simplesmente grupos individuais como mostra a figura 2.9. Em vez disso, não é raro encontrar no espaço multiespectral diversos grupos para a mesma superfície de solo, de água, de vegetação ou de qualquer outro tipo de cobertura terrestre. Esses grupos não são apenas resultado de diferenças específicas nos tipos de cobertura, mas também resultado de diferenças no conteúdo de umidade, tipos de solos sob a vegetação e influências topográficas. Conseqüentemente, é mais provável que um espaço multiespectral se assemelhe com o que é mostrado na figura 2.10, onde cada classe de informação é vista como sendo composta por várias subclasses chamadas de **classes espectrais**.

Nesse contexto, as técnicas de classificação de imagens multiespectrais de sensoriamento remoto têm como função principal efetuar a categorização de todos os pixels de uma imagem em uma de suas classes de informação, baseando-se no conjunto de classes espectrais existentes e, quando possível, em outras informações não-espectrais (como é mostrado adiante na seção 2.5.4).

Sob o ponto de vista de paradigmas de aprendizado, em sensoriamento remoto tradicional, as técnicas de classificação são divididas em três categorias principais: supervisionada, não-supervisionada e híbrida. As características, e principais vantagens e desvantagens, de cada

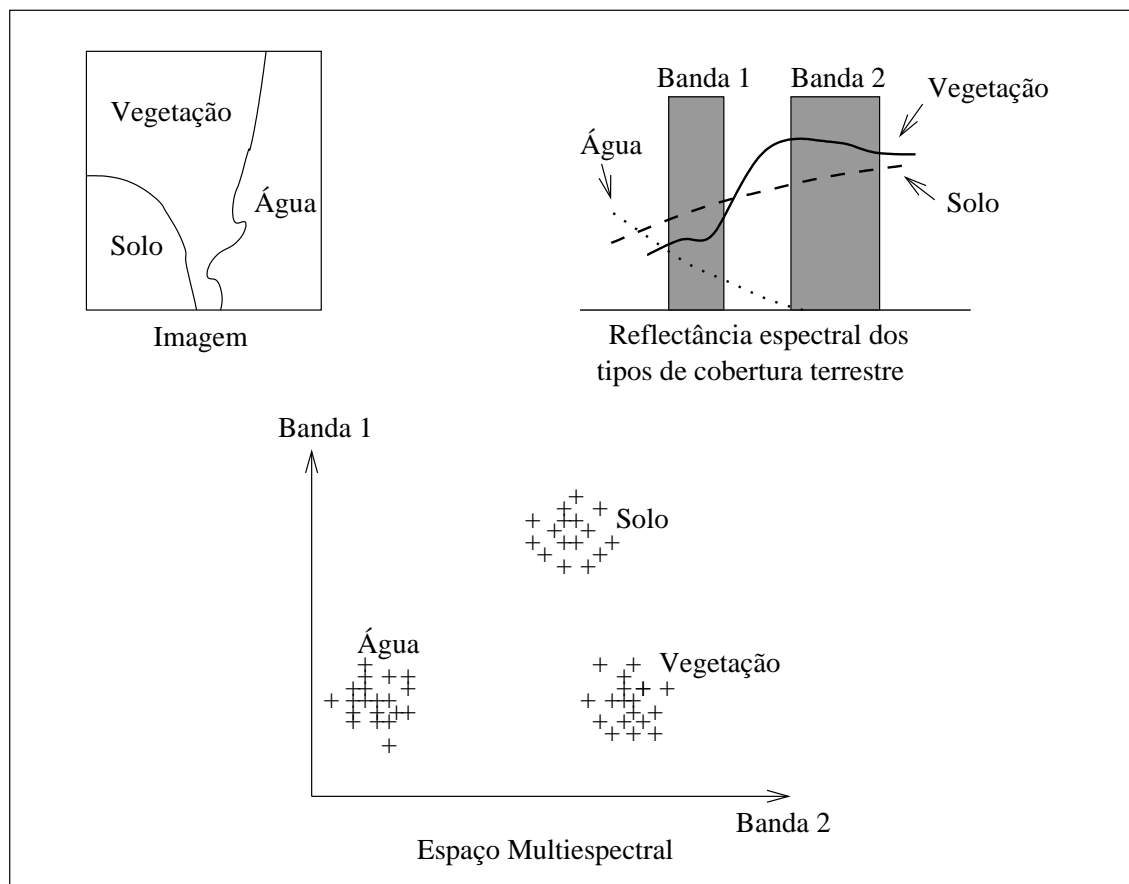


Fig. 2.9: Ilustração de um espaço multiespectral bidimensional mostrando sua relação com as reflectâncias espectrais dos tipos de cobertura terrestre.

uma dessas categorias de classificação são apresentadas nas próximas subseções.

### 2.5.1 Classificação supervisionada

Na classificação supervisionada de imagens de sensoriamento remoto um analista de imagens supervisiona o processo de categorização dos pixels especificando, para o algoritmo computacional, descritores numéricos das várias classes de cobertura terrestre presentes em uma cena. Para fazer isso, amostras representativas de classes conhecidas, chamadas de **áreas de treinamento**, são usadas para compilar uma “chave de interpretação” numérica que descreve os atributos (ou classes) espectrais de cada padrão de cobertura de interesse. Cada pixel da imagem é então comparado numericamente com cada categoria na chave de interpretação e rotulado com o nome daquela que ele mais se assemelha (Lillesand e Kiefer, 2000).

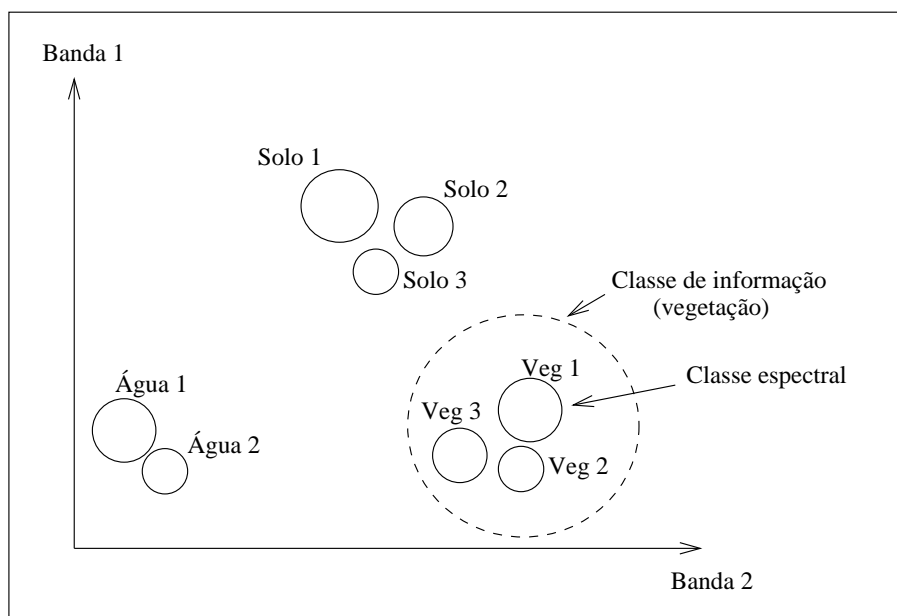


Fig. 2.10: *Representação de classes de informação por conjuntos de classes espectrais.*

Segundo Mather (2004), os algoritmos de classificação supervisionada podem ser divididos em dois tipos: **paramétricos** e **não-paramétricos**. Os algoritmos paramétricos assumem uma distribuição estatística ou uma função de probabilidade para cada classe de interesse, geralmente uma distribuição normal no espaço multiespectral, e então estimam os parâmetros dessa distribuição, tal como o vetor média e a matriz de covariância, para serem usados pelo algoritmo de classificação. Os algoritmos não-paramétricos, por outro lado, não fazem nenhuma suposição sobre a distribuição estatística das classes, ou seja, a forma das funções de probabilidade não são conhecidas previamente. As técnicas não-paramétricas são consideradas mais “robustas”, pois trabalham bem com imagens que possuem uma grande variedade de distribuições de classes (Gonçalves, 1997).

Os passos práticos essenciais de um processo de classificação supervisionada podem ser resumidos da seguinte forma:

1. Determinar o conjunto de tipos de cobertura terrestre sobre o qual a imagem será classificada, ou seja, determinar as classes de informação.
2. Escolher amostras representativas de pixels (áreas de treinamento) para cada uma das classes de informação. As amostras são selecionadas por um analista diretamente sobre

a imagem. Normalmente, várias amostras são selecionadas para uma mesma classe, para assegurar que os pixels a ela pertencentes sejam realmente representativos para a classe, ou seja, envolvam todos os seus atributos espectrais.

3. Utilizar as amostras das classes de informação para estimar os parâmetros do classificador particular (algoritmo) a ser utilizado.
4. Usando o classificador treinado, classificar todos os pixels da imagem dentro de uma das classes de informação desejadas.
5. Produzir mapas temáticos que apresentem sumários dos resultados da classificação.

As *vantagens* apresentadas pelas técnicas de classificação supervisionada se concentram basicamente nos seguintes pontos:

- Possibilidade do usuário controlar o conjunto de classes de informação para a área geográfica da imagem a ser classificada. Esta vantagem pode ser bastante significativa quando há a necessidade de comparar classificações de imagens de uma mesma área obtidas em datas diferentes, ou quando uma classificação deve ser compatível com aquelas de regiões vizinhas;
- O usuário não necessita associar as classes espectrais do mapa final com as categorias de informação de interesse, uma vez que esta tarefa é conduzida durante o processo de seleção de amostras de treinamento.

Por outro lado, segundo (Campbell, 2002), as *desvantagens* da classificação supervisionada são numerosas, dadas as seguintes razões:

- O usuário ou analista de imagens pode impor uma estrutura de classificação aos dados da imagem que não é compatível com a natureza de suas classes reais. As classes impostas pelo usuário podem não ser distintas ou bem definidas no espaço multidimensional, dificultando assim o processo de classificação da imagem;

- Os dados de treinamento são freqüentemente definidos primariamente com base nas classes de informação e, somente secundariamente, com base nas propriedades espectrais. Variações significativas nos atributos espectrais de uma classe, que podem ocorrer em virtude da natureza da própria classe ou de outros fatores como sombreamento topográfico e presença de outros materiais, podem, portanto, contribuir para uma seleção de amostras de treinamento pouco representativas para aquela classe de informação;
- De acordo com os passos da classificação supervisionada (descritos anteriormente), é necessário que o usuário tenha algum conhecimento sobre a área a ser classificada antes de iniciar o processo de classificação. Idealmente deveria haver locais específicos dentro da área que se deseja classificar, onde observações de campo tivessem sido efetuadas durante a passagem do satélite para adquirir a imagem. Esse tipo de observação é conhecida como **verdade terrestre**. No entanto, informações sobre verdade terrestre são difíceis de serem obtidas, especialmente quando a área de classificação é muito grande, complexa ou inacessível. Conseqüentemente, dependendo do grau de experiência do analista, o processo de seleção de amostras poderá conter erros e/ou imprecisões, comprometendo desse modo a qualidade da classificação resultante.
- A classificação supervisionada pode não ser capaz de reconhecer classes particulares não representadas pelos dados de treinamento, possivelmente porque essas classes não são de conhecimento do analista ou porque ocupam áreas muito pequenas na imagem.

Conforme relatado no capítulo 1, na literatura existem diferentes propostas de métodos para efetuar a classificação supervisionada de imagens de sensoriamento remoto, baseados em técnicas diversas, como árvores de decisão (Hansen *et al.*, 1996), lógica *fuzzy* (Wang, 1990), algoritmos genéticos (Pal *et al.*, 2001), máquinas de vetores suporte (Mitra *et al.*, 2004), etc.. No entanto, o classificador estatístico de Máxima Verossimilhanças e as redes neurais de *Perceptrons* multicamadas são os métodos supervisionados mais populares.

### 2.5.1.1 Classificador estatístico - Máxima Verossimilhança

Baseado em parâmetros estatísticos, o classificador de Máxima Verossimilhança parte do pressuposto de que as diferentes classes de informação da imagem são modeladas no espaço multiespectral por distribuições de probabilidade normal ou gaussiana multivariada.

A distribuição normal multivariada é especificada pelo seu vetor média e sua matriz de covariância. Conseqüentemente, se esses parâmetros são conhecidos para cada classe então é possível calcular o conjunto de probabilidades relativas de um padrão pertencer a uma das classes de interesse (Richards, 1995).

Assim, a partir dos parâmetros vetor média ( $\mathbf{m}_i$ ) e matriz de covariância ( $\Sigma_i$ ), estimados com base nos dados de treinamento, pode-se definir as funções discriminantes  $g_i(\mathbf{x})$ , para uma dada classe  $w_i$ , como sendo:

$$g_i(\mathbf{x}) = \ln(p(w_i)) - \frac{1}{2} \ln |\Sigma_i| - \frac{1}{2} (\mathbf{x} - \mathbf{m}_i)^T \Sigma_i^{-1} (\mathbf{x} - \mathbf{m}_i), \quad (2.1)$$

onde  $\mathbf{x}$  é o vetor de atributos do pixel em análise,  $p(w_i)$  é a probabilidade a priori para a classe  $w_i$ ,  $\mathbf{m}_i$  é o vetor média para a classe  $w_i$  e  $\Sigma_i$  é a matriz de covariância para a classe  $w_i$ .

Estimadas as funções discriminantes  $g_i(\mathbf{x})$  para todas as classes de interesse, determina-se então os valores  $g_i(\mathbf{x})$  para cada pixel  $\mathbf{x}$  da imagem, que são classificados de acordo com a seguinte regra de decisão:

$$x \in w_i, \quad \text{se } g_i(\mathbf{x}) > g_j(\mathbf{x}) \quad \text{para todo } i \neq j. \quad (2.2)$$

O método de classificação por máxima verossimilhança, assim como qualquer outro método paramétrico de classificação supervisionada, exige que as áreas de treinamento sejam bastante representativas para as classes de interesse. Caso isso não ocorra a imagem classificada poderá conter sérias imprecisões. Portanto, para que a classificação por máxima verossimilhança seja precisa o suficiente, é necessário um número razoavelmente elevado de pixels para cada conjunto de treinamento. Segundo Crósta (1992), esse número deve ser preferencialmente acima de uma centena, a fim de proporcionar uma base confiável para se obter estimativas razoáveis dos

elementos vetor média e matriz de covariância utilizados pelo método.

Mais detalhes sobre o classificador de Máxima Verossimilhança podem ser obtidos em Duda *et al.* (2000).

### 2.5.1.2 Classificador neural - *Perceptrons* multicamadas

As redes neurais de *Perceptrons* multicamadas, em geral, produzem mapeamentos de um espaço multidimensional para outro espaço multidimensional. O seu processo de aprendizagem realizado de maneira supervisionada consiste em ajustar os pesos sinápticos da rede de neurônios de forma que estes atinjam uma configuração que seja capaz de efetuar o mapeamento desejado.

Mais especificamente, um número finito  $p$  de amostras é fornecido à rede na forma de pares de vetores do tipo  $(\mathbf{x}^k, \mathbf{s}^k)$ , onde  $\mathbf{x}^k = [x_1^k, \dots, x_m^k]^t$  representa uma entrada e  $\mathbf{s}^k = [s_1^k, \dots, s_n^k]^t$  representa a saída desejada ( $k = 1, \dots, p$ ). No contexto de imagens de sensoriamento remoto, a entrada  $\mathbf{x}^k$  corresponde a um vetor de amostra de pixels e  $\mathbf{s}^k$  é o rótulo da classe de informação correspondente.

Para cada entrada  $\mathbf{x}^k$  a rede produz uma saída correspondente:

$$\mathbf{y}(\mathbf{W}, \mathbf{x}^k) = [y_1^k, \dots, y_n^k]^t,$$

onde  $\mathbf{y}(\mathbf{W}, \cdot)$  é a saída da rede calculada em função de sua matriz de pesos sinápticos  $\mathbf{W}$ .

Se o erro gerado pela saída da rede é definido como sendo o erro quadrático médio  $E$ , o processo de aprendizagem da rede MLP consiste, portanto, em encontrar a matriz de pesos  $\mathbf{W}$  tal que:

$$E = \min \frac{1}{2} \sum_{k=1}^p [\mathbf{y}(\mathbf{W}, \mathbf{x}^k) - \mathbf{s}^k]^t [\mathbf{y}(\mathbf{W}, \mathbf{x}^k) - \mathbf{s}^k]. \quad (2.3)$$

O procedimento mais popular para calcular os elementos de  $\mathbf{W}$  que minimizam o erro acima é o algoritmo *backpropagation*. Para ajustar os pesos da rede, o algoritmo *backpropagation* se baseia no método do Gradiente, bastante conhecido na teoria de otimização (Luenberger, 2003).

Dentre as vantagens das redes neurais de *Perceptrons* multicamadas sobre os classificadores estatísticos convencionais, como o classificador de Máxima Verossimilhança, está a não necessidade de um conhecimento a priori da distribuição de probabilidade dos dados, uma vez que as redes neurais têm a habilidade de aprender as propriedades de distribuição dos dados durante



a fase de treinamento, bem como a habilidade de generalizar e incorporar informações e conhecimentos não-estatísticos que podem ser potencialmente valiosos (Gonçalves *et al.*, 1998a).

Mais detalhes sobre a aplicação das redes neurais de *Perceptrons* multicamadas (e do algoritmo backpropagation) na classificação de imagens de sensoriamento remoto podem ser obtidas em Gonçalves (1997).

### 2.5.2 Classificação não-supervisionada

A classificação não-supervisionada é a técnica pela qual os pixels de uma imagem são atribuídos às classes espectrais sem que haja qualquer tipo de conhecimento sobre a existência ou nome dessas classes (Richards, 1995). Esse tipo de classificação é freqüentemente realizada usando métodos de análise de agrupamentos ou de *clustering*. Esses métodos examinam os pixels desconhecidos de uma imagem e os incorporam dentro de um conjunto de classes definidas em função dos agrupamentos naturais (*clusters*) dos níveis de cinza dos pixels da imagem. A premissa básica é a de que os níveis de cinza de pixels pertencentes a uma mesma classe de cobertura terrestre estejam próximos no espaço de atributos enquanto os níveis de cinza de pixels pertencentes à classes diferentes estejam comparativamente bem separados (Lillesand e Kiefer, 2000). As classes que resultam da classificação não-supervisionada são as classes espectrais presentes na imagem.

As *vantagens* da classificação não-supervisionada podem ser resumidas como seguem (Campbell, 2002):

- Não exige conhecimento prévio sobre a área a ser classificada, ou seja, o usuário não necessita possuir um conhecimento detalhado sobre as classes da imagem ou realizar trabalhos de campo (ir até o terreno) para selecionar amostras de treinamento;
- Minimiza a influência e/ou a chance de erro por parte do usuário ou analista de imagens. Para conduzir a classificação não-supervisionada, em geral o operador especifica apenas o número de classes desejado, e em alguns casos, controla o grau de compactação e dispersão dos agrupamentos. Diferentemente da classificação supervisionada, onde muitas hipóteses ou decisões detalhadas são requeridas, na classificação não-supervisionada o analista tem

menos oportunidade de cometer erros. Se o analista possui pressuposições a respeito da região a ser classificada, ele terá pouca oportunidade de influenciar a classificação da imagem;

- Detecta classes raras ou com baixa probabilidade de ocorrência na imagem. Essas classes de difícil reconhecimento são agrupadas separadamente, não sendo incorporadas em outras classes (como acontece geralmente em classificações supervisionadas);
- As classes definidas pela classificação não-supervisionada são freqüentemente mais homogêneas do que aquelas geradas pela classificação supervisionada.

É importante observar que o fato da classificação não-supervisionada não exigir grande interação do usuário no processo de classificação da imagem, principalmente em relação ao uso de suposições ou conhecimentos prévios sobre área a ser classificada, ela se coloca como uma ferramenta imprescindível para a exploração em grande escala do enorme volume de imagens de sensoriamento remoto que estão sendo adquiridas e armazenadas atualmente.

As desvantagens da classificação não-supervisionada se concentram apenas nos dois seguintes aspectos:

- Em geral, no final da classificação, as classes espectrais detectadas pelo classificador não-supervisionado devem ser relacionadas às classes de informação de interesse. Nem sempre há uma correspondência direta entre as classes detectadas pela classificação não-supervisionada com aquelas de interesse do usuário;
- Uma vez que as propriedades espectrais de classes de informações específicas podem mudar com o tempo, os relacionamentos entre as classes espectrais e de informação não são constantes e, portanto, os relacionamentos estabelecidos para uma imagem nem sempre podem ser estendidos para outras imagens adquiridas em datas diferentes.

Conforme já descrito no capítulo 1, embora existam vários métodos de análise de agrupamentos na área de reconhecimento de padrões, poucos deles são efetivamente possíveis de serem aplicados na classificação não-supervisionada de imagens de sensoriamento remoto. Na maior

parte dos casos, o tamanho das imagens e a dimensão dos seus espaços de atributos inibem a aplicação de um método particular de agrupamentos.

Em análise de agrupamentos de dados duas classes de métodos prevalecem: os métodos *hierárquicos* e os *particionais*.

Os métodos hierárquicos produzem uma sucessão de partições, cada qual correspondendo a um número diferente de agrupamentos. Esses métodos de agrupamentos não exigem que o usuário especifique a priori o número de classes e outros parâmetros adicionais, e ainda possibilitam a visualização do resultado da classificação por meio de um dendrograma, que mostra hierarquicamente o grau de semelhança entre os agrupamentos que são formados por fusões (ou divisões) em cada estágio sucessivo da análise. No entanto, os métodos hierárquicos apresentam algumas características que inviabilizam a sua aplicação na classificação de imagens de sensoriamento remoto: (a) requerem espaço de memória de ordem  $O(N^2)$ , onde  $N$  é o número de registros do conjunto de dados; (b) o dendrograma pode ser difícil de interpretar, principalmente quando o conjunto de dados é muito grande; (c) para determinar qual o nível de "corte" do dendrograma há a necessidade de se aplicar algum critério de decisão.

Diferentemente dos métodos hierárquicos, os métodos de agrupamentos particionais produzem uma partição dos dados em  $K$  agrupamentos, geralmente otimizando uma função objetivo. Existem várias formas de escolha dos critérios para efetuar a partição, assim como dos algoritmos de otimização que podem ser usados em conjunto. A principal vantagem desses métodos em relação aos métodos hierárquicos é a possibilidade de mudar a pertinência de objetos em relação a um agrupamento durante o processo de particionamento dos dados. Embora apresentem uma série de limitações (como será mostrado na próxima seção), a maioria dos sistemas computacionais voltados para o processamento digital de imagens de sensoriamento remoto executa a classificação não-supervisionada usando métodos de agrupamentos particionais, como o algoritmo  $K$ -médias, e uma variação dele, o ISODATA.

### 2.5.2.1 Algoritmo $K$ -médias

O algoritmo  $K$ -médias é um procedimento iterativo que busca minimizar a variância interna dos agrupamentos que são formados durante o processo de particionamento dos dados. O

algoritmo é composto por 4 passos básicos:

1. Inicializa-se  $K$  vetores  $\hat{\mathbf{m}}_i$  como o centro dos  $K$  agrupamentos de dados (o número  $K$  é fornecido pelo usuário e corresponde ao número desejado de classes para a imagem);
2. Calcula-se as distâncias de todos os vetores de dados  $\mathbf{m}_i$  a cada um dos  $K$  centros  $\hat{\mathbf{m}}_i$ . Cada vetor  $\mathbf{m}_i$  é associado àquele agrupamento que apresenta o centro  $\hat{\mathbf{m}}_i$  mais próximo;
3. Calcula-se os novos centros médios  $\mathbf{m}_i$  de cada agrupamento  $i$  produzido no passo 2.
4. Se a soma dos quadrados das distâncias ou erros (SD) é inferior a um valor  $\epsilon$  pré-definido, conforme mostra a equação abaixo, ou se o algoritmo atingiu um número máximo de iterações pré-estabelecido, então o procedimento é encerrado.

$$SD = \sum_{\forall \mathbf{m}_i} (\mathbf{m}_i - \hat{\mathbf{m}}_i)^2 < \epsilon \quad (2.4)$$

Nota-se através da equação 2.4, que o algoritmo K-médias busca minimizar a soma dos quadrados das distâncias em cada agrupamento, que equivale a minimizar o *erro quadrático médio* (EQM), que por sua vez, pode ser considerado como uma medida de variabilidade interna dos agrupamentos:

$$EQM = \frac{\sum_{\forall \mathbf{m}_i} (\mathbf{m}_i - \hat{\mathbf{m}}_i)^2}{(N - K).b} = \frac{SD}{(N - K).b} \quad (2.5)$$

onde  $N$  é o número de vetores de dados  $\mathbf{m}_i$  e  $b$  é o número de bandas espectrais da imagem.

A inicialização dos centros ( $\hat{\mathbf{m}}_i$ ) dos agrupamentos pode ser feita aleatoriamente ou adotando-se alguma concepção lógica, como iniciar agrupamentos espaçados equitativamente sobre o espaço multidimensional, ou ainda em regiões de correspondência espectral a possíveis classes de cobertura terrestre presentes na imagem.

Embora seja comumente utilizado na classificação não-supervisionada de imagens de sensoriamento remoto, o algoritmo K-médias apresenta limitações importantes:

- Sua função objetivo parte do pressuposto de que o número de classes,  $K$ , é conhecido previamente. Na hipótese de se ter escolhido um valor  $K'$  inadequado, o método irá impor, pelo uso de técnicas de otimização,  $K'$  agrupamentos aos dados;

- O algoritmo K-médias é muito sensível à escolha dos parâmetros iniciais, gerar diferentes classificações quando várias simulações são realizadas para um mesmo conjunto de dados. Devido a isso, seus parâmetros são muitas vezes determinados experimentalmente por tentativa e erro, procurando evitar que o algoritmo caia em mínimos locais. Esse processo certamente aumenta o nível de interação do usuário com o algoritmo e, conseqüentemente, aumenta o grau de subjetividade do processo de categorização da imagem;
- Alto custo computacional quando o conjunto de dados a ser analisado é muito grande (a cada iteração todos os pixels da imagem são comparados com todos os centros de agrupamentos);
- Sob o ponto de vista estatístico, supondo que cada agrupamento de dados produzido pelo algoritmo K-médias segue uma distribuição normal (esférica) com diferentes médias (mas variâncias iguais), a análise de agrupamentos realizada pelo algoritmo pode ser interpretada como uma estimativa de máxima verossimilhança para as médias (centros) dos grupos de dados. Isso acarreta uma outra desvantagem ao algoritmo, que é a de ser adequado somente para analisar conjuntos de dados que apresentem agrupamentos com formatos hiperesféricos e que tenham a mesma variância.

Avanços sobre o algoritmo K-médias e esforços para diminuir algumas de suas principais desvantagens podem ser encontrados na literatura. Su e Chou (2001) propuseram uma versão modificada do algoritmo K-médias que adota uma medida de distância não-métrica baseada na idéia de "simetria de pontos". O algoritmo proposto pode ser usado para agrupar conjuntos de dados com diferentes estruturas geométricas. Porém, o custo da flexibilidade em detectar agrupamentos diversos é o aumento da complexidade computacional do algoritmo. Likas *et al.* (2003) introduziram um algoritmo K-médias global consistindo de uma série de procedimentos usando o próprio K-médias, mas com o número de agrupamentos variando de 1 até  $K$ . Os autores declaram que o algoritmo é mais robusto em relação as suas condições iniciais e também propuseram estratégias para acelerar a sua convergência. Entretanto, o problema da complexidade computacional é evidente, uma vez que há a necessidade de executar o algoritmo K-médias diversas vezes para cada valor de  $K$ . Outros estudos recentes sobre o K-médias podem

ser encontrados em Khan e Ahmad (2004); Huang *et al.* (2005); Chung e Lin (2007).

### 2.5.3 Classificação híbrida

As classificações híbridas são aquelas em que o analista utiliza primeiramente um método de classificação não-supervisionada para identificar as classes espectrais presentes na imagem, bem como, para selecionar as amostras de treinamento para as classes de interesse e, posteriormente, utiliza um método supervisionado para realizar a classificação final da cena.

Os classificadores híbridos procuram incorporar as vantagens apresentadas por ambas as técnicas de classificação supervisionada e não-supervisionada, para tentar minimizar os problemas que surgem quando se utiliza apenas uma dessas duas técnicas (Richards, 1995; Gonçalves, 1997).

### 2.5.4 Incorporação de informações não-espectrais

No processo de classificação de imagens de sensoriamento remoto, diversos fatores podem contribuir para um decréscimo na acurácia da imagem temática produzida pelo classificador. De acordo com Lersch *et al.* (2007), entre aqueles que mais freqüentemente ocorrem, pode-se mencionar a presença de classes espectralmente muito semelhantes entre si, a não normalidade nas distribuições das classes individuais e a não uniformidade na irradiância solar ao longo da cena (causada por influências topográficas).

Uma estratégia que pode ser empregada para tentar diminuir a influência negativa destes fatores é a utilização de informações auxiliares (não-espectrais) no processo de classificação.

Várias abordagens podem ser utilizadas para incorporar dados auxiliares no processo de classificação de imagens, de modo a melhorar os resultados (Hutchinson, 1982; Silva Filho *et al.*, 2003; Lersch *et al.*, 2007). Entre elas, destacam-se: o uso de informações externas e o uso de medidas de textura na vizinhança de um pixel.

#### 2.5.4.1 Informações externas

Segundo (Mather, 2004), o termo externo é usado para descrever quaisquer tipos de dados que não sejam aqueles da imagem original ou que não tenham sido gerados a partir da imagem original. Exemplos incluem informações sobre elevação do terreno (dados topográficos), informação sobre o tipo do solo ou o resultado de classificações de outras imagens da mesma área espacial.

As informações externas mais utilizadas são aquelas derivadas de modelos digitais de terreno: altitude, declive e exposição (Woodcock *et al.*, 1997). Contudo, também se tem utilizado outros tipos de dados externos, como fatores ambientais que afetam, por exemplo, a distribuição de vegetação, entre eles, o vento e a insolação (Lees e Ritman, 1991).

Embora a incorporação de informações externas para melhorar o desempenho de classificadores espectrais seja vantajosa, é importante observar que, dependendo da área a ser classificada, nem sempre essas informações são acessíveis ou fáceis de serem adquiridas. Além disso, muitas informações externas não são medidas em uma escala contínua, o que dificulta a sua inclusão na forma de vetores (ou dimensões) adicionais no processo de classificação das imagens. Tipos de solo e resultados de classificações prévias são exemplos clássicos, ambos são representados na forma de categorias.

#### 2.5.4.2 Textura

A textura é uma importante característica utilizada na interpretação visual de imagens e, por esta razão, pode melhorar o desempenho de classificadores digitais, principalmente em imagens de satélites com alta resolução espacial. Entretanto, a textura de uma imagem é muito difícil de ser quantificada, uma vez que não há um consenso na definição deste termo, muito menos uma formulação matemática precisa. Diferentemente da informação espectral, que descreve a variação do nível de cinza de um pixel, a textura contém informações sobre a distribuição espacial dos níveis de cinza de uma região da imagem (Rennó, 1995; Martins, 2004).

Apesar das dificuldades, muitos métodos de análise de textura têm sido propostos na literatura (Chen, 1990; Chaudhuri e Sarkar, 1995; Pichler *et al.*, 1998; E. Vansteenkiste *et al.*, 1995).

O método mais comumente utilizado, proposto por (Haralick *et al.*, 1973), baseia-se no uso de estatísticas derivadas a partir da matriz de co-ocorrência dos níveis de cinza de uma imagem.

A matriz de co-ocorrência dos níveis de cinza representa a distância e as relações espaciais angulares sobre uma subregião de uma imagem de tamanho especificado. Cada elemento da matriz é uma medida de probabilidade de ocorrência de valores de níveis de cinza separados por uma dada distância numa dada direção (Mather, 2004).

A principal desvantagem no uso de medidas texturais baseadas na matriz de co-ocorrência é a grande exigência em memória e tempo computacional. Por exemplo, numa imagem de 8 bits (256 níveis de cinza), a matriz de co-ocorrência é composta por 65536 elementos, muitos deles com valor nulo.

Em geral, as medidas de textura são incorporadas no processo de classificação das imagens como bandas adicionais, usadas em conjunto com as bandas espectrais da imagem original.

## 2.6 Mecanismos de validação da acuidade do mapeamento

Uma imagem classificada (ou mapa temático) deve sempre apresentar um grau de confiabilidade. Portanto, uma avaliação de sua exatidão estatística é indispensável. Existem vários métodos para avaliar a acuidade do mapeamento. Os métodos mais utilizados têm como ponto de partida a construção de uma *matriz de confusão* (ou matriz de erros).

A matriz de confusão descreve as exatidões individuais de cada classe do mapa, em conjunto com erros de inclusão e de omissão. O erro de inclusão ocorre quando uma área é abrangida numa classe à qual ela não pertence. Já o erro de omissão acontece quando uma área é excluída de uma classe à qual ela pertenceria (Congalton e Green, 1999).

Por meio da matriz de confusão é possível utilizar técnicas de análise multivariada para determinar a concordância entre o levantamento de campo ("verdade terrestre") e o mapa temático produzido pela classificação. O *índice de concordância Kappa* é um método bastante utilizado na análise de exatidão de imagens classificadas. Ele inclui no seu cálculo os elementos da matriz de confusão:



$$kappa = \frac{\theta_1 - \theta_2}{1 - \theta_2}, \quad (2.6)$$

onde:

$$\theta_1 = \frac{\sum_{i=1}^r x_{ii}}{N} \quad e \quad \theta_2 = \frac{\sum_{i=1}^r (x_{i+} \cdot x_{+i})}{N^2}, \quad (2.7)$$

sendo:  $\left\{ \begin{array}{l} x_{ii} \text{ o número de elementos da } ii\text{-ésima célula da matriz de confusão,} \\ x_{i+} \text{ a soma dos elementos da } i\text{-ésima linha da matriz de confusão,} \\ x_{+i} \text{ a soma dos elementos da } i\text{-ésima coluna da matriz de confusão, e} \\ N \text{ o número total de elementos da imagem.} \end{array} \right.$

Baseando-se nos valores do índice Kappa, a tabela 2.5 (desenvolvida por Landis e Koch (1977)) é comumente utilizada para categorizar a qualidade de resultados de classificações de imagens de sensoriamento remoto.

Valor Kappa	Qualidade da classificação
< 0,00	Péssima
0,00 - 0,20	Ruim
0,20 - 0,40	Razoável
0,40 - 0,60	Boa
0,60 - 0,80	Muito boa
0,80 - 1,00	Excelente

Tab. 2.5: Qualidade da classificação associada aos valores do índice de concordância Kappa (adaptada de Landis e Koch (1977)).

# Capítulo 3

## Mapa Auto-organizável de Kohonen

### 3.1 Introdução

O Mapa Auto-organizável (SOM - *Self-organizing Map*), proposto por Kohonen (1982, 1997), é um tipo de rede neural artificial que realiza um mapeamento de um espaço de entrada  $p$ -dimensional contínuo para um conjunto finito de neurônios organizados em um arranjo normalmente unidimensional ou bidimensional. Baseado em aprendizado competitivo e não-supervisionado, o SOM tem sido freqüentemente utilizado em inúmeras aplicações, incluindo áreas como reconhecimento de padrões, processamento de sinais, compressão de dados, modelagem biológica e mineração de dados. Propriedades importantes como a capacidade de aproximar o espaço de entrada, ordenação topológica e casamento de densidade, aliadas à simplicidade do modelo e à facilidade de implementação do seu algoritmo de aprendizagem, justificam o sucesso do SOM e o colocam como um dos principais modelos de redes neurais na atualidade.

Uma vez que os métodos de classificação de imagens de sensoriamento remoto propostos nesse trabalho baseiam-se essencialmente no uso do SOM, este capítulo tem como objetivo descrever as características e propriedades dessa rede neural, bem como apresentar seu algoritmo de aprendizagem e considerações sobre seus parâmetros de treinamento. Esse capítulo apresenta também uma revisão sobre métodos que segmentam, de maneira automática, os neurônios de um SOM treinado, descrevendo, em particular, o método de segmentação do SOM proposto

por Costa e Andrade Netto (1999, 2001), o qual serviu de base para a construção dos métodos de classificação de imagens propostos nos capítulos 4 e 5 dessa tese.

## 3.2 Motivação neurobiológica

O córtex cerebral corresponde à camada mais externa do cérebro dos vertebrados, sendo rico em neurônios e o local do processamento neuronal mais sofisticado e distinto. A espessura do córtex humano varia de 1 a 4 mm, contendo cerca de 10 bilhões de neurônios e 60 trilhões de sinapses que desempenham um papel central em funções complexas do cérebro, como a memória, atenção, consciência, linguagem, percepção e pensamento (Shepherd, 1988).

Em um nível maior de complexidade, o córtex cerebral apresenta circuitos locais (cerca de 1 mm de tamanho) constituídos por neurônios com propriedades similares ou distintas. Estes agrupamentos neurais realizam operações características de uma região localizada no cérebro. Eles são seguidos por circuitos inter-regionais constituídos por caminhos, colunas e mapas topográficos, que envolvem múltiplas regiões localizadas em partes diferentes do cérebro.

Os *mapas topográficos* são organizados para responder informações sensoriais incidentes. Diferentes entradas sensoriais (motora, somestésica, visual e auditiva, por exemplo) são mapeadas sobre áreas correspondentes do córtex cerebral de uma forma ordenada. Particularmente, a ordem física dos sinais de entrada percebidos pelos tecidos dos órgãos sensoriais é projetada no córtex cerebral em ordem semelhante, resultando num mapeamento que *preserva a ordem topológica* do sinal recebido, apesar de algumas transformações (Haykin, 1994). A figura 3.1 mostra uma representação das várias regiões corticais especializadas no cérebro humano.

A formação de mapeamentos topologicamente corretos no córtex cerebral é atribuída a diversos mecanismos, dentre os quais a *auto-organização* recebeu uma atenção especial pela comunidade acadêmica em razão de fortes evidências biológicas. Os estudos pioneiros sobre auto-organização de von der Malsburg (1973) mostram que um modelo de córtex visual não poderia ser inteiramente predeterminado geneticamente. Em vez disso, um processo auto-organizável envolvendo aprendizagem sináptica deve ser responsável pela ordenação *local* de células locais sensitivas a características.

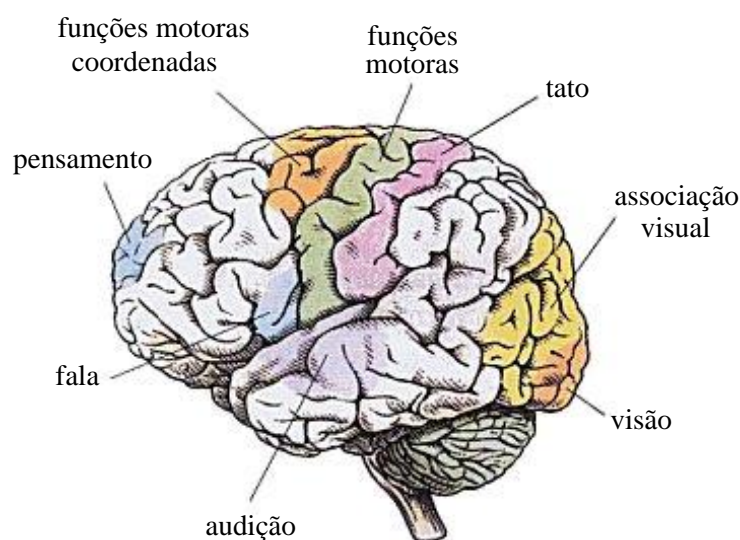


Fig. 3.1: *Representação das várias regiões corticais especializadas no cérebro humano.*

Nesse contexto, vários modelos de mapas topográficos ou mapas topologicamente corretos foram propostos. Dentre eles destacam-se os modelos de Willshaw e von der Malsburg (1976) e o de Kohonen (1982).

O modelo de Willshaw e von der Malsburg foi originalmente proposto sobre bases biológicas para explicar o problema do mapeamento retinotópico da retina para o córtex visual. A idéia básica desse modelo é a de codificar a proximidade geométrica de neurônios pré-sinápticos na forma de correlações na sua atividade elétrica, usando estas correlações para conectar neurônios pré-sinápticos vizinhos com neurônios pós-sinápticos vizinhos. Dessa forma, um mapeamento topologicamente ordenado é produzido por auto-organização. Entretanto, deve-se notar que o modelo de Willshaw e von der Malsburg é especializado em mapeamentos nos quais a dimensão de entrada é a mesma que a dimensão de saída.

O modelo de Kohonen, que não procura explicar detalhes neurobiológicos, captura as características essenciais dos mapas topográficos do cérebro, baseia-se no esquema de aprendizado competitivo e ainda se mantém tratável do ponto de vista computacional. Esse modelo é aparentemente mais geral que o modelo de Willshaw e von der Malsburg, uma vez que ele apresenta a capacidade de realizar a redução da dimensionalidade de entrada.

O modelo de Kohonen recebeu muito mais atenção na literatura que o modelo de Willshaw e von der Malsburg. Ele possui propriedades importantes (discutidas na seção 3.5) que o tornam

particularmente interessante para a compreensão e modelagem de mapas corticais do cérebro (Haykin, 1994).

### 3.3 Estrutura básica do SOM

O principal objetivo do SOM é realizar uma projeção não-linear de um espaço de entrada  $R^d$ , geralmente de dimensão elevada ( $d \gg 2$ ), em um conjunto discreto de neurônios, normalmente dispostos em um arranjo unidimensional ou bidimensional. Essa projeção é realizada adaptativamente de uma maneira topologicamente ordenada. Relações estatísticas complexas e não-lineares existentes entre os dados de entrada são transformadas em relações geométricas simples sobre uma grade de neurônios de menor dimensionalidade.

O SOM consiste basicamente de duas camadas de neurônios: uma camada de entrada e uma camada de saída (essa última comumente denominada de *camada de Kohonen*). As entradas da rede correspondem a vetores no espaço  $p$ -dimensional. Cada neurônio  $i$  da camada de Kohonen apresenta-se totalmente conectado com todas as entradas da rede, possuindo um vetor de pesos sinápticos associado, também no espaço  $p$ -dimensional,  $\mathbf{w}_i = [w_{i1}, w_{i2}, \dots, w_{ip}]^T$  (figura 3.2). Os neurônios na camada de Kohonen são conectados aos neurônios adjacentes através de uma relação de vizinhança que descreve a estrutura topológica do mapa. Conforme descrito anteriormente, essa estrutura geralmente é bidimensional, podendo apresentar uma vizinhança retangular ou hexagonal (figura 3.3). A escolha da topologia, assim como o tamanho do mapa (número de neurônios), depende da aplicação. No entanto, por razões de simplicidade, mapas retangulares são utilizados mais freqüentemente.

Durante o processo de aprendizagem do SOM, os vetores de dados de entrada são apresentados repetidamente à rede de neurônios em ordem aleatória. Para cada vetor de entrada apresentado haverá uma competição entre todos os neurônios pelo direito de representá-lo. O neurônio cujo vetor de pesos sinápticos for o mais próximo do dado de entrada, segundo uma métrica previamente definida (geralmente a Euclidiana), vence a competição. Esse neurônio é chamado de neurônio vencedor (ou BMU - *Best Matching Unit*) e este passo do processo de aprendizagem da rede é conhecido como *estágio competitivo*.

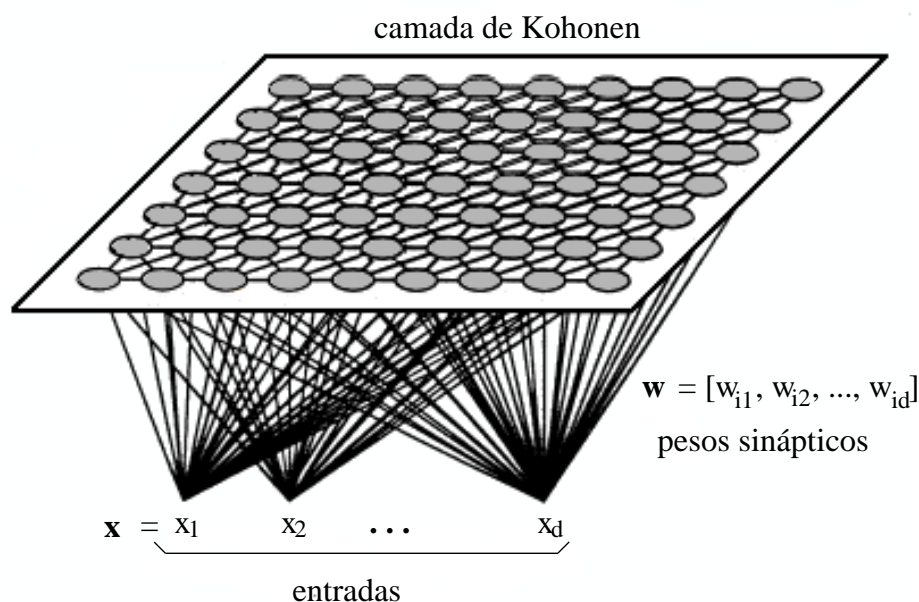


Fig. 3.2: *Estrutura básica do SOM.*

Uma vez determinado o neurônio vencedor para um dado padrão de entrada, esse neurônio é adaptado, isto é, o seu vetor de pesos sinápticos é ajustado de maneira que fique mais próximo do vetor de entrada, aumentando assim a chance de que esse mesmo neurônio vença a competição numa subsequente apresentação do mesmo padrão. Para viabilizar o requisito de que neurônios próximos no arranjo vençam para padrões que estejam próximos uns dos outros no espaço de entrada, neurônios pertencentes a uma vizinhança do neurônio vencedor também terão os vetores de pesos ajustados na direção do padrão de entrada, porém com menor intensidade (figura 3.3). A intensidade de ajuste do neurônio vizinho depende da distância dele ao neurônio vencedor. O conjunto de neurônios vizinhos que serão ajustados juntamente com o neurônio vencedor é definido por um raio de vizinhança, que normalmente abrange todo o arranjo de neurônios no início do treinamento, e depois decresce com o aumento do número de iterações de treinamento. Esse passo de ajuste de vizinhança é denominado de *estágio cooperativo*. O neurônio vencedor determina a localização espacial de uma vizinhança topológica de neurônios excitados, fornecendo assim a base para a cooperação entre os neurônios vizinhos.

Mais detalhes sobre os estágios de competição e de cooperação do SOM são descritos na

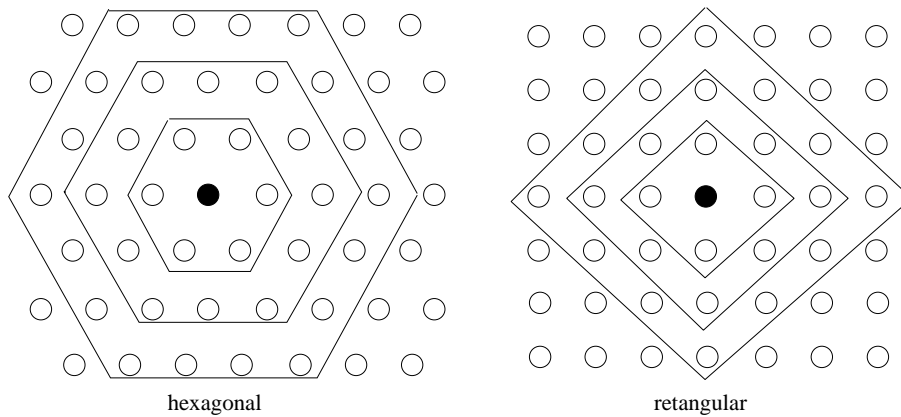


Fig. 3.3: Duas configurações de arranjo de neurônios do SOM e níveis de vizinhança ao redor do neurônio vencedor.

próxima seção, onde é apresentado o algoritmo de treinamento da rede.

### 3.4 Algoritmo de treinamento

O algoritmo responsável pelo treinamento não-supervisionado do SOM começa primeiramente inicializando os pesos sinápticos do arranjo de neurônios. Após a inicialização dos pesos, existem três passos básicos envolvidos na sua aplicação: amostragem, casamento por similaridade e atualização. Esses três passos são repetidos até que a rede neural não apresente mais variações significativas nos seus vetores de pesos sinápticos. Um resumo do algoritmo é apresentado como segue:

1. *Inicialização*: Escolha valores aleatórios para inicializar os vetores de pesos sinápticos  $\mathbf{w}_j$  da rede. Pode-se usar outras estratégias de inicialização apresentadas na seção 3.4.1. A única restrição nesse passo do algoritmo é que os pesos  $\mathbf{w}_j$  sejam diferentes para  $j = 1, 2, \dots, N$ , onde  $N$  é o número de neurônios do SOM.
2. *Amostragem*: Retire um vetor de amostra  $\mathbf{x}$  do espaço de entrada. O vetor  $\mathbf{x}$  representa o padrão de ativação que é aplicado à grade ou arranjo de neurônios.
3. *Casamento por similaridade*: Determine o neurônio vencedor  $i(\mathbf{x})$ , isto é, o neurônio com o melhor casamento na iteração  $t$  do treinamento, usando o critério da mínima distância

euclidiana:

$$i(\mathbf{x}) = \arg \min_j \|\mathbf{x}(n) - \mathbf{w}_j\|, \quad j = 1, 2, \dots, N. \quad (3.1)$$

Esse passo do processo de treinamento corresponde ao estágio *competitivo* da rede.

4. *Atualização*: Ajuste os vetores de pesos sinápticos de todos os neurônios usando a seguinte fórmula de atualização:

$$\mathbf{w}_j(t+1) = \mathbf{w}_j(t) + \eta(t)h_{j,i(\mathbf{x})}(t) (\mathbf{x}(t) - \mathbf{w}_j(t)) \quad (3.2)$$

onde  $\eta(t)$  é o parâmetro da taxa de aprendizagem e  $h_{j,i(\mathbf{x})}(t)$  é a função de vizinhança centrada em torno do neurônio vencedor  $i(\mathbf{x})$ . Ambos  $\eta(t)$  e  $h_{j,i(\mathbf{x})}(t)$  variam dinamicamente durante o treinamento. Essa etapa corresponde ao estágio *cooperativo* do processo de aprendizagem do SOM.

5. *Continuação*: Repita os passos de 2 a 4 até que a atualização de pesos não seja mais significativa.

### 3.4.1 Considerações sobre os parâmetros de treinamento

A configuração dos parâmetros de treinamento do SOM não é uma tarefa fácil de ser efetuada. Existem várias propostas de heurísticas e estratégias para escolher adequadamente os parâmetros da rede de Kohonen, muitas das quais são específicas para determinados tipos de problema ou aplicação (Kaski, 1997; Haese, 2001).

Os parâmetros que regulam o SOM são muitos e basicamente se dividem entre aqueles que controlam a estrutura do mapa (dimensão, tamanho, vizinhança e topologia do arranjo) e aqueles que controlam o treinamento propriamente dito (inicialização dos pesos, apresentação dos padrões, taxa de aprendizagem, raio e função de vizinhança, número de épocas).

A seguir são apresentadas algumas considerações sobre a escolha desses parâmetros. Maiores detalhes sobre a configuração de parâmetros de treinamento do SOM podem ser obtidos em



Costa (1999) e em Zuchini (2003).

### Dimensão e tamanho do mapa

Geralmente SOMs bidimensionais são escolhidos por permitirem a visualização do mapeamento dos dados na forma de um *display*. Embora seja a única configuração para a qual uma prova de convergência do SOM foi estabelecida (Erwin *et al.*, 1992), mapas unidimensionais são pouco explorados na prática, sendo utilizados basicamente em demonstrações sobre o comportamento do SOM.

O tamanho ou número de neurônios do SOM normalmente é fixado antes do início do treinamento. Segundo Costa (1999), para problemas de agrupamentos, o ideal é usar um número bem menor de neurônios do que a quantidade de dados no espaço de entrada. Em muitos casos práticos, pode ser útil dimensionar o mapa supondo que alguns neurônios serão inativos, isto é, não serão vencedores para nenhum padrão de entrada. Esses neurônios inativos podem funcionar como bordas de separação entre agrupamentos de dados representados pelos neurônios do SOM (Vesanto e Alhoniemi, 2000; Gonçalves *et al.*, 2007a).

Na literatura existem algumas propostas de cálculo para a determinação do número de neurônios da rede de Kohonen. Vesanto e Alhoniemi (2000), por exemplo, propõe a fórmula dada por  $5\sqrt{N}$  como uma estimativa razoável para o número de neurônios da rede, onde  $N$  é a quantidade de dados de entrada. Caso a quantidade de neurônios seja pequena (menor que 1000), Kaski (1997) propõe que o número de neurônios seja igual ao dos dados de entrada.

### Topologia da rede

Conforme descrito anteriormente, em SOMs bidimensionais, geralmente usa-se uma topologia retangular ou hexagonal para conectar os neurônios vizinhos. Por razões de simplicidade, a topologia retangular é a mais comumente empregada. Outras configurações possíveis são arranjos cilíndricos e toroidais, mas são pouco exploradas (Vesanto e Alhoniemi, 2000).

Existem algumas propostas na literatura que definem a relação de vizinhança do SOM ao longo do treinamento, como aquela apresentada em Kangas *et al.* (1990). A relação de vizinhança proposta permite a representação de estruturas alongadas, com agrupamentos desco-

nexos, sendo bastante rápida na convergência. No entanto, a mesma não garante a ordenação topológica e não atua muito bem em agrupamentos de dados hiperesféricos. Em Martinetz e Schulten (1991) é proposto um modelo onde os neurônios não apresentam uma região de vizinhança fixa para efetuar o ajuste dos pesos sinápticos. O modelo é capaz de mapear conjuntos de dados com estruturas bastante complexas, inclusive com diferentes dimensões e regiões desconexas. Entretanto, também não garante a ordenação topológica do mapeamento, possibilitando que apenas o conceito de similaridade seja avaliado no final do treinamento.

### Inicialização dos pesos

Embora o SOM seja relativamente robusto em relação à inicialização dos pesos sinápticos, métodos que buscam inicializar de forma adequada os valores dos pesos podem facilitar a convergência do algoritmo (Costa, 1999).

Normalmente, escolhe-se a forma aleatória (ou randômica) para a inicialização, ou seja, nenhum estado de organização inicial dos pesos é considerado. No entanto, Kohonen (1997) propôs a idéia de *inicialização linear*, a qual pode ser considerada mais eficiente do que a forma de inicialização aleatória, uma vez que permite uma convergência mais rápida do algoritmo de treinamento.

A inicialização linear de um SOM tenta pré-organizar o espaço de saída da rede utilizando-se dos  $K$  componentes principais da matriz de auto-correlação do conjunto de dados de entrada  $X$ . Mais detalhadamente, utilizando essa forma de inicialização, o arranjo de neurônios do SOM fica inicialmente definido ao longo de um subespaço linear correspondente aos  $K$  auto-vetores da matriz de auto-correlação de  $X$  que possuem os maiores auto-valores, sendo seu centróide coincidente com o centro de massa do conjunto de dados  $X$ .

### Função e raio de vizinhança

A função de vizinhança  $h_{j,i}(\mathbf{x})(t)$ , na equação 3.2, faz com que a vizinhança topológica em torno do neurônio vencedor decaia suavemente com a distância lateral no arranjo de neurônios do SOM. Considerando que  $h_{j,i}$  representa a vizinhança topológica centrada no neurônio vencedor  $i$ , contendo um conjunto de neurônios  $j$  excitados (cooperativos), e supondo que  $d_{j,i}$  rep-

resenta a distância lateral entre o neurônio vencedor  $i$  e o neurônio excitado  $j$ , então pode-se assumir que a vizinhança topológica  $h_{j,i}$  é uma função unimodal da distância  $d_{j,i}$ , desde que satisfaça duas exigências distintas (Haykin, 1994):

- A vizinhança topológica  $h_{j,i}$  é simétrica em relação ao ponto máximo definido por  $d_{i,j} = 0$ , ou seja,  $h_{j,i}$  alcança o seu valor máximo no neurônio vencedor  $i$  para o qual a distância  $d_{j,i}$  é zero.
- A amplitude da vizinhança topológica  $h_{j,i}$  decresce monotonamente com o aumento da distância lateral  $d_{i,j}$ , decaindo a zero para  $d_{j,i} \rightarrow \infty$ .

Uma escolha típica da função  $h_{j,i}$  que satisfaz as exigências acima é a função *gaussiana*, dada por:

$$h_{j,i}(x) = \exp\left(-\frac{d_{j,i}^2}{2\sigma^2}\right) \quad (3.3)$$

onde o parâmetro  $\sigma$  é a largura efetiva da vizinhança topológica, medindo o grau com o qual os neurônios excitados na vizinhança do neurônio vencedor participam do processo de aprendizagem. A função gaussiana é invariante à translação, ou seja, independe da localização do neurônio vencedor. Além disso, em termos qualitativos, a vizinhança topológica gaussiana é biologicamente mais apropriada do que a vizinhança do tipo *bolha* proposta pelo algoritmo original do SOM (Kohonen, 1982).

No caso de uma grade bidimensional, a distância lateral  $d_{j,i}$  entre o neurônio vencedor  $i$  e o neurônio excitado  $j$  é definida por:

$$d_{j,i}^2 = \|\mathbf{r}_j - \mathbf{r}_i\|^2 \quad (3.4)$$

onde o vetor discreto  $\mathbf{r}_j$  define a posição do neurônio excitado  $j$  e  $\mathbf{r}_i$  define a posição discreta do neurônio vencedor  $i$ , sendo ambos medidos no espaço de saída do SOM.

Uma outra característica do SOM é que o raio ou a largura da vizinhança topológica,  $\sigma$ , diminui com o tempo. Uma escolha usual para a dependência de  $\sigma$  com o tempo discreto  $t$  é o

decaimento exponencial dado por:

$$\sigma(t) = \sigma_0 \exp\left(-\frac{t}{\tau}\right), \quad t = 0, 1, 2, \dots, \quad (3.5)$$

onde  $\sigma_0$  é o valor de  $\sigma$  na inicialização do algoritmo SOM, e  $\tau$  é uma constante de tempo. Conseqüentemente, a função de vizinhança topológica do SOM assume uma forma variável no tempo, conforme mostra a equação:

$$h_{j,i(x)}(t) = \exp\left(-\frac{d_{j,i}^2}{2\sigma^2(t)}\right), \quad t = 0, 1, 2, \dots, \quad (3.6)$$

onde  $\sigma(t)$  é definido pela equação 3.5.

Portanto, quando o tempo ou o número de iterações  $t$  aumenta, a largura  $\sigma(t)$  decresce à uma taxa exponencial e a vizinhança topológica diminui de modo correspondente.

### Taxa de aprendizagem

Assim como a função de vizinhança, a taxa de aprendizagem  $\eta(t)$ , dada pela equação 3.2, também deve decrescer monotonicamente com o tempo. Normalmente,  $\eta(t)$  é inicializada com um valor próximo a 1, e decresce com o tempo de forma exponencial, linear ou inversamente proporcional a  $t$ .

### Apresentação dos padrões

Existem dois modos usuais de apresentar os dados de um conjunto de entrada  $X$  para o SOM. Esses dois modos de apresentação, denominados de *incremental* e *em lote*, alteram a maneira pela qual a atualização dos pesos sinápticos da rede é feita durante o seu treinamento. No algoritmo incremental, os pesos sinápticos são ajustados toda vez que um dado de entrada é apresentado à rede. Já no algoritmo em lote, as atualizações individuais dos pesos são preteridas e aplicadas somente após a apresentação de todos os elementos do conjunto de entrada  $X$ .

O algoritmo incremental tem como principal vantagem a possibilidade de uso de SOMs em problemas para os quais não se tem antecipadamente todos os dados de entrada disponíveis. No

entanto, o algoritmo incremental é sensível à ordem de apresentação dos dados e particularmente sensível à taxa de aprendizagem (Kohonen, 1997). O algoritmo em lote elimina esses problemas. Escolhe-se um valor pequeno e fixo para a taxa de aprendizagem ( $\eta(t)$ ), geralmente próximo a 0.1, e realiza-se o treinamento somando-se a média das contribuições de todos os padrões de entrada para cada neurônio. Na prática, esse deslocamento médio tem conduzido a melhores resultados que o do algoritmo incremental (Costa, 1999).

### Número de épocas de treinamento

Segundo Kohonen (1982), o treinamento do SOM ocorre em duas fases: uma de ordenação inicial e outra de convergência. Na primeira fase, ocorre a ordenação topológica dos vetores de pesos. Na segunda fase, ocorre um ajuste mais fino do mapa. É importante observar que Kohonen (1997) sugere que a inicialização linear dos pesos (mencionada na seção 3.4.1) pode eliminar a fase de ordenação inicial.

Várias provas de convergência para o SOM foram estudadas através de diferentes abordagens (Cottrell *et al.*, 1995; Flanagan, 1996), porém apenas no caso unidimensional foi possível demonstrar o processo de ordenação da rede. Uma análise matemática rigorosa da dinâmica do algoritmo de treinamento e uma prova de convergência para SOMs bidimensionais ainda não foram registradas na literatura. Diante disso, diferentes abordagens têm sido propostas para determinar o número de épocas de treinamento com o objetivo de se obter bons mapas (Vesanto e Alhoniemi, 2000; Azcarraga, 2000). No entanto, devido à ausência de fundamentação teórica sólida para o SOM, essas abordagens são muito dependentes da aplicação, não sendo adequadas para quaisquer tipos de dados. Conseqüentemente, na maioria dos casos de uso do SOM, o número de épocas de treinamento da rede é definido experimentalmente.

Segundo Costa (1999), condições gerais para se obter bons mapas incluem treinar a rede com um número de épocas relativamente alto, mantendo na fase inicial o raio de vizinhança e a taxa de aprendizagem bastante elevados, permitindo um decaimento suave desses parâmetros com o tempo.

## 3.5 Propriedades do SOM

Considerando que o treinamento do SOM tenha sido bem sucedido, o mapa de neurônios produzido pelo seu algoritmo de treinamento mostra características estatísticas importantes do espaço de entrada. As principais propriedades do mapeamento realizado pelo SOM podem ser resumidas como seguem:

- *Aproximação do espaço de entrada*: o mapa de neurônios do SOM, representado pelo conjunto de vetores de pesos sinápticos  $\mathbf{w}_j$ , fornece uma boa aproximação para o espaço de entrada. O SOM tem como objetivo básico armazenar um conjunto grande de vetores de entrada  $\mathbf{x}$  encontrando um conjunto menor de protótipos  $\mathbf{w}_j$ , provendo uma boa aproximação para o espaço de entrada original. A base teórica dessa estratégia está fundamentada na teoria da *quantização vetorial*, cuja motivação é a redução de dimensionalidade ou a compressão de dados.
- *Ordenação Topológica*: o mapeamento produzido pelo SOM é ordenado topologicamente. Ao realizar a projeção não-linear do espaço de entrada para o espaço de saída da rede, o algoritmo do SOM tenta preservar ao máximo a topologia do espaço original, ou seja, procura fazer com que neurônios vizinhos na grade retangular (espaço de saída) apresentem vetores de pesos que representem padrões vizinhos no espaço de entrada. Essa propriedade é uma consequência direta da equação 3.2, que força os vetores de pesos sinápticos do neurônio vencedor e dos seus neurônios vizinhos a se moverem em direção ao vetor de entrada  $\mathbf{x}$ .
- *Casamento de Densidade*: o SOM treinado reflete a distribuição de probabilidade dos dados no espaço de entrada. Regiões no espaço de entrada de onde os vetores de amostras  $\mathbf{x}$  são retirados com uma alta probabilidade de ocorrência são mapeados para domínios maiores no espaço de saída da rede, e portanto, com melhor resolução que regiões no espaço de entrada das quais vetores de amostras  $\mathbf{x}$  são retirados com uma baixa probabilidade de ocorrência.

Uma discussão mais detalhada sobre as propriedades do SOM pode ser obtida em Haykin (1994).

### 3.6 Métodos de visualização

Após o treinamento do SOM, métodos de visualização podem ser utilizados para analisar o mapeamento efetuado pela rede neural.

Para situações em que os vetores de pesos sinápticos possuem dimensão  $d$  menor ou igual a 3, pode-se usar os seus valores como coordenadas no espaço  $\mathcal{R}^d$  para visualizar a organização dos neurônios. A figura 3.7 mostra, por exemplo, os neurônios de um SOM treinado com dados no espaço bidimensional.

O *histograma de vencedores* também é outra forma de obter informações sobre o mapeamento do SOM, mais particularmente sobre a quantização vetorial produzida pela rede. Neste caso, os dados podem ser projetados no SOM atribuindo, para cada vetor de entrada  $x$ , o seu respectivo neurônio vencedor. A frequência de ocorrência de vencimento para um dado neurônio gera um histograma que refletirá o nível de atividade deste neurônio. A figura 3.8 mostra os neurônios do SOM apresentados na figura 3.7, porém com tamanho proporcional à sua atividade. A figura 3.9 mostra o histograma de vencedores daqueles mesmos neurônios, onde é possível visualizar o grau de atividade de cada um deles em sua respectiva posição na grade (ou arranjo) do SOM.

Para que se possa ter uma noção de como cada componente dos vetores de dados de entrada organizou-se no SOM treinado, pode-se utilizar algum método de coloração do SOM baseado nos valores dos seus *planos de componentes*. Para um dado componente  $j$  de um SOM bidimensional, gera-se uma imagem  $f(x, y)$  com as mesmas dimensões do mapa, onde cada pixel corresponderá ao valor do componente  $j$  na posição  $(x, y)$ . Este método de visualização do SOM permite avaliar a correlação entre componentes do espaço de entrada ao longo da rede neural (Kohonen, 1997; Vesanto, 1997).

Quando o objetivo de uso do SOM é a detecção de agrupamentos (como é o caso deste trabalho), a técnica mais freqüentemente empregada para visualizar o SOM treinado é o da

matriz de distâncias unificada ou *U-matrix*.

Uma vez que os métodos de classificação de imagens propostos neste trabalho baseiam-se no uso da *U-matrix*, ela é apresentada com maiores detalhes na próxima seção.

### 3.6.1 Matriz de distâncias unificada - *U-matrix*

O método de visualização do SOM treinado, denominado *matriz de distâncias unificada* ou *U-matrix*, foi desenvolvido por Ultsch (1993a,b) com o objetivo de permitir a detecção visual das relações topológicas dos neurônios na rede de Kohonen. A idéia básica do método é utilizar a mesma métrica que é empregada durante o treinamento do SOM para calcular distâncias entre os vetores de pesos sinápticos de neurônios adjacentes. Desse modo, a *U-matrix* apresenta a vantagem de permitir a visualização das relações entre os neurônios de um SOM até mesmo em problemas em que o vetor de entrada possui uma dimensão maior que 2 ou 3.

Considerando um mapa cujos neurônios possuem vizinhança retangular, a *U-matrix* é obtida através dos cálculos das distâncias  $dx$ ,  $dy$ ,  $dxy$  e  $du$  para cada neurônio, conforme ilustra a figura 3.4.

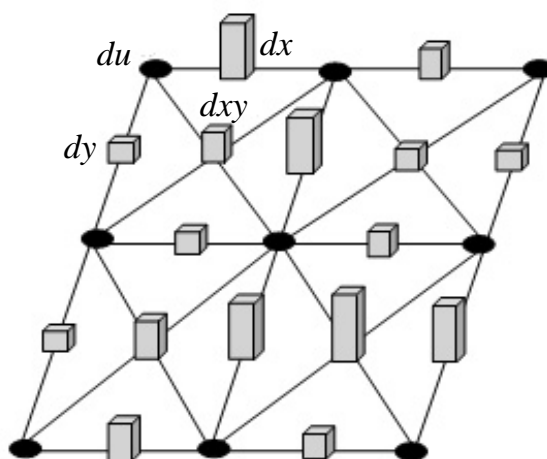


Fig. 3.4: Representação das distâncias utilizadas no cálculo da *U-matrix* em um arranjo de neurônios retangular. Os círculos representam os neurônios e os cubos representam as distâncias.



Adotando a métrica Euclidiana, as distâncias  $d_x$ ,  $d_y$  e  $d_{xy}$  podem ser definidas como:

$$d_x(x, y) = \sqrt{\sum_i (\mathbf{w}_{j_{x,y}} - \mathbf{w}_{j_{x+1,y}})^2} \quad (3.7)$$

$$d_y(x, y) = \sqrt{\sum_i (\mathbf{w}_{j_{x,y}} - \mathbf{w}_{j_{x,y+1}})^2} \quad (3.8)$$

$$d_{xy}(x, y) = \frac{1}{2} \left[ \frac{\sqrt{\sum_i (\mathbf{w}_{j_{x,y}} - \mathbf{w}_{j_{x+1,y+1}})^2}}{\sqrt{2}} + \frac{\sqrt{\sum_i (\mathbf{w}_{j_{x,y+1}} - \mathbf{w}_{j_{x+1,y}})^2}}{\sqrt{2}} \right] \quad (3.9)$$

onde  $(x, y)$  representa as coordenadas de um neurônio no arranjo do SOM e  $w_{j_{x,y}}$  representa o seu vetor de pesos sinápticos. O valor  $du$  da *U-matrix* pode ser calculado como a média, mediana, valor mínimo ou máximo das distâncias circunvizinhas. A equação 3.10 mostra o esquema de preenchimento dos elementos da *U-matrix* considerando um SOM retangular de tamanho  $M \times N$ :

$$\begin{bmatrix} du(0,0) & dx(0,0) & du(1,0) & \dots & du(M-1,0) \\ dy(0,0) & dxy(0,0) & dy(1,0) & \dots & dy(M-1,0) \\ du(0,1) & dx(0,1) & du(1,1) & \dots & du(M-1,1) \\ dy(0,1) & dxy(0,1) & dy(1,1) & \dots & dy(M-1,1) \\ \dots & \dots & \dots & \dots & \dots \\ du(0,N-1) & du(0,N-1) & du(1,N-1) & \dots & du(M-1,N-1) \end{bmatrix} \quad (3.10)$$

O resultado gerado a partir da aplicação da *U-matrix* sobre um mapa retangular de neurônios  $M \times N$  é uma imagem  $f(x, y)$ , de tamanho  $(2M-1) \times (2N-1)$ , na qual as coordenadas de cada pixel  $(x, y)$  são derivadas das coordenadas dos neurônios no arranjo, e a intensidade de cada pixel na imagem  $f(x, y)$  corresponde a uma distância calculada.

É possível também abstrair a imagem da *U-matrix* como uma função tridimensional, de modo que o valor do pixel na coordenada  $(x, y)$  seja representado por um ponto na coordenada

*z*. Neste caso, tem-se uma superfície 3D cuja topografia revela a configuração dos neurônios obtida após o treinamento. Superfícies baixas (ou vales), neste relevo topográfico, correspondem a neurônios que são similares, e portanto representam agrupamentos de neurônios. Superfícies altas (ou montanhas) refletem a dissimilaridade entre neurônios, sendo associadas a regiões ou neurônios em fronteiras (ou bordas) de agrupamentos. As figuras 3.10 e 3.11 mostram, respectivamente, a *U-matrix* do SOM da figura 3.7 na forma planar (2D) e na forma de superfície (3D).

## 3.7 Métodos de segmentação do SOM

Conforme descrito anteriormente, o objetivo principal do SOM é mapear um espaço de entrada  $p$ -dimensional contínuo para um conjunto finito de neurônios organizados em um arranjo normalmente unidimensional ou bidimensional. Ao realizar este mapeamento, o algoritmo de aprendizagem da rede busca reduzir a dimensionalidade do espaço de entrada ao mesmo tempo em que procura preservar ao máximo a topologia dos dados originais. Esta redução de dimensionalidade com preservação topológica permite ampliar a capacidade de análise de agrupamentos de dados pertencentes a espaços de dimensão elevada. Tal característica faz com que o SOM exerça grande atração em uma variedade de aplicações.

Entretanto, apesar de todos os seus benefícios, a aplicação do SOM tem sido limitada por algumas dificuldades relacionadas à interpretação do seu mapa de neurônios. Analisar a estrutura e as características dos dados através da simples visualização do SOM treinado não é uma tarefa trivial. Em geral, a determinação das fronteiras entre partes não relacionadas no espaço de saída da rede somente é possível quando a pessoa que examina a representação padrão do mapa de neurônios tem algum conhecimento prévio sobre os dados de entrada. Portanto, o mapeamento topologicamente ordenado produzido pelo SOM não é suficiente para efetuar análises de agrupamentos. De fato, como a saída do SOM para um dado padrão é apenas o índice do neurônio vencedor e o nível de ativação diretamente relacionado à quantização, a rede necessita de outras ferramentas para que se possa realizar uma interpretação apropriada do seu mapa de saída. Além disso, em tarefas de classificação automática ou de análise de

agrupamentos de dados, a detecção de agrupamentos através dos neurônios do SOM deve idealmente ser feita sem a intervenção do usuário.

Na literatura, existem diferentes propostas de algoritmos para analisar de maneira automática (ou semi-automática) os neurônios de um SOM treinado. Em Vesanto e Alhoniemi (2000), uma abordagem utilizando o algoritmo K-médias é empregada para particionar os neurônios do SOM após o seu treinamento. Os experimentos indicam que analisar agrupamentos de dados através dos protótipos do SOM ao invés de trabalhar diretamente com os dados originais é uma estratégia computacionalmente efetiva. Entretanto, como os próprios autores relatam, é importante observar que o uso específico do algoritmo K-médias impõe implicitamente suposições sobre as formas dos agrupamentos, sendo eficaz apenas para detectar agrupamentos com formatos hipersféricos. Além disso, o algoritmo exige que o usuário interaja no processo definindo previamente o número de classes desejadas, não caracterizando desse modo uma abordagem totalmente automática. Apesar dessas limitações, a forma padrão do algoritmo K-médias tem sido empregada para particionar o SOM em diferentes aplicações, como por exemplo, em Wang *et al.* (2002), García e González (2004), González e García (2006), e em Wen *et al.* (2007). Procurando contornar suas limitações, Vesanto e Alhoniemi (2000) propõem a execução do algoritmo K-médias considerando diferentes números de classes e aplicam o índice de validação de agrupamentos Davies-Bouldin (Davies e Bouldin, 1979) para determinar o melhor particionamento do SOM. Tal estratégia diminui a participação do usuário no processo, mas tendo em vista a necessidade de executar o algoritmo K-médias diversas vezes, o procedimento pode exigir um alto custo computacional quando aplicado em grandes conjuntos de dados. Outra desvantagem da abordagem empregada por Vesanto e Alhoniemi (2000) está relacionada com o índice de validação de agrupamentos empregado. O índice Davies-Bouldin é adequado apenas para avaliar conjuntos de dados que possuem formatos hipersféricos e bem compactados. Alternativamente aos métodos particionais (como o K-médias), Vesanto e Alhoniemi (2000) propuseram ainda a utilização de métodos de agrupamentos hierárquicos aglomerativos para segmentar o SOM. Apesar de apresentarem resultados satisfatórios, o critério adotado pelos autores para decidir sobre a fusão de agrupamentos também se baseia no índice de validação de agrupamentos Davies-Bouldin.

Na literatura, existem alguns outros trabalhos que também propõem a aplicação de métodos de agrupamentos hierárquicos para segmentar o SOM. No trabalho de Kiang (2001), o autor utilizou um método hierárquico aglomerativo empregando critérios de variância mínima na fusão de grupos de neurônios. O algoritmo necessita recalcular os centróides sempre que dois agrupamentos são unidos e apresenta boa aplicabilidade apenas para conjuntos de dados com formas hiperesféricas ou hiperelipsoidais. Em Yacoub *et al.* (2001), um método hierárquico é proposto para aglomerar os neurônios do SOM baseado em uma medida de similaridade que considera informações sobre o erro quadrático dos agrupamentos e a ordenação topológica do mapa. O método apresentou resultados satisfatórios na classificação de refletâncias de imagens oceânicas, porém emprega um parâmetro para controlar a influência das informações usadas no cálculo da medida de similaridade que depende fortemente do usuário. Wu e Chow (2004) propuseram também um método hierárquico para analisar a saída do SOM utilizando localmente um índice de validação de agrupamentos que se baseia em densidades inter e intra-agrupamentos. Comparado com outros métodos de agrupamentos clássicos, o algoritmo proposto em Wu e Chow (2004) utiliza mais informações sobre os dados além das medidas de distâncias tradicionais. Os resultados experimentais apresentados sobre quatro conjuntos de dados mostram que o algoritmo apresenta desempenho superior aos dos métodos de segmentação do SOM propostos por Murtagh (1995), Vesanto e Alhoniemi (2000) e Kiang (2001). Entretanto, é importante observar que embora a abordagem proposta por Wu e Chow (2004) realize a análise de agrupamentos dos dados através dos neurônios do SOM treinado, a decisão sobre a fusão de grupos de neurônios é efetuada através da aplicação do índice de validação diretamente sobre os dados originais (no espaço de entrada), procedimento que pode demandar um elevado tempo de processamento em aplicações que envolvem grandes volumes de dados. Finalmente, em Costa (2005), o autor aplicou o método das ligações simples com restrição de contigüidade para segmentar o SOM, e também propôs um mecanismo para estimar o número de agrupamentos existente no conjunto de dados baseado na estabilidade das fusões versus a distância entre os agrupamentos. Os resultados foram considerados satisfatórios para diferentes bases de dados.

Outros métodos de particionamento do SOM têm sido propostos baseando-se na segmentação da *U-matrix* (descrita na seção anterior). Em Costa (1999), um método de segmentação

automática do SOM foi desenvolvido usando técnicas de morfologia matemática para segmentar a matriz de distâncias unificada. O resultado do algoritmo são regiões conectadas de neurônios rotuladas, e que definem no espaço de entrada geometrias complexas e não-paramétricas. O método foi aplicado com sucesso em diversos conjuntos de dados com formatos e dimensões variadas (Costa e Andrade Netto, 1999, 2001). Soria-Frich e Köppen (2003) também propuseram um método de segmentação morfológica da imagem da *U-matrix* baseando-se em modelos de integral *fuzzy*. Apesar de apresentar resultados satisfatórios na classificação de imagens coloridas, o método proposto apresenta algumas desvantagens. Os agrupamentos de dados são representados por um único protótipo e o método exige que o usuário especifique previamente, além do número desejado de classes, um valor de limiar para segmentar a *U-matrix*, o qual influencia diretamente na estrutura dos agrupamentos. Em Opolon e Moutarde (2004), a segmentação do SOM é proposta através do pós-processamento da imagem da *U-matrix* utilizando técnicas elementares de processamento digital de imagens. Apesar do método proposto ser relativamente simples e apresentar um baixo custo computacional, o processo de segmentação da *U-matrix* não é automático, uma vez que utiliza parâmetros cujos valores são determinados experimentalmente. Além disso, os próprios autores reconhecem que o método necessita ser testado em conjuntos de dados reais e mais complexos.

Por fim, pode-se citar ainda os métodos de segmentação do SOM propostos por Costa e Andrade Netto (2005) e Tasdemir e Merényi (2005). Em Costa e Andrade Netto (2005), os autores propuseram um método de particionamento do SOM baseado em particionamento de grafos. Considerando o espaço de saída da rede como um grafo, o mapa de neurônios é particionado através da aplicação de regras heurísticas que eliminam conexões entre neurônios vizinhos consideradas inconsistentes. Embora o algoritmo utilize alguns limiares definidos de forma empírica, tanto o número de agrupamentos quanto os membros das classes são determinados automaticamente. Aplicações com sucesso do algoritmo foram obtidas em Silva *et al.* (2004) na análise exploratória de fragmentos urbanos utilizando dados geoespaciais. Em Tasdemir e Merényi (2005), o algoritmo proposto para segmentar o SOM utiliza informações sobre as densidades locais dos neurônios, explora a topologia dos mesmos e ainda emprega uma matriz de conectividade com informações de distâncias semelhante à *U-matrix*. Os resultados apresentados

destacam a eficiência do método na detecção de agrupamentos de dados com baixa densidade de probabilidade, mas também mostram que o método apresenta dificuldades na discriminação de alguns agrupamentos que apresentam alto grau de sobreposição.

Através desses trabalhos publicados, observa-se, portanto, a existência de diferentes propostas para realizar a segmentação automática ou semi-automática dos neurônios de um SOM treinado, cada uma delas com metodologias de agrupamentos diferenciadas (hierárquicas ou particionais) e com vantagens e desvantagens sob diferentes aspectos.

No presente trabalho, as metodologias de classificação de imagens de sensoriamento remoto propostas aplicam e aprimoram especificamente os métodos de segmentação do SOM baseados em técnicas de morfologia matemática (particularmente aquele proposto em Costa e Andrade Netto (2001)) e aqueles baseados em técnicas de agrupamentos hierárquicos. Como será apresentado adiante, esses métodos possuem características e propriedades que proporcionam uma série de vantagens às metodologias de classificação apresentadas nesta tese, as quais não são encontradas em métodos de classificação de imagens de sensoriamento remoto tradicionais.

Com o objetivo de facilitar o entendimento dos métodos de classificação de imagens propostos nesse trabalho, o método de segmentação automática do SOM proposto por Costa e Andrade Netto (2001) é descrito a seguir e um exemplo de aplicação sobre um conjunto de dados gerado artificialmente é apresentado.

### 3.7.1 Segmentação do SOM baseada em técnicas de morfologia matemática

Costa e Andrade Netto (1999, 2001) apresentaram um método de particionamento e rotulação automática do SOM baseado na segmentação da *U-matrix*. Toda análise e segmentação de um SOM treinado é feita inicialmente sobre a imagem da *U-matrix*, e posteriormente as informações correspondentes aos pixels são associadas aos neurônios. A *U-matrix* é segmentada por meio de um algoritmo de segmentação de imagens baseado em técnicas de morfologia matemática, denominado *watershed*. O algoritmo *watershed* foi proposto inicialmente por Beucher e Lantuéjoul (1979) e tem sido considerado como uma das ferramentas de segmentação mais

eficientes em processamento de imagens. Uma vez que a aplicação direta do algoritmo *watershed* sobre a imagem pode resultar numa sobre-segmentação da mesma, Costa e Andrade Netto (1999, 2001) utilizaram uma versão do algoritmo (proposta por Meyer e Beucher (1990)) que emprega marcadores para regular o processo de segmentação. Os marcadores são utilizados para identificar as regiões mais importantes da imagem. Eliminando os marcadores das regiões indesejáveis diminui-se automaticamente o número de partições na imagem final. Assim, apenas as regiões associadas aos mínimos descritos pelos marcadores terão rótulo, o que implica que o número de regiões finais da imagem segmentada será igual ao número de marcadores escolhidos.

Apesar de existirem diferentes métodos para escolha de bons marcadores para a imagem da *U-matrix*, Costa e Andrade Netto (1999, 2001) propuseram um método relativamente simples que tem gerado bons resultados em uma grande diversidade de problemas. Considerando a *U-matrix* de um SOM treinado dada pela imagem  $U$  com 256 níveis de cinza, os seguintes passos são efetuados para se obter a imagem de marcadores para  $U$ :

1. Filtragem: obter uma imagem  $U_1$  suavizando a imagem  $U$  através da remoção de pequenas depressões;
2. Para  $k = 1, \dots, f_{max}$ , onde  $f_{max}$  é o nível de cinza máximo da imagem  $U_1$ , criar as imagens binárias  $U_2^k$  correspondendo a conversões de  $U_1$ , usando  $k$  como valor de limiar;
3. Obter o número de regiões conectadas de  $U_2^k$ , para cada valor de  $k$ ,  $N_{rc}^k$ ;
4. Construir o gráfico  $k \times N_{rc}^k$  e encontrar no mesmo a maior seqüência contígua e constante de número de regiões conectadas  $N_{rc}^k$ , denotada por  $S_{max}$ .
5. A imagem de marcadores será a imagem  $U_2^j$ , onde  $j$  é o valor inicial da seqüência  $S_{max}$ .

Tendo em vista que a *U-matrix* apresenta geralmente um número considerável de rugosidades, o objetivo do passo 1 é suavizar muito discretamente a sua imagem original, resultando em uma imagem melhor para processamento. Costa (1999) sugere a remoção de pequenas depressões com áreas inferiores a 3 pixels. No entanto, a operação de filtragem também pode ser realizada através de filtros morfológicos usando um elemento estruturante de raio  $\rho$ , que

pode variar de acordo com o número de rugosidades da *U-matrix*. Quanto maior o raio  $\rho$ , mais forte será a filtragem (suavização). Caso a operação de filtragem não seja feita, o processo de limiarização (passo 2) irá produzir diversas regiões pouco significativas.

O passo 2 do algoritmo pode ser visto como um fatiamento da imagem em todos os seus níveis de cinza. Esses níveis são usados como limiares para binarizar a imagem. Para cada imagem binária derivada de uma operação de limiarização, o procedimento de rotulação atribui um código diferente a cada um dos componentes conectados, resultando em um número de regiões na imagem para cada valor de limiar  $k$  aplicado.

Definida a estratégia para determinar os marcadores da imagem da *U-matrix*, o algoritmo de segmentação do SOM, denominado por Costa (1999) de SL-SOM (*Self-Labeled SOM*), foi proposto considerando a execução dos seguintes passos:

1. Obter a *U-matrix*;
2. Determinar a imagem de marcadores da *U-matrix* (conforme o algoritmo descrito anteriormente);
3. Aplicar o algoritmo *watershed* sobre a *U-matrix* usando a imagem de marcadores obtida no passo 2;
4. Rotular as regiões conectadas da imagem segmentada no passo 3;
5. Copiar os rótulos obtidos no passo 4 para cada um dos neurônios associados aos pixels da *U-matrix*;
6. Caso ainda existam neurônios não rotulados, rotulá-los através do método do vizinho mais próximo, calculando as distâncias no espaço de pesos dos neurônios e atribuindo o código do neurônio rotulado mais próximo.

Com o objetivo de descobrir e representar subclasses de dados nos agrupamentos encontrados pelo algoritmo SL-SOM, Costa e Andrade Netto (1999, 2001) propuseram ainda uma extensão do modelo. Uma árvore dinâmica de redes SOM é construída automaticamente a partir de regiões do mapa que dão origem a outros mapas, os quais são treinados apenas com



subconjuntos de padrões que foram quantizados pela região do mapa inicial. Considerando o mapa inicial como mapa pai, cada região de neurônios  $R_k$  pode dar origem a um novo mapa, denominado mapa filho da região  $k$ ,  $M_k$ , que será treinado apenas com o subconjunto de dados  $X^k$ . O algoritmo SL-SOM aplicado no mapa pai é aplicado igualmente a cada um dos mapas filhos, resultando em (sub)regiões que podem dar origem a outros mapas filhos. Não se especifica *a priori* o número de níveis da árvore nem o número de submapas para cada rede, e a dinâmica de todo o processo pode ser pensada como um particionamento recursivo do conjunto de dados. Caso um submapa não possua partições ele é eliminado, e o subconjunto de dados fica representado pela região do mapa pai correspondente. Isto ocorre quando o sistema não consegue mais detectar subgrupos nos dados. No final, a estrutura da árvore obtida descreve relacionamentos hierárquicos entre as várias classes de padrões detectadas durante o processo. Essa extensão do algoritmo SL-SOM foi denominada por Costa e Andrade Netto (1999, 2001), de TS-SL-SOM (*Tree-Structured Self-Labeled SOM*). A figura 3.5 ilustra o processo de geração da árvore dinâmica de SOMs efetuado pelo TS-SL-SOM.

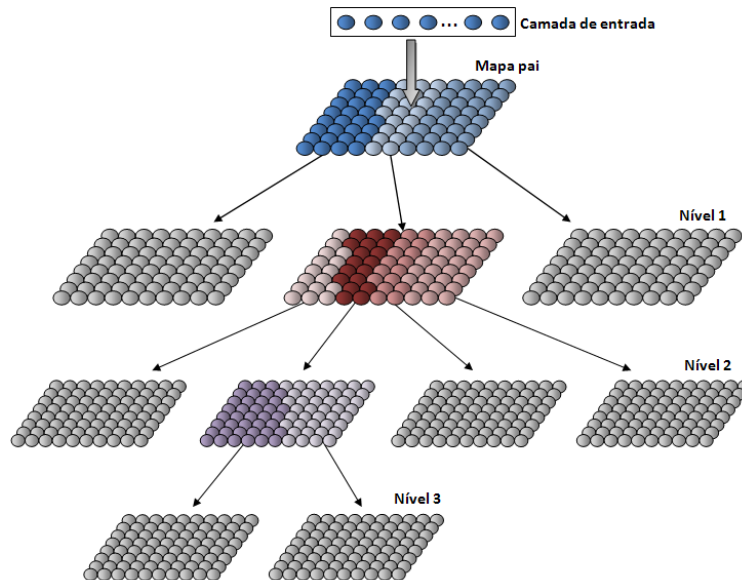


Fig. 3.5: Ilustração do processo de geração da árvore dinâmica de mapas pelo TS-SL-SOM (adaptado de Costa (1999)).

Conforme apresentado em Costa e Andrade Netto (1999, 2001), uma das principais vantagens do método SL-SOM está na utilização de conjuntos de protótipos para representar

agrupamentos, ou classes de objetos. Isso proporciona ao método a importante capacidade de se adaptar a diferentes geometrias no espaço de dados. Além disso, a extensão do modelo, o algoritmo TS-LS-SOM, possibilita a descoberta de agrupamentos mesmo quando o conjunto de dados apresenta classes de padrões com alto grau de sobreposição.

Para uma melhor compreensão do método de segmentação do SOM proposto por Costa e Andrade Netto (1999, 2001), resultados experimentais sobre um conjunto de dados gerado artificialmente são apresentados na próxima subseção. Maiores detalhes sobre os algoritmos SL-SOM e TS-SL-SOM podem ser obtidos em Costa (1999).

### 3.7.1.1 Exemplo de uso

O conjunto de dados abaixo (figura 3.6) foi gerado artificialmente com 1524 padrões, onde as classes 1 a 4 possuem, respectivamente, 73, 931, 61 e 459 padrões.

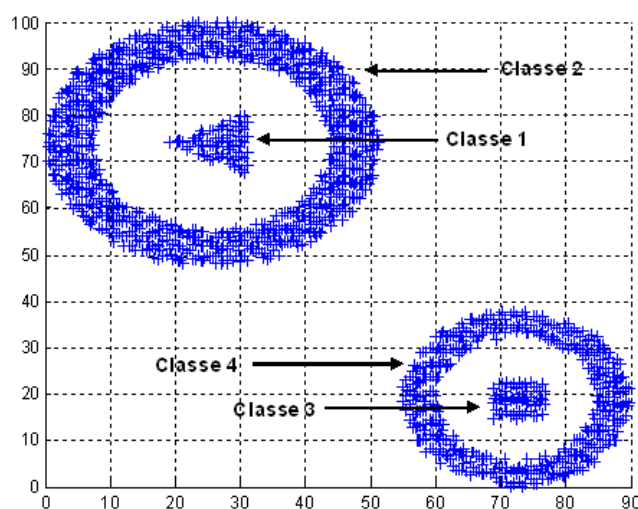


Fig. 3.6: *Conjunto de dados sintético.*

Neste experimento foi utilizado como mapa raiz do TS-SL-SOM um SOM com topologia retangular de tamanho 15x15. A função de vizinhança empregada foi a *gaussiana*, e os raios de vizinhança inicial e final foram 12 e 1, respectivamente. A inicialização dos pesos foi linear, a apresentação dos padrões em lote e o número de iterações foi 500. A configuração de neurônios após o treinamento é apresentada juntamente com os dados na figura 3.7.

A figura 3.8 mostra os neurônios do SOM com tamanho proporcional à sua atividade (fre-

quência de vencimento), e a figura 3.9 mostra o histograma de vencedores. A *U-matrix* é apresentada na forma planar na figura 3.10, e na forma de superfície na figura 3.11.

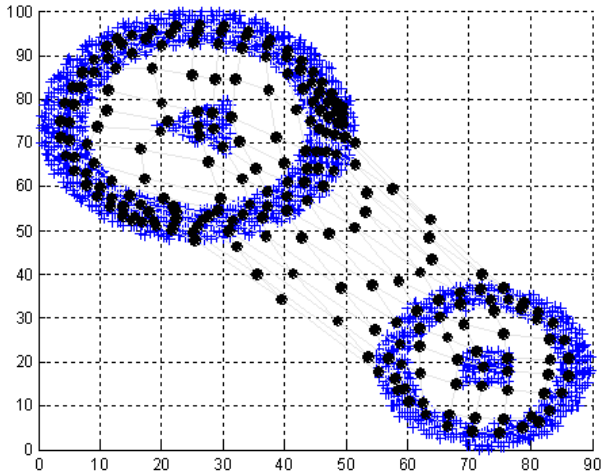


Fig. 3.7: *Dados e neurônios do SOM após o treinamento.*

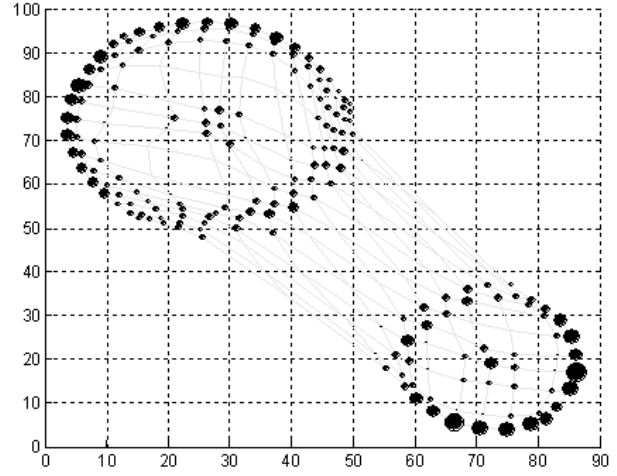
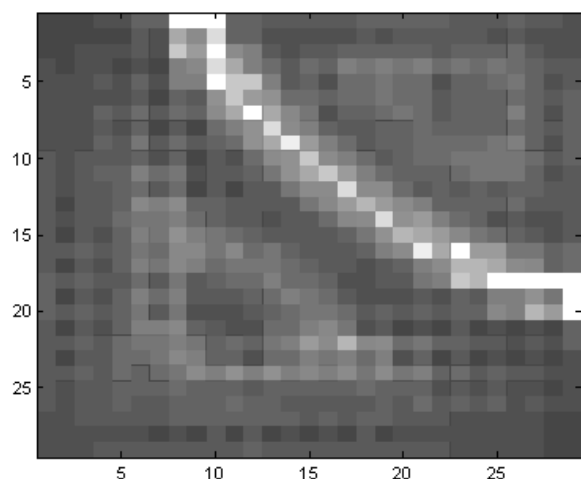
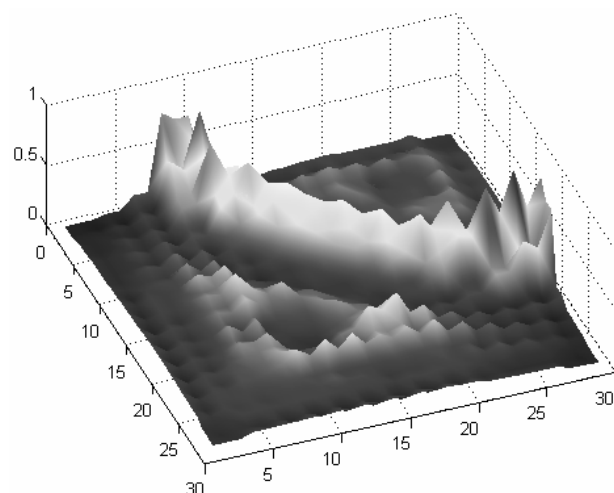


Fig. 3.8: *Neurônios do SOM com tamanho proporcional à sua atividade.*

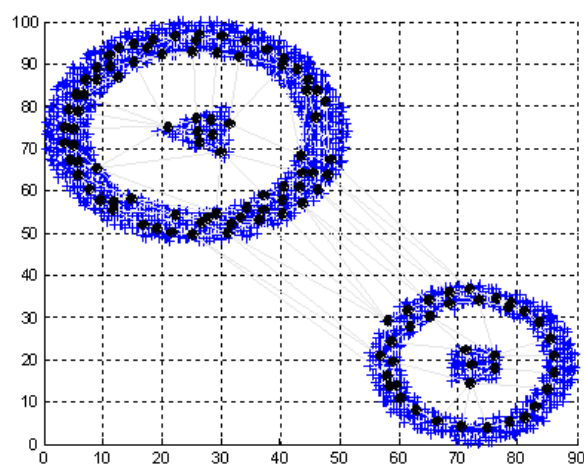
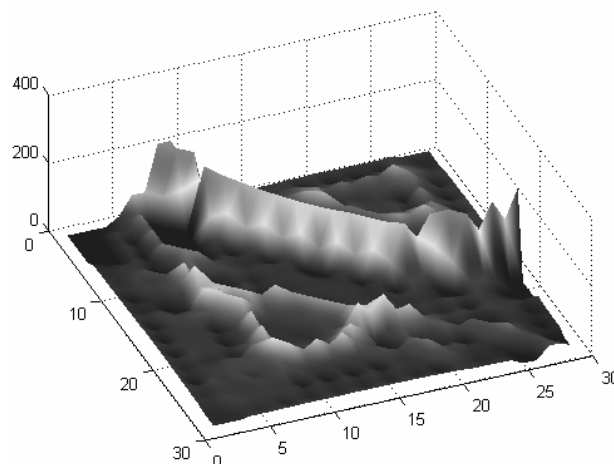
	2	4	6	8	10	12	14
2	16	18	17	23	19	12	17
4	11	15	10	12	15	11	19
6	14	11	10	11	12	10	17
8	13	10	10	10	12	16	15
10	15	10	10	10	10	10	10
12	13	10	10	10	10	10	10
14	12	10	10	10	10	10	10

Fig. 3.9: *Histograma de vencedores.*

Para salientar as bordas da *U-matrix* foram retirados os efeitos dos neurônios inativos seguindo a estratégia apresentada em Costa e Andrade Netto (1999, 2001). Esses neurônios, que apresentam um número de vencimentos inferior a um grau de ativação mínimo  $\varphi$ , são deslocados na direção dos neurônios ativos mais próximos. As figuras 3.12 e 3.13 mostram,

Fig. 3.10: *U-matrix na forma planar.*Fig. 3.11: *U-matrix na forma de superfície.*

respectivamente, a configuração dos neurônios e a *U-matrix* na forma de superfície depois desse procedimento.

Fig. 3.12: *Configuração dos neurônios do SOM após a eliminação do efeito dos neurônios inativos.*Fig. 3.13: *U-matrix na forma de superfície após a eliminação do efeito dos neurônios inativos.*

Aplicando o algoritmo para determinar a imagem de marcadores da *U-matrix* obtém-se o gráfico da figura 3.14, onde é possível visualizar o número de regiões conectadas ( $N_{rc}^k$ ) para cada valor de limiar da *U-matrix*. De acordo com o algoritmo SL-SOM, a imagem de marcadores corresponde à imagem binária obtida da limiarização da *U-matrix* com o valor inicial da seqüência mais estável de  $N_{rc}$ , que no caso foi  $k = 81$ , correspondendo a 2 (dois) agrupa-

mentos. A figura 3.15 mostra a imagem de marcadores para a  $U$ -matrix considerando o valor de limiar  $k = 81$ , e a figura 3.16 mostra as linhas de partição sobre a  $U$ -matrix encontradas pelo algoritmo *watershed*.

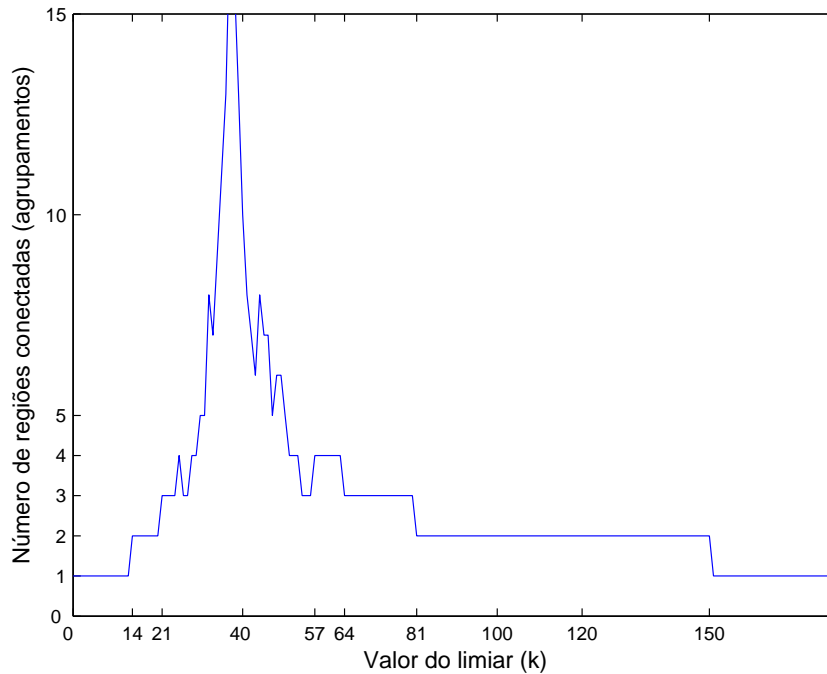


Fig. 3.14: Gráfico do valor de limiar ( $k$ ) da  $U$ -matrix versus o número de regiões conectadas (agrupamentos).

Após a rotulação da  $U$ -matrix segmentada (figura 3.17), os códigos das regiões foram passados aos neurônios correspondentes. A figura 3.18 mostra o SOM rotulado em duas grandes classes, as quais foram denominadas de C1 e C2. A partição C1 corresponde às classes de dados 1 e 2, e a C2 corresponde às classes de dados 3 e 4.

Dando continuidade ao processo de geração dinâmica da árvore de mapas (algoritmo TS-SL-SOM), dois novos SOMs (submapas) são treinados utilizando os conjuntos de dados correspondentes às classes C1 e C2: o submapa 1, filho da região 1 (classe C1), e o submapa 2, filho da região 2 (classe C2).

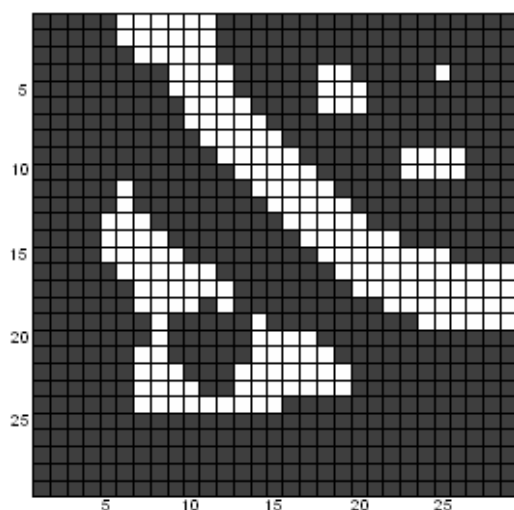


Fig. 3.15: *Imagem de marcadores para a U-matrix correspondente ao limiar  $k = 81$ .*

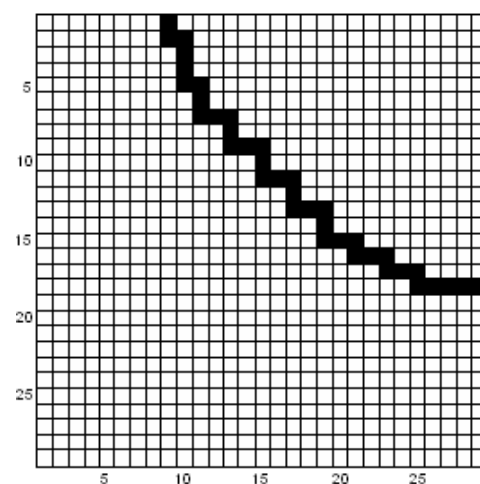


Fig. 3.16: *Partição encontrada pelo algoritmo watershed.*

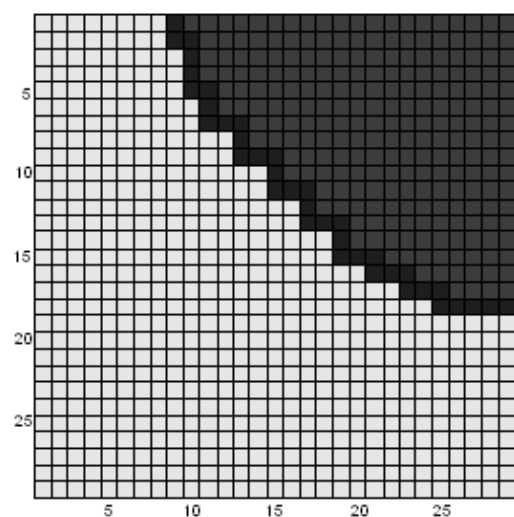


Fig. 3.17: *U-matrix rotulada após a segmentação pela watershed.*

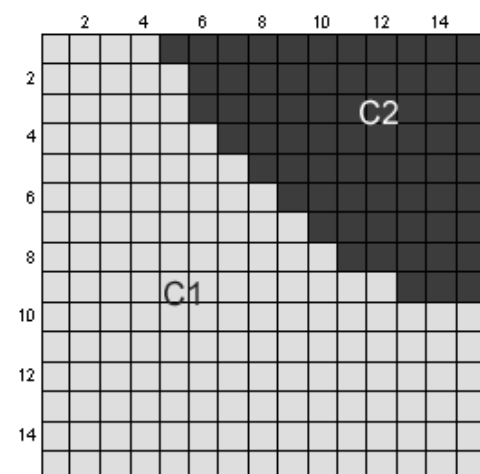


Fig. 3.18: *SOM rotulado.*

### Submapa 1

O submapa 1 foi treinado com os dados das classes 1 e 2 utilizando a mesma topologia e função de vizinhança que as do mapa pai. O seu tamanho foi 10x10, o raio da vizinhança inicial igual a 8 e o número máximo de iterações, executando o treinamento em lote, foi 500. A figura 3.19 mostra a configuração dos neurônios do submapa 1 (juntamente com os dados das classes 1 e 2) depois do treinamento e da eliminação do efeito dos neurônios inativos. O

histograma de vencedores é apresentado na figura 3.20, a  $U$ -matrix na forma planar e na forma de superfície são apresentadas nas figuras 3.21 e 3.22, respectivamente.

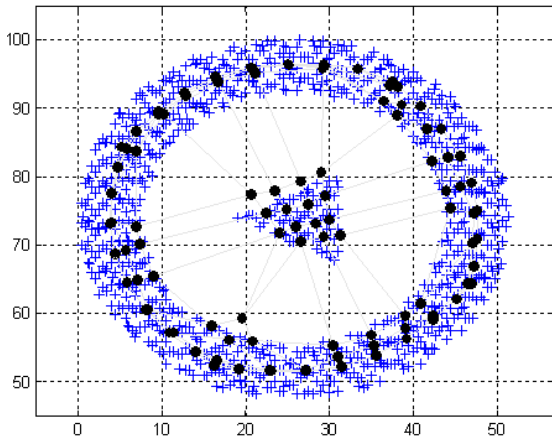


Fig. 3.19: *Neurônios do submapa 1 após o treinamento e eliminação dos neurônios inativos.*

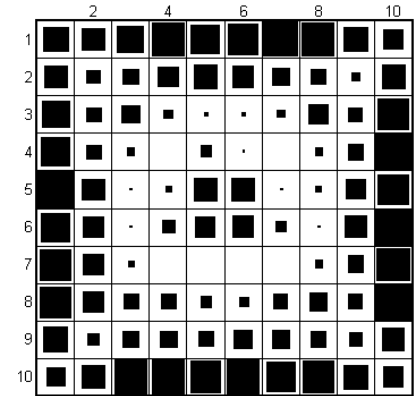


Fig. 3.20: *Histograma de vencedores do submapa 1.*

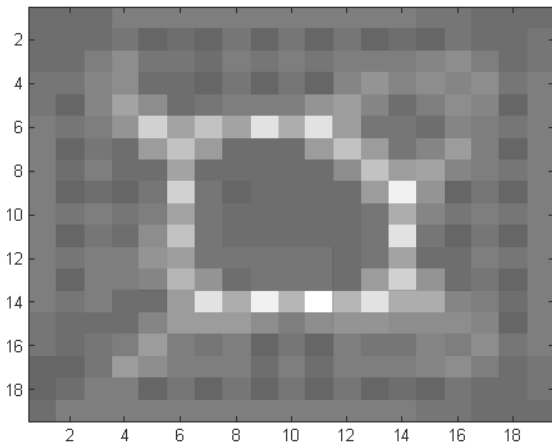


Fig. 3.21:  *$U$ -matrix do submapa 1 na forma planar.*

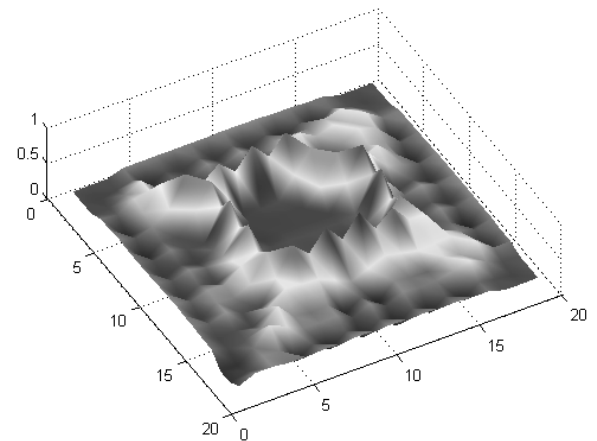


Fig. 3.22:  *$U$ -matrix do submapa 1 na forma de superfície.*

Para determinar a imagem de marcadores da  $U$ -matrix, o gráfico do número de regiões conectadas ( $N_{rc}^k$ ) para cada valor de limiar da  $U$ -matrix é obtido (figura 3.23). Conforme já descrito anteriormente, de acordo com o algoritmo SL-SOM, a imagem de marcadores corresponde à imagem binária obtida da limiarização da  $U$ -matrix com o valor inicial da seqüência mais estável de  $N_{rc}$ , que neste caso foi  $k = 34$ , correspondendo a 2 (dois) agrupamentos. A figura 3.24 mostra a imagem de marcadores encontrada para a  $U$ -matrix do submapa 1 e a

figura 3.25 mostra a partição encontrada pelo algoritmo *watershed*.

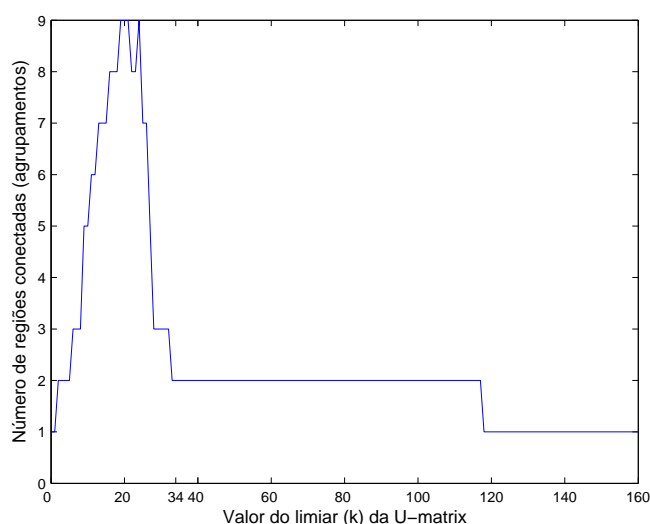


Fig. 3.23: Número de regiões conectadas versus limiar da *U-matrix* - submapa 1.

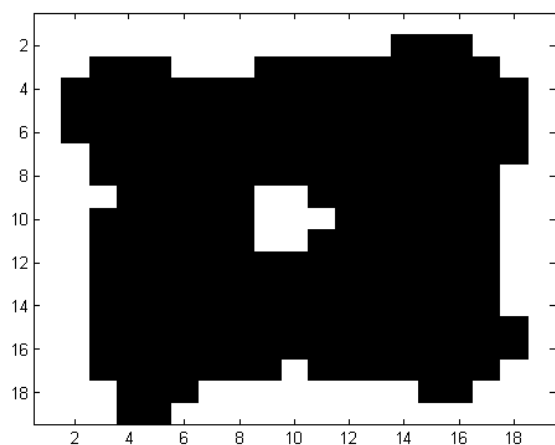


Fig. 3.24: Imagem de marcadores correspondente ao valor de limiar  $k = 34$  - submapa 1.

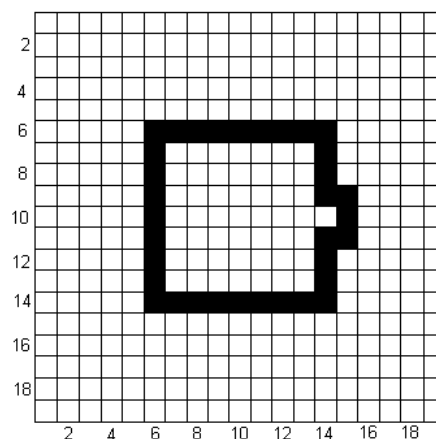


Fig. 3.25: Partição da *U-matrix* encontrada pelo algoritmo *watershed* - submapa 1.

Após a rotulação da *U-matrix* segmentada (figura 3.26), os códigos das regiões foram copiados para os neurônios correspondentes. A figura 3.27 mostra o submapa 1 particionado em duas classes, C1.1 e C1.2, correspondendo às classes reais de dados 1 e 2, respectivamente.

A figura 3.28 mostra os conjuntos de dados 1 e 2 classificados de acordo com a segmentação obtida para o submapa 1. A classificação final dos dados foi efetuada através do método dos vizinhos mais próximos comparando os vetores de pesos dos neurônios (rotulados) com os



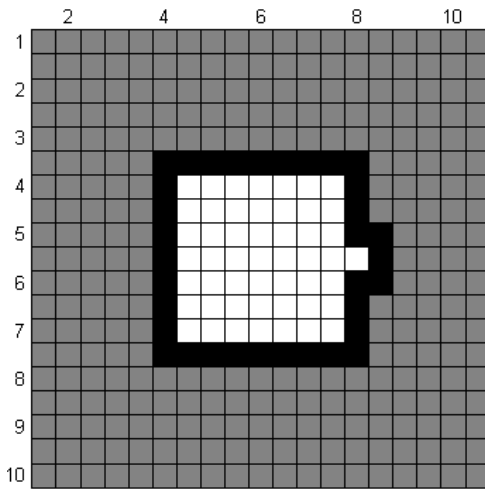


Fig. 3.26: *U-matrix segmentada - submapa 1.*

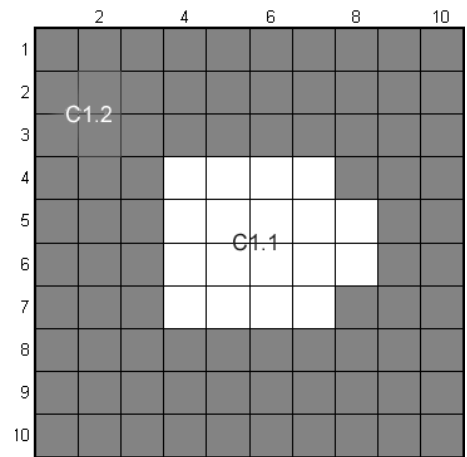


Fig. 3.27: *Submapa 1 rotulado.*

vetores de dados. Como pode ser visto, a classificação dos dados das classes 1 e 2 foi feita corretamente.

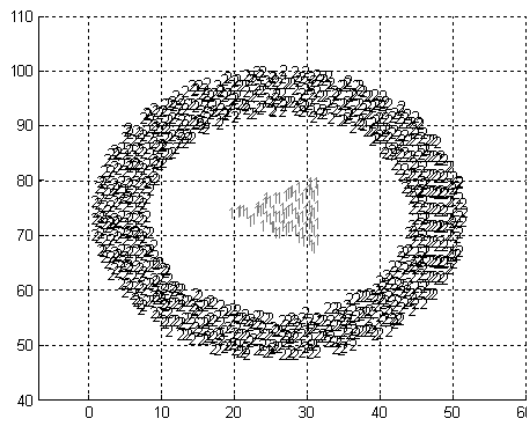


Fig. 3.28: *Conjuntos de dados 1 e 2 classificados.*

Outros dois submapas foram construídos a partir do submapa 1: o submapa 1.1, filho da região 1 (classe C1.1), o qual foi treinado com os dados da classe 1 e o submapa 1.2, filho da região 2 (classe C1.2), que foi treinado com os dados da classe 2. No entanto, como não houve detecção de mais de uma região ou classe, os dois submapas foram eliminados. De acordo com Costa e Andrade Netto (1999, 2001), quando há eliminação de um submapa em determinado nível da árvore, a região de neurônios que deu origem ao submapa é suficiente para representar o subconjunto de dados analisado.

## Submapa 2

O submapa 2, filho da região 2 (classe C2), foi treinado com os dados originais das classes 3 e 4. O tamanho do submapa foi 7x7 e a topologia foi a mesma do mapa pai. O raio de vizinhança inicial foi 5 e o número máximo de iterações apresentando os padrões em lote foi 500. A figura 3.29 mostra a configuração dos neurônios do submapa 2 (juntamente com os dados das classes 3 e 4) depois do treinamento e da eliminação do efeito dos neurônios inativos. O histograma de vencedores é apresentado na figura 3.30, a  $U$ -matrix na forma planar e na forma de superfície são apresentadas nas figuras 3.31 e 3.32, respectivamente.

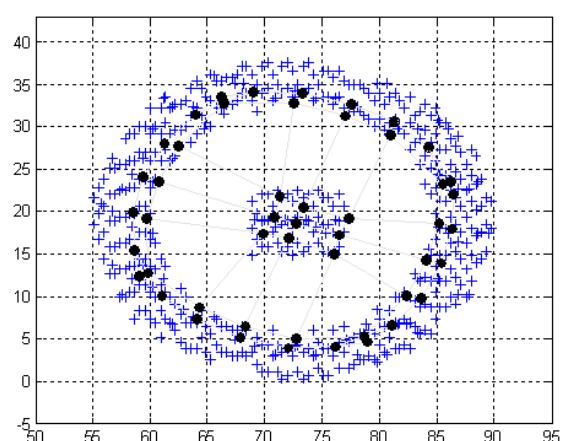


Fig. 3.29: *Neurônios do submapa 2 após o treinamento e eliminação dos neurônios inativos.*

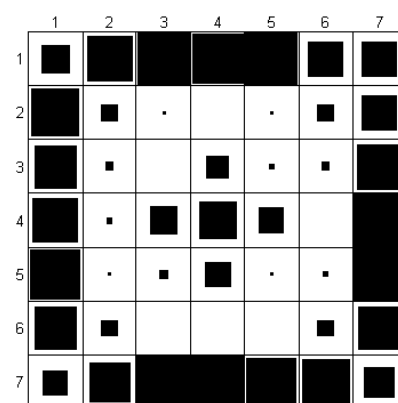


Fig. 3.30: *Histograma de vencedores do submapa 2.*

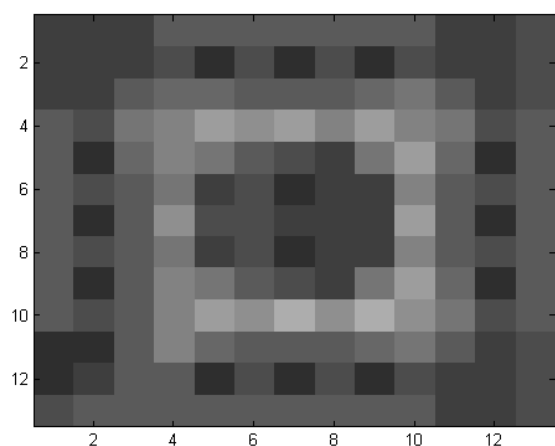


Fig. 3.31: *U-matrix do submapa 2 na forma planar.*

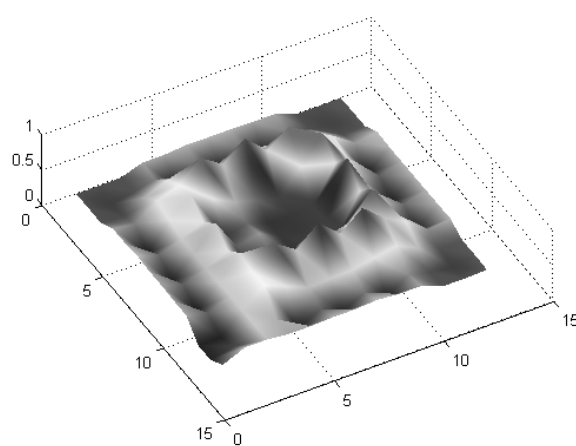


Fig. 3.32: *U-matrix do submapa 2 na forma de superfície.*

O gráfico do número de regiões conectadas ( $N_{rc}^k$ ) para cada valor de limiar da  $U$ -matrix (apresentado na figura 3.33) foi obtido com o objetivo de determinar a imagem de marcadores da  $U$ -matrix. Para a  $U$ -matrix do submapa 2, o valor inicial da seqüência mais estável do número de regiões conectadas,  $N_{rc}$ , foi  $k = 41$ , correspondendo a 2 (dois) agrupamentos. A figura 3.34 mostra a imagem de marcadores encontrada e a figura 3.35 mostra a partição efetuada pelo algoritmo *watershed*.

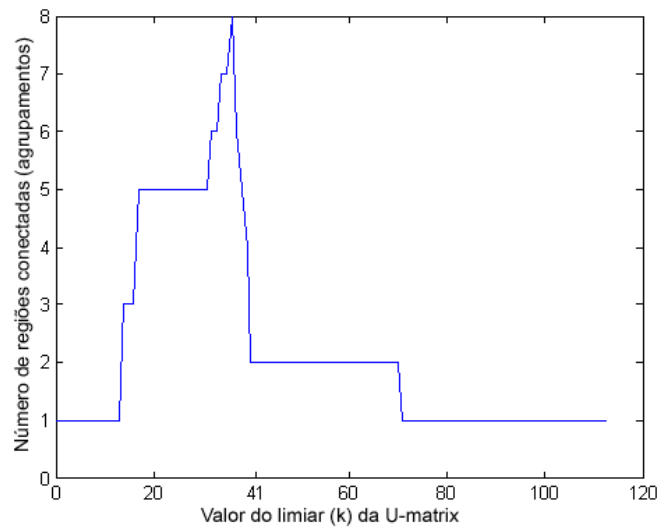


Fig. 3.33: Número de regiões conectadas versus limiar da  $U$ -matrix - submapa 2.

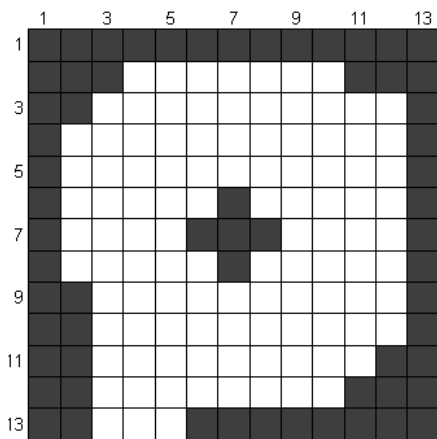


Fig. 3.34: Imagem de marcadores correspondente ao valor de limiar  $k = 41$  - submapa 2.

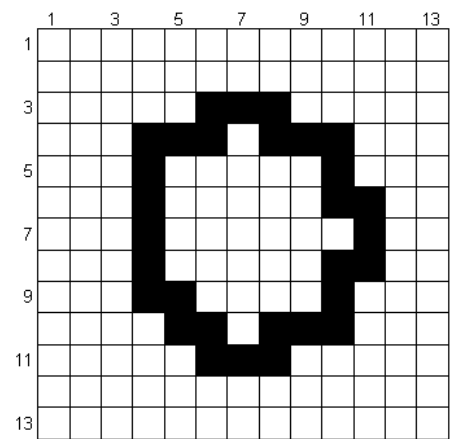


Fig. 3.35: Partição da  $U$ -matrix encontrada pelo algoritmo *watershed* - submapa 2.

Depois da rotulação da  $U$ -matrix segmentada (figura 3.36), os códigos das regiões foram

copiados para os neurônios correspondentes. A figura 3.37 mostra o submapa 2 rotulado em duas classes, C2.1 e C2.2, correspondendo às classes reais de dados 3 e 4, respectivamente.

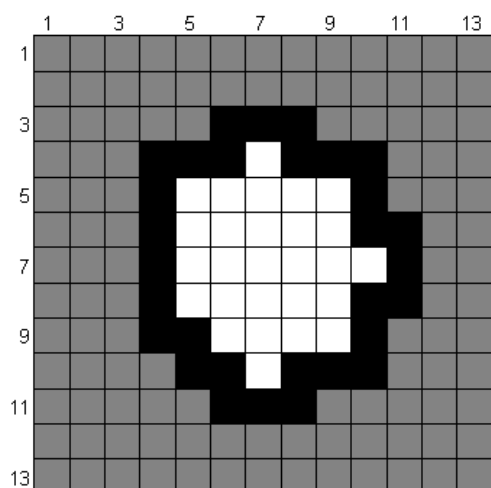


Fig. 3.36: *U-matrix segmentada - submapa 2.*

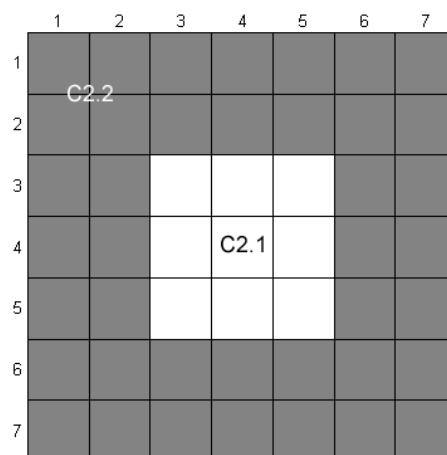


Fig. 3.37: *Submapa 2 rotulado.*

A figura 3.38 mostra os conjuntos de dados 3 e 4 classificados de acordo com a segmentação obtida para o submapa 2. A classificação final dos dados foi efetuada através do método dos vizinhos mais próximos. Como pode ser visto, a classificação dos dados das classes 3 e 4 também foi realizada corretamente.

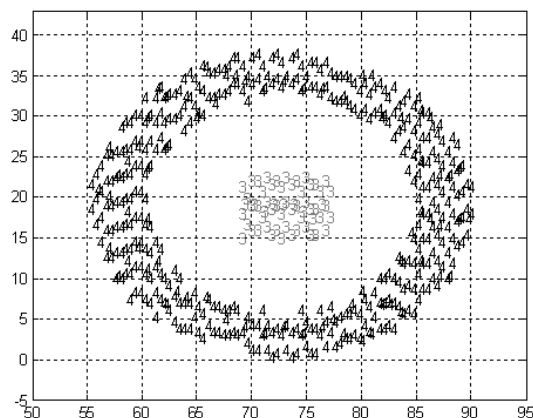


Fig. 3.38: *Conjuntos de dados 3 e 4 classificados.*

Na seqüência, mais dois submapas são construídos a partir do submapa 2: o submapa 2.1, filho da região 1 (classe C2.1), o qual foi treinado com os dados da classe 3, e o submapa 2.2, filho da região 2 (classe C2.2), que foi treinado com os dados da classe 4. No entanto, como

não houve a detecção de mais de um agrupamento ou classe nesses submapas, os mesmos foram descartados.

A figura 3.39 resume todo o processo de análise de agrupamentos efetuado sobre o conjunto de dados da figura 3.6, mostrando a árvore de mapas gerada pelo método proposto por Costa e Andrade Netto (1999, 2001). Foram construídos e treinados um total de 7 (sete) SOMs, sendo que os quatro últimos (submapas) foram eliminados por não terem apresentado mais de um agrupamento. O método detectou e classificou corretamente os 4 agrupamentos de dados existentes no conjunto de dados utilizado no experimento.

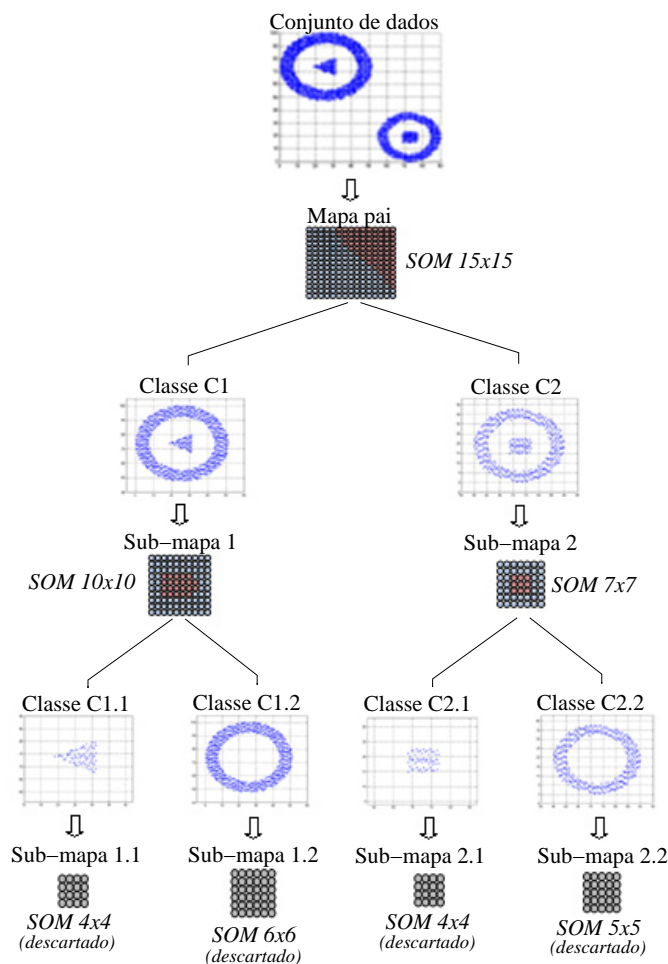


Fig. 3.39: Árvore de mapas gerada durante a aplicação do método proposto por Costa e Andrade Netto (1999, 2001) sobre o conjunto de dados da figura 3.6.

Outros resultados experimentais mostrando a aplicação do método em conjuntos de dados com densidades e geometrias variadas podem ser vistos em Costa (1999).

## 3.8 Literatura e modelos derivados do SOM

Na literatura é possível encontrar diversas abordagens variantes e modelos de redes neurais derivados do SOM.

Blackmore e Miikkulainen (1993) desenvolveram o método *Incremental Grid Growing* (IGG) com o objetivo de solucionar dois problemas do SOM convencional: (a) a necessidade de definir o tamanho do mapa a priori; e (b) detectar a borda dos agrupamentos. O IGG geralmente inicia com um mapa de neurônios de tamanho 2x2 e adiciona novos neurônios durante o processo de treinamento em posições do mapa onde há grande concentração de padrões. Desse modo, o IGG procura crescer em tamanho com o objetivo de representar melhor os dados. Conexões entre neurônios vizinhos são adicionadas ou eliminadas em função da distância entre os pesos sinápticos. Ao final do processo de treinamento, o efeito de adição de neurônios e conexões laterais, assim como o efeito de eliminação dessas últimas, podem gerar mapas completamente separados. Um dos grandes problemas desse método é fornecer um grande conjunto de parâmetros (além dos parâmetros básicos do SOM) antes do início do treinamento. O método é bastante susceptível à escolha dos seus parâmetros.

Seguindo o mesmo princípio do IGG, Fritzke (1991) propôs um modelo construtivo, o *Growing Cell Structure* (GCS), que consiste em um arranjo em 2D composto de neurônios (nós) conectados entre si em forma de triângulos. Inicialmente, três neurônios estão presentes e a atualização deles ocorre para o neurônio vencedor e seus vizinhos. O modelo insere novos neurônios na vizinhança daqueles cujos erros de representação são grandes, conectando-o ao arranjo. Diferentemente do IGG, estes novos neurônios não são adicionados no perímetro da rede e sim dentro do mapa. As deficiências do GCS incluem principalmente a escolha dos vários parâmetros. Kohonen *et al.* (1996) mostram que o GCS é muito mais propenso a variações nas escolhas dos parâmetros iniciais que o SOM convencional.

O uso de vários mapas de Kohonen com estrutura hierárquica foi proposto por Miikkulainen (1993). Nesse modelo, os mapas não são independentes entre si, mas possuem uma estrutura hierárquica rígida na forma de uma pirâmide. A idéia é que agrupamentos diferentes sejam mapeados em mapas diferentes no mais baixo nível da hierarquia. Um dos problemas do método é que o tamanho da pirâmide deve ser definido a priori, o que significa que não há nenhum

mecanismo de crescimento dinâmico de novos mapas. Além disso, todos os mapas em um determinado nível da pirâmide possuem o mesmo tamanho, mesmo que estejam representando quantidades muito diferentes de padrões.

O modelo TS-SOM, proposto por Koikkalein (1994), apresenta uma estratégia semelhante à do modelo de Miikulainen (1993). Com estrutura também piramidal e estática, onde cada neurônio possui quatro outros neurônios filhos, a rede é treinada seqüencialmente por níveis, do topo até a base da pirâmide. Pode-se usar pesos de um nível para inicializar os pesos do nível inferior, com o objetivo de reduzir o tempo de treinamento. Os resultados em simulações com dados uniformemente distribuídos foram semelhantes ao resultado do SOM convencional.

Castro e Von Zuben (1999) propuseram um algoritmo eficiente de poda, o PSOM (*Prunning SOM*), o qual procura remover neurônios pouco representativos (seguindo um critério estabelecido) e reiniciar o treinamento tomando como ponto de partida os parâmetros anteriores ao processo de poda. Entretanto, o algoritmo foi definido apenas para mapas unidimensionais e não é diretamente generalizável para dimensões maiores.

Muitos outros modelos derivados do SOM têm sido propostos na literatura (Wan e Fraser, 1994; Choi e Park, 1994; Alahakoon *et al.*, 2000; Kegang *et al.*, 2007; Ocsa *et al.*, 2007). A maioria deles busca contornar algumas dificuldades apresentadas pelo modelo convencional da rede de Kohonen, principalmente em relação à rigidez do arranjo de neurônios. No entanto, conforme ressalta Zuchini (2003), deve-se observar que as aparentes vantagens de métodos dinâmicos ou construtivos do SOM podem facilmente tornar-se seus principais problemas. Ao analisar essas versões construtivas do SOM nota-se que a maioria delas apresenta uma complexidade computacional superior ao do algoritmo original do SOM. Portanto, ao aplicar tais modelos é importante verificar se as melhorias e benefícios apresentados pelos mesmos compensam seus custos computacionais.

# Capítulo 4

## Classificação baseada em SOMs segmentados por morfologia matemática

### 4.1 Introdução

Conforme descrito na seção 1.2, esse trabalho tem como objetivo principal propor métodos de classificação não-supervisionada de imagens de sensoriamento remoto que explorem as características e propriedades do SOM.

A estratégia básica empregada pelos métodos de classificação propostos é realizar a análise de agrupamentos da imagem através do conjunto de protótipos de um SOM treinado, ao invés de trabalhar diretamente com os padrões originais da cena. Assim sendo, buscando contornar as dificuldades apresentadas pelo SOM no que se refere à interpretação do seu mapa de saída (mencionadas na seção 3.7), as metodologias de classificação de imagens apresentadas neste trabalho realizam o particionamento do mapa de neurônios da rede de Kohonen através de métodos de segmentação automática, diferenciando-se, portanto, da maioria das abordagens que empregam o SOM na análise de imagens de sensoriamento remoto, nas quais a determinação das fronteiras entre partes não relacionadas no espaço de saída da rede é realizada freqüentemente com a participação do usuário (que examina a representação padrão do mapa de neurônios). Além de preservar vantagens importantes inerentes às técnicas de classificação não-supervisionada (descritas na seção 2.5.2), a estratégia de particionar o SOM de maneira



automática contribui para a necessidade de exploração em grande escala de um volume cada vez maior de imagens coletadas por sensores remotos, uma vez que busca minimizar a intervenção do usuário no processo de classificação das cenas.

Conforme descrito na seção 3.7.1, na literatura existem algumas propostas de algoritmos para segmentar de maneira automática os neurônios de um SOM treinado. Todavia, o presente capítulo apresenta uma metodologia de classificação de imagens de sensoriamento remoto que procura testar a aplicabilidade do método de segmentação do SOM proposto por Costa e Andrade Netto (1999, 2001). Baseado em técnicas de morfologia matemática, este método proporciona vantagens significativas à metodologia de classificação proposta em comparação com métodos convencionais de classificação não-supervisionada.

Ao descrever a metodologia de classificação, o presente capítulo também apresenta e analisa procedimentos importantes relacionados com o processo de amostragem das imagens e com a configuração dos parâmetros de treinamento do SOM. Resultados experimentais mostrando a aplicação da metodologia de classificação proposta sobre uma imagem teste também são apresentados e uma análise comparativa com o algoritmo K-médias é realizada.

## 4.2 Metodologia de classificação de imagens de sensoriamento remoto usando SOMs segmentados por morfologia matemática

A figura 4.1 ilustra de maneira simplificada a metodologia de classificação de imagens de sensoriamento remoto proposta neste capítulo e a seguir são apresentadas as suas etapas principais:

1. *Amostragem da imagem*: amostras na forma de janelas de pixels são coletadas de maneira uniforme sobre toda a região da imagem e fornecidas como padrões de entrada para o SOM;
2. *Treinamento do SOM*: o treinamento não-supervisionado do SOM é realizado e um mapa de agrupamentos da imagem composto pelos neurônios da rede (protótipos) é obtido;

3. *Segmentação do SOM*: o SOM treinado é particionado e rotulado utilizando o método de segmentação automática do SOM proposto por Costa e Andrade Netto (1999, 2001);
4. *Classificação da imagem*: os protótipos do SOM rotulados no passo 3 são utilizados como referência para classificar todos os pixels da imagem original.

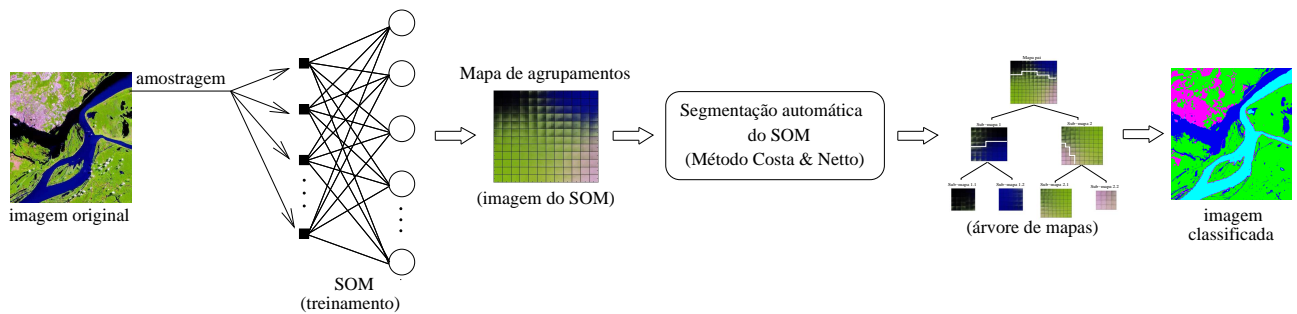


Fig. 4.1: Ilustração da metodologia de classificação proposta usando SOMs segmentados por morfologia matemática.

Cada uma das etapas da metodologia de classificação proposta é explicada com maiores detalhes nas próximas subseções. Vale destacar que os procedimentos e considerações apresentadas particularmente nas etapas de amostragem da imagem, treinamento do SOM e classificação final da imagem também são empregados nos métodos de classificação propostos nos próximos capítulos deste trabalho.

### 4.2.1 Amostragem da imagem

O primeiro passo da metodologia de classificação proposta consiste em coletar um conjunto de amostras da imagem para realizar o treinamento do SOM.

Com o aumento significativo da resolução espacial das imagens de satélite nos últimos anos, as abordagens de classificação pixel a pixel não têm mais alcançado resultados plenamente satisfatórios no reconhecimento de padrões de imagens coletadas por sensores modernos. Devido à heterogeneidade espectral existente dentro de uma mesma classe, amostras na forma de pixels isolados não são capazes de capturar a maioria dos padrões de cobertura terrestre, especialmente no caso de imagens que apresentam melhores resoluções (Blaschke *et al.*, 2007; Gonçalves *et al.*, 2008a). Diante disso, observando a necessidade de explorar novas técnicas que não

analisam apenas a assinatura espectral de pixels individuais, o presente trabalho realiza o processo de amostragem das imagens através de *janelas de pixels*. A idéia é incorporar na análise de agrupamentos da imagem informações de vizinhança (ou de contexto) dos pixels. Relacionamentos entre pixels vizinhos que sejam estáveis, consistentes e possíveis de serem previstos são utilizados para melhorar a qualidade da classificação resultante.

A figura 4.2 mostra uma imagem gerada artificialmente com quatro segmentos de classes de cobertura terrestre retirados de imagens do satélite IKONOS (com resolução espacial de 4 metros). As quatro classes apresentam padrões de textura bem diferenciados e fáceis de serem discriminados visualmente.

Com o objetivo de destacar as diferenças no processo de detecção de agrupamentos da imagem empregando uma amostragem pixel a pixel e uma amostragem através de janelas de pixels, dois SOMs compostos de 100 neurônios (arranjados numa grade retangular 10x10) foram treinados: um com amostras na forma de pixels isolados e o outro com amostras na forma de janelas 5x5 de pixels. A inicialização dos pesos foi linear, a apresentação dos padrões (amostras) em lote e o número de iterações foi 500 para os dois mapas. Os dois SOMs foram treinados com a mesma quantidade de amostras. O tempo de treinamento sob o Matlab em uma máquina com processador AMD Athlon XP 2600+ 1,91GHZ e 1GB de RAM foi de 28 segundos para o SOM treinado com janelas de pixels e de 13 segundos para o SOM treinado com pixels individuais.

A figura 4.3 mostra a *U-matrix* do SOM treinado com janelas de pixels, onde pode-se observar claramente quatro regiões ou "vales" bem definidos, que correspondem a quatro agrupamentos de dados. A figura 4.4 mostra a imagem dos protótipos do SOM<sup>1</sup> correspondente, na qual também pode-se visualizar claramente quatro agrupamentos que correspondem exatamente aos quatro "vales" da *U-matrix* da figura 4.3. É fácil perceber também que os quatro agrupamentos de protótipos na imagem do SOM treinado (figura 4.4) correspondem às quatro classes de cobertura terrestre da imagem original.

Por outro lado, a *U-matrix* calculada a partir do SOM treinado com pixels isolados (na figura 4.5) não apresenta depressões ou "vales" tão bem definidos, o que dificulta uma detecção automática de agrupamentos a partir da mesma. Além disso, se considerarmos a imagem dos

---

<sup>1</sup>O processo de construção da imagem dos protótipos do SOM é descrito na próxima subseção.



Fig. 4.2: Imagem construída com quatro classes de cobertura terrestre capturadas pelo satélite IKONOS (com resolução espacial de 4 metros).

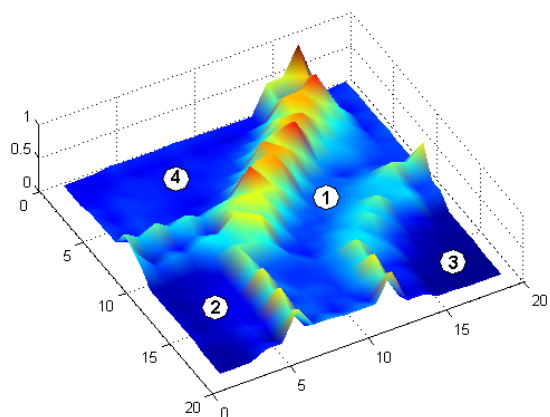


Fig. 4.3:  $U$ -matrix 3D do SOM treinado com janelas de pixels.

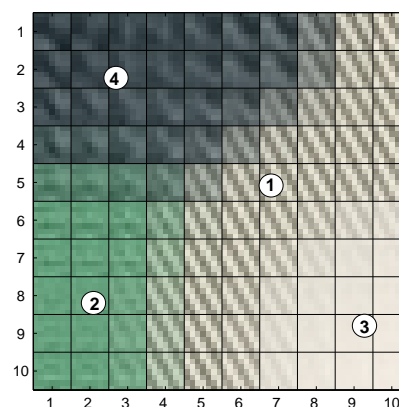


Fig. 4.4: Imagem dos protótipos do SOM treinado com janelas de pixels.

protótipos do SOM correspondente (apresentada na figura 4.6) e observarmos os protótipos que estão associados com aquelas regiões marcadas na imagem da  $U$ -matrix, verifica-se que eles não correspondem exatamente às classes de cobertura terrestre existentes na imagem original. Na realidade, é suficiente observar apenas a imagem do SOM treinado com pixels isolados (figura 4.6) para concluir que as classes existentes na imagem original não foram aprendidas pelo SOM usando esse tipo de amostragem. O SOM treinado com pixels isolados foi capaz de mapear todas as variações espectrais (cores) existentes na imagem, mas não conseguiu capturar padrões de cunho espacial (texturas), que são fundamentais para caracterizar e diferenciar as

quatro classes consideradas.

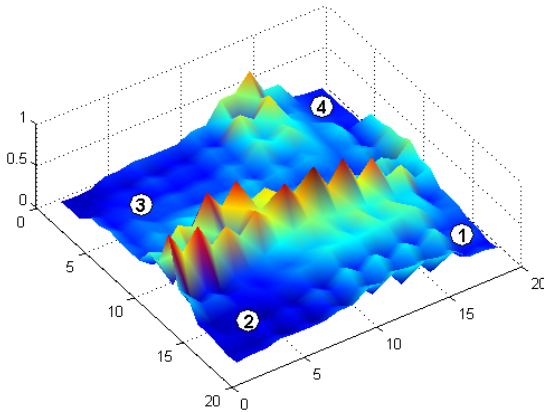


Fig. 4.5: *U-matrix 3D do SOM treinado com pixels isolados.*

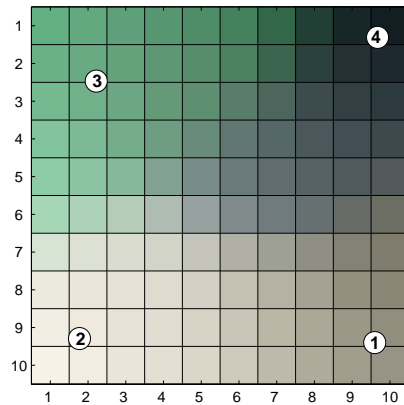


Fig. 4.6: *Imagem dos protótipos do SOM treinado com pixels isolados.*

As figuras 4.7 e 4.8 mostram, respectivamente, os resultados da classificação da imagem da figura 4.2 usando os protótipos do SOM treinado com janelas de pixels e os protótipos do SOM treinado com pixels isolados. Visualmente, nota-se uma diferença significativa nos resultados. A classificação realizada através dos protótipos do SOM treinado com janelas de pixels apresentou uma discriminação bastante satisfatória das quatro classes, com alguns erros apenas nas fronteiras ou bordas de transição. Já a imagem resultante da classificação pelo SOM treinado com pixels isolados apresentou um alto grau de confusão entre elas, uma vez que foi capaz de discriminar apenas atributos espectrais (tonalidades) diferentes, que na realidade pertencem a mais de uma classe.

Embora a estratégia de utilizar informações de pixels vizinhos apresente obviamente um maior custo computacional do que uma abordagem pixel a pixel, os resultados do experimento apresentado aqui, e de outros trabalhos existentes na literatura, mostram claramente que esse esforço é justificado por um aumento proporcional na acurácia da classificação (Gonçalves, 1997; Zhang *et al.*, 2003; Magnussen *et al.*, 2004). Na tentativa ainda de aperfeiçoar o desempenho das técnicas de classificação, procurando incorporar informações de vizinhança sem causar um aumento significativo no volume de processamento, algumas metodologias de classificação, como aquela apresentada em Bandyopadhyay (2005), têm proposto o uso de descritores estatísticos (como média e desvio-padrão) para representar as informações de vizinhança ao redor de um

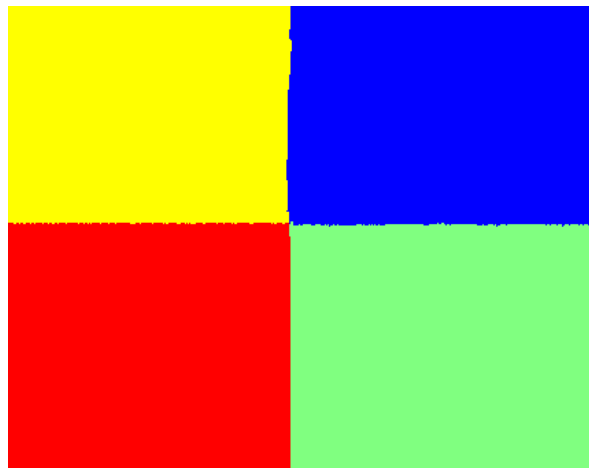


Fig. 4.7: Imagem da figura 4.2 classificada através dos protótipos do SOM treinado com janelas de pixels.

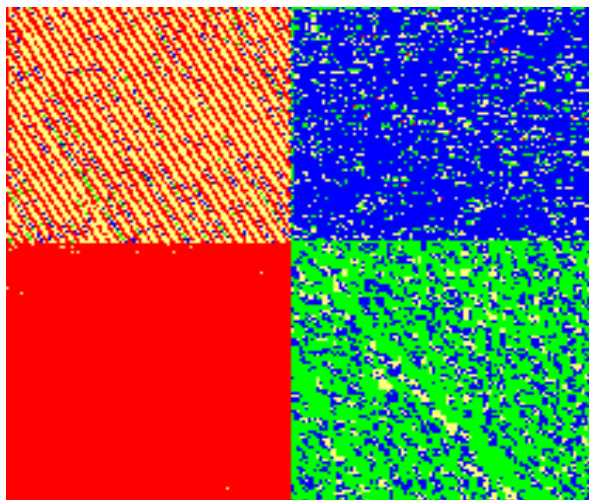


Fig. 4.8: Imagem da figura 4.2 classificada através dos protótipos do SOM treinado com pixels isolados.

pixel, evitando assim a utilização dos valores numéricos de todos os pixels que fazem parte de uma janela amostral. Embora essa abordagem diminua o custo computacional da análise da imagem, é importante observar que diante da diversidade de padrões de cobertura terrestre existentes, medidas estatísticas de primeira ordem, como média, variância e desvio-padrão, podem não ser suficientes para descrever ou representar de maneira adequada as informações contidas em uma amostra na forma de janelas de pixels. Conforme relatado em Rennó (1995), esse problema ocorre porque tais medidas não levam em conta o posicionamento espacial relativo

dos pixels, e portanto, uma janela amostral que contenha uma mesma quantidade de diferentes níveis de cinza que outra, diferenciando-se apenas em termos do posicionamento dos mesmos, irá apresentar os mesmos valores para as medidas estatísticas calculadas.

É certo que, além de medidas estatísticas básicas, existe na literatura uma série de outras técnicas que podem ser utilizadas para descrever ou representar a organização espacial dos níveis de cinza de um imagem ou de uma janela de pixels, como as medidas de Haralick (mencionadas na subseção 2.5.4.2), filtragem de Gabor (Jain e Farrokhnia, 1991) e decomposição em *wavelets* (Vower *et al.*, 1999). No entanto, a complexidade computacional desses métodos deve ser considerada, principalmente quando aplicados em grandes volumes de dados como as imagens de sensoriamento remoto.

Sendo assim, neste trabalho, optou-se por processar os valores de níveis de cinza de todos os pixels das janelas amostrais. Tal abordagem, além de não exigir nenhum cálculo explícito de medidas de textura ou de organização espacial dos pixels, proporciona algumas vantagens que serão apresentadas mais adiante nos experimentos e em outras abordagens de classificação apresentadas no trabalho.

Dois pontos importantes que decorrem do processo de amostragem por meio de janelas de pixels são a forma de extração das janelas amostrais e o tamanho das mesmas. Conforme mostra a figura 4.9, com o objetivo de garantir a captura de todos os padrões existentes em uma cena, inclusive daqueles com baixa probabilidade de ocorrência, a metodologia de classificação proposta coleta as janelas amostrais de maneira uniforme cobrindo toda a região da imagem, sem sobreposições e em intervalos regulares de aproximadamente 10 pixels. Todas as amostras são quadradas e possuem o mesmo tamanho.

O espaçamento em torno de 10 pixels, empregado para extrair as janelas amostrais das imagens, foi definido experimentalmente. Intervalos maiores podem ser utilizados, porém quanto maior o espaçamento adotado entre as janelas amostrais, maior é o risco de exclusão de padrões de cobertura terrestre que aparecem em pequenas áreas na imagem com baixa probabilidade de ocorrência.

A figura 4.10 mostra um exemplo de situação em que um intervalo muito grande entre as janelas amostrais pode prejudicar a detecção das classes de cobertura terrestre de uma

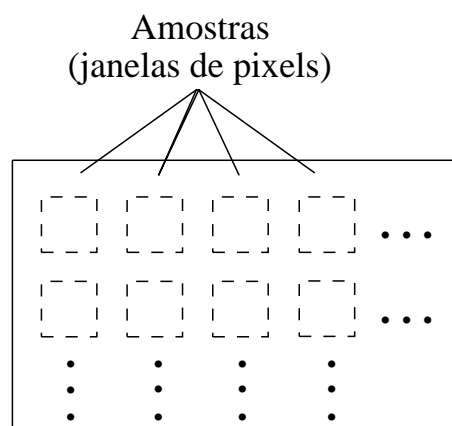


Fig. 4.9: *Processo de amostragem da imagem de entrada.*

imagem a partir de um SOM treinado. A figura mostra uma imagem do satélite Landsat/TM composta por diferentes classes de cobertura terrestre, incluindo a classe água (na cor preta), que aparece em pequenas áreas numa densidade menor que as dos outros padrões da imagem. Logo abaixo, a figura mostra também as imagens dos protótipos (neurônios) de três SOMs de tamanhos iguais (12x12), treinados a partir de três conjuntos diferentes de janelas amostrais. A diferença entre os conjuntos de amostras está justamente no espaçamento utilizado para coletar as janelas amostrais. A figura 4.10(b) mostra a imagem resultante de um SOM treinado a partir de um conjunto de janelas amostrais extraídas em intervalos (espaços) regulares de 10 pixels, onde é possível notar que há protótipos no canto superior esquerdo do mapa (de cor escura) correspondentes à classe água. Já as figuras 4.10(c) e (d) mostram imagens resultantes de SOMs treinados a partir de conjuntos de janelas amostrais extraídas em intervalos regulares de 20 e 30 pixels, respectivamente. Na imagem (c), ainda é possível visualizar pelo menos um protótipo (no canto superior esquerdo do mapa) com aparência espectral próxima ao do padrão de água. No entanto, na imagem (d), verifica-se claramente que o SOM não apresenta protótipos representantes dessa classe. Isso certamente ocorre devido ao espaçamento maior utilizado no processo de amostragem da imagem. O intervalo de 30 pixels de uma janela amostral à outra não possibilitou a inclusão de um número suficiente de amostras da classe água para treinar o SOM. Portanto, caso o SOM da figura 4.10(c) fosse utilizado para detectar todos os padrões existentes na imagem (a), a classe água não seria reconhecida.

Diante disso, pode-se observar ainda, que a definição do espaçamento entre as janelas amos-



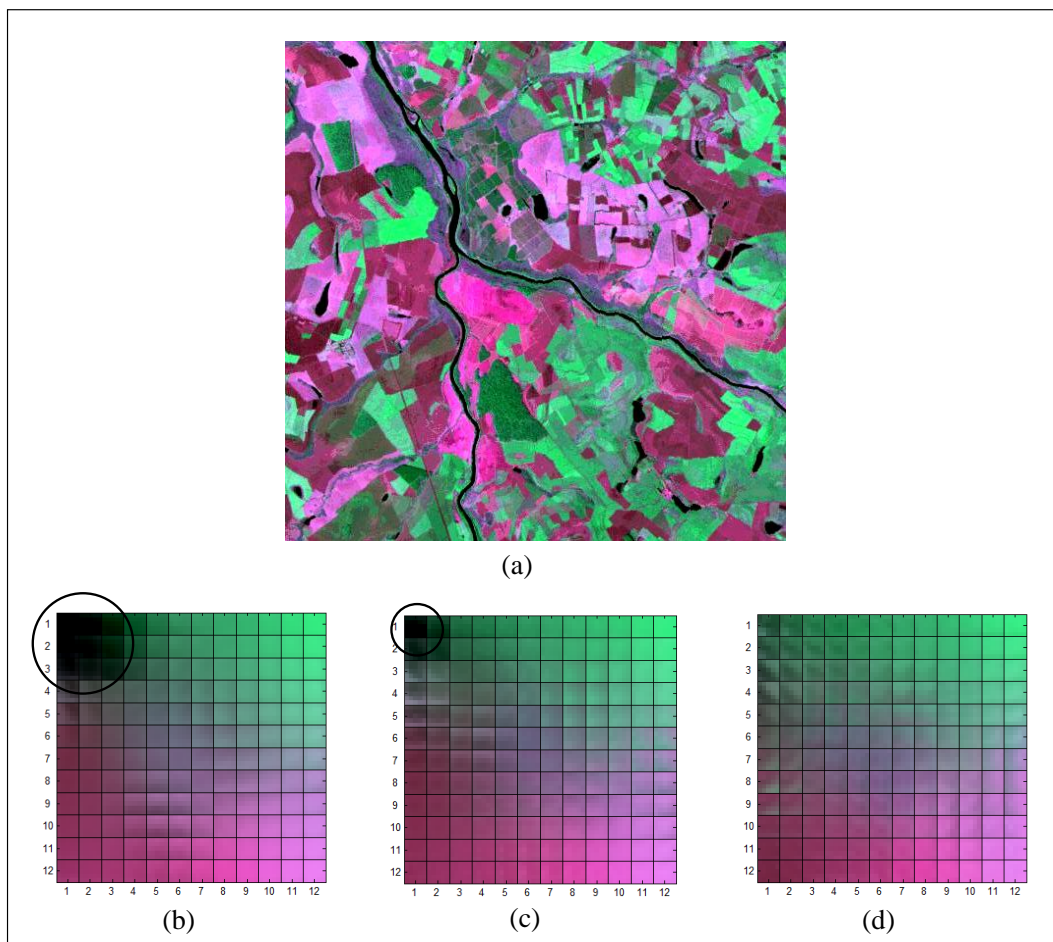


Fig. 4.10: *Efeito do espaçamento entre janelas amostrais no treinamento do SOM: (a) Imagem Landsat/TM. (b) Imagem de um SOM treinado com janelas amostrais coletadas em intervalos regulares de 10 pixels. Os protótipos correspondentes à classe água aparecem (circulados) no canto superior esquerdo do mapa. (c) Imagem de um SOM treinado com janelas amostrais coletadas em intervalos regulares de 20 pixels. Apenas um protótipo no mapa (circulado no canto superior esquerdo) corresponde à classe água. (d) Imagem de um SOM treinado com janelas amostrais coletadas em intervalos regulares de 30 pixels. O mapa não apresenta protótipos associados à classe água.*

trais da imagem pode depender também dos objetivos da classificação. Caso o usuário deseje uma classificação mais refinada ou detalhada da imagem, espaçamentos menores entre as janelas amostrais devem ser utilizados. Isso garante que uma porcentagem maior da área total da imagem será coberta pelo processo de amostragem, e portanto, classes com menor probabilidade de ocorrência na cena serão amostradas. Por outro lado, se o usuário pretende obter uma classificação mais superficial ou grosseira da cena, procurando detectar somente as classes

mais predominantes na imagem, espaçamentos maiores entre as janelas amostrais podem ser empregados.

Todavia, vale ressaltar também, que o custo computacional para realizar o treinamento do SOM será obviamente maior no caso de amostragens que tiverem um espaçamento menor entre as janelas amostrais, visto que, neste caso, um número maior de amostras será coletado a partir da imagem. De maneira contrária, quanto maior o espaçamento entre as janelas amostrais, menor será o tempo gasto para treinar os neurônios do SOM, uma vez que uma quantidade menor de amostras será extraída da imagem a ser classificada. Isso pode ser verificado retomando o exemplo da figura 4.10. O tempo gasto para treinar o SOM da figura 4.10(b) foi de 69 segundos, utilizando 2600 amostras, decorrentes do espaçamento de 10 pixels adotado para coletar as janelas amostrais. Já os treinamentos dos SOMs das figuras 4.10 (c) e (d), consumiram 21 e 15 segundos, usando 625 e 289 janelas amostrais, respectivamente. Como esperado, seus tempos de treinamento foram bem menores que o do SOM da figura 4.10(b), uma vez que utilizaram quantidades bem menores de amostras, determinadas em função do uso de espaçamentos maiores entre as janelas amostrais extraídas da imagem.

Ainda em relação ao processo de amostragem das imagens, verifica-se que a determinação do tamanho das janelas amostrais é outra tarefa difícil de ser automatizada. A escolha de um tamanho adequado para as janelas amostrais depende de alguns fatores importantes, entre eles, a resolução espacial e as características das classes de cobertura terrestre da imagem de entrada. Dependendo da imagem, janelas muito pequenas podem não capturar o padrão particular das classes, enquanto que janelas maiores podem incluir pixels de duas ou mais classes. Contudo, como será mostrado nos próximos capítulos deste trabalho, partindo de um tamanho não inferior a 5x5, o usuário possui alguma flexibilidade para definir a dimensão das janelas amostrais, o que permite descaracterizar essa tarefa como um ponto crítico do processo. Esse trabalho mostra que procedimentos para realizar a filtragem de amostras e de protótipos do SOM que incorporam mais de uma classe podem ser aplicados no processo de análise de agrupamentos das imagens. Além disso, também será possível verificar que protótipos que possuem mais de uma classe funcionam como bordas no mapa de neurônios do SOM, contribuindo para a separação dos agrupamentos.

### 4.2.2 Treinamento do SOM

Para treinar o SOM é necessário especificar alguns parâmetros que definem a estrutura do mapa e que controlam o treinamento propriamente dito. Com o objetivo de garantir um bom mapeamento dos padrões originais da imagem, a metodologia de classificação proposta define de modo particular os parâmetros da rede baseando-se nas considerações apresentadas na seção 3.4.1, em testes experimentais e em algumas peculiaridades da aplicação do SOM em imagens de sensores remotos. Todavia, outras estratégias também podem ser buscadas para se obter bons mapas (Kaski e Lagus, 1996).

A metodologia proposta utiliza os seguintes parâmetros para realizar o treinamento da rede:

- Inicialização dos pesos: linear,
- Apresentação dos padrões: em lote,
- Número de épocas de treinamento: 500,
- Tipo de função de vizinhança: gaussiana,
- Formato do arranjo de neurônios (topologia da rede): retangular.

Embora a inicialização aleatória seja a forma mais comum de inicializar os pesos do SOM, Kohonen (1997) defende a idéia da inicialização linear, a qual consiste em distribuir os neurônios de forma ordenada ao longo de um plano alinhado com os eixos das duas maiores variâncias no conjunto de dados. De acordo com Kohonen, a forma de inicialização linear permite que o algoritmo de aprendizagem da rede ingresse diretamente na fase de convergência, além de evitar torções indesejáveis no arranjo de neurônios.

A forma de apresentação dos padrões em lote torna o resultado do mapeamento do SOM insensível à seqüência de apresentação dos padrões. Conforme descrito na seção 3.4.1, isso pode ser feito somando-se a média das contribuições de todos os padrões de entrada para cada neurônio. Esse deslocamento médio tem conduzido a melhores resultados que o algoritmo convencional do SOM, no qual os neurônios são atualizados toda vez que um novo padrão de entrada é fornecido à rede.

Para determinar o número de épocas de treinamento, ou seja, o número de vezes que o conjunto inteiro de janelas amostrais da imagem é fornecido à rede, foram realizados experimentos com o SOM usando diferentes imagens de satélite. Testes experimentais realizados com imagens contendo em média 512x512 pixels, mostram que o mapa de saída da rede se torna estável antes das 500 épocas de treinamento.

Ao invés de usar a vizinhança do tipo "bolha", proposta no algoritmo original do SOM, a função de vizinhança do tipo gaussiana (descrita na seção 3.4.1) é utilizada aqui por evidenciar melhor os agrupamentos em termos das distâncias entre os neurônios (Kohonen, 1997).

No método proposto, a forma retangular foi escolhida para dispor o arranjo de neurônios da rede, não apenas por ser a mais comumente utilizada, mas também por ser mais adequada para gerar a imagem do mapa de neurônios da rede, o qual será descrito na próxima subseção.

De maneira similar ao processo de amostragem, onde o espaçamento entre as janelas amostrais depende das estruturas das classes e da precisão desejada para o reconhecimento dos padrões da imagem, o tamanho do mapa é um dos parâmetros livres do SOM que também depende particularmente da imagem de entrada e dos objetivos da classificação. Se o objetivo é detectar todos os padrões da imagem, inclusive aqueles com baixa probabilidade de ocorrência, mapas de tamanhos maiores devem ser empregados na análise, caso contrário, se o interesse se concentra apenas nos padrões mais predominantes da cena então um tamanho menor para o SOM pode ser utilizado.

Considerando novamente a imagem apresentada na figura 4.10(a), a figura 4.11 ilustra o efeito do tamanho do SOM no mapeamento dos padrões daquela cena. As figuras 4.11(a), (b) e (c) mostram imagens de SOMs de tamanhos 12x12, 9x9 e 6x6, respectivamente. Os três mapas foram treinados com a mesma quantidade de janelas amostrais coletadas a partir da imagem da figura 4.10(a). Visualizando os SOMs das figuras 4.11(a) e (b), nota-se que a classe água, que aparece na cor preta na imagem de entrada com uma baixa probabilidade de ocorrência, está representada por protótipos no canto superior esquerdo dos dois mapas. No entanto, no SOM da figura 4.11(c), a classe água não foi possível de ser mapeada de maneira representativa pelos protótipos da rede. Isso ocorre porque a quantidade de neurônios no mapa foi insuficiente para aprender todos os padrões da imagem. Portanto, caso o SOM de tamanho 6x6 fosse utilizado

para reconhecer todos os padrões da imagem, a classe água não seria detectada.

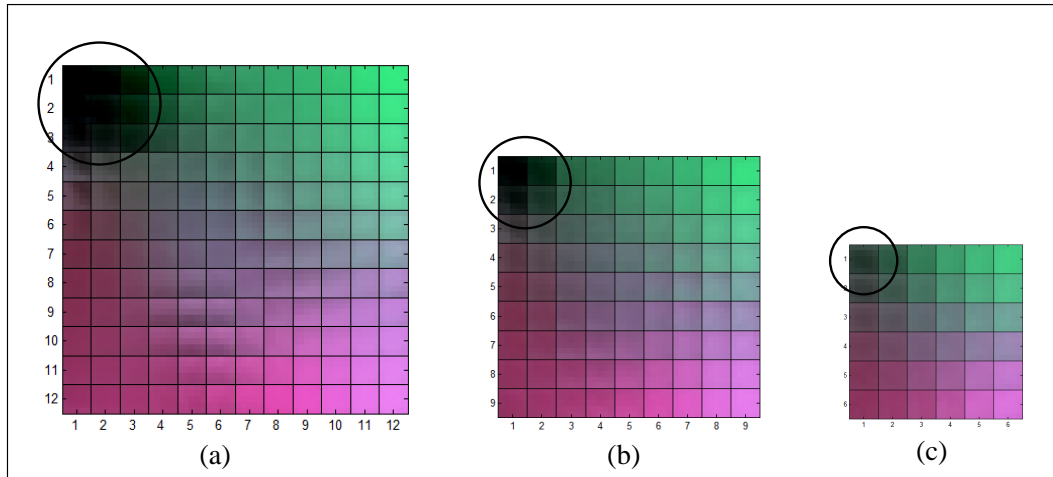


Fig. 4.11: *Efeito do tamanho do SOM no mapeamento dos padrões da imagem LANDSAT da figura 4.10: (a) SOM 12x12: protótipos correspondentes à classe água aparecem (circulados) no canto superior esquerdo do mapa. (b) SOM 9x9: o mapa também apresenta protótipos (circulados no canto superior esquerdo) correspondentes à classe água. (c) SOM 6x6: a classe água não aparece de maneira representativa no mapa.*

Testes experimentais realizados sobre diferentes imagens de sensoriamento remoto, compostas em média, por 512x512 pixels, mostram que SOMs com tamanhos variando de 8x8 a 25x25 neurônios têm sido suficientes para mapear os padrões de cobertura terrestre existentes nas cenas. Contudo, com exceção do tempo de processamento, o desempenho do método de classificação pode não ser significativamente afetado se mapas com dimensões maiores do que a necessária são utilizados. Como será visto adiante, assim como são propostas estratégias para tratar de maneira adequada amostras e protótipos que incorporam mais de uma classe da imagem, o mesmo acontece para protótipos inativos (protótipos que não possuem nenhum padrão de entrada associado a eles). Além disso, será possível concluir que os protótipos inativos também contribuem no processo de separação de agrupamentos, uma vez que geralmente fazem parte de regiões no mapa que correspondem à áreas de transição entre classes distintas.

Em todos os experimentos realizados neste trabalho, os SOMs foram treinados utilizando as rotinas básicas de inicialização e treinamento da biblioteca SOMTOOLBOX (para Matlab) proposta por Esa Alhoniemi, Johan Himberg, Juha Parhankangas and Juha Vesanto do Laboratório de Ciência da Informação e da Computação da Universidade de Helsinki.

Vale observar ainda que, o usuário pode empregar estratégias para diminuir o tempo de treinamento do SOM, como aquela apresentada em Costa (1999), cuja idéia básica é diminuir o tamanho do raio da busca pelo neurônio vencedor, ou então aquela apresentada em Gonçalves (1997), que divide o mapa em subgrupos de neurônios que são processados em paralelo utilizando um ambiente computacional com diversos nós processadores.

#### **4.2.2.1 Mapa de Características da Imagem**

A metodologia proposta possui uma particularidade em relação ao mapeamento da imagem através do SOM que deve ser destacada. Uma vez que os neurônios do SOM são dispostos espacialmente na forma de uma grade retangular, e considerando ainda que nesta rede neural os vetores de pesos de cada neurônio possuem as mesmas dimensões que os padrões de entrada (que neste caso são janelas de pixels), torna-se possível gerar uma imagem da grade de neurônios do mapa. Conforme mostra a figura 4.12, os neurônios ou protótipos do SOM (através dos seus vetores de pesos sinápticos) correspondem à janelas de pixels, o que permite a visualização dos mesmos na forma de uma imagem. Essa representação visual da grade de neurônios do SOM, após o seu treinamento não supervisionado, foi denominada de *Mapa de Características da Imagem* (MCI).

Como será apresentado adiante em alguns resultados experimentais, através do MCI é possível visualizar o mapeamento dos padrões da imagem de entrada realizado pelo SOM e suas respectivas propriedades, como a formação de agrupamentos, a aproximação de densidade e da topologia dos dados da imagem. Embora o MCI seja útil para visualizar a abstração dos dados construída pelo SOM, é importante salientar que nos métodos propostos neste trabalho, nenhuma interação do usuário é realizada através dele para realizar a classificação das imagens.

#### **4.2.3 Segmentação do SOM**

Um dos diferenciais da metodologia de classificação proposta aqui é a maneira pela qual o SOM treinado com os dados da imagem é interpretado. Diferentemente de outras abordagens de classificação de imagens de sensoriamento remoto, que também utilizam o SOM para a detecção de agrupamentos, mas que exigem a participação do usuário para segmentar o mapa

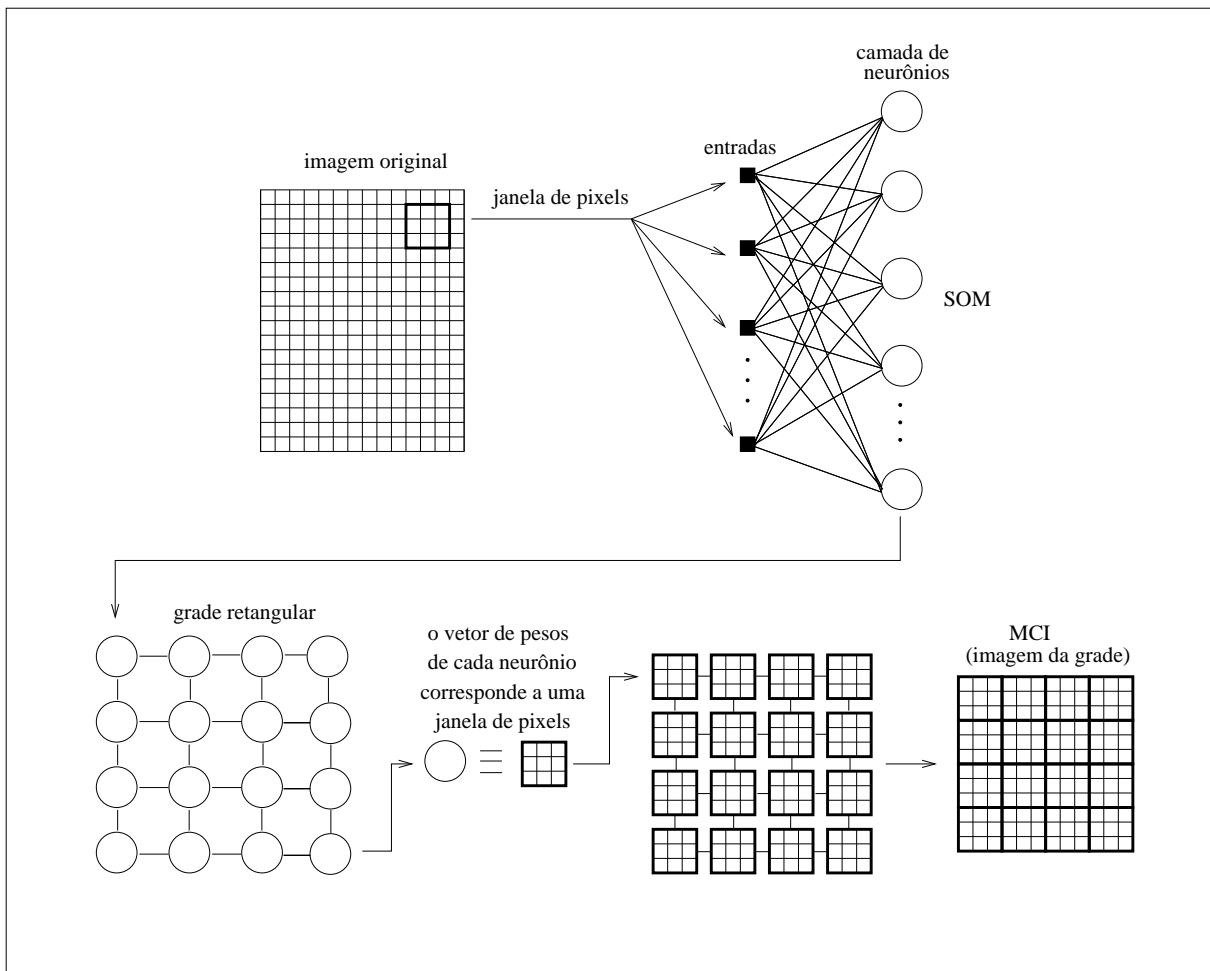


Fig. 4.12: Ilustração do processo de construção do Mapa de Características da Imagem (MCI).

de neurônios da rede, o método de classificação proposto realiza a segmentação do SOM treinado de maneira automática. A segmentação do mapa de neurônios do SOM e a determinação das classes da imagem de entrada são efetuadas sem a necessidade de participação do usuário.

A estratégia de segmentar o SOM de maneira automática ajuda a preservar algumas vantagens importantes apresentadas pelas técnicas de classificação não-supervisionadas, minimizando a influência ou a chance de erro por parte do usuário, o qual muitas vezes possui pressuposições a respeito da imagem a ser classificada que não são compatíveis com a estrutura real dos dados.

Na metodologia de classificação proposta, o SOM é particionado utilizando o método de segmentação proposto por Costa e Andrade Netto (1999, 2001). Baseado em técnicas de morfologia matemática, este método proporciona vantagens significativas à metodologia de classificação

proposta em comparação com métodos convencionais de classificação não-supervisionada. O método utilizado para segmentar o SOM permite detectar automaticamente os agrupamentos de dados da imagem sem assumir qualquer modelo de distribuição de probabilidade ou qualquer suposição sobre a estrutura dos dados. A representação distribuída dos agrupamentos por meio de grupos de protótipos do SOM provê à metodologia de classificação a potencialidade de discriminar padrões de cobertura terrestre com geometrias complexas e variadas. O método de segmentação do SOM empregado produz ainda uma árvore dinâmica de mapas que revela relacionamentos hierárquicos existentes entre as classes espectrais existentes nas imagens.

Vale lembrar que o método de segmentação do SOM proposto por Costa e Andrade Netto (1999, 2001) foi descrito na seção 3.7.1 (onde um exemplo de uso sobre um conjunto de dados sintético foi apresentado). Algumas particularidades da utilização do método dentro da metodologia de classificação proposta são descritas na seção 4.3, onde resultados experimentais sobre uma imagem de sensoriamento remoto são apresentadas.

No final desta etapa, o SOM apresenta-se segmentado em diferentes classes de protótipos, cada uma delas com um rótulo particular correspondendo à uma classe de cobertura terrestre existente na imagem de entrada.

#### **4.2.4 Classificação da imagem**

Após as etapas de treinamento e de segmentação automática do SOM, a imagem de entrada é então classificada utilizando os protótipos do SOM já rotulado. Janelas de pixels com dimensões iguais às das amostras de treinamento são coletadas da imagem original e comparadas com todos os protótipos do SOM. Essa comparação é realizada através das distâncias calculadas entre a janela de pixels considerada e cada um dos protótipos. O pixel central da janela de pixels corrente recebe o rótulo do protótipo do SOM que apresenta a menor distância a ela. A imagem é então inteiramente percorrida até que todos os pixels sejam classificados dessa forma.



## 4.3 Resultados Experimentais

### 4.3.1 Imagem teste

A metodologia de classificação proposta neste capítulo foi aplicada sobre uma imagem teste provida pela Divisão de Geração de Imagens do Instituto Nacional de Pesquisas Espaciais (DGI/INPE). A imagem é composta pelas bandas espectrais 3, 4 e 5 do satélite LANDSAT-5 (TM), obtida em 20/08/1988, possui 496x512 pixels e mostra a cidade de Manaus e o encontro dos rios Negro e Solimões no estado do Amazonas, Brasil. A figura 4.13 mostra uma composição colorida da imagem utilizando as três bandas espectrais. A cena apresenta quatro classes de cobertura terrestre predominantes: vegetação, área urbana e dois padrões de água (um mais escuro correspondente ao Rio Negro e outro de cor mais arroxeadada correspondente ao Rio Solimões). Pequenas áreas com vegetação rala ou desmatada também aparecem na cena, mas em uma densidade muito baixa.



Fig. 4.13: *Composição colorida das bandas 3(R), 4(G) e 5(B) da imagem teste.*

### 4.3.2 Aplicação da metodologia proposta

Na execução da primeira etapa da metodologia de classificação proposta, que consiste em realizar o processo de amostragem da cena, janelas de tamanho  $5 \times 5$  foram coletadas de maneira uniforme sobre toda a região da imagem, sem sobreposições e em intervalos regulares de 10 pixels, totalizando 2.499 amostras obtidas sem a intervenção do usuário. A figura 4.14 mostra o conjunto de janelas amostrais capturadas.

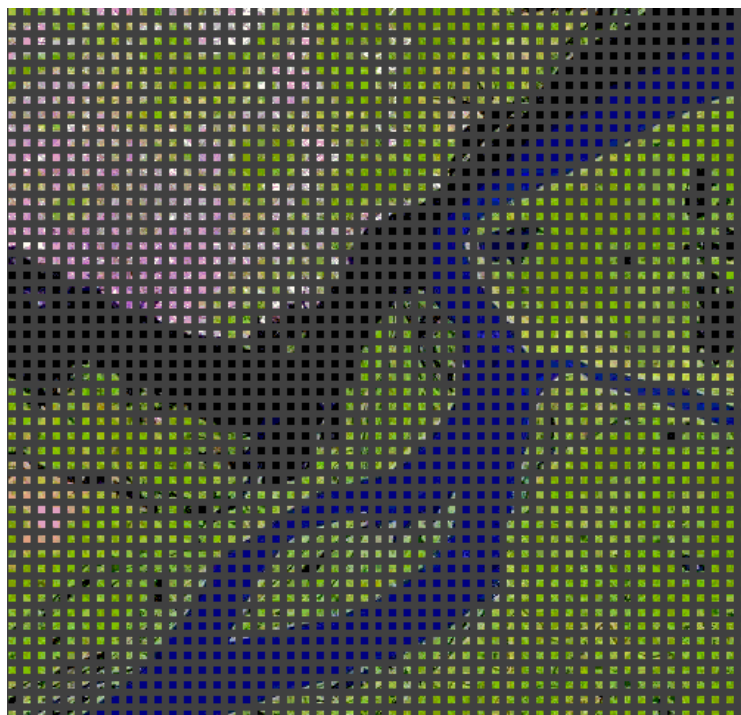


Fig. 4.14: *Janelas amostrais de tamanho  $5 \times 5$  coletadas a partir da imagem original.*

Um SOM inicial (mapa pai) composto por 144 neurônios, dispostos num arranjo retangular  $12 \times 12$ , foi treinado com todas as janelas amostrais. Os demais parâmetros de treinamento do SOM foram definidos de acordo com as especificações apresentadas na seção 4.2.2. O tempo de treinamento sob o Matlab em uma máquina com processador AMD Athlon XP 2600+ 1,91GHZ e 1GB de RAM foi de 72 segundos.

A figura 4.15 mostra o MCI obtido após o treinamento do SOM. Conforme descrito anteriormente, o MCI é uma imagem gerada a partir da grade de neurônios do SOM treinado. Através dele, é possível visualizar as propriedades do mapeamento realizado pela rede neural.

As quatro grandes classes de cobertura terrestre presentes na imagem original (figura 4.13) aparecem em forma de agrupamentos no MCI. No canto superior esquerdo se encontram os protótipos correspondentes ao padrão de água mais escuro, no canto superior direito estão presentes os protótipos referentes à outra classe de água (mais arroxeadada), no canto inferior direito (em rosa) se encontram aqueles que correspondem à área urbana, e do canto inferior esquerdo ao centro do MCI estão os protótipos associados à classe vegetação. Observando a imagem original, verifica-se que a classe vegetação é aquela que apresenta a maior área de ocupação na cena. Sendo assim, como o SOM reflete a distribuição de probabilidade dos dados de entrada, os protótipos associados à classe vegetação (em verde no MCI) se apresentam em maior número do que os protótipos das outras classes. A propriedade de ordenação topológica do SOM também pode ser notada. Os protótipos das duas classes de água presentes na imagem original aparecem próximos um dos outros na parte superior do MCI, isto porque os dados de entrada (amostras) correspondentes a essas duas classes possuem uma maior similaridade espectral em comparação com os atributos espectrais das outras duas classes de cobertura terrestre.

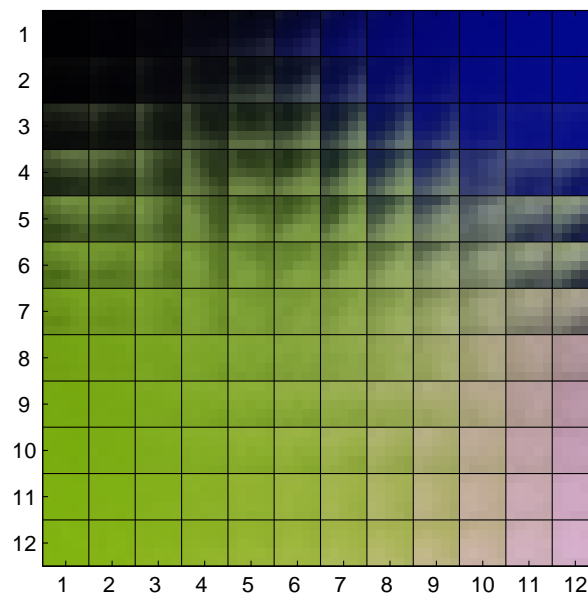


Fig. 4.15: *Imagem do MCI após o treinamento do SOM.*

Após o treinamento do SOM, o método de segmentação automática SL-SOM proposto por Costa e Andrade Netto (1999, 2001) foi aplicado sobre o mapa de neurônios. A figura 4.16

mostra a  $U$ -matrix na forma de superfície do SOM treinado.

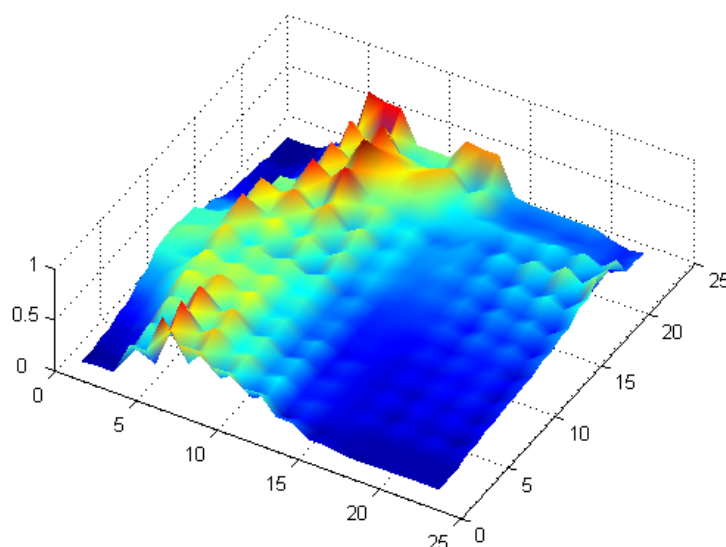


Fig. 4.16:  $U$ -matrix na forma de superfície do SOM treinado.

Aplicando o algoritmo para determinar a imagem de marcadores da  $U$ -matrix (descrito na seção 3.7.1) obtém-se o gráfico da figura 4.17, onde é possível visualizar o número de regiões conectadas ( $N_{rc}$ ) para cada valor de limiar ( $k$ ) da  $U$ -matrix. Com o objetivo de reduzir a quantidade muito grande de regiões pouco significativas, em razão do número elevado de rugosidades apresentado pela matriz de distâncias, foram aplicadas duas iterações do algoritmo de suavização da  $U$ -matrix nesta etapa da análise. Nesta abordagem, o algoritmo de suavização utilizado consiste na aplicação de um filtro morfológico de mediana de tamanho  $3 \times 3$ . Esse mesmo procedimento foi aplicado às  $U$ -matrizes dos submapas seguintes (gerados a partir deste mapa inicial).

De acordo com o algoritmo SL-SOM, a imagem de marcadores corresponde à imagem binária obtida da limiarização da  $U$ -matrix com o valor inicial da seqüência mais estável de  $N_{rc}$ , que neste caso foi  $k = 149$ , correspondendo a 2 (dois) agrupamentos. Embora seja possível visualizar quatro "vales" ou depressões nos quatro cantos da  $U$ -matrix (figura 4.16), nota-se também que há regiões mais altas ("montanhas") que se apresentam de forma acentuada dividindo a  $U$ -matrix em duas partes.

A figura 4.18 mostra a imagem de marcadores para a  $U$ -matrix considerando o valor de limiar

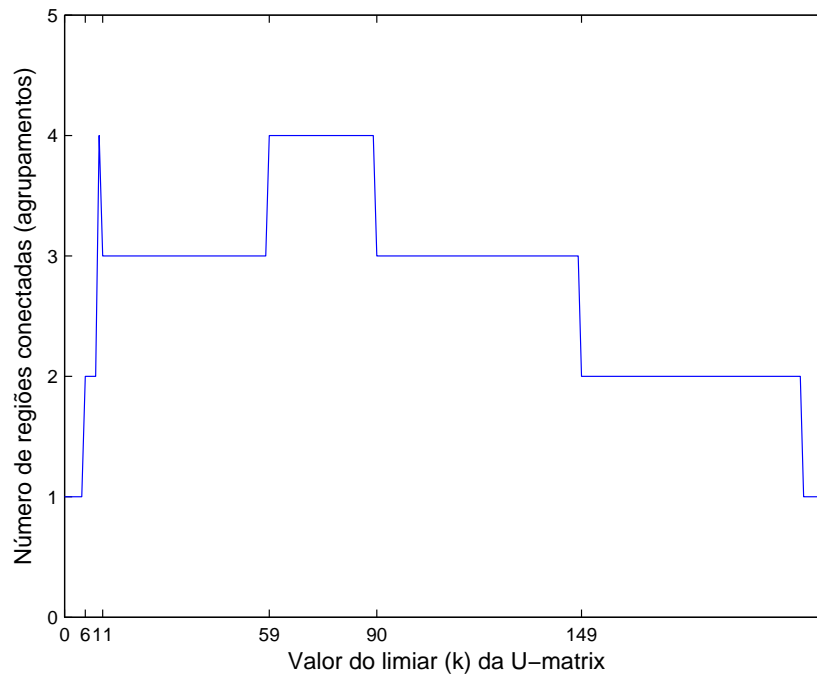


Fig. 4.17: Gráfico do valor de limiar ( $k$ ) da  $U$ -matrix versus o número de regiões conectadas (agrupamentos).

$k = 149$ , e a figura 4.19 mostra as linhas de partição encontradas pelo algoritmo *watershed*. A sobreposição das linhas de partição sobre a  $U$ -matrix também pode ser vista na figura 4.20.

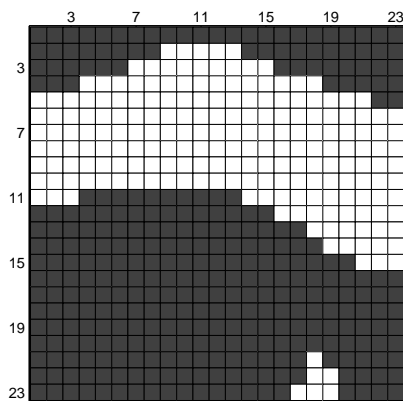


Fig. 4.18: Imagem de marcadores correspondente ao valor de limiar  $k = 149$ .

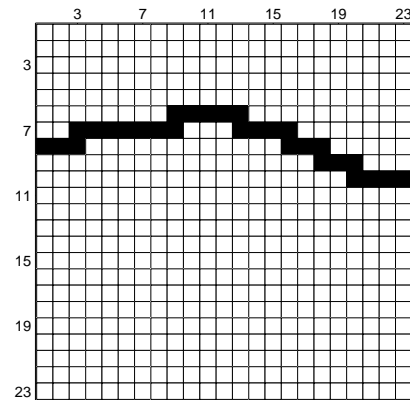


Fig. 4.19: Partição da  $U$ -matrix encontrada pelo algoritmo *watershed*.

Após a rotulação da  $U$ -matrix segmentada (figura 4.21), os códigos das regiões foram passados aos neurônios correspondentes. A figura 4.22 mostra o SOM rotulado em duas classes,

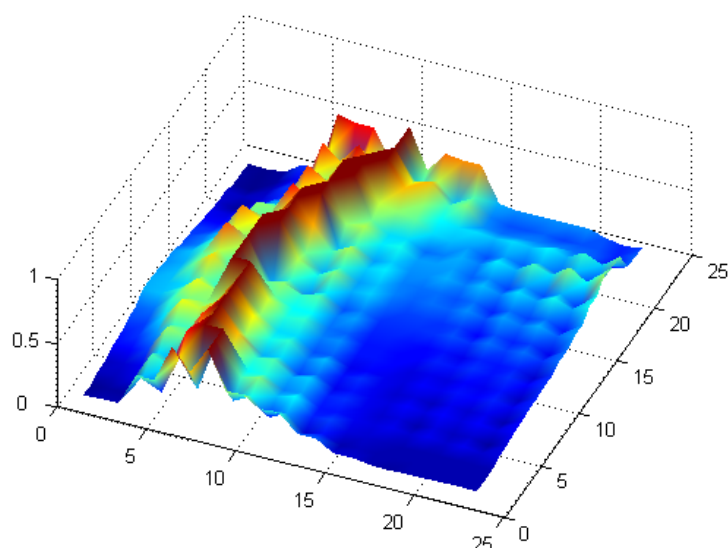


Fig. 4.20: Sobreposição das linhas de watershed sobre a  $U$ -matrix apresentada na figura 4.16.

denominadas de C1 e C2.

Comparando o MCI apresentado na figura 4.12 e o SOM rotulado (figura 4.22), verifica-se que a classe C1 inclui os protótipos dos dois padrões de água (que aparecem na parte superior do mapa), e a classe C2 inclui os protótipos dos padrões de área urbana (que aparecem em rosa no canto inferior direito do MCI) e também os de vegetação (que aparecem na cor verde).

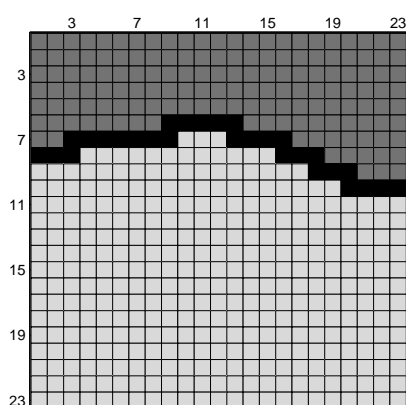


Fig. 4.21:  $U$ -matrix rotulada após a segmentação pela watershed.

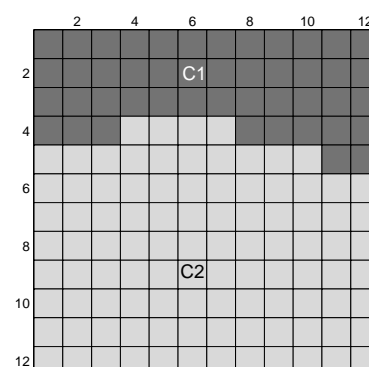


Fig. 4.22: SOM rotulado.

Dando continuidade ao processo de geração dinâmica da árvore de mapas, dois novos SOMs (submapas) são construídos e treinados utilizando os conjuntos de dados correspondentes às

classes C1 e C2: o submapa 1, filho da região 1 (classe C1), e o submapa 2, filho da região 2 (classe C2).

É importante destacar que a fórmula utilizada para a determinação do tamanho dos submapas é semelhante àquela proposta em Costa (1999)(p.201). Seja  $X$  o número de padrões utilizado para treinar o mapa pai e  $Y$  o número de padrões a ser utilizado para treinar o mapa filho (submapa), a fórmula usada nesta abordagem para determinar o tamanho  $Q$  dos submapas (em quantidade de neurônios) é dada por:

$$Q = \frac{Y}{X} \times N \quad (4.1)$$

onde  $N$  é o número de neurônios do mapa pai. Portanto, de acordo com a fórmula acima, o tamanho de um submapa é determinado de maneira proporcional ao tamanho do conjunto de padrões utilizado para treiná-lo e ao tamanho (número de neurônios) do seu mapa pai.

### Submapa 1

De acordo com a segmentação obtida para o mapa pai, o submapa 1 foi treinado somente com os dados pertencentes à classe C1 (os quais correspondem aos dois padrões de água da imagem). Nesse caso, o número de amostras foi igual a 776 e o tamanho utilizado para o submapa foi 7x7. O seu tempo de treinamento na mesma máquina utilizada para treinar o mapa pai foi de 38 segundos.

As figuras 4.23 e 4.24 mostram, respectivamente, a imagem do MCI e a *U-matrix* do submapa 1 após o treinamento. É possível visualizar os agrupamentos referentes aos dois padrões de água, o mais escuro na parte superior do MCI e o mais arroxeadado na parte inferior. Esses agrupamentos estão associados aos dois "vales" existentes na *U-matrix* correspondente. Na parte esquerda do MCI há alguns protótipos do SOM que correspondem a regiões de transição dos padrões de água com outros padrões de cobertura terrestre existentes na imagem (como área urbana e vegetação). Tais protótipos geram uma maior rugosidade nas áreas associadas a eles na imagem da *U-matrix*.

Conforme mostra a figura 4.25, para determinar a imagem de marcadores da *U-matrix*, o

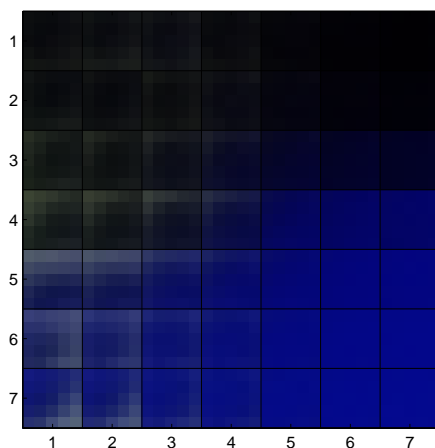


Fig. 4.23: Imagem do MCI do submapa 1 após o treinamento.

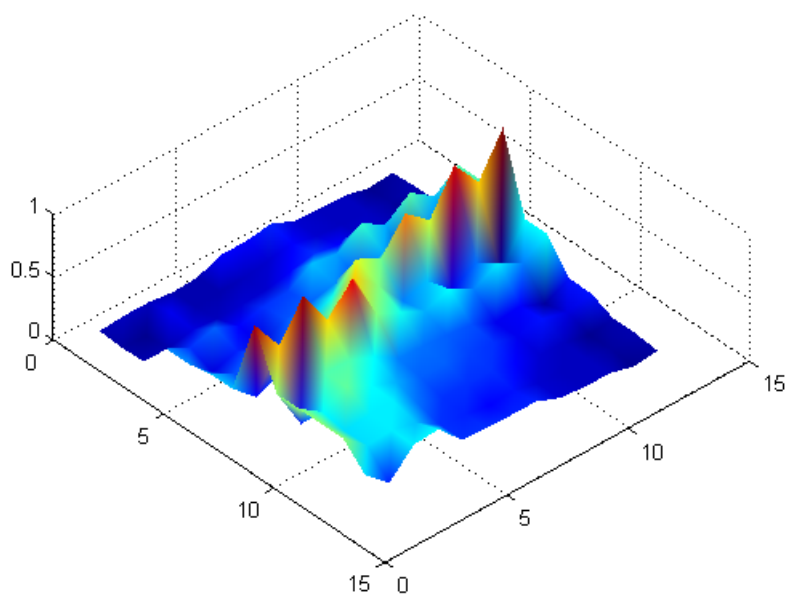


Fig. 4.24:  $U$ -matrix 3D do submapa 1 treinado.

gráfico do número de regiões conectadas ( $N_{rc}^k$ ) para cada valor de limiar da  $U$ -matrix foi construído. Neste caso, o valor inicial da seqüência mais estável de  $N_{rc}$  foi  $k = 19$ , correspondendo a 2 (dois) agrupamentos. A figura 4.26 mostra a imagem de marcadores encontrada para a  $U$ -matrix do submapa 1 e a figura 4.27 mostra a partição encontrada pelo algoritmo *watershed*. A sobreposição das linhas de partição sobre a  $U$ -matrix apresentada na figura 4.24 também pode ser vista na figura 4.28.

Após a rotulação da  $U$ -matrix segmentada (figura 4.29), os códigos das regiões foram co-



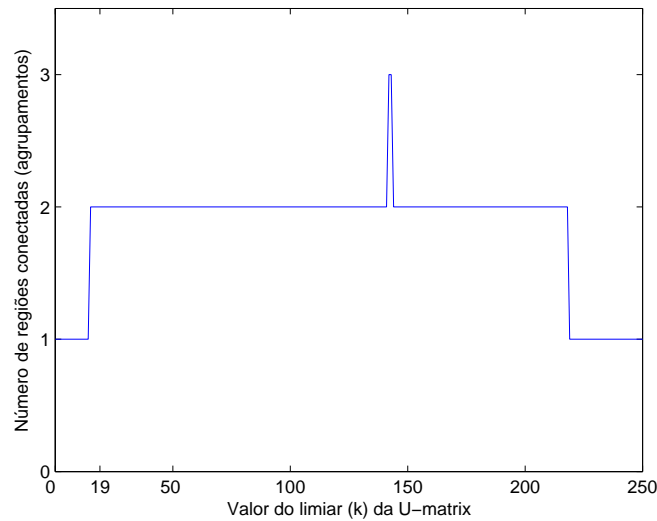


Fig. 4.25: Número de regiões conectadas versus limiar da U-matrix - submapa 1.

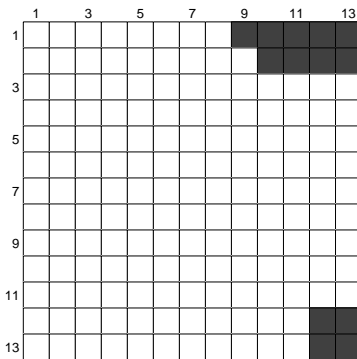


Fig. 4.26: Imagem de marcadores correspondente ao valor de limiar  $k = 19$  - submapa 1.

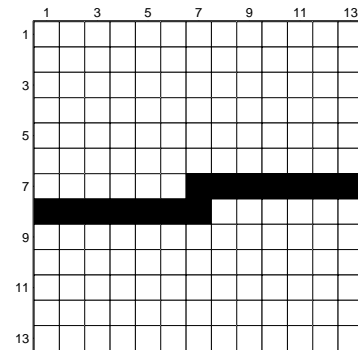


Fig. 4.27: Partição da U-matrix encontrada pelo algoritmo watershed - submapa 1.

piados para os neurônios correspondentes. A figura 4.30 mostra o submapa 1 particionado em duas classes,  $C1.1$  e  $C1.2$ , correspondendo aos dois padrões de água da imagem original.

Outros dois submapas foram construídos a partir do submapa 1: o submapa 1.1, filho da região 1 (classe  $C1.1$ ), o qual foi composto por um arranjo  $5 \times 5$  de neurônios e treinado com as 376 janelas amostrais da classe 1, e o submapa 1.2, filho da região 2 (classe  $C1.2$ ), composto também por um arranjo de neurônios  $5 \times 5$  e treinado com as 400 janelas amostrais da classe 2. No entanto, como não houve detecção de mais de uma região ou classe, os dois submapas foram eliminados. De acordo com Costa e Andrade Netto (1999, 2001), quando há eliminação de um

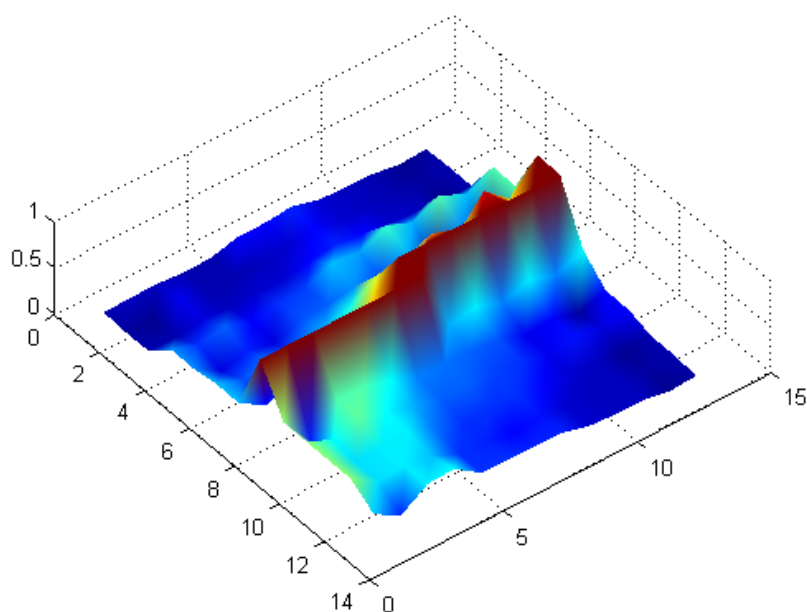


Fig. 4.28: Sobreposição das linhas de watershed sobre a  $U$ -matrix apresentada na figura 4.24.

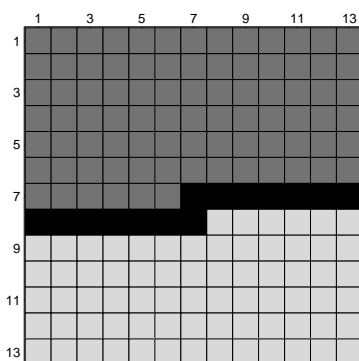


Fig. 4.29:  $U$ -matrix segmentada - submapa 1.

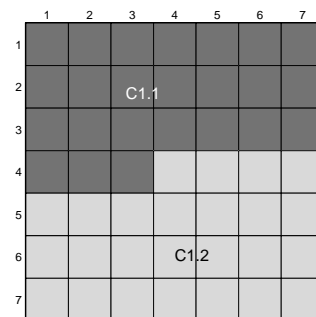


Fig. 4.30: Submapa 1 rotulado.

submapa em determinado nível da árvore, a região de neurônios que deu origem ao submapa é suficiente para representar o subconjunto de dados utilizado.

Vale observar que o critério utilizado para determinar se não há mais de um agrupamento (ou classe) em um submapa baseia-se no gráfico do valor de limiar ( $k$ ) da  $U$ -matrix versus o número de regiões conectadas (agrupamentos). De forma semelhante àquela adotada em Costa (1999), caso o número de níveis de cinza correspondentes a regiões conectadas diferentes de 1 seja inferior a 20% do número total de níveis de cinza da  $U$ -matrix, considera-se então que não

haverá segmentação do mapa (submapa). Alternativamente, conforme apresentado em Costa (1999)(p.221), outros critérios para decidir se haverá ou não segmentação dos submapas podem ser utilizados.

As figuras 4.31 e 4.32 mostram, respectivamente, os gráficos do valor de limiar ( $k$ ) versus o número de regiões conectadas das  $U$ -matrizes dos submapas 1.1 e 1.2. Pode-se observar que o número de níveis de cinza correspondentes a regiões conectadas diferentes de 1 é inferior a 20% do número total de níveis de cinza das  $U$ -matrizes dos dois submapas (2% para o submapa 1.1 e 0.5% para o submapa 1.2).

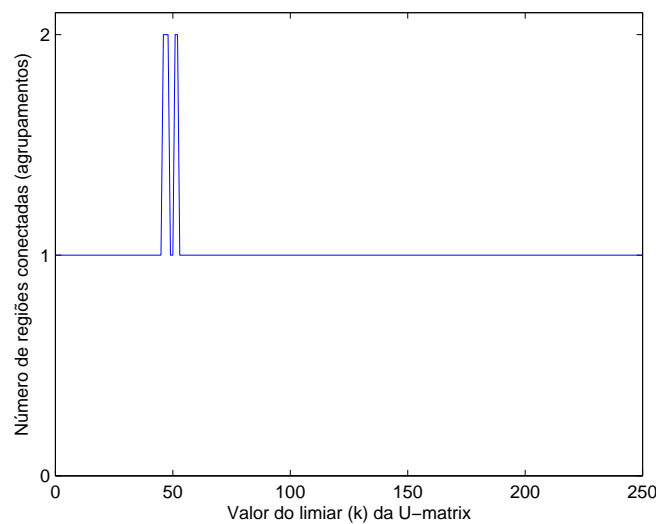


Fig. 4.31: *Número de regiões conectadas versus limiar da U-matrix - submapa 1.1.*

## Submapa 2

O submapa 2, filho da região 2 do mapa pai, foi treinado apenas com os dados associados à classe C2, os quais correspondem aos padrões de área urbana e vegetação da imagem teste. Nesse caso, o número de amostras foi igual a 1723 e o tamanho utilizado para o submapa foi 9x9. O seu tempo de treinamento na mesma máquina utilizada para treinar o mapa pai foi de 53 segundos.

A imagem do MCI e a  $U$ -matrix do submapa 2 após o treinamento são apresentadas nas figuras 4.33 e 4.34, respectivamente. É possível visualizar os agrupamentos referentes às classes

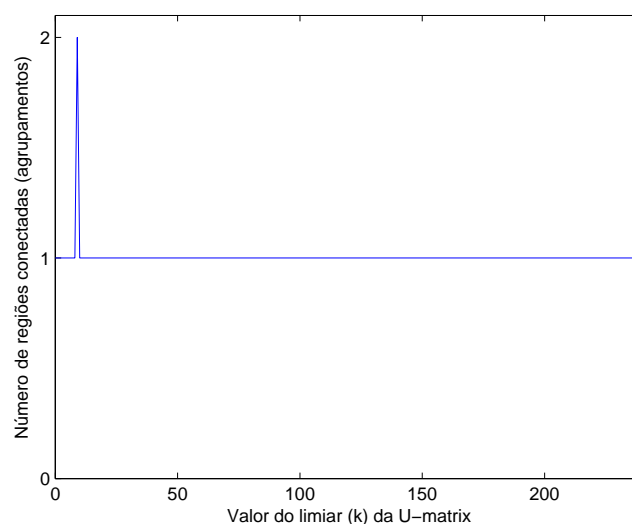


Fig. 4.32: *Número de regiões conectadas versus limiar da U-matrix - submapa 1.2.*

vegetação (em verde, ocupando quase toda a extensão do mapa) e área urbana (em rosa, no canto inferior esquerdo). Observando a imagem da *U-matrix* nota-se a existência de dois "vales" ou depressões: um pequeno, correspondente ao agrupamento do padrão de área urbana, e outro mais extenso, correspondente ao padrão de vegetação. A *U-matrix* obtida também apresenta superfícies mais altas com graus de rugosidade acentuados. Essas superfícies estão associadas à protótipos que representam regiões de transição entre os padrões de cobertura terrestre existentes na imagem.

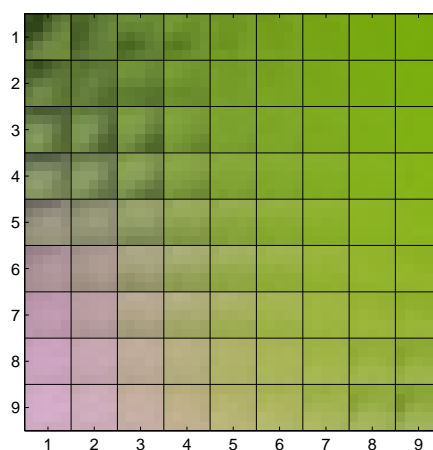


Fig. 4.33: *Imagem do MCI do submapa 2 após o treinamento.*

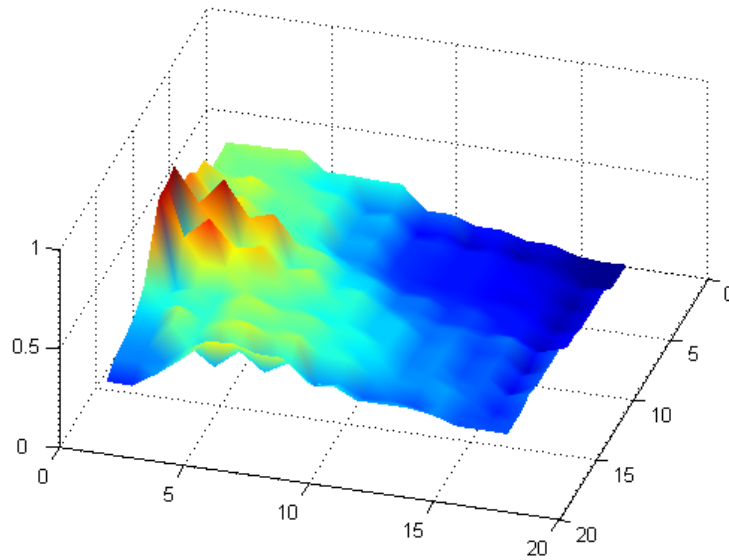


Fig. 4.34:  $U$ -matrix 3D do submapa 1 treinado.

Para determinar a imagem de marcadores da  $U$ -matrix, o gráfico do número de regiões conectadas ( $N_{rc}^k$ ) para cada valor de limiar da  $U$ -matrix foi construído (figura 4.35). Neste caso, o valor inicial da seqüência mais estável de  $N_{rc}$  foi  $k = 78$ , correspondendo a 2 (dois) agrupamentos. Vale ressaltar que, no gráfico da figura 4.35, a quantidade de níveis de cinza correspondente ao número de regiões conectadas igual a 2 é igual a 35% do número total de níveis de cinza da  $U$ -matrix (superior à taxa mínima estabelecida para dividir o mapa em mais de uma região).

A figura 4.36 mostra a imagem de marcadores encontrada para a  $U$ -matrix do submapa 2 e a figura 4.37 mostra a partição encontrada pelo algoritmo *watershed*. A sobreposição das linhas de partição sobre a  $U$ -matrix apresentada na figura 4.34 também pode ser vista na figura 4.38.

Após a rotulação da  $U$ -matrix segmentada (figura 4.39), os códigos das regiões foram copiados para os neurônios correspondentes. A figura 4.40 mostra o submapa 2 particionado em duas classes,  $C2.1$  e  $C2.2$ , correspondendo, respectivamente, aos padrões de área urbana e de vegetação.

Posteriormente, mais dois submapas foram construídos a partir do submapa 2: o submapa 2.1, filho da região 1 (classe  $C2.1$ ), o qual foi composto por um arranjo 8x8 de neurônios e

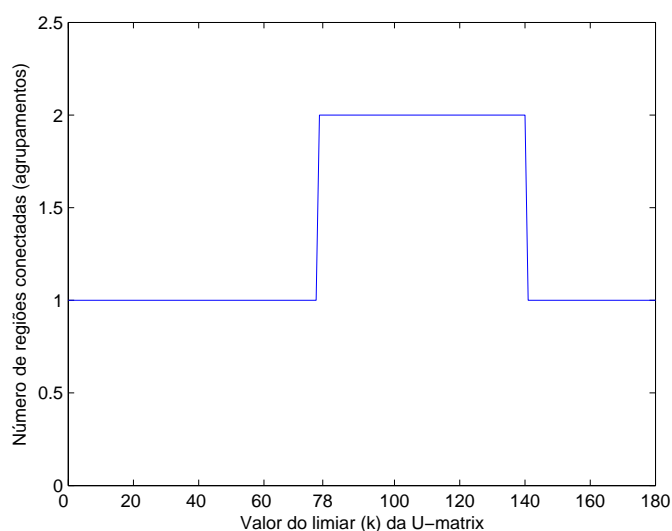


Fig. 4.35: Número de regiões conectadas versus limiar da *U-matrix* - submapa 2.

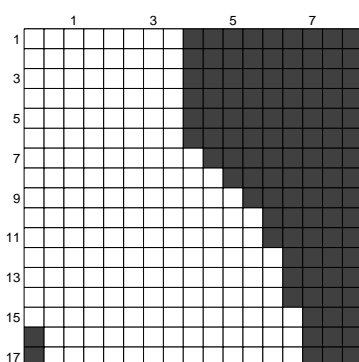


Fig. 4.36: Imagem de marcadores correspondente ao valor de limiar  $k = 78$  - submapa 2.

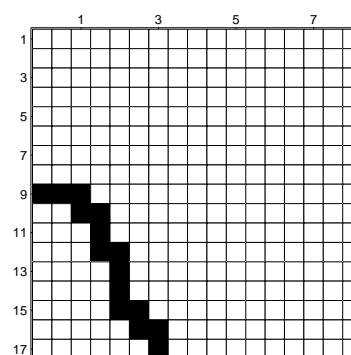


Fig. 4.37: Partição da *U-matrix* encontrada pelo algoritmo watershed - submapa 2.

treinado com as 1455 amostras da classe 1, e o submapa 2.2, filho da região 2 (classe  $C2.2$ ), composto por um arranjo  $4 \times 4$  de neurônios e treinado com as 268 amostras da classe 2.

As figuras 4.41 e 4.42 mostram, respectivamente, os gráficos do valor de limiar ( $k$ ) versus o número de regiões conectadas das *U-matrizes* dos submapas 2.1 e 2.2. O número de níveis de cinza associados às regiões conectadas diferentes de 1 no submapa 2.1 corresponde a 8% do número total de níveis de cinza da sua *U-matrix*, enquanto que para o submapa 2.2 esse número corresponde a 13%. Portanto, em ambos os casos o número de níveis de cinza associados às regiões conectadas diferentes de 1 é inferior a 20% do número total de níveis de cinza das *U-*

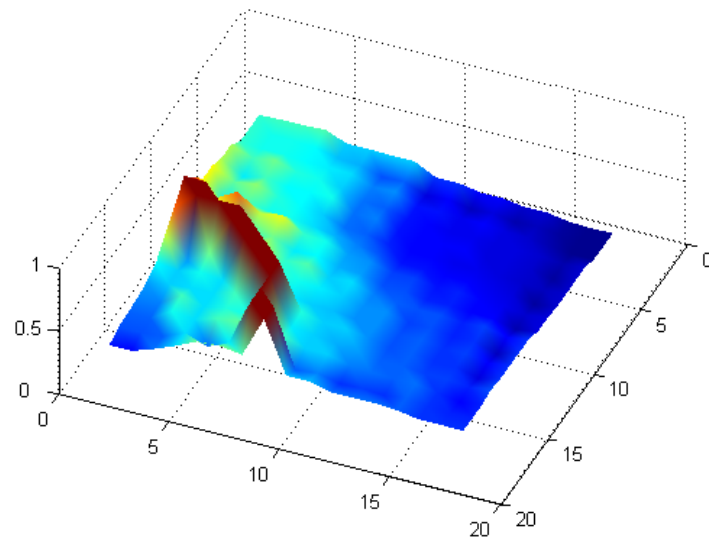


Fig. 4.38: Sobreposição das linhas de watershed sobre a  $U$ -matrix apresentada na figura 4.34.

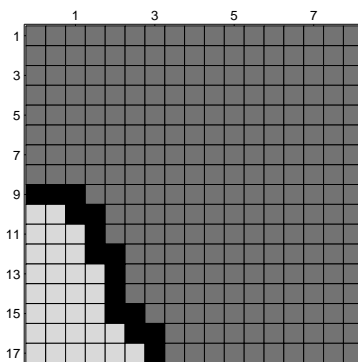


Fig. 4.39:  $U$ -matrix segmentada - submapa 2.

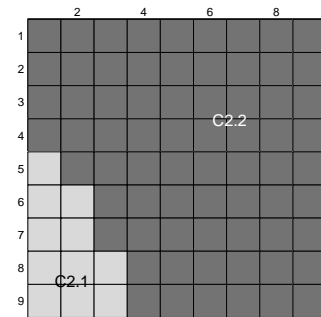


Fig. 4.40: Submapa 2 rotulado.

*matrizes* (taxa mínima estabelecida para permitir uma nova segmentação do submapa). Logo, os dois submapas foram descartados.

A figura 4.43 ilustra um resumo do processo de análise de agrupamentos da imagem teste, mostrando a árvore de mapas gerada pelo método de segmentação do SOM proposto por Costa e Andrade Netto (1999, 2001). Através dela, é possível ver de forma hierárquica a distribuição das classes existentes na imagem. Foram construídos e treinados um total de sete SOMs. O mapa pai, contendo protótipos representantes de todas as classes existentes

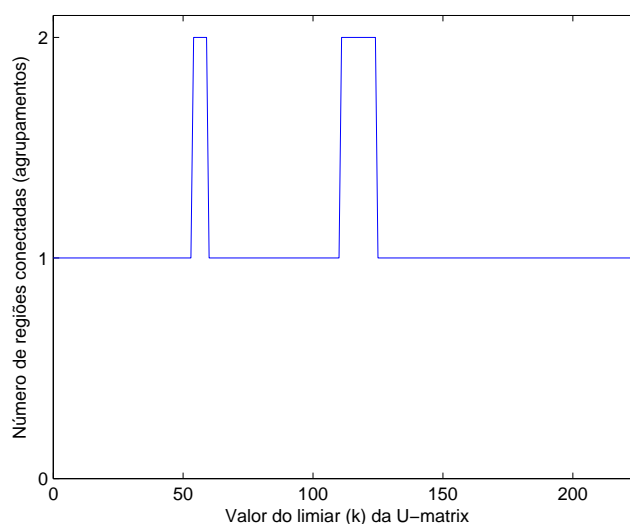


Fig. 4.41: *Número de regiões conectadas versus limiar da U-matrix - submapa 2.1.*

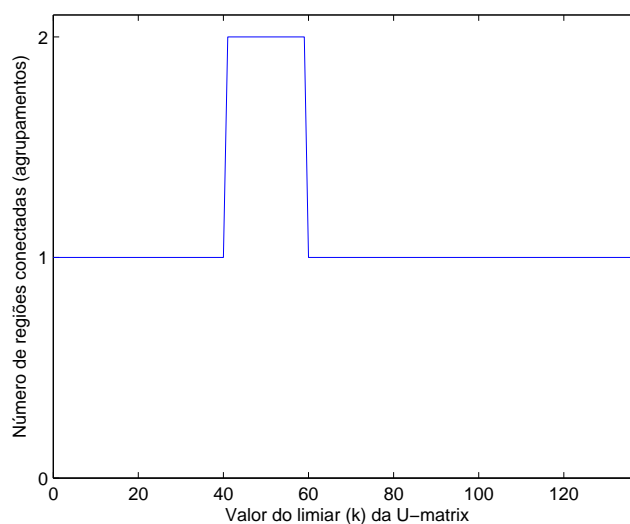


Fig. 4.42: *Número de regiões conectadas versus limiar da U-matrix - submapa 2.2.*

na imagem, foi particionado em dois submapas: o submapa 1, que foi particionado em duas classes de protótipos, uma classe representando o padrão de água de cor mais escura e a outra representando o padrão de água de cor mais arroxeadada, e o submapa 2, que também foi particionado em duas classes, uma delas composta por protótipos representantes da classe vegetação e a outra com protótipos representantes da classe área urbana. Os quatro últimos submapas, gerados a partir dos submapas 1 e 2, foram eliminados por não terem apresentado



mais de um agrupamento. O método detectou corretamente as 4 classes de cobertura terrestre predominantes na imagem teste.

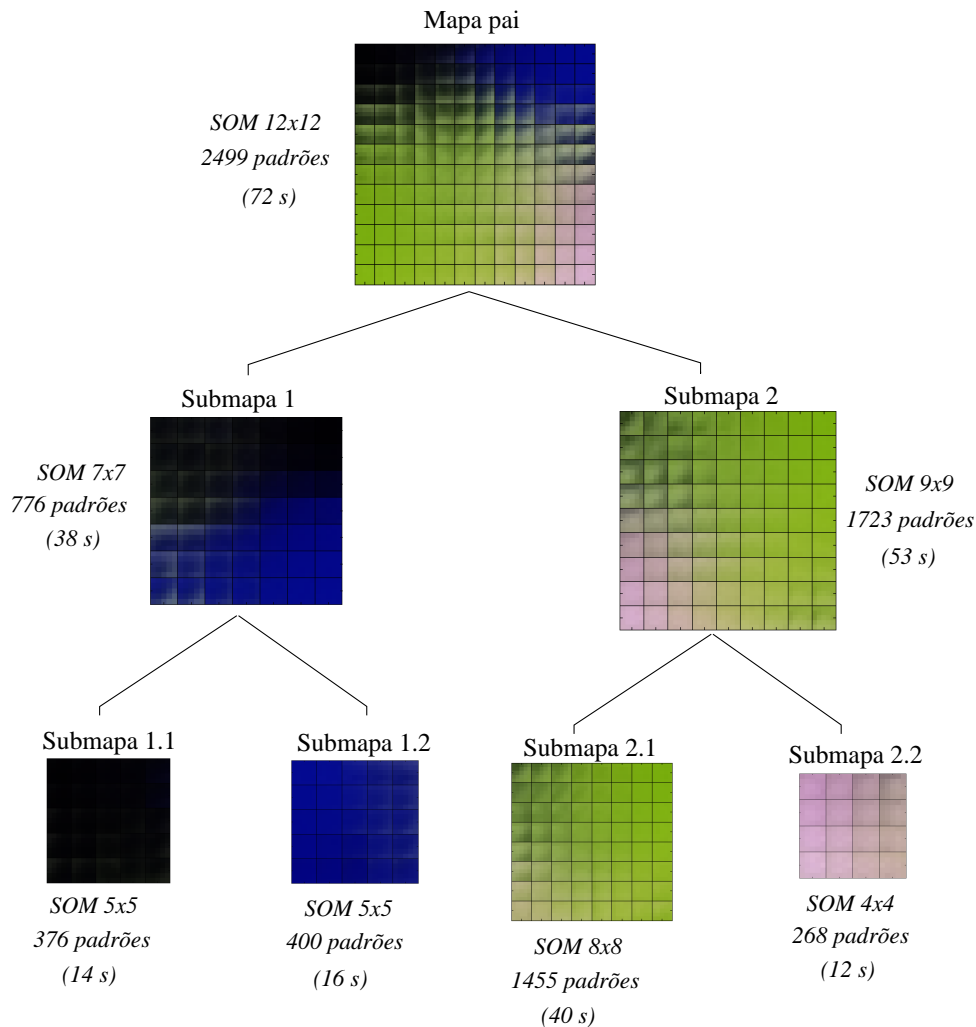


Fig. 4.43: Árvore de mapas gerada pelo método de segmentação do SOM proposto por Costa e Andrade Netto (1999, 2001) aplicado na análise de agrupamentos dos dados da imagem teste. Para cada mapa da árvore são apresentados o tamanho, número padrões e tempo de treinamento em segundos (s).

A última etapa da metodologia de classificação proposta consiste em realizar a classificação de todos os pontos da imagem teste utilizando como referência os conjuntos de protótipos dos SOMs que representam as classes de cobertura terrestre detectadas. A imagem foi então percorrida inteiramente considerando janelas de pixels de tamanho 5x5 (tamanho igual ao das janelas amostrais), as quais foram comparadas com os protótipos do SOM. Essa comparação é feita

através das distâncias calculadas entre a janela de pixels considerada e cada um dos protótipos do SOM. O pixel central da janela de pixels recebe o rótulo do protótipo que apresenta a menor distância a ela. A figura 4.44 mostra a imagem teste classificada pela metodologia proposta.

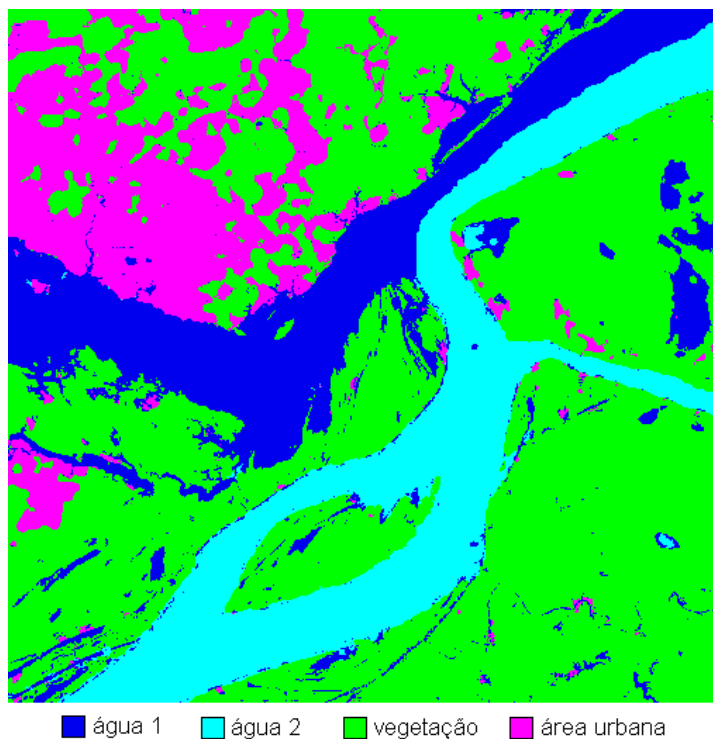


Fig. 4.44: *Imagem teste classificada pela metodologia proposta.*

### 4.3.3 Análise dos resultados

Analisando visualmente a imagem original (figura 4.13) e a imagem classificada (figura 4.44), verifica-se que o método proposto classificou adequadamente a imagem dentro das quatro classes de cobertura terrestre predominantes na cena. Algumas pequenas áreas de vegetação esparsa ou desmatada foram classificadas como área urbana, por serem espectralmente muito similares e por se apresentarem com baixa densidade em comparação com os outros padrões de cobertura terrestre.

Procurando realizar uma análise mais criteriosa e menos subjetiva da classificação obtida pela metodologia proposta e, considerando a ausência de verdade terrestre para a imagem teste,

outra classificação da mesma imagem foi realizada de maneira supervisionada e o seu resultado foi utilizado como referência (ou "verdade"). Conforme descrito na seção 2.5.1, na classificação supervisionada o usuário interage no processo selecionando as amostras das classes de interesse a partir da imagem a ser classificada.

Nesse experimento, a classificação supervisionada, considerando as quatro classes predominantes na imagem teste, foi feita através de uma rede neural de *Perceptrons* Multicamadas com o algoritmo de treinamento *Backpropagation* (descrita brevemente na seção 2.5.1.2). Essa classe de redes neurais tem sido amplamente empregada para realizar a classificação supervisionada de imagens de sensoriamento remoto por apresentar uma série de vantagens em relação aos classificadores estatísticos convencionais (Gonçalves *et al.*, 1998b; Liu *et al.*, 2003).

A matriz de confusão e o índice de concordância Kappa (descritos na seção 2.6) foram utilizados para medir a qualidade da classificação resultante. Eles foram calculados a partir da imagem classificada pelo método proposto e da imagem referência (resultante da classificação supervisionada).

Analisando os valores da matriz de confusão (tabela 4.1) e considerando que o valor obtido para o índice de concordância Kappa foi 0.90, pode-se então concluir, tomando como base a tabela 2.5, que a qualidade do resultado da classificação obtida pelo método proposto foi excelente.

<i>Classes</i>	água 1	água 2	vegetação	área urbana	<i>Total</i>
água 1	<b>93.45</b>	2.21	0.86	0.39	16.51
água 2	2.33	<b>97.78</b>	0.19	4.73	15.69
vegetação	4.22	0.00	<b>94.33</b>	6.55	54.32
área urbana	0.00	0.01	4.62	<b>88.33</b>	13.48
<i>Total</i>	100.00	100.00	100.00	100.00	100.00

Tab. 4.1: Matriz de confusão das classes (em porcentagem) entre a imagem classificada pelo método proposto e a imagem referência (classificada de modo supervisionado). As linhas da matriz mostram os erros de comissão (porcentagem de pixels que pertencem a outra classe, mas que foi classificado como a classe de interesse) e as colunas mostram os erros de omissão (porcentagem de pixels que pertencem a classe da referência/verdade, mas que foram classificados incorretamente)

A tabela 4.2 mostra o tempo de processamento consumido neste experimento por cada uma

das etapas da metodologia de classificação proposta. Vale ressaltar que os tempos das etapas de treinamentos e de segmentações do SOM apresentados na tabela equivalem, respectivamente, às somas dos tempos de treinamento e de segmentação de cada um dos sete SOMs gerados durante a aplicação do método.

Etapas do método proposto	Tempo consumido (em segundos)
Amostragem	9
Treinamentos dos SOMs	245
Segmentações dos SOMs	12
Classificação da imagem	38
<i>Total</i>	304

Tab. 4.2: Tempo consumido pelas etapas do método de classificação proposto.

O tempo total consumido pelo método proposto na classificação da imagem teste foi de 304 segundos. Observando os tempos de processamento de cada uma das etapas, nota-se que aquela referente aos treinamentos dos SOMs é a que apresenta o maior custo computacional. Particularmente neste experimento, o tempo gasto para treinar todos os SOMs que foram gerados durante a análise de agrupamentos da imagem foi equivalente a aproximadamente 80% do tempo total consumido pelo método de classificação. Isso ocorre devido ao número de mapas (ou submapas) que são gerados pelo método de segmentação do SOM utilizado para detectar os agrupamentos de dados existentes na imagem.

#### 4.3.3.1 Comparação com o algoritmo K-médias

Buscando ainda comparar o desempenho da metodologia proposta com métodos convencionais de classificação não-supervisionada de imagens de sensoriamento remoto, a imagem teste também foi classificada utilizando o algoritmo K-médias. Embora haja algumas variantes do algoritmo K-médias, a sua versão padrão (descrita na seção 2.5.2.1) foi utilizada nos experimentos por ser a mais comumente empregada pela comunidade de sensoriamento remoto.

A figura 4.45 mostra o resultado da classificação da imagem pelo algoritmo K-médias considerando o número de classes ( $k$ ) igual a quatro. O algoritmo K-médias foi executado utilizando

cinco inicializações diferentes para os centróides de cada classe e um número máximo de iterações igual a 100. A classificação considerada para a análise foi aquela que apresentou o valor mais baixo para a soma dos quadrados das distâncias dada pela equação 2.4.

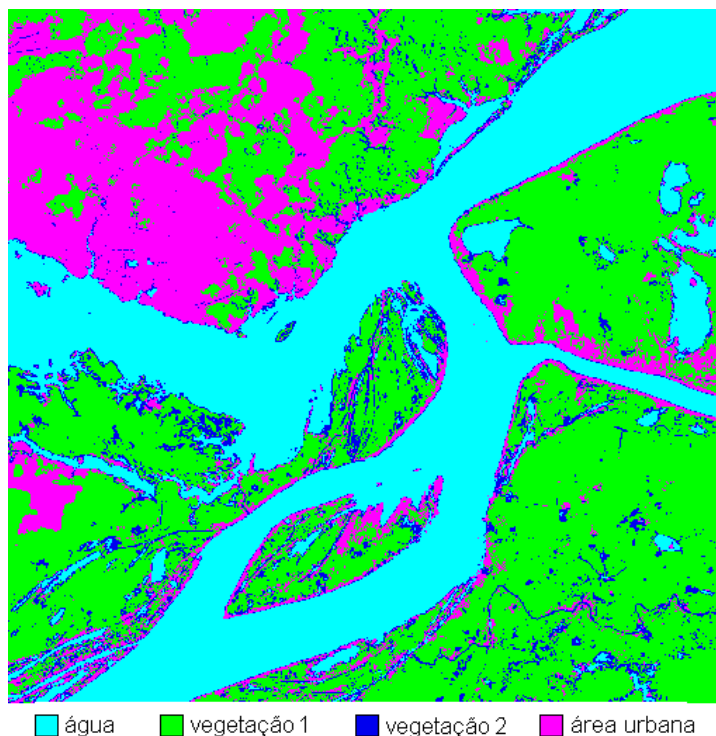


Fig. 4.45: Imagem teste da figura 4.13 classificada pelo algoritmo *K*-médias considerando  $k = 4$ .

Diferentemente do método proposto, o algoritmo *K*-médias não discriminou um dos padrões de água predominantes na imagem e classificou as áreas de vegetação da cena em duas categorias: uma mais abrangente (em verde) e outra aparecendo apenas em regiões de transição com outras classes (em azul escuro). O algoritmo também confundiu áreas de vegetação rala ou desmatada com padrões de área urbana, porém de forma bem mais acentuada que a metodologia de classificação proposta. Os dois padrões de água existentes na imagem teste somente são discriminados pelo algoritmo *K*-médias se considerarmos um número predefinido de classes maior ou igual a 5. No entanto, embora o algoritmo classifique a vegetação em várias categorias diferentes, a confusão entre as áreas de vegetação rala com os padrões de área urbana permanece para qualquer número de classes.

Realizando uma análise visual das imagens classificadas pelo método proposto e pelo algoritmo K-médias (figuras 4.44 e 4.45), observa-se que a imagem resultante da metodologia de classificação proposta apresenta um aspecto visual melhor. Enquanto a classificação pelo algoritmo K-médias possui uma aparência mais salpicada, a imagem produzida pelo método proposto se apresenta de forma mais homogênea em todas as áreas classificadas. Essa diferença no aspecto visual certamente se dá em virtude das formas pelas quais os dois métodos tratam a imagem. Enquanto o algoritmo K-médias utiliza uma abordagem pixel a pixel para realizar a classificação, o método proposto trabalha com janelas de pixels, as quais incorporam informações de vizinhança e, por isso, permitem obter uma classificação resultante de melhor qualidade.

O tempo consumido pelo algoritmo K-médias na classificação da imagem teste foi de 211 segundos, rodando sob o software MATLAB na mesma máquina utilizada para executar o método de classificação proposto.

Embora o algoritmo K-médias tenha apresentado um tempo de processamento inferior ao da metodologia proposta, é importante observar que os dois métodos possuem princípios e características bastante diferenciadas que devem ser consideradas na avaliação de suas complexidades de cálculo.

Na execução do algoritmo K-médias, todos os pontos da imagem são considerados seguindo uma abordagem pixel a pixel. Como a imagem teste possui três bandas espectrais, o algoritmo trabalha num espaço cujos dados possuem apenas 3 dimensões. Já o método proposto analisa a imagem através de um conjunto de amostras. No entanto, como as amostras no experimento realizado correspondem a janelas 5x5 de pixels, o método trabalha dentro de um espaço cujos dados têm 75 dimensões.

Apesar do algoritmo K-médias recalcular os centróides das classes em todas as iterações, o mesmo considera um único protótipo (representante) para cada uma delas. Já o método proposto trabalha com grupos de protótipos para representar cada uma das classes da imagem, conseqüentemente, a etapa final de classificação da cena, que consiste apenas em rotular os pixels de acordo com a classe do protótipo mais próximo, consome um maior tempo de processamento em decorrência do número de comparações que deve ser realizada entre cada padrão

da imagem e todos os protótipos das classes.

Além dessas diferenças em relação ao algoritmo K-médias, conforme descrito na próxima seção, o método proposto possui funcionalidades e vantagens importantes que certamente justificam e compensam o seu volume de processamento.

Outros resultados experimentais da metodologia de classificação proposta aqui, mostrando comparações com o algoritmo K-médias, podem ser encontrados em Gonçalves *et al.* (2005b).

## 4.4 Considerações finais

Neste capítulo foi apresentada uma metodologia de classificação de imagens de sensoria-mento remoto que realiza a análise de agrupamentos da imagem através da projeção de dados efetuada pelo SOM. Ao invés de trabalhar diretamente com os padrões originais da cena, o método proposto detecta as classes de cobertura terrestre da imagem através de SOMs treina-dos com janelas amostrais coletadas a partir da imagem a ser classificada.

Diferentemente de outros métodos de classificação, nos quais o SOM é utilizado pelo usuário como ferramenta de auxílio visual para descobrir as classes existentes na imagem, na metodo-logia de classificação proposta a detecção de agrupamentos da imagem através dos neurônios de um SOM treinado é realizada sem a necessidade de participação do usuário ou de um ana-lista de imagens. Para tanto, o método de segmentação automática do SOM proposto por Costa e Andrade Netto (1999, 2001) foi aplicado para particionar e rotular os SOMs após os seus treinamentos. Os resultados mostram que o método de segmentação é eficaz para parti-cionar os mapas de neurônios em diferentes classes de protótipos que representam os padrões de cobertura terrestre existentes na imagem original.

Além da eficácia e das potencialidades proporcionadas pelo método de segmentação do SOM aplicado, o processo de amostragem da imagem realizado por meio de janelas de pixels permite que a metodologia de classificação proposta apresente um desempenho superior aos das abordagens de classificação pixel a pixel. Testes experimentais mostram que não é possível discriminar classes de cobertura terrestre que apresentam padrões de cunho espacial (ou tex-turas diferenciadas) por meio de SOMs treinados com amostras coletadas na forma de pixels

individuais.

Ainda que a metodologia de classificação proposta possa demandar um tempo maior de processamento em função do número de SOMs que são gerados, em contrapartida, ela apresenta uma série de vantagens que não são proporcionadas pelos métodos convencionais de classificação não-supervisionada. Dentre essas vantagens destacam-se a não necessidade de definição prévia do número de classes da imagem a ser classificada, a potencialidade de descobrir agrupamentos com formas variadas (em virtude da representação distribuída por meio de grupos de protótipos) e a possibilidade de explorar visualmente os relacionamentos hierárquicos existentes entre as classes da imagem (através da árvore de mapas produzida pelo método de segmentação do SOM e das imagens dos arranjos de neurônios).

Resultados experimentais sobre uma imagem TM/LANDSAT mostram que a metodologia proposta pode ser mais eficaz na detecção automática dos padrões existentes na imagem em comparação com o algoritmo K-médias (comumente utilizado para realizar a classificação não-supervisionada de imagens de sensoriamento remoto).



# Capítulo 5

## Classificação baseada em SOMs segmentados por morfologia matemática e em índices de validação de agrupamentos

### 5.1 Introdução

Procurando diminuir o volume de processamento exigido pela metodologia de classificação apresentada no capítulo anterior, o presente capítulo propõe uma nova estratégia para analisar os agrupamentos de dados das imagens de sensoriamento remoto através do SOM. Utilizando SOMs segmentados por morfologia matemática em conjunto com índices de validação de agrupamentos, a estratégia apresenta-se como um aprimoramento do método de segmentação automática do SOM proposto por Costa e Andrade Netto (1999, 2001). A incorporação de índices de validação no processo de particionamento dos dados torna possível analisar a qualidade de diferentes configurações de agrupamentos que são geradas por diferentes segmentações de um SOM treinado. Resultados experimentais mostram que tal abordagem pode alcançar de maneira mais eficiente a melhor partição do conjunto de dados, uma vez que uma quantidade menor de SOMs é gerada durante o processo de detecção das classes.

Em virtude da grande diversidade de propostas de índices de validação de agrupamentos existentes na literatura, este capítulo apresenta um levantamento bibliográfico dos índices de

validação (*crisp*) mais comumente empregados na área de reconhecimento de padrões e dos mais recentemente propostos, buscando, dessa forma, pré-selecionar aqueles que apresentam algum potencial para serem aplicados na avaliação de agrupamentos de imagens de sensoriamento remoto.

Tendo em vista ainda, que a maioria dos índices de validação de agrupamentos apresentam complexidades computacionais elevadas, o que pode ser um complicador em aplicações que envolvem grandes volumes de dados (como é o caso das imagens de sensoriamento remoto), este capítulo também propõe uma estratégia para modificar os cálculos desses índices baseando-se na quantização vetorial produzida pelo SOM, reduzindo desse modo a complexidade computacional e viabilizando a aplicação dos mesmos.

Com o objetivo de validar a estratégia de modificação dos cálculos dos índices de validação e de determinar os mais apropriados para avaliar partições de dados de imagens de sensoriamento remoto, os índices pré-selecionados através da revisão bibliográfica realizada foram testados experimentalmente, em suas versões originais e modificadas, na avaliação de diferentes partições de conjuntos de dados com características variadas. Aquelos índices que apresentaram os melhores desempenhos na validação de todos os agrupamentos de dados considerados foram então empregados no método de classificação de imagens de sensoriamento remoto proposto neste capítulo.

Para gerar bordas mais definidas na *U-matrix* e, conseqüentemente, evitar a análise de partições de dados pouco significativas, este capítulo propõe ainda um procedimento baseado em medidas de textura para filtrar janelas amostrais que apresentam um alto grau de heterogeneidade espectral e degradam a superfície topográfica da *U-matrix*. Resultados experimentais mostram que a aplicação do procedimento pode melhorar o desempenho da metodologia de classificação proposta.

## 5.2 Análise de agrupamentos usando SOMs segmentados por morfologia matemática e índices de validação

Embora a metodologia de classificação apresentada no capítulo 4 tenha apresentado bons resultados, a necessidade de gerar e treinar vários SOMs durante o processo de análise de agrupamentos da imagem pode elevar substancialmente o seu tempo de processamento global. De fato, conforme descrito na seção 4.3.3, dentre todas as etapas da metodologia de classificação proposta, aquela que abrange os treinamentos dos SOMs é a que despense mais tempo de processamento.

Apesar de existirem abordagens alternativas voltadas exclusivamente para reduzir o tempo de processamento do algoritmo de aprendizagem da rede de Kohonen, o presente trabalho apresenta nesta seção uma outra estratégia de análise de agrupamentos através do SOM que pode demandar uma quantidade menor de mapas em comparação com a abordagem proposta por Costa e Andrade Netto (1999, 2001).

A estratégia de análise de agrupamentos proposta aqui procura encontrar a melhor partição para um conjunto de dados a partir da análise de diferentes segmentações realizadas sobre a *U-matrix*. Ao invés de definirmos apenas uma imagem de marcadores para segmentar a *U-matrix*, como proposto em Costa e Andrade Netto (1999, 2001), várias imagens de marcadores são determinadas e então diferentes segmentações da *U-matrix* são obtidas. Cada uma dessas segmentações é associada aos neurônios do SOM, obtendo-se assim diferentes partições do mapa de neurônios, as quais, conseqüentemente, definem diferentes configurações de agrupamentos para os dados. Para cada uma das diferentes configurações de agrupamentos é aplicado um índice de validação que serve como critério de decisão para a escolha da melhor partição. Aquela partição que apresenta o valor ótimo para o índice de validação utilizado é considerada a melhor. A figura 5.1 ilustra a estratégia de particionamento de dados proposta.

A estratégia apresentada pode ser vista como um aprimoramento do método proposto por Costa e Andrade Netto (1999, 2001) para realizar análises de agrupamentos de dados através do SOM. O algoritmo de segmentação de imagens, *watershed*, utilizado pelo método SL-SOM também é aplicado nesta nova abordagem para particionar o mapa de neurônios treinado.

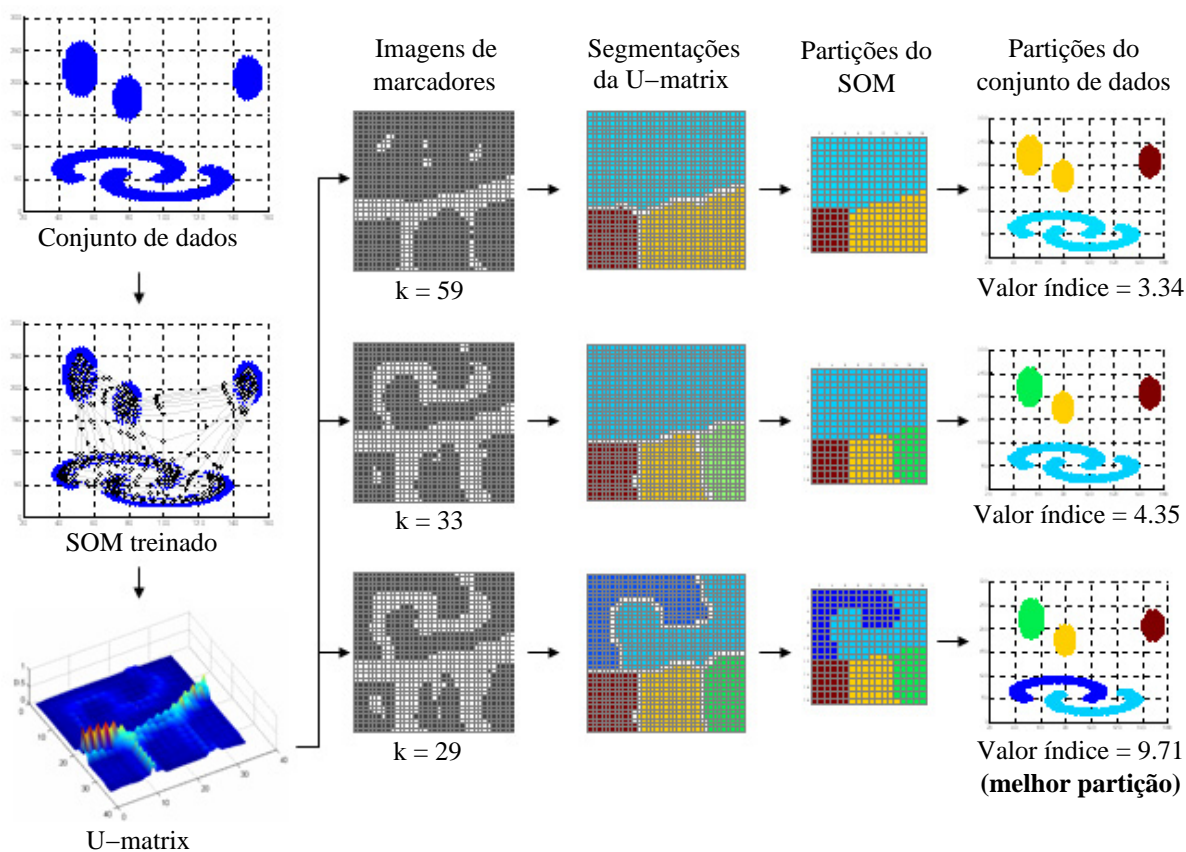


Fig. 5.1: Ilustração da estratégia de análise de agrupamentos proposta.

No entanto, a aplicação de índices de validação de agrupamentos na avaliação conjunta de diferentes partições do conjunto de dados possibilita que o método proposto aqui detecte mais rapidamente a melhor partição, sem a necessidade de gerar e treinar uma quantidade maior de SOMs.

A seguir, formaliza-se o procedimento para obter o conjunto de imagens de marcadores da *U-matrix* e o algoritmo da estratégia de análise de agrupamentos de dados proposta.

Considerando a *U-matrix* de um SOM treinado dada pela imagem  $U$  com 256 níveis de cinza, os seguintes passos são efetuados para obter o conjunto de imagens de marcadores para  $U$ :

1. Filtragem: obter uma imagem  $U_1$  suavizando a imagem  $U$  através da remoção de pequenas depressões;
2. Para  $k = 1, \dots, f_{max}$ , onde  $f_{max}$  é o nível de cinza máximo da imagem  $U_1$ , criar as imagens

binárias  $U_2^k$  correspondendo a conversões de  $U_1$ , usando  $k$  como valor de limiar;

3. Obter o número de regiões conectadas de  $U_2^k$ , para cada valor de  $k$ ,  $N_{rc}^k$ ;
4. Construir o gráfico  $k \times N_{rc}^k$  e, através dele, obter o conjunto  $S_k$  composto por todos os valores de  $k$  que correspondem a inícios de seqüências contíguas e constantes de números de regiões conectadas com tamanhos superiores a 3;
5. Obter o conjunto de todas as imagens de marcadores,  $S_m = \{U_2^{k_1}, U_2^{k_2}, \dots, U_2^{k_n}\}$ , onde  $k_1, k_2, \dots, k_n$  são os elementos de  $S_k$  obtidos no passo anterior.

Embora o procedimento usado para encontrar os marcadores da *U-matrix* seja semelhante ao apresentado na seção 3.7.1, nesta abordagem, os passos 4 e 5 determinam um conjunto de imagens de marcadores da *U-matrix* considerando todas as seqüências contíguas e constantes de regiões conectadas (com tamanhos superiores a 3), e não apenas uma única imagem de marcadores como proposto em Costa e Andrade Netto (1999, 2001).

Sendo assim, os passos da estratégia geral de análise de agrupamentos proposta aqui podem ser descritos como seguem:

1. Obtenha a *U-matrix* a partir do SOM treinado;
2. Encontre o conjunto de imagens de marcadores,  $S_m = \{U_2^{k_1}, U_2^{k_2}, \dots, U_2^{k_n}\}$ , para a *U-matrix*;
3. Para cada imagem de marcadores,  $U_2^{k_i}$ ,  $i = 1, 2, \dots, n$ , faça:
  - (a) Aplicar o algoritmo *watershed* sobre a *U-matrix*;
  - (b) Rotular as regiões conectadas da imagem segmentada no passo 3(a);
  - (c) Realizar a cópia dos rótulos obtidos no passo 3(b) para os neurônios do SOM associados a cada pixel da *U-matrix*;
  - (d) Rotular o conjunto de dados de entrada utilizando o SOM rotulado no passo 3(c);
  - (e) Aplicar um índice de validação de agrupamentos sobre o conjunto de dados rotulado no passo 3(d);

4. Escolher a partição de dados que apresenta o valor ótimo para o índice de validação de agrupamentos utilizado.

É importante observar que, de maneira similar ao método proposto por Costa e Andrade Netto (1999, 2001), o algoritmo descrito acima pode ser reaplicado caso existam ainda subclasses de dados nos agrupamentos encontrados.

Procurando facilitar a compreensão da estratégia de análise de agrupamentos apresentada, bem como das suas diferenças e vantagens em relação ao método original proposto por Costa e Netto (descrito na seção 3.7.1), o apêndice A.1 (inserido na mídia que acompanha essa tese) apresenta os resultados experimentais obtidos a partir da aplicação da estratégia proposta sobre o conjunto de dados apresentado na seção 3.7.1.1 (sobre o qual o método proposto por Costa e Netto também foi aplicado).

Outras considerações sobre as diferenças e vantagens do método de análise de agrupamentos apresentado aqui em relação ao algoritmo proposto por Costa e Andrade Netto (1999, 2001) são apresentadas na seção 5.4.1, onde mostra-se a aplicação do método na classificação de imagens de sensoriamento remoto. Outros resultados experimentais utilizando conjuntos de dados artificiais e reais podem ser vistos em Gonçalves *et al.* (2005a) e em Gonçalves e Andrade Netto (2005).

### 5.3 Análise, seleção e modificação dos cálculos de índices de validação

No contexto deste trabalho, os índices de validação são utilizados para avaliar partições de dados produzidas por um método que se baseia nos protótipos do SOM para realizar análises de agrupamentos. Embora o método proposto possa ser aplicado em conjuntos de dados de qualquer natureza, o foco principal deste trabalho é realizar a análise de agrupamentos (ou a classificação não-supervisionada) de imagens de sensoriamento remoto. Nesse sentido, a utilização de índices de validação que sejam particularmente adequados às características dos dados dessas imagens é de fundamental importância para o bom desempenho do método de

classificação proposto.

Contudo, quando se analisa a natureza ou as características das imagens de sensoriamento remoto, logo percebe-se que não é tarefa fácil encontrar um método para analisar seus agrupamentos de dados, bem como um índice de validação que seja totalmente adequado para avaliar as partições obtidas. Conforme relatado no trabalho de Tran *et al.* (2005), o tamanho das imagens, a alta dimensão dos seus espaços de atributos, a presença de ruídos, a sobreposição, o formato, a densidade e os tamanhos desiguais dos agrupamentos são alguns dos problemas enfrentados quando se realiza a detecção de agrupamentos em imagens de sensoriamento remoto. Portanto, selecionar um índice de validação que seja apropriado para esta aplicação implica, na prática, em selecionar um índice que seja apropriado para avaliar quaisquer tipos de agrupamentos.

Diante então da complexidade dos dados das imagens de sensoriamento, e considerando também a grande diversidade de propostas de índices de validação de agrupamentos, realizou-se neste trabalho um levantamento bibliográfico dos principais e dos mais recentes índices de validação (*crisp*) existentes na literatura, buscando dessa forma pré-selecionar aqueles que apresentam algum potencial para serem aplicados na avaliação de agrupamentos de imagens de sensoriamento remoto. Alguns trabalhos importantes, como os de Milligan e Cooper (1985) e de Maulik e Bandyopadhyay (2002), apresentam análises comparativas entre vários índices de validação, as quais certamente servem de base para selecionar aqueles que apresentam melhores desempenhos na avaliação de conjuntos de dados diversos.

Nesse sentido, conforme será apresentado na subseção 5.3.3, os índices pré-selecionados através da revisão bibliográfica realizada aqui foram testados experimentalmente na avaliação de diferentes partições de conjuntos de dados com características variadas. Os índices que apresentaram os melhores desempenhos na validação de todos os agrupamentos de dados considerados foram, então, empregados no método de classificação de imagens de sensoriamento remoto proposto neste capítulo.

Conforme destacado ainda em Halkidi e Vazirgiannis (2002), apesar da grande variedade de índices de validação propostos na literatura, a implementação da maioria deles é computacionalmente muito intensiva, especialmente em aplicações que apresentam um grande volume

de dados, como é o caso das imagens de sensoriamento remoto. Procurando contornar esse problema, o presente trabalho também apresenta na seção 5.3.2 uma proposta de modificação dos cálculos de índices de validação de agrupamentos baseando-se na abstração de dados produzida pelo SOM, buscando assim diminuir o volume de processamento dos mesmos. Resultados de experimentos realizados para avaliar a estratégia proposta de alteração nos cálculos dos índices são apresentados na subseção 5.3.3.

### 5.3.1 Levantamento bibliográfico e índices pré-selecionados

Quantas propostas de índices de validação de agrupamentos (*crisp*) existem na literatura? Conforme relatado em Bezdek e Pal (1998), a pouco mais de vinte anos atrás, Hubert e Arabie iniciaram um artigo sobre este assunto fazendo a seguinte colocação: "*Nós não tentaremos revisar esta literatura de forma abrangente, pois esta é uma tarefa que pode requerer o tamanho de uma monografia*" (Hubert e Arabie (1985), p. 193).

De fato, na literatura existe uma quantidade bastante expressiva de propostas de índices de validação de agrupamentos, a qual exige a adoção de algum critério quando se pretende realizar uma revisão mais concisa dos métodos existentes. Diante disso, a revisão bibliográfica apresentada nesta seção procurou concentrar-se principalmente naqueles índices de validação que têm sido empregados com mais frequência na área de reconhecimento de padrões, não deixando de investigar propostas mais recentes e de analisar também trabalhos importantes que enfocam comparações de desempenho entre índices de validação diferentes.

Ao realizar uma busca na literatura pelos índices de validação mais comumente empregados, nota-se de imediato que o índice de Dunn (proposto em Dunn (1974)) é um deles. Este índice é adequado para a avaliação de agrupamentos compactos e bem separados, e embora tenha sido proposto há um tempo considerável, provê um estrutura geral apropriada para a definição de variantes que sejam capazes de avaliar diferentes tipos de agrupamentos (Bezdek e Pal, 1998). No entanto, o índice de Dunn apresenta a desvantagem de ser muito sensível à presença de ruídos, podendo gerar validações erradas nesse caso. Na literatura existem algumas propostas para suprir essa deficiência do índice, entre elas destacam-se Pal e Biswas (1997), Bezdek e Pal (1998) e Azuaje (2002). Em Azuaje (2002), os autores utilizaram uma ferramenta que



possui dezoito versões diferentes do índice de Dunn, contendo combinações variadas de distâncias inter-agrupamentos e intra-agrupamentos. Em Pal e Biswas (1997), são apresentadas generalizações do índice de Dunn usando conceitos da Teoria de Grafos e, embora os autores ressaltem que o índice de Dunn não deva ser considerado totalmente ineficiente, os resultados mostram que as generalizações usando estruturas de grafos apresentam desempenhos superiores aos da versão original do índice, principalmente em conjuntos de dados que apresentam formas estruturadas como linhas e arcos, por exemplo. Em Bezdek e Pal (1998), os autores apresentam generalizações do índice de Dunn utilizando diferentes funções de distâncias para medir a dispersão dos elementos em cada agrupamento e a separação entre os grupos. De acordo com os resultados, cinco dos índices de Dunn generalizados apresentaram melhores desempenhos do que o índice de Dunn original. Comparando as generalizações de Dunn apresentadas em Pal e Biswas (1997) e em Bezdek e Pal (1998), verifica-se que aquelas apresentadas no trabalho de Bezdek e Pal exigem um menor volume de processamento do que aquelas baseadas em estruturas de grafos descritas por Pal e Biswas. Por outro lado, em Bezdek e Pal (1998), os autores ressaltam que as generalizações de Dunn propostas por eles não são adequadas para agrupamentos curvilíneos ou que se apresentam em cadeia (os autores citam as generalizações de Pal e Biswas (1997) como sendo adequadas para esses casos).

Outro índice de validação tão popular quanto o índice de Dunn é o índice Davies-Bouldin, proposto em Davies e Bouldin (1979). Este índice utiliza os centróides dos agrupamentos para avaliar a dispersão intra-agrupamento e a separação entre os agrupamentos, e apresenta desempenho satisfatório apenas quando é aplicado para avaliar conjuntos de dados que possuem agrupamentos com formatos hipersféricos. Tentando expandir a capacidade do índice para agrupamentos com formatos arbitrários, Pal e Biswas (1997) além de apresentarem generalizações baseadas na teoria de grafos para o índice de Dunn, também propuseram generalizações para o índice Davies-Bouldin. Os resultados experimentais mostram que as generalizações propostas apresentaram, em alguns casos, desempenhos superiores aos da versão original do índice na avaliação de diferentes tipos de agrupamentos. Mais recentemente, Kim e Ramakrishna (2005) propuseram também uma modificação para o cálculo do índice Davies-Bouldin alterando a sua fórmula principal e mantendo as medidas originais de dispersão e a separação

de agrupamentos. Os resultados do trabalho mostram que houve melhora no desempenho do índice para conjuntos de dados que apresentam agrupamentos com algum grau de sobreposição.

Em Milligan e Cooper (1985), é apresentada uma avaliação de trinta índices de validação utilizando quatro métodos de agrupamentos hierárquicos diferentes. De acordo com os resultados, os índices Calinski-Harabasz (Calinski e Harabasz, 1974),  $Je(2)/Je(1)$  (Duda e Hart, 1973) e C (Hubert e Schultz, 1976) foram aqueles que apresentaram os melhores desempenhos. No entanto, a avaliação realizada por Milligan e Cooper se baseou apenas em conjuntos de dados pequenos e com agrupamentos bem separados. De acordo com Sarle e Kuo (1993) e Halkidi *et al.* (2001), os índices que apresentaram bons desempenhos nessa avaliação são adequados apenas para agrupamentos compactos ou levemente alongados, e que tenham preferencialmente uma distribuição normal.

Já em Milligan e Cooper (1986), os índices de validação Rand, Jaccard, Fowlkes e Mallows, e as duas modificações do índice de Rand propostas por Morey e Agresti (1984) e por Hubert e Arabie (1985), foram avaliados na análise de agrupamentos hierárquicos. Os resultados indicaram que a modificação do índice Rand proposta por Hubert e Arabie (1985) foi a mais apropriada para o tipo de análise realizada, porém apresenta um elevado custo computacional devido aos cálculos de probabilidades exigidos.

A estatística de Hubert modificada, proposta em (Hubert e Arabie, 1985), também é comumente empregada para a validação de agrupamentos. No entanto, numa comparação apresentada em Bezdek e Pal (1998) entre este método e o índice Davies-Bouldin, os resultados mostram que, embora ambos os índices apresentem eficácias parecidas, a identificação visual das curvas de Hubert é muito subjetiva e dependente de escala, o que leva os autores a concluir que o índice Davies-Bouldin é preferível em relação a estatística de Hubert. Além disso, o método estatístico apresentado por Hubert não é apropriado para agrupamentos com formas arbitrárias.

Em Rousseeuw (1987), é proposto um outro índice de validação, o *Silhouette Width*, que calcula a medida de silhueta para cada padrão que faz parte de um agrupamento. As silhuetas mostram quais padrões estão bem situados dentro dos seus agrupamentos e quais estão fora de um agrupamento apropriado. No entanto, conforme descrito em Jörnsten (2004), o índice

apresenta alguns problemas e se desempenha bem somente quando os agrupamentos avaliados são aproximadamente esféricos, compactos e bem separados.

No trabalho de Boudraa (1999), é proposto o índice de validação denominado *Bcrit*, que consiste de uma combinação de duas medidas de validação, uma delas relacionada com as propriedades geométricas dos dados e a outra com a homogeneidade dos dados em cada agrupamento. O índice *Bcrit* apresenta um desempenho satisfatório para os conjuntos de dados analisados, porém Boudraa ressalta que o desempenho dele pode ser afetado quando os dados apresentam algum ruído.

Em Turi (2001), é apresentado um índice de validação que incorpora uma função multiplicadora para a razão entre as distâncias intra-agrupamentos e inter-agrupamentos. Nos resultados apresentados, o índice obteve um melhor desempenho que os índices Dunn e Davies-Bouldin, porém a função multiplicadora utilizada pelo índice apresenta um parâmetro que necessita ser especificado pelo usuário e que pode influenciar significativamente nos resultados.

Conforme relatado em Berry e Linoff (1996) e em Halkidi *et al.* (2002), grande parte das abordagens de validação de agrupamentos existentes obtêm melhores desempenhos quando os agrupamentos de dados são mais compactos e apresentam formatos hiperesféricos. Para aplicações que apresentam agrupamentos com sobreposições e formatos arbitrários, os critérios tradicionais de validação (variância, densidade e separação) não são suficientes para realizar uma avaliação adequada. É por isso que na literatura mais recente, novos índices de validação têm sido introduzidos adotando medidas e mecanismos para avaliar não apenas a compacidade e a separação dos agrupamentos, mas também a geometria dos mesmos. É o caso do trabalho apresentado por Halkidi e Vazirgiannis (2002), onde os autores propuseram um índice de validação, o CD<sub>bw</sub> (*Composing Density Between and Within Clusters*), que apresenta algumas características que o torna mais eficiente em relação a outros índices que se baseiam apenas em critérios de compacidade e de separação dos dados. As características geométricas dos agrupamentos são representadas pelo uso de vetores representativos, isso permite que o CD<sub>bw</sub> avalie de maneira correta estruturas que não tenham apenas formas hiperesféricas. Em Halkidi e Vazirgiannis (2002), os autores compararam o desempenho do CD<sub>bw</sub> com outros índices de validação, entre eles, o RMSSTD (Sharma, 1996), SD (Halkidi *et al.*, 2000) e o *S\_Dbw* (Halkidi

*et al.*, 2001). Em todos os experimentos realizados, o índice CDbw encontrou o número ótimo de agrupamentos adaptando-se corretamente aos conjuntos de dados, enquanto que os demais índices de validação falharam em alguns casos.

Veenman *et al.* (2002) propuseram um índice de validação que minimiza a variabilidade intra-agrupamento enquanto usa um critério de variância mínima para avaliar a união de dois agrupamentos. O índice apresenta bons desempenhos nos experimentos apresentados, mas da mesma forma que o índice proposto em Turi (2001), o índice de Veenman *et al.* (2002) possui um parâmetro que também deve ser definido pelo usuário e que influencia significativamente no resultado final das avaliações.

No trabalho de Bandyopadhyay e Maulik (2001), é proposto um índice de validação de agrupamentos chamado PBM, que consiste de um produto de três fatores calculados em função da dispersão, separação e quantidade dos agrupamentos. Em Bandyopadhyay e Maulik (2001), em Maulik e Bandyopadhyay (2002) e também em Pakhira *et al.* (2004) são apresentados testes experimentais comparando o índice PBM com outros índices de validação, entre eles, o índice de Dunn, Davies-Bouldin e Calinski-Harabasz. O índice PBM obteve um desempenho superior para todos os conjuntos de dados artificiais e reais considerados nos experimentos.

Em uma contribuição mais recente, Tasdemir e Merenyi (2007) propuseram um novo índice de validação para métodos de agrupamentos baseados em protótipos, chamado Conn. O índice se baseia em densidades inter e intra-agrupamentos que são determinadas através de uma matriz de conectividade entre os protótipos. De acordo com os autores, o índice pode ser aplicado em quaisquer conjuntos de dados, com agrupamentos de tamanhos e formatos variados e também com sobreposições. Em Tasdemir e Merenyi (2007), o SOM foi empregado para obter os protótipos dos agrupamentos, mas de acordo com os autores outros métodos de análise de agrupamentos podem ser utilizados. Os resultados mostram que o índice apresenta bons desempenhos em comparação com os índices de Dunn, Davies-Bouldin e CDbw. Entretanto, em configurações de agrupamentos onde existem protótipos não agrupados, o índice Conn não provê uma avaliação adequada.

Ao pesquisar e analisar as diversas propostas de índices de validação de agrupamentos existentes na literatura, conclui-se que dificilmente há um índice de validação que seja totalmente

adequado para avaliar quaisquer tipos de agrupamentos. Cada um deles apresenta critérios de validação diferenciados e vantagens e desvantagens sob diferentes aspectos.

Sendo assim, tendo como base a revisão bibliográfica efetuada aqui, realizou-se uma pré-seleção que resultou na escolha de sete índices de validação que apresentam algum potencial para avaliar agrupamentos de dados de imagens de sensoriamento. Esses índices foram escolhidos considerando principalmente as suas capacidades de avaliar agrupamentos com formatos variados, os seus bons desempenhos em análises comparativas com outros índices de validação (comprovados nos trabalhos abordados nesta revisão) e a não necessidade de interação com o usuário para a especificação de parâmetros. Os sete índices pré-selecionados foram:

- Índice de Dunn generalizado proposto por Pal e Biswas (1997) ( $\mathcal{V}_{D_{Pal}}$ ),
- Índice de Dunn generalizado proposto por Bezdek e Pal (1998) ( $\mathcal{V}_{D_{Bdk}}$ ),
- Índice de Davies-Bouldin generalizado proposto por Pal e Biswas (1997) ( $\mathcal{V}_{DB_{Pal}}$ ),
- Índice de Davies-Bouldin generalizado proposto por Kim e Ramakrishna (2005) ( $\mathcal{V}_{DB_{Kim}}$ ),
- Índice PBM proposto por Pakhira *et al.* (2004) ( $\mathcal{V}_{PBM}$ ),
- Índice CDbw proposto por Halkidi e Vazirgiannis (2002) ( $\mathcal{V}_{CDbw}$ ), e
- Índice Conn proposto por Tasdemir e Merenyi (2007) ( $\mathcal{V}_{Conn}$ ).

Tendo em vista que na literatura não há nenhum trabalho que apresente uma análise experimental e comparativa entre os sete índices de validação pré-selecionados nesta revisão, na seção 5.3.3 são apresentados testes experimentais aplicando os sete índices sobre conjuntos de dados com características variadas. O objetivo é verificar quais dentre eles apresentam os melhores desempenhos para diferentes padrões de agrupamentos.

Uma breve descrição dos setes índices de validação pré-selecionados é apresentada no apêndice A.2 (inserido na mídia que acompanha essa tese).

### 5.3.2 Modificação dos cálculos de índices de validação

Apesar da existência de um grande número de índices de validação de agrupamentos com características e propriedades variadas, a maior parte deles apresenta um problema comum: a alta complexidade computacional. Quando o número de agrupamentos e/ou o volume de dados a ser analisado é muito grande, o uso de índices de validação de agrupamentos se torna proibitivo. As imagens de sensoriamento remoto são um exemplo típico. Como descrito em Ji (2003), índices de validação de agrupamentos são freqüentemente ignorados na análise de imagens de sensoriamento remoto e não estão disponíveis na maioria dos pacotes ou softwares de processamento digital de imagens.

De fato, na literatura de sensoriamento remoto são raros os trabalhos que apresentam a aplicação direta de índices de validação na análise de agrupamentos de imagens. Quando isso é feito, nota-se que o índice de validação utilizado apresenta uma baixa complexidade de cálculo, mas não possui potencial para avaliar quaisquer tipos de agrupamentos. Nos trabalhos de Tran *et al.* (2003) e Marçal e Borges (2005), por exemplo, o índice de validação Davies-Bouldin foi empregado para avaliar diferentes partições de dados de imagens de satélite, uma vez que apresenta, segundo os autores, uma complexidade computacional adequada para a aplicação. No entanto, conforme mencionado na seção 5.3.1, o índice Davies-Bouldin é apropriado apenas para agrupamentos de dados que possuem formas hiperesféricas. Caso os agrupamentos avaliados apresentem outros padrões de forma ou geometria, a aplicação deste índice pode comprometer a interpretação dos resultados.

Diante disso, neste trabalho propõe-se uma maneira de reduzir o alto custo computacional exigido pelos índices de validação de agrupamentos, tornando possível aplicá-los em conjuntos de dados volumosos como as imagens de sensoriamento remoto. A idéia consiste basicamente em modificar os cálculos dos índices de validação usando a quantização vetorial produzida pelo SOM. Considerando que nos métodos de classificação de imagens propostos neste trabalho, a análise de agrupamentos não é feita a partir dos padrões originais da imagem, e sim por um conjunto reduzido de protótipos do SOM (que os representam), a aplicação de um índice de validação para avaliar as diferentes partições obtidas torna-se viável, uma vez que a quantidade de dados a ser avaliada é bem menor que a quantidade de padrões extraída diretamente da

imagem original. Ou seja, ao invés de usarmos o conjunto de vetores de dados amostrais da imagem, apenas os vetores de pesos sinápticos dos neurônios do SOM são utilizados para realizar os cálculos dos índices de validação.

Embora o SOM efetue uma boa aproximação do espaço de entrada, é certo que a estratégia de aplicar um índice de validação sobre os protótipos da rede, ao invés de aplicá-lo diretamente sobre os padrões originais, pode causar diferenças numéricas nos resultados. Para diminuir possíveis erros de aproximação, propõe-se que os índices sejam calculados utilizando não apenas os vetores de pesos ( $w_i$ ) dos protótipos do SOM, mas também o nível de atividade ( $h(w_i)$ ) de cada um deles, ou seja, o número de padrões de entrada que foram associados a cada protótipo pelo processo de mapeamento do SOM.

Um exemplo das modificações aplicadas nos cálculos dos índices de validação pode ser descrito através da equação 5.1. A fórmula, e variantes dela, são comumente usadas em alguns índices de validação de agrupamentos para calcular a distância ( $\delta_{ij}$ ) entre dois agrupamentos de dados  $C_i$  e  $C_j$  (Bezdek e Pal, 1998):

$$\delta_{ij} = \frac{1}{|C_i| |C_j|} \sum_{x \in C_i, y \in C_j} d(x, y) \quad (5.1)$$

Na fórmula acima,  $d(*, *)$  é uma medida de distância, e  $|C_i|$  e  $|C_j|$  representam o número de pontos nos agrupamentos de dados  $C_i$  e  $C_j$ , respectivamente. Portanto,  $\delta_{ij}$  é determinada basicamente através dos cálculos de distância entre todos os pares de pontos  $x$  e  $y$  pertencentes aos agrupamentos de dados  $C_i$  e  $C_j$ . Se  $|C_i|$  e  $|C_j|$  são muito grandes, a computação de  $\delta_{ij}$  torna-se excessivamente custosa.

Aplicando a proposta de modificação de cálculo dos índices, a computação da distância entre os agrupamentos  $C_i$  e  $C_j$ , equivalente a equação 5.1, fica formulada da seguinte maneira:

$$\delta_{ij}^{SOM} = \frac{1}{|C_i| |C_j|} \sum_{w_i \in W_i, w_j \in W_j} h(w_i) \cdot h(w_j) \cdot d(w_i, w_j) \quad (5.2)$$

na qual  $W_i$  e  $W_j$  são os conjuntos de protótipos do SOM que representam os agrupamentos de dados  $C_i$  e  $C_j$  respectivamente,  $d(*, *)$  é a mesma medida de distância utilizada na equação 5.1,

$h(w_i)$  é o nível de atividade do protótipo  $w_i$  pertencente a  $W_i$  e  $h(w_j)$  é o nível de atividade do protótipo  $w_j$  pertencente a  $W_j$ .

Enquanto  $\delta_{ij}$  baseia-se nas distâncias entre todos os pares de pontos pertencentes a  $C_i$  e  $C_j$ ,  $\delta_{ij}^{SOM}$  está baseada somente nas distâncias entre os pares de pontos pertencentes a  $W_i$  e  $W_j$ . Portanto, nota-se que  $\delta_{ij}^{SOM}$  é um cálculo aproximado de  $\delta_{ij}$ , mas com um custo de processamento muito mais baixo, já que as quantidades  $|W_i|$  e  $|W_j|$  são geralmente muito menores do que  $|C_i|$  e  $|C_j|$ . A inclusão dos níveis de atividade  $h(\cdot)$  dos protótipos no cálculo de  $\delta_{ij}^{SOM}$  ajuda a diminuir os erros ocasionados pelo uso da quantização vetorial produzida pelo SOM, uma vez que incorpora de maneira aproximada a densidade de pontos do espaço de entrada que são representados por cada protótipo da rede.

Seguindo essa proposta, foram modificados os cálculos dos índices de validação pré-selecionados na sessão 5.3.1, exceto a do índice Conn, uma vez que em sua versão original esse índice já se baseia somente no uso de protótipos.

Na próxima seção, as modificações dos índices são testadas na avaliação de diferentes partições de conjuntos de dados e os seus resultados são comparados com aqueles obtidos pelas suas versões originais.

O apêndice A.3, localizado na mídia que acompanha essa tese, apresenta a descrição das versões modificadas dos sete índices de validação pré-selecionados na sessão 5.3.1.

### 5.3.3 Testes experimentais

Com o objetivo principal de avaliar a proposta de modificação dos cálculos dos índices de validação, a presente seção mostra a aplicação das versões modificadas dos índices na avaliação de partições de diferentes conjuntos de dados. Em segundo plano, os resultados dos experimentos realizados também são considerados para verificar quais dentre os sete índices de validação pré-selecionados na seção 5.3.1 apresentam os melhores desempenhos.

Para a execução dos experimentos, a seguinte metodologia foi delineada:

1. Conjuntos de dados apresentando agrupamentos com características variadas (entre elas, o tamanho, a dimensionalidade, o formato, a densidade e os graus de sobreposição e de



- ruído) foram selecionados para os testes;
2. A análise de agrupamentos de cada um dos conjuntos de dados foi efetuada aplicando a estratégia proposta na seção 5.2. Sendo assim, um SOM foi treinado para cada conjunto de dados e diferentes partições para os mesmos foram obtidas a partir de diferentes segmentações da *U-matrix* correspondente;
  3. Os sete índices de validação de agrupamentos pré-selecionados na seção 5.3.1, e suas versões modificadas (com exceção a do índice Conn), foram aplicadas para avaliar as diferentes partições obtidas na etapa anterior;
  4. Uma comparação dos valores e dos tempos de processamento apresentados pelas versões originais e modificadas de cada um dos índices foi realizada, observando também quais índices indicaram corretamente a partição ótima dos conjuntos de dados analisados.

Dez conjuntos de dados foram utilizados nos experimentos (6 deles são sintéticos e 4 compostos por dados reais). A tabela 5.1 mostra o tipo, o tamanho (número de pontos), a dimensão e o número de agrupamentos de cada um dos conjuntos. A figura 5.2 mostra os gráficos dos 6 conjuntos de dados sintéticos. O conjunto *Gauss\_1* apresenta cinco agrupamentos bem separados, gerados através de distribuições gaussianas. O conjunto *Gauss\_2* possui os mesmos dados do conjunto *Gauss\_1* acrescido de 200 pontos (ruído) gerados de maneira aleatória e uniforme. O conjunto *Gauss\_3* também foi gerado com base no conjunto *Gauss\_1*, porém as variâncias das distribuições de cada agrupamento foram aumentadas diminuindo o grau de separação entre eles. Os conjuntos de dados *Rand\_1* e *Rand\_2* apresentam agrupamentos com formas geométricas variadas. Já o conjunto *Rand\_3* possui agrupamentos com formatos circulares e elipsoidais, mas com densidades diferentes. Dentre os conjuntos de dados reais estão os conjuntos *WBC* (*Wisconsin Breast Cancer*) e *Wine* (Blake e Merz, 1998). O conjunto *WBC* possui 683 pontos num espaço de 9 dimensões, os quais são divididos em duas classes (maligna e benigna) não linearmente separáveis. O conjunto de dados *Wine* tem 178 pontos num espaço de 13 dimensões com três classes conhecidas. Os conjuntos de dados *SR\_1* e *SR\_2* são formados por dados de imagens de sensoriamento remoto. O conjunto *SR\_1* possui 4176 vetores de dados que correspondem a janelas 3x3 de pixels. Essas janelas amostrais foram coletadas

de um segmento de imagem composto por 3 bandas espectrais do satélite CBERS-2/CCD. A imagem abrange uma área do estado do Amazonas (Brasil) e apresenta 3 classes de cobertura terrestre com sobreposições. O conjunto  $SR\_2$  apresenta 6994 janelas de pixels (5x5) coletadas de um segmento de imagem composto por 3 bandas espectrais do satélite LANDSAT/TM. A cena corresponde a uma área da região nordeste do estado de São Paulo e possui 4 classes que também se sobrepõem umas com as outras. Os segmentos de imagens de satélite a partir dos quais os conjuntos de dados  $SR\_1$  e  $SR\_2$  foram extraídos, podem ser visualizados no apêndice A.4 (localizado na mídia anexada a essa tese).

Nome	Tipo	Tamanho	Dimensão	Agrupamentos
<i>Gauss_1</i>	Sintético	1250	2	5
<i>Gauss_2</i>	Sintético	1450	2	5
<i>Gauss_3</i>	Sintético	1250	2	5
<i>Rand_1</i>	Sintético	1524	2	4
<i>Rand_2</i>	Sintético	1711	2	8
<i>Rand_3</i>	Sintético	1050	2	6
<i>WBC</i>	Real	683	9	2
<i>Wine</i>	Real	178	13	3
<i>SR_1</i>	Real	4176	27	3
<i>SR_2</i>	Real	6994	75	4

Tab. 5.1: Descrição dos conjuntos de dados utilizados nos testes.

A tabela 5.2 mostra os tamanhos dos SOMs utilizados para cada conjunto de dados e o número de partições diferentes obtidas a partir da aplicação da estratégia de análise de agrupamentos proposta na seção 5.2. Todos os SOMs foram treinados utilizando a mesma configuração de parâmetros descrita na seção 4.2.2.

Os experimentos foram realizados no software MATLAB em uma máquina com processador AMD Athlon XP 2600+ 1,91GHZ e 1GB de RAM. Os tempos de processamento consumidos pelas versões originais e modificadas dos índices na avaliação de todas as partições de cada conjunto de dados são mostrados nas tabelas 5.3 e 5.4, respectivamente. A última linha da tabela 5.4 mostra o *speedup* médio apresentado pela versão modificada de cada índice em relação à versão original considerando todos os conjuntos de dados. O apêndice A.5, inserido na mídia que acompanha essa tese, mostra mais detalhadamente os resultados desses experimentos.

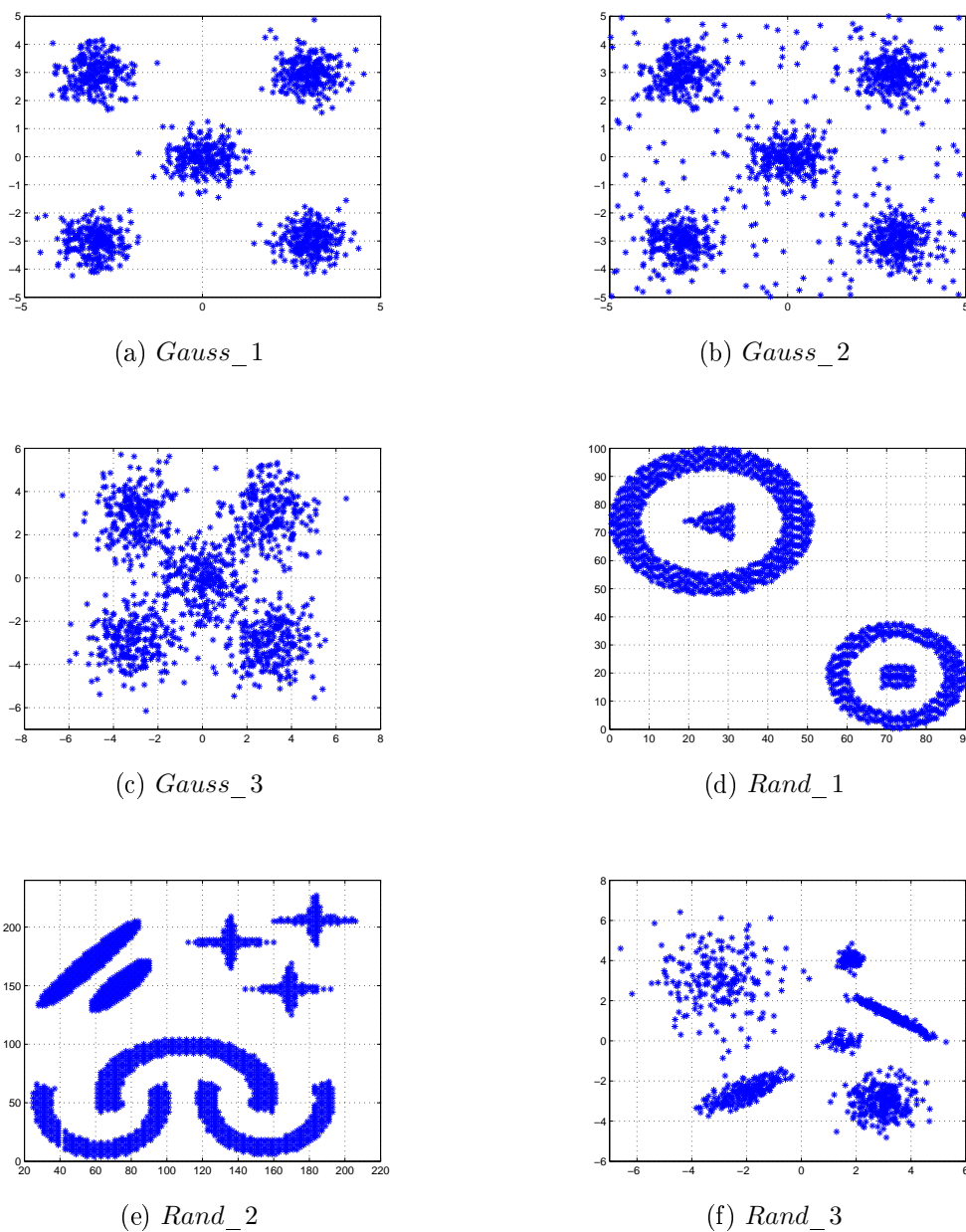


Fig. 5.2: Conjunto de dados sintéticos utilizados nos experimentos.

Analisando os tempos despendidos pelas versões originais dos índices (tabela 5.3), verifica-se que os índices  $\mathcal{V}_{DPal}$  e  $\mathcal{V}_{DBPal}$  são aqueles que consomem mais tempo de processamento. Isso ocorre em virtude da complexidade de cálculo do Grafo de Gabriel, que serve de base para a computação desses dois índices. Em seguida, em ordem decrescente de complexidade, estão os índices  $\mathcal{V}_{Conn}$ ,  $\mathcal{V}_{CDbw}$ ,  $\mathcal{V}_{DBdk}$ ,  $\mathcal{V}_{DBKim}$  e  $\mathcal{V}_{PBM}$ . Entre os sete índices de validação analisados,

Conjunto de dados	Tamanho do SOM	Número de partições
<i>Gauss_1</i>	12x12	4
<i>Gauss_2</i>	13x13	4
<i>Gauss_3</i>	12x12	4
<i>Rand_1</i>	14x14	5
<i>Rand_2</i>	15x15	6
<i>Rand_3</i>	10x10	6
<i>WBC</i>	6x6	3
<i>Wine</i>	5x5	3
<i>SR_1</i>	20x20	3
<i>SR_2</i>	22x22	4

Tab. 5.2: Tamanhos dos SOMs utilizados e o número de partições obtidas na aplicação da estratégia de análise de agrupamentos proposta na seção 5.2 em cada conjunto de dados.

Conjuntos de dados	$\mathcal{V}_{DPal}$	$\mathcal{V}_{DBdk}$	$\mathcal{V}_{DBPal}$	$\mathcal{V}_{DBKim}$	$\mathcal{V}_{PBM}$	$\mathcal{V}_{Cdbw}$	$\mathcal{V}_{Conn}$
<i>Gauss_1</i>	94.32	4.6175	95.16	4.1155	4.1543	11.614	7.0634
<i>Gauss_2</i>	125.34	7.5126	122.876	4.1036	4.1581	11.086	8.2047
<i>Gauss_3</i>	92.575	4.5492	98.781	4.113	4.1292	10.757	6.6651
<i>Rand_1</i>	250.33	8.0501	252.12	4.1407	4.1618	15.579	32.977
<i>Rand_2</i>	341.91	13.981	345.49	6.1652	6.219	18.764	45.005
<i>Rand_3</i>	80.663	5.3898	81.077	6.1677	6.2077	13.277	5.0906
<i>WBC</i>	20.686	1.6187	20.595	2.9252	2.41	2.1466	0.2543
<i>Wine</i>	0.3390	0.1175	0.3409	0.0773	0.1108	0.2297	0.0680
<i>SR_1</i>	8170.2	73.01	7410.7	15.232	16.08	118.14	155.27
<i>SR_2</i>	13090.3	281.96	13613.4	26.07	27.68	326.61	702.17

Tab. 5.3: Tempos de processamento (em segundos) consumidos pelas versões originais dos índices na avaliação de todas as partições de cada conjunto de dados.

os índices  $\mathcal{V}_{DBKim}$  e  $\mathcal{V}_{PBM}$  são, portanto, aqueles que exigem o menor custo de processamento, apresentando complexidades parecidas. Apesar de estar baseado apenas no uso de protótipos e não nos dados originais dos conjuntos, o índice  $\mathcal{V}_{Conn}$  apresenta um custo de processamento relativamente elevado em comparação com a maioria dos outros índices, principalmente quando o número de protótipos (tamanho do SOM) é maior, como nos casos dos conjuntos de dados *Rand\_2*, *SR\_1* e *SR\_2*. Tal comportamento é consequência da complexidade de cálculo da matriz de adjacência cumulativa (*CADJ*), a qual depende diretamente do número de protótipos considerado e serve de base para todos os cálculos intermediários do índice  $\mathcal{V}_{Conn}$ .

Conjuntos de dados	$\mathcal{V}_{DPal}^{SOM}$	$\mathcal{V}_{DBak}^{SOM}$	$\mathcal{V}_{DBPal}^{SOM}$	$\mathcal{V}_{DBKim}^{SOM}$	$\mathcal{V}_{PBM}^{SOM}$	$\mathcal{V}_{Cdbw}^{SOM}$
<i>Gauss_1</i>	0.2846	0.1963	0.2826	0.1025	0.1223	0.5023
<i>Gauss_2</i>	0.2861	0.1861	0.2939	0.1015	0.1015	0.4041
<i>Gauss_3</i>	0.2685	0.1691	0.212	0.0997	0.1034	0.4184
<i>Rand_1</i>	0.8873	0.3363	0.8491	0.1305	0.1286	1.2023
<i>Rand_2</i>	0.8777	0.4307	0.8893	0.1326	0.1636	1.3827
<i>Rand_3</i>	0.2442	0.1873	0.2304	0.1301	0.1424	0.4234
<i>WBC</i>	0.0815	0.0756	0.0801	0.0723	0.0731	0.0922
<i>Wine</i>	0.0783	0.0495	0.0764	0.0379	0.0609	0.0828
<i>SR_1</i>	3.5227	0.4258	3.3921	0.1108	0.1111	2.2541
<i>SR_2</i>	4.3849	1.344	4.334	0.1604	0.1956	5.2779
<b><i>Speedup (médio)</i></b>	<b>767.90</b>	<b>58.21</b>	<b>784.54</b>	<b>55.33</b>	<b>52.506</b>	<b>27.447</b>

Tab. 5.4: Tempos de processamento (em segundos) consumidos pelas versões modificadas dos índices na avaliação de todas as partições de cada conjunto de dados e média dos *speedups*.

De um modo geral, como era esperado, os índices consomem um tempo de processamento considerável para os conjuntos de dados que possuem alta dimensionalidade e um número grande de pontos (no caso, *SR\_1* e *SR\_2*). Cabe ressaltar que essa observação é válida até mesmo para os índices  $\mathcal{V}_{DBKim}$  e  $\mathcal{V}_{PBM}$  que exigiram tempos de processamento bem menores, pois em algumas análises de dados nas quais o número de partições a serem avaliadas é muito grande (como aquelas que são apresentadas no próximo capítulo), o tempo total que será gasto pelas versões originais desses dois índices também pode inviabilizar a aplicação da metodologia de classificação.

Por outro lado, as versões modificadas de todos os índices apresentaram tempos de processamento bastante reduzidos, bem inferiores aos das versões originais. Na última linha da tabela 5.4 pode-se observar que as taxas de *speedups* alcançadas com a aplicação da estratégia de modificação dos cálculos dos índices (proposta na seção 5.3.2) são significativas. Em particular, os índices  $\mathcal{V}_{DPal}^{SOM}$  e  $\mathcal{V}_{DBPal}^{SOM}$  apresentam taxas de *speedups* bem superiores as dos outros. Isso se dá em função do custo de processamento do Grafo de Gabriel, que diminui substancialmente diante da quantidade reduzida de pontos (ou protótipos) utilizada pelas versões modificadas desses índices.

Embora a redução nos tempos de processamento dos índices através de suas versões modificadas seja significativa para todos os conjuntos de dados, ela se destaca para os conjuntos

$SR\_1$  e  $SR\_2$  que apresentam uma maior dimensionalidade e um volume maior de pontos. Vale ressaltar que esses dois conjuntos são formados por dados de imagens de sensoriamento remoto, sendo, portanto, bastante similares aos tipos de dados que são objeto principal de aplicação das metodologias de classificação propostas neste trabalho.

Como era esperado, a estratégia de modificar os cálculos dos índices de validação apresenta erros de aproximação. As tabelas e gráficos apresentados no apêndice A.5 mostram as diferenças entre os valores produzidos pelas versões originais e modificadas de cada índice. No entanto, com exceção apenas de três casos (figuras A.5.3(a), A.5.10(a) e A.5.10(c)), a variação ou o comportamento dos valores das versões modificadas dos índices ocorre de maneira similar ao das versões originais para todas as partições avaliadas, não comprometendo, portanto, o processo de tomada de decisão a respeito da melhor partição de cada conjunto de dados.

Isso também pode ser observado através da tabela 5.5, que mostra, para cada conjunto de dados, o número de agrupamentos da melhor partição indicada pelas versões originais e modificadas de cada índice de validação. A tabela mostra, ainda, através dos símbolos "O" e "X", se a melhor partição indicada por cada índice é a correta ou não. O símbolo "O" denota que o índice de validação acertou na indicação da melhor partição e o símbolo "X" denota o contrário (que o índice errou). Os números e indicações, marcados em negrito na tabela, referem-se aos casos em que os comportamentos das duas versões dos índices foram diferentes. Analisando a tabela, é possível verificar que, praticamente para todos os conjuntos utilizados nos testes, as indicações sugeridas pelas versões modificadas dos índices foram as mesmas que às das versões originais.

Vale destacar ainda que mesmo naqueles três casos em que os comportamentos das duas versões dos índices foram diferentes (marcados em negrito na tabela 5.5), o número de agrupamentos sugerido pelos índices modificados foi o correto, enquanto suas versões originais sugeriram um número errado de classes. Isso ocorre em virtude da redução de ruído proporcionada pelo uso dos protótipos do SOM. Uma vez que os protótipos do SOM são "médias" locais dos dados, a avaliação dos agrupamentos por meio das versões modificadas dos índices se torna menos sensível às variações aleatórias dos conjuntos de dados. De fato, em função do uso de Grafo de Gabriel, particularmente os índices  $\mathcal{V}_{DPal}$  e  $\mathcal{V}_{DBPal}$  são muito suscetíveis à ruídos.

Índice	Conjuntos de dados										Acertos
	<i>Gauss_1</i>	<i>Gauss_2</i>	<i>Gauss_3</i>	<i>Rand_1</i>	<i>Rand_2</i>	<i>Rand_3</i>	<i>WBC</i>	<i>Wine</i>	<i>SR_1</i>	<i>SR_2</i>	
$\mathcal{V}_{D_{Pal}}$	5(O)	4(X)	<b>2(X)</b>	2(X)	2(X)	4(X)	2(O)	2(X)	2(X)	<b>3(X)</b>	2/10
$\mathcal{V}_{D_{Pal}}^{SOM}$	5(O)	4(X)	<b>5(O)</b>	2(X)	2(X)	4(X)	2(O)	2(X)	2(X)	<b>4(O)</b>	4/10
$\mathcal{V}_{D_{Bdk}}$	5(O)	5(O)	5(O)	2(X)	2(X)	4(X)	2(X)	3(O)	3(O)	3(X)	5/10
$\mathcal{V}_{D_{Bdk}}^{SOM}$	5(O)	5(O)	5(O)	2(X)	2(X)	4(X)	2(X)	3(O)	3(O)	3(X)	5/10
$\mathcal{V}_{DB_{Pal}}$	5(O)	5(O)	5(O)	2(X)	2(X)	4(X)	2(O)	3(O)	2(X)	<b>3(X)</b>	5/10
$\mathcal{V}_{DB_{Pal}}^{SOM}$	5(O)	5(O)	5(O)	2(X)	2(X)	4(X)	2(O)	3(O)	2(X)	<b>4(O)</b>	6/10
$\mathcal{V}_{DB_{Kim}}$	5(O)	5(O)	5(O)	4(O)	5(X)	6(O)	2(O)	3(O)	3(O)	5(X)	8/10
$\mathcal{V}_{DB_{Kim}}^{SOM}$	5(O)	5(O)	5(O)	4(O)	5(X)	6(O)	2(O)	3(O)	3(O)	5(X)	8/10
$\mathcal{V}_{PBM}$	5(O)	5(O)	5(O)	2(X)	2(X)	6(O)	2(O)	3(O)	2(X)	4(O)	7/10
$\mathcal{V}_{PBM}^{SOM}$	5(O)	5(O)	5(O)	2(X)	2(X)	6(O)	2(O)	3(O)	2(X)	4(O)	7/10
$\mathcal{V}_{CDbw}$	5(O)	5(O)	5(O)	4(O)	8(O)	6(O)	2(O)	3(O)	3(O)	4(O)	10/10
$\mathcal{V}_{CDbw}^{SOM}$	5(O)	5(O)	5(O)	4(O)	8(O)	6(O)	2(O)	3(O)	3(O)	4(O)	10/10
$\mathcal{V}_{Conn}$	5(O)	5(O)	5(O)	4(O)	2(X)	6(O)	2(O)	3(O)	2(X)	3(X)	7/10

Tab. 5.5: Número de agrupamentos e indicação de acerto (ou não) da melhor partição apontada por cada índice de validação para cada conjunto de dados. O símbolo "O" denota que o índice de validação acertou na indicação da melhor partição e o símbolo "X" denota que o índice errou. A última coluna da tabela mostra a proporção de acertos de cada índice considerando todos os conjuntos de dados. Os números e indicações marcados em negrito na tabela referem-se aos casos em que os comportamentos das duas versões dos índices foram diferentes.

Conforme mencionado no início desta seção, além de analisar a proposta de modificação de cálculo dos índices de validação, os testes experimentais apresentados aqui também são considerados para verificar quais dentre os sete índices de validação utilizados apresentam os melhores desempenhos na avaliação dos diferentes conjuntos de dados considerados nos experimentos. O critério utilizado para avaliar os desempenhos dos índices foi a quantidade de vezes que cada um deles indicou corretamente a melhor partição para os conjuntos de dados utilizados nos experimentos.

Analisando novamente a tabela 5.5, verifica-se que os índices  $\mathcal{V}_{D_{Pal}}$ ,  $\mathcal{V}_{D_{Bdk}}$  e  $\mathcal{V}_{DB_{Pal}}$  foram aqueles que apresentaram os piores desempenhos, acertando o número de agrupamentos de no máximo 50% dos conjuntos de dados analisados. Esses índices não avaliaram corretamente nenhum dos conjuntos de dados que possuem agrupamentos com formatos geométricos variados, especialmente os conjuntos *Rand\_1*, *Rand\_2* e *Rand\_3*. Os índices  $\mathcal{V}_{D_{Pal}}$ ,  $\mathcal{V}_{D_{Bdk}}$  e  $\mathcal{V}_{DB_{Pal}}$  também apresentaram indicações incorretas em boa parte dos conjuntos de dados reais, os

quais apresentam agrupamentos com sobreposições. Em seguida, observando ainda o número de acertos, têm-se os índices  $\mathcal{V}_{PBM}$  e  $\mathcal{V}_{Conn}$ , que em suas versões originais sugeriam corretamente o número de agrupamentos para 7 dos 10 conjuntos de dados analisados. O índice  $\mathcal{V}_{PBM}$  apresentou desempenho desfavorável para os conjuntos *Rand\_1* e *Rand\_2* que apresentam agrupamentos com formatos geométricos variados, e também errou a indicação para o conjunto *SR\_1* que apresenta um volume maior de dados com agrupamentos sobrepostos. Já o índice  $\mathcal{V}_{Conn}$  indicou erradamente o número de agrupamentos para os dois conjuntos de dados de sensoriamento remoto (*SR\_1* e *SR\_2*) e também para o conjunto *Rand\_2*. O índice  $\mathcal{V}_{DBKim}$  apresentou um desempenho semelhante ao do índice  $\mathcal{V}_{Conn}$ , sendo ligeiramente superior, uma vez que acertou o número de agrupamentos do conjunto *SR\_1*. Finalmente, dentre os sete índices de validação testados, o índice  $\mathcal{V}_{CDbw}$  foi o único que indicou corretamente o número de agrupamentos para todos os conjuntos de dados utilizados nos experimentos apresentados aqui.

Sendo assim, diante do levantamento bibliográfico sobre índices de validação realizado neste trabalho (apresentado na seção 5.3.1), dos resultados experimentais apresentados aqui e daqueles apresentados também em Gonçalves *et al.* (2005a), Gonçalves e Andrade Netto (2005) e em Gonçalves *et al.* (2006), onde a estratégia de modificação dos cálculos dos índices de validação é aplicada em outros conjuntos de dados, incluindo diferentes imagens de satélite, verifica-se que o índice  $\mathcal{V}_{CDbw}^{SOM}$  apresenta um desempenho destacável em comparação com outros índices de validação, sendo, portanto, adequado para realizar a avaliação de diferentes tipos de agrupamentos, inclusive de imagens de sensoriamento remoto.

No entanto, diante da redução considerável nos tempos de processamento dos índices de validação através de suas versões modificadas, principalmente para conjuntos de dados volumosos, verifica-se que é viável a aplicação de mais de um índice de validação (em suas formas modificadas) para avaliar diferentes partições de dados de imagens de sensoriamento remoto. Portanto, além do índice  $\mathcal{V}_{CDbw}^{SOM}$ , os índices  $\mathcal{V}_{DBKim}^{SOM}$  e  $\mathcal{V}_{PBM}^{SOM}$  são também utilizados dentro da metodologia de classificação de imagens proposta neste capítulo, uma vez que apresentam bons desempenhos de acordo com os experimentos mostrados aqui.

Vale observar ainda que, embora o índice  $\mathcal{V}_{Conn}$  tenha apresentado um desempenho similar



ao do índice  $\mathcal{V}_{PBM}^{SOM}$ , os seus tempos de processamento não são considerados adequados para a aplicação em conjuntos com grande densidade de dados. Além disso, conforme apontado em Tasdemir e Merenyi (2007), o índice  $\mathcal{V}_{Conn}$  não é adequado para análises de agrupamentos com SOMs que apresentem protótipos sem rótulos, situação que pode ocorrer dentro da metodologia de classificação proposta no próximo capítulo desse trabalho.

## **5.4 Metodologia de classificação de imagens usando SOMs segmentados por morfologia matemática e índices de validação de agrupamentos**

A metodologia de classificação de imagens de sensoriamento remoto proposta neste capítulo segue basicamente os mesmos passos daquela apresentada no capítulo anterior. No entanto, a estratégia de análise de agrupamentos usando SOMs segmentados por morfologia matemática e índices de validação (apresentada na seção 5.2) é aplicada.

A figura 5.3 ilustra os passos do método de classificação proposto. Primeiramente, o processo de amostragem da imagem é realizado por meio de janelas de pixels, as quais são utilizadas como padrões de entrada para o SOM. O treinamento não-supervisionado do SOM é efetuado e o MCI da imagem (descrito na seção 4.2.2) é obtido. Na seqüência, diferentemente do método apresentado no capítulo 4, diferentes partições para o SOM treinado são produzidas e, em seguida, avaliadas por meio de índices de validação de agrupamentos. De acordo com o estudo apresentado na seção anterior, as versões modificadas dos índices  $\mathcal{V}_{CDbw}$ ,  $\mathcal{V}_{DBK_{im}}$  e  $\mathcal{V}_{PBM}$  são empregadas nesta abordagem para avaliar a qualidade de cada uma das partições geradas, baseando-se apenas nos vetores de pesos sinápticos e nos graus de atividade dos protótipos do SOM treinado. No final, os protótipos da melhor partição do SOM, determinada de acordo com os valores dos índices, são então utilizados como referência para realizar a classificação final de todos os pixels da imagem.

Para avaliar o seu desempenho, a metodologia de classificação de imagens proposta neste capítulo é aplicada sobre a mesma imagem teste utilizada no capítulo 4 e sobre uma imagem

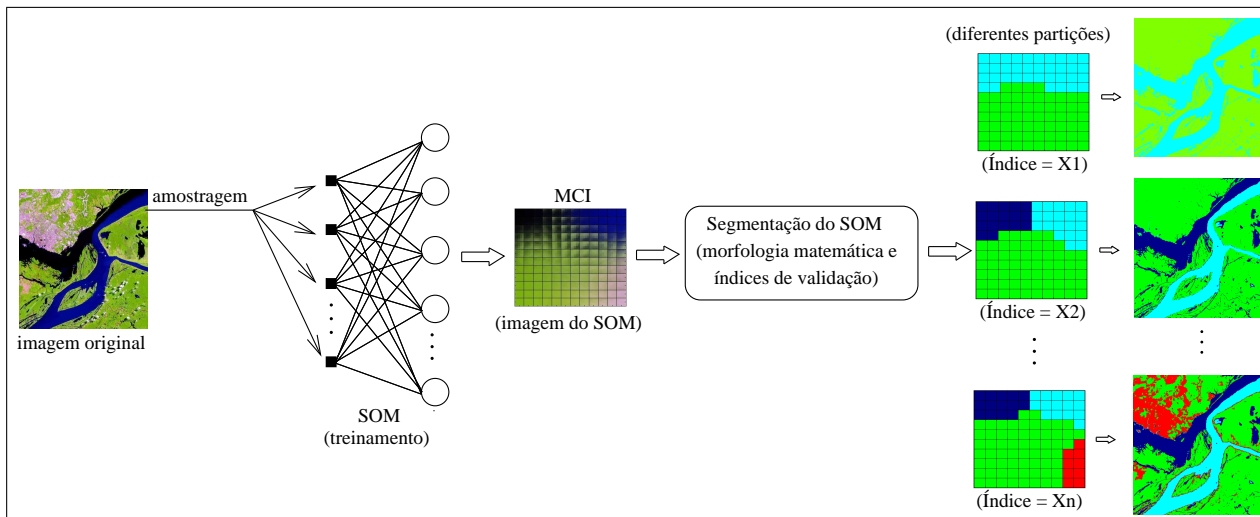


Fig. 5.3: Ilustração do método de classificação de imagens proposto usando SOMs segmentados por morfologia matemática e índices de validação.

do satélite CBERS. Os resultados das classificações são comparados com aqueles produzidos pelo algoritmo K-médias e por uma rede neural de *Perceptrons* multicamadas.

## 5.4.1 Resultados Experimentais

### 5.4.1.1 Imagem teste 1

Com o objetivo de comparar as diferenças de desempenho entre a metodologia de classificação de imagens proposta neste capítulo com aquela apresentada no capítulo 4, a primeira imagem utilizada nos testes é a mesma que foi empregada nos experimentos apresentados na seção 4.3.1. A imagem é composta pelas bandas espectrais 3, 4 e 5 do satélite LANDSAT-5 (TM), possui 496x512 pixels e apresenta quatro classes de cobertura terrestre.

Desse modo, o mesmo SOM que foi usado como mapa raiz (SOM inicial) no experimento mostrado na seção 4.3.2, também foi empregado aqui como ponto de partida para realizar a análise de agrupamentos da imagem. As figuras 4.15 e 4.16 mostram, respectivamente, a imagem do MCI e da *U-matrix* do mapa de neurônios treinado.

De acordo com a metodologia de classificação proposta neste capítulo, após o treinamento do SOM, a estratégia de particionamento do mapa de neurônios descrita na seção 5.2 deve ser aplicada. Sendo assim, o próximo passo consiste em obter o conjunto  $S_m$  de imagens de

marcadores para a  $U$ -matrix obtida. Para isso, é necessário determinar o conjunto  $S_k$  composto por todos os valores de  $k$  que correspondem à inícios de seqüências contíguas e constantes de números de regiões conectadas (com tamanhos superiores a 3).

A figura 5.4 mostra o gráfico do valor de limiar ( $k$ ) da  $U$ -matrix versus o número de regiões conectadas ( $N_{rc}$ ). O gráfico é igual àquele mostrado na figura 4.17, porém procurou-se destacar os valores de  $k$  que devem pertencer ao conjunto  $S_k$ .

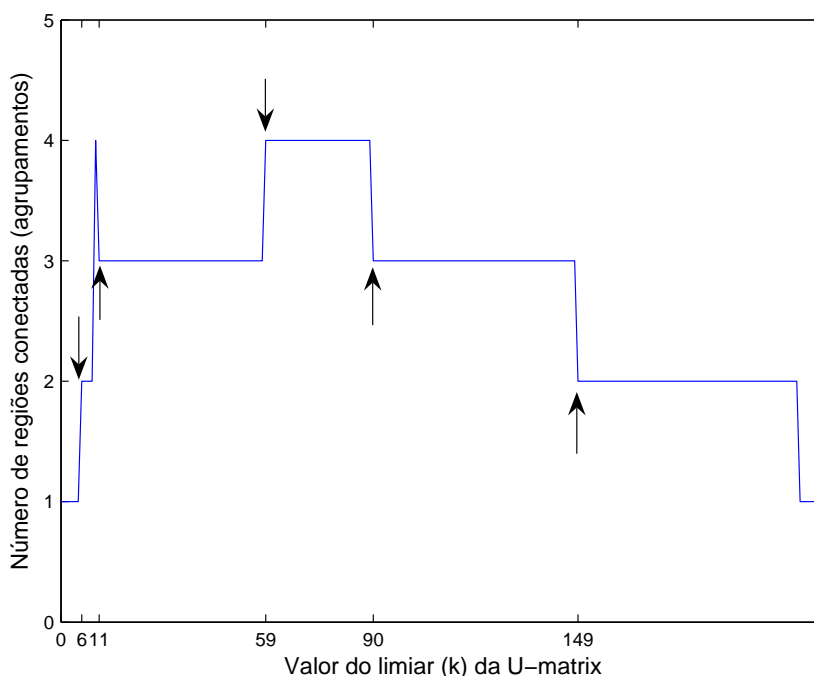


Fig. 5.4: Gráfico do valor de limiar ( $k$ ) da  $U$ -matrix versus o número de regiões conectadas destacando os valores de  $k$  que pertencem a  $S_k$ .

Analisando o gráfico, observa-se que os valores de  $k$  iguais a 6, 11, 59, 90 e 149 correspondem à inícios de seqüências contíguas e constantes de  $N_{rc}$  com tamanhos superiores a 3, ou seja,  $S_k = \{6, 11, 59, 90, 149\}$ . Com isso foi possível gerar então o conjunto  $S_m = \{U_2^6, U_2^{11}, U_2^{59}, U_2^{90}, U_2^{149}\}$  composto pelas imagens de marcadores da  $U$ -matrix, cada uma delas correspondendo a um dos cinco valores de  $k$  pertencentes a  $S_k$ .

Na seqüência, para cada uma das cinco imagens de marcadores de  $S_m$ , obteve-se as linhas de partição da  $U$ -matrix (através do algoritmo *watershed*), rotulou-se a  $U$ -matrix particionada e copiou-se os rótulos de cada elemento da  $U$ -matrix para os neurônios correspondentes no

SOM. A figura 5.5 mostra os resultados obtidos para cada uma dessas ações.

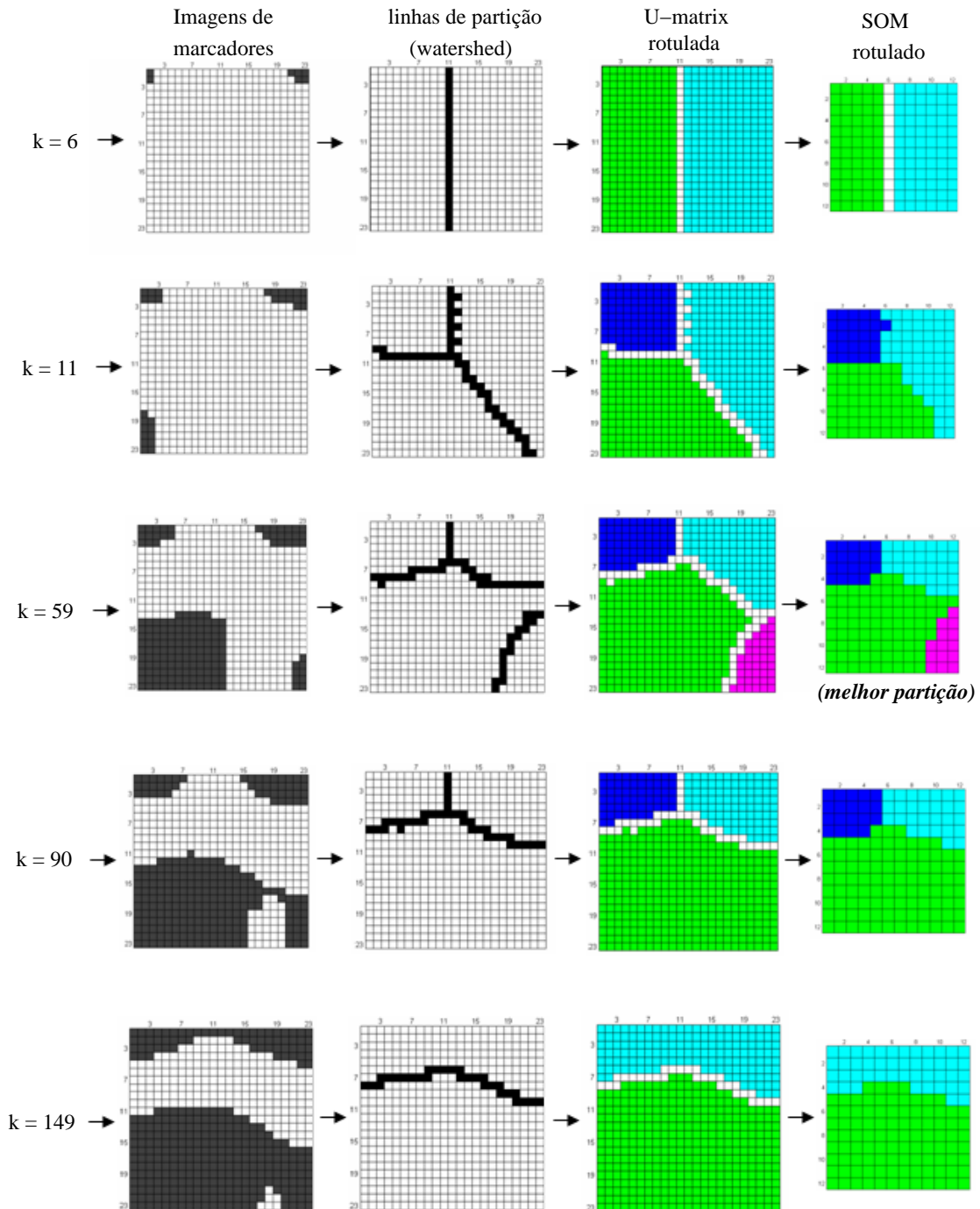


Fig. 5.5: Diferentes partições para o SOM da imagem teste 1.

Tendo produzido cinco partições diferentes para o SOM treinado, os índices de validação de agrupamentos (modificados)  $\mathcal{V}_{CDbw}^{SOM}$ ,  $\mathcal{V}_{DB_{Kim}}^{SOM}$  e  $\mathcal{V}_{PBM}^{SOM}$  foram então aplicados para avaliar a

qualidade de cada uma delas, buscando desse modo determinar qual a melhor partição. A tabela 5.6 mostra os valores dos índices para cada uma das partições obtidas (representadas pelo limiar  $k$ ), juntamente com o número de agrupamentos correspondente.

Considerando que as melhores partições são dadas pelos maiores valores dos índices  $\mathcal{V}_{CDbw}^{SOM}$  e  $\mathcal{V}_{PBM}^{SOM}$ , e pelos menores do índice  $\mathcal{V}_{DBKim}^{SOM}$ , verifica-se, portanto, que a melhor partição sugerida pelos três índices de validação é aquela referente ao limiar  $k = 59$ , a qual possui 4 agrupamentos.

Analisando o MCI do SOM treinado (figura 4.15) e a sua melhor partição (na figura 5.5), nota-se que o mapa foi segmentado adequadamente de acordo com as 4 classes de cobertura terrestre existentes na imagem, as quais são representadas pelos protótipos do SOM. A classe água 1 no canto superior esquerdo (padrão de água mais escuro), a classe água 2 no canto superior direito (mais arroxeadada), a classe área urbana no canto inferior direito (em rosa no MCI) e, por fim, a classe vegetação (em verde) com protótipos do canto inferior esquerdo ao centro do mapa.

Partição referente ao limiar $k$	Número de agrupamentos	Valores dos índices		
		$\mathcal{V}_{CDbw}^{SOM}$	$\mathcal{V}_{DBKim}^{SOM}$	$\mathcal{V}_{PBM}^{SOM}$
6	2	48.16	812.3	213.08
11	3	156.67	692.4	380.18
<b>59</b>	<b>4</b>	<b>777.49</b>	<b>332.7</b>	<b>728.79</b>
90	3	347.17	409.3	548.11
149	2	64.67	581.7	714.71

Tab. 5.6: Valores dos índices  $\mathcal{V}_{CDbw}^{SOM}$ ,  $\mathcal{V}_{DBKim}^{SOM}$  e  $\mathcal{V}_{PBM}^{SOM}$  para as 5 partições do SOM da imagem teste 1.

Em seguida, outros quatro submapas foram gerados a partir dos subconjuntos de dados de cada uma das classes encontradas, mas como nenhum deles gerou um número mínimo de duas regiões (ou agrupamentos), os mesmos foram descartados.

Executando a última etapa do método de classificação proposto, todos os pixels da imagem teste 1 foram classificados utilizando como referência os protótipos do SOM rotulado. A figura 5.6 mostra a imagem categorizada de acordo com as 4 classes de cobertura terrestre detectadas.

Ao comparar os resultados experimentais obtidos aqui com aqueles produzidos pela me-

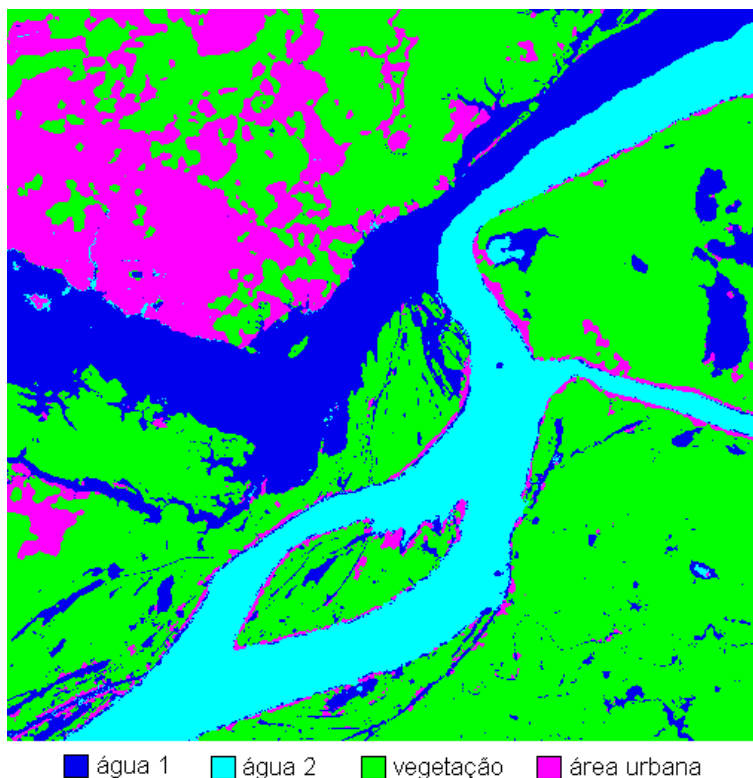


Fig. 5.6: Imagem teste 1 classificada pela metodologia de classificação proposta.

metodologia de classificação proposta no capítulo 4, verifica-se que um número menor de SOMs foi necessário para detectar os 4 agrupamentos de dados existentes na imagem teste 1. Desconsiderando os SOMs que não apresentaram nenhum agrupamento de dados, a metodologia proposta neste capítulo necessitou de apenas um único SOM para encontrar a partição correta do conjunto de dados da imagem. Já aquela do capítulo 4, necessitou de três mapas (dois a mais). Conforme mencionado anteriormente, essa diferença nas quantidades de SOMs analisados ocorre em função das estratégias de segmentação dos SOMs empregadas pelas duas metodologias. Na metodologia apresentada no capítulo 4, a partição escolhida para cada SOM é aquela que corresponde à imagem binária obtida da limiarização da *U-matrix* com o valor inicial da seqüência mais estável de número de regiões conectadas ( $N_{rc}$ ). Sendo assim, para o primeiro SOM da imagem teste 1, o valor inicial da seqüência mais estável foi  $k = 149$ , o qual corresponde a uma partição com apenas dois agrupamentos de dados. De acordo com os resultados experimentais mostrados na seção 4.3.2, cada um dos dois agrupamentos de dados inicialmente discriminados apresentavam ainda mais dois subgrupos de dados. Conseqüente-

mente, foi necessário gerar e analisar mais dois submapas para encontrar as quatro classes de cobertura terrestre existentes na imagem. Já no método de classificação proposto neste capítulo, que avalia diferentes segmentações para a *U-matrix*, e não somente a mais estável, foi possível encontrar a melhor partição dos dados logo na primeira iteração do algoritmo.

Para analisar as diferenças entre as duas metodologias em termos de tempo de processamento, a tabela 5.7 mostra os tempos consumidos por cada uma das etapas do método proposto neste capítulo na classificação da imagem teste 1. Vale observar que o tempo da etapa de treinamento do SOM (158 segundos) apresentado na tabela engloba, além do tempo de treinamento do SOM inicial, os tempos gastos no treinamento dos submapas que não apresentaram nenhum agrupamento de dados. O tempo total gasto pela metodologia de classificação proposta neste capítulo para classificar a imagem teste 1 foi de 215 segundos, aproximadamente um terço a menos do tempo total gasto pela metodologia de classificação apresentada no capítulo anterior (tabela 4.2). Como pode ser visto na subseção 4.3.3.1, esse tempo foi praticamente igual ao que foi consumido pelo algoritmo K-médias na classificação da mesma imagem.

Etapas do método proposto	Tempo consumido (em segundos)
Amostragem	9
Treinamentos dos SOMs	158
Segmentações dos SOMs	9
Avaliação das partições	3
Classificação da imagem	36
<i>Total</i>	215

Tab. 5.7: Tempo consumido pelas etapas da metodologia proposta na classificação da imagem teste 1.

Observando ainda os tempos mostrados na tabela 5.7, verifica-se que a etapa de avaliação das diferentes segmentações (ou partições) do SOM, a qual envolve a aplicação dos índices de validação, apresenta um tempo de processamento praticamente irrelevante diante do tempo total gasto pelo método. Tal fato ocorre em virtude da aplicação da estratégia proposta para modificar os cálculos dos índices de validação de agrupamentos. A tabela 5.8 mostra a diferença entre os tempos de processamento das versões originais e modificadas dos índices de validação na avaliação das partições da imagem teste desse experimento. Caso não fossem aplicadas as

versões modificadas dos três índices, e sim suas versões originais, o tempo total gasto para avaliar as cinco partições seria de aproximadamente 120 segundos. Essa quantidade de tempo seria maior do aquela que foi reduzida com a aplicação da metodologia de classificação proposta neste capítulo. Em outras palavras, se as versões originais dos índices de validação fossem utilizadas, a metodologia de classificação proposta aqui consumiria um tempo de processamento maior do que o tempo total gasto pela metodologia apresentada no capítulo 4.

Partição referente ao limiar $k$	Tempos dos índices					
	$\mathcal{V}_{CDbw}$	$\mathcal{V}_{CDbw}^{SOM}$	$\mathcal{V}_{DBK_{im}}$	$\mathcal{V}_{DBK_{im}}^{SOM}$	$\mathcal{V}_{PBM}$	$\mathcal{V}_{PBM}^{SOM}$
6	15.265	0.3926	3.0571	0.0299	3.9862	0.0399
11	16.311	0.4067	3.2357	0.0379	4.1100	0.0406
59	16.657	0.4838	3.0928	0.0367	4.0897	0.0397
90	17.768	0.3996	3.0885	0.0358	3.9887	0.0384
149	18.276	0.3982	3.0896	0.0374	4.0143	0.0382
$\Sigma$	<b>84.277</b>	<b>2.0809</b>	<b>15.5637</b>	<b>0.1777</b>	<b>20.1889</b>	<b>0.1968</b>
<i>speedup</i>	40.5		87.58		102.58	

Tab. 5.8: Tempos de processamento (em segundos) gastos pelas versões originais e modificadas dos índices de validação na avaliação das cinco partições do SOM da imagem teste 1. Na última linha da tabela, são mostrados os *speedups* de tempo das versões modificadas dos índices em relação as suas versões originais.

Os gráficos da figura 5.7 mostram ainda os valores das versões originais e modificadas dos três índices de validação para as cinco partições analisadas. Da mesma forma que nos testes experimentais mostrados na seção 5.3.3, as versões modificadas apresentaram praticamente o mesmo comportamento que suas versões originais.

Procurando avaliar também a qualidade da classificação produzida pela metodologia proposta aqui, a imagem resultante (figura 5.6) também foi comparada com o resultado da classificação supervisionada da imagem teste 1 apresentada na seção 4.3.3. Conforme relatado naquela seção, devido a ausência de verdade terrestre, a classificação da imagem teste 1 realizada de modo supervisionado através de uma rede neural MLP, serviu como referência para avaliar a qualidade das classificações obtidas. Sendo assim, a tabela 5.9 mostra os valores da matriz de confusão para a classificação produzida pela metodologia proposta neste capítulo. O valor obtido para o índice de concordância Kappa foi 0.93.



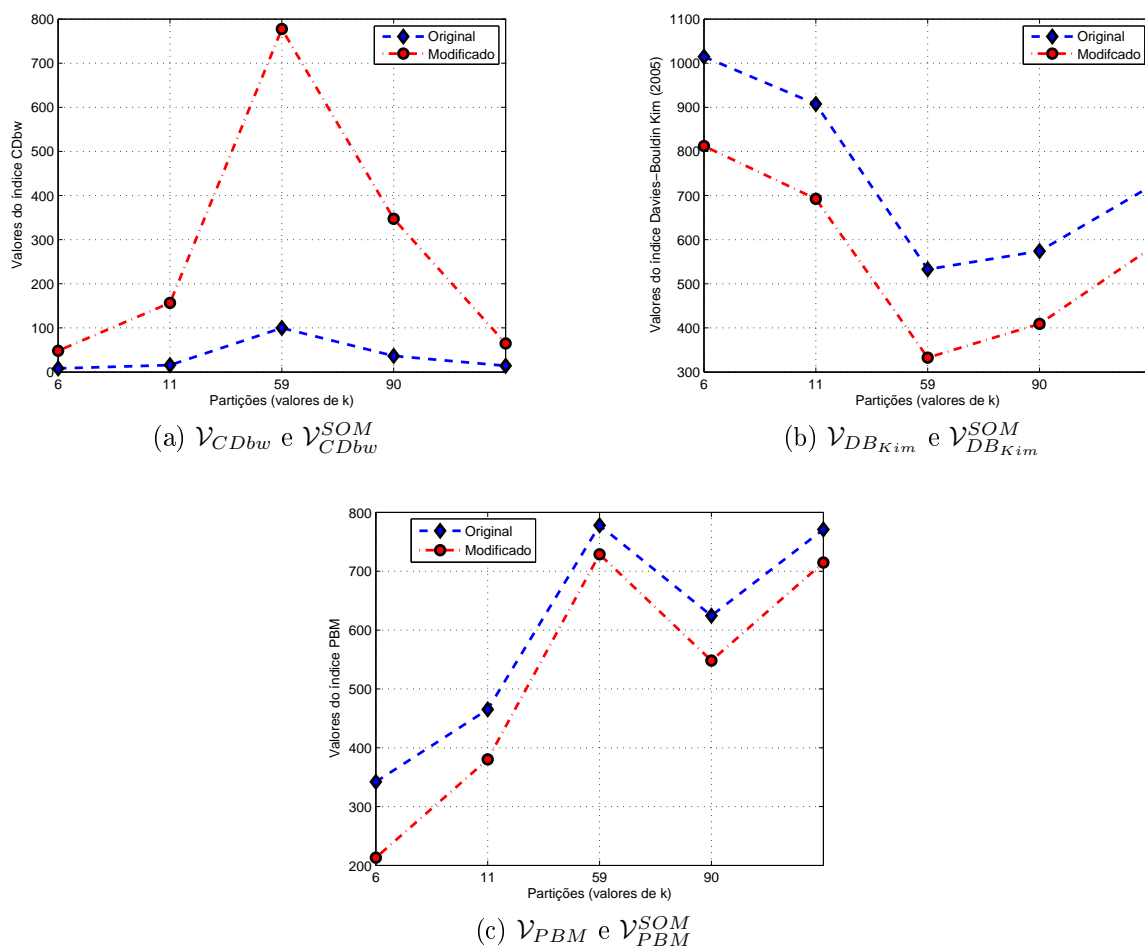


Fig. 5.7: Gráficos dos valores das versões originais e modificadas dos três índices de validação utilizados para avaliar as partições do conjunto de dados da imagem teste 1.

<i>Classes</i>	água 1	água 2	vegetação	área urbana	<i>Total</i>
água 1	<b>98.71</b>	2.20	0.86	0.67	17.56
água 2	0.91	<b>97.79</b>	0.01	0.84	15.15
vegetação	0.37	0.00	<b>93.90</b>	1.73	52.59
área urbana	0.01	0.01	5.43	<b>96.69</b>	14.69
<i>Total</i>	100.00	100.00	100.00	100.00	100.00

Tab. 5.9: Matriz de confusão das classes para a imagem classificada pela metodologia de classificação proposta. As linhas da matriz mostram os erros de comissão (porcentagem de pixels que pertencem a outra classe, mas que foi classificado como a classe de interesse) e as colunas mostram os erros de omissão (porcentagem de pixels que pertencem a classe da referência/verdade, mas que foram classificados incorretamente).

Comparando os valores das matrizes de confusão e do índice Kappa apresentados pelas duas metodologias de classificação, verifica-se que o resultado da classificação pelo método apresentado neste capítulo foi ligeiramente superior, principalmente para os padrões de água e de área urbana.

#### 5.4.1.2 Imagem teste 2

A segunda imagem teste utilizada nos experimentos é composta pelas bandas espectrais 2, 3 e 4 do satélite CBERS/CCD, possui 496x462 pixels e foi cedida pela Divisão de Geração de Imagens do Instituto Nacional de Pesquisas Espaciais (INPE). A cena, adquirida em 31/07/2000, abrange uma área ao redor da cidade de Brasília-DF e apresenta 5 classes de cobertura terrestre predominantes: água, vegetação, cultura agrícola, área irrigada (pivôs centrais) e cerrado. A figura 5.8 mostra uma composição colorida da imagem.



Fig. 5.8: *Composição colorida das bandas 3(R), 4(G) e 2(B) da imagem teste 2.*

Iniciando a aplicação da metodologia de classificação proposta neste capítulo, janelas amostrais de tamanho 5x5 foram coletadas de maneira uniforme sobre toda a região da cena, em intervalos regulares de oito pixels, totalizando 3.534 amostras obtidas sem a intervenção do

usuário.

Um SOM composto por 225 neurônios, dispostos num arranjo retangular 15x15, foi treinado com todos os dados amostrais de acordo com os parâmetros e procedimentos descritos na seção 4.2.2. O tempo de treinamento sob o Matlab em uma máquina com processador AMD Athlon XP 2600+ 1,91GHZ e 1GB de RAM foi de 109 segundos.

A figura 5.9 mostra o MCI obtido após o treinamento do SOM. É possível visualizar através dele, os 5 padrões de cobertura terrestre presentes na imagem original. No canto superior esquerdo do mapa (em cores verdes mais claras) estão os protótipos representantes do padrão de área irrigada (pivôs de irrigação), logo abaixo ao redor deles (na cor verde mais escura) estão os protótipos da classe vegetação, do canto inferior direito ao centro do mapa (na cor magenta) estão os protótipos da classe cultura agrícola, no canto inferior esquerdo do mapa (na cor magenta mais escura) estão os protótipos da classe cerrado, e logo acima deles (na cor preta) estão os protótipos que representam a classe água.

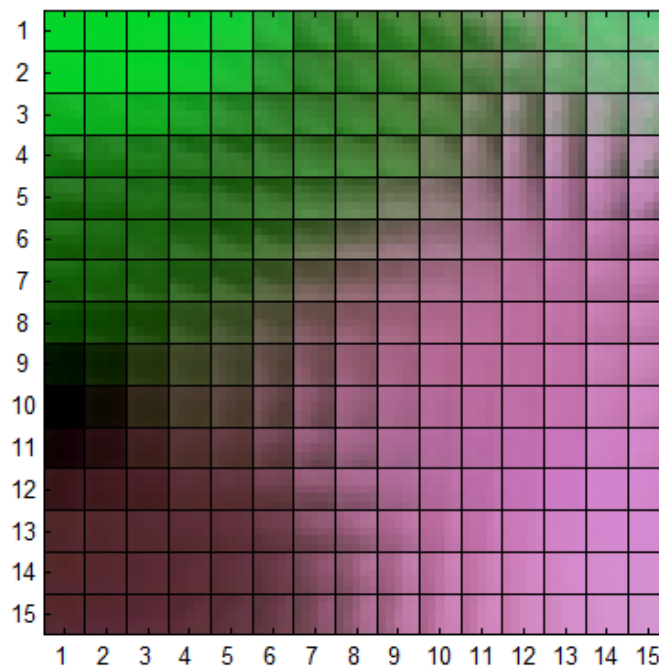


Fig. 5.9: Imagem do MCI da imagem teste 2 após o treinamento do SOM.

A figura 5.10 mostra a imagem em 3D da *U-matrix* do SOM treinado. É possível observar a existência de vales que estão associados aos agrupamentos de dados (classes) visualizadas no

MCI.

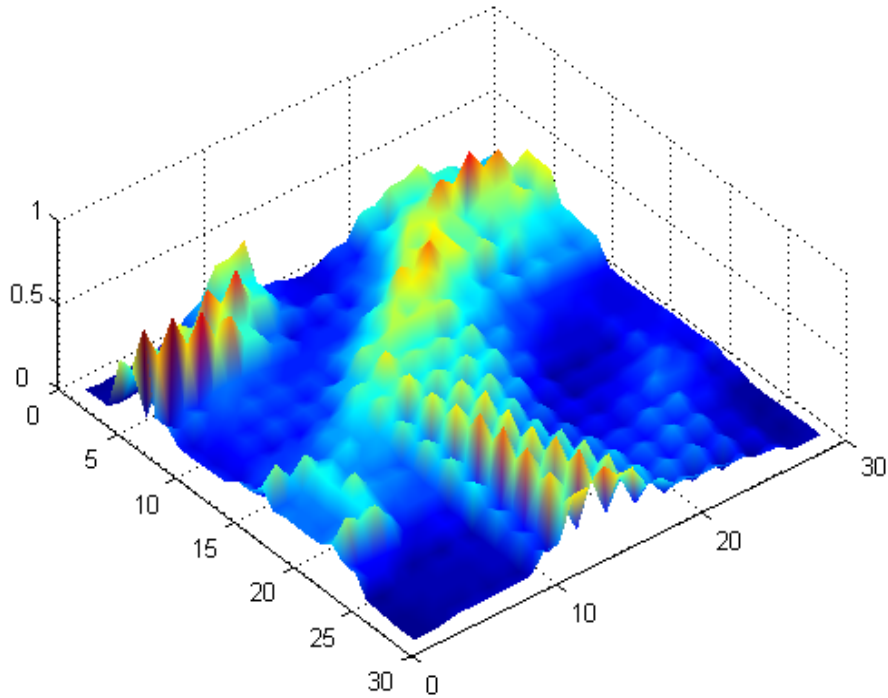


Fig. 5.10: *U-matrix* na forma de superfície do SOM da imagem teste 2.

Na seqüência, a estratégia de particionamento do mapa de neurônios descrita na seção 5.2 foi aplicada. O gráfico do número de regiões conectadas para cada valor de limiar  $k$  da *U-matrix*, apresentado na figura 5.11, foi então construído para determinar o conjunto  $S_k$ , cujos elementos são os valores de  $k$  que correspondem à inícios de seqüências contíguas e constantes de número de regiões conectadas ( $N_{rc}$ ) maiores que 3. Da mesma forma que nos experimentos anteriores, devido ao número elevado de rugosidades apresentadas pela matriz de distâncias, nesta etapa da análise foram aplicadas duas iterações do algoritmo de suavização da *U-matrix* com o objetivo de reduzir as rugosidades apresentadas pela mesma.

Analisando o gráfico da figura 5.11, observa-se que  $S_k$  é composto por 12 valores ( $S_k = \{17, 43, 57, 74, 91, 113, 128, 143, 150, 167, 183, 195\}$ ). Os 12 valores de  $k$ , que pertencem ao conjunto  $S_k$ , definem, portanto, 12 imagens de marcadores para a *U-matrix*.

Seguindo o método de classificação proposto, para cada uma das 12 imagens de marcadores, obteve-se as linhas de partição da *U-matrix* (através do algoritmo *watershed*), rotulou-se a *U-matrix* particionada e copiou-se os rótulos de cada elemento da *U-matrix* para os neurônios

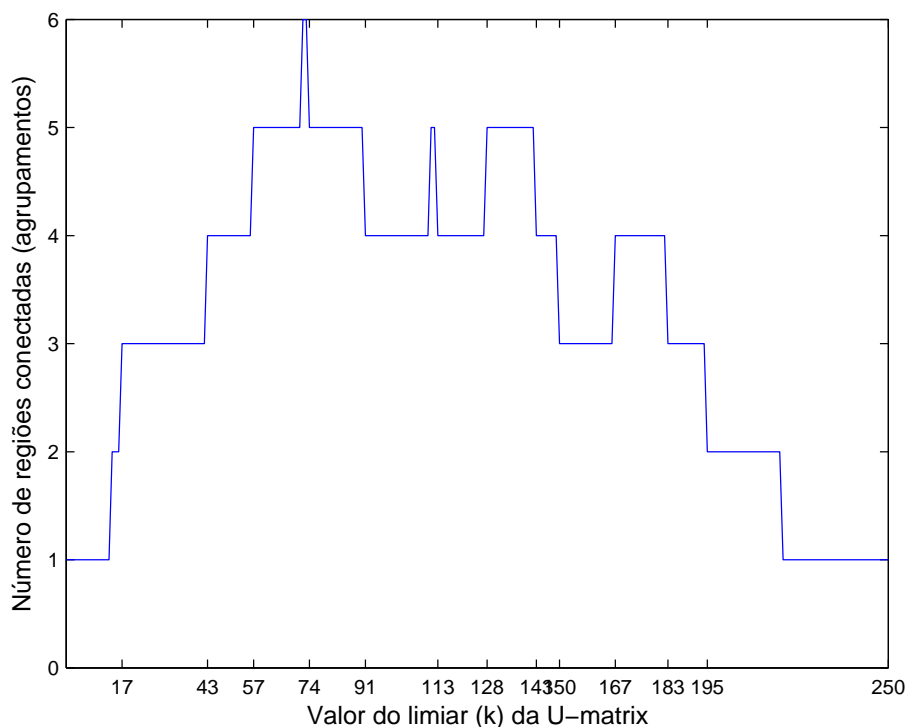


Fig. 5.11: Número de regiões conectadas versus limiar da U-matrix da imagem teste 2.

correspondentes no SOM. A figura 5.12 mostra as 12 partições do SOM da imagem teste 2 (cada uma delas correspondendo a um dos valores de  $S_k$ ).

Uma vez determinadas as 12 partições para o SOM treinado, os índices de validação de agrupamentos (modificados)  $\mathcal{V}_{CDbw}^{SOM}$ ,  $\mathcal{V}_{DBKim}^{SOM}$  e  $\mathcal{V}_{PBM}^{SOM}$  foram então aplicados para avaliar a qualidade de cada uma delas. A tabela 5.10 mostra os valores dos índices para cada uma das partições obtidas (representadas pelo limiar  $k$ ), juntamente com o número de agrupamentos correspondente.

Observando os valores da tabela 5.10, verifica-se que não houve consenso entre os três índices de validação na indicação da melhor partição. Os índices  $\mathcal{V}_{CDbw}^{SOM}$  e  $\mathcal{V}_{DBKim}^{SOM}$  indicaram como melhor partição aquela correspondente ao valor de  $k = 128$ . Já o índice  $\mathcal{V}_{PBM}^{SOM}$  apontou a partição referente ao valor de  $k = 150$  como a melhor.

Considerando que dos três índices utilizados dois deles sugeriram a mesma partição ( $k = 128$ ), e levando em conta que o índice  $\mathcal{V}_{CDbw}^{SOM}$  foi aquele que apresentou o melhor desempenho nos testes experimentais apresentados na seção 5.3.3, a partição indicada por ele e pelo índice

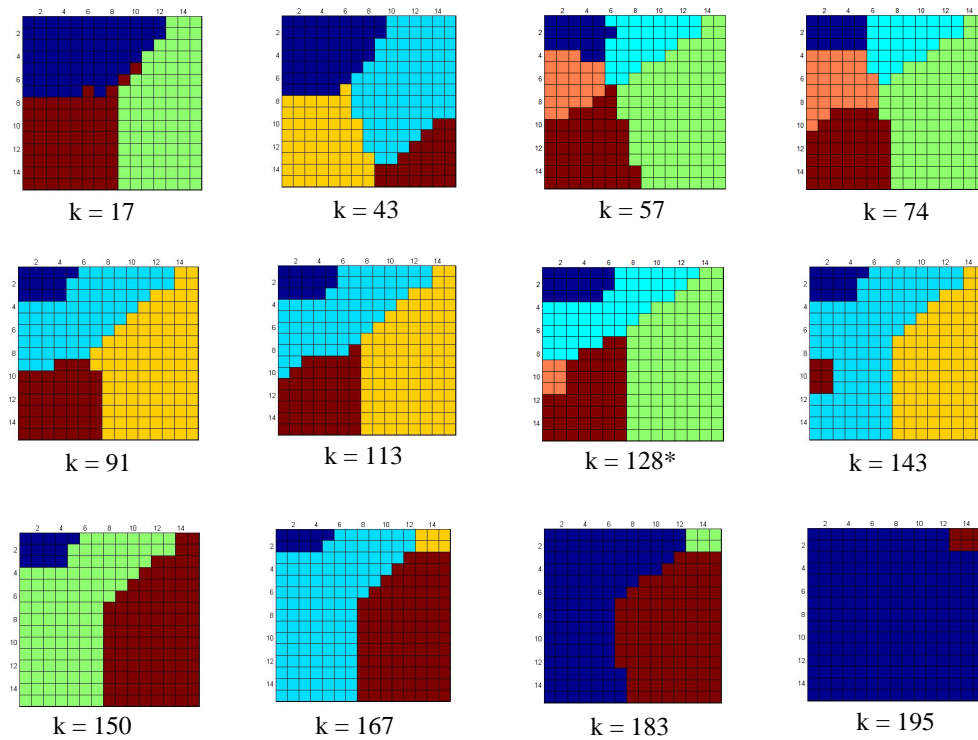


Fig. 5.12: Diferentes partições do SOM da imagem teste 2.

Partição referente ao limiar $k$	Número de agrupamentos	Valores dos índices		
		$\mathcal{V}_{CDBw}^{SOM}$	$\mathcal{V}_{DBKim}^{SOM}$	$\mathcal{V}_{PBM}^{SOM}$
17	3	142.29	414.83	648.62
43	4	892.53	404.22	539.39
57	5	1471.4	328.39	446.53
74	5	2094.6	306.29	461.08
91	4	2222.7	297.18	599.02
113	4	2072.2	286.90	599.40
128	5	<b>7146.3</b>	<b>236.23</b>	586.75
143	4	2849.7	299.13	744.84
150	3	1064.5	360.35	<b>799.70</b>
167	4	3350.0	364.49	538.20
183	3	90.793	406.09	421.93
195	2	24.598	647.78	216.86

Tab. 5.10: Valores dos índices  $\mathcal{V}_{CDBw}^{SOM}$ ,  $\mathcal{V}_{DBKim}^{SOM}$  e  $\mathcal{V}_{PBM}^{SOM}$  para as 12 partições do SOM da imagem teste 2.

$\mathcal{V}_{DBKim}^{SOM}$  foi então considerada aqui como a mais adequada para os dados da imagem teste 2.

Analisando o MCI do SOM treinado (figura 5.9) e a sua melhor partição (figura 5.12,

$k = 128$ ), nota-se que o mapa foi segmentado adequadamente de acordo com as cinco classes de cobertura terrestre existentes na imagem, as quais são representadas pelos protótipos do SOM.

Em seguida, outros cinco submapas foram gerados a partir dos subconjuntos de dados de cada uma das classes encontradas na partição  $k = 128$ , mas como nenhum deles apresentou mais de uma região (ou agrupamento), os mesmos foram descartados e o processo de análise de agrupamentos foi então finalizado.

No último passo do método de classificação proposto, os pixels da imagem teste 2 foram classificados utilizando como referência os protótipos do SOM rotulado. A figura 5.13 mostra a imagem categorizada de acordo com as 5 classes de cobertura terrestre detectadas.

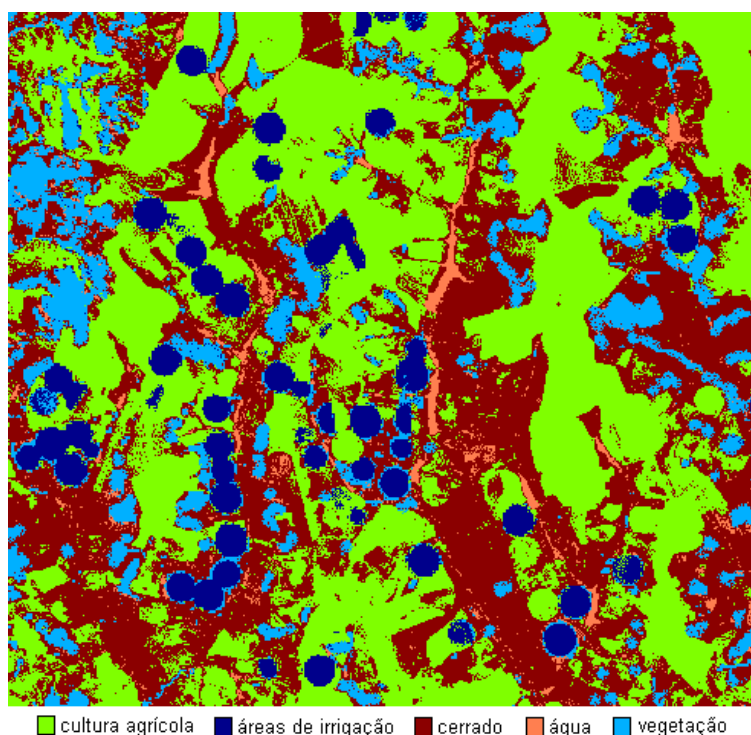


Fig. 5.13: *Imagem teste 2 classificada pela metodologia de classificação proposta.*

Analisando o processo de classificação da imagem teste 2 pela metodologia de classificação proposta nesse capítulo, nota-se que a avaliação de diferentes segmentações do SOM (inicial) possibilitou a detecção dos cinco agrupamentos de dados da imagem com apenas uma iteração do algoritmo. Ou seja, da mesma forma que no experimento da imagem teste 1, foi necessário apenas um único SOM para encontrar a melhor partição para os dados da imagem teste 2.

Caso a metodologia de classificação apresentada no capítulo 4 fosse aplicada, a partição referente ao limiar  $k = 17$  seria escolhida para segmentar o SOM inicial, uma vez que corresponde à seqüência mais estável de número de regiões conectadas em função do valor do limiar  $k$  da *U-matrix*. E como pode ser visto através da figura 5.12, a partição do SOM referente ao limiar  $k = 17$  apresenta apenas três classes, sendo que duas delas incluem ainda mais duas subclasses. Diante disso, é necessário que pelo menos dois outros SOMs sejam gerados e treinados para discriminar corretamente as cinco classes existentes na imagem teste 2, o que exige um tempo de processamento maior do que o do método proposto neste capítulo. O apêndice A.6 (localizado na mídia anexada a essa tese) mostra a árvore de mapas que é gerada a partir da aplicação da metodologia de classificação proposta no capítulo 4 sobre a imagem utilizada nesse experimento.

A tabela 5.11 mostra o tempo despendido por cada uma das etapas da metodologia proposta neste capítulo na classificação da imagem teste 2.

Etapas do método proposto	Tempo consumido (em segundos)
Amostragem	12
Treinamentos dos SOMs	207
Segmentações dos SOMs	13
Avaliação das partições	7
Classificação da imagem	39
<i>Total</i>	278

Tab. 5.11: Tempo consumido pelas etapas da metodologia proposta na classificação da imagem teste 2.

Buscando destacar a vantagem de se utilizar a estratégia proposta para modificar os índices de validação de agrupamentos, a tabela 5.12 mostra os tempos consumidos pelas versões modificadas dos três índices utilizados na avaliação das 12 partições do SOM (mostradas na figura 5.12). Os tempos das versões originais dos índices também são exibidos com o objetivo de evidenciar a inviabilidade de aplicação das mesmas dentro da metodologia de classificação proposta. Caso as versões originais fossem aplicadas para avaliar todas as 12 partições do SOM, o tempo gasto apenas nesta etapa (de avaliação) seria de aproximadamente 598 segundos, mais que o dobro do tempo que foi gasto para executar todas as etapas do método de classificação



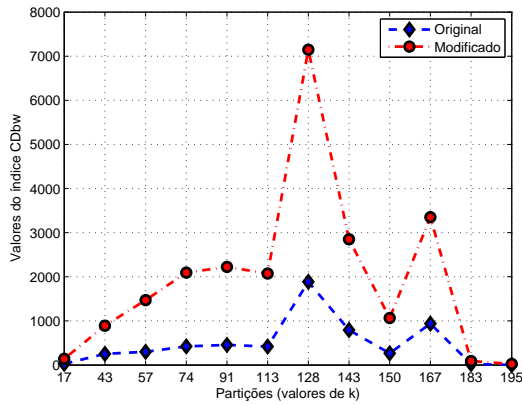
usando as versões modificadas dos índices.

Partição referente ao limiar $k$	Tempos dos índices					
	$\mathcal{V}_{CDbw}$	$\mathcal{V}_{CDbw}^{SOM}$	$\mathcal{V}_{DBKim}$	$\mathcal{V}_{DBKim}^{SOM}$	$\mathcal{V}_{PBM}$	$\mathcal{V}_{PBM}^{SOM}$
17	37.413	0.4842	5.2733	0.0404	6.1627	0.0401
43	38.839	0.4809	5.1456	0.0389	5.1376	0.0414
57	33.050	0.4822	5.142	0.0390	5.1380	0.0406
74	33.161	0.4840	5.1409	0.0489	5.1375	0.0415
91	34.696	0.4957	5.1382	0.0484	6.1318	0.0501
113	33.833	0.4930	5.1376	0.0478	6.1585	0.0578
128	33.226	0.5036	5.1458	0.0331	6.1394	0.0515
143	41.368	0.5024	5.1398	0.0480	6.0363	0.0498
150	41.933	0.4997	5.1399	0.0479	5.9512	0.0408
167	42.200	0.5081	5.1381	0.0497	5.8876	0.0521
183	48.970	0.5111	5.1393	0.0377	5.4359	0.0574
195	48.218	0.5405	5.1552	0.0475	6.1383	0.0504
$\Sigma$	<b>466.9</b>	<b>5.9854</b>	<b>61.8357</b>	<b>0.5273</b>	<b>69.4548</b>	<b>0.5735</b>
<i>speedup</i>	78		117.26		121.10	

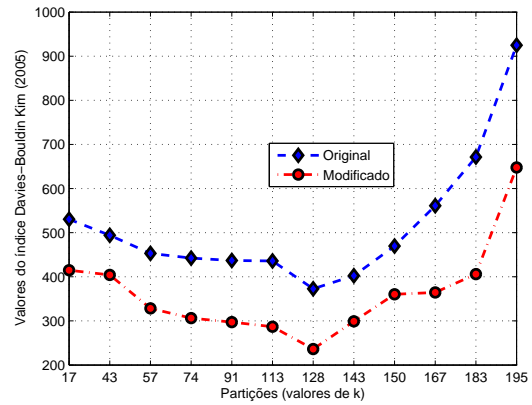
Tab. 5.12: Tempos de processamento (em segundos) gastos pelas versões originais e modificadas dos índices de validação na avaliação das 12 partições do SOM da imagem teste 2.

Os gráficos da figura 5.14 mostram os valores das versões originais e modificadas dos três índices de validação para as 12 partições analisadas. Através deles é possível verificar que as indicações das melhores partições dadas pelas versões modificadas dos índices coincidem com as indicações de suas versões originais, já que os valores de ambas apresentaram variações semelhantes em todas as partições.

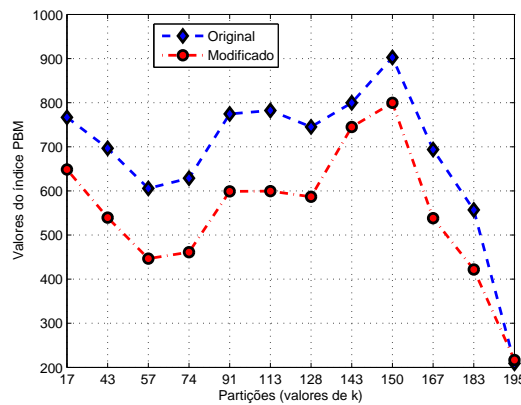
Do mesmo modo que foi feito para a imagem teste 1, devido a ausência de verdade terrestre, a imagem teste 2 foi classificada de maneira supervisionada por uma rede neural MLP. A imagem resultante serviu de referência para avaliar a qualidade da classificação produzida pela metodologia de classificação proposta. A tabela 5.13 mostra a matriz de confusão calculada considerando as duas classificações. Exceto para a classe vegetação, o percentual de acerto da imagem classificada pelo método proposto foi bastante elevado para cada classe de cobertura terrestre. Se observarmos novamente o MCI da imagem teste 2 (figura 5.9), verifica-se que a classe vegetação (na cor verde mais escuro) possui fronteira com os protótipos de todas as outras classes. Isso justifica o maior grau de confusão na classificação dos pixels dessa classe



(a)  $\mathcal{V}_{CDbw}$  e  $\mathcal{V}_{CDbw}^{SOM}$



(b)  $\mathcal{V}_{DBKim}$  e  $\mathcal{V}_{DBKim}^{SOM}$



(c)  $\mathcal{V}_{PBM}$  e  $\mathcal{V}_{PBM}^{SOM}$

Fig. 5.14: Gráficos dos valores das versões originais e modificadas dos três índices de validação utilizados para avaliar as partições do conjunto de dados da imagem teste 2.

de cobertura terrestre com os pixels das demais classes.

Em relação ao índice de concordância Kappa, o valor obtido foi igual a 0.81, que de acordo com a tabela 2.5, qualifica a classificação produzida pelo método proposto neste capítulo como excelente.

Procurando ainda comparar o desempenho da metodologia de classificação proposta com métodos convencionais de classificação não-supervisionada, a imagem teste 2 também foi classificada utilizando o algoritmo K-médias. A figura 5.15 mostra o resultado da classificação da imagem pelo algoritmo considerando o número de classes (k) igual a cinco. O algoritmo K-médias foi executado utilizando cinco inicializações diferentes para os centróides de cada classe e um número máximo de iterações igual a 150. A classificação considerada para a análise foi

<i>Classes</i>	cultura agrícola	área irrigada	água	cerrado	vegetação	<i>Total</i>
cultura agrícola	<b>96.87</b>	3.51	0.00	0.10	7.12	42.89
área irrigada	0.02	<b>96.49</b>	0.00	0.00	15.86	6.77
água	0.00	0.00	<b>100.00</b>	2.73	2.57	2.27
cerrado	3.11	0.00	0.00	<b>86.48</b>	21.16	37.39
vegetação	0.00	0.00	0.00	10.70	<b>53.29</b>	10.68
<i>Total</i>	100.00	100.00	100.00	100.00	100.00	100.00

Tab. 5.13: Matriz de confusão das classes (em porcentagem) entre a imagem classificada pelo método proposto e a imagem referência (resultante da classificação supervisionada). As linhas da matriz mostram os erros de comissão e as colunas mostram os erros de omissão.

aquela que apresentou o valor mais baixo para a soma dos quadrados das distâncias dada pela equação 2.4.

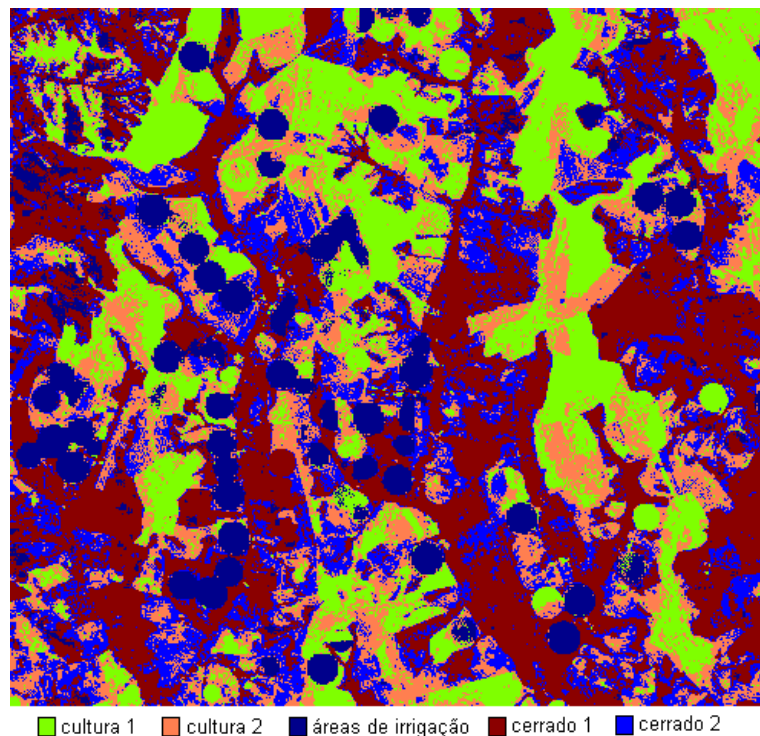


Fig. 5.15: Imagem teste 2 classificada pelo algoritmo K-médias considerando  $k = 5$ .

Diferentemente da imagem classificada pela metodologia proposta, na imagem resultante do algoritmo K-médias os pixels da classe água não foram reconhecidos, e tanto as áreas de cultura agrícola como as de cerrado foram divididas em duas subclasses. Vale destacar ainda que o algoritmo não discriminou corretamente os padrões mais importantes da cena, mesmo

inicializando os centróides dentro das cinco classes desejadas. As cinco classes de cobertura terrestre da imagem teste 2 são corretamente discriminadas pelo algoritmo K-médias somente quando um número elevado de classes é considerado. No entanto, para que o usuário tenha conhecimento disso, é necessário que ele interaja no processo e execute o algoritmo diversas vezes variando a quantidade de agrupamentos.

Com relação ao tempo de processamento, o algoritmo K-médias consumiu 268 segundos para classificar a imagem teste 2, quantidade de tempo ligeiramente inferior ao da metodologia de classificação proposta. Entretanto, conforme já discutido no capítulo anterior, os dois métodos apresentam princípios e características bastante diferenciadas, que devem ser levadas em conta quando suas complexidades computacionais são comparadas.

## 5.5 Geração de bordas mais definidas na *U-matrix*

Ao realizar uma breve interpretação do método de segmentação do SOM utilizado nas metodologias de classificação propostas até aqui, verifica-se que o seu funcionamento consiste basicamente em agrupar neurônios que estejam próximos no espaço de atributos (espaço de pesos sinápticos) e que ao mesmo tempo sejam adjacentes, ou vizinhos, na grade do SOM. Isso ocorre devido à utilização da *U-matrix*, que de acordo com a análise apresentada em Costa (1999), pode ser vista como um dendrograma de um método hierárquico de Ligações Simples (LS), onde apenas as distâncias entre os neurônios vizinhos na grade do SOM são calculadas na matriz de dissimilaridades (maiores detalhes sobre o método LS podem ser obtidos em Johnson (1967) e em Xu e Wunsch II (2005)).

Conforme descrito ainda em Costa (1999), o método de segmentação do SOM tem a mesma flexibilidade de detectar agrupamentos alongados ou de formas complexas que o LS possui, e que não há, por exemplo, em métodos que se baseiam no uso de centróides (como o K-médias). No entanto, ao mesmo tempo que o método LS apresenta essa vantagem, caso existam pontos entre duas regiões com grande densidade de pontos que caracterizem dois agrupamentos distintos, o método pode ligar os dois agrupamentos pela trajetória no espaço destes pontos de ligação.

Esse problema apresentado pelo método LS, conhecido como "efeito cadeia" ou problema

"amigos de amigos", também ocorre de maneira similar no SOM. Por exemplo, na figura 3.7 (no capítulo 3) é possível notar a existência de neurônios de ligação entre dois agrupamentos distintos de dados. É possível verificar também, através das figuras 3.8 e 3.9, que esses neurônios de ligação vencem para poucos ou nenhum padrão, pois se encontram em regiões de baixa densidade de pontos. Esse efeito "cadeia" ou de encadeamento de neurônios se reflete sobre a *U-matrix*, degradando as bordas que separam regiões (vales) que representam agrupamentos diferentes de neurônios. Essa degradação nas bordas pode dificultar o processo de segmentação da *U-matrix*.

Para contornar tal problema, Costa (1999) propôs uma maneira de acentuar as bordas da *U-matrix* procurando eliminar o efeito destes neurônios de ligação que apresentam baixa atividade. A idéia consiste em definir um grau de ativação mínimo que os neurônios devem possuir para serem considerados na análise do SOM. Os neurônios que não apresentarem esse nível mínimo de atividade são "empurrados" na direção dos neurônios ativos mais próximos. Para isso, basta copiar o vetor do neurônio ativo mais próximo para o neurônio com baixa atividade. Tal procedimento foi utilizado no exemplo de aplicação apresentado na seção 3.7.1.1 (o efeito da utilização do procedimento é mostrado na figura 3.12).

Porém, na aplicação das metodologias de classificação propostas neste trabalho, verifica-se que o efeito de encadeamento de neurônios no SOM apresenta algumas particularidades para os padrões de imagens de sensoriamento remoto. Após os treinamentos dos SOMs, nota-se que os protótipos ou neurônios de ligação, que são responsáveis pela degradação das bordas da *U-matrix*, são representantes (ou vencedores) de amostras de dados que incorporam mais de um padrão de cobertura terrestre, apresentando um alto grau de heterogeneidade espectral. Essas amostras de dados, denominadas aqui de *heterogêneas*, correspondem a regiões de transição entre classes de cobertura terrestre presentes na imagem, e são capturadas em decorrência do processo de amostragem da imagem, o qual, conforme descrito anteriormente, é realizado por meio de janelas de pixels.

Para um melhor entendimento do efeito dessas amostras heterogêneas na degradação das bordas da *U-matrix*, considere o segmento de uma imagem do satélite IKONOS mostrado na figura 5.16. A imagem apresenta três classes de cobertura terrestre homogêneas e espacialmente

bem separadas. São elas: solo arenoso (em cor mais clara, no alto da imagem), água (na cor azul) e mata (em cor mais escura, na parte inferior da imagem).

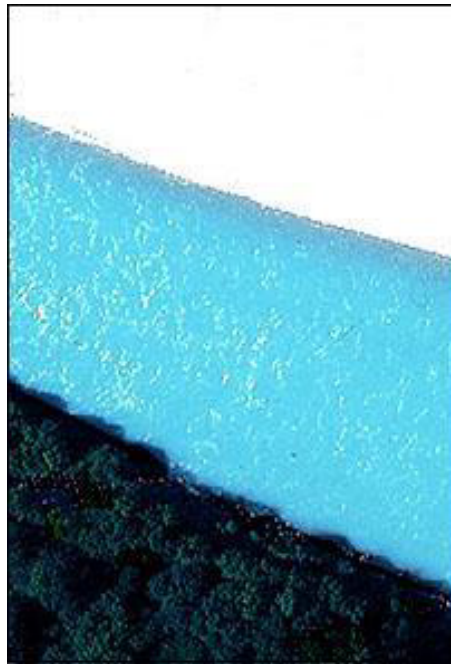


Fig. 5.16: *Segmento de imagem do satélite IKONOS.*

Seguindo o processo de amostragem definido nos métodos de classificação propostos, janelas amostrais de tamanho 10x10 foram coletadas de maneira uniforme sobre toda a região da imagem. A figura 5.17 ilustra as janelas amostrais coletadas, destacando (em vermelho) aquelas que são consideradas heterogêneas, ou seja, que apresentam mais de uma classe de cobertura terrestre. É possível observar que as amostras heterogêneas se situam exatamente nas regiões de fronteira (ou de transição) das três classes existentes na imagem. A figura 5.18 mostra a projeção de todas as amostras que foram coletadas usando apenas as duas componentes principais do seu espaço de atributos. Os pontos em vermelho correspondem às amostras heterogêneas e os pontos em azul correspondem às amostras homogêneas. Nota-se, claramente, que as amostras heterogêneas funcionam como pontos de ligação entre os três agrupamentos de dados (em azul), os quais correspondem às três classes de cobertura terrestre da imagem.

A figura 5.19 mostra o MCI de um SOM composto por 49 neurônios (arranjados em uma grade retangular 7x7), treinado com todas as amostras que foram coletadas da imagem da figura 5.16. Como era esperado, há neurônios ou protótipos no mapa (marcados em ver-

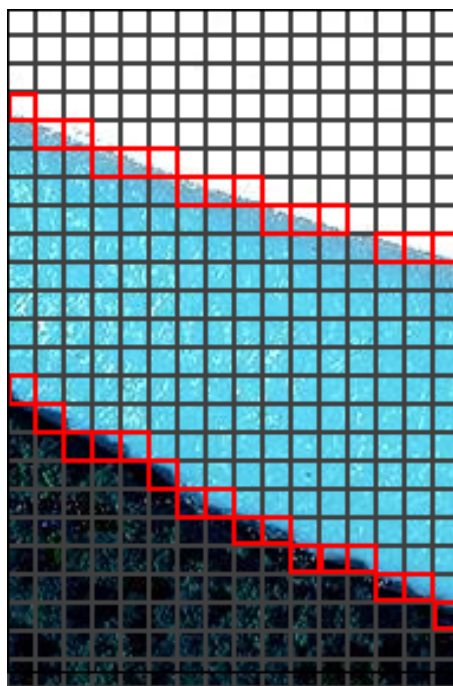


Fig. 5.17: *Janelas amostrais coletadas do segmento de imagem do satélite IKONOS. As amostras heterogêneas estão marcadas em vermelho.*

melho) que apresentam algum grau de heterogeneidade espectral. Esses protótipos obviamente surgem em função das janelas amostrais heterogêneas que são coletadas através do processo de amostragem da imagem.

A figura 5.20 mostra a projeção do conjunto de protótipos do SOM usando as duas componentes principais do espaço de atributos (ou espaço de pesos), onde os pontos em vermelho correspondem aos protótipos do SOM que apresentam algum nível de heterogeneidade espectral. Verifica-se, portanto, que o efeito de encadeamento também acontece no SOM, visto que os protótipos heterogêneos funcionam como pontos de ligação entre os neurônios (homogêneos) que representam as três classes de cobertura terrestre da imagem. Na *U-matrix* do SOM treinado (mostrada na figura 5.21), o efeito desses protótipos heterogêneos pode ser percebido através das degradações das bordas e das rugosidades existentes.

Procurando salientar o efeito das amostras heterogêneas no mapeamento do SOM e, conseqüentemente, na superfície topográfica da *U-matrix*, as figuras 5.22 e 5.23 mostram, respectivamente, o MCI e a *U-matrix* do SOM treinado sem considerar as amostras heterogêneas da imagem da figura 5.16. Comparando com o MCI e a *U-matrix* do SOM treinado com as

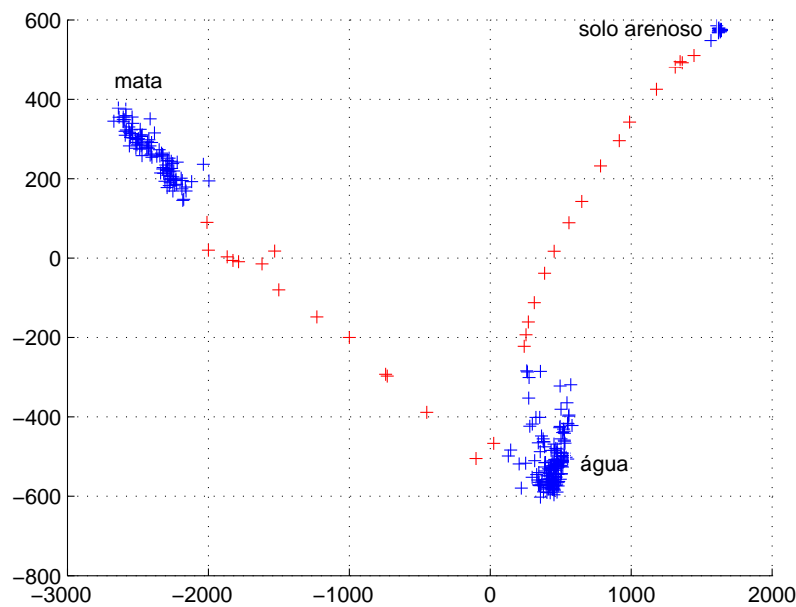


Fig. 5.18: *Projeção do conjunto de amostras usando as duas componentes principais. Os pontos em vermelho correspondem às amostras heterogêneas e os pontos em azul correspondem às amostras homogêneas.*

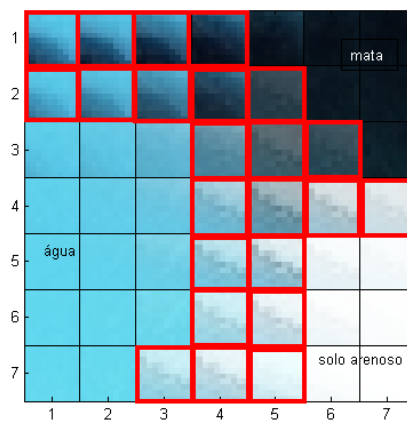


Fig. 5.19: *MCI do SOM treinado com as janelas amostrais apresentadas na figura 5.17. Os protótipos heterogêneos estão marcados em vermelho.*

amostras heterogêneas (figuras 5.19 e 5.21), verifica-se que há uma diferença significativa entre elas. A *U-matrix* do SOM treinado sem as amostras heterogêneas apresenta bordas de separação mais definidas e um número menor de rugosidades. Através dos MCIs também é possível observar as diferenças. O MCI do SOM treinado sem amostras compostas por mais de um



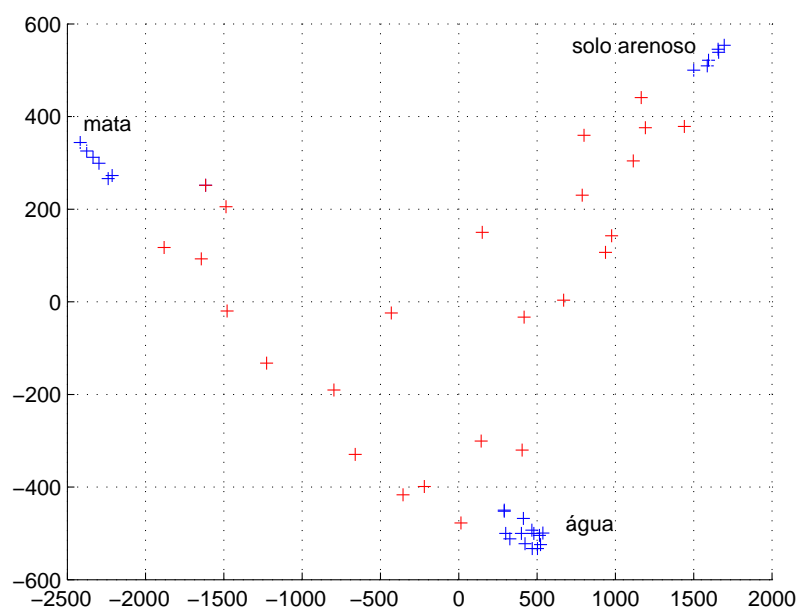


Fig. 5.20: *Projecção do conjunto de protótipos do SOM usando as duas componentes principais. Os pontos em vermelho são os protótipos heterogêneos.*

padrão de cobertura terrestre praticamente não apresenta protótipos heterogêneos. Nota-se a existência de alguns protótipos com características intermediárias entre as classes, mas que são decorrentes do processo de aprendizagem do SOM, que ajusta não somente os pesos do neurônio vencedor, mas também dos seus neurônios vizinhos.

Embora a exclusão de amostras heterogêneas do processo de treinamento do SOM nesse experimento, tenha contribuído para salientar as bordas da *U-matrix*, cabe relatar que a estratégia proposta por Costa (1999), de "empurrar" os neurônios de baixa atividade na direção dos neurônios ativos mais próximos, também serve, nesse caso, para amenizar o problema de degradação das bordas da *U-matrix*. Isso é possível, pois os protótipos do SOM da imagem da figura 5.16 que apresentam algum grau de heterogeneidade espectral estão, coincidentemente, localizados em regiões de baixa densidade de dados no espaço amostral. No entanto, é importante observar que o mesmo pode não acontecer para outras imagens de sensoriamento remoto, especialmente para aquelas que apresentam um número elevado de regiões de transição (ou de fronteira) e classes com baixa probabilidade de ocorrência.

Um exemplo é a imagem teste 2, apresentada na figura 5.8, que possui uma quantidade

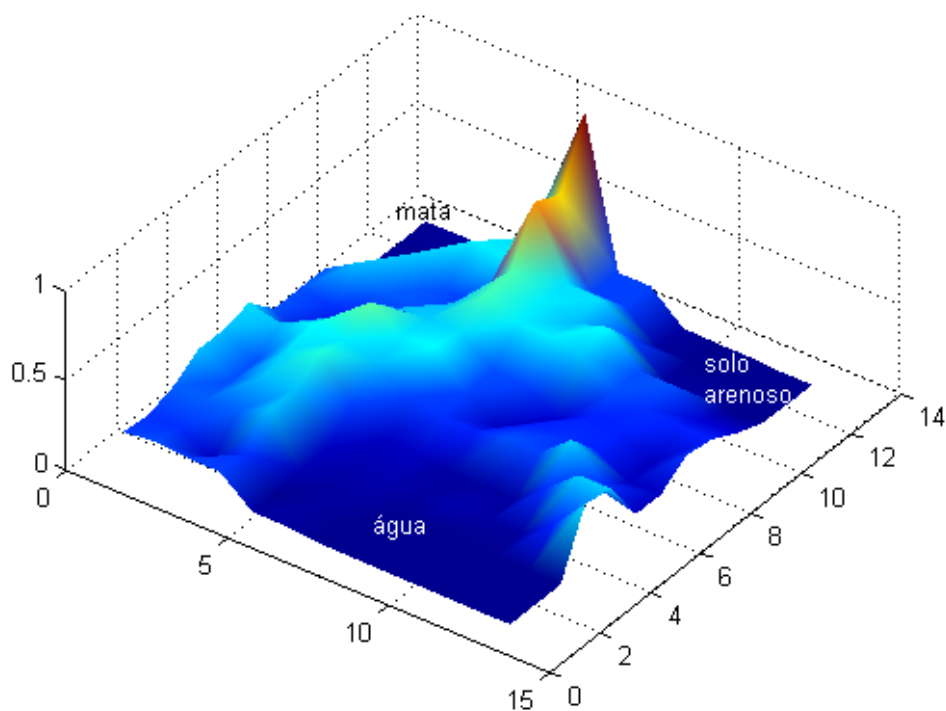


Fig. 5.21: *U-matrix* na forma de superfície do SOM treinado com todas as janelas mostrais coletadas da imagem da figura 5.16.

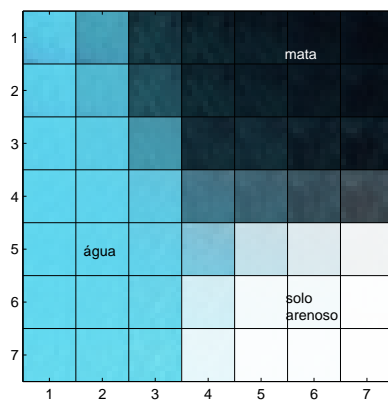


Fig. 5.22: *MCI* do SOM treinado sem as amostras heterogêneas.

muita grande de áreas de transição entre as classes e apresenta uma baixa densidade de regiões contendo o padrão água. Conseqüentemente, considerando que as janelas amostrais são coletadas de maneira uniforme sobre todo o espaço da cena, tem-se a partir daquela imagem, um número de amostras heterogêneas muito maior do que a quantidade de amostras da classe

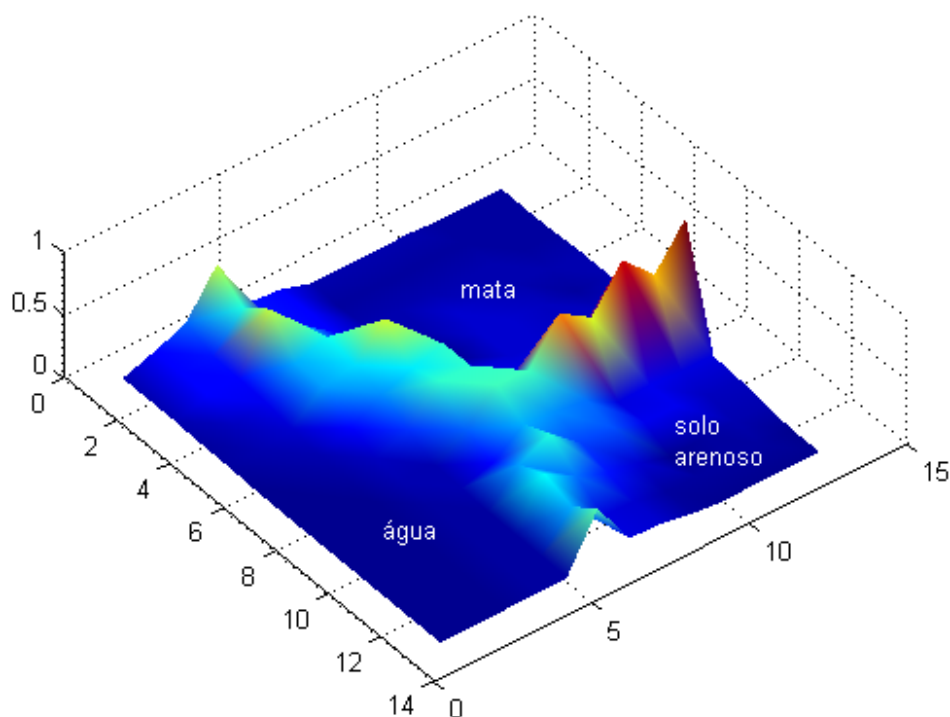


Fig. 5.23: *U-matrix* na forma de superfície do SOM treinado sem as amostras heterogêneas.

água. Logo, no MCI obtido a partir do SOM treinado com essas amostras (figura 5.9), há uma quantidade muito maior de protótipos representantes de amostras heterogêneas do que a de protótipos representantes do padrão água. Além disso, grande parte desses protótipos heterogêneos apresentam um grau de atividade (frequência de vencimentos) maior do que o dos protótipos da classe água. Devido a isso, ao aplicar a estratégia para salientar as bordas da *U-matrix*, proposta por Costa (1999), corre-se o risco de deslocar os protótipos correspondentes ao padrão água (que apresentam baixa atividade) na direção de protótipos representantes de outras classes. Por conseguinte, a classe água poderia não ser detectada ou reconhecida no processo de segmentação do SOM. O apêndice A.7, localizado na mídia que acompanha essa tese, mostra o histograma do SOM da figura 5.9, onde é possível verificar mais detalhadamente a possibilidade de exclusão dos protótipos da classe água, caso a estratégia de deslocar os protótipos com baixa atividade fosse aplicada.

Diante desse problema com a imagem teste 2, observa-se que os graus de atividade dos neurônios não podem servir exclusivamente de base para resolver o problema de degradação

das bordas da *U-matrix* quando se aplica o SOM para analisar imagens de sensoriamento remoto, especialmente em situações onde há interesse em detectar padrões de cobertura terrestre com baixa probabilidade de ocorrência. Conclui-se, então, que realizar a filtragem de amostras heterogêneas no processo de treinamento do SOM, conforme foi feito com a imagem da figura 5.16, é uma estratégia mais apropriada para salientar as bordas da *U-matrix* neste tipo de aplicação.

Assim sendo, para gerar bordas mais definidas na *U-matrix*, sem comprometer a natureza não-supervisionada dos métodos de classificação propostos, o presente trabalho propõe uma maneira de filtrar automaticamente janelas amostrais que apresentam um alto grau de heterogeneidade espectral, evitando que as mesmas sejam utilizadas no treinamento do SOM.

Considerando que as amostras coletadas a partir das imagens de entrada correspondem à janelas de pixels (pequenas imagens), a idéia consiste em aplicar uma medida estatística baseada na matriz de co-ocorrência de níveis de cinza, proposta por Haralick *et al.* (1973), para medir o grau de heterogeneidade espectral de cada amostra. As matrizes de co-ocorrência foram comentadas brevemente na seção 2.5.4.2.

Baseando-se em resultados experimentais (apresentados em Gonçalves *et al.* (2007a) e em Gonçalves *et al.* (2007b)), a *energia* (ou segundo momento angular), foi a medida estatística escolhida para medir o grau de heterogeneidade das amostras com base na matriz de co-ocorrência de seus níveis de cinza. Esta medida, descrita pela equação 5.3, retorna valores próximos de 1 quando a área de interesse apresenta textura uniforme (níveis de cinza próximos ou com alguma regularidade espacial) e valores que tendem a zero caso a área não seja uniforme.

$$ENE = \sum_i \sum_j (P(i, j)_{d, \theta})^2, \quad (5.3)$$

onde  $P(i, j)_{d, \theta}$  representa a frequência relativa com que dois pixels vizinhos  $i$  e  $j$ , separados por uma distância  $d$  na direção  $\theta$ , ocorrem em uma janela amostral. Nesta aplicação, a co-ocorrência dos pares de pixels foi feita de forma omnidirecional com uma distância  $d = 1$  e uma quantização de 4 bits para os níveis de cinza.

As janelas amostrais cujos valores de energia (ENE) satisfazem a relação dada abaixo são

consideradas heterogêneas e, conseqüentemente, filtradas.

$$ENE < \mu_{ENE} - \frac{1}{2}\sigma_{ENE}, \quad (5.4)$$

onde  $\mu_{ENE}$  e  $\sigma_{ENE}$  são, respectivamente, a média e o desvio padrão dos valores de energia de todas as amostras analisadas.

Vale observar que não foram empregadas medidas estatísticas de primeira ordem (como a variância ou desvio-padrão) para medir o grau de heterogeneidade das amostras, pois conforme explicado na sessão 4.2.1, apesar dessas medidas apresentarem uma complexidade computacional inferior, elas não incorporam informações espaciais sobre os pixels em seus cálculos, podendo gerar interpretações incorretas sobre a uniformidade espectral das amostras.

Com o objetivo de avaliar o procedimento proposto para realizar a filtragem de amostras heterogêneas, o mesmo foi aplicado nos conjuntos de janelas amostrais das imagens testes 1 e 2 (apresentadas nas figuras 4.13 e 5.8).

Dentre as 2499 janelas amostrais que foram extraídas da imagem teste 1 (figura 4.13), 635 foram classificadas como heterogêneas, uma vez que apresentaram valores de energia inferiores ao limiar da equação 5.4. Um SOM com as mesmas dimensões daquele utilizado no experimento com a imagem teste 1 (apresentado na seção 5.4.1), foi então treinado desconsiderando as amostras heterogêneas. A figura 5.24 mostra a *U-matrix* resultante. Comparando com a figura 4.16, que mostra a *U-matrix* do SOM treinado com todas as amostras (incluindo as heterogêneas), nota-se que houve uma diminuição do número de rugosidades e as superfícies mais baixas (que representam os agrupamentos) ficaram um pouco mais acentuadas.

Já na aplicação do procedimento de filtragem de amostras sobre a imagem da figura 5.8, que apresenta um número considerável de regiões de transição entre as classes, 1136 amostras das 3534 coletadas foram consideradas heterogêneas de acordo com o critério definido na equação 5.4. A figura 5.25 mostra a *U-matrix* resultante de um SOM treinado com as mesmas dimensões daquele do experimento com a imagem teste 2 (seção 5.4.1), mas sem utilizar as amostras heterogêneas. Comparando com a figura 5.10 que ilustra a *U-matrix* do SOM treinado com o conjunto completo de amostras, verifica-se, neste caso, que as bordas ficaram

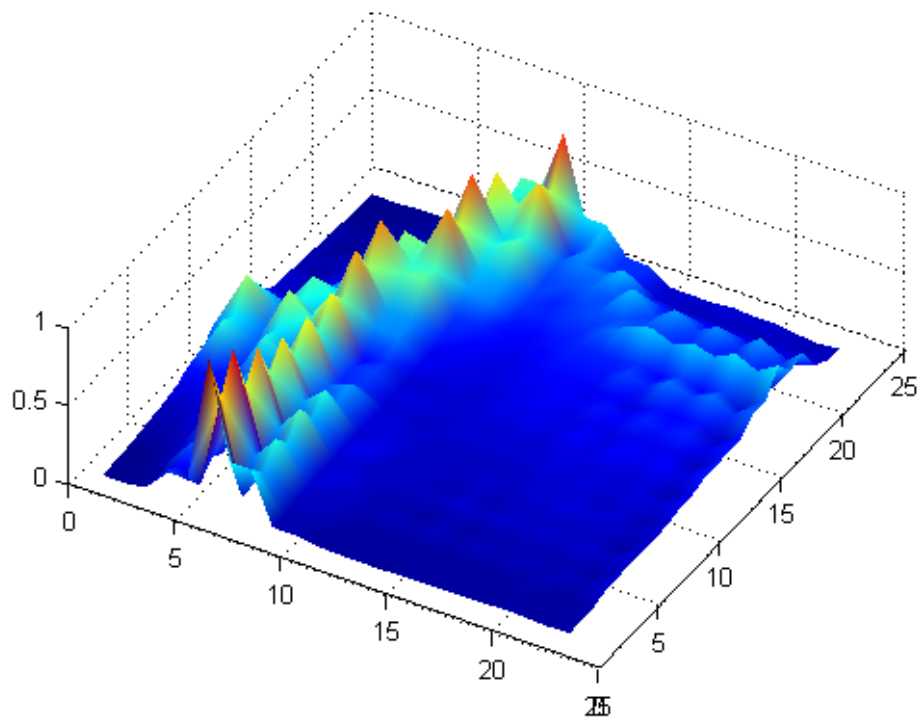


Fig. 5.24: *U-matrix* resultante do SOM treinado sem as amostras heterogêneas da imagem da figura 4.13.

bem mais definidas e a quantidade de rugosidades diminuiu consideravelmente.

Apesar da eficácia da estratégia proposta para acentuar as bordas da *U-matrix*, o tempo de processamento que é gasto para verificar o grau de heterogeneidade de todas as amostras deve ser considerado. No caso da imagem teste 1, foram gastos 83 segundos para analisar todas as 2499 amostras coletadas. Em contrapartida, como o SOM é treinado com um número menor de padrões, pois as amostras heterogêneas são descartadas, o seu tempo de treinamento diminuiu de 72 para 50 segundos. Apesar da redução no tempo de aprendizagem do mapa de neurônios, verifica-se que ela não foi suficiente para compensar o tempo consumido no processo de filtragem das amostras. O mesmo acontece para o caso da imagem teste 2. A diminuição no tempo de treinamento do SOM com a filtragem das amostras heterogêneas foi de 109 para 69 segundos (redução de 40 segundos), sendo que o tempo gasto para verificar o grau de heterogeneidade de todas as janelas amostrais foi de 101 segundos. Contudo, vale observar que embora as quantidades de amostras tenham sido diminuídas com a aplicação do processo de filtragem, os

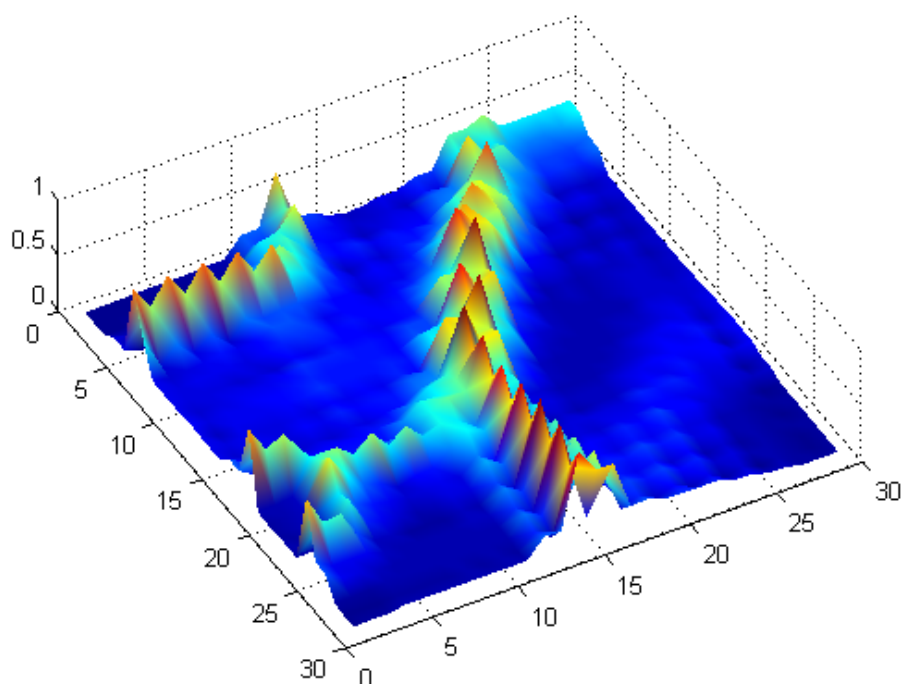


Fig. 5.25: *U-matrix* resultante do SOM treinado sem as amostras heterogêneas da imagem da figura 4.13.

tamanhos dos SOMs treinados não foram reduzidos na mesma proporção. Portanto, caso isso seja feito, os tempos de treinamento dos mapas podem ser ainda menores.

Outro ponto importante, que também deve ser observado na aplicação do procedimento proposto para realizar a filtragem de amostras, é a possibilidade de diminuir o número de partições que são analisadas na etapa de segmentação do SOM. Os experimentos apresentados anteriormente, mostram que realizar o processo de treinamento do SOM incorporando amostras consideradas heterogêneas resulta em uma *U-matrix* com bordas mais degradadas e com um número maior de rugosidades. Mesmo após a aplicação do processo de suavização, algumas regiões pouco significativas ainda permanecem na superfície da *U-matrix*, gerando partições do SOM que não necessitariam ser analisadas. Uma vez que a aplicação do processo de filtragem de amostras heterogêneas diminui o número de rugosidades e acentua as bordas da *U-matrix*, a chance de suprimir aquelas regiões (partições) pouco significativas se torna maior.

Isso pode ser observado no caso das duas imagens analisadas anteriormente. A figura 5.26 mostra o gráfico do número de regiões conectadas versus o limiar ( $k$ ) da *U-matrix* da imagem

teste 1, obtida a partir do SOM treinado sem amostras heterogêneas. Comparando com o gráfico do SOM treinado com as amostras heterogêneas (figura 5.4), nota-se que a seqüência estável de número de regiões conectadas com início no limiar  $k$  igual a 6 e um dos picos que aparece pouco antes do limiar  $k$  igual a 11 não aparecem no gráfico mostrado aqui (na figura 5.26). Como pode ser visto nos experimentos com a imagem teste 1 na seção 5.4.1, a seqüência estável referente ao limiar  $k$  igual a 6 gera um partição que é avaliada pelo método de classificação proposto. No entanto, observando essa partição na figura 5.5 e os valores apresentados para a mesma pelos índices de validação (tabela 5.6), conclui-se facilmente que ela não necessitaria ser considerada na análise.

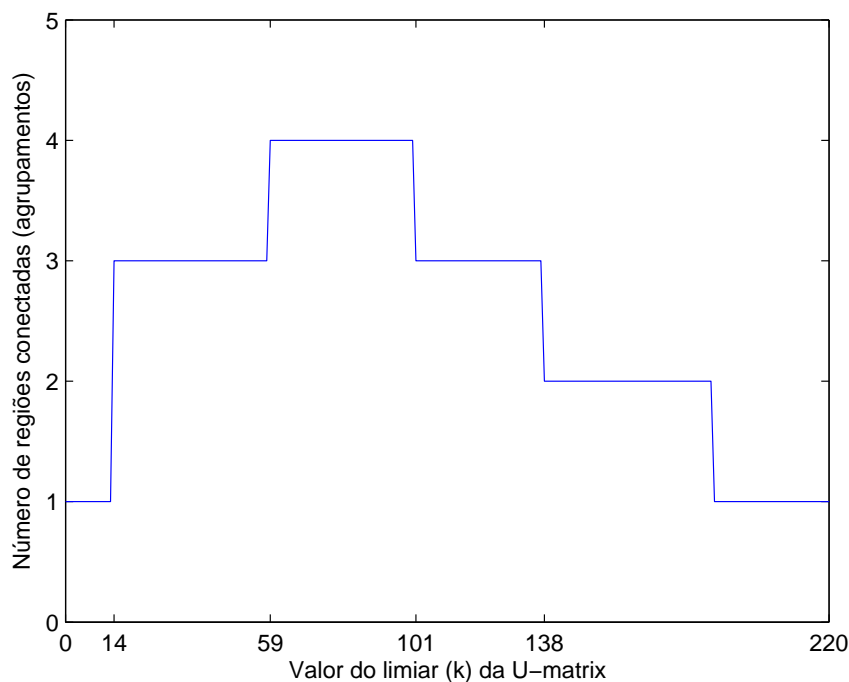


Fig. 5.26: Número de regiões conectadas versus limiar da *U-matrix* da imagem teste 1 derivada do SOM treinado sem amostras heterogêneas.

Já no caso da imagem teste 2, a redução do número de partições é maior. A figura 5.27 mostra o gráfico do número de regiões conectadas versus o limiar ( $k$ ) da *U-matrix* da imagem teste 2, obtida a partir do SOM treinado sem amostras heterogêneas. Há 7 seqüências estáveis de regiões conectadas, as quais geram 7 partições para serem analisadas. Neste caso, houve uma redução de 5 partições em relação ao total de partições apresentado pelo gráfico da



figura 5.11, o qual foi gerado a partir da *U-matrix* do SOM treinado com as amostras heterogêneas. Vale ressaltar que as partições mais significativas do SOM foram preservadas. No gráfico da figura 5.27, a partição referente ao valor de limiar  $k = 19$  corresponde àquela referente ao valor de limiar  $k = 17$  do gráfico da figura 5.11. Conforme explicado na seção 5.4.1.2, o limiar  $k = 17$  é escolhido para segmentar o SOM inicial caso o método de classificação proposto no capítulo 4 seja aplicado (esse valor de limiar corresponde à sequência mais estável de número de regiões conectadas). Já a partição referente ao valor de limiar  $k = 102$ , no gráfico da figura 5.27, corresponde àquela partição referente ao valor de limiar  $k = 128$  do gráfico da figura 5.11, que foi escolhida como melhor partição pelo método de classificação de imagens proposto neste capítulo.

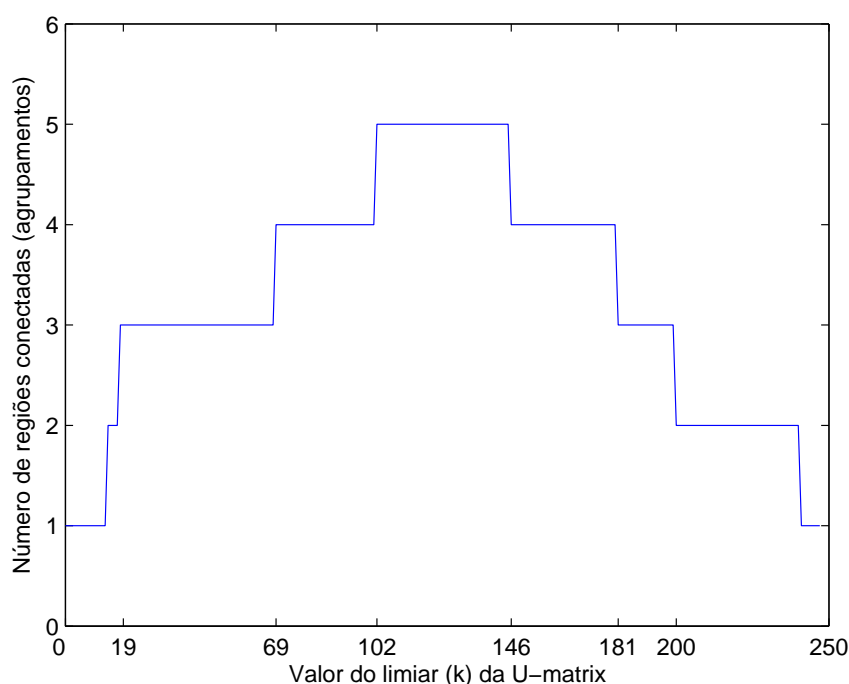


Fig. 5.27: Número de regiões conectadas versus limiar da *U-matrix* da imagem teste 2 derivada do SOM treinado sem amostras heterogêneas.

Diante desses resultados, pode-se concluir que a estratégia proposta para a filtragem de amostras heterogêneas não apenas reduz o número de rugosidades e gera bordas mais definidas na *U-matrix*, mas também, em consequência desses efeitos, pode suprimir partições pouco significativas, contribuindo inclusive para a diminuição do tempo global de processamento da

metodologia de classificação aplicada.

## 5.6 Abordagem paralela

Analisar diferentes partições para o arranjo de neurônios de um mesmo SOM é o ponto chave da estratégia de análise de agrupamentos empregada na metodologia de classificação de imagens proposta neste capítulo.

Vale destacar que tal estratégia apresenta um paralelismo inerente que pode ser explorado para reduzir o seu tempo de processamento. Dado um SOM treinado, quando o número de elementos de  $S_k$  é grande, ou seja, quando há muitas partições do SOM para serem analisadas, a execução do método de análise de agrupamentos em um ambiente multiprocessado pode ser uma alternativa apropriada para se obter uma resposta ainda mais rápida sobre a partição de melhor qualidade.

Nesse sentido, o paradigma mestre-escravo de computação paralela é sugerido como modelo para realizar a implementação paralela do método de análise de agrupamentos proposto. Nesse paradigma, as tarefas de computação são organizadas utilizando dois tipos de processos (ou programas): o processo mestre e o processo escravo. O processo mestre é o processo de controle, sendo responsável pela criação e inicialização de outros processos, e pela coleta e exibição de resultados. Os processos escravos efetuam as computações reais envolvidas de acordo com as suas cargas de trabalho alocadas pelo processo mestre.

Seguindo o modelo mestre-escravo, a programação paralela do método de análise de agrupamentos proposto pode ser implementada como segue:

1. O processo mestre determina todos os valores de  $k$  pertencentes a  $S_k$ , e os distribui entre os processos escravos;
2. Cada processo escravo, de acordo com o valor de  $k$  recebido, realiza o processo de particionamento do SOM e do conjunto de dados, e também aplica o índice de validação de agrupamentos para avaliar a partição efetuada;
3. O valor do índice obtido é enviado pelo processo escravo para o processo mestre;

4. Tendo recebido os valores do índice de validação de cada um dos processos escravos, o processo mestre determina qual partição é a melhor e, então, efetua o particionamento dos dados.

A figura 5.28 ilustra o modelo de implementação paralela sugerido para a estratégia de análise de agrupamentos proposta. Alternativamente, um número maior de partições pode ser analisada por cada processo escravo para aumentar a granularidade do processamento.

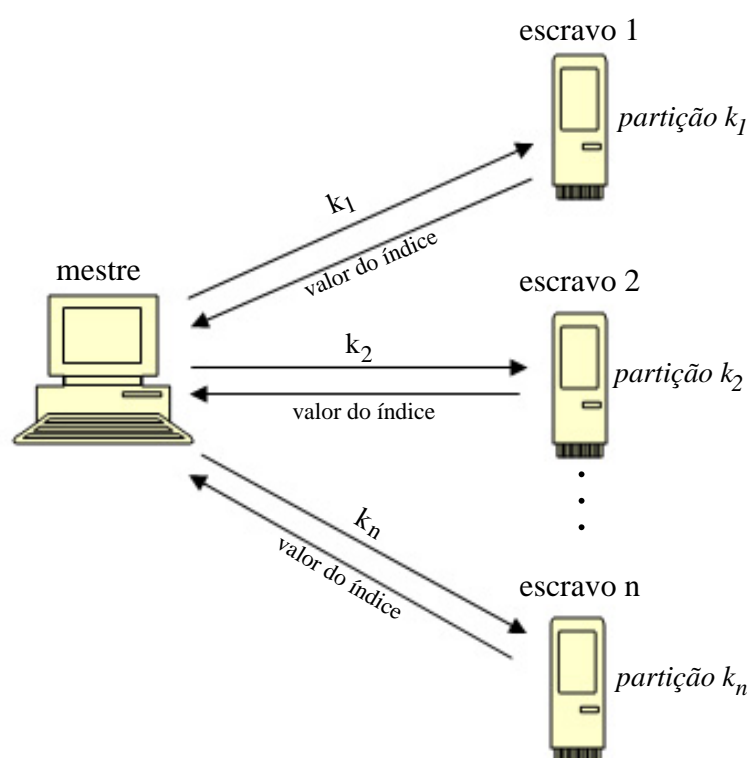


Fig. 5.28: Ilustração do modelo de implementação paralela sugerido para a estratégia de análise de agrupamentos empregada na metodologia de classificação proposta.

## 5.7 Considerações finais

Embora a metodologia de classificação apresentada no capítulo 4 apresente uma série de vantagens para realizar a classificação de imagens de sensoriamento remoto, as quais não são encontradas em métodos convencionais de classificação não-supervisionada, o seu custo de

processamento pode se tornar elevado se uma quantidade muito grande de SOMs for necessária para detectar os agrupamentos de dados das imagens. Dependendo dos tamanhos dos mapas utilizados, os tempos gastos com o treinamento dos neurônios podem ser expressivos em função do grande volume e da alta dimensionalidade dos dados de sensoriamento remoto.

Na tentativa de amenizar esse problema, neste capítulo foi proposta uma nova estratégia para realizar a detecção de agrupamentos existentes nas imagens de sensoriamento remoto. A estratégia proposta pode ser vista como um aprimoramento do método apresentado por Costa e Andrade Netto (1999, 2001) (aplicado na metodologia de classificação do capítulo 4). Ao invés de definirmos apenas uma imagem de marcadores para segmentar a *U-matrix*, várias imagens de marcadores são determinadas e, portanto, diferentes segmentações da *U-matrix* são obtidas. Desse modo, a estratégia proposta considera diferentes partições para o SOM treinado e avalia a qualidade dos agrupamentos que são gerados por eles através da aplicação de índices de validação de agrupamentos. Resultados experimentais sobre duas imagens testes mostram que a utilização dessa nova estratégia permite encontrar de forma mais eficaz os agrupamentos de dados existentes na imagem, uma vez que exige uma quantidade menor de SOMs do que o método apresentado no capítulo 4.

Os tempos de processamento gastos pela metodologia de classificação proposta neste capítulo na classificação das imagens testes foram praticamente iguais aos tempos consumidos pelo algoritmo K-médias. Além disso, as qualidades das classificações resultantes foram bastante superiores que as do método de classificação convencional. Vale observar ainda, que o método de classificação de imagens proposto neste capítulo herda quase todas as vantagens e potencialidades que foram apresentadas pelo método de classificação apresentado no capítulo 4.

Na busca por índices de validação adequados para avaliar a qualidade das diferentes partições de dados das imagens de sensoriamento remoto, no presente capítulo também foi apresentado um estudo comparativo entre diferentes índices de validação pré-selecionados com base na literatura existente. Uma série de testes experimentais, considerando diferentes partições de conjuntos de dados com características variadas, serviram de base para determinar aqueles índices que apresentam um maior potencial para avaliar agrupamentos de dados de imagens de sensoriamento remoto.

Levando em conta ainda, que a maioria dos índices de validação apresentam um custo de processamento elevado, principalmente quando aplicados em conjuntos de dados volumosos, também foi apresentada uma estratégia para modificar os seus cálculos com base na quantização vetorial produzida pelo SOM. Tal estratégia possibilitou reduzir o tempo de processamento gasto com a aplicação dos índices, viabilizando, portanto, a utilização dos mesmos na análise de dados de sensoriamento remoto. Os resultados experimentais mostram que apesar da existência de erros de aproximação, as versões modificadas dos índices de validação apresentam comportamentos semelhantes aos das suas versões originais na avaliação de diferentes partições de conjuntos de dados, não interferindo no processo de tomada de decisão sobre a escolha das melhores partições.

Por fim, tendo detectado que o processo de amostragem das imagens por meio de janelas de pixels incorpora amostras heterogêneas que degradam a superfície topográfica da *U-matrix*, neste capítulo foi apresentado também um procedimento alternativo, baseado em medidas de textura, para realizar de forma automática a filtragem dessas janelas amostrais heterogêneas. Resultados experimentais mostram que os SOMs treinados com o conjunto amostral filtrado, ou seja, sem as amostras heterogêneas, geram *U-matrizes* com bordas mais definidas, o que pode facilitar o processo de particionamento dos dados através dos protótipos dos SOMs.



# Capítulo 6

## Classificação baseada em SOMs segmentados hierarquicamente com informações espectrais e espaciais

### 6.1 Introdução

A estratégia de analisar os agrupamentos das imagens através de um conjunto de protótipos do SOM, ao invés de processar os padrões originais das cenas (ponto chave dos métodos de classificação propostos neste trabalho) reduz significativamente a complexidade da tarefa de classificação das imagens de sensoriamento remoto, tornando possível empregar métodos que normalmente são considerados computacionalmente inviáveis para este domínio de aplicação. No capítulo 5, por exemplo, baseando-se na quantização vetorial produzida pelo SOM, foi possível aplicar índices de validação de agrupamentos para avaliar as diferentes partições produzidas para os conjuntos de dados das imagens, e conforme ressaltado naquele mesmo capítulo, índices de validação de agrupamentos são freqüentemente ignorados na maioria dos pacotes ou softwares de processamento digital de imagens, em função da elevada complexidade computacional apresentada pelos mesmos.

Buscando explorar ainda mais as vantagens do mapeamento produzido pelo SOM, o presente capítulo propõe o uso combinado do SOM e de métodos de agrupamentos hierárquicos

para realizar a classificação não-supervisionada das imagens. Conforme descrito em Gonçalves *et al.* (2008c), métodos de agrupamentos hierárquicos apresentam algumas características que dificultam a sua aplicação no reconhecimento de padrões de imagens de sensoriamento remoto: (a) em geral requerem espaço de memória da ordem de  $O(N^2)$ , onde  $N$  é o número de registros do conjunto de dados; (b) dendrogramas são impraticáveis para grandes conjuntos de dados, podendo ser difíceis de interpretar; (c) para determinar qual a linha de "corte" do dendrograma (melhor partição) há a necessidade de se aplicar algum critério de decisão.

De fato, na literatura, não há muita pesquisa ou trabalhos que empregam métodos hierárquicos na classificação não-supervisionada de imagens de sensoriamento remoto e, em geral, versões comuns desses métodos são aplicadas quando o objetivo principal é basicamente comparar desempenhos com outros métodos não-supervisionados (Duda e Canty, 2002; Wilson *et al.*, 2002). Por essas razões, a grande maioria dos sistemas computacionais de sensoriamento remoto apresenta apenas métodos de agrupamentos particionais para efetuar a classificação não-supervisionada de imagens, como o K-médias e o ISODATA. Módulos de agrupamentos hierárquicos são muito raros nesses sistemas e, por isso, são normalmente utilizados de maneira desacoplada.

O presente capítulo mostra que ao realizar a análise de agrupamentos das imagens através do SOM, e não diretamente com os dados das cenas, a aplicação de métodos hierárquicos se torna viável. Na metodologia de classificação proposta neste capítulo, um método de agrupamento hierárquico é aplicado para particionar os vetores de protótipos do SOM, os quais representam os padrões originais da cena, gerando, desse modo, um dendrograma simplificado para a imagem de entrada. As relações de vizinhança do SOM são utilizadas para restringir possíveis fusões na construção do dendrograma, e a sua linha de "corte" (melhor partição) é determinada através da aplicação das versões modificadas de índices de validação propostas no capítulo 5.

Outro diferencial importante do método de agrupamento hierárquico proposto aqui para segmentar o SOM, está relacionado com o seu critério de fusão. As fusões entre protótipos (ou conjuntos de protótipos) do SOM não são decididas considerando somente a informação de distância espectral entre os agrupamentos, como é feito em abordagens convencionais. Informações espaciais sobre os agrupamentos também são incorporadas com o propósito de aumentar



a eficácia do processo de separação das classes, especialmente em situações em que os padrões de cobertura terrestre da imagem apresentam um alto grau de similaridade espectral.

Diferentemente da abordagem apresentada no capítulo 5, na qual um procedimento alternativo de filtragem de amostras heterogêneas é proposto para ser aplicado sobre o conjunto amostral das imagens (antes de realizar o treinamento do SOM), na metodologia de classificação proposta neste capítulo, o mesmo procedimento de filtragem é aplicado, mas apenas sobre os protótipos do SOM treinado. Resultados experimentais mostram que os protótipos heterogêneos (aqueles que representam mais de uma classe de cobertura terrestre), juntamente com os protótipos inativos, funcionam como bordas no mapa de neurônios do SOM treinado facilitando a discriminação dos agrupamentos.

Para testar a sua eficácia, a metodologia de classificação proposta neste capítulo foi aplicada sobre uma imagem IKONOS e os seus resultados são comparados com uma versão "janelada" do algoritmo K-médias e com os métodos de classificação propostos nos capítulos anteriores.

O presente capítulo mostra, ainda, resultados experimentais que comprovam a possibilidade de reutilização de SOMs para classificar padrões de imagens de sensoriamento remoto para os quais eles não foram treinados.

## **6.2 Metodologia de classificação de imagens usando SOMs segmentados hierarquicamente com informações espectrais e espaciais**

A metodologia de classificação proposta neste capítulo busca explorar, conjuntamente, as propriedades do SOM e de métodos de agrupamentos hierárquicos para realizar a classificação não-supervisionada de imagens de sensoriamento remoto. Seguindo a mesma estratégia utilizada pelos métodos de classificação propostos nos capítulos anteriores, a análise de agrupamentos da imagem é feita através dos protótipos do SOM. No entanto, a segmentação do mapa de neurônios é efetuada empregando um método de agrupamento hierárquico que utiliza mais informação em seu critério de fusão do que os métodos convencionais. A figura 6.1 mostra

de maneira simplificada a metodologia proposta e a seguir são apresentadas as suas etapas principais:

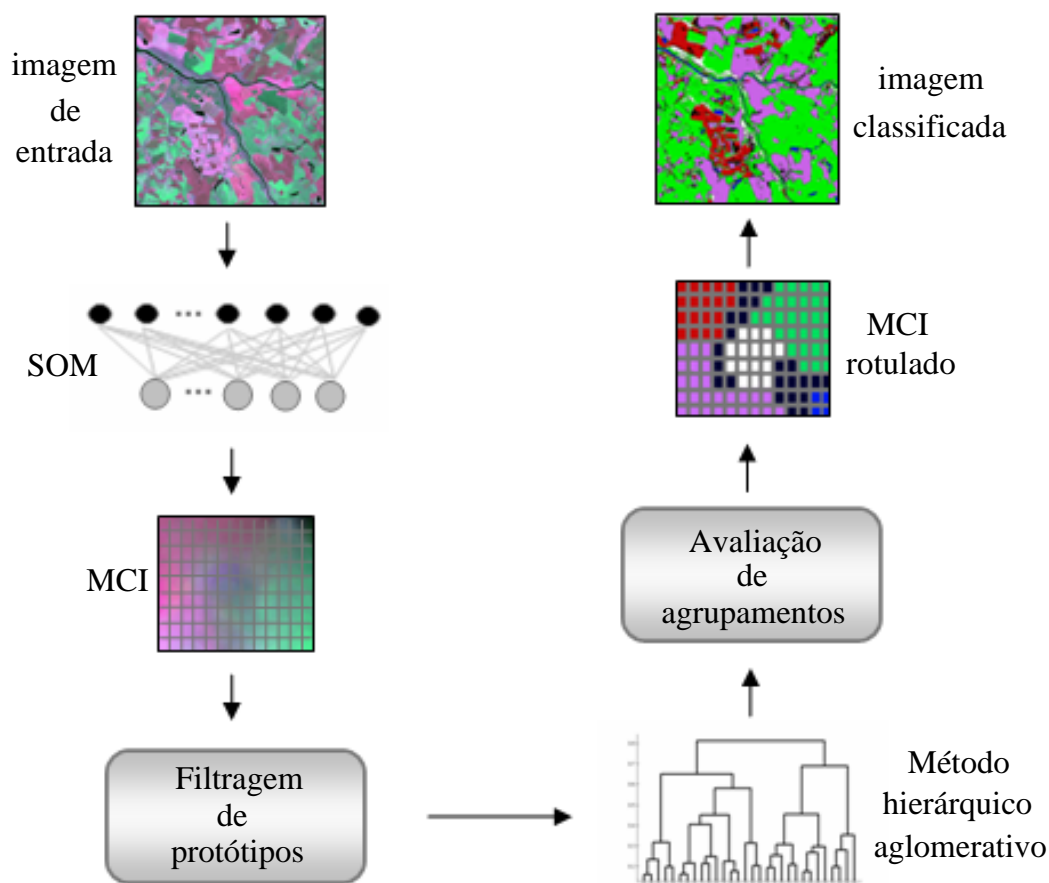


Fig. 6.1: Ilustração do método de classificação de imagens proposto usando SOMs segmentados hierarquicamente.

1. *Amostragem*: amostras na forma de janelas de pixels são coletadas de maneira uniforme sobre toda a região da imagem e fornecidas como padrões de entrada para o SOM.
2. *Treinamento do SOM*: o treinamento não-supervisionado do SOM é realizado e o MCI (Mapa de Características da Imagem) é obtido.
3. *Filtragem de protótipos*: os protótipos do SOM que apresentam atividade nula são descartados e aqueles que apresentam alto grau de heterogeneidade espectral (protótipos hete-

rogêneos) são excluídos do próximo passo.

4. *Análise hierárquica*: um método de agrupamento hierárquico aglomerativo, cujo critério de fusão se baseia nas relações de vizinhança dos protótipos na grade do SOM e, também, nas informações espaciais (geográficas) dos padrões que são representados por eles, é aplicado sobre o mapa de neurônios filtrado, gerando um dendrograma simplificado dos dados da imagem.
5. *Avaliação dos agrupamentos*: um índice de validação de agrupamentos calculado de forma modificada é aplicado em todos os níveis do dendrograma obtido no passo 4. O nível do dendrograma que apresenta o valor ótimo para o índice de validação é escolhido como a melhor configuração de agrupamentos para os protótipos do SOM (melhor partição).
6. *Classificação do SOM*: os protótipos do SOM são rotulados de acordo com as classes de protótipos determinadas no passo 5.
7. *Classificação da imagem*: os protótipos do SOM rotulados no passo 6, em conjunto com os protótipos heterogêneos (filtrados no passo 3), são utilizados como referência para classificar todos os pixels da imagem.
8. *Reclassificação*: cada um dos pixels da imagem que foram associados aos protótipos heterogêneos é reclassificado utilizando a classe do pixel vizinho que apresenta a menor distância (espectral) a ele.

Nas próximas subseções uma breve descrição das etapas 3, 4, 5, 7 e 8 do método proposto é apresentada. As demais etapas são realizadas da mesma forma que foram apresentadas nos métodos de classificação propostos nos capítulos anteriores.

### **6.2.1 Filtragem de protótipos**

A etapa 3 da metodologia de classificação proposta consiste em filtrar dois tipos de protótipos que geralmente aparecem no mapeamento dos padrões da imagem através do SOM. Esses protótipos são denominados aqui de *inativos* e *heterogêneos*.

Os protótipos *inativos* correspondem àqueles neurônios que apresentam atividade nula no processo de aprendizado competitivo do SOM e, portanto, não possuem nenhum padrão de entrada associado a eles. Esses protótipos são simplesmente descartados do processo não passando para as próximas etapas da análise.

Os protótipos *heterogêneos* são aqueles que apresentam alto grau de heterogeneidade espectral e estão associados normalmente a padrões de entrada que possuem mais do que uma classe de cobertura terrestre. Conforme descrito na seção 5.5, na maioria das vezes esses padrões (amostras heterogêneas) correspondem a regiões de transição entre classes de cobertura terrestre presentes na imagem e são capturados em decorrência da amostragem realizada por janelas de pixels.

Os protótipos considerados heterogêneos são excluídos do passo 4 da análise, onde é aplicado o método de agrupamento hierárquico. Essa exclusão é realizada com o objetivo de não permitir que esses protótipos e, conseqüentemente, os padrões associados a eles sejam atribuídos erroneamente a uma das classes que os compõe. Os protótipos heterogêneos podem ser vistos como padrões ruidosos ou discrepantes, e caso não sejam filtrados, o método hierárquico pode incorporá-los nos agrupamentos que serão produzidos ou retê-los em agrupamentos separados. Além disso, como em métodos hierárquicos a fusão de dados ou de agrupamentos em um determinado estágio não pode ser corrigida em estágios subseqüentes, interpretações erradas sobre as classes e/ou número de classes da imagem podem ser realizadas, comprometendo deste modo a exatidão dos resultados da classificação. Sendo assim, os protótipos heterogêneos são alternativamente reincorporados apenas nos procedimentos adotados nos passos 7 e 8 do método proposto.

Os procedimentos para medir o grau de heterogeneidade de cada protótipo do SOM treinado, e para filtrar aqueles que apresentam um alto grau de heterogeneidade espectral, são os mesmos que foram propostos na seção 5.5 para filtrar amostras heterogêneas (equações 5.3 e 5.4).

Vale destacar que o tempo de processamento consumido por esta etapa de filtragem de protótipos é compensado não apenas pela redução do número de protótipos que serão analisados no passo 4 do método, mas principalmente por um possível aumento na precisão de classificação da imagem. Adicionalmente, conforme mencionado anteriormente, protótipos inativos e hete-

rogêneos podem funcionar como "unidades de interpolação" (bordas) no mapa de neurônios, facilitando a separação das classes de agrupamentos.

### 6.2.2 Análise hierárquica

No quarto passo da metodologia de classificação proposta, um método de agrupamento hierárquico aglomerativo é aplicado sobre o conjunto de protótipos do SOM já treinado e filtrado. O objetivo desse passo é agrupar protótipos e conjuntos de protótipos em diferentes níveis de similaridade, buscando descobrir a estrutura (ou classes) dos mesmos.

O método hierárquico proposto para segmentar o SOM treinado apresenta duas características importantes. A primeira delas é a imposição de restrições às possíveis fusões dos protótipos do SOM. Um algoritmo hierárquico aglomerativo, uma vez aplicado sobre os agrupamentos de neurônios do SOM, deve respeitar as relações topológicas do espaço de saída da rede. Diferentemente da abordagem tradicional de métodos de agrupamentos hierárquicos, que consiste em comparar todos os pares de objetos ou grupos de objetos para decidir sobre uma fusão, a abordagem utilizada neste trabalho verifica a possibilidade de efetuar fusões somente entre pares de neurônios ou grupos de neurônios que sejam adjacentes (ou vizinhos) na grade do SOM. Conforme mostra a figura 6.2, dada uma grade  $m \times n$  de neurônios, o neurônio na posição (3,2), por exemplo, somente tem possibilidade de fusão com os neurônios vizinhos nas posições (2,1), (2,2), (2,3), (3,1), (3,3), (4,1), (4,2) e (4,3). Essa abordagem diminui significativamente o volume de processamento do método de agrupamento hierárquico aplicado.

Outra característica importante do método hierárquico aplicado para segmentar o SOM é que, além de considerar a informação de distância multiespectral para definir qual par de protótipos (ou de conjuntos de protótipos) será agrupado em cada iteração, o critério de fusão empregado também se baseia em informações espaciais dos pixels da imagem associados aos protótipos do SOM.

A informação de distância ( $D_{ij}$ ) no espaço multiespectral entre dois agrupamentos de protótipos  $i$  e  $j$  do SOM é calculada usando a métrica euclidiana e o método dos vizinhos mais próximos (ou de ligações simples). Os valores de  $D_{ij}$  são normalizados para o intervalo  $[0,1]$ .

As informações espaciais são calculadas através da utilização de dois índices propostos

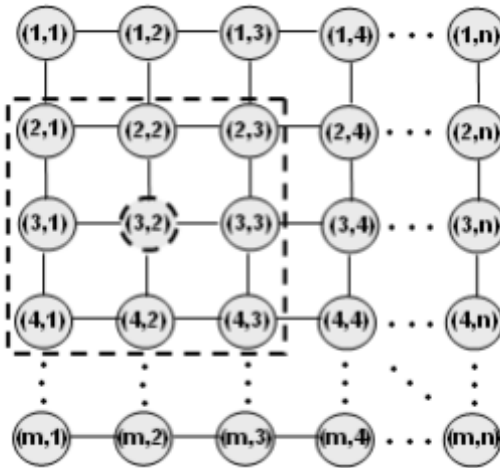


Fig. 6.2: Neurônios adjacentes na grade do SOM.

em Castro (2005). Esses índices, denominados de *índice de fronteira espacial* e *índice de compactidade espacial*, são adaptados e calculados nesta abordagem através da imagem classificada com base nos agrupamentos (ou classes) de protótipos do SOM.

O *índice de fronteira espacial* ( $IFE_{ij}$ ) calcula o tamanho da fronteira entre todos os pares de classes de protótipos  $(i,j)$  considerando oito vizinhos para cada pixel (quatro adjacentes e quatro oblíquos). A sua fórmula é dada como segue:

$$IFE_{ij} = 1 - \frac{1}{2} \left( \frac{f_{ij}}{\sum_{k=1}^N (k \neq i) f_{ik}} + \frac{f_{ij}}{\sum_{k=1}^N (k \neq j) f_{kj}} \right), \quad (6.1)$$

onde  $f_{ij}$  representa a quantidade de vezes que as classes de protótipos do SOM  $i$  e  $j$  são vizinhas na imagem e  $N$  é o número de classes no nível do dendrograma que está sendo analisado. Este índice gera valores dentro do intervalo numérico  $[0,1]$  e apresenta valores baixos quando a fronteira entre o par de classes analisado é significativa. A utilização deste índice se justifica pela idéia de que duas classes que possuem uma fronteira espacial significativa devem estar mais propensas a uma fusão do que classes que possuem pouca ou nenhuma fronteira espacial na imagem.

O *índice de compactidade espacial* ( $ICE_{ij}$ ), definido através da equação 6.2, baseia-se na quantidade de vezes ( $f_{ii}$ ) que pixels de uma mesma classe de protótipos  $i$  são vizinhos na imagem. Este índice penaliza a fusão de classes espacialmente compactas na imagem. Ele

apresenta valores baixos (dentro do intervalo  $[0,1]$ ) quando a classe é composta, em sua maioria, por pixels isolados.

$$ICE_{ij} = \frac{1}{2} \left( \frac{f_{ii}}{f_{ii} + 5 \sum_{k=1(k \neq i)}^N f_{ik}} + \frac{f_{jj}}{f_{jj} + 5 \sum_{k=1(k \neq j)}^N f_{kj}} \right). \quad (6.2)$$

O critério de fusão proposto estabelece que o par de classes de protótipos  $(i,j)$  do SOM que apresentar o valor mais baixo para o índice dado pela equação 6.3, chamado aqui de *índice de agrupamento espectro-espacial* ( $IA_{ij}$ ), é aquele que deve ser escolhido para a fusão.  $IA_{ij}$  é determinado para todos os pares  $i$  e  $j$  de protótipos (ou conjuntos de protótipos) do SOM e é dado pela média dos valores de  $D_{ij}$ ,  $IFE_{ij}$  e de  $ICE_{ij}$ , ou seja:

$$IA_{ij} = \frac{D_{ij} + IFE_{ij} + ICE_{ij}}{3}. \quad (6.3)$$

Sendo assim, o algoritmo do método de agrupamento hierárquico aglomerativo aplicado para segmentar o SOM pode ser descrito como segue:

1. Inicie com  $N$  agrupamentos, cada um contendo um único protótipo do SOM.
2. Determine o par de agrupamentos distintos de protótipos  $i$  e  $j$ , de modo que:

$$IA_{ij} = \min \{IA_{ml}\} \quad (6.4)$$

onde  $1 \leq m, l \leq N, m \neq l$ ,  $m$  e  $l$  vizinhos (adjacentes) no SOM e  $IA_{ij}$  é dado pela equação 6.3.

3. Forma-se um novo agrupamento de protótipos  $k$  pela união dos agrupamentos  $i$  e  $j$ .
4. Calculam-se os novos valores para o índice de agrupamento espectro-espacial  $IA_{kl}$ , entre o novo agrupamento  $k$  e os restantes vizinhos a ele.
5. Repita os passos 2, 3 e 4 até que todos os protótipos do SOM estejam em um único agrupamento.

No final desta etapa, tem-se um dendrograma que mostra de maneira hierárquica a relação de semelhança entre os protótipos e agrupamentos de protótipos do SOM. Lembrando que os protótipos do SOM representam todos os padrões de entrada, o dendrograma gerado reflete, portanto, as relações existentes entre os dados e agrupamentos de dados da imagem original, podendo ser visto então como um dendrograma simplificado da imagem.

### 6.2.3 Avaliação dos agrupamentos

Em cada nível do dendrograma obtido no passo anterior há uma configuração diferente de agrupamentos para os protótipos do SOM e, conseqüentemente, para o conjunto de dados da imagem. Portanto, como em qualquer método de agrupamento hierárquico, há a necessidade de se aplicar algum critério para verificar qual nível do dendrograma apresenta a melhor configuração de agrupamentos (melhor partição).

De acordo com Vesanto e Alhoniemi (2000), a solução mais típica é cortar o dendrograma no nível que apresenta a maior distância entre dois agrupamentos que foram unidos. Porém, essa abordagem ignora a informação de dispersão interna dos grupos. De fato, cada agrupamento pode ser composto por diversos agrupamentos menores e, portanto, uma medida de dispersão intra-agrupamentos pode ser importante para avaliar também o grau de compacidade dos grupos que são formados. Sendo assim, a idéia utilizada nesta abordagem, foi a de aplicar algum índice de validação de agrupamento que incorpore medidas de dispersão intra e inter-agrupamentos.

Na metodologia de classificação proposta, a versão modificada do índice de validação CD<sub>bw</sub> ( $\mathcal{V}_{CD_{bw}}^{SOM}$ ) foi escolhida para avaliar as configurações de agrupamentos de todos os níveis do dendrograma obtido. Conforme mostrado no capítulo 5, o índice CD<sub>bw</sub> baseia-se em informações de dispersão intra e inter-agrupamentos, e apresenta um desempenho superior em comparação com outros índices de validação na avaliação de agrupamentos de dados com características variadas, incluindo dados de imagens de sensores remotos.

O índice  $\mathcal{V}_{CD_{bw}}^{SOM}$  é então aplicado em cada nível do dendrograma de protótipos do SOM. O nível do dendrograma que apresenta o maior valor para o índice é considerado, portanto, aquele que apresenta a melhor partição para os dados da imagem.



### 6.2.4 Classificação da imagem

Para realizar a classificação da imagem, os protótipos heterogêneos que foram filtrados no passo 3 são agora reincorporados e considerados também como uma classe. Para isso, todos eles recebem um mesmo rótulo, que deve obviamente ser diferente daqueles usados para rotular as outras classes de protótipos detectadas no passo anterior. Na seqüência, a imagem de entrada é rotulada seguindo o mesmo procedimento adotado nos métodos de classificação propostos nos capítulos anteriores. Janelas de pixels da imagem com dimensões iguais a das amostras de treinamento são comparadas com todos os protótipos do SOM. Essa comparação é realizada através das distâncias calculadas entre a janela de pixels considerada e cada um dos protótipos. O pixel central da janela de pixels recebe o rótulo do protótipo que apresenta a menor distância a ela. A imagem é então inteiramente percorrida até que todos os pixels sejam classificados dessa forma.

### 6.2.5 Reclassificação

No último passo do método proposto, o resultado do processo de classificação realizado na etapa anterior é aprimorado. Os pixels da imagem que receberam o rótulo da classe de protótipos heterogêneos são reclassificados. Cada um desses pixels é comparado com os seus pixels vizinhos na imagem que não foram rotulados pela classe de protótipos heterogêneos. Eles recebem um novo rótulo que será igual ao do vizinho que apresenta a distância mínima a ele em termos de atributos espectrais. Esse procedimento impede que pixels que estejam em regiões de transição de classes de cobertura terrestre sejam classificados erroneamente.

## 6.3 Resultados Experimentais

Esta seção mostra um exemplo de aplicação da metodologia de classificação proposta neste capítulo sobre uma imagem do satélite IKONOS. A imagem apresenta uma resolução espacial de 4m, possui 350 linhas por 384 colunas e é composta por três bandas espectrais. A cena mostra pivôs de irrigação na região de Andaraí no estado da Bahia e apresenta 6 classes de cobertura terrestre: vegetação rala (pastagem), mata, dois tipos de solo exposto e duas classes

de cultura agrícola. A figura 6.3 mostra uma composição colorida da imagem utilizando as três bandas espectrais.



Fig. 6.3: *Composição colorida (3R 2G 1B + PAN) da imagem IKONOS (provida pela Enge-sat/Brasil, ©Space Imaging).*

Seguindo os passos da metodologia de classificação proposta, 1292 janelas amostrais de tamanho 9x9 foram coletadas em intervalos regulares de 10 pixels a partir da imagem original. Esse conjunto de janelas amostrais foi usado para treinar um SOM composto por 144 neurônios arranjados em uma grade retangular 12x12.

A figura 6.4 mostra o MCI após o treinamento da rede, onde é possível visualizar os agrupamentos que correspondem às classes de cobertura terrestre presentes na imagem. Como nos experimentos mostrados nos capítulos anteriores, também é possível visualizar através do MCI as propriedades de ordenação topológica e de aproximação de densidade do mapeamento produzido pelo SOM. Os conjuntos de protótipos das classes solo exposto 1 e 2 são vizinhos no MCI, uma vez que apresentam atributos espectrais semelhantes. O mesmo acontece com os conjuntos de protótipos das classes mata e cultura agrícola 1. Quanto à aproximação de

densidade dos dados de entrada, as classes cultura agrícola 1 e 2, que ocupam áreas menores na imagem, são representadas por um número menor de protótipos no MCI. Já as classes solo exposto 1 e 2, que ocupam a maior parte da cena, são representados por uma quantidade maior de protótipos.

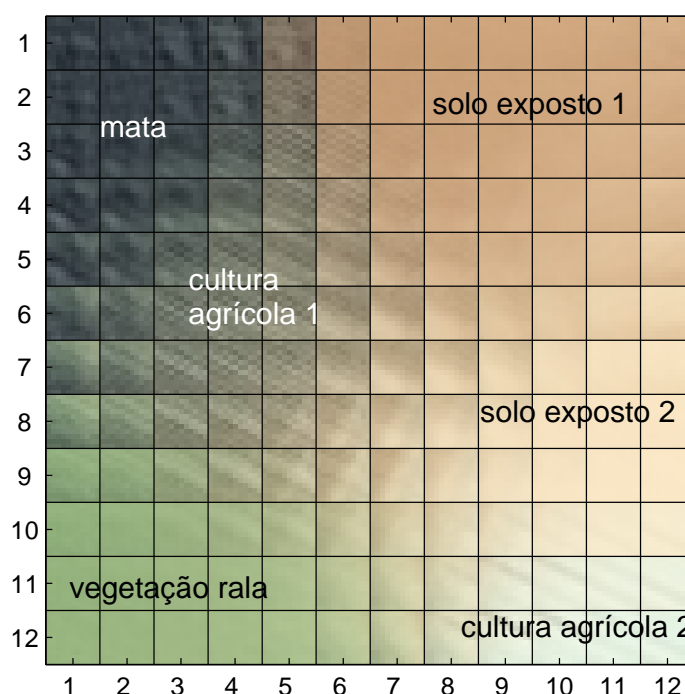


Fig. 6.4: MCI da imagem da figura 6.3 após o treinamento do SOM.

Após o treinamento do SOM, a etapa de filtragem de protótipos foi aplicada. Nesse experimento, 15 protótipos do SOM apresentaram atividade nula, ou seja, não apresentaram nenhum padrão de entrada associado a eles. A figura 6.5 mostra o histograma de vencimentos dos protótipos do SOM treinado. Os protótipos marcados em vermelho são os inativos.

Ainda na etapa de filtragem de protótipos, 18 neurônios do SOM apresentaram alto grau de heterogeneidade espectral de acordo com o critério estabelecido através das equações 5.3 e 5.4. A figura 6.6 mostra os valores de energia ( $ENE$ ) de cada um dos protótipos do SOM. Os 18 protótipos marcados em vermelho são aqueles cujos valores de energia não satisfizeram a relação dada pela equação 5.4.

Através das figuras 6.5 e 6.6, é possível observar ainda que, dos 18 protótipos considerados

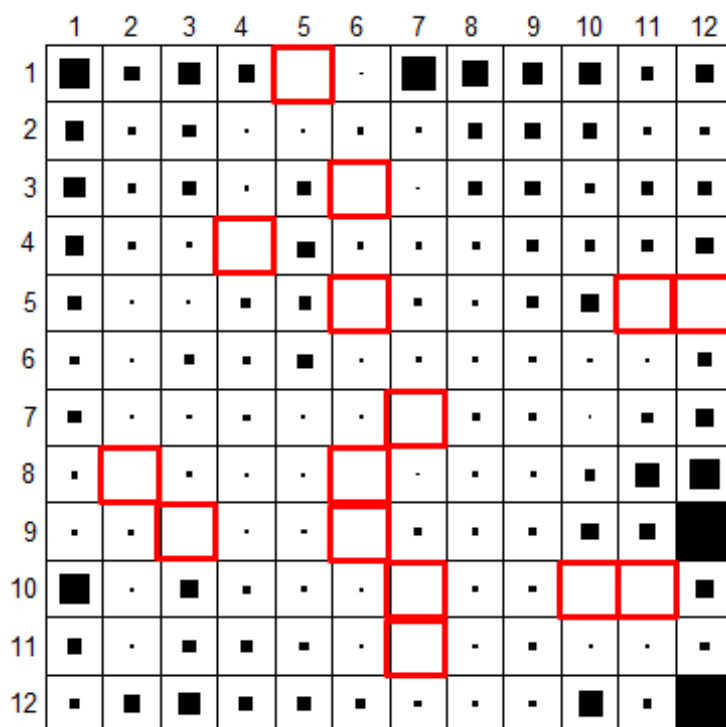


Fig. 6.5: Histograma de vencimentos do SOM treinado com os dados da imagem da figura 6.3. Os protótipos marcados em vermelho são aqueles que apresentaram atividade nula (inativos).

heterogêneos, 4 deles são inativos. Assim, dos 144 protótipos do SOM, 29 deles foram filtrados (15 inativos e 14 heterogêneos), restando, portanto, 115 protótipos para serem analisados na etapa seguinte.

Na seqüência, o método hierárquico, descrito na seção 6.2.2, foi aplicado sobre os protótipos do SOM filtrado. Conseqüentemente, um dendrograma consistindo de 115 níveis foi gerado, cada nível com uma configuração diferente de agrupamentos de protótipos.

A versão modificada do índice de validação CD<sub>bw</sub> ( $\mathcal{V}_{CD_{bw}}^{SOM}$ ) foi aplicada em todos os níveis do dendrograma para determinar qual a melhor configuração de agrupamentos dentre as 115 obtidas. O gráfico da figura 6.7 mostra os valores do índice  $\mathcal{V}_{CD_{bw}}^{SOM}$ . O gráfico apresenta somente os valores entre os níveis 97 e 114 do dendrograma. Para os níveis inferiores a 97, os valores do índice são menores do que aqueles apresentados no gráfico, ou até mesmo nulos, devido à existência de níveis que possuem agrupamentos com apenas um único protótipo. O número de agrupamentos ou de classes do nível 97 ao 114 decresce, respectivamente, de 19 até 2. Conforme

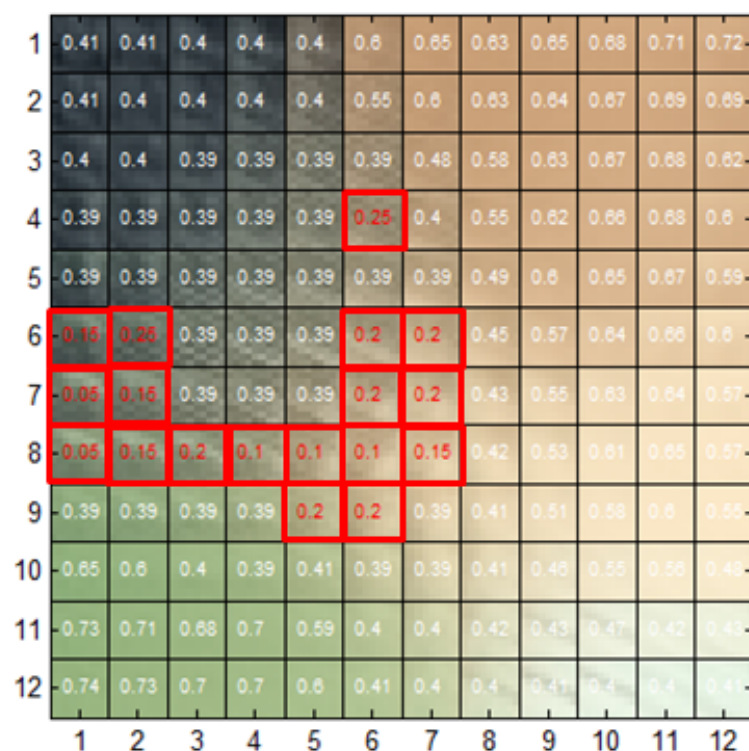


Fig. 6.6: Valores de energia ( $ENE$ ) de cada um dos protótipos do SOM. Os protótipos marcados em vermelho são aqueles cujos valores de energia não satisfizeram a relação dada pela equação 5.4.

pode ser visto no gráfico, o nível 110 foi aquele que apresentou o maior valor para o índice  $e$ , portanto, foi considerado a melhor partição para os protótipos do SOM da imagem da figura 6.3. O apêndice A.8, inserido na mídia anexada a essa tese, mostra o dendrograma produzido pelo método hierárquico utilizado, destacando a partição dos protótipos do SOM no nível 110.

A figura 6.8 mostra o arranjo de neurônios do SOM classificado de acordo com a partição do nível 110. Comparando com a imagem do arranjo (MCI) apresentado na figura 6.4, pode-se observar que o mapa de neurônios foi particionado em 6 grupos, correspondendo adequadamente às 6 classes de cobertura terrestre existentes na imagem original. Ainda na figura 6.8, também é possível visualizar os protótipos do SOM que foram filtrados (inativos e heterogêneos).

É importante destacar que, particularmente para a imagem utilizada neste experimento, uma segmentação adequada do SOM somente é alcançada devido ao critério de fusão empregado pelo método hierárquico proposto. Se apenas medidas de distâncias no espaço multiespectral

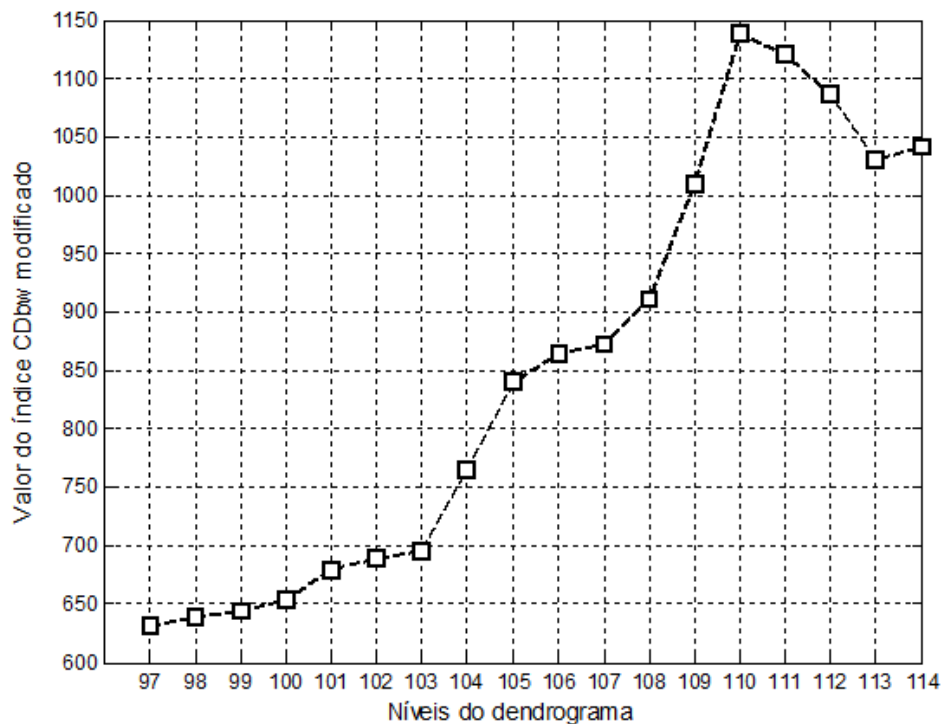


Fig. 6.7: Valores do índice  $\mathcal{V}_{CD_{bw}}^{SOM}$  para as configurações de agrupamentos dos níveis 97 ao 114 do dendrograma.

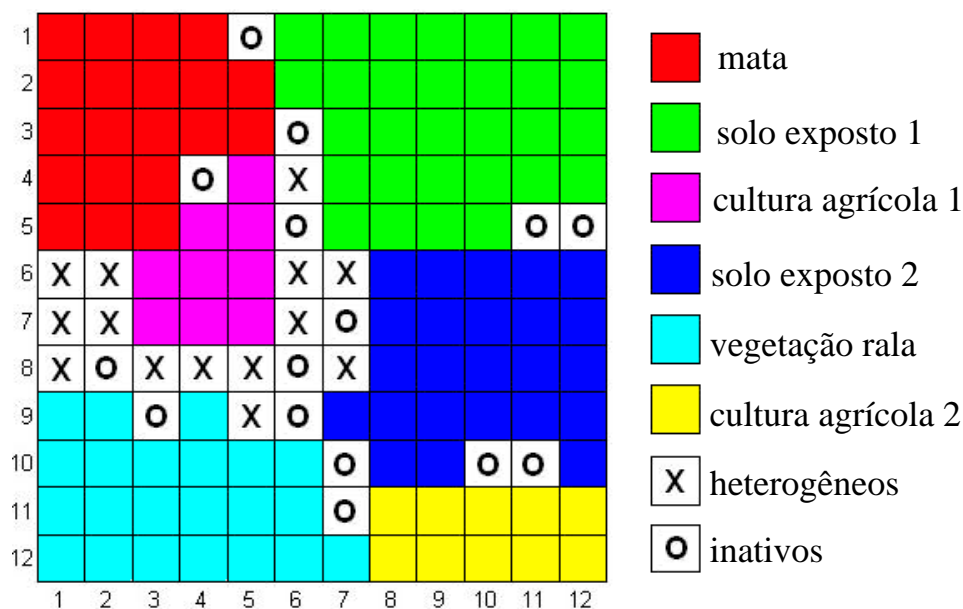


Fig. 6.8: SOM da imagem da figura 6.3 classificado.

fossem utilizadas para decidir sobre a fusão de agrupamentos, certamente classes distintas de protótipos poderiam ser erroneamente agrupadas. Por exemplo, os protótipos do SOM correspondentes às classes mata e cultura agrícola 1 são espectralmente muito similares. Como pode ser visto no MCI (figura 6.4), os protótipos dessas duas classes estão muito próximos uns dos outros no canto superior esquerdo do SOM. Essas duas classes de protótipos são corretamente discriminadas pelo método proposto em virtude dos índices  $IFE_{ij}$  e  $ICE_{ij}$  utilizados no critério de fusão, os quais incorporam informação espacial das classes na imagem. Esses índices apresentaram valores altos para as duas classes de protótipos, evitando, dessa forma, uma fusão prematura dos respectivos conjuntos de protótipos. De fato, como pode ser observado na imagem original (figura 6.3), as áreas correspondentes às classes de mata e cultura agrícola 1, situadas, respectivamente, nos cantos inferior esquerdo e canto superior direito da imagem, não têm fronteira em comum e são espacialmente compactas.

Procurando salientar a importância do uso de informações espaciais no critério de fusão do método hierárquico aplicado, o apêndice A.8 mostra o dendrograma resultante da aplicação do método hierárquico baseado apenas nas distâncias multiespectrais entre os agrupamentos de protótipos do SOM, sem considerar, portanto, informações de cunho espacial (como é feito pelo método hierárquico proposto aqui). No dendrograma mostrado, são apontadas fusões entre classes distintas de protótipos do SOM que comprometem uma separação correta das classes de informação da imagem.

Vale ressaltar ainda que, a segmentação adequada do SOM ocorre também em função do processo de filtragem de protótipos aplicado sobre ele após o seu treinamento. Os protótipos inativos e heterogêneos, que foram filtrados, agem como “bordas” na grade de neurônios do SOM contribuindo para a separação dos agrupamentos. Retomando o SOM classificado (figura 6.8), pode-se observar, por exemplo, que os protótipos da classe vegetação rala (na cor azul claro) se encontram quase que totalmente separados das outras classes em função apenas dos protótipos que foram filtrados (inativos e heterogêneos).

A figura 6.9 mostra o resultado da classificação da imagem usando todos os protótipos do SOM rotulado, exceto os protótipos inativos (as seis classes são representadas com as mesmas cores utilizadas na figura 6.8). De acordo com a última etapa do método de classificação

proposto, 19717 pixels da imagem associados aos protótipos heterogêneos foram reclassificados copiando o rótulo do pixel vizinho com a menor distância espectral.

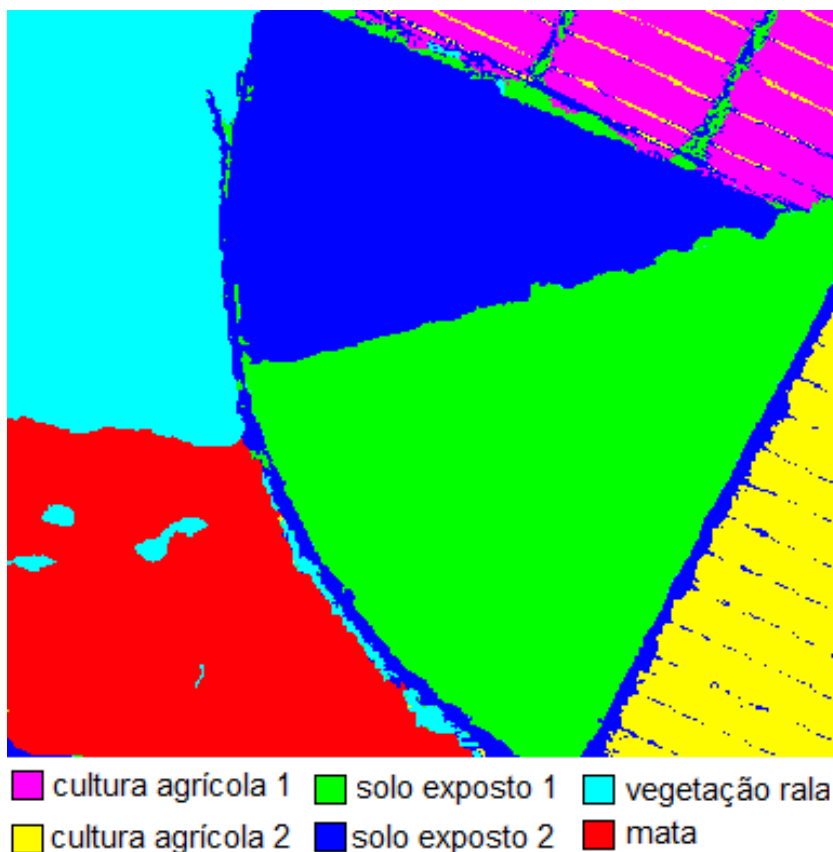


Fig. 6.9: Imagem da figura 6.3 classificada pelo método proposto.

Analisando visualmente a imagem classificada pelo método proposto (figura 6.9) e a imagem original (figura 6.3), pode-se notar que as classes de cobertura terrestre foram categorizadas adequadamente. Contudo, da mesma forma que nos experimentos realizados com os métodos de classificação propostos anteriormente, a imagem teste utilizada neste capítulo também foi classificada de maneira supervisionada através de uma rede neural MLP e o resultado da classificação foi usado como referência para avaliar a qualidade da classificação obtida pelo método proposto.

A tabela 6.1 mostra os valores da matriz de confusão para a classificação produzida pela metodologia proposta neste capítulo. O valor obtido para o índice de concordância Kappa foi 0.88, que de acordo com a tabela 2.5, qualifica a classificação produzida pelo método proposto neste capítulo como excelente.



<i>Classes</i>	vegetação	solo 1	solo 2	mata	cultura 2	cultura 1	<i>Total</i>
vegetação	<b>95.76</b>	0.54	0.15	0.92	0.03	0.16	15.81
solo 1	1.58	<b>89.05</b>	4.32	0.15	1.41	7.70	21.76
solo 2	0.19	3.59	<b>94.80</b>	0.01	0.02	0.80	28.91
mata	0.08	0.02	0.00	<b>83.44</b>	0.18	1.97	14.02
cultura 2	0.13	3.59	0.03	5.39	<b>93.33</b>	9.30	10.51
cultura 1	2.27	3.22	0.69	10.10	5.03	<b>80.06</b>	8.99
<i>Total</i>	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Tab. 6.1: Matriz de confusão das classes para a imagem classificada pela metodologia de classificação proposta. As linhas da matriz mostram os erros de comissão (porcentagem de pixels que pertencem a outra classe, mas que foi classificado como a classe de interesse) e as colunas mostram os erros de omissão (porcentagem de pixels que pertencem a classe da referência/verdade, mas que foram classificados incorretamente).

A tabela 6.2 mostra o tempo de processamento consumido neste experimento por cada uma das etapas da metodologia de classificação proposta. Os experimentos foram realizados em uma máquina com processador AMD Athlon XP 2600+ 1,91GHZ e 1GB de RAM.

Etapas do método proposto	Tempo consumido (em segundos)
Amostragem	7
Treinamento do SOM	56
Filtragem de protótipos	4
Análise Hierárquica	96
Avaliação dos agrupamentos	42
Classificação da imagem	26
Reclassificação	5
<i>Total</i>	236

Tab. 6.2: Tempo consumido pelas etapas do método de classificação proposto.

O tempo total consumido pelo método proposto na classificação da imagem teste foi de 236 segundos. Observando os tempos de processamento de cada uma das etapas, nota-se que, além do treinamento do SOM, as etapas de análise hierárquica e de avaliação de agrupamentos são aquelas que apresentam um maior custo computacional. Isso ocorre devido à quantidade de diferentes partições a serem analisadas, cada uma delas associada a um nível do dendrograma de protótipos do SOM. Particularmente, na etapa de análise hierárquica, o tempo de processa-

mento é alto em razão da necessidade de cálculo dos índices de informação espacial. Conforme descrito na seção 6.2.2, esses índices são calculados usando a imagem classificada com base nas classes de protótipos do SOM em cada nível do dendrograma.

### 6.3.1 Comparação com o algoritmo K-médias padrão e *janelado*

A figura 6.10 mostra o resultado da classificação da imagem IKONOS utilizada nos testes pelo algoritmo K-médias padrão. Analisando visualmente a imagem classificada, é possível notar claramente que o algoritmo estatístico não classificou adequadamente todos os padrões de cobertura terrestre existentes na cena. O algoritmo apresentou erros na classificação, principalmente para aquelas classes que possuem alta similaridade espectral, como, por exemplo, os pares de classes cultura agrícola 1 e mata, e também, solo exposto 2 e cultura agrícola 2. Comparando a imagem resultante do algoritmo K-médias padrão com a do método proposto neste capítulo (figura 6.9), percebe-se que há uma diferença significativa no reconhecimento dos padrões. A imagem classificada pelo algoritmo K-médias apresenta uma aparência salpicada e poucas regiões homogêneas. Isso ocorre porque, diferentemente do método proposto, o algoritmo realiza a classificação da imagem seguindo uma abordagem pixel a pixel, baseando-se somente nas informações de distância espectral entre os pixels da imagem. Sendo assim, o algoritmo estatístico é capaz de discriminar padrões espectrais (tonalidades de cores) presentes na imagem, mas não consegue capturar padrões de cunho espacial (textura), os quais são fundamentais para caracterizar e diferenciar as 6 classes de cobertura terrestre presentes na imagem.

Procurando proporcionar ao algoritmo K-médias a capacidade de capturar a heterogeneidade espectral de cada classe da imagem, o mesmo foi modificado para processar amostras na forma de janelas de pixels em vez de pixels isolados. A modificação proposta para o K-médias segue, portanto, os mesmos passos do algoritmo K-médias padrão descrito na seção 2.5.2.1. No entanto, ao invés de processar os dados da imagem na forma de pixels individuais, os dados (amostras) são processados na forma de janelas de pixels, as quais são coletadas do mesmo modo que nos métodos de classificação propostos nesse trabalho (figura 4.9). Essa versão modificada do algoritmo estatístico foi denominada aqui de algoritmo K-médias "janelado".

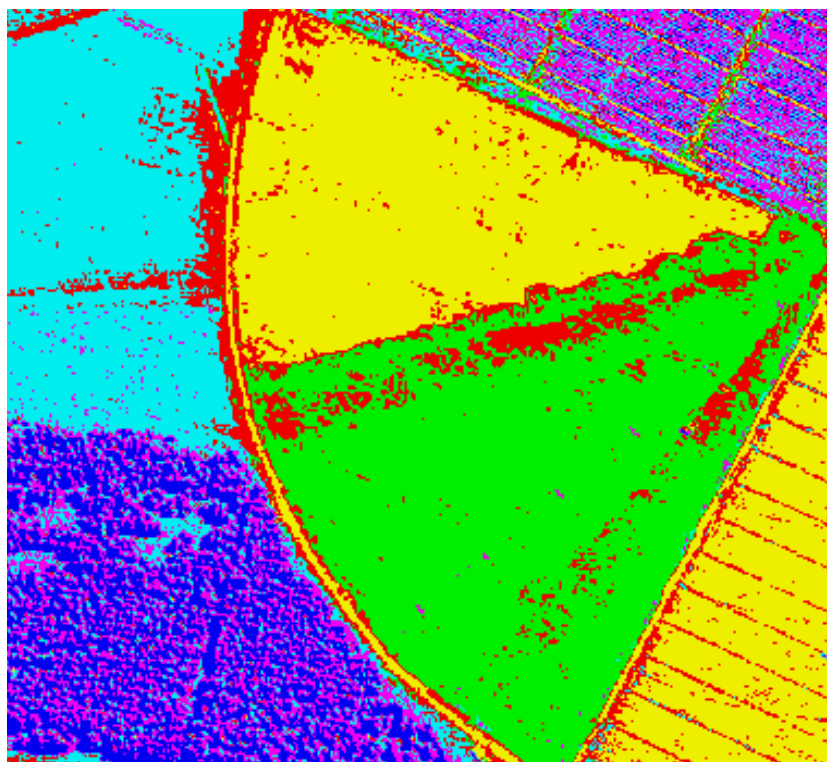


Fig. 6.10: *Imagem da figura 6.3 classificada pelo algoritmo K-médias padrão.*

A figura 6.11 mostra o resultado da classificação da imagem IKONOS pela versão janelada do algoritmo K-médias. O algoritmo foi executado empregando as mesmas janelas amostrais que foram utilizadas pela metodologia de classificação proposta neste capítulo.

Comparando o resultado da classificação do algoritmo K-médias padrão com o resultado da sua versão janelada, é possível observar um melhoramento na homogeneidade das regiões classificadas. No entanto, embora o algoritmo K-médias "janelado" consiga incorporar informações de vizinhança dos pixels através das janelas amostrais, o mesmo ainda não foi capaz de discriminar classes de cobertura terrestre que estão muito próximas umas das outras no espaço multiespectral. O algoritmo não classificou corretamente as regiões da imagem correspondentes às classes mata e cultura agrícola 1, apresentando um alto grau de confusão na discriminação dos dois padrões. O mesmo acontece para as classes cultura agrícola 2 e solo exposto 2. Além disso, em virtude do uso de janelas de pixels como dados de entrada para o algoritmo, algumas feições lineares, particularmente para as duas classes de cultura agrícola, não foram detectadas. Como pode ser visto na imagem resultante do método de classificação proposto (figura 6.9), tais

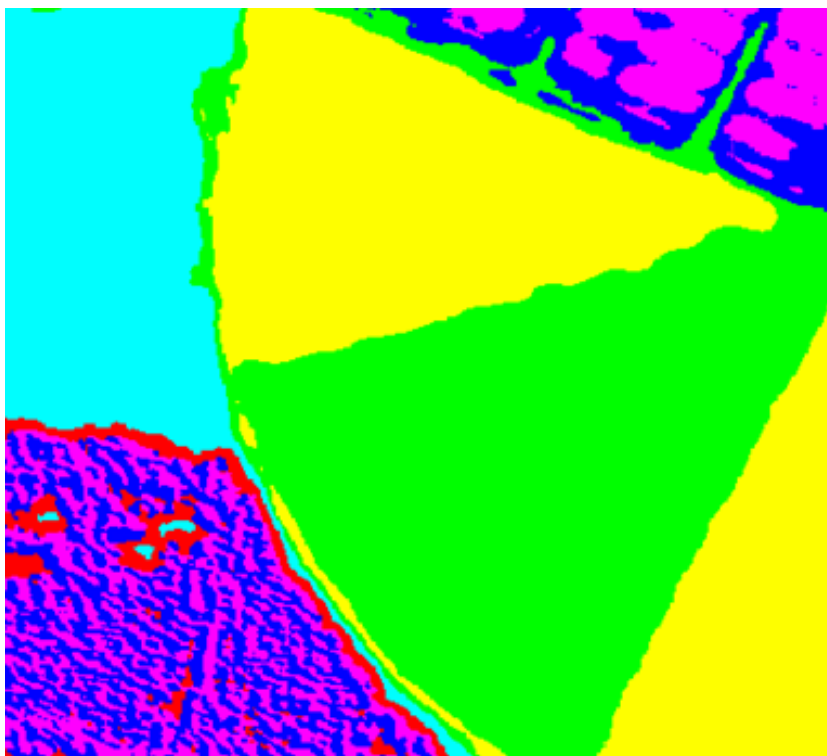


Fig. 6.11: Imagem da figura 6.3 classificada pelo algoritmo K-médias "janelado".

feições lineares são detectadas corretamente (em virtude da etapa de reclassificação de pixels situados em regiões de fronteira entre classes de cobertura terrestre distintas).

Os tempos consumidos pelos algoritmos K-médias padrão e "janelado" na classificação da imagem teste foram, respectivamente, 151 e 98 segundos, rodando sob o software MATLAB na mesma máquina utilizada para executar o método de classificação proposto. As duas versões do algoritmo K-médias foram executadas utilizando cinco inicializações diferentes para os centróides de cada classe e um número máximo de iterações igual a 100. As classificações consideradas para a análise foram aquelas que apresentaram os valores mais baixos para a soma dos quadrados das distâncias dada pela equação 2.4.

Embora as duas versões do algoritmo K-médias tenham apresentado tempos de processamento inferiores ao da metodologia proposta, é importante observar, conforme mencionado nos capítulos anteriores, que os métodos possuem princípios e características bastante diferenciadas, que devem ser consideradas na comparação de suas complexidades de cálculo. Além disso, apesar do algoritmo K-médias apresentar um custo de processamento menor que o do

método de classificação proposto neste capítulo, o mesmo apresenta um nível de exatidão inferior no resultado da classificação. Assim como foi feito na seção 6.3, a imagem resultante da classificação supervisionada foi adotada como referência para avaliar a qualidade das classificações produzidas pelo K-médias. Os índices de concordância Kappa para as classificações resultantes dos algoritmos K-médias padrão e "janelado" apresentaram valores iguais a 0.63 e 0.55, respectivamente. Tais valores são inferiores ao valor do índice de concordância Kappa apresentado pela classificação realizada através do método de classificação proposto.

### 6.3.2 Comparação com os métodos anteriores

Considerando, ainda, a imagem teste da figura 6.3, a presente seção apresenta uma comparação entre os resultados das classificações produzidas pelo método proposto neste capítulo e pelos métodos apresentados nos capítulos 4 e 5 desse trabalho.

Uma vez que os métodos propostos nos capítulos 4 e 5 baseiam-se no uso da *U-matrix*, apresentando resultados de classificação bastante semelhantes, a figura 6.12 mostra a classificação gerada apenas pelo método proposto no capítulo 5 (a classificação produzida pelo método do capítulo 4 apresentou poucas diferenças, quase imperceptíveis visualmente).

Analisando a imagem da figura 6.12 e comparando-a com a classificação resultante do método proposto nesse capítulo (figura 6.9), pode-se notar que as áreas da cena correspondentes às classes mata (no canto inferior esquerdo) e cultura agrícola 1 (no canto superior direito) receberam o mesmo rótulo (cor vermelha), ou seja, não foram reconhecidas como classes distintas. Algumas feições lineares também não foram preservadas no resultado da classificação. Por outro lado, as outras regiões da imagem correspondentes às classes de cobertura terrestre vegetação rala (em azul claro), solo exposto 1 (em verde), cultura agrícola 2 (em amarelo) e solo exposto 2 (em azul), foram classificadas adequadamente, de modo bastante similar à classificação produzida pelo método proposto nesse capítulo.

Para entender porque os métodos propostos nos capítulos anteriores não distinguiram corretamente as classes mata e cultura agrícola 1 (principal diferença entre as classificações obtidas), é importante lembrar que os mesmos não incorporam, no processo de classificação da imagem, informações de natureza geográfica (ou espacial) sobre os padrões, como é feito pelo método

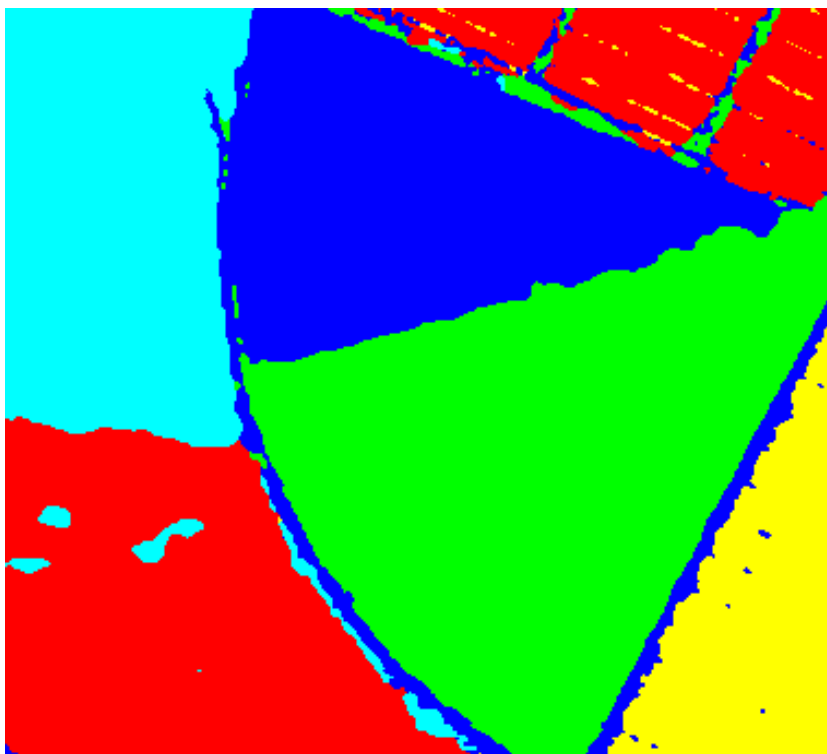


Fig. 6.12: Imagem da figura 6.3 classificada pelo método proposto no capítulo 5.

proposto neste capítulo, o qual utiliza os índices de fronteira e de compacidade espacial (descritos nas equações 6.1 e 6.2) para discriminar as classes das imagens.

Uma vez que se baseiam na utilização da *U-matrix*, os métodos propostos nos capítulos anteriores, consideram apenas informações de distância no espaço multiespectral das imagens para detectar os diferentes agrupamentos de dados existentes nas mesmas. Mesmo incorporando informações de vizinhança (através do uso de janelas de pixels), os métodos não detectaram dois agrupamentos de dados distintos associados às classes mata e cultura agrícola 1. Em outras palavras, na classificação da imagem da figura 6.3 pelos métodos propostos nos capítulos 4 e 5, os protótipos representantes das classes mata e cultura agrícola 1 foram considerados como pertencentes a um mesmo agrupamento.

Para entender mais claramente porque isso acontece, basta observar a figura 6.13, que mostra a *U-matrix* do SOM treinado com os dados da imagem teste (para facilitar a visualização, a *U-matrix* é mostrada na forma de superfície e já com a operação de suavização aplicada). Através da imagem 3D da matriz de distâncias do SOM, é possível notar seis depressões ou "vales"

que correspondem às seis classes de cobertura terrestre presentes na imagem. No entanto, dois desses "vales", os quais correspondem justamente às classes mata e cultura agrícola 1, além de não serem tão profundos ou tão bem definidos como os demais, não possuem uma borda de separação entre eles. Isso significa que os conjuntos de protótipos do SOM, que representam as duas classes, possuem uma similaridade considerável no espaço multiespectral da imagem. Portanto, a utilização apenas da informação de distância entres os protótipos do SOM (dada pela *U-matrix*) não é suficiente para separar as duas classes. Como o método proposto neste capítulo, incorpora no processo de discriminação dos padrões, informações de cunho geográfico (ou espacial), além das informações de distância, o mesmo conseguiu discriminar corretamente as classes mata e cultura agrícola 1. Observando a imagem teste (figura 6.3), as duas classes classes apresentam uma alta compacidade e baixo índice de fronteira espacial, fatores que penalizam uma junção das mesmas de acordo com o critério de fusão do método proposto nesse capítulo.

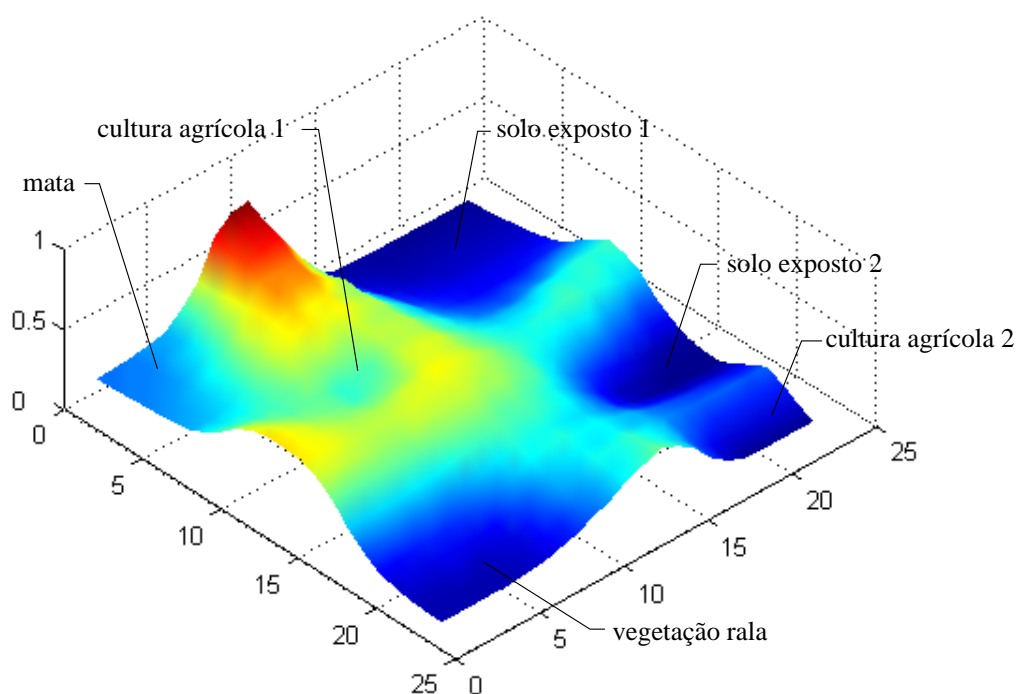


Fig. 6.13: *U-matrix* na forma de superfície do SOM da figura 6.4.

Uma outra forma de entender as diferenças nas classificações produzidas pelos métodos propostos, é interpretar a *U-matrix* como um dendrograma de um método hierárquico de Li-

gações Simples (conforme explicado no início da seção 5.5), onde apenas as distâncias entre os neurônios na grade do SOM são calculadas na matriz de dissimilaridades. Sendo assim, considerando que os neurônios representantes das classes mata e cultura agrícola 1 estão próximos uns dos outros no espaço multiespectral da imagem, apresentando uma distância pequena entre eles, ocorre uma fusão entre as duas classes. Já no método proposto nesse capítulo, a matriz de dissimilaridades do método hierárquico também é calculada considerando apenas os neurônios da grade do SOM, porém, além dos valores de distâncias, informações adicionais de compacidade e de fronteira espacial dos pixels da imagem (associados aos neurônios) são incorporadas no cálculo de dissimilaridades. O uso de mais informações no seu critério de fusão faz com que o método hierárquico proposto tenha um comportamento diferente do método hierárquico de Ligações Simples.

Utilizando os resultados da classificação supervisionada da imagem teste como referência para avaliar a qualidade das classificações produzidas, o índice de concordância Kappa apresentou valores iguais a 0.61 e 0.62 para as classificações dos métodos dos capítulos 4 e 5, respectivamente. Embora os valores do índice Kappa tenham sido ligeiramente inferiores ao valor do mesmo índice para a classificação resultante do algoritmo K-médias padrão, vale observar que as classificações produzidas pelos métodos propostos nos capítulos anteriores desse trabalho (figura 6.12) apresentam, visualmente, uma categorização mais consistente e menos confusa dos padrões da imagem do que aquela produzida pelo algoritmo K-médias padrão (figura 6.10). Embora a classificação produzida pelos métodos anteriores não tenha discriminado as classes mata e cultura agrícola 1, todas as regiões da imagem apresentam uma classificação mais homogênea para os diferentes padrões. Já a classificação resultante do algoritmo K-médias, além de não discriminar corretamente algumas classes da imagem, apresenta uma classificação mais salpicada e menos homogênea.

Em relação aos tempos de processamento, os métodos propostos nos capítulos 4 e 5 consumiram, respectivamente, 274 e 165 segundos para realizar a classificação da imagem da figura 6.3. Conforme mostrado na tabela 6.2, o tempo total consumido pelo método proposto nesse capítulo para classificar a mesma imagem foi 236 segundos. Como esperado, a metodologia de classificação do capítulo 4 foi a que apresentou o maior tempo de processamento para realizar



a classificação da imagem, uma vez que necessitou treinar e interpretar um número maior de SOMs (total de oito) para detectar as classes da imagem. Já o método proposto no capítulo 5 foi aquele que demandou o menor tempo para realizar a classificação. Com o auxílio dos índices de validação de agrupamentos, tal método conseguiu discriminar as classes da imagem analisando diferentes partições do SOM (inicial) apresentado na figura 6.4. Embora o método apresentado nesse capítulo também tenha utilizado apenas um único SOM para detectar os agrupamentos de dados da imagem, o critério de fusão do método hierárquico utilizado demanda um maior tempo de processamento em razão dos cálculos dos índices de fronteira e de compactidade espacial. Além disso, o método analisa um número muito maior de partições do SOM (nos diferentes níveis do dendrograma).

Outros resultados experimentais do método de classificação proposto nesse capítulo podem ser vistos em Gonçalves *et al.* (2009).

### 6.3.3 Exemplos de reuso

Procurando verificar a possibilidade de reutilização de SOMs que já foram empregados para classificar uma determinada área (ou cena) da superfície terrestre, para categorizar outras contendo padrões semelhantes, a presente seção mostra dois testes experimentais com as imagens apresentadas nas figuras 6.14 e 6.16.

A imagem da figura 6.14 corresponde a uma área geograficamente próxima a da imagem da figura 6.3 (com a mesma data de aquisição), contendo padrões de cobertura terrestres semelhantes, como os de cultura agrícola 1, vegetação rala, solo exposto 1 e solo exposto 2. A cena não apresenta padrões de cultura agrícola 2 e mata, presentes na imagem anterior. Ela apresenta, ainda, no canto inferior direito, uma área contendo uma vegetação mais escura, que não aparece na cena da imagem da figura 6.3.

Para classificar a imagem da figura 6.14, o SOM da figura 6.4, treinado com os padrões da imagem da figura 6.3 e rotulado pelo método de classificação proposto nesse capítulo (figura 6.8), foi reutilizado. Vale observar que os protótipos heterogêneos também foram utilizados e a etapa de reclassificação também foi executada.

A figura 6.15 mostra a imagem classificada. As áreas associadas aos quatro padrões de



Fig. 6.14: Imagem de uma área da superfície terrestre geograficamente próxima a da figura 6.3.

cobertura terrestre (cultura agrícola 1, vegetação rala, solo exposto 1 e solo exposto 2), que também aparecem na cena da figura 6.3, foram classificadas adequadamente. No entanto, aquela área no canto inferior direito da imagem, que apresenta um padrão de vegetação mais escura (que não existe na imagem anterior), foi classificada como cultura agrícola 1. Isso era esperado, tendo em vista que não há protótipos do SOM da imagem da figura 6.3 representantes dessa nova classe. Tal classe foi rotulada como cultura agrícola 1, por apresentar uma maior similaridade no espaço multiespectral com a mesma.

Procurando estabelecer uma avaliação mais criteriosa da classificação resultante, a imagem da figura 6.14 foi classificada de maneira supervisionada considerando 5 classes (dentre elas, o novo padrão de vegetação mais escura). Da mesma forma que nos experimentos anteriores, uma rede neural MLP com o algoritmo de treinamento *backpropagation* foi empregada para realizar a classificação supervisionada da cena. Aplicando o índice de concordância Kappa

entre a imagem da figura 6.15 e a imagem resultante da classificação supervisionada, o índice atingiu o valor 0.76. Esse valor, de acordo com a tabela 2.5, qualifica a classificação apresentada na figura 6.15 como muito boa.

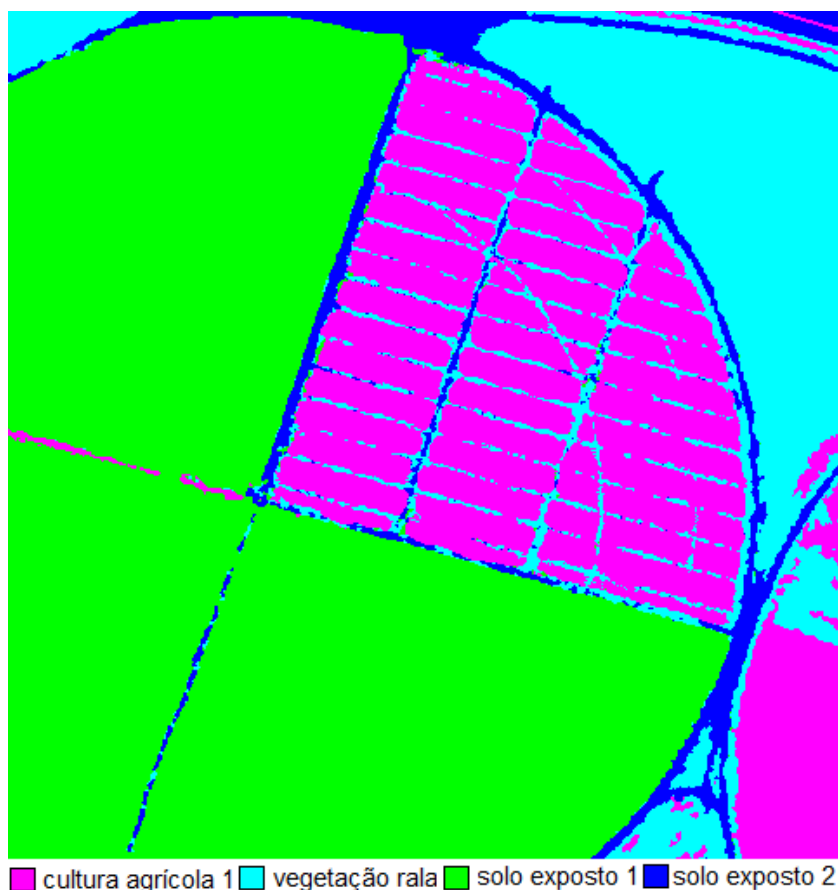


Fig. 6.15: Imagem da figura 6.14 classificada usando os protótipos do SOM rotulado da figura 6.8.

A figura 6.16 mostra a outra imagem utilizada para testar o potencial de reuso de SOMs já treinados e rotulados. A imagem mostra uma área do Arquipélago de Anavilhanas, englobando o município de Novo Airão, no estado do Amazonas. A área é geograficamente próxima àquela da imagem da figura 4.13, apresentando três padrões de cobertura terrestre semelhantes: vegetação, área urbana e água. As aquisições das cenas foram feitas com cinco dias de diferença uma da outra.

Para classificar a nova cena, o SOM da figura 4.15, treinado com os padrões da imagem da figura 4.13 e rotulado pelo método de classificação proposto no capítulo 5 (figura 5.5, partição



Fig. 6.16: Imagem de uma área da superfície terrestre geograficamente próxima a da figura 4.13.

$k = 59$ ), foi reutilizado.

A figura 6.17 mostra a imagem classificada. As áreas da imagem correspondentes aos três padrões de cobertura terrestre foram classificadas de maneira adequada. Algumas áreas de desmatamento (na cor rosa) foram classificadas como área urbana. Entretanto, isso também ocorreu na classificação da imagem da figura 4.13, tendo em vista que esses dois padrões de cobertura terrestre apresentam alta similaridade no espaço multiespectral.

Da mesma forma que nos experimentos anteriores, a imagem da figura 6.16 foi classificada de maneira supervisionada. O valor do índice de concordância Kappa entre o resultado dessa classificação com aquela produzida através do reuso do SOM foi 0.81. De acordo com a tabela 2.5, esse valor qualifica a classificação apresentada na figura 6.17 como excelente.

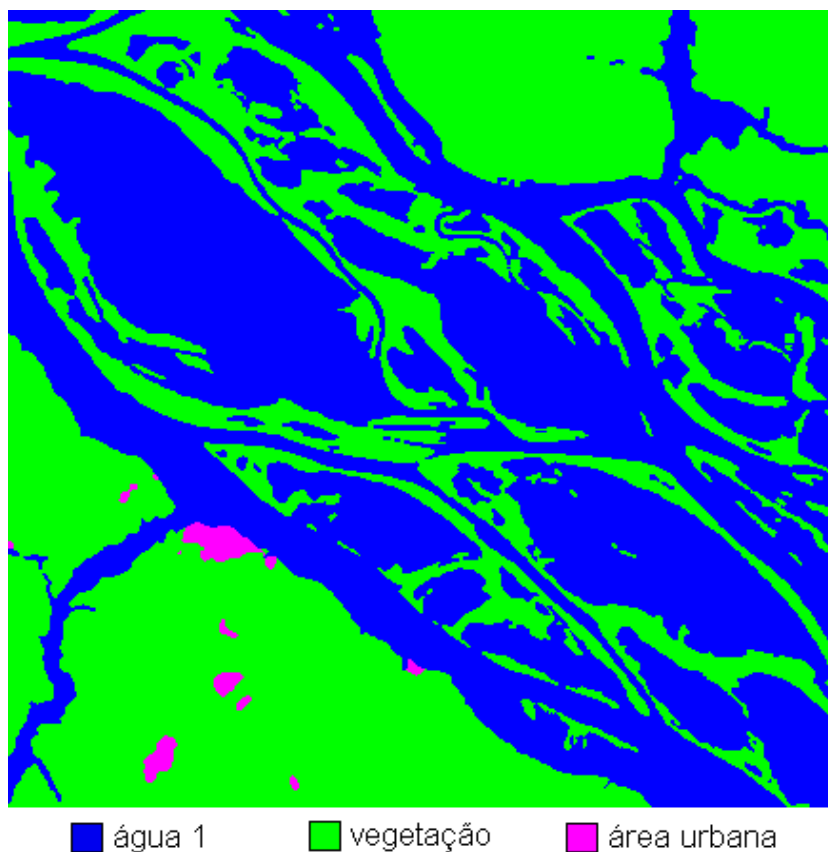


Fig. 6.17: Imagem da figura 6.16 classificada usando os protótipos do SOM rotulado da figura 4.15.

## 6.4 Considerações finais

Considerando que classificadores diferentes apresentam capacidades complementares, na metodologia de classificação de imagens proposta neste capítulo, o uso combinado do SOM e de métodos de agrupamentos hierárquicos não explora apenas as vantagens de ambos os métodos para realizar o processo de classificação das imagens, mas também permite que um método auxilie a contornar as desvantagens e limitações do outro. A representação dos dados da imagem original por meio de um conjunto reduzido de protótipos do SOM, torna viável o uso de métodos de agrupamentos hierárquicos, os quais são considerados impraticáveis para aplicações que envolvem grandes volumes de dados, como é o caso das imagens de sensoriamento remoto. Além disso, a propriedade de ordenação topológica do SOM restringe possíveis fusões entre os protótipos do mapa neural e, portanto, ajuda a reduzir a complexidade computacional do

método hierárquico aplicado. Por outro lado, o método hierárquico segmenta automaticamente o SOM treinado, sem a necessidade de intervenção do usuário, e ainda, organiza hierarquicamente os agrupamentos de protótipos do SOM que representam os padrões originais da imagem. A utilização de um método de agrupamento hierárquico permite que o usuário compreenda, em diferentes níveis de agrupamentos, as relações existentes entre os padrões de cobertura terrestre presentes na imagem. Isto pode ser muito útil em situações onde a estrutura da informação presente na imagem não é muito clara.

Um diferencial importante do método de classificação proposto está na utilização de um mecanismo de fusão eficiente, empregado pelo método de agrupamento hierárquico na etapa de segmentação do SOM. O critério de fusão adotado, além de respeitar as relações de vizinhança do SOM, não considera apenas a informação de distância entre os agrupamentos como é feito em métodos de agrupamentos tradicionais, mas incorpora também informações espaciais sobre os padrões da imagem que são representados pelos conjuntos de protótipos do SOM. Resultados experimentais mostram que essa integração de informações espectrais e espaciais é imprescindível para a discriminação de classes de cobertura terrestres que apresentam alta similaridade no espaço multiespectral da imagem. Métodos convencionais, como o algoritmo K-médias padrão e uma versão modificada do mesmo (apresentada neste capítulo), não foram capazes de classificar corretamente todas as classes de cobertura presentes numa imagem IKONOS, baseando-se apenas nas informações de distâncias entre os agrupamentos. O mesmo acontece com os métodos de classificação propostos nos capítulos anteriores, os quais se baseiam apenas nas informações de distâncias entre os protótipos do SOM através da *U-matrix*.

Na análise do SOM treinado, o método proposto emprega ainda um procedimento de filtragem de protótipos que colabora significativamente no processo de separação das classes através do mapeamento da imagem produzido pelo SOM. Protótipos inativos, e aqueles que apresentam alto grau de heterogeneidade espectral, são descartados durante a etapa de segmentação do SOM. Protótipos inativos e heterogêneos funcionam como bordas na grade de neurônios do SOM, contribuindo para o particionamento do mapa treinado.

É importante salientar ainda que, outro ponto importante que viabiliza a aplicação do método de classificação proposto é a utilização da estratégia de modificação dos cálculos dos

índices de validação proposta no capítulo anterior. A avaliação das diferentes configurações de agrupamentos de protótipos do SOM, em cada nível do dendrograma gerado pelo método hierárquico, somente é viável devido à baixa complexidade computacional exigida pelo índice de validação  $\mathcal{V}_{CDbw}^{SOM}$ .

Os resultados experimentais apresentados nesse capítulo mostram, também, a possibilidade de reutilização de SOMs para classificar outras imagens de satélite para as quais eles não foram treinados. Os resultados mostram que o reuso de SOMs para diferentes imagens que apresentam padrões de cobertura terrestres semelhantes é viável. Tal resultado apresenta uma importância significativa, tendo em vista a necessidade de exploração cada vez maior da enorme quantidade de imagens de sensoriamento remoto que são coletadas diariamente.

# Capítulo 7

## Conclusões

### 7.1 Considerações finais sobre os métodos propostos

Nas últimas décadas, um número cada vez maior de países e consórcios internacionais está investindo em programas espaciais direcionados ao lançamento e operação de satélites de levantamento e monitoramento de recursos naturais. Notadamente, neste momento em que a comunidade mundial começa a encarar seriamente a necessidade de acompanhar e estudar as conseqüências das mudanças globais que vêm ocorrendo na superfície do planeta, o sensoriamento remoto orbital se coloca como uma ferramenta estratégica para o futuro. As imagens coletadas por sensores remotos em nível orbital passaram a representar uma das poucas formas viáveis de monitoramento ambiental em escalas locais e globais, devido à rapidez, eficiência e visão sinóptica que as caracterizam. Em razão disso, sensores a bordo de satélites são aperfeiçoados rapidamente, procurando atender demandas cada vez mais sofisticadas por maiores resoluções dos dados gerados. Uma das principais conseqüências disso é o aumento significativo na quantidade de imagens geradas continuamente por sensores a bordo de inúmeras plataformas de satélite com resoluções espaciais, temporais e espectrais sem precedentes.

No entanto, enquanto a área de sensoriamento remoto apresentou nos últimos anos um enorme progresso em relação às tecnologias de captação de imagens em nível orbital, os métodos de reconhecimento automático de padrões de imagens de satélite não evoluíram no mesmo ritmo. É possível encontrar softwares e pacotes atuais voltados exclusivamente para o processamento



de imagens de sensoriamento remoto, com métodos, em sua maioria, desenvolvidos no final da década de 70. Diante disso, todo o avanço tecnológico que incide nos sistemas de imageamento fica sem sentido se não ocorrer simultaneamente a contrapartida em relação às técnicas de processamento das informações contidas nas imagens.

Diante desse cenário, verifica-se a necessidade de estudar e desenvolver novos métodos para a exploração automática e consistente das imagens adquiridas por sensores remotos. Considerando, sobretudo, o gigantesco volume de imagens disponíveis, conclui-se que uma exploração integral das mesmas requer que as estratégias de análise sejam as mais automatizadas possível, exigindo o mínimo de interação do usuário ou de um analista de imagens. E é por esta razão que as técnicas de análise de agrupamentos ou de classificação não-supervisionada exercem cada vez mais um papel especialmente importante nesta área de aplicação.

Embora seja possível detectar, na literatura de sensoriamento remoto algum progresso no desenvolvimento de novas metodologias de classificação de imagens de satélite, a grande maioria das inovações apresentadas se concentram em técnicas de classificação fundamentadas em paradigmas de aprendizado supervisionado. Mesmo diante da existência de um grande número de diferentes métodos de análise de agrupamentos na área de reconhecimento de padrões, dentre as poucas estratégias de classificação não-supervisionada que têm sido propostas na literatura, muitas delas baseiam-se somente em métodos particionais clássicos. Isso se confirma através dos softwares contemporâneos de análise de imagens de sensoriamento remoto, os quais, em sua maioria, ainda oferecem ao usuário métodos de classificação não-supervisionada tradicionais, como o K-médias e o ISODATA.

O presente trabalho buscou, portanto, contribuir com o progresso das técnicas de processamento de imagens de sensores remotos e com a necessidade de exploração em grande escala das mesmas, propondo novas metodologias de classificação não-supervisionada que particularmente exploram as características e propriedades do SOM.

A estratégia chave dos métodos de classificação propostos é realizar a análise de agrupamentos das imagens através do mapeamento produzido pelo SOM, ao invés de trabalhar diretamente com os padrões originais das cenas e, diferentemente de outras abordagens, nas quais o mapa de Kohonen é utilizado pelo usuário como ferramenta de auxílio visual para a detecção de agrupa-

mentos, nos métodos de classificação propostos, diferentes técnicas de segmentação automática foram aplicadas para realizar o particionamento dos neurônios do SOM treinado, não exigindo, portanto, que o usuário interaja diretamente com o mapa neural para descobrir as classes de cobertura terrestre existentes nas imagens.

Os experimentos realizados com diferentes imagens de satélite permitem concluir que os métodos de classificação propostos neste trabalho apresentam uma série de vantagens e potencialidades que os colocam como alternativas diferenciadas para a classificação não-supervisionada de imagens de sensoriamento remoto. Dentre elas, pode-se destacar:

- A não exigência de uma definição a priori do número de classes para realizar a classificação das imagens, o que não acontece na maioria dos métodos convencionais de classificação não-supervisionada, como o K-médias;
- Os métodos propostos não assumem nenhum modelo de distribuição de probabilidade ou qualquer suposição sobre a estrutura dos dados das imagens;
- O simples uso de janelas de pixels permite incorporar informações de contexto e textura sem nenhum cálculo explícito de medida. Essa abordagem contribui para a qualidade da classificação resultante, provendo um desempenho superior ao de abordagens de classificação pixel a pixel;
- A representação distribuída das classes por meio de grupos de protótipos do SOM proporciona aos métodos a potencialidade de descobrirem agrupamentos de dados com geometrias complexas e variadas. Métodos, como o K-médias, utilizam apenas um único protótipo (centróide) para representar cada classe e, por isso, são capazes de detectar adequadamente apenas agrupamentos que tenham formatos hiperesféricos;
- A análise de agrupamentos da imagem por meio de um conjunto reduzido de protótipos do SOM contribui para a redução de ruídos ("outliers") existentes nos dados originais. Os protótipos do SOM funcionam como médias locais dos dados, reduzindo desse modo a influência de padrões de pouca relevância;

- Diferentemente de métodos convencionais, como o K-médias e o ISODATA, que são bastante sensíveis a vários parâmetros dependentes do usuário (podendo gerar classificações diferentes quando são feitas várias simulações para um mesmo conjunto de dados), os métodos propostos nesse trabalho possuem apenas dois parâmetros que devem ser definidos pelo usuário: o tamanho das amostras e do SOM. Contudo, com exceção do tempo de processamento, o desempenho dos métodos pode não ser significativamente afetado se amostras e mapas com tamanhos maiores que os necessários forem utilizados. Estratégias para tratar de maneira adequada protótipos inativos e aqueles que incorporam mais de uma classe (chamados de heterogêneos) foram propostas. Além disso, conforme mostrado no capítulo 6, protótipos inativos e heterogêneos podem funcionar como bordas de separação entre as diferentes classes da imagem no SOM, fator que contribuiu para a detecção dos agrupamentos;
- O número de classes da imagem é determinado automaticamente, sem a necessidade de intervenção do usuário;
- Os métodos apresentados nos capítulos 2 e 4 permitem visualizar de maneira hierárquica a estrutura das classes presentes nas imagens, possibilitando que o usuário compreenda, em diferentes níveis de agrupamentos, as relações existentes entre os padrões de cobertura terrestre. Isso pode ser útil em aplicações nas quais a estrutura das informações presentes nas imagens não é claramente conhecida. Além disso, caso o usuário deseje, a imagem pode ser classificada considerando diferentes configurações de agrupamentos ou quantidades de classes;
- Nos métodos propostos, o usuário também tem a possibilidade de visualizar através do MCI (imagem da grade de neurônios do SOM) as relações topológicas e as distribuições de probabilidade das classes existentes na imagem. Adicionalmente, as árvores de mapas e/ou dendrogramas gerados podem ser melhor visualizados e compreendidos com o uso dos MCIs;
- Diferentemente dos métodos de agrupamentos tradicionais que utilizam apenas a informação de distância para decidir sobre a fusão de agrupamentos, o método de classificação

proposto no capítulo 6 incorpora, em seu critério de fusão de agrupamentos, informações de cunho espacial sobre os padrões da imagem. Isso dá ao método a vantagem de discriminar classes de cobertura terrestre que apresentam alta similaridade no espaço multiespectral;

- O método de classificação proposto no capítulo 4 também emprega uma abordagem particular para classificar pixels situados em regiões de transição entre classes. Este procedimento contribui para aumentar a precisão das classificações resultantes;
- Índices de validação apropriados são utilizados para avaliar as diferentes configurações de agrupamentos das imagens que são geradas pelos métodos propostos. A utilização de índices de validação adequados diminui o risco de interpretações incorretas sobre a escolha da melhor partição dos dados. Conforme apresentado no capítulo 5, na literatura é possível encontrar métodos de classificação não-supervisionada que empregam índices de validação que apresentam baixa complexidade computacional, mas que não são totalmente adequados para avaliar agrupamentos de dados de imagens de sensoriamento remoto.

É importante ressaltar ainda (conforme descrito no capítulo 2), que os métodos de classificação propostos nesse trabalho são indicados para imagens multiespectrais de sensoriamento remoto de média e alta resoluções espaciais, contendo quaisquer tipos de padrões de cobertura terrestre, exceto padrões de área urbana para o caso de imagens de alta resolução espacial.

Outro aspecto importante que também deve ser observado é que, embora tenham a mesma finalidade, os métodos de classificação propostos apresentam algumas características que os diferenciam um do outro e que devem ser consideradas pelo usuário na escolha do método mais apropriado para uma determinada tarefa ou situação.

Os métodos apresentados nos capítulos 4 e 6 permitem que o usuário visualize a estrutura hierárquica das classes de cobertura terrestre detectadas pelos mesmos, através de uma árvore de SOMs ou de um dendrograma simplificado dos dados da imagem. Conforme já mencionado anteriormente, essas visualizações podem ser bastante proveitosas para o caso de imagens nas quais o relacionamento existente entre as classes seja mais complexo ou mais difícil de ser compreendido. A árvore de SOMs que é gerada pelo método proposto no capítulo 4, permite

que o usuário tenha uma visualização mais concisa do relacionamento entre as classes, enquanto que o dendrograma gerado pelo método proposto no capítulo 6 possibilita uma visualização mais detalhada da estrutura dos dados da imagem, uma vez que o mesmo exibe o relacionamento existente entre cada um dos protótipos do SOM.

Apesar do método de classificação apresentado no capítulo 5 não proporcionar ao usuário a possibilidade de visualizar a estrutura hierárquica das classes das imagens, os resultados experimentais mostram que o mesmo apresenta um custo de processamento inferior aos dos outros métodos propostos nesse trabalho. Portanto, o método proposto no capítulo 5 pode ser mais adequado para situações em que se deseja obter resultados mais rápidos e, dentro das quais, não há a necessidade de ferramentas adicionais para compreender o relacionamento existente entre os padrões das imagens.

Dentre os três métodos de classificação propostos nesse trabalho, o método do capítulo 6 apresenta, ainda, uma maior sofisticação para interpretar imagens que possuam classes de cobertura terrestre com alta similaridade multiespectral. Embora a técnica de segmentação do SOM, utilizada nos métodos propostos nos capítulos 4 e 5, possibilite a descoberta de agrupamentos com alto grau de sobreposição no espaço multiespectral, a mesma se baseia apenas na informação de distância entre os agrupamentos. Já a técnica de segmentação do SOM, utilizada no método de classificação do capítulo 6, utiliza informações de cunho espacial, além da informação de distância, para discriminar as classes das imagens. Resultados experimentais apresentados naquele capítulo mostram que o uso dessas informações adicionais são potencialmente valiosas para distinguir classes de cobertura terrestre que são muito próximas no espaço de atributos das imagens.

## 7.2 Contribuições

Buscando explorar as características e as propriedades do SOM na construção de novos métodos de classificação não-supervisionada de imagens de sensoriamento remoto, a pesquisa desenvolvida ao longo deste trabalho apresenta contribuições que abrangem as áreas de redes neurais (especificamente o SOM), reconhecimento de padrões e processamento digital de

imagens de sensoriamento remoto. Tais contribuições podem ser descritas como segue:

- Revisão do estado da arte de métodos de classificação de imagens de sensoriamento remoto. Uma revisão das principais e mais recentes publicações relacionadas com a classificação de imagens de sensores remotos foi realizada, procurando dar uma ênfase maior àquelas que abordam métodos de classificação baseados no paradigma de aprendizado não-supervisionado;
- Constatação de que a estratégia de analisar os agrupamentos de dados de uma imagem, por meio de um conjunto reduzido de protótipos do SOM é uma abordagem computacionalmente efetiva para realizar a classificação de imagens de sensoriamento remoto. Tal estratégia, além de proporcionar uma detecção diferenciada e consistente das classes de cobertura terrestre existentes nas imagens, torna possível a utilização de métodos considerados computacionalmente inviáveis para o processamento de imagens de sensoriamento remoto, como métodos de agrupamentos hierárquicos e índices de validação de agrupamentos;
- Revisão bibliográfica de métodos de segmentação automática do SOM. Um levantamento dos principais trabalhos existentes na literatura que propõem e/ou aplicam técnicas para interpretar de maneira automática (ou semi-automática) os neurônios de um SOM treinado foi efetuado;
- Constatação de que o método de segmentação automática do SOM proposto por Costa e Andrade Netto (1999, 2001) é adequado para realizar o particionamento de SOMs treinados com dados de imagens de sensoriamento remoto;
- Proposição de uma nova estratégia de análise de agrupamentos de dados através do SOM. Usando SOMs segmentados por morfologia matemática e índices de validação de agrupamentos, a estratégia apresentada no capítulo 5 pode ser vista como um aprimoramento do método de análise de agrupamentos proposto por Costa e Andrade Netto (1999, 2001). Resultados experimentais mostram que a nova abordagem pode alcançar de maneira mais eficiente a melhor partição para o conjunto de dados analisado, uma vez que uma quantidade menor de SOMs é gerada durante o processo de detecção das classes;

- Levantamento bibliográfico das principais e mais recentes propostas de índices de validação de agrupamentos existentes na literatura;
- Proposição de uma estratégia baseada na quantização vetorial produzida pelo SOM para modificar os cálculos de índices de validação de agrupamentos. A estratégia proposta reduz o volume de processamento dos índices de validação, viabilizando a utilização dos mesmos em aplicações que envolvam grandes volumes de dados, como as de imagens de sensoriamento remoto;
- Análise experimental e comparativa do desempenho de diferentes índices de validação de agrupamentos. Uma série de testes experimentais aplicando sete índices de validação na avaliação de diferentes partições de conjuntos de dados variados foi executada. Os resultados dos testes apontam o índice  $\mathcal{V}_{CDbw}$ , seguido dos índices  $\mathcal{V}_{PBM}$  e  $\mathcal{V}_{DBK_{im}}$ , como os mais adequados para avaliar agrupamentos de dados com características diversificadas, incluindo dados de imagens de sensoriamento remoto;
- Constatação de que o uso combinado do SOM e de técnicas de agrupamentos hierárquicas não apenas provê uma exploração conjunta das vantagens de ambos os métodos na classificação de imagens de sensoriamento remoto, mas também permite que um método auxilie a contornar os problemas e limitações do outro. A representação da imagem por meio de um conjunto reduzido de protótipos do SOM torna viável o uso de métodos de agrupamentos hierárquicos, os quais são considerados impraticáveis para a análise de imagens de sensoriamento remoto. Além disso, a propriedade de ordenação topológica do SOM restringe possíveis fusões de protótipos, contribuindo, desse modo, para uma redução ainda maior da complexidade computacional do método hierárquico aplicado. Por outro lado, os métodos hierárquicos segmentam automaticamente o SOM treinado, e ainda organizam hierarquicamente os protótipos da rede neural, gerando dessa forma um dendrograma simplificado para a imagem analisada;
- Constatação de que a análise das imagens de sensoriamento remoto por meio de janelas de pixels produzem resultados de classificação superiores aos de abordagens pixel a pixel. Devido à heterogeneidade espectral apresentada pela maioria das classes de cobertura

terrestre, abordagens de classificação baseadas apenas na assinatura espectral de pixels individuais não são capazes de capturar a maioria dos padrões existentes nas imagens, principalmente daquelas que apresentam melhores resoluções;

- Proposição de um procedimento para filtragem de amostras e protótipos do SOM que incorporam mais de uma classe de cobertura terrestre. Resultados experimentais mostram que esses tipos de amostras e protótipos do SOM apresentam alto grau de heterogeneidade espectral e funcionam como pontos de ligação entre agrupamentos de dados distintos. Um procedimento de filtragem baseado em medidas de textura em imagens é proposto para realizar a filtragem dessas amostras e protótipos heterogêneos, contribuindo desse modo para facilitar o processo de discriminação das classes de uma imagem através do SOM;
- Constatação de que o uso de informação espacial a respeito dos padrões da imagem colabora para a separação de classes de cobertura terrestre que apresentam alto grau de similaridade espectral. Resultados experimentais mostram que nos casos em que a utilização apenas de medidas de distância multiespectral não é suficiente para discriminar determinadas classes de cobertura terrestre, o uso de informação espacial, particularmente do grau de compacidade e de fronteira espacial das classes, pode ser útil para a obtenção de resultados de classificação com maior grau de exatidão;
- Proposição de uma abordagem baseada em janelas de pixels para o algoritmo K-médias. Buscando proporcionar ao algoritmo K-médias a capacidade de capturar a heterogeneidade espectral apresentada pelas classes de cobertura terrestre, o mesmo foi modificado para processar amostras das imagens na forma de janelas de pixels, em vez de analisar pixels isolados como é proposto pelo algoritmo padrão. Essa versão modificada do algoritmo estatístico foi denominada de K-médias "janelado";
- Constatação da viabilidade de reutilização de SOMs para realizar a classificação de imagens de satélite que apresentam conjuntos de padrões de cobertura terrestres semelhantes.

É importante salientar, ainda, que o presente trabalho também contribui para uma maior difusão de métodos alternativos de classificação não-supervisionada de imagens na área de



sensoriamento remoto, provendo através dos mesmos vantagens importantes em relação às técnicas disponíveis nos pacotes ou softwares atuais de processamento de imagens de satélite. Conforme mencionado em Campbell (2002), não é possível declarar que um dado classificador de imagens é o "melhor" para qualquer situação, tendo em vista que as características de cada imagem e as circunstâncias para cada estudo variam grandemente. Portanto, a proposição de diferentes métodos de classificação, feita nesse trabalho, contribui no sentido de prover ao usuário ou analista de imagens uma gama maior de métodos para interpretar a grande diversidade de imagens de satélite que estão sendo coletadas e armazenadas atualmente em todo o mundo.

Finalmente, outro aspecto importante que também pode ser observado com o desenvolvimento desta pesquisa é que, embora os trabalhos já publicados que aplicam redes neurais na classificação de imagens de sensores remotos apresentem bons resultados em comparação com métodos convencionais, a maioria deles, além de abordarem muito superficialmente o processo de classificação em si, não exploram todo o potencial que as redes neurais podem oferecer. De fato, redes neurais de *Perceptrons* Multicamadas com algoritmo de treinamento *backpropagation* são utilizadas como se fossem a única alternativa para os classificadores estatísticos convencionais. O SOM, quando aplicado, quase sempre conta com a intervenção do usuário ou de um analista de imagens que realiza a interpretação do seu mapa de neurônios e detecta dessa forma as classes de interesse da imagem. Portanto, o presente trabalho também mostra, especialmente através da aplicação e segmentação automática do SOM, que outras alternativas no campo das redes neurais ainda podem ser melhor exploradas e aplicadas na análise de imagens de sensoriamento remoto.

### 7.3 Trabalhos futuros

Embora a eficiência dos métodos de classificação propostos nesta tese tenha sido comprovada através da realização de testes experimentais com diferentes imagens, verificou-se que algumas etapas deixam espaço para possíveis aprimoramentos e que novas idéias podem ser ainda investigadas e aplicadas. Possíveis extensões da pesquisa desenvolvida neste trabalho são

descritas a seguir:

- Ainda que os métodos de classificação propostos apresentem boas perspectivas de aplicação, é importante observar que o desempenho dos mesmos depende fortemente da qualidade do mapeamento realizado pelo SOM. O uso dos procedimentos e técnicas apresentados aqui supõe que o treinamento do SOM tenha sido sempre bem sucedido. Sendo assim, questões sobre as distorções topológicas ou sobre os erros de quantização produzidos pelo mapeamento da rede podem ser melhor investigadas. Na literatura é possível encontrar contribuições importantes sobre o assunto. Em Peres (2006), por exemplo, é apresentada uma abordagem heurística para estimar a dimensão topográfica ideal do espaço de saída do SOM usando a Teoria de Fractais e do Raciocínio Aproximado Fuzzy;
- Variações do SOM não foram empregadas nessa pesquisa, uma vez que planejou-se trabalhar com sua versão padrão, por ser a mais comumente utilizada. No entanto, variantes e novos desenvolvimentos do SOM podem naturalmente ser aplicados para aumentar a eficácia dos métodos de classificação propostos. Em particular, versões dinâmicas do SOM, as quais não requerem que o usuário especifique previamente o tamanho do mapa, podem ser empregadas como alternativa para a arquitetura estática da rede (Fritzke, 1991; Rauber *et al.*, 2002; Barbalho *et al.*, 2003);
- Investigar a proposta de aperfeiçoamento da *U-matrix* apresentada por Ultsch (2003). Ultsch propôs um aprimoramento da matriz de distâncias unificadas, a qual denominou de *U\*-matrix*, incorporando informação de densidade a respeito dos dados de entrada que estão associados aos protótipos do SOM. Segundo o autor, a *U\*-matrix* proporciona uma visualização mais sofisticada para conjunto de dados que apresentam alta dimensionalidade. No entanto, a construção da *U\*-matrix* exige que algum método de estimação de densidade do espaço de entrada do SOM seja aplicado, podendo gerar imposições sobre a estrutura dos dados, além de demandar uma maior complexidade de cálculo;
- Estudar a influência do erro de quantização do SOM no cálculo modificado dos índices de validação de agrupamentos. Embora a estratégia proposta para modificar o cálculo dos índices de validação tenha apresentado resultados satisfatórios, uma investigação mais

aprofundada sobre a relação da qualidade da quantização vetorial produzida pelo SOM e os erros de aproximação das versões modificadas dos índices pode ser feita;

- Dado o baixo tempo de processamento apresentado pelas versões modificadas dos índices de validação de agrupamentos, a possibilidade de utilização dessas versões para decidir sobre a fusão de protótipos ou grupos de protótipos do SOM pode ser estudada. Ao invés de se empregar apenas medidas de distância inter-agrupamentos nos critérios de fusão dos métodos de agrupamentos (como é feito usualmente), versões modificadas de índices de validação que se baseiam em medidas inter e intra-agrupamentos podem ser utilizadas para agregar mais informação ao processo de decisão sobre a junção ou não de dados;
- Estudar mecanismos que possam determinar automaticamente o tamanho ideal das janelas amostrais de treinamento, considerando o grau de complexidade dos padrões de cobertura terrestre e a resolução espacial da imagem de entrada. Adicionalmente, uma investigação sobre possíveis descritores para as janelas amostrais baseados em medidas de textura pode ser feita, buscando desse modo reduzir a dimensionalidade das amostras e representar de maneira adequada os diferentes padrões existentes nas imagens;
- Experimentar a atribuição de pesos diferentes para os índices que são utilizados no critério de fusão do método proposto no capítulo 6 (dado pela equação 6.3). Adotando tal estratégia, o usuário tem a possibilidade de realizar o processo de classificação das imagens dando maior importância para uma ou mais informações dentre as três que são incorporadas na análise dos agrupamentos (informação de distância, de compacidade espacial ou de fronteira espacial);
- Ainda que a utilização de informações espaciais a respeito dos padrões da imagem possa contribuir significativamente para a detecção de diferentes classes que apresentam a mesma aparência espectral, outros atributos, como forma e tamanho, podem ser incorporados para aumentar o poder de discriminação dos métodos de classificação propostos, principalmente para o caso de imagens de alta resolução espacial que possuem padrões de área urbana ou uma grande quantidade de edificações (Andrade *et al.*, 2002);

- Investigar a proposição de mecanismos de segmentação do SOM baseados em sistemas nebulosos. Com base em informações de distância, erro de quantização, topologia e densidade pode-se definir graus de "conectividade *fuzzy*" entre os neurônios do SOM. Índices de validação nebulosos também podem ser utilizados para avaliar as diferentes configurações de agrupamentos produzidas;
- Experimentar a utilização de *ensembles* de redes neurais com o objetivo de aumentar a acurácia das classificações resultantes;
- Uma vez que a estrutura hierárquica gerada pelos métodos de classificação propostos, em conjunto com o MCI, pode ser utilizada como interface de visualização dos dados das imagens, aplicações futuras podem incluir a implementação de mecanismos de visualização mais sofisticados que permitam ao usuário navegar pelos dendrogramas simplificados (ou árvores de mapas), podendo dessa forma compreender com maior facilidade a estrutura e relacionamento das classes existentes nas imagens.

## 7.4 Publicações

- GONÇALVES, M. L., ANDRADE NETTO, M. L.; COSTA, J. A. F., ZULLO JR, J. Automatic Remotely Sensed Data Clustering by Tree-Structured Self-Organizing Maps. *Proceedings of IEEE International Geoscience and Remote Sensing (IGARSS'05)*, 4p., Seoul, 2005.
- GONÇALVES, M. L., ANDRADE NETTO, M. L.; COSTA, J. A. F., ZULLO JR, J. Análise de Agrupamentos usando Mapas de Kohonen Segmentados por Morfologia Matemática e o Índice de Validação CDbw. *Anais do VII Congresso Brasileiro de Redes Neurais (CBRN'05)*, Natal, 2005.
- GONÇALVES, M. L., ANDRADE NETTO, M. L. Uma abordagem para Análise de Agrupamentos baseada em Mapas de Kohonen segmentados por Morfologia Matemática e Índices de Validação. *Relatório Técnico (RT-DCA05-02)*. Departamento de Engenharia de Computação e Automação Industrial. Faculdade de Engenharia Elétrica e de

Computação. Universidade Estadual de Campinas, 2005.

- GONÇALVES, M. L., ANDRADE NETTO, M. L.; COSTA, J. A. F., ZULLO JR, J. Data Clustering using Self-Organizing Maps segmented by Mathematic Morphology and Simplified Cluster Validity Indexes: an application in remotely sensed images. *Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN'06)*, p. 8854-8861, Vancouver, 2006.
- GONÇALVES, M. L., ANDRADE NETTO, M. L., COSTA, J. A. F. A Three-Stage Approach based on the Self-Organizing Map for Satellite Image Classification. *Lecture Notes in Computer Science*, v. 4669, p. 680-689, 2007.
- GONÇALVES, M. L., ANDRADE NETTO, M. L., COSTA, J. A. F. Explorando as Propriedades do Mapa Auto-organizável de Kohonen na Classificação de Imagens de Satélite. *Anais do VI Encontro Nacional de Inteligência Artificial (ENIA'07)*, p. 1132-1141, Rio de Janeiro, 2007.
- GONÇALVES, M. L., ANDRADE NETTO, M. L., COSTA, J. A. F., ZULLO JR, J.. An Unsupervised Method of Classifying Remotely Sensed Images using Kohonen Self-organizing Maps and Agglomerative Hierarchical Clustering Methods. *International Journal of Remote Sensing*, v. 29, p. 3171-3207, 2008.
- GONCALVES, M. L., ANDRADE NETTO, M. L., COSTA, J. A. F. Cluster Analysis of Land-cover Images using Automatically Segmented SOMs with Textural Information. *Lecture Notes in Computer Science*, v. 5326, p. 483-490, 2008.
- GONCALVES, M. L., ANDRADE NETTO, M. L., COSTA, J. A. F., ZULLO JR, J. Classificação Não-supervisionada de Imagens de Sensores Remotos utilizando Redes Neurais Auto-organizáveis e Métodos de Agrupamentos Hierárquicos. *Revista Brasileira de Cartografia*, v. 60, 13p., 2008.
- GONÇALVES, M. L., ANDRADE NETTO, M. L., COSTA, J. A. F. Land-Cover Classification Using Self-Organizing Maps Clustered with Spectral and Spatial Information. *Neural Computing and Applications*, 2009. (submetido)

# Referências Bibliográficas

- ALAHAKOON, D., HALGAMUGE, S.K. e SRINIVASAN, B., 2000, Dynamic Self-organizing Maps with controlled for knowledge discovery. *IEEE Transactions on Neural Networks*, **10**, 100–114.
- ANDRADE, A.F., CENTENO, J.S. e ARAKI, H., 2002, Utilização de parâmetros de forma como dado auxiliar na classificação de imagens IKONOS através de Redes Neurais Artificiais. In *Proceedings of the Simpósio Brasileiro de Geomática*, julho, pp. 342–349.
- ARAI, K., 1992, A supervised thematic mapper classification with a purification of training samples. *International Journal of Remote Sensing*, **13**, 2039–2049.
- AZCARRAGA, A.P., 2000, Assessing self-organization using order metrics. In *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks*, Como, Itália, July, pp. 159–164.
- AZUAJE, F., 2002, A cluster validity framework for genome expression data. *Bioinformatics*, **18**, 319–320.
- BACHMANN, C.M., 2003, Improving the performance of classifiers in high-dimensional remote sensing applications: an adaptive resampling strategy for error-prone exemplars (ARESEPE). *IEEE Transactions on Geoscience and Remote Sensing*, **41**, 2101–2112.
- BALL, G. e HALL, D., 1967, A clustering technique for summarizing multivariate data. *Behavior Science*, **12**, 153–155.

- BANDYOPADHYAY, S., 2005, Satellite image classification using genetically guided fuzzy clustering with spatial information. *International Journal of Remote Sensing*, **26**, 579–593.
- BANDYOPADHYAY, S. e MAULIK, U., 2001, Nonparametric genetic clustering: comparison of validity indices. *IEEE Transactions on Systems, Man, and Cybernetics*, **31**, 120–125.
- BARALDI, A., BLONDA, P., PARMIGGIANI, F. e SATALINO, G., 2000, Contextual clustering for image segmentation. *Optical Engineering*, **39**, 1–17.
- BARBALHO, J.M., COSTA, J.A.F., NETO, A.D.D. e ANDRADE NETTO, M.L., 2003, Hierarchical and dynamic SOM applied to image compression.. In *Proceedings of the International Joint Conference on Neural Networks*, 1, july, pp. 753–758.
- BERRY, M.J.A. e LINOFF, G., 1996, *Data mining techniques for marketing, sales and customer support*, 2a. edição (John Willey and Sons, Inc).
- BEUCHER, S. e LANTUÉJOUL, C., 1979, Use of watersheds in contour detection. In *Proceedings of the International Workshop Image Processing, Real-time Edge and Motion Detection/Estimation*, Rennes, França.
- BEZDEK, J.C. e PAL, N.R., 1998, Some new indexes of cluster validity. *IEEE Transactions on Systems, Man and Cybernetics*, **28**, 301–315.
- BLACKMORE, J. e MIIKKULAINEN, R., 1993, Incremental grid growing: encoding high-dimensional structure into two-dimensional feature map. In *Proceedings of the IEEE International Conference on Neural Networks*, San Francisco, USA.
- BLAKE, C.L. e MERZ, C.J., 1998, UCI repository of machine learning databases. *Department of Information and Computer Science, University of California at Irvine, CA*.
- BLASCHKE, T., GLÄSSER, C. e LANG, S., 2007, Processamento de imagens num ambiente integrado SIG/sensoriamento remoto - tendências e conseqüências. *Sensoriamento remoto e SIG avançados: novos sistemas sensores, métodos inovadores*. (São Paulo: Oficina de textos. T. Blaschke e H. Kux. pp. 11–18).

- BOUDRAA, A.O., 1999, Dynamic estimation of number of clusters in data sets. *Electronics Letters*, **35**, 1606–1608.
- BRITO, J.L.S. e ROSA, R., 2003, Introdução ao processamento digital de imagens. URL: <http://www.ig.ufu.br/lacar/>. Recuperado em 24/08/2008.
- CALINSKI, R.B. e HARABASZ, J., 1974, A Dendrite method for cluster analysis. *Comm. in Statistics*, **3**, 1–27.
- CAMPBELL, J.B., 2002, *Introduction to remote sensing* (New York: The Guilford Press).
- CARVALHO JR., O.A., CARVALHO, A.P.F. e GUIMARÃES, P.R.M.R.F., 2002, Classificação e eliminação dos ruídos em imagens hiperespectrais pela análise seqüencial da transformação por fração de ruído mínima. *Revista Brasileira de Geofísica*, **20**, 4p.
- CASTRO, L.N. e VON ZUBEN, F.J., 1999, An improving pruning technique with restart for the Kohonen Self-organizing feature map. In *Proceedings of the IEEE International Joint Conference on Neural Networks*, 3, Julho, pp. 1916–1919.
- CASTRO, M.L.M.S., 2005, “Classificação de imagens de satélite por aglomeração usando o Matlab”, Tese de mestrado, Faculdade de Ciências da Universidade do Porto, Porto, Portugal.
- CASTRO JR., R.M., 2001, Introdução ao sensoriamento remoto. URL: <http://www.ltc.ufes.br/>. Recuperado em 18/08/2008.
- CHAUDHURI, B.B. e SARKAR, N., 1995, Texture segmentation using fractal dimension. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **17**, 72–77.
- CHEN, C.H., 1990, Texture unit, texture spectrum, and texture analysis. *IEEE Transactions on Geoscience and Remote Sensing*, **28**, 509–512.
- CHEN, D. e STOW, D., 2002, The effect of training strategies on supervised classification at different spatial resolutions. *Photogrammetric Engineering and Remote Sensing*, **68**, 1155–1162.



- CHOI, D.I. e PARK, S.H., 1994, Self-creating and organizing neural networks. *IEEE Transactions on Neural Networks*, **5**, 561–575.
- CHUNG, K. e LIN, J., 2007, Faster and more robust point symmetry-based K-means algorithm. *Pattern Recognition*, **40**, 410–422.
- COHN, D., ATLAS, L. e LADNER, R., 1994, Improving generalization with active learning. *Machine Learning*, **15**, 201–221.
- CONGALTON, R.G. e GREEN, K., 1999, *Assessing the accuracy of remotely sensed data: principles and practices* (New York: Lewis Publishers).
- COSTA, J.A.F., 1999, “Classificação automática e análise de dados por redes neurais auto-organizáveis”, Tese de Doutorado, Departamento de Engenharia de Computação e Automação Industrial, Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas, SP, Brasil.
- COSTA, J.A.F., 2005, Segmentação do SOM por métodos de agrupamentos hierárquicos com conectividade restrita. In *Proceedings of the VII Congresso Brasileiro de Redes Neurais*, Natal, RN, Outubro, 2005.
- COSTA, J.A.F. e ANDRADE NETTO, M.L., 1999, Estimating the number of clusters in multivariate data by self-organizing maps. *International Journal of Neural Systems*, **9**, 195–202.
- COSTA, J.A.F. e ANDRADE NETTO, M.L., 2001, Clustering of complex shaped data sets via Kohonen maps and mathematical morphology. In *Proceedings of the SPIE, Data Mining and Knowledge Discovery*, 4384, pp. 16–27.
- COSTA, J.A.F. e ANDRADE NETTO, M.L., 2005, Segmentação do SOM baseada em particionamento de grafos. In *Proceedings of the VI Congresso Brasileiro de Redes Neurais*, São Paulo, junho, 2003, pp. 451–456.
- COTTRELL, M.F., FORT, J.C. e PAGES, G., 1995, Comment about analysis of the convergence properties of topology preserving neural networks. *IEEE Transactions on Neural Networks*, **6**, 797–799.

- CRACKNELL, A.P. e HAYES, L.W.B., 1991, *Introduction to Remote Sensing* (New York: Taylor & Francis).
- CRÓSTA, A.P., 1992, *Processamento digital de imagens de sensoriamento remoto* (Campinas: Instituto de Geociências/UNICAMP).
- DAVIES, D.L. e BOULDIN, D.W., 1979, A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **PAM1-1**, 224–227.
- DEAN, A.M. e SMITH, G.M., 2003, An evaluation of per-parcel land cover mapping using maximum likelihood class probabilities. *International Journal of Remote Sensing*, **24**, 2905–2920.
- DUDA, R.O. e HART, P.E., 1973, *Pattern classification and scene analysis*, 2a. edição (John Willey and Sons, Inc).
- DUDA, R.O., HART, P.E. e STORK, D.G., 2000, *Pattern classification* (New York: John Wiley & Sons).
- DUDA, T. e CANTY, M., 2002, Unsupervised classification of satellite imagery: choosing a good algorithm. *International Journal of Remote Sensing*, **23**, 2193–2212.
- DUNN, J.C., 1974, Well separated clusters and optimal fuzzy partitions. *Journal of Cybernetics*, **4**, 95–104.
- E. VANSTEENKISTE, E., GAUTAMA, S. e PHILIPS, W., 1995, Analysing multispectral textures in very high resolution satellite images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **17**, 72–77.
- EHLERS, M., 2007, Sensoriamento remoto para usuários de SIG - sistemas sensores e métodos: as exigências e a realidade. *Sensoriamento remoto e SIG avançados: novos sistemas sensores, métodos inovadores*. (São Paulo: Oficina de textos. T. Blaschke e H. Kux. pp. 19–32).
- EMBRAPA, 2004, EMBRAPA - satélites de monitoramento . URL: <http://www.sat.cnpem.embrapa.br>. Recuperado em 20/08/2008.

- EPIPHANIO, J.C.N., 2005, CBERS - Satélite sino-brasileiro de recursos terrestres. In *Proceedings of the XII Simpósio Brasileiro de Sensoriamento Remoto*, Goiânia, Brasil, abril, 2005, pp. 915–922.
- EPIPHANIO, J.C.N., 2006, Satélites de sensoriamento remoto. URL: <http://www.dsr.inpe.br/vcsr/html/>. Recuperado em 19/08/2008.
- ERBEK, F.S., ÖZKAN, C. e TABERNER, M., 2004, Comparison of maximum likelihood classification method with supervised artificial neural network algorithms for land use activities. *International Journal of Remote Sensing*, **25**, 1733–1748.
- ERWIN, E., OBERMAYER, K. e SCHULTEN, K., 1992, Self-organizing maps: ordering, convergence properties and energy functions. *Biological Cybernetics*, **67**, 47–55.
- ESPINOSA, C.H., REDONDO, M.F. e SOSPEDRA, J., 2004, Some experiments on ensembles of neural networks for hyperspectral image classification. *Lecture Notes in Computer Science*, **3213**, 677–684.
- FLANAGAN, J.A., 1996, Self-organization in Kohonen's SOM. *Neural Networks*, **9**, 1185–1197.
- FOODY, G.M., 1999, The significance of border training patterns in classification by a feed-forward neural network using back propagation learning. *International Journal of Remote Sensing*, **20**, 3549–3562.
- FOODY, G.M. e MATHUR, A., 2004, Toward intelligent training of supervised image classifications: directing training data acquisition for SVM classification. *Remote Sensing of Environment*, **93**, 107–117.
- FRITZKE, B., 1991, Unsupervised clustering with growing cell structures. In *Proceedings of the IEEE International Joint Conference on Neural Networks*, Seattle.
- FRIZZELLE, B.G. e MOODY, A., 2001, Mapping continuous distributions of land cover: a comparison of maximum-likelihood estimation and artificial neural networks. *Photogrammetric Engineering and Remote Sensing*, **67**, 693–705.

- GARCÍA, H.L. e GONZÁLEZ, I.M., 2004, Self-organizing map and clustering for wastewater treatment monitoring. *Engineering Applications of Artificial Intelligence*, **17**, 215–225.
- GIACINTO, G., ROLI, F. e FUMERA, G., 2000, Unsupervised learning of neural network ensembles for image classification. In *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks*, Como, Itália, July, 2000, pp. 155–162.
- GONÇALVES, M.L., 1997, “Uma arquitetura neural modular para classificação de imagens multiespectrais de sensoriamento remoto”, Tese de mestrado, Departamento de Engenharia de Computação e Automação Industrial, Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas, SP, Brasil.
- GONÇALVES, M.L., ANDRADE NETTO, M.L., e COSTA, J.A.F., 2005a, Análise de agrupamentos usando mapas de Kohonen segmentados por morfologia matemática e o índice de validação CD<sub>bw</sub>. In *Proceedings of the VII Congresso Brasileiro de Redes Neurais (CBRN'05)*, Natal, RN, Brasil, Outubro, 2005.
- GONÇALVES, M.L., ANDRADE NETTO, M.L., e COSTA, J.A.F., 2007a, Explorando as propriedades do Mapa Auto-organizável de Kohonen na classificação de imagens de satélite. In *Proceedings of the VI Encontro Nacional de Inteligência Artificial (ENIA'07)*, Rio de Janeiro, Brasil, Outubro, 2007, pp. 1132–1141.
- GONÇALVES, M.L., ANDRADE NETTO, M.L., e COSTA, J.A.F., 2007b, A three-stage approach based on the Self-Organizing Map for satellite image classification. *Lecture Notes in Computer Science*, **4669**, 680–689.
- GONÇALVES, M.L., ANDRADE NETTO, M.L., e COSTA, J.A.F., 2008a, Cluster Analysis of Land-cover Images using Automatically Segmented SOMs with Textural Information. *Lecture Notes in Computer Science*, **5326**, 483–490.
- GONÇALVES, M.L. e ANDRADE NETTO, M.L., 2005, Uma abordagem para análise de agrupamentos baseada em mapas de Kohonen segmentados por morfologia matemática e índices de validação. Relatório técnico, RT-DCA-02/05, Departamento de Engenharia de Computação e Automação Industrial, Faculdade de Engenharia Elétrica e de Computação, UNICAMP.

- GONÇALVES, M.L., ANDRADE NETTO, M.L. e COSTA., J.A.F., 2009, Land-Cover classification using Self-Organizing Maps clustered with spectral and spatial Information. *Neural Computing and Applications* (submetido).
- GONÇALVES, M.L., ANDRADE NETTO, M.L., COSTA, J.A.F. e ZULLO JR., J., 2005b, Automatic remotely sensed data clustering by tree-structured self-organizing maps. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS'05)*, Seoul, Korea, July, 2005, 4p.
- GONÇALVES, M.L., ANDRADE NETTO, M.L., COSTA, J.A.F. e ZULLO JR., J., 2006, Data clustering using self-organizing maps segmented by mathematic morphology and simplified cluster validity indexes. In *Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN'06)*, Vancouver, BC, Canada, July, 2006, pp. 8854–8861.
- GONÇALVES, M.L., ANDRADE NETTO, M.L., COSTA, J.A.F. e ZULLO JR., J., 2008b, An unsupervised method of classifying remotely sensed images using Kohonen self-organizing maps and agglomerative hierarchical clustering methods. *International Journal of Remote Sensing*, **29**, 3171–3207.
- GONÇALVES, M.L., ANDRADE NETTO, M.L. e ZULLO JR., J., 1996, A neural architecture for image classification using Kohonen feature extraction. In *Proceedings of the XVIII International Congress for Photogrammetry and Remote Sensing (ISPRS, Commission II)*, Viena, Julho, 1996, pp. 117–122.
- GONÇALVES, M.L., ANDRADE NETTO, M.L. e ZULLO JR., J., 1998a, Comparação de uma arquitetura neural com o algoritmo de máxima verossimilhança na classificação de uma cena LANDSAT/TM. In *Proceedings of the IX Simposio Brasileiro de Sensoriamento Remoto (SBSR'98)*, Santos, Setembro, 1998.
- GONÇALVES, M.L., ANDRADE NETTO, M.L. e ZULLO JR., J., 1998b, A neural architecture for the classification of remote sensing imagery with advanced learning algorithms. In *Proceedings of the IEEE Workshop on Neural Networks for Signal Processing (NNSP'98)*, Cambridge, Setembro, 1998, pp. 577–586.

- GONÇALVES, M.L., ANDRADE NETTO, M.L. e ZULLO JR., J., 1998c, Uma Arquitetura Neural para Reconhecimento de Padroes Multiespectrais de Imagens de Sensoriamento Remoto. In *Proceedings of the V Simposio Brasileiro de Redes Neurais (SBRN'98)*, Belo Horizonte, MG, Dezembro, 1998, pp. 47–52.
- GONÇALVES, M.L., ANDRADE NETTO, M.L., ZULLO JR., J. e COSTA, J.A.F., 2008c, Classificação não-supervisionada de imagens de sensores remotos utilizando Redes Neurais Auto-organizáveis e Métodos de Agrupamentos Hierárquicos. *Revista Brasileira de Cartografia*, **60-01**, Abril, 13p.
- GONZÁLEZ, I.M. e GARCÍA, H.L., 2006, End-point detection of the aerobic phase in a biological reactor using SOM and clustering algorithms. *Engineering Applications of Artificial Intelligence*, **19**, 19–28.
- GRAYS, T.I. e MCCRARY, D.G., 1981, Meteorological satellite data: a tool to describe the health of the world's agriculture. *NASA REPORT, AgRI STARS EQ-NI-04042*.
- HADJIMITSIS, D.G., EVANGELOU, I., LAZAKIDOU, A. e CLAYTON, C.R.I., 2003, Unsupervised Classification of Remotely Sensed Images Using Self-Organizing Maps for mapping land-cover changes. In *Proceedings of the Remote Sensing and Photogrammetry Society Annual Conference*, Nottingham, UK, Setembro, 2003.
- HAESE, K., 2001, Self-organizing feature maps with self-adjusting learning parameters. *IEEE Transactions on Neural Networks*, **9**, 1270–1278.
- HALKIDI, M., BATISTAKIS, Y. e VAZIRGIANNIS, M., 2001, On clustering validation techniques. *Journal of Intelligent Information Systems*, **17**, 107–145.
- HALKIDI, M., BATISTAKIS, Y. e VAZIRGIANNIS, M., 2002, Cluster validity methods: part I. *SIGMOD Record*, **31**, 40–45.
- HALKIDI, M. e VAZIRGIANNIS, M., 2002, Clustering validity assessment using multi representatives. In *Proceedings of the SETN Conference*, Thessaloniki, Grécia, Abril, 2002.

- HALKIDI, M., VAZIRGIANNIS, M. e BATISTAKIS, Y., 2000, Quality scheme assessment in the clustering process. In *Proceedings of the 4th European Conference on Principles of Data Mining and Knowledge Discovery*, Lyon, France, pp. 265–276.
- HANSEN, M., DUBAYAH, R. e DEFRIES, R., 1996, Classification trees: an alternative to traditional land cover classifiers. *International Journal of Remote Sensing*, **17**, 1075–1081.
- HARALICK, R.M., SHANMUGAM, K. e DINSTEN, 1973, Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, **6**, 610–621.
- HAYKIN, S., 1994, *Neural Networks: A Comprehensive Foundation* (Englewood Cliffs, NJ: Prentice-Hall).
- HUANG, J.Z., NG, M.K., RONG, H. e LI, Z., 2005, Automated variable weighting in K-means type clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **27**, 657–668.
- HUANG, K., 2002, The use of a newly developed algorithm of divisive hierarchical clustering for remote sensing image analysis. *International Journal of Remote Sensing*, **23**, 3149–3168.
- HUBERT, L. e SCHULTZ, J., 1976, Quadratic assignment as a general data analysis strategy. *British Journal of Mathematical and Statistical Psychology*, **29**, 190–241.
- HUBERT, L.J. e ARABIE, P., 1985, Comparing partitions. *Journal of Classification*, **2**, 193–218.
- HUTCHINSON, F.C., 1982, Techniques for combining Landsat and ancillary data for digital classification improvement. *Photogrammetric Engineering and Remote Sensing*, **48**, 123–130.
- INPE, 2007, Programa CBERS. URL: <http://www.cbbers.inpe.br>. Recuperado em 19/09/2007.
- JAIN, A.K. e FARROKHANIA, F., 1991, Unsupervised texture segmentation using Gabor filters. *Pattern Recognition*, **24**, 1167–1186.
- JI, C.Y., 2000, Land-use classification of remotely sensed data using Kohonen self-organizing feature map neural networks. *Photogrammetric Engineering and Remote Sensing*, **66**, 1451–1460.

- JI, M., 2003, Using fuzzy sets to improve cluster labeling in unsupervised classification. *International Journal of Remote Sensing*, **24**, 657–671.
- JOHNSON, S.C., 1967, Hierarchical clustering schemes. *Psychometria*, **32**, 241–254.
- JÖRNSTEN, R., 2004, Clustering and classification based on the L1 data depth. *Journal of Multivariate Analysis*, **90**, 67–89.
- KANGAS, J.A., KOHONEN, T. e LAAKSONEN, J.T., 1990, Variants of Self-organizing Maps. *IEEE Transactions on Neural Networks*, **1**, 93–99.
- KASKI, S., 1997, “Data exploration using Self-organizing Maps”, Tese de Doutorado, Helsinki University of Technology, Finland.
- KASKI, S. e LAGUS, K., 1996, Comparing self-organizing maps. *Lectures Notes in Computer Science*, **1112**, 809–814.
- KEGANG, L., YUANCHAO, L. e XIAOLONG, W., 2007, A dynamic SOM algorithm for clustering large-scale document collection. In *Proceedings of the VI International Conference on Advanced Language Processing and Web Information Technology*, China, pp. 15–20.
- KELLY, M., SHAARI, D., GUO, Q. e LIU, D., 2004, A comparison of standard and hybrid classifier methods for mapping hardwood mortality in areas affected by ‘sudden oak death’. *Photogrammetric Engineering and Remote Sensing*, **70**, 1229–1239.
- KHAN, S.S. e AHMAD, A., 2004, Cluster center initialization algorithm for K-means clustering. *Pattern Recognition Letters*, **25**, 1293–1302.
- KIANG, M.Y., 2001, Extending the Kohonen self-organizing map networks for clustering analysis. *Computational Statistics and Data Analysis*, **38**, 161–180.
- KIM, M. e RAMAKRISHNA, R.S., 2005, New indices for cluster validity assessment. *Pattern Recognition Letters*, **26**, 2353–2363.
- KOHONEN, T., 1982, Self-organizing formation of topologically correct feature maps. *Biological Cybernetics*, **43**, 59–69.



- KOHONEN, T., 1997, *Self-organizing Maps*, 2a. edição (Berlin:Springer-Verlag).
- KOHONEN, T., OJA, E., SIMULA, O. e KANGAS, J., 1996, Engineering applications of the self-organizing map. In *Proceedings of the IEEE*, 84, pp. 1358–1384.
- KOIKKALEIN, P., 1994, Progress with the tree-structured self-organizing map. In *Proceedings of the XI European Conference on Artificial Intelligence*, New York (John Wiley & Sons), pp. 211–215.
- KUX, H. e BLASCHKE, T., 2007, *Sensoriamento remoto e SIG avançados: novos sistemas sensores, métodos inovadores* (São Paulo: Oficina de Textos).
- LANDGREBE, D.A., 2005, Multispectral land sensing: where from, where to?. *IEEE Transactions on Geoscience and Remote Sensing*, **43**, 414–421.
- LANDIS, J.R. e KOCH, G.G., 1977, The measurement of observer agreement for categorical data. *Biometrics*, **33**, 159–174.
- LAWRENCE, R., BUNN, A., POWELL, S. e ZAMBOM, M., 2004, Classification of remotely sensed imagery using stochastic gradient boosting as a refinement of classification tree analysis. *Remote Sensing of Environment*, **90**, 331–336.
- LEES, B.G. e RITMAN, K., 1991, Decision-tree and rule-introduction approach to integration of remotely sensed and GIS data in mapping vegetation in disturbed or hilly environments. *Environmental Management*, **15**, 823–831.
- LERSCH, R., HAERTEL, V. e SHIMABUKURO, Y., 2007, Investigação sobre o uso de dados auxiliares no processo de classificação de imagens digitais empregando princípios da teoria da evidência. In *Proceedings of the XIII Simpósio Brasileiro de Sensoriamento Remoto*, Florianópolis, Abril, pp. 5871–5878.
- LIKAS, A., VLASSIS, N. e VERBEEK, J., 2003, The global K-means clustering algorithm. *Pattern Recognition*, **36**, 451–461.

- LILLESAND, T.M. e KIEFER, R.W., 2000, *Remote sensing and image interpretation*, 2a. edição (New York: John Wiley & Sons, Inc.).
- LIU, A., BICHENG, L., CHEN, G. e ZHANGIU, X., 2005, A new ART neural networks for remote sensing image classification. *Lecture Notes in Computer Science*, **3611**, 37–42.
- LIU, X., SKIDMORE, A.K. e OOSTEN, H.V., 2002, Integration of classification methods for improvement of land-cover map accuracy. *Journal of Photogrammetry and Remote Sensing*, **56**, 257–268.
- LIU, X., SKIDMORE, K.A. e OOSTEN, H.V., 2003, An experimental study on spectral discrimination capability of a backpropagation neural network classifier. *International Journal of Remote Sensing*, **24**, 673–688.
- LO, P.C. e CHOI, J., 2004, A hybrid approach to urban land use/cover mapping using Landsat 7 Enhanced Thematic Mapper Plus (ETM+) images. *International Journal of Remote Sensing*, **25**, 2687–2700.
- LUENBERGER, D.G., 2003, *Linear and nonlinear programming* (New York: Springer).
- MAGNUSSEN, S., BOUDEWYN, P. e WULDER, M., 2004, Contextual classification of Landsat TM images to forest inventory cover types. *International Journal of Remote Sensing*, **25**, 2421–2440.
- MARÇAL, A.R.S. e BORGES, J.S., 2005, Estimation of the natural number of classes of a multispectral image. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS'05)*, Seoul, Coréia, Julho, 2005, pp. 3788–3791.
- MARÇAL, A.R.S. e CASTRO, L., 2005, Hierarchical clustering of multispectral images using combined spectral and spatial criteria. *IEEE Geoscience and Remote Sensing Letters*, **2**, 59–63.
- MARTINETZ, T. e SCHULTEN, K., 1991, A Neural Gas network learns topologies. *Artificial Neural Networks*. (Amsterdam: Elsevier Science Publishers. T. Kohonen, K. Mäkisara, O. Simula e J. Kangas (editores)).

- MARTINS, S.P., 2004, “Classificação textural de imagens Radarsat-1 para discriminação de alvos agrícolas”, Tese de mestrado, Instituto Nacional de Pesquisas Espaciais - INPE, São José dos Campos, SP, Brasil.
- MATHER, P.M., 2004, *Computer processing of remotely-sensed images: an introduction* (New York: John Wiley & Sons).
- MAULIK, U. e BANDYOPADHYAY, S., 2002, Performance evaluation of some clustering algorithms and validity indices. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **24**, 1650–1654.
- MERKL, D. e RAUBER, A., 1997, Alternative ways for cluster visualization in self-organizing maps. In *Proceedings of the Workshop on Self-organizing Maps (WSOM'97)*, Finlândia, Junho, 1997, pp. 106–111.
- MEYER, F. e BEUCHER, S., 1990, Morphological segmentation. *Journal of Visual Comm. and Image Representation*, **1**, 21–46.
- MIKULAINEN, R., 1993, Integrated connectionist models: building AI systems on subsymbolic foundations. *Integrating symbol processors and connectionist networks for Artificial Intelligence and Cognitive Modeling* (New York: New York: Academic Press: Holvar, V. and Uhr, L. (editores)).
- MILLIGAN, G.W. e COOPER, M.C., 1985, An examination of procedures for determining the number of clusters in a data set. *Psychometrika*, **50**, 159–179.
- MILLIGAN, G.W. e COOPER, M.C., 1986, A study of the comparability of external criteria for hierarchical cluster analysis. *Multivariate Behavioral Research*, **21**, 441–458.
- MITRA, P., SHANKAR, U. e PAL, S.K., 2004, Segmentation of multispectral remote sensing images using active support vector machines. *Pattern Recognition Letters*, **25**, 1067–1074.
- MORAES, E.C., 2006, Fundamentos de sensoriamento remoto. URL: <http://www.dsr.inpe.br/vcsr/html/>. Recuperado em 18/08/2008.

- MOREIRA, M.A., 2005, *Fundamentos do sensoriamento remoto e metodologias de aplicação*, 3 (Viçosa, MG: Editora UFV).
- MOREY, L. e AGRETI, A., 1984, The measurement of classification agreement: an adjustment to the rand statistic for chance agreement. *Educational and Psychological Measurement*, **44**, 33–37.
- MURTAGH, F., 1995, Interpreting the Kohonen self-organizing feature map using contiguity-constrained clustering. *Pattern Recognition Letters*, **16**, 399–408.
- NOVO, E.M.L.M., 1989, *Sensoriamento remoto: princípios e aplicações* (São Paulo: Editora Edgard Blücher LTDA).
- OCSA, A., BEDREGAL, C. e CUADROS-VARGAS, E., 2007, DB-GNG: A constructive self-organizing map based on densilty. In *Proceedings of the IEEE International Joint Conference on Neural Networks*, Orlando, pp. 1953–1958.
- OPOLON, D. e MOUTARDE, F., 2004, Fast semi-automatic segmentation algorithm for Self-Organizing Maps. In *Proceedings of the European Symposium on Artificial Neural Networks (ESANN'04)*, Bruges, Bélgica, Abril, 2004, pp. 507–512.
- PACHÊCO, A.P. e RIBAS, N.S., 1998, Sensoriamento remoto aplicado ao uso da terra. *GEODESIA online*, **4**, 4p.
- PACHÊCO, A.P., 2001, Sensoriamento remoto multitemporal aplicado à cobertura vegetal da Mata Atlântica. *GEODESIA online*, **1**, 4p.
- PAKHIRA, M.K., BANDYOPADHYAY, S. e MAULIK, U., 2004, Validity index fr crisp and fuzzy clusters. *Pattern Recognition*, **37**, 487–501.
- PAL, M. e MATHER, P.M., 2002, A comparison of decision tree and backpropagation neural network classifiers for land use classification. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS'02)*, 1, pp. 503–505.

- PAL, M. e MATHER, P.M., 2005, Support vector machines for classification in remote sensing. *International Journal of Remote Sensing*, **26**, 1007–1011.
- PAL, N.R. e BISWAS, J., 1997, Cluster validation using graph theoretic concepts. *Pattern Recognition*, **30**, 847–857.
- PAL, S.K., BANDYOPADHYAY, S. e MURTHY, C.A., 2001, Genetic classifiers for remotely sensed images: comparison with standard methods. *International Journal of Remote Sensing*, **22**, 2545–2569.
- PASQUARIELLO, G., ANCONA, N., BLONDA, P., TARANTINO, C., SATALINO, G. e D'ADDABBO, A., 2002, Neural network ensemble and support vector machine classifiers for the analysis of remotely sensed data: a comparison. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS'02)*, 1, pp. 509–511.
- PERES, S.M., 2006, “Dimensão topológica e Mapas Auto-organizáveis de Kohonen”, Tese de Doutorado, Departamento de Engenharia de Computação e Automação Industrial, Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas, SP, Brasil.
- PICHLER, O., TEUNER, A. e HOSTICKA, B.J., 1998, An unsupervised texture segmentation algorithm with feature space reduction and knowledge feedback. *IEEE Transactions on image processing*, **7**, 53–61.
- RAUBER, A., MERKL, D. e DITTENBACH, M., 2002, The growing hierarchical self-organizing map: exploratory analysis of high-dimensional data. *IEEE Transactions on Neural Networks*, **13**, 1331–1341.
- RENNÓ, C.D., 1995, “Avaliação de medidas texturais na discriminação de classes de uso utilizando imagens SIR-C/X-SAR do perímetro irrigado do Bebedouro, Petrolina, PE”, Tese de mestrado, Instituto Nacional de Pesquisas Espaciais - INPE, São José dos Campos, SP, Brasil.

- RICHARDS, J.A., 1995, *Remote sensing digital image analysis: an introduction* (Berlin: Springer Verlag).
- RICHARDS, J.A., 2005, Analysis of remotely sensed data: the formative decades and the future. *IEEE Transactions on Geoscience and Remote Sensing*, **43**, 422–432.
- RODRIGUES, L.M., RESENDE, M.R. e TIMOSZCZUK, A.P., 2007, Sistema de imageamento hiperespectral de alta resolução embarcado em plataforma aérea de baixo custo. In *Proceedings of the XIII Simpósio Brasileiro de Sensoriamento Remoto*, Florianópolis, Abril, pp. 7079–7085.
- ROUSSEEUW, P.J., 1987, Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math*, **20**, 53–65.
- SARLE, W.S. e KUO, A.H., 1993, The MODECLUS procedure. Relatório técnico, Technical Report P-256, Cary, NC: SAS Institute Inc.
- SCHAALE, M. e FURRER, R., 1995, Land surface classification by neural networks. *International Journal of Remote Sensing*, **16**, 3003–3031.
- SHAH, C.A., ARORA, M.K. e VARSHNEY, P.K., 2004, Unsupervised classification of hyperspectral data: an ICA mixture model based approach. *International Journal of Remote Sensing*, **25**, 481–487.
- SHARMA, S.C., 1996, *Applied Multivariate techniques*, 3a. edição (John Willey and Sons, Inc).
- SHEPHERD, G.M., 1988, *Neurobiology*, 2a. edição (New York: Oxford University Press).
- SILVA, M.A.S., MONTEIRO, A.M.V. e MEDEIROS, J.S., 2004, Semi-automatic geospatial data clustering by self-organizing maps. In *Proceedings of the VIII Simpósio Brasileiro de Redes Neurais*, São Luiz, MA, outubro, 2004.
- SILVA FILHO, A.A., CENTENO, J.S. e CANDEIAS, A.L.B., 2003, Incorporação de dados auxiliares para melhorar a classificação da vegetação. In *Proceedings of the III Colóquio Brasileiro de Ciências Geodésicas*, Curitiba, Maio, p. 9p.

- SORIA-FRICH, A. e KÖPPEN, M., 2003, Clustering morphological of the SOM for multi-dimensional image segmentation. *Lecture Notes in Computer Science*, **2686**, 582–589.
- STATHAKIS, D. e VASILAKOS, A., 2006, Comparison of computational intelligence based classification techniques for remotely sensed optical image classification. *IEEE Transactions on Geoscience and Remote Sensing*, **44**, 2305–2318.
- SU, M. e CHOU, C., 2001, A modified version of the K-means algorithm with a distance based on cluster symmetry. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **23**, 674–680.
- SWAIN, P.H. e DAVIS, S.M., 1978, *Remote Sensing: the quantitative approach* (New York: McGraw-Hill).
- TASDEMIR, K. e MERENYI, E., 2007, A new cluster validity index for prototype based clustering algorithms based on inter- and intra-cluster density. In *Proceedings of the International Joint Conference on Neural Networks (IJCNN 2007)*, Orlando, Florida, USA, Agosto, 2007.
- TASDEMIR, K. e MERÉNYI, E., 2005, Considering topology in the clustering of self-Organizing maps. In *Proceedings of the V Workshop On Self-Organizing Maps (WSOM'05)*, Paris, França, Setembro, 2005, pp. 439–446.
- TRAN, T.N., WEHRENS, R. e BUYDENS, L.M.C., 2003, Sparef: a clustering algorithm for multispectral images. *Analytica Chimica Acta*, **490**, 303–312.
- TRAN, T.N., WEHRENS, R. e BUYDENS, L.M.C., 2005, Clustering multispectral images: a tutorial. *Chemometrics and Intelligent Laboratory Systems*, **26**, 2113 – 2133.
- TSO, B. e OLSEN, R.C., 2005, Combining spectral and spatial information into hidden Markov models for unsupervised image classification. *International Journal of Remote Sensing*, **26**, 2113 – 2133.
- TURI, R.H., 2001, “Clustering-based colour image segmentation”, Tese de Doutorado, Monash University, Australia.

- ULSTCH, A., 2003, U\*-matrix: a tool to visualize clusters in high dimensional data. Relatório Técnico 36, Technical Report, Department of Computer Science, University of Marburg.
- ULTSCH, A., 1993a, Knowledge extraction from self-organizing neural networks. *Information and Classification. O. Optiz et al. (Eds). Springer, Berlin*, pp. 301–306.
- ULTSCH, A., 1993b, Self-organizing neural networks for visualization and classification. *Information and Classification. O. Optiz et al. (Eds). Springer, Berlin*, pp. 307–313.
- VAN COILLIE, F.M.B., VERBEKE, L.P.C. e DE WULF, R.R., 2004, Previously trained neural networks as ensemble members: knowledge extraction and transfer. *International Journal of Remote Sensing*, **25**, 4843–4850.
- VEENMAN, C., REINDERS, M. e BACKER, E., 2002, A maximum variance cluster algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **24**, 1273–1280.
- VESANTO, J., 1997, “Data mining techniques based on the Self-Organizing Map”, Tese de mestrado, Helsinki University of Technology.
- VESANTO, J., 1999, SOM-based data visualization methods. *Intelligent Data Analysis*, **3**, 111–126.
- VESANTO, J. e ALHONIEMI, E., 2000, Clustering of the self-organizing map. *IEEE Transactions on Neural Networks*, **11**, 586–602.
- VILMANN, T., MERENYI, E. e HAMMER, B., 2003, Neural maps in remote sensing image analysis. *Neural Networks*, **16**, 389–403.
- VON DER MALSBERG, C., 1973, Self-organization of orientation sensitive cells in the striate cortex. *Kybernetik*, **14**, 85–100.
- VOWER, G.V., SCHEUNDERS, P. e DYCK, D.V., 1999, Statistical texture characterization from discrete wavelets representations. *IEEE Transactions on Image Processing*, **8**, 592–598.



- WAN, W. e FRASER, D., 1994, A Self-organizing neural network framework for unsupervised and supervised classification. In *Proceedings of the VII Australian Remote Sensing Conference*, Melbourne, Austrália.
- WANG, F., 1990, Fuzzy supervised classification of remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, **28**, 194–201.
- WANG, J., DELABIE, J., AASHEIM, H.C. e SMELAND, E., 2002, Clustering of the SOM easily reveals distinct gene expression patterns: results of a reanalysis of lymphoma study. *BMC Bioinformatics*, **3**, 9p.
- WANG, S.J. e ZHOU, C., 2002, Using remote sensing to estimate the change of carbon storage: a case study in the estuary of Yellow River delta. *International Journal of Remote Sensing*, **23**, 1565–1580.
- WEN, G., GUO, X., HUANG, D. e LIU, K., 2007, Application of self-organizing map in aerosol single particles data clustering. In *Proceedings of the International Joint Conference on Neural Networks (IJCNN 2007)*, Florida, Agosto, 2007.
- WILKINSON, G., 2005, Results and implications of a study of fifteen years of satellite image classification experiments. *IEEE Transactions on Geoscience and Remote Sensing*, **43**, 433–440.
- WILLSHAW, D.J. e VON DER MALSBERG, C., 1976, How patterned neural connections can be set up by self-organization. In *Proceedings of the Royal Society of London Series B*, 194, Florida, pp. 431–445.
- WILSON, H.G., BOOTS, B. e MILLWARD, A.A., 2002, A comparison of hierarchical and partitioned clustering techniques for multispectral image classification. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS'02)*, Toronto, Canadá, Junho, 2002, pp. 1624–1626.
- WOODCOCK, C.E., COLLINS, J.B., JAKABHAZY, V.D., LI, X., MACOMBER, A. e WU, Y.,

- 1997, Inversion of the Li-Strahler canopy reflectance model for mapping forest structure. *IEEE Transactions on Geoscience and Remote Sensing*, **35**, 405–414.
- WU, S. e CHOW, T.W.S., 2004, Clustering of the self-organizing map using a clustering validity index based on inter-cluster and intra-cluster density. *Pattern Recognition*, **37**, 175–188.
- XIANG, M., HUNG, C., PHAM, M., KUO, B. e COLEMAN, T., 2005, Parallelepiped multi-spectral image classifier using genetic algorithms. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS'05)*, Seoul, Coréia, Julho, 2005, 4p.
- XU, R. e WUNSCH II, D., 2005, Survey of clustering algorithms. *IEEE Transactions on Neural Networks*, **16**, 645–678.
- YACOUB, M., BADRAN, F. e THIRIA, S., 2001, A topological hierarchical clustering: application to ocean color classification. *Lecture Notes in Computer Science*, **2130**, 492–499.
- YOON, G.W., PARK, J.H. e CHOI, K.H., 2004, Land-cover supervised classification using user-oriented feature database. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS'04)*, September, 2004, pp. 2724–2726.
- ZHANG, Q., WANG, J., GONG, P. e SHI, P., 2003, Study of urban spatial patterns from SPOT panchromatic imagery using textural analysis. *International Journal of Remote Sensing*, **24**, 4137–4160.
- ZHONG, Y., ZHANG, L. e LI, P., 2005, Multispectral remote sensing image classification based on simulated annealing clonal selection. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS'05)*, Seoul, Korea, July, 2005, pp. 3745–3748.
- ZUCHINI, M.H., 2003, “Aplicações de mapas auto-organizáveis em mineração de dados e recuperação de informação”, Tese de mestrado, Departamento de Engenharia de Computação e Automação Industrial, Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas, SP, Brasil.

# Apêndice A

## Mídia com informações adicionais

# A.1 - Exemplo de uso do método de análise de agrupamentos proposto no capítulo 5

Este apêndice mostra a aplicação da estratégia de análise de agrupamentos proposta no capítulo 5 da tese sobre o conjunto de dados apresentado na seção 3.7.1. Através dos resultados apresentados aqui, é possível comparar o desempenho da estratégia proposta com o do método construído por Costa e Netto (1999, 2001) sobre o mesmo conjunto de dados.

Na aplicação da estratégia de análise de agrupamentos sobre o conjunto de dados mostrado na figura 3.6, foi gerado e treinado um SOM de tamanho e parâmetros de treinamento iguais ao daquele empregado como mapa raiz na aplicação do algoritmo Costa e Netto. É possível visualizar através das figuras 3.7 e 3.11 a configuração dos neurônios e a *U-matrix* do mapa após o seu treinamento.

De acordo com a estratégia proposta, uma vez obtida a *U-matrix* do SOM treinado, o próximo passo é obter o conjunto  $S_m$  de imagens de marcadores para a mesma. Para isso, é necessário determinar o conjunto  $S_k$ , o qual deve ser composto por todos os valores de  $k$  que correspondem à inícios de seqüências contíguas e constantes de números de regiões conectadas com tamanhos superiores a 3.

A figura A.1.1 mostra o gráfico do valor de limiar ( $k$ ) da *U-matrix* versus o número de regiões conectadas ( $N_{rc}$ ). O gráfico é igual àquele mostrado na figura 3.14, porém procurou-se destacar os valores de  $k$  que devem pertencer ao conjunto  $S_k$ .

Analisando o gráfico, observa-se que os valores de  $k$  iguais a 14, 21, 57, 64 e 81 corres-

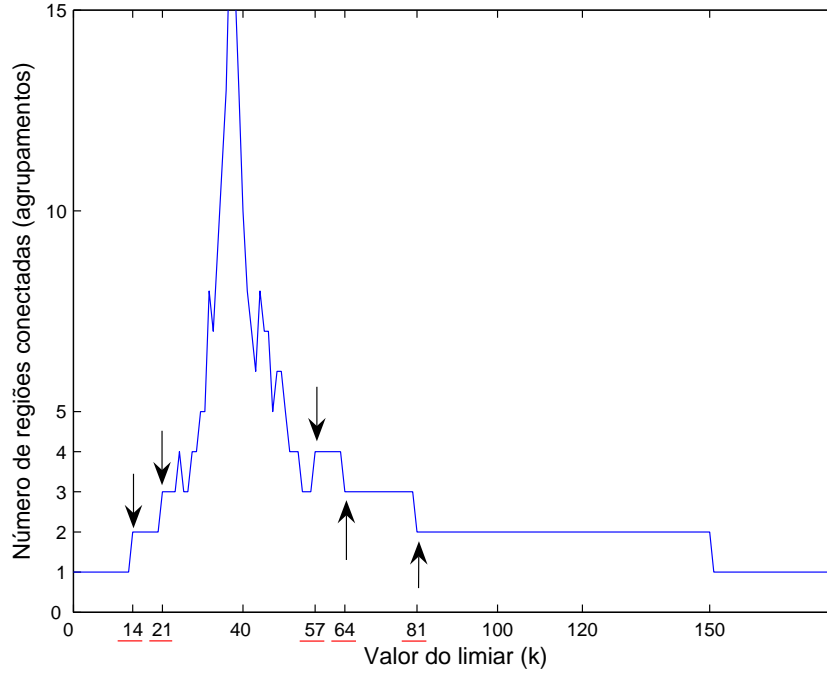


Fig. A.1.1: Gráfico do valor de limiar ( $k$ ) da  $U$ -matrix versus o número de regiões conectadas destacando os valores de  $k$  que pertencem a  $S_k$ .

pondem à inícios de seqüências contíguas e constantes de  $N_{rc}$  (com tamanhos superiores a 3), ou seja,  $S_k = \{14, 21, 57, 64, 81\}$ . Com isso, foi possível gerar então o conjunto  $S_m = \{U_2^{14}, U_2^{21}, U_2^{57}, U_2^{64}, U_2^{81}\}$  composto pelas imagens de marcadores da  $U$ -matrix, cada uma delas correspondendo a um dos cinco valores de  $k$  pertencentes a  $S_k$ .

Em seguida, para cada uma das cinco imagens de marcadores de  $S_m$ , obteve-se as linhas de partição da  $U$ -matrix (através do algoritmo *watershed*), rotulou-se a  $U$ -matrix particionada, copiou-se os rótulos de cada elemento da  $U$ -matrix para os neurônios correspondentes no SOM e rotulou-se o conjunto de dados utilizando como referência os protótipos do SOM já rotulado. A figura A.1.2 mostra os resultados obtidos para cada uma dessas ações.

Tendo produzido cinco partições diferentes para o conjunto de dados da figura 3.6, um índice de validação de agrupamentos foi então aplicado para avaliar a qualidade de cada uma delas, buscando desse modo determinar qual a melhor partição. Nesse experimento, optou-se em utilizar o índice de validação CD<sub>bw</sub> (*Composite Density between and Within Clusters*) proposto por Halkidi e Vazirgiannis (2002). A seção 5.3.3 mostra que este índice de validação

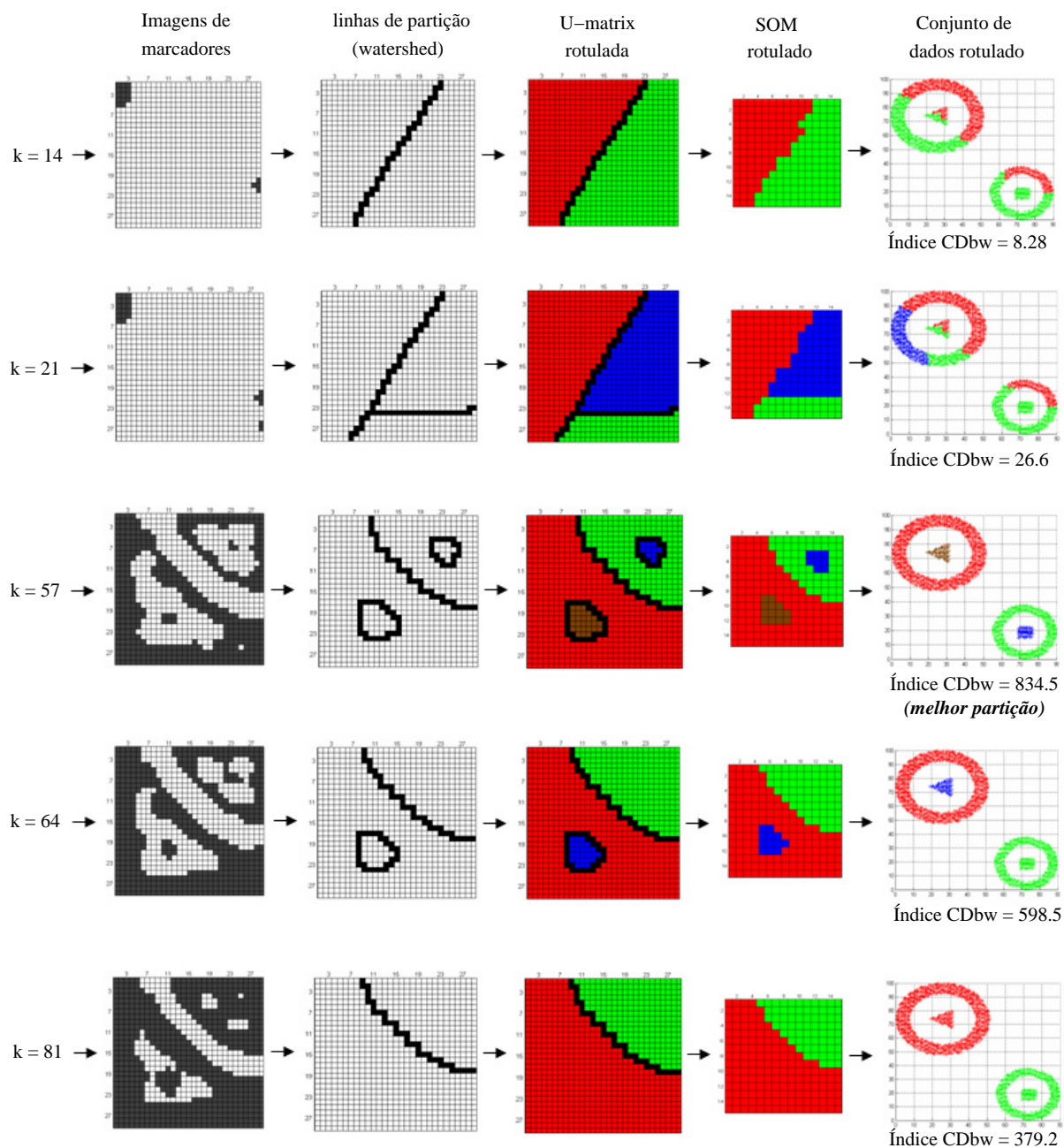


Fig. A.1.2: Diferentes partições para o conjunto de dados da figura 3.6.

apresenta, em comparação com outros índices, um bom desempenho para conjuntos de dados com geometrias complexas e variadas. A tabela A.1.1 mostra o valor do índice para cada uma das partições obtidas (representadas pelo limiar  $k$ ), juntamente com o número de agrupamentos correspondente.

Partição referente ao limiar $k$	Número de agrupamentos	Valor do índice CDbw
14	2	8.28
21	3	26.63
<b>57</b>	<b>4</b>	<b>834.5</b>
64	3	598.5
81	2	379.2

Tab. A.1.1: Valores do índice de validação CDbw para cada uma das partições obtidas para o conjunto de dados da figura 3.6.

Para o índice CDbw, quanto maior o seu valor, melhor é a qualidade da configuração dos agrupamentos avaliada. Sendo assim, de acordo com os valores do índice na tabela A.1.1, a partição referente ao limiar  $k$  igual 57 é considerada a melhor. De fato, se retomarmos a figura A.1.2, observa-se que para o valor de  $k$  igual a 57, o conjunto de dados foi classificado corretamente em 4 classes.

Na seqüência, outros quatro sub-mapas foram gerados a partir dos subconjuntos de dados de cada uma das classes encontradas, mas como não houve detecção de mais de uma região ou agrupamento em cada um deles, os mesmos foram eliminados, e a aplicação do método de análise de agrupamentos foi então finalizada.

Comparando o desempenho da estratégia de análise de agrupamentos proposta com o do algoritmo proposto por Costa e Netto, verifica-se que a estratégia apresentada detectou de maneira mais eficiente as 4 classes existentes no conjunto de dados da figura 3.6. Retomando a aplicação do algoritmo proposto por Costa e Netto (seção 3.7.1.1), observa-se que para o primeiro SOM treinado (o mesmo que foi analisado aqui), a partição escolhida foi aquela referente ao valor de  $k$  igual a 81, que corresponde a dois agrupamentos. Isso ocorre porque o método escolhe a imagem de marcadores da *U-matrix* correspondente ao valor de  $k$  que apresenta a maior seqüência de número de regiões conectadas. Desse modo, o método necessitou gerar e treinar mais dois mapas (ou submapas) para detectar as duas outras classes que não foram descobertas considerando o valor de  $k$  mais estável. Se observarmos novamente a tabela A.1.1, nota-se que a partição referente ao valor de  $k$  igual a 81 apresenta o terceiro maior valor para o índice CDbw, existindo portanto duas partições de melhor qualidade de acordo com o índice utilizado: a partição referente a  $k$  igual a 64, e a melhor de todas, aquela referente a  $k$  igual a

57. Desconsiderando aqueles submapas sobre os quais não foram detectados agrupamentos, ao aplicar o método proposto apresentado neste capítulo foi preciso treinar apenas um único SOM para detectar os 4 agrupamentos existentes no conjunto de dados da figura 3.6, enquanto que na aplicação do método proposto por Costa e Netto foi necessário treinar três SOMs diferentes.



## A.2 - Descrição dos índices de validação de agrupamentos pré-selecionados

Este apêndice mostra uma breve descrição das versões originais dos índices de validação de agrupamentos pré-selecionados através da revisão bibliográfica apresentada na seção 5.3.1. É com base nessa descrição que a estratégia de modificação dos índices de validação (explicada na seção 5.3.2) foi aplicada sobre cada um dos índices pré-selecionados.

- **Índice de Dunn generalizado proposto por Pal e Biswas (1997) ( $\mathcal{V}_{D_{Pal}}$ ):**

Em Pal e Biswas (1997) são propostas generalizações do índice de Dunn baseadas em três conceitos da teoria dos grafos: Árvore Geradora Mínima, Grafo de Vizinhança Relativa e Grafo de Gabriel. Dentre as três generalizações propostas, aquela que utiliza informações com base no Grafo de Gabriel apresentou um melhor desempenho para encontrar o número desejado de classes para todos os conjuntos de dados testados.

Na definição do Grafo de Gabriel, dois pontos  $x_i$  e  $x_j$  estão conectados se:

$$d^2(x_i, x_j) < d^2(x_i, x_k) + d^2(x_j, x_k), \forall k, k \neq i, k \neq j, \quad (1)$$

onde  $d(x_i, x_j)$  é a distância euclidiana entre  $x_i$  e  $x_j$ . Em outras palavras,  $x_i$  e  $x_j$  estão conectados no Grafo de Gabriel se nenhum ponto pertence ao disco que possui diâmetro  $d(x, y)$  e centro no ponto médio de  $x_i$  e  $x_j$ .

Supondo que um conjunto de dados  $X \subset \mathfrak{R}^n$  tenha sido particionado em  $K$  classes,

$C_1, C_2, \dots, C_k$ , define-se:

$$E^{GG}(C_i) = \{e_{ij}^{GG}, 1 \leq j \leq l_i, \text{ com } l_i = |E_i|; 1 \leq i \leq K\} \quad (2)$$

como o conjunto das arestas do grafo de Gabriel determinadas sobre o agrupamento  $C_i$ .

O diâmetro  $\Delta^{GG}$  do agrupamento  $C_i$  é dado como segue:

$$\Delta^{GG}(C_i) = \max_j \{e_i^{GG}, j = 1, 2, \dots, l_i\} \quad (3)$$

A separação  $\delta^{GG}$  entre dois agrupamentos  $C_i$  e  $C_j$  é dada pela distância entre os centros  $m_i$  e  $m_j$  das classes  $i$  e  $j$ :

$$\delta^{GG}(C_i, C_j) = \|m_i - m_j\|. \quad (4)$$

Portanto, o índice de Dunn generalizado com base no Grafo de Gabriel é definido como segue:

$$\mathcal{V}_{D_{Pal}} = \min_{1 \leq i \leq K} \left\{ \min_{1 \leq j \leq K, j \neq i} \left\{ \frac{\delta^{GG}(C_i, C_j)}{\max_{1 \leq k \leq K} \{\Delta^{GG}(C_k)\}} \right\} \right\}, \quad (5)$$

Uma boa partição é indicada por valores altos de  $\mathcal{V}_{D_{Pal}}$ .

- **Índice de Dunn generalizado proposto por Bezdek e Pal (1998) ( $\mathcal{V}_{D_{Bdk}}$ ):**

Em Bezdek e Pal (1998), os autores propuseram diversas generalizações para o índice de Dunn utilizando cinco funções  $\delta$  diferentes para medir a distância entre os agrupamentos e três funções  $\Delta$  para medir o diâmetro dos mesmos. Dentre as diferentes combinações de  $\delta$  e  $\Delta$ , a combinação das funções  $\delta_3$  e  $\Delta_3$  está entre aquelas que apresentaram os melhores resultados.  $\delta_3$  e  $\Delta_3$  são definidas como seguem:

$$\delta_3(C_i, C_j) = \frac{1}{|C_i||C_j|} \sum_{x \in C_i, y \in C_j} d(x, y) \quad (6)$$

e

$$\Delta_3(C_k) = 2 \left( \frac{\sum_{x \in C_k} d(x, m_k)}{|C_k|} \right), \quad (7)$$

onde  $C_i$  representa um agrupamento  $i$  e  $|C_i|$  a sua quantidade de pontos,  $d$  é uma medida de distância, e  $m_k = \frac{1}{|C_k|} \sum_{x \in C_k} x$ . E assim tem-se o índice de Dunn generalizado:

$$\mathcal{V}_{D_{Bdk}} = \min_{1 \leq i \leq K} \left\{ \min_{1 \leq j \leq K, j \neq i} \left\{ \frac{\delta_3(C_i, C_j)}{\max_{1 \leq k \leq K} \{\Delta_3(C_k)\}} \right\} \right\}, \quad (8)$$

sendo  $K$  o número total de agrupamentos.

Quanto maior o valor do índice  $\mathcal{V}_{D_{Bdk}}$ , melhor é a qualidade do particionamento de dados avaliado.

• **Índice de Davies-Bouldin generalizado proposto por Pal e Biswas (1997) ( $\mathcal{V}_{DB_{Pal}}$ ):**

Em Pal e Biswas (1997), também são apresentadas três generalizações do índice Davies-Bouldin usando as três estruturas de grafos empregadas para generalizar o índice de Dunn. Uma vez que as três generalizações apresentaram resultados bastante parecidos, optou-se em utilizar neste trabalho aquela que se baseia também no Grafo de Gabriel (da mesma forma que no índice  $\mathcal{V}_{D_{Pal}}$ ).

$\Delta^{GG}(C_i)$ , o máximo dos tamanhos das arestas de Gabriel de um agrupamento  $C_i$ , definida na equação 3 como sendo o diâmetro de  $C_i$ , é definida aqui como uma medida de dispersão interna da classe  $C_i$ . E a medida de similaridade ( $R^{GG}$ ) entre dois agrupamentos  $C_i$  e  $C_j$  é dada como segue:

$$R^{GG}(C_i, C_j) = \frac{\Delta^{GG}(C_i) + \Delta^{GG}(C_j)}{d_{ij}}, \quad (9)$$

onde  $d_{ij} = \|m_i - m_j\|$  é a distância entre os centros  $m_i$  e  $m_j$  dos dois agrupamentos  $C_i$  e  $C_j$ .

A generalização do índice Davies-Bouldin baseado no Grafo de Gabriel é, portanto,

definida da seguinte forma:

$$\mathcal{V}_{DBPal} = \frac{1}{K} \sum_{i=1}^K \max_{j, j \neq i} \{R^{GG}(C_i, C_j)\}, \quad (10)$$

onde  $K$  é o número de agrupamentos da partição.

Quanto menor o valor de  $\mathcal{V}_{DBPal}$ , melhor é a partição de dados avaliada.

- **Índice de Davies-Bouldin generalizado proposto por Kim e Ramakrishna (2005)** ( $\mathcal{V}_{DBKim}$ ):

Em Kim e Ramakrishna (2005), os autores propuseram uma modificação na fórmula geral do índice Davies-Bouldin dada pela equação abaixo:

$$\mathcal{V}_{DBKim} = \frac{1}{K} \sum_{i=1}^K \left( \frac{\max_{k=1, \dots, K, k \neq i} \{S(C_i) + S(C_k)\}}{\min_{l=1, \dots, K, l \neq i} \{d_{il}\}} \right), \quad (11)$$

onde  $K$  é o número total de agrupamentos,  $S(C_i)$  é a dispersão de um agrupamento  $C_i$  dada pela equação 12, e  $d_{ij} = d(m_i, m_j)$  é a medida de separação entre dois agrupamentos  $C_i$  e  $C_j$ , definidas conforme a versão original do índice.  $m_i$  e  $m_j$  são os centróides dos agrupamentos de dados  $C_i$  e  $C_j$ , respectivamente.

$$S(C_i) = \frac{1}{|C_i|} \sum_{x \in C_i} d(x, m_i). \quad (12)$$

Quanto menor o valor de  $\mathcal{V}_{DBKim}$ , melhor é a qualidade da partição de dados avaliada.

- **Índice PBM proposto por Pakhira *et al.*(2004)** ( $\mathcal{V}_{PBM}$ ):

A fórmula geral do índice de validação de agrupamentos PBM é definida da seguinte forma:

$$\mathcal{V}_{PBM} = \left( \frac{1}{K} \cdot \frac{E_0}{E_k} \cdot D_k \right), \quad (13)$$

onde  $K$  é o número de agrupamentos de um conjunto de dados  $X$ .  $E_0$  é a soma das

distâncias de cada ponto  $x_i$  ao centro geométrico  $m$  de  $X$ :

$$E_0 = \sum_{x_i \in X} \|x_i - m\|. \quad (14)$$

$E_k$  é a soma das distâncias intra-agrupamentos de cada um dos  $K$  agrupamentos pertencentes ao conjunto de dados  $X$ :

$$E_k = \sum_{k=1}^K \sum_{i=1}^{n_k} \|x_i - m_k\|, \quad (15)$$

onde  $n_k$  e  $m_k$  são, respectivamente, o número de pontos e o centróide do agrupamento  $C_k$ . E finalmente,  $D_k$  representa o maior valor para a medida de separação calculada entre os  $K$  agrupamentos:

$$D_k = \max_{i,j=1,\dots,K} \|m_i - m_j\|. \quad (16)$$

Quanto maior o valor de  $\mathcal{V}_{PBM}$ , melhor é a configuração de agrupamentos analisada.

- **Índice CDbw proposto por Halkidi e Vazirgiannis (2002) ( $\mathcal{V}_{CDbw}$ ):**

Seja  $X$  um conjunto de dados particionado em  $K$  agrupamentos e  $V_i = \{v_{i1}, v_{i2}, \dots, v_{ir_i}\}$  um conjunto de pontos representativos para o agrupamento  $C_i$ . A densidade intra-agrupamento é definida por:

$$Intra = \frac{1}{K} \sum_{i=1}^K \frac{1}{r_i} \sum_{j=1}^{r_i} density(\bar{v}_{ij})/stdev, \quad (17)$$

com  $density(\bar{v}_{ij}) = \sum_{l=1}^{n_i} f(x_l, \bar{v}_{ij})$ , onde  $x_l$  pertence ao agrupamento  $C_i$ ,  $\bar{v}_{ij}$  é a  $j$ -ésima representação do agrupamento  $C_i$  deslocada para o centro do agrupamento, e  $f(x_l, \bar{v}_{ij})$  é igual a 1, se  $\|x_l - \bar{v}_{ij}\| \leq stdev$ , ou 0, caso contrário.  $stdev$  corresponde à média dos desvios padrões de cada um dos  $K$  agrupamentos.

A densidade inter-agrupamento é dada por:

$$Inter = \sum_{i=1}^K \sum_{j=1, j \neq i}^K \frac{\|close\_rep(C_i) - close\_rep(C_j)\|}{\|stdev(C_i)\| + \|stdev(C_j)\|} \cdot density(u_{ij}), \quad (18)$$

onde  $close\_rep(C_i)$  e  $close\_rep(C_j)$  representam o par de pontos de representação mais próximos entre os agrupamentos de dados  $C_i$  e  $C_j$ ,  $u_{ij}$  é o ponto médio entre este par de pontos,  $stdev(C_i)$  e  $stdev(C_j)$  representam os desvios padrões dos agrupamentos  $C_i$  e  $C_j$ , e  $density(u_{ij}) = \sum_{k=1}^{n_i+n_j} f(x_k, u_{ij}) / (n_i + n_j)$ , onde  $x_k$  pertence ao agrupamento  $C_i$  ou  $C_j$ , e  $f(x_k, u_{ij})$  é igual a 1, se  $\|x_k - u_{ij}\| \leq (stdev(C_i) + stdev(C_j))/2$ , ou igual a zero, caso contrário.

A separação entre os agrupamentos é dada por:

$$Sep = \frac{\sum_{i=1}^K \sum_{j=1}^K \|close\_rep(C_i) - close\_rep(C_j)\|}{1 + Inter}. \quad (19)$$

O índice CDbw é dado então por:

$$\mathcal{V}_{CDbw} = Intra * Sep. \quad (20)$$

De acordo com Halkidi e Vazirgiannis (2002), uma boa partição de dados é indicada por valores altos de  $\mathcal{V}_{CDbw}$ .

- **Índice Conn proposto por Tasdemir e Merenyi (2007) ( $\mathcal{V}_{Conn}$ ):**

Seja  $CADJ$  uma matriz  $N \times N$ , onde  $N$  é o número de protótipos do conjunto de dados  $X$  particionado em  $K$  agrupamentos. De acordo com Tasdemir e Merenyi (2007), os pontos representativos ou protótipos de  $X$  podem ser determinados por algum método de agrupamentos, como por exemplo, o SOM.

A adjacência cumulativa,  $CADJ(i, j)$ , de dois protótipos  $v_i$  e  $v_j$ , é o número de vetores de dados para qual  $v_i$  é o protótipo mais próximo (BMU - *best matching unit*) e  $v_j$  o segundo mais próximo (segundo BMU). O nível de conectividade (similaridade) de dois

protótipos  $v_i$  e  $v_j$  é dado por:

$$CONN(i, j) = CADJ(i, j) + CAD(j, i). \quad (21)$$

A compacidade de um agrupamento  $C_k$ ,  $Intra\_Conn(C_k)$ , é a razão do número de vetores de dados em  $C_k$  que tem ambos os seus BMU e segundo BMU em  $C_k$ , pelo número de vetores de dados em  $C_k$ :

$$Intra\_Conn(C_k) = \frac{\sum_{i,j}^N \{CADJ(i, j) : v_i, v_j \in C_k\}}{\sum_{i,j}^N \{CADJ(i, j) : v_i \in C_k\}} \quad (22)$$

A conectividade intra-agrupamento de todos os  $K$  agrupamentos,  $Intra\_Conn$ , é a compacidade média, ou seja:

$$Intra\_Conn = \sum_{k=1}^K Intra\_Conn(C_k) / K. \quad (23)$$

A conectividade inter-agrupamento entre dois agrupamentos  $C_k$  e  $C_l$ ,  $Inter\_Conn(C_k, C_l)$ , é definida como a razão da conectividade entre  $C_k$  e  $C_l$  pela conectividade total dos protótipos em  $C_k$  que tem pelo menos uma conexão com um protótipo em  $C_l$ :

$$Inter\_Conn(C_k, C_l) = \frac{Conn(C_k, C_l)}{\sum_{i,j}^N \{CONN(i, j) : v_i \in V_{k,l}\}}, \quad (24)$$

onde:

$$Conn(C_k, C_l) = \sum_{i,j}^N \{CONN(i, j) : v_i \in C_k, v_j \in C_l\} \quad (25)$$

e

$$V_{k,l} = \{v_i : v_i \in C_k, \exists v_j \in C_l : CADJ(i, j) > 0\}. \quad (26)$$

A inter-conectividade de  $C_k$  para todos os outros agrupamentos,  $Inter\_Conn(C_k)$ , e a

conectividade (similaridade) média dos agrupamentos,  $Inter\_Conn$ , são definidas como seguem:

$$Inter\_Conn(C_k) = \max_{l, l \leq K} \{Inter\_Conn(C_k, C_l)\}. \quad (27)$$

e

$$Inter\_Conn = \sum_{k=1}^K Inter\_Conn(C_k) / K. \quad (28)$$

Assim,  $1 - Inter\_Conn$  é tomada como a medida de separação entre os agrupamentos. E finalmente, a fórmula geral do índice Conn é dada por:

$$\mathcal{V}_{Conn} = Intra\_Conn \times (1 - Inter\_Conn). \quad (29)$$

Quanto maior o valor de  $\mathcal{V}_{Conn}$ , melhor é a qualidade da partição avaliada.



## A.3 - Versões modificadas dos índices de validação de agrupamentos

Este apêndice descreve as versões modificadas dos índices de validação de agrupamentos pré-selecionados na seção 5.3.1, usando a estratégia de modificação proposta na seção 5.3.2.

Nas modificações dos cálculos dos índices descritas a seguir, considere  $X$  como um conjunto de dados particionado em  $K$  agrupamentos, os quais são determinados através da partição de um SOM treinado com os dados de  $X$ . Considere também  $W_i = \{w_i^1, w_i^2, \dots, w_i^{|W_i|}\}$  como o conjunto de protótipos do SOM que representam o agrupamento de dados  $C_i$  em  $X$  ( $i = 1, 2, \dots, K$ ). Cada protótipo  $w_i^j$  é dado então pelo vetor de pesos sinápticos do neurônio  $j$  pertencente a  $W_i$ . Nas propostas de modificação dos cálculos de cada um dos índices, o centróide  $m_i$  de um agrupamento de dados  $C_i$  é definido como segue:

$$m_i = \frac{1}{|C_i|} \sum_{w_i^j \in W_i} h(w_i^j) \cdot w_i^j, \quad (1)$$

onde  $h(w_i^j)$  é o grau de atividade do protótipo  $w_i^j$  e  $|C_i|$  é o número de pontos do agrupamento de dados  $C_i$ .

- **Índice  $\mathcal{V}_{D_{Pal}}$  modificado ( $\mathcal{V}_{D_{Pal}}^{SOM}$ ):**

Na modificação dos cálculos do índice  $\mathcal{V}_{D_{Pal}}$ , a distância  $\delta^{GG}$  entre dois agrupamentos de dados  $C_i$  e  $C_j$  é calculada em função dos conjuntos de protótipos do SOM,  $W_i$  e  $W_j$ :

$$\delta^{GG}(W_i, W_j) = \|m_i - m_j\|, \quad (2)$$

sendo  $m_i$  e  $m_j$  os centróides de  $C_i$  e  $C_j$  respectivamente, calculados de acordo com a fórmula dada na equação 1.

O diâmetro  $\Delta^{GG}$  de um agrupamento  $C_k$ , é dado pela maior aresta de Gabriel num grafo no qual os vértices são os protótipos do SOM pertencentes a  $W_k$  (que não apresentam atividade nula):

$$\Delta^{GG}(W_k) = \max_j \{e_i^{GG}, j = 1, 2, \dots, l_k\} \quad (3)$$

Sendo assim, para qualquer partição de dados, o índice  $\mathcal{V}_{D_{Pal}}$ , na sua forma modificada, fica definido da seguinte forma:

$$\mathcal{V}_{D_{Pal}}^{SOM} = \min_{1 \leq i \leq K} \left\{ \min_{1 \leq j \leq K, j \neq i} \left\{ \frac{\delta^{GG}(W_i, W_j)}{\max_{1 \leq k \leq K} \{\Delta^{GG}(W_k)\}} \right\} \right\}, \quad (4)$$

• **Índice  $\mathcal{V}_{D_{Bdk}}$  modificado ( $\mathcal{V}_{D_{Bdk}}^{SOM}$ ):**

Modificando a computação do índice  $\mathcal{V}_{D_{Bdk}}$ , a distância  $\delta_3$  entre dois agrupamentos de dados  $C_i$  e  $C_j$  e o diâmetro  $\Delta_3$  de um agrupamento  $C_k$  ficam definidos conforme seguem:

$$\delta_3(W_i, W_j) = \frac{1}{|C_i| |C_j|} \sum_{w_i^p \in W_i, w_j^q \in W_j} h(w_i^p) \cdot h(w_j^q) \cdot d(w_i^p, w_j^q) \quad (5)$$

e

$$\Delta_3(W_k) = 2 \left( \frac{\sum_{w_k^p \in W_k} h(w_k^p) \cdot d(w_k^p, m_k)}{|C_k|} \right), \quad (6)$$

onde  $W_i$  e  $W_j$  são os conjuntos de protótipos do SOM que representam, respectivamente, os agrupamentos de dados  $C_i$  e  $C_j$ ,  $h(w_i^p)$  representa o grau de atividade do protótipo  $w_i^p$  e  $m_k$  é calculado conforme a equação 1.

A fórmula geral do índice  $\mathcal{V}_{DBdk}$  modificado é dada como:

$$\mathcal{V}_{DBdk}^{SOM} = \min_{1 \leq i \leq K} \left\{ \min_{1 \leq j \leq K, j \neq i} \left\{ \frac{\delta_3(W_i, W_j)}{\max_{1 \leq k \leq K} \{\Delta_3(W_k)\}} \right\} \right\}. \quad (7)$$

• **Índice  $\mathcal{V}_{DBPal}$  modificado ( $\mathcal{V}_{DBPal}^{SOM}$ ):**

Na modificação dos cálculos do índice  $\mathcal{V}_{DBPal}$ , a medida de dispersão interna ( $\Delta^{GG}$ ) de um agrupamento  $C_i$  é calculada da mesma forma que na equação 3 do índice  $\mathcal{V}_{DBPal}$  modificado. Já a medida de similaridade ( $R^{GG}$ ) entre dois agrupamentos de dados  $C_i$  e  $C_j$  é definida como segue:

$$R^{GG}(W_i, W_j) = \frac{\Delta^{GG}(W_i) + \Delta^{GG}(W_j)}{d_{ij}}, \quad (8)$$

onde  $W_i$  e  $W_j$  são os conjuntos de protótipos do SOM que representam os agrupamentos de dados  $C_i$  e  $C_j$ , respectivamente.  $d_{ij} = \|m_i - m_j\|$  é a distância entre os centros  $m_i$  e  $m_j$  dos dois agrupamentos  $C_i$  e  $C_j$ , calculados de maneira modificada de acordo com a equação 1.

O índice  $\mathcal{V}_{DBPal}$  modificado é, portanto, definido da seguinte maneira:

$$\mathcal{V}_{DBPal}^{SOM} = \frac{1}{K} \sum_{i=1}^K \max_{j, j \neq i} \{R^{GG}(W_i, W_j)\}. \quad (9)$$

• **Índice  $\mathcal{V}_{DBKim}$  modificado ( $\mathcal{V}_{DBKim}^{SOM}$ ):**

Na versão modificada do índice  $\mathcal{V}_{DBKim}$ , a dispersão de um agrupamento  $C_i$  é definida da seguinte forma:

$$S(W_i) = \frac{1}{|C_i|} \sum_{w_i^p \in W_i} h(w_i^p) \cdot d(w_i^p, m_i), \quad (10)$$

onde  $W_i$  é o conjunto de protótipos do SOM que representa o agrupamento de dados  $C_i$ ,  $h(w_i^p)$  representa o grau de atividade do protótipo  $w_i^p$  e  $m_i$  é o centróide de  $C_i$  calculado conforme a equação 1.

A fórmula geral do índice  $\mathcal{V}_{DB_{Kim}}$  modificado é dada como segue:

$$\mathcal{V}_{DB_{Kim}}^{SOM} = \frac{1}{K} \sum_{i=1}^K \left( \frac{\max_{k=1, \dots, K, k \neq i} \{S(W_i) + S(W_k)\}}{\min_{l=1, \dots, K, l \neq i} \{d_{il}\}} \right), \quad (11)$$

na qual  $d_{ij} = d(m_i, m_j)$  é a medida de separação entre dois agrupamentos  $C_i$  e  $C_j$ , e  $m_i$  e  $m_j$  são os centróides dos agrupamentos de dados  $C_i$  e  $C_j$ , respectivamente (calculados de acordo com a equação 1).

• **Índice  $\mathcal{V}_{PBM}$  modificado ( $\mathcal{V}_{PBM}^{SOM}$ ):**

Na modificação dos cálculos do índice  $\mathcal{V}_{PBM}$ ,  $E_0$  é dado por:

$$E_0(W) = \sum_{w^j \in W} h(w^j) \cdot \|w^j - m\|, \quad (12)$$

onde  $W$  é o conjunto de todos os protótipos do SOM que representam o conjunto de dados  $X$ ,  $h(w^j)$  representa o grau de atividade de um protótipo  $w^j \in W$  e  $m$  é o centróide do conjunto de dados  $X$  calculado como segue:

$$m = \frac{1}{|X|} \sum_{w^j \in W} h(w^j) \cdot w^j, \quad (13)$$

onde  $|X|$  é quantidade de pontos do conjunto de dados  $X$ .

$E_k$  é definido como:

$$E_k(W) = \sum_{k=1}^K \sum_{j=1}^{|W_k|} h(w_k^j) \cdot \|w_k^j - m_k\|, \quad (14)$$

onde  $|W_k|$  é o número de protótipos do SOM que representam o agrupamento de dados  $C_k$ ,  $h(w_k^j)$  é o grau de atividade do protótipo  $w_k^j$  que pertence a  $W_k$  e  $m_k$  é o centróide de  $C_k$  calculado conforme a equação 1.

A máxima separação entre os  $K$  agrupamentos de dados é definida como:

$$D_k(W) = \max_{i,j=1, \dots, K} \|m_i - m_j\|, \quad (15)$$

onde  $m_i$  e  $m_j$  são os centróides dos agrupamentos de dados  $C_i$  e  $C_j$  (calculados de acordo com a equação 1).

O índice  $\mathcal{V}_{PBM}$  modificado é dado então por:

$$\mathcal{V}_{PBM}^{SOM} = \left( \frac{1}{K} \cdot \frac{E_0(W)}{E_k(W)} \cdot D_k(W) \right), \quad (16)$$

• **Índice  $\mathcal{V}_{CDbw}$  modificado ( $\mathcal{V}_{CDbw}^{SOM}$ ):**

Modificando os cálculos do índice  $\mathcal{V}_{CDbw}$ , a densidade intra-agrupamento fica definida como segue:

$$Intra(W) = \frac{1}{K} \sum_{i=1}^K \frac{1}{|W_i|} \sum_{j=1}^{|W_i|} density(\bar{w}_i^j) / stdev, \quad (17)$$

onde  $W_i$  é o número de protótipos do SOM que representam o agrupamento de dados  $C_i$ ,  $\bar{w}_i^j$  corresponde ao protótipo  $w_i^j$  levemente deslocado para o centróide do agrupamento  $C_i$  e:

$$density(\bar{w}_i^j) = \sum_{w_i^j \in W_i} h(w_i^j) \cdot f(w_i^j, \bar{w}_i^j), \quad (18)$$

onde  $f(w_i^j, \bar{w}_i^j)$  é igual a 1 se  $\|w_i^j - \bar{w}_i^j\| \leq stdev$ , ou 0 caso contrário.  $stdev$  corresponde à média dos desvios padrões de cada um dos  $K$  agrupamentos. O desvio padrão de um agrupamento de dados  $C_i$  é calculado como segue:

$$stdev(W_i) = \sqrt{\frac{\sum_{w_i^j \in W_i} h(w_i^j) \cdot (w_i^j - m_i)}{|C_i| - 1}} \quad (19)$$

onde  $W_i$  é o conjunto de protótipos do SOM que representam  $C_i$ ,  $h(w_i^j)$  é o grau de atividade do protótipo  $w_i^j$  e  $m_i$  é o centróide do agrupamento  $C_i$  (calculado de acordo com a equação 1).

A densidade inter-agrupamento é dada como segue:

$$Inter(W) = \sum_{i=1}^K \sum_{j=1, j \neq i}^K \frac{\|close\_rep(W_i) - close\_rep(W_j)\|}{\|stdev(W_i)\| + \|stdev(W_j)\|} \cdot density(u_{ij}), \quad (20)$$

onde  $close\_rep(W_i)$  e  $close\_rep(W_j)$  representam o par de protótipos do SOM mais próximos que representam os agrupamentos de dados  $C_i$  e  $C_j$ , respectivamente.  $u_{ij}$  é o ponto médio entre este par de pontos.  $stdev(W_i)$  e  $stdev(W_j)$  são calculados conforme a equação 19 e representam os desvios padrões dos agrupamentos  $C_i$  e  $C_j$ .  $density(u_{ij}) = \sum_{k=1}^{|C_i|+|C_j|} h(w_{ij}^k) \cdot f(w_{ij}^k, u_{ij}) / (|C_i| + |C_j|)$ , onde  $w_{ij}^k$  pertence ao conjunto de protótipos  $W_i$  ou  $W_j$ , e  $f(w_{ij}^k, u_{ij})$  é igual a 1, se  $\|w_{ij}^k - u_{ij}\| \leq (stdev(W_i) + stdev(W_j))/2$ , ou igual a zero, caso contrário.

A separação entre os agrupamentos fica definida como:

$$Sep(W) = \sum_{i=1}^K \sum_{j=1}^K \frac{\|close\_rep(W_i) - close\_rep(W_j)\|}{1 + Inter}. \quad (21)$$

E o índice  $\mathcal{V}_{CDbw}$  modificado é dado então por:

$$\mathcal{V}_{CDbw}^{SOM} = Intra(W) * Sep(W). \quad (22)$$

# A.5 - Tabelas e gráficos dos experimentos realizados com índices de validação de agrupamentos

Este apêndice tem como objetivo mostrar mais detalhadamente os resultados obtidos nos experimentos realizados com os setes índices de validação de agrupamentos pré-selecionados na seção 5.3.1. Para cada um dos conjuntos de dados utilizados nos testes (descritos na seção 5.3.3) são apresentadas tabelas e gráficos com os valores e tempos de processamento das versões originais e modificadas dos índices de validação.

O apêndice apresenta-se dividido em 10 subseções. Cada uma delas corresponde a um conjunto de dados considerado nos experimentos. As tabelas dentro de cada subseção apresentam os resultados obtidos para cada índice de validação. A primeira coluna das tabelas mostra as diferentes partições para o conjunto de dados considerado. Aquela marcada com asterisco representa a melhor partição de acordo com o índice aplicado. Essas partições foram determinadas aplicando a estratégia de análise de agrupamentos proposta na seção 5.2, e são representadas pelos valores do limiar  $k$  que definem as imagens de marcadores para a segmentação da *U-matrix* e do SOM correspondente. A segunda coluna das tabelas mostra o número de classes (agrupamentos) existentes em cada partição. A terceira e quarta coluna mostram, respectivamente, os valores obtidos pela versão original e modificada do índice de validação (aqueles marcados em negrito representam as melhores partições de acordo com o índice). A quinta coluna apresenta o erro absoluto entre esses valores. Os tempos de processamento das versões original e modificada do índice são apresentados (em segundos) na sexta e sétima co-

lunas, respectivamente. A última coluna mostra o *speedup* apresentado pela versão modificada do índice em termos de tempo de processamento.

Adicionalmente, para cada conjunto de dados, os valores das duas versões dos índices (original e modificada) são apresentados em forma de gráficos para facilitar a comparação entre os mesmos. Vale ressaltar também que, particularmente para o índice Conn ( $\mathcal{V}_{Conn}$ ), são apresentados os valores e os tempos de processamento apenas de sua versão original, tendo em vista que não há proposta de modificação para os cálculos deste índice por razões já explicadas na seção 5.3.2.

Conforme mencionado na seção 5.3.3, os experimentos foram realizados no software MATLAB em uma máquina com processador AMD Athlon XP 2600+ 1,91GHZ e 1GB de RAM.

## Conjunto de dados *Gauss*\_1

Partições	No. de classes	Valor $\mathcal{V}_{DPal}$	Valor $\mathcal{V}_{DPal}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DPal}$	Tempo(s) $\mathcal{V}_{DPal}^{SOM}$	Speedup
k=2	2	1.4454	1.7661	0.3207	20.372	0.0902	225.8
k=20*	5	<b>2.105</b>	<b>2.6688</b>	0.5637	5.227	0.0415	125.9
k=143	3	1.7853	2.2142	0.4289	21.366	0.0593	360.3
k=151	2	1.5589	2.0175	0.4586	47.335	0.0935	506.25
				$\mu = 0.443$	$\Sigma = 94.32$	$\Sigma = 0.2846$	331.4

Tabela A.5.1 - Valores e tempos de processamento de  $\mathcal{V}_{DPal}$  e  $\mathcal{V}_{DPal}^{SOM}$  para as partições do conjunto de dados *Gauss*\_1.

Partições	No. de classes	Valor $\mathcal{V}_{DBdk}$	Valor $\mathcal{V}_{DBdk}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DBdk}$	Tempo(s) $\mathcal{V}_{DBdk}^{SOM}$	Speedup
k=2	2	1.0434	1.0454	0.0019	1.1466	0.0560	20.5
k=20*	5	<b>2.4876</b>	<b>2.6761</b>	0.1885	1.6176	0.0598	27.0
k=143	3	1.065	1.064	0.0010	1.1695	0.0433	27.0
k=151	2	0.9407	0.9431	0.0023	0.6838	0.0371	18.4
				$\mu = 0.0484$	$\Sigma = 4.6175$	$\Sigma = 0.1963$	23.52

Tabela A.5.2 - Valores e tempos de processamento de  $\mathcal{V}_{DBdk}$  e  $\mathcal{V}_{DBdk}^{SOM}$  para as partições do conjunto de dados *Gauss*\_1.



Partições	No. de classes	Valor $\mathcal{V}_{DBPal}$	Valor $\mathcal{V}_{DBPal}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DBPal}$	Tempo(s) $\mathcal{V}_{DBPal}^{SOM}$	Speedup
k=2	2	1.3391	1.1013	0.2378	21.148	0.0808	261.7
k=20*	5	<b>0.5487</b>	<b>0.4853</b>	0.0633	3.142	0.0498	63.09
k=143	3	0.7866	0.6443	0.1422	21.712	0.06826	318.07
k=151	2	0.9749	0.7438	0.2310	49.16	0.08363	587.82
				$\mu = 0.1686$	$\Sigma = 95.16$	$\Sigma = 0.2826$	336.73

Tabela A.5.3 - Valores e tempos de processamento de  $\mathcal{V}_{DBPal}$  e  $\mathcal{V}_{DBPal}^{SOM}$  para as partições do conjunto de dados *Gauss\_1*.

Partições	No. de classes	Valor $\mathcal{V}_{DBKim}$	Valor $\mathcal{V}_{DBKim}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DBKim}$	Tempo(s) $\mathcal{V}_{DBKim}^{SOM}$	Speedup
k=2	2	5.8546	5.3362	0.5184	1.0307	0.0251	41.0
k=20*	5	<b>1.3984</b>	<b>1.1263</b>	0.2721	1.0284	0.02642	38.9
k=143	3	3.4707	3.3057	0.165	1.0271	0.02562	40.0
k=151	2	3.9757	3.7797	0.1959	1.0292	0.0252	40.8
				$\mu = 0.2878$	$\Sigma = 4.1155$	$\Sigma = 0.1025$	40.15

Tabela A.5.4 - Valores e tempos de processamento de  $\mathcal{V}_{DBKim}$  e  $\mathcal{V}_{DBKim}^{SOM}$  para as partições do conjunto de dados *Gauss\_1*.

Partições	No. de classes	Valor $\mathcal{V}_{PBM}$	Valor $\mathcal{V}_{PBM}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{PBM}$	Tempo(s) $\mathcal{V}_{PBM}^{SOM}$	Speedup
k=2	2	2.966	2.9544	0.0115	1.0559	0.0435	24.3
k=20*	5	<b>6.5897</b>	<b>5.4635</b>	1.1262	1.0319	0.0266	38.8
k=143	3	3.6328	3.5234	0.1093	1.0322	0.02559	40.3
k=151	2	3.379	3.4678	0.0887	1.0342	0.0265	39.0
				$\mu = 0.334$	$\Sigma = 4.1543$	$\Sigma = 0.1223$	33.96

Tabela A.5.5 - Valores e tempos de processamento de  $\mathcal{V}_{PBM}$  e  $\mathcal{V}_{PBM}^{SOM}$  para as partições do conjunto de dados *Gauss\_1*.

Partições	No. de classes	Valor $\mathcal{V}_{CDbw}$	Valor $\mathcal{V}_{CDbw}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{CDbw}$	Tempo(s) $\mathcal{V}_{CDbw}^{SOM}$	Speedup
k=2	2	59.549	64.354	4.8047	3.1281	0.1716	18.2
k=20*	5	<b>1149.9</b>	<b>1959.4</b>	809.59	2.1654	0.1113	19.4
k=143	3	869.77	993.39	123.62	2.8191	0.1070	26.3
k=151	2	636.23	241.45	394.78	3.5022	0.1122	31;2
				$\mu = 333.2$	$\Sigma = 11.6148$	$\Sigma = 0.5023$	23.12

Tabela A.5.6 - Valores e tempos de processamento de  $\mathcal{V}_{CDbw}$  e  $\mathcal{V}_{CDbw}^{SOM}$  para as partições do conjunto de dados *Gauss\_1*.

Partições	No. de classes	Valor $\mathcal{V}_{Conn}$	Tempo(s) $\mathcal{V}_{Conn}$
k=2	2	0.7153	1.7104
k=20*	5	<b>1.9466</b>	1.8717
k=143	3	0.9328	1.7598
k=151	2	0.8867	1.7215
			$\Sigma = 7.0634$

Tabela A.5.7 - Valores e tempos de processamento do índice  $\mathcal{V}_{Conn}$  para as partições do conjunto de dados *Gauss\_1*.

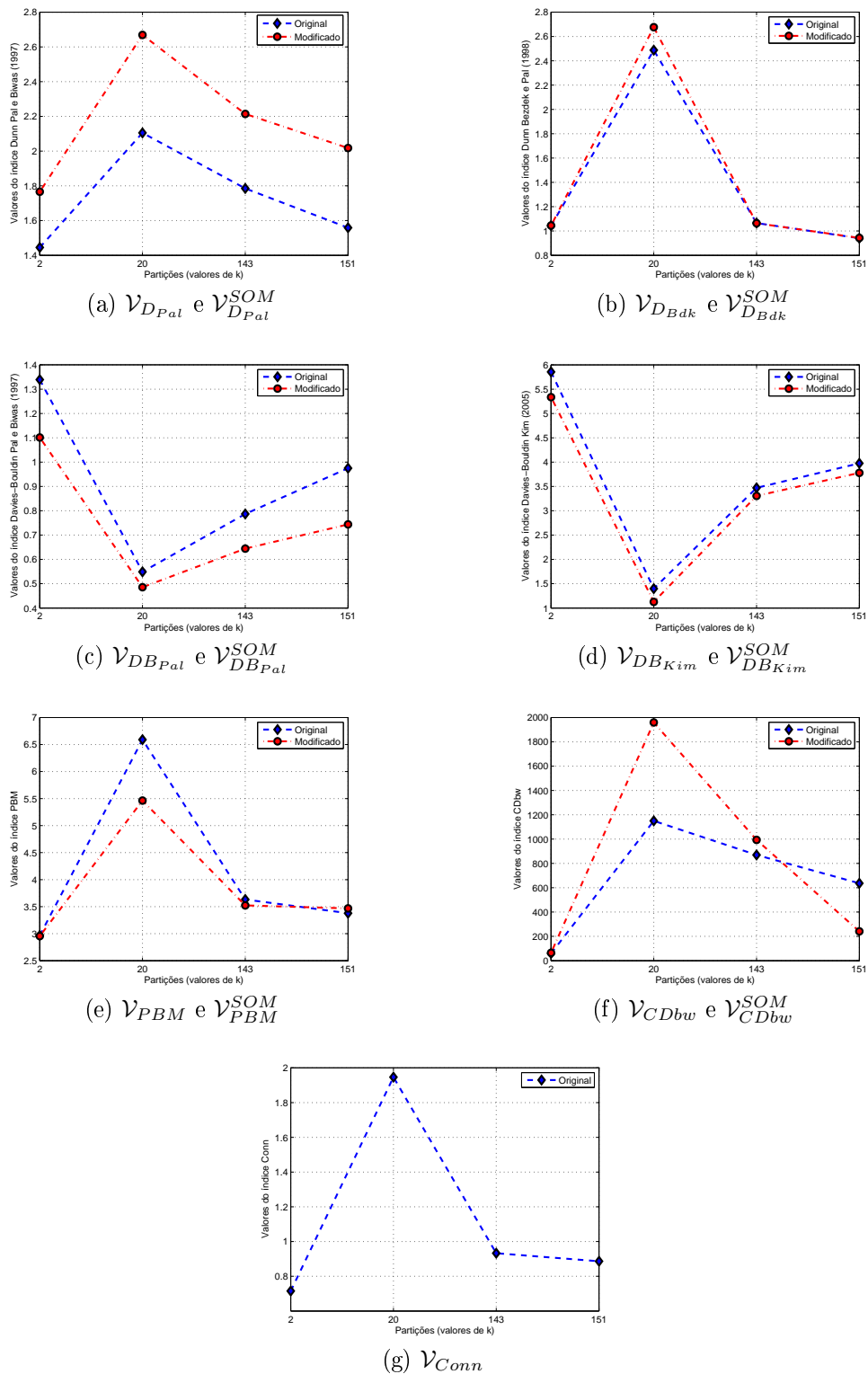


Fig. A.5.1 - Gráficos dos valores das versões originais e modificadas dos índices de validação para as partições do conjunto de dados Gauss\_1.

## Conjunto de dados *Gauss*\_2

Partições	No. de classes	Valor $\mathcal{V}_{DPal}$	Valor $\mathcal{V}_{DPal}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DPal}$	Tempo(s) $\mathcal{V}_{DPal}^{SOM}$	Speedup
k=7	3	2.6583	2.9413	0.2830	48.111	0.0679	708.5
k=13	4	3.1383	3.2529	0.1146	18.024	0.0714	252.4
k=22*	5	2.6736	2.9666	0.2929	15.0433	0.0610	246.6
k=149	4	<b>3.2635</b>	<b>3.3873</b>	0.1238	44.166	0.0858	514.7
				$\mu = 0.2035$	$\Sigma = 125.34$	$\Sigma = 0.2861$	438.09

Tabela A.5.8 - Valores e tempos de processamento de  $\mathcal{V}_{DPal}$  e  $\mathcal{V}_{DPal}^{SOM}$  para as partições do conjunto de dados *Gauss*\_2.

Partições	No. de classes	Valor $\mathcal{V}_{DBdk}$	Valor $\mathcal{V}_{DBdk}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DBdk}$	Tempo(s) $\mathcal{V}_{DBdk}^{SOM}$	Speedup
k=7	3	0.9689	0.9702	0.0013	1.7207	0.0459	37.5
k=13	4	1.7846	1.8395	0.0548	1.8877	0.0456	41.4
k=22*	5	<b>2.4739</b>	<b>2.8563</b>	0.3823	2.0481	0.0492	41.6
k=149	4	1.1804	1.1822	0.0017	1.8562	0.0453	40.9
				$\mu = 0.1101$	$\Sigma = 7.5126$	$\Sigma = 0.1861$	40.36

Tabela A.5.9 - Valores e tempos de processamento de  $\mathcal{V}_{DBdk}$  e  $\mathcal{V}_{DBdk}^{SOM}$  para as partições do conjunto de dados *Gauss*\_2.

Partições	No. de classes	Valor $\mathcal{V}_{DBPal}$	Valor $\mathcal{V}_{DBPal}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DBPal}$	Tempo(s) $\mathcal{V}_{DBPal}^{SOM}$	Speedup
k=7	3	0.6201	0.5582	0.0619	40.188	0.0611	657.7
k=13	4	0.6064	0.5468	0.0596	19.453	0.0729	266.8
k=22*	5	<b>0.5332</b>	<b>0.5118</b>	0.0214	21.364	0.0744	287.15
k=149	4	0.5526	0.5313	0.0213	41.871	0.0855	489.1
				$\mu = 0.0410$	$\Sigma = 122.876$	$\Sigma = 0.2939$	418.1

Tabela A.5.10 - Valores e tempos de processamento de  $\mathcal{V}_{DBPal}$  e  $\mathcal{V}_{DBPal}^{SOM}$  para as partições do conjunto de dados *Gauss*\_2.

Partições	No. de classes	Valor $\mathcal{V}_{DBKim}$	Valor $\mathcal{V}_{DBKim}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DBKim}$	Tempo(s) $\mathcal{V}_{DBKim}^{SOM}$	Speedup
k=7	3	4.7224	4.1103	0.6121	1.0200	0.0238	42.85
k=13	4	2.8487	2.6957	0.1530	1.0276	0.0269	38.2
k=22*	5	<b>1.6305</b>	<b>1.3117</b>	0.3188	1.0287	0.0254	40.5
k=149	4	2.0793	1.7972	0.2820	1.0271	0.0252	40.75
				$\mu = 0.3415$	$\Sigma = 4.1036$	$\Sigma = 0.1015$	40.42

Tabela A.5.11 - Valores e tempos de processamento de  $\mathcal{V}_{DBKim}$  e  $\mathcal{V}_{DBKim}^{SOM}$  para as partições do conjunto de dados *Gauss\_2*.

Partições	No. de classes	Valor $\mathcal{V}_{PBM}$	Valor $\mathcal{V}_{PBM}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{PBM}$	Tempo(s) $\mathcal{V}_{PBM}^{SOM}$	Speedup
k=7	3	2.7420	2.6739	0.0680	1.0580	0.0157	24.3
k=13	4	4.5824	4.0518	0.5306	1.0335	0.0254	38.8
k=22*	5	<b>7.6752</b>	<b>5.1907</b>	2.4845	1.0337	0.0262	40.3
k=149	4	5.5893	4.6318	0.9574	1.0327	0.0259	39.0
				$\mu = 1.0102$	$\Sigma = 4.1581$	$\Sigma = 0.0934$	44.52

Tabela A.5.12 - Valores e tempos de processamento de  $\mathcal{V}_{PBM}$  e  $\mathcal{V}_{PBM}^{SOM}$  para as partições do conjunto de dados *Gauss\_2*.

Partições	No. de classes	Valor $\mathcal{V}_{CDbw}$	Valor $\mathcal{V}_{CDbw}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{CDbw}$	Tempo(s) $\mathcal{V}_{CDbw}^{SOM}$	Speedup
k=7	3	200.43	196.44	3.9942	2.9903	0.1073	27.86
k=13	4	279.91	356.87	76.96	2.6521	0.1012	26.2
k=22*	5	<b>572.61</b>	<b>2641.7</b>	2069.1	2.4924	0.1044	23.87
k=149	4	505.77	1850.4	1344.7	2.9519	0.0911	32.4
				$\mu = 873.43$	$\Sigma = 11.0867$	$\Sigma = 0.4041$	27.43

Tabela A.5.13 - Valores e tempos de processamento de  $\mathcal{V}_{CDbw}$  e  $\mathcal{V}_{CDbw}^{SOM}$  para as partições do conjunto de dados *Gauss\_2*.

---

Partições	No. de classes	Valor $\mathcal{V}_{Conn}$	Tempo(s) $\mathcal{V}_{Conn}$
k=7	3	0.6997	2.4132
k=13	4	0.6576	1.9356
k=22*	5	<b>0.7548</b>	1.9254
k=149	4	0.6491	1.9304
			$\Sigma = 8.2047$

Tabela A.5.14 - Valores e tempos de processamento do índice  $\mathcal{V}_{Conn}$  para as partições do conjunto de dados *Gauss\_2*.

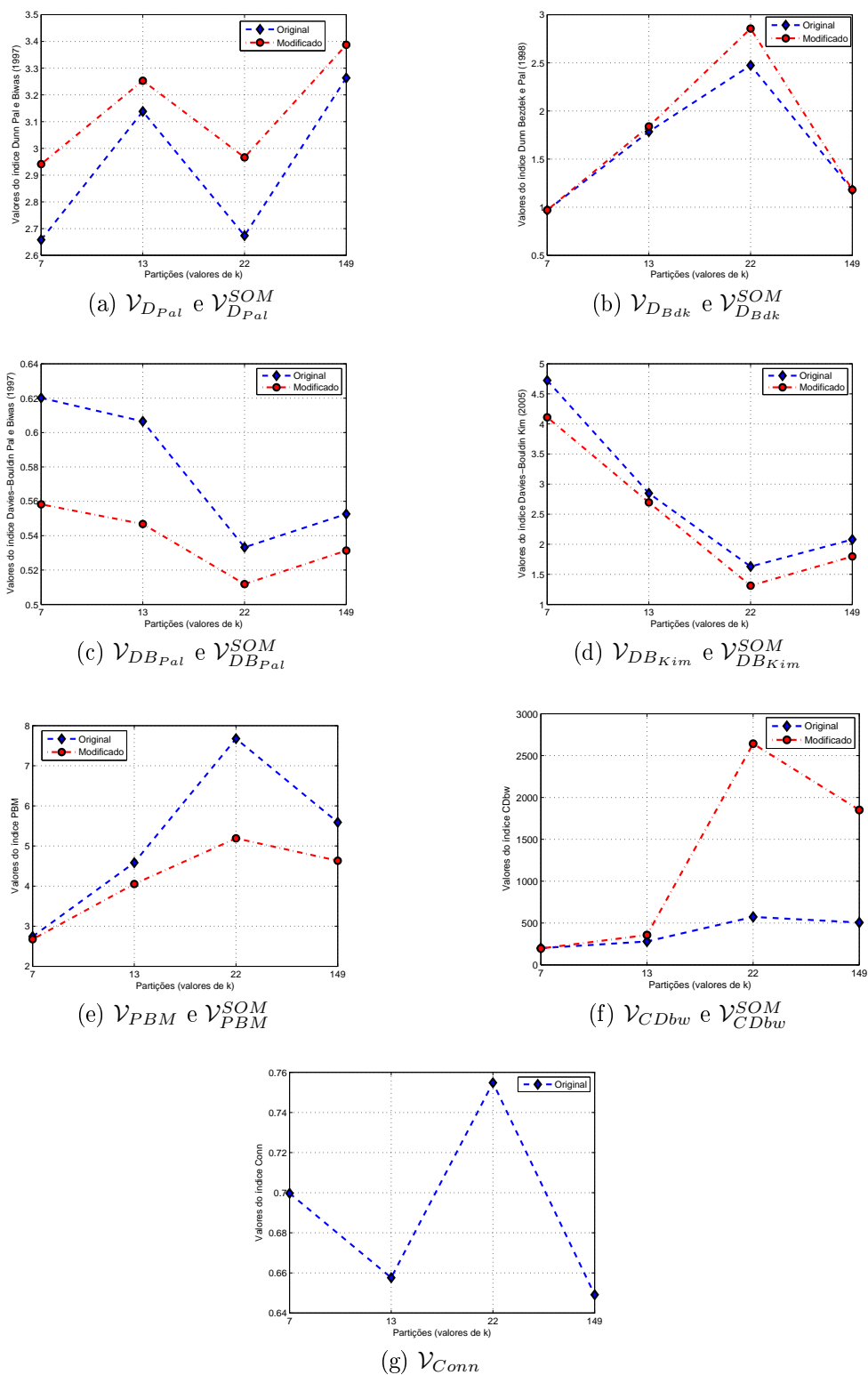


Fig. A.5.2 - Gráficos dos valores das versões originais e modificadas dos índices de validação para as partições do conjunto de dados Gauss<sub>2</sub>.

## Conjunto de dados *Gauss*\_3

Partições	No. de classes	Valor $\mathcal{V}_{DPal}$	Valor $\mathcal{V}_{DPal}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DPal}$	Tempo(s) $\mathcal{V}_{DPal}^{SOM}$	Speedup
k=17	2	2.3169	2.387	0.0700	26.751	0.0685	390.52
k=28	4	2.1509	2.2145	0.0635	6.4035	0.0415	154.30
k=57*	5	2.406	<b>3.3964</b>	0.9904	3.3424	0.0485	68.91
k=123	2	<b>2.4673</b>	2.5366	0.0693	56.078	0.1098	510.7
				$\mu = 0.2983$	$\Sigma = 92.575$	$\Sigma = 0.2685$	344.78

Tabela A.5.15 - Valores e tempos de processamento de  $\mathcal{V}_{DPal}$  e  $\mathcal{V}_{DPal}^{SOM}$  para as partições do conjunto de dados *Gauss*\_3.

Partições	No. de classes	Valor $\mathcal{V}_{DBdk}$	Valor $\mathcal{V}_{DBdk}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DBdk}$	Tempo(s) $\mathcal{V}_{DBdk}^{SOM}$	Speedup
k=17	2	1.0488	1.0533	0.0084	1.0015	0.0390	25.67
k=28	4	1.2433	1.2661	0.0227	1.3981	0.0476	29.37
k=57*	5	<b>1.6875</b>	<b>1.9035</b>	0.2160	1.5	0.0494	30.36
k=123	2	0.9182	0.9202	0.0020	0.6496	0.0330	19.68
				$\mu = 0.0623$	$\Sigma = 4.5492$	$\Sigma = 0.1691$	26.90

Tabela A.5.16 - Valores e tempos de processamento de  $\mathcal{V}_{DBdk}$  e  $\mathcal{V}_{DBdk}^{SOM}$  para as partições do conjunto de dados *Gauss*\_3.

Partições	No. de classes	Valor $\mathcal{V}_{DBPal}$	Valor $\mathcal{V}_{DBPal}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DBPal}$	Tempo(s) $\mathcal{V}_{DBPal}^{SOM}$	Speedup
k=17	2	0.7283	0.8006	0.0722	29.991	0.0533	562.68
k=28	4	0.6576	0.6225	0.0351	6.8603	0.0307	223.46
k=57*	5	<b>0.5930</b>	<b>0.4758</b>	0.1172	3.5833	0.0295	121.46
k=123	2	0.6643	0.6088	0.0555	58.347	0.0983	593.56
				$\mu = 0.0700$	$\Sigma = 98.781$	$\Sigma = 0.212$	465.94

Tabela A.5.17 - Valores e tempos de processamento de  $\mathcal{V}_{DBPal}$  e  $\mathcal{V}_{DBPal}^{SOM}$  para as partições do conjunto de dados *Gauss*\_3.



Partições	No. de classes	Valor $\mathcal{V}_{DBKim}$	Valor $\mathcal{V}_{DBKim}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DBKim}$	Tempo(s) $\mathcal{V}_{DBKim}^{SOM}$	Speedup
k=17	2	6.026	5.6477	0.3782	1.0303	0.0250	41.21
k=28	4	3.548	2.7605	0.7875	1.0284	0.0251	40.97
k=57*	5	<b>2.4808</b>	<b>1.9885</b>	0.4923	1.0276	0.0248	41.43
k=123	2	4.6763	4.304	0.3723	1.0266	0.0246	41.73
				$\mu = 0.5076$	$\Sigma = 4.113$	$\Sigma = 0.0997$	41.25

Tabela A.5.18 - Valores e tempos de processamento de  $\mathcal{V}_{DBKim}$  e  $\mathcal{V}_{DBKim}^{SOM}$  para as partições do conjunto de dados *Gauss\_3*.

Partições	No. de classes	Valor $\mathcal{V}_{PBM}$	Valor $\mathcal{V}_{PBM}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{PBM}$	Tempo(s) $\mathcal{V}_{PBM}^{SOM}$	Speedup
k=17	2	3.0923	3.1007	0.0083	1.0347	0.0266	38.9
k=28	4	4.2106	3.9465	0.2641	1.0316	0.0254	40.6
k=57*	5	<b>5.0754</b>	<b>4.3335</b>	0.7419	1.0314	0.0254	40.6
k=123	2	3.2763	3.2517	0.0245	1.0313	0.0259	39.8
				$\mu = 0.25973$	$\Sigma = 4.1292$	$\Sigma = 0.1034$	39.93

Tabela A.5.19 - Valores e tempos de processamento de  $\mathcal{V}_{PBM}$  e  $\mathcal{V}_{PBM}^{SOM}$  para as partições do conjunto de dados *Gauss\_3*.

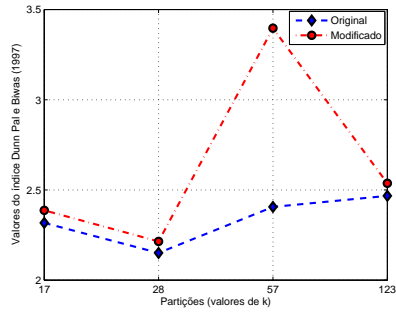
Partições	No. de classes	Valor $\mathcal{V}_{CDbw}$	Valor $\mathcal{V}_{CDbw}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{CDbw}$	Tempo(s) $\mathcal{V}_{CDbw}^{SOM}$	Speedup
k=17	2	29.634	78.4	48.766	2.9204	0.1023	28.5
k=28	4	71.065	317.88	246.81	2.2685	0.1010	22.4
k=57*	5	<b>292.12</b>	<b>801.32</b>	509.2	2.1413	0.1086	19.71
k=123	2	64.07	108.06	43.992	3.4276	0.1063	32.24
				$\mu = 212.19$	$\Sigma = 10.7578$	$\Sigma = 0.4184$	25.71

Tabela A.5.20 - Valores e tempos de processamento de  $\mathcal{V}_{CDbw}$  e  $\mathcal{V}_{CDbw}^{SOM}$  para as partições do conjunto de dados *Gauss\_3*.

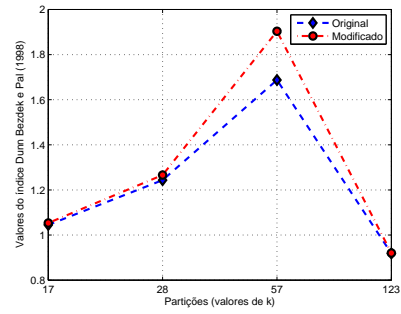
---

Partições	No. de classes	Valor $\mathcal{V}_{Conn}$	Tempo(s) $\mathcal{V}_{Conn}$
k=17	2	0.6961	1.6552
k=28	4	0.77	1.6538
k=57*	5	<b>0.8886</b>	1.6865
k=123	2	0.6436	1.6696
			$\Sigma = 6.6651$

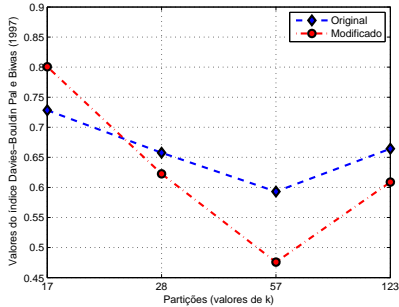
Tabela A.5.21 - Valores e tempos de processamento do índice  $\mathcal{V}_{Conn}$  para as partições do conjunto de dados *Gauss\_3*.



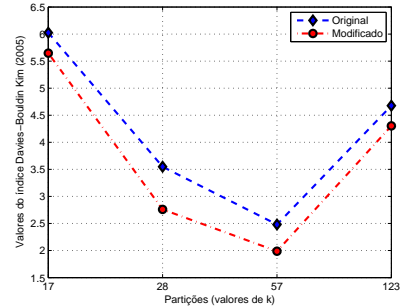
(a)  $\mathcal{V}_{DPal}$  e  $\mathcal{V}_{DPal}^{SOM}$



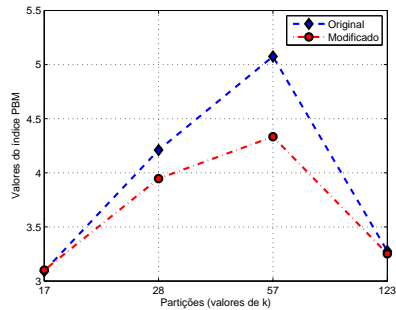
(b)  $\mathcal{V}_{DBdk}$  e  $\mathcal{V}_{DBdk}^{SOM}$



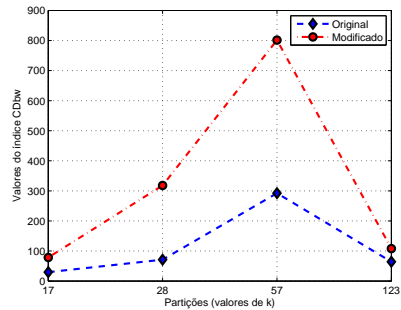
(c)  $\mathcal{V}_{DBPal}$  e  $\mathcal{V}_{DBPal}^{SOM}$



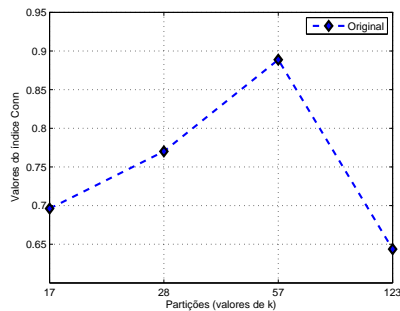
(d)  $\mathcal{V}_{DBKim}$  e  $\mathcal{V}_{DBKim}^{SOM}$



(e)  $\mathcal{V}_{PBM}$  e  $\mathcal{V}_{PBM}^{SOM}$



(f)  $\mathcal{V}_{CDbw}$  e  $\mathcal{V}_{CDbw}^{SOM}$



(g)  $\mathcal{V}_{Conn}$

Fig. A.5.3 - Gráficos dos valores das versões originais e modificadas dos índices de validação para as partições do conjunto de dados Gauss\_3.

## Conjunto de dados *Rand\_1*

Partições	No. de classes	Valor $\mathcal{V}_{DPal}$	Valor $\mathcal{V}_{DPal}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DPal}$	Tempo(s) $\mathcal{V}_{DPal}^{SOM}$	Speedup
k=14	2	1.0117	2.0129	0.9952	50.679	0.1972	256.99
k=21	3	1.2558	2.4847	1.2289	27.254	0.1167	233.53
k=57*	4	0.1175	0.0727	0.0448	53.156	0.1803	294.82
k=64	3	0.1124	0.0569	0.0555	53.695	0.1787	300.47
k=81	2	<b>4.709</b>	<b>6.307</b>	1.598	65.542	0.2143	305.84
				$\mu = 0.7845$	$\Sigma = 250.33$	$\Sigma = 0.8873$	282.12

Tabela A.5.22 - Valores e tempos de processamento de  $\mathcal{V}_{DPal}$  e  $\mathcal{V}_{DPal}^{SOM}$  para as partições do conjunto de dados *Rand\_1*.

Partições	No. de classes	Valor $\mathcal{V}_{DBdk}$	Valor $\mathcal{V}_{DBdk}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DBdk}$	Tempo(s) $\mathcal{V}_{DBdk}^{SOM}$	Speedup
k=14	2	0.7164	0.7125	0.0038	1.5269	0.0787	19.40
k=21	3	0.8272	0.8193	0.0079	1.9174	0.0735	26.08
k=57*	4	0.3539	0.3452	0.0086	1.6387	0.0625	26.21
k=64	3	0.5047	0.5047	0.0000	1.5924	0.0631	25.23
k=81	2	<b>1.7778</b>	<b>1.8012</b>	0.0233	1.3748	0.0583	23.58
				$\mu = 0.0087$	$\Sigma = 8.0501$	$\Sigma = 0.3363$	23.93

Tabela A.5.23 - Valores e tempos de processamento de  $\mathcal{V}_{DBdk}$  e  $\mathcal{V}_{DBdk}^{SOM}$  para as partições do conjunto de dados *Rand\_1*.

Partições	No. de classes	Valor $\mathcal{V}_{DBPal}$	Valor $\mathcal{V}_{DBPal}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DBPal}$	Tempo(s) $\mathcal{V}_{DBPal}^{SOM}$	Speedup
k=14	2	1.9397	0.9254	1.0143	54.139	0.1742	310.78
k=21	3	0.8777	0.4420	0.4357	28.595	0.1062	269.25
k=57*	4	6.0745	13.05	6.9756	51.949	0.1809	287.17
k=64	3	2.2584	11.962	9.7037	53.252	0.1784	298.49
k=81	2	<b>0.3613</b>	<b>0.2874</b>	0.0738	64.183	0.2092	306.80
				$\mu = 3.6406$	$\Sigma = 252.12$	$\Sigma = 0.8491$	296.92

Tabela A.5.24 - Valores e tempos de processamento de  $\mathcal{V}_{DBPal}$  e  $\mathcal{V}_{DBPal}^{SOM}$  para as partições do conjunto de dados *Rand\_1*.

Partições	No. de classes	Valor $\mathcal{V}_{DBKim}$	Valor $\mathcal{V}_{DBKim}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DBKim}$	Tempo(s) $\mathcal{V}_{DBKim}^{SOM}$	Speedup
k=14	2	64.851	60.305	4.5463	1.0271	0.0265	38.75
k=21	3	46.987	43.891	3.0955	1.0273	0.0264	38.91
k=57*	4	<b>21.375</b>	<b>30.795</b>	9.420	1.0302	0.0262	39.32
k=64	3	27.178	32.986	5.8075	1.0281	0.0255	40.31
k=81	2	35.094	34.06	1.0338	1.0277	0.0257	39.98
				$\mu = 4.7806$	$\Sigma = 4.1407$	$\Sigma = 0.1305$	31.73

Tabela A.5.25 - Valores e tempos de processamento de  $\mathcal{V}_{DBKim}$  e  $\mathcal{V}_{DBKim}^{SOM}$  para as partições do conjunto de dados *Rand\_1*.

Partições	No. de classes	Valor $\mathcal{V}_{PBM}$	Valor $\mathcal{V}_{PBM}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{PBM}$	Tempo(s) $\mathcal{V}_{PBM}^{SOM}$	Speedup
k=14	2	16.095	17.193	1.0986	1.0267	0.0153	67.1
k=21	3	30.649	29.168	1.4813	1.0335	0.0295	35.03
k=57*	4	35.222	34.006	1.2157	1.0346	0.0296	34.95
k=64	3	46.73	46.209	0.5212	1.0335	0.0272	37.99
k=81	2	<b>70.013</b>	<b>69.011</b>	1.0023	1.0334	0.0269	38.4
				$\mu = 1.0638$	$\Sigma = 4.1618$	$\Sigma = 0.1286$	32.36

Tabela A.5.26 - Valores e tempos de processamento de  $\mathcal{V}_{PBM}$  e  $\mathcal{V}_{PBM}^{SOM}$  para as partições do conjunto de dados *Rand\_1*.

Partições	No. de classes	Valor $\mathcal{V}_{CDbw}$	Valor $\mathcal{V}_{CDbw}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{CDbw}$	Tempo(s) $\mathcal{V}_{CDbw}^{SOM}$	Speedup
k=14	2	8.2837	31.094	22.811	3.0198	0.2770	10.9
k=21	3	26.635	75.943	49.308	2.5066	0.2330	10.75
k=57*	4	<b>834.53</b>	<b>815.2</b>	19.321	3.2083	0.2399	13.37
k=64	3	598.55	318.7	279.86	3.2903	0.2419	13.60
k=81	2	379.2	79.113	300.09	3.5545	0.2103	16.9
				$\mu = 134.28$	$\Sigma = 15.579$	$\Sigma = 1.2023$	12.95

Tabela A.5.27 - Valores e tempos de processamento de  $\mathcal{V}_{CDbw}$  e  $\mathcal{V}_{CDbw}^{SOM}$  para as partições do conjunto de dados *Rand\_1*.

Partições	No. de classes	Valor $\mathcal{V}_{Conn}$	Tempo(s) $\mathcal{V}_{Conn}$
k=14	2	0.6947	6.6197
k=21	3	0.5985	6.8793
k=57*	4	<b>0.975</b>	6.5432
k=64	3	0.8341	6.5318
k=81	2	0.7321	6.4033
			$\Sigma = 32.977$

Tabela A.5.28 - Valores e tempos de processamento do índice  $\mathcal{V}_{Conn}$  para as partições do conjunto de dados *Rand\_1*.

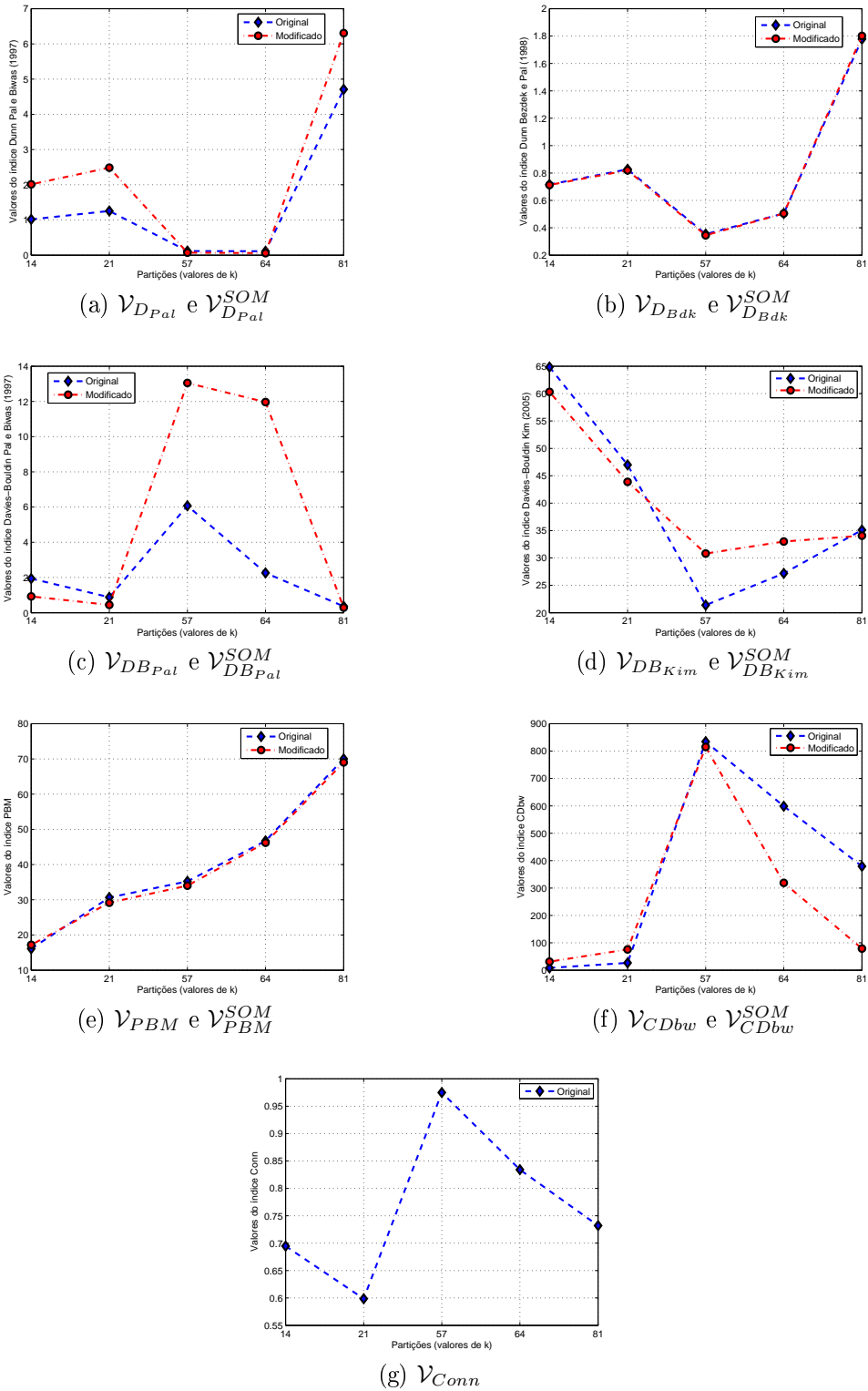


Fig. A.5.4 - Gráficos dos valores das versões originais e modificadas dos índices de validação para as partições do conjunto de dados *Rand\_1*.

## Conjunto de dados *Rand\_2*

Partições	No. de classes	Valor $\mathcal{V}_{DPal}$	Valor $\mathcal{V}_{DPal}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DPal}$	Tempo(s) $\mathcal{V}_{DPal}^{SOM}$	Speedup
k=4	3	1.0472	1.7895	0.7423	128.52	0.3148	408.25
k=35*	8	0.4465	0.4479	0.0014	7.992	0.0571	139.96
k=47	6	0.8312	0.9786	0.1473	20.622	0.0617	334.23
k=68	5	1.9776	2.1103	0.1327	61.663	0.1476	417.77
k=78	4	1.6186	1.296	0.3225	61.152	0.1482	412.63
k=104	2	<b>3.1593</b>	<b>3.3767</b>	0.2173	61.957	0.1481	418.34
				$\mu = 0.2606$	$\Sigma = 341.91$	$\Sigma = 0.8777$	389.55

Tabela A.5.29 - Valores e tempos de processamento de  $\mathcal{V}_{DPal}$  e  $\mathcal{V}_{DPal}^{SOM}$  para as partições do conjunto de dados *Rand\_2*.

Partições	No. de classes	Valor $\mathcal{V}_{DBdk}$	Valor $\mathcal{V}_{DBdk}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DBdk}$	Tempo(s) $\mathcal{V}_{DBdk}^{SOM}$	Speedup
k=4	3	0.5764	0.5680	0.0083	1.4268	0.0503	28.36
k=35*	8	0.4782	0.4546	0.0236	3.1666	0.0895	35.38
k=47	6	0.5263	0.5059	0.0203	2.9022	0.0878	33.05
k=68	5	0.5210	0.4967	0.0242	2.1826	0.0684	31.90
k=78	4	0.5252	0.5063	0.0188	2.171	0.0660	32.89
k=104	2	<b>1.0404</b>	<b>1.0396</b>	0.0008	2.1316	0.0684	31.16
				$\mu = 0.0160$	$\Sigma = 13.981$	$\Sigma = 0.4307$	32.46

Tabela A.5.30 - Valores e tempos de processamento de  $\mathcal{V}_{DBdk}$  e  $\mathcal{V}_{DBdk}^{SOM}$  para as partições do conjunto de dados *Rand\_2*.



Partições	No. de classes	Valor $\mathcal{V}_{DBPal}$	Valor $\mathcal{V}_{DBPal}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DBPal}$	Tempo(s) $\mathcal{V}_{DBPal}^{SOM}$	Speedup
k=4	3	0.4549	0.4753	0.0194	130.73	0.3035	430.74
k=35*	8	0.7632	1.0004	0.2372	8.0365	0.0583	137.84
k=47	6	0.7004	0.791	0.0906	20.017	0.0767	260.97
k=68	5	0.5266	0.5842	0.0576	61.474	0.1498	410.37
k=78	4	0.6392	0.7002	0.061	62.186	0.1479	420.45
k=104	2	<b>0.4368</b>	<b>0.4387</b>	0.0738	63.043	0.1529	412.31
				$\mu = 0.0899$	$\Sigma = 345.49$	$\Sigma = 0.8893$	388.49

Tabela A.5.31 - Valores e tempos de processamento de  $\mathcal{V}_{DBPal}$  e  $\mathcal{V}_{DBPal}^{SOM}$  para as partições do conjunto de dados *Rand\_2*.

Partições	No. de classes	Valor $\mathcal{V}_{DBKim}$	Valor $\mathcal{V}_{DBKim}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DBKim}$	Tempo(s) $\mathcal{V}_{DBKim}^{SOM}$	Speedup
k=4	3	82.376	97.982	15.606	1.0286	0.0262	39.25
k=35*	8	43.412	56.464	13.052	1.0288	0.0040	257.2
k=47	6	41.731	50.036	8.304	1.0278	0.0256	40.14
k=68	5	<b>40.249</b>	<b>45.375</b>	5.1264	1.0270	0.0256	40.11
k=78	4	56.54	54.338	2.2011	1.0266	0.0259	39.63
k=104	2	69.433	66.071	3.3621	1.0262	0.0251	40.88
				$\mu = 7.9419$	$\Sigma = 6.1652$	$\Sigma = 0.1326$	46.49

Tabela A.5.32 - Valores e tempos de processamento de  $\mathcal{V}_{DBKim}$  e  $\mathcal{V}_{DBKim}^{SOM}$  para as partições do conjunto de dados *Rand\_2*.

Partições	No. de classes	Valor $\mathcal{V}_{PBM}$	Valor $\mathcal{V}_{PBM}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{PBM}$	Tempo(s) $\mathcal{V}_{PBM}^{SOM}$	Speedup
k=4	3	49.086	46.478	2.608	1.0367	0.0265	39.12
k=35*	8	78.818	68.93	9.888	1.0352	0.0286	36.19
k=47	6	69.855	64.343	5.512	1.0341	0.0274	37.74
k=68	5	68.923	64.023	4.9	1.0374	0.0265	39.14
k=78	4	67.817	63.383	4.433	1.0413	0.0284	36.66
k=104	2	<b>86.059</b>	<b>80.099</b>	5.96	1.0340	0.0259	39.92
				$\mu = 5.5501$	$\Sigma = 6.219$	$\Sigma = 0.1636$	38.01

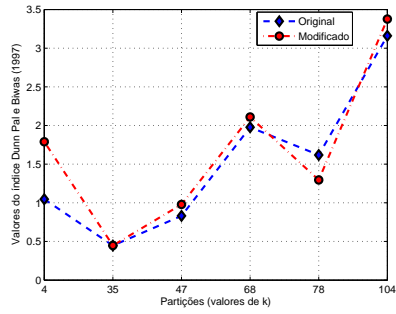
Tabela A.5.33 - Valores e tempos de processamento de  $\mathcal{V}_{PBM}$  e  $\mathcal{V}_{PBM}^{SOM}$  para as partições do conjunto de dados *Rand\_2*.

Partições	No. de classes	Valor $\mathcal{V}_{CDbw}$	Valor $\mathcal{V}_{CDbw}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{CDbw}$	Tempo(s) $\mathcal{V}_{CDbw}^{SOM}$	Speedup
k=4	3	26.188	66.579	40.391	4.4019	0.2344	18.77
k=35*	8	<b>1041.6</b>	<b>2668.8</b>	1627.2	1.9798	0.2302	8.6
k=47	6	674.18	1639.3	965.08	2.4381	0.2306	10.57
k=68	5	532.81	1194.7	661.94	3.3841	0.2412	14.03
k=78	4	383.92	308.17	75.746	3.3751	0.2258	14.94
k=104	2	207.81	120.48	87.333	3.285	0.2203	14.911
				$\mu = 576.28$	$\Sigma = 18.764$	$\Sigma = 1.3827$	13.57

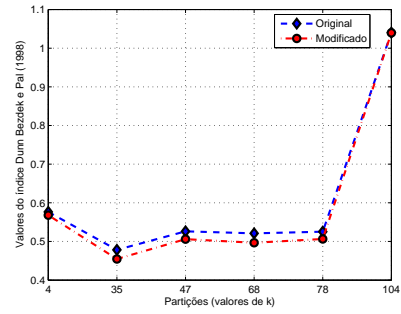
Tabela A.5.34 - Valores e tempos de processamento de  $\mathcal{V}_{CDbw}$  e  $\mathcal{V}_{CDbw}^{SOM}$  para as partições do conjunto de dados *Rand\_2*.

Partições	No. de classes	Valor $\mathcal{V}_{Conn}$	Tempo(s) $\mathcal{V}_{Conn}$
k=4	3	0.6318	8.0613
k=35*	8	0.9061	8.2763
k=47	6	0.8771	7.3156
k=68	5	0.8606	7.1409
k=78	4	0.4956	7.1266
k=104	2	<b>0.9801</b>	7.0847
			$\Sigma = 45.005$

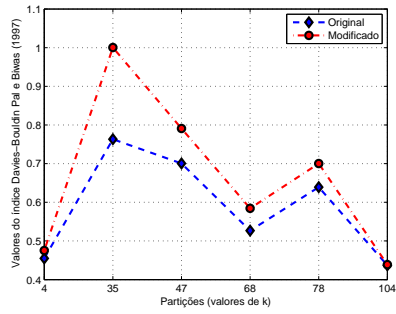
Tabela A.5.35 - Valores e tempos de processamento do índice  $\mathcal{V}_{Conn}$  para as partições do conjunto de dados *Rand\_2*.



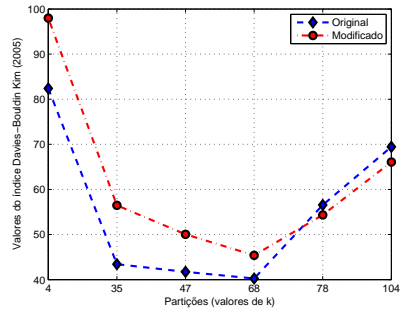
(a)  $\mathcal{V}_{DPal}$  e  $\mathcal{V}_{DPal}^{SOM}$



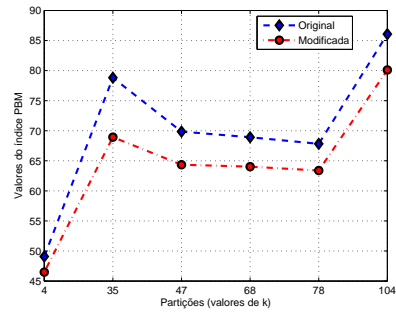
(b)  $\mathcal{V}_{DBdk}$  e  $\mathcal{V}_{DBdk}^{SOM}$



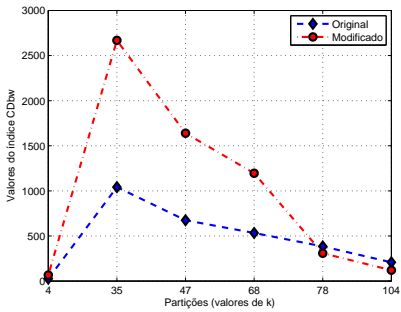
(c)  $\mathcal{V}_{DBPal}$  e  $\mathcal{V}_{DBPal}^{SOM}$



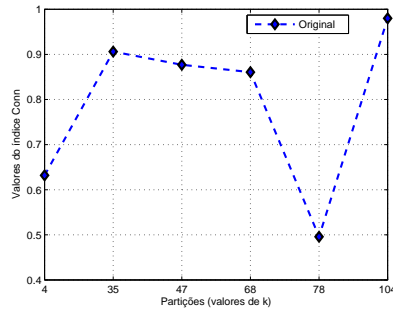
(d)  $\mathcal{V}_{DBKim}$  e  $\mathcal{V}_{DBKim}^{SOM}$



(e)  $\mathcal{V}_{PBM}$  e  $\mathcal{V}_{PBM}^{SOM}$



(f)  $\mathcal{V}_{CDbw}$  e  $\mathcal{V}_{CDbw}^{SOM}$



(g)  $\mathcal{V}_{Conn}$

Fig. A.5.5 - Gráficos dos valores das versões originais e modificadas dos índices de validação para as partições do conjunto de dados *Rand\_2*.

## Conjunto de dados *Rand\_3*

Partições	No. de classes	Valor $\mathcal{V}_{DPal}$	Valor $\mathcal{V}_{DPal}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DPal}$	Tempo(s) $\mathcal{V}_{DPal}^{SOM}$	Speedup
k=36	4	1.9369	1.9863	0.0493	3.2932	0.0424	77.66
k=74*	6	0.9263	0.9209	0.0054	2.5073	0.0312	80.36
k=79	5	1.5706	1.8083	0.2376	2.6163	0.0331	79.04
k=98	4	<b>2.298</b>	<b>3.1125</b>	0.8144	4.3857	0.0325	134.94
k=142	2	1.7922	3.0174	1.2252	33.921	0.0501	677.06
k=174	3	1.7922	3.0174	1.2252	33.94	0.0545	622.75
				$\mu = 0.5928$	$\Sigma = 80.663$	$\Sigma = 0.2442$	330.31

Tabela A.5.36 - Valores e tempos de processamento de  $\mathcal{V}_{DPal}$  e  $\mathcal{V}_{DPal}^{SOM}$  para as partições do conjunto de dados *Rand\_3*.

Partições	No. de classes	Valor $\mathcal{V}_{DBdk}$	Valor $\mathcal{V}_{DBdk}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DBdk}$	Tempo(s) $\mathcal{V}_{DBdk}^{SOM}$	Speedup
k=36	4	0.8273	0.8591	0.0317	1.0936	0.0274	39.91
k=74*	6	0.9534	0.9413	0.0120	1.115	0.0369	30.21
k=79	5	1.1932	1.4104	0.2171	1.139	0.0365	30.86
k=98	4	<b>1.6601</b>	<b>1.6936</b>	0.0335	1.08	0.0358	30.16
k=142	2	1.0788	1.0661	0.0127	0.4469	0.0289	15.46
k=174	3	1.0788	1.0661	0.0127	0.5152	0.0215	23.96
				$\mu = 0.0533$	$\Sigma = 5.3898$	$\Sigma = 0.1873$	28.77

Tabela A.5.37 - Valores e tempos de processamento de  $\mathcal{V}_{DBdk}$  e  $\mathcal{V}_{DBdk}^{SOM}$  para as partições do conjunto de dados *Rand\_3*.

Partições	No. de classes	Valor $\mathcal{V}_{DBPal}$	Valor $\mathcal{V}_{DBPal}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DBPal}$	Tempo(s) $\mathcal{V}_{DBPal}^{SOM}$	Speedup
k=36	4	0.6790	0.6516	0.0274	3.3171	0.0222	149.42
k=74*	6	0.9492	0.7821	0.167	2.5462	0.0338	75.331
k=79	5	0.7107	0.6794	0.0312	2.6437	0.0316	83.661
k=98	4	<b>0.6403</b>	<b>0.5175</b>	0.1228	4.4241	0.0323	136.96
k=142	2	0.6874	0.5884	0.099	33.997	0.0582	584.14
k=174	3	0.6775	0.5873	0.0902	34.149	0.0521	655.45
				$\mu = 0.0896$	$\Sigma = 81.077$	$\Sigma = 0.2304$	351.89

Tabela A.5.38 - Valores e tempos de processamento de  $\mathcal{V}_{DBPal}$  e  $\mathcal{V}_{DBPal}^{SOM}$  para as partições do conjunto de dados *Rand\_3*.

Partições	No. de classes	Valor $\mathcal{V}_{DBKim}$	Valor $\mathcal{V}_{DBKim}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DBKim}$	Tempo(s) $\mathcal{V}_{DBKim}^{SOM}$	Speedup
k=36	4	3.6609	3.0748	0.5860	1.0290	0.0236	43.60
k=74*	6	<b>2.0572</b>	<b>1.6539</b>	0.4032	1.0269	0.0024	427.87
k=79	5	2.1842	1.6549	0.5293	1.0302	0.0244	42.22
k=98	4	2.5248	1.9859	0.5389	1.0263	0.0273	37.59
k=142	2	4.6461	4.2164	0.4296	1.0260	0.0277	37.03
k=174	3	4.6351	4.1459	0.4892	1.0291	0.0244	42.17
				$\mu = 0.4960$	$\Sigma = 6.1677$	$\Sigma = 0.1301$	47.40

Tabela A.5.39 - Valores e tempos de processamento de  $\mathcal{V}_{DBKim}$  e  $\mathcal{V}_{DBKim}^{SOM}$  para as partições do conjunto de dados *Rand\_3*.

Partições	No. de classes	Valor $\mathcal{V}_{PBM}$	Valor $\mathcal{V}_{PBM}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{PBM}$	Tempo(s) $\mathcal{V}_{PBM}^{SOM}$	Speedup
k=36	4	4.9082	4.4199	0.4883	1.0522	0.0187	56.26
k=74*	6	<b>6.825</b>	<b>4.9608</b>	1.8642	1.0312	0.0247	41.74
k=79	5	6.8073	4.9146	1.8927	1.0309	0.0247	41.73
k=98	4	6.5441	4.883	1.6611	1.0305	0.0245	42.01
k=142	2	3.9161	4.2096	0.2935	1.0308	0.0248	41.56
k=174	3	3.7122	4.1086	0.3964	1.0318	0.0348	29.64
				$\mu = 1.0993$	$\Sigma = 6.2077$	$\Sigma = 0.1424$	43.59

Tabela A.5.40 - Valores e tempos de processamento de  $\mathcal{V}_{PBM}$  e  $\mathcal{V}_{PBM}^{SOM}$  para as partições do conjunto de dados *Rand\_3*.

Partições	No. de classes	Valor $\mathcal{V}_{CDbw}$	Valor $\mathcal{V}_{CDbw}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{CDbw}$	Tempo(s) $\mathcal{V}_{CDbw}^{SOM}$	Speedup
k=36	4	564.18	1101.5	537.32	1.9894	0.0993	20.03
k=74*	6	<b>1504.9</b>	<b>2461</b>	956.1	1.8587	0.0673	27.61
k=79	5	935.98	1895.5	959.52	1.8858	0.0631	29.88
k=98	4	427.07	1072.3	645.23	1.9506	0.0629	31.01
k=142	2	65.599	187.9	122.3	2.7749	0.0680	40.80
k=174	3	63.124	186.4	123.27	2.8079	0.0625	44.92
				$\mu = 557.29$	$\Sigma = 13.2773$	$\Sigma = 0.4234$	31.35

Tabela A.5.41 - Valores e tempos de processamento de  $\mathcal{V}_{CDbw}$  e  $\mathcal{V}_{CDbw}^{SOM}$  para as partições do conjunto de dados *Rand\_3*.

Partições	No. de classes	Valor $\mathcal{V}_{Conn}$	Tempo(s) $\mathcal{V}_{Conn}$
k=36	4	0.4013	0.7484
k=74*	6	<b>0.8120</b>	0.8727
k=79	5	0.5076	0.8579
k=98	4	0.3282	0.8725
k=142	3	0.2969	0.8689
k=174	2	0.2832	0.8699
			$\Sigma = 5.0906$

Tabela A.5.42 - Valores e tempos de processamento do índice  $\mathcal{V}_{Conn}$  para as partições do conjunto de dados *Rand\_3*.



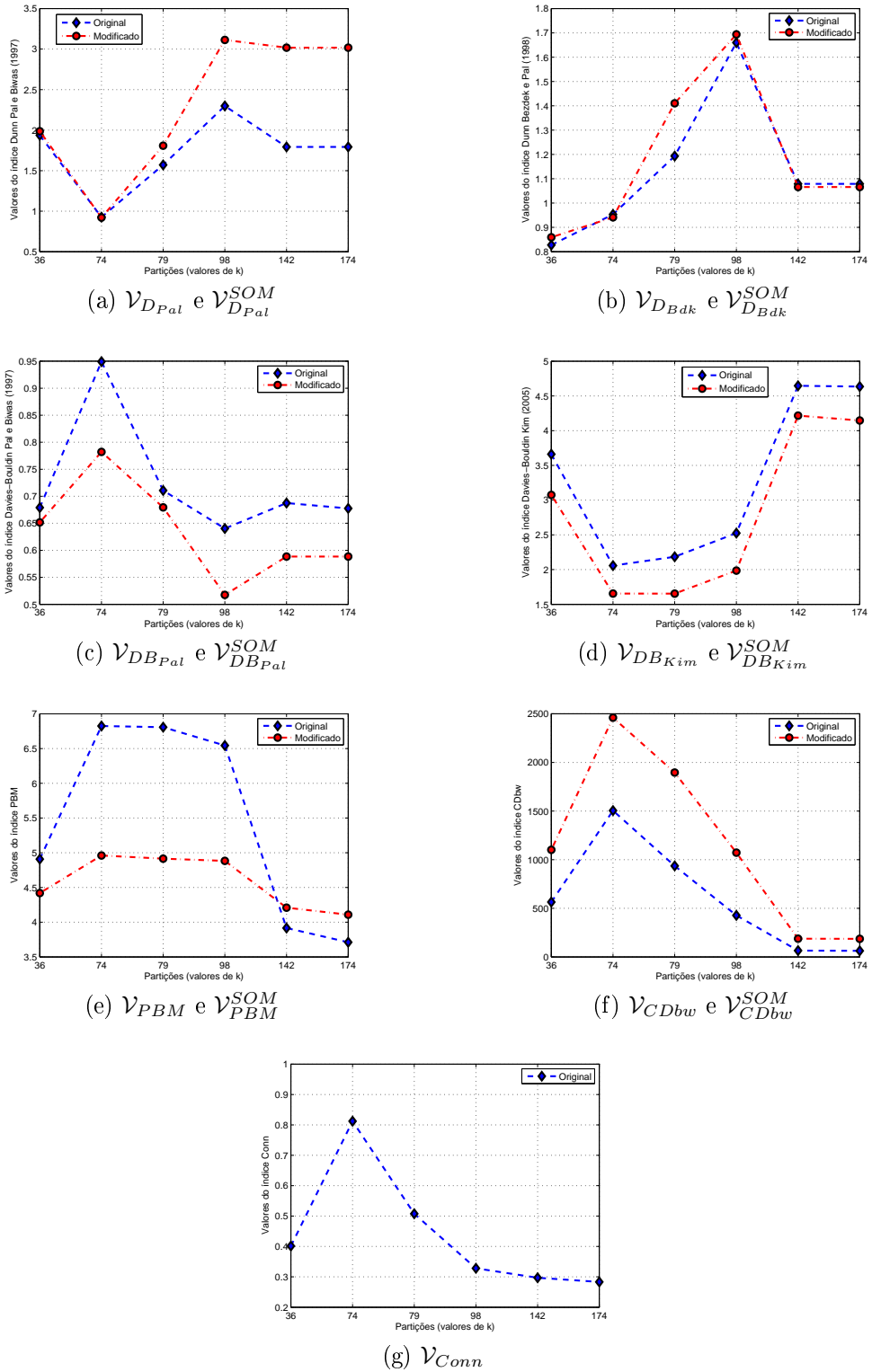


Fig. A.5.6 - Gráficos dos valores das versões originais e modificadas dos índices de validação para as partições do conjunto de dados *Rand\_3*.

## Conjunto de dados *WBC*

Partições	No. de classes	Valor $\mathcal{V}_{DPal}$	Valor $\mathcal{V}_{DPal}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DPal}$	Tempo(s) $\mathcal{V}_{DPal}^{SOM}$	Speedup
k=123	2	0.6839	2.1075	1.4235	9.4148	0.0284	331.50
k=148	3	0.5015	1.3361	0.8345	5.4969	0.0265	207.43
k=155*	2	<b>0.9635</b>	<b>2.8689</b>	1.9053	5.7748	0.0265	217.91
				$\mu = 1.3878$	$\Sigma = 20.686$	$\Sigma = 0.0815$	253.81

Tabela A.5.43 - Valores e tempos de processamento de  $\mathcal{V}_{DPal}$  e  $\mathcal{V}_{DPal}^{SOM}$  para as partições do conjunto de dados *WBC*.

Partições	No. de classes	Valor $\mathcal{V}_{DBdk}$	Valor $\mathcal{V}_{DBdk}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DBdk}$	Tempo(s) $\mathcal{V}_{DBdk}^{SOM}$	Speedup
k=123	2	<b>1.0988</b>	<b>1.9566</b>	0.8578	0.5105	0.0260	19.63
k=148	3	0.7949	1.2588	0.4638	0.6149	0.0252	24.4
k=155*	2	1.0345	1.7852	0.7506	0.4933	0.0243	20.30
				$\mu = 0.6907$	$\Sigma = 1.6187$	$\Sigma = 0.0756$	21.41

Tabela A.5.44 - Valores e tempos de processamento de  $\mathcal{V}_{DBdk}$  e  $\mathcal{V}_{DBdk}^{SOM}$  para as partições do conjunto de dados *WBC*.

Partições	No. de classes	Valor $\mathcal{V}_{DBPal}$	Valor $\mathcal{V}_{DBPal}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DBPal}$	Tempo(s) $\mathcal{V}_{DBPal}^{SOM}$	Speedup
k=123	2	2.5316	0.7573	1.7743	9.3787	0.0273	343.54
k=148	3	2.585	0.8862	1.6987	5.4834	0.0265	206.92
k=155*	2	<b>1.7855</b>	<b>0.6963</b>	1.0891	5.7324	0.0262	218.79
				$\mu = 1.5207$	$\Sigma = 20.595$	$\Sigma = 0.0801$	257.11

Tabela A.5.45 - Valores e tempos de processamento de  $\mathcal{V}_{DBPal}$  e  $\mathcal{V}_{DBPal}^{SOM}$  para as partições do conjunto de dados *WBC*.

Partições	No. de classes	Valor $\mathcal{V}_{DBKim}$	Valor $\mathcal{V}_{DBKim}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DBKim}$	Tempo(s) $\mathcal{V}_{DBKim}^{SOM}$	Speedup
k=123	2	12.641	7.9008	4.7407	0.9681	0.0241	40.17
k=148	3	12.069	7.342	4.7272	0.9701	0.0241	40.25
k=155*	2	<b>10.752</b>	<b>5.6844</b>	5.0672	0.9870	0.0240	41.12
				$\mu = 4.845$	$\Sigma = 2.9252$	$\Sigma = 0.0723$	40.45

Tabela A.5.46 - Valores e tempos de processamento de  $\mathcal{V}_{DBKim}$  e  $\mathcal{V}_{DBKim}^{SOM}$  para as partições do conjunto de dados *WBC*.

Partições	No. de classes	Valor $\mathcal{V}_{PBM}$	Valor $\mathcal{V}_{PBM}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{PBM}$	Tempo(s) $\mathcal{V}_{PBM}^{SOM}$	Speedup
k=123	2	6.322	6.4478	0.1256	0.8111	0.0260	31.19
k=148	3	9.5997	8.2435	1.3562	0.8048	0.0235	34.24
k=155*	2	<b>11.949</b>	<b>11.08</b>	0.8692	0.7941	0.0235	33.79
				$\mu = 0.7837$	$\Sigma = 2.41$	$\Sigma = 0.0731$	32.9

Tabela A.5.47 - Valores e tempos de processamento de  $\mathcal{V}_{PBM}$  e  $\mathcal{V}_{PBM}^{SOM}$  para as partições do conjunto de dados *WBC*.

Partições	No. de classes	Valor $\mathcal{V}_{CDbw}$	Valor $\mathcal{V}_{CDbw}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{CDbw}$	Tempo(s) $\mathcal{V}_{CDbw}^{SOM}$	Speedup
k=123	2	78.845	152.72	73.875	0.8199	0.0323	25.38
k=148	3	85.455	629.26	543.80	0.6502	0.0309	21.04
k=155*	2	<b>99.466</b>	<b>823.57</b>	724.10	0.6763	0.0289	23.40
				$\mu = 447.25$	$\Sigma = 2.1466$	$\Sigma = 0.0922$	23.28

Tabela A.5.48 - Valores e tempos de processamento de  $\mathcal{V}_{CDbw}$  e  $\mathcal{V}_{CDbw}^{SOM}$  para as partições do conjunto de dados *WBC*.

---

Partições	No. de classes	Valor $\mathcal{V}_{Conn}$	Tempo(s) $\mathcal{V}_{Conn}$
k=123	2	0.6678	0.0853
k=148	3	0.7046	0.0850
k=155*	2	<b>0.7452</b>	0.0839
			$\Sigma = 0.2543$

Tabela A.5.49 - Valores e tempos de processamento do índice  $\mathcal{V}_{Conn}$  para as partições do conjunto de dados *WBC*.

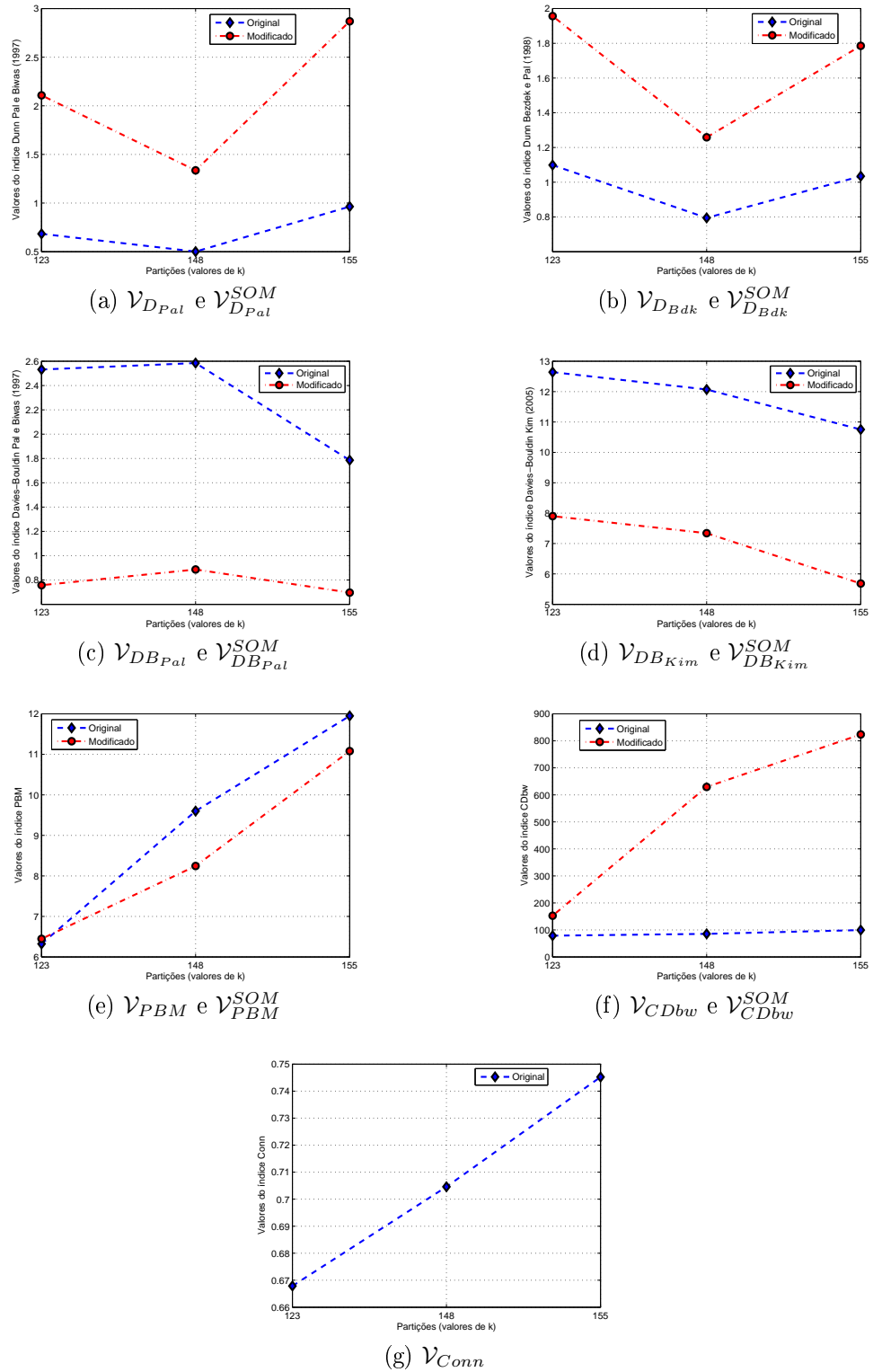


Fig. A.5.7 - Gráficos dos valores das versões originais e modificadas dos índices de validação para as partições do conjunto de dados WBC.

## Conjunto de dados *Wine*

Partições	No. de classes	Valor $\mathcal{V}_{DPal}$	Valor $\mathcal{V}_{DPal}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DPal}$	Tempo(s) $\mathcal{V}_{DPal}^{SOM}$	Speedup
k=57	2	0.4037	0.9945	0.5907	0.1069	0.0261	4.09
k=74*	3	0.6246	1.9181	1.2935	0.0709	0.0264	2.68
k=86	2	<b>0.6924</b>	<b>1.966</b>	1.2735	0.1610	0.0257	6.26
				$\mu = 1.0526$	$\Sigma = 0.3390$	$\Sigma = 0.0783$	4.32

Tabela A.5.50 - Valores e tempos de processamento de  $\mathcal{V}_{DPal}$  e  $\mathcal{V}_{DPal}^{SOM}$  para as partições do conjunto de dados *Wine*.

Partições	No. de classes	Valor $\mathcal{V}_{DBdk}$	Valor $\mathcal{V}_{DBdk}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DBdk}$	Tempo(s) $\mathcal{V}_{DBdk}^{SOM}$	Speedup
k=57	2	0.7727	0.7527	0.0199	0.0484	0.0247	1.95
k=74*	3	<b>0.8774</b>	<b>1.2179</b>	0.3404	0.0531	0.0239	2.22
k=86	2	0.8664	0.9989	0.1324	0.0159	0.0009	17.6
				$\mu = 0.1642$	$\Sigma = 0.1175$	$\Sigma = 0.0495$	2.37

Tabela A.5.51 - Valores e tempos de processamento de  $\mathcal{V}_{DBdk}$  e  $\mathcal{V}_{DBdk}^{SOM}$  para as partições do conjunto de dados *Wine*.

Partições	No. de classes	Valor $\mathcal{V}_{DBPal}$	Valor $\mathcal{V}_{DBPal}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DBPal}$	Tempo(s) $\mathcal{V}_{DBPal}^{SOM}$	Speedup
k=57	2	4.5379	1.479	3.0589	0.1087	0.0260	4.18
k=74*	3	<b>2.4425</b>	<b>0.7320</b>	1.7104	0.0713	0.0253	2.81
k=86	2	2.4681	0.7325	1.7356	0.1608	0.0258	6.23
				$\mu = 2.1683$	$\Sigma = 0.3409$	$\Sigma = 0.0764$	4.46

Tabela A.5.52 - Valores e tempos de processamento de  $\mathcal{V}_{DBPal}$  e  $\mathcal{V}_{DBPal}^{SOM}$  para as partições do conjunto de dados *Wine*.

Partições	No. de classes	Valor $\mathcal{V}_{DBK_{im}}$	Valor $\mathcal{V}_{DBK_{im}}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DBK_{im}}$	Tempo(s) $\mathcal{V}_{DBK_{im}}^{SOM}$	Speedup
k=57	2	6.2723	3.5347	2.7376	0.0264	0.0145	1.82
k=74*	3	<b>5.0031</b>	<b>1.2459</b>	3.7572	0.0256	0.0133	1.92
k=86	2	5.6223	2.1799	3.4424	0.0253	0.0101	2.50
				$\mu = 3.3124$	$\Sigma = 0.0773$	$\Sigma = 0.0379$	2.03

Tabela A.5.53 - Valores e tempos de processamento de  $\mathcal{V}_{DBK_{im}}$  e  $\mathcal{V}_{DBK_{im}}^{SOM}$  para as partições do conjunto de dados *Wine*.

Partições	No. de classes	Valor $\mathcal{V}_{PBM}$	Valor $\mathcal{V}_{PBM}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{PBM}$	Tempo(s) $\mathcal{V}_{PBM}^{SOM}$	Speedup
k=57	2	1.5417	1.8475	0.3058	0.0372	0.0201	1.85
k=74*	3	<b>2.2795</b>	<b>3.0402</b>	0.7607	0.0360	0.0203	1.77
k=86	2	2.2702	2.4865	0.2163	0.0376	0.0205	1.83
				$\mu = 0.4276$	$\Sigma = 0.1108$	$\Sigma = 0.0609$	1.81

Tabela A.5.54 - Valores e tempos de processamento de  $\mathcal{V}_{PBM}$  e  $\mathcal{V}_{PBM}^{SOM}$  para as partições do conjunto de dados *Wine*.

Partições	No. de classes	Valor $\mathcal{V}_{CDbw}$	Valor $\mathcal{V}_{CDbw}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{CDbw}$	Tempo(s) $\mathcal{V}_{CDbw}^{SOM}$	Speedup
k=57	2	3.6336	53.637	50.004	0.0795	0.0290	2.74
k=74*	3	<b>18.442</b>	<b>176.83</b>	158.29	0.0701	0.0274	2.55
k=86	2	3.8094	46.941	43.132	0.0800	0.0263	3.04
				$\mu = 83.843$	$\Sigma = 0.2297$	$\Sigma = 0.0828$	2.77

Tabela A.5.55 - Valores e tempos de processamento de  $\mathcal{V}_{CDbw}$  e  $\mathcal{V}_{CDbw}^{SOM}$  para as partições do conjunto de dados *Wine*.

Partições	No. de classes	Valor $\mathcal{V}_{Conn}$	Tempo(s) $\mathcal{V}_{Conn}$
k=57	2	0.7833	0.0230
k=74*	3	<b>0.7954</b>	0.0225
k=86	2	0.6117	0.0224
			$\Sigma = 0.0680$

Tabela A.5.56 - Valores e tempos de processamento do índice  $\mathcal{V}_{Conn}$  para as partições do conjunto de dados *Wine*.



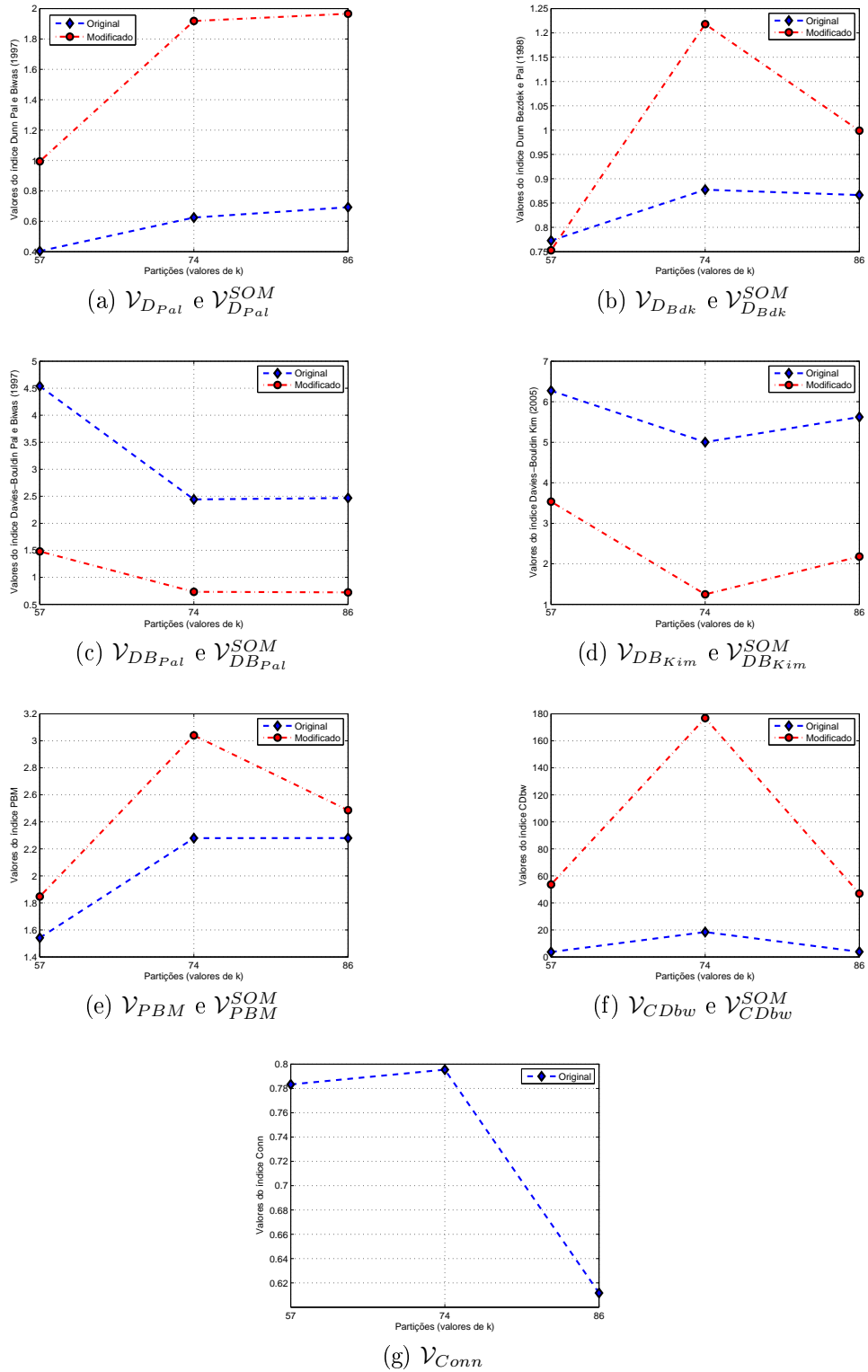


Fig. A.5.8 - Gráficos dos valores das versões originais e modificadas dos índices de validação para as partições do conjunto de dados Wine.

## Conjunto de dados $SR\_1$

Partições	No. de classes	Valor $\mathcal{V}_{DPal}$	Valor $\mathcal{V}_{DPal}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DPal}$	Tempo(s) $\mathcal{V}_{DPal}^{SOM}$	Speedup
k=14	4	0.5012	1.1626	0.6614	2365.7	0.5316	4450.1
k=30*	3	1.5265	3.0473	1.5208	1852.4	1.0537	1757.9
k=125	2	<b>2.6924</b>	<b>4.3707</b>	1.6783	3952.1	1.9374	2039.8
				$\mu = 1.2868$	$\Sigma = 8170.2$	$\Sigma = 3.5227$	2319.3

Tabela A.5.57 - Valores e tempos de processamento de  $\mathcal{V}_{DPal}$  e  $\mathcal{V}_{DPal}^{SOM}$  para as partições do conjunto de dados  $SR\_1$ .

Partições	No. de classes	Valor $\mathcal{V}_{DBdk}$	Valor $\mathcal{V}_{DBdk}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DBdk}$	Tempo(s) $\mathcal{V}_{DBdk}^{SOM}$	Speedup
k=14	4	0.3951	0.3845	0.0105	32.162	0.1928	166.81
k=30*	3	<b>2.5146</b>	<b>2.6489</b>	0.1342	24.523	0.1449	169.24
k=125	2	2.0512	2.1139	0.0626	16.325	0.0880	185.51
				$\mu = 0.0691$	$\Sigma = 73.01$	$\Sigma = 0.4258$	171.46

Tabela A.5.58 - Valores e tempos de processamento de  $\mathcal{V}_{DBdk}$  e  $\mathcal{V}_{DBdk}^{SOM}$  para as partições do conjunto de dados  $SR\_1$ .

Partições	No. de classes	Valor $\mathcal{V}_{DBPal}$	Valor $\mathcal{V}_{DBPal}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DBPal}$	Tempo(s) $\mathcal{V}_{DBPal}^{SOM}$	Speedup
k=14	4	0.3267	0.4958	0.1691	1842.7	0.4551	4.18
k=30*	3	0.3361	0.4918	0.1557	3564.5	1.0196	2.81
k=125	2	<b>0.1458</b>	<b>0.3894</b>	0.2436	2003.5	1.9174	6.23
				$\mu = 0.1894$	$\Sigma = 7410.7$	$\Sigma = 3.3921$	2184.7

Tabela A.5.59 - Valores e tempos de processamento de  $\mathcal{V}_{DBPal}$  e  $\mathcal{V}_{DBPal}^{SOM}$  para as partições do conjunto de dados  $SR\_1$ .

Partições	No. de classes	Valor $\mathcal{V}_{DBKim}$	Valor $\mathcal{V}_{DBKim}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DBKim}$	Tempo(s) $\mathcal{V}_{DBKim}^{SOM}$	Speedup
k=14	2	345.39	325.78	19.613	5.1019	0.0345	147.88
k=30*	3	<b>179.02</b>	<b>153.8</b>	25.225	5.0664	0.0389	130.24
k=125	2	222.58	210.51	12.071	5.0639	0.0374	135.39
				$\mu = 18.97$	$\Sigma = 15.2322$	$\Sigma = 0.1108$	137.47

Tabela A.5.60 - Valores e tempos de processamento de  $\mathcal{V}_{DBKim}$  e  $\mathcal{V}_{DBKim}^{SOM}$  para as partições do conjunto de dados  $SR_1$ .

Partições	No. de classes	Valor $\mathcal{V}_{PBM}$	Valor $\mathcal{V}_{PBM}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{PBM}$	Tempo(s) $\mathcal{V}_{PBM}^{SOM}$	Speedup
k=14	4	424.55	339.45	85.103	5.0960	0.0291	175.12
k=30*	3	743.59	611.23	132.36	4.9721	0.0409	121.56
k=125	2	<b>766.72</b>	<b>678.99</b>	87.73	6.0123	0.0411	146.28
				$\mu = 101.731$	$\Sigma = 16.08$	$\Sigma = 0.1111$	144.73

Tabela A.5.61 - Valores e tempos de processamento de  $\mathcal{V}_{PBM}$  e  $\mathcal{V}_{PBM}^{SOM}$  para as partições do conjunto de dados  $SR_1$ .

Partições	No. de classes	Valor $\mathcal{V}_{CDbw}$	Valor $\mathcal{V}_{CDbw}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{CDbw}$	Tempo(s) $\mathcal{V}_{CDbw}^{SOM}$	Speedup
k=14	4	56.224	832.61	776.39	32.751	0.7463	43.88
k=30*	3	<b>129.86</b>	<b>1046.5</b>	916.64	39.51	0.7244	54.54
k=125	2	90.725	926.59	835.87	45.882	0.7833	58.57
				$\mu = 842.96$	$\Sigma = 118.143$	$\Sigma = 2.2541$	52.41

Tabela A.5.62 - Valores e tempos de processamento de  $\mathcal{V}_{CDbw}$  e  $\mathcal{V}_{CDbw}^{SOM}$  para as partições do conjunto de dados  $SR_1$ .

Partições	No. de classes	Valor $\mathcal{V}_{Conn}$	Tempo(s) $\mathcal{V}_{Conn}$
k=14	4	0.7280	51.837
k=30*	3	0.2466	51.717
k=125	2	<b>0.9234</b>	51.714
			$\Sigma = 155.27$

Tabela A.5.63 - Valores e tempos de processamento do índice  $\mathcal{V}_{Conn}$  para as partições do conjunto de dados  $SR\_1$ .

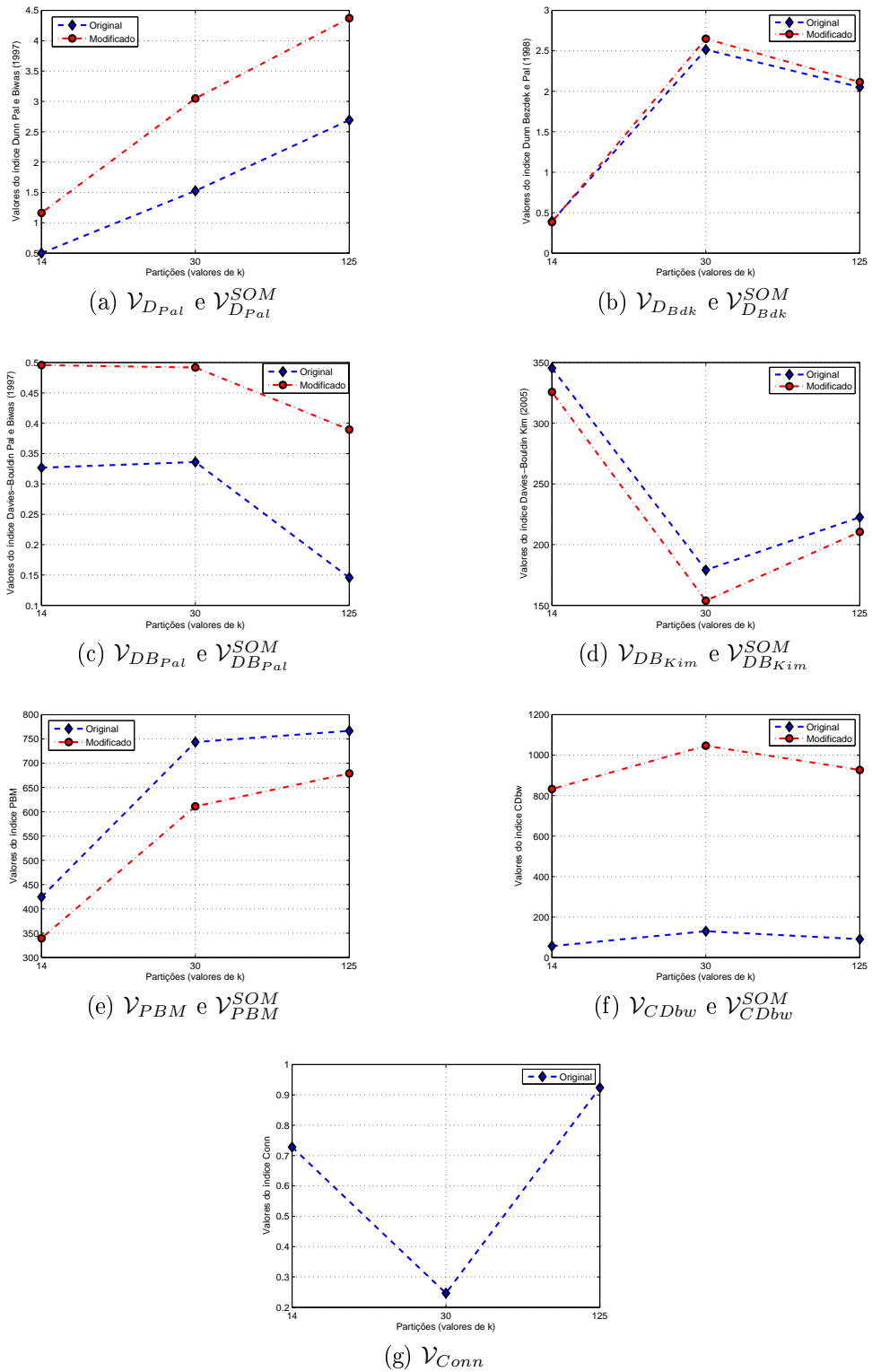


Fig. A.5.9 - Gráficos dos valores das versões originais e modificadas dos índices de validação para as partições do conjunto de dados  $SR\_1$ .

## Conjunto de dados $SR\_2$

Partições	No. de classes	Valor $\mathcal{V}_{DPal}$	Valor $\mathcal{V}_{DPal}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DPal}$	Tempo(s) $\mathcal{V}_{DPal}^{SOM}$	Speedup
k=21	2	0.3169	1.2799	0.963	2987.2	2.0261	1474.35
k=30*	4	0.8509	<b>1.8423</b>	0.9914	3432.7	0.5018	6840.77
k=56	5	0.4213	0.9521	0.5308	3108.5	0.3940	7889.59
k=77	3	<b>0.8976</b>	1.6129	0.7153	3561.9	1.463	2434.65
				$\mu = 0.8$	$\Sigma = 13090.3$	$\Sigma = 4.3849$	2985.3

Tabela A.5.64 - Valores e tempos de processamento de  $\mathcal{V}_{DPal}$  e  $\mathcal{V}_{DPal}^{SOM}$  para as partições do conjunto de dados  $SR\_2$ .

Partições	No. de classes	Valor $\mathcal{V}_{DBdk}$	Valor $\mathcal{V}_{DBdk}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DBdk}$	Tempo(s) $\mathcal{V}_{DBdk}^{SOM}$	Speedup
k=21	2	0.8402	0.8546	0.0144	57.508	0.2575	223.33
k=30*	4	1.0051	1.1528	0.1476	78.648	0.3949	199.15
k=56	5	0.8620	0.9472	0.0851	84.816	0.4274	198.44
k=77	3	<b>1.0973</b>	<b>1.1945</b>	0.0972	60.988	0.2639	231.10
				$\mu = 0.0861$	$\Sigma = 281.96$	$\Sigma = 1.344$	209.79

Tabela A.5.65 - Valores e tempos de processamento de  $\mathcal{V}_{DBdk}$  e  $\mathcal{V}_{DBdk}^{SOM}$  para as partições do conjunto de dados  $SR\_2$ .

Partições	No. de classes	Valor $\mathcal{V}_{DBPal}$	Valor $\mathcal{V}_{DBPal}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DBPal}$	Tempo(s) $\mathcal{V}_{DBPal}^{SOM}$	Speedup
k=21	2	0.7283	1.2688	0.5405	2842.7	1.9706	1442.55
k=30*	4	0.6576	<b>0.9954</b>	0.3378	4064.1	0.4979	8162.48
k=56	5	0.5930	1.1433	0.5503	3103.1	0.3783	8202.7
k=77	3	<b>0.5843</b>	1.0786	0.4943	3603.5	1.4872	2423
				$\mu = 0.4807$	$\Sigma = 13613.4$	$\Sigma = 4.334$	3141.07

Tabela A.5.66 - Valores e tempos de processamento de  $\mathcal{V}_{DBPal}$  e  $\mathcal{V}_{DBPal}^{SOM}$  para as partições do conjunto de dados  $SR\_2$ .

Partições	No. de classes	Valor $\mathcal{V}_{DBKim}$	Valor $\mathcal{V}_{DBKim}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{DBKim}$	Tempo(s) $\mathcal{V}_{DBKim}^{SOM}$	Speedup
k=21	2	750.23	663.63	86.597	6.1835	0.0399	154.97
k=30*	4	444.7	350.25	94.455	6.1850	0.0405	152.71
k=56	5	<b>440.59</b>	<b>343.24</b>	97.35	6.1801	0.0407	151.84
k=77	3	580.87	496.74	84.133	7.5214	0.0393	191.38
				$\mu = 90.634$	$\Sigma = 26.07$	$\Sigma = 0.1604$	162.53

Tabela A.5.67 - Valores e tempos de processamento de  $\mathcal{V}_{DBKim}$  e  $\mathcal{V}_{DBKim}^{SOM}$  para as partições do conjunto de dados  $SR\_2$ .

Partições	No. de classes	Valor $\mathcal{V}_{PBM}$	Valor $\mathcal{V}_{PBM}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{PBM}$	Tempo(s) $\mathcal{V}_{PBM}^{SOM}$	Speedup
k=21	2	219.56	218.52	1.0381	6.1717	0.0465	132.72
k=30*	4	<b>446.14</b>	<b>386.5</b>	59.638	7.1691	0.0444	161.46
k=56	5	364.35	286.67	77.677	7.1702	0.0452	158.63
k=77	3	411.58	379.19	32.388	7.1754	0.0595	120.59
				$\mu = 42.685$	$\Sigma = 27.68$	$\Sigma = 0.1956$	141.51

Tabela A.5.68 - Valores e tempos de processamento de  $\mathcal{V}_{PBM}$  e  $\mathcal{V}_{PBM}^{SOM}$  para as partições do conjunto de dados  $SR\_2$ .

Partições	No. de classes	Valor $\mathcal{V}_{CDbw}$	Valor $\mathcal{V}_{CDbw}^{SOM}$	Erro	Tempo(s) $\mathcal{V}_{CDbw}$	Tempo(s) $\mathcal{V}_{CDbw}^{SOM}$	Speedup
k=21	2	0.8731	3.2377	2.3646	106.03	1.3513	78.46
k=30*	4	<b>56.432</b>	<b>342.45</b>	286.018	62.34	1.2361	50.43
k=56	5	45.948	157.62	111.672	58.41	1.3848	42.17
k=77	3	10.858	21.898	11.04	99.831	1.3058	76.45
				$\mu = 102.77$	$\Sigma = 326.61$	$\Sigma = 5.2779$	61.88

Tabela A.5.69 - Valores e tempos de processamento de  $\mathcal{V}_{CDbw}$  e  $\mathcal{V}_{CDbw}^{SOM}$  para as partições do conjunto de dados  $SR\_2$ .

Partições	No. de classes	Valor $\mathcal{V}_{Conn}$	Tempo(s) $\mathcal{V}_{Conn}$
k=21	2	0.8053	172.86
k=30*	4	0.4802	168.70
k=56	5	0.3487	175.31
k=77	3	<b>0.8520</b>	185.30
			$\Sigma = 702.17$

Tabela A.5.70 - Valores e tempos de processamento do índice  $\mathcal{V}_{Conn}$  para as partições do conjunto de dados  $SR\_2$ .



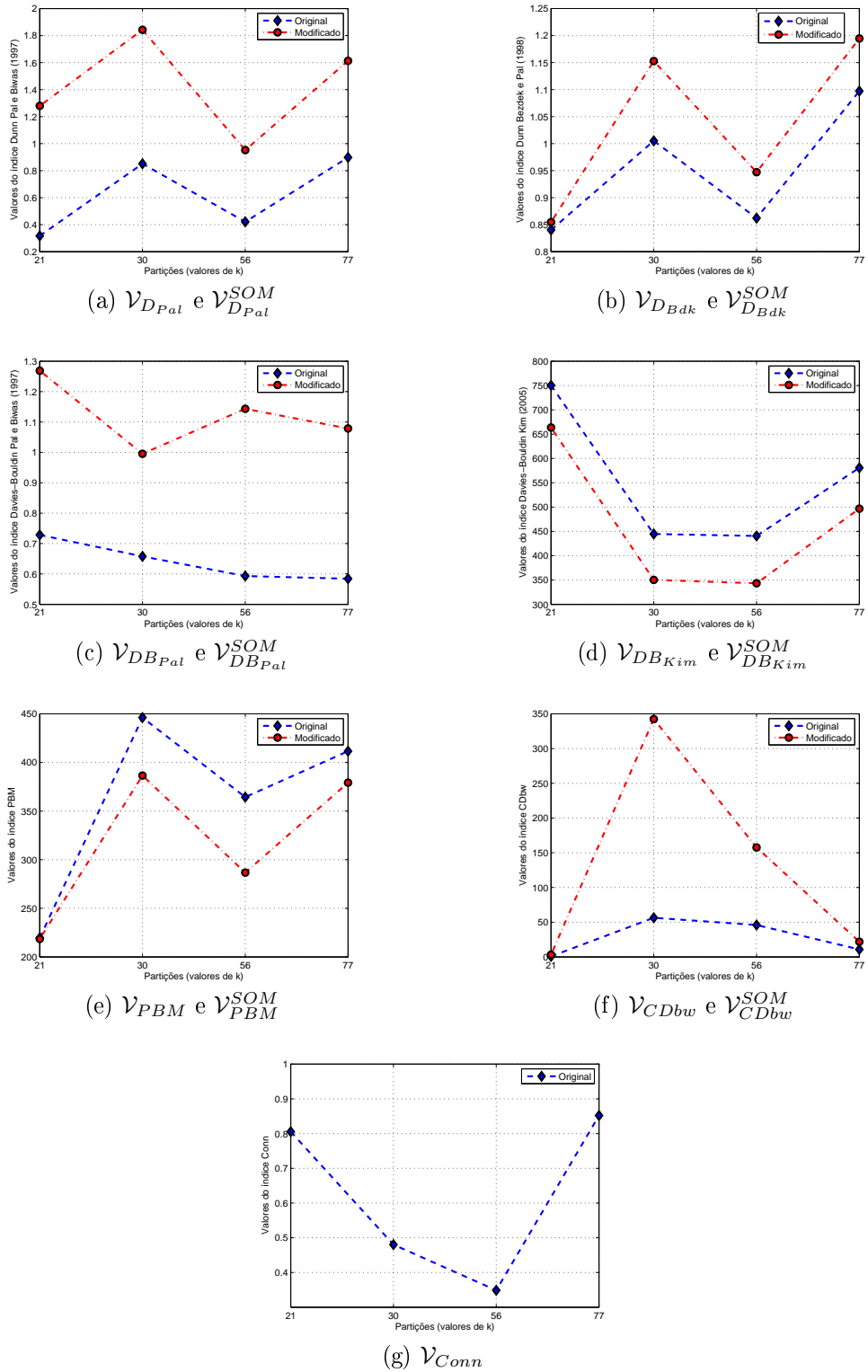


Fig. A.5.10 - Gráficos dos valores das versões originais e modificadas dos índices de validação para as partições do conjunto de dados  $SR\_2$ .

## A.6 - Árvore de mapas da imagem teste 2

Este apêndice mostra a árvore de mapas gerada a partir da aplicação da metodologia de classificação proposta no capítulo 4 sobre a imagem teste 2 (usada nos experimentos apresentados no capítulo 5).

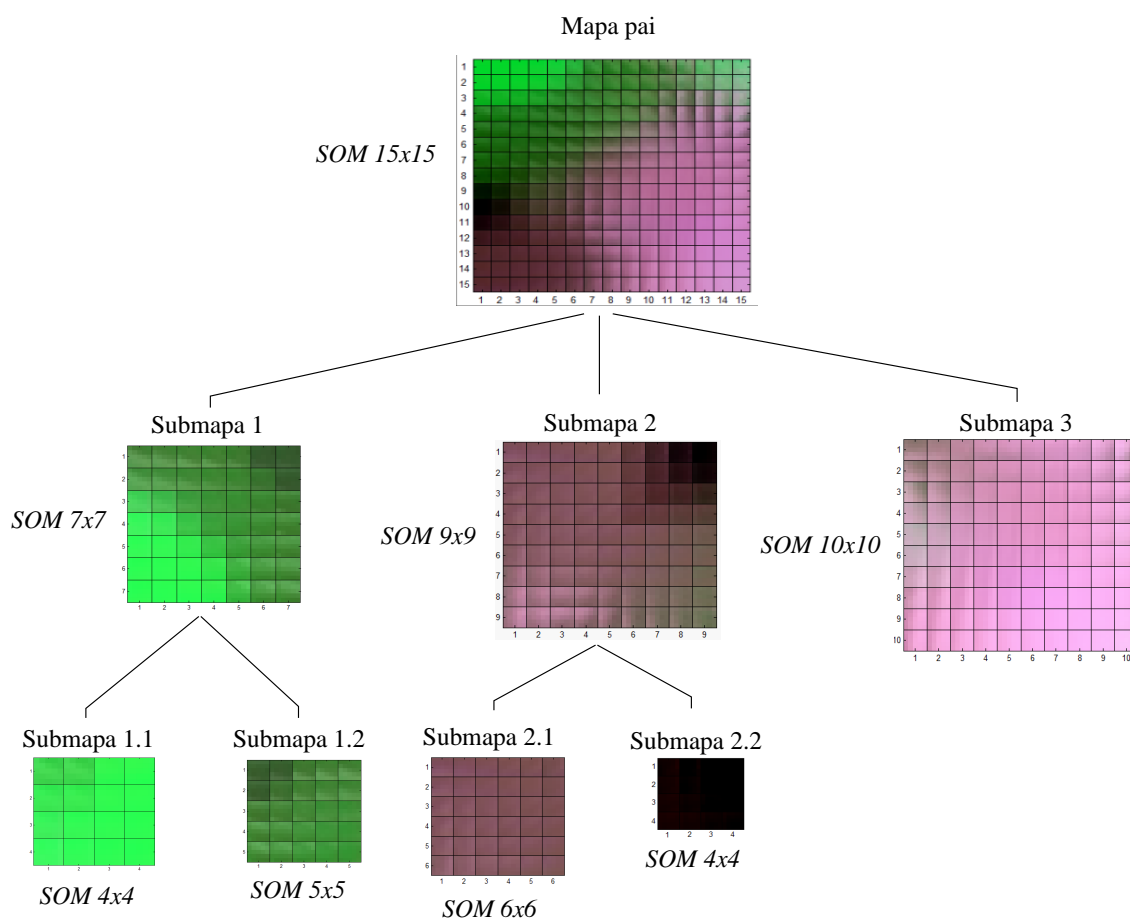


Fig. A.6.1: Árvore de mapas gerada a partir da aplicação da metodologia de classificação proposta no capítulo 4 sobre a imagem teste 2.

## A.7 - Histograma de vencimentos do SOM da imagem teste 2

Este apêndice mostra mais detalhes sobre a possibilidade de exclusão dos protótipos da classe água do SOM da imagem teste 2, caso a estratégia para salientar as bordas da *U-matrix*, proposta por Costa (1999), seja aplicada.

A figura A.7.1 mostra o histograma do SOM da imagem teste 2 (apresentado na figura 5.8). Os números em vermelho mostram as frequências de vencimentos (graus de atividade) dos protótipos considerados heterogêneos e os números em azul mostram a frequências de vencimento dos protótipos correspondentes à classe água. É possível observar que os protótipos heterogêneos aparecem em maior número que os protótipos da classe água. Além disso, pode-se notar também, que os graus de atividade dos protótipos da classe água são mais baixos que os dos protótipos heterogêneos. Vale observar que os protótipos heterogêneos foram determinados aplicando o método proposto na seção 5.5 deste trabalho.

Caso a estratégia para diminuir a degradação das bordas da *U-matrix*, proposta por Costa (1999), fosse aplicada, adotando o valor 10 como grau de atividade mínimo para que os protótipos sejam considerados na análise do mapa, os protótipos da classe água seriam eliminados e, portanto, não seriam detectados na etapa de segmentação do SOM.

As figuras A.7.2 e A.7.3 mostram, respectivamente, o MCI e a *U-matrix* do SOM da imagem teste 2 após o deslocamento dos neurônios com graus de atividade inferiores a 10. Comparando o MCI da figura A.7.2 com o MCI da figura 5.9, nota-se que os protótipos da classe água foram substituídos por protótipos de outras classes. Na imagem da figura A.7.2, observa-se que não há mais o "vale" próximo ao canto inferior esquerdo da superfície topográfica da *U-matrix*, o qual

aparece na imagem da  $U$ -matrix da figura 5.10 e que corresponde exatamente aos protótipos da classe água. O valor 10 considerado aqui, como o grau de atividade mínimo para os neurônios, é aproximadamente 37% inferior ao valor do grau de atividade *médio* dos protótipos do SOM analisado.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	35	18	18	12	21	17	15	11	10	15	16	13	5	6	7
2	15	15	18	17	14	15	17	14	15	12	8	8	4	0	13
3	21	19	16	21	15	13	15	16	12	10	13	14	16	15	23
4	18	17	15	12	14	14	12	22	14	16	15	16	13	16	21
5	19	13	17	15	17	14	17	14	18	17	13	15	12	14	15
6	15	14	18	15	16	18	21	14	16	15	16	14	13	19	17
7	16	16	13	15	17	16	17	15	13	16	13	17	15	13	18
8	18	14	11	15	19	17	16	16	8	17	12	16	17	13	15
9	7	5	19	15	14	17	15	15	16	14	13	16	16	13	14
10	9	6	18	17	17	16	19	13	17	18	12	14	18	17	17
11	6	4	11	21	15	18	14	12	13	17	16	17	14	21	17
12	16	12	15	17	18	17	15	16	12	15	11	12	13	15	16
13	15	15	16	18	21	21	22	21	16	7	0	8	15	16	24
14	16	15	12	13	13	12	18	17	13	18	14	14	18	17	25
15	34	23	18	17	16	19	22	24	19	21	37	36	35	34	33

Fig. A.7.1: Histograma de vencimentos dos neurônios do SOM da imagem teste 2 apresentado na figura 5.9. Os números em vermelho mostram a frequência de vencimentos dos protótipos considerados heterogêneos e os números em azul mostram a frequência dos protótipos correspondentes à classe água.

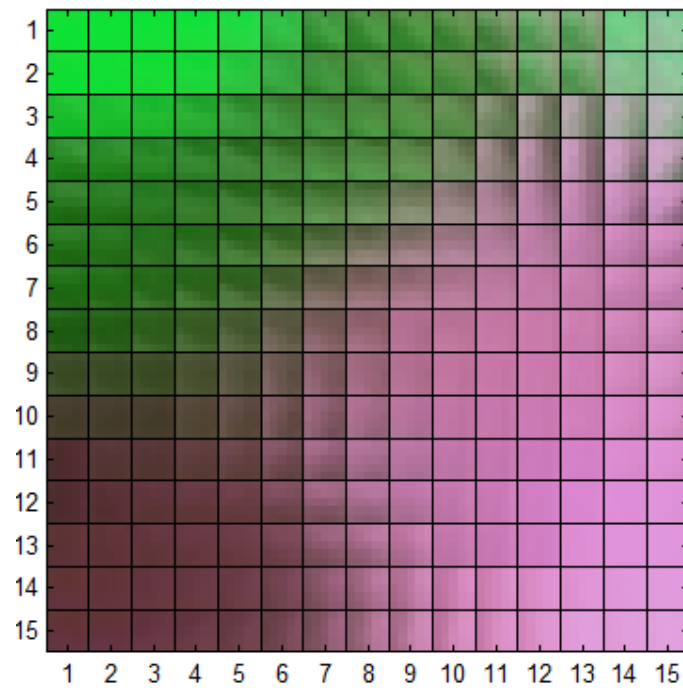


Fig. A.7.2: *MCI do SOM da imagem teste 2 após o deslocamento dos neurônios com graus de ativação inferiores a 10.*

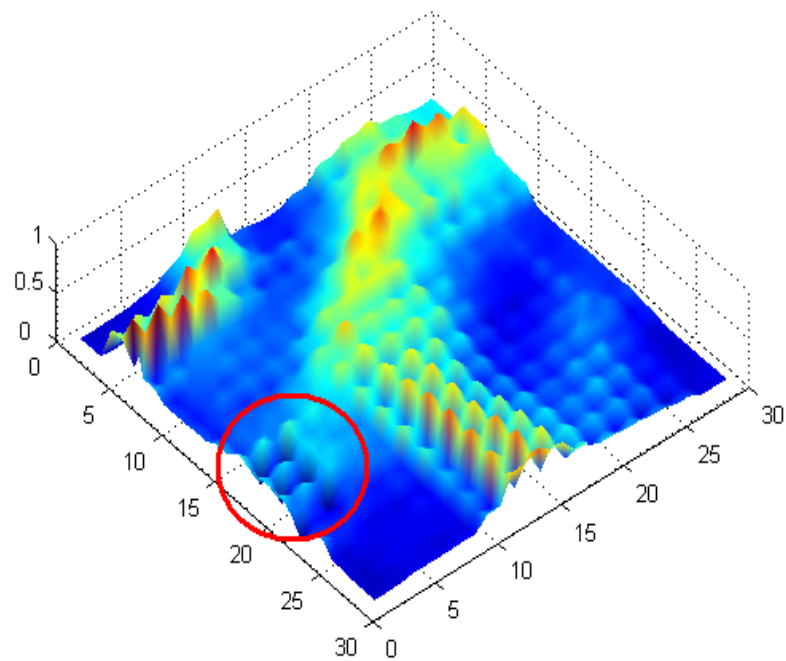


Fig. A.7.3: *U-matrix do SOM da imagem teste 2 após o deslocamento dos neurônios com graus de ativação inferiores a 10. O círculo em vermelho mostra a região onde não há mais o "vale" correspondente ao agrupamento de protótipos da classe água que aparece na U-matrix da figura 5.10.*

# A.8 - Dendrogramas do SOM da imagem IKONOS

Este apêndice mostra, através das figuras A.8.2 e A.8.3, os dendrogramas resultantes de dois métodos hierárquicos com critérios de fusão diferentes, aplicados sobre os protótipos do SOM da imagem IKONOS (apresentada na figura 6.3).

A figura A.8.2 mostra o dendrograma resultante da aplicação do método hierárquico aglomerativo proposto no capítulo 6. Conforme descrito na seção 6.2.2, esse método hierárquico apresenta um critério de fusão que, além de considerar a informação de distância multiespectral para definir qual par de protótipos (ou de conjuntos de protótipos) será agrupado em cada iteração, ele também se baseia em informações espaciais dos pixels da imagem associados aos protótipos do SOM. As informações espaciais são calculadas através da utilização dos índices descritos nas equações 6.1 e 6.2. No dendrograma apresentado na figura A.8.2, destaca-se a partição dos protótipos do SOM no nível 110 (melhor partição de acordo com a versão modificada do índice de validação CDbw).

A figura A.8.3 mostra o dendrograma resultante da aplicação de um método hierárquico aglomerativo que considera em seu critério de fusão apenas a distância multiespectral entre os agrupamentos. Na figura A.8.3 também são indicadas fusões entre classes distintas de protótipos do SOM que comprometem a separação correta das classes de informação da imagem, conforme aquela apresentada no dendrograma da figura A.8.2.

Deve-se observar que, os dendrogramas apresentados nas figuras A.8.2 e A.8.3 utilizam a enumeração dos protótipos do SOM apresentada na figura A.8.1 (mostrada a seguir). Vale lembrar que os protótipos filtrados (em preto, na figura A.8.1) não são considerados na aplicação

do método hierárquico de agrupamentos.

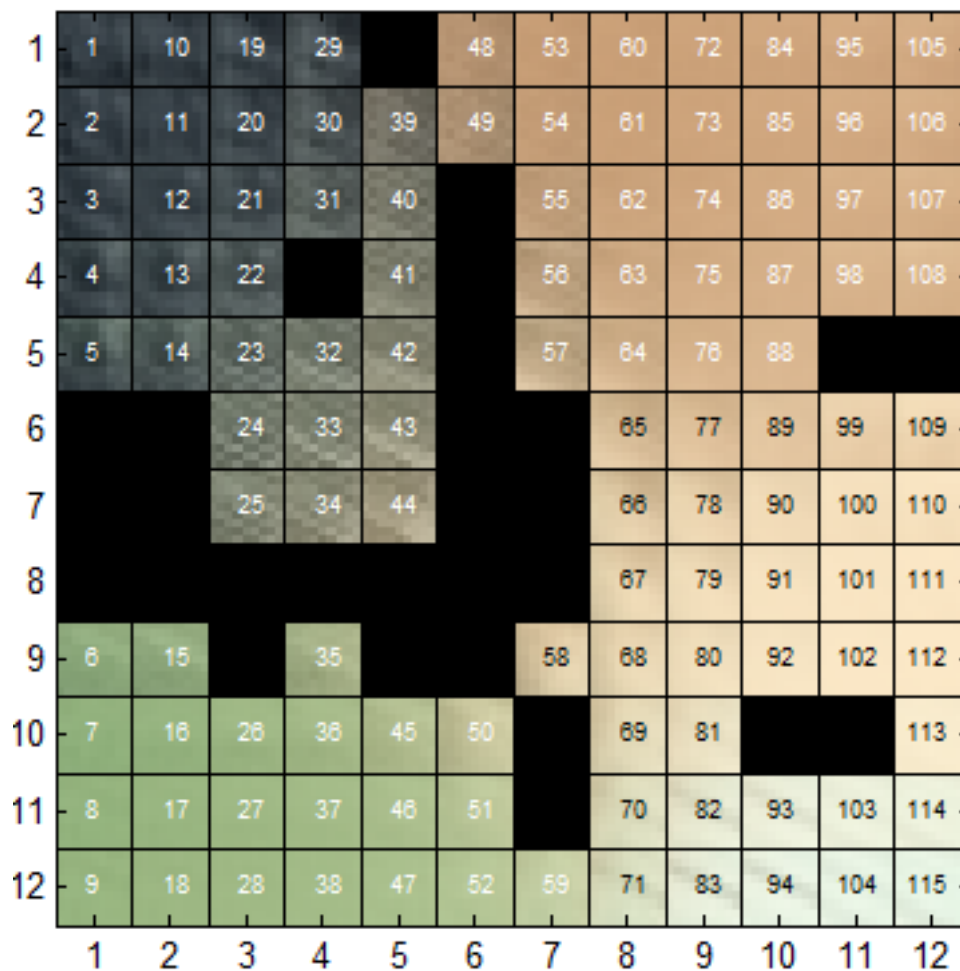


Fig. A.8.1: Protótipos do SOM enumerados. Os quadrados em preto correspondem aos protótipos que foram filtrados pelo método proposto no capítulo 6.



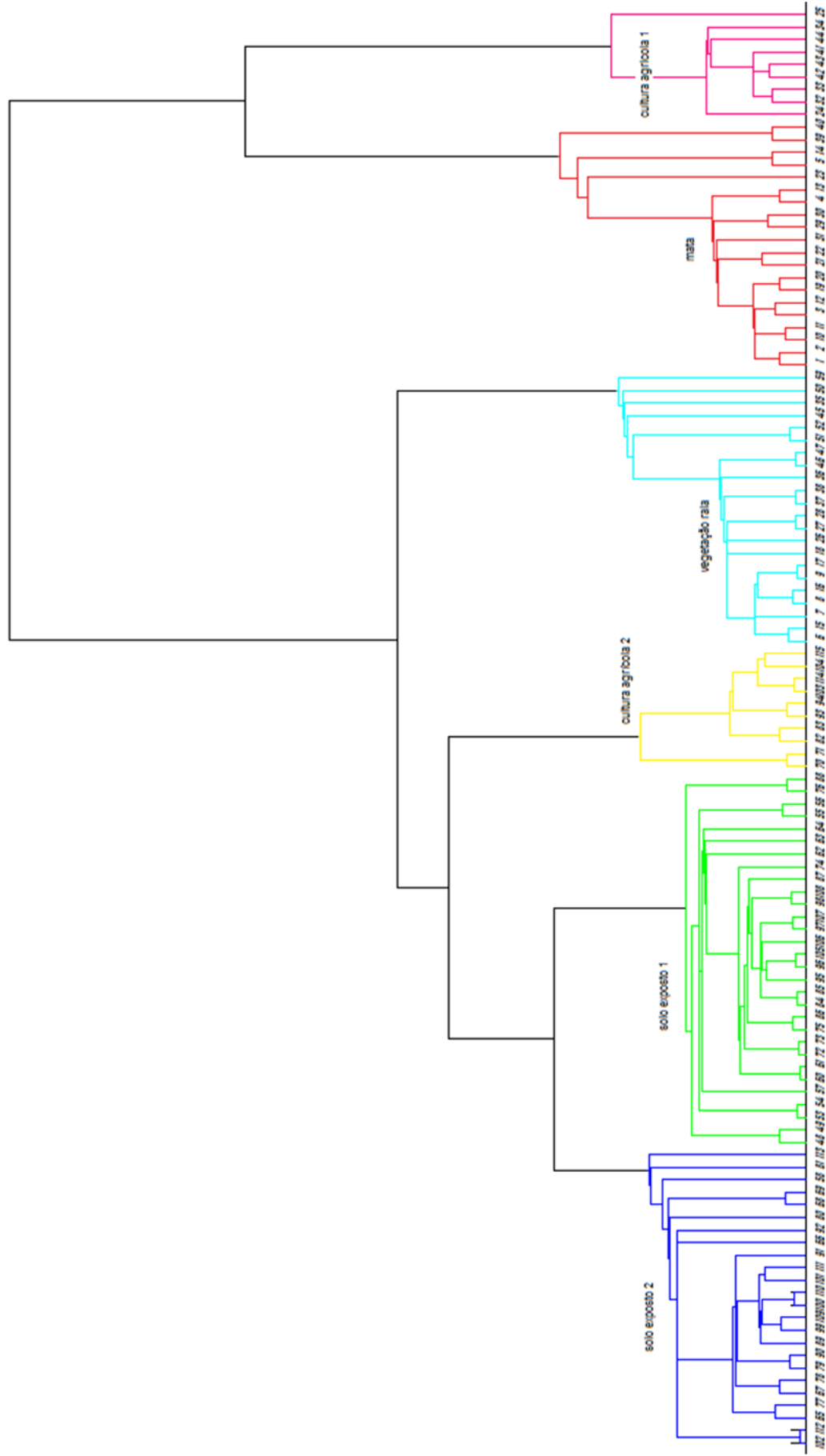


Fig. A.8.2: Dendrograma resultante da aplicação do método hierárquico aglomerativo proposto no capítulo 6 sobre os protótipos do SOM da imagem KONOS. O particionamento (colorido) do dendrograma refere-se à configuração de agrupamentos do nível 110 (melhor partição, segundo a versão modificada do índice CD<sub>bw</sub>). Os números na base do dendrograma correspondem aos protótipos do SOM (enumerados conforme a figura A.8.1).

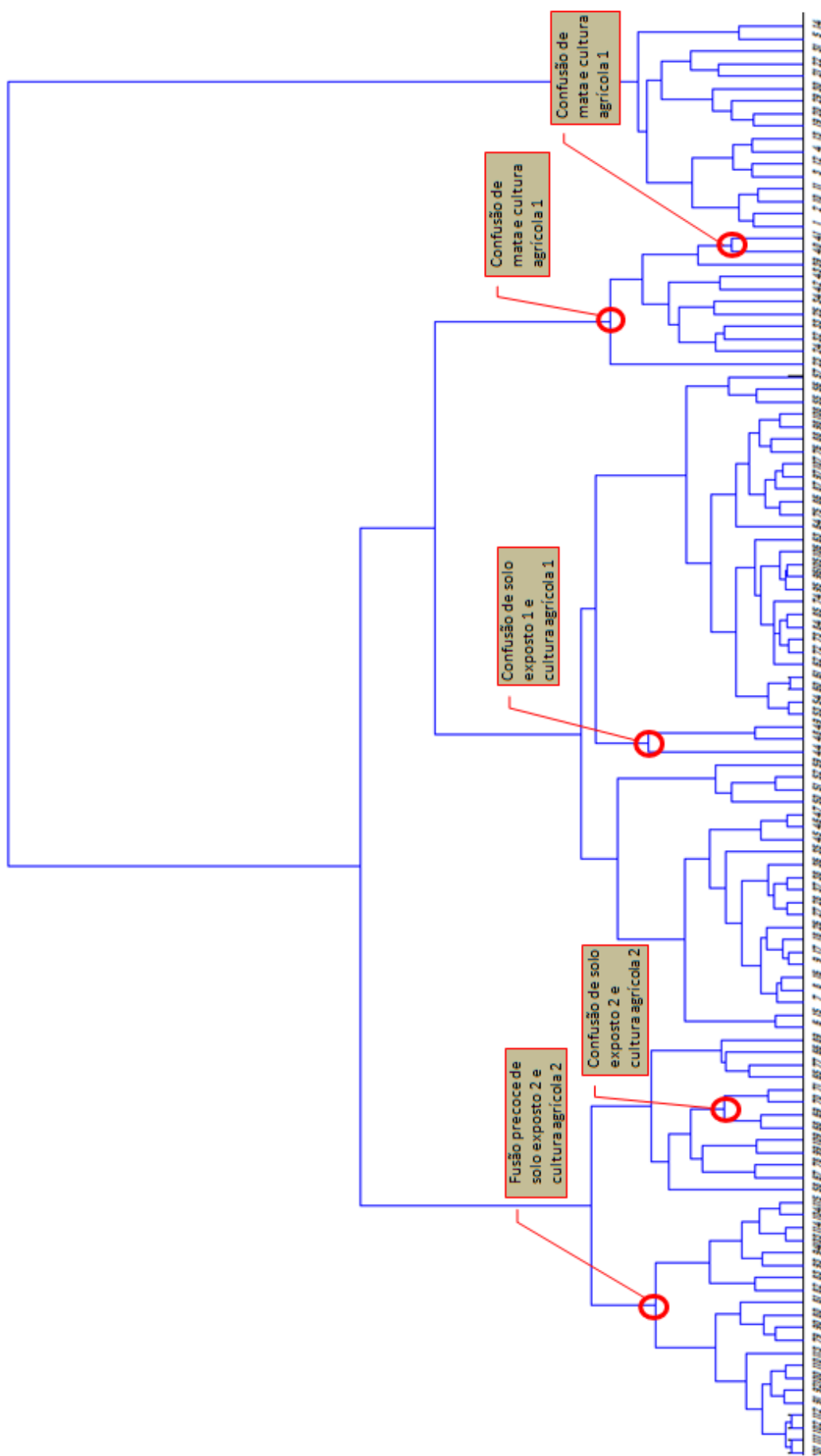
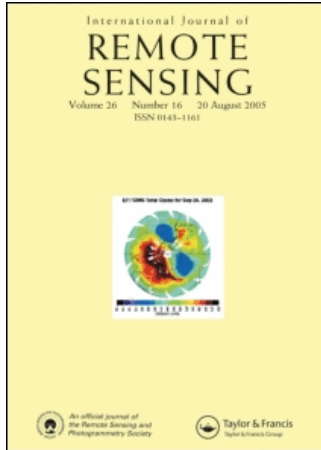


Fig. A.8.3: Dendrograma resultante da aplicação do método hierárquico aglomerativo sobre os protótipos do SOM baseado apenas na distância multiespectral. Os números na base do dendrograma correspondem aos protótipos do SOM (enumerados conforme a figura A.8.1).

This article was downloaded by:[Gonçalves, M. L.]  
On: 31 May 2008  
Access Details: [subscription number 793270168]  
Publisher: Taylor & Francis  
Informa Ltd Registered in England and Wales Registered Number: 1072954  
Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## International Journal of Remote Sensing

Publication details, including instructions for authors and subscription information:  
<http://www.informaworld.com/smpp/title~content=t713722504>

### An unsupervised method of classifying remotely sensed images using Kohonen self-organizing maps and agglomerative hierarchical clustering methods

M. L. Gonçalves<sup>a</sup>; M. L. A. Netto<sup>b</sup>; J. A. F. Costa<sup>c</sup>; J. Zullo Júnior<sup>d</sup>

<sup>a</sup> Department of Computer Science, Pontifical Catholic University of Minas Gerais, Pocos de Caldas, MG, Brazil

<sup>b</sup> School of Electrical and Computer Engineering, State University of Campinas, SP, Brazil

<sup>c</sup> Department of Electrical Engineering, Federal University of Rio Grande do Norte, Natal, RN, Brazil

<sup>d</sup> Agrometeorology Research Center, State University of Campinas, SP, Brazil

First Published: June 2008

To cite this Article: Gonçalves, M. L., Netto, M. L. A., Costa, J. A. F. and Zullo Júnior, J. (2008) 'An unsupervised method of classifying remotely sensed images using Kohonen self-organizing maps and agglomerative hierarchical clustering methods', International Journal of Remote Sensing, 29:11, 3171 — 3207

To link to this article: DOI: 10.1080/01431160701442146

URL: <http://dx.doi.org/10.1080/01431160701442146>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article maybe used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

## An unsupervised method of classifying remotely sensed images using Kohonen self-organizing maps and agglomerative hierarchical clustering methods

M. L. GONÇALVES\*†, M. L. A. NETTO‡, J. A. F. COSTA§ and  
J. ZULLO JÚNIOR¶

†Department of Computer Science, Pontifical Catholic University of Minas Gerais, Av. Padre Francis Cletus Cox, 1661, 37701-355, Poços de Caldas, MG, Brazil

‡School of Electrical and Computer Engineering, State University of Campinas, SP, Brazil

§Department of Electrical Engineering, Federal University of Rio Grande do Norte, Natal, RN, Brazil

¶Agrometeorology Research Center, State University of Campinas, SP, Brazil

(Received 26 January 2007; in final form 3 May 2007)

Unlike conventional unsupervised classification methods, such as *K*-means and ISODATA, which are based on partitional clustering techniques, the methodology proposed in this work attempts to take advantage of the properties of Kohonen's self-organizing map (SOM) together with agglomerative hierarchical clustering methods to perform the automatic classification of remotely sensed images. The key point of the proposed method is to execute the cluster analysis process by means of a set of SOM prototypes, instead of working directly with the original patterns of the image. This strategy significantly reduces the complexity of the data analysis, making it possible to use techniques that have not normally been considered viable in the processing of remotely sensed images, such as hierarchical clustering methods and cluster validation indices. Through the use of the SOM, the proposed method maps the original patterns of the image to a two-dimensional neural grid, attempting to preserve the probability distribution and topology of the input space. Afterwards, an agglomerative hierarchical clustering method with restricted connectivity is applied to the trained neural grid, generating a simplified dendrogram for the image data. Utilizing SOM statistic properties, the method employs modified versions of cluster validation indices to automatically determine the ideal number of clusters for the image. The experimental results show examples of the application of the proposed methodology and compare its performance to the *K*-means algorithm.

### 1. Introduction

Since the first satellites were launched for the purpose of searching for terrestrial resources, digital classification methods of remotely sensed images have acquired a growing importance in the automatic recognition of land cover patterns (Richards 2005). The general objective of the classification procedures is to categorize each pixel of an image into one of various land cover classes or themes. The resultant classified image is essentially a thematic map of the original image.

---

\*Corresponding author. Email: marcio@pucpcaldas.br

At present, while there is no substitute for a trained analyst, the enormous quantity of images that are being collected by sensor systems, which are more and more sophisticated and modern, require the development of innovative classification methodologies that allow an automatic and efficient identification of the great volume of data available in the images and which, at the same time, make the mapping process of terrestrial surfaces less subjective with greater potential for reuse in subsequent situations.

### 1.1 *Background on methods of classifying remotely sensed images*

In the literature, there are a large number of proposed algorithms for the classification of remotely sensed images. Different techniques derived from various research areas are being used as a basis for the development of these algorithms, for example, statistics (Dean and Smith 2003), neural networks (Atkinson and Tatnall 1997), classification trees (Hansen *et al.* 1996), fuzzy logic (Wang 1990), genetic algorithms (Pal *et al.* 2001), and support vector machines (Mitra *et al.* 2004).

Although the literature shows an abundance of approaches, some studies reveal that remote sensing research related to classification is still in need of more significant advances. In Wilkinson (2005), results and implications of a fifteen-year study of satellite image classification experiments show that, whilst a quantity of work demonstrates the advantages of one technique compared to another, the overall scenario reveals that no significant upward trend in results can be identified in the hundreds of experiments examined over the period observed. Moreover, as mentioned in Liu *et al.* (2002), the literature has also shown that no image classifier is perfect or has good performance for all types of remote sensing images. In fact, different classification algorithms generate different results even with the same data sets (Gonçalves 1998, Pal and Mather 2002, Pasquariello *et al.* 2002, Sunar Erbek *et al.* 2004). In view of that, and of the recent advances in remote sensor technology, efforts continue to be made in improving classification performance, and therefore research and proposals for different methods of classifying remotely sensed images continues.

From the point of view of a learning strategy, in traditional remote sensing the techniques for classification of images broadly fall into three categories: supervised, unsupervised, and hybrid (Lillesand and Kiefer 2000). Despite the significant number of different proposals, most of them perform the classification of images using the supervised approach. In this type of classification, human interaction allows the training of computational algorithms through the selection of samples of pixels for each class of interest. Conventional statistical classifiers such as the maximum likelihood method and multilayer perceptron neural networks are among the more popular supervised methods (Frizzelle and Moody 2001, Dean and Smith 2003, Sunar Erbek *et al.* 2004). However, in more recent literature there is a growing increase in the sophistication of supervised classification methods, including the use of some techniques that were not widely used within the remote sensing community. Lawrence *et al.* (2004), for example, explored advanced statistical approaches in the classification of remotely sensed data sets with a wide diversity of spectral, spatial and land cover patterns. Using stochastic gradient boosting as a refinement of classification tree analysis, the proposed method was more accurate than traditional classification tree methods, although the results appear to depend on the specific data. In Pal and Mather (2005), a support vector machine is considered for supervised pixel classification and the results show that the proposed algorithm

achieves a higher level of classification accuracy than maximum likelihood or neural classifiers. Assuming that different classifiers have complementary capabilities, several proposals of combined methods that take advantage of different algorithms have proven to be a powerful strategy in the increase of classification accuracy (Liu *et al.* 2002, Zhong *et al.* 2005, Stathakis and Vasilakos 2006). Notably, the combination of ensembles of supervised neural networks has also been proposed as an efficient approach to the development of high performance image classification systems. The experimental results reported in the literature show that neural network ensembles outperform the accuracy of the best single net (Giacinto *et al.* 2000, Espinosa *et al.* 2004, Van Coillie *et al.* 2004).

Although the supervised classification approach is the most commonly used, it must be emphasized that its performance strongly depends on the quality of the training samples. Each sample must be precisely chosen in such a way that it is the best possible representative for a class of interest; otherwise the classification results can contain errors or be imprecise (Swain and Davis 1978). As mentioned in Mitra *et al.* (2004), the primary problem in supervised classification is the pure availability of training samples, which can be obtained only from ground truths and by costly manual labelling. Furthermore, differences in training data sets in supervised classification can produce differences in classification results, mainly at high spatial resolutions (Chen and Stow 2002). For these reasons, a number of strategies to intelligently select and filter useful training samples from remotely sensed images have been reported in the literature (Arai 1992, Foody 1999, Foody and Mathur 2004). In particular, the active learning paradigm has been used extensively to reduce the data requisite of learning tasks. Instead of passively learning from random samples, the learner has the ability to iteratively select its own training data (Cohn *et al.* 1994, Bachmann 2003, Mitra *et al.* 2004). Alternatively, hybrid classification methods that combine the supervised and unsupervised approaches are also utilized to facilitate the identification of classes and the acquisition of training samples (Lillesand and Kiefer 2000). Most hybrid methods perform an initial stratification of the image by spectral clustering, associate the revealed clusters to user-defined classes, and then apply a supervised method to classify the entire image. The preceding use of an unsupervised method using clustering can help the analyst to identify classes that appear at a very low density in the image. These classes probably will not be recognized in an exclusively supervised classification process, especially in the case of images that have a very complex variation of patterns where there is no information regarding the terrestrial truth (Gonçalves *et al.* 1996, Kelly *et al.* 2004, Lo and Choi 2004).

Unlike supervised classification, the unsupervised classification of remote sensing images requires minimal initial training data and is based on the principle that the computational algorithm is capable of identifying by itself the classes of the image. This approach of classification plays an especially important role when very little *a priori* information about data is available, and for that reason continues to be a popular choice for analysts without ample field knowledge or for those wanting to avoid introduced bias in classification analysis (Duda and Canty 2002, Wang and Zhou 2002, Kelly *et al.* 2004).

Unsupervised classification is frequently performed through clustering methods. These methods examine the unknown pixels in an image and incorporate them into a set of classes defined through the natural clusters of the grey levels of the pixels. The basic premise is that the pixels that belong to the same land cover class have

similar attributes, while the pixels belonging to different classes are comparatively well separated. Cluster analysis provides a practical method for organizing a large set of data so that the retrieval of information may be made more efficiently. However, although there is a large quantity of different clustering methods in the pattern recognition area (Xu and Wunsch II 2005), only a limited quantity of them can be used in remote sensing applications. As pointed out in Tran *et al.* (2005), several problems are encountered when clustering remotely sensed images, but above all the image size and the feature dimension problems are those that often make a method inappropriate due to computation time and computer memory.

The most common clustering method applied to remotely sensed data is partitional. The majority of software or computational systems meant for the digital processing of remotely sensed images specifically contain two partitional clustering algorithms to perform unsupervised classifications: the  $K$ -means algorithm, and a variation of it, the ISODATA (Ball and Hall 1967). Despite being widely used, these partitional clustering methods have various limitations. The objective functions that they use begin with the assumption that the number of classes,  $K$ , is known *a priori*. In the hypothesis that an inadequate  $K'$  value has been chosen, the method will impose, through the use of optimization techniques,  $K'$  clusters to the data. The user must also manually specify various parameters in order to control the clustering process, for example, the initial centroids of each cluster, the maximum number of iterations, thresholds to perform the division, fusion, or exclusion of clusters.  $K$ -means and ISODATA are very sensitive to these parameters, which can generate different partitions when various simulations are carried out for the same data set. Facing this, the optimal value of these parameters is frequently encountered through trial and error. These needs certainly increase the level of interaction between the user and the computational algorithm, consequently increasing the degree of subjectivity of the categorization process of the image. Other, no less important limitations of partitional algorithms, such as  $K$ -means and ISODATA, are the high computational cost when the data to be analysed is very large (at each iteration, all of the pixels in the image are compared with all of the clusters centroids) and the existence of suppositions about the cluster forms. Generally only one prototype (centroid) is used to represent a cluster, thus these methods become adequate only for the analysis of clusters that have hyperspherical formats (Shah *et al.* 2004, Gonçalves *et al.* 2005).

Advances in  $K$ -means algorithms and attempts to diminish some of its main disadvantages can be encountered in the literature. Su and Chou (2001) proposed a modified version of the  $K$ -means algorithm that adopts a non-metric distance measure based on the idea of 'point symmetry'. The proposed algorithm can be used to group a given data set into a set of clusters of different geometrical structures, but the trade off for the flexibility in detecting clusters is an increase of computational complexity. In Wagstaff *et al.* (2001) a general method is proposed for incorporating background knowledge in the form of instance-level constraints into the  $K$ -means clustering algorithm. Experiments with random constraints on six data sets show significant improvements in accuracy. Likas *et al.* (2003) proposed a global  $K$ -means algorithm consisting of a series of  $K$ -means clustering procedures with the number of clusters varying from 1 to  $k$ . The authors claimed that the algorithm is independent of any starting conditions and also proposed accelerating strategies. However, the problem of computational complexity comes up due to the necessity of executing  $K$ -means a number of times for each value of  $k$ . Other

recent studies on  $K$ -means can be found in Khan and Ahmad (2004), Huang *et al.* (2005), and Chung and Lin (2007).

Another possible, although uncommon, way of performing unsupervised classification of remotely sensed images is through hierarchical clustering methods. Unlike partitional methods, hierarchical methods do not require the user to specify the number of clusters and other additional parameters beforehand, and therefore are not affected by initialization and local minima. Another significant advantage of these methods is that they make it possible to visualize the result of the classification by means of a dendrogram that illustrates in a hierarchical form the degree of similarity between the clusters that are formed by fusions (or divisions) at each successive stage of the analysis. However, hierarchical methods have some characteristics that prevent their application in the classification of remotely sensed images: (a) in general they require memory space of the order of  $O(N^2)$ , in which  $N$  is the number of records in the data set; (b) dendrograms are impractical for large data sets, and therefore the results can be difficult to interpret; (c) in order to determine the cut-off of the dendrogram (ideal number of clusters), some decision criteria must be applied (Tran *et al.* 2005, Xu and Wunsch II 2005).

Consequently, hierarchical clustering modules are rare in remote sensing software packages and for this reason are normally utilized in an uncoupled way. In the literature, there is not much research that makes use of hierarchic methods in the unsupervised classification of remotely sensed images, and, in general, common versions of these methods are applied in practice when the main objective is basically to compare performance with other unsupervised methods (Duda and Canty 2002, Wilson *et al.* 2002). However, a relatively recent progression of approaches attempting to make the use of hierarchical clustering methods viable in remote sensing applications can be seen. For instance, instead of applying agglomerative hierarchical methods directly on a pixel by pixel basis, some work has employed these methods only after the application of some other clustering method on the original image data. In Tran *et al.* (2003) the  $K$ -means algorithm was used to classify the image data before the application of a hierarchical method, and the ISODATA was used for the same reason in Marçal and Castro (2005) and in Marçal and Borges (2005). Although this strategy is effective at preventing the (expensive) application of hierarchic methods to large data sets, a possible disadvantage of it is that the results of the classification methods presented in these works can be influenced by the limitations of the  $K$ -means and ISODATA algorithms, which were mentioned earlier.

Clearly, the literature contains other proposals for unsupervised methods to classify remote sensing data that do not make use only of conventional clustering approaches. In Huang (2002), for example, a divisive hierarchical method based on concepts of hyperplanes and binary trees was proposed to serve as an alternative to ISODATA. Although the overall accuracy of the proposed hierarchical algorithm was slightly lower, it was much faster than the ISODATA in all the experimental tests that have been presented. Shah *et al.* (2004) proposed an unsupervised algorithm derived from independent component analysis that models class distributions with non-Gaussian densities and employs higher order statistics as opposed to conventional algorithms. Whilst the algorithms are highly complex and data processing is time consuming, unsupervised classification techniques based on Markov random fields that combine spectral and spatial information have also been investigated (Baraldi *et al.* 2000, Tso and Olsen 2005). In addition, researchers have



projected several unsupervised approaches to classify remote sensing images based on computational intelligence methods like neural networks, fuzzy systems, and evolutionary algorithms (Ji 2003, Bandyopadhyay 2005, Liu *et al.* 2005, Xiang *et al.* 2005).

In particular, a number of studies have investigated the use of the Kohonen neural network (more commonly known as SOM – self-organizing map) for the task of clustering and classification of remotely sensed images. Important properties such as the input space approximation, topological ordering and density matching, allied with the simplicity of the model and the ease of implementation of the learning algorithm make SOM a promising tool for this type of application. Ji (2000) showed that the SOM algorithm is capable of achieving higher classification accuracy as compared to that of the maximum-likelihood method and concluded that the neural network is a useful viable alternative for the classification of remotely sensed images. Vilmann *et al.* (2003) studied the effect of the application of several extensions of the SOM in the cluster analysis of multispectral and hyperspectral images. The authors showed excellent data clustering and classification, and concluded that some considerations about the consequence of the dimensionality reduction produced by SOM must still be investigated. In Hadjimitsis *et al.* (2003), an unsupervised classification method using SOM was proposed for mapping land cover changes. Preliminary results show that utilizing SOM for image classification is efficient and opens up significant areas for further research.

Although SOM has been considered as an effective tool for clustering in remote sensing images, it is worth noting that the detection of cluster boundaries through the output space of the network is not a trivial task. Moreover, as described in Ji (2000), the process of segmentation of the trained neural map can affect representation of the classes. Consequently, the majority of the classification methods based on SOM carry out the cluster detection with the help of an image analyst that manually segments the neural map using *a priori* knowledge of the data and some alternative tool to make cluster visualization easier in the SOM (Merkl and Rauber 1997). In this way, the analyst interacts with the cluster analysis process by visualizing and interpreting the trained map. However, in unsupervised classification approaches, which have the basic objective of discovering the inherent structure of the data, the detection of clusters through SOM neurons must be carried out without user interference. In the literature there are some proposed algorithms that seek to automatically interpret the neurons of a trained SOM (Costa and Netto 2001, Kiang 2001, Wu and Chow 2004, Gonçalves *et al.* 2006). Using this strategy, Sezgin *et al.* (2004) proposed an unsupervised learning algorithm for the classification of remote sensing images that first uses the SOM to obtain a large number of clusters, and then applies the *K*-means clustering algorithm to automatically segment the trained SOM, and to obtain the final clusters. Although the proposed method has used a clustering validity method to estimate the best number of clusters, approaches that use the *K*-means algorithm for SOM clustering are only feasible for hyperspherical-shaped clusters (Wu and Chow 2004).

## 1.2 Aim of this work

Unlike conventional unsupervised classification methods, such as *K*-means and ISODATA that are based on partitional clustering techniques, and presuming that different algorithms have complementary capabilities, the present work offers an unsupervised method of classifying remotely sensed images that attempts to exploit

the properties of Kohonen's self-organizing map (SOM) along with agglomerative hierarchical clustering methods. In the proposed methodology, the SOM is used to map the original patterns of the image to a two-dimensional neural grid. The objective is to quantize and represent the image patterns in a space of smaller dimension, seeking to preserve the probability distribution and topology of the input space. Afterwards, an agglomerative hierarchical clustering method is applied over the pre-trained SOM neural grid, generating a dendrogram of clustered neurons with different degrees of similarity. Each level of dendrogram obtained corresponds to a different clustering configuration of SOM neurons that can be utilized to represent the classes by which the original image will be classified.

In addition to enabling the use of a hierarchical method, the idea to represent the image through SOM neurons allows the discovery of clusters with complex and varied geometries, considering that the image classes are represented by groups of neurons rather than a single prototype. In addition to this, attempting to exploit all the properties of SOM and reduce the computational cost of the classification process even more, the proposed methodology employs a hierarchical clustering algorithm with restricted connectivity and clustering validation indices calculated in a modified manner to determine the ideal number of classes for the image.

Seeking to guarantee even greater efficiency in the classification process, the proposed method extracts information from the image by means of pixel windows, in order to incorporate contextual information. Following this approach, before the hierarchical clustering algorithm is applied, the method filters heterogeneous SOM prototypes, which represent patterns (pixel windows) corresponding to transition regions between different land cover classes. The patterns that were originally associated with these heterogeneous prototypes are particularly classified only at the end of the process, considering the neighbouring pixels that have already been labelled.

The remainder of the article is organized in the following form: §2 briefly describes the SOM and its properties, §3 presents the fundamentals of the agglomerative hierarchical clustering methods, while §4 explains the proposed classification methodology. Experimental results showing a comparison of the proposed method with the *K*-means algorithm are shown in §5, and §6 gives the conclusions and future perspectives.

## 2. The SOM

The SOM is certainly one of the most important models of neural networks at present, and is used in a variety of applications such as image processing (Ji 2000, Vilmann *et al.* 2003), precipitation estimations (Sorooshian *et al.* 2000, Hong *et al.* 2005, Hong *et al.* 2006), wind circulation (Cassano *et al.* 2006, Fayos and Fayos 2007), and seismic interpretation (Strecker and Uden 2002, Coléou *et al.* 2003).

The SOM is a type of artificial neural net based on competitive and unsupervised learning, i.e. no information about the input signal classes is used in the adjustment process for the synaptic weight in the net (Kohonen 1997). Despite having been developed to try to model the sensory areas of the human cerebral cortex and the lateral iterations between the neurons in such structures, the SOM model is considered extremely simple and easy to implement.

The network essentially consists of two layers: an input layer *I* and an output layer *U* with neurons generally organized in a two-dimensional topological array. The input to the net corresponds to a *p*-dimensional vector, *x*, generally in the space

$\mathfrak{R}^p$ . All of the  $p$  components of the input vector feed each of the neurons on the map. Each neuron  $i$  can be represented by a synaptic weight vector  $\mathbf{w}_i = (w_{i1}, w_{i2}, \dots, w_{ip})$ , also in the  $p$ -dimensional space. For each input pattern  $\mathbf{x}$  a winner neuron,  $c$ , is chosen, using the criterion of greatest similarity:

$$\|\mathbf{x} - \mathbf{w}_c\| = \min_i \|\mathbf{x} - \mathbf{w}_i\|, \quad (1)$$

where  $\|\cdot\|$  represents the Euclidian distance. The winner neuron ( $c$ ) weights, together with the weights of the neighbouring neurons, are adjusted according to the following equation:

$$\mathbf{w}_i(t+1) = \mathbf{w}_i(t) + h_{ci}(t)[\mathbf{x}(t) - \mathbf{w}_i(t)], \quad (2)$$

where  $t$  indicates the iteration of the training process,  $\mathbf{x}(t)$  is the input pattern and  $h_{ci}(t)$  is the nucleus of the neighbourhood around the winner neuron  $c$ . This last term is a decreasing function of time and distance between the neuron  $i$  and the winner neuron  $c$ , generally corresponding to the product of two components, the learning rate  $\alpha(t)$  and the neighbour function  $h$ :

$$h_{ci}(t) = \alpha(t)h(\|r_c - r_i\|, t), \quad (3)$$

where  $r_i$  and  $r_c$  are the positions of the neuron  $i$  and the winner neuron  $c$  in the topological array of the net respectively.

The SOM has some important properties. Once the training algorithm has converged, the computed feature map displays important statistical characteristics of the input space, summarized as follows (Haykin 1994):

- (i) *Approximation of the input space*: the basic objective of SOM is to store a large set of input vectors by finding a smaller set of prototypes that provides a good approximation to the original input space. This strategy is based on vector quantization theory, the motivation for which is dimensionality reduction, or data compression.
- (ii) *Topological ordering*: the features map computed by SOM is ordered topologically. Similar input vectors are mapped close to each other, while dissimilar ones are mapped far apart.
- (iii) *Density matching*: the SOM reflects the probability distribution of data in the input space. Regions in the input space in which the input patterns are taken with a high probability of occurrence are mapped onto larger domains of the output space, and thus have better resolution than regions in the output space from which input patterns are taken with a low probability of occurrence.

### 3. Agglomerative hierarchical clustering methods

In cluster analysis, two classes of methods prevail: the hierarchical and partitional methods. The partitional methods produce a partition of data in  $K$  clusters generally optimizing an objective function. Meanwhile, the hierarchical methods produce a succession of partitions, each one corresponding to a different number of clusters.

The most common of the hierarchical methods, the agglomerative techniques begin the data analysis considering the existence of  $N$  clusters, each one having exactly a single object of the data set. Successive fusion operations between the clusters are then performed until there is only a single cluster containing all of the  $N$

objects. The steps of an agglomerative hierarchical method can be described as follows:

1. Begin with  $N$  clusters,  $C_1, C_2, \dots, C_N$ , each one containing exactly a single object.
2. Determine the pair of distinct clusters  $(C_i, C_j)$ , so that:

$$D(C_i, C_j) = \min_{\substack{1 \leq m, l \leq N \\ m \neq l}} D(C_m, C_l), \tag{4}$$

where  $D(*,*)$  is a distance function.

3. Form a new cluster  $C_k$  through the unification of clusters  $C_i$  and  $C_j$ , i.e.  $C_k = C_i \cup C_j$ .
4. Calculate the new distances,  $D(C_k, C_l)$ , between the new cluster  $C_k$  and the others that are left.
5. Repeat steps 2, 3, and 4 until all of the objects are in a single cluster.

There are many variants of the agglomerative hierarchical methods based on different ways of calculating the distance  $D(*,*)$  between the cluster formed ( $C_k$ ) and all of the other clusters. The following formula, proposed by Lance and Williams *et al.* (1967), allows the different hierarchical clustering methods to be generalized:

$$D(C_i \cup C_k, C_j) = \alpha_i D(C_i, C_j) + \alpha_j D(C_k, C_j) + \beta D(C_k, C_i) + \gamma |D(C_i, C_j) - D(C_k, C_j)|, \tag{5}$$

where  $D(*,*)$  is a distance function and  $\alpha_i, \alpha_j, \beta$ , and  $\gamma$  are coefficients that assume values according to the hierarchical method used. The formula describes the distance between a cluster  $C_j$  and a new cluster formed by the fusion of the two others,  $C_i$  and  $C_k$ . Table 1 shows the values of the coefficients of equation (5) and the corresponding hierarchical methods.

The simplest and most popular methods of agglomerative hierarchical clustering are the single linkage method (or nearest neighbour method) and the complete linkage method (Xu and Wunsch II 2005). In the simple linkage method, the distance between two clusters,  $C_i$  and  $C_k$ , is determined by the closest two objects, one that belongs to  $C_i$  and the other to  $C_k$ :

$$D(C_i \cup C_k, C_j) = \min \{ D(C_i, C_j), D(C_k, C_j) \}. \tag{6}$$

In the opposite way, to measure the distance between two clusters  $C_i$  and  $C_k$ , the complete linkage method uses the distance of the pair of objects farthest away from

Table 1. Specifications of hierarchical clustering methods in accordance with the coefficients of equation (5).

Coefficients	Methods						
	Single linkage	Complete linkage	Average linkage	McQuitty	Median linkage	Centroid linkage	Ward
$\alpha_i$	0.5	0.5	$\frac{ i }{ i + j }$	0.5	0.5	$\frac{ i }{ i + j }$	$\frac{ i + k }{ i + j + k }$
$\beta$	0	0	0	0	-0.25	$-\frac{ i  j }{( i + j )^2}$	$-\frac{ k }{ i + j + k }$
$\gamma$	-0.5	-0.5	0	0	0	0	0

each other, one that belongs to  $C_i$  and the other to  $C_k$ :

$$D(C_i \cup C_k, C_j) = \max\{D(C_i, C_j), D(C_k, C_j)\}. \quad (7)$$

The results of an agglomerative hierarchical method are usually described by a similarity diagram, called a dendrogram. The dendrogram root node represents a complete data set and each leaf node is a single object. The intermediary nodes describe the extension of proximity between the object, and the height of the dendrogram usually expresses the distance between each pair of objects or clusters. The dendrogram thus organizes the degree of similarity between the clusters into a hierarchy, making it possible to obtain a two-dimensional representation of the similarity or dissimilarity of all of the data.

The dendrogram can be 'cut' at different levels to obtain any number of clusters. This property facilitates the comparison of different cluster configurations for the same data set. However, in order to determine the ideal number of clusters for a dendrogram, some criterion of decision must be employed.

As mentioned in §1.1, the main disadvantage of agglomerative hierarchical clustering methods is their computational complexity, which in general is at least  $O(N^2)$ . This high cost limits the direct application of these methods in large scale data sets, such as in the case of remotely sensed images.

#### 4. Proposed classification methodology

The methodology proposed in this work essentially attempts to exploit the properties of the SOM together with agglomerative hierarchical clustering methods in order to perform the unsupervised classification of remotely sensed images. Figure 1 shows a simplified version of the proposed method, the main steps of which are presented here:

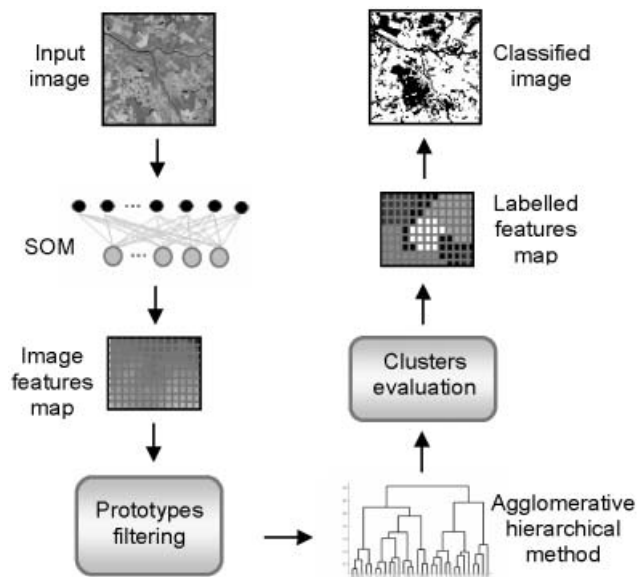


Figure 1. Illustration of the proposed classification methodology.

1. *Sampling*: samples in the form of pixel windows are collected uniformly across the entire region of the image and provide input patterns for SOM.
2. *SOM training*: unsupervised SOM training is performed and an image features map composed of the SOM neurons (prototypes) is obtained.
3. *Filtering of prototypes*: the SOM prototypes that have null activity are discarded and those that have a high degree of spectral heterogeneity (heterogeneous prototypes) are excluded in the next step.
4. *Agglomerative hierarchical analysis*: an agglomerative hierarchical clustering method with connectivity restrictions is applied to the already filtered SOM prototypes, generating a dendrogram of the map.
5. *Clusters evaluation*: a cluster validation index calculated in a modified form is applied at all of the levels of the dendrogram obtained in step 4. The level of the dendrogram that has the optimal value for the validation index is chosen as the best configuration of clusters for the SOM prototypes (ideal number of classes).
6. *SOM classification*: the SOM prototypes are labelled according to the prototype classes determined in step 5.
7. *Image classification*: the SOM prototypes labelled in step 6 as well as the heterogeneous prototypes (filtered in step 3) are utilized as a reference to classify all of the pixels of the image.
8. *Reclassification*: each one of the pixels of the image that have been associated to any heterogeneous prototype is reclassified using the neighbouring pixel class that has the least (spectral) distance from it.

In the following subsections, each of the steps of the proposed methodology is explained in greater detail.

#### 4.1 Sampling

The first step of the proposed methodology consists of collecting an image sample set in order to train the SOM. Unlike pixel by pixel approaches that only use the spectral information of individual points to find homogenous regions, the present work performs the sampling of the image through pixel windows. The idea is to incorporate in the classification process information about the neighbourhood (context) of the pixels, considering that isolated pixels are not able to represent the majority of cover land patterns, especially in the case of images that have higher spatial resolutions. Even though the strategy of using information from neighbouring pixels has a greater computational cost than in the pixel by pixel approach, in the literature, various work has shown that this effort is justified by a proportional increase in classification accuracy (Magnussen *et al.* 2004).

Two other important points related to the sampling process are the size of the sample windows and the way they are sampled. As shown in figure 2, the sample windows are collected uniformly across the entire region of the image, without overlapping and at regular intervals. All of the samples are square and have the same size.

Determination of the size of samples is not a process that can be easily automated. Depending on the characteristics of the image to be classified, windows that are too small might not capture the particular pattern of the land cover classes, while larger windows might include pixels from more than one class. Nevertheless, beginning with a size no smaller than  $3 \times 3$ , the user has, within the proposed methodology, a

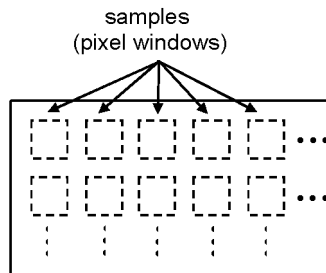


Figure 2. Sampling process of the input image.

certain degree of flexibility to define the dimensions of the sample windows, and this is not therefore a critical point in the process. The proposed method (through steps 3 and 8) processes in a special way the SOM prototypes that represent sample windows which incorporate more than one class (heterogeneous windows). In addition, prototypes that have more than one class can act as borders in the SOM features map contributing to the separation of clusters.

#### 4.2 SOM training

In order to train the SOM, some parameters must be specified to define the structure of the map and to specifically control the stated training. With the objective of guaranteeing good mapping of the original patterns, the proposed methodology defines, in a particular way, the neural net parameters based on the existing literature, on experimental tests, and some peculiarities of the application of SOM on remotely sensed images. However, other alternatives can also be sought out to obtain good maps (Kaski and Lagus 1996).

The proposed methodology utilizes the following parameters to train the SOM:

- Weights initialization: linear,
- Training mode: batch,
- Number of training epochs: 500,
- Type of neighbourhood function: Gaussian,
- Array format: rectangular.

Although the random initialization is the most common form to initialize SOM weights, Kohonen defends the idea of linear initialization, as this permits the learning algorithm to directly enter the convergence phase, as well as avoiding undesirable skews in the array (Kohonen 1997). Linear initialization consists of distributing the neurons in an ordered way along a plane aligned with the axes of the two largest variances in the data set.

The batch training mode makes the result of the SOM mapping insensitive to the presentation order of the input patterns. This can be done by adding the average of the contributions of all of the input patterns for each neuron. This average displacement has led to better results than the conventional SOM algorithm, in which the neurons are updated every time a new input pattern is provided to the net.

To determine the number of training epochs, i.e. the number of times the entire set of input patterns must be provided to the net, the disorder measure 'AUD' (*Average Unit Disorder*) proposed by Azcarraga (2000) was used in this approach. Essentially, AUD evaluates the degree of organization of the SOM during the training process through the differences of synaptic weights and geometric distance between neurons

in the rectangular grid. The SOM is considered sufficiently trained when the AUD measure does not change in value by more than a certain threshold. In this application, experimental tests have shown that the AUD measure becomes stable before the 500 training epochs, obviously considering the image sampling scheme adopted in this methodology. The reader may refer to Azcarraga (2000) for a detailed description of the AUD metric used here.

Instead of utilizing the ‘bubble’ type neighbourhood proposed in the original SOM algorithm, the Gaussian type neighbourhood is employed here because it better emphasizes the clusters in terms of distances between neurons (Kohonen *et al.* 1997). Returning to equation (3), the neighbourhood function  $h$ , which is Gaussian, is defined as follows:

$$h(\|r_c - r_i\|, t) = \exp\left[-\|r_c - r_i\|^2 / 2\sigma^2(t)\right], \quad (8)$$

where  $\sigma$  is defined as the neighbouring radius of the neurons in the SOM topological array.

In the proposed method, the rectangular shape was chosen to organize the array of the neurons of the net, not just because it is the most commonly used, but also because it is the best at generating an image of the SOM features map, which will be described in the next subsection.

The size of the map is one of the free parameters of the SOM that particularly depends on the input image and the objectives of the classification. If the objective is to detect all of the patterns in the image, including those with a low probability of occurrence, large-sized maps must be employed in the analysis; in the opposite case, if the interest is concentrated only on the predominant patterns in the scene, a smaller-sized SOM can be utilized. Nevertheless, the performance of the proposed classification methodology is not significantly affected if sufficiently large sizes for the SOM are utilized. While maps with larger dimensions than are necessary have a larger quantity of inactive neurons, i.e. neurons that do not have an associated input pattern, this event is not prejudicial within the proposed methodology. Just as the method deals with the heterogeneous prototypes in an adequate manner, the same occurs with inactive prototypes. In addition, such prototypes can also contribute to the class separation process, since generally they are part of regions in the map that correspond to transitions between clusters.

**4.2.1 Image features map.** The proposed methodology has a peculiarity in relation to image mapping through SOM that must be pointed out. Since the SOM neurons are spatially arranged in the shape of a rectangular grid, and still considering that in this neural net the weight vectors of each neuron have the same dimensions as the input patterns (pixel windows in this case), it makes it possible to generate an image of the neural grid of the map. As shown in figure 3, the SOM neurons or prototypes

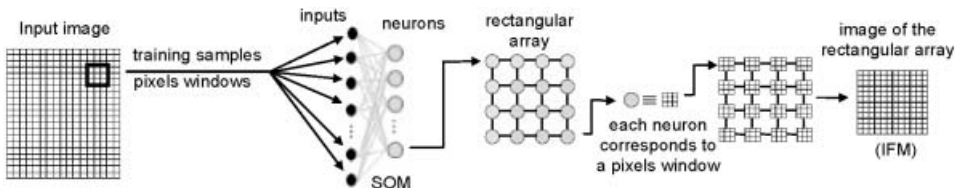


Figure 3. Illustration of the construction process of the IFM.



(through their synaptic weight vectors) correspond to pixel windows, which permit them to be visualized in the form of an image. This visual representation of the SOM neural grid, after its unsupervised training, is called here the *Image Features Map* (IFM).

As will be presented in the experimental results (§5), it is possible to visualize the mapping of the input image patterns computed by SOM and their respective properties by means of the IFM.

### 4.3 Prototypes filtering

The third step of the proposed methodology consists of filtering two types of prototypes that generally appear in the mapping of image patterns through SOM. These prototypes are called *inactive* and *heterogeneous*.

The inactive prototypes correspond to the neurons that have null activity in the SOM competitive learning process, i.e. they are not associated with any input patterns. These prototypes are simply discarded, not going on to the next steps of the analysis.

Heterogeneous prototypes are those that have a high degree of spectral heterogeneity and are normally associated with input patterns that incorporate more than one land cover class. Most of the time, these patterns correspond to transition regions between land cover classes present in the image and are captured as a consequence of the sampling through pixel windows. Prototypes that are considered heterogeneous are excluded in the fourth step of the proposed methodology, in which the hierarchical clustering method is applied. The objective of excluding these prototypes is to prevent them (and consequently the input patterns associated with them) from being erroneously attributed to one of the classes that are part of them. Heterogeneous prototypes can be seen as noisy or divergent patterns, and if they are not filtered the hierarchical method can incorporate them in the clusters that will be produced or retain them in separate clusters. As in hierarchical methods, data or cluster fusion at a determined stage cannot be corrected in subsequent stages, incorrect interpretations regarding the classes and/or the number of classes of the image can be made. Due to this, the heterogeneous prototypes are reincorporated only in the procedures adopted in steps 7 and 8 of the proposed method.

To measure the heterogeneity degree of each prototype, an index called the *Spectral Heterogeneity Index* (SHI) is applied. It is defined as:

$$\text{SHI} = \frac{1}{M} \sum_{i=1}^M \text{CV}_i, \quad (9)$$

where  $M$  is the number of spectral bands in the image considered in the analysis and  $\text{CV}_i$  is the coefficient of variation (relative dispersion) of the grey levels of all of the pixels in the spectral band  $i$  of the prototype. The prototypes whose SHI values satisfy the relationship given below are considered heterogeneous and are consequently filtered:

$$\text{SHI} > \mu_{\text{SHI}} + \frac{1}{2} \sigma_{\text{SHI}}. \quad (10)$$

Here  $\mu_{\text{SHI}}$  and  $\sigma_{\text{SHI}}$  are the average and the standard deviation of the SHI values of all of the prototypes respectively.

The processing time consumed by this filtering step is compensated not just by the reduction of the number of prototypes that will be analysed in step 4 of the method, but mainly by a possible increase in the precision of the image classification. In addition, as has been mentioned previously, inactive and heterogeneous prototypes can act as ‘interpolation units’ in the SOM grid, facilitating the separation of the clusters.

#### 4.4 Agglomerative hierarchical analysis

In the fourth step of the proposed methodology, an agglomerative hierarchical clustering method is applied to the trained and filtered SOM prototypes. The single linkage method (or nearest neighbour method) was used in this approach, although other hierarchical methods (as presented in table 1) can be employed.

An important characteristic of the hierarchical analysis strategy presented in this work is the imposition of restrictions to the possible SOM prototype fusions. If an agglomerative hierarchical clustering method is applied to the SOM neurons it must respect the topological relationships of the neural net output space. Unlike traditional hierarchical clustering methods, which consist of comparing all of the pairs of objects to decide on a fusion, the approach utilized in this work verifies the possibility of fusions only between adjacent (or neighbouring) neuron pairs in the SOM grid. As shown in figure 4, given a  $4 \times 4$  neurons grid, the neuron in position (3,2), for example, only has the possibility of fusion with neighbouring neurons in positions (2,2), (3,1), (3,3), and (4,2). This approach significantly diminishes the processing cost of the hierarchical clustering method applied.

At the end of this step, there is a dendrogram that shows, in a hierarchical way, the similarity levels between the SOM prototypes. Remembering that the SOM prototypes represent all of the input patterns, the obtained dendrogram reflects the relationships that exist between the original image data, and it can therefore be considered as a simplified dendrogram of the image.

#### 4.5 Clusters evaluation

At each level of the dendrogram obtained in the previous step there is a different cluster configuration for the SOM prototypes, and consequently for the image data set. Thus, as with any hierarchical clustering method, some decision criterion needs to be applied in order to verify which level of the dendrogram represents the ideal cluster configuration (or the ideal number of classes).

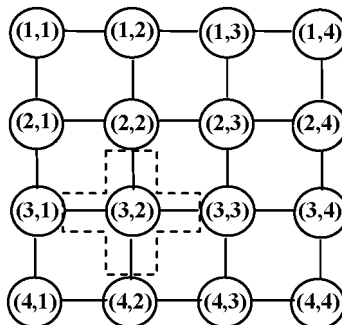


Figure 4. A  $4 \times 4$  neurons grid.

According to Vesanto and Alhoniemi (2000), the most typical solution is to cut the dendrogram where there is a large distance between two merged clusters. However, this approach ignores the fact that the within-cluster distance may be different for different clusters. In fact, each cluster can be made up of various smaller clusters and thus an intra-cluster dispersion measure can also be important in evaluating the compaction degree of the clusters that are formed. A common solution is to apply cluster validation indices that incorporate intra-cluster and inter-cluster dispersion measures. The strategy consists of applying the validation index at all levels of the dendrogram and choosing the level at which the index presents its optimum value among the values obtained.

On the other hand, it must be observed that the implementation of the majority of those cluster validation indices require a high computational cost. Due to this, when the number of clusters and the volume of the data set are very large, the utilization of these indices becomes prohibitive. Remotely sensed images are a typical example. As described in Ji (2003), validation indices are frequently ignored in the analysis of remotely sensed images and are not available in the majority of image processing packages.

However, since in the proposed methodology, the dendrogram is not constructed from the original patterns of the image, but is made up of a small set of SOM neurons that represent them, the application of a cluster validation index at each level of the dendrogram becomes viable, since the quantity of data to be evaluated is far less than the quantity of patterns directly extracted from the image.

Although the SOM performs a good approximation of the input space, it is true that the strategy of applying a validation index on the network prototypes, instead of applying them directly to the original pattern can cause some numerical differences in the results. To diminish possible approximation errors, the proposed method modifies the way that validation indices are calculated. The indices are calculated by utilizing not just the weight vectors of the SOM prototypes, but also the activity level of each one of them. The activity level of one prototype is equal to the number of input patterns that are associated to it by the SOM mapping process.

An example of the modifications applied to the calculations of indices can be described by equation (11). This formula, and variants of it, are commonly used in some cluster validation indices to calculate the distance ( $\delta_{ij}$ ) between two data clusters  $C_i$  and  $C_j$  (Bezdek and Pal 1998):

$$\delta_{ij} = \frac{1}{|C_i||C_j|} \sum_{x \in C_i, y \in C_j} d(x, y). \quad (11)$$

In the formula above,  $d(*,*)$  is a measure of distance, and  $|C_i|$  and  $|C_j|$  represent the number of points in data clusters  $C_i$  and  $C_j$  respectively. Therefore,  $\delta_{ij}$  is basically determined through calculations of the distance between all of the pairs of points  $x$  and  $y$  belonging to data clusters  $C_i$  and  $C_j$ . If  $|C_i|$  and  $|C_j|$  are very large, the calculation of  $\delta_{ij}$  becomes excessively costly.

Applying the strategy of modifying the index calculation, the computation of the distance between clusters  $C_i$  and  $C_j$ , equivalent to equation (11), becomes formulated in the following way:

$$\delta_{ij}^{\text{SOM}} = \frac{1}{|C_i||C_j|} \sum_{w_i \in W_i, w_j \in W_j} h(w_i)h(w_j)d(w_i, w_j), \quad (12)$$

in which  $W_i$  and  $W_j$  are SOM prototypes sets that represent the data from clusters  $C_i$  and  $C_j$  respectively,  $d(.,.)$  is the same distance measure used in equation (11),  $h(w_i)$  is the activity level of prototype  $w_i$  belonging to  $W_i$  and  $h(w_j)$  is the activity level of prototype  $w_j$  belonging to  $W_j$ .

Whereas  $\delta_{ij}$  is based on the distance between all of the pairs of points belonging to  $C_i$  and  $C_j$ ,  $\delta_{ij}^{\text{SOM}}$  is based only on the distance between the pairs of points belonging to  $W_i$  and  $W_j$ . Therefore, it can be observed that  $\delta_{ij}^{\text{SOM}}$  is an approximate calculation of  $\delta_{ij}$ , but with a much lower processing cost, since the quantities  $|W_i|$  and  $|W_j|$  are generally much smaller than  $|C_i|$  and  $|C_j|$ . The inclusion of the activity levels  $h(.)$  of the prototypes in the calculation of  $\delta_{ij}^{\text{SOM}}$  helps to diminish the effect of quantization errors caused by SOM mapping.

Another important point of this step of the proposed method is the choice of the validation index that will be applied in a modified form at all levels of the dendrogram. In the literature, there are various proposals of cluster validation indices (Milligan and Cooper 1985). The method proposed in this paper applies the CDbw validation index (*Composing Density Between and Within Clusters*) proposed in Halkidi and Vazirgiannis (2002). This index was chosen because it is based on two important concepts, the intra-cluster and inter-cluster density, and also because it is adequate to evaluate data clusters that have arbitrary and complex formats, which does not occur in the majority of other indices. In Sezgin *et al.* (2004), for example, the Davies–Bouldin validation index was used to estimate the best number of clusters in remotely sensed images from the SOM partitioning by the  $K$ -means algorithm. This same validation index was also employed in a way combined with another index in the unsupervised classification method proposed in Marçal and Borges (2005). As supported by Pal and Biswas (1997), the Davies–Bouldin validation index is suitable only for spherical clusters and is sensitive to noisy points.

#### 4.6 SOM classification

Having defined the dendrogram level that presents the best cluster configuration for the SOM prototypes, and thus for the image patterns, step 6 of the proposed method consists only of labelling these prototypes. The prototypes that belong to the same cluster correspond to a determined land cover class present in the original image and consequently receive a particular label.

According to the methodology proposed here, each discovered land cover class will therefore be represented by a group of SOM prototypes and not just one single prototype as occurs, for example, in the  $K$ -means algorithm.

#### 4.7 Image classification

In order to classify the image, the heterogeneous prototypes that were filtered out in step 4 are reincorporated and also considered as a class. For this purpose, all of them receive the same label, which obviously must be different from those used to designate the other classes of prototypes in the previous step.

Pixel windows are collected from the original image with equal dimensions from the training sample and are compared to all of the SOM prototypes. This comparison is performed through the distances calculated between the pixel windows and each of the prototypes. The central pixel of the pixel window receives

the label of the prototype that has the least distance from it. The image is then entirely run through until all of the pixels have been classified.

#### 4.8 *Reclassification*

In the last step of the proposed method, the results of the classification process performed in the previous step are improved. The pixels in the image that receive the label of the heterogeneous prototypes class are reclassified. Each one of these pixels is compared to its neighbouring pixels in the image that have not been labelled by the heterogeneous prototype class, and it receives a new label that will be the same as the neighbour that has the smallest distance from it in terms of spectral features. This procedure prevents the pixels that are in class transition regions from being classified erroneously.

### 5. Experimental results

This section shows examples of the application of the proposed methodology on two test images. The results are compared with those obtained by applying the *K*-means algorithm to the same images. It is worth emphasizing that the comparisons performed here refer to the results obtained by two algorithms with significantly different principles and characteristics and this work is not intended to decide which method is better, but to analyse the applicability and the benefits of the classification methodology proposed.

All of the experiments presented in this section were performed with MATLAB software and a microcomputer with an AMD Athlon<sup>™</sup> XP 2600+1.91 GHZ processor and 1 GB of RAM.

#### 5.1 *Experiment 1*

Figure 5 shows a colour composite of the image used in this experiment. The image ( $496 \times 512$  pixels) is composed of spectral bands 3, 4, and 5 of the Landsat-5 TM satellite and was acquired on 20 August 1988. It has an IFOV of 30 m. The study area covers the city of Manaus and the intersection of two rivers, Rio Negro and Rio Solimões in the state of Amazonas, Brazil. Four large land cover classes are present in the scene: urban area (pink), vegetation (green), and two water patterns (the darker blue pattern corresponds to Rio Negro and the more purple pattern corresponds to Rio Solimões). Small areas with sparse vegetation or deforestation also appear in the image, but with a very low density. This image was provided by the National Space Research Institute, (INPE), Brazil.

Application of the proposed methodology was initially performed with a sampling process of the scene. Sample windows of size  $5 \times 5$  were collected uniformly across the entire region of the image, without overlapping and at regular 10 pixel intervals, giving a total of 2500 samples obtained without user intervention.

A SOM composed of 100 neurons arranged in a  $10 \times 10$  rectangular grid was trained with all of the samples previously collected. The other parameters of the SOM were defined according to the specifications presented in §4.2.

Figure 6 shows the IFM obtained after the SOM training. As described in §4.2.1, the IFM is an image generated from the trained SOM neurons.

Each small square on the IFM corresponds to the image of a SOM prototype (or neuron). Therefore, it is possible to visualize the properties of the mapping produced by the neural net. The four land cover classes present in the original image appear in



Figure 5. Colour composite of test image 1.

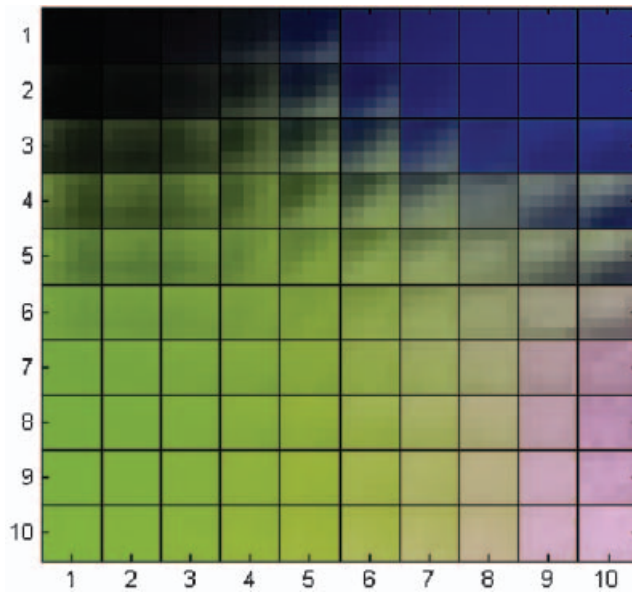


Figure 6. IFM obtained from test image 1 after SOM training.

the form of clusters on the IFM. In the upper left corner the prototypes associated with the darker water pattern are found, in the upper right corner the prototypes associated with the other water pattern (more purple) are present, in the lower right (in pink) those that correspond to the urban area are found, and in the lower left corner to the centre of the IFM there are prototypes associated with the vegetation class. If the original image is observed, it can be verified that the vegetation class occupies the most area in the scene. Due to this, since the SOM reflects the probability distribution of the input data, the prototypes associated with the vegetation class (in green on the IFM) are present in a greater number than the other prototype classes. The SOM's topological ordering property can also be seen. The prototypes of the two classes of water in the original image are close to one another in the higher part of the IFM, because the input patterns (samples) corresponding to these two classes are the most spectrally similar in comparison to the spectral attributes of the other two land cover classes. While the IFM is useful for visualizing the SOM mapping, it must be pointed out that in the method proposed here no interaction by the user is carried out.

After the SOM training, the next step to be executed is filtering inactive and heterogeneous prototypes. Figure 7(a) shows the activity level of each of the SOM prototypes, i.e. the number of input patterns (samples) that are associated with each of them through the SOM mapping. It is observed that the prototypes located in grid positions (1,5), (3,2), (4,8), and (6,10) have null activity levels and thus do not have any input pattern associated them. According to the proposed method, these prototypes are simply discarded, not going through to the next step of the analysis.

Figure 7(b) presents the values (in percentages) of the SHI values of each prototype. As shown in §4.3, the SHI measures the spectral heterogeneity of the SOM prototypes. In this way, since the values obtained in this experiment for the average ( $\mu_{SHI}$ ) and the standard deviation ( $\sigma_{SHI}$ ) of the SHIs were 11.89 and 8.93 respectively, the prototypes that have SHI values above 16.35 (the limit obtained according to the criteria given in equation (10)) are considered heterogeneous and

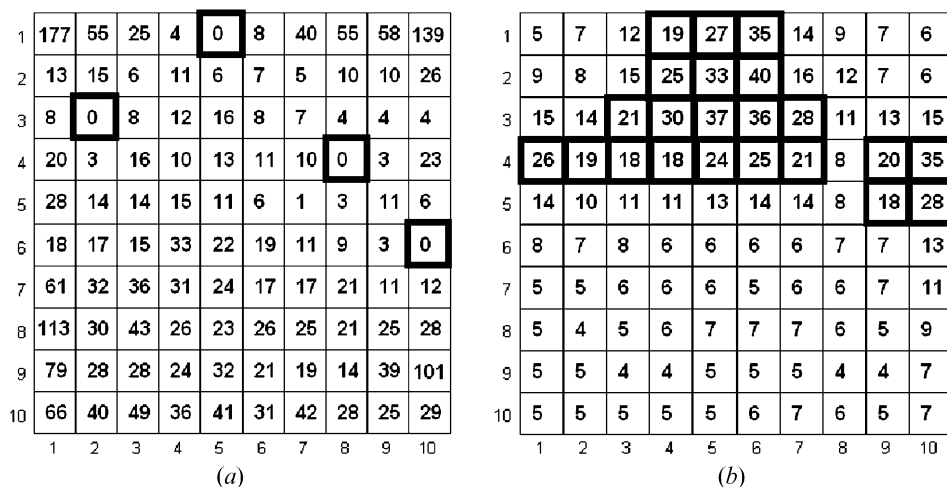


Figure 7. (a) Activity levels of the SOM prototypes (prototypes with null activity levels are marked in bold). (b) Values (in percentages) of the SHIs of the SOM prototypes (heterogeneous prototypes are marked in bold).

are filtered. These prototypes (marked in bold in figure 7(b)) will be reconsidered only in the last two steps of the proposed classification methodology.

Therefore, of the 100 total SOM prototypes, 25 of them (between heterogeneous and inactive) were filtered, with the remaining 75 prototypes to be analysed in the following two steps.

According to the proposed method, the following step consists of applying the agglomerative hierarchical method to the filtered SOM prototypes. As described in §4.4, the simple linkage method is used in this approach and the fusions of prototypes (or prototypes groups) are carried out respecting their topological neighbourhood in the rectangular SOM grid. Figure 8 shows the dendrogram obtained at the end of this step. The dendrogram has 75 levels, each one with a different configuration of SOM prototypes clusters that results in a different classification of the original image.

In addition to making it possible to classify the image at different cluster levels, the dendrogram obtained allows the user to analyse each one of the merges performed between the prototypes or prototypes groups, and their hierarchy. If it is restated that, in the proposed method, the user can visualize the image of each SOM prototype through the IFM, the analysis of the dendrogram becomes clearer and more efficient, making it easier to understand the relationships that exist between the different patterns present in the image.

It is important to note that this type of analysis is only possible due to the SOM mapping that was performed. The neural net 'abridges' the patterns of the original image by means of a small set of prototypes, making it viable to construct a simplified dendrogram. If the SOM was not utilized to represent the input data, the construction or interpretation of the dendrogram would become impractical, since there would be 2500 different cluster levels if all of the samples collected from the

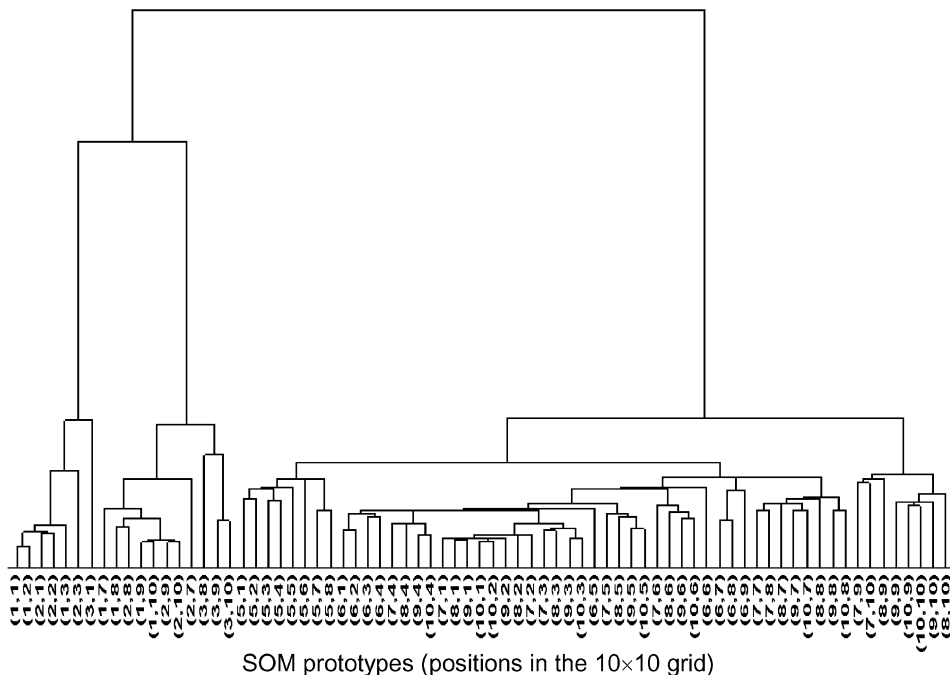


Figure 8. Dendrogram of the SOM prototypes.



image were considered, or a much higher number of levels if all of the points on the test image were considered.

Continuing with the application of the steps of the proposed method, the CDbw cluster validation index, calculated in a modified way (as presented in §4.5), is applied at all of the levels of the dendrogram to determine which cluster configuration of the 75 that were obtained is ideal. For the purpose of comparing the two forms of calculation of the indices, figure 9 shows the validation index values in the modified and the original form. The graph only shows the values between levels 61 and 74 of the dendrogram. For levels lower than 61 the index values are smaller than those on the graph, or even null, due to the existence of levels composed of clusters with a single prototype. The validation index value for level 75 is not shown either, as it has a single cluster that includes all of the data. The number of clusters or classes from level 61 to 74 decreases from 15 to 2 respectively.

As expected, the strategy of modifying the cluster validation index computation has approximation errors, but the variation of their values occurs in a similar way to the original version, not compromising the decision making process regarding the ideal number of data clusters. The main advantage of using the modified version of the index is its processing time. In this experiment, while the time to calculate the CDbw index in the original form (for the entire dendrogram) was 784 s, the modified version needed only 32 s. It is important to note that this significant difference occurs due to the volume of data considered for each of the ways to calculate the index. Calculation of the CDbw index in its original form was executed at each level of the dendrogram considering the 2500 samples collected in the image, while the modified version calculation considered only 100 SOM prototypes and their respective activity levels.

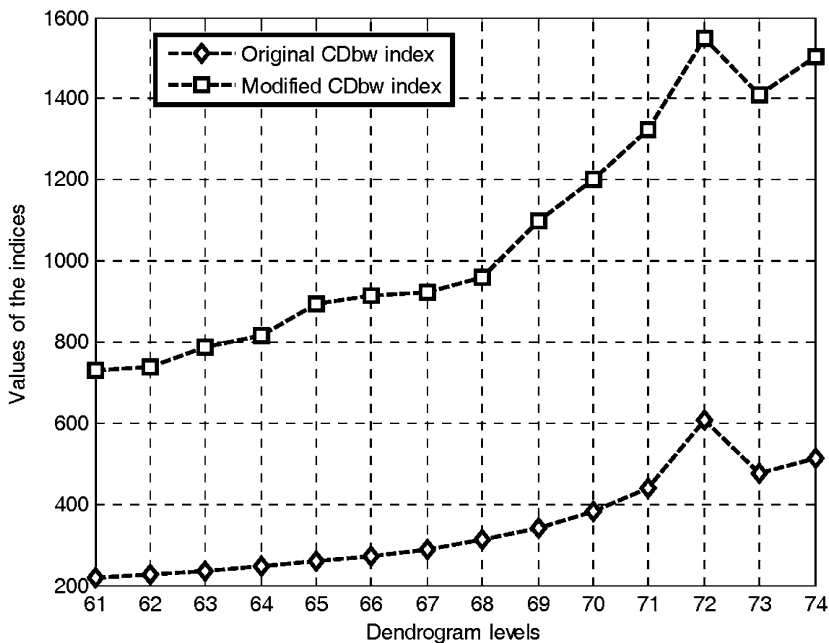


Figure 9. CDbw index values in the modified and the original form.

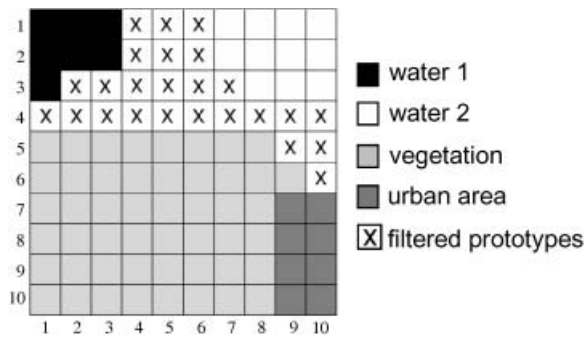


Figure 10. Classified IFM.

For the CD<sub>bw</sub> index, the greater its value, the better the result (Halkidi and Vazirgiannis 2002). Considering this, of all of the dendrogram levels obtained in this experiment, level 72 is the one that has the greatest value for the index (as can be seen in figure 9), and thus, the best cluster configuration for the SOM prototypes. Figure 10 shows the IFM classified in accordance with the clusters (classes) of level 72. Compared to the IFM image in figure 6, it can be seen that the SOM prototypes are grouped into four classes, adequately corresponding to the four large land cover classes present in the image (two water patterns, vegetation, and urban area). The squares marked with 'x' are the prototypes that were discarded in the filtering step of the proposed method. It is interesting to see that these prototypes act as 'interpolation units' and certainly contribute to the class separation process.

When the SOM prototype classification step has been completed, the last two steps in the method concentrate on the classification of all of the points in the original image. For this, the heterogeneous prototypes (previously filtered) are again incorporated in the process and also considered as a class. This way, in the penultimate step of the proposed method, five classes of SOM prototypes are utilized as a reference in order to perform the classification of the image, the four classes of the chosen dendrogram level (shown in figure 10) and the heterogeneous prototype class.

The test image 1 is entirely run by considering  $5 \times 5$  pixel windows (size equal to the sample windows) and by comparing to all of the SOM prototypes. This comparison is performed using distances calculated between considered pixel windows and each one of the prototypes. The central pixel of the pixel window receives the prototype label that has the shortest distance from it.

Finally, the last step of the method is applied. A total of 26 585 pixels (10% of the image) were associated with the heterogeneous prototype class. These pixels are reclassified and receive a new label equal to the neighbouring pixel that has the least distance from it (according to spectral attributes). Figure 11 shows the classification results for the original image using the proposed methodology (the four classes are represented by the same colours used in figure 10).

Therefore, according to the proposed methodology, the test image 1 was classified into four classes, which correspond to the four land cover patterns that dominate the scene. Some small areas of sparse vegetation or deforested were classified as urban class because they are spectrally very similar to the patterns of this class and have a lower density in comparison to the other land cover patterns.



Figure 11. Test image 1 classified by the proposed methodology.

Figure 12 shows the result of the image classification by the  $K$ -means algorithm considering the number of classes as being equal to four. The algorithm was executed using five different initializations for the centroids of each of the classes and a maximum number of iterations equal to 100.

Unlike the proposed method, the  $K$ -means algorithm did not differentiate one of the predominant water patterns in the image and classified the vegetation areas in two categories: one more extensive (a lighter grey) and another that appeared only at the transition areas with other classes (in black). The algorithm also confused the sparsely vegetated or deforested areas with urban area, but in a much more accentuated manner than the method proposed here. The two water patterns in the image are only differentiated by the  $K$ -means algorithm if the number of clusters ( $k$ ) is set to greater than or equal to five. However, while the algorithm classifies the vegetation in various different categories, the confusion between the sparsely vegetated areas and urban areas remains, regardless of the predefined number of classes.

Performing a visual analysis of the classification results obtained by the two methods (figures 11 and 12) it is possible to verify that the image classified by the proposed methodology has a superior visual aspect. While the classification of the image by the  $K$ -means algorithm appears more dotted, the image produced by the proposed method is more homogenous in all of the classified areas. The difference in appearance is certainly due to the way the two methods extract the information from the image. While the  $K$ -means algorithm applied here uses a pixel by pixel approach, the proposed method works with pixel windows that incorporate the neighbourhood information of pixels.

Attempting to perform a more systematic and less subjective analysis of the image classified by the proposed method, and considering the absence of terrestrial truth for the test image, the present work performed the classification of the test image in

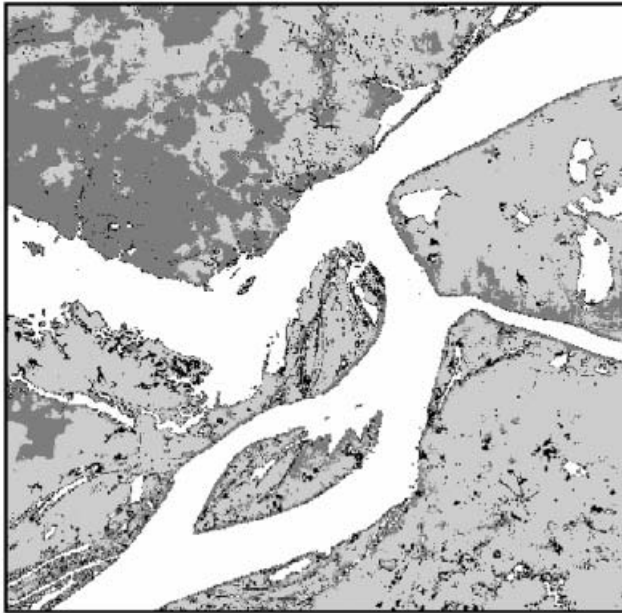


Figure 12. Test image 1 classified by the  $K$ -means algorithm considering  $k=4$  (the 4 classes are represented with the same colours used in figure 10).

a supervised manner and considers these results as a reference (or ‘true’). The supervised classification, considering the four predominant classes in the image, was carried out by means of a multilayer *Perceptrons* neural net with a *backpropagation* training algorithm. This class of neural nets has been widely employed to perform supervised classification of remotely sensed images (Gonçalves *et al.* 1996, Liu *et al.* 2003).

The confusion matrix and the Kappa agreement index were calculated by comparison of the image classified by the proposed method and the reference image (resulting from the supervised classification). Analysing the confusion matrix shown in table 2 and considering that the Kappa index reached the value of 0.07, the classification result of test image 1 by the method presented in this work can be considered to be extremely satisfactory.

## 5.2 Experiment 2

The second experiment was performed on a segment taken from an IKONOS image (provided by Engesat/Brazil, © Space Imaging). It has an IFOV of 4 m and is composed of three bands in the visible spectrum. The segment used in this

Table 2. Confusion matrix for test image 1 classified by the proposed method.

Classes	water 1	water 2	vegetation	urban area	Total
water 1	96.43	0.40	0.34	0.20	16.64
water 2	0.61	99.35	0.04	0.09	15.19
vegetation	2.89	0.02	98.89	6.73	53.29
urban area	0.07	0.23	0.73	92.98	14.88
Total	100.00	100.00	100.00	100.00	100.00



Figure 13. Colour composite of test image 2.

experiment has 480 lines and 360 columns. This imagery shows irrigation pivots in the region of Andaraí in the state of Bahia, Brazil, and presents five land cover classes: sparse vegetation, forest, two types of bare soil, and growing crop. Figure 13 shows the segment of the original image (denoted here as test image 2).

As stated in the procedures described in §4, 1728 sample windows of size  $7 \times 7$  were collected from test image 2 and used to train a SOM composed of 144 neurons arranged in a  $12 \times 12$  rectangular grid. Figure 14(a) shows the IFM obtained after SOM training, in which it is possible to visualize prototype clusters that correspond to the land cover classes present in the original image. As in experiment 1, topological ordering and density matching properties of the mapping produced by the neural net can be seen. Land cover classes with similar spectral attributes are mapped to neighbouring regions of the two-dimensional output grid and those that occupy larger areas in the original image are mapped to a larger number of prototypes of the grid.

After the SOM training, the prototype filtering process was applied. In this experiment 8 SOM prototypes presented null activity, and 12 presented a high degree of spectral heterogeneity, given that its SHI values exceeded the threshold defined in equation (10). Therefore, of the 144 total prototypes, 20 of them were filtered, leaving 124 prototypes to be analysed.

In the following step, the agglomerative hierarchical method was applied on the filtered SOM prototypes. Consequently, a dendrogram consisting of 124 levels was

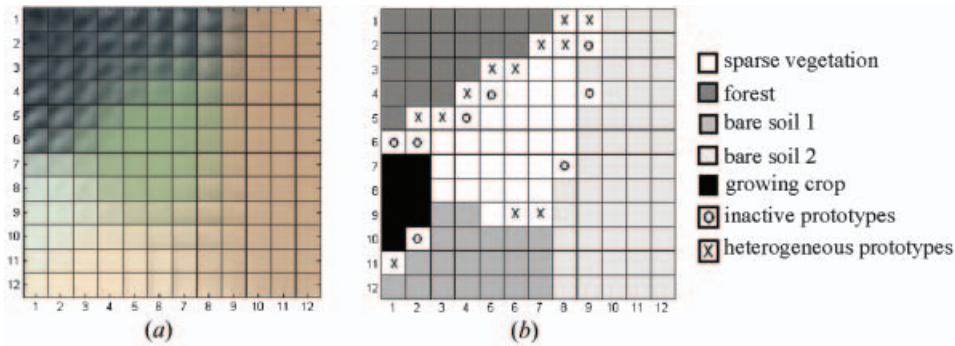


Figure 14. (a) IFM obtained from test image 2 after SOM training, and (b) IFM classified image.

generated, each level with a different configuration of SOM prototype clusters. In sequence, the modified version of the CDbw cluster validation index was applied at all of the levels of the dendrogram to determine which of the 124 cluster configurations that were obtained was ideal. Level 120 had the greatest value for the validation index, and thus, the best cluster configuration for the SOM prototypes. Figure 14(b) shows the IFM classified in accordance with the clusters (classes) of level 120. Compared to the IFM image in figure 14(a), it can be seen that the SOM prototypes are grouped into five classes, adequately corresponding to the five land cover classes present in the original image. The squares marked with 'o' and 'x' are the inactive and heterogeneous prototypes respectively that were filtered in the previous step.

The original image was classified comparing  $7 \times 7$  pixel windows with all of the labelled SOM prototypes (except the inactive prototypes). Conforming to the last step of the proposed method, a total of 6912 pixels associated with the heterogeneous prototype class were reclassified using the neighbouring pixel class that had the least (spectral) distance. Figure 15(a) shows the classification result of test image 2 using the proposed methodology (the five classes are represented by the same colours used in figure 14(b)).

Figure 15(b) shows the result of the classification of the same image by the  $K$ -means algorithm, and setting the number of classes to five. The algorithm was executed using various different initializations for the centroids of each of the classes and a maximum number of iterations equal to 100.

Performing a visual analysis of the results obtained by the two methods it is possible to clearly verify a large difference between the two classifications. The proposed method adequately classifies the different land cover patterns present in the original image without any significant confusion between them. On the other hand, the  $K$ -means algorithm does not correctly discriminate the patterns and presents a relatively erroneous classification. It does not differentiate the growing crop class (in the upper right corner of the image) from one of the bare soil patterns (in the upper left side of the image), and also does not classify the areas of the other land cover classes homogeneously. Just as in the classification results of experiment 1, the classification of the test image 2 by the  $K$ -means algorithm are more speckled, while the image produced by the proposed method is considerably more homogeneous in all of the classified areas. This difference in appearance is more accentuated in this experiment because test image 2 presents a higher spatial

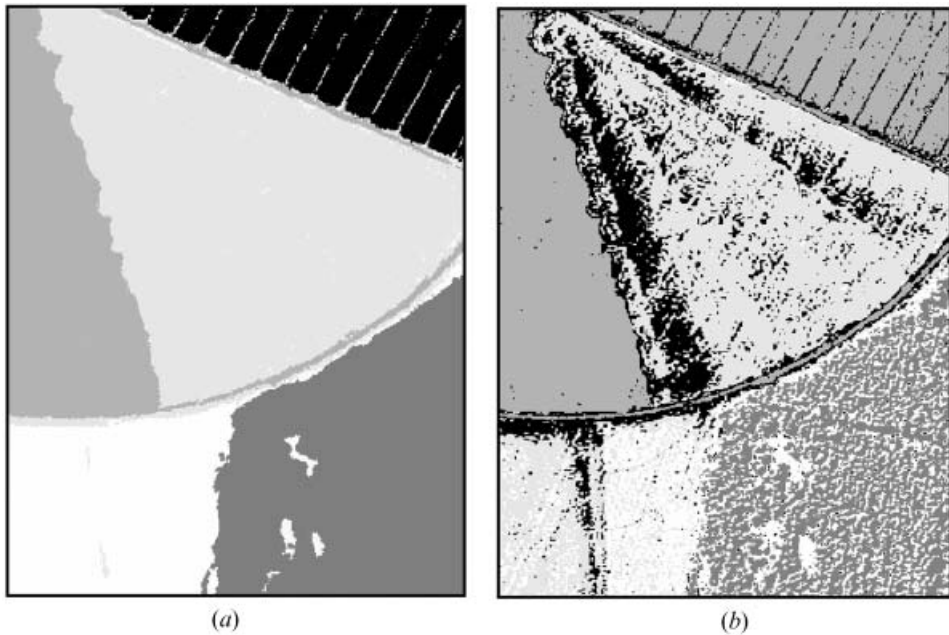


Figure 15. Test image 2 classified by: (a) the proposed method, and (b) the  $K$ -means algorithm considering  $k=5$ .

resolution than test image 1. In test image 2, the land cover classes have more well-defined textural features, and because the proposed method uses the neighbourhood information of the pixels, it produces a superior result to the  $K$ -means algorithm, which uses a pixel by pixel approach.

Although the quality of the classification produced by the proposed method is visually superior, the image used in this experiment was also classified in a supervised manner, and the result of this classification was used as a reference to evaluate the results obtained by the two methods. As in experiment 1, the supervised classification was carried out by means of a multilayer *Perceptrons* neural net. The Kappa index was then calculated individually for the images produced by the proposed method and by the  $K$ -means algorithm. Confirming the visual impression of the results, the image classified by the  $K$ -means algorithm had a Kappa value of 0.66, while the image classified by the proposed method had a value equal to 0.94.

### 5.3 Processing time

Table 3 shows the processing times spent in the two experiments by each of the steps of the proposed classification methodology. The global time consumed to perform the classifications of test images 1 and 2 were 154.9s and 167.7s respectively. Although test image 1 was larger than test image 2, the size of the samples, the size of the SOM and the number of dendrogram levels of experiment 2 were superior to those of experiment 1, and therefore the global time to classify test image 2 was greater. In both experiments, the training of the SOM, the cluster evaluation, and the final classification of the images were the steps that used the most processing time.

A comparison between the processing times consumed by the proposed methodology and by the  $K$ -means algorithm to classify the test images is not

Table 3. Processing time consumed in each step of the proposed method.

Steps	Consumed time (s)	
	Experiment 1	Experiment 2
Sampling	10	6
SOM training	40	53
Prototypes filtering	2	2
Agglomerative hierarchical analysis	0.5	1
Clusters evaluation	32	48
SOM classification	0.4	0.7
Image classification	66	55
Reclassification	4	2
Total	154.9	167.7

appropriate here. As previously mentioned, the two methods have principles and characteristics that are very different, and which should be considered in the evaluation of its calculation complexities.

An important difference between the methods is in the quantity of classes they analyse in order to classify the image. The proposed method evaluates different cluster configurations for the data, whereas the  $K$ -means algorithm performs the classification of the scene only for a single quantity of classes defined *a priori*. In experiment 1, for example, the time spent by the proposed method to evaluate 75 different cluster configurations was 32 s, using SOM prototypes to represent the image data and the CDbw validation index with modified calculation. If a similar strategy to the  $K$ -means algorithm was to be used, i.e. different classifications of the image performed by varying the number of classes ( $k$ ) and then applying the CDbw validation index to determine which was better, the processing time would be so high that the strategy would be considered impracticable. It is enough to take into consideration the time spent in applying the CDbw index to evaluate the classification produced by the  $K$ -means algorithm shown in figure 12. It took 1890 s, far more than the time spent by the proposed method to evaluate diverse cluster configurations for the image. Even though the CDbw validation index has a good level of computational complexity in comparison to others, the high number of points in remotely sensed images generates a very high processing volume to execute its calculations.

Taking these observations into account and also considering the possible benefits of the techniques and procedures employed in the methodology proposed (such as prototype filtering, hierarchical analysis, and the reclassification), we can conclude that the processing time consumed by the method presented here is perfectly admissible, certifying its application viability.

## 6. Conclusions and future perspectives

In this work, an unsupervised method of classifying remotely sensed images that attempts to exploit the properties of the SOM together with hierarchical clustering methods was presented. The key point of the proposed method is to perform the clusters analysis of the image through a set of SOM prototypes, instead of working directly with the original patterns of the scene. This approach significantly reduces the complexity of the analysis, making it possible to use methods that have not normally been considered viable for the processing of



remotely sensed images, for example, agglomerative hierarchical methods and cluster validation indices.

The proposed methodology presents advantages and potentialities that make it a promising alternative to performing the unsupervised classification of remotely sensed images. These include:

- (1) The method does not require a previous definition of the number of classes to perform the classification of the image. It does not occur in the majority of the conventional unsupervised classification methods.
- (2) The method has only two parameters that must be defined by the user (the size of the samples and the number of SOM neurons), and despite this it is very robust as far as these choices are concerned.
- (3) The simple use of pixel windows allows contextual information to be included without any explicit calculation of measure for it. This approach contributes to the quality of the resulting classification.
- (4) The utilization of an agglomerative hierarchical clustering method allows the user to observe the relationships that exist between the land cover patterns present in the image at different cluster levels. This approach can be extremely useful in situations where the information structure present in the image is not entirely known. In addition, if the user should desire, that image can be classified considering the different quantities of classes.
- (5) The distributed representation of the classes by means of prototype groups gives the method the potential to discover geometrically complex and varied data clusters. Methods such as *K*-means use a single prototype (centroid) to represent each class and because of this are only capable of adequately detecting clusters that have hyperspherical formats.
- (6) The method can determine, without any user intervention, the ideal number of clusters or classes for the image.
- (7) The method uses a particular approach to classify pixels situated in transition regions between classes. This procedure contributes to increasing the accuracy of the resulting classification.

In the proposed methodology, the combined use of the SOM and hierarchical clustering techniques not only makes use of the advantages of both the methods to perform the classification process of the images, but also allows each method to help overcome the drawbacks of the other. As mentioned previously, the representation of the original image by a smaller set of SOM prototypes makes possible the use of hierarchical clustering methods which are considered impractical for large data sets. Moreover, the topological ordering property of the SOM restricts the possible fusions of SOM prototypes, and therefore it helps to further reduce the computational complexity of the hierarchical clustering method applied. On the other hand, the hierarchical clustering method automatically segments the trained SOM without the necessity of user intervention, and still hierarchically organizes the groups of SOM prototypes that represent the patterns of the original image. In addition, it is important to emphasize that the proposed method is different to the majority of the unsupervised approaches for classifying remotely sensing images as it does not make use of partitional clustering methods such as *K*-means and ISODATA.

The strategy used to modify the calculation of cluster validation indices utilizing SOM prototypes also contributes to the good performance of the proposed method.

Despite approximation errors, the modified calculation of the indices does not compromise the decision making regarding the ideal number of classes of the images and still significantly collaborates in the reduction of global processing time. Furthermore, the CDbw cluster validation index employed in the proposed approach has great potential to evaluate data clusters that have arbitrary and complex formats, which does not occur in the majority of other indices, most of which are efficient only at evaluating data clusters that have hyperspherical shapes.

Even though the classification methodology proposed here has good perspectives for application, it still opens a lot of areas for research and advances that can be explored in future work. SOM variants were not used in this work because we wanted to select the most frequently used versions of the SOMs. However, these SOM variants and new SOM developments could naturally be applied to further increase the potential of the proposed method. For instance, incrementally growing versions of the SOM (which does not require the user to specify the size of the map beforehand) can overcome the static architecture of the neural network (Hong *et al.* 2006). Active learning algorithms, such as those proposed in Wetmore *et al.* (2005), can also be used to speed up the training of the SOM. In the same vein, variants of hierarchical clustering methods that consider additional information in the merging criteria (such as inter-cluster and intra-cluster distance, and spatial information of the image data) can also be explored here (Tran *et al.* 2005, Xu and Wunsch II 2005).

Moreover, it is important to observe that the performance of the proposed method highly depends on the quality of the SOM mapping. The use of the procedures and techniques presented here assumes that the SOM has been successfully trained. In this way, questions about topological distortions and density approximation produced by the mapping of the net can be better investigated. There are a number of contributions to this subject in the literature. In Peres and Netto (2004), for example, a technique to estimate the ideal topographical dimension of the SOM output space is presented using fractal theory and fuzzy logic. In particular, the SOM is known to show an asymptotic level density following a power law with a  $2/3$  magnification exponent (Ritter 1991, Claussen and Schuster 2002). Recent clustering models are based on the power law assumption, such as the network classification proposed by Hadzikadic *et al.* (2005), and many real networks in nature and society also follow a power law degree distribution (Faloutsos *et al.* 1999, Liljeros *et al.* 2001). Therefore, new methods and corresponding numerical procedures were recently introduced to estimate scaling exponents of power law degree distributions and hierarchical clustering functions for complex networks (Bo *et al.* 2006). Currently, the estimation of the scaling exponents of the power law is achieved by the employment of detrended fluctuation analysis (DFA), which has already proved its usefulness in several complex systems, such as surface air-pollutants (Varotsos *et al.* 2005), aerosol index (Varotsos *et al.* 2006), total ozone content (Varotsos 2005a,b), global tropospheric temperature (Varotsos and Kirk-Davidoff 2006), and carbon dioxide fluctuations (Varotsos *et al.* 2007).

Finally, in addition to the test images utilized in the experiments shown here, the proposed method has also been applied to other high and medium resolution images, with satisfactory results. However, in the case of high spatial resolution images that have urban area patterns or a large number of buildings, the method needs to incorporate other attributes, in addition to spectral attributes, to obtain better performance.

### Acknowledgements

The authors are grateful to the anonymous referees who provided supportive and constructive comments that have assisted in the refinement of this paper. The first author is also grateful to FEEC/UNICAMP and FIP/PUC Minas, Brazil.

### References

- ARAI, K., 1992, A supervised thematic mapper classification with a purification of training samples. *International Journal of Remote Sensing*, **13**, pp. 2039–2049.
- ATKINSON, P.M. and TATNALL, A.R.L., 1997, Introduction neural networks in remote sensing. *International Journal of Remote Sensing*, **18**, pp. 699–709.
- AZCARRAGA, A.P., 2000, Assessing self-organization using order metrics. In *Proceedings of IEEE-INNS-ENNS International Joint Conference on Neural Networks*, July 2000, Como, Italy, **6**, pp. 159–164.
- BACHMANN, C.M., 2003, Improving the performance of classifiers in high-dimensional remote sensing applications: an adaptive resampling strategy for error-prone exemplars (ARESEPE). *IEEE Transactions on Geoscience and Remote Sensing*, **41**, pp. 2101–2112.
- BALL, G. and HALL, D., 1967, A clustering technique for summarizing multivariate data. *Behavior Science*, **12**, pp. 153–155.
- BANDYOPADHYAY, S., 2005, Satellite image classification using genetically guided fuzzy clustering with spatial information. *International Journal of Remote Sensing*, **26**, pp. 579–593.
- BARALDI, A., BLONDA, P., PARMIGGIANI, F. and SATALINO, G., 2000, Contextual clustering for image segmentation. *Optical Engineering*, **39**, pp. 1–17.
- BEZDEK, J.C. and PAL, N.R., 1998, Some new indexes of cluster validity. *IEEE Transactions on Systems Man and Cybernetics*, **28**, pp. 301–315.
- BO, Y., WEN-QI, D. and ZHONG, C., 2006, New method to estimate scaling exponents of power-law degree distribution and hierarchical clustering function for complex networks. *Applied Mathematics and Mechanics*, **27**, pp. 1475–1479.
- CASSANO, E.N., LYNCH, A.H., CASSANO, J.J. and KOSLOW, M.R., 2006, Classification of synoptic patterns in the western Arctic associated with extreme events at Barrow, Alaska, USA. *Climate Research*, **30**, pp. 83–97.
- CHEN, D. and STOW, D., 2002, The effect of training strategies on supervised classification at different spatial resolutions. *Photogrammetric Engineering and Remote Sensing*, **68**, pp. 1155–1162.
- CHUNG, K. and LIN, J., 2007, Faster and more robust point symmetry-based K-means algorithm. *Pattern Recognition*, **40**, pp. 410–422.
- CLAUSSEN, J.C. and SCHUSTER, H.G., 2002, Asymptotic level density of the elastic net self-organizing feature map. *Lecture Notes in Computer Science*, **2415**, pp. 939–944.
- COHN, D., ATLAS, L. and LADNER, R., 1994, Improving generalization with active learning. *Machine Learning*, **15**, pp. 201–221.
- COLÉOU, T., POUPON, M. and AZBEL, K., 2003, Unsupervised seismic facies classification: a review and comparison of techniques and implementation. *The Leading Edge*, **22**, pp. 942–953.
- COSTA, J.A.F. and NETTO, M.L.A., 2001, Clustering of complex shaped data sets via Kohonen maps and mathematical morphology. In *Proceedings of the SPIE Conference on Data Mining and Knowledge Discovery*, April 2001, Orlando, FL, **4384**, pp. 16–27.
- DEAN, A.M. and SMITH, G.M., 2003, An evaluation of per-parcel land cover mapping using maximum likelihood class probabilities. *International Journal of Remote Sensing*, **24**, pp. 2905–2920.
- DUDA, T. and CANTY, M., 2002, Unsupervised classification of satellite imagery: choosing a good algorithm. *International Journal of Remote Sensing*, **23**, pp. 2193–2212.

- ESPINOSA, C.H., REDONDO, M.F. and SOSPEDRA, J., 2004, Some experiments on ensembles of neural networks for hyperspectral image classification. *Lecture Notes in Computer Science*, **3213**, pp. 677–684.
- FALOUTSOS, M., FALOUTSOS, P. and FALOUTOS, C., 1999, On power-law relationships of the Internet topology. *ACM SIGCOMM Computer Communications Review*, **29**, pp. 251–262.
- FAYOS, J. and FAYOS, C., 2007, Wind data mining by Kohonen neural networks, PLoS ONE. 2007, **2**, p. e210.
- FOODY, G.M., 1999, The significance of border training patterns in classification by a feedforward neural network using back propagation learning. *International Journal of Remote Sensing*, **20**, pp. 3549–3562.
- FOODY, G.M. and MATHUR, A., 2004, Toward intelligent training of supervised image classifications: directing training data acquisition for SVM classification. *Remote Sensing of Environment*, **93**, pp. 107–117.
- FRIZZELLE, B.G. and MOODY, A., 2001, Mapping continuous distributions of land cover: a comparison of maximum-likelihood estimation and artificial neural networks. *Photogrammetric Engineering and Remote Sensing*, **67**, pp. 693–705.
- GIACINTO, G., ROLI, F. and FUMERA, G., 2000, Unsupervised learning of neural network ensembles for image classification. In *Proceedings of IEEE-INNS-ENNS International Joint Conference on Neural Networks*, July 2000, Como, Italy, **3**, p. 155–159.
- GONÇALVES, M.L., 1998, A neural system for the patterns recognition of remote sensing images. In *Proceedings of The Fifth International Conference on Neural Information Processing (ICONIP'98)*, October 1998, Kitakyushu, Japan, **2**, pp. 360–364.
- GONÇALVES, M.L., NETTO, M.L.A. and ZULLO JÚNIOR, J., 1996, A neural architecture for image classification using kohonen feature extraction. In *Proceedings of XVIII International Congress for Photogrammetry and Remote Sensing (ISPRS, Commission II)*, July 1996, Vienna, Austria, **31**, pp. 117–122.
- GONÇALVES, M.L., NETTO, M.L.A., COSTA, J.A.F. and ZULLO JÚNIOR, J., 2005, Automatic remotely sensed data clustering by tree-structured self-organizing maps. In *Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS'05)*, July 2005, Seoul, Korea, **1**, p. 4.
- GONÇALVES, M.L., NETTO, M.L.A., COSTA, J.A.F. and ZULLO JÚNIOR, J., 2006, Data clustering using self-organizing maps segmented by mathematic morphology and simplified cluster validity indexes. In *Proceedings of IEEE International Joint Conference on Neural Networks*, July 2006, Vancouver, Canada, **1**, pp. 8854–8861.
- HADJIMITSIS, D.G., EVANGELOU, I., LAZAKIDOU, A. and CLAYTON, C.R.I., 2003, Unsupervised classification of remotely sensed images using self-organizing maps for mapping land-cover changes. In *Proceedings of the Remote Sensing and Photogrammetry Society Annual Conference*, September 2003, Nottingham, UK.
- HADZIKADIC, M., QIN, Z. and SUN, M., 2005, Network classification. *Journal of Experimental and Theoretical Artificial Intelligence*, **17**, pp. 35–44.
- HALKIDI, M. and VAZIRGIANNIS, M., 2002, Clustering validity assessment using multi representatives. In *Proceedings of SETN Conference*, April 2002, Thessaloniki, Greece.
- HANSEN, M., DUBAYAH, R. and DEFRIES, R., 1996, Classification trees: an alternative to traditional land cover classifiers. *International Journal of Remote Sensing*, **17**, pp. 1075–1081.
- HAYKIN, S., 1994, *Neural Networks: A Comprehensive Foundation* (Englewood Cliffs, NJ: Prentice-Hall).
- HONG, Y., HSU, K.L., SOROOSHIAN, S. and GAO, X.G., 2005, Self-organizing nonlinear output (SONO): a neural network suitable for cloud patch-based rainfall estimation from satellite imagery at small scales. *Water Resources Research*, **41**, Article no. W03008, 8 March 2005.

- HONG, Y., CHIANG, Y., LIU, Y., HSU, K.L. and SOROOSHIAN, S., 2006, Satellite-based precipitation estimation using watershed segmentation and growing hierarchical self-organizing feature mapping techniques. *International Journal of Remote Sensing*, **27**, pp. 5165–5184.
- HUANG, K., 2002, The use of a newly developed algorithm of divisive hierarchical clustering for remote sensing image analysis. *International Journal of Remote Sensing*, **23**, pp. 3149–3168.
- HUANG, J.Z., NG, M.K., RONG, H. and LI, Z., 2005, Automated variable weighting in *K*-means type clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **27**, pp. 657–668.
- Ji, C.Y., 2000, Land-use classification of remotely sensed data using Kohonen self-organizing feature map neural networks. *Photogrammetric Engineering and Remote Sensing*, **66**, pp. 1451–1460.
- Ji, M., 2003, Using fuzzy sets to improve cluster labeling in unsupervised classification. *International Journal of Remote Sensing*, **24**, pp. 657–671.
- KASKI, S. and LAGUS, K., 1996, Comparing self-organizing maps. In *Proceedings of International Conference on Artificial Neural Networks (ICANN'96), Lectures Notes in Computer Science*, **1112**, pp. 809–814.
- KELLY, M., SHAARI, D., GUO, Q. and LIU, D., 2004, A comparison of standard and hybrid classifier methods for mapping hardwood mortality in areas affected by 'sudden oak death'. *Photogrammetric Engineering and Remote Sensing*, **70**, pp. 1229–1239.
- KHAN, S.S. and AHMAD, A., 2004, Cluster center initialization algorithm for *K*-means clustering. *Pattern Recognition Letters*, **25**, pp. 1293–1302.
- KIANG, M.Y., 2001, Extending the Kohonen self-organizing map networks for clustering analysis. *Computational Statistics and Data Analysis*, **38**, pp. 161–180.
- KOHONEN, T., 1997, *Self-Organizing Maps*, Second Edition (Berlin: Springer-Verlag).
- KOHONEN, T., KASKI, S. and LAPPALAINEN, H., 1997, Self-organized formation of various invariant-feature filters in the adaptive-subspace SOM. *Neural Computation*, **9**, pp. 1321–1344.
- LANCE, G.N. and WILLIAMS, W.T.A., 1967, General theory of classificatory sorting strategies: 1. Hierarchical Systems. *Computer Journal*, **9**, pp. 373–380.
- LAWRENCE, R., BUNN, A., POWELL, S. and ZAMBOM, M., 2004, Classification of remotely sensed imagery using stochastic gradient boosting as a refinement of classification tree analysis. *Remote Sensing of Environment*, **90**, pp. 331–336.
- LIKAS, A., VLASSIS, N. and VERBEEK, J., 2003, The global *K*-means clustering algorithm. *Pattern Recognition*, **36**, pp. 451–461.
- LILJEROS, F., EDLING, C.R. and AMARAL, L.A.N., 2001, The web of human sexual contacts. *Nature*, **411**, pp. 907–908.
- LILLESAND, T.M. and KIEFER, R.W., 2000, *Remote Sensing and Image Interpretation* (New York: John Wiley & Sons Inc).
- LIU, X., SKIDMORE, A.K. and VAN OOSTEN, H., 2002, Integration of classification methods for improvement of land-cover map accuracy. *Journal of Photogrammetry and Remote Sensing*, **56**, pp. 257–268.
- LIU, X., SKIDMORE, A.K. and VAN OOSTEN, H., 2003, An experimental study on spectral discrimination capability of a backpropagation neural network classifier. *International Journal of Remote Sensing*, **24**, pp. 673–688.
- LIU, A., BICHENG, L., CHEN, G. and ZHANGUI, X., 2005, A new ART neural networks for remote sensing image classification. *Lecture Notes in Computer Science*, **3611**, pp. 37–42.
- LO, C.P. and CHOI, J., 2004, A hybrid approach to urban land use/cover mapping using Landsat 7 Enhanced Thematic Mapper Plus (ETM+) images. *International Journal of Remote Sensing*, **25**, pp. 2687–2700.

- MAGNUSSEN, S., BOUDEWYN, P. and WULDER, M., 2004, Contextual classification of Landsat TM images to forest inventory cover types. *International Journal of Remote Sensing*, **25**, pp. 2421–2440.
- MARÇAL, A.R.S. and BORGES, J.S., 2005, Estimation of the ‘natural’ number of classes of a multispectral image. In *Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS’05)*, July 2005, Seoul, Korea, **6**, pp. 3788–3791.
- MARÇAL, A.R.S. and CASTRO, L., 2005, Hierarchical clustering of multispectral images using combined spectral and spatial criteria. *IEEE Geoscience and Remote Sensing Letters*, **2**, pp. 59–63.
- MERKL, D. and RAUBER, A., 1997, Alternative ways for cluster visualization in self-organizing maps. In *Proceedings of the Workshop on Self-organizing Maps (WSOM97)*, June 1997, Finland, pp. 106–111.
- MILLIGAN, G.W. and COOPER, M.C., 1985, An examination of procedures for determining the number of clusters in a data set. *Psychometrika*, **50**, pp. 159–179.
- MITRA, P., SHANKAR, U. and PAL, S.K., 2004, Segmentation of multispectral remote sensing images using active support vector machines. *Pattern Recognition Letters*, **25**, pp. 1067–1074.
- PAL, N.R. and BISWAS, J., 1997, Cluster validation using graph theoretic concepts. *Pattern Recognition*, **30**, pp. 847–857.
- PAL, M. and MATHER, P.M., 2002, A comparison of decision tree and backpropagation neural network classifiers for land use classification. In *Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS’02)*, **1**, pp. 503–505.
- PAL, S.K., BANDYOPADHYAY, S. and MURTHY, C.A., 2001, Genetic classifiers for remotely sensed images: comparison with standard methods. *International Journal of Remote Sensing*, **22**, pp. 2545–2569.
- PAL, M. and MATHER, P.M., 2005, Support vector machines for classification in remote sensing. *International Journal of Remote Sensing*, **26**, pp. 1007–1011.
- PASQUARIELLO, G., ANCONA, N., BLONDA, P., TARANTINO, C., SATALINO, G. and D’ADDABBO, A., 2002, Neural network ensemble and support vector machine classifiers for the analysis of remotely sensed data: a comparison. In *Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS’02)*, **1**, pp. 509–511.
- PERES, S.M. and NETTO, M.L.A., 2004, A fractal fuzzy approach to clustering tendency analysis. *Lecture Notes in Computer Science*, **3171**, pp. 395–404.
- RICHARDS, J.A., 2005, Analysis of remotely sensed data: the formative decades and the future. *IEEE Transactions on Geoscience and Remote Sensing*, **43**, pp. 422–432.
- RITTER, H., 1991, Asymptotic level density for a class of vector quantization processes. *IEEE Transactions on Neural Networks*, **2**, pp. 173–175.
- SEZGIN, M., ERSOY, O.K. and YAZGAN, B., 2004, Segmentation of remote sensing images using multistage unsupervised learning. In *Proceedings of SPIE, Applications of Digital Image Processing XXVII*, August 2004, Colorado, USA, **5558**, pp. 616–623.
- SHAH, C.A., ARORA, M.K. and VARSHNEY, P.K., 2004, Unsupervised classification of hyperspectral data: an ICA mixture model based approach. *International Journal of Remote Sensing*, **25**, pp. 481–487.
- SOROOSHIAN, S., HSU, K., GAO, X., GUPTA, H.V., IMAM, B. and BRAITHWAITE, D., 2000, Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. *Bulletin American Meteorology Society*, **81**, pp. 2035–2046.
- STATHAKIS, D. and VASILAKOS, A., 2006, Comparison of computational intelligence based classification techniques for remotely sensed optical image classification. *IEEE Transactions on Geoscience and Remote Sensing*, **44**, pp. 2305–2318.
- STRECKER, U. and UDEN, R., 2002, Data mining of 3D poststack seismic attribute volumes using Kohonen self-organizing maps. *The Leading Edge*, **21**, pp. 1032–1037.

- SU, M. and CHOU, C., 2001, A modified version of the *K*-means algorithm with a distance based on cluster symmetry. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **23**, pp. 674–680.
- SUNAR ERBEK, F., ÖZKAN, C. and TABERNER, M., 2004, Comparison of maximum likelihood classification method with supervised artificial neural network algorithms for land use activities. *International Journal of Remote Sensing*, **25**, pp. 1733–1748.
- SWAIN, P.H. and DAVIS, S.M., 1978, *Remote Sensing: The Quantitative Approach* (New York: McGraw-Hill).
- TRAN, T.N., WEHRENS, R. and BUYDENS, L.M.C., 2003, Sparef: a clustering algorithm for multispectral images. *Analytica Chimica Acta*, **490**, pp. 303–312.
- TRAN, T.N., WEHRENS, R. and BUYDENS, L.M.C., 2005, Clustering multispectral images: a tutorial. *Chemometrics and Intelligent Laboratory Systems*, **77**, pp. 3–17.
- TSO, B. and OLSEN, R.C., 2005, Combining spectral and spatial information into hidden Markov models for unsupervised image classification. *International Journal of Remote Sensing*, **26**, pp. 2113–2133.
- VAN COILLIE, F.M.B., VERBEKE, L.P.C. and DE WULF, R.R., 2004, Previously trained neural networks as ensemble members: knowledge extraction and transfer. *International Journal of Remote Sensing*, **25**, pp. 4843–4850.
- VAROTSOS, C., 2005a, Power-law correlations in column ozone over Antarctica. *International Journal of Remote Sensing*, **26**, pp. 3333–3342.
- VAROTSOS, C., 2005b, Modern computational techniques for environmental data; application to the global ozone layer. *Lecture Notes in Computer Science*, **3516**, pp. 504–510.
- VAROTSOS, C. and KIRK-DAVIDOFF, D., 2006, Long-memory processes in ozone and temperature variations at the region 60 degrees S – 60 degrees N. *Atmospheric Chemistry and Physics*, **6**, pp. 4093–4100.
- VAROTSOS, C., ASSIMAKOPOULOS, M. and EFSTATHIOU, M., 2007, Technical note: long-term memory effect in the atmospheric CO<sub>2</sub> concentration at Mauna Loa. *Atmospheric Chemistry and Physics*, **7**, pp. 629–634.
- VAROTSOS, C., ONDOV, J. and EFSTATHIOU, M., 2005, Scaling properties of air pollution in Athens, Greece and Baltimore, Maryland. *Atmospheric Environment*, **39**, pp. 4041–4047.
- VAROTSOS, C., ONDOV, J., CRACKNELL, A., EFSTATHIOU, M. and ASSIMAKOPOULOS, M., 2006, Long-range persistence in global aerosol index dynamics. *International Journal of Remote Sensing*, **27**, pp. 3593–3603.
- VESANTO, J. and ALHONIEMI, E., 2000, Clustering of the self-organizing map. *IEEE Transactions on Neural Networks*, **11**, pp. 586–602.
- VILMANN, T., MERENYI, E. and HAMMER, B., 2003, Neural maps in remote sensing image analysis. *Neural Networks*, **16**, pp. 389–403.
- WAGSTAFF, K., ROGERS, S. and SCHROEDL, S., 2001, Constrained *K*-means clustering with background knowledge. In *Proceedings of 8th International Conference Machine Learning*, pp. 577–584.
- WANG, F., 1990, Fuzzy supervised classification of remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, **28**, pp. 194–201.
- WANG, S.J. and ZHOU, C., 2002, Using remote sensing to estimate the change of carbon storage: a case study in the estuary of Yellow River delta. *International Journal of Remote Sensing*, **23**, pp. 1565–1580.
- WETMORE, L., HEYWOOD, M.I. and ZINCIR-HEYWOOD, A.N., 2005, Speeding up the self-organizing features map using dynamic subset selection. *Neural Processing Letters*, **22**, pp. 17–32.
- WILKINSON, G., 2005, Results and implications of a study of fifteen years of satellite image classification experiments. *IEEE Transactions on Geoscience and Remote Sensing*, **43**, pp. 433–440.
- WILSON, H.G., BOOTS, B. and MILLWARD, A.A., 2002, A comparison of hierarchical and partitioning clustering techniques for multispectral image classification. In *Proceedings*

- of *IEEE International Geoscience and Remote Sensing Symposium (IGARSS'02)*, June 2002, Toronto, Canada, **3**, pp. 1624–1626.
- WU, S. and CHOW, T.W.S., 2004, Clustering of the self-organizing map using a clustering validity index based on inter-cluster and intra-cluster density. *Pattern Recognition*, **37**, pp. 175–188.
- XIANG, M., HUNG, C., PHAM, M., KUO, B. and COLEMAN, T., 2005, Parallelepiped multispectral image classifier using genetic algorithms. In *Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS'05)*, July 2005, Seoul, Korea, **1**, p. 4.
- XU, R. and WUNSCH II, D., 2005, Survey of clustering algorithms. *IEEE Transactions on Neural Networks*, **16**, pp. 645–678.
- ZHONG, Y., ZHANG, L. and LI, P., 2005, Multispectral remote sensing image classification based on simulated annealing clonal selection. In *Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS'05)*, July 2005, Seoul, Korea, **1**, pp. 3745–3748.



# Data Clustering using Self-Organizing Maps segmented by Mathematic Morphology and Simplified Cluster Validity Indexes: an application in remotely sensed images

Márcio L. Gonçalves, Márcio L. de Andrade Netto, José A. Ferreira Costa, Jurandir Zullo Jr.

**Abstract**—This paper presents a cluster analysis method which automatically finds the number of clusters as well as the partitioning of a data set without any type of interaction with the user. The data clustering is made using the Self-Organizing (or Kohonen) Map (SOM). Different partitions of the trained SOM are obtained from different segmentations of the U-matrix (a neuron-distance image) that are generated by means of mathematical morphology techniques. The different partitions of the trained SOM produce different partitions for the data set which are evaluated by cluster validity indexes. To reduce the computational cost of the cluster analysis process this work also proposes the simplification of cluster validity indexes using the statistical properties of the SOM. The proposed methodology is applied in the cluster analysis of remotely sensed images.

## I. INTRODUCTION

The advances in computer and electronic technologies and decreasingly cost of memory storage systems have been enabling large amounts of data to be available in many application areas. An example is the remote sensing of the earth surface from satellite or airborne scanners. Large volumes of remotely sensed images are being generated from an increasing number of sophisticated airborne and space borne sensor systems, and while there is no substitute for a trained analyst, exploitation of this data on a large scale requires consistent automatic data exploration tools [1], [2].

Clustering is a common technique for data analysis that can be used to reveal structures and to identify “natural” groupings on the image data. Clustering algorithms are under intense development and are often applied to many fields related to pattern recognition, biological sciences, medicine, social sciences, computer vision, etc. Normally, based on a similarity or dissimilarity measure, clustering consists of partitioning a data set into subsets or clusters, so that the data in each cluster have high similarity in

comparison to another, but are very dissimilar to data in other clusters.

Clustering methods fall into two types: hierarchical and partitioning approaches [3]. Hierarchical clustering builds a hierarchy of clusters from individual elements. The traditional representation of this hierarchy is a tree, with individual elements at one end and a single cluster with every element at the other. Hierarchical clustering works well with small data sets and can handle outliers very well but its computation is very expensive and therefore it is not feasible for large data sets. In the particular case of remotely sensed images, the improvement in sensor systems sensibilities has increased drastically the resolution in the spatial domain of the images. As a result, the size of images has increased too. The size of typical data set can easily get up to millions of pixels becoming impracticable the use of hierarchical clustering methods in terms of memory and processing time [3], [4]. Partitioning clustering approaches produce one partition with  $K$  groups usually by optimizing some objective criterion. Variants of  $K$ -clustering, such as  $K$ -means and ISODATA [5], are the partitioning clustering methods most widely used, mainly for satellite images. The major drawbacks of partitioning methods are the choice of the number of clusters in advance and the initial  $K$  clusters seeds.

Competitive neural networks also have been applied to data clustering [6]-[11]. The Self-Organizing (or Kohonen) Map (SOM) is the model most widely used [11]. However, in classical clustering methods based on the SOM usually the researcher sets the number of neurons equal to the expected number of clusters, or manually segments a two-dimensional map using some a priori knowledge of the data. For example, in [9] the trained SOM is used as a visual auxiliary tool which enables an image analyst to identify the number of clusters existing in a remotely sensed image. The analyst interacts in the clusters analysis process by visualizing and interpreting the trained map.

In this work we present a cluster analysis method that focuses the usage of SOM as a clustering tool and uses techniques for automatic partitioning and labeling of SOM networks in clusters of neurons that may be used to represent the data clusters. The partitioning of the trained SOM is based on the automatic segmentation method of SOM proposed by Costa & Netto [8]. Visualization of neuron's relations is made using the Unified distance matrix (U-matrix), which is a neuron-distance image. To segment the U-matrix the Costa & Netto algorithm applies the

Márcio L. Gonçalves is with the Pontifical Catholic University of Minas Gerais, MG, Brazil and with the State University of Campinas – UNICAMP, SP, Brazil (corresponding author to provide phone: +55 (35) 3729-9276; fax: +55 (35) 3729-9201; e-mail: marcio@pucpcaldas.br).

Márcio L. Andrade Netto is with the School of Electrical and Computer Engineering, State University of Campinas – UNICAMP, SP, Brazil (e-mail: marcio@dca.fee.unicamp.br).

José Alfredo F. Costa is with the Department of Electrical Engineering, Federal University of Rio Grande do Norte, Natal, RN, Brazil (e-mail: alfredo@dee.ufrn.br).

Jurandir Zullo Jr. is with Agrometeorology Research Center, State University of Campinas, UNICAMP, SP, Brazil (e-mail: jurandir@cpa.unicamp.br).

mathematical morphology segmentation method, watershed, using a markers image to regularize the segmentation process. This same strategy also is applied in our proposal, however instead of defining only one markers image to segment the U-matrix (as presented in [8]), several markers images are considered and, therefore, different segmentations of the U-matrix are obtained. Each one of these segmentations is associated with the neurons of the SOM, allowing determining different partitions for trained SOM and, consequently, different partitions for the data set. To evaluate the quality of these partitions and to determine automatically the best partition of the data set this work also proposes the use of simplified cluster validity indices. Instead of using all the points of the data set the validity indices are calculated using only the neurons of the trained SOM. The vector quantization of the data space performed by SOM learning process allows that the neurons of the map can be used as representative vectors (prototypes) of the data. This approach reduces considerably the processing time of the cluster analysis process and becomes possible its use in large data sets.

To test the efficiency of the proposed clustering method some experiments with remotely sensed images are presented. Besides being part of an application area extremely important for the real world we believe that clustering remotely sensed images is a significant test to validate cluster analysis methods. Differently of synthetic data sets the remotely sensed images generally do not present data with known probabilities distribution and well defined geometries. The large size of the images, higher-dimensional feature space, the existence of noises/outliers, overlapping clusters and unknown number of clusters are some of the main problems encountered when clustering remote sensing images [4], [10].

## II. THE SELF-ORGANIZING MAP

The SOM is one of the most popular artificial neural network algorithms. SOM is based on unsupervised competitive learning, which means that the training is entirely data-driven and that the neurons of the map compete with each other [11].

The SOM consists essentially of two layers of neurons: input-layer  $I$  and the unit layer  $U$ . Each neuron  $i$  is represented by an  $n$ -dimensional weight vector  $w_i = [w_{i1}, w_{i2}, \dots, w_{in}]^T$ , where  $n$  is equal to the dimension of the input vectors. The SOM defines a mapping from the high dimensional input data space onto a regular, usually, two-dimensional array of nodes. The neurons of the map are connected to adjacent neurons by a neighborhood relation dictating the structured of the map. In the 2-dimensional case the neurons of the map can be arranged either on a rectangular or hexagonal lattice.

The basic SOM algorithm is iterative. At each training step  $t$ , a sample data vector  $x(t)$  is randomly chosen from the training set. Distances between  $x(t)$  and all the weight vectors  $w_i$  are computed. The winner neuron, denoted by  $c$ , is the neuron with the weight vector closest to  $x(t)$ :

$$c = \arg \min_i \{ \|x(t) - w_i\| \} \quad (1)$$

where  $\|\cdot\|$  is the distance measure. The weight vectors of winner neuron as well as the neighboring nodes are moved closer to the input data vector. The SOM update rule for the weight vector of the unit  $i$  is

$$w_i(t+1) = w_i(t) + h_c(t)[x(t) - w_i(t)] \quad (2)$$

where  $h_c(t)$  is the neighborhood kernel around the winner unit  $c$  at step  $t$ .

### A. Properties of the SOM

The SOM has some important properties. Once the SOM algorithm has converged, the feature map computed displays important statistical characteristics of the input space, summarized as follows [6]:

1) *Vector quantization*: The basic aim of the SOM algorithm is to store a large set of input data  $x$  by finding a smaller set of prototypes  $w_i$  (weight vectors of neurons). The SOM algorithm is a vector quantization algorithm, which provides a good approximation to the input space.

2) *Topological Ordering*: The feature map computed by SOM algorithm is topologically ordered. Similar input vectors are mapped close to each other, while dissimilar ones are mapped far apart.

3) *Density Matching*: The SOM reflects variations in the statistics of the input distribution. Regions in the input space from which sample vectors are drawn with a high probability of occurrence are mapped onto larger domains of the output space.

### B. Interpretation of the trained SOM

Although the SOM presents important properties on the input data, the trained neural network requires some additional procedures to enable a suitable interpretation of the data clusters.

In the literature we can find some works proposing mechanisms and procedures to interpret the neurons of a trained SOM in automatic or semi-automatic way [8], [13]-[17]. In [13] there are two SOM layers for clustering. The number of the neurons on the second SOM layer is equal to the desired number of clusters. The task of the second layer is analogous to clustering of the SOM by  $K$ -means algorithm. The desired number of clusters is required and the proposed method is only feasible for hyper-spherical-shaped clusters. In [14] is proposed an agglomerative contiguity-constrained clustering method on the SOM. The method groups the output from SOM based on a minimal distance criterion to merge the neighboring node together. The algorithm presented in [15] extends the approach in [14] using the minimal variance criterion and obtains better clustering resulted. However, the algorithms presented in [14]-[15] need to recalculate the center after two clusters are merged and are appropriate only for clusters with hyper-spherical or hyper-ellipsoidal shapes. In [16] hierarchical agglomerative clustering and partitioning clustering using  $K$ -means are proposed to clustering of the SOM. The

experiments in [16] indicated that clustering the SOM instead of directly clustering the data is computationally effective approach. However, hierarchical clustering algorithms in [16] use only inter-cluster distance criterion to cluster the output neurons and, in a similar way that algorithm presented in [13], the batch K-means algorithm presented in [16] require the desired number of clusters in advance and is only feasible for hyper-spherical-shaped clusters. In [17] the proposed method to clustering of the SOM uses the clustering validity index locally to determine which pair of clusters to be merged. Compared with others classical clustering methods on the SOM, the algorithm proposed in [17] utilizes more information about the data in each cluster in addition to inter-cluster distances. The experimental results on the four data sets presented in [17] demonstrate that the proposed clustering algorithm has superior performance. In [8] the SOM is partitioned using mathematical morphology operations on the U-matrix (described in the next subsection). The obtained results are labeled and connected regions of neurons that defines, in the input space, complex and non-parametric geometries which approximately describes the shape of the clusters. The process of map partitioning is extended recursively. Each cluster of neurons gives rise to a new map, which are trained with the subset of data were classified to it. The advantages of the algorithm proposed in [8] include the automatic determination of the number of clusters and the fact that the discovered clusters can have a non-parametric geometry enabling complex shaped clusters to be detected.

As it will be shown in section III, the cluster analysis method presented in this work is based on the partitioning method of the SOM presented in [8]. In our approach the SOM is partitioned through the segmentation of the U-matrix by means of operations of mathematical morphology, however, a simplified cluster validity index is used to evaluate concomitantly the quality of different partitions of the SOM that are generated from different segmentations of the U-matrix. This strategy enables that the best partition for the data set can be obtained without the necessity to train a quantity of SOM networks as in the method proposed in [8].

### C. The U-matrix Method

The Unified Distance Matrix (U-matrix) method was developed by A. Ultsch [12] to detect nonlinearities in the resulting SOM mapping. The basic idea is to use the same metric that was used during the learning to compute distances between adjacent reference vectors. This method can be used to visualize the topological structure of the SOM unit layer and therefore also the topology of the  $N$ -dimensional input space.

The U-matrix has dimension  $(2X-1) \times (2Y-1)$  considering a two-dimensional SOM with a rectangular lattice topology of size  $(X \times Y)$ . The U-matrix is constructed on top of the map. Let  $n$  be a neuron on the map,  $NN(n)$  be the set of immediate neighbors on the map,  $w(n)$  the weight vector associated with neuron  $n$ , then:

$$U\_height(n) = \sum_{m \in NN(n)} d(w(n) - w(m)) \quad (3)$$

where  $d(\cdot)$  is a distance measure that make use of the metric used in the SOM algorithm to construct the map. The U-matrix is a display of the  $U\_heights$  on top the grid positions of the neurons on the map and can be visualized as a three dimensional landscape. Altitudes or the high places on the U-matrix encode dissimilarities between neurons and correspond to cluster borders while valleys represent to map units that are similar.

## III. THE PROPOSED CLUSTER ANALYSIS METHOD

The cluster analysis strategy proposed in this work determines the best partition for a data set from the analysis of different segmentations of the U-matrix. Costa and Netto [8] proposed an efficient method based on mathematical morphology to segment the U-matrix. The method applies the images segmentation algorithm, watershed [18], using a markers image to regularize the segmentation process. This same approach also is applied in our proposal, however, instead of using only one markers image to segment the U-matrix (as proposed in [8]), a quantity of markers images are considered and, therefore, different segmentations of the U-matrix are obtained. Each one of these segmentations is associated with the neurons of the SOM, allowing to determine different partitions for the map which define different partitions for the data set in the input space. To select the best one among different partitions, each of these can be evaluated using some cluster validity index computed in simplified way. The section IV describes as we simplify the calculation of cluster validity indexes.

Given the U-matrix image  $U$ , the following steps are performed to obtain the markers images set to the image  $U$ :

1. Filtering: create the image  $U_1$  by removing any pore with area less or equal than three pixels.
2. For  $k = 1, \dots, f_{max}$ , where  $f_{max}$  is the highest gray level of the image  $U_1$  {
  - 2.1. Create the binary image  $U_1^k$  that corresponds to the conversion  $U_1$  to a binary image using as threshold  $k$ .
  - 2.2. Obtain  $N_{rc}^k$ , the number of connected regions of  $U_1^k$  .}
3. Obtain the most persistent values of number of connected regions that correspond to the plateaus with sizes more than three contiguous gray levels in the plot of  $N_{rc}^k$  versus  $k$ .
4. Obtain the set of all markers images,  $S_m = \{U_2^{k_1}, U_2^{k_2}, \dots, U_2^{k_n}\}$ , where  $k_1, k_2, \dots, k_n$  are initial values of the plateaus chosen in the previous step.

Although the used procedure to find the markers of the U-matrix is similar to that one presented in [8], in our approach the steps 4 and 5 determine a set of markers images instead of only one image considering all the significant plateaus.

Thus, the general strategy for the partitioning of a data set  $X$  from a trained SOM can be summarized as follows:

1. Obtain the U-matrix using the trained SOM.
2. Find the set of all markers images to the U-matrix,  $S_m = \{U_2^{k_1}, U_2^{k_2}, \dots, U_2^{k_n}\}$  (as described previously).
3. For each markers image,  $U_2^{k_i}, i=1,2,\dots,n \{$ 
  - 3.1. Compute the watershed lines on the U-matrix.
  - 3.2. Assign a different label for each connected region (cluster of neurons) of the U-matrix.
  - 3.3. Copy the U-matrix labels to the corresponding neurons in the map.
  - 3.4. Label the data set  $X$  using the labeled SOM in the step 3.3 (each pattern in  $X$  is labeled with the same label of the neuron that is associated it).
  - 3.5. Apply a simplified cluster validity index to evaluate the data set  $X$  labeled (partitioned) in the step 3.4. }
4. Select the best partition for the data set  $X$  comparing the values of a cluster validity index calculated for each different partition of  $X$ .

The Fig. 1 illustrates the partitioning strategy proposed in this work. An important attribute of the proposed clustering algorithm is the distributed and multiple representation of neurons for the clusters in the input space. By using an image processing algorithm to segment the U-matrix, the obtained results are labeled and connected regions of neurons. Each region defines, in the input space, complex and non-parametric geometries which approximately describe the shape of the clusters. Therefore, the clustering algorithm presents a considerable potential to detect clusters with complex and varied shapes.

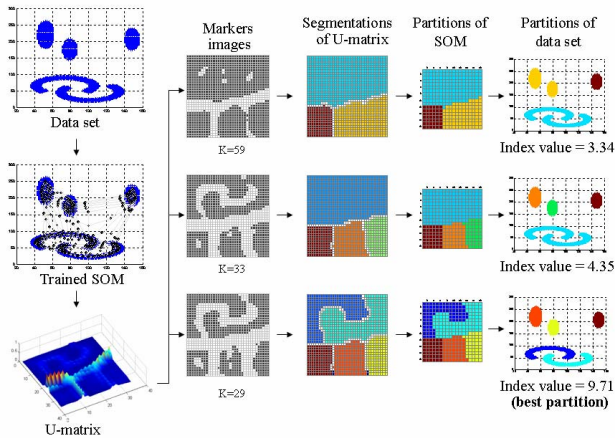


Fig. 1. Illustration of the general strategy for the partitioning of a data set.

#### IV. SIMPLIFICATION OF CLUSTER VALIDITY INDEXES

In a cluster analysis process in which different partitions for a data set are generated some evaluation criterion is necessary to compare the quality or correctness of the different partitions and to select the best of them. Different research communities have made a lot of efforts for clustering in the area of pattern recognition and, consequently, several cluster validity indexes have been proposed [19]-[22]. These indexes present varied characteristics and properties; however, the largest part of them presents a common drawback: the huge computation

effort required. When the number of clusters and the quantity of the input data are very large the use of cluster validity indices becomes prohibitive. The remotely sensed images are a typical example. As mentioned earlier, the spatial and spectral resolutions of these images has increased too, consequently, the size of representative data set and the dimension of features space can be extremely large. For this motive, cluster validity indexes are often ignored in remote sensing applications and are not available in most digital image processing packages [23].

In this paper we propose a way to reduce the high computational cost required by the cluster validity indexes becoming possible to apply them in large data sets. The idea consists in simplifying the calculation of cluster validity indexes using the data abstraction created by the SOM. As described before, the SOM provides a good approximation to the input space. A large set of input data can be represented by a smaller set of prototypes (weight vectors) of SOM, which approximately preserves the topology and the probability distribution of the input data. Thus, assuming that the partitioning of a data set has been executed through the partitioning of a trained SOM, the cluster validity indexes can be calculated using only the weight vectors ( $w_i$ ) and the frequency of *hits* ( $h(w_i)$ ) of the neurons instead of using all the points of the input data set. The frequency of *hits* of a prototype  $w_i$  is equivalent to the number of input patterns mapped onto neuron  $i$ .

Following the idea above, in this paper we present the simplification of four cluster validity indexes. Two well known indexes: Generalized Dunn ( $V_D$ ) [19] and Davies-Bouldin ( $V_{DB}$ ) [20], and two indexes proposed more recently which have presented good performances: PBM ( $V_{PBM}$ ) [21] and CDbw ( $V_{CDbw}$ ) [22]. Certainly others indexes can be simplified using the same approach presented here.

Consider an input data set  $X$ . Suppose that a SOM has been trained with the data of  $X$  and its map of neurons partitioned in  $K$  clusters. Consider also that the data set  $X$  has been partitioned in  $K$  clusters using the labeled SOM, in way that each input pattern belongs to the same cluster as its nearest neuron. Following the approach proposed here, instead of using all the points of a data cluster  $C_i$  of  $X$  the calculations of the validity indexes are made using only the prototypes of the cluster of neurons  $W_i$  which is associated to  $C_i$ . The formalization of these calculations is given as follows.

- *Simplified Generalized Dunn Index* ( $V_D^{SOM}$ ):

$$V_D^{SOM} = \min_{1 \leq i \leq K} \left\{ \min_{1 \leq j \leq K, j \neq i} \left\{ \frac{\delta(C_i, C_j)}{\max_{1 \leq k \leq K} \{\Delta(C_k)\}} \right\} \right\}, \quad (4)$$

where the distance  $\delta$  between the data clusters  $C_i$  and  $C_j$ , and the diameter  $\Delta$  of a data cluster  $C_k$  are given by:

$$\delta(C_i, C_j) = \frac{1}{N_i \cdot N_j} \sum_{p=1}^{r_i} \sum_{q=1}^{r_j} h(w_i^p) \cdot h(w_j^q) \cdot d(w_i^p, w_j^q) \quad (5)$$

and

$$\Delta(C_k) = 2 \cdot \left( \sum_{p=1}^{r_k} h(w_k^p) \cdot d(w_k^p, z_k) \right) / N_k, \quad (6)$$

where  $N_i$  and  $N_j$  are, respectively, the numbers of points of  $C_i$  and  $C_j$ .  $r_i$  and  $r_j$  represent, respectively, the numbers of prototypes of the clusters of neurons  $W_i$  and  $W_j$  (which are associated to  $C_i$  and  $C_j$ ).  $w_i^p$  represents the prototypes of  $W_i$ , and  $w_j^q$  represents the prototypes of  $W_j$ . As mentioned before,  $h(w_i^p)$  is the frequency of *hits* of the prototype  $w_i^p$ ,  $d$  is a distance function, and  $z_k$  is the centroid of data cluster  $C_k$  defined as:

$$z_k = \frac{1}{N_k} \sum_{j=1}^{r_k} h(w_k^j) \cdot w_k^j. \quad (7)$$

- *Simplified Davies-Bouldin Index* ( $V_{DB}^{SOM}$ ):

$$V_{DB}^{SOM} = \frac{1}{K} \sum_{i=1}^K \max_{j \neq i} \left\{ \frac{S_{i,q} + S_{j,q}}{d_{ij,t}} \right\} \quad (8)$$

The dispersion  $S_{i,q}$  of the data cluster  $C_i$  and the separation  $d_{ij,t}$  between the data clusters  $C_i$  and  $C_j$  are defined as:

$$S_{i,q} = \left\{ \frac{1}{N_i} \sum_{j=1}^{r_i} h(w_i^j) \cdot \|w_i^j - z_i\|_q \right\}^{1/q} \quad (9)$$

and

$$d_{ij,t} = \|z_i - z_j\|_t, \quad (10)$$

where  $q, t > 1$ ,  $q$  is a integer;  $z_i$  and  $z_j$  are, respectively, the centroids of  $C_i$  and  $C_j$  which are calculated as (7).  $N_i$  is the number of points of  $C_i$ ,  $h(w_i^j)$  is the frequency of *hits* of the prototype  $w_i^j$ , and  $r_i$  is the number of prototypes of the cluster of neurons  $W_i$ .

- *Simplified PBM Index* ( $V_{PBM}^{SOM}$ ):

$$V_{PBM}^{SOM} = \left( \frac{1}{K} \cdot \frac{E_0}{E_k} \cdot D_K \right), \quad (11)$$

with  $E_0 = \sum_{j=1}^P h(w_j) \cdot \|w_j - z\|$  and  $E_k = \sum_{k=1}^K \sum_{i=1}^{r_k} h(w_k^i) \cdot \|w_k^i - z_k\|$ ,

where  $h(w_k^i)$ ,  $z_k$  and  $r_k$  follow the same definitions of the previous indexes.  $P$  is the total number of neurons or prototypes ( $w_j$ ) of SOM, and  $z$  is the centroid of all the  $N$  points in  $X$  defined of the following way:

$$z = \frac{1}{N} \sum_{j=1}^P h(w_j) \cdot w_j \quad (12)$$

The term  $D_k$  in (11) is the maximum separation between the  $K$  data clusters, given by:

$$D_K = \max_{i,j=1,\dots,K} \|z_i - z_j\| \quad (13)$$

where  $z_i$  and  $z_j$  are defined as (7).

- *Simplified CDbw Index* ( $V_{CDbw}^{SOM}$ ):

$$V_{CDbw}^{SOM} = Intra * Sep. \quad (14)$$

*Intra* represents the intra-cluster density, which is given by:

$$Intra = \frac{1}{K} \sum_{i=1}^K \frac{1}{r_i} \sum_{j=1}^{r_i} density(v_i^j), \quad (15)$$

where  $v_i^j$  corresponds to the prototype  $w_i^j$  lightly dislocated in direction to the centroid of the data cluster  $C_i$ , which is calculated as (7),  $density(v_i^j) = \sum_{l=1}^{r_i} h(w_i^l) \cdot f(w_i^l, v_i^j)$ , where  $r_i$  e  $h(w_i^l)$  are defined as in the previous indexes, and  $f(w_i^l, v_i^j)$  is equal to 1, if  $\|w_i^l - v_i^j\| \leq stdev$ , or equal to 0, otherwise. *stdev* is the average standard deviation of the data clusters, which is given by:

$$stdev = \frac{1}{K} \sqrt{\sum_{i=1}^K stdev_i}, \quad (16)$$

where  $stdev_i$  is the standard deviation of the data cluster  $C_i$ , defined as:

$$stdev_i = \sqrt{\frac{\sum_{j=1}^{r_i} h(w_i^j) \cdot (w_i^j - z_i)}{(N_i - 1)}}, \quad (17)$$

being that  $r_i$ ,  $N_i$  and  $z_i$  follow the same definitions of the previous indexes.

The density in the between-cluster region is defined as follows:

$$Inter = \sum_{i=1}^K \sum_{j=1}^K \frac{\|close\_rep_i - close\_rep_j\|}{\|stdev_i\| + \|stdev_j\|} \cdot density(m_i^j), \quad (18)$$

where *close\_rep<sub>i</sub>* and *close\_rep<sub>j</sub>* are the closest pair of prototypes of the clusters of neurons  $W_i$  and  $W_j$ ;  $m_i^j$  is the middle point between this pair of points, and  $density(m_i^j) = \sum_{k=1}^{r_i+r_j} h(w_{ij}^k) \cdot f(w_{ij}^k, m_i^j) / (N_i + N_j)$ , where  $w_{ij}^k$  represents all the prototypes of  $W_i$  and  $W_j$ . The function  $f(w_{ij}^k, m_i^j)$  is equal to 1, if  $\|w_{ij}^k - m_i^j\| \leq (\|stdev_i\| + \|stdev_j\|) / 2$ , or equal to 0, otherwise.

The separation between clusters, *Sep* in (14), is given by:

$$Sep = \frac{\sum_{i=1}^K \sum_{j=1}^K \|close\_rep_i - close\_rep_j\|}{1 + Inter}. \quad (19)$$

Once that the simplified versions of the indexes follow basically its original algorithms a good partition is indicated by high values of  $V_D^{SOM}$ ,  $V_{PBM}^{SOM}$  and  $V_{CDBw}^{SOM}$ , and by low values of  $V_{DB}^{SOM}$ .

## V. EXPERIMENTAL RESULTS

In the experiments presented in this paper the proposed clustering method and the simplifications of the cluster validity indexes are applied to two remotely sensed images. The test image 1 (Fig. 2(a)) is composed by spectral bands 3, 4 and 5 of the TM/LANDSAT-5 satellite (provided by Image Generation Division of National Institute for Space Research - INPE, Brazil). The image has 512x512 pixels, and it shows the Manaus city and the crossing of rivers Negro and Solimões in the Amazonas state, Brazil. The test image 2 (Fig. 2(b)) is an IKONOS image also composed of three spectral bands (© Space Imaging, provided by Engesat/Brazil). The image has 1883x1607 pixels, and it shows irrigation pivots in the region of Andaraí in the Bahia state, Brazil. In the two experiments only three spectral bands for each image had been used because of the non-availability of the other bands.



Fig. 2. (a) Test image 1 (Landsat/TM). (b) Test image 2 (Ikonos).

A set of samples was collected from each one of the images. The samples were randomly and uniformly collected on all the space region of the images. The samples of test image 1 correspond to individual pixels, and the samples of test image 2 correspond to 5x5 pixels windows. The data set of test image 1 consisted of 10000 3-D samples, while the data set of test image 2 consisted of 40000 75-D samples.

In the two experiments were used two-dimensional SOMs. The weight initialization was linear and the training was done with the batch algorithm. The neighborhood function was Gaussian and during ordering the neighbor radius decreased with the training epochs, which in two cases were set 500. The simulations were performed in an AMD Athlon™ XP 2600+ processor with 512MB of fast PC2700 DDR memory using *Mathworks Matlab*. Some functions of the *SOM toolbox* were also used (available at <http://www.cis.hut.fi/projects/>).

For the data set of the test image 1 the size of SOM network was set 15 x 10 (i.e., 150 neurons in a two-dimensional lattice). The U-matrix was calculated from the trained SOM (Fig. 3(a) shows a 3D view of the U-matrix). The methodology presented in the section III determines

different partitions for a given U-matrix using different markers images. It uses information like the number of connected regions ( $N_{rc}^k$ ) for each gray level ( $k$ ) of the U-matrix image for a useful gray level range, which in turn is related to the distances between neighboring neurons. For the U-matrix image of this experiment (test image 1) were determined eight markers images. The markers images were obtained after thresholding the U-matrix by the lower gray levels from significant plateaus in the plot of  $N_{rc}^k$  versus  $k$  (Fig. 3(b)). In this case, the values of these gray levels  $k$  were 7, 24, 37, 94, 107, 114, 121, and 163. Therefore, in accordance with the proposed method eight different segmentations of U-matrix image were performed using the watershed method, one for each marker image. After, eight different partitions for the SOM were determined from the segmented U-matrix images and, consequently, eight different partitions for the data set of test image 1 were obtained.

For the data set of the test image 2 was applied the same procedure, however the size of SOM network used here was set 20 x 20. The Fig. 3(c) shows a 3D view of the U-matrix generated from the trained SOM. In this case, seven significant plateaus were detected in the plot  $N_{rc}^k$  versus  $k$  (Fig. 3(d)). The gray levels  $k$  correspondents to seven plateaus were 4, 45, 57, 68, 72, 91, and 115. Following the methodology proposed seven different segmentations of U-matrix were determined from the markers images using the watershed method and, therefore, seven different partitions for the data set of test image 2 were obtained.

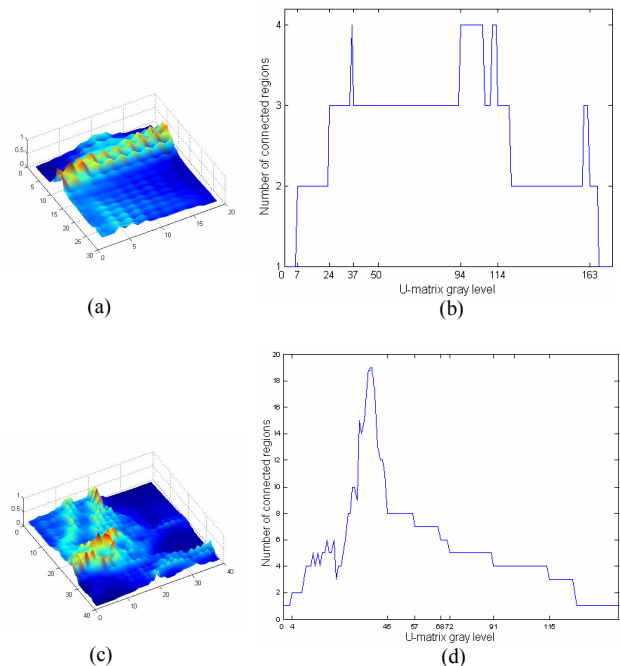


Fig. 3. (a) 3D view of the U-matrix for SOM of test image 1. (b) Number of connected regions versus image threshold of U-matrix (test image 1). (c) 3D view of the U-matrix for SOM of test image 2. (d) Number of connected regions versus image threshold of U-matrix (test image 2).



To evaluate the different partitions generated for the two data sets and to determine automatically the best partitions the methodology presented in this work considers the use of some cluster validity index simplified in conformity with the approach presented in the section IV. Among the four cluster validity indexes that were treated in the section IV we believe that the  $V_{CDbw}$  index is the more indicated to evaluate data clusters of remotely sensed images, because of the characteristics of the data and in virtue of the properties of the index [17],[22]. However, the original and simplified versions of all the four cluster validity indexes considered in this work were applied to evaluate the different partitions generated for the two data sets used in the experiments. The objective of this is to compare the results of the simplifications proposed to the indexes with the results of the original indexes.

Parallely to this, an image analyst was consulted to evaluate the two test images used in the experiments and to indicate the best partition for each one of the images among the different generated partitions. In accordance with the analyst's point of view, the test image 1 presents four predominant land cover classes and the best partition among the eight different partitions is that one generated by gray level  $k = 94$  corresponding to four clusters ( $N_{rc}^k = 4$ ). For the test image 2 the best partition among the seven different partitions is that one generated by gray level  $k = 45$  corresponding to eight clusters ( $N_{rc}^k = 8$ ).

The plots of the original and simplified versions of the cluster validity indexes calculated for each partition of the data sets 1 and 2 are shown in the Figs. 4 and 5, respectively. Although this work does not have the objective to analyze the performances of different cluster validity indexes, but of the simplifications proposed for these indices, it is observed that only the versions of the CDbw index correctly indicated the best partitions for the two data sets.  $V_{CDbw}$  and  $V_{CDbw}^{SOM}$  presented the higher values for the partition  $k = 94$  in the analysis of the data set 1, and higher values for the partition  $k = 45$  in the analysis of the data set 2 (coinciding with the indications of the image analyst).  $V_{PBM}$  and  $V_{PBM}^{SOM}$  correctly indicated the best partition only for the data set 2. The  $V_D$  and  $V_{DB}$  indexes, and its simplified versions  $V_D^{SOM}$  and  $V_{DB}^{SOM}$  did not indicated correctly the best partitions in the analyses of the two data sets (these indexes are more appropriate for the analysis of hyperspherical clusters [19],[20]).

Comparing the values of the original and simplified versions of the indexes (Figs. 4 and 5), it is possible to observe that exist approximation errors, as was waited. However, the simplified indices presented similar behavior to the original versions for all the data partitions. Therefore, in the two experiments presented here the simplified indices can perfectly be used to evaluate the different partitions. The diverse partitions are interpreted by the simplified indexes of

the same way that the original indexes, allowing a correct decision-making.

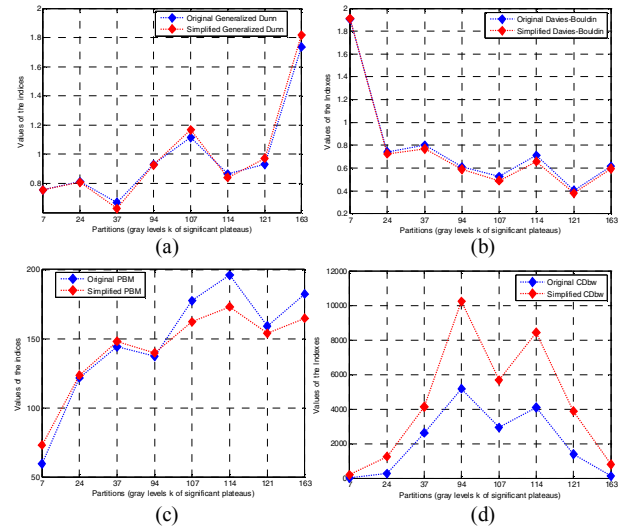


Fig. 4. Plots of the original and simplified versions of the indexes for the partitions of data set of test image 1: (a) Generalized Dunn. (b) Davies-Bouldin. (c) PBM. (d) CDbw. The blue markers represent the values of the original index and the red markers represent the simplified indices.

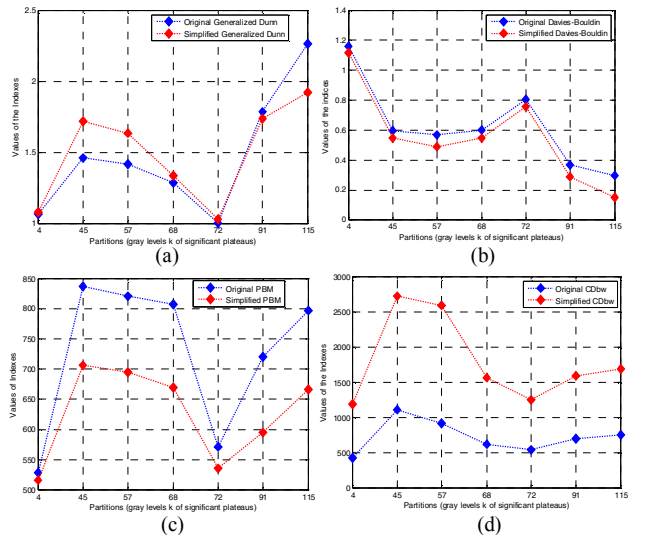


Fig. 5. Plots of the original and simplified versions of the indexes for the partitions of data set of test image 2: (a) Generalized Dunn. (b) Davies-Bouldin. (c) PBM. (d) CDbw. The blue markers represent the values of the original index and the red markers represent the simplified indices.

The table I shows the total processing time (in seconds) consumed by original and simplified versions of the four validity indexes used to evaluate all the partitions that were generated for the data sets of test images 1 and 2. The table I also shows how much each simplified index is faster than the corresponding original index (speedup). It is possible to observe that the reduction of the computational cost is very significant, mainly for the data set of the test image 2 which presents a well bigger data volume.

TABLE I  
PROCESSING TIME OF THE INDEXES AND SPEEDUP

Indexes	Data set of test image 1		Data set of test image 2	
	time	speedup	time	speedup
$V_D$	591	2674	18868	9434
$V_D^{SOM}$	0.221		2	
$V_{DB}$	8	133	83	576
$V_{DB}^{SOM}$	0.06		0.144	
$V_{PBM}$	7	65	64	304
$V_{PBM}^{SOM}$	0.107		0.21	
$V_{CDbw}$	954	1191	32158	4020
$V_{CDbw}^{SOM}$	0.801		8	

## VI. CONCLUSIONS AND FINAL REMARKS

This work presents a strategy for clusters analysis using SOM networks segmented by mathematical morphology and simplified cluster validity indexes. The strategy can be seen as an improvement of the clustering method proposed in [8]. The incorporation of a cluster validity index in the data partitioning process makes possible to analyze the quality of clusters that are generated by different segmentations of a trained SOM. The parallel evaluation of different partitions for a data set using a cluster validity index allows determining more efficiently the best partition without the necessity to train a quantity of SOM networks as required in [8]. Using some cluster validity index appropriate for the data set to be partitioned, the strategy proposed here automatically finds the number of clusters as well as the partitioning of a data set without any type of interaction with the user. The application of mathematical morphology techniques to segment the U-matrix and the representation of data clusters by means of multiple prototypes supply to the proposed method the advantage to discover clusters with varied and complex shapes.

With the objective to reduce the computational cost of the proposed method, this work also proposes the simplification of cluster validity indexes using only the weight vectors and the frequency of *hits* of the neurons of the SOM to represent the clusters of the original data set. The experiments indicated that this approach is computationally effective, mainly for large data sets. Moreover, although this has not been studied here, the use of prototypes reduces the effect of outliers in data sets, therefore, the simplification of cluster validity indexes sensitive to noisy points (as the Dunn's index) can provide validation results better than the original versions of these indexes.

The proposed methodology has been applied to remotely sensed images and the results are satisfactory. However, some questions that were not treated in this work need to be analyzed in future works. All the analysis presented here suppose that the training of the SOM has been well accomplished, but how to obtain a trained SOM with minimum topographical distortions and suitable space quantization is still an open question and can be investigated. Although the CD<sub>bw</sub> index has presented the best performance for the remotely sensed data considered in our experiments, criteria to decide which cluster validity index is more appropriate to analyze remotely sensed images clustering

also must be studied. To confirm the success of the application of the proposed method in the analysis of remotely sensed images comparisons still must be made with clustering methods conventionally used for this type of application.

## REFERENCES

- [1] J. A. Richards, "Analysis of remotely sensed data: the formative decades and the future," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, pp. 422–432, March 2005.
- [2] Z. Sun, D. Huang, Y. Cheung, J. Liu, and G. Huang, "Using FCMC, FVS, and PCA techniques for feature extraction of Multispectral images," *IEEE Geosc. and Rem. Sens. Lett.*, vol. 2, no. 2, pp. 108–112, April 2005.
- [3] R. Xu and D. Wunsch, "Survey of clustering algorithms," *IEEE Trans. on Neural Networks*, vol. 16, no. 3, pp. 645–678, May 2005.
- [4] T. N. Tran, R. Wehrens, and L. M. C. Buydes, "Clustering multispectral images: a tutorial," *Chemometrics and Intelligent Laboratory Systems*, vol. 77, pp. 3–17, 2005.
- [5] G. H. Ball and D. J. Hall, "ISODATA, a novel method of data analysis and pattern classification," Stanford Research Institute, Menlo Park, CA, NTIS Report AD 699616, 1965.
- [6] S. Haykin, *Neural Network: A Comprehensive Foundation*. New York: Prentice-Hall, 2nd edition, 1999.
- [7] J. A. F. Costa and M. L. A. Netto, "Estimating the number of clusters in multivariate data by self-organizing maps," *International Journal of Neural Systems*, vol. 9, pp. 195–202, 1999.
- [8] J. A. F. Costa and M. L. A. Netto, "Clustering of complex shaped data sets via kohonen maps and mathematical morphology," in *Proceedings of the SPIE, Data Mining and Knowledge Discovery*. B. Dasarathy (Ed.), 2001, vol. 4384, pp. 16–27.
- [9] M. L. Gonçalves, M. L. A. Netto and J. Zullo, "A neural architecture for the classification of remote sensing imagery with advanced learning algorithms," in *Proc. of the IEEE Signal Processing Society Workshop*, 1998, pp. 577–585.
- [10] M. L. Gonçalves, M. L. A. Netto, J. A. F. Costa and J. Zullo, "Automatic remotely sensed data Clustering by tree-structured self-organizing," in *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, IGARSS'05*, Korea, 2005.
- [11] T. Kohonen, *Self-Organizing Maps*, 2nd Edition, Berlin: Springer Verlag, 1997.
- [12] A. Ultsch, "Self-organizing neural networks for visualization and classification," in *O. Opitz et al. (Eds). Information and Classification*, Berlin, Springer-Verlag, pp. 307–313, 1993.
- [13] J. Lampinen and E. Oja, "Clustering properties of hierarchical self-organizing maps," *J. of Math. Im. and Vis.*, vol. 2, pp. 261–272, 1992.
- [14] F. Murtagh, "Interpreting the kohonen self-organizing feature map using contiguity-constrained clustering," *Pattern Recognition Letters*, vol. 16, pp. 399–408, 1995.
- [15] M. Y. Kiang, "Extending the kohonen self-organizing map networks for clustering analysis," *Computational Statistics & Data Analysis*, vol. 38, pp. 161–180, 2001.
- [16] J. Vesanto and E. Alhoniemi, "Clustering of the self-organizing map," *IEEE Trans. on Neural Networks*, vol. 11, pp. 586–600, May 2000.
- [17] S. Wu and T. W. S. Chow, "Clustering of the self-organizing map using a clustering validity index based on inter-cluster and intra-cluster density," *Pattern Recognition*, vol. 37, pp. 175–188, 2004.
- [18] A. Bleau and L. J. Leon, "Watershed-based segmentation and region merging," *Comp. Vis. Image Underst.*, vol. 77, pp. 317–370, 2000.
- [19] J. C. Bezdek and N. R. Pal, "Some new indexes of cluster validity," *IEEE Trans. on Syst. Man and Cybern.*, vol. 28, pp. 301–315, 1998.
- [20] D. Davies and D. Bouldin, "A cluster separation measure," *IEEE Trans. Patt. Rec. and Mach. Intell.*, vol. PAMI-1, pp. 224–227, 1979.
- [21] M. K. Pakhira, S. Bandyopadhyay, and U. Maulik, "Validity index for crisp and fuzzy clusters," *Pattern Recog.*, vol. 37, pp. 487–501, 2003.
- [22] M. Halkidi and M. Vazirgiannis, "Clustering validity assessment using multi representatives", in *Proceedings of SETN Conference*, Thessaloniki, Greece, 2002.
- [23] M. Ji, "Using fuzzy sets to improve cluster labeling in unsupervised classification," *Int. J. of Remote Sensing*, vol. 24, pp. 657–671, 2003.



# Automatic Remotely Sensed Data Clustering by Tree-Structured Self-Organizing Maps

Márcio L. Gonçalves<sup>†</sup>, Márcio L. de Andrade Netto<sup>‡</sup>, José A. Ferreira Costa<sup>§</sup>, Jurandir Zullo Jr.<sup>#</sup>

<sup>†</sup> PUC Minas - 37701-355 - Poços de Caldas – MG – Brazil – E-mail: marcio@pucpcaldas.br

<sup>‡</sup> School of Electrical and Computer Engineering - State University of Campinas – SP – Brazil

<sup>§</sup> Department of Electrical Engineering – Federal University of Rio Grande do Norte – Natal – RN – Brazil

<sup>#</sup> Center of Meteorological and Climatic Research Applied to Agriculture - State University of Campinas – SP -Brazil

**Abstract**—This work presents a clusters analysis method which automatically finds the number of clusters as well as the partitioning of data set in a remotely sensed image without any type of assistance of an image analyst. The data clustering is made using the self-organizing (or Kohonen) map (SOM) and the techniques proposed by Costa & Netto [5] for automatic partition of trained SOM networks and for generating a hierarchy of maps based on the detected data clusters. The proposed clustering method has been applied on a LANDSAT/TM image and its performance was compared with that of K-means algorithm, conventionally used for remotely sensed images.

## I. INTRODUCTION

The automatic analysis of remotely sensed data has become an increasingly important topic over the last decades. Large volumes of remotely sensed multispectral data are being generated from an increasing number of sophisticated airborne and spaceborne sensor systems, and while there is no substitute for a trained analyst, exploitation of this data on a large scale requires the automated detection of specific features of interest [1][2].

Particularly the classification of pixels for partitioning different land-cover regions is an important problem in the realm of satellite imagery. Remotely sensed images usually contain a large quantity of classes with highly overlapping and nonlinear class boundaries. This interclass confusion introduces errors into the resulting spectral classification. Therefore, it is necessary the creation and development of efficient classification algorithms that make an appropriate modeling of spectral classes in remotely sensed images.

Cluster analysis is a form of pattern recognition often used in remote sensing applications to perform image decomposition. Clustering algorithms are designed to make the best possible use of the overall spectral content of an image, with the condition that the resulting classes remain as thematically uniform as possible. Several clustering methods have been developed for unsupervised classification of remotely sensed images [3]. However, in most of the cases, assumptions about the number of classes present in the image and about the geometric structure of the given data set are made. For example, partitioning methods such as *K*-means, and hierarchical techniques such as centroid and furthest neighbor are only feasible for clusters with hyper-spherical shapes, while others such as the nearest neighbor method are adequate only for elongated clusters. Some partitioning algorithms use

heuristics for splitting, merging or deleting clusters. Examples include ISODATA that rely on user supplied thresholds to perform these operations [3][7].

In this work we perform the unsupervised clustering of remote sensing data, which automatically finds the number of clusters as well as the partitioning of data in an image, and do not to assume any particular distribution of the data set or any other assumption about the type of structure present in data. For this purpose, the self-organizing (or Kohonen) map (SOM) [4] is used as a clustering tool together with the automatic segmentation method of SOM proposed by Costa & Netto [5], which enable a meaningful cluster's interpretation in the trained map. Although Kohonen maps had been applied for remote sensing data clustering, usually the researcher sets the number of neurons equal to the expected number of clusters, or manually segments a two-dimensional map using some a priori knowledge of the data with the assistance of an image analyst [6]. Applying the techniques proposed in [5], this work perform an automatic partitioning and labeling SOM networks in clusters of neurons that may be used to represent the data clusters in remotely sensed images, without any kind of interaction with the user. Mathematical morphology operations, such as watershed, are performed on the U-matrix, which is a neuron-distance image. The obtained results are labeled and connected regions of neurons that defines, in the input space, complex and non-parametric geometries which approximately describes the shape of the clusters. The process of map partitioning is extended recursively. Each cluster of neurons gives rise to a new map, which are trained with the subset of data were classified to it. The algorithm produces dynamically a hierarchical tree of maps, which explains the cluster's structure in levels of granularity. The distributed and multiple prototypes cluster representation enables the discoveries of clusters even in the case when we have images with highly overlapping spectral classes.

## II. THE SELF-ORGANIZING MAP

The SOM is one of the most popular artificial neural network algorithms. SOM is based on unsupervised competitive learning, which means that the training is entirely data-driven and that the neurons of the map compete with each other. It has been widely studied as a software tool for visualization of high-dimensional data. The various benefits the SOM offers include an approximation of the probability density function of the training data while preserves the

topological structure of them, nearby patterns in the sample data set are associated with nearby neurons in the SOM grid [4].

The SOM consists essentially of two layers of neurons: input-layer I and the unit layer U. Each neuron  $i$  is represented by an  $n$ -dimensional weight vector  $w_i = [w_{i1}, w_{i2}, \dots, w_{in}]^T$ , where  $n$  is equal to the dimension of the input vectors. The SOM defines a mapping from the high dimensional input data space onto a regular, usually, two-dimensional array of nodes. The neurons of the map are connected to adjacent neurons by a neighborhood relation dictating the structured of the map. In the 2-dimensional case the neurons of the map can be arranged either on a rectangular or hexagonal lattice.

The basic SOM algorithm is iterative. At each training step  $t$ , a sample data vector  $x(t)$  is randomly chosen from the training set. Distances between  $x(t)$  and all the weight vectors  $w_i$  are computed. The winner neuron, denoted by  $c$ , is the neuron with the weight vector closest to  $x(t)$ :

$$c = \arg \min_i \{ \|x(t) - w_i\| \} \quad (1)$$

where  $\|\cdot\|$  is the distance measure. The weight vectors of winner neuron as well as the neighboring nodes are moved closer to the input data vector. The SOM update rule for the weight vector of the unit  $i$  is

$$w_i(t+1) = w_i(t) + h_c(t)[x(t) - w_i(t)] \quad (2)$$

where  $h_c(t)$  is the neighborhood kernel around the winner unit  $c$  at step  $t$ .

#### A. The U-matrix Method

The Unified Distance Matrix (U-matrix) method was developed by A. Ultsch [8] to detect non-linearities in the resulting SOM mapping. The basic idea is to use the same metric that was used during the learning to compute distances between adjacent reference vectors. This method can be used to visualize the topological structure of the SOM unit layer and therefore also the topology of the  $N$ -dimensional input space.

The U-matrix has dimension  $(2X-1) \times (2Y-1)$  considering a two-dimensional SOM with a rectangular lattice topology of size  $(X \times Y)$ . The U-matrix is constructed on top of the map. Let  $n$  be a neuron on the map,  $NN(n)$  be the set of immediate neighbors on the map,  $w(n)$  the weight vector associated with neuron  $n$ , then

$$U\_height(n) = \sum_{m \in NN(n)} d(w(n) - w(m)) \quad (3)$$

where  $d(\cdot)$  is a distance measure that make use of the metric used in the SOM algorithm to construct the map. The U-matrix is a display of the  $U\_heights$  on top the grid positions of the neurons on the map and can be visualized as a three dimensional landscape. Altitudes or the high places on the U-matrix encode dissimilarities between neurons and correspond to cluster borders while valleys represent to map units that are similar.

### III. PROPOSED METHODOLOGY

The clusters analysis method of remotely sensed data proposed in this work is described as follow.

Initially a training patterns set is randomly and uniformly obtained from all over the image and presented as input vectors to the SOM. These training patterns can be isolated pixels or pixel windows to enable incorporating of neighborhood attributes. The SOM is then trained and after its reference vectors are segmented using the automatic partitioning method of SOM networks proposed by Costa & Netto [5].

The Costa & Netto algorithm applies mathematical morphology operations, such as watershed, to segment the U-matrix. Markers are used to identify significant clusters and are automatically found by performing a multi-level scan of connected regions of the U-matrix. Labels are assigned for each connected region (cluster of neurons) of the U-matrix and are copied to the corresponding neurons in the map.

The process of map partitioning is extended recursively. Each cluster of neuron gives rise to a new map, which are trained with the subset of data that were classified to it. The algorithm produces dynamically a hierarchical tree of maps, which explains the cluster's structure in levels of granularity. It is not necessary to specify the number of sub-maps for a given map in a given height of the tree. The parameters of the child map are functions of the parameters of the parent map and of the subgroup of data used to train that map. A pruning strategy cuts sub-maps (leave nodes) which do not present further partitions [5].

The Costa & Netto algorithm automatically produces labeled sets of neurons that are related to the clusters in the input space, which corresponds to the spectral classes present in the image. The labeled sets of neurons (prototypes) are then used jointly with the nearest neighbor rule to label all the image pixels. During all the process the image analyst has only the task to associate the labeled clusters with information classes (or ground cover types).

In the next subsections are presented the steps of automatic partitioning algorithm of SOM networks and the approach used to find watershed markers for the U-matrix. More details about these algorithms and the generation of hierarchy of labeled SOM networks can be directly obtained in [5].

#### A. Automatic Partitioning and Labeling of trained SOM

The general approach for automatic partitioning and labeling of a trained SOM using watershed transform could be summarized as follow [5]:

1. Obtain the U-matrix using the trained map;
2. Find the image markers for the U-matrix (see next subsection);
3. Compute the watershed lines on the U-matrix using the markers obtained in the step 2;
4. Assign a different label for each connected region (cluster of neurons) of the U-matrix;
5. Copy the U-matrix labels to the corresponding neurons in the map.

#### B. Finding Watershed Markers for U-matrix

Given the U-matrix image  $U$  with 256 gray levels, the following steps summarize the simple but efficient approach proposed in [5] to find good watershed markers for the U-matrix:

1. Filtering: create the image  $U_l$  by removing any pore with area less or equal than three pixels;

2. For  $k=1$  to  $f_{max}$ , where  $f_{max}$  is the highest gray level of the image, create the binary images  $U_j^k$  that corresponds to the conversions  $U_j$  to a binary image using as threshold  $k$ .
3. Compute  $N_{rc}^k$ , the number of connected regions of  $U_j^k$ , for each value  $k$ ;
4. Obtain the most persistent value of number of regions  $\eta$  (clusters of neurons) that corresponds to the highest (and significant) plateau in the plot of  $N_{rc}^k$  versus  $k$ .
5. Take as image marker the binary image  $U_j^v$ , where  $v$  is the value  $k$  of the plateau chosen in the step 4.

#### IV. EXPERIMENTAL RESULTS

A TM/LANDSAT image composed by spectral bands 3, 4 and 5 was used to test the efficiency of the proposed clusters analysis method. The image has 512x492 pixels, and it shows the Manaus city and the crossing of rivers Negro and Solimões (Amazonas state, Brazil) [9]. The Fig. 1a shows a colour composite image from the three spectral bands. The scene presents 4 predominant land cover classes: water 1 (class 1), water 2 (class 2), vegetation (class 3) and urban area (class 4).

The training patterns set was composed by 5000 pixels randomly and uniformly obtained from all over the image and presented as input data to the root SOM. The Fig. 1b shows the training patterns in the three-dimensional space where it is possible to visualize four “clouds of points” (clusters) that corresponds to the four land cover classes present in the image. Also it is possible to observe that the clusters have different shapes and they do not have well defined boundaries. These factors increase the difficulty of the partitioning process of the classes.

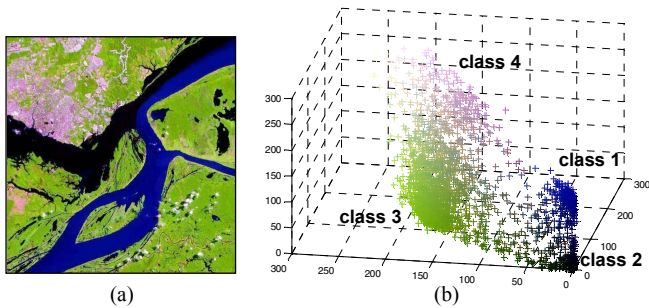


Figure 1. (a) TM/Landsat image; (b) Training patterns in the 3D space.

The size of root SOM network was set 15x10. Weight initialization was linear and the training was done with the batch algorithm [4]. The neighborhood function was Gaussian and during ordering the neighbor radius decreased with the training epochs. The number of training epochs was set 500.

Fig. 2a shows a 3D view of the U-matrix for the trained root map. Following the proposed method, the number of clusters for U-matrix and the corresponding image marker are determined automatically. It uses information like the number of connected regions ( $N_{rc}^k$ ) for each gray level of the image  $U_j^k$  (Fig. 2b), which in turn is related to the distances between neighboring neurons. The image markers for this map are presented in Fig. 3a. Markers act as seeds for the watershed

transform and are obtained after thresholding the U-matrix by lowest gray level  $k$  from largest contiguity range of  $N_{rc}^k$ , which in this case it was the value 37 that corresponds to 3 connected regions. The segmented U-matrix is illustrated in Fig. 3b while the corresponding labeled SOM is shown in Fig. 3c.

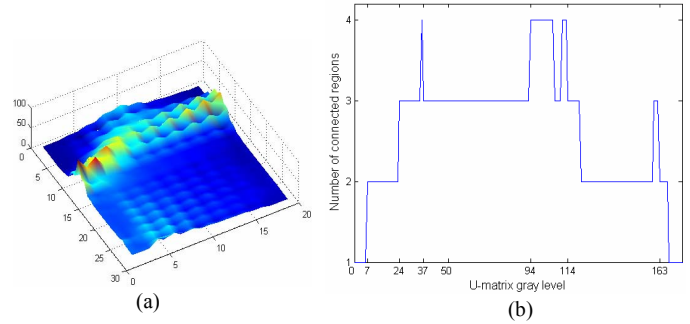


Figure 2. (a) 3D view of the U-matrix for trained root map; (b) Number of connected regions versus image threshold.

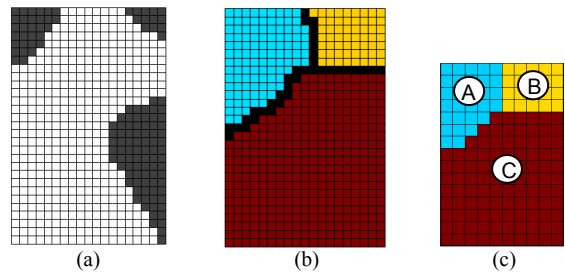


Figure 3. (a) Image markers; (b) Segmented U-matrix after watershed transform; (c) Labeled SOM after U-matrix segmentation.

Note that although a 4-cluster solution would be the optimal, the algorithm detected 3 clusters in the root map. The clusters of neurons “A” e “B” correspond to classes 1 and 2, respectively, while the cluster of neurons “C” correspond to the classes 3 and 4 joined. The classes 3 and 4 present a more accentuated overlapping of points hindering the detection of the two corresponding clusters in immediate way.

Following the hierarchical SOM strategy, only the cluster of neurons “C” generates a new map, which was trained with the subset of data that were classified to it. The sub-maps derived of the clusters of neurons “A” and “B” were pruned because they did not generated new clusters of neurons. The new sub-network generated from cluster “C” was configured with 11x9 neurons and trained of the same manner as that of root map. A 3D view of the U-matrix for trained sub-network is illustrated in Fig. 4 where the image markers and labeled sub-network are also shown. Note that two news clusters of neurons were detected. The cluster “C1” correspond to class 3 and cluster “C2” correspond to class 4. The process was applied again to clusters of neurons “C1” and “C2”, but the derived sub-maps were pruned because they did not generated new clusters of neurons.

Finalized the clusters analysis by proposed methodology, the reference vectors (prototypes) of each cluster of neurons found were used jointly with the nearest neighbor rule to label the training samples set and also all image pixels (Fig. 5). Although the spectral classes present geometries that are not well defined, the method proposed in this work detected and

partitioned correctly the corresponding clusters of mode totally unsupervised, without knowledge of the desired number of classes and without any type of assistance of the user.

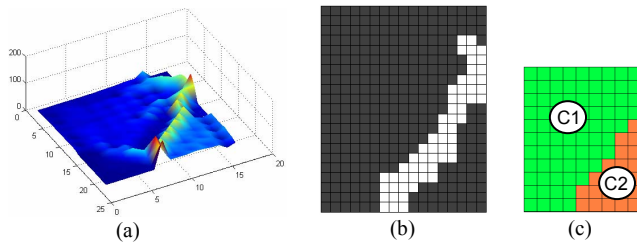


Figure 4. (a) 3D view of the U-matrix for sub-map generated from cluster of neurons “C”; (b) Image markers; (c) Labeled sub-map.

To validate the efficiency of the proposed clusters analysis method the K-means clustering algorithm (conventionally used for remotely sensed images) was also applied on the test image. For this method the number of clusters has to be set in advance. Considering that the 4-cluster solution would be the optimal for the test image, the K-means algorithm was executed with this desired number of clusters. The results of this experiment show that the K-means algorithm did not detected the classes 1 and 2 separately. Both the classes had belonged to the same cluster (Fig. 6). The classes 1 and 2 only are detected by K-means algorithm when the desired number of clusters is defined as more than 4, but in these cases the algorithm determine sub-classes that are not considered important by analysis of a image expert. It is important to mention also that this process of analysis requires an intense interaction of the expert so that correct interpretations of the results can be made.

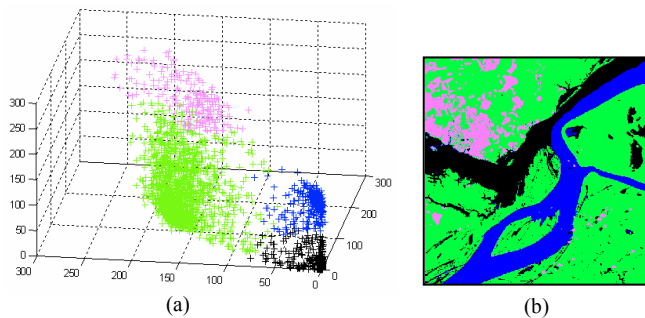


Figure 5. (a) Training samples labeled by proposed method; (b) Image labeled by proposed method (each color represents a different class).

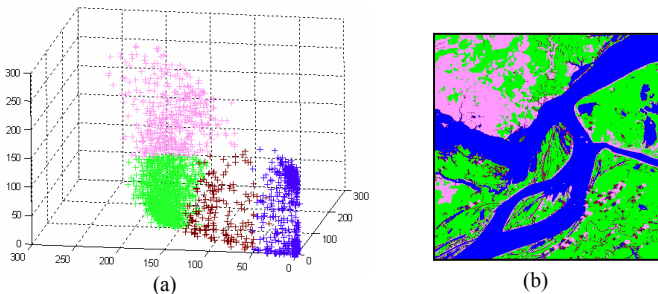


Figure 6. (a) Training samples labeled by K-means; (b) Image labeled by K-means (each color represents a different class).

## V. CONCLUSIONS

This work presents a cluster analysis method for remotely sensed images which automatically finds the number of clusters as well as the partitioning of image data set. The method proposed uses the SOM as a clustering tool together with the automatic segmentation algorithms of SOM proposed by Costa & Netto [5]. Although Kohonen maps had been applied for remote sensing data clustering, usually an image analyst manually segments a two-dimensional map using some a priori knowledge of the data. The method proposed here is entirely data-driven and it does not employ any type of assistance of an image analyst or the user to find and to partition the clusters of neurons of SOM and consequently the data clusters. The algorithm produces dynamically a hierarchical tree of maps, which explains the cluster’s structure in levels of granularity. Differently of the traditional clustering methods, the representation of the clusters structures is given by a set of neurons. Not only the hierarchy of maps can detect cluster and sub-clusters but also the distributed and multiple cluster representation enables the discoveries of clusters even in the case when we have images with highly overlapping spectral classes and data clusters with complex and arbitrary geometries. In the traditional clustering algorithms, such as K-means, the number of clusters has to be set in advance, and if a single prototype is used for representing the cluster structure, only hyperspherical shaped clusters are detected. This lead to errors in recovering the inherent structure of the clusters.

The proposed method was applied on others LANDSAT/TM and CBERS images. The comparison of its performance with those of other methods used conventionally for clustering of remote sensing data, such as K-means and ISODATA, allowed to confirm its efficiency.

## REFERENCES

- [1] L. Soh, C. Tsatsoulis, “Segmentation of satellite imagery of natural scenes using data mining,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 2, pp. 1086-1099, March 1999.
- [2] N. R. Harvey, J. Theiler, S. P. Brumby, S. Perkins, J. J. Szymanski, J. J. Bloch, R. B. porter, M. Galassi, A. C. Young, “Comparison of GENIE and conventional supervised classifiers for multispectral image feature extraction,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 2, pp. 393-404, February 2002.
- [3] H. G. Wilson, B. Boots, A.A. Millward, “A comparison of hierarchical and partitional clustering techniques for multispectral image classification,” in *Proceedings of the International Geoscience and Remote Sensing Symposium*, pp. 1624-1626, Toronto, Canada, 2002.
- [4] T. Kohonen, *Self-Organizing Maps*, 2nd Edition, Berlin: Springer Verlag, 1997.
- [5] J. A. F. Costa, M. L. A. Netto, “Clustering of complex shaped data sets via Kohonen maps and mathematical morphology,” in *Proceedings of the SPIE, Data Mining and Knowledge Discovery*. B. Dasarthy (Ed.), vol. 4384, pp. 16-27, 2001.
- [6] M. L. Gonçalves, M. L. A. Netto, J. Zullo Jr., “A neural architecture for the classification of remote sensing imagery with advanced learning algorithms,” in *Proc. of the IEEE Signal Processing Society Workshop - VIII Neural Networks for Signal Processing*, pp. 577-585, 1998.
- [7] T. Y. Young, King-Su Fu, *Handbook of Pattern Recognition and Image Processing*, London, Academic Press Inc., 1986.
- [8] A. Ultsch, “Self-organizing neural networks for visualization and classification,” in *Information and Classification*, Berlin, Springer-Verlag, pp. 307-313, 1993.
- [9] <http://www.dgi.inpe.br/html/imagens.htm>, Image Generation Division – National Institute for Space Research, Brazil.

# Explorando as Propriedades do Mapa Auto-organizável de Kohonen na Classificação de Imagens de Satélite

Márcio L. Gonçalves<sup>1,2</sup>, Márcio L. de Andrade Netto<sup>2</sup>, José Alfredo F. Costa<sup>3</sup>

<sup>1</sup>Curso de Ciência da Computação - PUC Minas *campus* de Poços de Caldas  
Av. Padre Francis Cletus Cox, 1661 – Cep: 37701-355 – Poços de Caldas, MG

<sup>2</sup>Faculdade de Engenharia Elétrica e de Computação  
Universidade Estadual de Campinas (UNICAMP) – Campinas, SP

<sup>3</sup>Departamento de Engenharia Elétrica  
Universidade Federal do Rio Grande do Norte (UFRN) – Natal, RN

marcio@pucpcaldas.br, marcio@dca.fee.unicamp.br, alfredo@dee.ufrn.br

**Abstract.** *This work presents a methodology for the land-cover classification of satellite images which explores the self-organizing map (SOM) properties. The proposed method performs a cluster analysis of the image data using a three-stage approach. Firstly, the original patterns of the image are clustered using the SOM, after that the SOM is filtered, and then segmented through a hierarchical clustering method which uses the neighborhood relation of the neurons and incorporates spatial information in the merging criterion. The experimental results show an application example of the proposed methodology on an IKONOS image.*

**Resumo.** *Este trabalho apresenta uma metodologia que explora as propriedades do mapa auto-organizável (SOM) para realizar a classificação de padrões de cobertura terrestre em imagens de satélite. O método proposto efetua uma análise de agrupamentos dos dados da imagem empregando uma abordagem em três estágios. Primeiramente, os padrões originais da imagem são agrupados usando o SOM, em seguida o SOM é filtrado, e então segmentado através de um método de agrupamento hierárquico que utiliza as relações de vizinhança dos neurônios e incorpora informações espaciais em seu critério de fusão. Os resultados experimentais mostram um exemplo de aplicação da metodologia proposta sobre uma imagem IKONOS.*

## 1. Introdução

O mapa auto-organizável (SOM - *self-organizing map*), proposto por Kohonen (1997), tem sido frequentemente utilizado em inúmeras aplicações, incluindo áreas como reconhecimento de padrões, processamento de sinais, compressão de dados, modelagem biológica e mineração de dados. O SOM é um tipo de rede neural artificial baseada em aprendizado competitivo e não-supervisionado. Propriedades importantes como a capacidade de aproximar o espaço de entrada, ordenação topológica e casamento de densidade, aliadas a simplicidade do modelo e a facilidade de implementação do seu algoritmo de aprendizagem justificam o sucesso do SOM e o colocam como um dos principais modelos de redes neurais na atualidade.



Este trabalho apresenta uma metodologia que explora as características e propriedades do SOM para realizar a classificação não-supervisionada de padrões de cobertura terrestre em imagens de sensoriamento remoto. Desde o lançamento dos primeiros satélites voltados para a exploração de recursos terrestres, os métodos digitais de classificação de imagens de sensoriamento remoto têm adquirido uma importância crescente no reconhecimento automático de padrões da superfície terrestre (Richards, 2005). Atualmente a enorme quantidade de imagens que estão sendo coletadas por sistemas sensores cada vez mais modernos e sofisticados requerem o desenvolvimento de metodologias de classificação inovadoras, as quais possibilitem uma exploração automática e eficiente do grande volume de informações disponíveis nas imagens e, ao mesmo tempo, tornem o processo de mapeamento de características da superfície terrestre menos subjetivo e com maior potencial de repetição em situações subseqüentes.

O método de classificação proposto aqui realiza uma análise de agrupamentos dos dados da imagem empregando uma abordagem constituída de três estágios de processamento através do SOM. Primeiramente, o SOM é utilizado para agrupar os padrões originais da imagem em um conjunto reduzido de protótipos (neurônios) dispostos em um arranjo retangular bidimensional. Em seguida, o SOM passa por um processo de filtragem onde protótipos que representam regiões de transição entre diferentes classes de cobertura terrestre são descartados do próximo estágio de processamento e aqueles que apresentam atividade nula são eliminados. No terceiro e último estágio da análise, um método de agrupamento hierárquico que respeita as relações de vizinhança dos neurônios e que incorpora informações espaciais em seu critério de fusão é utilizado para segmentar o SOM gerando um dendrograma de seus protótipos. Cada nível do dendrograma obtido corresponde a uma configuração diferente de agrupamentos de neurônios do SOM que pode ser utilizada para representar as classes sobre as quais a imagem original será classificada.

## 2. O SOM

O SOM é uma rede neural artificial que define um mapeamento de um espaço de entrada  $p$ -dimensional contínuo para um conjunto finito de protótipos (neurônios), que são dispostos em um arranjo topológico, geralmente bidimensional (Kohonen, 1997).

A rede consiste essencialmente de duas camadas de neurônios: uma camada de entrada  $I$  e uma camada de saída  $U$  (camada de Kohonen). A entrada da rede corresponde a um vetor  $p$ -dimensional,  $\mathbf{x}$ , geralmente no espaço  $\mathcal{R}^p$ . Todas as  $p$  componentes do vetor de entrada alimentam cada um dos neurônios do mapa. Cada neurônio  $i$  pode ser representado então por um vetor de pesos  $\mathbf{w}_i = [w_{i1}, w_{i2}, \dots, w_{ip}]^T$ , também no espaço  $p$ -dimensional.

Para cada padrão de entrada um neurônio é escolhido o vencedor,  $c$ , usando o critério de maior similaridade:

$$\|\mathbf{x} - \mathbf{w}_c\| = \min_i \{\|\mathbf{x} - \mathbf{w}_i\|\}, \quad (1)$$

onde  $\|\cdot\|$  representa a distância Euclidiana. Os pesos do neurônio vencedor, juntamente com os pesos dos seus neurônios vizinhos, são ajustados de acordo com a seguinte equação:

$$\mathbf{w}_i(t+1) = \mathbf{w}_i(t) + h_c(t)[\mathbf{x}(t) - \mathbf{w}_i(t)], \quad (2)$$

onde  $t$  indica a iteração do processo de treinamento,  $\mathbf{x}(t)$  é o padrão de entrada e  $h_c(t)$  é o núcleo de vizinhança em torno do neurônio vencedor.

Uma vez garantida a convergência do seu algoritmo de aprendizagem, o arranjo de neurônios do SOM reflete características estatísticas importantes do espaço de entrada. As principais propriedades da rede podem ser resumidas como seguem (Haykin, 2001):

1) *Aproximação do espaço de entrada*: o SOM tem como objetivo básico armazenar um conjunto grande de vetores de entrada encontrando um conjunto menor de protótipos (vetores de pesos sinápticos  $\mathbf{w}_i$ ) de modo a fornecer uma boa aproximação para o espaço de entrada original. A base teórica dessa estratégia está fundamentada na teoria da quantização vetorial, cuja motivação é a redução de dimensionalidade ou a compressão de dados.

2) *Ordenação Topológica*: ao realizar o mapeamento não-linear dos vetores de entrada para o arranjo de neurônios da rede, o algoritmo do SOM tenta preservar ao máximo a topologia do espaço original, ou seja, procura fazer com que neurônios vizinhos no espaço de saída apresentem vetores de pesos que representem padrões vizinhos no espaço de entrada.

3) *Casamento de Densidade*: o mapeamento efetuado pelo SOM reflete a distribuição de probabilidade dos dados no espaço de entrada original. Regiões do espaço de entrada de onde os vetores de amostra  $\mathbf{x}$  são retirados com uma alta probabilidade de ocorrência são mapeadas para domínios maiores no espaço de saída, e portanto com melhor resolução que regiões no espaço de entrada de onde vetores de amostra  $\mathbf{x}$  são retirados com uma baixa probabilidade de ocorrência.

### 3. Classificação Não-supervisionada de Imagens de Sensoriamento Remoto

O processo de classificação não-supervisionada de imagens de sensoriamento remoto baseia-se no princípio de que o algoritmo computacional é capaz de identificar por si só as classes dentro do conjunto de dados da imagem. Esse tipo de classificação é freqüentemente realizado através de métodos de agrupamentos (*clustering*). Embora exista uma grande quantidade de diferentes métodos de agrupamentos na área de reconhecimento de padrões (Xu and Wunsch II, 2005), a maioria dos softwares ou sistemas computacionais voltados para o processamento digital de imagens de sensoriamento remoto realiza a classificação não-supervisionada baseada em métodos de agrupamentos particionais, como o K-médias e o ISODATA (Ball and Hall, 1967).

Apesar de serem amplamente utilizados esses métodos particionais de agrupamentos apresentam diversas limitações. As funções objetivo usadas por eles partem do pressuposto de que o número classes é conhecido a priori. O usuário também deve especificar manualmente vários parâmetros para controlar o processo de agrupamento, entre eles: os centróides iniciais de cada agrupamento, o número máximo de iterações, limiares para realizar a divisão, fusão ou exclusão de agrupamentos. O K-médias e o ISODATA são muito sensíveis a esses parâmetros, podendo gerar partições diferentes quando são feitas várias simulações para um mesmo conjunto de dados. Essas exigências certamente aumentam o nível de interação do usuário com o algoritmo computacional, aumentando conseqüentemente o grau de subjetividade do processo de

categorização da imagem. Outras limitações não menos importantes de algoritmos particionais são: o alto custo computacional quando o conjunto de dados a ser analisado é muito grande (a cada iteração todos os pixels da imagem são comparados com todos os centros de agrupamentos) e a existência de suposições sobre a forma dos agrupamentos. Geralmente apenas um protótipo (centróide) é utilizado para representar um agrupamento, portanto esses métodos tornam-se adequados apenas para analisar agrupamentos que apresentam formatos hiperesféricos.

Uma outra forma possível, mas não usual, de realizar a classificação não-supervisionada de imagens de sensoriamento remoto é através de métodos hierárquicos de agrupamentos. Diferentemente dos métodos particionais, os métodos hierárquicos não exigem que o usuário especifique previamente o número de classes e outros parâmetros adicionais. Outra vantagem significativa desses métodos é a de possibilitar a visualização do resultado da classificação por meio de um dendrograma, que ilustra de forma hierárquica o grau de semelhança entre os agrupamentos que são formados por fusões (ou divisões) em cada estágio sucessivo da análise. Entretanto, os métodos hierárquicos apresentam algumas características que inviabilizam a sua aplicação na classificação de imagens de sensoriamento remoto: (a) em geral requerem espaço de memória de ordem  $O(N^2)$ , onde  $N$  é o número de registros do conjunto de dados; (b) os resultados podem ser difíceis de interpretar, principalmente para conjuntos de dados grandes; (c) para determinar qual a linha de “corte” do dendrograma (qual o número ideal de agrupamentos) há a necessidade de se aplicar algum critério de decisão.

## 4. Metodologia Proposta

O método não-supervisionado de classificação de imagens de satélite proposto aqui tem como princípio básico realizar a análise de agrupamentos da imagem através de um conjunto de protótipos do SOM ao invés de trabalhar diretamente com os padrões originais da cena. Para tanto, uma abordagem dividida em três estágios de processamento através do SOM é utilizada com o objetivo de descobrir agrupamentos de protótipos representativos para cada uma das classes de interesse da imagem. Os três estágios de processamento, que consistem basicamente no treinamento, filtragem e segmentação do SOM, são descritos nas subseções seguintes.

### 4.1 Treinamento

No primeiro estágio de processamento um conjunto de amostras coletadas a partir da imagem original é utilizado para realizar o treinamento do SOM. Diferentemente de abordagens pixel-a-pixel que utilizam apenas a informação espectral de pontos individuais para encontrar regiões homogêneas, o método proposto realiza a amostragem da imagem através de janelas de pixels. A idéia é incorporar no processo de classificação informações de vizinhança (contexto), tendo em vista que pixels isolados não são capazes de representar a maioria dos padrões de cobertura terrestre presentes nas imagens de satélite. As janelas amostrais são coletadas de maneira uniforme sobre toda a região da imagem, sem a intervenção do usuário, sem sobreposições e em intervalos regulares. Todas as amostras são quadradas e possuem o mesmo tamanho.

Para realizar o treinamento do SOM é necessário especificar alguns parâmetros que definem a estrutura do mapa e que controlam o treinamento propriamente dito. Com



o objetivo de garantir um bom mapeamento dos padrões originais, a metodologia proposta define de modo particular os parâmetros da rede baseando-se na literatura existente, em testes experimentais e em algumas peculiaridades da aplicação do SOM em imagens de sensores remotos. Todavia, outras alternativas também podem ser buscadas para se obter bons mapas. Uma boa revisão sobre os parâmetros de treinamento do SOM pode ser obtida em Costa (1999).

A metodologia proposta utiliza os seguintes parâmetros para realizar o treinamento da rede: inicialização linear dos pesos, apresentação dos padrões em lote, número de épocas de treinamento igual a 500 (quinhentos), tipo de função de vizinhança gaussiana e formato retangular para o arranjo bidimensional de neurônios.

O tamanho do mapa (número de neurônios) é um dos parâmetros livres do SOM que depende dos objetivos desejados para a classificação. Se o objetivo é detectar todos os padrões da imagem, inclusive aqueles com baixa probabilidade de ocorrência, mapas de tamanhos maiores devem ser empregados na análise, caso contrário, se o interesse se concentra apenas nos padrões mais predominantes da cena então um tamanho menor para o SOM pode ser utilizado.

No final da etapa de treinamento do SOM tem-se, portanto, um conjunto reduzido de protótipos (neurônios) dispostos em um arranjo retangular bidimensional, os quais representam todos os padrões de entrada coletados a partir da imagem original.

#### 4.2 Filtragem

O segundo estágio da abordagem proposta consiste em filtrar dois tipos de protótipos que geralmente aparecem no mapeamento dos padrões da imagem através do SOM. Esses protótipos, chamados aqui de *inativos* e *heterogêneos*, podem funcionar como bordas (ou “unidades de interpolação”) no arranjo retangular de neurônios do SOM contribuindo para a separação de agrupamentos.

Os protótipos inativos correspondem àqueles neurônios que apresentam atividade nula no processo de aprendizado competitivo do SOM, ou seja, não possuem nenhum padrão de entrada associado a eles. Esses protótipos são simplesmente eliminados da análise.

Os protótipos heterogêneos são aqueles que apresentam alto grau de heterogeneidade espectro-textural e estão associados normalmente a padrões de entrada que possuem mais do que uma classe de cobertura terrestre. Na maioria das vezes esses padrões correspondem a regiões de transição entre classes de cobertura terrestre presentes na imagem e são capturados em decorrência da amostragem realizada por meio de janelas de pixels. Os protótipos considerados heterogêneos são excluídos do estágio seguinte da análise. Essa exclusão é realizada com o objetivo de não permitir que esses protótipos e, conseqüentemente, os padrões da imagem associados a eles sejam atribuídos erroneamente a uma das classes que os compõe. Os pixels da imagem associados aos protótipos heterogêneos são classificados de maneira especial somente no final do processo usando as informações dos rótulos de seus pixels vizinhos.

O grau de heterogeneidade espectro-textural de cada protótipo do SOM é calculado através de uma medida estatística baseada na matriz de co-ocorrência de níveis de cinza proposta por Haralick et al. (1973). Considerando que os protótipos do

SOM possuem as mesmas dimensões que as amostras de dados utilizadas para treinar a rede, as quais nesta abordagem correspondem a janelas de pixels, conseqüentemente eles também podem ser vistos como janelas de pixels (pequenas imagens). Sendo assim, é possível medir a probabilidade de co-ocorrência de níveis de cinza  $i$  e  $j$  (separados por uma distância  $d$  numa direção  $\theta$ ) em cada um dos protótipos do SOM. A energia (ou segundo momento angular) foi a medida estatística aplicada aqui para medir a heterogeneidade dos protótipos com base na matriz de co-ocorrência de níveis de cinza de cada um deles. Esta medida, descrita pela equação (3), apresenta valores próximos de 1 quando a área de interesse apresenta textura uniforme (níveis de cinza próximos) e valores que tendem a zero caso a área não seja uniforme.

$$ENE = \sum_i \sum_j P(i, j)_{d, \theta}, \quad (3)$$

onde  $P(i, j)_{d, \theta}$  representa a freqüência relativa com que dois pixels vizinhos ( $i$  e  $j$ ), separados por uma distância  $d$  na direção  $\theta$ , ocorrem em um protótipo do SOM.

Os protótipos cujos valores de energia (ENE) satisfazem a relação dada abaixo são considerados heterogêneos e, conseqüentemente, são filtrados:

$$ENE > \mu_{ENE} + \frac{1}{2} \sigma_{ENE}. \quad (4)$$

$\mu_{ENE}$  e  $\sigma_{ENE}$  são, respectivamente, a média e o desvio padrão dos valores de energia de todos os protótipos do SOM.

### 4.3 Segmentação

No último estágio do processo de análise de agrupamentos da imagem um método de agrupamento hierárquico aglomerativo é aplicado para agrupar protótipos semelhantes no SOM já treinado e filtrado.

O método hierárquico utilizado aqui apresenta duas características importantes. A primeira delas é a imposição de restrições às possíveis fusões dos protótipos do SOM. Diferentemente da abordagem tradicional de métodos de agrupamentos hierárquicos, que consiste em comparar todos os pares de objetos ou grupos de objetos para decidir sobre uma fusão, a abordagem utilizada neste trabalho verifica a possibilidade de efetuar fusões somente entre pares de protótipos ou grupos de protótipos que sejam adjacentes (ou vizinhos) na grade do SOM. Essa estratégia já foi utilizada com sucesso em Costa (2005). Outra característica importante é que além de utilizar a informação de distância o critério de fusão empregado também se baseia em informações espaciais dos pixels da imagem associados aos protótipos do SOM.

A informação de distância ( $D_{ij}$ ) entre dois agrupamentos de protótipos  $i$  e  $j$  do SOM é calculada usando a métrica euclidiana e o método de vizinhos mais próximos (ou de ligações simples). Os valores de  $D_{ij}$  são normalizados para o intervalo  $[0,1]$ .

As informações espaciais são calculadas através de dois índices, denominados *índice de fronteira espacial* e *índice de compacidade espacial*. Esses índices, desenvolvidos por Marçal e Castro (2005), são calculados nesta abordagem através da imagem classificada com base nos agrupamentos (ou classes) de protótipos em cada

nível do dendrograma gerado pelo método hierárquico.

O índice de fronteira espacial ( $IFE_{ij}$ ) calcula o tamanho da fronteira entre todos os pares de classes (i,j) considerando oito vizinhos para cada pixel (quatro adjacentes e quatro oblíquos). A sua fórmula é dada como segue:

$$IFE_{ij} = 1 - \frac{1}{2} \left( \frac{f_{ij}}{\sum_{k=1(k \neq i)}^N f_{ik}} + \frac{f_{ij}}{\sum_{k=1(k \neq j)}^N f_{kj}} \right), \quad (5)$$

onde  $f_{ij}$  representa a quantidade de vezes que as classes i e j são vizinhas na imagem e N é o número de classes no nível do dendrograma que está sendo analisado. A utilização deste índice se justifica pela idéia de que duas classes que possuem uma fronteira espacial significativa devem estar mais propensas a uma fusão do que classes que possuem pouca ou nenhuma fronteira espacial na imagem.

O índice de compactidade espacial ( $ICE_{ij}$ ) se baseia na quantidade de vezes ( $f_{ii}$ ) que pixels de uma mesma classe i são vizinhos na imagem. Este índice que busca penalizar a fusão de classes compactas é dado da seguinte forma:

$$ICE_{ij} = \frac{1}{2} \left( \frac{f_{ii}}{f_{ii} + 6 \sum_{k=1(k \neq i)}^N f_{ik}} + \frac{f_{jj}}{f_{jj} + 6 \sum_{k=1(k \neq j)}^N f_{jk}} \right). \quad (6)$$

Da mesma forma que o índice  $IFE_{ij}$ , o  $ICE_{ij}$  resulta valores no intervalo [0,1].

O critério adotado aqui estabelece que o par de classes de protótipos (i,j) do SOM que apresentar o valor mais baixo resultante do cálculo da média entre  $D_{ij}$ ,  $IFE_{ij}$  e  $ICE_{ij}$  é aquele que deve ser escolhido para a fusão.

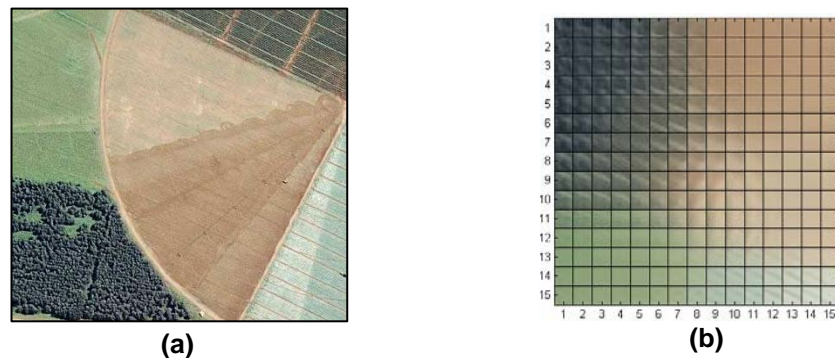
No final deste estágio tem-se então um dendrograma de protótipos do SOM. Cada nível do dendrograma apresenta uma configuração diferente de agrupamentos (ou classes) de protótipos que pode ser utilizada para realizar a classificação da imagem.

## 5. Resultados Experimentais

Uma imagem de 385×350 pixels composta por três bandas espectrais do satélite IKONOS (cedida pela Engesat/Brazil, © Space Imaging) foi utilizada para testar a metodologia de classificação proposta. A cena mostra parte de uma área com pivots de irrigação na região de Andaraí – BA e apresenta 06 classes de cobertura terrestre predominantes: vegetação rasteira, mata, dois padrões de solo nu e dois tipos de culturas agrícolas. A Figura 1(a) mostra uma composição colorida da imagem utilizada nos testes.

De acordo com os procedimentos descritos na subseção 4.1, 1292 janelas amostrais de tamanho 9×9 foram coletadas a partir da imagem teste e utilizadas para treinar um SOM composto por 225 neurônios dispostos em uma grade retangular 15×15. A Figura 1(b) ilustra a imagem de cada um dos protótipos do SOM dispostos na grade

após o término do treinamento. Através dela é possível visualizar a formação de agrupamentos que correspondem aos padrões de cobertura terrestre presentes na imagem e verificar as propriedades de ordenação topológica e de casamento de densidade do SOM. Classes de cobertura terrestre com atributos espectrais semelhantes foram mapeadas em regiões vizinhas no mapa e aquelas que ocupam áreas maiores na imagem também foram mapeadas para um número maior de protótipos na grade do SOM.



**Figura 1. (a) Composição colorida da imagem utilizada nos experimentos. (b) Imagem dos protótipos do SOM na grade retangular 15x15 após o treinamento. OBS: as imagens são apresentadas em escalas diferentes.**

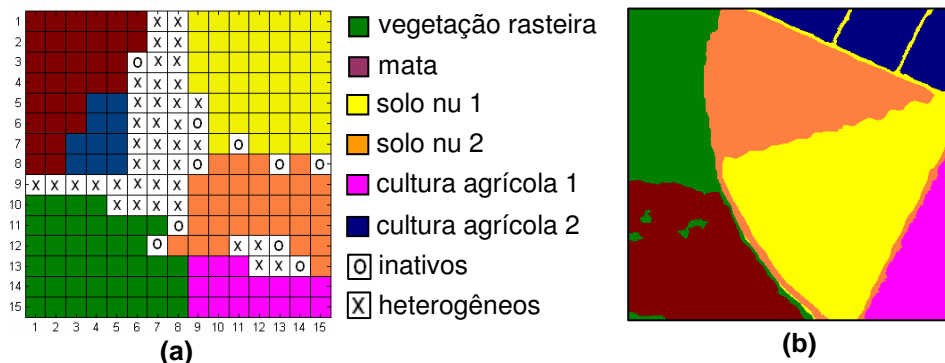
No estágio seguinte da análise, 49 protótipos do SOM foram filtrados. Dentre eles, 10 apresentaram atividade nula e 39 demonstraram um alto grau de heterogeneidade espectro-textural, uma vez que apresentaram valores de energia (ENE) acima do limiar apresentado na equação (4). As matrizes de co-ocorrência de níveis de cinza foram calculadas para cada protótipo nas quatro direções ( $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  e  $135^\circ$ ) e a uma distância de um pixel ( $d = 1$ ).

Após o estágio de filtragem, o método de agrupamento hierárquico (descrito na subseção 4.3) foi aplicado sobre os 176 protótipos restantes. Com isso, um dendrograma contendo 176 níveis de configurações de agrupamentos diferentes para o SOM filtrado foi gerado.

Utilizando a informação de que a imagem apresenta 6 classes de cobertura terrestres predominantes (citadas anteriormente), aquele nível do dendrograma que apresentou 6 agrupamentos (ou classes) de protótipos para o SOM foi escolhido para realizar a classificação da cena. A Figura 2(a) mostra a grade de protótipos do SOM segmentada em 6 agrupamentos. Os protótipos marcados com ‘o’ e com ‘x’ correspondem, respectivamente, aos protótipos inativos e heterogêneos filtrados na etapa anterior. É possível observar que tais protótipos funcionam como “bordas” separando algumas classes de protótipos do SOM.

Para realizar a classificação final, a imagem foi percorrida inteiramente tomando janelas de pixels com dimensões  $9 \times 9$  (iguais aos das janelas amostrais) e comparado-as com todos os protótipos do SOM. Essa comparação foi feita através das distâncias calculadas entre a janela de pixels considerada e cada um dos protótipos. Ao pixel central de cada janela de pixels da imagem foi atribuído o rótulo do protótipo que apresentou a menor distância a ela. Em seguida, cada um dos pixels associados aos protótipos heterogêneos do SOM foi classificado atribuindo um rótulo igual ao do pixel

vizinho (já classificado) com a menor distância a ele em termos de atributos espectrais. A Figura 2(b) mostra a imagem teste classificada pelo método proposto.



**Figura 2. (a) Grade de protótipos do SOM segmentada em 6 agrupamentos. (b) Imagem teste classificada pelo método proposto.**

Com o objetivo de validar a classificação resultante do método proposto, e considerando ainda a ausência de “verdade terrestre” para a imagem teste, foi realizada uma classificação supervisionada da imagem através de uma Rede Neural de *Perceptrons* Multicamadas, e o resultado da mesma foi utilizado como referência (ou “verdade”) para o cálculo do índice de concordância *kappa* (normalmente utilizado para avaliar a precisão de classificação de imagens de satélite). O valor obtido para o índice calculado com as duas imagens foi de 0.82, o que nos permite concluir que a qualidade da classificação produzida pelo método proposto foi bastante satisfatória.

## 6. Conclusões e Considerações Finais

Este trabalho mostra que o SOM pode ser uma ferramenta eficaz para realizar a análise de agrupamentos de imagens de satélite. Ao invés de trabalhar diretamente com os padrões originais da cena o método proposto analisa os dados da imagem através de um conjunto de protótipos do SOM. Essa abordagem reduz significativamente a complexidade da análise tornando possível a utilização de técnicas que são normalmente consideradas inviáveis para o processamento de imagens de sensoriamento remoto, como por exemplo, métodos de agrupamentos hierárquicos. Além disso, o método apresenta outras vantagens e potencialidades que o colocam como uma alternativa promissora para a classificação de imagens de satélite. Entre elas, podemos destacar:

- A não exigência de uma definição a priori do número de classes para realizar a classificação da imagem, o que não acontece na maioria dos métodos convencionais de classificação não-supervisionada de imagens de satélite.

- A representação distribuída das classes por meio de grupos de protótipos dá ao método a potencialidade de descobrir agrupamentos de dados que possuem geometrias complexas e variadas. Métodos como o K-médias utilizam apenas um único protótipo (centróide) para representar cada classe e por isso são capazes de detectar adequadamente apenas agrupamentos que tenham formatos hipersféricos.

- A utilização de um método de agrupamento hierárquico permite que o usuário compreenda em diferentes níveis de agrupamentos as relações existentes entre os padrões de cobertura terrestre presentes na imagem. Isto pode ser muito útil em situações onde a estrutura da informação presente na imagem não é muito clara. Além



disso, caso o usuário deseje, a imagem pode ser classificada considerando diferentes quantidades de classes.

- O método proposto apresenta um mecanismo de fusão eficiente utilizado pelo método de agrupamento hierárquico empregado para segmentar o SOM, o qual respeita as relações de vizinhança do SOM evitando que todos os pares de agrupamentos sejam analisados para se decidir sobre uma nova fusão. Além disso, as fusões não são decididas considerando somente a informação de distância entre os agrupamentos. Informações espaciais sobre os agrupamentos também são incorporadas e certamente contribuem no processo de separação das classes.

Como trabalho futuro, pretende-se aplicar versões modificadas de índices de validação de agrupamentos, conforme proposto em (Gonçalves et al., 2006), para decidir de maneira automática qual o nível do dendrograma do SOM apresenta o número ideal de agrupamentos para realizar a classificação da imagem. Pretende-se também realizar comparações de desempenho e complexidade com métodos convencionais de classificação de imagens de sensoriamento remoto.

## Referências

- Ball, G. and Hall, D. (1967), “A Clustering Technique for Summarizing Multivariate Data”, *Behavior Science*, v. 12, p. 153-155.
- Costa, J. A. F. (1999), *Classificação Automática e Análise de Dados por Redes Neurais Auto-organizáveis*. Tese (Doutorado em Engenharia Elétrica). Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas-SP.
- Costa, J. A. F. (2005), “Segmentação do SOM por Métodos de Agrupamentos Hierárquicos com Conectividade Restrita”, In: *Anais do VII Congresso Brasileiro de Redes Neurais*, outubro 2005, Natal, RN.
- Gonçalves, M. L., Netto, M. L. A., Costa, J. A. F. and Zullo Jr., J. (2006), “Data Clustering using Self-Organizing Maps segmented by Mathematic Morphology and Simplified Cluster Validity Indexes”, In: *Proceedings of IEEE International Joint Conference on Neural Networks*, Vancouver, July 2006, p. 8854-8861.
- Haralick, R. M., Shanmugam, K. and Dinstein, I. (1973), “Textural Features for Image Classification”, In: *IEEE Transactions on Systems, Man and Cybernetics*, v. 3, n. 6, p. 610-621.
- Haykin, S. (2001), *Redes neurais: princípios e práticas*. Porto Alegre: Bookmann, 2 ed..
- Kohonen, T. (1997), *Self-Organizing Maps*, Springer Verlag (Berlim), 2nd edition.
- Marçal, A. R. S. and Castro, L. (2005), “Hierarchical Clustering of Multispectral Images Using Combined Spectral and Spatial Criteria”, In: *IEEE Geoscience and Remote Sensing Letters*, v. 2, n. 1, p. 59-63.
- Richards, J. A. (2005) “Analysis of Remotely Sensed Data: The Formative Decades and the Future”, In: *IEEE Transactions on Geoscience and Remote Sensing*, v. 43, n. 3, p. 422-432.
- Xu, R. and Wunsch II, D. (2005), *Survey of Clustering Algorithms*. In: *IEEE Transactions on Neural Networks*. v. 16, n. 3, p. 645 - 678.

# Análise de agrupamentos usando Mapas de Kohonen segmentados por morfologia matemática e o índice de validação CDbw

Márcio L. Gonçalves<sup>1</sup>, Márcio L. de Andrade Netto<sup>2</sup>, José Alfredo F. Costa<sup>3</sup>, Jurandir Zullo Jr.<sup>4</sup>

<sup>1</sup>PUC Minas Campus de Poços de Caldas - E-mail: marcio@pucpcaldas.br

<sup>2</sup>DCA/FEEC/UNICAMP – E-mail: marcio@dca.fee.unicamp.br

<sup>3</sup>DEE – Centro de Tecnologia - UFRN – E-mail: alfredo@ufrnet.br

<sup>4</sup>CEPAGRI/UNICAMP – E-mail: jurandir@cpa.unicamp.br

## Abstract

*This work presents a new strategy of automatic clusters analysis using Kohonen's Maps segmented by mathematical morphology and the clustering validity index CDbw (Composing Density Between and Within Clusters)[12]. The strategy is presented as an improvement of the method proposed in [2]. The incorporation of a clustering validity index in the data partitioning process makes possible to analyze the quality of clusters that are generated by different segmentations of a trained SOM. Experimental results show that the new strategy allows to determine in more efficient way the correct partitioning of a data set.*

## 1. Introdução

O Mapa Auto-organizável de Kohonen (SOM – *Self-Organizing Map*) tem sido utilizado constantemente em diversas aplicações de engenharia cobrindo áreas como reconhecimento de padrões, análise de imagens e texto, quantização vetorial, regressão, problema do caixeiro viajante e mineração de dados [1-4,6,8,10,13].

Embora o SOM seja empregado largamente como uma ferramenta de visualização de dados que apresentam dimensões elevadas, a sua aplicação em tarefas de classificação automática de dados ou de análise de agrupamentos tem sido significativa. Diferentemente de métodos supervisionados de reconhecimento de padrões, nos quais um conjunto rotulado de amostras é usado para treinar o classificador, o SOM realiza um mapeamento, via aprendizado não-supervisionado, de um espaço  $p$ -dimensional contínuo para um conjunto discreto de vetores de referência, ou neurônios, dispostos na forma

de um arranjo espacial regular, normalmente bidimensional. O SOM funciona como uma rede elástica ocupando o espaço  $p$ -dimensional de forma a representar da melhor maneira as regiões com maior densidade de pontos.

Uma vez que em um SOM tradicional a única informação de saída quando se apresenta um padrão,  $x$ , são os índices  $(i,j)$  do neurônio vencedor,  $c$ , e o erro de quantização, que pode ser dado pela distância  $d(x,c)$ , a rede necessita de ferramentas adicionais para que uma interpretação adequada sobre o seu mapa de saída seja realizada. Diferentemente de outras aplicações, nas quais o SOM é usado como uma ferramenta de auxílio visual para indicar tendências de agrupamentos, em tarefas de análise de agrupamentos (ou classificação automática), cujo objetivo básico é descobrir a estrutura inerente dos dados, a detecção de agrupamentos através dos neurônios do SOM deve ser feita sem a intervenção de um usuário.

Na literatura podemos encontrar alguns trabalhos propondo mecanismos e procedimentos para analisar e interpretar de maneira automática os neurônios de um SOM treinado. Costa e Netto [2] apresentaram um método de segmentação do SOM aplicando técnicas de morfologia matemática sobre uma imagem de distâncias entre neurônios adjacentes. O resultado do algoritmo são regiões conectadas de neurônios rotuladas, e que definem no espaço de entrada geometrias complexas e não paramétricas. Vesanto e Alhoniemi [5] utilizaram métodos hierárquicos aglomerativos e o algoritmo K-médias para detectar agrupamentos de neurônios em um SOM treinado. No entanto, além do processo de descoberta dos agrupamentos exigir muita interação com o usuário, não caracterizando, portanto, um método totalmente automático, os autores utilizaram como critério para

fusão ou separação de grupos de dados, o índice Davies-Bouldin [7], que baseia-se nos centróides dos agrupamentos, não permitindo uma avaliação adequada de estruturas não hiperesféricas. Costa e Netto [3] também propuseram um método de segmentação do SOM baseado em particionamento de grafos. A estratégia consiste em eliminar ramos inconsistentes que correspondem a ligações entre neurônios vizinhos que não satisfazem pré-requisitos relacionados com informações de distância entre neurônios, erro de quantização e atividade. Silva et al. [8] aplicaram com sucesso o método de segmentação baseado em particionamento de grafos juntamente com índices de validação de agrupamentos na análise de dados geoespaciais. No trabalho de Kiang [9], o autor utilizou um método hierárquico aglomerativo empregando critérios de variância mínima para junção ou divisão de grupos de neurônios. O algoritmo necessita recalculer os centróides sempre que dois grupos são unidos e apresenta boa aplicabilidade apenas para agrupamentos com formas hiperesféricas ou hiperelipsoidais. Recentemente, Wu e Chow [14] propuseram também um método hierárquico aglomerativo para analisar a saída do SOM utilizando um índice de validação de agrupamentos. A aplicação do método em algumas bases de dados artificiais e reais mostra que o seu desempenho é superior em relação a outras técnicas de agrupamentos.

Através desses trabalhos publicados observa-se, portanto, a existência de diferentes propostas para analisar agrupamentos de dados a partir da segmentação automática ou semi-automática dos neurônios de um SOM treinado, cada uma delas com metodologias de agrupamentos diferenciadas (hierárquicas ou particionais) e com vantagens e desvantagens sob diferentes aspectos. Especificamente, pode-se observar também que a incorporação de índices de validação de agrupamentos pode ajudar no processo, certamente quando índices adequados são utilizados.

Nesse sentido, o presente trabalho apresenta um método alternativo para a análise de agrupamentos empregando a técnica de segmentação do SOM baseada em morfologia matemática, proposta por Costa e Netto [2], em conjunto com o índice de validação de agrupamentos CDbw [12], que trata de maneira eficiente agrupamentos com formatos arbitrários. O método proposto apresenta um aprimoramento da técnica de análise de agrupamentos apresentada em [2], uma vez que a incorporação de um índice de validação permite avaliar a qualidade de diferentes configurações de agrupamentos dos dados que são geradas por diferentes partições do conjunto de neurônios do SOM,

possibilitando assim determinar de forma mais eficiente aquela que corresponde a partição mais adequada.

Este artigo apresenta-se organizado da seguinte forma: a seção 2 descreve brevemente o algoritmo SOM e a seção 3 apresenta o índice de validação de agrupamentos CDbw. A seção 4 apresenta a estratégia de análise de agrupamentos proposta. A seção 5 mostra alguns resultados experimentais, e a seção 6 apresenta as conclusões obtidas.

## 2. Mapa Auto-organizável de Kohonen

O SOM é uma rede neural artificial que define um mapeamento de um espaço de entrada  $p$ -dimensional contínuo para um conjunto finito de vetores referências (neurônios), que são dispostos em um arranjo topológico, geralmente bidimensional [1].

A rede consiste essencialmente de duas camadas de neurônios: uma camada de entrada  $I$  e uma camada de saída  $U$  (camada de Kohonen). A entrada da rede corresponde a um vetor  $p$ -dimensional,  $\mathbf{x}$ , geralmente no espaço  $\mathfrak{R}^p$ . Todas as  $p$  componentes do vetor de entrada alimentam cada um dos neurônios do mapa. Cada neurônio  $i$  pode ser representado então por um vetor de pesos  $\mathbf{w}_i = [w_{i1}, w_{i2}, \dots, w_{ip}]^T$ , também no espaço  $p$ -dimensional.

Para cada padrão de entrada um neurônio é escolhido o vencedor,  $c$ , usando o critério de maior similaridade:

$$\|\mathbf{x} - \mathbf{w}_c\| = \min_i \{\|\mathbf{x} - \mathbf{w}_i\|\}$$

onde  $\|\cdot\|$  representa a distância Euclidiana. Os pesos do neurônio vencedor, juntamente com os pesos dos seus neurônios vizinhos, são ajustados de acordo com a seguinte equação:

$$\mathbf{w}_i(t+1) = \mathbf{w}_i(t) + h_c(t)[\mathbf{x}(t) - \mathbf{w}_i(t)]$$

onde  $t$  indica a iteração do processo de treinamento,  $\mathbf{x}(t)$  é o padrão de entrada e  $h_c(t)$  é o núcleo de vizinhança em torno do neurônio vencedor.

Uma propriedade importante do SOM, derivada da quantização vetorial gerada pelo algoritmo, é aproximar a densidade dos dados de entrada pela densidade de neurônios em um mapa treinado. Isso possibilita obter informações dos agrupamentos analisando as relações geométricas dos neurônios. Ultsch [15] descreveu a matriz de distâncias unificadas, a  $U$ -matrix, que permite obter uma imagem das distâncias entre neurônios adjacentes. Essa imagem pode ser visualizada como uma superfície 3D, na qual "vales" correspondem a neurônios que são similares. Valores altos na  $U$ -matrix, que correspondem as



“montanhas” na imagem da superfície 3D, refletem a dissimilaridade entre neurônios vizinhos e podem ser associados a regiões de fronteiras de agrupamentos. Em Costa e Netto [2], os autores apresentaram uma técnica eficiente baseada em morfologia matemática para segmentar a *U-matrix* e obter informações sobre os agrupamentos de neurônios. Conforme será descrito na seção 4, a mesma técnica é utilizada no presente trabalho para realizar diferentes segmentações na *U-matrix*.

### 3. Índice de Validação de Agrupamentos

Na literatura podemos encontrar diferentes métodos para investigar a validação de agrupamentos de dados [7,12,16]. A implementação da maioria dos algoritmos de validação exige alto custo computacional, principalmente quando o número de agrupamentos e a quantidade dos dados de entrada são muito grandes. Alguns índices são dependentes dos dados e/ou do número de agrupamentos. Existem índices que utilizam os centróides de subconjuntos de dados em seus algoritmos, enquanto que outros usam todos os pontos de cada subconjunto. Quando os agrupamentos apresentam geometrias ou formatos arbitrários, o uso de um conjunto de vetores para representar cada agrupamento gera melhores resultados que a utilização de apenas um único vetor (centróide) como representante de cada grupo de dados, além de apresentar um custo computacional menor em comparação com os algoritmos que consideram todos os pontos de cada agrupamento [16].

Halkidi e Vazirgiannis [12] propuseram um índice de validação, o CDbw (*Composing Density Between and With Clusters*), que apresenta algumas características que o torna mais eficiente em relação a outros índices. O método baseia-se em dois conceitos importantes: a densidade intra-agrupamento e a separação entre os agrupamentos. A característica geométrica do agrupamento é representada pelo uso de vetores representativos, isso permite que o índice avalie de maneira correta estruturas não hiperesféricas, o que não ocorre em outros índices [7]. Além disso, a complexidade computacional do método é  $O(n)$  [12], que é aceitável para grandes conjuntos de dados.

O cálculo do índice CDbw é descrito a seguir. Para um conjunto de dados particionados em  $c$  agrupamentos, define-se um conjunto de pontos representativos  $V_i = \{v_{i1}, v_{i2}, \dots, v_{ir}\}$  para o agrupamento  $i$ , onde  $r$  representa o número de pontos de representação do agrupamento  $i$ . A densidade intra-agrupamento é definida por:

$$Intra(c) = \frac{1}{c} \sum_{i=1}^c \frac{1}{r_i} \sum_{j=1}^{r_i} density(v_{ij}), c > 1$$

com  $density(v_{ij}) = \sum_{l=1}^{n_i} f(x_l, v_{ij})$ , onde  $x_l$  pertence ao  $i$ -ésimo agrupamento,  $v_{ij}$  é a  $j$ -ésima representação do agrupamento  $i$ , e  $f(x_l, v_{ij})$  é dado por 1, se  $\|x_l - v_{ij}\| \leq stdev$ , ou 0 caso contrário.  $stdev$  corresponde a média dos desvios padrões de cada um dos  $c$  agrupamentos.

A densidade inter-agrupamento é dada por:

$$Inter(c) = \sum_{i=1}^c \sum_{j=1}^c \frac{\|close\_rep(i) - close\_rep(j)\|}{\|stdev(i)\| + \|stdev(j)\|} density(v_{ij}),$$

onde  $c > 1$ ,  $close\_rep(i)$  e  $close\_rep(j)$  representam o par de pontos de representação mais próximos entre o agrupamento  $i$  e  $j$ ,  $v_{ij}$  é o ponto médio entre este par de pontos,  $stdev(i)$  e  $stdev(j)$  representam os desvios padrões dos agrupamentos  $i$  e  $j$ , e  $density(v_{ij}) = \sum_{k=1}^{n_i+n_j} f(x_k, v_{ij})$ , onde  $x_k$  pertence ao agrupamento  $i$  ou  $j$ , e  $f(x_k, v_{ij})$  é dado por 1 se  $\|x - v\| \leq (\|stdev(i)\| + \|stdev(j)\|)$ , ou 0 caso contrário.

A separação entre os agrupamentos é definida por:

$$Sep(c) = \sum_{i=1}^c \sum_{j=1}^c \frac{\|close\_rep(i) - close\_rep(j)\|}{1 + Inter(c)}, c > 1$$

O índice CDbw é dado então por:

$$CDbw(c) = Intra(c) * Sep(c).$$

Experimentos em [12] demonstram que uma boa partição dos dados é indicada por valores altos do índice.

No método de análise de agrupamentos proposto neste trabalho (seção 4), o índice CDbw foi utilizado para avaliar a qualidade de diferentes configurações de agrupamentos dos dados. Diferentemente da proposta apresentada em [12], que encontra os vetores de referência dos agrupamentos de forma iterativa a partir do conjunto de dados, em nossa abordagem os próprios vetores de pesos dos neurônios do SOM são utilizados como os vetores de referência dos seus respectivos agrupamentos, simplificando assim o cálculo do índice CDbw.

### 4. Metodologia proposta

A estratégia de análise de agrupamentos proposta neste trabalho procura encontrar a melhor partição para um conjunto de dados amostrais a partir da análise de diferentes segmentações realizadas sobre a *U-matrix* (descrita brevemente na seção 2).

Costa e Netto [2] propuseram um método eficiente baseado em morfologia matemática para realizar a segmentação da *U-matrix*. O método emprega o algoritmo de segmentação de imagens, *watershed*, utilizando uma imagem de marcadores para regularizar o processo de segmentação. Essa mesma estratégia também é aplicada em nossa proposta, porém ao invés de definirmos apenas uma imagem de marcadores para segmentar a *U-matrix* (como apresentado em [2]), várias imagens de marcadores são determinadas e, portanto, diferentes segmentações da *U-matrix* são obtidas. Cada uma dessas segmentações é associada aos neurônios do SOM, obtendo-se assim diferentes partições do mapa de neurônios que definem diferentes configurações de agrupamentos dos dados. Para cada uma das diferentes configurações de agrupamentos é aplicado o índice de validação CDbw, que serve como critério de decisão para a escolha da melhor partição dos dados. Aquela partição que apresenta o maior valor para o índice CDbw é considerada a melhor.

Considerando a *U-matrix* de um SOM treinado dada pela imagem  $U$  com 256 níveis de cinza, os seguintes passos são efetuados para se obter o conjunto de imagens de marcadores para  $U$ :

1. Filtragem: obter uma imagem  $U_1$  suavizando a imagem  $U$  através da remoção de pequenas depressões com área inferior a 3 pixels;
2. Para  $k = 1, \dots, f_{max}$ , onde  $f_{max}$  é o nível de cinza máximo na imagem  $U_1$ , criar as imagens binárias  $U_2^k$  correspondendo a conversões de  $U_1$  usando  $k$  como valor de limiar;
3. Obter o número de regiões conectadas de  $U_2^k$ , para cada valor de  $k$ ,  $N_{rc}^k$ ;
4. Construir o gráfico  $k \times N_{rc}^k$ , e através dele obter o conjunto  $S_k$  composto por todos os valores de  $k$  que correspondem a inícios de seqüências contíguas e constantes de números de regiões conectadas com tamanhos superiores a 3;
5. Obter o conjunto de todas as imagens de marcadores,  $S_m = \{U_2^{k_1}, U_2^{k_2}, \dots, U_2^{k_n}\}$ , onde  $k_1, k_2, \dots, k_n$  são os elementos de  $S_k$  obtidos no passo 4.

Embora o procedimento usado para encontrar os marcadores da *U-matrix* seja semelhante ao apresentado em [2], em nossa abordagem os passos 4 e 5 determinam um conjunto de imagens de marcadores da *U-matrix* considerando todas as seqüências contíguas e constantes de regiões conectadas. Em todos os experimentos foram considerados apenas as seqüências com tamanhos maiores que 3.

Sendo assim, a estratégia geral de particionamento de um conjunto de dados, a partir de um SOM treinado com sucesso, pode ser resumida como segue:

1. Obtenha a *U-matrix* a partir do SOM treinado;
2. Encontre o conjunto,  $S_m = \{U_2^{k_1}, U_2^{k_2}, \dots, U_2^{k_n}\}$ , de imagens de marcadores para a *U-matrix*;
3. Para cada imagem de marcadores,  $U_2^{k_i}$ ,  $i=1,2,\dots,n$ , faça:
  - 3.1. Aplicar o algoritmo watershed sobre a *U-matrix*;
  - 3.2. Rotular as regiões conectadas da imagem segmentada no passo 3.1;
  - 3.3. Realizar a cópia dos rótulos obtidos no passo 3.2 para os neurônios do SOM associados a cada pixel da *U-matrix*;
  - 3.4. Rotular o conjunto de dados de entrada utilizando o SOM rotulado no passo 3.3;
  - 3.5. Aplicar o índice de validação de agrupamentos CDbw sobre o conjunto de dados rotulados obtido no passo 3.4, usando os vetores de pesos do SOM como vetores referência dos agrupamentos.
4. Escolher a partição de dados que apresenta o maior valor para o índice CDbw.

Para ilustrar a aplicabilidade da metodologia proposta resultados experimentais com dados de sensoriamento remoto são apresentados na próxima seção.

## 5. Resultados Experimentais

Uma imagem composta pelas bandas espectrais 3, 4 e 5 do satélite Landsat-5 TM [17], mostrando a cidade de Manaus, e o encontro dos Rios Negro e Solimões foi utilizada nos experimentos. Uma composição colorida das 3 bandas da imagem, que possui resolução espacial de 512x496 pixels, é apresentada na figura 1.



Figura 1 – Composição colorida da imagem de satélite.

A cena apresenta 4 grandes classes de cobertura terrestre: vegetação (classe 1), área urbana (classe 2),

água 1 (classe 3) e água 2 (classe 4). Uma amostragem de 2500 pixels selecionados de maneira aleatória e uniforme foi feita sobre a imagem. O gráfico da figura 2 apresenta as amostras de pixels plotadas no espaço 3D onde é possível visualizar as 4 classes de padrões. É possível notar também que as classes não apresentam fronteiras bem definidas, tornando mais difícil o processo de particionamento das mesmas.

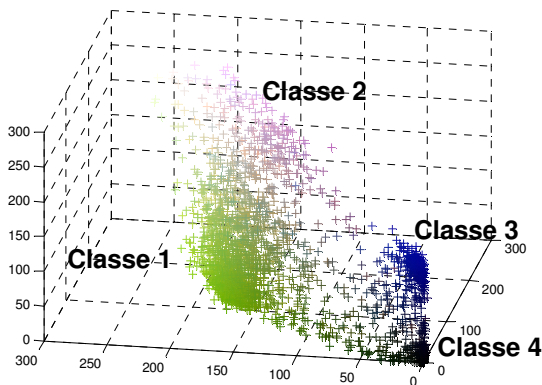


Figura 2 – Pontos amostrais da imagem.

Um SOM retangular com tamanho 15x10 foi treinado. Foram realizadas 500 épocas de treinamento e os raios de vizinhança inicial e final usados foram 12 e 1, respectivamente. A figura 3 mostra a *U-matrix* na forma de superfície obtida a partir do conjunto de neurônios do SOM treinado. O gráfico do número de regiões conectadas versus o limiar  $k$  da *U-matrix* é apresentado na figura 4.

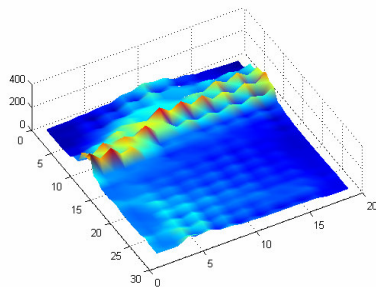


Figura 3 – *U-matrix* (3D).

Conforme pode ser observado no gráfico da figura 4, foram obtidos oito valores de  $k$  (7, 24, 37, 94, 107, 114, 121, 163) que correspondem a inícios de seqüências contíguas e constantes de números de regiões conectadas com tamanhos superiores a 3. Para cada um dos valores de  $k$  foi determinada então a imagem de marcadores da *U-matrix* e aplicada a estratégia de particionamento dos dados. No gráfico da figura 5 é possível visualizar o valor do índice CDbw que foi calculado para cada uma das partições de dados geradas a partir dos oito valores de  $k$ . A partição

correspondente ao valor de  $k$  igual a 94 foi aquela que apresentou o maior valor para o índice CDbw e foi considerada, portanto, a melhor partição. Observando o gráfico da figura 4 nota-se que o número de agrupamentos é 4 para o valor de  $k$  igual a 94, o que coincide com a quantidade de classes presentes na imagem.

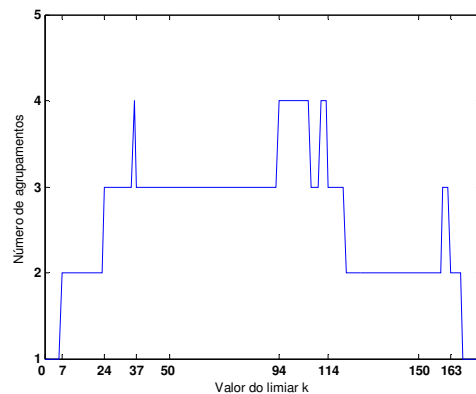


Figura 4 – Gráfico do número de regiões conectadas versus o limiar  $k$  da *U-matrix*.

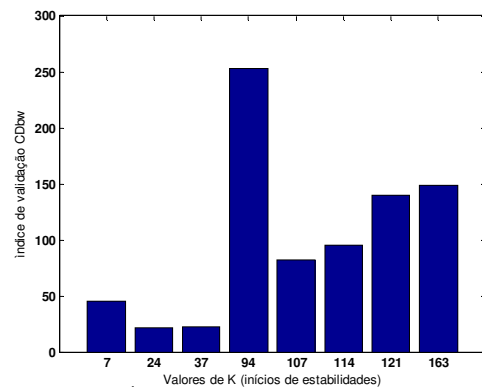


Figura 5 – Índice CDbw para as partições geradas a partir dos valores de  $k$  estáveis.

A figura 6 ilustra o particionamento dos pontos amostrais que foi gerado a partir do valor de  $k$  que apresentou o maior índice CDbw (cada cor representa um agrupamento diferente). Embora as classes de padrões apresentem estruturas não muito bem definidas, o algoritmo proposto neste trabalho, que não teve acesso a nenhuma informação sobre o número de agrupamentos e nenhuma intervenção do usuário (ou de um especialista em imagens), conseguiu de forma totalmente não-supervisionada identificá-las e particioná-las com êxito. O tempo gasto, desde o treinamento do SOM até a classificação final dos dados foi de aproximadamente 260 segundos, em uma máquina com processador AMD Athlon™ XP 2600+ 1,91GHZ e 512 de RAM.

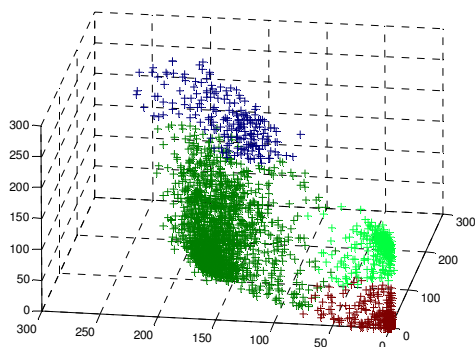


Figura 6 – Partição dos dados obtida a partir de  $k = 94$ .

## 6. Conclusões

Este trabalho apresentou uma nova estratégia de análise automática de agrupamentos usando mapas de Kohonen segmentados por morfologia matemática e o índice de validação de agrupamentos CDbw. A estratégia adotada apresenta-se como um aprimoramento do método proposto por Costa e Netto [2]. Ao invés de definirmos apenas uma imagem de marcadores para segmentar a *U-matrix* (como apresentado em [2]), várias imagens de marcadores são determinadas e, portanto, diferentes segmentações da *U-matrix* são obtidas. A estratégia apresentada considera então diferentes partições do SOM e avalia a qualidade dos agrupamentos que são gerados por eles aplicando um índice de validação adequado, podendo alcançar de maneira mais eficiente o particionamento ideal do conjunto de dados. No exemplo mostrado nesse artigo, vê-se na figura 4 que o número de agrupamentos é igual a 3 para o valor de  $k$  igual a 37, que corresponde ao início da seqüência que apresenta maior estabilidade, abordagem adotada em Costa e Netto [2]. Os dados pertencentes as classes 1 e 2 não foram particionados e fizeram parte do mesmo agrupamento, o que exige mais uma iteração do algoritmo para encontrar todas as 4 classes de padrões. Através da abordagem apresentada aqui foi possível encontrar de forma mais eficaz a partição correta do conjunto de dados (figura 6), avaliando diferentes configurações de agrupamentos que foram geradas por diferentes partições do SOM.

A estratégia de particionamento de dados foi aplicada também a outras bases de dados artificiais e reais obtendo resultados satisfatórios [11]. O bom desempenho da estratégia deve-se também a adequação do índice de validação CDbw à representação distribuída dos agrupamentos através dos neurônios do SOM, permitindo encontrar uma boa representação para agrupamentos que apresentam geometrias variadas e complexas.

## 7. Referências Bibliográficas

- [1] Kohonen, T., *Self-Organizing Maps*, 2<sup>nd</sup> Edition, Berlin: Springer Verlag, 1997.
- [2] Costa, J.A.F., Netto, M. L. A., “Clustering of complex shaped data sets via Kohonen maps and mathematical morphology”, *Proceedings of the SPIE, Data Mining and Knowledge Discovery*, vol. 4384, pp. 16-27, 2001.
- [3] Costa, J.A.F., Netto, M.L.A., “Segmentação do SOM Baseada em Particionamento de Grafos”, *Congresso Brasileiro de Redes Neurais*, São Paulo, 2003.
- [4] Kohonen, T., Kaski, S., Lagus, K., Salojärvi, J., Honkela, J., Paatero, V., Saarela, A., “Self Organization of a Massive Document Collection”, *IEEE Transactions on Neural Networks*, vol. 11, (3), pp. 574-585, 2000.
- [5] Vesanto, J., Alhoniemi, E., “Clustering of the Self-organizing Map”, *IEEE Transactions on Neural Networks*, v. 11, (3), pp. 586-602, 2000.
- [6] Suganthan, P. N., “Pattern Classification using multiple hierarchical overlapped self-organising maps”, *Pattern Recognition*, vol. 34, pp. 2173-2179, 2001.
- [7] Davies, D. L., Bouldin, D. W., “A Cluster separation measure”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. PAMI-1, pp. 224-227, 1979.
- [8] Silva, M. A. S., Monteiro, A. M. V., Medeiros, J. S., “Semi-Automatic Geospatial Data Clustering by Self-Organizing Maps”, *Simpósio Brasileiro de Redes Neurais*, São Luiz, MA, 2004.
- [9] Kiang, M. Y., “Extending the Kohonen self-organizing map networks for clustering analysis”, *Computational Statistics & Data Analysis*, v. 38, pp. 161-180, 2001.
- [10] Gonçalves, M. L., Netto, M. L. A., Costa, J. A. F., Zullo Jr., J., “Automatic Remotely Sensed Data Clustering by Tree-structured Self-Organizing Map”, *Proc. of IEEE International Geosciences and Remote Sensing Symposium – IGARSS’05*, Korea, 2005.
- [11] Gonçalves, M. L., Netto, M. L. A., “Uma abordagem para análise de agrupamentos baseada em Mapas de Kohonen segmentados por Morfologia Matemática e índices de validação”. TR-002/05 DCA-FEEC-UNICAMP, 2005.
- [12] Halkidi, M., Vazirgiannis, M., “Clustering validity assessment using multi representatives”, *Proceedings of SETN Conference*, Thessaloniki, Grécia, 2002.
- [13] Vieira, F.C., Dória Neto, A.D., e Costa, J.A.F., “An Efficient Approach To The Travelling Salesman Problem Using Self-Organizing Maps”, *International Journal of Neural Systems*, vol. 13, pp. 59-66, 2003.
- [14] Wu, S., Chow, T. W. S., “Clustering of the self-organizing map using a clustering validity index based on inter-cluster and intra-cluster density”, *Pattern Recognition*, v. 37, pp. 175-188, 2004.
- [15] Ultsch, A., “Self-organizing Neural Networks for Visualization and Classification”, *Information and Classification*, Berlin, Springer-Verlag, pp. 307-313, 1993.
- [16] Bezdek, J. C., Pal, N. R., “Some new indexes of cluster validity”, *IEEE Trans. System, Man, and Cybern.*, v. 28 (3), pp. 301-315, 1998.
- [17] <http://www.dgi.inpe.br/html/imagens.htm>, Divisão de Geração de Imagens – INPE, recuperado em 15/02/05.

# CLASSIFICAÇÃO NÃO-SUPERVISIONADA DE IMAGENS DE SENSORES REMOTOS UTILIZANDO REDES NEURAIS AUTO-ORGANIZÁVEIS E MÉTODOS DE AGRUPAMENTOS HIERÁRQUICOS

*Unsupervised classification of remote sensors images using self-organizing neural networks and hierarchical clustering methods*

**Márcio Leandro Gonçalves<sup>1,2</sup>**  
**Márcio Luiz de Andrade Netto<sup>2</sup>**  
**Jurandir Zullo Jr.<sup>3</sup>**  
**José Alfredo Ferreira Costa<sup>4</sup>**

<sup>1</sup>**Pontifícia Universidade Católica de Minas Gerais – PUC Minas**  
**Departamento de Ciência da Computação**  
Av. Padre Francis Cletus Cox, 1661, CEP 37701-355, Poços de Caldas-MG, Brasil  
marcio@pucpcaldas.br

**Universidade Estadual de Campinas - UNICAMP**

<sup>2</sup>**Faculdade de Engenharia Elétrica e de Computação – FEEC**

<sup>3</sup>**Centro de Pesquisas Meteorológicas e Climáticas Aplicadas à Agricultura – CEPAGRI**  
Cidade Universitária Zeferino Vaz, CEP 13083-970, Campinas – SP, Brasil  
marcio@dca.fee.unicamp.br, jurandir@cpa.unicamp.br

<sup>4</sup>**Universidade Federal do Rio Grande do Norte - UFRN**  
**Departamento de Engenharia Elétrica – DEE**  
Campus Universitário - Lagoa Nova, CEP 59072-970, Natal – RN, Brasil  
alfredo@dee.ufrn.br

## RESUMO

Este trabalho apresenta uma nova metodologia para a classificação não-supervisionada de imagens de sensoriamento remoto. Diferentemente dos métodos convencionais de classificação não-supervisionada, como o K-médias e o ISODATA, os quais se baseiam somente em técnicas de agrupamentos particionais, a metodologia proposta realiza a classificação automática de imagens através de uma abordagem inovadora empregando o Mapa Auto-organizável de Kohonen (SOM – *Self-Organizing Map*) em conjunto com um método de agrupamento hierárquico aglomerativo. O ponto chave do método proposto é executar o processo de análise de agrupamentos através de um conjunto de protótipos do SOM ao invés de trabalhar diretamente com os padrões originais da imagem. Essa estratégia reduz significativamente a complexidade da análise dos dados tornando possível a utilização de técnicas normalmente consideradas inviáveis para o processamento de imagens de sensoriamento remoto, como métodos de agrupamentos hierárquicos e índices de validação de agrupamentos. Através do SOM, o método proposto mapeia os padrões originais da imagem para uma grade de neurônios bidimensional procurando preservar a distribuição de probabilidade e a topologia dos mesmos. Posteriormente, um método de agrupamento hierárquico aglomerativo com restrições de conectividade é aplicado sobre a grade de neurônios já treinada, gerando um dendrograma simplificado para os dados da imagem. Cada nível do dendrograma apresenta uma configuração diferente de agrupamentos de neurônios (ou protótipos) do SOM que pode ser utilizada para representar as classes sobre as quais a imagem original será classificada. Aplicando versões modificadas de índices de validação de agrupamentos o método determina automaticamente o número ideal de agrupamentos da imagem não exigindo que o usuário defina previamente a quantidade de classes para realizar o processo de classificação. Os resultados experimentais mostram um exemplo de aplicação da metodologia proposta sobre uma imagem teste e compara o seu desempenho com o do algoritmo K-médias.

**Palavras chaves:** processamento digital de imagens, sensoriamento remoto, classificação não-supervisionada, redes neurais artificiais.

## ABSTRACT

This work presents a new methodology for the unsupervised classification of remotely sensed images. Differently of the traditional methods of unsupervised classification, such as K-means and ISODATA, which use only partitional clustering techniques, the proposed methodology accomplishes the automatic classification of images through an innovative approach applying the Kohonen Self-Organizing Map (SOM) together with an agglomerative hierarchical clustering method. The key point of the proposed method is to execute the clustering process through a set of prototypes of the SOM instead of analyzing directly the original patterns of the image. This approach significantly reduces the complexity of the analysis becoming possible the use of techniques that normally are considered impracticable for the digital processing of remotely sensed images, such as hierarchical clustering methods and cluster validity indexes. Using the SOM, the proposed method maps the original patterns of the image to a set of neurons arranged in a two-dimensional lattice searching to preserve the probability distribution and the topology of the input space. Subsequently, an agglomerative hierarchical clustering method with restricted connectivity is applied on the lattice of neurons previously trained, generating a simplified dendrogram for the image data. Each level of the dendrogram corresponds to a different configuration of clusters of neurons (or prototypes) of the SOM that can be used to represent the classes on which the original image will be classified. Applying modified versions of cluster validity indexes the method automatically determines the ideal number of clusters of the image not demanding that the user previously defines the quantity of classes to execute the classification process. The experimental results show an application example of the proposed method on a test image and its performance is compared with the K-means algorithm.

**Keywords:** digital image processing, remote sensing, unsupervised classification, artificial neural networks.

## 1. INTRODUÇÃO

Desde o lançamento dos primeiros satélites voltados para a exploração de recursos terrestres, os métodos digitais de classificação de imagens de sensoriamento remoto têm adquirido uma importância crescente no reconhecimento automático de padrões da superfície terrestre (RICHARDS *et al.*, 2005). Atualmente a enorme quantidade de imagens que estão sendo coletadas por sistemas sensores cada vez mais modernos e sofisticados requer o desenvolvimento de metodologias de classificação inovadoras, as quais possibilitem uma exploração automática e eficiente do grande volume de dados disponíveis nas imagens e, ao mesmo tempo, tornem o processo de mapeamento de características da superfície terrestre menos subjetivo e com maior potencial de repetição em situações subsequentes.

Na literatura existe uma grande variedade de propostas de algoritmos para a classificação de imagens de sensoriamento remoto. Diferentes técnicas provenientes de diversas áreas de pesquisa estão sendo utilizadas como base para o desenvolvimento desses algoritmos, entre elas: a estatística (DEAN *et al.*, 2003), redes neurais (GONÇALVES *et al.*, 1996), máquinas de vetores-suporte (PAL *et al.*, 2005), algoritmos genéticos (ALIXANDRINI *et al.*, 2002), lógica nebulosa (WANG, 1990) e árvores de decisão (HANSEN *et al.*, 1996).

Em sensoriamento remoto tradicionalmente as técnicas de classificação de imagens são agrupadas em três categorias gerais: supervisionada, não-supervisionada e híbrida (LILLESAND *et al.*, 2000). Particularmente, a classificação não-supervisionada baseia-se no princípio de que o algoritmo

computacional é capaz de identificar por si só as classes dentro de um conjunto de dados. Esse tipo de classificação é frequentemente realizado através de métodos de agrupamentos (*clustering*).

Embora exista uma grande quantidade de diferentes métodos de agrupamentos na área de reconhecimento de padrões (XU *et al.*, 2005), a maioria dos softwares ou sistemas computacionais voltados para o processamento digital de imagens de sensoriamento remoto realiza a classificação não-supervisionada baseada em métodos de agrupamentos particionais, como o K-médias e o ISODATA (BALL *et al.*, 1967).

Apesar de serem amplamente utilizados esses métodos particionais de agrupamentos apresentam diversas limitações. As funções objetivo usadas por eles partem do pressuposto de que o número de agrupamentos ou classes,  $K$ , é conhecido a priori. Na hipótese de se ter escolhido um valor  $K'$  inadequado o método irá impor, pelo uso de técnicas de otimização,  $K'$  agrupamentos aos dados. O usuário também deve especificar manualmente vários parâmetros para controlar o processo de agrupamento, entre eles: os centróides iniciais de cada agrupamento, o número máximo de iterações, limiares para realizar a divisão, fusão ou exclusão de agrupamentos. O K-médias e o ISODATA são muito sensíveis a esses parâmetros, podendo gerar partições diferentes quando são feitas várias simulações para um mesmo conjunto de dados. Outras limitações não menos importantes desses algoritmos particionais são: o alto custo computacional quando o conjunto de dados a ser analisado é muito grande (a cada iteração todos os pixels da imagem são comparados com todos os centros de agrupamentos) e a existência de suposições sobre a forma dos agrupamentos. Geralmente apenas um protótipo

(centróide) é utilizado para representar um agrupamento, portanto esses métodos tornam-se adequados apenas para analisar agrupamentos que apresentam formatos hiperesféricos (GONÇALVES *et al.*, 2005).

Outra forma possível, mas não usual, de realizar a classificação não-supervisionada de imagens de sensoriamento remoto é através de métodos hierárquicos de agrupamentos. Diferentemente dos métodos particionais, os métodos hierárquicos não exigem que o usuário especifique previamente o número de agrupamentos e outros parâmetros adicionais. Outra vantagem significativa desses métodos é a de possibilitar a visualização do resultado da classificação por meio de um dendrograma, que ilustra de forma hierárquica o grau de semelhança entre os agrupamentos que são formados por fusões (ou divisões) em cada estágio sucessivo da análise. Entretanto, os métodos hierárquicos apresentam algumas características que inviabilizam a sua aplicação na classificação de imagens de sensoriamento remoto: (a) em geral requerem espaço de memória de ordem  $O(N^2)$ , onde  $N$  é o número de registros do conjunto de dados; (b) os resultados podem ser difíceis de interpretar, principalmente para conjuntos de dados grandes; (c) para determinar qual a linha de “corte” do dendrograma há a necessidade de se aplicar algum critério de decisão (DUDA *et al.*, 2002). Conseqüentemente, na literatura é quase imperceptível a existência de trabalhos que aplicam métodos hierárquicos de agrupamentos na análise de imagens de sensoriamento remoto.

Diante desse cenário, o presente trabalho apresenta uma nova metodologia para a classificação não-supervisionada de imagens de sensoriamento remoto. Na metodologia proposta o Mapa Auto-organizável de Kohonen (SOM – *Self-Organizing Map*) é utilizado para mapear os padrões originais da imagem para uma grade de neurônios bidimensional. O objetivo é quantizar e representar os padrões da imagem em um espaço de menor dimensão, buscando preservar a distribuição de probabilidade e a topologia dos mesmos. Posteriormente, um método de agrupamento hierárquico aglomerativo é aplicado sobre a grade de neurônios do SOM já treinado, gerando um dendrograma de neurônios agrupados com diferentes graus de similaridade. Cada nível do dendrograma obtido corresponde a uma configuração diferente de agrupamentos de neurônios do SOM que pode ser utilizada para representar as classes sobre as quais a imagem original será classificada.

Além de viabilizar a aplicação de um método hierárquico, a idéia de representar a imagem através dos neurônios de um SOM possibilita a descoberta de agrupamentos com geometrias complexas e variadas, tendo em vista que as classes da imagem são representadas por grupos de neurônios e não apenas por um único protótipo. Além disso, procurando explorar todas as propriedades do SOM e reduzir ainda mais o volume de processamento do processo de classificação, a metodologia proposta emprega um algoritmo de

agrupamento hierárquico com conectividade restrita e índices de validação de agrupamentos calculados de maneira modificada para determinar o número ideal de classes da imagem.

Buscando garantir ainda uma maior eficiência no processo de classificação, o método proposto extrai as informações da imagem por meio de janelas de pixels, a fim de incorporar informações de contexto. Seguindo essa abordagem, antes de se aplicar o algoritmo de agrupamento hierárquico o método filtra protótipos heterogêneos do SOM, os quais representam padrões (janelas de pixels) correspondentes a regiões de transição entre diferentes classes de cobertura terrestre. Os padrões originais associados a esses protótipos heterogêneos são classificados de forma particular somente no final do processo, considerando as classes dos pixels vizinhos que já foram rotulados.

O restante do artigo está organizado da seguinte forma: a seção 2 descreve sucintamente o SOM, a seção 3 apresenta uma breve explanação sobre métodos de agrupamentos hierárquicos aglomerativos, enquanto a seção 4 apresenta a metodologia de classificação proposta. Resultados experimentais mostrando uma comparação do método proposto com o algoritmo K-médias são mostrados na seção 5, e a seção 6 apresenta as conclusões e considerações finais.

## 2. MAPA AUTO-ORGANIZÁVEL DE KOHONEN

O Mapa Auto-organizável de Kohonen (SOM – *Self-Organizing Map*) é um tipo de rede neural artificial baseada em aprendizado competitivo e não-supervisionado, i.e., nenhuma informação sobre as classes dos sinais de entrada é utilizada no processo de ajuste dos pesos sinápticos da rede (KOHONEN, 1997).

A rede consiste essencialmente de duas camadas: uma camada de entrada  $I$  e uma camada de saída  $U$  com neurônios geralmente dispostos em um arranjo topológico bidimensional. A entrada da rede corresponde a um vetor  $p$ -dimensional,  $\mathbf{x}$ , geralmente no espaço  $\mathcal{R}^p$ . Todas as  $p$  componentes do vetor de entrada alimentam cada um dos neurônios do mapa. Cada neurônio  $i$  pode ser representado por um vetor de pesos sinápticos  $\mathbf{w}_i = [w_{i1}, w_{i2}, \dots, w_{ip}]^T$ , também no espaço  $p$ -dimensional.

Para cada padrão de entrada  $\mathbf{x}$  um neurônio é escolhido o vencedor,  $c$ , usando o critério de maior similaridade:

$$\|\mathbf{x} - \mathbf{w}_c\| = \min_i \{\|\mathbf{x} - \mathbf{w}_i\|\} \quad (1)$$

onde  $\|\cdot\|$  representa a distância Euclidiana. Os pesos do neurônio vencedor, juntamente com os pesos dos seus neurônios vizinhos, são ajustados de acordo com a seguinte equação:

$$\mathbf{w}_i(t+1) = \mathbf{w}_i(t) + h_{ci}(t)[\mathbf{x}(t) - \mathbf{w}_i(t)] \quad (2)$$

onde  $t$  indica a iteração do processo de treinamento,  $\mathbf{x}(t)$  é o padrão de entrada e  $h_{ci}(t)$  é o núcleo de vizinhança ao redor do neurônio vencedor  $c$ . Este último termo é



uma função decrescente com o tempo e com a distância do neurônio  $i$  ao neurônio vencedor  $c$ , e geralmente corresponde ao produto de duas componentes: a taxa de aprendizado  $\alpha(t)$  e a função de vizinhança  $h(d,t)$ :

$$h_{ci}(t) = \alpha(t) \cdot h(\|r_c - r_i\|, t) \quad (3)$$

onde  $r_i$  e  $r_c$  são, respectivamente, as posições do neurônio  $i$  e do neurônio vencedor no arranjo topológico da rede.

Uma propriedade importante do SOM, derivada da quantização vetorial gerada pelo algoritmo, é a de aproximar o espaço de entrada buscando preservar a ordenação topológica e a densidade dos dados. Isso possibilita obter informações dos agrupamentos de dados originais analisando as relações geométricas dos neurônios do mapa treinado (HAYKIN, 1999).

### 3. MÉTODOS DE AGRUPAMENTOS HIERÁRQUICOS AGLOMERATIVOS

Mais comuns dentre os métodos de agrupamentos hierárquicos, as técnicas aglomerativas iniciam a análise dos dados considerando a existência de  $N$  agrupamentos, cada um deles contendo exatamente um único objeto do conjunto de dados. Sucessivas operações de fusão entre os agrupamentos são então efetuadas até que exista um único agrupamento contendo todos os  $N$  objetos. Os passos de um método hierárquico aglomerativo podem ser descritos da seguinte forma:

1. Inicie com  $N$  agrupamentos  $C_1, C_2, \dots, C_N$  cada um contendo exatamente um único objeto.
2. Determine o par de agrupamentos distintos  $(C_i, C_j)$ , de modo que:

$$D(C_i, C_j) = \min_{\substack{1 \leq m, l \leq N \\ m \neq l}} D(C_m, C_l) \quad (4)$$

onde  $D(*,*)$  é uma função de distância.

3. Forma-se um novo agrupamento  $C_k$  pela união dos agrupamentos  $C_i$  e  $C_j$ , i.e.,  $C_k = C_i \cup C_j$ .
4. Calculam-se as novas distâncias,  $D(C_k, C_l)$ , entre o novo agrupamento  $C_k$  e todos os outros restantes.
5. Repita os passos 2, 3 e 4 até que todos os objetos estejam em um único agrupamento.

Existem diversas variantes de métodos hierárquicos aglomerativos baseadas em diferentes formas de calcular a distância  $D(*,*)$  entre o agrupamento formado ( $C_k$ ) e todos os outros agrupamentos. Os métodos de agrupamentos hierárquicos aglomerativos mais simples e populares são o método de Ligações Simples (ou de vizinhos mais próximos) e o de Ligações Completas (XU *et al.*, 2005).

Os resultados de um método hierárquico aglomerativo são usualmente descritos por um diagrama de similaridade, chamado de dendrograma. O

dendrograma hierarquiza o grau de semelhança entre os agrupamentos tornando possível obter uma visão bidimensional da similaridade ou dissimilaridade de todos os dados. Esta propriedade facilita a comparação de diferentes configurações de agrupamentos para um mesmo conjunto de dados. Entretanto, para se determinar o nível apropriado do dendrograma ou o número ideal de agrupamentos algum critério de decisão deve ser empregado.

### 4. METODOLOGIA PROPOSTA

A metodologia proposta neste trabalho essencialmente busca explorar em conjunto as propriedades do SOM e dos métodos de agrupamentos hierárquicos aglomerativos para realizar a classificação não-supervisionada de imagens de sensoriamento remoto. A Fig. 1 ilustra de maneira simplificada o método proposto e a seguir são apresentadas as suas etapas principais:

1. *Amostragem*: amostras na forma de janelas de pixels são coletadas de maneira uniforme sobre toda a região da imagem e fornecidas como padrões de entrada para o SOM.
2. *Treinamento do SOM*: o treinamento não-supervisionado do SOM é realizado e um mapa de características da imagem composto pelos neurônios da rede (protótipos) é obtido.
3. *Filtragem de protótipos*: os protótipos do SOM que apresentam atividade nula são descartados e aqueles que apresentam alto grau de heterogeneidade espectral (protótipos heterogêneos) são excluídos do próximo passo.
4. *Análise hierárquica aglomerativa*: um método de agrupamento hierárquico aglomerativo com restrições de conectividade é aplicado sobre os protótipos do SOM já filtrado, gerando um dendrograma do mapa.
5. *Avaliação dos agrupamentos*: um índice de validação de agrupamentos calculado de forma modificada é aplicado em todos os níveis do dendrograma obtido no passo 4 para determinar o número ideal de agrupamentos.
6. *Classificação do SOM*: os protótipos do SOM são rotulados de acordo com as classes de protótipos determinadas no passo 5.
7. *Classificação da imagem*: os protótipos do SOM rotulados no passo 6 em conjunto com os protótipos heterogêneos (filtrados no passo 3) são utilizados como referência para classificar todos os pixels da imagem.
8. *Refinamento*: cada um dos pixels da imagem que foram associados a algum protótipo heterogêneo é reclassificado utilizando a classe do pixel vizinho que apresenta a menor distância (espectral) a ele.

Nas próximas subseções cada um dos passos da metodologia proposta é explicado com maiores detalhes.



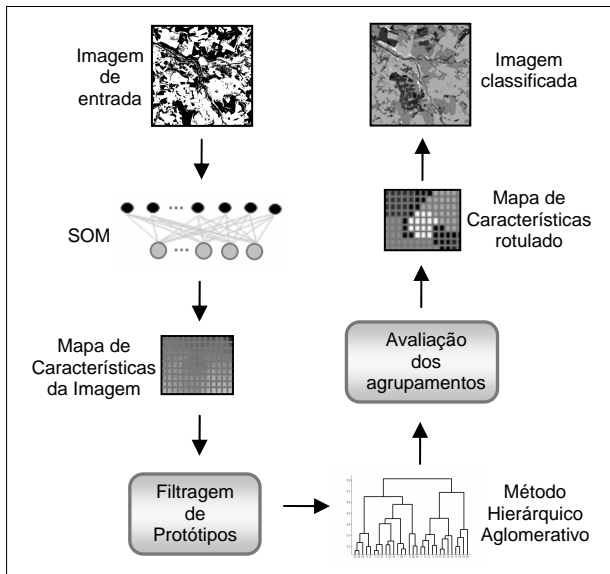


Fig. 1 – Ilustração da metodologia de classificação.

#### 4.1 Amostragem

O passo 1 da metodologia proposta consiste em coletar um conjunto de amostras da imagem para realizar o treinamento do SOM. Diferentemente de abordagens pixel a pixel que utilizam apenas a informação espectral de pontos individuais para encontrar regiões homogêneas, o presente trabalho realiza a amostragem da imagem através de janelas de pixels. A idéia é incorporar no processo de classificação informações de vizinhança (contexto), tendo em vista que pixels isolados não são capazes de representar a maioria dos padrões de cobertura terrestre, especialmente no caso de imagens que apresentam resoluções espaciais mais altas. Embora a estratégia de utilizar informações de pixels vizinhos apresente um maior custo computacional do que uma abordagem pixel a pixel, na literatura vários trabalhos têm mostrado que esse esforço é justificado por um aumento proporcional na acurácia da classificação (MAGNUSSEN *et al.*, 2004).

Outros dois pontos importantes relacionados com o processo de amostragem são a forma de extração das janelas amostrais e o tamanho das mesmas.

As janelas amostrais são coletadas de maneira uniforme sobre toda a região da imagem, sem sobreposições e em intervalos regulares (de no máximo 10 pixels). Todas as amostras são quadradas e possuem o mesmo tamanho.

A determinação do tamanho das amostras não é uma tarefa simples de ser automatizada. Dependendo das características da imagem a ser classificada, janelas muito pequenas podem não capturar o padrão particular das classes, enquanto janelas maiores podem incluir pixels de mais do que uma classe. No entanto, partindo de um tamanho não inferior a 5x5, o usuário possui dentro da metodologia proposta um certo grau de flexibilidade para definir a dimensão das janelas amostrais, não sendo portanto essa tarefa um ponto crítico do processo. Caso sejam capturadas janelas

amostrais que incorporem mais de uma classe, o método proposto (através dos passos 3 e 8) trata de maneira diferenciada os protótipos do SOM que representam essas janelas heterogêneas. Além disso, protótipos que possuem mais de uma classe podem funcionar como bordas no mapa de características do SOM contribuindo para a separação dos agrupamentos.

#### 4.2 Treinamento do SOM

Para realizar o treinamento do SOM é necessário especificar alguns parâmetros que definem a estrutura do mapa e que controlam o treinamento propriamente dito. Com o objetivo de garantir um bom mapeamento dos padrões originais, a metodologia proposta define de modo particular os parâmetros da rede baseando-se na literatura existente, em testes experimentais e em algumas peculiaridades da aplicação do SOM em imagens de sensores remotos. Todavia, alternativas também podem ser buscadas para se obter bons mapas. Uma boa revisão sobre os parâmetros de treinamento do SOM pode ser obtida em COSTA, (1999).

A metodologia proposta utiliza os seguintes parâmetros para realizar o treinamento da rede:

- Inicialização dos pesos: linear
- Apresentação dos padrões: em lote
- Número de épocas de treinamento: 500
- Tipo de função de vizinhança: gaussiana
- Formato do arranjo: retangular.

O tamanho do mapa é um dos parâmetros livres do SOM que depende particularmente da imagem de entrada. O principal propósito para escolher um tamanho apropriado para o SOM é realizar um bom mapeamento dos padrões da imagem no arranjo topológico de neurônios. Contudo, o desempenho da metodologia de classificação proposta não é significativamente afetado se tamanhos suficientemente grandes para o SOM forem utilizados. Embora mapas com dimensões maiores do que a necessária apresentem uma quantidade maior de neurônios inativos, i.e., neurônios que não possuem nenhum padrão de entrada associado a eles, este evento não é prejudicial dentro da metodologia proposta. Assim como o método trata de maneira adequada protótipos heterogêneos, o mesmo acontece com os protótipos inativos.

No final da etapa de treinamento do SOM tem-se, portanto, um conjunto de protótipos (neurônios) que representam todos os padrões de entrada coletados a partir da imagem original.

##### 4.2.1 Mapa de Características da Imagem

A metodologia proposta possui uma particularidade em relação ao mapeamento da imagem através do SOM que deve ser destacada. Uma vez que os neurônios do SOM são dispostos espacialmente na forma de uma grade retangular, e considerando ainda

que nesta rede neural os vetores de pesos de cada neurônio possuem as mesmas dimensões que os padrões de entrada (que neste caso são janelas de pixels), torna-se possível gerar uma imagem da grade de neurônios do mapa. Conforme mostra a Fig. 2, os neurônios ou protótipos do SOM (através dos seus vetores de pesos sinápticos) correspondem a janelas de pixels, o que permite a visualização dos mesmos na forma de uma imagem. Essa representação visual da grade de neurônios do SOM, após o seu treinamento não supervisionado, foi denominada aqui de *Mapa de Características da Imagem* (MCI).

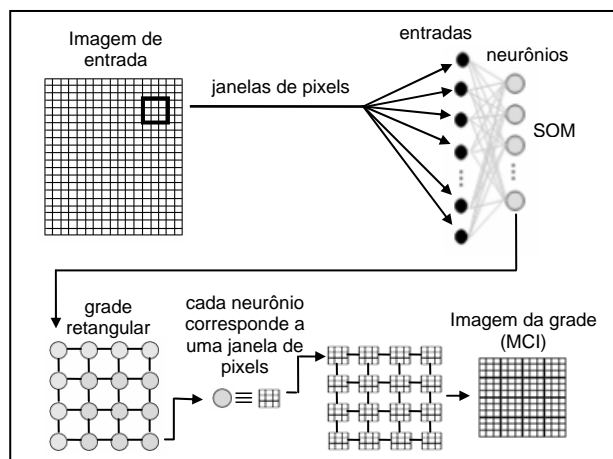


Fig. 2 – Ilustração do processo de construção do Mapa de Características da Imagem.

Conforme será apresentado nos resultados experimentais (seção 5), através do MCI é possível visualizar o mapeamento dos padrões da imagem de entrada realizado pelo SOM e suas respectivas propriedades.

### 4.3 Filtragem de protótipos

O passo 3 da metodologia proposta consiste em filtrar dois tipos de protótipos que geralmente aparecem no mapeamento dos padrões da imagem através do SOM. Esses protótipos são chamados aqui de *inativos* e *heterogêneos*.

Os protótipos inativos correspondem àqueles neurônios que apresentam atividade nula no processo de aprendizado competitivo do SOM, e por isso não possuem nenhum padrão de entrada associado a eles. Esses protótipos são simplesmente descartados do processo não passando para as próximas etapas da análise.

Os protótipos heterogêneos são aqueles que apresentam alto grau de heterogeneidade espectral e estão associados normalmente a padrões de entrada que possuem mais do que uma classe de cobertura terrestre. Na maioria das vezes esses padrões correspondem a regiões de transição entre classes de cobertura terrestre presentes na imagem e são capturados em decorrência da amostragem realizada por janelas de pixels. Os protótipos considerados heterogêneos são excluídos do passo 4 da análise, onde é aplicado o método de

agrupamento hierárquico. Essa exclusão é realizada com o objetivo de não permitir que esses protótipos e, conseqüentemente, os padrões associados a eles sejam atribuídos erroneamente a uma das classes que os compõe. Os protótipos heterogêneos são reincorporados apenas nos procedimentos adotados nos passos 7 e 8 do método proposto.

Para medir o grau de heterogeneidade de cada protótipo é aplicado um índice, denominado aqui de *Índice de Heterogeneidade Espectral* (IHE), definido da seguinte forma:

$$IHE = \frac{1}{M} \sum_{i=1}^M CV_i \quad (5)$$

onde  $M$  é o número de bandas espectrais da imagem consideradas na análise e  $CV_i$  é o coeficiente de variação (dispersão relativa) dos níveis de cinza de todos os pixels na banda espectral  $i$  do protótipo. Os protótipos cujos IHE's satisfazem a relação dada abaixo são considerados heterogêneos e, conseqüentemente, serão filtrados:

$$IHE > \mu_{IHE} + \frac{1}{2} \sigma_{IHE} \quad (6)$$

$\mu_{IHE}$  e  $\sigma_{IHE}$  são, respectivamente, a média e o desvio padrão dos IHE's de todos os protótipos.

O tempo de processamento consumido por esta etapa de filtragem de protótipos é compensado não apenas pela redução do número de protótipos que serão analisados no passo 4 do método, mas principalmente por um possível aumento na precisão de classificação da imagem. Protótipos inativos e heterogêneos podem funcionar como "unidades de interpolação" no mapa de neurônios facilitando a separação das classes de agrupamentos.

### 4.4 Análise Hierárquica Aglomerativa

No quarto passo da metodologia proposta, um método de agrupamento hierárquico aglomerativo é aplicado sobre o conjunto de protótipos do SOM já treinado e filtrado. O objetivo desse passo é agrupar protótipos e conjuntos de protótipos em diferentes níveis de similaridade buscando descobrir a estrutura (ou classes) dos mesmos. O método de Ligações Simples (ou dos vizinhos mais próximos) foi utilizado nesta abordagem, porém outros métodos hierárquicos podem ser empregados.

Uma característica importante da estratégia de análise hierárquica apresentada neste trabalho é a imposição de restrições às possíveis fusões dos protótipos do SOM. Um algoritmo hierárquico aglomerativo uma vez aplicado sobre os agrupamentos de neurônios do SOM deve respeitar as relações topológicas do espaço de saída da rede. Diferentemente da abordagem tradicional de métodos de agrupamentos hierárquicos, que consiste em comparar todos os pares de objetos ou grupos de objetos para decidir sobre uma fusão, a abordagem utilizada neste trabalho verifica a

possibilidade de efetuar fusões somente entre pares de neurônios ou grupos de neurônios que sejam adjacentes (ou vizinhos) na grade do SOM.

No final desta etapa tem-se um dendrograma que mostra de maneira hierárquica a relação de semelhança entre os protótipos e agrupamentos de protótipos do SOM. Lembrando que os protótipos do SOM representam todos os padrões de entrada, o dendrograma gerado reflete, portanto, as relações existentes entre os dados e agrupamentos de dados da imagem original, podendo ser visto então como um dendrograma simplificado da imagem.

#### 4.5 Avaliação dos agrupamentos

Em cada nível do dendrograma obtido no passo anterior há uma configuração diferente de agrupamentos para os protótipos do SOM, e conseqüentemente, para o conjunto de dados da imagem. Portanto, como em qualquer método de agrupamento hierárquico, há a necessidade de se aplicar algum critério para verificar qual nível do dendrograma representa a configuração ideal de agrupamentos (ou o número ideal de classes).

Uma saída usual é a aplicação de índices de validação de agrupamentos que incorporem medidas de dispersão intra e inter-agrupamentos (MAULIK *et al.*, 2002). A estratégia consiste em aplicar o índice de validação em todos os níveis do dendrograma e escolher aquele nível em que o índice apresenta o valor ótimo dentre todos os valores obtidos.

Entretanto, deve-se observar que a implementação da maioria dos índices de validação de agrupamentos exige um alto custo computacional. Devido a isso, quando o número de agrupamentos e o volume do conjunto de dados são muito grandes a utilização desses índices torna-se proibitiva. As imagens de sensoriamento remoto são um exemplo típico. Conforme descrito em JI (2003), índices de validação são freqüentemente ignorados em aplicações de sensoriamento remoto e por isso não estão disponíveis na maioria dos pacotes de processamento de imagens.

No entanto, como na metodologia proposta o dendrograma obtido no passo anterior não é construído em função dos padrões originais da imagem, e sim, por um pequeno conjunto de neurônios do SOM que os representam, a aplicação de um índice de validação de agrupamentos em cada nível do dendrograma se torna viável, pois a quantidade de dados a ser avaliada neste caso é bem inferior a do conjunto de padrões extraídos diretamente da imagem.

Embora o SOM realize uma boa aproximação para o espaço de entrada, é certo que a estratégia de aplicar um índice de validação sobre os protótipos da rede, ao invés de aplicá-lo diretamente sobre os padrões originais da imagem, poderá causar alguma diferença numérica nos resultados. Para diminuir possíveis erros de aproximação, o método proposto modifica a forma de calcular os índices de validação. Os índices são calculados utilizando não apenas os vetores de pesos

dos protótipos do SOM, mas também o nível de atividade de cada um deles. O nível de atividade de um protótipo é igual ao número de padrões de entrada que estão associados a ele pelo processo de mapeamento do SOM.

Para exemplificar a modificação aplicada nos cálculos dos índices considere a fórmula descrita na equação (7). Esta fórmula e variantes da mesma são comumente utilizadas em alguns índices de validação de agrupamentos para calcular a distância ( $\delta_{ij}$ ) entre dois agrupamentos de dados  $C_i$  e  $C_j$  (BEZDEK *et al.*, 1998).

$$\delta_{ij} = \frac{1}{|C_i||C_j|} \sum_{x \in C_i, y \in C_j} d(x, y) \quad (7)$$

Na fórmula acima,  $d(*,*)$  é uma medida de distância, e  $|C_i|$  e  $|C_j|$  representam o número de pontos nos agrupamentos de dados  $C_i$  e  $C_j$ , respectivamente. Como se pode notar,  $\delta_{ij}$  é determinada basicamente através dos cálculos das distâncias entre todos os pares de pontos  $x$  e  $y$  pertencentes aos agrupamentos de dados  $C_i$  e  $C_j$ . Se  $|C_i|$  e  $|C_j|$  são muito grandes o cálculo de  $\delta_{ij}$  torna-se custoso demais.

Aplicando a estratégia apresentada aqui, o cálculo da distância entre os agrupamentos  $C_i$  e  $C_j$ , equivalente ao apresentado na equação (7), fica formulado da seguinte maneira:

$$\delta_{ij}^{SOM} = \frac{1}{|C_i||C_j|} \sum_{w_i \in W_i, w_j \in W_j} h(w_i) \cdot h(w_j) \cdot d(w_i, w_j) \quad (8)$$

onde  $W_i$  e  $W_j$  referem-se a quantidade de protótipos dos mesmos,  $d(*,*)$  é a mesma medida de distância usada na equação (7),  $h(w_i)$  é o nível de atividade do protótipo  $w_i$  pertencente a  $W_i$  e  $h(w_j)$  é o nível de atividade do protótipo  $w_j$  pertencente a  $W_j$ .

Enquanto  $\delta_{ij}$  se baseia nas distâncias entre todos os pares de pontos pertencentes a  $C_i$  e  $C_j$ ,  $\delta_{ij}^{SOM}$  se baseia nas distâncias apenas entre os pares de pontos pertencentes a  $W_i$  e  $W_j$ . Logo, nota-se que  $\delta_{ij}^{SOM}$  é um cálculo aproximado de  $\delta_{ij}$ , porém com um custo bem menor de processamento, pois as quantidades  $|W_i|$  e  $|W_j|$  geralmente são bem menores que  $|C_i|$  e  $|C_j|$ . A inclusão dos níveis de atividade  $h(.)$  dos protótipos no cálculo de  $\delta_{ij}^{SOM}$  ajuda a diminuir o efeito dos erros de quantização decorrentes do mapeamento realizado pelo SOM.

Um outro ponto importante desta etapa do método proposto é a escolha do índice de validação que será aplicado de forma modificada em todos os níveis do dendrograma. Na metodologia proposta neste trabalho foi aplicado o índice de validação CDbw (*Composing Density Between and Within Clusters*) proposto em HALKIDI *et al.* (2002). Além de se basear em dois conceitos importantes, a densidade intra-agrupamento e a separação entre os agrupamentos, este índice foi escolhido por apresentar características que o

tornam capaz de avaliar de maneira adequada agrupamentos de dados que tenham formatos arbitrários e complexos, o que não acontece com a maioria dos outros índices de validação existentes na literatura.

#### 4.6 Classificação do SOM

Uma vez definido o nível do dendrograma que apresenta a melhor configuração de agrupamentos para os protótipos do SOM e, portanto, para os padrões da imagem, o passo 6 do método proposto consiste apenas em rotular esses protótipos. Os protótipos pertencentes a um mesmo agrupamento recebem um rótulo particular e correspondem a uma determinada classe de cobertura terrestre presente na imagem.

De acordo com a metodologia proposta, cada classe de cobertura terrestre descoberta será representada então por um grupo de protótipos do SOM e não apenas por um protótipo individual.

#### 4.7 Classificação da Imagem

Para realizar a classificação da imagem, os protótipos heterogêneos que foram filtrados no passo 4 são agora reincorporados e considerados também como uma classe. Para isso, todos eles recebem um mesmo rótulo, que obviamente deve ser diferente daqueles usados para rotular as outras classes de protótipos no passo anterior.

Janelas de pixels da imagem com dimensões iguais a das amostras de treinamento são comparadas com todos os protótipos do SOM. Essa comparação é realizada através das distâncias calculadas entre a janela de pixels considerada e cada um dos protótipos. O pixel central da janela de pixels recebe o rótulo do protótipo que apresenta a menor distância a ela. A imagem é então inteiramente percorrida até que todos os pixels sejam classificados dessa forma.

#### 4.8 Refinamento

No último passo do método proposto, o resultado do processo de classificação realizado na etapa anterior é aprimorado. Os pixels da imagem que receberam o rótulo da classe de protótipos heterogêneos são reclassificados. Cada um desses pixels é comparado com os seus pixels vizinhos na imagem que não foram rotulados pela classe de protótipos heterogêneos. Eles recebem um novo rótulo que será igual ao do vizinho que apresenta a distância mínima a ele em termos de atributos espectrais.

### 5. RESULTADOS EXPERIMENTAIS

Esta seção apresenta um exemplo de aplicação da metodologia proposta sobre uma imagem teste. Os resultados são comparados com aqueles obtidos a partir da aplicação do algoritmo K-médias sobre a mesma imagem.

Todos os experimentos apresentados nesta

seção foram realizados em uma máquina com processador AMD Athlon™ XP 2600+ 1,91GHZ e 1GB de memória RAM.

A Fig. 3 mostra uma composição colorida da imagem utilizada nos testes. A imagem é composta pelas bandas espectrais 3, 4 e 5 do satélite Landsat-5 (TM), possui 496x512 pixels e mostra a cidade de Manaus e o encontro dos Rios Negro e Solimões. A cena apresenta 4 grandes classes de cobertura terrestre: área urbana, vegetação e dois padrões de água (um mais escuro correspondente ao Rio Negro e outro de cor mais arroxeada correspondente ao Rio Solimões). Pequenas áreas com vegetação rala ou desmatada também aparecem na cena, mas com uma densidade muito baixa. Esta imagem foi cedida pelo Instituto Nacional de Pesquisas Espaciais (INPE).



Fig. 3 – Composição colorida da imagem teste.

Aplicando a metodologia proposta foi realizado inicialmente o processo de amostragem da cena. Janelas amostrais de tamanho 5x5 foram coletadas de maneira uniforme sobre toda a região da imagem, sem sobreposições e em intervalos regulares de 10 pixels, totalizando cerca de 2500 amostras obtidas sem a intervenção do usuário.

Um SOM composto de 100 neurônios dispostos em uma grade retangular de dimensão 10x10 foi treinado com todas as amostras coletadas anteriormente. Os demais parâmetros do SOM foram definidos de acordo com as especificações apresentadas na seção 4.2.

A Fig. 4 ilustra o MCI obtido após o treinamento do SOM. Conforme descrito na seção 4.2.1, o MCI é uma imagem gerada a partir da grade de neurônios do SOM já treinado.

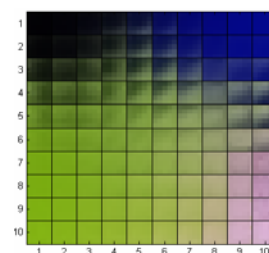


Fig. 4 – MCI obtido após o treinamento do SOM.

Cada pequeno quadrado no MCI corresponde à

imagem de um protótipo (ou neurônio) do SOM. Através deles é possível visualizar as propriedades do mapeamento realizado pela rede neural. As quatro grandes classes de cobertura terrestre presentes na imagem original aparecem em forma de agrupamentos no MCI. No canto superior esquerdo se encontram os protótipos correspondentes ao padrão de água mais escuro, no canto superior direito estão presentes os protótipos referentes à outra classe de água (mais arroxeada), no canto inferior direito (em rosa) se encontra aqueles que correspondem à área urbana, e do canto inferior esquerdo ao centro do MCI estão os protótipos associados à classe vegetação. Se observarmos a imagem original, verifica-se que a classe vegetação é aquela que apresenta a maior área de ocupação na cena. Sendo assim, como o SOM reflete a distribuição de probabilidade dos dados de entrada, os protótipos associados à classe vegetação (em verde no MCI) se apresentam em maior número do que os protótipos das outras classes. A propriedade de ordenação topológica do SOM também pode ser notada. Os protótipos das duas classes de água presentes na imagem original aparecem próximos um dos outros na parte superior do MCI, isto porque os dados de entrada correspondentes a essas duas classes possuem uma maior similaridade espectral em comparação com os atributos espectrais das outras duas classes de cobertura.

Após o treinamento do SOM, a próxima etapa a ser executada é a filtragem dos protótipos inativos e heterogêneos. A Fig. 5a mostra o nível de ativação de cada um dos protótipos do SOM, i.e., o número de padrões de entrada (amostras) que estão associados a cada um deles através do mapeamento efetuado pelo SOM. Observa-se que os protótipos localizados nas posições da grade (1,5), (3,2), (4,8) e (6,10) apresentam níveis de ativação nulos e, portanto, não possuem nenhum padrão de entrada associado a eles. De acordo com o método proposto, esses protótipos são simplesmente descartados do processo não passando para as próximas etapas da análise.

A Fig. 5b apresenta os valores (em porcentagem) dos IHE's de cada protótipo. Conforme apresentado na seção 4.3, o IHE mede o grau de heterogeneidade espectral dos protótipos do SOM. Os protótipos com IHE's cujos valores satisfazem o critério dado em (6) são considerados heterogêneos e devem ser filtrados. Desse modo, como os valores obtidos nesse experimento para a média ( $\mu_{IHE}$ ) e para o desvio-padrão ( $\sigma_{IHE}$ ) dos IHE's foram, respectivamente, 11.89 e 8.93, os protótipos que apresentam valores de IHE's acima de 16.35 (limiar obtido de acordo com o critério dado em (6)) são então considerados heterogêneos e por isso são filtrados. Esses protótipos (marcados em negrito na Fig. 5b) serão reconsiderados somente nos dois últimos passos da metodologia de classificação.

Sendo assim, do total de 100 protótipos do SOM, 25 deles (entre inativos e heterogêneos) foram filtrados, restando 75 protótipos para serem analisados nas duas próximas etapas.

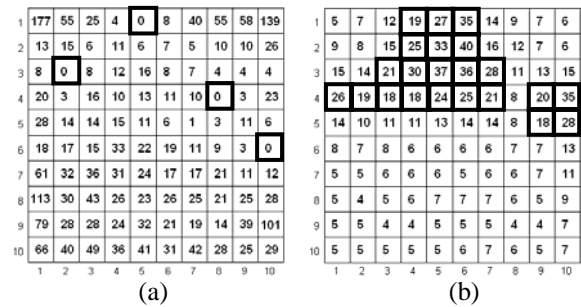


Fig. 5 – (a) Níveis de ativação dos protótipos do SOM (protótipos com níveis de ativação nulos estão marcados em negrito). (b) Valores dos IHE's dos protótipos do SOM (protótipos considerados heterogêneos estão marcados em negrito).

De acordo com o método proposto, a etapa seguinte consiste em aplicar o método hierárquico aglomerativo sobre os protótipos do SOM filtrado. Conforme descrito na seção 4.4, o método das Ligações Simples é utilizado nesta abordagem e as fusões entre os protótipos (ou grupos de protótipos) são efetuadas respeitando a vizinhança topológica dos mesmos na grade retangular do SOM. O dendrograma possui 75 níveis, cada um deles apresentando uma configuração diferente de agrupamentos de protótipos do SOM que, por conseguinte, produz uma classificação diferente para a imagem original.

Além de possibilitar a classificação da imagem em diferentes níveis de agrupamentos, o dendrograma obtido permite que o usuário analise de maneira particular cada uma das uniões realizadas entre os protótipos ou grupos de protótipos, e também a hierarquia das mesmas. Se ressaltarmos ainda que no método proposto o usuário pode visualizar a imagem de cada protótipo do SOM através do MCI, a análise do dendrograma se torna muito mais clara e eficiente, facilitando a compreensão das relações existentes entre os diferentes padrões presentes na imagem.

É importante observar que esse tipo de análise somente é possível devido ao mapeamento realizado pelo SOM. A rede neural “resume” os padrões originais da imagem por meio de um pequeno conjunto de protótipos viabilizando assim a construção de um dendrograma simplificado. Se porventura o SOM não fosse utilizado para representar os dados de entrada, a construção do dendrograma ou a interpretação do mesmo se tornaria impraticável, pois o mesmo teria 2500 níveis de agrupamentos caso fossem consideradas todas as amostras coletadas da imagem.

Dando continuidade na execução das etapas do método proposto, o índice de validação de agrupamentos CDbw, calculado de forma modificada, é aplicado em todos os níveis do dendrograma para determinar qual a configuração de agrupamentos ideal dentre as 75 obtidas. Com o objetivo de comparar as duas formas de cálculo, a Fig. 6 mostra os valores do índice de validação em suas versões modificada e original. O gráfico apresenta somente os valores entre os níveis 61 e 74 do dendrograma. Para os níveis

inferiores a 61 os valores do índice são menores do que aqueles apresentados no gráfico, ou até mesmo nulos, devido à existência de níveis que possuem agrupamentos com apenas um único protótipo. O número de agrupamentos ou de classes do nível 61 ao 74 decresce, respectivamente, de 15 até 2.

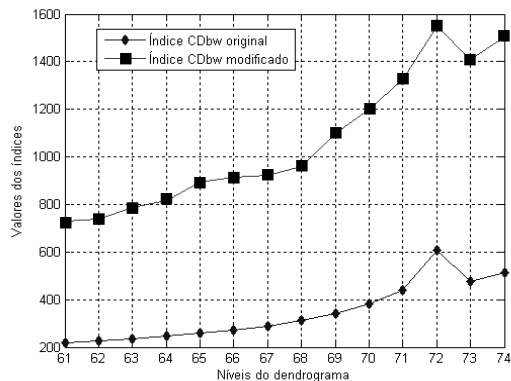


Fig. 6 – Valores do índice CDbw nas suas formas modificada e original.

Como esperado, a estratégia de modificar a computação de índices de validação de agrupamentos (descrita na seção 4.5) apresenta erros de aproximação, porém a variação dos valores é semelhante ao da versão original, não comprometendo a tomada de decisão sobre o número de agrupamentos ideal dos dados. A vantagem principal em usar a versão modificada do índice está no seu tempo de processamento. Neste experimento, enquanto o tempo gasto para calcular o índice CDbw na forma original (para todo o dendrograma) foi de 784 segundos, a versão modificada exigiu apenas 32 segundos. É importante observar que esta diferença significativa ocorre devido ao volume de dados considerados por cada uma das formas de cálculo do índice. O cálculo do índice CDbw na sua forma original foi executado em cada nível do dendrograma considerando as 2500 amostras coletadas da imagem, enquanto o cálculo da versão modificada considera apenas os 100 protótipos do SOM e os seus respectivos níveis de ativação.

Para o índice CDbw, quanto maior o seu valor melhor é a configuração de agrupamentos avaliada. Sendo assim, dentre todos os níveis do dendrograma obtido neste experimento o nível 72 é aquele que apresenta o maior valor para o índice (conforme mostra o gráfico da Fig. 6), e portanto, o melhor conjunto de agrupamentos para os protótipos do SOM. A Fig. 7 mostra o MCI classificado de acordo com as classes de agrupamentos do nível 72. Comparando com a imagem do MCI na Fig. 4, nota-se que os protótipos do SOM são agrupados em 4 classes, correspondendo adequadamente às 4 grandes classes de cobertura terrestre presentes na imagem. Os quadrados marcados com um “X” no MCI classificado são os protótipos descartados na etapa de filtragem do método proposto. É interessante notar que tais protótipos funcionam como “unidades de interpolação” e certamente contribuíram para o processo de separação da maioria das classes.

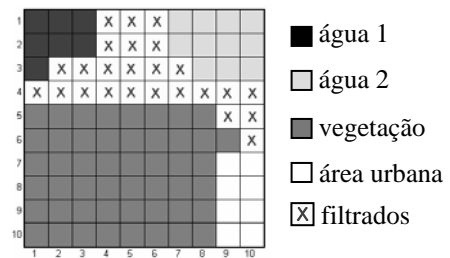


Fig. 7 – MCI classificado de acordo com a melhor configuração de agrupamentos do dendrograma.

Terminada a etapa de classificação dos protótipos do SOM, os dois últimos passos do método concentram-se na classificação de todos os pontos da imagem original. Para isso, os protótipos heterogêneos (filtrados anteriormente) voltam a ser incorporados no processo e considerados também como uma classe. Deste modo, no penúltimo passo do método proposto, 5 classes de protótipos do SOM são utilizadas como referência para realizar a classificação da imagem, as 4 classes do nível escolhido no dendrograma (apresentadas na Fig. 7) e a classe dos protótipos heterogêneos. A imagem é percorrida inteiramente considerando janelas de pixels com dimensões 5x5 (iguais aos das janelas amostrais) e comparado-as com todos os protótipos do SOM. O pixel central da janela de pixels recebe o rótulo do protótipo que apresenta a menor distância a ela.

Por fim, a última etapa do método, denominada aqui de “refinamento”, é aplicada. Um total de 26585 pixels da imagem foi associado à classe de protótipos heterogêneos. Esses pixels são reclassificados e recebem um novo rótulo, igual ao do pixel vizinho que apresenta a menor distância a ele (em termos de atributos espectrais). A Fig. 8 mostra o resultado da classificação da imagem original pela metodologia proposta (as 4 classes são representadas pelas mesmas cores utilizadas na Fig. 7).

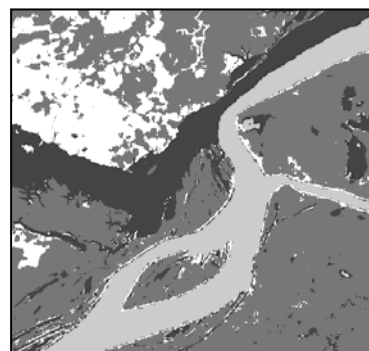


Fig. 8 – Resultado da classificação da imagem teste pela metodologia de classificação proposta.

Portanto, de acordo com a metodologia proposta, a imagem teste foi classificada em 4 classes, as quais correspondem aos 4 padrões de cobertura terrestre de maior predominância na cena. Algumas pequenas áreas de vegetação rala ou desmatada presentes na imagem foram classificadas como área

urbana, por serem espectralmente muito semelhantes aos padrões dessa classe e por se apresentarem com baixa densidade em comparação com os outros padrões de cobertura.

A Fig. 9 mostra o resultado da classificação da imagem teste pelo algoritmo K-médias considerando o número de classes igual a 4. O algoritmo foi executado usando cinco inicializações diferentes para os centróides de cada uma das classes e um número máximo de iterações igual a 100.

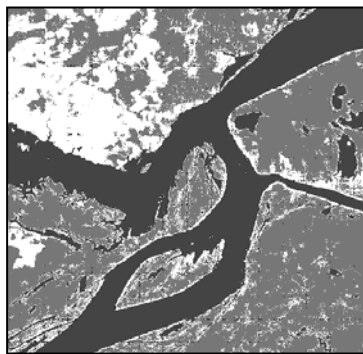


Fig. 9 – Resultado da classificação da imagem teste pelo algoritmo K-médias considerando  $k=4$ .

Diferentemente do método proposto, o algoritmo K-médias não discriminou um dos padrões de água predominantes na imagem e classificou as áreas de vegetação da cena em duas outras categorias. O algoritmo também confundiu áreas de vegetação rala ou desmatada com padrões de área urbana, porém de forma bem mais acentuada que o método proposto aqui. Os dois padrões de água existentes na imagem teste somente são discriminados pelo algoritmo K-médias se considerarmos um número de agrupamentos maior ou igual a 5. No entanto, embora o algoritmo classifique a vegetação em várias categorias diferentes, a confusão entre as áreas de vegetação rala com os padrões de área urbana permanece para qualquer número de classes.

Realizando uma análise visual das imagens classificadas pelos dois métodos (Figs. 8 e 9), observa-se que a imagem resultante da metodologia proposta apresenta um aspecto visual melhor. Enquanto a classificação pelo algoritmo K-médias possui uma aparência mais salpicada, a imagem produzida pelo método proposto se apresenta de forma mais homogênea em todas as áreas classificadas. Essa diferença no aspecto visual certamente se dá em virtude das formas pelas quais os dois métodos tratam a imagem. Enquanto o algoritmo K-médias utiliza uma abordagem pixel a pixel para realizar a classificação, o método proposto trabalha com janelas de pixels, as quais incorporam informações de vizinhança e por isso permitem obter uma classificação resultante de melhor qualidade.

Procurando realizar uma análise mais criteriosa e menos subjetiva da classificação obtida pela metodologia proposta, e considerando a ausência de verdade terrestre para a imagem teste, o presente trabalho realiza a classificação da imagem de modo

supervisionado e considera os resultados da mesma como referência (ou “verdade”). A classificação supervisionada, considerando as 4 classes predominantes na imagem, foi feita através de uma Rede Neural de *Perceptrons* Multicamadas com o algoritmo de treinamento *Backpropagation*. Essa classe de redes neurais tem sido amplamente empregada para realizar a classificação supervisionada de imagens de sensoriamento remoto (GONÇALVES, 1997).

A matriz de confusão e o índice de concordância *Kappa* foram calculados a partir da comparação entre a imagem classificada pelo método proposto e a imagem referência (resultante da classificação supervisionada). Analisando a matriz de confusão apresentada na Tabela 1 e considerando que o valor alcançado para o índice de concordância *Kappa* foi igual a 0.96, podemos considerar que o resultado da classificação da imagem teste pelo método apresentado neste trabalho foi bastante satisfatória.

TABELA 1 – MATRIZ DE CONFUSÃO DAS CLASSES PARA A IMAGEM CLASSIFICADA PELO MÉTODO PROPOSTO (%).

Classes	água1	agua2	veg.	urbana	Total
água1	96.43	0.40	0.34	0.20	16.64
água2	0.61	99.35	0.04	0.09	15.19
veg.	2.89	0.02	98.89	6.73	53.29
urbana	0.06	0.23	0.73	92.98	14.88
Total	100	100	100	100	100

Por fim, a Tabela 2 mostra o tempo de processamento consumido (em segundos) neste experimento por cada uma das etapas da metodologia de classificação proposta.

TABELA 2 – TEMPO CONSUMIDO PELO MÉTODO DE CLASSIFICAÇÃO PROPOSTO.

Etapas do método	Tempo consumido
Amostragem	10
Treinamento do SOM	40
Filtragem de protótipos	2
Análise hierárquica	0.5
Avaliação	32
Classificação do SOM	0.4
Classificação da imagem	66
Refinamento	4
Total	154.9

O tempo total gasto pelo método para realizar a classificação da imagem foi de aproximadamente 155 segundos. O treinamento do SOM, a avaliação dos agrupamentos e a classificação final da imagem foram as etapas que consumiram mais tempo de processamento.

Já o tempo consumido pelo algoritmo K-médias na classificação da mesma imagem (considerando  $k$  igual a 4) foi de aproximadamente 93 segundos, realizando cinco inicializações diferentes para os centróides das classes (como feito nos experimentos).

Embora o algoritmo K-médias tenha apresentado um tempo de processamento inferior ao da metodologia proposta, os dois métodos possuem princípios e características muito diferentes, que devem ser consideradas na avaliação de suas complexidades de cálculo.

Uma diferença importante entre os dois métodos está na quantidade de classes analisada por eles para realizar a classificação da imagem. O método proposto avalia diferentes configurações de agrupamentos para os dados, enquanto que o algoritmo K-médias realiza a classificação da cena apenas para uma única quantidade de classes definida a priori. O tempo gasto pelo método proposto para avaliar 75 configurações diferentes de agrupamentos foi de 32 segundos. Se resolvêssemos utilizar uma estratégia semelhante para o algoritmo K-médias, ou seja, realizar diferentes classificações da imagem variando o número de classes ( $k$ ) e depois aplicar o índice de validação CDbw para decidir qual delas é a melhor, o tempo de processamento seria tão elevado ao ponto de tal estratégia ser considerada impraticável. Basta levarmos em conta o tempo gasto na aplicação do índice CDbw para avaliar a classificação produzida pelo algoritmo K-médias apresentada na Fig. 9. O tempo consumido foi de 1890 segundos, muitíssimo superior ao tempo gasto pelo método proposto para avaliar diversas configurações de agrupamentos para imagem.

Diante dessas observações e considerando ainda os possíveis benefícios das outras técnicas e procedimentos empregados na metodologia proposta (como a filtragem de protótipos, a análise hierárquica e o refinamento da classificação), conclui-se que o tempo de processamento consumido pelo método apresentado é perfeitamente admissível, o que ressalta ainda mais a viabilidade de aplicação do mesmo.

## 6. CONCLUSÕES E CONSIDERAÇÕES FINAIS

Neste trabalho foi apresentada uma nova metodologia para a classificação não-supervisionada de imagens de sensoriamento remoto. O ponto chave do método proposto é realizar a análise de agrupamentos da imagem através de um conjunto de protótipos do SOM ao invés de trabalhar diretamente com os padrões originais da cena. Essa abordagem reduz significativamente a complexidade da análise tornando possível a utilização de técnicas que são normalmente consideradas inviáveis para o processamento de imagens de sensoriamento remoto, como por exemplo, métodos de agrupamentos hierárquicos e índices de validação de agrupamentos.

O método proposto apresenta uma série de vantagens e potencialidades que o colocam como uma alternativa bastante diferenciada para a classificação não-supervisionada de imagens. Entre elas, podemos destacar:

- A não exigência de uma definição a priori do número de classes para realizar a classificação da

imagem, o que não acontece na maioria dos métodos convencionais de classificação não-supervisionada;

- O método possui apenas dois parâmetros que devem ser definidos pelo usuário (o tamanho das amostras e do SOM), e mesmo assim é bastante robusto quanto à escolha dos mesmos;

- O simples uso de janelas de pixels permite incorporar informações de contexto e textura sem nenhum cálculo explícito de medida. Essa abordagem contribui para a qualidade da classificação resultante;

- A utilização de um método de agrupamento hierárquico aglomerativo permite que o usuário compreenda em diferentes níveis de agrupamentos as relações existentes entre os padrões de cobertura terrestre presentes na imagem. Além disso, caso o usuário deseje, a imagem pode ser classificada considerando diferentes quantidades de classes;

- A representação distribuída das classes por meio de grupos de protótipos dá ao método a potencialidade de descobrir agrupamentos de dados que possuem geometrias complexas e variadas. Métodos como o K-médias utilizam apenas um único protótipo (centróide) para representar cada classe e por isso são capazes de detectar adequadamente apenas agrupamentos que tenham formatos hiperesféricos;

- O método pode determinar sem nenhuma intervenção do usuário o número ideal de agrupamentos ou de classes da imagem;

- O método classifica de maneira diferenciada pixels situados em regiões de transição entre classes. Esse procedimento contribui para aumentar a exatidão da classificação resultante.

Tendo em vista que a maioria dos métodos convencionais de classificação não-supervisionada de imagens de sensoriamento remoto se baseiam somente em métodos de agrupamentos particionais (como o K-médias e o ISODATA), a possibilidade de utilização de um método de agrupamento hierárquico para analisar a imagem é certamente um dos benefícios mais importantes proporcionado pelo uso do SOM nesta abordagem. Além disso, a imposição de restrições às possíveis fusões dos protótipos do SOM diminui ainda mais a complexidade computacional do método de agrupamento hierárquico aplicado.

A estratégia de modificar o cálculo de índices de validação de agrupamentos utilizando os protótipos do SOM também contribui para o bom desempenho do método. Apesar dos erros de aproximação, o cálculo modificado do índice não compromete a tomada de decisão sobre o número ideal de classes da imagem e ainda colabora significativamente na redução do tempo de processamento global.

Além da imagem teste utilizada nos experimentos mostrados aqui, o método proposto também foi aplicado em outras imagens de média e alta resolução com números variados de classes, tendo apresentado resultados satisfatórios.

Finalmente, é importante ressaltar ainda que, apesar dos trabalhos já desenvolvidos que aplicam as Redes Neurais na classificação de imagens de sensores



remotos apresentem bons resultados em comparação com os métodos convencionais, a maioria desses trabalhos, além de abordar muito superficialmente o processo de classificação em si, não explora todo o potencial que as redes neurais podem oferecer. De fato, redes neurais de *Perceptrons* Multicamadas com algoritmo de treinamento *backpropagation* são utilizadas como se fossem a única alternativa para os classificadores estatísticos convencionais. Portanto, o presente trabalho também mostra, através da aplicação do SOM, que alternativas no campo das redes neurais podem ainda ser exploradas e aplicadas na análise de imagens de sensoriamento remoto.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ALIXANDRINI, M. J. JR., CENTENO, J. A. S. Experimento da Aplicação de Algoritmos Genéticos para a Integração de Dados Espectrais e Textura na Classificação de Imagens de Alta Resolução. In: **Anais do Simpósio Brasileiro de Geomática**, pp. 212-217, 2002.
- BALL, G., HALL, D. A Clustering Technique for Summarizing Multivariate Data. **Behavior Science**, v. 12, pp. 153-155, 1967.
- BEZDEK, J. C., PAL, N. R. "Some new indexes of cluster validity", **IEEE Transactions on Systems Man and Cybernetics**, v. 28, pp. 301-315, 1998.
- COSTA, J. A. F. **Classificação Automática e Análise de Dados por Redes Neurais Auto-organizáveis**. 1999. Tese (Doutorado em Engenharia Elétrica). Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas-SP.
- DEAN, A. M., SMITH, G. M. An Evaluation of Per-Parcel Land Cover Mapping using Maximum Likelihood Class Probabilities. **International Journal of Remote Sensing**, v. 24, n. 14, July 2003.
- DUDA, T., CANTY, M. Unsupervised Classification of Satellite Imagery: choosing a good algorithm. **International Journal of Remote Sensing**, v. 23, n. 11, June 2002.
- GONÇALVES, M. L., NETTO, M. L. A., ZULLO JUNIOR, J. A. Neural Architecture for Image Classification using Kohonen Feature Extraction. In: **Proceedings of XVIII International Congress for Photogrammetry and Remote Sensing (ISPRS, Comission II)**, v. 31, pp. 117-122, Viena, 1996.
- GONÇALVES, M. L. **Uma Arquitetura Neural Modular para a Classificação de Imagens Multiespectrais de Sensoriamento Remoto**. 1997. Dissertação (Mestrado em Engenharia Elétrica). Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas-SP.
- GONÇALVES, M. L., NETTO, M. L. A., COSTA, J. A. F., ZULLO JUNIOR, J. Automatic Remotely Sensed Data Clustering by Tree-Structured Self-Organizing Maps. In: **Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS'05)**, v.1, 4 p., Seoul, July 2005.
- HALKIDI, M., VAZIRGIANNIS, M., "Clustering validity assessment using multi representatives". In: **Proceedings of SETN Conference**, Thessaloniki, Grécia, 2002.
- HANSEN, M., DUBAYAH, R., DEFRIES, R. Classification Trees: An Alternative to Traditional land Cover Classifiers. **International Journal of Remote Sensing**, v. 17, n. 5, 1996.
- HAYKIN, S. **Neural Networks: A Comprehensive Foundation**. New York: Prentice-Hall, 2nd edition, 1999.
- JI, M. "Using fuzzy sets to improve cluster labeling in unsupervised classification," **International Journal of Remote Sensing**, v. 24, pp. 657-671, 2003.
- KOHONEN, T. **Self-Organizing Maps**, 2nd Edition, Berlin: Springer Verlag, 1997.
- LILLESAND, T. M., KIEFER, R. W. **Remote Sensing and Image Interpretation**, John Wiley & Sons, inc., 2000.
- MAGNUSSEN, S., BOUDEWYN, P., WULDER, M. Contextual Classification of Landsat TM Images to Forest Inventory Cover Types. **International Journal of Remote Sensing**, v. 25, n. 12, pp. 2421-2440, 2004.
- MAULIK, U., BANDYOPADHYAY, S. Performance Evaluation of Some Clustering Algorithms and Validity Indices. **IEEE Transactions on Pattern Analysis and Machine Intelligence**. v. 24, n. 12, December 2002.
- PAL, M., MATHER, P. M. Support Vector Machines for Classification in Remote Sensing. **International Journal of Remote Sensing**, v. 26, n. 5, March 2005.
- RICHARDS, J. A. Analysis of Remotely Sensed Data: The Formative Decades and the Future. **IEEE Transactions on Geoscience and Remote Sensing**, v. 43, n. 3, March 2005.
- WANG, F. Fuzzy Supervised Classification of Remote Sensing Images. **IEEE Transactions on Geoscience and Remote Sensing**, v. 28, n. 2, March 1990.
- XU, R., WUNSCH II, D. Survey of Clustering Algorithms. **IEEE Transactions on Neural Networks**. v. 16, n. 3, May 2005.

# Cluster Analysis of Land-Cover Images Using Automatically Segmented SOMs with Textural Information

Márcio L. Gonçalves<sup>1,2</sup>, Márcio L.A. Netto<sup>2</sup>, and José A.F. Costa<sup>3</sup>

<sup>1</sup>Department of Computer Science, PUC Minas, Poços de Caldas, MG, Brazil

<sup>2</sup>School of Electrical Engineering, State University of Campinas, Brazil

<sup>3</sup>Department of Electrical Engineering, Federal University of Rio Grande do Norte, Brazil  
marcio@pucpcaldas.br, marcio@dca.fee.unicamp.br,  
alfredo@dee.ufrn.br

**Abstract.** This work attempts to take advantage of the properties of Kohonen's Self-Organizing Map (SOM) to perform the cluster analysis of remotely sensed images. A clustering method which automatically finds the number of clusters as well as the partitioning of the image data is proposed. The data clustering is made using the SOM. Different partitions of the trained SOM are obtained from different segmentations of the U-matrix (a neuron-distance image) that are generated by means of mathematical morphology techniques. The different partitions of the trained SOM produce different partitions for the image data which are evaluated by cluster validity indexes. Seeking to guarantee even greater efficiency in the image categorization process, the proposed method extracts information from the image by means of pixel windows, in order to incorporate textural information. The experimental results show an application example of the proposed method on a TM-Landsat image.

**Keywords:** data clustering, self-organizing maps, image processing, remote sensing.

## 1 Introduction

The advances in computer and electronic technologies and decreasingly cost of memory storage systems have been enabling large amounts of data to be available in many application areas. An example is the remote sensing of the earth surface from satellite or airborne scanners. Large volumes of remotely sensed images are being generated from an increasing number of sophisticated airborne and space borne sensor systems, and while there is no substitute for a trained analyst, exploitation of this data on a large scale requires consistent automatic data exploration tools [8].

The self-organizing map (SOM), proposed by Kohonen [6], has been widely used in a variety of applications, including areas as data compression and data mining [7]. Important properties as the input space approximation, topological ordering and density matching, allied with the simplicity of the model and the easiness to implement its learning algorithm justify the success of the SOM and place it as one of the main models of neural nets in the present time.

This work presents a methodology that explores the characteristics and properties of the SOM to perform the cluster analysis of remotely sensed images. In the proposed method, the SOM is used to map the original patterns of the image to a 2-dimensional neural grid. The objective is to quantize and represent the image patterns in a space of smaller dimension, seeking to preserve the probability distribution and topology of the input space. Afterwards, different partitions of the trained SOM are obtained from different segmentations of the U-matrix [9], which are generated by means of mathematical morphology techniques. Each different partition of the U-matrix corresponds to a different clustering configuration of SOM neurons that can be utilized to represent the patterns by which the original image will be categorized. A cluster validity index is applied to determine automatically the best partition of the image data.

Seeking to guarantee even greater efficiency in the categorization process, the proposed method extracts samples from the image by means of pixel windows, in order to incorporate textural information. Following this approach, the method filters heterogeneous samples which represent patterns (pixel windows) corresponding to transition regions between different land cover classes.

The remainder of the paper is organized in the following form: section 2 describes succinctly the SOM; section 3 presents the proposed clustering methodology, while section 4 shows an application example of the proposed approach on a TM-LANDSAT image, and section 5 gives the conclusions and final considerations.

## 2 SOM

SOM is a type of artificial neural net based on competitive and unsupervised learning, i.e., no information about the input signal classes is used in the adjustment process for synaptic weight in the net [6]. The network essentially consists of two layers: an input layer  $I$  and an output layer  $U$  with neurons generally organized in a 2-dimensional topological array. The input to the net corresponds to a  $p$ -dimensional vector,  $\mathbf{x}$ , generally in the space  $\mathcal{R}^p$ . All of the  $p$  components of the input vector feed each of the neurons on the map. Each neuron  $i$  can be represented by a synaptic weight vector  $\mathbf{w}_i = [w_{i1}, w_{i2}, \dots, w_{ip}]^T$ , also in the  $p$ -dimensional space.

For each input pattern  $\mathbf{x}$  a winner neuron,  $c$ , is chosen, using the criterion of greatest similarity:

$$\|\mathbf{x} - \mathbf{w}_c\| = \min_i \{\|\mathbf{x} - \mathbf{w}_i\|\} \quad (1)$$

where  $\|\cdot\|$  represents the Euclidian distance. The winner neuron weights, together with the weights of the neighboring neurons, are adjusted according to the following equation:

$$\mathbf{w}_i(t+1) = \mathbf{w}_i(t) + h_{ci}(t)[\mathbf{x}(t) - \mathbf{w}_i(t)] \quad (2)$$

where  $t$  indicates the iteration of the training process,  $\mathbf{x}(t)$  is the input pattern and  $h_{ci}(t)$  is the nucleus of neighborhood around the winner neuron  $c$ .

Once the SOM algorithm has converged, the 2-dimensional array of neurons displays important statistical properties, such as the approximation of the input space, topological ordering, and density matching.

Although the SOM presents attractive properties on the input data, the trained neural network requires additional procedures to enable a suitable interpretation of the data clusters. An example is the Unified Distance Matrix (U-matrix) method, which was developed by A. Ultsch [9] to detect non-linearities in the resulting SOM mapping. The basic idea is to use the same metric that was used during the learning to compute distances between adjacent reference vectors. This method can be used to visualize the topological structure of the SOM unit layer and therefore also the topology of the N-dimensional input space. The U-matrix can be visualized as a three dimensional landscape. Altitudes or the high places on the U-matrix encode dissimilarities between neurons and correspond to cluster borders while valleys represent to map units that are similar.

### 3 Proposed Clustering Methodology

The methodology proposed in this work essentially attempts to exploit the properties of SOM to perform the cluster analysis of remotely sensed images. The key strategy of the clustering method proposed here is to perform the clusters analysis of the image through a set of SOM prototypes instead of working directly with the original patterns of the scene. The method proposed basically presents four processing stages: sampling of the input image, training and segmentation of the SOM, and final categorization of the image.

In the following subsections each of the steps of the proposed methodology is explained in greater detail.

#### 3.1 Sampling

The first step of the proposed methodology consists in collecting an image sample set in order to train the SOM. Unlike pixel by pixel approaches that only use the spectral information of individual points to find homogenous regions; the present work performs the sampling of the image through pixel windows. The idea is to incorporate in the clustering process information about the neighborhood (context) of the pixels, considering that isolated pixels are not able to represent the majority of cover land patterns, especially in the case of images that have higher spatial resolutions. The sample windows are collected uniformly across the entire region of the image, without overlappings and at regular intervals. All of the samples are square and have the same size.

Seeking to guarantee greater efficiency in the clustering method, the proposed method filters heterogeneous samples which represent patterns (pixel windows) corresponding to transition regions between different land cover classes. Heterogeneous samples are those that have a high degree of spectral heterogeneity and are normally associated with input patterns that incorporate more than one land cover class.

The spectral heterogeneity degree of each sample is computed from Haralick's co-occurrence matrix [4]. Since the samples are pixel windows, it makes it possible to generate an image of each sample and to calculate the co-occurrence probability of all pairwise combinations of grey levels in each one of them. The energy (sometimes called uniformity) was the measure chosen to calculate the spectral heterogeneity of

each sample from co-occurrence matrix. This measure, described through the equation (3), gets values next to 1 when the area of interest presents uniform texture (similar grey levels), and values that tend to zero when the area is not uniform.

$$ENE = \sum_i \sum_j (P(i, j)_{d, \theta})^2 \quad (3)$$

where  $P(i, j)_{d, \theta}$  is the co-occurrence probability of two grey levels  $i$  and  $j$ , separate to a distance  $d$  in the direction  $\theta$ . The prototypes whose ENE's satisfy the relationship given below are considered heterogeneous and are consequently filtered:

$$ENE < \mu_{ENE} - \frac{1}{2} \sigma_{ENE}. \quad (4)$$

Here  $\mu_{ENE}$  and  $\sigma_{ENE}$  are, respectively the average and the standard deviation of the ENE's of all of the samples.

### 3.2 SOM Training

In order to train the SOM, some parameters must be specified to define the structure of the map and to specifically control the stated training. With the objective of guaranteeing good mapping of the original patterns, the proposed methodology defines in a particular way the neural net parameters based on the existing literature, on experimental tests, and some peculiarities of the application of SOM on remotely sensed images. However, since the SOM can be sensitive to choice of its training parameters, other alternatives can also be sought out to obtain good maps [5].

The proposed methodology utilizes the following parameters to train the SOM: linear initialization of weights, batch training mode, gaussian neighborhood function and rectangular shape to organize the two-dimensional array of neurons of the net.

The size of the map is one of the free parameters of SOM that particularly depends on the input image and the objectives of the clustering. If the objective is to detect all of the patterns in the image, including those with low probability of occurrence, large-sized maps must be employed in the analysis; in the opposite case, if the interest is concentrated only on the predominant patterns in the scene, a smaller-sized SOM can be utilized.

### 3.3 Segmentation of the SOM

At the third processing stage of the proposed approach the trained SOM is segmented. The strategy used in this work to interpret the SOM determines the best partition for the trained SOM from the analysis of different segmentations of the U-matrix. The strategy used can be seen as an improvement of the clustering method proposed in [2]. Costa and Netto [2] proposed an efficient method based on mathematical morphology to segment the U-matrix. The method applies the images segmentation algorithm, watershed [1], using a markers image to regularize the segmentation process. This same approach also is applied in our proposal, however, instead of using only one markers image to segment the U-matrix (as proposed in [2]), a quantity of markers images are considered and, therefore, different segmentations of the U-matrix are

obtained. Each one of these segmentations is associated with the neurons of the SOM, allowing to determine different partitions for the map which define different partitions for the image data. To select the best one among different partitions, each of these is evaluated using the CDbw cluster validity index proposed in [3].

Given the U-matrix image  $U$  obtained from the trained SOM, the following steps are performed to obtain the markers images set to the image  $U$ :

1. Filtering: create the image  $U_1$  by removing any pore with area less or equal than three pixels.
2. For  $k = 1, \dots, f_{max}$ , where  $f_{max}$  is the highest gray level of the image  $U_1$  {
  - 2.1. Create the binary image  $U_1^k$  that corresponds to the conversion  $U_1$  to a binary image using as threshold  $k$ .
  - 2.2. Obtain  $N_{rc}^k$ , the number of connected regions of  $U_1^k$  .}
3. Obtain the most persistent values of number of connected regions that correspond to the plateaus with sizes more than three contiguous gray levels in the plot of  $N_{rc}^k$  versus  $k$ .
4. Obtain the set of all markers images,  $S_m = \{ U_2^{k_1}, U_2^{k_2}, \dots, U_2^{k_n} \}$ , where  $k_1, k_2, \dots, k_n$  are initial values of the plateaus chosen in the previous step.

Although the used procedure to find the markers of the U-matrix is similar to that one presented in [2], in our approach the steps 4 and 5 determine a set of markers images instead of only one image considering all the significant plateaus.

Thus, the general strategy for the partitioning of the trained SOM can be summarized as follows:

1. Obtain the U-matrix using the trained SOM.
2. Find the set of all markers images to the U-matrix,  $S_m = \{ U_2^{k_1}, U_2^{k_2}, \dots, U_2^{k_n} \}$  (as described previously).
3. For each markers image,  $U_2^{k_i}, i=1,2,\dots,n$  {
  - 3.1. Compute the watershed lines on the U-matrix.
  - 3.2. Assign a different label for each connected region (cluster of neurons) of the U-matrix.
  - 3.3. Copy the U-matrix labels to the corresponding neurons in the map.
  - 3.4. Apply the CDbw index to evaluate the SOM partitioned in the step 3.4. }
4. Select the best partition for the SOM comparing the values of the CDbw index obtained for each different partition of the neurons map.

### 3.4 Image Catgorization

In order to categorize the image, pixel windows are collected from the original image with equal dimensions from the training sample and are compared to all of the SOM prototypes. This comparison is performed through the distances calculated between the pixel windows and each of the prototypes. The central pixel of the pixel window receives the label of the prototype that has the least distance from it. The image is then entirely run through until all of the pixels have been labeled.

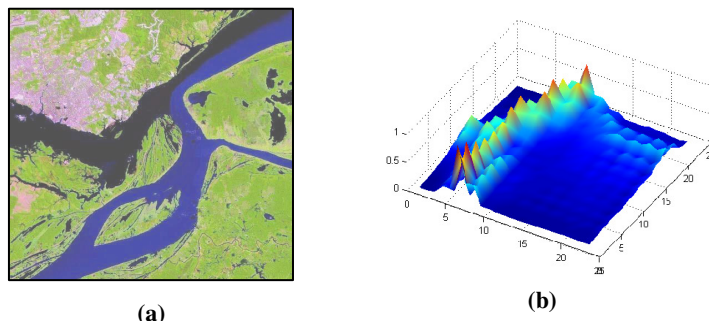
## 4 Experimental Results

This section shows an application example of the proposed methodology on a test image. The image used in the experiments is composed of spectral bands 3, 4, and 5 of the Landsat-5 TM satellite and was acquired on 20 August 1988. It has an IFOV of 30 m. The study area covers the city of Manaus and the intersection of two rivers, Rio Negro and Rio Solimões in the state of Amazonas, Brazil. Four large land cover patterns are present in the scene: urban area, vegetation, and two water patterns (the darker pattern corresponds to Rio Negro and the more purple pattern corresponds to Rio Solimões). This image was provided by the National Space Research Institute, (INPE), Brazil. The Fig.1 (a) shows a color composite of the test image.

Application of the proposed methodology was initially performed with a sampling process of the scene. Sample windows of size 5x5 were collected uniformly across the entire region of the image, without overlappings and at regular 10 pixel intervals, given a total of 2500 samples obtained without user intervention. Applying the samples filtering procedure, of the 2500 total samples, 635 of them presented a high degree of spectral heterogeneity, given that its ENE values exceeded the threshold defined in the equation (4), remaining then 1865 samples to be used in the SOM training stage.

A SOM composed of 144 neurons arranged in a 12x12 rectangular grid was trained with all of the 1865 samples. The other parameters of the SOM were defined according to the specifications presented in section 3.2.

The U-matrix was calculated from the trained SOM (Fig. 1(b) shows a 3D view of the U-matrix). The strategy proposed to segment the SOM (presented in the subsection 3.3) determines different partitions for the U-matrix using different markers images. It uses information like the number of connected regions ( $N_{rc}^k$ ) for each gray level ( $k$ ) of the U-matrix image for a useful gray level range, which in turn is related to the distances between neighboring neurons. For the U-matrix image of this experiment were determined four markers images. The markers images were obtained after thresholding the U-matrix by the lower gray levels from significant plateaus in the

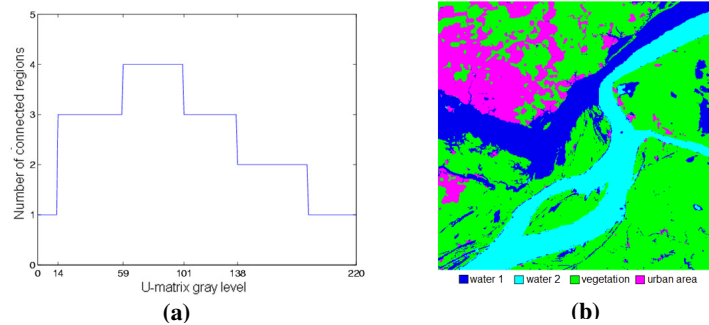


**Fig. 1.** (a) Color composite of the image used in the tests. (b) 3D view of the U-matrix for trained SOM.

plot of  $N_{rc}^k$  versus  $k$  (Fig. 2(a)). In this case, the values of these gray levels  $k$  were 14, 59, 101, and 138. Therefore, in accordance with the proposed method four different segmentations of U-matrix image were performed using the watershed method, one for each marker image. After, four different partitions for the SOM were determined from the segmented U-matrix images.

To evaluate the different partitions generated for the trained SOM, the CDbw cluster validity index was applied. The CDbw index presented the higher value for the partition  $k = 49$ , which corresponds to 4 clusters. For the CDbw index, the greater its value, the better the partition [3].

Figure 2(b) shows the test image categorized in accordance with the prototypes of the labeled SOM considering the partition  $k = 49$ .



**Fig. 2.** (a) Number of connected regions versus image threshold of U-matrix. (b) Test image categorized by proposed method.

Attempting to evaluate the categorization generated by the proposed method, and considering the absence of terrestrial truth for the test image, the present work performed the classification of the image in a supervised manner using a multilayers Perceptrons (MLP) neural net, and considered these results as a reference (or “true”) to calculate the Kappa agreement index (normally used to evaluate the accuracy of satellite image categorization). The MLP net was trained with a sample set collected from original image by an image analyst. The Kappa index obtained here was 0.93, which allows concluding that the categorization result of the test image by method presented in this work was very satisfactory.

## 5 Conclusions and Final Considerations

In this work, a cluster analysis method of remotely sensed images that attempts to exploit the properties of the SOM was presented. The key point of the proposed method is to perform the clusters analysis of the image through a set of SOM prototypes instead of working directly with the original patterns of the scene. This approach significantly reduces the complexity of the analysis and presents advantages



that make it as a promising alternative to carry out the data clustering of remotely sensed images. Among these, we can point out:

- The proposed method does not require a previous definition of the number of clusters to perform the categorization of the image. It does not occur in the majority of the conventional unsupervised categorization methods;
- The distributed representation of the patterns by means of prototype groups gives the method the potential to discover geometrically complex and varied data clusters. Methods such as K-means use a single prototype (centroid) to represent each pattern and because of this are only capable of adequately detecting clusters that have hyperspherical formats;
- The simple use of pixel windows allows textural information to be included without any explicit calculation of measure for it. This approach contributes to the quality of the resulting categorization.

In addition to the test image utilized in the experiments shown here, the proposed method has also been applied to other high and medium resolution images, with satisfactory results.

## References

1. Bleau, A., Leon, L.J.: Watershed-based segmentation and region merging. In: *Comp. Vis. Image Underst.*, vol. 77, pp. 317–370 (2000)
2. Costa, J.A.F., Netto, M.L.A.: Clustering of Complex Shaped Data Sets via Kohonen Maps and Mathematical Morphology. In: *Proceedings of the SPIE Conference on Data Mining and Knowledge Discovery*, Orlando, FL, vol. 4384, pp. 16–27 (2001)
3. Halkidi, M., Vazirgiannis, M.: Clustering validity assessment using multi representatives. In: *Proceedings of SETN Conference*, Thessaloniki, Greece (2002)
4. Haralick, R.M., Shanmugam, K., Dinstein, I.: Textural Features for Image Classification. *IEEE Trans. on Systems, Man and Cybernetics* 3(6), 610–621 (1973)
5. Kaski, S., Lagus, K.: Comparing self-organizing maps. In: *Vorbrüggen, J.C., von Seelen, W., Sendhoff, B. (eds.) ICANN 1996. LNCS*, vol. 1112, pp. 809–814. Springer, Heidelberg (1996)
6. Kohonen, T.: *Self-Organizing Maps*, 2nd edn. Springer, Berlin (1997)
7. Kohonen, T., Oja, E., Simula, O., Visa, A., Kangas, J.: Engineering Applications of the Self-Organizing Map. *Proceedings of the IEEE* 84(10), 1358–1384 (1996)
8. Richards, J.A.: Analysis of Remotely Sensed data: the formative decades and the future. *IEEE Transactions on Geoscience and Remote Sensing* 43, 422–432 (2005)
9. Ultsch, A.: Self-organizing neural networks for visualization and classification. In: *Information and Classification*, pp. 307–313. Springer, Berlin (1993)

# A Three-Stage Approach Based on the Self-organizing Map for Satellite Image Classification

Márcio L. Gonçalves<sup>1,2</sup>, Márcio L.A. Netto<sup>2</sup>, and José A.F. Costa<sup>3</sup>

<sup>1</sup> Department of Computer Science, PUC Minas, Poços de Caldas, MG, Brazil

<sup>2</sup> School of Electrical Engineering, State University of Campinas, Brazil

<sup>3</sup> Department of Electrical Engineering, Federal University of Rio Grande do Norte, Brazil

marcio@pucpcaldas.br, marcio@dca.fee.unicamp.br,

alfredo@dee.ufrn.br

**Abstract.** This work presents a methodology for the land-cover classification of satellite images based on clustering of the Kohonen's self-organizing map (SOM). The classification task is carried out using a three-stage approach. At the first stage, the SOM is used to quantize and to represent the original patterns of the image in a space of smaller dimension. At the second stage of the method, a filtering process is applied on the SOM prototypes, wherein prototypes associated to input patterns that incorporate more than one land cover class and prototypes that have null activity are excluded in the next stage or simply eliminated of the analysis. At the third and last stage, the SOM prototypes are segmented through a hierarchical clustering method which uses the neighborhood relation of the neurons and incorporates spatial information in its merging criterion. The experimental results show an application example of the proposed methodology on an IKONOS image.

**Keywords:** Pattern recognition, self-organizing maps, image processing, remote sensing.

## 1 Introduction

The self-organizing map (SOM), proposed by Kohonen [5], has been widely used in a variety of applications, including areas as pattern recognition, data compression, biological modeling, signal processing, and data mining [6]. Important properties as the input space approximation, topological ordering and density matching, allied with the simplicity of the model and the easiness to implement its learning algorithm justify the success of the SOM and place it as one of the main models of neural nets in the present time.

This work presents a methodology that explores the characteristics and properties of the SOM to perform the unsupervised classification of remotely sensed images. Since the first satellites were launched for the purpose of searching for terrestrial resources, the digital classification of remotely sensed images has acquired a growing importance in the automatic recognition of the land cover patterns [8]. Presently, the enormous quantity of images that are being collected by sensor systems that are more and more sophisticated and modern require the development of innovative classification

methodologies, which allow an automatic and efficient identification of the great volume of information available in the images and at the same time makes the mapping process of terrestrial surfaces less subjective and with greater potential for reuse in subsequent situations.

The classification method proposed here performs a cluster analysis of the image data employing an approach consisting of three processing stages through SOM. Firstly, the SOM is used to map the original patterns of the image to a reduced set of prototypes (neurons) arranged in a two-dimensional grid. Afterwards, a filtering process is applied on the SOM prototypes, wherein prototypes associated to input patterns that incorporate more than one land cover class are excluded in the next stage of the analysis, and prototypes that have null activity are simply discarded. At the last processing stage, an agglomerative hierarchical clustering method is applied over the filtered SOM, generating a dendrogram of clustered prototypes with different degrees of similarity. Each level of dendrogram obtained corresponds to a different clustering configuration of SOM prototypes that can be utilized to represent the classes by which the original image will be classified.

The remainder of the paper is organized in the following form: section 2 describes the SOM and its properties, section 3 presents a brief explanation about the unsupervised classification of satellite images, while section 4 explains the proposed classification methodology. An application example of the proposed approach on an IKONOS image is shown in the section 5, and section 6 gives the conclusions and final considerations pointing out the advantages of the proposed method on conventional classification methods.

## 2 SOM

SOM is a type of artificial neural net based on competitive and unsupervised learning. The network essentially consists of two layers: an input layer  $I$  and an output layer  $U$  with neurons generally organized in a 2-dimensional topological array. The input to the net corresponds to a  $p$ -dimensional vector,  $\mathbf{x}$ , generally in the space  $\mathcal{R}^p$ . All of the  $p$  components of the input vector feed each of the neurons on the map. Each neuron  $i$  can be represented by a synaptic weight vector  $\mathbf{w}_i = [w_{i1}, w_{i2}, \dots, w_{ip}]^T$ , also in the  $p$ -dimensional space.

For each input pattern  $\mathbf{x}$  a winner neuron,  $c$ , is chosen, using the criterion of greatest similarity:

$$\|\mathbf{x} - \mathbf{w}_c\| = \min_i \{\|\mathbf{x} - \mathbf{w}_i\|\} \tag{1}$$

where  $\|\cdot\|$  represents the Euclidian distance. The winner neuron weights, together with the weights of the neighboring neurons, are adjusted according to the following equation:

$$\mathbf{w}_i(t+1) = \mathbf{w}_i(t) + h_{ci}(t)[\mathbf{x}(t) - \mathbf{w}_i(t)] \tag{2}$$

where  $t$  indicates the iteration of the training process,  $\mathbf{x}(t)$  is the input pattern and  $h_{ci}(t)$  is the nucleus of neighborhood around the winner neuron  $c$ .

Once the SOM algorithm has converged, the 2-dimensional array of neurons displays important statistical characteristics of the input space, summarized as follows:

i) *Approximation of the input space*: the basic objective of SOM is to store a large set of input vectors by finding a smaller set of prototypes that provides a good approximation to the original input space.

(ii) *Topological ordering*: the SOM algorithm attempts to preserve as well as possible the topology of the original space, i.e., it tries to make the neighboring neurons in the 2-dimensional array (output space) present weight vectors that represent neighboring patterns in the input space.

(iii) *Density Matching*: the SOM reflects the probability distribution of data in the input space. Regions in the input space where the input patterns  $\mathbf{x}$  are taken with a high probability of occurrence are mapped onto larger domains of the output space, and thus with better resolution than regions in the input space from which input patterns  $\mathbf{x}$  are taken with a low probability of occurrence.

In the literature there are some proposed algorithms that seek to automatically (or semi-automatically) interpret and segment the neurons of a trained SOM [1, 2, 9, 10]. As will be presented in section 4, the SOM segmentation strategy employed in this work utilizes an agglomerative hierarchical clustering method that incorporate more information about the data clusters in its merging criterion beyond the usual inter-cluster distance information.

### 3 Unsupervised Classification of Remotely Sensed Images

The unsupervised classification of remote sensing images is based on the principle that the computational algorithm is capable of identifying by itself the classes of the image. This type of classification is frequently performed through clustering methods.

Although there are a large quantity of different clustering methods in the pattern recognition area [11], the majority of software or computational systems meant to the digital processing of remotely sensed images perform unsupervised classification based on the partitional clustering methods, such as K-means and ISODATA.

Despite being widely used, these partitional clustering methods have various limitations. The objective functions that they used begin with the assumption that the number of classes,  $K$ , is known a priori. In the hypothesis that an inadequate  $K'$  value has been chosen, the method will impose, through the use of optimization techniques,  $K'$  clusters to the data. The user must also manually specify various parameters in order to control the clustering process, among them: the initial centroids of each cluster, the maximum number of iterations, thresholds to perform the division, fusion, or exclusion of clusters. K-means and ISODATA are very sensitive to these parameters, which can generate different partitions when various simulations are done for the same data set. Facing this, the optimal value of these parameters is frequently encountered through trial and error. These needs certainly increase the level of interaction between the user and the computational algorithm, consequently increasing the degree of subjectivity of the categorization process of the image.

Other no less important limitations of partitional algorithms, such as K-means and ISODATA, are: the high computational cost when the data to be analyzed is very large (at each iteration, all of the pixels in the image are compared with all of the clusters centroids) and the existence of suppositions about the cluster forms. Generally only one prototype (centroid) is used to represent a cluster, thus these methods become adequate only for the analysis of clusters that have hyperspherical formats.

Another possible, though uncommon, way of performing unsupervised classification of remotely sensed images is through hierarchical clustering methods. Unlike partitional methods, hierarchical methods do not require the user to specify the number of clusters and other additional parameters beforehand. Another significant advantage of these methods is that they make it possible to visualize the result of the classification by means of a dendrogram which illustrates in a hierarchical form the degree of similarity between the clusters that are formed by fusions (or divisions) at each successive stage of the analysis. However, hierarchical methods have some characteristics that prevent their application in the classification of remotely sensed images: (a) in general they require memory space in the order of  $O(N^2)$ , in which  $N$  is the number of records in the data set; (b) the results can be difficult to interpret, mainly for large data sets; (c) in order to determine the cutoff of the dendrogram (ideal number of clusters) some decision criteria must be applied [11].

## 4 Proposed Methodology

The key point of the unsupervised classification method of satellites images proposed here is to perform the clusters analysis of the image through a set of SOM prototypes instead of working directly with the original patterns of the scene. For this, an approach consisting of three processing stages through SOM is utilized with the objective to discover representative clusters of prototypes for each land-cover class of interest. The three processing stages, which consist basically of the training, filtering and segmentation of the SOM, are described in the following subsections.

### 4.1 Training

At the first processing stage a sample set collected from the original image is used to train the SOM. Unlike pixel by pixel approaches that only use the spectral information of individual points to find homogenous regions, the proposed method performs the sampling of the image through pixel windows. The idea is to incorporate in the classification process information about the neighborhood (context) of the pixels, considering that isolated pixels are not able to represent the majority of cover land patterns, especially in the case of images that have higher spatial resolutions. The sample windows are collected uniformly across the entire region of the image, without overlappings and at regular intervals. All of the samples are square and have the same size.

In order to train the SOM, some parameters must be specified to define the structure of the map and to specifically control the stated training. With the objective of guaranteeing good mapping of the original patterns, the proposed methodology

defines in a particular way the neural net parameters based on the existing literature, on experimental tests, and some peculiarities of the application of SOM on remotely sensed images. However, since the SOM can be sensitive to choice of its training parameters, other alternatives can also be sought out to obtain good maps [4].

The proposed methodology utilizes the following parameters to train the SOM: linear initialization of weights, batch training mode, gaussian neighborhood function and rectangular shape to organize the two-dimensional array of neurons of the net.

The size of the map is one of the free parameters of SOM that particularly depends on the input image and the objectives of the classification. If the objective is to detect all of the patterns in the image, including those with low probability of occurrence, large-sized maps must be employed in the analysis; in the opposite case, if the interest is concentrated only on the predominant patterns in the scene, a smaller-sized SOM can be utilized. Nevertheless, the performance of the proposed classification methodology is not significantly affected if sufficiently large sizes for the SOM are utilized. Although maps with larger dimensions than are necessary have a larger quantity of inactive neurons, as shows the next processing stage, this event is not prejudicial within the proposed methodology.

## 4.2 Filtering

The second stage of the proposed approach consists of filtering two types of prototypes that generally appear in the mapping of image patterns through SOM. These prototypes, denominated here as *inactive* and *heterogeneous*, can act as borders (or “interpolation units”) in the SOM grid contributing to the separation of clusters.

The inactive prototypes correspond to the neurons that have null activity in the SOM competitive learning process, i.e., they are not associated with any input pattern. These prototypes are simply eliminated of the analysis.

Heterogeneous prototypes are those that have a high degree of spectral-textural heterogeneity and are normally associated with input patterns that incorporate more than one land cover class. Most of the time, these patterns correspond to transition regions between land cover classes present in the image and are captured in consequence of the sampling through pixel windows. Prototypes that are considered heterogeneous are excluded in the third stage of the proposed approach, in which the hierarchical clustering method is applied. The objective of excluding these prototypes is to prevent them (and consequently the input patterns associated with them) from being erroneously attributed to one of the classes that are part of them. Heterogeneous prototypes can be seen as noisy or divergent patterns, and if they are not filtered the hierarchical method can incorporate them in the clusters that will be produced or retain them in separate clusters. And because in hierarchical methods, data or cluster fusion at a determined level can not be corrected in subsequent levels, incorrect interpretations regarding the classes and/or the number of classes of the image can be made. The input patterns associated with these heterogeneous prototypes are particularly classified only at the end of the analysis, considering the neighboring pixels that have already been labelled.

The spectral-textural heterogeneity degree of each prototype of the SOM is computed from Haralick’s co-occurrence matrix [3]. Since the weight vectors of the SOM prototypes have the same dimensions as the input patterns (that in this case are pixel windows), it makes it possible to generate an image of each prototype of the net and to calculate the co-occurrence probability of all pairwise combinations of grey levels in each one of them. The *energy* (sometimes called uniformity) was the measure chosen to calculate the spectral-textural heterogeneity of each prototype from co-occurrence matrix. This measure, described through the equation (3), gets values next to 1 when the area of interest presents uniform texture (similar grey levels), and values that tend to zero when the area is not uniform.

$$ENE = \sum_i \sum_j P(i, j)_{d, \theta} \tag{3}$$

where  $P(i, j)_{d, \theta}$  is the co-occurrence probability of two grey levels  $i$  and  $j$ , separate to a distance  $d$  in the direction  $\theta$ .

The prototypes whose  $ENE$ ’s satisfy the relationship given below are considered heterogeneous and are consequently filtered:

$$ENE > \mu_{ENE} + \frac{1}{2} \sigma_{ENE} . \tag{4}$$

Here  $\mu_{ENE}$  and  $\sigma_{ENE}$  are, respectively the average and the standard deviation of the  $ENE$ ’s of all of the prototypes.

### 4.3 Segmentation

At the last processing stage of the proposed approach, an agglomerative hierarchical clustering method is applied to the trained and filtered SOM prototypes.

The hierarchical clustering method utilized here has two important characteristics. The first one of them is the imposition of restrictions to the possible SOM prototype fusions. Unlike traditional hierarchical clustering methods, which consist of comparing all of the pairs of objects to decide on a fusion, the approach utilized in this work verifies the possibility of fusions only between adjacent (or neighboring) prototype pairs in the SOM grid. Another important characteristic is that beyond using inter-cluster distance information the employed merging criterion also incorporates space information of the image pixels associated to the SOM prototype clusters.

The distance information ( $D_{ij}$ ) between two prototype clusters  $i$  and  $j$  is calculated using the euclidean metric and the nearest neighbor method (or single linkage method). The values of  $D_{ij}$  are normalized within the interval [0,1].

The space information is calculated through two indices, denominated *spatial boundary index* and *spatial compactness index*. These indices, developed by Marçal and Castro [7], are computed here from classified image using the SOM prototype clusters (classes) in each level of the dendrogram generated by hierarchic method.

The spatial boundary index ( $B_{ij}$ ) calculates the boundary length between all class pairs ( $i, j$ ) considering eight neighbors for each pixel (four adjacent and four oblique). Its formula is given as follows:

$$B_{ij} = 1 - \frac{1}{2} \left( \frac{b_{ij}}{\sum_{k=1(k \neq i)}^N b_{ik}} + \frac{b_{ij}}{\sum_{k=1(k \neq j)}^N b_{kj}} \right) \tag{5}$$

where  $b_{ij}$  is the number of boundary counts for the class pair  $(i,j)$  and  $N$  is the number of classes in the dendrogram level that is being analyzed. The idea behind this index is that two classes that have a significant common boundary should be more likely to merge than classes with very little or no common boundaries [7].

The spatial compactness index ( $C_{ij}$ ), defined through the equation (6), is based on the number of self-boundary counts for each class ( $b_{ii}$ ). This index penalizes the merger of compact classes. In the same way that the index  $B_{ij}$ , the index  $C_{ij}$  results values within the interval  $[0,1]$ .

$$C_{ij} = \frac{1}{2} \left( \frac{b_{ii}}{b_{ii} + 6 \sum_{k=1(k \neq i)}^N b_{ik}} + \frac{b_{jj}}{b_{jj} + 6 \sum_{k=1(k \neq j)}^N b_{jk}} \right) \tag{6}$$

The merging criterion adopted here establishes that the pair of prototype classes  $(i,j)$  that to present the lowest value resultant of the average computed between  $D_{ij}$ ,  $B_{ij}$  and  $C_{ij}$  is selected for merger.

At the end of this stage there is a dendrogram that shows in a hierarchical way the similarity levels between the SOM prototypes. Each level of dendrogram obtained corresponds to a different clustering configuration of SOM prototypes that can be utilized to represent the classes by which the original image will be classified.

## 5 Experimental Results

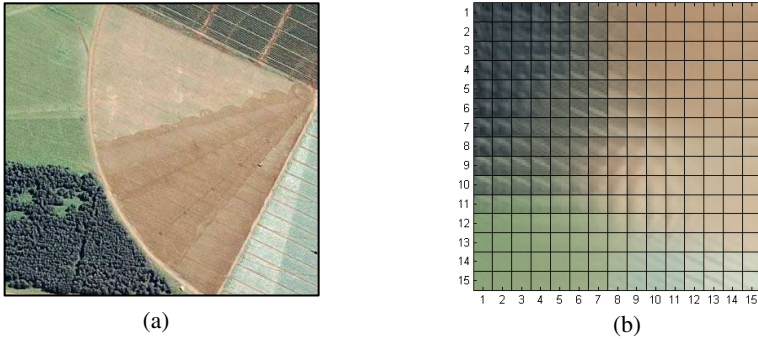
This section shows an application example of the proposed methodology on a test image. The image used in the experiments (provided by Engesat/Brazil, © Space Imaging) has  $385 \times 350$  pixels and is composed of three spectral bands of the IKONOS satellite. The study area shows irrigation pivots in the region of Andaraí in the Bahia state, Brazil. Six large land cover classes are present in the scene: sparse vegetation, forest, two types of exposed soil, and two stages of coffee plantation. The Fig. 1(a) shows a color composite of the test image.

In accordance with the procedures described in the subsection 4.1, 1292 sample windows of size  $9 \times 9$  were collected from test image and used to train a SOM composed of 225 neurons arranged in a  $15 \times 15$  rectangular grid.

The Fig. 1(b) shows the images of each SOM prototype (neuron) arranged in the rectangular grid after the training. By means of them it is possible to visualize prototype clusters that correspond to the land cover classes present in the original image. It is also possible to observe the *topological ordering* and *density matching*



properties in the mapping produced by the neural net. Land cover classes with similar spectral attributes are mapped to neighboring regions of the two-dimensional output grid and those that occupy bigger areas in the original image are mapped to a bigger number of prototypes of the grid.



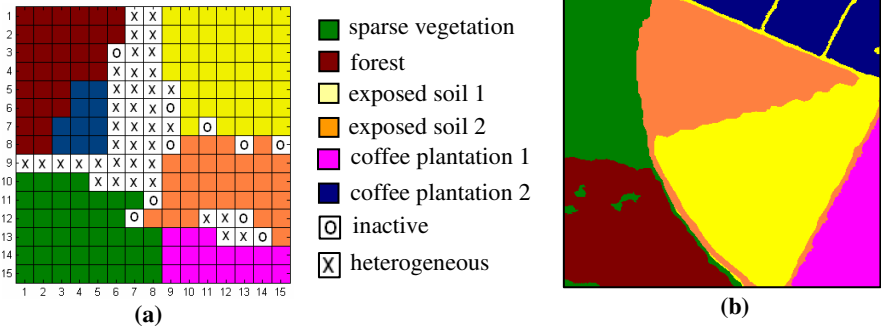
**Fig. 1.** (a) Color composite of the image used in the tests. (b) Image of the SOM prototypes arranged in the rectangular grid after the training. The images shown in (a) and (b) are in different scales.

After the SOM training, the prototypes filtering process was applied. In this experiment, 10 SOM prototypes presented null activity and 39 presented a high degree of spectral-textural heterogeneity, given that its values of energy (ENE) exceeded the threshold defined in the equation (4). Thus, of the 225 total prototypes, 49 of them were filtered, remaining 176 prototypes to be analyzed in the following stage.

In the last processing stage, the agglomerative hierarchical method (described in the section 4.3) was applied on the filtered SOM prototypes. Consequently, a dendrogram consisting of 176 levels was generated, each level with a different configuration of SOM prototypes clusters.

Since the image presents six large land cover classes (previously cited), the level of the dendrogram composed by six SOM prototype clusters was chosen to effect the classification of the scene. Fig. 2(a) shows the SOM grid segmented in 6 clusters. The squares marked with “o” and “x” are, respectively, the inactive and heterogeneous prototypes that were filtered in the previous stage.

To perform the classification of all pixels of the test image, the image was entirely run through considering  $9 \times 9$  pixel windows (size equal to the sample windows) and comparing with all of the SOM prototypes. This comparison was performed using distances calculated between considered pixel windows and each one of the prototypes. The central pixel of the pixel window received the prototype label that presented the shortest distance from it. In the sequence, each one of the pixels of the image that were associated to any heterogeneous prototype was reclassified using the neighboring pixel class that has the least (spectral) distance from it. Fig. 2(b) shows the classification results for the test image using the proposed methodology.



**Fig. 2.** (a) Filtered and segmented SOM grid. (b) Result of the test image classification by proposed methodology.

Attempting to evaluate the classification generated by the proposed method, and considering the absence of terrestrial truth for the test image, the present work performed the classification of the image in a supervised manner using a multilayers *Perceptrons* (MLP) neural net, and considered these results as a reference (or “true”) to calculate the Kappa agreement index (normally used to evaluate the accuracy of image classification). The MLP net was trained with a sample set collected from original image by an image analyst. The Kappa index obtained here was 0.82, which allows to conclude that the classification result of the test image by method presented in this work was very satisfactory.

## 6 Conclusions and Final Considerations

In this work, an approach consisting of three processing stages through SOM was proposed to perform the unsupervised classification of satellite images. The key point of the proposed method is to perform the clusters analysis of the image through a set of SOM prototypes instead of working directly with the original patterns of the scene. This approach significantly reduces the complexity of the analysis, making it possible to use methods that have not normally been considered viable for the processing of remotely sensed images, such as agglomerative hierarchical methods. Moreover, the proposed method presents advantages that make it as a promising alternative to carry out the classification of remotely sensed images. Among these, we can point out:

(a) The method does not require a previous definition of the number of classes to perform the classification of the image. It does not occur in the majority of the conventional unsupervised classification methods.

(b) The distributed representation of the classes by means of prototype groups gives the method the potential to discover geometrically complex and varied data clusters. Methods such as K-means use a single prototype (centroid) to represent each class and because of this are only capable of adequately detecting clusters that have hyperspherical formats.

(c) The method uses a particular approach to classify pixels situated in transition regions between classes. This procedure contributes to increasing the accuracy of the resulting classification.

(d) The utilization of an agglomerative hierarchical clustering method allows the user to observe the relationships that exist between the land cover patterns existing in the image at different cluster levels. It can be very helpful in applications where the structure of the information present in the image is not clearly known.

(e) The proposed method employs an efficient merge mechanism that incorporates more information about the data in each cluster. Traditional clustering methods use only inter-cluster distance information to decide on the merging of clusters. Integration of spatial characteristics helps to increase the understanding of some classes confusable with others.

In addition to the test image utilized in the experiments shown here, the proposed method has also been applied to other high and medium resolution images, with satisfactory results.

As future works, it is intended to apply modified versions of cluster validation indices, as considered in [2], to automatically determine the cutoff of the dendrogram (ideal number of clusters) for the image. Comparisons of performance with conventional classification methods and sensitivity analysis also must be executed.

## References

1. Costa, J.A.F., Netto, M.L.A.: Clustering of Complex Shaped Data Sets via Kohonen Maps and Mathematical Morphology. In: Proceedings of the SPIE Conference on Data Mining and Knowledge Discovery, Orlando, FL, vol. 4384, pp. 16–27 (2001)
2. Gonçalves, M.L., Netto, M.L.A., Costa, J.A.F., Zullo Júnior, J.: Data Clustering using Self-Organizing Maps Segmented by Mathematic Morphology and Simplified Cluster Validity Indexes. Proceedings of IEEE International Joint Conference on Neural Networks 1, 8854–8861 (2006)
3. Haralick, R.M., Shanmugam, K., Dinstein, I.: Textural Features for Image Classification. IEEE Trans. on Systems, Man and Cybernetics 3(6), 610–621 (1973)
4. Kaski, S., Lagus, K.: Comparing self-organizing maps. In: Vorbrüggen, J.C., von Seelen, W., Sendhoff, B. (eds.) Artificial Neural Networks - ICANN 96. LNCS, vol. 1112, pp. 809–814. Springer, Heidelberg (1996)
5. Kohonen, T.: Self-Organizing Maps, 2nd edn. Springer, Berlin (1997)
6. Kohonen, T., Oja, E., Simula, O., Visa, A., Kangas, J.: Engineering Applications of the Self-Organizing Map. Proceedings of the IEEE 84(10), 1358–1384 (1996)
7. Marçal, A.R.S., Castro, L.: Hierarchical Clustering of Multispectral Images using Combined Spectral and Spatial Criteria. In: IEEE Geoscience and Remote Sensing Letters 2, 59–63 (2005)
8. Richards, J.A.: Analysis of Remotely Sensed Data: the Formative Decades and the Future. In: IEEE Transactions on Geoscience and Remote Sensing 43, 422–432 (2005)
9. Vesanto, J., Alhoniemi, E.: Clustering of the Self-organizing Map. In: IEEE Transactions on Neural Networks 11, 586–602 (2000)
10. Wu, S., Chow, T.W.S.: Clustering of the Self-organizing Map using a Clustering Validity Index based on Inter-cluster and Intra-cluster Density. Pattern Recognition 37, 175–188 (2004)
11. Xu, R., Wunsch II, D.: Survey of Clustering Algorithms. In: IEEE Transactions on Neural Networks 16, 645–678 (2005)