

UMA NOVA TÉCNICA DE CLASSIFICAÇÃO DE IMAGENS BASEADA EM LÓGICA NEBULOSA NÃO ITERATIVA

WELLINGTON PINHEIRO DOS SANTOS², MOZART DE SIQUEIRA CAMPOS ARAÚJO FILHO¹,
ANTÔNIO JOSE BARROS NETO¹, CLEDJA KARINA ROLIM DA SILVA¹, PABLO VIANA DA SILVA¹,
JOÃO DINIZ DA SILVA³, ALEJANDRO C. FRERY¹

¹CIn – Universidade Federal de Pernambuco
Caixa Postal 7851, 50732-970 - Recife – PE, Brasil
{msca,ajbn,ckrs,pvs,rcf,frery}@cin.ufpe.br

² Departamento de Eletrônica e Sistemas - CTG - Universidade Federal de Pernambuco
Caixa Postal 7800, 50711-970 - Recife - PE, Brasil
wellington-ps@bol.com.br

³ Departamento de Engenharia Cartográfica - CTG - Universidade Federal de Pernambuco
CEP: 50740-530 - Recife - PE, Brasil
jdiniz@ufrpe.br

Abstract. This paper describes a new technique for the classification of hyperspectral imagery using Fuzzy Logic with non-interactive trapezoidal membership functions. The fuzzy classifier is implemented using a rule-based approach, where each class is defined as a set of sub rules and only one sub rule associated to each band within a class. This technique was implemented as an extension feature for the ENVI platform. Results were obtained using an AVIRIS image of Moffett Field, CA acquired from NASA's JPL Web Database. The classification error was negligible, combining precision and effective computing.

Keywords: remote sensing, image processing, fuzzy logic.

1. Introdução

Os recentes avanços na tecnologia dos imageadores hiperespectrais, que trabalham com 50, 100, 200 ou até mais bandas do espectro eletromagnético, possibilitaram uma significativa melhoria na resolução espectral, quando comparada com os sensores das séries SPOT e Landsat que utilizam sensores multiespectrais de apenas até seis bandas.

O surgimento da **Espectroscopia de Imageamento** ou **Sensoriamento Remoto Hiperespectral** permitiu representar, para cada *pixel* na imagem, diferentes tipos de alvos na cena (solo, vegetação e água), podendo-se extrair um espectro de refletância diretamente da imagem com uma resolução espectral próxima da obtida em condições ideais de laboratório. Quanto maior o número de bandas, maior é a possibilidade de se obter informações para uma reconstituição espectral que pode ser usada para a identificação de temas na cena.

Embora a resposta espectral de cada tema seja influenciada por vários fatores atuando em conjunto (matéria orgânica, concentração de elementos químicos, umidade), a refletância de certos elementos, em alguns intervalos espectrais, se sobrepõe à do tema em si, o que possibilita a sua identificação remota através de dados de alta resolução espectral.

Este trabalho descreve uma nova técnica de classificação de imagens hiperespectrais que utiliza Lógica Nebulosa com funções de pertinência trapezoidais, bem como a sua implementação eficiente em um sistema de produção. É apresentado também um estudo de caso utilizando imagens do sensor AVIRIS (*Airborne Visible InfraRed Imaging Spectrometer*), que fornece imagens da radiância espectral de alvos em 224 bandas de aproximadamente 0.01 μm (micrômetros) ou 10nm (nanômetros) entre as regiões do espectro de 0.4 μm e 2.45 μm . O sensor AVIRIS é um sensor hiperespectral aerotransportado operado pelo JPL (*Jet Propulsion Laboratory*) da NASA, que tem como finalidade a utilização em

várias áreas científicas, tais como botânica, geologia, hidrologia, oceanografia e ciências atmosféricas. A imagem utilizada foi obtida em 2002 e corresponde a área de Moffett Field-EUA (ver **Figura 1**), localizada geograficamente no retângulo inserido entre o ponto de latitude +37,44982 graus N/S e longitude -122,01176 graus L/O e o ponto de latitude +37,44999 graus N/S e longitude -121,90756 graus L/O.

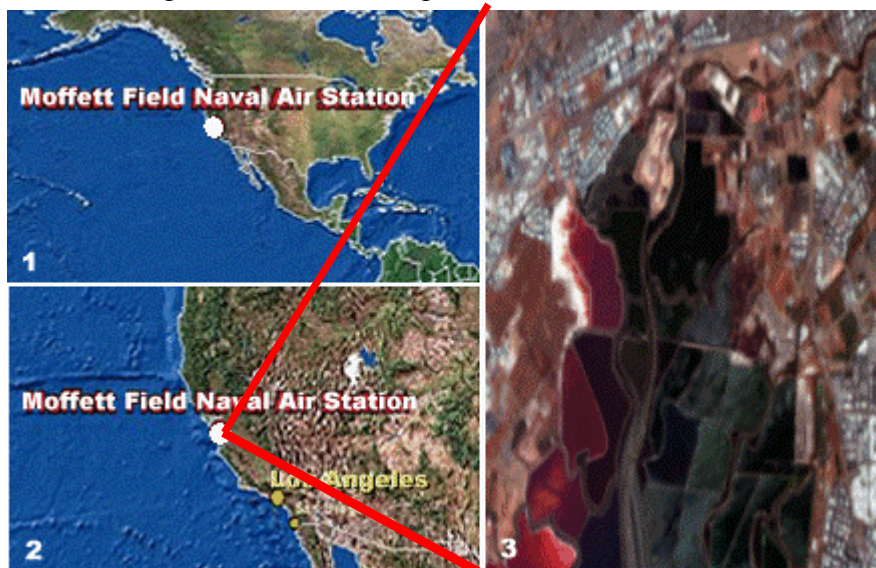


Figura 1: Visões aéreas de Moffett Field, CA, retirada de National e AVIRIS.

2. Seleção das Bandas Espectrais e das Regiões de Interesse

Para efetuar a classificação de regiões de uma imagem, tais como vegetação, água, área urbana e outras, não é necessário utilizar todas as 224 bandas espectrais que o sensor AVIRIS dispõe. Por isso, optou-se por escolher algumas bandas dentre aquelas que se encontrassem na faixa espectral do visível, ou seja, nas faixas do azul (B), 0.44 – 0.49 μm , do verde (G), 0.49 – 0.56 μm , e do vermelho (R), 0.63 – 0.78 μm .

A seleção das regiões de interesse foi realizada por meio de critérios visuais utilizando uma composição RGB das bandas escolhidas, associando as bandas 35, 16 e 9 da imagem AVIRIS às bandas R, G e B, respectivamente. Tal composição não apenas evidencia visualmente as características que diferenciam as regiões de interesse, como também resulta numa imagem nas cores naturais da área em estudo.

Dentro da faixa visível oito bandas são suficientes para classificar as regiões de interesse. Assim, sendo I o centro da faixa da banda, as bandas do AVIRIS selecionadas foram (1) Azul: bandas 8 ($I = 0.45122 \mu\text{m}$) e 10 ($I = 0.47076 \mu\text{m}$); (2) Verde: bandas 13 ($I = 0.50014 \mu\text{m}$), 16 ($I = 0.52958 \mu\text{m}$) e 19 ($I = 0.55907 \mu\text{m}$); (3) Vermelho: bandas 28 ($I = 0.64765 \mu\text{m}$), 35 ($I = 0.68244 \mu\text{m}$) e 43 ($I = 0.75909 \mu\text{m}$).

3. Proposta: Algoritmos de Classificação Nebulosos

Os algoritmos de classificação propostos são nebulosos e não iterativos baseados em funções de pertinência trapezoidais. Estes algoritmos permitem classificar imagens hiperespectrais a um baixo custo computacional.

Uma imagem digital $f : [0, \dots, m-1] \times [0, \dots, n-1] \rightarrow W^p$ pode ser definida como uma função que mapeia uma grade $m \times n$ num espaço real de dimensão p . Diz-se que f é uma imagem de tamanho $m \times n$ composta de p bandas. É comum representar a imagem f por

uma matriz $m \times n \times p$. Cada elemento p -dimensional $f(i, j)$, onde $0 \leq i \leq m-1$ e $0 \leq j \leq n-1$, é considerado um vetor de atributos p -dimensional.

Seja $\mathbf{X} = (x_1, x_2, \dots, x_p) \in \mathbf{R}^p$ um vetor de p atributos reais correspondente a um pixel qualquer da imagem f , onde a componente x_j corresponde ao valor do referido pixel na j -ésima banda. Seja também $\Omega = (C_1, C_2, \dots, C_M)$ o conjunto das M classes. Cada classe C_k está associada a uma *função discriminante real* $g_k : \mathbf{R}^p \times \Omega \rightarrow \mathbf{R}$ que retorna o grau de pertinência do vetor de atributos \mathbf{X} à classe C_k , ou seja, “o quanto” de \mathbf{X} pertence a C_k . A classificação é uma decisão tomada segundo o critério da classe vencedora (Duda, Hart e Stork, 2001), como segue:

$$g_k(\mathbf{X}) = \max(g_1, g_2, \dots, g_M) \Rightarrow \mathbf{X} \in C_k \quad (1)$$

Em caso de empate, é indiferente tomar a decisão por uma ou outra dentre as que empataram.

Considerando $M-1$ classes de interesse, define-se a M -ésima classe como sendo a *classe de rejeição*. Cada classe C_k é caracterizada por um conjunto de N regras nebulosas (Roubos et. al., 2001) caracterizadas por conjuntos A_{kj} , onde $1 \leq k \leq M-1$ e $1 \leq j \leq N$:

$$\{x_1 \in A_{k1}\} \wedge \{x_2 \in A_{k2}\} \wedge \dots \wedge \{x_N \in A_{kN}\} \Rightarrow \mathbf{X} \in C_k \quad (2)$$

Por sua vez, cada conjunto A_{kj} está associado a uma *função de pertinência nebulosa* $u_{kj} : \mathbf{R} \rightarrow [0,1]$. As funções de pertinência representam o *grau de pertinência* do valor x_j ao conjunto A_{kj} , isto é, quanto a componente x_j está no conjunto A_{kj} . Observa-se que, no caso específico de classificação de imagens de sensoriamento remoto, o número de regras nebulosas que definem uma classe como da forma dada na equação (2) é igual ao número p de bandas presentes na imagem.

Bezdek (1981) afirma que a Lógica Nebulosa pode ser considerada uma extensão natural da Álgebra Booleana, onde a operação “e”, denotada “ \wedge ”, é equivalente à operação de mínimo e pode ser associada à intersecção entre conjuntos, enquanto que a operação “ou”, denotada “ \vee ”, é a extração de máximos, associada à união de conjuntos. Assim, a função discriminante g_k da classe C_k é definida como segue, onde $1 \leq k \leq M-1$:

$$g_k(\mathbf{X}) = \bigwedge_{j=1}^N u_{kj}(x_j) \equiv \min(u_{k1}(x_1), u_{k2}(x_2), \dots, u_{kN}(x_N)) \quad (3)$$

A classe C_M é a *classe de rejeição* e sua função discriminante g_M é calculada utilizando que C_M é definida como $C_M = \overline{C_1 \cup C_2 \cup \dots \cup C_{M-1}}$, logo

$$g_M(\mathbf{X}) = 1 - \max(g_1(\mathbf{X}), g_2(\mathbf{X}), \dots, g_{M-1}(\mathbf{X})). \quad (4)$$

As funções de pertinência u_{kj} aqui propostas são trapezoidais e definidas como

$$u_{kj}(x) = \begin{cases} 0, & x < a_{kj} \vee x > d_{k,j} \\ \frac{(x - a_{kj})}{(b_{kj} - a_{kj})}, & a_{kj} \leq x < b_{kj} \\ 1, & b_{kj} \leq x \leq c_{kj} \\ \frac{(x - d_{kj})}{(c_{kj} - d_{kj})}, & c_{kj} < x \leq d_{kj} \end{cases} \quad (5)$$

A proposta de funções trapezoidais é feita para capturar possíveis assimetrias nos histogramas dos dados em cada banda, através dos parâmetros a_{kj} , b_{kj} , c_{kj} e d_{kj} que, por sua vez, podem ser calculados por diversos métodos.

Neste trabalho propomos dois métodos não iterativos para obtenção desses parâmetros: o *Método do Trapézio Simétrico* e o *Método da Mediana*. Tais métodos permitem a obtenção dos parâmetros das funções de pertinência trapezoidais de forma não iterativa, a partir de estatísticas extraídas de cada banda das amostras (regiões de interesse).

Seja $\{x_j^k(i): 1 \leq i \leq N_k\}$ a amostra de tamanho N_k da j -ésima banda da região de interesse R_k correspondente à classe C_k . As estatísticas de interesse são a *média amostral* \bar{m}_{kj} , o *desvio padrão amostral* s_{kj} , a *mediana amostral* h_{kj} , o *valor máximo observado* e o *valor mínimo observado*.

No *Método do Trapézio Simétrico* a função de pertinência trapezoidal é centrada na média amostral e os parâmetros são obtidos da forma que segue:

$$a_{kj} \equiv \bar{m}_{kj} - 2s_{kj} \quad (6)$$

$$b_{kj} \equiv \bar{m}_{kj} - s_{kj} \quad (7)$$

$$c_{kj} \equiv \bar{m}_{kj} + s_{kj} \quad (8)$$

$$d_{kj} \equiv \bar{m}_{kj} + 2s_{kj} \quad (9)$$

No *Método da Mediana* a função trapezoidal é transformada numa função triangular com ponto de máximo na mediana, sendo os extremos a_{kj} e d_{kj} fixados no valor mínimo e máximo da amostra, respectivamente. Portanto, os parâmetros são calculados como:

$$a_{kj} \equiv \min(x_{1j}, x_{2j}, \dots, x_{N_jj}) \quad (10)$$

$$d_{kj} \equiv \max(x_{1j}, x_{2j}, \dots, x_{N_jj}) \quad (11)$$

$$b_{kj} = c_{kj} = h_{kj} \quad (12)$$

A maior motivação para a implementação de um método baseado na mediana é gerar um algoritmo nebuloso que não parta da premissa que os dados nas bandas das regiões de interesse são simetricamente distribuídos a custo aceitável.

4. Implementação em Software

Para implementação foram usadas as ferramentas ENVI e IDL (Research Systems Inc). ENVI, *Environment for Visualizing Images*, é uma ferramenta para processamento de imagens, que possibilita, entre outras funções, a visualização e gerenciamento de imagens multi- e hiper-espectrais, bem como a criação e o gerenciamento de regiões de interesse. Além disto, inclui funções básicas de processamento de imagens, tais como funções de análise espectral, classificação e filtros e correções geométricas. IDL, *Iterative Data Language*, é uma linguagem matricial, e que por isto oferece operações otimizadas, o que facilitou muito a implementação dos métodos aqui propostos. ENVI foi desenvolvido sob IDL, por esta razão a integração destas duas ferramentas possibilitou o compartilhamento de funções de ambas.

4.1. Algoritmos de classificação

Obtidos os pixels das regiões de interesse, inicia-se o treinamento que, para o primeiro algoritmo, consiste apenas no cálculo dos dois primeiros momentos amostrais dos valores dos pixels e, a partir destes momentos, todos os pixels são classificados baseado no cálculo da *função de pertinência* definida na equação (5).

O método do trapézio simétrico, por não considerar a assimetria dos dados, é bem mais simples, do ponto de vista computacional, e rápido de ser implementado e executado, se comparado ao método da mediana que requer da ordem de $N_k \ln N_k$ operações para calcular cada mediana. Para simples ilustração de como a eficiência pode vir a ser um gargalo no processo de treinamento, o fato de se classificar 8 bandas de uma imagem de 512 linhas por 512 colunas requer 2097152 avaliações da função de pertinência, o que pode ser inaceitável.

Para otimizar os algoritmos, pode ser feito o pré-processamento de todos os possíveis valores para cada banda (denotada j) e região de interesse (k) aproveitando o fato de que os valores podem ser representados por até um inteiro de 16 bits. Esta abordagem tem um requerimento de memória de $2^{\text{numero_de_bits}} jk$, o que é aceitável. Isto permite fazer o cálculo da função de pertinência para todos os possíveis valores do dado a serem classificados durante o treinamento. O resultado é armazenado em uma *tabela de busca*, tornando a antes custosa avaliação da função de pertinência em uma simples consulta. Embora isso sobrecarregue um pouco a fase de treinamento, diminui notadamente o peso da fase de classificação. A **Figura 2** mostra em diagrama de blocos o sistema implementado.

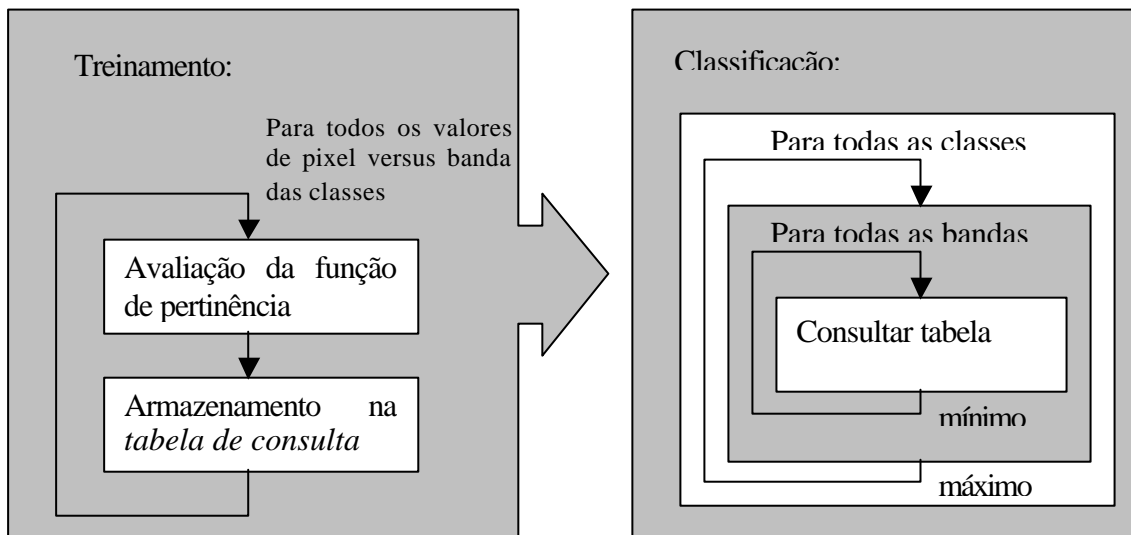


Figura 2: Visualização do processo de treinamento utilizando Tabela de Consulta

Ao comparar os métodos propostos com o de máxima verossimilhança Gaussiana do ponto de vista do custo computacional, os primeiros são muito mais eficientes pelo fato de requererem apenas operações aritméticas simples e de comparação. Os métodos propostos podem também ser implementados numa linguagem de programação matricial, como IDL.

Entretanto, a implementação imperativa em uma linguagem interpretada como IDL tornaria o tempo de execução inaceitável. Em testes feitos em um PC com processador Pentium III de 1GHz com 512 Mb de RAM este algoritmo levou *cinco minutos e vinte e dois segundos* para uma imagem de tamanho 512x512 e uma banda apenas. Apesar de termos ganho tempo com a melhoria da *tabela de busca*, executando este algoritmo em apenas *dois minutos e trinta segundos*, este tempo mostra-se ainda grande.

Com isto nos sentimos impelidos a usar melhor o ferramental que uma linguagem matricial oferece para melhorarmos a eficiência de nosso algoritmo. Na **Tabela 1** definimos um possível pseudocódigo de nosso algoritmo.

```

Para cada uma das ROI's faça
  para cada uma das bandas faça
    pegue a banda
    para todos os elementos da banda calcule a função de pertinência
    guarde este valor em res
    minimize a função de pertinência para esta banda baseado no valor anterior
  fim
  maximize a função de pertinência para banda baseado no valor anterior e guarde em max
fim

para cada uma das ROI's faça
  compare res e com max e para todos os pixels iguais classifique nesta ROI
fim

```

Tabela 1: Algoritmo matricial para classificação nebulosa

No pseudocódigo mostrado na **Tabela 1** devemos entender que as operações sublinhadas representam operações atômicas para linguagens matriciais e, com isso, podem ser modeladas de tal modo que a execução do algoritmo seja mais rápida. Desta forma o algoritmo proposto se torna eficiente, demandando apenas *dez segundos* para classificar as oito bandas em três classes da imagem citada anteriormente.

5. Resultados

A avaliação visual da imagem AVIRIS utilizada foi possível pela montagem de uma composição colorida utilizando as bandas 35 (R), 16 (G) e 9 (B), percebendo-se a existência de doze classes de interesse. Entretanto, os algoritmos de classificação nebulosa implementados irão considerar a existência de 17 classes sendo a última a classe de rejeição. As classes identificadas são mostradas na **Tabela 2**.

Classe	Legenda
Vegetação costeira (margem do rio)	VG
Vegetação: verde	VR
Hidrografia (açude): azul	H1
Área construída	AC
Hidrografia	HD
Solo exposto	S1
Construção	CN
Cultura sem sombra	CS
Cultura com sombra	CC
Solo salino	SS
Solo exposto	S2
Solo exposto (outro tipo)	S3
Cultura 2	C2
Cultura 3	C3
Cultura 4	C4
Cultura 5	C5
Não classificada	NC

Tabela 2: Classes de interesse identificadas na imagem AVIRIS

Foram realizadas as classificações pelo método da Máxima Verossimilhança Gaussiana (MVG), pelo método do paralelepípedo e também pelos dois métodos propostos (método do trapézio simétrico e método da mediana), sendo obtidos seus respectivos coeficientes *kappa* e a taxa de acerto total (**Tabela 3**).

Método de Classificação	Taxa de Acerto	Índice Kappa
MVG	94,52%	0,94
Paralelepípedo	57,11%	0,55
Trapézio Simétrico	73,18%	0,71
Mediana	72,09%	0,70

Tabela 3: Taxas de acerto e coeficientes *kappa* dos métodos de classificação utilizados

Pode-se perceber da **Tabela 3** que, para os dados considerados (bandas, regiões de treinamento e de teste), o método que apresentou o melhor desempenho de classificação segundo o índice *kappa* foi o método MVG ($k = 0,94$ e taxa de acerto total de 94,52%). Entretanto, este é o método de maior custo computacional e também o que demanda maior tempo de execução, tanto na etapa de treinamento quanto na de classificação. No que segue será constatado que esta qualidade de classificação é realizada ao custo de ignorar pelo menos uma feição na imagem.

O método do paralelepípedo, ao contrário do anterior, apresentou os menores custo computacional e tempo de execução, tanto na classificação quanto no treinamento. Entretanto, a sua performance de classificação foi sofrível uma vez que sua taxa de acerto não chegou aos 60% e o índice *kappa* se mostrou relativamente pequeno.

Os dois métodos propostos apresentaram desempenhos bastante razoáveis (ver **Tabela 3**), pois seus índices *kappa* ultrapassaram 0,7 e a taxa de acerto total fica entre 72% e 74%. Além disso, o custo computacional associado à implementação dos métodos propostos é bem menor do que o da MVG, apesar de ser maior do que o do método do paralelepípedo. Portanto, uma primeira observação pode levar a concluir que os métodos propostos têm um desempenho intermediário entre o do método do paralelepípedo e o da MVG.

Porém, visualmente pode-se perceber que a redução da taxa de acerto dos métodos propostos decorre da alta incidência de pixels não classificados (pertencentes à classe de rejeição, na cor preta na carta temática), e não a erros de classificação, fato constatado ao se comparar as matrizes de confusão dos métodos aqui analisados. Os métodos de classificação do trapézio simétrico e da mediana se mostraram bastante rigorosos e seletivos. Esta afirmação pode ser verificada na figura que segue.

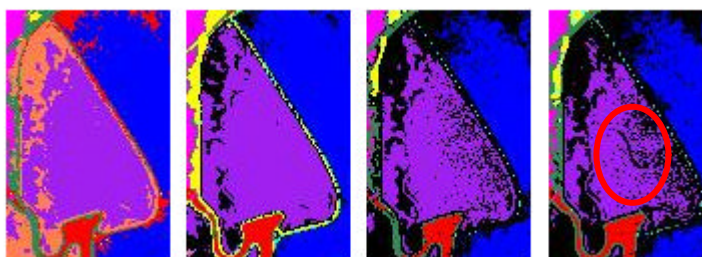


Figura 3: Destaque na região “Cultura sem sombra” (lilás) obtidas das classificações utilizadas

Na **Figura 3** é mostrada uma pequena região classificada segundo os métodos da MVG, paralelepípedo, trapézio simétrico e método da mediana, nesta ordem a partir da esquerda. A área de cor lilás corresponde à classe “Cultura sem sombra”, enquanto que a área em vermelho é “Vegetação costeira”, amarelo é “Área construída” e azul é a classe “Açude”. Pode-se notar que, enquanto a MVG associou algumas áreas à classe “Vegetação costeira” e o método do paralelepípedo reconheceu as mesmas áreas como “Área construída”, os métodos propostos (dois últimos a direita) atribuíram a classe de rejeição às mesmas regiões (cor preta), por perceberem que são classes diferentes daquelas previamente selecionadas.

Além disso, é possível perceber no centro da primeira imagem a partir da direita a existência de uma feição similar a um rio, o que dificilmente pode ser observado na etapa de determinação das classes de interesse. Esta feição é claramente reconhecida pelo método da mediana e frisada na **Figura 3** com a circunferência vermelha.

6. Conclusões

Pode-se concluir que é possível, utilizando métodos computacionalmente simples baseados em Lógica Nebulosa, construir classificadores com performance e taxa de acerto bastante razoáveis. Os métodos propostos se mostraram capazes de distinguir entre as classes selecionadas e a classe de rejeição, evitando falsa aceitação e indicando para o analista a existência de outras classes que não as previamente observadas quando da etapa de treinamento. Além disso, os métodos propostos permitiram otimizar a execução das etapas de treinamento e classificação, uma vez que utilizam em sua maioria operações lineares e comparações. A classificação foi implementada empregando uma abordagem matricial, permitindo a utilização das funções de comparação otimizadas da linguagem IDL.

Referências

- AVIRIS. Airborne Visible InfraRed Imaging Spectrometer. site Comercial. JLP-Jet Propulsion Laboratory, California. Disponível em: <http://popo.jpl.nasa.gov/html/aviris.spectroscopy.html>, 2001. Consultado em 29 de setembro de 2002.
- Bezdek, J. C. *Pattern Recognition with Fuzzy Objective Algorithms*. New York: Plenum Press, 1981.
- Duda, R.; Hart, P., Stork, D.G. *Pattern Classification*, segunda edição, New York: John Wiley & Sons, 2001.
- National Geographic Maps & Geography Online. Disponível: site *National Geographic*. URL: <http://www.nationalgeographic.com/maps>. Consultado em 30 de outubro de 2002.
- Roubos, J. A.; Setnes, M; Abonyi, J. *Learning Fuzzy Classification Rules from Labeled Data*. International Journal of Information Sciences, 2001.
- Research Systems Inc. Disponível: site *Research Systems Inc*. URL: <http://www.researchsystems.com>. Consultado em 26 de setembro de 2002.