

A Lógica Difusa e Métodos de Análise em Sistemas de Informação Geográfica

Prof. Dr. Czeslau L. Barczak ¹
 Prof. Dr. Miguel A. Verdinelli ²
 Prof. Dr^a. Martha E. P. Verdinelli ³
 Eng. Luiz Fernando G. de Figueiredo ⁴

¹ UFSC - Departamento de Engenharia Mecânica/GRUCON
 Campus Trindade 88010-970 Florianópolis SC
 ✉ czeslau@mbox1.ufsc.br

² UFSC - Departamento de Zootecnia/CCA
 Campus Trindade 88010-970 Florianópolis SC
 ✉ miguelv@eps.ufsc.br

³ UNISUL
 Rua Padre Roma, 404 Centro - Florianópolis SC
 ✉ martha@hu.ufsc.br

⁴ UFSC - Departamento de Engenharia de Produção
 Campus Trindade 88010-970 Florianópolis SC
 ✉ lff@eps.ufsc.br

Conteúdo	A Análise de Informações em Sistemas de Informação Geográfica Características dos Sistemas de Classificação Classificadores Dicotômicos Incertezas na Classificação em Imagens de Sensoriamento Remoto Classificadores Difusos Conjuntos Difusos e Funções de Pertinência Características dos Sistemas Difusos Estudo de Caso Conclusões Referências Bibliográficas
-----------------	--

Resumo: Neste artigo, a estrutura fundamental da lógica difusa é apresentada, fazendo-se uma análise de sua adequação como um método eficiente de busca aplicável a sistemas de informação geográfica (SIG). Suas potencialidades e limitações como ferramenta para construir sistemas de busca e de decisão são analisadas brevemente.

Palavras Chaves: SIG, Classificação, Lógica Difusa

Abstract: In this paper, the basic structure of fuzzy logic and its adequacy as an efficient search method for use in Geographic Information System (GIS) are presented. Its potentials and its limitations as a tool for the implementation of search and decision systems are briefly analysed.

Keywords: GIS, Classification, Fuzzy Logic

A Análise de Informações em Sistemas de Informação Geográfica

Um Sistema de Informação Geográfica - SIG compreende a aquisição, armazenamento, apresentação e análise de informações geográficas muito diversificadas. Implementado em computadores, tem como elementos principais dois tipos de arquivos. Um banco de dados contém uma descrição da geografia da superfície, ou seja, dados sobre a forma e posição de cada local da superfície. Um segundo banco de dados contém os atributos da superfície terrestre, isto é, dados que descrevem as características ou qualidades de cada local. Os bancos podem ser independentes ou ambos os tipos de dados.

Utilizando-se um SIG pode-se converter as informações disponíveis para formas adequadas a usos específicos, seja na forma de relatórios, tabelas e similares seja na forma de mapas, contendo os dados requisitados por um usuário. Os requisitos podem vir a ser extremamente variados e complexos, necessitando de processamentos elaborados de análise sobre os dados daqueles bancos.

Os Sistemas de Informação Geográfica tem sido usados principalmente como suporte para o gerenciamento de recursos ambientais, muitas vezes, com base em algum tipo de sistema de decisão. Os sistemas ambientais são sistemas dinâmicos não lineares e variantes no tempo, portanto, de extrema complexidade. Sujeitos, ainda, à ação humana, constituem um problema eivado de incerteza e que envolve elevado risco na tomada de decisões. A análise dos dados acumulados é, portanto, uma fase essencial deste processo.

Embora a visão e o cérebro humanos apresentem melhores capacidades para a classificação e interpretação de informações como as adquiridas por sensoriamento remoto, os processos computacionais são mais rápidos. Por isso, diversos procedimentos estão sendo implementados em "software" procurando-se desenvolver sistemas de classificação, e também de decisão, que possam auxiliar o homem na solução de problemas de alocação de recursos e gerenciamento do meio ambiente de modo a não esgotá-los ou torna-lo inabitável.

Características dos Sistemas de Classificação

Nos estudos de ecossistemas complexos por meio do sensoriamento remoto, única técnica que pode ser usada para medir as características físicas e biológicas das espécies em regiões de grande área (Ramsey III 1996), as assinaturas espectrais de vegetação obtidas por satélites podem fornecer informações para a análise das estruturas, distribuição e funções das diversas espécies. Conforme é mostrado na Figura 1 a título de ilustração, ao se classificar por métodos dicotômicos áreas costeiras que contêm vegetação típica de mangue, haverá uma dicotomia entre as faixas de vegetação distinta e ainda a de água. No entanto, em um mapa elaborado desse modo não se poderá caracterizar as áreas comuns, onde ocorre uma mistura de objetos de diferente natureza: água e arbórea de mangue, arbórea de mangue e herbácea, herbácea e arbustiva, etc.

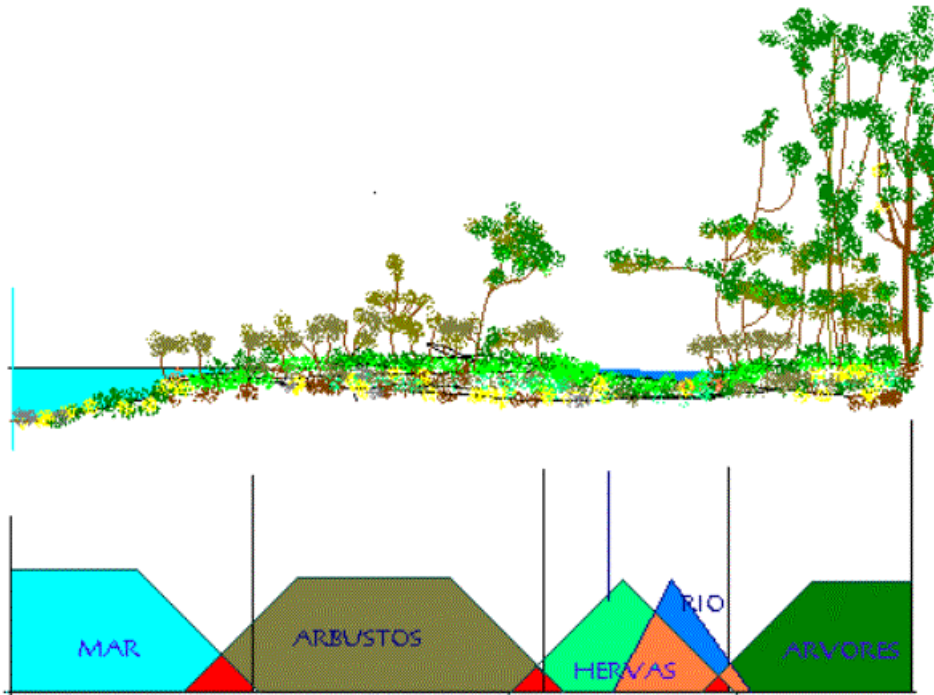


Fig. 1: Perfil Esquemático de Vegetação com Influência Flúvio-Marinha (Mangue)

Pode-se introduzir um número maior de classes porém os custos de análise podem tornar-se proibitivos. Mesmo que se defina mais classes, uma mistura de classes ou de subclasses continuará existindo na imagem a ser classificada. Classificadores como o de máxima verosimilhança ou de máxima probabilidade, de probabilidade condicional ou de Bayes, de mínima distância, hierárquico e outros estão baseados no mesmo método de representação da informação e apresentam, portanto, os mesmos problemas de decisão no sentido de que cada pixel ou pertence ou não pertence a uma dada classe.

Nos gráficos da figura 1 mostra-se uma tentativa de definição de classes onde se destaca um grau de pertinência das áreas à certas classes de objetos, uma tentativa de definição de funções de pertinência.

Classificadores Dicotômicos

Várias ferramentas utilizadas em análise e na construção de bancos de dados ou na elaboração de mapas, envolvem operações lógicas (E, OU, NÃO), geralmente realizadas com operadores lógicos dicotômicos, como na lógica de Boole. Classificadores dicotômicos permitem construir um mapa com pixels sim-não, isto é, cada pixel contém a informação relativa à existência de determinado atributo ou de sua não existência. A informação está representada no método um-pixel-uma-classe que não pode descrever adequadamente mistura de objetos de distintas classes (Wang 90).

Se o universo de discurso da informação é representado por X com elementos genéricos

$$x: X = \{x\} \quad (1)$$

na lógica clássica, a função de pertinência f_A de um conjunto A de X pode ser vista como uma função característica tal que

$$f_A(x) = 1 \text{ se e apenas se } x \in A. \quad (2)$$

Em outras palavras, se um objeto pertence a uma dada classe A então $f_A(x) = 1$; em caso contrário, ou seja, se não pertence à classe A então $f_A(x) = 0$.

Incertezas na Classificação em Imagens de Sensoriamento Remoto

Na classificação de objetos de imagens da superfície terrestre há duas fases distintas de processamento. Na fase de definição de classes, são delimitadas regiões da imagem, a cada região assinalada uma designação e uma representação de cada uma por meio do valor (centróide) da região. Na segunda fase, cada objeto da imagem é colocado em uma classe segundo os atributos típicos observados no objeto e na classe escolhida.

Nesta classificação supervisionada são definidas áreas que serão usadas para treinamento do classificador, com base em imagens de grande contraste, a partir das quais se procura extrair assinaturas que caracterizem as classes. Um procedimento de classificação pode ser escolhido e aplicado à imagem (Eastman 1997). Entretanto, os procedimentos de classificação podem não reconhecer o fato de que existe variação de refletância, isto é, não se inclui o fato de que algumas classes tem maior variação que outras. A normalização pode ser um modo de contornar tal problema. O procedimento de máxima verosimilhança é baseado em probabilidades mas depende da qualidade e homogeneidade dos locais (pixels) usados como padrões para treinamento do sistema. Em alguns métodos, paralelogramos e retângulos são usados para delimitar as áreas mas podem se sobrepor deixando áreas de interpretação ou classificação ambíguas, pois pertencerão simultaneamente a distintas classes.

Incerteza em uma medição sobre a pertinência de um pixel a uma certa classe pode existir mesmo que cada classe seja definida de modo dicotômico (abrupto). A origem da incerteza pode ser distinta e ser difusa apesar de não se usar lógica difusa.

Um sistema classificador dicotômico permite determinar se um pixel pertence a uma classe ou a outra. A lógica associada ao classificador é Aristotélica (ou Booleana), isto é, o pixel pertence (verdade) ou não pertence (falso) àquela classe.

Existem dois tipos de incertezas na classificação. Uma está relacionada com a pertinência do objeto a uma determinada classe de objetos. Um segundo tipo refere-se à incerteza de que determinado objeto pertence aos conjuntos de classes definidas. No primeiro caso, um pixel contém uma mistura de diversas categorias de atributos ou assuntos e que não pode ser separado e incluído em uma dada classe. No segundo caso, um pixel pode não conter elementos ou assuntos que sejam reconhecidos pelo sistema classificador. Como estas questões estão relacionadas com a informação contida no pixel e não um defeito no sistema classificador, uma solução é um levantamento no local para fins de identificação ou mesmo de criação de novas

Como resultado, a lógica dicotômica não é suficiente para o processamento das informações, incorporando erros que produzem interpretações e decisões errôneas, como nas áreas onde diversos casos relacionados a fronteiras de distintos atributos estão presentes.

Classificadores Difusos

Um sistema classificador difuso expressa o grau de pertinência de um pixel a uma e à outra ou outras classes.

No campo dos sistemas de informação geográfica os conjuntos difusos (Fuzzy Sets) podem ser utilizados para o processamento de problemas de mapeamento de adequação de áreas, isto é, determinar a que classe de solo pertence uma área ou se uma área pode ser utilizada para um determinado fim.

Diferentemente dos modelos matemáticos e dos sistemas peritos, na modelagem de um sistema difuso estuda-se, primeiramente, o comportamento do sistema real como um todo. As variáveis que afetam o comportamento podem envolver informações endógenas, internas ao sistema e sobre as quais se pode ter algum controle ou exógenas, independentes do sistema. Busca-se identificar as variáveis que, efetivamente, interferem no comportamento do sistema. E em um universo de variáveis qualitativas e quantitativas, contínuas e discretas, densas ou esparsas, precisas ou imprecisas, deve-se identificar aquelas que constituem a causa do comportamento (entradas) e aquelas que descrevem o comportamento (saídas).

Conjuntos Difusos e Funções de Pertinência

Um modelo difuso processa variáveis difusas. Conjuntos difusos devem ser definidos. Descrevem o mapeamento do processo realizado pelas regras que constituem a base de conhecimento do sistema difuso.

As variáveis de entrada devem sofrer um processo de *fuzzyficação*, isto é, devem ser convertidas de valores determinísticos para valores difusos usando-se, para isso, as funções de pertinência.

Em um sistema de classificação supervisionada áreas de treinamento do algoritmo são demarcadas. Os pixels podem conter essencialmente objetos de determinada classe mas podem também pertencer em menor grau a uma outra ou a outras classes de objetos, conforme Figura 1. Um algoritmo baseado em lógica difusa será caracterizado por conjuntos difusos.

Um conjunto difuso B em um universo de discurso X como definido por (1) é caracterizado por uma função de pertinência f_B que associa a cada x um número real em [0, 1], onde $f_B(x)$ representa o grau de pertinência de x em B. Se o valor $f_B(x)$ se aproxima de 1 significa que o elemento x pertence mais ao conjunto B. Se, ao contrário, $f_B(x)$ se aproxima de 0 o elemento x pertence menos ao conjunto B. Não há, portanto, uma definição dicotômica dos limites de um conjunto pois o grau de pertinência pode variar continuamente entre 0,0 e 1,0. Além disso, um elemento pode pertencer parcialmente a um único ou a vários conjuntos difusos.

Um classificador politômico deve utilizar uma representação em partições, um mapa ou matriz para cada classe onde cada pixel recebe um grau de pertinência. Para informações obtidas por sensoriamento remoto, pode-se usar partições difusas em um espaço espectral às quais um pixel poderá pertencer segundo um certo grau a uma ou várias faixas espectrais (Wang 90). Definindo-se o espaço espectral por

$$X = \{x_1, x_2, \dots, x_n\} \quad (3)$$

uma partição difusa do espaço espectral será representada por uma família de conjuntos difusos F_1, F_2, \dots, F_k . Um elemento x_i no conjunto difuso F_j terá um grau de pertinência u_{ij} , ou seja

$$u_{ij} = f_{F_j}(x_i) \quad (4)$$

Interessante observar que uma partição dicotômica pode ser obtida de uma partição difusa trocando-se os valores máximos em cada coluna da matriz por 1 e todos os outros por 0.

Características dos Sistemas Difusos

Embora tenham limitações, os sistemas difusos apresentam peculiaridades interessantes cujas principais características são:

Estimam funções entrada-saída;

Estimam funções sem o uso de modelos matemáticos;

Aprendem por experiência, a partir de dados numéricos e/ou de dados lingüísticos;

Processam informações em uma estrutura paralela;

Podem ser implementados em software, em circuitos digitais e analógicos, em hologramas, em dispositivos óticos;

Processam informações de conhecimento estruturado em forma de regras;

Processam informações imprecisas, vagas, nebulosas, incompletas;

São muito robustos;

Sistemas difusos adaptativos podem ser construídos e associados a outras técnicas de busca ou classificação como redes neurais e algoritmos genéticos, por exemplo.

A construção de um sistema difuso não é, em si, um processo complexo. Mesmo aplicado na modelagem de sistemas reais complexos, envolve algumas tarefas básicas mais ou menos comuns a qualquer aplicação:

Escolha das variáveis antecedentes e conseqüentes (de entrada e de saída);

Definição dos conjuntos difusos (funções de pertinência);

Estabelecimento das relações entre as variáveis (regras);

Interpretação da solução (obtenção de valor determinístico).

Simulação do sistema (ensaios);

Para se chegar aos itens apontados, algumas decisões preliminares precisam ser tomadas. Qualquer modelo de um sistema real, difuso ou de outra natureza, envolve algum tipo de simplificação. Um sistema real complexo pode ser decomposto em múltiplos subsistemas, permitindo a construção de uma estrutura que facilita a análise e o projeto embora possa, algumas vezes, causar a ocultação de alguma característica ou relação importante. Seguindo as técnicas usuais de modelagem, primeiramente procura-se definir qual é a informação que flui para dentro do sistema difuso, quais são as transformações que essa informação sofre e qual o resultado que eventualmente o sistema difuso fornece. Ou seja, uma relação de causa e efeito é estabelecida.

Estudo de Caso

A região costeira da Praia da Pinheira, em Santa Catarina, foi escolhida para ilustrar a utilização de classificador difuso. Na geração das figuras mostradas no texto, foram usadas as imagens das faixas 1, 2, 3 e 4 obtidas por satélite LANDSAT, abrangendo o espectro visível e o infravermelho próximo. As áreas de treinamento foram escolhidas em imagem da faixa 4 fazendo-se um realce para melhor contraste e está mostrada em 256 tons de cinza. Os locais escolhidos como áreas de treinamento são de constituição conhecida e estão ilustradas na mesma figura 2.

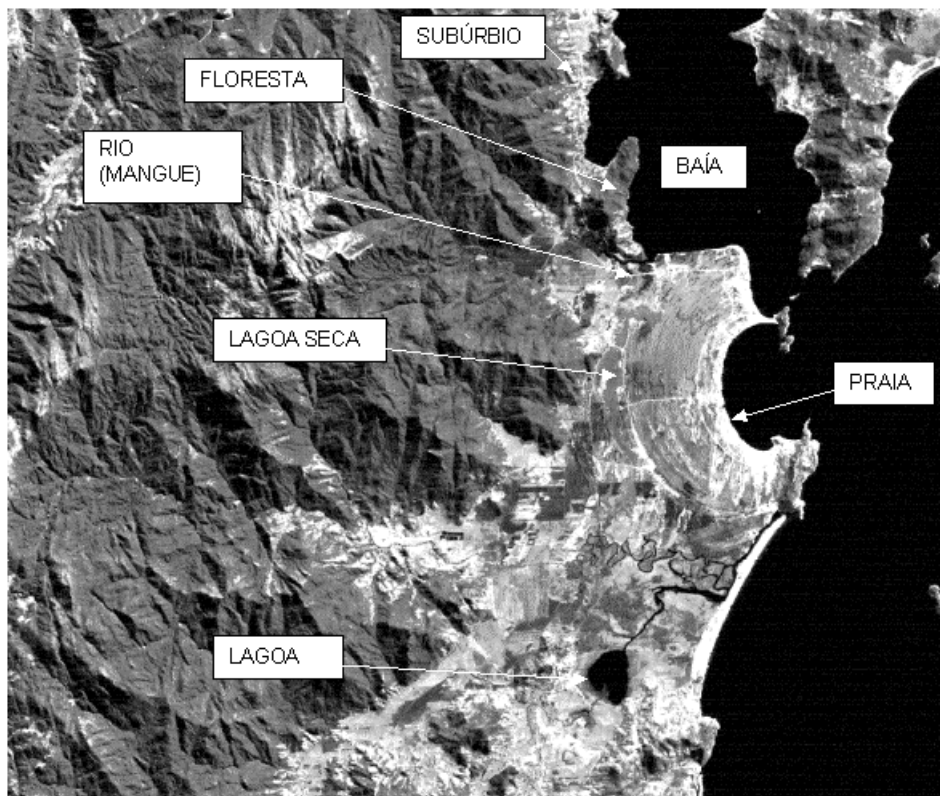


Fig. 2: Praia da Pinheira SC - Imagem LANDSAT da Banda 4

Na figura 3 mostra-se a classificação feita por verossimilhança, um procedimento dicotômico. Nota-se grandes áreas classificadas como subúrbio mas que são áreas do Parque da Serra do Tabuleiro (em verde claro) e áreas extensas classificadas como lagoa seca (em vermelho) embora esta seja limitada a uma área estreita e bastante menor entre a lagoa (em azul escuro na parte inferior da imagem) e a área escolhida como elemento de rio ou mangue, mais acima. A causa possivelmente está na falta de cuidado na escolha dos locais de treinamento e no reduzido número de classes.

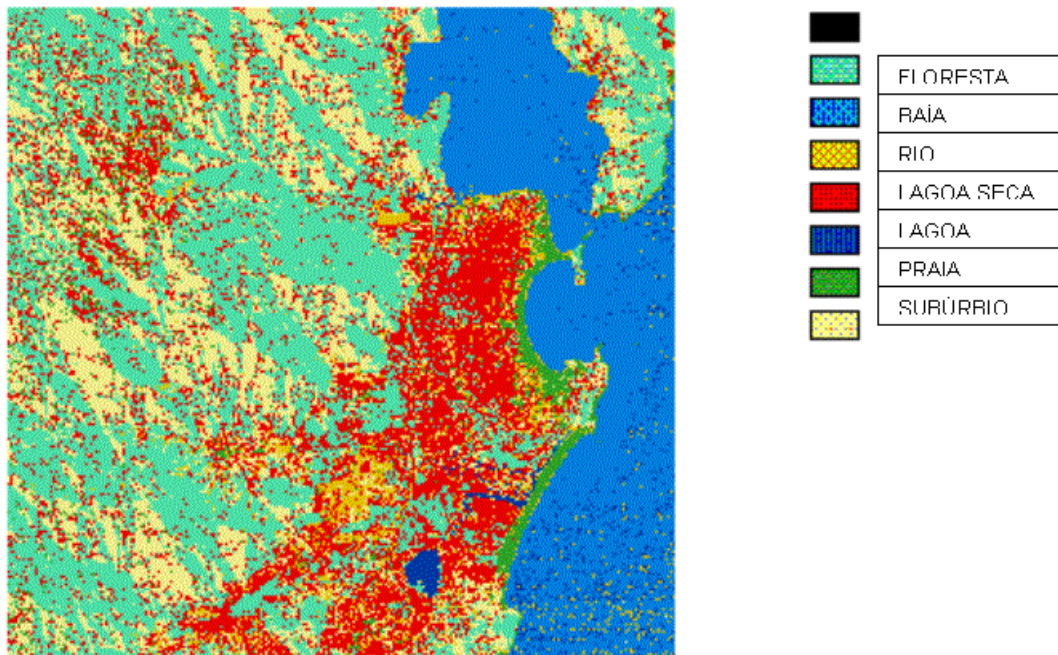
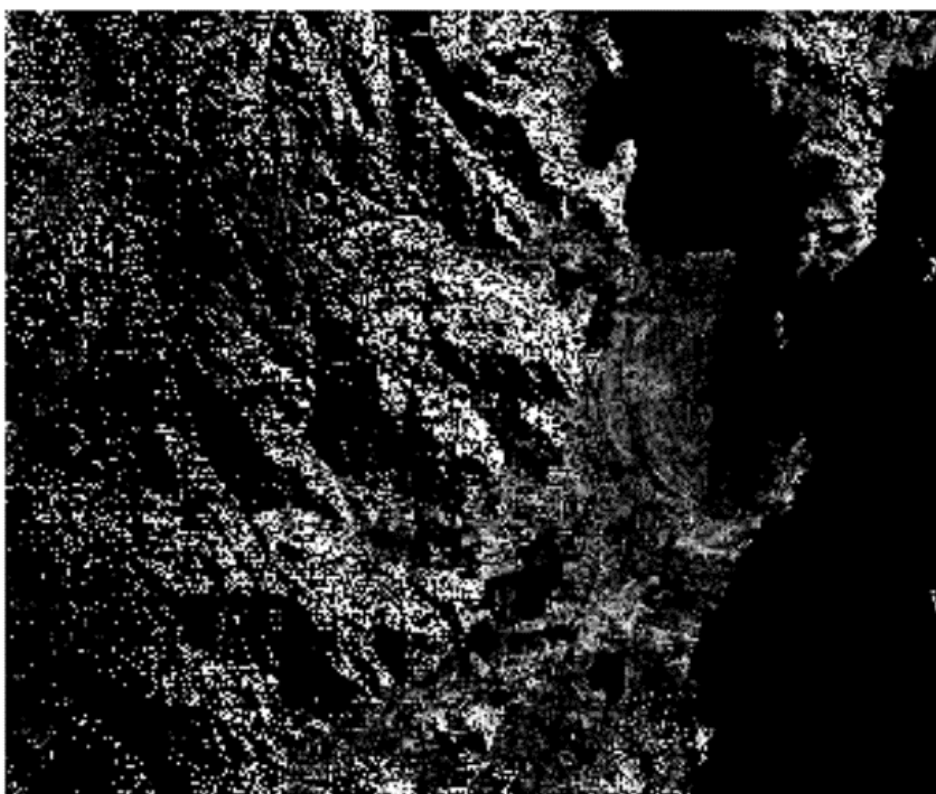


Fig. 3: Praia da Pinheira - Classificação por Verossimilhança

Nas figuras 4 e 5 estão mostradas as classificações por procedimento que utiliza a lógica difusa (em 16 tons de cinza: Pixel Preto=0.0; Pixel Branco=1.0). Os locais de treinamento são os mesmos usados pelo procedimento de máxima verossimilhança. Como as classes constituem conjuntos difusos, objetos poderão estar incluídos em mais de uma classe. Por esta razão é gerada uma imagem para cada classe. As classes não são mutuamente exclusivas como na classificação dicotômica, isto é, não se pode usar diferentes cores para cada classe em uma mesma imagem. Apenas as classes Floresta e Lagoa Seca são mostradas como exemplo.

Fuzzy Image Classification for flo



Fuzzy Image Classification for Isc

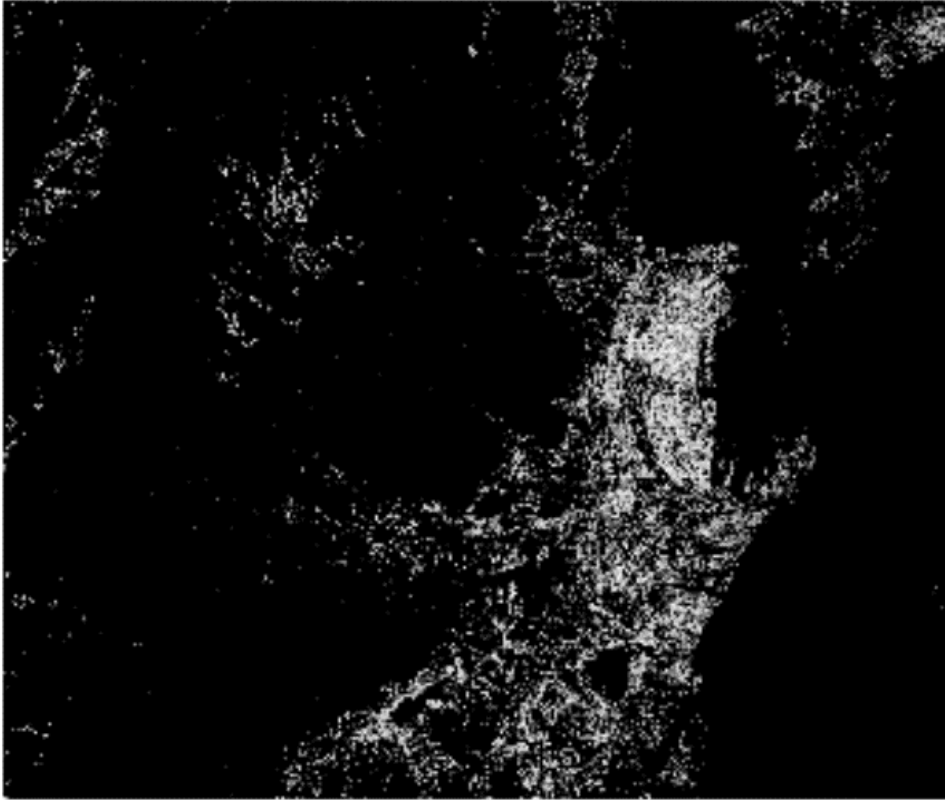


Fig. 5: Praia da Pinheira - Classificação Difusa de Lagoa Seca

Nota-se que muitas áreas são classificadas como Floresta e também como Lagoa Seca, tal como ocorreu na classificação dicotômica. O mesmo ocorreu na classificação difusa relativa às outras classes que não estão mostradas. Na figura 6 mostra-se um mapa de incerteza na classificação difusa na qual as áreas em branco indicam elevado índice de incerteza.

Classification Uncertainty

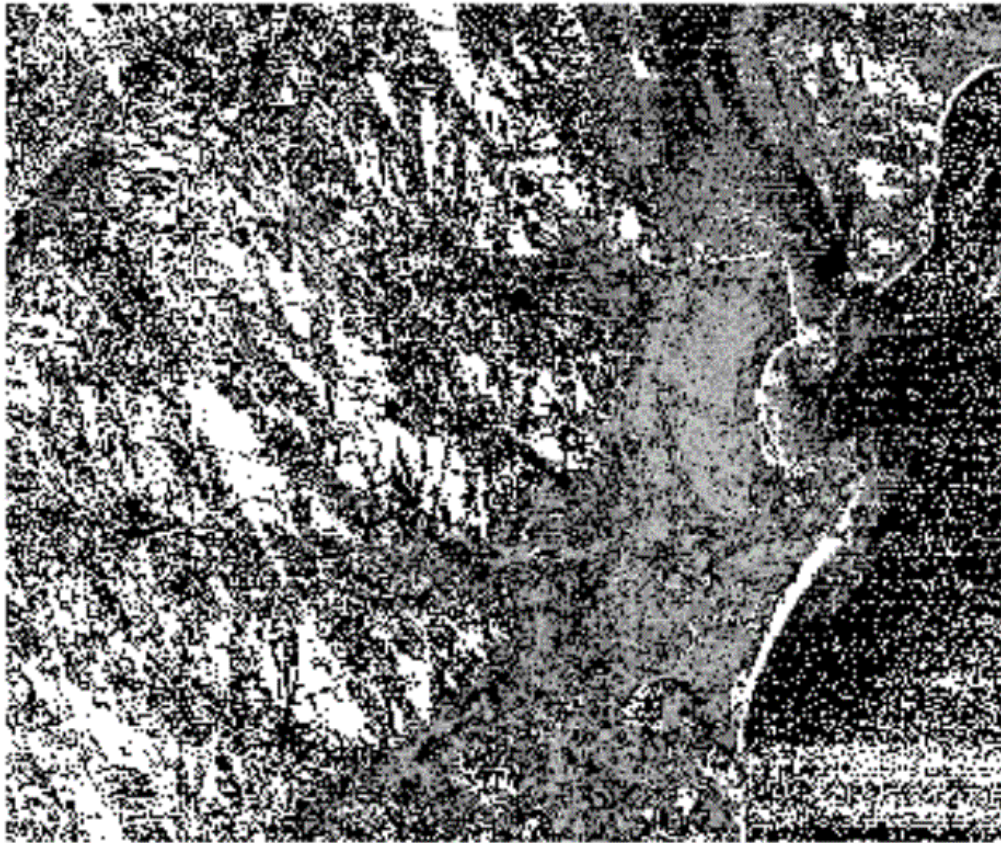


Fig. 6: Mapa de Incerteza na Classificação Difusa

Há inúmeras limitações na aplicação de classificadores dicotômicos e politômicos (Eastman 1997; Zhu 1997; Centeno 1995). O algoritmo de lógica difusa utilizado neste artigo (Eastman 1997) é baseado na lógica de mínima distância portanto a lógica subjacente difere da usada pelo de máxima verosimilhança e os resultados de ambas técnicas não podem ser comparados em termos de eficiência pois usam métodos distintos.

Algumas limitações específicas dos modelos difusos podem, entretanto, ser esclarecidas:

O desenvolvimento de um modelo a partir de um sistema difuso é muito difícil;

Sistemas difusos necessitam de muitas simulações e ajustes até se tornarem operacionais;

Sistemas difusos estão limitados pela expansão exponencial de possibilidades n^m , onde n é o número de variáveis e m o número de conjuntos difusos (funções de pertinência) definidos para cada variável.

A aplicação de sistemas difusos de classificação apresenta os mesmos tipos de problemas dos classificadores dicotômicos relacionados às incertezas associadas aos procedimentos de interpretação das imagens fornecidas pelos sistemas de sensoriamento remoto. Ambigüidade pode originar-se de diferentes fontes, não apenas de conjuntos difusos (Fuzzy Sets). Pixels de distintas refletâncias serão obtidos em uma imagem de solo, por exemplo, de plantação não uniforme no espaço de algum vegetal. Embora exista uma variação entre os pixels, que poderão ter graus de refletância desde 0,0 (mínimo número de plantas/pixel) até 1,0 (máximo número de plantas/pixel) não há nenhuma ambigüidade quanto ao fato de que se trata da mesma planta. É uma única classe de vegetal e o problema é o de classificar distintos tipos de pixels do mesmo tipo de objeto. As imagens obtidas por sensoriamento remoto fornecem informações sobre a reflectância dos objetos da superfície terrestre, em diversos comprimentos de onda da energia refletida por esses objetos, esperando-se que exista alguma diferença nas imagens para distintos objetos. Um pixel que tenha a mesma reflectância de certa classe, em todas as faixas de frequência, será classificado como pertencendo a essa classe. Entretanto nada garante que se trate do mesmo tipo de objeto; pode tratar-se de distintos objetos com a mesma reflectância em todas as frequências pesquisadas.

Muitos outros detalhes relativos à classificação e os problemas que lhe estão afetos podem ser discutidos e ampliados. As técnicas politômicas apresentam, no entanto, potencialmente melhores condições de aplicação que aquelas baseadas em lógica dicotômica para resolver as questões inerentes porque permitem representar objetos de natureza complexa, vaga, imprecisa, ambígua. Embora existam diversas implementações de sistemas difusos em computadores, excetuando-se as áreas relacionadas a controle de sistemas reais, ainda se faz pouco uso deles.

Referências Bibliográficas

Centeno, J. A.; Haertel, V. : Adaptive Low-Pass Fuzzy Filter for Noise Removal, Photogrammetric Engineering and Remote Sensing, Vol. 61, No. 10, October 1995, p 11267-1272.

Eastman, J.R. : Idrisi for Windows, Clark University, Worcester, 1997.

Ramsey III, E.W.; Jensen, J.R. : Remote Sensing of Mangrove Wetlands: Relating Canopy Spectra to Site-Specific Data, Photogrammetric Engineering and Remote Sensing, Vol. 62, No. 8, August 1996, p 939-948.

Wang, F. : Fuzzy Supervised Classification of Remote Sensing Images, IEEE Transactions on Geoscience and Remote Sensing, Vol. 28, No. 2, 1990, p 194-201.

Wang, F. : Improving Remote Sensing Image Analysis through Fuzzy Information Representation, Photogrammetric Engineering and Remote Sensing, Vol. 56, No. 8, August 1990, p 1163-1169.

Zhu, A. : Measuring Uncertainty in Class Assignment for Natural Resource Maps under Fuzzy Logic, Photogrammetric Engineering and Remote Sensing, Vol. 63, No. 10, October 1997, p 1195-1202.