



## SISTEMA DE CLASSIFICAÇÃO FUZZY NÃO SUPERVISIONADA APLICADA A DETECÇÃO DE BIOMAS EM IMAGENS DE SATÉLITES

Carlos Eduardo Guerra /UFPA/Matemática  
Nayara Safira da Silva Caldas UFPA/Matemática

Copyright 2008, SBGf - Sociedade Brasileira de Geofísica

*Este texto foi preparado para a apresentação no IV Simpósio Brasileiro de Geofísica, Brasília, 14 a 17 de novembro de 2010. Seu conteúdo foi revisado pelo Comitê Técnico do IV SimBGf, mas não necessariamente representa a opinião da SBGf ou de seus associados. É proibida a reprodução total ou parcial deste material para propósitos comerciais sem prévia autorização da SBGf.*

**ABSTRACT** - This paper aims to show a system for unsupervised classification, using as selection tool classes, a Fuzzy Clustering algorithm (FCM) combined with a fuzzy system for classification. Among the most common classification techniques like the binary, which classifies the pixel as the relevance or not in the range of gray levels, however, this technique has some limitations for the separation of two or more classes. According to Meirelles and House (2007), the separation between different biomes does not occur abruptly and there is a transition zone that can not be fully characterized as belonging to a particular class. The algorithm FCM (Fuzzy C-means clustering) groups the pixel's in clusters which can be obtained intervals classes that fed into the Fuzzy system. Finally, the fuzzy classifier is applied to LANDSAT TM, with the aim of discriminating features of type savanna. The system can be used to discriminate against other objects, such as lakes, forests and urban areas, may also be useful for studying changes in land cover due to human occupation, deforestation, among others.

**RESUMO** – Este trabalho tem por objetivo mostrar um sistema de classificação não supervisionada, utilizando, como ferramenta de seleção classes, um algoritmo Fuzzy Clustering(FCM) combinado com um sistema Fuzzy para classificação. Dentre as técnicas mais comuns classificação destaca-se a binária, que classifica o pixel quanto a pertinência ou não em um intervalo de níveis de cinza; no entanto, esta técnica apresenta restrições para a separação de duas ou mais classes. Segundo Meirelles e Câmara (2007), a separação entre biomas distintos não ocorre de maneira abrupta, existindo uma zona de transição que não pode ser caracterizada totalmente como pertencente a uma classe específica. O algoritmo FCM(Fuzzy C- means clustering) agrupa os pixel's em clusters os quais podem-se obter os intervalos de classes que alimentarão o sistema Fuzzy. Finalmente, o classificador Fuzzy será aplicado em imagens LANDSAT TM, com o objetivo de discriminar um tipo específico de bioma, as savanas. O sistema pode ser utilizado para discriminar outros objetos, tais como lagos, florestas e zonas urbanas, podendo também ser útil no estudo de

mudanças na cobertura vegetal devido à ocupação humana, desmatamentos, outros biomas, dentre outras.

### 1 INTRODUÇÃO

Os estudos de imagens de satélite são de grande importância para o monitoramento de regiões, caracterizando mudanças no solo e na cobertura vegetal. No entanto, para que essas imagens forneçam dados significativos muitas vezes faz-se necessário a utilização de ferramentas do processamento digital de imagens, dentre as quais podemos destacar os classificadores.

Os procedimentos de classificação são frequentemente utilizados nos estudos ambientais, com o objetivo de discriminar objetos de interesses (biomas, zonas urbanas, corpos d'água, etc.). Segundo Fonseca (2008), a classificação consiste em dividir o espaço de atributos da imagem em regiões correspondentes as classes temáticas. Dentre os tipos de classificações, temos a classificação supervisionada e a não supervisionada. Genericamente, na classificação não-supervisionada, cada pixel da imagem é associado a uma classe sem que o usuário tenha conhecimento prévio do número ou da identificação das diferentes classes presentes na área imageada. A classificação é feita através de agrupamentos baseados em similaridades (FONSECA, 2008). Na classificação supervisionada, é necessário um conhecimento prévio da área estudada para se definir as fronteiras de classe.

Nos processos de classificação de imagens, utiliza-se normalmente ferramentas de limiarização, que tem como objetivo, o agrupamento de pixel's em regiões com características específicas, pertencentes a um intervalo comum de níveis de cinza ou limiar (Caldas & Guerra, 2009).

Conforme o algoritmo, do classificador a ser empregado, os classificadores podem ser divididos em duas abordagens: **Classificadores "pixel a pixel"**, que utilizam apenas a informação espectral, isoladamente, de cada pixel para achar regiões homogêneas; e os

**classificadores por regiões**, que utilizam, além de informação espectral de cada "pixel", a informação espacial que envolve a relação entre os "pixels" e seus vizinhos (Florenzano, 2002). Os classificadores por região requerem a utilização de segmentadores para poder agrupar os pixels e separar as regiões. Neste trabalho, aplicaremos uma rotina de clusterização (FCM) para seleção automática (não supervisionada) de classes. As classes obtidas, em forma de intervalos, serão utilizadas em um sistema Fuzzy, que realizará uma classificação pixel a pixel. Finalmente, aplicaremos o sistema classificador completo em imagens TM, para a detecção de savanas na região de Santarém.

Savanas são ecossistemas com vegetação esclerofila mista; de espécies arbóreo-arbustivas esparsas, cujo estrato inferior é dominado por espécies herbáceas. Segundo Holanda et al. (2009), as savanas na região de Santarém estão inseridas numa região de tensão ecológica entre florestas e apresentam intensas mudanças na sua cobertura vegetal, associadas ao regime pluviométrico, dinâmica natural da sua vegetação e a ação antrópica (agricultura, ocupação, etc.).

## 2 OBJETIVOS

O objetivo deste trabalho é discriminar objetos, do tipo savanas (na Região de Santarém), valendo-se de uma metodologia de classificação não supervisionada, que utiliza o algoritmo Fuzzy C- means Clustering (FCM), para seleção automática de intervalos de classes, aliado a um sistema Fuzzy de classificação.

## 3 MATERIAIS E MÉTODOS

### 3.1 Fundamentos teóricos

Definimos uma imagem de satélite por um mapeamento  $\psi$  de  $\mathbf{A}$  em  $\mathbf{Z}$ , onde  $\mathbf{A}$  é uma matriz com elementos  $a_{ij}$ , que indicam a posição de cada pixel.

Os valores de  $\psi(a_{ij})$  assumidos pelo mapeamento são denominados por níveis de cinza.

$$\psi : a_{ij} \rightarrow \psi(a_{ij}), a_{ij} \in \mathfrak{R} \text{ e } \psi(a_{ij}) \in \mathbf{Z}. \quad (1)$$

Um sistema de classificação binária é uma função do tipo:

$$\tau : \psi_{ij} \rightarrow \{0,1\}, \quad (2)$$

que levam os pixel's da imagem em níveis cinza a valores 0 ou 1. Nesse caso, os pixel's são mapeados de acordo com limiares definidos segundo as características espectrais dos alvos que se deseja separar (Gonzales, 2003).

Um sistema de classificação binária tem como base a lógica clássica, i.é, a atribuição de valores fixos para relacionar elementos a um conjunto, baseia-se no princípio do terceiro excluído. Atribui-se valor 1 e 0 para a pertinência e não pertinência, respectivamente. O procedimento binário normalmente começa por uma análise prévia do histograma de uma das componentes espectrais da imagem; e define-se um limiar, ou seja,

um intervalo de níveis de cinza, que corresponda à feição que se deseja extrair. Aos pixel's pertencentes ao intervalo alvo atribui-se valor 1 e aos demais, 0.

Seja  $A$  um conjunto no universo  $X$ , a pertinência ou não de  $x \in X$  em  $A$  pode ser caracterizada, pela lógica clássica por:

$$\mu_A(x) = \begin{cases} 1 & \text{se } x \in A \\ 0 & \text{se } x \notin A \end{cases} \quad (3)$$

### 3.1.1 Lógica Fuzzy

Um conjunto de abordagens que permitem tratar incertezas, como um especialista da área de sensoriamento remoto o faria, são denominados por sistemas especialistas, dentre os quais destacam-se os algoritmos genéticos, computação evolutiva, as redes neurais e a lógica fuzzy (Barron, 1993; Aguiar, 1999 e Gomide et. al., 1994).

O conceito de conjunto Fuzzy foi introduzido por Lotfi A.Zadeh, dando importantes contribuições à lógica e a inteligência artificial. Na lógica Fuzzy, os elementos podem assumir graus de pertinência entre os valores 0 e 1. A função de pertinência é assim definida:

$$\mu_A(x) \rightarrow [0,1], \quad (4)$$

ou seja, um mapeamento dos elementos do universo de discurso  $X$  no intervalo  $[0, 1]$ . Determinamos o conjunto fuzzy  $F$  pela união dos pares ordenados  $(x, \mu(x))$ .

**Definição 3.1.1** Os operadores usados para elaboração das sentenças são:

- $\wedge$  operador de conjunção (and).
- $\vee$  operador de disjunção (or).
- $\Rightarrow$  operador de implicação (if-then).
- $\Leftrightarrow$  operador de dupla implicação (if and only if).

**Definição 3.1.2** Dados os conjuntos fuzzy  $W$  e  $Q$  e as funções  $W_\mu$  e  $Q_\mu$ , definimos o operador Norma-

Tmínimo por:

$$T\text{-min}(W,Q)=\min(W_\mu, Q_\mu)=W_\mu \wedge Q_\mu \quad (5)$$

**Definição 3.1.3** O operador Conorma-T máximo :

$$T\text{-max}(W,Q)=\max(W_\mu, Q_\mu)=W_\mu \vee Q_\mu \quad (6)$$

### 3.1.2 Fuzzy Clustering

A metodologia FCM foi desenvolvido por Dunn(1973) e melhorado por Bezdek em 1981. O "Fuzzy C-means Clustering" é um algoritmo que organiza um determinado conjunto de objetos em grupos, clusters. Um cluster é um conjunto de objetos agrupados segundo características em comum e que os diferem de outros clusters. Esse algoritmo baseia-se na minimização da função:

$$J_m = \sum_{i=1}^N \sum_{j=1}^C u_{ij}^m \|x_i - c_j\|^2, 1 \leq m < \infty, \quad (7)$$

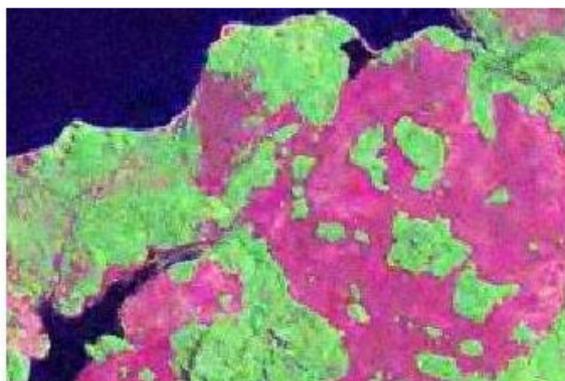
onde  $m$  é qualquer número maior que 1,  $u_{ij}$  é o grau de pertinência de  $x_i$  no cluster  $j$ ,  $x_i$  é o  $i$ -ésimo dado,  $c_j$  é o centro do cluster e  $\|\cdot\|$  é qualquer norma que expresse similaridades entre o centro e os dados processados.

### 3.2 Metodologia

A metodologia que implementa o classificador consiste das seguintes etapas:

#### 3.2.1 Seleção da área estudada e imagens

O alvo a ser estudado são as savanas da região de Santarém-PA. Primeiramente, escolhemos as imagens a serem processadas, no que se refere às bandas espectrais e sensores típicos para estudo de vegetação. Na **Figura 1** e **Figura 2** temos uma composição falsa cor e um recorte das bandas do vermelho, infravermelho próximo e infravermelho médio.



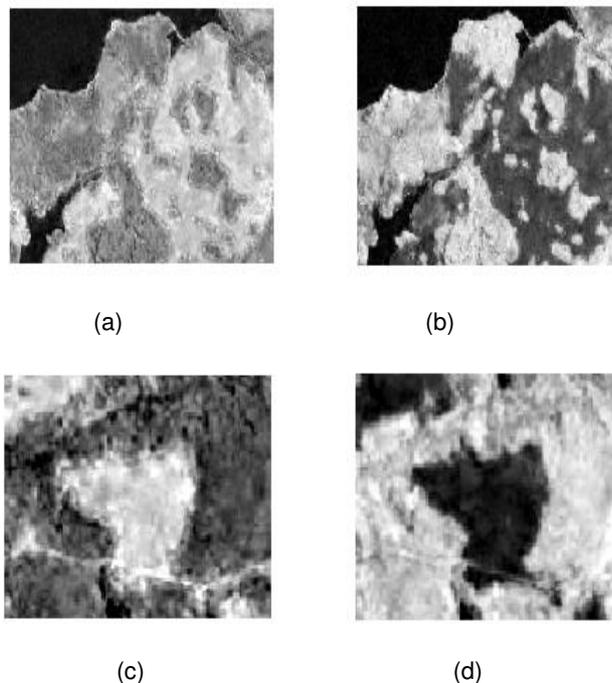
**Figura 1:** Composição colorida TM (banda 3,4 e 5) de uma região de Savanas (Região do Eixo-Forte em Santarém) com baixa cobertura Vegetal.



**Figura 2:** Recorte da região de Savanas do eixo forte em Santarém.

As bandas espectrais escolhidas como entrada para o classificador foram o vermelho e infravermelho

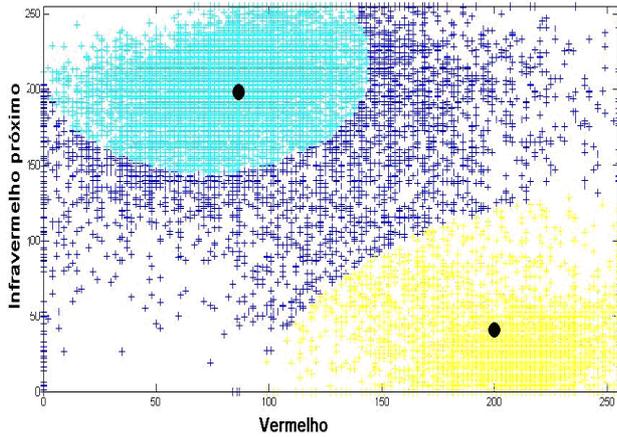
próximo, pois segundo Ponzoni & Shimabukuro (2007), estas banda caracterizam melhor as estruturas com baixa e alta cobertura vegetal. Na **Figura 2.(a)** e **2.(c)** as zonas mais claras correspondem as feições com baixa cobertura vegetal, i.é, savanas, estradas, desmatamentos, solos, dentre outras; enquanto que na **Figura 2.(b)** e **2.(d)** estas mesmas estruturas são representadas pelas as zonas mais escuras.



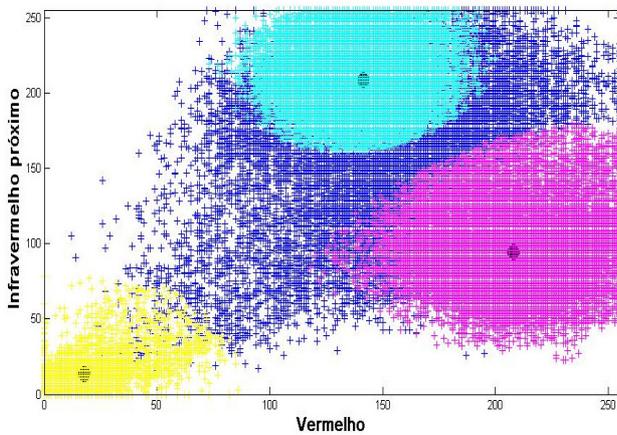
**Figura 3:** Na figura (a) e (c), Imagem TM na banda 3(vermelho) e banda 4 (b) e (d), infravermelho próximo, referentes as imagens das **Figuras (1) e (2)**.

#### 3.2.2 Aplicação do FCM e determinação de intervalos de classe.

Após a escolha das componentes a serem trabalhadas, as imagens são transformadas em vetores que armazenam os valores de níveis de cinza. Para a imagem da **Figura 1** o número de clusters a serem determinados pelo FCM é 3, pois é visível a presença de 3 estruturas bem definidas na imagem, que correspondem a savana (magenta), cobertura vegetal(verde) e a água(azul escuro). Na **Figura 2** o número de Clusters são apenas 2, visto que há predominância de savanas e cobertura vegetal. Os clusters podem ser observados na **Figura 4** e **Figura 5**. Os pontos pretos em destaques correspondem aos centros de cada cluster. O eixo horizontal e vertical são os valores dos pixel's, em nível de cinza, para as componentes vermelha e infravermelho próximo, respectivamente.



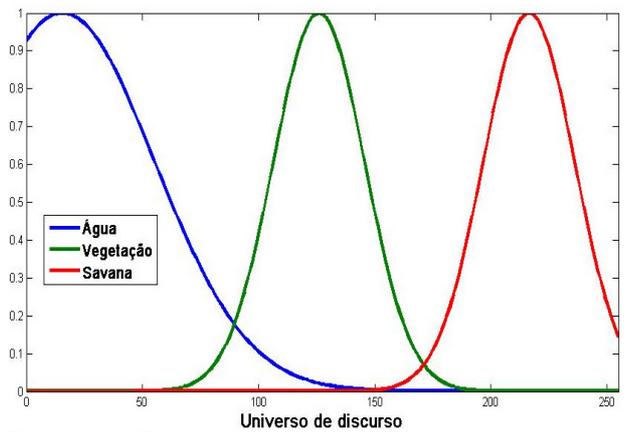
**Figura 4:** Agrupamento dos pixel's em clusters. O cluster azul claro corresponde às zonas com cobertura vegetal e o amarelo às zonas de savana.



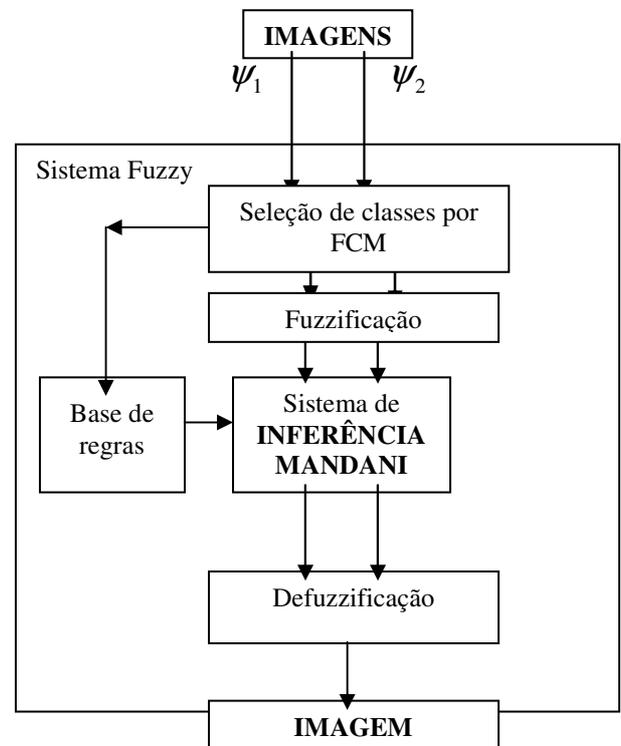
**Figura 5:** O cluster amarelo corresponde à água, o azul claro, as zonas com cobertura vegetal densa e o magenta, savanas.

**3.2.3 Implementação do Sistema Fuzzy a partir da extração de classes do FCM.**

Definidos os clusters, observamos que estes possuem formas geométricas similares as de elipses, cujas origens são os centros dos clusters. Como exemplo, a projeção do eixo maior sobre a horizontal nos fornece o suporte das funções de pertinência da primeira variável de entrada, i.e., a componente vermelha do sistema fuzzy de classificação. A projeção sobre a vertical corresponde ao limite das funções para a segunda variável de entrada, i.e., a componente infravermelho próximo. Na **Figura 6**, temos como exemplo as funções de pertinência resultante do algoritmo FCM para a componente do vermelho. O diagrama a seguir mostra como funciona o sistema classificador. As entradas  $\psi_1$  e  $\psi_2$  são a primeira e segunda componente das imagens escolhidas.  $\psi_r$  é a imagem resultante após a classificação. O sistema Fuzzy, mostrado no diagrama, possui duas variáveis de entrada, que correspondem as imagens a serem processadas (bandas 3 e 4 para o exemplo).



**Figura 6:** Funções de pertinência com suportes obtidos dos intervalos de classes calculados pelo algoritmo FCM.



O universo de discurso para cada uma das variáveis é o intervalo [0, 255] que corresponde aos valores de níveis de cinza. Para a imagem da **Figura 1**, o estudo das bandas 3 e 4 revelam a existência de três classes distintas: água, vegetação densa e as savanas. Após a aplicação do algoritmo FCM para três classes, obtém-se 3 intervalos que servirão de suporte para as funções de pertinência de cada variável de entrada. Na **Tabela 1**, estão os valores correspondentes aos intervalos de classe e o centro dos clusters. Para a segunda imagem, **Figura 2**, as componentes espectrais evidenciam apenas duas classes distintas: vegetação densa e savanas. O resultado obtido pelo FCM são dois intervalos suporte, que podem ser visualizados na **Tabela 2**.

**Tabela 1**-Intervalos e centro de classes para a banda 3.

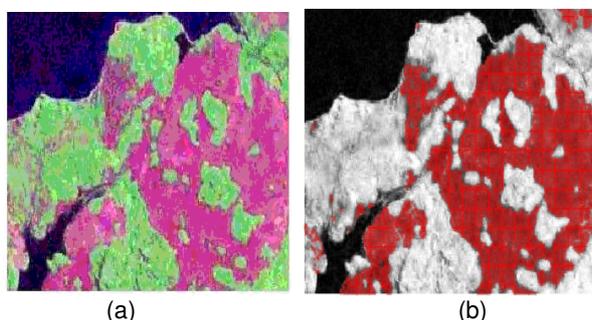
CLASSE	SUORTE	CENTRO (B3, B4)
ÁGUA	0-75	(18, 13)
VEGETAÇÃO	60-190	(141, 208)
SAVANA	125-255	(208, 94)

**Tabela 2**-Intervalos e centro de classes para a banda 3.

CLASSE	SUORTE	CENTRO (B3, B4)
VEGETAÇÃO	0-140	(86, 198)
SAVANA	95-255	(200, 41)

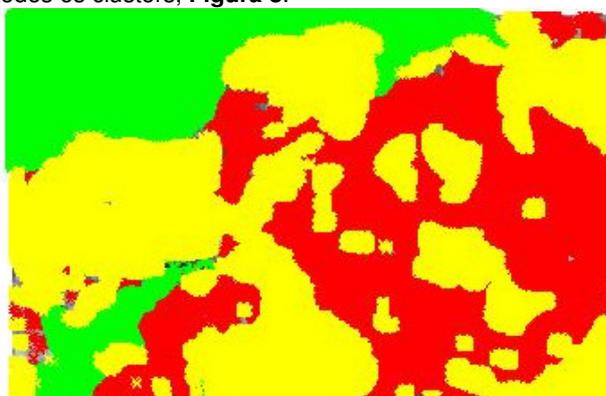
#### 4 RESULTADOS

Para melhor verificação do bom funcionamento do algoritmo FCM, sobrepomos os pontos de cada cluster em uma das bandas da imagem original (**Figura 1**), o que pode ser visto na **Figura 7**.



**Figura 7:** Na figura (a) a composição colorida TM. Em (b), os pontos do cluster, referente à zona de savana, sobreposto sobre a banda do infravermelho próximo.

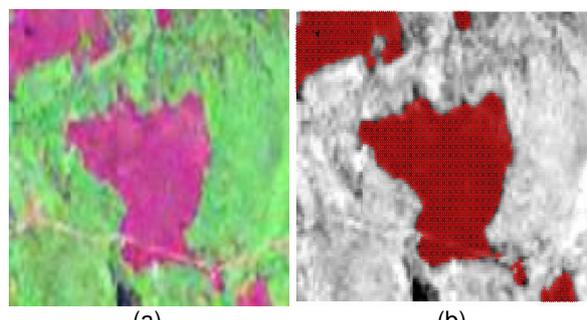
O algoritmo mostrou-se eficiente para separar as zonas de savanas. Nota-se também, a presença de alguns conjuntos de pontos não que não fazem parte do alvo savanas. Estas zonas apresentam comportamento espectral semelhantes às savanas para a banda do infravermelho próximo, i.é., zonas de baixa refletância. Uma análise completa do funcionamento do algoritmo dá-se pela visualização de todos os clusters, **Figura 8**.



**Figura 8:** Os 3 Clusters referente à imagem da **Figura 1**.

Na **Figura 9** temos a visualização do cluster "savana" referente a imagem da **Figura 2**. Observe que

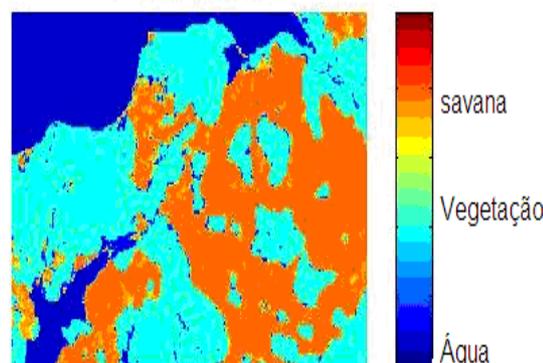
a sombra da nuvem na parte inferior da imagem e a estrada que corta a savana no centro (**Figura 9.a**) não fazem parte desta classe.



**Figura 9:** Vizualização do cluster "Savana" após aplicação do algoritmo FCM na imagem da **Figura 1**.

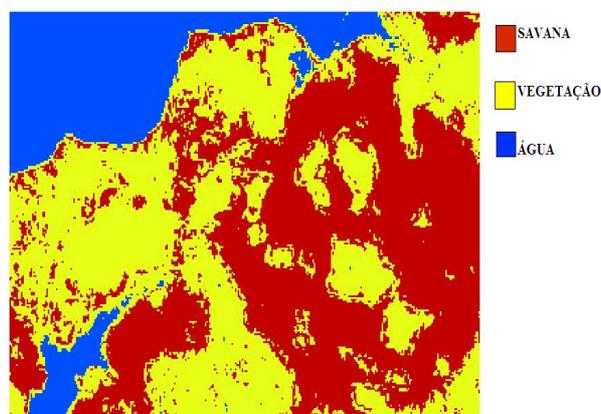
Como dito anteriormente, a utilização do FCM foi a base para montar um sistema Fuzzy de classificação, em que as funções de pertinências foram construídas através de intervalos extraídos dos clusters obtidos. Assim, podemos considerar o classificador como não supervisionado.

Classificação Fuzzy



**Figura 10:** Resultado obtido pela classificação Fuzzy, referente a imagem da **Figura 1**.

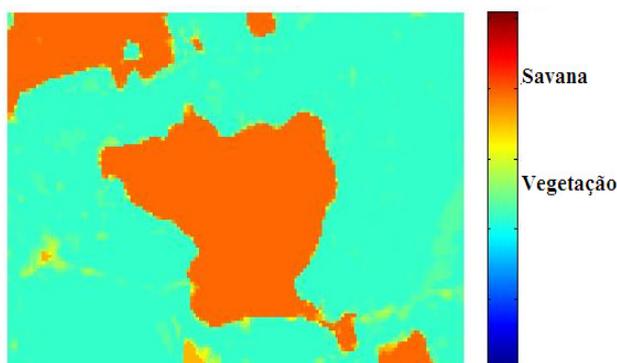
Na **Figura 10** temos o resultado de classificação Fuzzy para 3 classes.



**Figura 11:** Classificação Binária para três classes.

Comparando este resultado com a classificação binária (figura 11) podemos claramente observar a transição

dos biomas floresta e savanas, através das bordas amarelas que acompanham as feições em laranja. Na classificação binária, a inserção dos pixels em cada classe é rígida como pode ser observado na **Figura 11**. A **Figura 12** mostra um exemplo da aplicação do algoritmo sobre um recorte de figura (**Figura 2**), focalizando uma savana típica. Nota-se claramente a presença da borda de transição ao redor da savana. Por outro lado, uma sombra da nuvem aparece com uma coloração amarela. Embora o FCM não tenha incluído a nuvem na classe savana, a introdução dos intervalos de classes, como suporte das funções de pertinência, introduz uma margem de incerteza adicional que pode levar o classificador a considerar estas zonas como uma feição da classe.



**Figura 12:** Classificação Fuzzy para 2 intervalos de classes.

## 5 CONCLUSÃO

De uma forma geral, o sistema classificador funcionou satisfatoriamente. As classes majoritárias (óbvias) foram detectadas facilmente pelo algoritmo FCM. Quanto aos alvos savanas, o sistema obteve resultados satisfatórios, pois elas eram feições bem diferenciadas na imagem. O FCM foi capaz de extrair automaticamente os intervalos de classe, entretanto, o sistema Fuzzy de classificação apresenta uma dificuldade para tratar algumas feições que geram incertezas adicionais (sombas). Este problema poderá ser resolvido, em trabalhos futuros, onde passaremos a trabalhar com regiões segmentadas, e não apenas com pixel's individuais. Comparado ao classificador binário, o sistema fuzzy proposto é muito mais robusto e isto se deve à possibilidade de agregar incertezas. Como estamos trabalhando com dados reais, faz-se necessário um método mais flexível a variações, que leve em consideração mudanças não bruscas entre biomas.

## 7 AGRADECIMENTOS

Agradecemos a UFPA/LASERS-campus Santarém pelo apoio e suporte dado para a realização deste trabalho, e ao INPE pelo fornecimento das imagens.

## 8 REFERÊNCIAS

- MEIRELLES, M. S. P.; CÂMARA, G.; ALMEIDA, C. M, **Geomática - Modelos e Aplicações Ambientais**. Brasília, DF: EMBRAPA Informação tecnológica, 2007. 593 p.
- GOMIDE, F.A. CAMPOS; GUDWIN, RICARDO R., 1994, **Modelagem, Controle, Sistemas e Lógica Difusa**; SBA Controle & Automação/ vol.4.
- AGUIAR, H., OLIVEIRA, JR., 1999. **Lógica Difusa – Aspectos Práticos e Aplicações**; Editora Interciência, 1a Edição.
- BARRON, J. J., 1993, **A Lógica Fuzzy em Foco**, Byte No 06, pág. 94 – 100.
- FLORENZANO, TEREZA GALLOTI. **Imagens de Satélites para Estudos Ambientais**. Oficina de Textos, São Paulo, 2002.
- GONZALEZ, R. C. E WINTS, P. **Digital image processing**. 2ed., Massachusetts, Addison Wesley, 1987.
- FONSECA, E.L. **Métodos para Classificação de Imagens Digitais Multiespectrais: Classificadores supervisionados e não supervisionados**. 2008.
- HOLANDA, A.S;GUERRA, C.E & SUEMITSU,C.**Uso de imagens NDVI e Modelo de Mistura Espectral para o monitoramento de Savanas no município de Santarém-PA**.SBGF, 2009.
- CALDAS, N.S & GUERRA, C.E. **Limiarização Fuzzy aplicada a Imagens de Satélite**. Siges, 2009.
- PONZONI, F.J. & Y.E. SHIMABUKURO. **Sensoriamento Remoto no Estudo de Vegetação**. São Paulo. 2007.
- Florenzano, Tereza Galloti. **Imagens de Satélites para Estudos Ambientais**. Oficina de Textos, São Paulo, 2002.
- Gonzalez, Rafael C.; Woods, Richard E. e Eddins, Steven L. **Digital Image Processing Using MATLAB**. Prentice Hall, 2003.