

Comparação de métodos de classificação de imagens, visando o gerenciamento de áreas citrícolas

Ana Paula Barbosa¹
Alessandra Fagioli da Silva²
Célia Regina Lopes Zimback²
Sérgio Campos²

¹ Universidade Estadual Paulista - UNESP/FCA
Caixa Postal 237 - 18610-307 - Botucatu - SP, Brasil
anap_barbosa@yahoo.com.br

² Universidade Estadual Paulista - UNESP/FCA
Caixa Postal 237 - 18610-307 - Botucatu - SP, Brasil
alefagioli@hotmail.com; czimback@gmail.com; seca@fca.unesp.br

Abstract. This study aimed to compare the efficiency of methods of classification of orbital images in areas cultivated with citrus, using GIS techniques to the planning and management located in the areas of citrus production. The study area is located in Bariri City, São Paulo State. It noted the evaluation results of the classification accuracy were satisfactory, which the CLUSTER classification had excellent quality (0.9276) and very good (0.6485), the MAXVER algorithm had excellent ratings, with kappa 0.8338 and 0.8818, and the Fuzzy classification submitted a rating between very good (0.7260) and good (0.5235) for 2003 and 2007, respectively. The methods used in the discrimination of areas cultivated with citrus showed different efficiencies in the images classification. In general, the classifications of 2003 images showed the best performance than 2007 images scores, as they were less influenced by the reflectance of other classes.

Palavras-chave: *Fuzzy*, LANDSAT satellite, Images classifiers, *Fuzzy*, satélite LANDSAT, classificadores de imagens

1. Introdução

A citricultura no Estado de São Paulo é de grande importância econômica e social, pois apresenta participação significativa na geração de empregos, na ocupação de terra, formação de renda, ativação dos elos da cadeia produtiva, captação de divisas e desenvolvimento regional de outros setores da economia (prestação de serviços, transportes, comércio, etc.).

Apesar da significativa produção, verifica-se no Brasil a carência de ferramentas para um monitoramento constante das áreas citrícolas, permitindo a tomada de decisões rápidas e abrangentes, de forma a favorecer a manutenção desta exploração em níveis internacionalmente competitivos (SANCHES et. al, 2005; ASN, 2004).

Assim, compreende-se que o estudo de pomares de citros a partir de sensores de resolução espacial equivalente ao TM/LANDSAT exige o entendimento de diversos fatores. Isto porque ocorrem variações na resposta espectral destes alvos em função de alterações no substrato, altura e cobertura do terreno pelas árvores e características geométrica da disposição das plantas e das condições de observação do sensor (COVRE, 1989).

Os métodos de classificação digital podem ser agrupados em função da presença ou não de uma fase de treinamento onde o analista interage com o computador, isto é, podem ser divididos em categorias, de acordo com o tipo de treinamento: supervisionados, não-supervisionados e híbridos, que podem ser rígidos ou relativos. (DAINESE, 2001; RIZZI, 2004; EASTMAN, 2006).

O uso de classificadores não supervisionados é indicado quando, *a priori*, o analista não possui conhecimento da área de estudo, buscando eliminar a subjetividade do processo de obtenção das amostras de áreas. o classificador CLUSTER, que efetua a classificação de imagens compostas que combinam as bandas de informação mais úteis (EASTMAN, 2006).

O classificador de máxima verossimilhança (MAXVER) é executado assumindo igual probabilidade de ocorrência e custo de erro de classificação para todas as classes. Porém, se houver um conhecimento *a priori* de que as probabilidades não são iguais para todas as classes, é possível especificar fatores de peso para classes particulares. Esta variante da regra de decisão é conhecida como regra de decisão Bayesiana. Se não houver conhecimento *a priori* das probabilidades, recomenda-se não especificá-las, utilizando os pesos predefinidos da equação (NOGUCHI, 2004).

A lógica de classificação baseada nos conjuntos *fuzzy* leva em consideração a natureza imprecisa e heterogênea do mundo real. Pode ser usada em combinação com algoritmos de classificação supervisionada e não supervisionada (CAMPOS, 2005).

Ao invés de ser designado a uma única classe, entre todas as classes possíveis, cada *pixel* na classificação *fuzzy* tem *m* valores de graus de participação, cada um associado a sua probabilidade (ou correlação) com as classes de interesse (FOODY, 1992).

São várias as possibilidades para a estimativa da exatidão de um mapa temático, dentre elas a estimativa do parâmetro Kappa (*k*), fundamentado em critérios da análise multivariada discreta, que é utilizado como medida de concordância entre o mapa e a referência que se tem adotado para a estimativa da exatidão.

Este trabalho teve por objetivo a comparação da eficiência de métodos de classificação de imagens orbitais em áreas cultivadas com citros, Município de Bariri/SP, utilizando técnicas de geoprocessamento visando o planejamento e o gerenciamento localizado das áreas de produção de citros.

2. Metodologia de Trabalho

A área de estudo utilizada nesta pesquisa corresponde a uma fazenda produtora de citros, localizada no município de Bariri/SP, situada entre as latitudes 22° 01'S a 22° 59'S e longitudes 48° 47' WGr a 48° 45' WGr.

Na propriedade são cultivadas as variedades de citros: Hamlin, Natal, Pera, e Valência, totalizando 28 talhões, dos quais, no ano de 2006, cinco talhões da variedade Pera foram erradicados por razões fitossanitárias.

Foram utilizadas imagens no formato digital do satélite LANDSAT-5 TM, órbita/ponto 221/75, em composição falsa cor RGB 543, referente às datas de passagem 16/06/2003 e relacionadas na 26/05/2007.

As imagens foram preparadas com realce, retificação geométrica e redução da dimensionalidade para a realização da classificação.

Cinco classes de uso do solo foram localizadas na imagem: citros, solo exposto, cana-de-açúcar, cultura anual e pastagem. Em seguida, foram digitalizados polígonos em torno dessas áreas, atribuindo um identificador para cada tipo de cobertura.

Depois de identificadas as classes espectrais referentes à classe de uso citros, elas foram agrupadas, desconsiderando-se as demais classes, tendo como resultado uma imagem binária da área de citros para cada classificação, onde as classes espectrais correspondentes à citros receberam o valor um (1) e as demais o valor zero (0).

No processo de classificação CLUSTER, as bandas para classificação foram selecionadas e foi indicado o nível de generalização para obtenção de uma imagem com maior número de agrupamentos. Foram eliminados os agrupamentos não significativos, obtendo-se o agrupamento de cinco clusters representativos para avaliação.

Na classificação MAXVER, as classes foram definidas com número mínimo de 30 *pixels* para cada classe de treinamento, cujos *pixels* nelas contidos foram analisados e assinaturas espectrais foram criadas para cada objeto. Finalmente, executou-se o classificador, onde todos os *pixels* foram classificados com igual probabilidade para cada classe de uso.

Neste método foram desenvolvidas assinaturas *fuzzy*, que consideram que cada *pixel* tem potencialmente algum grau de associação a mais de um conjunto (classe), de acordo com o

grau com que misturas destas classes existem dentro do *pixel*. Dessa maneira, não houve necessidade de definir áreas de treinamento homogêneas.

O passo seguinte foi a criação de uma matriz de partição *fuzzy*, para indicação do grau de associação de cada área de treinamento a cada classe, onde foram adotados os seguintes pesos discriminados na Tabela 5.

Tabela 1. Matriz de partição *fuzzy*, com valores de graus de associação das classes

Classes	Citros	Solo exposto	Cana-de-açúcar	Cultura anual	Pastagem
Citros	0,7	0	0	0	0
Solo exposto	0,1	0,9	0,1	0	0,1
Cana-de-açúcar	0	0,1	0,7	0,1	0
Cultura anual	0	0	0,1	0,8	0,2
Pastagem	0,2	0	0,1	0,1	0,7

Os pesos foram determinados levando-se em consideração o aspecto da cultura na imagem (reflectância), pois o citros pode apresentar certa confusão com pastagem e solo exposto quando em estádios mais novos; cana-de-açúcar com solo exposto e cultura anual; e pastagem, por apresentar diversos graus de degradação, pode ser confundida com citros, cana-de-açúcar e cultura anual. Novas imagens foram associadas a conjuntos *fuzzy* e foram processadas com probabilidades a priori iguais, com z-score de 1.

Para cada data foi gerado um mapa de referência, com os quais foram feitas as tabulações cruzadas entre as imagens classificadas e a referência, obtendo-se as matrizes de erros dos métodos de classificação e o índice Kappa.

3. Resultados e Discussão

A avaliação do desempenho das classificações de imagens requer a elaboração de um mapa de referência, que foi obtido a partir da classificação em tela da área de estudo, para os anos 2003 e 2007 (Figura 1). Foram determinadas seis (6) classes de uso: citros, solo exposto, cana-de-açúcar, cultura anual, pastagem e vegetação nativa. As demais classes de uso não foram contempladas devido às pequenas áreas presentes nas imagens analisadas.

Os métodos de classificação foram analisados com base nos resultados obtidos através da interpretação visual, representando a verdade terrestre.

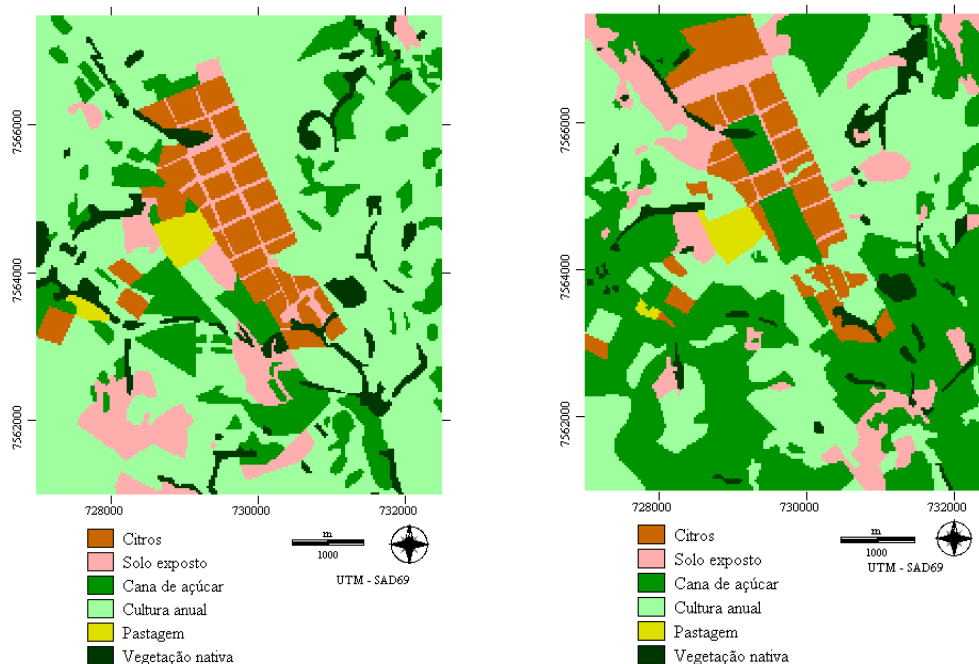


Figura 1. Mapa resultante da interpretação visual para os anos 2003 e 2007.

O método não supervisionado CLUSTER tem como produto a identificação de classes espectrais distintas na imagem, as quais podem ser identificadas como tipos de cobertura por meio da associação com dados de referência terrestre (Figura 2).

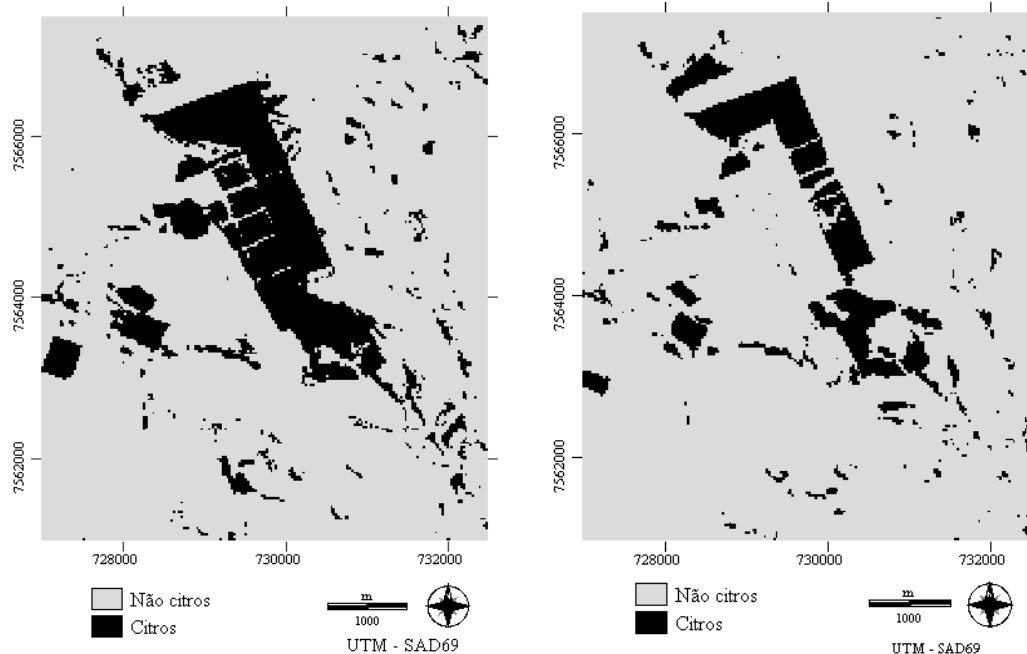


Figura 2. Mapas resultantes da classificação CLUSTER para os anos 2003 e 2007.

Na classificação não supervisionada CLUSTER da imagem de 2003, a classe citros foi agrupada satisfatoriamente, notando-se que houve baixa influência da refletância de outras classes, visto que foram poucas as áreas identificadas como citros que não corresponderam a esta classe.

Observando o mapa obtido por este classificador referente à imagem de 2007, quando comparado ao mapa de referência (Figura 1), foi possível detectar nitidamente que os pomares citrícolas sofreram influência de outras classes na sua resposta espectral, pois parte do pomar não foi classificado como tal, embora as áreas agrupadas como citros tenham sido detectadas com definição, discriminando-as das demais classes.

Alguns erros de designação do classificador CLUSTER foram observados nas áreas de transição entre culturas anuais, solo exposto e o citros. Embora o citros tenha sido bem classificado pelo classificador, os talhões erradicados sofreram forte influência da refletância das classes solo exposto.

O classificador MAXVER está implementado nos principais programas de processamento digital de imagens digitais, por ser considerado um dos mais eficientes classificadores e é, conseqüentemente, um dos mais utilizados.

Pelo fato deste classificador utilizar áreas de treinamento definidas pelo analista para agrupar diferentes classes, pode-se verificar na Figura 3, que a classe citros foi apropriadamente agrupada tanto em 2003 quanto em 2007, quando comparados com as classificações visuais.

Esse classificador apresentou desenvolvimento satisfatório, inserindo nas classes determinadas os *pixels* a elas correspondentes. No entanto, observando os mapas de referência, verifica-se que houve confusão do classificador ao agrupar diferentes áreas que apresentavam comportamento espectral semelhante à desta classe.

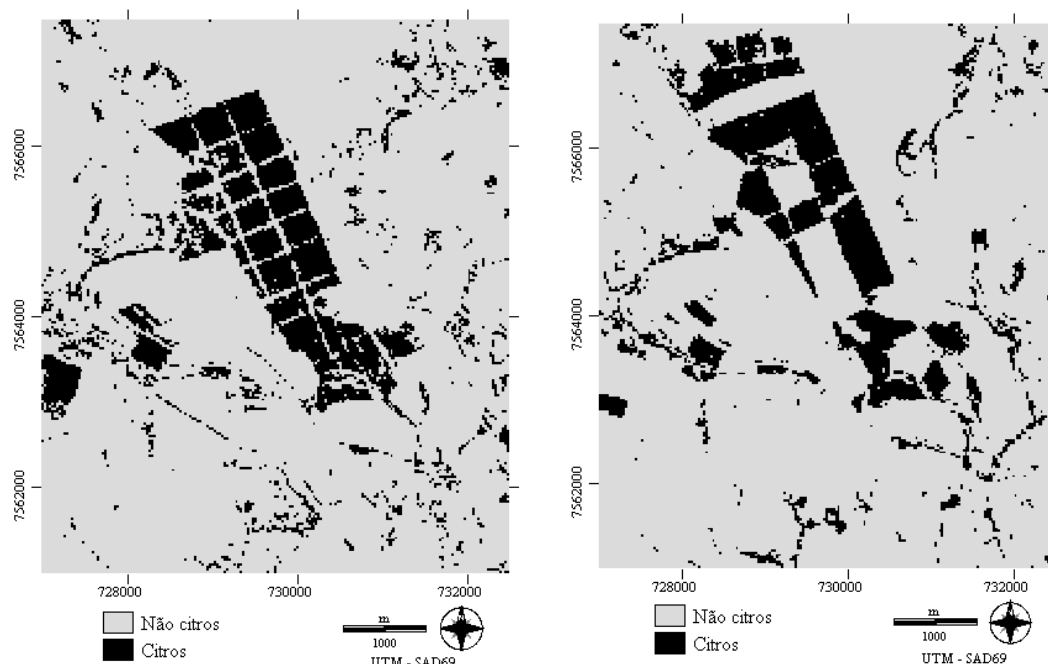


Figura 3. Mapas resultantes da classificação MAXVER para os anos 2003 e 2007.

Por fim, em ambas as datas, algumas classes encontradas nas imagens foram suprimidas, pois não alcançaram os 30 *pixels* necessários para serem classificadas.

A classificação relativa *Fuzzy* foi executada para discriminar áreas cultivadas com citros de outras culturas, como mostrado na Figura 4, a fim de se obter uma classificação mais adequada e que considere a complexidade da natureza em sua heterogeneidade.

Foi possível verificar na imagem classificada de 2003 que a matriz de partição representou de maneira satisfatória as classes de citros, uma vez que as áreas detectadas erroneamente, referentes a outras culturas, foram mínimas.

Na imagem de 2007, houve uma redução na qualidade do classificador, dado pelo fato que algumas classes de citros não foram detectadas, por apresentar características semelhantes às de outras classes de uso, como a cana-de-açúcar, por exemplo.

Isso ocorreu pelo fato de se utilizar a mesma matriz de partição para a determinação das classes de uso, mostrando que o comportamento das culturas, no momento da atribuição dos pesos, e a amostragem das classes influenciam os valores de graus de participação, bem como a associação das classes de interesse com sua probabilidade.

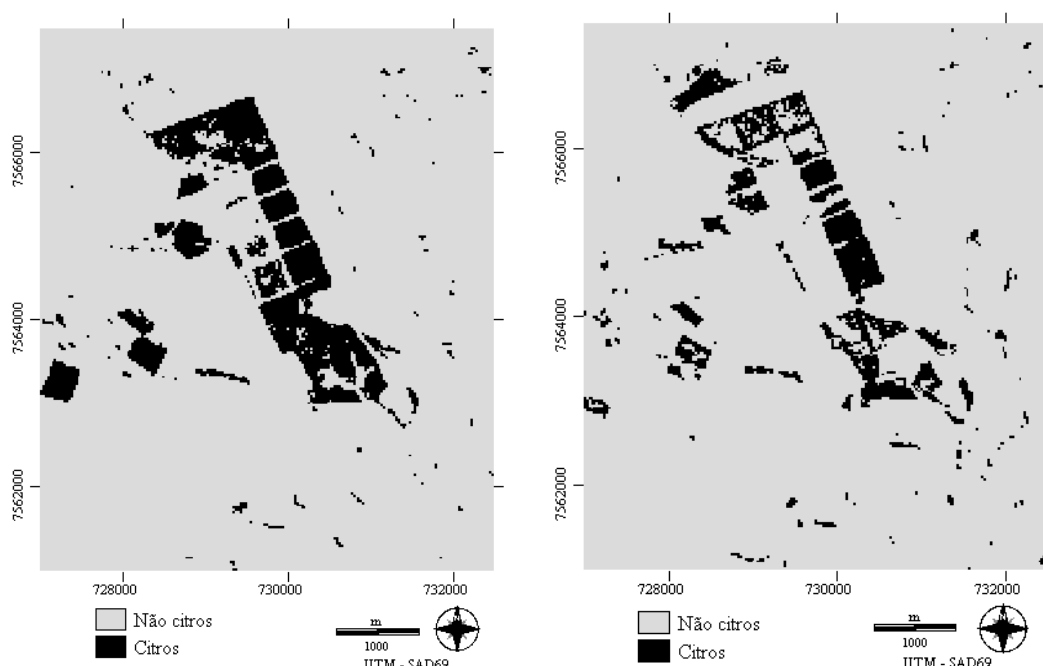


Figura 4. Mapas resultantes da classificação *Fuzzy* para os anos 2003 e 2007.

Na Tabela 1 constata-se que os resultados na avaliação de acurácia da classificação foram satisfatórios, nos quais as classificações do algoritmo CLUSTER tiveram qualidade excelente (0,9276) e muito boa (0,6485) para as imagens de 2003 e 2007, respectivamente. É interessante notar que a classificação pode apresentar distintas acurácias em paisagens e épocas diferentes, em razão do desenvolvimento das plantas e comportamento espectral de alvos da superfície terrestre.

Tabela 1. Comparação das classificações CLUSTER, MAXVER e FUZZY pelo Índice Kappa

Método	Ano	Kappa	Qualidade de classificação
Cluster	2003	0.9276	Excelente
	2007	0.6485	Muito boa
Maxver	2003	0.8338	Excelente
	2007	0.8818	Excelente
<i>Fuzzy</i>	2003	0.7260	Muito boa
	2007	0.5235	Boa

O algoritmo MAXVER apresentou classificações excelentes, com Kappa de 0,8338 para a imagem de 2003 e 0,8818 para a imagem de 2007. De fato, a classificação MAXVER apresentou nível excelente nas duas épocas analisadas, validando o exposto por Shimabukuro (1996).

Já o índice gerado pelo método de classificação relativa *Fuzzy* teve o menor valor do índice para a classificação de 2007, comparado aos demais métodos, o qual apresentou qualidade de classificação entre muito boa (0,7260) e boa (0,5235), para os anos de 2003 e 2007, respectivamente.

Foi notado, durante o período de escolha e geração dos mapas de classificação *Fuzzy*, que os mapas finais foram bastante condicionados pelos graus de pertinência entre as classes de

culturas. É muito provável que, com novas simulações, ocorram mudanças e alguma melhoria na classificação.

4. Conclusões

O método de classificação não supervisionado CLUSTER mostrou vantagem sobre os demais métodos, pois classificou as áreas de citros com qualidade e rapidez, não necessitando de intervenção do analista.

O método de classificação supervisionada MAXVER apresentou o melhor desempenho, entretanto, exigiu um elevado tempo na determinação das amostras de treinamento.

Ao comparar todos os métodos verifica-se que, em termos de tempo, o classificador CLUSTER foi o mais rápido, seguido dos classificadores MAXVER e *Fuzzy*. Com relação à qualidade da classificação, o MAXVER mostrou-se superior aos demais em ambas as datas, seguido pelos classificadores CLUSTER e *Fuzzy*.

Agradecimentos

À Empresa Citrovita®, pelo interesse no trabalho e por disponibilizar a área do estudo realizado; ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pelo apoio financeiro, durante a dissertação de mestrado e ao Grupo de Estudos e Pesquisas Agrárias Georreferenciadas – GEPAG pelo apoio e infraestrutura.

Referências Bibliográficas

- AGÊNCIA SEBRAE de notícias (ASN). Produção integrada chega à citricultura. 2004. Capturado em 18 out. 2004. [online]. Disponível em: <<http://asn.interjornal.com.br/site/noticia.kmf?noticia=23004&canal=199&total=57&indice=10>>. Acesso em: 18 dez. 2008.
- CAMPOS, M. A. A. Padrão e dinâmica de floresta tropical, através de classificação orientada a objeto e da análise da paisagem com imagens LANDSAT. Tese (Doutorado em Engenharia Florestal). Universidade Federal do Paraná. Curitiba. 2005, 122 f.
- COVRE, M. Influencia de parâmetros culturais de citros sobre os dados TM/Landsat. 1989. (INPE-4856-TDL/367). Dissertação (Mestrado em Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São Jose dos Campos. 1989. 241 f.
- DAINESE, R. C. Sensoriamento remoto e geoprocessamento aplicado ao estudo temporal do uso da terra e na comparação entre classificação não supervisionada e análise visual. 2001. Dissertação (Mestrado em Agronomia) – Faculdade de Ciências Agrônômicas, Universidade Estadual Paulista, Botucatu, 2001. 186 f.
- EASTMAN, J.R. **Idrisi 15: The Andes Edition**. Worcester, MA: Clark University, 2006.
- FOODY, G. M. A Fuzzy sets approach to the representation of vegetation continua from remotely sensed data: an example from lowland heath. **Photogrammetric Engineering and Remote Sensing**, Bethesda, v. 58, n. 2, p. 221-225, 1992.
- NOGUCHI, M. T. G. Um sistema baseado em regras *fuzzy* para classificação supervisionada de imagens multiespectrais de alta resolução. Dissertação (Mestrado em Informática), Universidade Federal do Paraná, Curitiba, 2004. 150 f.
- RIZZI, R. Geotecnologias em um sistema de estimativa da produção de soja: estudo de caso no Rio Grande do Sul. Tese (Doutorado em Sensoriamento Remoto), Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2004. 212 f.
- SANCHES, I. DEL´ARCO., GÜRTLER, S., FORMAGGIO, A. R. Discriminação de variedades de citros em imagens CCD CBERS-2. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 12, 2005, Goiânia. **Anais...** São José dos Campos : INPE, 2005. p. 277-284.
- SHIMABUKURO, M. T. Avaliação de classificadores espectrais de imagens Landsat-TM em áreas rurais densamente ocupadas. 1996. Dissertação (Mestrado em Biologia Vegetal) – Universidade Estadual de Campinas, Campinas, 1996. 193 f.