

Mapeamento da cobertura da terra por meio de análise orientada a objetos: estudo de caso para uma área costeira do município de Raposa/MA

Taíssa Caroline Silva Rodrigues¹
Thiago Batista dos Santos¹
Yosio Edemir Shimabukuro¹

¹Instituto Nacional de Pesquisas Espaciais - INPE
Caixa Postal 515 - 12227-010 – São José dos Campos - SP, Brasil.
(taissageo, batista, yosio)@dsr.inpe.br

Abstract. Remote sensing technology has been increasingly used in decision making and in studies involving land cover. With the availability of new sensors presenting better spatial, spectral and radiometric resolutions, there is a large increase in the wealth of detail which leads to the application of new techniques for classification of images. The coastal environments exhibit intense and dynamic variation over short distances, in mangrove ecosystems and sandbanks, which face environment favorable to their development. With greater wealth of information and detail in the images, there is a complexity in automatic mapping, requiring the use of new techniques that can help and optimize the choice of the best attributes to be used in the final classification. To tackle this problem, we used in the present work the technique of data mining that assists in selecting the best attributes to differentiate the classes of interest. This work generated automatically a map of land cover of the region in the municipality of Raposa / MA. The final product presented a Kappa coefficient of 0.85 and 88.40% overall agreement.

Palavras-chave: WorldView II, data mining, coastal environments, eCognition, WorldView II, mineração de dados, região costeira.

1 Introdução

O sensoriamento remoto, devido à crescente disponibilidade de sensores de alta resolução espacial e ao surgimento de novas técnicas para a extração de informação destas imagens, tem sido uma ferramenta cada vez mais utilizada na tomada de decisão e nos estudos envolvendo cobertura do terra (BLASCHKE e KUX, 2007).

Com uma maior riqueza de detalhes há uma complexidade no mapeamento automático dessas imagens. Dessa maneira existe a necessidade de aplicações de novas técnicas que possam auxiliar no processo de classificação. Neste sentido, tem-se a inserção das técnicas de mineração de dados, que pode tornar a seleção de atributos e limiarização (processos que demandam tempo) muito mais rápida e igualmente adaptável para qualquer área em que se queira mapear a cobertura da terra de forma automática.

Com a utilização de técnicas de mineração de dados no processo de classificação orientada a objetos, cabe ao intérprete apenas analisar e elaborar a rede hierárquica com base na seleção de atributos resultantes dessas técnicas. Este procedimento é mais pertinente nos casos em que se têm inúmeros atributos, tanto espectral, quanto de textura e forma. Exemplo dessa condição é o sistema eCognition (Definiens, 2008).

O emprego dessas técnicas podem agilizar e potencializar a exatidão do mapeamento de ambientes estuarinos e de manguezais. Essas áreas apresentam intensa dinâmica ambiental e em curtas distâncias, são encontrados ecossistemas de mangues e restingas, que deparam com ambiente propício ao seu desenvolvimento. A distinção de alvos, como mangues, restingas, dunas e banco de areias em imagens de satélites é complexa devido à resposta espectral similar desses alvos.

Diante este cenário e da disponibilidade de imagens WorldView II da região costeira do município de Raposa no estado do Maranhão, este trabalho tem como

objetivo gerar automaticamente um mapa de cobertura da terra, por meio de árvores de decisão elaboradas com base em seleção de atributos.

2 Área de Estudo

A área de estudo está localizada no município de Raposa, no extremo nordeste do Maranhão, limitada pelas coordenadas 02° 21' a 02° 32' de latitude sul e 44° 00' a 44° 12' de longitude oeste (Figura 01).

O município caracteriza-se por um intenso processo de alteração da paisagem natural. A fisiografia costeira da região compõe-se de extensas baixadas litorâneas com formações de praias arenosas, dunas móveis, paleodunas, manguezais, restingas, marismas e depósitos de vasas modelados por uma extensa rede de canais, normalmente preenchidos pela preamar (Feitosa, 1998).

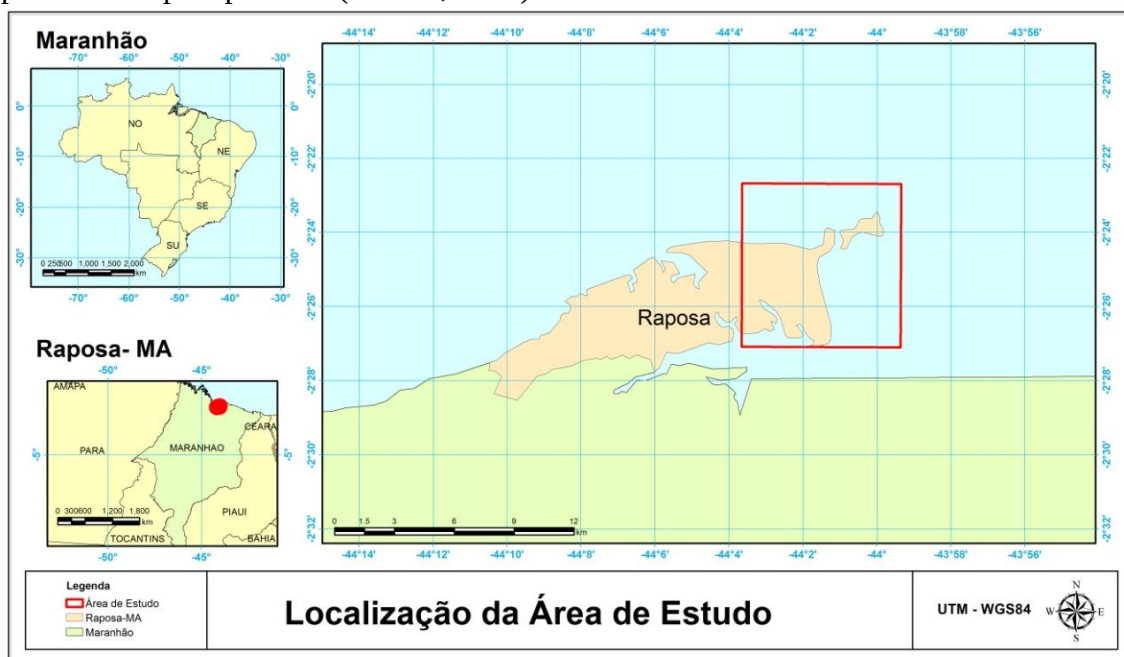


Figura 1: Localização da área de estudo no município de Raposa, Maranhão.

O ecossistema de mangues e restingas compõe-se de grandes áreas com diferentes níveis de maturação e de espécies vegetais. Esses dois tipos de vegetação encontraram na área, ambiente propício, pois as restingas são tipicamente de ambientes arenosos com pouca disponibilidade de água no solo durante a estação seca, o que favorece a rarefação e o caráter efêmero da cobertura vegetal. Enquanto que os mangues localizam-se em ambientes de encontro de águas doce e salgada, exposto diretamente à linha da costa e submetido diariamente à ação das marés de água salgada ou salobra (FEITOSA, 1998). Esses são sistemas funcionalmente complexos, altamente resistentes e, portanto, estáveis, caracterizados por vegetação e sedimentos típicos e apresentando grande importância socioeconômica e ambiental (MOCHEL et al., 2006).

3 Metodologia

3.1 Materiais

Para a realização deste trabalho, foram utilizados os seguintes dados:

1) Uma imagem do sensor WorldView II: multiespectral, com resolução espacial de 2 metros e oito bandas espectrais (*coastal, blue, green, yellow, red, re-edge, near infra-red 1 e near infra-red 2*). As imagens foram obtidas em 10 de junho de 2010, com uma resolução radiométrica de 11 bits.

2) Foram utilizados os seguintes softwares:

SPRING 5.2, utilizado para a realização dos índices;

eCognition 8.0, utilizado no processo de aquisição do conjunto de treinamento e na implementação da árvore de decisão;
ENVI 4.8; utilizado para aplicar o Modelo Linear de Mistura Espectral;
WEKA 3.6.8, utilizado na Mineração dos dados e criação da árvore de decisão;
Terra View 4.2, utilizado para gerar as amostras para a validação.

3.2 Métodos

3.2.1 Modelo Linear de Mistura Espectral (MLME)

A composição espectral do elemento de resolução de imagens é a combinação das respostas dos alvos que estão dentro do IFOV (*instantaneous field of view*). Dessa forma, cada pixel da imagem, que pode assumir qualquer valor dentro da escala de cinza, contém informações da proporção e da resposta espectral de cada componente presente na unidade de resolução do terreno (SHIMABUKURO e SMITH, 1991).

A radiação eletromagnética, captada pelo sistema sensor, pode ter sido espalhada somente pela vegetação e retornado ao sensor diretamente, ter sofrido espalhamento múltiplo sem ter interagido com o solo, ter interagido somente com o solo ou com o solo e a vegetação, além de não captar nenhuma energia devido ao sombreamento presentes na vegetação (PONZONI e SHIMABUKURO, 2009).

Dependendo das características intrínsecas de cada alvo, a mistura espectral pode ser linear ou não linear, porém, a modelagem linear é amplamente utilizada pela comunidade científica e tem apresentado resultados consistentes (PONZONI e SHIMABUKURO, 2009).

O papel do modelo linear de mistura espectral (MLME) é estimar a proporção de cada alvo no elemento de resolução da imagem. Para tanto, é necessário escolher amostras puras, pois servirão de referência para estimar a contribuição de cada alvo na resposta espectral captada pelo sistema sensor.

Neste trabalho o modelo linear foi aplicado nas imagens WorldView II com intuito de realçar a vegetação de restinga e de mangue. Na área de estudo ambos ecossistemas são limítrofes e cobrem quase toda a região.

3.2.2 Seleção de Atributos

3.2.2.1 Amostragem

Essa etapa do trabalho foi realizada no programa eCognition 8.0. Primeiramente, foi feita a segmentação das imagens. Para tanto, utilizou-se o algoritmo “*multiresolution segmentation*” adotando-se 70 como fator de escala e peso 1 para todas as bandas, inclusive para a fração vegetação do Modelo Linear de Mistura Espectral (MLME). Após a segmentação efetuou-se a coleta de amostras para a construção do conjunto de treinamento. Posteriormente, tais amostras foram transformadas para formato nativo do WEKA, software responsável pela mineração dos dados.

Foram definidas cinco classes (água, banco de areia, dunas, mangue e restinga). Para cada uma delas foram coletadas 30 amostras bem distribuídas na cena com intuito de representar as diferentes variações internas de cada classe (cor, textura, forma, tonalidade e brilho) de forma a evitar possíveis incertezas na classificação.

3.2.2.2 Mineração de Dados

O sistema eCognition possui inúmeros atributos, além da possibilidade, da criação de diversos outros pelo usuário. Para a realização deste trabalho foram escolhidos atributos como: média e o desvio padrão das oito bandas e da fração vegetação derivada do MLME, área, comprimento e atributos de textura (homogeneidade e contraste).

Geralmente, a mineração de dados (também chamada de descoberta de conhecimento), é o processo de análise de dados a partir de diferentes perspectivas e resumindo-a em informações úteis, ou seja, é o processo de encontrar correlação ou padrões entre dezenas de campos relacionais em um grande banco de dados (YANG et. al., 2008).

A partir da coleta das amostras e da escolha dos atributos a serem exportados foi realizado a mineração de dados no software WEKA. Foi escolhido o emprego dessas técnicas de mineração, pois elas podem indicar, entre os atributos disponíveis, aqueles mais pertinentes para a descrição das classes de cobertura da terra de forma automática e sem nenhuma subjetividade por parte do usuário na sua escolha e na definição dos limiares (NOVACK, 2009).

A partir disso foi fornecida uma árvore de decisão, em que o programa indicou, segundo o conjunto amostral coletado nas imagens, quais foram os melhores atributos que melhor distinguiram as classes propostas. Conforme Han e Kamber (2001), a árvore de decisão constitui-se em um fluxograma estruturado como uma árvore, em que cada nó interno denota um teste em um atributo, cada ramo representa um resultado do teste, e cada nó-folha (ou nó terminal) possui um rótulo da classe.

Pode-se perceber que a árvore de decisão utilizou como principais atributos: o desvio padrão, as médias das imagens fração do modelo linear de mistura espectral e das bandas 7 e 8 (*near infra-red* 1 e 2) assim como a média da banda 3 (*green*) e média da banda 4 (*yellow*). O que foi confirmado por estudos realizados por Souza et al. (2011), que afirmaram que as novas bandas do sensor WorldView II, principalmente a banda “*yellow*”, conseguiu separar significativamente vegetações diferentes em áreas costeiras.

3.2.3 Classificação

Após o processo de mineração de dados e da criação da árvore de decisão, foram implementados no software eCognition 8.0 as regras e atributos utilizados para a construção da árvore.

Para isso, foi criado um novo processo no programa, com o nome de Classificação, em que foram inseridos todos os processos filhos, propostos pela árvore de decisão. Para classificação utilizou-se o algoritmo “*classification*” em que todas as regras são inseridas em todas as cinco classes definidas (água, banco de areia, dunas, mangue e restinga). Com isso, todos os processos foram executados, finalizando assim a classificação da cena.

Após a etapa de classificação, foi inserido um novo processo com o algoritmo “*merge region*”, em que os objetos de cada classe são agrupados em objetos maiores. Finalmente, os dados foram exportados em formato shapefile para ser feito o processo de validação no SPRING 5.2.

3.2.4 Validação do Mapa de Cobertura da Terra

A validação do mapa de cobertura da terra foi feito a partir de pontos amostrais extraídos das imagens WorldView II originais. No total, mil pontos amostrais foram coletados. A coleta foi feita com o uso do plugin gerador de pontos amostrais do Terra View 4.2. Esse plugin garante o princípio da aleatoriedade e a boa distribuição espacial dos pontos.

Com os pontos de referência foi possível gerar uma matriz de confusão, a partir da qual foram calculadas as medidas de exatidão global, exatidão do produtor, exatidão do usuário e índice Kappa. Esse índice demonstra o quão coerente é o mapa temático em relação aos pontos de referência (CONGALTON e GREEN, 2009).

4 Resultados

4.2 Árvore de Decisão

Após a realização da mineração dos dados com o algoritmo J48, foi criada a árvore de decisão, que apresentou os atributos que melhor auxiliariam na classificação da imagem.

Para cada classe, a árvore de decisão inseriu um grupo de atributos. As classes, água, banco de areia e dunas, foram as que menos precisaram de atributos para sua distinção.

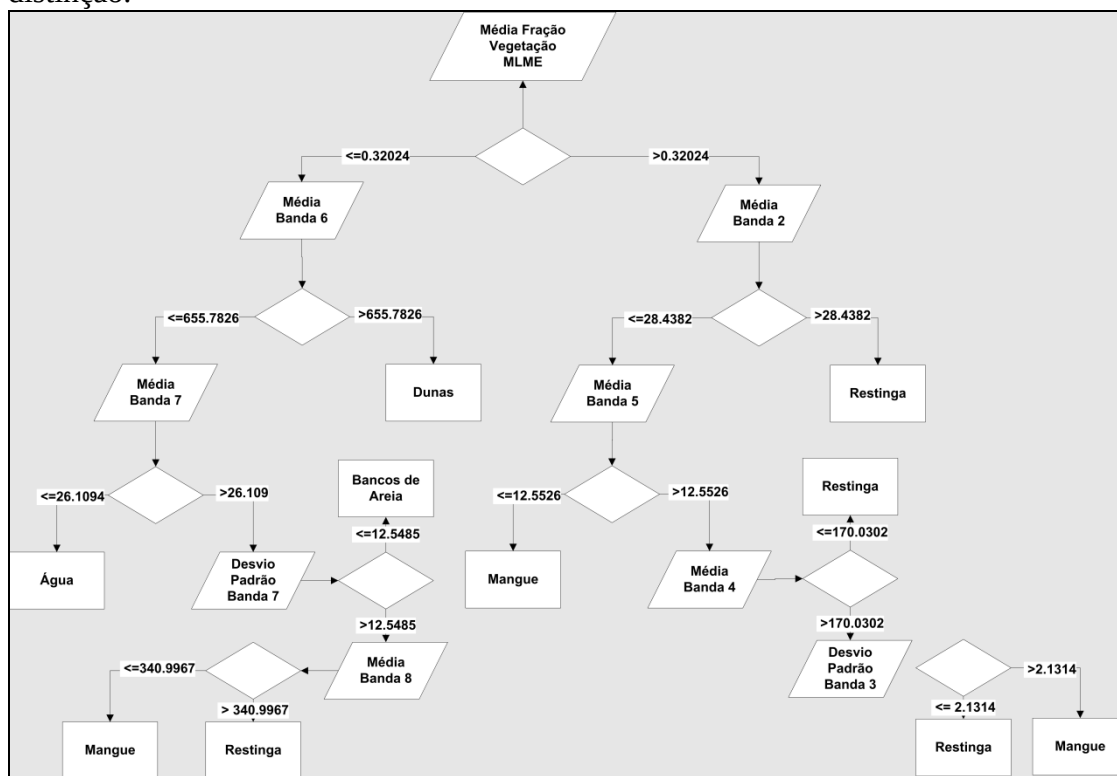


Figura 2: Árvore de decisão construída com base na seleção de atributos.

Para a classe água, além do MLME, utilizou-se apenas as médias das bandas 6 (*red-edge*) e 7 (*near infra-red 1*). A nova banda do sensor, banda 6, separou melhor água com maior quantidade de sedimentos enquanto que a banda 7 distinguiu água mais puras que apresentam baixa reflectância nessa banda. As áreas de banco de areia foram diferenciados da classe água somente pelo atributo desvio padrão da banda 7. A resposta espectral desses alvos é semelhante, pois os bancos de areia estão submersos em águas marinhas (poucos centímetros de profundidade), ou permanecem úmidos na maré baixa.

Para a classificação da classe de dunas, foi utilizado além da média do MLME a média da banda 6, que diferentemente da água, as dunas apresentam nessa faixa do espectro, maiores valores de reflectância, o que facilita sua separação dos outros alvos.

Já para as classes mangue e restinga foram necessários um maior número de nós e atributos para separação, devido à aproximação de suas características. Para isso foram utilizadas as médias do MLME das bandas 2 (*blue*), 4 (*yellow*), 5 (*red*), 7 e 8 (*near infra-red 1 e 2*) e o desvio padrão da banda 3 (*green*).

O uso da nova banda do sensor WorldView II, assim como o MLME, auxiliaram na distinção dos dois tipos de vegetação encontrados na área de estudo. Isso confirmou os estudos realizados anteriormente por Souza et al. (2011), em que afirmaram que as novas bandas do sensor WorldView II, principalmente a banda “*yellow*”, apresentam importantes informações que ajudam na separação significativa dos tipos de vegetação encontrados em áreas costeiras.

4.2 Validação do Mapa de Cobertura da Terra

Na figura 3 são apresentados o mapa de cobertura da terra e as amostras utilizadas no processo de análise da exatidão temática.

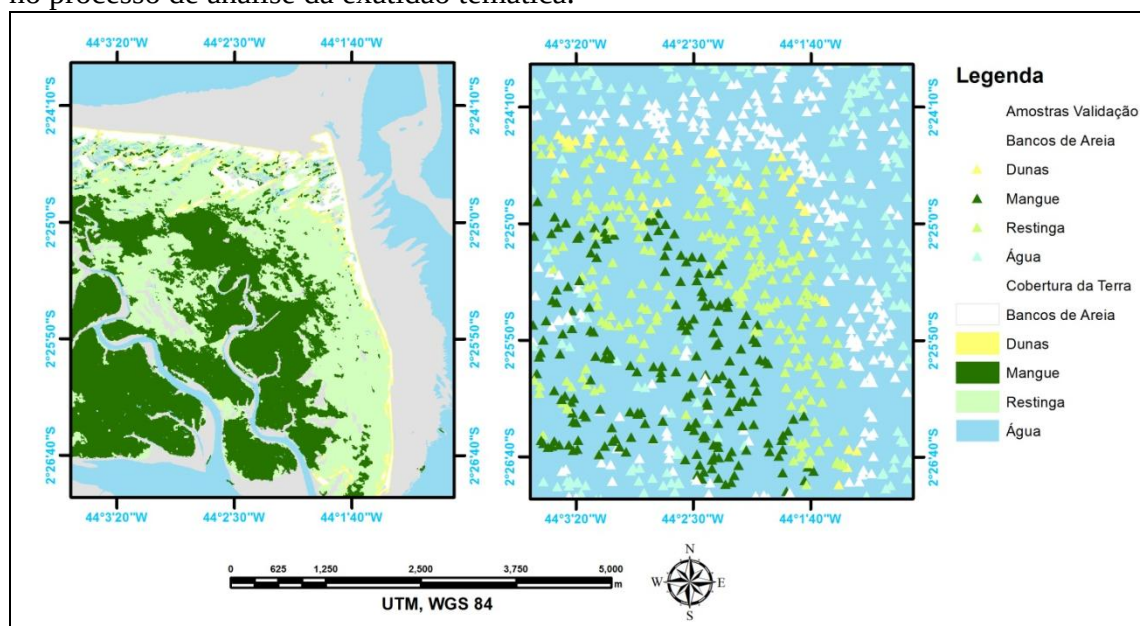


Figura 3: À esquerda, o mapa de cobertura da terra gerado e à direita a distribuição dos pontos de validação.

Na tabela 1 são apresentados o índice Kappa e as exatidões global, do produtor e do usuário.

Tabela 1: Análise de exatidão dos tipos de cobertura da terra identificados

Classes	Área (ha)	E* Produtor	E* Usuário	E* Global	I* Kappa
Banco de Areia	891,70	93,53	93,53	88,40	0,85
Mangue	736,20	87,08	78,79		
Restinga	653,71	77,82	86,92		
Dunas	138,61	85,71	95,45		
Água	3.841,07	95,11	91,85		
Total	6.261,29	-----	-----		

O mapa gerado obteve o coeficiente de concordância Kappa de 0,85 e concordância global de 88,40%. No escopo geral foi um resultado satisfatório, porém algumas deficiências foram encontradas ao analisar a exatidão individual das classes.

As classes mangue e restinga apresentaram menor exatidão quando comparadas as demais. As confusões no mapeamento ocorreram somente entre essas classes. Tal problema manifestou-se em áreas de transição entre esses ecossistemas (mangue e a restinga). Regiões mais próximas dos canais de maré possuem vegetação mais exuberante, com uma vegetação mais densa. Porém, quanto mais distante dos canais, a vegetação de mangue tende a ficar mais esparsa e com menor porte arbóreo. Devido a essas características, áreas de manguezais, foram classificadas como restinga e vice-versa.

Também foi detectada a ocorrência de incertezas no mapeamento de bancos de areia. Lagoas secas inseridas entre as dunas foram erroneamente mapeados como banco de areias, isso provavelmente ocorreu devido à umidade existente nessas áreas. Outra fonte de inexatidões foi o mapeamento errôneo de dunas como banco de areias. Essa

situação ocorreu próxima à linha de costa, em áreas limítrofes entre as dunas e os bancos de areia no mar.

5 Considerações Finais

A geração de mapas de cobertura da terra para regiões costeiras, por meio de classificação orientada a objetos e utilizando imagens originais e as imagens fração do MLME, foi eficaz. A seleção de atributos contribuiu para a elaboração da árvore de decisão implementada, gerando economia de tempo sem comprometer a exatidão temática do mapa final. As poucas incertezas embutidas no mapeamento estão relacionadas a semelhanças nos atributos dos alvos presentes na cena estudada. Uma possível solução para melhorar ainda mais a exatidão temática seria o uso de segmentações e classificação hierárquica. Nesse processo os alvos podem ser rotulados em classes genéricas e serem refinados por meio de uso de filtros de classe e com uso de regras de vizinhança.

Agradecimentos

Agradecemos a empresa Digital Globe pelo fornecimento das imagens do sensor WorldView II utilizadas neste trabalho.

6 Referências Bibliográficas

- BLASCHKE, T.; KUX, H. (editores) **Sensoriamento remoto e SIG avançados: novos sistemas e sensores inovadores**, 2ª. Edição, São Paulo: Oficina de Textos, 2007, 304 p., 2007.
- CONGALTON, R. G.; GREEN, K. **Assessing the accuracy of remotely sensed data: principles and practices** (J. G. Lyon, Ed.) **The Photogrammetric Record**. v. 25, p.204–205. CRC. doi: 10.1111/j.1477-9730.2010.00574_2.x, 2009.
- FEITOSA, A. C. Morfodinâmica da área costeira a nordeste da Ilha do Maranhão. In: Simpósio Nacional de Geomorfologia, 2.. **Geosul**, Florianópolis, Nov. 1998, v. 14, n.27, p. 252-255.
- HAN, J.; KAMBER, M. **Data mining: concepts and techniques**. San Francisco: Morgan Kaufmann Publishers, 2001.
- MOCHEL, F. R.; COSTA, M. R. P.; ALCÂNTARA, E. H.; AMORIM, A. J. E. Avaliação das potencialidades e fragilidades das áreas de manguezal para a implementação do ecoturismo usando ferramentas de sensoriamento remoto em Curupu – MA. **Revista Caminhos de Geografia**. São Luís, v. 22, n.17, p. 237 – 243, 2006.
- NOVACK, T. **Classificação da cobertura da terra e do uso do solo urbano utilizando o sistema InterIMAGE e imagens do sensor Quickbird**. 2009. 214p. INPE-16594 TDI/1580. Dissertação (Mestrado em Sensoriamento Remoto). INPE – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2009. Disponível em: <http://mtc.m18.sid.inpe.br/col/sid.inpe.br/mtcm18%4080/2009/08.31.21.23/doc/publicacao.pdf> Acesso em: 01 nov 2012.
- PRADO, F. A. **Sistema hierárquico de classificação para mapeamento da cobertura da terra nas escalas regional e urbana**. 2009. 167 p. Dissertação (Mestrado em Ciências Cartográficas) – Universidade Estadual Paulista (UNESP), Presidente Prudente, 2009.
- PONZONI, F. J.; SHIMABUKURO, Y. E. **Sensoriamento remoto no estudo da vegetação**. p.144. São José dos Campos - SP: Editora Parêntese, 2009.
- PINHO, C. M. D.; KUX, H.; ALMEIDA, C. M. **Elaboração de rede semântica para classificação de cobertura do solo de ambientes intra-urbanos: o Caso de São José dos Campos -SP, Brasil**. In: XIII Simpósio Brasileiro de Sensoriamento Remoto, 21-26 abr 2007, Florianópolis, Brasil. **Anais...** São José dos Campos: INPE, 2007. p. 637-644. Disponível em: <http://marte.dpi.inpe.br/col/dpi.inpe.br/sbsr@80/2006/11.15.21.54.40/doc/637-644.pdf> . Acesso em: 03 de nov de 2012.
- SHIMABUKURO, Y. E.; SMITH, J. A. The least-squares mixing models to generate fraction images derived from remote sensing multispectral data. *Transaction on Geoscience and Remote Sensing*, v. 29, n. 1, p. 16–20. doi: 10.1109/36.103288, 1991.
- YANG, T. L.; BAI, P.; GONG, Y. S. **Spatial Data Mining Features between general Data Mining**. International Workshop on Education Technology and Training and International Workshop on Geoscience and Remote Sensing, 2008.