

## **Análise da expansão urbana na Microrregião dos Lagos (RJ) através da integração entre GEOBIA e mineração de dados**

Marcus Vinícius Alves de Carvalho\*  
Victor Buznello de Vasconcellos Maluf\*  
Fernando Souza Damasco\*  
Isabel Santos de Lima Gomes\*  
Liliane Ronquette da Silveira Santos\*  
Luisa Schneider Moreira Dias\*  
Raúl Sánchez Vicens\*

\* Universidade Federal Fluminense - UFF  
Laboratório de Geografia Física - LAGEF / Instituto de Geociências - IGEO  
Av. General Milton Tavares de Souza, s/nº - Campus da Praia Vermelha  
Gragoatá - Niterói, RJ - Brasil  
CEP: 24.210-346  
marcus\_carvalho@id.uff.br  
vic.bvm@gmail.com  
fernandodamasco@id.uff.br  
isabeluerj@gmail.com  
lilianeronquette@id.uff.br  
luisa.smd@hotmail.com  
rsvicens@gmail.com

**Abstract.** The objective of this paper is to evaluate the integration between GEOBIA and Data Mining applied to the mapping of urban sprawl from TM/Landsat-5 images in three periods: 1991, 2000 and 2010 (periods between census). This study achieved a good Kappa Index: 0.66.

**Palavras-chave:** change detection, urban sprawl, data mining, GEOBIA, detecção de mudanças, expansão urbana, mineração de dados, GEOBIA.

### **1. Introdução**

O avanço da urbanização em nosso país tem provocado sérios impactos no modo de vida da população. O contínuo aumento da impermeabilização do solo, da densidade demográfica e de construções em locais impróprios, dentre outros, são sempre acompanhados de graves problemas sociais e econômicos, além de impactos negativos na infraestrutura urbana e no meio ambiente (ARAÚJO, 2006).

Assim, cada vez mais é necessária a constante aquisição de informações atualizadas sobre os ambientes urbanos para a eficaz tomada de decisão por parte dos gestores e planejadores. Tradicionalmente, os métodos de mapeamento de áreas urbanas, fazendo uso de sensoriamento remoto, contam com dados provenientes de levantamentos aerofotográficos. Porém, no caso de monitoramento do crescimento e expansão urbana, são necessários conjuntos de dados multitemporais, não sendo as fotografias aéreas produtos facilmente disponíveis nestas condições. Assim, devido ao seu caráter sinótico, multitemporal, multiespectral e de menor custo (em comparação aos levantamentos aerofotogramétricos), os produtos de sensoriamento remoto orbital têm se mostrado de relevante utilidade e eficiência neste tipo de análise, alcançando resultados bastante satisfatórios (VIEIRA, 1993).

Por meio da análise de dados de Sensoriamento Remoto de épocas distintas é possível realizar a detecção de mudanças na cobertura e uso do solo. Esta análise permite inferir sobre a dinâmica das transformações em vários outros aspectos, como intensidade, tipo de mudança (substituição ou conversão) e as taxas de mudança detectadas em uma região durante um

determinado período. Assim, o presente trabalho tem por objetivo analisar a expansão urbana na Microrregião dos Lagos no nos anos de 1991, 2000 e 2010 (período entre Censos). A análise se dará de forma de dois em dois anos, ou seja, 1991-2000 e 2000-2010.

## **2. Fundamentação teórica**

### **2.1. Detecção de mudanças em dados de sensores remotos**

Detecção de mudanças, do inglês *change detection*, é uma abordagem em Sensoriamento Remoto utilizada para identificar diferenças no estado de um objeto ou fenômeno através de sua observação em épocas diferentes (VERONA, 2002). Sensores remotos são muito utilizados no monitoramento de mudanças na superfície terrestre, fornecendo informações consistentes e medições repetidas em escala apropriada para muitos processos que provocam mudanças que incluem agentes naturais, antrópicos e mudanças climáticas (KENNEDY et al., 2007).

Segundo JIANYA et al. (2008), o processamento de detecção de mudança configura-se como uma área de pesquisa realmente promissora em Sensoriamento Remoto. Nos últimos anos, um grande progresso tem sido observado com o intuito de superar obstáculos tecnológicos, mediante o desenvolvimento de novas plataformas e sensores. A disponibilidade de grandes acervos de imagens históricas configura-se como importante fator, uma vez que, torna a detecção de alterações e modelagem possível. Tal desenvolvimento estimula investigação mais aprofundada no desenvolvimento de métodos de processamento de imagem e novas abordagens na manipulação de dados de imagem na dimensão temporal.

Dados de sensoriamento remoto de melhor resolução em diferente intervalo de tempo auxiliam na análise da taxa de mudanças, bem como os fatores causais ou condutores de mudanças (DAI & KHORRAM, 1999). Por isso, tem um papel importante no planejamento em diferentes escalas espaciais e temporais. Para Macleod e Congalton (1998) alguns aspectos devem ser considerados para o emprego de tal técnica: detectar as mudanças ocorridas, identificar a natureza da mudança, medir a extensão da mudança, definir o padrão espacial e a intensidade da mudança.

### **2.2. GEOBIA - Análise de Imagem Baseada em Objeto Geográfico**

Os classificadores que fazem uso da abordagem conhecida como Análise de Imagem Baseada em Objeto Geográfico buscam aprimorar o desempenho dos processos automatizados de classificação de imagens de Sensores Remotos. Eles possuem recursos que permitem sistematizar e reproduzir o conhecimento do intérprete humano. BLASCHKE et al., (2000) destacam que é possível que a semântica seja desenvolvida apoiada em parâmetros físicos e de conhecimento sobre relacionamentos espaciais dos objetos.

Segundo Hay e Castilla (2008), GEOBIA é uma subdisciplina da Ciência da Informação Geográfica (*GIScience*) dedicada ao desenvolvimento de métodos automatizados para a significativa divisão de imagens de sensoriamento remoto em objetos-imagem, e para avaliar suas características por meio de escalas espaciais, espectrais e temporais, com o intuito de gerar novas informações geográficas em SIG. DEFINIENS (2007) destaca que a utilização do conceito de objeto é peça-chave neste tipo de análise de imagens, pois se parte do princípio de que a informação semântica necessária para a interpretação de uma imagem não está presente apenas no pixel, e sim em objetos da imagem e nas relações existentes entre eles.

### **2.3. Mineração de dados**

Mineração de dados é o nome dado ao conjunto de técnicas que permite a extração de conhecimentos a partir de grandes volumes de dados. O processo de transformação de dados em informações e conhecimento é conhecido como descoberta de conhecimento em base de

dados (*KDD - Knowledge Discovery in Databases*). A mineração de dados trata da conversão de dados ou informações (que podem ter sido pré-processadas) em padrões (SANTOS, 2012).

Segundo Goldschmidt e Passos (2005), existem diferentes métodos de Mineração de Dados (*Data Mining*), entre eles destacam-se: métodos baseados em redes neurais, métodos baseados em algoritmos genéticos, métodos baseados em instâncias, métodos estatísticos, métodos específicos, métodos baseados em indução de árvores de decisão e métodos baseados em lógica nebulosa.

A mineração de dados geográficos envolve a aplicação de ferramentas computacionais para revelar padrões interessantes em objetos e eventos distribuídos no espaço geográfico e ao longo do tempo. Estes padrões podem envolver as propriedades espaciais de objetos individuais e eventos (como, por exemplo, a forma, extensão) e as relações espaço-temporais entre os objetos e eventos, além dos atributos não-espaciais de interesse na mineração de dados tradicional (MILLER; HAN, 2009).

### 3. Área de estudo

A área de estudo é conhecida como Microrregião dos Lagos está inserida na Região das Baixadas Litorâneas no Estado do Rio de Janeiro e é constituída por sete municípios, a saber: Araruama, Armação dos Búzios, Arraial do Cabo, Cabo Frio, Iguaba Grande, São Pedro da Aldeia e Squarema (Figura 1).

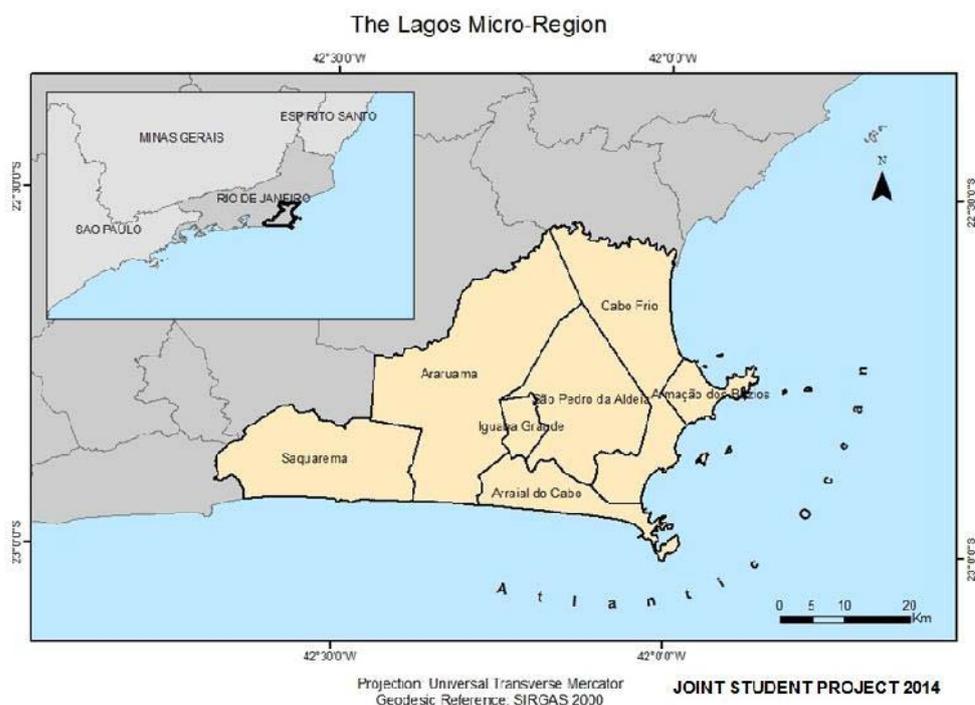


Figura 1 – Municípios que compõem a área de estudo.

### 4. Metodologia

As imagens TM/Landsat 5 foram adquiridas gratuitamente no Centro de Dados de Sensoriamento Remoto (CDSR) da Divisão de Geração de Imagens (DGI) do Instituto Nacional de Pesquisas Espaciais (INPE). As imagens são datadas de 05/12/1991, 01/04/2000 e 08/02/2010.

A primeira etapa consistiu na normalização radiométrica das imagens. O processo de normalização radiométrica é uma técnica de calibração relativa que consiste na regressão linear entre imagens multiespectrais em uma série temporal em relação a uma imagem de

referência (PONZONI, 2009). Esta técnica diminui diferenças radiométricas entre imagens, causadas por inconsistências de condições de aquisição, ao invés de mudanças reais de cobertura da terra (YUAN e ELVIDGE, 1996; YANG e LO, 2000). Para tal foram utilizados os softwares ENVI (correção pela subtração do pixel escuro) e ArcGIS (geração das imagens normalizadas através da ferramenta *Raster Calculator*).

A Etapa seguinte consistiu na correção geométrica das imagens. Para tal fez-se uso do módulo *Registration* (opção *Image to Image*) da plataforma ENVI. Assim, tomou-se como base uma imagem OLI/Landsat 8 no site do Serviço Geológico dos Estados Unidos (USGS) para corrigir espacialmente a imagem TM/Landsat-5. Vale destacar que as imagens OLI e TIRS já têm sido disponibilizadas ortorretificadas pelo USGS.

Em seguida, gerou-se o Modelo Linear de Mistura Espectral (MLME). Segundo INPE (2014), o MLME baseia-se na premissa que quando um sensor remoto observa a cena, a radiância detectada é a integração (mistura) de todos os objetos, denominados componentes da mistura, contidos no elemento de cena. Deste modo, por meio da extração das respostas individuais de cada alvo a partir de amostras de pixels puros desses alvos, também chamados de *endmembers*, pode-se modelar o peso com que cada alvo está contribuindo para o sinal de cada pixel e, assim realizar uma composição deste sinal em imagens-fração (ROBERTS et al., 1998). O software utilizado nesta etapa foi o SPRING (CÂMARA et al., 1996).

Para a detecção de mudança optou-se pela geração de “imagens-amplitude”. As imagens amplitudes consistem na soma dos valores máximos de imagens de diferentes datas subtraídas pela soma dos valores mínimos de imagens de diferentes datas. Exemplo: (Maximum values band 3 (red) 2010 + Maximum values band 3 (red) 2000) - (Minimum values band 3 (red) 2010 + Minimum values band 3 (red) 2000). Para isso fez-se uso da ferramenta *Mosaic to New Raster* do ArcGIS. Realizou-se isso para cada banda do sensor TM/Landsat-5 (exceto para a banda do Termal).

Em seguida as imagens-amplitude mais as imagens-fração (solo, sombra e vegetação) foram divididas de dois em dois anos: 1991-2000 e 2000-2010. Assim, foram gerados dois projetos no eCognition Developer. Todas as imagens foram segmentadas de uma única vez com o intuito de segmentar os objetos que compõem a cena mais as mudanças. O algoritmo utilizado foi o *Multiresolution Segmentation* e os parâmetros foram: fator de escala: 10 / forma: 0,4 / compacidade: 0,5. Chegou-se a estes valores de forma empírica. O resultado da segmentação foi apreciado através de avaliação visual.

O 1º nível de segmentação gerado foi copiado duas vezes em níveis abaixo. Assim o primeiro nível correspondeu às mudanças captadas através da imagem amplitude. O segundo nível correspondeu ao que seria urbano em 2000 e o terceiro nível, o que seria urbano em 1991 (isto para o caso de 1991-2000).

A etapa que segue consistiu na seleção de amostras para treinar o algoritmo de Mineração de Dados. Neste trabalho foi utilizado o CART (*Classification And Regression Tree*) que faz uso da técnica de indução por árvore de decisão. No primeiro nível (imagem amplitude) constatou-se que os pixels com valores mais altos correspondiam às mudanças. Assim, foram coletadas amostras do que seria mudança (pixels mais claros) e também coletou-se amostras do que seria não-mudança (pixels não claros). Coletou-se amostras também nos níveis 2 e 3, só que neste caso para o que seria urbano e não-urbano.

Novamente realizou-se uma cópia do 1º nível e fez-se uso de algoritmo para remover a classificação deste nível. Ele foi nomeado de “Expansão”.

Assim, a partir das classificações nos níveis 1 (mudança e não-mudança), 2 (urbano e não-urbano) e 3 (urbano e não-urbano), foi elaborada uma regra de relacionamento entre os níveis para classificar aquilo que seria a Expansão. Exemplo: + *If is there super-objects like “Mudança” three levels above AND If is there super-objects like “Não\_Urbano\_1991” one level above AND If is there super-objects like “Urbano\_2000” two levels above = Expansão.*

A última etapa consistiu na avaliação das classificações por meio do cálculo do Índice *Kappa* (COLGATON & CREEN, 1999).

### 5. Resultados e discussão

A utilização das imagens amplitude apresentou-se como um bom recurso para a detecção das mudanças no terreno, porque separa bem a mudança, alvos mais claros, da não-mudança, alvos menos claros (Figura 3). No entanto, devido a área de estudo estar localizada em uma região litorânea, foram detectadas mudanças em alvos bem dinâmicos como dunas e linhas de costa (ação das ondas). O recurso utilizado para contornar isto foi justamente a classificação do que seria urbano e não-urbano.

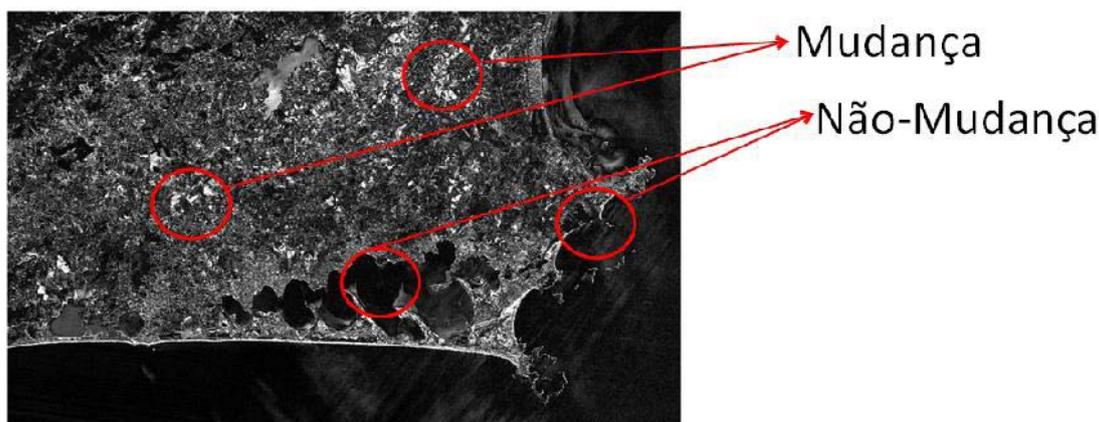


Figura 3 – Exemplo de imagem amplitude utilizado no trabalho.

O atributo selecionado pelo CART para separar a mudança da não-mudança tanto em 1991-2000 quanto em 2000-2010 foi o quartil de 75% (Figura 4).

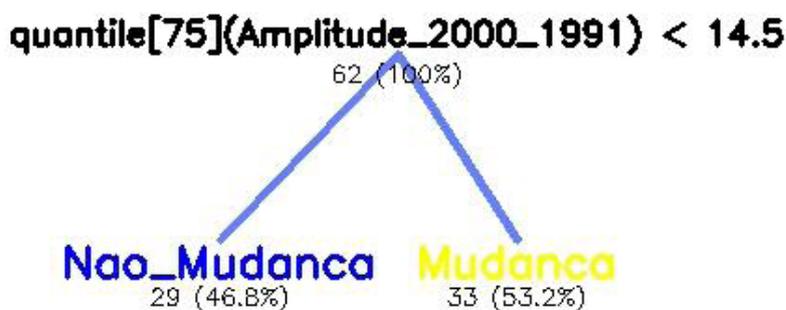


Figura 4 – Árvore de decisão empregada no primeiro nível hierárquico.

O emprego do MLME neste tipo de mapeamento mostrou-se como um bom recurso, visto que conseguiu separar urbano de não-urbano (Figura 5).

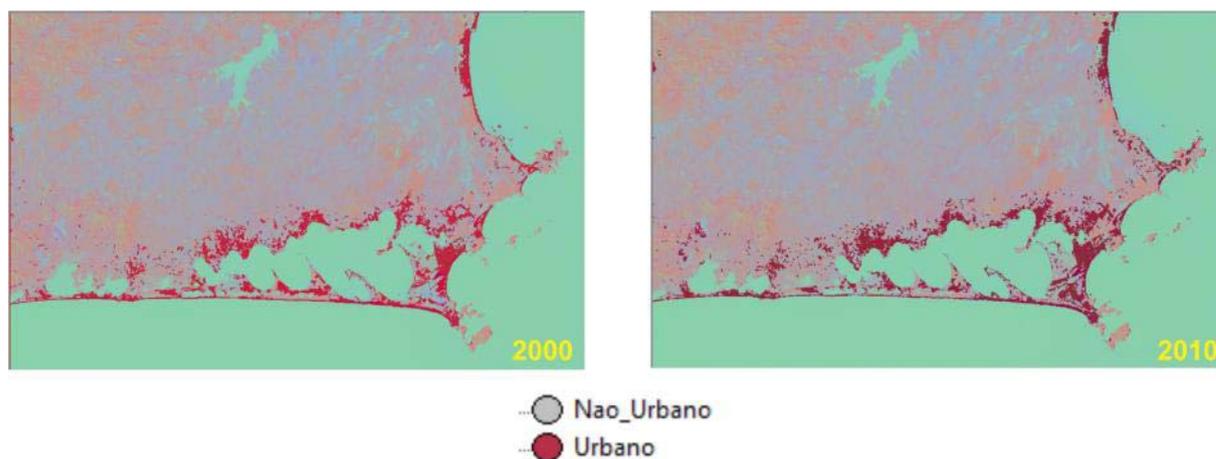


Figura 5 – Classificação gerada a partir do modelo de mistura para separar urbano de não-urbano para os anos de 2000 e 2010.

O mapa temático da expansão urbana (resultado final) para os períodos de 1991-2000 e 2000-2010 é apresentado na Figura 6.

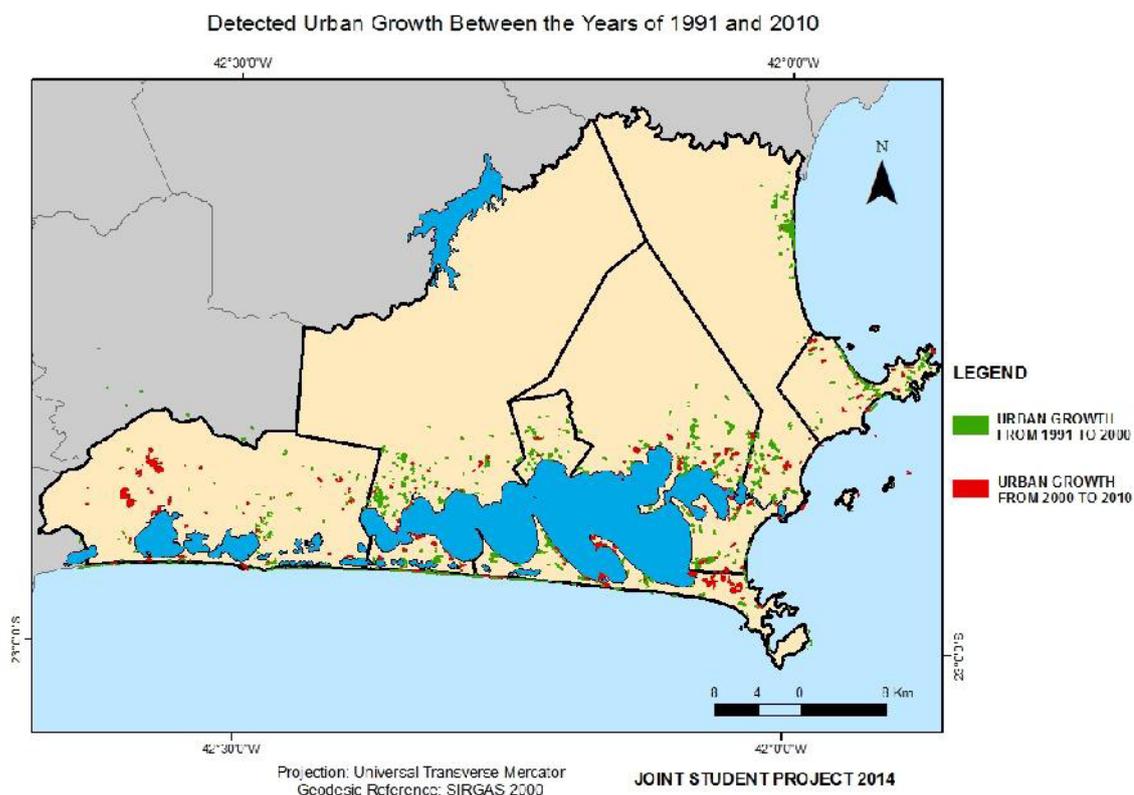


Figura 6 – Mapeamento temático da expansão urbana para o período em análise.

Por meio do cálculo do Índice *Kappa*, constatou-se que este mapeamento obteve o valor de 0,66. O que por sua vez, representa um bom resultado.

## 6. Conclusão

Mediante os resultados apresentados, podemos concluir que o objetivo do presente trabalho foi atendido, o de analisar a expansão urbana por meio da integração entre GEOBIA e Mineração de Dados.

O valor de Índice *Kappa* corrobora a adoção do MLME como atributo para o mapeamento da expansão urbana e da estratégia de relacionamento entre os níveis hierárquicos de objetos.

Assim sendo, a metodologia é passível de replicação. Sugere-se a aplicação da mesma em outras regiões do Estado do Rio de Janeiro onde a tipologia de urbanização é diferente. Sugere-se também a aplicação da metodologia em imagens provenientes de sensores com maior resolução espacial e a avaliação de outros índices provenientes de imagens orbitais e de outros algoritmos de segmentação.

## 7. Agradecimentos

Os autores expressam seus agradecimentos à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão de bolsa de estudos ao primeiro, segundo, quinto e sexto autor respectivamente.

## 8. Referências bibliográficas

ARAÚJO, E. H. G. Análise multi-temporal de cenas do satélite QuickBird usando um novo paradigma de classificação de imagens e inferências espaciais. Estudo de caso: Belo Horizonte, MG. 2006. 175 p. (INPE-13956- TDI/1062). Dissertação (Mestrado em Sensoriamento Remoto) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 2006.

BLASCHKE, T.; LANG, L.; LORUP, E.; STROBL, J.; ZEIL, P. Object-oriented image processing in an integrated GIS/Remote sensing environment and perspectives for environmental applications. In: CREMERS, A.; GERVE, K. (eds.). Environmental information for planning, politics and the public. Metropolis: Verlag, Marburg, 2000. v. 2, p. 555-570.

CÂMARA, G.; SOUZA, R.C.M.; FREITAS, U. M.; GARRIDO, J. C. P. Spring: Integrating Remote Sensing and GIS with Object-Oriented Data Modelling. Computers and Graphics, v.15, n.6, p.13-22, 1996.

CONGALTON, R.; GREEN, K. Assessing the Accuracy of remotely sensed data: principles and practices. Boca Raton, FL.: CRC/Lewis Press, 1999. 137 p.

DAI, X.L., KHORRAM, S. The effects of image misregistration on the accuracy of remotely sensed change detection, IEEE Tbars. on Geoscience and Remote Sensing, 36(5):1566-1577, 1998.

DEFINIENS. DEFINIENS Professional 7: Reference Book. Munich: DEFINIENS (The Imaging Intelligence Company), 2007. 122 p.

GOLDSCHMIDT, R.; PASSOS, E. Data mining - um guia prático. Rio de Janeiro: Elsevier, 2005. 257 p. ISBN (85-352-1877-7).

HAY, G.J.; CASTILLA, G. Geographic Object-based Image Analysis (GEOBIA): a new name for a new discipline. In: BLASCHKE, T.; LANG, S.; HAY, G.J. (Eds) Object-based image analysis- spatial concepts for knowledge-driven remote sensing applications. Berlin: Springer-Verlag, 2008.

JIANYA, G., HAIGANG, S., Guori, M., QIMING, Z. A Review of Multi-temporal Remote Sensing data Change Detection Algorithms. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, XXXVII(b7), pp. 757-762, 2008.

KENNEDY, R. E.; COHEN, W. B.; SCHROEDER, T. A. Trajectory-based change detection for automated characterization of forest disturbance dynamics. Remote Sensing of Environment, v. 110, n. 3, pp. 370-386, 2007.

MACLEOD, R. D., CONGALTON, R. G. A Quantitative Comparison of Change-Detection Algorithms for Monitoring Eelgrass from Remotely Sensed Data. *Photogrammetric Engineering & Remote Sensing*, 64(3), pp.207-216, 1998.

MILLER, J. H.; HAN, J. *Geographic data mining and knowledge discovery*. 2. ed. EUA: Taylor & Francis Group, 2009.

ROBERTS, S., HUSMEIER, D., REZEK, I., PENNY, W. Bayesian approaches to gaussian mixture modelling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1998, 20(11):1133-1142.

SANTOS, R. *Introdução à mineração de dados com aplicações em ciências ambientais e espaciais*, 2012. Disponível em: < [http://www.lac.inpe.br/ELAC/arquivos/MiniCurso\\_03ELAC2012.pdf](http://www.lac.inpe.br/ELAC/arquivos/MiniCurso_03ELAC2012.pdf)>.

VERONA, J. D. *Classificação e monitoramento fenológico foliar da cobertura vegetal na região da Floresta Nacional do Tapajós - Pará, utilizando dados multitemporais do sensor Thematic Mapper (TM) do Landsat*. São José dos Campos: INPE, 2002. 159p. - (INPE-9556-TDI/832).

VIEIRA, M. A. *Análise de imagem orientada a objeto e mineração de dados aplicadas ao mapeamento da cultura da cana-de-açúcar*. 2010. 110 p. (sid.inpe.br/mtc-m19@80/2010/04.08.17.43-TDI). Dissertação (Mestrado em Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2010.