

Sistema multiagente para mineração de imagens de satélite

Nicksson Ckayo Arrais Freitas¹
Ciro Daniel Gurgel de Moura²
Marcelino Pereira dos Santos Silva²

¹ LES – Laboratório de Engenharia de Software, Departamento de Informática, Universidade do Estado do Rio Grande do Norte (UERN) – Mossoró, RN – Brasil
nickssonarrais@gmail.com

² Programa de Pós-Graduação em Ciência da Computação, Universidade do Estado do Rio Grande do Norte (UERN) / Universidade Federal Rural do Semi-Árido (UFERSA) – Mossoró, RN – Brasil
{ciro.dgm, prof.marcelino}@gmail.com

Abstract. Image Mining is a field with huge potential and relevant challenges. For satellite images, this technique and resources as well algorithms can bring appropriate responses to important problems. However, due to this approach limitations, multiagent systems present features that, properly applied, can bring advances on pattern analysis in satellite images. In this context, the goal of this research is to present a methodology and a software for satellite image mining through multiagent systems. The prototype was developed in Java and uses resources from TerraView and GeoDMA.

Palavras-chave: remote sensing, data mining, classification, data mining agents, sensoriamento remoto, mineração de dados, classificação, agentes de mineração de dados.

1. Introdução

Uma das melhores formas de se obter informações sobre o nosso planeta é através dos dados obtidos pelos inúmeros satélites de sensoriamento remoto. E com a disponibilidade desses dados é possível realizar estudos em várias áreas.

Com a grande quantidade de informações que os satélites de sensoriamento remoto adquirem com o passar do tempo os bancos de dados dessas imagens são os que mais crescem. Esses bancos de dados de sensoriamento remoto são uma coleção de fotos da paisagem, que fornecem uma oportunidade única de compreender como, quando e onde mudanças ocorreram em nosso mundo.

A mineração de dados em bancos de imagens auxilia na exploração dessas informações através de técnicas, algoritmos e sistemas capazes de extrair conhecimento (informações estratégicas), relacionamentos espaciais e padrões interessantes não explícitos em grandes acervos de imagens.

Sistemas de agentes ou multiagentes e o processo de DCBD são duas áreas de pesquisa cada vez mais inter-relacionadas, que proporcionam benefícios para ambas as partes. Para Silva e Ralha (2011), essas duas áreas, a princípio, seguem com metas e objetivos distintos, porém há vários aspectos de ambas as áreas que coincidem, tais como: interação usuário-sistema, papéis humanos, modelagem dinâmica, fatores de domínio, fatores organizacionais e sociais, entre outros.

Nesse contexto, o objetivo deste trabalho é apresentar uma metodologia e respectiva ferramenta para a mineração de imagens através de sistemas multiagentes. O protótipo foi desenvolvido em Java e trabalha em conjunto com o software TerraView e o seu plugin, o GeoDMA.

2. Metodologia de trabalho

Esta seção apresenta os principais métodos, técnicas e ferramentas que foram utilizados durante o desenvolvimento deste trabalho.

2.1. Mineração de Imagens

A mineração de dados é uma das etapas no processo de Descoberta de Conhecimento em Banco de Dados (DCBD), como nos mostra a Figura 1. E essa etapa é responsável pela aplicação de algoritmos específicos para extrair padrões de dados (Fayyad, 1997). Para Ramos e Lobo (2004), os tipos de padrões a descobrir dependem da tarefa de mineração de dados empregue: previsão ou descrição. Os objetivos da previsão e descrição são concretizados recorrendo às tarefas primárias de mineração de dados: classificação, regressão, agrupamento ou clustering, sumarização, modelização das dependências e alterações e detecção de desvios.

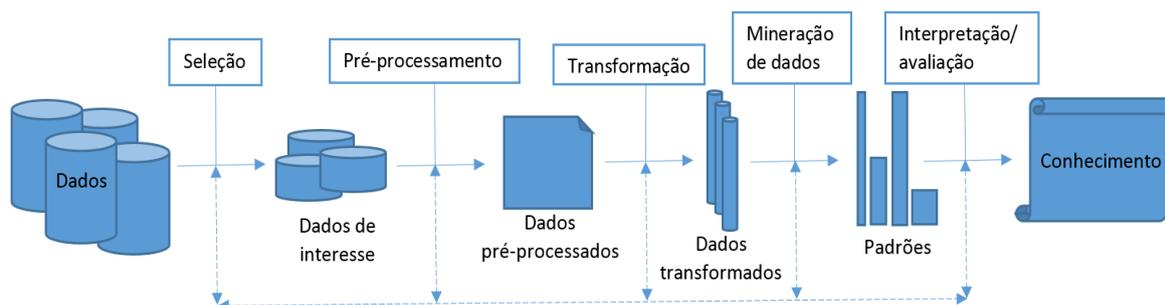


Figura 1. Processo de descoberta de conhecimento em banco de dados. Fonte: Adaptado de Fayyad (1997).

Na mineração de imagens, as imagens de um acervo (banco de imagens) são recuperadas segundo critérios inerentes à aplicação. A seguir, uma fase de pré-processamento aumenta a qualidade dos dados, os quais são então submetidos a uma série de transformações e de extração de características que geram importantes informações a respeito das imagens. A partir destas informações, a mineração pode ser realizada através de técnicas específicas, com o intuito de descobrir padrões significativos. Esses padrões resultantes são então avaliados e interpretados para a obtenção do conhecimento final, que pode ser aplicado no entendimento de problemas, na tomada de decisões ou outras aplicações estratégicas (Silva, 2006).

Dessa forma, a mineração de imagens de sensoriamento remoto trata especificamente do desafio de captar padrões, processos e agentes presentes no espaço geográfico, a fim de extrair conhecimento específico para entender ou para tomar decisões relativas a um conjunto de temas relevantes, incluindo a mudança da terra, variações climáticas e estudos de biodiversidade. Eventos como padrões de desmatamento, correlações de mudança climática e dinâmica de espécies são exemplos de preciosos conhecimentos contidos em repositórios de imagens de sensoriamento remoto (SILVA, CÂMARA e ESCADA, 2008).

2.2. Sistemas Multiagentes

Um agente é capaz de perceber seu ambiente por meio de sensores e de agir sobre este ambiente por intermédio de atuadores. Os quatro tipos básicos de agentes são o agente reativo simples, o agente reativo com estado, o agente baseado em objetivo e o agente baseado em utilidade (Russel e Norving, 2010). Um Sistema Multiagente (SMA) consiste de um número de agentes que interagem uns com os outros de forma cooperativa ou competitiva. No caso mais geral, os agentes em um SMA estarão representando ou agindo em nome de usuários ou proprietários com diferentes objetivos e motivações (Wooldridge, 2009).

Segundo Sycara (1998), as características de SMA são tais que: cada agente possui uma visão limitada; não há controle global do sistema; trabalham com dados descentralizados; a computação é assíncrona, de maneira a permitir a comunicação entre entidades heterogêneas. Para Weiss (2000) e Wooldridge (2009) uma das razões para o desenvolvimento da área é que

os conceitos subjacentes à SMA não são restritos a um único domínio de aplicação, sendo interessantes no desenvolvimento e análise de modelos e teorias de interatividade nas sociedades humanas. Isto ocorre devido ao SMA ser considerado uma metáfora natural para o entendimento e construção de uma ampla faixa do que podem ser rudemente denominados de sistemas artificiais sociais.

Aplicando os conceitos de agentes e SMA neste trabalho, cada agente realiza uma busca por padrões nas imagens de acordo com seu perfil e, dessa forma, o SMA minera a imagem como um todo.

2.3. GeoDMA

A ferramenta utilizada para minerar imagens é o GeoDMA (Geographical Data Mining Analyst). A ferramenta surgiu a partir das ideias propostas por Silva et al. (2005) e realiza todas as fases de processamento necessárias para manipular dados de sensoriamento remoto, incluindo os processos de segmentação, extração e seleção de atributos, treinamento, classificação e análise exploratória dos dados.

Segundo Korting et al. (2009), o GeoDMA funciona como um *plugin* para o sistema TerraView. Isto significa que toda a estrutura que manipula e visualiza bancos de dados geográficos é proporcionada pelo TerraView. Então o GeoDMA é executado em conjunto com o TerraView, e produz os resultados que são exibidos na sua tela principal.

A Figura 2 apresenta um diagrama geral do GeoDMA. As etapas de processamento começam a partir da definição dos dados de entrada, passando pela extração de características e a aplicação de algoritmos de mineração de dados para extrair e fornecer informações sobre observação da Terra (Korting, 2012).



Figura 2. GeoDMA: diagrama das principais etapas de processamento para a análise de imagens. Fonte: Adaptado de Korting (2012).

Dessa forma, uma imagem é inserida no Terraview e, utilizando o GeoDMA, é possível segmentar a imagem e extrair atributos espaciais e espectrais de cada um desses segmentos. Após a etapa de extração, na etapa de treinamento seleciona-se alguns dos segmentos e seus respectivos atributos, rotulando cada um desses segmentos com uma determinada classe. Com as amostras selecionadas nesta etapa, o próximo passo é a classificação, cujo algoritmo escolhido para este trabalho foi a árvore de decisão C4.5 (Quinlan, 1993), uma vez que o GeoDMA pode carregar uma árvore já existente ou salvar uma nova para que ela possa ser usada em futuras classificações.

2.4 Ferramenta para mineração de imagens através de sistemas multiagentes

A ferramenta foi desenvolvida em java, que é uma linguagem de alto nível orientada a objetos, que pode funcionar em qualquer máquina que possua uma máquina virtual java. Além de ser uma linguagem robusta e facilitadora para o desenvolvedor, Java é *multithread* o que possibilita várias tarefas em apenas um processo.

A Figura 3 apresenta o diagrama do processo. (1) O TerraView gerencia o banco de dados e manipula as imagens; (2) O GeoDMA segmenta as imagens e extrai os atributos espaciais e espectrais de cada um dos segmentos; (3) Esses atributos são armazenados no banco de dados; (4-a) A ferramenta então é conectada ao mesmo banco de dados utilizado pelo TerraView, (4-b) as tabelas que possuem os atributos obtidos são selecionados para criar o ambiente dos agentes; (4-c) neste ponto, o usuário cria ou escolhe os agentes que serão utilizados visando

atingir os seus resultados esperados. Cada agente possui sua árvore de decisão para que os segmentos possam ser minerados pelo mesmo. Finalmente, (4-c) com os dados classificados, (5) o usuário visualiza o resultado da mineração no TerraView.

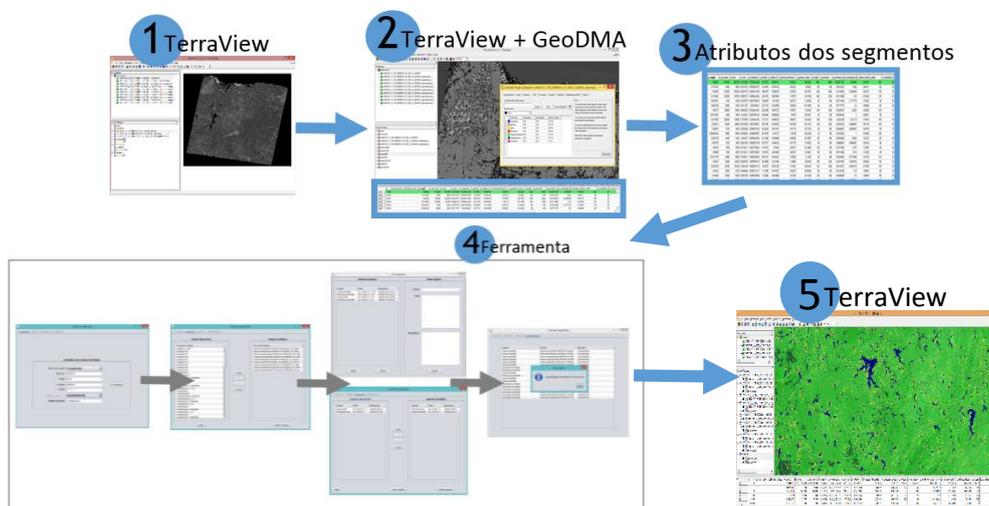


Figura 3. Diagrama do processo.

Os agentes são do tipo reativos simples, ou seja, são agentes que selecionam ações com base na sua percepção atual (Russel e Norving, 2010). A partir da sua árvore, cada agente possuirá um perfil que lhe permitirá minerar padrões específicos em novas imagens. Então, digamos que o agente A1 tenha o perfil de minerar corpos d'água. Na interface do protótipo o usuário terá recursos para selecioná-lo e atribuir ao mesmo a tarefa de minerar em novas imagens objetos que sejam mananciais, lagoas, açudes ou outros objetos pertinentes ao seu perfil.

Dessa forma, o agente A1 é capaz de identificar todos os segmentos que satisfaçam sua árvore de decisão, superando a limitação do GeoDMA, pois os agentes não minerarão apenas os segmentos de uma imagem, mas todos os segmentos das imagens que forem alocadas em seu ambiente.

3. Resultados e Discussão

O estado do Rio Grande do Norte possui grandes áreas de risco e algumas frações do território estão sofrendo com processos de desertificação. Um dos primeiros conceitos sobre Desertificação foi utilizado pelo engenheiro francês A. Aubreville, em 1949, que definiu o termo por “um conjunto de fenômenos que conduz determinadas áreas a se transformarem em desertos ou a eles se assemelharem”.

O mapa de ocorrência da desertificação no território potiguar deixa explícita a classe de intensidade segundo as regiões afetadas (Figura).

Como forma de demonstrar a metodologia e o protótipo desenvolvido, foi realizado um estudo buscando padrões que indiquem a existência de corpos d'água em imagens de satélite. Para isso, usou-se uma série histórica de imagens do satélite Landsat 5 TM dos anos de 1987, 1999 e 2008, e a cena 215/64.

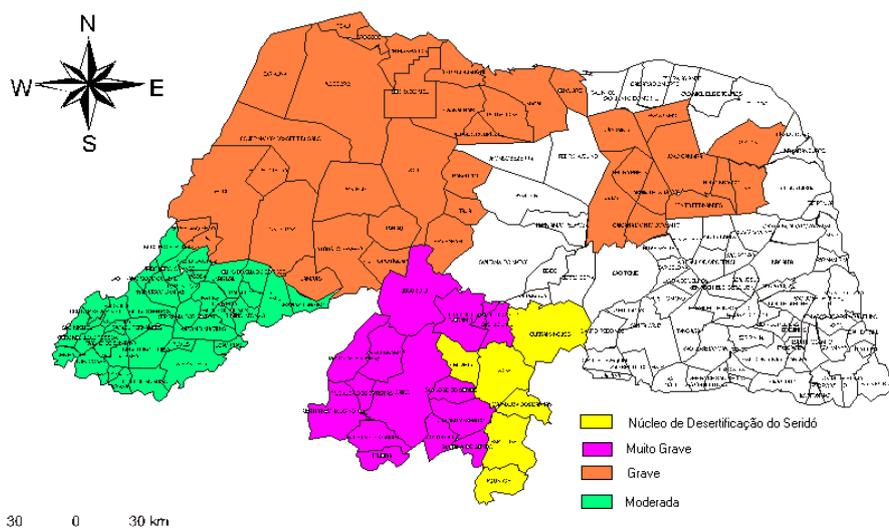


Figura 4. Mapa de Ocorrência da Desertificação no Rio Grande do Norte. Fonte: MMA (2005).

A Figura 5 apresenta o recorte que foi realizado abrangendo 13 municípios: Jucurutu, Timbaúba dos Batistas, São Fernando, Florânea, Cruzeta, São José do Seridó, Acari, São Vicente, Tenente Laurentino Cruz, Lagoa Nova, currais novos, Bodó e Cerro Corá. Esses municípios foram escolhidos pois pertencem a região de grande susceptibilidade ao fenômeno da desertificação.

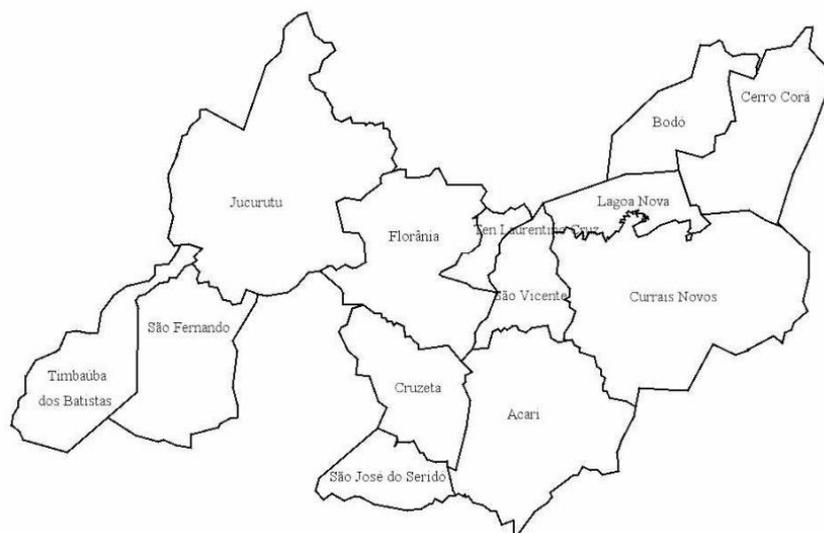


Figura 5. Mapa dos municípios (limites) analisados na região do Seridó - RN.

Para criar o perfil capaz de encontrar corpos d'água, foi utilizada a cena 214/64 e a banda selecionada foi a 5, pois corpos d'água são diferenciados de áreas com vegetação, áreas de cultivo, bem como solo exposto.

Utilizando o GeoDMA, as imagens foram segmentadas com a técnica de crescimento de regiões e os valores para similaridade e área, foram 10 e 15, respectivamente. A etapa de treinamento foi realizada apenas na imagem de 1987. Após a etapa de treinamento, uma árvore de decisão foi gerada pelo GeoDMA para classificar a imagem de 1987. Essa árvore gerada serviu de perfil para um agente que foi o responsável pela classificação das outras duas imagens.

A Figura 6 apresenta a imagem de 1987, classificada com o GeoDMA. As Figuras 7 e 8, apresentam as outras duas imagens, 1999 e 2008, respectivamente, classificadas com a ferramenta. Da mesma forma como foi possível classificar utilizando um agente, outros agentes poderiam ter sido criados com diferentes perfis (lavoura, solo exposto etc.) e usados para minerar as imagens.

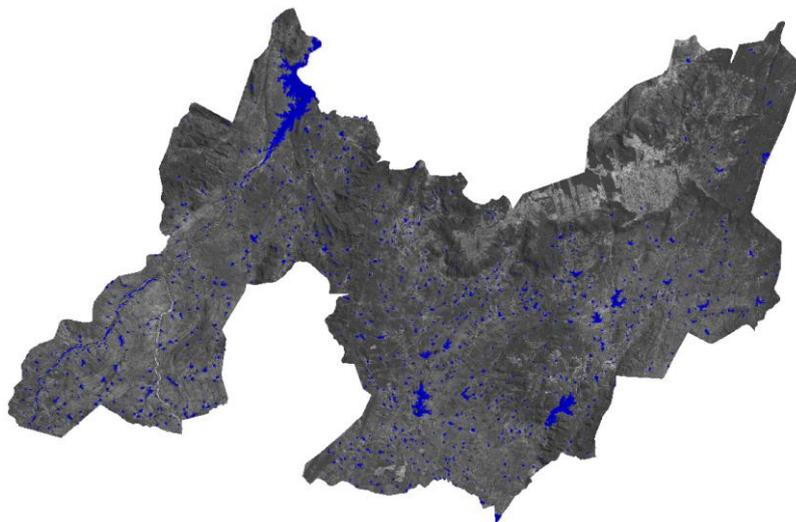


Figura 6. Imagem de 1987 classificada utilizando o GeoDMA.

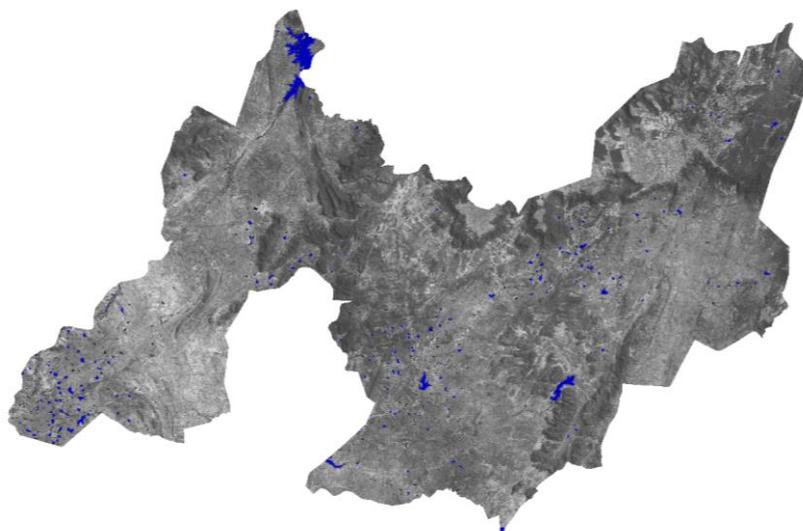


Figura 7. Imagem de 1999 classificada pela ferramenta.

4. Conclusões

Este trabalho propõe uma metodologia e ferramenta que utiliza as características de agentes para auxiliar no processo de extração de conhecimento de imagens de satélite. Esses agentes são capazes de minerar dados geográficos a partir das características espaciais e espectrais dos objetos presentes nas imagens de sensoriamento remoto. O SMA, através dos diferentes perfis dos seus agentes, tem a funcionalidade, após devidamente treinados, de minerar objetos de diferentes padrões (solo exposto, vegetação densa, corpos d'água, dentre outros) em um repositório de imagens.

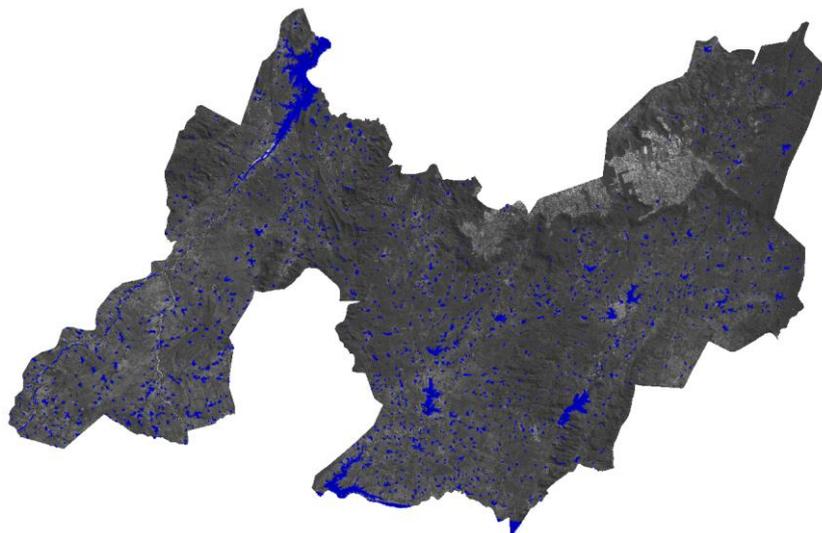


Figura 8. Imagem de 2008 classificada pela ferramenta.

Agradecimentos

Primeiramente a Deus. Agradecemos à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo auxílio financeiro no desenvolvimento desta pesquisa.

Referências bibliográficas

- Aubreville, A. (1949). *Climats, forets et desertification de l'Afrique tropicale*. Paris, Geogr. Marit. & Col., 351 p.
- Fayyad, U. Data Mining and Knowledge Discovery in Databases: Implications for Scientific Datadabase. Proc. of the **9th International Conference on Scientific and Statistical Database Management**, pp. 2-11, 1997.
- Korting, T. S. **GeoDMA: a toolbox integrating data mining with object-based and multi-temporal analysis of satellite remotely sensed imagery**. 2012. 123f. Tese (Doutorado em Sensoriamento Remoto) – INPE, São José dos Campos, 2012.
- Korting, T. S.; et al. GeoDMA - Um sistema para mineração de dados de sensoriamento remoto. In: **SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO**, 14, 2009, Natal. Anais do XIV Simpósio Brasileiro de Sensoriamento Remoto. Natal: INPE, 2009. p. 7813-7820.
- MMA. Secretaria de Recursos Hídricos. **Programa de ação nacional de combate à desertificação e mitigação dos efeitos da seca**. Ministério do Meio Ambiente (MMA). [S.l.], 2004. 220p.
- Quinlan, J. R. **C4.5: Programs for Machine Learning**. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1993. ISBN 1-55860-238-0.
- Ramos, C.; LOBO, F. **Descoberta de Conhecimento em Bases de Dados**. Dos Algarves - A Multidisciplinary E-Journal Revista da ESGHT/UALG. Edição 12. Campus da Penha: UALG, junho 2004. ISSN: 0873-7347. Disponível em: <<http://www.dosalgarves.com/revistas/N12/7rev12.pdf>>. Acesso em: 30.out. 2013.
- Russel, S. Norving, P. **Artificial Intelligence: A Modern Approach**. 3ª. Ed. Prentice Hall. 2010. ISBN-10: 0-13-604259-7.
- SANTOS, A. R. et al. **Sensoriamento Remoto no ArcGis 10.2.2 passo a passo: Processamento de Imagens Orbitais** – Volume 1. Alegre: CAUFES, 2014. 107 p. Disponível em <<http://www.mundogeomatica.com.br/LivroSR102.htm>>
- Silva, C. V. S. e Ralha, C. G. **Detecção de Cartéis em Licitações Públicas com Agentes de Mineração de Dados**. RESI - Revista Eletrônica de Sistema de Infomação. v. 10, n. 1, artigo 8. 2011. ISSN: 1677-3071. Disponível em: <<http://revistas.facecla.com.br/index.php/reinfo/article/view/754/pdf>>. Acesso em: 24.out.2013.

Silva, M. P. S. et al. Mining patterns of change in remote sensing image databases. In: ICDM '05: Proceedings of the **Fifth IEEE International Conference on Data Mining**. Washington, DC, USA: IEEE Computer Society, 2005. p. 362–369. ISBN 0-7695-2278-5.

Silva, M. P. S. **Mineração de dados em bancos de imagens**. 2006. 123f.. Tese (Doutorado em Computação Aplicada) - INPE, São José dos Campos, 2006.

SILVA, M. P. S.; CÂMARA, G.; ESCADA, M. I. S. Image Mining: Detecting Deforestation Patterns Through Satellites, p. 55-75, 2008. **Data Mining Applications for Empowering Knowledge Societies**. IGI Global. 2008.

Sycara, K. P.. **Multiagent systems**. AI Magazine, 19(2):79–92, 1998.

Weiss, G. **Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence**. The MIT Press, 2000.

Wooldridge, M. **An Introduction to MultiAgent Systems**. John Wiley & Sons, LTD, Chichester, England, 2nd edition, 2009.