

Análise de classificadores para mapeamento de uso e cobertura do solo

Pétala Bianchi Augusto-Silva¹
Larissa Patrício Valério¹
Thiago Batista dos Santos¹
Enner Herenio de Alcântara²
José Luiz Stech¹

¹ Instituto Nacional de Pesquisas Espaciais - INPE
CEP - 12227-010 - São José dos Campos - SP, Brasil
{petala, larissa, batista, stech}@dsr.inpe.br

² Universidade Estadual Paulista Júlio de Mesquita Filho - UNESP
CEP - 19060-900 – Presidente Prudente - SP, Brasil
enner@fct.unesp.br

Abstract. Land Use and Land Cover (LULC) maps have been developed in order to guide decision-making upon spatial data. This allows the construction of indicators for assessing the support capacity of the environment. In the present work, a map of land cover and use was developed with a TM-Landsat-5 image beyond a multispectral classification scheme. Methods of classification available in the software SPRING were tested to evaluate their performance in organize the study area based on five thematic classes defined by FAO. A reference map was developed by an interpreter based on visual classification aided with high-resolution images from *Google Earth*. Sample points were selected for comparison between the reference map and the ones automatically classified, and the performance of the classifying methods was evaluated based on the percentage of rights. The *Isoseg* has the best percentage (83,79%) of rights, but since it is a non-supervised classifier, it separates the scene into much more themes than the others, so we have to remap these themes into the original classes of FAO. That's why *Bhattacharya* was considered the best method with 77,25% of rights. A class named "antrophic" was the one with the worst performance probably because some objects can get mixed up with the spectral response of soil prepared for cultivation. The methods in general can perform well, but in fact they should be used with the aid of interpreter knowledge and knowing the precision it's required for the work.

Palavras-chave: Land cover and use maps, remote sensing, classifiers, mapa de uso e cobertura, sensoriamento remoto, classificadores.

1. Introdução

A elaboração de mapas de uso e cobertura tem como fundamento a interpretação, análise e registro de informações no que concerne a observação da paisagem em foco (IBGE, 2006). Representando uma ferramenta de planejamento e orientação à tomada de decisão, estes mapas auxiliam a construção de indicadores ambientais para avaliação da capacidade suporte do meio. Desta forma, os mapas de uso e cobertura atuam fornecendo subsídios para as análises e avaliações dos impactos ambientais.

O levantamento de uso e cobertura da terra consiste na elaboração de mapas indicativos da distribuição geográfica dos usos através da identificação de padrões homogêneos da cobertura terrestre (IBGE, 2006). A consciência de que as modificações implementadas no uso e cobertura da terra implicam em impactos no clima local, global e no funcionamento dos ecossistemas, trouxe uma nova linha de pesquisa para as ciências naturais e da terra (LAMBIN et al., 2003). Sendo assim, o monitoramento de uso e cobertura envolve capacitação multidisciplinar e serve de subsídio para diversos órgãos, sejam eles governamentais, privados, ou de pesquisa.

Os termos uso e cobertura são utilizados em conjunto, apesar de terem significados distintos. Cobertura se refere aos componentes bióticos e abióticos que revestem a superfície da terra (e a subsuperfície imediata) englobando solos e rochas, vegetação e florestas, água ou neve, refletindo o estado biofísico do sistema terrestre (PRAKASAM, 2010) e dependendo do

ponto de vista, pode incluir a infraestrutura urbana, pois esta também cobre a superfície do globo (CCRS, 2007). O uso por sua vez faz referência às atividades humanas executadas diretamente sobre a terra, descrevendo de que maneira a cobertura se modifica, como agricultura e pastagem, construções, lazer e turismo, mineração, desflorestamento, envolvendo a modificação dos ambientes naturais (PRAKASAM, 2010).

O advento do sensoriamento remoto com o auxílio de tecnologias computacionais impulsionou os levantamentos de uso e cobertura (CHILAR, 2000). A perspectiva do observador foi estendida em amplas escalas espaciais e temporais, tornando o acompanhamento das mudanças de forma rápida e com baixo custo. As observações efetuadas pelos sensores orbitais proporcionam medidas espaciais dos atributos biofísicos em superfície, tais como a cobertura de vegetação, conteúdo de água e de matéria orgânica no solo, e heterogeneidades presentes nas paisagens naturais e antrópicas (LAMBIN et al., 2003). Alocar informações que permitem entender o passado, monitorar o presente e modelar as condições futuras de uso e cobertura da terra, torna-se ferramenta principal de manejo e conservação dos recursos naturais (PRAKASAM, 2010).

Com o avanço das geotecnologias, têm-se usado classificadores para identificar as distintas classes de uso e cobertura do solo. Estes classificadores possuem por finalidade, reconhecer padrões em uma imagem representativa da superfície da terra, de acordo com temas que sejam de interesse do usuário. Sendo assim, os classificadores procuram simular o comportamento de um intérprete, ao reconhecer áreas homogêneas na imagem (CÂMARA et al., 1996).

Neste trabalho empregou-se a distribuição dos níveis de cinza em três canais do sensor TM a bordo do satélite Landsat-5: o canal 5 foi empregado no vermelho; o 4 no verde e o 3 no azul seguindo a metodologia descrita por Galo et al. (2002) para facilitar a discriminação dos alvos. A classificação utiliza os níveis de cinza de cada canal sendo realizado por métodos supervisionados ou não supervisionados (automáticos). O método supervisionado é realizado quando o usuário tem disponíveis amostras que permitam identificar na imagem o alvo de interesse (CÂMARA et al., 1996). Os classificadores não supervisionados usam algoritmos que calculam distâncias estatísticas e limiares para reconhecer padrões na cena sem a intervenção de um intérprete. Desta forma os níveis de cinza associados a cada pixel são identificados em termos de um tipo de cobertura da superfície imageada, denominado classe (CRÓSTA, 1992). Neste caso o máximo de interação do usuário se dá na escolha do número de classes a serem separadas.

Câmara et al. (1996) versam sobre outra diferenciação feita entre classificadores: classificadores por pixels ou por regiões. Os classificadores por pixels utilizam apenas a informação espectral de cada pixel para achar regiões homogêneas. Já os classificadores por regiões utilizam, além de informação espectral de cada pixel, a informação espacial que envolve a relação entre os pixels e seus vizinhos. Há um pré-reconhecimento de padrões definido pela segmentação da cena e, posteriormente, é estimada uma distância estatística e um limiar de separação entre esses padrões que será usado na identificação dos alvos.

Neste contexto a motivação deste trabalho foi analisar o desempenho de uma série de classificadores disponíveis no software SPRING (Sistema de Processamento de Informações Georreferenciadas, CÂMARA et al., 1996), na identificação de classes de uso e cobertura do solo usando-se uma imagem TM-Landsat-5. Os classificadores foram utilizados com a finalidade de organizar a área de estudo com base em cinco classes temáticas de uso e cobertura da terra definidas de acordo com o Sistema de Classificação de Cobertura do Solo (SCCS), da FAO (*Food and Agriculture Organization of the United Nations*; DI GREGÓRIO, 2005), conforme descrito na Tabela 1.

Tabela 1: Tabela de Referência para a Definição de Classes de Uso e Cobertura. Adaptado de Nascimento et al. (2009)

Áreas cultivadas e gerenciadas	Áreas onde a vegetação natural foi removida ou modificada e substituída por outros tipos de cobertura vegetal de origem antrópica, essa vegetação requer cuidados humanos para mantê-la em longo prazo.
Vegetação natural e semi-natural	Áreas onde a cobertura vegetal está em equilíbrio com os fatores bióticos e abióticos, e vegetação não plantada pelo homem, mas influenciada por suas atividades, respectivamente. Em vegetação semi-natural encontram-se os pastos.
Superfícies artificiais e áreas associadas	Referem-se a áreas que possuem cobertura artificial resultante de atividades humanas como, por exemplo, cidades e vias de transporte.
Áreas expostas	Áreas que não apresentam uma cobertura artificial resultante de atividades humanas. Essas áreas incluem áreas com menos de 4% de cobertura vegetal, como rochas e desertos.
Corpos d'água naturais e artificiais	Áreas naturalmente cobertas por água, como lagos e rios, e áreas cobertas por água devido à construção de artefatos, como reservatórios, canais e lagos artificiais.

2. Metodologia de Trabalho

2.1 Área de Estudo

A área de estudo deste trabalho compreende a região entre os municípios de Promissão SP e Buritama-SP, localizada nos arredores da barragem do reservatório hidrelétrico de Promissão, pertencente à cascata de reservatórios do rio Tietê (Figura 1).

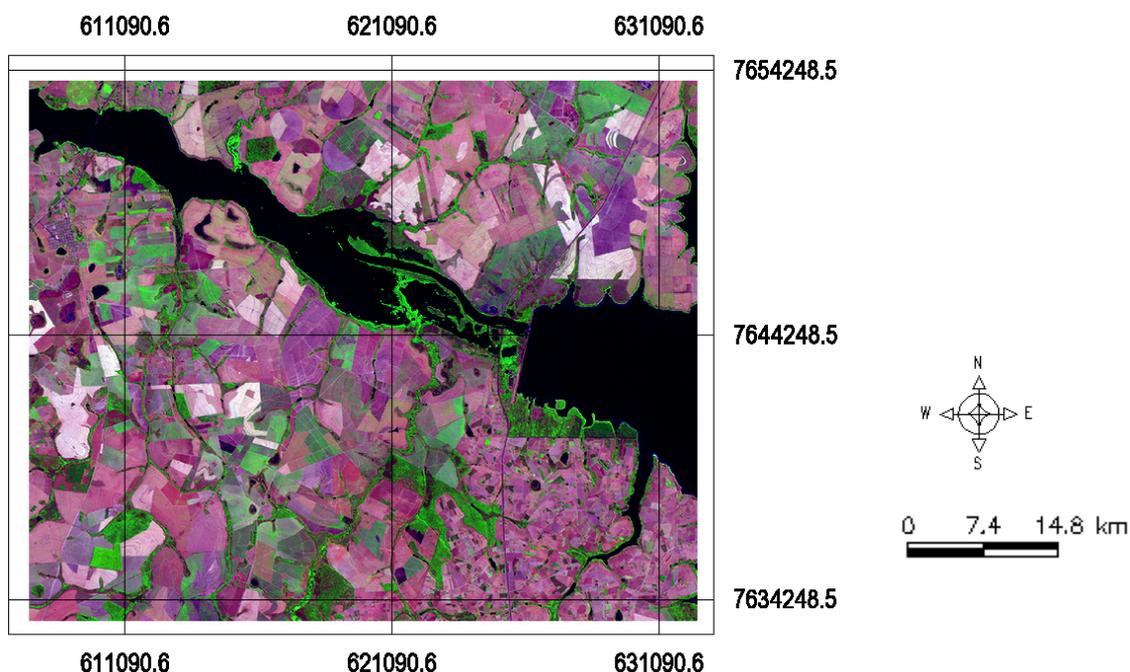


Figura 1: Localização da área de estudo. Composição colorida (5R4G3B) de imagens do TM-Landsat-5, órbita/ponto: 221/75; data de passagem: 26/09/2011.

2.2 Material

2.2.1 Dados Orbitais

Para a execução desse trabalho foi adquirida uma cena do sensor TM-Landsat-5 referente à área de estudo (órbita/ponto: 221/75; data de passagem: 26/09/2011). A cena foi adquirida no catálogo online do INPE (Instituto Nacional de Pesquisas Espaciais). Além desta cena, uma imagem GEOCOVER foi obtida no catálogo da NASA (*National Aeronautics and Space Administration*) para que fosse realizada a correção geométrica da cena TM-Landsat-5.

2.2.2 Aplicativo

O aplicativo computacional utilizado neste trabalho foi o SPRING (Câmara et al. 1996), versão 5.1.8, disponível para os usuários no sítio digital do INPE.

2.3 Método

Inicialmente a cena TM-Landsat-5 passou por um pré-processamento padrão, ou seja, as bandas utilizadas foram restauradas e registradas. A restauração é importante para corrigir as distorções inseridas pelo sensor na geração das imagens (GONZALEZ e WOODS, 2000). Já a correção geométrica é necessária para posicionar as bandas de acordo com uma referência (CENTENO, 2009), no caso, a GEOCOVER do catálogo da NASA. Após esse processo, as bandas passaram por um realce visando à melhoria da qualidade visual da imagem de forma a facilitar a sua interpretação (IBGE, 2001).

Para a obtenção de um mapa de referência, com o qual as diferentes classificações seriam comparadas, a composição pré-processada foi então classificada por um intérprete com o auxílio de imagens de alta resolução disponíveis no aplicativo *Google Earth*. Esta classificação visual levou em conta o conhecimento do intérprete para discriminar alvos em uma composição de acordo com a resposta esperada de cada alvo, ou seja, levando em conta o comportamento espectral das classes de alvo.

Em seguida, as bandas 543 pré-processadas foram segmentadas testando heurísticamente diversos valores para os parâmetros de similaridade e de área. Escolheu-se a segmentação que usou similaridade igual a 15 e área igual a 40, que obteve melhor desempenho em separar os objetos da cena de acordo com os objetivos específicos deste trabalho. Esta segmentação serviu de base para as classificações por regiões que foram realizadas.

Os classificadores Bhattacharya, Histograma, Iseseg, K-médias, Maxver e Distância Euclidiana, todos disponíveis no SPRING, foram aplicados às bandas pré-processadas. Os resultados finais dos classificadores foram comparados com o mapa de referência editado pelo intérprete. Na comparação foram utilizados 400 pontos amostrais aleatórios espalhados por toda área de estudo. Pôde-se então analisar o desempenho de cada classificador pela obtenção da porcentagem de acerto e índice kappa de cada um. De acordo com Landis e Koch (1977) índices kappa que estão entre 0,61 e 0,80 são considerados robustos. Essa análise foi realizada no próprio SPRING, em uma ferramenta de análise chamada validação.

3. Resultados e Discussão

Os resultados referentes à validação dos classificadores encontram-se sumarizados na Tabela 2, e os classificadores estão divididos em por regiões ou por pixel, como apresentado no SPRING.

Sobre o percentual de acertos, para a classificação por pixels o Maxver foi o que obteve melhor desempenho (76,06%). Já para a classificação por regiões, a maior porcentagem de acertos foi do Iseseg (83,79%). Com relação ao índice kappa, os únicos classificadores que não obtiveram um índice robusto foram o classificador por histograma e o K-Médias.

Tabela 2: Desempenho geral dos classificadores

Tipo	Classificação	% de Acertos	Kappa
Classificadores por Região	Bhattacharya	77,25	0,6584
	Histograma	69,58	0,5575
	Isoseg	83,79	0,6996
Classificadores por Pixels	K-Médias	60,10	0,4534
	Maxver	76,06	0,6212
	Dist. Euclidiana	74,31	0,6113

Deve ser ressaltado que o classificador Isoseg é um classificador não-supervisionado (automático). No caso desta área de estudo, este classificador separou a área em 37 temas distintos que foram remapeados para os temas definidos pela FAO a partir da intervenção do intérprete. Já para os classificadores Histograma e K-médias o usuário deve definir a quantidade de temas que devem ser gerados, que neste caso foram os quatro temas da FAO. Por esse motivo é que todos os classificadores, até mesmo os supervisionados, têm um desempenho inferior ao Isoseg. Ao classificar as imagens, o Isoseg define o número de temas em função das distâncias estatísticas e do limiar de separação estimados. Esta característica apresentou-se vantajosa, pois com um número muito maior de temas, ele possibilitou isolar e classificar áreas espectralmente similares que em outros classificadores foram englobadas em outras classes reduzindo a exatidão dos produtos gerados.

Diante da grande heterogeneidade presente na área, principalmente na classe cultivo que detém estágios muito distintos de evolução do plantio, a fase de treinamento dos classificadores supervisionados ficou limitada devido ao pequeno número de classes de interesse. Visando elucidar essa problemática a Tabela 3 mostra como o desempenho do classificador Histograma e o K-médias pode ser melhorado pelo aumento do número de temas e posterior remapeamento destes. Nota-se um significativo incremento na classificação realizada pelo K-médias e Histograma, o que corrobora com esta teoria.

Tabela 3: Desempenho dos classificadores não supervisionados com mais temas de entrada.

Classificação	% de Acertos	Kappa
Histograma com 10 temas	81,80	0,6880
Isoseg com 37 temas	83,79	0,6996
KMédias com 10 temas	85,79	0,7495

Deve-se ser ressaltar que em todas as classificações testadas o maior erro apareceu na classe de áreas antrópicas (Tabela 4), que é em geral englobada pela classe cultivo. Isso se dá pelo fato de que uma área classificada como antrópica (uma estrada de terra, por exemplo), ter uma resposta espectral similar a uma área de cultivo em estágio inicial (com grande parcela de solo exposto). Para testar essa hipótese, o nível digital de áreas que apresentaram esse tipo de confusão foram analisados e notou-se que, nas três bandas usadas para a composição colorida (bandas 3, 4 e 5), os valores de nível digital são semelhantes para ambas as classes.

Tabela 4: Desempenho dos classificadores para a classe antrópica.

Tipo	Classificação	% de Acertos para a Classe Antrópica
Classificadores por Região	Bhattacharya	26,74
	Histograma	17,30
	Histograma com 10 temas	2,94
	Isoseg	30,00
Classificadores por Pixels	KMédias	11,38
	KMédias com 10 temas	32,26
	Maxver	18,48
	Dist. Euclidiana	19,28

Baseando-se então nos resultados obtidos para a classe de uso antrópico, o melhor classificador foi o K-médias com 10 temas, que obteve um desempenho de 32,26% de acertos. Esse desempenho superior apresentado para esta classe se deve essencialmente à natureza do K-médias como classificador por pixels, visto que o ato de segmentar a imagem pode não isolar adequadamente essa classe, que tem uma distribuição espacial reduzida em comparação com as outras para a área de estudo.

Para que os classificadores não supervisionados tenham um desempenho bom, estes devem ter um número elevado de temas de entrada, os quais devem ser remapeados pelo usuário/intérprete. Assim, em casos em que se disponha de tempo para o remapeamento de classes, fica evidente que os classificadores não supervisionados tendem a ter um desempenho muito bom. Já considerando como entrada apenas os quatro temas em que havia interesse, o melhor classificador foi o Bhattacharya, classificador supervisionado por regiões, que teve um percentual de acertos global de 77,25%, que para a finalidade deste estudo foi considerado alto (Figura 2).

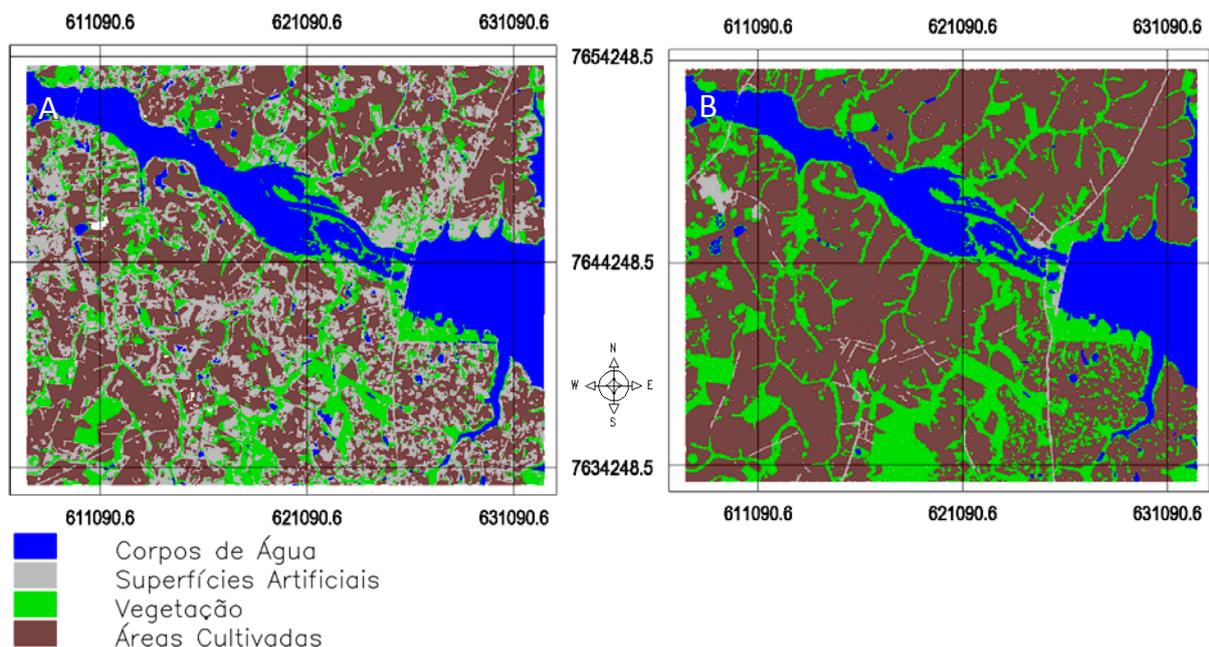


Figura 2: A-Classificação Bhattacharya; B- Mapa de referência editado manualmente.

5. Conclusão

No desenvolver deste trabalho foi possível testar o desempenho de diferentes classificadores disponíveis no software SPRING para mapear o uso e a cobertura de uma área. Considerando como entrada os quatro temas de interesse, o melhor classificador foi o Bhattacharya.

Dessa forma, conclui-se que a classificação visando o mapeamento do uso e cobertura do solo pode ser bem sucedida utilizando-se classificadores. O usuário deve ter em mente suas necessidades e, a partir delas definir o desempenho mínimo que ele espera que seja alcançado e, se o uso apenas dos classificadores for insuficiente para atingir seus objetivos, a intervenção de um usuário/intérprete para remapear e editar classificações errôneas ainda é menos trabalhosa do que a edição manual de uma cena inteira.

6. Referências

CAMARA G, SOUZA RCM, FREITASUM, GARRIDO J, 1996. SPRING: Integrating remote sensing and GIS by object-oriented data modelling. **Computers & Graphics**, 20: (3) 395-403.

CCRS, 2007. **Fundamentals of Remote Sensing**. Canada Centre for Remote Sensing, Natural Resources Canada, 258 p.

CENTENO, J. A. S. **Sensoriamento remoto e processamento de imagens digitais**. Curitiba: Ed. Curso de pós-graduação em ciências geodésicas: Universidade Federal do Paraná. 219 p., 2009.

CHILAR, J., 2000 Land cover mapping of large areas from satellites: status and research priorities. **International Journal of Remote Sensing**, v. 21, n. 6 & 7, p. 1093–1114.

CRÓSTA, A. P. **Processamento digital de imagens de sensoriamento remoto**. Campinas: UNICAMP, 170p., 1992.

DI GREGÓRIO, A., 2005. Land Cover Classification System (LCCS), version 2: Classification Concepts and User Manual. **FAO Environment and Natural Resources Service Series**, N° 8 - FAO, Rome. 208p.

GALO, M.L.B.T.; VELINI, E.D.; TRINDADE, M.L.B.; SANTOS, S.C.A. (2002) Uso do Sensoriamento Remoto Orbital no Monitoramento da Dispersão de Macrófitas nos Reservatórios do Complexo Tietê. **Planta Daninha**: Minas Gerais: Viçosa. V. 20, p 7-20.

GONZALEZ, R. C.; WOODS, R. E. **Processamento de imagens digitais**. 264 p., 2000.

IBGE. Introdução ao processamento digital de imagens. **Manuais Técnicos em Geociências** n° 9, Instituto Brasileiro de Geografia e Estatística, 91 p., 2001.

IBGE, 2006. Manual Técnico de Uso da Terra, 2ª edição. **Manuais Técnicos em Geociências** n° 7, Instituto Brasileiro de Geografia e Estatística, 91 p.

Landis, J.; Koch, G. G. The measurements of observer agreement for categorical data. *Biometrics*, v. 33, n. 1, p. 159-174, Mar 1977.

LAMBIN, E. F., GEIST, H. J., LEPERS, E., 2003. Dynamics of Land-use and Land-cover change in tropical region. **Annual Review of Environmental Resources**, v. 28, p. 205–241.

NASCIMENTO, R.F.F.; ALCÂNTARA, E.H.; KAMPEL, M.; STECH, J.L.; NOVO, E.M.L.M.; FONSECA, L.M.G., 2009. O algoritmo Support Vector Machines (SVM): avaliação da separação ótima de classes em imagens CCD-CBERS-2 Anais XIV Simpósio Brasileiro de Sensoriamento Remoto, Natal, Brasil, 25-30 abril, INPE, p. 2079 - 2085.

PRAKASAM, C., 2010. Land use and land cover change detection through remote sensing approach: A case study of Kodaikanal taluk, Tamil nadu. **International Journal of Geomatics and Geosciences**, v.1, n. 2, p. 150-158.