

Utilização de sensoriamento remoto mediante classificação supervisionada de imagens digitais pelo método *Support Vector Machine* para o reconhecimento de padrões de uso do solo

Thays Marcela Hamulak¹, Selma Regina Aranha Ribeiro².

¹Trabalho executado com recursos da Fundação Araucária.

² Estudante do curso de Licenciatura em Geografia; Universidade Estadual de Ponta Grossa; Ponta Grossa, Paraná; tha.m.hamulak@hotmail.com;

² Professora Doutora adjunta ao Departamento de Geociências; Universidade Estadual de Ponta Grossa; Ponta Grossa, Paraná; selmar@gmail.com;

RESUMO: O sensoriamento remoto está cada vez mais presente no cotidiano das pessoas, e nos últimos anos vem se mostrando indispensável na questão de monitoramento de uso e ocupação da terra, principalmente por utilizar métodos que facilitam o trabalho de classificação de uso por métodos de análise de imagens digitais que otimizam o tempo. Neste trabalho foram realizados os pré-processamentos cabíveis e a classificação supervisionada pelo método algoritmos máquina de vetores de suporte (SVM – *Support Vector Machine*), para separar as classes de uso do solo presentes na área de São Mateus do Sul – Paraná – Brasil. Utilizando – se para o estudo, uma cena do satélite Landsat 5 sensor TM. Notou-se que a classificação resultou a categorização de classes de forma satisfatória quanto ao estudo proposto, visto que as informações de uso da terra geradas coincidem com as observadas em campo. Sendo assim, este trabalho apresenta a importância dos processamentos de imagens digitais e do sensoriamento remoto para o reconhecimento de padrões de uso e ocupação da terra.

INTRODUÇÃO

A aplicação do sensoriamento remoto vem sendo difundida em grande escala no que diz respeito ao mapeamento do solo, cuja rápida modificação no século XX tornou ineficiente os sistemas de tradicionais de aerolevante devido ao custo de aquisição. Nas últimas três décadas o sensoriamento remoto se tornou a fonte primária de informações sobre a cobertura do solo, em decorrência dos avanços tecnológicos nos sistemas sensores e da ampliação do número de países que operam Sistemas de Observação da Terra (Foody e Mathur, 2004).

Os satélites e sensores da série Landsat e SPOT de resolução espacial de 30m e 20 m, respectivamente, geram imagens adequadas para a caracterização do uso do solo a nível regional, estes que vem sendo utilizados em vários projetos como por exemplo o CORINE (*Co-ordination of Information on the Environment*) Land Cover da Agência Europeia do Ambiente da União Europeia (Instituto do Ambiente, 2005).

A classificação consiste em gerar uma nova imagem, mais simples, a partir dos valores do contador digital de uma imagem com várias bandas, onde cada pixel é associado a uma categoria; onde o produto final é uma única imagem com poucos valores, e esses produtos podem ser considerados mapas temáticos. (Centeno, 2004)

Os métodos de classificação podem ser divididos em duas categorias: supervisionada e a não supervisionada. A supervisionada o usuário define as áreas de interesse na imagem, e na não supervisionada usa critérios de análises estatísticas dos dados de amostragem, tais como, média, desvio padrão, variância e outros. (LIU, 2006)

No trabalho foi realizada a transformação geométrica e correção atmosférica, estes procedimentos são considerados processamentos iniciais e fundamentais de uma imagem. Após esses processamentos foi realizada a classificação pelo método Máquina de Vetores de Suporte (SVM) este método desenvolvido por Vapnik, (1995), baseado em estudos iniciais pelo mesmo autor no ano de 1971, estabelece uma série de princípios que devem ser seguidos na obtenção de classificadores com boa generalização, definidos como a sua capacidade de prever corretamente a classe de novos dados do mesmo domínio em que o aprendizado ocorreu.

O objetivo do presente trabalho é demonstrar a importância de práticas de sensoriamento remoto no reconhecimento de padrões de uso do solo, tendo uma forma de demonstrar mediante a classificação, que visa justamente caracterizar e separar essas classes conforme as suas representações no terreno.

MATERIAL E MÉTODOS

No presente trabalho utilizou-se uma cena do satélite LANDSAT 5 sensor TM, resolução espacial de 30m, e resolução radiométrica de 8 bits e resolução temporal de 16 dias, num recorte da cidade de São Mateus do Sul – Paraná – Brasil, ponto 221 órbita 78, do dia 28/10/2011 adquiridas

gratuitamente no *site* do INPE (Instituto Nacional de Pesquisas Espaciais).

Na cena foram realizados os processamentos iniciais no programa de processamento de imagens digitais (PDI) ENVI 4.7, estes processamentos constituíram na transformação geométrica que tem por objetivo corrigir os erros de geometria ocasionados na aquisição de imagens, e a correção atmosférica, que visa minimizar as interferências ocasionadas na imagem devido à presença e a variação da atmosfera.

Posteriormente foi realizada a classificação supervisionada pelo método algoritmo *Support Vector Machine* (SVM), (VAPNIK,1995), este método de classificação é não paramétrica, ou seja, a forma da função densidade probabilidade associada aos dados não é conhecida.

A classificação foi realizada coletando aproximadamente 200 amostras de treinamento, e classificando a imagem em sete classes de uso do solo sendo elas: corpos d'água, área urbana, vegetação rala, vegetação densa, área de extração de xisto, e duas classes de solo exposto.

Foi realizada a matriz de confusão, este procedimento mostra o número de classificações corretas versus as classificações preditas para cada classe, sobre um conjunto de amostras de treinamento de cada classe.

RESULTADOS E DISCUSSÃO

A **figura 1** representa a cena da cidade de São Mateus do Sul, combinação colorida verdadeira falsa cor R(5)G(4)B(3), representativas dos comprimentos de onda do espectro eletromagnético. A banda 5 representa o infravermelho médio, a banda 4 infravermelho próximo e a banda 3 a do vermelho do espectro eletromagnético.



Figura 1: Imagem da cidade de São Mateus do Sul, combinação colorida verdadeira falsa cor R(5)G(4)B(3).

Na **figura 1**, notam-se várias classes de ocupação do solo, com maior evidência a mina de extração de xisto. Outra classe de ocupação bastante significativa no local é a de solo exposto destinado à agricultura.

A classificação das imagens digitais é o reconhecimento de padrões dos *pixels* e a separação dos mesmos em classes específicas. A **figura 2** representa a classificação supervisionada pelo método *Support Vector Machine*.

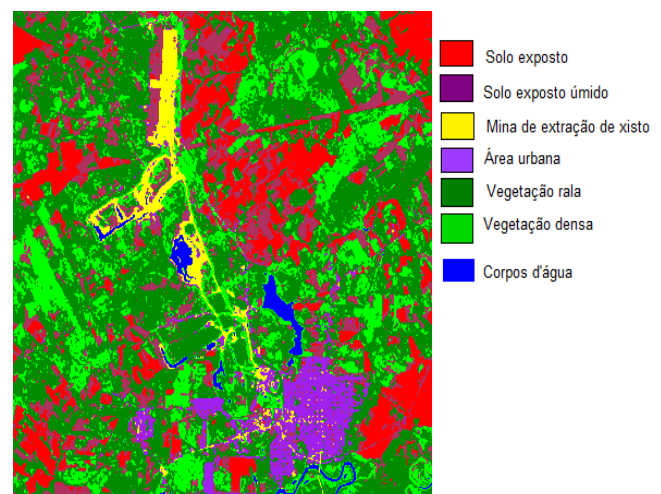


Figura 2: imagem classificada pelo método SVM.

Realizando-se uma análise visual na imagem, nota-se que houve separação satisfatória na classe da mina de extração de xisto, notando –se a definição da área onde há maior índice de exploração, representada pela cor amarela, os corpos hídricos também foi bem separada, essa classe é representada pela cor azul, notando-se então que na imagem existe o lago, provindo de uma antiga área de extração de xisto, existem também na mina áreas que se encontram com grande quantidade de água, e também o Rio Iguazu ao sul da área urbana que no caso da imagem classificada a área urbana apresenta-se com a cor lilás.

Na área urbana ocorreram confusões na classificação, principalmente com a área de extração de xisto, essa confusão se dá, devido aos diferentes objetos presentes na área urbana e também devido à difração na luz, que no caso ocorre de modo linear e difusa, provocando confusão espectral, desta maneira a classe urbana normalmente é extraída separadamente em cenas orbitais.

Outras classes que se obteve separação considerada satisfatória foi nas de solo exposto, e de vegetação rala e vegetação densa, diferenças

essas que devem ser observadas na **figura 1** comparando com a **figura 2**.

A separabilidade entre as classes foi obtida mediante o cálculo da divergência transformada (Swain e Davis, 1978). Este procedimento dita que quanto maior o valor da divergência transformada, melhor será a o resultado da classificação das classes de uso. O resultado da divergência transformada está apresentado na **tabela 1** os resultados variam de 0 a 2,0, onde o 0 é o pior valor, ou seja, menor separabilidade, e 2 melhor separabilidade.

Tabela 1 – Divergência transformada da imagem classificada *Support Vector Machine*

Solo exposto	1,97
Solo exposto úmido	1,97
Área de extração de xisto	1,99
Vegetação densa	1,99
Vegetação rala	1,99
Área urbana	1,90
Corpos d'água	2,00

Mediante os valores apresentados na divergência transformada, nota-se que a melhor classificação ocorreu nos corpos d'água, obtendo valor 2,0 que indica uma melhor acuracidade, porém na classe área urbana obteve-se uma valor de 1,90, valor que demonstra grande confusão espectral, o que nota-se pela imagem representada na **figura 2**, onde há incidências da cor amarela (área de extração de xisto) dentro da área urbana (lilás).

Foi realizada também a matriz de confusão para que assim obtivesse o valor a acurácea da classificação, este procedimento mostra o número de classificações corretas versus as classificações preditas para cada classe, sobre um conjunto de amostras de treinamento de cada classe. A matriz de confusão está disposta na **tabela 2**.

Vendo os resultados da divergência transformada e da matriz de confusão, notou-se que as classes que obtiveram confusão espectral, foram a vegetação rala, e densa, onde confundiu 6,67% dos pixels da vegetação rala, classificando assim como vegetação densa, tendo a vegetação rala acurácea de 93,33. Outra classe que se notou grande confusão espectral é a área urbana, onde se obteve acurácea de 94,59% e confusão espectral de 2,70 com a classe extração de xisto e solo exposto úmido.

Comparando a imagem e os resultados estatísticos notou-se que os resultados estatísticos

condizem com os resultados apresentados na imagem, visualizando-se pequenas confusões espectrais, porém obtendo-se uma classificação satisfatória, quando ao uso do solo da cidade de São Mateus do Sul.

CONCLUSÃO

A utilização do sensoriamento remoto e a classificação de imagens digitais são pertinentes aos estudos de uso e ocupação do solo, pois aperfeiçoa possíveis saídas a campo, pois fornece uma visão prévia do local.

Na presente imagem foi conveniente o uso da classificação, pois nos diversos usos do solo, como a extração de xisto, tipos de vegetação, solos expostos destinados à agricultura.

Tendo em vista que esse algoritmo de classificação é não paramétrico e não linear pode-se incluir a imagem em questão bandas auxiliares como NDVI (*Normalized Difference Vegetation Index*) ou um modelo digital do terreno (MDT) para que dessa maneira as separações das classes de uso sejam mais satisfatórias e em áreas de confusões espectrais possam ser separadas com maior acuracidade.

AGRADECIMENTOS

Agradecemos a Fundação Araucária pela concessão de bolsa de iniciação científica.

REFERÊNCIAS BIBLIOGRÁFICAS

- CENTENO, Jorge Antonio Silva. Sensoriamento remoto e processamento de imagens digitais. Editora Curso de Pós Graduação em Ciências Geodésicas. Curitiba, 2004 Cap. 5.
- FOODY, G. M.; MATHUR, A. A Relative Evaluation of Multiclass Image Classification by Support Vector Machines. IEEE Transactions on Geoscience and Remote Sensing, v.42, n.6, 2004.
- INSTITUTO DO MEIO AMBIENTE – IA, 2005, CLC2000 Portugal Technical Report. Disponível em: < www.igeo.pt/instituto/cegig/gdr/projectos/clc2000/ >. Acesso em: 22 de maio de 2013.
- LIU, Willian Tse Horng. Aplicações de sensoriamento remoto / Willian Tse Horng Liu – Campo Grande: Ed. UNIDERP, 2006. Cap 14 pág. 733
- SWAIN, P. e DAVIS, S. Remote Sensing: The quantitative Approach, Mc.Graaw-Hill, New York, 1978.



Tabela 2 – Matriz de confusão da classificação supervisionada pelo método Support Vector Machine

Classes	Veg. densa	Veg. rala	Área urbana	Extração de xisto	Corpos hídricos	Solo exposto	Solo exposto úmido
Não classificada	0	0	0	0	0	0	0
Vegetação densa	100	0	0	0	0	0	0
Vegetação Rala	6,67	93,33	0	0	0	0	0
Área Urbana	0	0	94,59	2,70	0	0	2,70
Área de extração de xisto	0	0	2,70	94,59	0	0	2,71
Corpos hídricos	0	0	0	0	100	0	0
Solo exposto	0	0	0	0	0	100	0
Solo exposto úmido	0	0	2,70	0	0	2,70	94,60