

A ANÁLISE BAYESIANA NA CLASSIFICAÇÃO SUPERVISIONADA DE IMAGENS: APLICAÇÃO NA DETERMINAÇÃO DE CLASSES DE USO DO SOLO (MOGI DAS CRUZES – SP)

ROSÂNGELA LEAL SANTOS^{1,2}
ARLETE TIEKO OHATA¹
JOSÉ ALBERTO QUINTANILHA¹

¹PTR/EPUSP - Universidade de São Paulo
Caixa Postal: 61 548 CEP 05508-900 – São Paulo - SP, Brasil
arlete.ohata@poli.usp.br, jaquinta@usp.br

²DTEC – Universidade Estadual de Feira de Santana
Departamento de Tecnologia - BR 116 – Km 3 - 44020-340 – Feira de Santana- SP, Brasil
rosaleal@usp.br

Abstract. This article describes an application using bayesian classification algorithms to determine land use classes. In this research was a used BAYCLAS, BELCLAS and Dempster-Shafer modified algorithm to process LANDSAT 7TM image satellite, implemented algorithms in IDRISI 32 software. We made a analysis between them observing with one has better performance to recognize different patterns and to discriminate most clearly as possible the land use classes. With these elements analyzed we will define a methodology to apply in a role area.

Keywords: Bayesian classification, remote sensing, image processing, land use.

1. Introdução

Técnicas de classificação de objetos consistem, basicamente, em separá-los colocando-os em grupos previamente definidos. Podemos dizer que a classificação é um processo de extração de informação em imagens com o objetivo de reconhecer padrões e objetos homogêneos. No processamento de imagem, atributos relacionados a intensidade dos pixels são naturalmente escolhidos para representarem objetos ou regiões das imagens num classificador (Lilesand & Kiefer, 1987).

O tipo de classificação pode ser dividida em duas categorias: supervisionada, que precisa de dados de treinamento para o classificador; e não -supervisionada que funciona sem a etapa de treinamento, utilizando outros critérios de semelhança dos objetos, como por exemplo, a distância entre os mesmos. No presente trabalho iremos tratar de um método de classificação supervisionada em particular, o método bayesiano, e como ele utiliza o conhecimento prévio do analista, sobre a área que se deseja analisar.

1.1 Classificação de imagens e reconhecimento de padrões

Classificação é o processo de interpretar imagens de sensoriamento remoto de forma automatizada. Como consequência, a classificação é, talvez, o mais importante aspecto do processamento de imagens para o GIS (Jensen, 1996). Uma característica importante na classificação de imagens é que nenhuma abordagem individual é ótima, de modo que métodos e abordagens múltiplas devem ser utilizados combinado -se várias modalidades de sensores, pré-processamentos e métodos de classificação (Johnson, & Wichern, 1998)

1.2 Classificação supervisionada

Pode-se dizer, resumidamente, que um sistema de classificação envolve três etapas: aquisição de dados e pré-processamento, representação dos dados e tomada de decisão. O classificador toma decisões baseando-se no aprendizado realizado a partir de um conjunto de treinamento, o qual contém exemplos de padrões de todas as classes existentes no sistema (Ripley, 2000). Um classificador é uma função que possui como entrada padrões desconhecidos e, como saída, rótulos que identificam a que classe tais padrões provavelmente pertencem. O que difere um classificador de outro é a forma como esse cria as fronteiras de decisão a partir dos exemplos de treinamento.

1.2.1 O método bayesiano

Alguns algoritmos utilizados como classificadores de imagem podem fornecer probabilidades a posteriori a partir da entrada de um padrão \underline{X} , determinando que este pode pertencer a uma dada classe de dados \mathbf{w}_i , onde $P(\underline{X} \in \mathbf{w}_i | \underline{X})$, $I=1..M$. (Giacinto et al., 2000). Dentre eles, o que mais se destaca é o algoritmo baseado nos métodos bayesianos. O método Bayesiano utiliza a probabilidade para quantificar as incertezas em inferências. O resultado da coleta de suas amostras é a elaboração de uma distribuição de probabilidades expressando nossa opinião a respeito de como provavelmente diferentes são as predições. Este tipo de classificador permite incorporar um dado muito importante na classificação supervisionada: o conhecimento prévio do analista da área de estudo (Gonzales & Hoods, 1992).

2. Objetivo

O objetivo deste trabalho é utilizar um classificador bayesiano e outras técnicas correlatas de classificação relativa para definir as classes de uso do solo da região de Mogi das Cruzes (SP).

3. Área de Estudo

A área de estudo corresponde a um projeto em coordenadas planas, no sistema UTM (WGS84), abrangendo um retângulo envolvente de coordenadas (X1) 366148,953 E e 7401361,500 N e (X2) 381118,953 E e 7388401,500 N. Compreende a região de Mogi das Cruzes, município que faz parte da Região Metropolitana de São Paulo, no Estado de São Paulo. Faz parte do cinturão verde que abastece de hortifrutigranjeiros o município de São Paulo, pertencente ao ecossistema da Mata Atlântica, onde verifica-se uma boa preservação da vegetação original. Entretanto, esta área apresenta áreas de exploração do subsolo, através da mineração de areia, conflitando abertamente com as áreas de horticultura além de fazer limite com a área urbana. Estes tão diferentes usos do solo e do subsolo tem gerado conflitos de interesses e de exploração nesta região, agravada pela atual crescente expansão da atividade mineradora em detrimento das áreas agrícolas, e a expansão urbana dos últimos anos.

4. Metodologia

4.1 Material

Imagem Landsat TM 7 219/76 (22/04/2002); Bandas 1,2,3,4,5,7.
Software ENVI 3.0

Software IDRISI 32.

4.2 Procedimentos

- Recorte da área (redução do tamanho da imagem com áreas significativas)
- Conversão de áreas classificadas pelo algoritmo Isodata em vetores para servir de áreas de treinamento
- Determinação das classes (4 classes de vegetação, 2 de água, 1 de solo, 1 área urbana, 3 áreas de transição)
- Treinamento
- Especificação das probabilidades

4.2.1 Algoritmos utilizados

- **Máxima verossimilhança** - o qual se baseia no teorema de Bayes, como probabilidade condicional, o qual expressa a relação de evidência, o conhecimento prévio e a probabilidade de que determinada hipótese é verdadeira;
- **Bayclass** - um avaliador da probabilidade de um dado pixel pertencer a uma dada classe, gerando a matriz de incerteza da classificação e as matrizes de probabilidade para cada classe;
- **Belclass** - variante da teoria bayesiana, conhecida como teoria Dempster-Shafer, onde determina a confiança estimada de cada pixel pertencer a cada classe
- **Dempster-Shafer modificado** - declara o grau de incerteza da classificação, através da imagem da plausibilidade, ou seja, o grau com que uma dada hipótese de classificação não pode ser descartada.

• 6. Resultados e Discussões

A comparação entre as diferentes imagens geradas permite a determinação de uma classificação mais segura, apoiada numa decisão em que estão embutidos os possíveis erros decorrentes da etapa de treinamento. Entretanto, no presente trabalho, esta etapa foi bastante otimizada pela utilização dos vetores obtidos pela classificação não supervisionada.

7. Conclusões

Na maioria dos casos os analistas não fazem considerações sobre a probabilidade relativa de um pixel pertencer ou não a uma dada classe, assumindo que cada classe possui igual probabilidade sobre este pixel. Através da probabilidade a priori, se permite incorporar o conhecimento prévio na forma de imagens de probabilidade, permitindo que esta varie de um local para outro. Acredita-se que a não utilização das probabilidades a priori se deve a uma certa inabilidade do usuário comum de poder especificá-la corretamente. A utilização das matrizes de incerteza, da confiança e da plausibilidade, nos permite uma classificação mais eficiente, com uma matriz de erro com grãção de padrões com pequena variação intra-classe e grande variação inter-classe e uma redução do número de pixels não classificados.

Referências

GIORGIO, G; ROLI, F.; BRUZZONE, L. Combination of neural and statistical algorithms for supervised classification of remote-sensing images. *Pattern Recognition Letters* 21, 2000. PP. 385-397

GONZALES, R.C. & HOODS, R.E. *Digital image processing*. Addison-Wesley, reading, Massachusetts, 1992.

JENSEN, J.R. *Introductory digital image processing: a remote sensing perspective*. Prentice Hall, Upper Saddle River, NJ, 1996.

JOHNSON, R. A.; WICHERN, D. W.; *Applied Multivariate Statistical Analysis*, Prentice-Hall, 4th edition, 1998.

LILESAND, T.M. & KIEFER, R.W. *Remote sensing and image interpretation*. 2ª New York : Ed. John Wiley & Sons, 1987.

RIPLEY, B D . *Pattern recognition and neural networks*. Cambridge University, Press, Cambridge, 2000