

INTEGRAÇÃO DE IMAGEM DE SATÉLITE E MAPA TOPOGRÁFICO POR MEIO DA TEORIA DA EVIDÊNCIA

Selma Regina Aranha Ribeiro¹
Alzir Felipe Buffara Antunes²
Jorge da Silva Centeno³

- 1 Universidade Federal do Paraná – selma@geoc.ufpr.br
2 Universidade Federal do Paraná – felipe@geoc.ufpr.br
3 Universidade Federal do Paraná – centeno@geoc.ufpr.br

RESUMO

O sensoriamento remoto tem sido hoje a principal fonte de dados para o mapeamento temático em grandes e médias escalas. A maioria dos temas são oriundos da classificação digital derivada em geral pelo método clássico da máxima verossimilhança.

Haja vista, a dificuldade de se classificar algumas feições do terreno devido a respostas espectrais serem próximas entre si, resulta em uma diminuição na confiabilidade da classificação automatizada. A fim de mitigar este problema, pode-se integrar outras fontes de dados tal como mapa topográfico, visando o aumento de variáveis para a tomada de decisão dentro dos algoritmos estatísticos.

Dentro do processo de integração de dados de imagem de satélite e mapa topográfico, a teoria da evidência aparece como uma ferramenta útil no incremento da acurácia da classificação digital. Esta teoria agrega as feições contidas em ambas as fontes, de forma complementar ou através de inferências estatísticas.

Neste trabalho foram integradas imagem de satélite pré-classificada com mapa topográfico afim de se obter um mapa temático que considere as informações contidas em ambas as fontes, com o intuito de melhorar a classificação digital através da integração de dados de diferentes fontes: Mapa Topográfico e Imagem Multiespectral Landsat TM7. E como objetivo específico, avaliar a utilização da teoria da evidência e comparar a classificação pelo método da máxima verossimilhança sobre imagem orbital, com a classificação oriunda da teoria da evidência.

Palavras-chave: Sensoriamento Remoto, Integração de dados, Teoria da Evidência

ABSTRACT

The remote sensing has been today the main source of data for the thematic mapping in big and averages scales. Most of the themes is originating from of the digital classification derived in general by the classic method of the maximum likelihood.

Have seen, the difficulty of being classified some features of the land due to ghastly answers be to each other close, it results in a decrease in the reliability of the automated classification. In order to mitigate this problem, it can be integrated other sources of data just as topographical map, seeking the increase of variables inside for the taking of decision of the statistical algorithms.

Inside of the process of integration of data of satellite image and topographical map, the theory of the evidence appears as an useful tool in the increment of the accuracy of the digital classification. This theory joins the features contained in both sources, in a complementary way or through statistical inferences.

In this work satellite image pré-classified with kindred topographical map of obtaining a thematic map were integrated that it considers the information contained in both sources, with the goal of improving the digital classification through the integration of data of different sources: Topographical map and Imagem Multiespectral Landsat TM7. And as specific objective, to evaluate the use of the theory of the evidence and to compare the classification for the method of the maximum likelihood on image orbital, with the classification originating from of the theory of the evidence.

Keywords : Remote Sensing, Integration of data, Theory of the Evidence

1. INTRODUÇÃO

O sensoriamento remoto tem sido hoje a principal fonte de dados para o mapeamento temático em grandes e médias escalas. A maioria dos temas são oriundos da classificação digital derivada em geral pelo método clássico da máxima verossimilhança. A dificuldade de se classificar algumas feições do terreno devido a respostas espectrais serem próximas entre si, resulta em uma diminuição na confiabilidade da classificação automatizada. A fim de mitigar este problema, pode-se integrar outras fontes de dados tal como mapa topográfico, visando o aumento de variáveis para a tomada de decisão dentro dos algoritmos estatísticos.

Dentro do processo de integração de dados de imagem de satélite e mapa topográfico, a teoria da evidência aparece como uma ferramenta útil no incremento da acurácia da classificação digital. Esta teoria agrega as feições contidas em ambas as fontes, de forma complementar ou através de inferências estatísticas.

2. OBJETIVO

2.1. GERAL

Melhorar a classificação digital através da integração de dados de diferentes fontes: Mapa Topográfico e Imagem Multiespectral Landsat TM7.

2.1.1. Específico

- a)- Avaliar a utilização da teoria da evidência ;
- b)- Comparar a classificação pelo método da máxima verossimilhança sobre imagem orbital, com a classificação oriunda da teoria da evidência.

3. REVISÃO BIBLIOGRÁFICA

Segundo RICHARDS (1999), a integração de dados significa a possibilidade de junção de diferentes componentes ou fontes de dados afim de se obter um melhor resultado.

Integração significa que diferentes componentes serão agrupados e os mesmos cooperaram para a obtenção de um melhor resultado, que não seria obtido se estes dados fossem utilizados separadamente (BALTSAVIAS, 1999).

A integração de dados de diferentes origens, associada a informações oriundas de imagem capturadas por técnicas de Sensoriamento Remoto (imagem digital/ orbital), tem se mostrado uma ferramenta de grande utilidade em mapeamentos temáticos. Cientistas da área das Ciências da Terra, tais como (WARNER e SHANK, 1997); (HARRIS e MURRAY, 1990); (RIBEIRO, 2002) e (ANDRADE, 2003), podem ser citados como referência em estudos de integração de dados.

A teoria da evidência, ou seja, teoria matemática na qual as fontes de dados são tratadas separadamente e suas contribuições combinadas de forma a permitir uma inferência conjunta relativa, isto é, determinar o correto atributo para um determinado pixel, isto requer que as variáveis originais sejam numéricas (RICHARDS, 1999).

Esta teoria está baseada na probabilidade de determinado pixel pertencer a uma classe, levando-se em conta a incerteza ou ignorância, sendo que o somatório das probabilidades e incertezas devem ser igual a unidade. A incerteza faz com que as probabilidades (função m1) em favor de determinada classe tenha um decréscimo devido a insegurança da assertiva (função m2).

$$m1 (< p1, p2, p3, >), m2(< p1, p2, p3, q >)$$

onde: θ significa a incerteza e p as probabilidades de se pertencer a determinada classe.

Considerando as duas fontes de dados utilizadas, o mapa topográfico e a imagem de satélite, a combinação destas duas fontes de dados é dada pela seguinte soma ortogonal (MOON, 1990; GONG, 1996):

$$m1 \wedge m2(D) = S (Ai \nabla Bj = D) m1(A).m2(B) / k$$

$$k=[1- S (Ai \nabla Bj = 0) m1(A).m2(B)]$$

onde: m1 e m2 são as evidências para as duas fontes de dados;

k um fator de escala, que por ser constante pode ser desconsiderado.

Desta forma a expressão acima pode ser reescrita da seguinte forma:

$$m(w) = m1(w) \times m2(w) + m1(w) \times q2(w) + m2(w) \times q1(w)$$

onde:

$m(w)$ = a evidência para a classe w ;

$m1(w)$ = a probabilidade do pixel da fonte 1 pertencer a classe w ;

$m2(w)$ = a probabilidade do pixel da fonte 2 pertencer a classe w ;

$\theta1(w)$ = a ignorância para a fonte 1;

$\theta2(w)$ = a ignorância para a fonte 2.

4. METODOLOGIA

4.1. FONTE DE DADOS

- Mapa Topográfico de GUARATUBA, 1:50.000 Folha SG.22-X-D-V MI-2858-4, IBGE(1992);
- Imagem Landsat TM7, Bandas 2,4,7, Jan/2000(Figura 1).

Figura 1- Área de Estudo- Guaratuba- PR, mapa topográfico e imagem Landsat TM7



4.2. DESCRIÇÃO DO MÉTODO

4.2.1. Obtenção de Evidências a partir da imagem (Landsat TM7)

Sobre a imagem Landsat TM7 foram selecionadas amostras representativas das seguintes classe:

- a) Baixios – áreas rasas, em geral composta de areia e lodo;
- b) Reflorestamento- áreas reflorestada com pinus;
- c) Mangues- vegetação típica de ambientes lacustres;
- d) Vegetação- generalização de todas as áreas verdes;
- e) Solo Exposto- inclui todos os antropismos;
- f) Mar- inclui a área de baía, mar aberto e rio.

Procurou-se que as mesmas classes observadas na imagem fossem também identificadas no mapa topográfico.

Então, a partir de amostragem foram obtidas as matrizes variância- covariância para as classes envolvidas. Utilizando o algoritmo de MAXVER, foram determinadas as funções densidade de probabilidade de acordo com a expressão abaixo:

$$P(X/W_i) = \frac{1}{N} \frac{1}{\sqrt{|S_i|}} \exp\left\{-\frac{1}{2}(X - \bar{m}_i)^t S_i^{-1} (X - \bar{m}_i)\right\}$$

onde:

$P(X/W_i)$ = probabilidade de um pixel X pertencer a uma classe W;

S_i = matriz variância-covariância;

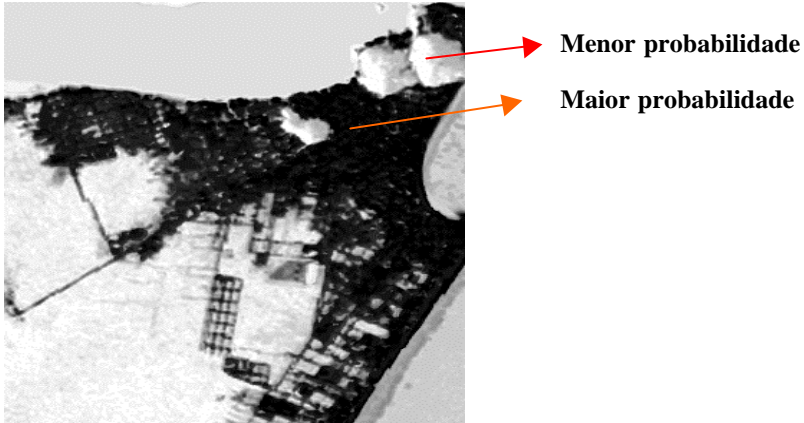
X= vetor, nível de cinza;

W_i = classe;

N= número de classes.

Cada uma das matrizes variância-covariância representativa de cada classe amostrada, gerou uma imagem probabilística, onde os tons mais escuros mostram a maior probabilidade do pixel pertencer a determinada classe, e a medida que os tons se tornam mais claros a probabilidade diminui. A Figura 2 mostra a imagem densidade de probabilidade para a classe solo exposto (que inclui as áreas urbanas, vias e descampado).

Figura 2-Imagem probabilística de solo exposto



Para a consecução da teoria da evidência, faz-se mister o conhecimento das incertezas (θ) das classes. Para a determinação de (θ) para os dados da imagem, utilizou-se como referência a qualidade da classificação individual das classes, através do coeficiente **Kappa** de acurácia, que considera todos os valores da matriz de erros oriunda da verdade de campo (imagem original) e a imagem classificada por MAXVER.

A determinação de θ é dada pela seguinte expressão:

$$p(w_i) = p(w_i) * K; \quad q = 1 - p(w_i)$$

onde:

K -coeficiente kappa;

$p(w_i)$ – probabilidade do pixel pertencer a determinada classe;

θ - incerteza.

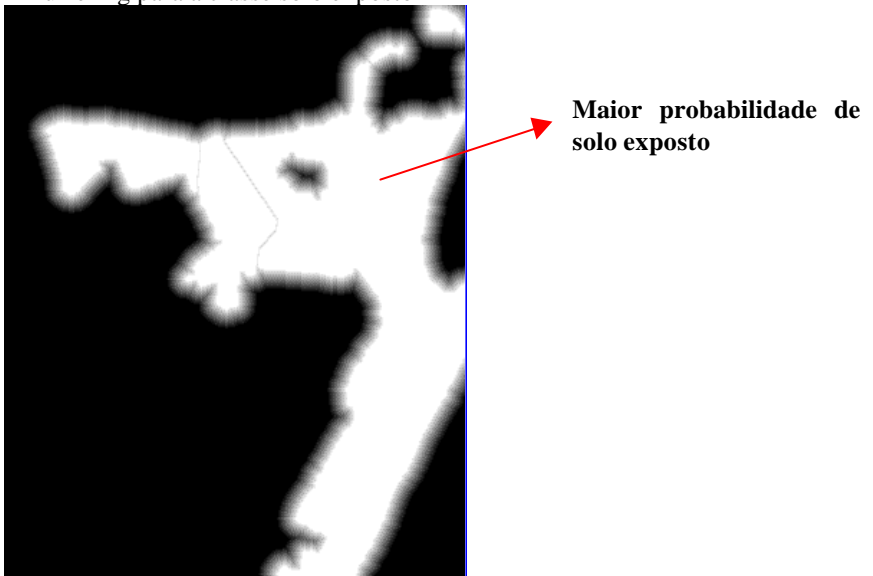
E os valores de K para as classes:

Kbaixio=	0,940
Kreflorestamento	0,770
Kmangue=	0,940
KVegetação=	0,877
KSolo Exposto=	0,882
Kmar=	0,940

4.2.2. Evidências oriundas do mapa

A imagem do mapa topográfico rasterizado, foi segmentada de acordo com as diferentes classes supra citadas. Cada segmento da imagem (classes), foi referenciado a uma determinada imagem probabilística. Afim de que, a probabilidade de cada uma das classes pudessem também variar de 0 a 1, foi efetuado um procedimento de distanciamento, ou seja, a probabilidade de determinada feição existir fora dos seus limites desenhados no mapa. Estas distâncias levam em conta a tendência de uma classe se expandir ou se retrain, além dos seus limites. De acordo com o ilustrado na Figura 3.

Figura 3 – Buffering para a classe solo exposto



Estas distâncias ou “Buffers” foram determinadas de acordo com critérios específicos, de reconhecimento e interpretação espacial, tal como se segue abaixo:

Distâncias (em pixel)

Classes:	Dentro	Fora
BAIXIO	10	10
REFLO.	15	0
S. EXP	0	15
VEG	15	0

MANGUE	2	2
MAR	0	0

Com relação a determinação das incertezas para o mapa, chegou-se através de observações a seguinte expressão:

q = (1 – Máx(Pwi))/N;

onde:
N - número de classes;
Máx(Pwi) - máxima probabilidade

4.2.3. Integração das Evidências

A matriz abaixo, mostra a distribuição de ocorrência de determinada classe para as duas fontes distintas de dados. Onde nas linhas e colunas limítrofes tem-se a ignorância ou as incertezas (Quadro1).

Quadro1 – Matriz de evidências

		MAPA						
		W1	W2	W3	W4	W5	W6	q
I M A G E M	W1							
	W2							
	W3							
	W4							
	W5							
	W6							
	q							

Uma vez composta a matriz das evidências, determina-se m(w)- a decisão em favor de determinada classe. A evidência é determinada de acordo com a expressão:

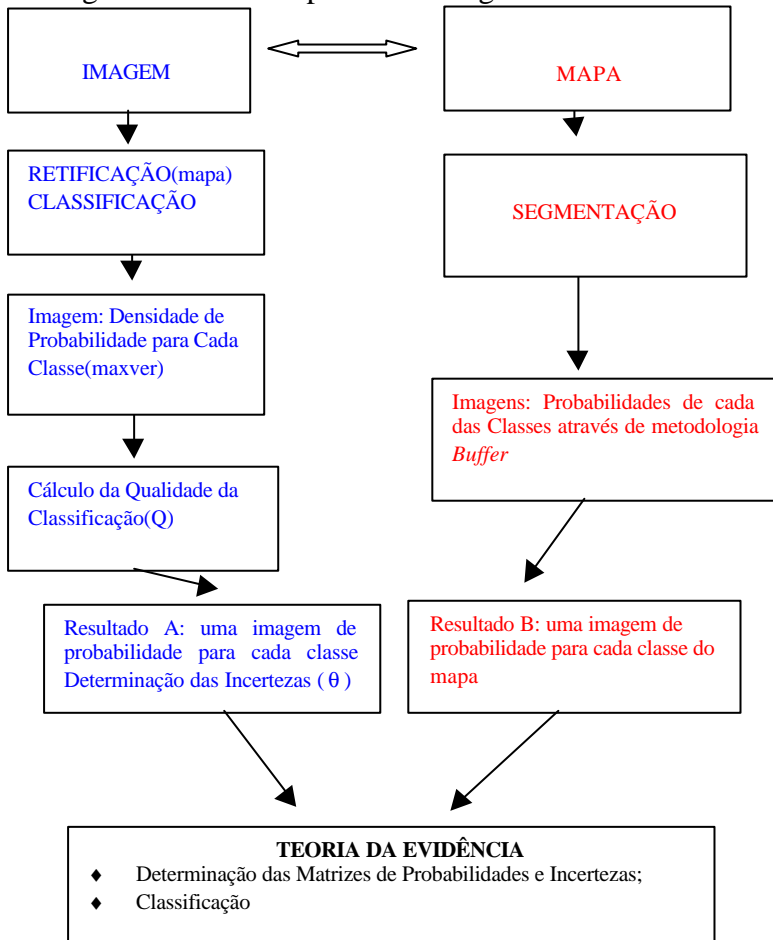
m(w)= m₁(w) x m₂(w) + m₁(w) x q₂(w) + m₂(w) x q₁ (w)

onde :
m(w)= a evidência para a classe w;

$m_1(w)$ = a probabilidade do pixel da fonte 1 pertencer a classe w ;
 $m_2(w)$ a probabilidade do pixel da fonte 2 pertencer a classe w ;
 $\theta_1(w)$ = a ignorância para a fonte 1;
 $\theta_2(w)$ = a ignorância para a fonte 2.

O fluxograma abaixo resume todos os procedimentos efetuados até a determinação das evidências e posteriormente a imagem classificada considerando as duas fontes de dados.

Fluxograma com as etapas metodológicas



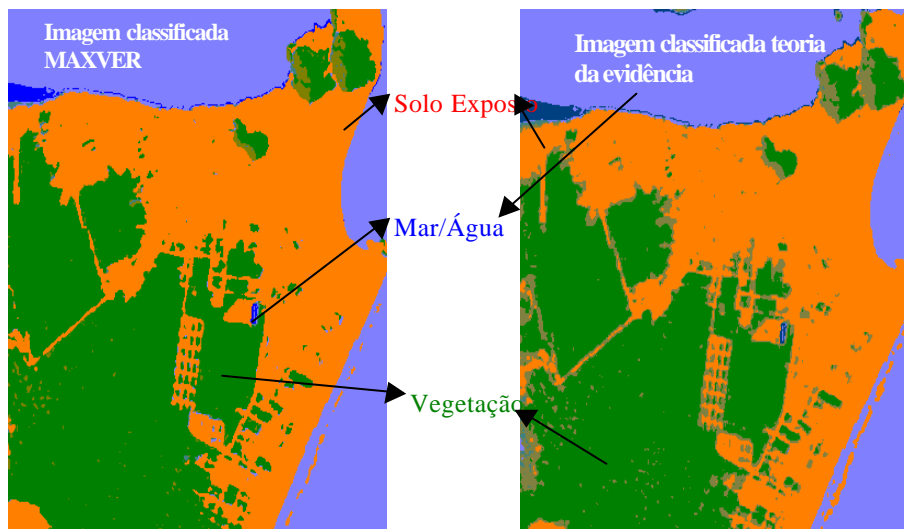
5. RESULTADOS

A imagem oriunda do mapa topográfico está de certa forma desatualizada, haja vista que o mesmo foi editado em 1992 (IBGE, 1992) e a imagem Landsat é de janeiro de 2000. Observa-se na Figura 1 que a ocupação urbana na imagem é bem maior que no mapa topográfico, considerando a defasagem temporal.

A Figura 4 (imagem classificada por meio da teoria da evidência) demonstra que a classificação pela teoria da evidência, manteve as informações atualizadas da imagem classificada MAXVER, ou seja, existia maior probabilidade de ocorrer solo exposto a vegetação.

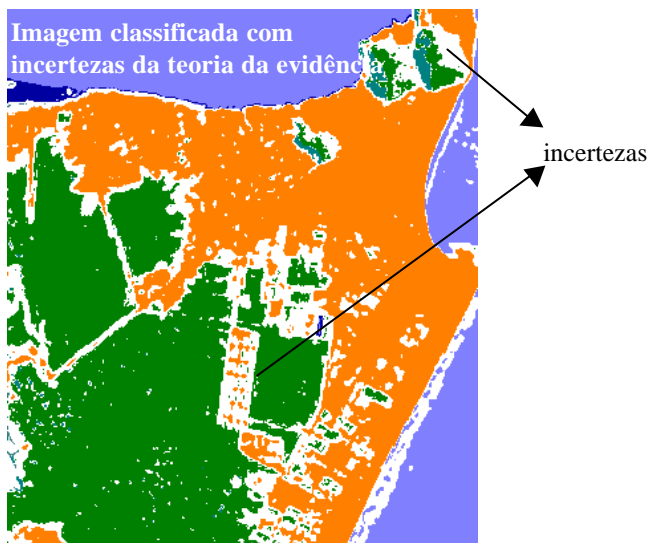
Analogamente, observa-se na imagem MAXVER a ocorrência da classe água no local onde deveria estar presente a classe solo exposto. Este equívoco então, foi corrigido através dos dados provenientes do mapa topográfico, que foram preponderantes, e a classe água foi substituída por solo exposto e vegetação.

Figura 4 – Imagens classificadas



A Figura 5 abaixo, mostra a imagem classificada e as incertezas como uma classe complementar. Apresenta as regiões (em branco) onde existem diferenças entre as duas fontes de dados. Posteriormente foram atribuídas a essas regiões a classes de maior probabilidade de ocorrência, considerando-se as incertezas do pixel pertencer a determinada classe.

Figura 5- Imagem classificada com as incertezas ou ignorância



6. CONCLUSÕES E RECOMENDAÇÕES

Através comparação da imagem classificada por MAXVER e o mapa topográfico, chegou-se as seguintes conclusões:

- Houve uma melhora significativa utilizando-se as duas fontes de dados no processo de classificação digital;
- A teoria da evidência mostrou-se uma ferramenta útil na classificação digital, pois considera as informações contidas em ambas as fontes de dados, considerando as probabilidades e incertezas.

Para uma avaliação mais criteriosa dos resultados obtidos, seria recomendável que fossem feitas verificações em campo para se determinar o índice de acurácia em ambas as imagens, e compará-las em termos de qualidade.

Em adição seria interessante que sempre fossem utilizadas fontes compatíveis em termos temporais, com a finalidade de se poder extrair sempre a melhor informação em ambas as fontes.

7. REFERÊNCIAS

- BALTSAVIAS, E. & HAHN, M. 1999. **Integration of Image Analysis and GIS**. International Archives of Photogrammetric and Remote Sensing, v. 32.
- ANDRADE, A. F. 2003. **Integração de variáveis espectrais e forma na classificação de imagens de alta resolução utilizando redes neurais artificiais**. Dissertação de mestrado- UFPR. 120p.
- GONG, P. 1996. **Integrated analysis of spatial data from multiple sources: using evidential reasoning and artificial neural network techniques for geological mapping**. PE&RS v.62n.5p.513-523.
- HARRIS, R. J.; MURRAY, R. 1990. **IHS transform for the integration of RADAR imagery with other remotely sensing data**. Photogrammetric Engineering & Remote Sensing. V. 61, n. 12, p. 1631- 1641.
- IBGE.1992. **Mapa topográfico de Guaratuba folha SG.22-X-D-V MI-2858-4**. Rio de Janeiro
- MOON, M.W. 1990. **Integration of geophysical and geological data using evidential belief function**. IEEE Transactions on geosciences and remote sensing. V.28, n.4.
- RIBEIRO, S.R. A. CENTENO, J.A . S. 2002. **Integração de imagens de satélite e dados de complementares para a delimitação de unidades de paisagem usando uma abordagem baseada em regiões**. Boletim de Ciências Geodésicas- Curso de Pós-Graduação em Ciências Geodésicas –UFPR. P. 47-57.
- RICHARDS, John & JIA, Xiuping. 1999. **Remote Sensing Image Analysis, an Introduction**. Spring-Verlag. Berlim.
- WARNER, T. A .; SHANK, M. 1997. **An evaluation of the potential fuzzy classification of multispectral data using artificial neural networks**. Photogrammetric Engineering & Remote Sensing. v.63, n.11, p.1285-1294.