

EXTRAÇÃO DE ATRIBUTOS ESPACIAIS EM IMAGENS DIGITAIS MULTIESPECTRAIS

L. V. Dutra e N. D. A. Mascarenhas

Instituto de Pesquisas Espaciais

Conselho Nacional de Desenvolvimento Científico e Tecnológico

Caixa Postal 515 - 12200 - São José dos Campos, SP, Brasil

RESUMO

Neste trabalho são desenvolvidos métodos para extrair atributos espaciais de imagem, através de filtragem linear e não-linear. Portanto, com um volume maior de informação, é obtida uma precisão maior de classificação. Métodos de seleção de atributos são empregados, porque não é possível usar todos os atributos gerados, devido a restrições físicas, custo computacional e disponibilidade de amostras de treinamento. Os experimentos foram realizados em uma parte de uma cena do satélite LANDSAT-C sobre a região de cana de açúcar perto de Ribeirão Preto. Os quatro canais originais foram filtrados por um filtro do tipo média sobre uma região 5 x 5 (filtro passa-baixa), derivando assim outros quatro canais. Mais quatro canais foram obtidos, utilizando-se operadores derivativos (passa-altas), sendo posteriormente amaciados e seus histogramas equalizados com um modelo gaussiano. Desses doze canais obtidos, quatro foram escolhidos segundo o critério de máxima distância JM média, e outros por máxima distância JM mínima. Foi observado que entre os canais escolhidos houve sempre, no mínimo, um canal que carregava informação espacial e que a precisão da classificação aumentou para as mesmas classes, com relação ao caso onde só os canais originais são usados.

ABSTRACT

In this work, methods to extract spatial features from images through linear and non-linear filtering are developed. Therefore, with a greater volume of information, one obtains a greater precision in the classification. Feature selection methods are also employed, since it is not possible to use all the generated features, because of physical restrictions, computational cost and availability of training samples. The experiments were performed in part of a LANDSAT-C satellite scene over the sugar cane growing region of Ribeirão Preto. The four original channels were filtered by a filter of the average type over a 5 x 5 region (low pass filter), deriving therefore four other channels. Four more channels were obtained through the derivative operators (high pass) being afterwards smoothed and its histogram equalized with a gaussian shape. From the twelve channels obtained in this way, four were chosen according to the criterion of maximum average JM distance and the other four according to the maximum minimum JM distance. It was observed that among the chosen channels there was always at least one carrying spatial information and that the precision of classification increased for the same classes, with respect to the case where only the original channels were used.

1. INTRODUÇÃO

O uso das imagens obtidas por satélites de recursos terrestre veio facilitar consideravelmente a tarefa de analisar, com rapidez e economia, grandes áreas agrícolas, possibilitando o levantamento de dados, tais como os de áreas cultivadas por espécie e as projeções de safras.

Os métodos de partição das imagens multi-espectrais em regiões, associada a cada região uma classe ou cultura, foram desenvolvidos com base na teoria estabelecida para a resolução dos problemas gerais de classificação de padrões.

Um sistema básico de classificação de padrões pode ser decomposto em 3 partes: o transdutor, o extrator de atributos e o classificador (Figura 1).

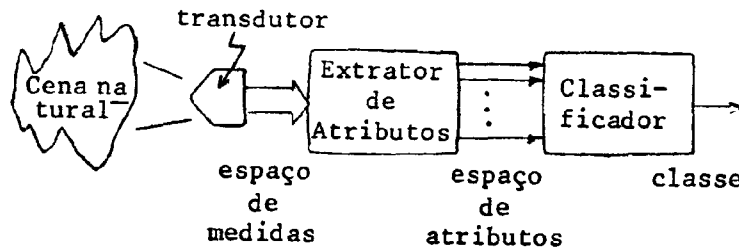


Fig. 1 - Sistema de classificação de padrões

O *transdutor* recebe a radiação que não chega do objetivo natural, em todos comprimentos de onda, mas capta apenas a informação contida em algumas bandas estreitas no espectro de frequência. Normalmente, essas bandas são tão localizadas desde o visível até o infravermelho. Essa informação, chamada "espectral", é digitalizada tanto espacialmente, quanto em intensidade (o nível de cinza, ou seja, o valor medido para a radiação em determinada banda só pode assumir valores inteiros).

Obtém-se, assim, para representar uma determinada cena natural, uma matriz de pontos, onde cada ponto é na realidade um vetor, cujos componentes são os valores das "cores" espectrais, que o transdutor adquiriu para a pequena área correspondente à sua resolução.

Chama-se essa matriz *imagem multiespectral*. Chama-se *canal* a cada um dos componentes dos vetores. Cada ponto da imagem é chamado "pixel".

Conforme os objetos que se quer, o conjunto de medidas espectrais adquiridas pode não ser o mais conveniente para distinguir as classes de interesse. O *extrator de atributos* se encarrega de transformar os dados de medida em outra forma mais adequada, de acordo com as classes envolvidas.

O *classificador* associa o objeto a uma classe, de acordo com os atributos apresentados. Ele toma, portanto, uma decisão baseada nas evidências (atributos) apresentadas.

2. A IMPORTÂNCIA DA ADEQUADA EXTRAÇÃO DE ATRIBUTOS

Um dos principais problemas no projeto de sistemas de classificação é a escolha, ou, dependendo do caso, a extração de atributos que sejam os melhores para distinguir as classes de interesse.

O número de padrões disponíveis para treinamento e/ou o custo computacional dos algoritmos podem requerer que se reduza a dimensionalidade dos vetores de atributos para uso no classificador. Para isso, podem-se utilizar

técnicas conhecidas como *seleção de atributos*, seja por simples escolha de algumas das medidas disponíveis, seja por processos mais complexos como a chamada *rotação espectral*, por meio da qual faz-se uma transformação sobre o espaço de medidas, de maneira que o novo espaço de atributos gerado possa dispor de qualidades desejáveis, tais como, os eixos não serem correlacionados, e a informação se concentrar em algumas determinadas componentes do novo espaço. Uma transformação que apresenta essas qualidades é a chamada Karhunen-Loève.

Os sistemas de classificação mais comuns utilizam os atributos gerados segundo os processos descritos acima. Desta maneira, despreza-se a informação, chamada *espacial*, de cada uma das imagens componentes da imagem multiespectral, já que cada "pixel", representado pelo seu vetor, é classificado independentemente dos seus vizinhos. A informação espacial é determinada por relações existentes entre "pixels" vizinhos.

Uma das formas de extração de atributos espaciais é a da utilização de filtros espaciais, uma vez que o conteúdo de frequência de uma imagem é uma forma de expressão do relacionamento espacial entre "pixels", e os filtros podem ressaltar um ou outro aspecto desse relacionamento.

Utilizando-se filtros, no entanto, é possível a produção de um grande número de atributos do tipo espacial. Neste caso, será necessário selecionar atributos para que se possa diminuir os custos computacionais, e garantir uma boa estimativa dos parâmetros necessários à classificação.

3. EXTRAÇÃO DE ATRIBUTOS ESPACIAIS POR FILTROS

Filtros são operadores, lineares ou não, que podem ressaltar o conteúdo de certas faixas de frequência de uma imagem, sendo utilizados também para detecção de bordas, linhas ou manchas.

No caso de filtragem linear e em aplicação para a extração de atributos, a imagem resultado é obtida da original pela convolução desta com diversos filtros *curtos* (poucos pontos), chamados *máscaras*, ou seja:

$$[r(i,j)] = [h(i,j)] * [f(i,j)], \quad (1)$$

onde:

- $[r(i,j)]$ - imagem resultado
- $[h(i,j)]$ - máscara
- $[f(i,j)]$ - imagem original
- * - denota convolução

A Figura 2 apresenta algumas máscaras possíveis para a realização de filtragem "passa baixa".

As Figuras 3 e 4 apresentam máscaras que são aproximações dos operadores derivativos laplaciano e bi-laplaciano. Tais filtros reforçam o conteúdo de alta frequência da imagem.

Os filtros não-lineares são, normalmente, realizados por processos em 2 estágios: primeiro, uma convolução linear, e, em seguida, uma operação pontual como, por exemplo, módulo ou função de transferência não-linear. A Figura 5 apresenta 2(duas) funções de transferência, normalmente utilizadas neste trabalho, para uso subsequente à filtragem pelo operador laplaciano.

$$\begin{array}{ccc} \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} & \frac{1}{25} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} & \frac{1}{15} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 1 \end{bmatrix} \\ (a) & (b) & (c) \end{array}$$

Fig. 2 - Filtros passa baixa

- (a) média em uma região 3 x 3
- (b) média em região 5 x 5
- (c) filtro gaussiano passa-baixa.

$$\begin{array}{cc} \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} & \begin{bmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{bmatrix} \\ (a) & (b) \end{array}$$

Fig. 3 - Aproximações digitais do operador, laplaciano no $\nabla^2(\cdot) = \partial^2/\partial x^2 + \partial^2/\partial y^2$

$$\begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix}$$

Fig. 4 - Aproximação digital do operador bi-laplaciano no $\partial^4/\partial x^2 \partial y^2$

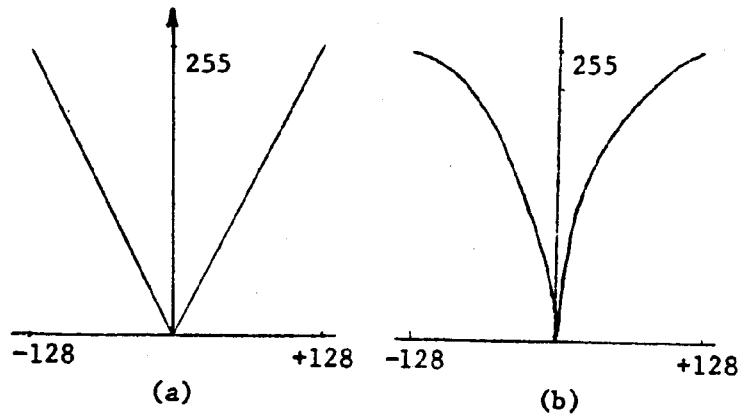


Fig. 5 - Curvas de transferência não lineares para uso posterior à filtragem linear.

- (a) módulo multiplicado por 2,
- (b) tipo raiz quadrada.

O uso da função de transferência do tipo raiz quadrada visa redistribuir os níveis de cinza resultantes da operação $|\nabla^2(\cdot)|$, visto que a maior parte desses valores concentra-se na região dos níveis baixos.

Define-se, também, um operador heurístico chamado "variação", cuja saída é proporcional à rugosidade do relevo em uma pequena região 3 x 3. Considerando-se uma vizinhança 3 x 3 de um ponto, podem-se rotular os pontos dessa vizinhança segundo a convenção:

```
a b c
d x e
f g h
```

A função denominada variação (Schachter et alii, 1979) é a soma da variação vertical horizontal. Assim:

$$VTV = |a-d| + |d-f| + |b-x| + |x-g| + |c-e| + |e-h| \quad (1)$$

$$HTV = |a-b| + |b-c| + |d-x| + |x-e| + |f-g| + |g-h| \quad (3)$$

$$\text{Variação total} = V.T. = VTH + HTV \quad (4)$$

4. O ALGORITMO DE CLASSIFICAÇÃO POR MÁXIMA VEROSSIMILHANÇA

O método escolhido neste algoritmo, para a classificação, é o do tipo estatístico paramétrico. Supõe-se que o comportamento das classes envolvidas pode ser descrito por funções de densidade de probabilidade gaussianas.

Faz-se a classificação ponto a ponto, ou seja, supõe-se que cada "pixel" é não-correlacionado com o vizinho, porque a classe a ser

associada ao "pixel" não depende dos vetores representativos dos "pixels" vizinhos. Em algoritmos de classificação, chamados *decisão composta*, a classe do ponto não vai depender somente do vetor que o representa, mas, também, dos valores dos "pixels" vizinhos, introduzindo, assim, a informação espacial como elemento a influenciar a decisão. No caso do presente trabalho, embora a classificação seja feita ponto a ponto, uma ou mais das dimensões do vetor que representa o "pixel" é proporcional a algum tipo de informação espacial em torno desse ponto. A segunda maneira oferece, no entanto, um eficiência computacional muito maior.

O algoritmo de classificação chamado de *decisão por máxima verossimilhança* pode ser descrito resumidamente (Velasco et alii, 1978): decide-se que um padrão \vec{x} pertence à classe ω_i , se uma função $g_i(\vec{x})$, chamada função discriminante para a classes i , satisfizer às seguintes condições:

$$g_i(\vec{x}) > g_j(\vec{x}) \quad i \neq j \quad i, j = 1, 2, \dots, M, \quad (5)$$

onde M é o número de classes

$$g_i(\vec{x}) > T_i \quad \text{onde } T_i \text{ é o limiar para a classe } i.$$

A função discriminante, para o caso gaussiano, é dada por:

$$g_i(\vec{x}) = \ln P(\omega_i) - \frac{1}{2} \ln |\Sigma_i| - \frac{1}{2} (\vec{x} - \vec{\mu}_i)^T \Sigma_i^{-1} (\vec{x} - \vec{\mu}_i), \quad (6)$$

onde μ_i é o vetor média para classe i e Σ_i é a matriz de covariância para a classe i .

Faz-se a avaliação do classificador, categorizando-se, novamente, as próprias áreas de treinamento, utilizadas para estimar os parâmetros de cada classe, e também categorizando-se as outras áreas, que não de treinamento, mas que têm classificação conhecida "a priori", chamadas áreas testes.

Os erros que o classificador vai cometer ao categorizar pontos de identidade conhecida darão uma idéia da probabilidade de erro envolvida no processo.

5. SELEÇÃO DE ATRIBUTOS

Analisar-se-á, com brevidade nesta seção como escolher N canais dos M disponíveis, depois de um processo de extração de atributos.

O critério natural para a escolha do subconjunto com N elementos é o da minimização da probabilidade de erro. Acontece que a aplicação direta desse critério envolve custos computacionais elevados. Optou-se, portanto, pela utilização de uma classe de funções que indicam a "distância normalizada" entre funções densidade de probabilidade. Quanto maior a distância entre funções, menor a probabilidade de erro envolvida entre as classes.

Utilizar-se-á, neste trabalho, a chamada "Distância de Bhattacharyya" (Distância-B) e sua relacionada "Distância Jeffreys-Matusita" (Distância JM).

A distância-B é dada por:

$$B = - \ln \rho, \quad (7)$$

onde ρ é o coeficiente de Bhattacharyya, da do por:

$$\rho = \int (p(\vec{x}/\omega_1) p(\vec{x}/\omega_2))^{1/2} dx \quad (8)$$

A distância JM é dada por:

$$d_{JM}^2 = 2(1 - \rho) \quad (9)$$

É possível derivar limitante superiores a inferiores da probabilidade de erro, (P_E), em função do coeficiente de Bhattacharyya (Chen, 1973):

$$\frac{1}{4} \rho^2 \leq P(\omega_1) P(\omega_2) \rho^2 \leq \frac{1}{2} (1 - \sqrt{1 - 4P_1 P_2 \rho^2}) \leq P_E \leq \sqrt{P_1 P_2} \quad \rho \leq \frac{1}{2} \rho \quad (10)$$

Para densidades gaussianas, a distância B é dada por:

$$B = \frac{1}{8} (\vec{\mu}_1 \vec{\mu}_2)^T \left(\frac{\Sigma_1 + \Sigma_2}{2} \right) (\vec{\mu}_1 - \vec{\mu}_2) + \frac{1}{2} \ln \left\{ \frac{\frac{1}{2} (\Sigma_1 + \Sigma_2)}{|\Sigma_1|^{1/2} |\Sigma_2|^{1/2}} \right\} \quad (11)$$

onde: $\vec{\mu}_1$ e $\vec{\mu}_2$ são vetores médias, e

Σ_1 e Σ_2 são matrizes de covariância para as classes 1 e 2.

6. RESULTADOS

O primeiro experimento foi efetuado com uma imagem do satélite LANDSAT, sobre a área de Ribeirão Preto, de órbita 178, ponto 27, de abril de 1978.

Oa doze atributos utilizados foram os seguintes:

- os canais de 1 a 4 são canais originais de 4 a 7 do LANDSAT;
- os canais de 5 a 8 foram obtidos da seguinte maneira: cada canal original de 4 a 7 do LANDSAT foi processado pelo filtro laplaciano da Figura 3b; reescalado pela curva do tipo raiz quadrada da Figura 5b e, em seguida, cada canal resultante foi suavizado pela máscara da Figura 2b;
- os canais de 9 a 12 foram obtidos dos originais pela filtragem através da máscara da Figura 2b (suavização).

Obtiveram-se, assim, 12 canais para escolher quatro, que é a capacidade máxima de manipulação do imageador automático I-100, do laboratório de Interpretação Automática do INPE, em São José dos Campos.

Os canais escolhidos, pelo critério do subconjunto com máxima distância JM média, e do subconjunto de canais com máxima distância JM mínima entre os pares de canais, foram os mesmos, que são:

- 6 - laplaciano suavizado do canal 5 do LANDSAT;
- 9 - média 5 x 5 do canal 4 do LANDSAT;
- 10 - média 5 x 5 do canal 5 do LANDSAT;
- 12 - média 5 x 5 do canal 7 do LANDSAT.

As classes usadas são mostradas na Tabela 1.

TABELA 1

CLASSES DE INTERESSE

	CLASSE	Nº DE PONTOS -AREA DE TREINAMENTO	Nº DE PONTOS -AREA TESTE
1	cana	252	108
2	cana nova	216	108
3	pasto	108	72
4	água	72	36
5	infraestrutura	72	36
6	mata	72	36

O segundo experimento foi estudado para a mesma imagem do experimento anterior. Extraíram-se também mais 8 atributos, além dos 4 originais.

Os canais de 1 a 4 e de 9 a 12 são os mesmo anteriores. Os atributos de 5 a 8 foram agora obtidos pelo filtro de variação (Equação 4), aplicado a cada canal original. Em seguida, cada um deles foi suavizado, realizando-se uma média, como indica a Figura 2b.

Obtiveram-se da mesma forma, outros 12 canais, dos quais foram escolhidos quatro, pelo mesmos critérios do experimento anterior. Também nesse caso, o subconjunto escolhido é o mesmo, considerando-se os 2 critérios disponíveis

Os canais escolhidos foram:

- 7 - variação do canal 6, suavizada pelo filtro de média da Figura 2b;
- 9 - média 5 x 5 do canal 4;
- 10 - média 5 x 5 do canal 5;
- 11 - média 5 x 5 do canal 7.

Para comprovar a eficiência desse método, foram obtidas as matrizes de classificação das áreas de treinamento, essas matrizes apresentam, de forma sucinta, o resultado da classificação das áreas de identificação conhecida - áreas de treinamento - quando é aplicado a essas áreas o algoritmo de classificação com os parâmetros obtidos a partir delas. Os erros cometidos ao classificar incorretamente os pontos de identidade conhecida permitem estimar os erros envolvidos e também as matrizes de classificação para áreas teste (áreas de identidade conhecida, mas que não foram usadas na extração dos parâmetros para a classificação).

A partir dessas matrizes foi possível obter o "desempenho médio", definido como a média das percentagens de classificação correta de cada área de treinamento, ponderada pelo número de pontos de cada uma. A "abstenção média" foi definida como sendo a percentagem média de abstenção das áreas de treinamento, ponderada pelo número de pontos de cada área. A "confusão média" foi definida como sendo o erro médio ponderado de classificação das áreas.

As Tabelas 2 e 3 apresentam, para as áreas de treinamento e teste, os valores obtidos para o desempenho médio, a confusão média e a abstenção média, considerando-se os canais originais e os escolhidos pelos 2 (dois) experimentos, com limiar igual a 5 para a classificação.

Os resultados demonstraram uma boa melhoria no desempenho médio, e uma sensível diminuição na confusão média, pelo uso dos atributos espaciais. Na maioria dos casos, porém, a abstenção média aumentou, resultado que não é tão nocivo quanto à classificação errônea. Isso indica que o poder discriminatório do algoritmo de classificação, com base nos novos canais, aumentou, ou seja, a classificação ficou mais aguda em virtude das variâncias em jogo terem diminuído.

Observando-se diretamente as matrizes de classificação dos novos canais, notou-se que, nos casos em que as áreas de treinamento estavam localizadas próximas às fronteiras entre diferentes regiões, a percentagem de classificação correta diminuiu. Isto se deve ao efeito de "borramento", ou difusão, que ocorre nas fronteiras entre regiões, quando se usa operadores lineares como as médias.

TABELA 2

RESULTADOS DAS MATRIZES DE CLASSIFICAÇÃO PARA AS ÁREAS DE TREINAMENTO

	CANAIS ORIGINAIS	CANAIS ESCOLHIDOS -1ª EXPERIÊNCIA	CANAIS ESCOLHIDOS -2ª EXPERIÊNCIA
Desempenho médio (D_m)	94.8%	99.5%	98.7%
Confusão média (C_m)	4.8%	0.1%	0.0%
Abstenção média (A_m)	0.4%	0.4%	1.3%

TABELA 3

RESULTADOS DAS MATRIZES DE CLASSIFICAÇÃO PARA AS ÁREAS TESTE

	CANAIS ORIGINAIS	CANAIS ESCOLHIDOS -1ª EXPERIÊNCIA	CANAIS ESCOLHIDOS -2ª EXPERIÊNCIA
D_m	77.8%	80.3%	89.4%
C_m	17.4%	3.3%	1.0%
A_m	4.8%	16.4%	9.6%

O campo classificado mostrou-se mais homogêneo, e menos ruidoso do que no caso em que os canais originais foram usados; as fronteiras entre regiões não foram classificadas, o que indica que o método desenvolvido permitiu detectar melhor as regiões de classificação normalmente confusas, como é o caso das regiões de transição entre campos de diferentes texturas, e não classificá-las, ao invés de classificá-las erradamente.

Novos testes serão efetuados, com outros operadores, para extrair atributos de relevo. Pode-se estudar também a utilização de operadores de suavização, ou passa-baixa, não-lineares, de maneira a evitar o efeito de borramento das fronteiras. Estudos mais profundos devem ser feitos para analisar o problema dos erros de classificação na fronteira entre regiões.

A rotação espectral poderá ser usada também como pré-processamento, antes das extrações de atributos espaciais. Os resultados

preliminares, no entanto, demonstram que o método pode ser útil na resolução de vários problemas de sensoriamento remoto.

Outras referências relacionadas com o assunto são: Haralick (1979), Panda (1977) e Pickett (1970).

7. REFERÊNCIAS BIBLIOGRÁFICAS

- CHEN, C.H. *Statistical pattern recognition*. Rochelle Park, N.J., Spartan Books, 1973.
- HARALICK, R.M. Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67(5):786-804, May 1979 (survey).
- PANDA, D.P. *Statistical analysis of some edge operators*. Maryland, M.D., University of Maryland, 1977. (Tech. Report n° 558).
- PICKETT, R.M. Visual analysis of texture in the detection and recognition of objects. In: LIPKIN, B.S.; ROSENFELD, A., ed. *Picture Processing and Psychopictorics*, New York, N.Y., Academic Press, 1970. p. 289.308.

SCHACHTER, B.J.; DAVIS, L.S.; ROSENFELD, A.
Some experiments in image segmentation by
clustering of local features values.
Pattern Recognition, 11(1):19-28, 1979.

VELASCO, F.R.D.; PRADO, L.O.C.; SOUZA, R.C.M.
Sistema MAXVER: manual do usuário. São Jo
sé dos Campos, INPE, jul. 1978. (INPE-1315
-NTI/110).