

RESPOSTAS ESPECTRAIS DE SOJA E ALGODÃO A PARTIR DE DADOS DIGITAIS DE TM-LANDSAT

Sherry Chou Chen
Instituto de Pesquisas Espaciais - INPE
Caixa Postal 515
12201 - São José dos Campos - SP

RESUMO

O sucesso da classificação digital do TM-Landsat depende fundamentalmente do conhecimento das respostas espectrais representativas dos alvos de estudo. Dentre os alvos naturais, as respostas espectrais das culturas são mais heterogêneas devido às diferenças nas variedades, estágios fenológicos, tratos culturais etc. Considerando a importância da produção agrícola do Estado do Paraná na economia do País, um trabalho foi realizado para estudar as respostas espectrais da soja e do algodão a partir dos dados digitais do TM-Landsat. Estas informações são indispensáveis na identificação de culturas através de dados de sensoriamento remoto com a finalidade de fornecer estatísticas agrícolas confiáveis em tempo hábil. Neste trabalho, 120 segmentos foram sorteados aleatoriamente na área de estudo que abrange 3 cenas da órbita 223 do TM-LANDSAT na Estado do Paraná. Através de um sistema de classificação por máxima verossimilhança, 7 classes espectrais distintas de soja e 5 de algodão foram obtidas usando dados digitais dos canais 3, 4 e 5 sendo apresentadas as estatísticas destas 12 classes. O resultado deste trabalho demonstra que, existem diferenças relativas das respostas espectrais de soja, algodão, pastagem e mata natural. Estas diferenças espectrais observadas a partir dos dados de TM-Landsat podem ser usadas juntamente com outras informações texturais e contextuais num sistema especialista para classificar as culturas. Para proceder à classificação convencional de culturas usando dados espectrais, o programa k-médias implementado no SITIM-150 deve ser modificado. O programa de "agrupamento" precisa possuir a capacidade de analisar um grande volume de segmentos armazenados no disco e definir as classes espectrais homogêneas antes de iniciar a classificação MAXVER. A modificação de k-médias do SITIM-150, no sentido de agrupar os "pixels" de diversas áreas de treinamento da mesma cultura em várias classes espectrais homogêneas tornará a classificação espectral convencional mais eficiente.

ABSTRACT

The success of digital classification depends fundamentally on the spectral representativeness of training samples to the study targets. Among natural targets spectral responses of crops are the most heterogeneous due to differences on varieties, phenological stages, culture practices etc. Viewing on the importance of the agriculture productions of Paraná State to the national revenue, a study was carried out to characterize the spectral responses of soybeans and cotton using Landsat-TM digital data. This spectral information is essential for crop identification using remote sensing data to provide agriculture statistic in a reasonable lead time. In this study 120 segments were sampled randomly from three full scenes of DA 223 of Paraná State. Based on the maximum likelihood decision rule 7 and 5 spectral classes were formed for soybeans and cotton respectively using Landsat-TM digital data of channels 3, 4 and 5, and statistics of these 12 classes were presented. Study results show that there are relative differences in spectral response among soybean, cotton, pasture and natural forest. These spectral differences observed from Landsat-TM channels can be included together with textural and contextual parameters in an expert system for crop classification. For the conventional crop classification based on spectral data the k-means program implemented in the SITIM-150 should be modified to provide the capability of analysing large volume stored information from sampled segments to defined homogeneous spectral classes for further classificatio

1. INTRODUÇÃO

Os objetos naturais na superfície da Terra podem ser discriminados através de três atributos de dados de sensoriamento remoto, ou seja, espectral, espacial e temporal. A aplicação da informação temporal na discriminação requer um conjunto de dados multitemporais que, muitas vezes, é impossível a obtenção devido à alta frequência de cobertura de nuvens. Este problema é especialmente grave quando a intenção é classificar as culturas de verão no sul do País, onde o período apropriado para adquirir os dados de satélite coincide com a época chuvosa. A ênfase na utilização do atributo espacial na classificação digital foi iniciada após o lançamento dos satélites de segunda geração (i.e. TM-Landsat e SPOT), devido aos resultados não satisfatórios de classificação digital com base somente nas informações espectrais (Hopkins et al., 1988; Haack et al., 1987; Azevedo et al., 1984; Markham et al., 1981). Entretanto, os algoritmos desenvolvidos até agora, que utilizam o atributo espacial, ainda não avançaram ao estágio operacional. Por isso, os conhecimentos sobre as características espectrais de alvos naturais ainda são fundamentais na análise e na interpretação de dados de sensoriamento remoto.

Entre os alvos naturais, as respostas espectrais das culturas são mais heterogêneas devido às diferenças nas variedades, estágios fenológicos, orientação e espaçamento de fileira, tratamentos culturais etc. Estas heterogeneidades são mais ressaltadas quando envolvem uma grande área geográfica, como no caso do projeto de previsão de safras a nível estadual. Considerando a heterogeneidade das respostas espectrais de culturas na classificação digital, o procedimento de classificação supervisionada é normalmente aplicado. Neste caso, as áreas de treinamento que representam todas as variações espectrais das culturas de interesse devem ser amostradas na área de estudo, pois a representatividade é imprescindível para o sucesso da classificação digital (Mather, 1987; Swain, 1978). Por outro lado, coletar as informações de culturas numa grande área geográfica não é uma tarefa trivial, além de ser dispendiosa.

Por isso, em 1986 foi iniciado um projeto de cooperação INPE-IBGE para realizar o primeiro estudo espectral de culturas, de modo sistemático, a partir de dados digitais do TM-Landsat. Isso

quer dizer que as características espectrais definidas neste estudo servem para discriminar as culturas, estimar as áreas plantadas e fornecer as estatísticas agrícolas de um sistema operacional em nível estadual.

Os principais objetivos deste estudo são: a) conhecer as respostas espectrais representativas de culturas para classificação digital; e b) verificar o ambiente de trabalho do SITIM para um projeto operacional de previsão de safras.

2. MÉTODOS

Num sistema operacional, para estimar as áreas cultivadas com base nos dados TM-Landsat, o Estado é normalmente dividido em vários distritos de análise (DA). Um DA é a área do Estado imageada pelo o sensor do Landsat no mesmo dia e na mesma órbita. As estatísticas espectrais representativas dos alvos de interesse são definidas separadamente para cada DA; evitando assim as variações espectrais causadas pelos dados multitemporais. O Estado do Paraná é dividido em cinco DAs e os DAs podem conter duas ou três imagens inteiras do Landsat (Figura 1). O DA 223 foi escolhido para este estudo devido à disponibilidade de dados TM-Landsat sem nuvens e à concentração de cultura de verão neste distrito.

a) Organização de Dados

O Distrito de Análise 223 é localizado entre as longitudes 52° a 54° W e as latitudes 22° 30' a 26° 30' S. No DA 223, 120 áreas de treinamento, i.e. segmentos, foram amostradas. Todos os segmentos foram visitados no campo no início de 1987 para coletar as informações que serviram para o estudo espectral de culturas. O detalhe do método estatístico para a seleção destas amostras encontra-se no trabalho de Mueller et al. (1988).

No ambiente do SITIM, o meio mais eficiente de trabalhar com pequenas áreas é transportar as informações espectrais da fita CCT para o disquete de 5.1/4". Por isso, uma janela de 256 x 256 "pixels", em volta de cada segmento foi extraída das CCTs da passagem de 10/02/87. As informações espectrais dos canais 3, 4 e 5 de todos os "pixels" dentro da janela foram copiadas para o disquete. Posteriormente, os limites dos talhões de cada segmento, com diferentes coberturas vegetais, foram

digitalizados. Assim, para cada segmento, as informações espectrais do TM e os tipos de culturas dos talhões rotulados foram armazenadas em um único disquete para serem analisadas.

b) Procedimento usado para analisar as características espectrais de soja e algodão

Dada a heterogeneidade das respostas espectrais de culturas, a classificação supervisionada, baseada na máxima verossimilhança (MAXVER), é a mais indicada. Porém, "a priori" à regra de decisão, exige-se a distribuição Gaussiana dos dados. Por isso, um problema encontrado para iniciar a classificação é a determinação do número de classes espectrais homogêneas, as quais, muitas vezes, são desconhecidas pelos usuários que por isso as definem subjetivamente (Moik, 1980). Este problema é mais grave para dados do TM, devido ao maior grau de variação espectral, quando se compara com os dados do MSS em consequência da melhor resolução espacial.

Então, neste estudo, foi aplicado o procedimento híbrido de classificação que define primeiramente o número de classes homogêneas através de um método não-supervisionado (K-médias). Após esta definição, os vetores das médias e as matrizes de covariância destas classes são usados na regra de decisão do MAXVER. Como o K-médias do SITIM-150 é programado para analisar um grupo de "pixels" escolhido pelo cursor do monitor de imagem (i.e., a tela), cujo espaço máximo é de 1024 x 1024 "pixels", é impossível mostrar todos os talhões da mesma cultura simultaneamente na tela para realizar o agrupamento espectral. Diante desta limitação, a definição das classes espectrais foi feita da seguinte maneira: no primeiro segmento, o cursor foi colocado dentro de um talhão para adquirir "pixels puros" e representativos desta classe de informação (i.e. soja, algodão, pastagem ou mata), formando as primeiras classes espectrais para cada alvo presente no segmento. Assim, o primeiro arquivo de estatísticas espectrais foi criado e usado posteriormente para classificar o primeiro segmento através do algoritmo MAXVER. O resultado desta classificação foi visualizado e avaliado na tela do SITIM. Se algum talhão do segmento ou parte dele não foi classificado, então os "pixels" desta área foram usados para gerar a segunda classe espectral do arquivo anterior e a classificação MAXVER foi executada novamente no

segmento 1. Este processo foi repetido até que todos os talhões do segmento 1 fossem devidamente classificados. Depois, o arquivo final obtido no segmento 1 serviu para iniciar a classificação do segundo segmento. Desta maneira, o processo iterativo foi aplicado sucessivamente para todos os 120 segmentos do DA 223, e as classes espectrais homogêneas foram definidas de acordo com o critério de máxima verossimilhança.

3- RESULTADOS E DISCUSSÃO

Ao todo, existiam 111 talhões de soja e 88 de algodão nos 120 segmentos amostrados. A maioria dos talhões encontra-se na cena 77, o que indica alta intensidade de uso do solo para fins agrícolas (Tabela 1). Esses talhões foram usados para caracterizar as respostas de soja e algodão com base no critério de máxima verossimilhança. A Tabela 2 mostra, surpreendentemente, que sete classes espectrais foram definidas para soja e cinco para o algodão usando a imagem TM-Landsat de 10/02/87. Analisando as estatísticas espectrais da Tabela 2 (i.e. as médias e variâncias), observou-se que, de maneira geral, as diferenças espectrais da soja e do algodão foram mais nítidas nos canais 4 e 5 do que no canal 3. A pastagem (P) foi o alvo que mais confundiu com a soja (S) especialmente entre as classes P2 e S3. A soja 3 (S3) foi um plantio tardio. No dia da passagem do satélite, as plantas deste talhão ainda eram pequenas e suas respostas espectrais foram muito influenciadas pela presença do solo. Entretanto, a proporção da classe S3 na área de estudo não foi expressiva; consequentemente sua omissão ou inclusão na classificação não causou disparidades nos resultados.

Água, solo e vegetação são os três principais alvos na superfície da Terra. Como a vegetação é o único alvo cuja reflectância no canal do infravermelho próximo (por ex. canal 4 do sensor TM) é muito mais alta do que o solo e a água, a sua discriminação é relativamente trivial usando-se a informação deste canal.

Por outro lado, a discriminação das subclasses de vegetação (i. e. pastagem e tipos de culturas) é mais sutil. um fato interessante é a existência de valores de limiar para a soja e o algodão. O conhecimento do especialista sobre estes limiares pode ser usado na segmentação de imagens para separar os alvos de soja, algodão, pastagem e mata.

As regras de decisão preliminares do sistema de especialista, com base na Tabela 2 e os coeficientes de variação da Tabela 3, excluindo as classes P2 e S3, são:

se $C4 \geq 106$, $C5 < 87$
e c.v. do $C3 = 6$ então é algodão;

se $C4 \geq 106$, $C5 < 87$
e c.v. do $C3 > 6$ então é pasto;

se $C4 > 106$ e $C5 > 87$
então é soja;

se $C4 < 106$ e $C5 < 87$
então é mata.

Uma árvore de classificação baseada no atributo espectral é apresentada na Figura 2. A classificação digital usando somente os limiares espectrais é limitada. Porém, estes limiares juntamente com outras informações texturais ou contextuais (i.e. características espectrais dos "pixels" vizinhos) podem ser usados no processo de segmentação de imagem, melhorando o desempenho da classificação convencional.

As seguintes modificações do software do SITIM-150 foram recomendadas para torná-lo mais versátil para atender a todos os usuários:

- Programa Traço

Uma vez digitalizados vários polígonos, os usuários necessitam da soma das áreas de um subconjunto ou de todos os polígonos digitalizados, sem repetir o processo de digitalização.

- Arquivo MAXVER

Na classificação MAXVER, o vetor das médias e a matriz de covariância são as principais estatísticas usadas na regra de decisão, e o conhecimento destas estatísticas é indispensável para os usuários. Por estas razões, na saída da impressora, as matrizes de covariância e correlação devem ser apresentadas, em vez das matrizes de transformação e autocorrelação que são raramente aproveitadas pelos usuários. Mesmo a matriz de covariância pode ser calculada com base na matriz de autocorrelação; porém, as informações contidas na saída do programa devem ser as mais diretas possíveis.

- Filtragem Espacial

Apresentação da área em hectares ou número de "pixels" classificados após a aplicação de filtragem espacial-UNITOT.

- K-média

Permite aplicar o algoritmo K-médias num arquivo de amostras armazenadas no disco.

4. CONCLUSÕES

As respostas espectrais dos 120 segmentos, distribuídos na região agrícola mais importante do Estado do Paraná foram analisados a partir de dados TM-Landsat de 10/02/87. As presenças das culturas de verão (i.e. soja, milho, algodão, arroz e feijão), contidas nestes segmentos, foram verificadas no campo e seus dados digitais correspondentes usados para caracterizar as culturas. Após o levantamento de campo, constatou-se um número insuficiente de amostras para representar arroz e feijão. Por outro lado, a multiplicidade dos estágios fenológicos do milho, que variavam desde recém-plantado até talhões totalmente secos, limitou o desempenho de classificação unitemporal. Por estas razões, somente a soja e o algodão foram estudados. Os resultados da análise digital mostram que as respostas espectrais de culturas são bastante heterogêneas. Existem 7 e 5 classes espectrais para soja e algodão respectivamente. Os canais 4 e 5 do TM-Landsat são mais úteis para separar estas duas culturas, enquanto a informação do canal 3 foi substancial para discriminar entre algodão e pastagem. Os valores limiares foram encontrados e podem ser usados juntamente com outras informações texturais ou contextuais no processo de segmentação de imagem, melhorando o desempenho da classificação convencional. As sugestões nas modificações de alguns programas do SITIM-150 foram feitas e suas implementações tornarão o "software" mais eficiente e versátil para os usuários.

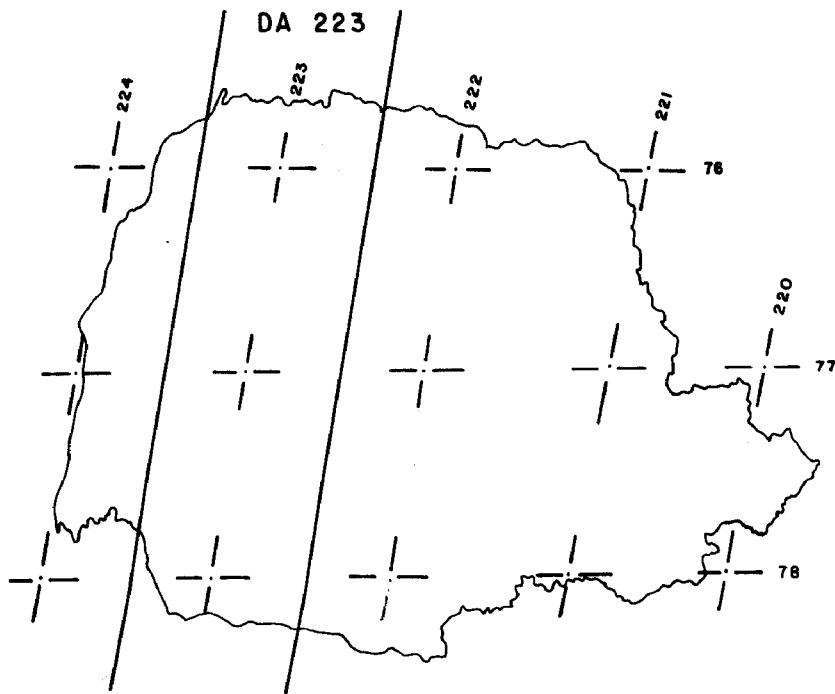


Fig. 1 - Localização das imagens do satélite TM-LANDSAT no Estado do Paraná.

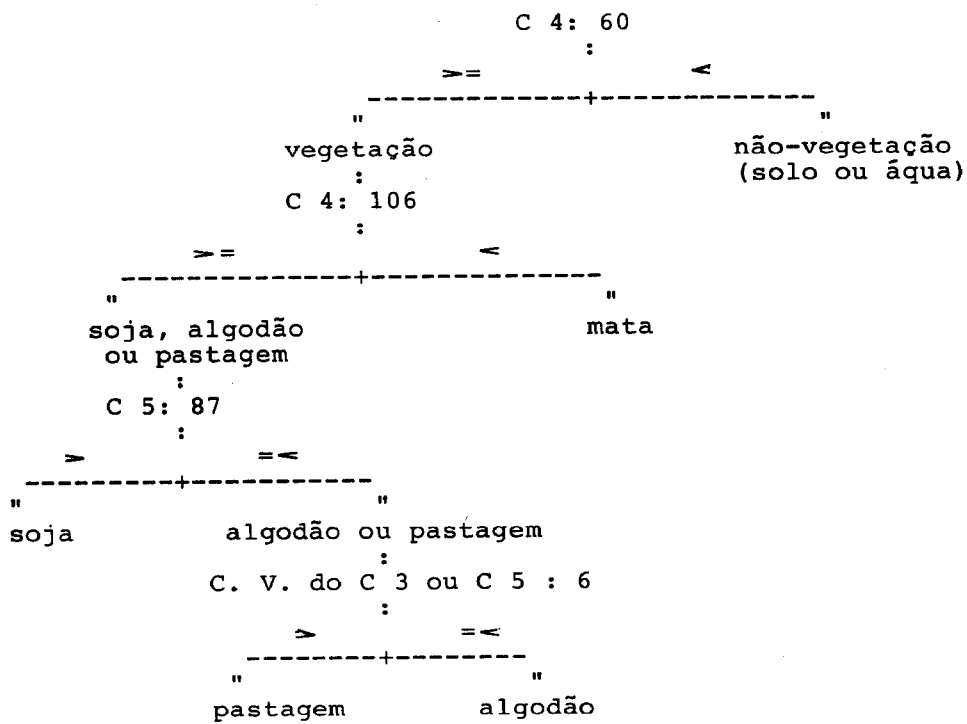


FIGURA 2. A ÁRVORE DE CLASSIFICAÇÃO

TABELA 1
DISTRIBUIÇÃO DE TALHÕES DE SOJA E ALGODÃO NOS
120 SEGMENTOS AMOSTRADOS

CENA QUADRANTE		Nº DE TALHÕES COM	
		SOJA	ALGODÃO
76	A	0	1
	B	0	5
	C	0	11
	D	15	5
	Subtotal	15	22
77	A	21	27
	B	24	11
	C	8	10
	D	4	18
	Subtotal	57	66
78	A	20	0
	B	9	0
	C	1	0
	D	9	0
	Subtotal	39	0
Total		111	88

TABELA 2

CLASSES ESPECTRAIS DE SOJA E ALGODÃO E SUAS CARACTERÍSTICAS USANDO
DADOS DIGITAIS DO LANDSAT-TM

CLASSE ESPECTRAL	CANAL TM LANDSAT					
	3		4		5	
	\bar{X}	S ²	\bar{X}	S ²	\bar{X}	S ²
Soja 1	21.27	0.50	137.18	18.58	107.64	8.56
Soja 2	19.75	0.69	154.14	19.91	107.40	6.05
Soja 3	22.16	2.33	112.98	31.22	92.74	18.32
Soja 4	20.74	2.28	153.87	17.93	116.28	9.43
Soja 5	18.28	0.98	129.21	44.71	87.24	26.47
Soja 6	24.25	0.99	117.36	9.47	106.75	4.73
Soja 7	20.87	1.01	145.41	12.95	99.13	15.02
Algodão 1	20.52	0.73	111.26	75.55	59.84	9.49
Algodão 2	22.97	1.04	106.90	67.29	68.29	9.93
Algodão 3	25.16	1.60	127.53	19.75	74.63	17.40
Algodão 4	21.79	0.74	147.06	11.41	86.08	17.33
Algodão 5	23.18	1.03	109.91	22.36	78.33	6.66
Pastagem 1	22.76	2.48	106.36	42.35	85.12	62.37
Pastagem 2	24.79	2.34	111.46	26.13	94.04	19.61
Mata 1	16.84	3.21	88.11	78.41	58.30	20.41
Mata 2	14.70	0.89	67.16	59.69	29.89	22.39
Mata 3	19.60	0.93	96.19	21.57	66.03	7.52

5. REFERÊNCIAS

AZEVEDO, W., J.S. BUIS, R.C. WRIGLEY. Changes in Classification Accuracy due to varying Thematic Mapper and multispectral scanner spatial, spectral and radiometric resolution. In: PROC. 18 INT'L SYMP. REMOTE SENSING OF ENVIRON. pp. 27-44, 1984.

HAACK, B., N. BRYANT, S. ADAMS. An assessment of LANDSAT MSS and TM data for urban and near-urban land-cover digital classification. Remote Sensing of Environ. 21(2): 201-213, 1987.

HOPKINS, P.F., A.L. MACLEAN, T.M. Lillesand Assessment of Thematic Mapper imagery for forest applications under lake states conditions Photogramm. Eng. and Remote Sensing 54(1): 61-68, 1988.

MATHER, P.M. Computer Processing of Remotely - Sensed Images. John Wiley & Sons. New York, N.Y. 1987. 329 p.

MARKHAM, B.L., J.R.G. TOWNSHEND. Land cover classification accuracy as a function of sensor spatial resolution. In: PROC. 15 INT'L SYMPOSIUM ON REMOTE SENSING OF ENVIRON., pp. 1075-1091, 1981.

MOIK J.G.. Digital Processing of Remotely Sensed Images. NASA special Publications 431, Washington DC, 1980.

MUELLER, C.C., G. SILVA, A.G. VILALOBOS. Sistema de informações agropecuárias - Pesquisas do Paraná 1987. Relatório de Trabalho, IBGE, 1988.

SWAIN, P.H. Fundamentals of Pattern Recognition in Remote Sensing. Remote Sensing: The Qualitative Approach (P.H. Swain e S.H. Davis. eds) MacGraw-Hill, New York, N.Y. pp. 137-187, 1978.

TABELA 3
COEFICIENTES DE VARIAÇÃO DAS CLASSES ESPECTRAIS

CLASSE ESPECTRAL	CANAL TM-LANDSAT		
	3	4	5
Soja 1	3.6	3.1	2.7
Soja 2	4.2	2.9	2.3
Soja 3	6.9	5.0	4.6
Soja 4	7.0	2.8	2.6
Soja 5	5.4	5.2	5.9
Soja 6	4.1	2.6	2.0
Soja 7	4.8	2.5	3.9
Algodão 1	4.2	7.8	5.2
Algodão 2	4.4	7.8	4.6
Algodão 3	5.0	3.5	5.6
Algodão 4	4.0	2.3	4.8
Algodão 5	4.4	4.3	3.3
Pastagem 1	6.9	6.1	9.3
Pastagem 2	6.2	4.6	4.7
Mata 1	10.6	10.0	7.8
Mata 2	6.4	11.5	15.8
Mata 3	4.9	4.8	4.2