
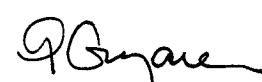



1. Publicação nº <i>INPE-2482-PRE/166</i>	2. Versão	3. Data <i>Julho, 1982</i>	5. Distribuição <input type="checkbox"/> Interna <input checked="" type="checkbox"/> Externa <input type="checkbox"/> Restrita
4. Origem <i>DIN/DPI</i>	Programa <i>SAFRAS</i>		
6. Palavras chaves - selecionadas pelo(s) autor(es) <i>PREVISÃO DE SAFRAS ESTIMATIVA DE ÁREA</i> <i>IMAGENS DE SATÉLITE</i>			
7. C.D.U.: <i>528.711.7:631.55</i>			
8. Título <i>INPE-2482-PRE/166</i>		10. Páginas: <i>12</i>	
<i>UM PROCEDIMENTO AUTOMÁTICO PARA CLASSIFICAÇÃO DE ÁREAS AGRÍCOLAS EM IMAGENS DE SATÉLITE, PARA USO EM PREVISÃO DE SAFRAS</i>		11. Última página: <i>07</i>	
9. Autoria <i>Gilberto Câmara Neto</i> <i>Flávio Roberto Dias Velasco</i> <i>Misae Odo Bueno de Oliveira</i>		12. Revisada por  <i>Getúlio Batista</i>	
Assinatura responsável 		13. Autorizada por  <i>Nelson de Jesus Parada</i> Diretor	
14. Resumo/Notas <i>Experimentos de previsão de safras em grande escala que utilizam imagens de satélite, tais como o projeto LACIE (feito nos E.U.A.), indicaram a importância de procedimentos de classificação adequados; tais procedimentos objetivam a medida da proporção de culturas agrícolas em segmentos de imagens de satélite, com a conseqüente agregação dos resultados em um sistema integrado de previsão de safras. Com este objetivo, vem sendo desenvolvido no INPE um procedimento interativo que envolve analista e computador, com quatro fases distintas: a) rotulação e treinamento; b) classificação; c) estimativa de proporção; d) avaliação e aprovação. Este procedimento tem ainda as seguintes características: 1) uso de várias passagens de satélites sobre a mesma área (classificação multitemporal); 2) necessidade de um mínimo de informação de campo; 3) capacidade de verificação entre classificação automática (feita por computador) e rotulação do analista. Neste trabalho apresenta-se uma descrição do fluxo de processamento, bem como discutem-se os principais algoritmos envolvidos. Dar-se-á ênfase, ainda, aos resultados obtidos e às perspectivas futuras dessa tecnologia.</i>			
15. Observações <i>Trabalho submetido para apresentação na 34ª Reunião da SBPC, 06 a 14 de julho de 1982, Campinas, São Paulo.</i>			



### ABSTRACT

Large area crop inventory experiments based on satellite data, such as the LACIE project, carried out in the US, have shown the importance of selecting adequate classification procedures; these procedures have the objective of assessing crop areal proportion in an MSS image and the results are aggregated using statistical methods to provide stratum estimates. A procedure is being developed at INPE for that application, divided into four parts: a) labeling; b) classification; c) proportion estimation; d) evaluation. The procedure also has the following characteristics: 1) multitemporal classification; 2) the need for a minimum field information; 3) verification capability between automatic classification and analyst labeling. In this work, a description of the processing steps is given, and the main algorithms involved are discussed. An outlook on the future of this technology is also presented.



## SUMÁRIO

	<u>Pág.</u>
1. INTRODUÇÃO .....	1
2. DESCRIÇÃO DO PROCEDIMENTO .....	3
2.1 - Rotulação - Treinamento .....	3
2.2 - Classificação .....	4
2.3 - Estimação da Proporção de cultura no segmento .....	5
2.4 - Avaliação - Aprovação .....	6
3. CONCLUSÃO .....	6
REFERÊNCIAS BIBLIOGRÁFICAS .....	7

## 1. INTRODUÇÃO

O Instituto de Pesquisas Espaciais (INPE) vem desenvolvendo um sistema de estimativa de safras agrícolas para as culturas de cana-de-açúcar, soja e trigo, a nível nacional, baseado, principalmente, em informações colhidas por satélites. As atividades do sistema de previsão de safras são as seguintes:

- estimativa de área plantada;
- estimativa de produtividade;
- integração da área e produtividade para a obtenção da produção das culturas.

A metodologia proposta para a estimativa de área de uma dada cultura, para fins de previsão de safras, é a de amostragem estratificada, que envolve os seguintes passos:

- 1) A região em estudo é dividida em zonas homogêneas, distribuídas de acordo com características de clima, solos e técnicas agrícolas comuns.
- 2) Cada zona homogênea é particionada em estratos.
- 3) Para cada estrato são escolhidas, de maneira aleatória, unidades amostrais denominadas segmentos.
- 4) A proporção da cultura em cada segmento é determinada.
- 5) As proporções estimadas para os segmentos são agregadas e é estimada a área da cultura no estrato.
- 6) As áreas dos estratos obtidos no passo 5 são agregadas e assim é obtida a área da cultura para a zona; a união das áreas estimadas para as zonas fornece a área total da cultura para a região.

O presente trabalho apresenta um procedimento para determinação da proporção de cultura nos segmentos (unidades amostrais). O procedimento deverá envolver a participação de analistas (agrônomos, foto-intérpretes) e computadores, e funcionará de modo interativo, com quatro fases distintas:

- rotulação e treinamento;
- classificação;
- estimativa de proporção;
- avaliação e aprovação.

Este procedimento tem as seguintes características:

- (a) Com o intuito de aumentar a discriminabilidade da cultura, o procedimento usa várias passagens do satélite. Para isto, segmentos obtidos em várias datas distintas serão digitalmente sobrepostos ("registrados").
- (b) O uso de informação de campo ("verdade terrestre") deverá ser minimizado, em função do alto custo geralmente associado a este tipo de dado.
- (c) O procedimento dispõe ainda de capacidade de verificação entre a classificação automática (feita por computador) e a rotulação do analista.

Experimentos anteriores, como o LACIE (Experimento para Previsão de Safras de Grandes Áreas), realizado nos E.U.A., evidenciaram a importância do uso de métodos adequados para estimativa de proporção das culturas nas unidades amostrais. O procedimento desenvolvido é baseado na experiência do LACIE (Heydorn et alii 1978), como também em trabalhos anteriores realizados no INPE, nas áreas de Reconhecimento e Classificação de Imagens.

## 2. DESCRIÇÃO DO PROCEDIMENTO

O Procedimento de determinação da proporção de cultura nos segmentos consiste em quatro fases distintas: rotulação - treinamento, classificação, estimação da proporção de cultura no segmento e avaliação-aprovação, com o fluxo de processamento ilustrado na Figura 1.

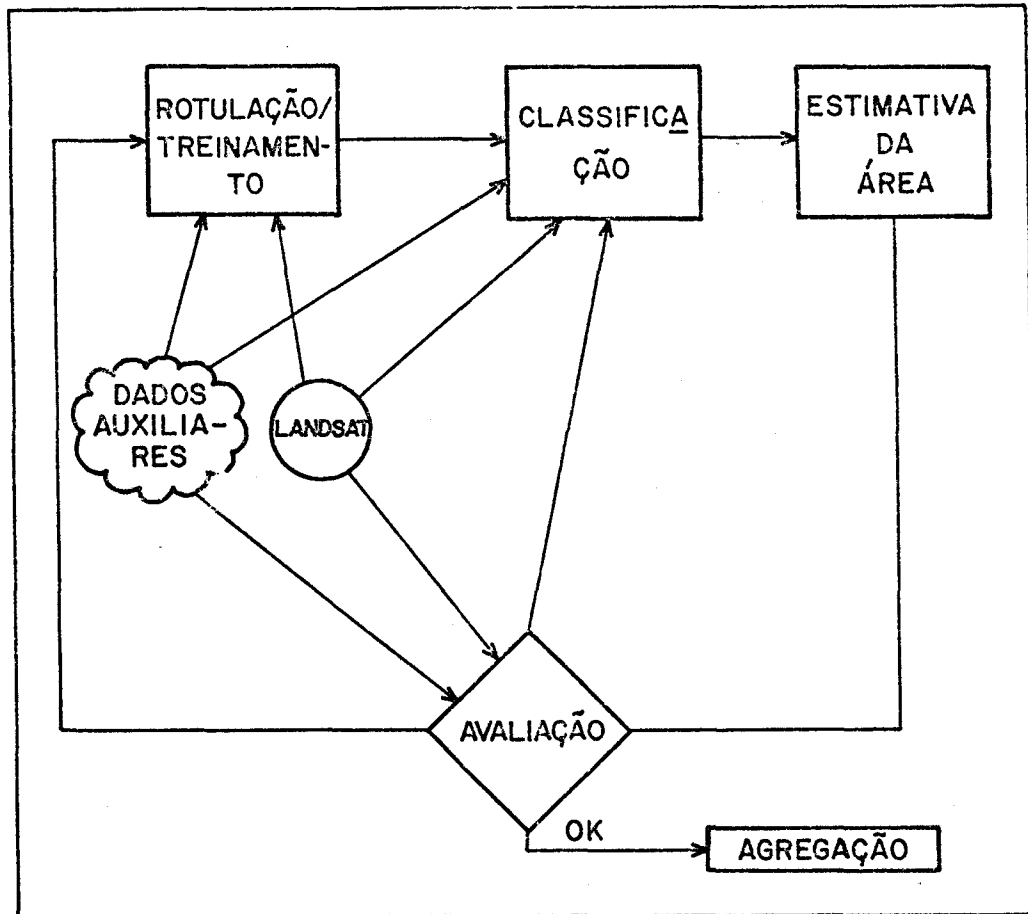


Fig. 1 - Fluxo de processamento.

### 2.1 - ROTULAÇÃO - TREINAMENTO

Nesta fase, o analista manualmente rotula pontos de imagem, com dois objetivos: (1) fornecer observações necessárias para estimar parâmetros para a classificação; e (2) dar subsídios para testar a



qualidade das estimativas. Para a rotulação, são escolhidos aleatoriamente dois subconjuntos de pontos em um segmento (denominados tipo 1 e tipo 2); os pontos do tipo 1 são utilizados para "treinar" o classificador e os do tipo 2, para a estimativa de erro e correção de tendência da classificação.

O analista poderá atribuir apenas dois rótulos aos pontos: cultura "x" (por exemplo, soja) e "outros" (exemplos, pastagem); estas também serão as classes resultantes da classificação. Alguns refinamentos são introduzidos nesse procedimento proposto pelo INPE, em relação à "Procedure 1" do LACIE (Heydorn et alii, 1978). Mais especificamente, as imagens a serem rotuladas poderão, a critério do analista, não corresponder às cenas originalmente geradas pelo satélite LANDSAT, mas sim a novos gerados a partir das imagens originais por transformações e filtragens; duas possibilidades se apresentam: a obtenção de novos canais representativos de "brilho" (cobertura do solo) e "verde" (crescimento da vegetação), por meio da transformação de "tasseled cap" (Kauth and Thomas, 1976); e a extração de atributos espaciais por meio de filtros locais, que procuram levar em conta a informação da textura da imagem para a classificação (Dutra e Mascarenhas, 1980).

Além disso o analista disporá, a exemplo do que acontece no LACIE (Spiers and Patterson, 1978), de informação adicional a respeito dos pontos a serem rotulados, com a finalidade de ressaltar características do desenvolvimento da cultura. Esta informação consiste: (1) na aplicação da transformação de "tasseled cap" aos pontos de interesse e determinação das coordenadas de "brilho" e "verde"; (2) na determinação do espalhamento espectral desses pontos para cada passagem e da trajetória espectral de cada ponto (valores de "brilho" e "verde" em cada passagem).

## 2.2 - CLASSIFICAÇÃO

A operação de classificação separa cada ponto do segmento em duas classes: cultura "x" e "outros", de acordo com a regra de máxi

ma verossimilhança. Adicionalmente, supõe-se que as classes cultura "x" e "outros" são divididas em subclasses, espectrais, considerando-se que a distribuição de cada uma delas seja gaussiana. O número de subclasses e os pontos que compõem cada subclasse são determinados através de um algoritmo de aglomeração ("clustering") do tipo "vizinho-mais-próximo", que utiliza os pontos do tipo 1 como sementes. A partir dos pontos de cada agregamento, são calculados os parâmetros das distribuições (média e matriz de covariância) de cada subclasse. As subclasses são então rotuladas (cultura x e "outros") pelo critério de mínima distância euclidiana. O rótulo da subclasse é o rótulo do ponto de tipo 1 mais próximo.

Por fim, todos os pontos do segmento são classificados segundo a regra de máxima verossimilhança.

### 2.3 - ESTIMAÇÃO DA PROPORÇÃO DE CULTURA NO SEGMENTO

O resultado da classificação é usado para a estimação de proporção de cultura no segmento. Nesta fase, os "pontos-teste" rotulados pelo analista na fase de rotulação-treinamento são utilizados para corrigir eventuais tendências ("bias") na classificação.

De maneira simplificada, o procedimento de correção de tendência obtém o estimador  $\hat{p}$  para a proporção de cultura no segmento a partir da equação:

$$\hat{p} = \hat{p}_{11} \cdot \lambda + \hat{p}_{10}(1 - \lambda), \quad (1)$$

onde  $\hat{p}_{11}$  é a razão entre o número de pontos do tipo 2, denominados classe 1 (cultura "x") pelo analista e classificados como classe 1 pela máquina, e o número de pontos do tipo 2, classificados pelo computador como classe 1; e  $\hat{p}_{10}$  é a razão entre o número de pontos do tipo 2, denominados classe 1 pelo analista e classificados como classe 0 ("outros") pela máquina, e o número de pontos do tipo 2, classificados como classe 0 pelo computador;  $\lambda$  é a razão entre o número de pontos classificados como classificados como classe 1 pela máquina e o número de pontos do segmento.

## 2.4 - AVALIAÇÃO - APROVAÇÃO

O resultado da estimação é submetido ao analista para aprovação; a seu critério, um ou mesmo todos os passos do procedimento poderão ser refeitos. Para a avaliação são também utilizados mapas de classificação. Uma vez obtida a proporção de cultura para cada segmento, os dados são agregados para fornecer a área plantada em uma determinada região.

## 3. CONCLUSÃO

O procedimento apresentado deve ser encarado como uma primeira aproximação de uma metodologia satisfatória para estimativa de áreas agrícolas em imagens de satélite; com base na experiência adquirida, futuros desenvolvimentos procurarão melhorar o desempenho do procedimento. Apontam-se aqui algumas áreas críticas, que certamente merecem estudos posteriores:

- a) O algoritmo de aglomeração para obtenção das subclasses ("vizinho-mais-próximo") é muito simples para ser ótimo. O desenvolvimento de algoritmos mais sofisticados, que utilizem informação espacial juntamente com a espectral, deverá melhorar o desempenho da fase de classificação.
- b) A rotulação de pontos isolados pelo analista pode ser uma fonte de erro, devido a diversos fatores como a mistura de classes dentro de um ponto. Uma alternativa seria fazer com que o analista rotulasse "regiões homogêneas" obtidas previamente através de aglomeração.
- c) A hipótese gaussiana para a distribuição das subclasses e o algoritmo de classificação por máxima verossimilhança podem não ser os mais adequados para o problema, dada a heterogeneidade das características disponíveis. Algoritmos mais apropriados à natureza dos dados serão, certamente, objeto de futuras pesquisas.

## REFERÊNCIAS BIBLIOGRÁFICAS

- DUTRA, L.V.; MASCARENHAS; N.D.A. *Extração de atributos espaciais em imagens multiespectrais*. São José dos Campos, INPE, 1980. (INPE-1885-RPE/229).
- HEYDORN, R.P.; BIZZELL, R.M.; QUIREIN, J.A.; ABOTTEEN, K.M.; SUMNER, C. A. Classification and mensuration of LACIE segments. In: LACIE SYMPOSIUM, Houston, 1978. *Proceedings*. Houston, NASA/JSC, 1978, p. 73-86.
- KAUTH, R.; THOMAS, G. The tasseled cap: a graphic description of spectral-temporal development of agricultural crop as seen in LANDSAT. In: MACHINE PROCESSING OF REMOTELY SENSED DATA, 3., West Lafayette, 1976. *Proceedings*. West Lafayette, Purdue University, 1976, p. 413-441.
- SPIERS, B.E.; PATTERSON, R.L. Ancillary data acquisition for LACIE. In: LACIE SYMPOSIUM, Houston, 1978. *Proceedings*. Houston, NASA/JSC, 1978, p. 157-162.