

# ANÁLISE DE MISTURAS ESPECTRAIS - UMA NOVA TÉCNICA DE CLASSIFICAÇÃO DE IMAGENS E SUAS APLICAÇÕES NO MAPEAMENTO E MONITORAMENTO DOS RECURSOS NATURAIS

Luciano J. de O. Accioly

Embrapa-CNPS-UEP Recife, Rua Antonio Falcão, 402 - Boa Viagem, Recife-PE. 51020-240

Fone: (081) 325-5988. Fax: (081) 325-0230. E-mail: [accioly@elogica.com.br](mailto:accioly@elogica.com.br)

**Resumo:** A análise de misturas espectrais (AME) é uma técnica de processamento de imagens utilizada para a extração de informações qualitativas e quantitativas ao nível de subpíxel para feições comumente encontradas nas imagens. Nesse trabalho preliminar a AME foi aplicada a uma janela (5 x 7 píxels) extraída de uma imagem do Landsat 5 TM (cena 216-66) de Mirandiba-PE. A recomposição dos píxels foi obtida com base nos três primeiros fatores. Três potenciais membros finais (endmembers) foram selecionados a partir do gráfico da dispersão dos pontos obtido com as bandas 3 (vermelho) e 4 (infravermelho próximo). A curva espectral de cada um dos potenciais membros finais é apresentada.

**Abstract:** Spectral Mixture Analysis (SMA) has been used to extract qualitative and quantitative information about common features in an imagery at the subpixel level. In this preliminary work SMA was applied to a image window (5 x 7 pixels) extracted from a Landsat 5 TM imagery (scene 216-66) of Mirandiba-PE. Pixel recomposition was obtained on the base of the first 3 factors. Three potential endmembers were selected on the base of the scatterplot obtained with band 3 (red) and band 4 (near infrared). The spectral curve of each one of the potential endmembers is presented.

## Introdução

Dependendo do campo de visada do sensor e da complexidade espacial dos elementos presentes em uma dada cena a radiância captada e, conseqüentemente, o sinal digital proveniente desse sensor, na maioria das vezes, representa a integração das radiâncias de dois ou mais componentes presentes no campo de visada do sensor. Esses componentes, também conhecidos como membros finais, são os elementos básicos utilizados nas análises de misturas. As misturas espectrais podem conter componentes diversos cujas contribuições para o sinal integrado recebido variam na dimensão espacial e temporal. Esses componentes podem incluir diferentes tipos de solo, diferentes tipos de vegetação, rochas diversas, resíduos orgânicos, entre outros elementos comumente encontrados nos limites do campo de visada de um dado sensor. A análise de misturas (também conhecida como análise fatorial ou análise de correspondências) é uma técnica de inversão matemática útil para a interpretação de problemas ligados aos campos da química e da ótica cujos dados são de natureza complexa e de dimensão multivariada (Huete, 1986).

Quando aplicada ao sensoriamento remoto a análise de misturas tem sido utilizada para decompor as curvas espectrais em espectros dos membros finais presente em um dado píxel. Além disso, com a análise de misturas é possível medir a contribuição relativa de cada um dos membros finais (Accioly, 1997; Accioly et al., 1998; Huete, 1986). Dessa forma a análise de misturas espectrais (AME) representa uma técnica especial de classificação de imagens que permite a classificação dos alvos ao nível de subpíxel. Accioly (1997) apresenta um quadro com as principais diferenças entre a classificação de imagens usando o método da máxima verossimilhança (MaxVer) e a AME. Muitas dessas diferenças estão implícitas nos conceitos de classe (extraídos da MaxVer) e de membros finais (extraídos da AME). Por exemplo, as classes são definidas usando-se amostras de treinamento cujas dimensões devem satisfazer os requerimentos da estatística e são, teoricamente, homogêneas. Na AME um membro final pode ser representado por um único píxel e, como este membro final é, normalmente, modelado com base em curvas espectrais provenientes de membros finais de referência, esse píxel não precisa ser necessariamente puro.

As primeiras técnicas utilizadas para a decomposição de píxels foram desenvolvidas no início dos anos setenta utilizando dados do sensor Landsat MSS. Naquela época foram sugeridas duas metodologias para estimar as proporções de componentes em um dado píxel: a classificação usando a probabilidade máxima (Horwitz et al., 1971) e a regressão linear (Richardson et al., 1975). O método da regressão linear foi inicialmente aplicado para extrair a contribuição das reflectâncias de plantas, solos e sombra de áreas cultivadas utilizando-se dados do sensor Landsat MSS. O modelo proposto por Richardson et al. (1975) foi utilizado para estudos das relações das frações de plantas, sombreamento e índice de área foliar (IAF) em 23 áreas cultivadas com a reflectância composta ( $R_c$ ) extraída a partir de imagens do Landsat MSS. Forshi (1994) revisou os diferentes métodos utilizados para a decomposição da resposta espectral de píxel com misturas de diferentes componentes e concluiu que os métodos estatísticos utilizados para estimar as frações desses componentes são bem sucedidos quando as classes de componentes apresentam uma boa separação.

Em anos recentes, com a perspectiva da popularização em um futuro próximo de dados provenientes de sensores hiper-espectrais aliada ao desenvolvimento de softwares e hardwares cada vez mais poderosos e acessíveis a um custo cada vez menor, vários modelos que atacam o problema da decomposição espectral ao nível de subpíxel tem sido sugeridos. Um dos modelos mais utilizados em dias atuais, e que tem servido como base para várias adaptações foi proposto por Smith et al. em 1987. Além de atacar o problema da decomposição espectral e avaliação das frações dos membros finais presentes em um dado píxel o modelo integra dados espectrais de laboratório no aprimoramento da calibração para a correção dos efeitos da atmosfera sobre o sinal gerado pelo sensor.

A análise de misturas espectrais como tem sido comumente utilizada, no entanto, exige a utilização de bibliotecas espectrais obtidas em laboratório ou ao nível de campo. Embora bibliotecas espectrais referentes a diversos minerais estejam disponíveis para os pesquisadores através da Internet (como, por exemplo, a biblioteca espectral do USGS) o mesmo não ocorre quando se trata de vegetação, solos e outros componentes cuja variabilidade está muito associada às condições edafoclimáticas locais. Dessa forma, a utilização da técnica da AME nesse trabalho na forma como vem sendo aplicada, exigiria a elaboração de uma biblioteca espectral, o que inviabilizaria o trabalho do ponto de vista de custos (um espectrorradiômetro para cobrir a faixa do sensor Landsat 5 TM está em torno de cem mil dolares), de difusão da tecnologia e de tempo. As modificações propostas nesse trabalho na técnica da análise de mistura proporcionarão atacar o problema espacial da diversidade de ambientes (como a presença de matas ciliares ao lado de vegetação de caatinga), o problema das variações espaço/temporais de ecossistemas do semiárido nordestino e de suas relações com a degradação ambiental nessa região. As adaptações propostas nesse trabalho representam sensível redução nos custos do monitoramento de mudanças ambientais (pois elimina a necessidade de uma biblioteca espectral construída com dados de campo e de laboratório) e sua simplicidade permitirá a fácil assimilação e utilização por técnicos ligados ao planejamento ambiental.

### Metodologia

A adaptação proposta nesse trabalho ao método de análise de misturas espectrais toma como base os trabalhos apresentados por Huete (1986; 1991) e Accioly (1997; 1998) que por sua vez se fundamentaram nos trabalhos de análise fatorial apresentados por Malinowski (1982) direcionados para a química analítica. Essa adaptação se fundamenta na utilização de alvos “puros” extraídos da própria imagem e utilização de modelos simples para a correção atmosférica.

A base do modelo da análise fatorial a ser utilizado nesse subprojeto considera que a resposta espectral de um dado píxel da imagem Landsat 5 TM pode ser representado pela combinação linear das somas balanceadas das respostas espectrais do membros finais (alvos puros) presentes nesse píxel:

$$P_c = \sum_{i=1}^N F_{i,c} E_c \quad (1)$$

Com a imposição de que

$$E_c = \sum_{i=1}^N F_{i,c} \quad (2)$$

Onde  $d_{i,k}$  representa a resposta espectral do píxel  $k$  na banda  $i$ ,  $n$  é o número de membros finais,  $r_{i,j}$  é a resposta espectral do membro final  $j$  na banda  $i$ , e  $c_{i,k}$  e a contribuição do membro final  $j$  para resposta espectral do píxel  $k$ . Usando-se a notação de matrizes a equação (1) pode ser representada por:

$$[D] = [R][C] \quad (3)$$

onde  $[D]$  é a matriz contendo os dados experimentais,  $[R]$  é a matriz das curvas espectrais básicas dos membros finais (ou matriz de eigenspectra) e  $[C]$  e a matrix de vetores de eigen formada pelas contribuições relativas de cada membro final.

A análise de principais componentes é inicialmente utilizada para decompor a matriz  $[D]$  em duas matrizes abstratas;  $[R]_A$  e  $[C]_A$  de tal forma que

$$[D] = [R]_A [C]_A \quad (4)$$

Matematicamente a solução da equação (3) é dada por:

$$[Z]_0 [C]_A = [S] [C]_A \quad (5)$$

Onde  $[Z]_0$  e a matriz simétrica de covariança ( $[Z]_0 = [D]^T [D]$ ) e  $[S]$  é a matriz contendo os “eigenvalues”. A matriz abstrata  $[R]_A$  é então construída de acordo com:

$$[R]_A = [D] [C]_A^T \quad (6)$$

Embora os “eigenvalues” sejam normalmente utilizados para a definição do numero de fatores ou membros finais fisicamente significativos o melhor teste para esse número é dado pela reconstrução da matriz original de dados

[D] com base na equação (1). Nesse caso a reconstrução é feita passo a passo, incorporando um fator de cada vez ao modelo. No final o número de fatores ou membros finais será aquele em que os desvios (obtidos pela diferença entre os dados originais e os reconstruídos) estejam dentro do que é estatisticamente aceitável. O número de fatores no caso desse trabalho será limitado pelo número de bandas da imagem utilizada (6 para imagens provenientes do Landsat 5 TM). A solução até esta etapa é tida como abstrata por ser meramente matemática, ou seja, nesse estágio não sabemos o significado físico dos membros finais que são importantes para serem incluídos no modelo.

Após a determinação do número de fatores, a próxima etapa é achar o significado físico desses fatores com base nas assinaturas espectrais dos alvos tidos como “puros”. Como foi dito anteriormente os alvos puros serão extraídos da própria imagem.

A localização dos alvos puros será feita em duas etapas. Na primeira etapa os pixels teoricamente “puros” serão visualizados nos diagramas de dispersão dentro do espaço bi e multidimensional. Softwares como o Erdas 8.3.1 integrados com softwares estatísticos como o S-Plus permitem uma excelente visualização dos dados nesses diagramas.

Uma vez conhecidos os pixels “puros” que se suspeita sejam relevantes para a decomposição da imagem os mesmos serão testados através de um procedimento que transforma a matriz abstrata  $[R]_A$  na matriz de significado físico ou real  $[R]_{\text{Real}}$  que contem as curvas espectrais provenientes dos pixels puros. Essa operação de transformação de  $[R]_A$  em  $[R]_{\text{Real}}$  é dada por (Malinowski e Howery, 1980):

$$[R]_{\text{real}} = [R]_A [T] \quad (7)$$

$$T_l = \sqrt{\lambda_l} [R]_A^T R_l \quad (8)$$

Onde  $[T]$  é a matriz de transformação,  $T_l$  é o vetor coluna contendo os valores de quadrados mínimos utilizado na transformação de cada um dos  $n$  membros finais e  $R_l$  é o vetor coluna contendo os dados espectrais dos membros finais que se suspeita sejam relevantes para a decomposição da matriz de dados originais  $[D]$ .

Para se determinar se a curva espectral dos pixels “puros” ou membros finais que serão testados está presente nos dados originais serão computadas as assinaturas espectrais previstas,  $R_p$ , dada por:

$$R_p = [R]_A T_l \quad (9)$$

Se cada elemento da assinatura espectral prevista se igualar ao correspondente elemento da assinatura espectral que está sendo testada (proveniente dos pixels “puros”) dentro de um erro estatisticamente admissível se conclui que o membro final suspeito se encontra, de fato, nos dados originais. Nesse caso o vetor coluna  $T_l$  passará a ser incluído na matriz de transformação  $[T]$  que converte  $[R]_A$  em  $[R]_{\text{real}}$ . O teste dos alvos (“target test” conforme Malinowski e Howery, 1980) continuará até que todos os membros finais sejam identificados.

A primeira fase desse trabalho, cujos resultados serão aqui apresentados, inclui apenas os testes do modelo teórico em um pequena janela de uma imagem do Landsat 5 TM coletada em 29 de outubro de 1995 (cena 216-66). Essa janela corresponde a 35 pixels extraídos de uma área localizada no município de Mirandiba-PE.

### Resultados Parciais

A Figura 1A apresenta a composição colorida para os 35 pixels provenientes dos dados originais. A decomposição espectral utilizando a análise fatorial mostrou, através da técnica de reconstrução progressiva dos dados originais, que três fatores são necessários para explicar a variância dos dados originais. As dispersões dos níveis de cinza considerando-se as bandas 3 (vermelho) e 4 (infravermelho próximo) do Landsat 5 TM é apresentada na Figura 2. Com base na Figura 2 foram extraídos 3 pixels, considerados potenciais membros finais (endmembers), com os quais obtivemos as transformações indicadas na equação 8. A curva espectral de cada um desses membros finais (endmembers) é apresentada na Figura 3. A curva espectral do membro final 2 (endmember 2) é característica da vegetação verde. A reconstrução da imagem original com base em um, dois e três fatores é apresentada na Figura 1 (B, C e D, respectivamente). O trabalho prosseguirá, ainda nessa etapa, buscando minimizar os erros nas transformações através da identificação dos membros finais que mais se adequam ao modelo e interpretando as suas características físicas.

### Bibliografia

- ACCIOLY, L.J. de O. **Applying spectral mixture analysis (SMA) for soil information extraction on the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) data**. Tucson-AZ, The University of Arizona, 1997. 284p. Tese de Doutorado.
- ACCIOLY, L.J. de O.; HUETE, A.R.; BATCHILY, K.. Using mixture analysis for soil information extraction from an AVIRIS scene at the Walnut Gulch Experimental Watershed-Arizona. In. **Anais do IX Simpósio Brasileiro de Sensoriamento Remoto**, Santos-SP, 1998. Publicado em CD-rom.

- FOSCHI, P.G. 1994. A geometric approach to a mixed pixel problem: Detecting subpixel woody vegetation. **Remote Sens. Environ.** 50:317-327.
- HORWITZ, H.M.; HALEPKA, R.F.; HYDE, P.D.; MORGENSTERN, J.P. Estimating the proportions of objects within a single resolution element of a multispectral scanner. p.1307-1320. *In Proceedings of 7th International Symposium on Remote Sensing of Environment*. Ann.Arbor, MI, Environmental Research Institute of Michigan. 1971.
- HUETE, A.R.; ESCADAFAL, R.. Assessment of biophysical soil properties through spectral decomposition techniques. **Remote Sen. Environ.** 35:149-159. 1991.
- HUETE, A.R.. Separation of soil-plant spectral mixtures by factor analysis. **Remote Sen. Environ.** 19:237-251. 1986.
- MALINOWSKI, E.R.. Obtaining the key set of typical vectors by factor analysis and subsequent isolation of component spectra. **Analytica Chimica Acta**, 134:129-137. 1982.
- MALINOWSKI, E.R.; HOWERY, D.G. **Factor analysis in chemistry**. John Wiley e Sons, New York, NY. 1980. 365p.
- RICHARDSON, A.J.; WIEGAND, C.L.; GAUSMAN, H.W.; CUELLAR, J.A.; GERBERMANN, A.H.. Plant, soil, and shadow reflectance components of row crops. **Photogramm. Eng. Remote Sens.** 41:1401-1407. 1975.
- SMITH, M.O.; ROBERTS, D.A.; SHIPMAN, H.M.; ADAMS, J.B.; WILLIS, S.C.; GILLESPIE, A.R. 1987. Calibrating AIS images using the surface as a reference. p. 63-69. *In Proceedings of 3rd Airborne Imaging Spectrometer Data Analysis Workshop*. JPL Publicaions. Jet Populsion Laboratory, Pasadena, CA. 2-4 June, 1987.

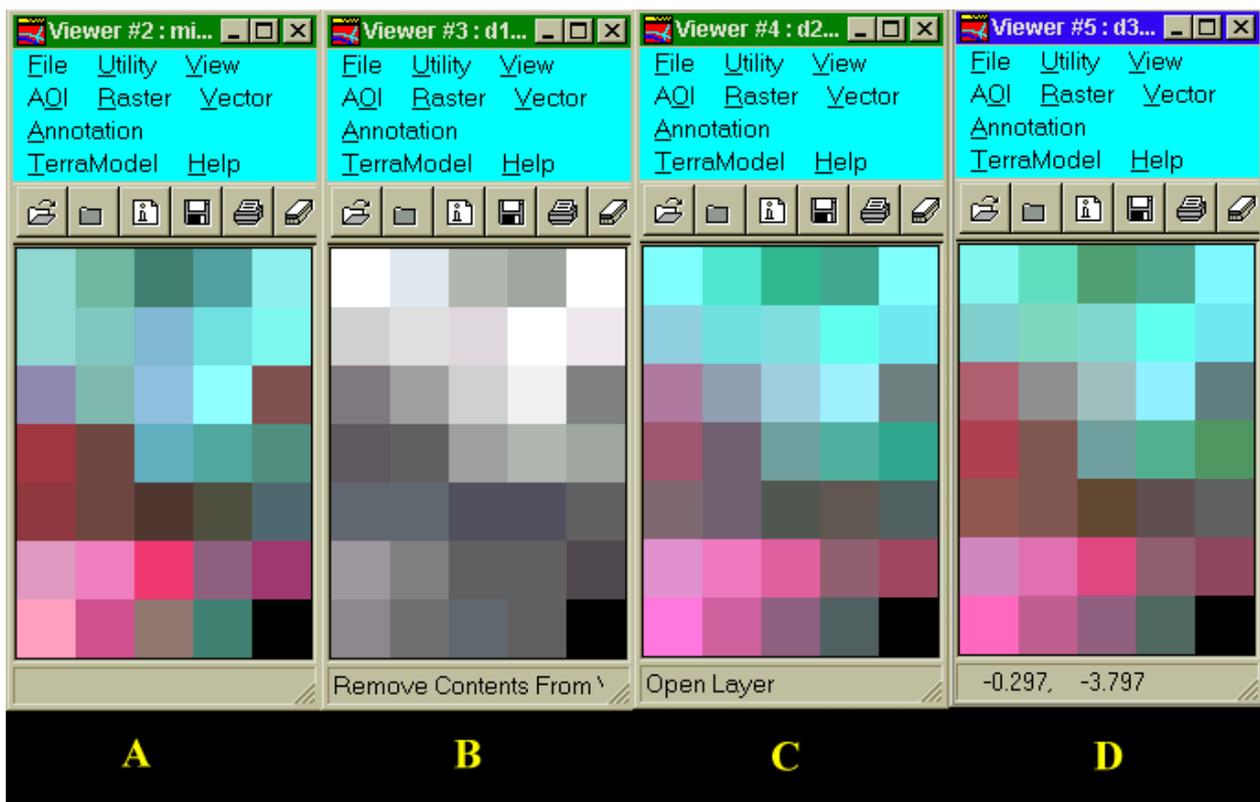


Figura 1 - Reconstrução da imagem original (A) com base em um fator (B), dois fatores (C) e três fatores (D). Composição colorida: banda 2 (azul), banda 3(verde) e banda 4 (vermelho).

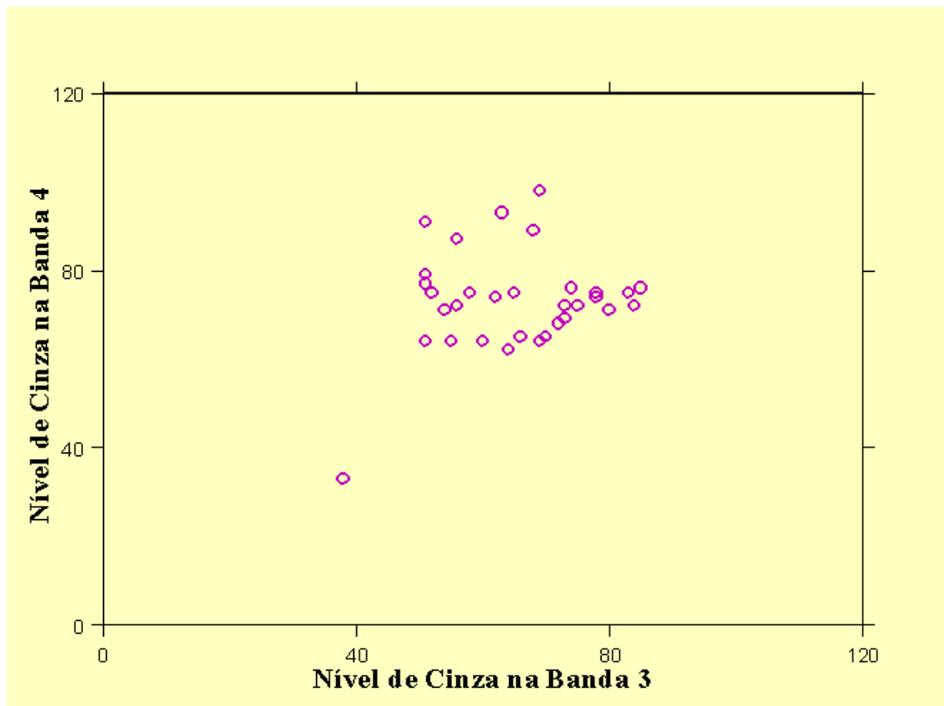


Figura 2 – Diagrama de dispersão dos níveis de cinza considerando-se as bandas 3 (vermelho) e 4 (infravermelho)

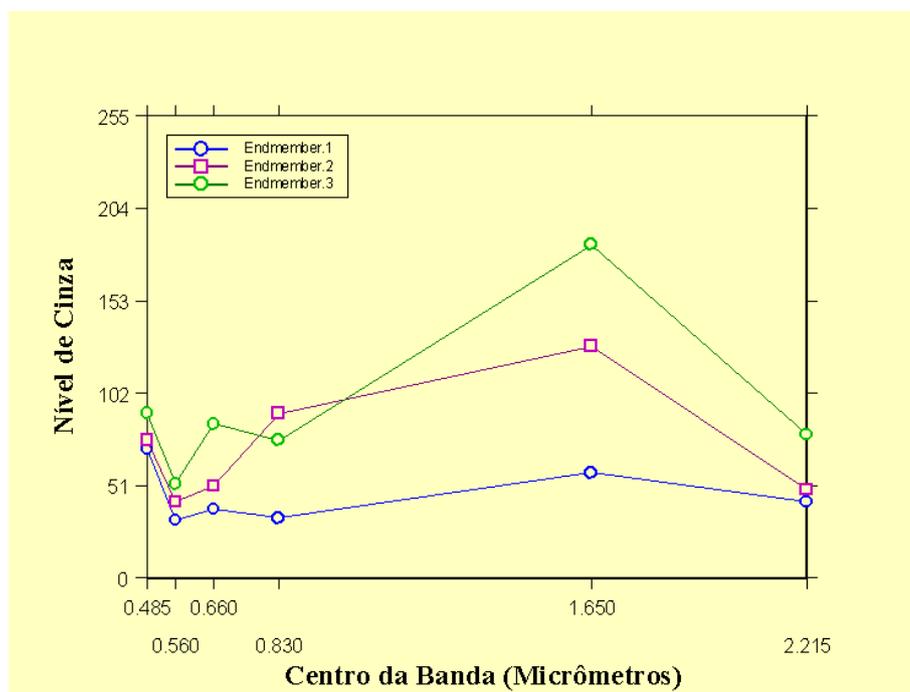


Figura 3 – Curvas espectrais de potenciais membros finais (endmembers)