

Análise de Espectros de Reflectância por Principais Componentes

LÊNIO SOARES GALVÃO*/IAG / USP, ÍCARO VITORELLO/INPE

ABSTRACT

Principal Components Analysis (PCA) are used in this study to exemplify the possibility of the quantitative evaluation of high resolution reflectance spectra. To attain such goal, a simple experiment is carried out through moisture variations in three kinds of sand material. The results confirm that the first component is associated to the albedo, and the second is related to the general slope, or more precisely to the near-infrared/visible ratio. To enhance the absorption feature information attached to the remaining components, we have used the technique of vertically scaling the spectra to unity. With this example, we have demonstrated the potential of PCA applications in a large number of problems of geologic interest such as the analysis of rock components, lithostratigraphic correlations and spectra modelling.

INTRODUÇÃO

Até o presente momento, o Sensoriamento Remoto tem se caracterizado por medidas de baixa resolução espectral ao longo da faixa 400-2500 nm. Este é o caso das imagens TM-Landsat 5, adquiridas em bandas espectrais da ordem de 60 (Banda 3) a 270 nm (Banda 7) de largura, ou mesmo dos radiômetros multibandas construídos para obter dados de reflectância em faixas espectrais aproximadamente equivalentes às imagens orbitais. Entretanto, devido ao interesse crescente de se desenvolver sistemas imageadores capazes de adquirir um maior número de imagens em bandas espectrais mais estreitas (10 a 20 nm de largura) e conseqüentemente de se reconstituir feições de absorção úteis para discriminação e/ou identificação de alvos terrestres, espectralradiômetros têm sido cada vez mais utilizados para produzir curvas de reflectância de alta resolução espectral. Por outro lado, a alta dimensionalidade do conjunto de dados destas curvas, em geral com mais de 800 pontos espectrais ou comprimentos de onda amostrados entre 400 e 2500 nm, tem resultado também numa grande quantidade de informação redundante. Desta forma, a questão básica advinda deste fato é: como eliminar tal redundância e analisar espectros de reflectância sem uma perda substancial de seu conteúdo de informação?

Neste contexto, este trabalho discute aspectos advindos de nossa experiência com a aplicação da técnica Principais Componentes (APC) para análise de espectros

de reflectância. O propósito essencial do uso desta técnica é tentar descrever as relações de covariância existentes em um conjunto de dados através de um pequeno número de fatores que preserve a maior parte do conteúdo de informação original. Detalhes sobre a fundamentação teórica e exemplos de aplicação de APC na área de geociências poderão ser encontrados em Davis (1973) e Joreskog et al. (1976). Para efeito demonstrativo de seu uso com dados de reflectância, um pequeno experimento com variação de umidade de areias de praia foi efetivado e os resultados serão brevemente discutidos.

METODOLOGIA DE TRABALHO

Suponha, como exemplo, a existência de um conjunto de 100 amostras de rochas e que para cada uma delas foi obtida uma curva de reflectância na faixa do visível e infravermelho refletido (400-2500 nm). Solicita-se, então, que sejam estabelecidas as relações de similaridade-separabilidade espectral entre estas amostras. A primeira constatação óbvia para quem vai interpretar os dados é a dificuldade de se comparar qualitativamente um tão grande número de curvas. Daí a necessidade de técnicas como Principais Componentes.

O primeiro passo para análise destes dados por tal técnica é amostrar os espectros, extraindo-se valores médios de reflectância compreendendo um pequeno número de intervalos de comprimento de onda. Estas são as denominadas variáveis "banda". Como regra, estes valores quando plotados graficamente têm que ser capazes de reconstituir de forma bastante razoável o formato geral das curvas, incluindo suas feições de absorção. Normalmente, um número de 10 a 15 variáveis "banda" calculadas em intervalos espectrais de 10 a 20 nm de largura é suficiente para caracterizar cada espectro. Estas variáveis constituirão o conjunto de dados sobre o qual será aplicada APC. Os cálculos iniciais dos programas estatísticos disponíveis normalmente envolvem a obtenção da matriz de correlação ou covariância do conjunto de dados. Ela é a matriz-base para a obtenção dos autovalores e subsequentemente dos autovetores. Os autovetores associados aos autovalores diferentes de zero podem ser vistos como um conjunto de vetores linearmente independentes, cujas posições expressam as direções de máxima variância dos dados. A variância por sua vez é dada pelos autovalores. Em outras palavras, o autovetor associado ao maior autovalor

determina a direção de máxima variância. O associado com o segundo maior autovalor localiza a direção de máxima variância ortogonal ao primeiro; e assim por diante, conforme destacaram Joreskog et al. (1976).

De nossa experiência com dados de reflectância, a obtenção de todos os possíveis autovetores é desnecessária e, como opção, pode-se analisar apenas aqueles com autovalores maiores que 1. De fato, os dois primeiros eixos PC são responsáveis por mais de 90 % da variância total dos dados. Nós temos observado que o primeiro autovetor normalmente expressa variações espectrais devidas ao albedo dos alvos, ou seja, ao fato deles possuírem uma baixa ou alta reflectância total entre 400 e 2500 nm ou, simplificarmente, de serem alvos mais escuros ou mais claros. Por outro lado, o segundo autovetor geralmente está relacionado com a "declividade espectral" das curvas, ou seja, com a variação relativa entre dados do infravermelho e visível. Os demais tendem a realçar aspectos menores das curvas, como, por exemplo, bandas de absorção. Neste contexto, para examinar o conteúdo de informação expresso por um determinado autovetor basta verificar os fatores de peso ("loading factors") dos elementos que o constituem. Desta forma, se uma variável for importante para explicar um determinado autovetor o seu elemento autovetorial correspondente terá um fator de peso maior que os verificados para as demais variáveis. Quando os elementos de um autovetor são multiplicados pelas variáveis originais correspondentes e somados no contexto de cada observação, escores PC são obtidos.

Quando o interesse é analisar de forma mais detalhada as variações de reflectância devidas às bandas de absorção, nós temos usado o artifício de escalar os espectros para a unidade, tomando como ponto-base um determinado comprimento de onda. Desta forma, a primeira componente relativa ao albedo é removida e as demais, mais diretamente associadas às feições de absorção, são realçadas.

UM EXEMPLO DE APLICAÇÃO

Dados de reflectância foram obtidos sobre três amostras de areia de praia em diferentes estágios de umidade. O espectrorradiômetro utilizado foi o IRIS/GER, que obtém dados em intervalos de 2 nm na faixa 400-1000 nm e de 4 nm na faixa 1000-2500 nm. O Fator de Reflectância de um alvo é obtido a partir da comparação entre a energia refletida pelo mesmo e a de uma placa-padrão de reflectância conhecida.

As curvas de reflectância 1, 2 e 3 na Figura 1 compreendem areias secas de coloração cinza-escuro, vermelho-escuro e amarelo-claro, respectivamente, enquanto que as curvas A, B e C indicam diferentes

estágios de umidade para cada amostra (úmida, saturada, supersaturada). Para caracterizar cada uma das 12 curvas mostradas nesta figura, valores médios de reflectância foram calculados para 11 intervalos espectrais de 10 nm de largura, cujo posicionamento está indicado na curva 2 da Figura 1b. Sobre este conjunto de variáveis "banda", APC foi efetivada.

De acordo com os autovalores resultantes, os dois primeiros eixos PC são responsáveis por cerca de 93 % da variância dos dados (82 e 11 % para PC1 e PC2, respectivamente). Analisando-se a matriz de autovetores, pôde-se observar que o autovetor 1 está relacionado com as modificações no albedo das amostras, possuindo fatores de peso aproximadamente constantes para todos os comprimentos de onda analisados. Por outro lado, o autovetor 2 exibiu fatores de peso negativos e positivos em comprimentos de onda superiores e inferiores a 1400 nm, respectivamente, expressando variações espectrais relativas introduzidas pela água sobre os dados do infravermelho e visível.

Estes aspectos podem ser melhor ilustrados através da projeção dos escores PC mostrada na Figura 2. No que se refere ao albedo, observe que à medida que as amostras se tornam mais úmidas a reflectância total das curvas diminui. Isto corresponde a um deslocamento dos pontos da direita para esquerda do gráfico da Figura 2, o que pode ser facilmente acompanhado na Figura 1. No que se refere ao eixo PC2, ele reflete as mudanças na declividade espectral das curvas pelo forte efeito espectral da água, não apenas provocando o rebaixamento relativo dos dados de reflectância no infravermelho, como também o surgimento e o subsequente desaparecimento gradual das bandas de absorção d'água em 1400 e 1900 nm. Observe na porção superior direita da Figura 2, relativa a amostra 3, a tendência da razão de reflectância 1900/500 nm, comprimentos de onda aproximadamente equivalentes aos picos negativos e positivos dos fatores de peso do autovetor 2, diminuir à medida que as amostras se tornam mais úmidas.

Para analisar de forma mais detalhada o efeito da variação de umidade sobre estas duas feições, todas as curvas foram escaladas para a unidade em 1300 nm e 10 variáveis "banda" foram extraídas ao longo da faixa 1300-2200 nm. APC foi aplicada e o exame da matriz de autovetores revelou a variabilidade relativa ao formato dos espectros escalados associada a PC1. Entretanto, os fatores de peso associados a PC2 e PC3 expressaram as variações nas profundidades das bandas de absorção em 1900 e 1400 nm, respectivamente, conforme pode ser visto também através da projeção dos escores PC ilustrada na Figura 3. Compare, por exemplo, as curvas relativas a amostra 2 na Figura 1b e o mostrado na Figura 3. Quando comparado com a curva espectral da amostra seca (curva 2), o

estágio A (úmido) provoca uma diminuição na profundidade da feição em 1900 nm, ocorrendo o contrário com a absorção em 1400 nm. Isto corresponde na Figura 3 a um deslocamento do ponto 2 da direita para esquerda do eixo PC2 e da base para o topo do eixo PC3 até o ponto A. No estágio B (saturado), apenas a banda de absorção em 1900 nm sofre uma redução mais forte em sua profundidade e o movimento do ponto A para o B relativo a amostra 2 na Figura 3 se verifica principalmente ao longo do eixo PC2 (direita para esquerda). Finalmente, no estágio C (supersaturado) a absorção em 1400 nm torna-se mais ampla e rasa e a de 1900 nm desaparece. Na Figura 3, esta descrição é concordante com o deslocamento do ponto B para o C do topo para a base em PC3 e da direita para esquerda em PC2, respectivamente. Raciocínio similar pode ser seguido para as demais amostras.

CONSIDERAÇÕES FINAIS

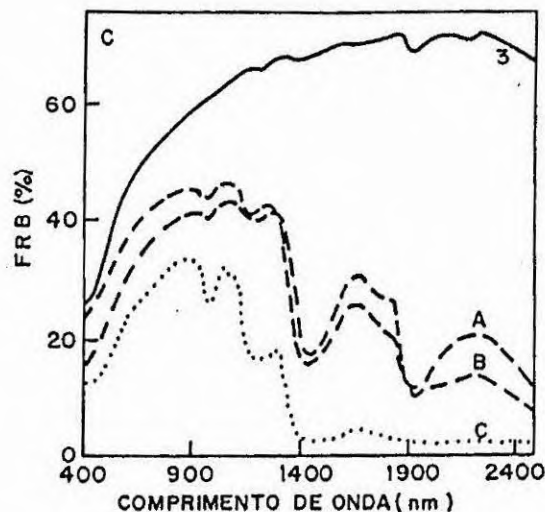
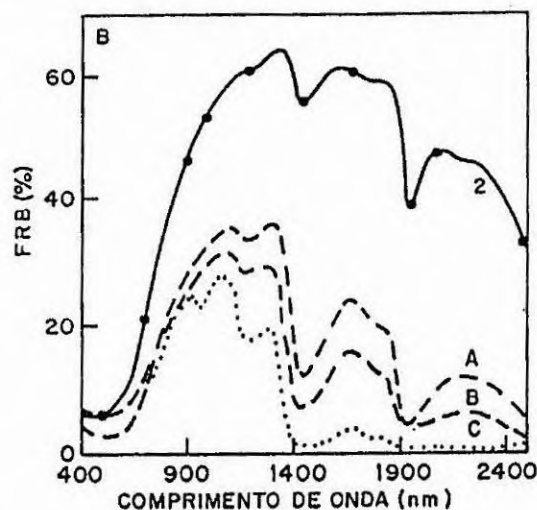
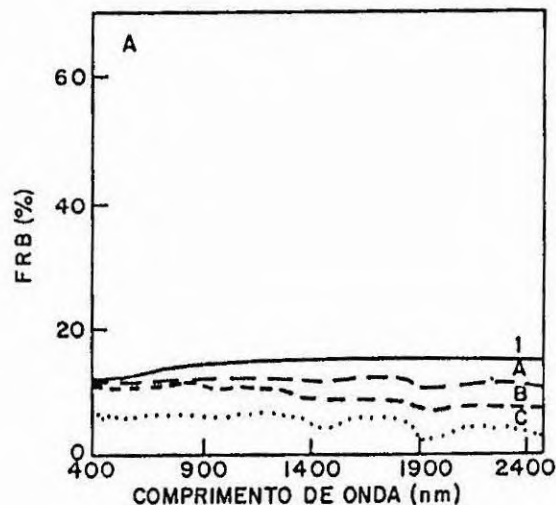
Com o auxílio da técnica Principais Componentes, nós temos analisado as relações de similaridade-separabilidade espectral de mais de trezentos espectros de reflectância de rochas e solos de diferentes ambientes geológicos. Aplicações mais específicas resultantes deste processo têm incluído o correlacionamento litoestratigráfico de seqüências de superfície e de subsuperfície (testemunhos de perfuração), bem como o estudo da influência das variações da composição química de rochas sedimentares sobre os dados de reflectância. Em todos os casos, Análise Principais Componentes tem sido uma técnica útil e muito interessante para análise de espectros.

BIBLIOGRAFIA

Davis, J.C. Statistics and data analysis in geology. New York, NY, John Wiley, 1973. 550 p.

Joreskog, K.G.; Klován, J.E.; Reyment, R.A. Geological factor analysis. New York, NY, Elsevier Scientific Publishing, 1976. 178 p.

FIGURA 1 - Fator de Reflectância Bidirecional (FRB) de areias cinza-escuro em (1a), vermelho-escuro em (1b) e amarelo-claro em (1c) em diferentes estágios de umidade. As curvas 1, 2 e 3 foram obtidas sobre amostras secas. As curvas A, B e C indicam a obtenção de dados espectrais sobre amostras úmidas, saturadas e supersaturadas, respectivamente. Na Figura (1b) estão indicadas a posição das 11 variáveis "banda" utilizadas para APC.



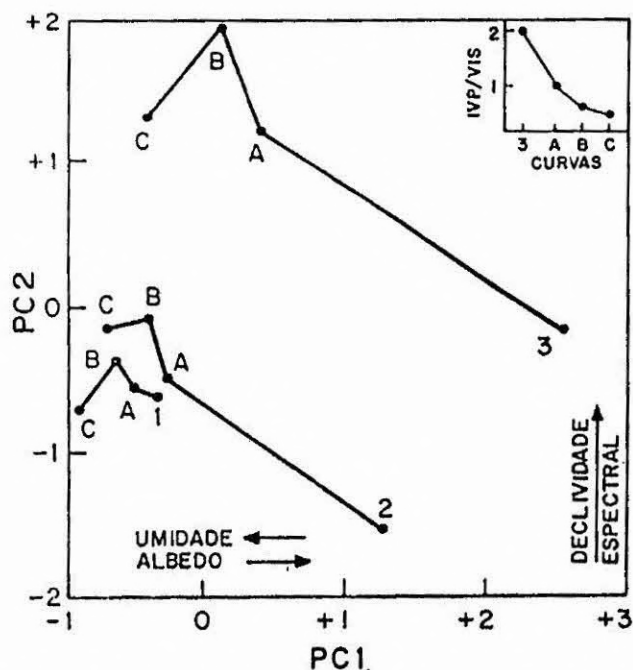


FIGURA 2 - Projeção dos escores sobre os dois primeiros eixos principais componentes obtidos a partir de 11 variáveis "banda", cujo posicionamento espectral está mostrado na Figura 1b. Dados da razão de reflectância 1900/500 nm para a amostra 3 estão ilustrados na porção direita superior do gráfico. O restante da legenda é similar ao da Figura 1.

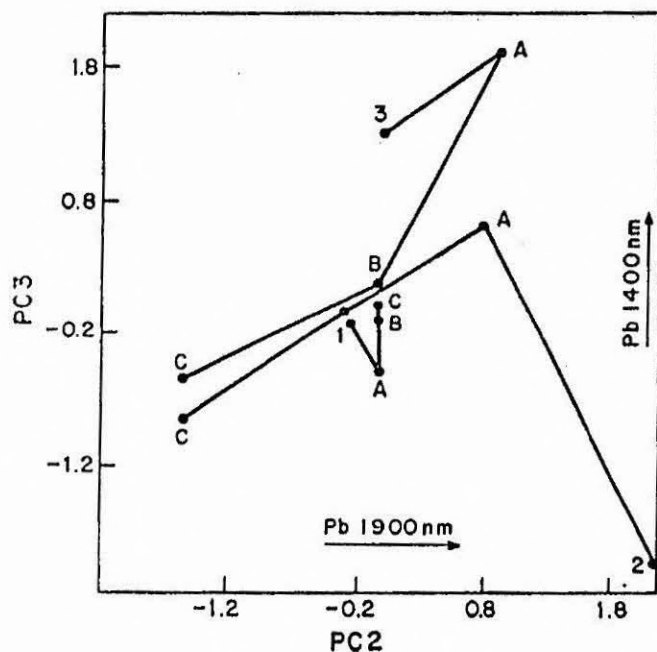


FIGURA 3 - Projeção dos escores PC2 e PC3, obtidos a partir de 10 variáveis "banda" distribuídas ao longo da faixa 1300-2200 nm. As variáveis foram calculadas após as curvas terem sido escaladas para a unidade em 1300 nm. Pb indica a profundidade das bandas de absorção. O restante da legenda é similar ao apresentado na Figura 1.