SPECTRAL VARIATIONS IN A COLLECTION OF AVIRIS IMAGERY

John C. Price

USDA Agricultural Research Service
Beltsville Agricultural Research Center
Beltsville, Maryland 20705

1. INTRODUCTION

The advent of AVIRIS presents to the scientific community the first opportunity to examine high quality hyperspectral image data over areas of interest (Vane, 1987), while recent developments promise even better data in the future. With these data we may address questions such as: 1) Is such high spectral resolution (0.01μm) necessary?, and 2) How can we take advantage of the AVIRIS experiance in specifying an improved future Landsat satellite? In this paper we describe the use of relatively broadband measurements (20-50 nm) to represent spectral variability (Price, 1990, 1991). The problem of generality is addressed through consideration of 28 AVIRIS scenes, although analysis of additional areas is clearly desirable. Section 2 defines the expansion of spectra in terms of basis functions, section 3 describes the application to AVIRIS data, and the conclusion, section 4, relates these preliminary results to possible future remote sensing instrumentation such as the planned Landsat 8.

2. DESCRIPTION BY BASIS FUNCTIONS

From inspection most visible near infrared reflectance spectra vary in a relatively smooth, continuous fashion, implying that correlations exist between measurements at nearby wavelengths. Thus a measurement in a limited spectral range provides information about values over a wider spectral range. Let $\mathbf{x}^\alpha(\boldsymbol{\lambda}) = (x_1^\alpha, x_2^\alpha, \ldots x_n^\alpha)$ represent a measured spectrum for the n = 224 wavelength values $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3, \ldots \lambda_n)$ of the AVIRIS instrument, with superscript α denoting an individual spectrum (pixel). We describe reflectance spectra, i. e. the ratio of reflected energy to the solar constant, as this tends to equalize the wavelength contributions, which are otherwise heavily weighted toward the visible region of the spectrum where reflected radiation is strongest. We represent spectra by a set of spectral basis functions $\boldsymbol{\varphi}$:

$$\mathbf{x}^\alpha \approx \sum_{i=1}^{M} S_i^\alpha \, \boldsymbol{\varphi}_i(\boldsymbol{\lambda}), \qquad (1)$$

where the $\boldsymbol{\varphi}(\lambda_1, \lambda_2, \lambda_3, \ldots \lambda_n)$ are spectral shapes obtained by statistical analysis of AVIRIS spectra, and the coefficients S_i^α are wavelength integrals relating to the original spectra $\mathbf{x}^\alpha$. Each basis function $\boldsymbol{\varphi}_i$ has an associated spectral interval $[\lambda_i(\min), \lambda_i(\max)]$ representing the domain of integration for specifying the coefficients S_i. Each φ has essentially unit value in its spectral interval (more precisely has mean value of 1.0 in this interval), then decreases according to the degree of wavelength correlation in the ensemble of measured spectra. The expansion represents successive approximations to the original spectra. Evidently the number of basis functions M which is required to describe the $\mathbf{x}^\alpha$ to within small residuals must be much less than 224, or else the expansion is not useful. Thus let $\delta\mathbf{x}_i^\alpha$ be the difference between a measured spectrum and the sum to order i of the expansion (eq. 1), and let S_i^α be the mean of $\delta\mathbf{x}_i^\alpha$ over a selected spectral interval $[\lambda_i(\min), \lambda_i(\max)]$

$$S_i^\alpha = \frac{1}{[\lambda_i(\max)-\lambda_i(\min)]} \int_{\lambda_i(\min)}^{\lambda_i(\max)} \delta\mathbf{x}_i^\alpha \, d\lambda \tag{2}$$

At the beginning $\delta\mathbf{x}_1 = \mathbf{x}$. Since S_i is correlated with the value of $\delta\mathbf{x}$ over a wider spectral interval, we define the basis function φ_i by

$$\varphi_i(\lambda) = < \delta\mathbf{x}_i S_i > / < (S_i)^2 >. \tag{3}$$

where the brackets represent an average over the ensemble of N spectra, e. g. $< \mathbf{x} > = 1/N \sum_{\alpha=1}^{N} \mathbf{x}^\alpha$. From the definition of S, the normalization of φ is given by

$$\frac{1}{[\lambda_i(\max)-\lambda_i(\min)]} \int_{\lambda_i(\min)}^{\lambda_i(\max)} < \delta\mathbf{x}_i S_i > / < (S_i)^2 > d\lambda = 1 \tag{4}$$

so that φ has a mean value of one within the integration interval. At each iteration level i the residual vector $\delta\mathbf{x}_i^\alpha$ is approximated by $\delta\mathbf{x}_i^\alpha \approx S_i^\alpha \varphi_i(\lambda)$, leaving a new residual $\delta\mathbf{x}_{i+1}^\alpha$. Then the procedure moves to $\delta\mathbf{x}_{i+1}$. From the definition $\delta\mathbf{x}_i$ and all higher order residuals have the value 0 somewhere within the wavelength interval $[\lambda_i(\min), \lambda_i(\max)]$. Thus successive residuals $\delta\mathbf{x}_i$ pass through zero at more and more wavelength values as the order of the expansion increases, and the magnitude of the residuals $\int(\delta\mathbf{x})^2 d\lambda$ decreases. The criterion for terminating the expansion of basis functions is been based on the examination of the residuals. Let percent error R, given M basis functions, be defined by

$$R(M) = 100\% \cdot \left< \int \left(x - \sum_{i=1}^{M} s_i \varphi_i\right)^2 d\lambda \right> / \left< \int (x^2)\, d\lambda \right> \qquad 15)$$

where integration extends over the AVIRIS wavelength interval. The residual R provides an indication of the remaining signal plus noise after subtraction of the basis function series. Given enough basis functions the residual is dominated by instrument noise. Thus at a mean signal to noise ratio of 30:1, the residual R is $(1/30)^2 = 1/900 \approx 0.1\%$. The description of spectra in terms of reflectance as opposed to radiance greatly increases the importance to longer wavelengths where the signal to noise of AVIRIS is poorer.

3. APPLICATION TO AVIRIS SPECTRA

Table 1 lists the data sets used in these calculations, and the general type of scene. The selection attempts to span the natural variability represented throughout the AVIRIS data sets, with the exception of clouds. Only after the fact was it noticed that #17 and 27 cover largely the same area. Scenes were selected using the quick look data at the Jet Propulsion Laboratory. I am indebted to the staff for their cooperation.

Table 1. **AVIRIS Data Scenes and Identifiers**

	Flight	Run	Scene	Catalog name	type of scene
1	920602A	9	8	Moffett Field	suburb, shallow water
2	920826B	3	2	Maricopa farm	agriculture
3	920828B	13	1	Los Alamos	geology, town
4	920827B	2	5	Rodgers Dry Lake	geology
5	921119B	9	8	Tampa Bay	city, water
6	920603B	2	3	Cuprite	geology
7	920826B	5	1	Camp Pendleton	water, military base
8	920615B	2	3	Harvard Forest	forest
9	920612B	2	5	Indian Pines	agriculture, forest
10	920531C	6	1	Death Valley	geology
11	920603B	14	3	Cima volcanic field	geology
12	920708B	1	2	Gainesville, FL	lake, vegetation
13	920819B	2	1	Denver	suburb, agriculture
14	920828B	2	5	San Juan Mtns	snow, geology
15	920602A	6	2	Jasper Ridge	suburb, vegetation
16	920826B	4	1	Fort Huachuca	a.f.base, geology
17	920616B	2	1	Spruce forest	forest, clear cuts
18	920820B	6	1	Pleasant Grove	agriculture
19	921117D	2	30	Jackson, TN	agriculture, forest
20	921119B	5	4	Tampa Bay	island, shallow water
21	920826B	6	2	San Joaquin	agriculture
22	920531C	2	1	Owens Valley	geology
23	920820B	7	1	Dunnigan, CA	agriculture
24	920819B	10	14	San Berdardino	agriculture
25	920708B	5	1	Gainesville, FL	town, vegetation
26	920621C	2	1	Blackhawk Island	ag, forest, water
27	920615B	8	1	Spruce forest	forest, clear cuts
28	920820B	8	3	Davis, Webster	agriculture, town

Since 28 AVIRIS images represent 4 gigabytes of data, some care was needed in devising an efficient analysis strategy. First a scene showing considerable variability in surface types (Moffet Field) was analyzed. It was found that 9 variables (spectral intervals) described the signal variation to within 0.1% (Table 2, column 1). Then a 1% sample of the 28 scenes (every 100th pixel) was analyzed, with results as shown in column 2 of Table 2. Finally, the basis function expansion to level 20 was used to select "bad" pixels, with the worst 1% being saved from each scene. Then these were added to the original 1% sample, so that the effect of poorly described spectra was exaggerated by a factor of 50. Even this requires only approximately 20 basis functions to describe the AVIRIS signal very well.

Table 2. **Variables Required to Meet Varying Accuracies**

Accuracy (1-R)	# variables Moffet field (1 scene)	# variables 1% sample (28 scenes)	#variables 1% sample +1% bad (28 scenes)
90%	1	1	1
99%	3	3	4
99.9%	9	8	12
99.95%	-	15	19
99.96%	-	20	23
99.97%	-	28	31

Finally, each of 34 coefficient images (S_i) for each of the 28 scenes was studied visually. Small variance images (i>25) still showed signals at a level of a few tenths of a percent. Possible explanations include instrument noise, spectral misregistration, broad atmospheric variations between images (aersols), and true scene to scene variability. Little or no evidence was found for isolated surface types with extraordinary spectral features.

4. CONCLUSION AND OUTLOOK

From examination of 28 scenes it appears that approximately 20-25 measurements are adequate to define the spectral variability of the 20 meter data from AVIRIS. This result will simplify treatment of atmospheric effects, assist in the identification of remotely sensed spectra from a spectral library, and may guide the design of future instrumentation, such as the number of spectral channels to consider for Landsat 8.

5. REFERENCES

Vane, G. T., "First Results of the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS), SPIE, Vol 834 Imaging Spectroscopy, 166 (1987)

Price, J. C., "Information content of soil spectra", Remote Sensing Environment, 33, 113 (1990).

Price, J. C., "Variability of high resolution crop reflectance spectra", Int. J. of Remote Sensing, 13, 2593 (1992).