\2

A NEW AND FAST METHOD FOR SMOOTHING SPECTRAL IMAGING DATA

Bo-Cai Gao*, Ming Liu**, and Curtiss O. Davis*

*Code 7212, Naval Research Laboratory, Washington, DC 20375
**SAIC, Chantilly,VA 22021

1. INTRODUCTION

The Airborne Visible Infrared Imaging Spectrometer (AVIRIS) acquires spectral imaging data covering the 0.4 - 2.5 μm wavelength range in 224 10-nm-wide channels from a NASA ER-2 aircraft at 20 km. More than half of the spectral region is affected by atmospheric gaseous absorption. Over the past decade, several techniques (Goetz *et al.*, 1997) have been used to remove atmospheric effects from AVIRIS data for the derivation of surface reflectance spectra. An operational atmosphere removal algorithm (ATREM) (Gao *et al.*, 1993), which is based on theoretical modeling of atmospheric absorption and scattering effects, has been developed and updated (Gao and Davis, 1997) for deriving surface reflectance spectra from AVIRIS data. Figure 1 shows an example of a reflectance spectrum derived with ATREM from AVIRIS data acquired over Cuprite, Nevada in June, 1995. Due to small errors in assumed wavelengths and errors in line parameters compiled on the HITRAN database, small spikes (particularly near the centers of the 0.94- and 1.14-μm water vapor bands) are present in this spectrum. Similar small spikes are systematically present in entire ATREM output cubes. These spikes have distracted geologists who are interested in studying surface mineral features. A method based on the "global" fitting of spectra with low order polynomials or other functions for removing these weak spikes has recently been developed by Boardman (this volume). In this paper, we describe another technique, which fits spectra "locally" based on cubic spline smoothing, for quick post processing of ATREM apparent reflectance spectra derived from AVIRIS data. Results from our analysis of AVIRIS data acquired over Cuprite mining district in Nevada in June of 1995 are given. Comparisons between our smoothed spectra and those derived with the empirical line method are presented.

2. METHODOLOGY

In order to describe our smoothing technique, we first describe the commonly used cubic spline "fitting" technique, then we decribe the cubic spline "smoothing" technique.

2.1 Cubic Spline Fitting

The cubic spline fitting technique is a powerful numerical method and has been widely used in engineering and scientific computing. For example, Numerical Recipes (Press *et al.*, 1989) provides standard subroutines, using cubic spline fitting method, for interpolating data between points. In order to describe mathematically the cubic spline fitting technique, we consider an interval $a \le x \le b$, and subdivide it by a mesh of points corresponding to the location of the data at $a = X_0 < X_1 < ... < X_{j-1} < X_j ... < X_J = b$. An associated set of the observed data is prescribed by $y_0, y_1, ..., y_j, ..., y_J$. We seek an interpolating function *h(x)*, which is defined in the interval $[a,b]$. Its first and second derivatives are continuous on $[a,b]$ and it coincides

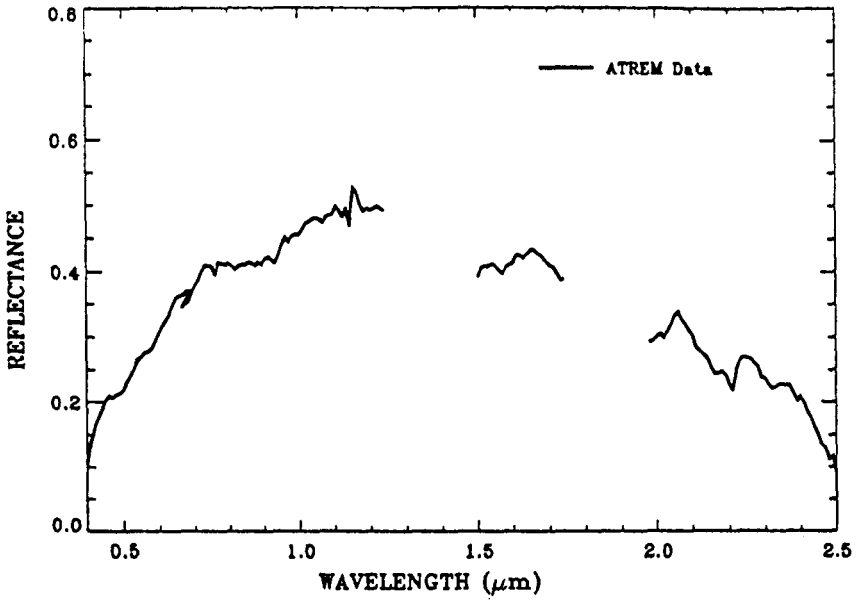

Fig. 1. An example of a reflectance spectrum derived with ATREM from AVIRIS data acquired over Cuprite, Nevada in June, 1995.

with a cubic polynomial in each subinterval $X_{j-1} \le x \le X_j$, and satisfies the relationship $h_j = h(X_j) = y_j$. Figure 2 illustrates the function *h(x)*. As adapted from Ahlberg (Ahlberg *et al.*, 1967), the function *h(x)* in the interval $X_{j-1} \le x \le X_j$ can be expressed as (for convenience, we assume the problem of equally spaced samples with a step size of Δ):

$$h(x) = \begin{cases} H_1(x) & X_0 \le x \le X_1 \\ \vdots & \vdots \\ H_j(x) & X_{j-1} \le x \le X_j \\ \vdots & \vdots \\ H_J(x) & X_{J-1} \le x \le X_J \end{cases}$$

where

$$H_j(x) = s_{j-1}\frac{(X_j - x)^3}{\Delta} + s_j\frac{(x - X_{j-1})^3}{\Delta} + \left[h_{j-1} - s_{j-1}\right]\frac{(X_j - x)}{\Delta} + \left[h_j - s_j\right]\frac{(x - X_{j-1})}{\Delta} \qquad (1)$$

$$X_j = X_0 + j\Delta$$

$$h_j = h(X_j); j = 0,1,2,..J$$

$\{s_j\}$, the spline coefficients, can be interpreted as the normalized second derivatives.

The polynomials (1) in adjacent segments are continuous at the knots:

$$H_j(X_j) = h_j = H_{j+1}(X_j).$$

The first derivative is continuous at the knot provided that

$$s_{j-1} + 4s_j + s_{j+1} = h_{j-1} - 2h_j + h_{j+1}. \qquad (2)$$

The second derivative is continuous at the knots

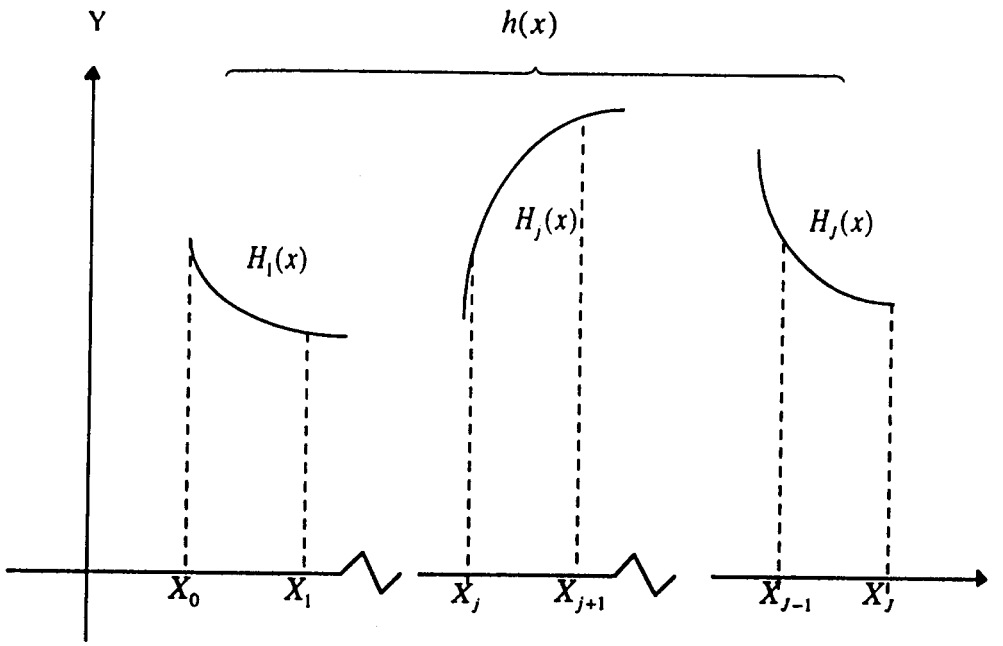

Figure 2: An illustration of the interpolating function *h(x)* .

$$H_j''(X_j) = H_{j+1}''(X_j) = \frac{6}{\Delta^2} s_j.$$

The polynomials (1) are determined by specification of $\{s_j\}$. The selection of these spline coefficients can involve any number of imposed weak constraints that characterize the spline fitting. One of the constraints is the minimization of the second derivative. Because

$$\int_{X_{j-1}}^{X_j} \left[H''(x)\right]^2 dx = \frac{12}{\Delta^3}\left[s_{j-1}^2 + s_j^2 + s_{j-1}s_j\right],$$

it follows that the quantity $\left[s_{j-1}^2 + s_j^2 + s_{j-1}s_j\right]$ is to be minimized in any kind of variational selections of $\{s_j\}$.

The simplest quadratic form to minimize is

$$\sum_{j=1}^{J}\left[s_{j-1}^2 + s_j^2 + s_{j-1}s_j\right].$$

However, this is not enough to guarantee continuity of the derivatives at the knots. A method to incorporate the condition

$$s_{j-1} + 4s_j + 4s_{j+1} - \left(h_{j-1} - 2h_j + h_{j+1}\right) = 0$$

must be found. This is done by introducing Lagrangian multipliers.

The simple spline formulation for the minimization is

$$E\left[\{s_j\}\right] = \sum_{j=1}^{J}\left[s_{j-1}^2 + s_j^2 + s_{j-1}s_j\right] + 2\sum_{j=1}^{J-1}\lambda_j\left[s_{j-1} + 4s_j + s_{j+1} - \left(h_{j-1} - 2h_j + h_{j+1}\right)\right] \tag{3}$$

where λ_js are the Lagrangian multipliers. These conditions are exactly satisfied upon completion of the minimization so that zeros are in effect added to the quantity to be minimized.

The procedure of solving those λ_js, therefore, the spline coefficients $\{s_j\}$, and the interpolating splines $\{h_j\}$, is similar to that of spline smoothing to be described in the next section.

2.2 Cubic Spline Smoothing

In the Spline fitting technique described above, the $\{h_j\}$ are taken to represent errorless data or observations, and the spline passes each point y_j. However, there can be circumstances that the observations are contaminated and unwanted noises are present. For example, in our case raw spectra exhibit coherent sawtooth like "noises". Under these circumstances, the data integrity condition should be relaxed. This can be done by adding a weak constraint term, $\sum_{j=0}^{J}\left[h_j - y_j\right]^2$ to Eq. (3), where y_j is the observed data, and only a "best fit" should be sought. The smoothed spline $\{h_j\}$ does not necessarily pass original observed data $\{y_j\}$, unlike the case of the spline fitting. An appropriate discrepancy sum can be formed as

$$\begin{aligned} E\left[\{s_j\},\{h_j\}\right] = \quad & \tau^2\sum_{j=1}^{J}\left[s_{j-1}^2 + s_j^2 + s_{j-1}s_j\right] + \sum_{j=0}^{J}\left[h_j - y_j\right]^2 \\ & +2\sum_{j=1}^{J-1}\lambda_j\left[s_{j-1} + 4s_j + s_{j+1} - \left(h_{j-1} - 2h_j + h_{j+1}\right)\right], \end{aligned} \tag{4}$$

where τ^2 is an adjustable weighting factor. As it increases, the tension of the spline smoothing increases, *i.e.*, the curve "flattens out". On the other hand, as it decreases, the observed data are reproduced more closely at the expense of increased curvature.

The variations on the spline coefficients are tabulated as:

$$\begin{array}{lcll} \delta s_0 & \tau^2(2s_0 + s_1) + 2\lambda & = 0 & \\ \delta s_1 & \tau^2(s_0 + 4s_1 + s_2) + 2\lambda_1 + 8\lambda_2 + 2\lambda_3 & = 0 & \\ \vdots & \vdots & \vdots & \\ \delta s_j & \tau^2(s_{j-1} + 4s_j + s_{j+1}) + 2\lambda_{j-1} + 8\lambda_j + 2\lambda_{j+1} & = 0 & j = 2,\ldots,(J-2) \\ \vdots & \vdots & \vdots & \\ \delta s_{J-1} & \tau^2(s_{J-2} + 4s_{J-1} + s_J) + 2\lambda_{J-2} + 8\lambda_{J-1} & = 0 & \\ \delta s_J & \tau^2(s_{J-1} + s_J) + 2\lambda_{J-1} & = 0 & \end{array} \tag{5}$$

The variations on the multipliers lead to

$$\delta\lambda_j: \quad s_{j-1} + 4s_j + s_{j+1} = h_{j-1} - 2h_j + h_{j+1}; \quad j = 1,(J-1) \tag{6}$$

Since the spline does not pass the data $\{y_j\}$, the $\{h_j\}$ are no longer fixed; their variations are listed below:

$$
\begin{array}{lcll}
\delta h_0 & 2(h_0 - y_0) - 2\lambda_1 & = 0 & \\
\delta h_1 & 2(h_1 - y_1) + 4\lambda_1 - 2\lambda_2 & = 0 & \\
\vdots & \vdots & \vdots & \\
\delta h_j & 2(h_j - y_j) - 2\lambda_{j-1} + 4\lambda_j - 2\lambda_{j+1} & = 0 & j = 2,\ldots,(J-2) \\
\vdots & \vdots & \vdots & \\
\delta h_{J-1} & 2(h_{J-1} - y_{J-1}) - 2\lambda_{J-2} + 4\lambda_{J-1} & = 0 & \\
\delta h_J & 2(h_J - y_J) - 2\lambda_{J-1} & = 0 &
\end{array}
\tag{7}
$$

Combining terms in (7), we have

$$
\left[h_{j-1} - 2h_j + h_{j+1}\right] - \left[y_{j-1} - 2y_j + y_{j+1}\right] = \lambda_{j-2} - 4\lambda_{j-1} + 6\lambda_j - 4\lambda_{j+1} + \lambda_{j+2} \tag{8}
$$

Each of the combinations $\left[s_{j-1} + 4s_j + s_{j+1}\right]; j = 1,\ldots,(J-1)$ in (5) can be replaced by their equivalents from (6) to obtain the following equations,

$$
\begin{array}{lcl}
\delta s_1: & \tau^2(h_0 - 2h_1 - h_2) + 8\lambda_1 + 2\lambda_2 & = 0 \\
\delta s_j: & \tau^2(h_{j-1} - 2h_j - h_{j+1}) + 2\lambda_{j-1} + 8\lambda_j + 2\lambda_{j+1} & = 0 \\
\delta s_{J-1}: & \tau^2(h_0 - 2h_1 - h_2) + 2\lambda_{J-2} + 8\lambda_{J-1} & = 0
\end{array}
\tag{9}
$$

The $\left[h_{j-1} - 2h_j + h_{j+1}\right]$ in (9) can be replaced by the groupings in (8)

$$
\begin{array}{lcll}
\delta s_1 & \tau^2\left[\Delta^2 \aleph^2 y_1 + (6\lambda_1 - 4\lambda_2 + \lambda_3)\right] + 8\lambda_1 + 2\lambda_2 & = 0 & \\
\delta s_2 & \tau^2\left[\Delta^2 \aleph^2 y_2 + (-4\lambda_1 + 6\lambda_2 - 4\lambda_3 + \lambda_4)\right] + 2\lambda_1 + 8\lambda_2 + 2\lambda_3 & = 0 & \\
\vdots & \vdots & \vdots & \\
\delta s_j & \tau^2\left[\Delta^2 \aleph^2 y_j + (\lambda_{j-2} - 4\lambda_{j-1} + 6\lambda_j - 4\lambda_{j+1} + \lambda_{j+2})\right] + 2\lambda_{j-1} + 8\lambda_j + 2\lambda_{j+1} & = 0 & j = 3,\ldots,(J-3) \\
\vdots & \vdots & \vdots & \\
\delta s_{J-2} & \tau^2\left[\Delta^2 \aleph^2 y_{J-2} + (\lambda_{J-4} - 4\lambda_{J-3} + 6\lambda_{J-2} - 4\lambda_{J-1})\right] + 2\lambda_{J-3} + 8\lambda_{J-2} + 2\lambda_{J-1} & = 0 & \\
\delta s_{J-1} & \tau^2\left[\Delta^2 \aleph^2 y_{J-1} + (6\lambda_{J-3} - 4\lambda_{J-2} + \lambda_{J-1})\right] + 2\lambda_{J-2} + 8\lambda_{J-1} & = 0 &
\end{array}
\tag{10}
$$

where $\aleph^2 y_j = \dfrac{y_{j-1} - 2y_j + y_{j+1}}{\Delta^2}$

The $\{\lambda_j\}$ are then found as solutions of

$$
\mathbf{A}\,\lambda = \mathbf{y}'' \tag{11}
$$

where **A** is the pentadiagonal matrix

$$\begin{pmatrix} c & b & a & & & & & & \\ b & c & b & a & & & & & \\ a & b & c & b & a & & & & \\ & a & b & c & b & a & & & \\ & & a & b & c & b & a & & \\ & & & . & . & . & . & . & \\ & & & & . & . & . & . & . \\ & & & & & . & . & . & . \\ & & & & & & a & b & c & b & a \\ & & & & & & & a & b & c & b \\ & & & & & & & & c & b & c \end{pmatrix} \tag{12}$$

Here the matrix elements are constants and given by

$$\begin{aligned} a &= \tau^2 \\ b &= (2-4\tau^2) \\ c &= (8+6\tau^2) \end{aligned}$$

and

$$y'' = -\Delta^2\tau^2 \begin{pmatrix} \aleph^2 y_1 \\ \aleph^2 y_2 \\ . \\ \aleph^2 y_j \\ . \\ \aleph^2 y_{J-1} \end{pmatrix} \tag{13}$$

Given the $\{\lambda_j\}$, the $\{h_j\}$ can be determined directly from (8), and the spline coefficients $\{s_j\}$ can be found as a tridiagonal solution of (5).

3. POST PROCESSING OF ATREM REFLECTANCE CUBE

The small spikes, as seen in Fig. 1, are systematically present in all spectra in an ATREM output data cube. We hope to make mild "gain" adjustments to remove these small spikes during our post processing of the ATREM data cube. Specifically, we hope to find a gain function, $g(\lambda)$, which contains all the weak spikes and which has a mean value close to 1. The multiplication of $g(\lambda)$ to the ATREM output spectra should allow the removal of the systematic small spikes. Several steps are involved in our post processing of an ATREM output data cube. They are:

a). The cubic spline smoothing technique described in Section 2 is applied to each of the spectra in the ATREM data cube. As a result, an intermediate "smoothed" data cube is produced. Because the cubic spline smoothing technique fits a spectrum "locally", most of the smoothed spectra at this stage match quite well with the ATREM spectra. If the spectra were fit with low order Legendre polynomials "globally", only a minor fraction of the smoothed spectra would match well with the ATREM spectra.

b). The average reflectance, ρ_{avg}, is calculated for each of the spectra in the ATREM output data cube.

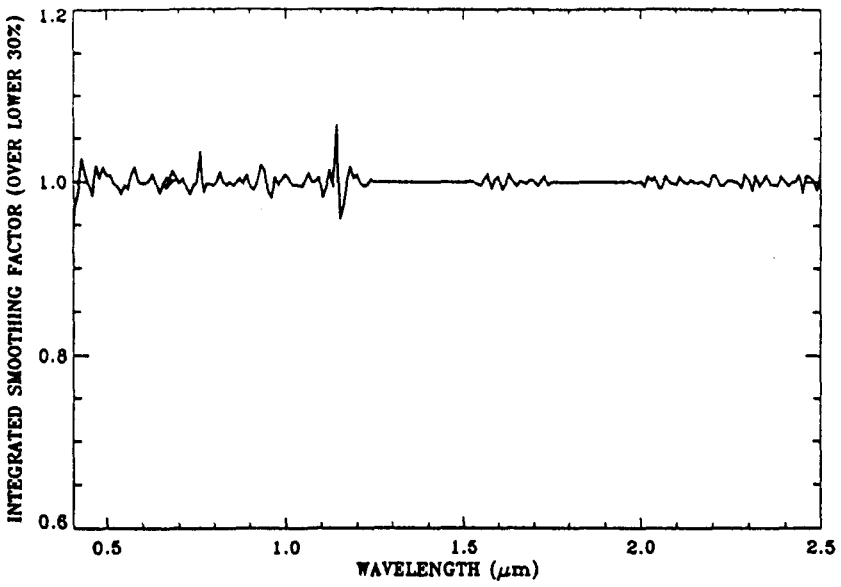

Fig. 3. A sample gain spectrum.

c). For each pixel, the standard deviation, σ, between the ATREM spectrum and the "smoothed" spectrum is calculated.

d). For an AVIRIS scene, a scatter plot of σ / ρ_{avg} v.s. ρ_{avg} is made. Pixels with σ / ρ_{avg} values in the lower twenty percentile are identified.

e). For each of the pixels identified in Step d, a ratio spectrum ("smoothed" spectrum / ATREM spectrum) is calculated. Our desired gain spectrum, $g(\lambda)$, is obtained by averaging all the ratio spectra. Figure 3 shows an example of a gain spectrum, which contains a number of weak spikes in the 0.4 - 2.5 μm spectral region.

f). The gain spectrum is applied to each of the spectra in the ATREM output data cube to obtain the "final" smoothed data cube.

Our algorithm for smoothing the ATREM output data cube is fast. It takes about 3 minutes on a SGI machine with a 150 MHz processor to process one complete data cube.

4. PRELIMINARY RESULTS

Results from one set of AVIRIS data acquired over the Cuprite Mining District in Nevada in June, 1995 are described below. Figure 4 shows a comparison among an ATREM reflectance spectrum over a single pixel , the smoothed spectrum, and the reflectance spectrum obtained with the well known empirical line method (Conel *et al.*, 1987). For clarity, the spectra in Fig. 4 are vertically displaced. The general shapes of these spectra in the 0.4 - 1.26 μm, 1.5 - 1.75 μm, and 2.0 - 2.5 μm wavelength intervals are very similar. Major mineral features in the 2.0 - 2.5 μm region are seen in all the spectra. The un-smoothed ATREM spectrum has quite a few weak spikes. These spikes are largely removed in the smoothed spectrum. The spectrum derived with the empirical line method shows weak inverse water vapor features near 0.94 and 1.14 μm. This indicates that the method results in a slight over-correction of atmospheric water vapor absorption effects for this pixel.

Figure 5 shows six ATREM reflectance spectra (vertically displaced for clarity). These spectra have distinct mineral absorption features in the 2.0 - 2.5 μm spectral region. Weak spikes (for example near 1.14 μm) are systematically present in all the spectra. Figure 6 shows the corresponding smoothed spectra, which look very similar to laboratory-measured reflectance spectra, particularly in the 2.0 - 2.5 μm spectral region. Weak spikes are all removed. A broad iron feature near 0.9 μm is seen nicely in one spectrum - the 4th spectrum from top. Figure 7

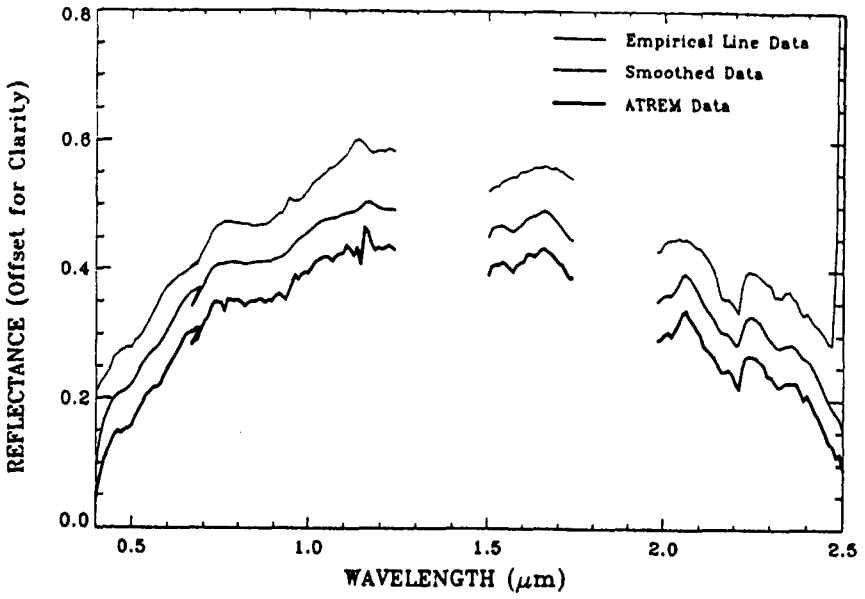

Fig. 4. A comparison among an ATREM reflectance spectrum, a smoothed spectrum, and a reflectance spectrum obtained with the empirical).

shows six spectra derived from the AVIRIS data with the empirical line method. Mineral features in the 2.0 - 2.5 μm region are recovered quite well with this method. However, water vapor features in the 0.94 and 1.14 μm regions are either over- or under- corrected. The broad iron feature in the 4th spectrum from the top is not clearly seen due to the over-correction of atmospheric water vapor absorption effects. By comparing Figures 5, 6 and 7, it is seen that major mineral features are preserved during our smoothing of the ATREM reflectance cube.

5. SUMMARY

We have described a technique, which fits spectra "locally" based on cubic spline smoothing, for quick post processing of apparent reflectance spectra derived from AVIRIS data using the ATREM code. Results from our analysis of AVIRIS data acquired over Cuprite mining district in Nevada in June of 1995 are presented. Very good agreement between our results and those of empirical line method in the 2.0 - 2.5 μm spectral region is obtained. It appears that the use of ATREM code for retrieving surface reflectance spectra from AVIRIS data plus the use of the smoothing technique described in this paper should yield surface reflectance spectra that look very similar to laboratory-measured reflectance spectra.

6. ACKNOWLEDGMENTS

The authors are grateful to J. W. Boardman of Analytical Imaging Geophysics, Boulder, Colorado for useful discussions, and to R. O. Green of Jet Propulsion Laboratory and Kathy Heidebrecht of Center for the Study of Earth from Space, University of Colorado at Boulder for providing AVIRIS data used in this study. This research is partially supported by a grant from NASA Headquarters in Washington, DC to Naval Research Laboratory in Washington, DC.

7. REFERENCES

Ahlberg, J. H., E. N. Nilson, and J. L. Walsh: The Theory of Splines and their Applications. Academic Press, New York, 9-13, 1967.

Conel, J. E., R. O. Green, G. Vane, C. J. Bruegge, R. E. Alley, and B. Curtiss, Airborne imaging spectrometer-2: Radiometric spectral characteristics and comparison of ways to compensate for the atmosphere, in *SPIE Proceedings, Vol. 834*, 140-157, 1987.

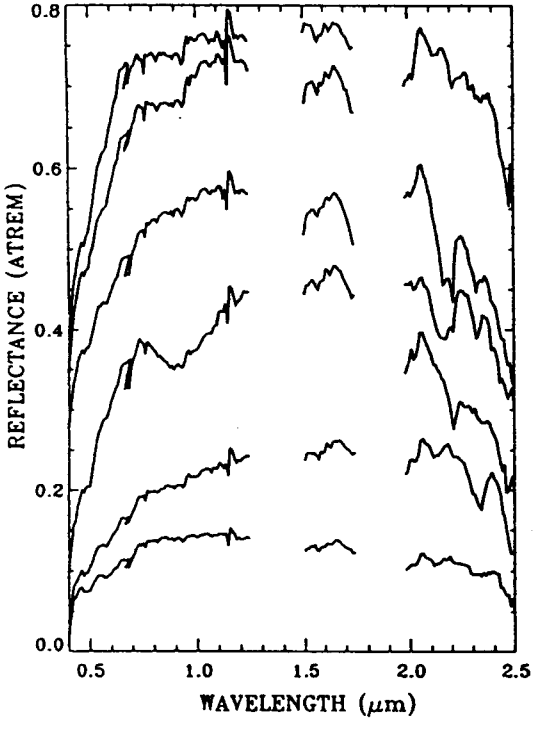

Fig. 5. Six reflectance spectra (displaced vertically for clarity) derived with ATREM from AVIRIS data acquired over Cuprite, Nevada in June, 1995.

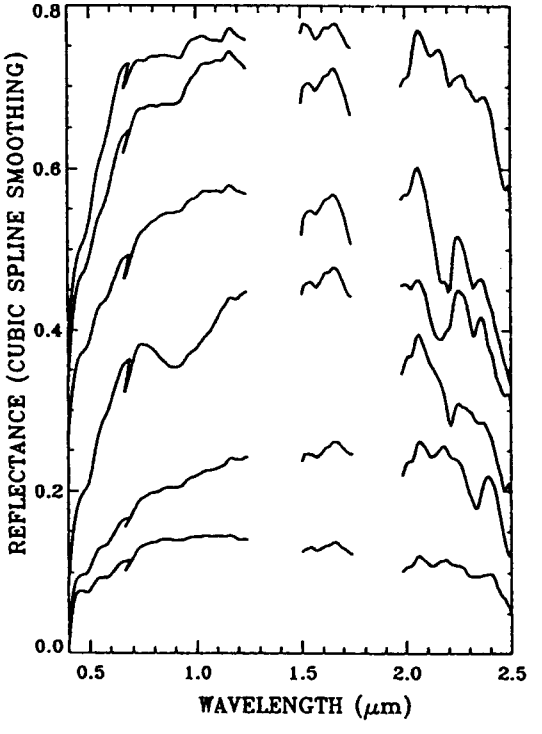

Fig. 6. Six smoothed reflectance spectra corresponding to those in Fig. 5.

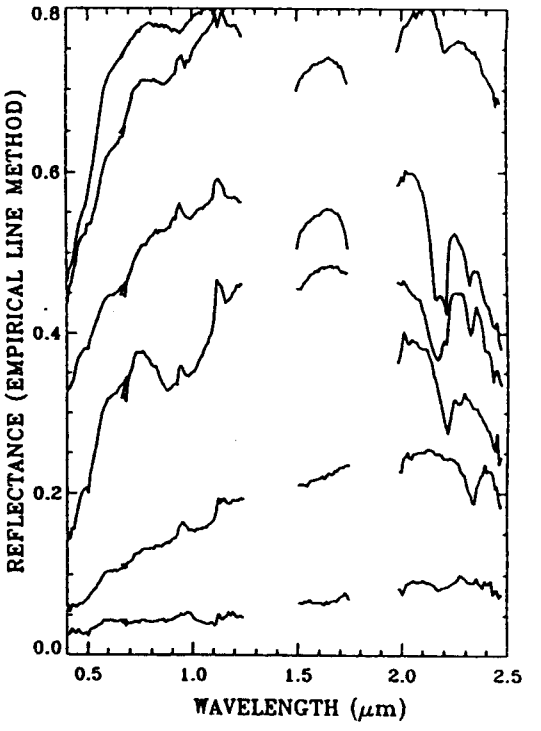

Fig. 7. Six reflectance spectra derived with the empirical line method.

Gao, B.-C., K. H. Heidebrecht, and A. F. H. Goetz, Derivation of scaled surface reflectances from AVIRIS data, *Remote Sens. Env., 44*, 165-178, 1993.

Gao, B.-C., and C. O. Davis, Development of a line-by-line-based atmosphere removal algorithm for airborne and spaceborne imaging spectrometers, in*SPIE Proceedings*, *Vol. 3118*, 132-141, 1997.

Goetz, A. F. H., J. W. Boardman, B. Kindel, and K. B. Heidebrecht, Atmospheric corrections: On deriving surface reflectance from hyperspectral imagers, , in*SPIE Proceedings*, *Vol. 3118*, 14-22, 1997.

Press, W. H., B. P. Flannery, S. A. Teukolsky and W. T. Vetterling, Numerical Recipes. Cambridge University Press, Cambridge CB2 1RP, 86-89, 1989.