Semi-automated Mineral Identification Algorithm for Ultraviolet, Visible and Near Infrared Reflectance Spectroscopy

Montero S., Irene C.¹, and Brimhall, George H, Dept. of Earth and Planetary Science, Univ. of California, Berkeley and Earth Resources Center Digital Mapping Laboratory; Berkeley, CA 94720-4767

Abstract

Identification and mapping of surface minerals illustrate and clarify the geochemical processes responsible for their spatial distribution. When the geochemical processes are of such a nature that efficient and fast mineral mapping is crucial, reflectance spectroscopy of the ultraviolet, visible and infrared regions of light is an ideal method for this purpose. In order to enhance the efficiency of this method a fast identification algorithm for reflectance spectra was developed for identification of minerals that show absorption features in reflectance spectra covering the 0.35 to 2.5 micrometers range. The algorithm uses field-acquired files containing intensity of reflected light as a function of wavelength, and derives basic information about spectral features before comparing these features to a library of reference spectra. The reference spectral library was created from laboratory spectra of selected and well-characterized endmember and mixture samples, whose spectral features were defined in a similar manner that for subsequent unknowns. Spectral absorption (band) characterization starts by fitting line segments over the spectrum to create a continuum or upper hull as defined by Clark and Roush (1984) which is removed by division. Processing of the continuum-removed spectrum proceeds with smoothing with a Savitsky-Golay least squares filter and computation of first and second derivatives following formulas developed by Madden (1978). The window sizes for the smoothing filter, first, and second derivatives are pre-assigned based on the known spectral features of the bands of interest. Spectral parameters such as band center, bandwidth, bandwidth at half maximum and band symmetry are found using the first and second derivatives. These parameters are compared to similarly acquired band information from the various reference spectra and a similarity vector is generated for each comparison. Mineral identification results from the minimization of the similarity vector. Mineral identification is output as text files, which is then used by a real-time digital mapping system, to create surface mineralogy maps contemporaneously with survey progress. Minerals specially suited to identification by this method include iron minerals such as sulfates, oxihydroxides and oxides and clays.

1. Introduction

Hyperspectral reflectance spectroscopy (high altitude or ground-based) of iron minerals has been proven invaluable to characterize the surface of the earth for purposes as varied as mineral exploration (Cudahy and Ramanaidou, 1997) or abandoned site remediation (Swayze, Smith et al., 2000).

Hyperspectral Reflectance spectroscopy applied to minerals and rocks makes use of the

¹ Corresponding author: irene_s@uclink4.berkeley.edu

selective absorption of light by electromagnetic processes that arise from the chemical composition, mineral structure, and physical state of surface minerals (refer to Hunt and Salisbury, 1970 and Burns, 1993 and references therein). If captured as intensity of reflected light as a function of wavelength, the process of absorption yields distinctive patterns that can be used to identify the minerals present as well as the environment in which they formed (figure 1).





Hyperspectral reflectance spectroscopy is particularly useful in the mapping and study of the formation and surface pathways of acid rock drainage, a consequence of the oxidation of pyrite above the water table. Acid rock drainage associated with abandoned mines is also known as Acid Mine Drainage (AMD) and it is responsible for elevated sulfate and metal concentration in adjacent streams and aquifers, as well as acidic soils and waters (California Dept. of Conservation, 2000).

Inherent to the process of AMD generation is the precipitation of secondary iron minerals (Bigham, 1994; Nordstrom and Alpers, 1999a). AMD in untreated, pyritic waste rock piles generates a metal and sulfate-rich acidic solution that flows outward following topographic and hydrological gradients depositing distinctive mineral assemblages at each step in its evolution. Each iron secondary mineral is an indicator of the metal, sulfate and pH levels of the AMD solution at the time the mineral precipitated, although subsequent mineralogical changes do occur. Mapping of the surface distribution of iron secondary minerals allows for the rapid identification of AMD sources, facilitates the evaluation process regarding the extent of the environmental problem, and illustrates

sources, sinks, and pathways for AMD. In practice traditional mapping, surveying, and geochemical methods put undue stresses on the resources available for the study of abandoned mines. Classical geochemistry methods, although effective in identifying acid indicator minerals, do not offer enough flexibility in field studies. High altitude remote sensing methods do not have the spatial or spectral resolution to resolve their presence. The seasonality of these systems is rarely addressed, and many times completely missed.



Figure 2. Idealized model of the spatial distribution of secondary iron minerals in a AMD generating waste rock pile, as viewed from above, and as a function of pH and distance from source. C= copiapite ((Fe²⁺,Fe³⁺)₄(SO₄)₆(OH)₂•20H₂0); J= jarosite (KFe₃(SO₄)₂(OH)₆); G= goethite (FeOOH); H= hematite (Fe₂O₃) Given the complexity (fig. 3) and large number of abandoned sites throughout the world it is

imperative to develop new ways to identify and characterize AMD sources rapidly and accurately.



Figure 3. Abandoned mine site, California (Red-purple areas are piles of hematite-stained metallurgical waste, yellow areas are low pH zones marked by the presence of jarosite, and white areas are rapidly eroding spots with very low pH marked by the precipitation of copiapite)

In order to support these efforts, provide definitive identification of minerals from reflectance spectra, and do so in a time frame short enough that it allows for data analysis in the field we have created a Fast Mineral Identification algorithm (FSTSpecID) to use in combination with a Geomapper® equipped digital mapping system (Brimhall and Vanegas, 2000). If combined with accurate digital mapping methods supported by real-time differential GPS and laser range finder, reflectance spectra taken from low altitude or ground platforms can be translated into detailed maps of surface mineralogy of abandoned mines and their surroundings (refer to figure 4). Analysis of these maps while still in the field offers investigators valuable insights about mineral distribution patterns, location of acid indicator minerals, and the opportunity to adapt mapping strategy to site conditions, saving thus valuable time and resources.



Figure 4. Combined digital mapping system with Fast Mineral Identification Algorithm (FSTSpecID)

The instrument used for ground and low-altitude hyperspectral mapping is a portable spectrometer measuring intensity of reflected light in the range from 0.35 μ m to 2.5 μ m (spanning ultraviolet, visible and near infrared light) with a spectral resolution under 0.03 μ m. Spatial resolution on the ground varies between 30 cm² and 4 m² depending on distance to target (Analytical Spectral Devices, 1999). Concurrent with spectral data acquisition a digital mapping system, consisting of a DGPS, laser range finder, and portable computer running Geomapper® (Brimhall and Vanegas, 2000) locates sampling points with submeter accuracy.

2. Fast Spectral Identification Algorithm (FSTSpecID)

FSTSpecID is designed to read binary or ASCII reflectance spectra files, process them as to identify iron minerals if detectable, and output results that can be used by the mapping software to generate maps of the area in the least amount of time possible. This mineral identification program makes use of spectral features caused by electrical transition processes originating in the Fe atom and that are detectable in reflectance spectrum in the range from 0.35 μ m to 1.5 μ m. (Montero and Brimhall, in preparation). In iron sulfates, oxides, and oxyhydroxides two regions display characteristic features that allow for identification. In this paper we refer to these regions as NUV and NIR (figure 5).



Figure 5. Reflectance spectra of selected secondary iron minerals in the range from 0.35mm to 1.5mm.

The approach followed in the design and implementation of this algorithm is geared at doing data processing and identification in the field, in contrast with the majority of programs available for the identification of minerals from reflectance spectra. To expedite data processing differentiation of the spectrum as a function of wavelength is used to locate and characterize absorption features of interest. Although much success has been experienced by workers identifying minerals based on the location of the absorption center (Scheinost et al., 1998) we find that given the many variables in degree of crystallization, mixing, and grain size found in rocks in the field it is not possible to identify iron minerals based on absorption center alone. The main engine of this algorithm compares not only absorption centers but also other parameters measurable in the absorption features of iron minerals in order to match unknown spectra to reference spectra in FSTSpecID's reference library.

2. 1 The reference library

Reflectance spectra used to create the reference library were acquired using the same portable spectrometer used for field studies. Clean samples of mineral endmembers and mixtures were illuminated with solar light or with a pair of quartz halogen lights while placed on a dark table or a dark box, respectively. Spectra of mixtures included in the reference library were acquired either from naturally occurring binary or tertiary mixtures of minerals or from mechanically formed binary mixtures of well characterized endmembers. Mechanical mixtures were formed by drying and crushing with agate mortar and pestle well known, pure mineral samples that were later sorted with vibrating sieves, and mixed according to particle size.

Reference spectra of mixtures are included in the library rather than attempt to incorporate an unmixing algorithm into the program. Minerals that make up the rocks encountered in the AMD environment are intimately mixed and typically have varying grains sizes, indices of absorption and reflection. These apparently simple combinations of minerals to make up a rock result in a complex, non-linear mixing of their signatures in reflectance spectra (Sherman, et al., 1982) that demands much more computing time and power to resolve than desired in field-oriented work. Given the focus of this work on AMD environments, mixtures included in this library include only mixtures of low temperature iron minerals and some simple natural assemblages of iron minerals and phyllosilicates Mg-Fe silicates.



Figure 6. Selected samples of reference spectra showing features of interest in the NUV and NIR



Figure 7. Flow diagram FSTSpecID data processing and mineral identification

2.2 FSTSpecID processing.

FSTSpecID is designed to operate as a stand-alone, graphic interactive program for use in portable or desktop computers running either Windows[©] or DOS[©] operative environments. Instead of functioning as a black-box, the graphic interface allows the user to stop processing at any time for inspection of the spectrum on the screen. Without user-induced interruptions the program typically takes less than 2 seconds computer processing time to complete the processing of one spectrum from input of spectral file to output of mineral identification output file (fig.7).

2.2.1 Continuum removal.

In order to remove background absorption and isolate particular absorption features for analysis and identification the continuum removal method of Clark and Roush (1984) is first applied to unknown spectra. This mathematical function is defined by finding pivot points or high reflectance values in the spectrum. (defined after calculating 1st and 2nd derivatives of the spectrum without smoothing), calculating line segments between the pivot points, and refined by iteration until the

continuum completely drapes the spectrum. Finally the original, unprocessed spectrum is divided by the calculated continuum (Fig. 8)



Figure 7. Process of continuum calculation and removal

2.2.2 Smoothing and derivation

Once the continuum is removed by division, the program proceeds to smooth the spectrum and calculate first and second derivatives from the smooth spectrum by means of convolution with a Savitzky and Golay second order polynomial filter (Savitzky and Golay, 1964) calculated following formulas created by Madden (1978). Prior to differentiation smoothing of the data is done in order to ameliorate the effect of random and atmospheric noise on the shape and

location of absorption bands. Our approach to minimize noise caused by variable atmospheric conditions and by changes in temperature inside the instrument is to use a well warmed instrument and to optimize the reflectance spectrometer to current conditions very frequently (i.e. less than once every five minutes, depending on atmospheric conditions) during data acquisition. Noise introduced by variable atmospheric conditions typically takes the form of high frequency noise concentrated in the regions around 0.9 μ m, 1.4 μ m, and 1.9 μ m which coincides with the regions of light absorption by atmospheric water (which is the most temporally variable atmospheric component). On the other hand, noise introduced by dark current or by changes in viewing geometry during data acquisition typically takes the form of changes in the overall value of reflectance. These sources of noise are closely monitored during acquisition in the field and constitute a minor source of noise in the final spectra.

Random noise related to instrument performance constitutes a minor fraction of the resulting spectrum, with signal-to-noise ratio typically measuring well above 500 under solar illumination conditions (Analytical Spectral Devices, 1999). However, the effect of random noise in differentiated spectra is heavily felt. Differentiation enhances high frequency noise while suppressing low frequency background noise (Demetriades-Shah et al., 1990) with the result that signal-to-noise ratio decreases with increasing derivative. Although previous workers have demonstrated the use of high order derivatives in reflectance spectra and their advantages for characterization of absorption features (Hugenin and Jones, 1986; Tsai and Philpot, 1998) we find that the increased noise levels after the third differentiation make it very difficult to discern true features from noise in field-acquired data. As a consequence of these findings we limit the use of derivatives to the use of the first (dR/d?) and second $(d^2R/d?^2)$ derivatives.

Our approach to design the type and size of the smoothing filter is based on striking a balance between the need to eliminate random, high frequency noise in the data and the need to avoid oversmoothing. As noted by Kawata and Minami (1982) oversmoothing can introduce noise in the form of band distortion and changes in the location of the center of the absorption feature. In FSTSpecID a smoothing filter of predetermined width controls the degree of smoothing and is created and convoluted to each spectrum value. To avoid oversmoothing in critical areas for mineral identification method we have set the filter window to be no greater than ½ the full width at half maximum (FWHM) of the smallest feature of interest. However, the need to avoid oversmoothing needs to be balanced by the need to account for atmospheric induced noise in the NIR region, and so knowledge regarding the instrument used, of atmospheric interference, and the character of the absorption features to be identified in the electrical region have to be taken into consideration in determining the size of the filter window. The very commonly used method of Savitzky and Golay assumes that all noise is random and has similar characteristics over the entire spectrum (Tsai and Philpot, 1998) and in this study it is preferred over variable filters.

2.2.3 Characterization of features in the NUV and NIR regions

The centers of absorption features are found where dR/d? equals zero and d^2R/d ?² is positive. The NUV and NIR regions are then reviewed to determine whether any of the features fall inside any of these regions. If this routine finds a feature that is comparable to the narrow NUV feature that is characteristic in many iron sulfates it signals the program to compare the NIR region preferentially to iron sulfates. Otherwise all reference spectra are compared without any set preference. The parameters that are analyzed by the program are (see figure 9):

- Center of absoprtion (NIR): Wavelength of the most intensely absorbed radiation of feature.
- Depth of absorption feature or spectral contrast (DEPTH_NIR): difference between lowest and highest point of feature in units of reflectance.
- Full width at half maximum (FWHM_NIR)
- Symmetry of absorption (SYM_NIR): Ratio of the left side to right side of feature at FWHM.



Figure 9. Features of NIR absorption.

2.2.4 Comparison and identification vector

During the comparison routine the program cycles though the reference spectra and determines whether there might be a feature in the unknown spectrum in the NIR region that is similar to the NIR feature in the current reference spectrum. By comparison ?NIR, ?FWHM, and

? SYM are calculated and used to create the similarity vector (fig. 10). FSTSpecID simply minimizes the length of that vector and determines that the closest match between an unknown and a reference spectra is that that yields the smallest identification vector. The three best matches as well as the NIR features of the unknown spectrum are output to an ASCII files along with file information such as file name and time and date of acquisition (fig. 11). The name of the file and the time of acquisition can then be combined with the location data to yield a map of the area based on the minerals identified by reflectance spectroscopy.



Figure 10. Comparison of parameters and similarity vector (in gray).

7:02	Depth NIR:	0.3621%R	_
Ins	FWHM:	0.2410microns	
rons	Symmetry Fac	tor: 1.5185	
Goet2592	FIT Vector Val	ue:1.172527783244	
GtQz0002	FIT Vector Value	e: 1.703615799330	
Shwert02	FIT Vector Value	e: 2.902360727544	
15 pts	Processing time(CPUsec): 2.546457637	Ŧ
15 pts	Elapsed time (US	ERsec): 65.23466778	
	7:02 ns rons Goet2592 GtQz0002 Shwert02 15 pts 15 pts	7:02 ns Depth NIR: rons FWHM: Symmetry Fac Goet2592 FIT Vector Value GtQ z0002 FIT Vector Value Shwert02 FIT Vector Value 15 pts Processing time(Elapsed time (US	7:02 ns ronsDepth NIR:0.3621%R FWHM:FWHM:0.2410microns Symmetry Factor:Symmetry Factor:1.5185Goet2592 GtQ z0002 Shwert02FIT Vector Value:15 pts 15 ptsProcessing time(CPUsec):2.546457637 Elapsed time (USERsec):65.23466778

Figure 11. Captured screen of final FSTSpecID output. First line is time of acquisition, second through fourth lines are the parameters found in the NIR absorption feature, fifth through eight lines the name of the three best matches from the spectral reference library (in this case the unknown spectrum is identified as Goethite), and the last two lines include processing parameters, processing time, and time employed by user to look at spectrum.

3. Conclusions

The ease of use, graphical interface optimized for use in field conditions, and speed of processing by FSTSpecID demonstrates that this is indeed a valid and effective method for mapping surface minerals associated with AMD environments and obtain usable information very rapidly.

We find, after verification of results by FSTSpecID of field spectra against rock samples analyzed with the aid of X-ray powder diffraction techniques, that the characterization of the shape of absorption features in the reflectance spectrum of rocks in the range from 0.35 μ m to 1.5 μ m is a definitive way to identify minerals from their reflectance spectra, even in the uncontrolled, complex environment of abandoned mines. Furthermore, we find that comparison of spectra acquired in different illumination conditions yield definitive identification of minerals after continuum removal and second order differentiation. This is an important validation of the effectiveness of comparing absorption features of unknown spectra against the features in reference spectra.

4. Acknowledgments

This research was supported by NASA Mission to Planet Earth grant NAG5-6515 entitled "Center for Assessment and Monitoring of Forest and Environmental Resources (CAMFER) and the Earth Resources Center (ERC)", and Earth Resources Center funds from UC Berkeley Office of the Vice Chancellor for Research

References

- Analytical Spectral Devices, 1999. Technical Guide. Hatchell, D., managing ed. Boulder, 15 sec. ftp://ftp.sni.net/asd/TECHGUIDE/
- Analytical Spectral Devices, 1997. Field SpecTM User's guide.
- Bigham, J. M., 1994. Mineralogy of Ochre Deposits formed by Sulfide Oxidation.In: J. L. Jambor and D. W. Blowes, ed. Environmental Geochemistry of Sulfide Mine Wastes. Toronto, Mineralogical Association of Canada. Short Course Handbook: 105-133.
- Brimhall, G. H. and A. Vanegas, 2000. Digital mapping of geology and ore deposits with Geomapper. Geol. Soc. Amer., Abstracts with programs 32: A-514.
- Burns, R. G., 1993. Origin of electronic spectra of minerals in the visible to near-infrared region.In:C. M. Pieters and P. A. J. Englert, ed. Remote Geochemical Analysis: Elemental and Mineralogical Composition. Cambridge, Cambridge University Press. 4: 3-29.
- California Dept. of Conservation, 2000. California's Abandoned Mines, a Report on the Magnitude and Scope of the Issue in the State. Volume 1. Sacramento, California Department of Conservation,

Abandoned Mine Lands Unit.

- Clark, R. N. and T. L. Roush, 1984. Reflectance Spectroscopy: Quantitative Analysis Techniques for Remote Sensing Applications. Journal of Geophysical Research 89: 6329-6340.
- Cudahy, T. J. and E. R. Ramanaidou, 1997. Measurement of the hematite:goethite ratio using visible and near-infrared reflectance spectrometry in channel iron deposits, Western Australia. Australian Journal of Earth Sciences 44: 411-420.
- Demetriades-Shah, T. H., M. D. Steven, et al., 1990. High resolution derivative spectra in remote sensing. Remote Sensing of Environment 33: 55-64.
- Hugenin, R. L. and J. L. Jones, 1986. Intelligent information extraction from reflectance spectra: Absorption band positions. J. of Geophysical Res. 91: 9585-9598.
- Hunt, G. R. and J. W. Salisbury, 1970. Visible and Near-Infrared Spectra of Minerals and Rocks: I Silicate Minerals. Modern Geology 1: 283-300.
- Kawata, S. and S. Minami, 1982. Adaptative smoothing of spectroscopic data by a linear meansquare estimation. Applied Spectroscopy 38: 49-58.
- Madden, H., 1978. Comments on the Savitzky-Golay convolution method for least-squares smoothing and differentiation of digital data. Analytical Chemistry 50: 1383-1386.
- Nordstrom, D. K. and C. N. Alpers, 1999a. Geochemistry of acid mine waters.In: G. S. Plumlee and M. S. Logson, ed. The Environmental Geochemistry of Mineral Deposits. Part A: Processes, Techniques, and Health Issues. Denver, Soc. of Econ. Geologists. 6A: 133-155.
- Savitzky, A. and M. J. E. Golay, 1964. Smoothing and differentiation of data by simplyfied least squares procedures. Analytical Chemistry 36: 1627-1639.
- Scheinost, A. C., A. Chavernas, et al., 1998. Use and limitations of second-derivative difusse reflectance spectroscopy in the Visible and Near-infrared range to identify and quantify Fe oxide minerals in soils. Clay and Clay Minerals 46: 528-536.
- Sherman, D. M., R. G. Burns, et al., 1982. Spectral characteristics of the iron oxides with applications to the Martian Bright Region Mineralogy. J. of Geophysical Res. 87: 10169-10180.
- Swayze, G. A., K. S. Smith, et al., 2000. Using imaging spectroscopy to map acidic mine waste. Environmental Science and Technology 34: 47-54.
- Tsai, F. and W. Philpot, 1998. Derivative analysis of hyperspectral data. Remote Sensing of Environment 66: 41-51.