1999

Simulation of spectra signatures of chemical leachates from landfills

Jason Hamel

SIMG-503
Senior Research

Simulation of Spectra Signatures of Chemical Leachates from Landfills

Final Report

Jason Hamel
Center for Imaging Science
Rochester Institute of Technology
May 1999

Table of Contents

Simulation of Spectral Signatures

Jason Hamel

Table of Contents

Simulation of Spectral Signatures

Jason Hamel

Abstract

Chemical leachates from landfills can turn into dangerous hazards if they are not identified and properly disposed. Currently, these chemicals are identified in the laboratory with samples hand collected from contaminated sites. This is economical for small sites but quickly gets expensive and time consuming for larger facilities that can cover hundreds of acres or multiple sites. This study examines the feasibility of using new hyperspectral detectors and computer processing to autonomously identify the presence of these leachates and greatly simplifying the monitoring process for these large sites.

The effect of even low concentrations of many chemicals requires the identification of leachates. In most cases, the spectral signature is extremely subtle and cannot be directly detected from the background soil spectra. The approach used here will identify secondary effects of the leachates on surrounding features such as vegetation stress or effects on soil moisture that might indicate contamination. This study is exploring the detectability of various states of vegetation health and soil moistures resulting from these contaminants by processing the spectra with several spectral matching algorithms. The ability to classify various health levels will determine if this monitoring method has useful applications in the field.

With the levels of health and soil moisture used in this research, the linear spectral unmixing (LSU) and orthogonal subspace projection (OSP) algorithms performed the best. They not only could identify the major constituents of mixed pixels but also generated fractions maps listing the amount of material with a specific level of health or moisture. Since their classification results were very accurate in identifying the materials with simple thresholding for the OSP and LSU algorithms, the material fractions were also included in evaluating the performance of LSU and OSP algorithms.

Table of Contents

Center for Imaging Science
Rochester Institute of Technology
Rochester, NY 14623-5604

This report is accepted in partial fulfillment of the requirements of the course SIMG-503 Senior Research.
Title: Simulation of Spectra Signatures of Chemical Leachates from Landfills
Author: Jason Hamel
Project Advisor: Rolando Raqueño
SIMG 503 Instructor: Joseph P. Hornak

Table of Contents

Simulation of Spectral Signatures

Jason Hamel

INTRODUCTION

Materials production in the United States has produced many waste products that are difficult to destroy and too dangerous to allow into the surrounding environment. These waste products often end up in landfills that are capped with clay to keep rainwater out and slow the process of the waste leaking into the surrounding land. Unfortunately, these landfills were designed to last only 40 years and then be replaced with new technology [1]. Many of these landfill sites have surpassed this limit and nothing has been developed to address this problem. The passage of time has resulted in many contaminants, or leachates, spreading into the surrounding land and water tables.

These chemical leachates can have a serious effect on the environment. Monitoring efforts currently consist of manually sampling a site and analyzing the soil in a laboratory; a procedure which can get very costly and time consuming for large sites. Remote sensing offers the ability to easily monitor large sites and reduce the need of manual sampling of high risk areas. Unfortunately, dangerous concentrations can build up long before the subtle spectral signature of the chemical is detectable in the background soil spectral signature that is captured by remote sensing detectors. The current generation of hyperspectral sensors, with hundreds of spectra channels, show the promise to identify secondary spectral effects of the chemicals on vegetation and on soil features, like water content, that result from contamination.

The ability to detect and classify the various levels of vegetation health and soil moisture could highlight areas with a high possibility of contamination and alert the monitoring agency to initiate a remediation procedure. Knowledge about the performance of classification algorithms in this situation would help select an appropriate procedure to quickly and easily identify these contaminated regions. This research will examine the ability of current hyperspectral sensors to detect the secondary spectra effects by simulating the spectra of vegetation at various levels of health and various levels of soil water moisture. These spectra will be converted into spectral bands that are used by a hyperspectral remote sensing system. The accuracy of computer classification algorithms to identify each spectra will be calculated based on known ground truth spectra. The need for atmospheric correction will be derived from analyzing the results of classifying spectra both with and without and atmosphere.

BACKGROUND

Capping Technology

The need for monitoring leachates derives from how waste products are disposed. The simplest method is to put material into landfills and cover it. To minimize spreading or leaching, rainwater is prevented from seeping into the landfill and dissolving the chemicals as it slowly percolates into ground water basins. This is accomplished by putting a clay cap over the landfill site after it has been filled [2]. The two main types of caps used today are the older kaolin clay caps and the newer geosynthetic caps. Each has a 2-foot thick foundation layer of compacted soil just above the waste pile to support the rest of the cap. Then the capping layers are placed which vary for the different capping technologies. Lastly, the top of the cap is a 2-foot thick vegetation layer. It is made of dirt that supports grasses planted on it to prevent erosion.

The capping layer for the kaolin cap, working up from the foundation layer, is a 1 foot layer of rock for venting gas, a 2 foot layer of kaolin clay to block water, and a 1 foot layer of rock to drain the blocked water. Just above this is the vegetation layer [2]. The geosynthetic cap has a high strength geosynthetic reinforcement layer just above the foundation layer to hold the rest of the capping materials on a strong, flat area. The layers places on the reinforcement layer are a synthetic gas vent layer, a geosynthetic clay liner layer made of bentonite clay to block water, a flexible membrane layer to also block water for double protection, and lastly a geosynthetic drainage layer. All of these layers are a total of about 4 inches thick. On top of this is the vegetation layer. Geosynthetic capping allows for 3-5 orders of magnitude reduction in water penetration over kaolin clay caps [3].

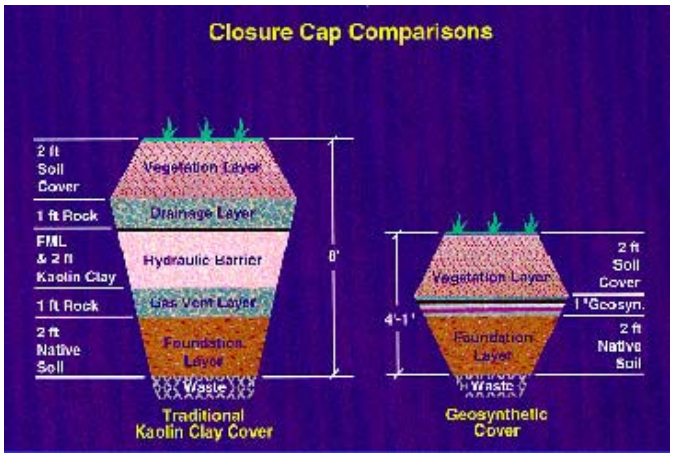


Figure 1: Visual diagram of the layers used in traditional and geosynthetic capping [2].

The problems with these technologies are that they only prevent rainwater from accelerating the seepage of chemicals. They do not actually completely isolate and contain toxic substances from the environment. The geosynthetic capping technology has only recently been developed and functions solely to block rainwater like the older capping technology; it just performs this function better. Because the chemicals can leach out over time, landfills containing toxic materials must be monitored to insure the chemical concentrations do not reach a dangerous level outside of the landfill itself.

Hyperspectral Sensors

The interest in remotely monitoring landfills has developed due to an increase in the number of hyperspectral imagers that are now operating as well as the presence of some multispectral imagers that have been used previously with some degree of success. These imagers detect the light reflecting off of the earth, objects, and atmosphere. The detectors can sample this light in specific, usually narrow, regions of the electromagnetic spectrum. By combining a number of these narrow contiguous groups, it is possible to record the wavelengths of reflected light from 0.4 microns (blue light) to 2.5 microns (infrared light).
These narrow regions are also called channels or bands. A multispectral imager often has tens of bands while a hyperspectral detector can have hundreds of bands. The reflectance recorded from objects (given a sufficient number of narrow bands in a wavelength range) can be unique for different substances and materials which is why this known as a spectral signature. The older detectors, like DAEDALUS, collected 16 bands between the visible and short wave infrared regions of the spectrum. This was not enough to detect the signatures generated by subtle contaminants though it can identify gross changes in material. The newer HYDICE sensor with 210 bands from 0.413-2.504 microns and AVIRIS with 224 bands from 0.4-2.45 microns show more potential with their increased spectral resolution. Specific information about these sensors can be found in many common remote sensing publications and web resources [4].

Previous Studies

Previous studies have tried to use remote sensing to study landfills. One study, conducted at the Department of Energy's Savannah River Site in Aiken, South Carolina used the TRW Visible/Near infrared Imaging Spectrometer (TRWIS) hyperspectral sensor, which has 85 available bands from 466 to 880 nm, to collect data over the landfill area [5]. This study identifies areas of vegetation and used a normalized difference vegetation index (NDVI) to build biomass maps. The NDVI uses the simple equation:

$$NDVI = (Band\ 78 - Band\ 43) / (Band\ 78 + Band\ 43)$$

to calculate this. The main problem with this study was that there was no ground calibration panels or ground spectral measurements collected concurrently with the sensor's collection flight. Thus, images could not be compared from day to day, and atmospheric correction could not be applied to the images. Without ground truth and atmospheric correction, it was difficult to correlate the spectral signatures from the sensor and biophysical measurements such as

leaf area index (LAI). As a test program, this study shows that hyperspectral data might be useful in monitoring landfill areas by searching for the secondary effects of leachates. The research described by this paper will expand the basic findings of this study.

A similar study was conducted in the U.K. with SPOT HRV and Landsat TM monitoring sugar beet fields [6]. This study also used a vegetation index called optimized soil adjusted vegetation index (OSAVI) as well as listing several other indices. The calculation of this metric is as simple as NDVI above:

$$OSAVI = (R_{nir} - R_r) / (R_{nir} + R_r + 0.16)$$

The R is the spectral reflectance of the material for near infrared (nir) and red (r) bands of the detector being used. The (0.16) is a soil adjustment coefficient selected to minimize the background in the image (i.e. soil). The main problem with this is that the focus is on the general vegetation cover. It does not make use of all of the spectral data in an image, just a few bands. Thus, while it shows general vegetation cover, it can not distinguish between healthy or unhealthy vegetation. Also, the soil background is minimized and this is often a large component of the spectra in a pixel if vegetation is not very thick.

METHODS

This research project can be divided up into basic tasks. Task 1 is to generate the basic vegetation and soil spectra and combine these into mixed pixels to stress the classification algorithms. Task 2 is to convert the spectra into digital counts that would be returned from a hyperspectral imager looking at a spot on the ground. Task 3 is to add noise to the digital counts (produces a more realistic data set) and then invert the digital counts back to reflectance, as would be done with real data. Task 4 is to classify the basic reflectance set, the noisy reflectance set, and the AVIRIS digital count set of pixels. The general flow of this procedure is shown in Figure 2. The steps of the process are explained below.

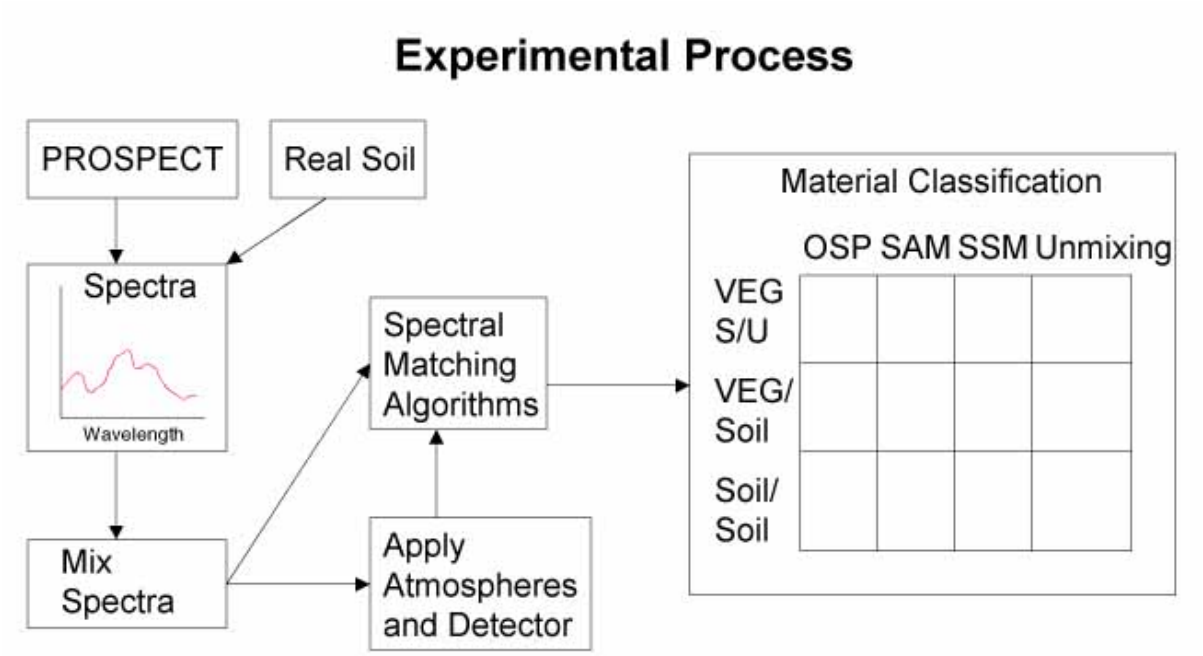


Figure 2: The basic experimental procedure for this research

Spectra

The research described here used simulated vegetation spectra and ground measured soil spectra as its data instead of actual imagery. This allows for very well known truth data when analyzing the accuracy of the classification algorithms used on our data set. The soil spectra were collected over 5.5 hours after a rain storm while the ground dried out. Specific water content measurements were not taken so the spectra can not be labeled with specific soil moisture content. Only relative labels of wet, moist, and dry soil can be used.

Prospect

The vegetation spectra were generated using PROSPECT, a radiative transfer model of the optical properties of plant leaves [7]. Spectra from 400 nm to 2500nm are generated using four basic inputs: chlorophyll concentration, equivalent water thickness, incident angle of light onto the leaf, and a structural parameter based on the cellular structure of the leaf. For this research, the incident angle was held at 45 degrees and the structural parameter was held at 1.832 (soybean leaves). The chlorophyll concentration and water thickness were varied to produce a health leaf (chlorophyll = 70.0 micrograms/cm^2 and water thickness of 0.1000 cm) and a stressed leaf (chlorophyll = 10.0 micrograms/cm^2 and water thickness of 0.0010 cm).

Expanded Data Set

The five spectra mentioned above, health leaf, stressed leaf, wet soil, moist soil, and dry soil, were mixed to provide a more realistic data set. A mixed pixel was created by combining the original spectra in 50% mixtures. This is done by:

$$\text{Reflectance}_{\text{mixed}} = 0.5\text{reflectance}_1 + 0.5\text{reflectance}_2$$

This procedure created 10 additional mixed pixels: dry/moist, dry/wet, healthy/dry, healthy/stressed, healthy/moist, healthy/wet, stressed/dry, stressed/moist, stressed/wet, and moist/wet pixels. The similarity of these spectra should stress each classification algorithm. This is especially true for the soil spectra, which have a very similar spectral shape before they are mixed.

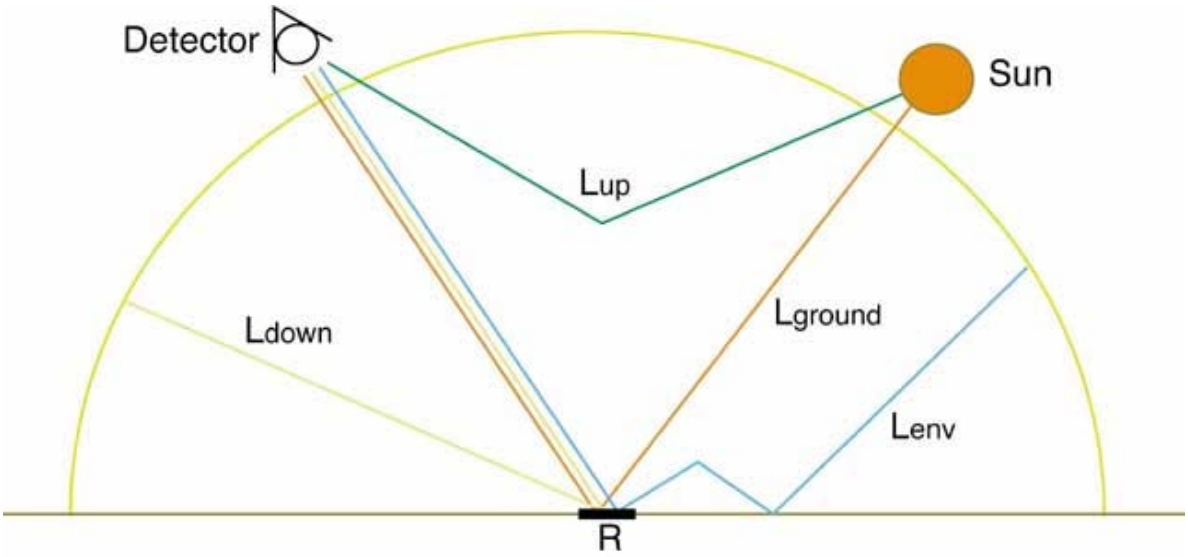
Atmosphere and Sensors

A major factor in any remote sensing image is the atmosphere. The detector is usually in a plane or satellite far above the target it is looking at. The light from the sun passes through the atmosphere to hit the target and then passes through it again to reach the sensor. The various propagation effects can be simulated using MODTRAN, the MODerate resolution TRANsmission atmospheric model developed by Spectral Sciences Inc. and the USAF Phillips Laboratory [8]. Version 4.0 of this model can generate the atmospheric effects for wavelengths from 0.4 to 20 microns, which is within the range output by PROSPECT, making this models use ideal for the current study.

The transformation from reflectance on the ground to radiance reaching the detector is calculated by:

$$L_{\text{sensor}} = (L_{\text{ground}} + L_{\text{down}})*R/(1-R*S) + L_{\text{env}}*R + L_{\text{up}}$$

The following diagram shows the radiance paths for each of the above terms:



where L_{sensor} is the radiance reaching the detector, L_{ground} is the radiance reflected from the sun to the detector, L_{down} is the downwelled radiance from the sky reflecting off the target, R is the reflectance of the object, S is the spherical albedo, L_{env} is the radiance that multiple scatters from the surrounding environment onto the target and then to the detector, and L_{up} is the upwelled radiance or light that reaches the detector without hitting the target. MODTRAN calculates all these variables except for R, which was generated as mentioned in previous sections above. The atmosphere used in this research had a visibility of 10.0 km, elevation of 0.315 km, and water vapor of 2.25.

Sensor Effects

There are a variety of hyperspectral sensors available today. Due to availability of data, the AVIRIS (Airborne Visible and InfraRed Imaging Spectrometer) hyperspectral sensor will be used in this simulation [9]. This sensor has 224 bands from 400 nm to 2500 nm. ENVI's spectral resampling function was used to convert all generated spectral data into AVIRIS's specific spectral bands. The second main sensor effect used for this simulation was the gain factor applied to each channel to turn a radiance value into a digital count. By applying the supplied gain factors to the radiances calculated above, the approximate AVIRIS digital counts for each pixel were determined.

Realistic Simulation

Real experimental measurements do not provide perfectly clean spectra. There is always some random noise that enters the measurements due to the electronics of the sensor, environmental conditions, etc. This was taken into account by generating some standard gaussian random noise with a standard deviation of 1 and adding it to the spectrum of each AVIRIS digital count pixel. This noise is not a true representation of AVIRIS noise but its magnitude is what might be expected with AVIRIS's signal to noise ratio of 500. This is essentially the AVIRIS image a user would receive. This data is converted to radiance by the same gain factors mentioned above and then the above transformation is inverted to solve for the reflectance of the pixel at the ground. We are simplifying this case because the atmosphere that will be applied in this inversion is already known since it is the atmosphere that was applied to generate the digital counts in the first place. In a real situation, outside information like radiosonde data is used to generate an approximate atmosphere that will then be used for the inversion. The result of this transformation will contain artifacts due to errors in estimates of the atmospheric parameters and spectral mismatch of the MODTRAN model and the actual sensor.

Classification

A variety of classification algorithms will be applied to the generated spectra during this simulation to determine the detectability of the various levels of vegetation health and soil water content. Six algorithms where selected, of

which the spectral signature matching (SSM) and the orthogonal subspace projection (OSP) algorithms was coded up in IDL, an Interactive Data Language available from Research Systems Inc. [10]. The mathematics for these two routines are listed in following sections. The spectral angle mapper (SAM), minimum distance (MD), binary encoding (BE), and linear spectral unmixing (LSU) algorithms were performed using ENVI's built in routines.

The endmembers for the classification will be the five basic vegetation and soil spectra. The reflectance spectra will be used for the classification of the original 15 reflectance pixels as well as the 15 noisy reflectance pixels resulting from the inversion procedure. The five basic spectra will then be propagated through the atmosphere and converted to digital counts to be used as endmembers while classifying the digital count spectra.

Spectral Signature Matching

The SSM uses an encoding algorithm to match spectra. The following is reproduced with only slight changes from Mazer et al [11]. A single pixel is an L-dimensional vector:

$$\vec{X}_{ij} = [X_{ij}(1), X_{ij}(2),..., X_{ij}(\gamma),...,X_{ij}(L)]^T$$

where L is the number of spectral channels and indices (i,j) refer to the spatial location of the pixel. μ_{ij} is the spectral mean of pixel (i,j):

$$\mu_{ij} = \left\lceil \frac{1}{L} \right\rceil \sum_{\gamma=1}^{L} X_{ij}(\gamma)$$

An L-bit binary code vector is constructed as:

$$\vec{Y}_{ij}^{a} = H\left\{\vec{X}_{ij} - \mu_{ij}\right\}$$

where H(u) is the unit step operator defined by:

$$H(u) = \begin{cases} 1, u \geq 0; \\ 0, u < 0. \end{cases}$$

The vector is constructed is a binary representation of the spectral amplitude but information about local slope at each measured wavelength is not present. Therefore, an additional L-bit code vector is constructed as:

$$\vec{Y}_{ij}^{b}(\gamma) = \begin{cases} 1, [X_{ij}(\gamma+1) - X_{ij}(\gamma-1)] \geq 0, \\ 0, [X_{ij}(\gamma+1) - X_{ij}(\gamma-1)] < 0, \end{cases} \quad \gamma = 1,2,...,L.$$

The two code vectors are then concatenated to form a single, 2L-bit code vector $\vec{Y}_{ij}$ which is taken to be the binary code word representing the spectrum of the pixel (i,j). $\vec{Y}_{mn}$ is the 2L-bit code for the reference spectra, calculated like above. The similarity measure used to determine spectral signature matches is the Hamming distances computed from:

$$D_h\left(\vec{Y}_{ij}, \vec{Y}_{mn}\right) = \sum_{\gamma=1}^{2L} Y_{ij}(\gamma)(XOR)Y_{mn}(\gamma)$$

which is just a 2L sum of bit-wise exclusive-or operations. With this measure, a similar match will have a low output value like the SAM algorithm described below. Thus, if displayed as an image, the collection of Hamming distances would form an image that have bright pixels corresponding to materials that were dissimilar from what was being identified. A classified image is produced for each material which is classified.

Orthogonal Subspace Projection

The orthogonal subspace projection algorithm is the second algorithm that will be implemented. The following outline of the mathematics of the OSP is taken directly from Ientilucci [12] with only slight modification. This algorithm is used to identify a substance in a mixed pixel, or a pixel which contains various materials. The unwanted

spectral signatures are called background while the desired spectral signature is called an endmember. The mixed pixel can be described as a column vector:

$$r_i = M\alpha_i + n_i$$

where r is an γ x 1 column vector of the i^{th} mixed pixel. γ is the number of bands in the image. The matrix M is γ x p where p is the number of distinct endmembers in a pixel and $\gamma > p$ for an overdetermined system. Each column is a linearly independent column representing a spectral signature. α_i is a p x 1 column vector where the j^{th} element represents the fraction of the j^{th} signature present in the i^{th} mixed pixel. n is a γ x 1 column vector of additive white Gaussian noise.

For brevity, the i subscript will be dropped with the assumption that all following calculations will be done on a per pixel basis. $M\alpha$ can be written so as to separate the desired spectral signature from the undesired signatures:

$$M\alpha = d\alpha_p + U\Gamma$$

where d is a γ x 1 column vector containing the desired signature and α_p is a 1 x 1 vector with the fraction of the desired signature. Matrix U is composed of the remaining columns from M with dimension γ x (p-1) and where Γ is a column vector containing the remaining (p-1) fractions of α.

The operator P eliminates the effects of U using a least squares optimal interference rejection operator which is a γ x γ matrix:

$$P = (I - UU^{\#})$$

where $U^{\#}$ is the pseudo inverse of U, denoted by $U^{\#} = (U^T U)^{-1} U^T$. I is the identity matrix. If this operator is applied to r, we get:

$$Pr = Pd\alpha_p + PU\Gamma + Pn$$

P operating on $U\Gamma$ reduces contribution of U to about zero. Now a operator x^T needs to be applied to maximize the signal-to-noise ratio. With some simple math, this turns out to be:

$$x^T = kd^T$$

where k is an arbitrary scalar. Thus, an overall classification operator is a 1 x γ vector:

$$q^T = d^T P$$

which first nullifies the unwanted spectral signatures and then maximizes the desired signal. The output of this algorithm will be negative that of the SAM and the SSM. It is bright where a material is located but is dark where it is not. Like the SSM, this algorithms produces a classified image for every material type which is classified.

ENVI Classification Algorithms

The remaining four algorithms are found in ENVI, the Environment for Visualizing Images image processing software by Research Systems [13], and were chosen as standards against which the coded algorithms mentioned above will be compared. This software is specifically designed to work with hyperspectral imagery and most common imaging formats used to transmit airborne and satellite data. The algorithms selected were the Spectral Angle Mapper (SAM), Binary Encoding (BE), Minimum Distance (MD), and Linear Spectral Unmixing (LSU).

The SAM algorithm uses a reference spectra, r, and the spectra found at each pixel, t. The basic comparison algorithm is:

$$\cos^{-1}\left[\frac{\vec{t}\cdot\vec{r}}{\|\vec{t}\|\cdot\|\vec{r}\|}\right]$$

which can be written as:

$$\cos^{-1}\left[\frac{\sum_{i=1}^{nb} t_i r_i}{\sqrt{\sum_{i=1}^{nb} t_i^2}\sqrt{\sum_{i=1}^{nb} r_i^2}}\right]$$

where nb is the number of bands in the image [14]. The formula above treats each spectra as a vector with nb dimensions. The algorithms calculates the angle between the vector of the test pixel's spectra and the vector of the

reference spectra. The greater the difference between spectra, the larger the value this function returns. A threshold can be set such that a pixel with an angle less then the threshold is classified as the material being identified. This algorithm is useful because treating the spectra as vectors allows the spectra being compared to have different relative brightnesses. This derives from basic geometry were two vectors pointing in the same direction with different magnitudes will have an angle difference of zero degrees. Different relative image brightnesses just act as scalars to vectors and will not change to reported angle between the vectors. This property can be very useful when dealing with uncalibrated remote sensing data.

The BE algorithm uses the same basic algorithm as the SSM mentioned above but is a version coded by Research Systems [13]. It is included as a comparison to the version coded for this project.

The MD algorithm calculates the spectral mean of each endmember class [15]. The spectral mean of each test pixel is calculated and compared to the endmember spectral means. The smallest difference determines which material the pixel gets classified as. This method was chosen over a gaussian maximum likelihood classifier due to the limited number of pixels being classified. This small data set does not allow for accurate calculation of the covariances of each class.

The LSU algorithm was chosen because it is designed to specifically deal with mixed pixels [13]. Given a set of endmembers, it determines the fraction of each endmember present in each pixel. This ability makes the LSU much more versatile than the above algorithms when analyzing the spectra in a pixel.

RESULTS/DISCUSSION

Spectra Results

The following graphs show the results of generating the five basic vegetation and soil spectra. Figure 3 shows the reflectance spectra of the vegetation as generated by PROSPECT and spectra of the soil as measured on the ground with a spectrometer. The regions just beyond 1.3 microns and 1.8 microns are areas of water absorption. This absorption greatly affects the signal to noise ratio of these bands and carries almost no useful spectral data for this research. Therefore these bands were set to zero so the noise would not interfere with the classification of the materials.

The difference between the healthy and stressed vegetation is attributable to the chlorophyll and water concentrations which were entered as input into the PROSPECT model [16]. Water absorbs in the near infrared and shortwave infrared regions of the spectra (0.8-2.5 microns). Thus, the stressed leaf, which has less water, absorbs less radiation and has a higher reflectance in the IR portions of the spectrum. Chlorophyll absorbs the most radiation around the blue and red regions of the spectrum (0.4 microns and 0.7 microns). Thus, the lower chlorophyll concentrations in the stressed leaf absorbs less radiation and has a higher reflectance in the visible portions of the spectrum.

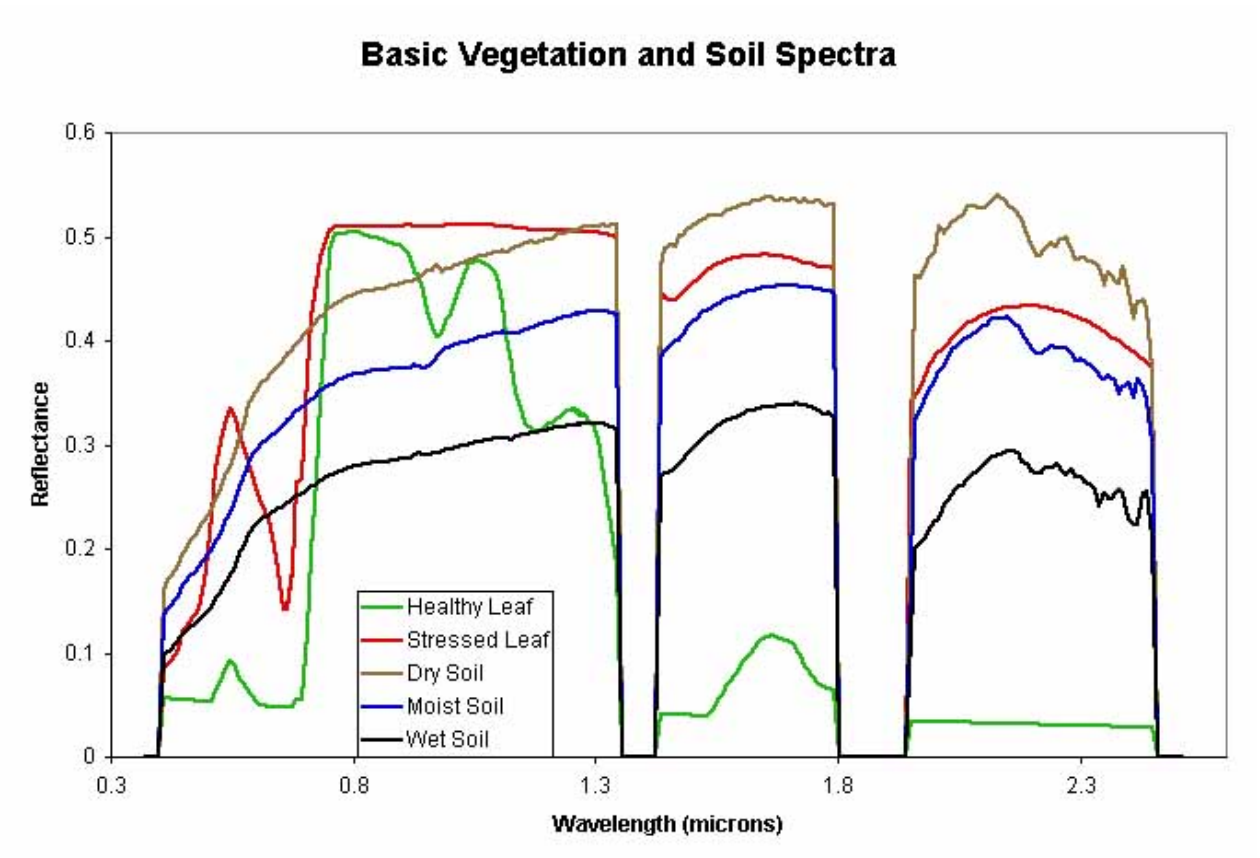


Figure 3- The five basic spectra that were used in this research.
The vegetation spectra were generated by PROSPECT.
The soil spectra were measured as ground truth.

Figure 4 shows the digital count spectra of the vegetation and soil after being propagated through the specified atmosphere and the AVIRIS digital count gain factors. The areas just beyond 1.3 microns and 1.8 microns are still set to zero at the water absorption regions where little signal is observed because of attenuation.

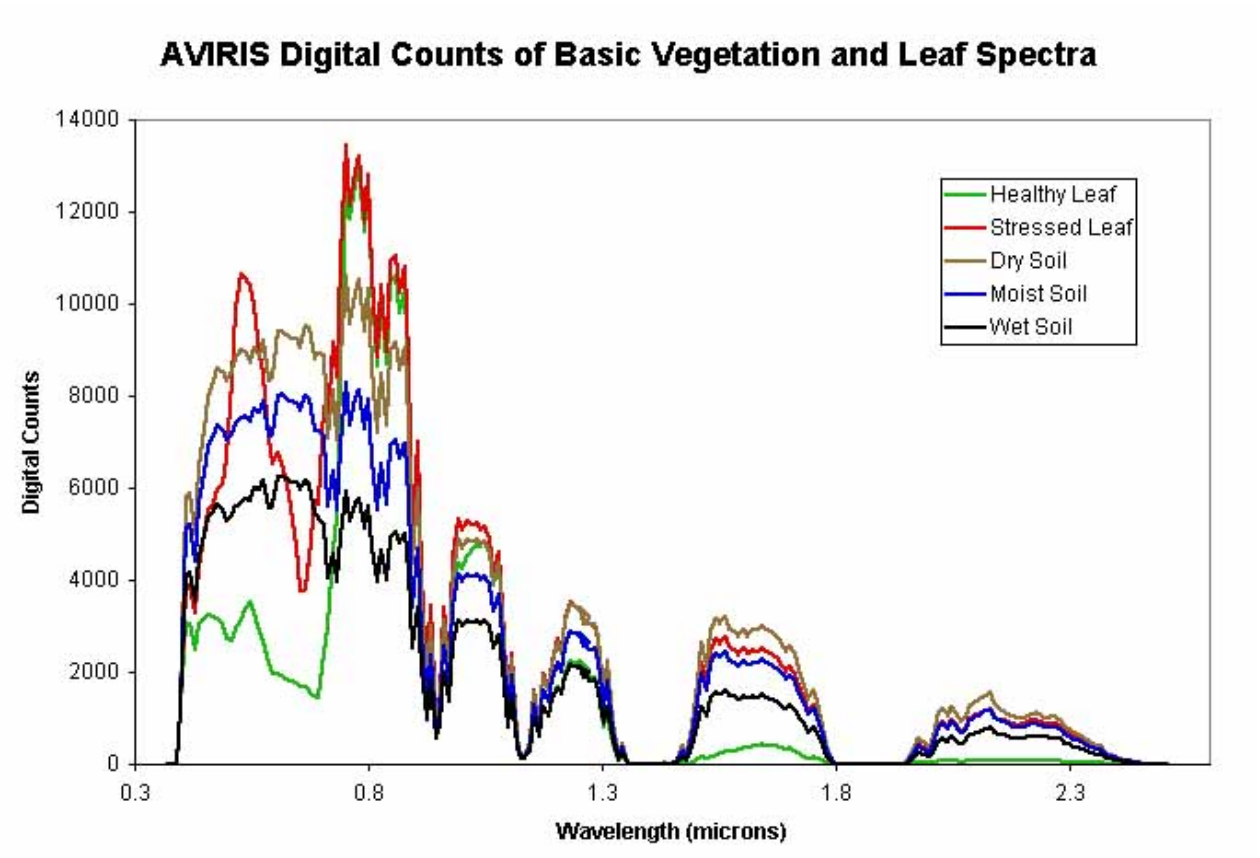


Figure 4-The five basic vegetation and soil spectra after propagation through an atmosphere and conversion into AVIRIS digital counts.

Figure 5 shows the reflectance spectra after noise was added to the digital count images and the now noisy pixels were inverted from digital counts to reflectance units again. The additional regions of the spectra which go to zero around 0.6 microns and 1.0 microns are due to the inversion process with the specific atmosphere mentioned above. In this simulated atmosphere, the spherical albedos were set to zero values but are in effect very small values. This effect will most likely not be present in clearer atmospheres. The effect of the noise can be seen in the addition of slight variations in what used to be very smooth spectral values (examine the spectra for the healthy leaf between 1.9 and 2.5 microns). The artifacts created by the spectral mismatch of MODTRAN with the sensor model can be seen at the edges of some of the water absorption features where the worst effects were removed.

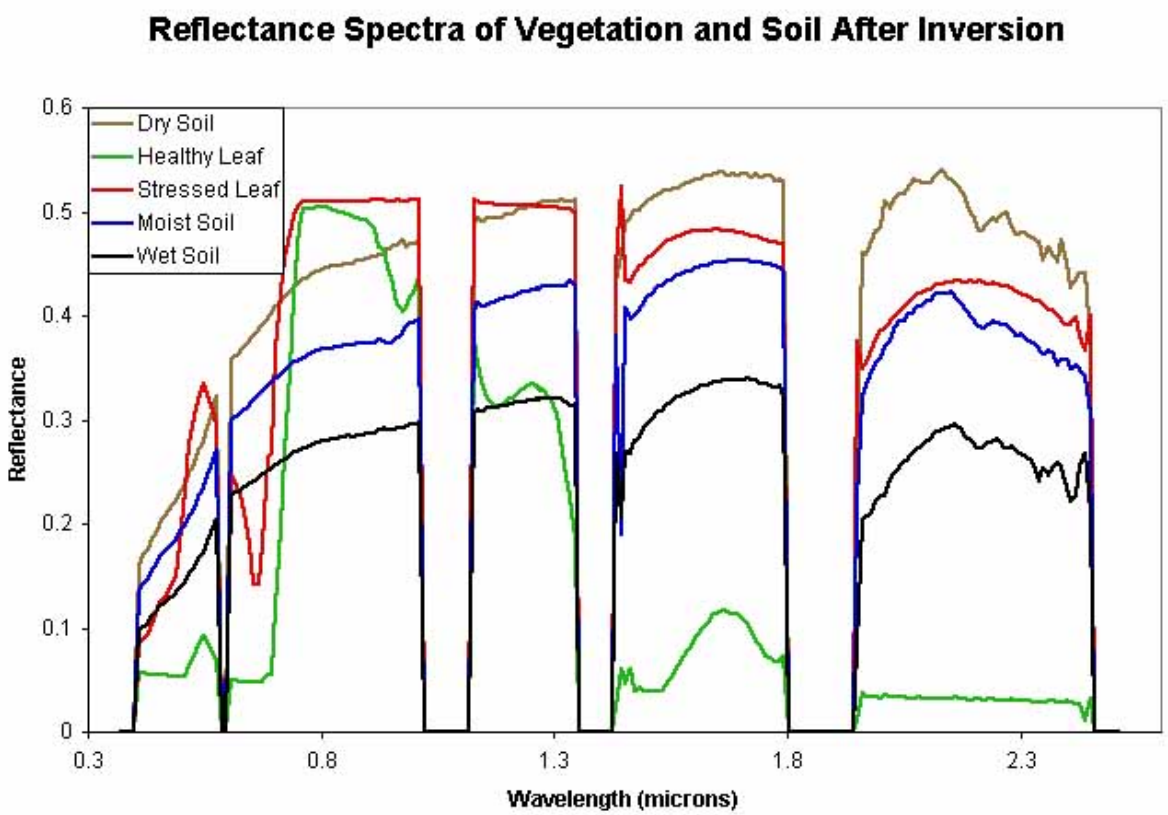


Figure 5- The reflectance of the five basic spectra after noise was added and the digital counts were inverted back to reflectance values.

Note that in the data sets actually used, there are an additional 10 spectra that lie between the 5 spectra shown in the above figures which are averages of two of the shown spectra. These were not shown because they complicate the figures without showing any additional detail beyond the fact that many of the spectra have very similar spectral features.

Classification Results

The results for the classification of the original 15 ground reflectance spectra, the 15 AVIRIS sensor digital count spectra with atmosphere added, and the 15 noisy reflectance spectra which were retrieved from inverting the atmosphere are broken up into two groups. The LSU and OSP algorithms both produce fraction maps for each endmember which list the fraction of each endmember present in a mixed pixel (i.e. it identifies the amount of each material present in a mixed pixel). Knowing the actual fractions in each pixel, the error can be calculated as:

$$SSE = \frac{1}{N}\sum(Image_{i,j} - Truth_{i,j})^2$$

where i is an endmember (1 to 5 in this study), j is a pixel (1 to 15 in this study), and N is the total number of pixels (15 in this study). The output of this formula, applied to the classification results for each of the three data sets, is shown in Table 1.

Classifier	Ground Reflectance	Sensor DC with Atmosphere	Retrieved Reflectance
LSU	4.81×10^{-11}	0.239	0.032
OSP	2.32×10^{-6}	0.237	0.038

Table 1-The sum of squared errors for the LSU and OSP algorithms on each pixel set.

The remaining four classification algorithms were not as useful as the LSU and OSP. The SAM, MD, BE, and SSM algorithms can search for mixed pixels if endmembers made up of the appropriate pixel mixtures are included in the spectral library. This would not increase the size of the library used in this research much because of the relatively small number of material conditions which were being analyzed. The classification of real imagery could include so many factors that the spectral library would quickly become excessively large. Thus, the classifications were performed with the basic unmixed spectra as one would want to do with a real application. This resulted in a much lower performance of these algorithms as compared to the LSU and OSP. The accuracy was calculated by assigning a correct classification of a pixel if an algorithm correctly identified the material present in the pure pixel or if it correctly identified one of the materials present in a mixed pixel. The percent correct is simply the sum of the number of correctly identified pixels over the total number of pixels. The results for each of the four algorithms on each data set is found in Table 2.

Classifier	Ground Reflectance	Sensor DC with Atmosphere	Retrieved Reflectance
SAM	66.67%	40.00%	66.67%
MD	66.67%	80.00%	66.67%
BE	86.67%	66.67%	86.67%
SSM	93.33%	80.00%	93.33%

Table 2-The percent correct results of the SAM, MD, BE, and SSM algorithms on each pixel set.

The classification results clearly show a sharp drop in performance for every algorithm except the MD algorithm when classifying the pixels which contain atmosphere (the middle column of Tables 1 and 2). This increased error is probably due to the compression of the spectra in the shortwave region. The fifteen spectra will show very little difference because of the narrow range of digital counts covered. For the majority of the classification routines, this severely lowered their performance. For future studies and application on real imagery, it is recommended that the atmosphere be removed (or its effects corrected) before classification of the scenes.

The results for the reflectance data set show that the LSU and OSP algorithms both performed far better than the other four algorithms. Listing the percent correct for these algorithms was not useful because these algorithms both show perfect classification with minimal thresholding. The toughest endmembers to classify often had the correct fractions present at twice the level of incorrect fractions. Most of the endmember classifications reported six times the amount of correct fractions over incorrect fractions. This made identifying the major constituents of each pixel very easy with simple thresholds. Thus, instead of percent of correct classification, the sum of squared error indicates how correctly each fraction of each material was calculated. The very low results describe how well these 2 algorithms performed. These results suggest that future studies should probably work with algorithms which produce fraction maps to calculate more useful information.

CONCLUSIONS

The results of this research certainly indicate that atmospheric correction allows more accurate determinations of vegetation health and soil moisture. The amount of spectral information present in hyperspectral data sets allows for the identification of a variety of environmental stress levels, especially with the LSU and OSP algorithms. The accuracy of these algorithms also suggests that this study did not determine the limit of detectability of these sensors. There are many areas where more factors can be included to make this simulation both more realistic and more difficult for the algorithms tested.

The first area is to improve the noise simulation. Gaussian noise with standard deviation of 1 is not representative of the AVIRIS sensor (even assuming excellent signal to noise). Much more accurate values could be derived by analyzing the dark field data supplied with AVIRIS scenes. Because the noise used in this study was somewhat low, this change will probably introduce more variation in the observed retrieved reflectance data set. This was not implemented in this study because of time constraints.

The second area to study is the effect of the atmosphere. This research used a single atmosphere with a 10 km visibility. The effect of a clearer atmosphere or one that is even worse (MODTRAN can produce atmospheres with

visibilities down to 1 km) on the classification algorithms performance is not known. As in the case of the MD classifier, this could improve the results for some of the algorithms. Money and time could be saved if specific atmospheric conditions can be identified where atmospheric correction has very little effect on the performance of the classification algorithms. Also, the effect of clouds need to be addressed in operational remote sensing where collections are not constrained to perfect collections conditions. Even clouds not in the scene can still have a large impact on the reflectance of objects just by reflecting more light onto the object. Including these effects would make the simulation more realistic and provide more accurate reflectance retrievals.

Increasing the number of the vegetation health levels would also be useful. This study used the extremes of high and low vegetation health to determine if these extreme levels are distinguishable. Since these levels were easily distinguished by both the LSU and the OSP algorithms more subtle levels need to be studied. Adding additional health levels will not only stress the classification algorithms more but it will improve the quality of the simulation by making a more realistic range of vegetation healths; plants are not just really healthy or essentially dead. Just raising the number of levels to 4 or 5 shows the potential to greatly stress the algorithms. While the two vegetation levels were clearly distinguishable, the algorithms were starting to have trouble with the 3 soil spectra. Determining the exact number of levels where detectability starts to break down should be a major focus for future studies. Also future investigations should target a specific plant type to analyze because most plants will die with chlorophyll and water concentration which are too low or too high, such as those used in this study.

Along with increasing the number of vegetation layers, it would be worthwhile to change the degree of mixing observed in the mixed pixels. This research only used a 50% mixture (or an average of two spectra). Other combinations like 25% of one material and 75% of another (as well as many other combinations) will allow us to study the accuracy of the OSP and LSU algorithms fraction map outputs. These algorithms can also be stressed by adding 3 or even 4 material types into a mixed pixel. Both of these procedures should help in determining better limits on the detectability of the various vegetation and soil parameters.

Lastly, this study considered only single simulated pixels. There was not sufficient time or available data to validate these results on actual collected data. Subsequent AVIRIS flights with good spectral ground truth data could either support or refute the conclusions that were found in this simulated case. A complication is that it will be difficult to collect real data which shows the vegetation health and soil moisture variations that were explored here on just a basic level. The more complex simulation factors mentioned above would be even harder to validate. An intermediate step would be to incorporate the basic pixel spectra implemented here into a simulated scene as created by a radiation model like DIRSIG (Digital Imaging and Remote Sensing group Image Generation model). This software builds a three dimensional scene and can incorporate the various atmosphere and sensor effects used in this research as well as a variety of other effects encountered with real image collection (like the effects of imaging from unstable platforms like a plane). This would be an intermediate simulation step which has shown to be accurate in past studies of other remote sensing phenomena.

While this research shows that the LSU and OSP algorithms are the most accurate classifiers for the conditions studied, it is unlikely that a single detection algorithms will have the best performance for all situations and conditions that are experienced in real world imagery. It is much more probable that a set of classification algorithms will be put together in which each algorithms performs best for a certain set of conditions. In the future, by knowing what this classification set is and the condition under which the imagery was taken, it should be possible to accurately identify a variety of materials under a variety of conditions.

Table of Contents

Simulation of Spectral Signatures

Jason Hamel

References

1. Applications of Hyperspectral Data. Technical Proposal, MTP98-015/BP98IT-035, NRA SSC-02-98, Digital Imaging and Remote Sensing Group, Center for Imaging Science, Rochester Institute of Technology, Rochester, NY.

2. Environmental Restoration Briefing Book: Burial Ground Complex, Department of Energy's Savannah River Site Environmental Restoration Program, Technologies & General Information, Aiken, South Carolina, URL: http://www.srs.gov/general/srenviro/erd/gen/bgc.pdf

3. Environmental Restoration Briefing Book: Geosynthetic Cap Closure Technology, Department of Energy's Savannah River Site Environment Restoration Program, Technologies and General Information, Aiken, South Carolina, URL: http://www.srs.gov/general/srenviro/erd/gen/geocap.pdf

4. Imaging Spectroscopy: Link to Hyperspectral & Multi-spectral (Sensor, Algorithms, & Data Processing Applications) Resources on the Web, OKSI Opto-Knowledge Systems, Inc., URL: http://www.techexpo.com/WWW/opto-knowledge/IS_resources.html

5. Dr. Halkard E. Mackey, Westinghouse Savannah River Company, Aiken, SC and John R. Jensen, Kan He, and Derek Graves, Department of Geography, University of South Carolina, Columbia, SC. Technological Transfer of High Spectral Resolution Remote Sensing Imaging Spectrometer on the Savannah River Site for Environmental Applications (October 9, 1995), URL: http://www.cla.sc.edu/GEOG/rslab/vip/west/west.html

6. M.D. Steven, The Sensitivity of the OSAVI Vegetation Index to Observational Parameters, *Remote Sensing of Environment*, **63**:49-60 (1998).

7. S. Jacquemoud and F. Baret, PROSPECT: A Model of Leaf Optical Properties Spectra, *Remote Sensing of Environment*, **34**:75-91 (1990).

8. MODTRAN, Spectral Sciences Inc., URL: http://www.spectral.com/modelhome.html

9. AVIRIS hyperspectral sensor, URL: http://makalu.jpl.nasa.gov/aviris.html

10. IDL, Research Systems Inc., URL: http://www.rsinc.com/idl/index.cfm

11. A.S. Mazer, M. Martin, M. Lee, and J.E. Solomon, Image Processing Software for Imaging Science Spectrometry Data Analysis, *Remote Sensing of Environment*, **24**:201-210 (1988).

12. E. Ientilucci, Hyperspectral Image Classification Using Orthogonal Subspace Projections, Student Paper, Center for Imaging Science, Rochester Institute of Technology, May 13 (1997).

13. ENVI, Research Systems Inc., URL: http://www.rsinc.com/envi/index.cfm

14. R.H. Yuhas, A.F.H. Goetz, and J.W. Boardman, Discrimination Among Semi-arid Landscape Endmembers Using the Spectral Angle Mapper (SAM) Algorithm, *Summaries 3rd Annual JPL Airborne Geoscience Workshop*, Jet Propulsion Laboratory, Pasadena, CA, June 1992, pp.147-149.

15. J.A. Richards, *Remote Sensing Digital Image Analysis: An Introduction*. 2nd Edition. Springer-Verlag, Berlin, 1993.

16. J.A. Howard, *Remote Sensing of Forest Resources: Theory and Application*, First Edition. Chapman & Hall, London, 1991.

Table of Contents | Thesis

List of Symbols

Symbol	Definition
AVIRIS	Airborne Visible-Infrared Imaging Spectrometer
BE	Binary Encoding classification algorithm
ENVI	Environment for Visualizing Images hyperspectral software package
HYDICE	Hyperspectral Digital Imagery Collection Experiment
LSU	Linear Spectral Unmixing classification algorithm
MD	Minimum Distance classification algorithm
MODTRAN	MODerate resolution TRANsmission atmospheric model
NDVI	Normalized Difference Vegetation Index
OSAVI	Optimized Soil Adjusted Vegetation Index
OSP	Orthogonal Subspace Projection classification algorithm
SAM	Spectral Angle Mapper classification algorithm
SSM	Spectral Signature Matching classification algorithm

Table of Contents | Thesis