USING REMOTELY SENSED IMAGERY TO IDENTIFY PARKING LOT SEALANT SURFACE TYPES

THESIS

Presented to the Graduate Council of
Texas State University-San Marcos
in Partial Fulfillment
of the Requirements

for the Degree

Master of SCIENCE

by

Mohan Rao, B.A.

San Marcos, Texas
May 2009

USING REMOTELY SENSED IMAGERY TO IDENTIFY

PARKING LOT SEALANT SURFACE TYPES

Committee Members Approved:

_____
Nate Currit, Chair

_____
Mark Fonstad

_____
Peter Van Metre

Approved:

_____
J. Michael Willoughby
Dean of the Graduate College

ACKNOWLEDGEMENTS

Every aspect of this thesis, from conception to the printing of these pages was made possible because many individuals chose to give their time, energy and enthusiasm to this effort. My thanks and appreciation goes out to every one of them. I want to especially highlight the contributions of the following:

- Dr. Nate Currit, Dept. of Geography Assistant Professor at Texas State University-San Marcos and committee chair, for his patience, consistency, and invaluable guidance from the beginning,
- Dr. Peter Van Metre, USGS Texas Water Science Center Research Hydrologist and committee member, for providing the motivation for this study and for proactively supporting this effort in so many ways,
- Dr. Mark Fonstad, Dept. of Geography Associate Professor at Texas State University-San Marcos and committee member, for his crucial insights and feedback,
- Ms. Allison Glass-Smith, Dept. of Geography Graduate Student Advisor at Texas State University-San Marcos, for enthusiastically providing an accurate navigational map of the thesis process,
- Mr. Thomas Bashara, Environmental Investigator at the City of Austin, for his dedication to protect Austin's water quality and for sharing his encyclopedic knowledge of parking lot sealants and their distribution across the city,

- Mr. Thomas Ennis, Environmental Engineer at the City of Austin, for his dedication to protect Austin's water quality and for sharing his encyclopedic knowledge of parking lot sealants and their distribution across the city,
- Larry Hill, for being a true friend and for helping obtain spectral data well past lunch time,

and a special thanks to

- Elizabeth Kubala, for being patient and supportive for many years, and for taking accurate and articulate field notes for many hours.

This manuscript was submitted on April 7, 2009.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

INTRODUCTION

Over the last 25 years, the United States has experienced strong population and automobile usage growth. This, coupled with increased popularity of "big box" commercial centers with expansive parking lots has resulted in a large amount of land turned into parking lot surfaces. Parking lot surfaces are routinely coated with a variety of sealants for perceived aesthetic and durability benefits. Unfortunately, some of the materials used to seal parking lots can be toxic (Bryer et al., 2006, Scoggins et al., 2007). The commonly used coal tar based sealants are especially toxic because the rainwater runoff from coal tar surfaces can contain more than 100 times the Probable Effect Concentration guideline level for polycyclic aromatic hydrocarbons (PAHs) (Mahler et al., 2005). Before the banning of coal tar sealant application in Austin, the City of Austin estimates that about 600,000 gallons of coal tar sealants were applied annually in the greater Austin area alone (Van Metre et al., 2006).

As awareness of the environmental consequences of PAH increases, we can expect new government and citizen lead initiatives for remedial actions aimed at reducing the sealants' harmful effects. As of April 2009, the City of Austin, Wisconsin's Dane County, and Washington D.C. have instituted a ban on coal tar sealant use. An initial step in such a remediation effort is to produce a geographic inventory of seal-coated parking lots within an area of interest. Such an inventory would allow for detection and prioritization of hot spots based on objective spatial factors such as aerial coverage,

distance to nearest stream, and overlap with groundwater recharge zones. The PAH content varies widely from one type of sealant to another (Mahler et al., 2005), therefore, a parking lot sealant inventory will need to distinguish among the sealant types if we are to accurately identify potential high-level PAH sources. The traditional way to identify parking lot sealant type is to manually scrape off a sample and have the sample chemically analyzed. Although the lab analysis is fairly straight-forward (Kershaw, 1996, Van Metre et al., 2009), they are expensive and the manual field collection process is laborious. Since a large parking lot may have only a portion of its surface sealed, and the sealed portion may have a variety of sealants, it is necessary to sample large parking lots in multiple places in order to assess the PAH-rich sealant coverage area accurately. Even a medium sized city would involve collecting, analyzing, tracking, and georeferencing thousands of samples.

Remote sensing techniques have the potential to greatly simplify obtaining a census of sealed parking lots. Detecting sealant types by analyzing remotely sensed imagery can replace the need for large scale chemical analysis. In addition, remotely sensed imagery can be directly fed to GIS systems for spatial analysis, reducing manual data entry effort (Benz et al., 2004, Ward et al., 1999). Besides simplifying spatial analysis, remote sensing analysis can also assist in tracking temporal changes in parking lot surfaces by analyzing pixel-by-pixel changes in images of a particular location taken at different times (Walsh et al., 2001).

Two classification techniques, Maximum Likelihood Classification (MLC) and Classification Tree Analysis (CTA), are used to differentiate coal tar surfaces from non-coal tar surfaces in Landsat and Hyperion images. Of the two classification techniques,

MLC is regarded as the more accurate but also more computationally demanding method. Of the two remote sensing platforms, Landsat images are far cheaper and more readily available than Hyperion images. Hyperion images, on the other hand, have much greater radiometric and spectral resolutions than Landsat images.

Differentiating dark and relatively small objects such as parking lots is a challenging remote sensing application. To gauge the usefulness of our image classification processes, an accuracy assessment of the two methods and the two remote sensing platforms will be made. Comparing the capabilities of the two classification methods and the two platforms could help us point the way to an adequately robust set of remote sensing tools for our application.

CHAPTER 1

CONCEPTUAL FRAMEWORK

The environmental and health effects of polycyclic aromatic hydrocarbons (PAHs) have been widely documented in the scientific literature (Eisler, 1987, Fabacher et al., 1991, O'Conner and Huggett, 1988). The U.S. Environmental Protection Agency (2008) characterizes PAHs as "highly potent carcinogens that can produce tumors in some organisms at even single doses… their effects are wide-ranging within an organism and have been found in many types of organisms, including non-human animals, birds, invertebrates, plants, amphibians, fish, and humans."

PAHs are a family of chemicals formed when hydrocarbons and other organic substances are burned incompletely. There are over 100 varieties of PAHs and they normally occur as complex mixtures. In a pure form, PAHs are mostly colorless, white, or pale yellow-green. But unless distilled into a pure form in a laboratory, PAHs are almost always found in very dark substances such as coal tar pitch (a byproduct of producing coke from coal), asphalt, roofing tar, and crude oil (US Dept. of Health and Human Services, 1995). PAHs with smaller molecular weights are more volatile than PAHs with larger molecular weights.

Although there are a variety of urban sources of PAHs such as tire dust, motor oil, and asphalt, coal tar sealed parking lots may be the dominant source of PAHs in urban runoff and streams (Mahler et al., 2005, Van Metre et al., 2009). The average PAH content in runoff from coal tar-sealed surfaces was 65-times higher than from other parking lot surface types (Mahler et al., 2005). The adverse wildlife impact of coal tar sealed parking lots raises the need to build an inventory of coal tar sealed surfaces with their location, surface area, and if possible, the "quality" of the surface in terms of wear and coal tar content. Remote sensing analysis may be able to play an important role in building such an inventory. Remote sensing techniques can be used to cost-effectively survey large geographic areas and is routinely used to classify land cover types on a continent-wide scale (Jensen 2005). Remote sensing analysis is used to identify a wide variety of surface materials such as vegetation types, soils, rocks, ice, water, buildings, and pavement (Christensen et al., 2000, Dalton et al., 2004).

To aid remotely sensed mineralogical surveys, the United States Geological Survey (USGS) maintains a publicly available spectral library of a large variety of minerals (Clark et al., 2007). A spectral library of parking lot surfaces, similar to the USGS spectral library of minerals, can potentially be employed to distinguish different types of parking lot surfaces captured in remotely sensed images. In particular, remote sensing techniques may help distinguish coal tar sealed parking lots from non-coal tar sealed parking lots.

Despite the potentially adverse health impact of coal tar sealed parking lots, there do not appear to be any published reports describing attempts to identify coal tar sealed parking lots by analyzing remotely sensed imagery. However, astronomers, oil

prospectors, and land remediation specialists have attempted to identify coal tar and other PAH-bearing molecules through remote sensing techniques. By analyzing spectral absorption bands, astronomers have detected PAH molecules in moons orbiting Jupiter and Saturn and in interstellar gas clouds (Bernstein et al., 2005, Cloutis, 1989). Closer to home, oil prospectors have utilized the unique spectral characteristics of coal tar to prospect for oil bearing tar sands in the Canadian Rockies (Cloutis, 1989), and to detect oil contamination in soils (Winkelmann, 2005). The sampled wavelength region is between 400nm to 2,500nm (Winkelmann, 2005). In hyperspectral images, the sampled spectrum is divided into 100 to 300 bands with a bandwidth of two to ten nm per band (Winkelmann, 2005).

Although PAH-bearing coal tar is very dark over all the spectral bands used in remote sensing platforms, coal tar's reflectance is somewhat higher in the near-UV around 350nm (Cloutis, 1989). Imagery containing near UV-bands is not available via satellite based sensors due to severe atmospheric scattering. However, some airborne sensors do include UV-A bands (315nm to 400nm) and are primarily used for imaging oil spills on water (Jha et al., 2008). Therefore, airborne sensors capable of detecting near-UV radiation may be useful in differentiating coal tar sealed parking lots.

Since parking lots are found mainly in urban areas, the mapping and discrimination of parking lot surface types is an urban application. Urban mapping presents unique challenges for remote sensing analysis because of small features (<10m) and also due to large variations in materials and textures (Herold et al., 2004). Various studies suggest a spatial resolution of 5m or better for accurate mapping of urban land cover (Herold et al., 2003). Adequate spectral resolution and wide spectral coverage are also essential.

Herold (2003) indicates that wavelengths around 498nm (AVIRIS channel 14), 538nm (AVIRIS channel 18), 580nm (AVIRIS channel 22), 640nm (AVIRIS channel 28), 740nm (AVIRIS channel 41), and 2330nm (AVIRIS channel 207) describe 99% of the spectral variability in all datasets. While AVIRIS and Hyperion sensors' band centers are within 5nm of each of the six wavelengths mentioned above, the Ikonos and Landsat TM bands lie outside or near the edge of the suitable bands. Herold (2003) also shows that natural surfaces require less spectral resolution than urban landscapes.

Within an urban setting, dark asphalt and parking lot surfaces present their own set of remote sensing challenges. Producer's accuracies for parking lots are uniformly low at just under 40% for AVIRIS, Ikonos, and Landsat TM (Herold et al., 2003). Producer's accuracy for dark asphalt road is slightly better for AVIRIS at 55% (Herold et al., 2003). Ikonos and Landsat TM show modest improvements in user's accuracies for both parking lot and dark road asphalt (Herold et al., 2003). However, AVIRIS shows marked improvements for parking lot and dark asphalt user's accuracies at 95% and 85% respectively (Herold et al., 2003). The markedly higher AVIRIS user's accuracies for mapping dark asphalt (roads and parking lots) suggest the usefulness of higher spectral and spatial resolutions. The uniformly low reflectance of parking lot surfaces point to the need for high radiometric resolution. We would therefore expect Landsat TM scenes, with 8-bit radiometric resolution, to perform poorly when compared to Ikonos (11-bit), AVIRIS (16-bit), and Hyperion (16-bit) scenes. Another complication in discriminating urban surfaces is introduced by changes in spectral properties over time. Asphalt surfaces (and perhaps sealed parking lot surfaces) become more reflective across all bands as they age (Herold and Roberts, 2005). Due to parking lot size and reflectance

properties, differentiating coal tar sealed parking lots from other kinds of parking lots will likely require high spatial, spectral, and radiometric resolutions.

CHAPTER 2

DATA AND METHODOLOGY

The process of distinguishing coal tar from other kinds of paved surfaces in remotely sensed imagery begins with collecting spectral samples of coal tar and non-coal tar parking lot surfaces. The spectral samples are used as pure-pixel training data to classify pixels in a remotely sensed image as either "coal tar" or "non-coal tar." Finally, a portion of a classified image is compared against known surface types to assess the accuracy of the classification process. The entire methodology can be summarized into six basic steps (Table 2.1).

Surface type spectroscopy

Spectral signatures of coal tar sealed and non-coal tar sealed parking lots are the only primary data that will be collected. A minimum of 15 coal tar sites and at least another 15 non-coal tar sites will be sampled. Sampling will occur across a wide swath of Austin. The spectral device used is FieldSpec Pro, a backpack held instrument manufactured by Analytical Spectral Devices, Inc. To keep lighting conditions consistent, all sampling will take place within a period of two weeks on sunny to mostly sunny days. To correct for changing light conditions, the spectrometer will be recalibrated by retaking White Reference Measurement (WRM) at least every five

minutes. The spectrometer will be set to take one reading per sample. To increase sampling accuracy, 20 readings will be taken for each WRM. The spectrometer's sensor will be held approximately 1m above ground (Figure 2.1). Notes associated with each sampling will include time, location address, and surface type. Sampling locations and their surface types will be provided by the City of Austin based on their lab analysis of parking lot surfaces. Spectral signatures of coal tar sealed surfaces, surfaces sealed with non-coal tar based sealants, plain asphalt and concrete will be sampled. Since sealed surfaces erode significantly within three years after application, sealed surfaces with a wide range of wear will be sampled.

The software provided with FieldSpec Pro produces two kinds of output files: an image file (JPEG, TIFF, and BMP formats) showing graphs of spectral values and a delimited text file listing user-specified mnemonic for the sampled location, band, and reflectance value. The text file is formatted in way that can be conveniently imported into a database or spreadsheet for further analysis.

FieldSpec Pro measures spectral irradiance with wavelengths from 350nm to 1050nm. The spectral resolution is 1nm. The radiometric values are recorded to a precision of 3 decimal places and range from 0.000 to 1.000. The fiber optic cable has a 25 degree field-of-view (FOV). Thus at 1m height, the FOV has a 22 cm radius.

Image and ancillary data

Besides spectral signatures, the data sets needed for the study include Landsat and EO-1 Hyperion scenes covering all or most of Austin, a list of parking lots and their

characteristics provided by the City of Austin, and a GIS polygon layer that describes the taxable structures (in particular, parking lots) within Austin's city limits.

The Landsat scene for Austin, provided by the US Geological Survey (USGS) National Map Seamless Server was acquired on October 12, 2007. The EO-1 Hyperion scene for Austin, provided by USGS Earth Resources Observation and Science (EROS) Center, was acquired on March 30, 2003. Both scenes have a spatial resolution of 30m. The spectral overlap with FieldSpec Pro's range (350nm to 1050nm) is somewhat broader for Hyperion (from 430nm to 1050nm, with no gaps) when compared to Landsat (from 450nm to 900nm, with 10nm, 20nm, and 90nm gaps between four bands). While the spatial resolution and spectral coverage for the two platforms are comparable, Hyperion's spectral and radiometric resolutions are far greater than Landsat's. Hyperion bands are 9nm wide with 69 bands overlapping FieldSpec Pro's spectral range. The corresponding number of bands for Landsat is just four and each band is much wider (450nm - 520nm, 530nm - 610nm, 630nm - 690nm, and 780nm – 900nm). Hyperion has a 16-bit radiometric resolution with 65,536 possible values while Landsat has an 8-bit resolution with 256 possible values. Hyperion is currently regarded as an experimental platform.

In addition to the Hyperion and Landsat datasets, a City of Austin TCAD 2003 GIS layer will be acquired. This layer contains polygons for taxable structures within the Austin metropolitan area. Each polygon has a set of associated attributes including the type of structure represented by the polygon. One of the structure types is "parking lot." Thus the TCAD 2003 layer can be used to clip just the parking lots in a scene and greatly reduce the amount of data that would need to be processed.

Another dataset acquired for this project is a list of parking lot locations and their surface characteristics. This dataset is provided by the City of Austin's Watershed Protection Department and is used to identify coal tar sealed surfaces in an accuracy assessment. The Watershed Protection Department determined the parking lot characteristics by chemically analyzing the parking lots' surface samples.

Methodology

Once data acquisition is complete, the data will be processed in four steps: data reduction, spectral signature band convolution, spectral classification training, and accuracy analysis.

Data reduction

Data reduction decreases the amount of data that need to be processed. In this application, data reduction will be accomplished in three ways: scene filtering, band exclusion, and Principle Component Analysis (PCA). Remotely sensed scenes are typically measured in gigabytes and data reduction makes the processing of large scenes more manageable by reducing disk space and computing requirements. Apart from reducing the quantity of data, data reduction also reduces the number of data processing steps. For example, the exclusion of each Hyperion band also eliminates the tasks associated with processing the band's data.

Scene filtering removes the parts of a scene that are not relevant to the application. For our application, the goal is to distinguish coal tar parking lots from non-coal tar parking lots. Therefore all non-parking lot surfaces can be ignored. The City of

Austin TCAD 2003 polygons attributed as parking lots are used to clip the parking lot surfaces from a scene.

Band exclusion removes unusable portions of a scene's spectral data. An EO-1 Hyperion image has 220 bands ranging from 430nm to 2400nm. Since the FieldSpec Pro spectral range is from 350nm to 1050nm, all bands in a Hyperion image from 1050nm to 2400nm can be ignored. Thus, a 220 band Hyperion image is reduced to sixty-nine 9nm wide bands between 430nm and 1051nm.

Some of the 69 bands of a Hyperion image between 430nm and 1050nm are likely to be highly correlated with each other. Highly correlated bands are redundant in that any one band contains nearly all the information present in the other bands it is highly correlated with. One of the outputs of PCA analysis is a correlation matrix among the image bands. This matrix will be used to identify Hyperion bands that are the least correlated with each other. The set of bands selected for analysis will have at most a correlation coefficient of 0.96 from any of the other selected bands.

DN to reflectance conversion

The pixel values of Landsat and Hyperion imagery is a *digital number* (DN) representing the amount of light energy received by the remote sensing device when the scene was captured. The values are scaled to the minimum and maximum values that can be recorded by the receiving device. Landsat values range from 0 to 255 and Hyperion values range from 0 to 65,535. The DN values must be converted to reflectance values between 0 and 1 before they can be compared to the training site reflectance values collected by the spectrometer. The DN to reflectance conversion is a two-step process.

The DN values are first converted to radiance values which are in turn converted to reflectance values. The obtained Hyperion and Landsat images appear free of clouds and haze, so no haze reduction was performed. The standard atmospheric model was used to calculate reflectance. There are no units associated reflectance values. A zero value denotes full absorption and one denotes full reflectance.

Radiance is a measure of the incoming light's energy intensity. The radiance units are W/(m^2*Sr*μm) or watts per square meter per steradian per micrometer. To convert Hyperion DN values to radiance, DN values in bands 15 – 56 are divided by 40 and DN values in bands 78 – 89 are divided by 80. For Landsat, the following formula was used to convert DN values to radiance (provided by USGS EROS):

$$Radiance\lambda = Bias\lambda + (Gain\lambda * DN\lambda)$$

where λ is the Landsat band number.

The gain and bias for each band are:

band 1: gain: 0.762824, bias: -1.52

band 2: gain: 1.442510, bias: -2.84

band 3: gain: 1.039880, bias: -1.17

band 4: gain: 0.872588, bias: 1.51

The formula to convert radiance to reflectance is (provided by USGS EROS):

$$\rho = (\pi * L\lambda * d^2) / (ESUN\lambda * \cos\theta s)$$

where ρ is reflectance value (unitless)

L is radiance value

λ is band number

d is the earth to sun distance in astronomical units (based on scene's date)

ESUN is the solar irradiance in $W/(m^2 * \mu m)$

s is solar zenith angle in degrees (based on scene's date and time).

d and s were obtained from published astronomical charts. ESUN values for each band were obtained from USGS web site for Hyperion and Landsat FAQ sheet.

Spectral signature band convolution

The Fieldspec Pro spectral reflectance values, recorded with 1nm resolution, are convolved to relevant Hyperion or Landsat bands. Variations in a signature's reflectance values within a band are lost. A large number of narrow bands will result in smaller information loss than fewer, wider bands. Therefore, Hyperion's finer spectral resolution, compared to Landsat, will likely result in greater classification accuracy for Hyperion data. For example, a signature's spike followed by a dip of similar amplitude spanning 450nm to 520nm (Landsat's blue band) would be lost in a Landsat image but would be largely preserved in a Hyperion image.

Scene values bias correction

The reflectance values of a sample obtained from Fieldspec Pro are for a homogenous surface type. The sample values are used as training sites representing "pure pixels" for scene classification. For a variety of reasons, scene pixel values may not match sample values. One reason for the mismatch may be due to inherent bias in the

instruments. Another cause for mismatch is a 30m by 30m patch of an actual parking lot is almost never a homogenous surface. Parking lots typically contain cars, concrete curbs, and stripe markings. Since most things are more reflective than asphalt and parking lot sealants, we would expect pixel values of parking lot scenes to be higher than the values for pure asphalt or sealants.

For scene classification to work properly, the sample and scene values for a particular class need to match. If sample and scene values do not match, then a correction must be applied either to the sample values or to the scene pixel values. A correction can be applied by either adding or multiplying values by some factor. Multiplication is preferable to addition because multiplication preserves key distribution curve characteristics such as mean to SD ratio. Preservation of mean to SD ratio of the distribution curve is necessary for some classification techniques.

For this application, we choose to adjust samples values rather than a scene's pixel values since sample images are smaller than scene images. The adjustment ratio is calculated by dividing the median value for a scene's band with the corresponding median value for the samples. If the variation in adjustment factors among the bands is small, the process can be simplified by multiplying all sample values by an equal amount. Some possible extensions and refinements of this process are explored in the Discussion section.

Spectral classification training and analysis

Two classification methods will be applied on the Landsat and Hyperion scenes: Decision Tree (also known as Classification Tree Analysis, or CTA) and Maximum

Likelihood Classification (MLC). Both methods are known as "hard" classifiers because all pixels are categorized into exactly one class.

Classification Tree Analysis is a rule based classification process where each pixel is subject to a set of rules that classify a pixel to a specific surface category. The training data define the category set. The training data are also used to derive the classification rules. A notable advantage of CTA is it does not depend on sample values to be distributed in any particular way (Jensen 2005). Specifically, sample values need not be normally distributed.

The Maximum Likelihood Classification is based on Bayesian probability theory. This classification method assumes the training data values for each category have a normal (Gaussian) distribution (Jensen 2005). The mean and standard deviation (SD) of pixel values for each category within training sites are computed. The mean and SD values are used to estimate the probability of a scene's pixel belonging to each class category. Each pixel in the scene is then classified to the category with the maximum probability of belonging. When sample values are distributed normally, MLC is often more accurate than other classification methods (Shafri et al., 2007), but the accuracy comes at a high ($O(N^2)$) computational cost (Tso and Mather 2001, 273). The high computation cost can become particularly cumbersome when classifying hyperspectral images because of their large number of bands.

For our application, there are two surface categories or classes: "coal tar" and "non-coal tar." Our training data are the spectral signatures collected using the FieldSpec Pro spectrometer. Each training data pixel represents a contiguous set of Fieldspec Pro reflectance values for a site that has been convolved to a Hyperion or Landsat band. The

scene classifications will be executed using IDRISI Andes. IDRISI Andes expects training data as raster images and does not support importing spectral data directly from a spectrometer. Thus, the usual classification methodology will be altered to use training images created from the spectral signatures instead of pixels representing "pure class" training areas.

Accuracy Assessment

Classification of a remotely sensed scene is rarely error-free. Errors may be introduced in a variety of ways: imperfect optics of the remote sensing platform, poor atmospheric conditions during data capture, scene creation processing errors, and errors introduced by the classification process itself. As discussed in the Conceptual Framework section, differentiating different kinds of parking lot surfaces demands high spatial, spectral, and radiometric resolutions. The resolutions available in Landsat and Hyperion datasets may not be adequate for our application. In addition, the probabilistic nature of both CTA and MLC methods is likely to introduce some number of misclassified pixels. Accuracy assessment quantifies the correctness of the scene *and* the classification process. A scene's pixel may be misclassified in two ways: error of *commission* and error of *omission*. For our application, an error of commission occurs when a non-coal tar area is classified as coal tar. Conversely, an error of omission occurs when a coal tar area is not classified as coal tar.

An *error matrix* (Table 3.4) shows errors of commission and omission for each classification category in a tabular format. The numbers on the diagonal represent the count of correctly categorized pixels for each category. The off-diagonal numbers are

counts of misclassified pixels for each true-category/classified-category pair. Additionally, an error matrix may also provide a statistical measure of overall accuracy (for example, the kappa index of agreement). To assess the accuracy of our Landsat and Hyperion scenes, along with the accuracy of our CTA and MLC methods, an error matrix for each classification method and platform combination (CTA/Landsat, CTA/Hyperion, MLC/Landsat, and MLC/Hyperion) will be computed. To produce an error matrix, the classified image created for each method/platform combination is compared against a *truth* image (also known as a *reference* image). The truth image contains a polygon for each parking lot whose location and surface type is provided by the City of Austin. Each polygon in the truth image is digitized over an USGS DOQQ image and categorized as either "coal tar" or "non-coal tar." The truth image will be in raster format to enable a pixel-for-pixel comparison with the classified images. IDRISI Andes ERRMAT facility will be used to produce the error matrices.

The classification and error assessment methodology is shown in a graphical form below (Figures 2.2 through 2.6).

Table 2.1 Methodology outline. The methodology used to distinguishing coal tar from other kinds of paved surfaces in remotely sensed imagery can be summarized into six basic steps.

Step number	Process summary
1	Collect a statistically meaningful number of spectral signatures of coal tar and non-coal tar paved surfaces.
2	Use the collected signatures to define training data for coal tar and non-coal tar surfaces. The convolved reflectance values based on the training data signatures are used instead of identifying "pure" image pixels to classify paved surfaces.
3	Obtain the latest available Landsat and Hyperion images of the Austin, Texas Metropolitan area.
4	Classify paved surfaces using the Classification Tree Analysis (CTA) and Maximum Likelihood Classification (MLC) methods. Use the two classification methods on the Landsat and Hyperion imagery.
5	Obtain a list of known coal tar and non-coal tar paved surfaces from the City of Austin. The paved surfaces must be located within the classified area.
6	Assess the accuracy of the two classification methods (CTA and MLC) and the two platforms (Landsat and Hyperion) by comparing the classified images against known parking lot surfaces and creating an error matrix for each classified image.

Figure 2.1 Taking a spectrometer reading. The sensor is held about 1m above ground (center). The spectrometer is controlled via the laptop computer (left). The note taker (right) is recording the parking lot's location and surface characteristics.

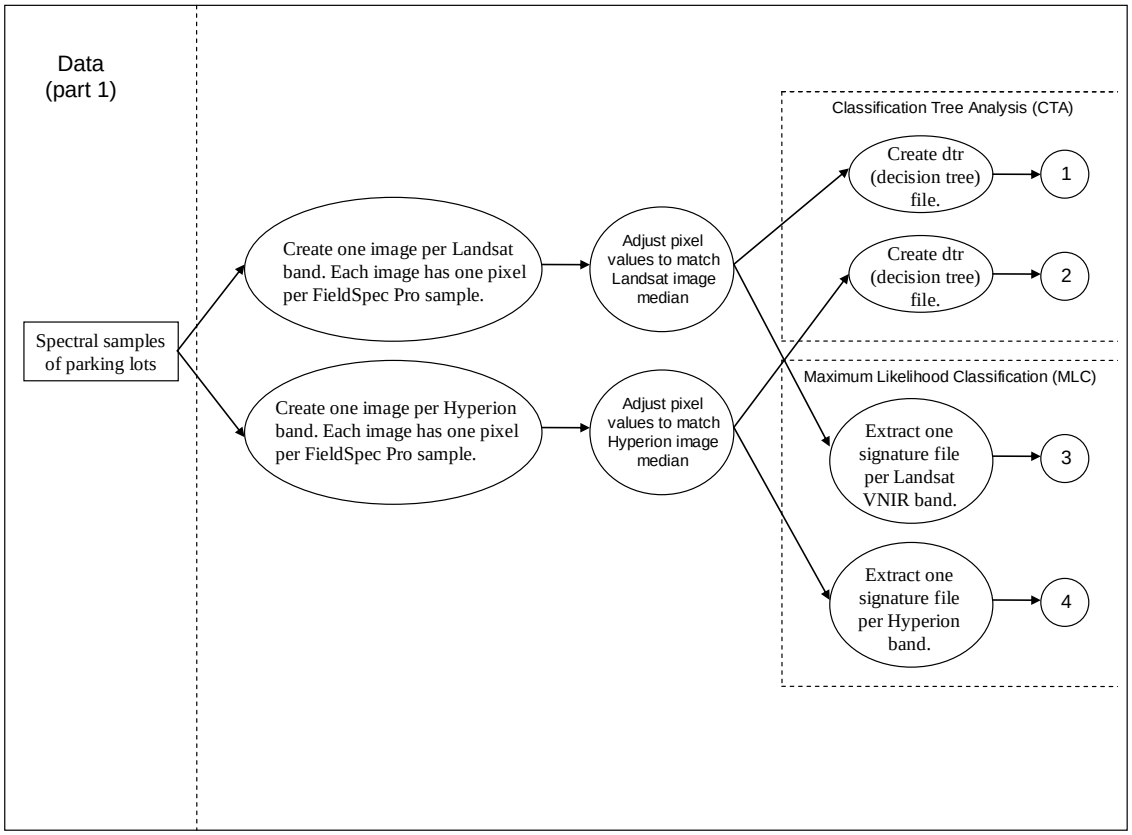


Figure 2.2 Data and methodology in a graphical form. Rectangles represent either an input or an output file. Ovals represent a process. Numbered circles are continuation labels to Figure 2.4.

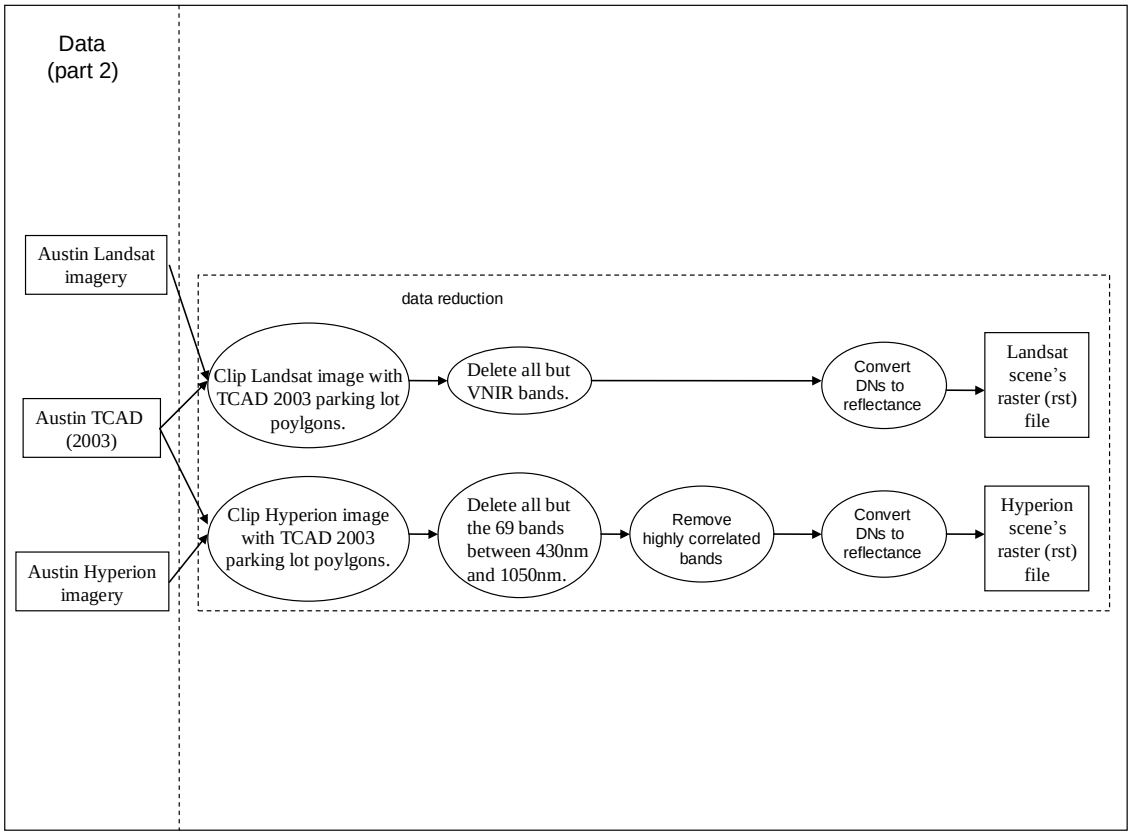


Figure 2.3 Data and methodology in a graphical form. Rectangles represent either an input or an output file. Ovals represent a process.

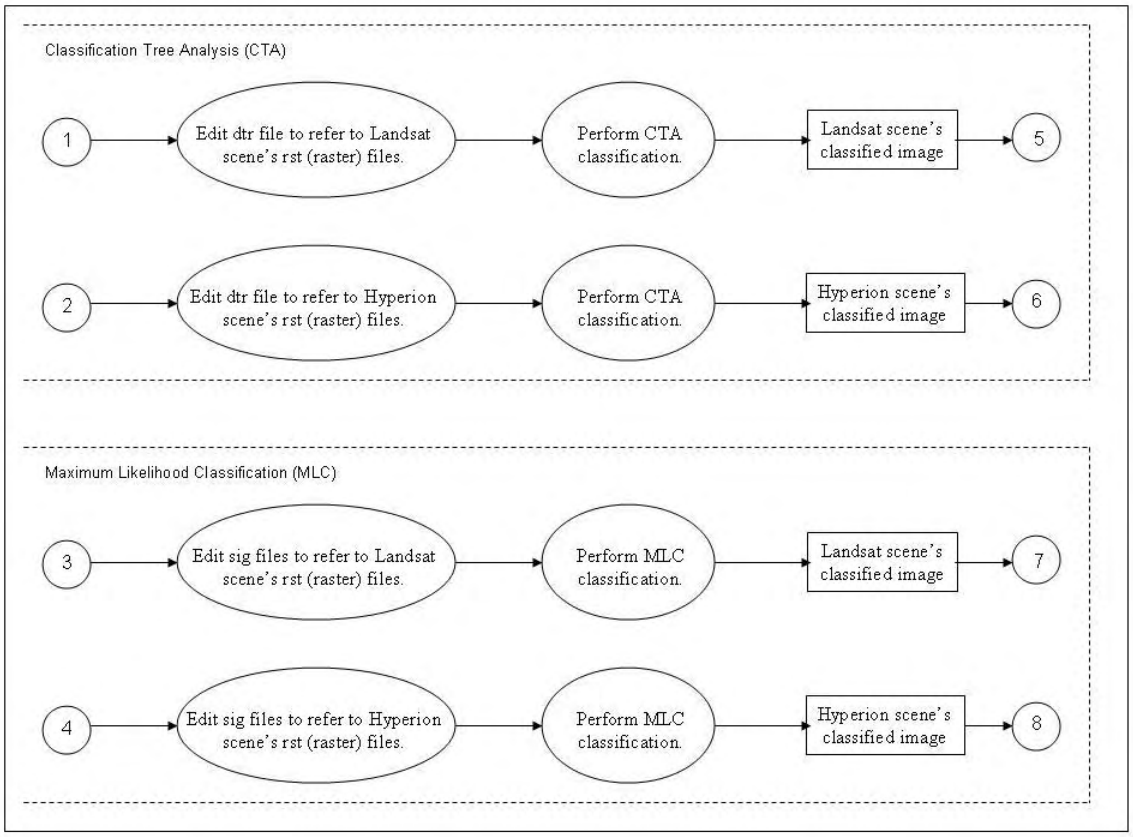


Figure 2.4 Data and methodology in a graphical form. Rectangles represent either an input or an output file. Ovals represent a process. Numbered circles 1 through 4 are continuation labels from Figure 2.2. Numbered circles 5 through 8 are continuation labels to Figure 2.6.

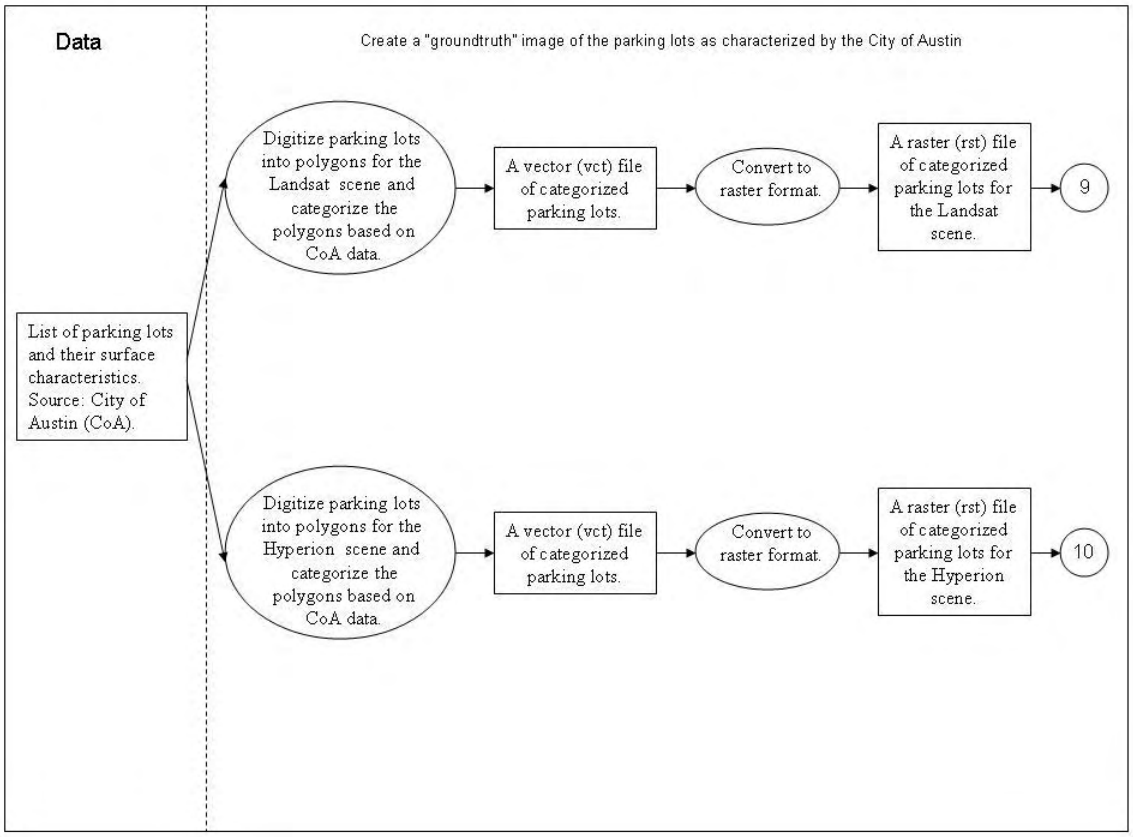


Figure 2.5 Data and methodology in a graphical form. Rectangles represent either an input or an output file. Ovals represent a process. Numbered circles are continuation labels to Figure 2.6.

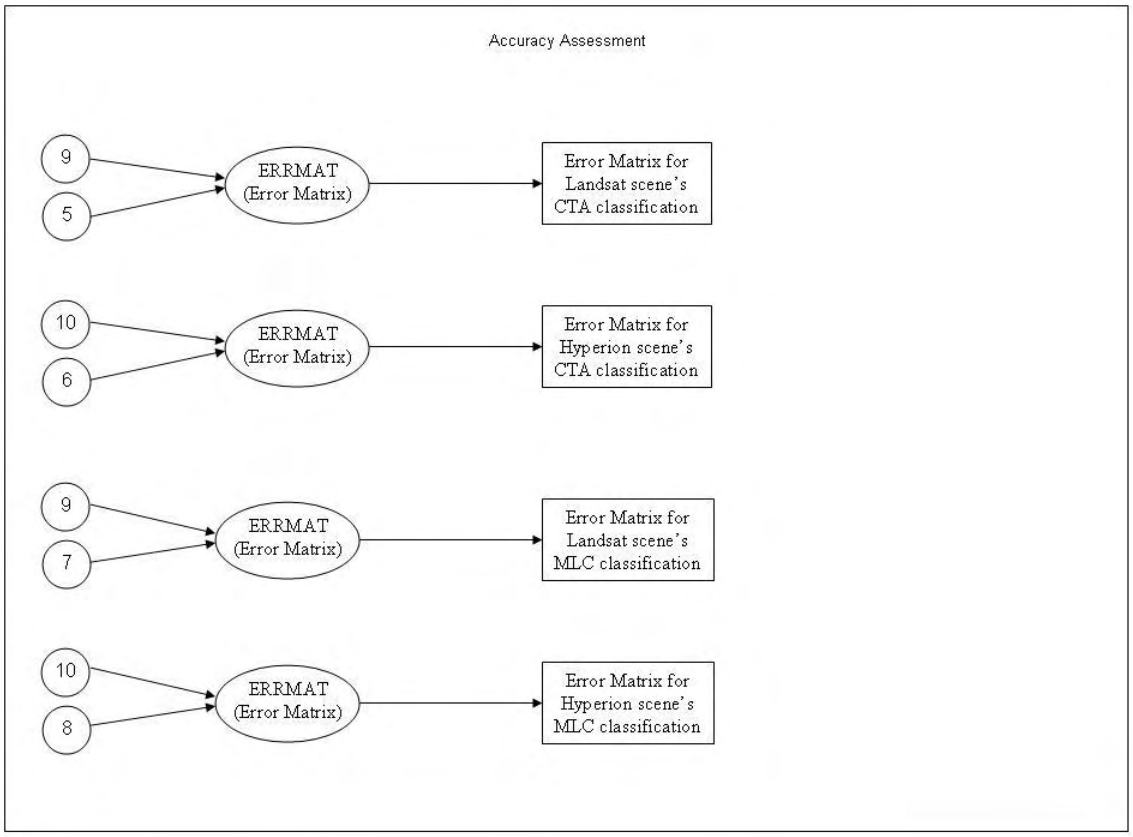


Figure 2.6 Data and methodology in a graphical form. Rectangles represent either an input or an output file. Ovals represent a process. Numbered circles 5 through 8 are continuation labels from Figure 2.4. Numbered circles 9 and 10 are continuation labels from Figure 2.5.

CHAPTER 3

RESULTS

Hyperion band selection

Ten Hyperion spectral bands are selected for performing the classification based on the band exclusion and principal components analyses (Table 3.1). The correlation coefficients among the selected bands are less than 0.95. Since neighboring Hyperion bands are typically highly correlated with each other, there is generally a spectral gap between the selected bands. However, the consecutive bands 35, 36, and 37 are included because their correlation coefficients are smaller than 0.95. VNIR band 56 (920 – 929nm) and SWIR band 78 (929 – 938nm) are also selected despite being spectral neighbors since their correlation coefficient falls below the 0.95 threshold.

Reflectance values

The reflectance values of the Hyperion image (Table 3.1) and the Landsat image (Table 3.2) are closely matched. The distributions of reflectance values for the Hyperion (Figure 3.1) and the Landsat (Figure 3.2) images have a positive skew. The positive skew is progressively stronger in both platforms for shorter wavelengths. The reflectance values along the right tail (the exceptionally bright pixels) for all bands in both platforms are similar to the reflectance values of worn asphalt and concrete samples measured by

Fielspec Pro. The median reflectance values for all bands in both platforms average about three times the corresponding Fieldspec Pro values (Table 3.3). Therefore, Fieldspec Pro reflectance values cannot be compared to Hyperion and Landsat image values without a correction process that matches the training site sample values to the scene values. This correction process is described in the Methodology section.

Parking lot reflectance values measured by Fieldspec Pro

Compared to concrete and weathered asphalt, fresh asphalt and parking lot sealed surfaces are consistently dark across Fieldspec Pro's spectral range (Figure 3.4). Concrete's reflectance is 0.21 at 350nm and rises gradually to 0.35 at 600nm and stays between 0.35 and 0.36 between 600nm and 1050nm. Weathered asphalt's reflectance ranges from 0.12 at 350nm to 0.30 at 1050nm. Sealed surfaces' reflectance values vary from 0.05 and 0.06 for the entire Fieldspec Pro's spectral range.

Parking lot surface wear and weathering increase surface reflectance (Figure 3.5). Unworn Carbonplex sealed surface reflectance ranges from 0.07 at 350nm to 0.04 at 1050nm. But a highly worn Carbonplex spectral curve resembles the spectral curve of the underlying asphalt. Foreign substances on parking lots can significantly change surfaces characteristics (Figure 3.6). Wet algae atop Carbonplex decrease reflectance from 0.07 to 0.01 at 410nm and increase reflectance from 0.06 to 0.14 at 1050nm.

Fresh dark asphalt is increasingly reflective at longer wavelengths (Figure 3.7). In comparison, coal tar sealant's spectral profile (Figure 3.9) is flat across Fieldspec Pro's spectral range. Although sealants are consistently dark from 350nm to 1050nm, there are variations in their spectral profiles (Figure 3.8). Coal tar sealant's spectral profile has a

distinctive dip centered somewhere between 480 and 500nm (Figure 3.9). On average, the reflectance and SD values for coal tar surfaces (Figure 3.9) are lower than for non-coal tar surfaces (Figure 3.10).

Accuracy assessment

The Hyperion and Thematic Mapper images were classified as coal tar or non-coal tar using classification tree and maximum likelihood algorithms (Figures 3.11 and 3.12). The accuracies of these classification techniques are assessed using ground reference data provided by the City of Austin.

Based on the ground reference data, a Hyperion raster image representing the known (or *true*) parking lots is constructed. The image contains 899 pixels classified as "coal tar" and 340 pixels classified as "non-coal tar." Of the 899 true "coal tar" pixels, the CTA technique classified 776 pixels (86.3%) in the Hyperion image as "coal tar" and 123 pixels (13.68%) as "non-coal tar." Of the 340 true "non-coal tar" pixels in the reference image, the CTA technique classified 269 pixels (79.1%) as "non-coal tar" and 71 pixels (20.88%) as "coal tar" (Table 3.4). Omission and commission error rates for "coal tar" pixels are 0.137 and 0.084 respectively. Omission and commission error rates for "non-coal tar" pixels are 0.209 and 0.314 respectively. The overall error rate is 0.157. The overall Kappa Index of Agreement (KIA) is 0.625.

Applying the MLC technique on the Hyperion image, 7.7% of the 899 reference "coal tar" pixels were classified as "coal tar" and 92.3% were classified as "non-coal tar." Of the 340 "non-coal tar" pixels in the reference image, 78.2% were classified as "non-coal tar" and 21.8% were classified as "coal tar" (Table 3.5). Omission and commission

error rates for “coal tar” pixels are 0.923 and 0.517 respectively. Omission and commission error rates for “non-coal tar” pixels are 0.218 and 0.757 respectively. The overall error rate is 0.730. The overall KIA is -0.083.

The Landsat raster image representing the known (or *true*) parking lots contains 88 pixels classified as “coal tar” and 33 pixels classified as “non-coal tar.” Of the 88 true “coal tar” pixels, the CTA technique classified 31 pixels (35.2%) in the Landsat image as “coal tar” and 57 pixels (64.8%) as “non-coal tar.” Of the 33 true “non-coal tar” pixels in the reference image, the CTA technique classified 25 pixels (75.8%) as “non-coal tar” and 8 pixels (24.2%) as “coal tar” (Table 3.6). Omission and commission error rates for “coal tar” pixels are 0.648 and 0.205 respectively. Omission and commission error rates for “non-coal tar” pixels are 0.242 and 0.695 respectively. The overall error rate is 0.537. The overall Kappa Index of Agreement (KIA) is 0.075.

Applying the MLC technique on the Landsat image yields only 5 “coal tar” pixels and none of the 5 pixels overlap one of the reference parking lots. No error matrix is produced.

Spatial analysis

Only the areas of Austin with parcels attributed as “parking lot” in the TCAD 2003 layer were classified. Not surprisingly, the densest cluster of parking lots is in the downtown area. Parking lot parcel density in Austin’s periphery and in the lower-income areas east of IH-35 is lower than in downtown. It is important to note that the peripheral areas of the city do not necessarily have a lower density of parking lot surfaces. Parking lots serving shopping malls and office complexes are not designated as parking lot

parcels and are thus not included in this study. This study, therefore, only considers a subset of all parking lots in the Austin area.

The Hyperion (Figure 3.13) and Landsat (Figure 3.14) CTA classified images reveal a greater concentration of coal tar sealed parking lots west of IH-35 than east of the highway. Assuming parking lot parcels' distribution and surface type is representative of the overall parking lot population in Austin, the CTA classified images show that not only is the density of parking lots higher west of IH-35, but the proportion of coal tar sealed parking lots is also higher west of the highway.

The error/residual maps of classified scenes do not reveal any glaring geographic patterns. However, the Hyperion (Figure 3.15) and Landsat (Figure 3.16) error maps for CTA classification do show that a large proportion of classification errors occur along parking lot edges. The edge errors are likely due to mixed pixels covering a portion of a parking lot's edge and some of the surrounding area. The low accuracy of MLC classification of the Hyperion image is reflected in the corresponding error map (Figure 3.17).

Table 3.1 Selected Hyperion bands. Ten Hyperion bands were selected for this study. Each band is about 10nm wide. The darkest and brightest pixel values for each band are shown. The numbers in parenthesis in the rightmost column are standard deviation values.

Hyperion band number	Wavelength range (nm)	Minimum value	Maximum value	Mean (SD)
17	524 - 533	0.101	0.451	0.162 (0.046)
22	574 - 583	0.097	0.453	0.157 (0.05)
35	707 - 716	0.096	0.542	0.175 (0.067)
36	717 - 726	0.099	0.524	0.191 (0.062)
37	727 - 736	0.104	0.532	0.205 (0.06)
42	779 - 788	0.107	0.572	0.238 (0.064)
56	920 - 929	0.081	0.437	0.203 (0.057)
78	929 - 938	0.077	0.469	0.218 (0.063)
81	958 - 967	0.031	0.326	0.152 (0.044)
89	1040 - 1049	0.061	0.580	0.276 (0.081)

Table 3.2 Selected Landsat bands. Four Landsat bands were selected for this study. The darkest and brightest pixel values for each band are shown. The numbers in parenthesis in the rightmost column are standard deviation values.

Landsat band	Wavelength range (nm)	Minimum value	Maximum value	Mean (SD)
1 (blue)	450 - 520	0.084	0.350	0.131 (0.038)
2 (green)	530 - 610	0.073	0.476	0.127 (0.046)
3 (red)	630 - 690	0.052	0.548	0.123 (0.056)
4 (near IR)	780 - 900	0.120	0.607	0.236 (0.056)

Table 3.3 Hyperion and Landsat mean and median values. Scene mean and median reflectance values for four selected Hyperion bands followed by mean and median parking lot sample reflectance values for the band's wavelength range. The rightmost column is the ratio of Hyperion scene against sample median values.

Hyperion band number	Scene mean value	Scene median value	Spectral sample mean	Spectral sample median	Band median divided by spectral sample median
17	0.162	0.147	0.080	0.054	2.72
35	0.175	0.151	0.092	0.058	2.60
42	0.238	0.216	0.095	0.062	3.48
89	0.276	0.245	0.102	0.061	4.02

Table 3.4 Error matrix for CTA classification of Hyperion image. The bolded numbers represent correctly classified pixels representing coal tar sealed and non-coal tar sealed parking lot surfaces.

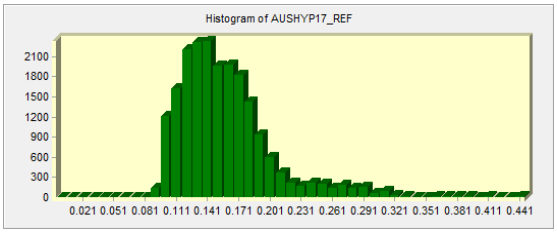
	Non-coal tar	Coal tar	Total (mapped)	Error of Commission
Non-coal tar	**269**	123	392	0.3138
Coal tar	71	**776**	849	0.0838
Total (true)	340	899	1239	
Error of Omission	0.2088	0.1368		Overall error: 0.1566

Table 3.5 Error matrix for MLC classification of Hyperion image. The bolded numbers represent correctly classified pixels representing coal tar sealed and non-coal tar sealed parking lot surfaces.

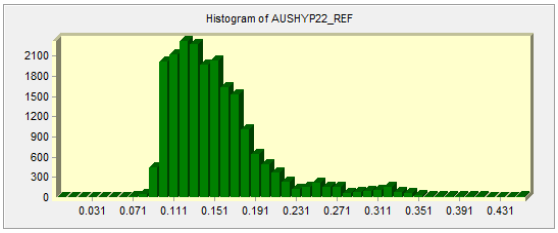
	Non-coal tar	Coal tar	Total (mapped)	Error of Commission
Non-coal tar	**266**	830	1096	0.7573
Coal tar	74	**69**	143	0.5175
Total (true)	340	899	1239	
Error of Omission	0.2176	0.9232		Overall error: 0.7296

Table 3.6 Error matrix for CTA classification of Landsat image. The bolded numbers represent correctly classified pixels representing coal tar sealed and non-coal tar sealed parking lot surfaces.

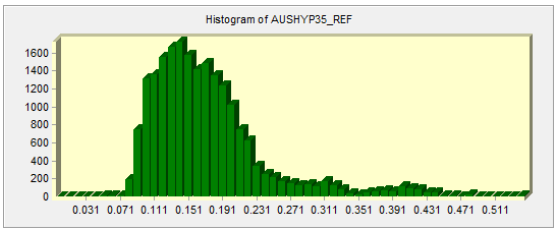
	Non-coal tar	Coal tar	Total (mapped)	Error of Commission
Non-coal tar	**25**	57	82	0.6951
Coal tar	8	**31**	39	0.2051
Total (true)	33	88	121	
Error of Omission	0.2424	0.6477		Overall error: 0.5372



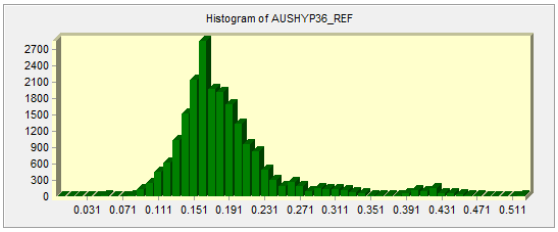
(a) Hyperion Band 17



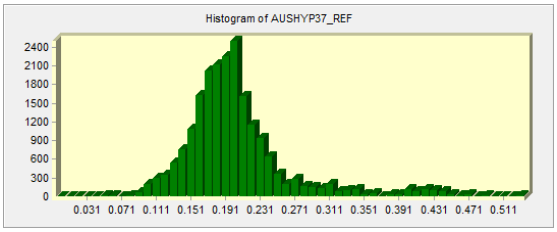
(b) Hyperion Band 22



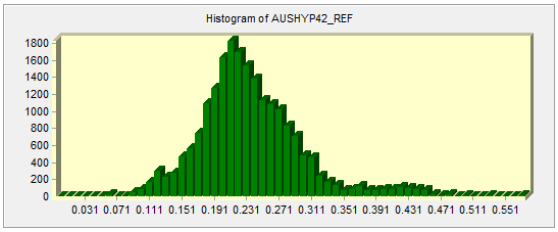
(c) Hyperion Band 35



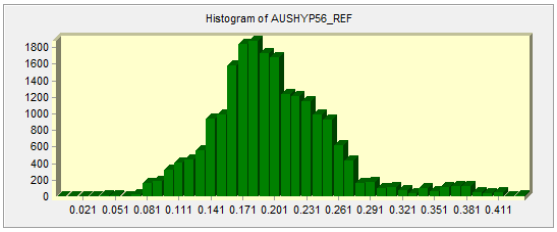
(d) Hyperion Band 36



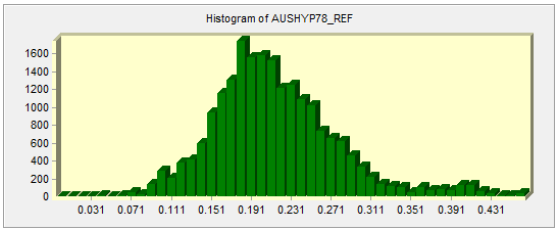
(e) Hyperion Band 37



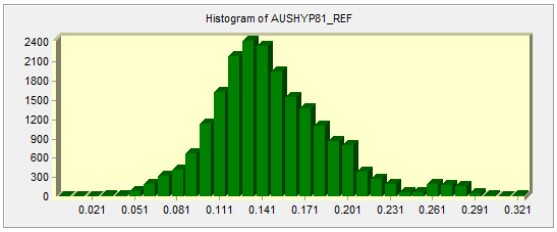
(f) Hyperion Band 42



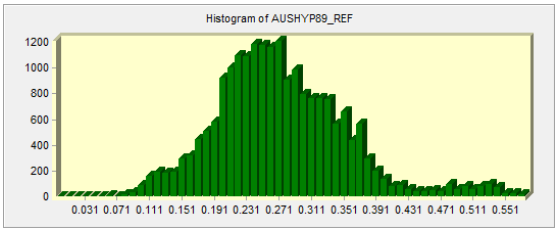
(g) Hyperion Band 56



(h) Hyperion Band 78

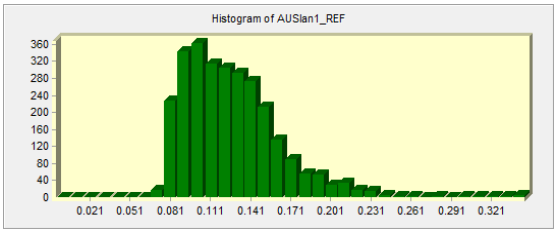


(i) Hyperion Band 81

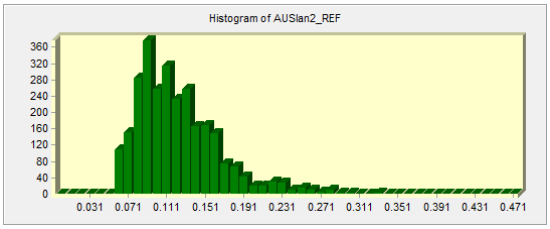


(j) Hyperion Band 89

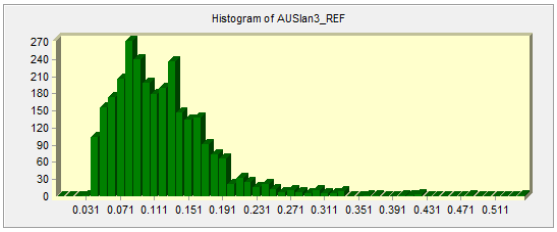
Figure 3.1 Histogram of pixel values of the Hyperion scene for the 10 selected bands. The distribution curves for bands 17, 22, and 35 have a positive skew.



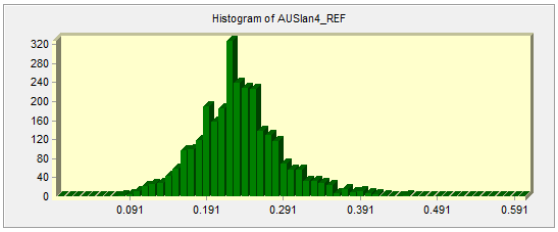
(a) Landsat Band 1



(b) Landsat Band 2



(c) Landsat Band 3



(d) Landsat Band 4

Figure 3.2 Histogram of pixel values of the Landsat scene for the four selected bands. The distribution curves for bands 1, 2, and 3 have a positive skew. Band 3 is bimodal.

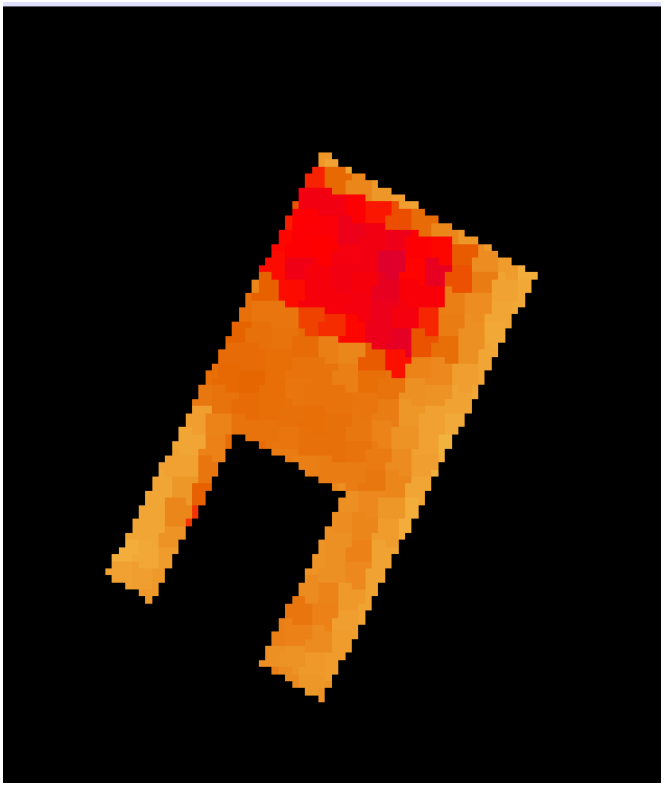


Figure 3.3 Detail from reflectance raster image derived from Hyperion scene's band 35. High reflectance areas are shown in red hues and low reflectance areas are shown in yellow. The black background's pixel values are 0.

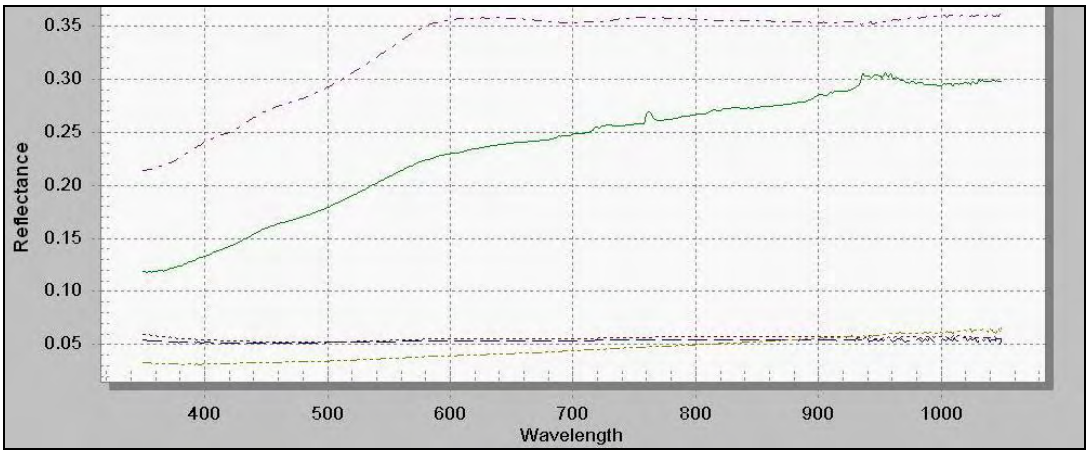


Figure 3.4 Reflectance values obtained from Fieldspec Pro for selected surfaces. The top line is for concrete, the middle line is for weathered asphalt, and the remaining lines are for dark surfaces found in parking lots. The reflectance ranges for the dark surfaces are narrow compared to concrete and weathered asphalt.

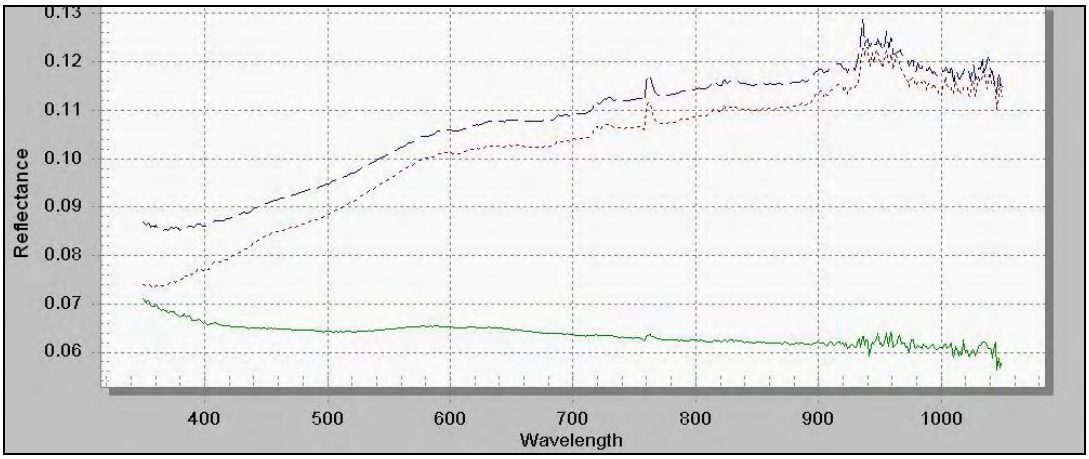


Figure 3.5 Reflectance values obtained from Fieldspec Pro for selected surfaces. Reflectance curves for dark unsealed asphalt (top), highly worn CarbonPlex sealant (middle), and "fresh" CarbonPlex sealant (bottom). As sealant surfaces erode, their spectral reflectance curves start to resemble the reflectance curves of underlying layers.

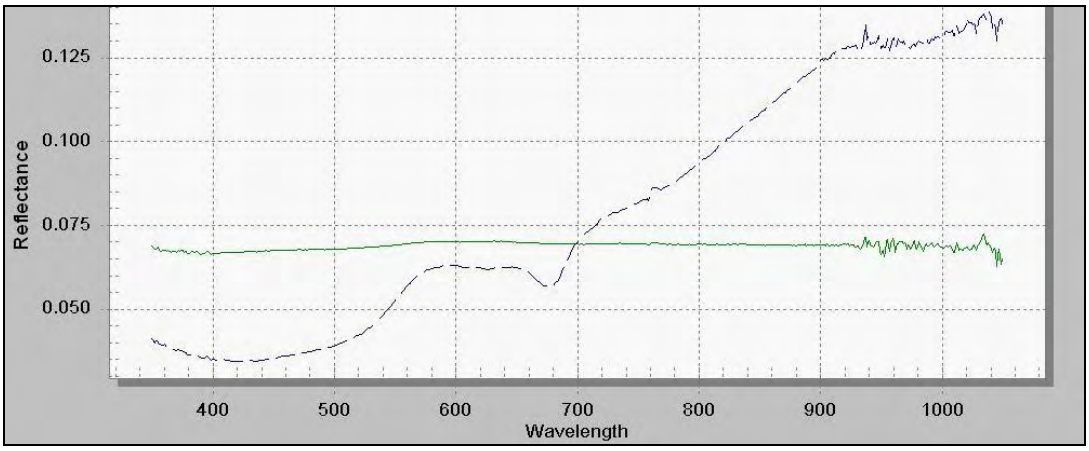


Figure 3.6 Reflectance values obtained from Fieldspec Pro for selected surfaces. Reflectance curves for debris-free CarbonPlex surface (solid line) and for CarbonPlex surface partially covered with wet algae (dashed line).

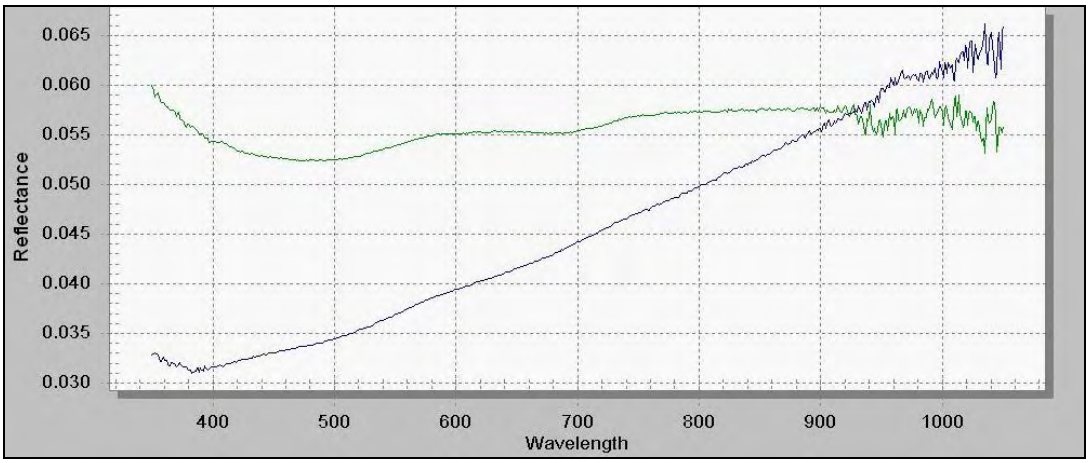


Figure 3.7 Reflectance values obtained from Fieldspec Pro for selected surfaces. Reflectance curves for coal tar sealant (green) and dark asphalt (diagonal line). The two surfaces look nearly identical to the untrained eye.

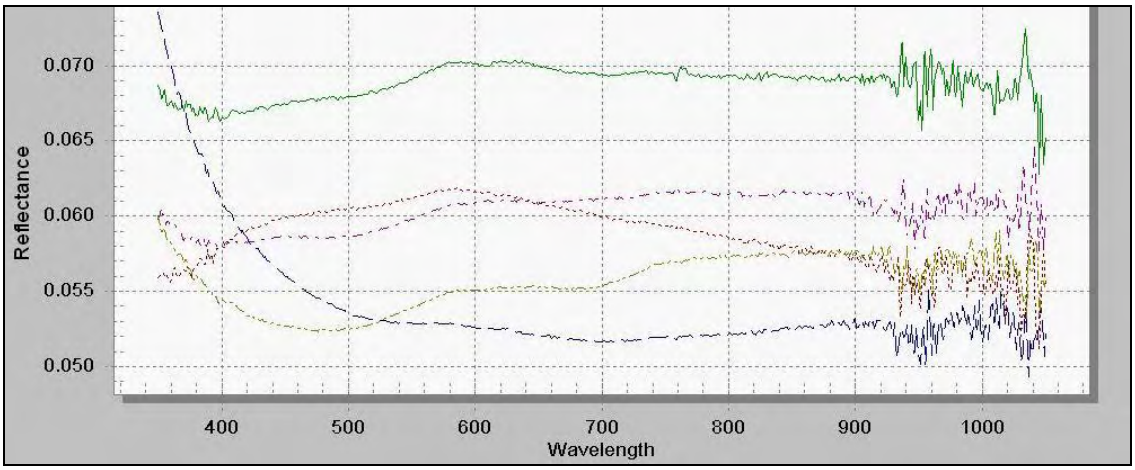


Figure 3.8 Reflectance values of various parking lot sealant surfaces. The curve characteristics show a greater variety in the 350nm to 500nm part of the spectrum.

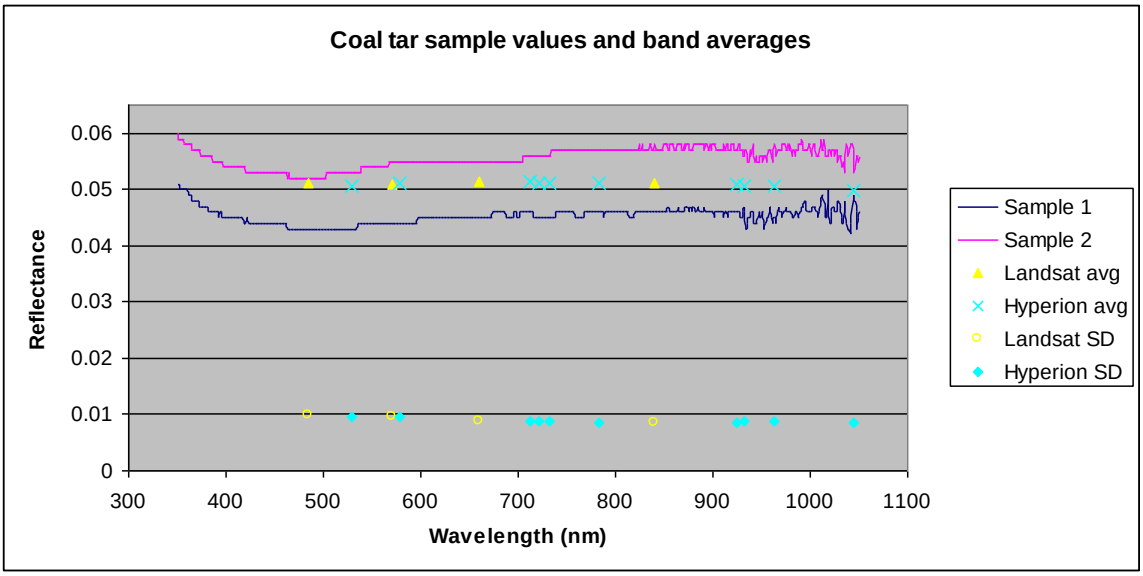


Figure 3.9 Reflectance values of two coal tar sealed surfaces (Sample 1 and Sample 2). Coal tar sealants have a distinctive dip centered around 480nm to 500nm. The four yellow triangles represent the average reflectance values of all coal tar samples for the four Landsat bands. The yellow open diamonds represent the standard deviation (SD) values of all coal tar samples for the four Landsat bands. The ten blue Xs represent the average reflectance values of all coal tar samples for the ten Hyperion bands. The blue solid diamonds represent the standard deviation (SD) values of all coal tar samples for the ten Hyperion bands.

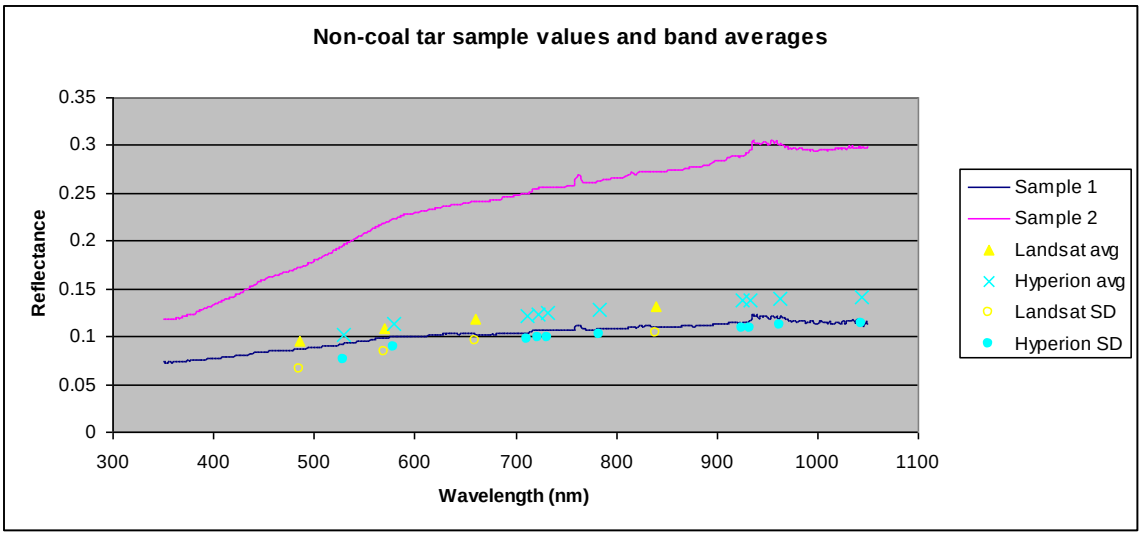


Figure 3.10 Reflectance values of two non-coal tar surfaces (Sample 1 and Sample 2). The spectral characteristics of most non-coal tar surfaces are closer to Sample 1 than Sample 2. The four yellow triangles represent the average reflectance values of all non-coal tar samples for the four Landsat bands. The yellow open diamonds represent the standard deviation (SD) values of all non-coal tar samples for the four Landsat bands. The ten blue Xs represent the average reflectance values of all non-coal tar samples for the ten Hyperion bands. The blue solid diamonds represent the standard deviation (SD) values of all non-coal tar samples for the ten Hyperion bands.

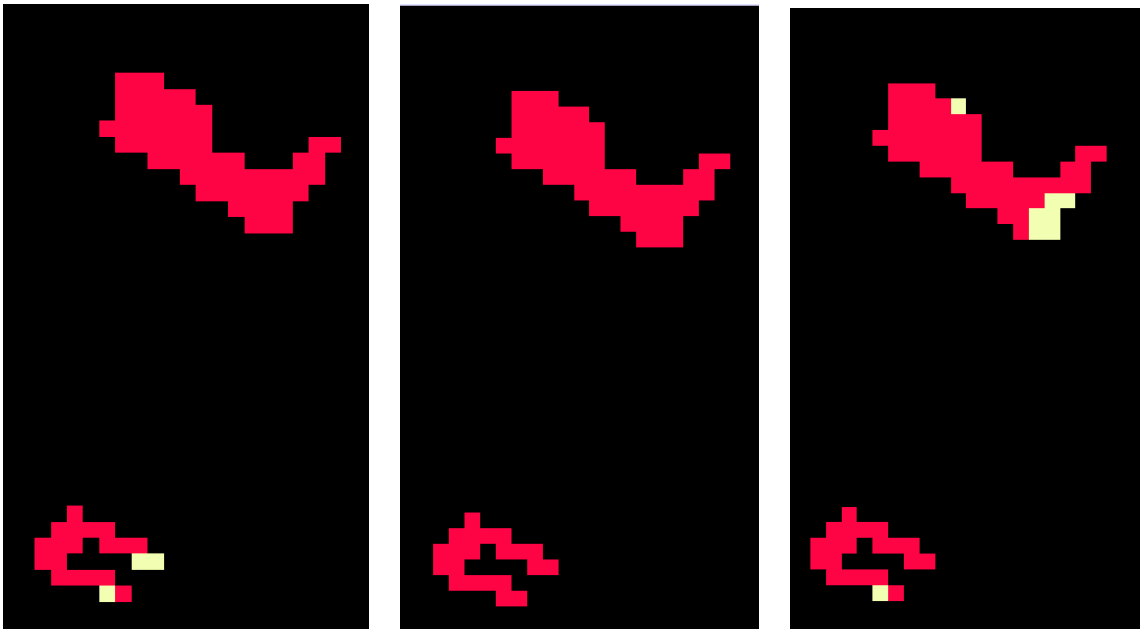


Figure 3.11 Detail from Hyperion classified scenes and "truth" or reference image. The reference image is derived from reference parking lot data. The scenes have two possible classes: "coal tar" (yellow) or "non-coal tar" (red). From left to right: scene classified by the Classification Tree Analysis technique, reference image where the two parking lots shown here are known non-coal tar, and scene classified by the Maximum Likelihood Classification technique.

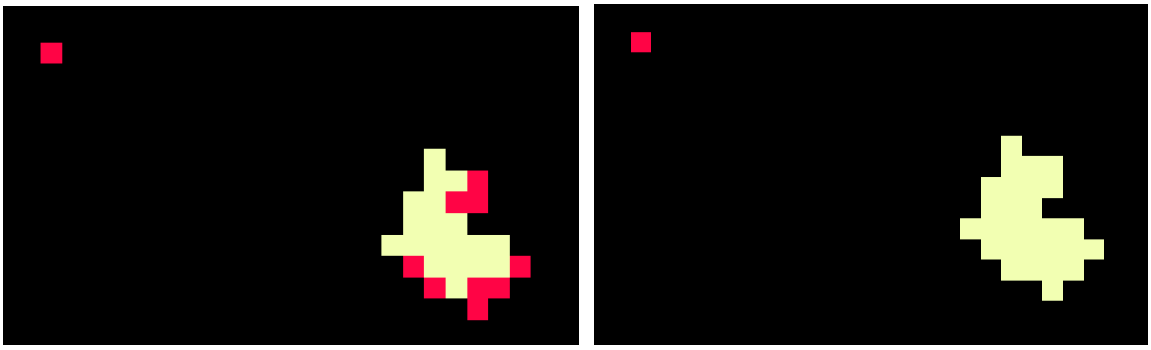


Figure 3.12 Detail from Landsat classified scene and "truth" or reference image. The reference image is derived from reference parking lot data. The scenes have two possible classes: "coal tar" (yellow) or "non-coal tar" (red). The scene on the left is classified by the Classification Tree Analysis technique. The scene on the right is the reference image where the larger parking lot in the lower right is a known coal tar and the smaller parking lot in the upper left is a known non-coal tar.

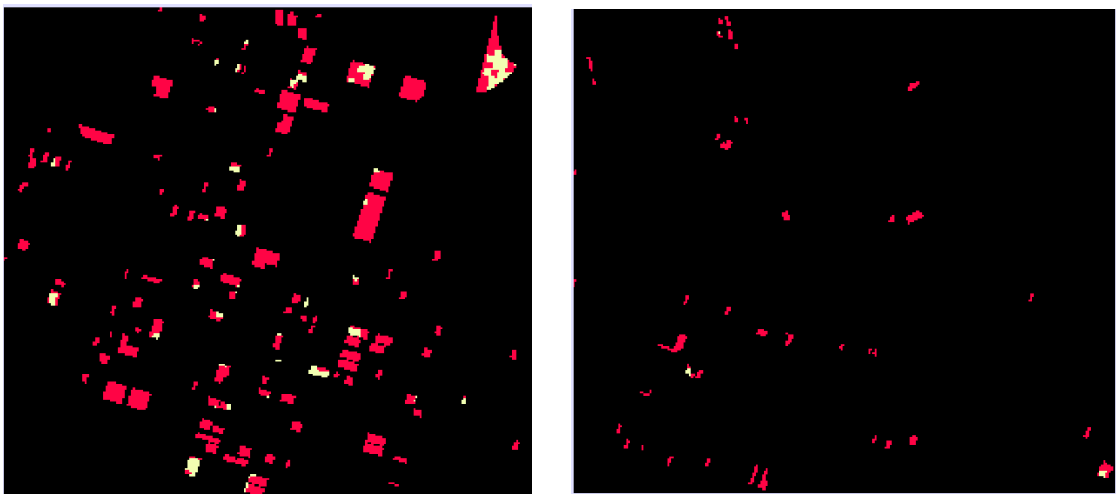


Figure 3.13 Detail from CTA classification of the Hyperion image. There are two possible classes: "coal tar" (yellow) or "non-coal tar" (red). The scene on the left is for downtown Austin west of IH-35 and the scene on the right is for an area directly east of downtown and east of IH-35. Both scenes have similar areal coverage.

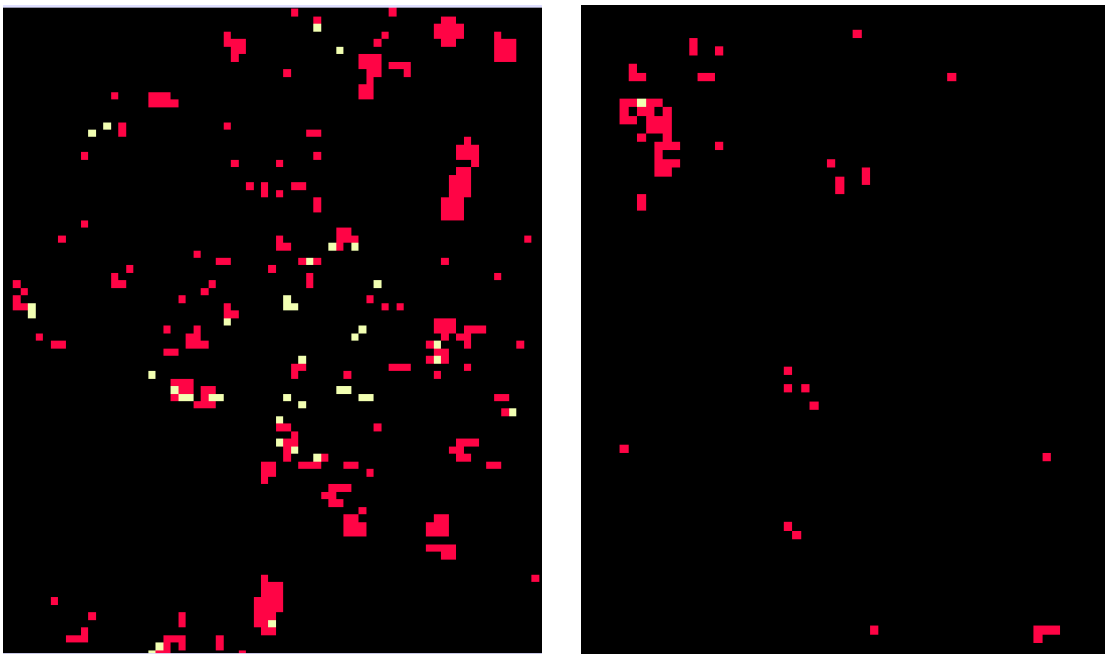


Figure 3.14 Detail from CTA classification of the Landsat image. There are two possible classes: "coal tar" (yellow) or "non-coal tar" (red). The scene on the left is for downtown Austin west of IH-35 and the scene on the right is for east of downtown and east of IH-35. Both scenes have similar areal coverage.

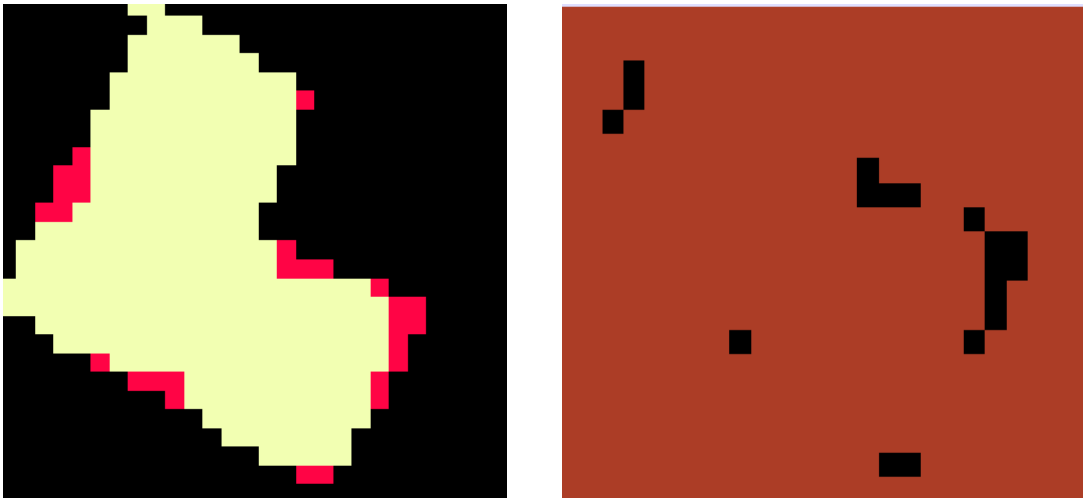


Figure 3.15 CTA/Hyperion error map. Detail from the CTA/Hyperion classified image (left) and the error map for the same area. Yellow pixels represent coal tar surface and red pixels represent non-coal tar surface. Black pixels in the error map represent misclassified areas. The misclassified areas are mostly along the parking lot edges.

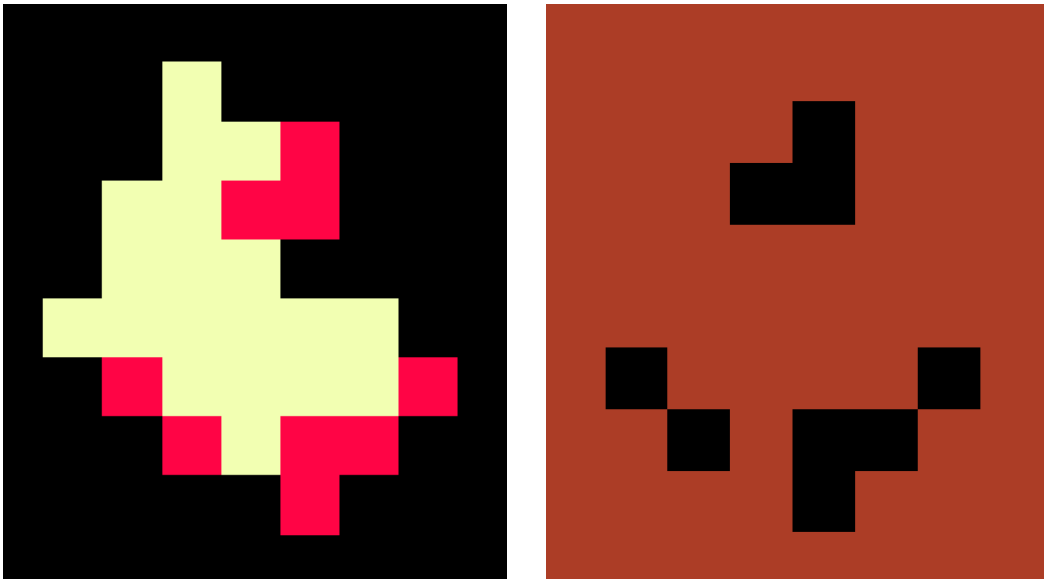


Figure 3.16 CTA/Landsat error map. Detail from the CTA/Landsat classified image (left) and the error map for the same area. Yellow pixels represent coal tar surface and red pixels represent non-coal tar surface. Black pixels in the error map represent misclassified areas. The misclassified areas are mostly along the parking lot edges.

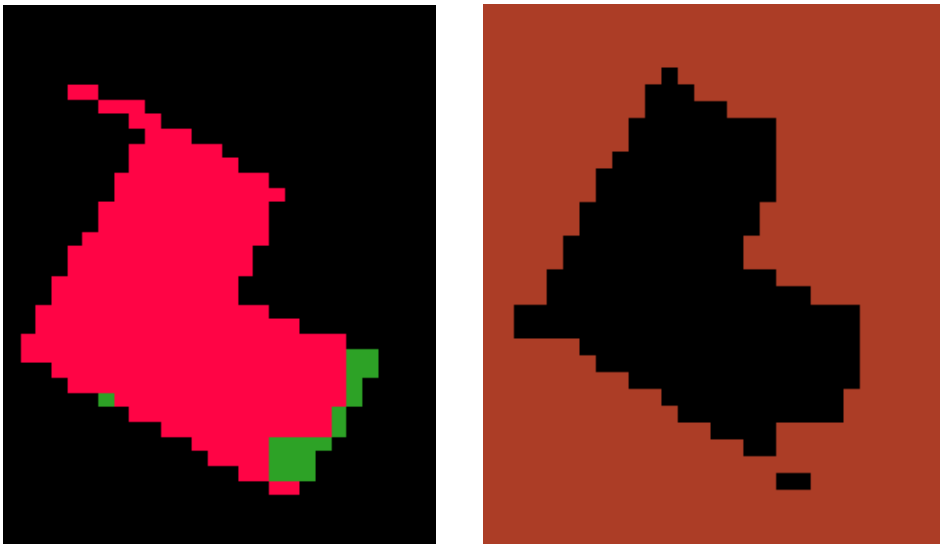


Figure 3.17 MLC/Hyperion error map. Detail from the MLC/Hyperion classified image (left) and the error map for the same area. Red pixels represent non-coal tar surface and green pixels represent coal tar surface. Black pixels in the error map represent misclassified areas. The coal tar sealed parking lot is almost entirely misclassified.

CHAPTER 4

DISCUSSION

As mentioned in the Conceptual Framework section, distinguishing different types of parking lot surfaces is a challenging remote sensing application. The uniformly low reflectance of parking lot sealants and the relatively small areal extents of parking lots require high spatial, spectral, and radiometric resolutions. Not surprisingly, classification of Hyperion imagery, with far greater spectral and radiometric resolutions than Landsat imagery, yielded the most accurate results. More surprisingly, CTA clearly outperformed MLC on both Hyperion and Landsat imagery. CTA classification of Hyperion imagery (CTA/Hyperion) yielded an overall accuracy of 84.34%, or conversely, an error rate of 15.66%. In comparison, MLC/Hyperion overall accuracy was just 27.04%, even lower than CTA/Landsat's overall accuracy of 46.28%. MLC/Landsat did not classify a single pixel in the reference parking lots (parking lots with known surface types) as coal tar. The preponderance of error in MLC classification is due to coal tar being classified as non-coal tar (false-negatives). A partial explanation for MLC's poor performance compared to CTA may be due to pixel value distributions. Histograms of Hyperion (Figure 3.1) and Landsat (Figure 3.2) scenes show positive skew distribution curves. The positive skew is especially pronounced for the shorter wavelength bands. In addition, several bands are multi modal. For example, the

distribution of Landsat's band three (Figure 3.2(c)) has two peaks and MLC requires that data be at least unimodal, if not normally distributed.

A number of uncontrollable factors can lower classification accuracy. For example, parking lot wear and aging change surface characteristics. Coal tar and other sealant materials erode relatively quickly, especially in highly trafficked areas. While a reference parking lot might be classified as coal tar sealed, some sections of the parking lot could have eroded to bare asphalt. Therefore, pixels that have been correctly classified as non-coal tar might be incorrectly assessed as false-negatives. Another source of confusion can occur when a coal tar sealed lot is covered by a non-coal tar sealant (such re-topping is common in Austin, perhaps to comply with local ordinances prohibiting coal tar sealants). Tire wear can eventually expose the underlying coal tar in some areas. In this case, pixels that are correctly classified as coal tar would be assessed as false-positives.

Uncontrollable factors like the ones mentioned above point out the need to have an intimate knowledge of the reference parking lots, both in terms of the lots' current condition and their history. It is at least necessary to know the type of surface a reference parking lot had when the remotely sensed scene was captured. Fortunately, the City of Austin's Watershed Protection and Development Review Department was kind enough to furnish such information. To the extent possible, the reference parking lot categorizations match the surface types of the actual lots when the images were acquired. However, the reference parking lots are assumed to belong to a single class and no attempt was made to identify and re-class worn areas. Though difficult to quantify, this simplification may result in overstating classification errors.

Perhaps the biggest source of confusion is parked cars. Cars are typically far more reflective than parking lots, so even a single car can significantly change a pixel's reflectance value. The Hyperion image was captured at 10:46am (local time) on a Sunday (March 30, 2003). The Landsat image was captured at 11:57am (local time) on a Wednesday (April 25, 2007). Except for church parking lots, Hyperion likely captured the scene when parking lots were generally empty. In contrast, the Landsat scene was collected when many cars were likely within the sensor's field of view. For readily available Landsat data, a future refinement to our methodology could partially get around the parked car problem by analyzing multiple images of a particular parking lot and re-compose a new image by mosaicing the darkest (and hence the most car-free) pixels. This assumes that at some point or another, a 30m x 30m patch of parking lot would be largely empty, even if the scenes are always captured at the same time of the day (both Landsat and Hyperion have sun-synchronous orbits).

Even with cars, more than half of a parking lot's surface likely remains exposed in an overhead image. There is always space between cars and aisle sections are largely empty. Thus, assuming an identical car configuration, a dark parking lot reflects less light than a lighter parking lot. The cars in effect introduce a positive bias to reflectance values and are the likely cause for higher median values in remotely sensed imagery when compared to "pure" sample values. Pixels with high reflectance values, perhaps due to concrete structures or highly worn aisle areas where cars do not contribute a positive bias, are similar to the reflectance values of concrete obtained from Fieldspec Pro. But the bulk of the scene pixels have values that are about three times higher than Fieldcpec Pro values for dark surfaces. The remotely sensed scenes do have a handful of

pixels that match the dark surface Fieldspec Pro values. These values may be due to the occasional 30m x 30m patches of parking lots that are free of cars. These "dark" pixels suggest there was no inherent bias in either the spectrometer or the remote sensing platforms. To remove what appears to be a car-induced bias, the Fielspec Pro values in training sites were multiplied by three to closely match scene pixel median values. This bias-reduction technique could be made more robust by introducing a goal-seeking iterative approach whose goal is to find the multiplication factor that yields the highest classification accuracy. A further refinement could obtain different multiplication factors for each band. Another way to potentially improve accuracy would be to obtain spectrometer values atop a ladder with one or more cars occupying varying proportions of the field of view.

Three hardware limitations are worth mentioning. The Fieldspec Pro spectrometer used for this project has a spectral range from 350nm to 1050nm. Hyperion and Landsat gather data up to 2,400nm and 2,350nm respectively at 30m resolution. But we removed potentially valuable bands beyond 1050nm from the scenes for lack of training data. At the other end of the spectrum, we had to discard Fieldspec Pro data from 350nm to 430nm because neither Hyperion nor Landsat provide data for wavelengths shorter than 430nm and 450nm respectively. This spectral gap is unfortunate since coal tar sealants have a distinctive decrease in reflectivity from 350nm to 480nm (Figure 3.9). Hyperion's spectral data does in fact nominally start at 350nm, but the platform's first eight bands (350nm – 430nm) are un-calibrated and contain no data. Hyperion is still considered an experimental platform and it is possible the data from the first eight bands become available in the future. Finally, 30m spatial resolution

probably makes achieving accuracies much above what has been obtained in this study impossible. If the mixed pixels around the edges are discarded, few parking lots would have more than just a handful of "pure" pixels. In fact, a large proportion of classification errors are near parking lot edges (Figure 3.15 and Figure 3.16). These errors are likely due to the mixed pixels along a parking lot's perimeter. Also, few parking lots have large patches without foreign objects such as cars, trees, and concrete structures. Images with greater spatial resolution would be better suited for this application.

Despite various limitations, the CTA/Hyperion classification accuracy appears promising. Our goal was to distinguish coal tar sealed parking lots from other types of parking lots. To that end, we set out to develop a remote sensing tool to help us guide our search for coal tar sealed surfaces. The CTA classification of the Hyperion image has a false-positive rate of 0.1368 and a false-negative rate of 0.2088. Such accuracies can help guide us to large clusters of coal tar sealed parking lots for further on-ground analysis while skipping over areas with few coal tar sealed lots. The resulting CTA classified image can be input into a GIS tool for further analysis to help prioritize on-ground verification. Even the CTA/Landsat image can be of use. Although the CTA classification of the Landsat image has an unacceptably high false-negative rate (0.6477), the false-positive rate is low (0.2424). Therefore, the Landsat images can be used as an initial screen, where significant clusters of pixels classified as coal tar guide us to areas where more expensive Hyperion imagery could be obtained.

CONCLUSIONS

Coal tar based sealants are widely used in the United States to resurface parking lots. Although the sealant industry's marketing literature tout the aesthetic appeal of freshly coated parking lots, in recent years, a number of studies have cataloged the potential environmental and health dangers of coal tar sealed surfaces. Polycyclic aromatic hydrocarbons (PAHs) are a major component of coal tar and have been classified as a carcinogen by the Environmental Protection Agency. Recent hydrological studies have implicated coal tar sealants as the major source of PAHs in some urban streams. Despite the potential environmental hazards of coal tar sealants, the high cost of on-ground cataloging has prevented a nationwide census of parking lot surface types. We explore the feasibility of using relatively inexpensive satellite imagery to identify and locate coal tar sealed surfaces. The goal is to classify parking lots in a remotely sensed image as either "coal tar" or "not coal tar." The study area is Austin, Texas.

Differentiating parking lot surface types is a challenging remote sensing application that requires fine spatial, spectral, and radiometric resolutions. The remote sensing platforms chosen for this study are Landsat and Hyperion. Landsat images are inexpensive and are readily available, but the platform's spectral and radiometric resolutions are course when compared to hyperspectral imagery. On the other hand, Hyperion imagery, while more expensive than Landsat's, has much higher spectral and radiometric resolutions. The spatial resolution for the two platforms is 30m. Although

satellite imagery with finer spatial resolution is available, they are either more expensive than Landsat, or have coarser spectral and radiometric resolutions than Hyperion.

Pixel values for the two platforms were delivered as digital numbers (DNs). The DNs were transformed into reflectance values using the standard USGS EROS atmospheric models. To make such transformations practical, imagery data was reduced in a number of ways. Bands that do not overlap the spectrometer's spectral range were discarded, and all image files were transformed into integers. The real to integer transformation alone reduced IDRISI Andes file sizes by a factor of four. Finally, highly correlated Hyperion bands were discarded, which reduced the number of bands from 69 to 10.

Two classification techniques, Classification Tree Analysis (CTA) and Maximum Likelihood Classification (MLC) were employed. Training data were collected with a field spectrometer capable of measuring reflectance values at 1nm interval from 350nm to 1050nm. Spectral samples from 14 coal tar parking lots and 19 non-coal tar parking lots were collected from various sites across Austin. The median pixel values in Hyperion and Landsat imagery for parking lots (with cars and other bright objects on them) were three times greater than the "pure parking lot" training sample median values. For the CTA and MLC classifications to work properly, the training sample values were adjusted to match the image pixel values by multiplying the sample values by three.

The accuracy of the classified images is assessed against reference data provided by the City of Austin. CTA classification of the Hyperion image provides the most accurate results with an overall accuracy of 84.34%. The overall accuracy of CTA classification of the Landsat image is 46.28%. MLC classification of the Hyperion image

has an overall accuracy of just 27.04%. MLC classification of the Landsat image yields only five pixels classified as "coal tar." CTA classification clearly outperforms MLC classification for both platforms.

Factors that contribute to classification errors include less-than-optimal spatial resolutions of Hyperion and Landsat imagery, parked cars, lack of data from some key spectral regions, and changes in parking lot characteristics due to wear and aging. However, we believe the CTA classification of the Hyperion image is accurate enough to guide us to significant expanses of coal tar sealed surfaces. As an initial screen, even the CTA classification of Landsat imagery can be useful since only 24.24% of non-coal tar pixels were incorrectly classified as coal tar. Thus, a cluster of coal tar pixels in a CTA classified image likely signifies a true concentration of coal tar surfaces. For the Austin area, the classified images show a substantially higher concentration of coal tar sealed parking lots west of IH-35 than in the east.

REFERENCES

Benz, U., P. Hofmann, G. Willhauck, I. Lingenfelder, and I. Heynen. 2004. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing*, no. 58: 239-258.

Bernstein, M. P., S. A. Sandford, and L. J. Allamandola. 2005. The Mid-Infrared Absorption Spectra of Neutral Polycyclic Aromatic Hydrocarbons in Conditions Relevant to Dense Interstellar Clouds. *The Astrophysical Journal Supplement Series* 161, no. 1: 53-64.

Bryer, P. J., J. N. Elliot, and E. J. Willingham. 2006. The effects of coal tar based pavement sealer on amphibian development and metamorphosis. *Ecotoxology* 15, no. 3: 241-247.

Christensen, P. R., J. L. Bandfield, V. E. Hamilton, D. A. Howard, M. D. Lane, J. L. Piatek, S. W. Ruff, and W. L. Stefanov. 2000. A thermal emission spectral library of rock-forming minerals. *Journal of Geophysical Research* 105, no. E4: 9735–9739.

Clark, R. N., G. A. Swayze, R. Wise, E. Livo, T. Hoefen, R. Kokaly, and S. J. Sutley. 2007. *USGS digital spectral library splib06a: U.S. Geological Survey, Digital Data*, no. 231.

Cloutis, E. A. 1989. Spectral reflectance properties of hydrocarbons: remote sensing implications. *Science* 245, no. 4914:165-168.

Dalton, J. B., D. J. Bove, C. S. Mladinich, and B. W. Rockwell. 2004. Identification of Spectrally similar materials using the USGS Tetracorder algorithm: the calcite-epidote-chlorite problem. *Remote Sensing of Environment* 89, no. 4: 455-466.

Eisler, R. 1987. Polycyclic aromatic hydrocarbon hazards to fish, wildlife, and invertebrates: a synoptic view. *U.S. Fish and Wildlife Service Biological Report* 85, no. 1.11.

Fabacher, D. L., J. M. Besser, C. J. Schmitt, J. C. Harshbarger, P. H. Peterman, and J. A. Lebo. 1991. Contaminated sediments from tributaries of the Great Lakes: chemical characterization and cancer-causing effects in medaka (Oryzias latipes). *Archives of Environmental Contamination and Toxicology*, no. 20: 17-35.

Herold, M., M. E. Gardner, and D. A. Roberts. 2003. Spectral resolution requirements for mapping urban areas. *IEEE Transactions on Geoscience and Remote Sensing* 41, no. 9: 1907-1919.

Herold, M., D. A. Roberts, M. E. Gardner, and P. E. Dennison. 2004. Spectrometry for urban area remote sensing –Development and analysis of a spectral library from 350 to 2400nm. *Remote Sensing of Environment* 91, no. 3: 304 -319.

Herold, M. and D. Roberts. 2005. Spectral characteristics of asphalt road aging and deterioration: implications for remote-sensing applications. *Applied Optics* 44, no. 20: 4327-4334.

Jensen, John R. 2005. *Introductory Digital Image Processing* 3rd Edition. Pearson Prentice Hall.

Jha, M. N., J. Levy, and Y. Gao. 2008. Advances in remote sensing for oil spill disaster management: state-of-the-art sensors technology for oil spill surveillance. *Sensors* 8, no. 1: 236-255.

Kershaw, J. R. 1996. Fluorescence spectroscopic analysis of benzo[a]pyrene in coal tar and related products. *Fuel* 75, no. 4: 522-524.

Mahler, B. J., P. C. Van Metre, T. J. Bashara, J. T. Wilson, and D. A. Johns. 2005. Parking lot sealcoat: an unrecognized source of polycyclic aromatic hydrocarbons. *Environmental Science and Technology* 39, no. 15: 5560-5566.

O'Conner, J. M. and R. J. Huggett. 1988. Aquatic pollution problems, North Atlantic coast, including Chesapeake Bay. *Aquatic Toxicology* 11, no. 2: 163-190.

Scoggins, M., N. L. McClintock, and L. Gosselink. 2007. Occurrence of polycyclic aromatic hydrocarbons below coal-tar-sealed parking lots and effects on stream benthic macroinvertebrate communities. *Journal of North American Benthological Society* 26, no. 4: 694-707.

Shafri, H., A. Suhaili, and S. Mansor. 2007. The performance of Maximum Likelihood, Spectral Angle Mapper, Neural Network and Decision Tree classifiers in hyperspectral Image Analysis. *Journal of Computer Science*, June issue.

Tso, B. and P. Mather. 2001. *Classification methods for remotely sensed data* 1st Edition. Taylor and Francis.

U.S. Environmental Protection Agency. 2008. *Ecological Toxicity Information*. Available online: http://www.epa.gov/R5Super/ecology/html/toxprofiles.htm#pahs.

Van Metre, P. C. and B. J. Mahler. 2003. The contribution of particles washed from rooftops to contaminant loading to Urban streams. *Chemosphere* 52, no. 10: 1727-1741.

Van Metre, P. C., B. J. Mahler, M. Scoggins, and P. A. Hamilton. 2006. Parking lot sealcoat: A major source of polycyclic aromatic hydrocarbons (PAHs) in urban and suburban environments. *U.S. Geological Survey, Fact Sheet,* no. 2005-3147.

Van Metre, P. C., B. J. Mahler and J. T. Wilson. 2009. PAHs underfoot: contaminated dust from coal-tar sealcoated pavement is widespread in the United States. *Environmental Science & Technology* 43, no. 1: 20-25.

Walsh, S. J., T. W. Crawford, W.F. Welsh, and K. A. Crews-Meyer. 2001. A multiscale analysis of LULC and NDVI variation in Nang Rong district, northeast Thailand. *Agriculture, Ecosystems and Environment* 85, no. 1: 47-64.

Ward, M. H., J. R. Nuckols, S. J. Weigel, S. K. Maxwell, K. P. Cantor, and R. S. Miller. 2000. Identifying populations potentially exposed to pesticides using remote sensing and Geographic Information System. *Environmental Health Perspectives* 108, no. 1: 5-12.

Winkelmann, K. H. 2005. On the applicability of imaging spectrometry for the detection and investigation of contaminated sites with particular consideration given to the detection of fuel hydrocarbon contaminants in soil. Ph. D. diss., Brandenburg University at Cottbus.

VITA

Mohan Rao was born in New Delhi, India. At age 10, he and his family moved to Lima, Peru. At age 18, he moved to Austin, Texas to attend the University of Texas at Austin and graduated with a B.A. in Computer Science in 1980. Subsequently, he worked as a software developer for 13 years and helped develop, among other things, a high-performance database management system tailored for use in supercomputers. He worked for another 12 years as an IT consultant, principally as a database performance specialist. In 2006, he joined the Masters of Applied Geography program at Texas State University-San Marcos. He currently works as a Geographer at the U.S. Geological Survey's Texas Water Science Center in Austin, Texas.

Permanent Address: mvr@grandecom.net

This thesis was typed by Mohan Rao.