Using the Semi-automated Plug-In Tool for Feature Identification, Recognition, and Extraction (SPITFIRE) to Extract Roads and Surface Material Types from AVIRIS Imagery

Brian S. Penn
Boeing-Autometric
Colorado Springs, Colorado
brian.s.penn@boeing.com

K. Eric Livo
United States Geological Survey
Denver, Colorado

Introduction

Effective tactical troop deployment can be negatively impacted by a lack of current and accurate lines of communication (LOC) information. For example, little, if any NIMA and ADRG data exist for locations such as East Timor, Bosnia, Somalia, or Afghanistan. Because of funding constraints and the sheer size of the problem, this situation is unlikely to change in the near term and will continue to unnecessarily complicate the already difficult task of troop deployment and logistical support for troop deployment. Additionally, while some engagements occur in rural areas, the trend is toward engagements in urban areas.

Over the past several years, a variety of military, government, and civilian organizations have developed tools for automated feature extraction to significantly speed up the task of identifying and extracting LOC information from imagery for use in mapping applications (Maillard and Cavayas, 1989; Zlotnick and Carnine, 1993; Gee, 1994; Kahrig et. al., 1994; Lersch, Iverson, and West, 1994; Bullock et al., 1996; Couloigner and Ranchin, 2000). While some of these efforts have had limited success, constraints in computing power and the availability of data containing the requisite spatial as well as spectral characteristics have significantly hampered achievement of a complete solution.

The primary goal of this system described here is to quickly extract LOC for use by military forces. They don't need the quintessential map, they need roads, and they need them fast. For this reason the threshold for success is based on the Pareto Principle or the 80/20 rule. The Pareto Principle states that the first 80% of whatever is being done is achieved with the first 20% of the total effort. The converse is also true: 80% of the effort is expended completing the final 20% of a task. This is the metric used to determine the success for LOC extraction.

Spatial resolutions of satellite imagery data have improved to the point that it is now possible to automatically map LOCs for most environments. These data are readily available in differing spatial and spectral resolutions and can be merged to provide the range of data products necessary to produce LOC maps. Moreover, computer speed continues to increase dramatically, permitting even computationally intensive tasks like satellite image processing to be performed on a desktop computer. Consequently, the primary technological barriers to developing fast, efficient, and effective LOC recognition systems no longer exist.

Hyperspectral imagery (HSI) offers significant advantages over panchromatic and multispectral imagery. Hyperspectral imagery is not a new type of imagery; it really is just more of the same data at finer spectral resolution. Previously interpreters were limited to broad bands of the electromagnetic spectrum (EMS). These bands "smeared" subtle features in the integrated signatures representing information about an imaged material. These details are only apparent when the wide bands of the EMS are split into 100s and 1000s of samples. This type of resolution requires a hyperspectral sensor. HSI data provides unprecedented capability for extracting previously hidden information from imagery.

Hyperspectral imagery complements high spatial resolution imagery by providing high spectral resolution data. Instead of having several very wide bands ranging in the tens of micrometers (μm), hyperspectral sensors have many bands with sampling measured in the tens of nanometers (nm). Sensors like the Jet Propulsion Laboratory's Airborne Visible & Infrared Imaging Spectrometer (AVIRIS) are regularly providing high-fidelity hyperspectral images composed of 224 samples at 10 nanometer (nm) intervals over the EMS from 400 to 2500 nm wavelengths.

The Semi-automated Plug-In Tool for Feature Identification, Recognition, and Extraction (SPITFIRE) extracts lines of communication (LOC), i.e., roads and rivers, from digital imagery of urban environments and generates centerline vectors in ArcView shapefile format. SPITFIRE is capable of quickly extracting LOCs from digital imagery ranging from panchromatic to hyperspectral imagery (HSI). When HSI data is provided, SPITFIRE can also provide surface material types. This paper explores the techniques used to extract LOCs and determine material types using AVIRIS HSI data.

BACKGROUND

During fiscal year 1999, Army Space Command in Colorado Springs, Colorado drafted a proposal to the Military Exploitation of Reconnaissance and Intelligence Technology (MERIT) Board to develop a prototype tool for LOC extraction. The successful completion of the multispectral phase of SPITFIRE led to the development of a hyperspectral version implemented as a plug-in to Research System's Environment for Visualizing Images (ENVI) during fiscal year 2001. The basic ideas for extracting LOCs were carried over from Phase I. The significant difference in Phase II involved using HSI data to identify road surface material types.

To take full advantage of the information in hyperspectral data, it was necessary to develop a hybrid knowledge-based system to identify surface material types. The tremendous range in urban surface material types led to the development of general categories of surface material types used for LOC extraction in urban environments, i.e., concrete, asphalt, bare soil, and water. These general categories provided the basis for material identification using a combination of rules, neural networks, and fuzzy logic. The hyperspectral component of SPITFIRE was developed and tested using an AVIRIS scene acquired in July of 1997 west of Denver over the Front Range of Colorado. The AVIRIS scene was atmospherically corrected by the U.S.G.S. at the Spectroscopy Laboratory in Lakewood, Colorado.

EXTRACTING LINES OF COMMUNICATION

Lines of communication, i.e., roads, rivers, etc. have certain spatial and spectral characteristics. The chief requisite among these characteristics needed is spectral contrast between the road surface material and the off-road material. Finding the LOCs requires proper image preparation. Image preparation and extraction of the LOC is a three-step process:

1) Finding the bands with the highest amount of contrast between the road surface and the non-road surface;
2) Reduce the amount of information to only include the roads of interest;
3) Connect adjoining pixels of interest along the LOC.

Wavelengths with the maximal contrast can be found either empirically, using a statistical function like the Bhattacharyya distance (Chulhee and Landgrebe, 1993), or by performing a change of basis on the data set, i.e., a principle component transformation. Visual inspection of the average spectra of two areas of interest is a simple way to find those wavelengths providing the best contrast (Figure 1). Only the visual inspection technique and principle component transformations are covered here.

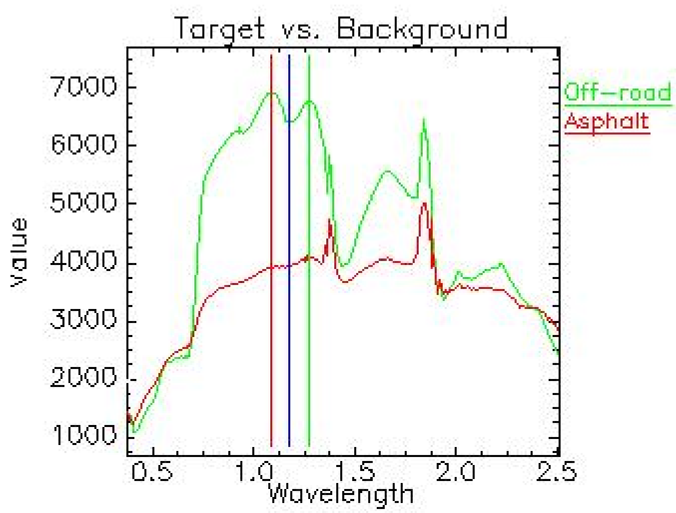


Figure 1 – Average spectra showing largest contrast between road surface and non-road surface

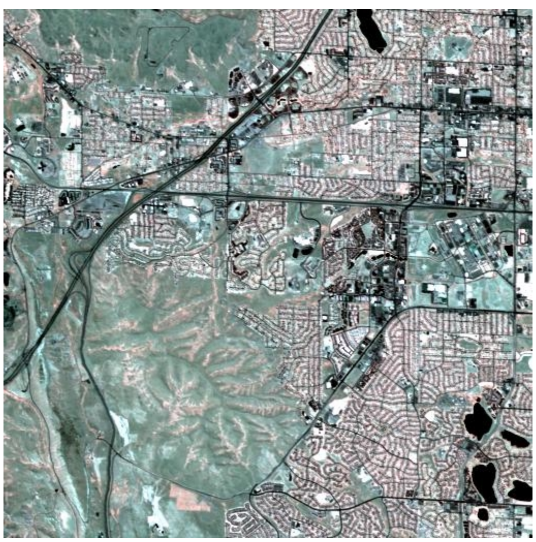
PRINCIPLE COMPONENT TRANSFORMATION

Finding a maximal contrast between road surface materials and non-road materials can be achieved by performing a principle component (PC) transformation. PC transformations are useful because of the high positive correlation normally found between bands in imagery data. The number of PC bands generated is equal to the total number of bands, *n*, in the image. A PC transformation rotates the coordinate axes around such that the first PC is parallel to the maximum amount of variance in the imagery data set. The second PC is orthogonal to the first PC and is parallel to that portion of the data containing the second largest amount of variance. The is process is continued until the n^{th} axis is perpendicular to the $(n-1)^{th}$ and contains the next least amount of variance. The last couple of PCs generally contain mostly noise.

It was empirically determined that subtracting the 4^{th} PC from 2^{nd} PC generated a good contrast between the road surface and non-road surface materials. The complexity and time required to perform a PC transformation compared to simply using a visual test to find the bands with the highest contrast makes doing a PC transformation untenable.

LOC EXTRACTION

Once the bands with the highest contrast were found, it was obvious that even with the contrast there was too much high-frequency information in the imagery. This problem of high-frequency information was addressed by applying a low-pass filter to the image (Figure 2).



a

b

Figure 2 – a) False-color image of bands shown in Figure 1;
b) Figure 2a smoothed with 3x3 low-pass filter.

Once the image is prepared, the LOC extraction process is performed with a simple flood-fill. A flood-fill should be familiar to anyone using a computer paint program. A flood-fill works by selecting the starting point of the feature to fill with the fill tool. Because the fill procedure is done using simple statistics it is quite rapid (Figure 3). The primary benefit of this approach is that it uses both the spectral and spatial nature of LOCs to constrain the bounds of the LOC.

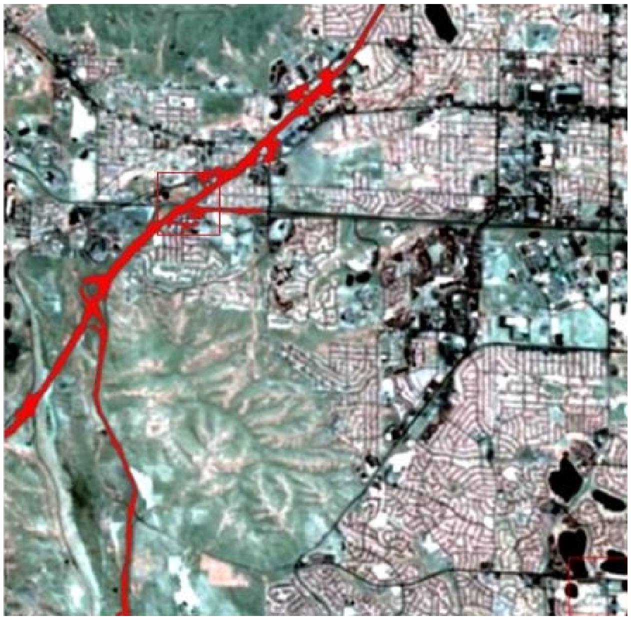


Figure 3 – LOCs found using flood-fill method

LIMITATIONS OF LOC EXTRACTION

Spatial resolution restricts the size of LOCs that can be effectively extracted. A single lane road is approximately 3-4 m wide, a two-lane road is 6 – 8 m, and a four-lane road with a turning lane is approximately 25 m wide. While it is possible to extract subpixel features, this can lead to inaccuracies. The AVIRIS sensor when flown at 20 km has an instantaneous field of view (IFOV) of 1 milliradian yielding a nominal pixel size of ≈ 20 m. At this pixel size four-lane roads are slightly larger than individual pixels making them easy to unambiguously identify and extract. Using the success metric of the Pareto Principle described above, most four-lane and some two-lanes can be extracted quickly. Extracting smaller LOCs can be easily achieved by flying the sensor lower with the same IFOV yielding pixels sufficiently small to extract one- and two-lane roads.

Once the flood-fill finds the appropriate LOC, the next task is to generate centerline vectors. Using Able Software’s R2V, commercial off-the-shelf raster to vector conversions tool, the LOCs are converted to vectors and exported as shapefiles for use in GIS systems.

Material Identification

The primary reason for using hyperspectral data is to extract material types. Urban areas are spectrally complex (Gardner et al., 2001; Small, 2001). For most researchers in HSI spectral complexity equates to collecting vast amounts of spectra for every possible substance that might be encountered and storing these spectra in a reference library. This is equivalent to carrying around an album containing a picture of every person you have ever met and comparing everyone you meet to this photo album! Humans don't do this; they have the capability of generalizing about faces. An example of our ability to generalize is demonstrated by the number of people who see a face on the surface of Mars. There are techniques pioneered in artificial intelligence that permit satisfactory generalization such as neural networks and fuzzy logic.

Representation

In artificial intelligence (AI), it is generally understood that problem complexity is usually a function of representation. All intelligent human tasks involve some form of problem solving which is fundamentally a search problem. Recasting problems into different representations is just another way of constraining the search space. Seemingly intractable problems in one representation when recast into a different domain are easily solved (Raphael, 1976; Rich, 1983; Winston, 1984). In the current context, using a flood-fill to extract LOCs is an example of how recasting a problem into another domain makes it easily solvable. Another particularly powerful example is Fourier transforms. Fourier transforms convert signals from the time domain into the frequency domain. In the frequency domain all signals can be represented as a series of sine waves that are much easier to work with than other waveforms.

It should be clear that how the problem is perceived or represented determines the ease by which it is solved. That is the controlling mechanism behind the solution presented below for road surface material identification.

The alternative to collecting individual spectrum of all surfaces to identify materials is to do what comes naturally to humans, i.e., generalize. As was alluded to above regarding the recognition of faces, one of the skills our brains naturally possess is the ability to generalize. The first task in developing a system capable of generalizing about spectral properties of road surface material types is to collect representative spectra of the surficial material types of interest. Once the spectra were collected, a small library of average spectrum for each material type was developed (Figure 4).

Reflectance rules for surface material types

Roadway surfaces are generally composed of aggregate (bare soil), asphalt, and/or concrete. Other material types include vegetation and water. Several heuristics (rules of thumb, i.e., generalizations) were developed. One rule was that bare soil and concrete have significantly higher albedos than water and asphalt. Using this rule, bright reflectors are easily separated from dark reflectors. Average albedos were useful enough for this first pass. Any spectrum with an average albedo below 10% was considered

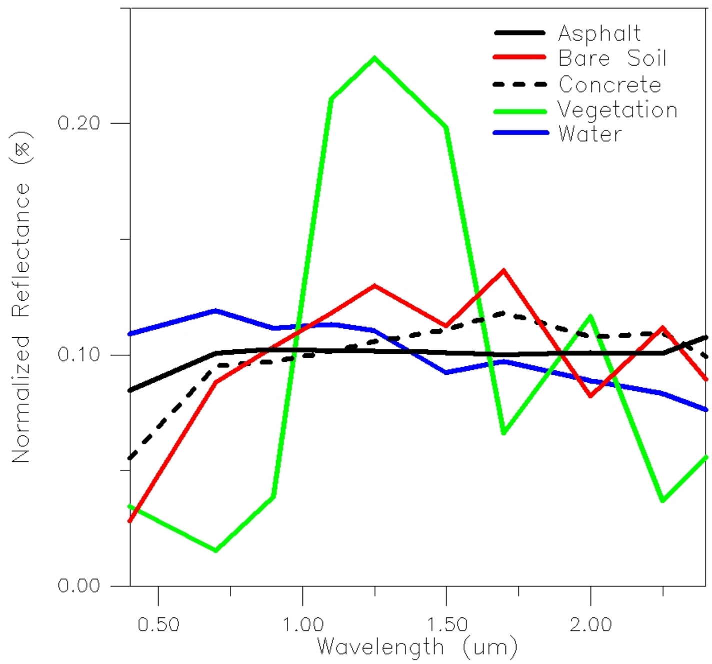


Figure 4 – Normalized averages for spectra collected using and ASD-FR hand held spectrometer of primary surface material types resample to ten wavelengths.

either water or fresh asphalt. Water and fresh asphalt are distinguished by the fact that water may have a slight spike in the blue wavelengths and the overall spectrum has a negative slope. Fresh asphalt is practically flat over the range of wavelengths from 400-2500 nanometers (nm). Spectra with reflectance values greater than 25% were considered either pure concrete or pure soil. Since concrete and bare soil don't have any unique absorption features it was necessary to use the overall spectral shape to distinguish between them. It was determined empirically that all 224 channels of AVIRIS data were more data than necessary to distinguish between these material types. The necessary spectral information was reduced to 10 bands. These 10 bands equally spaced over 400-2500 nm were enough to discriminate between bare soil and concrete. The problem then reduced to one of pattern matching. One of the best tools for pattern matching and recognition is a neural network (Lippman, 1987, 1989; Fukushima, 1988; King, 1989).

NEURAL NETWORK

A feed-forward neural network was designed with ten input nodes and two output nodes (Rummelhart, 1986). The network was trained using back propagation. Several hundred training sets of spectra of concrete and bare soil were presented to the neural network during the training process and convergence generally occurred within one thousand iterations.

FUZZY RULE

Spectra of road surface material types with average reflectance values between 10% and 25% were determined to be mixtures of either asphalt and concrete or asphalt and bare soil. It was possible to distinguish whether the mixture was either concrete or bare soil by the shape of the curve, but not how much was present. To address this question, a fuzzy rule was developed and applied to the spectrum. Fuzzy rules are based on fuzzy set theory (Zadeh, 1965). The fuzzy rule is show in Figure 5.

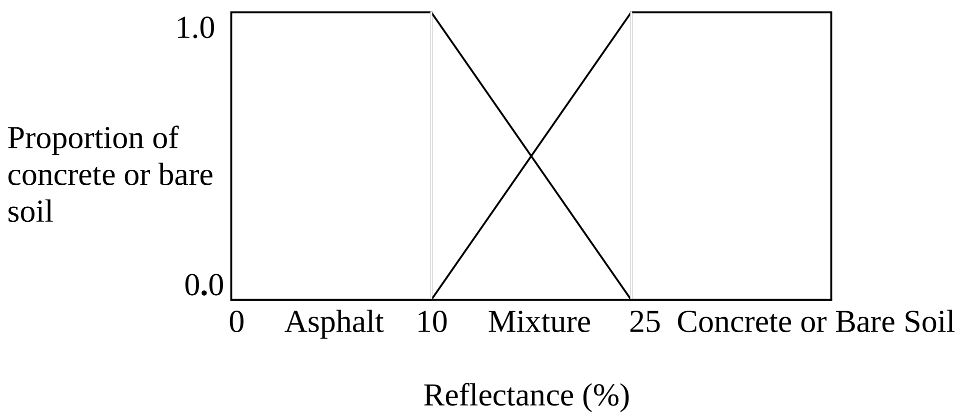


Figure 5 – Fuzzy rule for determining proportions
of asphalt and bare soil or concrete

The rule in Figure 5 is over-simplified. The diagonal lines representing relative proportions should really be exponential curves concave downward with steep slopes nearly parallel to the y-axis and rolling over near 0.85 and becoming asymptotic with the upper x-axis. This shape is more consistent with what is understood regarding non-linear mixtures between dark and light materials and the effect dark materials have on the reflectance values of lighter materials. The exact shape of these curves is the subject of current research.

The general reflectance rules, neural network, and fuzzy rule were combined into a hybrid knowledge-based system and integrated into the SPITFIRE plug-in module to ENVI.

CONCLUSIONS

The Semi-automated Plug-In Tool for Feature Identification, Recognition, and Extraction (SPITFIRE) was designed and implemented to extract lines of communication from digital imagery. The initial version operated on multispectral imagery. SPITFIRE was later migrated to ENVI to operate on hyperspectral imagery.

Extracting LOCs involves locating wavelengths with sufficient spectral contrast between the desired road surface and non-road surface (background). Once these wavelengths are found, then a simple flood-fill algorithm is used to follow the road. A raster to vector tool is then used to find centerlines and generate vectors as shapefiles for ease of incorporation in to a Geographic Information Systems like Arc/Info or Arc/View.

The chief benefit of using hyperspectral imagery is the ability to discriminate and identify material types. SPITFIRE achieves road surface material identification using a hybrid knowledge-based system using a combination of rules, neural networks, and fuzzy logic.

The basis for this approach of classifying road surface materials is generalization and pattern recognition. This approach obviates the need to develop large spectral libraries because it mimics what humans do in terms of pattern recognition. A smaller set of characteristic spectra are still needed to develop generalizations about the different surface material types.

REFERENCES

Bullock, M.E., Fairchild, S.R., Patterson, T.J., and Haxton, R. 1996. Automated map generation and update from high-resolution multispectral imagery, SPIE vol. 2758, 91-103.

Chulhee L. and Landgrebe, D. A. 1993. "Decision Boundary Feature Extraction for Non-Parametric Classification," IEEE Transactions on System, Man, and Cybernetics, Vol. 23(2), 433-444.

Couloigner, I. and Ranchin, T. 2000. Mapping of Urban Areas: A Multiresolution Modeling Approach fro Semi-Automatic Extraction of Streets, Photogrammetric Engineering & Remote Sensing, vol. 66 (7), 867-874.

Fukushima, K. 1988. Neural Networks for Visual Pattern Recognition, IEEE Computer, vol. 21(3), 65-75.

Gardner, M. E., Roberts, D. A., Funk, C., and Noronha V. 2001. Road Extraction from AVIRIS using Spectral Mixture and Q-Tree Filter Techniques in Proceedings of the 10th JPL Airborne Earth Science Workshop, Green, R. O. (ed.), JPL Publication 02-1, 145-150.

Gee, S.J. 1994. Lines of Communication in LANDSAT TM Imagery, SPIE vol. 2231, 27-35.

Kahrig, A.H., Abeloe, K.A., Teschan, P.E. 1994. Model-based approach to Landsat-TM scene linear LOC segment detection using morphology, SPIE vol. 2231, 7-18.

King, T. 1989. Using Neural Networks for Pattern Recognition, Dr. Dobb's Journal, January, 14-28.

Lersch, J.R., Iverson, A.E., and West, K.F. 1994. Multispectral Lines of Communication, SPIE vol. 2231, 36- 47.

Lippman, R. P. 1987. An Introduction to Computing with Neural Networks, IEEE Transactions in Acoustics, Speech, and Signal Processing Magazine, April, 4-22.

----------------- 1989. Pattern Classification Using Neural Networks, IEEE Communications Magazine, November, 47-64.

Maillard, P. and Cavayas, F. 1989. Automatic map-guided extraction of roads from SPOT imagery for cartographic database updating, International Journal of Remote Sensing, vol. 10(11), 1775-1787.

Raphael, B. 1976. The Thinking Computer: Mind inside Matter, W.H. Freeman and Company, San Francisco, CA, 322 p.

Rich, E. 1983. Artificial Intelligence. McGraw-Hill, New York, New York, 436 p.

Rummelhart, D.E., Hinton, G.E., and Williams, R.J. 1986. Learning Internal Representations by Error Propagation, *in* Parallel Distributed Processing: Explorations in the microstructure of cognition-Volume 1: Foundations, Feldman, J.A., Hayes, P.J., and Rummelhart, D.E. (eds.). MIT Press, Cambridge, MA, 318-362.

Small, C. 2001. Spectral Dimensionality and Scale of Urban Radiance in Proceedings of the 10th JPL Airborne Earth Science Workshop, Green, R. O. (ed.), JPL Publication 02-1, 375-385.

Winston, P.H. 1984. Artificial Intelligence, 2nd edition, Addison-Wesley, New York, New York, 524 p.

Zadeh, L.A. 1965. Fuzzy Sets, Information and Control, vol. 8, 338-353.

Zlotnick, A., and Carnine, P.D. 1993. Finding Road Seeds in Aerial Images, CVGIP: Image Understanding, vol. 57(2), 243-260.