INTELLIGENT TARGET DETECTION IN HYPERSPECTRAL IMAGERY

Ayanna Howard, Curtis Padgett, Kenneth Brown
Jet Propulsion Laboratory, California Institute of Technology
4800 Oak Grove Drive, Pasadena, CA 91109-8099

ABSTRACT

Many applications that use hyperspectral imagery focus on detection and recognition of targets that occupy a portion of a hyperspectral pixel. We address the problem of sub-pixel target detection by evaluating individual pixels belonging to a hyperspectral image scene. We begin by clustering each pixel into one of **n** classes based on the minimum distance to a set of **n** cluster prototypes. These cluster prototypes have previously been identified using a modified clustering algorithm based on prior sensed data. Associated with each cluster is a set of linear filters specifically designed to separate signatures derived from a target embedded in a background pixel from other typical signatures belonging to that cluster. The filters are found using directed principal component analysis which maximally separates the two groups. Each pixel is projected on this set of filters and the result is fed into a trained neural network for classification.

A detailed description of our algorithm will be given in this paper. We outline our methodology for generating training and testing data, describe our modified clustering algorithm, explain how the linear filters are designed, and provide details on the neural network classifier. Evaluation of the overall algorithm demonstrates that for pixels with embedded targets taking up no more than 10% of the area, our detection rates approach 99.9% with a false positive rate of less than 10^{-4}.

1.0 INTRODUCTION

The use of multi-spectral imaging systems in such areas as geology, meteorology, and defense has sparked an interest for the development of innovative automatic detection and classification schemes. Various techniques have been developed in order to effectively address this issue. Kim et al [Kim,91] used a decision tree classifier in order to sort various hyperspectral signatures into numerous classes using as few features as possible. Jia et al [Jia,94] grouped hyperspectral data into groups based on maximum likelihood classification. Yu et al [Yu,97] detected known low-contrast objects in a nonstationary background using a generalized maximum-likelihood ratio principal. Sabol et al [Sabol,92] quantified the conditions necessary for sub-pixel spectral detection. Harsanyi et al [Harsanyi,94] used orthogonal subspace projection to reduce the dimensionality of the hyperspectral imagery.

Some of the main limitations with these techniques are the processing time requirements for sub-pixel detection and the compromise between a low false alarm vs. detection rate. Due to the high dimensionality of hyperspectral imagery, an optimal number of spectral bands must be extracted which have the greatest ability to discriminate embedded targets from typical background pixels. The larger the number of bands utilized, the longer the processing time required, and the less "real-time" the process becomes. In addition, a low false positive rate coupled with a high detection rate is essential, especially in such areas as military target detection were wide area surveillance (many thousand of spectra) is needed. In these tasks, a very low false positive rate is desired to increase the confidence that the targets identified are real.

Our research objective is to develop an algorithm for autonomous detection of sub-pixel target objects embedded in hyperspectral data. The evaluation of these algorithms is based on inserting actual target signatures into real scenes of hyperspectral images. The scenes used are generated by the Airborne Visible InfraRed Imaging Spectrometer (AVIRIS), an optical sensor which generates 224 bands of spectral information for each pixel, with each pixel covering a 20 m^2 surface area.

The target data inserted into the background was extracted from the Forest Radiance I experiment provided by SITAC. This database contains reference signatures for a number of military targets, which we shall call our target set. The scenario we are examining involves examining ground images from AVIRIS to determine if the spectral signature from any of the 20 m^2 pixels contains, as a constituent element, a member of the given target set. In real time, we will reduce the data dimensionality of a scene using an optimal set of linear filters and spatially locate targets in the scene with a neural network classifier. Figure 1 provides an overview of our approach for detecting a known set of targets in a hyperspectral image. The rest of this paper describes the methodology used to investigate sub-pixel target detection in detail.

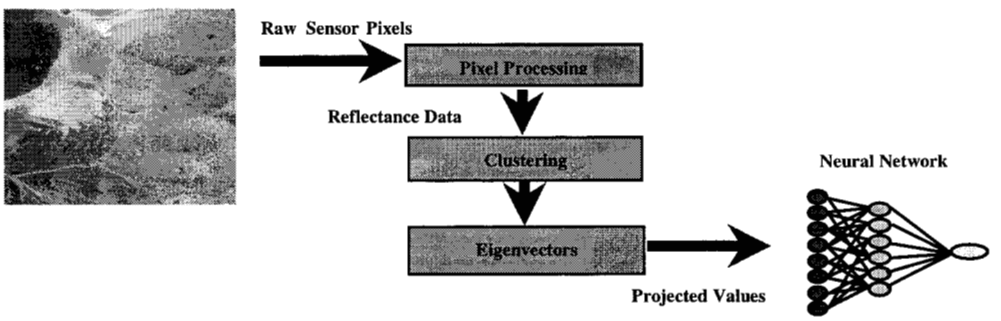

Figure 1: The data processing path for each pixel from the sensed image.

2.0 TECHNIQUE

2.1 AVIRIS DATA AND TARGET DESCRIPTION

The Airborne Visible InfraRed Imaging Spectrometer (AVIRIS) is an optical sensor that delivers calibrated images of the upwelling radiance in 224 spectral channels, or bands, with wavelengths from 400-2500nm. AVIRIS collects spectra sequentially by using a whiskbroom scan mechanism. The radiance from an approximately 20 m^2 patch on the ground is dispersed thru four grating spectrometers to obtain a spectrum consisting of 224 channels [Vane,93]. The AVIRIS data is inverted to units of spectral reflectance using a radiative transfer model estimate of atmospheric path radiance and reflected radiance [Vane,93]. We thus classify a background pixel **b** as a vector of 224 bands, representing the spectral reflectance data from a 20 m^2 ground patch area. Due to the atmospheric correction routines, some of the 224 bands of spectral reflectance data do not possess any interesting (or valid) information. We thus default these bands to zero before they are processed by our algorithm (Figure 3).

The hyperspectral image scenes used in this research effort are taken from the Cuprite (Fig. 1), and ARM (Fig. 6) sites. Target spectra were obtained from ground truth measurements conducted at the FR1 experiment. These measurements give us target spectra data which must be transformed into reflectance data for import into our algorithm. This is accomplished by comparing target spectra to a known reflectance standard. The reflectance of the target is determined by computing the ratio of the signals between target **T** and the known standard **S** multiplied by the reflectance of the standard $\mathbf{R_s}$ (Fig. 2).

$$R_f = T/S * R_s$$

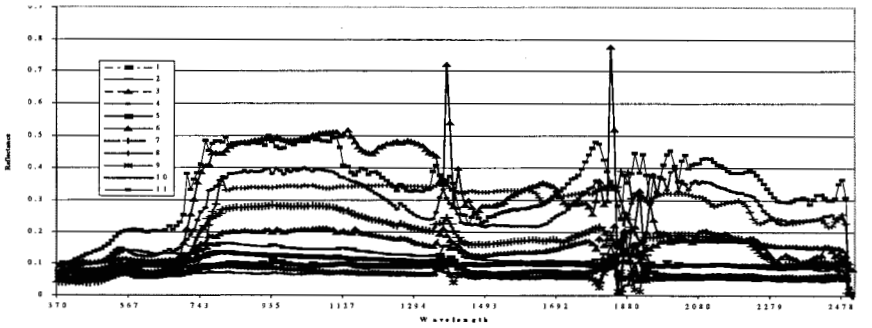

Figure 2. Average target signatures obtained from the FR1 experiment.

Passing the target spectra through this equation gives us the target reflectance data, a 224 vector band of information, which we refer to as the target signature.

In order to mix targets into an AVIRIS scene, we form the composite signature:

$$x*t + (1-x)*b+n \quad (1)$$

where **x** is the fractional mixing level, **t** is the target signature, **b** is a pixel belonging to a typical Aviris scene, and **n** is randomly generated zero mean Gaussian noise.

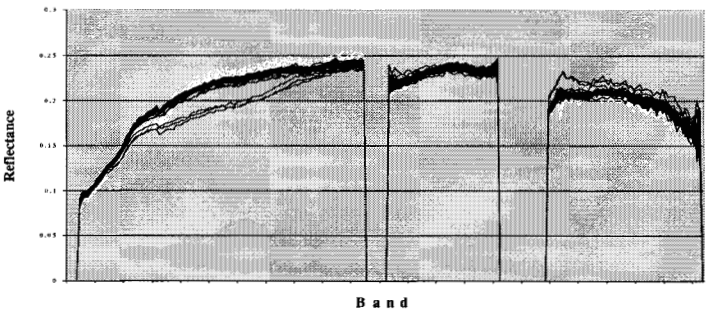

Figure 3. Randomly sampled background pixels (black) and targets mixed at 10% (white) drawn from the same cluster.

Figure 3 gives a comparative example of randomly drawn background pixels belonging to a single cluster versus an embedded 10% target signature mixed with background pixels drawn from the same distribution. As shown, separating the mixed target examples from the background elements is an extremely difficult task. The next section describes our approach for detecting the target elements as shown in Figure 3.

2.2 ALGORITHM DESCRIPTION

Given a set of targets **T**, the goal of the algorithm is to detect, in real time, any target $\mathbf{t} \in \mathbf{T}$ present in any pixel belonging to a hyperspectral scene. We begin by clustering the pixels belonging to a hyperspectral scene into one of **n** classes based on its distance to a set of **n** cluster prototypes. These cluster prototypes have previously been identified using a modified clustering algorithm based on prior sensed data. Associated with each cluster is a set of

linear filters specifically designed to separate signatures dervied from a target embedded in a background pixel from other typical signatures belonging to that cluster. Each pixel is projected on this set of filters and the result is fed into a trained neural network for classification.

In order to accomplish this task, knowledge must be extracted through the following preprocessing steps:

1) Identify a set of background cluster prototypes used to segement the hyperspectral scene.
2) Derive a set of linear filters for each background type used to optimally separate targets embedded in background pixels from other typical background pixels.
3) Train a set of *expert* neural network classifiers that receive as input the corrected pixels projected on their respective linear filter set and respond with 1 when targets are embedded in the pixels and –1 otherwise.

2.2.1 Clustering

To effectively simplify the distribution of data classified by an expert neural network, we partition the incoming pixel signatures into a number of predetermined groups by using the prototypes $\mathbf{P_i}$ of a clustering algorithm. The clustering algorithm is run on previously aquired data that reflects the distribution of the scene being analyzed.

The clustering algorithm employed is a modified version of a standard clustering technique outlined in Duda and Hart [Duda,73]. The standard algorithm uses a standard least squares criterion to minimize the distance between each of **n** randomly selected groups. These groups are initially constructed by selecting typical background pixels and randomly distributing them among each group. The criterion minimized by the standard clustering algorithm is:

$$\text{cost} = \sum_i \sum_j \| p_j - \mathbf{P_i} \| \tag{2}$$

where i is one of **n** clusters and p_j is a pixel in that cluster. The clustering algorithm iterates through each pixel and determines if moving the pixel to another group reduces the overall cost. If it does, the pixel is moved to the other group and the associated averages of each prototype cluster are recalculated. This continues until the moving of pixels no longer reduces the overall cost. The resultant cluster prototypes are then employed by our algorithm to segment the scene.

The clustering algorithm, as described, is independent of the detection problem. It does not take into account any information that we might have concerning the target set. The distribution size of clusters that contain backgrounds which closely resemble targets could be quite wide, thus making it more difficult to detect target pixels in them. As targets are typically camouflaged to resemble the background, this is a potentially serious problem. Similarly, it makes little sense to have narrow distributions for clusters whose background elements highly contrast with the targets, thus making them easier to detect. What is needed are clusters with narrower distributions for backgrounds that resemble targets in our set and wider distributions for those backgrounds that contrast with them. To accomplish this, we modify the criterion given in (2) to reflect target knowledge. The change in (2) consists of simply weighting each pixel by a term reflecting its closeness with elements in the target set. The modified criterion is given by:

$$\text{cost} = \sum_i w_j \sum_j \| p_j - \mathbf{P_i} \| \tag{3}$$

where

$$w_j = 1/\sum_s \| p_j - t_s \| \tag{4}$$

and t_s is an element of **T**, the target set. Pixels that are close to targets will be weighted more in the cost of the clustering algorithm than those further away, allowing the clustering algorithm to naturally provide more resource groups to those background types.

2.2.2 Filter Sets

The filtering step is an orthogonal sub-space projection of each pixel. It is used to optimally linearly separate the embedded target background pixels of each group from those pixels without targets. This is a standard technique used to reduce the dimensionality of the pixel (from 224 to 32 dimensions) while preserving as much of the signal as possible. The filters associated with a given prototype are derived from the distribution of its background pixels (***noise***) and the distribution of potential targets embedded in that background (***signal***). These two groups can be optimally separated to maximize the *signal to noise* ratio between them using directed principal components analysis (DPCA).

To characterize the distribution for cluster i, the covariance matrix, $\mathbf{R}_i$, is found for background pixels belonging to the cluster which do not contain targets and the covariance matrix, $\mathbf{S}_i$, is derived for those pixels mixed with target signatures. We are interested in finding a set of orthogonal basis vectors $\mathbf{W}_i$, that maximizes the expected signal to noise ratio of these two distributions. The generalized eigenvector solution:

$$S_i W_i = \lambda R_i W_i \tag{5}$$

accomplishes this. The set of filters defined by $\mathbf{W}_i$ is the directed components used in our algorithm to steer the eigenvector solution away from dimensions of high noise variance in a linearly optimal fashion.

2.2.3 Classification

The next step in our algorithm involves classifying each pixel's projection on $\mathbf{W}_i$ with a neural network. The networks are trained with data drawn from the two distributions used to determine $\mathbf{W}_i$, $\mathbf{R}_i$ and $\mathbf{S}_i$. The expert network for class i is required to respond with 1 for elements drawn from $\mathbf{S}_i$ and –1 from those drawn from $\mathbf{R}_i$. We use a simple feed forward model employing 10 sigmoidal hidden units trained with back propogation to get the desired result. The output can then be thresholded to achieve the desired detection and false positive rate by examining the receiver operator curves.

3.0 IMPLEMENTATION

We evaluated the overall performance of the algorithm using the described target set and two AVIRIS scenes (Cuprite copper mine in New Mexico and Midwestern farmland). Based on the target-background composite equation (1), we are interested in examining the algorithm's performance as the size (**x**) of the target relative to the background pixel varies, and as target measurement and sensor noise (**n**) changes.

3.1 METHODOLOGY

The two scenes consisted of over ¾ of a million pixels of which less than 10% were used in developing a set of training data. Testing data consisted of randomly drawn pixels from the remaining scene. Target pixels were generated by randomly selecting signatures from the target set and mixing them with arbitrary background pixels. The training data was then evaluated with the clustering technique to realize the prototypes (**P**). Figure 5 shows the prototypes generated by a sample clustering and Figure 6 shows the ARM scene segmented with those prototypes. A sub-sample of the training data (1000 examples each) was used to generate the covariance matrixes $\mathbf{R}_i$ and $\mathbf{S}_i$. The generalized eigenvector solutions $\mathbf{W}_i$, to these matrixes were then solved using a Matlab routine.

Training data for the neural network was again drawn from the set of training pixels. In addition, a portion of the training data for the network was used to halt training (a hold out set) as described in Haykin [Haykin,94]. Training of the networks used 5000 examples, ½ target and ½ background pixels. The hold out set consisted of 2500

examples not trained upon. It is used to stop training in order to prevent over learning on the training data which tends to decrease generalization.

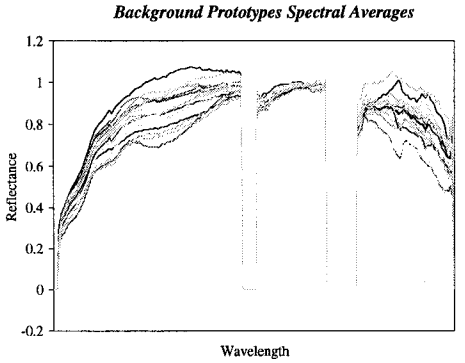

Figure 5. Example prototypes obtained from the clustering algorithm.

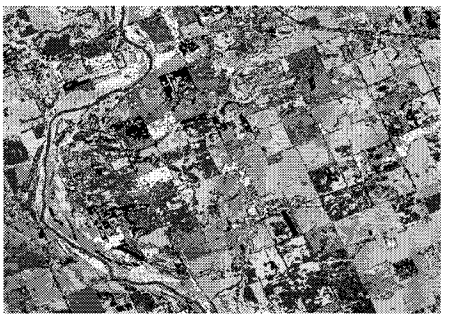

Figure 6. The ARM scene segmented with the prototypes found in Figure 5.

3.1.1 Detection As Target Size Varies

One of the major goals in this project was to determine how large the target had to be in relation to the pixel. For search or characterization scenarios, larger pixels for a given size target represents more ground coverage and allows for a wider search to be conducted in the same amount of time and for the sensor to be on a higher (and presumably safer) flying platform.

In the first experiment, we examined four mixing ratios of target and background. The algorithm was evaluated with 16 clusters derived using the unmodified clustering algorithm. No noise was added to the targets or the background pixels. Figure 7 plots the receiver operator curves (detection vs. false positives) for mix percentages of 2%, 5%, and 10%. 25% target mix resulted in 100% detection (not shown) with no false positives. The plots shown consist of averaging the results from 8 of the 16 clusters (over 50,000 pixels sampled).

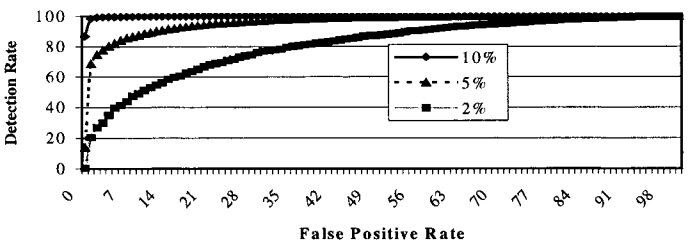

Figure 7. Receiver operator curves for neural network output as mix (x) of target in background varies as a percentage of pixel size.

3.1.2 Adding Noise to the Targets and Sensor

The filter sets are specifically designed to minimize the impact of noise in the detection process. However, significant noise is bound to impact the detection rates. The next two figures examine the impact of noise on the target set and the sensor. Figure 8 (Left) gives the receiver operator curves for the original cluster with SNR of 50 and 10. In this case, Gaussian noise was added to the background pixel. This simulates poor calibration, noise in calculating reflectance, or simple sensor noise. The impact on the algorithm is substantial at the 5% mix percentage shown.

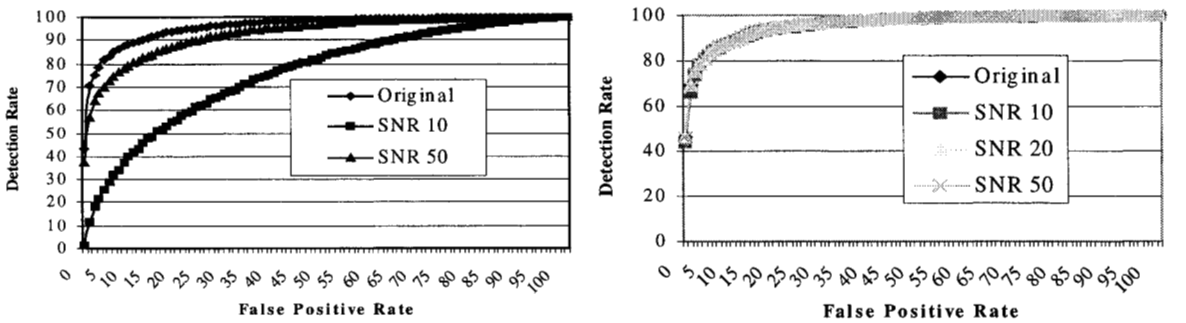

Figure 8. ROC for neural network output as Gaussian noise is added to background pixel signatures (Left) and target signatures (Right).

Figure 8 (Right) provides the impact of noise degradation on the target set. This represents less than perfect knowledge about the target signatures or its properties. Again, Gaussian noise is added to targets at 10, 20, and 50 SNR. The impact of noise on the targets is quite modest as compared to the quality of the sensed image. This suggests that the investment should be made in sensor calibration, noise reduction, and correction procedures rather than in collecting more accurate target spectra.

4.0 CONCLUSION

A novel detection algorithm and our evaluation methodology are described here. The detection algorithm was shown to perform detection at a rate of over 99% with false positives less than 10^{-4} on a set of targets mixed at 10% with background pixels. For larger targets, the detection rates approach 100% (at 25% mix, the algorithm was perfect).

5.0 ACKNOWLEDGEMENTS

The research described in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

6.0 REFERENCES

Duda, H. and Hart, M. *Pattern Classification and Scene Analysis.* John Wiley & Sons. New York. 1973.

Green, R.O., Conel, J.E., Roberts, D.A. "Estimation of Aerosol Optical Depth, Pressure Elevation, Water Vapor and Calculation of Apparent Surface Reflectance from Radiance Measured by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) Using A Radiative Transfer Code" *SPIE Imaging Spectrometry of the Terrestrial Environment,* vol. 1937, pp 2-11, 1993.

Harsanyi, J. and Chang, C. "Hyperspectral Image Classification and Dimensionality Reduction: An Orthogonal Subspace Projection Approach", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 32, pp. 779-785, July 1994.

Haykin, S. *Neural Networks: A Comprehensive Foundation.* Macmillan College Publishing Company. New York. 1994.

Jia, X., and Richards, J.A., "Efficient Maximum Likelihood Classification for Imaging Spectrometer Data Sets," *IEEE Trans. on Geo. and Rem. Sens.* 32 (2) 274 – 281 (1994).

Kim, B. and Landgrebe, D. "Hierarchical Classifier Design in High-Dimensional, Numerous Class Cases", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 29, pp. 518-528, July 1991.

Reed, I.S., and Yu, X., "Adaptive Multiple-Band CFAR Detection of an Optical Pattern with Unknown Spectral Distribution," *IEEE Trans. on Acoust, Speech and Sig. Proc.*38(10)1760–70 ,1990.

Sabol, D.E., Adams, J.B., Smith, M.O., "Quantitative Subpixel Spectral Detection of Targets in Multispectral Images," *J. of Geophys. Res.* 97 (E2) 2659–2672,1992.

Vane, G., Green, R.O., Chrien, T., et al. "The Airborne Visible/Infrared Imaging Spectrometer (AVIRIS)", *Remote Sens. Environ.* Vol. 44, pp. 127-143, 1993

Yu, X., Hoff, L.E., Reed, I.S., et. al., "Automatic Target Detection and Recognition in Multiband Imagery: A Unified ML Detection and Estimation Approach," *IEEE Trans. on Image Proc.* 6 (1) 143–156,1997.