Anomaly Detection in Hyperspectral Imagery using Statistical Mixture Models

C. J. Willis
BAE Systems Advanced Technology Centre,
Great Baddow, Chelmsford, CM2 8HN

Abstract

The majority of anomaly detection processes used for hyperspectral image data are based on pixel-by-pixel whitening and thresholding operations using local area statistics. This paper discusses an alternative approach in which a mixture model is fitted to the whole of the image. This mixture model may be used to segment the image into component memberships and these may, in turn, be used for anomaly detection.
In this paper the mixture model is generated for the whole scene using the stochastic expectation maximization (SEM) algorithm. The maximum a-posteriori probability (MAP) mixture component for each pixel is then determined. The pixel may then be examined using a conventional statistical hypothesis test to see whether it is plausible that it was drawn from the distribution of the identified component, at a given significance level.
This anomaly detection process has been examined using both synthetic and real hyperspectral imagery and results are presented here for synthesized imagery that includes military target pixels. These results include the component map of the imagery and anomalous pixel maps at given significance levels.

Keywords: hyperspectral imagery, anomaly detection, Gaussian mixture model, stochastic expectation maximization

Introduction

Hyperspectral imaging sensors provide a powerful means for detecting and recognizing ground targets in typical military contexts. They are particularly important where attempts at camouflage and concealment have been employed and provide a means of combating these factors. In some cases knowledge of the spectra of objects of interest is available, and this knowledge may be used to build detection and recognition processes based on matched filtering and supervised classification processes. This paper considers the situation in which no such prior knowledge is available, though the objects of interest (targets) *are* known to be present in only a small proportion of the pixels of the image. Where this is the case anomaly detection processes may be used to detect target regions. These processes attempt to identify those pixels that are unlike the majority of their peers in one way or another. For hyperspectral data the differences exploited are usually based on comparison of the shape of the spectral signature, of a supposed background region, to the test sample. If the test sample is found to be very different from the background it is marked as an anomaly. The key points here are how the background region is defined and how the difference between the test sample and background is measured.
The majority of anomaly detection processes for hyperspectral imagery are based on the RX approach described Reed & Yu [1]. In the RX algorithm the background is taken to be an annular region surrounding a central test pixel. Sufficient statistics are extracted from this annular

region to characterize a multivariate Gaussian distribution, these being the mean vector and covariance matrix of the spectra of the pixels in this region. Where the multivariate Gaussian model is used the Mahalanobis distance of a sample from that Gaussian is known to be distributed chi-squared. Therefore, the Mahalanobis distance of the test pixel to the background Gaussian distribution is measured and hypothesis tested. The test uses the chi-squared distribution to assess whether the sample satisfies the null hypothesis that it is from the same distribution as the background region. If it fails this hypothesis test at the user-specified threshold it is marked as an anomaly. There has been considerable development of RX-based algorithms since the original publication. A more recent development and application of the algorithm is described in Stellman *et al.* [2].

There are several issues associated with the assumptions of the RX algorithm and its derivatives. The first and most significant is that the pixels of the background region represent a single ground cover type and may therefore be adequately modeled by a single multivariate Gaussian. This assumption puts pressure on the implementor to reduce the size of the background region in order that this region does not span too large an area of the image, and is therefore more likely to hail from a single ground cover type. This, however, conflicts with the requirement to make this region as large as possible such that many background pixels are available to calculate the large number of parameters in the required statistics. Wherever the compromise is met there will still be areas of the image which border two or more distinct ground cover regions. In these areas the statistics of the background region will represent a mixture of types and might be statistically significantly different from one of more of the individual types. In this case the RX method will effectively edge-detect in that region, generating a false alarm. The second assumption is that if a pixel is anomalous in its local area then it is anomalous in the whole scene. It is possible to imagine an image in which certain ground cover types are common in the scene, but have a low density in any given area, an example might be scrub plants in a desert environment. In this case the RX algorithm is likely to detect the scrub plants as being anomalous to their surrounding desert regions and therefore generate a potentially large number of false alarms.

The anomaly detection process discussed in this paper is designed to avoid these problems by treating the whole image at one time. In this approach a single mixture model, consisting of several multivariate Gaussian distributions, is fitted to the full image at a single step. Each pixel is then assigned to its most likely component within this mixture model. It is then tested against the statistics of the selected component to examine the null hypothesis that it generated this pixel spectrum.

Naturally this approach has problems of its own. An implicit assumption when extracting the mixture model is that the total of any individual target spectrum in the image represents only a small proportion of the total image area. Therefore the model extraction step must be capable of managing significant image volumes at a time. This, and the fact that the process is both more involved than the equivalent RX based method and essentially iterative in nature, places considerable demands on the computational subsystem.

Background

The scene model examined here is based on a statistical model of the whole image. Alternative whole scene model types are available including those based on endmember sets. For a discussion of the range of techniques see Stein *et al.* [3]. The specific statistical model examined here uses a mixture of Gaussians model to represent the scene. This is described by the following equation

$$f(x \mid \Phi) = \sum_{i=1}^{K} \alpha_i f_i (x_i \mid \phi_i)$$

where the values x are the observed spectral pixels, assumed a function of the model and its parameters Φ. Here K is the number of mixture components, the α_i are the mixture proportions and ϕ_i are the individual component statistics. The x_i are the elements of x which are found to belong to mixture component i. In this paper the mixture components are assumed to be multivariate Gaussian distributions. For this case the ϕ_i are comprised of the mean vector μ_i and the covariance matrix Σ_i only. In order to use the model for anomaly detection we must infer K, the α_i, μ_i, Σ_i and x_i.

The probability density function for the multivariate Gaussian distribution is

$$p(x \mid \mu, \Sigma) = \frac{1}{(2\pi)^{n/2} \mid \Sigma \mid^{1/2}} \exp(-\Delta^2 / 2)$$

where $\Delta^2 = (x-\mu)^T \Sigma^{-1} (x-\mu)$ is the Mahalanobis distance of the sample and n is the dimensionality of the problem. This PDF is used as the f_i in the function above.

Procedure

The mixture based anomaly detection process is based on the following steps:

1. Spectral pre-processing using, for example, Principal Components Analysis (PCA);
2. Mixture model extraction using a variant of the Stochastic Expectation Maximization (SEM) algorithm for the training data set;
3. Calculation of mixture membership for the test data set using the Maximum A-posteriori Probability (MAP) approach;
4. Calculation of the mixture-element posterior Mahalanobis distance statistic for each test sample, and;
5. Comparison of the test statistic with the appropriately parameterized hypothesis test distribution to identify those test elements that do not satisfy the null hypothesis.

Throughout this paper pre-processing using PCA (step 1) is always used. This step is arranged such that the number of principal components accounting for 99.9% of the variance of the training data are used in mapping the imagery into the new subspace.

Here, for steps 3, 4 and 5, the test data is identical to the training data. This represents the scenario in which collected data is processed “live” whilst the sensor over-flies the region of interest rather than the situation in which the target detection processes are pre-loaded with data derived from a corresponding, preferably target free, associated region. There might be some advantage to considering the latter approach in an operational scenario, not least because the online component of the image processing operations is significantly reduced.

The anomaly detection procedure in step 5 uses the Mahalanobis distance of each pixel to its classified mixture component

$$\Delta_i^2 = (x_i - \mu_i)^T \Sigma_i^{-1} (x_i - \mu_i) .$$

This may be compared with the appropriate cumulative density function (CDF), in this case of the chi-squared distribution, in order to identify anomalies.

Synthetic Hyperspectral Imagery

Results are presented for a synthetic hyperspectral image produced using the GCI toolkit. The preparation of this data is discussed in Bishop *et al.* [5]. This image is one of a series synthesized at a range of sensor altitudes (translating to resolution on the ground), times of day and seasons. The image and anomaly ground truth is presented in Figure 1. The dataset is 256 pixels square and was originally synthesized with 420 bands ranging from 0.4 to 2.5 microns, that is through the visible, near infrared and short wave infrared parts of the spectrum. For the current study a dataset of 210 bands has been analyzed by simply selecting alternate bands across the full spectral range. The principal components analysis step reduces this to a six dimensional subspace within which the mixture extraction and anomaly detection steps are applied.

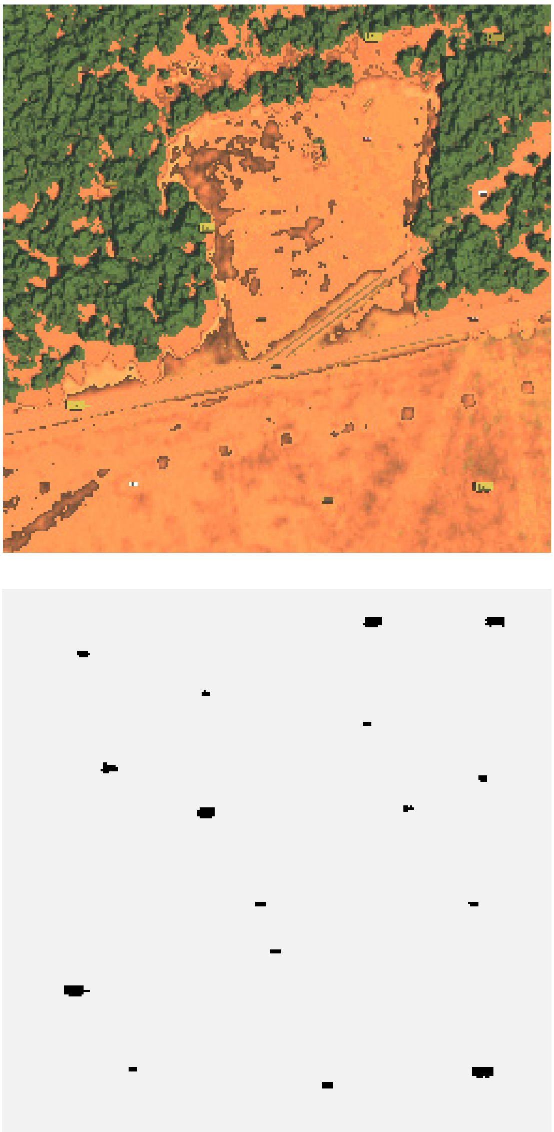

Figure 1 GCI Toolkit synthesized hyperspectral image in a near true colour representation (top) and associated Ground Truth Anomaly Map (bottom)

The image displayed at the top of the figure is an approximate representation of true colour and shows two principal ground cover types, forest and open ground. Embedded into the image are sixteen target-like objects with position and extent as displayed in the anomaly map at the bottom of the figure. It may be noted that though it is a reasonably simple image with regard to ground cover types, there is considerable structure, with texture and shading in the forested regions, deep shadow in the vicinity of the target objects, and regional spectral variations in the open ground.

Figure 2 shows the results from the Gaussian mixture model based anomaly detection process applied to the dataset displayed in Figure 1. In this example the process was initiated with a maximum of five components in the mixture model and a minimum component size of 5% of the image. The SEM process converged to a model with three components shown at the top of the figure. The three components of the mixture are assigned: one representing the forested region (green); one the majority of the open ground (burnt orange), and; the final (yellow) component representing deep shadows, several of the target regions and the road or track-way. Interestingly the latter item does not appear distinctly different from the other open ground regions in the true-colour image representation displayed above, but is obviously sufficiently spectrally different to be picked out by the mixture modelling process.

The anomaly map displayed at the bottom of Figure 2 shows both the detected anomalies and ground truth at a significance of 0.01%. Here false alarm detections (false positives) are displayed in green, missed detections (false negatives) in magenta and true detections in black. The figure illustrates the successful detection of eleven of the sixteen embedded target objects with the larger proportion of most of these objects having been detected. There are also several significant regions of false alarm detections produced by the algorithm and a large number of smaller single pixel false alarms. It is possible that post processing operations might be used to filter out some of these areas

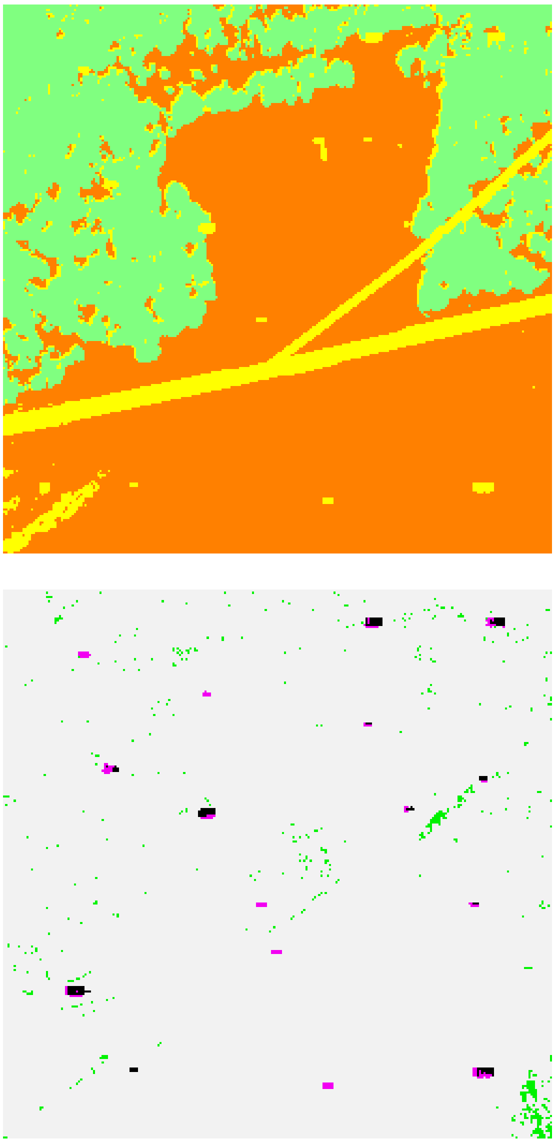

Figure 2 Image Component Map, 3 components (top) and Anomaly Map, 0.01% significance (bottom)

Conclusions

The paper examines an anomaly detection method based on Gaussian mixture models. The mixture model, derived using the stochastic expectation maximization method, is used to provide a representation of the typical scene content. This model is used to identify anomalous scene pixels using statistical hypothesis testing methods. The method has been evaluated using a synthetic hyperspectral image generated for the purpose of testing anomaly detection processes. This dataset comes complete with embedded target regions that have spectral properties that should make them anomalous to the rest of the natural scene. For this image a map of these target object regions is available so that evaluations of anomaly detection processes may be carried out. Results for the component map generated by the mixture extraction process discussed in the paper generally follow the pattern that might be expected by examining the image by eye. Regions of distinctly different ground cover are placed in separate components of the mixture model, and the same components are used across these ground cover regions. Results for the anomaly detection process are presented at a single hypothesis test threshold. At this point the results show the successful detection of eleven of the sixteen embedded target objects at a moderate false alarm rate.
Results have also been produced using this method for real hyperspectral imagery from a range of sensors, some containing known target objects. Generally similar results have been produced on these data, suggesting that the method is robust across a range of sensors and ground cover scenarios.

References

1. Reed, I. S., Yu, X., "Adaptive Multiple Band CFAR Detection of an Optical Pattern with Unknown Spectral Distribution", IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. 38, pp. 1760-1770, 1990.

2. Stellman, C. M., Hazel, G. G., Bucholtz, F., Michalowicz, J. V., Stocker, A., Schaaf, W., “Real-time Hyperspectral Detection and Cuing”, Optical Engineering, Vol. 39, No. 7, pp. 1928-1935, July 2000.
3. Stein, D. W. J., Beaven, S. G., Hoff, L. E., Winter, E. M., Schaum, A. P., Stocker, A. D., “Anomaly Detection from Hyperspectral Imagery”, IEEE Signal Processing Magazine, January 2002.
4. Landgrebe, D. A., Biehl, L., “AVIRIS Indian Pines Reflectance Data: 92AV3C”, available as a part of the documentation for the MultiSpec hyperspectral imagery analysis environment from http://dynamo.ecn.purdue.edu/~biehl/MultiSpec/documentation.html, last update May 2003.
5. Bishop, G. J., Geatches, R. M. and Willis, C. J., Unpublished MoD Report, December 2003.

Acknowledgements

The work reported in this paper was funded by the Electro-Magnetic Remote Sensing (EMRS) Defence Technology Centre, established by the UK Ministry of Defence and run by a consortium of SELEX Sensors and Airborne Systems, Thales Defence, Roke Manor Research and Filtronic.