

Hyperspectral Anomaly Detection and Un-mixing in Fat-Tailed Clutter

Mark Bernhardt, Jamie Heather and Oliver Watkins
Waterfall Solutions Limited
Parklands, Guildford, Surrey, GU2 9JX, UK

Abstract

Detection of anomalies in hyperspectral clutter is an important task in military surveillance. Most algorithms for unsupervised anomaly detection make either explicit or implicit assumptions about hyperspectral clutter statistics: for instance that the abundance is either normally distributed or elliptically contoured. In this paper we investigate the validity of such claims. We show that while non-elliptical contouring is not necessarily a barrier to anomaly detection, it may be possible to do better. In this paper we develop a method for visualizing the degree of elliptical contouring of hyperspectral data. Having observed that most data fails to be elliptically contoured, we develop a new method for anomaly detection that has good performance on non-EC data.

Keywords: Hyperspectral Imaging, Clutter Statistics, Anomaly Detection

Introduction

Anomaly detection is the task of finding pixels in Hyperspectral Image (HSI) data that are not typical compared either to their local spatial neighbourhood or to the overall image statistics. The archetypal anomaly detection method is the famous RX [1] algorithm. This algorithm makes what is essentially a multivariate Gaussian assumption about the background statistics in order to compute anomalous pixels. More recently it has been observed [2-5] that HSI data has ‘fatter tails’ than a Gaussian. That is, the probability of natural pixels occurring far from the mean (in terms of Mahalanobis distance) is higher than the Gaussian model would suggest. The radar community uses the term ‘spikey’ for the analogous phenomena in radar clutter statistics. Recent work [2-7] has modified the Gaussian assumption in order to improve the modeling of HSI clutter. One promising approach is to widen the class of distributions used for the modeling to the Elliptically Contoured (EC)

distributions, of which the multivariate Gaussian is a special case. The general form for the probability density function of an EC distribution is

$$f_n(\mathbf{x}) = (2\pi)^{-\frac{n}{2}} |C|^{-\frac{1}{2}} h_n(d)$$

where n is the dimensionality of the random vectors \mathbf{x} and C is the covariance matrix. The requirements on the function h are that it is positive and a monotonically decreasing function of d where d is the squared Mahalanobis distance given by

$$d = (\mathbf{x} - \boldsymbol{\mu})^T C^{-1} (\mathbf{x} - \boldsymbol{\mu})$$

where $\boldsymbol{\mu}$ is the mean of the random vectors, \mathbf{x} . Clearly when we set $h_n(d) = \exp(-d/2)$ the multivariate Gaussian is recovered, so these functions generalise the multivariate Gaussian in a straightforward and natural way ensuring that the equi-probability contours are ellipsoids. Given that hyperspectral data is probably fat-tailed

rather than Gaussian it is important to consider the importance of this to anomaly detection algorithms like RX that make a Gaussian assumption. RX actually computes d , as the measure of how anomalous a pixel is, or the pixel strangeness. This means that it is really only the RX threshold that is altered between Gaussian or fat-tailed data. In practice RX is run in a ‘Constant False Alarm Rate’ (CFAR) mode with feedback from the number of detections to control the threshold (equivalently just rank the anomalies and pick the top N). This means that the threshold will automatically adapt to take account of the tail statistics (Gaussian or not). Thus, RX should be optimal for all EC data, not just multivariate Gaussian data. As part of an on-going project, work has been undertaken by Waterfall Solutions to try to understand the nature of hyperspectral background statistics in more depth. In particular, simple Generative models have been developed which model the process of competition in vegetation which leads to particular distributions of abundance. This in turn leads to particular hyperspectral statistics. The hope is that this may give some insight into the type of distributions that one ought to consider for modeling this data. Some of this work is reported in section 2. One intriguing aspect of this work is that most of the models lead to fat-tailed clutter, in agreement with observations of real data. In addition however, many of the models developed also produced results that were not EC. This led us to investigate whether real data was indeed EC or not. This is discussed in some detail in Section 3 of this paper. Having devised some tests and tools to understand this question for real data (and learning that it is in general not well described by a single EC distribution), we attempted to devise a new anomaly detector to overcome the shortcomings of RX. This is described in Section 3.

Elliptical Contouring

To our knowledge, testing for elliptical contouring has not been applied in the area of hyperspectral clutter data before. However, there are some tests that have been developed which can, in principle, be applied. One such test has been implemented based upon [8, 9] and has been verified using various synthetic data sets that we created. When this test was applied to hyperspectral data it invariably indicated that it was not EC. While this is interesting, it does not by itself give any guide as to the nature of the lack of EC, or what can be done about this in anomaly detection. In order to answer these questions we developed a technique for visualising the EC nature of high dimensional data. Clearly, a good place to start is with polar coordinates, however it is also necessary to first transform the data into a lower dimensional sub-space through standard Principal Component Analysis (PCA). Since this is a linear transform it does not alter the EC properties of the data. The visualisation approach works for any number of principal components, so choosing the number is not critical – one can choose a larger number and then reduce it. Having done this we may convert this data into polar coordinates:

Given a pixel in (now in d dimensions after PCA) $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_d]$, the polar coordinates are as follows:

$$r = |\mathbf{x}|$$

$$\cos(\theta_1) = \frac{x_1}{r}$$

$$\cos(\theta_i) = \frac{x_i}{x_{i-1} \tan(\theta_{i-1})}$$

where the subscript i refers to indices greater than 2. If the data is EC then its projection onto the unit d-sphere after PCA will have a uniform density on the surface of the sphere. Note that the converse of this is not true. However what is important is

that if the data appears non-uniform it definitely *is not* EC.

One might assume that HS data will be uniformly distributed across these angles if it is uniformly distributed over the sphere. This is not, in fact, the case. This is simply the effect of the ‘poles’. On a sphere with a uniform density of points, the marginal equatorial distribution is uniform in angle, but clearly the marginal distribution with respect to the latitude coordinate is not, and shows a reduction in density at latitude $+90^\circ$ and -90° corresponding to the poles. It would be convenient for visualisation purposes if EC data such gave a uniform distribution in all of the coordinates. This is straightforward to accomplish through a non-linear transformation of the angles.

We have plotted transformed angular locations against each other, producing sets of scatter plots such as those shown in figures 1-2. These diagrams show the radial and (transformed) angular coordinates in a pair-wise scatter-plot

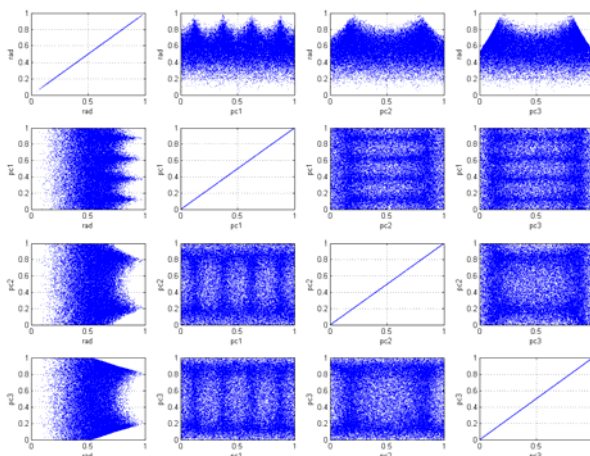


Figure 1 Scatter plots of transformed angular locations of synthetic data uniformly distributed on a cube. Note the regular patterns in the plots, in which the original structure may be clearly seen.

It is clear from these visualisations that the method allows the detection of deviations from EC in both real and synthetic data.

The AVIRIS data in Figure 2 is particularly interesting as it shows clear evidence of non-EC distributions. Indeed, much of the hyperspectral data from a range of sensors that we have analysed shows deviation from EC behaviour

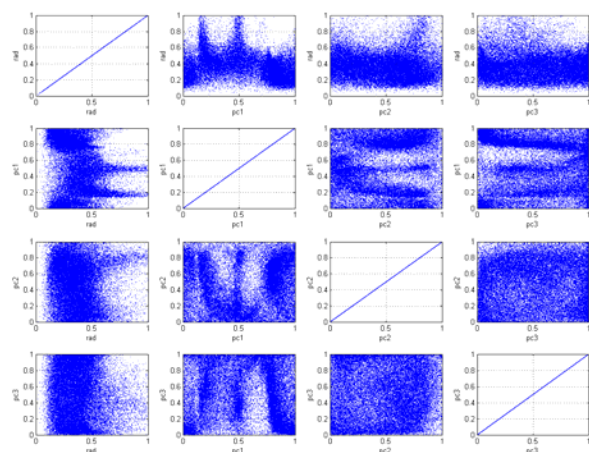


Figure 2 Scatter plots of transformed angular locations of vegetation background data from the AVIRIS sensor [10].

A New Anomaly Detector

Inspection of the plots in the previous section leads to some observations on the nature of the data:

- The data is clearly not elliptically contoured.
- There are areas of higher ‘angular density’ (those directions in which there are many pixels) and other areas of much lower ‘angular density’.
- There are clear ‘spikes’ of higher radii, at particular angular locations.

These observations guide thinking on the nature of RX. The ‘spikes’ of higher radii are not, in general, anomalous pixels as there are far too many of them; it is simply

that pixels in that particular spectral direction tend to lie further from the mean. RX will, on the other hand, declare these pixels to be highly anomalous. Is it therefore more appropriate to measure anomalousness in terms of the number of angular near-neighbours of each pixel? Such an approach would take no account of the radius of each pixel and would therefore be able to detect anomalies which would not be picked up by RX.

The naïve method for determining the number of pixels lying within a cone of slope φ would represent an $O(N^2)$ computation which would render the algorithm rather slow. However by judicious use of Mont-Carlo sampling techniques this algorithm can be made to run in linear time in the number of pixels.

As $\frac{x^T x_i}{|x||x_i|} = \cos(\theta)$, where θ is the angle

between the vectors x and x_i , it is clear that if $x^T x_i > \cos(\theta)$ then the pixel located at x_i is located within the cone. Extension of this to many dimensions is trivial, although determining the angle φ is not. The following algorithm will act as a ‘directional detector’:

For each pixel in image **Determine** number of angular near neighbours.

End

Set strangeness equal to the reciprocal of then number of near neighbours.

Alert for those pixels with the highest strangeness.

While it may seem odd to introduce the reciprocal step here, it is used simply in order for the most anomalous pixel to have the highest strangeness, in common with other anomaly detectors.

How does one choose the size of the multi-dimensional cone for which we determine inclusion or otherwise? We wish to be able to determine a quantity which will give

consistent results for all data dimensionalities. The solution we use in this case is to determine a solid angle which we expect to encompass some fixed proportion of the data. So far in this study a suitable value has been found to be 0.1, meaning that for EC data, we expect 10% of the pixels to lie within the cone emanating from the origin.

The algorithm described in this section operates in time of order N^2 , so anomaly detection in large datasets can be slow. To alleviate this problem we adopt a randomised algorithm which will typically produce results of a similar quality to those produced by the full-complexity algorithm outlined earlier.

Tests on the relative performance of the two algorithms show that the strangeness rankings of the two tend to be similar.

Results

A useful tool in comparing anomaly detectors against each other is to plot the ranks of each pixel’s strangeness measure against each other on a scatter plot. By observing to which side of the diagonal line ($y=x$) the majority of anomalies lie on this plot, we may gain a clear understanding of which anomaly detector has performed better. The anomaly detection results presented here are for the vegetation part of an AVIRIS cube. The anomalies are pixels from an urban part of the same cube that have been embedded at different mixing fractions. We present here four sets of results for anomalies which are strongly embedded, moderately embedded, weakly embedded and of mixed embedding strength. Figures 3-6 show the results obtained.

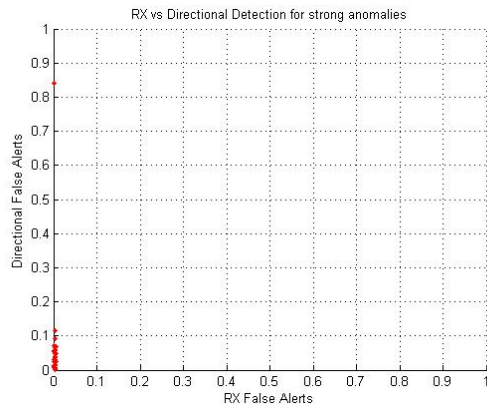


Figure 3 *RX vs Directional Detection for strong (pure pixel) anomalies. Note that in general RX has very low false alert rates and hence the anomalies tend to lie above the ($y=x$) line. This is as expected because these very strong anomalies are the easiest to detect.*

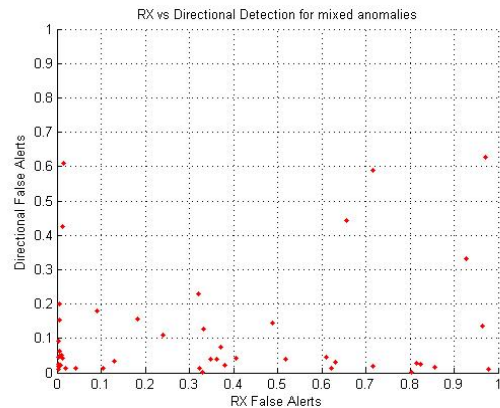


Figure 6 *The anomalies in this picture are of a variety of strengths from 0.1 to 0.7. Clearly the directional detector is superior in these circumstances.*

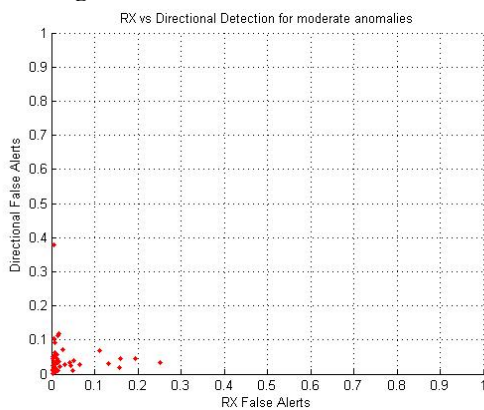


Figure 4 *Although for medium strength (0.6) anomaly embedding RX still outperforms directional detection it is clear that some anomalies will be detected substantially earlier using the direction detector.*

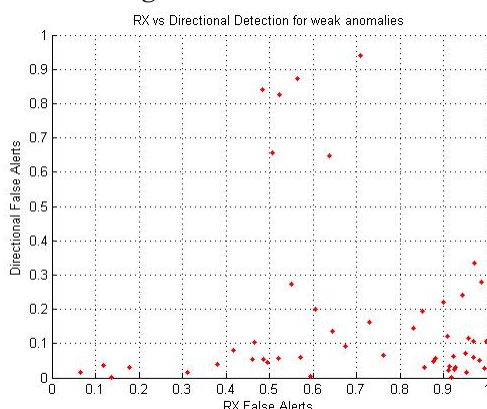


Figure 5 *For very weak anomalies (0.2 embedding strength) the direction detector is clearly superior, with almost every anomaly being detected earlier by the new detector.*

Summary

In this paper we have noted that real and synthetic data show a lack of elliptical contouring. This led us to analyse this feature in more detail. A new detection algorithm has been developed which takes a new approach to anomaly detection: detecting pixels based on the rarity of their ‘spectral direction’ with respect to the mean of their background, rather than their Mahalanobis distance from the mean of their background. This detector has had some success in detecting synthetic anomalies in real backgrounds, particularly strongly sub-pixel anomalies (exactly those which most algorithms find hard to detect). This algorithm presents a promising new direction in sub-pixel anomaly detection in hyperspectral imagery. Further enhancements are possible including the combination of Mahalanobis distance with spectral direction. These are currently being developed.

References

1. I.S. Reed, X. Yu, "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution", *IEEE Trans, Acoustics, Speech and Signal Processing*, vol. 38 no. 10, pp1760-1770, IEEE, 1990.
2. D.G. Manolakis, D. Marden, J.P. Kerekes, G.A. Shaw, "Statistics of hyperspectral imaging data", *Algorithms for Multispectral, Hyperspectral and Ultraspectral Imagery VII*, S.S. Shen, M.R. Descour, vol. 4381, pp308-316, SPIE, Orlando, 2001.
3. D. Marden, D. Manolakis, "Algorithms for hyperspectral imaging data exploitation using non-Gaussian elliptically contoured distributions", *Imaging Spectroscopy VIII*, S.S. Shen, vol. 4816, SPIE, Seattle, WA, 2002.
4. D. Manolakis, G. Shaw, "Detection algorithms for hyperspectral imaging applications", *IEEE Signal Proc. Magazine*, pp29-43, IEEE, 2002
5. M. Bernhardt, W. Oxford, P. Clare, V.A. Wilkinson, D.G. Clark, "Statistical detection algorithms in fat-tailed hyperspectral background clutter". *Proc. SPIE Vol. 5573* pp215-225. Nov. 2004.
6. P. Clare, M. Bernhardt, W. Oxford, S. Murphy, P. Godfree, V. Wilkinson, "A new approach to anomaly detection in hyperspectral images", *Algorithms and Technologies for Multispectral, Hyperspectral and Ultraspectral Imagery IX*, S.S. Shen, P.E. Lewis, vol. 5093, pp17-27, SPIE, Orlando, 2003.
7. M. Bernhardt, J.P. Heather. M.I. Smith. "New Models for hyperspectral anomaly detection and un-mixing". *Proc. SPIE Vol.5806* pp720-730. Jun 2005
8. Li-X Zhu and G. Neihaus, "Conditional tests for elliptical symmetry", *Journal of Multivariate Analysis*, Vol. 84, pp284-290, 2003.
9. Li-X Zhu and G. Neihaus, "Nonparametric Monte-Carlo tests for multivariate distributions", *Biometrika* Vol 87, No. 4, pp919-928, 2000.
10. <http://aviris.jpl.nasa.gov/>

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 SELEX Sensors and Airborne Systems, Thales Defence, Roke Manor Research and Filtronic.