LA-UR- 12-00188

Approved for public release;
distribution is unlimited.

Title:	Detection of spectrally sparse anomalies in hyperspectral imagery
Author(s):	James Theiler, ISR-3 Brendt Wohlberg, T-5 Los Alamos National Laboratory
Intended for:	Southwest Symposium on Image Analysis and Interpretation Santa Fe, NM 22-24 April 2012

Los Alamos
NATIONAL LABORATORY
EST. 1943

Form 836 (7/06)

Detection of Spectrally Sparse Anomalies in Hyperspectral Imagery

James Theiler
Space Data Systems Group
Intelligence and Space Research Division
Los Alamos National Laboratory
Los Alamos, NM 87544, USA
Email: jt@lanl.gov

Brendt Wohlberg
Applied Mathematics and Plasma Physics Group
Theoretical Division
Los Alamos National Laboratory
Los Alamos, NM 87544, USA
Email: brendt@lanl.gov

***Abstract*—A variant is presented of the classic problem of anomaly detection in hyperspectral imagery. In this variant, the anomalous signatures are assumed to be additive and to exhibit spectra that are sparse – that is, only a few of the many hyperspectral channels are significantly nonzero.**

When the background data are Gaussian, and there is no structure in the anomalous signatures, then the optimal detector is given by a Mahalanobis distance and exhibits contours that are ellipsoids. When the desired signature is known, then the solution is given by a matched filter that is specifically optimized for that signature; the contours are parallel planes whose orientation depends on the covariance matrix of the background and the desired signature. We address an in-between problem, one for which the detailed signature is not known, but a more generic description of the structure is available.

We propose that this solution might have application to the detection of gaseous plumes, when the chemistry of the gas is unknown. Such plumes have approximately additive effect on their backgrounds, and – especially in the thermal infrared "fingerprint region" – tend to have very sparse absorption and emission spectra.

***Keywords*-hyperspectral imagery, signal processing, anomaly detection, plume detection, sparse modeling**

I. Introduction

Gaseous plumes, particularly in the infrared, exhibit very distinctive signatures of absorption[1] as a function of wavelength. Hyperspectral imagery enables analysts to exploit these distinctive signatures, and to detect specific gaseous chamicals even at very low concentrations, usign matched filters that are tailored both the the specific structure of the chemical signature and to the statistics of the background clutter. [1]–[8]

In general, the better the model of the background clutter, the smaller the deviations from that model that can be reliably detected.

All of these algorithms, however, assume that the gas signature is known. If the chemical itself is not known, then the usual approach is to attack the data with a large library of gas spectra. If the distorting effects of the atmosphere are not accounted for, then these algorithms achieve reduced effectiveness.

[1]Depending on the temperature of the plume relative to its background, the plume may be in absorption or emission; but the spectral shape is the same in both cases.

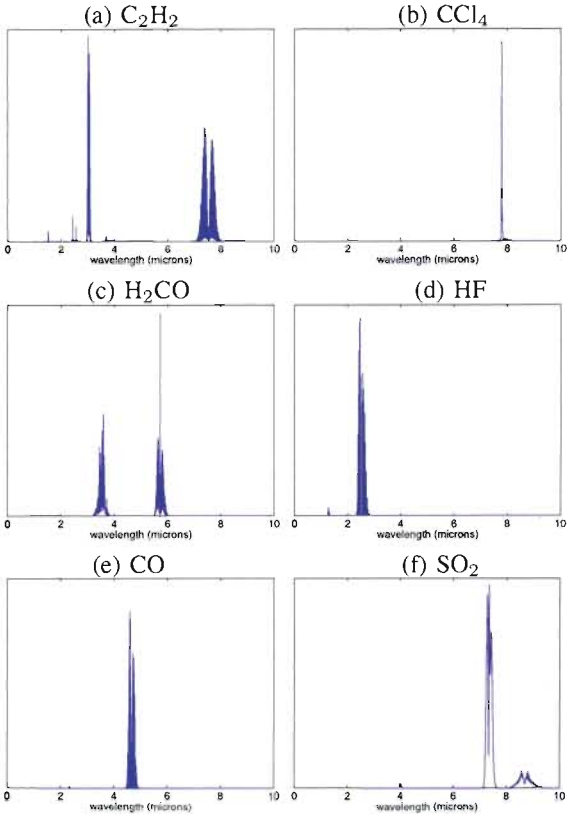

Figure 1. Infrared gas absorption spectra for six diffrent gases, illustrating the sparse structure that is typically inherent in gaseous chemical spectra.

An entirely different approach for detecting targets in clutter is to treat the targets as completely unknown, but as unusual compared to the rest of the background. Again, we seek deviations from the background clutter, but in this case we do not know the direction of that deviation. This is the anomaly detection approach, and most algorithms for finding anomalies are variants of the so-called RX algorithm, which is based on Mahalanobis distance [9]–[12].

It is in the nature of anomaly detection that the anomalies are not well-defined, and when they are modeled, then they are generally modelled (with few exceptions [13], [14]) with a uniform distribution.

One property exhibited by nearly all gaseous chemicals is that the spectrum is composed of a relatively sparse forest of narrow lines; several examples – which were obtained from the HITRAN high-resolution transmission molecular absorption database [15] – are illustrated in Fig. 1. We propose here to exploit this property and search for anomalies that have this spectrally sparse character.

In a hypothesis testing framework, we write

$$\text{Null } H_o: \quad \mathbf{x} = \mathbf{z} \tag{1}$$

$$\text{Alternative } H_1: \quad \mathbf{x} = \mathbf{z} + \mathbf{t} \tag{2}$$

Here, $\mathbf{z}$ is the background and is distributed with some distribution $B(\mathbf{z})$ that is in practice learned from the data. And $\mathbf{t}$ is an additive target. We consider three kinds of targets: in the first case $\mathbf{t}$ is almost completely known (it is known up to a scalar multiplier), in the second case, $\mathbf{t}$ is completely unknown. Our interest for this paper is a third case in which the only thing known about $\mathbf{t}$ is that it is spectrally sparse; that is, the vector-valued $\mathbf{t}$ has mostly zero-valued elements.

We remark that the equation $\mathbf{x} = \mathbf{z} + \mathbf{t}$ has the flavor of a low-rank plus sparse decomposition. [16], [17] In our work here, the "low-rank" component is treated as Gaussian.[2]

II. Derivation of RX for anomaly detection

One way to derive the RX detector is with the Generalized Likelihood Ratio Test (GLRT), in which the anomalous signal is treated as a nuisance parameter, and then one maximizes the likelihood over all possible values of the nuisance parameters. The likelihood of observing $\mathbf{x}$ when $\mathbf{t}$ is the anomalous target signal is proportional to

$$(2\pi)^{-d/2}|R|^{-1/2}\exp[-(\mathbf{x}-\mathbf{t})^T R^{-1}(\mathbf{x}-\mathbf{t})/2] \tag{3}$$

where d is the dimension of the data (number of spectral channels in the hyperspectral image), R is the covariance, and $|R|$ is the determinant of R. When $\mathbf{t}$ is known, the likelihood ratio

$$\mathcal{L}(\mathbf{x}) = \frac{\exp[-(\mathbf{x}-\mathbf{t})^T R^{-1}(\mathbf{x}-\mathbf{t})/2]}{\exp[-\mathbf{x}^T R^{-1}\mathbf{x}/2]} \tag{4}$$

leads to the matched filter

$$\mathcal{M}(\mathbf{x}) = (\mathbf{t}^T R^{-1}\mathbf{t})/2 + \log \mathcal{L}(x) = \mathbf{t}^T R^{-1}\mathbf{x} \tag{5}$$

which is linear in $\mathbf{x}$.

When $\mathbf{t}$ is not known, we can use a GLRT to write

$$\mathcal{L}(\mathbf{x}) = \frac{\max_{\mathbf{t}} \exp[-(\mathbf{x}-\mathbf{t})^T R^{-1}(\mathbf{x}-\mathbf{t})/2]}{\exp[-\mathbf{x}^T R^{-1}\mathbf{x}/2]} \tag{6}$$

[2]For hyperspectral data, the Gaussian model typically includes many small eigenvalues, so the low-rank is not entirely out of place, here.

The numerator achieves its maximum when $\mathbf{t} = \mathbf{x}$, which leads to

$$\mathcal{L}(\mathbf{x}) = \frac{1}{\exp[-\mathbf{x}^T R^{-1}\mathbf{x}/2]} \tag{7}$$

and therefore $\mathcal{A}(\mathbf{x}) = \mathbf{x}^T R^{-1}\mathbf{x}$ is a measure for anomalousness.

III. Spectrally sparse anomalies

In the derivation of $\mathcal{A}(\mathbf{x})$ in the previous section, $\mathbf{t}$ was unrestricted, and we found that the likelihood was maximized when $\mathbf{t} = \mathbf{x}$.

We suggest two approaches for restricting the target to sparse signatures.

1. Strictly constrain the target to have a fixed number of nonzero elements. Let $\mathcal{T}_k$ correspond to the set of targets $\mathbf{t}$ with k or fewer of the components nonzero; that is, $\mathcal{T}_k = \{\mathbf{t} \mid ||\mathbf{t}||_0 \leq k\}$. The likelihood ratio then becomes

$$\mathcal{L}(\mathbf{x}) = \frac{\max_{\mathbf{t}\in\mathcal{T}_k} \exp(-(\mathbf{x}-\mathbf{t})^T R^{-1}(\mathbf{x}-\mathbf{t})/2)}{\exp(-\mathbf{x}^T R^{-1}\mathbf{x}/2)} \tag{8}$$

or equivalently:

$$\mathcal{A}(\mathbf{x}) = \mathbf{x}^T R^{-1}\mathbf{x} - \min_{\mathbf{t}\in\mathcal{T}_k}(\mathbf{x}-\mathbf{t})^T R^{-1}(\mathbf{x}-\mathbf{t}) \tag{9}$$

is the measure for anomalousness. The case $k = d$, where d is the number of spectral channels, leads to the standard RX formulation.

The experiments described in Section IV uses a matching pursuit algorithm to greedily add components to $\mathbf{t}$ until k components have been added.

2. Penalize the likelihood function to favor sparse signatures. Here, we can employ an L1 instead of an L0 metric, and consequently achieve a convex optimization. Rather than restrict $\mathbf{t}$ to a fixed number of nonzero elements, we "nudge" it toward sparsity by altering Eq. (6) with a penalty factor:

$$\mathcal{L}(\mathbf{x}) = \frac{\max_{\mathbf{t}} \exp[-(\mathbf{x}-\mathbf{t})^T R^{-1}(\mathbf{x}-\mathbf{t})/2]\exp[-\lambda||\mathbf{t}||_1]}{\exp[-\mathbf{x}^T R^{-1}\mathbf{x}/2]} \tag{10}$$

This leads to

$$\mathcal{A}(\mathbf{x}) = \mathbf{x}^T R^{-1}\mathbf{x} - \min_{\mathbf{t}} \left[(\mathbf{x}-\mathbf{t})^T R^{-1}(\mathbf{x}-\mathbf{t}) + \lambda||\mathbf{t}||_1\right] \tag{11}$$

as a measure of anomalousness which can be computed as a straightforward quadratic programming problem. The limit $\lambda \to 0$ leads to the standard RX formulation.

IV. Experiment

For the experiment shown in Fig. 3, Fig. 4, and Fig. 5, we used the Indian Pines dataset, but truncated to the first 128 spectral channels (removing the channels where atmospheric absorption was strong). We also removed a 19-pixel wide stripe (on the left of the image) to avoid a single-pixel anomaly that overpowered the rest of the image. To this background image, along a small 2×21 pixel slice of the image, a gradient of plume was added, weaker on the left

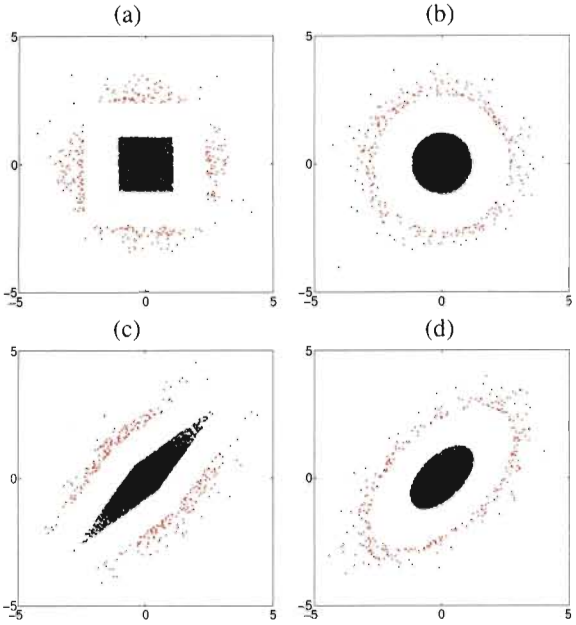

Figure 2. In this simple $d = 2$ dimensional model, we plot the most anomalous (top three percent) and least anomalous (below median in anomalousness) points from 10^4 total points drawn at random from a Gaussian distribution. In (a,b) the distribution is Gaussian with covariance given by the identity; in (c,d) the Gaussian has a correlation of 0.3 in the off-diagonal component of the covariance matrix. We take $k = 1$ in the leftmost panels (a,c), and $k = 2$ in (b,d). We note that $k = d$ corresponds to the standard RX algorithm and in that case elliptical contours of anomalousness are seen. For $k = 1$, we see diamond-shaped contours, corresponding to matched filters with respect to the axis directions.

Figure 3. Truncated Indian Pines dataset. 145×126 pixels, and 128 channels. This broadband image is the sum over all the channels, and although it includes the simulated plume, that plume is too weak to be observed in this projection.

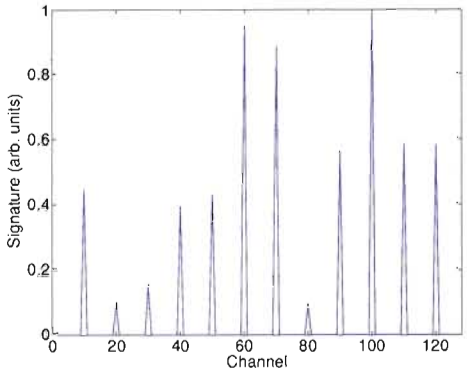

Figure 4. Sparse spectral structure of simulated gas used in Fig. 5.

and stronger on the right. The detection of this plume shows up in Fig. 5 as a dark horizontal stripe.[3] The more sensitive detection observed in Fig. 5(a) is expected since the matched filter in Eq. (5) makes explicit use of the spectrum $\mathbf{t}$. By contrast, the anomaly-based approaches in Fig. 5(b,c) do not use explicit knowledge about the simulated gas spectrum shown in Fig. 4. What Fig. 5 further illustrates is that the spectrally sparse anomaly detector in Eq. (9) provides a more sensitive detection than the traditional RX anomaly detector. Thus, we are able to exploit the sparsity of the signal in Fig. 4 without knowing the details of that signature.

V. Discussion

We have described preliminary efforts to extend standard anomaly detection to the case of a sparse additive target, and noted that this may have applications in plume detection.

One limitation of this work is the restriction to Gaussian backgrounds, which are not particularly realistic. Natural extensions are to heavy-tailed backgrounds [18], [19], to more arbitrary global representations [20], or to local background estimation [21].

Acknowledgment

This work was supported by the Laboratory Directed Research and Development (LDRD) program at Los Alamos National Laboratory.

References

[1] R. C. Carlson, A. F. Hayden, and W. B. Telfair, "Remote observations of effluents from small building smokestacks using FTIR spectroscopy," *Applied Optics*, vol. 27, pp. 4952–4959, 1988.

[2] C. L. Bennett, M. R. Carter, D. J. Fields, and F. D. Lee, "Infrared hyperspectral imaging results from vapor plume experiments," *Proc. SPIE*, vol. 2480, pp. 435–444, 1995.

[3]The authors do not know the cause of the light vertical stripe that is observed in Fig. 5(b,c), but speculate that it is a data artifact, caused perhaps by some kind of data interpolation that led to some unsually un-anomalous pixel values.

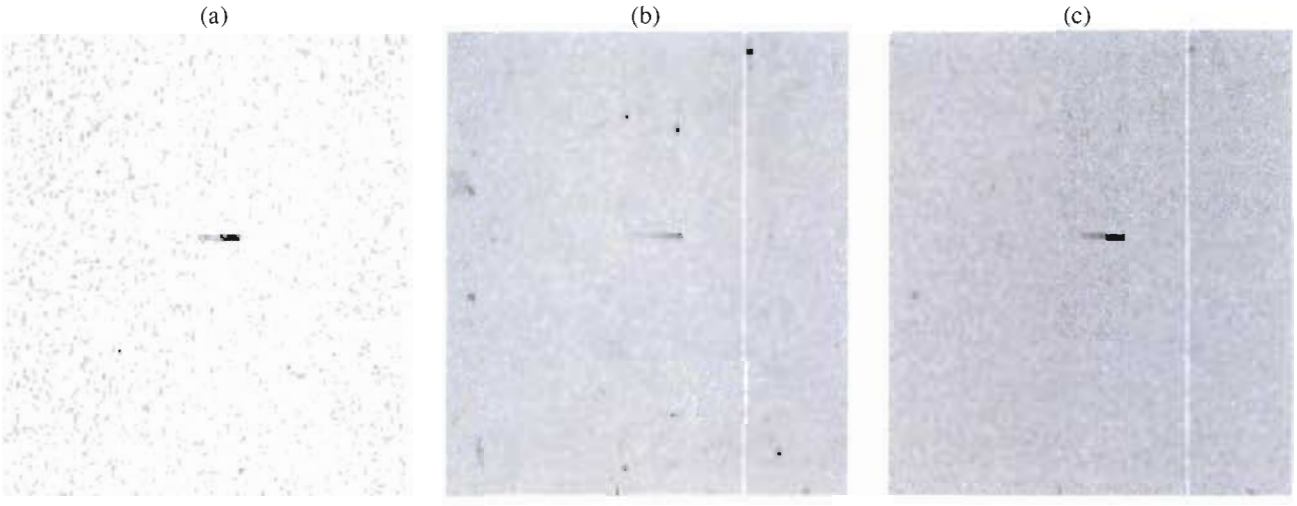

Figure 5. Detecting a weak plume signature, seen here as a small dark horizontal stripe near the center of the image. **(a)** Matched-filter detector. **(b)** Standard RX anomaly detector: **(c)** Spectrally sparse anomaly detector with $K = 10$;

[3] P. V. Villeneuve, H. A. Fry, J. Theiler, B. W. Smith, and A. D. Stocker, "Improved matched-filter detection techniques," *Proc. SPIE*, vol. 3753, pp. 278–285, 1999.

[4] S. J. Young, "Detection and quantification of gases in industrial-stack plumes using thermal-infrared hyperspectral imaging," Tech. Rep. ATR-2002(8407)-1, The Aerospace Corporation, 2002.

[5] B. R. Foy, R. R. Petrin, C. R. Quick, T. Shimada, and J. J. Tiee, "Comparisons between hyperspectral passive and multispectral active sensor measurements.," *Proc. SPIE*, vol. 4722, pp. 98–109, 2002.

[6] B. R. Foy, H. A. Fry, and B. D. McVey, "Approaches to chemical plume detection in hyperspectral infrared imaging," Tech. Rep. LA-CP-03-0103, Los Alamos National Laboratory, 2003.

[7] B. R. Foy and J. Theiler, "Scene analysis and detection in thermal infrared remote sensing using Independent Component Analysis," *Proc. SPIE*, vol. 5439, pp. 131–139, 2004.

[8] J. Theiler, B. R. Foy, and A. M. Fraser, "Characterizing non-Gaussian clutter and detecting weak gaseous plumes in hyperspectral imagery," *Proc. SPIE*, vol. 5806, pp. 182–193, 2005.

[9] I. S. Reed and X. Yu, "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 38, pp. 1760–1770, 1990.

[10] D. W. J. Stein, S. G. Beaven, L. E. Hoff, E. M. Winter, A. P. Schaum, and A. D. Stocker, "Anomaly detection from hyperspectral imagery," *IEEE Signal Processing Magazine*, vol. 19, pp. 58–69, Jan 2002.

[11] A. Schaum, "Hyperspectral anomaly detection: Beyond RX," *Proc. SPIE*, vol. 6565, pp. 656502, 2007.

[12] S. Matteoli, M. Diani, and G. Corsini, "A tutorial overview of anomaly detection in hyperspectral images," *IEEE A&E Systems Magazine*, vol. 25, pp. 5–27, 2010.

[13] J. Theiler and D. M. Cai, "Resampling approach for anomaly detection in multispectral images," *Proc. SPIE*, vol. 5093, pp. 230–240, 2003.

[14] I. Steinwart, D. Hush, and C. Scovel, "A classification framework for anomaly detection," *J. Machine Learning Research*, vol. 6, pp. 211–232, 2005.

[15] L. Rothman and S. Tashkun, "HITRAN on the web," Harvard-Smithsonian Center for Astrophysics (CFA), Cambridge, MA, USA; and V.E. Zuev Insitute of Atmosperic Optics (IAO), Tomsk, Russia. http://hitran.iao.ru/.

[16] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?," *J. ACM*, vol. 58, pp. 11:1–11:37, 2011.

[17] B. Wohlberg, R. Chartrand, and J. Theiler, "Local principal component pursuit for nonlinear datasets," *Proc. ICASSP*, 2012, To appear.

[18] S. M. Adler-Golden, "Improved hyperspectral anomaly detection in heavy-tailed backgrounds," *Proc. IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, vol. 1, 2009.

[19] J. Theiler and D. Hush, "Statistics for characterizing data on the periphery," *Proc. IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 4764–4767, 2010.

[20] H. Kwon and N. M. Nasrabadi, "Kernel RX: a new nonlinear anomaly detector," *Proc. SPIE*, vol. 5806, pp. 35–46, 2005.

[21] N. Gorelnik, H. Yehudai, and S. R. Rotman, "Anomaly detection in non-stationary backgrounds," *Proc. IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, vol. 2, 2010.