

Change Detection for Hyperspectral Sensing in a Transformed Low-dimensional Space

February 2010

Bernard R. Foy
James Theiler

Los Alamos National Laboratory
Los Alamos, NM 87545
LA-UR 10-00455

Keywords: Hyperspectral imaging, change detection, machine learning, clutter, infrared

ABSTRACT

Change detection in hyperspectral imagery is the process of comparing two spectral images of the same scene acquired at different times, and finding a small set of pixels that has the largest apparent spectral change. We present an approach that operates in a two-dimensional space rather than in the original high-dimensional space of the images, which can be greater than 100 spectral channels. The coordinates in the 2-D space are related to Mahalanobis distances for the combined (“stacked”) data and the individual hyperspectral scenes. Several previously developed change detection algorithms can be represented as straight lines in this space, including the hyperbolic anomalous change detector, based on Gaussian scene clutter, and the EC-uncorrelated detector based on heavy-tailed (elliptically contoured) clutter. We show that adaptive machine learning methods can produce new change detectors with good performance that can avoid problems associated with the curse of dimensionality. We investigate, in particular, the utility of the support vector machine for learning boundaries in this 2-D space, using two classes of data to represent pervasive and simulated changes.

1.0 Introduction

Hyperspectral Imaging (HSI) has wide utility for both military and civilian purposes. Change detection is an application of hyperspectral sensors that is important for finding possible new features of interest in a cluttered scene, either a target of interest or activity such as disturbed earth or new structures. Acquiring datasets at two different times, and possibly even with two different sensors, a comparison can reveal the appearance or disappearance of objects with distinct spectra.¹ Unlike target detection, however, the goal in change detection is to detect an entity using no prior information on its spectrum. A common feature of existing change detection algorithms for spectral data is to reduce the multi-dimensional spectral vector (corresponding to one scene pixel) to a scalar quantity, whose magnitude indicates the likelihood of that pixel representing a substantial change.² In this paper, we introduce a scheme in which two of these scalar quantities are produced, and a change detection decision boundary is “learned” from the data in this 2D space. Two classes of data can be generated to facilitate this process: one class consists of pixels with only minor changes (e.g. environmental changes, illumination changes, instrumental noise, etc.), and a second class consisting of major changes whose detection is desired (e.g. grassy vegetation changing to vehicle or structure). Classification algorithms can then be used to discriminate between the minor and major changes. The Support Vector Machine (SVM) can be useful in this regard, and it can be implemented not in the original high-dimensional space of the HSI data but rather in a lower dimensional space derived from the original data. We apply this concept to two examples of HSI data in different spectral regions, the long-wavelength infrared (LWIR) and the visible/near-infrared (VNIR).

2.0 Framework for Anomalous Change Detection

We would like to compare two hyperspectral images, the x -image and the y -image, and find a small set of pixels for which the x -to- y change is unusual compared to the changes exhibited by the rest of the pixels. We recognize at the outset that *all* of the pixels exhibit some degree of change, as a result of environmental or instrumental factors or both.

Let $\mathbf{x} \in \mathbf{R}^{d_x}$ denote the observed radiance spectrum observed at one pixel in the x -image, and $\mathbf{y} \in \mathbf{R}^{d_y}$ be the corresponding pixel in the y -image. We assume that the images are registered, *i.e.* that corresponding pixels \mathbf{x} and \mathbf{y} correspond to the same location in the scene, but we acknowledge that this registration is not always precise.³ The number of spectral channels in the

two images are denoted by d_x and d_y , respectively.

In the machine learning framework introduced in Ref. ⁴, the full set of data is modeled as random samples from a probability distribution $P(\mathbf{x}, \mathbf{y})$. The simplest form of change detection would be straight anomaly detection for this distribution: find the pixels for which $P(\mathbf{x}, \mathbf{y})$ is small, *i.e.* where the pair \mathbf{x}, \mathbf{y} is on the “tail” of the distribution. But that would identify pixels where \mathbf{x} and \mathbf{y} are individually unusual (*e.g.* low or high radiance), whereas we would really like to find where the *relationship* between \mathbf{x} and \mathbf{y} is unusual. If we write $P(\mathbf{x})$ as the distribution just of the pixels in the x-image, then this $P(\mathbf{x})$ will be the marginal distribution of $P(\mathbf{x}, \mathbf{y})$. We can similarly write $P(\mathbf{y})$ as the distribution of pixels in the y-image. Then the product $P(\mathbf{x})P(\mathbf{y})$ describes a distribution of \mathbf{x} and \mathbf{y} values that are independent of each other. When $P(\mathbf{x})P(\mathbf{y})$ is small, that means that either \mathbf{x} or \mathbf{y} (or both) are individually unusual. When the ratio

$$\frac{P(\mathbf{x}, \mathbf{y})}{P(\mathbf{x})P(\mathbf{y})} \tag{1}$$

is small, it signifies that $P(\mathbf{x}, \mathbf{y})$ is small compared to $P(\mathbf{x})P(\mathbf{y})$, which enables us to isolate the notion of anomalous change from that of straight anomaly.

In seeking a function $\mathcal{A}(\mathbf{x}, \mathbf{y})$ which quantifies the “anomalousness” of the change that has occurred at this pixel location, we can take a function of this ratio

$$\mathcal{A}(\mathbf{x}, \mathbf{y}) = f\left(\frac{P(\mathbf{x}, \mathbf{y})}{P(\mathbf{x})P(\mathbf{y})}\right) \tag{2}$$

where f is a monotonically decreasing function of its argument. When the ratio is small, the anomalousness is large.

2.1. Gaussian clutter

The ratio in Eq.(1) takes a simple form when the distribution is modeled as a multivariate Gaussian, a distribution determined completely by the mean spectral vector and a covariance matrix:

$$\begin{aligned}\mu &= \langle \mathbf{z} \rangle \\ Z &= \langle (\mathbf{z} - \mu)(\mathbf{z} - \mu)^T \rangle\end{aligned}\tag{3}$$

where the angle brackets signify a mean over the distribution (in practice, computed from data), and the superscript T denotes a matrix transpose. The density of the distribution at a point $\mathbf{z} \in \mathcal{R}^d$ is given by:⁵

$$P(\mathbf{z}) = (2\pi)^{-d/2} |Z|^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{z} - \mu)^T K^{-1}(\mathbf{z} - \mu)\right]\tag{4}$$

For the change detection problem, we construct “stacked” vectors \mathbf{z} from the pixel pair \mathbf{x}, \mathbf{y} as:

$$\mathbf{z} = \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \in \mathcal{R}^{d_x+d_y}\tag{5}$$

The mean vector for the “stacked” data is:

$$\mu = \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix} = \begin{bmatrix} \langle \mathbf{x} \rangle \\ \langle \mathbf{y} \rangle \end{bmatrix}\tag{6}$$

The covariance of \mathbf{z} is given by:

$$Z = \begin{bmatrix} X & C^T \\ C & Y \end{bmatrix}\tag{7}$$

where

$$\begin{aligned}X &= \langle (\mathbf{x} - \mu_x)(\mathbf{x} - \mu_x)^T \rangle \\ Y &= \langle (\mathbf{y} - \mu_y)(\mathbf{y} - \mu_y)^T \rangle \\ C &= \langle (\mathbf{y} - \mu_y)(\mathbf{x} - \mu_x)^T \rangle\end{aligned}\tag{8}$$

As shown in Ref. 6, we can combine Eq. (4) with analogous expressions for $P(\mathbf{x}), P(\mathbf{y})$ to express the ratio in Eq. (1):

$$\frac{P(\mathbf{x}, \mathbf{y})}{P(\mathbf{x})P(\mathbf{y})} = \left[\frac{|Z|}{|X||Y|} \right]^{-1/2} \exp\left[-\frac{1}{2}(\xi_z - \xi_x - \xi_y)\right]\tag{9}$$

Here, we denote squared Mahalanobis distances for the individual and “stacked” data sets as:

$$\begin{aligned}
 \xi_x &= (\mathbf{x} - \mu_x)^T X^{-1}(\mathbf{x} - \mu_x) \\
 \xi_y &= (\mathbf{y} - \mu_y)^T Y^{-1}(\mathbf{y} - \mu_y) \\
 \xi_z &= (\mathbf{z} - \mu)^T Z^{-1}(\mathbf{z} - \mu).
 \end{aligned} \tag{10}$$

The “prefactor” in Eq. (9) does not depend on the individual pixel spectra \mathbf{x}, \mathbf{y} , so a simple expression for anomalousness is obtained by taking the log, yielding:

$$\mathcal{A}(\mathbf{x}, \mathbf{y}) = -2 \log \left[\frac{P(\mathbf{x}, \mathbf{y})}{P(\mathbf{x})P(\mathbf{y})} \right] - \log \left[\frac{|Z|}{|X||Y|} \right] = \xi_z - \xi_x - \xi_y \tag{11}$$

This has been referred to as the “hyperbolic anomalous change detector”, HACD, due to the hyperbolic decision boundaries.⁷ The quantities ξ_x, ξ_y are related to anomalousness in each dataset taken separately, using the RX definition of anomaly.⁸ The quantity ξ_z can be interpreted as measuring anomalousness in a collective sense: pixel pairs with large ξ_z have some unusual nature in \mathbf{x}, \mathbf{y} , or both.

The Chronochrome change detector has been derived using Wiener filtering. It can be expressed in similar notation, however. Two forms exist, depending on whether one attempts to predict \mathbf{x} from \mathbf{y} , or the reverse, predict \mathbf{y} from \mathbf{x} . They are:

$$\begin{aligned}
 \mathcal{A}(\mathbf{x}, \mathbf{y}) &= \xi_z - \xi_x \\
 \text{or} \\
 \mathcal{A}(\mathbf{x}, \mathbf{y}) &= \xi_z - \xi_y
 \end{aligned} \tag{12}$$

The relationship to HACD is that the sum $\xi_x + \xi_y$ is replaced by the individual anomalousness in either x or y . Finally, if one disregards both ξ_x and ξ_y , then the measure of change detection is the RX anomaly for the stacked data, which we can label RX-ACD:

$$\mathcal{A}(\mathbf{x}, \mathbf{y}) = \xi_z \tag{13}$$

2.2. Heavy-tailed clutter

For HSI data, it has been reported that heavy-tailed distributions are often more appropriate than

Gaussian.^{5,9,10} An example of a heavy-tailed distribution that is elliptically contoured (EC) and that seems to work well is the multivariate t-distribution:

$$P(\mathbf{z}) = |\mathbf{Z}|^{-1/2} H(d, \xi) = |\mathbf{Z}|^{-1/2} \frac{\Gamma\left(\frac{d+\nu}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right) \pi^{d/2} (\nu-2)^{d/2}} \left(1 + \frac{\xi}{\nu-2}\right)^{-(d+\nu)/2}. \quad (14)$$

The parameter ν determines the extent to which the tail is heavier than Gaussian, and as $\nu \rightarrow \infty$, the distribution becomes Gaussian. Ref. 7 showed that this leads to an anomalousness measure given by:

$$\begin{aligned} \mathcal{A}(\nu; x, y) &= (d_x + d_y + \nu) \log(\xi_z + \nu - 2) \\ &\quad - (d_x + \nu) \log(\xi_x + \nu - 2) \\ &\quad - (d_y + \nu) \log(\xi_y + \nu - 2) \end{aligned} \quad (15)$$

This was shown to perform well on simulated and experimental data in the vis-SWIR spectral region in Ref. 7. Ref. 6 further showed that one can approximate the ratio in Eq. (1) for the case of EC data by replacing the denominator, $P(\mathbf{x})P(\mathbf{y})$, with a distribution $P_u(\mathbf{x}, \mathbf{y})$ that treats \mathbf{x} and \mathbf{y} as uncorrelated instead of independent. The covariances X and Y from Eq. (8) are kept, but the cross-covariance C is set to zero. This leads to the ‘‘EC uncorrelated’’ change detector:

$$\mathcal{A}(\nu; \mathbf{x}, \mathbf{y}) = \frac{\xi_z + \nu - 2}{\xi_x + \xi_y + \nu - 2} \quad (16)$$

In the fat-tailed limit $\nu \rightarrow 2$, this is simply:

$$\mathcal{A}(\mathbf{x}, \mathbf{y}) = \frac{\xi_z}{\xi_x + \xi_y} \quad (17)$$

Equations 11, 16, and 17 comprise three expressions for anomalous change detection that depend on the individual anomalousness only through the sum $\xi_x + \xi_y$. In HACD, we take the difference between the ‘‘stacked’’ image anomalousness ξ_z and this sum, whereas in the EC-uncorrelated fat-tailed limit, we take the ratio. In the more general EC-uncorrelated detector of Eq. 14, we use a modified ratio.

3.0 A 2D Visualization

The simple dependence of these three detectors on ξ_z and the sum $\xi_x + \xi_y$ allows us to display them in the two-dimensional space consisting of ξ_z vs $\xi_x + \xi_y$. The data can be displayed as a scatter plot, where the image pair appears as a cloud of points. By simulating changes on this image pair, a second cloud of points can be produced, as shown in Fig. 6 of Ref. 6. A similar plot is shown here in Fig. 1. The cloud of red points was produced by randomly swapping pixel positions in the second image, and computing the Mahalanobis distances as if the covariances had not changed. For the HACD algorithm in Eq. (11), a contour of constant anomalousness in this 2D space is a line with the equation:

$$\xi_z = \text{const} + \xi_x + \xi_y \quad (18)$$

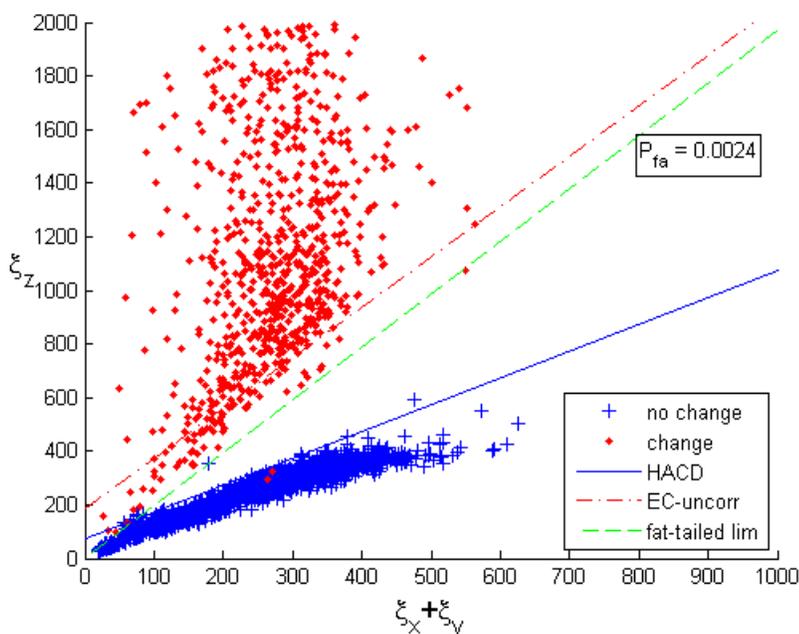


Figure 1. Scatter plot of image-pair data. The blue points correspond to the two input images. The red points correspond to the simulated changes resulting from scrambled pixel positions in the second image. Decision boundaries for three change detection algorithms are plotted as lines, for the same false alarm rate indicated. The two images are synthetic EC data using $\nu = 3$.

The constant represents the intercept in the plot, which increases with increasing detection threshold. The line shown in the plot is then a decision boundary for change detection: points above are classified as having significant change, and points below are regarded as having negligible change, possibly arising from pervasive differences between the images such as

illumination or mis-registration. Similar lines are obtained for the EC-uncorrelated detector (Eq. (14)) and its fat-tailed limit Eq. (15). Note also that a horizontal line on the plot, $\xi_z = \text{const}$, corresponds to an RX anomaly detection performed only on the stacked image \mathbf{z} .

The distribution of red points in Fig. 1 is worth noting. Most of the data are located near the center portion of the plot, but they are distributed upwards with a rapidly declining density. A histogram of ξ_z values is shown in Fig. 2a. This plot is associated not with simulated data, but experimental hyperspectral data described in the next section. The peak in the histogram is at the lowest value plotted, and the points form a very long tail out to high values. If one first takes the log of the data and then compiles the histogram, a very different shape is obtained, Fig. 2b. The peak in the histogram is closer to the “middle” of the span of values, and the tail at high values is less pronounced. Scatter plots for the same HSI data are shown in Fig. 2c, 2d.

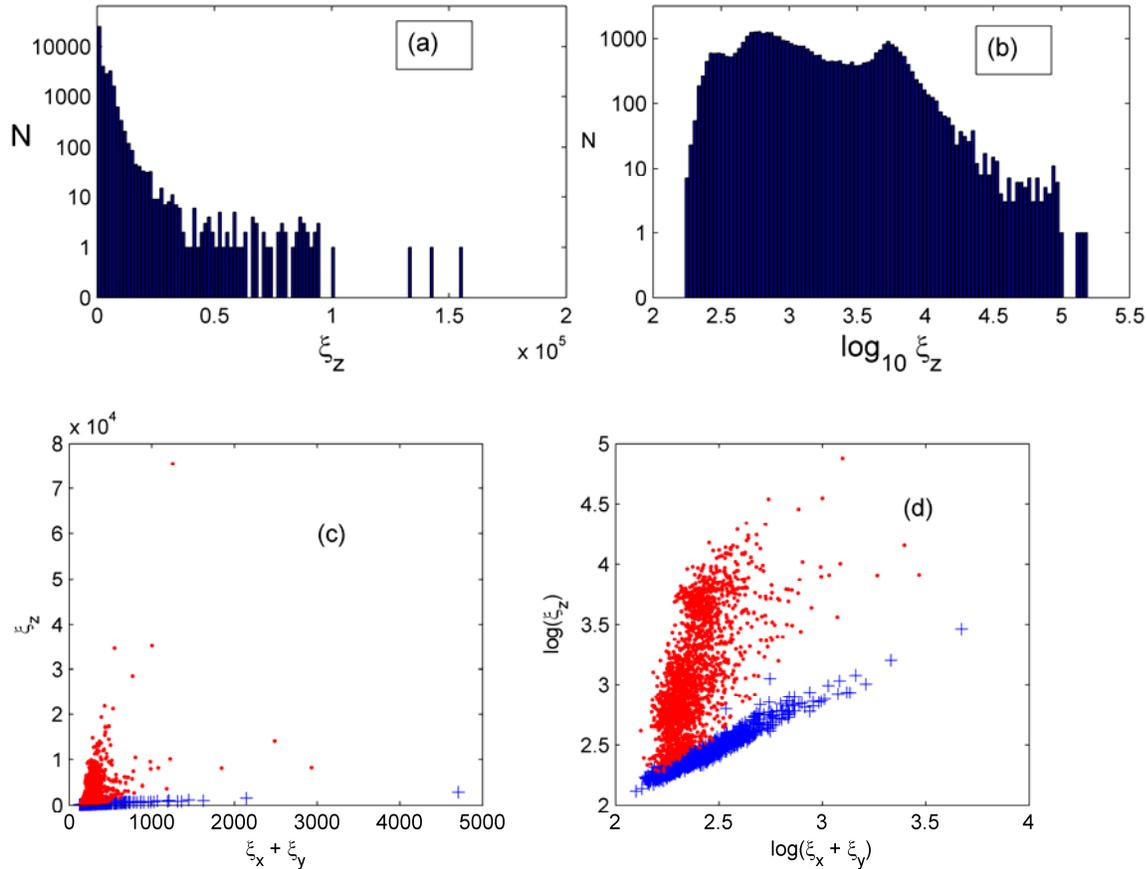


Figure 2. (a,b) Histograms of collective anomaliousness, ξ_z . Computed from the ground-based LWIR image pair. (a) On a linear scale. The histogram is sharply peaked near zero. (b) Histogram of the log of ξ_z . The distribution has a more regular shape: the peak value is closer to the middle. (c,d) 2D scatter

plots on linear scale (c) and log scale (d).

3.1. Experimental Hyperspectral Data

We acquired long-wave infrared (LWIR) hyperspectral data in experiments at Los Alamos National Laboratory.¹¹ An imaging sensor examined a cluttered scene repeatedly over the course of several days, two examples of which are shown in Fig. 3. A scene is comprised of 300X128 spatial pixels, and a spectrum from 742 to 1333 cm^{-1} with 128 spectral channels is acquired at each pixel. The scenes in Fig. 3 are from one selected spectral channel (near 900 cm^{-1}), obtained at two different times about 25 minutes apart. During this time, thermal changes occur in the scene from a variety of effects, mostly solar heating as the sun rises in the sky. Restricted access to the site helped to prevent major changes in the scene from motion of large objects like vehicles. The changes occurring over the course of 25 minutes are largely “natural”, i.e. objects heat up in the sun or cool down in the breeze, and tree branches sway in the wind. These are the kinds of unavoidable effects that might occur while a facility is under surveillance and important changes (like movement of large vehicles, opening of facility doors, disturbance of ground, etc.) are occurring. We would like to detect these “important” changes amidst “natural” changes. The images were found to be well registered to each other, to about a tenth of a pixel, by examining simple difference images such as shown in the Figure. There were some small artifacts arising from bad pixels on the detector array. These were ameliorated by first finding the bad pixels, which are manifested as repeated noisy pixels on the same row in the image,¹² and then by replacing the spectral data at those pixels using a data imputation approach.¹³ This process has negligible effects on change detector performance, however, and won’t be discussed further.

The two-dimensional scatter plot for the experimental LWIR data is shown in Fig. 5. Straight lines indicate decision boundaries for HACD, the EC-uncorrelated detector, and the fat-tailed limit as in Fig. 1. The plot motivates a different way of delineating changed pixels in the image pair via a curved boundary, which should have more flexibility to divide the two sets of points with higher accuracy than a straight line. An example curve is shown in the figure.

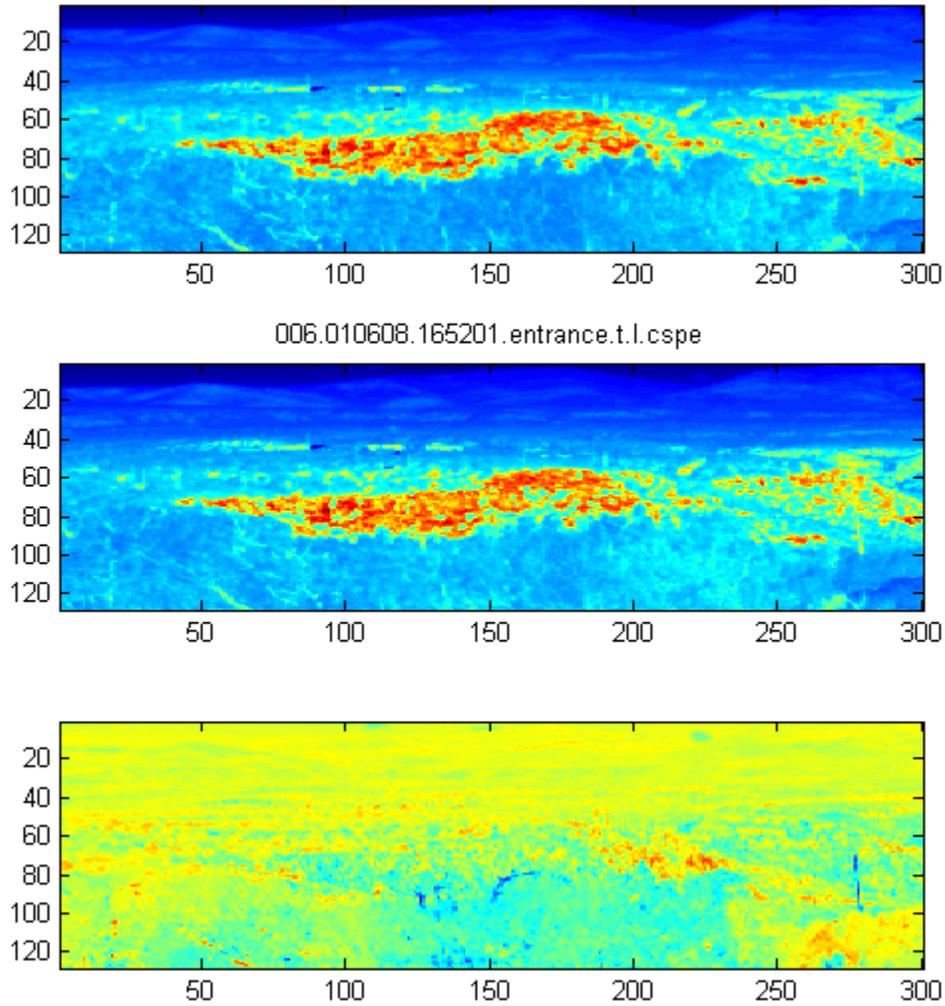


Figure 3. LWIR image pair. Single-band images from a hyperspectral dataset. The top image shows a scene with natural and man-made clutter, acquired at 900 cm^{-1} at LANL. The second image was acquired 25 minutes later. The bottom image is a simple difference between the top two images, exhibiting mostly thermal changes.

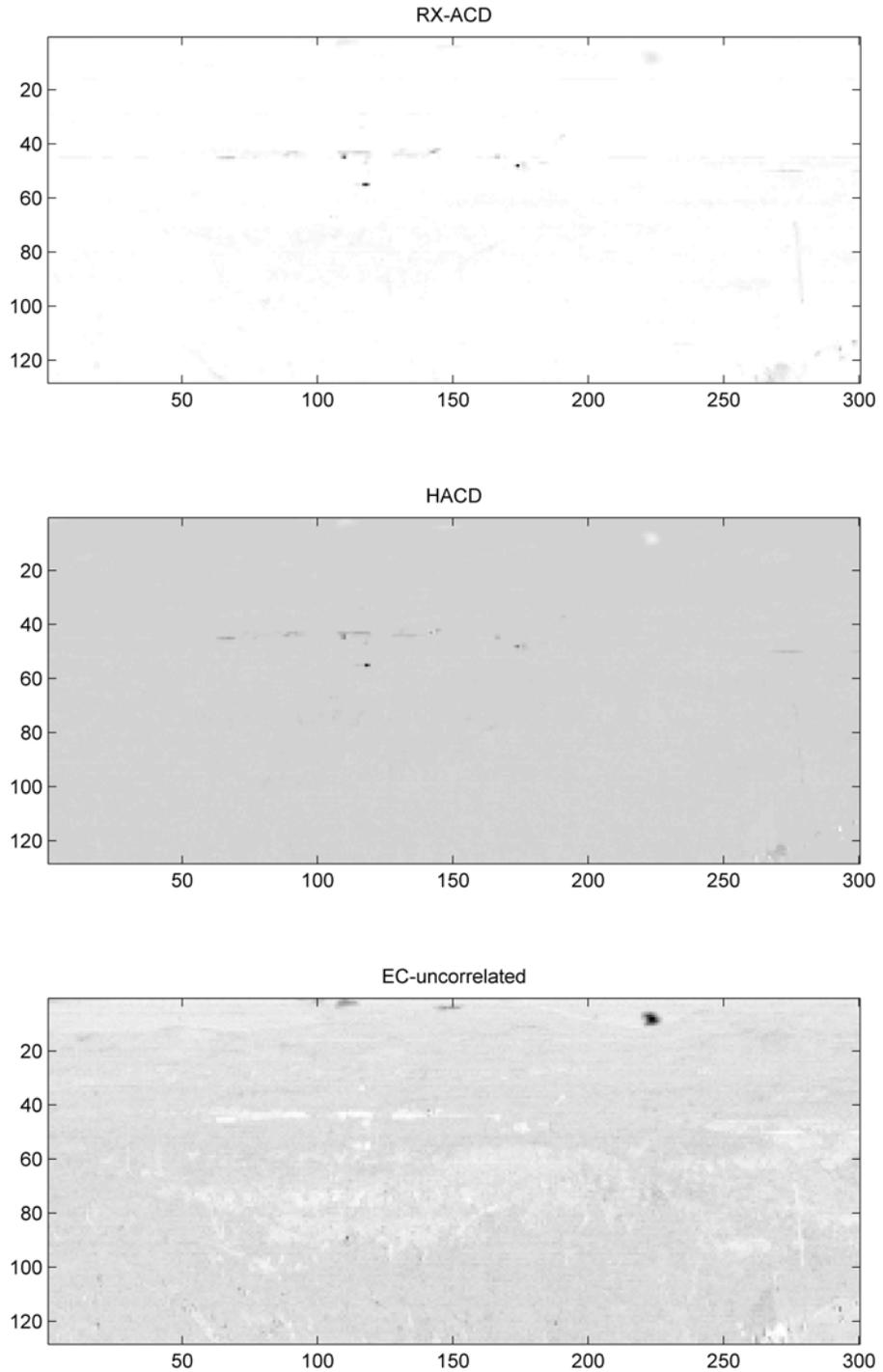


Figure 4. Images from change detection algorithms. For RX-ACD (top), the collective anomalies are highlighted as dark spots. They appear mostly at metal objects in the scene. For HACD (middle), individual anomalies appear as dark spots (large negative values), and anomalous changes appear in white (large positive values). For EC-uncorrelated (bottom), anomalous changes appear as dark spots. Note that HACD and EC-uncorrelated detect a cloud in the sky near the upper right image.

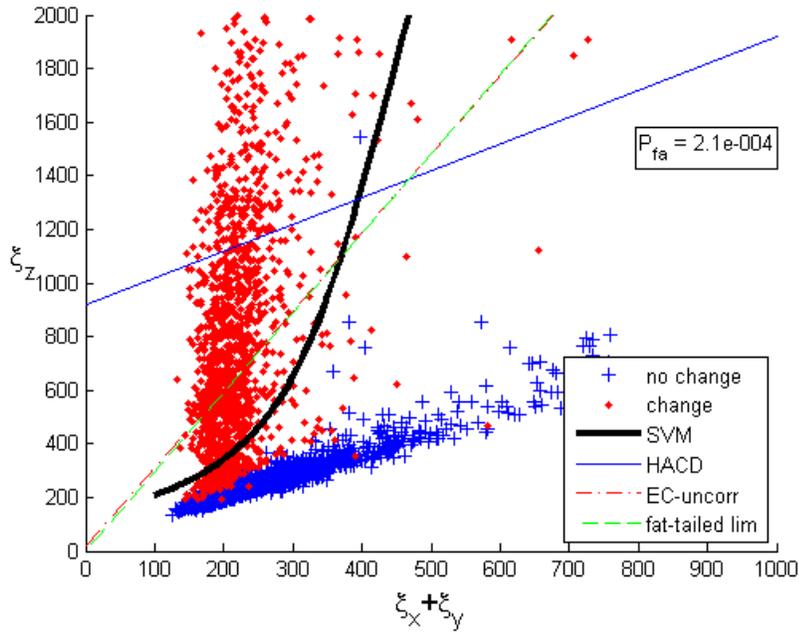


Figure 5. Scatter plot for LWIR image pair. The smooth curve is a decision boundary from an SVM calculation, using a rbf kernel with $\gamma = 10$. Decision boundaries for the other detectors are shown at the same false alarm rate, $P_{fa}=2.1 \times 10^{-4}$. For clarity, only a few points randomly selected from the data are shown.

4.0 Classification in ξ space

One can envision many ways to construct a curved boundary in this reduced-dimensional space of the hyperspectral data. The problem is really a binary classification problem, where two classes are being used to describe levels of change between the scenes. The blue crosses in the lower region of the plot are associated with “natural” changes that occur with most pixels in the scene, while the red points refer to more dramatic changes associated with object movements. (We simulate these changes by creating a large number of pixel location swaps, which amounts to a resampling of the data, as described in Ref. 14.)

Consider first the lower set of points in blue on the plot. Since this is LWIR data, these changes are caused by heating of objects by solar radiation during the 25-minute interval between datasets. Such heating, though, is differential: rock outcroppings warm up more quickly than tree branches, for example, because of differences in solar absorption, heat capacity, and exposure to wind, which cools the objects. In the vis-SWIR spectral region, there can be similar widespread differences due to changing solar illumination and differential shadowing of objects, or due to scene-wide changes in atmospheric interference. Such changes are pervasive in nature, in the sense that they affect virtually all pixels but in varying amounts. Instrumental effects (noise,

misregistration, calibration drift, etc.) are another possible source of pervasive differences.

For the second class of data, the red points in the higher region of the plot, the pervasive differences are supplemented by much more substantial changes that accompany, for example, the substitution of pavement for grass, or a vehicle for a patch of dirt. Note that the plots in Figs. 2 and 4 show only the low- ξ_z region of the 2-D space. Many points are distributed up to very large ξ_z values that correspond to changes with high spectral contrast. The large ξ_z points are mostly irrelevant to the classification problem, because the region of overlap is restricted to low ξ_z values. The significant amount of overlap presents a challenge to classification algorithms potentially useful for the problem, and also to the HACD and other algorithms.

A convenient tool for obtaining a curved decision boundary is the Support Vector Machine. SVM produces boundaries for arbitrary distributions of labeled data, in contrast to HACD (which assumes Gaussian) and EC-uncorrelated (which assumes EC data). SVM optimizes a surrogate loss function that approximates classification error, but is convex.¹⁵ A variety of kernels may be used in the algorithm, but here we use just the Gaussian radial basis function (rbf) kernel. Working in the 2-D space has several advantages: the SVM converges quickly, and the decision boundary can be visualized. One can directly compare decision boundaries for the algorithms.

The black curve plotted in Fig. 4 is from an SVM, computed using the libSVM software.¹⁶ We found that implementing SVM in a logarithmic version of the 2-D space, i.e. $\log(\xi_z)$ vs. $\log(\xi_x + \xi_y)$ was preferable, because of the better-behaved density of points as indicated in Fig. 2. The boundary is re-plotted in linear coordinates in the figure. At large ξ_z , it becomes close to linear, but in the most important region of overlap between classes, it is decidedly curved.

The curved boundary has performance advantages, as shown in Fig. 4. Receiver operating characteristic (ROC) curves were computed for multiple runs of SVM and the other algorithms by splitting the data randomly into half for training, and half for testing. Splitting the sample helps to insure that the SVM classifier performs well outside of the idiosyncrasies of the training set on which it is based, and helps evaluate performance variability and reliability. The results here were obtained by keeping the two SVM parameters fixed: $C = 5$ and $\gamma = 10$. (Here, C controls the amount of regularization, and γ specifies the width of the Gaussian in the rbf kernel). One SVM run produces a decision boundary with some resultant false alarm rate that cannot be set ahead of

time. By varying the weights of the two classes in the computation, we can obtain a series of decision boundaries at different false alarm rates, and trace out the ROC curve shown in the figure. Using different SVM kernel parameters, we might also produce decision boundaries that curl closely around individual data points in the plot, but we expect those classifiers to perform poorly on the test set – they are overfitted to the training set. One can see in the figure that multiple training/testing splits of the data give consistent performance.

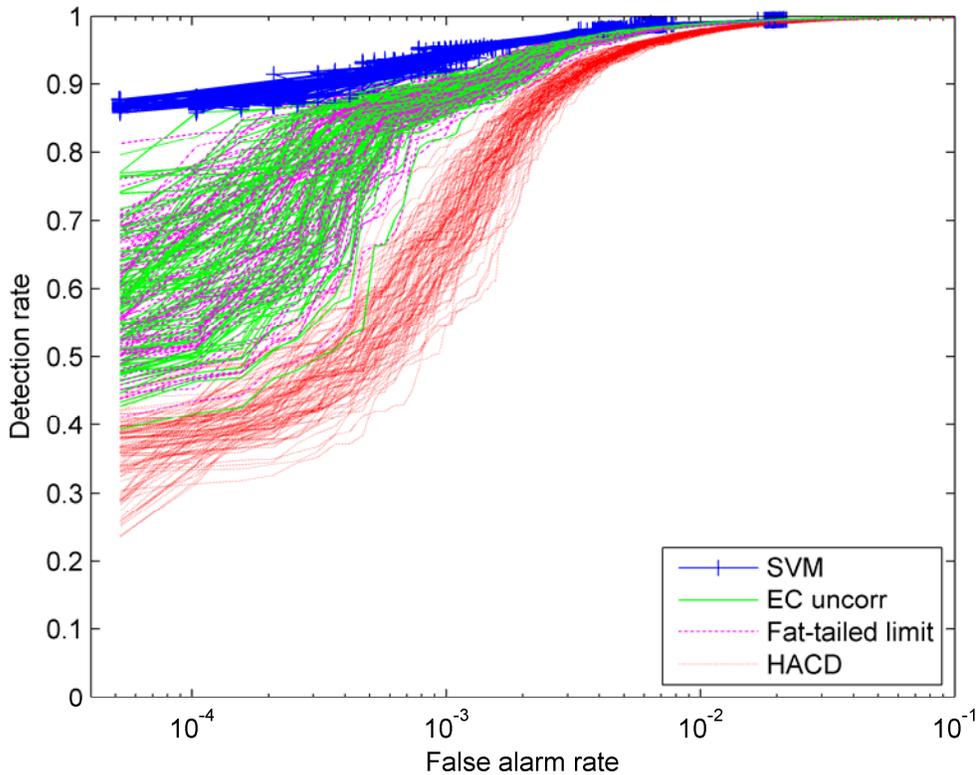


Figure 6. Performance curves for the four different detection algorithms. The LWIR image-pair data were split into testing and training halves. The results from multiple splits are plotted to show the variability of performance from run to run.

Note also that the various test/train splits result in highly variable performance of HACD, EC-uncorrelated, and fat-tailed limit. In Fig. 5, we show only the mean performance curves of these algorithms for clarity. EC-uncorrelated and fat-tailed limit are almost identical, and each outperforms HACD by a substantial margin. Although there is considerable variability in the SVM performance, it consistently beats the mean performance of the detectors that have linear boundaries in the 2-D space.

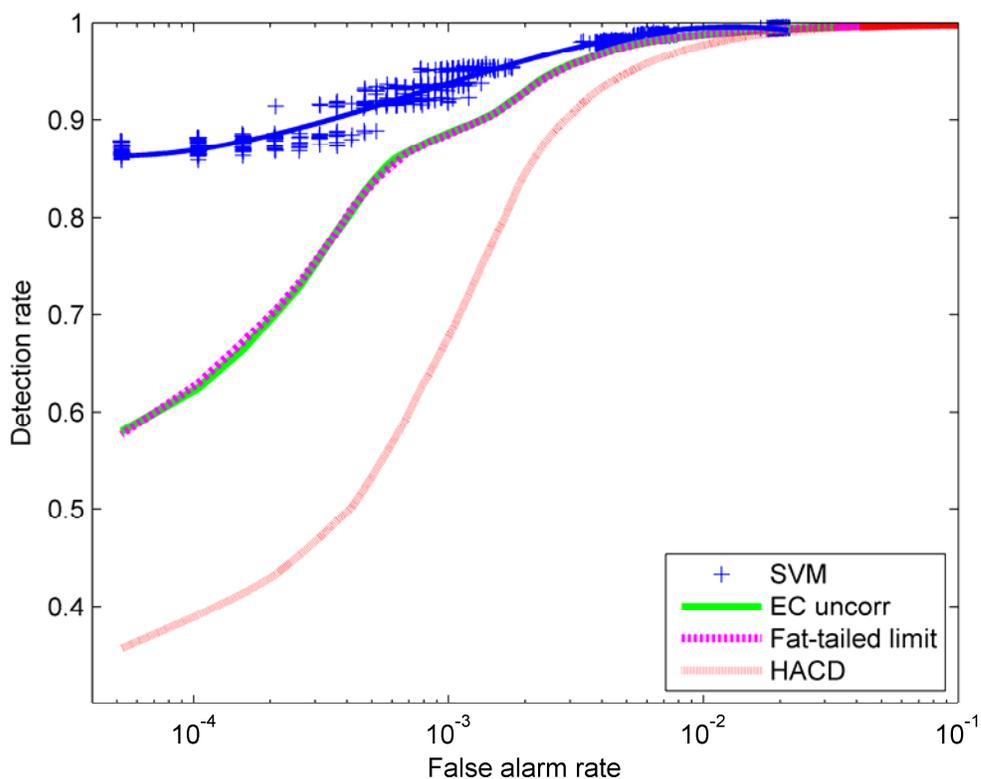


Figure 7. Same results as Fig. 6, except that the mean performance curves are shown for the three linear detectors. A smooth curve is drawn through the SVM points as a guide.

4.1. Performance on VNIR data

Hyperspectral imaging in the VNIR spectral region (400 – 900 nm) is of interest for many remote sensing applications, but is influenced by a different set of physical phenomena and provides another useful regime for comparison with the LWIR. In the VNIR, temperature across the scene does not affect the observation, but changing solar illumination characteristics (e.g. angle, cloudiness, shadowing) have a much more direct influence and can be the main source of pervasive difference. A useful dataset for evaluation of change detection algorithms has been published previously by Eismann et al.¹⁷ We selected one HSI dataset acquired in August, and a second in October as the image pair for analysis. Distinct seasonal changes occurred naturally from the first to the second scene. The 2-D scatter plot for this image pair is shown in Fig. 7. The images are 800 X 1024 pixels, and were originally acquired with 140 spectral channels. As described in Ref. 7, we used Canonical Correlation Analysis (CCA)¹⁸ to reduce the dimension of the data, so the values in Fig. 7 were obtained from the resultant 10-dimensional data. Additional

instrumental details and results of change detection algorithms are discussed in Refs. 3, 7, 17.

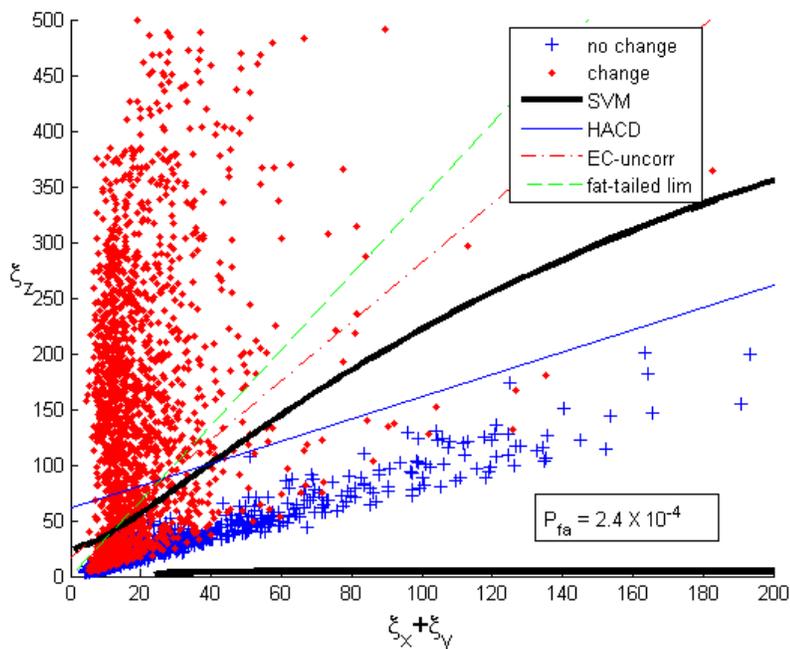


Figure 8. Scatter plot for VNIR image pair. Change detector boundaries are plotted as lines and a curve.

The SVM algorithm used on the VNIR data was identical to that used for the LWIR data. Regularization and kernel parameters were $C = 5$ and $\gamma = 10$. (The number of samples was somewhat larger, so a randomly selected subset was used for computational speedup.) We believe that the insensitivity of performance to parameter values is a consequence of working in the reduced 2-D space. It indicates that parameter selection is not expected to be a burdensome task. Class weights were adjusted incrementally to yield a series of SVM decision boundaries with increasing false alarm rate. Performance of the various algorithms, based on simulated scene changes, is shown in Fig. 8. Overall, the results are similar to the LWIR example. SVM, carried out in the logarithmic version of the 2-D space, outperforms the other three algorithms. One caveat is that there is some variability in performance as different train/test splits of the data are used. On occasion, the SVM will perform about the same as the next leading algorithm (fat-tailed limit, in this case). But by using repeated splits of the data, improved performance is seen in an average sense.

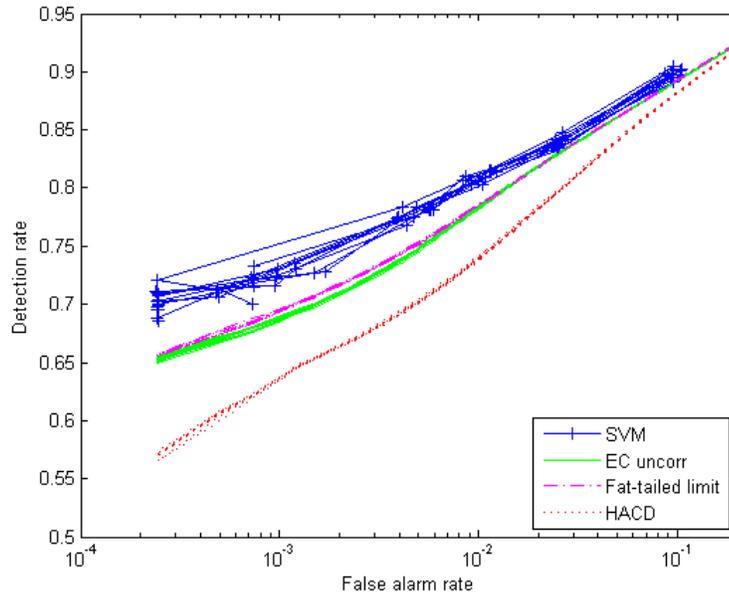


Figure 9. Performance curves for VNIR data of Ref. 17.

5.0 Summary

The approach to change detection that we propose in this paper contains two important conceptual elements: a novel dimension reduction of the HSI data, and a machine learning step that shapes the decision boundary to the data of interest. Dimension reduction is not new in HSI exploitation: PCA and CCA are commonly used in pursuit of various goals. The particular dimension reduction that we employ, however, uses coordinates that are themselves associated with change detection quantities. Implementing the SVM in 2-D space is convenient because the optimization is efficient and the results can be directly visualized. The latter is helpful in checking for artifacts of various types and in preventing overfitting of the data, which can result from inappropriate SVM parameters. We have shown that the same SVM parameters give good results for HSI data in two different regimes (VNIR and LWIR). The proposed algorithm outperforms others by reasonable margins. This two-step approach to this important problem in HSI exploitation is similar in spirit to one proposed for target detection, where the spectral signature of interest is known ahead of time and it is being sought in a single datacube.¹⁹

5.0 Acknowledgements

This work was supported by the Los Alamos Laboratory Directed Research and Development program. We thank Joseph Meola and Michael Eismann of the Air Force Research Laboratory for providing the VNIR hyperspectral data.

6.0 References

- ¹ R.J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, "Image change detection algorithms: a systematic survey," *IEEE Trans. Image Processing*, **14**, pp. 294-307, 2005.
- ² A. Schaum and A. Stocker, "Long-interval chronochrome target detection," *Proc. 1997 Intl. Symp. Spectral Sensing Research*, 1998.
- ³ J. Meola and M. T. Eismann, "Image misregistration effects on hyperspectral change detection," *Proc. SPIE* **6966**, p. 69660Y, 2008.
- ⁴ J. Theiler and S. Perkins, "Proposed Framework for anomalous change detection," *ICML Workshop on Machine Learning Algorithms for Surveillance and Event Detection*, pp. 7-14, 2006.
- ⁵ S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory, Vol. II*, Prentice Hall, New Jersey, 1998.
- ⁶ J. Theiler and C. Scovel, "Uncorrelated versus Independent Elliptically-Contoured Distributions for Anomalous Change Detection in Hyperspectral Imagery," *Proc. SPIE* **7426**, p.72460T, 2009.
- ⁷ J. Theiler, C. Scovel, B. Wohlberg, and B.R. Foy, "Elliptically-contoured distributions for anomalous change detection in hyperspectral imagery," *Geosci. Remote Sensing Lett.* 2010, in press.
- ⁸ 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 Proc.*, **38**, pp. 1760-1770, 1990.
- ⁹ D. B. Marden and D. Manolakis, "Modeling hyperspectral imaging data," *Proc. SPIE* **5093**, pp. 253-262, 2003.
- ¹⁰ A. Schaum, E. Allman, J. Kershenstein, and D. Alexa, "Hyperspectral change detection in high clutter using elliptically contoured distributions," *Proc. SPIE* **6565**, p. 656515, 2007.
- ¹¹ B.R. Foy, R.R. Petrin, C.R. Quick, T. Shimada, and J.J. Tiee, "Comparisons Between Hyperspectral Passive and Multispectral Active Sensor Measurements," *SPIE Proc.* **4722**, 98-109 (2002).
- ¹² B.D. McVey and K.L. Mitchell, LANL, private communication.
- ¹³ T. Schneider, "Analysis of incomplete climate data: estimation of mean values and covariance matrices and imputation of missing values," *J. Climate* **14**, pp. 853-71, 2001.
- ¹⁴ J. Theiler, "Quantitative comparison of quadratic covariance-based anomalous change detectors," *Applied Optics* **47** (2008) F12-F26.
- ¹⁵ B. Schölkopf and A. J. Smola, *Learning with Kernels* (MIT Press, Cambridge, MA, 2002).
- ¹⁶ C.-C. Chang and C.-J. Lin, *LIBSVM: a library for support vector machines* (2001). Software available at www.csie.ntu.edu.tw/~cjlin/libsvm.
- ¹⁷ M. T. Eismann, J. Meola, and R. Hardie, "Hyperspectral change detection in the presence of diurnal and seasonal variations," *IEEE Trans. Geoscience and Remote Sensing*, vol. 46, pp. 237-249, 2008.
- ¹⁸ A. A. Nielsen, K. Conradsen, and J. J. Simpson, "Multivariate alteration detection (MAD) and MAF post-processing in multispectral bi-temporal image data: new approaches to change detection studies," *Remote Sensing of the Environment*, vol. 64, pp. 1-19, 1998.
- ¹⁹ B.R. Foy, J. Theiler, and A.M. Fraser, "Decision Boundaries in Two Dimensions for Target Detection in Hyperspectral Imagery," *Opt. Express* **17**, 17391-17411 (2009).