

The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology

<http://dms.sagepub.com/>

Towards the mitigation of correlation effects in anomaly detection for hyperspectral imagery

Jason P. Williams, Trevor J. Bihl and Kenneth W. Bauer

The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology published online 21 February 2013
DOI: 10.1177/1548512912473679

The online version of this article can be found at:

<http://dms.sagepub.com/content/early/2013/01/25/1548512912473679>

Published by:



<http://www.sagepublications.com>

On behalf of:



The Society for Modeling and Simulation International

Additional services and information for *The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology* can be found at:

Email Alerts: <http://dms.sagepub.com/cgi/alerts>

Subscriptions: <http://dms.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

>> [OnlineFirst Version of Record - Feb 21, 2013](#)

[What is This?](#)

Towards the mitigation of correlation effects in anomaly detection for hyperspectral imagery

Journal of Defense Modeling and Simulation: Applications, Methodology, Technology
I–II
© 2013 The Society for Modeling and Simulation International
DOI: 10.1177/1548512912473679
dms.sagepub.com
SAGE

Jason P Williams, Trevor J Bihl and Kenneth W Bauer

Abstract

This research involves simulating remote sensing conditions using previously collected hyperspectral imagery (HSI) data. The Reed–Xiaoli (RX) anomaly detector is well-known for its unsupervised ability to detect anomalies in hyperspectral images. However, the RX detector assumes uncorrelated and homogeneous data, both of which are not inherent in HSI data. To address this difficulty, we propose a new method termed linear RX (LRX). Whereas RX places a test pixel at the center of a moving window, LRx employs a line of pixels above and below the test pixel. In this paper, we contrast the performance of LRx, a variant of LRx called iterative linear RX (ILRx), the recently introduced iterative RX (IRx) algorithm, and the support vector data description (SVDD) algorithm, a promising new HSI anomaly detector. Through experimentation, the line of pixels used by ILRx shows an advantage over RX and IRx in that it appears to mitigate the deleterious effects of correlation due to the spatial proximity of the pixels; while the iterative adaptation taken from IRx simultaneously eliminates outliers allowing ILRx an advantage over LRx. Such innovations to the basic RX algorithm allow for the reduction of bias and error in the estimation of the mean vector and covariance matrix, thus accounting for a portion of the spatial correlation inherent in HSI data.

Keywords

Anomaly detection, hyperspectral imagery, iterative linear RX, normalized difference vegetation index, RX detector, support vector data description, remote sensing

I. Introduction

Remote sensing involves studying a given object without initiating physical contact;^{1,2} of particular interest are passive remote sensing systems which rely on natural sources of illumination. Hyperspectral imagery (HSI) systems are passive systems which collect spectrally contiguous data across a large swath of the electromagnetic spectrum, permitting material identification through fine spectral sampling. One of the fundamental problems faced by practitioners in this area is analyzing the highly correlated data streams that are output from these models.³ Computer models, such as discrete-event simulations, are used to aid in understanding real-world processes. Simulation analysts must deal with temporal correlation. In this paper, we are concerned with highly correlated data of both a spatial and spectral nature. Specifically, we will address the spatial correlation problem.

Typically, HSI encompasses the visible to infrared regions of the spectrum, containing anywhere from more than 20 to 250 plus spectral bands, whereas standard digital cameras capture three coarsely sampled bands: red, green, and blue. The vast amount of data contained in HSI affords a great opportunity to detect anomalies in an image using standard multivariate statistical techniques, as each material reflects individual wavelengths of the spectrum

Air Force Institute of Technology, Wright-Patterson Air Force Base (WPAFB), OH, USA

Corresponding author:

Jason P Williams, Air Force Institute of Technology, 2950 Hobson Way, Wright-Patterson Air Force Base, OH 45433, USA.
Email: jason.williams.2@ramstein.af.mil

differently. However, the large amount of data contained within each image often requires dimensionality reduction/feature selection techniques to be employed such that analysis algorithms operate on lower dimensional, uncorrelated data as described by Landgrebe.⁴

Anomaly detection refers to the location of spectral data that does not belong within a given set. It can be used in numerous applications such as financial fraud detection, computer security, and military surveillance.⁵ In HSI applications, anomaly detection is used to find objects of interest within the image by locating pixels statistically different from the non-anomaly pixels, referred to as the background. Three broad categories of anomaly detection methods exist:⁵ supervised, semi-supervised, and unsupervised detection. Supervised detection requires a set of training data that includes both the background and anomaly data prior to analysis. Semi-supervised detection also requires a training set; however, it only requires background data. Differences between images, e.g. the desert and forest images in this paper, present a problem that effects supervised or semi-supervised methods when applied to HSI. Therefore, it is difficult to train a detector on one image and test it against another. The standard work around for semi-supervised detection is to select a random set of data. This practice is successful because the set of anomalies in the data set is assumed to be sparse; hence, the random selection should provide a representative sample of the true background. Unsupervised detection does not require a training set and is therefore more appropriate when analyzing HSI data.

The literature on anomaly detection in HSI has increased following the publication of Reed and Yu's paper on the Reed–Xiaoli (RX) detector in 1990,⁶ to include various articles with modifications or additions to the RX detector,^{1,7–10} classification and discrimination methods,^{1,11,12} different fusion techniques,^{13,14} and overview articles^{15–17} of detection algorithms, including RX. Related work includes a number of additional detectors, such as: support vector data description (SVDD)^{18–21}, multiple window detectors,^{8,22,23} and various mixture models.^{1,17,24,25} More recently, work has been conducted using synthetically generated or simulated data to supplement the low number of hyperspectral images with available truth masks that are typically accessible to researchers.^{7, 26–28}

In practice, the RX method, when applied to hyperspectral data, is likely to have a high false positive detection percentage because the underlying statistics assume the data being analyzed follows a Gaussian distribution. However, Banerjee et al.²⁰ showed that HSI is not often unimodal. Further, to compound the non-Gaussian difficulty, an image, by its very nature, is correlated and heterogeneous. However, RX is still used in practice because

it offers fast processing times, is intuitively easy to understand, and is algorithmically simple.

The purpose of this paper is to present modifications to the standard RX algorithm. A new method, called linear RX (LRX), has the ability to overcome some of the correlation problems hindering RX⁶ and iterative RX (IRX).¹⁰ This paper contrasts the performance of LRX and, another new method, its variant iterative linear RX (ILRX), to RX and IRX. In addition, to further test the benefit of the new algorithms, both algorithms are tested against the global SVDD algorithm, a promising new supervised HSI detector.^{20, 21}

The remainder of this paper is organized as follows: Section 2 presents a description of the algorithms contrasted in this research, Section 3 details the methodology used to compare the five anomaly detectors, Section 4 provides the experimental results, and in Section 5, the paper is concluded.

2. Algorithms

This section of the paper describes how each of the five anomaly detection algorithms are contrasted and implemented, and explains the use of the normalized difference vegetation index (NDVI) in post-processing to realize improved results from the detectors. Due to the large amount of data contained within a given hyperspectral image, it is standard practice, prior to applying an anomaly detector, to reduce the dimensionality of the image by running principal component analysis (PCA)²⁹ on the whole data set, retaining the p largest principal components (PCs).

2.1. RX detector

The RX detector, introduced by Reed and Yu,⁶ detects anomalies utilizing a moving window approach, where the pixel in the center is scored by comparison to the remaining pixels in the window. The window is shifted, one pixel at a time, across a row of pixels with the new center pixel being scored at each step, as displayed in Figure 1(a), where the small square in the center of the box represents the test pixel and the box around the test pixel represents the pixels compared with the test pixel to generate an RX score. This process is continued until all possible pixels have been analyzed. Each test pixel, x , is given a score based upon a generalized likelihood ratio test which simplifies to equation (1) if the pixels within the test window are assumed to be normally distributed with mean vector of the background pixels, μ , and covariance matrix Σ . It should also be noted that as the number of pixels in the window approaches to infinity, the RX score becomes the

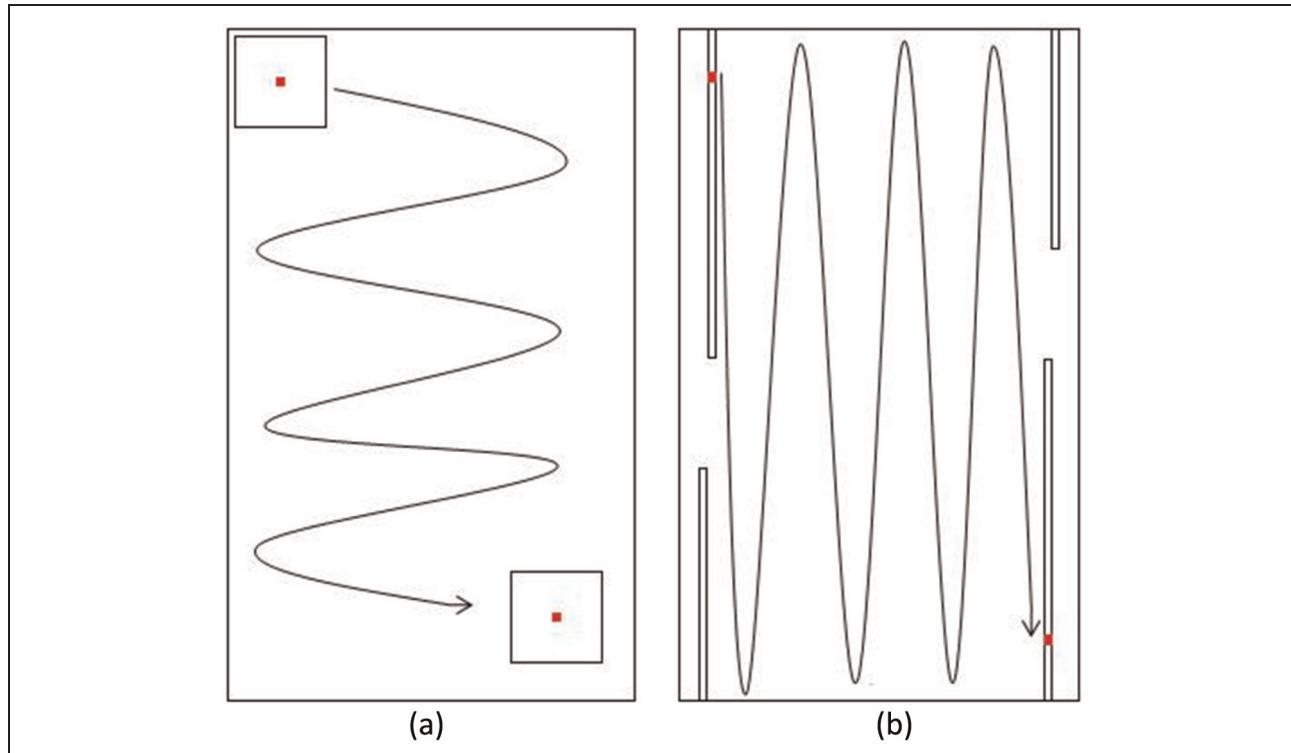


Figure 1. (a) RX window vs. (b) LRX line.

squared Mahalanobis distance between the test pixel and the mean vector of the background pixels,

$$RX(x) = (x - \mu)^T \left[\left(\frac{N}{N+1} \right) \sum + \left(\frac{1}{N+1} \right) (x - \mu)(x - \mu)^T \right]^{-1} (x - \mu) \quad (1)$$

Pixels with an *RX* score greater than $\chi^2_{\alpha, (N-1)}$, where α represents the corresponding significance level of the chi-squared distribution, are labeled anomalous by the RX detector.

2.2. IRX detector

The IRX detector¹⁰ is an extension of the standard RX detector; IRX extends the RX detector through an iterative process, where each iteration sees IRX calculating an improved estimate of the mean vector and covariance matrix of the background pixels.

The IRX algorithm is processed using the following steps:

1. Each iteration begins by running the standard RX algorithm to calculate an RX score for each

testable pixel in the image; however, to improve accuracy, pixels selected as anomalies in the previous iteration are excluded from the data used to estimate the mean vector and covariance matrix of the background. Note: At the start of the algorithm the set of anomalies is empty.

2. Using the RX scores calculated in step (1), a pixel is considered anomalous if its RX score is greater than $\chi^2_{\alpha, (N-1)}$. This ends a given iteration, allowing for pixels to enter and exit the set of anomalies.
3. The algorithm ends if the set of anomalies determined in step (2) is identical to the set of anomalies from the previous iteration or the maximum number of iterations has been reached. Otherwise, the algorithm iterates again from step (1).

2.3. LRX and ILRX detectors

LRX and ILRX are similar to RX and IRX, respectively; however, instead of a window being moved through the image, they employ a vertical line of pixels above and below the test pixel. If the number of pixels above or below the test pixel exceeds the height of the image, the required pixels are taken from the bottom of the previous column or from the top of the following column (Figure 1(b)). The line is used to increase the average

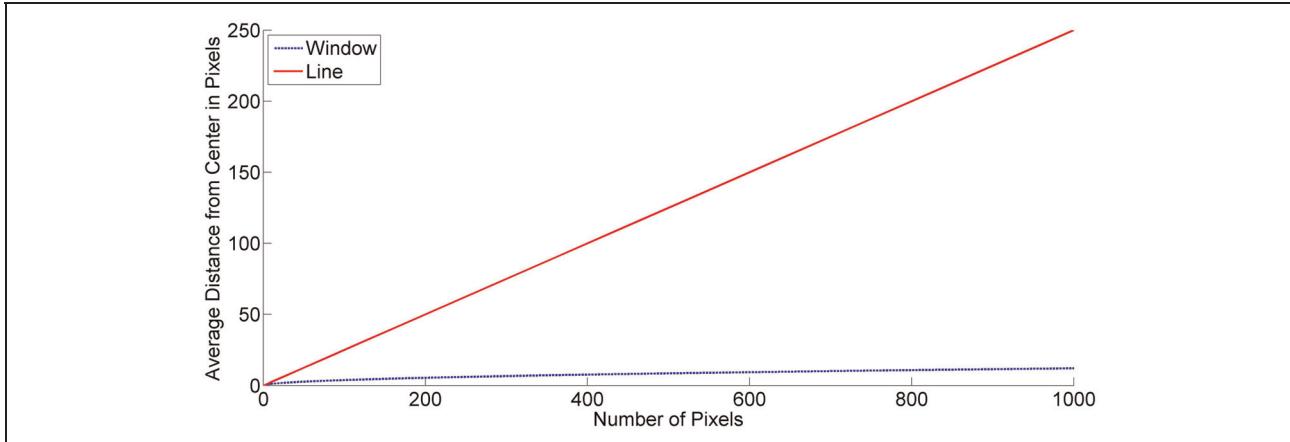


Figure 2. Average distance between pixels.

distance between the pixels used to estimate the mean vector and the covariance matrix. This can be seen in Figure 2, which shows the average distance between pixels using a window and a line. Increasing the average distance between pixels mitigates the deleterious effects of correlation due to the spatial proximity of the pixels. Such a step allows for the reduction of bias and error in the estimation of the mean vector and the covariance matrix. A possible concern for such an approach might be that the reduction in the contribution to the bias due to spatial correlation may be offset by the contribution to the bias due to image non-stationarity. This issue is discussed later in the paper, and as we demonstrate below, this was not a concern for the images we tested.

2.4. SVDD

Banerjee et al.^{20,21} extended the SVDD algorithm by Tax and Duin^{18,19} into an HSI anomaly detector. SVDD is a one-class classifier, where points are considered in class or out of class, where the support of the distribution is considered as the minimally enclosing hypersphere in the feature space. In operation, SVDD takes a training set, $T = \{x_i, i = 1, \dots, M\}$, of M background pixels, where x is randomly selected from the image as the training data. SVDD then attempts to determine the minimum volume hypersphere, $S = \{x : \|x - a\|^2 < R^2\}$, as the L^2 norm or Euclidean distance, with radius R and center a that contains the set of M randomly chosen pixels. This is obtained by solving the following minimization problem

$$\min(R) \text{ subject to } x_i \in S, \quad i = 1, \dots, M \quad (2)$$

The radius R and center a of the hypersphere are determined by optimizing the Lagrangian, L , with respect to the weights, or support vectors, α_i

$$L(R, a, \alpha_i) = R^2 - \sum_i \alpha_i \{R^2 - (\langle x_i, x_i \rangle - 2\langle a, x_i \rangle + \langle a, a \rangle)\} \quad (3)$$

where $\langle \cdot \rangle$ represents the dot product of the operation of the two vectors.

After optimizing, the kernel trick, which transforms data to a different dimensional space for simpler computations without ever explicitly calculating the mapping, can be applied which leads to the SVDD statistic for a given pixel y

$$SVDD(y) = R^2 - K(y, y) + 2 \sum_i \alpha_i K(y, x_i) \quad (4)$$

where $K(x, y)$ is the kernel mapping defined by

$$K(x, y) = \exp(-\|x - y\|^2 / \sigma^2) \quad (5)$$

and variable σ^2 is a radial basis function parameter used as a scaling factor to determine the size of the hypersphere, hence adjusting how well the SVDD algorithm generalizes the incoming data.

When applied to HSI, the SVDD algorithm is processed in the following steps:²⁰

1. Randomly select M pixels from the image.
2. Estimate an optimal value for σ^2 by determining a value that will minimize the false positive rate or the number of background pixels classified as targets.
3. Estimate the parameters (R, a, α_i) needed to model the hypersphere.
4. Determine whether $SVDD(y) \geq t$ for some user-defined threshold t for every pixel in the image.

The SVDD algorithm is considered in this research because it is a novel and promising state of the art detector, and as a semi-supervised method, it allows for an interesting performance contrast relative to the other unsupervised methods.

2.5. NDVI

It is fairly common to get false positives, i.e. pixels flagged as anomalies that are truly background pixels, when attempting to find anomalies, generally man-made objects, in HSI using one of the previously described methods. One relatively simple way to reduce false positives is to implement some form of pre-processing or post-processing. Since we are attempting to locate anomalies without prior knowledge, one applicable post-processing method is applying NDVI, as introduced by Rouse et al.,³⁰ to remove pixels that are likely to be vegetation.

NDVI gauges whether or not a given pixel is green vegetation by using the absorptive cutoff of chlorophyll between the visible and near infrared spectrum. It does this by comparing the intensity of the visible bands to the intensity of the near infrared bands, since the reflectance in the near infrared bands is considerably larger for vegetation. The measure is given by

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (6)$$

where *NIR* equals the radiance value of the near infrared spectral band and *Red* is the radiance value of the red spectral band.^{1,2,30,31} Prior to locating the anomalies, an NDVI threshold for the image and an NDVI score for each pixel is calculated. The NDVI threshold was determined by plotting the NDVI scores for all of the images and setting a threshold. Subsequently, all pixels with an NDVI score above the threshold are classified as vegetation. Once the anomaly detector has been run, regardless of the indications, all declared vegetation pixels are classified as background. Since the desert images display NDVI scores that are, for the most part, below the selected threshold, very few pixels will be classified as vegetation; hence, very few potential false positives are deleted.

3. Methodology

The five anomaly detectors were compared using six images from the forest radiance I and desert radiance II collection events, from the hyperspectral digital imagery collection equipment (HYDICE) push-broom, aircraft mounted sensor. The HYDICE sensor collects spectral data in 210 bands between 397 nm and 2500 nm, including visible and infrared data. Due to atmospheric absorption effects, only 145 bands were used in the analysis of the

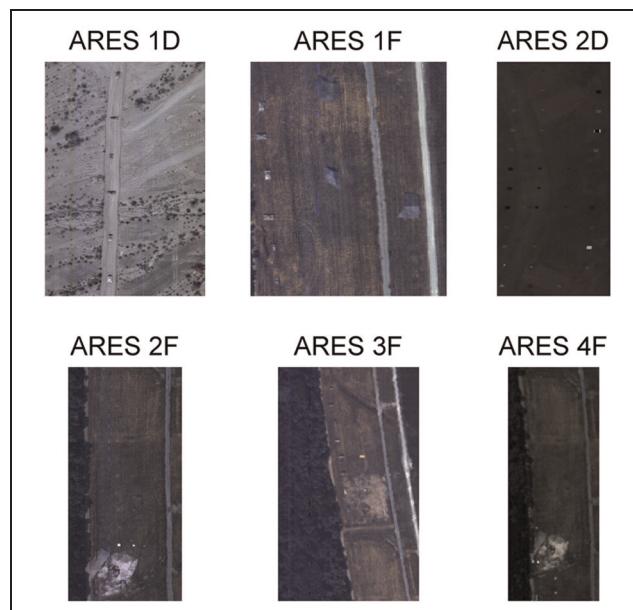


Figure 3. HYDICE images.

Table 1. HYDICE image data.

| Image | Size | Total Pixels | Anomalies | Anomalous Pixels |
|---------|-----------|--------------|-----------|------------------|
| ARES 1D | 291 x 199 | 57,909 | 6 | 235 |
| ARES 1F | 191 x 160 | 30,560 | 10 | 1007 |
| ARES 2D | 215 x 104 | 22,360 | 46 | 523 |
| ARES 2F | 312 x 152 | 47,424 | 30 | 307 |
| ARES 3F | 226 x 136 | 30,736 | 20 | 145 |
| ARES 4F | 205 x 80 | 16,400 | 29 | 109 |

images. A description of each image is shown in Table 1, and the images are displayed in Figure 3.

Due to the small number of images available and the need to train and validate each of the algorithms, all six images, presented in Table 1 and Figure 3, were divided in half to create a top and bottom portion of the image. Then, a top or bottom of each image, chosen randomly, was selected for the training set of images and the other half was used in the validation set. The training set included the top of ARES 2D and ARES 4F and the bottom of ARES 1D, ARES 1F, ARES 2F, and ARES 3F. It may be a stretch to try and draw too much from a comparison of five algorithms and only six images; however, we believe our experiments point to the clear potential of the new technique. Furthermore, while splitting the images in half to double our data may not be the best method for creating training and test sets, it is certainly better than using the same images for training and test. We acknowledge that the image halves are spectrally correlated due to shared

Table 2. Algorithm parameter settings.

| Algorithm | Number of PCs | Number of Pixels | Number of Iterations | σ^2 | NDVI | Total Data Points Collected (Per Image) |
|-----------|---------------|---------------------------|----------------------|-----------------|--------|---|
| RX | 3, 4,..., 10 | $17^2, 19^2, \dots, 25^2$ | 1 | — | Yes/No | 80 |
| IRX | 3, 4,..., 10 | $17^2, 19^2, \dots, 25^2$ | 10, 20,..., 50 | — | Yes/No | 400 |
| LRX | 3, 4,..., 10 | $0.5*H, 1*H, 1.5*H, 2*H$ | 1 | — | Yes/No | 64 |
| ILRX | 3, 4,..., 10 | $0.5*H, 1*H, 1.5*H, 2*H$ | 10, 20,..., 50 | — | Yes/No | 320 |
| SVDD | 3, 4,..., 10 | $172, 192, \dots, 252$ | 1 | 10, 20,..., 300 | Yes/No | 2400 |

H represents the height of the image in pixels.

Table 3. Training data results (TPF at FPF = 0.1).

| Algorithm | ARES 1D | ARES 1F | ARES 2D | ARES 2F | ARES 3F | ARES 4F | Average |
|-----------|---------|---------|---------|---------|---------|---------|---------|
| RX | 0.8673 | 0.3410 | 0.9933 | 0.9455 | 0.9535 | 0.8649 | 0.8276 |
| IRX | 1.0000 | 0.4615 | 1.0000 | 1.0000 | 0.9744 | 0.9444 | 0.8967 |
| LRX | 0.9118 | 0.7916 | 0.9474 | 0.7632 | 0.8308 | 0.7990 | 0.8406 |
| ILRX | 1.0000 | 1.0000 | 1.0000 | 0.9912 | 1.0000 | 0.9500 | 0.9902 |
| SVDD | 0.9558 | 0.9588 | 0.9880 | 0.9386 | 0.9846 | 0.8750 | 0.9501 |

weather, viewing conditions, etc. However, correlation in the spatial domain appears to be minimal.

Each of the algorithms was tested across a large combination of parameter settings in order to find the optimal settings for the algorithm. The parameters were the number of PCs to retain, the number of pixels to use in the window/line using RX methods or the size of the training set for SVDD, the number of iterations to use for the iterative methods, whether NDVI was used in post-processing, and the parameter value, σ for SVDD. This information is summarized in Table 2. The last column is displayed to show how many combinations of parameters were collected for each method on a single image.

The algorithms' anomaly detection performance on the selected test set was compared through the use of receiver operating characteristic (ROC) curves.³² Since the RX algorithms' test statistics are based upon the chi-squared distribution, the significance level α was varied to serve as the threshold for the ROC curves. Similarly, the user-defined anomaly threshold t was varied in the SVDD algorithm to generate ROC curves. Owing to the fact that a large number of settings for each algorithm were examined, visual inspection of the ROC curves was not feasible. Therefore, the individually tested setting combinations for each algorithm were scored using the Neyman–Pearson technique.³³ Specifically, the true positive fraction (TPF) for the anomalous pixels detected in each of the six images was averaged when the corresponding false positive fraction (FPF) is equal to 0.1. A FPF of 0.1 was chosen because it was deemed that if the FPF exceeded 0.1, the algorithm would no longer be of any practical use due to over-saturation of misclassified data.

After the setting combination with the highest average TPF at a FPF = 0.1 was determined for each anomaly detector, its performance was validated by taking the best settings for each individual algorithm and running them on the six validation images.

An artifact of the RX and IRX methods, as described by Reed and Yu⁶ and Taitano et al.,¹⁰ is an area of pixels that form a border around the image which cannot be tested due to the requirement of the window. Methods to allow the algorithms to test the border pixels can be implemented, such as using only the part of the window that is within the image or moving the test pixel from the center of the window when it is against the border of the image. However, in this research, the RX and IRX algorithms as originally designed were compared and the border pixels that could not be tested were not considered in the performance evaluation.

4. Results

Relative to the training data, the results for the best settings of each algorithm by image and overall average are shown in Table 3. It can be seen that LRx achieves equivalent performance to RX in most images; ARES 1F is an exception where the spatially large objects appear to confound the RX algorithm, yet are detected by LRx. With iterations, ILRX was the best performing algorithm or tied with IRX in all cases, except ARES 2F.

The corresponding best tested setting for each of the algorithms is displayed in Table 4. “Yes” or “No” in the NDVI column for SVDD implies the algorithm achieved the same results whether or not NDVI was used

Table 4. Best tested parameter setting from training data.

| Algorithm | Number of PCs | Number of Pixels | Number of Iterations | σ^2 | NDVI |
|-----------|---------------|------------------|----------------------|------------|-----------|
| RX | 9 | 232 | 1 | — | Yes |
| IRX | 9 | 252 | 20 | — | Yes |
| LRX | 9 | 1*H | 1 | — | Yes |
| ILRX | 10 | 2*H | 30 | — | Yes |
| SVDD | 10 | 252 | 1 | 60 | Yes or No |

H represents the height of the image in pixels.

Table 5. Validation data results (TPF at FPF = 0.1).

| Algorithm | ARES 1D | ARES 1F | ARES 2D | ARES 2F | ARES 3F | ARES 4F | Average |
|--------------|---------|---------|---------|---------|---------|---------|---------|
| RX | 0.9016 | 0.2075 | 0.9920 | 0.8282 | 0.9545 | 0.8864 | 0.7950 |
| IRX | 1.0000 | 0.3186 | 1.0000 | 0.9495 | 1.0000 | 1.0000 | 0.8780 |
| LRX | 0.9645 | 0.4902 | 0.9890 | 0.8342 | 0.9104 | 0.7669 | 0.8259 |
| ILRX (Best) | 1.0000 | 0.9449 | 1.0000 | 0.9741 | 1.0000 | 1.0000 | 0.9865 |
| ILRX (Worst) | 1.0000 | 0.7394 | 1.0000 | 0.9646 | 1.0000 | 0.9744 | 0.9464 |
| SVDD | 0.9180 | 0.9850 | 0.9983 | 0.8641 | 0.9330 | 0.8200 | 0.9197 |

in post-processing. It should be noted that ILRX was the most robust of the algorithms tested. During training, eleven different parameter settings realized the same values shown in Table 3. No other algorithm had multiple parameter settings that obtained the optimal results. Since multiple setting combinations were found for ILRX, they were all tested on the validation images.

The results from the validation images are displayed in Table 5, to include the best and worst tested parameter settings of the eleven training combinations validated for ILRX. It can be seen that ILRX is still the top performer overall regardless of whether the best or worst training settings were implemented. Furthermore, the ILRX algorithm received the smallest drop in average TPF when the settings were tested on the validation images, as compared to the training images.

Figure 4 shows the ROC curves for each of the six validation images comparing TPF to FPF using the best tested settings for each algorithm from the training images, as displayed in Table 4. In every case, IRX performs better than RX, and ILRX performs better than LRX; hence, the comments below focus on IRX, ILRX, and SVDD.

IRX did well on all of the images except when there are large anomalies, such as the ones highlighted in ARES 1F. This is because the window, as it moves through a large anomaly, becomes dominated by the local anomalous pixels rather than the general background of the image. This defeats the purpose of the window, which is to give a good estimate of the true background of the image. As a result, the pixel being analyzed appears similar to the other pixels in the window and is not classified as an anomaly. ILRX mitigates this problem through its use of a vertical line

which only contains a small portion of even a large anomaly and considerably more background pixels.

ILRX had the highest performance or was on par with the other detectors in all of the images. It has slight problems with the rock formations in ARES 2F and 4F that IRX does not detect due to the window effect of large images; however, this is difficult to discern from the ROC curve due to ILRX detecting most of the anomalies at a relatively low threshold.

SVDD consistently performed better than both of the non-iterative methods; however, it was inconsistent with regard to its performance against the iterative methods. Also, the fact that it is a semi-supervised method that randomly selects training data can lead to less than optimal performance from the detector. The only image where SVDD outperformed the other algorithms is ARES 1F, where IRX has trouble with large anomalies, and ILRX has difficulty with vertical roads.

Figure 5 shows a representation of the image and the pixels classified as anomalies, or anomalous pixel maps, for IRX, ILRX, and SVDD on the validation images ARES 1F and 4F; note that the aspect ratio of the IRX mask is smaller because it was not used to test the borders of the image. The anomalous pixel maps were generated at the first knee in the ROC curve so that they were not overwhelmed by false positive pixels. The corresponding TPF and FPF are displayed below each of the images. It can be easily seen in ARES 1F that SVDD is realizing superior results, primarily because IRX is not locating the large anomalies, and ILRX in addition to finding almost all of the anomalous pixels is having some difficulties with the roads. In ARES 4F, both of the RX methods are giving

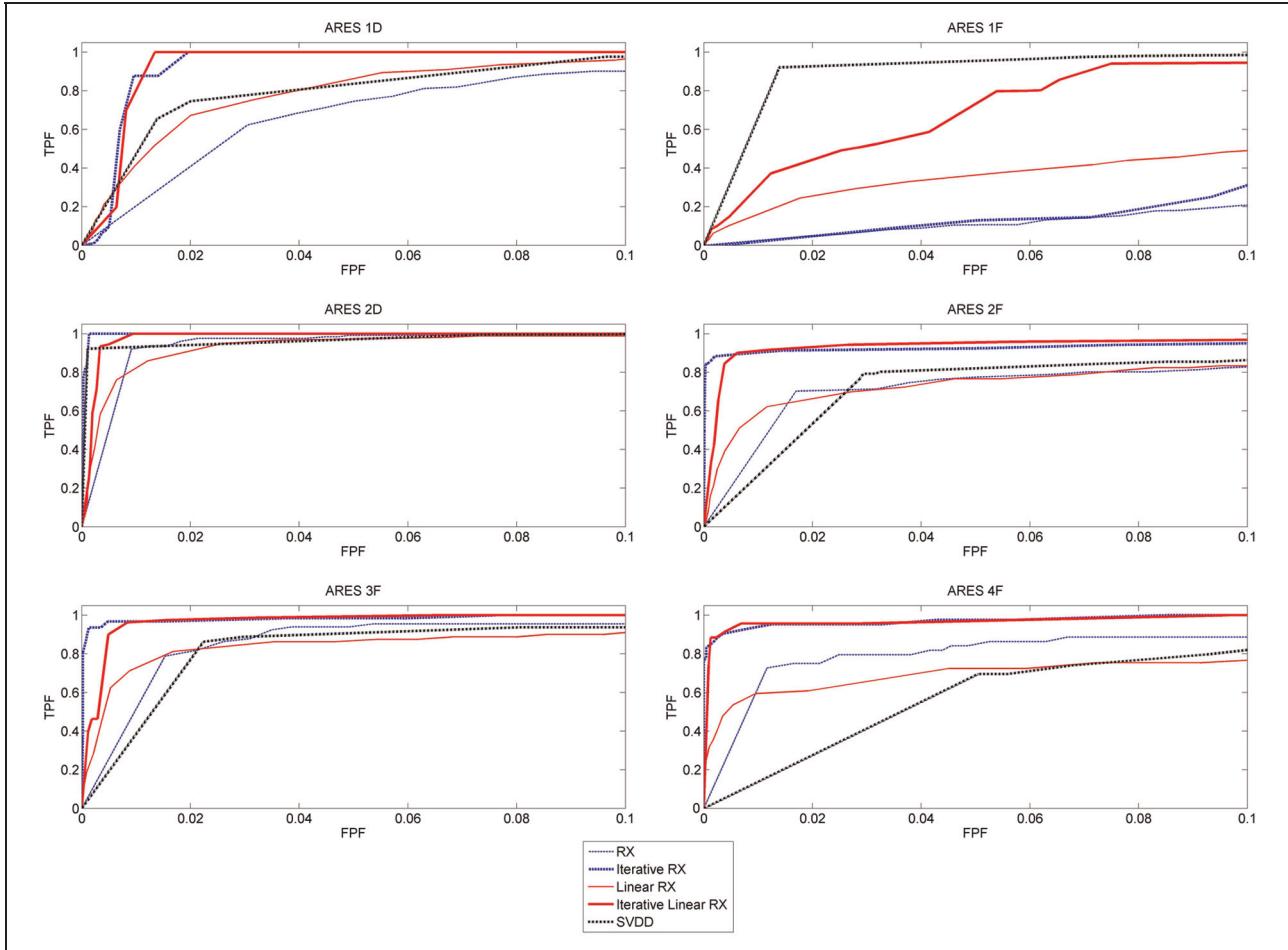


Figure 4. ROC curves of best tested parameter settings on validation images.

high-quality results and the SVDD algorithm is getting inundated by the large rock formation.

The embellishments to RX follow a reasoned pattern. IRX allows for the exclusion of outliers in the local mean vector and covariance matrix calculations, thereby promoting a more accurate assessment of the target pixel.¹⁰ LRX mitigates the correlation difficulties related to RX by establishing the mean vector and covariance matrix that is, on average, further from each other than the standard RX window. The possible concern that the reduction in the contribution to the bias due to spatial correlation may be offset by the contribution to the bias due to image non-stationarity was not realized here. We believe this is due to the following factors. If one considers the non-stationarity in the image as being characterized by distinct pixel clusters then the variation between these clusters appears to be significantly less than the variation between the background pixels, in general, and the target pixels. Furthermore, the running covariance matrix estimate

calculated across the background pixels appears to be fairly robust to the heterogeneity as evident by the algorithms performance. The notion of using separate estimates from the individual clusters is the subject of current research. Finally, ILRX exploits the innovations of both IRX and LRX. Taken together, these innovations make ILRX a very competitive algorithm.

5. Conclusions

This paper presented LRX and ILRX updates to the newly introduced IRX algorithm. Through experimentation, the line of pixels used by ILRX shows an advantage over RX and IRX in that it can help mitigate the deleterious effects of correlation due to the spatial proximity of the pixels while the iterative adaptation taken from IRX simultaneously eliminates outliers. Such steps allow for the reduction of bias and error in the estimation of the mean vector and covariance matrix, thus accounting for a

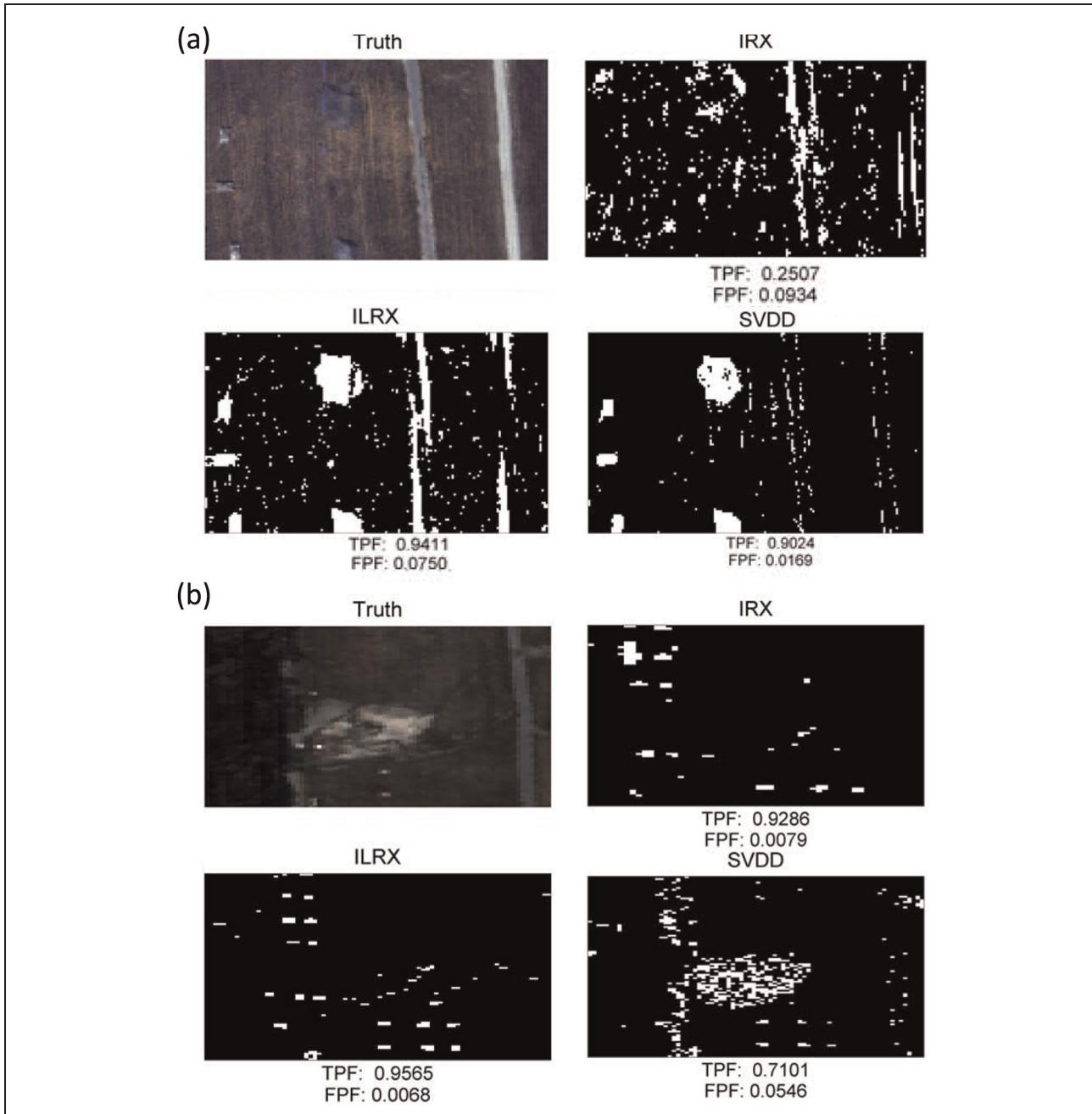


Figure 5. Anomalous pixels maps: (a) ARES IF and (b) ARES 4F.

portion of the spatial correlation inherent in HSI data. Using the HYDICE images, ILRX has been shown to be a very promising unsupervised anomaly detection algorithm.

Funding

This work was supported by the Secretary of the Air Force.

References

1. Eismann MT. *Hyperspectral remote sensing*. Bellingham, WA, USA: SPIE Press, 2012, pp. 1–29, 409–546.
2. Schott JR. *Remote sensing: the image chain approach*. New York: Oxford University Press, 1997, pp. 1–22.
3. Banks J, Carson JS, Nelson BL et al. *Discrete-event system simulation*. New York: Prentice Hall, 2009, pp. 305–482.

4. Landgrebe DA. Hyperspectral image data analysis. *IEEE Signal Proc* 2002; 19(1): 17–28.
5. Chandola V, Banerjee A and Kumar V. Anomaly detection: a survey. *ACM Comp Surv* 2009; 41(3): 15:1–58.
6. Reed IS and Yu X. Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution. *IEEE Trans Acous Speech Signal Proc* 1990; 38(10): 1760–1770.
7. Hsueh M and Chang C-I. Adaptive causal anomaly detection for hyperspectral imagery. In: *IEEE international geoscience and remote sensing symposium*, 2004, vol. 5, pp. 3222–3224.
8. Yanfeng G, Ying L, and Ye Z. A selective kernal PCA algorithm for anomaly detection in hyperspectral imagery. In: *IEEE international conference on acoustics, speech, and signal processing*, 2006, vol. 2, pp. 725–728.
9. Liu W and Chang C-I. Multiple-window anomaly detection for hyperspectral imagery. In: *IEEE international geoscience and remote sensing symposium*, 2008, vol. 2, pp. 41–44.
10. Taitano YP, Geier BA and Bauer KW. A locally adaptable iterative RX detector. *EURASIP Adv Signal Proc* 2010; 2010(341908): 1–10.
11. Chang C-I and Ren H. An experiment-based quantitative and comparative analysis of target detection and image classification algorithms for hyperspectral imagery. *IEEE Trans Geosci Remote Sens* 2000; 38(2): 1044–1063.
12. Chang C-I and Chiang S-S. Anomaly detection and classification for hyperspectral imagery. *IEEE Trans Geosci Remote Sens* 2002; 40(6): 1314–1325.
13. Acito N, Corsini G, Diani M et al. Reducing computational complexity in hyperspectral anomaly detection: a feature level fusion approach. In: *IEEE international geoscience and remote sensing symposium*, 2006, pp. 1804–1807.
14. Nasrabadi NM. A nonlinear kernel-based joint fusion/detection of anomalies using hyperspectral and SAR imagery. In: *IEEE international conference on image processing*, 2008, pp. 1864–1867.
15. Manolakis D and Shaw G. Detection algorithms for hyperspectral imaging applications. *IEEE Signal Proc* 2002; 19(1): 29–43.
16. Stein DWJ, Beaven SG, Hoff LE et al. Anomaly detection from hyperspectral imagery. *IEEE Signal Proc* 2002; 19(1): 58–69.
17. Smetek TE and Bauer KW. A comparison of multivariate outlier detection methods for finding hyperspectral anomalies. *Military Oper Res* 2008; 13(4): 19–43.
18. Tax DMJ and Duin RPW. Support vector domain description. *Patt Recogn Lett* 1999; 20(11–13): 1191–1199.
19. Tax DMJ and Duin RPW. Support vector data description. *Mach Learn* 2004; 54(1): 45–66.
20. Banerjee A, Burlina P and Diehl C. A support vector method for anomaly detection in hyperspectral imagery. *IEEE Trans Geosci Remote Sens* 2006; 44(8): 2282–2291.
21. Banerjee A, Burlina P and Meth R. Fast hyperspectral anomaly detection via SVDD. In: *IEEE international conference on image processing*, 2007, vol. 4, pp. 101–104.
22. Kwon H, Der SZ and Nasrabadi NM. Adaptive anomaly detection using subspace separation for hyperspectral imagery. *Opt Engr* 2003; 42(11): 3342–3351.
23. Liu W and Chang C-I. A nested spatial window-based approach to target detection for hyperspectral imagery. In: *IEEE international geoscience and remote sensing symposium*, 2004, vol. 5, pp. 266–268.
24. Grossman JM, Bowles J, Haas D et al. Hyperspectral analysis and target detection system for the adaptive spectral reconnaissance program. In: *SPIE conference on algorithms for multispectral and hyperspectral imagery IV*, 1998, vol. 3372, pp. 2–13.
25. Clare P, Bernhardt M, Oxford W et al. A new approach to anomaly detection in hyperspectral images. In: *SPIE conference on algorithms and technology for multispectral, hyperspectral, and ultraspectral imagery IX*, 2003, vol. 5093, pp. 17–28.
26. Shi M and Healey G. Using multiband correlation models for the invariant recognition of 3-D hyperspectral textures. *IEEE Trans Geosci Remote Sens* 2005; 43(5): 1201–1209.
27. Gaucel JM, Guillaume M and Bourennane S. Whitening spatial correlation filtering for hyperspectral anomaly detection. In: *IEEE international conference on acoustics, speech, and signal processing* 2005; 5: 333–336.
28. Bellucci JP, Smetek TE and Bauer KW. Improved hyperspectral image processing algorithm testing using synthetic imagery and factorial designed experiments. *IEEE Trans Geosci Remote Sens* 2010; 48(3): 1211–1223.
29. Farrell MD and Mersereau RM. On the impact of PCA dimension reduction for hyperspectral detection of difficult targets. *IEEE Geosci Remote Sens Lett* 2005; 2(1): pp. 192–195.
30. Rouse JW, Haas RH, Schell JA et al. Monitoring vegetation systems in the Great Plains with third ERTS. In: *ERTS Symposium*, 1973, paper no. SP-351, pp. 309–317. NASA.
31. Landgrebe DA. *Signal theory methods in multispectral remote sensing*. Hoboken, NJ, USA: John Wiley & Sons, 2003, pp. 273–322.
32. Fawcett T. An introduction to ROC analysis. *Patt Recogn Lett* 2006; 27(8): 861–874.
33. Kay SM. *Fundamentals of statistical signal processing: estimation theory*. Englewood Cliffs, NJ, USA: PTR Prentice Hall, 1993, pp. 157–218.

Author biographies

Jason P Williams received a BS in Mathematics from the University of Illinois at Urbana-Champaign in 2001, and a MS and PhD in Operations Research from the Air Force Institute of Technology in 2007 and 2012, respectively. He currently serves in the United States Air Force at Ramstein Air Base, Germany. His research interests are in the areas of multivariate statistics and response surfaces.

Trevor J Bihl is a Research Associate at the Air Force Institute of Technology's Center of Operational Analysis. He received a BS and MS in Electrical Engineering from Ohio University at Athens, Ohio. His research interests include multivariate statistics, control theory, and electrical power systems.

Kenneth W Bauer is a Professor of Operations Research at the Air Force Institute of Technology where he teaches classes in applied statistics and pattern recognition. His research interests lie in the areas of automatic target recognition and multivariate statistics. He received

his BS from Miami University at Ohio in 1976, his MEA from the University of Utah in 1980, his MS from the Air Force Institute of Technology in 1981, and his PhD from Purdue University in 1987.