Research Article
A Locally Adaptable Iterative RX Detector

Yuri P. Taitano, Brian A. Geier, and Kenneth W. Bauer Jr.

Air Force Institute of Technology, 2950 Hobson Way, Wright Patterson AFB, OH 45433-7765, USA

Correspondence should be addressed to Kenneth W. Bauer Jr., kenneth.bauer@afit.edu

We present an unsupervised anomaly detection method for hyperspectral imagery (HSI) based on data characteristics inherit in HSI. A locally adaptive technique of iteratively refining the well-known RX detector (LAIRX) is developed. The technique is motivated by the need for better first- and second-order statistic estimation via avoidance of anomaly presence. Overall, experiments show favorable Receiver Operating Characteristic (ROC) curves when compared to a global anomaly detector based upon the Support Vector Data Description (SVDD) algorithm, the conventional RX detector, and decomposed versions of the LAIRX detector. Furthermore, the utilization of parallel and distributed processing allows fast processing time making LAIRX applicable in an operational setting.

1. Introduction

In many experiments, the RX detector is modified in a preprocessing fashion [1–10] in order to minimize the false alarm rate while attaining a reasonable true positive rate. In most cases the modifications that are proposed can be generally described as dimensionality reduction variants coupled with RX [5, 10], window adjustments for covariance estimates [2–4], and the RX detector coupled with an entropy and a nonparametric approach [1, 6]. In cases where a new anomaly detection methodology is proposed, the RX detector is often used as a performance benchmark [9, 11–14].

The literature on anomaly detection in HSI is quite extensive [1–5, 7–12, 14–18] with major contributions appearing rapidly after Reed and Yu [19, 20]. In anomaly detection, the goal has always been to distinguish background from potential targets in an automatic fashion while jointly minimizing false alarms and maximizing true positives.

The RX detector is prone to high false alarms because the local Gaussian assumption is largely inaccurate [11]. The purpose of this paper is to propose a refinement of the RX detector by taking into account the anomaly dominance upon first and second order statistic estimation. That is, we wish to force stability upon the subsequences, locally defined with respect to a window size, in order to reduce the bias and

error when estimating the mean and covariance, respectively. The subsequences are refined by removing anomalies, in an iterative fashion, from consideration in local statistics estimation. Even so, the refined subsequence that is used to estimate a mean vector, μ, and the covariance matrix, $\sum$, is likely to still be nonGaussian; but, as is demonstrated subsequently, it often provides a better false alarm rate than the conventional RX because its estimates are not as contaminated by anomalies.

To illustrate the potential of this idea, consider the following "abbreviated" image created from a desert image (see Figure 1). The image is "abbreviated" in that the targets have been moved closer together than in the original image by simply eliminating columns of image pixels. This creates a situation with a very nonhomogeneous background. This small image is 63×49 pixels. A 25×25 window will be used in all subsequent processing. In Figure 1(a) the truth mask shows the known objects of interest. Figure 1(b) shows the RX scores for the HSI image, Figure 1(c) shows the RX scores for the 1st 10 principal components (instead of using the entire HSI data-cube), Figure 1(d) shows the output of LAIRX for 2 iterations (called LAIRX (2) in subsequent discussion), using PCA for the input. In Figures 1(b), 1(c), and 1(d) RX scores are displayed such that anomalies are expected to "fire" as bright and the background should "fire" as dark. Figure 1(b) shows that many background

pixels may be declared as anomalies (depending on the threshold) for RX. Figure 1(c) shows the benefit of using principal components as input to RX. Figure 1(d) shows a large reduction in lighter pixels in the background using one iterative refinement of the covariance calculation used in RX.

2. Methods

The data used in our experiment are from the ARES desert and forest radiance collections. In our analysis we only consider two classes, background and certain man-made targets. The goal of our analysis is to distinguish the latter from many sources of background variation, such as brush, roads, forest, large rock formations and other natural anomalies.

2.1. Locally Adaptive Iterative RX Detector (LAIRX). Reed and Yu derived an anomaly detector using a Generalized Likelihood Ratio Test (GLRT), which was later dubbed the RX detector [19, 20]. The detector GLRT is simplified by assuming that background pixel vectors are *iid* normal with estimated mean, $\hat{\mu}$, and covariance matrix, $\hat{\Sigma}$, that is,

$$\mathrm{RX}(\mathbf{x}) = (\mathbf{x} - \hat{\mu})^T \left[\frac{n}{n+1}\hat{\Sigma} + \frac{1}{n+1}(\mathbf{x} - \hat{\mu})(\mathbf{x} - \hat{\mu})^T \right]^{-1}$$
$$\times (\mathbf{x} - \hat{\mu}) \xrightarrow[n \to \infty]{} (\mathbf{x} - \hat{\mu})^T \hat{\Sigma}^{-1}(\mathbf{x} - \hat{\mu}). \quad (1)$$

In short, an incoming pixel vector, $\mathbf{x}$, is the center of a neighborhood of size n which is checked for irregularity via the distance formulated above. That is, the pixel vector is checked to see if it lies outside the hyper ellipse whose location and shape are determined via $\hat{\mu}$ and $\hat{\Sigma}$, respectively. An anomalous vector is declared given that $\mathrm{RX}(\mathbf{x}) > \chi^2_{\alpha,p}$, where $\chi^2_{\alpha,p}$ is the αth quantile of a χ^2 distribution with p degrees of freedom. For more information see [18].

The idea, following the philosophy of the RX method, is to place a window about each pixel in an image and use local image statistics to determine whether or not the point is anomalous. It is evident that such a method can suffer from at least 3 potential complications. First, the window pixel vectors are almost never statistically independent. Second, such vectors are not typically identically distributed. Further, outliers (the things we are looking for) can seriously compromise the integrity of the local statistics, particularly the estimated covariance matrix. The first and second complications are the subject of current research. In this paper, we examine the third complication. A look at outlier effects and some remedies is given in [18]. The basic approach given in this paper is laid out in [21]. Here, we propose to deal with the outlier effects in an iterative fashion. As we process the basic RX algorithm across the image, we maintain a catalog of anomalous pixels, this is the 1st iteration. Indeed, if we simply quit after processing the image once, we would have simply run the RX method. A second iteration is applied, only this time we withhold the anomalous pixels from consideration in calculating the local statistics. So long as we find new anomalies the iterations

continue; otherwise, the algorithm terminates. Hence, in LAIRX we allow the RX detector to be iteratively refined with respect to the estimation of μ and Σ while keeping track of detected anomalies. LAIRX has the following steps:

Step 1. Reduce the dimensionality to a set of p principal components via a global estimate of Σ.

Step 2. Apply the RX detector to the data matrix using a pixel process window. If this is not the 1st iteration, withhold anomalous pixels identified in the previous iteration from the local estimation of μ and Σ.

Step 3. Identify those RX scores that exceed $\chi^2_{\alpha,p}$. These are referred to as anomalies. This step ends an iteration.

Step 4. If the set of pixels identified as anomalies in Step 3 are identical to the set of pixels identified as anomalies in the previous iteration then go to Step 5, otherwise; return to Step 2.

Step 5. Map detected anomalies to the image space.

Once LAIRX has terminated we are assume that the respective window sequences' anomaly indicator has converged almost surely to some target given the $\chi^2_{\alpha,p}$ cut-off. The subsequences associated with the target may or may not be Gaussian but the refined μ and Σ estimates should result in a higher true positive rate coupled with a relatively low false alarm rate when compared to the conventional RX detector because the sequence is forced to be more *iid* and, hence, estimation bias is reduced. The global SVDD algorithm was chosen as a competing algorithm because of its recent promise as an efficient and powerful anomaly detector.

In what follows, the LAIRX detector is compared to the SVDD algorithm, the conventional RX detector and a decomposed version of the LAIRX detector. This later comparison was made in an effort to more fully understand the performance of the LAIRX detector. Each anomaly detector was applied to four images; see Figure 2, containing multiple vehicles, varying land formations, sage brush and a road. These images vary in the difficulty they present to the algorithms as reflected in the SNR plots of Figure 3.

2.2. Parallel Implementation of LAIRX. The RX detector is naturally a computationally inefficient task since M matrix inversions are required, $M \approx 52000$ in our analysis. This computational burden was an inspiration for the application of SVDD in HSI anomaly detection [22]. However, most implementations of the RX detector are not optimized. We decided to optimize the RX algorithm via parallel and distributed processing on a dual quad-core machine. The implementation is very simple given the Matlab Parallel Computing Toolbox but is described nonetheless.

2.2.1. Basic Setup. The hyperspectral image is formulated as a data matrix where columns are wavelengths and rows are pixels, respectively. A window is moved row-wise across the data matrix at a single row increment where each center
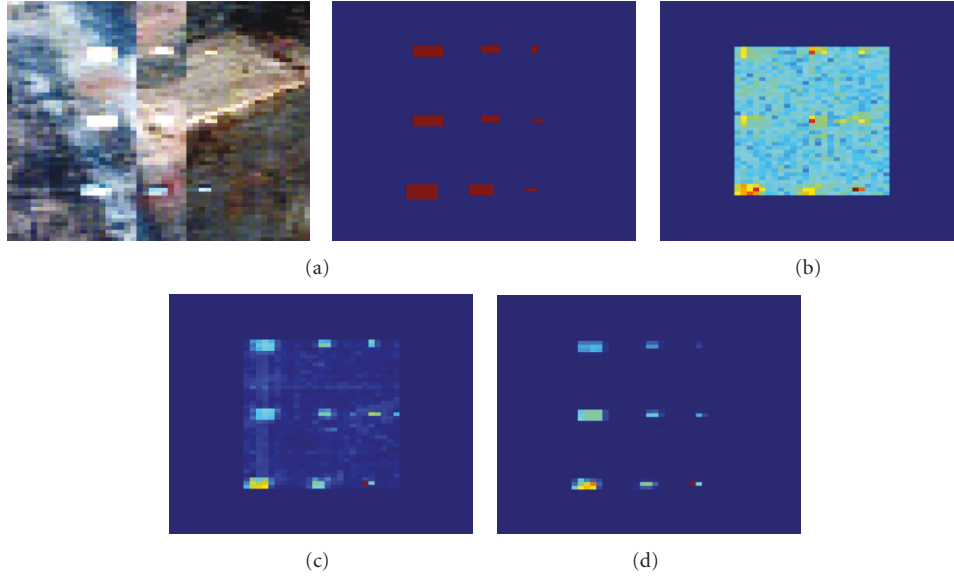
Figure 1: Comparison of RX and LAIRX. (a) Original test image with truth mask, (b) RX scores, (c) RX using 1st 10 Principal Components, (d) LAIRX scores.
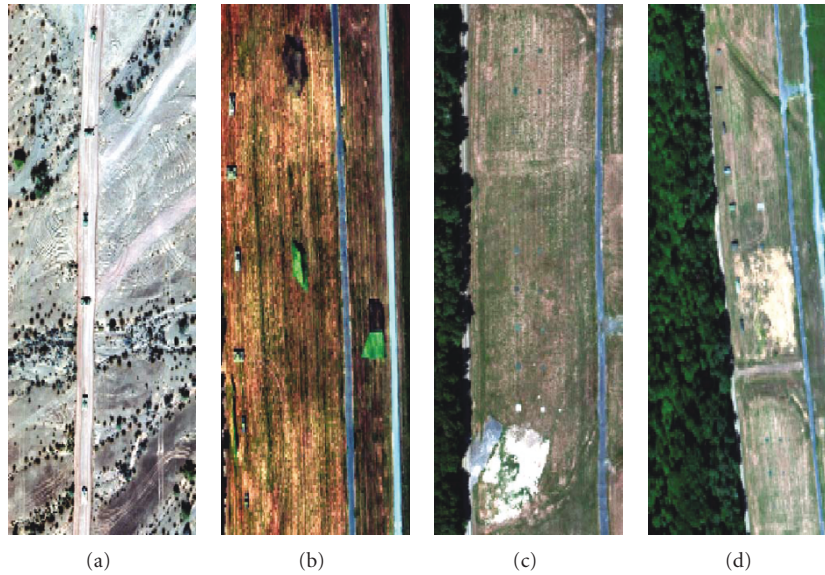

Figure 2: RGB of image scenes. (a) ARES1D, (b) ARES1F, (c) ARES2F, (d) ARES3F.

pixel serves as input to the RX detector. The data held locally in each window is used to estimate μ and Σ. The components of μ and Σ are based on a set of p principal components from the wavelength bands, where $p = 10$ in our analysis. The number of principal components is dependent on the data collection environment and should be determined via exploratory data analysis. From our analysis, we found that the desert radiance images required only two principal components while the forest radiance images required about ten.

2.2.2. Parallelization Scheme. (1) Generate all possible window indices where each column indexes a window of data.

The row midpoint of these data matrices are the center pixels for that particular window. This step makes RX and LAIRX an "embarrassingly parallel" problem.

(2) The data partitions are batched in batches of size G where G is the number of available processor cores, $G = 8$ in our setup.

(3) Each batched data partitions are processed using the RX detector simultaneously.

(4) Results are pooled and used for subsequent analysis.

2.3. Global SVDD Anomaly Detector. The Support Vector Data Description (SVDD) algorithm was originally proposed by Tax and Duin as a means for exploring one class
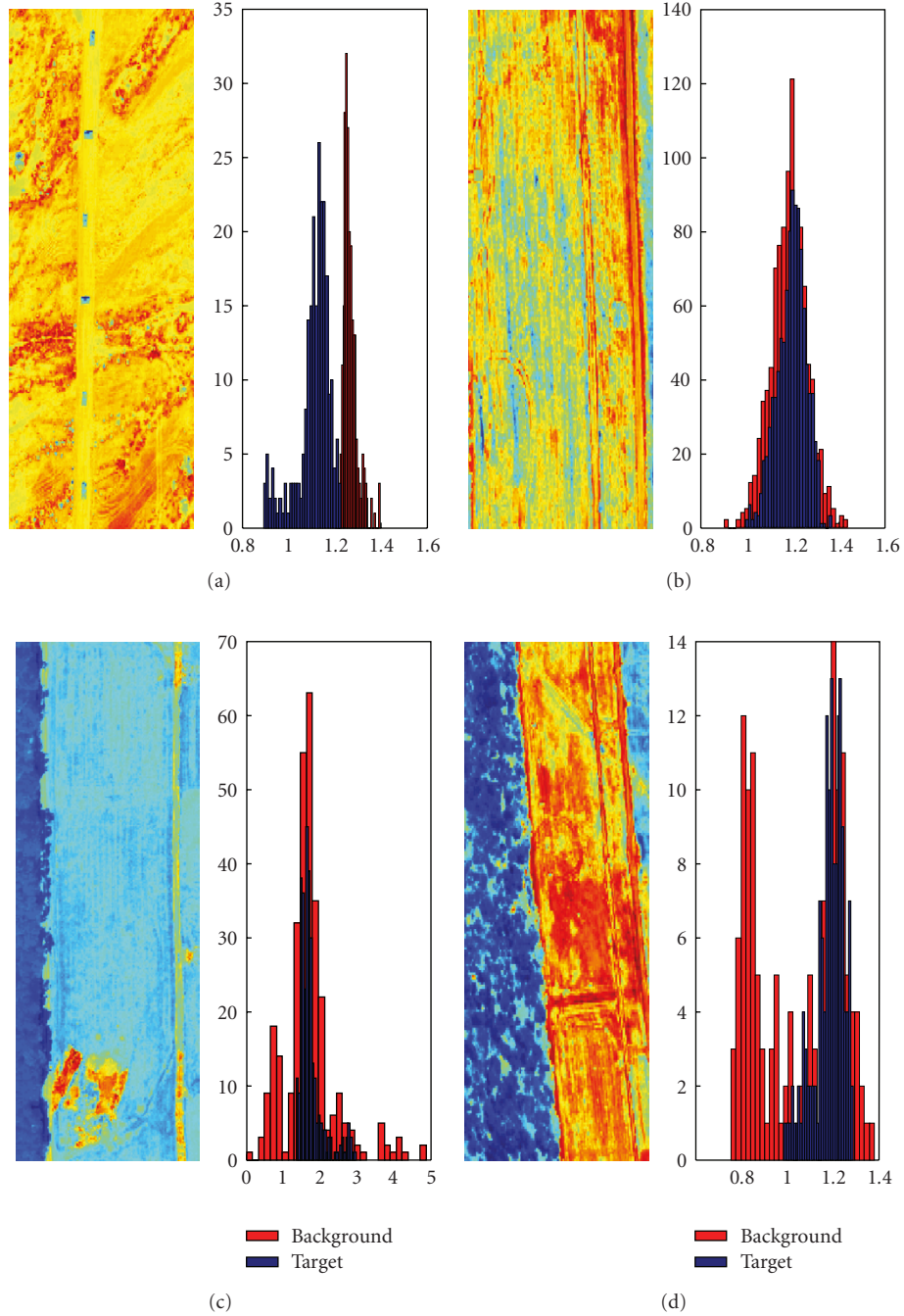
(a)

(b)

(c)

(d)

FIGURE 3: Signal to noise ratio (SNR) pixel map, that is, μ/σ, displayed with background-target SNR histogram. (a) ARES1D, (b) ARES1F, (c) ARES2F, (d) ARES3F.

distributions [23, 24]. The SVDD anomaly detector applied to HSI was proposed by Banerjee et al. [22]. In general, the goal of the SVDD algorithm is to find the hypersphere with minimum volume about a set of random vectors. In our case these random vectors are the background pixels with each dimension corresponding to a different spectral band, that is, for a set of pixels, $\mathbf{T} = \{x_i, i = 1,\dots,M\}$, we seek the minimum enclosing hypersphere, $\mathbf{S} = \{x : \|x - a\|^2 < R^2\}$,

that contains $\mathbf{T}$. This is a constrained optimization problem stated as.

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

The center a and radius R are found by optimizing the following Lagrangian:

$$L(R, a, \alpha_i) = R^2 - \sum_i \alpha_i \{R^2(\langle x_i, x_i \rangle - 2\langle a, x_i \rangle + \langle a, a \rangle)\}. \quad (3)$$
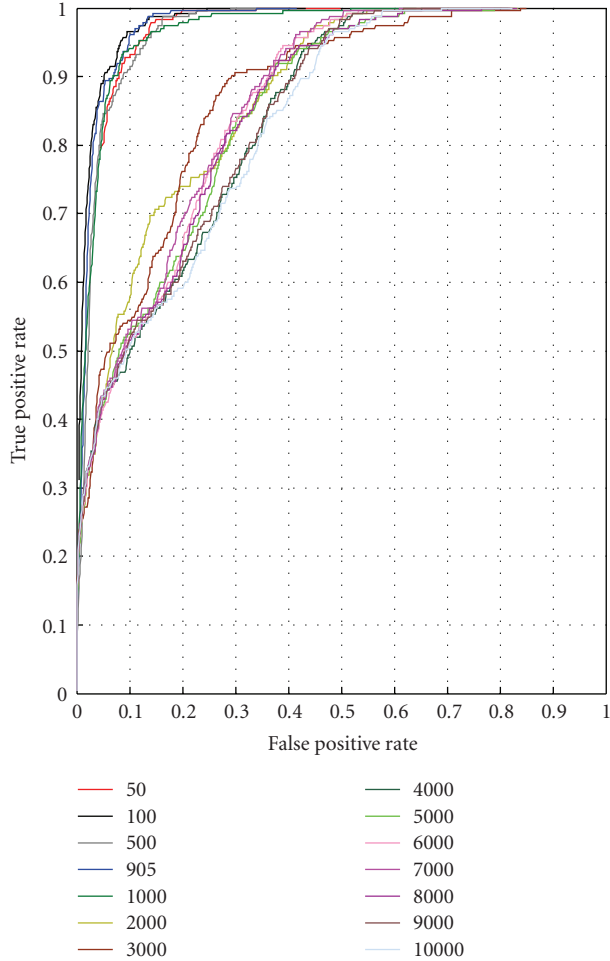
FIGURE 4: σ selection for ARES1D image. Effect of bandwidth on SVDD performance. The highest performance is achieved when $\sigma \approx 905$; σ was incrementally varied from $50 \rightarrow 10,000$.

The next step is to apply the kernel trick with a kernel function $\Phi(x)$, which is usually taken to be Gaussian [11, 22–24]. Once L is optimized with respect to α_i and after incorporating the kernel function, $K(y, x)$, we get the following SVDD statistic [11]

$$
\begin{aligned}
\text{SVDD}(y) &= R^2 - K(y, y) + \sum_i \propto_i K(y, x_i) \\
&= C + 2 \sum_i \propto_i K(y, x_i),
\end{aligned}
\tag{4}
$$

where

$$
K(x, y) = \exp\left(\frac{-\|x - y\|^2}{\sigma^2}\right).
\tag{5}
$$

The radial basis function (RBF) parameter σ^2 controls how well the SVDD generalizes to unseen data. The choice of this parameter is driven by the data and must be chosen empirically. In our analysis, $\mathbf{y}$ is a pixel vector and $\mathbf{x}$ is a support vector obtained via SVDD given background spectra.

For HSI anomaly detection, Banerjee et al. propose a fast global SVDD detector [11]. The algorithm that they propose is fast because it uses a small subset of the data to build the SVDD model. The algorithm is as follows:

(1) Randomly select a set of N background pixels from the training set.

(2) Estimate an optimal value for σ^2 using a cross-validation or minimax method given the set of background spectra.

(3) Estimate the SVDD parameters (a, α_i, R) to model the region of support for the background given a random subset of the background spectra and σ^2.

(4) For each pixel in the data matrix perform the decision test:

  (i) if SVDD(y) is less than the detection threshold t, the pixel is part of the background.
  (ii) else, declare the pixel as an anomaly.

As you can see in Figure 4, there is a performance effect as σ^2 is varied. The only free parameter in the RBF is σ^2. We wish to find the optimal σ^2 that is able to fully describe the nontarget/background spectra.

A common method is to estimate σ^2 while minimizing the false positive rate (Pfa) via an argument based on leave-one-out cross validation [11, 22–25], which gives the following upper bound on Pfa:

$$
Pfa \le E\left[\frac{\#SV}{N}\right],
\tag{6}
$$

N is the number of samples selected to train the SVDD model and $\#SV$ are the number of support vectors required to describe the data. Based upon the above inequality, Banerjee et al. [22] propose an approximate minimax estimate for σ^2 as

$$
\begin{aligned}
\hat{\sigma} &= \min_\sigma \frac{1}{M} \sum_{i=1}^{M} PFa_i \\
&\approx \min_\sigma \left\{ \frac{1}{M} \sum_{i=1}^{M} E\left[\frac{\#SV_i}{N}\right] \right\} \\
&\approx \min_\sigma \left\{ \frac{1}{M} \sum_{i=1}^{M} \left[\frac{\#SV_i}{N}\right] \right\},
\end{aligned}
\tag{7}
$$

where M is the number of replicates. Based upon the above result, the algorithm to obtain the global estimate for σ^2 is the following.

(1) Generate M equal sets of training data by randomly selecting pixels from the background.

(2) For each set of training data, the SVDD decision boundary is determined using different values for σ^2.

(3) For each value of σ^2, the average fraction of support vectors, $(1/M) \sum_{i=1}^{M} [\#SV_i/N]$, is computed over all of the training sets.
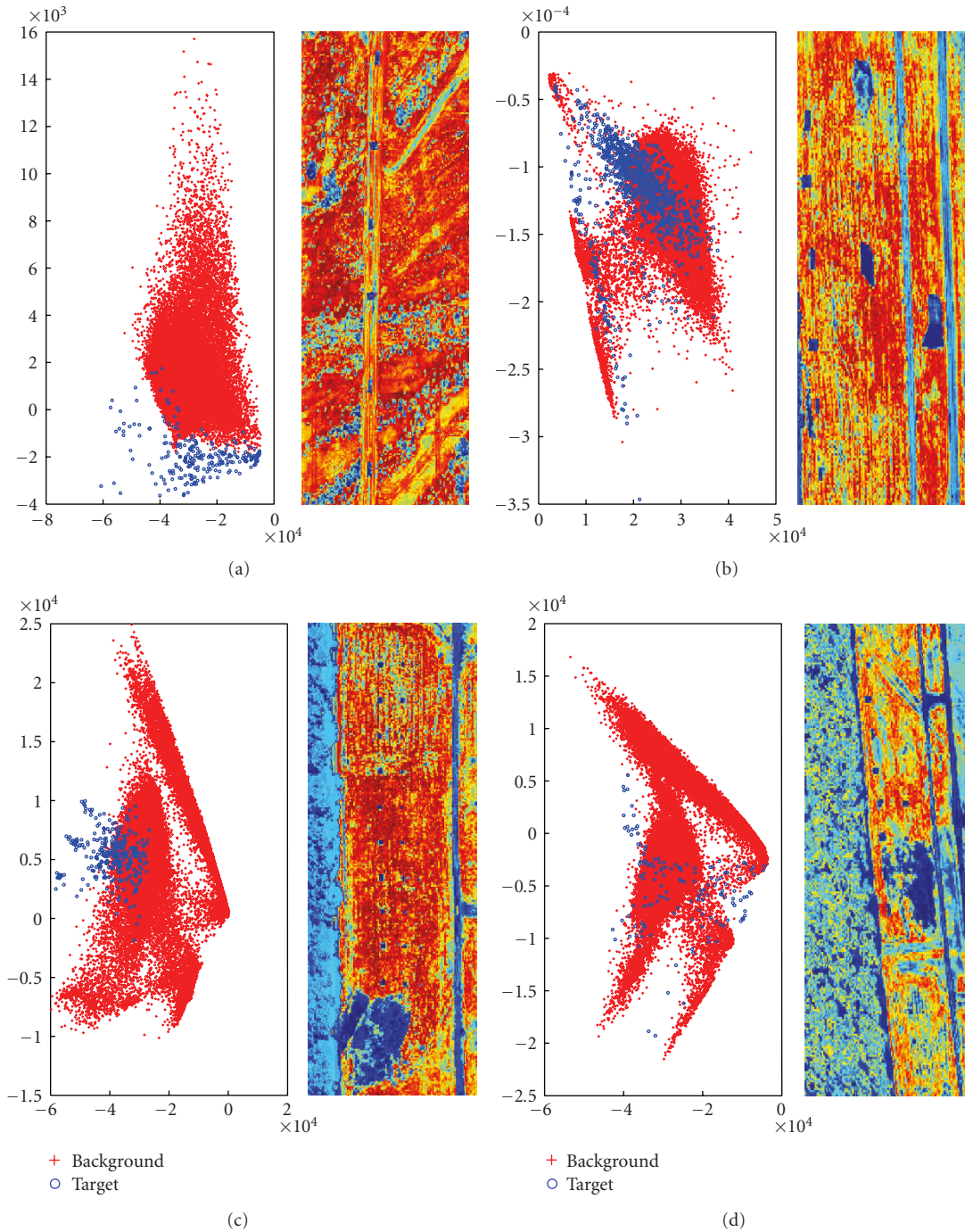
Figure 5: 2-D Principal Component Analysis view and SVDD pixel map. (a) ARES1D, (b) ARES1F, (c) ARES2F, (d) ARES3F. *Note: Blue is target and red is background.*

(4) The σ^2 that produces the smallest average fraction of support vectors is the minimax estimate.

We discovered that the average fraction of support vectors is at a minimum when $\sigma^2 \geq 905$. Therefore, our minimax estimate is 905 because this is the smallest σ^2 that allows us to effectively describe our data. If $\sigma^2 > 905$ then our detection results may be poor because the resulting SVDD model is overly general. Note that when using the minimax approach, 60% of the training data is 300 pixel vectors while there are 149 features. The balance between sample size and number of variables is causing the minimax estimation to converge at a minimum that allows a fairly high *Pfa* while maintaining an effective characterization of the background spectra. This demonstrates that the bandwidth parameter selection is robust to small sample sizes relative to the number of dimensions or spectral bands.
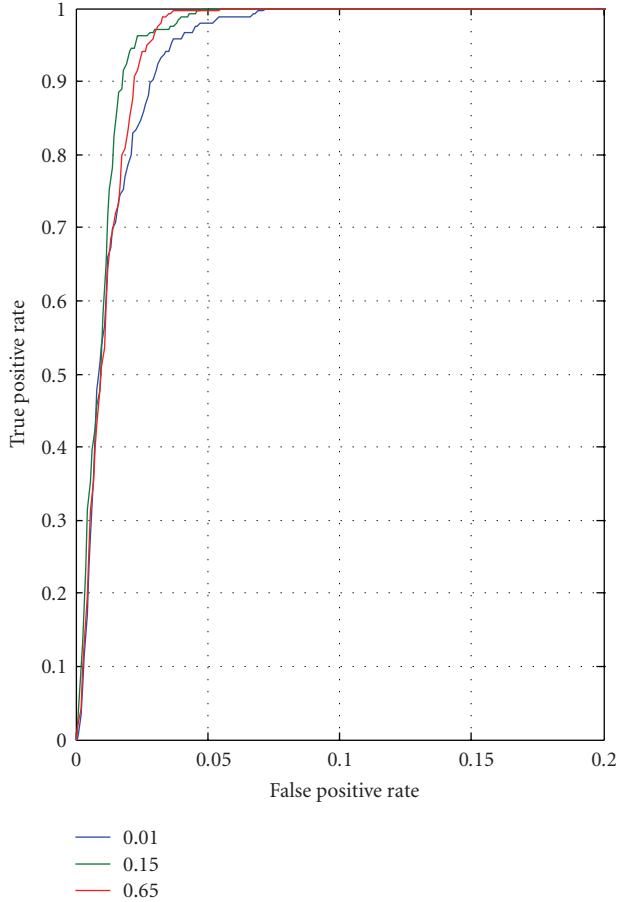
FIGURE 6: Rejection Criteria Effect: ROC curves for LAIRX statistic map thresholded at 0.01, 0.15, and 0.65, respectively.

Figure 5 is SVDD pixel map with an accompanying 2D principal component scatter plots. These will contrasted in subsequent analysis.

2.4. Experiment Description. In our analysis we included all wavelength bands that were not contaminated by atmospheric absorption, which resulted in 149 bands. The global SVDD approach is supervised in the sense that it requires as input a known set of background spectra while LAIRX and RX are unsupervised. For SVDD, the sample size used when sampling from the background spectra was $N = 500$. This size was chosen based upon computing limitations while accommodating the high dimensional feature space. For LAIRX, the maximum number of iterations and principal components was 50 and 10, respectively. Based on our exploratory data analysis 10 principal components was sufficient. A window size of 25×25 was employed.

Each anomaly detector was applied to four images; see Figure 2, containing multiple vehicles, varying land formations, sage brush and a road. ARES1D is desert radiance while the other three are forest radiance. Three of the images area contains less than 1% target pixels while one image has about 3.4% target pixels.

The goal of our analysis was to compare a promising anomaly detector (global SVDD), a benchmark detector (RX) and our proposal (LAIRX). The RX algorithm is run in two different modes. RX-FULL is run on the full 149 bands of the image cube; whereas, RX-PCA is run on the 10 retained principal components. Finally, in an effort to examine the iteration sensitivity of LAIRX we created a version denoted LAIRX(2) that only performs one iterative refinement (in other words, 2 RX-type iterations). In order to make inferences about these competing algorithms we investigate the runtime, percentage of target pixels, area under the ROC curve restrained to false positive cases falling below 0.20, and the true positive rate at a fixed false positive rate of 0.05. In what follows (Table 1), AUC denotes the "area under the (ROC) curve (false positive rates from 0–.2)" and TPR is the "true positive rate" at a fixed a false positive rate of 0.05.

At the onset, we believed that performance gains would be substantial given that we force our multidimensional data sequences within each window to be more *iid* even though these data sequences are likely to not be locally Gaussian. Even if the multivariate populations are not Gaussian, by iteratively refining the anomaly detection we are introducing robustness to our final RX($\mathbf{x}$) score map, upon which a threshold is applied to determine the final classification. A drawback to LAIRX is that you need to know a good rejection rate for the iterative refinement which is largely data dependent as you can see in Figure 6. This problem is a topic of future research.

3. Results

In what follows, ROC curves are presented for each image in Figure 7 and summary statistics are offered in Table 1.

For ARES1D, the background and target pixels are linearly separable given the first and second principal components (see Figure 5(a)). Additionally, the SNR is showing a clear separation with lower SNR values present for the target pixels, which is encouraging (see Figure 5(b)). We would expect with these observations that the RX-type classifiers should do very well and that the nonlinear classifier, SVDD, should also perform reasonably well. When viewing the ROC curve for ARES1D, you can see that PCA is beneficial for the RX derivative methods. The SVDD algorithm is performing well but as you can see in the SVDD statistic map there are many locally clustered areas which display the same in value as the target pixels.

The image ARES1F poses difficulties for all the algorithms tested. ARES1F which has a very noisy SNR statistic map, as you can see in Figure 3(b). This image in particular, highlights the benefits of LAIRX's iterative approach.

The image scene is similar for both ARES2F and ARES3F (Figure 2). As you can see in Figure 3, the SNR map is showing a reasonable segmentation for both of these images and the distributions are somewhat separable. Table 1 and Figure 7 show that performance is good across the board with LAIRX being superior. It is interesting to point out that the nonlinear technique is performing poorly in contrast to
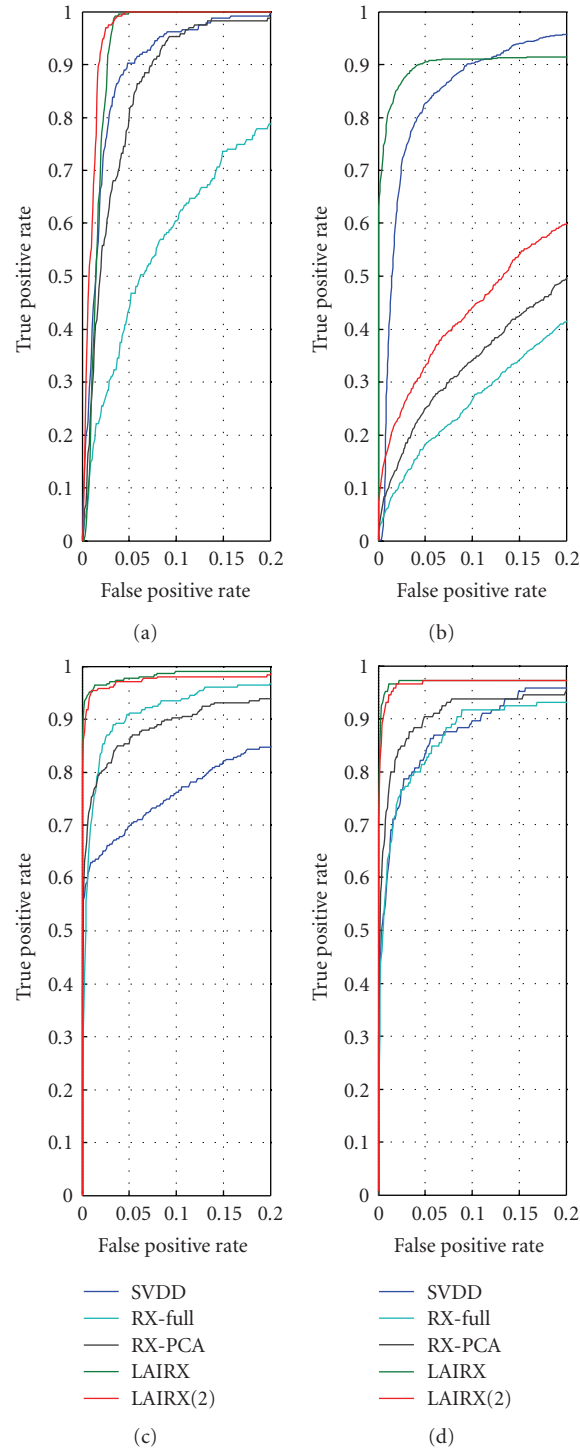
FIGURE 7: Receiver Operating Characteristic (ROC) curves for each image. (a) ARES1D, (b) ARES1F, (c) ARES2F, (d) ARES3F.

the other algorithms. Additionally, the data associated with these images is separable in a lower dimensional subspace, as you can see in Figure 3, which indicates that the RX type approach should perform well. Also, note that for ARES2F and ARES3F the SVDD statistic maps depicted in Figure 5 is showing that large natural anomalies have a similar value as the targets, which leads to a higher false positive rate.

4. Discussion

We have presented an unsupervised automatic target detection algorithm which builds upon the conventional RX detector by direct manipulation of the RX algorithm. As a practical matter, the LAIRX detector must have data pre-processed as principal components before detection which

Table 1: Tabular results by image.

Image	Algorithm	Runtime (min)	Image Prop	AUC	TPR
	SVDD	1.46	291 × 199 pixels	0.18	0.90
	RX-Full	3.93		0.11	0.44
ARES1D	RX-PCA	**0.1992**	0.41 % Targets	0.17	0.80
	LAIRX	48.54		0.37	**1.00**
	LAIRX(2)	4.12	6 Targets	**0.45**	**1.00**
	SVDD	1.05	191 × 160 pixels	0.16	**0.83**
	RX-Full	2.11		0.05	0.18
ARES1F	RX-PCA	**0.10**	3.40% Targets	0.06	0.25
	LAIRX	16.11		**0.30**	0.80
	LAIRX(2)	1.99	10 Targets	0.14	0.26
	SVDD	1.61	312 × 152 pixels	0.15	0.70
	RX-Full	3.14		0.18	0.91
ARES2F	RX-PCA	**0.16**	0.66% Targets	0.18	0.85
	LAIRX	19.46		0.47	**1.00**
	LAIRX(2)	3.24	30 Targets	**0.48**	**1.00**
	SVDD	1.10	226 × 136 pixels	0.17	0.83
	RX-Full	1.93		0.17	0.82
ARES3F	RX-PCA	**0.09**	0.48% Targets	0.18	0.90
	LAIRX	11.99		0.44	**0.97**
	LAIRX(2)	1.99	20 Targets	**0.47**	0.96

hinders real-time viability while the global SVDD builds a model based upon background spectra and then classifies raw pixel vectors as anomalous as they are received. However, the global SVDD is a supervised algorithm given that a set of background spectra and RBF spread parameter are specified which may limit its real-time viability in a dynamic operational setting.

For the types of images analyzed here, our results have shown that LAIRX is a reasonable competitor to the SVDD algorithm and that a linear technique can perform well in a nonlinear environment after statistic estimation modification. We have also demonstrated, see Table 1 and Figure 7, that the algorithmic steps taken to create LAIRX interact in a way that lead to higher true positives coupled with low false positives. By introducing iterative refinement, we are getting better performance because the first and second order statistic estimation have less bias and error. Our method has demonstrated potential in an image scene with sparse vehicle activity. Whether or not similar results follow in a densely populated target environment remains to be seen.

Both LAIRX and LAIRX(2) are implemented in a parallel and distributed fashion which makes these algorithms computationally efficient. The processing time for LAIRX and LAIRX(2) are constrained by the number of available processor cores. In general, LAIRX is somewhat slow as you can see in Table 1 but the ready availability of cluster machines and even affordable 8 processor core machines make LAIRX a viable algorithm in an operational setting. In contrast, the runtime of the global SVDD algorithm is fixed from an algorithmic perspective of parallel and distributed computing.

References

[1] W. Di, Q. Pan, Y.-Q. Zhao, and L. He, "Multiple-detector fusion for anomaly detection in multispectral imagery based on maximum entropy and nonparametric estimation," in *Proceedings of the 8th International Conference on Signal Processing (ICSP '06)*, vol. 3, January 2006.

[2] M. Hsueh and C.-I. Chang, "Adaptive causal anomaly detection for hyperspectral imagery," in *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS ' 04)*, vol. 5, pp. 3222–3224, January 2004.

[3] W. Liu and C.-I. Chang, "A nested spatial window-based approach to target detection for hyperspectral imagery," in *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS '04)*, vol. 5, pp. 266–268, January 2004.

[4] W. Liu and C.-I. Chang, "Multiple-window anomaly detection for hyperspectral imagery," in *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS '08)*, vol. 2, pp. 41–44, January 2008.

[5] F. Mei, C. Zhao, L. Wang, and H. Huo, "Anomaly detection in hyperspectral imagery based on kernel ICA feature extraction," in *Proceedings of the 2nd International Symposium on Intelligent Information Technology Application (IITA '08)*, vol. 1, pp. 869–873, January 2008.

[6] F. Mei, C. Zhao, H. Huo, and Y. Sun, "An adaptive kernel method for anomaly detection in hyperspectral imagery," in *Proceedings of the 2nd International Symposium on Intelligent Information Technology Application (IITA '08)*, vol. 1, pp. 874–878, January 2008.

[7] N. M. Nasrabadi, "A nonliner kernel-based joint fusion/detection of anomalies using hyperspectral and SAR imagery," in *Proceedings of the IEEE International Conference on Image Processing (ICIP '08)*, pp. 1864–1867, 2008.

[8] N. Renard and S. Bourennane, "Dimensionality reduction based on tensor modeling for classification methods," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 4, pp. 1123–1131, 2009.

[9] G. Yanfeng, Z. Ye, and L. Ying, "Unmixing component analysis for anomaly detection in hyperspectral imagery," in *Proceedings of the IEEE International Conference on Image Processing (ICIP '06)*, pp. 965–968, 2006.

[10] Y. Gu, Y. Liu, and Y. Zhang, "A selective kernel PCA algorithm for anomaly detection in hyperspectral imagery," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '06)*, vol. 2, pp. 725–728, January 2006.

[11] A. Banerjee, P. Burlina, and R. Meth, "Fast hyperspectral anomaly detection via SVDD," in *Proceedings of the 14th IEEE International Conference on Image Processing (ICIP '07)*, vol. 4, pp. 101–104, January 2007.

[12] J.-M. Gaucel, M. Guillaume, and S. Bourennane, "Whitening spacial correlation filtering for hyperspectral anomaly detection," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '05)*, vol. 5, pp. 333–336, January 2005.

[13] S. Tiwari, S. Agarwal, and A. Trang, "Texture feature selection for buried mine detection in airborne multispectral imagery," in *Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS '08)*, vol. 1, pp. 145–148, 2008.

[14] S. M. Thornton and J. M. F. Moura, "The fully adaptive GMRF anomaly detector for hyperspectral imagery," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '00)*, vol. 1, pp. 37–40, 2000.

[15] C.-I. Chang and S.-S. Chiang, "Anomaly detection and classification for hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 6, pp. 1314–1325, 2002.

[16] C.-I. Chang and H. Ren, "An experiment-based quantitative and comparative analysis of target detection and image classification algorithms for hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 2, pp. 1044–1063, 2000.

[17] M. Shi and G. Healey, "Using multiband correlation models for the invariant recognition of 3-D hyperspectral textures," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 5, pp. 1201–1209, 2005.

[18] T. E. Smetek and K. W. Bauer Jr., "A comparison of multivariate outlier detection methods for finding hyperspectral anomalies," *Military Operations Research*, vol. 13, no. 4, pp. 19–44, 2008.

[19] I. S. Reed and X. Yu, "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 10, pp. 1760–1770, 1990.

[20] X. Yu, L. E. Hoff, I. S. Reed, A. M. Chen, and L. B. Stotts, "Automatic target detection and recognition in multiband imagery: a unified ML detection and estimation approach," *IEEE Transactions on Image Processing*, vol. 6, no. 1, pp. 143–156, 1997.

[21] Y. P. Taitano, *Hyperspectral imagery target detection using the iterative RX detector*, M.S. thesis, School of Operation Science, Air Force Institute of Technology, Wright-Patterson AFB, Ohio, USA, March 2007, AFIT/GOR/ENS/07-25.

[22] A. Banerjee, P. Burlina, and C. Diehl, "A support vector method for anomaly detection in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 8, pp. 2282–2291, 2006.

[23] D. M. J. Tax and R. P. W. Duin, "Support vector domain description," *Pattern Recognition Letters*, vol. 20, no. 11–13, pp. 1191–1199, 1999.

[24] D. M. J. Tax and R. P. W. Duin, "Support vector data description," *Machine Learning*, vol. 54, no. 1, pp. 45–66, 2004.

[25] V. N. Vapnik, *Statistical Learning Theory*, Wiley Series on Adaptive and Learning Systems for Signal Processing, Communications, and Control, John Wiley & Sons, New York, NY, USA, 1998.