

# Different Approaches for Improved Covariance Matrix Estimation in Hyperspectral Anomaly Detection

Stefania Matteoli, Marco Diani, and Giovanni Corsini

**Abstract**— Anomaly detection in remotely sensed hyperspectral images has proven to be valuable in many applications, such as search-and-rescue operations, and hazardous material detection. The benchmark anomaly detection algorithm is the Reed-Xiaoli (RX) Detector, which is based on the Local Multivariate Normality of background. RX algorithm, along with its many modified versions, has been widely explored, and the main concerns identified are related to local background covariance matrix estimation. Besides the well-known small-sample size problem, other limitations have been found affecting covariance matrix estimation, e.g. local background non-homogeneity and contamination from adjacent targets. Since these critical aspects are deeply different in nature, like the situations from which they arise, they have been typically discussed within different frameworks, disregarding the possible existing connections while developing different approaches to solution. On the other hand, the aforementioned critical aspects may occur together in the same operating scenario, and all of them have to be taken into consideration when approaching anomaly detection, since they may strongly affect detection performance. Therefore, an analysis of the possible existing links seems crucial in order to assess if existing algorithms, maybe designed ad-hoc to solve a specific problem, can handle more complex situations. As part of our research into this framework, in this work, the aforementioned limitations have been investigated from an anomaly detection perspective, and the corresponding approaches to improved covariance matrix estimation have been analyzed by using real hyperspectral data.

**Index Terms**—Hyperspectral imagery, anomaly detection, covariance matrix, regularization, small-sample size, robust statistics, quasi-local covariance matrix, segmented covariance matrix, cluster-based approach.

## I. INTRODUCTION

Hyperspectral sensors collect data in hundreds of narrow contiguous spectral bands with a very dense spectral sampling, thus providing a powerful means for the discrimination of different materials on the basis of their unique *spectral signatures*. This has led to several applications for the analysis of remotely sensing data, such as classification and target detection. Specifically, the latter refers

to methodologies aimed at searching into the image for “rare” objects on the basis of their known spectral signature. Target detection has been proven valuable in many applications, such as hazardous material detection, and mine detection. In this work, we are interested in anomaly detection, which is a target detection problem in which no knowledge about the spectrum of the target of interest is available. The goal of anomaly detection is, in fact, to detect those pixels that are anomalous with respect to the background, so as to identify regions of interest that deserve further and more refined processing since they may contain potential targets.

Within the anomaly detection framework, Reed-Xiaoli detector (RX) [1] is considered the benchmark algorithm for detecting anomalies in hyperspectral images. It relies upon the assumption of Local Multivariate Normality (LMN) of background, and requires the background covariance matrix to be estimated from a set of pixels (Estimation Window, EW) in the neighborhood of the pixel under test. RX, along with many detectors developed following its footsteps, has been widely explored in the hyperspectral anomaly detection literature, and the main concerns identified are related to local background covariance matrix estimation and subsequent inversion.

The most well known problem is the small sample-size: the local background covariance matrix has to be estimated from a small number of very high-dimensional samples, thus resulting in a badly-conditioned and unstable estimate that strongly affects detection performance. A possible solution that disregards dimensionality reduction, not always recommended in rare target detection applications, is covariance matrix regularization [2]. This has been widely explored in classification scenarios, where labeled training data are typically available that allow the regularization parameters to be selected through cross-validation. In rare target detection settings, ground-truth data are very seldom available, and regularization parameters have to be estimated from the image itself.

Besides the well-known small-sample size problem, other limitations have been found, and different approaches to solution have been proposed.

Covariance matrix estimation effectiveness is also undermined by background local non-homogeneity. In fact, multiple background classes are likely to be included in the EW, thus violating background LMN and compromising the

The authors are with the Department of Information Engineering, University of Pisa, Pisa, Italy, (e-mail: {stefania.matteoli, m.diani, g.corsini}@iet.unipi.it).

detection process. This can be mitigated by considering global background statistics for covariance matrix computation. Different solutions can be mentioned, such as detection based on image segmentation [3-5] and “quasi local” [5] approaches to covariance matrix estimation. These approaches mainly rely upon global statistics, and may include different degrees of local information into the process.

When targets are placed in close proximity, target pixels (outliers) are likely to be included in the local background EW, thus causing target “contamination” effects. Covariance matrix estimation can be made less sensitive to the presence of outliers in the EW through appropriate methods aimed at excluding possible outlier pixels from covariance estimation [6].

As is evident, these critical aspects are deeply different in nature, like the situations from which they arise. As a consequence, in the literature, they have been typically dealt with and discussed within different frameworks, disregarding possible existing connections also when developing different approaches to solution. More importantly, the aforementioned critical aspects are likely to occur together in the same operating scenario, and all of them have to be taken into consideration when approaching anomaly detection, since they may be detrimental to detection performance. Therefore, an analysis of these limitations, also aimed at indentifying possible existing links, seems crucial in order to asses if existing algorithms, maybe designed ad-hoc to solve a specific problem, can handle also more complex situations.

As part of our research into this framework, in this paper, the aforementioned critical aspects along with their impact on the anomaly detection process are investigated, and the corresponding proposed solutions for improved covariance matrix estimation are discussed. It is worth mentioning that, due to the intrinsic difference among the problems investigated, this work does not attempt to cross-compare these techniques; rather, the behavior of each method is analyzed in terms of its capability in solving the problems for which it was designed and its ability in handling complex situations that may encompass also a combination of the aforementioned limitations.

In order to accomplish this task, real data from an airborne hyperspectral sensor are used in the analysis. Different cases of study are examined in order to illustrate and assess algorithms behavior: among these, a target “contaminated” by the presence of similar targets in its surroundings, and one placed in a strongly non-homogeneous local background. Performance is examined for each case of study, and results obtained by including improved covariance matrix estimation approaches into RX detection scheme are compared to standard RX detection. Algorithms limitations, along with future directions for refinement and improvement, are highlighted as well.

The paper is structured as follows: Section II introduces the problems involved with covariance matrix estimation and the proposed solutions. Methodology and experimental design are dealt with in Section III. Section IV reports experimental

results and discussion. A summary and conclusions are outlined in Section V.

## II. IMPROVED COVARIANCE MATRIX ESTIMATION FOR DETECTION APPLICATIONS

RX detector [1] is considered as the benchmark algorithm for detecting anomalies in multi/hyperspectral images. Its assumption is the Local Multivariate Normality (LMN) of background, reinforced through a spatial demeaning of the image. RX then computes the Mahalanobis distance between the pixel under test and the local background:

$$RX(\mathbf{x}) = \mathbf{x}^T \cdot \hat{\mathbf{C}}_b^{-1} \cdot \mathbf{x} \quad (1)$$

where  $\mathbf{x}$  is the  $L$ -dimensional vector representing the pixel under test, and  $\mathbf{C}_b$  is the  $L \times L$  background covariance matrix, being  $L$  the number of spectral channels.  $\mathbf{C}_b$  is typically estimated from a set of pixels (Estimation Window, EW) in the neighborhood of the pixel under test. The choice of the EW size is not trivial, as it will appear evident by reading the following subsections.

RX, along with its many improved/modified versions, has been widely explored, and the main concerns identified are related to  $\mathbf{C}_b$  estimation and subsequent inversion. These critical aspects are here examined, along with the corresponding solutions proposed within the anomaly detection framework. It is worth mentioning, as it will appear clear in the following sections, that these critical aspects are intrinsically different in nature, and so are the solutions proposed to cope with them.

### A. Small-sample size

In order to have a reliable  $\mathbf{C}_b$  estimation, a sufficiently wide EW size (possibly capturing a number of samples about 10 times the number of spectral bands [7]) would be desirable. However, small EW sizes are typically adopted, so as not to violate the LMN assumption of RX. Therefore,  $\mathbf{C}_b$  has generally to be estimated from a small number of very high-dimensional samples, and the resulting estimate is badly-conditioned and unstable [2], and this strongly affects detection performance. In dealing with this, one can try to reduce data dimensionality. However, in target detection applications this is not always recommended, since targets (i.e. rare vectors) are most likely to be addressed by the lower order (and therefore discarded) components identified by approaches such as the Principal Component Analysis (PCA) [8]. Alternatively,  $\hat{\mathbf{C}}_b$  can be regularized, thus stabilizing its inverse, by adding a scaled identity matrix  $\mathbf{I}$  before inversion, which suppress the influence of the lowest eigenvalues:

$$\hat{\mathbf{C}}_b^{REG} = (\hat{\mathbf{C}}_b + \beta \cdot \mathbf{I}) \quad (2)$$

This was shown to be equivalent to including a *regularization-term* in the design of the detector [9, 10], which results in:

$$REGULARIZED - RX(\mathbf{x}) = \mathbf{x}^T \cdot (\hat{\mathbf{C}}_b + \beta \cdot \mathbf{I})^{-1} \cdot \mathbf{x} \quad (3)$$

In classification applications, labeled training data are

typically available that allow for  $\beta$  selection through cross-validation/leave-one-out approaches. This availability is at the basis of all the work carried out for stabilizing  $\hat{\mathbf{C}}_b$  in classification settings [2, 11, 12]. In target detection, also due to the rarity of targets, labeled training data are generally not available, and  $\beta$  in equation (2) has to be estimated from the image itself. This parameter was shown to be related to sensor-noise level in the image [13], and to be reasonably estimated by identifying the flattest region of the sorted-covariance eigenvalues curve, e.g. by computing the *median* eigenvalue:  $\beta = \text{median}\{\lambda_i\}_{i=1}^L$ , where  $\{\lambda_i\}_{i=1}^L$  are the eigenvalues, and  $\beta$  is the cut-off eigenvalue.

### B. Local non-homogeneity of background

$\hat{\mathbf{C}}_b$  effectiveness is also undermined by background *local* non-homogeneity. In fact, given the constraint on the EW size (that has to be greater than the number of spectral bands in order to assure a non singular  $\hat{\mathbf{C}}_b$ ), multiple background classes are likely to be included in the EW, thus violating background LMN and compromising the detection process. This can be avoided by taking into consideration global background statistics.

One possible solution is clustering the image so as to include the global information concerning the scene, inferred by the resulting segmentation map, into the detection scheme:

$$\text{SEGMENTED-RX}(\mathbf{x}) = (\mathbf{x} - \hat{\boldsymbol{\mu}})^T \cdot \hat{\mathbf{C}}_b^{\text{SEG}^{-1}} \cdot (\mathbf{x} - \hat{\boldsymbol{\mu}}) \quad (4)$$

where  $\hat{\mathbf{C}}_b^{\text{SEG}}$  and  $\hat{\boldsymbol{\mu}}$  are the estimated covariance matrix and mean vector, respectively. These parameters can be estimated by accounting for different degrees of both local and global information, and several slightly different versions of this method exist [3-5], depending on how local and global information is integrated in the estimation procedure. For instance,  $\hat{\boldsymbol{\mu}}$  can either be the local mean vector estimated in the neighborhood of  $\mathbf{x}$  [4, 5] or be computed on the spectrally closest (in a Mahalanobis distance sense) cluster. As regards  $\mathbf{C}_b$  estimation, a pure global approach [3] employs the pixels of the cluster that is spectrally closest to  $\mathbf{x}$ , whereas more sophisticated *neighbor-guided* approaches [4, 5] include locality (e.g. segmentation labels of the neighboring pixels), in addition to global clustering information, in order to select the cluster(s) on which perform covariance estimation (e.g.  $\mathbf{C}_b$  can be either estimated over the cluster that is most common in the pixel neighborhood, or by computing a mixture of the covariance matrices of the clusters represented in the pixel neighborhood). This cluster-based approach is basically a *global* technique, into which different levels of local information can be incorporated; therefore it is expected to handle situations of even strong local background non-homogeneity.

Another approach, which makes use of both locality and global statistics of the image, is the one using the so-called

*Quasi Local* covariance matrix [5]:

$$\begin{cases} \text{QUASI LOCAL-RX}(\mathbf{x}) = \mathbf{x}^T \cdot \mathbf{E} \cdot \mathbf{V}_{\text{QL}} \cdot \mathbf{E}^T \cdot \mathbf{x} \\ \mathbf{V}_{\text{QL}} = \boldsymbol{\Lambda}_{\text{QL}}^{-1} = \text{diag} \left\{ v_i^{-1} \right\}_{i=1}^L, v_i = \left\{ \max(\lambda_i, \sigma_{\text{LOC}i}^2) \right\} \end{cases} \quad (5)$$

where  $\mathbf{E}$  is the matrix containing eigenvectors of global covariance matrix estimated on the whole image, and  $v_i$  is the largest of the global variance, given by the eigenvalue  $\lambda_i$ , and the local variance  $\sigma_{\text{LOC}i}^2$ , computed in the neighborhood of  $\mathbf{x}$ , as regards the  $i$ -th principal component. In this way, for each pixel the estimated  $\hat{\mathbf{C}}_b^{\text{QL}} = \mathbf{E} \cdot \mathbf{V}_{\text{QL}} \cdot \mathbf{E}^T$  is based both on global behavior and local characteristics of background. Thanks to the selection of the largest value between  $\lambda_i$  and  $\sigma_{\text{LOC}i}^2$ , the arising of false alarms around pixels characterized by abnormally small local variance is avoided.

### C. Target contamination by adjacent targets

Another critical aspect involved in  $\mathbf{C}_b$  estimation is the inclusion of anomalous pixels in the EW, which can occur when targets are placed in close proximity, and smaller EW sizes cannot be adopted without incurring in the small-sample size problem. In such a situation, the estimated  $\hat{\mathbf{C}}_b$ , computed employing also target pixels, does not reflect the background distribution accurately, thus resulting in a *masking effect* that makes the detection process more vulnerable [6]. In [6], an approach based on the Minimum Covariance Determinant (MCD) method [14] was proposed to mitigate target contamination in RX detection. It consist in an iterative technique aimed at finding a subset of pixels, within the EW, characterized by the minimum hyperellipsoid volume (which is proportional to the square root of the covariance determinant [14]). These pixels, which probably do not contain outliers, are the ones to be used for robust covariance estimation.  $\hat{\mathbf{C}}_b^{\text{MCD}} = \hat{\mathbf{C}}_b|_{\text{MCDsubset}}$  should therefore result less sensitive to possible outliers within the EW. The resulting detector

$$\text{MCD-RX}(\mathbf{x}) = \mathbf{x}^T \cdot \hat{\mathbf{C}}_b^{\text{MCD}^{-1}} \cdot \mathbf{x} \quad (6)$$

was proven to be effective for detection of multiple anomalous pixels of the same kind in a simple 50 x 50 three-bands hyperspectral image [6], and a warning was given about its application to a higher number of bands.

## III. EXPERIMENTAL DESIGN

In this section, the hyperspectral image employed for testing is described, and the experimental set-up for the aforementioned improved covariance matrix estimation approaches are presented. Since the improved ways for  $\mathbf{C}_b$  estimation examined here are deeply different in nature, so as the problems they try to mitigate, this work does not attempt to cross-compare these techniques; rather, the behavior of each method is analyzed in terms of its capability in solving the problems for which it was designed.

### A. Data Set Description

The hyperspectral image employed in this work was acquired in May 2008 by the SIM-GA hyperspectral sensor,

designed and manufactured by Galileo Avionica. The image considered here refers to 512-bands data acquired in the the Visible-Near InfraRed (VNIR) range, with an average spectral sampling of 2 nm, and a nominal Ground Sampling Distance (GSD) of about 0.7m. The actual image used for testing refers to a portion of the whole flight-line and was processed by performing water-vapor absorption and noisy bands removal, and a spectral binning aimed at increasing Signal to Noise Ratio (SNR), thus resulting in a 89-bands by 131-samples by 212-lines data cube. During measurement campaign, target panels characterized by different size and material were placed in the scene, and a ground-truth targets map was constructed. The portion analyzed is interesting since it encompasses different cases of study to be examined in order to illustrate and analyze algorithm behavior. Among these, targets contaminated by the presence of targets of the same type in their surroundings, and a target placed in a strongly non-homogeneous background. In the former case, some target pixels mixed to background were extracted from the actual target location and implanted in its surroundings in order to generate a “contamination effect” effective for testing. In the latter, a panel was placed, during the acquisition campaign, in the lawn very close to a road. A true-color representation, with highlighted these cases of study, is reported in Fig. 1.

### B. Experimental set-up

For local algorithms, the EW is defined through a set of three parameters  $EW = d_g, d_m, d_c$ , where (i)  $d_g$  is the semi-dimension of the inner guard window aimed at excluding potential target pixels from local background characterization; (ii)  $d_m$  is the semi-dimension of the outer EW for local demeaning; (iii)  $d_c$  is the semi-dimension of the outer EW for  $C_b$  estimation. The number of samples for  $C_b$  estimation are therefore  $N_s = (2d_c + 1)^2 - (2d_g + 1)^2$ .

#### 1) Regularized-RX experiments

Experiments for regularization testing were carried out by using EW: 5, 6, 7, so as to reproduce a situation of strong small-sample size. In fact,  $C_b$  had to be estimated by using  $N_s = 104$  samples, resulting in a ratio of  $N_s/L = 1.17$ .  $\beta$  was estimated for the whole image as the median of the global covariance eigenvalues:  $\beta = \lambda_{45} = 1.92 \cdot 10^{-7}$ .

#### 2) Segmented-RX and Quasi Local-RX experiments

Segmentation was performed by means of SEM algorithm [15], and the resulting thematic map, made up of 3 clusters, was used for detection through equation (4). Specifically, given the simplicity of the scenario analyzed, not encompassing many classes, and since the targets of interest can be also considered global anomalies with respect to such background, the mere global application of equation (4) was adopted for convenience, i.e. mean vector and covariance matrix were estimated over the cluster closest to the pixel under test. A comparison with more sophisticated approaches aimed at including local information is subject of ongoing work. Segmented-RX results were compared to RX ones, obtained with different sizes for EW: (i) EW= 5,6,7; (ii) EW=



Fig. 1. True-color representation of the image analyzed. An EW is superimposed over each target. Targets enclosed by red, blue, and yellow EWs are contaminated by the presence of targets of the same type in their surroundings. The target encircled by the pink EW is placed in a non-homogeneous local background.

6,7,8; (iii) EW= 6,8,8; (iv) EW= 6,8,9. Results will be examined both globally and specifically for the target placed near to the road, which exemplifies a local non-homogeneity situation. In this case, EW= 5,6,7 is the largest EW size that allows road’s pixels not to be included in local background characterization, but, at the same time, it reproduces a strong small-sample size condition. By increasing EW size, more and more road’s pixels fall into EW, thus increasing local background non-homogeneity. It is worth mentioning that Segmented-RX is obviously expected to yield good results also as regards the contaminated targets, because of its intrinsic *global* nature.

The aforementioned EW configurations were also tested with Quasi Local-RX algorithm, whose results will be examined for all the targets. In fact, this approach takes advantage of both global statistics and locality, thus being promising, not only for handling local background non-homogeneity, but also to mitigate target contamination effects. Furthermore, this method should allow a small EW size to be adopted as well, since EW is used only for local variances computation, whereas eigenvector matrix, the one affected by small-sample size, is computed on the whole image.

#### 3) MCD-RX experiments

MCD algorithm for free-outliers subset selection within the EW was tested as proposed in [6]. As pointed out by the author, its application to a high number of bands image was found complex, not only due to the computational burden. Numerical problems were found, due to the high-dimensional samples, e.g. in singularity and determinant-convergence checks. Moreover, several aspects to be improved were found, such as algorithm initialization and cardinality of the subset, which could be refined on the basis of contextual information (such as percentage of maximum expected outliers captured by the EW), and which are subject of ongoing research. In addition, as said in section II C, testing of this algorithm in the literature is only at an early stage [6]. Nonetheless, embodying MCD algorithm within RX detection was found promising since it embeds potentialities to mitigate target contamination effects. Therefore, MCD-RX was tested, and compared to RX, on a slightly simplified case, that is on the first 5 principal components of the image analyzed, which was enough to show MCD-RX’s potential. Moreover, testing was performed adopting a EW= 6,8,12 ( $N_s = 456$ ), so as to obtain a stronger contamination effect for the central targets, which are the ones placed in close proximity. Results will be examined specifically for the contaminated targets.

#### IV. EXPERIMENTAL RESULTS

In this section criteria employed to analyze algorithm behavior and evaluate detection performance are presented, and experimental results are reported and discussed.

##### A. Evaluation Criteria

Detection performance are comparatively evaluated by means of the Receiver Operating Characteristics (ROC) curves, on the basis of the ground-truth targets map available. ROC curves plot the Fraction of Detected Target (FoDT) versus the False Alarm Rate (FAR), computed over the operating scenario analyzed. Here, *pixel-based* ROC are reported, i.e. FoDT is computed as the ratio of the number of target pixels properly detected to the total number of target pixels. In certain situations, measures of FAR at a given Detection rate (FAR@D) [16] are reported, since they can be more effective to highlight algorithm behavior and show detection performance on specific targets. Test statistics around targets are examined as well, in order to have also an immediate visual representation of algorithm effectiveness on specific cases of study.

##### B. Regularized-RX results

Results for regularization are examined with respect to overall detection performance, since detection of all the targets is jeopardized by the strong small-sample size situation due to the employment of a 104-samples EW. ROC curves, represented in Fig. 2, show that much better results were obtained by using regularization. In order to achieve FoDT=0.79, RX provides FAR=1.37·10<sup>-2</sup>, whereas its regularized version yields FAR=2.06·10<sup>-4</sup>. Test statistics around the targets are also shown, for both algorithms, in Fig. 5 (a-f). The better behavior of the Regularized-RX appears evident, from simple visual inspection, for all the targets, except for the three targets placed in close proximity, where only a moderate improvement is observable (Fig. 5 (e)). This latter is due to the target contamination, whose effect, even though not strong thanks to the small EW size adopted, affects the detection process, and which has to be solved resorting to others of the aforementioned approaches.

##### C. Segmented-RX and Quasi Local-RX results

Segmented-RX has been presented here as an effective method capable of efficiently handling local background non-homogeneity. Good results are also expected as regards target contamination effect. The ROC curves shown in Fig. 3. highlight the better performance provided by the segmentation approach as compared to RX performed with different EW sizes. This is particularly evident when focusing on the range of operating interest (FoDT>0.7 & FAR<10<sup>-2</sup>), where Segmented-RX yields much better detection performance than any implementation of RX. Fig. 5 (g-i) show test statistic portions around the targets. All the targets clearly emerge from the local background, even the contaminated ones and the one placed close to the road.

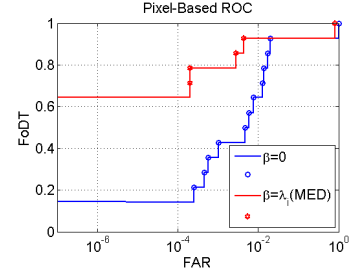


Fig. 2. Pixel-Based ROC for a 104-samples EW. The blue curve refers to the RX case ( $\beta=0$ , no regularization), whereas the red curve refers to Regularized-RX ( $\beta$  estimation through the median eigenvalue).

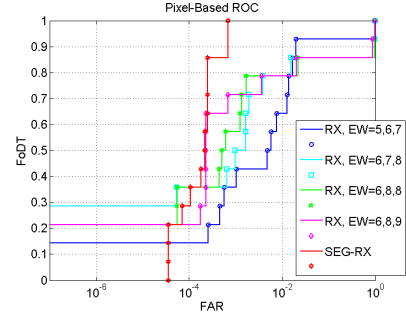


Fig. 3. Pixel-Based ROC for RX and Segmented-RX, compared by using different EW sizes for RX.

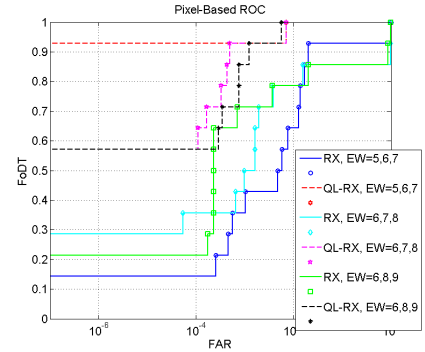


Fig. 4. Pixel-Based ROC for RX and Quasi Local-RX, compared using different EW sizes.

ROC for Quasi Local-RX are reported in Fig. 4. For all the EW configurations experimented, Quasi Local-RX results are better than the corresponding RX ones. This is especially true for EW= 5,6,7, for which the Quasi Local approach yields much better results than RX, and in general the best results for the Quasi Local approach, with FAR=0 for FoDT≤0.93. In fact, such a small EW allows the local background to be accurately characterized through those pixels that are physically closest to the pixel under investigation (and hence probably belonging to the same distribution) without incurring in small-sample size problems, since the eigenvector matrix is computed globally on the whole image. The corresponding test statistic portions around the targets are depicted in Fig. 5 (l-n), which undoubtedly show the high detection capability on all the targets, even, as expected though, on the contaminated ones.

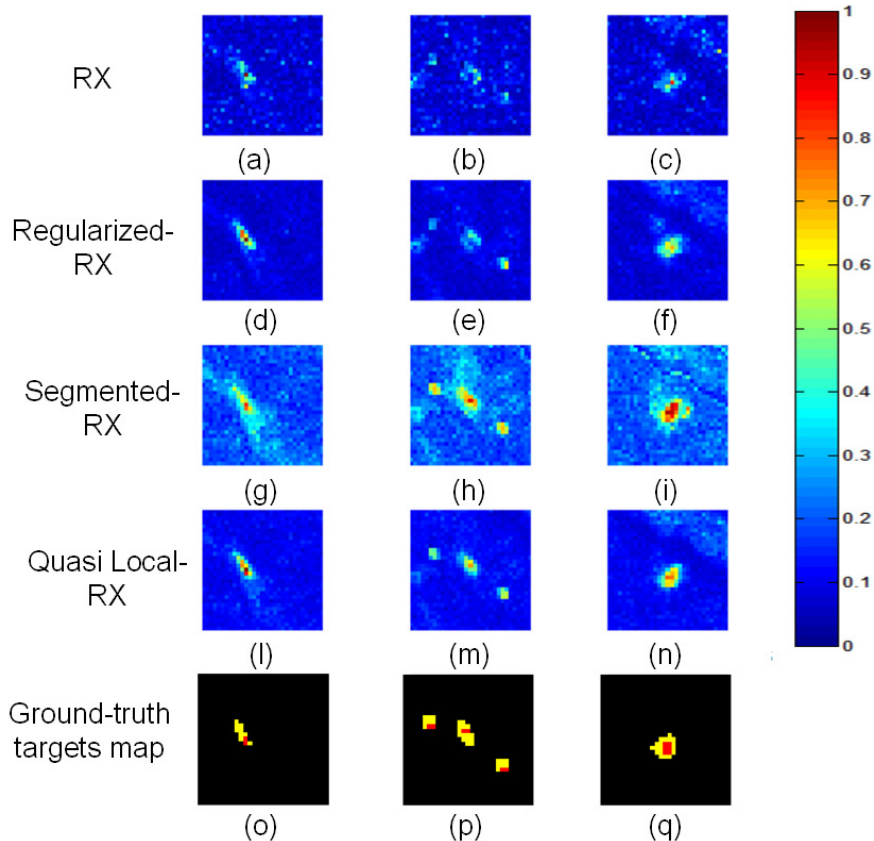


Fig. 5.  $30 \times 30$  test statistic portions centered around each set of targets, compared to ground-truth targets map. Each test statistic is normalized to its maximum value so as to obtain a fair visual comparison. Statistics reported are obtained with  $EW=5,6,7$  (local algorithms). (a, b, c) RX for left, center, and right targets; (d, e, f) Regularized-RX for left, center, and right targets; (g, h, i) Segmented-RX for left, center, and right targets; (l, m, n) Quasi Local-RX for left, center, and right targets; (o, p, q) ground-truth map for left, center, and right targets: core full-pixels are represented in red, whereas outer mixed pixels are highlighted in yellow. MCD-RX test statistic is not reported here, since it was obtained with  $EW=6, 8, 12$  so as to reproduce a situation of strong target contamination.

#### D. MCD-RX results

FAR@D measures for MCD-RX are reported in Table I. Results in terms of FAR@D allow the specific situation of the contaminated targets to be examined. As is evident, MCD-RX detection of the contaminated targets provides FAR@Ds lower ( $1\div 2$  orders of magnitude) than the ones yielded by RX, whereas no significant difference is appreciable for the other targets. MCD-RX's ability in mitigating target contamination effect is confirmed by looking at the test statistics around the contaminated targets in Fig. 6 (b-c). By adopting  $EW=6,8,12$ , RX totally fails in detecting the three targets, which are hardly perceptible in the test statistic, and easily mistaken with the surrounding background. On the other hand, these targets are clearly visible in MCD-RX test statistic, which exhibits higher scores in target locations. In order to have an insight into MCD-RX behavior, RX test statistic values (i.e. Mahalanobis distances) are plotted against MCD-RX ones (i.e. robust MCD distances) and shown in the scatter plot of Fig. 6 (a). As is evident, due to contamination, both background and target pixels exhibit similar Mahalanobis distances from the pixels captured by the EW, whereas measures of MCD distance are well distinguished for background and target pixels, and higher for these latter. It is worth mentioning that MCD-RX test statistics are not reported in Fig. 5, since they are obtained

with  $EW=6, 8, 12$  (which makes the contamination effect stronger), whereas all the test statistics in Fig. 5 refer to the  $EW=5, 6, 7$  configuration.

#### V. CONCLUSION AND FUTURE WORK

A work has been presented that examines critical aspects involved in covariance matrix estimation within the anomaly detection framework. Several limitations have been identified in the literature that arise from different situations. These limitations are deeply different in nature, like the situations from which they arise. However, all these critical aspects have to be accounted for, since they may occur together in reality, and they may strongly degrade anomaly detection performance. Since investigation into this framework has typically been performed disregarding possible existing connections, and approaches to solution have been developed independently, these limitations have been here investigated from an anomaly detection perspective, and the corresponding approaches to improved covariance matrix estimation have been analyzed on the same operating scenario by using real hyperspectral data that encompass interesting cases of study.

Regularized-RX was tested by employing a small EW size, so as to induce a strong small-sample size condition. Much better results have been obtained through regularization, as

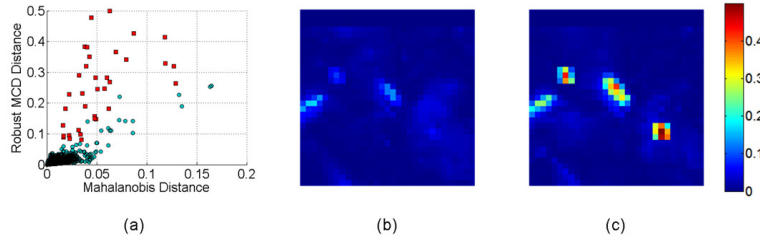


Fig. 6. MCD-RX performed with  $EW=6,8,12$ . (a) scatter plot of Mahalanobis distances (RX test statistic values) against Robust MCD distances (MCD-RX test statistic values). Background pixels are represented by the blue circles, whereas target pixels are denoted by red squares; (b-c) 30 x 30 test statistic portions around contaminated targets: (b) RX test statistic, (c) MCD-RX test statistic. Each test statistic is normalized to its maximum value as in Figure 5. However, in order to improve clarity of the images, contrast was enhanced by letting the colormap span only the values that are present in the 30 x 30 portions.

Table I. FAR@D computed for RX and MCD-RX. Results for contaminated targets are highlighted by the dashed box.

Algorithm	FAR@D				
MCD-RX	0	$7.70 \cdot 10^{-4}$	$6.42 \cdot 10^{-4}$	$3.21 \cdot 10^{-4}$	$2.10 \cdot 10^{-3}$
RX	0	$5.10 \cdot 10^{-3}$	$1.60 \cdot 10^{-3}$	$1.54 \cdot 10^{-2}$	$1.03 \cdot 10^{-3}$

compared to standard RX detection, even though detection of the contaminated targets did not show such a high improvement. This latter problem, in fact, can be mitigated either resorting to *ad-hoc* methods designed to neglect outliers in the EW, such as the MCD-RX, or by accounting for global statistics of the image.

The employment of global information has been shown also for approaches aimed at coping with local background non-homogeneity. Segmented-RX results have been compared to RX experiments carried out by using several different EW sizes. In the range of operating interest ( $FoDT > 0.7$  &  $FAR < 10^{-2}$ ), Segmented-RX has been shown to yield much better performance than any implementation of RX. As expected, Segmented-RX provided good results also in detecting the contaminated targets. The employment of more sophisticated segmentation-based approaches, which merge both global and local information, to be included in this analysis is still subject of ongoing work. As regards the Quasi Local-RX, it provided better results than RX, for any EW configurations tested. More specifically, the smallest EW size adopted has shown to yield the best performance, being able to accurately capture local background characteristics from the immediate neighboring of the pixel under test, without incurring in small-sample size problems.

Finally, MCD-RX has been shown to be a promising approach to handle contamination due to adjacent targets. Several aspects to be improved and refined were found though, which deserve further investigation, and which will be subject of future research. Therefore, MCD-RX was applied to a slightly simplified scenario, which has proven enough to show MCD potential in mitigating even strong target contamination effects, as the one obtained through the employment of a wide EW.

As already mentioned, due to the intrinsic difference among the problems investigated, this work does not attempt to provide a cross-comparison of these techniques. However,

considerations concerning how each algorithm can perform in practical applications, specifically as regards its actual utility in operating scenarios, can be made. As to Regularized-RX, the resulting algorithm is operationally easy to use also by non-experts, since it does not require complicate selection of free parameters (the regularization parameter has to be estimated from the image itself, procedure that can be embedded in the algorithm core) nor a specific supervision, and it can be easily applied in real-time. This is not the case of the Segmented-RX, which requires the segmentation of the image to be performed. Such a procedure makes a real-time application of this algorithm not feasible, unless accepting performance degradations due to on-line approximated applications of image clustering. Furthermore, the number of clusters has to be selected a-priori, and this strongly affects detection performance, and also requires operation with “man-in-the-loop”. On the other hand, when the clustering procedure is optimized, such an algorithm manages to handle both local background non-homogeneity and target contamination effects, and it does not suffer from the small-sample size, since, typically, thematic classes forming clusters are quite numerous. As regards the Quasi Local-RX, no free-parameters and no user interaction is required. However, its simple structure does not allow for a strong improvement margin. The MCD-RX has shown to need further investigation and refinements. Trying to improve the weak points found without increasing the computational burden will be a challenging task of future work.

Finally, this work has highlighted several interesting topics that will be subject of future research. Specifically, future work will be devoted not only to refinement/improvement of the various algorithms, but will also focus on analyzing more deeply the existing connections among these methods, also examining the possibility of jointly exploit the most interesting features of each algorithm in a single detection scheme, with the final aim of improving anomaly detection performance.

Lastly, it is worth noting that, although the analysis has been carried out focusing on anomaly detection, many of the outcomes of this research also apply to target detection algorithms that rely upon local background covariance matrix estimation.

## REFERENCES

- [1] I. S. Reed, X. Yu, "Adaptive Multiple-Band CFAR Detection of An Optical Pattern with Unknown Spectral Distribution," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 38, no. 10, pp. 1760-1770, 1990.
- [2] J. Hoffbeck, D. Landgrebe, "Covariance matrix estimation and classification with limited training data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 7, pp. 763-767, 1996.
- [3] M. J. Carlotto, "A Cluster-Based Approach for Detecting Man-Made Objects and Changes in Imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 2, pp. 374-387, 2005.
- [4] J. E. West, D. W. Messinger, J. R. Schott, "Comparative evaluation of background characterization techniques for hyperspectral unstructured matched filter target detection," *J. Applied Remote Sens.*, vol. 1, pp. 1-15, 2007.
- [5] C. E. Cafer, J. Silverman, O. Orthal, D. Antonelli, Y. Sharoni, S. R. Rotman, "Improved covariance matrices for point target detection in hyperspectral data," *Opt. Eng.*, vol. 47, no. 7, pp. 1-13, 2008.
- [6] E. Lo, J. Ingram, "Hyperspectral anomaly detection based on minimum generalized variance method," in *Proc. SPIE*, vol. 6966, pp. 696603-1-7, 2008.
- [7] J. A. Richards, X. Jia, *Remote Sensing Digital Image Processing*, Springer-Verlag, 1993.
- [8] S. Prasad, L. M. Bruce, "Limitations of principal component analysis for hyperspectral target recognition," *IEEE Geosci. Remote Sens. Letters*, vol. 5, no. 4, pp. 625-629, 2008.
- [9] N. M. Nasrabadi, "Regularized Spectral Matched Filter for Target Recognition in Hyperspectral Imagery," *IEEE Signal Proc. Letters*, vol. 15, pp. 317 - 320, 2008.
- [10] N. M. Nasrabadi, "Regularization for spectral matched filter and RX anomaly detector," in *Proc. SPIE*, vol. e 6966, pp. 1-12, 2008.
- [11] A. Berge, A. C. Jensen, A. S. Solberg, "Sparse Inverse Covariance Estimates for Hyperspectral Image Classification," vol. 45, no. 5, pp. 1399-1407, May 2007.
- [12] A. C. Jensen, A. Berge, A. H. S. Solberg, "Regression Approaches to Small Sample Inverse Covariance Matrix Estimation for Hyperspectral Image Classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 10, pp. 2814-2822, Oct. 2008.
- [13] A. F. Hayden, D. R. Twede, "Observations on the relationship between eigenvalues, instrument noise, and detection performance," in *Proc. SPIE*, Vol. 4816, pp. 355-362, 2002.
- [14] P. J. Rousseeuw, K. V. Driessen, K.V., "A fast algorithm for the minimum covariance determinant estimator," *Technometrics*, vol. 41, no. 3, pp. 212-223, 1999.
- [15] P. Masson, W. Pieczynski, "SEM Algorithm and Unsupervised Statistical Segmentation of Satellite Images", *IEEE Trans Geosci. Remote Sens.*, vol. 31, no.3, pp. 618-633, 1993.
- [16] P. Clare, M. Bernhardt, W. Oxford, S. Murphy, P. Godfree, V. Wilkinson, "A new approach to anomaly detection in hyperspectral images," in *Proc. SPIE*, vol. 5093, pp. 17-28, 2003.