Comparative evaluation of hyperspectral anomaly detectors in different types of background.

Dirk Borghys [a]

Ingebjorg Kasen[b]

Veronique Achard [c]

Christiaan Perneel [d]

[a]Royal Military Academy, Dept. CISS, Brussels, Belgium
[b]Norwegian Defence Research Establishment (FFI), Kjeller, Norway
[c]French Aerospace Laboratory (ONERA), Toulouse, France
[d]Royal Military Academy, Dept. of Math., Brussels, Belgium

ABSTRACT

Anomaly detection (AD) in hyperspectral data has received a lot of attention for various applications. The aim of anomaly detection is to detect pixels in the hyperspectral data cube whose spectra differ significantly from the background spectra. Many anomaly detectors have been proposed in literature. They differ by the way the background is characterized and by the method used for determining the difference between the current pixel and the background. The most well-known anomaly detector is the RX detector that calculates the Mahalanobis distance between the pixel under test (PUT) and the background. Global RX characterizes the background of the complete scene by a single multi-variate normal distribution. In many cases this model is not appropriate for describing the background. For that reason a variety of other anomaly detection methods have been developed. This paper examines three classes of anomaly detectors: sub-space methods, local methods and segmentation-based methods. Representative examples of each class are chosen and applied on a set of hyperspectral data with different backgrounds. The results are evaluated and compared.

Keywords: anomaly detection, sub-pixel anomalies, urban scenes, sub-space anomaly detection, local anomaly detection, segmentation-based anomaly detection

1. INTRODUCTION

Many types of anomaly detectors (ADs) have been proposed in literature.[1] The most frequently used anomaly detector is the Reed-Xiaoli (RX) detector[2] that is often used as a benchmark to which other methods are compared. The RX detector characterizes the background by its spectral mean μ_B and covariance matrix C_B. The actual detector calculates the Mahalanobis distance between the pixel under test r and the background:

$$D_{RX} = (r - \mu_B)^T C_B^{-1} (r - \mu_B) \tag{1}$$

The global RX detector characterizes the background of the complete scene by a single multi-variate normal distribution. In many scenes this model is not adequate. For that reason several variations of the global RX detector have been proposed in literature.[1,3–10] They can be sub-divided into three classes: sub-space methods, local methods and segmentation-based methods. In complex scenes, i.e. scenes with a large variability of background materials, the latter category has shown to be very effective and several segmentation-based anomaly detectors (SBAD), not necessarily based on RX, have recently been proposed.[11–15] The aim of the current paper is to compare the results obtained by different types of anomaly detectors in scenes characterized

Further author information: (Send correspondence to Dirk Borghys)
Dirk Borghys: E-mail: Dirk.Borghys@elec.rma.ac.be, Telephone: +32 2 742 6660

by different types of background. In particular two rural scenes with sub-pixel anomalies, a rural scene with some targets in shadow and an urban scene were considered. Representative examples of each of the three previously mentioned classes of anomaly detectors were included in the comparison. The evaluation is based on Receiver-Operator Characteristic (ROC) curves. For spatially fully resolved targets the probability of false alarms at first detection was also considered. For the urban scene, besides an objective evaluation, a more subjective evaluation is also given.

The rest of the paper is organized as follows: section 2 presents the used datasets, in section 3 the examined anomaly detection methods are briefly presented, section 4 presents the different pre-processing methods that have been applied to the data. The last two sections of the paper present the results and the conclusions.

2. OVERVIEW OF THE DATASET

The analysis was performed on a set of hypercubes of scenes with various backgrounds and representative of three scenarios:

- a rural environment with sub-pixel targets (CAM and OSLO)
- a rural environment with some of the targets in shadow (BJO)
- an urban environment (OSLO2)

Table 1 presents an overview of the used dataset. The first two datacubes are real hyperspectral images in which a matrix of anomalies was inserted artificially. Fig. 1 shows these images with full-pixel anomalies. The results shown in this paper were obtained with 10% mixing ratio sub-pixel anomalies for the CAM scene. For OSLO1 the mixing ratio was varied from 100% to 10%. The inserted anomalies are spectra of a green paint (CAM) and a green fabric (OSLO1).

The BJO image (fig. 2, left) was acquired over a natural scene with an agricultural region and a small forest near the village of Bjoerkelangen in Norway. Fourteen targets composed of different types of material and with different colors were laid in the scene during the image acquisition. Targets T3-T7 were in shadow. T3 in deep shadow between the trees, the four others in the shadow at the edge of the forest. Table 2) presents the dimensions and material types of the different targets.

The OSLO2 scene (fig. 2, right) is part of the center of Oslo. In this scene four targets (T1-T4) were laid out. Their respective dimensions in the image are: T1: 5×10, $T2 : 5 \times 9$, T3: 2×6 and T4: 6×7 pixels. Targets T2 and T3 are pieces of green fabric and the other two of a blue plastic. T1-T3 were laid out on the grass in a park, T4 was put on an asphalt background in the shadow from a building.

The images BJO, OSLO1 and OSLO2 were not geo-rectified before processing and all processing was applied to radiance data, i.e. without applying any atmospheric correction.

Name	Site	Sensor name	Nr of bands	Waveband (in μm)	Approx. spat. res.	Image size	Ntgts	Tot tgt size (in pixels)
CAM	Camargue (Fr)	Hymap	126	0.44-2.45	4 m	150×100	45	45
OSLO	Oslo (No)	HySpex	80	0.410-0.984	25 cm	286×287	81	81
BJO	Bjoerkelangen (No)	HySpex	80	0.410-0.984	25 cm	700×1600	14	574
OSLO2	Oslo (No)	HySpex	80	0.410-0.984	25 cm	700×1600	4	45

Table 1. Overview of the used dataset

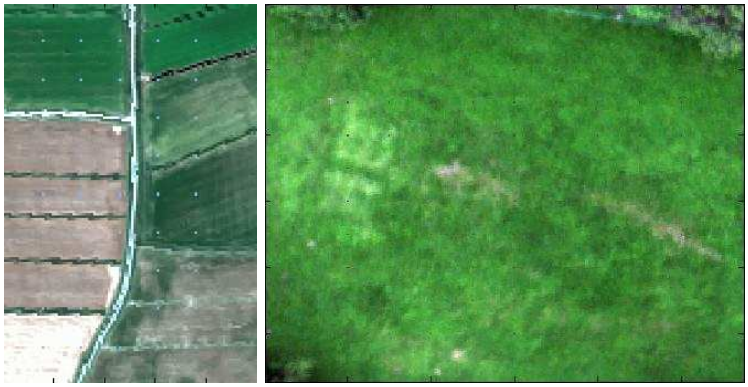

Figure 1. RGB color composite of the CAM (left) and OSLO1 (right) datacubes

Figure 2. RGB composite of the BJO (left) and OSLO2 (right) datacubes

T1	T2	T3	T4	T5	T6	T7
green car	red car	cloths	cloths	painted boards		
				paint 1	paint 2	paint 3
8x22	5x11	4x3	2x3	5x7	4x4	3x3

T8	T9	T10	T11	T12	T13	T14
cloths	cloths	painted boards				people
		paint 1	paint 2	paint 3	paint 4	
2x7	3x4	8x10	5x7	4x7	5x8	7x16

Table 2. Material types and target sizes (in pixels) for the BJO image

3. ANOMALY DETECTION METHODS

Besides global RX, representative examples of three categories of anomaly detectors are examined in the paper. Fig. 3 presents an overview of the detectors in the three classes. As can be seen from the figure, many of the investigated methods are RX-based, but for the sub-space detection methods and mainly for the SBAD methods, some anomaly detectors that are not related to RX have also been included. The different detectors are briefly described below.

3.1 Sub-space methods

The sub-space methods are global and have in common that they apply Principal Component Analysis (PCA) or Singular Value Decomposition (SVD) to the datacube and that detection is run on a subset of PCA/SVD bands. They are thus global anomaly detectors applied on a spectral sub-set (sub-space). For all of the sub-space methods the only parameter is the number of PCA or SVD bands that are considered to represent the

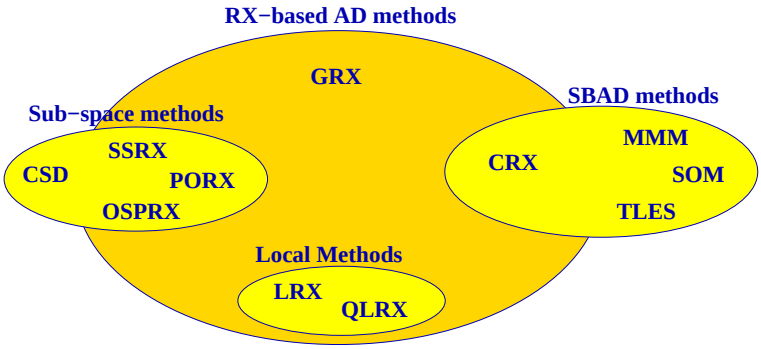

Figure 3. Schematic overview of the examined anomaly detection methods

background. If this number is too high, targets will disappear in the background, if it is too low, too many false alarms will remain. Finding the optimal value for the dimension of the background sub-space remains a current research topic.

3.1.1 Subspace RX (SSRX)

In SSRX the RX is applied on a limited number of PCA bands. The first PCs are discarded in SSRX.

3.1.2 RX after orthogonal subspace projection (OSPRX)

In OSPRX the background is defined by the first components of a Singular Value Decomposition. These first components define the background subspace and the data are projected onto the orthogonal subspace before applying the RX detector.[1,16]

In the current paper the SVD of the global spectral covariance matrix C is determined:

$$[U, S, V] = svd(C)$$

with U and V respectively the left and right singular vectors of the decomposition and S the diagonal matrix with decreasing singular values. The projection operator P_{SVD} is defined as a function of the first N_{bck} left singular vectors (columns of U), corresponding to the highest singular values (in S): $W = U(1..N_{Bands}, 1..N_{bck})$:

$$P_{SVD} = W(W^T W)^{-1} W^T$$

and the result of the detector is

$$d_{OSPRX}(r) = r^T(I - P_{SVD})r$$

with I the $N_{Bands} \times N_{Bands}$ identity matrix (N_{Bands}=number of spectral bands in the original image).

3.1.3 RX after "partialling out" the clutter subspace (PORX)

In this method the effect of the clutter in a pixel is removed (partialled out) component-wise by predicting each of its spectral components as a linear combination of its high-variance principal components. The detector applies the RX detector on the residual. Details of the method can be found in Ref. 9.

3.1.4 Complimentary subspace detector (CSD)

The CSD is not an RX-based method. In the CSD the highest variance principal components (PCs) are again used to define the background subspace and the other PCs to define the target subspace (the complimentary subspace).[6] The PUT is then projected on the two subspaces and the anomaly detector is the difference of the projection onto the target subspace and the background subspace.

$$d_{CSD}(r) = r^T P_t r - r^T P_b r$$

with:

$$P_b = W_{PCA}(1..N_{bck}) W^T_{PCA}(1..N_{bck})$$
$$P_t = W_{PCA}(N_{bck}..N_{tot}) W^T_{PCA}(N_{bck}..N_{tot})$$

3.2 Local methods

3.2.1 Local RX (LRX)

In LRX the covariance matrix C_B and mean μ_B spectrum of the background are estimated locally in a window around the PUT. A double sliding window is used: a guard window and an outer window are defined and the background statistics are determined using the pixels between the two (see fig. 4).

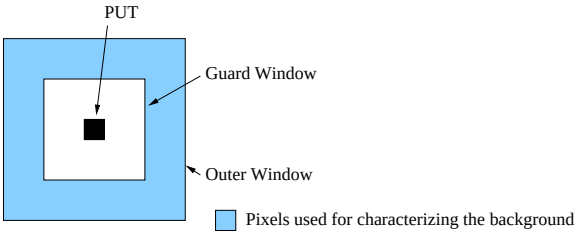

Figure 4. Sliding dual window used in the local AD methods

3.2.2 Quasi-local RX (QLRX)

Quasi-local RX (QLRX)[10] offers a compromise between the global and local RX approach. In QLRX the global covariance matrix C is decomposed using eigenvector/eigenvalue decomposition:

$$C = U\Lambda U^T$$

The eigenvectors are kept in the RX, but the eigenvalues are replaced by the maximum of the local variance and the global eigenvalue

$$\Lambda_{QL} = [max(\lambda^i_{loc}, \lambda^i_{glob})]$$

This means that the score of the detector will be lower at locations of the image with high variance (e.g. edges) than in more homogeneous areas. Spectral statistical standardization is applied as a pre-processing step. The local variance is determined in a double sliding window as in LRX.

3.3 Segmentation-based methods

In complex scenes, the hypothesis of a single multi-variate normal distribution of background spectra is normally not verified, not even locally. For that reason several segmentation-based anomaly detectors have been proposed in literature. In this paper four of these methods have been included in the analysis.

3.3.1 Class-conditional RX (CRX)

In CRX the image is first segmented, the covariance matrix and mean within each class i (i.e. $C_{B,i}$ and $\mu_{B,i}$) are determined. The Mahalanobis distance between the PUT and each of the classes is calculated. The final result is the minimum of these distances.

$$D_{CRX} = min_i \left[(r - \mu_{B,i})^T C^{-1}_{B,i} (r - \mu_{B,i}) \right]$$

In the current paper K-means clustering is used and the parameters of the method are the minimum number of pixels allowed in each class and the maximum number of classes used in the clustering. NC, the number of classes, follows from these parameters

3.3.2 Method based on Multi-variate normal Mixture Models (MMM)

A Stochastic Expectation Maximization (SEM) algorithm[17] is used for fitting a multi-variate normal mixture model to the image for describing the background. The anomaly detector detects pixels that have a low probability according to the estimated model. The MMM was applied after spectral binning[18] and square root transformation of the data in order to make the noise signal independent.[19] The parameters of the method are the number of mixture components and the termination threshold for the iterative parameters estimation method. The idea behind the mixture model is to use the mixture components (the multi-variate normal pdfs) as base pdfs in a row expansion of the true pdf, which can then, in principle, have any shape.

3.3.3 Two-Level End-member Selection method (TLES)

The principle of the TLES method[14] is the following: a small scanning window (50×50 pixels) runs over the image and at each position of the window the principal background spectra are determined using a segmentation method based on end-member selection. Endmembers that correspond to a minimum percentage (MP) of the image tile are stored. At the end of the process an endmember selection is again applied on the stored endmembers and linear unmixing is applied on the image. Anomalies correspond to pixels with a large residue after unmixing. In Ref. 14 N-FINDR was used as the endmember selection method. In the current paper the minimum volume simplex analysis algorithm (MVSA)[20] was used because it was found to give better results. Parameters of the method are the threshold MP and the number of end-members kept in the two stages of the algorithm.

3.3.4 Method based on Self-Organizing Maps (SOM)

A trained SOM is considered as a representation of the background classes in the scene. Anomalies are determined by computing the spectral distances of the pixels from the SOM units.[13] The SOM was applied on the first PCA components and run using a square map consisting of NsxNs hexagonal cells. It was optimized sequentially and its parameters are Ns and the number of PCA bands used.

4. PRE-PROCESSING METHODS

Before applying the actual anomaly detectors, some pre-processing methods were applied to the data. The pre-processing technique used depends on the dataset and on the anomaly detector. The different types of pre-processing are briefly described below.

4.1 Dimension reduction by spectral binning

The binning consists in averaging over groups of neighboring bands, down to a resolution of about 30 nm, as legitimated in Ref. 18 and references therein.

4.2 Kurtosis dimension reduction

As anomaly detection aims to search for outliers, a projection that enhances outliers applied as a pre-processing can improve detection performance. It has been shown that Kurtosis is very sensitive to outliers. In Ref.21 data are projected on the (first) eigenvectors of the Kurtosis matrix K:

$$K = \Sigma^{-1} \frac{1}{N} \sum_{i=1}^{N} (X_i - \mu)^t \Sigma^{-1} (X_i - \mu) (X_i - \mu) (X_i - \mu)^t$$

whereX_i is the i^{th} element of X, the matrix of observations, μ and $\Sigma = cov(X)$ are the mean and covariance matrix of X.

4.3 Spectral whitening

If the eigenvalues and eigenvectors of the covariance matrix of the complete image are resp. Λ and U and μ is the average spectral vector of the image, then the spectral whitening of the pixel r is given by:[6]

$$r_W = U\Lambda^{-1}U^T(r - \mu)$$

Whether whitening is beneficial for the anomaly detector depends on the AD method and the datacube. For CSD spectral whitening is always applied. The other sub-space methods were applied with and without whitening and the best result obtained is reported in this paper.

4.4 Spectral statistical standardization

The spectral statistical standardization converts each spectral band to have a zero mean and a standard deviation of 1. This is necessary for the QLRX in order to make sure the global eigenvalues and the local variances can be interchanged in the algorithm.

4.5 Shadow detection

In the case of images that exhibit large contrast between shadow and sunny parts, a classification step in two classes shadow/sunny can be inserted prior to the selection of spectral features and AD. The classification method applied here relies on the fact that radiance in NIR bands is very low in shadow, because there is no direct irradiance and the Rayleigh scattering effects that decrease with the wavelength power 4, induce low diffuse irradiance, too. Therefore, the histogram of NIR radiances of the image exhibits a first mode for low values that corresponds to shadow parts of the image. Fig. 5(left) shows the histogram in the NIR region, and on the right the detected shadows for the BJO dataset are shown. This shadow/sunny classification has been applied prior to GRX, the sub-space detectors and CRX for the BJO image and prior to GRX and CRX for the OSLO2 image. In the CRX method the shadow class counts as an additional class.

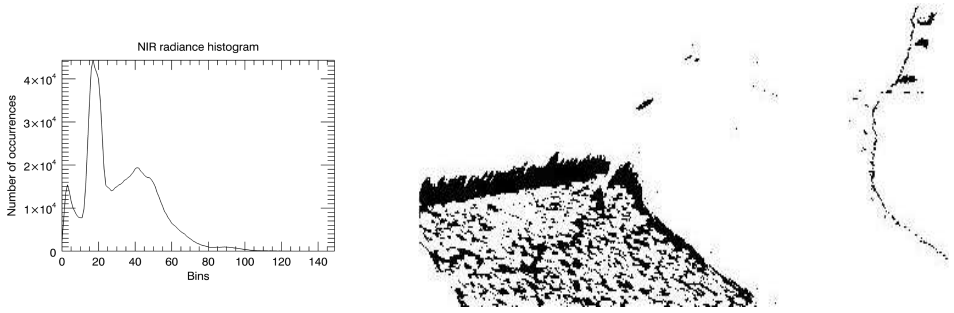

Figure 5. Histogram of NIR band and result of shadow detection for the BJO dataset

5. EVALUATION METHOD

Experimental ROC (Receiver Operating Characteristics) curves, showing the detection rate (DR) vs. the false alarm rate (FAR), are used to evaluate the results obtained with the various detectors. For the images with resolved targets, a pixel based ROC curve is calculated for each target, whereas for the images with sub-pixel targets, a ROC curve is calculated based on all the targets in the image. DR is plotted vs. the logarithm of the FAR (the resulting curve is referred to as a logROC), and the area under the logROC curve (the logAUC) is calculated and used as the measure of performance. The reason for using a logarithmic FAR scale is that it ensures equal weight across the range of FAR values.

For extended targets (in BJO and OSLO2), ROC curves give the detection performance for each pixel of the target with respect to the false alarms. For defense and security applications it is also of interest to assess the performance at the first detection of a target. In this paper, for extended targets, we therefore also determined the false alarm rate at the first detection for each of the targets.

6. RESULTS AND DISCUSSION

6.1 Implementation issues

The different examined AD methods depend on different parameters. Because the aim of the paper is to compare the different algorithms in the different scenes, the parameters are optimized for each detector/scene combination. For each algorithm, parameters were varied within a reasonable range and the results shown are the best results obtained for the algorithm.

For the different algorithms, different pre-processing steps were applied. Table 3 presents an overview of the used pre-processing method for each detector.

Moreover for the rural scene with targets in shadow (BJO), results with and without taking into account the shadow mask will be shown for GRX and all of the sub-space methods as well as for CRX, where the shadow mask is used as a supplementary class in the segmentation. For the urban scene (OSLO2), the shadow mask was applied only for GRX and CRX.

AD method	Applied pre-processing
GRX	Kurtosis Dimension Reduction
SSRX	Spectral Whitening (optional)
PORX	Spectral Whitening (optional)
OSPRX	Spectral Whithening (optional)
CSD	Spectral Whithening
LRX	Kurtosis Dimension Reduction
QLRX	Spectral Statistical Standardization
CRX	Kurtosis Dimension Reduction
TLES	None
SOM	None, but applied on first PCA bands
MMM	Spectral binning, square root transformation

Table 3. Overview of the type of pre-processing used for each detector

Sub-space methods		Local methods		SBAD methods	
Method	logAUC	Method	logAUC	Method	logAUC
GRX	0.732	LRX	**1.0**	CRX	**1.0**
SSRX	0.742	QLRX	0.723	MMM	**1.0**
PORX	0.569			SOM	0.116
OSPRX	**0.931**			TLES	0.129
CSD	0.868				

Table 4. logAUC results for the CAM scene with a mixing ratio of 10%

6.2 Results for sub-pixel detection in a rural environment

6.2.1 Results for CAM

Table 4 shows the logAUC results for the different detectors obtained in the CAM dataset with a 10% mixing ratio. The best results are obtained with LRX, CRX and MMM. They all achieve a logAUC of 1.0, which means that all targets have been detected with a false alarm rate that is smaller than 1/image size, i.e. $FAR < 6.6 * 10^{-5}$. Of the sub-space methods OSP-RX gives the best results.

6.2.2 Results for OSLO1

Fig. 6 shows the logAUC results of the different detectors versus the mixing ratio for the OSLO1 datacube. In this experiment the mixing ratio was varied from 100% (full pixel anomaly) to 10%. Results of global RX-based methods are shown as solid lines, the LRX and QLRX as dot-dash lines and the segmentation-based methods as dashed lines. CRX and LRX clearly give the overall best results. The next best results are obtained by MMM. Of the global RX-based methods the OSPRX and SSRX give the best results.

6.3 Results for a rural environment with some of the targets in shadow

Fig. 7(left) shows a graphical representation of the logAUC for each of the detectors and for each target for the BJO scenes. The colors represent the value of the logAUC. The figure on the right shows the logarithm of the probability of false alarms at first detection (logFARAt1stDet) for each of the targets. The colormap in both figures is such that red corresponds to the best performance. The two figures are globally consistent. From the figures it is clear that T2 is the most easily detectable target. It is also clear that the targets in shadow are more difficult to detect than the others. T3, hidden in the forest, is the most difficult to detect.

For the targets in shadow MMM gives the overall best results, followed by GRX, SSRX and CRX, all three after shadow masking (GRX_SM, SSRX_SM and CRX_SM), and LRX. Taking into account the shadow mask improves the performance of CSD and PORX slightly. For OSPRX the shadow mask increases the FAR.

For the targets that are not in shadow areas the MMM and CRX give the globally best results. For these targets, applying the shadow mask to the GRX, CRX and sub-space methods in general degrades the results by

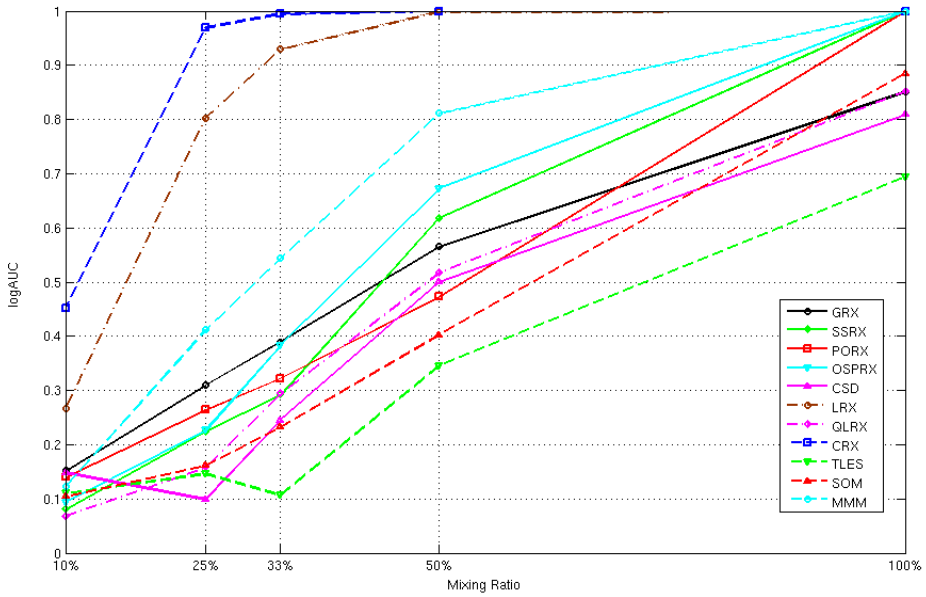

Figure 6. LogAUC vs. mixing ratio for the different detectors in OSLO1

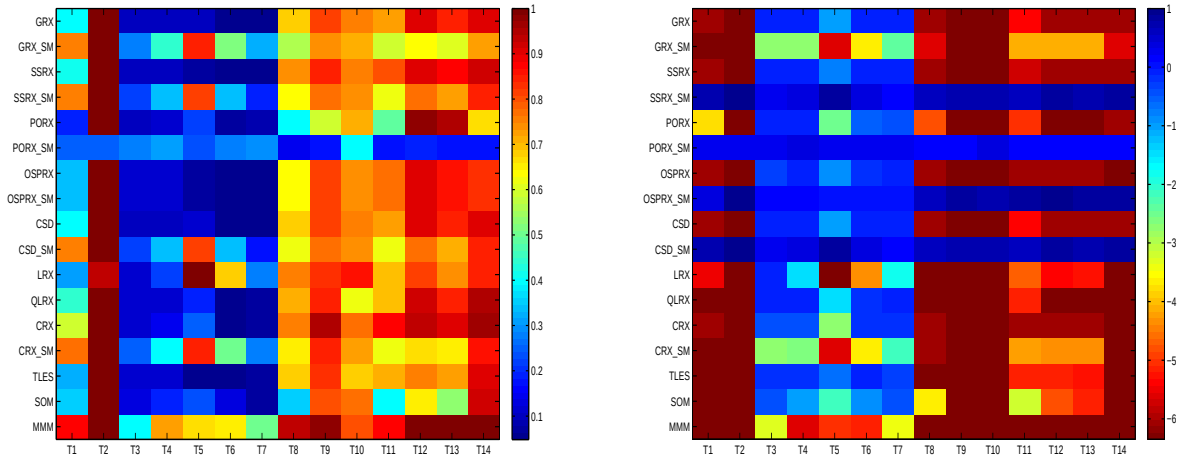

Figure 7. Graphical representation of detection results for BJO for each detector. Left: per target logAUC, right: per target FAR at first detection

increasing the FAR. The only exception is for GRX and CRX for target T1. MMM gives the most consistent results over all targets. Only for T5 some of the other methods and in particular LRX give better results than MMM. LRX also gives the best results for T10, which is painted using the same paint as T5. The local methods in general suffer from the variability in target size.

6.4 Results for the urban scene

Fig. 8(left) shows the logAUC results for the OSLO2 scene. It can be seen that the values of logAUC are much lower than for the other data cubes. The maximum value obtained here is 0.54. This is due to the complexity of the scene: the targets inserted into the scene are not the only anomalies. In an urban environment many objects can present an anomalous spectrum, e.g. cars, special roof materials, etc. The comparison therefore

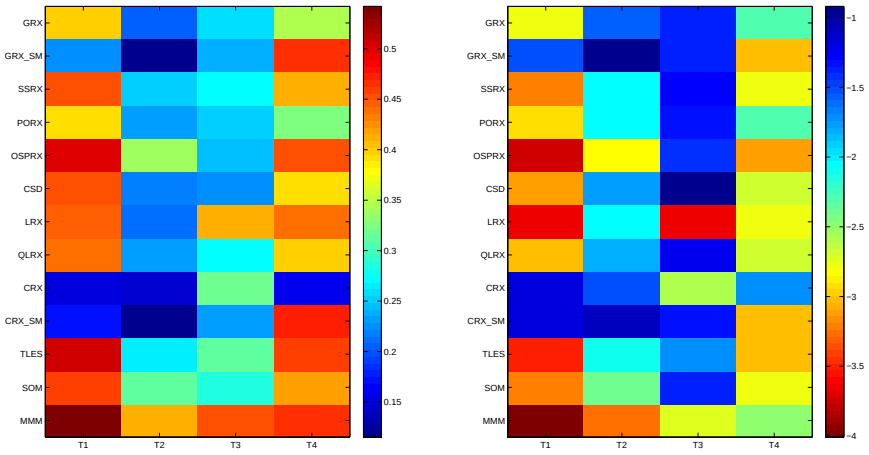

Figure 8. Graphical representation of detection results for OSLO2 for each detector. Left: per target logAUC, right: per target FAR at first detection

only shows how well the different anomaly detection methods cope with urban "clutter". Fig. 8(right) shows the logFARAt1stDet for each target. From the two figures it can be seen that MMM gives the overall best results, followed by TLES, OSPRX, LRX and SOM. From the global RX methods, OSPRX gives the best results. The shadow mask presents a benefit for T4, which is the only target in shadow. For this target GRX_SM and CRX_SM and OSPRX give the best results, GRX_SM and CRC_SM are slightly better than OSPRX for the logAUC criterion, while OSPRX gives slightly better results than the two others for the logFARAt1stDet.

As mentioned above an urban scene presents many objects that may have an anomalous spectrum and that thus will be considered a false alarm in the above evaluation. It is therefore of interest to give an idea of these false alarms for each of the detectors. Fig. 9 presents the results of the three best detectors (MMM, OSPRX and TLES), superimposed on a grayscale image of the OSLO2 scene. The results shown are thresholded detection results with the threshold set to the lowest first detection level for the true targets (i.e. the threshold for which at least one pixel of the most difficult target is detected). The figure shows that target 1 is the most easy and that it was detected completely by the three detectors. T2 and T3 have been completely detected by MMM, while the two other detectors (at the selected thresholds) detect only a part of the interior. On the contrary T4 has been completely detected by TLES and OSPRX, while MMM only detects a part of its interior. The figure also shows that many of the "false alarms" are indeed cars and structures on the roofs of buildings. Many cars show up as being detected by OSPRX and TLES (cyan), some were only detected by MMM. The TLES also detects some vegetation as being anomalous. The scene contains one extra large rectangular structure on the pavement, marked "FA1" that was detected as anomalous by all three detectors. This subjective evaluation shows that the three detectors that perform best according to the "objective" evaluation, are quite complimentary. Examining further the properties of their results may lead to interesting ideas on fusion of anomaly detectors.

7. CONCLUSIONS

This paper evaluates the performance of anomaly detection methods in scenes with different backgrounds and types of targets: agricultural scenes with sub-pixel targets, an agricultural scene with some of the targets in shadow and an urban scene. Three classes of anomaly detectors were considered besides the global RX: sub-space methods, local methods and segmentation-based anomaly detection (SBAD) methods.

For sub-pixel anomaly detection in scenes of low complexity (rural and non shadow) , LRX, CRX give the best results, followed by MMM. From the investigated global RX-based methods OSPRX gives the best results.

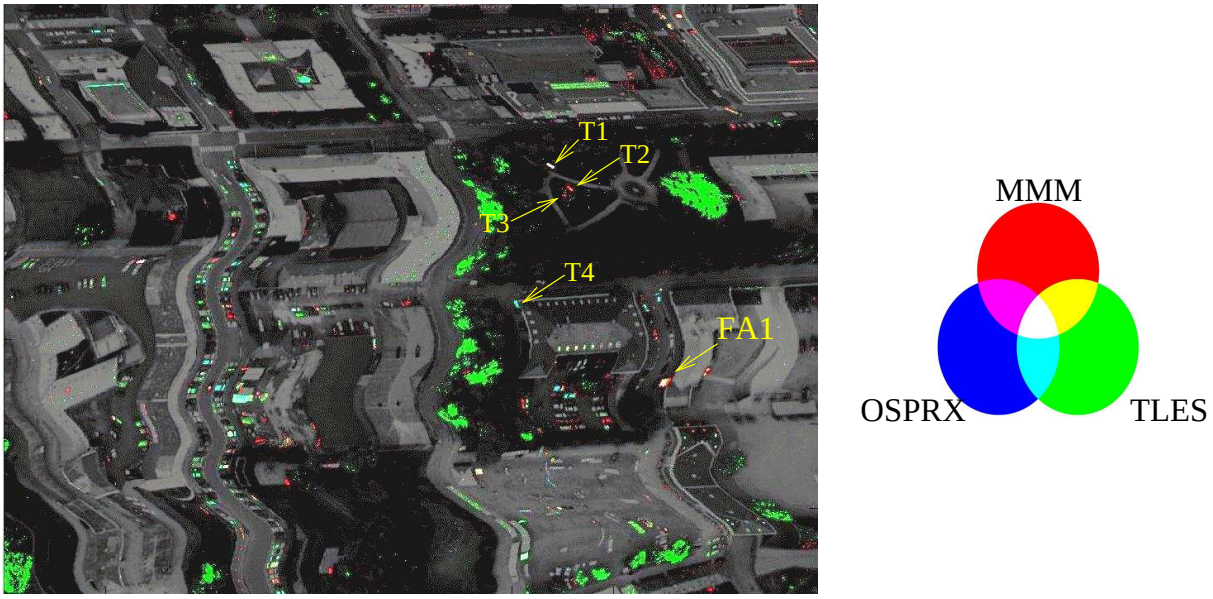

Figure 9. Color composite of the detection results for MMM, TLES and OSPRX at lowest 1st detection threshold, superimposed on B/W image of OSLO2 (R: MMM results, G: TLES results, B: OSPRX results)

For the rural scene with part of the targets in shadow, MMM gives the best overall results. The targets in shadow are also well detected by GRX and CRX after applying shadow mask. However, applying the shadow mask degrades performance of these two detectors for targets that are not in shadow.

In the urban environment the SBAD methods give generally better results than the RX-based methods. Of the global RX-based methods, OSPRX gives the best results. The overall best result for the urban scene is obtained by MMM, followed by TLES, LRX, OSPRX and SOM. Subjective evaluation of the detection results, shows that the best performing detectors give complimentary results and that "false alarms" are mainly due to objects with anomalous spectra in the scene such as cars and parts of buildings. Further investigation of this complementarity may lead to efficient detector fusion.

In this work it became evident that pre-processing has an important impact on the results. This will be further investigated.

8. ACKNOWLEDGMENTS

The authors would like to acknowledge Salvatore Resta at the University of Pisa for providing the ground truth for the OSLO2 data set. The OSLO1, OSLO2 and BJO datasets were provided by the Norwegian Defence Research Establishment (FFI). The CAM scene was provided by ESA (DAISEX experiment) through a collaboration with ONERA.

REFERENCES

[1] Matteoli, Diani, M., and Corsini, G., "A tutorial overview of anomaly detection in hyperspectral imagery," *IEEE A&S Systems Magazine, Part3: Tutorials* **21** (June 2010).

[2] Reed, I. and Yu, X., "Adaptive multiband cfar detection of an optical pattern with unknown spectral distribution," *IEEE ASSP* **38**, 1760–1770 (Oct 1990).

[3] Stein, D., Beaven, S., Hoff, L., Winter, E., Schaum, A., and Stocker, A., "Anomaly detection from hyperspectral imagery," *IEEE Signal Proc. Mag.* **38**, 58–69 (Jan 2002).

[4] Chang, C. and Chiang, S., "Anomaly detection and classification for hyperspectral imagery," *IEEE-TGRS* **40**, 1314–1325 (June 2002).

[5] Schaum, A., "Advanced methods for multivariate anomaly detection," in [*Proc. IEEE Aerospace Conference*], 1–7 (March 2007).

[6] Schaum, A., "Hyperspectral anomaly detection: Beyond RX," in [*Proc. SPIE Algorithms and Technologies for Multispectral, Hyperspectral and Ultraspectral Imagery XII*], **6565** (2007).

[7] Kwon, H. and Nasrabadi, N., "Kernel RX-algorithm: A nonlinear anomaly detector for hyperspectral imagery," *IEEE-TGRS* **43**, 388–397 (Feb 2005).

[8] Lo, E. and Ingram, J., "Hyperspectral anomaly detection based on minimum generalized variance method," in [*Proc. SPIE Algorithms and Technologies for Multispectral, Hyperspectral and Ultraspectral Imagery XV*], **6966** (2008).

[9] Lo, E. and Schaum, A., "A hyperspectral anomaly detector based on partialling out a clutter subspace," in [*Proc. SPIE Algorithms and Technologies for Multispectral, Hyperspectral and Ultraspectral Imagery XV*], **7334** (2009).

[10] Caefer, C., Silverman, J., Orthal, O., Antonelli, D., Sharoni, Y., and Rotman, S., "Improved covariance matrices for point target detection in hyperspectral data," *Optical Engineering* **47** (July 2008).

[11] Blumberg, D., Ohel, E., and Rotman, S., "Anomaly detection in noisy multi- and hyperspectral images of urban environments," in [*Proc. ISPRS 3rd URBAN Symp.)*], (March 2005).

[12] Duran, O. and Petrou, M., "A time-efficient method for anomaly detection in hyperspectral images," *IEEE-TGRS* **45**, 3894–3904 (Dec 2007).

[13] Duran, O. and Petrou, M., "Spectral unmixing with negative and superunity abundances for subpixel anomaly detection," *IEEE-GRSL* **6**, 152–156 (Jan 2009).

[14] Borghys, D., Truyen, E., Shimoni, M., and Perneel, C., "Anomaly detection in hyperspectral images of complex scenes.," in [*Proc. 29th Earsel Symposium*], (June 2009).

[15] Tarabalka, Y., Haavardsholm, T., Kasen, I., and Skauli, T., "Real-time anomaly detection in hyperspectral images using multivariate normal mixture models and gpu processing," *Int. J. on Real-Time Proc.* **4**, 287–300 (2009).

[16] Chang, C., "Orthogonal subspace projection (OSP) revisited: A comprehensive study and analysis," *IEEE-TGRS* **43**, 502–518 (March 2005).

[17] Masson, P. and Pieczynski, W., "SEM algorithm and unsupervised segmentation of satellite images," *IEEE-TGRS* **31**, 618–633 (Mar 1993).

[18] Kasen, I., Rodningsby, A., Haavardsholm, T., and Skauli, T., "Band selection for hyperspectral target detection based on a multinormal mixture anomaly detection algorithm," in [*Proceedings SPIE*], **6966** (2008).

[19] Skauli, T., "Sensor noise informed representation of hyperspectral data, with benefits for image storage and processing," *Optics Express* **19**(14) (2011).

[20] Li, J. and Bioucas-Dias, J., "Minimum volume simplex: A fast algorithm to unmix hyperspectral data," in [*Proc. IEEE IGARSS*], **III**, 250–253 (Aug 2008).

[21] Pea, D., Prieto, F. J., and Viladomat, J., "Eigenvectors of a kurtosis matrix as interesting directions to reveal cluster structure," *J. Multivariate Analysis* **101**(9), 1995–2007 (2010).