

ANOMALY DETECTION IN COMPLEX ENVIRONMENTS: EVALUATION OF THE INTER- AND INTRA-METHOD CONSISTENCY

D. Borghys, E. Truyen, M. Shimoni

Royal Military Academy
Signal & Image Centre
Brussels, Belgium

C. Perneel

Royal Military Academy
Dept. of Appl. Mathematics
Brussels, Belgium

ABSTRACT

Many anomaly detection methods, depending on various parameters, have been proposed in literature. Given the diversity of available anomaly detectors, from an operational viewpoint it is interesting to determine an efficient strategy to find the best suited detector for a given application. This is not obvious, especially in scenes with a highly structured background. The work presented here proposes a generic approach to the problem by examining the following questions: How different are the results of the various anomaly detectors? Are the parameters influencing the results significantly? Are there classes of methods sufficiently similar so that one can test only one of each class and see which results are most adequate for a given application? What are the spectral/spatial characteristics of the differences between methods? Can one predict which detector will give the best results for a given application? The current paper tries to answer the first three questions by comparing results of different types of anomaly detectors applied to different complex (urban, industrial and harbor) scenes. The comparison is not in absolute terms because it does not rely on a priori ground truth. In stead the detectors are compared relative to one another, the aim being to evaluate the similarities between the performance of the detectors and the dependency of their results on the used parameters, i.e. the inter- and intra method consistency.

Index Terms— Anomaly detection, hyperspectral, clustering, segmentation

1. INTRODUCTION

Anomaly detection in hyperspectral data has received a lot of attention for various applications. The aim of anomaly detection is to detect pixels in the hyperspectral datacube whose spectra differ significantly from the background spectra. In anomaly detection, in contrast to target detection, no a priori knowledge about the target is assumed [1]. Anomaly detection methods in general estimate the spectra of the background (locally or globally) and then detect anomalies as pixels with a large spectral distance w.r.t. the determined background spectra. Many types of anomaly detectors have been

proposed in literature, each depending on several parameters. The Reed Xiaoli (RX) algorithm [2], which is the benchmark anomaly detector for hyperspectral imagery, models the local background by a multi-variate normal distribution. Sub-space detectors form a very interesting family of anomaly detectors. In [3] several linear and non-linear (kernelized) sub-space detectors are compared and the Kernel Principal Component Analysis (KPCA) [4] based detector was found to be very promising for the investigated scenes. All of these methods detect anomalies by considering the spectral difference between the current pixel and its immediate surroundings. An alternative approach for anomaly detection consists in applying a scene segmentation prior to the actual anomaly detection. Several types of segmentation-based anomaly detection (SBAD) methods have been proposed [5]-[8]. These methods seem promising for anomaly detection in complex environments [5, 8] because they estimate a set of background spectra globally over the image. In this paper the similarities between different anomaly detectors and the dependence of their results on their parameters. Three classes of detectors are examined. The first class is specifically designed for complex environments and is based on a two-level clustering scheme. The second class consists of global image segmentation-based methods. The third class is based on local statistics. From each class several detectors were selected for the comparison. The test dataset consists of six datacubes, acquired by four different airborne sensors, showing scenes of diverse complexity.

2. DATASET

The presented analysis was performed on a set of 6 hypercubes of scenes with various complexity, acquired by 4 different airborne sensors. Table 1 presents the main characteristics of the dataset. The first column is the name by which the scenes will be referred further in this paper.

3. ANOMALY DETECTION METHODS

The three classes of anomaly detectors examined in this paper are briefly described below. The main parameters of each

Name	Site	Sensor name	Nr of bands	Waveband (μm)	Spatial resolution (m)	Scene Description
OBP	Oberpfaffenhofen(Ge)	Hymap	126	0.44-2.45	4	Airfield with aerospace industry
NEU	Neugilching (Ge)	Hymap	126	0.44-2.45	4	Semi-urban village
ANT	Port of Antwerp (Be)	AHS160	70	0.455-12.8	2.5	Part of petro-chemical port
KMT	Kalmthout (Be)	AHS160	70	0.455-12.8	2.5	Small airfield in vegetated area
PAV	Pavia (It)	Rosis	102	0.430-0.834	1.3	City
BJO	Bjoerkelangen (No)	HySpex	80	0.410-0.984	0.20	Rural village

Table 1. Overview of the dataset

method are also given because they are used in the study of intra-method consistency (section 4).

3.1. Class 1: Two-level clustering methods

In these methods [8] local segmentation results are combined to find characteristic spectra of the background in the scene. The anomaly detector is then based on a distance classifier w.r.t. these spectra. The image is scanned with a small (50×50) window. At each position, the corresponding image tile is segmented using a clustering based on the spectral angle as distance measure. The mean spectra of only the clusters that represent more than a minimum percentage (MPCT) of the image tile are stored. At the end of the process the collected set of spectra is clustered again and the image is classified w.r.t. the found cluster centres, using a distance classifier. In this paper the spectral angle (TLCM-SA) and an extended binary encoding (TLCM-BE) [9] were used as distance classifiers. The parameters of this method are the number of classes determined in the image tiles (NC1), the number of spectra retained in the final clustering (NC2) and MPCT.

3.2. Class 2: Global segmentation-based methods

Two methods based on global image segmentation are included in the comparison. The first method was proposed by Blumberg [5] (BLUM). A principal component analysis (PCA) is performed on the complete hypercube. The first 2 PCA bands are used to construct a 2-D histogram, which is segmented by locating its local maxima and attributing a given percentage of background pixels to each local maximum [5, 10]. Parameters are the number of bins to construct the 2D histogram (NBIN) and the size of the structuring element (NBH) used in the determination of the local maxima [8]. The second method is based on a self-organizing map (SOM). A trained SOM is considered as a representation of the background classes in the scene. In [6] the resulting U-matrix is segmented based on its local minima. In the current paper anomalies are determined by computing the spectral distances of the pixels from the SOM units. The SOM

was applied on the first three PCA and run using a square map consisting of $N_s \times N_s$ hexagonal cells. It was optimized sequentially and its only parameter is N_s . Contrary to [6] no spatial sub-sampling was applied here.

3.3. Class 3: Local statistics-based methods

3.3.1. Local RX (LRX)

In LRX [2] the Mahalanobis distance is calculated between the current pixel and its local neighborhood defined by an outer window (OWS) and separated from the current pixel by a guard window (GWS). In the implementation singular value decomposition is used for calculating the pseudo-inverse of the covariance matrix and the low values in the diagonal matrix are set to zero (cut-off is 99% of the trace).

3.3.2. PCA and KPCA based sub-space detectors

In the PCA-based detector [3] the spectra of individual pixels are projected on the first PCA bands obtained using the spectra from the outer window. The same dual-window is used as with RX. The cut-off percentage for the PCA bands was 99%. The kernelized version of this method was also examined [3, 4]. In the current paper a Gaussian radial basisfunction is used as kernel function and an adequate value for the standard deviation of the Gaussian σ_g is determined in each pixel as a function of the estimated standard deviation of the spectral values in the outer window σ_{OW} . It was empirically determined as $\sigma_g = 100 \sigma_{OW}$.

3.3.3. Local histogram-based method (LHIS)

The Euclidean distances are determined between the spectra of the current pixel and the pixels in its immediate neighborhood [8]. If the current pixel is sufficiently different from its background, the histogram of these distances has a mean far from zero. The mean is used as the detection value. Like with RX, a dual window is used.

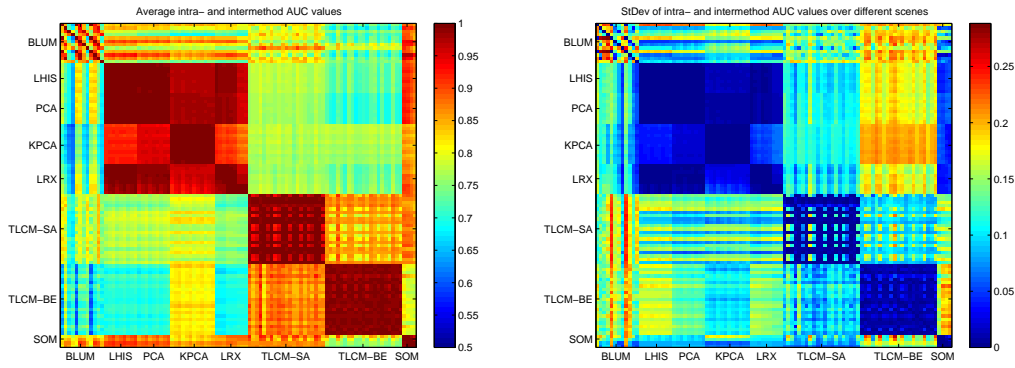


Fig. 1. AUC for all methods and parameters, averaged over the six scenes. Left: average AUC, right: standard deviation.

4. RESULTS AND DISCUSSION

4.1. Experimental setup

In order to examine the level of dependence of each method on its parameters, the detectors were run with different parameters and results compared. The same range of parameters for each anomaly detection method was applied to each of the datacubes because our main aim is to investigate the similarity between anomaly detectors and their behavior w.r.t. their parameters. If one is interested in detecting specific types of anomalies, parameters should be set accordingly and may depend on the image characteristics (e.g. on the spatial resolution). The sets of parameter values used in the analysis for the different detectors is presented in table 2. For instance for BLUM 3×4 settings of parameters were examined.

Method	Parameter	Parameter Range
BLUM	NBH	3,5,7
	NBIN	50,100,200,500
LRX,LHM	GWS	$3 \times 3, 5 \times 5, 7 \times 7$
	OWS	$9 \times 9, 11 \times 11, 13 \times 13$
TLCM-SA	MPCT	5,10
TLCM-BE	NC1	3,6,9
	NC2	6,10,20
SOM	Ns	5, 10, 20, 40

Table 2. Parameter sets used for the different detectors

4.2. Comparison method and analysis of results

The aim of this research is to compare the results of various detectors for different parameter sets, independently of a priori ground truth. Therefore the comparison will always be between one detector with a given set of parameters and the other detectors and sets of parameters. The 1% detection threshold for each of the detectors and for a given parame-

ter set is used to generate a "ground truth". The results of the other detectors and/or parameter sets are evaluated with respect to this ground truth. ROC (Receiver Operator Characteristic) curves are determined and the area under the curves (AUC) is used as comparison value. This analysis is made for the six different scenes. Each of the matrices was analyzed separately. The mean and standard deviation over all scenes was also determined and is shown in fig. 1. Individual lines and columns correspond to a single parameter setting for each of the methods. A line in the matrix shows the AUC obtained by the different methods and parameters when the results of the detector mentioned on the line (with a given parameter setting) is used as "ground truth". The columns represent the results of a detector when the other detectors serve to generate the ground truth. The left and right matrices in fig. 1 correspond resp. to the average similarity and the scene dependence of the different methods for each parameter setting.

In fig. 2 the AUC values were averaged over all parameters for each combination of detectors. This gives the average consistency within each detector w.r.t. its parameters and the average consistency between methods. The standard deviation matrix is also shown as an indication of the reliability of the conclusions. The matrices of the average AUC values

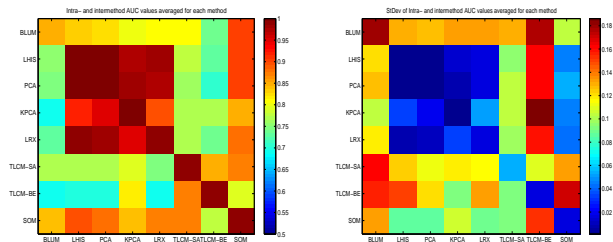


Fig. 2. AUC values for different detector combinations averaged over all scenes and all parameters. Left: average AUC, right: standard deviation.

are roughly symmetric. This means that the AUC of method x w.r.t. ground truth extracted by method y is similar to the AUC obtained for method y w.r.t. ground truth of method x .

For the examined parameter ranges, all methods, except the method of Blumberg exhibit a very high internal consistency. The minimum AUC is 0.985 ± 0.05 for all other methods when the parameters are varied (fig 2). For BLUM it is 0.85 ± 0.2 . The main parameter that influences its the results is NBIN. Although the TLMC methods show some variation with parameters (fig. 1), the average internal AUC is still very high (0.985 ± 0.05).

There are classes of methods that are globally very similar. The most similar are PCA, LHM and LRX: between PCA and LHM $AUC(PCA, LHM) : 0.995 \pm 0.003$ and $AUC(PCA, LHM, LRX) = 0.94 \pm 0.02$. KPCA is somewhat different to these methods, but these four methods are globally the most similar ($AUC = 0.89 \pm 0.05$). The two types of TLMC also give similar results ($AUC : 0.85 \pm 0.08$), but the SA method is more similar to other methods and mainly to the SOM ($AUC : 0.87 \pm 0.09$) than the BE method ($AUC : 0.79 \pm 0.17$). The correspondence between TLMC_BE and the other methods is very scene dependent. For most scenes the method of Blumberg gives globally similar results to the SOM method ($AUC : 0.86 \pm 0.13$). The high standard deviation is due to the ANT site with $AUC : 0.79 \pm 0.14$.

5. CONCLUSIONS

Results of 8 anomaly detectors were compared relatively to one another using 6 hypercubes of diverse complex environments, acquired by 4 different sensors. Three classes of detectors were considered. The first class (CL1) uses a two-level clustering method (TLMC) based on two types of spectral distances (SA and BE). CL2 is based on global scene segmentation and comprises the Blumberg method (BLUM) and a method using a SOM. The third class (CL3) is based on local statistics and consists of local-RX, a PCA- and KPCA-based sub-space detector and a method based on the local spectral distance. The results show most methods (except BLUM) to be have little variation w.r.t. their parameters in the considered parameter range. Results also show that the detectors within each of the classes give similar results, but different from results of the other classes. Most similar are methods in CL3. In CL2 results of BLUM, as well as its similarity to SOM, vary with its parameters and are also scene dependent. This should be examined further. The CL1 methods are less similar than the detectors in the other two classes. Other spectral distances in TLMC will be investigated. Further work will examine the spectral/spatial characteristics of the differences between detector classes.

6. ACKNOWLEDGMENTS

The OBP, NEU, ANT and KMT data acquisitions were sponsored by the Belgian Science Policy. For OBP and NEU the flights were operated by the German DLR and for ANT and KMT by the Spanish INTA. PAV data were provided by Prof Gamba of Pavia Univ. and acquired in the HySens campaign by DLR. Dr Skauli of the Norwegian Defense Research Institute (FFI) acquired and provided the BJO dataset.

7. REFERENCES

- [1] D.W.J. Stein, S.G Beaven, L.E. Hoff, E.M. Winter, A.P. Schaum, and A.D. Stocker, "Anomaly detection from hyperspectral imagery," *IEEE Signal Proc. Mag.*, vol. 38, pp. 58–69, Jan 2002.
- [2] I.S. Reed and X. Yu, "Adaptive multiband cfar detection of an optical pattern with unknown spectral distribution," *IEEE ASSP*, vol. 38, no. 10, pp. 1760–1770, Oct 1990.
- [3] H. Goldberg and N. Nasrabadi, "A comparative study of linear and nonlinear anomaly detectors for hyperspectral imagery," in *Algorithms and Techn. for Multi-, Hyper-, and Ultraspectral Imagery XIII*. 2007, vol. 6565, SPIE.
- [4] B. Scholkopf, A.J. Smola, and K.R. Muller, "Kernel principal component analysis," *Neural Computation*, vol. 10, pp. 1299–1319, 1999.
- [5] D.G. Blumberg, E. Ohel, and S.R. Rotman, "Anomaly detection in noisy multi and hyper spectral images of urban environments," in *Proc. ISPRS 3rd URBAN Symp.*, Tempe,AZ, US, March 2005.
- [6] O. Duran and M. Petrou, "A time-efficient method for anomaly detection in hyperspectral images," *IEEE-TGRS*, vol. 45, no. 12, pp. 3894–3904, Dec 2007.
- [7] M.J. Carlotto, "A cluster-based approach for detecting man-made objects and changes in imagery," *IEEE-TGRS*, vol. 43, no. 2, pp. 374–387, Feb 2005.
- [8] D. Borghys, E. Truyen, M. Shimoni, and C. Perneel, "Anomaly detection in hyperspectral images of complex scenes.," in *Proc. 29th Earsel Symposium*, Chania, June 2009.
- [9] A.S. Mazer, M. Martin, M. Lee, and J.E. Solomon, "Image processing software for imaging spectrometry data analysis," *Remote Sensing of Environment*, vol. 24, pp. 201 – 210, 1988.
- [10] J. Silverman, S.R. Rotman, and C.E. Cafer, "Segmentation of multi-dimensional infrared imagery from histograms," *Infrared Physics and Technology*, vol. 45, pp. 191–200, 2004.