ITERATIVE ANOMALY DETECTION

Yulei Wang[1,2,3], *Bai Xue*[1,4], *Lin Wang*[5], *Hsiao-Chi Li*[6], *Li-Chien Lee*[4], *Chunyan Yu*[1], *Meiping Song*[1,2], *Sen Li*[1] *and Chein-I Chang*[1]

[1]Information and Technology College, Dalian Maritime University, Dalian, China
[2] State Key Laboratory of Integrated Services Networks, Xian, China
[3] Key Laboratory of Spectral Imaging Technology, Chinese Academy of Sciences, Xian, China
[4]Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore County, MD USA
[5]School of Physics and Optoelectronic Engineering, Xidian University, Xian, China
[6]Department of Computer Science and Information Engineering, Fu Jen Catholic University, New Taipei City, 242, Taiwan, ROC

ABSTRACT

Anomaly detection (AD) is designed to find targets that are spectrally distinct from their surrounding neighborhood. Unfortunately, commonly used anomaly detectors generally do not take into account its surrounding spatial information. This paper derives an iterative version of anomaly detection, iterative anomaly detection (IAD) to address this issue. Its idea is to use a Gaussian filter to capture spatial information of the anomaly detection map and then feeds back the Gaussian filtered AD map to create a new data cube. The whole process is repeated over again in an iterative manner. When IAD is terminated anomaly representatives are identified and can be used as desired target signatures to implement constrain energy minimization (CEM) so as to classify all detected anomalies. Accordingly, IAD can be considered as anomaly classification.

Index Terms—Anomaly detection (AD), Iterative AD (IAD)

1. INTRODUCTION

Anomaly detection (AD) has received considerable interest in recent years due to advances of hyperspectral imaging technology which enables to uncover many unknown subtle targets that cannot be inspected visually or detected by prior knowledge. Since AD must be performed in a completely blind environment, two issues must be addressed, 1) how to find anomalies without using any prior knowledge and 2) how to distinguish one anomaly from another once anomalies are detected. As for the first issue many efforts have been reported including RXD developed by Reed and Xiaoli Yu in [1] which uses the sample global covariance matrix **K**, referred to RX detector (RXD) or K-AD along with its variants developed in [2], specifically RXD which uses the sample correlation matrix **R**, referred to as R-AD. However, all these anomaly detectors suffer from a common issue which requires an appropriate threshold value to segment anomalies from the background because AD-detection maps are generally real-valued. Regarding the second issue, a recent work on anomaly discrimination and categorization in [3,Chapter 15]. But it stops short on how to classify anomalies.

This paper expands AD to an iterative version of AD, called iterative anomaly detection (IAD) which implements AD in an iterative manner. Since the presence of an anomaly is closely rated to its surrounding neighborhood its spatial information is crucial to detection of its existence. The proposed IAD takes advantage of a spatial filter to gather spatial information for AD. Its idea is to apply a Gaussian filter to an AD-detection map whose spatial information can be captured by a specific Gaussian window. Then the Gaussian-filtered AD map is further fed back as a new hyperspectral band image to increase the current hyperspectral image cube by one dimension for the next round process of AD. In order to differentiate detected anomalies, an unsupervised target detection algorithm is further implemented in conjunction with AD during each iteration to find distinguish anomalies among detected anomalies. To terminate IAD an automatic stopping rule is introduced, in which case a set of anomaly representatives is generated and each representative is used to specify a particular anomaly class. These anomaly representatives are used as desired target signatures for a subpixel target detection technique, constrain energy minimization (CEM) in [2] to find real-valued anomaly classes. The final anomaly classification is then accomplished by Otsu's method [4] to produce anomaly classification maps.

2. ATGP

This section briefly reviews two major anomaly detectors commonly used in the literature. One is RXD. Since it takes advantage of global data sample covariance matrix **K** to on behalf of background suppression, it is referred to as K-AD, denoted by $\delta^{\text{K-AD}}(\mathbf{r})$ and specified by

$$\delta^{\text{K-AD}}(\mathbf{r}) = (\mathbf{r}-\boldsymbol{\mu})^T \mathbf{K}^{-1}(\mathbf{r}-\boldsymbol{\mu}) \quad (1)$$

where $\mathbf{r}$ is a vector of a data sample, $\boldsymbol{\mu}$ is the mean vector of data samples and $\mathbf{K}$ is the covariance matrix of global data

sample. Eq. (1) is actually the well-known Mahalanobis distance.

Another type of anomaly detector is developed in [1,8]. It is different from K-AD in the sense that the global sample data covariance matrix **K** in (1) is replaced by the global sample data correlation matrix **R**. It is defined by

$$\delta^{\text{R-AD}}(\mathbf{r}) = \mathbf{r}^T \mathbf{R}^{-1} \mathbf{r} \tag{2}$$

where $\delta^{\text{R-AD}}(\mathbf{r})$ is used to indicate **R** being used to differentiate $\delta^{\text{K-AD}}(\mathbf{r})$ used in (1).

3. ITERATIVE AD

Since anomalies are generally spatially correlated to their surrounding data sample vectors, it would be beneficial to include spatial contextual information into AD. Unfortunately, K-AD in (1) and R-AD in (2) do not take into account local spatial correlation. In order to capture such local spatial properties this section develops a new version of AD, to be called iterative AD (IAD) which includes Gaussian filters to smooth AD-generated detection maps and further feed the Gaussian-filtered detection maps back to form a new hyperspectral image cube for AD to be re-implemented over again in an iterative manner. The following algorithm details its ideas where AD can be either R-AD or K-AD.

Iterative Anomaly Detection

1. Initial condition:
   Let $\Omega^{(0)} = \{\mathbf{B}_l\}_{l=1}^{L}$ be the original hyperspectral image cube. Let $\{\mathbf{r}_i^{(0)} = (r_{i1}, \cdots, r_{iL})^T\}_{i=1}^{N}$ be data sample vectors in $\mathbf{\Omega}^{(0)}$. Let δ_0^{AD} be an AD (K-AD/R-AD) operating on $\mathbf{\Omega}^{(0)}$ and $k = 0$.
2. Let $k = k+1$. At the k^{th} iteration, use a Gaussian filter to blur $|\mathbf{B}|_{\text{AD}}^{(k)}$, where $|\mathbf{B}|_{\text{AD}}^{(k)}$ is the absolute value of the detection map produced by $\delta_{k\text{-}1}^{\text{AD}}$, $\mathbf{B}_{\text{AD}}^{(k)}$. The resulting image is denoted by Gaussian-filter $|\mathbf{B}|_{\text{GFAD}}^{(k)}$.
3. Form $\mathbf{\Omega}^{(k)} = \mathbf{\Omega}^{(k-1)} \bigcup \{|\mathbf{B}|_{\text{GFAD}}^{(k)}\}$. Find representatives for anomaly classes $\text{T}^{(k)}$ using the algorithms proposed in Section IV.
4. Check if $\text{T}^{(k)}$ satisfies a given stopping rule. If no, back to step 2. Otherwise, go to step 5.
5. $\text{T}^{(k)}$ is the data set with desired target signatures and IAD is terminated.

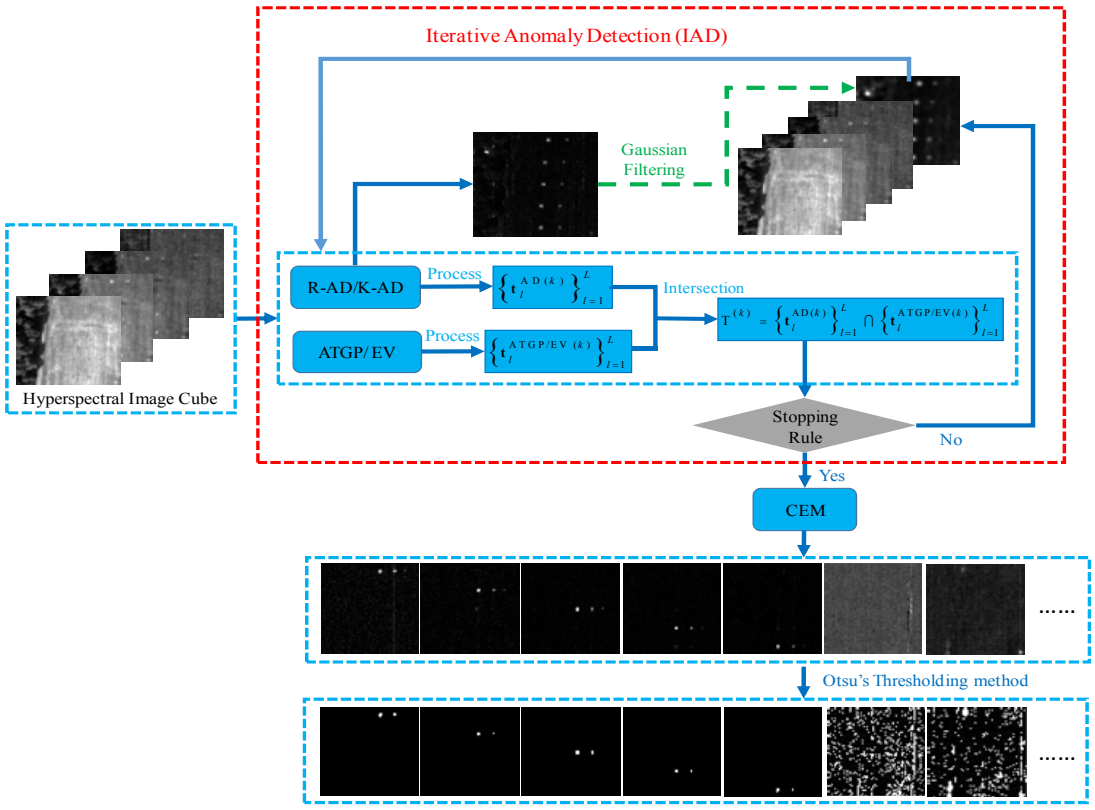


Figure 1. Graphic representation for illustration of target classification by TD-IAD

In order to terminate IAD it will require a stopping rule. This section derives an idea which takes advantage of ATGP for R-AD (or EV for K-AD) to find spectrally distinct anomalies from Gaussian-filter AD-detection maps iteratively by IAD until the number of such spectrally distinct anomalies converges. The details of step-by-step implementation of the stopping rule are described as follows.

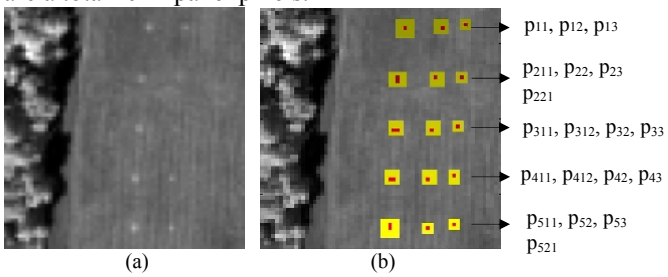
Stopping Rule for IAD

1. Use δ_k^{AD} in IAD to find L maximum gray-scales detection values in descending order in terms of vector length from the corresponded hyperspectral image $\mathbf{\Omega}^{(k)}$, i.e., $\| \mathbf{t}_1^{\mathrm{AD}(k)} \| \geq \| \mathbf{t}_2^{\mathrm{AD}(k)} \| \geq \cdots \geq \| \mathbf{t}_L^{\mathrm{AD}(k)} \|$
2. Apply ATGP for R-AD to $\mathbf{\Omega}^{(k)}$ to generate L target pixels, denoted by $\left\{ \mathbf{t}_l^{\mathrm{ATGP/EV}(k)} \right\}_{l=1}^{L}$.
3. Find $\mathrm{T}^{(k)} = \left\{ \mathbf{t}_l^{\mathrm{AD}(k)} \right\}_{l=1}^{L} \cap \left\{ \mathbf{t}_l^{\mathrm{ATGP/EV}(k)} \right\}_{l=1}^{L} = \left\{ \mathbf{t}_j^{\mathrm{A}(k)} \right\}_{j=1}^{\widetilde{L}}$ where $\widetilde{L} \leq L$ is the number of different classes of anomalies.

If $\mathrm{T}^{(k+1)} = \mathrm{T}^{(k)}$, then the algorithm is terminated. Otherwise, go to step 1.

4. REAL IMAGE EXPERIMENTS

The image scene shown in Fig. 2(a) is an airborne Hyperspectral Digital Imagery Collection Experiment (HYDICE) data with details available in [2-3]. There are 15 panels with three different sizes, $3m \times 3m$, $2m \times 2m$, $1m \times 1m$ with its ground truth provided in Fig. 2(b) where the center and boundary pixels of objects are highlighted by red and yellow respectively. In particular, R panel pixels are denoted by p_{ij} with rows indexed by $i = 1, \cdots, 5$ and columns indexed by $j = 1,2,3$ except that the panels in the 1st column with the 2nd, 3rd, 4th, 5th rows which are two-pixel panels, denoted by p_{211}, p_{221}, p_{311}, p_{312}, p_{411}, p_{412}, p_{511}, p_{521}. The 1.56m-spatial resolution of the image scene suggests that most of the 15 panels are one pixel in size. As a result, there are a total 19 R panel pixels.

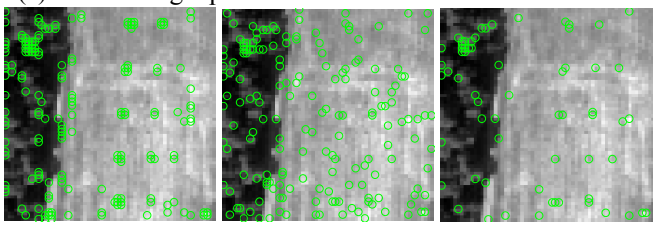


(a) (b)

Figure 2. (a) A HYDICE panel scene which contains 15 panels; (b) Ground truth map of spatial locations of the 15 panels

Fig. 2(b) shows the precise spatial locations of these 19 R panel pixels where red pixels (R pixels) are the panel center pixels and the pixels in yellow (Y pixels) are panel pixels mixed with the BKG.

Fig. 3(a-b) shows the results of the first iteration of R-IAD carried out by $\delta_0^{\mathrm{R-AD}}$ and ATGP where 169 anomalous pixels, $\left\{ \mathbf{t}_j^{\mathrm{R-AD}(1)} \right\}_{j=1}^{169}$ found in Fig. 3(a) by $\delta_0^{\mathrm{R-AD}}$, 169 ATGP-generated target pixels, $\left\{ \mathbf{t}_j^{\mathrm{ATGP}(1)} \right\}_{j=1}^{169}$ found in Fig. 3(b) respectively and their intersection, $\mathrm{T}^{(0)} = \left\{ \mathbf{t}_j^{\mathrm{A}(0)} \right\}_{j=1}^{64} = \left\{ \mathbf{t}_j^{\mathrm{R-AD}(0)} \right\}_{j=1}^{169} \cap \left\{ \mathbf{t}_j^{\mathrm{ATGP}(0)} \right\}_{j=1}^{169}$ shown in Fig. 3(c) with 64 target pixels $\widetilde{L} = 64$.



(a) $\left\{ \mathbf{t}_j^{\mathrm{R-AD}(0)} \right\}_{j=1}^{169}$ (b) $\left\{ \mathbf{t}_j^{\mathrm{ATGP}(0)} \right\}_{j=1}^{169}$ (c) $\left\{ \mathbf{t}_j^{\mathrm{A}(0)} \right\}_{j=1}^{64}$

Figure 3. Resulyts of the first iteration of IAD for HYDICE

Let $\mathbf{B}_{\mathrm{R-AD}}^{(1)}$ be the detection map produced by $\delta_0^{\mathrm{R-AD}}$ at the first iteration and the original hyperspectral image cube be denoted by $\mathbf{\Omega}^{(0)}$. A Gaussian filter with the window size of 11×11 and $\sigma = 0.1$ was applied to the absolute values of $\mathbf{B}_{\mathrm{R-AD}}^{(1)}$ in Fig. 3(a), $|\mathbf{B}|_{\mathrm{R-AD}}^{(1)}$ to produce a new Gaussian-filtered band image $|\mathbf{B}|_{\mathrm{GFR-AD}}^{(1)}$ which is further included into $\mathbf{\Omega}^{(1)}$ to create a new hyperspectral image cube $\Omega^{(1)} = \Omega^{(0)} \bigcup \left\{ |\mathbf{B}|_{\mathrm{GFR-AD}}^{(1)} \right\}$ for next iteration carried by $\delta_1^{\mathrm{R-AD}}$. The same process was repeated over again until it satisfied the stopping rule described in Section IV. The 19 target pixels found in Fig. 4 were further used as desired signatures for CEM to produce 19 real valued CEM-detection maps shown in Fig. 5. In order to perform target classification, the Otsu method was then used to segment targets out of background as shown in Fig. 6.

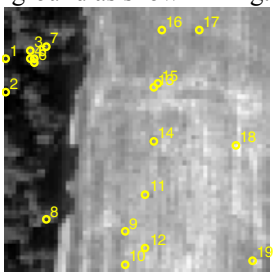


Figure 4. Final 19 targets locations for HYDICE using the iteration method

According to Fig. 6, it is very obvious that the panel pixels in five rows were classified where panel pixels in the 4th, 5th, 2nd, 3rd and 1st rows were classified in the 11th, 12th, 13th, 14th, 16th in separate and individual classification maps. It is interesting to note that the panel pixels in row 2 were classified in 13th and 15th classification maps because the 13th and 15th target pixels found by R-AD and ATGP were p_{221}, p_{212} in the same row, row 2. Similarly, the 16th and 17th target pixels were found to be p_{11} and p_{12} belonging to the same row 1. As a result, the 16th and 17th classification maps were used to classify panel pixels in row 1.

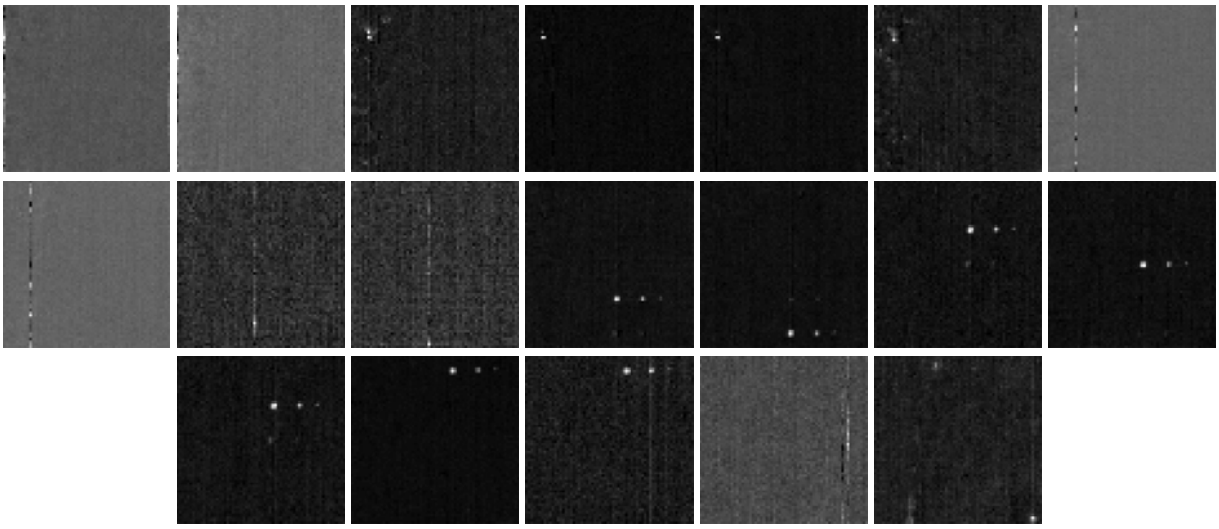


Figure 5. Original gray scale detection results of CEM algorithm using 19 signatures in Fig. 4.

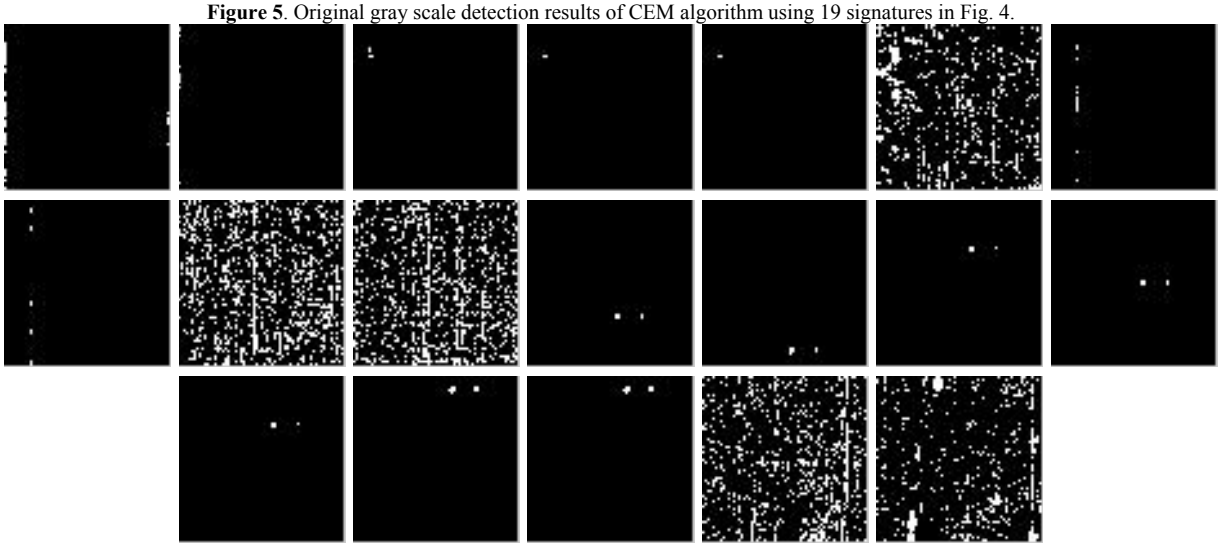


Figure 6. Otsu segmentation results for CEM results in Fig. 5

5. CONCLUSION

Anomaly detection (AD) has been widely studied in the literature, but how to use AD to perform target classification has not been explored. This paper derives an iterative version of AD (IAD) in conjunction with CEM, to classify targets of interest in an unsupervised manner. The development of IAD is to incorporate Gaussian-filtered spatial information iteratively via feedback loops, then uses ATGP to distinguish and find representatives among detected anomalies so that each representative specifies a different anomaly class. These anomaly representatives are further used by CEM as desired target signatures to perform target classification.

Acknowledgment

This work is supported by Fundamental Research Funds for Central Universities under Grants (3132016331, 3132017124 and 3132017080). The work of Y. Wang is supported by the Open Research Fund of Key Laboratory of Spectral Imaging Technology, Chinese Academy of Sciences (LSIT201707D). The work of L. Wang was supported by the 111 Project (B17035). The work of M. Song is supported by National Nature Science Foundation of China (61601077). The work of M. Song and Y. Wang are also supported by State Key Laboratory of Integrated Services Networks.

6. REFERENCES

[1] I.S. Reed and X. Yu, "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE Trans. on Acoustic, Speech and Signal Process.*, vol. 38, pp. 1760-1770, 1990.

[2] C.-I Chang, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*, Kluwer Academic Publishers, New York, 2003.

[3] C.-I Chang, *Real Time Progressive Image Processing: Endmember Finding and Anomaly Detection*, 2016.

[4] N. Otsu, “A threshold selection method from gray-level histgram,” *IEEE Trans on Systems, Man, and Cybernetics*, vol. SMC-9, no. 1, pp. 62-66 January, 1979