

Anomaly Detection Using Spectral Mismatch Between Anomaly Pattern and its Neighborhood

Denisova A.Yu.

Samara State Aerospace University
denisova_ay@geosamara.ru

Myasnikov V.V.

Samara State Aerospace University
vmyas@geosamara.ru

ABSTRACT

In this paper we present a novel algorithm for anomaly detection in multichannel images. Proposed algorithm uses spectral mismatch criterion to describe anomalous properties of small image regions. The idea behind the criterion is that the brightness of the anomalous region can't be represented as a function of pixels comprising that region. In our paper, we consider a local pattern of anomaly and its neighborhood, and we use a linear function to approximate the anomaly at each image position. In contrast to existing global and local RXD algorithms our approach allows more adaptive and noise resistant detection of anomalies. Experimental results are presented for hyperspectral remote sensing images.

Keywords

Hyperspectral images, anomaly detection, spectral mismatch criterion.

1. INTRODUCTION

Anomaly detection is one of the common tasks of digital image processing. Generally anomalies can be described as the image regions that do not correspond to normal behavior of some valuable image characteristics. The emergence of anomalous data may be caused from different reasons including some noise or registration errors, but for anomaly detection problem it is crucial to find such portions of data that correspond to real-world features or their parts that is not typical to the environ reflected on input image.

There are different definitions of term "anomaly" which are used in various applications and depend on particular data models. An comprehensive description of variety of anomaly detection tasks can be found in [Cha09a]. In this article we consider only the problem of anomaly detection for hyperspectral images. Hyperspectral images have hundreds of image channels that correspond to narrow spectral bands, that is why every pixel is presented as vector in a multidimensional space. In hyperspectral image analysis anomaly is usually considered a small image

region with spectral description sufficiently different from its neighborhood's.

One of the first anomaly detection algorithms, proposed in [Ree90a] by I.S. Reed and X.Yu, was RX-detector or RXD. Anomaly measure computed in RXD is Mahalanobis distance between current pixel vector and the average pixel vector of image. Thus anomaly is defined as a pixel which distance to the average value is the largest taking into account correlation between spectral channels of image. This algorithm demonstrates good results for images with simple single signature background, but for more complex background it is not effective. This fact and also possibility of defining the term "anomaly" in different way led up to many modifications of the RX-algorithm and other new algorithms. The examples of RXD modifications and some new algorithms are considered accordingly in [Sch07a], [Soo07a] and [Mes11a], [Ban06a], [Gu08a], [Bas07a].

In accordance with the classification of anomaly detection algorithms proposed in [Bor11a] and [Bor12a] all methods can be divided into three groups:

- subspace methods, that use global dimensionality reduction transformation for all image pixels. Usually principal component analysis or singular value decomposition is used;
- local algorithms that estimate background properties of each pixel neighborhood;

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

- algorithms with preliminary segmentation that aim to decompose image into regions with different background properties. Anomaly detection is performed inside these regions.

Depending on the specific task one or several of the aforementioned approaches can be used. Some algorithms may include RXD as the final processing step. This fact and along with continuous development of new algorithms, which cannot be classified into groups described above (for example, graph algorithms [Mes11a] or topological algorithms [Bas07a]), shows that such classification is very subjective.

The new algorithms proposed in this paper differ from others in definition of anomaly and image model exploited in it. They use local spatial pattern of the anomaly region and its neighborhood to incorporate assumptions about anomaly's size and form. The term "anomaly" mathematically is described by spectral mismatch criterion which is an error of anomaly candidate region approximation by its neighbourhood. In first spectral mismatch anomaly detection algorithm (SMAD) it is supposed, that image can be considered as stationary random field. The algorithm uses global spectral-spatial mismatch criterion. Because stationary random field model is used, coefficients of approximation of an anomaly-candidate region by its neighborhood are assumed to be the same for every analyzed fragment. An approximation error computed using such coefficients is the value of spectral mismatch criterion at each point and is the anomaly measure in this case.

In adaptive spectral mismatch anomaly detection algorithm (Adaptive SMAD) anomaly value is defined to be proportional to approximation error, when pixels of a potential anomaly are represented by pixels of its surroundings. Approximation coefficients are computed locally for every position of anomaly spatial pattern on the image. There is also a modification of the algorithm that employs pixel normalization.

Because both of the proposed algorithms use spectral mismatch criterion to measure anomaly of the region they can be grouped into class of spectral mismatch anomaly detection algorithms.

Proposed algorithms are compared with RXD (its global and local versions) and their superiority is shown.

2. SPECTRAL MISMATCH ANOMALY DETECTORS

Spectral mismatch algorithms compute anomaly value for each location of sliding window [Soi09a], that represent anomaly region and its neighborhood pattern.

Window is divided into two regions: interior region is interpreted as anomaly candidate and exterior region is interpreted as surroundings of potential anomaly (interior and exterior pixel sets do not intersect). Mentioned pair of pixel regions sequentially passes all possible positions on image (for example, in line-by-line scanning mode) and at each position with coordinates of central window pixel (n_1, n_2) a total "anomaly" value is computed. Total "anomaly" value for window is a result of aggregation of "anomaly" values for each pixel inside interior region. Note that aggregation can be made in different ways, for example, sum, minimum, maximum, median and so on.

Let us denote $I(n_1, n_2)$ – set of interior window pixels and $J(n_1, n_2)$ – set of exterior window pixels, where (n_1, n_2) is an image coordinate of window center, see Fig.1. The ordering of pixels within interior and exterior sets is not sufficient.

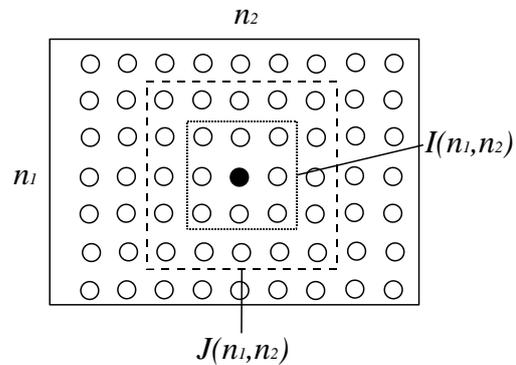


Figure 1. Interior and exterior pixel sets within processing window

Denote by $v_i, i \in I(n_1, n_2)$ and $v_j, j \in J(n_1, n_2)$ hyperspectral pixels from $I(n_1, n_2)$ and $J(n_1, n_2)$ set correspondingly.

Spectral mismatch value $\varepsilon_i^2(n_1, n_2)$ for interior pixel $v_i, i \in I(n_1, n_2)$ at window position (n_1, n_2) is defined as an error of representing interior pixel with the linear combination of pixels in exterior pixel set $J(n_1, n_2)$:

$$\varepsilon_i^2(n_1, n_2) = \left\| v_i - \sum_{j \in J(n_1, n_2)} \alpha_{ij}(n_1, n_2) v_j \right\|^2 \quad (1)$$

where $\|\cdot\|$ - some vector norm (in our case L_2 - norm), and $\alpha_{ij}(n_1, n_2)$ – coefficients of linear combination of exterior pixels that should be estimated from the image. Depending on the approach used to estimate these coefficients two algorithms can be considered.

In the first algorithm, spectral mismatch anomaly detector (SMAD), coefficients are supposed to be the same for all image. This assumption is equivalent to the following condition:

$$\alpha_{ij}(n_1, n_2) \equiv \alpha_{ij}, \quad i \in I(n_1, n_2), j \in J(n_1, n_2). \quad (2)$$

Coefficients defined in such way correspond to stationary random field image model. In this case maximum of an error Eq.1 is located at the points with sufficiently non stationary behavior.

In second algorithm, Adaptive SMAD, expression Eq.1 is used directly. It means that chosen pixel from interior set is represented as a linear combination of exterior pixels. If an error of such a representation is high, the pixel or region is interpreted as anomaly.

Below both algorithms are described and formulas for the coefficients are written.

SMAD

For spectral mismatch anomaly detector coefficients $\alpha_{ij}(n_1, n_2)$ are considered to be constant α_{ij} for all image. Their values are computed to achieve a minimum of square errors sum:

$$\varepsilon^2 = \sum_{(n_1, n_2)} \varepsilon^2(n_1, n_2), \quad (3)$$

where

$$\varepsilon^2(n_1, n_2) = \sum_{i \in I(n_1, n_2)} \left(\bar{v}_i - \sum_{j \in J(n_1, n_2)} \alpha_{ij} \bar{v}_j \right)^2. \quad (4)$$

Coefficients can be obtained as the solutions of the following system of linear algebraic equations:

$$\begin{aligned} \sum_{(n_1, n_2)} \bar{v}_i^T(n_1, n_2) \bar{v}_k(n_1, n_2) &= \\ &= \sum_{j=0}^{J-1} \alpha_{kj} \sum_{(n_1, n_2)} \bar{v}_i^T(n_1, n_2) \bar{v}_j(n_1, n_2) \end{aligned} \quad (5)$$

where $\bar{v}_k(n_1, n_2) \in I(n_1, n_2)$ and $k = 0, \dots, I-1$, $\bar{v}_t(n_1, n_2) \in J(n_1, n_2)$ and $t = 0, \dots, J-1$.

The coefficients α_{ij} in SMAD need to be computed only once since they are the same for each anomaly pattern position. That's why they globally define best linear approximation for all possible image regions according to required pattern. After coefficients α_{ij} have been obtained from Eq.5 for each pattern position "anomaly" value can be measured using Eq.4.

Adaptive SMAD

For Adaptive SMAD algorithm coefficients $\alpha_{ij}(n_1, n_2)$ must be different at every possible

window position (n_1, n_2) . These coefficients are found from orthogonal projection of chosen interior pixel vector \bar{v}_i into the space linearly spanned [Kos97a] by exterior region pixels. Let us denote this projection as $\hat{\bar{v}}_i$. Then error Eq.1 will look as follows:

$$\varepsilon_i^2(n_1, n_2) = \|\bar{v}_i\|^2 - \|\hat{\bar{v}}_i(n_1, n_2)\|^2, \quad i \in I(n_1, n_2), \quad (6)$$

where $\hat{\bar{v}}_i(n_1, n_2) = P_{\perp}^{(n_1, n_2)} \bar{v}_i$ is the projection of vector-pixel \bar{v}_i from set $I(n_1, n_2)$ on linear envelope of vectors from $\{\bar{v}_j\}_{j \in J(n_1, n_2)}$.

Projection operator $P_{\perp}^{(n_1, n_2)}$ is calculated to minimize mean square error of vector \bar{v}_i represented through the pixels from exterior set $J(n_1, n_2)$:

$$P_{\perp}^{(n_1, n_2)} = V(V^T V)^{-1} V^T \quad (7)$$

where $V = [\bar{v}_0 \bar{v}_1 \dots \bar{v}_j \dots \bar{v}_{J-1}]$ is matrix formed from pixels from exterior set $\bar{v}_j \in J(n_1, n_2)$ (to simplify formulae we will omit arguments (n_1, n_2) of a matrix below). It is evident that $\alpha = (V^T V)^{-1} V^T$. As pixels from set $J(n_1, n_2)$ can be linearly dependent among themselves, it is necessary to select a subset of linearly independent vectors or to provide projector regularization. In our work projector with regularization is used:

$$\hat{P}_{\perp}^{(n_1, n_2)} = V(V^T V + \beta I)^{-1} V^T, \quad (8)$$

where $\beta > 0$ is regularization parameter, I – identity matrix.

Total value of the spectral mismatch criterion at the current image point is evaluated as the following expression:

$$\varepsilon^2(n_1, n_2) = \sum_{i \in I(n_1, n_2)} \left(\bar{v}_i - \sum_{j \in J(n_1, n_2)} \alpha_{ij}(n_1, n_2) \bar{v}_j \right)^2. \quad (9)$$

where $\alpha_{ij}(n_1, n_2)$ are the representation coefficients for current pattern position.

An optional modification of Adaptive SMAD algorithm includes preliminary normalization of all image pixels to meet the following condition:

$$\|v_i(n_1, n_2)\| = 1. \quad (10)$$

In this case value of error Eq.6 can be written as follows:

$$\begin{aligned} \varepsilon_i^2(n_1, n_2) &= 1 - \cos^2(v_i, \hat{v}_i(n_1, n_2)) \\ &= \sin^2(v_i, \hat{v}_i(n_1, n_2)), \quad i \in I(n_1, n_2) \end{aligned} \quad (11)$$

where sine (or cosine) is calculated for angle between interior pixel v_i and its projection into linear subspace defined by exterior pixels. It is obvious, that error value Eq.11 unambiguously (and monotonously) depends on the specified angle.

3. EXPERIMENTAL RESEARCH

In experiments we used synthetic hyperspectral images. Images were size 256×256 pixels and 100 spectral channels corresponding to wavelengths ranging from 0.8 to 2.5 micrometer with step 0.017. Images were formed as linear combination of four "background" signatures (ACTINOLITE_AM3000, ILLITE_IL101, SEPIOLITE_SEP3101, BUDDINGTONITE_NHB2301) and two "anomaly" signatures (HEMATITE_FE2602, SIDERITE_COS2002) taken from IGCP-264 Library - CSES Beckman Spectrometer [Cla93a]. Coefficients for background and anomaly signatures were generated as stationary random fields with exponential correlation function.

Research was conducted on three synthetic images ("PIC-1", "PIC-2", "PIC-3") with correlation coefficients ρ 0.999, 0.98 and 0.45, respectively. At every image point sum of the coefficients of the linear combination was equal to one and coefficients were nonnegative. Test images were generated according to hyperspectral data linear mixture model described in [Cha02a], [Cha13a], [Cha07a]. Anomalies embedded into images were square plates with size 7×7 , 5×5 and 3×3 pixels. The examples of test images with built in anomalies are shown in Fig. 2. First two images were used to compare performance with global and local RXD without dimensionality reduction. First image illustrates situation with simple constant background, the second one has more complex background.

To compare algorithms the following experiment was done. At every test image channel additive independent zero-mean white noise with gauss distribution was added. Images with added noise were processed independently by two RXD modifications and proposed SMAD algorithm. Square window pattern of 5×5 pixel size was used with square interior region of 3×3 pixel size. The result of processing is shown in Fig. 3 and Fig. 4, signal to noise ratios for images were 1000, 100 and 10. Dark pixels correspond to high values of spectral

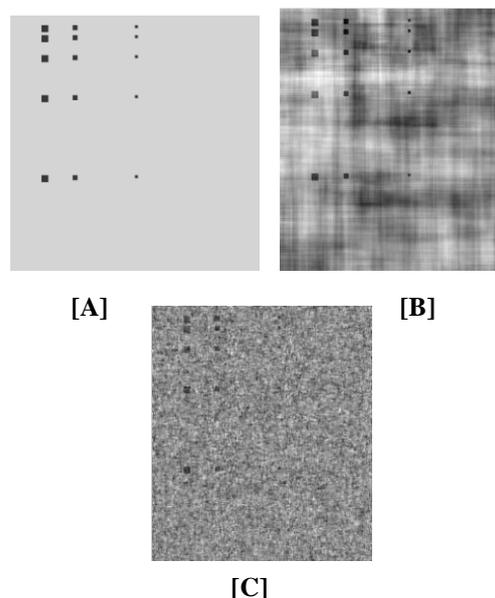


Figure 2. Examples of test images: [A] "PIC-1", [B] "PIC-2", [C] "PIC-3".

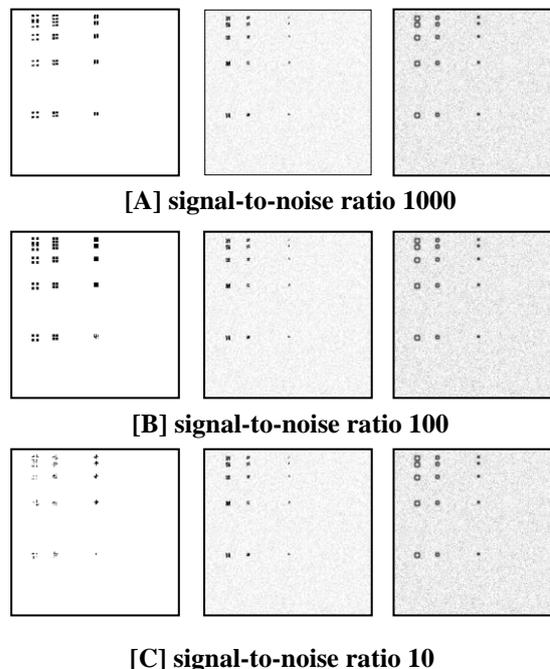


Figure 3. Experimental results for "PIC-1". From the left to right: SMAD, global RXD, local RXD (5×5 window size)

mismatch value and as consequence "anomaly" region.

As we can see from Fig.3 both algorithms performs well for simple background which is close to constant (correlation coefficient is 0.999). But it is required PCA transformation before RXD to avoid fluctuations arising from RXD processing of image "PIC-1". For the experiment shown on Fig. 3 the results of RXD algorithm and its modification were

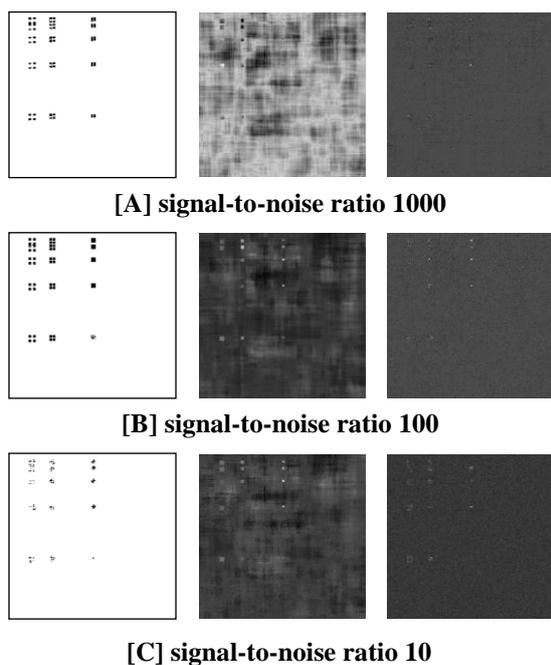


Figure 4. Experimental results for "PIC-2"
From the left to right: SMAD, global RXD, local RXD (5×5 window size).

obtained for first two principal components of PIC-1 image. As for SMAD algorithm it does not require preliminary PCA transformation.

It can be seen from the results shown in Fig. 4 that for more complex background with correlation coefficient 0.98 SMAD works significantly better than RXD. For this example RXD didn't detect any anomalies while SMAD marked all of them. Thus SMAD is very noise resistant and detects anomaly from complex background better than RXD (see the results for "PIC-2").

Experiment with "PIC-3" shows the influence of the parameter selection on SMAD result. The result of processing a square window pattern with square interior anomaly-candidate region using SMAD algorithm is shown in figure 5. Sizes of window and its interior region in pixels were, respectively, 5×5 and 3×3, 7×7 and 5×5, 9×9 and 7×7. So we can see that bigger anomaly size is detected better with larger window pattern. It should be noted that window size becomes more critical parameter for images with low correlation, for "PIC-3" correlation coefficient was 0.45.

The example of Adaptive SMAD detection for image "PIC-1" with signal-to-noise ratio 100 is shown in Fig. 6. Regularization parameter was set to $0,01\lambda_{\max}$, where λ_{\max} is the largest eigenvalue of matrix $V_{(n_1, n_2)}^T V_{(n_1, n_2)}$.

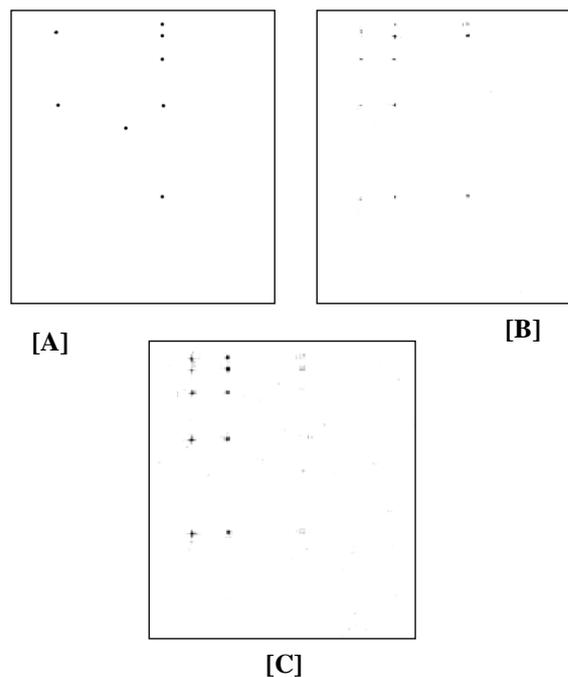


Figure 5. - SMAD results for "PIC-3" with window and interior region sizes respectively: [A] 5×5 and 3×3, [B] 7×7 and 5×5, [C] 9×9 and 7×7.

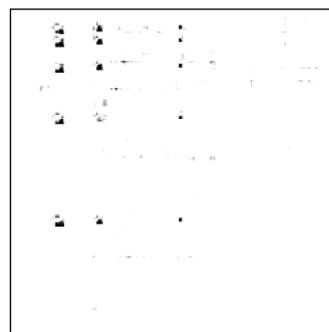


Figure 6. Adaptive SMAD result for "PIC-1". Window size and interior region size respectively 5×5 and 3×3.

Apparently, the Adaptive SMAD algorithm also yields significantly better results than RXD algorithm, although (unlike SMAD) it doesn't assume any model of the image. Adaptive SMAD has some disadvantages compared to SMAD algorithm. It is computationally expensive and generally the projection operator used in it is unstable and requires regularization. So the practical use of Adaptive SMAD algorithm has certain difficulties.

Figure 7 illustrates an example of using spectral mismatch algorithms and RXD modifications for real hyperspectral remote sensing image. We used AVIRIS Moffett field image, one of its spectral bands is shown in figure 7[A]. AVIRIS has about two

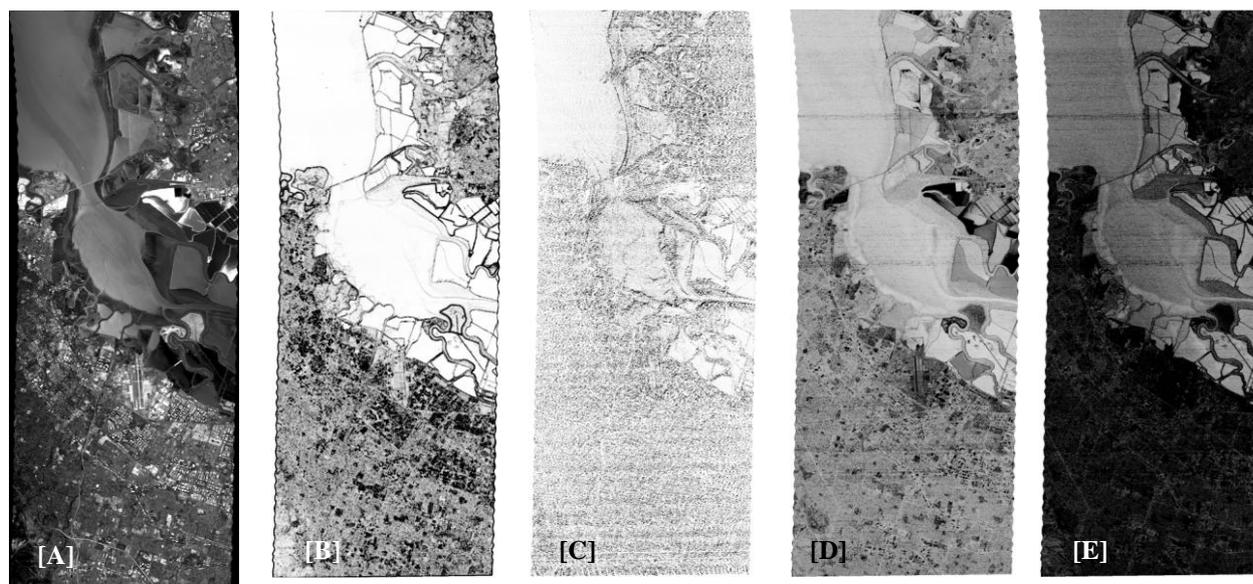


Figure 7. The results for AVIRIS Moffett field image for different anomaly detectors:[A] Original image, 550 nm band, [B] SMAD, [C] Adaptive SMAD, [D] global RXD, [E] local RXD (5×5 window)

hundred spectral channels from 400 to 2500 nanometers, some of this bands has significant noise. In our experiments we used all spectral bands of image including corrupted by noise bands. The same object in different spectral bands may look differently because of reflectance properties of its material. That is why in some spectral bands it may disappear or appears no contrast in some spectral bands. For example, in figure 7[A] bridge over the river has low contrast with water.

Fig.7[B]-[E] shows the results of all anomaly detection techniques, the darker pixels are more anomalous than the lighter. Spectral mismatch algorithms were used with 5×5 pixels square window with 3×3 interior region. Local RXD algorithm had window size 5×5. As we can see, proposed SMAD algorithm underlines borders of objects as anomalies. So the key characteristic for this algorithm is difference of spectral signatures between image objects. This fact allows algorithm to discriminate one image object from another or from background. Global RXD algorithm identified as anomalies objects which brightness was mostly different from the average brightness of the image. It is too weak condition, and we can see that only white in original image 7[A] objects were detected as anomalies by global RXD.

As for Adaptive SMAD algorithm, it demonstrates effective detection relief features in the river basin. Because of small anomaly pattern such objects as buildings were not detected as anomalies. Local RXD algorithm detected entire river bank as

anomalous region that seems to be incorrect or and it makes further analysis too difficult.

It should be noted that proposed algorithms are less affected by noise than RXD. For example, on both RXD images a noise stripe can be seen in upper part of image, this stripe is absent for spectral mismatch detectors results.

4. CONCLUSION

Two new algorithms were presented in the paper, namely, spectral spatial mismatch anomaly detector (SMAD) and adaptive spectral mismatch detector (Adaptive SMAD). Their performance was studied on synthetic and real hyperspectral remote sensing images. The results of experimental comparison with basic global and local RXD algorithms were presented and advantage of proposed methods was shown.

A short comparative analysis was also provided. particularly, it has been shown that SMAD is a noise resistant algorithm and allows confident detection of anomalies on images holding on stationary random field model even in case of low signal-to-noise ratio. Adaptive SMAD algorithm has no limitation due to the absence of any underlying image model but is more computationally expensive and requires regularization parameter selection. Proposed detectors were shown to be more effective than RXD.

5. ACKNOWLEDGEMENTS

This work was financially supported by the Russian Scientific Foundation (RSF), grant no. 14-31-00014

“Establishment of a Laboratory of Advanced Technology for Earth Remote Sensing”.

6. REFERENCES

- [Ban06a] Banerjee, A., Burlina, P. and Diehl, C. A support vector method for anomaly detection in hyperspectral imagery. *Geoscience and Remote Sensing, IEEE Transactions on*. V. 44(8), pp. 2282-2291, 2006.
- [Bas07a] Basener, D., Ientilucci, E. and Messinger, D. W. Anomaly detection using topology. *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XIII, SPIE*, V. 6565, 2007.
- [Bor11a] Borghys, D., Achard, V., Rotman, S.R., Gorelik, N., Perneel, C. and Schweicher, E. Hyperspectral anomaly detection: A comparative evaluation of methods. *General Assembly and Scientific Symposium, 2011 XXXth URSI*, pp. 1-4, 2011.
- [Bor12a] Borghys, D., Kasen, I., Achard, V. and Pernee, C. Hyperspectral Anomaly Detection: Comparative Evaluation in Scenes with Diverse Complexity. *Journal of Electrical and Computer Engineering*. V. 2012, pp. 16, Article ID 162106, 2012.
- [Cha09a] Chandola, V., Banerjee, A. and Kumar, V. Anomaly detection: A survey. *ACM Computing Surveys (CSUR)*, V. 41(3), pp. 72, 2009.
- [Cha02a] Chang, C.I. and Shao-Shan, C. Anomaly detection and classification for hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, V. 40(6), pp. 1314-1325, 2002.
- [Cha13a] Chang, C.I. *Hyperspectral Data Processing: Algorithm Design and Analysis*. John Wiley & Sons, 1164 p., 2013.
- [Cha07a] Chang, C.I. *Hyperspectral data exploitation: theory and applications*. Wiley-Interscience, 456 p., 2007.
- [Cla93a] Clark, R. N., Swayze, G. A., Gallagher, A. J., King, T. V. V. and Calvin, W. M. The U. S. Geological Survey, Digital Spectral Library: Version 1: 0.2 to 3.0 microns, U.S. Geological Survey Open File Report 93-592. 1340 p, 1993.
- [Gu08a] Gu, Y., Liu, Y. and Zhang, Y. A selective KPCA algorithm based on high-order statistics for anomaly detection in hyperspectral imagery. *Geoscience and Remote Sensing Letters*. V. 5(1), pp. 43 -47, 2008.
- [Kos97a] Kostrikin, A. I., and Manin, I. I. *Linear algebra and geometry*. Gordon and Breach Science Publishers, p. 313, 1997.
- [Mes11a] Messinger, D. W. and Albano, J. A graph theoretic approach to anomaly detection in hyperspectral imagery. *Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 2011 3rd Workshop on*. pp. 1-4, 2011.
- [Ree90a] Reed, I.S. and Yu, X. Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution. *IEEE Transactions on Acoustics, Speech, and Signal Processing*. V. 38(10), pp. 1760–1770, 1990.
- [Sch07a] Schaum, A. P. Hyperspectral anomaly detection beyond RX. *Proceedings of the SPIE Algorithms and Technologies for Multispectral, Hyperspectral and Ultraspectral Imagery XII*. V. 6565, 2007.
- [Soi09a] Soifer, V. A. *Computer Image Processing, Part II: Methods and algorithms*. VDM Verlag, p.584, 2009.
- [Soo07a] Soofbaf, S. R., Fahimnejad, H., Valadan Zoej, M. J. and Mojaradi, B. Anomaly detection algorithms for hyperspectral imagery *Proceedings, Remote Sensing and Image Processing, Presented at the Map of the World Forum*. pp. 1-8, 2007.