

GLOBAL UNSUPERVISED ANOMALY EXTRACTION AND DISCRIMINATION IN HYPERSPECTRAL IMAGES VIA MAXIMUM ORTHOGONAL-COMPLEMENT ANALYSIS

Oleg Kuybeda, David Malah, and Meir Barzohar

Department of Electrical Engineering, Technion IIT, Haifa 32000, Israel
email: koleg@technion.ac.il, malah@ee.technion.ac.il, meirb@visionsense.com

ABSTRACT

In this paper we address the problem of global unsupervised detection, discrimination, and population estimation of anomalies in hyperspectral images. The proposed approach, denoted as Anomaly Extraction and Discrimination Algorithm (AXDA), detects anomalies via analysis of a signal-subspace obtained by the recently developed Maximum Orthogonal-Complement Algorithm (MOCA). MOCA is unique in providing an unsupervised combined estimation of signal-subspace that includes anomalies, and its rank. The main idea of AXDA is to iteratively reduce the anomaly vector subspace-rank, making the related anomalies to be poorly represented. This helps to detect them by a statistical analysis of the $\ell_{2,\infty}$ -norm of data residuals. As a by-product, AXDA provides also an anomaly-free robust background subspace and rank estimation.

1. INTRODUCTION

In this work we address the problem of global unsupervised anomaly detection, as well as anomaly discrimination and population estimation of anomalies of the same type, in hyperspectral image cubes. The considered sceneries are composed of reflected spectra of abundant natural ground materials such as vegetation, soil, minerals, etc., along with anomalies such as localized man-made objects. E.g., small buildings, vehicles, etc. The wealth of spectral information in hyperspectral images provides plentiful amount of data for classification tasks. One such task relates to anomaly detection, in which hyperspectral pixels have to be classified into either background material spectra class or anomaly material spectra class. Since most often, neither prior anomaly signatures nor their statistical model, are known, anomaly detection methods first model the background and then detect anomalies by finding pixels that are not well-described by the background model. It turns out that the problem of background pixels modelling is a critical and a subtle task. As a matter of fact, it poses a two-fold problem: On one hand, the model has to be general enough in order to accurately represent the wealth of background material spectra, so as to avoid false alarms due background pixel deviations from the model. On the other hand, the model has to be concise enough (e.g., in terms of its order/rank), limiting its ability to adapt to anomalies, and leaving anomalies to disagree with the model, which is essential for a high probability of detection.

A variety of background modelling methods appears in the literature. One type of these methods is based on estimating the underlying probability density function (pdf) of the background signature, and applying a threshold to the likelihood of tested pixels. The Reed-Xiaoli (RX) algorithm [3], is a benchmark anomaly detector for hyperspectral imagery. RX models the background distribution by a single Gaussian and requires to apply a threshold to the Mahalanobis distance to detect an anomaly [3]. Unfortunately,

in many environments, it has been shown empirically that a single Gaussian modelling provides an inadequate representation of the underlying background distribution [8]. This is especially true when the local background contains multiple classes of terrain. Furthermore, the need to specify a rather arbitrary threshold value poses additional difficulties for applying RX. To properly characterize nonhomogeneous backgrounds, researchers have employed a Gaussian Mixture Model (GMM) [6], [8], which constitutes a global background model. Anomaly detection may be achieved by applying the generalized likelihood ratio test (GLRT) to the model. The authors of [8] denote the related approach by GMRX. While GMRX provides good performance, it is limited by the simplicity of Gaussian components. GMRX is further limited by the need to know or estimate a priori the number of terrain classes in the image. Some works, like [7], propose a nonlinear version - the Kernel-RX algorithm, in which RX is applied in an extended high-dimensional feature space associated with the original input via a certain nonlinear mapping function. Another approach to local background modelling corresponds to the so-called large-margin techniques, such as support vector machines (SVMs), which detect anomalies by directly estimating a decision boundary with maximal separability [5]. The authors of [5] propose to determine the minimal enclosing hypersurface that contains a training set of background data pixels. A training set is sampled from a window enclosing the tested pixel, excluding pixels belonging to its adjacent neighborhood (which makes this method local) that is supposed to be large enough to contain a maximum-size anomaly. The anomaly is detected by thresholding the distance from the tested pixel to the obtained hypersurface.

In this paper we adopt the following linear mixing model from [9]:

$$\mathbf{x}_i = \mathbf{A}\mathbf{s}_i + \mathbf{z}_i, \quad i = 1, \dots, N, \quad (1)$$

where, $\mathbf{x}_i \in \mathbb{R}^p$ is the observed pixel, $\mathbf{z}_i \in \mathbb{R}^p$ is additive Gaussian noise, columns of \mathbf{A} are the pure materials spectra (endmembers) and $\mathbf{s}_i \in \mathbb{R}^r$, $r \leq p$, their corresponding abundances. Using this model, a class of Generalized Likelihood Ratio Test (GLRT) - based algorithms to detect anomalies in hyperspectral images is proposed in [9]. The authors of [9] assume a priori known target signatures, whereas the background subspace is estimated via SVD for a priori known background subspace rank. Using Orthogonal Subspace Projection, the authors of [4] propose first to select target pixels from the data, then to detect the presence of anomaly signatures in the data pixels. Unfortunately, the approaches of [4] and [9] lack a systematic way to estimate the anomaly and the background subspace dimensionality. Moreover, they require ancillary information to identify signatures belonging to anomalies.

A recent rank-estimation approach, named MOCA [2], was shown to be a signal-subspace and rank estimation approach that preserves anomalies. The reason to this is that MOCA is based on a norm (denoted as $\ell_{2,\infty}$ -norm), which is

much more sensitive to individual pixel contributions. The proposed approach for anomaly detection, being based on MOCA, inherits this property.

2. MAXIMUM ORTHOGONAL-COMPLEMENT ALGORITHM (MOCA) PRINCIPLES

Originally, MOCA was developed for signal-subspace and rank estimation in high-dimensional noisy signals that may contain anomaly vectors, which MOCA intends to preserve. It is important to note that MOCA assumes \mathbf{z}_i to be Gaussian and *spectrally white* with a known variance. In hyperspectral data we assume the noise to be statistically independent between spectral bands. Therefore, in order to make the noise i.i.d., the noise std in each band was estimated and normalized to 1 by scaling the data. As a matter of fact, MOCA proposes an appropriate compromise between the following two approaches: The first approach is based on selecting the signal-subspace basis vectors directly from the data as presented in [4]. This approach is good for representing anomalies, since it is capable of selecting anomalies from the data. However, due to noise in the obtained basis vectors, it may perform poorly in representing background pixels. The second approach is based on SVD, which, as shown in [2], represents well the background pixels. Yet, it may perform poorly in representing anomalies [2]. Thus, the signal-subspace estimated by MOCA admits the following form: $\hat{\mathcal{S}}_k = \text{range} [\Psi_{k-h}|\Omega_h]$. I.e., $\hat{\mathcal{S}}_k$ is the space linearly spanned by columns of matrices Ω_h and Ψ_{k-h} , where Ω_h is a matrix composed of h linearly independent columns selected from the data, k is the estimated signal rank and Ψ_{k-h} is a matrix with $k-h$ orthogonal columns. This matrix is obtained via SVD of the data residuals $\mathcal{P}_{\Omega_h^\perp} \mathbf{x}_i, i=1, \dots, N$, where $\mathcal{P}_{\Omega_h^\perp}$ is a projection onto $(\text{range } \Omega_h)^\perp$. The matrix Ω_h provides a natural way to identify an anomaly subspace and to discriminate anomalies, since it collects individual data pixels that are poorly represented by the PCA-based subspace and, therefore, compose a good basis for anomaly subspace (see [2] for details). Whereas, columns of Ψ_{k-h} span the background pixels projections lying in the $(\text{range } \Omega_h)^\perp$. The signal-subspace rank estimator is based on examining the maximal data residual norms for an increasing sequence of rank values k . It is assumed that for k large enough [2], the set of all data-vector residual norms can hypothetically be divided into two subsets as follows:

$$\begin{aligned} \Gamma_k &\triangleq \{\text{squared norms of background vector residuals}\} \\ \Delta_k &\triangleq \{\text{squared norms of the remaining data-vectors}\}. \end{aligned}$$

Once the value of the maximum data-residual squared-norm $\eta_k = \max_{j=1, \dots, N} \|\mathcal{P}_{\hat{\mathcal{S}}_k^\perp} \mathbf{x}_j\|^2$ becomes available (with N denoting the number of pixels), the following two hypotheses are formulated:

$$\begin{aligned} H_0 &: \eta_k \text{ belongs to } \Gamma_k, \\ H_1 &: \eta_k \text{ belongs to } \Delta_k. \end{aligned} \quad (2) \quad (3)$$

The rank estimator \hat{r} is set to be equal to the minimal value of k for which the following condition is satisfied:

$$p(H_0|\eta_k) \geq p(H_1|\eta_k), \quad (4)$$

which means that the optimal rank is reached when there is a higher likelihood that the maximum data-residual squared norm η_k is governed by the noise statistics (i.e., it doesn't include significant signal contributions). The details of how to calculate $p(H_0|\eta_k)$ and $p(H_1|\eta_k)$ are found in [2].

3. ANOMALY EXTRACTION AND DISCRIMINATION ALGORITHM (AXDA)

In this section we propose an Anomaly Extraction and Discrimination Algorithm (AXDA) that employs signal-subspace and rank estimation results obtained by the above described MOCA. Let's recall that the signal-subspace basis Φ produced by MOCA admits the following form:

$$\Phi_{\hat{r}} = [\Psi_{\hat{r}-h}|\Omega_h], \quad (5)$$

where \hat{r} is the estimated signal-subspace rank, the sub-matrix Ω_h consists of h linearly independent columns selected from the data matrix \mathbf{X} and the sub-matrix $\Psi_{\hat{r}-h}$ consists of $\hat{r}-h$ principal components of $\mathcal{P}_{\Omega_h^\perp} \mathbf{X}$. It is important to note that the matrix Ω_h doesn't contain *all anomaly vectors* among its columns, but can be used to represent the anomaly vectors subspace. Since the background and anomaly subspaces are often non-orthogonal in hyperspectral images, the direct projection of the data onto range Ω_h could lead to a high false-alarm rate due to background contributions. Therefore, in order to detect and discriminate all anomaly pixels, we propose a new algorithm that extracts all anomalies in the data and associates them with Ω_h columns (depending on their contribution as described below). As stated earlier, the proposed algorithm is denoted as Anomaly Extraction and Discrimination Algorithm (AXDA).

3.1 AXDA

At this point, we are ready to describe AXDA in detail, as shown in Fig. 1.

1. Initialization

AXDA starts by initializing $j = h$, the number of anomaly vectors in Ω_h and $s = \hat{r}$, the determined rank of the signal subspace $\text{range} [\Psi_{s-j}|\Omega_j]$, and the maximum-residual norm denoted by η_s , all as obtained from MOCA.

2. Reduction of anomaly subspace basis

It is important to note that initially, the number $j = h$ of anomaly vectors in the partition $\Phi_{\hat{r}} = [\Psi_{\hat{r}-h}|\Omega_h]$, obtained by MOCA, is optimal for the given signal-subspace rank. That is, a decrease of h for a given \hat{r} would result in an increased maximum-residual norm and, possibly, of other residual norms of data-vectors. We intentionally alter this optimality by dropping the last column of Ω_j , providing Ω_{j-1} . This operation is designed to detect anomalies related to the last column of Ω_j .

3. Calculation of a new background vector representation basis Ψ_{s-j+1} , corresponding to the new anomaly subspace basis Ω_{j-1}

In the previous step, we reduced the rank of Ω_j . However, at this point, we still keep the total signal-subspace rank s intact. The decision whether to update s , which is carried out in the block (14), is based on the next steps. In order to retain s now, a new background vector representation basis Ψ_{s-j+1} is calculated by applying SVD on $\mathcal{P}_{\Omega_{j-1}^\perp} \mathbf{X}$ that matches the reduced-rank matrix Ω_{j-1} .

4. Calculation of data residual-norms in the obtained residual-subspace:

$$r_i = \|\mathcal{P}_{[\Psi_{s-j+1}|\Omega_{j-1}]^\perp} \mathbf{x}_i\|^2. \quad (6)$$

5. Detection of anomaly vectors belonging to the dropped column j

In this block we identify indices of all anomaly vector residuals that exceed the noise level η_s , which is equal to the maximum residual-norm initially obtained from MOCA.

6. Decision about the next operation, based on previous block results

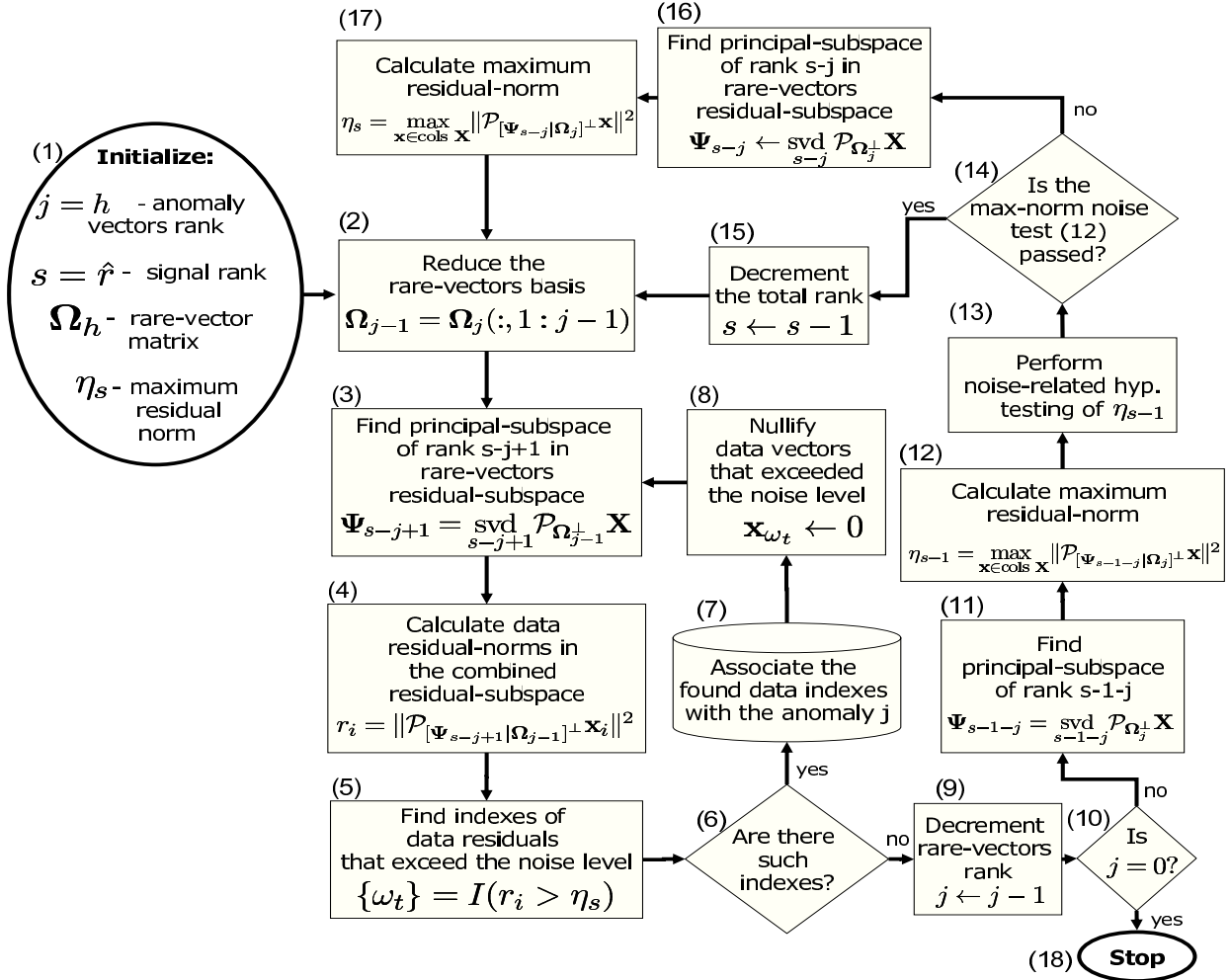


Figure 1: Detailed description of Anomaly Extraction and Discrimination Algorithm (AXDA). The notation in block (2) is a MATLAB[®] notation.

In this block we decide about the next operation based on whether anomaly vectors were found in the previous block. If there are such indices, then we perform the inner loop, in which we deplete the found anomaly vectors, recalculate Ψ_{s-j+1} , and try to detect more anomaly vectors. Otherwise, the depletion of anomaly vectors in this iteration is completed and other operations (blocks (9) - (17)) of the current iteration are performed.

7. Association of found anomaly vectors to j -th column of Ω_j

This block belongs to the inner loop of anomaly vectors depletion. We associate all data vectors indices (found in the block (5)) to the dropped column j and store them. Therefore, the corresponding anomaly vectors are denoted as j -associated anomaly vectors.

8. Depletion of found j -associated anomaly vectors from input data and recalculation of Ψ_{s-j+1}

Since $(\text{range } \Omega_{j-1})^\perp$ contains j -associated anomaly vector contributions, the subspace corresponding to Ψ_{s-j+1} (obtained earlier via SVD in Block (3)) is expected to be diverted in a way that aims to reduce these contributions along with the background vector residual-norms. As a result, not all anomaly vectors corresponding to the dropped j -th column of Ω_j may be detected via thresholding their corresponding norms by η_s in Block (5). In order to remedy this situation, we deplete the j -associated

anomaly vectors in the data (detected in Block (5)) and perform operations of blocks (3) - (5) again in order to obtain a more precise estimation of the background vectors subspace Ψ_{s-j+1} , which is not diverted by the j -associated anomaly vectors found in Block (5).

9. Decrementing of anomaly subspace rank

Once all j -associated anomaly vectors are depleted, the rank of the anomaly subspace can be reduced by one. However, this does not necessarily mean that the total signal representation rank should also drop by one. This can be explained as follows: The background and anomaly subspaces in the hyperspectral images are not orthogonal. Therefore, if one reduces the rank of $[\Psi_{s-j}|\Omega_j]$ by removing a column from Ω_j , one might transfer a significant amount of background contribution to the complementary subspace $(\text{range } [\Psi_{s-j}|\Omega_{j-1}])^\perp$, which means that the reduced-rank subspace basis $[\Psi_{s-j}|\Omega_{j-1}]$ might not represent well the signal-subspace of the data after the j -associated anomaly-vectors depletion. Therefore, the decrementing of anomaly subspace rank j does not necessarily entails decrementing the total signal-basis rank s . Thus, to decide if the total signal-basis rank s should be also decremented, we again employ, in the next blocks, the maximum-norm hypothesis testing (4). Due to the algorithm construction (see blocks (4),(5),(6)), it is guaranteed that at the input to

this block, the subspace $(\text{range} [\Psi_{s-j+1}|\Omega_{j-1}])^\perp$ doesn't contain signal contributions. It is left to determine if the same holds true for the subspace $(\text{range} [\Psi_{s-j}|\Omega_{j-1}])^\perp$, which corresponds to the reduced total signal rank $s-1$. We start by setting $j \leftarrow j-1$. In the next blocks we perform steps necessary for deciding if to decrement also the total signal-subspace rank s .

10. **Termination condition block**

If the anomaly subspace rank has reached 0, then terminate (block (18)). Otherwise, continue to block (11).

11. **Calculation of Ψ_{s-1-j} corresponding to a reduced-rank signal-subspace**

In order to decide if decrementing j should also entail the decrementing of the total signal-subspace s , one has to obtain the reduced-rank subspace $[\Psi_{s-1-j}|\Omega_j]$ and test the corresponding data residuals. Therefore, in this block, we calculate Ψ_{s-1-j} by: $\Psi_{s-1-j} = \text{SVD } \mathcal{P}_{\Omega_j^\perp} \mathbf{X}$.

12. **Calculation of maximum data residual-norm η_{s-1} in the obtained residual-subspace**

$$\eta_{s-1} = \max_{\mathbf{x} \in \text{cols } \mathbf{X}} \|\mathcal{P}_{[\Psi_{s-1-j}|\Omega_j]^\perp} \mathbf{x}\|^2$$

13. **Performing noise-related hypothesis testing of η_{s-1}**

In this block we assess if η_{s-1} contains signal-contribution. For this purpose we apply the test of equation (4).

14. **Decision if to reduce the total signal-subspace rank s**

If η_{s-1} meets the noise-hypothesis, meaning that the subspace $(\text{range} [\Psi_{s-1-j}|\Omega_j])^\perp$ doesn't contain signal contributions (i.e., the basis $[\Psi_{s-1-j}|\Omega_j]$ represents well the signal-subspace), then s should be decremented. Otherwise, leave s intact and continue to a new iteration via blocks (16) and (17).

15. **Decrementing the total signal-subspace rank s**

$s \leftarrow s-1$, and continue to a new iteration starting at block (2).

Comments

1. Once the new value of s is determined, we approach a nominal state (at block (2)), where the anomaly vectors matrix rank is decremented by 1, and the signal-subspace basis $[\Psi_{s-j}|\Omega_j]$ (with the updated values of j and s) is "MOCA-optimal" with respect to the modified data-matrix \mathbf{X} . In order to extract other anomaly vectors, corresponding to the rest of Ω_j columns, until the complete depletion of all anomaly vectors, steps 2 - 15 are repeated. The iterations stop when there are no more columns in the anomaly-basis matrix Ω_j , i.e., $j = 0$.
2. It is important to note that at the end of the AXDA procedure, the signal-subspace basis is composed solely of Ψ_s ($s \leq \hat{r}$), which constitutes the MOCA-optimal basis of the *background* vectors. Hence AXDA equips us also with an anomaly-free (in other words "robust") estimate of the background-subspace and rank.

4. EXPERIMENTS WITH REAL HYPERSPECTRAL DATA

In this section we evaluate the performance of MOCA, followed by AXDA postprocessing, by applying them on real hyperspectral data. For an analysis of the effect of noise on MOCA, using self designed synthetic data experiments with different signal to noise ratios, the reader is referred to [2]. To demonstrate the results, the proposed approach was applied to 6 real hyperspectral image cubes collected by an AISA airborne sensor configured to 65 spectral bands, uniformly covering VNIR range of 400nm - 1000nm wavelengths. At 4 km altitude pixel resolution corresponds to $(0.8m)^2$. The obtained image cubes are $b \times r \times c = 65 \times 300 \times 479$ hy-

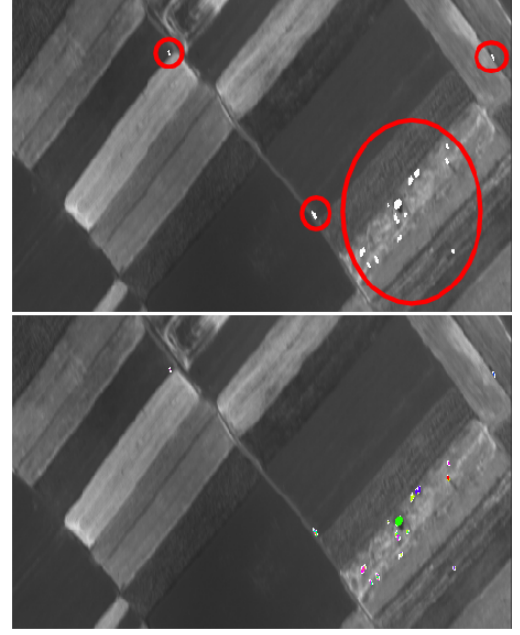


Figure 2: AXDA results at the *nominal* operating point; top image - manually identified ground-truth anomalies (marked in white and encircled in red); bottom image - anomalies (marked in color), overlayed on the white ground-truth pixels.

perspectival images, where b, r and c denote the number of hyperspectral bands, the number of rows and the number of columns in the image, respectively.

In Fig. 2 one can see results of anomaly detection and discrimination. Shown are images containing the 30th-band of the first hyperspectral cube out of 6 used for the evaluation. All 6 images are not shown here just because of space limitations. The top image contains ground-truth anomalies (marked in white and encircled by red ellipses), which were manually identified using side information collected from high resolution RGB images of the corresponding scenes. The ground truth anomalies consist of vehicles and small agriculture facilities, which occupy few-pixel segments.

The bottom image contains anomalies (marked in color) detected by AXDA, overlayed on the white ground-truth pixels. All anomaly pixels of the same type are marked by the same color. There are no missed anomalies in the presented image. The corresponding obtained signal subspace ranks determined by MOCA are $\hat{r} = 15, 10, 10, 16, 15, 11$, anomaly dimensionalities determined by AXDA are $h = 9, 2, 5, 8, 7, 4$, and the dimensionalities of anomaly-free background are $s = 11, 10, 8, 12, 13, 10$, respectively for all 6 examined images. Note, that according to the discussion in step 9 of AXDA, presented in section 3, it is possible that $s \geq \hat{r} - h$. Thus, AXDA allows discrimination of anomalies according to corresponding anomaly endmembers (constituent materials spectra) found by MOCA, although the accuracy of this discrimination is not evaluated in this paper and is under investigation.

In Fig. 3, we compare between GMRX [8], MSD [1] and the proposed AXDA in terms of Receiver Operation Characteristic (ROC) curves. For the purpose of ROC curves generation, all 6 hyperspectral images were used, in which the total number of anomaly segments count is 25. An anomaly

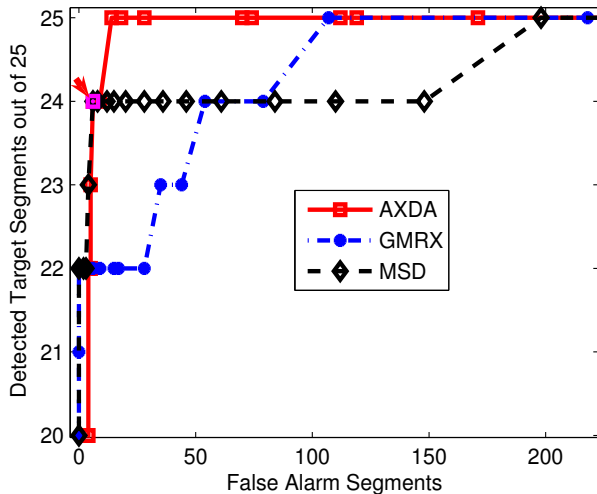


Figure 3: ROC curves corresponding to GMRX, MSD and AXDA. The nominal operating point of AXDA is marked in magenta color and is pointed out by the arrow. This point corresponds to 24 detected anomalies and 6 false alarm segments.

is considered as detected if at least one of the detected pixels hits the corresponding marked segment. All pixels detected by the algorithms were grouped into connected objects using 8-connected object labelling. If an object doesn't intersect a marked anomaly, it is considered a false alarm object. Although AXDA has naturally a single (nominal) operating point dictated by the noise statistical properties, we introduced a rather compelling parameter γ to the equation of block (5) in Fig. 1, in order to obtain multiple operating points for AXDA, which now reads as: $\{\omega_i\} = I(r_i > \gamma\eta_s)$. In words, the noise-related threshold value η_s (measured in a previous iteration) is multiplied by the factor γ in order to produce a new threshold value. The lower the factor γ is, the more data vectors will be treated as anomaly-vectors and be associated to the dropped column j of Ω_j . In our simulation, we have used 30 values of γ , which were uniformly sampled from $[0.8, 1.2]$. The position of nominal operating point of AXDA (for $\gamma = 1$) is pointed out by a red arrow. As can be seen from the figure, the nominal operating point provides a high detection rate (24 detected anomalies) with a significantly low false alarm rate (see Fig. 3). The GMRX algorithm was initialized by an excessive number of Gaussians using the k-means algorithm for initializing the Gaussian parameters. During the EM iterations of the GMRX, too small clusters, and hence unreliable, were eliminated. The MSD algorithm was provided an anomaly-free estimation of the background basis Ψ_s estimated by AXDA, which uses the anomaly subspace basis Ω_h provided by MOCA, since MOCA and AXDA combined are unique in their ability to perform an unsupervised determination of both anomaly and background subspaces and their ranks. Fig. 3 clearly shows that for the examined images AXDA has a better performance than GMRX and MSD, in most of the range of the tested parameters.

5. SUMMARY AND CONCLUSION

In this work we have proposed an algorithm for anomaly detection, discrimination and population estimation of anomalies of the same type, called AXDA. The algorithm is based on a signal-subspace and rank that are estimated by MOCA [2]. By its construction, the signal basis consists of two

groups of basis vectors. One group spans the subspace of anomalies. The second group is designed to represent background pixel residuals belonging to the subspace that is complementary to the subspace of the anomalies. The proposed AXDA identifies anomaly pixels that belong to the subspace spanned by the first group of the basis vectors. In experiments with real hyperspectral image cubes AXDA was shown to have a better performance than GMRX and MSD, in most of the range of the tested parameters. It is also important to note, that in contrast to MSD and GMRX, AXDA is equipped with an unsupervised determination of the nominal operating point. AXDA also has a capability to discriminate between different types of anomalies, though the accuracy of this discrimination, as well as the accuracy of population estimation of anomalies of the same type, are not evaluated in this paper and are under investigation. Moreover, AXDA also allows an anomaly-free (robust) estimation of the background-subspace and rank.

REFERENCES

- [1] L. L. Sharf and B. Friedlander "Matched Subspace Detectors," *IEEE Trans. Signal Proc.*, vol. 42, no. 8, pp. 2446 – 2457, Aug. 1994.
- [2] O. Kuybeda, D. Malah and M. Barzohar "Rank Estimation and Redundancy Reduction of High-Dimensional Noisy Signals with Preservation of Rare Vectors," *IEEE Trans. Signal Proc.*, vol. 55, no. 12, pp. 5579 – 5592, Dec. 2007.
- [3] I. S. Reed and X. Yu "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 1, pp. 1760 – 1770, Oct. 1990.
- [4] H. Ren and C.I. Chang "Automatic spectral target recognition in hyperspectral imagery," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, pp. 1232-1249, Oct. 2003.
- [5] A. Banerjee, P. Burlina, and C. Diehl "A Support Vector Method for Anomaly Detection in Hyperspectral Imagery" *IEEE Trans on Geoscience and Remote Sensing*, vol. 44, no. 8, pp. 2282 – 2291, Aug 2006
- [6] S. G. Beaven, D. Stein, and L. E. Hoff "Comparison of Gaussian mixture and linear mixture models for classification of hyperspectral data," in *Proc. IGARSS, Honolulu, HI*, pp. 1597 – 1599, Jul. 2000
- [7] H. Kwon and N. M. Nasrabadi "Kernel RX-Algorithm: A Nonlinear Anomaly Detector for Hyperspectral Imagery," *IEEE Trans on Geoscience and Remote Sensing*, vol. 43, no. 2, pp. 388 – 397, Feb. 2005.
- [8] D. Stein, S. Beaven, L. E. Hoff, E. Winter, A. Shaum, and A. D. Stocker "Anomaly detection from hyperspectral imagery," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 58 – 69, Jan. 2002.
- [9] D. Manolakis, C. Siracusa, and G. Shaw "Hyperspectral Subpixel Target Detection Using the Linear Mixing Model", *IEEE Trans. Geoscience and Remote Sensing*, vol. 39, no. 7, pp. 1232-1249, July 2001.