Anomaly detection using morphology-based collaborative representation in hyperspectral imagery

Maryam Imani

Faculty of Electrical and Computer Engineering, Tarbiat Modares University, Tehran, Iran

ABSTRACT

A nonparametric anomaly detection method is proposed in this paper which does not consider any probability density function for data, and so can work well for real hyperspectral image containing complicated and non-normal background. Since the predominant part of an image is composed by background pixels and because of similarity of neighboring pixels in a local region, each background pixel can be approximated from its surrounding samples. To this end, the collaborative representation with a simple and closed form solution is used. To more benefit from the spatial information of hyperspectral image, the morphological filters are applied for extraction of contextual features and a decision fusion rule is used to utilize the spatial information of all principal components of data. The experimental results show the superior performance of the proposed detector compared to several state-of-the-art anomalous target detection methods. The area under ROC curve (AUC) values achieved by the proposed detector are 95.53, 96.34 and 99.72 for San Diego hyperspectral image, Indian hyperspectral image and the WorldView-2 multispectral image, respectively.

Introduction

A large amount of spectral information of the Earth’s surface is provided by hyperspectral imagery. For example, famous Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) covers a wide range of electromagnetic spectrum (380–2500 nm) with 10 nm spectral resolution. So, various materials with minor spectral differences can be distinguished using more details provided by hyperspectral imagery than multispectral one. Hyperspectral image by providing the continuous spectral curves of material on the image scene creates an information cube containing spectral and spatial dimensions. This rich source of information simplifies object identification methods such as classification and target detection (Chein & Shao-Shan, 2002; Imani & Ghassemian, 2016, 2017; Plaza et al., 2009). Target detection, which is a binary classification problem, can be done supervised or unsupervised. In supervised target detection, the spectral signatures of objects of interest are known (Chang, Du, Zhang, & Zhao, 2017; Du & Zhang, 2014a). But, in unsupervised target detection, known as anomaly detection, there is no prior knowledge about the spectral signatures of objects or targets of interest (Du & Zhang, 2014b). Anomalies have some main characteristics. They occur with low probabilities and the spectral signatures of them are unknown. The spectral of anomalous targets (human-made objects) is very different respect to other materials on the ground surface. Anomaly detection has many applications in both civilian and military problems (Schweizer & Moura, 2000; Stein et al., 2002). The Reed-Xiaoli (RX) detector (Reed & Yu, 1990) is the most widely used anomaly detection method, which assumes a Gaussian probability density function for background clutter. The background pixels compose the predominant part of image. The RX detector measures each pixel with which probability belongs to the background. In many real hyperspectral images containing mixture and complicated background, the assumed Gaussian distribution in RX detector is not hold. So, the estimate of background statistics such as mean vector and covariance matrix in order to anomaly detection through a Mahalanobis distance with normal distribution assumption may not be efficient in many real applications.

Several strategies have been proposed for estimates of background statistics in RX (Carlotto, 2005). In Global-RX, all pixels of image contribute in estimate of mean and covariance matrix. In Local-RX, a local neighborhood is considered around the pixel under test for estimate of background statistics. The anomalous targets or noisy signals may contaminate the background information and so degrade the estimate accuracy of background statistics. This causes wrongly detection of some background or noisy signals as anomalies. To decrease the negative effects of contamination of anomalous signatures in estimation...
of background statistics, some improved versions of RX such as Weighted RX and linear filter-based RX (LF-RX) (Guo et al., 2014) have been proposed. While in RX, the same weight is assigned to all pixels for estimate of mean and covariance, in the Weighted RX, high weights are assigned to the samples which belong to the background with high probability and low weights are assigned to the samples which belong to the anomalies with high probability. The LF-RX method also performs filtering of noise and anomalies to retain the background data and refine the estimates of its statistics. Another version of RX uses the uniform target detector (UTD) (Chang & Chiang, 2002) to remove noise and background in RX detector. RX-UTD is obtained by subtracting UTD from RX. Subspace RX (SSRX) (Schaum, 2007) and some other subspace methods (Farrell & Mersereau, 2005; Ranney & Soumekh, 2006) assume that anomalous targets and background pixels can be represented in different subspaces. SSRX chooses the representative eigenvectors of hyperspectral covariance matrix for background feature extraction. The Kernel-RX (Kwon & Nasrabadi, 2005), the robust nonlinear anomaly detector (Zhao, Du, & Zhang, 2014) and some other kernel-based methods (Goldberg, Kwon, & Nasrabadi, 2007; Kwon & Nasrabadi, 2010; Sakla, Chan, Ji, & Sakla, 2011), as the nonlinear versions of RX, have been introduced for projection of RX into a high dimensional feature space using the kernel trick.

The compressed sensing theory and the sparse coding have been developed and applied in many pattern recognition problems (Dong, Li, Zhang, & Shi, 2011; Elad, 2010; Guo, Wittman, & Osher, 2009). The compressed sensing theory represents that the most of natural signals can be represented using only a few elements from a dictionary or basis compactly. Sparsity models have been successfully used to incorporate the contextual information for classification and object recognition in high dimensional data such as hyperspectral remote sensing images (Li, Du, et al., 2015; Zhao, Du, & Zhang, 2017). The background joint sparse representation detection (BJSRD) (Li, Zhang, et al., 2015) uses the benefits of sparse representation of hyperspectral images for estimate of background and anomaly detection. Since in a hyperspectral image, the pixels in a local neighborhood window have similar spectral features and assumed to belong to limited subspaces, a normal pixel under test can be predicted by the selected background dictionary.

In addition to sparse representation, the collaborative representation has been also proposed recently for hyperspectral classification. Collaborative representation means that a sample can be approximated by a subset of related samples especially by using some surrounding spatial neighborhoods in a local window. The nearest regularized subspace (NRS) classifier (Li, Tramel, Prasad, & Fowler, 2014) uses the collaborative representation for approximation of each testing pixel of image via the linear combination of training samples. NRS is a pixel-wise classifier without considering the spatial information contained in the neighborhood locations. The joint collaborative representation (JCR) method (Li & Du, 2014) is an extend version of NRS which incorporates the contextual features within the classification process. The weighted JCR (WJCR) method (Xiong, Ran, Li, Zou, & Du, 2015) extends JCR with assigning different weights to the surrounding pixels in a neighborhood local window. All of NRS, JCR and WJCR methods use the collaborative representation for classification of hyperspectral image in a supervised manner for a multi-class problem. In this paper, we propose a collaborative representation method for anomaly detection, which is a two-class and unsupervised target detection problem. To increase incorporation of contextual information in addition to spectral one, the collaborative representation is applied to the morphological filtered image instead of the original one.

In mathematical morphology, erosion and dilation are the fundamental operators where they are applied to a given image with a known shape, called a structuring element (SE). Applying the erosion operator provides an output image which represents where the SE hits the objects in the scene. Also, the output image of dilation operator shows areas that the SE fits the objects in the image (Benediktsson et al, 2005). All other morphological operators such as opening and closing can be expressed in terms of erosion and dilation. The opening operator dilates an eroded image in order to recover it. In contrast, the closing operator erodes a dilated image to recover the initial shape of dilated image structures.

Depending on the interaction between the size of structure and the SE size, some image structures may have a high response for a given SE size while they have a lower response for other SE sizes. Since the size of the interested structures for detection is not exactly known, it is a good idea to use a multiscale approach based on a range of different SE sizes instead of using a single size SE. By this approach, a morphological profile is built and exploration of different hypothetical spatial domains becomes possible (Dalla Mura, Benediktsson, Waske, & Bruzzone, 2010).

The proposed method, morphology-based collaborative representation (MCR), predicts the anomalous targets by subtracting the background, approximated by collaborative representation, from the original hyperspectral image. This process is repeated for each morphological profile (MP) acquired from each principal component of data and the final residual image is obtained by using a
decision fusion rule. The experimental results show the superior performance of the proposed MCR method compared to several state-of-the-art anomaly detection methods.

The reminder of this study is organized as follows. Section 2 introduces the proposed MCR method with more details. The experimental results and conclusions are discussed in section 3 and 4, respectively.

MCR

The most of anomaly detectors such as RX and its variants are parametric methods which need to estimate the background statistics such as mean vector and covariance matrix. But our proposed method in this paper, called MCR, is a nonparametric anomaly detection method that does not consider any probability density function for background data and thus can work efficiently for real complicated backgrounds where a multivariate normal distribution cannot well fit to data.

Anomalies are targets of interest with different spectral signatures respect to local background pixels and also occur with low probabilities. So, the predominant of image is composed by the background pixels. Therefore, each background pixel can be approximated by its spatial neighbors. The MCR method utilizes the useful spatial neighborhood information for estimation of background through the collaborative representation. Each background pixel is linearly approximated from the neighborhood pixels. The estimated background image is subtracted from the original hyperspectral data, and finally anomalies are predicted from the residual image. Since the hyperspectral image, in addition to rich spectral information contains high geometrical features for modeling the objects in the scene, the collaborative representation is applied on the morphological profile (MP) containing the useful spatial features extracted from the data. The estimate of background from the spatial features of neighboring pixels through the collaborative representation can improve the estimate accuracy. For extraction of spatial features from data, the morphological filters including opening and closing transformations are applied to the data. To this end, at first, the principal component analysis (PCA) is implemented and $q$ principal components (PCs) of data with the highest variance (energy) are chosen, i.e. $PC_1, PC_2, \ldots, PC_q$. Then, from each single band image $PC_i$; $i = 1, \ldots, q$, a MP is achieved by applying the opening and closing operators by reconstruction as follows:

$$MP_i(PC) = \{\varphi_i^1(PC), \ldots, \varphi_i^m(PC), \ldots, \varphi_i^n(PC)\},$$

where $\varphi_i^\lambda(y_i^\lambda)$; $\lambda = 1, \ldots, n$ are the morphology closing (opening) operators by reconstruction and $\lambda$ is the structuring element (SE). So, with applying $n$ opening and $n$ closing morphological operators, a MP consist of $m = 2n + 1$ bands containing spatial information is provided from each PC. Applying morphological filters on the data removes the salt and pepper noise and prevents to wrongly detect noisy pixels as anomalies. The geometrical characteristics of the sliding window, SE, determines the degree of processing of image. The opening filter by reconstruction removes the bright connected components compared to grey level of neighbors everywhere that SE does not fit in. This operator merges the flat regions without distorting the edges of region and with preserving the geometrical characteristics of spatial structures. By duality, the closing filter by reconstruction removes dark connected components from the image.

The multiband MP image provided from each PC is denoted by $X_i = \{x_i^k\}_{k=1}^N$; $i = 1, \ldots, q$ in $R^m$ where $q$ is the number of PCs, i.e. $q$ MPs are produced from $q$ PCs, $N$ is the number of total pixels in the scene image, $m = 2n + 1$ is the number of bands in each MP and $n$ is the number of closing/opening operators. Around each test pixel in each MP, $y_i$; $i = 1, \ldots, q$, a local window containing $p$ neighboring pixels, denoted by $X_{p,i} = \{x_1^i, x_2^i, \ldots, x_p^i\}$; $i = 1, \ldots, q$, is considered. The objective is that the test pixel $y$ is approximated by its spatial neighbors $X_{p,i}$. To this end, the following objective function is defined:

$$\arg\min_{\mathbf{w}_i} \|y_i - X_{p,i} \mathbf{w}_i\|^2 + \lambda \|\Gamma_y \mathbf{w}_i\|^2; i = 1, \ldots, q$$

(2)

where $\mathbf{w}_i$ is the coefficient vector, $\Gamma_y$ is a biasing Tikhonov matrix specific to each test sample $y_i$ and $\lambda$ is a global regularization parameter that balances the minimization between the first term (residual) and the second term (regularization). Specifically, a diagonal $\Gamma_y$ matrix is used in the form of:

$$\Gamma_y = \begin{bmatrix} \|y_i - x_{1,1}\|^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \|y_i - x_{p,1}\|^2 \end{bmatrix}; i = 1, \ldots, q$$

(3)

where $x_{1,1}, \ldots, x_{p,1}$ are the columns of $X_{p,1}$. According to the defined regularization matrix, larger coefficients are assigned to those surrounding pixels which are more similar to the central pixel and vice versa. In other words, depending upon the similarity or distance of the neighboring pixel to the central pixel, the penalty values on different elements in the coefficient vector are different. The smaller distance means that the neighboring pixel is more similar to the central pixel, and so, a large coefficient is assigned to it. To solve the optimization problem in (2), the derivative with regard to $\mathbf{w}_i$ is taken and set to zero. The solution is:
After estimation of the coefficient vector, the background in each of \( q \) MPs is estimated by:

\[
\hat{\mathbf{y}}_i = \mathbf{X}_{p,i} \mathbf{w}_i; \quad i = 1, \ldots, q
\]

Then, in order to anomaly detection, the residual images \( \text{res}_i; \ i = 1, \ldots, q \) are computed by subtracting the predicted background from the original data as follows:

\[
\text{res}_i = \| \mathbf{y}_i - \hat{\mathbf{y}}_i \|_2 = \| \mathbf{y}_i - \mathbf{X}_{p,i} \mathbf{w}_i \|_2; \ i = 1, \ldots, q
\]

The residual images are computed from all MPs obtained from \( q \) PCs. To provide the information of all PCs, the majority voting (MV) rule is used to fuse the results of all components as follows:

\[
\text{res}_f = \text{mode}_{i=1 \ldots q} \text{res}_i
\]

where \( \text{mode}(\cdot) \) computes the mode of values and \( \text{res}_f \) is the output of detector or detection map. In an anomaly detection problem, two common hypotheses \((H_0 \text{ and } H_1)\) are:

\[
H_0 : x = b \text{ (Target absent)}
\]

\[
H_1 : x = at + b \text{ (Target present)}
\]

where \( a = 0 \) under \( H_0 \) and \( a > 0 \) under \( H_1 \). \( t \) is the target and \( b \) is the background signal following a normal distribution. Finally, the anomalous targets are determined by comparison of final residual image with an appropriate threshold \( \tau \) is follows:

\[
H_1, \quad \text{res}_f \leq \tau, \quad H_0
\]

If the residual value is larger than the threshold \( \tau \), then the pixel \( \mathbf{y} \) is claimed to an anomaly, if not, it is assigned to the background clutter.

The flowchart of the proposed MCR method is shown in Figure 1. Briefly, At first \( q \) PCs of data is obtained by applying the PCA transformation on the hyperspectral data. Then, a MP is acquired from each PC by repeated use of \( n \) opening and \( n \) closing filters by reconstruction on the PC image. The result is \( MP_n(\mathbf{PC}_i); i = 1, \ldots, q \). After that, the collaborative representation (CR) is applied on each morphological filtered image for estimate of background. From each approximated background, the residual image is computed. The final residual image is obtained by applying the MV rule for decision fusion. The final residual image is used for anomaly detection through comparison of it with a threshold value.

**Experiments**

Two real hyperspectral images are used for doing experiments. The first one is an urban scene of San Diego airport in CA, USA collected by AVIRIS. This image with a spatial resolution of 3.5 m per pixel contains 224 original spectral bands spanning the wavelength range of 0.37–2.51 µm. The spectral channels corresponding to water absorption regions and low signal-to-noise ratios (1–6, 33–35, 107–113, 153–166, and 221–224) are removed and 189 spectral bands are retained for doing our experiments. The background consists of different land covers such as road, ground, roof, airpark, building, shadow, and airplanes.

Some human-made objects with different spectral signature respect to background are considered as anomalies. The target map is obtained by allocating one instead of interested targets and zero instead of background pixels. The experiments are done on a region with 80 × 80 pixels of this data. Band 19 of San Diego hyperspectral image and the region under test is shown in Figure 2.

The Indian Pines hyperspectral image collected by AVIRIS over Northwestern Indiana in June of 1992 consists of the agricultural and forest areas. It has 145 × 145 pixels with 20 m spatial resolution per pixel. It comprises 224 spectral bands which the

---

Figure 1. Flowchart of the proposed MCR method (HSI: hyperspectral image, PC: principal component, MP: morphological profile, CR: collaborative representation, res: residual image, MV: majority voting).
number of them is reduced to 200 by removing noisy and water absorption bands. The wavelength range is from 0.4 to 2.5 µm with a nominal spectral resolution of 10 nm. There are different land covers consisting of 16 classes such as corn, grass-trees, alfalfa, soybean and so on in this dataset. The Oats class which only consists of 20 pixels of this image is considered as target. The grey level image of band 9 of Indian and the used target map are shown in Figure 3. There are many difficulties in separation of different classes in Indian pines image. This dataset is acquired in June. So, the most of crops and plants were in early stages of growth. The Indian pines is a study area where some of the crops planted in previous years remain on the surface of the ground as a dirt. Based on the amount of dirt remains on the earth, parts of the study area are divided into three categories: the no-till, zone-till and complete-till areas. There are different materials in soils with the same plants. So, the measured spectrums for a plant are very different from each other in different areas. These factors result in multi-modal probability distribution for Indian image with a large overlap between different classes. Thus, the separability between different classes is low. Based on aforementioned reasons, the classification and target detection of Indian is naturally hard. To show the distribution of each hyperspectral image in two dimensional feature spaces, the PCA transform is applied to it. Two first principal components of PCA transform indicated by PC1 and PC2 are used to obtain the histogram of datasets. The histogram of San Diego and Indian hyperspectral images are shown in Figure 4.

To assess and comparison of different detectors different tools are used: receiver operating characteristic (ROC) curve, area under curve (AUC) value, separation analysis, and detection maps. The ROC curve represents the relationship between the probability of detection (PD) and the false alarm rate (FAR), which given by:

$$PD = \frac{N_{cd}}{N_t}, \quad FAR = \frac{N_{fd}}{N}$$

(11)

where $N$ is the total number of pixels in the image, $N_t$ is total number of anomalous pixels, $N_{cd}$ denotes the number of correctly detected pixels, and finally $N_{fd}$ represents the number of falsely detected anomalous pixels. The ROC curve represents the relationship between the probability of detection (PD) and the false alarm rate (FAR). A detector that obtains a higher PD compared than others, at the same FAR,
is superior to them. AUC value is a number between 0 and 1. If the AUC value be closer to 1, it means that the detector has higher performance. In other words, the ROC curve that is closest to the upper left corner of coordinates, i.e. its associated AUC value is the highest, represents the best detector.

The proposed MCR method is compared to some popular and state-of-the-art anomaly detection methods such as Global-RX, Local-RX, Weighted RX, LF-RX, Kernel-RX, SSRX, RX-UTD, BJSRD and WJCR. In Local-RX, Weighted RX, LF-RX and MCR methods a dual-window strategy is applied in order to model the local background around each pixel under test to reduce interference. The dual windows prevent from mixing anomalous target pixels with the outer neighboring pixels. A rectangular local neighborhood is set on the outside of the guard window. The length of dual window, $w_{\text{in}} \times w_{\text{in}}$ for inner window and $w_{\text{out}} \times w_{\text{out}}$ for outer window, are considered in the range of {3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25} where $w_{\text{out}} > w_{\text{in}}$ and the best parameters are chosen for each method through doing experiments. The inner and the outer window sizes of the dual window are set to $7 \times 7$ and $13 \times 13$, respectively. Four percent of total pixels of image are randomly chosen to form the background data for computing the Gram matrix and the kernel map of the test pixel in the Kernel-RX method. The parameters of BJSRD are set as follows. The upper bound of the sparsity level in the sparse representation of the background data is chosen as $L = 3$. The size of guard and outside local window is set as $5 \times 5$ and $17 \times 17$, respectively, and the size of the local dictionary is chosen as $19$. Also, the residual threshold in SOMP algorithm used in BJSRD is set as $1e^{-7}$. Please note that WJCR is a supervised method which needs to a training set. To this end, 10% of available target and background samples available in target map are used as training samples.

The proposed MCR method uses the collaborative representation for estimate of background. Then, anomalies are predicted through the residual image, obtained by subtracting the approximated background image from the original hyperspectral data. As said, the collaborative representation is applied to the morphological filtered image instead of the original one. The effect of applying morphological filters in the performance of MCR is assessed. To this end, in Figure 5, the proposed detector is assessed in two different versions for San Diego dataset. The first version is called collaborative representation (CR) where the morphological profiles, i.e. $MP_n(PC_1), MP_n(PC_2), \ldots, MP_n(PC_q)$, are deleted from the block diagram of Figure 1 and the collaborative representation is applied straightforwardly.
to the original hyperspectral image. In the improved version, the proposed morphological-based collaborative representation (MCR) detector, the collaborative representation is applied to the morphological profile produced from each principal component of hyperspectral image according Figure 1. Morphological filters can extract valuable spatial information from the image structures. So, the MCR detector, which utilizes both of spectral and spatial information, outperforms the CR detector, which only uses the spectral information of hyperspectral data. The area under ROC curve (AUC) values for both methods are obtained: AUC = 95.53 for MCR and AUC = 94.76 for CR. The McNemars test is used to assess the statistical significance of differences between two detectors. Comparison of two methods by using McNemars test is done based on calculation of parameter \( Z_{12} \) as follows:

\[
Z_{12} = \frac{f_{12} - f_{21}}{\sqrt{f_{12} + f_{21}}} \quad (12)
\]

where \( f_{12} \) is the number of samples correctly detected by method 1 and incorrectly by method 2. The difference in the accuracy between detectors 1 and 2 is said to be statistically significant if \( |Z_{12}| > 1.96 \). The sign of \( Z_{12} \) indicates whether detector 1 is more accurate than detector 2 (\( Z_{12} > 0 \)) or vice versa (\( Z_{12} < 0 \)). If consider MCR as detector 1 and CR as detector 2, the parameter \( Z_{12} = 4.58 \) is obtained. It means that MCR is more accurate than CR and also this difference is statistically significant.

In the latter experiment, the effect of the number of principal components \( q \) obtained by the PCA transformation and also the number of opening and closing morphological operators \( n \) are assessed for San Diego and the results are shown in Table 1. According to Table 1, the worst results are obtained for \( q = 1 \) which means that the first principal component (PC1) is not enough to provide a reasonable detection result. The use of more principal components increases the probability of correct detection. The use of four PCs, i.e. PC1, PC2, PC3, and PC4 achieves the best results. Moreover, with increasing the number of opening and closing morphological operators \( n \) from \( 2n + 1 = 3 \) to \( 2n + 1 = 27 \), the AUC value is increased and after that it is decreased. It means that from one hand, the use of low number of morphological filters is not enough to obtain

| \( q \) | AUC | AUC |
|----|----|----|
| 1  | 75.56 | 79.63 |
| 2  | 90.64 | 93.97 |
| 3  | 92.83 | 94.66 |
| 4  | 92.98 | 94.94 |
| 5  | 92.29 | 94.38 |

Table 1. The effects of parameters of principal component analysis (PCA) and morphological profile (MP) on the performance of MCR.

\( q \): The number of principal components obtained by the PCA transformation

\( n \): The number of opening and closing morphological operators

The reported values are AUC values.

Figure 6. ROC curves of MCR and the competitor detectors for San Diego hyperspectral image.
different image structures of data; and from the other hand, the use of much number of filters provide redundant spatial features which may decrease the performance of detector. So, according to the obtained results, the best detection results are obtained by $q = 4$ and $n + 1 = 27$.

**Table 2.** AUC values and running computation times of different anomaly detection methods for San Diego hyperspectral image.

| Method         | Global RX | Local RX | Weighted RX | LF-RX | BJSRD | Kernel RX | SSRX | RX-UTD | MCR |
|----------------|-----------|----------|-------------|-------|-------|-----------|------|--------|-----|
| Area under ROC curve (AUC) | 46.33     | 68.71    | 81.66       | 65.84 | 87.44 | 87.38     | 80.17| 48.22  | 95.53|
| Computation time (second)     | 0.33      | 124.47   | 255.75      | 77.66 | 2754.90| 348.85    | 0.27 | 0.29   | 555.93|

**Figure 7.** False-color detection maps of different detectors for San Diego hyperspectral image.

**Figure 8.** Binary detection maps of different detectors for San Diego hyperspectral image.
The ROC curves of competitor detectors for San Diego are shown in Figure 6. The AUC values and the running computation times of different anomaly detection methods are reported in Table 2. As seen, MCR is superior to other methods in terms of anomaly detection performance. It provides the highest AUC. After that, the BJSRD and Kernel-RX methods provide good results. Then, the Weighted-RX and SSRX methods are preferred methods. The high performance of MCR is expected because it does not need to estimate the background statistics and utilizes the benefits of collaborative representation for estimate of background clutter. Moreover, it uses the spatial information of hyperspectral data, in addition to spectral information, by applying the morphological filters on data. The collaborative representation of pixels filtered by morphological operators significantly improves the performance of detector. BJSRD utilizes the benefits of sparse representation for approximation of background. Kernel-RX by mapping data from the original feature space to a higher dimensional feature space increases the discrimination between the nonlinear separable anomalous targets and backgrounds. The use of efficient representative components with the most energy in SSRX also provides reasonable detection results. The least efficiency is obtained by Global-RX.

Among different detectors, the BJSRD method needs the highest computation time. This is because of solving the sparse optimization problem.

Figure 9. Background-anomaly separation analysis of different detectors for San Diego hyperspectral image.

Figure 10. ROC curves of MCR and the competitor detectors for Indian hyperspectral image.
The simultaneous orthogonal matching pursuit (SOMP) (Tropp, Gilbert, & Strauss, 2006) is used for computation of coefficient matrix. After BJSRD, the proposed MCR method needs more computation time respect to other methods. Solution of optimization problem including a distance-weighted regularization matrix spends time. SSRX, RX-UTD and Global-RX are the fastest detectors.

The false color detection maps achieved by MCR and other detectors are shown in Figure 7. The output of each detector is a value that is a measure of belonging of each pixel to the anomalies. In each map, the output value is shown with a color where a color nearer to red represents that the pixel belongs to anomalous, while a color nearer to blue represents that the pixel is background. As we can see from this figure, in the Kernel-
RX detection map, there are many pixels with a high detection value despite not being anomalous. This indicates that Kernel-RX cannot avoid the presence of noisy pixels or ones corresponding to rare materials wrongly detected as anomalies. There are also several samples wrongly detected as anomalies in other detectors. It can be seen that the proposed MCR detector gives a detection map where the anomalous targets are obvious. By considering a threshold on the false color detection maps, the binary detection maps are provided. In other words, false color detection map is output of formula (7) while binary detection map is output of formula (10). The binary detection maps for San Diego dataset are shown in Figure 8.

Detection output range of the anomalous targets and the background for each detector are shown in Figure 9 for San Diego image. These separability maps are used to further investigate the separability between anomalous targets and background. The boxes show the distribution of the target pixels' values. There are background and target boxes for each detector. The boxes enclose the main parts of the pixels, excluding the smallest 10% and the biggest 10%. The extreme values are shown with the lines at the bottom and top of each box and the line in the middle of the box illustrates the mean of pixels. The position of the boxes represents the compactness and tendency of the distribution of the pixels. In other words, the separability of the background and anomalies are reflected by the position of background and anomalies boxes respect together. As seen from this figure, the anomaly targets and background are well separated from each other by the MCR method especially in the middle 80% of the main pixels.

The ROC curves, false color detection maps, binary detection maps and Background-anomaly separation analysis for Indian hyperspectral image are shown in Figure 10–13, respectively. The AUC values and computation times are also reported in Table 3.

![Figure 13](image)

**Figure 13.** Background-anomaly separation analysis of different detectors for Indian hyperspectral image.

| Method       | Global RX | Local RX | Weighted RX | LF-RX | BJSRD | Kernel RX | SSRX | RX-UTD | WJCR | MCR |
|--------------|-----------|----------|-------------|-------|-------|-----------|------|--------|------|-----|
| Area under ROC curve (AUC) | 51.77 | 63.84 | 62.07 | 59.46 | 66.14 | 62.04 | 44.76 | 52.52 | 75.09 | 96.34 |
| Computation time (second) | 0.55 | 300.48 | 609.70 | 162.03 | 11,801.51 | 17,712.84 | 0.75 | 0.60 | 178.11 | 2110.77 |

![Table 3](image)

**Table 3.** AUC values and running computation times of different anomaly detection methods for Indian hyperspectral image.

![Figure 14](image)

**Figure 14.** Multispectral data: pseudo color image (left), target map (right).
In addition to hyperspectral images, the performance of the proposed MCR detector compared to other methods is assessed on a multispectral image. A 128×128 pixels image with three channels containing red, green and blue, which is acquired by WorldView-2, is used for doing experiments in this paper. WorldView-2 is a high-resolution satellite which provides panchromatic images at 0.46 m resolution and 8-band MS images with spatial resolution of 1.84 m. The used MS image in this paper is downloaded from the website [http://www.datatang.com/data/43234](http://www.datatang.com/data/43234) where it is originated from Digital Globe arranged by Beijing key laboratory of digital media, Beihang university. The available MS images in this website contain only these three spectral bands. See pseudo color image and corresponding target map in Figure 14. The ROC curves, false-color detection maps, binary detection maps and background-anomaly separation analysis for multispectral image are shown in Figure 15–18, respectively.

![ROC curves of MCR and the competitor detectors for multispectral image.](image1)

**Figure 15.** ROC curves of MCR and the competitor detectors for multispectral image.

![False-color detection maps of different detectors for multispectral image.](image2)

**Figure 16.** False-color detection maps of different detectors for multispectral image.
The AUC values and computation times are also reported in Table 4. As we can see from the background-anomaly separation analysis in Figure 18, the proposed MCR method provides a large gap between target and background pixels that means it can completely separate targets from background. According to the obtained results, it is found that the proposed MCR detector can efficiently detect the anomalous targets with a tolerable computation time in both hyperspectral and multispectral images.

**Conclusion**

The morphological-based collaborative representation (MCR) method was proposed in this paper for anomaly detection. This nonparametric method unlike conventional detectors such as RX and its different variants does not enforce a multivariate normal distribution for data. MCR instead of estimate of background statistics utilizes the spatial information contained in the hyperspectral image.
for background estimation. To this end, the useful contextual information of data is extracted by applying the morphological operators to data, and then background data is estimated by using the collaborative representation. The majority voting rule fuses the residual images obtained from the principal components of data. According to the experimental results, MCR provides the best ROC curve and the highest AUC value among some popular and state-of-the-art anomaly detection methods such as Global-RX, Local-RX, Weighted RX, LF-RX, Kernel-RX, SSRX, RX-UTD, BJSRD and WJCR. Moreover, the separability between anomaly and background pixels in MCR is high and the computation time of MCR compared to high computational methods such as BJSRD is reasonable.

Disclosure statement

No potential conflict of interest was reported by the author.

ORCID

Maryam Imani http://orcid.org/0000-0002-1924-9776

References

Benediktsson, J.A., Aever Palmason, J., & Sveinsson, J.R. (2005). Classification of hyperspectral data from urban areas based on extended morphological profiles. IEEE Transactions on Geoscience and Remote Sensing, 43(3), 480–491. doi:10.1109/TGRS.2004.824787

Carlotto, M.J. (2005). A cluster-based approach for detecting man-made objects and changes in imagery. IEEE Transactions on Geoscience Remote Sensing, 43(2), 374–387. doi:10.1109/TGRS.2004.841481

Chang, C.-I., & Chiang, S.-S. (2002). Anomaly detection and classification for hyperspectral imagery. IEEE Transactions on Geoscience Remote Sensing, 40(6), 1314–1325. doi:10.1109/TGRS.2002.800280

Chang, S., Du, B., Zhang, L., & Zhao, R. (2017). IBRS: An iterative background reconstruction and suppression framework for hyperspectral target detection. IEEE Journal of Selected Topics Applications Earth Observation Remote Sensing, 10(7), 3406–3417. doi:10.1109/JSTARS.2017.2676120

Chein, I.C., & Shao-Shan, C. (2002). Anomaly detection and classification for hyperspectral imagery. IEEE Transactions on Geoscience Remote Sensing, 40(6), 1314–1325. doi:10.1109/TGRS.2002.800280

Dalla Mura, M., Benediktsson, J.A., Waske, B., & Bruzzone, L. (2010). Morphological attribute profiles for the analysis of very high resolution images. IEEE Transactions on Geoscience and Remote Sensing, 48(10), 3747–3762. doi:10.1109/TGRS.2010.2048116

Dong, W., Li, X., Zhang, L., & Shi, G. (2011). Sparsity-based image denoising via dictionary learning and structural clustering. Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on 20–25 June, 2011. Proceedings IEEE Conference Computation Vision Pattern Recognition, Colorado Springs, CO, USA, 457–464.

Du, B., & Zhang, L. (2014a). Target detection based on a dynamic subspace. Pattern Recognition, 47(1), 344–358. doi:10.1016/j.patcog.2013.07.005

Du, B., & Zhang, L. (2014b). A discriminative metric learning based anomaly detection method. IEEE Transactions on Geoscience and Remote Sensing, 52(11), 6844–6857. doi:10.1109/TGRS.2014.2303895

Elad, M. (2010). Sparse and redundant representations: From theory to applications in signal and image processing. Berlin, Germany: Springer.

Farrell, M.D., & Mersereau, R.M. (2005). On the impact of PCA dimension reduction for hyperspectral detection of difficult targets. IEEE Geoscience Remote Sensing Letters, 2(2), 192–195. doi:10.1109/LGRS.2005.846011

Goldberg, H., Kwon, H., & Nasrabadi, N.M. (2007). Kernel eigenspace separation transform for subspace anomaly detection in hyperspectral imagery. IEEE Geoscience Remote Sensing Letters, 4(4), 581–585. doi:10.1109/LGRS.2007.903083

Guo, Q., Zhang, B., Ran, Q., Gao, L., Li, J., & Plaza, A. (2014). Weighted- RXD and linear filter-based RXD: Background statistics estimation for anomaly detection in hyperspectral imagery. IEEE Journal of Selected Topics Applications Earth Observation Remote Sensing, 7(6), 2351–2366. doi:10.1109/JSTARS.2014.2302446

Guo, Z., Wittman, T., & Osher, S. (2009). L1 unmixing and its application to hyperspectral image enhancement. Proceedings SPIE, 7334(73341M-1–73341M-9).

Imani, M., & Ghasemian, H. (2016). Spectral-spatial classification of high dimensional images using morphological filters and regression model. 6th International Conference on Intelligent & Advanced Systems (ICIAS2016), Kuala Lumpur, Malaysia.

Imani, M., & Ghasemian, H. (2017). Local histogram and discriminative learning-based hyperspectral data classification. Remote Sensing Letters, 8(1), 86–95. doi:10.1080/2150704X.2016.1232868

Kwon, H., & Nasrabadi, N.M. (2005). Kernel RX-algorithm: A nonlinear anomaly detector for hyperspectral imagery. IEEE Transactions Geoscience Remote Sensing, 43(2), 388–397. doi:10.1109/TGRS.2004.841487

Kwon, H., & Nasrabadi, N.M. (2010). Kernel RX-algorithm: A nonlinear anomaly detector for hyperspectral imagery. IEEE Transactions Geoscience Remote Sensing, 48(11), 4099–4109.

Li, J., Zhang, H., Zhang, L., & Ma, L. (2015). Hyperspectral anomaly detection by the use of background joint sparse representation. IEEE Journal of Selected Topics Applications Earth Observation Remote Sensing, 8(6), 2523–2533. doi:10.1109/JSTARS.2015.2437073

Li, W., & Du, Q. (2014). Joint within-class collaborative representation for hyperspectral image classification. IEEE Journal of Selected Topics Applications Earth Observation Remote Sensing, 7(6), 2200–2208. doi:10.1109/JSTARS.2014.2306956

Li, W., Du, Q., & Zhang, B. (2015). Combined sparse and collaborative representation for hyperspectral target detection. Pattern Recognition, 48, 3904–3916. doi:10.1016/j.patcog.2015.05.024

Li, W., Tramel, E.W., Prasad, S., & Fowler, J.E. (2014). Nearest regularized subspace for hyperspectral image classification. IEEE Transactions Geoscience Remote Sensing, 52(1), 477–489. doi:10.1109/TGRS.2013.2241773

Plaza, A., Benediktsson, J.A., Boardman, J.W., Brazile, J., Bruzzone, L., Camps-Valls, G., … Trianni, G. (2009).
Recent advances in techniques for hyperspectral image processing. Remote Sensing Environment, 113, S110–S122. doi:10.1016/j.rse.2007.07.028
Ranney, K.I., & Soumekh, M. (2006). Hyperspectral anomaly detection within the signal subspace. IEEE Geoscience Remote Sensing Letters, 3(3), 312–316. doi:10.1109/LGRS.2006.870833
Reed, I.S., & Yu, X. (1990). Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution. IEEE Transactions Acoustics Speech Signalling Processing, 38(10), 1760–1770. doi:10.1109/29.60107
Sakla, W., Chan, A., Ji, J., & Sakla, A. (2011). An SVDD-based algorithm for target detection in hyperspectral imagery. IEEE Geoscience Remote Sensing Letters, 8(2), 384–388. doi:10.1109/LGRS.2010.2078795
Schaum, A.P. (2007). Hyperspectral anomaly detection beyond RX. Proceedings SPIE, 6565, 656502.
Schweizer, S.M., & Moura, J.M.F. (2000). Hyperspectral imagery: Clutter adaption in anomaly detection. IEEE Transactions Information Theory, 46(5), 1855–1871. doi:10.1109/18.857796
Stein, D.W.J., Beaven, S.G., Hoff, L.E., Winter, E.M., Schaum, A.P., and Stocker, A.D. (2002). Anomaly detection from hyperspectral imagery. IEEE Signal Processing Magazine, 19(1), 58–69.
Tropp, J.A., Gilbert, A.C., & Strauss, M.J. (2006). Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit. Signal Process, 86(3), 572–588. doi:10.1016/j.sigpro.2005.05.030
Xiong, M., Ran, Q., Li, W., Zou, J., & Du, Q. (2015). Hyperspectral image classification using weighted joint collaborative representation. IEEE Geoscience Remote Sensing Letters, 12(6), 1209–1213. doi:10.1109/LGRS.2015.2388703
Zhao, R., Du, B., & Zhang, L. (2014). A robust nonlinear hyperspectral anomaly detection approach. IEEE Journal of Selected Topics Applications Earth Observation Remote Sensing, 7(4), 1227–1234. doi:10.1109/JSTARS.2014.2311995
Zhao, R., Du, B., & Zhang, L. (2017). Hyperspectral anomaly detection via a sparsity score estimation framework. IEEE Transactions Geoscience Remote Sensing55(6), 3208-3222.