DSmT Based RX Detector for Hyperspectral Imagery

Lin He, Peipei Zhang,Weitong Ruan
College of Automation Science and Engineering
South China University of Technology
Guangzhou, China

***Abstract*—Anomaly detection is very useful for hyperspectral detection with no a priori information of spectral signature. However, most prevailing anomaly detectors are applied directly to the all-bands data without considering the characteristics of hyperspectral imagery in local spectrum range and some computation problem incurred from high dimensional data. In this paper, a Dezert-Smarandache Theory (DSmT) based anomaly detector is presented to handle these problems. The all-bands data are first partitioned into several lower dimensional band-subsets. Then the generalized basic belief assignment of DSmT reasoning are constructed according to target signal to noise ratio (TNR) of different band subsets and probability densities of the detection value from different band subsets. Experimental results from the real hyperspectral imagery of Operative Modular Imaging Spectrometer I (OMIS-I) show that our method outperforms the benchmark RX detector (RXD).**

Keywords-hyperspectral imagery; anomaly detection; DSmT;band-subset

I. INTRODUCTION

Hyperspectral imagery records the electromagnetic spectrum information in high spectral resolution which can be utilized to detect target. In hyperspectral anomaly detection, we have no *a priori* information about the spectral signature of the target and the background. The most reasonable method is to treat pixel whose spectral signature is significantly different from that of background as anomaly (target) [1]. The most commonly used anomaly detector in hyperspectral imagery are RX detector (RXD) based on generalized likelihood ratio test (GLRT), low probability detection (LPD) and unified target detection (UTD) [2-4], etc. All these anomaly detectors are based on covariance matrix of all spectral bands (CMASB). However, CMASB based anomaly detector has some inherent weakness, such as concealing subtle differences between materials in some local spectrum ranges, low precision of estimating the covariance matrix, and even causing the singularity problem in calculation.

In order to handle these problems, we present a DSmT based anomaly detection method in this paper. The whole hyperspectral data are first partitioned into several band-subsets. This partition reveals the local spectral variation in hyperspectral data [5-6]. Also, since the spectral dimension of each band-subset is much lower, the singularity problem may be avoided when applying CMASB based anomaly detectors to each band-subset, and a better estimation of covariance matrix can be obtained. After band-subset partition, we apply the RXD to each band-subset. Then, the thing remained is to determine how to effectively merge the primary detection results from individual band-subsets to improve the final detection performance. In this paper, we present a novel DSmT based anomaly detector to integrate the discriminabilities from different band-subsets.

II. GLRT BASED RXD

Target detection can be stated as a problem of hypothesis testing [1][2]. The likelihood ratio test is most commonly used method for hypothesis testing. However, in some practical hyperspectral detection, two likelihood functions of the target and the background may depend on some unknown parameters. Thus, the generalized likelihood ratio test (GLRT) can be utilized as follows

$$\Lambda_G(\mathbf{x}) = \frac{L\left(\mathbf{x}_n \mid \hat{\boldsymbol{\theta}}_{1,MLE}, H_1\right)}{L\left(\mathbf{x}_n \mid \hat{\boldsymbol{\theta}}_{0,MLE}, H_0\right)} = \frac{\sup_{\boldsymbol{\theta}_1 \in \Theta_1} L\left(\mathbf{x}_n \mid \boldsymbol{\theta}_1, H_1\right)}{\sup_{\boldsymbol{\theta}_0 \in \Theta_0} L\left(\mathbf{x}_n \mid \boldsymbol{\theta}_0, H_0\right)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{1}$$

where $\boldsymbol{\theta}_0$ and $\boldsymbol{\theta}_1$ are the parameters of the likelihood functions of background and target, respectively. $\mathbf{x}_n$ is the observed spectrum of the nth pixel in hyperspectral imagery, and the null hypothesis H_0 and the alternative hypothesis H_1 denote target absence and target presence, respectively. Let $\mathbf{S} = [s_1, s_2, ..., s_N]$ denote the shape pattern of target and is known, $\mathbf{t} = [t_1, t_2, ..., t_B]$ be the spectrum of target and $\mathbf{b} = [b_1, b_2, ..., b_B]$ be the spectrum of background. An observed hyperspectral can be expressed as:

$$\begin{aligned} H_0 &: \mathbf{x}_n = \mathbf{b} + \mathbf{v} \\ H_1 &: \mathbf{x}_n = \mathbf{b} + \mathbf{t} \cdot s_n + \mathbf{v} \end{aligned} \tag{2}$$

If $\mathbf{b}$ and $\mathbf{t}$ follows $N(\boldsymbol{\mu}_b, \boldsymbol{\Sigma}_b)$ and $N(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t)$ respectively, where covariance matrix of background $\boldsymbol{\Sigma}_b$ and that of target $\boldsymbol{\Sigma}_t$ are identical to $\boldsymbol{\Sigma}$, and mean vector of background $\boldsymbol{\mu}_b$ is know, we have the hypothesis testing

$$D_M = \frac{[(\mathbf{X} - \boldsymbol{\mu}_b)\mathbf{S}^T]^T [(\mathbf{X} - \boldsymbol{\mu}_b)(\mathbf{X} - \boldsymbol{\mu}_b)^T]^{-1} ((\mathbf{X} - \boldsymbol{\mu}_b)\mathbf{S}^T)}{\mathbf{S}\mathbf{S}^T} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{3}$$

which is multipixel target version of RXD [2][9]. For single-pixel target version, (3) is simplified to

This work is supported by Specialized Research Foundation for the Doctoral Program of Higher Education of China under Grant 200805611063, and Guangdong Nature Science Foundation under Grant 9251064101000012 and 9451064101002925, the Fundamental Research Funds for the Central Universities, SCUT, under grants 2012ZM0098, and Guangdong Undergraduate Innovative Experiment Project under grants S1010561085.

$$D_S = (\mathbf{X}-\boldsymbol{\mu}_b)^T[(\mathbf{X}-\boldsymbol{\mu}_b)(\mathbf{X}-\boldsymbol{\mu}_b)^T]^{-1}(\mathbf{X}-\boldsymbol{\mu}_b) \underset{H_0}{\overset{H_1}{\gtrless}} \eta \quad (4)$$

If $\boldsymbol{\Sigma}_b$ and $\boldsymbol{\mu}_b$ are estimated within a neighborhood of every test pixel, RXDs shown in (3) and (4) are generalized to detectors that are sensitive to local anomaly.

III. BANDS PARTITION AND SNR OF BAND SUBSETS

RXD is an anomaly detector directly applied to all-band hyperspectral data without considering electromagnetic characteristics in local spectrum, sensor noise level and the transfer characteristics in imaging. Partitioning the whole hyperspectral data into several band-subsets can solve these problems, as well as some computational problems. A widely used method to partition the complete set of bands into several band-subsets is to use correlation coefficients between bands [5], because correlation coefficients measure the similarity of different bands. For the hyperspectral imagery with B bands, the correlation coefficients matrix is defined as

$$\mathbf{R}_{corr}(i,j) = \mathbf{C}(i,j) / \sqrt{\mathbf{C}(i,i)\mathbf{C}(j,j)} \quad (5)$$

for the ith and jth band, where $\mathbf{C}$ is the $L \times L$ covariance matrix with respect to the data matrix $\mathbf{Y}$. Fig. 1 shows the correlation coefficients corresponding to all pairs of bands of OMIS-I hyperspectral imagery. The brighter the pixel is, the larger the correlation coefficient value is.

In terms of Fig. 1, the complete set of bands can be partitioned into several non-overlapped band-subsets. Since correlation coefficients between different bands in same band subset are roughly identical, TNRs of bands in each band-subset can be viewed as being identical. Assume that spectrums of target and background are linearly independent and the number of category of these spectrums is less than that of spectral band. For covariance matrix of hyperspectral data $\boldsymbol{\Sigma}_i$ in ith band subset, there is[7]

$$\boldsymbol{\Sigma}_i = \mathbf{U}_i \boldsymbol{\Lambda}_i \mathbf{U}_i^T + \sigma_i^2 \mathbf{I} \quad (6)$$

where $\mathbf{I}$ is the identity matrix, $\mathbf{U}_i$ is the orthonormal matrix, $\boldsymbol{\Lambda}_i$ is the diagonal and singular matrix and σ_i^2 is the noise level of the nth band data. Then we can obtain the TNR corresponding to the ith band-subset [7].

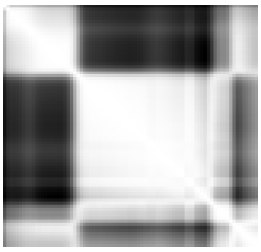

Figure 1. Correlation coefficient matrix for 80 bands OMIS-I imagery

IV. DSMT BASED RX DETECTOR

The Dezert-Smarandache Theory (DSmT) of plausible and paradoxical reasoning proposed in recent years can be considered as an extension of the classical Dempster-Shafer theory (DST) [8]. Both the probability and the DST are under the assumption of exclusivity and exhaustivity of hypotheses. Assume there is a discernment frame $\Theta = \{\theta_1, \theta_2\}$ which is composed of two exhaustive elements. Probability deals with basic probability assignments (BPA) $m(\cdot) \in [0,1]$ such that

$$m(\theta_1) + m(\theta_2) = 1 \quad (7)$$

The DST, on the other hand, deals with basic belief assignments (BBA) $m(\cdot) \in [0,1]$ such that

$$m(\theta_1) + m(\theta_2) + m(\theta_1 \cup \theta_2) = 1 \quad (8)$$

The DSmT theory is only under the assumption of exhaustivity of hypotheses, and deals with the generalized BBA (GBBA) $m(\cdot) \in [0,1]$ such that

$$m(\theta_1) + m(\theta_2) + m(\theta_1 \cup \theta_2) + m(\theta_1 \cap \theta_2) = 1 \quad (9)$$

One of the distinguished advantages of DSmT is the notion of hyper-power set [8]. If let $\Theta = \{\theta_1, \theta_2, ..., \theta_n\}$ be a finite set of n exhaustive elements, hyper-power set D^Θ in the DSmT framework is defined as the set of all composite propositions built from elements of Θ with union operator $\cup$ and intersection operator $\cap$ such that: (1) $\phi, \theta_1, \theta_2, ..., \theta_n \in D^\Theta$;(2) If $A, B \in D^\Theta$, then $A \cap B \in D^\Theta$ and $A \cup B \in D^\Theta$; (3) No other elements belong to D^Θ, except those obtained by using rule (1) or (2). From Θ, we let the mapping $m(\cdot): D^\Theta \to [0,1]$ be as

$$m(\phi) = 0 \text{ and } \sum_{A \in D^\Theta} m(A) = 1 \quad (10)$$

The generalized belief and plausibility function are

$$Bel(A) = \sum_{B \subseteq A, B \in D^\Theta} m(B) \quad (11)$$

$$Pl(A) = \sum_{B \cap A = \phi, B \in D^\Theta} m(B) \quad (12)$$

Two independent sources of evidences over the same frame Θ with belief functions $Bel_1(\cdot)$ and $Bel_2(\cdot)$ associated with GBBA $m_1(\cdot)$ and $m_2(\cdot)$ correspond to the conjunctive consensus of the sources. It is given by

$$\forall C \in D^\Theta,\ m(C) = \sum_{A \cap B = C, A, B \in D^\Theta} m_1(A) m_2(B) \quad (13)$$

Each band-subset split from all-bands hyperspectral data can be treated as an information source. The data of such an information source contains information of local spectral signature. Then, the features for target detection can be extracted from each source. These features reflect the local discriminability of spectral signatures for our detection task.

Let θ_t and θ_b denote target and background, respectively. The value of RXD calculation (before thresholding) of each band subset can be viewed as an evidence over the basic frame $\{\theta_t, \theta_b\}$. Then the corresponding hyper-power set is $\{\theta_t, \theta_b, \theta_t \cap \theta_b, \theta_t \cup \theta_b\}$. Considering that the performance of hyperspectral detector is often significantly affected by the TNR, we define the GBBAs of target, background, uncertainty

and paradox corresponding to the ith band subset respectively as

$$m_i(\theta_t) = TNR_i \cdot \int_{-\infty}^{RXD_i} p_{RXD}(x)dx \quad (14)$$

$$m_i(\theta_b) = TNR_i \cdot (1 - \int_{-\infty}^{RXD_i} p_{RXD}(x))dx \quad (15)$$

$$m_i(\theta_{uncer}) = 1 - TNR_i \quad (16)$$

$$m_i(\theta_{para}) = 0 \quad (17)$$

where θ_{uncer} is $\theta_1 \cup \theta_2$, and θ_{para} is $\theta_1 \cap \theta_2$, $p_{RXD}(x)$ is the probability density function (PDF) of RXD value of ith band-subset and RXD_i is the RXD value of ith band subset. $m_i(\theta_t)$, $m_i(\theta_b)$, $m_i(\theta_{uncer})$ and $m_i(\theta_{para})$ denote supports for target, background, uncertainty of target and background, as well as paradox of target and background from ith band-subset, respectively. Then the support for target from ith and jth band-subset is

$$\begin{aligned} R_{i,j} &= (m_i, m_j) \\ &= m_i(\theta_t)m_j(\theta_t) + m_i(\theta_t)m_j(\theta_{uncer}) + m_i(\theta_{uncer})m_j(\theta_t) \\ &+ m_i(\theta_t)m_j(\theta_{para}) + m_i(\theta_{para})m_j(\theta_t) + m_i(\theta_{uncer})m_j(\theta_{para}) \\ &+ m_i(\theta_{para})m_j(\theta_{uncer}) \\ &\approx m_i(\theta_t)m_j(\theta_t) + m_i(\theta_t)m_j(\theta_{uncer}) + m_i(\theta_{uncer})m_j(\theta_t) \end{aligned} \quad (18)$$

where (,) denotes the operator corresponding to calculation in (13).

For K ($K > 2$) band subsets, we utilize the generalized fusion rule to construct decision function as follows

$$R = (m_K, (m_{K-1}, (..., (m_2, m_1)))) \begin{cases} \geq \lambda & target \\ < \lambda & background \end{cases} \quad (19)$$

to make decision for target detection, where λ is the threshold for decision. Since $\{\theta_t, \theta_b, \theta_t \cap \theta_b, \theta_t \cup \theta_b\}$ is closed under $\cup$ and $\cap$ operators of $\{\theta_t, \theta_b\}$, this combination rule of detection guarantees that R is a proper generalized belief assignment. This combination rule of detection is commutative and associative, and can be used for the fusion of detection value of K band subset.

V. EXPERIMENTAL RESULTS

The data used in this experiment were recorded by OMIS-I over Zaoyuan, China in 2001, from the flight height of about 1,200 meters. The radiance imagery has been calibrated to the reflectance imagery. A portion of the image data, i.e., a scene of size 100×100 pixels is used in our experiments. There are several kinds of natural materials and two car target in the scene. We name these two targets as target 1 and target 2, respectively.

According to Fig. 1, we partition the all-band data into four band subsets: data from the 1st to the 23rd band, data from the 24th to the 66th band, data from the 67th to the 72ed band, and data from the 73rd to the 80th band. These four subband datasets are referred as 1st to 4th band-subset respectively. Next, we adopt covariance matrix of four band-subsets to obtain the TNRs corresponding to these band-subset. These TNRs are 155.2974, 158.1240, 223.9299 and 102.2755, respectively.

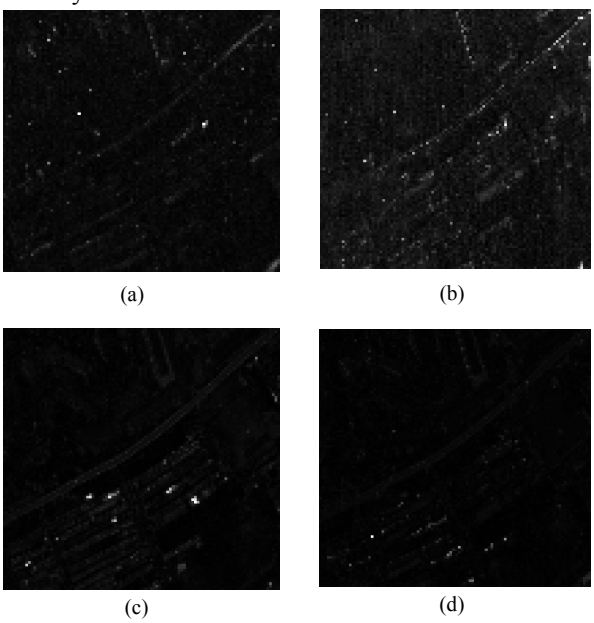

Figure 2. Output of RXD of four band subsets. (a) 1st band subset; (b) 2nd band subset; (c) 3rd band subset; (d) 4th band subset.

The RXD is then applied to each band-subset. Fig. 2(a)-2(d) show the RXD results (before thresholding) of the 4 band subsets respectively. The sparseness of small targets in the whole scene allows the use of the whole hyperspectral data to estimate the statistical characteristic the background.

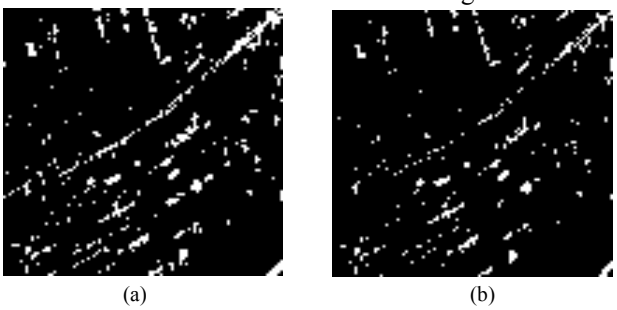

Figure 3. Detection results of RXD applied to all-band hyperspectral data. (a) target 1; (b) target 2.

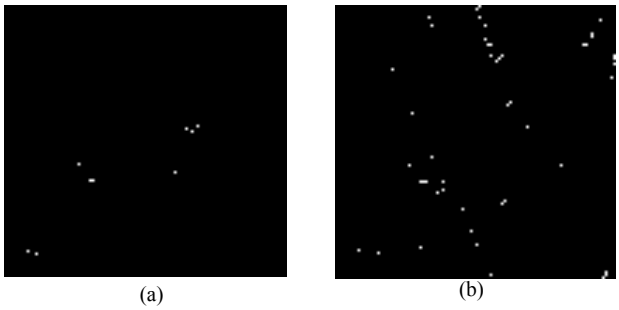

Figure 4. Detection results of DSmt based fusion (a) target 1; (b) target 2.

Fig. 3(a)- Fig. 3(b) show detection results of RXD applied to full-band data. Fig. 4(a)- Fig. 4(b) show detection results utilizing (19). For Fig. 3(a) and Fig. 4(a), the thresholds are set by the value that can detect target 1 with the fewest false alarms, respectively. For Fig. 3(b) and Fig. 4(b), the thresholds are set by the value that can detect target 2 with the fewest false alarms, respectively. Compare Fig 3 with Fig. 4, it can be clearly seen that our method can detect target 1 and target 2 with significantly fewer false alarm than the benchmark RXD which is applied on all-band data. This demonstrates the feasibility and effectiveness of our proposed method.

VI. CONCLUSION

In this paper, we have exploited the DSmT based hyperspectral anomaly detection on real OMIS-I data. Theoretical analysis and the experimental results have shown the effectiveness of our DSmT based detection method and its better performance compared with benchmark RXD applied to all-band hyperspectral data.

REFERENCES

[1] Stein D. W. J., Beaven S. G., Hoff L. E., et al., "Anomaly Detection from Hyperspectral Imagery", IEEE Signal Processing Magazine, 19(1), pp. 58-69, 2002.

[2] Yu, X.L., Reed, I.S., Stocker, A.D., "Comparative performance analysis of adaptive multispectral detectors", IEEE Trans. Signal Process. 41 (8), 2639-2656, 1993.

[3] Chein-I Chang, Shao-Shan Chiang, "Anomaly detection and classification for hyperspectral imagery", IEEE Transactions on Geoscience and Remote Sensing, 40(6), pp. 1314-1325, 2002.

[4] He L, Pan Q, Di W and Li Y Q, "Anomaly detection in hyperspectral imagery by maximum entropy and nonparametric estimation", Pattern Recognition Letters, 29(9): pp. 1392-1403, 2008.

[5] Jia X P, Richards J A, "Segmented principal components transformation for efficient hyperspectral remote-sensing image display and classification", IEEE Trans. Geoscience and Remote Sensing, 37(1), pp. 538-542, 1999.

[6] Di W., Pan Q., He L., and Cheng Y M. "Anomaly Detection in Hyperspectral Imagery by Fuzzy Integral Fusion of Band-subsets", Photogrammetric Engineering & Remote Sensing, 74(2), pp. 201-214, 2008.

[8] Smarandache F. and Dezert J., Advances and Applications of DSmT for Information Fusion, American Research Press, Rehoboth, 2006.