Hyperspectral Anomaly Detection via Global and Local Joint Modeling of Background

Journal:	*Transactions on Signal Processing*
Manuscript ID	Draft
Manuscript Type:	Regular Paper
Date Submitted by the Author:	n/a
Complete List of Authors:	Wu, Zebin; Nanjing University of Science and Technology, School of Computer Science and Engineering; Grenoble Institute of Technology, GIPSA-Lab Zhu, Wei; Duke University, Mathematics Chanussot, Jocelyn; Grenoble Institute of Technology, Xu, Yang; Nanjing University of Science and Technology, School of Computer Science and Engineering; Grenoble Institute of Technology, GIPSA-Lab Osher, Stanley; Dept. of Mathematics, 7619F Math Sciences Building
EDICS:	32. MDS-APPL Applications of multidimensional signal processing < MDS MULTIDIMENSIONAL SIGNAL PROCESSING, 145. SSP-DETC Detection < SSP STATISTICAL SIGNAL PROCESSING

Hyperspectral Anomaly Detection via Global and Local Joint Modeling of Background

Zebin Wu, *Senior Member, IEEE,* Wei Zhu, Jocelyn Chanussot, *Fellow, IEEE,* Yang Xu, *Member, IEEE,* and Stanley Osher

***Abstract*—Anomaly detection is a hot topic in hyperspectral signal processing. The key point of hyperspectral anomaly detection is the modeling of the background. In this paper, we propose a novel anomaly detection method via global and local joint modeling of background. Based on the observation that the local 3D patch belonging to the background in hyperspectral image (HSI) usually lies in a low dimensional manifold, we propose to reconstruct the background part of a HSI from its subsample by scalable low dimensional manifold modeling (SLDMM). Thus the background of HSI can be well characterized in both global and local aspects. Taking into consideration that the SLDMM reconstructs the background part at a low sampling ratio, we propose a multiple random sampling reconstruction strategy to further improve the detection accuracies and robustness. The final background is generated by the mean of backgrounds reconstructed from the multiple random sampling and the anomalies are contained in the residual between the observed HSI and the mean background. Experimental results on three real data sets demonstrate that the proposed anomaly detection method outperforms other state-of-the-art hyperspectral anomaly detection methods.**

***Index Terms*—Hyperspectral image, anomaly detection, low dimensional manifold model, multiple random sampling.**

I. Introduction

ANOMALY detection is one of the most critical signal and image processing tasks in hyperspectral imaging[1–5]. Hyperspectral images(HSIs) provide hundreds of images in wavelengthes covering the visible, near-infrared, and short-wave infrared bands [6–9]. The reliable and nearly continuous spectra in HSI allows accurate measures of the captured scene and provides discriminative information of the ground materials in pixel level. Different materials usually have different electromagnetic energy at different and specific wavelengths. Using this property, we can use HSI to detect anomalous material which has a significantly different spectral signature from their neighboring background clutter pixels in the captured scene [10–13]. Since the last two decades, HSI anomaly detection has been widely used in military and civilian applications .

Many methods have been proposed for HSI anomaly detection [14–16] which can be categorized into two groups: global methods and local methods. In global methods, the HSI is processed as integrated data. We detect the anomalous pixels according to the background statistic calculated from the whole data. The Reed-Xiaoli (RX) [17] detector is regarded as a classical global method. Given the assumption that the background follows a multivariate normal distribution, it estimates the probability of a pixel belonging to the background. In global methods, the background statistic is estimated from the whole image. However, the Gaussian distribution assumption in RX detector could not capture the complexity of the background which is composed of multiple materials. Other improved methods based on RX detector were proposed. The Gaussian mixture model methods (GMMM) [12] uses the mixture of multivariate Gaussian distributions to model the background statistic information. The cluster-based anomaly detection method [18] first segments the image into different homogeneous parts and then detects anomalies in each part. In [19], the weighted-RX and linear-filter-based RX methods are introduced to improve the background information estimation. Kernel based methods such as kernel-RX [20, 21] and support vector data description [22, 23] were proposed to extend the original spectral space to a higher dimensional feature space. In addition, some non-RX based methods are well developed. To separate the anomalies and background pixels, a robust anomaly degree metric [24] is proposed using discriminative information. Kang et. al. proposed an attribute and edge-preserving filters based anomaly detection method [25] which makes full use of the spatial correlations among adjacent pixels. Oleg et al. [26] proposed a dimensionality reduction method that can preserve the anomaly pixels in HSI. The subspace is obtained by minimizing the maximal-norm of misrepresentation residuals. In [27], a random-selection-based anomaly detector was proposed which randomly selected representative background pixels and employed sufficient number of random selections. By this way, the background statistic is purified. Moreover, subspace based methods have attracted more attentions in HSI anomaly detection. The robust principal component analysis (RPCA) based anomaly detector assumes the background is low-rank. Sun et al. [28] proposed the ran-

Z. Wu is with the School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China, with Nanjing Robot Research Institute Co. Ltd, Nanjing, 211135, China, and also with the Grenoble Images Parole Signal Automatique Laboratory, Centre National de la Recherche Scientifique, Grenoble Institute of Technology (Grenoble INP), 38402 Grenoble, France (Email: zebin.wu@gmail.com).

W. Zhu is with Mathematics Department, Duke University, USA((Email: zhu@math.duke.edu).

J. Chanussot is with the Grenoble Images Parole Signal Automatique Laboratory, Centre National de la Recherche Scientifique, Grenoble Institute of Technology (Grenoble INP), 38402 Grenoble, France(Email: jocelyn.chanussot@gipsa-lab.grenoble-inp.fr, jocelyn@hi.is).

Y. Xu is with the School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, 210094, China, and also with the Grenoble Images Parole Signal Automatique Laboratory, Centre National de la Recherche Scientifique, Grenoble Institute of Technology, 38402 Grenoble, France (Email: xuyangth90@njust.edu.cn).

S. Osher is with the Department of Mathematics, University of California, Los Angeles, CA 90095, USA (Email:sjo@math.ucla.edu).

domized subspace learning based method by means of random techniques. Further, Xu et al. [29] proposed the low rank and sparse representation model to separate the background and anomalies. In this model, the background is assumed to lie in the union of multiple subspaces which is more suitable for HSI. Niu et at. [30] improved the LRASR method by a learned dictionary. Then, tensor methods have also been applied in anomaly detection. Xu et al. [31] recovered HSI and detected the anomalies from compressive data by combining tensor nuclear norm and RX detector. In [32], Tucker decomposition is used to factorize the HSI cube into background part and anomaly part.

In the local methods, a sliding double concentric window is used to compute the neighboring background statistic information for the test pixel. The local RX (LRX) detector [33] uses the pixels in the out window to estimate the background statistic. Sparse representation based [2, 34, 35] and collaborative representation based [36] detectors assumes the background dictionary consisting of the the pixels in the out window region, so the background pixels can be well represented by the background dictionary while the anomalies can not. The local summation anomaly detection method [37] combines the multiple local distributions from neighboring local window with spectral-spatial feature. Multiple-window anomaly detection [38] is designed to capture the local spectral variations.

From the previous works, it can be observed that the key point of hyperspectral anomaly detection lies in the modeling of the background. Low rank constraint is a common way to model the background structure. However, low rank can only capture the global correlation of the scene's pixels and assumes the pixels lie in the common subspace spanned by the same bases. In this case, the local structures of HSI are ignored by assuming all the local patches follow the same subspace. In practice, both global low dimensional structure of the background pixels and the variability of local patches should be considered. In this paper, we propose to model the background of HSI in a global and local joint way, by scalable low dimensional manifold model (SLDMM). Unlike the traditional subspace method which models the background pixels as a whole, we extract the local 3D patches and discover their local low dimensional structures. Based on the observation that 3D patches of a HSI's background part typically sample a collection of low dimensional manifolds, we can use the dimension of the patch manifold as a regularization term in a variation functional when reconstructing the background part of a HSI[39–41] . Thus both the global and local low dimensional structures of the background are taken into consideration. The resulted Euler-Lagrange equation can be solved by the point integral method (PIM)[42, 43], or the weighted nonlocal Laplacian [44]. Since a hyperspectral image is a collection of 2D images of the same scene, the spatial similarity matrix can be shared across all the bands which allows to design a fast algorithm to solve the reconstruction problem. In this paper, the reconstruction of background part is depended on the random sampling of the original HSI. To make the detection result robust to random sampling, we propose a multiple random sampling reconstructions strategy and take the mean result of several reconstruction as the final reconstructed background part. Experimental results show that the scalable low dimensional manifold model based detection method achieves higher detection accuracies and is more robust compared to other stated-of-the-art anomaly detection methods.

The remainder of the paper is organized as follows: Section II describes the motivation for the proposed method. In Section III, the proposed hyperspectral anomaly detection method is presented. Section IV experimentally assesses the proposed method and conclusions are reported in Section V.

II. Motivation

The state-of-the-art anomaly detection methods usually rely on the background modeling. The anomalies can be obtained by subtracting the background directly or extracting the background information from the observed data. In global anomaly detection methods, it is assumed that the background part of a HSI is composed of limited ground materials which can be represented in a subspace, and the background is modeled as a whole data. In local anomaly detection methods, it is assumed that the local 3D patch of background is smooth and have similar spectral characteristics. An anomaly pixel is determined by comparing with the neighboring pixels. However, to the best of our knowledge, there is no method having the ability of modeling the background in both global and local aspects. Take into consideration that the background part lies in a low dimensional subspace globally and is composed of smooth local 3D patches, it is important to include both global and local information in the background modeling. Therefore, we propose to model the HSI background by SLDMM, which can efficiently reconstruct the background part at a low sampling ration.

However, since the background is generated based on random sampling, it is possible that some information of the anomalies are sampled at the very first initialization. In this case, the accuracy and robustness of anomaly detection algorithm are greatly affected. It is necessary to design effective strategy to handle this issue. Therefore, a multiple random sampling strategy is proposed to avoid the anomalies contained in the background.

To sum up, Fig.1 graphically illustrates the proposed method.

III. Proposed method

A. SLDMM based background modeling

In the proposed method, the background is modeled by the scalable low dimensional manifold model (SLDMM). Let a hyperspectral image be represented as $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{M\times N\times B}$, where M, N, and B represent the rows, columns and band number of the HSI. For any $\boldsymbol{q} \in \overline{\Omega} = [m] \times [n]$, where $[m] = \{1, 2, \ldots, M\}$, $[n] = \{1, 2, \ldots, N\}$, we define a 3D patch $\mathcal{P}_{\boldsymbol{q}}(\boldsymbol{\mathcal{X}})$ as a 3D block of size $d_1 \times d_2 \times B$ of the original HSI $\boldsymbol{\mathcal{X}}$, where d_1 and d_2 represent the patch size. The 3D patch set $\mathcal{P}(\boldsymbol{\mathcal{X}})$ is defined as the collection of all 3D patches:

$$\mathcal{P}(\boldsymbol{\mathcal{X}}) = \{\mathcal{P}_{\boldsymbol{q}}(\boldsymbol{\mathcal{X}}) : \boldsymbol{q} \in \overline{\Omega}\} \subset \mathbb{R}^d, d = d_1 \times d_2 \times B. \quad (1)$$

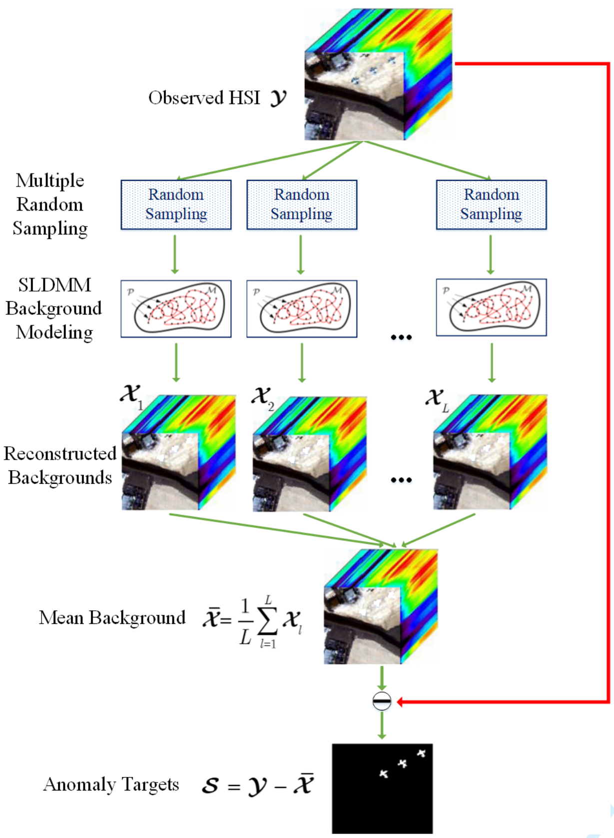

Fig. 1. Flow chart of the proposed anomaly detection method

As discussed in [45], the point cloud $\mathcal{P}(\boldsymbol{\mathcal{X}})$ is typically close to a collection of low dimensional smooth manifolds $\mathcal{M} = \cup_{l=1}^{L}\mathcal{M}_l$ embedded in $\mathcal{R}^d$. We call the collection of manifolds 3D patch manifolds of $\boldsymbol{\mathcal{X}}$. The patch set $\mathcal{P}(\boldsymbol{\mathcal{X}})$ has a trivial 2D parameterization which is given as $\boldsymbol{q} \mapsto \mathcal{P}_{\boldsymbol{q}}(\boldsymbol{\mathcal{X}})$. In this sense, the patch set is locally a 2D sub-manifold embedded in $\mathbb{R}^d$. However, this parameterization is globally not injective, thus it will lead to high curvature variations and self-intersections. For HSI $\boldsymbol{\mathcal{X}}$ without anomalies, the patch manifold $\mathcal{M}(\boldsymbol{\mathcal{X}})$ is always of low dimensionality. In the following analysis, we use $\boldsymbol{\mathcal{X}}$ as a discrete sampling of the continuous function $\widetilde{\boldsymbol{\mathcal{X}}} : [0,1]^3 \to \mathbb{R}$. Specifically, $\boldsymbol{\mathcal{X}}(i,j,k) = \widetilde{\boldsymbol{\mathcal{X}}}(x_i, y_j, t_k)$, where, $(x_i, y_j, t_k) = (i\Delta_x, j\Delta_y, k\Delta_t)$. In the linear mixing model (LMM), only a small collection of constituent elements (endmembers) $e_l \in L^2([0,1]), l = 1, 2, \ldots, K$ are able to generate the entire image $\widetilde{\boldsymbol{\mathcal{X}}} \in L^2([0,1]^3)$. Thus, we have

$$\widetilde{\boldsymbol{\mathcal{X}}}(x,y,t) = \sum_{l=1}^{K}\beta_l(x,y)e_l(t),\ \beta_l(x,y) \geq 0. \tag{2}$$

Denote $\boldsymbol{q} = (x,y)$, then the patch $\mathcal{P}_{(x,y)}\boldsymbol{\mathcal{X}} \in \mathbb{R}^d$ can be written as:

$$\begin{aligned}\mathcal{P}_{(x,y)}\boldsymbol{\mathcal{X}}(i,j,k) &= \widetilde{\boldsymbol{\mathcal{X}}}(x+x_i, y+y_j, t_k)\\ &= \sum_{l=1}^{K}\beta_l(x+x_i, y+y_i)e_l(t_k).\end{aligned} \tag{3}$$

The size of the patches $d_1 \times d_2$ is chosen small enough to be consistent with the spatial resolution of HSI. Thus, the abundance $\beta_l(x+x_i, y+y_i))$ can be approximated by Taylor expansion:

$$\begin{aligned}&\mathcal{P}_{(x,y)}\boldsymbol{\mathcal{X}}(i,j,k)\\ &\approx \sum_{l=1}^{K}\left(\beta_l(x,y) + \frac{\partial\beta_l}{\partial x}(x,y)\cdot x_i + \frac{\partial\beta_l}{\partial y}(x,y)\cdot y_i\right)e_l(t_k)\\ &= \sum_{l=1}^{K}\left(\beta_l(x,y) + i\frac{\partial\beta_l}{\partial x}(x,y)\Delta_x + j\frac{\partial\beta_l}{\partial y}(x,y)\Delta_y\right)e_l(k\Delta_t).\end{aligned} \tag{4}$$

Therefore the underlying local 3D patch manifold $\mathcal{P}(\boldsymbol{\mathcal{X}})$ can be approximated by a manifold of dimension 3K. We can infer that the HSI without anomalies can be reconstructed by using the low dimensionality of the patch manifold as a prior knowledge.

Generally, the background of HSI is composed of limited ground materials. Thus we assume all the 3D patches of the background lie in the same low dimensional manifold. With the 3D patch manifold, we can reconstruct the background part of the HSI from its incomplete observation $\boldsymbol{\mathcal{O}} \in \mathbb{R}^{M\times N\times B}$. Here we assume for any spectral band $t \in [B]$, $\boldsymbol{\mathcal{O}}$ is known on a random subset $\Omega^t \subset \overline{\Omega}$, with a sampling ratio r. According to [41, 45], the dimension of the 3D patch manifold is used as a regularizer to reconstruct $\boldsymbol{\mathcal{X}}$ from $\boldsymbol{\mathcal{O}}$:

$$\begin{aligned}&\min_{\substack{\boldsymbol{\mathcal{X}}\in\mathbb{R}^{M\times N\times B}\\ \mathcal{M}\subset\mathbb{R}^d}} \int_{\mathcal{M}}\dim(\mathcal{M}(\boldsymbol{x}))d\boldsymbol{x} + \lambda\sum_{t=1}^{B}\|\boldsymbol{\mathcal{X}}^t - \boldsymbol{\mathcal{O}}^t\|_{L^2(\Omega^t)^2}\\ &\text{subject to :} \qquad \mathcal{P}(\boldsymbol{\mathcal{X}}) \subset \mathcal{M},\end{aligned} \tag{5}$$

where $\boldsymbol{\mathcal{X}}^t$ is the t-th spectral band of the HSI $\boldsymbol{\mathcal{X}}$, $\mathcal{M}(\boldsymbol{x})$ denotes the smooth manifold $\mathcal{M}_l$ to which $\boldsymbol{x}$ belongs. Also $\int_{\mathcal{M}}\dim(\mathcal{M}(\boldsymbol{x}))d\boldsymbol{x} = \sum_{l=1}^{L}|\mathcal{M}_l|\dim(\mathcal{M}_l)$ is the L^1 norm of the local dimension. According to Proposition 3.1 in [39], the first term in Eq. (5) can be written as the L^2 norm of the coordinate function $\alpha_i^t : \mathcal{M} \to \mathbb{R}$. Moreover, Eq. (5) can be written as:

$$\begin{aligned}&\min_{\substack{\boldsymbol{\mathcal{X}}\in\mathbb{R}^{M\times N\times B}\\ \mathcal{M}\subset\mathbb{R}^d}} \sum_{i=1}^{d_s}\sum_{t=1}^{B}\|\nabla_{\mathcal{M}}\alpha_i^t\|^2_{L^2(\mathcal{M})} + \lambda\sum_{t=1}^{B}\|\boldsymbol{\mathcal{X}}^t - \boldsymbol{\mathcal{O}}^t\|_{L^2(\Omega^t)^2}\\ &\text{subject to :} \qquad \mathcal{P}(\boldsymbol{\mathcal{X}}) \subset \mathcal{M},\end{aligned} \tag{6}$$

where $d_s = d_1 \times d_2$ is the spatial dimension, α_i^t is the coordinate function that maps every point $\boldsymbol{x} = (\boldsymbol{x}_i^t)_{i,t} \in \mathcal{M}$ into its (i,t)-th coordinate p_i^t. Since Eq. (5) is nonconvex, we solve it by alternating the direction of minimization with respect to $\boldsymbol{\mathcal{X}}$ and $\mathcal{M}$. Assume we are in the k-th step and $\mathcal{M}^{(k)}$ and $\boldsymbol{\mathcal{X}}^{(k)}$ are given satisfying $\mathcal{P}(\boldsymbol{\mathcal{X}}^{(k)}) \subset \mathcal{M}^{(k)}$. Thus, we have

1. Fixing $\mathcal{M}^{(k)}$, update $\boldsymbol{\mathcal{X}}^{(k+1)}$ by solving:

$$\begin{aligned}&\min_{\boldsymbol{\mathcal{X}}}\sum_{i,t}\|\nabla_{\mathcal{M}^{(k)}}\alpha_i^t\|^2_{L^2(\mathcal{M}^{(k)})} + \lambda\sum_{t=1}^{B}\|\boldsymbol{\mathcal{X}}^t - \boldsymbol{\mathcal{O}}^t\|_{L^2(\Omega^t)^2}\\ &\text{subject to : } \alpha_i^t(\mathcal{P}\boldsymbol{\mathcal{X}}^{(k)}(\boldsymbol{q})) = \mathcal{P}_i^t\boldsymbol{\mathcal{X}}(\boldsymbol{q}), \quad \boldsymbol{q} \in \overline{\Omega}\end{aligned} \tag{7}$$

where $\mathcal{P}_i^t\boldsymbol{\mathcal{X}}(\boldsymbol{q})$ is the (i,t)-th element in the patch $\mathcal{P}_{\boldsymbol{q}}\boldsymbol{\mathcal{X}}$.

2. Update the manifold $\mathcal{M}^{(k+1)}$ as the image under the perturbed coordinate function $\boldsymbol{\alpha}$:

$$\mathcal{M}^{(k+1)} = \boldsymbol{\alpha}(\mathcal{M}^{(k)}) \tag{8}$$

The manifold update (8) can be easily implemented and Eq. (7) can be implemented using the weighted nonlocal Laplacian (WNLL))[44]. It discretizes the Dirichlet energy $\|\nabla_{\mathcal{M}^{(k)}}\alpha_i^t\|^2_{L^2_(\mathcal{M}^{(k)})}$ as

$$\begin{aligned}
&\frac{|\overline{\Omega}|}{|\Omega_i^t|}\sum_{\boldsymbol{q}\in\Omega_i^t}\sum_{\boldsymbol{s}\in\overline{\Omega}}\overline{w}(\boldsymbol{q},\boldsymbol{s})\left(\alpha_i^t(\mathcal{P}\boldsymbol{\mathcal{X}}^{(k)}(\boldsymbol{q}))-\alpha_i^t(\mathcal{P}\boldsymbol{\mathcal{X}}^{(k)}(\boldsymbol{s}))\right)^2\\
&+\sum_{\boldsymbol{q}\in\overline{\Omega}\setminus\Omega_i^t}\sum_{\boldsymbol{s}\in\overline{\Omega}}\overline{w}(\boldsymbol{q},\boldsymbol{s})\left(\alpha_i^t(\mathcal{P}\boldsymbol{\mathcal{X}}^{(k)}(\boldsymbol{q}))-\alpha_i^t(\mathcal{P}\boldsymbol{\mathcal{X}}^{(k)}(\boldsymbol{s}))\right)^2
\end{aligned} \tag{9}$$

where $\Omega_i^t = \{\boldsymbol{q}\in\overline{\Omega} : \mathcal{P}_i^t\boldsymbol{\mathcal{X}}^{(k)}(\boldsymbol{q})\,\text{is sampled}\}$ is a spatially translated version of Ω^t, $|\overline{\Omega}|/|\Omega_i^t| = 1/r$ is the inverse of the sampling rate. $\overline{w}(\boldsymbol{q},\boldsymbol{s}) = w(\mathcal{P}\boldsymbol{\mathcal{X}}^{(k)}(\boldsymbol{q}),\mathcal{P}\boldsymbol{\mathcal{X}}^{(k)}(\boldsymbol{s}))$ is the similarity between the patches as

$$w(\boldsymbol{u},\boldsymbol{v}) = \exp\left(-\frac{\|\boldsymbol{u}-\boldsymbol{v}\|^2}{\sigma(\boldsymbol{u})\sigma(\boldsymbol{v})}\right), \tag{10}$$

where $\sigma(\boldsymbol{u})$ is the normalizing factor. Combining the WNLL discretization and Eq. (7), the update of $\boldsymbol{\mathcal{X}}$ in (7) can be discretized as

$$\begin{aligned}
\min_{\boldsymbol{\mathcal{X}}}\quad &\lambda\sum_{t=1}^{B}\|\boldsymbol{\mathcal{X}}^t-\boldsymbol{\mathcal{O}}^t\|_{L^2(\Omega^t)}\\
&+\sum_{i,t}[\sum_{\boldsymbol{q}\in\overline{\Omega}\setminus\Omega_i^t}\sum_{\boldsymbol{s}\in\overline{\Omega}}\overline{w}(\boldsymbol{q},\boldsymbol{s})((\mathcal{P}_i^t\boldsymbol{\mathcal{X}}(\boldsymbol{q}))-(\mathcal{P}_i^t\boldsymbol{\mathcal{X}}(\boldsymbol{s})))^2\\
&+\frac{1}{r}\sum_{\boldsymbol{q}\in\Omega_i^t}\sum_{\boldsymbol{s}\in\overline{\Omega}}\overline{w}(\boldsymbol{q},\boldsymbol{s})((\mathcal{P}_i^t\boldsymbol{\mathcal{X}}(\boldsymbol{q}))-(\mathcal{P}_i^t\boldsymbol{\mathcal{X}}(\boldsymbol{s})))^2].
\end{aligned} \tag{11}$$

According to [41, 45], for any given $t\in[B]$, the similarity matrix $\overline{w}$ is the same since it is built on 2D coordinates $\boldsymbol{q},\boldsymbol{s}\in\overline{\Omega}$,. Thus, we only need to solve the following problem:

$$\begin{aligned}
\min_{\boldsymbol{\mathcal{X}}}\quad &\lambda\|\boldsymbol{\mathcal{X}}^t-\boldsymbol{\mathcal{O}}^t\|_{L^2(\Omega^t)}\\
&+\sum_{i=1}^{d_s}\Big[\sum_{\boldsymbol{q}\in\overline{\Omega}\setminus\Omega_i^t}\sum_{\boldsymbol{s}\in\overline{\Omega}}\overline{w}(\boldsymbol{q},\boldsymbol{s})((\mathcal{P}_i\boldsymbol{\mathcal{X}}^t(\boldsymbol{q}))-(\mathcal{P}_i\boldsymbol{\mathcal{X}}^t(\boldsymbol{s})))^2\\
&+\frac{1}{r}\sum_{\boldsymbol{q}\in\Omega_i^t}\sum_{\boldsymbol{s}\in\overline{\Omega}}\overline{w}(\boldsymbol{q},\boldsymbol{s})((\mathcal{P}_i\boldsymbol{\mathcal{X}}^t(\boldsymbol{q}))-(\mathcal{P}_i\boldsymbol{\mathcal{X}}^t(\boldsymbol{s})))^2\Big].
\end{aligned} \tag{12}$$

The Euler-Lagrange equation of Eq. (12) is:

$$\begin{aligned}
0=&\mu\sum_{i=1}^{d_s}\mathcal{P}_i^*I_{\Omega_i^t}\Big[\sum_{\boldsymbol{s}\in\overline{\Omega}}\overline{w}(\boldsymbol{q},\boldsymbol{s})(\mathcal{P}_i\boldsymbol{\mathcal{X}}^t(\boldsymbol{q})-\mathcal{P}_i\boldsymbol{\mathcal{X}}^t(\boldsymbol{s}))\Big]\\
&+\sum_{i=1}^{d_s}\mathcal{P}_i^*\Big[\sum_{\boldsymbol{s}\in\overline{\Omega}}2\overline{w}(\boldsymbol{q},\boldsymbol{s})(\mathcal{P}_i\boldsymbol{\mathcal{X}}^t(\boldsymbol{q})-\mathcal{P}_i\boldsymbol{\mathcal{X}}^t(\boldsymbol{s}))\\
&+\mu\sum_{\boldsymbol{s}\in\Omega_i^t}\overline{w}(\boldsymbol{q},\boldsymbol{s})(\mathcal{P}_i\boldsymbol{\mathcal{X}}^t(\boldsymbol{q})-\mathcal{P}_i\boldsymbol{X}^t(\boldsymbol{s}))\Big]\\
&+\lambda I_{\Omega^t}(\boldsymbol{\mathcal{X}}^t-\boldsymbol{\mathcal{O}}^t),\quad\forall\boldsymbol{q}\in\overline{\Omega}
\end{aligned} \tag{13}$$

where $\mu = 1/r-1$, $\mathcal{P}_i^*$ is the adjoint operator of $\mathcal{P}_i$, I_{Ω^t} is the projection operator that sets $\boldsymbol{\mathcal{X}}^t(\boldsymbol{q})$ to zero for $\boldsymbol{q}\notin\Omega^t$. Using the notation $\boldsymbol{q}_{\widehat{j}}$ to denote j-th component (in the spatial domain) after $\boldsymbol{q}$ in a patch. It has been verified in [41] $\mathcal{P}_i\boldsymbol{\mathcal{X}}^t(\boldsymbol{q}) = \boldsymbol{\mathcal{X}}^t(\boldsymbol{q}_{\widehat{i-1}})$, and $\mathcal{P}_i^*\boldsymbol{\mathcal{X}}^t(\boldsymbol{q}) = \boldsymbol{\mathcal{X}}^t(\boldsymbol{q}_{\widehat{1-i}})$. Similar to [41] and [45], Eq. (13) is rewritten as

$$\begin{aligned}
0=&\mu I_{\Omega_i^t}\Big[\sum_{\boldsymbol{s}\in\overline{\Omega}}\sum_{i=1}^{d_s}\overline{w}(\boldsymbol{q}_{\widehat{1-i}},\boldsymbol{s}_{\widehat{1-i}})(\boldsymbol{\mathcal{X}}^t(\boldsymbol{q})-\boldsymbol{\mathcal{X}}^t(\boldsymbol{s}))\Big]\\
&+\sum_{i=1}^{d_s}\Big[\sum_{\boldsymbol{s}\in\overline{\Omega}}2\overline{w}(\boldsymbol{q}_{\widehat{1-i}},\boldsymbol{s}_{\widehat{1-i}})(\boldsymbol{\mathcal{X}}^t(\boldsymbol{q})-\boldsymbol{\mathcal{X}}^t(\boldsymbol{s}))\\
&+\mu\sum_{\boldsymbol{s}\in\Omega_i^t}\overline{w}(\boldsymbol{q}_{\widehat{1-i}},\boldsymbol{s}_{\widehat{1-i}})(\boldsymbol{\mathcal{X}}^t(\boldsymbol{q})-\boldsymbol{X}^t(\boldsymbol{s}))\Big]\\
&+\lambda I_{\Omega^t}(\boldsymbol{\mathcal{X}}^t-\boldsymbol{\mathcal{O}}^t),\quad\forall\boldsymbol{q}\in\overline{\Omega}
\end{aligned} \tag{14}$$

Denote $\widetilde{w}(\boldsymbol{q},\boldsymbol{s}) = \sum_{i=1}^{d_s}\overline{w}(\boldsymbol{q}_{\widehat{1-i}},\boldsymbol{s}_{\widehat{1-i}})$, Eq. (14) is equal to

$$\begin{aligned}
0=&2\sum_{\boldsymbol{s}\in\overline{\Omega}}\widetilde{w}(\boldsymbol{q},\boldsymbol{s})(\boldsymbol{\mathcal{X}}^t(\boldsymbol{q})-\boldsymbol{\mathcal{X}}^t(\boldsymbol{s}))+\lambda I_{\Omega^t}(\boldsymbol{\mathcal{X}}^t-\boldsymbol{\mathcal{O}}^t)\\
&+\mu I_{\Omega_i^t}\Big[\sum_{\boldsymbol{s}\in\overline{\Omega}}\widetilde{w}(\boldsymbol{q},\boldsymbol{s})(\boldsymbol{\mathcal{X}}^t(\boldsymbol{q})-\boldsymbol{\mathcal{X}}^t(\boldsymbol{s}))\Big]\\
&+\mu\sum_{\boldsymbol{s}\in\Omega^t}\widetilde{w}(\boldsymbol{q},\boldsymbol{s})(\boldsymbol{\mathcal{X}}^t(\boldsymbol{q})-\boldsymbol{X}^t(\boldsymbol{s})),\quad\forall\boldsymbol{q}\in\overline{\Omega}
\end{aligned} \tag{15}$$

Eq. (15) is a linear system for $\boldsymbol{\mathcal{X}}^t\in\mathbb{R}^{MN}$, and the coefficient matrix is not symmetric. To simplify the problem, the similarity matrix $\overline{w}(\boldsymbol{q},\boldsymbol{s})$ is truncated to 20 nearest neighbors. Thus, Eq. (15) is a sparse linear system and can be solved by the generalized minimal residual method (GMRES).

Firstly, background part is usually piece-wise smooth in spatial direction. Pixels in a small window can be seen as local 3D patch which consists of same materials and have similar spectral characteristics. Thus, these local 3D patches are considered to lie on a low dimensional manifold. However, since the spectral signature of anomalies are different from their neighboring background clutter pixels, the local 3D patch which contain the anomalies are not in the low dimensional manifold. Thus it is possible to reconstruct the background by SLDMM with high precision, without containing anomalies.

Secondly, the objective function of SLDMM is the sum of all the 3D patches' manifold. To minimize the objective function, the low dimensional manifold of all 3D patches are considered together which can well characterize the global low dimensional structure of the background.

Last but not least, it is worth noticing that random sampling of the observed HSI is performed in the first step for SLDMM. Although the anomalous pixels take only a small fraction of the whole scene, it is possible that there is anomalous information contained in the samples. Hence, if the sampling ratio is high, it is more likely to contain anomalies in the reconstructed HSI. To avoid this situation, the sampling ratio should be low. However, it is still important to maintain high reconstruction accuracy to well characterize all the background part. Fortunately, SLDMM is proven to have the ability of reconstructing the HSI at a low sampling ratio with high accuracy. Thus, it is reasonable to reconstruct the background part of HSI by SLDMM which can well model the background from both global and local aspects.

B. Strategy of Multiple random sampling reconstruction

It is worth noting that, although we can use random sampling with low ratio for the initialization of SLDMM, it is possible that some information of anomalies are included in the sampling pixels, which greatly affects the effectiveness of our proposed anomaly detection method. Thus, we design a multiple random sampling reconstruction strategy to further improve the accuracy and robustness of the detection method,.

When employing SLDMM for anomaly detection, it is important to make sure that the sampled values are randomly distributed in the whole scene. However, in this case, it is unavoidable to sample the information of anomaly pixels. If we randomly sample the pixels, the anomalous pixels may be sampled. Fortunately, the anomalies are sparsely distributed in the scene and the probability of sampling the anomalous information is very low. Bearing in mind that random technology can greatly improve the detection efficiency[24, 25], we believe the impact of sampling anomalous information can be further descreased with enough times of random sampling and multiple reconstruction. The proposed strategy of multiple random sampling reconstruction are presented in the following steps:

Step 1: Randomly sample the observed HSI $\boldsymbol{\mathcal{Y}}$ L times in a low sampling ratio and obtain the subsampled image $\boldsymbol{\mathcal{O}}_l, l = 1, 2, \ldots, L$.

Step 2: Employ SLDMM to reconstruct each subsample's background part $\boldsymbol{\mathcal{X}}_l$.

Step 3: Compute the final background part by the mean of the L reconstructed HSIs, i.e. $\overline{\boldsymbol{\mathcal{X}}} = \sum_{l=1}^{L} \boldsymbol{\mathcal{X}}_l$.

Step 4: The anomaly part is computed as $\boldsymbol{\mathcal{S}} = \boldsymbol{\mathcal{Y}} - \overline{\boldsymbol{\mathcal{X}}}$ and the anomalies for pixel (i, j) is determined by L_2 norm of the spectrum in $\boldsymbol{\mathcal{S}}$:

$$\boldsymbol{T}(i, j) = \|\boldsymbol{\mathcal{S}}(i, j, :)\|_2. \tag{16}$$

If it is larger than a threshold, pixel (i, j) is claimed to be an anomalous pixel. Then, the proposed method of hyperspectral anomaly detection via scalable low dimensional manifold model (SLDMM-AD) can be summarized as **Algorithm 1**.

Algorithm 1 SLDMM-AD for HSI anomaly detection

Input: The observed HSI $\boldsymbol{\mathcal{Y}} \in \mathbb{R}^{M\times N\times B}$, times of random sampling L

For l=1:L

Random sampled HSI $\boldsymbol{\mathcal{O}}_l$ from $\boldsymbol{\mathcal{Y}}$.

Initial guess $\boldsymbol{\mathcal{X}}_l^{(0)}$

while not convergence **do**

Extract the patch set $\mathcal{P}\boldsymbol{\mathcal{X}}_l^{(k)}$ from $\boldsymbol{\mathcal{X}}_l^{(k)}$.

Compute the similarity matrix on the spatial domain

$$\overline{w}_l(\boldsymbol{q}, \boldsymbol{s}) = w(\mathcal{P}\boldsymbol{\mathcal{X}}_l^{(k)}(\boldsymbol{q}), \mathcal{P}\boldsymbol{\mathcal{X}}_l^{(0)}(\boldsymbol{s})), \quad \boldsymbol{q}, \boldsymbol{s} \in \overline{\Omega}. \tag{17}$$

Assemble the new similarity matrix

$$\widetilde{w}_l(\boldsymbol{q}, \boldsymbol{s}) = \sum_{j=1}^{d_s} \overline{w}_l(\boldsymbol{q}_{\widehat{1-j}}, \boldsymbol{s}_{\widehat{1-j}}) \tag{18}$$

For every spectral band t, update $(\boldsymbol{\mathcal{X}}_l^t)^{(k+1)}$ as the solution of Eq. (15) using GMRES.

$k \leftarrow k + 1$

end while

$\boldsymbol{\mathcal{X}}_l = \boldsymbol{\mathcal{X}}_l^{(k)}$.

End for

Compute mean background: $\overline{\boldsymbol{\mathcal{X}}} = \sum_{l=1}^{L} \boldsymbol{\mathcal{X}}_l$

Compute the anomaly part: $\boldsymbol{\mathcal{S}} = \boldsymbol{\mathcal{Y}} - \overline{\boldsymbol{\mathcal{X}}}$

Output: The detection map $\boldsymbol{T}$ according to Eq. (16)

IV. Experimental Results

A. Data Set Description

In this paper, three real hyperspectral data sets are used to evaluate our proposed method. Two of the data sets are collected by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) sensor over San Diego, CA, USA. After removing the low SNR and water absorption bands, there are 186 bands available in the source image. In the experiment, we use the first 128 bands. Two groups of airplanes appeared in the captured scene are the anomalous targets to be detected. The first group are in the up-left corner of the image. A sub-region of size 128×128 is chosen to be the the first experimental data set which is denoted as SanDiego-1. There are 57 anomaly pixels in total which form three airplanes in the data set. The false color image is shown in Fig. 2(a) and the ground-truth image is shown in Fig. 2(b) where the white pixels represent the anomaly pixles. The second group of airplanes lies in the middle of the complete San Diego image. A sub-region of size 128×128 is chosen as the experimental data set. It contains 120 anomaly pixels forming three airplanes. We denote this data set as SanDiego-2. The false color image and groundtruth map are shown in Fig. 2(b) and (e), respectively.

The third data set is acquired by the Hyperion imaging sensor [24] in 2008. It captured an agricultural area in the State of Indiana, USA. There are 149 bands available after removal of the low SNR and uncalibrated bands. We choose a sub-region of size $128 \times 128 \times 128$ containing 12 anomalous pixels as the experimental data set. The false color image of the scene and the ground-truth map are shown in Fig. 2(c) and (f), respectively.

B. Experimental Details

To evaluate the performance of the proposed method, we demonstrate the global RX(GRX), collaborative representation-based detector (CRD) [36], RPCA-RX, low rank and sparse representation (LRASR) [29]. The RPCA-RX first decomposes the observed HSI into a low-rank component and a sparse error component using RPCA. Then the resulted detection result is obtained by using the RX detector applied on the sparse error component. Similarly, LRASR decomposes the tested HSI into background and anomalies part, but the background is assumed to lie in multiple subspaces. Thus the background part can be

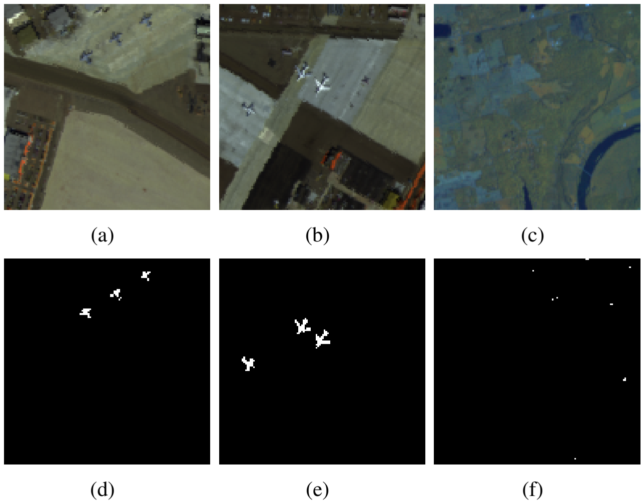

Fig. 2. False color images and ground-truth maps. (a) and (d) SanDiego-1, (b) and (e) SanDiego-2, (c) and (f) Indiana.

represented by a dictionary multiplying the corresponding coefficient matrix.

Empirically, it is easier for SLDMM to converge with a reasonable initialization of $\boldsymbol{\mathcal{X}}^{(0)}$. We use the result of the low-rank matrix completion algorithm APG [46] as an initialization. The initialization obtained by APG can also be considered as the background reconstructed by low-rank matrix completion model. Similar to the proposed method, we use the residual between the tested HSI and the background obtained by APG as the anomaly part. Here we denote this anomaly detection method as APG-AD.

The receiver operating characteristic (ROC) curves [47] are used to measure the performance of the compared methods. A better method would lie nearer to the upper leftmost corner and result in a larger area under the curve. Meanwhile an quantitative index– area under ROC curve (AUC) is calculated for numerical comparison. A higher AUC value indicates the detector has a better detection performance.

In our experiments, the patch size is set to 2×2 and the sampling ratio is 5%. Details about the parameters will be discussed later.

C. Detection Performance

In this section, the anomaly detection performance of the proposed SLDMM-AD is evaluated. All the compared methods and the proposed method are performed on SanDiego-1, SanDiego-2, and Indiana data sets. Fig. 3 illustrates the ROC curves of the proposed SLDMM-AD and other compared methods on the three data sets. For the SanDiego-1 data set, the LRASR curve has the largest probability of detection with a small false alarm rate less than 0.01 and the proposed SLDMM-AD has the second highest probability of detection. With the increasing of false alarm rate, our method achieves the highest probability of detection. This indicates that most part of the anomalies can be detected easily by LRASR but not all the anomalies. Whereas the proposed SLDMM-AD can detect almost all the anomalies at a low false alarm rate. For the SanDiego-2 data set shown in Fig. 3(b), SLDMM-AD has the lowest false alarm rate at 100% probability of detection. And at most of the false alarm rate period, SLDMM-RX has the highest probability of detection. LARSR is close to our method at low false alarm rate but it cannot detect all the anomalies when the false alarm rate is less than 0.1. For the Indiana data set shown in Fig. 3(c). GRX and CRD performs slights better than the proposed method when the false alarm rate is less than 0.005. But the SLDMM-AD reaches 100% at the lowest false alarm rate. Among all the three data sets, the proposed method are better than APG-AD at all false alarm rate. This indicates that the proposed method can better reconstruct the background part than the low-rank model.

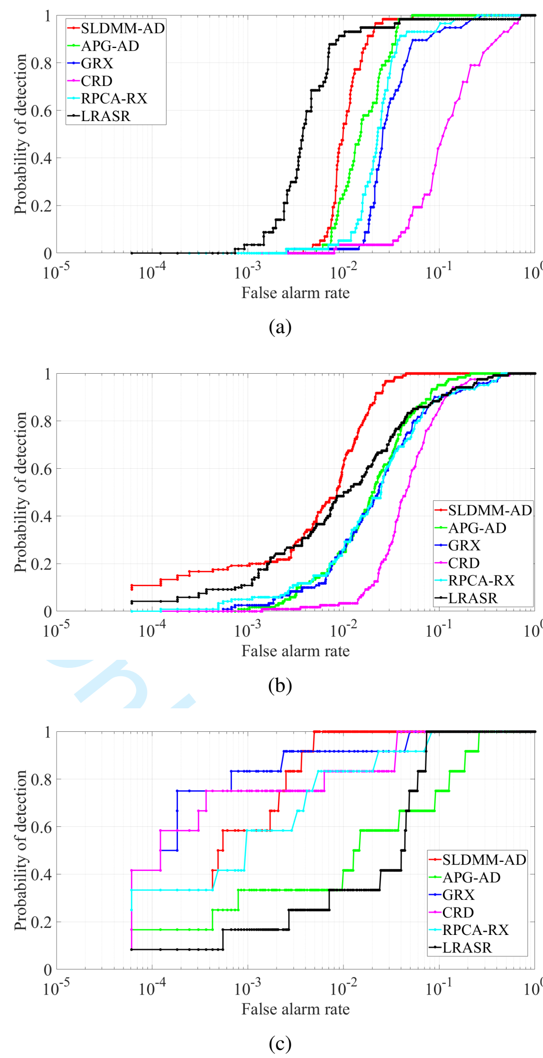

Fig. 3. The logarithmic receiver operating characteristic (ROC) curves on the three HSI data sets. (a) SanDiego-1; (b)SanDiego-2; (c)Indiana.

Table I shows the AUC results of SLDMM-AD and other compared method on the three data sets. The best scores are emphasized in bold for each data set. Fig. 4-6 show the detection maps of all six methods on the SanDiego-1, SanDiego-2 and Indiana data sets with normalized anomaly values between 0 and 1. The SLDMM-AD has the highest AUCs on the three data sets among all the methods. Although the performance of LRASR is relatively stable, it does not provide the best AUC scores in any case. The APG-AD performs well in SanDiego-1 data set. But in SanDiego-2 and Indiana data sets, the SLDMM-AD shows obvious advantage over APG-AD. It can be inferred that the SLDMM-AD is more robust in background modeling. Since SLDMM directly sums the low dimension manifold of every 3D patch instead of modeling the whole image, SLDMM is more flexible than low-rank model. From the detection maps in Fig. 4-6, we can see the proposed method can detect the anomalies and suppress the background pixels simultaneously. In general, the detection maps are consistent with the AUC values reported in Table I. Thus, it is able to conclude that the SLDMM-AD shows competitive performances for hyperspectral anomaly detection.

D. Parameters Analysis

In this section, we investigate the impacts from the patch size and sampling ratio in the detection performance of SLDMM-AD on the three data sets. First, we manually set the range of patch size on the three data sets as $[1, 2, 3, 4, 5, 6]$. Fig. 7 shows the AUC scores of SLDMM-AD on the three data sets with the changing patch size. The results shown that the best patch size is 2×2 which gives the highest AUC scores in all the data sets. Since we assume the local 3D patching lying in the low dimensional manifold, a larger patch size will accordingly increase the dimension of the manifold which will have impacts on the reconstruction of background. Patch size of 1×1 abandon the spectral-spatial structures in HSI which is important in HSI analysis. Thus, we set the patch size as 2×2 in our experiments.

Then we evaluate the impact of sampling ratio in the detection performance of SLDMM-AD. We manually set the range of the sampling ratio as $[0.02, 0.05, 0.08, 0.10, 0.13, 0.15]$. Fig. 8 shows the AUC scores of SLDMM-AD on the tested data sets with the changing sampling size. As can be seen from Fig. 8, for SanDiego-1 data set, the best sampling ratio is 0.05. For the other data sets, when the sampling ratio is larger than 0.05, the AUC scores trend to be stable. In addition, the AUC scores trends to decrease slowly when the sampling ratio is larger than 0.10. Although larger sampling ratio can increase the reconstruction accuracy, the probability of sampling anomaly information will be higher. Therefore, we set the sampling ratio as 0.05 in the experiments.

V. Conclusions

In this paper, a novel hyperspectral anomaly detection method via global and local joint modeling of background is proposed. The background modeling is the key factor for hyperspectral anomaly detection. Assuming that the local 3D patches of background lie in a low dimensional manifold, the background of HSI is effectively characterized via SLDMM. The objective function is to minimize the manifold dimension and can be efficiently solved by a scalable weighted nonlocal Laplacian algorithm. A strategy of multiple random sampling reconstruction is proposed to further improve the robustness and accuracy of hyperspectral anomaly detection. Experimental results on three real HSI data sets show that the proposed method outperforms the other state-of-the-art anomaly detection methods.

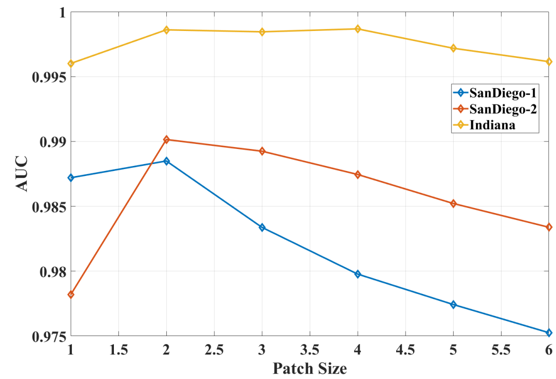

Fig. 7. The curves of area under curve (AUC) with different patch size on the three HSI data sets

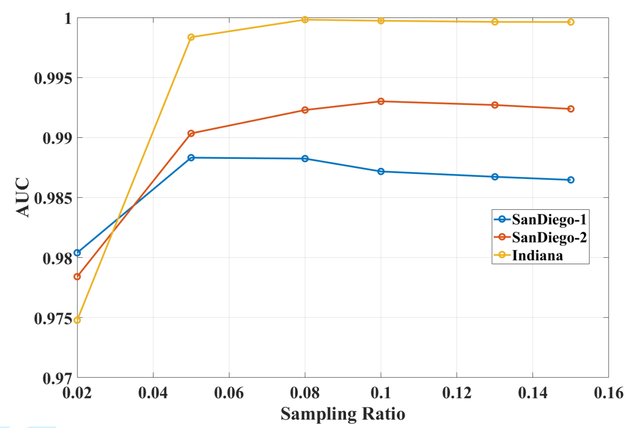

Fig. 8. The curves of area under curve (AUC) with different sampling ratio on the three HSI data sets

References

[1] O. Kuybeda, D. Malah, and M. Barzohar, "Rank estimation and redundancy reduction of high-dimensional noisy signals with preservation of rare vectors," *IEEE Transactions on Signal Processing*, vol. 55, no. 12, pp. 5579–5592, 2007.

[2] R. Bacher, C. Meillier, F. Chatelain, and O. Michel, "Robust control of varying weak hyperspectral target detection with sparse nonnegative representation," *IEEE Transactions on Signal Processing*, vol. 65, no. 13, pp. 3538–3550, 2017.

[3] H. Li and J. H. Michels, "Parametric adaptive signal detection for hyperspectral imaging," *IEEE transactions on signal processing*, vol. 54, no. 7, pp. 2704–2715, 2006.

[4] S. Yang and Z. Shi, "Hyperspectral image target detection improvement based on total variation," *IEEE Transactions on Image Processing*, vol. 25, no. 5, pp. 2249–2258, 2016.

[5] M. Farrell and R. M. Mersereau, "On the impact of covariance contamination for adaptive detection in hyperspectral imaging," *IEEE Signal Processing Letters*, vol. 12, no. 9, pp. 649–652, 2005.

TABLE I
THE COMPARISON IN AREA UNDER CURVE (AUC) BETWEEN SLDMM-AD AND OTHER FIVE COMPARATIVE METHODS ON THE TESTED DATA SETS

Algorithm	SLDMM-AD	APG-AD	GRX	CRD	RPCA-RX	LRASR
SanDiego-1	**0.9882**	0.9816	0.9589	0.8286	0.9689	0.9822
SanDiego-2	**0.9903**	0.9678	0.9490	0.9359	0.9480	0.9610
Indiana	**0.9986**	0.9385	0.9958	0.9935	0.9906	0.9652

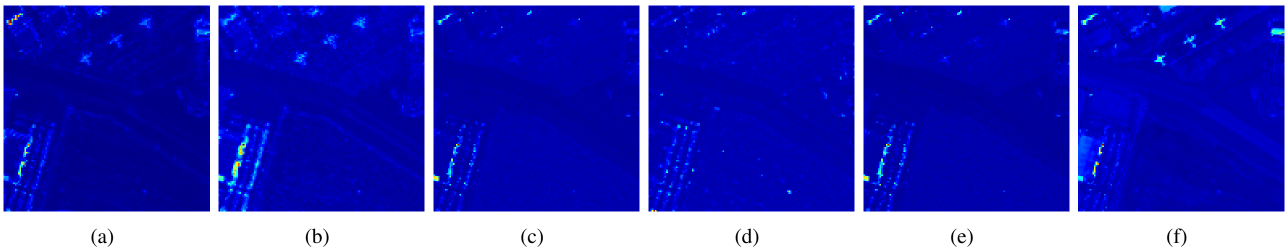

Fig. 4. The normalized detection maps of SanDiego-1 data set (a) SLDMM-AD, (b) APD-AD, (c) GRX, (d) CRD, (e) RPCA-RX, (f) LRASR.

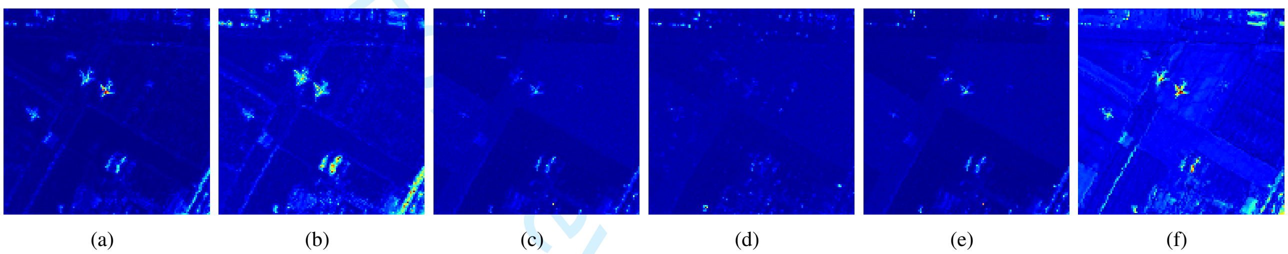

Fig. 5. The normalized detection maps of SanDiego-2 data set (a) SLDMM-AD, (b) APD-AD, (c) GRX, (d) CRD, (e) RPCA-RX, (f) LRASR.

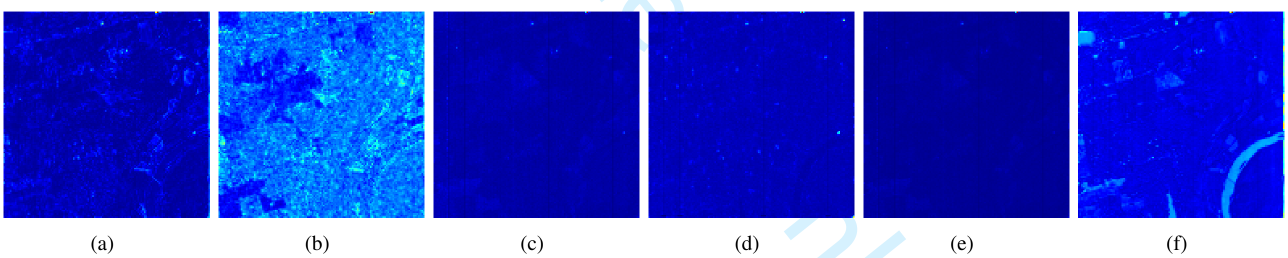

Fig. 6. The normalized detection maps of Indiana data set (a) SLDMM-AD, (b) APD-AD, (c) GRX, (d) CRD, (e) RPCA-RX, (f) LRASR.

[6] Y. Xu, Z. Wu, J. Chanussot, M. Dalla Mura, A. L. Bertozzi, and Z. Wei, "Low-rank decomposition and total variation regularization of hyperspectral video sequences," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 3, pp. 1680–1694, 2018.

[7] R. M. Willett, M. F. Duarte, M. A. Davenport, and R. G. Baraniuk, "Sparsity and structure in hyperspectral imaging: Sensing, reconstruction, and target detection," *IEEE signal processing magazine*, vol. 31, no. 1, pp. 116–126, 2014.

[8] P.-A. Thouvenin, N. Dobigeon, and J.-Y. Tourneret, "Hyperspectral unmixing with spectral variability using a perturbed linear mixing model," *IEEE Transactions on Signal Processing*, vol. 64, no. 2, pp. 525–538, 2016.

[9] Y. Shao, N. Sang, and C. Gao, "Representation space-based discriminative graph construction for semisupervised hyperspectral image classification," *IEEE Signal Processing Letters*, vol. 25, no. 1, pp. 35–39, 2018.

[10] B. Du, Y. Zhang, L. Zhang, and D. Tao, "Beyond the sparsity-based target detector: A hybrid sparsity and statistics-based detector for hyperspectral images," *IEEE Transactions on Image Processing*, vol. 25, no. 11, pp. 5345–5357, 2016.

[11] D. Manolakis and G. Shaw, "Detection algorithms for hyperspectral imaging applications," *IEEE signal processing magazine*, vol. 19, no. 1, pp. 29–43, 2002.

[12] D. W. Stein, S. G. Beaven, L. E. Hoff, E. M. Winter, A. P. Schaum, and A. D. Stocker, "Anomaly detection from hyperspectral imagery," *IEEE signal processing magazine*, vol. 19, no. 1, pp. 58–69, 2002.

[13] D. Manolakis, E. Truslow, M. Pieper, T. Cooley, and M. Brueggeman, "Detection algorithms in hyperspectral imaging systems: An overview of practical algorithms," *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 24–33, 2014.

[14] S. M. Schweizer and J. M. F. Moura, "Hyperspectral

imagery: clutter adaptation in anomaly detection," *IEEE Transactions on Information Theory*, vol. 46, no. 5, pp. 1855–1871, Aug. 2000.

[15] N. M. Nasrabadi, "Regularized spectral matched filter for target recognition in hyperspectral imagery," *IEEE Signal Processing Letters*, vol. 15, pp. 317–320, 2008.

[16] N. Nasrabadi, "Hyperspectral target detection: An overview of current and future challenges," *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 34–44, 2014.

[17] I. S. Reed and X. Yu, "Adaptive multiple-band cfar detection of an optical pattern with unknown spectral distribution," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 10, pp. 1760–1770, 1990.

[18] M. J. Carlotto, "A cluster-based approach for detecting man-made objects and changes in imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 2, pp. 374–387, 2005.

[19] Q. Guo, B. Zhang, Q. Ran, L. Gao, J. Li, and A. Plaza, "Weighted-rxd and linear filter-based rxd: Improving background statistics estimation for anomaly detection in hyperspectral imagery," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 6, pp. 2351–2366, 2014.

[20] H. Kwon and N. M. Nasrabadi, "Kernel rx-algorithm: A nonlinear anomaly detector for hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 2, pp. 388–397, 2005.

[21] H. Goldberg, H. Kwon, and N. M. Nasrabadi, "Kernel eigenspace separation transform for subspace anomaly detection in hyperspectral imagery," *IEEE Geoscience and Remote Sensing Letters*, vol. 4, no. 4, pp. 581–585, 2007.

[22] A. Banerjee, P. Burlina, and C. Diehl, "A support vector method for anomaly detection in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 8, pp. 2282–2291, 2006.

[23] W. Sakla, A. Chan, J. Ji, and A. Sakla, "An svdd-based algorithm for target detection in hyperspectral imagery," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, no. 2, pp. 384–388, 2011.

[24] B. Du and L. Zhang, "A discriminative metric learning based anomaly detection method," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 11, pp. 6844–6857, 2014.

[25] X. Kang, X. Zhang, S. Li, K. Li, J. Li, and J. A. Benediktsson, "Hyperspectral anomaly detection with attribute and edge-preserving filters," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 10, pp. 5600–5611, 2017.

[26] O. Kuybeda, D. Malah, and M. Barzohar, "Anomaly preserving l2 optimal dimensionality reduction over a grassmann manifold," *IEEE Transactions on Signal Processing*, vol. 58, no. 2, pp. 544–552, 2010.

[27] B. Du and L. Zhang, "Random-selection-based anomaly detector for hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 5, pp. 1578–1589, 2011.

[28] W. Sun, L. Tian, Y. Xu, B. Du, and Q. Du, "A randomized subspace learning based anomaly detector for hyperspectral imagery," *Remote Sensing*, vol. 10, no. 3, p. 417, 2018.

[29] Y. Xu, Z. Wu, J. Li, A. Plaza, and Z. Wei, "Anomaly detection in hyperspectral images based on low-rank and sparse representation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 4, pp. 1990–2000, Apr. 2016.

[30] Y. Niu and B. Wang, "Hyperspectral anomaly detection based on low-rank representation and learned dictionary," *Remote Sensing*, vol. 8, no. 4, p. 289, 2016.

[31] Y. Xu, Z. Wu, J. Chanussot, and Z. Wei, "Joint reconstruction and anomaly detection from compressive hyperspectral images using mahalanobis distance-regularized tensor rpca," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 5, pp. 2919–2930, 2018.

[32] X. Zhang, G. Wen, and W. Dai, "A tensor decomposition-based anomaly detection algorithm for hyperspectral image," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 10, pp. 5801–5820, 2016.

[33] W. Sun, C. Liu, J. Li, Y. M. Lai, and W. Li, "Low-rank and sparse matrix decomposition-based anomaly detection for hyperspectral imagery," *Journal of Applied Remote Sensing*, vol. 8, no. 1, p. 083641, 2014.

[34] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Simultaneous joint sparsity model for target detection in hyperspectral imagery," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, no. 4, pp. 676–680, 2011.

[35] S. Paris, D. Mary, and A. Ferrari, "Detection tests using sparse models, with application to hyperspectral data," *IEEE Transactions on Signal Processing*, vol. 61, no. 6, pp. 1481–1494, 2013.

[36] W. Li and Q. Du, "Collaborative representation for hyperspectral anomaly detection," *IEEE Transactions on geoscience and remote sensing*, vol. 53, no. 3, pp. 1463–1474, 2015.

[37] B. Du, R. Zhao, L. Zhang, and L. Zhang, "A spectral-spatial based local summation anomaly detection method for hyperspectral images," *Signal Processing*, vol. 124, pp. 115–131, 2016.

[38] W.-M. Liu and C.-I. Chang, "Multiple-window anomaly detection for hyperspectral imagery," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 2, pp. 644–658, 2013.

[39] S. Osher, Z. Shi, and W. Zhu, "Low dimensional manifold model for image processing," *SIAM Journal on Imaging Sciences*, vol. 10, no. 4, pp. 1669–1690, 2017.

[40] Z. Shi, S. Osher, and W. Zhu, "Generalization of the weighted nonlocal laplacian in low dimensional manifold model," *Journal of Scientific Computing*, vol. 75, no. 2, pp. 638–656, 2018.

[41] W. Zhu, B. Wang, R. Barnard, C. D. Hauck, F. Jenko, and S. Osher, "Scientific data interpolation with low dimensional manifold model," *Journal of Computational Physics*, vol. 352, pp. 213–245, 2018.

[42] Z. Li, Z. Shi, and J. Sun, "Point integral method for solving poisson-type equations on manifolds from point

clouds with convergence guarantees," *Communications in Computational Physics*, vol. 22, no. 1, pp. 228–258, 2017.
[43] Z. Shi and J. Sun, "Convergence of the point integral method for the poisson equation with dirichlet boundary on point cloud," *arXiv preprint arXiv:1312.4424*, 2013.
[44] Z. Shi, S. Osher, and W. Zhu, "Weighted nonlocal laplacian on interpolation from sparse data," *Journal of Scientific Computing*, vol. 73, no. 2-3, pp. 1164–1177, 2017.
[45] W. Zhu, Z. Shi, and S. Osher, "Scalable low dimensional manifold model in the reconstruction of noisy and incomplete hyperspectral images," *arXiv preprint arXiv:1605.05652*, 2016.
[46] K.-C. Toh and S. Yun, "An accelerated proximal gradient algorithm for nuclear norm regularized linear least squares problems," *Pacific Journal of optimization*, vol. 6, no. 615-640, p. 15, 2010.
[47] J. Kerekes, "Receiver operating characteristic curve confidence intervals and regions," *IEEE Geoscience and Remote Sensing Letters*, vol. 5, no. 2, pp. 251–255, 2008.