

联合空间预处理与谱聚类的协同稀疏 高光谱异常检测

成宝芝^{1,2}, 赵春晖³, 张丽丽^{2,3}, 张健沛¹

¹哈尔滨工程大学计算机科学与技术学院, 黑龙江 哈尔滨 150001;

²大庆师范学院机电工程学院, 黑龙江 大庆 163712;

³哈尔滨工程大学信息与通信工程学院, 黑龙江 哈尔滨 150001

摘要 针对利用稀疏表示进行高光谱图像异常目标检测效率不高的问题, 基于高光谱图像成像原理和图像结构, 充分利用高光谱图像的空间特性和光谱特性, 并在它们之间建立协同处理机制, 提出了联合空间预处理与谱聚类的协同稀疏高光谱图像异常目标检测算法。该算法首先对高光谱图像空间特性进行分析, 并结合光谱特性进行空间预处理, 使得处理后的高光谱图像更易于异常目标的检测; 利用建立在谱图划分思想基础上的谱聚类方法进行波段子集划分, 谱聚类方法具有收敛于全局最优解、聚类速度快的特点; 利用提出的新的空间和光谱协同稀疏差异指数方法对每个子集进行异常目标检测, 该协同稀疏方式充分考虑了高光谱图像的空间特性和光谱特性, 通过对每个波段子集检测结果进行叠加, 得到最终异常检测结果。利用真实的 AVIRIS 高光谱图像和合成的高光谱图像对算法进行仿真实验和结果分析, 结果表明该算法具有稳健性, 同时检测精度高, 虚警率低。

关键词 遥感; 高光谱图像; 异常目标检测; 空间预处理; 谱聚类; 协同稀疏

中图分类号 TP751.1

文献标识码 A

doi: 10.3788/AOS201737.0428001

Joint Spatial Preprocessing and Spectral Clustering Based Collaborative Sparsity Anomaly Detection for Hyperspectral Images

Cheng Baozhi^{1,2}, Zhao Chunhui³, Zhang Lili^{2,3}, Zhang Jianpei¹

¹ College of Computer Science and Technology, Harbin Engineering University, Harbin, Heilongjiang 150001, China;

² School of Mechatronics Engineering, Daqing Normal University, Daqing, Heilongjiang 163712, China;

³ College of Information and Communication Engineering, Harbin Engineering University,

Harbin, Heilongjiang 150001, China

Abstract In order to overcome the low efficiency of anomaly detection for hyperspectral images based on sparse representation, a joint spatial preprocessing and spectral clustering based collaborative sparsity anomaly detection algorithm is proposed, which makes full use of the spatial and spectral properties of hyperspectral images, and establishes a cooperative processing mechanism between the spatial and spectral properties, according to the imaging principle and structure of the hyperspectral imagery. The spatial properties of the hyperspectral images are analyzed, and the spatial preprocessing is combined with the spectral properties, which makes the anomalous targets in hyperspectral images easier to be detected. Then, the spectral clustering method based on spectrogram division is used to divide the band subsets, and the spectral clustering method has the features of convergence to the global optimal solution and fast speed. The anomalous targets in each band subset are detected with the proposed new space and spectral collaborative sparsity divergence index method. This collaborative sparsity method considers the spatial and spectral properties of the hyperspectral imagery. Final anomaly detection result is obtained by the superposition of the results of each band subset. The real AVIRIS and synthetic hyperspectral imagery data sets are used for simulations. Simulation results demonstrate that the proposed algorithm is robust, and has higher precision and lower false alarm probability.

收稿日期: 2016-10-13; **收到修改稿日期:** 2016-10-20

基金项目: 国家自然科学基金(61571145)、黑龙江省博士后基金(LBH-Z14062)

作者简介: 成宝芝(1976—), 男, 博士, 副教授, 主要从事高光谱图像处理方面的研究。E-mail: chengbaozhigy@163.com

Key words remote sensing; hyperspectral images; anomaly detection; spatial preprocessing; spectral clustering; collaborative sparsity

OCIS codes 280.4750; 300.6390; 100.4999

1 引 言

高光谱图像是通过高光谱传感器获取地表物质反射的电磁波谱成像得到的,是一个包括二维空间信息和一维光谱信息的 3D 数据集,光谱维含有几十或者成百的连续光谱通道数据,高光谱图像以最大可能的信息量实现对获取地表物质的清晰表达。因此,近年来高光谱图像在地表物质分类、解混、目标检测等方面得到了广泛应用。特别是在不需要先验信息的高光谱图像异常目标检测方面,无论是理论研究还是实际应用中都成为热点。

1990 年以来,高光谱异常目标检测理论研究和应用经历了几个主要阶段,即从线性方法,到非线性核方法,再到稀疏表示的引入。1990 年,Reed 等^[1]首先提出了线性 RX 算法,RX 算法是高光谱图像异常目标检测领域最经典的方法之一,RX 算法通过构造似然比检测算子、估计背景协方差矩阵等得到异常目标检测结果。但是由于高光谱图像具有高维性,算法的检测精度和有效性都较差。针对线性方法的不足,研究者提出了非线性核方法,代表性方法有 Kernel RX(KRX)和支持向量数据描述(SVDD)的异常检测算法。2005 年,Kwon 等^[2]提出了 KRX 异常检测算法,Kernel RX 算法将原始数据通过非线性函数映射到高维特征空间,KRX 算法减小了高光谱图像非线性特性对异常目标检测的影响,但算法的复杂度增大,核函数及其参数的选取对算法的有效性影响较大,因此,研究者在其基础上又提出了若干改进算法。SVDD 算法是由 Banerjee 等^[3]提出的,SVDD 算法进行异常检测是通过分类器完成的,利用最小超球面估计支持区的训练数据,进而寻找到能包围具有共性的一类样本的最小封闭超球体,从而利用判决准则使目标类得到区分^[4],实现异常检测。但是,SVDD 也存在着不足之处,主要是算法在核参数的选择上没有提出合理科学的解决方法,导致核参数选择易受人为因素影响。

近几年,稀疏理论被引入高光谱图像异常目标检测的理论研究和实际应用中,得到了很好的检测效果。2011 年,Chen 等^[5]首先将稀疏表示引入高光谱图像目标检测中,利用压缩感知理论将 l_0 范数最小值问题转化为标准线性规划问题,对目标检测的效果优于一些传统方法;2013 年,赵春晖等^[6]从波段的冗余性和相关性出发,研究了高光谱图像的空间相关性,利用稀疏表示理论,构建了 4-邻域稀疏模型,综合考虑将目标像素及其 4-邻域像素的稀疏表示,提高了检测算法的效果和效率;2014 年,Yuan 等^[7]将稀疏表示用于高光谱异常目标检测,提出了基于局部稀疏差异指数的异常目标检测算法,该算法设定了光谱差异指数和空间差异指数,然后将两个差异指数进行加权,得到一个判决公式,得到了较好的异常检测结果;2016 年,Xu 等^[8]利用低秩矩阵和稀疏表示进行高光谱图像异常目标检测,一个低秩表示被用来模拟背景部分,一个稀疏诱导正则化项被增加用来表示系数,从而对背景和异常目标进行分离,该算法的异常检测结果也较好。Li 等^[9]提出了背景联合稀疏表示的异常目标检测算法,进一步改进了异常检测精度。Peng 等^[10]基于稀疏理论提出了利用稀疏核学习进行特性选择,进而得到异常检测结果。唐意东等^[11]针对高光谱图像目标检测问题,将核方法引入稀疏表示中,构建了自适应核联合表示的异常检测算法。相比于稀疏表示强调系数向量的稀疏性,该方法更强调背景字典中每一原子的作用,并且系数求解更加简单直接。黄远程等^[12]在稀疏表达异常目标检测中引入盲分解方法,首先针对异常检测先验信息不足的问题,采用盲分解方法建立描述背景的冗余字典,然后采用稀疏回归计算每个像元的重建误差,应用局部近邻分析来增强目标在图像邻域的高群表达,从而获得最终的异常检测特征。张丽丽等^[13]基于稀疏差异指数推导出联合核协同方法,用于异常目标检测,也得到了较好的检测结果;赵春晖等^[14]利用稀疏理论,提出了分段式正交匹配追踪的高光谱图像目标检测,对稀疏表示方法用于高光谱图像异常目标检测提出了新的思路。综上所述,基于稀疏表示的异常目标检测解决方案,在一定范围内优于传统的非稀疏方式的异常检测算法,但是也具有一定的局限性,在检测精度和算法有效性等方面有待于进一步优化。

高光谱传感器空间分辨率受到技术条件限制,观测的地物情况的复杂性,以及各类空间环境的干扰,使得高光谱传感器形成的图像在空间特性方面需要进一步处理,以便增加图像后续应用的效果,同时,高光谱

图像数据具有稀疏特性。现有的基于稀疏表示的高光谱图像异常目标检测算法缺少空间特性的有效利用,以及空间特性和光谱特性的有机结合。基于以上分析,提出了联合空间预处理和谱聚类(SPSC)的协同稀疏高光谱图像异常目标检测算法。该算法先通过一种新型的空间预处理方式对高光谱图像进行处理;然后将整个图像按照波段分成若干个子集,这个过程采用谱聚类方法对图像进行聚类处理;进而对每个波段子集利用协同稀疏方法分别完成异常目标检测;最后,叠加每个波段子集异常检测结果,得到最终检测结果。用1组真实高光谱图像和2组合成的高光谱图像数据对提出的算法进行了仿真验证。

2 高光谱图像异常目标检测稀疏表示分析

稀疏表示理论可以很好地解决高维信号处理问题,近年来,在高光谱图像处理上得到广泛应用,稀疏表示理论模型为

$$\mathbf{X} = \mathbf{D}\boldsymbol{\alpha}, \quad (1)$$

式中 $\mathbf{X} = (x_1, \dots, x_i, \dots, x_N) \in \mathbf{R}^N$ 是波段为 L 的高光谱数据, $\mathbf{D} = (d_1, \dots, d_i, \dots, d_K) \in \mathbf{R}^{N \times K}$ ($N \ll K$) 为过完备字典, $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_i, \dots, \alpha_K) \in \mathbf{R}^K$ 为系数,根据稀疏表示理论, $\boldsymbol{\alpha}$ 是稀疏的,即只有少数几个分量是非零的。

高光谱图像异常目标检测可以被视为一种特殊的二元分类,即图像中只有背景类和目标类。因为图像中的像元光谱可以由字典原子的线性组合表示,可以根据表示像元光谱所使用的字典原子属于背景子字典还是目标子字典来判断该像元的归属^[15]。对于任意的待检测像元 \mathbf{X} ,可以假设它处于背景子空间和目标子空间共同张成的空间里,其光谱可以近似地表示为背景子字典 \mathbf{D}_b 和目标子字典 \mathbf{D}_t 中对应的训练样本的线性组合:

$$\begin{aligned} \mathbf{X} &\approx \alpha_1^b d_1^b + \alpha_2^b d_2^b + \dots + \alpha_{N_b}^b d_{N_b}^b + \alpha_1^t d_1^t + \alpha_2^t d_2^t + \dots + \alpha_{N_t}^t d_{N_t}^t = \\ &\underbrace{[d_1^b d_2^b \dots d_{N_b}^b]}_{\mathbf{D}_b} \underbrace{[\alpha_1^b \alpha_2^b \dots \alpha_{N_b}^b]}_{\boldsymbol{\alpha}_b} + \underbrace{[d_1^t d_2^t \dots d_{N_t}^t]}_{\mathbf{D}_t} \underbrace{[\alpha_1^t \alpha_2^t \dots \alpha_{N_t}^t]}_{\boldsymbol{\alpha}_t} = \\ &\underbrace{[\mathbf{D}_b \mathbf{D}_t]}_{\mathbf{D}} \begin{bmatrix} \boldsymbol{\alpha}_b \\ \boldsymbol{\alpha}_t \end{bmatrix} = \mathbf{D}\boldsymbol{\alpha}, \end{aligned} \quad (2)$$

式中字典 \mathbf{D} 是由背景子字典 \mathbf{D}_b 和目标子字典 \mathbf{D}_t 构成的 $L \times (N_b + N_t)$ 的矩阵; $\boldsymbol{\alpha}$ 是对应的未知稀疏系数权向量, $\boldsymbol{\alpha}_b$ 为背景系数权向量, $\boldsymbol{\alpha}_t$ 为目标系数权向量。

对于稀疏表示理论下的高光谱图像异常目标检测问题,不同于其他传统的RX异常检测、KRX异常检测等,不再假设目标和背景分布特性;同时由于背景像元和目标像元本身光谱特性的差别,稀疏系数权向量 $\boldsymbol{\alpha} = \boldsymbol{\alpha}_b + \boldsymbol{\alpha}_t$ 。若 \mathbf{X} 是一个背景像元,则 $\boldsymbol{\alpha}_b$ 是稀疏的, $\boldsymbol{\alpha}_t$ 是一个零向量;若 \mathbf{X} 是一个目标像元,则 $\boldsymbol{\alpha}_b$ 是一个零向量,而 $\boldsymbol{\alpha}_t$ 是稀疏的。因此,根据像元 \mathbf{X} 的稀疏表示其系数向量 $\boldsymbol{\alpha}$ 的非零系数位置,就可以判别该像元是背景还是目标^[6]。求解高光谱像元 \mathbf{X} 的稀疏表示的系数向量 $\boldsymbol{\alpha}$ 就是求解(3)式的最优化问题。

$$\boldsymbol{\alpha} = \operatorname{argmin} \|\boldsymbol{\alpha}\|_1, \text{ subject to } \mathbf{D}\boldsymbol{\alpha} = \mathbf{X}. \quad (3)$$

由于向量 $\boldsymbol{\alpha}$ 的稀疏特性,可以用求解 l_1 范数的最小值问题来解决该最优化问题。

3 研究方法

3.1 空间预处理

高光谱图像是由空间维和光谱维组成的三维立体图像,由于高光谱传感器空间获取技术的限制和空间环境的复杂性,高光谱图像存在着混合像元等情况,异常目标检测效果受到空间特性和光谱特性的共同影响。因此,在对高光谱图像进行异常目标检测之前进行信息预处理是很有必要的,高光谱图像预处理有很多方式,如降维、数据去噪和光谱混合像元分解等,本文采用空间预处理方式进行高光谱图像修正^[16],使得处理后的高光谱图像更易于异常目标的检测。

假设 \mathbf{Y} 是波段为 L 的高光谱图像, \mathbf{M} 是一个 $L \times P$ 的光谱特征矩阵, $\mathbf{M} = [m_1, \dots, m_P]$ 为 P 个端元向量, $\mathbf{E} = [e_1, \dots, e_P]^T$ 为每一列端元向量占有的丰度值向量,高光谱图像数据模型可以写成的矩阵形式为

$$\mathbf{Y} = \mathbf{M}\mathbf{E} + \mathbf{n}, \quad (4)$$

式中 \mathbf{n} 为 L 维噪声或者误差, P 个端元向量和对应的丰度值都是未知量。

定义 $\zeta(i, j)$ 为一个标量权重,

$$\zeta(i, j) = \sum_{r=i-d}^{i+d} \sum_{s=j-d}^{j+d} \beta(r-i, s-j) \cdot \gamma(r-i, s-j), \quad (5)$$

式中

$$\gamma(r-i, s-j) = \gamma[Y(r, s)Y(i, j)], \quad (6)$$

γ 是中心像元 $Y(i, j)$ 和相邻像元 $Y(r, s)$ 之间的相似度计算, $i, j, r, s \in \mathbf{R}^{L \times P}$ 。 d 是空间处理窗口的半径, 在实际计算中, 处理的空间区域是以 $Y(i, j)$ 为中心像元, 尺寸为 $l \times l$ 大小的正方形, 且 l 取奇数, 因此, $d = (l-1)/2$ 。 $\zeta(i, j)$ 的值是通过 β 加权 γ 得到的, β 是一个标量值, 在 $l \times l$ 限定的空间区域内, 可以任意设定 β 值从而使 γ 值加权得到 $\zeta(i, j)$ 。

基于标量权重 $\zeta(i, j)$, 再定义每一个像元的空间衍生加权因子为

$$\eta(i, j) = [1 + \sqrt{\zeta(i, j)}]^2, \quad (7)$$

式中 $\eta(i, j) \geq 1$ 。

由此可得

$$\mathbf{Y}' = \frac{1}{\eta(i, j)}(\mathbf{Y} - \bar{\mathbf{T}}) + \bar{\mathbf{T}}, \quad (8)$$

式中 \mathbf{Y}' 是 \mathbf{Y} 利用加权的空间信息得到的处理后的高光谱图像, 通过空间预处理, 高光谱图像的光谱信息也得到了改善, 将作为后续处理的原数据; $\bar{\mathbf{T}}$ 是高光谱单形体中心, 它是所有像元向量的均值。

3.2 谱聚类

谱聚类方法是一种新的聚类方式, 它建立在谱图划分理论上, 能够收敛于全局最优解^[17], 对样本空间没有要求, 可以是任意形状。本质上谱聚类是一种点对聚类算法, 在谱聚类中, 聚类问题实质转化为图的最优划分问题^[18-19]。

谱聚类算法实现步骤如下。

1) 利用数据样本集(包含 N 个点)构造一个无向权图 \mathbf{G} , \mathbf{G} 的每一个节点对应一个数据点, 用邻接矩阵 \mathbf{W} 表示无向权图 \mathbf{G} , \mathbf{W} 为 $N \times N$ 的相似矩阵, 即

$$\mathbf{W} = \begin{bmatrix} \omega_{11} & \omega_{12} & \cdots & \omega_{1n} \\ \omega_{21} & \omega_{22} & \cdots & \omega_{2n} \\ \vdots & \vdots & & \vdots \\ \omega_{n1} & \omega_{n2} & \cdots & \omega_{nn} \end{bmatrix}。 \quad (9)$$

2) 归一化拉普拉斯矩阵。首先定义 $\boldsymbol{\varphi}$ 为 \mathbf{W} 每一行数据之和构成的对角矩阵, 即

$$\boldsymbol{\varphi} = \begin{bmatrix} \varphi_{11} & 0 & \cdots & 0 \\ 0 & \varphi_{22} & \cdots & 0 \\ \vdots & \vdots & \varphi_{ii} & \vdots \\ 0 & 0 & \cdots & \varphi_{nn} \end{bmatrix}, \quad (10)$$

式中 $\varphi_{ii} = \omega_{i1} + \omega_{i2} + \cdots + \omega_{in} + \cdots + \omega_{in}$ 。由此, 拉普拉斯矩阵 \mathbf{L} 定义为

$$\mathbf{L} = \mathbf{D} - \mathbf{W}, \quad (11)$$

归一化后的拉普拉斯矩阵 \mathbf{L}' 表示为

$$\mathbf{L}' = \mathbf{D}^{-1/2} \mathbf{L} \mathbf{D}^{-1/2} = \mathbf{I} - \mathbf{D}^{-1/2} \mathbf{W} \mathbf{D}^{-1/2}。 \quad (12)$$

3) 求出 \mathbf{L}' 的前 K 个最小特征值及其对应的特征向量 $\{\mathbf{v}_k\}_{k=1}^K$, 将这 K 个特征向量构成一个 $N \times K$ 的矩阵 $\mathbf{v} = [\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k]$ 。

4) 利用 k-means 聚类算法对特征向量进行聚类, 聚类的结果中每一行所属的类别就是原来 \mathbf{G} 中的 N 个数据点分别所属的类别。

3.3 联合空间预处理与谱聚类的协同稀疏异常检测

经过空间预处理之后的高光谱图像, 通过谱聚类将分为若干个波段子集, 然后在每个子集内利用协同稀

疏方法进行异常目标检测^[7],稀疏指数采用将光谱和空间稀疏指数协同加权的方法得到。光谱的稀疏差异指数表示为

$$S_{\text{spectral}} = \sum_{i=1}^N \|\alpha_i^* - \sum_{i=1}^N \alpha_i^* / N\|_2^2 / N, \quad (13)$$

式中 α^* 为权向量。

空间协同的稀疏差异指数表示为

$$S_{\text{spatial}} = \frac{1}{P} \sum_{j=1}^P S_{\text{spatial}}^j, \quad (14)$$

式中 S_{spatial}^j 为第 j 个波段的空间稀疏差异指数, P 为波段总数, 有

$$S_{\text{spatial}}^j = \sum_{i=1}^N \|\beta_i^* - \sum_{i=1}^N \beta_i^* / N\|_2^2 / N, \quad (15)$$

式中 β^* 是第 j 个波段的协同表示的权向量, N 是其维数, 即为第 j 个波段中局部背景字典原子个数。

将(13)、(14)式获得的光谱的稀疏差异指数和空间的稀疏差异指数进行协同稀疏, 得到一个新的光谱和空间协同稀疏加权的差异指数为

$$S_{\text{new}} = \frac{\tau S_{\text{spectral}} + (1 - \tau) S_{\text{spatial}}}{S_{\text{spectral}} * S_{\text{spatial}}^T}, \quad (16)$$

式中 τ 为加权系数, S_{new} 同时反映了光谱相关性和空间相关性。根据每个像元的 S_{new} 值设定阈值, 进行异常目标检测。

联合空间预处理与谱聚类协同稀疏的高光谱图像异常目标检测算法步骤为: 1) 高光谱图像数据 \mathbf{Y} 按照 3.1 节提出的空间预处理方法进行数据处理, 得到空间特性和光谱特性都改善的新的的高光谱图像数据 \mathbf{Y}' ; 2) 对空间预处理后的高光谱图像数据 \mathbf{Y}' 按照 3.2 节所述谱聚类方法进行聚类识别。得到整个高光谱图像数据的波段子集; 3) 对每个波段子集利用(16)式提出的新的协同稀疏差异指数 S_{new} 进行异常目标检测; 4) 最后, 叠加每个波段子集的异常检测结果, 并设定阈值, 得到检测后的二值图像。

4 仿真实验结果与分析

4.1 实验数据源

为了验证提出的 SPSC 的协同稀疏算法的有效性。利用真实的和合成的 3 组高光谱图像进行仿真分析, 第 1 组是美国圣地亚哥海军机场的 AVIRIS 高光谱图像, 如图 1(a) 所示为其第 15 波段的图像, 有 126 个波段, 波长范围为 $0.4 \sim 1.8 \mu\text{m}$, 空间大小为 $100 \text{ pixel} \times 100 \text{ pixel}$, 空间分辨率 3.5 m , 图 1(b) 所示为图 1(a) 区域对应的 38 个异常目标分布情况。

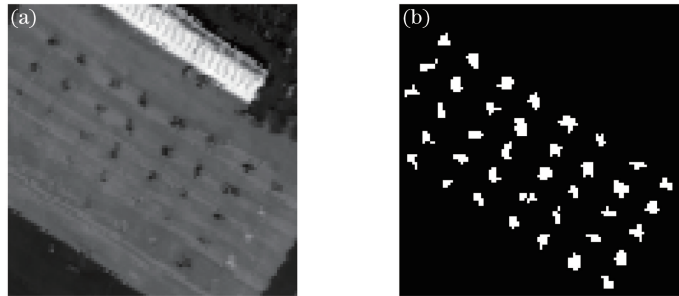


图 1 仿真用高光谱图像和对应的目标分布。(a)第 15 波段 $100 \text{ pixel} \times 100 \text{ pixel}$ 的原始高光谱图像; (b) 对应的目标分布

Fig. 1 Hyperspectral imagery used for simulation and the corresponding targets. (a) Original hyperspectral image of $100 \text{ pixel} \times 100 \text{ pixel}$ in the 15th band; (b) corresponding target distribution

第 2 组图像是一个合成的高光谱图像, 采用的高光谱数据源为 AVIRIS 图像, 如图 2 所示。该合成图像为大数据量高光谱图像, 空间尺寸为 $200 \text{ pixel} \times 200 \text{ pixel}$, 波段数目达到 189 个, 图像在复杂背景信息中含有多个目标点, 根据 Chang 等^[20]所述的图像合成过程可以得到, 合成图像的背景如果没有附加噪声则含有的目标点为 25 个, 这些点以整个图像的中心为排列中心按照 5×5 矩阵的形式排列, 图 2 所示为加入噪声之

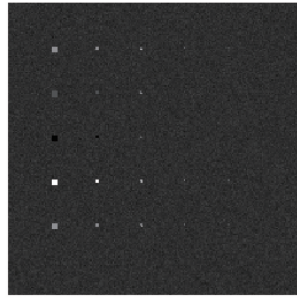


图 2 合成的异常目标检测图像

Fig. 2 Synthetic hyperspectral image for anomalous target detection

后的图像,此时图像可识别的点为 21 个。

第 3 组图像是利用 2001 年 NASA EO-1 卫星获取的博茨瓦纳奥卡万戈三角洲的真实高光谱图像合成的,该图像具有 30 m 的空间分辨率,去除低信噪比波段和被水蒸气吸收的波段后剩余 145 个波段作为高光谱图像处理使用。为了验证 SPSC 算法在大数据量的高光谱图像中的异常检测性能,截取了空间尺寸为 $400 \text{ pixel} \times 250 \text{ pixel}$ 的高光谱图像数据,将其作为异常目标检测的背景图像,然后在该背景图像上叠加目标信号和白噪声,形成包含 7 个异常目标的合成高光谱图像。图 3 所示为利用真实的博茨瓦纳高光谱图像合成的仿真图像,其中图 3(a)为第一波段的合成高光谱图像,图 3(b)为 7 个异常目标的实际分布。

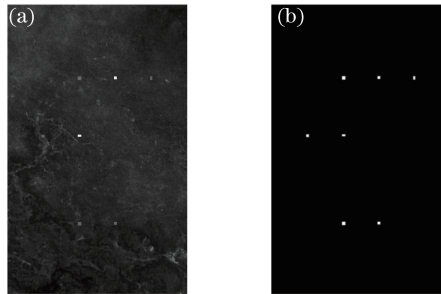


图 3 利用真实的博茨瓦纳高光谱图像合成的仿真图像。(a) 第一波段的合成高光谱图像;(b) 7 个异常目标的实际分布

Fig. 3 Synthetic simulation image based on the real Botswana hyperspectral image.

(a) Synthetic hyperspectral image in the 1st band; (b) actual distribution of the seven anomalous targets

对于 SPSC 算法,为了分析和对比它的有效性和虚警率,同时仿真了利用稀疏表示处理异常目标检测问题比较经典的局部稀疏检测(LSD)算法^[7],以及经典的非线性 KRX 算法和线性 RX 算法,仿真使用的软件是 Matlab R2008b,Windows XP 操作系统,主频为 1.86 GHz,内存为 1 GB。

4.2 真实高光谱图像的仿真分析

对图 1 所列的原始高光谱图像进行空间预处理,预处理之后就可以进行谱聚类得到波段子集,然后对每个波段子集利用(16)式进行协同稀疏分解,最后得到异常目标检测结果,此时设定仿真检测到的总像元数目为 300 个。空间预处理环节需要确定所选空间尺寸的大小,在仿真中共选取 $1 \text{ pixel} \times 1 \text{ pixel}$, $3 \text{ pixel} \times 3 \text{ pixel}$, $5 \text{ pixel} \times 5 \text{ pixel}$, $7 \text{ pixel} \times 7 \text{ pixel}$, $9 \text{ pixel} \times 9 \text{ pixel}$ 五种情况,实验发现 $1 \text{ pixel} \times 1 \text{ pixel}$ 不符合实际图像要求,故去除。其他四种情况的仿真结果二值图像如图 4 所示。

同时,对这四种空间情况利用接收机操作特性(ROC)曲线进行评价^[21],ROC 曲线在检测性能评价中具有重要的作用,其定义为检测概率 P_d 与虚警概率 P_f 之间的变化关系,其中, $P_d = N_{\text{hit}}/N_{\text{target}}$, $P_f = N_{\text{false}}/N_{\text{total}}$, P_d 为检测到的真实目标像元数目 N_{hit} 与地面真实目标像元数目 N_{target} 的比值; P_f 为检测到的虚警像元数目 N_{false} 与整幅图像像元数目总和 N_{total} 的比值^[22]。ROC 是衡量检测算法性能优劣的一个重要指标,ROC 曲线如图 5 所示。

通过图 4 和图 5 四种不同空间大小情况下异常检测结果的分析,能够发现当空间尺寸取 $7 \text{ pixel} \times 7 \text{ pixel}$ 进行高光谱图像预处理时,图 5,图 6,表 1 SPSC 算法异常检测性能最好。因此,选取 $7 \text{ pixel} \times 7 \text{ pixel}$ 大小的空间尺寸对原始高光谱图像进行空间预处理,处理后得到的图像如图 6 所示。

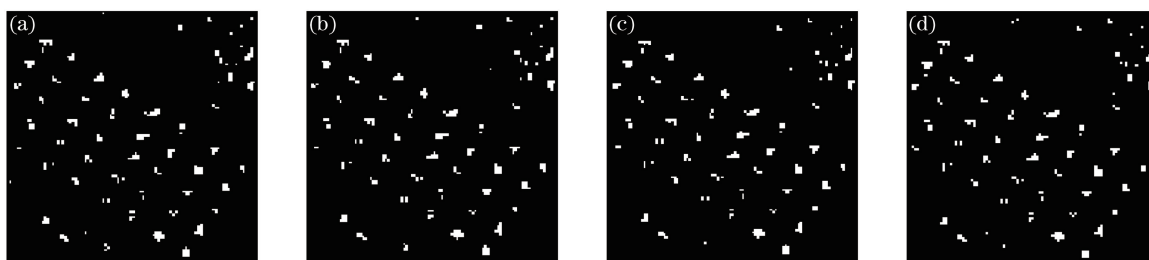


图 4 四种不同空间大小的异常检测结果。(a) 3 pixel×3 pixel; (b) 5 pixel×5 pixel; (c) 7 pixel×7 pixel; (d) 9 pixel×9 pixel

Fig. 4 Anomaly detection results of four different spatial sizes. (a) 3 pixel×3 pixel; (b) 5 pixel×5 pixel; (c) 7 pixel×7 pixel; (d) 9 pixel×9 pixel

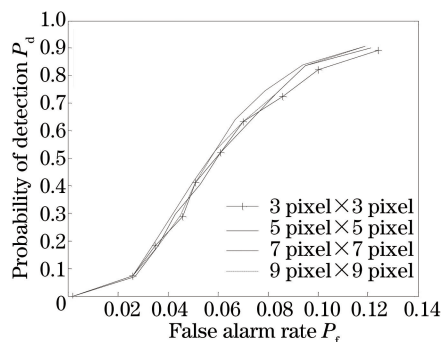


图 5 四种不同空间大小的异常检测结果的 ROC 比较
Fig. 5 ROC comparison of the anomaly detection results of four different spatial sizes

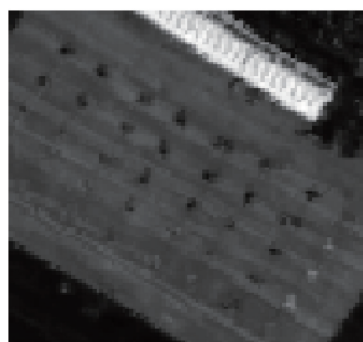


图 6 空间预处理后的高光谱图像
Fig. 6 Hyperspectral imagery after spatial preprocessing

表 1 六个波段子集的划分

Table 1 Division of the six band subsets

Serial No.	1	2	3	4	5	6
Band No.	1-12	13-29	30-60	61-72	73-93	94-126

对空间预处理后的高光谱图像进行谱聚类, 根据实际图像情况, 利用谱聚类将图像分为六个波段子集, 如表 1 所示。

在每个波段子集内利用(16)式进行协同稀疏异常检测, 有两个关键的问题需要考虑, 一是双窗口大小的选择, 二是加权系数 τ 的选取。通过对双窗口大小和加权系数 τ 的实验对比, 选取内 5 pixel×5 pixel(外 15 pixel×15 pixel), $\tau=0.1$ 的情况, 作为 SPSC 算法中稀疏分解使用的参数。检测结果如图 7 所示, 同时也仿真了 LSD 算法, KRX 算法和 RX 算法进行比较, 设定此时仿真检测到的总的像元数目为 400。

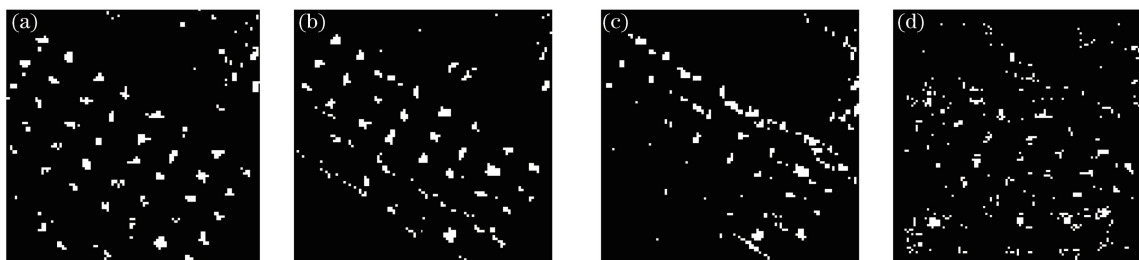


图 7 四种算法的异常检测结果。(a) SPSC 算法; (b) LSD 算法; (c) KRX 算法; (d) RX 算法
Fig. 7 Anomaly detection results of the four algorithms. (a) SPSC algorithm; (b) LSD algorithm; (c) KRX algorithm; (d) RX algorithm

从图 7 可以得到, SPSC 算法相比于其他算法对异常目标的检测精度较高, 虚警率较低。LSD 算法对于高光谱图像异常检测也具有较好的稳健性和有效性, 但是检测精度不高, 说明算法本身还需要进一步改进;

KRX 算法是经典的非线性异常检测算法,但是算法本身仅考虑了光谱信息存在的问题,忽略了空间信息对图像的影响,以及波段的冗余性和相关性对算法效果的限制,导致检测异常目标能力不强;RX 算法是一种基于广义似然比检验的恒虚警率线性异常目标检测,但是高光谱图像波段间存在非相关性,因此检测精度和虚警率都不理想。

图 7 从定性的角度对各类算法的检测性能进行了比较,下面从定量的角度进一步分析。算法能检测到的异常目标数、总像元中目标数和虚警数是三个重要的性能评定指标,因此,设定检测的总的像元数目为 400,进行性能比对。结果如表 2 所示,SPSC 算法能检测到 38 个异常目标,目标数为 307 pixel,虚警数为 93 pixel;LSD 算法能检测到 35 个异常目标,目标数为 263 pixel,虚警数为 137 pixel;KRX 算法能检测到 34 个异常目标,目标数为 199 pixel,虚警数为 201 pixel;RX 算法能检测到 23 个异常目标,目标数为 159 pixel,虚警数为 241 pixel。由表 2 的定量比较可得到,SPSC 算法检测性能最优,该算法的漏检率降低,虚警率低,算法检测精度明显优于其他几种算法。

表 2 为检测像元总数固定时得到的检测结果,为了说明像元总数变化时算法的检测性能情况,根据实际情况选取像元数为 50~300,如图 8 所示,随着检测的像元总数的增加,SPSC 算法检测能力越来越强,对目标的识别能力也远远优于其他算法,同时也能发现算法具有较强的稳健性。

表 2 四种检测算法定量比较
Table 2 Quantitative comparison of the four detection algorithms

Algorithm	Anomaly number	Target number /pixel	False alarm number /pixel
SPSC	38	307	93
LSD	35	263	137
KRX	34	199	201
RX	23	159	241

利用 ROC 指标对 SPSC 算法、LSD 算法、KRX 算法和 RX 算法的检测性能进行分析,从图 9 曲线中能够发现,SPSC 算法的性能与检测的像元总数有紧密联系,当检测的像元总数较少时,SPSC 算法性能稍低;当检测的像元总数较多时,对虚警率要求不高,当虚警率达到 0.055 以上时,检测性能变得突出,这与前文各项指标对 SPSC 算法的分析是一致的。

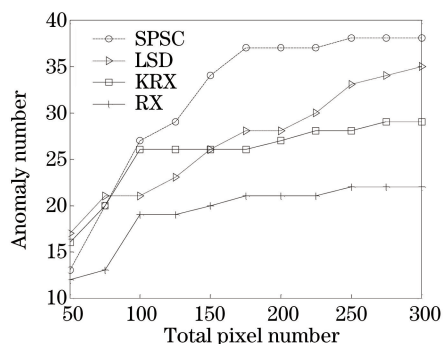


图 8 四种算法检测到异常目标数目的比较
Fig. 8 Comparison of the detected anomaly number by the four algorithms

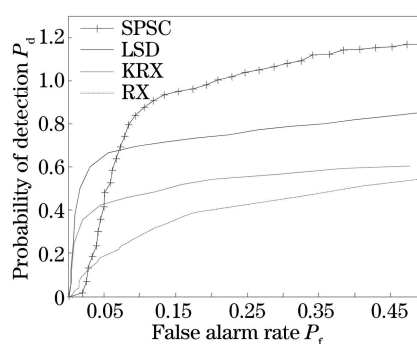


图 9 四种异常检测算法的 ROC 比较
Fig. 9 ROC comparison of the four anomaly detection algorithms

由此,说明算法采用的空谱结合处理策略与协同机制是可行的,通过对高光谱图像预先进行空间信息和光谱信息处理,使得图像更有利于从复杂背景中识别异常目标信息。

4.3 合成高光谱图像的仿真实验分析

合成的高光谱图像如图 2 和图 3 所示,利用 SPSC 算法对其进行目标检测,即先通过空间预处理,然后再进行谱聚类,最后利用波段子集内协同稀疏得到异常检测结果。其他几种用于对比的算法也在相同条件下进行了相应的仿真。

4.3.1 第 2 组合成图像的仿真分析

图 2 所示合成图像的空间尺寸为 200 pixel×200 pixel,波段数是 189 个。从图 10 的二值仿真图像性能

定量比较可得,对于大数据背景下异常目标个数的检测,SPSC 算法检测性能最优,在 21 个能够被检测的目标点中有 16 个可以被发现;LSD 算法是一种局部稀疏异常指数检测算法,只能检测到 8 个异常目标;传统的 RX 异常检测算法虽然能检测到 14 个异常目标,但是该算法检测的稳健性不强;KRX 算法是一种基于核方法的非线性异常检测算法^[23-25],但对于大数据量的图像处理效果最差,仅检测到 5 个异常目标。所以,从合成的高维图像来看,SPSC 算法具有稳健性强、检测精度高、虚警率低的特点,但是异常目标检测计算时间较长。

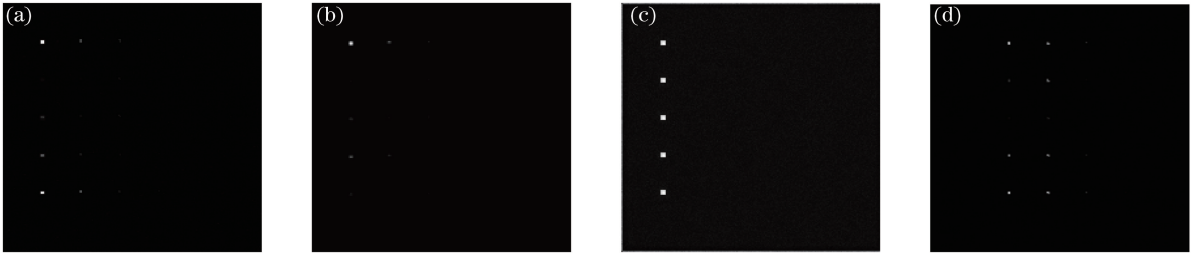


图 10 四种算法的异常检测结果。(a) SPSC 算法; (b) LSD 算法; (c) KRX 算法; (d) RX 算法

Fig. 10 Anomaly detection results of the four algorithms. (a) SPSC algorithm;
(b) LSD algorithm; (c) KRX algorithm; (d) RX algorithm

4.3.2 第 3 组合成图像的仿真分析

第 3 组合成高光谱图像如图 3 所示,它的空间尺寸为 $400 \text{ pixel} \times 250 \text{ pixel}$,波段数是 145,是一个典型的高维大数据量高光谱图像。利用 SPSC 算法得到的异常检测结果如图 11 所示,同时也仿真了 LSD 算法, KRX 算法和 RX 算法。截取的高光谱合成图像达到了 $400 \times 250 \times 145 = 14500000 \text{ pixel}$,而且图像背景为 NASA EO-1 卫星获取的真实地物场景高光谱图像,因此背景数据信息较复杂。为了使合成的图像更接近真实场景存在的异常目标信息,叠加异常目标数据时引入了高斯白噪声^[26]。通过仿真结果比较可知 SPSC 算法检测结果明显优于其他几种算法,从仿真图像中可以看出,对于人为添加到复杂背景上的 7 个异常目标,在引入高斯白噪声的情况下,SPSC 算法仍然能检测到 6 个异常目标。而 LSD 算法、KRX 算法和 RX 算法对这种复杂背景下的大数据量高光谱图像异常目标检测能力明显不足,导致检测效果不佳。

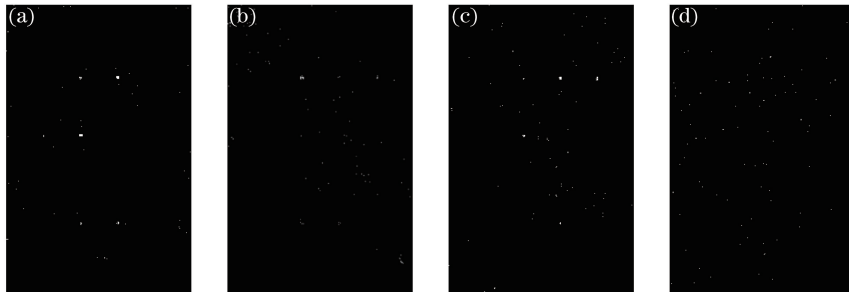


图 11 第 3 组合成高光谱图像的异常检测仿真结果。(a) SPSC 算法; (b) LSD 算法; (c) KRX 算法; (d) RX 算法

Fig. 11 Anomaly detection simulation results of the 3rd synthetic hyperspectral image. (a) SPSC algorithm;
(b) LSD algorithm; (c) KRX algorithm; (d) RX algorithm

通过对以上 3 组高光谱图像的仿真分析,SPSC 算法无论对真实图像还是合成图像,特别是复杂背景下的高维大数据高光谱图像,都有较好的检测性能,具有稳健性强的特点。

5 结 论

稀疏表示用于高光谱图像异常目标检测是研究热点,有许多需要解决的问题。在分析高光谱图像异常目标检测原理和高光谱图像特点的基础上,提出了 SPSC 的协同稀疏异常目标检测算法,该算法利用了高光谱三维立体图像的特点,对高光谱图像空间特性进行了分析,并在空间维特性和光谱维特性之间建立联动机制,从而完成了高光谱图像的空间预处理,使处理后的图像更适合于异常目标检测。谱聚类方法的引入使波段子集划分更加合理和有效,而且谱聚类具有聚类速度快的特点,提高了算法整体的时效性。基于空间和光

谱特性的协同稀疏检测方法,将高光谱图像的空间特性和光谱特性进行了充分协同利用。与 LSD 算法、KRX 算法、RX 算法相比,SPSC 算法在检测精度和虚警率方面都明显优于其他算法,而且算法结构简单。但是 SPSC 算法在处理高维的高光谱图像数据时,需要优化算法的处理机制,使得算法处理速度进一步提高。

参 考 文 献

- [1] Reed I S, Yu X. Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution[J]. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1990, 38(10): 1760-1770.
- [2] Kwon H, Nasrabadi N M. Kernel RX-algorithm: a nonlinear anomaly detector for hyperspectral imagery[J]. IEEE Transactions on Geoscience and Remote Sensing, 2005, 43(2): 388-397.
- [3] Banerjee A, Burlina P, Diehl C. A support vector method for anomaly detection in hyperspectral imagery[J]. IEEE Transactions on Geoscience and Remote Sensing, 2006, 44(8): 2282-2291.
- [4] Chen Derong, Gong Jiulu, He Guanglin, *et al.* A RFS-SVDD algorithm for hyperspectral global anomaly detection[J]. Journal of Astronautics, 2010, 31(1): 228-232.
 谌德荣, 宫久路, 何光林, 等. 高光谱图像全局异常检测 RFS-SVDD 算法[J]. 宇航学报, 2010, 31(1): 228-232.
- [5] Chen Y, Nasrabadi N M, Tran T D. Sparse representation for target detection in hyperspectral imagery[J]. IEEE Journal of Selected Topics in Signal Processing, 2011, 5(3): 629-640.
- [6] Zhao Chunhui, Li Xiaohui, Zhu Haifeng. Hyperspectral imaging target detection algorithm based on spatial 4 neighborhoods for sparse representation[J]. Journal of Harbin Engineering University, 2013, 34(9): 1171-1178.
 赵春晖, 李晓慧, 朱海峰. 空间 4-邻域稀疏表示的高光谱图像目标检测[J]. 哈尔滨工程大学学报, 2013, 34(9): 1171-1178.
- [7] Yuan Z Z, Sun H, Ji K F, *et al.* Local sparsity divergence for hyperspectral anomaly detection[J]. IEEE Geoscience and Remote Sensing Letters, 2014, 11(10): 1697-1701.
- [8] Xu Y, Wu Z B, Li J, *et al.* Anomaly detection in hyperspectral images based on low-rank and sparse representation[J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(4): 1990-2000.
- [9] Li J Y, Zhang H Y, Zhang L P, *et al.* Hyperspectral anomaly detection by the use of background joint sparse representation[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2015, 8(6): 2523-2533.
- [10] Peng Z M, Gurram P, Kwon H, *et al.* Sparse kernel learning-based feature selection for anomaly detection[J]. IEEE Transactions on Aerospace and Electronic Systems, 2015, 51(3): 1698-1716.
- [11] Tang Yidong, Huang Shucai, Ling Qiang, *et al.* Adaptive kernel collaborative representation anomaly detection for hyperspectral imagery[J]. High Power Laser and Particle Beams, 2015, 27(9): 091008.
 唐意东, 黄树彩, 凌 强, 等. 高光谱图像自适应核联合表示异常检测[J]. 强激光与粒子束, 2015, 27(9): 091008.
- [12] Huang Yuancheng, Zhong Yanfei, Zhao Yehe, *et al.* Joint blind unmixing and sparse representation for anomaly detection in hyperspectral image[J]. Geomatics and Information Science of Wuhan University, 2015, 40(9): 1144-1150.
 黄远程, 钟燕飞, 赵野鹤, 等. 联合盲分解与稀疏表达的高光谱图像异常目标检测[J]. 武汉大学学报·信息科学版, 2015, 40(9): 1144-1150.
- [13] Zhang Lili, Zhao Chunhui, Cheng Baozhi. A joint kernel collaborative representation based approach for anomaly target detection of hyperspectral images[J]. Journal of Optoelectronics & Laser, 2015, 26(11): 2154-2161.
 张丽丽, 赵春晖, 成宝芝. 基于联合核协同的高光谱图像异常目标检测[J]. 光电子·激光, 2015, 26(11): 2154-2161.
- [14] Zhao Chunhui, Jing Xiaohao, Li Wei. Hyperspectral image target detection algorithm based on StOMP sparse representation[J]. Journal of Harbin Engineering University, 2015, 36(7): 992-996.
 赵春晖, 靖晓昊, 李 威. 基于 StOMP 稀疏方法的高光谱图像目标检测[J]. 哈尔滨工程大学学报, 2015, 36(7): 992-996.
- [15] Song Xiangfa, Jiao Licheng. Classification of hyperspectral remote sensing image based on sparse representation and spectral information[J]. Journal of Electronics & Information Technology, 2012, 34(2): 268-272.
 宋相法, 焦李成. 基于稀疏表示及光谱信息的高光谱遥感图像分类[J]. 电子与信息学报, 2012, 34(2): 268-272.
- [16] Zortea M, Plaza A. Spatial preprocessing for end member extraction[J]. IEEE Transactions on Geoscience and Remote Sensing, 2009, 47(8): 2679-2693.

- [17] Luxburg U V. A tutorial on spectral clustering[J]. *Statistics and Computing*, 2007, 17(4): 395-416.
- [18] Cai Xiaoyan, Dai Guanzhong, Yang Libin. Survey on spectral clustering algorithms[J]. *Computer Science*, 2008, 35(7): 14-18.
蔡晓妍, 戴冠中, 杨黎斌. 谱聚类算法综述[J]. *计算机科学*, 2008, 35(7): 14-18.
- [19] Zheng Xiumeng, Chen Fucui, Huang Ruiyang. Research on collaborative recommendation algorithms based on parallel spectral clustering[J]. *Journal of University of Science and Technology of China*, 2016, 46(1): 82-86.
郑修猛, 陈福才, 黄瑞阳. 基于并行化谱聚类的协同推荐算法研究[J]. *中国科学技术大学学报*, 2016, 46(1): 82-86.
- [20] Chang C I, Jiao X L, Wu C C, *et al.* Component analysis based unsupervised linear spectral mixture analysis for hyperspectral imagery[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2011, 49(11): 4123-4137.
- [21] Gao G. A parzen-window-kernel-based CFAR algorithm for ship detection in SAR images[J]. *IEEE Geoscience and Remote Sensing Letters*, 2011, 8(3): 557-561.
- [22] Zou J Y, Li W, Du Q. Sparse representation-based nearest neighbor classifiers for hyperspectral imagery[J]. *IEEE Geoscience and Remote Sensing Letters*, 2015, 12(12): 2418-2422.
- [23] Zhang H Y, Zhai H, Zhang L P, *et al.* Spectral-spatial sparse subspace clustering for hyperspectral remote sensing images[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, 54(6): 3672-3684.
- [24] Du B, Zhao R, Zhang L P, *et al.* A spectral-spatial based local summation anomaly detection method for hyperspectral images[J]. *Signal Processing*, 2016, 124: 115-131.
- [25] Ye Zhen, Bai Lin, Nian Yongjian. Hyperspectral image classification algorithm based on gabor feature and locality-preserving dimensionality reduction[J]. *Acta Optica Sinica*, 2016, 36(10): 1028003.
叶 珍, 白 璘, 粘永健. 基于 Gabor 特征与局部保护降维的高光谱图像分类算法[J]. *光学学报*, 2016, 36(10): 1028003.
- [26] Xue Peng, Wang Zhibin, Zhang Rui, *et al.* Highly efficient measurement technology based on hyperspectropolarimetric imaging[J]. *Chinese J Lasers*, 2016, 43(8): 0811001.
薛 鹏, 王志斌, 张 瑞, 等. 高光谱全偏振成像快捷测量技术研究[J]. *中国激光*, 2016, 43(8): 0811001.