

基于空间密度聚类的改进 KRX 高光谱异常检测

刘春桐, 马世欣*, 王浩, 汪洋, 李洪才

火箭军工程大学导弹发射与定向瞄准技术军队重点实验室, 陕西 西安 710025

摘要 高光谱遥感影像包含了丰富的光谱信息, 对于地物具有极强的分辨能力, 从而促进了不需任何先验信息的高光谱异常目标探测技术的发展。KRX(Kernel RX)异常探测算法巧妙地利用核函数将 RX 算法映射至高维特征空间, 加强了光谱中非线性信息的运用, 具有较强的可分辨性, 显著改善了低维空间的光谱不可分问题。然而, 也暴露了 KRX 算法中病态 Gram 矩阵求逆误差大, 异常检测效率低等缺点。为实现理论上 KRX 算法的强探测性能, 提出一种基于新型聚类方法的改进 KRX 探测算法(DC-KRX)。(1)由于空间邻域像元具有较强的光谱相似性, 会造成 Gram 矩阵病态, 严重影响了异常探测效果, 因此背景虚检现象严重。针对病态 Gram 矩阵的求逆误差问题, 算法改进了 KRX 算子, 对 Gram 矩阵进行奇异值分解, 选取特征值较大的主成分, 保证了 Gram 矩阵的求逆精度, 待测像元的探测结果采用 $l-2$ 范数表示, 检测效果提高明显;(2)在改进 KRX 的基础上, 提出了空间聚类 KRX 算法。空间像元之间具有光谱强相关性, 既造成了 Gram 矩阵的病态, 数据的冗余也影响了探测效率。实验发现, 通过聚类算法可以合并像元于聚类中心, 减少空间维度, 提高计算效率; 同时, 聚类中心按照聚类大小被赋予不同的权重, 保证了探测精度;(3)另一方面, 选用合适的聚类算法是一个难点。聚类 KRX 算法对于聚类算法的精度和实时性要求较高, 比较发现, 一种基于密度峰值快速搜索(DC)的新型聚类算法具有较好的聚类性能。算法采用欧式距离计算任意像元的相似度, 利用局部密度和邻域距离作为聚类中心的联合判断准则, 对结果进行排序得到聚类中心。实验发现, 该聚类算法计算速度快, 且能够对任意形状的进行聚类, 非常适合于维度较高, 成分复杂的高光谱图像, 且适用于较高次数的重复聚类。DC-KRX 算法提供了一种空间聚类预处理的高光谱异常探测新思路, 最后, 与国际主流探测算法对比发现, 该算法表现了较好的探测性能。同时, 时效性对比分析发现, 聚类前后算法的检测效率提高了 30% 以上, 有效改善了 KRX 算法的实时性。

关键词 高光谱图像; 异常检测; 密度聚类; 奇异值分解; Kernel RX

中图分类号: TP751.1 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2019)06-1878-07

引言

近几十年来, 随着高光谱成像技术的飞速发展, 高光谱成像的空间分辨率和光谱分辨率, 以及光谱成像的实时性得到了显著提高, 针对高光谱图像的相关技术逐渐占据了遥感领域的研究前沿。高光谱图像不同于一般意义上的 RGB 彩色图像, 其包含着成百上千个波段, 可以精确地描述地物的空间和光谱特性, 因此, 成为高光谱遥感领域中地物精确探测的一个重要手段^[1-3]。

高光谱目标探测技术是高光谱遥感领域的一个重要应用, 而异常探测技术由于不需要知道目标像元和背景的任何

先验信息, 成为高光谱目标探测技术领域研究的重点^[4]。最典型的异常探测算法是 Reed 等在 1990 年针对多光谱目标探测提出的 Reed-Xiaoli (RX) 异常探测算法^[5], 作为一种基准算法, RX 算法得到了广泛应用。但是, 全局 RX 算法存在探测概率不高, 虚警现象无法克服等诸多问题, 因此, 基于局部领域滑动窗的 RX 算法(local RX, LRX)被提出^[6]。LRX 算法能够自适应地根据目标像元的周围像元构造背景正态模型, 显著提高了探测效果。而 RX 算法及其改进算法只是利用了高光谱图像的低阶信息, 且是一种基于马氏距离的线性模型。为了充分利用高光谱信息, Kwon 等^[7]提出了一种非线性核方法(kernel RX, KRX), 通过利用非线性核函数将原来空间的光谱信号映射到高维特征空间, 有效解决了线性空

收稿日期: 2018-05-03, 修订日期: 2018-10-16

基金项目: 国家自然科学基金项目(41574008)资助

作者简介: 刘春桐, 1972 年生, 火箭军工程大学导弹工程学院教授

* 通讯联系人 e-mail: aheadb@sina.com

e-mail: liuchuntong2001@163.com

间的光谱不可分问题,大大降低了虚警率。

KRX 算法虽然在一定程度上提高了异常检测的性能,但是探测精度和探测效率往往取决于协方差的求逆运算,为避免此问题, Banerjee 等^[8]提出了一种支持向量数据描述的异常探测算法 (support vector data description, SVDD)。SVDD 不需要对目标和背景的分布作任何假设,主要通过计算包含一组数据的最小超球体边界来对数据的分布进行描述,位于边界以外的视为异常点;虽然避免了求逆运算,但是对于代表稀疏权重的拉格朗日算子的求解往往比较困难,且时间复杂度与背景像元的个数成正比。近些年来,联合协同理论 (collaborative representation detection, CRD)^[9]被用于异常目标探测,目标像元往往可以表示为背景像元的线性组合,通过将协同表示结果与实际结果作差来作为异常目标的判断依据,取得了不错的探测效果,且实时性得到了显著增强。

针对 KRX 异常探测的算法复杂度较高,实时性不强等问题^[10],提出了一种基于密度聚类的核探测算法 (density cluster KRX, DC-KRX)。算法着眼于空间像元的光谱特征,对空间像元采用密度聚类算法聚类,根据类别大小分配权重因子,从而构造聚类 KRX 探测算子;采用奇异值分解改进探测算子中协方差阵的求逆运算,最终将探测算子转化为 $l-2$ 范数表示。实验部分,将该方法与传统 KRX 和 CRD 等高光谱异常探测方法比较,探讨了空间聚类对于算法实时性与探测效果的影响。

1 传统异常探测算法

1.1 传统 KRX 算法

实际异常探测中,光谱各波段之间具有很强的相关性。因此,不同于 RX 算法, KRX 算法将三维高光谱数据 $X = \{x_i\}_{i=1}^M$ 映射至高维特征空间 $\Phi(X)$,充分挖掘了像元的非线性光谱特性,在分离目标和背景方面取得了不错的探测效果。KRX 算法的表达式为

$$\text{KRXD}(r) = (\Phi(r) - \hat{\mu}_\Phi)^T X_{\Phi C} K_C^{-1} X_{\Phi C}^T (\Phi(r) - \hat{\mu}_\Phi) \underset{H_0}{\overset{H_1}{>}} \eta \quad (1)$$

其中,特征空间的背景均值 $\hat{\mu}_\Phi = \frac{1}{M} \sum_{i=1}^M \Phi(x_i)$,中心化的光谱数据为 $X_{\Phi C}(x_i) = \Phi(x_i) - \hat{\mu}_\Phi$,中心化核矩阵 $K_C = X_{\Phi C}^T X_{\Phi C}$ 。

利用核函数的方法计算特征空间的点积运算,极大地减少了高维数据的计算复杂度,进一步推导可以得到 KRXD 算子

$$\delta_{\text{KRXD}}(r) = (k_r^T - k_\mu^T)^T \hat{K}_b^{-1} (k_r^T - k_\mu^T) \underset{H_0}{\overset{H_1}{>}} \eta \quad (2)$$

式中, \hat{K}_b 为 Gram 矩阵。

1.2 改进 KRX 算法

高光谱数据邻域像元之间具有很强的光谱相似性,这样计算得到的 Gram 矩阵线性相关性过大,容易产生病态矩

阵。因此,传统 KRX 算法往往受到特征空间协方差阵求逆运算的制约,大量病态的协方差矩阵也难以通过广义逆求得可靠的结果。而采用奇异值分解代替协方差阵的求逆运算,结果的准确性将显著提高。

$$\hat{K}_b = U \Lambda V^T \quad (3)$$

其中, \hat{K}_b 为对称 Gram 阵, Λ 为对角阵。由对称阵的性质可知, $U=V$ 。

选取非零的特征值及对应的特征向量用于求解病态矩阵的逆矩阵,可以提高准确性。 t 为非零特征值的个数,假设 t 满足

$$D(t+1, t+1) < 10^{-6} \quad (4)$$

则有

$$\begin{cases} \hat{V} = V(:, 1:t) \\ \hat{D} = D(1:t, 1:t) \end{cases} \quad (5)$$

由此,异常探测算子可以表示为

$$\hat{\delta}_{\text{KRXD}}(r) = (k_r^T - k_\mu^T)^T \hat{V} \hat{\Lambda}^{-1} \hat{V}^T (k_r^T - k_\mu^T) \quad (6)$$

式(6)可以写成 $l-2$ 范数的表达形式

$$\hat{\delta}_{\text{KRXD}}(r) = \|\hat{\Lambda}^{-1/2} \hat{V}^T (k_r^T - k_\mu^T)\|_2 \underset{H_0}{\overset{H_1}{>}} \eta \quad (7)$$

2 聚类异常探测算法

传统 KRX 算法往往建立在局部窗模型,计算效率取决于窗的大小,即参与计算的背景像元的个数。而高光谱图像的邻域像元存在谱相关性,通过空间聚类,可以减少空间维度而不损失背景的特征信息,提高算子中 K_b , k_r 和 k_μ 等因子的计算效率。

2.1 密度聚类算法

传统的聚类算法如 Kmeans 聚类方法,是一种动态自更新的聚类方法,该方法的优点是计算效率比较高,收敛速度比较快,可以提高算法的实时性。但是由于 Kmeans 采用聚类结果的平均值作为聚类中心,改变了原有像元的光谱特征,得到的结果不具有可靠性,且收敛的结果往往取决于初始聚类中心的选取,因此,每次结果的差异性较大,鲁棒性较差^[11, 13]。

其他诸如启发式聚类算法和全局智能优化聚类方法等算法的复杂度较高^[12-13],不适用于对于较高次数重复聚类的高光谱异常探测。因此,必须找到一种优异而快速的高效聚类方法。

不同于传统的聚类算法, Rodrigue 等^[13]提出了一种基于密度峰值快速搜索的聚类算法 (CFSDP),简称密度聚类算法 (density cluster, DC)。该算法计算速度快,且能够对任意形状分布进行聚类,非常适合对于维度较高,成分复杂的高光谱图像。

考虑待聚类的高光谱图像 $X_b = [x_1, x_2, \dots, x_M]^T$,密度聚类算法实现步骤如下。

(1) 计算任意两点的欧式距离,构造距离矩阵 $D = \{d_{ij}\}_{i,j=1}^M$,即

$$d_{ij}^2 = (x_i - x_j)^T (x_i - x_j) = \|x_i - x_j\|_2^2 \quad (8)$$

(2) 计算局部密度和邻域距离

假设聚类中心都是密度较大的点，不妨定义局部密度

$$\rho = \sum_i e^{-(d_i/d_c)^2} \quad (9)$$

其中， d_c 表示截断距离。理论上，选取的 d_c 应使得每个数据点的周围点个数的平均值约为数据点总数的 1%~2%。

实验发现，任意两个聚类中心之间应具有一定的空间距离。则可定义邻域距离 s

$$s_{q_i} = \begin{cases} \min_{j < i} \{d_{q_i q_j}\}, & i \geq 2 \\ \max_{j \geq 2} \{d_{q_i q_j}\}, & i = 1 \end{cases} \quad (10)$$

式中， $\{q_i\}_{i=1}^M$ 表示 $\{\rho_i\}_{i=1}^M$ 降序排列后的位置序列。

(3) 建立聚类中心的联合判断准则

聚类中心往往是局部密度和邻域距离较大的点，因此，两个指标对于聚类中心的判定具有决定性作用。建立联合判断准则

$$\gamma_i = \rho_i s_i, \quad i = 1, \dots, M \quad (11)$$

只需对 $\{\gamma_i\}_{i=1}^M$ 进行降序排列，选取靠前的 N 个数据点作为聚类中心。

图 1 中给出了密度聚类算法(density cluster, DC)的流程图。

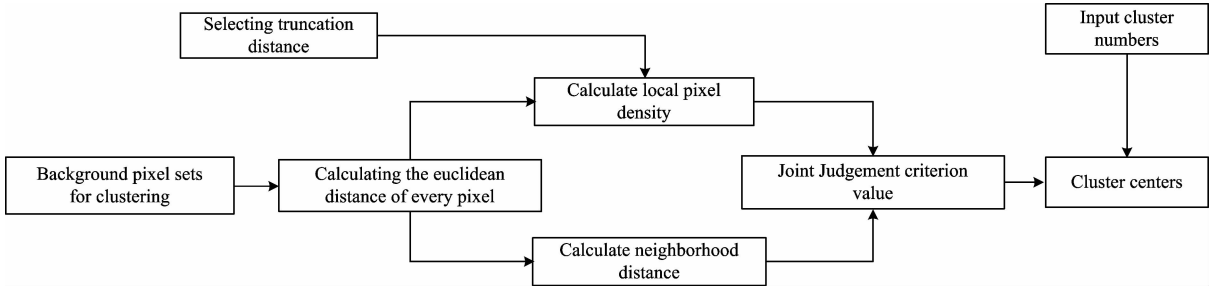


图 1 密度聚类算法流程图
Fig. 1 Flow chart of density cluster method

2.2 聚类 KRX 算法

假设有 N 个聚类中心构成 N 类，则建立样本数据集

$$I_S = \{(z_1, s_1), (z_2, s_2), \dots, (z_n, s_n)\} \quad (12)$$

其中， z_i 和 s_i 分别为第 i 类的聚类中心和对应的像元个数。

不同的聚类中心对应的类集大小不同，单纯的用聚类中心代替所有点不能完全表现背景像元的光谱特性，为体现不同聚类中心的差异性，更好地反映背景光谱，分别计算权重向量和权重矩阵作为聚类权重因子代入到探测表达式中。

权重向量

$$w_i = s_i / \sum_{i=1}^n s_i \quad (13)$$

权重矩阵

$$S = \text{diag}(s_1, s_2, \dots, s_n) / \sum_{i=1}^n s_i \quad (14)$$

由此，可以得到聚类 KRX 异常探测算法(density cluster kernel cluster, DC-KRX)的表达式为

$$\hat{\delta}_{\text{DC-KRXD}}(r) = \|\hat{\Lambda}^{-1/2} \hat{V}^T (\hat{\mu} - \hat{\nu})\|_2 \underset{H_0}{\overset{H_1}{>}} \eta \quad (15)$$

其中

$$\hat{\mu} = K_z w - l_{M \times 1} w^T K_b w \quad (16)$$

$$\hat{\nu} = \nu - l_{M \times 1} w^T \nu, \quad \nu_i = k(x_i, r) \quad (17)$$

$$\hat{K}_z = K_z - K_z w l_{1 \times M} - l_{M \times 1} w^T K_z + l_{M \times 1} w^T K_z w l_{1 \times M} \quad (18)$$

$l_{1 \times M}$ 和 $l_{M \times 1}$ 为所有元素均为 1 的 M 维向量， K_z 为聚类 Gram 核矩阵， \hat{K}_z 为聚类 Gram 矩阵。

聚类 KRX 算法在对 \hat{K}_z 进行奇异值分解时加入了权重矩阵 S ，可以得到 $\hat{\Lambda}$ 和 \hat{V}

$$\hat{K}_z S \hat{V} = \hat{V} \hat{\Lambda} \quad (19)$$

3 实验验证与分析

3.1 可行性分析

为了验证本算法的优越性，选取两幅高光谱图像用于实验分析。两幅图像均采用 AVIRIS 高光谱成像仪拍摄的真实美国圣地亚哥海军机场的真实高光谱数据。图像的原始大小为 400×400 ，共有 224 个波段，从中截取具有 3 个异常目标的 100×100 大小的高光谱数据作为区域 1，以及 38 个飞机异常目标的 100×100 大小的高光谱数据作为区域 2。去掉水汽吸收和噪声较大的波段以及第一和最后一个波段，剩下 201 个波段用于异常目标探测。如图 2 所示区域 1 和区域 2 为用于实验的高光谱图像。

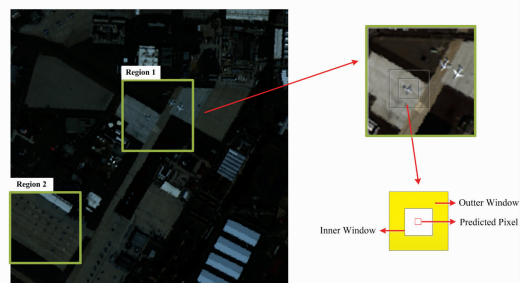


图 2 海军机场及截取的伪彩色图像
Fig. 2 Naval airfields and pseudo-color images for test

如图 3 所示，高光谱图像的邻域像元具有光谱强相似性。选取 Region1 中 (36, 72) 处像元为目标，设定窗的大小为 17×13 ，采用密度聚类算法对目标像元的外窗进行聚类，设定聚类中心的个数为 4，则聚类效果如图 3(c) 所示，不难

看出, 该聚类算法具有较好的可行性, 聚类中心在目标周围分布均匀, 且聚类的彩色图[图 3(c)]反映出聚类效果较好, 充分表现了背景像元的光谱特性。图 4(a)为聚类前的外窗像元光谱曲线, 可以看出, 光谱曲线的聚类特征明显。图 4(b)为 4 个聚类中心的光谱曲线, 表明类间像元的光谱差异性; 而图 4(c-f)为聚类结果的光谱曲线(分别为类别 1~4), 表明了类内像元光谱的相似性。因此, 在进行高光谱异常探测时, 基于光谱相似性的聚类预处理可以集中复杂的背景信息, 减少数据的冗余。而实际操作中, 往往将聚类个数设定为像元总数的 20%~30%, 以保证异常目标探测的效果。

3.2 算法性能分析

(1) 聚类算法性能

本节详细分析了密度聚类算法在高光谱空间聚类的优越性, 并将三种不同的聚类算法进行了对比。数据采用图 1 中的 Region1。对于局部异常探测算法, 窗的大小取决于异常目标的大小。与 Region2 相比, Region1 的目标更大, 因此设置外窗的大小为 17×17 , 内窗的大小为 13×13 。

第一种聚类方法采用 Unique cluster, 是将图像中具有

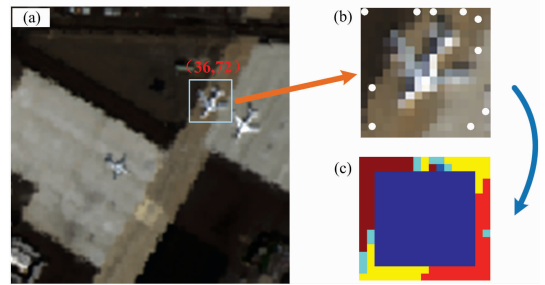


图 3 采用密度聚类算法的聚类效果

(a): 目标区域; (b): 聚类中心分布; (c): 聚类结果

Fig. 3 Cluster performance using density clustering algorithm

(a): Target region; (b): Cluster centers distribution; (c): Clustering results

相似光谱的像素进行合并; 第二种聚类方法采用 K-means cluster; 第三种聚类方法采用 Density Cluster。探测结果如图 5 所示。

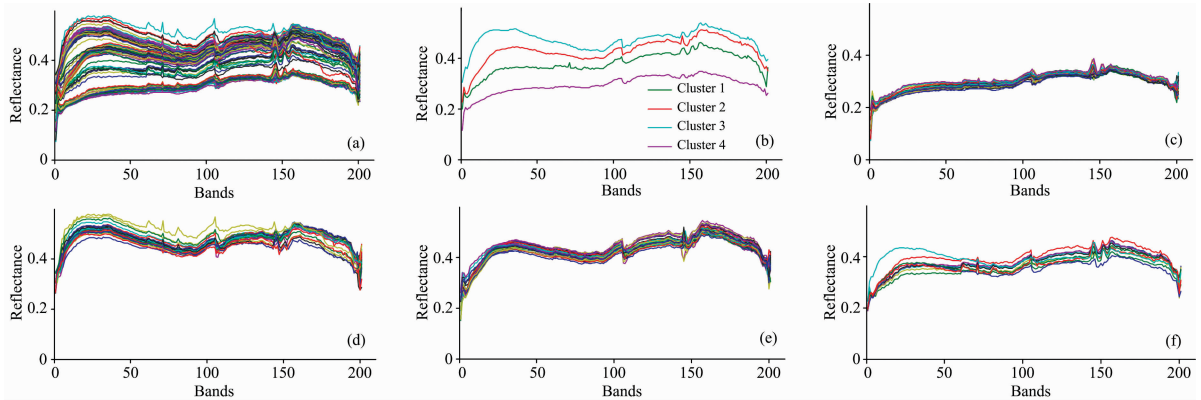


图 4 (a) 以点 (36, 72) 为中心的待聚类像元光谱曲线; (b) 4 个聚类中心的光谱曲线; (c-f) 类别 1~4 的像元光谱曲线

Fig. 4 (a) Reflectance spectra of pixels around the point (36, 72); (b) Reflectance spectra of four clusters; (c-f) Reflectance spectra of class one to class four

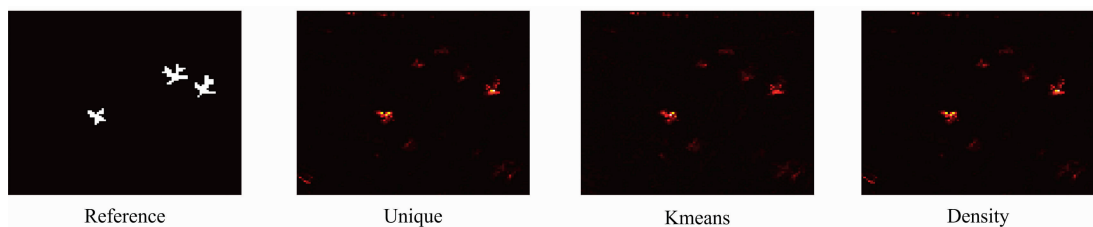


图 5 不同聚类算法的探测结果

Fig. 5 Detection results of different cluster algorithms

从图 5 可以看出, 不同聚类算法在结果表现上差别不大。这种情况下, 常用 AUC Value 作为不同探测算法比较的重要指标。同时, 也比较了不同算法的时间复杂度, 如表 1 所示。可以看出, Unique 聚类方法对于异常探测目标效果的提高比较明显, 但是聚类点数较多, 算法的时效性不强。而 Kmeans 聚类方法由于聚类中心的随机性, 算法的稳定性不强, 表中显示的指标为 Kmeans 聚类 5 次得到的平均值。不

同于其他两种聚类方法, Density 聚类算法有效提高了算法的时效性, 且对于异常点的探测概率高于不采用聚类方法的 KRX 算法。实验运行平台为 2.5 GHz 的 Intel Core i5-2450M 双核处理器, 睿频可达 3.1 GHz, 运行内存为 16 GB, 硬盘为 230 G 固态硬盘。仿真软件版本为 Matlab 2014a。为防止电流脉冲和噪声干扰的影响, 计算时间均为程序运行 5 次所取的平均值。

表 1 不同聚类算法的性能比较

Table 1 Performance comparison of different cluster methods

	AUC value	Cluster time/s	Detection time/s
Unique	0.971 6	1.572	58.817
Kmeans	0.966 0	170.216	187.151
Density	0.971 6	26.031	53.650

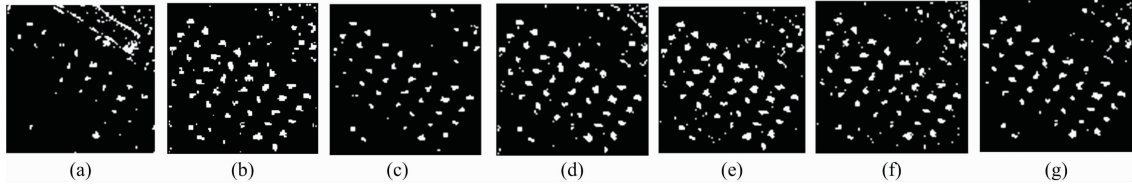


图 6 不同异常探测算法的探测结果

(a): 传统 KRX; (b): 局部 RX; (c): CRD; (d): 改进-KRX; (e): U-KRX; (f): K-KRX; (g): DC-KRX

Fig. 6 Detection results of different anomaly detection algorithms

(a): Original-KRX; (b): LRX; (c): CRD; (d): Improved-KRX; (e): U-KRX; (f): K-KRX; (g): DC-KRX

从图 6 可以看出, DC-KRX 算法[图 6(g)]表现出了不错的探测效果。相对于传统的 KRX 算法[图 6(a)], 改进的 KRX 算法[图 6(d)]通过奇异值分解巧妙地解决了病态矩阵的不可逆问题, 相比于利用广义逆算法求解协方差阵的逆, 该算法极大地改善了探测概率, 并有效地降低了虚警, 算法的效率也得到了显著提升。另一方面, 基于聚类的 KRX 算法[图 6(e, f, g)], 相对于国际上较为领先的算法(如 CRD

(2)DC-KRX 算法探测性能

本节将 DC-KRX 算法与探测性能较好的 CRD 和 LRX 等算法进行比较, 同时考虑不同聚类算法对于 KRX 算法的影响。实验图像采用比较普遍的美国圣地亚哥海军机场数据, 即图 1 中 Region 2。

所有算法基于局部滑动窗的模型, 这里统一采用相同的滑动窗, 考虑到异常目标大小, 将内窗大小设定为 5×5 , 外窗大小设定为 13×13 。得到异常探测的结果如图 6 所示。

等), 展现了较好的探测性能, 而本工作提出的 DC-KRX 算法, 在探测效果和实时性方面, 要优于其他算法。同时, 图 6(e)和(f)也给出了基于不同的聚类方法(Unique 和 K-means)下的 KRX 探测结果。

为更充分地表现 DC-KRX 算法的探测性能, 图 7 中给出了不同探测算法的接收机曲线和 AUC 指标。从 AUC 指标可以看出, DC-KRX 的探测性能明显优于其他探测算法。

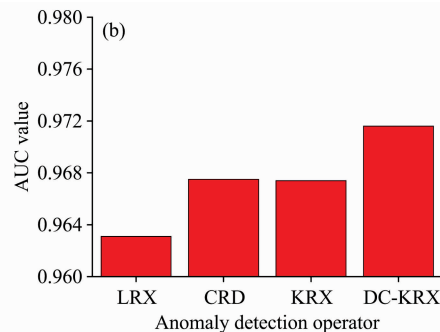
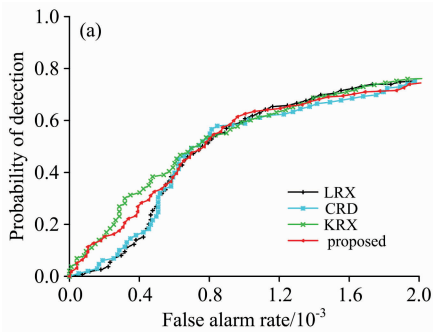


图 7 不同异常目标探测算法的性能比较

Fig. 7 Receiver operator curves and AUC value of different detection methods

表 2 不同窗口对应的像元数和聚类中心数

Table 2 Number of pixels and cluster centers corresponding to different windows

Window size	Pixel number	Cluster number
0×0	0	0
3×5	16	5
5×7	24	10
3×7	40	10
5×9	56	20
3×9	72	20
5×11	96	25
3×11	112	25
5×13	144	30
3×13	160	30

3.3 时效性分析

Density 聚类与传统的 K-means 聚类算法相比, 具有较强的时效性。为充分说明算法在时间复杂度上的优战性, 比较了不同窗口下不同聚类个数算法所用的时间。

调整内窗和外窗的尺寸, 对应单个像元计算时所需的聚类中心个数如表 2 所示。

基于此, 分析对于不同尺寸滑动窗条件下的不同聚类方法的时间复杂度, 以及不同探测算法的时间复杂度。(图 8 中数据为实验 5 次所取的平均值)。

从图 8 可以看出, Density 聚类方法明显优于 K-means 聚类方法。

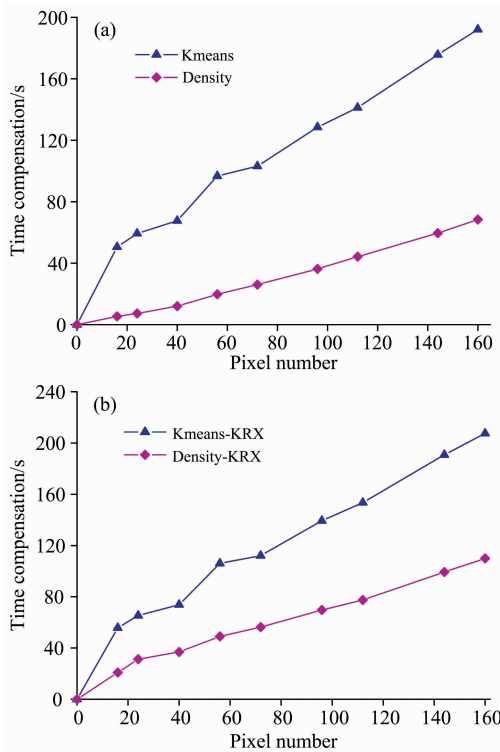


图 8 不同聚类算法时效性比较

Fig. 8 Comparison of timeliness of different clustering algorithms

表 3 中给出了 DC-KRX 算法与 CRD-DW-STO 和 KRX 算法的对比, 窗口统一设置为 5×13 , 给出了不同算法的时间复杂度以及增速比。可以看出 DC-KRX 算法与其他两种算

表 3 DC-KRX 和 CRD-DW-STO, KRX 探测算法的性能比较
Table 3 Performance comparison of DC-KRX and CRD-DW-STO, KRX Detection algorithms

Detection algorithm	AUC value	Time /s	Speed-up ratio/%
DC-KRX (50 clusters)	0.968 9	52.367	—
CRD-DW-STO	0.967 5	77.155	32.13
KRX	0.967 5	86.463	39.43

References

[1] Bioucas-Dias J M, Plaza A, Camps-Valls G, et al. IEEE Geoscience & Remote Sensing Magazine, 2013, 1(2): 6.
 [2] Nasrabadi N M. IEEE Signal Processing Magazine, 2014, 31(1): 34.
 [3] ZHANG Bing(张 兵). Journal of Remote Sensing(遥感学报), 2016, 20(5): 1062.
 [4] Kang X, Zhang X, Li S, et al. IEEE Transactions on Geoscience & Remote Sensing, 2017, 55(10): 5600.
 [5] Reed I S, Yu X. IEEE Transactions on Acoustics Speech & Signal Processing, 1990, 38(10): 1760.
 [6] Molero J M, Garzón E M, García I, et al. IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing, 2013, 6(2): 801.
 [7] Kwon H, Nasrabadi N M. IEEE Transactions on Geoscience & Remote Sensing, 2005, 43(2): 388.
 [8] Banerjee A, Burlina P, Diehl C. IEEE Transactions on Geoscience & Remote Sensing, 2006, 44(8): 2282.
 [9] Li W, Du Q. IEEE Transactions on Geoscience & Remote Sensing, 2015, 53(3): 1463.

法相比, 时间增速比都在 30% 以上, 探测效率提高的非常明显。

为了进一步验证 DC-KRX 算法在计算效率上的优越性, 图 9 中给出了 DC-KRX 和 KRX 算法在计算单个像元所需要的时间, 从图中可以看出, 由于避免了复杂的求逆运算, 采用了高效的聚类方法, 探测效率明显提高。

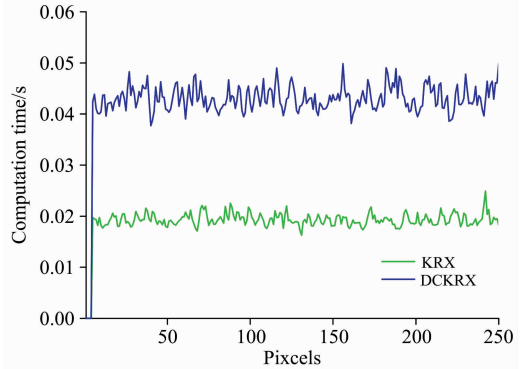


图 9 聚类前后的时间复杂度分析(每 5 个像素求和)

Fig. 9 Analysis of time complexity before and after clustering (sum of every 5 pixels)

4 结 论

为解决 KRX 算法的准确性和时效性不强的问题, 提出了一种基于空间密度聚类的改进 KRX 算法。算法改进了 KRX 探测算子中的协方差求逆运算, 解决了病态矩阵的求逆误差, 提高了探测概率; 对于实时性方面, 充分考虑了高光谱影像的空间相关性及光谱相似性, 综合背景像元的信息于聚类中心, 降低空间维度, 并将聚类中心和权重的组合运算代替全部像元的整体运算, 改善了探测算法的时效性; 同时, 算法比较不同聚类算法的优缺点, 将一种新型高效的聚类算法用于空间聚类, 该聚类算法具有较好的实时性, 且聚类结果比较理想, 较好地保证了异常探测的效果。在实际实验过程中, 该算法表现出较强的鲁棒性, 为提高异常探测效率提供了一种聚类预处理的新思路。但是, 该算法过分依赖聚类算法的准确度和计算效率, 在后续的研究中, 寻找更加高效和准确的聚类算法, 将有助于提高 DC-KRX 算子的实时性和准确性。

- [10] ZHAO Chun-hui, LI Xiao-hui, WANG Yu-lei(赵春晖, 李晓慧, 王玉磊). Journal of Electronic Measurement and Instrumentation(电子测量与仪器学报), 2014, 28(8): 803.
- [11] ZHU Guang-hui, HUANG Sheng-bin, YUAN Chun-feng, et al(朱光辉, 黄圣彬, 袁春风, 等). Chinese Journal of Computer(计算机学报), 2017, 40(77): 1.
- [12] ZHENG Jian-wei, ZHU Wen-bo, WANG Wan-liang, et al(郑建伟, 朱文博, 王万良, 等). Journal of Computer-Aided Design&-Computer Graphics(计算机辅助设计与图形学学报), 2018, 30(1): 116.
- [13] YANG Shu-ying(杨淑莹). Implementation of Pattern Recognition and Intelligent Computing(模式识别与智能计算). Beijing: Publishing House of Electronics Industry(北京: 电子工业出版社), 2015. 4.
- [14] Rodriguez A, Laio A. Science, 2014, 344(6191): 1492.

A Density-Based Cluster Kernel RX Algorithm for Hyperspectral Anomaly Detection

LIU Chun-tong, MA Shi-xin*, WANG Hao, WANG Yang, LI Hong-cai

Key Laboratory of Missile Launching and Orientation Aiming Technology of PLA, Rocket Force University of Engineering, Xi'an 710025, China

Abstract Hyperspectral remote sensing image contains abundant spectral information, which has strong ability to distinguish ground objects, thus promoting the development of hyperspectral anomaly detection technology without any prior information. Kernel-RX algorithm uses the kernel function to ably RX algorithm mapped to high-dimensional feature space, which has a strong ability to solve the spectrum inseparable problem in the low dimensional space. However, it also reveals the disadvantages such as large inverse error in ill-conditioned matrix and low efficiency. In order to realize the strong detection performance of KRX algorithm in theory, this paper proposes an improved KRX detection technology based on new clustering algorithm. (1) Due to the strong spectral similarity of spatial neighborhood pixels, the Gram matrix is ill-conditioned, which seriously affects the detection performance of anomalies, so the phenomenon of background error detection is serious. In order to solve the problem of the inverse error of ill-conditioned Gram matrix, the algorithm improves the KRX operator. By decomposing the singular value of Gram matrix and selecting the principal component with larger eigenvalue, the algorithm ensures the inverse accuracy of Gram matrix. In the end, the detection result of the pixel to be measured is expressed by l_2 norm. The experiment shows that the detection effect is improved obviously. (2) Based on the improved KRX, a spatial clustering KRX algorithm is proposed. There is strong a correlation between spatial pixels, which not only leads to the ill-condition of Gram matrix, but also affects the detection efficiency. The experimental results show that combining pixels in the clustercenters can reduce the spatial dimension and improve the computational efficiency. At the same time, the clustering center is given different weight factors according to the size of the cluster, which ensures the detection accuracy. (3) On the other hand, it is difficult to select an appropriate clustering algorithm. Clustering KRX algorithm requires high accuracy and real-time performance. It is found that a new clustering algorithm based on the peak density fast search algorithm has better clustering performance. The Euclidean distance is used to calculate the similarity of any two pixels, and the Local Density and Neighborhood Distance are used to calculate the clustering center. The clustering center is obtained by sorting the results of the Joint Judgement Criterion. The clustering results show that this clustering algorithm is fast and can cluster arbitrary shape distribution, which is very suitable for hyperspectral images with high dimension and complex components, and can be used for repeated clustering with high frequency of anomaly detection. In conclusion, DC-KRX algorithm provides a new idea of hyperspectral anomaly detection based on spatial clustering preprocessing. Finally, the algorithm is compared with the advanced method. The results show our method has a strong detection performance. And, it is found that the detection efficiency of clustering algorithm is improved by more than 30%, which greatly improves the real-time performance of KRX algorithm.

Keywords Hyperspectral imagery; Anomaly detection; Density cluster; Singular value decomposition; Kernel RX

(Received May 3, 2018; accepted Oct. 16, 2018)