

文章编号:1004-4213(2010)12-2224-5

# 基于稳健背景子空间的高光谱图像异常检测<sup>\*</sup>

蒲晓丰,雷武虎,黄涛,王迪

(电子工程学院 脉冲功率激光技术国家重点实验室,合肥 230037)

**摘要:**在 RX 算法中,局部背景协方差矩阵估计会由于背景受到异常像元的“污染”而不能准确反映背景分布,从而导致检测性能下降.针对这一问题,提出一种基于稳健背景子空间的异常检测算法.利用空间秩深度度量背景中每个样本相对于整个背景样本分布空间的位置,将“游离”于背景分布空间之外的样本看作是潜在的异常样本,并将其映射到背景分布空间之内.在此基础上,通过估计背景的协方差矩阵,利用主成分分析构造能更精确地刻画背景的子空间,在该子空间进行了基于马氏距离的检测异常.模拟和真实数据验证了该算法的有效性.

**关键词:**异常检测;背景子空间;空间秩深度;映射;高光谱图像

**中图分类号:**TP751.1

**文献标识码:**A

**doi:**10.3788/gzxb20103912.2224

## 0 引言

高光谱图像的异常检测是在缺少背景先验知识的前提下检测出图像中与背景存在光谱差异的小目标.由 Reed 和 Xiao Yu 提出的 RX 算法<sup>[1]</sup>得到广泛应用,并成为了高光谱图像异常检测的基准算法.它基于背景服从局部多元正态分布,通过估计局部区域中背景的协方差矩阵计算异常和背景的马氏距离来检测异常.但是,在实际应用过程中,局部区域中常常会包含少量异常像元,使得对异常比较敏感的背景协方差矩阵并不能够准确反映背景分布,产生“掩饰”效应(masking effect)<sup>[2]</sup>,从而导致该算法的检测性能受到很大的影响.文献<sup>[2]</sup>利用最小广义方差法(Minimum Covariance Determinant, MCD)获取稳健的背景协方差矩阵估计,但应用在高维数据时会受到限制<sup>[3]</sup>.文献<sup>[4]</sup>通过去除局部区域中前期检测的异常样本来避免其对协方差的影响,但会加剧协方差矩阵估计小样本的问题.

针对这一问题,本文提出一种基于稳健背景子空间的异常检测算法,利用协方差矩阵的特征值-特征向量来重新表述马氏距离,通过将“游离”在背景数据云团之外的少量异常数据映射到背景数据云团内,利用主成分分析(Principal Components Analysis, PCA)构造能更好刻画背景的子空间,并在该子空间中检测异常.

## 1 RX 算法

RX 算法是建立在多元正态分布的基础上,整幅图像中的数据是看作是多个多元正态分布的组合,而局部区域内的数据可以用一个多元正态分布近似.该算法假定异常目标光谱和背景的协方差未知,以检测像元为中心,确定两个同心嵌套的局部检测窗口,假设背景分布于外窗口,计算检测像元和背景的马氏距离检测异常为

$$RX(x) = (x - \hat{u})^T C^{-1} (x - \hat{u}) \quad (1)$$

式中, $x$  是待检测像元的  $L$  维光谱向量, $\hat{u}$  是背景的均值, $C$  是背景的协方差矩阵.

## 2 用特征值-特征向量表述马氏距离

利用 PCA 对协方差矩阵  $C$  对角化

$$C = U \Lambda U^T \quad (2)$$

式中: $U$  是特征向量按其对应的特征值由大到小的顺序排列而成的矩阵,且  $UU^T = I$ ;  $\Lambda$  是相对应的特征值构成的对角矩阵; $I$  是  $L$  维单位矩阵.将马氏距离式(1)改写为由特征值-特征向量的表述方式<sup>[5-6]</sup>

$$RX(x) = \sum_{i=1}^L \frac{\{\phi_i^T (x - \hat{u})\}^2}{\lambda_i} \quad (3)$$

式中, $\phi_i$  和  $\lambda_i$  分别是  $C$  的第  $i$  个特征向量和特征值.令

$$RX_i(x) = \frac{\{\phi_i^T (x - \hat{u})\}^2}{\lambda_i} \quad (4)$$

式(4)可以被认为是在特征向量  $\phi_i$  上的马氏距离.但是,异常数据的干扰会导致特征向量发生偏转,从而影响了各特征向量上的马氏距离.故而,抑制背景中的异常数据,并在其基础上通过 PCA 构造的背景

<sup>\*</sup>安徽省自然科学基金(070415217)资助

Tel:0551-5767748

Email:pxf\_555@sina.com

收稿日期:2010-08-03

修回日期:2010-09-17

子空间能够更好地描述背景. 这样, 只使用前  $D(D < L)$  个主向量上的马氏距离, 式(3)可以写为

$$RX'(x) = \sum_{i=1}^D \frac{\{\phi_i'^T(x - \hat{u}')\}^2}{\lambda_i} \quad (5)$$

式中,  $\phi_i'$  和  $\lambda_i'$  分别是第  $i$  个无异常干扰的主向量和特征值.  $\hat{u}'$  为背景均值.

### 3 稳健子空间方法

基于 PCA 构建稳健的背景子空间, PCA 的优化目标函数是使样本重建的均方误差最小

$$J_{\text{pca}}(U) = \sum_{i=1}^N x_i - x_i'^2 = \sum_{i=1}^N x_i - UU^T x_i^2 \quad (6)$$

式中,  $x_i \in X$  是  $L$  维的样本点,  $N$  是样本个数,  $U$  是前  $D(D < L)$  个特征向量组成的特征空间,  $Y = U^T X$  是将  $L$  维数据映射到低维子空间中. 设数据集  $X$  中含有  $M$  个少量异常数据, 则式(6)可以写为

$$J_{\text{pca}}(U) = \sum_{i=1}^{N-M} x_i - UU^T x_i^2 + \sum_{j=1}^M x_j - UU^T x_j^2 \quad (7)$$

异常数据往往“游离”于服从某一  $L$  维分布的数据云团, 最小化式(7)就是弱化异常数据对均方误差的影响. 其最直接的方法是检测出异常数据并去除, 但是这种方法并不适用于小样本的情况. 为此, 引入空间秩深度<sup>[7]</sup>的概念来度量局部区域中每个样本数据相对整个数据云团的空间位置, 将游离于整个数据云团之外的数据映射到数据云团内, 消除异常的影响.

#### 3.1 空间秩深度

文献[7]基于空间秩给出了一个数据深度, 设  $x$  是  $L$  维的样本点 ( $x \in R^d$ ), 其相对于服从  $L$  维  $P_n$  分布的数据云团  $Y = [y_1, y_2, \dots, y_N], x \in R^d (x \notin Y)$  的空间秩深度可以写为

$$D(x, P_n) = 1 - \frac{1}{N} \sum_{i=1}^N \frac{x - y_i}{x - y_i} \quad (8)$$

式中,  $y_i$  是数据云团  $Y$  中第  $i$  个样本,  $N$  是数据云团包含的样本个数. 其主要具有如下的性质:

1) 偏离云团无穷远处深度为 0;

2) 单调性: 数据越靠近数据云团的表面, 深度越小, 这保证了异常数据具有较小的深度值.

#### 3.2 映射关系

计算局部区域中每个样本点  $x_i (x_i \in X, i = 1, 2, \dots, n)$  相对于其它  $n-1$  个样本点构成的数据云团的深度. 将深度值小于预先设定的阈值  $D_{\text{anomaly}}$  的  $M$  个样本点  $\bar{X} = [\bar{x}_1, \bar{x}_2, \dots, \bar{x}_M]$  看作是潜在的异常数据; 将深度值大于预先设定的阈值  $D_{\text{background}} (D_{\text{background}} > D_{\text{anomaly}})$  的  $K$  个样本点看作是“干净”的背景数据. 利用式(9)将样本点  $\bar{x}_i \in \bar{X}$  映射为  $z_i$ , 使得  $z_i$  的空

间秩深度  $D(z_i, P_n)$  大于  $D_{\text{background}}$ , 这相当于将游离于数据云团之外的样本映射到数据云团中.

$$z_i = F(\bar{x}_i, X^b, \hat{u}) = \bar{x}_i + t \cdot (u^b - \bar{x}_i) \quad (9)$$

式中,  $u^b = \frac{1}{K} \sum_{j=1}^K x_j^b, x_j^b \in X^b, \hat{u} = \frac{1}{n} \sum_{i=1}^n x_i, t = \frac{(\hat{u} - \bar{x}_i)^T (u^b - \bar{x}_i)}{(u^b - \bar{x}_i)^T (u^b - \bar{x}_i)}$ .  $\hat{u}$  是所有样本的均值,  $u^b$  是“干净”的背景数据的均值, 即  $z_i$  可以被看作是  $\hat{u}$  在直线  $\bar{x}_i u^b$  上的投影点.

### 4 算法的实现

基于本文的讨论, 算法采用同心嵌套矩形窗口选择样本, 外窗用于选择背景像元, 内窗为待检测像元. 内、外窗大小的选择主要依据异常目标的大小及光谱维数来确定. 具体实现步骤归纳为:

1) 输入高光谱图像 ( $L$  个波段), 对数据进行归一化处理;

2) 计算每个背景样本的空间秩深度, 并与预先设定的阈值  $D_{\text{anomaly}}$  和  $D_{\text{background}}$  进行比较, 选择出潜在的异常数据和“干净”的背景数据. 利用式(9)对异常数据进行映射, 并将处理后的数据替代异常数据, 得到新的背景样本集  $X'$ ;

3) 计算  $X'$  的协方差矩阵  $C'$ , 对  $C'$  进行 PCA, 利用前  $D(D < L)$  个较大特征值对应的特征向量构造背景子空间  $V = [\phi_1', \phi_2', \dots, \phi_D']$ ;

4) 在子空间  $V$  中, 利用式(5)计算待检测像元的马氏距离, 若大于阈值  $\eta$ , 则认为是异常像元, 否则, 认为是背景;

5) 以内窗大小为步长对整幅图像进行遍历, 即可得到异常像元分布的二值图像.

### 5 实验分析

为验证算法的有效性, 通过实验仿真来说明.

#### 5.1 实验一

为检验空间秩深度探测异常数据的能力, 随机产生 510 个二维数据点作为数据集, 其中, 正常数据共 500 个服从  $N((0, 0), I)$  分布, 异常数据共 10 个服从  $N((3, 3), I)$  分布, 其分布如图 1(a). 根据式(8)计算每个数据点的相对整个数据云团的深度, 不同深度范围的数据分布如图 1(b). 通过设定阈值  $D_{\text{anomaly}} = 0.1$  和  $D_{\text{background}} = 0.9$ , 选择出潜在的异常数据集和“干净”的正常数据, 利用式(9)对异常数据集进行映射, 其映射数据、异常数据和“干净”数据的分布如图 1(c).

从图 1(b)中可以看出, 空间秩深度能够很好地反映出每个数据相对整个数据云团的位置, 深度越

大,越靠近云团的中心,反之亦然.而且,深度等高线在一定程度上也可以反映出数据云团的实际的形状分布.同时,分布在深度范围为  $0.0 < D(x) < 0.1$  的数据共有 6 个,且全部是异常数据,表明了利用空间秩深度探测潜在的异常数据是可行的.图 1(c) 则表明,通过映射,异常数据全部分布在数据云团内部,优化了式(7)的目标函数.

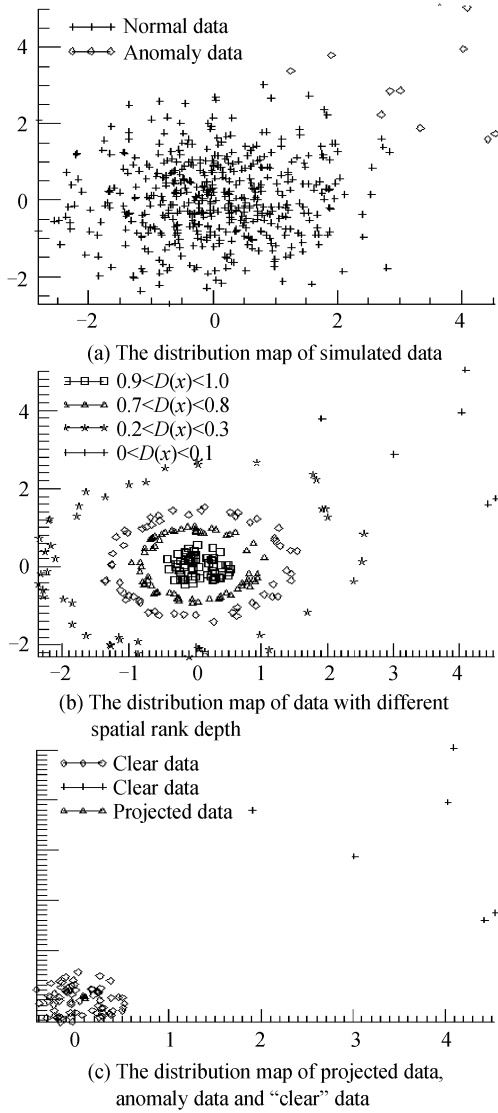


图 1 不同情况下仿真数据的分布图  
Fig. 1 Distribution maps of simulated data in different situations

5.2 实验二

为检测算法的有效性,利用真实高光谱数据(如图 2)进行了仿真.高光谱图像共 60 个波段,含有  $100 \times 180$  个像素,波长范围为  $0.441 \sim 1.3208 \mu\text{m}$ . 利用  $15 \times 15$  大小的矩形框选择不包含异常像元的区域作为背景样本集,如图 3.

在高光谱图像中选取 25 个异常像元替换背景集中 25 个背景样本,使得该局部背景被异常像元“污染”,分别利用 RX 算法本文算法计算被“污染”

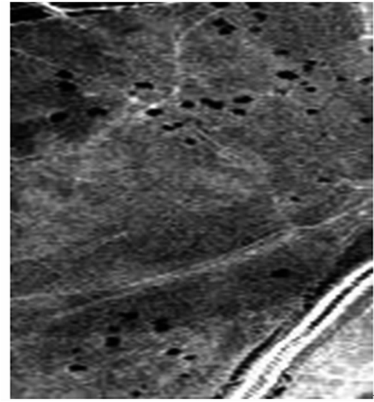


图 2 第一波段图像  
Fig. 2 The first image

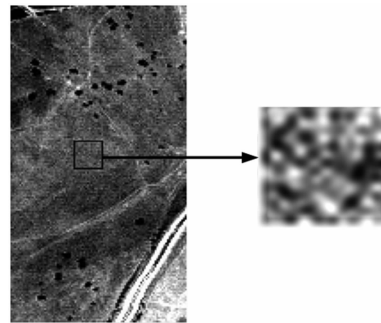


图 3 不含异常的局部背景  
Fig. 3 The local background without anomaly

的背景中每个样本的马氏距离,如图 4 和表 1. 其中,参量  $D_{\text{anomaly}} = 0.1, D_{\text{background}} = 0.5, D = 10$ .

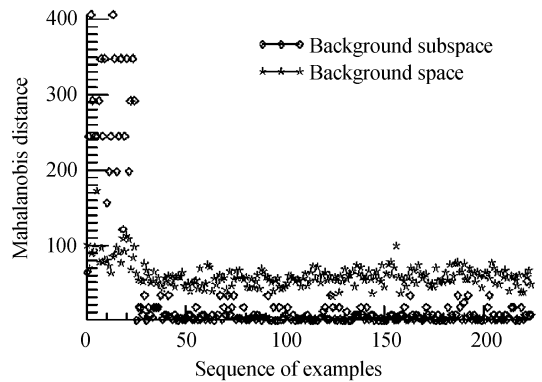


图 4 样本在不同背景空间上的马氏距离  
Fig. 4 Mahalanobis distances of examples on different background spaces

表 1 样本集在不同背景空间上的最大、最小马氏距离  
Table 1 The maximal and minimum mahalanobis distances of the example set on different background spaces

| Mahalanobis distance | In the background subspace | In the background space |
|----------------------|----------------------------|-------------------------|
| Max(A)               | 406.813 68                 | 172.136 42              |
| Min(A)               | 63.540 931                 | <b>62.644 439</b>       |
| Max(B)               | 51.689 918                 | <b>99.148 458</b>       |
| Min(B)               | 0.130 898 90               | 32.162 181              |

(注:Max(A)、Min(A)分别表示异常的最大、最小的马氏距离;Max(B)、Min(B)分别表示背景的最大、最小的马氏距离).

定义相对应马氏距离较小的异常像元为较弱异常(一般为亚像元).从图 4 和表 1 可以看出,在本文算法中,异常和背景的马氏距离存在明显的分离;而在 RX 算法中,异常和背景分离的程度并不明显,且存在一定的交叠(表 1 中加粗字体所示),这对于检测较弱异常带来困难,因为当分割阈值选取较大时,异常对背景的影响会“淹没”待检测的较弱异常,丢失异常目标;而当分割阈值选取较小时,异常的干扰会导致马氏距离发生偏离,容易造成异常周围的背景被当作异常检测出来.通过比较表明,本文算法构造的背景子空间能够有效抑制异常对背景的干扰,有助于检测较弱异常.

将两个算法应用于整个高光谱图像(如图 2),其中,两个算法使用相同大小的嵌套窗口,外窗为  $15 \times 15$ ,内窗为  $5 \times 5$ ,  $D_{\text{anomaly}} = 0.1$ ,  $D_{\text{background}} = 0.5$ ,  $D = 10$ ,异常检测结果如图 5.

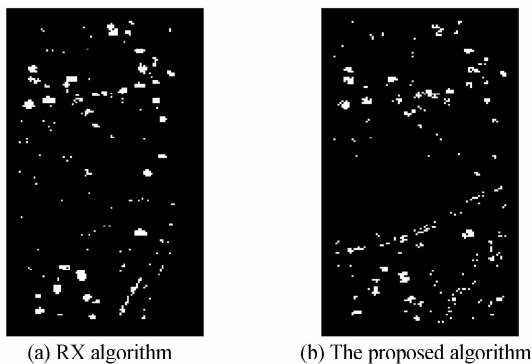


图 5 两种不同算法的检测结果

Fig. 5 Detection results of two different algorithms

由于整幅图像背景的复杂性,不能准确获取目标的详细分布及数目,为定量分析算法,从图像中选取  $50 \times 50$  大小的背景(不包含异常)和 94 个异常像元,并将异常像元随机插入到选取的背景中,应用两种算法对异常进行检测.利用 ROC 曲线描述两个算法的检测概率  $P_d$  和虚警概率  $P_f$  之间的变化关系,将检测概率  $P_d$  定义为检测到真实目标像素个数  $N_i$  与异常总数的比值;虚警概率  $P_f$  定义为检测的虚警像素个数  $N_b$  与总像素个数的比值.

$$P_d = N_i / 100, P_f = N_b / 2500 \quad (10)$$

利用不同的阈值分割得到的检测结果得到不同条件下的  $P_d$  和  $P_f$ ,图 6 给出了 RX 算法和基于稳健背景子空间的算法的 ROC 曲线的比较.可以看出:在虚警概率  $P_f < 0.012$  时,本文算法的性能略差于 RX 算法,这是因为当局部背景没有或被极少异常“污染”时,偏离背景较大的背景像元被当作异常检测出来,引起虚警上升,但这也从反面说明,该算法估计的背景子空间对于异常和偏离较大的背景像元非常敏感;当虚警概率  $P_f > 0.012$  时,本文算法的

性能明显优于 RX 算法, RX 算法的检测概率从 0.86 上升到 0.94,其相应的虚警概率从 0.0087 上升到 0.052,而本文算法虚警概率则只从 0.011 上升到 0.026,其主要原因是 RX 算法中异常对背景的干扰造成较弱的异常像元与背景的马氏距离存在着交叠现象,为了检测出较弱的异常,从而引起虚警概率骤然上升.

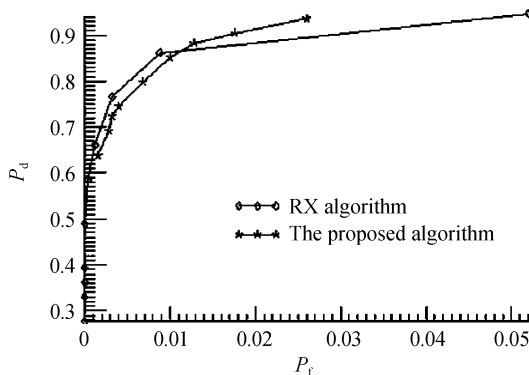


图 6 两种算法的 ROC 曲线

Fig. 6 ROC curves obtained by two algorithms

## 6 结论

提出了一种基于稳健背景子空间的异常检测算法,并对其有效性进行了阐述,通过与 RX 算法进行比较表明,该算法对于较弱异常具有较强的敏感性,有效抑制了局部背景中异常干扰的影响,但是当背景中不含有(或极少)异常时,容易将偏离背景较远的背景样本检测为异常,引起一定的虚警.尽管,当数据量比较大时,算法的计算量比较大,但是,相对于有效检测较弱异常来说,这些计算量还是值得的.

### 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 Trans on Acoustics, Speech, and Signal Processing*, 1990, **38**(10): 1760-1770.
  - [2] LO E, INGRAM J. Hyperspectral anomaly detection based on minimum generalized variance method [C]. *SPIE*, 2008, **6966**: 696603-1-7.
  - [3] MATTEOLI S, DIANI M, CORSINI G. Improved covariance matrix estimation: interpretation and experimental analysis of different approaches for anomaly detection application [C]. *SPIE*, 2009, **7477**: 74770K-1-9.
  - [4] HSUEH M, CHANG C I. Adaptive causal anomaly detection for hyperspectral imagery [C]. *2004 IEEE International Geoscience and Remote Sensing Symposium (IGARSS '04)*, 2004: 3222-3224.
  - [5] TAKESHITA T, NOZAWA S, KIMURA F. On the bias of mahalanobis distance due to limited sample size effect [C]. *IEEE, Proceedings of Second International Conference Document Analysis and Recognition*, 1993: 171-174.
  - [6] LI Guo-hong, SHI Peng-fei. Modified mahalanobis distance by compensation for error of non-dominant eigenvalues and asymmetrical distribution [J]. *Acta Electronica Sinica*, 2007, **35**(4): 747-750.
- 李国宏, 施鹏飞. 基于次特征值误差补偿和非对称分布的马

氏距离改进算法[J]. 电子学报, 2007, 35(4): 47-750.

*Statistics & Probability Letters*, 2003, 65(3): 17-225.

[7] GAO Yong-hong. Data depth based on spatial rank [J].

## Robust Background Subspace Based Anomaly Detection Algorithm for Hyperspectral Imagery

PU Xiao-feng, LEI Wu-hu, HUANG Tao, WANG Di

(*State Key Laboratory of Pulsed Power Laser Technology, Electronic Engineering Institute, Hefei 230037, China*)

**Abstract:** In RX anomaly detection algorithm, when background being contaminated from anomaly pixels, the local background covariance matrix estimation can not reflect background distribution accurately, which results in low detection capacity. To overcome this problem, a new method based on the robust background subspace was proposed. Utilizing the spatial rank depth, the position of every sample relative to the distribution space of whole background samples could be measured. Samples which locating at the edge of the distribution space were regarded as anomaly, and being mapped into the distribution space. In this way, the local background covariance matrix was estimated, and the principal component analysis as background space was obtained which can characterize background more accurately. An anomaly detection model was constructed on this subspace using mahalanobis distance. The effectiveness of the proposed method is validated by experimental results from simulated and real data.

**Key words:** Anomaly detection; Background subspace; Spatial rank depth; Mapping; Hyperspectral imagery



**PU Xiao-feng** was born in 1984. He is currently pursuing the M. S. degree at Electronic Engineering Institute, and his research interests focus on pattern recognition and hyperspectral imagery process.