

基于波段子集特征融合的高光谱图像异常检测*

贺霖 潘泉 赵永强 郑纪伟 魏坤

(西北工业大学自动化学院, 西安 710072)

摘要 高光谱图像分析中,对未知环境下伪装目标的检测识别具有较大难度,因为缺乏背景与目标的先验光谱信息.针对这一问题,提出一种高光谱图像异常检测算法.将高光谱图像分成波段子集进行特征提取,利用对图像中噪声程度及目标、背景之间可分性敏感的特征样本高阶统计量构造基本置信指派函数,通过 D-S 证据推理实现特征层智能融合异常检测.理论分析及仿真实验结果表明了算法的有效性.

关键词 高光谱图像处理;目标检测;特征融合;证据推理;波段子集

中图分类号 TP391 **文献标识码** A

0 引言

高光谱图像成像波段密集且连续,具有很高的光谱分辨率,与只利用空间信息的传统图像识别技术^[1~4]相比,高光谱图像处理在检测、识别地面低空间分辨率目标方面具有独特的优势.异常检测算法不利用光谱的先验知识,直接检测出与周围景物光谱存在明显差异的光谱信号所在位置作为异常点,所以异常检测算法对于未知场景中光谱特性未知目标的检测识别具有重要意义.常用异常检测算法主要有 RX 算法^[5,6]和 LPD 算法^[7]等.RX 算法将不符合场景统计特性的点判为异常;LPD 算法将场景中出现概率低的点判为异常.

利用高光谱图像对目标进行检测的常用算法通常会涉及通过投影等方法将原数据变换到具有更好可分性的数据空间,在此变换过程中存在一些问题:①变换过程经常包含对各个波段中背景与目标之间光谱差异的加权或乘方加权等运算,由于地物光谱的复杂性,由不同的光谱特性所得到的最终运算结果可能相近;②由于成像系统各波段成像噪声程度不同及成像系统波谱工作范围可能包含某些大气吸收波段,因此各波段成像质量不同.考虑这些因素,本文算法先将一组多维的高光谱数据按波段划分成多组波段子集数据并分别对每一子集进行特征提取,这样使得高光谱数据的可分性投影到了各子波段上,然后结合提取出的特征及特征的高阶统计量,并利用证据推理构造融合规则,最终实现融合检测.

以上处理可拉大目标与背景的差别,尽量体现出数据的内在可分性,从而有效检测目标.

1 RX 异常检测和 D-S 证据推理

1.1 RX 异常检测

RX 算法建立在待处理样本数据服从多元正态分布的基础上.对 P 维向量 x ,建立假设模型

$$H_0: x \sim N(\mu_b, \Gamma) \quad (1)$$

$$H_1: x \sim N(\mu_t, \Gamma) \quad (2)$$

式中: H_1 、 H_0 分别表示目标有、无, μ_b 、 μ_t 分别为目标有、无时对应的均值向量, Γ 为正态分布的协方差阵.根据 Neyman-Pearson 准则及渐近近似条件,可得到检验函数

$$RX(x) = (x - \mu)^T \Gamma^{-1} (x - \mu) \quad (3)$$

及相应的判定形式

$$RX(x) > T, \text{ 则 } x \text{ 为异常} \quad (4)$$

$$RX(x) \leq T, \text{ 则 } x \text{ 不为异常} \quad (5)$$

式中: $\Gamma = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)(x_i - \mu)^T$, μ 为样本均值, N 为高维样本数, T 为虚警率确定的阈值.

1.2 D-S 证据推理

D-S 证据推理^[8,9]建立在辨识框架 Θ 的基础上.设 $m: R \rightarrow [0, 1]$ 为基本置信指派,则证据由若干证据体 $(A, m(A))$ 组成,其中 A 为 Θ 的幂集中的元素.假定辨识框架 Θ 上性质不同的 n 个证据分别提供置信指派 m_1, m_2, \dots, m_n ,按照 Dempster 证据组合规则, n 个证据的正交和为组合结果,得到

$$m(A) = m_1(A) \oplus m_2(A) \cdots \oplus m_n(A) =$$

$$\frac{1}{1 - K} \sum_{A_1 \cap A_2 \cap \dots \cap A_n = A} \prod_{i=1}^n m_i(A_i) \quad (6)$$

式中: $K = \sum_{A_1 \cap A_2 \cap \dots \cap A_n = \emptyset} \prod_{i=1}^n m_i(A_i)$

由上面的新基本置信指派可以得到新证据体.

*国家自然科学基金(60172037)、航空科学基金(03D53032)、武器装备预研基金(51401040204HK0359)和西北工业大学科技创新基金资助

Tel: 029-88495954-802 Email: helin@mail.nwpu.edu.cn

收稿日期: 2005-04-28

2 波段子集特征提取和高阶矩可分性度量

高光谱图像各波段在成像过程中受波段响应特性、大气吸收和成像系统噪声等因素的影响不同,使得图像数据呈现很大的复杂性,但相近波段图像的特性有较大相关性,某个包含数个连续波段的波段子集包含了图像数据中的某些局部光谱特性. 将所有波段的图像划分为若干波段子集,每个波段子集包含数个连续波段的数据,对每个波段子集内的数据使用式(3)进行特征提取,得到相应的特征图,这样使得波段子集内所含的数个波段的图像数据压缩为一幅特征图.

数据样本的高阶统计量体现数据中包含小目标信息的情况,特征空间样本高阶矩值大的波段子集包含更重要的目标信息. 作为图像特征空间数据的三阶和四阶统计量,样本偏度和峰度为

$$h_{skewness} = \frac{1}{mn} \sum_{j=1}^m \sum_{i=1}^n (x_{ij} - \bar{x})^3 / \left(\frac{1}{mn} \sum_{j=1}^m \sum_{i=1}^n (x_{ij} - \bar{x})^2 \right)^{3/2} \quad (7)$$

$$h_{kurtosis} = \frac{1}{mn} \sum_{j=1}^m \sum_{i=1}^n (x_{ij} - \bar{x})^4 / \left(\frac{1}{mn} \sum_{j=1}^m \sum_{i=1}^n (x_{ij} - \bar{x})^2 \right)^2 \quad (8)$$

式中 $\bar{x} = \frac{1}{mn} \sum_{j=1}^m \sum_{i=1}^n x_{ij}$, m, n 分别为图像的长和宽.

样本偏度衡量样本分布的对称性,样本峰度衡量样本分布的平坦度. 因此,样本偏度、峰度值既可以体现噪声对真实数据的影响程度(噪声程度越大,样本峰度和偏度值越小),也可以体现目标与背景的可分性(可分离程度越大,样本峰度和偏度值越大). 样本偏度和峰度越大的数据越有利于将目标与背景进行分离.

3 波段子集特征融合检测

结合本文所要解决的问题,设辨识框架 $\Theta = \{h_1, h_2\}$, 其中: h_1 表示目标, h_2 表示背景. 使用由波段子集提取出的特征和特征对应的样本高阶矩构造用于推理的基本置信指派函数,并考虑以下因素: ①对于每个波段子集对应的特征图,假设该特征图中元素为从某正态分布总体取出的样本,由于小目标出现的概率小,可将特征图的分布视为背景像素特征的分布,即背景像素点特征服从正态分布; ②特征图对应的高阶统计量体现出了图像中的噪声程度和目标、背景的可分离性程度. 故针对高光谱图像某位置处的光谱信号向量 x , 构造第 i 个证据对该

位置处所属类别的基本置信指派函数为

$$m_i(h_2) = W_i \cdot \exp\left(-\frac{[(x_i - \mu_i)^T \Gamma_i (x_i - \mu_i)]^2}{M_i^2}\right) \quad (9)$$

$$m_i(h_1) = W_i \cdot [1 - \exp\left(-\frac{[(x_i - \mu_i)^T \Gamma_i (x_i - \mu_i)]^2}{M_i^2}\right)] \quad (10)$$

$$m_i(\Theta) = 1 - m_i(h_2) - m_i(h_1) = 1 - W_i \quad (11)$$

式中, x_i 是向量 x 在第 i 个波段子集内的分量构成的向量, Γ_i, μ_i 分别是第 i 个波段子集内的高光谱数据的样本协方差和均值向量, M_i 是第 i 个特征图的样本二阶中心矩, W_i 由下式确定

$$W_i = \frac{H_i}{\sum_{j=1}^k H_j} \quad (12)$$

式中 H_i 是第 i 个波段子集对应特征图的样本高阶统计量, k 为总共划分的波段子集数. 则 $m_i(h_1)$ 是第 i 个证据对空间某点为目标的支持程度, $m_i(h_2)$ 是第 i 个证据对该点为背景的支持程度, $m_i(\Theta)$ 是第 i 个证据对目标、背景判决表现出的不确定程度.

经过对基本置信指派函数的分析,确定以下用于检测目标点的规则: ①某点是目标的基本置信指派值应大于某一阈值; ②某点是目标的基本置信指派值应大于其背景的基本置信指派值; ③某点是目标的基本置信指派值应大于其不确定性基本置信指派值; ④不确定性基本置信指派应小于某一阈值.

4 检测算法步骤

本文算法如图 1, 步骤为: ①将待处理高光谱图像数据按波段分为若干个波段子集数据, 每个波段子集包含若干个连续波段, 波段子集之间连续但不重叠; ②用 RX 算法对每个波段子集进行特征提取,

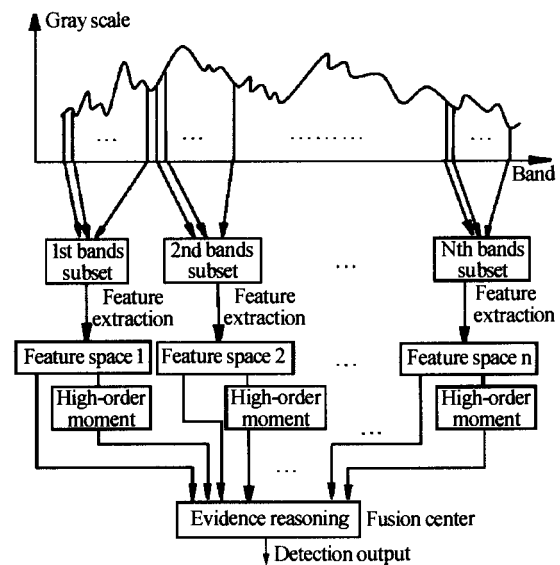


图 1 基于波段子集特征和证据推理的异常检测算法
Fig. 1 Anomaly detection algorithm based on features of band subset and evidence reasoning

使得每个波段子集对应一幅特征图；③对每一波段子集的特征图计算样本高阶统计量；④使用②③中的结果构造用于证据推理的基本置信指派函数；⑤使用多证据合成方法得到所有高光谱图像像素空间位置对应的基本置信指派值。⑥使用基于基本置信指派值的分类规则得到检测结果。

5 实验仿真与分析

试验对象是一套机载 HYDICE 高光谱图像数据，含有 210 个波段，光谱成像谱段范围为可见光/近红外波段，光谱分辨率约为 10nm。在图像中有 5 个人造物异常点，这些异常点具有与背景不同的光谱特征。图 2 给出了 HYDICE 高光谱成像仪得到的具有代表性的不同波段的图像(将数据规范化到 0~255 范围内显示)。图 2(a)第 4 波段图像成像良好；图 2(b)和图 2(c)的 106 和 149 波段分别处于不同的大气吸收带内，所成的图像几乎无有用信息；图 2(d)的 204 波段图像受噪声污染较为严重。图 2 中的图像表明了高光谱图像数据不同波段成像的复杂性。

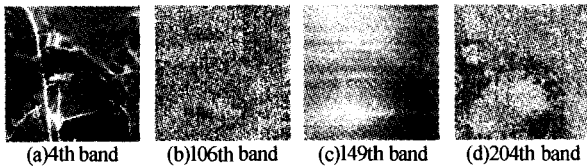


图 2 HYDICE 不同波段的图像
Fig. 2 HYDICE images of different bands

由于相邻波段图像的特性有较大相关性，可将连续相邻的数个波段划分为一个波段子集，划分的原则是要求同一波段子集内的若干幅图像具有类似的成像特性，而不同波段子集提取出的目标、背景可分性特征的差异较大，适于采用证据推理进行融合检测。对于本文实验图像，考虑以上原则，将待检测高光谱图像分为 21 个波段子集，每个波段子集含有 10 个相邻波段。对每个波段子集使用 RX 算法进行特征提取，得到 21 个特征图，其中包含图 2 中第 4、106、149、204 波段图像的第 1、11、15、21 波段子集对应的特征图如图 3。对 21 个特征图分别利用式(7)、(8)计算样本偏度和峰度值得到表 1 中的数据，表 1 中的 W_i 一栏的数据由式(11)计算得出。从表 1 中的数据可以看出，由 101~110 波段、141~150 波段、201~210 波段这三个波段子集的特征得到的偏度值和峰度值明显偏低，这正是由于 HYDICE 图像的 101~110 波段、141~150 波段处在大气吸收波段，而 201~210 波段内的图像几乎完全被噪声淹没；一些波段子集得出了较高的峰度和偏度值，如 1~10 波段、11~20 波段，说明这些波段子集受噪声影响较少，且目标和背景的光谱特性在这些波段范

围内体现了较大的可分离性。

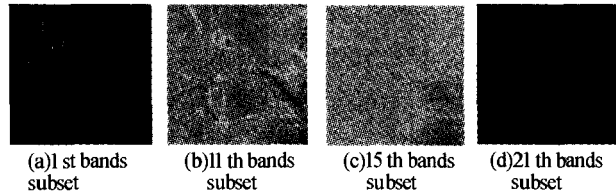


图 3 不同波段子集对应的特征图
Fig. 3 Features of different band subsets

表 1 不同波段子集特征的样本偏度和峰度

波段子集	样本偏度	W_i	样本偏度
1	120.7	0.458	7.8×10^6
2	33.9	0.129	1.8×10^7
3	29.2	0.111	2.0×10^7
4	14.2	0.054	1.6×10^6
5	8.7	0.033	3.4×10^3
6	10.4	0.040	2.0×10^3
7	1.9	0.007	2030
8	2.6	0.01	4244
9	4.1	0.016	15016
10	5.1	0.019	41292
11	0.4	0.001	351
12	1.1	0.004	778
13	6.1	0.023	78822
14	4.3	0.016	15596
15	0.2	0.001	651
16	2.7	0.01	4778
17	4.4	0.016	26209
18	5.2	0.02	52473
19	3.8	0.014	13877
20	4.2	0.016	20090
21	0.5	0.002	467

将特征图中数据和表 1 中的样本偏度数据代入式(9)、(10)、(11)，获取每个波段子集的基本置信指派值，使用式(6)将 21 组证据合成，获取每个像素点的最终基本置信特征。运用证据推理的分类判定准则得到最终检测结果。表 2 为用本文算法和 RX 算法对高光谱图像中不同目标进行检测得到的虚警率。从该表中可以看出，本文算法和 RX 算法对不同的目标敏感程度不同。比如，RX 算法检测到目标 3 时虚警率最小，检测到目标 5 时虚警率最大。本文算法检测到目标 1 时虚警率最小，检测到目标 4 时虚警率最大。在相同虚警率下，本文算法可以检测出 RX 算法检测不出的目标，且总体检测性能没有下降。

表 2 不同算法检测到不同目标时的虚警率 ($\times 10^{-5}$)

目标编号	本文算法	RX 算法
1	4.25	296.5
2	78.25	228
3	12	15
4	2348.5	18
5	254	3741

除了可用特征图的样本偏度构造基本置信指派

函数,也可使用样本峰度或是样本峰度与偏度的组合形式来进行基本置信指派函数的构造. 波段子集的划分要合适,如果波段子集数增多,则运算量增大;如果波段子集过少,每个子集内所含波段数过多,就不能充分体现某些噪声波段对成像的影响,不能充分体现噪声、目标的可行性,进而影响检测结果.

6 结论与问题

本文算法是一种以波段子集特征及其高阶统计量为基础运用证据推理的融合检测算法. 在对高光谱图像中小目标的检测中,本文算法能够自动去除和削弱受大气吸收和噪声影响较严重的波段的图像对检测结果的影响,通过波段子集特征的融合检测出 RX 算法不能检测出的目标,算法具有有效性. 但该算法也存在一些问题. 推理过程中基本置信指派函数的构造受到主观因素的影响,所依赖的数学模型可能并不完全符合所处理数据的内在规律,这样增加了检测结果的不确定性,对具有某些光谱特性的目标难以检测到,可以将该算法与 RX 算法或 LPD 算法结合使用,提高整体检测效果.

参考文献

- 薛彬,赵葆常,杨建峰,等. 改进的线性混合模型用于高光谱分离实验模拟. 光子学报, 2004, **33**(6):689~692
Xue B, Zhao B C, Yang J F, et al. *Acta Photonica Sinica*, 2004, **33**(6):689~692
- 唐自力,马彩文,刘波,等. 单站光测图像确定空间目标三

- 维姿态. 光子学报, 2004, **33**(12): 1480~1485
Tang Z L, Ma C W, Liu B, et al. *Acta Photonica Sinica*, 2004, **33**(12):1480~1485
- 程咏梅,潘泉,张洪才,等. 信息融合图像识别算法及其在三维飞机图像识别中的应用研究. 航空学报, 2004, **25**(2):176~179
Cheng Y M, Pan Q, Zhang H C, et al. *Acta Aeronautica et Astronautica Sinica*, 2004, **25**(2):176~179
- 柴饶军,马彩文. 图像序列中目标关键帧快速搜索算法. 光子学报, 2004, **33**(10):1233~1235
Chai R J, Ma C W. *Acta Photonica Sinica*, 2004, **33**(10):1233~1235
- Yu X, Reed I S, Stocker A D. Comparative performance analysis of adaptive multispectral detectors. *IEEE Transactions on Signal Processing*, 1993, **41**(8):2639~2656
- Yu X, Hoff L E, Reed I S, et al. Automatic target detection and recognition in multispectral imagery: A unified ML detection and estimation approach. *IEEE Transactions on Image Processing*, 1997, **6**(1):143~156
- Joseph C H. Detection and Classification of Subpixel Spectral Signature in Hyperspectral Image Sequences. Ph. D. dissertation, University Maryland, Baltimore County, 1993. 56~80
- Dempster A P. Upper and lower probabilities induced by a multi-valued mapping. *Ann Mathematical Statistics*, 1967, **38**(2):325~339
- Shafer G. A Mathematical Theory of Evidence. Princeton:Princeton University Press, 1976

Anomaly Detection in Hyperspectral Imagery Based on Feature Fusion of Band Subsets

He Lin, Pan Quan, Zhao, Yongqiang, Zhen Jiwei, Wei Kun

College of Automation, Northwestern Polytechnic University, Xi'an 710072

Received date:2005-04-28

Abstract Detecting camouflaged targets in an unknown environment presents a great challenge in hyperspectral image analysis since the prior knowledge about targets and background is not available. A anomaly detection method for hyperspectral imagery was proposed for this problem. Features were extracted from subband sets of hyperspectral imagery, then fusion algorithm for detection was implemented by D-S evidence reasoning while basic belief assignment function was constructed involving high-order moments of features. Theoretical analysis and results of experiment verify the effectiveness of the algorithm.

Keywords Hyperspectral imagery processing; Target detection; Feature fusion; Evidence reasoning; Band subsets

He Lin was born in Oct. 1973. He received his M. S. degree in Chongqing University in 2003. Now he is a Ph. D. candidate in Northwestern Polytechnical University. His current research interest is multidimensional signal processing, information fusion and multispectral/hyperspectral imagery processing.

