

文章编号:1001-9014(2006)06-0473-05

基于光谱解译的高光谱图像奇异检测算法

谷延锋, 刘颖, 贾友华, 张晔

(哈尔滨工业大学 信息工程系, 黑龙江 哈尔滨 150001)

摘要: 奇异检测是高光谱图像的重要应用之一。针对目前研究中存在的问题, 提出了一种新的基于光谱解译的奇异目标检测算法, 用于高光谱图像处理。该算法利用光谱解译技术有效地实现了目标信息和复杂背景分离, 很好地抑制了背景对检测的干扰。解译后的误差数据仅包含丰富的目标信息且更好地服从高斯分布。利用主成分分析对解译误差数据进行变换, 根据高阶统计量, 定义局部平均奇异度来选择对于奇异检测最有效的主分量, 并利用 RX 算子完成最终检测。为验证算法的有效性, 利用真实的 AVIRIS 数据进行了仿真实验。结果表明该算法能够较大地改进经典 RX 算法的检测性能。

关键词: 高光谱图像; 奇异检测; 主成分分析

中图分类号: TP751.1 **文献标识码:** A

ANOMALY DETECTION ALGORITHM OF HYPERSPETRAL IMAGES BASED ON SPECTRAL ANALYSES

GU Yan-Feng, LIU Ying, JIA You-Hua, ZHANG Ye

(Department of Information Engineering, Harbin Institute of Technology, Harbin 150001, China)

Abstract: Anomaly detection is an important application of hyperspectral images. Aiming at the problems in current researches, a new anomaly detection algorithm based on spectral analyses was proposed for hyperspectral images. The algorithm used spectral analyses technique to effectively separate target information from complicate backgrounds and greatly suppressed background interferences for detection. And those error data obtained after spectral unmixing with background endmembers only include abundant target information and better follow Gaussian distribution. Principal component analysis was used to transform the error data, and local average singularity was defined to select the most effective principal component for anomaly detection according to high-order statistics. Final detection was realized with conventional RX detector. To validate the effectiveness of the proposed algorithm, numerical experiments were conducted on real AVIRIS data. Experimental results show that the proposed algorithm greatly outperforms the conventional RX algorithm.

Key words: hyperspectral images; anomaly detection; principal component analysis

引言

成像光谱仪在对目标空间特征成像的同时, 对每个空间像素经过色散形成几十个乃至几百个窄波段以进行连续的光谱覆盖, 从而形成具有纳米级光谱分辨率的遥感数据立方体, 即高光谱图像。由于具有较高的光谱分辨率, 高光谱图像比多光谱图像在遥感地物分类、目标检测识别等方面具有更广泛的应用^[1]。

目标检测和奇异检测是高光谱图像应用的主要方面。由于目标的先验光谱信息在实际处理中很难

得到, 因此, 不需事先已知目标先验信息的奇异检测更符合实际需求, 逐渐成为了研究的热点^[2,3]。但相对于目标检测技术来说, 奇异检测技术的发展较为缓慢, 奇异检测没有目标光谱信息可以利用, 主要依靠计算局部区域的统计变化来检测奇异目标^[4,5]。经典的奇异检测算法是源自多光谱图像的 RX 算法, RX 算法最早是由 Reed 和 Xiaoli Yu 两人提出的^[6,7]。RX 算法在一些简化的假设条件下构造似然比检测算子。直接利用 RX 算法对高光谱图像进行处理将会产生较高的虚警概率, 这主要是由三方面原因造成的: 一、RX 算法中所采用的局部统计模型

收稿日期: 2005-08-15, 修回日期: 2006-05-15

Received date: 2005-08-15, revised date: 2006-05-15

基金项目: 国家自然科学基金(60472048 和 60402025)资助项目

作者简介: 谷延锋(1977-), 男, 黑龙江佳木斯人, 博士/副教授, 主要从事高光谱图像处理、核机器学习算法方面的研究。

假定数据是空间不相关的,或者是空间白化的,且数据要求服从局部正态分布,这种假设在真实场景中常常得不到满足;二、当 RX 算法直接用于高光谱图像处理时,需要计算样本的协方差矩阵,其维数随着波段数目的增加而迅速增加,这将会带来巨大的计算量^[8];三、高光谱图像波段之间具有很强的分块相关特性,这直接影响着 RX 算法的性能.因此,将 RX 算法直接用于原始高光谱图像处理并不合理.为此, Schweizer^[8]等人提出了一种基于三维高斯马尔可夫随机场(Gaussian Markov Random Field, GM-RF)的奇异目标检测算法.该算法主要有三方面的优点:一、该算法利用三维高斯马尔可夫随机场来描述高光谱数据,更符合高光谱数据图谱合一的特点;二、该算法直接对样本协方差矩阵的逆矩阵进行描述,不需计算样本协方差矩阵本身;三、多变量高斯数据的概率密度函数可以用样本协方差矩阵的逆矩阵来表示.但是,该方法也存在着一个较大的缺陷,即三维高斯马尔可夫随机场无法对较小空间尺寸的区域进行描述,这决定它不能较好地检测仅占几个像素的奇异小目标,且统计模型参数的估计也比较复杂.在 RX 算法的基础上, Kwon^[9]等人进一步地提出了一种基于核的非线性 RX 改进算法,该算法能够更好地利用高光谱图像波段之间的非线性统计特性,但仍然没有较好地解决背景干扰问题.在国内,李智勇等人开展了高光谱图像小目标检测研究,并提出了一种基于主成分分量分析的检测算法^[10].综合目前国内外研究的现状,严重的背景干扰和高光谱图像所具有的高数据维仍是奇异检测所面临的主要问题.

针对现有问题,本文提出了一种基于光谱解译(Spectral Unmixing)的高光谱图像奇异检测新算法.该算法能够有效地抑制背景干扰,提取包含最多目标信息的主分量以降低维数并实现有效的检测.为验证该算法的有效性,最后利用真实的 AVIRIS 数据进行了奇异检测实验.

2 基于光谱解译的奇异检测算法

在实际检测场景中未知目标光谱的情况下,特别是当目标的空间尺寸相对于高光谱图像的空间分辨率较小时,也就是说目标在高光谱图像中仅占较少的像素(几个或者十几个),目标有着较强的稀疏性和奇异性.因此,可以合理地假设利用背景端元所获取的解译数据中几乎不包含奇异目标信息.这样,关于奇异目标的信息必然存在于光谱解译后的误差数据中,

也即,奇异检测可以在解译误差数据上进行.

本文算法首先利用纯像素自动提取技术来获取复杂的背景信息,然后利用背景光谱信息进行解译处理,将奇异目标信息集中在解译误差数据中,利用主成分分析技术对变换后的误差数据进行特征提取,并定义平均局部奇异度,据此来选择具有较大奇异性的特征分量,结合经典的 RX 奇异检测算子,构造一个有效的基于光谱解译的高光谱奇异检测算法.

2.1 混合像素光谱解译

光谱解译(也称作解混叠)是指在观测像素中预测多种类别物质比例信息的过程.光谱解译可以实现小于一个像素的物质光谱信息的预测和获取,因此可以用来进行已知光谱特性的目标识别处理.光谱解译主要包括三部分内容:混合模型的选取、模型求解和光谱端元(Endmember)的自动提取.针对本文的具体情况,选用线性模型.

设 $\hat{\mathbf{r}}$ 是一个 $L \times 1$ 的列向量,它描述的是高光谱图像中观测到的某像素的光谱向量.设 \mathbf{S} 是一个 $L \times p$ 的矩阵,它描述的是该高光谱图像中存在的光谱端元向量矩阵,它可以记为 $[s_1, s_2, \dots, s_p]$, 其中 s_j 是一个 $L \times 1$ 列向量, p 表示的是某一类纯地物的光谱特性曲线,表示的是图像中所存在的纯地物的总数.设 $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_p)^T$ 是一个 $p \times 1$ 的列向量,即各个纯地物在某像素中的混合比例.线性光谱解译(Linear Spectral Unmixing)假定混合像素中的各类纯地物是按照一定比例线性混合的.这样像素向量 $\hat{\mathbf{r}}$ 可以利用如下线性模型表达

$$\hat{\mathbf{r}} = \mathbf{S}\boldsymbol{\alpha} + \mathbf{n} \quad (1)$$

其中, \mathbf{n} 是存在于像素 $\hat{\mathbf{r}}$ 各个波段中的噪声向量,通常假设为高斯白噪声.在理想情况下, $\mathbf{r} = \mathbf{S}\boldsymbol{\alpha}$, $\boldsymbol{\alpha}$ 是观测测量的真实值.设 $\hat{\mathbf{r}} = \mathbf{r} + \mathbf{n}$, 则令观测误差的协方差矩阵为 $\mathbf{V} = E(\mathbf{nn}^T)$.

为了保证获得的比例具有一定的物理意义,通常给线性模型(1)加上两个约束条件,即非负性约束和归一化约束:

$$\sum_{j=1}^p \alpha_j = 1 \quad (2)$$

$$\alpha_j \geq 0 \quad (j = 1, 2, \dots, p) \quad (3)$$

对于服从式(2)和(3)两个约束的线性模型求解,一些文献都给出了可行的方法^[11,12].

针对光谱端元提取,在本文算法中,仅对背景物质类别进行解译,而且背景通常具有较大面积的分布,因此无需采用复杂的光谱端元自动提取技术就

可以很好地选择出背景光谱端元,这也是本算法的一个优点.此外,根据进行线性光谱解译的理论前提,解译误差数据能够较好地服从高斯分布,这对后续特征提取和检测来说是很有用的.

2.2 奇异特征的提取和选择

在光谱解译之后,本文算法采用主成分分析对误差数据进行特征提取.变换之后的数据维数同误差数据相同,但是能量更为集中,这样使得奇异目标的能量都集中在前面的少数几个成分中.为了选择能够包含丰富奇异目标信息的主分量,本文算法考虑了RX算法中对于奇异性的一个有效度量.在RX算法的假设当中,认为局部区域内背景均一,那么背景向量将服从高斯分布.理论上,对于高斯概率密度分布,扭曲度和峭度(即三阶矩和四阶矩)为零.但是如果内部有奇异点,则高斯分布被破坏,表现为扭曲度和峭度变大.通常,扭曲度用来作为分布对称性的度量,峭度被用来度量分布的重尾特性,峭度和扭曲度主要用来衡量局部区域的奇异性.因此,计算整个主分量图的峭度和扭曲度是不合理的.本文算法利用扭曲度和峭度定义了局部平均奇异度,记为 \bar{N}_A ,用它来衡量变换后的主分量的奇异性.在本文中,局部平均奇异度定义为单个主分量中奇异性大于给定阈值的局部区域数目,其中阈值是随着局部区域自适应设定的.大于给定阈值的局部区域数目越大,说明该主分量集中了越多的奇异目标信息.根据定义,局部平均奇异度的具体计算过程如下:

- (1) 首先令 $\bar{N}_A = 0$;
- (2) 在主分量图内划分局部窗口,计算局部的扭曲度 $skew$ 和峭度值 kur ;
- (3) 将(2)得到的结果分别同阈值 T_s 和 T_k 进行比较,如果 $|skew| > T_s$ 并且 $|kur| > T_k$,则 $\bar{N}_A = \bar{N}_A + 1$,否则保持不变;
- (4) 判断整个图像属否处理完毕,如果完毕,输出 \bar{N}_A ,否则返回(2)继续处理下一个窗口.

在上述处理过程中,最后输出的 \bar{N}_A 就是该主分量的局部平均奇异度.阈值 T_s 和 T_k 分别取值为 $\sigma_s\theta_s$ 和 $\sigma_k\theta_k$,其中 σ_s 和 σ_k 分别代表三阶矩和四阶矩的标准偏差,通常取值为 $\sqrt{6/N}$ 和 $\sqrt{24/N}$, N 是局部窗口大小, $\theta_s \gg 1$ 且 $\theta_k \gg 1$,通常取值为11^[6].可以看出,阈值 T_s 和 T_k 随局部窗口的大小自适应变化的.

2.3 RX检测算子

RX算法是源自于多光谱图像奇异检测的经典算法,适用于具有较少波段的遥感数据源.在RX算法中通常假定目标光谱和背景的协方差未知,以图

像中的每个像素为中心,确定局部检测窗,建立二值假设: H_0 代表目标不存在的假设, H_1 代表目标存在的情况,并假定两种情况下数据的分布具有相同的方差和不同的均值,分别服从高斯分布 $N(\mu_b, C_b)$ 和 $N(\mu_s, C_b)$,其中 μ_b 和 μ_s 分别为局部窗口内的背景均值和奇异目标均值, C_b 为背景方差.RX算子简化的判决式是

$$RX(r) = (r - \hat{\mu}_b)^T \hat{C}_b^{-1} (r - \hat{\mu}_b) \begin{cases} \geq \eta & \text{接受 } H_1 \\ < \eta & \text{接受 } H_0 \end{cases}, \quad (4)$$

在上式中, r 是待检测像素, η 是判决阈值, $\hat{\mu}_b$ 和 \hat{C}_b 分别是 μ_b 和 C_b 的统计估值.

2.4 本文检测算法的步骤

现在,我们将本文提出算法的具体实现步骤归纳为如下:

- (1) 输入高光谱图像;
- (2) 利用端元提取算法获取背景类别的光谱;
- (3) 根据背景端元光谱,利用基于线性模型的方法对高光谱图像进行混合像素解译,获得背景端元在图像中的混合比例;
- (4) 根据(2)和(3)所获取的背景端元光谱和相应混合比例,由原始高光谱图像中去除背景类别信息,得到解译误差数据;
- (5) 对解译误差数据进行白化处理,并利用主成分分析对白化的误差数据进行特征提取;
- (6) 针对变换后的主成分分量,根据2.2节所提供的具体步骤,计算局部平均奇异度;
- (7) 挑选出具有最大平均局部奇异度的主成分分量,利用RX算子对其进行检测,输出检测结果.

需要注意,计算局部平均奇异度和利用RX算子进行检测时,所采用的局部窗口大小一般是相同的.

3 实验结果及分析

为了验证本文提出算法的有效性,利用真实的

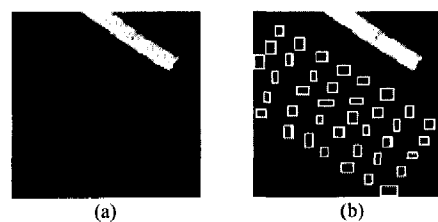


图1 用于奇异检测的高光谱图像第6波段和相应的目标分布(a)高光谱图像第6波段(b)真实目标分布
Fig.1 The 6th band of hyperspectral images for anomaly detection and the corresponding ground distribution of targets. (a) the 6th band of hyperspectral images (b) real distribution of targets

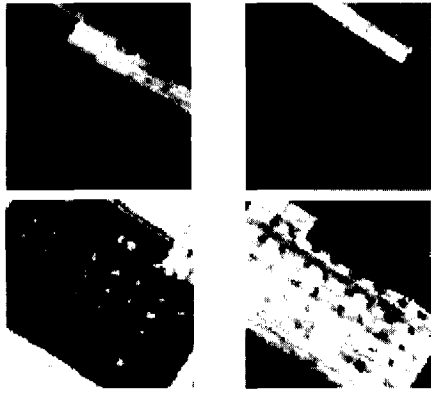


图2 4类背景解译分量
Fig.2 The unmixed components of four backgrounds

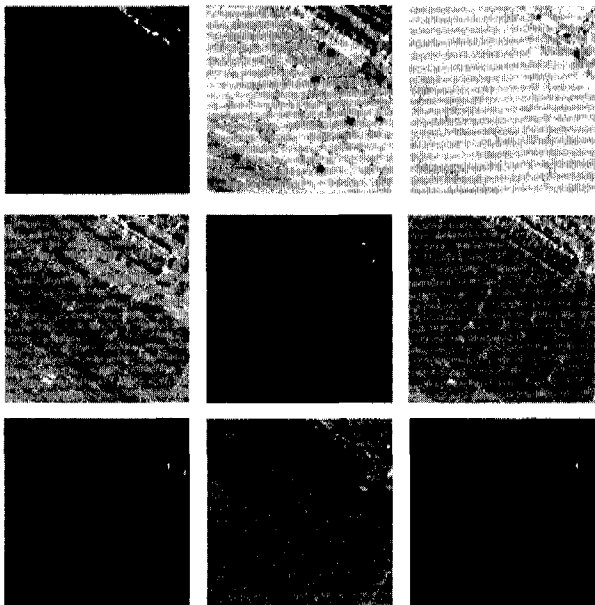


图3 解译误差数据的前9个变换的主成分分量
Fig.3 The first-nine principal components transformed from the analysed error data

AVIRIS 高光谱数据进行了仿真实验. 图1 给出了检测区域和相应奇异目标真实分布, 图1(a)给出的是原始高光谱数据的第6波段, 以此来显示图像场的情况. 在该检测区域内, 包含奇异目标共38个. 在进行光谱解译时, 选择了四类背景. 图2给出的是对四类背景解译所获得的分量图像(每类混合比例), 由该图可以看出, 在分量图像中, 同目标相混合的背景大部分都被分离出来. 图3给出了由主成分变换所得到的解译误差数据的前9个主成分分量.

由图3可以明显地看出, 第7个主分量(按照由左至右、由上至下的顺序)很好地将几乎全部的奇异目标信息都包含在其中. 图4给出了前30个主成分分量的局部平均奇异度变化曲线, 局部窗口大小

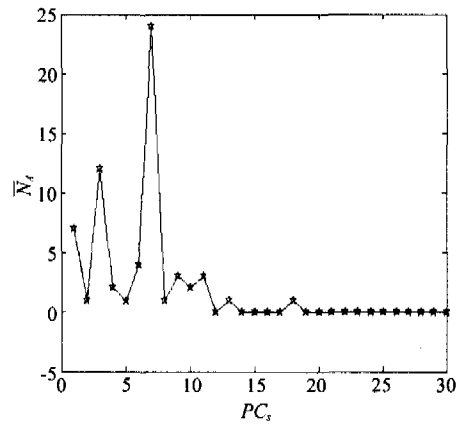


图4 前30个主成分分量的局部平均奇异度
Fig.4 Local average singularity of the first-thirty principal components

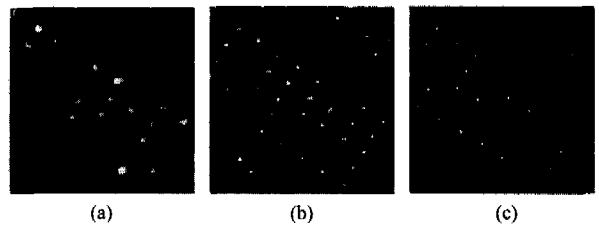


图5 3种方法检测结果的比较(a) GMRF (b) M-RX (c) 本文方法
Fig.5 Comparison of detection results with three methods (a) GMRF (b) M-RX (c) the proposed algorithm

N 选择为 11×11 . 由变化曲线可以看出, 第7个分量具有最大的奇异性, 从第11个主分量以后, 局部平均奇异度逐渐减小, 接近于0. 这证明了本文方法所定义的局部平均奇异度的有效性.

在检测算法中, 根据所选用的数据空间大小和分辨率, 将背景检测窗口和局部平均奇异度计算的窗口大小均设定为 11×11 . 图5给出了不同方法所获得的检测结果, 除本文方法外, 实验主要考察了GMRF算法和传统的多光谱RX算法(简记为M-RX), 其中M-RX是按照多光谱图像的波长选择相应的几个高光谱图像波段作为检测算法的输入. 由图5可以看出本文方法所获得的检测结果要明显好于其它两种方法, M-RX算法不能很好地抑制背景干扰, 产生较多虚警, GMRF方法不能检测出全部目标, 仅适用于检测具有一定空间尺寸的目标. 本文方法能够很好地抑制背景干扰, 这意味着在要求检测的目标数目相同的情况下, 本文方法将有更低的虚警. 由于GMRF方法并不适用于检测仅占很少像素的奇异目标, 实验仅对M-RX方法和本文方法进行

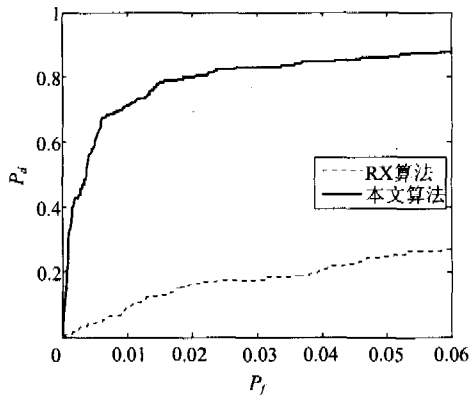


图6 本文算法同RX算法的ROC比较

Fig.6 ROC comparison between the proposed algorithm and RX algorithm

量化比较.

接收机操作特性(ROC)用于描述检测概率 P_d 同虚警概率 P_f 之间的变化关系,能够提供算法检测性能的定量分析.根据地面真实图可以获得每个目标所包含的像素区域坐标.通过考察检测到的奇异像素是否落入真实图的目标像素区域可以来判定检测得到的是真实目标还是虚警.我们将检测概率定义为检测到的像素数目同地面真实目标像素数目的比值,将虚警概率定义为检测到的虚警数目同整个观测场内像素总数的比值.图6给出了2种算法ROC特性的比较,可以看出,本文所提出的基于光谱解译的奇异检测算法(在图6中标识为“本文算法”)能够有效地改进经典RX算法的性能,具有良好的检测性能和较低的虚警.

4 结语

针对现有算法不能很好地解决背景干扰问题,本文提出了一种基于光谱解译的高光谱图像奇异目标检测新方法.该方法仅对背景光谱进行提取,降低了解译处理的复杂度,并实现了背景光谱信息的分离,很好地抑制了背景干扰,使得奇异目标信息都存在于解译误差数据中;利用主成分分析对误差数据进行变换,有效地集中了奇异目标信息;根据高阶统计量,定义局部平均奇异度,能够很好地选择具有最

大平均奇异度的主分量.仿真实验结果表明,该方法要明显好于经典的多光谱RX检测算法,具有良好的检测性能和较低的虚警.

REFERENCES

- [1] Landgrebe David. Hyperspectral data analysis [J]. *IEEE Signal Processing Magazine*, 2002, 1: 17—28.
- [2] Manolakis D, Shaw G. Detection algorithms for hyperspectral imaging applications [J]. *IEEE Signal Processing Magazine*, 2002, 1: 29—43.
- [3] Manolakis Dimitris, Siracusa Christina, Shaw Gary. Hyperspectral subpixel target detection using the linear mixing model [J]. *IEEE Trans. Geosci. Remote Sensing*, 2001, 39 (7): 1392—1409.
- [4] David W J S, Scott G B, Lawrence E H. Anomaly detection from hyperspectral imagery [J]. *IEEE Signal Processing Magazine*, 2002, 1: 58—69.
- [5] Chang Chein-I, Chiang Shao-Shan. Anomaly detection and classification for hyperspectral imagery [J]. *IEEE Trans. Geosci. Remote Sensing*, 2002, 40(6): 1314—1325.
- [6] Reed I S, Yu X. Adaptive multiband CFAR detection of an optical pattern with unknown spectral distribution [J]. *IEEE Transactions Acoustics, Speech, and Signal Processing*, 1990, 38(10): 1760—1766.
- [7] Yu X, Hoff L E, Reed I S. Automatic target detection and recognition in multiband imagery: a unified ml detection and estimation approach [J]. *IEEE Trans. Geoscience and Remote Sensing*, 1997, 6: 143—156.
- [8] Schweizer Susan M, Moura José M F. Hyperspectral imagery: Clutter adaptation in anomaly detection [J]. *IEEE Tran. On Information Theory*, 2000, 46(5): 1855—1871.
- [9] Kwon H, Nasrabadi N M. Kernel RX-algorithm: A nonlinear anomaly detector for hyperspectral imagery [J]. *IEEE Trans. Signal Processing*, 2005, 43(2): 388—397.
- [10] LI Zhi-Yong, KUANG Gang-Yao, YU Wen-Xian, et al. Algorithm on small target detection base on principal component of hyperspectral imagery [J]. *J. Infrared Millim. Waves* (李智勇, 匡纲要, 郁文贤, 等. 基于高光谱图像主成分分量的小目标检测算法研究. *红外与毫米波学报*), 2004, 23(4): 286—290.
- [11] Du Qian, Chang C-I. Linear mixture analysis-based compression for hyperspectral image analysis [J]. *IEEE Trans. Geosci. Remote Sensing*, 2004, 42(4): 875—891.
- [12] ZHOU Hao, WANG Bin, ZHANG Li-Ming. New scheme for decomposition of mixed pixels of remote sensing images [J]. *J. Infrared Millim. Waves* (周昊, 王斌, 张立明. 一种新的遥感图像混合像元分解方法. *红外与毫米波学报*), 2005, 24(6): 463—466.