

文章编号: 0253-2239(2009)03-0844-04

基于多重分形谱的高光谱数据特征提取

刘小刚^{1,2} 赵慧洁^{1,2} 李娜^{1,2*}

¹北京航空航天大学仪器科学与光电工程学院,北京 100083

²北京航空航天大学精密光机电一体化技术教育部重点实验室,北京 100083

摘要 针对单一分形维数在高光谱数据处理中的不足,提出了一种基于多重分形谱的光谱信号奇异性特征提取方法,引入多重分形谱表征光谱曲线的奇异性特征。该方法根据分形测度将光谱曲线进行划分,用光谱概率测度计算配分函数,通过尺度指数的 Legendre 变换实现光谱曲线多重分形谱的提取,根据各类地物间的类别可分性准则 Bhattacharyya 距离选择有效特征,最后利用地物分类实验来验证该方法的有效性。实验结果表明,多重分形谱用于分类时分类精度达 95.2%,当其维数为原数据波段数的 10% 时,总体分类精度仍可达 82.2%。多重分形谱表征了具有相同奇异性的波段子集的分形维数,准确的描述了光谱曲线的奇异性分布特点,该方法能够有效地实现高光谱数据的特征提取。

关键词 高光谱遥感; 多重分形谱; 配分函数法; 奇异性指数

中图分类号 TP751.1 **文献标识码** A **doi**: 10.3788/AOS20092903.0844

Feature Extraction Based on Multifractal Spectrum for Hyperspectral Data

Liu Xiaogang^{1,2} Zhao Huijie¹ Li Na^{1,2}

¹*School of Instrument Science and Opto-Electronics Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100191, China*

²*Key Laboratory of Precision Opto-Mechatronics Technology, Ministry of Education, Beijing University of Aeronautics and Astronautics, Beijing 100191, China*

Abstract Considering the fractal dimension deficiency to process hyperspectral data, a singularity feature extraction method was proposed, and the multifractal spectrum was used to characterize the singularity feature of spectra. In this method, the spectral curves were divided to several segments according to fractal measure, and the partition function was generated with the spectral probability measure. The multifractal spectrum was extracted with the Legendre transformation of scale exponent. Effective features of multifractal spectrum were selected based on discriminable rule of Bhattacharyya distance. Classification experiments of hyperspectral data are carried out to prove the value of multifractal spectrum, and the classification accuracy reaches 95.2%. With 10% of the original spectra's dimension, the accuracy reaches 82.2%. The fractal dimension of spectral subset with the same singularity exponent is characterized by multifractal spectrum, and the singularity distribution of spectra are expressed sufficiently. As a conclusion, the method is appropriate to extract the features of hyperspectral data.

Key words hyperspectral remote sensing; multifractal spectrum; partition function definition; singularity exponent

1 引言

高光谱图像是目前国内外在遥感图像处理领域

的研究热点之一,与多光谱图像相比,高光谱图像具有多得多的光谱分辨率,使得在多光谱图像下不能

收稿日期: 2008-06-19; 收到修改稿日期: 2008-09-12

基金项目: 中国地质调查局项目(1212010816033)和中国地质调查局科研项目(1212030616010)资助课题。

作者简介: 刘小刚(1982-),男,硕士研究生,主要从事高光谱遥感数据处理应用方面的研究。E-mail: lxxg8848@126.com

导师简介: 赵慧洁(1966-),女,教授,博士生导师,主要从事光电探测、高光谱遥感、仿真计算等方面的研究。

E-mail: hjzhao@buaa.edu.cn

* 通信联系人。E-mail: wherein@163.com

解决的问题在高光谱图像下可以得到解决,在目标探测识别^[1,2]、异常检测^[3]等问题中得到广泛的应用。然而高光谱图像具有较高光谱分辨率的优越性是以其较大的数据量及较高的数据维为前提的,这就为高光谱图像数据的分析与处理带来了较大的困难,因此通过特征提取和特征选择算法降低数据维数是非常必要的。分形理论在遥感图像解译中,常常被用来研究遥感影像的空间分形特征^[4],近年国内外开始将其用于波段选择进行数据降维^[5],并针对光谱曲线进行分形特征^[6,7]分析。分形维数是定量的表示自相似的随机形状和现象的最基本的量,对于现实中存在的物体,只有在某种被限制的观测尺度的范围内,自相似性才成立,仅利用分形维数一个数字去描述所有的复杂形状和现象是不够的^[8,9]。随着多重分形理论研究的深入,人们已将它成功地应用到其它学科中,如金融时间序列分析、地质学、气象学、图像处理等领域,其原理是基于配分函数下的多重分形形式体系。对于那些非均匀、不规则的现象,当一个维数无法描述其全部的细致特征时,就需要用多重分形的连续谱来表示。近 20 年来,有关指纹识别、图像分析^[10]、人脸识别和目标检测^[11]的研究都有不同程度的发展。而多重分形在遥感图像纹理特征提取方面已得到应用^[12]。

非孤立的奇异性出现在大多数分形信号中,针对光谱曲线的多重分形特征展开研究尚处于起步探索阶段,本文提出一种光谱曲线多重分形特征提取方法用于高光谱数据的处理,包括光谱曲线划分,配分函数计算和多重分形谱特征提取和基于类间可分性的特征选择。多重分形根据光谱曲线内在的奇异性对其进行自动划分,在无标度区间内更精细的描述了分形集的局部尺度行为,地物分类实验结果表明了该方法在高光谱数据特征提取方面的有效性。

2 多重分形与多重分形谱

2.1 多重分形谱定义

设信号 $s(nb) \subset F$, F 是 nb 维欧氏空间 R^{nb} 上的 nb 维子集,它是测度 δ 的支集。在某种划分下, (F, δ) 产生的分形集可以表成若干分形子集的并,且每一分形子集有不同的分形维数,则 (F, δ) 称为多重分形^[13]。

把信号 $[s(nb), \delta]$ 划分为 N 个尺度最大为 e 的单元 $S_i (i = 1, 2, \dots, N)$, 设 δ_i 是 S_i 的线度大小, P_i 为 S_i 的概率测度,且 $\sum_{i=1}^N P_i = 1$ 。即信号概率测度 P 组

成的集合被划分为一系列满足下面的幂函数的子集 $P_i (i = 1, 2, \dots, N)$:

$$P_i \propto \delta_i^{\alpha_i}, \quad (1)$$

这里的 δ_i 是划分一维小盒子的尺寸; α_i 称为 Lipschitz 指数,也称奇异性指数。 α_i 与所在区域有关,它反映分形上测度的奇异程度。根据 α 值的大小可以将 $S(n, b)$ 分为许多小子集,使得每个子集中的小单元具有相同的 α 值,然后计算这个子集内的单元数 $N_\alpha(\delta)$, 将子集内的线段数 $N_\alpha(\delta)$ 和 δ 的关系式定义为

$$N_\alpha(\delta) \propto \delta^{-f(\alpha)}, \delta \rightarrow 0 \quad (2)$$

$f(\alpha)$ 表示有相同的 α 值的子集的分形维数,称为多重分形谱。一个复杂的分形体,它的内部可以分为一系列不同 α 值所表示的子集, $f(\alpha)$ 就给出了这一系列子集的分形特征。

2.2 多重分形谱的性质

$\alpha \sim f(\alpha)$ 是描述多重分形局部特征的一套基本语言,具有如下性质:

- 1) $f(\alpha)$ 是描述多重分形子集维数的连续谱,如果研究对象是单分形的,则 $f(\alpha)$ 为一定值;如果研究对象是多重分形的,则 $f(\alpha)$ 一般呈单峰图像;
- 2) $f(\alpha) \geq 0$;
- 3) 多重分形谱函数 $f(\alpha)$ 是关于 α 的凸函数。

多重分形考虑了盒子内信号之间的差别(不但考虑到信号所覆盖的盒子的数目,而且考虑到盒子内信号的多少),用一个多重分形谱描述分形不同层次的特征,全面反映了光谱信号的概率分布。多重分形谱的宽度可以定量表征光谱信号的振动程度,因此多重分形提供了一种能深入地研究分形物体性质的手段。

3 光谱曲线多重分形谱的提取与特征选择

3.1 多重分形谱特征提取

3.1.1 概率测度

图 1 所示为高光谱数据中某一地物光谱曲线,其中横坐标为波长,纵坐标为光谱辐亮度。沿横轴方向划分为 $N(\delta)$ 个尺寸为 δ 的一维小盒子, $S_i(\delta)$ 表示盒子尺寸为 δ 时第 i 个小盒子内所有光谱波段的辐亮度数值之和,则 i 个小盒子内光谱曲线的概率测度可表示为

$$P_i(\delta) = \frac{S_i(\delta)}{\sum_{i=1}^{N(\delta)} S_i(\delta)}, \quad i = 1, 2, \dots, N(\delta), \quad (3)$$

其中 $\sum S_i(\delta)$ 为所有光谱波段辐亮度之和。

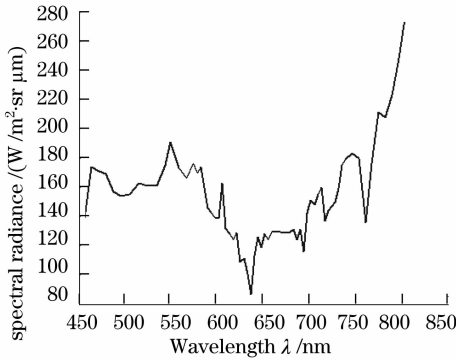


图 1 高光谱数据光谱曲线

Fig. 1 Spectral curve of hyperspectral data

3.1.2 配分函数

定义配分函数 $\chi_q(\delta)$ 为光谱概率测度 $P_i(\delta)$ 的 q 阶矩^[14]:

$$\chi_q(\delta) = \sum_{i=1}^{N(\delta)} P_i^q, i = 1, 2, \dots, N(\delta) \quad (4)$$

q 称为权重因子, 不同的 q 表示不同大小的概率测度 P_i 在配分函数 $\chi_q(\delta)$ 中所具有的比重, 从而突出特定的概率测度 P_i 对 $\chi_q(\delta)$ 的贡献。当 $q \gg 1$ 时, 具有较大概率测度 P_i 的大波动特征对 $\chi_q(\delta)$ 的贡献占优势; 相反地, 当 $q \ll -1$ 时, 具有较小概率测度 P_i 的小波动特征对 $\chi_q(\delta)$ 的贡献占优势。

在无标度的自相似区域内 $\chi_q(\delta)$ 与 δ 之间存在着如下的幂率关系:

$$\chi_q(\delta) \sim \delta^{\tau(q)}. \quad (5)$$

如果上式成立, $\text{lb}\chi_q(\delta_i)$ 作为 $\text{lb}\delta_i$ 的函数拟合得到 $\tau(q)$:

$$\text{lb}\chi_q(\delta_i) \approx \tau(q)\text{lb}\delta_i + C(q), \quad (6)$$

式中 $\tau(q)$ 为尺度指数, $C(q)$ 表示拟合直线的截距。

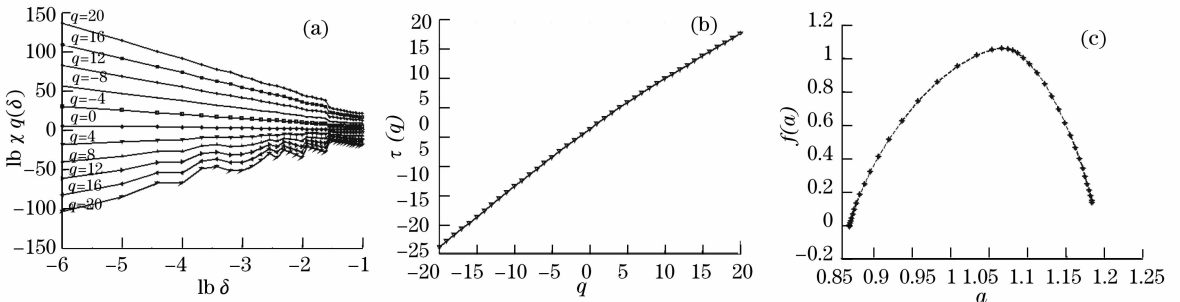


图 2 光谱曲线光谱多重分形谱计算结果 (a) 分形尺度与配分函数的对数关系; (b) 尺度指数; (c) 多重分形谱

Fig. 2 Extraction result of multifractal spectrum of spectra. (a) Logarithmic relation between fractal scale and partition function; (b) scale exponent; (c) multifractal spectrum

3.1.3 多重分形谱

多重分形谱 $f(\alpha)$ 和尺度指数 $\tau(q)$ 之间通过统计物理中的 Legendre 变换^[15] 联系起来, 并有如下关系式:

$$\begin{cases} \alpha(q) = \frac{d\tau(q)}{dq}, \\ f(\alpha) = q\alpha(q) - \tau(q) = q \cdot \frac{d\tau(q)}{dq} - \tau(q) \end{cases}, \quad (7)$$

由此得到了光谱曲线的多重分形谱 $\alpha \sim f(\alpha)$ 的关系。 $q = 0$ 时, $f(\alpha)$ 是盒维数; $q = 1$ 时, $f[\alpha(1)] = \alpha(1)$ 是信息量维数, $q = 2$ 时, $\tau(2)$ 是容量维数。

3.2 基于类别可分性的特征选择

提取光谱曲线的多重分形谱后, 所提取的谱维数取决于 q 的取值个数二者满足关系

$$D = 2 \cdot q_{\max} + 1. \quad (8)$$

为了从多重分形谱中获得最有效的光谱特征, 利用地物类别可分性准则进行特征的选择。选用的类别可分性准则是 Bhattacharyya 距离, 定义如下:

$$BH_{12} = \frac{1}{8}(\mu_2 - \mu_1)^T \left[\frac{\sum_1 + \sum_2}{2} \right]^{-1} (\mu_2 - \mu_1) + \frac{1}{2} \ln \left[\frac{|\sum_1 + \sum_2|/2}{\sqrt{|\sum_1| |\sum_2|}} \right], \quad (9)$$

其中 μ_1, μ_2 和 \sum_1, \sum_2 分别代表两类地物的统计均值和方差。Bhattacharyya 距离描述了不同地物类别间的可分性, Bhattacharyya 距离值越大, 地物间的可分性就越好。

3.3 光谱曲线多重分形分析

光谱信号表现出非平稳性和随机性, 对于图 1 所示的光谱曲线, 按照上述算法在 Matlab 6.5 中编程求得多重分形谱及相关参数如图 2 所示。

图 2(a)给出了 $\text{lb } \chi_q(\delta)$ 与 $\text{lb } \delta$ 的关系图。可以看出随着 $|q|$ 的增大,关系图出现了略微的浮动,但观察整个尺度的变化范围, $\text{lb } \chi_q(\delta)$ 与 $\text{lb } \delta$ 之间呈现了较好的线性关系,因此 $\chi_q(\delta)$ 与 δ 满足幂律关系,图线保持较好的线性衰减且汇聚于一点,说明在规定的尺度变化范围内光谱曲线具有较高的标度不变性,也就是说光谱曲线具有分形的特性。

若 $\tau(q)$ 关于 q 是一条直线,则研究对象是单分形的;若 $\tau(q)$ 关于 q 是凸函数,则研究对象有多重分形特征。图 2(b) 所示尺度指数中,可以看出 $\tau(q)$ 是一个上凸的函数,即 $\tau(q)$ 与 q 之间存在着非线性的关系,这表明光谱曲线确实具有多重分形的特性。

图 2(c)给出了光谱曲线的多重分形谱,为一单峰曲线,并且宽度较大,说明光谱的奇异性变化较大。图 2(a)~图 2(c)体现了信号的奇异性,与光谱曲线的特征是一致的,说明多重分形谱准确反映了光谱的奇异特征及其分布特点。

4 实验结果与分析

为了验证多重分形谱用于高光谱数据特征提取的有效性,以分类精度和 kappa 系数作为评价标准,用上海技术物理研究所研制的 PHI 成像光谱仪拍摄的江苏方麓茶场航空数据进行实验,原数据大小为 $210 \times 150 \times 80$,去除坏波段以及信噪比低的波段后,采用的数据大小为 $210 \times 150 \times 64$,波段区间 $455 \sim 805 \text{ nm}$,数据包含 6 类主要地物,分别为水稻(C4)、竹(T6)、茶(T7)、红薯(V2)、香菜(V13)和水(W2),图 3 给出了其中的第 30 个波段原始影像。



图 3 高光谱影像(第 30 波段)

Fig. 3 Hyperspectral image (band 30)

随机选取 6 类地物总数目为 3870 的样本用于分类实验,如表 1 所示。在分类实验中使用了最小距离分类器。

分类实验中,为了比较分析,利用多重分形谱进行分类的同时,也利用了信息量维数和关联维数进

行分类,两种方法分别称之为 DIC 和 DCC,多重分形谱的权重因子 q_{\max} 限定在 $5 \sim 50$ 。

表 1 实验光谱样本个数

	Train samples	Test samples	Total
C4	400	800	1200
T6	140	280	420
T7	140	280	420
V2	100	250	350
V13	80	200	280
W2	400	800	1200
Total	1260	2610	3870

图 4 显示了 q_{\max} 为 3、5、10、15、20、25、30、35、40、45、50 的多重分形谱用于分类实验的分类精度比较。 q_{\max} 为 3、5、15、30 时的分类精度和 kappa 系数分别为 82.2% 和 0.777、94.7% 和 0.931、95.2% 和 0.938、93.8% 和 0.920; q_{\max} 取 15 之前,精度随 q_{\max} 值的增加而提高,达到 15 后,多重分形谱的维数的提高一定程度上降低了分类精度。为了特征提取速度,可以适当选择较小的 q_{\max} 值。利用关联维数和信息量维数进行分类的分类精度和 kappa 系数分别为 57.7% 和 0.483、80.1% 和 0.743,可见多重分形谱分类的精度要显著高于单一分形维数。

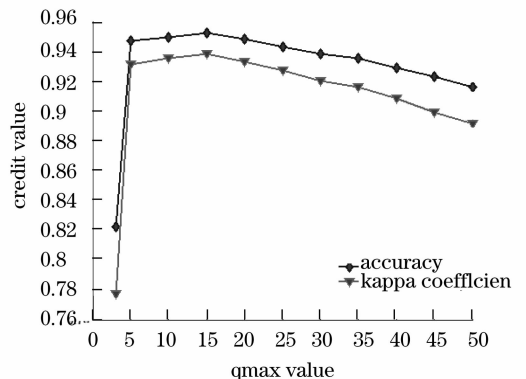


图 4 不同 q_{\max} 下的多重分形谱分类精度比较

Fig. 4 Comparison of classification accuracy with variable q_{\max} of multifractal spectrum

5 结 论

光谱特征提取是高光谱数据处理的主要任务之一。目前基于单一分形维数的分析方法无法有效地表征光谱包含的信息,本文验证了光谱曲线具有多重分形特征,并采用多重分形谱表征光谱特性。多

重分形谱根据光谱曲线内在的奇异性对光谱空间进行自动的划分,用多重分形谱表征具有同一奇异性指数的光谱子集的共同性质,通过在不同的奇异性范围内对光谱波动特征进行放大或压缩,实现了对光谱曲线的奇异性及分形维数的双重描述。高光谱数据分类实验结果表明基于多重分形谱的特征提取方法能够有效地应用于在高光谱数据处理中。光谱曲线的多重分形分析还处于探索阶段,对其多重分形谱的物理含义有待进一步研究。

参 考 文 献

- Zhu Hai, Jiang Lu, Liang Bo *et al.*. Retrieving Optical Concealment Depth of Underwater Target by Remote Sensing [J]. *Chin. J. Lasers*, 2007, **34**(5): 699~702
朱海,姜璐,梁波等.水下目标光学隐蔽深度遥感获取方法[J].中国激光,2007,34(5):699~702
- He Lin, Pan Quan, Di Wei *et al.*. A small-target detector based on single likelihood test for hyperspectral imagery [J]. *Acta Optica Sinica*, 2007, **27**(12): 2155~2162
贺霖,潘泉,邸韡等.一种基于单似然检验的高光谱图像小目标检测器[J].光学学报,2007,27(12):2155~2162
- Li Na, Zhao Huijie, Jia Guorui *et al.*. Anomaly detection based on extracted mathematical morphology for hyperspectral imagery [J]. *Acta Optica Sinica*, 2008, **28**(8): 1480~1484
李娜,赵慧洁,贾国瑞等.基于扩展数学形态学的高光谱图像异常检测[J].光学学报,2008,28(8):1480~1484
- Zhang Huaguo, Huang Weigen, Zhou Changbao. Fractal characterization of IKONOS imagery [J]. *Multispectral and Hyperspectral Remote Sensing Instruments and Applications, Proc of SPIE*, 2003, **4897**:292~301
- Du Huaqiang, Zhao Xianwen, Fan Wenyi. Fractal dimensions being an exponent of bands selection for hyper-spectral remote sensing data [J]. *Remote Sensing Technology and Application*, 2004, **19**(1):5~9
杜华强,赵宪文,范文义.分形维数作为高光谱遥感数据波段选择的一个指标[J].遥感技术与应用,2004,19(1):5~9
- Xiong Yuhong, Wen Zhiyu, Zhang Liuqiang *et al.*. Application of fractal theory to spectral signal recognition [J]. *Spectroscopy and Spectral Analysis*, 2006, **26**(4): 772~774
熊宇虹,温志渝,张流强等.分形理论在光谱识别中的应用[J].光谱学与光谱分析,2006,26(4):772~774
- Zhan Yunjun, Hu Guangdao, Yuan Yanbin. Hyperspectral RS image classification based on fractal and rough set [C]. *Second International Conference on Space Information Technology, Proc. of SPIE*, 2007, **6795**:67954F-1~67954F-6
- Xia Guifen, Zhao Baojun. Detection of air target based on multifractal analysis in a laser radar [J]. *Chin. Opt. Lett.*, 2007, **5**(1): 51~53
- Ping Qingwei, Xia Guifen. Laser radar target detection based on fractional brownian motion model [J]. *Chin. J. Lasers*, 2008, **35**(1): 106~110
平庆伟,夏桂芬.基于分数布朗运动模型的激光雷达目标检测[J].中国激光,2008,35(1):106~110
- Wang Gang, Xiao Liang, Jiang ZhiJing *et al.*. Sub-pixel multifractal method for image processing application [J]. *Acta Optica Sinica*, 2006, **26**(9): 1345~1349
王刚,肖亮,姜志静等.亚像素多重分形方法在图像处理中的应用[J].光学学报,2006,26(9):1345~1349
- Wang Gang, Xiao Liang, He Anzhi. An improved computing method for the image edge detection [J]. *Chin. Opt. Lett.*, 2007, **5**(2): 79~81
- Xia Yong, Zhao Rongchun, Feng D D. Multifractal estimation for remote sensing image segmentation [C]. *CSP 04, 2004 7th International Conference*, 2004, **1**(31): 773~776
- Zhang Jizhong. *Fractal Dimension* [M]. Beijing: Tsinghua University Press, 1995, 363~373
张济忠.分形[M].北京:清华大学出版社,1995.363~373
- Zhou Weixing, Wang Yanjie, Yu Zhunhong. Geometrical characteristics of singularity spectra of multifractals II. Partition function definition [J]. *Journal of East China University of Science and Technology*, 2000, **26**(4):390~395
周炜星,王延杰,于遵宏.多重分形奇异谱的几何特性II.配分函数法[J].华东理工大学学报,2000,26(4):390~395
- Mohamed Abadi, Enguerran Grandchamp. Legendre spectrum for texture classification [C]. *Signal Processing, The 8th International Conference*, 2006, **2**:16~20