

## 地形表面的自相似程度与分类感知

李旭涛 彭复员 曹汉强 朱光喜  
(华中科技大学电子与信息工程系 武汉 430074)

**摘要:** 地形的自相似特性在整个尺度空间上并非恒定的, 而与度量的尺度范围相关。该文通过将自相似特性曲线划分为若干个线性区间, 依据信息熵的定义提出了一个新的参数——自相似程度指数 SSD(Self-Similarity Degree) 来表征地形自相似的程度。进而采用距离度量和模糊  $c$  均值聚类分析, 表明 SSD 是地形识别的一个有效特征, 此特征的加入丰富了传统单一 Hurst 特征对分类对象的描述, 提高了地形分类感知的精度。

**关键词:** 分形; 自相似特性; 度量码尺; 模式分类

中图分类号: TP391.41

文献标识码: A

文章编号: 1009-5896(2007)06-1480-03

## Self-similarity Degree of Terrain Surface and Class Perception

Li Xu-tao Peng Fu-yuan Cao Han-qiang Zhu Guang-xi

(Dept. of Electronics and Information Engineering, Huazhong University of Science and Technology, Wuhan 430074, China)

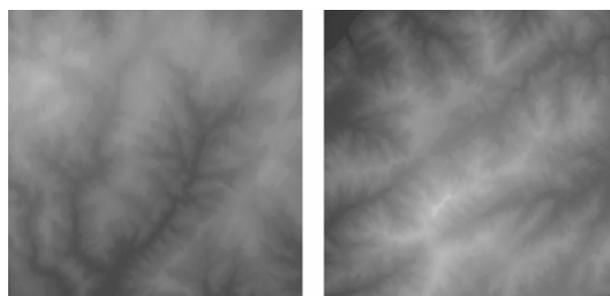
**Abstract:** Terrain surface is not always so perfect that it keeps invariable self-similar characteristic in whole scale space. To describe the degree of the self-similarity, the self-similarity curve is divided into several linear parts and then a new parameter called Self-Similarity Degree (SSD) is presented in the similitude of information entropy. Furthermore, the simulation in which distance measurement and fuzzy  $c$  mean clustering are adopted and it shows that the new parameter is the effective feature for terrain recognition. As the presented feature provides more information than traditional single Hurst feature, the precision of terrain classing is improved.

**Key words:** Fractal; Self-similarity; Measure scale; Model classing

### 1 引言

地形的分类与识别是地理环境感知中的重要课题, 而提取恰当的地形特征是实现正确分类的基础。地形表面是长期自然演化的产物, 具有非线性和自相似性的特点, 因此传统的基于统计的特征量如均值、方差等难以准确描述地形的本质特征。上世纪 80 年代初 Pentland 提出了自然表面粗糙度的分形描述<sup>[1]</sup>, 以 Hurst 指数  $H$  的大小表征纹理的复杂程度, 给出了一种描述自然表面的非线性本质特征的方法。我们知道, 自相似性亦即标度的不变性是分形体的重要特征<sup>[2]</sup>。若自然表面具有良好的自相似性, 其自相似性曲线应该在整个度量空间上保持良好的线性, 如图 1 所示的地形表面 1, 其自相似曲线见图 2(a)。然而, 实际的自然表面的标度往往仅在某些度量区间上保持不变, 这也就是说在整个度量空间上自然表面的自相似性实际上并不那么“完美”<sup>[3]</sup>, 如地形表面 2, 其自相似曲线见图 2(b)。因此, 实际表面的 Hurst 指数很少保持恒定, 其随度量尺度范围的不同而不同。关于如何度量信号自相似性的问题, 目前尚缺乏深入的研究。

以单一的 Hurst 指数为特征的分类往往过于简单, 如图 1 所示的两类地形, 由于其  $H$  值均接近于 0.56, 造成分类的



(a) 地形表面 1 (b) 地形表面 2

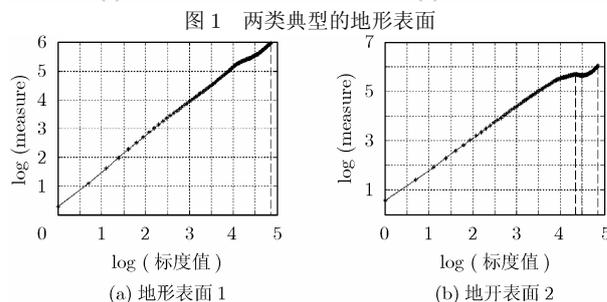


图 2 自相似特性曲线

困难。对于地形感知而言, 需要提取多尺度条件下的特征量, 这也符合人对物体感知由远而近, 由粗而细的过程。如果比较其多尺度的自相似特性, 则上述两类地形有着显著差异, 如图 2 所示, 其中地形表面 2 的大尺度 Hurst 指数明显异于小尺度 Hurst 指数。

本文基于分数布朗运动(FBM)模型<sup>[2,4]</sup>,研究了地形表面自相似特性,将地形表面的自相似指数看作依度量尺度范围而变化的变量  $H(s)$ 。并将自相似性曲线划分为若干个线性区间,依据信息熵的定义提出了自相似程度SSD(Self-Similarity Degree)指数,该指数的大小表征了信号自相似的程度。基于距离度量的仿真表明SSD值作为地形分类的特征,能够提供传统单一的Hurst指数所不能描述的信息,从而以SSD和 $H$ 为特征,能够提高地形感知分类的精度。

## 2 FBM 自相似模型

分数布朗运动(FBM)模型是描述统计自相似的重要模型。在大量自然现象的描述中FBM模型已经得到了有效的应用,FBM模型将自然表面的粗糙现象归结为随机行走的结果。Hurst指数为 $H$ ,初值  $B_H(0) = B_0$ ,  $t > 0$  的FBM定义为

$$\begin{aligned} B_H(t) - B_H(0) &= \frac{1}{\Gamma(H+1/2)} \int_{-\infty}^t k(t-t') dB(t') \\ &= \frac{1}{\Gamma(H+1/2)} \left\{ \int_{-\infty}^0 (|t-s|^{H-1/2} \right. \\ &\quad \left. - |s|^{H-1/2}) dB(s) + \int_0^t |t-s| dB(s) \right\} \quad (1) \end{aligned}$$

其中  $0 < H < 1$ , 分形维数  $D = E + 1 - H$ 。  $E$  为集合  $S$  的欧氏维数,对于二维地形表面,  $E = 2$ 。由定义可见,FBM 是一个非平稳过程<sup>[4]</sup>,其功率谱与频率  $\omega$  之间满足  $|\omega|^{2H+1}$  的关系。但其一阶差分却是平稳的,称为分数高斯噪声(FGN),其功率谱与频率  $\omega$  之间满足关系  $|\omega|^{1-2H}$ 。一般,一个分数布朗函数由零均值高斯增量的随机过程描述:

$$\text{var}(B_H(t_2) - B_H(t_1)) = C(t_2 - t_1)^{2H} \quad (2)$$

对于二维表面:

$$\langle |f(x,y) - f(x + \Delta x, y + \Delta y)|^2 \rangle = C |\Delta x^2 + \Delta y^2|^H \quad (3)$$

其中  $C$  为常数。

## 3 自相似程度的度量

为刻画地形表面的自相似性,本文采用 FBM 模型的矩估计法,选择二阶矩为测度  $\mu(\epsilon)$ ,  $\epsilon$  代表相应的度量尺度,尺度的范围为  $1 \leq \epsilon \leq L/2$ 。首先令  $\mu(\epsilon)$  为集合  $S$  的测度,  $\epsilon$  为度量的尺度,其变化范围  $[\epsilon_{\min}, \epsilon_{\max}]$ ,对于具有  $L \times L$  个数据点的二维地形表面,  $\epsilon_{\min} = 1$ ,  $\epsilon_{\max} = L/2$ ;然后计算尺度和测度的对数,在双对数坐标上绘制出自相似性曲线,如图 2。 $x$  轴代表对数度量尺度,  $y$  轴代表相应的对数方差测度。通过最小均方回归计算出自相似性曲线的斜率,此斜率代表了地形表面的自相似性指数。理想的自相似集的自相似性曲线在整个度量范围内呈一条直线,然而实际的表面通常在整个度量范围内很难保持良好的线性,而是呈现出分段线性的特点。因此可以令

$$p_i = s_i / S, \quad i = 1, 2, \dots, n \quad (4)$$

代表第  $i$  个线性区域  $s_i$  同整个度量空间  $S$  的比率,由信息熵的定义,相应地提出一个量称为自相似程度 SSD,

$$\text{SSD} = - \sum_{i=1}^n p_i \log p_i \quad (5)$$

来表征实际表面自相似的多变性。SSD 值越小,意味着存在一个占优的线性区域并且表面的自相似性越好;相反值越大,线性分区较多,其自相似性就越不确定,相应的地形表面的自相似性越差。对于如何划分自相似曲线的线性区间,提出如下的简便算法:

步骤 1 令  $c(x)$ ,  $x \in [1, L/2]$  代表自相似曲线,确定曲线上的极值点(极大值点和极小值点);

步骤 2 对于区间  $[\text{start}, \text{end}]$ , 区间起始点为  $\text{start}$ , 区间终点为  $\text{end}$ 。令初始区间为  $[1, 3]$ , 采用线性最小均方回归估计区间内数据点的线性特性。若估计误差大于设定的误差限(本文设为 0.03), 则选取区间的中间点作为一个新的分区起始点,同时区间终点向后递增 2 个点,重复步骤 2, 直到区间的终点为  $L/2$  时为止;若估计误差小于设定的误差限,则执行步骤 3;

步骤 3 区间起始点  $\text{start}$  保持不变, 区间终点  $\text{end}$  向后递增 2 点, 然后执行步骤 2, 直到区间的终点为  $L/2$  时为止;

步骤 4 以计算出的一系列区间起点、终点和曲线的极值点共同确定曲线的线性区域的划分。

应用上述方法,计算图 1 所示的两地形,线性分区如图 2 所示。由式(5)计算得到的两类地形的自相似程度指数,地形表面 1 为  $\text{SSD}=0.0112$ , 地形表面 2 为  $\text{SSD}=1.2891$ 。地形是复杂的物理过程长期演化的结果,其形成的因素十分复杂,呈现出显著的非线性特点。地形的形成因素越单一,则其自相似性越好,自相似程度指数相应较小,如地形表面 1 主要是由于雨水和风的侵蚀形成的;反之自相似程度指数较大,如地形表面 2 不仅同雨水和风的侵蚀有关,而且还存在大尺度的板块的运动挤压。其大尺度条件下的自相似性异于小尺度条件下的自相似性( $H$  值分别为 0.5565 和 0.5915)表明大小尺度条件下影响地形形成的因素不同,而  $H < 0$  的情况是由于二维 FBM 的各向同性造成的,说明在这个尺度范围内地形表面呈现出有向的纹理。因此,自相似程度指数表征了信号多尺度条件下自相似性的变化,与分形维数(或 Hurst 指数)一样是地形非线性本质的反映,能够用作地形分类的特征。

## 4 特征的距离度量与分类仿真

对于模式分类,具有小的平均类内距离  $S_b$  和大的类间距离  $S_w$  的特征才具有好的可分性<sup>[5,6]</sup>。对于  $c$  个模式类  $\omega_1, \omega_2, \dots, \omega_c$  的特征矢量集为  $\{X^{(i,k)}, i = 1, 2, \dots, c; k = 1, 2, \dots, N_i\}$ , 其中  $X^{(i,k)} = (x_1^{(i,k)}, x_2^{(i,k)}, \dots, x_m^{(i,k)})^t$  为  $\omega_i$  中第  $k$  个  $m$  维特征矢量,  $N_i$  为  $\omega_i$  类中的特征矢量的数目。因此距离准则

定义为

$$J_A = \frac{S_b}{S_w} = \frac{\frac{1}{c} \sum_{i=1}^c \sum_{l=1}^m |\mu_l^{(i)} - \mu_l|^2}{\frac{1}{c} \sum_{i=1}^c \sum_{l=1}^m \frac{1}{N_i - 1} \sum_{k=1}^{N_i} |x_l^{(i,k)} - \mu_l^{(i)}|^2} \quad (6)$$

其中  $\mu_l^{(i)} = \frac{1}{N_i} \sum_{k=1}^{N_i} x_l^{(i,k)}$ ,  $\mu_l = \frac{1}{c} \sum_{i=1}^c \frac{1}{N_i} \sum_{k=1}^{N_i} x_l^{(i,k)}$ 。若  $J_A$  值大则表明特征的可分性好。

为衡量自相似程度指数 SSD 作为分类特征的性能,采用距离准则对采用传统单一的 Hurst 指数特征和采用 SSD 特征进行比较。选取了 50 幅大小为  $256 \times 256$  的地形作为训练样本,采用模糊  $c$  均值聚类算法,在取不同的聚类数目( $c=2, 3, 4, 5, 6$ )时,计算特征 SSD 的  $J_A$  值,并与单一 Hurst 指数为特征的  $J_A$  值比较,结果如图 3 所示。在不同聚类数目时,特征 SSD 的  $J_A$  值除  $c=5$  外基本上高于单一 Hurst 指数为特征的  $J_A$  值,而且分类数越大其  $J_A$  值越高,表明特征 SSD 对样本集具有较强的分类能力,因此 SSD 适合于用作地形分类的特征。

进一步采用单一 Hurst 指数  $H$  与 SSD 特征共同组成特征矢量 ( $H, SSD$ ) 用于地形的分类。取不同的聚类数目( $c=2, 3, 4, 5, 6$ )时其分类精度(正确分类的百分比)与以单一  $H$  为特征的分类精度的比较如图 4 所示。可见新的特征矢量在取不同分类数时,具有高于采用单一 Hurst 特征的分类精度。

## 5 结束语

地形表面是一类复杂的自然表面,本文提出了描述实际

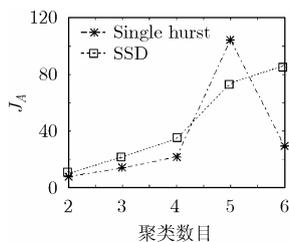


图 3  $J_A$  值的比较

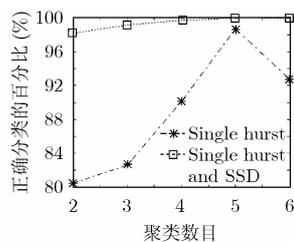


图 4 分类精度

地形自相似程度的方法,并提出了称为自相似性程度 SSD 的指数来表征自相似的程度。验证了以 SSD 为特征的地形分类感知的合理性,与传统的采取单一  $H$  指数为特征的分类相比,以 SSD 和  $H$  组成的新的分类特征具有较高的分类精度。通过本文的研究,表明了实际表面的 Hurst 指数  $H$  在整个度量空间上不再是一个常数,而是一个与空间相关的量  $H(s)$ 。这为进一步研究基于各向异性多尺度自相似模型的地形重建方法建立了基础。

## 参考文献

- [1] Pentland A. Fractal-based description of natural scenes. *IEEE Trans. on Pattern Anal. Machine Intell.*, 1984, 6(12): 666-674.
- [2] Falconer K J. *Fractal Geometry: Mathematical Foundations and Applications*. England: John Wiley and Sons, 1990, Chapter 16.
- [3] Li Xu-tao, Cao Han-qiang, and Zhu Guang-xi, *et al.* Complicated self-similarity of terrain surface. The Fourth International Symposium on Multispectral Image Processing and Pattern Recognition 2005, Wuhan China, Oct. 2005, Proc. of SPIE Vol.6044: 1G-1 - 1G-11.
- [4] Voss R F. Random Fractal Forgeries. in *Fundamental Algorithms for Computer Graphics*, R.A.Earnshaw, Ed. New York: Springer-Verlag, 1985: 805-836.
- [5] 王首勇,朱光喜.应用复双谱对角切片的雷达杂波多目标特征提取. *信号处理*, 2002, 18 (6): 500-504.
- [6] 孙即祥. *现代模式识别*. 长沙:国防科技大学出版社, 2002: 14-25.

李旭涛: 男, 1971 年生, 博士生, 研究方向为现代信号处理与模式识别.

彭复员: 女, 1945 年生, 教授, 博士生导师, 主要研究方向为模式识别及图像处理.

曹汉强: 男, 1953 年生, 教授, 博士生导师, 主要研究方向为图像处理及信息安全.

朱光喜: 男, 1945 年生, 教授, 博士生导师, 主要研究方向为无线通信系统.