

测量分维的矢量计盒算法研究

韩 杰 陆桂华

(河海大学水文水资源和水利工程科学国家重点实验室, 南京 210098)

摘 要 分形理论目前在众多学科领域中取得了广泛的应用, 研究这些对象主要通过确定它们的分维, 其中计盒算法是一种最常用的方法。传统的计盒算法是基于栅格(位图)文件的方法, 由于其存在图像放大后失真、过程繁琐和迭代次数有限等缺陷, 因此为了准确简单方便地进行分维计算, 开发了一种以矢量文件为载体的矢量计盒算法, 并详细阐述了这种算法的数据结构、处理流程和主要函数, 同时以 Koch 曲线、骨肿瘤边界及水系证明了矢量计盒算法的准确性和优越性。实践表明, 该算法有 3 个优点: (1) 图像不会随着放大缩小而失真; (2) 可完全进行计算机操作, 且简单可靠; (3) 从某种意义上讲, 由于迭代的次数可以是无穷的, 所以能非常精确地确定图形的标度空间, 并可获得精确的分维值。因此矢量计盒算法是一种方便、实用而且精确的维数计算方法。

关键词 分形 计盒算法 栅格 矢量 Koch 曲线 骨肿瘤边界 水系

中图法分类号: TP391.4 文献标识码: A 文章编号: 1006-8961(2008)03-0525-06

Research on the Vector Box-counting Algorithm in Fractal Dimension Measurement

HAN Jie, LU Gui-hua

(State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, HoHai University, NanJing 210098)

Abstract The fractal theory developed by the French scientist Mandelbort in 1970s is beneficial in many areas. It greatly expands and deepens our knowledge on irregular geometric bodies. Fractal theory quantifies these phenomena mainly by ascertaining their fractal dimensions. Box-counting algorithm is the one most practical and also most frequently adopted method. The traditional box-counting algorithm is based on the grid document and has some serious shortcomings, such as the distortion of the image being enlarged, the trivialness of the process and the finite of the iterative degree, etc. The vector box counting algorithm developed in this paper takes vector document as the carrier and has three advantages. First, the image will not be distorted after being enlarged. Second, the process is completely handled by computer, simple and reliable. Third, to some degree, the iterative degree can be infinite. Therefore, it can ascertain precisely the scaling space of the graph and acquire accurate fractional dimension value. This paper expounds the data structures, the process of disposing and the main functions in detail. What's more, it proves the precision and advantages of the vector box counting algorithm by making use of Koch curve, osteoma boundary and river system. The result shows that the vector box counting algorithm is a convenient, useful and precise way of dimensional calculating method.

Keywords fractal, box-counting algorithm, grid, vector, Koch curve, osteoma boundary, channel systems

1 引言

分形理论由法裔美国科学家曼得尔布罗特(Mandelbort)于 20 世纪 70 年代创立的, 它的产生, 不仅加深

了人们对自然界不规则几何体的认识, 而且可使人们能够借助于事物内部的自相似性(self similarity)来洞察隐藏在混乱现象中的精细结构, 从而为人们从局部认识整体, 从有限认识无限提供了全新的思路和描述

工具^[1]。通常人们在研究这些现象时,发现它们大部分是无序的、不稳定的和随机的,并不存在任何线性关系,而分形则可以直接从纷繁复杂的系统本身入手,揭示非线性系统中的有序和无序、确定性和随机性的统一^[2]。分形理论目前已经在包括地学、医学、物理学等众多学科领域取得了广泛的应用。

利用分形理论来定量描述这些现象主要是通过确定它的分维。目前,关于分维的测量有很多种方法,例如圆规维数、计盒算法、Sandbox 法、面积-回转半径法、变换(variation)法、密度相关函数法等等,其中计盒算法是诸多方法最为常用与实用的方法之一,它能非常精确地确定不规则曲线的分维,比如海岸线、水系、肿瘤的边界等。计盒算法可以表述为:首先将 2 维空间划分成边长为 ξ 的盒子,同时数出整个曲线所占据的盒子数 $N(\xi)$;然后依次改变盒子边长为 ξ 的 $1/2, 1/4, 1/8, 1/16, \dots$,再数出整个曲线所占据的盒子数 $N(\xi)$;接着将一系列的盒子数 $N(\xi)$ 和边长 ξ 的数据取对数;最后绘制 $\ln N(\xi) \sim \ln \xi$ 图,如果能得到一条直线,则说明盒子数 N 和边长 ξ 有以下关系: $N(\xi) \sim (1/\xi)^D$,直线的斜率 D 即是图形的分维^[3,4]。

2 栅格计盒算法介绍与主要缺陷

目前计盒算法的计算机实现普遍是通过对栅格图像文件进行以下处理:(1)图片预处理,即先利用扫描仪将图片扫描为某种格式的位图文件(*.BMP、*.GIF等),然后进行增加对比度、清除不需要的部分等工作;(2)记录,即先选取分维计算的图像范围,然后进行适当的细化处理,用非可视化网格在范围内逐步改变网络尺寸(以像素点为单位)或采用分割算法来获得二值图像,并统计各尺寸下有图像充填的盒子数;(3)分析计算,即对统计结果进行线性分析,斜率即分维值。

上述位图图像文件是由多个像素组成的,并带有每个像素颜色的数据。它的一个缺点是当图像放大时,由于每个像素的信息量是固定的,所以在超过一定范围后再看位图,就会出现马赛克现象。这种马赛克现象必然在水系维数计算中,随着迭代次数的增加而越来越明显,其首先造成的后果是对水系维数计算的精度产生一定的影响;其次,上述计算过程随着迭代次数的增加会显得越来越繁琐。另外一个很重要的原因是,由于实际的分形和数学上理想的分形不同,它只存在于有限的范围内,这就更加需要在一定

的情况下增加迭代的次数,以便能够得到充分的分析,而这些问题都是栅格计盒算法难以解决的。

3 矢量计盒算法的计算机处理

矢量图形存储的是一些描述信息,其所描述的都是关于曲线的点、颜色、曲率等信息,而不是特定像素的信息。它的描述方式是:在一个填充区域和一片颜色相似的区域,它是通过一种数学方式建立一个固定的函数关系来描述图像的。而这些矢量文件在计算机技术高速发展的今天已经不难得到,比如水系、海岸线等等,而且这种矢量计盒算法有以下 3 个优点:(1)图像不会随着放大缩小而失真;(2)可完全用计算机操作,简单可靠;(3)由于从某种意义上讲,迭代的次数可以是无穷的,所以能非常精确地确定图形的标度空间,并能获得精确的分维值^[5]。

本方法可以用计算机高级语言来实现,比如 Basic, C++ 等等,一般可在 VB (visual basic)、VC (Visual C++) 等软件中实现。下面重点介绍一下本方法的实现过程:

矢量计盒算法的整体思路是根据设定的迭代次数,对覆盖图形的矩形不断细分,并逐个判断细分所得的矩形是否被图形所覆盖。数据准备阶段分下面 3 个步骤:

- (1) 取矢量图形文件,如 .ai, .cdr 矢量格式的文件;
- (2) 读取的数据存入结构体 *pathgroup* 中,该结构体用于描述一系列曲线的集合,即一系列首尾相接的三次 Bezier 曲线的集合;
- (3) 计算出整个图形的轮廓矩形,存储在矩形列表中。

数据处理过程(如图 1 所示)如下:

- (1) 取矩形列表 b_i 中的第 1 个作为当前矩形;
- (2) 调用函数,判断当前矩形是否被图形 *pathgroup* 所覆盖;
- ① 取图形 *pathgroup* 的第 1 条曲线作为当前曲线;
- ② 判断当前曲线是否覆盖当前矩形;
- ③ 如果当前曲线覆盖当前矩形,则认为图形覆盖当前矩形;否则,取下一条曲线作为当前曲线,重复以上①、②、③步骤,直到有曲线覆盖矩形或所有曲线均不覆盖矩形;
- (3) 如果当前矩形被图形所覆盖,则将当前迭

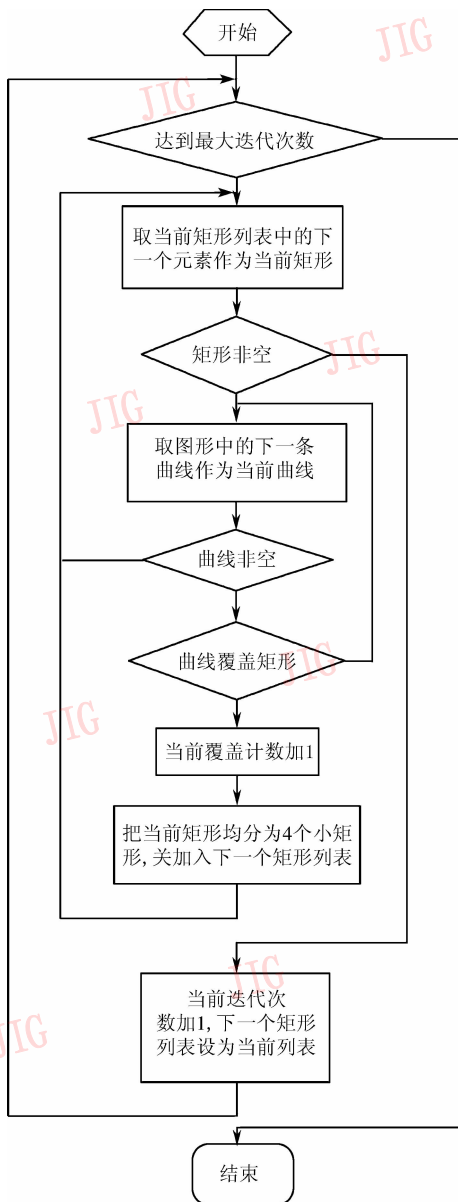


图 1 矢量计盒算法流程
Fig. 1 The flow chart of vector box-counting algorithm

代被覆盖矩形数 $boxarray$ 加 1, 同时把当前矩形分割为 4 个相同的小矩形, 并将它们加入到下一迭代的矩形列表 $pBoxListNext$ 中, 则下一次迭代过程, 图形必然至少覆盖这 4 个小矩形中的一个;

(4) 如果当前矩形不被图形所覆盖, 则随后迭代过程中, 由该矩形细分所得到的所有小矩形均不被图形覆盖, 故不再将该矩形细分, 随后的迭代过程也不再考虑由该矩形所派生的子矩形, 以减少计算量;

(5) 取矩形列表 b_i 中的下一个矩形作为新的当前矩形;

(6) 重复步骤(2) ~ (5), 至矩形列表 b_i 中的所有矩形均被处理;

(7) 当前迭代完成, 当前迭代的被覆盖矩形数为 $boxarray$;

(8) 迭代次数 k 增加 1, 准备下一次迭代;

(9) 把新的迭代次数 $k + 1$ 和新的矩形列表 $pBoxListNext$ 作为参数, 重复调用上述函数, 计算各次迭代所覆盖的矩形数, 直至达到最大迭代次数。

矢量计盒算法的关键是如何判断矢量图形是否与当前矩形相交。由于矢量图形是一系列曲线的集合, 这些曲线由一系列首尾相接的 3 次 Bezier 曲线组成^[6-8], 因此, 若要判断矢量图形是否覆盖矩形, 只要将其转化为判断组成矢量图形的 Bezier 曲线组是否覆盖当前矩形即可。为了判断 Bezier 曲线是否覆盖当前矩形, 可以用高精度的多边形去逼近 Bezier 曲线, 以便把问题转化为多边形是否覆盖矩形, 然后进一步转化为构成多边形的线段是否覆盖矩形; 用此方法判断线段是否覆盖当前矩形, 非常简单快捷。

三次 Bezier 曲线的描述函数如下:

$$\begin{aligned}
 Q(t) &= \sum_{i=0}^3 P_i B_{i,3}(t) \\
 &= (1-t)^3 P_0 + 3t(1-t)^2 P_1 + \\
 &\quad 3t^2(1-t) P_2 + t^3 P_3 \\
 &= B_{0,3}(t) P_0 + B_{1,3}(t) P_1 + B_{2,3}(t) P_2 + \\
 &\quad B_{3,3}(t) P_3 \quad t \in [0, 1]
 \end{aligned}$$

其中, P_0, P_1, P_2, P_3 是 4 个控制点。本文把 Bezier 曲线在 $t=0.5$ 的地方分开为两段 Bezier 曲线, 再以同样的方法对生成的 Bezier 曲线进行分割, 这样递归分割, 直到 Bezier 曲线上的分割点足够密集; 然后把把这些分割点依次以线段相连, 这样就得到一个足够近似于原 Bezier 曲线的多边形。其分割过程描述如下:

(1) Bezier 曲线的 4 个控制点分别为 P_0, P_1, P_2, P_3 , 在 P_0P_1, P_1P_2, P_2P_3 直线上分别取中点 Q_0, Q_1, Q_2 ;

(2) 取 Q_0Q_1, Q_1Q_2 的中点 S_0, S_1 ;

(3) 当取线段 S_0S_1 的中点 F_0 , 该点即为 Bezier 曲线上的点 ($t=0.5$), 则原 Bezier 曲线就被分为两段 Bezier 曲线: $P_0Q_0S_0F_0$ 和 $F_0S_1Q_2P_3$;

(4) 重复步骤(1) ~ (3), 直到任一 Bezier 曲线的两个端点的距离小于某阈值 A ;

(5) 把上面分割所得到的一系列 Bezier 曲线的端点依次以线段相连, 就形成一个高度逼近原 Bezier 曲线的多边形, 然后即可通过线段判断是否覆盖当前矩形。

4 实例分析

现在以几个实例来检验一下矢量计盒算法的效果,计算过程中只需要读入矢量文件即能在 VC++ 中进行计算。

4.1 Koch 曲线

Koch 曲线是从一条线段开始,将线段中间的 1/3 部分去掉而代之以一个等边三角形的另外两边,首先形成山丘形图形,然后在新的图形中,又将图中每一线段中间的 1/3 部分都用一个等边三角形的两

条边代替,再次形成新的图形,如此迭代即形成 Koch 分形曲线,它是一个严格意义的分形曲线,图 2 (a) 是 Koch 曲线经过 4 次迭代产生的图形,它的精确维数是 $\ln 4 / \ln 3$, 约为 1.26^[1]。通过矢量计盒算法计算得到的 Koch 曲线的维数值为 1.262 2 (见图 2(b)), 并且相关性系数 R^2 达到 0.998 1, 从而证明了矢量计盒算法完全能够达到维数估计的精度要求。设想如果采用栅格计盒算法,当 Koch 曲线经过更多次迭代后,必然产生马赛克现象,那么将对维数的估计产生很大的影响。

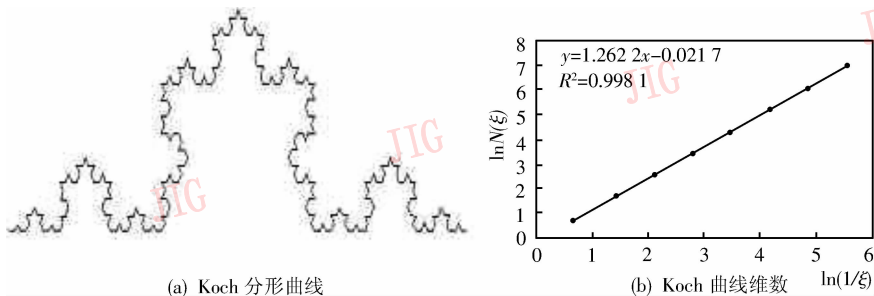


图 2 Koch 曲线及矢量计盒算法维数计算结果
Fig. 2 Koch curve and the fractal dimension by vector box-counting algorithm

4.2 骨肿瘤边界

大家知道,骨肿瘤边界的分形维数有助于诊断骨肿瘤的病变程度,即分形维数值越大,肿瘤边缘组织浸润性生长能力越强,边缘越不规则,恶性程度越高,预后越差;反之,分形维数越小,恶性程度越低,预后越好^[9]。图 3(a) 是良性骨肿瘤边界,图 3(b) 是恶性骨肿瘤边界,图 3(c) 上图是良性骨肿瘤边界的矢量

计盒算法的结果,下图是恶性骨肿瘤边界的矢量计盒算法计算的分形维数,恶性肿瘤分形维数值为 1.040 9,良性肿瘤分形维数值为 1.014 3。从维数计算结果和相关性上可见,矢量计盒算法能很好地反映骨肿瘤边缘组织的浸润性生长状态,因此能很好地进行骨肿瘤的识别分类。分维计算结果表明,矢量计盒算法简洁方便,完全可以得到精确的分维值。

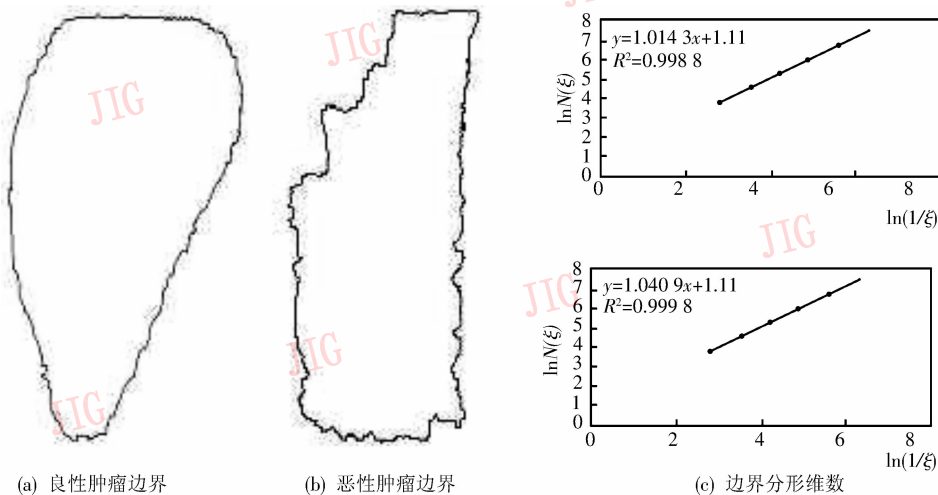


图 3 骨肿瘤边界及矢量计盒算法维数计算结果
Fig. 3 Osteoma boundary and the fractal dimension by vector box-counting algorithm

4.3 水系

众所周知,水系是由地表水的侵蚀、搬运和堆积作用形成的地貌景观,也是所在地区内外地质营力作用的产物。分形理论用于对水系进行分析具有很大的优势,因其能够对水系特征进行定量描述,比传统的用水系切割密度描述具有无比的细致性^[10-12]。为了更好地

说明不同地区水系的分维值,选取了云南省石屏县与盈江县 1:50 000 水系分布图(如图(4)所示)进行分维研究,其中石屏县水系稀疏,而盈江县则水系稠密,若选用栅格计盒算法进行计算,则相当繁琐,而用矢量计盒算法则解决了这一难题,它不仅精确地确定图形的标度空间,并能获得精确的分维值。

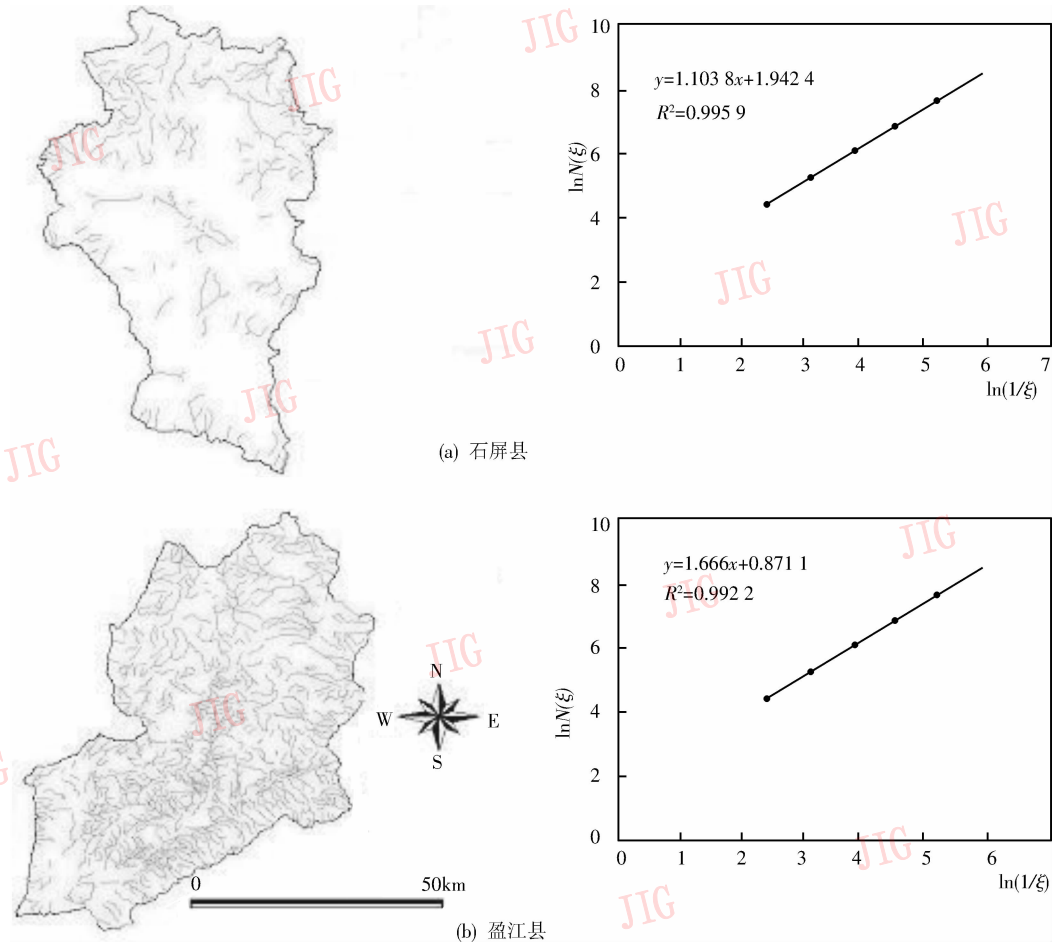


图4 水系图(左)及矢量计盒算法维数计算结果(右)

Fig. 4 Channel system and the fractal dimension by vector box-counting algorithm

5 结论

本文在讨论栅格计盒算法缺陷的基础上,开发了一种基于矢量的计盒算法,并用 Koch 曲线、骨肿瘤边界及水系来证明了矢量计盒算法的准确性和优越性。实例分析结果表明,矢量计盒算法是一种方便、实用而且精确的维数计算方法。

参考文献 (References)

1 Li Hou-qiang. Fractal and Fractional Dimension [M]. Chengdu:

Sichuan. Education Press, 1990. [李后强. 分形与分维[M]. 成都:四川教育出版社,1990.]

2 Qin Yao-chen, Liu Kai. Advancement of applied studies of fractal theory in geography [J]. Progress in Geography, 2003, 22(4): 426~436. [秦耀辰,刘凯. 分形理论在地理学中的应用研究进展[J]. 地理科学进展,2003,22(4):426~436.]

3 Chen Jiang-feng, Liu Yong, Zhang Guo-wang. Computer processing and fractal dimension calculation of natural water system [J]. Coal Geology and Exploration, 2001, 29(4):45~47. [陈江峰,刘勇,张国旺. 自然水系的计算机处理与分维计算[J]. 煤田地质与勘探, 2001, 29(4):45~47.]

4 Liang Dong-fang, Li Yu-liang, Jiang Chun-bo. Research on the box counting algorithm in fractal dimension measurement [J]. Journal of

- Image and Graphics, 2002, 7(3):246 ~ 250. [梁东方, 李玉良, 江春波. 测量分维的“数盒子”算法研究[J]. 中国图象图形学报, 2002, 7(3):246 ~ 250.]
- 5 Chen Guang-mao. Fractal Theory and Application[M]. Chengdu: Sichuan Education Press, 1989. [程光钺. 分形理论及其应用[M]. 成都: 四川大学出版社, 1989.]
 - 6 Wang Bing-tuan, Zhang Zhi-gang, Wang Jun-tuan. Best approximation of Bezier surface[J]. Beijing Journal of University of Science and Technology 1996, 18(5):491 ~ 494. [王兵团, 张志刚, 王军团等. Bezier曲面的最佳逼近[J]. 北京科技大学学报, 1996, 18(5):491 ~ 494.]
 - 7 Wang Xiao-dong. Bezier curve tree[J]. Journal of Computer Aided Design & Computer Graphics, 1994, 6(4):266 ~ 270. [王晓东. Bezier曲线树[J]. 计算机辅助设计与图形学学报, 1994, 6(4):266 ~ 270.]
 - 8 Shen Hai-ou, Chen Shu-zhen, Sun Xiao-an. Quadratic rational bezier curve fitting[J]. Journal of Wuhan University, 1997, 43(1):119 ~ 123. [沈海鸥, 陈淑珍, 孙晓安. 曲线的二次有理 Bezier 曲线拟合[J]. 武汉大学学报, 1997, 43(1):119 ~ 123.]
 - 9 Chen Shao-qing, Wu Zhao-xia, Cheng Jing-zhi. Fractal characteristics of osteoma Boundary[J]. Journal of Image and Graphics, 2000, 6(9):790 ~ 793. [陈少卿, 吴朝霞, 程敬之. 骨肿瘤边界分形特性的研究[J]. 中国图象图形学报, 2000, 6(9):790 ~ 793.]
 - 10 Hong Shi-zhong, Hong Shi-ming. Study on fractal dimension in geomony area[J]. Exploration of Nature, 1988, 24(2):33 ~ 40. [洪时中, 洪时明. 地质领域中的分维研究[J]. 大自然探索, 1988, 24(2):33 ~ 40.]
 - 11 Liu De-ping. The application of fractal theory in hydrologic process shape characteristic analysis. [J]. Journal of Hydraulic Engineering, 1998, 2(2):20 ~ 25. [刘德平. 分形理论在水文过程形态特征分析中的应用[J]. 水利学报, 1998, 2(2):20 ~ 25.]
 - 12 Feng Ping, Feng Yan. Calculation on fractal Dimension of River Morphology[J]. Acta Geographica Sinica, 1997, 52(4):324 ~ 329. [冯平, 冯焱. 河流形态特征的分维计算方法[J]. 地理学报, 1997, 52(4):324 ~ 329.]