DOI: 10.13657/j.cnki.gxkxyxb.1998.03.004

第 14卷第 3期 广西科学院学报 Vol. 14, No. 3
1998年 8月 Journal of Guangxi Academy of Sciences August 1998

分形及其在图像压缩编码中的应用

Fractal and It's Application in Image Compression Coding

曾 毅

Zeng Yi

(南京邮电学院 南京 210003)

(Nanjing College of Posts and Telecommunications, Nanjing, 210003)

摘要 简要介绍了分形、分数维数的概念,分形用于图像压缩的基本原理,以及基于分形的若干图像压缩编码方法。

关键词 分形 迭代函数系统 局部迭代函数系统 压缩仿射变换 人眼视觉系统

中图法分类号 TN 911.7

Abstract The conceptions of fractal and fractal demension, the mathematical foundations for fractal image compression technique, and some kinds of image compression coding methods based on fractal geometry are introduced.

Key words fractal, iterated function system, partitioned iterated function system, compression affine transformation, human visual system

1 分形及其量度

自然界中的所有形状和人类迄今所考虑的一切图形, 大致可分为两种: 一种是具有特征长度的图形, 比如球、自行车、正方形等等; 另一种则是不具备特征长度的图形, 例如云彩、雪花、海岸线等等。后一类图形的特征是局部与整体之间存在着一定的相似性, 我们称之为自相似性。Benoit B. Mandelbrot将这种具有自相似性的图形称为分形。分形具有如下特征[1]: (1) 具有精细结构, 即有任意小比例的细节; (2) 不规则性, 不能用传统的几何语言描述其整体或局部; (3) 通常有某种自相似的形式, 可能是近似的或是统计的; (4) 其分形维数(以某种方式定义的) 一般大于其拓扑维数; (5) F可以以非常简单的方法来定义, 可能由迭代产生。

分形物体的维数可由分数维数来描述,有各种分数维数的定义方法,如豪斯道夫维数、信息维数、容量维数等, 一般常用的是 Hausdorff (豪斯道夫) 维数, 其定义为[2]:

假定 $D>0$,用直径小于 $X>0$的可数个数的球覆盖集合 E 若假定 $d_1, d_2, \cdots, d_k$为各球的直径,则 D维 Haustorff测试可表示为:

1998-02-26收稿

$$M_D(E) = \lim_{\varepsilon \to 0} \inf_{d_k} \sum_k (d_k)^D ,$$

此量从 0向无限大迁移时，称 D为集合 E的 Hausdorff维数，且对于任意给定图形，Hausdorff维数是唯一的。

其它分数维数的测量方法见文献 [2]。

自 1975年 Mandelbrot提出分形概念以来，它在自然科学和社会科学的许多领域得到广泛应用。在图像学方面，自 80年代中后期 Barnsley时提出分形图像压缩的概念以来，分形图像压缩编码以其高压缩比而倍受人瞩目，分形图像编码的物理依据是自然界具有统计意义上的自相似结构 (分形特征)，其数学基础是迭代函数系统和拼贴定理

2　分形图像编码的理论基础

2.1　迭代函数系统 (IFS)

设 D为 n维欧氏空间 R^n上的闭子集，若存在实常数 λ，$0 < \lambda < 1$，使得

$$d(w(x), w(y)) \leqslant \lambda d(x,y) \qquad x, y \in D \tag{1}$$

其中 $d(x,y)$为欧氏距离

则称映射 $W: D \to D$为 D上的压缩映射，若 (1) 式取 = 号，则称为相似映射。若 $\{w_i | i = 1, 2, \cdots, m\}$是压缩映射组，则称 $\{w_i: D \to D\}$为压缩映射集，或称迭代函数系统 (iterated function system)。

若 $W = \{w_i | i = 1, 2, \cdots, m\}$是 $D \subset R^n$上的压缩映射集，即

$$W_i(D) = \{w_i(x), x \in D\}, W(D) = \bigcup_{i=1}^{m} w_i(D),$$

则存在唯一的吸引集 A为 W的迭代分形集，且 A是 W的不变集，即

$$A = W(A) \approx \bigcup_{i=1}^{m} w_i(A)$$

这可以从两方面来解释：一方面，任何集合 F经过压缩变换 W的反复作用都会最终收敛至 A；另一方面，不论从 F中哪一点 X_0出发，只要依相同顺序进行 $W_{i_k}(\cdots W_{i_1}(x_0))$ $(i_1, \cdots, i_k, \cdots$为任意序列，且 $1 \leqslant i_k \leqslant m)$的迭代运算，最终都会收敛至 A中同一点 $A(x_0)$，也就是说，吸引集 A具有自相似结构和吸引性，若 w_i为随机函数时，A具有统计自相似结构。同时 A与 $W_i(A)$之间具有自相似或自仿射性质。

2.2　拼贴定理

设 $W = \{w_i | i = 1, 2, \cdots, m\}$是 R^n上的 IFS，压缩比为 λ，$0 < \lambda < 1$，F是 W的不变集，对于 R^n上任意非空紧致集 E有：

$$d(E, F) = d(E, \bigcup_{i=1}^{m} W_i(E)) / (1 - \lambda)$$

其中 $d(.,.)$是集合的 Hausdorff距离。

拼贴定理的重要推论是 R^n上任意非空紧致集均可用自相似集来逼近，即：

对任意给定的误差 $\varepsilon > 0$和任意非空紧致集 $E \subset R^n$，存在 $m(\varepsilon) < \infty$及压缩映射集 $W = \{w_i | i = 1, 2, \cdots, m\}$，$F$为此压缩映射集的不变集，使 $d(E, F) < \varepsilon$

拼贴定理保证了对任一图像 E，在误差失真 ε水平上，总能找到 F与其相似，且 F是迭代函数系统的不变集。因此只要找到适当的 $\{w_i\}$，便可丢弃原图像，在收端用收到的 $\{w_i\}$决定唯一的……

3 几种分形图像编码方法

由前面可看出，分形图像编码利用了自然界图像普遍存在的自相似结构，并用压缩映射关系来表示这种自相似结构，当这些映射关系满足一定条件时，可通过迭代重建图像。因此，分形编码过程实际上就是寻找映射关系的过程，而解码过程则是迭代过程。

1988年， M. Barnsley提出了 IFS方法。1990年， A. Jacquin提出了基于 IFS的分块编码方法。同时， J. Jang和 S. Rajala提出基于分形和人眼视觉系统的分割编码，下面简要介绍这三种方法。

3.1 IFS法

所谓 IFS法，就是根据前述的压缩原理为每幅图像寻求其 IFS代码。寻求 IFS代码通常有两种方法：一种是利用图像存在的自相似性直接求出 IFS代码；另一种方法是将图像分成若干小物体并对每个物体在 IFS库中寻找其代码。

3.1.1 直接寻求法

分形图像具有自相似性，我们可将图像分割成几个子图，每个子图都与原图具有一定的分形结构，对每个子图 X 利用人机交互方式经过压缩仿射变换 W_i(包括平移，旋转，伸缩与仿射)，使 $W_i(x)$ 能覆盖 X 的一部分，同时还得到一个重生信息度量 $P_i = W_i(x)$面积 $/X$面积，当足够多的 $W_i(x)$ 能完全覆盖 X 时，$\{W_i, P_i | i = 1,\cdots,N\}$ 称为分割部分 X 的 IFS码。过程要求 $W_i(x)$ 的重叠尽可能小，且 N 为最小。

一般采用二维空间 R 上的压缩仿射变换 W_i，可表示为：

$$W_i\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a_i & b_i \\ c_i & d_i \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e_i \\ f_i \end{pmatrix}$$

，$\{a_i, b_i, c_i, d_i, e_i, f_i\}$ 即为 W_i 的仿射参数集。

3.1.2 IFS库查寻法

有些图像从表面很难看出分形特征，此时可采用一些图像处理技术，如色彩分离、边缘检测、频谱分析等办法将图像分为几块或分为一些小物体。一块中可能包括一片叶子，一朵云，或是它们的简单组合，这类典型物体的 IFS码可在 IFS库中查到，IFS库中包含的并不是大量的分形图形，而是这些分形的 IFS代码。当图像中所要查寻的物体是 IFS库中分形的压缩，旋转或平移时，我们可将 IFS库中分形与原图中分形进行匹配从而得到原图中分形的 IFS代码。但这种方法存在的问题是：IFS库中代码也需要通过 (1) 法来求出，且如何才能自动精确地将图像分割成合理的块结构。

3.1.3 IFS法的特点

(1) 引入了全局和局部相关去除信息冗余的思想，为实现高压缩比提供了可能；

(2) 引入了仿射变换 (affine transformation)，扩大了相似性范围；

(3) 仿射变换必须是压缩的，否则不存在吸引集，图像无法恢复；

(4) 压缩过程是局部区域在全图上的搜索过程，运算量极大，例如在以双 68020-CPU并带 Auraro图形终端的 Masscomp 5600工作站上，编码一幅复杂的彩色静止图像约需 100 h，解码需 30 min[1]。因此，IFS法的高压缩比是以其编码的费时费力为代价的。

3.2 基于 IFS的分块编码

IFS法虽然压缩比很高 (Barnsley利用此法达到 10 000∶1的压缩比)，但编码费时费力。

此法的基本思想是: 某些图像, 比如人脸, 本身没有自相似性, 因此不可能有自相似 /仿射映射。但是图像的某一部分与另一部分相似 (如左脸与右脸差别不大)。这样我们可通过对图像的某些部分进行适当变换, 再由这些变换恢复成图像来逼近原图像。当然, 由于不是全图的仿射变换, 此法会引起一定的误差, 但此法可以自动实现, 且编码时间大大小于 IFS法。

编码步骤是先将原图像 f 分为互不重叠的小方块,称为分类块 R (range block)。且原图又分为稍大一些的,可重叠的范畴块 D (domain block) (大方块与小方块的尺度比称尺度压缩因子,一般尺度压缩因子小一些,容易实现 R,D 的匹配,但解码时误差较大;反之,编码时较难匹配,但解码误差较小)。然后寻找 R_i 的匹配块 D_i,匹配过程即通过变换 W_i 使 $W_i(D)$ 逼近 R_i。

变换 W_i 包括几何变换和亮度变换,几何变换采用仿射变换:

$$\begin{pmatrix} x \\ y \end{pmatrix} = S\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} R_x \\ R_y \end{pmatrix},$$

S 为尺度变换因子, R_x, R_y 为 D_i 的起始位置。$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 为变换矩阵,固定取表 1中的值。

表 1　A矩阵定义及取值

序号	A矩阵	变换意义
0	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	恒等变换
1	$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$	Y轴对称
2	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	X轴对称
3	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$	180°旋转
4	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$	相对 45°线对称
5	$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$	90°旋转
6	$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$	270°旋转
7	$\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$	相对 135°线旋转

亮度变换为: $Z = P_i \times Z + Q_i$

Z 为原图像亮度值, P_i, Q_i 为线性变换系数。

将每一 R_i 对应的 B_x, R_y, P_i, Q_i 及 A 矩阵序号记录下来,即为所求的 PIFS(Partitioned IFS) 代码。所以此法也可称 PIFS法,只要变换系统比原系统简单,便可实现压缩。

一般 R_i, $W_i(D_i)$ 很难完全匹配,因此定义一个平方根误差函数 (RMS) 来衡量 R_i 和 $W_i(D_i)$ 的近似程度,并取 RMS 最小的 $W_i(D_i)$ 做为 R_i 的最佳匹配。

平方根误差函数定义为:

对两个 $B \times B$ 大小的图像块 u,v, 有:

$$d(u,v) = \sqrt{\sum_{(i,j)=0}^{B-1} [u(i,j) - v(i,j)]^2}$$

房育栋等[5]人在 PIFS的基础上提出了快速局部迭代函数系统算法。该算法基于一种区域相关的假设,即设变换 U_i: $D_i \to R_i$, U_{i+1}: $D_{i+1} \to R_{i+1}$,若 D_{i+1} 在 D_i 邻域,则 R_{i+1} 在 R_i 邻域内概率较大,也就是说映射块在邻域的可能性很大,为此可在前一图像块 D_{i-1} 的最佳匹配映射位置 $(R_{x_{i-1}}, R_{y_{i-1}})$ 处设立窗口,在此窗口内搜索当前 D_i 的最佳匹配映射,当误差要求 (Error) 范围较大时, 该算法可大大缩短搜索时间 (参见图 1)。

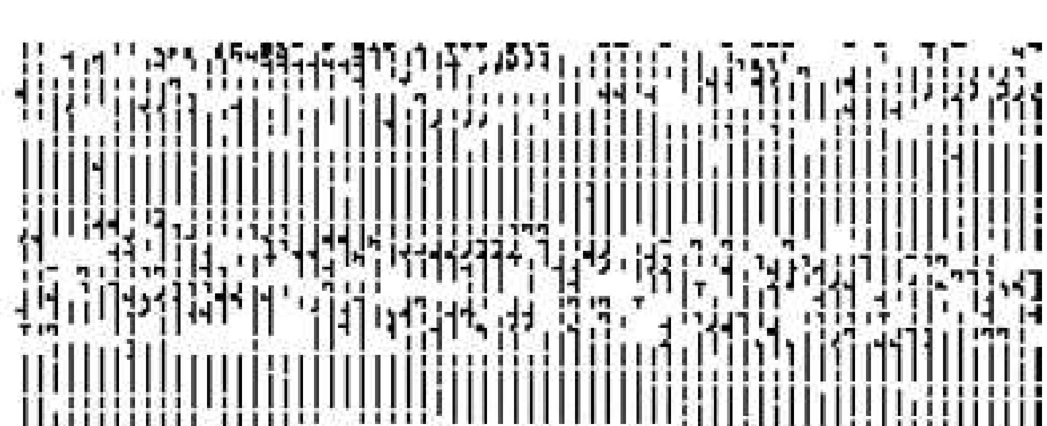

图 1　快速搜索次序[5]

3.3　基于 PIFS的其它方法

Jacquin提出的 PIFS算法使分形图像编码进入了一个新的时代。由于分形编码在保持图

能较好地保持图像低频成分;而且 PIFS方法在寻找最挂匹配时运算量较大,而采用内插则只用对插值图像的部分子块进行分形编码,因此这两方面结合可望使编码所需时间以及压缩比有显著提高。有关这方面的方法可参看文献 [6]。

3.4 基于分形维数的分割编码

人眼视觉系统 (HVS) 并非十分完美,有些图像失真它是感觉不到的,分割图像编码就利用了人眼的视觉特性及图像分析技术,将图像分割成一些区域,每个区域内的图像各有一些类似的特征,编码时对人眼较为敏感的特征加以强调,可实现高的压缩比,但传统分割编码的缺点在于只能把图像分为等密度的区域,而对于结构复杂的区域无法细分。针对这一缺点, Jang和 Rajala提出了在分数维数的基础上利用人眼可观察到的粗视化程度,把图像分为包含相似结构的区域集合的分割编码方法[4]。

我们已经知道,分数维数 D可以很好地描述分形物体的复杂程度,比如,人的动脉的分形维数约为 2.7,按理查森经验公式挪威海岸线的分形维数约为 1.52[7]。D越高,分形越复杂。

因此,可先计算图像的分形维数[2]。根据分形维数将图像分为几个结构阶层 (textual class),阶层Ⅰ观察到不变的图像密度值,其分形维数 $< D_1$;阶层Ⅱ观察到光滑结构,其分形维数介于 D_1, D_2之间;阶层Ⅲ观察到粗结构,其分形维数 $> D_2$ (D_1, D_2的值凭经验选取)。

编码时,对属于阶层Ⅰ的区域对其密度值和轮廓信息进行编码;对属于阶层Ⅱ的区域要着重考虑,因为此阶层位于人眼的敏感区域内,包含较多的信息;阶层Ⅲ的分形维数较高,属于它的区域一般包含图像的高频分量 (图像细节),可以高比例压缩。这样处理后,在原图像 256灰度级的情况下,可用 0.2 bit/像素进行编码,且重建图像质量良好[4]。

4 结束语

分形概念的提出使人类有了描述自然界更加有利的工具,分形图像编码是继 Shanno信息论为基础的预测编码、变换编码、子带编码等之后的又一新的编码技术,它以迭代规则和随机规则为特征,可实现高的压缩比。特别是 IFS的提出,具有重大意义,在交互式电视的应用上世有着重大的发展潜力。PIFS若与小波变换、自适合变换码本大小等处理方法相结合,可望大大改善它的失真度及编码时间等问题。

参考文献

1 余兆明. 数字电视和高清晰度电视. 北京: 人民邮电出版社, 1997. 61~123.
2 高安秀树. 分数维. 沈步明, 常子文译. 北京: 地震出版社, 1989. 52~55.
3 Jacquin A E. Image coding based on a fractal theory of iterated contractive image transformations. IEEE Trans on Image Processing, 1992, IP-1 (1): 18~30.
4 Guojun Lu. Fractal image compression. Signal Processing, Image Communication, 1993, 5: 327~343.
5 房育栋, 余英林. 快速分形图像压缩编码. 电子学报, 1996, 24 (1): 18~20.
6 王舟, 余英林. 一种新的分形图像压缩编码方法. 通信学报, 1996, 17 (3): 70~74.
7 张效忠. 分形. 北京: 清华大学出版社, 1995. 99~130.
8 吴敏金. 分形与图像压缩编码. 通信学报, 1993, 14 (2): 16~64.
9 祁大勇, 韩月秋. 分形与图像压缩. 电视技术, 1996, 9: 80~83.