

图像压缩的分形 理论和方法

陈仲仪 编著



清华大学出版社



致 读 者

本书由国防科技图书出版基金资助出版。

国防科技图书出版工作是国防科技事业的一个重要方面。优秀的国防科技图书既是国防科技成果的一部分,又是国防科技水平的重要标志。为了促进国防科技事业的发展,加强社会主义物质文明和精神文明建设,培养优秀科技人才,确保国防科技优秀图书的出版,国防科工委于 1988 年初决定每年拨出专款,设立国防科技图书出版基金,成立评审委员会,扶持、审定出版国防科技优秀图书。

国防科技图书出版基金资助的对象是:

1. 学术水平高,内容有创见,在学科上居领先地位的基础科学理论图书;在工程技术理论方面有突破的应用科学专著。

2. 学术思想新颖,内容具体、实用,对国防科技发展具有较大推动作用的专著;密切结合科技现代化和国防现代化需要的高新技术内容的专著。

3. 有重要发展前景和有重大开拓使用价值,密切结合科技现代化和国防现代化需要的新工艺、新材料内容的科技图书。

4. 填补目前我国科技领域空白的薄弱学科和边缘学科的科技图书。

5. 特别有价值的科技论文集、译著等。

国防科技图书出版基金评审委员会在国防科工委的领导下开展工作,负责掌握出版基金的使用方向,评审受理的图书选题,决定资助的图书选题和资助金额,以及决定中断或取消资助等。经评审给予资助的图书,由国防工业出版社列选出版。

国防科技事业已经取得了举世瞩目的成就。国防科技图书承

担着记载和弘扬这些成就,积累和传播科技知识的使命。在改革开放的新形势下,国防科工委率先设立出版基金,扶持出版科技图书,这是一项具有深远意义的创举。此举势必促使国防科技图书的出版,随着国防科技事业的发展更加兴旺。

设立出版基金是一件新生事物,是对出版工作的一项改革。因而,评审工作需要不断地摸索、认真地总结和及时地改进,这样,才能使有限的基金发挥出巨大的效能。评审工作更需要国防科技工业战线广大科技工作者、专家、教授,以及社会各界朋友的热情支持。

让我们携起手来,为祖国昌盛、科技腾飞、出版繁荣而共同奋斗!

国防科技图书出版基金
评审委员会

国防科技图书出版基金 第三届评审委员会组成人员

名誉主任委员	怀国模		
主任委员	黄 宁		
副主任委员	殷鹤龄	高景德	陈芳允
	曾 铎		
秘 书 长	崔士义		
委 员	于景元	王小谟	尤子平
(以姓氏笔划为序)	冯允成	刘 仁	朱森元
	朵英贤	宋家树	杨星豪
	吴有生	何庆芝	何国伟
	何新贵	张立同	张汝果
	张均武	张涵信	陈火旺
	范学虹	柯有安	侯正明
	莫梧生	崔尔杰	

目 录

第一章 绪论

§ 1.1 分形与混沌: 孪生兄弟	1
§ 1.2 为什么分形方法能压缩图像	6
§ 1.3 用分形方法进行图像压缩的基本思路	9

第二章 理想图像的数学模型

§ 2.1 人类视觉	14
§ 2.2 理想图像	17
§ 2.3 理想图像的数学模型	20
§ 2.4 离散图像	24
2.4.1 离散黑白图像	24
2.4.2 离散灰阶图像	25

第三章 图像压缩分形理论的数学基础

§ 3.1 基本集合论	26
3.1.1 集合	26
3.1.2 集代数	26
3.1.3 笛卡儿积、向量和、数乘	28
§ 3.2 度量空间、映射、变换	30
§ 3.3 欧氏空间的仿射变换	32
3.3.1 线性变换和仿射变换	32
3.3.2 一维空间的仿射变换	32
3.3.3 欧几里得平面上的仿射变换	33
3.3.4 三维实空间的仿射变换	35
3.3.5 欧几里得平面上线性变换的范	35
§ 3.4 度量空间的拓扑性质和变换	36
§ 3.5 压缩映射定理	44
3.5.1 完备度量空间的压缩映射定理	44
3.5.2 不完备度量空间的压缩映射定理	47

§ 3.6	拼贴定理	48
3.6.1	完备度量空间的拼贴定理	48
3.6.2	度量空间的拼贴定理	50
§ 3.7	码的稳定性	51
§ 3.8	迭代变换理论与分形几何的联系	53
第四章	迭代函数系统(IFS)理论	55
§ 4.1	图像的空间	55
§ 4.2	H 空间上的压缩映射	58
§ 4.3	迭代函数系统(IFS)	60
§ 4.4	二维空间 R^2 中仿射变换的 IFS	63
§ 4.5	计算 IFS 吸引子的复印机算法	64
§ 4.6	Hausdorff 空间的拼贴定理	64
§ 4.7	用 IFS 的黑白分形图像压缩	65
第五章	灰阶图像的迭代函数系统(IFS)方法	68
§ 5.1	测度论的基本概念	68
5.1.1	测度	69
5.1.2	Borel 集、Borel 测度	69
§ 5.2	随机的迭代函数系统	71
§ 5.3	灰阶图像的复印机压缩算法	74
§ 5.4	寻找 Markov 算子的基本思路	75
第六章	基于迭代变换理论(ITT)的分形编码方法	77
§ 6.1	迭代变换理论(ITT)	77
§ 6.2	ITT 编码的数学基础	79
6.2.1	对象支持的分割	79
6.2.2	距离函数	80
6.2.3	二维 Borel 测度空间 P 上的 Markov 算子	81
6.2.4	测度的编码和解码	83
6.2.5	分辨率	85
§ 6.3	基于 ITT 的数字图像编码系统	86
6.3.1	图像编码概述	86
6.3.2	图像分割、距离和变换	87
§ 6.4	数字图像的分块式分形编码程序	92
6.4.1	程序概述	92
6.4.2	域块库的构造	93

6.4.3 变换库的构造及搜索方法	94
§ 6.5 系统特性	96
6.5.1 比特率的计算	96
6.5.2 图像重构	98
6.5.3 分块式分形码的特性	99
第七章 研究进展	101
§ 7.1 与向量量化(VQ)压缩方法的关系	101
§ 7.2 与小波压缩方法的关系	103
§ 7.3 分形图像压缩方法的新进展	106
7.3.1 不必搜索的分形图像编码	107
7.3.2 增大区块和减少搜索时间的分形方法	108
7.3.3 近似正交基的分形压缩方法	110
§ 7.4 分形方法的进一步分析	110
7.4.1 改进的拼贴定理	111
7.4.2 分形图像压缩的收敛性	113
第八章 结语	116
后记	121

第一章 绪 论

图像压缩的分形方法要解决这样一个问题:给定一幅离散的原始图像,怎样用计算机产生一幅分形图像,它看上去与原图像很相似,并可以用比原图像更少的位数表示。

围绕着这个中心问题,我们在本章简述:(1)什么是分形;(2)为什么分形方法可以解图像压缩;(3)用分形方法进行图像压缩的基本思路。

§ 1.1 分形与混沌:孪生兄弟

本世纪 70 年代,由于计算科学的飞速发展,非线性科学出现了一对孪生兄弟:混沌与分形。

研究表明,宇宙中许多表面上服从决定论定律的简单系统,其行为仍然是很难预测的,由此产生了混沌理论。大自然中许多常见的不规则的复杂现象,如山峦和云团的外形,曲折的海岸线等。它们处处无规则可言,但整体与其局部有相似性,因而导致了分形理论。

混沌事件在时间标度上表现了相似的变化模式,分形在空间标度上表现出相似的结构模式,它们表明混沌与分形之间有密切的关系。混沌是演化的科学,分形是存在的科学;混沌是过程的科学,分形是状态的科学,有人称分形几何就是混沌几何。

实际上,分形结构往往是混沌动态过程的产物。只要混沌过程在一些地方形成了某种环境(如海岸、大气、地质断层等),就很可能留下分形结构(海岸、云、岩层等)。

混沌学说的奠基人 Feigenbaum 在 1992 年说,混沌和分形之

所以迟至本世纪 70 年代中期才蓬勃发展,并非由于以前人们的智力不足,也不只是因为计算技术的发展,而是因为人们的智力与计算技术的完善的有机结合。我们因此对长年观测到的现象有了新的理解,它并非只停留在计算机上,而是将它公式化,变为我们可以操作的概念,这是具有新知识的人凭藉计算机威力迅速取得的结果。这些观念往往是带根本性的。如果没有计算结果的支持,人们即使明显地感觉到它们存在,也不知如何得到它们。计算技术的巨大发展,为科学的发现提供了更深层次的信息,瓜熟蒂落,水到渠成,混沌与分形这对孪生兄弟便应运而生。

自然界的形状和各种图形可分为两类,一类是有特征长度的图形,一类是没有特征长度的图形。

有特征长度的事物如房屋、汽车、足球、人……,房子的宽高,汽车的长度,足球的直径,人的身长都是特征长度,这些事物的形状可以用线段,圆这些基本要素去逼近,这些线和面几乎都是光滑的,几乎处处可以求微分。

没有特征长度的形状是海岸线、云……,如果没有人工参照物,很难测量其尺度,仔细观察其局部可以发现许多细节,将细节放大,又发现局部与整体相似。没有特征长度的图形,其重要性质是自相似。典型例子是 Koch 曲线,它是将一根直线线段等分为三段,中间一段以两段等长线段构成的凸折线代替,这过程无穷地重复下去构成的(见图 1.1)。

分形,就是那些没有特征长度的图形的总称,Mandelbrot 说“1975 年,我由描述碎石的拉丁文 Fractus,创造出 Fractal(分形)一词,分形是几何外形,它与欧几里德外形相反,是没有规则的,首先它们处处无规则可言,其次它们在各种尺度上都有同样程度的不规则性。”

分形还没有确切的简明的定义,但是分形的集合具有如下特征:

- (1)该集有精细结构,即在任意小的比例尺度内包含整体;
- (2)无论从局部和整体来看,分形集都是很不规则的,无法用

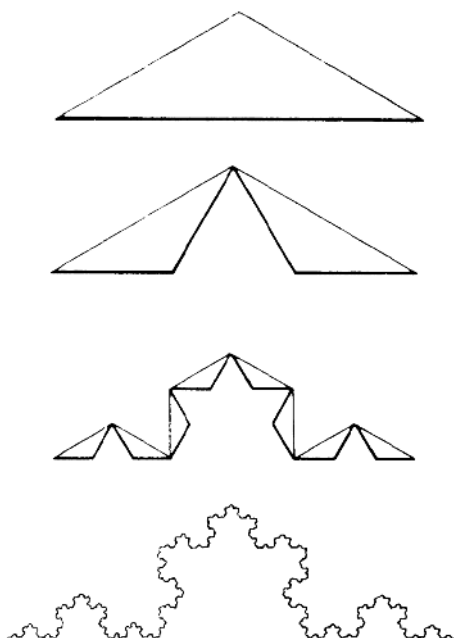


图 1.1 Koch 曲线。将线段等分为三段，中间一段以两段等长线段构成的凸折线代替，此过程无限地重复下去，即可构成 Koch 曲线

传统的几何语言来描述：

(3) 通常分形集都有某种自相似性，可能是近似的，也可能是统计意义；

(4) 通常分形集的“分形维数”比它的拓扑维数要大；

(5) 许多情况下，分形集是非常简单的，或者是递归的。

欧几里得几何和微积分方法不适合于分形的研究。分形的主要工具是它的维数，粗略地说，维数表示一个集合占有多大空间，

它定量地表述分形的形状和复杂性。

(1)最简单的分形维数是相似性维数 D_s

如图 1.2 的线段、正方形、立方体……,当线度减少一半时,相似形的个数 $N(1/2)$ 为

线段 $N(1/2) = 2 = 2^1$

正方形 $N(1/2) = 4 = 2^2$

立方体 $N(1/2) = 8 = 2^3$

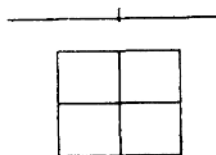


图 1.2 线段、正方形的维数

当把某图形线度缩为原来的 $1/a$,该图形是由 $N(1/a)$ 个与原图形相似的图形所构成, $N(1/a) = a^{D_s}$.

正常的线、面、体、维数 D_s 分别是 1、2、3,都是整数。对 Koch 曲线, $N(1/3) = 4$ 。

(2) Hausdorff 维数 D_H

假定维数 $D_H > 0$,用直径小于 ε 的可数个球覆盖集合 E ,此时,若 $d_1, d_2, \dots, d_k$ 为各球的直径,则 D_H 维 Hausdorff 测度可用下式定义:

$$M_D(\varepsilon) = \lim_{\varepsilon \rightarrow 0} \inf \sum_{d_i < \varepsilon} d_i^D \quad (1.1)$$

当该测度从 0 向无限大迁移时,在间断点上的 D 称为集合 E 的 Hausdorff 维数,以 D_H 表示。

[例 1]

直线段的测度是线长, $D_H = 1$,即一维测度是直线段长度;二维测度是面积;三维测度是体积。

[例 2]

Koch 曲线。

Koch 曲线为集合 E ,用 $\varepsilon = 1/3, d_i < 1/3$ 的多个小球覆盖,比如七个小球(图 1.3)。覆盖的方法应该完全没有浪费(没有重叠,即下确界是 $\inf$)。

如用一维测度,每当把单位长度缩为原先的 $1/3$ 时就出现 4

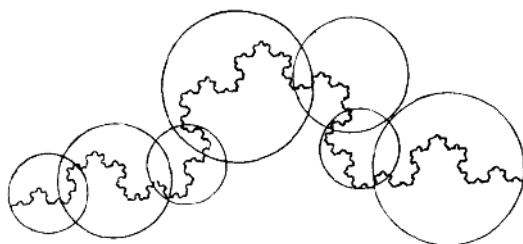


图 1.3 小球覆盖 Koch 曲线

个同样图形,即是说,若覆盖球直径为 $1/3$,近似图形长度就伸长 $(4/3)$ 倍,若球直径为 $(1/3)^n$,则总长度就为 $(4/3)^n$ 。所以,球的直径趋于零时,一维测度(线长)趋于无限大,由于 Koch 曲线是线,不难推知其二维测度(面积)为零。在 $D=1$ 与 $D=2$ 之间测度从无穷大变为 0 的过程中,在间断点 $\lg 4/\lg 3$ 处有有限值,该点的维数(D 值)就是 Hausdorff 维数(图 1.4)。

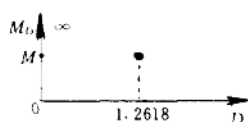


图 1.4 Hausdorff 维数

$$D_H = \lg 4 / \lg 3 = 1.2618$$

该值又可证明如下,用多个大小相等的小球,不重叠地完全覆盖 Koch 曲线,球的直径为 $(1/3)^n$,个数便为 4^n 。从定义出发,该曲线的 Hausdorff 测度

$$M_{D_H}(\varepsilon) = \lim_{N \rightarrow \infty} (4)^N (1/3)^{N \cdot D_H} = M(\text{有限值}) \quad (1.2)$$

$$3^{N \cdot D_H} \cdot M = 4^N$$

$$\lg M + N \cdot D_H \lg 3 = N \lg 4$$

$$\text{当 } N \rightarrow \infty, N D_H \lg 3, N \lg 4 \gg \lg M$$

$$\therefore \lim_{N \rightarrow \infty} N \cdot D_H \lg 3 = N \lg 4$$

$$D_H = \lg 4 / \lg 3 = 1.2618 \quad (1.3)$$

分形维数还有多种定义方法,分形维数表示分形的复杂程度,以曲线为例,直线段的分形维数为 1,最简单。维数越接近 2,形状越复杂,占据的面积越大,最复杂的 Peano 曲线 $D_H=2$ 。它能覆盖整个平面。

分形是一类看上去复杂性颇高的特别的几何图形,它与欧几里得几何的根本区别是,其基本要素不是直线、圆之类的直观元素,而是用算法和程序集来表示,要凭藉计算机显示出它的几何形态。

用语言可类比这两种不同的几何理论。

印欧语言以一个有限的字母表(英语是 26 个字母)为基础,串接成各种单词,欧几里得几何则拥有能构造各种复杂图形的少量基本要素(直线、圆等)。

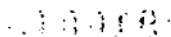
汉语等一类语言由本身就有含义的象形符号组成,这类语言拥有的符号或元素成百上千,可认为是无限的。分形几何采用大致相同的方式组建而成,它由无数的元素组成,每个元素都是完整和唯一的,它们由算法定义,算法在分形几何语言中起基本单位的作用。

§ 1.2 为什么分形方法能压缩图像

根据分形理论,不少复杂的图形,从计算的观点看来,其信息含量并不大,可以用简单的程序来产生。

比如蕨类植物的图形,看上去比较复杂,用一般图形表示法要用上万个数据,但采用分形方法,仅需 24 个数据,用迭代函数系统(IFS)方法便可在计算机上产生这种图形(图 1.5)。

这说明,通过迭代这全反馈的动态过程,用参数不多的算法就可以在计算机上显示出相当复杂的自然图形,复杂的图形寓于简单的算法之中。这是可以采用分形方法进行图像压缩的主要根据之一。



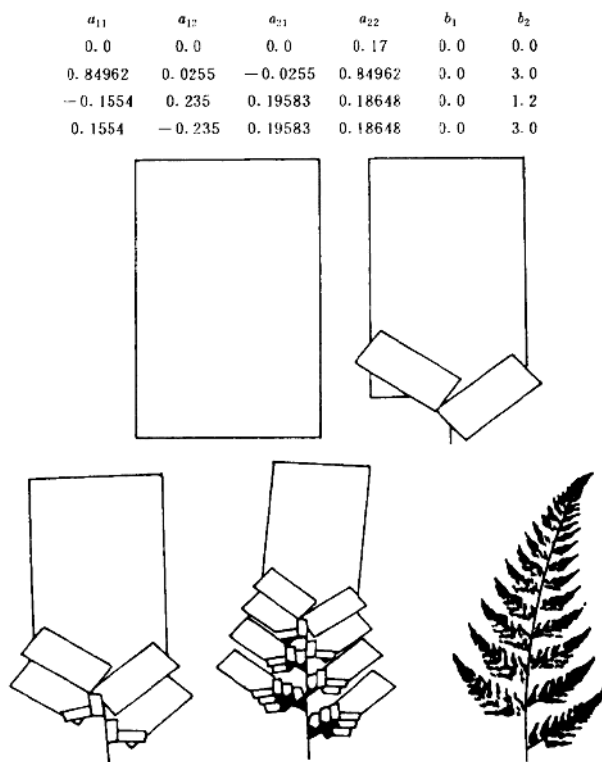


图 1.5 蕨类植物的图形。以任意图形(此处为矩形方块)为初始图形,以一组四个仿射变换作多次迭代变换,便得到极限图形,它由 24 个数据唯一确定,而通常这些复杂图形要占有上万个数据

科学家们发现,自然界许多事物的发展过程,如生长、凝聚、进化等,形成多种分形结构。

以人的循环系统为例,从主动脉到毛细血管,形成一类连续分布,直至细到血球细胞只能排单行滑行,它们的分支是一种分形结

9910015

构,就象 Koch 曲线把无限长的曲线挤进有限的面积一样,循环系统把巨大的表面挤进有限的体积中。大自然所发明的分形结构非常有效。在多数组织中,一个细胞到血管的最远距离,很少超过三至四个细胞的大小,这样紧凑有效的结构中,血管所占的空间还不到人体的百分之五。

理论生物学家推测,在形态发生的过程中,分形结构不仅是常见的,而且是普遍的。

树木的枝叶呈分形,用以获取阳光、空气和抵抗风力。在植物的种子里,只有一定的信息为这种植物编码,所以植物的复杂程度是有限的。它不过是在生长过程中新陈代谢(不断迭代)而形成的复杂分形形态。

因此,分形意味着自然界是许多复杂形态中潜藏着有组织结构。大自然的美,就在于她在本质上是简单的。如果我们能找到这些有效的信息,我们就能简洁地表述自然界的图景。

这是我们能够采用分形方法进行图像压缩的另一个重要根据。

当然,这些称为分形的自然体,没有一个是真正的分形。它们只是在一定的尺度内表现得非常像分形,在这种尺度范围内可以看作分形。“自然分形”与数学上的“分形集”是有区别的,本质上不存在真正的分形,就像并不存在真正的线和圆一样。但是,这并不妨碍我们用欧几里德几何去描述物体的形状,也不妨碍我们用分形方法去进行图像压缩。

用分形方法进行图像压缩的先驱是 Barnsley。

1979 年, Barnsley, 这位牛津大学培养出来的数学家,去意大利科西嘉岛开会时,遇见 Feigenbaum,知道了混沌理论的基本概念。他的直觉告诉他那些正是分形的对象,他深入研究 Mandelbrot 分形集,发现自然界许多事物隐藏着分形结构,可以用很少几条规则把它们翻译出来,这些模型比欧几里得的几何模型有意义得多。于是,他在美国乔治亚理工学院教授数学的同时,致力于寻找这些分形的秩序。

分形与混沌理论

80年代中期, Barnsley 从数学上找到了重要的理论根据, 率先提出用迭代函数系统(IFS)这种分形方法进行图像压缩, 为图像压缩提供了一条与以往完全不同的新思路。在他的指导下, 他的得意门生 Jacquin 首次提出分块的迭代变换理论算法。这是第一个利用计算机进行图像的分形压缩的自动算法, 对分形图像压缩方法的实用化, 起了奠基的作用。

§ 1.3 用分形方法进行图像压缩的基本思路

迭代函数系统, 其实是一组压缩仿射变换 $[W_1, W_2, \dots, W_N]$ 每个压缩仿射变换 W 形如:

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = W \begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_n \\ y_n \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \quad (1.4)$$

$$d(W(X_1), W(X_2)) \leq s \cdot d(X_1, X_2) \quad (1.5)$$

$$0 \leq s < 1$$

x_i, y_i 是点 X_i 的坐标, $a_{11}, a_{12}, a_{21}, a_{22}, b_1, b_2$, 是变换常数, $d(X, Y)$ 表示点 X 与点 Y 的距离, s 是小于 1 的正常数, 称为压缩因子, 压缩仿射变换其实是由缩小、旋转、反射、位移等基本运算构成的变换, 其迭代运算得到的极限图像与初始图像无关, 它由一组仿射变换唯一确定。前面说过的 24 个参数即一组四个仿射变换确定了一幅蕨类植物的图像。

图像压缩是一个逆问题, 给定一个图像, 如何找出这一组压缩仿射变换的参数。

我们以一维的压缩变换为例, 来说明用分形方法进行压缩的原理。

设 $f(X)$ 是实数轴某区间 $[a, b]$ 内的压缩变换

$$f(x) = \alpha x + \beta \quad |\alpha| < 1 \quad (1.6)$$

对区间 $[a, b]$ 内的点, $x_{n+1} = f(x_n) = \alpha x_n + \beta$ 此迭代变换产生一个收敛点序列, $\{x_0, x_1, \dots, x_n, \dots\}$ 收敛于不动点 x_f 。

根据不动点的定义, $x_f = f(x_f)$, 由该变换唯一确定, 而不管初

始点 x_0 如何选取, 很易证明,

$$f(x) = \alpha x + (1 - \alpha)x_f \quad (1.7)$$

只要 $|\alpha| < 1$, 该函数迭代的结果必然是 x_f 。换句话说, 式(1.6)给不动点 x_f 编码, 确定了

$$x_f = \beta / (1 - \alpha) \quad (1.8)$$

假设我们有一个具有约定编译码方式的通讯系统。为了把数据 x_f 从发射端送到接收端, 在发射端只需按式(1.6)把 α 和 β 两个数送出, 接收端收到 α, β 后, 通过迭代运算便把

$$x_f = \lim_{n \rightarrow \infty} f^{(n)}(x_0)$$

译出。

以此为例, 我们可以构造任何一个收敛于 x_f 的压缩变换

$$f(x) = [x + (g - 1)x_f]/g, \quad |g| > 1 \quad (1.9)$$

压缩因子 $\alpha = 1/g, \beta = (g - 1)x_f/g$, 不动点为 x_f 。

重要的是不动点是唯一的。发射端送出 α, β , 接收端通过迭代便可以唯一地解出 x_f 。

当然, 在一维的实数时, 找出相应的压缩变换相当简单, 在二维图像的情况下, 构造压缩变换就困难得多, 但拼贴定理告诉我们, 可以做到这一点。

仍以一维实数为例, 我们要对 x_f 编码, 假定很难找到以 x_f 为不动点的函数 $f(x)$, 但可以找到另一个压缩变换函数 $g(x)$, 它的不动点 $x_g = g(x_g)$, 近似等于 x_f 。其压缩因子为 $|s_g| < 1$ 。我们看看, 当不动点与 g 对不动点的变换相差不大, 即 $d(x_f, g(x_f)) < \varepsilon$ 时, $g(x)$ 的不动点与 $f(x)$ 的不动点究竟相差多大? 按三角不等式:

$$\begin{aligned} d(x_g, x_f) &\leq d(x_g, g(x_f)) + d(g(x_f), x_f) \\ &= d(g(x_g), g(x_f)) + d(g(x_f), x_f) \\ &\leq |s_g| d(x_g, x_f) + d(g(x_f), x_f) \\ d(x_g, x_f) &\leq \frac{1}{1 - |s_g|} \cdot d(g(x_f), x_f) = \frac{\varepsilon}{1 - |s_g|} \quad (1.10) \end{aligned}$$

这结果在图像范围内称为拼贴定理。为什么叫拼贴定理? 因为我们将此结果推广到数字图像时, 我们想寻找一个图像的压缩