

基于分形编码拓扑特性的图像检索

陈添丁¹⁾ 刘济林²⁾ 楼伟进³⁾

¹⁾(杭州商学院信息与电子工程学院, 杭州 310035) ²⁾(浙江大学信息与电子工程系, 杭州 310027)

³⁾(浙江大学计算机科学系, 杭州 310027)

摘要 为了更有效、更准确地进行图像检索,提出了一种利用分形编码这项重要的拓扑特性来处理图像索引的新方法,即将图像经分形编码,首先得到每张图像的迭代函数,然后将其伴随图像存入数据库中,成为该图像的索引文件最后对数据库进行搜索时,则通过对此索引文件的比对来找出与查询图像相似的图像。反观使用其他方法建立的图像索引数据库,则无法证明其建立的索引文件具有上述特质。实验显示,图像经过分形编码所表现出的几何性质以及独特的有效性和鲁棒性,证明该方法是一个更有效率、准确度高的检索方法。

关键词 分形编码 迭代函数 索引文件 检索

中图法分类号: TP391.3 文献标识码: A 文章编号: 1006-8961(2004)01-0056-06

A Image Retrieval Model of Fractal-based Encoding Topological Characteristic

CHEN Tian-ding¹⁾, LIU Ji-lin²⁾, LOU Wei-jin³⁾

¹⁾(College of Information and Electronics engineering, Hangzhou University of Commerce, Hangzhou 310035)

²⁾(Department of Information and Electronics engineering, Zhejiang University, Hangzhou 310027)

³⁾(Department of Computer Science, Zhejiang University, Hangzhou 310027)

Abstract Fractal code is approved an effective method to compress digital image. It proposes a new method to analyze digital image retrieval by fractal iterative function system. By image fractal code and obtained its iterative function, then the image and its iterative function are saved database becoming index file of the image. When database is searched and image index file is compared so that user retrieval images similar to query image. The index files of based on fractal code have three characteristics, the first, similar images have similar iterative functions so can produce similar index files; the second, similar index files can retrieval similar images; the third, no similar images have no similar iterative functions, vice versa. Compare to other methods are used to create database, their index files have not these characteristics. The fractal code produces large of data and needs an effective method to retrieval. So it combines to fractal function information, based on discriminant analysis estimating images similarity, so that determine correlation sequence of all images in database similar to query image. Experiment shows this retrieval method specialties, efficiency and robustness.

Keywords Fractal code, Iterative function system, Index file, Retrieval

1 引言

目前以图像为键值来进行图像数据库索引的技术大致可分为根据图像的颜色、形状、内容来建立图像数据库索引文件的多种方法。虽然分形使用在图像压缩上,已被证实有良好效果^[1,2],但作为一个新方法来处理图像搜索,尚未见有效的研究。其压缩方法是先将原图像分割成数个方块,然后对于每一个

方块,应在原图像中找到一个较大方块及一个变换函数,以使得较大方块经此函数处理后,会与较小方块相似,再结合所有方块转换函数即成为一个函数集。根据不动点(Fixed Point)性质,若使用此函数集不断进行迭代变换(Iterated Transformation),则无论初始图像如何,终将得到同一归结图,即最终都将收敛到同一张图像,而且归结图只与分形迭代函数有关,而起始图像则与最后归结图互不相关。也就是说,使用分形编码来压缩图像,其转换方式决定了译

码后的图像,即结果与初始图像无关。由此可知,真正能影响结果的因素是此函数集,或称迭代函数系统(Iterated Function System, IFS),本文运用这项重要特性,提出了一个新方法来进行图像搜索,即将图像经分形编码,首先得到每张图像的迭代函数,然后将其伴随图像存入数据库中,使之成为该图像的索引文件,日后从数据库中搜索图像时,就可通过对索引文件的比对来找出与使用者选定的查询图像相似的图像。这样以分形迭代函数系统为基础的编码算法,就可利用分形几何原意的拓扑特性来研究有识别特点的检索,虽然其他检索方法有特定擅长的处理条件,但无法保证其使用的方式能符合下述要求:(1)根据相似图像能产生相似索引文件;(2)根据相似索引文件能检索到相似图像。

2 相关研究与分形索引理论

2.1 相关研究

随着信息科技的发展,数据类型不再仅限于文字,还出现大量多媒体影音信息,因此,近年来虽有很多学者投入关于图像检索的研究,但其研究模式仍离不开分析图像、特征比对,其中,图像特征主要为颜色、形状、纹理、运动物体等。

有一些组织对如何能找出最相似的相关图像开展了研究,如 QBIC^[3]是 IBM 公司研究中心研发的图像搜索匹配系统,此系统可以从大型图像数据库或视频图像中找出用户希望查寻的相似图像。

如果以图像的内容组成元素为搜索特征数据库,则搜索就是以图像中各种组成物的比例来判定图像的相似度,若有两张图像,其组成物以相似比例组成,则意味着这两张图像相似^[4,5]。在图像中,可依人们对事物的了解,将组成物分成为背景、纹理及目标,其判定方法如表 1 所示,由表 1 可见,背景有较少量边缘存在,并有较高颜色相似度;纹理有较少或正常颜色数目,但却有较高比例边缘存在和较低颜色相似度;目标物体有一般或较多颜色数,且包含边缘数目正常,但颜色相似度较低。

表 1 不同图像内容有不同特征

种类	背景	纹理	目标
色彩的数目	正常	少或正常	正常或多
边缘信息	较少的长线	很多短线或少的长线	正常
相似颜色	高	低	低

有了各种图像内容组成的比例关系后,就可再

以组成物的相对位置来增加图像的判别度。若加入更多不同类型的特征,则将使建立的数据库索引更具判别力,但加入更多的特征,也将造成系统更大负担,再者,由于现实生活中对许多事物并无明确定义,因此对人类而言,虽可以依感觉来处理或判断,但这些信息若交由计算机处理前,则需先量化,以表 1 为例,多少比例才可称为多,颜色数目多少才能归类为背景?这些问题成了首先需要解决的问题,若无法处理这些问题,仍要根据内容建立索引,这便成为一项艰巨任务。

2.2 理论论证

迭代函数系统被定义为由多个具有收敛特征的变换函数所构成的集合,可表示为: $T = \{t_i: R^2 \rightarrow R^2 | i = 1, 2, \dots, m\}$ 。而变换的收敛性是以 d 为计算两点距离的函数, t 为一变换, p_1, p_2 为空间中两个点。若 t 满足: $d(t(p_1), t(p_2)) \leq sd(p_1, p_2)$, 其中 $0 \leq s < 1$, 则 t 满足收敛性,对于空间中任意一个点,共经过多次变换,其将趋近一固定位置,则称为不动点^[6~8]。

分形迭代函数满足收敛性,由于其是满足收敛性的函数,可使空间中任意两点距离缩短,因此,此函数又可称为压缩函数,而满足收敛性的函数,其将具唯一的不动点(依 Banach 不动点定理),即无论初始图像如何,只要经分形迭代函数的足够次数操作,均将得到唯一归结图。因分形函数也是一满足收敛性的函数,故可知其归结图唯一存在。

由此可知,一套良好的资料检索系统,需有良好的索引文件及有效的相关排名输出计算方法,但由于唯有良好的索引文件才能产生好的正确相关的排名输出,因此索引文件决定了数据库的搜索效能。良好的索引文件应符合以下两个性质:(1)高相关度信息有高相关索引文件;(2)索引文件相关度高,其信息相关度也高,但是并非所有图像数据库索引方式皆可满足此两性质,可以颜色建立索引为例说明,如红玫瑰与黄玫瑰虽形态相似,但颜色不相似;又如,同为白色的云与车子并不相似。利用分形技术建立的图像索引,应具有如下特性:(1)由于相似图像有相似分形迭代函数,因此能产生相似索引文件;(2)由于相似归结图有相似分形迭代函数,因而相似索引文件能检索到相似图像。关于分形之所以具备成为一个良好索引的条件,文献^[6~8]已给予详细论证。

定义 1 函数 f 为距离空间 (K, d) 上映射至它自身的函数,若存在点 $a \in K$, 使得 $a = f(a)$, 则称 a 为 f 的不动点。

若函数满足不动点定理,则此函数的不动点即可以由此函数表示。同理,由于分形函数符合不动点定理,因此其归结图可以由此函数表示,以达到图像压缩的目的,若此函数无法符合不动点定理,则自然无法得到归结图,可以说,不动点定理是分形编码技术的精髓。

定义 2 给定任意 $\varepsilon(\varepsilon > 0)$, 若点列 $\{x_n\}$ 对应存在 n_0 , 当 $m \geq 0, n \geq n_0$, 使得 $d(x_n, x_m) < \varepsilon$, 则称 $\{x_n\}$ 为柯西点列。

定义 3 距离空间 (K, d) 中, 若每一柯西(Cauchy)点列都会收敛到 K 中, 则称 K 是完备距离空间(Complete metric space)。

这样就可引出 Banach 不动点定理, 即在完备距离空间中, 若满足收敛性函数, 则必有唯一的不动点。

若一距离空间具备完备性, 则能确保其结果也在此空间中, 反之不然, 因此, 若分形迭代函数所在空间为完备距离空间, 则可知其所得的结果也会在此空间中, 即不致造成结果不存在的现象。

有些研究虽提出了一些建立图像数据库索引文件的方法, 但无一能证明其具良好图像数据库索引文件的性质, 而本文通过引用的两个定理来说明用分形函数建立的图像数据库索引文件则具上述两个性质, 这在研究图像数据库索引领域是一大突破。

定义 4 f, g 两函数为距离空间 (K, d) 上映射至它自身的函数, 对任意 $x \in K$, 若满足 $d(f(x), g(x)) \leq \varepsilon$, ε 为一很小值, 则称 f, g 两函数靠近。如图 1 所示, 两函数很靠近, 其形状虽可能不同, 但任何一点的值皆很接近。



图 1 两函数相互靠近

定理 1 设 g, h 两函数在距离空间 (K, d) 上映射至它自身, 且满足收敛性, 其不动点分别为 f_g, f_h , 则对于 K 中任意两点, 此两函数有近似斜率, 且斜率不靠近 1, 即在适当区间内, 若 g, h 靠近, 则 $f_g,$

f_h 靠近; 反之, 若 f_g, f_h 靠近, 则 g, h 靠近。

定理 2 设 g, h 为距离空间 (K, d) 上映射至它自身的两函数, 且满足收敛性, 其不动点分别为 f_g, f_h , 则对于 K 中任意两点, 此两函数有近似斜率, 即, 若 g, h 不靠近, 则 f_g, f_h 不靠近; 若 f_g, f_h 不靠近, 则 g, h 不靠近。

2.3 定理应用

将定理 1 应用于 Jacquin 分形编码上, 由于分形迭代函数具有收敛性, 若知其存在唯一不动点, 则此不动点就是归结图, 且两分形迭代函数若相似, 则其归结图相似; 反之, 若两归结图相似, 则两分形迭代函数也相似。由此可知, 使用分形编码建立的索引文件, 具有良好索引文件的性质。

此外, 定理 2 说明, 分形函数可保证不相似图像, 其分形迭代函数不相似; 反之亦然, 因此, 使用分形迭代函数来建立索引文件, 除可找到相似图像外, 也可使不相似图像被排除在搜索结果之外。

Jacquin 在文献[9]中使用的分形编码函数包含了缩小、旋转、翻转及不动点函数 $f(x) = kx + c$, 其中缩小、旋转、翻转等操作将产生不同起始点, 而使用的函数 $f, f(x) = kx + c, 0 \leq k < 1$, 满足收敛性, 依 Banach 不动点定理, 其不动点唯一存在, 由于其初始点改变并不影响不动点, 故在编码时, 应尽可能地使 k 不靠近 1, 且应使斜率(k)范围小, 即取一足够小区间 $[a, b] \subset [0, 1], k \in [a, b]$, 依定理 1 及定理 2, Jacquin 使用的方法也具有如下性质:

(1) 若两分形函数相似, 则其归结图也相似; 若两分形函数不相似, 则其归结图也不相似;

(2) 若两图像相似, 则其分形函数也会相似; 若两图像不相似, 则其分形函数也不相似。

这样, 由于定理 1, 2 说明分形具有建立良好图像数据库索引文件的性质, 因此可以说使用分形编码方法建立图像数据库索引文件最佳。

2.4 分形编码

Jacquin 为使每张图像能自变换(Self-transformable), 就使用了使图像结构简单化的策略, 所以在文献[9, 10]中提出的分形编码方法, 就先将图像分割成数个大小为 $B \times B$ 的不重叠方块, 称为 Range 块, 而将同一张图像, 再分割成数个大小为 $D \times D$ 的重叠方块, 称为 Domain 块, 且 $D > B$, 经适当大小分割后, 由于所得的 Range 方块具有较简单的结构, 因此易于从同一张图像中找到一 Domain 方块, 使其缩小、翻转后相似(如图 2 所示)。

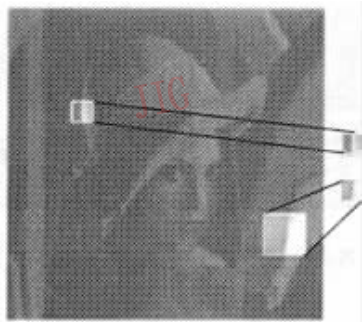


图 2 同一张图像中找出相似的 Range、Domain 方块

本文取 Domain 块大小为 $8 \times 8 \text{Pixel}^2$, Range 块大小为 $4 \times 4 \text{Pixel}^2$, 所以一张图像 R、G、B 3 个平面, 每个平面可分成 4 800 个 Range 方块, 72 384 个 Domain 方块, 再配合文献[9, 10]中提出 8 个变换, 则每张图像最多需进行 $3 \times 4\ 800 \times 72\ 384 \times 8$ 次比对, 每次比对的主要目的是对于每一个 Range 方块 R_i , 需找到一个 Domain 方块 D_j 及一变换函数 τ_a , 使得 R_i 与 $\tau_a(D_j)$ 有小差异(如图 3 所示)。

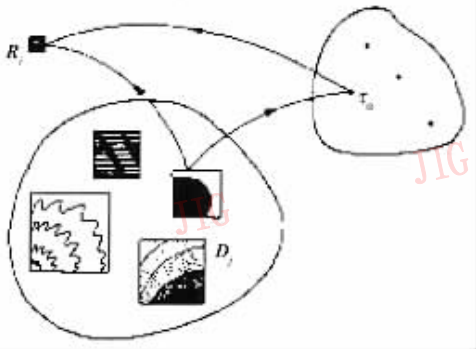


图 3 由多个 Domain 及 8 个变换中选出最佳组合

得到最佳 Domain 块及变换组合后, 为使其误差更小, 可将每个 $\tau_a(D_j)$ 再以函数 g 做修正, $Z = g(z) = sz + o$, z 为 $\tau_a(D_j)$ 上每个像素的颜色, $s = \{0.5, 0.6, 0.7, 0.8, 0.9, 1\}$, $o = \frac{R_i - \tau_a(D_j)}{B \times B}$ 。

在 Jacquin 提出分形压缩编码法[1]后, 陆续有许多研究者提出改良方法, 其中以 Zhao 和 Yuan 提出的以二次多项式取代文献[11]中的一次多项式最具代表性, 其形式为: $Z = g(z) = k_1 z + k_2 z^2 + o$, 文献[11]虽证明了这方式可使误差最佳化, 但这方法仍无法保证不动点唯一存在[6], 由于其所得的归结图可能因初始图像而异, 即其一可能为误差最小的图像(与原图像最相似), 另一图像则为误差最大的图像(与原图像最不相似), 这将使得文献[11]的分形编码算法无法满足收敛性, 也无法符合良好图像数

据库索引文件的条件, 因此在使用分形编码建立图像数据库索引文件时, 应先检查其是否符合定理 1 的条件, 只有在条件成立时, 方可确定建立的索引文件具有良好图像数据库索引文件的条件, 也才能确保由数据库中检索到的图像是相似的。

为求更佳的编码效果, 可参考文献[11]的作法将迭代函数 g 改为二次表示, 即

$Z = g(z) = k_1 z + k_2 z^2 + o$, 其误差 e 为

$$e = \sum_{i=1}^K [Z_i - g(z_i)]^2 = \sum_{i=1}^K [Z_i - k_1 z_i - k_2 z_i^2 - o]^2$$

可对其进行微分, 使误差最小化:

$$\begin{cases} \frac{\partial e}{\partial k_1} = -2 \sum_{i=1}^K (Z_i - k_1 z_i - k_2 z_i^2 - o) z_i = 0 \\ \frac{\partial e}{\partial k_2} = -2 \sum_{i=1}^K (Z_i - k_1 z_i - k_2 z_i^2 - o) z_i^2 = 0 \\ \frac{\partial e}{\partial o} = -2 \sum_{i=1}^K (Z_i - k_1 z_i - k_2 z_i^2 - o) = 0 \end{cases}$$

其得到的迭代函数 g 的各项系数为

$$k_1 = \frac{(\bar{Z}\bar{z} - \bar{Z}\bar{z})(\bar{z}^2)^2 - \bar{z}^4) - (\bar{Z}\bar{z}^2 - \bar{z}^2\bar{Z})(\bar{z}^2\bar{z} - \bar{z}^3)}{(\bar{z}^2\bar{z} - \bar{z}^3)^2 - ((\bar{z}^2)^2 - \bar{z}^2)(\bar{z}^2)^2 - (\bar{z}^2)^2}$$

$$k_2 = \frac{(\bar{Z}\bar{z} - \bar{Z}\bar{z}) + k_1((\bar{z}^2)^2 - \bar{z}^2)}{\bar{z}^2\bar{z} - \bar{z}^3}$$

$$o = \bar{Z} - k_1\bar{z} - k_2\bar{z}^2$$

其中, $\bar{Z} = \frac{1}{K} \sum_{i=1}^K Z_i$, $\bar{z} = \frac{1}{K} \sum_{i=1}^K z_i$, $\bar{z}^2 = \frac{1}{K} \sum_{i=1}^K z_i^2$,

$\bar{z}^3 = \frac{1}{K} \sum_{i=1}^K z_i^3$, $\bar{z}^4 = \frac{1}{K} \sum_{i=1}^K z_i^4$, $\bar{Z}\bar{z} = C \frac{1}{K} \sum_{i=1}^K Z_i z_i$,

$\bar{Z}\bar{z}^2 = \frac{1}{K} \sum_{i=1}^K Z_i z_i^2$ 。

以二次多项式代替一次多项式虽可由上述证明其有最小误差, 但二次多项式将造成两个不动点[6], 其一为误差最佳化(最小)后的不动点, 而另一个不动点却非所要, 所以使用文献[11]算法编码后, 若其误差较文献[1]方法大, 则选择文献[1]的编码结果。图 4 就是依此思想对电影“卧虎藏龙”一个镜头进行迭代的结果(PSNR 为 27.53dB, 压缩比为 8.68)。

3 实验评估方法及结果

本文提出新的图像检索方法, 就是以分形迭代函数来进行图像搜索, 即先以分形迭代函数建立图像数据库索引, 再以此索引为特征, 配合其他分形迭代函数包含的信息, 并将其作为搜索键值来求出相似图像。

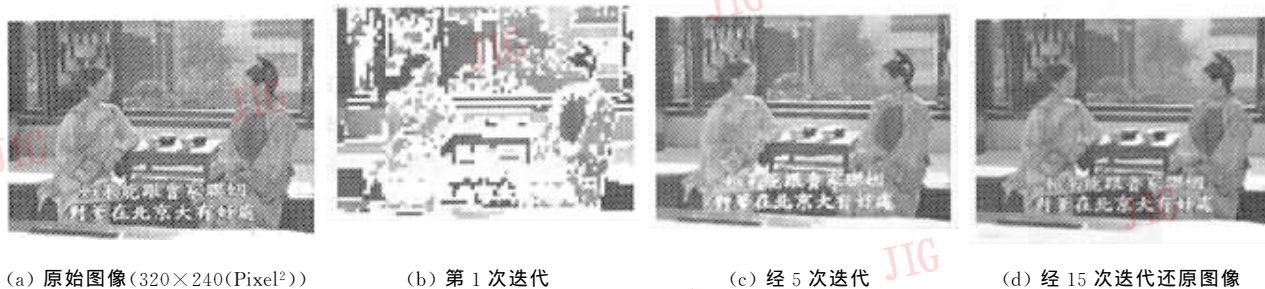


图4 图像的分形迭代结果

分形图像数据库查询模型一般分为两大部份：第1部份为编码部份，即使用分形理论编码，对将存入数据库的图像进行编码；第2部份为比对部份，即从数据库中找出与搜索图像(查询图像)相似的图像。本文选择电影“卧虎藏龙”来建立图像数据库。

图5为本方法检索的结果，其中 S 为相似度，用相似距离度量，检索实验是在Linux平台上进行，实验数据为1000幅(编号为0~999)图像的数据库，图5左上方一幅为输入的搜索键值图像(查询图像)。



图5 本分形方法的检索结果

表2 查询图像与不同亮度的同一图像的相似度

图像亮度比例(%)	-100	-60	-20	+20	+60	+100
查询结果的相似度	0.16	0.45	0.73	0.72	0.50	0.24

4 结论

以往分形编码被广泛使用在图像压缩研究上，本文则将分形编码运用于图像检索上，并证明其具

由图5检索结果可知，相似距离越小的图像，在整体上皆与输入图像越相似。另外，改变查询图像的亮度(如表2所示)，其检索结果也表明，依分形迭代函数系统的拓扑特性来检索图像，图像特征的改变(如亮度、颜色)对检索结果的影响是不大的，这也说明，此方法的适应性好，因为其特征索引是依索引文件进行的。用户进行数据库搜索后，可再依其特征相似程度来排列信息，以使用户从中选出需要的资料。

有成为良好图像数据库索引文件的性质。

这样，不仅可确保以分形编码方法建立的索引文件能取得相似图像，并且可避免找出不相似图像的情况产生。分形编码本身具有成为良好索引文件的性质，而这性质却是其他方法先天的缺点，因此，未来研究可放在使分形编码所得到的函数与人类直觉更接近，并用以增大图像数据库等方面。由于以分形迭代函数系统为基础的编码算法，其根据分形的拓扑几何特性来研究图像的检索是很有意义的，因

此可望建立一套确保搜索结果与人类直觉相符的图像检索系统。

参 考 文 献

- 1 Jacquin A E. A fractal theory of iterated markov operators with applications to digital image coding [D]. Thesis, Georgia Institute of Technology, Atlanta, Georgia, USA, 1989.
- 2 Fisher Yuval. Fractal image compression: Theory and application[M]. New York:Springer-Verlag, 1996.
- 3 Faloutsos C, Flickner M, Niblack W, *et al.* Efficient and effective querying by image content[R]. Research Report #RJ 9203 (81511), IBM Almanden Research Center, San Jose, CA, USA, Aug. 1993.
- 4 Lai Z, Jim C, Hsu Fu-Te. Image retrieval using semantic classification and partial match [A]. In: The 13th IPPR Conference on Computer and Vision, Graphics and Image Processing[C], Taipei China, 2000,1-6.
- 5 Xu Hui, Liao Mengyang. Cluster-based texture matching for image retrieval[A]. In: Proceedings. International Conference on Image Processing[C], 1998:766~769.
- 6 Alan Beardon F. Iteration of retional functions[M]. New York: Springer-Verlag, 1991.
- 7 Dunford N, Schwartz J T. Linear operators[M]. New York: Wiley, 1966.
- 8 Ortega J M, Rheinboldt W C. Iterative solution of nonlinear equations in several variables [M]. San Diego, CA, USA: Academic Press, 1970.

- 9 Jacquin A E. Image coding base on a fractal theory of iterated contractive image transformations [J]. IEEE Transactions on Image Processing, 1992,1(1):18~30.
- 10 Jacquin A E. Fractal image coding: A review[J]. Proceeding of the IEEE, 1993,81(10):1451~1465.
- 11 Zhao Yao, Yuan Baozong. A new affine transformation: Its theory and application to image coding[J]. IEEE Transactions on Circuits and System for Video Technology, 1998,8(3):269~274.



陈添丁 2002 年获浙江大学信息与通信工程博士学位,现为杭州商学院讲师。主要研究领域是多媒体通信、无线通信与安全、图像处理、图像(视频)检索。



刘济林 1947 年生,博士生导师。主要研究领域是图像处理、机器人视觉与导航、ITS、MPEG 系列及视频检索。

楼伟进 男,1972 年生,2003 年获浙江大学计算机系计算机应用专业博士学位,现为浙江省移动通信有限公司高级工程师。