快速分形图像编码的一种特征方法

何传江[1] ,蒋海军[2] ,黄席樾[2]
(1. 重庆大学数理学院;2. 重庆大学自动化学院,重庆 400044)

摘 要: 快速分形图像编码的特征向量法是最具创新性、最有前途的方法之一,但它有几个缺点、特别是特征向量的高维数性. 针对这个问题,本文提出减少分形编码时间的一种可选的特征方法. 作为它的应用,本文先定义图像块的新特征 ——叉迹,然后提出一个基于叉迹的快速分形算法. 这个算法把 Range-Domain 子块匹配问题转化为叉迹意义下的邻域搜索问题. 对 256 ×256 Lena 图像的实验显示,与基于全搜索的基本分形算法比较,依赖于搜索邻域大小,该算法既能在峰值信噪比相同的情况下实现加快 3 倍多,也能在主观质量有一定下降的成本下实现加快 100 倍以上.

关键词: 分形图像编码; 图像压缩; 叉迹

中图分类号: TN919.81 **文献标识码**: A **文章编号**: 0372-2112 (2004) 11-1864-04

A Feature Method for Fast Fractal Image Encoding

HE Chuan-jiang[1] ,JIANG Hai-jun[2] ,HUANG Xi-yue[2]
(*1. College of Mathematics and Science;2. College of Automation, Chongqing University, Chongqing* 400044, *China*)

Abstract: Feature vector method for fast fractal image encoding is considered as one of the most innovative and promising approaches,but it suffers from several drawbacks,especially high dimensionality of feature vectors. Thus,an alternative feature method to reduce fractal encoding time is proposed. As one of its applications,cross trace-based fast fractal algorithm is presented,where the cross trace is a newly-defined feature of an image block. The proposed algorithm converts the range-domain block matching problem to the neighborhood search problem in the sense of cross trace. A simulation on popular 256 ×256 Lena image shows that,depending on the search window size,the proposed algorithm not only can achieve the speed-up of over 3 times with the same PSNR (peak signal-to-noise ratio) as the baseline fractal algorithm with the full search,but also can obtain the speed-up of 100 times or more at the cost of tolerable degradation of the decoded image quality.

Key words: fractal image coding;image compression;cross trace

1 引言

图像是自然界某个片断的真实再现,通常具有某种分形特征,特别是不同尺度下的局部自相似性(可以看成一种结构相关). 分形编码就是利用这种局部自相似性来消除图像数据中存在的跨尺度冗余的新一代压缩技术. 十余年来,分形编码以新颖的思想、高压缩比、分辨率无关性、解码快等优点受到国内外广泛关注[1,2].

在分形编码中,图像由揭示局部自相似性、并行作用使图像近似不变的一组仿射映射表示(编码),然后由分形码描述的压缩变换迭代作用于任意初始图像来重构(解码). 图像的分形码是对这些仿射映射构成的压缩变换的描述(与使用 DCT 的频域描述截然不同). 压缩映象原理指出,图像空间(完备度量空间)中的压缩变换存在唯一不动图像,而且可以由这个变换迭代作用于任何初始图像来生成. 于是,我们可以用压缩变换表示这幅不动图像. 此外,拼贴定理表明,使图像近似不变的压缩变换的不动图像是原始图像的近似,于是可用这个压缩变换近似表示待编码图像. 这就是分形编码的创新性编码思想.

然而,如何寻求这样的压缩变换? 这是分形编码最基本的问题. Barnsley 提出第一个方法,并申请了美国专利. 但这个方法不太实用,因为编码时间太长,需要借助专门的硬件设备和人机交互. 后来,Jacquin 提出了一个相对实用的方法,实现了计算机自动编码过程,即分形块编码(fractal block coding)[3]. 这种编码方法,其计算复杂性仍然很高.

目前,人们已提出了许多加快编码的方法[1,2],新的方案仍在不断出现[4~6]. 总体上,这些方法可以分成两类:"无损"加快与有损加快. "无损"加快是相对基本分形算法(全搜索)而言保持图像质量不变的一种加快编码方法,缺点是加快的倍数不高(通常 2 倍左右[7,8]). 有损加快是一种牺牲图像质量以换取编码时间减少的编码方法. 因为分形编码是一种基于拼贴定理的编码方法,这种方法本身是有损的,所以,有损

收稿日期:2003-06-09;修回日期:2004-04-26

加快方法得到了人们更多的关注,成为快速分形编码研究的主流.

在快速分形编码方法中,较为有效的有分类法和特征向量法(feature vector method)[9,10]. 分类法是一种按灰度变化情况对码本进行某种分类,每个 R 块只在同类的码本子类中搜索最好匹配块. 特征向量法是把匹配搜索转换为特征向量的最近邻搜索. 文献[10]对分类法和特征向量法作了对比研究,结果显示特征向量法比分类法更有效. 但是,特征向量法有几个公认的缺点,特别是特征向量的高维问题. 此外,它要求的字体太强(详见第 2.2 节),针对这些问题,本文提出一个可行的解决方案.

论文具体安排如下:第 2 节介绍特征向量法的基本思想,阐明本文方法的原理;第 3 节简介基本分形算法;第 4 节给出本文方法的理论依据与算法描述;第 5 节给出实验结果和结论.

2　特征向量法与本文算法原理

2.1　特征向量法

特征向量法是快速分形编码最具创新性、最有前途的方法之一(Saupe 提出)[9,10]. 首先构造子块 X 的特征向量 $\phi(X)$ (即 X 的规范化,详见第 4 节),然后证明子块 R 与 D 的亮度仿射变换之间的 MSE(Mean Square Error)的极小化问题等价于

$$d(R, D) = \min\{\ \|\phi(D) + \phi(R)\|,\ \|\phi(D) - \phi(R)\|\} \tag{1}$$

的极小化问题. 其中,$\|\cdot\|$是向量 2-范数. 于是,子块 R、D 的匹配问题转换为特征向量的最近邻搜索问题. 在计算机科学中,欧氏空间的最近邻搜索问题得到了较为透彻的研究,已有许多富有成效的算法. 特征向量法采用了其中著名的 kd-tree 最近邻搜索法.

特征向量法有几个缺点[9]. 最大的缺点是子块 R、D 的特征向量的维数很高,这样就产生特征向量的存储问题. 此外,高维时 kd-tree 最近邻搜索法的性能并不好(维数越高,性能越差). Saupe 提出的解决办法有两个:一是通过象素平均或间距采样(decimation)减少 R、D 的维数;二是量化 R、D 特征向量的各分量(8b 足够). 这样处理的代价是搜索到的特征向量 $\phi(R)$ 的最近邻只能是近似的. 于是,Saupe 提出在搜索到的特征向量 $\phi(R)$ 的最近邻附近再搜索 5～10 个近邻,从中选择 MSE 最小者.

2.2　本文算法原理

特征向量法的原理,实际上是通过构造图像块适当的特征向量,使得特征向量接近(某个度量 d 下)的 R 和 D,在 MSE 度量下也接近,反之亦然. 具体说,

$$d(F_R, F_D) \le \delta \Leftrightarrow \mathrm{MSE}(R, D) \le \varepsilon \tag{2}$$

其中,F_X 表示图像块 X 的特征向量,$\delta > 0$、$\varepsilon > 0$ 是适当阈值. 显然,如何选取特征向量以及特征向量空间中的搜索方法是特征向量法的关键.

关于上述特征向量法,我们认为,能够找到满足(2)的特征向量固然不错,但还有许多好的图像特征(或特征向量)无法满足如此强的条件. 为此,放宽条件(2)如下

$$d(F_R, F_D) > \delta \Rightarrow \mathrm{MSE}(R, D) > \varepsilon \tag{3}$$

或等价地(逆否),

$$\mathrm{MSE}(R, D) \le \varepsilon \Rightarrow d(F_R, F_D) \le \delta \tag{4}$$

(3)的含义是,特征向量在度量 d 下不接近的 R 和 D,在 MSE 度量下也不接近. (4)的含义是,在 MSE 度量下接近的 R 和 D,其特征向量在度量 d 下也接近. 换言之,R 和 D 的特征向量接近是 R 和 D 在 MSE 度量下接近的必要条件. 因此,在 MSE 度量下匹配的 R 和 D,在特征空间中一定是近邻(度量 d 意义下).

为了实现上述想法,本文提出一种快速分形编码方法. 具体做法如下:首先定义叉迹的概念(作为图像块的一种特征),并证明一个联系 MSE 与叉迹的不等式,使得(3)成立. 然后根据(3)把匹配度量从 MSE 转化为叉迹,尽管这种方法得到的 D 并不能保证是 R 在 MSE 意义下的最佳匹配块,但这样的最佳匹配块在叉迹意义下距离 D 也不能太远. 鉴于此我们把码本按叉迹大小赋予序结构,然后在赋序码本中使用二分法对 D 的 k 邻域进行搜索.

3　基本分形算法

本文讨论基本分形编码,并从矢量量化(VQ)编码的观点叙述其算法. 图像被分割成大小两类子块:range 块(R 块)和 domain 块(D 块),其中 R 块互不重叠且覆盖整幅图像,D 块可以重叠且边长为 R 块的两倍. D 块经四邻域象素平均收缩为 R 块的大小,这种子块的全体构成码本(记为 Ω). 然后每个 R 块 R 由其最佳匹配块 $D \in \Omega$ 的亮度变换版本来近似,即 $R \approx s \cdot D + o \cdot \mathbf{1}$,其中$\mathbf{1}$是亮度值均为1的常值块,$s$、$o$ 分别调整 D 的对比度和亮度. 此外,为了改进图像质量,一般还要对码本块 D 进行 8 个等距变换.

编码阶段,对于每个 R 块 R_i,为了寻求其最佳匹配块 $D \in \Omega$,需要求解下面的极小化问题:

$$\|R_i - (s_i \cdot D_{m(i)} + o_i \cdot \mathbf{1})\| = \min_{\substack{s,o \in R \\ |s| < 1, j}} \|R_i - (s \cdot D_j + o \cdot \mathbf{1})\| \tag{5}$$

其中 $\|\cdot\|$是向量 2-范数,$m(i)$、s_i 和 o_i 分别表示 R_i 的最佳匹配块 $D_{m(i)} \in \Omega$ 的序号和该匹配块对比度和亮度的最优调整因子. 此外,对比度因子要求满足约束 $|s| < 1$ 是为了理论上保证解码迭代序列收敛. 三元组$(m(i), \bar{s}_i, \bar{o}_i)$称为 R_i 的分形码,其中 $\bar{s}_i$、$\bar{o}_i$ 是量化值. 此外,如果考虑 8 个等距变换,则分形码中还应包括等距变换序号. 全体 R 块的分形码就组成原始图像的分形码,它描述了一个使图像近似不变的压缩仿射变换.

解码是相对简单的迭代过程,由压缩映象原理给出的迭代算法完成,即分形码描述的压缩变换迭代作用于任何初始图像来生成. 具体地说,分形解码按下面的方式进行(未考虑等距变换):

$$R_i^{(k)} = s_i \cdot D_{m(i)}^{(k-1)} + o_i \cdot \mathbf{1},\ \mu^{(k)} = \cup_i R_i^{(k)} \tag{6}$$

其中,$R_i^{(k)}$ 是第 k 次迭代图像 $\mu^{(k)}$ 的 R 块,$D_{m(i)}^{(k-1)}$ 是源于第($k-1$)次迭代图像 $\mu^{(k-1)}$ 的码块($m(i)$ 指明 $D_{m(i)}^{(k-1)}$ 是 $R_i^{(k)}$ 的最佳匹配块). 不难看出,分形解码中每次迭代图像所需要的码本都是由上一次迭代图像提供的,而第一次迭代图像所需

要的码本则由初始图像提供. 这是分形编码不同于 VQ 编码的地方. 此外,迭代解码也是分形编码的新颖之处.

4 快速算法

首先定义图像块的叉迹,给出联系均方误差与叉迹的一个不等式,然后给出本文算法的描述.

定义 1 子块 $X = (x_{i,j})_{B\times B}$ 的规范化定义为

$$\hat{X} = (X - \bar{X}\cdot 1)/\|X - \bar{X}\cdot 1\| \tag{7}$$

其中, $\bar{X}$ 是 X 的均值.

显然, $\langle \hat{X}, \mathbf{1}\rangle = 0$ 且 $\langle \hat{X}, \hat{X}\rangle = \|\hat{X}\|^2 = 1$, 这里, $\langle\cdot,\cdot\rangle$ 表示向量内积. 顺便指出,图像块 X 的规范化就是特征向量法中的特征向量.

定义 2 子块 $X = (x_{i,j})_{B\times B}$ 的叉迹(cross trace)定义为

$$C_{tr}(X) = \sum_{i=1}^{B}(\hat{x}_{i,i} + \hat{x}_{i,B-i+1}),\quad \hat{x}_{i,j} = (\hat{X})_{i,j} \tag{8}$$

也就是说,图像块 X 的叉迹就是其规范化的块主次对角象点亮度之和. 主次对角线形成"十字叉",又矩阵的主对角元之和称为迹(trace),这就是"叉迹"名称的由来.

设 C 是一个主次对角元为 1、其余元为 0 的 $B\times B$ 矩阵,根据叉迹定义,显然有 $C_{tr}(X) = \langle \hat{X}, C\rangle$.

定理 设 $R, D \in R^n (n = B\times B)$ 及

$$E(R,D) = \min_{a,b\in R}\|R - a\cdot D - b\cdot 1\| \tag{9}$$

则下面的不等式成立

$$E(R,D) \ge \frac{\sigma_R}{4}\left||C_{tr}(R)|^2 - |C_{tr}(D)|^2\right| \tag{10}$$

其中, $\sigma_R = \|R - \bar{R}\|/B$.

证明可以在文献[11, p. 96 - 98]中找到.

不等式(10)表明,只有对于非平滑块 R(即 σ_R 大于预定阈值),(3)和(4)才成立. 但是,如果 σ_R 很小,则说明 R 的亮度变化不大,通常是图像中的平滑部分可以近似视为常值块,并用 $\bar{R}\cdot 1$ 近似之. 因此,可以事先排除掉平滑块 R,从而避免了上述问题. Jacquin 在其奠基算法[3]中就采用了这种源于 VQ 编码的策略.

基于上述分析,本文算法的具体步骤如下:

步骤 1 设定平滑性阈值 τ.

步骤 2 分割图像为 $B\times B$ 固定块,记为 R(R 块). 以纵横方向步长均为 s 生成尺寸为 $2B\times 2B$ 的 D 块池,即相邻 D 块间在水平方向或垂直方向有 s 个像素的重叠. 对于每个 D 块,采用 4-邻域像素值平均得到 $B\times B$ 块,这样的子块集合就构成码本 Ω.

步骤 3 计算每个子块 R 和 $D\in\Omega$ 的叉迹,以及 $\sigma(R)$ 和 $\sigma(D)$. 按 $|C_{tr}(D)|$ 大小对码本 Ω 进行升序排列.

步骤 4 对于每个 R_i.

(1) 如果 $\sigma_{R_i} < \tau$, 则直接用 $\bar{R}_i\cdot 1$ 近似之;

(2) 如果 $\sigma_{R_i} \ge \tau$, 则按二分法在赋序码本 Ω 中搜索叉迹意义下的最佳匹配块 $D_{m(i)}$.

步骤 5 在赋序码本 Ω 中定义以匹配块 $D_{m(i)}$ 为中心的 k-邻域,在其中选择与 R_i 有最小 $E(R,D)$ 的码本块 D_m. 然后考虑 8 个等距变换从中选择最佳等距变换序号 l. 序号 m 和 l、参数 s 和 o 就组成 R 的分形码 $(m, l, \bar{s}, \bar{o})$. 其中, $\bar{s}, \bar{o}$ 是量化值(均匀量化).

步骤 6 对于每个 R 块 R,重复以上步骤直至所有 R 块.

算法说明:

(1) 参数 s 和 o 由最小二乘法得到,即

$$s = \langle R - \bar{R}\cdot 1, D - \bar{D}\cdot 1\rangle / \|D - \bar{D}\cdot 1\|^2$$
$$o = \bar{R} - s\cdot\bar{D} \tag{11}$$

(2) 上面的 s 未必满足约束 $|s| < 1$, 因此需要作截断处理. 实验表明[11],综合考虑压缩比与图像质量,参数 s、o 分别按 5 b 和 7 b 量化可得到最佳效果. 因此,我们的截断方案如下:若 $s > 1$, 取 $s = 31/32$, 若 $s < 0$, 取 $s = 0$.

(3) 为了独立显示本方案的效果,本文没有采用已提出的大量分形编码策略. 当然,在本文算法的基础上,融入这些编码策略将获得更好的图像质量和更快的编码速度.

5 实验结果与结论

我们选择 256 ×256Lena 测试图像(8b 量化)为实验对象. 计算在 IBM ThinkPad(700 MHz CPU、内存 128M)上进行,程序用 C++编写. 测试性能参数是编码时间(秒,源于系统时钟)和 PSNR(dB,峰值信噪比). 在实验中,我们选取 R 块大小为 4 ×4,生成 D 块池的滑窗步长 s 选取为 8,参数 s、o 分别按 5b、7b 量化.

本文算法依赖于两个控制参数 k 和 τ. 实验表明,当 $\tau\ge 5$ 时,Lena 的肩膀等平滑部分出现"块效应", τ 值越大,块效应越明显(图 1). 但 $\tau\le 4$ 时块效应消失,因此,我们选取 $\tau = 4$ 作为缺省值.

图 1 $\tau\ge 5$ 时出现"块效应"($k = 100$). 左图: $\tau = 5$; 右图: $\tau = 6$

表 1 给出了基本算法与本文算法的对比实验结果. 可以看出,如果邻域大小 $k = 350$, 则本文算法实现无损加快 3.1 倍. 作为比较,我们引述文献[7,8]算法的实验结果:对于 256 ×256 Lena 图像,按 4 ×4 像素块分割,文献[7]实现无损加快 2.32 倍;文献[8]实现无损加快 0.56 倍. 可见,文献[7,8]算法加快编码至多 2 倍左右. 本文算法的编码时间由邻域阈值 k 控制,是可调的. 此外,与文献[7,8]不同,在实时性要求高、质量次之的应用场合,本文算法可选取较小的 k 值,以获得较高的编码速度(图 2). 此外,本文算法理论分析简单(不等式的证明较为复杂),实现容易,而算法[7,8]都需要复杂的理论分

图 2 基本算法(左)与本文算法(右, $\tau = 4$, $k = 5$)的解码图像对比. 本文算法加快 140 倍(0.68 秒),图像质量仍然可以接受.

析,实现也很复杂.

因此,综合考虑质量、时间和算法复杂性,本文算法为加快分形编码提供了一个好的候选.在实时性要求高、质量次之的场合,本文算法有较好的应用前景.此外,本文定义的叉迹也许在其它应用中有用.

表 1　实验数据($\tau=4$,BFC 表示基本分形编码)

邻域 k	50	100	150	200	250	300	350	*BFC*
PSNR	27.52	27.98	28.22	28.38	28.48	28.6	28.66	28.66
下降	1.14	0.68	0.42	0.28	0.18	0.06	0	-
时间	5.22	10.5	14.6	19.78	23.26	27.17	30.9	95.6
加快	18.3	9.1	6.55	4.83	4.11	3.52	3.1	-

参考文献:

[1] B Wohlberg, G Jager. A review of the fractal image coding literature [J]. IEEE Trans. Image Process. 1999, 8(12): 1716 - 1729.

[2] 赵耀,王红星,袁保宗. 分形图像编码研究的进展[J]. 电子学报, 2000, 28(4): 95 - 101.

[3] A E Jacquin. Image coding based on a fractal theory of iterated contractive image transformations [J]. IEEE Trans. Image Process. 1992, 1(1): 18 - 30.

[4] 谭郁松,周兴铭. 一种新型图像分形压缩的改进算法[J]. 电子学报, 2003, 31(11): 1739 - 1742.

[5] C M Lai, K M Lam, W C Siu. A fast fractal image coding based on kick-out and zero contrast conditions[J]. IEEE Trans. Image Process. 2003, 12(11): 1398 - 1403.

[6] C He (何传江), S X Yang, X Huang. Variances-based accelerating scheme for fractal image encoding[J]. IEE Electronics Letters. 2004, 40(2): 115 - 116.

[7] J H Jeng, T K Truong, J R Sheu. Fast fractal image compression using the Hadamard transform [J]. IEE Proc.-Vis. Image Signal Process. 2000, 147(6): 571 - 573.

[8] H Hartenstein, D Saupe. Lossless acceleration of fractal image encoding via the fast Fourier transform[J]. Signal Processing: Image Communication, 2000, 16(4): 383 - 394.

[9] C S Tong, M Wong. Adaptive approximate nearest neighbor search for fractal image compression[J]. IEEE Trans. Image Process, 2002, 11(6): 605 - 615.

[10] M Polvere, M Nappi. Speed-up in fractal image coding: comparison of methods[J]. IEEE Trans. Image Process, 2000, 9(6): 1002 - 1009.

[11] 何传江. 分形图像编码技术的算法研究[D]. 重庆:重庆大学博士学位论文, 2004.

作者简介:

何传江　男,1964 年生于贵州遵义,先后于 1985 年、1988 年获得四川大学数学学士和硕士学位,2004 年获得重庆大学自动化学院博士学位,现为重庆大学数理学院数学系副教授、硕士导师,主要研究领域为分形、小波和图像处理等. Email: chuanjianghe@sina.com.

蒋海军　男,1975 年生于重庆江津,1998 年获得中国矿业大学工学学士学位,现为重庆大学自动化学院硕士研究生,主要研究领域为分形编码和图像处理等. Email: jhjcqu@163.com.

黄席樾　男,1943 年出生于重庆奉节,1967 年毕业于四川大学无线电系,1987 年获得日本东北大学博士学位,现为重庆大学自动化学院教授,博士生导师,主要研究领域为智能控制,机器视觉等.