

基于分形编码的图像相似匹配研究

洪安祥, 陈 刚, 吴炯锋, 李均利

(浙江大学计算机图像图形研究所, 杭州玉泉 310027)

摘 要: 本文以自适应四叉树分割的分形图像编码为基础, 研究了极坐标表达下分形码之间的匹配, 并在此基础上提出了图像间分形码峰值信噪比的概念和算法, 以达到量化图像间相似性的目的. 实验证明, 这种基于自适应分形编码的图像相似性匹配方法和人的主观判断非常接近, 非常适合于用来构建和管理大型图像数据库.

关键词: 分形编码; 图像匹配; 分形码峰值信噪比

中图分类号: TN911.21 **文献标识码:** A **文章编号:** 0372-2112(2002)05-0624-04

Image Matching Based on Fractal Image Coding

HONG An-xiang, CHEN Gang, WU Jiong-feng, LI Jun-li

(Institute of Imaging and Computer Graphics, Zhejiang University, Hangzhou, Zhejiang 310027, China)

Abstract: In adaptive quadtree fractal image encoding scheme, a polar coordinates fractal-based coding algorithm is presented. The concept and algorithm of peak signal noise ratio (PSNR) based on the fractal codes is put forward and is used to match the similarities between images. Experiments show that the matching scheme in this paper is much close to human perception. And this technology is very suitable to construct and to manage a huge image database.

Key words: fractal coding; image matching; PSNR based on fractal codes

1 引言

近年来, 用分形方法进行图像处理, 诸如图像编码、区域分割、图像检索等的研究日益广泛和深入. 本文研究用分形码存储的图像库中图像间相似性的判别及检索方法.

A. D. Sloan 最早把分形方法用于图像检索, 他提出了合并编码的方法. 对一幅图标图像, 要寻找图像库中与之相似的图片, 采用合并编码, 把图标图像(Icon)和图像库中的图标图像直接拼凑起来进行编码. 对于任何图像的 Range 块可以在任意的图中寻找块来匹配, 而两幅图之间的相似性则以一个图像中的子块作为另一个图像中 Range 块的 Domain 块的数目来度量. 虽然 A. D. Sloan 用合并编码的方法进行图像相似匹配是可行的, 但由于分形编码费时及分形码不稳定的特点使得这种方法离实际应用还有很大距离.

后来, A. D. Zhang 等提出基于九叉树的检索匹配策略. 由于九叉树结构的优点及采取了固定 Range 块大小并作像素裁剪的方法, 使得合并编码的缺点得以较大改善, 而且 A. D. Zhang 等也提出了度量相似性的模型, 使图像间的相似性有了量化方法. 但由于固定 Range 块大小而使记录的分形码未很好地表征图像特征, 对图像的裁剪又要花费很多计算时间, 而且他们的相似性度量模型里有许多参数要人为设定, 没有确定的标准, 最终使得该方法在实际应用上有些不适应.

本文在自适应四叉树分割的分形编码基础上对图像块进

行匹配, 不仅有简单明了的匹配算法, 弥补了 A. D. Zhang 等只能对固定 Range 块大小进行处理以及要进行裁剪而消耗大量时间等方面的不足, 而且借鉴图像峰值信噪比的模型给出了分形码匹配的峰值信噪比模型, 很好地量化了图像间的相似性.

2 分形码结构

假设读者已有分形图像编码的基础知识, 那么用块方法进行四叉树分割编码, 经简化得到的码字结构应该包括以下信息:

- (1) Domain 块位置信息
- (2) 四叉树分割层次信息
- (3) 亮度平移值 t_i

Domain 块位置采用相对极坐标(r, θ)的方式记录. 即对直角坐标为(i, j)位置的 Range 块, 若它对应的 Domain 块位置像素坐标为(x, y)时, 以(i, j)为原点, 求得(x, y)点的极坐标(r, θ)如下:

$$\begin{cases} r = \sqrt{(x-i)^2 + (y-j)^2} \\ \theta = \arctg \frac{y-j}{x-i} \end{cases}$$

容易知道, 在保证足够精度的前提下, 极坐标表示和普通表示是等价的. 以(r, θ)来标记 Domain 块的位置.

对于四叉树层次信息, 引进分形码权的概念.

我们觉得大小不同的块对应的分形码的重要性应该有些不同. 当把图像作为一个整体, 和另一个图像作相似性比较时, 大多数情况下更关注大区域的状况, 即块越大越体现图像的整体信息. 因此, 定义分形码的权为它所对应的 Range 块占图像总面积的比例.

定义 1 对大小为 $M \times N$ 的离散数字图像进行分形编码, 设当前 Range 块的大小为 $r \times r$, 则定义该 Range 块对应的分形码的权为

$$w = r \times r / M \times N$$

注意到, 分形码的权和它所在的四叉树层次是相对应的. 若以原图像为四叉树的根, 则第一层(根节点算第 0 层)节点对应的权为 $1/4$, 下一层为 $1/16$, 也就是说此时, 若一个 Range 块对应于四叉树的第 i 层, 则它的分形码的权 w 为 $1/4^i$.

至此定义我们的分形码结构为: $\{r_i, \theta_i, t_i, w_i\}$, 其中 r_i, θ_i 为对应 Range 块的极坐标, t_i 为亮度平移值, w_i 为权. 于是一幅图的分形码为:

$$\begin{pmatrix} (r_1, \theta_1, t_1, w_1) \\ \dots\dots\dots \\ (r_i, \theta_i, t_i, w_i) \\ \dots\dots\dots \\ (r_K, \theta_K, t_K, w_K) \end{pmatrix}$$

3 匹配算法

对两幅图 X, Y 和它们的分形码, 不妨记为:

$$X = \begin{pmatrix} (r_1, \theta_1, t_1, w_1) \\ \dots\dots\dots \\ (r_i, \theta_i, t_i, w_i) \\ \dots\dots\dots \\ (r_K, \theta_K, t_K, w_K) \end{pmatrix} \quad Y = \begin{pmatrix} (r'_1, \theta'_1, t'_1, w'_1) \\ \dots\dots\dots \\ (r'_j, \theta'_j, t'_j, w'_j) \\ \dots\dots\dots \\ (r'_L, \theta'_L, t'_L, w'_L) \end{pmatrix}$$

其中 X 有 K 个分形码, Y 有 L 个分形码. 在介绍匹配算法前, 先了解这样一个基本性质:

如图 1, 当 Range 块 R 对应的匹配是 Domain 块 D 时, 看它们四叉树剖分后 $R1, R2, R3, R4$ 和 $D1, D2, D3, D4$ 的对应关系. 由图易知, 若 $R1, R2, R3, R4$ 分别以 $D1, D2, D3, D4$ 为相应的 Domain 块时, 它们的分形码和 R 对应 D 的分形码是一样的. 另一方面, 一般情况下, 当 R 和 D 的匹配误差小于阈值时, 小块间的匹配误差也会小于阈值, 只有当误差的分布很不均匀时才有例外(而事实上, 编码过程中, 当误差分布不均匀时, 这样的 R 和 D 匹配应该重新选取). 因此, 在我们的分形码匹配过程中, 用已知大的 R 块的码字来近似它剖分所得小块的

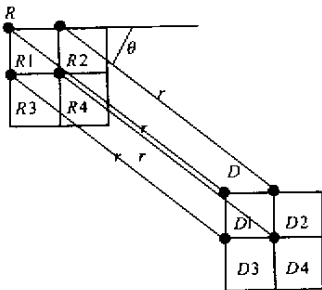


图 1 极坐标分形码对应示意图

码.

下面是 X 和 Y 的分形码匹配算法:

(1) 设定两个分形码的权累加器 $W_X = 0$ 和 $W_Y = 0$;

(2) 从 X 中取出第一条分形码 $(r_1, \theta_1, t_1, w_1)$, 从 Y 中取出第一条分形码 $(r'_1, \theta'_1, t'_1, w'_1)$, 则 $W_X = W_X + w_1, W_Y = W_Y + w'_1$. 比较 r_1 和 r'_1, θ_1 和 θ'_1, t_1 和 t'_1 是否分别一致(若一致就加以统计, 统计的方式和内容我们在图像基于分形码的相似性度量中加以说明);

(3) 假设这时已经完成 $\{r_{i-1}, \theta_{i-1}, t_{i-1}, w_{i-1}\}$ 和 $\{r'_{j-1}, \theta'_{j-1}, t'_{j-1}, w'_{j-1}\}$ 的比较, 分不同情况继续进行比较:

- 若 $W_X > W_Y$, 则从 Y 中取出下一条分形码 $(r'_j, \theta'_j, t'_j, w'_j)$ 和 $\{r_{i-1}, \theta_{i-1}, t_{i-1}, w_{i-1}\}$ 作比较且 $W_Y = W_Y + w'_j$;

- 若 $W_X < W_Y$, 则从 X 中取出下一条分形码 $\{r_i, \theta_i, t_i, w_i\}$ 和 $\{r'_{j-1}, \theta'_{j-1}, t'_{j-1}, w'_{j-1}\}$ 作比较且 $W_X = W_X + w_i$;

- 若 $W_X = W_Y < 1$, 则分别从 X 和 Y 中取出下一条分形码 $\{r_i, \theta_i, t_i, w_i\}$ 和 $\{r'_j, \theta'_j, t'_j, w'_j\}$ 作比较, 令 $W_X = W_X + w_i, W_Y = W_Y + w'_j$;

(4) 重复第(3)步一直进行直至 $W_X = W_Y = 1$, 即匹配完毕.

由四叉树结构及分形码权与 Range 块层次信息的关系可以知道, 在第(3)步中, 当 $W_X > W_Y$ 时增加 W_Y (反之增加 W_X), 若干步后总能得到 $W_X = W_Y$.

4 相似性度量

现在我们研究基于分形码的图像相似性度量, 先来观察分形特征码 $\{r_i, \theta_i, t_i, w_i\}$ 的参数含义.

称 r_i 为相似频率参数, 因为 r_i 表示 Domain 块和 Range 块间的相对距离. 注意到 Domain 块的是由 $\|R_i - s_i \tilde{D}_i - t_i\|_2 < \epsilon$ 确定的. 换句话说是根据是否和 Range 块“足够相似”而定. 当 r_i 小时说明 Range 块在自身附近找到了匹配对象, r_i 大则表示只有较远的地方才有和当前 Range 块相似的块. 当一幅图分形码中大部分的 r_i 都较小时说明 Range 块在自身附近找到匹配 Domain 块的机会较大, 即相似块在自身附近出现的频率较高.

称 θ_i 为方向性参数. 由极坐标辐角的意义知 θ_i 表示图像相似块出现在哪个方向. 和 r_i 的意义相似, 若分形码中大部分 θ_i 的取向比较一致时, 图像纹理本身可能就有较大的方向性.

至于亮度平移因子 t_i , 当图像间只是对对比度发生变化时, 它们的分形码 $\{r_i, \theta_i, t_i, w_i\}$ 中, 除了 t_i 不同外, 其余三个参数基本一致. 也就是说 t_i 体现了图像对比度的差异.

最后一个参数分形码的权则粗略的表示了每条分形码的重要性.

每个参数的含义虽各有不同, 但孤立起来看是容易发生偏差的, 因此必须全面考虑. 下面给出一个分形码相似性度量模型.

首先定义两条分形码间的距离. 设分形码 $\{r_i, \theta_i, t_i, w_i\}$ 中 r_i, θ_i, t_i 的重要性参数分别为 a, b, c ($a, b, c \geq 0$), 这三个

参数分别表示 r_i, θ_i, t_i 对相似性评判影响的大小, $\{r_i, \theta_i, t_i, w_i\}$ 和 $\{r_j, \theta_j, t_j, w_j\}$ 间的距离应该有如下基本特征:

- 当 r_i 和 r_j, θ_i 和 θ_j, t_i 和 t_j 三组量对应相等时, 距离为 0;
- 当三组量全不相等时, 距离应该较大;
- 当权 w_i 和 w_j 大小不一致时, 由我们分形码的匹配规则知, 对距离而言较小的那一个更重要些。

于是对分形码 $frac_i = \{r_i, \theta_i, t_i, w_i\}$ 和 $frac_j = \{r_j, \theta_j, t_j, w_j\}$ 定义它们之间的距离如下:

$$d(frac_i, frac_j) = (a \times \text{sig}(\theta_i - \theta_j) + b \times \text{sig}(r_i - r_j) + c \times \text{sig}(t_i - t_j)) \times \min\{\sqrt{w_i}, \sqrt{w_j}\}$$

其中 $\text{sig}(x)$ 为特征函数,

$$\text{sig}(x) = \begin{cases} 1, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

一般地, 虽然 r_i, θ_i, t_i 对图像自相似性特性有各不相同的作用, 但从统计意义上看, 这三个量的地位差不多相等, 故我们取 $a = 1, b = 1, c = 1$, 于是, 分形码间的距离简化成:

$$d(frac_i, frac_j) = (\text{sig}(\theta_i - \theta_j) + \text{sig}(r_i - r_j) + \text{sig}(t_i - t_j)) \times \min\{\sqrt{w_i}, \sqrt{w_j}\}$$

根据分形码匹配算法, 对两幅图的分形码集合 $\{frac_i^1\}_{i=1}^k$ 和 $\{frac_j^2\}_{j=1}^l$ 按下式计算它们的均方误差:

$$MSE = \frac{\sum_i \sum_j d(frac_i^1, frac_j^2)^2}{\sum_k w_k^1 \sum_l w_l^2}$$

注意, MSE 中分子求和时并不是分别对所有 i, j 求和, 其中 i, j 要满足一定的对应关系, 即 i, j 是匹配算法中所指的两条相匹配分形码的下标, 若两条分形码在算法中不进行相互匹配, 则它们也不参与求和。

由于 $\sum_k w_k^1 = 1$ 和 $\sum_l w_l^2 = 1$, 所以

$$MSE = \sum_i \sum_j d(frac_i^1, frac_j^2)^2$$

定义分形码的峰值信噪比 $PSNR_{frac}$ 为:

$$PSNR_{frac} = 10 \log_{10} \frac{(\max_{i,j} \{d(frac_i^1, frac_j^2)\})^2}{MSE}$$

若 $MSE = 0$, 定义 $PSNR_{frac} = \infty$.

由 $d(frac_i, frac_j)$ 的定义 (取 $a = 1, b = 1, c = 1$) 可知, $\max_{i,j} \{d(frac_i^1, frac_j^2)\} \leq 3$, 所以此时可以定义:

$$PSNR_{frac} = 10 \log_{10} \frac{9}{MSE}$$

用分形码的峰值信噪比 $PSNR_{frac}$ 可以度量两组分形码间的相似性. 和普通图像中的峰值信噪比类似, $PSNR_{frac}$ 是个在 0 和 $+\infty$ 间取值的数, 当

- $PSNR_{frac} = 0$ 时, $MSE = 9$, 也就是说此时所有参与匹配的分形码的 r_i, θ_i, t_i 中没有一个量是对应相等的. 我们认为这两幅图完全无关;

- $PSNR_{frac} = \infty$ 时, $MSE = 0$, 所有分形码对应相等, 此时两幅图是完全一样的 (应该是在分形编码误差允许范围内完全一样);

- $PSNR_{frac}$ 为一个正数时, 随着它的增大, 两幅图的相似程度越大. 一般的, 当 $PSNR_{frac}$ 在 30 左右两幅图就很相近了.

5 实验结果

下面看几个例子:



图 2 景色图像, 从左到右分别记这四幅图标为:

Icon1, Icon2, Icon3, Icon4

它们的二叉树编码解码图为:

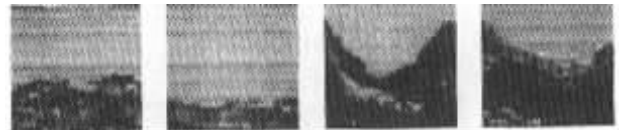


图 3 从左到右分别为 Icon1, Icon2, Icon3, Icon4 的解码图

它们分别有 169, 124, 193, 193 条分形码. 匹配的结果如下表:

表 1 一些分形码的匹配结果

匹配的图标	r_i 一致权重	θ_i 一致权重	t_i 一致权重	r_i 和 θ_i 一致权重	r_i, θ_i, t_i 都一致权重	$PSNR_{frac}$
Icon1 和 Icon2	54.30%	51.95%	55.08%	51.56%	50.78%	32.21
Icon1 和 Icon3	16.41%	14.84%	35.16%	13.67%	9.77%	28.06
Icon1 和 Icon4	16.80%	4.84%	38.67%	13.28%	10.16%	27.93
Icon2 和 Icon3	13.28%	11.33%	35.55%	10.16%	8.59%	27.82
Icon2 和 Icon4	19.14%	14.84%	39.84%	13.67%	11.33%	27.71
Icon3 和 Icon4	34.77%	12.50%	31.25%	12.11%	10.55%	29.12

从 $PSNR_{frac}$ 的值看这四幅图相互之间是比较相似的. 而且 Icon1 和 Icon2 最相似 ($PSNR_{frac}$ 达 32.21), Icon3 和 Icon4 之间也很相似 ($PSNR_{frac}$ 有 29.12), 这和人们视觉判断基本上是一致的.

取四个非常相似的纹理图如图 4:

可算得

$$PSNR_{frac}(\text{Icon5}, \text{Icon6}) = 34.88, PSNR_{frac}(\text{Icon5}, \text{Icon7}) = 39.04, PSNR_{frac}(\text{Icon5}, \text{Icon8}) = 39.05, PSNR_{frac}(\text{Icon6}, \text{Icon7}) = 34.94,$$

$$PSNR_{frac}(\text{Icon6}, \text{Icon8}) = 34.88, PSNR_{frac}(\text{Icon7}, \text{Icon8}) = 39.00,$$

下面我们再取四个不同的图像如图 5:

可得到:

$$PSNR_{frac}(\text{Icon9}, \text{Icon10}) = 2.296, PSNR_{frac}(\text{Icon9}, \text{Icon11}) = 2.525,$$

$$PSNR_{frac}(\text{Icon9}, \text{Icon12}) = 2.485, PSNR_{frac}(\text{Icon10}, \text{Icon11}) = 2.932,$$

$$PSNR_{frac}(\text{Icon10}, \text{Icon12}) = 2.875, PSNR_{frac}(\text{Icon11}, \text{Icon12}) =$$

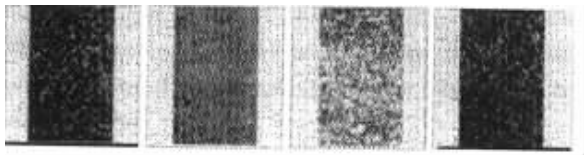


图 4 四种相似的纹理图从左到右依次称
Icon5 ,Icon6 ,Icon7 ,Icon8

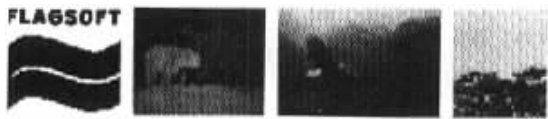


图 5 四种不同的图像从左到右依次称
Icon9 ,Icon10 ,Icon11 ,Icon12

3.31 ,

从分形码峰值信噪比的结果我们可以看出,图 4 中四个纹理图相互之间是非常相似的,而图 5 中的图像之间是不相似的,这和人的视觉判断是相符的。

最后估计一下匹配算法的计算复杂性。

假设两组分形码分别有 m, n 条,不妨设 $m \geq n$,则要做 m 次匹配,每次匹配至多要 6 次加法、3 次比较、2 次乘法,匹配结束后计算 $PSNR_{fract}$ 要进行一次对数运算,所以整个过程的计算量是很小的。用主频为 500MHz 的微机,一般做一次两幅 256×256 大小的图像的分形码匹配只需 0.6 秒左右。

综上所述,基于分形编码的图像相似性量化匹配方法不仅效果好,而且计算复杂性不高,适合用于大型多媒体数据库的图像检索,有很好的使用价值。

参考文献:

[1] Zhang A D ,Cheng B. A fractal-based clustering approach in large visual database systems[J]. Multimedia Tools and Applications ,1996 225 - 244 .

- [2] Zhang A D ,Cheng B. An approach to query-by-texture in image database system[A]. SPIE[C].2606 .
- [3] Guo J ,Zhang A D. Image decomposition and representation in large image database systems[J]. Journal of visual communication and image representation ,1997 8(2) :167 - 181 .
- [4] Lasfar A ,Mouline S ,Aboutajdine D ,Cherifi H. Content-based retrieval in fractal coded image databases[A]. Proceedings of 15th International Conference on Pattern Recognition[C].1 :1031 - 1034 .
- [5] Hamzaoui R ,Hartenstein H ,Saupe D. Local iterative improvement of fractal image codes[J]. Image and Vision Computing ,2000(18) :565 - 568 .
- [6] Wang Z ,Chi Z ,Feng D. Content-based image retrieval using block-constrained fractal coding and nona-tree decomposition[J]. Vision Image and Signal Processing ,IEEE Proceeding ,2000(1) 9 - 15 .
- [7] 张旦 陈刚 金以文.基于金字塔正交小波分解的快速分形图像编码[J].电子学报 ,1998 26(8) 37 - 42 .
- [8] 李均利 陈刚 吴炯峰.基于模糊积分的一类图像度量[J].高校应用数学学报 A 辑 ,2001 ,16(1) 8 - 14 .

作者简介:

洪安祥 男,1971 年生于湖北黄梅,浙江大学应用数学系博士研究生,主要研究方向为数字图像处理、图像压缩和图像检索技术。

陈 刚 男,1963 年生于上海,浙江大学理学院教授,1994 年在浙江大学获理学博士学位,并从事博士后研究工作,专长于应用数学、非线性系统分析、数字图像处理等方面的研究,承担国家和省部级多项研究课题,共发表学术论文和著作 50 余篇(册),现兼任宁波市科技园区管理委员会副主任、浙江大学计算机图像图形研究所副所长等职。