

一种分形域基于内容的图像检索方法¹

杨守义*** 罗伟雄*

* (北京理工大学电子工程系 北京 100081)

** (郑州大学电子工程系 郑州 450052)

摘要 基于内容的图像检索是多媒体、网络通信及计算机等应用研究领域的一项关键技术。该文提出了一种在分形压缩域直接进行基于内容的图像检索方法。该方法不需要对查询图像进行分形变换,因此可以提高检索速度,降低检索复杂度。仿真结果表明,使用该文提出的方法,能够有效地进行分形域基于内容的图像检索,比较大程度地降低了检索时间,优于试验中其他3种方法。

关键词 分形图像编码, 基于内容的图像检索, 迭代函数系统

中图分类号 TN911.73

1 引言

基于内容的图像检索(Content-Based Image Retrieval, CBIR)是一种新的检索技术,它除了利用传统的数据库对图像描述的文字信息进行存贮管理外,还利用图像的颜色、纹理、形状等特征进行检索,融合了传统的模式识别技术与多媒体良好的人机交互技术,有着广泛的应用前景。

当前,已有很多成功的基于目标形状、颜色、纹理和空间结构等可视化特征的检索原型或系统^[1-4],但可视化特征的提取仍然是图像分析领域中没有很好解决的问题。

由于数字图像数据量大,在图像的存储、传输过程中,都是要进行某种形式的图像压缩编码的。图像的压缩编码增加了基于内容图像检索的复杂度,因为查询图像(Iconic image)一般是在空间域给出的。通常的方法是将图像变换到空间域,与给定的查询图像进行比较,得到最相似的图像。

实际上,图像的编码本身在某种意义上来说就是图像特征的抽取过程,因为图像的特征抽取实际上是寻求一种尽可能紧凑的特征码值来描述图像最本质的内容;而图像压缩编码的目的也是要尽可能地减少或消除图像中的冗余信息,从而以尽可能紧凑的代码表征原图像。所以可以说,图像压缩编码的过程实际上就是图像特征抽取的过程,图像的某种压缩编码就可以看成是抽取得到的图像特征。因此,我们可以对查询图像进行相应的压缩编码,然后在压缩域直接进行图像的检索,而不必把图像解压缩到空间域再来进行,这样就可以提高图像检索的速度。

分形图像压缩编码是近年来非常受到关注的一种图像压缩编码技术。它利用图像的自相似性,用图像块的仿射变换系数来表征图像,具有压缩比高、解码简单等优点,是很有前途的一种新的图像压缩编码方法。

A.D.Sloan 首先将分形图像编码技术应用于基于内容的图像检索中^[5]。A.D.Zhang 等人在 A.D.Sloan 的基础上做了进一步的研究,分别提出了基于联合分形编码(Joint Fractal Coding)图像检索策略和基于九叉树分解和分形编码的图像检索策略^[6-8]。

以上这些基于分形编码的图像检索策略都要对查询图像进行分形编码,而分形编码是很耗时的,虽然很多学者在提高分形编码速度方面做了大量的工作^[9,10],但结果仍是不太如人意的。这也就限制了这些基于分形编码的图像检索方法的检索速度。

在这篇文章中,我们提出了一种在分形域进行图像检索的方法。这种方法充分利用分形编码数据的性质,不需要对查询图像进行分形编码,而只要对查询图像在空间域做简单的处理。由于不需要对查询图像进行分形编码,就可以提高分形域基于内容的图像检索速度。

2 分形图像压缩编码原理

分形图像编码的基本理论是迭代函数系统(IFS)和拼贴定理^[11]。

¹ 2001-09-18 收到, 2002-04-19 改回

对于二维的灰度级图像, 其分形压缩变换为

$$W_i(R_i) = s(R_i)S^{(n+1)}\tau(R_i)D(R_i) + o(R_i) \quad (1)$$

式中 $D(R_i) \in \{D_1, \dots, D_m\}$ 是域块, $\tau(R_i) \in \{\tau_1, \dots, \tau_8\}$ 是 8 种对称变换, $S^{(n+1)}: \mathbb{R}^{2^{2(n+1)}} \rightarrow \mathbb{R}^{2^{2n}}$ 是下抽样算子, $s(R_i) \in \{s_1, \dots, s_n\} \subset \mathbb{R}$ 是缩放系数 (scaling factor), $o(R_i) \in \{o_1, \dots, o_l\} \subset \mathbb{R}$ 是偏移量 (offset), R_i 是值块.

分形图像压缩编码, 就是要对每一个值块, 在域块库中找到其相似块和相应的仿射变换, 则对应每一个值块的域块位置信息 $D(R_i)$ 和仿射变换系数中的对称变换 $\tau(R_i)$ 、缩放系数 $s(R_i)$ 以及偏移量 $o(R_i)$ 一起组成了一幅图像的分形变换编码, 可以代表该图像进行存储或传输. 或者说, 这组仿射变换码就是对该图像提取的一种特征.

对于一个给定的查询图像, 如果要在用分形编码进行压缩的图像库中进行检索, 可以对查询图像做同样的分形变换, 然后在分形压缩图像库中, 把查询图像的分形码与库中图像的分形码进行比较, 从而找出最相似的图像. 但分形编码是很耗时的. 当然, 也可以把图像库中的图像进行分形解码, 然后与查询图像进行比较, 虽然分形解码比较快, 但由于要对库中所有的图像进行解码, 所以仍然是比较费时的.

3 图标的分形域直接比较检索法

事实上, 图像的分形编码数据里就显式地包含着图像的空间域信息. 由 (1) 式可知, 对于一个给定的值块, 有

$$\overline{M}[W_i(R_i)] = \overline{M}[s(R_i)S^{(n+1)}\tau(R_i)D(R_i)] + o(R_i) \quad (2)$$

式中 $\overline{M}[\bullet]$ 表示取均值操作. 另一方面, 为了使域块 $D(R_i)$ 经变换 W_i 后尽量与值块 R_i 匹配, 应有

$$\overline{M}[W_i(R_i)] = \overline{R}_i \quad (3)$$

式中 \overline{R}_i 为值块 R_i 的均值. 由 (2) 式和 (3) 式得

$$\overline{M}[W_i(R_i)] = \overline{M}[s(R_i)S^{(n+1)}\tau(R_i)D(R_i)] + o(R_i) = \overline{R}_i \quad (4)$$

因此

$$o(R_i) = \overline{R}_i - \overline{M}[s(R_i)S^{(n+1)}\tau(R_i)D(R_i)] \quad (5)$$

也就是说, 对某一值块 R_i 的仿射变换中的偏移量 $o(R_i)$, 实际上是该值块的均值与相对应的域块应变换操作后的均值之差. 因此, 由各值块的 $o(R_i)$ 值所组成的图像 \overline{f}_{ij} (称为偏移量图像), 与原图像 f_{ij} 有着紧密的联系. 由于对应于原图像 f_{ij} 的一个值块 (有 $N \times N$ 个像素), 在图像 \overline{f}_{ij} 中只对应一个像素, 因此图像 \overline{f}_{ij} 的大小是原图像 f_{ij} 的 $1/(N \times N)$, 其每一像素的灰度值是对应的原图像的值块的灰度平均值与相似域块变换后的灰度平均值之差.

由于偏移量图像对应于原图像的 $N \times N$ 抽取 (水平方向 N 点抽取一点, 垂直方向 N 点抽取一点), 所以对于给定的查询图像 I_{ij} , 我们也可以做同样的抽取 (即其 $N \times N$ 的一个块求平均得到一个点), 得到查询图像的抽取图像 \overline{I}_{ij} , 然后用抽取图像 \overline{I}_{ij} 与偏移量图像 \overline{f}_{ij} 进行比较, 按某种准则计算其距离, 从而得到最相似的图像, 该图像即是对给定查询图像在分形压缩图像库中找到的检索图像.

当进行两幅图像的比较时, 可以以两幅图像的对应像素灰度差的平方和作为两幅图像的距离度量, 即

$$d_{\overline{f}\overline{I}} = \sum_{i=0}^{U-1} \sum_{j=0}^{V-1} [\overline{f}(i+p, j+q) - \overline{I}(i, j)]^2 \quad (6)$$

式中 $d_{\bar{f}\bar{I}}$ 是两幅图像 \bar{f} 和 \bar{I} 之间的距离度量, $\bar{f}(i, j)$ 和 $\bar{I}(i, j)$ 分别表示 (i, j) 点图像 \bar{f} 和图像 \bar{I} 的像素灰度值, U 和 V 是图像 \bar{I} 的水平方向像素数和垂直方向像素数, 因为图像 \bar{I} 总是小于至多等于图像 \bar{f} 的大小 (给定的查询图像 I_{ij} 总是小于或等于图像 f_{ij} 的), 所以在把图像 \bar{I} 与图像 \bar{f} 比较时, (p, q) 是图像 \bar{f} 的左上角开始位置坐标。

但是, 图像 \bar{f} 并不是原图像 f_{ij} 的平均图像, 而是平均图像与域块变换图像之差, 它并不能很好地代表原图像, 所以在检索时, 如果图像数据库中的某些图像比较接近的话, 检索结果的准确性可能就会变差。如果我们能够在图像的分形压缩编码中插入值块的平均灰度, 那么在图像检索时就可以得到原图像的平均图像了。不过, 这样就会增加分形编码数据, 降低分形编码的压缩比。

事实上, 可以只要值块的平均灰度而省略偏移量 $o(R_i)$, 因为偏移量 $o(R_i)$ 可以由值块的平均灰度求得。根据 J. M. Beaumont 的分形编码方案^[12], 对于第 i 个值块, 有

$$\frac{S^{(n+1)}\tau(R_i)D(R_i) - \overline{M}[D(R_i)]}{\sigma_d} = \frac{R_i - \overline{M}[R_i]}{\sigma_r} \quad (7)$$

式中 σ_d 和 σ_r 分别是值块和域块的标准方差。上式可以改写为

$$R_i = \frac{\sigma_r}{\sigma_d} S^{(n+1)}\tau(R_i)D(R_i) + \overline{M}[R_i] - \frac{\sigma_r}{\sigma_d} \overline{M}[D(R_i)] = s(R_i)S^{(n+1)}\tau(R_i)D(R_i) + o(R_i) \quad (8)$$

因此有

$$o(R_i) = \overline{M}[R_i] - \frac{\sigma_r}{\sigma_d} \overline{M}[D(R_i)] = \overline{M}[R_i] - s(R_i)\overline{M}[D(R_i)] \quad (9)$$

由 (9) 式可以看出, 只要由值块的均值 $\overline{M}[R_i]$ 、域块的均值 $\overline{M}[D(R_i)]$ 以及缩放系数 $s(R_i)$ 就可以求出偏移量 $o(R_i)$ 来。而域块的均值可以由拼贴成该域块的各值块的均值求得。所以, 我们可以用值块的均值来代替偏移量, 作为图像的分形编码。图 1 是采用这种分形编码后, 从分形编码中得到的 Lenna, san256 和 collic 图像的平均图像 \bar{f} (值块大小为 4×4)。

当给定了查询图像后, 图像检索的步骤如下:

- (1) 从分形压缩图象库中图像的分形码中得到原图像的平均图像;
- (2) 把查询图像按照分形编码时的平均方法一样进行平均抽取, 得到查询图像的平均图像;
- (3) 把二者进行比较, 并按照 (6) 式的距离准则进行判别, 从而找出距离最近的图像作为检索结果, 或者是给定一个距离阈值, 把小于距离阈值的图像作为检索结果。

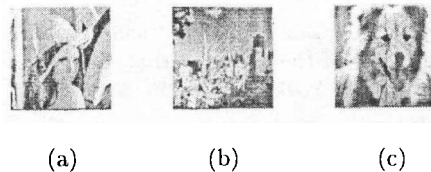


图 1 由 J. M. Beaumont 编码方案得到的 Lenna(a), san256(b), collic(c) 的平均图像

4 仿真结果

使用本文提出的方法, 对 256×256 大小的 Lenna 图像、san256 图像、collic 图像分别进行了检索, 查询图像是从原图像中取出的一部分, 表 1 是对 256×256 大小的 Lenna 图像用本文方法检索时间与其它 3 种方法检索时间的比较。由表 1 可以看出, 本文所提出的基于内容的图像检索方法, 可以较大幅度地减少检索时间, 提高检索效率。

表 1 几种基于内容分形压缩图像检索方法比较

检索策略	Zhang 基于九叉树分解和分形码的检索策略 ^[6]	基于直接像素匹配的检索策略 ^[8]	Wang 基于九叉树分解的检索策略 ^[8]	本文方法
所用时间 (s)	303.3	25.9	4.7	1.9

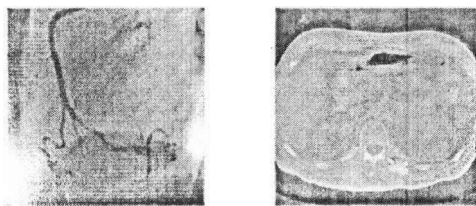


图 2 两幅医学图像

图 2 是两幅医学图像, 使用本文方法对包括这两幅图像在内的 10 幅医学图像的分形压缩图像在分形域直接进行检索, 使用 Pentium II 350 微机, 对一个给定的查询图像, 检索时间是 19.2s, 对所有给定的查询图像都能够得到正确的检索结果。

5 结论及讨论

在这篇文章中, 我们提出了一种在分形域基于内容的图像检索方法。这种方法不需要对查询图像进行分形编码, 而只要对查询图像进行平均求出其平均图像, 因而可以降低图像的检索时间, 提高图像检索效率。对 Lenna 等图像和一些医学图像的试验结果表明, 与其它几种方法相比, 图像的检索时间至少可以减少一半。但是, 由于该方法要对查询图像进行抽取得到平均图像, 因此当给定的查询图像小时, 得到的抽取平均图像就会太小, 又可能会影响到图像的检索准确性。所以这方面还是有待于进一步探讨的问题。另外, 有关该方法的精度和鲁棒性等方面的分析, 将在以后的工作中作进一步探讨。

参 考 文 献

- [1] M. D. Marsicoi, L. Cinque, S. Levialdi, Indexing pictorial documents by their contents: A survey of current techniques, *Image and Vision Computing*, 1997, 15(2), 199-141.
- [2] J. Ashley, *et al.*, Automatic and semi-automatic methods for image annotation and retrieval in QBIC, in *SPIE Storage and Retrieval for Image and Video Database III*, San Jose, California, Feb. 1995, 2420, 24-35.
- [3] P. M. Kelly, M. Cannon, D. R. Hush, Query by image example: the CANDID approach, in *SPIE Storage and Retrieval for Image and Video Database III*, San Jose, California, Feb. 1995, 2420, 238-248.
- [4] E. Remias, G. Sheikholeslami, A. D. Zhang, Block-oriented image decomposition and retrieval in image database systems, In *Proc. of the 1996 International Workshop on Multimedia Database Systems*, Blue Mountain Lake, New York, Aug. 1996, 85-92.
- [5] A. D. Sloan, Retrieving database contents by image recognition: New fractal power, *Advanced Imaging*, 1994, 9(5), 5-15.
- [6] Aidong Zhang, Diao Cheng, Raj Acharya, An approach to query-by-texture in image database systems, In *Proc. of the SPIE Conference on Digital Image Storage and Archiving Systems*, Philadelphia, USA, Oct. 1995, 338-349.
- [7] Aidong Zhang, Diao Cheng, Raj Acharya, Raghu Menon, Comparison of wavelet transforms and fractal coding in texture-based image retrieval, In *Proc. of the SPIE Conference on Visual Data Exploration and Analysis III*, San Jose, Jan. 1996, 1753-1762.
- [8] 王志勇等, 分形编码在图像检索中的应用, *电子学报*, 2000, 28(6), 19-23.
- [9] F. Barnsley, *Fractal Everywhere*, San Diego, Academic press, 1988, 55-78.

- [10] Fisher, Fractal Image Compression: Theory and Application, New York, Spring-Verlag, Inc., 1995, 89-115.
- [11] A. E. Jacauin, A novel fractal block-coding technique for digital images, Proc., ICASSP, USA, 1990, 2225-2228 .
- [12] J. M. Beaumont, Advances in block based fractal coding of still pictures, Proc. of the IEE Colloquium, The Application of Fractal Techniques in Image Processing, 1990.

CONTENT-BASED IMAGE RETRIEVAL IN FRACTAL DOMAIN

Yang Shouyi* ** Luo Weixiong*

**(Electronic Engineering Dept., Beijing Institute of Technology, Beijing 100081, China)*

*** (Electronic Engineering Dept., Zhengzhou University, Zhengzhou 450052, China)*

Abstract Content-based image retrieval is a key technique in multimedia, network communication and computer applications, etc. In this paper, a content-based image retrieval method is proposed, which can be executed in fractal domain and does not need to do fractal coding of the iconic image. Thus the retrieval complexity and retrieval time can be reduced greatly. Experimental results have shown that the method proposed is very effective and better than the other three retrieval methods.

Key words Fractal image coding, Content based image retrieval, Iteration function system

杨守义: 男, 1965 年生, 副教授, 博士生, 主要研究方向: 图像压缩编码、图像处理等。
罗伟雄: 男, 1940 年生, 教授, 博士生导师, 研究方向为 GPS 系统, 信号处理。