

一种适用于网络应用的压缩图象快速检索方法

李晓华, 沈兰荪, 贾克斌

(北京工业大学信号与信息处理研究室, 北京 100022)

摘要: 随着多媒体技术和因特网的发展, 基于内容的图象检索已经成为多媒体处理中的关键技术, 而基于压缩域的图象检索由于其低复杂性逐渐形成一个新的研究热点. 针对当前网上图像浏览与检索的应用要求, 本文提出了一种基于 JPEG2000 压缩码流的压缩图象快速检索方法. 该方法直接从 JPEG2000 压缩码流包头中抽取码块零位平面数, 码块编码通道数及码块编码长度信息, 并进而构造特征矢量用于图象检索. 初步实验结果验证了该方法的有效性和快速性.

关键词: JPEG2000; 基于内容; 图象检索; 网上检索

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112 (2002) 12A-2016-04

A Fast Indexing Method on Compressed Images for Internet Applications

LI Xiao-hua, SHEN Lan-sun, JIA Ke-nan

(Signal & Information Processing Lab., Beijing Polytechnic University, Beijing 100022, China)

Abstract: Multimedia and INTERNET is increasingly developing in recent years. It is crucial to develop indexing techniques for searching images and video based on their content. Due to the lower computational complexity, compressed domain indexing techniques are becoming popular. In this paper, an efficient compressed image indexing scheme based on JPEG2000 framework is proposed. The target application is the access and interaction with huge amount of visual data on internet. Here, the information on the number of zero bit-planes in code-blocks, number of coding passes in code-blocks and length of the data from code-blocks is directly extracted from the packet header of a JPEG2000 compressed image without decompressing the bitstream. A feature vector is built based on the three-part information and is used as the image index. Preliminary experimental results show that the proposed scheme is efficient and effective.

Key words: JPEG2000; content-based; image indexing; internet retrieval

1 引言

近年来, 随着多媒体技术和因特网的发展, 基于内容的图象检索已经成为一个研究热点, 并成为多媒体数据库, 数字图书馆, 网上浏览检索, 网络信息安全等重大研究项目中的关键技术^[1]. 基于内容检索的基础是快速有效提取各种表征图像内容的特征, 如颜色, 纹理, 形状, 语义等. 一般来说, 特征的选择遵循两个原则^[2]: (i) 该特征具有较好的区分不同图像的能力, 即具有较好的检索效率; (ii) 特征的计算及相似性匹配具有较低的计算复杂度(或较小的内存需求量), 即高效性. 现有的特征提取技术大部分是基于图像的非压缩数据, 对于广泛存在的压缩格式图像, 一般需要先进行解压操作, 而解压过程是比较费时的, 也就是说, 不符合高效性的原则. 如果可以直接从压缩数据中提取特征, 就可以大大减少系统的处理时间, 另外由于处理的是大为减少的压缩数据, 资源占用量也会减少. 这样就兼顾了特征选择的两个原则, 特别适合网上检索与浏览. 与集中式多媒体数据库不同, 网上图像浏览与检索对于特征提取及相似性匹配的速度具有更高的要求, 因为集中式多媒体数据库可以事先离线抽取特征并存储在一个特征库中, 不需要频繁更新; 而网上图象则不同, 每时每刻都会有新

图象的加入, 需要实时提取特征. 据统计, 一些骨干网上的流量高达每秒数 G 字节, 这些信息一般都以压缩码流的形式在网上传输, 这就需要针对压缩图像的快速内容分析及特征抽取技术来实时的监控并处理这些信息.

本文提出了一种新的特征提取方法. 这种方法直接从 JPEG2000 压缩码流中提取特征量, 由于省去了解码过程, 处理速度大大提高, 特别适用于网上图像的浏览与检索.

2 JPEG2000 简介

JPEG2000 是推出的最新静止图象编码标准, 它不仅在压缩性能方面明显优于现有的 JPEG 标准, 还具有很多 JPEG 无法提供或无法有效提供的新功能, 比如, 同时支持有损和无损压缩、大幅图像的压缩、渐进传输、感兴趣区编码、良好的传输鲁棒性、码流随机访问等^[3].

JPEG2000 编解码过程方框图如图 1 所示. 编码器首先对原始图像做小波变换, 根据小波系数的特点对其进行量化. 将量化后的系数划分成小的编码单元-码块, 对每个码块进行独立的嵌入式编码, 然后按照率失真最优原则将所有码块的嵌入式位流分层组织. 最后按照一定的码流格式对各层打包, 输出压缩码流.

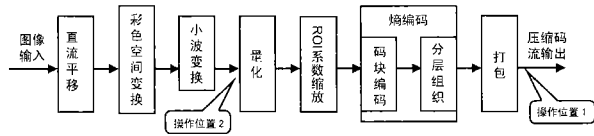


图1 JPEG2000 编码过程方框图

JPEG2000 在打包的过程中,在包头中放置了有关码块的编码信息.本文的算法将直接在最终的压缩码流处,即操作位置1(见图1)进行操作,利用包头中的各种信息构建特征量用于检索.

JPEG2000 包头信息

在 JPEG2000 码流结构中,把编码结果按分辨率、质量层进行打包.这里分辨率与小波变换的分解层相对应,假定对图象进行3级小波分解,就得到4级分辨率,如图2所示,子带 LL3 被称为第1级分辨率, LH3、HL3 和 HH3 称为第2级分辨率, LH2、HL2 和 HH2 称为第3级分辨率, LH1、HL1 和 HH1 称为第4级分辨率.质量层是指根据最优率失真原则对各个分辨率码流进行的分层,分层是在码块级进行的,也就是需要标记每一码块中所有编码通道所属的层.

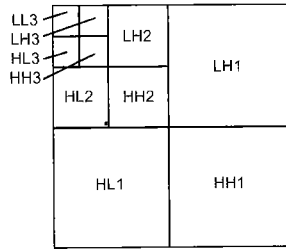


图2 图像小波分解示意图

每个包由两部分组成:包头及包体.包体中是特定分辨率、特定质量层的编码码流,即某一分辨率在某一质量层的熵编码码流;而包头中主要包含下面几方面的信息^[4]:(1)码块零位平面数 P ; (2)码块中第一个编码通道所属质量层; (3)码块在每层中的编码通道数; (4)码块对每层码流的贡献长度;

这四部分信息中的第一、二部分用标签树的策略进行编码,第三部分用查表的方式进行编码,第四部分直接用其二进制数表示.

3 基于 JPEG2000 压缩码流的图象检索

本文基于这样一种假定,如果子带能量越大,则在某一确定的压缩比下,子带参与编码的通道数就越多,子带对整个码流长度的贡献也就越大.这样就可以利用包头信息计算特征量来支持图象检索.

3.1 提取包头信息

由 §2 可知,包头中包含 4 方面的码块编码信息,这里主要利用(1)、(3)、(4)三方面的信息构建特征矢量.通过简单的解包头操作,就可得到各个子带所有码块的编码信息:

$$\{H_1^k(i), H_2^k(i, j), H_3^k(i, j)\}, \quad k = 1, 2, \dots, 3M + 1; \\ i = 1, 2, \dots, N_k; j = 1, 2, \dots, L \quad (1)$$

其中: $H_1^k(i)$ 代表第 k 个子带中第 i 个码块的零位平面数; $H_2^k(i, j)$ 代表第 k 个子带中第 i 个码块在 j 质量层的编码通道数; $H_3^k(i, j)$ 代表第 k 个子带中第 i 个码块对 j 质量层码流的贡献长度; M 是小波分解级数; N_k 是第 k 个子带中码块的数量; L 是压缩码流的分层数目;

3.2 基于包头信息构建特征量

下面介绍如何利用 §3.1 中的 $H_1^k(i), H_2^k(i, j), H_3^k(i, j)$ 构建特征量 V 用于图像检索.

I 计算码块非零位平面数 $H_1^k(i)$. 码块非零位平面数反映了码块中小波系数幅度最大值的数量级. $H_1^k(i)$ 可以由下式计算:

$$H_1^k(i) = M_k - H_1^k(i) \quad (2)$$

其中: M_k 是第 k 个子带小波系数的最大编码位平面数,该值可由压缩码流中主头的有关信息求得^[4].

II 计算每一子带中所有码块非零位平面数的均值和方差:

$$M^k = (\sum_{i=0}^{N_k} H_1^k(i)) / N_k \quad (3)$$

$$\sigma^k = \sqrt{(\sum_{i=0}^{N_k} (H_1^k(i) - M^k)^2) / N_k} \quad (4)$$

III 计算每一子带中加权编码通道数:

$$N^k = \sum_{j=1}^L w_j \sum_{i=1}^{N_k} H_2^k(i, j) \quad (5)$$

其中 w_j 表示层加权系数. JPEG2000 中对每一码块的小波系数采用的是比特平面编码方式,这样系数的高位就被先编码,在分层的压缩码流中,即使不同层包含同样数目的编码通道或编码长度,但他们对图像内容的贡献是不同的,属于码流前部质量层的编码通道或编码长度对图像的内容将具有更大的贡献,因此,这里引进一个加权系数 w_j , 来反映各层中编码通道或编码长度对图像内容的贡献程度,本文采用下面的公式计算 w_j :

$$w_j = W_a \times 2^{-j} \quad (6)$$

其中 W_a 是预先设置的参数.

IV 统计每一子带的加权编码长度

$$L^k = \sum_{j=1}^L w_j \sum_{i=1}^{N_k} H_3^k(i, j) \quad (7)$$

其中层加权系数 w_j 由式(6)计算.

综合 II、III、IV 三部分的统计结果,可以得到一个长度为 $4 \times (3M + 1)$ 的特征量 V :

$$V = \{M^k, \sigma^k, N^k, L^k, k = 1, 2, \dots, 3M + 1\} \quad (8)$$

3.3 特征量归一化

由上述方法得到的特征量由四块分量组成,他们表征不同的编码信息,幅度大不相同,在进行距离匹配时,会产生较大的偏差,通常需要通过特征归一化来消除这种偏差.这里通过对图象库中的所有图象进行统计,得到 M^k, σ^k, N^k, L^k 的最大值 $M_{max}, \sigma_{max}, N_{max}, L_{max}$, 然后对式(8)中的 V 进行归一化:

$$V = \{M^k / M_{max}, \sigma^k / \sigma_{max}, N^k / N_{max}, L^k / L_{max}\}, \\ k = 1, 2, \dots, 3M + 1 \quad (9)$$

最后得到归一化的特征量:

$$V = \{v_1, v_2, \dots, v_N\}, N = 4 \times (3M + 1) \quad (10)$$

3.4 相似性匹配

I 对检索输入图像 q 计算其特征量:

$$V^q = \{v_1^q, v_2^q, \dots, v_N^q\} \quad (11)$$

II 通过式(12)计算检索输入图像 q 和图象库中一幅图像 I 之间的距离 D :

$$D(q, I) = \sum_{i=1}^{i=N} (v_i^q - v_i^I)^2 \quad (12)$$

III 计算库中所有图像和检索输入图像 q 之间的距离.

IV 选出距离最小的若干幅图像作为检索输出图像.

4 实验及结果

通过大量实验对所提出新方法的性能进行了验证,并与有关方法进行了对比.本文的实验是在我们自己搭建的压缩域图像检索系统(CDIRS)上进行的.该系统的数据库包含1000幅图像,内容涉及人物、动物、建筑、纹理以及风景等许多方面,以保证实验结果的可信度和可靠性(每一幅图像在库中都有10幅相似图像,包含光照、平移、尺度及旋转变化).图象库中的图像均为 512×512 大小的灰度图像,以 JPEG2000 压缩格式存放.压缩时小波分解层数为3,码流分为10层,最小压缩比为30.

实验是从两个角度进行的:检索结果的准确性(effectiveness)和特征量计算及相似性匹配的快速性(efficiency).

4.1 特征量计算及相似性匹配的快速性

在这里把本文提出的方法和文献[5]中介绍的方向梯度相角直方图方法、文献[7]中介绍的基于小波子带能量法进行对比,因为文献[5],文献[7]中的方法都可以看作是基于小波压缩域的图象检索技术.方向梯度相角直方图特征量的提取包括以下几个步骤^[5]:(i)对压缩图象进行熵解码和反量化,得到小波变换图像;(ii)利用小波系数提取多尺度边缘图像^[6];(iii)构造每一尺度下的方向梯度相角直方图.子带能量法特征量的提取包括2个步骤^[7]:(i)对压缩图象进行熵解码和反量化,得到小波变换图像;(ii)计算各子带小波系数能量分布的均值和方差,即特征量.

本文方法直接在压缩码流上提取特征,只需解开包头即可统计特征矢量,省去了熵解码和反量化过程,计算时间大为缩减.表1列出了三种方法提取图象库中1000幅图像的特征

矢量所用的时间(同一台机器上).小波分解层数为3时,按本文方法特征矢量维数是40,方向梯度相角直方图方法的特征矢量维数是256,子带能量法特征矢量维数是20.相似性匹配用的都是欧式距离,匹配的速度与特征矢量的维数成正比.

表1 三种方法性能对比

特征提取方法	提取1000幅图像特征所用时间	检索效率
本文方法	47.98秒(解包头9.6秒)	80.8%
方向梯度相角直方图	182.76秒(熵解码+反量化121.3秒)	78.2%
子带能量法	163.47秒(熵解码+反量化121.3秒)	83.0%

从表1可以看出,本文方法的优势主要表现在速度方面,这主要是因为本文方法是在操作位置1(见图1)提取特征,基本不需要解码;而方向梯度相角直方图方法及子带能量法是在操作位置2(见图1)提取特征,需要先做熵解码和反量化.速度方面的优势使得本文方法特别适用于网上图象浏览与检索的要求.

值得指出的是,本文方法与传统的像素域(非压缩域)检索技术相比,在速度方面具有更大的优势.这是因为,与方向梯度相角直方图方法及子带能量法只需部分解码(熵解码和反量化)不同,像素域技术通常需要对压缩图象全解码(熵解码、反量化及小波反变换),再进行特征提取,而小波反变换的计算量在整个解压过程中占主要部分(60%左右).

4.2 检索结果的准确性

检索结果的准确性采用检索效率 η_R 来衡量, η_R 采用文[8]中介绍的评价准则:假设图像数据库内的图像总数为 P ,对其中每一幅图像 $i(1 \leq i \leq P)$,人工列出图像库中与其内容相似的 $N_i(1 \leq N_i \leq P)$ 幅图像.对每一次输入的查询图像 q ,检索输出相似度最大的 $N_q + t$ 幅图像, t 为预先设定的检索余量,如果输出图像中有 n_q 幅检索成功的图像,则 η_R 定义

$$\eta_R = \sum_{q=0}^P n_q / \sum_{q=0}^P N_q \quad (13)$$

该 η_R 值的大小直观地表征了检索效率的高低.

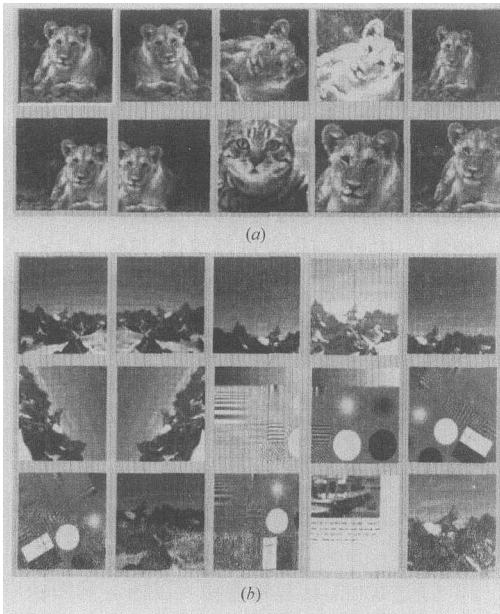


图3 本文方法检索结果实例

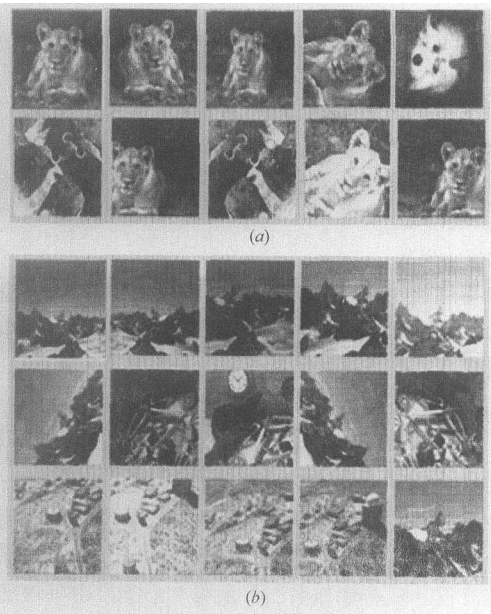


图4 文献[5]方法检索结果实例

检索图像直接从图像库中选取,图 3 为本文方法检索结果的 2 组实例,结果图像的相似度从左到右,从上到下依次排列,其中左上角的是例子图像.图 3(a)表明用本文提出的方法进行图像检索时,在相似度最大的 10 幅结果图像中(即检索余量 $t=0$),库中与例子图像相似的 10 幅中的 9 幅被成功检出,图 3(b)表明了相似度最大的 15 幅结果图像中(即检索余量 $t=5$),库中与例子图像相似的 10 幅中的 9 幅被成功检出.图 4 为方向梯度相角直方图方法的 2 组检索结果实例,图 4(a)表明用文献[5]中介绍的方法进行图像检索时,在相似度最大的 10 幅图像中,库中与例子图像相似的 10 幅中的 7 幅被成功检出,图 4(b)表明了相似度最大的 15 幅结果图像中,库中与例子图像相似的 10 幅中的 8 幅被成功检出.图 5 为子带能量法检索结果的 2 组实例.另外还可以看出,相对于方向梯度相角直方图方法,本文方法得到的结果中被正确检索出来的图像更靠前.

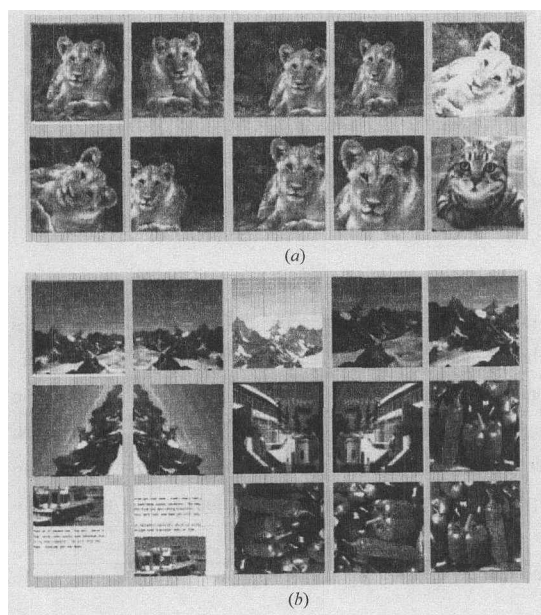


图 5 子带能量法检索结果实例

统计检索效率的实验中我们取检索余量 $t=5$.表 1 给出了三种方法的检索效率.可以看出,本文方法在速度方面具有明显优势,同时,检索正确率与方向梯度相角直方图法和子带能量法基本相当.

5 结论及讨论

本文提出了一种基于 JPEG2000 压缩码流的图像检索方法,对于 JPEG2000 压缩格式图像,可以在基本不解码的情况下,直接从包头中提取信息构建表征图像内容的特性量,初步实验结果验证了该方法的有效性和快速性.本文方法的最大特点是在保持检索效率基本不降低的情况下,大大减少特征计算的复杂度.利用这种检索技术,对于以压缩格式存放的图

象数据库,无需预先形成额外的特征库,可以在检索过程中快速计算特征量,从而对图像数据库的动态管理非常方便,特别利于进行基于互联网的图像检索与查询.

我们从实验结果中发现,与子带能量法相似,本文方法对纹理图像的检索效果尤其好,这主要是因为该方法统计了小波变换各子带的多种特性,与基于子带能量的纹理分类方法^[7]具有相似的内容描述,即纹理特征描述,所以该文提出的方法亦可用于压缩域纹理图像快速分类,这将是我们将要进行的研究工作.

参考文献:

- [1] Smeulders A W M, Worring M, Santini S, et al. Content-based image retrieval at the end of the early years[J]. IEEE transactions on pattern analysis and machine intelligence, 2000, 22(12): 1349 - 1379.
- [2] Mandal M K, Idris F, Panchanathan S. A critical evaluation of image and video indexing techniques in the compressed domain[J]. Image and Vision Computing, 1999, 17: 513 - 529.
- [3] Christopoulos C, Skodras A, Ebrahimi T. The JPEG2000 still image coding system: an overview[J]. IEEE Transactions on Consumer Electronics, 2000, 46(4): 1103 - 1127.
- [4] ISO/IEC 15444 - 1, JPEG 2000 Part-1 standard[S]. 2000.
- [5] 魏海, 沈兰荪. 小波变换域内基于方向梯度相角直方图的图像检索算法[J]. 电路与系统学报, 2001, 6(2): 20 - 24.
- [6] 魏海, 沈兰荪. 反对称双正交小波应用于多尺度边缘提取的研究[J]. 电子学报, 2002, 30(3): 313 - 316.
- [7] Nicu Sebe, M S Lew. Wavelet based texture classification[A]. Proc. of the 15th International Conference on Pattern Recognition[C]. Spain: Barcelona, 2000. 947 - 950.
- [8] Flickner M, Sawhney H. Query by image and video content: The QBIC system[J]. Computer, 1995, 28: 23 - 32.

作者简介:



李晓华 女, 1973 年 2 月生, 陕西韩城人, 1998 年毕业于西安交通大学, 获工学硕士学位, 现为北京工业大学模式识别与智能系统专业博士生, 主要研究领域为基于内容图像检索, 图像压缩等.



沈兰荪 男, 1938 年 6 月生, 江苏省苏州人, 北京工业大学教授, 博士生导师, 主要研究领域为图像处理, 视频编码与传输, 医学图像处理等, 已出版《图像编码与异步传输》、《视频编码与低速率传输》等专著.