

基于 DCT 压缩域的图像检索方法

黄祥林, 宋 磊, 沈兰荪

(北京工业大学信号与信息处理研究室, 北京 100022)

摘 要: 本文提出了一种基于 DCT(Discrete Cosine Transform)压缩域的图像检索方法. 对于 DCT 编码的图像数据, 在不需要完全解码的情况下, 直接抽取图像的内容特征进行图像检索. 首先, 重组 DCT 域的频率系数, 使其具有方向性、多分辨率等特点, 并利用这些特点提取图像的大致轮廓. 接着统计图像轮廓的连通直方图(CRH: Connected-Region Histogram), 进行图像检索. 并利用 DC 图的灰度直方图对检索结果进行重新排序. 这种检索方法对灰度、旋转、平移等都具有一定的鲁棒性, 具有较好的检索效果.

关键词: 基于内容的图像检索; DCT; 压缩域图像检索; 连通直方图 (CRH); 灰度直方图

中图分类号: TN919 **文献标识码:** A **文章编号:** 0372-2112 (2002) 12-1786-04

Image Retrieval Based on DCT Compressed-Domain

HUANG Xiang-lin, SONG Lei, SHEN Lan-sun

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

Abstract: In this paper, An image retrieval method is proposed based on DCT compressed-domain. With the partial decoding of DCT-encoded image data, the image features are extracted directly from the compressed domain. At first, reorder the DCT parameters in compressed-domain to get the directionality and multi-resolution for extracting the sketch of image. And then retrieved images by the CRH (Connected-Region Histogram) of image sketch. At last, we can sort the retrieved result by the gray histogram of DC map. The method is robust to gray, rotation, translation, and so on. The experimental results show that the proposed method is effective in image retrieval.

Key words: content-based image retrieval (CBIR); DCT; image retrieval in compressed-domain; connected-region histogram; gray histogram

1 引言

随着图像数据的急剧增加, 如何进行图像检索变得愈来愈重要. 由于多种原因^[1,2], 传统的基于文本的图像检索技术已无法满足实际要求. 因此, 人们对基于内容的图像检索技术进行了广泛的研究, 并且已经有了许多研究成果和实验系统^[1-3].

大量的图像数据都以压缩格式进行存储和传输, 但目前对图像检索技术的研究大多是基于非压缩图像数据的^[1-3]. 对压缩图像数据进行检索时, 要进行图像的“全解压”操作, 显然限制了检索系统性能的提高, 因此, 人们开始进行基于压缩域的图像检索技术研究^[1,4,5]. 由于 DCT 在图像编码中的重要地位(如 JPEG、H. 261/263、MPEG-1/2 等均采用 DCT 变换), 进行基于 DCT 压缩域的图像检索技术研究极具现实意义.

在 DCT 压缩域进行图像检索, 主要是分析并提取能反映图像内容的 DCT 系数的特性. 目前大致有两种方法^[4,6,7], 第一种方法直接对 DC 图进行处理, 因为 DC 图是整幅图像的缩图, 所以可采用灰度直方图等常用的图像检索技术, 由于没有利用 DCT 域的低中频信息, 检索效果受到一定的影响. 第二种方法通过计算 DCT 系数的能量直方图等信息来检索图像, 其检索效果也不理想.

本文提出了一种新的 DCT 压缩域图像检索算法, 对灰

度、旋转、平移等都具有一定的鲁棒性. 首先对 DCT 域系数进行重组得到具有方向性和多分辨率的图像特征, 并利用此特征抽取图像的大致轮廓. 然后采用一种基于轮廓的连通直方图进行图像检索, 并采用 DC 图的灰度直方图对最终结果进行排序.

2 DCT 域的多分辨率特性^[8]

JPEG、H. 261/263、MPEG-1/2 等图像压缩标准都采用了 8×8 的 DCT 变换, 其基本过程如图 1. 首先把原始图像分成若干 8×8 的像素块, 对每个像素块进行 DCT 变换, 然后进行量化、Zigzag 排序、RLE 编码、熵编码等操作, 最终形成压缩码流. 因此, 只要对压缩码流进行熵解码后就能得到 DCT 的系数.

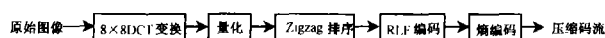


图 1 基于 DCT 变换的压缩编码算法的基本编码框图

为了构造 DCT 系数所具有的多分辨率特性, 这里参照对 8×8 图像进行三级小波分解的情况^[9], 分解后得出了 10 个子带: LL3、LH3、HL3、HH3 四个子带均只有 1 个系数, LH2、HL2、HH2 三个子带均含 4 个系数, 而 LH1、HL1、HH1 三个子带均含 16 个系数. 根据这种方式, 对 DCT 变换的系数也进行 10 个区域的划分: B4、B3_1、B3_2、B3_3、B2_1、B2_2、B2_3、B1_1、B1_2、B1_3, 其所含系数的个数分别为 1、4、16, 与小波的各个

收稿日期: 2001-07-17; 修回日期: 2001-09-29

基金项目: 国家自然科学基金 (No. 60172045); 国家 863 计划 (2001AA114140, 2002AA119010)

子带相对应,如图 2 所示.

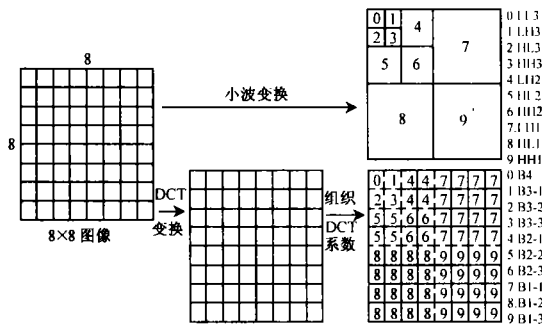


图 2 8×8 图像的小波变换与 DCT 变换

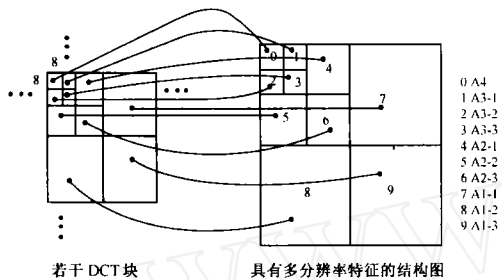


图 3 整幅图像中 DCT 块系数区域的组织

对于整幅图像,把所有的 8×8 DCT 块中相对应的区域组织在一起(如图 3),即:所有的 B4 组成区域 A4,所有的 B3_1 组成区域 A3_1,……,所有的 B1_3 组成区域 A1_3.实际上,区域 A4 就是整个图像的低频近似(即所有 DCT 直流系数 B4 所组成的 DC 图),正好与三级小波的 LL3 相对应;而区域 A3_1、A2_1、A1_1 主要代表了原图像中不同频率段上水平方向的频率变化(竖直方向的边缘、轮廓、纹理等),与三级小波的 LH3、LH2、LH1 相对应;区域 A3_2、A2_2、A1_2 则主要集中了原图像中各频率段上垂直方向频率变化(水平方向的边缘、轮廓、纹理等),与三级小波的 HL3、HL2、HL1 相似;区域 A3_3、A2_3、A1_3 基本上原图像中各频率段上对角方向的频率变化(对角方向的边缘、轮廓、纹理等),与三级小波的 HH3、HH2、HH1 一致.因此,通过这种方法对 DCT 系数进行重新组织,就形成了与三级小波分解基本相似的具有方向性的多分辨率结构的形式.图 4(a)是一幅原始图像,其三级小波分解图像如图 4(b),经 8×8 DCT 变换后重新组织系数得到的多分辨率形式如图 4(c).由图可知,它们之间确实具有很强的对应关系.

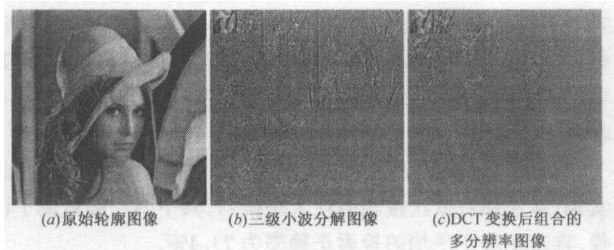


图 4 DCT 变换与小波分解的多分辨率结构图比较

3 提取图像的大致轮廓

对 DCT 域的频率系数进行重组后得到了具有方向性和多分辨率的特征.在图 3 中,区域 A1_1、A1_2、A1_3 主要代表了原图像中精细变化,人眼对其并不敏感.而反应图像主要内容的区域为 A4、A3、A2,因此,这是提取内容特征的重点区域.在区域 A2 中,A2_1、A2_2、A2_3 分别反映了水平方向、垂直方向、对角方向的频率变化.图像中水平、竖直、对角方向的边缘、轮廓、纹理等表现为 A2_1、A2_2、A2_3 区域中的频率系数幅值较大,因此适当对 A2 区域中的大系数进行处理,便可得到图像的大致轮廓.

设 (x, y) 表示图像区域中象素位置的坐标,将 A2_1、A2_2、A2_3 区域内同一位置的系数中的最大值组成一个与 A2_1 区域同样大小的 $I(x, y)$ 图像块,即:

$$I(x, y) = \max\{A2_1(x, y), A2_2(x, y), A2_3(x, y)\} \quad (1)$$

这样,图像块 $I(x, y)$ 就包含了图像中各个方向的频率成分,只要对 $I(x, y)$ 图像块进行简单的量化处理,就基本上能得到图像的大致轮廓.其量化方法如下:

$$M1 = \max\{I(x, y)\}, M2 = \min\{I(x, y)\}, M3 = Ave\{I(x, y)\}$$

$$M = (M1 - M2)/3$$

则有

$$\text{If } I(x, y) > M3 + M \text{ or } I(x, y) < M3 - M \text{ then}$$

$$I(x, y) = 255$$

$$\text{Else } I(x, y) = 0$$

End If

上面对 A2 区域的处理就得到了图像的大致轮廓 $I(x, y)$,对 A3 区域也进行同样的轮廓提取,而对于 A4 区域(即 DC 图)则采用简单的 Roberts 算子提取大致轮廓.通过对 A2、A3、A4 区域的处理可得如图 5 所示的大致轮廓图,它是图 4(c)在 DCT 域下的多分辨率大致轮廓图.得到了图像的大致轮廓后,就可根据此轮廓图特征进行压缩域的图像检索,下面主要讨论轮廓特征的匹配算法.



图 5 DCT 域的图像多分辨率大致轮廓图

4 建立轮廓的连通直方图

轮廓特征的匹配算法有许多,这里提出了一种基于连通直方图的方法,它具有灰度、旋转、平移等不变性的特性,并且计算比较简单.

4.1 定义

对于一给定分辨率为 $X \times Y$ 的轮廓图像 $I(x, y)$,首先计算其连通区域,再计算具有相同连通区域面积(用象素点数表示)的概率.连通直方图 $H^*(k)$ 可通过下式计算:

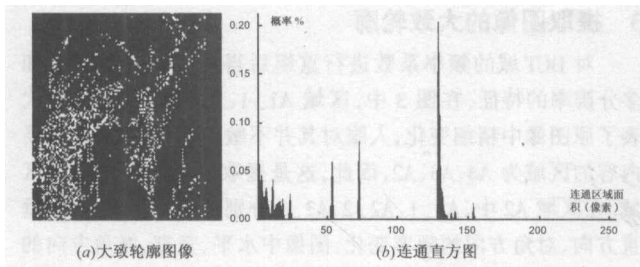


图 6 大致轮廓图像的连通直方图

$$H^*(k) = \sum_{l(a,s)} \delta^k \quad (2)$$

$$\text{其中 } \delta^k = \begin{cases} 1, & \text{如果 } C = k \\ 0, & \text{否则} \end{cases}$$

k 表示连通区域的面积, $k = 1, \dots, M; M \leq X \times Y/S$, S 是连通区域面积的量化步长. C 表示某个连通区域的面积.

在实际应用中,为了消除图像尺寸的影响,要对 $H^*(k)$ 进行归一化处理,得到归一化的连通直方图曲线 $H(k)$:

$$H(k) = H^*(k)/A \quad (3)$$

其中 $A = \sum_k H^*(k)$. 图 6(a) 是一幅大致轮廓图像, 图 6(b) 是其连通直方图.

4.2 权函数

对于一般图像的大致轮廓,较大的连通区域更能引起视觉注意,它主要集中在连通直方图的右侧.而对于那些点状噪

子图像是未经压缩的,因此要对其进行 DCT 变换等处理,以期得到与 JPEG 码流一致的数据.对于 JPEG 图像库中的压缩数据,熵解码后进行 DCT 系数的重组,经处理得到大致轮廓后进行连通直方图统计,并与例子图像的连通直方图进行比较.此时的结果只是基于图像的轮廓检索得到的,并未用到图像的其他信息(如灰度等).为了使检索结果更加符合人

眼视觉特性,这里采用了基于灰度的排序方法,即采用 DC 图的灰度直方图对检索结果进行排序,尽量使灰度相近并且具有相似轮廓的图像排列在一起,使其最终输出更具准确性.

图 7 中的距离测量均采用计算简单、效果较好的 L1 距离^[1,2],设图像 p 的直方图为 $[H(k)]_p$, 图像 q 的直方图为 $[H(k)]_q$, 则两直方图之间的距离为:

$$d(p, q) = \sum_k | [H(k)]_p - [H(k)]_q | \quad (5)$$

由于是在压缩域进行,处理时需要考虑一些特殊情况,文献[10]中的 3.4 节对此有详细的讨论,这里不再赘述.

6 实验及结论

为了验证本文的检索算法,我们作了大量的实验.

在实验中,计算连通直方图的各种参数如下:

(a) 根据各区域轮廓的重要性,在 A4 的连通直方图中取

音与高频纹理来说,虽然在直方图中所占比例较大,但是对图像的内容却贡献较少!因此,这里引进一个权函数 $W(k)$,来反映各种大小的连通区域对图像内容的贡献程度,此时的加权直方图 $H_w(k)$ 为:

$$H_w(k) = H(k) \times W(k) \quad (4)$$

对于不同类型的图像,可设计不同的权函数 $W(k)$.例如对于高频纹理图像,一般强调较小的连通区域,而对于普通图

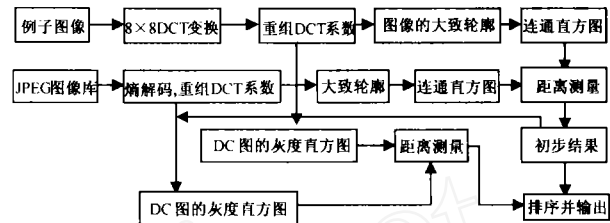


图 7 基于连通直方图的图像检索框图

像则强调较大的连通区域.

5 基于压缩域的图像检索

根据上面的讨论,设计了一种对压缩图像不进行完全解码的情况下实现图像检索的算法,采用例子图像查询(QBE)方式,其具体的检索流程如图 7 所示.在图 7 的检索框图中,假设例



图 8 JPEG 图像库中的部分图像

$M = 256$, 而其它两个分辨率的连通直方图的 $M = 128$.

(b) 对于各个分辨率轮廓的连通直方图,公式(2)中连通区域面积的量化步长 $S = 3$.公式(4)中的权函数 $W(k)$ 为线性函数:

$$W(k) = (50/M)k + (1/50) \quad (6)$$

对三个连通直方图加权后,再组合成一个 512 维的直方图进行处理.实验均采用 512×512 大小的灰度图像,共 500 幅.其中一些是经过放大、缩小、旋转、改变光照等获得的,图 8 是几幅典型的图像.

例子图像直接从图像库中选取,图 9 为检索结果的几组实例,从左到右依次按相似性大小排列,其中第一幅是例子图像.在本实验中平均的检索正确率为 73.4%.

本文提出了一种基于 DCT 压缩域的图像检索方法,对于以 DCT 为编码基础的压缩图像(如 JPEG 等),可以在不完全

解码(只进行熵解码)的情况下重组 DCT 系数得到具有方向性和多分辨率的图像特性,并提取图像多分辨率下的大致轮廓,用连通直方图的方法统计其特征来进行图像检索,最后利用 DC 图的灰度特征对结果进行排序,使其结果更符合人眼的视觉特性。这种检索技术无需预先形成额外的特征库,因此对图像数据库的动态管理非常方便,特别利于进行基于互连网的图像检索与查询。

由于连通直方图是对图像轮廓连通区域进行统计,故不受图像中物体的光照、旋转及其它一些几何变化的影响,有利于图像检索,实验结果显示这种基于压缩域的图像检索算法具有较好的效果。本文的实验主要是针对灰度图像,对于彩色图像,可先对 Y 分量进行处理,而在最后的结果排序中综合 Cr 、 Cb 的 DC 图的灰度直方图就可以了。

参考文献:

- [1] 黄祥林,沈兰荪.基于内容的图像检索技术研究[J].已被电子学报录用.
- [2] F Idris, S Panchanathan. Review of image and video indexing techniques [J]. Journal of visual communication and image representation, 8(2):146 - 166.
- [3] Y A Aslandogan, C T Yu. Techniques and systems for image and video retrieval[J]. Knowledge and Data Engineering, 1999, 111:56 - 63.
- [4] 沈兰荪,魏海等.压缩域图像处理技术研究[J].北京工业大学学报,2000,26(3):24 - 32.
- [5] 魏海,沈兰荪.基于分类矢量量化的图像压缩和检索算法[J].电子学报,2001,29(7):933 - 936.
- [6] J A Lay, G Ling. Image retrieval based on energy histograms of the low frequency DCT coefficients [J]. Acoustics, Speech, and Signal Processing, 1999, 6: 3009 - 3012.
- [7] A Vellaikal, C J Kuo. Joint spatial-spectral indexing for image retrieval [J]. Image Processing, 1996, 3: 867 - 870.
- [8] 黄祥林,沈兰荪.一种具有旋转不变性的压缩域纹理图像分类方法[J].电子与信息学报,已录用.

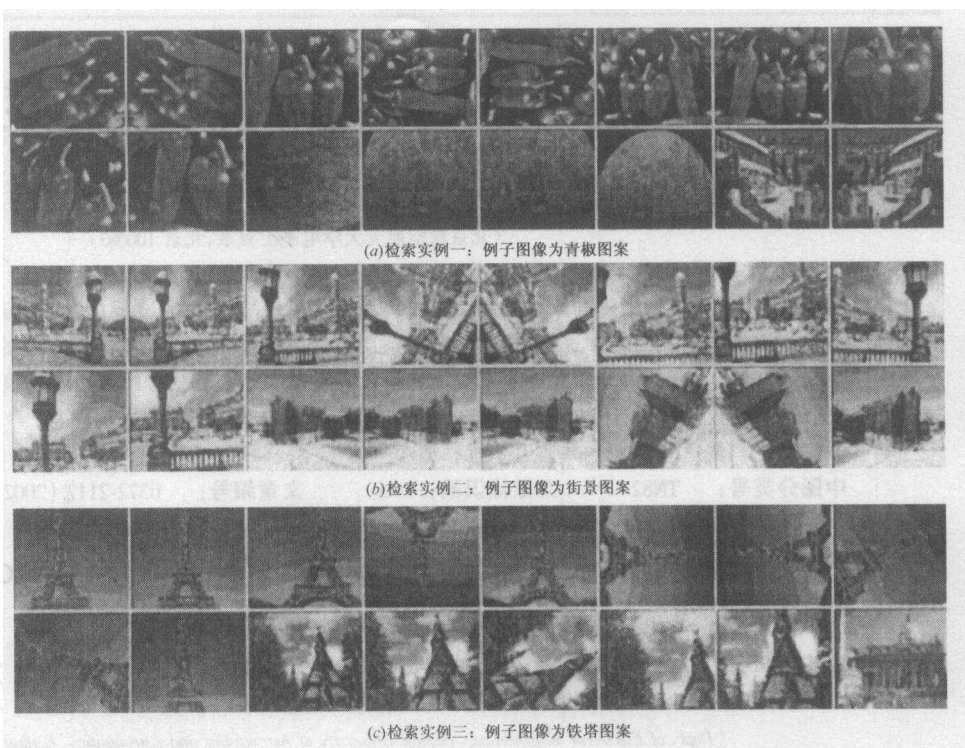


图 9 基于 DCT 压缩域的图像检索实例

- [9] 沈兰荪,魏海等.基于子带/小波分解的视频编码可分级性研究[J].电子学报,2000,28(7):126 - 127.
- [10] 黄祥林,沈兰荪.基于 DCT 压缩域的纹理图像分类[J].电子与信息学报,2001,23(11):1 - 6.

作者简介:



黄祥林 男,1967年5月出生于湖南省石门县.博士,现在北京工业大学工作.主要研究方向为基于内容的图像检索、压缩域图像/视频处理等.曾获国家科技进步三等奖、地矿部科技成果二等奖各一次,已发表论文20余篇.



宋 磊 女,1974年10月出生于河南省郑州.1998年毕业于北京工业大学自动化系,现为北京工业大学信号与信息处理专业硕士研究生,主要研究方向为图像处理、基于内容的图像检索等.