

基于块边缘特征直方图的图像检索^{*})

邱开金 肖国强 张为群

(西南大学计算机与信息科学学院 重庆 400715)

摘 要 本文提出一种利用块边缘特征来进行图像检索的方法。首先对图像中每个 8×8 的块分别提取其块边缘特征,利用块边缘特征构建三维或四维直方图进行图像检索。所有数据在 DCT 压缩域中进行处理,从而降低了计算复杂度。实验结果表明,该方法与传统的基于 DCT 压缩域的图像检索方法相比,具有检索准确度高和检索速度快的特点。

关键词 块边缘特征,块边缘直方图,基于内容的图像检索,DCT

An Image Retrieval Algorithm Based on Block-Edge-Pattern Histogram

QIU Kai-Jin XIAO Guo-Qiang ZHANG Wei-Qun

(Faculty of Computer & Information Science, Southwest University, Chongqing 400715)

Abstract In this paper, we propose a new image retrieval algorithm based on block edge patterns. Block edge patterns are extracted respectively from each 8×8 block of image. Then a 3-D or 4-D histogram is constructed by utilizing block edge patterns to retrieve images inside image databases. Because the edge orientation of each block can be estimated in discrete cosine transform (DCT), it will reduce the complexity of evaluating edge orientation. In comparison with existing representative techniques, the experimental results show the superiority of the proposed algorithm in terms of retrieval precision and processing speed.

Keywords Block edge patterns, Edge pattern histograms, Content-based image retrieval, Discrete cosine transform (DCT)

随着多媒体和互联网的高速发展,数字图像的数量也在以惊人的速度增长,如何在这海量的图像中快速而又准确地搜索所需要的图像,一直是人们普遍关注的问题,因此凡是能提高图像查询速度和查询准确度的图像检索算法都具有重要的实用价值。

目前,由于时间和空间的限制,大量的图像都以压缩方式存储和传输,其中 JPEG 是最常用的图像压缩格式。直接从 DCT 压缩域提取特征可以减少系统的处理时间和计算复杂度,从而提高图像的检索速度,因此,直接从 DCT 域提取特征进行图像检索是目前基于内容的图像检索技术的研究热点^[1~7]。从 DCT 域提取特征的方法很多,但近年来这方面的研究主要集中在纹理、颜色、形状等方面。这些特征能够比较有效地表征图像的内容,但总体性能和速度仍有待提高。

图像的边缘信息是人类理解图像内容的重要线索,在图像分析和处理中有重要价值,使用边缘信息来查询图像能提高图像的检索性能。直接从 DCT 域中提取边缘信息,传统的方法是采用卷积和滤波^[6],因此,其计算量很大,如果将其应用于图像检索,则会影响检索速度。

MPEG-7 标准中定义了一个边缘直方图描述子(histogram descriptor)^[9],该描述子统计水平方向、垂直方向、45 度方向、135 度方向及无方向五种边缘特征。目前,已有文献报导直接从 DCT 系数中提取块边缘特征的研究成果,如 Chang 等人^[8]提出了一种快速地提取 DCT 域中每个块的边缘特征的方法,他把块边缘特征分为五类:无边缘、水平边缘、45°边

缘、垂直边缘、135°边缘,并且提取这五种块边缘特征的计算复杂度也较低。但直接利用这五种块边缘特征进行图像查询,则图像特征信息太少,检索性能差,所以很少有人直接用这五种块边缘特征检索图像。

针对这些问题,本文提出一种基于块边缘特征直方图的图像检索算法。首先提取图像的三种块边缘模式,它们分别是无边缘、水平边缘、垂直边缘。其次,以相邻三个块作为整体来构造三维直方图,或以相邻四个块作为整体来构造四维直方图。算法的基本思想是通过增加更多的位置信息和空间信息来提高图像查询的查准率。实验结果表明,利用本文提出的算法进行图像检索具有较高的图像查询速度和查准率。

本文第 1 节介绍如何提取图像块的边缘特征;第 2 节详细描述基于块边缘特征直方图的图像检索算法,包括如何构造块边缘特征的三维和四维直方图;第 3 节给出了实验结果,并对其进行了分析;最后对全文进行总结。

1 块边缘特征的提取

1.1 块边缘特征数目的确定

文[8]中,Chang 等人提出了用五种块边缘特征来表示图像的边缘,这五种块边缘特征能够较好地表征图像的边缘性质,表现图像的内容。但是根据我们的研究发现,适当地减少块边缘特征的种类,对图像边缘的表征并不会产生明显的影响,从而可以降低计算复杂度,提高查询效率。

我们先看这样一个现象:如图 1 所示,将一幅图像的边缘

^{*})重庆市科委自然科学基金资助项目(CSTC-2004BB8481);教育部“春晖计划”资助项目(Z2004-55011)。邱开金 硕士生,研究方向为图形图像处理及信息安全;肖国强 教授,主要研究方向为信号与信息处理和无线网络通信;张为群 教授,主要研究方向为软件工程。

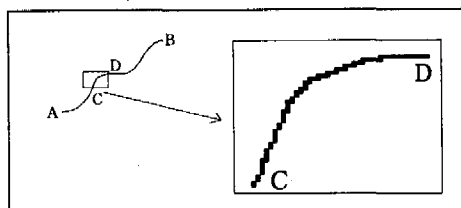


图1 图像的边被放大8倍的示意图

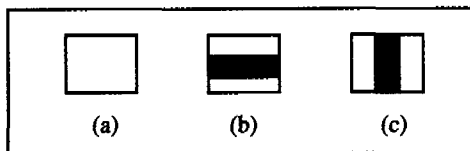


图2 三种边缘特征(a)NE(b)HE(c)VE

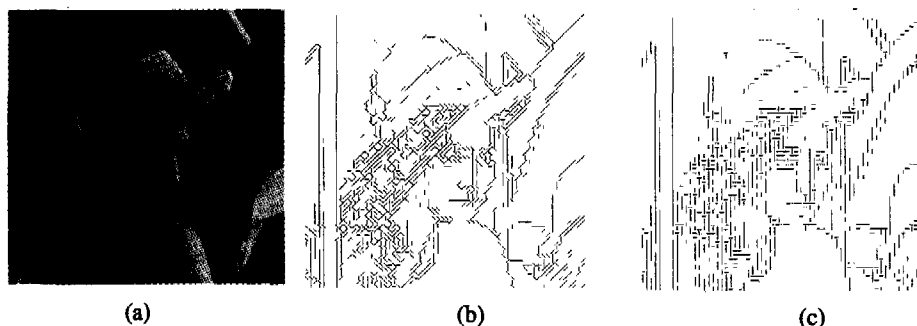


图3 (a)Lena 图像 (b)五种边缘特征构建的边缘图像 (c)三种边缘特征构建的边缘图像

1.2 三种块边缘特征的提取

Chang 等人^[8]提出了一种提取五种块边缘特征的算法,该算法根据 DCT 系数的特点,对计算公式进行了简化,从而得到一组提取五种块边缘特征的简单计算公式。本文将以文^[8]的公式为基础来提取图像的三种块边缘特征,算法的伪码描述如下。

```
for each block
input  $X(i, j)$ ;
 $\delta E = |X(0, 1)|$ ;  $\delta HE = |X(1, 0)|$ ;
if ( $\text{Max}(\delta VE, \delta HE) < \lambda$ ) then output ('NE');
else
    if ( $\delta VE \geq \delta HE$ ) then output ('VE');
    else output ('HE');
```

其中, $X(i, j)$ 是 8×8 块的 DCT 系数, $i, j = 0, 1, \dots, 7$ 。 δVE 和 δHE 是用于判断块边缘特征的变量, λ 是用户设定的阈值,实验结果表明 λ 设定为 50 较合适。

2 算法描述

如何利用块边缘特征作为特征值来表示图像的内容,从而进行基于图像内容的检索,传统的作法是利用直方图。直方图是表示随机变量概率分布的好方法,不管什么分布,或分布的维数如何都可用直方图很好地表示,如常用的颜色直方图、灰度直方图等,这里我们用三种块边缘特征来构建直方图。

为直观起见,我们用符号“N”、“H”、“V”分别表示一个 8×8 块的无边、水平和垂直边缘特征。块边缘特征直方图用函数 $h_j(c_1, c_2, \dots, c_j, \dots, c_n)$ 表示,其中 c_j 是第 j 类块边缘特征在一幅图像中出现的频率, $j = 1, 2, \dots, n$ 。 n 为块边缘特征总数。用 $V_r^w(m)$ 表示从 m 个元素中取 r 个元素的可重复选排列数,用下式计算

$$V_r^w(m) = m^r \quad (1)$$

放大后,可以看到图像的边呈锯齿状,这些锯齿由水平方向和垂直方向的模块组成。因此我们尝试只用水平和垂直方向的块来构造图像的边,即只用三种块边缘特征来表示一幅图像的边,如图 2 所示,这三种块边缘特征分别是无边 (NE)、水平边缘 (HE) 和垂直边缘 (VE)。

图 3 是分别用三种块边缘特征和文^[8]提出的用五种块边缘特征,对同一幅图像提取边缘图像进行对比的实验结果。从图中可以看出,这两种方法所提取的图像边缘的轮廓差异非常小,使用三种边缘特征构建的边缘图像,与使用五种边缘特征一样,都能反映图像的边缘性质。因此本文采用三种块边缘特征来构建图像边缘,从而进一步构造出图像的直方图特征向量进行图像检索。后面的实验结果证明,这种方法是可行的。

由图像的块边缘特征构建直方图的方法如下:

一维块边缘特征直方图 (One-Dimension Block Edge Histogram, 1D-BEH) 是以一个图像块为单位,对块边缘特征进行统计,因此,可以得到 3 种块边缘特征的组合模式。即从集合 $s = \{N, H, V\}$ 中,取 1 个元素进行有重复的选排列,其排列数为 $n = V_1^3(3) = 3^1 = 3$ 。同理,二维块边缘特征直方图 (2D-BEH) 是以两个相邻图像块为单位,对块边缘特征进行统计,可得到 9 种块边缘特征的组合模式。

三维块边缘特征直方图 (3D-BEH) 是以相邻三个块为单元,对块边缘特征进行统计,如图 4(a) 所示。可以得到的块边缘特征的组合模式数即为从集合 s 中取 3 个元素进行有重复的选排列的数目,于是由 (1) 式可知, $n = V_3^3(3) = 3^3 = 27$ 种,即 $C_1 = 'NNN'$, $C_2 = 'NNH'$, $C_3 = 'NNV'$, $C_4 = 'NHN'$, $C_5 = 'NHH'$, $\dots$, $C_{26} = 'VVH'$, $C_{27} = 'VVV'$ 。通过统计一幅图像中 C_j ($j = 1, 2, \dots, 27$) 出现的频率,可构成一个 27 个元素的 3D-BEH。

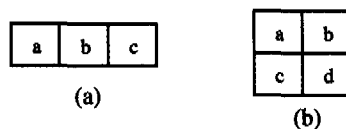


图4 (a)3D-BEH 水平排列 (b)4D-BEH 块的排列

同理,四维块边缘特征直方图 (4D-BEH) 是以相邻四个块为单元,对块边缘特征进行统计,如图 5(b) 所示。这里 $r = 4$ (a, b, c, d 四块),于是可以得到的块边缘特征的组合模式为 $n = V_4^3(3) = 3^4 = 81$ 种,即:

$$C_1 = \begin{matrix} N & N \\ N & N \end{matrix} \quad C_2 = \begin{matrix} N & N \\ N & H \end{matrix} \quad C_3 = \begin{matrix} N & N \\ N & V \end{matrix} \quad C_4 = \begin{matrix} N & N \\ H & N \end{matrix}$$

$$C_3 = \begin{matrix} N & N \\ H & H \end{matrix} \quad C_6 = \begin{matrix} N & N \\ H & V \end{matrix} \cdots \cdots C_{80} = \begin{matrix} V & V \\ V & H \end{matrix} \quad C_{81} = \begin{matrix} V & V \\ V & V \end{matrix}$$

通过统计一幅图像中 $C_j (j=1,2,\cdots,81)$ 出现的频率,从而形成一个 81 个元素的 4D-BEH.

表 1 不同情况下块边缘特征的组合模式数目

	1D-BEH	2D-BEH	3D-BEH	4D-BEH	5D-BEH	6D-BEH
三种块边缘特征	$3^1=3$	$3^2=9$	$3^3=27$	$3^4=81$	$3^5=243$	$3^6=729$
五种块边缘特征	$5^1=5$	$5^2=25$	$5^3=125$	$5^4=625$	$5^5=3125$	$5^6=15625$

从表 1 看出,1D-BEH 和 2D-BEH 的块边缘特征的组合模式数目较少,有利于降低算法的复杂度,但由于位置信息量太少,图像查询的准确性不高.而四维以上的块边缘特征虽位置信息多,但它的块边缘特征的组合模式数目很大,所以选用三维或四维块边缘特征直方图较为合适.

从表 1 还可以看出,用文[8]中的五种边缘特征来构造直方图,那么它的各种边缘特征直方图的组合模式数目大大超过用三种边缘特征的数目.因此选用三种块边缘特征来构建直方图,可以减少块边缘特征的组合模式数,降低计算量,从而提高图像的查询速度.

把 3D-BEH 和 4D-BEH 进行归一化处理后,作为图像查询或检索的特征向量.对于查询图像 q 和图像库中的图像 x ,两者的图像特征向量的距离用下式计算

$$d(x,q)=\sum_{j=1}^n |k_x(j)-k_q(j)| \quad (2)$$

其中, $k_x(j)$ 和 $k_q(j)$ 分别为图像 x 和 q 对应的归一化特征向量, n 是特征向量的维数.图像查询或检索的过程即是寻求查询图像与图像库中图像的特征向量之间距离的最小值.

3 实验结果与分析

文[7]中提出一种从 DCT 压缩域提取纹理特征进行图像查询的方法.它在每个 8×8 块的 DCT 系数中,计算出均值和方差,然后把均值分成互不重叠的 4 个区间,把方差分成互不重叠的 7 个区间,两者结合从而构成一个 28 维的特征向量去查询图像.其均值(m_1)和方差(σ)按下式计算

$$m_1 = \frac{1}{8} X(0,0) \quad (3)$$

$$\sigma = \sqrt{\frac{1}{64} \left(\sum_{u=0}^7 \sum_{v=0}^7 X^2(u,v) - X^2(0,0) \right)} \quad (4)$$

为了评估本文提出的算法的性能,我们建立了 10200 幅图像的数据库,其中包括自然风景、汽车、人物、花草、动物、物品、卡通人物等.库中的图像均为 24 位真彩色图像,采用 jpg 格式.

实验过程如下:先从 10200 幅图像中抽取 100 幅作为查询图像,这 100 幅查询图像由汽车、人物、花草、风景、动物各 20 幅构成.用查询图像在图像库中进行检索,并取出与查询图像的特征向量之间距离最小的 12 幅图像作为检索的结果.对每一幅查询图像的检索结果,用下式计算图像的查准率 (Precision)^[10]

$$\text{Precision} = \left(\sum_{q=1}^k n_q \right) / M \quad (5)$$

其中, n_q 是在一次检索结果中与查询图像相关的图像数, k 为每一类查询图像的总数, M 是每一类图像检索出的图像总数,实验中 $M=12 \times 20=240$.

用文[7]的算法和本文提出的算法分别进行图像查询的实验,并用(5)式计算其各自的查准率,实验结果如表 2 所示.

从表 2 中可以看出,采用块边缘特征的 3D-BEH 和 4D-BEH 算法,比文[7]的算法的图像检索的平均查准率分别提高 32.4%和 35.3%.采用 4D-BEH 比采用 3D-BEH 算法的

通过对块边缘特征直方图的建立过程可以看出,对不同数目的块边缘特征构成不同维数的直方图时,得到的块边缘特征的组合模式数目将有很大的变化,如表 1 所示.

平均查准率约有提高.由此可见,在图像的特征向量中增加图像的位置信息和空间信息后,使图像的检索精度得到明显提高.

表 2 查准率

	汽车	花草	人物	风景	动物	平均查准率
文[7]的算法	0.17	0.17	0.74	0.38	0.25	0.34
3 D EBH	0.67	0.25	0.74	0.33	0.24	0.45
4 D EBH	0.64	0.27	0.76	0.39	0.22	0.46

本文提出的算法中,从 DCT 域提取三种块边缘特征不需要乘法和加法的计算,仅需进行绝对值运算,而文[7]中计算均值和方差如式(3)和(4)所示,它需要 67 个乘法和 128 个加法.因此直接从 DCT 域提取三种块边缘特征比文[7]中提取图像特征的算法的复杂度低,从而用本文提出的算法进行图像检索时,其检索速度将得到大幅度的提高.

总结 目前,对互联网和动态数据库中图像查询的需求与日俱增,提高图像查询速度和准确性是图像检索研究的一个重要方面.从实际应用出发,本文提出了一种有效的图像查询算法,该算法与其它传统的从 DCT 压缩域提取图像特征的算法相比,具有极低的计算复杂度和较高的图像检索精度和查询速度,有效地提高图像检索的性能,适合于互联网和动态数据库的图像查询.

由于用块边缘特征直方图构造的图像特征向量,仅利用了图像块的局部相对位置信息,未能包含图像块的全部位置信息,因此,图像检索精度还不够理想.为了进一步提高图像的检索精度,在构造图像特征向量时,应增加更多的图像块的相对位置信息和空间信息,这也是我们在基于内容的图像检索方面进一步的研究工作.

参 考 文 献

- Shneier M, Abdel-Mottaleb M. Exploiting the JPEG compression scheme for image retrieval. IEEE Trans. Pattern Anal. Mach. Intell, 1996, 18(8): 849~853
- Chang Shih-Fu. Compressed domain techniques for image/video indexing and manipulation. IEEE International Conference on Image Processing, 1995. 314~317
- Ngo Chong-Wah, Pong Ting-chuen. Exploiting image indexing techniques in DCT domain. Pattern Recognition, 2001, 34: 1841~1851
- Lay Jose A, Ling Guang. Image retrieval based on energy histograms of the low frequency DCT coefficients [A]. In: Proc of International Conference on Acoustics, Speech and Signal Processing [C], Phoenix, Arizona, USA, 1999, 6: 3009~3012
- Sim D G, Kim H K, Park R H. Fast texture description and retrieval of DCT-based compressed image [J]. Electronic Letters, 2001, 37(1): 18~19
- Kim T Y, Han J H. Model-based discontinuity evaluation in the DCT domain [J]. Signal Processing, 2001, 81(4): 871~882
- Feng Guocan, Jiang Jianmin. JPEG compressed image retrieval via statistical features. Pattern Recognition, 2003, 36: 977~985
- Chang H S, Kang K. A compressed domain scheme for classifying block edge patterns. IEEE Transactions on image processing, 2005, 14(2)
- Park D K, Jeon Y S, Won C S. Efficient use of local edge histogram descriptor [A]. In: Proc. of the ACM Workshops on Multimedia [C], Los Angeles, [s. n.], 2000, 51~54
- Lee H Y, Lee H K, Ha Y H. Spatial color descriptor for image retrieval and video segmentation. IEEE Transaction on Multimedia, 2003, 5(3)