

基于前景提取的复杂背景图像检索算法

冯喆 夏虎 傅彦 周俊临

(电子科技大学计算机科学与工程学院互联网科学中心 成都 611731)

摘要 基于内容的图像检索提供给使用者一种更直观、更精准的检索方式。用户在进行此类检索时,往往更关注图像的主体部分。为了消除背景信息对检索效果的影响,提出一种基于前景提取的复杂背景图像检索算法。实验证明,在 H-S 颜色直方图、LBP 纹理特征以及颜色纹理混合特征上,该算法可以得到较优化的性能。

关键词 前景提取,图像检索,颜色直方图,纹理特征

中图分类号 TP391 文献标识码 A

Complex Background Image Retrieval Based on Foreground Extraction

FENG Zhe XIA Hu FU Yan ZHOU Jun-lin

(Web Sciences Center, School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China)

Abstract Content-based image retrieval provides users a more intuitive and more accurate retrieval method. Users in such retrieval, tend to be more concerned about the main part of the image. In order to impair the influence of the background information on retrieval result, this paper presented an algorithm of complex background image retrieval based on foreground extraction. The experiment indicates that on H-S color histogram, LBP texture feature, and color-texture mixed characteristic, this algorithm can obtain more optimized performance.

Keywords Foreground extraction, Image retrieval, Color histogram, Texture features

1 引言

随着互联网技术和电子商务的迅速发展,数字图像每日都在以惊人的速度增长,如何在日益增长的海量图像中找到用户所感兴趣的图像成为当前研究的热门问题。传统的基于文本的图像检索方式忽略了图像自身包含的各种信息,用户并不能通过几个关键词准确地描述他们所想要查找的图像,基于内容的图像检索(CBIR)应运而生。

用户在进行基于内容的图像检索的时候,往往关注的是图像中的主体部分,如模特身上的某件衣服,桌上的某个杯子等。然而复杂的背景信息对颜色、纹理等特征向量取值都会有比较大的影响,从而影响检索的准确性。显然,只针对前景部分做特征提取和匹配能提高检索的准确率。目前,大多数的前景提取方法都需要人工进行交互,这对于图像库中海量的图像是行不通的。由 Rother 等人提出的 Grab cut 算法^[1]是在 Graph cuts 算法^[2,3]基础上进行的改进,它只需较少的用户交互即可得到较满意的分割效果。本文在 Grab cut 的基础上,提出一种基于前景提取的复杂背景图像检索算法,此方法无需人工交互即可对图像进行前景提取,之后针对前景部分提取颜色、纹理等特征并进行相似度计算和排序。实验

证明在 H-S 颜色直方图、LBP 纹理特征以及颜色纹理混合特征上,该算法可以得到较优化的性能。

2 相关研究

20 世纪 90 年代人们提出了基于内容的图像检索(CBIR)方法,与基于文本的图像检索不同的是,它利用了图像本身的内容,检索质量依赖于图像与图像之间内容的相关性。

基于内容的图像检索中最重要的一环是图像的特征提取。从人类视觉的角度看,颜色是人类感知物体最直观的特征,Rao 等人提出的基于颜色直方图的方法^[4]反映了图像颜色的统计分布,描述的是图像整体的颜色特征。纹理特征^[5]是物理表面共有的内在特性,用来表示物体表面组织结构的排列。Tamura 等人^[6]提出了一种纹理特征的表达方法,它包含 6 个分量,对应于纹理特征的 6 种属性,即粗糙度(coarseness)、对比度(contrast)、方向度(directionality)、线像度(line-likeness)、规整度(regularity)和粗略度(roughness)。

图像的全局统计特征虽然在一定程度上符合部分用户的检索意图,然而大多数情况下,用户并不关心图像全局相似与否,而更多关注图像中具有语义的区域。周静等人^[7]提出的基于用户感兴趣区域的检索方法,需要用户手工选择关

到稿日期:2013-05-21 返修日期:2013-06-20 本文受国家自然科学基金(61103109,11105024,61003231),中央高校基本科研业务费(ZYGX2011J057,ZYGX2012J071,ZYGX2012J085),四川省科技项目(2010HH0002,2011GZ0106,2011Z20001,2012RZ0002,2012RZ0003)资助。

冯喆(1989-),男,硕士生,主要研究方向为数据挖掘、图像检索;夏虎(1981-),男,博士,助理研究员,主要研究方向为数据挖掘、复杂网络,E-mail: xiahu@uestc.edu.cn(通信作者);傅彦(1969-),女,教授,博士生导师,主要研究方向为数据挖掘、异常检测;周俊临(1981-),男,博士,副教授,主要研究方向为数据挖掘、异常检测、推荐系统。

注的区域,再利用一个滑动窗口在图像库中的每张图片上选择一定的区域进行比对。这种方法虽然提高了准确率,但是图像库中的每一张图片都分成了多个区域,增大了计算量,同时由于需要用户的交互也使系统更加繁琐。

3 基于前景提取的图像检索

Grab cut 算法^[1]是一种基于图割理论的前景提取算法。本文提出的基于前景提取的图像检索改进了 Grab cut 算法,提取过程中无需用户的交互,之后在此基础上提取图像特征,并与图像库中所有图像的特征进行相似度计算和排序,最终得到检索结果。

3.1 图割理论

图像的分割问题实际上是图像中每个像素的二值化标号问题。二值化向量 $A = (A_1, A_2, A_3, \dots, A_{|P|})$ 中每一维代表的是该像素的取值, P 是所有像素点的集合,“bkg”代表的是背景标号,“obj”代表的是前景标号。文献^[2]给出了一个能量泛函 $E(A)$:

$$E(A) = \lambda R(A) + B(A)$$

其中

$$R(A) = \sum_{p \in P} R_p(A_p)$$

$$B(A) = \sum_{\{p, q\} \in N} B\{p, q\} \cdot \delta(A_p, A_q)$$

$$\delta(A_p, A_q) = \begin{cases} 1, & A_p = A_q \\ 0, & A_p \neq A_q \end{cases}$$

其中, N 表示的是 P 中相邻像素点对的集合。 $R(A)$ 表示的是图像分割的区域信息 (regional term), 它的含义是每个像素点赋予标号“bkg”或者“obj”的代价。而 $B(A)$ 表示的是分割的边界信息 (boundary term), $B\{p, q\}$ 代表相邻点对 $\{p, q\}$ 不连续所付出的代价。当像素点 p, q 相似的时候, $B\{p, q\}$ 很大, 反之 $B\{p, q\}$ 趋近于 0。如此, 图像分割转换成用组合优化的方法最小化能量泛函 $E(A)$ 的问题。通过构造一个带权值的图, 采用图论中的最大流/最小割理论^[8,9] 可以得到 $E(A)$ 最小化的最优解。

3.2 基于前景提取的图像检索算法

Grab cut 算法^[1]是在图割基础上的一次改进, 它通过多次迭代计算提高了图像的分割精度, 并且降低了用户的交互量。但是我们要对图像库中大量的图像做前景提取, 每一张都靠人工去标记前景和背景区域是不可取的。我们提出了一种无交互的 Grab cut 方法, 它可以模拟用户人工标记的阶段。

图像的前景一般集中于图像的中间部分, 我们从位于 4 个角的像素点出发, 向内部进行单点扩散。算法分为统计阶段和扩散阶段。统计阶段首先取图像 4 个角的一块区域作为样本区间。对每个样本区间, 将第一个像素点作为第一类, 记作 C_1 , 并把这个像素的 RGB 颜色分量值作为 C_1 的特征值。依次遍历此区间的每一个像素点, 对每个点而言, 计算它与类集合 C 中每个类的 RGB 特征值的差值, 若它和类 C_k 的差值在阈值以内, 就把它归为此类, 并把 C_k 的计数加 1, 反之则新建一个类 C_{n+1} , 以此像素点的 RGB 颜色分量作为 C_{n+1} 的特征值。最后取类集合 C 中计数最靠前的 5 个类作为统计结果。

传统的 Grab Cut 算法^[1]需要用户手工对图像的掩码 MASK 进行标记, 我们的算法在扩散阶段会进行自动标记。此阶段仍然从 4 个角的第一个像素点开始, 默认第一个点具有背景标号“bkg”, 对于它扩散方向上相邻的点, 判断 RGB 颜色分量是否和它本身或者和此区域统计阶段的 5 个类特征值的差值在阈值以内, 若是, 则把这个点也标记为“bkg”, 反之不做任何操作。扩散阶段一直到访问到所有的像素点为止。图 1 展示了从左上角开始的扩散过程, 箭头代表扩散的方向。

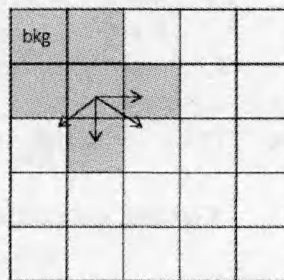


图 1 从图像左上角开始的扩散

对掩码 MASK 标记好之后, 就可以用 Grab cut 算法进行前景提取了 (如图 2 所示)。图 2 中 (a)~(d) 都具有较好的提取效果, (e) 的提取效果则不理想, 但是我们将其用作图像检索时, 和未处理的图像相比, 由于前景未被破坏或破坏较少, 因此并不会降低检索的准确率, 第 5 节的实验也证明了这一点。

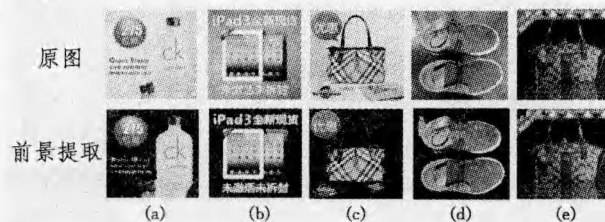


图 2 对图像进行前景提取

对于图像库中的每一幅图像, 都采用上述的方法进行前景提取, 之后对前景部分进行特征提取生成特征向量并保存。对于待查图像, 也在前景提取之后提取特征向量, 并与图像库中所有图像的特征向量进行相似度计算并排序, 最后返回前 N 幅图像。算法流程如下:

```

输入: 待查图像 A
foreach A 的 4 个角 1/20×1/20 区域
    根据 RGB 统计出像素最多的 5 个类, 记为 C=(C1, C2, C3, C4, C5)
foreach A 的角点 p
    标记 p 为“bkg”
    for (p 不是最后一个像素点; p=p 的下一个像素)
        if (p 被标记“bkg”)
            if (p 扩散方向上相邻点 q 与它的 RGB 差在阈值内 or q 与 C 的特征 RGB 差在阈值内)
                then 标记 q 为“bkg”
用 Grab cut 算法对 A 提取前景, 记为 A'
对 A' 提取特征向量 H, 计算 H 与图像库中图像特征的相似度, 记为 (Sim1, Sim2, ..., SimN)
Sort (Sim1, Sim2, ..., SimN)
Return 前 n 张图片
    
```

4 特征提取

4.1 H-S 颜色直方图

颜色是图像最重要的底层特征,也是 CBIR 系统中使用最多的特征。HSV 颜色模型相对于最常见的 RGB 模型而言,更符合人眼的直观感觉,其中 H 代表色调(Hue),范围是 0 到 360,S 代表饱和度(Saturation),范围是 0 到 1,V 代表亮度(Value),范围也是 0 到 1。

本文中只提取 H 和 S,不提取 V,主要是由于有光照的影响,同个物体拍出不同的亮度可能性很大,因此亮度并不适合用来区分不同物体。我们将 H 平均分为 16 份,S 平均分为 8 份(如图 3 所示)。

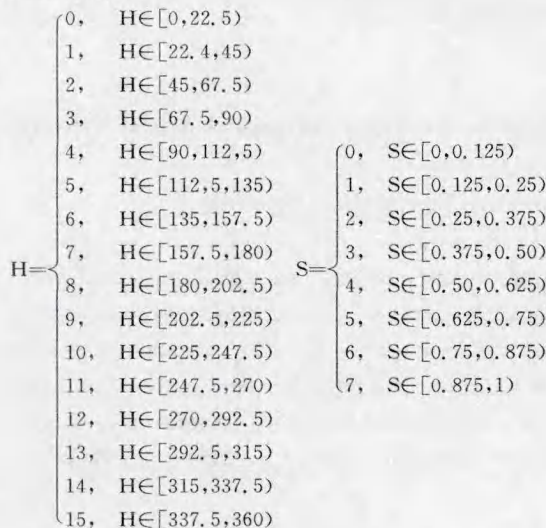


图 3 H 平均分为 16 份,S 平均分为 8 份

H-S 直方图实际上就是 H 分量和 S 分量的离散型联合概率分布。每一个联合分量的取值作为图像颜色特征向量的一维,最终可以生成一个 $16 \times 8 = 128$ 维的特征向量。图 4、图 5 分别为图 2(a)上、下图相应的 H-S 颜色直方图,图上标出了区分比较大的区域,直观上就可以发现背景区域对特征向量的取值影响很大。



图 4 图 2(a)上图对应的 H-S 颜色直方图



图 5 图 2(a)下图对应的 H-S 颜色直方图

由于每幅图像的大小不一样,因此在进行相似度计算之前必须对特征向量进行归一化操作,对于原始图像只需将 128 维数据都归一化到 $[0, 1]$ 之间。而对于前景提取过的图像,第一维纯黑色为背景信息,计算时不予考虑,所以去除第一维后再将剩下的 127 维进行归一化操作。

4.2 LBP 纹理特征

纹理特征,反映的是图像像素点某个邻域内颜色或者灰度的变化情况,Ojala 等人提出的 LBP(局部二进制模式)算子^[10],具有旋转不变性和单调变换不变性。对一幅图像提取 LBP 纹理特征,首先需要将彩色图像转换为灰度图像,灰度的计算公式如下:

$$Gray = 0.3R + 0.59G + 0.11B$$

式中, R, G, B 分别代表红绿蓝 3 个通道的值。LBP 算子定义在一个 3×3 的窗口内,将中心点邻域内的 8 个像素的灰度和它的灰度作比较,若比它大,则取值为 1,反之为 0。这样对于图片上的每个点,都会形成一个 8bit 的编码,它是一个 0 到 255 之间的无符号整数(如图 6 所示)。

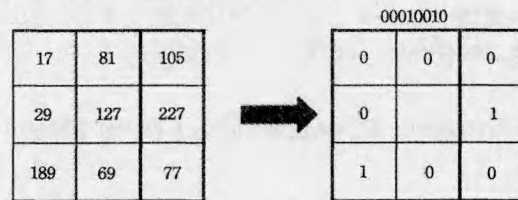


图 6 LBP 纹理算子

依次遍历所有像素点,就会得到与源图像大小相等的 LBP 矩阵,统计这个矩阵的直方图,得到的向量就是这幅图片的基于 LBP 算子的纹理特征向量。

4.3 颜色纹理混合特征

图像一种类型的底层特征可能还无法准确描述图像的内容,因此我们将 4.1 节和 4.2 节中提取的颜色直方图和纹理特征混合,混合特征的相似度计算公式为:

$$S = S_{color} W_{color} + S_{texture} W_{texture}$$

式中, S_{color} 和 $S_{texture}$ 是颜色和纹理特征向量计算出来的相似度, W_{color} 和 $W_{texture}$ 代表特征所占的权重。本文实验对于两种特征均采用曼哈顿距离度量相似度,权值各取 0.5,因此计算时可以直接将 2 个特征向量进行连接, $H = (H_{color}, H_{texture})$ 。

5 实验及评价

本文实验图像来自国内某 C2C 电子商务网站,选取了 15 种商品,每种商品选取了不同背景的 11 幅图片,部分图片中还有 LOGO、文字等干扰信息(如图 7 所示)。

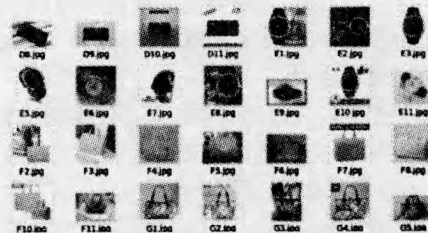


图 7 实验图片集

5.1 特征实验结果

图 8 展示了颜色纹理混合特征在直接用原始图片检索和基于前景提取的检索两种情况下相对于单个特征的提高比。图的横坐标表示前 N 张返回的图片,纵坐标表示平均准确率。每种特征在两种方法中都对所有图片进行循环交叉测试,即对每张图片都取最相似的前 5 到前 10 张图片,计算准

(下转第 126 页)

[10] 章毓晋. 基于内容的视觉信息检索[M]. 北京: 科学出版社, 2003:57-130,401-421

[11] 刘宇红,刘桥,任强. 基于模糊聚类神经网络的语音识别方法[J]. 计算机学报,2006,29(10):1894-1900

[12] 章毓晋. 图像处理和分析[M]. 北京: 科学出版社, 2003(12): 180-215

[13] 韩晓微,晏磊,原忠虎,等. 基于BP神经网络的颜色模糊量化方法[J]. 系统仿真学报. 2006,18(10):3007-3010

[14] Yu W,Li X O. Fuzzy identification using fuzzy neural networks with stable learning algorithms[J]. IEEE Transactions on fuzzy

[15] Yong Y,Chi L H,Tsan M C. Intelligent Fabric Hand predication system with Fuzzy Neural Network[J]. IEEE Transactions on Au,R. System, Man, and Cybernetics, Part C: App. and Reviews,2010,6(40):619-629

[16] 汪卫星. 基于颜色和形状特征图像检索技术研究[D]. 长沙: 长沙理工大学,2010

[17] Computer Vision Test Images[OL]. <http://www.cs.cmu.edu/~cil/v-images.html>

[18] 刘盈盈,石跃祥,莫浩瀚. 基于改进的动态聚类算法在彩色图像分割中的应用[J]. 计算机工程与应用,2008,44(29):191-192

(上接第 115 页)

准确率(查准率),最后求所有图片准确率的平均值。从以上实验结果可以看出,使用混合特征在两种情况下较单个特征准确率均有提高,其中对于纹理特征的提高比较大,对于颜色直方图特征的提高比较小。

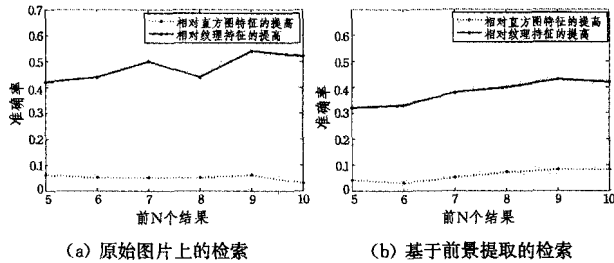


图 8 使用混合特征的提高比

5.2 算法实验结果

图 9 展示了直接用原始图片检索和基于前景提取的检索在不同的特征上准确率的对比。对于颜色直方图、LBP 纹理特征、颜色纹理混合特征,与直接使用原始图片检索的准确率相比较,本文提出的基于前景提取的复杂背景图像检索算法均有较大的提升。

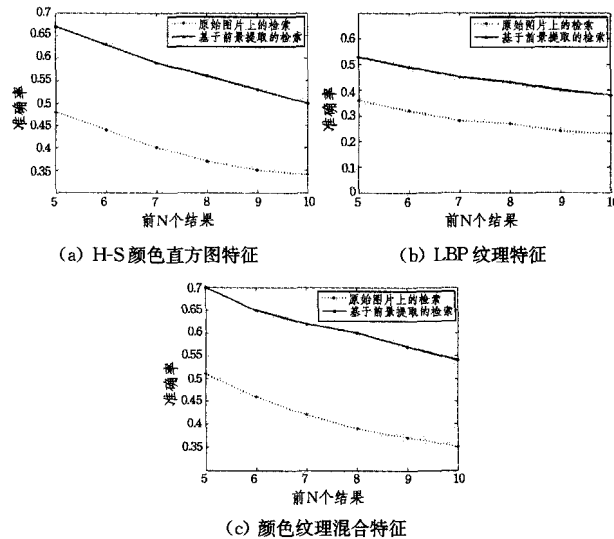


图 9 使用不同特征的检索准确率

结束语 基于内容的图像检索是一种“以图搜图”的检索方式,用户往往关注的是图像中主体部分的相关性,而复杂的背景会影响检索结果的准确性。本文提出基于前景提取的图像检索算法,其能够在无人工交互的情况下提取图像的前景部分,之后针对前景部分做特征提取和相似度计算得到检索结果。通过实验证明,对于颜色直方图、LBP 纹理特征、颜色

纹理混合特征,该算法相对于直接用原始图片的检索准确率有较大的提升。

同时我们发现,对于某些图片,采用从 4 个角开始的扩散方法自动标记背景区域对某些图片并不适合,有些图片会把位于角落的前景也标记为背景,有些图片提取出的前景仍然包含大部分的背景,下一步的工作我们会考虑从图像边界向内部扩散的方法。另外,基于前景提取的图像检索算法虽然可以实现复杂背景的图像检索,但部分图像经过前景提取后仍然存在 Logo、文字等区域,这些区域大部分属于噪声信息,对图像检索的准确率会造成一定的影响,下一步的工作会将这些区域去除后再提取图像的特征向量。

参 考 文 献

[1] Rother C, Kolmogorov V, Blake A. Grabcut: Interactive foreground extraction using iterated graph cuts[J]. ACM Transactions on Graphics (TOG), ACM,2004,23(3):309-314

[2] Boykov Y Y, Jolly M P. Interactive graph cuts for optimal boundary & region segmentation of objects in ND images[C]// Proceedings of the International Conference on Computer Vision,ICCV 2001. 2001:105-112

[3] Kolmogorov V, Zabini R. What energy functions can be minimized via graph cuts? [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2004,26(2):147-159

[4] Rao A,Srihari R K,Zhang Z. Spatial color histograms for content-based image retrieval[C]//Proceedings of the 11th IEEE International Conference on Tools with Artificial Intelligence, 1999. IEEE,1999:183-186

[5] Manjunath B S, Ma W Y. Texture features for browsing and retrieval of image data[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,1996,18(8):837-842

[6] Tamura H, Mori S, Yamawaki T. Textural features corresponding to visual perception[J]. IEEE Transactions on Systems, Man and Cybernetics,1978,8(6):460-473

[7] 周静,郝红卫. 基于用户感兴趣区域的图像检索方法[J]. 计算机应用研究,2007,24(9):282-284

[8] Ford D R, Fulkerson D R. Flows in networks [M]. Princeton: Princeton university press,2010

[9] Goldberg A V, Tarjan R E. A new approach to the maximum-flow problem[J]. Journal of the ACM (JACM),1988,35(4): 921-940

[10] Ojala T, Pietikainen M, Maenpaa T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2002,24(7):971-987