

基于目标区域的语义图像检索

党长青

DANG Chang-qing

唐山学院 信息工程二系,河北 唐山 063020

Department No.2 of Information Engineering, Tangshan College, Tangshan, Hebei 063020, China

E-mail:dcqts@126.com

DANG Chang-qing. Image retrieval based on high-level semantics of region. Computer Engineering and Applications, 2008, 44(20): 185-187.

Abstract: A new image retrieval method using high-level semantics is presented in this paper. First, the proposed approach employs a segmentation algorithm based on color and space to divide images into regions, and low-level feature for the color, shape, position, and texture of each region are subsequently extracted. The cluster image object using these features and then the semantic feature is gained; FCM is used to cluster image. In image retrieval, the querying image is compared to cluster center, then retrieved in the class with the minimal distance. The experiment results show that the proposed approach has an excellent precision and has reduced the "semantic gap" between the visual feature and semantic visual.

Key words: image retrieval; image segmentation; FCM; Clustering

摘要:提出了一种基于高层语义的图像检索方法,该方法首先将图像分割成区域,提取每个区域的颜色、形状、位置特征,然后使用这些特征对图像对象进行聚类,得到每幅图像的语义特征向量;采用模糊C均值算法对图像进行聚类,在图像检索时,查询图像和聚类中心比较,然后在距离最小的类中进行检索。实验表明,提出的方法可以明显提高检索效率,缩小低层特征和高层语义之间的“语义鸿沟”。

关键词:图像检索;图像分割;模糊C均值;聚类

DOI:10.3778/j.issn.1002-8331.2008.20.056 **文章编号:**1002-8331(2008)20-0185-03 **文献标识码:**A **中图分类号:**TP391

1 引言

随着多媒体和网络技术的迅速发展,图像数据的来源不断扩大,基于内容的图像检索技术正越来越成为人们研究的热点。其基本思想就是从图像中提取颜色、纹理、形状及区域等视觉特征,给定一幅图像,根据设计的相似度公式,计算查询图像和数据库中图像相似度的大小显示给用户^[1,5]。

目前基于目标区域的方法主要有两种:一种是人为地将图像分为几个区域的方法。该方法采用将图像人为划分为几个不同区域,对不同区域进行特征提取的方式进行检索,文献[1,2]提到了两种基于分块的彩色图像检索方法;另一种是基于图像分割的方法,如文献[4-5]提到的方法,这些基于分割的方法在一定程度上解决了基于目标区域进行图像检索问题,但是由于经典的分割方法得到的分割区域与人眼的视觉感知存在较大的差异,因此在利用这些分割方法进行基于目标区域的图像检索方面还存在一定的局限性。

本文提出了一种基于目标区域的语义图像检索方法,首先采用基于颜色与空间的图像分割算法将图像分割成不同的区域^[6],然后提取每个区域的低层索引特征,以颜色、位置、形状的

低层特征相似度对图像对象进行聚类,得到每一幅图像的语义向量,然后使用FCM聚类算法对图像进行聚类。在图像检索时,先计算查询图像和聚类后每个聚类中心的距离,然后在距离最小的聚类里去检索相似的图像。实验证明,本文的方法比传统的基于内容的图像检索方法的性能更好。

2 低层特征的提取

为了获取有意义的目标区域,必须对图像进行分割,因为图像分割在理论上仍缺乏简单易行又准确可靠的通用方法,至今没有确切的数学定义能和人的感觉完全一致,为了克服现有经典分割方法难以提取具有视觉一致性目标区域的问题,采用一种基于颜色与空间的图像分割方法^[6],该方法可以较好地实现图像中有意义的目标区域的提取。

通过图像分割得到了一些有意义的目标区域,称之为对象。假设通过图像分割,得到 n 个有意义的图像区域,用 $\{r_i\}_{i=1}^n$ 表示,分别提取归一化的颜色、位置、形状特征来表示。

2.1 颜色特征

一般认为R、G、B颜色空间与人眼的感知差异很大,本文

使用较好符合人眼感知特性的 HSV 颜色模型。首先将图像中每一像素的 r, g, b 值转换为 h, s, v 值 ($h \in [0, 360], s \in [0, 1], v \in [0, 1]$) 提取区域 r_i 的 H, S, V 三个分量的特征, 分别用 F_{i1}, F_{i2}, F_{i3} 表示如下:

$$\begin{aligned} F_{i1} &= \frac{1}{360 \cdot A_i} \sum_{p \in r_i} h(p) \\ F_{i2} &= \frac{1}{A_i} \sum_{p \in r_i} p \in r_i s(p) \\ F_{i3} &= \frac{1}{A_i} \sum_{p \in r_i} p \in r_i v(p) \end{aligned} \quad (1)$$

其中 A_i 为区域 r_i 中像素的个数, p 表示区域中的像素点。

2.2 位置特征

提取归一化的区域空间中心作为区域图像的位置特征, 分别用 F_{i4}, F_{i5} 表示。

$$F_{i4} = \frac{1}{A_i \cdot x_{\max}} \sum_{p \in r_i} p_x, F_{i5} = \frac{1}{A_i \cdot y_{\max}} \sum_{p \in r_i} p_y \quad (2)$$

其中 $p = [p_x, p_y]^T$ 表示像素点的位置坐标, $p_x = 1, \dots, x_{\max}, p_y = 1, \dots, y_{\max}, x_{\max}, y_{\max}$ 为图像区域的尺寸。

2.3 形状特征

提取归一化的形状大小和离心率作为区域图像的形状特征, 分别用 F_{i6} 和 F_{i7} 表示。

$$F_{i6} = \frac{A_i}{x_{\max} \cdot y_{\max}} \quad (3)$$

为了计算区域 r_i 的离心率, 定义协方差矩阵 C_i :

$$C_i = \frac{1}{A_i} \sum_{p \in r_i} (p - S(r_i))(p - S(r_i))^T \quad (4)$$

其中 $S(r_i) = \frac{1}{A_i} \sum_{p \in r_i} p$ 为区域的中心。定义 ρ_k 和 $\mu_k, k=1, 2$ 分别为 C_i 的特征值和特征向量, 即:

$$C_i u_k = \rho_k u_k (\rho_1 \geq \rho_2)$$

特征向量 u_1 为区域的方向, 而 u_2 与 u_1 正交, 区域的离心率

$\varepsilon_i = 1 - \frac{\rho_1}{\rho_2}$, 对它归一化后得到离心率 F_{i7} :

$$F_{i7} = e^{\varepsilon_i} \quad (5)$$

得到区域 r_i 的特征向量为 $F_i = (F_{i1}, F_{i2}, \dots, F_{i7})$ 。

3 基于高层语义的图像检索

在本文中, 首先根据相似的对象对图像进行聚类, 因为在图像中有些对象比其它的对象重要, 因此, 必须为图像中的对象分配权值。为了确定图像对象的权值, 首先根据颜色、位置、形状对图像对象进行聚类, 然后根据信息检索中基于术语频率和倒排文档频率计算对象的权值。最后每一幅图像用一个语义向量进行描述, 向量中的每一项表示在图像中出现的对象的权值。然后使用模糊 C 均值聚类算法对图像进行聚类。

3.1 对象聚类

在图像分割后, 得到分割后的图像对象, 假设图像库中有 N 幅图像, 在经过分割后, 得到 M 个对象。然后根据对象之间的相似性对 M 个对象进行聚类, 得到 t 个聚类, R_1, R_2, \dots, R_t 。在本文中, 对象的相似性是根据每个对象之间的颜色、位置和

形状相似性得到。这里采用高斯函数计算任意对象 i, j 之间各特征的相似程度。通过调整各特征的方差 σ_i 可以方便地归一化具有不同量纲的特征变量, 并且相似度总在 $[0, 1]$ 之间。下面分别介绍颜色、位置和形状相似度, 然后在结合三个相似度计算对象之间的相似度。

3.1.1 颜色相似度

区域的颜色特征用该区域的平均色代表, 为了方便颜色相似度的计算, 将色度坐标统一转化为柱坐标下的欧式空间坐标 (c_1, c_2, c_3) 表示:

$$c_1 = F_{i2} \cdot \cos(F_{i1}), c_2 = F_{i2} \cdot \sin(F_{i1}), c_3 = F_{i3} \quad (6)$$

$$S_1 = \exp\left(-\frac{(c_{i1} - c_{j1})^2 + (c_{i2} - c_{j2})^2 + (c_{i3} - c_{j3})^2}{3\sigma_1^2}\right) \quad (7)$$

3.1.2 位置相似度

$$S_2 = \exp\left(-\frac{(F_{i4} - F_{j4})^2 + (F_{i5} - F_{j5})^2}{2\sigma_2^2}\right) \quad (8)$$

3.1.3 形状相似度

$$S_3 = \exp\left(-\frac{w_s(F_{i6} - F_{j6})^2 + w_\rho(F_{i7} - F_{j7})^2}{2\sigma_3^2}\right), w_s + w_\rho = 1 \quad (9)$$

图像中任意两个对象 i, j 之间的相似度是各特征间相似度的加权平均:

$$S_{object}(i, j) = \sum_{r=1}^3 w_r S_r, \sum_{r=1}^3 w_r = 1 \quad (10)$$

其中 w_1 为颜色的权值, w_2 为位置的权值, w_3 为形状的权值。

为了对图像对象进行聚类, 使用一个阈值 T_{obj} , 在实验中 $T_{obj} = 0.7$ 。如果两个对象之间的相似性比 T_{obj} 大, 则这两个对象必须聚到同一个类中, 换句话说, 同一类中对象之间的相似度都大于 T_{obj} 。结合三个相似度计算对象之间的相似度, 然后对图像对象聚类, 最后得到 t 个聚类, R_1, R_2, \dots, R_t 。

3.2 图像向量模型

图像聚类是基于图像的相似度, 为了计算图像的相似度, 构建了一幅图像的一个特征向量 $W_l(w_{1,l}, w_{2,l}, \dots, w_{i,l}, \dots, w_{t,l})$ 。 $w_{i,l}$ 是图像 l 中聚类 R_i 的权重, 向量的维数跟聚类数目相等。在向量 W_l 中, 可能某个聚类的权重为 0, 这是因为在聚类过程中, 这幅图像没有对象分到某个类。采用信息检索领域的思想计算这个向量, 假设图像总的数目为 N, n_i 为聚类 R_i 中出现的图像数目。定义归一化的术语频率 $f_{i,j}$:

$$f_{i,j} = \frac{freq_{i,l}}{\max_h freq_{h,l}} \quad (11)$$

其中 $freq_{i,l}$ 是聚类 R_i 在图像 l 中出现的次数, $\max_h freq_{h,l}$ 是聚类 $\{R_i\}, i=1, \dots, t$ 在图像 l 中出现的最大的次数。假设当一个聚类出现在大多数的图像中, 则它在区分相关图像和不相关图像时是不重要的。所以需要为 R_i 定义一个倒排文档频率 idf_i :

$$idf_i = \log \frac{N}{n_i} \quad (12)$$

为了平衡上面的两个因素, 得到图像对象的权值 w_{il} :

$$w_{il} = tf_{il} \times idf_i = tf_{il} \times \log \frac{N}{n_i} \quad (13)$$

3.3 基于模糊 C 均值图像聚类

FCM 聚类是一种基于目标函数的聚类方法, 它把聚类归

结成一个带约束的非线性规划问题,通过优化求解获得数据集的模糊划分和聚类。其基本思想是通过反复修改聚类中心和分类矩阵 U 来实现动态的迭代聚类,使得被划分到同一簇的对象之间相似度最大,而不同簇之间的相似度最小。

给定观察空间中的一个有限样本集 $x_k \in R^N (k=1,2,\dots,K)$, 则 FCM 的价值函数(或目标函数)的形式如下:

$$J_m = \sum_{i=1}^C \sum_{k=1}^K \mu_{ik}^m (x_k - c_i)(x_k - c_i)^T \quad (14)$$

其中, C 是类别数, x_k 为第 k 个样本, c_i 为第 i 类的聚类中心, $m \in [1, \infty)$ 是影响聚类模糊度的一个加权指数, $\mu_{ik} \in [0, 1]$ 是 x_k 属于第 i 类的隶属度,且有:

$$0 \leq \sum_{i=1}^C \mu_{ik} \leq 1, \forall k=1,2,\dots,K \quad (15)$$

隶属度迭代公式:

$$\mu_{ik} = \frac{1}{\sum_{j=1}^C \left(\frac{d_{ik}}{d_{jk}} \right)^{\frac{2}{m-1}}} \quad (16)$$

其中 d_{ik} 是第 k 个样本与第 i 个聚类中心的 Euclid 距离,用以度量数据点和聚类中心的相似性。

聚类中心的迭代公式:

$$c_i = \frac{\sum_{k=1}^K \mu_{ik}^m x_k}{\sum_{j=1}^K \mu_{ij}^m} \quad (17)$$

FCM 算法步骤如下:

(1) 给定聚类类别数 $C (2 \leq C \leq K)$ 和 C 个聚类中心, K 是样本个数, 设定迭代停止阈值 ε , 设置算法迭代计数器 $t=1$ 。

(2) 用值在 $[0, 1]$ 间的随机数初始化隶属矩阵 U , 使其满足式(15)中的约束条件。

(3) 用式(16)计算每一个样本的隶属度 $\mu_{ik} (i=1,2,\dots,C; k=1,2,\dots,K)$ 。

(4) 用式(17)计算 C 个聚类中心 $c_i (i=1,2,\dots,C)$ 。

(5) 根据(14)式计算目标函数, 如果 $|J(t) - J(t-1)| > \varepsilon, t \leftarrow t+1$, 转步骤(3), 否则算法停止, 输出隶属度矩阵和聚类中心。

3.4 图像检索

在图像检索时,首先对查询图像进行分割得到有意义的图像对象,然后使用 3.2 小节的方法计算得到图像的语义向量 $W_q (w_{1,q}, w_{2,q}, \dots, w_{i,q}, \dots, w_{l,q})$, 接着计算查询图像和分类后得到的聚类中心的距离 $\|W_q - c_i\|, i=1,2,\dots,C$, 找出最小距离的聚类 $\arg \min_c \|W_q - c_i\|$ 。在最小距离的聚类中进行检索, 检索时使用余弦相似度计算查询图像和这类图像之间的相似性, 查询图像 q 和检索图像 I 之间的相似度计算公式如下:

$$\text{sim}_{\text{img}}(q, I) = \frac{\sum_{h=1}^l w_{h,q} \times w_{h,I}}{\sqrt{\sum_{h=1}^l w_{h,q}^2} \times \sqrt{\sum_{h=1}^l w_{h,I}^2}} \quad (18)$$

根据相似度的大小对检索后的图像进行排序。

5 实验结果分析

基于上述思想,在 Windows XP 操作系统上实现一个图像

检索系统,利用该系统在 P4-2.8 G, RAM 512 M 的 PC 机上对一个具有 4 000 幅的彩色图像数据库进行处理,应用本文提出的分割算法得到 29 684 个区域。对检索算法的性能评价的重要指标是查准率(Precision)、查全率(Recall)。查准率定义为检索出的图像中相关图像的数目占的比例,查全率定义为检索出的相关图像的数目占数据库中所有相关的图像数目的比例。实验中对“花”进行检索,第一张为查询图片,检索部分结果如图 1 所示。



图 1 基于高层语义的图像检索结果

为了更好地评估系统性能与其它方法作了比较。文献[1]基于颜色的方法,文献[2]基于颜色和空间的方法,文献[4]也是一种基于目标区域的图像检索方法。比较结果如表 1 所示,从表中可以看出,本文提出的方法的查准率、查全率比这些方法都要高。通过实验对比可知,本文的算法具有较好的性能。

表 1 四种不同方法的性能比较

	文献[1]	文献[2]	文献[4]	Proposed
平均查准率	0.522	0.534	0.671	0.894
平均查全率	0.395	0.421	0.582	0.836

6 总结

为了提高图像检索的效率,缩小低层特征和高层语义间的语义鸿沟,本文提出了基于目标区域的语义的图像检索方法。实验结果表明所提出方法的查全率与查准率比传统的基于内容的图像检索高,系统具有较好的鲁棒性和优良的检索性能。

参考文献:

- [1] 张恒博,欧宗瑛.一种基于色彩和灰度直方图的图像检索方法[J].计算机工程,2004,30(10):20-22.
- [2] 刘芳,王涛,周登文,等.基于颜色-空间二维直方图的图像检索[J].计算机工程与应用,2002,38(12):85-88.
- [3] 党长青,宋凤娟,刘树明.基于多特征融合和相关反馈的图像检索[J].计算机工程与应用,2008,44(3):186-188.
- [4] Nascimento M A, Sridhar V. Effective and efficient region-based image retrieval[J]. Journal of Visual Languages and Computing, 2003 (23):772-775.
- [5] Byoung Chul K O, Lee H S, Byun H. Region-based image retrieval system using efficient feature description, pattern recognition [C]// Proceedings of the 15th International Conference on Barcelona, Spain, 2000, 4: 283-286.
- [6] Vasileios Mezaris, Ioannis Kompatsiaris. An ontology approach to object-based image retrieval [C]// Proc IEEE ICIP, Rochester, NY, USA, 2003.