

基于内容的多特征融合图像检索

罗 军, 况 奔

LUO Jun, KUANG Hang

重庆教育学院, 重庆 400067

Chongqing Education College, Chongqing 400067, China

E-mail: kanghangpaper@163.com

LUO Jun, KUANG Hang. Content-based image retrieval using combination features. *Computer Engineering and Applications*, 2009, 45(1): 153-155.

Abstract: A new method combining color shape and spatial feature for efficient image retrieval is proposed in this paper. Color feature is extracted by the histogram, shape feature is extracted by zernike moments and spatial analysis is performed by computing the bar centre and the standard deviation of the distribution of colors. In order to improve the retrieval performance, a relevance feedback based on support vector machines is proposed. It is demonstrated that the image retrieval based on multi-feature is superior than which based on single feature.

Key words: image retrieval; similarity measurement; color histogram; Zernike moments; spatial information

摘 要: 提出一种组合颜色、形状和空间信息的图像检索方法, 用颜色块的颜色直方图表示图像的颜色特征; 使用 Zernike 矩表示图像的形狀特征, 通过提取颜色块的质心、分布方差聚散度等特征得到图像的空间特征。为了进一步提高图像检索的精确度, 提出一种基于支持向量机的相关反馈方法。实验结果表明, 基于组合特征的图像检索方法优于基于单一特征的检索方法。

关键词: 图像检索; 相似度计算; 颜色直方图; Zernike 矩; 空间信息

DOI: 10.3778/j.issn.1002-8331.2009.01.048 文章编号: 1002-8331(2009)01-0153-03 文献标识码: A 中图分类号: TP391

1 引言

随着计算机技术和网络技术的迅猛发展, 以及多媒体的推广应用, 每天都有成千上万的图片产生, 如何从如此海量的图像库中快速而准确地找到满足用户需求的图像就成为迫切需要研究的热点问题。因此, 人们提出了基于内容的图像检索技术^[1-3]。这种技术包括提取图像的可视特征: 颜色、纹理、形状、位置和相互关系等, 以及将数据库中图像和查询样本图像在特征空间进行相似匹配, 检索出与样本相似的图像。

颜色是描述图像内容的最直接的视觉特征。已经有许多文献提出了各种基于颜色的图像索引技术: Swain^[4]首先提出了使用颜色直方图作为图像内容的索引特征, 即首先将颜色空间划分成若干固定的子空间, 然后对每幅图像统计属于各个子空间的像素的数目, 图像之间的相似性测度则采用直方图求交的方法。直方图除了具有计算简单的特性外, 同时还具有对平移和旋转不敏感的优点。但是, 全局直方图无法捕捉颜色组成之间的空间关系; 同样, 全局矩方法也丢失了图像的空间信息。

由于形状表示了图像中有意义的区域或相关对象, 因此基于形状的图像索引技术允许用户查询具有相似对象的数据库图像^[3-4]。Zernike 具有比较稳定的性质, 他对光照、旋转不敏感。空间信息也是一种有用的检索特征, 目前, 许多研究人员着手研究颜色的空间索引技术。基于图像空间的固定划分方法, 即人为地将图像划分成适当的分块, 然后为每个分块提取相应的

局部颜色特征。例如, 分块的颜色直方图和分块的颜色矩^[2,5]。QBIC 从预定义的颜色集中为每个分块选择主色, Hsu 和 Chua^[6]则按照某种方法为整幅图像提取 C 种主色, 然后确定各分块属于主色中的哪种。S.Lim, G Lu^[7]采用类似于地理学的统计方法来描述颜色的空间信息分布。这种统计方法总的来说, 仅仅是针对某一种颜色来进行空间信息统计, 却忽略了各种颜色之间的互相联系。

由于各种颜色之间的相关联系也是图像内容的一个重要信息, 通过计算不同颜色间的相对距离来弥补的仅仅是传统直方图的缺陷, 再结合图像形状和空间信息, 提出一种相似度计算方法将不同的特征信息进行融合, 为了进一步提高图像检索的精确度, 提出支持向量机的相关反馈方法, 在相关反馈过程中, 采用主动学习算法对支持向量机进行训练。实验结果表明, 提出的方法具有良好的检索性能。

2 特征提取

2.1 颜色特征

颜色特征在图像检索中占有重要位置, 通常采用 RGB 三色表示。但为使表示法更符合人眼视觉特征, 把 RGB 空间模型转换为 HSV 空间模型进行颜色特征的统计。

设一幅 $m \times n$ 的图片 I , 这幅图片的颜色信息被量化为 71 色^[2]。定义 $M_i = \{(x, y) \in I, i = C_i\}$ 为 C_i 中像素集合, 其中 C_i 为第 i

作者简介: 罗军(1974-), 男, 讲师, 主要研究方向: 现代教育技术, 人工智能; 况奔(1974-), 男, 讲师, 主要研究方向: 图像检索, 模式识别, 数据库。

收稿日期: 2007-12-25 **修回日期:** 2008-03-24

种颜色。传统的颜色直方图是每一个 bin 中的像素占总像素个数的比率。表示如下:

$$H'_g(i) = \frac{\|M_i\|}{m \times n} \quad (1)$$

其中, $\|M_i\|$ 表示集合 M_i 中的元素个数。

在图像中各种颜色之间有着密切的联系。选择量化后颜色之间的像素统计的相对距离 $d_{ij} = d(H_g(i), H_g(j))$ 为因素来考虑各个颜色之间的相对关系。

$$d_{\max} = \max(d_{ij}), \alpha_{ij} = 1 - \frac{d_{ij}}{d_{\max}} \quad (2)$$

从上式可以看出, $(\alpha_{ij})_{m \times n}$ 即为各颜色相关的对称矩阵。把 $(\alpha_{ij})_{m \times n}$ 按行累加求均值如下:

$$\omega(i) = \frac{1}{71} \sum_j \alpha_{ij} \quad (3)$$

把式(3)的 $\omega(i)$ 作为传统直方图的权重:

$$H'_c(i) = H'_g(i) \times \omega(i) \quad (4)$$

式(4)即为图像的颜色特征。

2.2 基于 Zernike 矩形状特征

Zernike 矩不变量只具有旋转不变性, 为了使其同时具有平移和缩放不变性, 首先使用 Kim H^[8] 提出的图像规格化方法对形状进行预处理, 并使之映射到一个紧凑的单位圆内, 圆心即为形状的质心。函数 $f(x, y)$ 的 Zernike 矩 z_{nm} 的定义如下:

$$z_{nm} = \frac{n+1}{\pi} \int \int V_{nm}^*(x, y) f(x, y) dx dy \quad (5)$$

$$V_{nm}(x, y) = R_{nm}(x, y) \exp(jn \tan^{-1} \frac{y}{x}) \quad (6)$$

其中, $V_{nm}(x, y)$ 表示 n 阶 Zernike 矩的基函数, $x, y \in [0, 1], n$ 为非负整数, m 为满足 $n-|m|$ 为偶数且 $|m| \leq n$ 的整数。

实值多项式表示如下:

$$R_{nm}(x, y) = \sum_{s=0}^{(n-|m|)/2} \frac{(-1)^s (n-s)!}{s! [\frac{n+|m|}{2}-s]! [\frac{n-|m|}{2}-s]!} x(x^2+y^2)^{\frac{n-s}{2}} \quad (7)$$

在数字图像域, Zernike 矩的表达形式如下:

$$z_{nm} = \frac{n+1}{\pi} \sum_x \sum_y V_{nm}^*(x, y) f(x, y), x^2+y^2 \leq 1 \quad (8)$$

$|z_{nm}|$ 被称为 Zernike 矩不变量。他有以下优点: 旋转不变性, Zernike 矩的模对于旋转不变, 冗余度小, 由于变换基是正交的, 因此信息表达的冗余性小, 所提取特征的相关性和冗余性小^[8]; 容易构造高阶不变矩, 可提供更多特征用于识别; 健壮性、鲁棒性好, 抗噪能力强^[8]。

2.3 空间特征提取

通过颜色聚类, 把颜色量化成 71 种颜色, 对于各种颜色的质心、标准偏差和相对于质心的聚散程度即是我们感兴趣的空間特征, 每一种颜色计算其质心如下:

$$\bar{x}(i) = \frac{1}{\|M_i\|} \sum_{x,y \in M_i} x, \bar{y}(i) = \frac{1}{\|M_i\|} \sum_{x,y \in M_i} y \quad (9)$$

质心: $\mu^l(i) = (\bar{x}(i), \bar{y}(i))$ 。

在对每种颜色计算质心后, 用每种颜色中的像素对其质心的相对距离的倒数来表示颜色对质心的聚散程度。

$$\text{聚散度: } s^l(i) = \sum_{(x,y) \in M_i} \frac{1}{d(p, \bar{p}_i)} \quad (10)$$

其中 p 为量化后颜色 i 中的像素点坐标, \bar{p}_i 为这种颜色的质

心。这样图像的颜色质心和空间信息都被考虑在内。

$$\text{分布方差: } \sigma^l(i) = \frac{\sqrt{\sum_{(x,y) \in M_i} d(p, \bar{p}_i)^2}}{M_i - 1} \quad (11)$$

$$\text{空间特征: } H'_s(i) = (\mu^l(i), s^l(i), \sigma^l(i)) \quad (12)$$

3 相似度计算

在基于内容的图像检索中, 距离度量函数也是影响图像检索效率的一个十分重要的因素。理想的距离度量函数应该与人类视觉感知一致, 即距离度量函数 $D(I, J)$ 的值越小, 表明两幅图像在人看来越相似。

采用直方图求交的方法计算两幅图像之间的颜色相似度:

$$d_{color}(I, J) = 1 - \sum_{i=1}^{71} \frac{\min(H'_c(i), H'_c(i))}{\max(H'_c(i), H'_c(i))} \quad (13)$$

形状的相似度采用欧式距离进行计算, 首先将形状特征向量归一化, 设图像 I 和图像 J 的特征向量分别为 f_i 和 f_j , 那么图像 I 和图像 J 距离计算公式:

$$d_{shape} = 1 - \text{norm}(d(f_i, f_j)) \quad (14)$$

其中, $d(f_i, f_j)$ 是两幅图像形状向量之间的欧式距离, $\text{norm}(d(f_i, f_j))$ 为 $d(f_i, f_j)$ 归一化的结果。

空间特征距离计算公式如下:

$$d_{spatial}(I, J) = \sum_{i=1}^{71} (d(s^l(i), s^l(i)) \times \frac{\sqrt{2} - d(\mu^l(i), \mu^l(i)) + \frac{\min(\sigma^l(i), \sigma^l(i))}{\max(\sigma^l(i), \sigma^l(i))}}{\sqrt{2}}) \quad (15)$$

不同特征之间的权向量为 $\omega = (\omega_1, \omega_2, \omega_3)^T$ 满足 $\sum_{i=1}^3 \omega(i) = 1$

且有:

$$D = (d_{color}(I, J), d_{shape}(I, J), d_{spatial}(I, J)) \quad (16)$$

则两幅图的相似度为:

$$\text{sim}(I, J) = \frac{1}{D \cdot \omega} \quad (17)$$

4 相关反馈

相关反馈是近年来在图像检索中较为重要的研究方法, 在一定程度上影响图像检索的查全率, 特别是在基于内容的图像检索中为解决低层特征和高层语义之间的差异而造成的“语义鸿沟”问题提供了方法。

4.1 支持向量机

简单的说, SVM 就是通过在原空间或经投影后的高维空间中构造最优分类面, 将给定的属于两个类别的训练样本分开, 构造超平面的依据是两类样本离超平面的距离最大化。首先介绍线性可分情况下 SVM 的原理。

设线性可分样本集 $(x_i, y_i), 1 \leq i \leq N, x_i \in R^d, y_i \in \{-1, 1\}$ 是类别标号, d 维空间中线性判别函数的一般形式为 $g(x) = w \cdot x + b$, 相应的分类面方程为 $w \cdot x + b = 0$ 。

将 $g(x)$ 进行归一化, 使所有的 x_i 都满足 $|g(x)| \geq 1$, 即离分类面最近的样本 $|g(x)| = 1$, 这样分类间隔就等于 $2 / \|w\|$ 。求解最优分类面就等效于最小化 $\|w\|$ 目标函数为:

$$\min \Phi(w) = \frac{1}{2} \|w\|^2 = \frac{1}{2} (w \cdot w) \quad (18)$$

式(18)满足下面的约束条件:

$$y_i(w \cdot x_i + b) \geq 1, i=1, \dots, N \quad (19)$$

采用 Lagrange 乘子法引入乘子 $\alpha=(\alpha_1, \dots, \alpha_N)$, 求解该二

次规划问题, 可以得到最优分类面, 其中 $w = \sum_i \alpha_i y_i x_i$, x_i 是位于分类间隔面上的样本, 这些训练样本被称为支持向量, 分类函数为:

$$f(x) = \text{sign} \left[\sum_i \alpha_i y_i x_i \cdot x + b \right] \quad (20)$$

对于线性不可分情况, SVM 一方面引入松弛变量和惩罚因子, 使目标函数变为:

$$\Phi(w, \xi) = \frac{1}{2} (w \cdot w) + C \left[\sum_1^N \xi_i \right] \quad (21)$$

另一方面, SVM 通过非线性变换将输入空间变换到高维空间, 然后在新空间中求解最优分类面, 线性可分情况下的点积运算变为 $k(x, y) = (\phi(x) \cdot \phi(y))$, 这样最终得到的分类函数为

$$f(x) = \text{sign} \left[\sum_i \alpha_i y_i \cdot k(x_i, x) + b \right] \quad (22)$$

4.2 基于 SVM 的相关反馈

由于图像检索的相关反馈过程可以看作模式识别中的二分类问题, 因此可以将 SVM 算法用于相关反馈的学习和检索过程。具体做法是将检索结果图像作为训练样本, 由用户标记出正例样本(其他为负例样本), 作为有类别标号的训练样本由 SVM 进行学习, 构造出适合表示用户查询意图的分类器, 然后用该模型对图像库中的所有图像进行分类。详细的学习和检索算法如下:

(1) 使用结合颜色、形状和空间信息的检索方法对查询图像进行检索。

(2) 用户对返回的 N_1 幅图像进行标记, 得到相关图集 I^+ 和不相关图集 I^- 。

(3) 构造训练样本集 (x_i, y_i) , x_i 为返回图像的特征向量。

$$x_i \in I^+ \cup I^-, y_i = \begin{cases} +1 & \text{if } x_i \in I^+ \\ -1 & \text{if } x_i \in I^- \end{cases} \quad (23)$$

(4) 用 SVM 算法对训练样本进行学习并构造分类器, 在训练时候采用主动学习算法, 选择距离分类超平面最近而与当前训练样本集中具有最大差异性的样本训练分类器, 保证了分类器的推广能力同时保持训练集中样本的多样性, 通过计算未标记的样本同训练集中样本的余弦距离来度量其差异性。提取颜色、形状、空间信息特征组成一个特征向量 $V=(v_1, v_2, \dots, v_n)$ 表示图像。假设初始分类面的参数对为 (w_0, b_0) , $U_i=(u_{i1}, u_{i2}, \dots, u_{in}) \in U$ 表示第 i 个未标记的样本, 设当前训练集中存在 N 个已标记样本, 第 j 个样本的特征向量为 $V_j=(v_{j1}, v_{j2}, \dots, v_{jn}) \in V, j=1, \dots, N$, 利用公式(24)计算第 i 个样本到分类面的距离。

$$d_i = \frac{|w_0^T U_i + b_0|}{\|w_0\|} \quad (24)$$

利用余弦距离计算向量 U_i 和 U_j 之间的计算公式如下:

$$d(U_i, U_j) = \frac{\sum_{k=1}^n u_{ik} v_{jk}}{\sqrt{\sum_{k=1}^n u_{ik}^2 \sum_{k=1}^n v_{jk}^2}} \quad (25)$$

未标记样本的差异性定义为:

$$\text{difference} = \arg \min_{U_i \in U} (\lambda \times d_i + (1-\lambda) \times (\max)_{V_j \in V} d(U_i, V_j)) \quad (26)$$

其中 λ 为平衡因子, 对以上两种距离进行平衡调整。根据公式(26), 算法从集合 U 中选择具有最小度量值的未标记样本加入训练集, 根据以上步骤重复选择样本加入训练集, 以此来训练分类器, 可提高分类器的性能。最后得到分类函数如下:

$$f(x) = \sum_i \alpha_i y_i K(x, x_i) + b \quad (27)$$

(5) 对所有图像的距离值按 $f(x)$ 排序, 返回检索图像。

5 实验结果分析

用本文方法进行了基于示例图像的图像检索实验, 其中图像库来自 <http://wang1.lst.psu.edu/> 中的 1 000 幅, 内容分为 10 类: Africa、Beach、Buildings、Buses、Dinosaurs、Elephants、Flowers、Horses、Mountains、Glaciers、Food, 每类图像包含 100 幅。本文采用查全率(recall)、查准率(precision)和检索速度来评价图像检索的性能。下面对单一特征和综合特征进行对比实验, 实验结果如下: 图 1 是基于 Zernike 矩经过 1 次相关反馈后返回的 14 幅图像, 检索时间为 3.223 s, 图 2 是基于颜色与空间信息的检索结果, 检索时间为 3.792 s, 图 3 是本文方法经过一次相关反馈后的检索结果, 检索时间为 4.251 s。从对比实验可知综合特征比单一特征检索性能更好。图 1~3 中第一幅图像为检索示例图像。

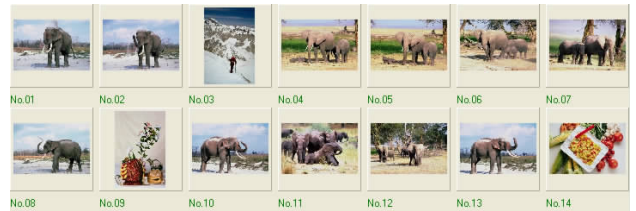


图 1 基于 Zernike 矩的检索结果



图 2 基于颜色与空间信息的检索结果



图 3 本文方法的检索结果

为进一步说明方法的检索性能, 将文献[2]、文献[3]、文献[4]的方法与文中方法做对比实验, 使用 10 类图像的平均查全率和查准率评价四种检索方法的检索性能, 实验结果如表 1 所示。从表 1 可以看出提出的检索方法的平均查全率和查准率比其他三种方法要高, 检索性能更好。

(下转 197 页)