

一种基于可拓理论的图像检索方法^①

An Image Retrieval Method Based on Cotorgy

刘 海 朱小平 (广东科学技术职业学院 软件学院 广东 珠海 519090)

夏明波 (国防科技大学 计算机学院 湖南 长沙 410073)

摘 要: 图像检索技术发展已经有一定的历史, 基于文本的检索技术逐渐成熟, 但是基于内容的图像检索技术还在探索阶段, 本文在简单的介绍图像检索技术的基础上, 提出一种可拓方法应用于图像检索的的图像检索方法, 并简单的分析了基于颜色, 纹理和图像轮廓的检索算法, 最后通过实验比较了各算法的效果。

关键词: 图像检索 可拓方法 物元 轮廓 直方图 纹理

1 前言

图像检索技术由基于文本的检索技术发展发展到基于内容的检索技术, 但是是一种好的, 能在性能和功能方面很好的满足用户需求的检索系统还未出现。目前得到广泛应用的还是基于文本的检索系统, 而基于文本的检索技术并不能很好的全面的理解一幅图像的含义, 还大量的涉入了人工操作。基于内容的技术技术是图像检索技术的发展方向, 如果能突破目前人们对图像理解的定势思维, 该技术可能会得到长足的发展。本文从可拓学理论出发, 利用可拓学处理矛盾事物这一特点, 试图寻找一种图像检索的新的理论方法。文章组织如下: 第 2 节讲述了可拓学的一些基本知识, 并就如何利用可拓学知识运用于图像检索领域提出了问题解决的初步方法; 第 3 节结合图像的基于颜色, 纹理和形状的检索, 设计了一个可拓检索的系统; 第 4 节比较了几种检索方法的效果。

2 可拓方法在图像检索中的应用

2.1 可拓工程方法

可拓学研究的对象是现实世界中的矛盾问题, 研究方向就是探讨处理矛盾问题的规律和方法。可拓方法的理论支柱是物元理论和可拓集合理论, 物元理论是以形式化的语言来描述事物的可变性及其变换, 可拓集合理论可以量化的描述事物的可变性。我们把

可拓方法应用于决策, 识别, 判断等领域的方法, 称为可拓工程方法。其特色为:

(1) 物元模型化 物元模型是将现象或问题给以物元表述的物元关系式。在可拓方法中, 是采用一定的形式化工具, 利用物元模型来制定开拓方案和评价方案, 创造可操作、可推理的方法, 以至于利用计算机去解决他们。

(2) 定性定量相结合 在利用可拓方法解决问题过程中, 既可以利用物元的可拓性确定定性方法, 又可以利用可拓集合论, 通过关联函数进行定量计算。所以是一种定性定量综合的较好的形式。

(3) 非封闭性 可拓方法本身不是封闭的, 它可以也必然和其他方法结合使用, 取长补短。

图像的检索事实上是这样一个问题, 给定一个图像或者图像的描述, 在数据库中检索出与给定图像或图像描述非常相似的图像群体, 而要判别是否相似是非常困难的, 可拓方法中物元的概念为这种问题提供了新的途径。在以下的讨论中我们检索图像是基于这样的假设:

(1) 用户给定要检索的图像;

(2) 图像所需的各种特征已经利用某种方法获得。

2.2 利用可拓方法检索图像

假设: 给定一图像 p , 特定的图像数据库 W 。求解: 数据库中和给定图像 p 类似的图像集合 P_0 。

^① 基金项目: 广东省自然科学基金项目(05006349)

收稿时间: 2008-08-24

Step1: 物元模型的一般形式: 物元分析法将人、事和物统称为事务, 把事物名称 M 、特征 c 和量值 v 作为描述事物的基本要素, 即三元, 因而可用三元组 $R=(M,c,v)$ 来对事物进行描述。

确定图像 p 这一对象形成的物元

$$R(p)=(p,C,v)=\left(\begin{array}{c} p \quad c_1, v_1 \\ \quad c_2, v_2 \\ \quad \dots\dots \\ \quad c_n, v_n \end{array}\right)$$

p : 要检索的图像对象; c_i : 特征, 比如颜色特征, 纹理特征, 空间关系特征以及相关的文本特征等等; v_i : 相应于某一特征值

Step2: 确定经典域和节域

经典域: 某一事件的特征 c_j 发生时所规定的范围。

可以用下式来表示:

$$R_0=(P_0,C,V_0)=\left(\begin{array}{c} P_0 \quad c_1, V_{01} \\ \quad c_2, V_{02} \\ \quad \dots\dots \\ \quad c_n, V_{0n} \end{array}\right)$$

其中: $V_{0i}=[v_i-a_i, v_i+a_i]$, a_i 为一事先给定的常数, 不同的特征显然该常数可能不同; v_i 为图像 p 特征 c_i 的特征值。

节域是某一事件的特征 c_j 发生时所能容许的范围, 这里表示为:

$$R_1=(P_1,C,V_1)=\left(\begin{array}{c} P_0 \quad c_1, V_{11} \\ \quad c_2, V_{12} \\ \quad \dots\dots \\ \quad c_n, V_{1n} \end{array}\right)$$

其中: $V_{1i}=[d_i, e_i]$ 为特征 c_i 可取的一切可能值。

Step3: 确定各种特征的权系数

有些特征是必须满足的, 比如检索一篮球图片, 可能圆形或椭圆形这一轮廓特征就是必须满足的条件。对于必须满足的特征作一记号为 q 。其他特征用户根据重要程度设定权值。除必须满足的特征外其他特征权值分别为: $\lambda_1, \lambda_2, \dots, \lambda_n$, 且:

$$\sum_{i=1}^n \lambda_i = 1$$

Step4: 判别数据库图像 p_0 和 p 的相关程度。

确定数据库图像 p_0 所形成的物元

$$R(p_0)=(p_0,C,v_0)=\left(\begin{array}{c} p_0 \quad c_1, v_{01} \\ \quad c_2, v_{02} \\ \quad \dots\dots \\ \quad c_n, v_{0n} \end{array}\right)$$

判断必须满足的特征: 如果 p_0 满足必须满足的

特征, 则继续下去, 否则, p_0 和 p 不相关, 也就是说 p_0 不是用户所要查询的图像, 结束。

C, 根据公式计算相关程度:

定义: 距

$$\rho(v_{0i}, V_{0i}) = |v_{0i} - v_i| - a_i (i=1, 2, \dots, n)$$

$$d_i + e_i \quad e_i - d_i$$

$$\rho(v_{0i}, V_{1i}) = \left| v_{0i} - \frac{d_i + e_i}{2} \right| - \frac{e_i - d_i}{2}$$

计算特征关联函数值公式:

$$K_i(v_{0i}) = \begin{cases} -\rho(v_{0i}, V_{0i}) / |V_{0i}|, & v_{0i} \in V_{0i} \\ \rho(v_{0i}, V_{1i}) / (\rho(v_{0i}, V_{1i}) - \rho(v_{0i}, V_{0i})), & v_{0i} \notin V_{0i} \end{cases}$$

计算关联程度:

$$K(p_0) = \sum_{i=1}^n \lambda_i K_i(v_{0i})$$

这里 $K(p_0)$ 表示了数据库中一特定图像 p_0 和目标图像 p 之间的相关程度。

如果 $K(p_0) > 0$ 则说明 p_0 和 p 是相关的, 并且 $K(p_0)$ 越大表明 p_0 和 p 越相似。 $K(p_0) < 0$ 则表示两者不相关。

3 系统设计

实验中, 用到的图像特征为图像的颜色直方图 (Color-Histogram) 表示颜色特征, 图像的边缘方向直方图 (Marginal-Directional-Histogram) 算法来确定图像的轮廓特征, 图像的纹理特征采用了共生矩阵的能量和熵等方法。

在求边缘的轮廓特征时对图像灰度灰度预处理, 采用的算法是

灰度减光处理:

$F(x,y) = (30r+59g+11b)/100$; (r, g, b 表示图像的红, 绿, 蓝值)

$$G(x,y) = a + (\ln f(x,y) + 1) / b;$$

其中, $G(x,y)$ 表示输出图像, $f(x,y)$ 表示输入图像, a, b, c 为一个常数, 具体值更具情况适当调整, 试验中数据取 $a=50, b=1/36$ 。这种算法可以适当的减低光线对图像的影响。

轮廓提取采用了 sobel 算子^[7]:

$$dx = f(x-1, y+1) + 2f(x, y+1) + f(x+1, y+1) - f(x-1, y-1) - 2f(x, y-1) - f(x+1, y-1)$$

$$dy = -f(x-1, y-1) + f(x+1, y-1) - 2f(x-1, y) + 2f(x+1, y) - f(x-1, y+1) + f(x+1, y+1)$$

计算边缘方向直方图:

$$\text{计算边缘方向: } f(x,y) = \arctan(dx/dy)$$

$$\text{弧度转化角度: } f(x,y) = f(x,y) * 180 / \Pi + 90;$$

用直方图统计边缘方向值 f , 并归一化, 得到一维的边缘直方图 $H[0,180]$. H 是图像边缘特征匹配的依据。

特征匹配采用边缘直方图的绝对值距离。

求纹理的算法:

共生矩阵(accrete-matrix-texture) 是用来描述纹理中灰度基元之间空间联系的基础, 基于共生矩阵的模型充分利用了纹理中灰度分布的性质。基于共生矩阵, 可定义许多基于统计法的纹理特性。

若将图像的灰度级定为 N 级, 那么共生矩阵为 $N \times N$ 矩阵, 可表示为 $M_{(\Delta x, \Delta y)}(h, k)$, 其中位于 (h, k) 的元素 $M_{(h, k)}$ 的值表示一个灰度为 h 而另一个灰度为 k 的两个相距为 $(\Delta x, \Delta y)$ 的像素对出现的次数。设 S 为目标区域 R 中具有特定空间联系的像素对的集合, 则共生矩阵 P 可定义为:

$$P(g_1, g_2) = \frac{\#\{(x_1, y_1), (x_2, y_2) \in S | f(x_1, y_1) = g_1 \& f(x_2, y_2) = g_2\}}{\#S}$$

利用上述灰度共生矩阵可以定义诸如反差、能量、熵和相关等纹理描述符:

反差(或称为主对角线的惯性矩):

$$CON = \sum_h \sum_k (h - k)^2 m_{hk}$$

对于粗纹理, 由于的数值较集中于主对角线附近, 此时 $(h-k)$ 的值较小, 所以相应的 CON 值也较小。相反, 对于细纹理则相应的 CON 值较大。

能量(或称为角二阶矩):

$$ASM = \sum_h \sum_k (m_{hk})^2$$

这是一种对图像灰度分布均匀性的度量。当 m_{hk} 的数值分布较集中于主对角线附近时, 其相应的 ASM 值较大; 反之, ASM 值则较小。

熵:

$$ENT = -\sum_h \sum_k m_{hk} \log m_{hk}$$

当灰度共生矩阵中各 m_{hk} 数值相差不大且较分散时, ENT 值较大; 反之, 若 m_{hk} 的数值较集中时, ENT 值较小。

相关:

$$COR = \frac{\left[\sum_h \sum_k hkm_{hk} - \mu_x \mu_y \right]}{\sigma_x \sigma_y}$$

其中 $\mu_x, \mu_y, \sigma_x, \sigma_y$ 分别为 $m_x, m_y = \sum_k m_{hk}$ 的均值和标准差, 是矩阵 M 中每列元素之和; $m_x = \sum_k m_{hk}$ 是矩阵 M 中每行元素之和。相关量是用来描述矩阵中行或列元素之间相似程度的, 它是灰度线性关系的度量。

系统设计模块图:

界面: 用户操作可视化界面;

可拓方法处理模块: 处理可拓方法检测功能;

颜色直方图处理模块: 处理颜色直方图特征;

边缘轮廓处理模块: 处理图像的边缘特征;

纹理处理模块: 处理图像的纹理特征;

图像预处理模块: 对图像的一些操作, 比如灰度化, 减低光的影响等。

检测模块: 处理用户的选择条件, 以及比较图像之间的特征距离等功能。

4 实验结果

图像库共有十种类别的图像分别为人, 马, 草原, 天空, 大海, 桌子, 树, 篮球, 鱼, 电脑, 其中每种类别 10 幅图。

试验中比较了各种类别的图像检索的效果, 然后取平均值。其中可拓方法检索图像, 参数的设置根据不同类别的图像采用了不同的设置, 并且参数设置后尽量使得检索效果更好。

几种检索方法的效果如图 3 所示, 横坐标是取出检索的图像个数, 纵坐标表示检索的查准率。从图 3 可以看出, 基于颜色的检索, 平均查准率会高于边缘轮廓和纹理的检索。可拓方法的检索会高于前三种检索方法。当然以上结论和所选取的图像类别有一定的关联。

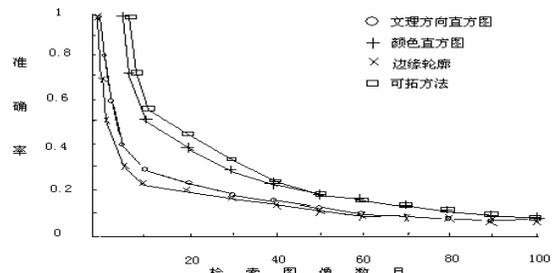


图 3 各种检索方法效果比较图

参考文献

- 1 温小兵, 康耀红. Internet 图像检索技术综述. 海南大学学报(自然科学版), 2006, 24(2).
- 2 张锋利. 基于内容的图像检索方法研究[硕士学位论文]. 长春: 吉林大学, 2005.
- 3 蔡文, 杨春燕, 林伟初. 可拓工程方法, 1997.
- 4 苏新宁. 信息检索理论与技术. 北京: 科技文献出版社, 2004.
- 5 董卫军, 周明全, 狄国华, 黎晓. 基于内容的图像检索技术研究. 计算机工程, 2005, 31(10).
- 6 丘行航, 杨群生, 战荫伟. 基于GMEM 聚类的 EMD 图像检索. 第十三届全国图像图形学学术会议论文集.
- 7 任平红, 陈鑫. 基于改进的边缘直方图的图像检索方法. 计算机技术与发展, 2007, 17(8).