

基于视觉点特征的图像检索技术研究

孙 思¹ 赵 珊¹ 魏从刚²

(河南理工大学计算机学院 焦作 454000)¹ (河南理工大学机械与动力工程学院 焦作 454000)²

摘 要 基于视觉点特征的图像检索是图像检索(Content Based Image Retrieval, 简称 CBIR)的一个子集,指利用图像中对图像的内容有显著影响的一些点特征对图像进行查询,试图在理解图像内容的基础上,检索出与示例相类似的图像。目前,CBIR 技术在商标查询、罪犯比对等领域有着很重要的应用。本文主要集中在对图像的感兴趣点特征提取技术上,针对边缘方向直方图法存在的问题,提出一种基于 Canny 边缘提取及轮廓序列矩法,首先对图像进行预处理,然后采用 Canny 边缘提取方法提取图像的轮廓,最后,将该轮廓的 3 个矩作为最后的特征向量进行检索。实验证明,提出的算法检索效果较好。

关键词 图像检索,边缘方向直方图,不变矩,轮廓序列矩

中图分类号 TP393 文献标识码 A

Research on Visual Features Based on Image Retrieval

SUN Si¹ ZHAO Shan¹ WEI Cong-gang²

(School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo 454000, China)¹

(School of Mechanical and Power Engineering, Henan Polytechnic University, Jiaozuo 454000, China)²

Abstract At present, Content Based points Retrieval is a subset of the CBIR. It tries to inquiry the trademarks similar to the sample image, based on the comprehension of image contents. CBIR in trademark has wide prospect in the stage of similar or same trademarks examination, even in the field of trademarks inquiry. This paper does mainly research on the extracting of the points characters and counting the similarity in images. Firstly, the image shape is extracted by Canny algorithm, then, the moments are introduced to image retrieval. Experiments show that the proved method has effective performance.

Keywords Image retrieval, Edge directions histogram, Invariant moments, Contour sequence moments

科学技术的快速发展,特别是计算机技术和网络技术的发展,使得现代信息处理的对象和方法都有了很大的变化。数字图书馆^[1]、数字博物馆^[2]、点播视频服务^[3]、远程教育^[4]等新的服务形式和信息交流手段的产生,对人们的生活方式和社会发展产生了重要的影响。如何有效地组织、管理和检索大规模的图像数据库,成为当前热门研究课题。基于视觉和语义的图像检索技术利用图像的颜色、形状、纹理、空间关系、语义等特征对图像进行查询,试图在理解图像内容的基础上,检索出与示例相类似的图像。它突破了传统的基于文本检索的局限,避免了用文字标签表达图像内容所带来的二义性,直接对图像进行分析,客观地抽取图像的特征,由计算机代替了大量的人工劳动,从而成为了国内外研究的热点。

本文将就图像预处理、对灰度图像的形状特征提取和相似性匹配、算法性能比较和原型系统实现这 4 个部分展开对基于视觉和语义的图像检索的论述。

1 基于内容的图像检索原理及技术

1.1 基本原理

基于视觉和语义的图像检索的基本系统结构如图 1 所示。该系统包括以下 4 个主要模块:输入模块、数据库模块、

查询模块和检索模块。

实际上,这些问题在文本检索中也存在,在 Internet 中检索时,输入某个关键字后,会查出很多信息,这些信息中有用户需要的,也有不需要的,这时要求用户根据自己的意愿进一步挑选^[5]。

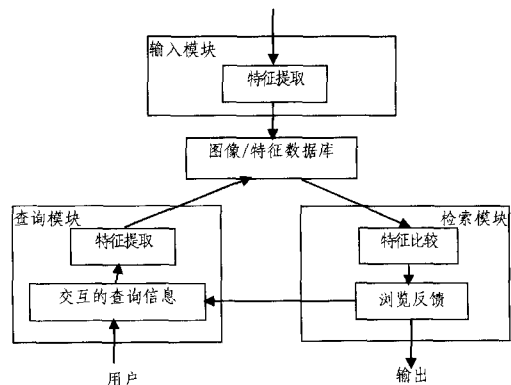


图 1 基于视觉和语义的图像检索的基本系统结构图

1.2 关键技术

图像的内容是通过图像的特征来反映的,图像的特征可分为两大类,即低层物理特征和高层语义特征。图像检索所

本文受教育部科学技术研究重点项目(210128)资助。

孙 思(1973—),女,硕士生,工程师,主要研究方向为基于内容图像检索;赵 珊(1975—),女,博士,副教授。

用到的基本特征主要有颜色、纹理、形状、图像字符、空间关系、图像的语义等。在基于视觉和语义的图像检索中,两幅图像是否相似,是指它们的特征是否相似。特征相似关系的理论研究通常使用的是几何模型,即将图像的特征看作是坐标空间中的点,两个点的接近程度通常用它们之间的距离表示,也即它们之间的不相似程度。对于几何相似距离常用的距离函数主要是 Minkowski 距离、欧氏距离、城区距离和马氏距离等。距离度量函数的定义通常要满足距离公理的自相似性、最小性、对称性、三角不等性等条件。即设 d 为距离函数, s_1, s_2, s_3 为 3 个刺激,则下面的计量定理成立:(1)自相似: $d(s_1, s_1) = d(s_2, s_2) = 0$;(2)最小性: $d(s_1, s_2) \geq d(s_1, s_1) \geq 0$;(3)对称性: $d(s_1, s_2) = d(s_2, s_1)$;(4)三角不等性: $d(s_1, s_2) + d(s_2, s_3) \geq d(s_1, s_3)$ 。设 $I = (x_{i1}, x_{i2}, \dots, x_{ik}), J = (x_{j1}, x_{j2}, \dots, x_{jk})$ 为 k 维特征向量 $i, j = 1, 2, \dots$, 则(1)Minkowsky 距离: $d(I, J) = [\sum_{k=1}^k |x_{ik} - x_{jk}|^q]^{1/q}$, 当 $q=2$ 时为欧式距离; $d(I, J) = [\sum_{k=1}^k |x_{ik} - x_{jk}|^2]^{1/2}$; $q=1$ 时为城区距离; $d(I, J) = \sum_{k=1}^k |x_{ik} - x_{jk}|$;(2)马氏距离:马氏距离的计算是以样本的统计特征为基础的,距离计算考虑了样本的概率分布特性,马氏距离定义为特征向量与某一模式类的相似性度量,计算公式为: $d^2 = (x - \mu) \Sigma^{-1} (x - \mu)^T$, 式中, x 为待测特征向量, μ 为模式类的均值向量, Σ^{-1} 为模式类协方差矩阵的逆;当各特征互不相

关时,则协方差矩阵为对角阵: $\Sigma = \begin{bmatrix} \sigma_{11}^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_{kk}^2 \end{bmatrix}$, 则 $\Sigma^{-1} =$

$$\begin{bmatrix} \frac{1}{\sigma_{11}^2} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\sigma_{kk}^2} \end{bmatrix};$$

特征向量 x 与模式类的马氏距离为: $d^2 = (x -$

$\mu) \Sigma^{-1} (x - \mu)^T = \sum_{i=1}^k \frac{(x_i - \mu_i)^2}{\sigma_{ii}^2}$, 从以上对比可知, Minkowsky 距离在计算两个甚至多个特征间距离时较为简单且具代表性。因此,在本文的实验中,相似性测度采用该距离。

1.3 检索算法的评价

基于内容的图像检索不同于传统的图像理解和目标识别技术,它是一种信息检索技术,能够从图像存储源中以用户可以接受的响应时间检索出所需的数据。它不必考虑对象的分类和识别,更关注的是基于内容的快速查找和发现。换句话说,基于内容的图像检索属于信息过滤技术,在检索过程中,由用户最终决定选用哪些图像。由于基于内容的图像检索是面向图像数据库查询的,而现有的检索表达手段并不能充分反映检索要求,因此在应用中往往并不要求查询结果是图像库中的所有相似图像,而允许有漏检。

检索算法的好坏直接影响检索效果,所以算法的评价是一个不容忽视的问题,可以采用精确度和检索率曲线(P-R Graph)评价方法来衡量各种特征提取方法和相似性匹配方法的优劣,评价时应该选用同一个图像库、相同的示例,即在相同条件下进行比较,这样的比较结果才有意义。

设 S 为图像库中所有和查询图像相关的图像集, R 为所

有检索到的图像集合, s 为一次查询中检索到的所有相关图像的数目, u 为一次检索过程中检索到的不相关的图像数目, v 为图像库中和检索图像相关但在检索中未被检索到的图像的数目,这样精确度和检索率可表示为,

$$recall = P(R|S) = \frac{P(S \cap R)}{P(S)} = \frac{s}{s+u}$$

$$precision = P(R|S) = \frac{P(S \cap R)}{P(R)} = \frac{s}{s+u}$$

2 算法实现、改进与结果分析

2.1 边缘方向直方图

边缘信息对于具有明显边缘的图像的视觉识别特别重要。边缘直方图统计图像中像素点边缘的方向信息。以边缘的方向为横轴,以某个方向上的像素点个数为纵轴,构成一维直方图。边缘上每一像素都对应此处一个边缘梯度方向,边缘可以看作由特定方向的边缘像素点构成,边缘像素的幅值信息可以完全被忽略,对边缘像素点的梯度方向的统计能够表征物体的形状。边缘直方图基于图像边缘信息的统计特性能够较好地反映目标的形状和边缘信息。

在理论上,利用边缘方向直方图法进行计算时,图像的旋转仅仅将直方图的数据进行了平移,若按所有可能的情况对直方图进行平移,便可以完全解决图像旋转不变性的问题,但计算量太大,效率很低,可行性也很小。因此,我们可以放弃对边缘统计其方向直方图的想法,考虑采用另一种方式将边缘方向直方图法和不变矩法更好地融合在一起。

2.2 基于 Canny 边缘提取及轮廓序列矩

2.2.1 图像预处理

由于计算机中大部分彩色图像的显示都是采用 RGB 空间表示的,为便于图像检索,通常要将图像进行预处理,预处理包括图像灰度化和图像平滑处理。因此,在图像的预处理阶段,我们首先将彩色商标图像转化为灰度图像。诚然,对图像进行灰度化处理将会损失图像的信息,但对本文的研究影响不大。彩色图像换算为灰度图像的方法很多,我们采用其中较常用的一种换算方法进行^[6]:

$$I = 0.30R + 0.59G + 0.11B \quad (1)$$

式中, I 表示图像像素的灰度值, R, G, B 分别表示彩色图像像素的红、绿、蓝 3 色的值。

2.2.2 图像的边缘特征提取

首先,采用 Canny 算子对图像进行边缘检测。



图 2 采用 Canny 算子对图像进行边缘检测的结果图

待处理的商标选用珠海格力电器股份有限公司的格力商标。彩色图、未加噪的灰度图和受到少许高斯噪声污染后的灰度图,分别如图 2(a) - (c) 所示,其图像大小为 120×215 。

图 2(d)、(e) 分别表示采用 Canny 算子对图 2(b) 和图 2(c) 进行边缘检测的结果。采用高斯平滑滤波器对图像进行平滑时, $\sigma=1$, 滤波器长度取值 7; 阈值 T_1 取 0.35。并且, 为了方便比较图像的效果, 我们对边缘图像进行了反色处理。实验结果表明, 即使在存在少量高斯噪声的情况下, 采用 Canny 算子也仍然可以有效地检测出图像的边缘。

2.2.3 基于 Canny 算子的轮廓矩特征的提取

上述特征仅把点当作孤立的单元, 没有考虑点整体的分布情况和图像中物体的形状结构。研究表明, 人的视觉系统对图像中物体的形状更敏感, 在许多情况下, 人们只凭物体的形状信息就可以识别物体, 这是形状有别于颜色、纹理等视觉特征的关键所在。因此, 准确地提取图像中物体的形状特征并将其与其它图像特征一起有效地应用于图像检索中具有重要意义, 必将大大提高图像检索的准确率。因此, 为了提高图像检索的效率, 考虑边缘点所隐含的形状信息是必要的。为此, 提取 7 个 Hu 不变量作为图像的形状特征。

设得到的二值图像为 $f(x, y)$, 则 7 个 Hu 不变量如下:

$$\begin{aligned} \phi_1 &= \eta_{20} + \eta_{2} \\ \phi_2 &= (\eta_{20} - \eta_{2})^2 + 4\eta_1^2 \\ \phi_3 &= (\eta_{30} - 3\eta_{2})^2 + (3\eta_{21} - \eta_{3})^2 \\ \phi_4 &= (\eta_{30} + \eta_{2})^2 + (\eta_{21} + \eta_{3})^2 \\ \phi_5 &= (\eta_{30} - 3\eta_{2})(\eta_{30} + \eta_{2})[(\eta_{30} + \eta_{2})^2 - 3(\eta_{21} + \eta_{3})^2] + \\ &\quad (3\eta_{21} - \eta_{3})(\eta_{21} + \eta_{3})[3(\eta_{30} + \eta_{2})^2 - (\eta_{21} + \eta_{3})^2] \\ \phi_6 &= (\eta_{20} - \eta_{2})[(\eta_{30} + \eta_{2})^2 - (\eta_{21} + \eta_{3})^2] + 4\eta_1(\eta_{30} + \\ &\quad \eta_{2})(\eta_{21} + \eta_{3}) \\ \phi_7 &= (3\eta_{21} - \eta_{3})(\eta_{30} + \eta_{2})[(\eta_{30} + \eta_{2})^2 - 3(\eta_{21} + \eta_{3})^2] + \\ &\quad (3\eta_{21} - \eta_{3})(\eta_{21} + \eta_{3})[3(\eta_{30} + \eta_{2})^2 - (\eta_{21} + \eta_{3})^2] \end{aligned}$$

式中, $\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}}$ 为规格化后的 $p+q$ 阶中心矩且 $\gamma = \frac{p+q}{2} + 1$, $p+q=2, 3, \dots$,

$$\mu_{pq} = \sum_x \sum_y (x-\bar{x})^p (y-\bar{y})^q f(x, y)$$

$$\bar{x} = \frac{m_{10}}{m_{00}}, \bar{y} = \frac{m_{01}}{m_{00}} \text{ 表示图像的区域重心, 而且 } m_{pq} = \sum_x \sum_y$$

$x^p y^q f(x, y)$ 。

2.3 相似性度量

设提取的形状特征向量为 $M = [\phi_1, \phi_2, \dots, \phi_7]$, 其中 ϕ_i 为第 i 个形状不变量。则两图像间的相似性距离定义如下:

$$d(Q, I) = \sqrt{(\phi_1^Q - \phi_1^I)^2 + (\phi_2^Q - \phi_2^I)^2 + \dots + (\phi_7^Q - \phi_7^I)^2}$$

2.4 实验

本实验所用到的图像库中图像共有 200 幅, 对于纯文字商标和混合商标, 输出了相似的 7 幅图像, 对于纯图形商标输出了与待检索图像最相似的前 21 幅图像, 相似度按输出位置从左到右, 再由上到下的顺序依次递减。采用改进的对 Canny 算子进行边缘提取后求取轮廓序列的矩的方法(为描述方便, 以下称为基于 Canny 边缘提取及轮廓序列矩法), 将 J0009 作为查询的范例商标, 对以上图像库中的图像进行检索, 实验证明, 提出的算法检索取得了较好的视觉效果。检索结果如图 3 所示。

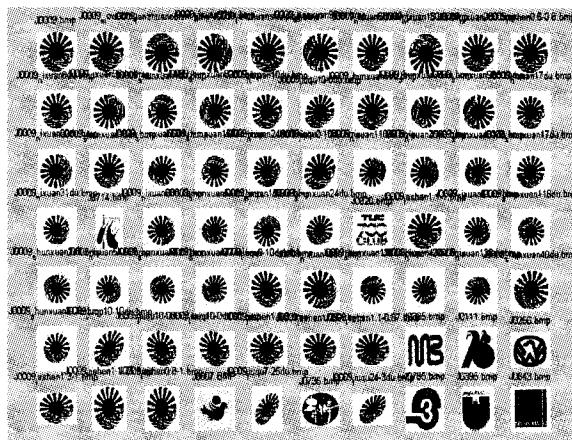


图 3 基于 Canny 边缘提取及轮廓序列矩法的检索结果

结束语 虽然本文在基于视觉特征点的图像检索领域做了一些实际的研究工作, 实现了对灰度图像快速、准确的特征提取和检索, 但是还有许多内容值得我们进一步深入探讨。本文新提出的基于 Canny 边缘提取结合 7 个边缘不变矩的算法, 首先要对预处理后的图像采用 Canny 算子进行特征提取, 而双阈值如何选取是该算法所存在的固有的问题, 迄今仍然没有一种通用的标准, 为使图像检索的效果更好, 该问题还有待进一步的研究。

图像检索的目的是为了找出近似的或者相同的商标, 为商标排查工作提供判断依据。虽然在技术手段上其对于更为客观地反映商标特征可起到硬性指标的作用, 然而视觉相似是个较主观的概念, 在基于视觉和语义检索中一般采用相关反馈技术获取在图像检索中与用户的交互信息, 以提高对图像特征分析和相似度判断的准确性。而本实验系统仅提供了支持范例图像进行查询的方式, 实现了图像特征提取、显示等基本功能, 因而, 该实验系统还有待完善。

参考文献

- [1] Fox E A, et al. Special Issue on Digital Library. Communications of the ACM, 1995, 38(4): 22-75
- [2] Mintzer F. Images on the path to the digital museum (plenary talk)[J]. IEEE ICIP, 2001, 2: 554
- [3] 章毓晋. 浅谈 Divx——一种新的视频服务[J]. 世界广播电视, 1998, 12(12): 93
- [4] Eriksson T, Goller A, Muchin S. A comparison of online communication in distance education and in conventional education[C]// Proc. Frontiers in Education, 2001, 1: T2F/20- T2F/25
- [5] 邵虹, 崔文成, 赵宏. 基于视觉和语义的图像检索技术研究[J]. 小型微型计算机系统, 2003, 24(10): 1845-1848
- [6] 贾永红. 计算机图像处理与分析[M]. 武汉: 武汉大学出版社, 2001
- [7] Castelli V, Bergman L D. Image Databases-Search and Retrieval of Digital Imagery[M]. John Wiley & Sons, Inc, 2002
- [8] Jacobs C E, et al. Fast multiresolution image querying [C]// Proc. SIGGAPH'95. 1995: 277- 286
- [9] Androustos D, Plataniotis K N, Venetsanopoulos A N. Image retrieval using the directional detail histogram[J]. SPIE, 1997, 3312: 129-137
- [10] 韩殿元. 基于梯度倒数加权和 中值滤波的图像平滑改进算法[J]. 潍坊学院学报, 2008, 06: 9-11