

一种融合语义的图像检索技术研究

彭太乐^{1,2}, 蒋建国¹, 魏仕民², 沈克^{1,2}

(1. 合肥工业大学 计算机与信息学院, 安徽 合肥 230009;

2. 淮北煤炭师范学院 计算机系, 安徽 淮北 235000)

摘要:针对基于内容的图像检索系统检索效率不高的情况,从目标图像外轮廓特征的提取、图像外轮廓特征的描述方式、图像轮廓的语义描述方式和相似度测定4个方面对基于内容的图像检索的研究状况进行了分析和研究。综合图像的外轮廓特征及轮廓的语义描述来实现对图像数据库的检索。实验结果表明,该方法提高了图像检索的精度和效率。

关键词:图像检索;语义;轮廓特征

中图分类号:TP391.4

文献标识码:A

文章编号:1673-629X(2008)03-0102-03

Research on Image Retrieval Technique Using Semantic Information

PENG Tai-le^{1,2}, JIANG Jian-guo¹, WEI Shi-min², SHEN Ke^{1,2}

(1. School of Computer & Information, Hefei University of Technology, Hefei 230009, China;

2. Department of Computer Science, Huaibei Coal Industry Teachers College, Huaibei 235000, China)

Abstract: To improve the efficiency of image retrieval based on the content, a new method is proposed using the feature extraction of figure, the description of figure features, the description of semantic information of images and the measurement of similarity degree. Achieves the retrieval of image database by combining the characteristic of out contour and semantic description. The experiments show that this method improves the precision and efficiency of image retrieval.

Key words: image retrieval; semantic information; figure features

0 引言

随着信息技术的迅猛发展,多媒体信息急剧膨胀。图像作为一种内涵丰富、表现直观的多媒体信息,长期以来一直受到人们的关注。随着各种各样的应用需求的出现,比如,商家利用丰富的图像来介绍产品,在医疗领域,医生利用医学成像技术产生的各类图像来分析病理,使图像信息资源的检索显得日益重要。

目前基于图像内容的检索技术成为计算机视觉与数据库研究领域的一个热点问题。

基于图像内容的检索技术(CBIR)是指利用图像的物理特征(颜色、纹理、形状等)从数据库中查找到与指定特征匹配的图像。但是,由于提取的视觉特征不

够直观,目前尚且没有和人的视觉特征一致的图像特征表示方法。仅依据图像的视觉特征建立索引无法体现用户检索的主观性,检索的效率和精度都不高。为了解决这些缺点,基于语义的图像检索技术应运而生。该技术面临3个方面的困难^[1]:

- ①必须提供图像语义的有效描述方式;
- ②必须有提取图像语义描述的方法;
- ③语义检索系统的语义处理方法。

为此,文中将从4个方面对图像语义的建模和表示、图像语义的抽取方法和语义检索系统的设计等进行探讨。

在对图像的轮廓进行描述的基础上提出了一种融合轮廓及图像语义的检索算法并设计了一个实验系统。首先,提取图像的轮廓,然后从图像的轮廓语义出发,建立图像的轮廓特征和高层语义之间的联系。为了提高图像检索的效率,分别从低层视觉特征和高层语义层次出发,运用不同的图像匹配技术来实现检索,取得了明显的效果。最后,实验证明该方法对于含有大量复杂图像的检索效果明显。

收稿日期:2007-06-19

基金项目:国家自然科学基金(60573026);教育部科学技术研究重点项目(206068);安徽省教育厅省学术技术带头人科研资助项目(2005hzb24);安徽省自然科学基金(070412052)

作者简介:彭太乐(1974-),男,安徽萧县人,硕士,讲师,研究方向为数字图像处理技术;蒋建国,博士生导师,博士,教授,研究方向为分布式智能控制系统、数字图像分析与处理;魏仕民,教授,博士,研究方向为网络与信息安全。

1 基于轮廓特征的检索方法

1.1 轮廓特征提取

由于实验系统要求输入图像为 BMP 格式,对于非 BMP 格式的图像,在进行操作前要先进行格式转换。对获取的图像进行去噪、增强、几何调整、二值化处理。

文中采用数学形态学方法来去噪。在 MATLAB 环境下,对图像进行最大程度的腐蚀,对于图像中细小的噪音,采用开运算滤除噪音。选择合适的结构元素大小,先进行一次腐蚀运算,滤除噪声点,再进行一次膨胀运算,来填补空洞,也就是一次开运算,能够滤除大部分的噪声点,得到较为清晰的目标。

轮廓是图像的一个重要特征。对利用轮廓特征的图像检索,人们已提出了许多不同的方法,比较典型的有几何参数法^[2]、不变矩法^[3]、边界跟踪法、小波轮廓表示法、局部特征频率法、隐马尔可夫法^[4]等。图像轮廓的提取一般有基于区域和基于边界的方法。其中,边缘检测的方法尤为广泛。利用 CANNY 算子和零交叉法进行边缘检测,以改进的边界跟踪法提取图像的边缘信息,进一步提取图像的轮廓特征向量。影响边界跟踪效果的主要因素有三个:第一个因素是图像分割的效果。图像分割的质量的优劣、区域界限定位的精度直接影响到图像的区域描述以及图像的分析理解;第二个因素是图像跟踪起始点的选择,起始点的选取直接影响到轮廓提取的速度和最终效果,文中以启发式信息来确定选择起始点;第三个是跟踪遵循的规则,不同的规则适用不同的算法,以不同的轮廓跟踪算法提取轮廓,效果也截然不同。

为了精确标识图像边缘曲线的轮廓特征,引入了凹凸度和平均张角来描述图像的轮廓特征。首先求出图像边缘曲线的所有拐点 $I(x_k, y_k)$, 如图 1 所示, $I(x_k, y_k)$ 表示 I 分量对应像素点。设拐点数 n , 以任一拐点 P_M 为起点,按顺时针方向连接相邻拐点 P_N , 计算 P_MP_N 中间各点 P_i 到直线 P_MP_N 之间的距离 d_i , 求取最大垂直距离 α_1 , 设直线 P_MP_N 的方程为:

$$A_1x + A_2y + A_3 = 0$$

$$\text{则 } d_i = \frac{|A_1x_i + A_2y_i + A_3|}{\sqrt{A_1^2 + A_2^2}}$$

再以 P_N 为起点,按顺时针方向连接相邻的下一拐点 P_S , 求取两点之间的曲线到两点连线的最大距离 α_2 , 依此类推, 求出 $\alpha_3, \alpha_4, \dots, \alpha_n$ 。

定义 1(凹凸度)

$$\text{roughness} = \frac{\text{MAX}(\frac{|A_1x_i + A_2y_i + A_3|}{\sqrt{A_1^2 + A_2^2}})}{\sqrt{(x_j - x_{j+1})^2 + (y_j - y_{j+1})^2}}$$

其中 $j, j+1$ 为相邻拐点。

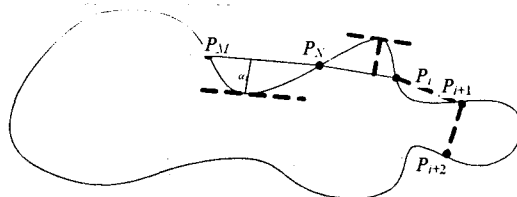


图 1 给定轮廓

设定:

M, N 为轮廓上连续的两个拐点。

$K_c(M)$ 是 P_MP_N 的长度

$M = (x_M, y_M)$

$N = (x_N, y_N)$

P_i 为 P_MP_N 上非 M, N 的轮廓点。

对于 P_MP_N 中的 P_i , 可求得 P_i 对 M, N 所张角的余弦值为:

$$\cos(P_i) = \frac{(x_i - x_M)(x_i - x_N) + (y_i - y_M)(y_i - y_N)}{\sqrt{[(x_i - x_M)^2 + (y_i - y_M)^2][(x_i - x_N)^2 + (y_i - y_N)^2]}}$$

轮廓段 P_MP_N 的平均余弦值为:

$$\text{Avecos}(P_MP_N) = \frac{1}{K_c(M)} \sum_{P_i} \cos(P_i), P_i \in P_MP_N$$

定义 2(平均张角) 以任一拐点 P_M 为起始点, 顺时针连接拐点 P_M, P_N, P_i 是 P_MP_N 上的轮廓点, 则平均

$$\text{张角: } \beta_i = \arccos\left(\frac{1}{K_c(X)} \sum_{P_i} \cos(P_i)\right)$$

其中 X 为拐点。

文中根据凹凸度和平均张角的概念提出了基于拐点分段的轮廓特征提取算法。

1.2 具体算法步骤

Step1 取范例图像进行预处理。

对获取的具有多拐点轮廓的图像进行格式转换、去噪、增强、二值化、轮廓提取、提取骨架。

Step2 利用改进的边界跟踪法提取图像的外轮廓并对图像进行编码, 依次提取曲线各点的坐标及拐点。

Step3 以任一拐点开始, 求取凹凸度 roughness 和平均张角 β , 分别依次求出 $(\text{roughness}_1, \cos(\beta_1))$ 、 $(\text{roughness}_2, \cos(\beta_2))$, \dots , 以 $((\text{roughness}_1, \cos(\beta_1))$ 、 $(\text{roughness}_2, \cos(\beta_2))$, \dots) 为特征矩阵。

Step4 从图像数据库读取图像数据进行预处理, 转向 Step 2。

Step5 根据相似度函数进行相似性比较。

文中采用非几何距离作为相似度度量的准则。设两幅图像 P 和 Q 的特征矩阵分别为 A 和 B , 维数分别

为 M, N 。特征矩阵的维数 M, N 可以不同。

$$\text{SIMI}(P, Q) = \frac{\text{MAX}(\text{CON}(|A_i - B_j| < \delta))}{\text{MAX}(M, N)}$$

其中 $\text{CON}(|A_i - B_j| < \delta)$ 表示以特征矩阵 A 的第 i 列向量或以特征矩阵 B 的第 j 列向量开始连续满足 $|A_i - B_j| < \delta$ 的列数, δ 为给定的经验值。其中 SIMI 越大, 两幅图像越相似。

2 图像语义的提取

基于内容的图像检索的特点是基于一幅图像底层特征的匹配是一种相似性匹配, 检索的效率不高。无法体现用户检索的主观性。这就要求在检索系统中要实现图像视觉特征进行语义标识, 支持广泛的用户要求, 对整个语义表示和处理的过程进行必要的抽象。这个抽象的过程, 通过建立图像语义模型, 整合现有的系统和技术来实现。

图像语义层次分为三层, 即底层特征层、对象层和概念层^[5]。由于图像语义的含义取决于图像底层固有的视觉特征, 从视觉特征到对象的空间关系和主观意识语义, 这个纵向的发展过程是图像语义中最主要、最基本的内容。层次结构是建立好语义模型的基础。典型的层次化语义模型如图 2 所示。

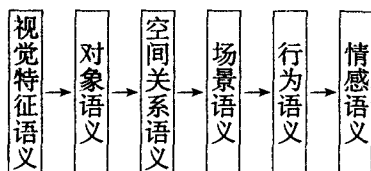


图 2 层次化语义模型

结合图像的轮廓语义信息进行检索, 依照对图像的语义描述, 进行相似性检索, 尽管每个用户对图像的描述不同, 即具有主观性, 但表达的主题意思相近。图像的语义, 即对图像内涵的解释。简单语义通常就是图像的主题词表示, 复杂语义则是对图像内容的叙事型描述。文中采用简单语义的标记方式, 即图像语义。根据领域专业知识, 提取出代表图像基本轮廓语义的关键字集合, 并用关键字集合作为图像的语义特征向量 L ^[6,7]。

对任意两图像之间的相似性匹配就转化为语义特征向量之间的匹配。实际应用中相似性匹配可以使用欧氏距离、内积距离等来表示。

在文中, 对任意两幅图像 P 和 Q , 利用距离测量公式^[7]表示:

$$\text{SIM}(P, Q) = 1 - \frac{\sum_{i=1}^n W_{A_i} \times W_{B_i}}{\sqrt{(\sum_{i=1}^n W_{A_i}^2)(\sum_{i=1}^n W_{B_i}^2)}}$$

其中, n 为关键字数。

3 系统检索的流程

基于以上的算法描述, 给出一个 CHIR 的实验系统, 系统进行检索的算法流程图如图 3 所示。

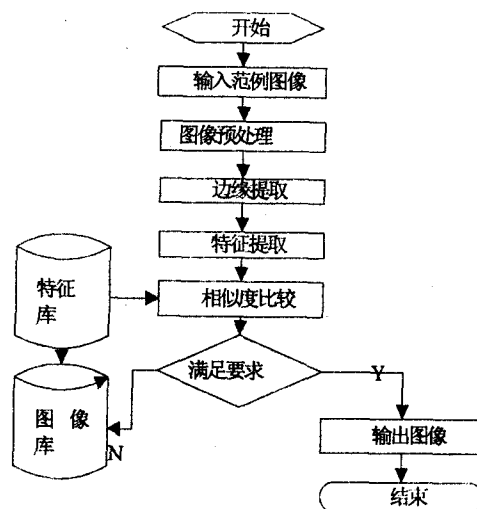


图 3 基于内容的检索系统流程图

4 实验结果分析

为了验证文中所述算法, 根据提出的方法, 进行了充分的实验。本实验使用的图像有 410 张来自互连网, 其余的 40 张为手绘图。其中, 蝴蝶图像 50 张, 其他图像 400 张。由这些图像组成图像库。

文中以查准率 (Precision) 和查全率 (Recall) 来评价检索算法的优劣。查准率指返回的结果集中有效图像的比率, 用于系统排除无关图像的能力。查全率指返回的结果集中有效图像占数据库中所有相似图像的数量比率, 用于测量系统检索相关图像的能力。查全率和查准率越高, 说明该检索算法效果越好。公式:

$$\text{查全率} = \frac{\text{检索出的相关图像}}{\text{所有相关图像}}$$

$$\text{查准率} = \frac{\text{检索出的相关图像}}{\text{检索出的所有图像}}$$

由检索结果知: 结合语义进行检索有较高的检索效率, 也体现了用户检索的主观性。下面列出 6 次检索的检索效果, 如表 1 所示。

表 1 实验效果统计表

δ 值	查全率	查准率	δ 值	查全率	查准率
$\begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$	0.52	0.081	$\begin{pmatrix} 0.02 \\ 0.02 \end{pmatrix}$	0.4	0.083
$\begin{pmatrix} 0.1 \\ 0.5 \end{pmatrix}$	0.44	0.071	$\begin{pmatrix} 0.015 \\ 0.005 \end{pmatrix}$	0.38	0.097
$\begin{pmatrix} 0.05 \\ 0.03 \end{pmatrix}$	0.4	0.08	$\begin{pmatrix} 0.002 \\ 0.002 \end{pmatrix}$	0.34	0.106

总的来说, δ 值越小, 检索的查准率就越高, 查全率就会下降。对于 $\delta = \begin{pmatrix} 0.002 \\ 0.002 \end{pmatrix}$ 时的部分检索结果如

(下转第 108 页)

2.4 检测结果分析

由于方法一相对于方法二来说相对复杂,并且需要确定亮度 Y 的值,所以文中采用模型二来提取手形,即基于 (Y, Cr, Cb) 空间下的肤色检测方法二,设定 Cr 和 Cb 值的范围为 $120 \leq Cr \leq 172$ 且 $77 \leq Cb \leq 118$ 。最后利用数学形态学方法腐蚀和膨胀去噪。其肤色检测结果见图 2。

由实验知,当背景相对简单时,肤色检测出的手势区域就比较清晰明了。如图 2(a),检测结果受其他方面的制约,比如图片的分辨率、光照、清晰度等都有影响,需要说明的是,当手形图片质量较高时,检测的效果也相对较好,但处理速度也相应减慢,所以一个比较合适的图片质量对检测结果也有很大的影响。当背景相对复杂并且图片质量不高时,肤色检测的效果就相对要差,见图 2(b) 和 (c)。但此方法对背景几乎没有什么限制,并且易于实施。

另外,在手语分类中已经提到“视觉类”手势与身体和脸部有相对位置关系的词汇,这类词有跟脸部接触的,也有跟脸部有一定距离的,而肤色检测后最主要的区域是手和脸,如果检测一些手跟面部相交的手语词时,要把手和脸的图像分割开来,对于肤色检测来说,是比较难以克服的问题,还需要进一步的改进。

3 结 论

文中主要内容是利用肤色检测分割复杂背景下的

手势区域,根据分割结果,可以获取到手的形状、位置、方向等参数,便于利用 HMM 模型或是其他的技术进行手势识别。但还有一些难以解决的问题,需要进一步的研究和改进。

参考文献:

- [1] Starner T, Weaver J, Pentland A. Real-time American sign language recognition using desk and wearable computer based video[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20 (12): 1371-1375.
- [2] Jones M J, Rehg J M. Statistical color model with application to skin detection [R]. UN: Cambridge Research Laboratory Technical Report Series, CRL 98/11, 1998.
- [3] 吴相豪, 申铨京. 基于像素的三种肤色检测模型的比较与研究[J]. 计算机应用研究, 2003, 20(精扩本): 430-432.
- [4] Yang Ming-Hsuan, Ahuja N. Gaussian mixture model for human skin color and its application in image and video databases [C]//Proceedings of SPIE'99. San Jose, CA: [s. n.], 1999: 458-466.
- [5] 段立娟, 崔国勤, 高文, 等. 多层次特定类型图像过滤方法[J]. 计算机辅助设计与图形学学报, 2002, 14(5): 404-409.
- [6] 范晓, 申铨. 基于 IE 浏览器的色情图片过滤器[J]. 吉林大学学报: 信息版, 2004, 22(6): 631-637.
- [7] 任海兵, 祝远新, 徐光, 等. 连续动态手势的时空表观建模及识别[J]. 计算机学报, 2000, 23(8): 824-828.

(上接第 104 页)

图 4 所示。



图 4 基于轮廓特征的图像检索结果

5 结束语

基于内容的图像检索系统是一个复杂的系统,图像的各种视觉特征都会影响系统检索的效果。文中仅从图像的外轮廓和外轮廓的语义两个方面探讨了图像检索的效果。

在检索的过程中若综合考虑图像的其他特征则系统检索的精度和效度都会进一步提高。

参考文献:

- [1] 王惠锋, 孙正兴, 王箭. 语义图像检索研究进展[J]. 计算机研究与发展, 2002, 39(5): 513-521.
- [2] Niblack W, Barber R. The QBIC project: querying images by content using color, texture and shape[M]. [s. l.]: SPIE, 1993: 173-187.
- [3] Mehre B W. Content based image retrieval Using a Composite color-shape approach[J]. Information Processing & Management, 1998, 33(30): 319-337.
- [4] Trzegnies C, Bandera A. Planar Shape indexing and retrieval based on Hidden Markov Models[J]. PRI. 2002, 23 (20): 1143-1151.
- [5] 陈剑赞, 老松扬, 吴玲达. 基于内容的图像检索的发展最新趋势[J]. 计算机工程与应用, 2002, 38(10): 47-49.
- [6] 蔡菲, 蔡珣, 史同广, 等. 一种基于形状特征的图像检索方法[J]. 计算机应用与软件, 2005, 22(12): 98-99.
- [7] 王李冬, 邵晓英, 巴特尔, 等. 综合纹理特征和语义信息的医学图像检索[J]. 计算机应用, 2005, 25(10): 2385-2386.