

DOI: 10.3969/j.issn.1673-4785.201401045

网络出版地址: <http://www.cnki.net/kcms/detail/23.1538.TP.20150320.1036.001.html>

基于手绘草图的图像检索技术研究进展

辛雨璇, 闫子飞

(哈尔滨工业大学 机电工程学院, 黑龙江 哈尔滨 150001)

摘要:通过对现阶段基于草图的图像检索相关技术的研究, 尝试构建了领域无关的基于手绘草图的图像检索系统框架, 并分别对手绘草图预处理、草图特征表示、草图匹配及图像反馈等系统阶段所涉及的相关技术及其发展进行梳理, 进而对基于手绘草图的图像检索系统的相关应用进行总结, 展望了手绘草图检索系统在自然人机交互、普适计算、大数据背景下的研究趋势。

关键词:手绘草图; 特征表示; 图像检索; 人机交互; 大数据

中图分类号: TP391.41 **文献标志码:** A **文章编号:** 1673-4785(2015)02-0167-11

中文引用格式: 辛雨璇, 闫子飞. 基于手绘草图的图像检索技术研究进展[J]. 智能系统学报, 2015, 10(2): 167-177.

英文引用格式: XIN Yuxuan, YAN Zifei. Research progress of image retrieval based on hand-drawn sketches[J]. CAAI Transactions on Intelligent Systems, 2015, 10(2): 167-177.

Research progress of image retrieval based on hand-drawn sketches

XIN Yuxuan, YAN Zifei

(School of Mechatronics Engineering, Harbin Institute of Technology, Harbin 150001, China)

Abstract: This paper build a framework for the domain-independent image retrieval system based on hand-drawn sketches by researching the existing sketch-based image retrieval related technologies. The relevant technologies and their development of the involved system stages such as: hand-drawn sketch preprocessing, sketch feature interpretation, sketch matching, and image feedback were also outlined. Applications relating to sketch-based image retrieval were summarized. Finally, the coming trends of the hand-drawn sketch-based image retrieval system under the background of natural user interface, pervasive computing and big data were forecasted.

Keywords: hand-drawn sketches; feature interpretation; image retrieval; human-computer interaction; big data

互联网的发展和普及拓宽了人们获取信息的途径, 也使得人们处于海量信息世界之中。这些信息不仅包括文字信息, 也包括视觉信息。视觉感知是人类从客观世界获取信息的主要来源, 也是人们认知的一种重要方式。如何从海量图像数据库中快速有效地进行信息检索, 建立更为有效的图像描述, 是人们一直关心的问题。20 世纪 70 年代提出了基于图像文本标注的检索方式, 即检索图像的文字表示。

这种检索方式, 由于人工标注的主观性^[1], 导致检索效果并不理想。20 世纪 90 年代初研究者们又提出基于图像内容的检索方式, 通过提取图像的底层特征^[1], 如颜色、纹理等, 增加了图像检索的有效性, 而随着便捷化, 小型化无线设备的发展和“数字水墨”或电子纸的出现, 笔式交互成为了新型人机交互方式之一。人机交互界面由桌面环境模拟了笔纸的环境, 促进了笔式交互的进程, 也为利用手绘草图进行图像检索奠定了基础。基于手绘草图的图像检索也属于基于图像内容检索范畴。草图与图像不同, 草图是人依靠记忆和模仿来进行信息的表达, 并且勾勒草图是人与生俱来的能力, 古代人就会画简

收稿日期: 2014-01-22. 网络出版日期: 2015-03-20.

基金项目: 国家自然科学基金资助项目(61102037); 哈尔滨工业大学科研创新基金资助项目(HIT. NSRIF. 2015057).

通信作者: 闫子飞. E-mail: cszfyan@gmail.com.

图,因此利用手绘草图进行图像检索更接近自然的人机交互方式。然而图像间的相似性不仅仅建立在颜色、纹理等特征^[1]的基础上,Eakins^[2]根据检索的复杂性将用户查询分成了 3 个层次,其中第 1 层次为视觉相似检索,第 2、3 层为语义图像检索。视觉相似检索主要利用图像的形状信息检索,语义图像检索的目的则是使计算机在检索图像时拥有接近人的理解水平。目前基于内容的图像检索结合相关领域语义知识已经取得大量的研究成果^[3],但都分别针对某一特定领域(即领域相关)处理,典型的包括:IBM 的 QBIC^[4]系统、MIT 的 PhotoBook^[5]系统等。基于草图的图像检索也取得了很大进展。孙正兴^[6]对基于草图的人机交互技术早期的研究进展进行了归纳和总结,并提出了基于草图的人机交互系统模型,此模型高度概括了当时绝大多数的基于草图的人机交互技术及应用,但这些应用大部分依赖于领域知识,只能应用到特定的领域。而本文针对基于草图的人机交互技术的图像检索领域,主要利用视觉信息,并结合语义知识检索,尝试建立领域

无关即通用的图像检索框架;并对该框架所涉及的相关技术进行了较详细的归纳;对手绘草图检索系统现阶段的应用进行总结;并展望了手绘草图检索系统在草图特征表示、多通道输入、检索机制和系统评价的建立等方面在未来的研究趋势。

1 手绘草图检索系统概述

在对大量基于手绘草图检索系统分析的基础上,本文提出了手绘草图检索系统框架图,基于手绘草图的图像检索系统框架一般可分为手绘输入模块、手绘草图预处理模块、草图特征表示模块、草图特征匹配模块、图像反馈模块、数据库模块等 6 个阶段,如图 1 所示。现有手绘草图系统的草图特征表示阶段根据特征提取单元不同一般可分为基于笔划描述的草图特征表示、基于组合图形的草图特征表示和基于形状特征的草图特征表示。图像反馈模块主要体现在对查询结果进行反馈、通过机器学习方法度量用户绘制意图等方面。

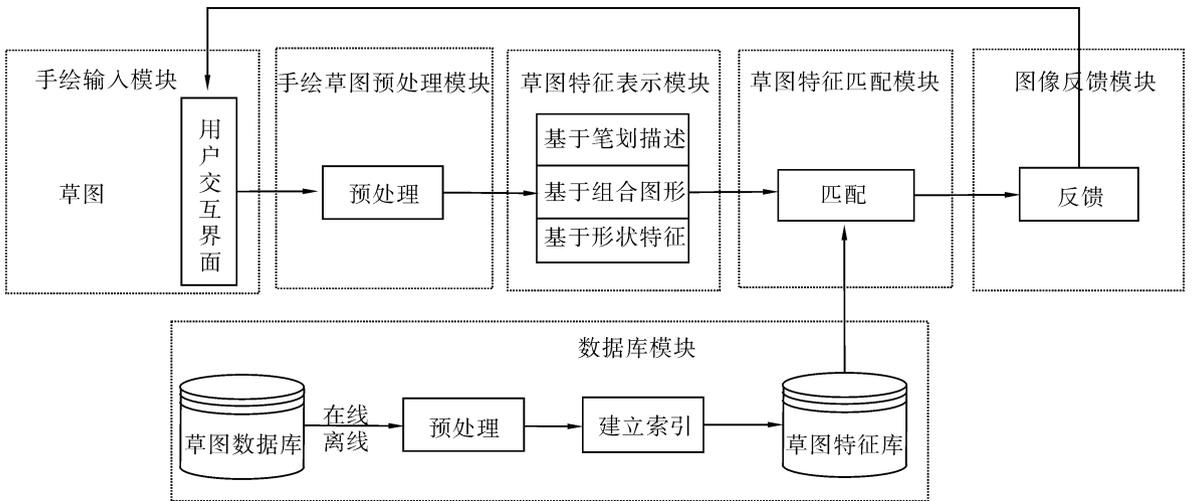


图 1 基于手绘草图的图像检索系统框架图

Fig.1 The system framework of hand-drawn sketch-based image retrieval

1.1 手绘草图的预处理

对手绘草图的预处理主要用于消除由用户绘图习惯不同以及草图固有的随意性而引发的噪声。噪声处理包括曲线闭合处理、冗余点去除、笔划合并等。不同的手绘草图检索系统采用不同的草图预处理方法,为草图特征表示阶段做准备。下面分别介绍基于笔划描述的草图特征表示、基于组合图形的草图特征表示和基于形状特征的草图特征表示所对应的预处理方法。

基于笔划描述的草图预处理阶段主要对草图笔划进行处理:采样、冗余点去除、平滑处理,并计算笔划的特征值,比如方向等信息^[7-8]。李雪峰等^[9]还

添加了笔划合并以及多旋转检测。笔划合并和多旋转检测都考虑了多笔划输入的情况。汪文睿等^[10]还提出了对草图笔划进行简化的思想,并将对带阴影区域的草图的处理称为草图简化,通过测量笔划密度进行阴影区域的判断。

基于组合图形的草图特征表示,主要考虑图元间的位置关系,因此在预处理阶段,首先将图形分割为基本图元(圆形、三角形等),然后对基本图元进行噪声处理。孙正兴等^[11]认为对基本图元的噪声处理主要包括冗余点去除、聚点消除、端点校正和闭包计算。冗余点通过设定某个结点到相邻结点的阈值进行判断。而聚点是那些点密度较大的点,通过

计算草图点密度并设定密度的阈值进行聚点的判断,并用重心代替消除。端点校正则是处理起点和终点难重合的情况。处理方法为:将顶点边反向延长交叉,交叉部分比例小于阈值则删除。闭包计算,利用平面点集的闭包算法将凹多边形变为凸多边形,为后期图元拟合做准备。

而基于形状特征的草图特征表示大部分以提取轮廓特征为主,因此,在预处理阶段,常采用八连通区域自适应追踪算法来获得草图边界轮廓^[12]。龚健等^[12]在此基础上改进,提出基于四向连通种子填充的封闭性判断算法进行封闭性判断。然后对于封闭和非封闭的区域,分别用不同的算法提取轮廓。此外,Eitz等^[13]还提出应用Canny边缘检测,并设定像素位置距离标准差去除不属于草图的线。然而Canny算法仅利用了图像的亮度信息,而Berkeley^[14]边缘检测算法综合利用了图像的亮度、色度和纹理三方面的信息,计算出每个像素作为边界的概率并用图中每个像素点的灰度值表示,得到概率图。这种表示方式与人对图像轮廓的理解更相近。

1.2 草图的特征表示

草图特征表示旨在描述预处理后草图的特征。本文根据特征提取单元不同,将草图特征表示方法分为:基于笔划描述的草图特征表示、基于组合图形的草图特征表示和基于形状特征的草图特征表示。

1.2.1 基于笔划描述的草图特征表示

由于笔式交互以笔划为绘制基本单元,因此,可以提取笔划的特征来表示草图。对笔划的特征提取分为离线特征提取和在线特征提取。离线特征提取即在用户抬笔后,特征提取才开始。早期基于笔划的特征表示,不仅是离线特征提取,还需要训练用户手势。手绘草图匹配的难易与用户草图绘制的自由度相关,对用户输入限制越多的草图,匹配越准确。如1991年Rubine等^[15]提出手势特征提取工具GRANDMA,通过学习指定的单笔划手势,构造手势识别器来表示特征,该方法采用包围盒对角线的长度及倾斜度、起点与终点的距离等11个几何特征和笔划的最大速度值、起点到终点的时间值2个动态特征来描述单笔划图形。Rubin的工作后来被Long^[16]延伸,他提出新的特征集合,但仍然需要严格的训练,即用户要用同样的方式(逆时针画圆和顺时针认为不同)绘图。为了减少对用户输入的限制,Hse等^[17]提出用Zernike矩描述子来描述用户输入的笔划,这种方法与笔划顺序、完成同一个图形的笔划数量和方向无关,同时满足平移、缩放和旋转的不变性。此后孙正兴等^[18]提出使用笔划的空间

关系进行笔划的自动成组。在草图特征表示方面,他采用朴素贝叶斯分类器,在离线情况下对草图和样本分割,得到基本图元,然后计算一个六维度特征向量,最后对用户草图得到的特征向量和样本特征向量分布进行比较,完成笔划集的特征表示。

在线特征提取与用户的书写绘画过程同时开始、同时结束。用户可以及时了解计算机是否正确认识自己的意图。随着用户的输入,新像素点加入,特征也随之更新,计算机应找到与当前特征最匹配的图形。早期的研究,只能匹配几种几何图形^[19]或算法复杂度高^[20]。李俊峰等^[21]提出增量式意图提取的草图识别算法采用增量式意图提取的方式理解用户的勾画意图。增量式意图提取通过不断收集用户的输入信息,并分析历史记录,然后对现有信息进行分析,进而根据当前信息修正以前生成的意图段落,是一个迭代的问题求解过程。实验证明,增量式意图提取通过迭代的修正笔划,可以识别多种输入图形。近期,微软公开了一种基于手绘草图的在线图像检索引擎——MindFinder^[22],它将由笔划形成的线条所包含的形状信息转化为一种由像素坐标与方向角共同表示的边缘像素词典,简化对形状特征的描述的同时保持了轮廓的空间信息。之后,上海交通大学^[23]在MindFinder基础上做了改进,它认为提取的笔划的方向角特征中包含冗余信息,不能准确地描述草图的轮廓信息。通过将方向图组合为轮廓显著性图,并分别对主要区域和感兴趣区域进行查找可以克服MindFinder不能找到位置、大小相差较大的场景图像的情况。然而基于笔划描述的特征表示方法,无论在进行离线、在线特征提取后都难以找到图形的唯一表示,并且对于笔划复杂的图形,适应性较差,仍需进一步提高笔划特征的可分辨性。

1.2.2 基于组合图元的草图特征表示

在用户的实际应用中,所要检索的图形有可能是种类繁多的一个或多个图元构成的复杂图形。而多个图元构成的复杂草图可以利用组合图元来表示草图特征。

基于组合图元的草图特征表示,在预处理阶段已经完成了图元分割、噪声的处理。特征表示阶段,对基本图元进行识别,最终利用图元间的空间关系进行检索。图元识别是寻找与输入草图最相近并接近用户输入的图形。对于图元的识别已经有很多研究。如Revankar等^[24]提出用独立的几何模型去识别和修正手绘的几何草图,用图表示线的关系,并设

定连通性、相对方向、相等和平行性的阈值判断来确定线与线的关系去修正手绘笔划,使其更接近用户想画的图形。Sezgin 等^[25]提出的系统用探测最高曲率和最低速度点的方式去识别图形。但是 Sezgin 的系统仅可以识别出简单的几何图元,包括直线、圆和由直线和曲线组成的简单合成图形。而 Yu 等^[26]提出的领域无关的草图识别系统,对图元的识别扩展到了折线、椭圆、弧和螺旋结构。他认为仅用曲率判定,容易被噪音所误导,因此引入特征面积的概念,并将方向、曲率和特征面积结合将用户所绘制草图近似为标准的图元。后来 Paulson 等^[27]提出一个精准的手绘草图识别系统 PaleoSketch,他认为像 Sezgin 和 Yu 提出的那些简单图形的识别器在手绘草图领域的通用性不强,很难识别由图表和草图构成的复杂符号。所以 Paulson 在 Sezgin 和 Yu 的基础上,又引入了 2 个新的特征值,NDDE (normalized distance between direction extremes) 和 DCR (direction change ratio), NDDE 是用总笔划长度除以最高方向值和最低方向值之间的距离,得到方向相反线的百分比;DCR 则是最大方向变化量占平均方向变化量的百分比,用于更好地区分折线和曲线,在识别阶段获得较好的效果。此外,孙正兴等^[11]和团队在其开发的 Smart Sketchpad 系统中提出用引力模型(认为点与点有相互吸引的趋势),通过设定阈值将点合并的方法,对基本图元分类。并提出图元内规整和图元外规整的方法。例如,判断三角形两边近似等长则规整为等腰三角形的方法为内规整。利用图元间相邻信息,将相近图元规整为相同大小的方法为外规整。该系统使用的是识别图元的基本方法。李雪峰等^[9]则又对图元进行了扩充,支持直线、折线、圆弧、曲线、同心螺旋线、异心螺旋线、椭圆、圆等 8 种基本图元以及多种图元组成的复杂图形。识别出基本图元后,可以将不同复杂层次的图形元素抽象,从而获得统一的表示。例如组合图形属性包括图形元素类型、坐标、尺寸等。组合图形的空间关系包括图形元素的相对位置关系、相对方位、相对旋转等。张莉莎等^[28]则提出将复杂图形描述转化为不同信息粒度的属性和空间关系表示。但草图本身具有模糊性和不确定性,加上线条数目繁多,难以对其施加判断和约束。此外,组合图形越复杂,空间关系的信息维数越多,计算量越大。因此基于组合图元的草图检索还需要依赖相关领域知识和上下文信息简化对复杂组合图形的表示过程。

1.2.3 基于形状特征的草图特征表示

基于组合图形的草图特征表示,着眼于草图中

图元的空间关系。而基于形状特征的特征表示,着眼于草图自身的外在形状特征。早期的研究采用闭包盒大小、最大内接三角形、或傅里叶描述子等作为形状特征表示^[29-31],检索效率低,时间复杂度高。之后很多研究主要在草图轮廓的基础上提取轮廓的全局特征或局部特征作为草图的形状特征表示。全局特征着眼于整幅图像,能更好地描述图像中物体的相对位置。如 Chee 等^[32]采用 MPEG-7 标准中提出的边缘直方图(edge histogram descriptor, EHD),具有描述图像像素变化方向的能力。它通过统计每个子图块含有垂直、水平、45°、135°及无方向 5 种边缘特征形式的个数,形成了五维特征向量,因此 EHD 利用 5 个方向的直方图提取特征,更注重图像的整体信息。但是 EHD 对于子图和方向的划分十分粗略,描述轮廓的能力有限。此外,李曼舞等^[33]提出对草图预处理后,对质心距离形状描述子进行傅里叶变换并除以直流成分生成傅里叶边界描述子描述草图的形状特征。高竹红等^[34]在李曼舞的基础上,对生成的傅里叶描述子进行傅里叶反变换再得到图像轮廓,将轮廓点到质心点进行连接得到轮廓结构图,构造邻接矩阵并提取矩阵的特征向量进行检索。而吴明珠等^[35]提取轮廓成对几何直方图作为区域描述子,成对几何直方图是链码编码直方图的一种扩展,是由角度和距离 2 个维度构成的二维直方图。钱晶等^[36]提出使用一种仿射变换自适应骨架提取算法来提取对象骨架,构造骨架树描述符进行图像表示。最后将图像的骨架特征与轮廓特征结合进行检索。此后,钱晶等^[37]在之前只采用形状特征表示的基础上,又结合了颜色和纹理特征进行表示。为了加快检索速度,对每一个图像特征都使用了粗尺度和细尺度特征。先用粗尺度快速计算区域相似度,丢弃不匹配的区域;再用细尺度特征完成相似度的精确计算。

特别地,Eitz 等^[13]提出著名的特征袋(bag-of-features)描述子,通过对形状内容描述子、星点描述子、改进的标准方向梯度直方图描述子分别进行实验和评估发现,改进的标准方向梯度直方图描述子表现最优。通过随机提取图像中的感兴趣点,以每个感兴趣点的邻域为单元,计算梯度方向直方图并取其主方向作为该感兴趣点的边缘特征,最终形成了一个特征袋。Lukas 等^[38]认为之前的系统都缺少在互联网环境下检索图像的能力,他提出将方向梯度直方图和离散距离变换(discrete distance transformation, DDT)相结合的方法实现在线检索,通过对小数据库和较大数据库进行实验都取得了良好的

效果。然而,基于形状特征的草图特征表示难以区分形状相似的不同物体,因此,还需要结合语义对图形分类或者利用草图识别提高检索精度。

1.3 草图的特征匹配

为了加快检索速度,需要对图像库也进行预处理,形成图像特征库。将草图的特征和图像特征库进行匹配,即特征匹配阶段。对图像库预处理首先缩小图像库中图像大小,然后一般采用与输入草图同样的处理方法对库中的图像进行预处理和特征提取。匹配时,需要根据不同的草图特征表示方法找到合适的相似度计算方法。基于笔划描述的草图检索方法,大部分采用转换为计算特征向量之间相似性度量的方法,其中欧式距离由于计算简单、效果好,被许多系统采用^[12,36-37]。也可以表示为点的序列,计算2个序列中相对位置相同项占所有项的比例,即相关系数法^[13,39-40]。基于组合图元的特征表示,在识别基本图元后,首先要进行图元成组。然后进行部分结构相似性计算和整体相似性计算。部分结构相似性计算包括图形构成相似性计算和图形骨架间的相似性计算^[11]。实质上就是把图元构成和骨架结点都转换为多维向量或结点数组的计算。还可以基于空间关系计算组合图元的相似度,将空间关系转化为空间拓扑图和层次结构图,用拉普拉斯图谱转换为特征向量^[41-42]。最后利用欧式距离进行相似度计算^[41]。对于复杂的组合图形的检索方法,可以在匹配阶段,找到源空间关系图和目标空间关系图的映射关系,但是匹配复杂度很高。也可以采取约束的部分枚举算法^[28],满足顶点匹配约束或边匹配约束的序列可以从当前状态略去无效序列直接到达后继状态,节省了大量的计算。此外,基于形状特征的草图特征表示除了转化为向量特征表示,还包括直方图特征表示和骨架特征表示。基于直方图特征的草图匹配可以采用复杂度很低的直方图相交算法来计算直方图的距离^[35],数值越小,相似度越高。基于骨架特征的草图匹配,对骨架树的分支分别计算相似度,对于子节点需要由分支相似距离计算节点相似距离,最后由上述结果根据各自权重相加计算骨架树相似距离^[36-37]。MindFinder^[22]中,提出倒角匹配(chamfer matching)^[43]和定向的倒角匹配(oriented chamfer matching)方法^[44],但前者时间复杂度高,后者内存开销大。因此,该文将利用定向的倒角匹配生成的草图线条距离图转变为有 N 个方向通道的击中映射图(hit map),并验证待匹配图像中是否有某个边缘像素(edgel)与草图线条映射图中的点相似。这种方法称为可索引的定

向的倒角匹配(indexable oriented chamfer matching)^[22],节约了物理花销并具有局部形状不变性。

1.4 图像反馈

利用视觉信息检索,能够增加检索的有效性,却造成视觉信息和高层语义之间的鸿沟。为了架起高层语义和底层特征的桥梁,需要对草图检索结果进行语义提取。反馈是语义提取的一个重要方面,主要包括对查询结果进行反馈、通过机器学习方法度量用户的绘制意图等方面。早期的反馈多采用查询点移动策略^[45]或权值再计算。通过重新排列检索结果来改善检索效率,虽然这种反馈对系统而言无本质的提高,但是用户在搜索结果上的行为数据是分析用户心理的重要数据来源,如何基于这些数据提高排序质量,也是一个值得研究的问题。

后来的反馈系统引入机器学习方法。例如引入SVM学习方法^[11,46],一些检索系统采用经典的二值SVM或单类SVM分类器。然而二值SVM忽略正反例数量不对称的情况,而单类SVM则对反例信息不适用。因此应用这2种经典的SVM得到的检索效果都不理想。梁爽等^[46]在此基础上提出了有偏SVM,通过学习用户对草图的理解和主观评价,并实时捕捉用户的查询兴趣,使搜索的结果更加接近用户的意图。袁贞明等^[47]通过计算草图结构中贝叶斯网络拓扑结构的最大后验概率,根据用户输入的笔划信息对笔划进行动态最优分组,保证了笔划输入的连续性,提高了分组效率。裴继红等^[48]提出带反馈机制的闭环隐马尔可夫模型,采用带压缩率调整因子的特征压缩算法,在计算各个后验概率后进行满意度判断,以确定是否调整压缩因子。裴继红将这种方法应用于手绘图形的识别,通过实验证明带反馈机制的闭环隐马尔可夫比开环隐马尔可夫有更高的识别率。然而目前机器学习的方法并没有人工标注图像的方法发展的成熟,因此可以采用人工标注的方法辅助检索,缩小检索范围,例如Sketch2Photo^[52]首先用关键字查找,然后再提取草图特征进一步查找。总之,利用人工标注图像语义的方法极大地缩小了视觉信息和语义的鸿沟,仍然被广泛采用。

2 手绘草图检索系统现阶段的应用

现有基于手绘草图的检索技术的相关文章大多可归为于本文所总结系统框架的某个或某几个模块的研究范畴。通过对之前研究的分析,提出手绘草图的手绘草图检索系统框架,旨在建立领域无关的手绘草图检索系统,从而将草图检索系统应用于更广阔的领域。本文现将基于手绘草图的检索系统现

阶段的应用总结为五方面内容。

1) 只利用检索功能,检索出用户需要的图像。目前草图检索已经实现从 2-D 图像到 3-D 图像,从小数据库到百万数据库的检索。如 Mindfinder^[49]系统已经实现利用手绘草图从在线的百万数据图片中进行快速检索,并用三维检索结构去减少检索时间,如图 2。而 MagicBrush^[50]系统还考虑了颜色信息,实现利用带颜色的草图在百万数据库中的实时检索,如图 3 所示。离线情况也可以将搜索完成的图像利用系统进行合成,用于照片的制作^[51-52],如图 4 所示。还可以通过简单的图形表示人脸位置和人脸属性,例如指定是否戴帽子或指定肤色^[53],实现对面脸图像的检索,如图 5 所示。此外,Eitz 等^[54]已经实现利用手绘草图对三维模型的检索,如图 6。

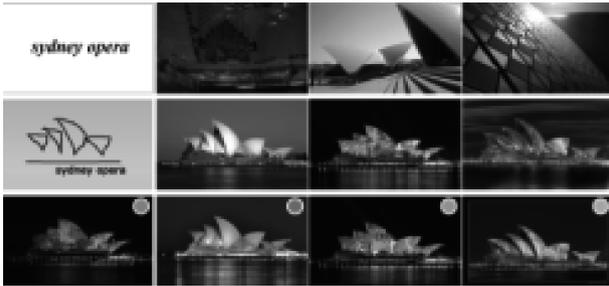


图 2 利用 Mindfinder 查找悉尼歌剧院

Fig.2 Find the Sydney Opera house with Mindfinder

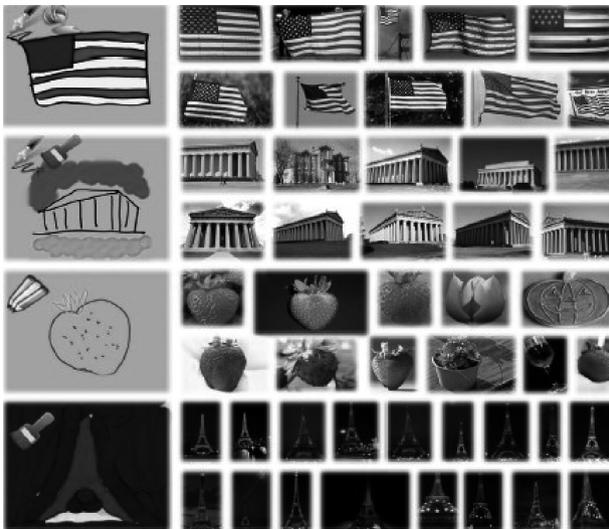


图 3 MagicBrush: 利用带颜色草图进行检索

Fig.3 MagicBrush: sketch retrieval by color sketches

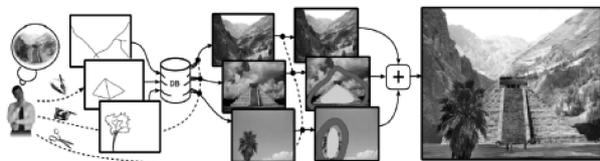


图 4 利用草图在数据库中查询并合成图片

Fig.4 Retrieve in the database and synthesis of a picture by sketches



图 5 基于人脸位置和属性的图像搜索

Fig.5 Image retrieval based on human face location and property

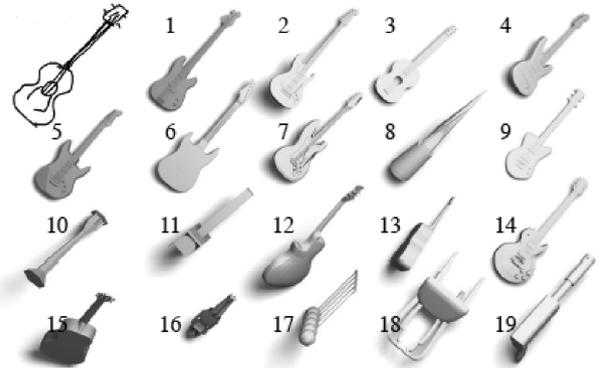


图 6 利用草图进行三维模型的检索

Fig.6 Three-dimensional model retrieval by sketches

2) 扩展检索功能,用草图检索系统管理图像数据。例如肖秦琨等^[55]等用手绘图作为查询条件运用 Zernike 矩函数提取图像特征,完成对商标的管理和检索,用于新商标的设计和检验。周良等^[56]提出利用图匹配方法对工程图纸检索,便于工程图纸的查找和管理。Tencer 等^[57]提出利用草图对文档中的插图进行检索,如图 7 所示,对于图文相混合的大量文档,更便于感兴趣区域的快速定位。



图 7 利用草图对文档中插图进行检索

Fig.7 Using sketches to retrieve the illustrations in a document

3) 利用草图模糊性和不明确性的特点开发特定应用。可以根据现场目击者和受害者的回忆,手绘犯罪嫌疑人头像,然后通过手绘草图检索系统和人脸识别技术,在犯罪嫌疑人的图像库中进行搜索从而缩小嫌疑人的范围,以便协助公安机关破案

和法院诉讼^[58]。

4) 用于开发面向儿童的认知训练系统。Sketch2Cartoon 系统^[59]提出新颖的剪贴画搜索引擎 ClipartFinder,能够实时检索互联网上一百万幅剪贴画的卡通图像,并将寻找到的矢量图进一步编辑从而构成自己的图画,如图 8 所示。2012 年上市的基于 iOS 平台的社交游戏“你画我猜”(Draw something),通过画线条图进行猜词,进一步拓展了面向儿童的应用,如图 9 所示。

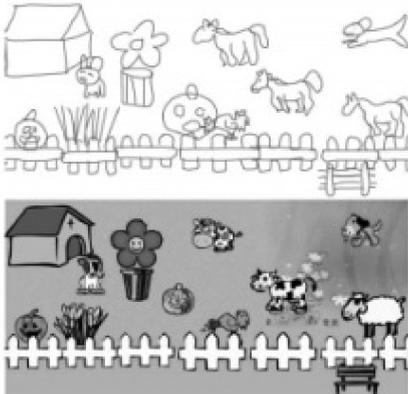


图 8 利用草图构建儿童画

Fig.8 The construction of children's drawing by sketches

5) 利用手绘草图进行三维建模。陈雪锦^[60]用二维设计草图生成具有真实感的三维几何模型,如建筑模型或树模型,帮助用户方便地捕捉构思,进行抽象的思考和形象的描述。如图 10 所示。刘凯等^[61]利用草图进行三维人脸的建模,可快速生成特定的人脸模型。如图 11 所示。



图 9 你画我猜游戏界面

Fig.9 The interface of drawing something



图 10 利用草图生成树木和房屋的真实感模型

Fig.10 Realistic models of trees and houses generated by sketches

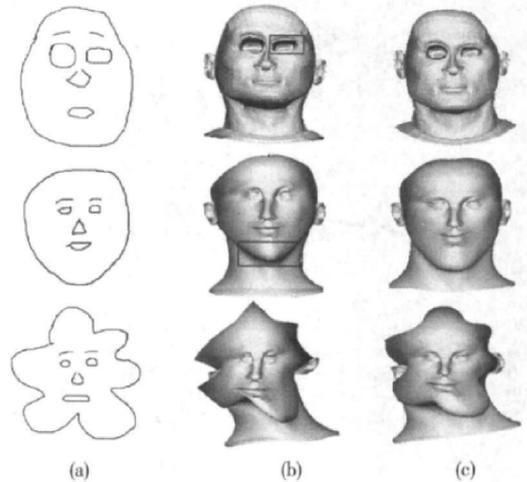


图 11 利用草图进行三维人脸重建

Fig.11 3-D face reconstruction by sketches

3 展望

本文通过对基于手绘草图的图像检索的系统框架进行梳理发现,以下 3 个方面可能成为未来手绘草图检索技术的研究趋势。

3.1 多通道的交互方式

笔纸交互是自然的交互方式,然而在手绘草图领域,最自然的交互方式是不妨碍甚至推进用户进行创作的过程,并非只采用鼠标模拟笔进行绘图。因此,在未来的手绘草图系统中,可以引入触摸交互、语音交互、手势交互与笔纸交互相融合,实现多通道的交互环境。在触摸交互方面,随着平板电脑的出现,可以利用多点触摸技术在触摸屏上精准地绘图。在语音交互方面,可以运用声音改变笔的颜色、绘图的场景等。例如,王丹力等^[62]设计了儿童交互式智能讲故事系统,儿童可以使用笔和语音的多通道交互技术,完成故事的创作。在手势交互方面,可以利用摄像头、数据手套进行空中涂鸦。也可以借鉴体感游戏的设计理念,利用 Kinect 开发基于手势识别的“手绘”系统。由于基于多通道交互的手绘草图系统,丰富了创作方法,减小了交互障碍,极大释放了用户创作的本能,必将成为普适计算时代人机交互的一个重要研究方向。

3.2 大数据背景下的草图检索

3.2.1 更合理的特征表示方法

利用手绘草图进行检索的难点在于草图本身的模糊性和不准确性,因此,手绘草图系统研究的关键在于对草图的理解与表征。本文将现有特征表示方法归纳为基于笔划描述的草图特征表示、基于形状特征的草图特征表示和基于组合图元的草图特征表示。但是对于构图复杂的草图,现有方法仍难以全

面合理地描述其特征。因此,找到更合理有效的草图特征描述方法仍需进一步研究。有些学者已经开始了新的尝试,如 Wang 等^[63]等人认为目前大部分研究只关注对图像特征的描述,而忽略了人的视觉注意机制,从而提出利用角径向分区的分割方法,根据用户手绘细致程度的不同求得空间权值分配图,并将权值分配图与图像特征相结合,提取分割后各块的特征与权值组成多维特征向量进行相似度计算。通过实验表明,将权值图应用于图像特征对实验效果有明显改善。可见,更合理的特征描述不能只关注草图本身的构成,还要结合人的视觉机制、人对图形的认知,以及其他反应人的输入意图的主观因素,形成共同描述草图的特征表示方法。

3.2.2 基于大数据的索引结构

早期的草图检索系统均建立在小型数据库的基础上,采用线性的索引结构。而随着互联网数据的海量激增,信息检索技术需要满足人们在大数据背景下的检索需求。目前已经提出一些索引方法来加速检索。例如将用于文本检索的倒排索引运用到草图检索系统^[13,23,51]中来,满足基于较大数据库快速检索的要求。Lukas 等^[38]利用最近邻算法加速检索。MinderFinder^[22]提出 Edgel Index 的索引结构,配合上文提到的可索引的定向倒角匹配算法,实现了在 两百万幅图片上检索仅需几百毫秒的检索效率。MagicBrush^[50]提出类似于倒排索引的方法实现六百四十万幅图片上的实时检索。然而对于十亿级数据库 (billion-level), MinderFinder 需要几分钟的时间和 5TB 的内存空间。因此微软亚洲研究院的 Sun Xinghai 等^[64]提出了采用 KLSH,即采用中心点算法 (K-Medoids) 代替 K-means 算法的局部敏感哈希 (Locality sensitive hashing, LSH) 的检索结构,实现了二十亿数据库的实时检索。然而对于更大规模 (big data) 的数据库,如何实现快速检索的需求,仍在探索阶段。

3.2.3 利用草图语义改进草图检索

传统基于草图的图像检索方法,虽然已经不再限制笔划方向等绘画要求,但也需要所绘制的草图接近用户意图。而对于那些在几何形状上离用户意图很远的草图,传统基于草图的图像检索方法则不具有适应性。Eitz 等^[65]针对此种情况,提出语义草图检索的方法,采用将用户所绘制草图分类,然后将类别作为关键字进行检索。然后在此基础上可以选择继续利用形状特征等传统图像检索方法进行图像检索。Sun 等^[66]也提出将草图识别用于加强草图检索。可见,将草图的语义用于草图检索可以放宽用户绘图的相似性要求、增加检索的准确性,也为未

来基于草图的图像检索系统的开发开拓了思路。

3.2.4 统一的搜索评价标准

随着手绘草图检索系统的增多,如何判定检索系统的优劣成为人们关心的问题。对检索系统的评价分为检索性能评价和检索效果评价,检索性能包括检索速度、所需内存空间大小及是否满足平移、旋转、缩放不变性等。好的检索系统占用最低的内存空间,实现海量数据库的实时响应。检索效果则为检索出的图片与所画草图的相似程度,一般通过绘制“查准率-查全率曲线”进行评价,如^[33-35,67]等。然而查准率是实验者对检索结果的主观判断,因此 Mathias Eitz^[13]提出把检索算法对图像的排序与人的主观感受进行比较的评价标准,用二者之间的相关性作为检索效果的评价标准。然而检索系统的优劣,应该从特征描述的有效性、系统的稳定性和可移植性、检索性能和效果以及人的主观需求综合考虑,因此如何形成统一的针对基于手绘草图检索系统的评价标准仍是一个值得深入研究的问题。

参考文献:

- [1] RUI Y, THOMAS S, CHANG S. Image retrieval: past, present and future [J]. *Journal of Visual Communication and Image Representation*, 1999, 10(1): 39-62.
- [2] EAKINS J. Automatic image content retrieval—Are we getting anywhere [C]//*Proceedings of 3rd International Conference on Electronic Library and Visual Information Research*. Newcastle, United Kingdom, 1996: 123-135.
- [3] VELTKAMP R C, TANASE M. Content-based image retrieval systems; a survey. UU-CS-2000-34 [R]. Dept of Computing Science, Utrecht University, 2000.
- [4] NIBLACK C, BARBER R, EQUITZ W, et al. The QBIC project: querying images by content using color, texture, and shape [C]//*Proceedings of SPIE Electronic Imaging: Science and Technology*. San Jose, Canada, 1993: 173-187.
- [5] PENTLAND A, PICARD R W, SCLAROFF S. Photobook: content-based manipulation of image databases [J]. *International Journal of Computer Vision*, 1996, 18(3): 233-254.
- [6] 孙正兴,冯桂焕,周若鸿.基于草图的人机交互技术研究进展 [J]. *计算机辅助设计与图形学学报*, 2005, 9(17): 1889-1899.
SUN Zhengxing, FENG Guihuan, ZHOU Ruohong. Techniques for sketch-based user interface: review and research [J]. *Journal of Computer Aided Design and Computer Graphics*, 2005, 9(17): 1889-1899.
- [7] 尹婷婷,王丹力,戴国忠.基于多笔划的草图绘制工具 [C]//*第 5 届全国普适计算学术会议论文集*.北京,中国, 2009: 308-312.
YIN Tingting, WANG Danli, DAI Guozhong. A sketch

- drawing tool based on multiple strokes recognition [C]//Proceedings of the 5th National Conference on Pervasive Computing. Beijing, China, 2009: 308-312.
- [8] SEZGIN T, STAHOVICH T, DAVIS R. Sketch based interfaces: early processing for sketch understanding [C]//Proceedings of 2001 Workshop Perceptive User interface. New York, USA, 2001: 1-8.
- [9] 李雪峰, 周良. 基于自适应顶点探测的手绘草图识别 [C]//2008年中国高校通信类院系学术研讨会论文集. 南京, 中国, 2009: 315-320.
- LI Xuefeng, ZHOU Liang. Sketch recognition based on adaptive vertex detection [C]//2008 Proceedings of the Symposium on Communication Colleges and Universities in China. Nanjing, China, 2009: 315-320.
- [10] 汪文睿, 周良. 基于层次的草图检索框架 [J]. 中国制造信息化, 2006, 35(19): 78-81.
- WANG Wenrui, ZHOU Liang. The framework for sketch searching based on hierarchy [J]. Manufacture Information Engineering of China, 2006, 35(19): 78-81.
- [11] 孙正兴, 徐晓刚, 孙建勇, 等. 支持方案设计的手绘图形输入工具 [J]. 计算机辅助设计与图形学学报, 2003, 15(9): 1145-1152.
- SUN Zhengxing, XU Xiaogang, SUN Jianyong, et al. Sketch-based graphic input tool for conceptual design [J]. Journal of Computer Aided Design and Computer Graphics, 2003, 15(9): 1145-1152.
- [12] 龚健, 费广正, 石民勇, 等. 基于手绘草图轮廓检索的简笔绘图系统 [C]//2008 系统仿真技术及其应用学术会议论文集. 北京, 中国, 2008: 761-764.
- GONG Jian, FEI Guangzheng, SHI Minyong, et al. A painting system based on freehand sketch contour retrieval [C]//Proceedings of 2008 Symposium on System Simulation Technology and Its Applications. Beijing, China, 2008: 761-764.
- [13] EITZ M, HILDEBRAND K, BOUBEKEUR T, et al. Sketch-based image retrieval: benchmark and bag-of-features descriptors [J]. IEEE Transactions on Visualization and Computer Graphics, 2011, 17(11): 1624-1634.
- [14] MARTIN D R, FOWLKES C C, MALIK J. Learning to detect natural image boundaries using local brightness, color, and texture [J]. IEEE Transactions on Pattern Analysis and Machine, 2004, 26(5): 530-549.
- [15] RUBINE D. Specifying gestures by example [C]//Proceedings of the 18th Annual Conference on Computer Graphics and Interactive Techniques. New York, USA, 1991: 329-337.
- [16] LONG A C, JAMES J, LANDAY A, et al. Visual similarity of pen gestures [C]//Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York, USA, 2001: 360-367.
- [17] HSE H, NEWTON A. Sketched symbol recognition using Zernike moments [C]//Proceedings of the 17th International Conference on Pattern Recognition. [S.l.], 2004: 367-370.
- [18] 谢强, 冯桂焕, 孙正兴. 基于上下文的在线草图识别方法 [J]. 计算机科学, 2007, 34(3): 216-220.
- XIE Qiang, FENG Guihuan, SUN Zhengxing. Online sketch recognition using contexts [J]. Computer Science, 2007, 34(3): 216-220.
- [19] TANDLER P, PRANTE T. Using incremental gesture recognition to provide immediate feedback while drawing pen gestures [C]//Proceedings of ACM Symposium on User Interface Software and Technology. New York, USA, 2001: 18-25.
- [20] ARVO J, NOVINS K. Fluid sketches: continuous recognition and morphing of simple hand-drawn shapes [C]//Proceedings of the ACM Symposium on User Interface Software and Technology. New York, USA, 2000: 73-80.
- [21] 李俊峰, 张习文, 敖翔, 等. 具有实时反馈的草图交互识别方法 [J]. 计算机辅助设计与图形学学报, 2005, 17(11): 2453-2458.
- LI Junfeng, ZHANG Xiwen, AO Xiang, et al. An interactive sketch recognition approach with real time feedback [J]. Journal of Computer Aided Design and Computer Graphics, 2005, 17(11): 2453-2458.
- [22] CAO Y, WANG C, ZHANG L, et al. Edgel index for large-scale sketch-based image search [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. [S.l.], 2011: 761-768.
- [23] ZHOU R, CHEN L, ZHANG L. Guess what you draw: interactive contour-based image retrieval on a million-scale database [C]//Proceedings of the 20th International Conference on Multimedia. New York, USA, 2012: 1343-1344.
- [24] REVANKAR S, YEGNANARAYANA B. Machine recognition and correction of freehand geometric line sketches [C]//Proceedings of IEEE International Conference on Systems, Man, and Cybernetics. New York, USA, 1991: 87-92.
- [25] SEZGIN T. Feature point detection and curve approximation for early processing of free-hand sketches [D]. Cambridge, USA: Massachusetts Institute of Technology, 2001: 3-74.
- [26] YU B, CAIN S. A domain-independent system for sketch recognition [C]//Proceedings of the 1st International Conference on Computer Graphics and Interactive Techniques. New York, USA, 2003: 141-146.
- [27] PAULSON B, HAMMOND T. Accurate primitive sketch recognition and beautification [C]//Proceedings of Intelligent User Interfaces. New York, USA, 2008: 1-10.
- [28] 张莉莎, 孙正兴, 周若鸿, 等. 一种基于图匹配的复杂草图识别方法 [J]. 计算机科学, 2004, 31(4): 147-150.

- ZHANG Lisha, SUN Zhengxing, ZHOU Ruohong, et al. A method of graph-based composite sketchy graphics recognition[J]. *Computer Science*, 2004, 31(4):147-150.
- [29] GROSS M. The electronic cocktail napkin: a computational environment for working with design diagram[J]. *Design Studies*, 1996, 17(1): 53-69.
- [30] CHEN C, XIE S. Freehand drawing system using a fuzzy logic concept[J]. *Computer-Aided Design*, 1996, 28(2): 77-89.
- [31] SCIASIO E, MONGIELLO M. Query by sketch and relevance feedback for content-based image retrieval over the web[J]. *Journal of Visual Languages and Computing*, 1999, 10(6): 565-584.
- [32] CHEE S, PARK K, PARK S. Efficient use of MPEG-7 edge histogram descriptor[J]. *ETRI Journal*, 2002, 24(1): 35-42.
- [33] 李曼舞, 孙正兴, 顾庆东. 基于轮廓的手绘草图检索研究[J]. *计算机应用*, 2005, 25(2): 344-347.
LI Manwu, SUN Zhengxing, GU Qingdong. Study on contour-based freehand sketch retrieval[J]. *Computer Applications*, 2005, 25(2): 344-347.
- [34] 高竹红, 汤进, 罗斌. 基于结构图的手绘草图检索[J]. *计算机技术与发展*, 2008, 18(3): 33-35.
GAO Zhuhong, TANG Jin, LUO Bin. Sketch retrieval based on structural graph[J]. *Computer Technology and Development*, 2008, 18(3): 33-35.
- [35] 吴明珠. 利用成对几何直方图改进基于轮廓的图像检索[J]. *电脑知识与技术*, 2007, 16(1): 1140-1147.
WU Mingzhu. Improvement of contour-based image retrieval using pair wise geometric histogram [J]. *Computer Knowledge and Technology*, 2007, 16(1): 1140-1147.
- [36] 钱晶. 按例和按绘的图像检索研究[D]. 合肥: 中国科学技术大学, 2009: 69-100.
QIAN Jing. Research on example and sketch based image retrieval[D]. Hefei: University of Science and Technology of China, 2009: 69-100.
- [37] 钱晶, 汪增福. 一种基于规范化形状的手绘草图检索算法[J]. *中国科学技术大学学报*, 2010, 10(40): 1043-1053.
QIAN Jing, WANG Zengfu. A method for freehand sketch retrieval based on affine adaptive skeleton[J]. *Journal of University of Science and Technology of China*, 2010, 10(40): 1043-1053.
- [38] LUKAS T. A new framework for online sketch-based image retrieval in web environment[C]//Proceedings of the 11th International Conference on Information Sciences, Signal Processing and their Applications. Montreal, Canada, 2012: 1430-1431.
- [39] KENDALL M. A new measure of rank correlation[J]. *Biometrika*, 1938, 6(30): 81-93.
- [40] KENDALL M, GIBBONS J. Rank correlation methods [M]. London: Griffin, 1948: 1-160.
- [41] 李彬, 梁爽, 孙正兴. 基于空间关系的手绘草图检索[J]. *计算机科学*, 2005, 32(12): 227-231.
LI Bin, LIANG Shuang, SUN Zhengxing. Sketch retrieval based on topological relations [J]. *Computer Science*, 2005, 32(12): 227-231.
- [42] SOUSA P, FONSECA M. Sketch-based retrieval of drawings using topological proximity [C]//Proceedings of the 14th International Conference on Distributed Multimedia Systems. Lisboa, Portugal, 2008: 276-281.
- [43] BORGEFORS G. Hierarchical chamfer matching: a parametric edge matching algorithm[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1988, 10(6): 849-865.
- [44] STENGER B, THAYANANTHAN A, TORR P, et al. Model-based hand tracking using a hierarchical Bayesian filter[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, 28(9): 1372-1384.
- [45] ROCCHIO J. Relevance feedback in information retrieval [M]. [S.l.]: Prentice Hall, 1971: 313-323.
- [46] 梁爽, 孙正兴. 面向草图检索的相关反馈方法[J]. *计算机辅助设计与图形学学报*, 2006, 18(11): 1753-1757.
LIANG Shuang, SUN Zhengxing. BSVM-based relevance feedback for sketch retrieval[J]. *Journal of Computer Aided Design and Computer Graphics*, 2006, 18(11): 1753-1757.
- [47] 袁贞明, 金贵朝, 张佳. 基于贝叶斯网络的在线草图识别算法[J]. *计算机工程*, 2010, 36(5): 32-34.
YUAN Zhenming, JIN Guichao, ZHANG Jia. Online sketch recognition algorithm based on Bayesian network [J]. *Computer Engineering*, 2010, 36(5): 32-34.
- [48] 裴继红, 李翠芸, 龚忻. 一种新的隐马尔可夫模型及其在手绘图形识别中的应用[J]. *计算机学报*, 2005, 28(10): 1745-1752.
PEI Jihong, LI Cuiyun, GONG Xin. A novel hidden Markov model and its application to recognize hand-drawn graphic symbol[J]. *Chinese Journal of Computers*, 2005, 28(10): 1745-1752.
- [49] CAO Y, WANG H, WANG C, et al. MindFinder: interactive sketch-based image search on millions of images[C]//Proceedings of the 18th International Conference on Multimedia. Santiago, USA, 2010: 1605-1608.
- [50] SUN X, WANG C, SUD A, et al. MagicBrush: image search by color sketch[C]//Proceedings of the 21st ACM International Conference on Multimedia. New York, USA, 2013: 475-476.
- [51] EITZ M, HILDEBRAND K, BOUBEKEUR T, et al. PhotoSketch: a sketch based image query and compositing system[C]//ACM SIGGRAPH Talk Program. New Orleans, USA, 2009: 60.
- [52] CHEN T, CHENG M, TAN P, et al. Sketch2Photo: inter-

- net image montage [J]. ACM Transactions on Graphics, 2009, 28(5): 789-790.
- [53] LEI Y, CHEN Y, CHEN B, et al. Photo search by face positions and facial attributes on touch devices [C]//Proceedings of the 19th International Conference on Multimedia. New York, USA, 2011: 651-654.
- [54] EITZ M, RICHTER R, BOUBEKEUR T, et al. Sketch-based shape retrieval [J]. ACM Transactions on Graphics, 2012, 31(4): 31.
- [55] 肖秦琨,张楠,高嵩. 一种基于手绘图的商标检索算法 [J]. 西安工业大学学报, 2012, 32(5): 373-378.
XIAO Qinkun, ZHANG Nan, GAO Song. A trademark retrieval algorithm based on sketch [J]. Journal of Xi'an Technological University, 2012, 32(5): 373-378.
- [56] 周良, 谢强, 丁秋林. 基于图匹配的工程图纸检索 [J]. 南京航空航天大学学报, 2008, 40(3): 354-359.
ZHOU Liang, XIE Qiang, DING Qiulin. Engineering drawing retrieval based on graph matching [J]. Journal of Nanjing University of Aeronautics and Astronautics, 2008, 40(3): 354-359.
- [57] TENCER L, RENAKOVA M, CHERIET M. Sketch-based retrieval of document illustrations and regions of interest [C]//Proceedings of 12th International Conference on Document Analysis and Recognition. Washington, DC, USA, 2013: 728-732.
- [58] 马月娜, 马力克. 基于边缘检测的人脸模拟画像检索 [J]. 计算机应用与软件, 2009, 26(8): 238-240.
MA Yuena, MA Like. Simulated face portrait retrieval based on edge detection [J]. Computer Applications and Software, 2009, 26(8): 238-240.
- [59] WANG C, ZHANG J, YANG B, et al. Sketch 2 Cartoon: composing cartoon images by sketching [C]//Proceedings of the 19th ACM International Conference on Multimedia. New York, USA, 2011: 789-790.
- [60] 陈雪锦. 手绘现实: 手绘草图的真实感解释 [D]. 合肥: 中国科学技术大学, 2008: 88-90.
CHEN Xuejin. Sketching reality: realistic interpretation from freehand sketching [D]. Hefei: University of Science and Technology of China, 2008: 88-90.
- [61] 刘凯, 孙正兴, 张尧焯. 采用草绘轮廓的3维人脸建模方法 [J]. 中国图像图形学报, 2011, 16(6): 1102-1110.
- LIU Kai, SUN Zhengxing, ZHANG Yaoye, et al. Method of 3D face modeling from freehand portrait sketch [J]. Journal of Image and Graphics, 2011, 16(6): 1102-1110.
- [62] 王丹力, 詹志征, 戴国忠. 儿童交互式智能讲故事系统 [J]. 计算机辅助设计与图形学学报, 2011, 7(23): 1186-1193.
- WANG Danli, ZHAN Zhizheng, DAI Guozhong. Interactive and intelligent storytelling system for children [J]. Journal of Computer Aided Design and Computer Graphics, 2011, 7(23): 1186-1193.
- [63] WANG Y, YU M. Query by sketch: an asymmetric sketch-vs-image retrieval system [C]//Proceedings of the 4th International Congress on Image and Signal Processing. Shanghai, China, 2011: 1368-1372.
- [64] SUN X, WANG C, XU C, et al. Indexing billions of images for sketch-based retrieval [C]//Proceedings of the 21st ACM International Conference on Multimedia. New York, USA, 2013: 233-242.
- [65] EITZ M, HAYS J, ALEXA M. How do humans sketch object [J]. ACM Transaction on Graphics, 2012, 7(31): 1-10.
- [66] SUN Z, WANG C, ZHANG L, et al. Query-adaptive shape topic mining for hand-drawn sketch recognition [C]//Proceedings of the 20th ACM International Conference on Multimedia. New York, USA, 2012: 519-528.
- [67] CARLOS A F, ARAJUJO A A, CRUCIANU M, et al. Sketch-finder: efficient and effective sketch-based retrieval for large image collections [C]//Proceedings of the 2013 XXVI Conference on Graphics, Patterns and Images. Arequipa, Peru, 2013: 234-241.

作者简介:



辛雨璇,女,1988年生,硕士研究生,主要研究方向为多媒体技术与人机交互技术。



闫子飞,女,1980年生,讲师,博士。主要研究方向为图像处理、模式识别、非物质文化遗产的数字化保护,基于手绘草图的认知界面设计等。主持和参与多项国家自然科学基金项目;发表学术论文10余篇,其中,被SCI检索1篇,EI检索9篇。

术论文10余篇,其中,被SCI检索1篇,EI检索9篇。