

基于二维手绘草图的三维形状检索

谢晓华

(中国科学院深圳先进技术研究院 深圳 518055)

摘 要 随着互联网上三维模型数量的迅猛增多,需要发展便捷、可靠的基于内容的三维形状(模型)检索引擎。手绘草图具有以视觉形象表达概念的能力,符合人类传递信息的习惯,因而成为三维形状检索的重要交互手段。然而,直接表达人类意识的二维手绘草图与标准的三维形状表达之间常常存在较大的语义鸿沟,这给基于手绘草图的三维形状检索带来很多技术难题。文章对基于手绘草图的三维形状检索相关背景以及技术进展进行了介绍,重点介绍了用上下文信息填补语义鸿沟的最新方法。

关键词 草图;形状检索;上下文;轮廓;搜索引擎;特征提取;计算机图形学;计算机视觉
中图分类号 TP 391 **文献标志码** A

Three Dimensional Shape Retrieval Based on Two Dimensional Sketches

XIE Xiaohua

(Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China)

Abstract With the rapid increase of 3D model on the internet, it is necessary to develop a convenient and reliable content-based 3D shape (model) search engine. Sketch has the ability to express concept as a visual icon, which is in line with the human habit of passing information, and thus becomes an important interactive means in 3D shape retrieval. However, there is often a larger semantic gap between the standard 3D shape representation and the 2D sketch which expresses human consciousness, which brings a lot of technical challenges to the sketch-based shape retrieval. In this paper, the relevant background and technological advances on the sketch-based 3D shape retrieval were introduced, and the latest methods that use context information to fill the semantic gap were particularly focused on.

Keywords sketch; shape retrieval; context; contour; search engines; feature extraction; computer graphics; computer vision

1 引 言

随着三维建模相关技术的进步尤其是三维扫描设备的成熟及大众化推广,互联网空间已经积累了大量可分享的三维模型,且模型数量呈持续剧增之势。譬如,用户可以直接从谷歌三维仓库

(3D Warehouse)下载海量三维模型;著名的 3D 模型交易平台 TurboSquid 目前已拥有超过 30 万个模型。为充分利用已有的三维模型,需要发展便捷、高效、可靠的三维形状(模型)检索引擎。三维形状检索即通过特定的交互操作从数据库中查找出符合用户意向的三维形状。其中,检索的对象可以是整个模型(如一张椅子),也可以是模

收稿日期:2014-09-11 修回日期:2014-12-31

基金项目:国家自然科学基金项目(61202223)

作者简介:谢晓华(通讯作者),博士,副研究员,研究方向为计算机视觉、图像处理与图像识别, E-mail:sysuxiexh@gmail.com。

型的一部分(如椅子的把手)。

基于交互方式(用户输入的内容),常见的三维形状检索技术主要有以下几类:(1)基于文本信息的检索,即根据用户输入的文本与模型对应的文本标签之间的匹配度进行检索,如谷歌公司的 3D Warehouse 所用的搜索引擎;(2)基于特定命令符号的检索^[1],即用定义好的笔画或笔画组合来表示一些基本形状,譬如用“ ”表示一个三维球体;(3)基于三维模型的检索^[2],即用户输入一个三维模型,系统查找与之相似的模型, Bronstein 等^[3]对此类方法进行了很好的总结;(4)基于照片的检索^[4];(5)基于草图的检索,即基于二维草图(手绘或线绘)检索模型,其中以手绘草图为主。

基于文本信息以及特定命令符号的方法都要提前对每个三维模型进行手工标注。人工标注需要消耗大量的人力和时间,且人工标注往往具有主观性和片面性,导致标记标准的不统一。此外,基于特定命令符号的方法只适用于对少量基本图元的搜索,因为要求用户记住大量的命令符号并不现实。

为了达到更加便捷可靠的检索,研究人员开始关注基于内容的形状检索。“基于内容”这一概念之前被提出用于图像检索^[5]。基于内容的含义就是试图利用反映对象自身(如图像、视频、三维模型等)视觉特征的内容自动建立特征索引并通过对视觉特征的相似性匹配来查找用户所要搜索的目标,这更贴近于人们在现实生活中靠直

觉印象使用信息的方式^[1]。上述(3)、(4)和(5)三种形状检索方式都属于这个范畴。

基于三维模型或照片的方法要求用户手头拥有与意向目标类似的模型或者相应照片,所以该方法的应用范围相对比较有限。相比之下,基于草图的模型检索交互更加便捷。基于草图的人机交互方式能模拟传统的“纸-笔”环境,符合人类记录信息的习惯。草图以视觉形象表达和传递概念,且具有表达模糊概念的能力,能非常直接地记录用户的意图,故而基于草图能极大提高人机交互效率。尤其随着触屏设备的普及,基于手绘草图的检索技术得到极大关注。微软亚洲研究院在 2010 年推出的图像搜索引擎 MindFinder^[6]就支持用户通过输入草图在大规模数据库上进行图像实时检索。

2000 年,德国国家资讯技术研究中心(GMD)的 Löffler^[7]发布了一个用“草图+文本”作为输入的三维模型搜索系统。之后,多个完全支持手绘草图的三维形状搜索引擎和方法^[1,7-17]相继被提出,且越来越受到研究人员的关注。

作为一种直接记录人类意识的符号,手绘草图具有风格变化、抽象、不精确、不稳定等特征^[18]。柏林工业大学 Eitz 等^[9]曾经组织过一次大规模的在线调查。调查结果表明,人们针对同类目标绘制的草图往往具有多种相差甚远的表现形式(如图 1),实际上同一个人在不同时间手绘同个物体所产生的草图都有可能存在较大不同。这给计算机通过草图理解用户的意图增加了非常大的困

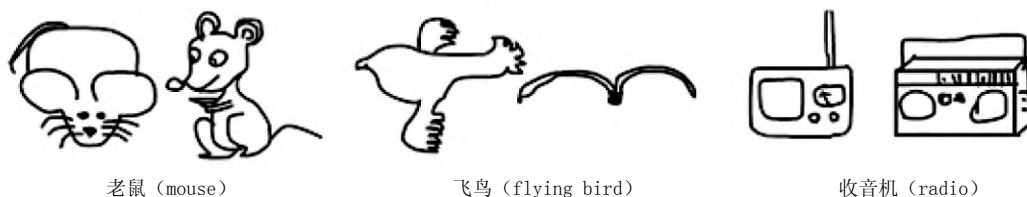


图 1 人们根据给定的关键字手绘得到的不同种代表性草图^[9]

Fig. 1 A sample of the representative sketches by human corresponding to different key words^[9]

难。导致这些现象的主要原因有：(1)用二维草图表达三维模型时，由于数据维数的减少必然存在模糊性；(2)由于不同用户绘画功底和世界观的不同，针对同个目标所绘制的草图存在很大区别；(3)在很多情况下用户对要绘制的目标的几何形式尚不明确，导致所绘制的草图具有随机性。

图 2 展示的是一盏台灯模型的三维形状、三维形状在特定视角下的二维渲染轮廓图以及用户的手绘草图。可见，三者样式之间存在巨大的差异。换言之，表现人类意识的手绘草图与符合计算机规范语言的形状表达之间存在明显的语义鸿沟。如何填补这个人机信息表达鸿沟是实现基于草图的形状检索系统的关键。

最近几年，研究人员在基于手绘草图的形状检索方面进行了大量研究。相关研究主要面向两



图 2 3D 模型(左)、3D 模型渲染得到的 2D 轮廓线(中)及用户的手绘草图(右)

Fig. 2 A sample of 3D model (left), the rendered 2D contour (middle), and the user sketches (right)

个目标：一是如何精准地理解草图所表达的信息并与三维形状建立有效的关联；二是在用户意图模糊的情况下如何给出合理建议从而启发用户。本文对上述相关技术进展进行介绍，其中重点介绍了利用上下文信息填补语义鸿沟的最新方法。

2 基于手绘草图的三维形状检索框架

二维手绘草图和三维形状存在数据维数上的不一致，两者之间并不能直接进行相似度计算。为解决这个问题，一类方法是根据二维草图推测出其所表达的三维形状，然后与数据库中的三维形状进行匹配；另外一类方法是将数据库中的三维形状渲染为多个视角下的二维形状图，然后用二维形状图与草图进行匹配。从可用性上考虑，基本都采用第二类方法，因为第一类方法中直接从二维草图推测三维形状本身就具有非常大的挑战性。

图 3 给出了基于草图的三维形状检索的一般框架。对数据库中的三维模型集，要事先对它们进行坐标归一化等预处理。对每个模型，计算得到若干采样视角下的二维表达(也称二维形状图像)，对每张二维形状图像进行特征提取，最终建立特征与原始三维形状之间的索引。

对用户输入的手绘草图，系统对草图进行滤波等预处理，然后提取图像特征。最后，根据图

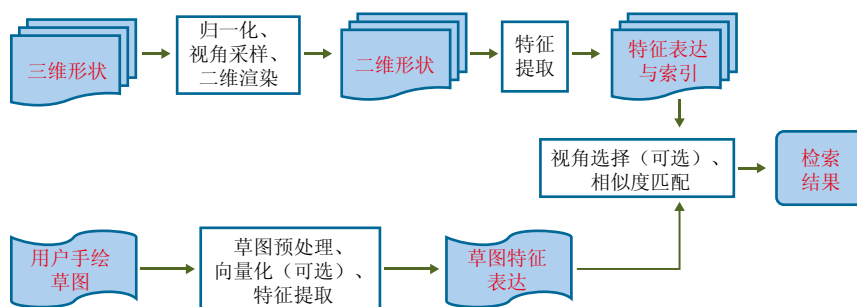


图 3 基于手绘草图的三维模型检索框架

Fig. 3 The flow chart of retrieving 3D model based on 2D sketches

像特征之间的匹配程度以及索引,把检索结果返回给用户。

值得指出的是,在图 3 中,不同操作之间并非孤立而是存在紧密关联,甚至相邻步骤有时会融为一体、难分你我。特别地,不同的特征提取算法要求不同的数据预处理方法,譬如用到拓扑描绘子的时候要对草图进行逼近式多边形向量化操作^[19,20]。下面将对图 3 中的不同步骤进行详细介绍。

3 数据预处理与特征提取

三维形状检索在技术实现上与自然图像检索有非常大的区别。三维形状不像自然图像那样具有丰富的反射、阴影、纹理和颜色信息,可用信息相对比较单一,只有几何特征,但这也使得三维形状检索不受光照变化、遮挡等外在因素的影响。此外,形状检索对数据的预处理以及特征表达都有特殊的要求和方法。

3.1 草图预处理

一个丰富的草图获取系统不仅记录用户手绘的最终草图,还记录用户绘制过程中笔的动态轨迹、笔画样式和用户的用力情况。采用更多信息会更有助于理解用户的意图,但也要求更复杂的处理与理解系统。现有的三维形状检索系统基本仅采用最终的静态草图作为用户的输入。

用户输入的草图经常具有随意性,譬如本该相交的笔划并未相交,本该闭合的曲线没有闭合,在笔画的端点处出现交叉等。这些导致手绘草图与形状渲染得到的二维形状表达风格之间存在较大不同。此外,不同用户绘制风格也存在很大不同。所以需要对草图进行必要的预处理。常采用的预处理包括对笔画进行重采样、滤波、直线和曲线拟合等,以达到消除冗余点和聚点、补合断点、平滑笔画等目的。有时候甚至使用基本图元(包括圆、椭圆、多边形、直线、曲线、圆

弧等)对草图进行逼近重表达。

柏林工业大学 Eitz 等^[9]的研究表明,人们在手绘草图的时候,习惯于采用从粗到细的绘画顺序,即先用长笔画绘制目标的主体轮廓,然后再添加细节。所以不少人也提出要对草图的主要轮廓和细节进行分开预处理。在 Saavedra 等^[21]提出的系统中就只提取草图的外轮廓用于识别。

3.2 形状预处理与二维表达

为了解决形状表达与草图之间的维数不一致问题,需要获取形状的二维表达,在这之前要对三维模型进行坐标系归一化^[22]、各向同性^[23]等预处理。获取三维形状的二维表达常用的方法是在特定视角下对形状进行线渲染,得到二维形状图像。常见的二维渲染有侧影、轮廓线、边界聚合、显著脊^[24]以及显著特征线^[25]。其中,DeCarlo 等^[26]提出的启发式轮廓线得到广泛的应用(见图 4)。除上述预处理外,若有必要,还要对渲染得到的二维形状图像进行向量化^[27]、加粗等操作。通过二维线渲染,三维模型的特征提取问题就被转化为二值图像特征提取问题。

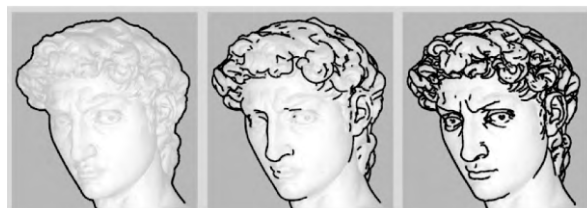


图 4 侧影(左)、轮廓线(中)及轮廓线+启发式轮廓线(右)

Fig. 4 A sample of the rendered silhouette (left), occluding contours (center), and the integration of occluding contours with suggestive contours (right)

在对形状进行渲染的时候需要指定视角,对每个视角渲染得到一张二维形状图,因而需要对三维空间进行视角采样。一般上,采样的视角越多,就能获得对三维形状越丰富的信息表达,但是过于稠密的采样会导致信息表达出现冗余,使得后续的计算量增大(因为用户输入的草图与每一张二维形状图都要进行相似度计算)。

普林斯顿大学 Funkhouser 等^[1]指出人们手绘特定目标所采用的视角集合是非常一致的,从而提出用三维目标在 13 个视角下的轮廓投影图来表示该目标。Shin 等^[12]、Daras 等^[10]和 Fonseca 等^[19,20]的系统则分别采用了 16、32 和 20 个视角采样。Eitz 等^[8]用支持向量机模型学习选定 42 个视角。Saavedra 等^[21]的方法仅用到侧影,由于侧影随视角改变所发生的变化相对比较缓慢,所以只选取了 6 个视角。Li 等^[17]则提出根据每个模型的视觉复杂性度量,用聚类的方法自适应选定若干个代表性的视角。Xie 等^[28]的系统用到 169 个视角采样,但是他们的交互界面能够自动获取用户手绘目标时所选定的视角,所以采用大量视角采样并不会导致计算量的增大(因为用户草图只需要与选定视角下的形状图进行相似度计算即可)。

3.3 形状特征提取与匹配

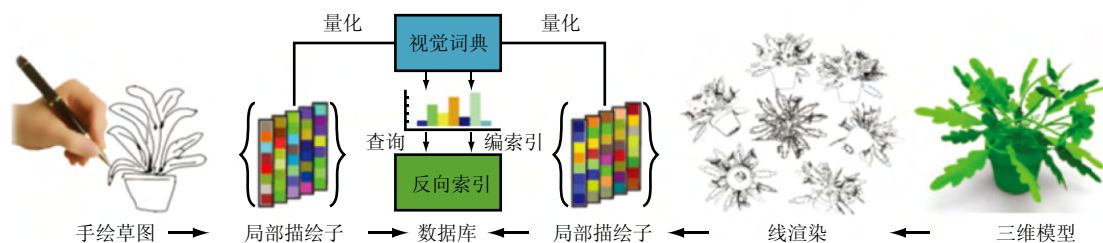
形状图像(为方便表达,本节把手绘草图和三维形状渲染得到的二维形状图像都称为形状图像)并不像自然图像那样具有丰富的颜色、纹理等特征,只有少量的几何或结构信息。如何从少量的几何线中得到目标形状的有效特征表达并进行可靠匹配,是计算机视觉以及计算机图形学领域一个传统的研究课题。

在形状特征表达方面,研究人员已经提出了大量方法^[8,11,14,15,29-36],这些方法主要目的是构造表达能力强、区分度明显、且基本满足平移不变性、旋转不变性以及缩放不变性的特征描绘子。简单而言,已有的形状描绘子基本可以分为两类:(1)具有明显几何语义描述的算子,如边界长度、曲率、曲率尺度空间^[29]、形状树、轮廓灵活性、外接矩形、区域面积、区域纵横比、主方向、紧凑度、欧拉数、形状数、骨架特征、直线段描述、样条拟合曲线描述、内角直方图描述、高斯参数曲线描述、曲率尺度空间和分层拓扑^[19,20]等;(2)基于统计的

视觉特征描绘子,如尺度不变特征变换(Scale Invariant Feature Transform, SIFT)^[37,38]、极图谱、梯度方向直方图(Histogram of Oriented Gradient, HoG)^[30,31,39,40]、边缘直方图^[41]、球谐函数^[1]、弥散张量^[11]、矩特征^[42,43]、傅里叶描述子^[35,36]、基于 Gabor 滤波的局部线特征(Gabor Local Line-based Feature, GALIF)^[8]、形状上下文^[34]等。其中,形状上下文是极为有效的形状特征描绘子,已经衍生出包括内距离形状上下文^[44]在内的多种形状描绘子。

在第(2)类特征描绘子中,有相当一部分是局部特征描绘子,如 SIFT、HoG、GALIF 等。基于局部视觉特征描绘子的方法主要采用“词袋模型”^[45]来对图像进行编码。词袋模型即采用 K-means 聚类等学习算法从大量图像块中学习得到一组“视觉单词”构成“视觉字典”(或称“视觉词袋”),对每张图像统计其中的视觉单词出现的频率,最后用由这些单词频率组成的直方图来表示这张图像。基于词袋模型编码,图像匹配问题被转为直方图匹配问题,且能够采用反向索引^[46]的方法对大型数据库实现快速查询。原始的词袋模型忽略特征的空间位置信息,为弥补这点不足,Bronstein 等^[47]对传统的词袋模型进行改进,在原有模型框架上增加一个词典来记录视觉单词之间的相对空间关系,目前该方法已经被当成图像匹配基准方法。柏林工业大学 Eitz 等^[8]于 2012 年提出一个采用词袋模型实现基于手绘草图的形状检索方法(图 5),其中提出一个基于 Gabor 滤波的局部线特征,该方法近年获得了极大的关注。

Eitz 等^[39]曾经在 2011 年针对基于手绘草图的图像检索应用对相关特征描绘子进行比较并提出一个检索基准,他们特别指出应用原始的 HoG 特征于草图检索效果并不佳,从而提出一个简化版的 HoG 算子用于形状检索。美国德州大学 Li 等^[18,21,48]先后组织了“SHREC 2012”和

图5 Eitz 等^[8]提出的基于手绘草图的三维形状检索引擎Fig. 5 Overview of the 3D shape search engine presented by Eitz et al.^[8]

“SHREC 2013”两次评测，评测内容就是基于手绘和线绘草图的三维模型检索方法的性能。他们的评测结果表明，与傅里叶描绘子以及拓扑描绘子相比，形状上下文^[34]、GALIF^[8]、稠密SIFT^[38]和简化版的HoG^[9,39]能取得更好的形状检索结果；与基于全局特征的方法相比，基于局部描绘子的方法(结合词袋模型)更加容易实现实时检索；此外，多种特征融合也能提升模型检索的准确率，譬如Furuya等^[49]的方法结合运用改进后的GALIF和稠密SIFT特征，在所有参加评测的方法中取得最佳结果。

对不同种形状特征表达，也对应有不同的距离度量方法，Velkamp等^[50]对此有很好的总结。对两个形状图像采样得到的两个点集之间的匹配度，可以用瓶颈距离、Hausdorff距离、地动距离^[51]或者弹性匹配距离进行度量。对参数化为多边形或者折线之间的匹配，可以用旋转方程、反射距离进行度量。对参数化的曲线之间，可以用Fréchet距离进行匹配度量。针对向量化的轮廓线，可以采用基于采样的形状匹配算法^[52]。对基于视觉特征的形状描绘子方法，一般得到直方图、频谱等形式表达，这时特征之间的匹配度适用欧式距离、曼哈坦距离、明科夫基距离、切比雪夫距离、卡方距离、余弦相似度或直方图交等进行度量。对直方图特征，在应用距离度量之前，经常会采用词频-反文档频率^[53]方法对直方图进行加权归一化。Grauman等^[54]还指出可以通过度量学习的方法学习得到每个视觉单词对应的

最优权重。

值得注意的是，如3.2节所述，每个三维模型实际上拥有跟采样视角数目一样多的二维形状图像。在用户手绘草图所用视角未知的情况下，一般的做法是将用户草图与某个模型对应的所有形状图像逐个进行匹配，并选取其中最大的相似值作为草图与该模型的相似值。为了减少计算量同时降低检索错误，Li等^[55]提出在匹配时用视角上下文^[56]方法实现草图与三维模型的对齐，从而避免将草图与所有采样视角下的形状图像进行匹配。

4 上下文信息的应用

手绘草图本身具有信息表达的模糊性，加之有时候用户对自己绘画的目标的几何形式还不是非常明确，在这种情况下单纯对用户输入的草图本身进行分析理解仍然很难得到合理的形状检索结果。最近有研究人员开始关注运用上下文信息来辅助基于草图的形状检索，希望进一步提高形状检索的精准度和合理性。

上下文是一个广泛的概念，是对现实世界的高层语义表达，描述了相邻目标或者相邻状态之间的显性或隐性关联。上下文信息在计算机信息处理与理解上具有重要的作用。2002年在美国麻省理工学院智能信息检索研究中心召开的未来信息检索挑战国际会议上，上下文检索被认为是信息检索的长期挑战。上下文信息在计算机视

觉领域尤其在基于图像的目标检测^[57-60]、目标识别^[61,62]和显著性检测^[63]方面都已经获得成功的应用。麻省理工学院的 Torralba^[64]曾经提出一个名为“上下文挑战”的评测项目：在图像中，将要检测的目标(如一个人脸或者一辆小车)遮住，要求相关算法只根据周围内容(如人的身体或者道路)推断被遮住的目标为何物。实验结果表明，即使在如此挑战的条件下，相关算法对被遮物体的推测准确率还相当高。

上下文信息的利用最近也已经引起三维模型领域研究人员的关注。2010 年斯坦福大学 Fisher 等^[65]提出一个基于三维场景中物体位置关系的三维模型检索方法。他们首先从三维场景数据库中学习到不同物体之间的位置关系(譬如台灯在桌子上面、椅子在桌子的旁边)。当用户在某三维场景中输入一个边界盒，系统会根据边界盒和场景中已有模型的位置关系来推荐搜索结果(见图 6)。

清华大学 Xu 等^[66]提出了一种基于上下文从



图 6 空间位置上下文辅助下的三维模型检索^[65](左边为用户输入的边界盒以及关联模型(桌子)，右边为前 24 个检索结果)

Fig. 6 Context query results(Left: the user's query boxes with a desk as supporting object; right: the top 24 search results for each query^[65].)

草图构建三维场景的方法。该方法从三维场景库中学习得到一系列包含特定语义信息的结构组(如桌子周围放置四把椅子，电视对着沙发，床头柜置于床头两边)，并利用这些语义信息来加强基于草图的三维模型协同搜索和摆放。该方法支持一次输入整个场景的草图来生成对应的三维场景模型。

中国科学院深圳先进技术研究院 Xie 等^[28]提出一个基于草图的交互式三维模型设计系统(sketch2design)。该系统基于“部件合成建模”框架，即根据用户的手绘草图从数据库中检索相关部件组成新的模型。在该方法中，用户每次绘制一个部件(譬如一把椅子的靠背)的草图，系统从数据库中检索多个候选部件模型推荐给用户，用户选定其中一个后系统会把选定的部件自动组装到目标模型。图 7 是该系统的用户界面，图 8 给出该系统所用到的三维部件搜索框架。在进行部件检索的时候，系统不仅考虑用户输入的草图与候选部件的形状相似度，还附加两种上下文语义约束：(1)候选部件与目标模型中相邻部件之间(如图 8 中 p 与 q' 之间)应具有局部几何纹理相似性；(2)候选部件在源模型中的相邻部件与其在目标模型中的相应部件之间(如图 8 中 q 与 q' 之间)应具有全局几何相似性。约束(1)保证所

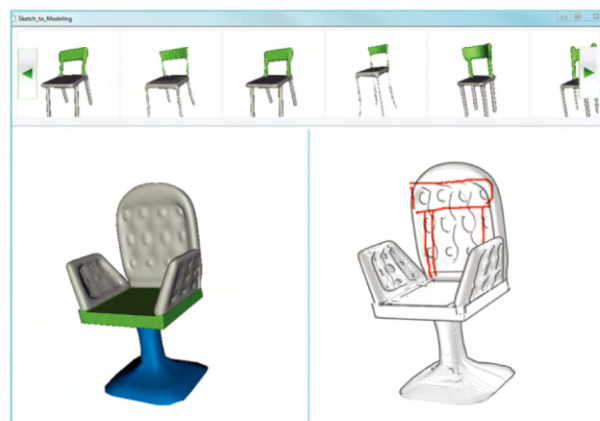


图 7 “sketch2design”系统用户交互界面^[28]

Fig. 7 User interface of the sketch2design system^[28]

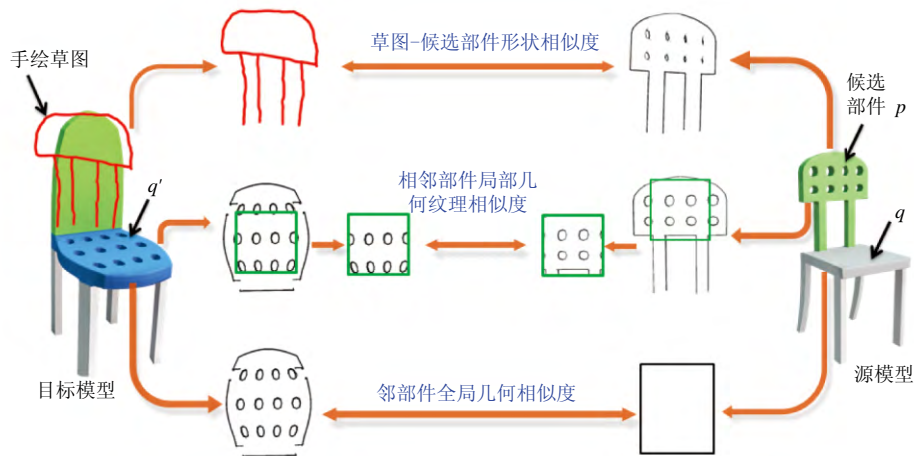


图 8 “sketch2design”系统中基于草图的三维部件检索框架(每个候选部件的检索得分由图中所示的三个相似度加权构成)

Fig. 8 Illustration of 3D part retrieval in the sketch2design system (In this approach, the relevance score for each candidate part is a weighted sum of three similarities.)

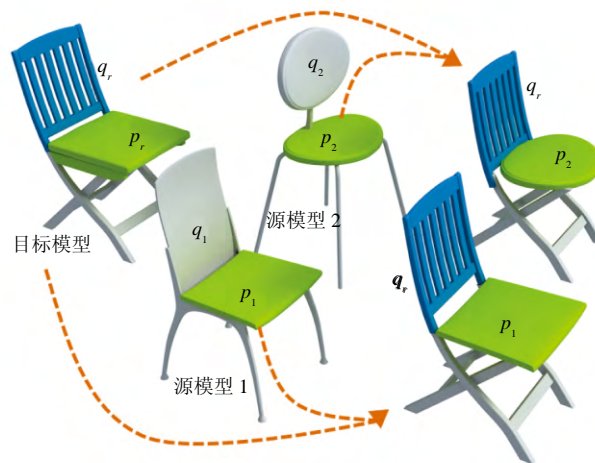


图 9 邻部件全局几何一致性约束在部件检索中的作用示例^[28](基于该约束条件,当目标模型中靠背 q_r 已经被固定的情况下,在检索椅子坐垫的时候,系统会倾向于推荐 p_1 而非 p_2)

Fig. 9 Illustration of the effect of overall geometric style consistency constraint on part retrieval (Based on this constraint, when retrieving a chair seat for the object model with the chair back q_r fixed, system tends to suggest p_1 rather than p_2 .)

生成的新模型的不同部件之间具有几何样式一致性(譬如都是条纹状或者网格状),约束(2)使得候选部件与已固定的相邻部件之间的拼接更加合理。如图 9,在目标模型中椅子靠背已经被固定,现在要检索一个椅子坐垫。对两个候选坐垫 p_1 和 p_2 ,根据约束(2),因为 $s(q_r, q_1) > s(q_r, q_2)$,

其中, s 表示全局几何相似度,系统会更加倾向于推荐 p_1 ,而事实上,“ $q_r + p_1$ ”的结合也显然比“ $q_r + p_2$ ”结合看起来更加合理。图 10 给出用不同种方法进行三维部件检索的例子比较。结果表明,在考虑到上述两种上下文约束之后,检索结果更加符合志愿者投票的结果。



图 10 基于手绘草图的三维部件检索结果^[28]: (a)表明目标模型中椅子坐垫已经被固定,要根据用户的手绘草图检索椅子靠背; (b)志愿者投票得分最高的 8 个候选部件; (c)没有考虑上下文信息的检索结果; (d)考虑了“sketch2design”系统中两种上下文约束得到的检索结果

Fig. 10 Illustration of 3D part retrieval from sketches^[28]: (a) the reference model and its contours with user's sketch, where the seat of chair has been fixed; (b) the top eight matching parts voted by volunteers; (c) the retrieved results according to only the user's sketch; (d) the results by considering both the user's sketch and contextual information

已有的工作表明,在基于手绘草图的三维形状检索中,有效利用上下文信息,不仅使得检索结果更加精准,还能在用户意向不明确的情况下得到更加合理的推荐。

5 总结与未来展望

草图是一种直接而快速的思想表达方式,因而发展基于草图的三维形状检索系统具有显著的应用前景。草图的模糊特性使得草图传递概念具有艺术性,另一方面,也给计算机理解草图带来巨大的挑战。本文对基于草图的三维形状检索相关技术进行了全面介绍。从形状检索性能来看,现有系统与人类理解草图的能力还相差甚远,还有很大提升空间。

在未来的研究中,可以考虑运用更接近人类感知原理的机器学习模型来自动对草图进行特征选择、特征提取以及语义理解。

草图与三维模型在形态上属于异质数据,存在明显的语义鸿沟,这给两者之间的匹配带来较大困难。有效利用上下文信息是克服这一困难的重要途径,即使在用户意图尚不明确的情况下都可以通过上下文辅助给出合理的推荐,从而进一步启发用户。现有方法主要通过学习场景中物体之间或者同一物体中不同部件之间的关系来实现语义理解。康奈尔大学 Jiang 等^[67]最新的研究指出,人类生活空间中物体之间的关系受某种隐含因素的影响。譬如,一个显示器和一个键盘之间的位置经常隐含着一个人的特定行为:一个人看着显示器敲打着键盘。所以,他们提出通过学习隐在的“人类上下文”(人-物关系)来辅助三维场景中的物体标记。在未来的研究中,要挖掘利用更加丰富的上下文信息来理解和检索三维模型。

上下文信息的深入挖掘应用给三维模型分析也提出了更高的要求。为挖掘更加丰富的上下文



图 11 三维模型的多视觉属性分解

Fig. 11 A sample of decomposition of multiple visual components from a 3D model

线索，有必要对三维模型进行多尺度、多视角分析，譬如，除了进行类似于“场景 - 物体 - 部件 - 零件”这样的多尺度模块分解，还可以进行“模型 - 整体轮廓 - 几何条纹 - 局部几何纹理 - 特殊形状”这样的多视觉属性提取(图 11)。总之，模型分析、模型检索、三维建模和三维应用往往存在紧密关联，在具体应用中，要综合多方面的相互影响，系统性地设计三维形状检索方案。

参 考 文 献

- [1] Zeleznik RC, Herndon KP, Hughes JF. SKETCH: an interface for sketching 3D scenes [C] // Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques, 1996: 163-170.
- [2] Funkhouser T, Min P, Kazhdan M, et al. A search engine for 3D models [J]. ACM Transactions on Graphics, 2003, 22(1): 83-105.
- [3] Bronstein AM, Bronstein MM, Guibas LJ, et al. Shape Google: geometric words and expressions for invariant shape retrieval [J]. ACM Transactions on Graphics, 2011, 30(1): 1.
- [4] Aono M, Iwabuchi H. 3D shape retrieval from a 2D image as query [C] // Signal & Information Processing Association Annual Summit and Conference, 2012: 1-10.
- [5] Kato T. Database architecture for content-based image retrieval [C] // Proceedings of the SPIE: International Society for Optics and Optical Engineering, 1992, 1662: 112-123.
- [6] Cao Y, Wang H, Wang CH, et al. Mindfinder: interactive sketch-based image search on millions of images [C] // Proceedings of the International Conference on Multimedia, 2010: 1605-1608.
- [7] Löffler J. Content-based retrieval of 3D models in distributed web databases by visual shape information [C] // Proceedings of the IEEE International Conference on Information Visualization, 2000: 82-87.
- [8] Eitz M, Richter R, Boubekeur T, et al. Sketch-based shape retrieval [J]. ACM Transactions on Graphics, 2012, 31(4): 31.
- [9] Eitz M, Hays J, Alexa M. How do humans sketch objects [J]? ACM Transactions on Graphics, 2012, 31(4): 44.
- [10] Daras P, Axenopoulos A. A 3D shape retrieval framework supporting multimodal queries [J]. International Journal of Computer Vision, 2010, 89(2-3): 229-247.
- [11] Yoon SM, Scherer M, Schreck T, et al. Sketch-based 3D model retrieval using diffusion tensor fields of suggestive contours [C] // Proceedings of the International Conference on Multimedia, 2010: 193-200.
- [12] Shin HJ, Igarashi T. Magic canvas: interactive design of a 3-D scene prototype from freehand sketches [C] // Proceedings of Graphics Interface, 2007: 63-70.
- [13] Lee J, Funkhouser T. Sketch-based search and composition of 3D models [C] // Proceedings of Eurographics Workshop on Sketch-Based Interfaces and Modeling, 2008: 97-104.
- [14] Saavedra JM, Bustos B, Scherer M, et al. STELA: sketch-based 3D model retrieval using a structure-based local approach [C] // Proceedings of Interna-

- tional Conference on Multimedia Retrieval, 2011: 26.
- [15] Shao TJ, Xu WW, Yin KK, et al. Discriminative sketch-based 3D model retrieval via robust shape matching [J]. Computer Graphics Forum, 2011, 30(7): 2011-2020.
- [16] Li B, Lu YJ, Fares R. Semantic sketch-based 3D model retrieval [C] // International Conference on Multimedia and Expo Workshops, 2013: 1-4.
- [17] Li B, Lu YJ, Johan H. Sketch-based 3D model retrieval by viewpoint entropy-based adaptive view clustering [C] // Proceedings of Eurographics Workshop on 3D Object Retrieval, 2013: 49-56.
- [18] Li B, Lu YJ, Godil A, et al. A comparison of methods for sketch-based 3D shape retrieval [J]. Computer Vision and Image Understanding, 2014, 119: 57-80.
- [19] Fonseca MJ, Ferreira A, Jorge JA. Sketch-based Retrieval of Vector Drawings [M]. London: Springer, 2011: 181-201.
- [20] Sousa P, Fonseca MJ. Sketch-based retrieval of drawings using spatial proximity [J]. Journal of Visual Languages and Computing, 2010, 21(2): 69-80.
- [21] Li B, Lu Y, Godil A, et al. SHREC'13 track: large scale sketch-based 3D shape retrieval [C] // Proceedings of the Sixth Eurographics Workshop on 3D Object Retrieval, 2013: 89-96.
- [22] Kazhdan M, Funkhouser T, Rusinkiewicz S. Rotation invariant spherical harmonic representation of 3D shape descriptors [C] // Proceedings of the Eurographics/ACM SIGGRAPH Symposium on Geometry Processing, 2003: 156-164.
- [23] Kazhdan M, Funkhouser T, Rusinkiewicz S. Shape matching and anisotropy [J]. ACM Transactions on Graphics, 2004, 23(3): 623-629.
- [24] Judd T, Durand F, Adelson E. Apparent ridges for line drawing [J]. ACM Transactions on Graphics, 2007, 26(3): 19.
- [25] Hildebrandt K, Polthier K, Wardetzky M. Smooth feature lines on surface meshes [C] // Proceedings of the Third Eurographics Symposium on Geometry Processing, 2005: 85-90.
- [26] DeCarlo D, Finkelstein A, Rusinkiewicz S, et al. Suggestive contours for conveying shape [J]. ACM Transactions on Graphics, 2003, 22(3): 848-855.
- [27] Fleishman S, Cohen-Or D, Silva CT. Robust moving least-squares fitting with sharp features [J]. ACM Transactions on Graphics, 2005, 24(3): 544-552.
- [28] Xie XH, Xu K, Mitra NJ, et al. Sketch-to-design: context-based part assembly [J]. Computer Graphics Forum. 2013, 32(8): 233-245.
- [29] Mokhtarian F, Suomela R. Robust image corner detection through curvature scale space [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(12): 1376-1381.
- [30] Hu R. Free-hand sketch based image and video retrieval [C] // Proceedings of the International Conference on Multimedia, 2010: 1697-1698.
- [31] Hu R, Barnard M, Collomosse J. Gradient field descriptor for sketch based retrieval and localization [C] // IEEE International Conference on Image Processing, 2010: 1025-1028.
- [32] Hou S, Ramani K. Classifier combination for sketch-based 3D part retrieval [J]. Computers and Graphics, 2007, 31(4): 598-609.
- [33] Mai F, Chang CQ, Hung YS. Affine-invariant shape matching and recognition under partial occlusion [C] // IEEE International Conference on Image Processing (ICIP), 2010: 4605-4608.
- [34] Belongie S, Malik J, Puzicha J. Shape context: a new descriptor for shape matching and object recognition [C] // NIPS, 2000: 831-837.
- [35] Zhang DS, Lu GJ. Shape-based image retrieval using generic Fourier descriptor [J]. Signal Processing: Image Communication, 2002, 17(10): 825-848.
- [36] Zhang DS, Lu GJ. A comparative study of Fourier descriptors for shape representation and retrieval [C] // Proceedings of Asian Conference on Computer Vision, 2002: 652-657.
- [37] Lowe DG. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [38] Ohbuchi R, Furuya T. Scale-weighted dense bag of visual features for 3D model retrieval from a partial view 3D model [C] // IEEE International Conference on Computer Vision Workshops, 2009: 63-70.
- [39] Eitz M, Hildebrand K, Boubekur T, et al. Sketch-based image retrieval: benchmark and bag-of-features descriptors [J]. IEEE Transactions on Visualization and Computer Graphics, 2011, 17(11): 1624-1636.
- [40] Dalal N, Triggs B. Histograms of oriented gradients for human detection [C] // IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005, 1: 886-893.
- [41] Eitz M, Hildebrand K, Boubekur T, et al. A

- descriptor for large scale image retrieval based on sketched feature lines [C] // Eurographics Symposium on Sketch-Based Interfaces and Modeling, 2009: 29-38.
- [42] Khotanzad A, Hong YH. Invariant image recognition by Zernike moments [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1990, 12(5): 489-497.
- [43] Revaud J, Lavoué G, Baskurt A. Improving Zernike moments comparison for optimal similarity and rotation angle retrieval [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(4): 627-636.
- [44] Ling HB, Jacobs DW. Shape classification using the inner-distance [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(2): 286-299.
- [45] Sivic J, Zisserman A. Video Google: a text retrieval approach to object matching in videos [C] // IEEE International Conference on Computer Vision, 2003: 1470-1477.
- [46] Baeza-Yates RA, Ribeiro-Neto B. Modern Information Retrieval [M]. New York: ACM Press, 1999.
- [47] Bronstein AM, Bronstein MM. Spatially-sensitive affine-invariant image descriptors [C] // Proceedings of the European Conference on Computer Vision: Part II, 2010: 197-208.
- [48] Li B, Schreck T, Godil A, et al. SHREC'12 track: sketch-based 3D shape retrieval [C] // Eurographics Conference on 3D Object Retrieval, 2012: 109-118.
- [49] Furuya T, Ohbuchi R. Ranking on cross-domain manifold for sketch-based 3D model retrieval [C] // IEEE International Conference on Cyberworlds, 2013: 274-281.
- [50] Veltkamp RC. Shape matching: similarity measures and algorithms [C] // IEEE International Conference on Shape Modeling and Applications, 2001: 188-197.
- [51] Rubner Y, Tomasi C, Guibas LJ. The earth mover's distance as a metric for image retrieval [J]. International Journal of Computer Vision, 2000, 40(2): 99-121.
- [52] Shao TJ, Xu WW, Yin KK, et al. Discriminative sketch-based 3D model retrieval via robust shape matching [J]. Computer Graphics Forum, 2011: 2011-2020.
- [53] Salton G, Buckley C. Term-weighting approaches in automatic text retrieval [J]. Information Processing and Management, 1988, 24(5): 513-523.
- [54] Jain P, Kulis B, Grauman K. Fast image search for learned metrics [C] // IEEE Conference on Computer Vision and Pattern Recognition, 2008: 1-8.
- [55] Li B, Johan H. Sketch-based 3D model retrieval by incorporating 2D-3D alignment [J]. Multimedia Tools and Applications, 2013, 65(3): 363-385.
- [56] Li B, Johan H. View Context: A 3D Model Feature for Retrieval [M]. Berlin Heidelberg: Springer, 2010: 185-195.
- [57] Divvala SK, Hoiem D, Hays JH, et al. An empirical study of context in object detection [C] // IEEE Conference on Computer Vision and Pattern Recognition, 2009: 1271-1278.
- [58] Torralba A. Contextual priming for object detection [J]. International Journal of Computer Vision, 2003, 53(2): 169-191.
- [59] Hoiem D, Efros AA, Hebert M. Putting objects in perspective [J]. International Journal of Computer Vision, 2008, 80(1): 3-15.
- [60] Rabinovich A, Vedaldi A, Galleguillos C, et al. Objects in context [C] // IEEE International Conference on Computer Vision, 2007: 1-8.
- [61] Torralba A, Murphy KP, Freeman WT, et al. Context-based vision system for place and object recognition [C] // IEEE International Conference on Computer Vision, 2003: 273-280.
- [62] Shotton J, Winn J, Rother C, et al. Textonboost: joint appearance, shape and context modeling for multi-class object recognition and segmentation [C] // European Conference on Computer Vision, 2006: 1-15.
- [63] Goferman S, Zelnik-Manor L, Tal A. Context-aware saliency detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(10): 1915-1926.
- [64] The context challenge [OL]. [2014-12-30]. <http://web.mit.edu/torralba/www/carsAndFacesInContext.html>.
- [65] Fisher M, Hanrahan P. Context-based search for 3D models [J]. ACM Transactions on Graphics, 2010, 29(6): 182.
- [66] Xu K, Chen K, Fu H, et al. Sketch2scene: sketch-based co-retrieval and co-placement of 3D models [J]. ACM Transactions on Graphics, 2013, 32(4): 123.
- [67] Jiang Y, Koppula H, Saxena A. Hallucinated humans as the hidden context for labeling 3D scenes [C] // IEEE Conference on Computer Vision and Pattern Recognition, 2013: 2993-3000.