

非刚性三维模型检索特征提取技术研究*

李海生^{1,2}, 孙莉^{1,2}, 武玉娟^{1,2}, 吴晓群^{1,2}, 蔡强^{1,2}, 杜军平³



¹(北京工商大学 计算机与信息工程学院, 北京 100048)

²(食品安全大数据技术北京市重点实验室(北京工商大学), 北京 100048)

³(北京邮电大学 计算机学院, 北京 100876)

通讯作者: 李海生, E-mail: lihsh@th.btbu.edu.cn

摘要: 三维模型特征描述符是一种简洁且信息量丰富的表示方式,特征提取是许多三维模型分析处理任务的关键步骤.近年来,针对非刚性三维模型特征提取技术的研究引起了人们的广泛关注.首先,汇总了常用的非刚性三维模型基准数据集和算法评价标准;然后,在广泛调研大量文献和最新成果的基础上,将非刚性三维模型特征分为人工设计的特征描述符和基于学习的特征描述符两大类,并分别加以介绍,对每类方法所包含的典型算法,尤其是近几年基于深度学习的特征提取算法的基本思想、优缺点进行了分析、对比和总结;最后进行总结,并对未来可能的发展趋势进行了展望.

关键词: 非刚性三维模型;特征提取;深度学习

中图法分类号: TP391

中文引用格式: 李海生,孙莉,武玉娟,吴晓群,蔡强,杜军平.非刚性三维模型检索特征提取技术研究.软件学报,2018,29(2): 483-505. <http://www.jos.org.cn/1000-9825/5379.htm>

英文引用格式: Li HS, Sun L, Wu YJ, Wu XQ, Cai Q, Du JP. Survey on feature extraction techniques for non-rigid 3D shape retrieval. Ruan Jian Xue Bao/Journal of Software, 2018, 29(2): 483-505 (in Chinese). <http://www.jos.org.cn/1000-9825/5379.htm>

Survey on Feature Extraction Techniques for Non-Rigid 3D Shape Retrieval

LI Hai-Sheng^{1,2}, SUN Li^{1,2}, WU Yu-Juan^{1,2}, WU Xiao-Qun^{1,2}, CAI Qiang^{1,2}, DU Jun-Ping³

¹(School of Computer and Information Engineering, Beijing Technology and Business University, Beijing 100048)

²(Beijing Key Laboratory of Big Data Technology for Food Safety (Beijing Technology and Business University), Beijing 100048)

³(School of Computer Science, Beijing University of Posts and Telecommunications, Beijing 100876)

Abstract: Shape descriptor is a concise and informative representation. Feature extraction is a key step in many 3D shape analysis tasks. In recent years, feature extraction technologies of non-rigid 3D shape have attracted a lot of attentions. This paper firstly introduces the evaluation criteria and the datasets which are commonly used as benchmark in non-rigid 3D shape feature extraction. Secondly, based on extensive research on the existing literatures and the latest achievements, the paper categorizes the non-rigid 3D shape descriptors into two types: Hand-Crafted shape descriptors and learning based shape descriptors. The basic ideas, advantage and disadvantage of typical algorithms belong to each category, especially the most recent feature extraction algorithms based on deep learning are analyzed, compared and summarized. Finally, some potential future work is discussed.

Key words: non-rigid 3D shape; feature extraction; deep learning

* 基金项目: 国家自然科学基金(61320106006, 61532006, 61602015); 北京市自然科学基金(4162019); 北京市科技计划(Z161100 001616004)

Foundation item: National Natural Science Foundation of China (61320106006, 61532006, 61602015); Beijing Natural Science Foundation (4162019); Beijing Science and Technology Project (Z161100001616004)

收稿时间: 2017-05-19; 修改时间: 2017-07-16; 采用时间: 2017-08-29; jos 在线出版时间: 2017-10-09

CNKI 网络优先出版: 2017-10-09 16:20:58, <http://kns.cnki.net/kcms/detail/11.2560.TP.20171009.1620.009.html>

人类的视觉系统所感知到的世界是立体的、三维的,相对于图像、声音和视频等多媒体数据来说,三维模型及其组成的三维场景能够更全面、真实地展现物体的信息,是一种更符合人类视觉感知的数据形式,被广泛地应用在虚拟现实、工程制图、城市规划、影视娱乐、文物复原和生物医学等领域中^[1].

近年来,新的三维扫描设备和建模工具的出现,使三维建模变得简单、快捷、高效;显卡技术的发展,使三维模型的显示和处理速度得到大幅的提高;网络技术的发展则进一步推动了三维模型的传播和应用^[2].互联网上已涌现出越来越多的三维模型^[3],合理重复使用现有的三维模型,对其进行编辑和修改以符合设计需要,不仅能够节约成本、减少构建模型的工作量,而且还能推动三维模型的流通,并对电子商务、虚拟现实等行业的发展有着很强的促进作用.这需要高效的三维模型检索系统作为支撑,三维模型特征提取引起了国内外学者的广泛关注^[4].

与图像检索发展过程类似,三维模型检索的发展也分为两个阶段,即基于文本的检索和基于内容的检索^[5].基于文本的检索方式需要对模型进行人工标注^[6,7],具有主观性和片面性,并且在大数据背景下不具备鲁棒性和可扩展性;基于内容的检索方式是以模型本身作为处理对象,不再需要人工标注,减少了人为干预,同时提高了检索效果,成为三维模型检索领域的主流技术^[8,9].一个典型的基于内容的三维模型检索系统框架如图 1 所示,主要包括模型预处理、特征提取和相似度比较这 3 个步骤^[3].其中,预处理主要包括把模型转换到一个标准统一的坐标空间,或对模型进行简化、去噪等处理,为下一步的特征提取做准备;特征提取是依据物理或数学原理分析模型的内蕴属性,构建模型的特征描述符;相似度比较是依据模型特征描述符判定模型间的匹配程度,进而获取多个相似的三维模型.通过以上分析可以看出,如何对三维模型提取信息量丰富且具有识别力的特征描述符,是三维模型检索任务要解决的关键性问题.

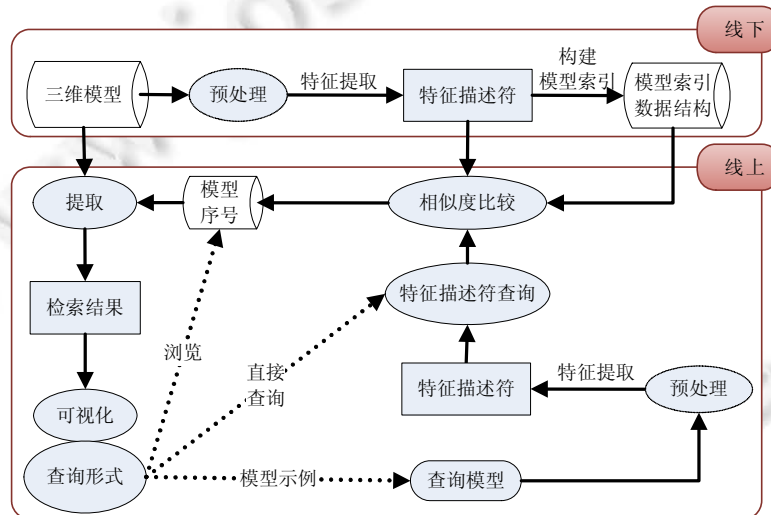


Fig.1 Conceptual framework for content-based 3D shape retrieval system^[3]

图 1 基于内容的三维模型检索系统概念框架^[3]

三维模型有刚性模型和非刚性三维模型之分:刚性模型在受力前后内部点的相对位置不发生改变,对应的变换主要有旋转、平移、缩放和映射等;非刚性三维模型具有铰链结构,模型内部点的相对位置会因受力而发生改变,模型变换种类多样、形式复杂,更加接近真实世界中物体的变化^[10].针对刚性模型,已有较成熟的检索系统,例如美国普林斯顿大学设计的三维模型搜索引擎^[11,12]、德国康斯坦兹大学开发的三维模型匹配与检索系统^[13,14]和日本国立多媒体研究所设计的 Ogden IV 系统^[15,16]等.针对非刚性三维模型,在提取其特征描述符时需要考虑模型获取技术和模型自由度所引发的模型变换问题.如图 2 所示,非刚性三维模型的变换可以被分为 12 类,主要有等距变换、拓扑变换、噪声、散粒噪声、大孔、小孔、采样、光栅化、部分缺失、视图、仿射变换和尺寸变换等^[17].面对丰富多样的非刚性变换,一方面,刚性模型特征提取算法在非刚性三维模型上并不能得到

同样有效的结果;另一方面,现有的非刚性三维模型特征提取算法都有其适应性和局限性,针对非刚性三维模型特征提取技术的研究成为计算机图形学、计算机视觉等领域的研究热点^[18]。

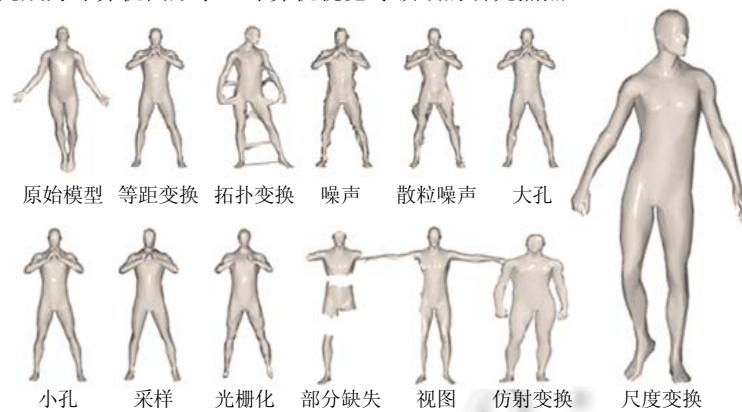


Fig.2 Transformations of the human shape^[17]

图2 人体模型的不同变换形式^[17]

本文主要综述了基于内容的非刚性三维模型检索的近期学术发展成果,重点是对非刚性三维模型特征提取技术进行综合分析并比较,主要贡献如下:

- (1) 汇总统计了非刚性三维模型检索的实验基准数据集和算法评价指标;
- (2) 系统地对非刚性三维模型特征提取技术进行了综合研究和分类,对每一类方法中的典型算法按照时间先后进行阐述、分析和比较;
- (3) 重点对如何应用深度学习技术处理非刚性三维模型数据的方法进行分析,并对不同方法所获取的效果进行了比较;
- (4) 讨论了非刚性三维模型特征提取技术目前存在的主要难点和今后的发展趋势。

本文第1节对非刚性三维模型检索实验所用的基准数据集和算法评价指标进行总结,第2节对已有的非刚性三维模型检索特征提取技术进行阐述分析和归纳比较,第3节对非刚性三维模型检索特征提取技术所面临的主要难点和今后的发展趋势进行总结和探讨,最后总结全文。

1 非刚性三维模型检索实验数据集和算法评价指标

1.1 实验数据集

标准的非刚性三维模型数据集包括 McGill 模型集^[19]、TOSCA 模型集^[20,21]、Sumner 模型集^[22]、Shapegoogle 模型集^[5]和 SHREC 模型集^[23]等,对以上非刚性三维模型数据集更为详细的介绍和分类见表 1。为了利用深度学习技术提取三维模型特征,普林斯顿大学等机构提供了两个大规模的三维模型数据集,分别为 ShapeNet 数据集^[24]和 ModelNet 数据集^[25]。其中,ShapeNet 包含了约 300 万个模型,其子集 ShapeNetCore 包含了来自 55 类的 51 300 个模型;ModelNet 数据集包含了来自 662 类的 12 7915 个三维模型,其子集 Model10 包含了来自 10 类的 4 899 个三维模型,ModelNet40 包含了来自 40 类的 12 311 个三维模型。

Table 1 Non-Rigid 3D shapes benchmark

表 1 非刚性三维模型基准数据集

名称	描述	发布网址	分类&总数	部分模型示例
McGill ^[19]	McGill 数据集规模比较小,但是选取的不同姿势的各类模型非常具有代表性	http://www.cim.mcgill.ca/~shape/benchMark/	由 255 个非刚性三维模型组成,分为 10 个类,类中对象的最大数量为 30,最小数量为 20	

Table 1 Non-Rigid 3D shapes benchmark (Continued)

表 1 非刚性三维模型基准数据集(续)

名称	描述	发布网址	分类&总数	部分模型示例	
TOSCA ^[20,21]	高分辨率 (每个模型大约由50 000个顶点组成)	TOSCA 数据集规模比较小,但是选取的不同姿势的各类模型非常具有代表性	http://tosca.cs.technion.ac.il/data/toscahires-mat.zip	共包含 9 个类,80 个非刚性三维模型,其中,人体模型为 3 类:两类不同的男性模型分别是 7 个与 20 个不同的姿势;一类女性模型为 12 个不同姿势,另外还有不同姿势下的 4 只猩猩、8 匹马、9 只狗、11 只猫、6 匹半人马、3 匹狼等模型	
	低分辨率 (每个模型大约由3 000个顶点组成)		http://tosca.cs.technion.ac.il/data/nonrigid3d.zip	共包含 12 个类,148 个非刚性三维模型,其中,人体模型为 3 类:两类不同的男性模型分别是 15 个与 20 个不同的姿势;一类女性模型为 24 个不同姿势,另外还有不同姿势下的 21 只猩猩、17 匹马、15 头狮子、1 只狗、9 只猫、6 只海马、6 匹半人马、3 匹狼、1 条鲨鱼等模型	
Sumner ^[22]	Sumner 数据集规模比较小,而且模型形变程度比较严重.每个模型大概有 3 000 个~ 5 000 个顶点	http://people.csail.mit.edu/sumner/research/deftransfer/data.html	共 100 个模型,包含男性、马、狗、鸟和骆驼这 5 个类别,对这些类别做了 4 种变换:尺寸变换、局部尺寸变换、取部分变换、拓扑变换		
ShapeGoogle ^[5]	ShapeGoogle 数据集共包含两部分:一部分是用来做检索的非刚性三维模型,并且分类明确;另一部分是不属于任何类的无关模型,用来扩充模型集和测试检索算法的性能.每个模型大概有 1 200 个~1 500 个顶点	http://www.lix.polytechnique.fr/~maks/shapegoogle_code.zip	用来做检索的部分数据集包含 10 个类别,分别是马、猫、狗、半人马、男性、女性、狮子、大象、骆驼和鸟.这 10 个类别经过了 6 种变换:等容等距变换、拓扑变换、等容等距+拓扑变换、噪声变换、部分缺失变换、采样变换.10 个类别经过 6 种变换后,共有 596 个模型.用来扩充数据集规模的无关模型 463 个		

Table 1 Non-Rigid 3D shapes benchmark (Continued)

表 1 非刚性三维模型基准数据集(续)

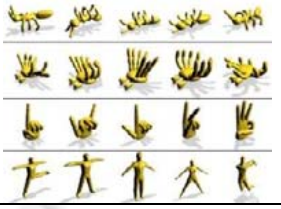

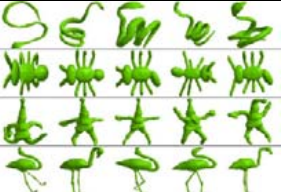






名称	描述	发布网址	分类&总数	部分模型示例
SHREC 2010 ^[26] (检索)	模型为水密的三角形网格,不含有拓扑结构误差,是标准的非刚性三维模型数据集	http://www.itl.nist.gov/iad/vug/sharp/contest/2010/NonRigidShapes/	共包含 200 个姿态各异的非刚性三维模型,平均分为 10 类	
SHREC 2010 ^[27] (鲁棒性)	用于测试非刚性三维模型描述符应对模型发生不同程度形变的鲁棒性	http://tosca.cs.technion.ac.il/book/shrec_robustness_2010.html	共含有 13 个类别 728 个模型,其中,13 个为原始模型,并对每个原始模型模拟了 11 种形变,在每类变换中,设定了 1 个~5 个形变强度等级	
SHREC 2011 ^[28] (检索)	模型为水密的三角形网格,不含有拓扑结构误差,是标准的非刚性三维模型数据集	http://www.itl.nist.gov/iad/vug/sharp/contest/2011/NonRigid/	共包含 600 个姿态各异的非刚性三维模型,平均分为 30 类	
SHREC 2011 ^[17] (鲁棒性)	用于测试非刚性三维模型描述符应对模型发生不同程度形变的鲁棒性	http://tosca.cs.technion.ac.il/book/shrec_robustness.html	共含有 13 个类别 793 个模型,其中,13 个为原始模型,并对每个原始模型模拟了 12 种形变,在每类变换中,设定了 1 个~5 个形变强度等级	
SHREC 2014 ^[29]	Real	http://www.cs.cf.ac.uk/shaperetrieval/shrec14/index.html	Real 数据集中的模型是通过直接扫描真实的人体构建的,共包含 400 个模型,平均分为 40 类	
	Synthetic		Synthetic 数据集中的模型是用三维模型建模软件生成的,共包含 300 个模型,平均分为 15 类	
SHREC 2015 ^[30] (检索)	模型为水密的三角形网格,不含有拓扑结构误差,是标准的非刚性三维模型数据集	http://www.icst.pku.edu.cn/zlian/shrec15-non-rigid/data.html	共包含 1 200 个非刚性三维模型,平均分为 50 类	

Table 1 Non-Rigid 3D shapes benchmark (Continued)

表 1 非刚性三维模型基准数据集(续)

名称	描述	发布网址	分类&总数	部分模型示例
SHREC 2015 ^[31] (标准形)	用于比较基于模型“标准形”的特征描述符的识别能力,一些模型包含孔、洞,一些模型含有自相交结构	http://www.cs.cf.ac.uk/shaperetrieval/shrec15/SHREC15.zip	包含两个子集,即训练集和测试集,每个数据集都含有 100 个非刚性三维模型,平均分为 10 类	
SHREC 2016 ^[32] (拓扑噪声)	带有拓扑噪声的形变模型,用于模型匹配实验	http://vision.in.tum.de/~laehner/shrec2016/	该数据集包含 1 个没有发生拓扑变化的原始非刚性三维模型及其 25 个拓扑形变版本	
SHREC 2016 ^[33] (部分结构)	原始模型来源于 TOSCA,对每个模型均提取部分结构组成部分缺失的非刚性三维模型数据集,用于测试模型部分匹配	http://www.dais.unive.it/~shrec2016/dataset.php	包括 76 个近似等体积的模型,被分为 8 个类,经过重新划分,缺失部分以常规分割和不规则孔两种方式引入,并产生两个不同的数据集.其中,常规分割产生的数据集包括 320 个部分模型,不规则孔产生的数据集包括 279 个模型	
SHREC 2017 ^[34] (点云结构)	对三维实体进行扫描获取非刚性三维点云数据集,用以评估检索算法在非刚性点云数据集上的有效性以及应对扫描时模型结构自遮挡而产生部分缺失等的鲁棒性	https://www.cs.york.ac.uk/cvpr/pronto/dataset.html	该数据集包含 100 个非刚性点云模型,平均分为 10 类,每类模型分别扫描了 10 种不同的姿势	

1.2 算法评价指标

欧洲图形学会每年都举办一次三维模型检索竞赛(3D shape retrieval contest,简称 SHREC)^[23],在历年的竞赛中,Shilane 等人^[35,36]提出的普林斯顿形状基准(Princeton shape benchmark,简称 PSB)是衡量算法优劣的主要方式,已成为事实上的通用评估标准.PSB 主要包括以下评价指标.

- (1) 查准率(precision rate,简称 P)和查全率(recall rate,简称 R):查准率是指检索结果中同类模型比率,查全率是指检索结果中同类模型占数据集中该类模型总量的比率.
- (2) PR 曲线:PR 曲线由查准率与查全率之间的函数关系生成,在一定范围内,查准率与查全率呈现出反比关系,因此,PR 曲线与坐标轴形成区域的面积大小与检索能力成正比,面积越大,则检索性能越好.
- (3) 最近邻方法(nearest neighbor,简称 NN):假设检索某类模型,返回的模型总数为 N ,其中相关模型数量为 K ,则

$$NN=K/N \quad (1)$$

- (4) 第 1 层级(first tier,简称 FT)和第 2 层级(second tier,简称 ST):假设对类型 C 模型进行检索,数据集中的所有相关模型的总量是 $|C|$,而检索返回的相关模型数量为 K ,则

$$\begin{cases} FT = K / (|C| - 1) \\ ST = K / (2(|C| - 1)) \end{cases} \quad (2)$$

- (5) E 度量(E _measure,简称 E):考虑前 32 个检索结果并计算 Precision-Recall(P-R),则

$$E=2/(1/P+1/R) \quad (3)$$

- (6) 折扣的累积结果(discounted cumulative gain,简称 DCG):DCG 体现了正确匹配模型的顺序,定义如下所示.

$$DCG_i = \begin{cases} G_i, & i = 1 \\ DCG_{i-1} + \frac{G_i}{\lg_2(i)}, & i > 1 \end{cases} \quad (4)$$

$$DCG = \frac{DCG_k}{1 + \sum_{j=2}^{|C|} \frac{1}{\lg_2(j)}} \quad (5)$$

该定义首先满足一个假设条件:在检索结果的排序列表中,几乎不考虑排在列表后面的模型.假设把有序列表 R 转换成列表 G ,如果模型 R_i 是正确匹配的,即与查询模型属于同一个类别,则 G_i 的值为 1;否则为 0.

2 非刚性三维模型检索特征提取技术

一个理想的非刚性三维模型特征描述符应该具备以下几个特点.

- (1) 唯一性.一个形状的三维模型应该与它的特征描述符保持一一对应关系.
- (2) 显著的辨别力.特征描述符能够对模型的内蕴属性进行编码,同时剔除不重要的细节部分,能够辨识出视觉上很相似但并不属于同一类别的两个模型,并且给出准确的相似度量.
- (3) 丰富的信息量.特征描述符能够保证蕴含尽量多的三维模型特征信息,能够刻画模型的不同细节层次.
- (4) 较好的鲁棒性和稳定性.特征描述符对模型的刚性变换和非刚性变换保持不变,面对原始模型中微小的非等距变换、噪声、孔洞、退化等扰动保持稳定.
- (5) 简单、高效性.特征描述符维度尽量低,提取时间尽量短,满足查询系统实时性的要求.

本文依据对模型内容表征角度的不同,将非刚性三维模型检索特征提取技术分为两大类:一是人工设计的特征描述符(hand-crafted shape descriptor)提取技术,即人们根据三维模型表面或立体形状的几何性质来设计模型特征提取算法;另一类是基于学习的特征描述符(learning based shape descriptor)提取技术,即利用机器学习的方法,通过数据驱动自动学习模型的特征描述符.对于每一个大类,又可将其中的方法进一步细分为若干子类,如图 3 所示.

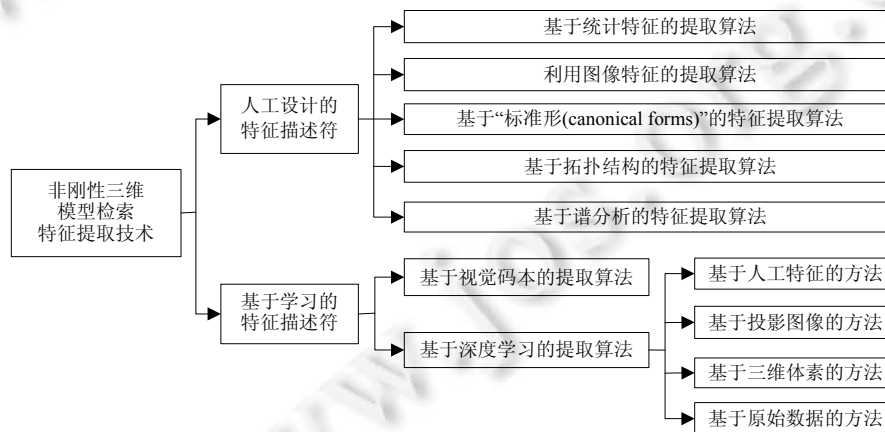


Fig.3 Classification of feature extraction techniques for non-rigid 3D shape retrieval

图 3 非刚性三维模型检索特征提取技术分类

2.1 人工设计的特征描述符提取技术

(1) 基于统计特征的提取算法

这类算法将模型匹配问题简化为概率分布问题,从几何信息分布的角度去描述模型的结构特征,非常直观、易于理解.但是也有两点不足:一是往往聚焦在模型的某一种几何特征上,对模型的描述不够充分;二是作为全局模型描述符,它很可能会忽视掉一些重要的局部信息.因此一般情况下,基于统计特征的提取算法应用于模型的粗检索过程^[37].

其思想来源于 2002 年 Osada 等人^[38]提出的形状分布(shape distribution,简称 SD)算法,即首先选择一个合适的形状函数来度量三维模型,例如模型表面任意两点之间的欧氏距离;然后计算形状函数的分布,并随机采样构造分布直方图;最后,构建分布直方图的分段线性函数作为特征描述符去刻画三维模型的全局几何特性.但是这种方法所用到的欧氏距离对于非刚性三维模型的等距变换不具有有一致性.

为了更加准确地反映非刚性三维模型的内蕴属性,2008 年,Mahmoudi 等人^[39]定义了模型的测地距离、扩散距离、曲率加权距离等几何信息,并统计其分布直方图作为特征描述符,用于非刚性三维模型分析、匹配.测地距离等虽然把模型表面顶点之间的空间连通性考虑在内,但是对于局部噪声比较敏感,且对于高分辨率的模型来说,计算时间比较长.

为了缩短计算时间,2010 年,Ruggeri 等人^[40]利用拉普拉斯特征映射的临界点作为锚点,并且将最远点采样与统计标准组合,生成一组具有等距变换不变性的采样点;2011 年,Sipiran 等人^[41]引入 Harris 算子检测出三维网格模型上的显著点,并统计模型表面采样点间的测地距离分布作为特征描述符.实验结果表明:采样机制的引入不仅提高了计算效率,而且特征描述符具有等距变换不变性,适用于非刚性三维模型检索.

2011 年,Pan 等人^[42]基于泊松方程提出了泊松直方图特征描述符.该方法将泊松方程以离散化的形式定义到三维模型体素上,根据狄利克雷边界条件,边界体素处的势函数值保持不变;然后,基于泊松方程的解得到每个内部体素的势函数值;最后统计这些值的分布直方图.由于不依赖于坐标系,这一特征对刚性变换具有鲁棒性.它不仅能够很好地捕获模型的结构信息,而且对拓扑变换、边界噪声不敏感,也适用于非刚性三维模型分析.

2012 年,Ohkita 等人^[43]提出根据模型表面局部范围内定向点的法向量及坐标值,为每个采样点关联一组局部特征四元组,并统计得到一个四维联合直方图;最后整合采样点的局部统计特征,得到模型的一个特征向量.该方法可以处理多种形式表示的三维模型,包括多边形网格和定向点集合等,能够捕获模型的局部信息,对铰链结构的形变具有不变性.但该方法还需要进一步探索采样点数、局部范围半径等参数对检索结果的影响.

2014 年,Pickup 等人^[29]在 SHREC2014 上提出利用模型的表面积作为特征描述符.虽然这一方法不可避免地忽视掉了很多局部信息,不适用于局部特征匹配,但是实验结果表明:这一特征对等距变换不敏感,计算简单,实时性好,适用于对大规模、高分辨率的非刚性三维模型数据集进行粗检索.

(2) 利用图像特征的提取算法

相对于三维模型来说,图像特征提取技术效果较好,应用广泛.因此,如何利用图像特征提取技术进行三维模型检索,得到了越来越多学者的关注.目前,这类算法主要有两个思路:一是立足于二维图像特征提取技术的算法本质,对其加以改进,延伸到三维模型领域;二是对三维模型进行投影,对投影视图或投影轮廓应用二维图像提取技术.第 1 种思路不需要对模型进行降维处理,符合人们的认知经验;第 2 种思路降低了特征提取的复杂度,但容易丢失模型的空间信息,影响检索准确率.

对于第 1 种思路,在诸多的图像特征描述符中,Lowe^[44]所提出的尺度不变特征转换(scale invariant feature transform,简称 SIFT)描述符应用最广.2008 年,Castellani 等人^[45]在此基础上对三维模型设计了一种局部特征描述符.算法主要分为两个阶段:一是对三维模型进行多尺度分析,并检测尺度变换下仍然稳定的特征显著点;二是利用隐马尔可夫模型对特征显著点做进一步的统计分析.其优点是能够很好地检测出特征显著区域,对旋转、不规则采样、分辨率变化以及对含孔洞的模型都有很好的鲁棒性.2011 年,Castellani 等人^[46]又对其做了进一步改进,通过引入等测地线圆环,提出了多环隐马尔可夫模型.首先,计算每一条等测地线圆环上每一个点的多个几何属性;然后,利用这些几何属性训练出隐马尔可夫模型的隐含参数作为这一条测地线圆环的特征向量.

与前期工作相比,在建立局部支撑区域时,利用等测地线圆环代替了螺旋测地曲线,对非刚体的姿态变换以及局部区域的剧烈形变都有很好的鲁棒性,但是对原始模型的点采样敏感。

2009年,Zaharescu 等人^[47]在 SIFT 特征和方向梯度直方图(histogram of oriented gradients,简称 HOG)特征的启发下,提出一种网格方向梯度直方图特征描述符。首先对三角网格均匀采样,然后确定特征点的局部支撑区域并创建特征点的局部坐标系,最后进行空间梯度直方图的统计。该方法将 SIFT 特征引入到流形曲面,能够有效地对三维模型进行特征点检测;另外设计了自适应性的支撑区域大小确定方法,使得描述符具有旋转不变性。但该方法在实际应用中往往过分依赖于三维模型表面三角网格的质量,要求理想的均匀采样,故有一定的局限性。

2010年,Maes 等人^[48]提出了一种网格尺度不变特征转换描述符,将其应用于三维人脸识别。随后,Smeets 等人在 2011 年将其应用于非刚性三维模型检索,并在 SHREC 2011^[28]中取得了较高的检索准确率。该特征首先将网格模型检测尺度空间极值点作为局部特征点;其次,对每一个特征点赋予一个方向,使得特征具有旋转不变性;然后构建局部特征向量,使用模型索引和倾斜角的级联直方图描述每个尺度空间极值的邻域;最后,通过比较特征在空间中的角度进行特征匹配,并基于匹配的数量比较两个模型之间的相似性。实验结果显示,这一描述符对非刚性三维模型有很好的表达能力。

对于第 2 种思路,1999 年,Johnson 等人^[49]提出了旋转图像特征(spin images,简称 SI)。该特征在流形空间内定义,将三维模型特征匹配问题转化为二维投影图像的模板匹配问题。尽管这种方法对于噪声和三角化具有很好的鲁棒性,但是对非刚性三维模型形变比较敏感。因此,2006 年,Liu 等人^[50]通过改进 SI 特征应用于非刚性部分模型检索。2010 年,Wang 等人^[51]进一步提出本征自旋图像描述符,首先将 N 维空间中的局部本征模型分解为两个子空间,然后将模型顶点投影到每个子空间上。这一特征不仅具有等距变换不变性,而且对非参数化的自旋图像有着高度的表现力。

2008 年,Ohbuchi 等人^[52]对模型的多角度投影视图提取显著的局部视觉特征,并结合词袋模型提出一种特征描述符,称为 BF-SIFT。首先,将模型位置、尺度和姿势进行归一化;其次,利用包围球均匀选取多个视点,对模型进行投影,再对投影得到的深度图像提取 SIFT 特征;然后,利用词袋模型对局部视觉特征进行矢量化;最后,将视觉词汇的频率直方图作为模型的特征描述符,算法流程如图 4 所示。

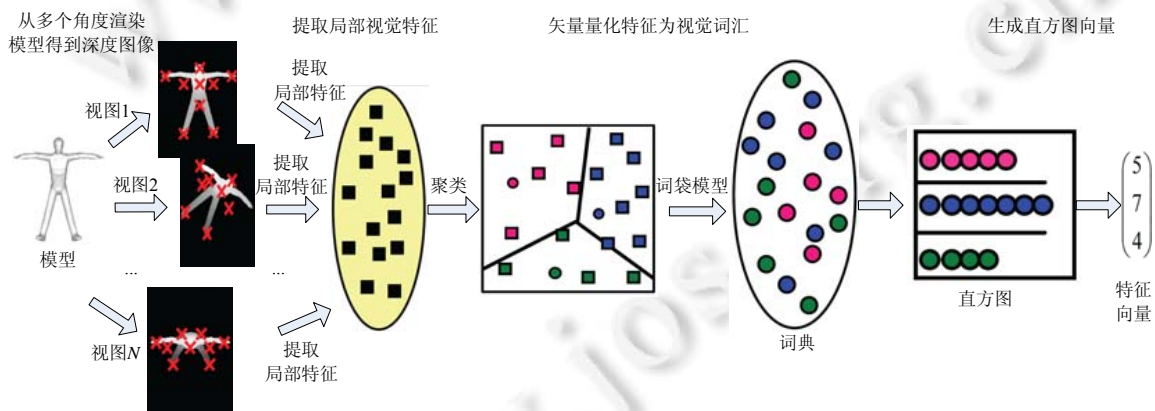


Fig.4 Flow of the BF-SIFT feature extraction ^[52]

图 4 BF-SIFT 特征提取流程^[52]

BF-SIFT 特征描述符具有尺度变换、旋转变换及等距变换不变性等优点,但仍存在码本获取时间较长的问题。Furuya 等人^[53]在 2009 年又利用密集采样和快速编码机制改进了 BF-SIFT 特征,使得计算更为高效,对模型细节的描述能力更强。这一算法也在 SHREC 2010^[26]中取得了最优的检索结果。

(3) 基于“标准形”的特征提取

对于非刚性三维模型而言,由于具有形状弯曲、姿势多变等非刚性变换特点,对其进行预处理的方法比刚性模型的预处理方法更为复杂.这类算法的主要目的是将不同姿势的模型归一化到一个标准空间中,从而把非刚性三维模型匹配问题转化成刚性模型匹配问题.目前主要是采用空间嵌入的方式,在等距变换的条件下,将三维模型上的点统一映射到另一个坐标系空间中.

2003年,Elad等人^[54]首次提出利用多维尺度(multi-dimensional scaling,简称MDS)分析算法将模型由测地距离空间转换到欧式空间,把不同姿势的三维模型归一化到统一的姿态.MDS是一种等距变换,能够保持原有模型的内蕴属性不变.

2010年,Lian等人^[55]利用MDS分析算法和主元分析(principal component analysis,简称PCA)算法计算模型的“标准形”,然后对模型的“标准形”提取BF-SIFT特征,最后利用时钟匹配技术计算模型的相似度.其中,“标准形”的计算主要有3步:首先对模型进行简化,然后应用MDS嵌入算法将简化后的模型映射到三维欧氏空间中,最后利用PCA算法将模型进行调整对齐得到最终的“标准形”.实验结果表明:相对于直接在原始的非刚性三维模型上提取BF-SIFT特征,标准化之后的三维模型上的BF-SIFT特征能够显著提高检索准确率.

2013年,Lian等人^[56]提出了一种保特征的“标准形”,具体过程如图5所示.首先,将非刚性三维模型分割成一个个接近于刚性的子部分;然后,对这些子部分进行适当的旋转和平移,使它们的姿势在位置和方向上与MDS规范形式有着良好的对应关系;最后,对准和平滑子部分之间的边界来获得模型的“标准形”.丰富的对比实验结果表明:这种方法所得到的“标准形”能够保留重要的局部细节信息,对模型整体的等距变换具有不变性,提高了检索准确率.该方法也有局限性:对拓扑噪声敏感;分割结果决定了“标准形”的效果;计算成本相对较高.

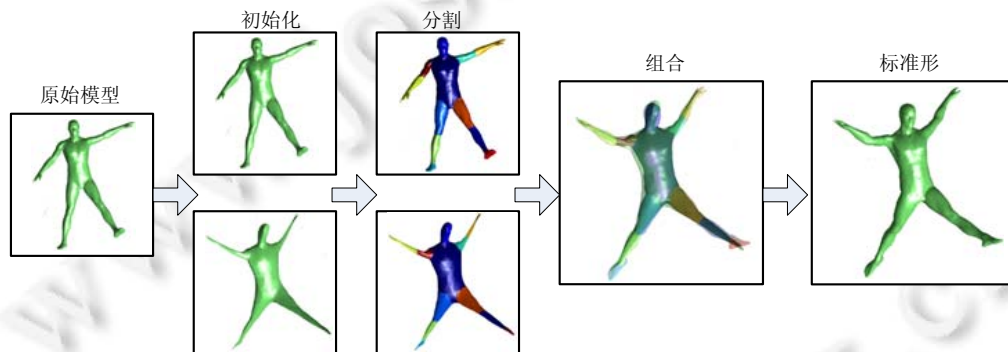


Fig.5 Flow of constructing feature-preserved canonical form^[56]

图5 构造保特征的“标准形”的流程^[56]

基于MDS嵌入的方式在处理高分辨模型时速度较慢,难以满足实时性的要求.2015年,Pickup等人^[57]提出了一种在线性时间复杂度下计算模型“标准形”的方法:首先,基于模型顶点的保形因子^[58]选取特征点;然后计算特征点之间最大的欧氏距离.该方法不需要对模型简化,能够在合理时间内处理高分辨率的模型.并且对不同分辨率的模型,特征点的位置一般是稳定的,因此会形成相同的标准化姿态.但是该方法也有两点限制:一是模型特征点的数目被固定,在实际应用中可能会丢失一部分特征点;二是对于模型的噪声扰动比较敏感.

为了比较不同的模型“标准形”获取方法,Pickup等人^[31]在SHREC 2015中组织了一次基于“标准形”的非刚性三维模型检索竞赛,提出并比较了基于改进的MDS嵌入、基于全局点特征、基于骨架驱动和基于欧氏距离等多种构造模型“标准形”的方法.实验结果显示:这些算法的检索效果相似,其中一些传统的基于MDS的方法比后来出现的算法检索准确率稍高一些,但是耗时较长.因此,如何选择还取决于用户个人的需求.

(4) 基于拓扑结构的特征提取

这类算法的主要思想就是将三维模型匹配问题转换成拓扑结构匹配问题,只要两个模型具有相同的拓扑结构,就认为它们是相似的.通常,三维模型的拓扑结构反映了模型各部分之间的连通性,可以有效地描述模型

的全局结构和局部细节,对旋转、拉伸、简化、细分及姿态形变等具有一定的稳定性.但是这类算法对拓扑噪声敏感,对模型的要求比较苛刻,且计算复杂度比较高.

表示三维模型拓扑结构的方法主要有两类:一类是基于 Reeb 图的方法,另一类是基于骨架图的方法.

Reeb 图通常是:定义一连续函数 μ ,计算模型上每个顶点处的 μ 函数值,将函数值相同且属于同一连通区域的顶点看作是等价的,用一个结点表示;进一步根据模型区域的连通性连接这些结点,就得到了一个可以表示原始模型拓扑结构的图,称为 Reeb 图.选择不同的函数 μ 会得到不同的 Reeb 图,进而会影响模型相似性比较的结果,目前, μ 常被设置为高度函数、Morse 函数等.Hilaga 等人^[59]采用测地距离和测地线构建 Morse 函数,提出了多分辨率 Reeb 图.Biasotti 等人^[60]提出扩展 Reeb 图来进行部分模型匹配.Barra 等人^[61]通过聚合扩展 Reeb 图簇的高斯核进行相似度比较.

骨架图是描述三维模型的拓扑结构特征的另一种重要方法.该方法在保持三维模型的拓扑特征、方向、其他局部特征的基础上,还可以直观、紧凑有效地表示三维模型.Sundar 等人^[62]使用参数细化算法构建模型的骨架图,并提出基于图匹配技术对模型的骨架图进行比较.Sfikas 等人^[63,64]基于离散化的保形因子构建非刚性三维模型的骨架图来应对模型形变,并且组合模型的几何特性来提高检索效果.

基于拓扑结构的特征提取方法对模型的拓扑噪声敏感.然而在三维模型扫描和建模过程中有很多因素会引起模型的拓扑结构发生变化,尤其是在非刚性三维模型上,由于铰链结构的存在,不同部分很容易相互接触,而相互接触的部分在扫描时会被连接为一个整体,从而改变了模型的拓扑结构.为了应对这一问题,有很多研究者致力于寻求拓扑鲁棒的模型特征提取算法,例如,2010 年,Bronstein 等人^[65]利用 Gromov-Hausdorff 体系中的扩散距离替代传统的测地距离去度量模型,从而提高了应对模型拓扑变化的鲁棒性.

值得一提的是,Löhner 等人^[32]在 SHREC 2016 中组织了具有拓扑噪声的非刚性三维模型匹配竞赛,并介绍了等距嵌入、格林函数嵌入对齐和随机森林这 3 种算法.2012 年,Sahillioglu 等人^[66]提出利用期望值最大化算法获取等距变换模型间的对应关系.等距嵌入算法在此基础上结合双调和距离实现拓扑鲁棒的模型匹配.格林函数嵌入对齐算法是指利用定义在流形上的拉普拉斯-贝尔特拉米(Laplace-Beltrami,简称 LB)算子的格林函数进行模型匹配.随机森林算法是指利用学习的方法获取模型的对应关系,首先,基于训练集获取决策树集合;然后,每一个决策树对测试模型分配一个定义在离散标签集合上的概率分布,标签标识着模型点之间的对应关系.实验结果表明:前两种算法会得到更高的最大匹配准确率,而随机森林算法在应对不同的拓扑变换时有着更稳定的匹配结果.然而,拓扑鲁棒的模型匹配问题还远未解决,仍需要人们进一步探索更为有效的方法^[32].

(5) 基于谱分析的特征提取

在计算机视觉领域,人们通常将模型表面看作是黎曼流形,对模型描述就是在黎曼空间内进行一系列操作^[67].在这一思想的指导下,研究者们开始探索如何将非刚性三维模型定义为一个流形函数,并利用流形函数的特征值和特征向量对模型进行描述,这种基于流形特征的方法也称为谱方法.基于谱方法的特征提取算法是建立在谱图理论^[68]基础上,利用一系列方法构建模型的 LB 算子^[69],通过 LB 算子的特征函数来描述模型的内蕴属性,适用于非刚性三维模型检索任务.这类方法也有其局限性:一是计算量比较大,如果对模型进行简化处理,又会影响检索的准确度;二是缺乏空间相关性信息,在增加空间信息的同时,又会加大计算复杂性以及空间开销;三是对形状上相似但属于不同类别的模型,区分能力很难达到要求.

2006 年,Reuter 等人^[70]最早基于谱分析提出了一种谱特征 ShapeDNA.首先定义了三维模型的 LB 算子,并利用有限元方法计算出特征值;然后对特征值序列由小到大排序并归一化;最后,将前 k 个特征值组成的向量作为模型特征描述符.其优点是表示简单,易于比较,具有尺度不变性和等距变换不变性,能够表达模型的内蕴属性,适合于非刚性三维模型检索.但是,该特征描述的是模型的全局信息,不适于模型的局部分析和局部匹配.

2007 年,Rustamov^[71]在 Lévy^[72]提出的离散的 LB 算子特征分解的基础上将模型转换到谱嵌入空间,提出了一种全局点特征(global point signature,简称 GPS).模型上给定点 p 的 GPS 特征是一个无限维的向量:

$$GPS(p) = \left(\frac{1}{\sqrt{\lambda_1}} \phi_1(p), \frac{1}{\sqrt{\lambda_2}} \phi_2(p), \frac{1}{\sqrt{\lambda_3}} \phi_3(p), \dots \right) \quad (6)$$

其中, λ_i 是第 i 个特征值, ϕ_i 是其对应的特征向量. 全局点特征能够对三维模型表面顶点的位置信息进行内蕴的描述, 而不依赖于任何外界的坐标系, 能够很好地应对姿势弯曲等非刚性变换. 但该方法也有一些局限性: 一是离散的 LB 算子缺乏在边界点的定义与证明; 二是 LB 算子无法对退化的模型生成特征描述符; 三是模型的嵌入空间的基不相同. Rustamov^[73] 在 2009 年进一步提出基于模板的特征描述符, 即随机抽取一些模型作为形状模板, 建立模型到这些形状模板的映射, 随后在形状模板中计算 GPS 特征. 这种方法虽然解决了嵌入空间不统一的问题, 但是却引入了附加计算及如何选取模板的问题.

2009 年, Smeets 等人^[74] 利用特征值分解和奇异值分解等数学方法对模型的测地距离矩阵进行谱分解, 其中, 分解得到的特征值和奇异值表达了模型的内蕴信息, 特征向量和奇异值向量表达了模型的对应关系. 将前 50 个最大的特征值和奇异值组成向量作为模型的特征描述符, 具有等距变换不变性. 该特征也有两点不足: 一是比较耗时; 二是由于计算的是模型表面点对之间的最短路径, 因此对模型的拓扑变换比较敏感.

将该特征与 meshSIFT 特征进行融合^[28], 利用最大最小归一化方法将各单一特征对应的距离矩阵归一化, 并对归一化后的距离矩阵加权求和, 用于相似度比较. 基于融合特征进行非刚性三维模型检索, 准确率会得到进一步的提高.

2009 年, Sun 等人^[75] 通过模拟模型表面热扩散^[76] 的过程提出热核特征描述符 (heat kernel signature, 简称 HKS). 黎曼流形上的热扩散方程可以表示为

$$h_t(x) = \sum_{i=0}^m e^{-\lambda_i t} \phi_i^2(x) \quad (7)$$

其中, λ_i 是第 i 个特征值, ϕ_i 是其对应的特征向量. $h_t(x)$ 被称为热核, 描述了流形曲面顶点 x 处经过 t 时间的热量扩散后剩余的热量值. HKS 特征刻画了模型的内蕴属性, 对非刚性三维模型的等距变换、拓扑噪音及局部失真等情形具有鲁棒性. 但 HKS 也有不足: 一是高频信息被抑制, 对模型的细节描述能力有限; 二是时间参数和对数离散化方法的选择缺乏充分的理论证明; 三是对尺度变换敏感. 2010 年, Bronstein 等人^[77] 利用傅里叶变换中幅值的时移不变性来使 HKS 特征具有尺度变换不变性, 提出了尺度无关的热核特征 (scale invariant heat kernel signature, 简称 SIHKS); Raviv 等人^[78] 设计了一种基于模型体系的热核特征提取算法, 通过深入模型内部, 获得其内部信息, 更充分地描述模型特征; 2011 年, Fang 等人^[79] 在热传导过程中量化模型表面的温度分布, 提出了热量均值描述符 (heat mean signature, 简称 HMS).

2011 年, Aubry 等人^[80] 在量子力学框架内利用波动方程对模型进行分析, 提出了波核特征 (wave kernel signature, 简称 WKS). WKS 被定义为对数能量尺度下的标量:

$$\begin{cases} \text{WKS}(x, \cdot): R \rightarrow R \\ \text{WKS}(x, e) = C_e \sum_i \phi_i^2(x) e^{-\frac{(e - \log \lambda_i)^2}{2\sigma^2}}, C_e = \left(\sum_i e^{-\frac{(e - \log \lambda_i)^2}{2\sigma^2}} \right)^{-1} \end{cases} \quad (8)$$

其中, λ_i 是第 i 个特征值, ϕ_i 是其对应的特征向量. WKS 特征表示量子粒子在模型给定顶点位置、不同能量尺度下的平均概率分布. 与 HKS 相比, WKS 存在很多优势: 一是 WKS 通过对频率而不是对时间进行参数化, 可以清晰地将不同频率以及不同空间尺度的信息分离开; 二是 WKS 能够捕获相当量的高频信息, 因此能够对模型形状进行更为细节的刻画^[81]; 三是通过稳定性理论分析选择合适的参数化, 能够更好地应对模型的非刚性形变, 检索效果更好.

2011 年, Hammond 等人^[82] 基于谱图小波生成核提出了谱图小波变换. 与傅里叶变换相比, 小波变换是一个时间和频域的局部变换, 通过平移和缩放等运算对函数进行多尺度细化分析, 能够有效地从信号中提取信息.

在此基础上, 2013 年, Li 等人^[83] 提出了谱图小波描述符 (spectral graph wavelet signature, 简称 SGWS), 用于非刚性三维模型分析和检索. SGWS 是一个基于三次样条生成核的多分辨率描述符, 可以捕获三维模型的全局和局部几何特征, 不仅具有等距变换不变性, 而且具有带通滤波器和低通滤波器的优点. 在 SHREC 2010^[26] 和 SHREC 2011^[28] 上的实验结果表明, SGWS 的检索准确率高于扩散几何框架内的光谱特征.

2.2 基于学习的特征描述符提取技术

基于人工设计的模型特征提取方法过程繁杂耗时,难以实现对大数据量的三维模型自动获取特征描述符.视觉码本是一种基于低层特征描述符进行机器学习的模型.它只是一种浅层学习模型,对三维模型的中层结构和高层语义信息的表示力度有限.深度学习的发展为解决这一问题提供了新的思路.自 2014 年开始,深度学习开始应用于三维模型特征提取,并取得了一系列突破性的成果.

(1) 基于视觉码本的提取算法

词袋模型(bag-of-words,简称 BoW)最早用于文本检索中,特征袋模型(bag-of-feature,简称 BoF)是它的一种改进,被应用到图像视频检索中.主要步骤为:首先提取低层特征,并对低层特征集合利用聚类算法生成词汇,构成词典;然后,利用词典对每一个样本进行特征量化,将词汇的频率分布作为特征描述符.与图像类比,BoF 同样可以用来描述三维模型,即将低层三维模型特征通过聚类方法生成码本,然后将三维模型表示成一组码本分布频率值的无序集合.算法流程如图 6 所示^[84].

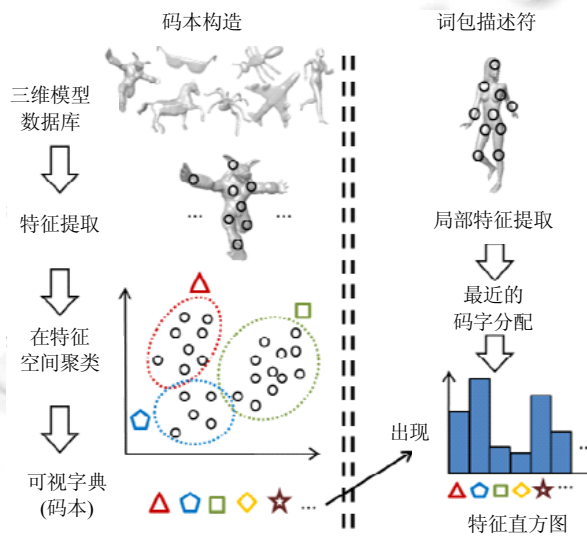


Fig.6 Flow chart of the BoF-based algorithm^[84]

图 6 基于特征袋模型的算法流程^[84]

2009 年,Ovsjanikov 等人^[85]提出了一种具有空间敏感性的特征袋模型,并将其应用到非刚性三维模型检索中.首先,利用 K -means 算法对提取的 HKS 特征进行聚类,获得几何词汇;然后,将距离相近的几何词汇联系起来构成几何短语;最后,同时统计几何词汇和几何短语出现的频率,隐性地引入空间关系到特征描述符中.

2011 年,Bronstein 等人^[5]在此基础上做了改进:在设定的训练集上,利用相似性敏感哈希算法对特征描述符之间的汉明距离进行学习得到最佳参数;然后,将模型的词袋特征映射到汉明空间得到新的特征描述符,称为 ShapeGoogle 特征.该方法通过对模型距离进行优化,不仅降低了特征描述符的维度,也提高了检索准确率.

2014 年,Litman 等人^[86]利用监督学习方法对传统的 BoF 框架进行了改进:首先,使用稀疏编码方式代替 K -means 聚类算法进行字典学习;然后,基于任务特定目标对字典进行双层优化.该方法几乎可以学习任何种类和不同强度的变换,不足之处在于迭代优化构建稀疏编码的时间比较长,因此还需要设计加速算法.

2014 年,Li 等人^[87]对 HKS,SIHKS,HMS 和 WKS 等应用于非刚性三维模型检索的谱特征描述符进行了详细的阐述和对比,着重研究将这些低层局部谱特征聚合成简洁高层整体特征的编码方法,并对比分析了编码过程中不同的码本尺寸等参数设置对检索结果的影响.实验结果表明:SIHKS 特征对非刚性三维模型的刻画分辨能力更强,同时,带有空间信息的编码方式能够更进一步地提高检索准确率.

2017 年,Han 等人^[88]提出了一种新的增强空间信息的三维模型表征方法,称为空间上下文相关性词袋模型

(bag of spatial context correlation,简称 BoSCC).该方法首先通过马尔可夫链以多尺度方式对每个顶点的空间上下文进行建模,然后基于模型表面局部区域内顶点的空间上下文之间的相关性对该区域的几何和空间信息同时编码,最后构建模型空间上下文相关模式频率直方图.BoSCC 能够解决传统的 BoW 框架在对三维模型特征编码时需要解决的 4 个主要问题,即网格分辨率不同、顶点拓扑不规则、表面取向不确定及刚性和非刚性形变等.在 McGill^[19],SHREC 2010^[26]等数据集上的实验结果表明,BoSCC 优于最先进的空间信息增强 BoW 方法.

(2) 基于深度学习的提取算法

现有的三维模型深度特征表示方法又可以分为基于人工特征的方法、基于投影图像的方法、基于三维体素的方法以及基于原始数据的方法^[89].

基于人工特征的方法^[89],其主要思想是对人工设计的低层特征采用深度学习框架自动学习出高层特征.即首先在三维模型上提取人工特征,进而将这些特征作为深度神经网络的输入,用于学习高层特征表示.其优势在于可以充分利用现有的低层特征描述符及深度学习模型;其局限是仍然依赖人工特征的选择与参数优化,在某种程度上损失了深度学习的优势,无法从根本上解决人工特征存在的问题.

2015年,Fang 等人^[90]提出了一种深度形状描述符(deep shape descriptor,简称 DeepSD).首先提取模型 HKS 特征;然后,在此基础上对模型顶点进行聚类,提出热量形状描述符(heat shape descriptor,简称 HeatSD);最后,将 HeatSD 特征集合作为深度神经网络的输入,经过学习得到 DeepSD,具体流程如图 7 所示.在 McGill^[19]和 ShapeGoogle^[5]数据集上进行实验,得到 DeepSD 特征的 PR 曲线,比 HeatSD 特征和 ShapeGoogle 特征的 PR 曲线有明显的提升.并且 DeepSD 特征对模型噪声、模型部分缺失也具有较好的鲁棒性和稳定性.

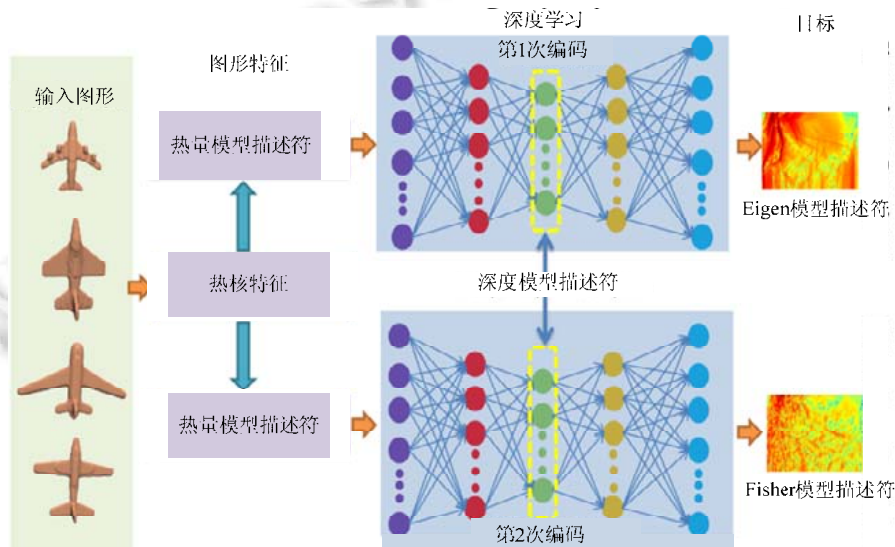


Fig.7 Process of learning deep shape descriptor^[90]

图 7 学习深度模型描述符流程^[90]

2015年,Xie 等人^[91]基于有识别力的深度自动编码器提出了一种高层模型特征学习架构,如图 8 所示.首先,提取多尺度的模型分布特征(如 HKS)作为自动编码器的输入;然后,对隐含层中的神经元施加 Fisher 判别标准自动学习模型的特征;最后,将所有自动编码器的隐含层中的神经元连接组成一个向量,作为模型的特征描述符用于匹配和检索.在 McGill^[19]和 ShapeGoogle^[5]数据集上进行实验,该特征的检索结果比传统特征有明显的提高,且对模型的拓扑变换具有很好的鲁棒性.

2016年,Ghodraty 等人^[92]提出了一种多级特征学习架构.首先提取三维模型的 SGWS 特征矩阵,矩阵的每一列描述了模型的局部特性;其次,利用 BoF 架构基于这些局部描述符构造中间层特征描述符,在 BoF 架构中,特征

编码方式为局部约束线性编码(locality-constrained linear coding,简称 LLC),同时引入双调和距离去度量 BoF 向量之间的空间关系;最后,通过深度自动编码器学习高层特征,具体流程如图 9 所示.在 SHREC 2014^[29]和 SHREC 2015^[30]标准数据集上的实验结果显示:高层学习特征的辨识能力与传统谱特征(ShapeDNA,WKS,SIHKS 等)和视觉码本特征相比有显著的提高,检索准确率接近于 1.多级特征学习架构提取了一个简洁、几何信息丰富且鲁棒的深度模型感知描述符,但参数学习的过程也增加了时耗,这一局限性在实时性要求较高时需要注意.

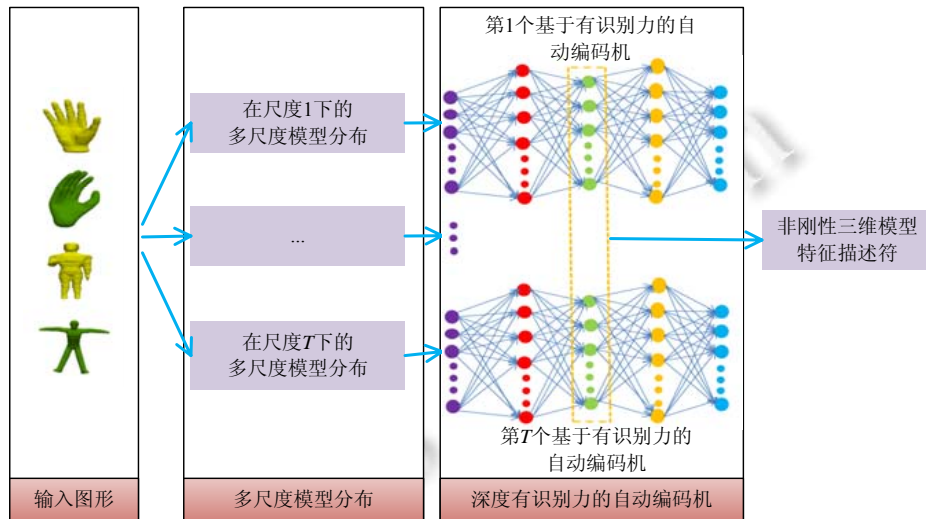


Fig.8 Framework of the learning shape descriptor based on discriminative auto-encoder^[91]

图 8 基于有识别力的自动编码器学习模型特征描述符的框架图^[91]

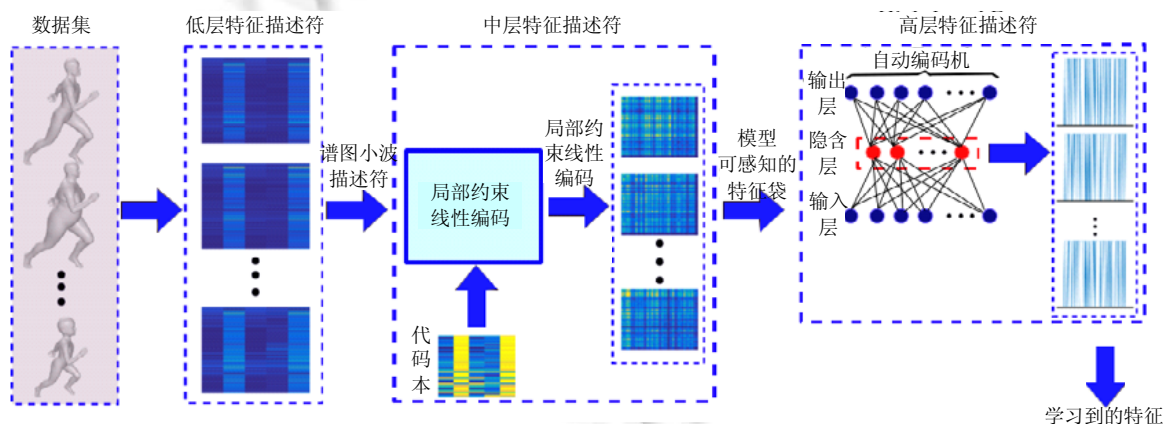


Fig.9 Diagram of the multi-level feature learning methodology^[92]

图 9 多层次特征学习方法示意图^[92]

2016年,Dai 等人^[93]提出一个新的学习框架,用以提取简洁的数据驱动模型特征描述符:首先提取 SIHKS 特征来描述模型的顶点;然后使用 LLC 对模型的每个顶点进行编码,形成全局模型表示;最后使用多对一编码器 (many-to-one encoder,简称 MOencoder)进行学习,并将隐含层作为模型的高层特征描述符.具体流程如图 10 所示.在 McGill^[19],ShapeGoogle^[5]和 SHREC 2014^[29]等标准数据集上进行了对比实验,结果显示:相比于传统的 SIHKS 特征,MOencoder 能够大幅度提高检索准确率.该方法可以强制地将相同类的模型从相同的目标输出,不同类的模型从不同的目标输出,因此,高层特征描述符面对非刚性三维模型的各种结构变化具有较强的鲁棒性.

基于投影图像的方法^[89],其主要思想是将三维模型转换成一系列二维图像,再采取深度学习提取图像特征作为最终的模型描述符.其优势在于,可以充分利用二维图像领域性能优越的网络架构,并且存在海量图像数据供深度学习模型进行预训练;其不足之处是变换过程本身改变了三维形状的局部和全局结构,降低了特征鉴别力.2016年,Sinha 等人^[94]提出了几何图像(geometry image)的概念,并利用卷积神经网络对几何图像学习,获取模型的内蕴特征,用于三维模型检索.具体过程为:首先将三维模型参数化到球形表面,进而将球形表面投影到八面体后展开成二维平面,并采用主曲率或 HKS 在平面的分布获得二维图像(称为几何图像),最后采用卷积神经网络从几何图像中学习特征表示.在 TOSCA^[20,21]和 SHREC 2011^[28]数据集上的实验结果显示:几何图像能够灵活地编码模型的几何属性和内蕴信息,具有等距变换不变性,相对于基于人工特征的深度学习体系,该方法所获取的特征表达能力更强,检索准确率也更高.

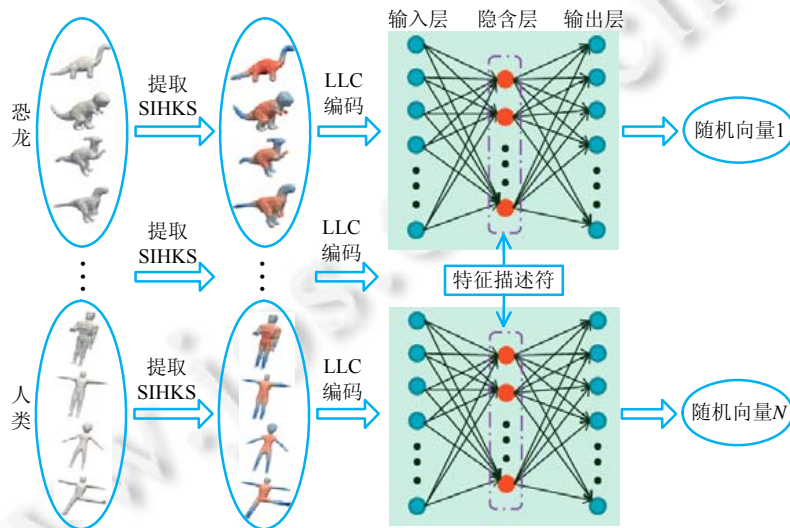


Fig. 10 Flow of the feature learning based on many-to-one encoder^[93]

图 10 基于多对一编码的特征学习流程图^[93]

基于三维体素的方法^[89],其主要思想是将三维模型进行体素化表示,再构建三维深度学习模型提取特征.这类方法的优势在于,三维体素完整地保留了三维模型的形状信息,有利于提高特征的鉴别力.其不足之处是,为了使网络训练不过于复杂,三维体素的分辨率不能太高,而较低的分辨率限制了所学特征的鉴别力;三维模型表面所占的体素比例不高,使得体素化结果较稀疏;相对于传统的学习网络,三维的深度学习架构的计算复杂度大幅度提高.Wu 等人^[95]将三维模型表示为二值概率分布,即若体素在三维表面内,则值为 1,否则为 0.并采用卷积深度置信网络学习三维体素与标签之间的联合分布.Li 等人^[96]将三维模型表示成体素场以克服三维体素表示的稀疏性问题,并采用场探索滤波器(field probing filter)取代卷积神经网络中的卷积层来学习特征.在 ModelNet40^[25]标准数据集上的实验结果显示:该方法在显著降低计算时间的同时,能够提取鲁棒的全局描述符.

基于原始数据的方法^[89],其主要思想是设计能适应原始三维数据特点的深度学习模型.Han 等人^[97]提出了一种网格卷积受限玻尔兹曼机(mesh convolutional restricted Boltzmann machine,简称 MCRBM)以实现三维模型的非监督特征学习.该方法首先在三维模型上均匀放置固定点数的节点,并在这些节点上用局部函数能量分布来表达三维形状局部区域的几何和结构信息,进而采用网格卷积深度置信网络(mesh convolutional deep belief network,简称 MCDBN)来学习高层特征.在 McGill^[19]和 SHREC 2007^[98]等数据集上应用该特征进行全局模型检索、部分模型检索和模型匹配,实验结果表明,该特征比传统的谱特征及视觉码本框架下的特征具有更强的辨识力.

上述方法从 4 种不同的角度将深度学习技术应用到三维模型数据上,均对本领域的研究者具有很重要的

启发意义.例如,为了提高模型检索算法的准确率,Bu 等人^[99]结合基于投影图像的方法和基于三维体素的方法提出了一种多模态数据特征学习架构,在 McGill^[19],SHREC 2011^[28]等数据集上的结果表明,组合算法的检索准确率高于任一单一算法.此外,深度学习技术在三维模型上的应用主要受限于处理速度和存储空间,并且需要应对模型数量级不断增加的情况.对于该现状,一项具有启发性的工作是,Furuya 等人^[100]构建了一种二进制的深度局部特征聚合网络(binanzed deep local feature aggregation network,简称BDLAN)架构.BDLAN以定向点集表示的三维模型作为输入,引入深度语义哈希算法,生成简洁的二进制向量作为模型特征.相对于实值特征向量,二进制编码不仅可以通过汉明距离进行非常快速的比较,而且可以减少存储空间,既能提高检索算法的时间效率,又能实现检索算法的可扩展性.

3 非刚性三维模型特征提取的主要难点和发展趋势

3.1 主要难点

- (1) 当一个非刚性三维模型处于不同姿态时,虽然从语义角度来看还是同一物体,但从几何角度看,模型已发生了较大的形变.扫描不同姿势下的非刚性三维模型,所获得的数据会有不同程度的噪声,如部分数据因遮挡而缺失、不同部分因粘连使得拓扑结构发生改变等,因此,处理非刚性三维模型需要弥补由铰链结构产生的自由度,同时还应满足对部分缺失、拓扑变换等形变具有较高的鲁棒性.
- (2) 结合多学科的知识,提取模型的内蕴属性.比如探索新的数学分析工具,观察自然物理现象,借鉴其他领域比如文本、图像上的思想,并将它们迁移应用到三维模型的分析 and 处理上.
- (3) 虽然已有的工作提出了不少方法,但是现有的特征描述符往往侧重于描述模型的某种性质,这使得这些特征描述符只在特定的数据集中有较好的效果,至今还没有任何一个特征描述符可以在所有的非刚性数据集中得到满意的结果.
- (4) 神经网络具有优异的特征学习能力,学习得到的特征对数据有更强的刻画能力.但是现有的神经网络一般适应于 2D 和 2.5D,对于三维模型来说,需要将模型按照一定的策略转换成二维的数据.因此,如何对非刚性三维模型进行投影或按照哪些标准选择低层特征,成为本方向的研究难点.
- (5) 对于大规模三维模型,尤其是高分辨率的三维模型数据集来说,应用神经网络提取特征的另一个难点是训练时间较长,即使使用 GPU 加速,也难以满足实时性的要求.因此,如何构造深度学习网络,以便能够在可接受的时间内训练三维模型,也成为亟待解决的问题.

3.2 发展趋势

- (1) 引入计算共形几何领域的相关理论提取非刚性三维模型的内蕴信息

实验结果表明:将热扩散理论、薛定谔波动方程、泊松方程、谱图小波变换等数学工具应用于非刚性三维模型分析上,可以获取模型的内蕴属性,对模型形变也具有一定的鲁棒性.因此,未来的工作仍然要探索和发现新的、更加有效的数学工具,并将其迁移应用到非刚性三维模型特征提取上来,以应对模型的各种非刚性形变.在众多的数学理论和分析工具中,计算共形几何^[101,102]这一新兴学科的出现,带给人们很多启发.

基于计算共形几何,所有三维几何处理的问题都可以转换为球面、欧氏平面和二维双曲空间这 3 种标准空间中的二维问题.因此,在共形几何理论指导下,将非刚性三维模型映射成平面几何图像,并保留其内蕴几何属性,进一步基于映射得到的图像进行检索,不仅可以简化问题的难度、提高效率,还可以促进计算机图形学和计算共形几何领域的发展.

- (2) 多种特征描述符组合的研究

随着计算机硬件设备的发展,三维模型搜索会成为未来搜索引擎的方向.互联网上流通的三维模型具有结构复杂、格式多样的特点,而单一的特征提取算法往往只对某些种类的模型有效.多特征融合可以突破单一特征的限制性.具体来说,多特征融合是指通过对模型设计多种不同的特征提取算法,从多个角度对模型进行描述,获取多种特征描述符,进一步通过对这些描述符的合理组合,实现特征互补,获得应用更广泛的特征描述符.

虽然从实验测试结果来看,融合特征比单一特征的效果提高很多,但是从物理意义和数学证明上很难对特征融合算法的合理性做出具有信服力的解释.而且从不同特征空间提取的特征具有不同的属性,对模型的表达角度和表达程度也有差异,每个特征都有其各自的适用环境和优缺点,因此,被融合的统一特征的选择以及在融合过程中权重的学习,也是有待解决的问题.

(3) 应用深度学习技术提取非刚性三维模型的特征

目前,应用深度学习技术处理三维模型是本领域的主流趋势,但是与图像数据不同,三维模型是非结构化的,不能直接输入到深度神经网络中.采用深度学习技术提取模型特征首先要解决三维模型的结构化表示问题^[89].针对该问题,有几种不同的解决方式:(1) 基于人工设计的低级特征,采用深度学习模型提取高级特征;(2) 将三维模型转化为系列二维图像,再采用深度学习模型提取图像特征;(3) 将三维模型体素化得到结构化表示,再构建三维深度学习模型提取特征;(4) 设计能够适应原始三维模型数据特点的深度学习模型.前 3 种方式在应用深度学习之前都需要对模型进行处理,而这一步骤不可避免地会丢失模型的部分信息,而且增加了模型的处理时间.因此,开发设计能够处理原始三维模型数据的神经网络结构更具有前沿性和挑战性.此外,在应用深度学习处理三维模型时,解决计算复杂性和参数选择的不确定性,也是本方向的一个研究趋势.

(4) 构建大规模的非刚性三维模型标准数据集

大数据是深度学习成功的基石.与 ImageNet^[103]等千万量级的二维图像数据集相比,现有的三维模型数据集很小,ModelNet^[25]与 ShapeNet^[24]两个数据集加和也仅仅约有 300 万个模型.较小的数据集会对深度学习模型的设计与训练有更高的要求.这一制约对非刚性三维模型来说尤为严重,目前还没有出现一个标准的大规模的非刚性三维模型数据集.因此,构建大规模的非刚性三维模型标准数据集,也是本领域发展的必然趋势.

4 结 语

三维模型特征提取技术的发展对于三维模型检索、匹配等工作至关重要,对计算机视觉和计算机图形学的发展有着重要的影响.具有铰链结构的非刚性三维模型比刚性三维模型更为复杂,对表征模型属性的特征描述符有更高的要求,需要具有等距变换不变性等特点.本文汇总统计了目前互联网上存在的标准的非刚性三维模型数据集及国际上通用的模型检索评价指标,重点回顾了近 10 年来出现的非刚性三维模型检索特征提取技术,将模型特征划归为两大类:人工设计的特征和通过学习获取的特征,并对每一类中的典型算法按照时间顺序进行了详细的说明、分析和比较.最后,立足于国内外的研究现状,综合阐述了非刚性三维模型检索特征提取技术的主要难点及未来的发展趋势.希望本文的工作能够对本领域及相关领域研究者有所帮助.

References:

- [1] Yang YB, Lin H, Zhu Q. Content-Based 3D model retrieval: A survey. Chinese Journal of Computers, 2004,27(10):1297-1310 (in Chinese with English abstract). [doi: 10.3321/j.issn:0254-4164.2004.10.001]
- [2] Zhang XD. Research on shape-based feature extraction method of 3D model [Ph.D. Thesis]. Beijing: China University of Petroleum, 2010 (in Chinese with English abstract).
- [3] Tangelder JW, Veltkamp RC. A survey of content based 3D shape retrieval methods. Multimedia Tools and Applications, 2008, 39(3):441-471. [doi: 10.1007/s11042-007-0181-0]
- [4] Biasotti S, Cerri A, Bronstein A, Bronstein M. Recent trends, applications, and perspectives in 3D shape similarity assessment. Computer Graphics Forum, 2016,35(6):87-119. [doi: 10.1111/cgf.12734]
- [5] Bronstein AM, Bronstein MM, Guibas LJ, Ovsjanikov M. Shape Google: Geometric words and expressions for invariant shape retrieval. ACM Trans. on Graphics (TOG), 2011,30(1):623-636. [doi: 10.1145/1899404.1899405]
- [6] Goldfeder C, Allen P. Autotagging to improve text search for 3D models. In: Proc. of the 8th ACM/IEEE-CS Joint Conf. on Digital Libraries. 2008. 355-358. [doi: 10.1145/1378889.1378950]
- [7] Fisher M, Hanrahan P. Context-Based search for 3D models. ACM Trans. on Graphics (TOG), 2010,29(6):81-95. [doi: 10.1145/1882261.1866204]

- [8] Funkhouser T, Min P, Kazhdan M, Chen J, Halderman A, Dobkin D. A search engine for 3D models. *ACM Trans. on Graphics (TOG)*, 2003,22(1):83–105. [doi: 10.1145/588272.588279]
- [9] Yang Y, Lin H, Zhang Y. Content-Based 3-D model retrieval: A survey. *IEEE Trans. on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 2007,37(6):1081–1098. [doi: 10.1109/TSMCC.2007.905756]
- [10] Bronstein AM, Bronstein MM, Kimmel R. *Numerical Geometry of Non-Rigid Shapes*. New York: Springer-Verlag, 2009. [doi: 10.1007/978-0-387-73301-2]
- [11] 3D model search engine. <http://shape.cs.princeton.edu>
- [12] Min P, Halderman JA, Kazhdan M, Funkhouser TA. Early experiences with a 3D model search engine. In: *Proc. of the Web3D Symp.* 2003. 7–18. [doi: 10.1145/636593.636595]
- [13] 3D model similarity search engine. <http://merkur01.inf.uni-konstanz.de/CCCC>
- [14] Vranic DV. An improvement of rotation invariant 3D-shape based on functions on concentric spheres. In: *Proc. of the Int'l Conf. on Image Processing*. 2003. 757–760. [doi: 10.1109/ICIP.2003.1247355]
- [15] Ogden IV system. <http://www.nime.ac.jp/~motofumi/Ogden/>
- [16] Suzuki MT. A search engine for polygonal models to support development of 3D e-learning applications. In: *Proc. of the 10th Int'l Conf. on the World Wide Web*. 2001. 182–183.
- [17] Boyer E, Bronstein AM, Bronstein MM, Bustos B, Darom T, Horaud R, Hotz I, Keller Y, Keustermans J, Kovnatsky A, Litman R, Reininghaus J, Sipiran I, Smeets D, Suetens P, Vandermeulen D, Zaharescu A, Zobel V. SHREC 2011: Robust feature detection and description benchmark. In: *Proc. of the 4th Eurographics Conf. on 3D Object Retrieval*. 2011. 71–78. [doi: 10.2312/3DOR/3DOR11/071-078]
- [18] Lian Z, Godil A, Bustos B, Daoudi M, Hermans J, Kawamura S, Kurita Y, Lavoue G, Nguyen HV, Ohbuchi R, Ohkita Y, Ohishi Y, Porikli F, Reuter M, Sipiran I, Smeets D, Suetens P, Tabia H, Vandermeulen D. A comparison of methods for non-rigid 3D shape retrieval. *Pattern Recognition*, 2013,46(1):449–461. [doi: 10.1016/j.patcog.2012.07.014]
- [19] Siddiqi K, Zhang J, Macrini D, Shokoufandeh A, Bouix S, Dickinson S. Retrieving articulated 3-D models using medial surfaces. *Machine Vision and Applications*, 2008,19(4):261–275. [doi: 10.1007/s00138-007-0097-8]
- [20] Bronstein AM, Bronstein MM, Kimmel R. Efficient computation of isometry-invariant distances between surfaces. *Siam Journal on Scientific Computing*, 2006,28(5):1812–1836. [doi: 10.1137/050639296]
- [21] Bronstein AM, Bronstein MM, Kimmel R. *Numerical Geometry of Non-Rigid Shapes*. New York: Springer-Verlag, 2009. [doi: 10.1007/978-0-387-73301-2]
- [22] Summer RW, Popović J. Deformation transfer for triangle meshes. *ACM Trans. on Graphics (TOG)*, 2004,23(3):399–405. [doi: 10.1145/1015706.1015736]
- [23] Eurographics Workshop on 3D Object Retrieval. <http://www.shrec.net/>
- [24] ShapeNet. <https://www.shapenet.org/>
- [25] ModelNet. <http://modelnet.cs.princeton.edu/>
- [26] Lian Z, Godil A, Fabry T, Furuya T, Hermans J, Ohbuchi R, Shu C, Smeets D, Suetens P, Vandermeulen D, Wuhler S. SHREC 2010 track: Non-Rigid 3D shape retrieval. In: *Proc. of the 3rd Eurographics Conf. on 3D Object Retrieval*. 2010. 101–108. [doi: 10.2312/3DOR/3DOR10/101-108]
- [27] Bronstein AM, Bronstein MM, Castellani U, Falcidieno B, Fusiello A, Godil A, Guibas LJ, Kokkinos I, Lian Z, Ovsjanikov M, Patané G, Spagnuolo M, Toldo R. SHREC 2010: Robust large-scale shape retrieval benchmark. In: *Proc. of the 3rd Eurographics Conf. on 3D Object Retrieval*. 2010. 71–78. [doi: 10.2312/3DOR/3DOR10/071-078]
- [28] Lian Z, Godil A, Bustos B, Daoudi M, Hermans J, Kawamura S, Kurita Y, Lavoué G, Nguyen HV, Ohbuchi R, Ohkita Y, Ohishi Y, Porikli F, Reuter M, Sipiran I, Smeets D, Suetens P, Tabia H, Vandermeulen D. SHREC 2011 track: Shape retrieval on non-rigid 3D watertight meshes. In: *Proc. of the 4th Eurographics Conf. on 3D Object Retrieval*. 2011. 79–88. [doi: 10.2312/3DOR/3DOR11/079-088]
- [29] Pickup D, Sun X, Rosin PL, Martin RR, Cheng Z, Lian Z, Aono M, Hamza AB, Bronstein A, Bronstein M, Bu S, Castellani U, Cheng S, Garro V, Giachetti A, Godil A, Isaia L, Han J, Johan H, Lai L, Li B, Li C, Li H, Litman R, Liu X, Liu Z, Lu Y, Sun L,

- Tam G, Tatsuma A, Ye J. Shape retrieval of non-rigid 3D human models. In: Proc. of the 7th Eurographics Conf. on 3D Object Retrieval. 2014. 101–110. [doi: 10.2312/3dor.20141056]
- [30] Lian Z, Zhang J, Choi S, ElNaghy H, El-Sana J, Furuya T, Giachetti A, Guler RA, Lai L, Li C, Li H, Limberger FA, Martin R, Nakanishi RU, Neto AP, Nonato LG, Ohbuchi R, Pevzner K, Pickup D, Rosin P, Sharf A, Sun L, Sun X, Tari S, Una G, Wilson RC. SHREC 2015 track: Non-Rigid 3D shape retrieval. In: Proc. of the 8th Eurographics Conf. on 3D Object Retrieval. 2015. 107–120. [doi: 10.2312/3dor.20151064]
- [31] Pickup D, Sun X, Rosin PL, Martin RR, Cheng Z, Nie S, Jin L. SHREC 2015 track: Canonical forms for non-rigid 3D shape retrieval. In: Proc. of the 8th Eurographics Conf. on 3D Object Retrieval. 2015. 1–8. [doi: 10.2312/3dor.20151063]
- [32] Lähner Z, Rodolà E, Bronstein MM, Cremers D, Burghard O, Cosmo L, Dieckmann A, Klein R, Sahillioglu Y. SHREC 2016: Matching of deformable shapes with topological noise. In: Proc. of the 9th Eurographics Conf. on 3D Object Retrieval. 2016. [doi: 10.2312/3dor.20161088]
- [33] Cosmo L, Bronstein MM, Torsello A, Rodolà E, Cremers D, Sahillioglu Y. Partial matching of deformable shapes. In: Proc. of the 9th Eurographics Conf. on 3D Object Retrieval. 2016. 61–67. [doi: 10.2312/3dor.20161089]
- [34] Limberger FA, Wilson RC, Aono M, Audebert N, Boulch A, Bustos B, Giachetti A, Godil A, Saux B, Li B, Lu Y, Nguyen HD, Nguyen VT, Pham VK, Sipiran I, Tatsuma A, Tran MT, Velasco-Forero S. SHREC 2017 track: Point-Cloud shape retrieval of non-rigid toys. In: Proc. of the 9th Eurographics Conf. on 3D Object Retrieval. 2017. 1–11. [doi: 10.2312/3dor.20171056]
- [35] The Princeton shape benchmark. <http://shape.cs.princeton.edu/benchmark/>
- [36] Shilane P, Min P, Kazhdan M, Funkhouser T. The Princeton shape benchmark. In: Proc. of the Conf. on the Shape Modeling Applications 2004. 2004. 167–178. [doi: 10.1109/SMI.2004.1314504]
- [37] Li PJ. Research on feature extraction algorithms for 3D model retrieval [Ph.D. Thesis]. Beijing: Beijing University of Posts and Telecommunications, 2013 (in Chinese with English abstract).
- [38] Osada R, Funkhouser T, Chazelle B, Dobkin D. Shape distributions. *ACM Trans. on Graphics (TOG)*, 2002,21(4):807–832. [doi: 10.1145/571647.571648]
- [39] Mahmoudi M, Sapiro G. Three-Dimensional point cloud recognition via distributions of geometric distances. *Graphical Models*, 2008,71(1):22–31. [doi: 10.1016/j.gmod.2008.10.002]
- [40] Ruggeri MR, Patané G, Spagnuolo M, Saupe D. Spectral-Driven isometry-invariant matching of 3D shapes. *Int'l Journal of Computer Vision*, 2010,89(2):248–265. [doi: 10.1007/s11263-009-0250-0]
- [41] Sipiran I, Bustos B. Harris 3D: A robust extension of the Harris operator for interest point detection on 3D meshes. *The Visual Computer*, 2011,27(11):963–976. [doi: 10.1007/s00371-011-0610-y]
- [42] Pan X, You Q, Liu Z, Chen QH. 3D shape retrieval by poisson histogram. *Pattern Recognition Letters*, 2011,32(6):787–794. [doi: 10.1016/j.patrec.2011.01.003]
- [43] Ohkita Y, Ohishi Y, Furuya T, Ohbuchi R. Non-Rigid 3D model retrieval using set of local statistical features. In: Proc. of the 13th IEEE Int'l Conf. on the Multimedia and Expo Workshops (ICMEW). 2012. 593–598. [doi: 10.1109/ICMEW.2012.109]
- [44] Lowe DG. Object recognition from local scale-invariant features. In: Proc. of the 3rd Int'l Conf. on Computer Vision. 1999. 1150. [doi: 10.1109/ICCV.1999.790410]
- [45] Castellani U, Cristani M, Fantoni S, Murino V. Sparse points matching by combining 3D mesh saliency with statistical descriptors. *Computer Graphics Forum*, 2008,27(2):643–652. [doi: 10.1111/j.1467-8659.2008.01162.x]
- [46] Castellani U, Cristani M, Murino V. Statistical 3D shape analysis by local generative descriptors. *IEEE Trans. on Pattern Analysis & Machine Intelligence*, 2011,33(12):2555–2560. [doi: 10.1109/TPAMI.2011.85]
- [47] Zaharescu A, Boyer E, Varanasi K, Horaud R. Surface feature detection and description with applications to mesh matching. In: Proc. of the 22nd IEEE Conf. On CVPR. 2009. 373–380. [doi: 10.1109/CVPR.2009.5206748]
- [48] Maes C, Fabry T, Keustermans J, Smeets D, Suetens P, Vandermeulen D. Feature detection on 3D face surfaces for pose normalisation and recognition. In: Proc. of the 4th Int'l Conf. on Biometrics: Theory Applications and Systems. IEEE, 2010. 1–6. [doi: 10.1109/BTAS.2010.5634543]
- [49] Johnson AE, Hebert M. Using spin images for efficient object recognition in cluttered 3D scenes. *IEEE Trans. on Pattern Analysis & Machine Intelligence*, 1999,21(5):433–449. [doi: 10.1109/34.765655]

- [50] Liu Y, Zha H, Qin H. Shape topics: A compact representation and new algorithms for 3D partial shape retrieval. In: Proc. of the 19th IEEE Conf. on CVPR. 2006. 2025–2032. [doi: 10.1109/CVPR.2006.278]
- [51] Wang XL, Liu Y, Zha H. Intrinsic spin images: A subspace decomposition approach to understanding 3D deformable shapes. In: Proc. of the Conf. on 3DPVT, Vol.10. 2010. 17–20.
- [52] Ohbuchi R, Osada K, Furuya T, Banno T. Salient local visual features for shape-based 3D model retrieval. In: Proc. of the Int'l Conf. on Shape Modeling and Applications. 2008. 93–102. [doi: 10.1109/SMI.2008.4547955]
- [53] Furuya T, Ohbuchi R. Dense sampling and fast encoding for 3D model retrieval using bag-of-visual features. In: Proc. of the ACM Int'l Conf. on Image and Video Retrieval. 2009. 1–8. [doi: 10.1145/1646396.1646430]
- [54] Elad A, Kimmel R. On bending invariant signatures for surfaces. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2003, 25(10):1285–1295. [doi: 10.1109/TPAMI.2003.1233902]
- [55] Lian Z, Godil A, Sun X, Zhang H. Non-Rigid 3D shape retrieval using multidimensional scaling and bag-of-features. In: Proc. of the IEEE Int'l Conf. on Image Processing. 2010. 3181–3184. [doi: 10.1109/ICIP.2010.5654226]
- [56] Lian Z, Godil A, Xiao J. Feature-Preserved 3D canonical form. *Int'l Journal of Computer Vision*, 2013, 102(1):221–238. [doi: 10.1007/s11263-012-0548-1]
- [57] Pickup D, Sun X, Rosin PL, Martin RR. Euclidean-Distance-Based canonical forms for non-rigid 3D shape retrieval. *Pattern Recognition*, 2015, 46(8):2500–2512. [doi: 10.1016/j.patcog.2015.02.021]
- [58] Ben-Chen M, Gotsman C. Characterizing shape using conformal factors. In: Proc. of the 1st Eurographics Conf. on 3D Object Retrieval. 2008. 1–8. [doi: 10.2312/3DOR/3DOR08/001-008]
- [59] Hilaga M, Shinagawa Y, Kohmura T, Kunii TL. Topology matching for fully automatic similarity estimation of 3D shapes. In: Proc. of the 28th annual Conf. on Computer Graphics and Interactive Techniques. 2001. 203–212. [doi: 10.1145/383259.383282]
- [60] Biasotti S, Marini S, Mortara M, Patanè G, Spagnuolo M, Falcidieno B. 3D shape matching through topological structures. *LNCS*, 2003, 2886:194–203. [doi: 10.1007/978-3-540-39966-7_18]
- [61] Barra V, Biasotti S. 3D shape retrieval using kernels on extended Reeb graphs. *Pattern Recognition*, 2013, 46(11):2985–2999. [doi: 10.1016/j.patcog.2013.03.019]
- [62] Sundar H, Silver D, Gagvani N, Dickinson S. Skeleton based shape matching and retrieval. In: Proc. of the Int'l Conf. on Shape Modeling. 2003. 130–139. [doi: 10.1109/SMI.2003.1199609]
- [63] Sfikas K, Pratikakis I, Theoharis T. ConTopo: Non-Rigid 3D object retrieval using topological information guided by conformal factors. In: Proc. of the 1st Eurographics Conf. on 3D Object Retrieval. 2011. 25–32. [doi: 10.2312/3DOR/3DOR11/025-032]
- [64] Sfikas K, Theoharis T, Pratikakis I. Non-Rigid 3D object retrieval using topological information guided by conformal factors. *The Visual Computer*, 2012, 28(9):943–955. [doi: 10.1007/s00371-012-0714-z]
- [65] Bronstein AM, Bronstein MM, Kimmel R, Mahmoudi M, Sapiro G. A gromov-hausdorff framework with diffusion geometry for topologically-robust non-rigid shape matching. *Int'l Journal of Computer Vision*, 2010, 89(2-3):266–286. [doi: 10.1007/s11263-009-0301-6]
- [66] Sahillioglu Y, Yemez Y. Minimum-Distortion isometric shape correspondence using EM algorithm. *IEEE Trans. on Pattern Analysis & Machine Intelligence*, 2012, 34(11):2203–2215. [doi: 10.1109/TPAMI.2012.26]
- [67] Kuang ZZ. Research on non-rigid-body model retrieval based on content [Ph.D. Thesis]. Shanghai: China University of Petroleum (East China), 2013 (in Chinese with English abstract).
- [68] Spielman DA. Spectral graph theory and its applications. In: Proc. of the 48th Annual IEEE Symp. on Foundations of Computer Science. 2007. 29–38. [doi:10.1109/FOCS.2007.56]
- [69] Reuter M, Wolter FE, Peinecke N. Laplace-Spectra as fingerprints for shape matching. In: Proc. of the ACM Int'l Conf. on Symp. on Solid and Physical Modeling. 2005. 101–106. [doi: 10.1145/1060244.1060256]
- [70] Reuter M, Wolter FE, Peinecke N. Laplace-Beltrami spectra as 'Shape-DNA' of surfaces and solids. *Computer-Aided Design*, 2006, 38(4):342–366. [doi: 10.1016/j.cad.2005.10.011]
- [71] Rustamov RM. Laplace-Beltrami eigenfunctions for deformation invariant shape representation. In: Proc. of the Eurographics Symp. on Geometry Processing. 2007. 225–233. [doi: 10.1145/1281991.1282022]

- [72] Lévy B. Laplace-Beltrami eigenfunctions towards an algorithm that understands geometry. In: Proc. of the Int'l Conf. on Shape Modeling and Applications. 2006. 13–13. [doi: 10.1109/SMI.2006.21]
- [73] Rustamov RM. Template based shape descriptor. In: Proc. of the 2nd Eurographics Conf. on 3D Object Retrieval. 2009. 1–7. [doi: 10.2312/3DOR/3DOR09/001-007]
- [74] Smeets D, Fabry T, Hermans J, Vandermeulen D, Suetens P. Isometric deformation modelling for object recognition. In: Proc. of the Int'l Conf. on Computer Analysis of Images and Patterns. Springer-Verlag, 2009. 757–765. [doi: 10.1007/978-3-642-03767-2_92]
- [75] Sun J, Ovsjanikov M, Guibas L. A concise and provably informative multi-scale signature based on heat diffusion. Computer Graphics Forum, 2009,28(5):1383–1392. [doi: 10.1111/j.1467-8659.2009.01515.x]
- [76] Jones PW, Schul R. Manifold parametrizations by eigenfunctions of the Laplacian and heat kernels. Proc. of the National Academy of Sciences of the United States of America, 2008,105(6):1803–1808. [doi: 10.1073/pnas.0710175104]
- [77] Bronstein MM, Kokkinos I. Scale-Invariant heat kernel signatures for non-rigid shape recognition. In: Proc. of the 23rd IEEE Conf. On CVPR. 2010. 1704–1711. [doi: 10.1109/CVPR.2010.5539838]
- [78] Raviv D, Bronstein MM, Bronstein AM, Kimmel R. Volumetric heat kernel signatures. In: Proc. of the ACM Workshop on 3D Object Retrieval. 2010. 39–44. [doi: 10.1145/1877808.1877817]
- [79] Fang Y, Sun M, Ramani K. Temperature distribution descriptor for robust 3D shape retrieval. In: Proc. of the 24th IEEE Conf. On CVPR Workshops. 2011. 9–16. [doi: 10.1109/CVPRW.2011.5981684]
- [80] Aubry M, Schlickewei U, Cremers D. The wave kernel signature: A quantum mechanical approach to shape analysis. In: Proc. of the IEEE Int'l Conf. on ICCV Workshops. 2011. 1626–1633. [doi: 10.1109/ICCVW.2011.6130444]
- [81] Bronstein AM. Spectral descriptors for deformable shapes. arXiv preprint arXiv: 1110.5015, 2011.
- [82] Hammond DK, Vandeheyne P, Gribonval R. Wavelets on graphs via spectral graph theory. Applied and Computational Harmonic Analysis, 2011,30(2):129–150. [doi: 10.1016/j.acha.2010.04.005]
- [83] Li C, Hamza AB. A multiresolution descriptor for deformable 3D shape retrieval. The Visual Computer, 2013,29(6-8):513–524. [doi: 10.1007/s00371-013-0815-3]
- [84] Laga H, Schreck T, Ferreira A, Godil A, Pratikakis I, Velkamp R. Bag of words and local spectral descriptor for 3D partial shape retrieval. In: Proc. of the 4th Eurographics Conf. on 3D Object Retrieval. 2011. 41–48. [doi: 10.2312/3DOR/3DOR11/041-048]
- [85] Ovsjanikov M, Ovsjanikov M, Bronstein AM, Bronstein MM, Guibas LJ. Shape google: A computer vision approach to isometry invariant shape retrieval. In: Proc. of the IEEE Conf. On ICCV Workshops. 2009. 320–327. [doi: 10.1109/ICCVW.2009.5457682]
- [86] Litman R, Bronstein A, Bronstein M, Castellani U. Supervised learning of bag-of-features shape descriptors using sparse coding. In: Proc. of the Symp. on Geometry Processing. 2014. 127–136. [doi: 10.1111/cgf.12438]
- [87] Li C, Hamza AB. Spatially aggregating spectral descriptors for nonrigid 3D shape retrieval: A comparative survey. Multimedia Systems, 2014,20(3):253–281. [doi: 10.1007/s00530-013-0318-0]
- [88] Han Z, Liu Z, Vong CM, Liu YS, Bu S, Han J, Chen CLP. BoSCC: Bag of spatial context correlations for spatially enhanced 3D shape representation. IEEE Trans. on Image Processing, 2017,26(8):3707–3720. [doi: 10.1109/TIP.2017.2704426]
- [89] Guo YL. Depth feature representation of 3D shape data. CCF Computer Visual Panel, 2017,(2):8–11 (in Chinese with English abstract).
- [90] Fang Y, Xie J, Dai G, Wang M, Fan Z, Xu T, Wang E. 3D deep shape descriptor. In: Proc. of the 28th IEEE Conf. On CVPR. 2015. 2319–2328. [doi: 10.1109/CVPR.2015.7298845]
- [91] Xie J, Fang Y, Zhu F, Wong E. Deepshape: Deep learned shape descriptor for 3D shape matching and retrieval. In: Proc. of the 28th IEEE Conf. on CVPR. 2015. 1275–1283. [doi: 10.1109/CVPR.2015.7298732]
- [92] Ghodrati H, Hamza AB. Deep shape-aware descriptor for nonrigid 3D object retrieval. Int'l Journal of Multimedia Information Retrieval, 2016,5(3):151–164. [doi: 10.1007/s13735-016-0103-x]
- [93] Dai G, Xie J, Zhu F, Fang Y. Learning a discriminative deformation-invariant 3D shape descriptor via many-to-one encoder. Pattern Recognition Letters, 2016,83(3):330–338. [doi: 10.1016/j.patrec.2016.04.005]
- [94] Sinha A, Bai J, Ramani K. Deep learning 3D shape surfaces using geometry images. In: Proc. of the Conf. on ECCV. 2016. 223–240. [doi: 10.1007/978-3-319-46466-4_14]

- [95] Wu Z, Song S, Khosla A, Yu F, Zhang L, Tang X, Xiao J. 3D shapenets: A deep representation for volumetric shapes. In: Proc. of the 28th IEEE Conf. on CVPR. 2015. 1912–1920. [doi: 10.1109/CVPR.2015.7298801]
- [96] Li Y, Pirk S, Su H, Qi CR, Guibas LJ. FPNN: Field probing neural networks for 3D data. In: Proc. of the Advances in Neural Information Processing Systems. 2016. 307–315.
- [97] Han Z, Liu Z, Han J, Vong CM, Bu S, Chen CLP. Mesh convolutional restricted Boltzmann machines for unsupervised learning of features with structure preservation on 3-D meshes. IEEE Trans. on Neural Networks & Learning Systems, 2017,28(10):2268–2281. [doi: 10.1109/TNNLS.2016.2582532]
- [98] Giorgi D, Biasotti S, Paraboschi L. Shape retrieval contest 2007: Watertight models track. SHREC Competition, 2007,8(7):1–11.
- [99] Bu S, Wang L, Han P, Liu Z, Li K. 3D shape recognition and retrieval based on multi-modality deep learning. Neurocomputing, 2017,259(11):183–193. [doi: 10.1016/j.neucom.2016.06.088]
- [100] Furuya T, Ohbuchi R. Deep semantic Hashing of 3D geometric features for efficient 3D model retrieval. In: Proc. of the Int'l Conf. On Computer Graphics. 2017. 1–6. [doi: 10.1145/3095140.3095148]
- [101] Gu XD. Computational Conformal Geometry. Somerville: Int'l Press, 2008.
- [102] Gu XF, Lei N. Introduction to conformal geometry. College Mathematics, 2016,32(3):1–13 (in Chinese with English abstract). [doi: 10.3969/j.issn.1672-1454.2016.03.001]
- [103] ImageNet. <http://www.image-net.org/>

附中文参考文献:

- [1] 杨育彬,林琚,朱庆.基于内容的三维模型检索综述.计算机学报,2004,27(10):1297–1310. [doi: 10.3321/j.issn:0254-4164.2004.10.001]
- [2] 张晓东.三维模型的形状特征提取方法研究[博士学位论文].北京:中国石油大学,2010. [doi: 10.1007/s11042-007-0181-0].
- [37] 李朋杰.面向三维模型检索的特征提取算法研究[博士学位论文].北京:北京邮电大学,2013.
- [67] 匡振中.基于内容的非刚体模型检索的研究[博士学位论文].上海:中国石油大学(华东),2013.
- [89] 郭裕兰.三维形状数据的深度特征表示.CCF 计算机视觉专委简报,2017,(2):8–11.
- [102] 顾险峰,雷娜.计算共形几何简介.大学数学,2016,32(3):1–13. [doi: 10.3969/j.issn.1672-1454.2016.03.001]



李海生(1974—),男,山东宁津人,博士,教授,CCF 高级会员,主要研究领域为计算机图形学,可视化.



吴晓群(1984—),女,博士,讲师,CCF 专业会员,主要研究领域为计算机图形学.



孙莉(1992—),女,硕士生,CCF 学生会员,主要研究领域为计算机图形学.



蔡强(1969—),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为计算机图形学,智能信息处理.



武玉娟(1992—),女,硕士生,CCF 学生会员,主要研究领域为计算机图形学.



杜军平(1963—),女,博士,教授,博士生导师,CCF 杰出会员,主要研究领域为智能信息处理,人工智能.