学 校 代 码 10459

学号或申请号

密 级



博 士 学 位 论 文

集体行为

作 者 姓 名：

导 师 姓 名：

学 科 门 类： 工 科

专 业 名 称： 软件工程

培 养 院 系： 信息工程学院

完 成 时 间： 2017.03

A thesis submitted to

Zhengzhou University

for the degree of doctor

**Research on the Recognition and Simulation of**

**Collective Behavior**

Software Engineering

School of Information Engineering

Mar. 2017

**原创性声明**

本人郑重声明：所呈交的学位论文，是本人在导师的指导下，独立进行研究所取得的成果。除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的科研成果。对本文的研究作出重要贡献的个人和集体，均已在文中以明确方式标明。本声明的法律责任由本人承担。

学位论文作者： 日期： 年 月 日

**学位论文使用授权声明**

本人在导师指导下完成的论文及相关的职务作品，知识产权归属郑州大学。根据郑州大学有关保留、使用学位论文的规定，同意学校保留或向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅；本人授权郑州大学可以将本学位论文的全部或部分编入有关数据库进行检索，可以采用影印、缩印或者其他复制手段保存论文和汇编本学位论文。本人离校后发表、使用学位论文或与该学位论文直接相关的学术论文或成果时，第一署名单位仍然为郑州大学。保密论文在解密后应遵守此规定。

学位论文作者 日期： 年 月 日

摘 要

集体行为是连续、有序的个体呈现出的宏观行为模式，广泛地存在于细菌菌落、动物群、人群、车流等各种群体系统中。

**关键字：** 集体行为，密度估计

Abstract

Collective behavior refers to macroscopic patterns consisted of continuous and ordered agents, which exists widely in diverse crowd systems, such as bacterial colony, animals group, human crowd and traffic flow.

**Keywords:** Collective Behavior, Density Estimation

目录

[摘 要 I](#_Toc477443116)

[Abstract III](#_Toc477443117)

[目录 VI](#_Toc477443118)

[图目录 X](#_Toc477443119)

[表目录 XIII](#_Toc477443120)

[第 1 章 绪论 1](#_Toc477443121)

[1.1 研究背景 1](#_Toc477443122)

[1.2 研究现状 2](#_Toc477443123)

[1.2.1 交叉学科中的集体行为研究 2](#_Toc477443124)

[1.2.2 集体行为识别 3](#_Toc477443125)

[1.2.3 集体行为仿真 5](#_Toc477443126)

[1.3 研究内容及意义 7](#_Toc477443127)

[1.4 论文章节安排 8](#_Toc477443128)

[第 2 章 基于一致性密度聚类的集体行为识别算法 10](#_Toc477443129)

[2.1 引言 10](#_Toc477443130)

[2.2 背景知识 12](#_Toc477443131)

[2.3 问题描述 13](#_Toc477443132)

[2.4 基于一致性密度聚类的集体行为识别方法 14](#_Toc477443133)

[2.4.1 一致性密度估计 14](#_Toc477443134)

[2.4.2 一致性密度聚类算法 16](#_Toc477443135)

[2.4.3 一致性合并算法 18](#_Toc477443136)

[2.4.4 算法时间复杂度分析 19](#_Toc477443137)

[2.5 实验结果 20](#_Toc477443138)

[2.6 实验对比及分析 21](#_Toc477443139)

[2.6.1 与密度峰值聚类算法的对比实验 21](#_Toc477443140)

[2.6.2 与其它集体行为识别方法的对比实验 23](#_Toc477443141)

[2.6.3 参数讨论 26](#_Toc477443142)

[2.7 本章小结 27](#_Toc477443143)

[第 3 章 基于动态核密度的复杂交互集体行为识别方法 28](#_Toc477443144)

[3.1 引言 28](#_Toc477443145)

[3.2 背景知识 30](#_Toc477443146)

[3.2.1 核密度估计 30](#_Toc477443147)

[3.2.2 并查集 31](#_Toc477443148)

[3.3 基于动态核密度的复杂交互集体行为识别方法 32](#_Toc477443149)

[3.3.1 动态核密度估计 32](#_Toc477443150)

[3.3.2 动态核密度峰值聚类算法 34](#_Toc477443151)

[3.3.3 基于分层并查集的一致性合并算法 36](#_Toc477443152)

[3.3.4 算法时间复杂度分析 38](#_Toc477443153)

[3.4 实验结果及分析 39](#_Toc477443154)

[3.4.1 视频监控中的集体行为识别 39](#_Toc477443155)

[3.4.2 微生物集体行为识别 43](#_Toc477443156)

[3.4.3 仿射运动分割 45](#_Toc477443157)

[3.5 本章小结 46](#_Toc477443158)

[第 4 章 基于一致性协方差的集体行为识别方法 47](#_Toc477443159)

[4.1 引言 47](#_Toc477443160)

[4.2 背景知识 49](#_Toc477443161)

[4.2.1 协方差矩阵 49](#_Toc477443162)

[4.2.2 黎曼流形 49](#_Toc477443163)

[4.3 基于一致性协方差的集体行为描述子 50](#_Toc477443164)

[4.3.1 一致性协方差描述子 50](#_Toc477443165)

[4.3.2 相似度计算 52](#_Toc477443166)

[4.4 基于空间加权的密度峰值聚类算法 52](#_Toc477443167)

[4.4.1 密度估计 52](#_Toc477443168)

[4.4.2 一致性近邻 54](#_Toc477443169)

[4.4.3 聚类指派 54](#_Toc477443170)

[4.5 实验结果 56](#_Toc477443171)

[4.6 实验对比及分析 57](#_Toc477443172)

[4.7 本章小结 59](#_Toc477443173)

[第 5 章 基于情绪传播模型的动态人群路径规划 60](#_Toc477443174)

[5.1 引言 60](#_Toc477443175)

[5.2 背景知识 63](#_Toc477443176)

[5.3 问题描述 64](#_Toc477443177)

[5.4 情绪传播模型 64](#_Toc477443178)

[5.4.1 基于人格特征模型的情绪偏好 65](#_Toc477443179)

[5.4.2 基于集体行为子群组的情绪传播算法 67](#_Toc477443180)

[5.5 人群动态路径规划 72](#_Toc477443181)

[5.5.1 环境初始化 72](#_Toc477443182)

[5.5.2 情绪偏好与路径选择的映射 73](#_Toc477443183)

[5.6 实验结果 74](#_Toc477443184)

[5.7 实验对比及分析 77](#_Toc477443185)

[5.7.1 情绪偏好的初始化 77](#_Toc477443186)

[5.7.2 不同情绪状态下路径选择行为的对比 78](#_Toc477443187)

[5.7.3 与不同局部碰撞避免方法的兼容性实验 79](#_Toc477443188)

[5.7.4 与其它人群路径规划方法的对比实验 81](#_Toc477443189)

[5.8 本章小结 82](#_Toc477443190)

[第 6 章 总结和展望 84](#_Toc477443191)

[6.1 工作总结 84](#_Toc477443192)

[6.2 下一步工作 85](#_Toc477443193)

[参考文献 86](#_Toc477443194)

[个人简历、在学期间学术成果 93](#_Toc477443195)

图目录

[图 1.1计算机图形学主要研究方向。 1](file:///E:\毕业\PHD_thesis\第一章-绪论.docx#_Toc521600155)

[图 1.2富含语义的三维建模过程示意图。 2](file:///E:\毕业\PHD_thesis\第一章-绪论.docx#_Toc521600156)

[图 1.3三维扫描设备图。左图为MetraSCAN 3D手持高精度扫描设备，右图为Twinstant Mobile大型扫描设备。 4](file:///E:\毕业\PHD_thesis\第一章-绪论.docx#_Toc521600157)

[图 1.4基于多视角图像建模方法。左图为基于SfM的建模方法，右图为基于SLAM的建模方法。 5](file:///E:\毕业\PHD_thesis\第一章-绪论.docx#_Toc521600158)

[图 1.5基于交互式的三维建模方法。 6](file:///E:\毕业\PHD_thesis\第一章-绪论.docx#_Toc521600159)

[图 1.6基于视频序列的运动提取方法。左：Yan[1]等人方法。右：Jacquet[2]等人方法。 7](file:///E:\毕业\PHD_thesis\第一章-绪论.docx#_Toc521600160)

[图 1.7 基于模型分析的铰链连接类型。 8](file:///E:\毕业\PHD_thesis\第一章-绪论.docx#_Toc521600161)

[图 1.8研究内容示意图。 9](file:///E:\毕业\PHD_thesis\第一章-绪论.docx#_Toc521600162)

表目录

**未找到图形项目表。**

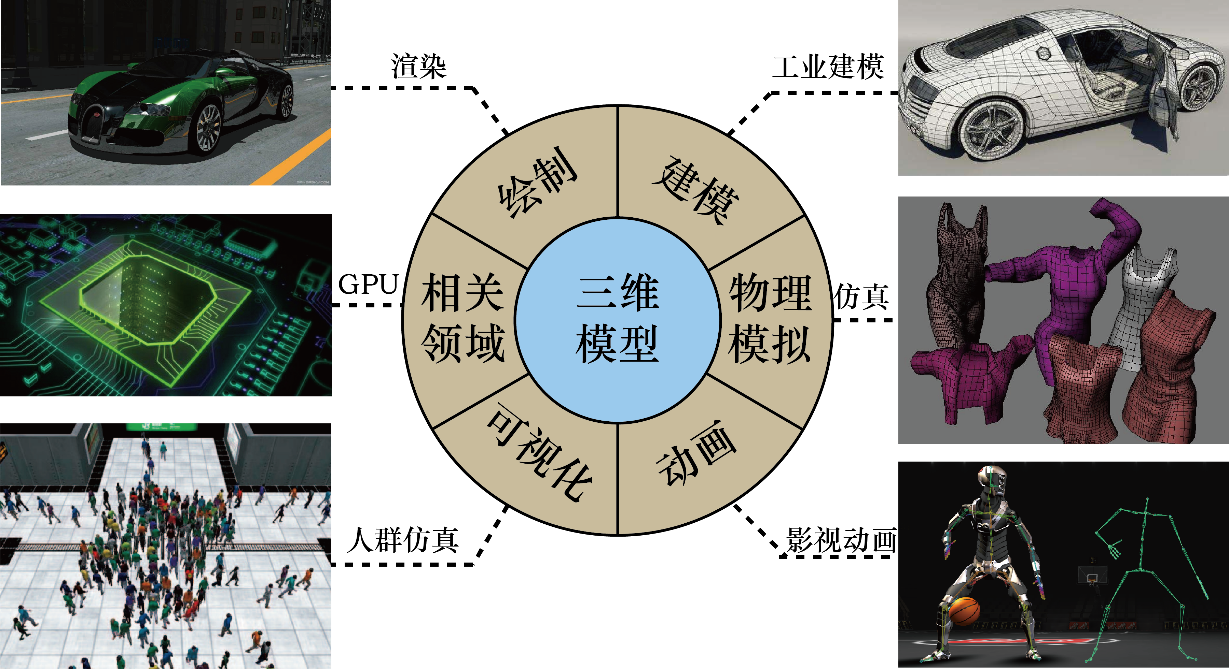
# 绪论

* 1. 研究背景

计算机图形学作为计算机领域的重要学科，其主要研究方向包含：绘制、建模、物理模拟、动画、可视化等延展的相关领域。计算机图形学的研究内容是可视对象的数字和成与操作，是一门面向“数字信息表达”的学科，而该领域的核心数字信息即为三维几何模型。所以，三维模型在精度上、语义完善等方面的提升都将促进图形学中各个领域的发展，并且，图形学中各个领域的发展也迫切要求三维建模方法具有更高的精度、算法效率和普适性。

三维建模作为计算机图形学领域研究的重要组成部分之一，它广泛应用于工业制造、产品设计、影视动画制作等多个领域。三维建模过程是对几何信息的获取、处理、优化及分析，生成“立即可用（readily usable）”的完备三维模型是该领域研究人员长期追求的目标。近年来随着传感设备的精确化、实时化和低成本化，极大推动了三维模型数据获取的发展和普及。这也为三维建模领域提供了更广阔的发展和更严峻的要求。传统的三维建模主要关注于模型的静态几何结构的获取，而对三维模型的分析与理解等操作都作为单独的课题对模型进行后续的优化操作，如模型分割[[3](#_ENREF_3)]、运动提取[[4](#_ENREF_4)]、模型拆解[[5](#_ENREF_5)]、快速渲染[[6](#_ENREF_6)]等相关研究。为了使建模的结果达到“立即可用”的标准，所建模型就需要富含了更多的语义信息，从而使模型具有更广阔的应用领域。例如，获取模型的运动信息，可将模型直接应用于物理模拟的仿真领域；优化模型网格信息，可提高多分辨率（LOD）网格模型渲染速度；对模型进行拆解优化，可得到适宜三维打印的模型部件组。所以在建模过程中，要求算法不仅仅提取模型的几何信息，同时应该挖掘模型更深层次的语义信息。

图 1.1计算机图形学主要研究方向。



目前，该领域的研究者主要关注新颖的知识驱动和数据驱动三维建模方法。依靠人的先验知识实现三维模型的描述、提取和语义分析，其优点是显示得将人类知识集成到三维建模过程中，提高了人为操作的建模效率，但对于大规模的建模任务，人为指定的规则往往缺乏大数据朴实的适应性，并且完善的规则制定也需要完备的领域知识，所以很难广泛应用于建模方法中。数据驱动的三维建模方法是从数据中学习结构和语义的关联，具有很好的泛化能力和相对简单的方法设计，近年来也收到了更多的关注，但完全基于数据驱动的三维建模方法还不成熟，模型数据规模的不足和语义信息定义的不完善，也是其主要面临的瓶颈问题。综上所述，富含语义的三维模型建模的研究在计算机图形学领域具有重要意义，也是学术界公认的主要瓶颈之一。

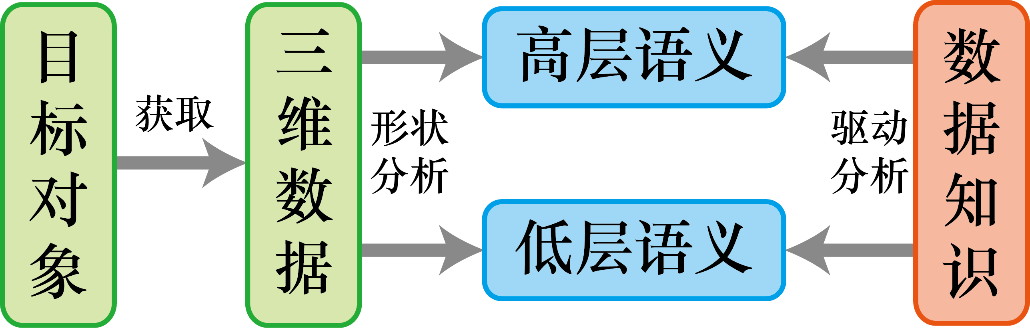


图 1.2富含语义的三维建模过程示意图。

传统的三维建模方法是通过采集设备获取目标对象的三维数据。但为了使模型富含语义信息而做到“立即可用”的效果，建模过程中需要将三维数据通过形状分析获取语义。在获取语义的过程中，需要基于用户先验知识驱动或数据驱动对数据进行分析，如图1.2所示。一个可直接适用于众多领域研究的模型需要有一套完备的语义信息做支撑，其中高层语义不仅包含模型的三维几何形状，同时还需包含形状的模型的分割信息，几何对称性，部件运动信息等等。目前的研究对模型中的高层语义信息的获取比较重视，因为具有这些高层信息的三维模型可以直接应用于运动仿真，三维渲染，模型识别，3D打印等计算机图形学热点领域。三维模型的低层语义信息，包括几何形状表示，网格拓扑结构等信息。低层语义信息的改进有助于精确地表达三维模型，加快模型的计算效率，优化模型的存储结构等。高层语义和低层语义都是蕴含在三维模型内部，需要在建模过程中提取、处理及分析的重要信息，是使模型做到“立即可用”的数据基础。

针对当前三维建模技术的语义信息获取途径匮乏，本文对于富有语义的三维建模方法进行了深入的研究，提出了一系列获取模型高层和低层语义的三维建模方法和模型优化方法的新技术，实现了所建模型富含精确的几何结构，运动参数，和多分辨率的网格拓扑结构，从而扩展了所建模型的应用范围，将三维建模系统与运动仿真、3D打印、多分辨率渲染、模型存储等计算机图形学中众多应用进行无缝对接，真正做到了“即建即用”的三维建模方法。

接下来，本文首先回顾计算机图形学领域的三维建模方法及模型优化技术的研究现状，然后指出本文的研究内容及意义，最后介绍本文的章节安排。

* 1. 研究现状
     1. 三维建模方法

三维建模作为计算机图形学中的重要组成部分，受到该领域内研究工作者的广泛关注。无论是商业领域还是研究领域都在三维建模方向展开了深入的研究。根据三维模型采集数据的方式不同，三维建模方法可以分为基于专业软件的、基于扫描设备的、基于图像的和基于交互式的建模方法。这些建模方法因为数据的采集方式不同，所以有各自的优缺点，也应用于不同的领域。

**基于专业软件的三维建模方法**。该方法是专业用户通过商用建模软件设计出所需要的三维模型。该方法在工业界应用非常广泛，如游戏产业、影视动漫、工业制造、多媒体等。通过专业软件建模是目前最普遍的建模方式，无论是大规模的城市场景，还是精细的工业机械零件，都可以由专业的建模软件来完成。常用的建模软件包括3Dmax[[7](#_ENREF_7)]、Maya[[8](#_ENREF_8)]、SolidWorks[[9](#_ENREF_9)]、AutoCAD[[10](#_ENREF_10)]、Blender[[11](#_ENREF_11)]、MotionBuilder[[12](#_ENREF_12)]等。这些专业建模软件为用户提供全面的建模操作工具，用户可以自由调整多个观察视角，操作鼠标或画板绘制编辑模型，通过离散网格（如三角形网格）或光滑曲面（如Nurbs曲面）表达三维物体表面。专业软件不仅能够绘制出静态模型的几何形状，并且通过设置关键帧，来控制三维物体的运动动画，来完善三维模型的高层运动信息。不仅如此，专业软件同时也包含了对模型低层网格信息的简化、平滑等丰富的优化操作。同时，建模软件也具有专业的渲染功能，其渲染结果直接应用于影视、游戏等产业。其中SolidWorks[[9](#_ENREF_9)]、AutoCAD[[10](#_ENREF_10)]等工业建模软件集成了众多机械零件和产品设计中的标准模块来辅助用户建立机械零件模型，并能自动分析运动仿真结果，其软件可以更高效的辅助工程师设计出工业级三维机械模型。总的来说，基于专业软件建模是工业界最普遍的建模方法，其优点的是工具众多，模型全面，应用范围广，但其缺点也同样显著：其一，专业的软件服务于专业的用户，想做出精致复杂的三维模型，需要具备专业的三维知识与熟练的软件操作能力，无法直接服务于广大普通用户；其二，普通用户习惯于在二维平面上绘制，直接在三维空间中建模对三维知识要求很高，对于复杂模型的建模往往工作量巨大，进一步提高了专业软件三维建模的门槛。

**基于扫描设备的三维建模方法。**该方法通过扫描设备采集真实世界中的场景物体表面的三维信息，结合后续的三维重建算法可以得到完整的三维模型。通常为了重建出完整的三维模型表面信息，需要三维扫描设备对该物体从多个视角进行连续扫描，通过密集的多视角信息重建出物体的稠密三维点云信息。三维扫描仪根据其扫描精度、范围、距离、原理的不同，其价格也差别非常大。例如MetraSCAN 3D[[13](#_ENREF_13)]是Creaform公司推出的便携式光学CMM（Coordinate Measuring Machine）3D扫描仪，如图1.3左图所示，该设备通过光学坐标测量系统扫描物体的三维信息。其测量物体的范围介于1米至3.5米，精度最高可达到0.064毫米，但其价格非常昂贵。由于多个光学扫描设备在数据采集时会相互干扰和扫描距离的限制，大型的扫描设备通常采用结构光和RGB相机结合的技术，多台扫描仪同时进行数据采集，如Twinstant Mobile[[14](#_ENREF_14)]扫描设备，如图1.3右图所示。其扫描最大高度为2.18米，精度为0.7毫米。该设备体积庞大，建模流程和配置复杂，价格昂贵，只适用于工业领域内的人体建模。目前民用级别的扫描设备也越来越普遍，如微软公司发布的Kinect 2.0[[15](#_ENREF_15)]设备，该设备采用结构光扫描方式，扫描范围为0.8米至4.0米，精度为厘米级别，由于其低精度和低帧率的原因，通过反复扫描物体才能得到物体的稠密点云数据，且粗糙的点云数据也限制了该设备的应用范围。总的来说，高精度、大场景的三维扫描设备的价格非常昂贵，只能应用于专业的建模领域，难以应用于普通用户，而低成本的扫描设备，精度低、扫描场景小，很难得到广泛应用。



图 1.3三维扫描设备图。左图为MetraSCAN 3D手持高精度扫描设备，右图为Twinstant Mobile大型扫描设备。

**基于图像的三维建模方法**。该方法根据输入图像个数的不同分为基于多视角图像建模和基于单视角图像建模。基于多视角图像建模是在若干个不同视角下对真实物体进行拍照，通过多张图像间的特征点对应匹配[[16](#_ENREF_16)]和照相机的参数信息[[17](#_ENREF_17)]，根据立体视觉原理[[18](#_ENREF_18)]，确定匹配点在空间中的三维坐标，从而得到三维模型的点云信息。而基于单视角图像建模的研究仅从单张二维图像和深度信息中获取模型数据信息，往往三维数据不完整，通常采用数据驱动的方式[[19](#_ENREF_19)]在已有的数据库中寻找相匹配的模型信息来完善所建模型。

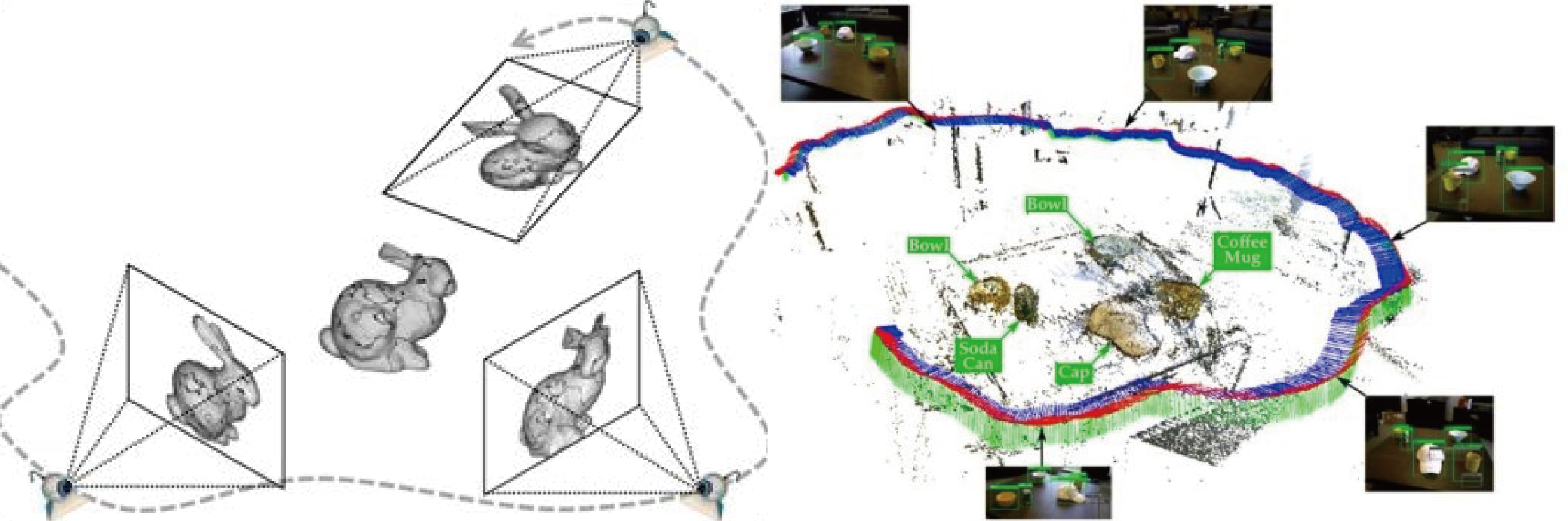
目前基于多视角建模方法主要有两种主流路线，其一，为基于运动恢复结构[[20](#_ENREF_20)]（Structure from Motion，SfM）的技术路线，该方法通常被用于从多视角图像中重建出三维模型的点云数据，如图1.4左图所示。其二，为基于同时定位与地图构建[[21](#_ENREF_21)]（Simultaneous Localization And Mapping，SLAM）的技术路线，该方法主要面向大规模室外场景基于视频序列的实时三维重建，如图1.4右图所示。SfM的基本流程为检测每张图像的特征点，然后在多张图像间对特征点进行匹配[[22](#_ENREF_22), [23](#_ENREF_23)]，使用RANSAC（RANdom Sample Consensus）方法[[24](#_ENREF_24)]得到照相机矩阵参数，最后通过迭代增量的SfM方法重建出三维模型的点云信息。SfM针对小场景物体广泛采用KLT追踪器[[23](#_ENREF_23), [25](#_ENREF_25)]对特征进行匹配，一些方法采用图像局部的像素块[[26](#_ENREF_26), [27](#_ENREF_27)]作为图像特征或者其他的图像特征描述符号[[28-31](#_ENREF_28)]。对于大场景的多视角图像数据，SfM方法可以高效地处理百万级别的图像[[32](#_ENREF_32)]，其主要的处理思路包括：分块处理[[33](#_ENREF_33)]，多层次处理[[34](#_ENREF_34)]，离线处理[[35](#_ENREF_35)]，结构化处理[[36](#_ENREF_36)]等。对于实时建模方面，SLAM技术是指自身方位的定位，即计算采集设备在空间中的位置及朝向，并同时构建三维环境的地图，即重建出周围环境的三维点云数据。Klein[[37](#_ENREF_37)]等人提出了并行跟踪和建图的构架，Mur-Artal等人[[38](#_ENREF_38)]能够在大尺度场景下实时运行，RDSLAM[[39](#_ENREF_39)]能够在动态场景下稳定工作，RKSLAM[[40](#_ENREF_40)]可以实时运行在移动设备上。

图 1.4基于多视角图像建模方法。左图为基于SfM的建模方法，右图为基于SLAM的建模方法。

基于多视角图像的建模方法在重建大规模场景时不需要用户操作，即可以重建出精确的三维场景信息，这是该建模方法的最大优势。但是该方法对于输入的视频或照片要求较高，为了提高重建精度和成功率，则需要稳定和缓慢地移动采集设备进行场景拍摄，一旦重建失败，则需要重新采集数据。

图 1.5基于交互式的三维建模方法。

**基于交互式的三维建模方法**。该方法是通过用户绘制二维草图生成三维模型。如文献[[41-45](#_ENREF_41)]中开发了通过草图直接重建三维模型系统，该建模方法不仅能重建简单的单个物体，同时能构造出合理的三维多边形表面、曲面甚至复杂的大型建筑。交互式建模方法也可以和图像信息相结合，如修饰基于图像的建模结果或直接在二维图像上绘制草图。Jiang[[46](#_ENREF_46)]等人和Arikan[[47](#_ENREF_47)]等人通过用户手动交互改进从图像中提取的三维模型，如图1.5右图所示。Olsen[[48](#_ENREF_48)]等人通过用户在图像中绘制出模型的边界、特征等信息，挤压或膨胀出模型的三维形状。Andre[[49](#_ENREF_49)]等人提出了一种基于交互式的复杂模型的建模系统，用户可以将复杂物体拆分开，逐步绘制出物体的所有细节。Oh[[50](#_ENREF_50)]等人通过用户在模型照片中标出深度和图层信息，辅助系统从图像中提取三维模型，如图1.5左图所示。本人发表在ACM Transactions on Graphics的基于多视角图像交互建模方法[[51](#_ENREF_51)]，通过用户在二维图像中仅需三笔勾勒出模型的二维轮廓，计算得到三维模型信息。

基于交互式的建模方法与基于专业软件的方法相比，虽然都有用户参与，但参与程度大大降低，并且也不需要用户具备很强的专业知识。与基于扫描设备的建模方法相比，该方法不需要复杂的采集设备。相比较多视角建模方法，通过人工参与，大大提高了可建模型的复杂程度。所以交互式建模方法具有门槛低，灵活性高，应用广等特点。

* + 1. 三维模型运动提取方法

在三维建模过程中，不仅要构建静态的三维模型，并且模型部件间蕴含的运动语义信息也需要被提取。现实场景中大部分铰接模型（articulated model）中的部件都是通过铰链连接，则模型的运动语义信息即为部件间的连接类型和部件的刚体运动参数。通过获取运动语义信息，可以获得对三维运动模型的功能性理解[[52](#_ENREF_52)]。

**基于多帧重建的运动提取方法。**该方法通常需重建出视频中多帧的点云数据，通过计算点云的运动轨迹，优化出对应模型部件的运动参数。Tresadern [[53](#_ENREF_53)]等人采用SfM方法重建模型得到每一视频帧的点云数据，在视频序列中跟踪特征点，匹配对应的点云，通过时序点云运动信息优化得到物体运动中的旋转轴。在此基础上，Yan[[1](#_ENREF_1)]等人提出的方法是首先对非刚性运动部分在运动关节处进行分割，然后计算出每个部件刚性运动参数，组件出模型的运动链，如图1.6左图所示。Jacquet [[2](#_ENREF_2)]等人致力于用单个运动矩阵描述两个相连接部件的运动，以简化计算过程，该单个运动矩阵可以包含所有的相关运动参数，如图1.6右图所示。Chang [[54](#_ENREF_54)]等人提出了全局关节运动配准算法，从动态视频序列帧中重建出具有关节运动的三维模型。这些运动提取方法很强程度上依赖在时序上重建出的点云数据的匹配和运动信息，连续视频帧的三维重建非常耗时，运动物体的遮挡也会很大程度影响时序上点云间的匹配，从而影响提取的运动参数的精度。

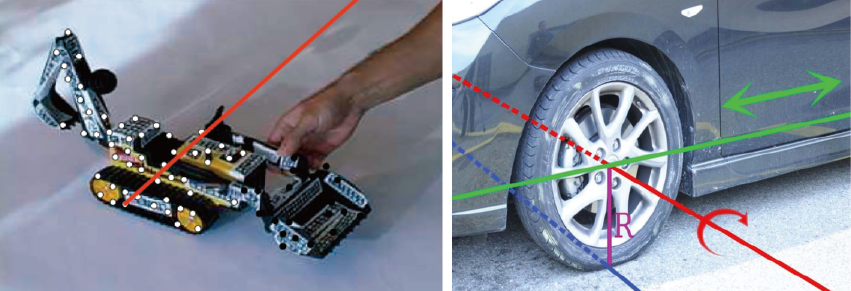


图 1.6基于视频序列的运动提取方法。左：Yan[[1](#_ENREF_1)]等人方法。右：Jacquet[[2](#_ENREF_2)]等人方法。

**基于模型分析的运动提取方法。**利用先验知识，在给定的静态三维模型上进行分析，赋予模型中部件的运动语义信息，该运动语义信息不仅包含模型人为设定的运动姿态，还可以演化出模型能够执行的运动姿态。

提取模型运动语义信息，可以服务于用户的交互式操作。Savva [[55](#_ENREF_55), [56](#_ENREF_56)]等人的工作可以识别出场景中模型与用户交互的运动信息。Hu[[57](#_ENREF_57)]等人通过分析模型动态交互信息提取静态物体的运动功能，如一个杯子如何装满咖啡和被人饮用的运动。Hermans[[58](#_ENREF_58)]等人的工作通过学习模型的动态交互操作，可以使机器人学会平移和旋转该物体。这些方法主要关注于分析模型与用户的交互式运动信息，而没有探测物体自身的相对运动信息。

一些研究工作可直接从模型的几何信息中提取模型运动参数。如早期Gelfand[[59](#_ENREF_59)]等人从模型的几何形状推测出模型中部件可活动的运动信息，但可推测的运动信息有限，该工作也主要适用于模型的分割算法。Xu[[60](#_ENREF_60)]等人归纳了铰接模型中铰链连接的几种运动形式，如图1.7左图所示，通过分析模型中的铰接类型推测模型的运动信息。Mitra[[61](#_ENREF_61)]等人通过分析机械模型的几何结构和空间配置推测出机械模型中各单个部件可能产生的运动，并根据分析结果推测组装后的完整模型的运动信息。Guo[[62](#_ENREF_62)]等人用相似的方法得到模型的运动信息，将该信息应用于三维机械模型的组装。最近一些工作将模型几何信息与视频信息进行结合，更精准的提取模型的运动参数。如Li[[63](#_ENREF_63)]等人通过分析视频中的模型运动部分来捕捉关节对象的运动信息，不需要重建多帧模型的数据，直接解决了关节运动参数的优化问题。本人工作[[51](#_ENREF_51)]在Xu[[60](#_ENREF_60)]等人归纳的铰接类型上进行了扩展，如图1.7右图所示，通过对重建模型分析建立铰链类型候选集，建立视频序列图像与模型运动部件的匹配函数，优化目标函数得到精确的模型部件运动参数，仿真出与视频中相同的模型运动姿态。

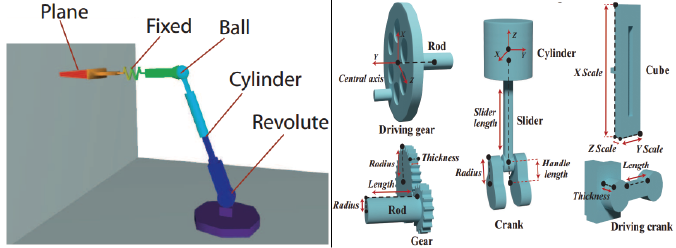


图 1.7 基于模型分析的铰链连接类型。

* 1. 研究内容及意义

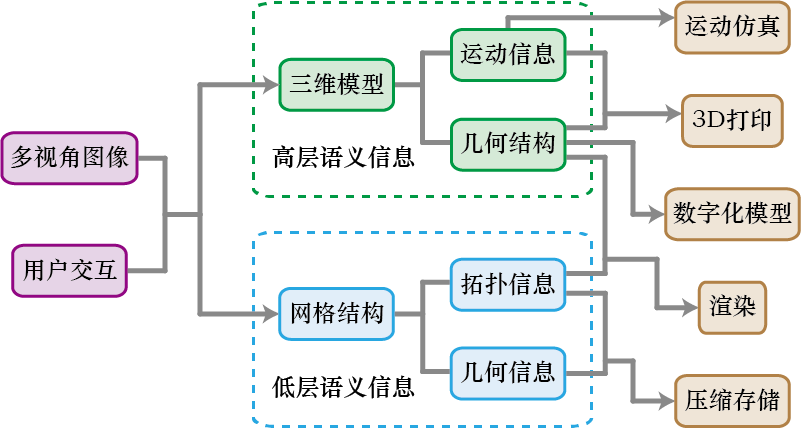
上述研究现状展现了基于不同采集数据的方式的三维建模方法，每种建模方法都各有优缺点，其方法也应用在不同的领域。在几种三维建模方法中，基于交互式的建模方法最为灵活，它可以与多种采集设备相结合进行三维重建。本文的目标是基于多视角图像的交互式建模方法，吸收了基于多视角建模和交互式建模两种建模方法的优势，使所建的模型更广泛，类型更多，结构更复杂，并通过对模型语义信息的优化，得到具有完备语义信息的三维模型，使模型能够更直接、更广泛地应用于其他研究及工业领域。

图 1.8研究内容示意图。

如图1.6所示，本文希望通过多视角图像信息与用户交互信息得到模型的三维数据。该数据分为两个层次语义信息，模型的高层语义信息包括模型的几何数据和运动信息，而模型的低层语义信息包括网格的拓扑结构和几何结构。目前的建模方法只注重重建出模型的几何数据，该数据只能够表达模型的静态几何信息，本文重建出的模型运动信息可使模型能够直接应用于运动仿真、3D打印等领域。优化模型的网格结构对于三维模型也是非常有必要的，该优化能使模型有更好的数据存储结构，本文希望模型具有层次网格结构，该结构能使模型直接应用于多分辨率渲染与模型网格压缩中。综上所述，本文开展了以下四个方面的研究。

1. 提出一种基于多视角的交互式建模方法，该方法不需要用户具备专业的建模知识，且操作简单，能在短时间内重建出复杂的三维模型。
2. 提出一种基于视频序列的模型运动仿真方法，该方法可以从视频中自动提取模型的运动信息，使重建的模型具有更完备的语义信息。
3. 提出一种模型运动参数优化方法，该方法可以直接应用在模型的拆解、装配和装箱问题上，解决了模型压缩及降低3D打印成本等问题。
4. 提出一种实时精确的网格半规则化方法，该方法可将模型网格结构优化成层次网格结构，使模型具有适用于多分辨率渲染与压缩的低层数据结构。
   1. 论文章节安排

本文针对重建出具有完备语义的三维模型建模方法展开了相关研究，各章节组织如下：

在第1章绪论中，介绍了三维建模领域的研究背景，回顾了研究现状，分析了各种三维建模方法的优缺点，结合存在的问题概述了本文的研究内容。

在第2章中，本文提出了一种基于多视角交互式建模方法，该方法的交互方式为用户通过三笔在二维图像上绘制出模型的轮廓线，通过多视角生成的点云信息与用户绘制的轮廓线得到初始模型，再通过自动探测的模型部件间的几何约束优化模型参数，得到精确模型。该方法通过用户的简单交互式操作得到精确的三维模型。

在第3章中，为了获取模型高层运动运动语义信息，提出了一种基于视频序列的机械模型仿真方法，该方法基于机械运动先验知识驱动，通过优化三维模型剪影与图像边界匹配的能量函数，得到精确的模型运动参数，得到与视频中真实物体运动形态一致的模型仿真结果。

在第4章中，提出了一种基于运动参数优化的机械模型装箱算法。该方法利用模型运动的特性，提高了模型拆解和装箱算法的效率和空间利用率。通过对模型部件运动参数的优化可以使模型更好的应用于3D打印等领域。

在第5章中，提出了一种实时精确的网格半规则化算法。该网格优化算法通过分层建立独立集使算法并行化，实时得到精确的半规则网格。具有层次网格结构的模型，可以更好的应用于多分辨率渲染、模型压缩等领域。

在第6章中，总结本文的工作并展望下一步工作。本文工作所有的相关参考文献，致谢和本人攻读博士期间所发表的文章在最后给出。

# 基于

* 1. 引言

集体行为是最普遍的自然现象之一，

。

* 1. 背景知识

提出的

* 1. 本章小结

# 识别方法

* 1. 引言

已有的集体行为识别方法中

* 1. 本章小结

集体行为中同时蕴含着局部和全局行为模式。

# 总结和展望

* 1. 工作总结
  2. 下一步工作

参考文献

[1] D. Helbing, P. Molnár. Social force model for pedestrian dynamics[J]. Physical Review E, 1995, 51(5), 4282.

[2] T. Vicsek, A. Czirók, E. Ben-Jacob, I. I. Cohen, O. Shochet. Novel type of phase transition in a system of self-driven particles[J]. Physical Review Letters, 2006, 75(6), 1226.

个人简历、在学期间学术成果

（一）个人简历

（二）在学期间学术成果

（三）参与工作

[1] J. Yan, M. Pollefeys. A factorization-based approach for articulated nonrigid shape, motion and kinematic chain recovery from video[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2008, 30(5), 865-877.

[2] B. Jacquet, R. Angst, M. Pollefeys. Articulated and Restricted Motion Subspaces and Their Signatures[C]. Computer Vision and Pattern Recognition, 2013, 1506-1513.

[3] G. Marsaglia. Choosing a Point from the Surface of a Sphere[J]. Annals of Mathematical Statistics, 1972, 43(2), 645-646.

[4] L. Zhu, W. Xu, J. Snyder, Y. Liu, G. Wang, B. Guo. Motion-guided mechanical toy modeling[J]. Acm Transactions on Graphics, 2012, 31(6), 127.

[5] G. Tian, M. Zhou, P. Li. Disassembly sequence planning considering fuzzy component quality and varying operational cost[J]. IEEE Transactions on Automation Science and Engineering, 2018, 15(2), 748-760.

[6] E. Gobbetti, E. Bouvier. Time-critical multiresolution scene rendering[C]. Visualization '99. Proceedings, 2018, 123-130.

[7] *Autodesk 3DS Max Software* Available: <http://www.autodesk.com/products/3ds-max/overview>

[8] *Autodesk Maya Software*. Available: <http://www.autodesk.com/products/maya/overview>

[9] *Wiki:SolidWorks*. Available: https://en.wikipedia.org/wiki/SolidWorks

[10] *Autodesk AutoCAD Software*. Available: <http://www.autodesk.com/products/autocad/overview>

[11] *Blender Foundation*. Available: https://[www.blender.org](http://www.blender.org)

[12] *Autodes MotionBuilder*. Available: <http://www.autodesk.com/products/motionbuilder/overview>

[13] *METRASCAN 3D*. Available: https://[www.creaform3d.com/en/metrology-solutions/optical-3d-scanner-metrascan](http://www.creaform3d.com/en/metrology-solutions/optical-3d-scanner-metrascan)

[14] *Twinstant Mobile*. Available: https://web.twindom.com/

[15] R. A. Newcombe, S. Izadi, O. Hilliges, D. Molyneaux, D. Kim, A. J. Davison*, et al.* KinectFusion: Real-time dense surface mapping and tracking[C]. IEEE International Symposium on Mixed and Augmented Reality, 2012, 127-136.

[16] J. Luo, O. Gwun. A Comparison of SIFT, PCA-SIFT and SURF[J]. International Journal of Image Processing, 2013, 3(4), 143-152.

[17] Z. Zhang. A flexible new technique for camera calibration[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2000, 22(11), 1330-1334.

[18] R. Hartley, A. Zisserman. Multiple view geometry in computer vision[C]. 2000, 1865 - 1872.

[19] J. Rock, T. Gupta, J. Thorsen, J. Y. Gwak, D. Shin, D. Hoiem. Completing 3D object shape from one depth image[C]. Computer Vision and Pattern Recognition, 2015, 2484-2493.

[20] N. Snavely, S. M. Seitz, R. Szeliski. Photo Tourism: Exploring Photo Collections In 3D[J]. Acm Transactions on Graphics, 2006, 25(3), págs. 835-846.

[21] F. Endres, J. Hess, N. Engelhard, J. Sturm, D. Cremers, W. Burgard. An evaluation of the RGB-D SLAM system[C]. IEEE International Conference on Robotics and Automation, 2012, 1691-1696.

[22] D. G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints[J]. International Journal of Computer Vision, 2004, 60(2), 91-110.

[23] C. Tomasi. Detection and tracking of point features[J]. Technical Report, 1991, 91(21), 9795-9802.

[24] M. A. Fischler, R. C. Bolles. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography[J]. Readings in Computer Vision, 1987, 24(6), 726-740.

[25] B. D. Lucas, T. Kanade. An iterative image registration technique with an application to stereo vision[C]. International Joint Conference on Artificial Intelligence, 1981, 674-679.

[26] D. Scaramuzza, F. Fraundorfer. Visual Odometry [Tutorial][J]. Robotics & Automation Magazine IEEE, 2011, 18(4), 80-92.

[27] E. Royer, M. Lhuillier, M. Dhome, J. M. Lavest. Monocular Vision for Mobile Robot Localization and Autonomous Navigation[J]. International Journal of Computer Vision, 2007, 74(3), 237-260.

[28] K. Mikolajczyk, C. Schmid. A performance evaluation of local descriptors[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2005, 27(10), 1615-1630.

[29] J. Matas, O. Chum, M. Urban, T. Pajdla. Robust wide-baseline stereo from maximally stable extremal regions[J]. Image & Vision Computing, 2004, 22(10), 761-767.

[30] H. Bay, A. Ess, T. Tuytelaars, L. V. Gool. Speeded-Up Robust Features (SURF)[J]. Computer Vision & Image Understanding, 2008, 110(3), 346-359.

[31] G. Zhang, Z. Dong, J. Jia, T. T. Wong, H. Bao. Efficient non-consecutive feature tracking for structure-from-motion[C]. European Conference on Computer Vision, 2010, 422-435.

[32] J. M. Frahm, P. Fite-Georgel, D. Gallup, T. Johnson, R. Raguram, C. Wu*, et al.* Building Rome on a cloudless day[C]. European Conference on Computer Vision, 2010, 368-381.

[33] D. Steedly, I. Essa, F. Dellaert. Spectral partitioning for structure from motion[C]. IEEE International Conference on Computer Vision, 2003, 996.

[34] W. Burgard, O. Brock, C. Stachniss. Mapping Large Loops with a Single Hand-Held Camera[C]. Robotics: Science & Systems Iii, June, Georgia Institute of Technology, Atlanta, Georgia, Usa, 2007, 297-304.

[35] K. Ni, D. Steedly, F. Dellaert. Out-of-Core Bundle Adjustment for Large-Scale 3D Reconstruction[C]. IEEE International Conference on Computer Vision, 2007, 1-8.

[36] N. Snavely, S. M. Seitz, R. Szeliski. Skeletal graphs for efficient structure from motion[J]. 2008, 1(1--8.

[37] G. Klein, D. Murray. Parallel Tracking and Mapping for Small AR Workspaces[C]. IEEE and ACM International Symposium on Mixed and Augmented Reality, 2007, 1-10.

[38] R. Mur-Artal, J. M. M. Montiel, J. D. Tardós. ORB-SLAM: A Versatile and Accurate Monocular SLAM System[J]. IEEE Transactions on Robotics, 2017, 31(5), 1147-1163.

[39] W. Tan, H. Liu, Z. Dong, G. Zhang, H. Bao. Robust monocular SLAM in dynamic environments[C]. IEEE International Symposium on Mixed and Augmented Reality, 2013, 209-218.

[40] H. Liu, G. Zhang, H. Bao. Robust Keyframe-Based Monocular SLAM for Augmented Reality[C]. IEEE International Symposium on Mixed and Augmented Reality, 2017, 1-10.

[41] T. Igarashi, S. Matsuoka, H. Tanaka. Teddy: a sketching interface for 3D freeform design[C]. Conference on Computer Graphics and Interactive Techniques, 1999, 409-416.

[42] B. Xu, W. Chang, A. Sheffer, A. Bousseau, J. Mccrae, K. Singh. True2Form:3D curve networks from 2D sketches via selective regularization[J]. Acm Transactions on Graphics, 2014, 33(4), 1-13.

[43] X. Chen, S. B. Kang, Y. Q. Xu, J. Dorsey, H. Y. Shum. Sketching reality:Realistic interpretation of architectural designs[J]. Acm Transactions on Graphics, 2008, 27(2), 1-15.

[44] L. Olsen, F. F. Samavati, M. C. Sousa, J. A. Jorge. Sketch-based modeling: A survey[J]. Computers & Graphics, 2009, 33(1), 85-103.

[45] M. Eitz, R. Richter, T. Boubekeur, K. Hildebrand, M. Alexa. Sketch-based shape retrieval[J]. Acm Transactions on Graphics, 2012, 31(4), 1-10.

[46] N. Jiang, P. Tan, L. F. Cheong. Symmetric architecture modeling with a single image[C]. ACM SIGGRAPH Asia, 2009, 113.

[47] M. Arikan, M. Wimmer, S. Maierhofer. O-snap:Optimization-based snapping for modeling architecture[J]. Acm Transactions on Graphics, 2013, 32(1), 1-15.

[48] L. Olsen, F. F. Samavati. Image-assisted modeling from sketches[J]. 2010, 225-232.

[49] A. Andre, S. Saito. Single-view sketch based modeling[C]. Eighth Eurographics Symposium on Sketch-Based Interfaces and Modeling, 2011, 133-140.

[50] B. M. Oh, M. Chen, J. Dorsey. Image-based modeling and photo editing[C]. ACM Siggraph’01. ACM, 2001, 433-442.

[51] M. Xu, M. Li, W. Xu, Z. Deng, Y. Yang, K. Zhou. Interactive mechanism modeling from multi-view images[J]. ACM Transactions on Graphics, 2016, 35(6), 236.

[52] R. Hu, B. Wu, B. Wu, A. Shamir, A. Shamir, H. Zhang. Learning how objects function via co-analysis of interactions[J]. Acm Transactions on Graphics, 2016, 35(4), 47.

[53] P. Tresadern, I. Reid. Articulated Structure from Motion by Factorization[C]. Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, 2005, 1110-1115 vol. 2.

[54] W. Chang, M. Zwicker. Global registration of dynamic range scans for articulated model reconstruction[J]. Acm Transactions on Graphics, 2011, 30(3), 1-15.

[55] M. Savva, A. X. Chang, P. Hanrahan, M. Fisher, M. Nießner. SceneGrok: inferring action maps in 3D environments[J]. Acm Transactions on Graphics, 2014, 33(6)

[56] M. Savva, A. X. Chang, P. Hanrahan, M. Fisher. PiGraphs: learning interaction snapshots from observations[J]. Acm Transactions on Graphics, 2016, 35(4), 139.

[57] K. Hu, S. D. Rajasekaran, H. Kang, Y. Yoshiyasu, B. Benes, L. J. Guibas. Understanding and exploiting object interaction landscapes[J]. Acm Transactions on Graphics, 2017, 36(4), 52b.

[58] T. Hermans, F. Li, J. M. Rehg, A. F. Bobick. Learning contact locations for pushing and orienting unknown objects[C]. Ieee-Ras International Conference on Humanoid Robots, 2013, 435-442.

[59] N. Gelfand, L. J. Guibas. Shape segmentation using local slippage analysis[C]. Symposium on Geometry Processing, 2004, 214-223.

[60] W. Xu, J. Wang, K. K. Yin, K. Zhou, V. D. P. Michiel, F. Chen*, et al.* Joint-aware manipulation of deformable models[J]. Acm Transactions on Graphics, 2009, 28(3), 1-9.

[61] N. Fish, M. Averkiou, O. V. Kaick, O. Sorkine-Hornung, D. Cohen-Or, N. J. Mitra. Meta-representation of shape families[J]. Acm Transactions on Graphics, 2014, 33(4), 34.

[62] J. Guo, D. M. Yan, E. Li, W. Dong, P. Wonka, X. Zhang. Illustrating the disassembly of 3D models[J]. Computers & Graphics, 2013, 37(6), 574-581.

[63] H. Li, G. Wan, H. Li, A. Sharf, K. Xu, B. Chen. Mobility fitting using 4D RANSAC[J]. Computer Graphics Forum, 2016, 35(5), 79-88.