**天津工业大学**

本 科 生 毕 业 设 计（论 文）

**基于****特征点匹配的点云自动拼接系统设计**

**学 号：　 1610410319**

**姓 名：　　　　师 聪**

**专 业：　　　　自动化**

**学 院： 电气工程与自动化学院**

**指导教师：　　　　宋丽梅**

**职 称：　　　　教 授**

**完成日期：　 2020年5月**

**天津工业大学2020届本科毕业设计（论文）任务书**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **题目** | **基于特征点匹配的点云自动拼接系统设计** | | | | | | |
| **学生姓名** | **师聪** | **学号** | **1610410319** | **专业班级** | | **自动化1603** | |
| **学院名称** | **电气工程与自动化学院** | | | | **课题类型** | | **工程研究** |
| **课题意义** | 点云拼接是三维重建中非常重要的一个研究方向，经过几十年的研究，点云拼接取得了非常多的研究成果，但仍存在算法复杂度、计算时间和精度难以兼顾的问题。基于此，提出了一种算法简单、适用性强以及实时性好的自由视角三维点云拼接算法尤为重要。  通过特征点匹配来找到不同角度采集到图像的匹配点，再有二维映射到三维求取旋转平移矩阵来实现自动拼接，可以做到适用性强以及实时性好，实现自由视角三维点云拼接。 | | | | | | |
| **任务要求** | 2019年12月至2020年2月 查阅相关文献，学习相点云算法的理论基础知识以及VS和PCL库的使用。  2020年3月 设计实验方案并搭建实验所需要用到的硬件平台。  2020年4月 模拟实际过程中可能遇到的问题并设计相关实验，采集实验所需要用到的图像并完成实验。  2020年5月 完成论文撰写。  2020年6月 修改论文，准备答辩。 | | | | | | |
| **推荐参考资料** | 1.柴江龙,樊彦国,王斌,韩志聪.尺度及主方向改正的ORB特征匹配算法[J/OL].计算机工程与应用:1-13[2018-10-04].  2.徐建鹏,卜凡亮.三维重建系统下的特征点处理与位姿恢复优化算法[J/OL].计算机应用研究,2019(11):1-3[2018-10-04].  3.Yanyan Qin. Image Feature Points Matching via Improved ORB[A]. IEEE Beijing Section.Proceedings of 2014 IEEE International Conference on Progress in Informatics and Computing[C].IEEE Beijing Section:IEEE BEIJING SECTION,2014:5.  4.Ye, Feng,Hong, Zheng,Lai, Yizong,Zhao, Yuting,Xie, Xianzhi. Multipurification of matching pairs based on ORB feature and PCB alignment case study[D]. JOURNAL OF ELECTRONIC IMAGING, 2018, 27(3).  5.Li Zhuo,Liu Jieyu,Li Hui,Zhou Xiaogang,Li Weipeng. Feature detection and description algorithm based on ORB-LATCH[J]. Journal of Computer Applications, 2017, 37(6):1759-1781. | | | | | | |
| **起止日期** | 2019年12月至2020年6月 | | | | | | |
| **备注** |  | | | | | | |

**指导教师 教研室主任**  **主管院长**

**天津工业大学2020届本科毕业设计（论文）开题报告表**

**学院： 电气工程与自动化学院 2020 年 1 月 2 日**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **姓名** | **师聪** | | **学号** | **1610410319** | **专业班级** | | **自动化1603** | |
| **题目** | **基于特征点匹配的点云自动拼接系统设计** | | | | | | **指导教师** | **宋丽梅** |
| **一、与本课题有关的国内外研究情况、课题的目的、意义及主要内容：**  **研究情况：**Y.Ye等人提出了基于 L 型平面的转台标定方法，得到转台的转轴中心以及转轴方向。Lai 等人将转台进行过修改，但待测物的大小、形状和使用范围同样受到限制。Kim、戴静兰等人提出了一种基于主成分分析(PCA)的拼接方法。李强等人提出了一种基于多特征匹配的点云拼接算法。付鲲等人提出了一种基于曲率信息的人工蜂群点云配准算法，该方法的特征点根据曲率信息获取，并结合人工蜂群算法的特点。  **目的：**利用特征点对物体实现三维拼接。  **意义：**三维重建技术在诸多领域都有广泛应用：文物保护及数字化方面，面向博物馆数字化技术的联合技术，利用结构光三维重建技术，可以快速实现文化遗产的高清记录与复制，对世界文化遗产的保护与传承，对典籍文物进行修复。生活应用领域，虚拟试衣设备中采集用户的三维数据，通过三维人体参数化建模技术可以快速得到人体的三维模型，消费者通过云平台可以在数据库中找到与扫描模型最为接近的衣服型号。  **主要内容：**利用双目标定采集数据，再利用orb算法提取特征点，最后用拼接算法实现特征点自动拼接，完成对一个三维物体的拼接还原。 | | | | | | | | |
| **二、进度及预期结果：** | | | | | | | | |
| **起止日期** | | **主要内容** | | | | **预期结果** | | |
| 2019年12月  2020年3月  2020年4月  2020年5月  2020年6月 | | 查阅相关文献，学习相点云算法的理论基础知识以及VS和PCL库的使用。  设计实验方案并搭建实验所需要用到的硬件平台。  模拟实际过程中可能遇到的问题并设计相关实验，采集实验所需要用到的图像并完成实验。  完成论文撰写。  修改论文，准备答辩。 | | | | 完成  完成  完成  完成  完成 | | |
| **完成课题的现有条件** | | 本人已经查阅相关文献，学习特征值提取算法，有一定的理论基础，并具备一定的编程能力。 | | | | | | |
| **审查意见** | | **指导教师： 年 月 日** | | | | | | |
| **学院意见** | | **主管领导： 年 月 日** | | | | | | |

天津工业大学本科毕业设计（论文）评阅表

（论文类）

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **题目** | | **基于特征点匹配的点云自动拼接系统设计** | | | | | | | | |
| **姓名** | **师聪** | | **学号** | **1610410319** | **专业班级** | 自动化/自动化1603 | | **指导**  **教师** | | 宋丽梅 |
| **评审项目** | | 指标 | | | | | 满分 | | 评分 | |
| **选题** | | 能体现本专业培养目标，使学生得到较全面训练。题目大小、难度适中，学生工作量饱满，经努力能完成。 | | | | | 10 | | 9 | |
| 题目与生产、科研等实际问题结合紧密。 | | | | | 10 | | 8 | |
| **课题调研、 文献检索** | | 能独立查阅文献以及从事其他形式的调研，能较好地理解课题任务并提出实施方案；有分析整理各类信息，从中获取新知识的能力。 | | | | | 15 | | 12 | |
| **论文撰写** | | 结构严谨，理论、观点、概念表达准确、清晰。 | | | | | 10 | | 8 | |
| 文字通顺，用语正确，基本无错别字和病句，图表清楚，书写格式符合规范。 | | | | | 10 | | 10 | |
| **外文应用** | | 能正确引用外文文献，翻译准确，文字流畅。 | | | | | 5 | | 5 | |
| **论文水平** | | 论文论点正确，论点与论据协调一致，论据充分支持论点，论证过程有说服力。 | | | | | 15 | | 13 | |
| 有必要的数据、资料支持，数据、资料翔实可靠，得出的结论有可验性。 | | | | | 15 | | 12 | |
| 论文有独到见解或有一定实用价值。 | | | | | 10 | | 9 | |
| **合计** | | | | | | | 100 | | 86 | |
| **意见及建议：**  该论文选题具有一定的实际意义，基于点云拼接的三维重建技术提出了一种基于特征点匹配的点云自动拼接系统。该系统借助拼接软件开发平台对双目相机进行标定，求取相机内参、外参和畸变系数，具有一定的可行性，然后选用ORB和SIFT相结合的特征提取算法，使得提取速度大幅提高，相比 ORB 不仅精度也大幅提高，而且也解决了尺度不变性的问题。最够使用极线约束(RANSAC)对错误点进行筛选剔除完成三维点云的拼接。通过计算拼接标准偏差为0.0394mm。该论文写作基本满足科技论文写作规范，数据资料可靠，具有一定的实用价值。  **评阅人签名： 年: 月: 日** | | | | | | | | | | |

**天津工业大学 2020 届本科毕业设计（论文）成绩评定表**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **学生姓名** | **师聪** | | **学号** | | **1610410319** | | | **专业班级** | | **自动化/自动化1603** |
| **题目** | **基于特征点匹配的点云自动拼接系统设计** | | | | | | | | | |
| **学院名称** | **电气工程与自动化学院** | | | | | **“查重”检测结果的文字复制比** | | |  | |
| **1．毕业设计（论文）指导教师评语及成绩：**  该论文选题具有一定的实际意义，基于点云拼接的三维重建技术提出了一种基于特征点匹配的点云自动拼接系统。该系统借助拼接软件开发平台对双目相机进行标定，求取相机内参、外参和畸变系数，具有一定的可行性，然后选用ORB和SIFT相结合的特征提取算法，使得提取速度大幅提高，相比 ORB 不仅精度也大幅提高，而且也解决了尺度不变性的问题。最够使用极线约束(RANSAC)对错误点进行筛选剔除完成三维点云的拼接。通过计算拼接标准偏差为0.0394mm。该论文写作基本满足科技论文写作规范，数据资料可靠，具有一定的实用价值。 | | | | | | | | | | |
| **指导教师**  **成绩86**  **年 月 日** | | | | | | | | | | |
| **2．毕业设计（论文）答辩委员会评语及成绩：**  汇报毕业设计过程条理较清楚，描述了一种三维点云拼接算法，通过特征点匹配和空间转换来求取对应的旋转平移矩阵来实现自动拼接。答辩过程中能较准确的回答老师提出的问题，对所涉及的方法及流程较熟悉，体现了较饱满的工作量。 | | | | | | | | | | |
| **答辩主席（或组长）**  **成绩82**    **年 月 日** | | | | | | | | | | |
| **3．毕业设计（论文）总成绩：** | | | | | | | | | | |
| a.指导教师  给定成绩 | | b.评阅教师  给定成绩 | | c.毕业答辩成绩 | | | 总成绩  (a×0.5+b×0.2+c×0.3) | | | |
| 86 | | 88 | | 82 | | | 85.2 | | | |

# 摘　　要

近年来，基于点云拼接的三维重建技术已经发展为机器视觉领域的核心技术之一。三维点云自动拼接技术一直是机器视觉、逆向工程、摄影测量、模式识别等领域的重点前沿研究课题。国内外最常见的拼接算法有转台标定法和特征点标记法，但被测物体的形状、大小会受到限制，费时费力，而且重建的模型也会失去完整性。在人工智能迅速发展的前提下，继续研究更高精度的三维重建技术具有十分重要的意义。基于此，本文提出了一种基于特征点匹配的点云自动拼接系统。

本文中的自动拼接系统可分为如下几个部分：第一步，为了在实验中获取三维点云数据与二维待测物体的采集图像。首先介绍了硬件组成与实现功能、拼接系统软件开发平台；然后采用张正友标定法对双目相机进行标定，求取相机内参、外参和畸变系数；最后，完成数据图像的采集并进行存储。第二步，为了选择特征匹配算法。首先从理论上、算法上对目前主流的特征点匹配算法ORB、SIFT、SURF进行分析；然后，选用ORB和SIFT相结合的特征提取算法，融合后的算法相比SIFT特征点提取速度大幅提高，相比 ORB 不仅精度也大幅提高，而且也解决了尺度不变性的问题。第三步，为了剔除错误特征点对并求解两幅点云之间的*R, t*矩阵。本文使用极线约束(RANSAC)对错误点进行筛选剔除，在得到正确匹配点对之后，使用SVD(奇异性分解)求出两幅点云之间的*R, t*矩阵从而完成了三维点云自动拼接算法。

最后，为了验证算法的合理性。将两幅点云图在计算机中的VS 2013软件开平台上进行自动拼接，通过计算给出了拼接的平均距离为0.0385mm，拼接标准偏差为0.0394mm。该结果表明：本模型可提高点云拼接精度要求。最后，结合实际情况，本算法模型可以进一步推广到工程领域中立体重建与虚拟现实技术等领域。

关键词：双目相机；点云拼接；三维重建；特征点匹配

# **Abstract**

In recent years, 3D reconstruction technology based on point cloud stitching has developed into one of the core technologies in the field of machine vision. Three-dimensional point cloud automatic Mosaic technology has been a key frontier research topic in machine vision, reverse engineering, photogrammetry, pattern recognition and other fields. The most common stitching algorithms at home and abroad include the turntable calibration method and the feature point marking method, but the shape and size of the object to be measured will be limited, which is time-consuming and laborious, and the reconstructed model will also lose its integrity. Under the premise of rapid development of artificial intelligence, it is very important to continue to study the 3d reconstruction technology with higher precision. Based on this, this paper proposes an automatic point cloud Mosaic system based on feature point matching.

The automatic splicing system in this paper can be divided into the following parts: The first step is to obtain the three-dimensional point cloud data and the acquisition image of the two-dimensional object to be measured in the experiment. Firstly, the hardware composition and realization function, the software development platform of the splicing system are introduced. Then, Zhang Zhengyou calibration method is used to calibrate binocular camera, and the internal parameters, external parameters and distortion coefficient are obtained. Finally, the data and images are collected and stored. The second step is to select the feature matching algorithm. Firstly, the main feature point matching algorithms including ORB, SIFT and SURF are analyzed theoretically and algorithmically. Then, the feature extraction algorithm combined with ORB and SIFT is selected. Compared with SIFT, the extraction speed of the fusion algorithm is greatly improved, and the accuracy of the ORB is greatly improved. Besides, the scale invariance problem is also solved. The third step is to eliminate the wrong characteristic point pairs and solve the *R, t* matrix between the two-point clouds. In this paper, the correct and wrong points are screened and eliminated by using the polar line constraint (RANSAC). After the correct matching point pairs are obtained, the *R* and *t* matrix between the two-point clouds are obtained by SVD (singular decomposition), so as to complete the automatic three-dimensional point cloud Mosaic algorithm.

Finally, in order to verify the rationality of the algorithm. Two-point cloud images are automatically spliced on the open platform of vs 2013 software in the computer. The average distance of splicing is 0.0385mm, and the standard deviation of splicing is 0.0394mm. The results show that this model can improve the accuracy requirement of point cloud stitching. Finally, combined with the actual situation, this algorithm model can be further extended to the field of engineering three-dimensional reconstruction and virtual reality technology.

**Keywords**: Binocular camera; Point cloud stitching; Feature point matching; Three-dimensional reconstruction;

目　　录

[摘　　要 i](#_Toc40981504)

[Abstract vi](#_Toc40981505)

[目　　录 viii](#_Toc40981506)

[第一章 绪论 1](#_Toc40981507)

[1.1课题的研究背景及意义 1](#_Toc40981508)

[1.2国内外研究现状 2](#_Toc40981509)

[1.3本文工作及内容安排 4](#_Toc40981510)

[第二章 系统硬件平台与软件开发环境 5](#_Toc40981511)

[2.1拼接系统硬件组成 5](#_Toc40981512)

[2.2拼接系统软件开发平台 6](#_Toc40981513)

[2.3双目相机的标定 6](#_Toc40981514)

[2.4双目相机标定结果 8](#_Toc40981515)

[第三章 特征点匹配筛选与*R, t*矩阵的求解 10](#_Toc40981516)

[3.1点云数据的获取 10](#_Toc40981517)

[3.2特征点提取匹配技术 11](#_Toc40981518)

[3.2.1 ORB算法介绍 11](#_Toc40981519)

[3.2.2 SURF算法介绍 14](#_Toc40981520)

[3.2.3 SIFT算法介绍 17](#_Toc40981521)

[3.3三种方法的比较与特征点匹配算法的选择 19](#_Toc40981522)

[3.4极线约束(RANSAC)筛选正确匹配点 20](#_Toc40981523)

[3.5奇异值分解求*R, t*矩阵 22](#_Toc40981524)

[第四章 实验 25](#_Toc40981525)

[4.1点云数据的获取 25](#_Toc40981526)

[4.2特征点匹配算法比较 26](#_Toc40981527)

[4.3极线约束RANSAC算法 27](#_Toc40981528)

[4.4奇异值分解求*R, t*矩阵 28](#_Toc40981529)

[第五章 总结与展望 30](#_Toc40981530)

[5.1总结 30](#_Toc40981531)

[5.2展望 32](#_Toc40981532)

[参考文献 33](#_Toc40981533)

[附 录 35](#_Toc40981534)

[致　　谢 72](#_Toc40981557)

# 第一章 绪论

## 1.1课题的研究背景及意义

近年来，随着机器视觉的不断发展，基于点云拼接的三维重建已经发展成为机器视觉非常重要的核心关键技术之一。三维点云拼接技术已经直是人工智能的诸多分支领域中的重要前沿研究主题。例如在某一项目中，根据已知的图形运用相关的技术对目标物体进行影像反求时，因为被测物体的几何形状及测量方式等物体自身硬性条件的限制, 采集勘测设备(如双目相机等)需要从多个角度对目标物体进行造型、定位和测量，然后拼接多组从不同角度测量的点云数据信息，拼接后，所有点云图像将会被整合在一个坐标系中，这就是点云拼接问题。与之对应的三维重建技术,就是在计算机中利用三维数据重建现实中的三维场景或物体,最终在计算机上对其进行仿真模拟，达到一种将真实物体进行缩小存储的效果。通常情况下，三维点云拼接技术也可以被称为点云配准技术, 简而言之它可以将在两个坐标系中测量的两幅点云数据进行转换拼接(*R，t*转换)。一个物体的三维重建，可以通过软件运用几何建模的方式手动制作。国内外相应的建模软件种类颇多，但是这类利用软件来制作三维模型的时间周期长，过程琐碎复杂，并且操作者需要具备一定的专业技术，获取的数据一般用于游戏的动画设计或者三维还原展示等等。对物体进行三维重建也可以使用基于机器视觉原理的方法。它是通过相机来拍摄采集物体的图像数据，之后通过相关技术来达到三维重建的目的，其重建精度和效率都比较高。一般使用三维重建算法比较多的可以分为结构光方法、立体视觉法、SFM、TOF等方法[1- 2]，这些算法可以成功结合到目标设备中方便使用。例如在已经被广泛使用桌面扫描仪中结合了结构光与立体视觉的方法，通过扫描对象的立体模型信息就可以通过扫描设备容对其重建。

三维重建技术在电子竞技游戏以及虚拟现实技术等领域都得到了比较广泛应用[3]，例如美国的微软公司在2016年推出的 Kinect，是玩家在娱乐的同时也锻炼了身体，一经推出就广受好评的游戏产物。同时，三维重建技术在文物保护方面也得到了广泛的应用[4]，比如在唐太宗陵墓的某些出土文物中，我们可以利用三维建模技术，再加上一定的技术手段对已经破碎的器具进行立体拼接，制成立体仿真图像,让博物馆游客可以立体欣赏已经破损的文化瑰宝，欣赏的同时也避免了直接触碰所造成的二次损伤。利用结构光三维重建技术，可以有效地对发掘过程中的“破损”古物进行拼接，可以有效地实现文化遗产的记载、已知古籍的修复、文化遗产瑰宝的保护传承。在生活应用需求上，虚拟试衣设备通过结构光技术收集用户的人体三维数据，从而快速得到人体三维立体模型，消费者可以云端自由选择衣服与型号。最终通过显示屏全方位显示最终效果。三维重建技术同时在工业检测中也发挥着重要作用，例如比亚迪，在进行汽车零件加工时，利用结构光可以快速找出瑕疵品，以减少不必要的损失。在国家促进人工智能与自动控制领域的大前提下，继续对更高精度三维重建的方法进行深层次研究，对进一步促进结构三维重建技术的应用具有重要意义。

## 1.2国内外研究现状

基于相机等硬件平台，在获取被测物体表面各采样点的空间坐标后，可以得到一组点的集合，称为“点云”。它表示的是在同一个空间参考系下目标物体特征与分布的大量点的集合。硬件设备每次采集的数据都是在当前的坐标系下完成的，为了得到相对更加丰富的物体三维数据，我们要在两幅点云之间确定一个相对合适的空间坐标转换的关系，通过转化之后，我们已知的两幅点云就可以成功融合在同一坐标系之中，更好地显现出物体的特性，这就是所谓的三维点云数据拼接。该技术在虚拟技术和影像反求等领域有着十分广泛的应用。国内外目前常见的点云拼接技术一般可以分为两步：粗拼接和精确拼接。顾名思义，前者的意思就是将多个坐标系下的多幅点云数据“大致对准”到同一目标坐标系下,但是结果经常无法满足精度要求，为了获得更优的结果，我们在此基础上进行下一步的优化，类似于显微镜的调节，我们以上述方法所得的结果为基础，再次进行精确拼接，尽量使点云图之间的拼合缝隙达到一个理想值，从而满足一个较高的精度要求。

国内外主要使用的点云拼接算法主要有以下几种：Song. Z等人提出基L型平面的转台标定法[5]：已知转台的转轴中心和方向，每次进行实验的时候，将物体放在转台上，只需要控制转台旋转的角度就能够得到不同视角的点云数据，只要记录转台相应的旋转角度，就能将多个点云一起拼接到同一个坐标系下(见图1-1)。虽然这种方法求取旋转平移矩阵简单，精度和效率都比较高，但是也存在比较大的缺点，被测物体只能在已经标定好的转台上进行测量，要受转台位置的影响，而且待测物的大小重量也要在转台承受范围内，同样受到转台的限制。尽管后来有人对转台进行过修改，但被测物体的形状、大小同样受到限制。罗先波等人采用特征点标记法[6]，比如在被测物体表面粘贴不同颜色小球，两个不同视角中同时出现待测物体上的相同标记点(见图1-2)时，通过在不同坐标中寻找小球位置对应点的方法可以求出点云之间的旋转矩阵分量与平移矩阵分量，实现三维点云数据的自动精确拼接。此方法对于三维特征点不明显的物体效果比较好。但毕竟需要人工参与，会费时费力，尤其待测物相对比较大的时候或者脆性材料时，无法随意粘贴标记点。此外由于标记点的遮挡，重建的模型也会失去完整性。因此在实际操作中此方法同样有很大的局限性。Besl提出了一种点云间的三维配准算法ICP算法[7]，它可以较好的计算使两幅点云之间的欧式距离达到最小的转换关系。近年来ICP算法也得到了很多学者的研究。但是ICP算法依然存在一些问题，首先，它需要一个比较好的初始值，一旦初始值无法确定，迭代结果可能出现问题；此外，ICP算法在处理大量点云数据时肯定会比较缓慢，因此需要事先精简点云，而且计算时两片点云间需要有重叠部分。

|  |  |
| --- | --- |
| 图1-1 转台拼接法 | 图1-2 特征点标记法 |

总之，由于受自身仪器或被测物体的影响，现有的拼接算法都不同程度存在自身局限性，对不同环境要求下的拼接需求经常不能达到。无辅助拼接技术不受设备自身限制，在使用的过程中自动化程度高、适用广，因而成为点云拼接的研究热点。

## 1.3本文工作及内容安排

本文在已有的三维点云拼接基础之上，提出了一种算法简单，实用性强，实时性好的自由视角的三维点云拼接算法。算法通过特征点匹配找到不同角度采集的图像对应匹配点对，再由二维空间映射到三维空间来求取对应的旋转平移矩阵来实现自动拼接。

第一章，绪论部分。主要介绍了三维点云自动拼接算法的研究背景、研究价值、研究意义，以及三维点云自动拼接技术的适用场合。并分析了三维点云拼接技术的国内外研究现状和进展，最后给出了本文主要的工作及文章的结构安排。

第二章，硬件平台与系统软件开发环境。主要介绍了系统硬件系统的主要组成部分与系统软件开发编译平台，并介绍了应用广泛的相机标定方法，以及相机成像模型与标定过程。

第三章，特征点匹配算法与*R, t*矩阵的求解。比较现有常用的特征点匹配算法orb，sift，surf之后，经过严谨的分析，选用orb与sift相结合的算法作为特征点匹配的主要算法。 但是所匹配的特征点对中可能含有部分错误值，因此我们引入RANSAC极线约束从上文的匹配结果中筛选出正确且优质的特征点，进一步对通过已知的特征点对进行奇异性分解经过最小二乘估计迭代求出两幅点云之间的*R, t*矩阵，从而实现三维点云自动拼接。

第四章，实验分析。对前文中提到过的算实现验证。给出两张采集数据图片，利用三维点云自动拼接算法进行试验，对结果进行分析。

第五章，对本文主要工作进行总结。将本文的工作中的各种情况充分分析研究，得出具有指导意义的结论并分析文章中算法的不足之处，对下一步要做的工作进行预测展望。

# 第二章 系统硬件平台与软件开发环境

## 2.1拼接系统硬件组成

根据相关文献可知，在采集物体信息时，非接触式结构光法是一种不与目标物体之间进行直接接触就可以采集到目标物体的深度信息。而且此方法对系统的硬件 (相机，投影仪等) 环境要求相对不高，并且具有成本相对不高、扫描频率快、技术层面也相对容易等优点，因此在实际测量中经常被使用。因此本文利用此方法获取被测物体的三维点云数据。在实验室实际环境中硬件系统主要由投影仪、双目相机、电脑三部分组成(见图2-1)。

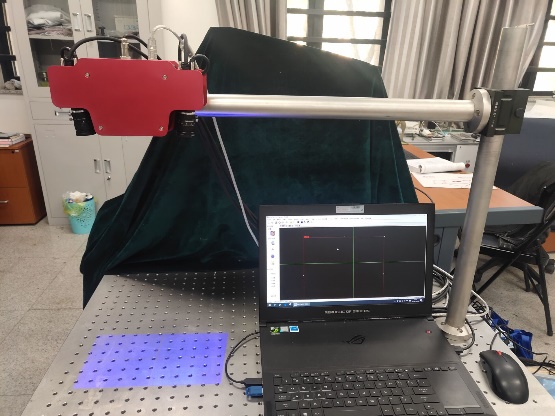




图2-1 硬件系统实物图

系统硬件平台主要实现功能为：

(1)、在接收到相应的信号之后，投影仪就会开始动作，首先投影仪会发出强光原信号，接着该信号会被光学透镜进行转换进而变为一种矩形光栅，接着光栅经过传递又会直接打到被测量物体的表面，由于光栅的作用，被测物体表面就会形成特殊的条纹光。

(2)、因为在进行相机标定的时候，条纹光不能只打几次，为了尽量减小误差，投影仪会打出十八次条纹光，每次都会对相机进行外触发，当全部完成后，此条纹被双目相机拍摄记录且被进一步采集，经过模数转换为目标信号传入目标PC。

(3)、通过计算机软件开发环境编写相关应用程序，对所得到的二维图像信息进行处理，进行一定的技术手段处理之后，可以将该图像的信息转化为被测量物体的三维点云信息，之后经过点云转换关系进行求解。

## 2.2拼接系统软件开发平台

VS2013是微软公司开发的产品，它类似于一个“软件包”,有助于开发者更快、更好、更安全、更便捷的完成软件开发任务。同时是目前比较流行的软件集成开发平台，其安装完成配置成功之后的系统界面如图2-2所示。(图中的软件开发平台是基于DELL灵越7000笔记本，inter i7-7200U处理器的Windows10系统下)它包括多重功能，包含了在软件开发过程中使用到的大部分工具。OpenCV是上世纪末期由Inter公司创建的，它可以在不同系统上运行，是一个开源的计算机视觉库。因为它是由C++语言所构成，所以相对其他库而言效率较高且轻量级，它的出现可以轻松实现机器视觉的相关算法。本文中将利用OpenCV中的算法来实现特征点匹配的相关算法。PCL(点云库)就是建立的一个C++开发库，它纳入了多种操作三维点云数据的处理算法，具有高效数据结构，涉及到了有关点云的方方面面。文中将使用PCL点云库来采集被测物体的三维点云数据。

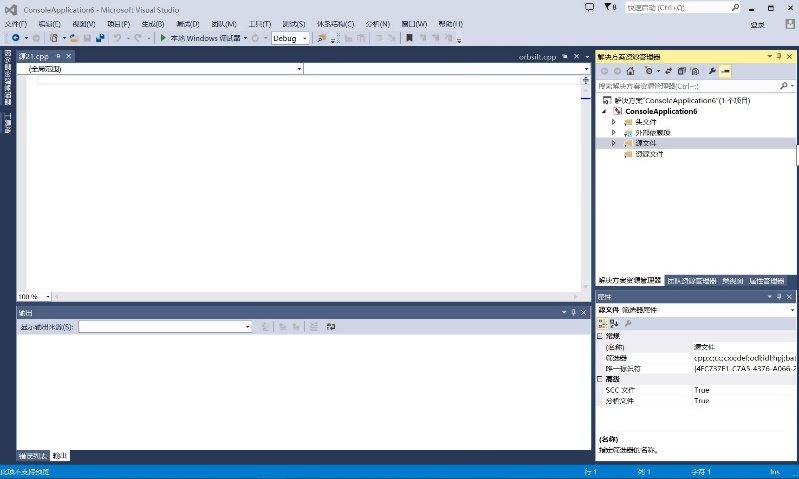


图2-2 开发平台界面

系统软件是基于 Visual Studio 2013平台进行开发，从而实现点云自动拼接的功能，并调用OpenCV库，PCL库来实现算法的基本功能。

## 2.3双目相机的标定

为了获得点云数据就必须进行双目相机的标定。而相机标定的目的就是为了确定相机自身相关的参数，相机参数中包含相机内参，，，，畸变系数，，，，，本文中采用张正友标定法[8]对相机进行标定。

张正友教授提出的张正友标定法是一种采用棋盘格来对相机进行标定的方法，既克服了标定物的高精度要求，又提高了精度，更易于操作。在使用此法进行标定时，设定棋盘格中的角点作为检测对象，坐标用表示，对应图像中的像素点为，坐标为，根据相机模型可得：

**** (2-1)

其中，，分别为相机在*x，y*方向上的等效焦距，*R*、*t*——相机外参数，*R*是3×3旋转矩阵，*t*是3×1平移矩阵，，为平移常量。

由于棋盘格是人为特制的，每个黑白方格的角点是已知的坐标且处于同一平面，因此将****代入公式2-1中，可得：

 (2-2)

其中 分别为R的第*i*列；K为相机内参。

根据平面的单应性，令：

 (2-3)

其中为标量常系数；*H*为单应性矩阵。

这样我们就可以通过单应性矩阵*H*对相机内外参数进行约束，*H*矩阵则可以通过棋盘坐标和图像像素坐标对应点来进行求解，最终通过计算得到相机的参数，基本上实现了相机的标定。

但求得的参数相对而言精度不够，为了对标定的参数进一步修正，拟合出比较完美的系统，张正友还提出了利用最小二乘法进行最大似然估计，来求取非线性方程的最优解，公式为：

 (2-4)

## 2.4双目相机标定结果

我们首先打印棋盘格标定纸，黑白方格边长为30cm，通过双目摄像头对多幅图像进行拍摄以供标定使用。

标定实验过程如下：

(1)、将打印好的棋盘格纸用胶水进行固定，等固定到板上之后，使整个棋盘格位于同一平面。

(2)、每采集一幅图片时都要变换一次相机和标定板之间位置，多次重复。

(3)、对左右多幅图像都要进行角点检测并进行匹配。

(4)、计算得到相机的固定参数和*R, t*参数，就可以完成相机的标定。

标定过程中在棋盘格上采集数据的部分图像(见图2-3)与标定的最终参数结果如表2-1所示：

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |

图2-3 双目相机标定过程部分图像

表2-1 双目相机标定结果

|  |  |  |
| --- | --- | --- |
| 参数 | 左相机 | 右相机 |
| 内参矩阵 |  |  |
| 径向畸变 |  |  |
| 切向畸变 |  |  |
| 外参旋转矩阵*R* |  |  |
| 外参平移矩阵*t* |  |  |

通过张正友标定法将双目相机标定好，就为以后获取点云数据与待拼接物体的二维图像做好了铺垫。

# 第三章 特征点匹配筛选与*R, t*矩阵的求解

## 3.1点云数据的获取

双目视觉相机获取三维点云的基本原理是分别从两个视觉相机的角度来观测相同的目标物体，将左右不同视角的两幅图像的数据都采集到，通过视差原理[9] 计算像素视差，最终就可以获得被测量目标物体的三维点云数据图。本文使用 VS2013为开发平台，结合 Open CV 库和点云库(PLC) 对原始的点云图进行提取。本文中使用的多频外差与立体匹配的方法实现对三维点云信息的提取。

多频外差是基于时间相位展开的算法，此方法在被测物体表面上投影不同频率的条纹光栅，并且在投影过程中，每种频率条纹都要进行相应的相移变换。这种基于时间相位而展开的算法，对每一个像素点相位信息的计算都是独立的，有利于抑制干扰，因此多频外差法也变成了重要的结构光算法之一。该算法对双目相机的标定，点云数据的获取具有重要作用。

立体匹配是根据双目立体视觉实现物体三维重建中复杂度和困难度最高的部分，立体匹配的结果将会直接影响物体三维重建精度与准确度，所以立体匹配算法也就成为了当前学术界中的热议话题。立体匹配顾名思义就是从待匹配的图中找到与参考图中的对应像素点。然后通过左右相机的位置来构建三角测量关系，构建完成之后通过该原理进一步就可以对像素点对应的三维信息进行恢复。立体匹配中存在准则、基元与算法三大主要问题。在基于双目视觉的测量系统中，由于结合了立体视觉测量的原理，必须首先找到左右匹配点，之后就计算空间点的三维坐标之间的转换矩阵参数，进而实现三维点云转换的目的。上述过程会产生一个相对比较完整地，已经通过了欧式距离检验过的匹配点集，之后分别在左右两幅图像中使用RANSAC约束的原则对像点之间进行逐一匹配。这样我们就可以把面匹配转换成线匹配，这样一来运算时复杂度会大大降低，相对于面匹配而言，搜索范围缩小了很多，极大程度上提高了匹配的速度，与此同时提高了匹配精度。

空间中的目标具有明显的特征点，将特征点与物体的空间分布以散点形式表示出来之后，在相机坐标系下，点云图中每个点的信息都代表了图像中对应像素点所对应的三维坐标，之后再将二维点与三维点一一对应起来存储于同一表格中。

## 3.2特征点提取匹配技术

图像中的特征点提取相关算法是机器视觉应用中的基础算法且应用广泛，常用于描述与识别物体、图像反求、三维重建等众多AI相关领域。特征匹配就是提取图像的描述参数或特征属性，用来代表整幅图像的特征属性，通过参数间的相似性配准从而实现原图像间的相互匹配。

所有的特征匹配算法中影响匹配效率和精度的重要因素是特征描述的选择，为了实现稳像算法的实时性条件，需要选用计算相对简单的特征点描述，但是特征点计算的时间复杂度与准确度之间是互斥关系，无法实现双赢。选用运算效率相对高的特征描述，在运算过程中难免会存在一定的错误点对，因此我们需要对所匹配的全部特征点通过一定的技术手段进行再次筛选，就可以剔除了部分误匹配的错误特征点，保留符合筛选条件的 “合格”匹配特征点，同时尽量提高计算效率。接下来文章中将主要介绍并分析现在常用的几个匹配算法，他们分别是ORB算法、SURF算法及SIFT算法。

### 3.2.1 ORB算法介绍

为了解决匹配同名点对数量及单一特征点的分布不均问题，我们引入一种 名为ORB 的特征点检测算法[10]。ORB算法主要是由两部分组成，分别是FAST角点检测方法与 BRIEF特征描述子。但是究其本质而言，ORB算法就是一种“改进版”的FAST算法。ORB算法直接提取特征点可能会出现运算时间慢，误差大等各种各样的问题，因此在对图片特征点进行提取的时候，我们应通过灰度质心法获取主方向，在获取完成之后，基于此再来获取特征描述子，获取描述子的时候由于自身缺点的影响，我们需要对描述子进行改进，通过对某一个点的像素对比可能已经无法满足要求，因此再改进的时候可以通过特征点的某一邻域内进行灰度平均值的方法来改善算法自身的缺点，使其具有旋转不变性，而且对噪声不敏感。

(1)、FAST特征检测

FAST特征点：将图像中的某一点与其领域内像素进行对比，如果该点的像素值比周围连续N(设定阈值)个像素点的值都大个单位或都小个单位，那么这个点就可以称之为FAST特征点[11]。

此邻域大小通常可定义为以3个像素为半径R的圆周形区域，一共16个像素点，N的取值通常为9，称之为FAST-9算法。圆心处的像素点可以通过遍历这16个像素点进行对比，以此来确定目标圆心点是否可以作为FAST特征点进行存储，将运算过程可视化处理为图像，如图3-1所示。

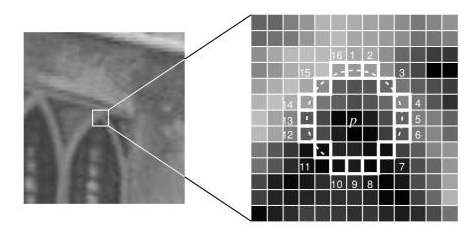


图3-1 FAST特征点检测原理

在实际计算时，一般要选择一幅检测图像中的任意一个像素点作为圆心，在中心像素处，绘制一个半径为3个像素的圆形区域。这就是像素的邻域。如图3-1所示，在图中的点p是要确定的目标，整个灰度包围区域以点p为中心，半径为3 (半径值可根据实际情况进行确定，比如6,9等)的圆邻域，为判定*ｐ*点是否为特征角点，该像素点的邻域中共有16个像素点，如图3-1所示，我们首先可以给出一个阈值*t*，接下来我们需要进行灰度值对比，用*p*点处的值与该像素点邻域内全部的像素点逐一进行对比，若至少有N(自定)个点与*p*点相比超过了*t*，则将该点作为目标数据保存，上述全部过程可以表示为角点响应函数，见公式3-1与3-2。

 (3-1)

 (3-2)

在ORB算法中N值通常取9。具体判定*p*点是否为FAST特征点的步骤：

*a*、首先计算上、下两方向的灰度值与*p*点灰度值之差的绝对值，若两个值都小于*t*，说明该不符合要求，将其进行剔除，否则就将该点作为候选点进行保留，进行下一步计算。

*b*、对于上述筛选通过的候选点*p*，将其再次与1、5、9、13四个点处的灰度值分别进行差值计算，如果差值的绝对值超过3个大于设定阈值则将保留*p*点作为候选点再次进行考察，否则*p*点排除。

*c*、经过前两次的筛选，已经选择出来少数相对正确的点，最后将这些点与邻域全部点之间按照(1)中的方法逐个计算，如果计算结果中75%以上的点都满足*t*约束，则*p*点就会被认定为目标点，否则*p*点排除。

虽然此方法速度快，但是并不具有多尺度特性，ORB算法为保证尺度不变性，要额外的对图像进行高斯金字塔滤波，得到多尺度图像，在我们进行特征匹配时在所有尺度上都要进行FAST特征检测。

(2)、BRIEF特征描述

SURF和SIFT算法的特征描述维度相对极高，每个特征描述符占用字节数相对较大，因此SURF和SIFT特征描述占用了大量的存储空间。BRIEF特征描述二进制表示法为基础，每一位占用1比特，相对而言大大降低空间复杂度[12]。

BRIEF在特征点邻域取一个正方形区域，在这个区域内选取*n*个像素点，然后特征点将会与这*n*个像素点一一进行灰度值比较。我们定义如下公式：

 (3-3)

其中，*X，Y*分别表示特征点的坐标、等待比较的*n*个像素点的坐标；*i*表示BRIEF特征描述中的第*i*个比特位。

一个描述对应的只占用*n*个比特位。*n*常取128、256，(N值的选择与计算机的性能相关，经常在一般计算过程中选为256)因此一个描述占用16、32字节。

BRIEF算子没有旋转不变性，在实际运算中会存在很大的问题，因此ORB算法如果相对其这一功能进行改进，就必须引入特征点的主方向这一步骤，然后旋转坐标系直至与主方向重合。ORB利用灰度质心法来确定特征点主方向，定义的灰度矩如下：

 (3-4)

其中，*x*，*y*为该点邻域内像素点坐标；为该点灰度值。

之后我们就可以利用公式3-4来确定灰度质心，灰度矩心为：

 (3-5)

其中、及分别由公式(3-5)计算可以得到。

因此ORB算法中特征点的主方向角度为：

 (3-6)

BRIEF将*n*个像素点全部进行存储，生成了一个2n的矩阵**Q**：

 (3-7)

之后，我们就可以利用主方向角度对旋转角度进行计算：

 (3-8)

之后将坐标系进行旋转到主方向重合，坐标矩阵为：

 (3-9)

在新坐标系之中，相对于之前BRIEF算子而言，新计算的描述符就具有旋转不变性。因此我们在进行特征点匹配时，使用空间上的间隔关系来计算两个比特串形式的特征描述之间的具体差异，同时统计两比特串中不同比特位的数目，当数目小于特定阈值时，我们就可以认为两者匹配，反之则为不匹配。

### 3.2.2 SURF算法介绍

SURF算法实质上对基于SIFT的 “加速版”，相对于SIFT算法而言，算法的鲁棒性得到了增强，而且SURF是具有尺度不变特性[13] 的局部特征算法。SIFT算法有128维，而SURF算法仅仅具有64维，再进行运算的过程中，其运行速度比起后者快一个数量级，在对多幅图像同时处理的情况下，前者的鲁棒性相对更好，因此近些年来经常取代SIFT实现更快的运算，该算法常被用在物体重建、图像反求等AI相关技术领域。其运算过程主要包含了如下几个部分。

(1)、构造多尺度空间

SURF算法是一种具有尺度不变特性的局部特征算法，在对图像进行处理的时候，在特征点提取之前首先要建立一个尺度空间[14]。SURF算法与其他的算法不同，他不是使用的高斯差分函数，而是Hessian矩阵构建尺度空间。如图3.2所示，为SURF算法尺度图像不变，卷积核尺寸变化情况。

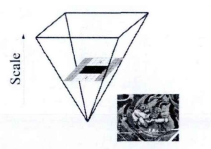


图3.2 尺度空间金字塔结构

为了保证尺度不变性，SURF算法中使不同方差的高斯函数与目标图像之间进行卷积，进而得到不同尺度空间函数，Hessian矩阵为:

 (3-10)

其中，下脚标中的*x, y*表示函数分别对其进行求偏导。

利用*H*矩阵就可以对目标点进行判断，在SURF算法中，不同的octave层图像尺寸是相同的，SURF算法实现多尺度的手段是通过图像卷积与尺度各异的高斯模糊核来实现的，这种方法既实现了算法过程中对采集数据过程的简化，同时又提高了算法的运算效率。

(2)、特征点定位

SURF算法特征点的定位就是采用尺度空间中的27个像素点中最大点或最小点作为特征点。同时利用Hessian矩阵对边沿效应的特征点进行筛选，同时用三维二次函数插值对特征点位置进行准确定位。

(3)、特征点主方向的确定

SURF算法与SIFT算法统计梯度直方图的方式不同，进而导致两张算法之间的运行效率有所不同，SURF是统计指定特征点邻域内Harr小波特征，以划分整个邻域为6个子区域，将扇形子区域内存在的所有特征点的水平、垂直Harr小波特征一起进行加权求和，接着旋转扇形区域，更换子区域按照上述步骤再次计算加权求和值，一直重复求和运算，直至圆形内所有区域都已经进行过了运算，将所有扇形得到的值两两依次进行比较，选择出上述求和值最大的那个扇形，该方向就是所求主方向。如图3-3所示，在图中可以看出我们要选取的方向应该是最右边那幅图中长箭头对应的方向。

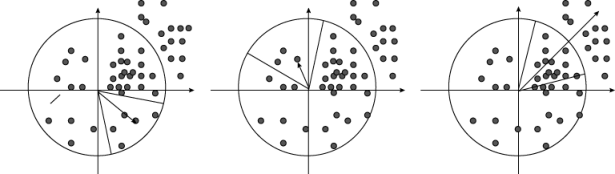


图3-3 求取特征点主方向

(4)、SURF算法的特征描述

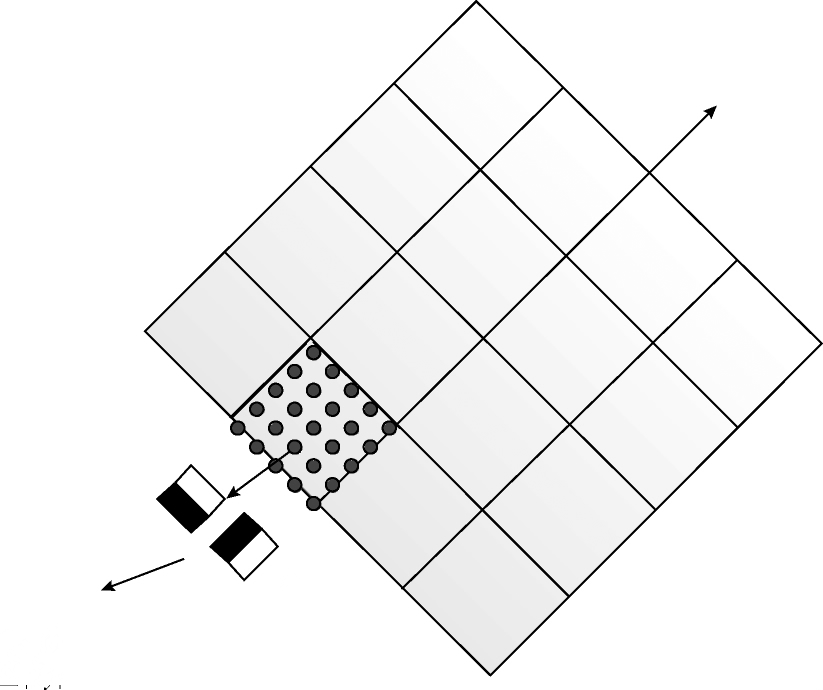
SURF算法的方向有两种因素共同影响，分别是特征点主方向与特征描述算子。后者的步骤主要是先在目标特征点周围截取一个“口”字形区域，然后将该区域进行等额划分，等份分为子区域，加起来一共有16个区域，将它们连接起来可以得到特征描述，因此SURF的描述子为444维。上述过程转换为图像语言如图3-4所示，在每个子区域内共有像素垂直于特征点主方向的、平行于特征点主方向的总计25个Harr小波响应，分别计算他们的绝对值之和与方向之和，并且对、、和进行记录，进而可以得到该特征点的描述符。





图3-4 计算特征描述符

在获得特征描述子之后，SURF匹配算法进而可以采用的欧氏距离进行比较，通过以上步骤我们可以得到一个比值，此值是欧式距离中最近邻和次近邻间距离比值，将此值与预先设定值比较，通过比较结果就可以判断两个特征点之间是否匹配成功。

### 3.2.3 SIFT算法介绍

SIFT算法即尺度不变的特征变换匹配算法，于90年代末由David Lowe提出并进行完善[15]。 它是一种局部特征检测的算法，具有旋转不变性、尺度不变性、光照不变性。SIFT特征点在图片中普遍存在，其检测也相对容易，且不容易受到被测物体自身或者采集设备等硬性因素的影响，因此广泛应用于基于机器视觉的各种任务中。SIFT特征检测主要分为如下几个步骤:

(1)、 构造多尺度空间。

SIFT特征提取算法的第一阶段在多个图像尺度和位置层面上对特征点进行初步的搜索，首先对目标图像进行初步高斯模糊:

 (3-11)

其中，*x, y*分别表示目标像素点的横纵坐标，\*为卷积操作符。

高斯函数[16]表达式为:

 (3-12)

将图像与高斯差分函数卷积就能提高特征点的检测效率，假设已知高斯差分尺度空间:

 (3-13)

其中，*k*表示高斯差分函数的平滑图像效率。

由于采集到的每幅图像都会得到尺度空间函数，所以高斯差分尺度空间可以直接通过两相邻尺度空间作差得出。

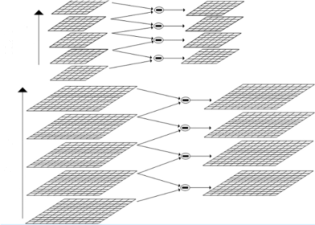


图3-5 高斯差分尺度空间计算图

如图3-5所示，从下往上，原始图像与高斯函数迭代得到左边每大层(first octave)中的每一小层，然后对图像进行处理之后作为另一个大层((next octave),以此类推。右边为相邻两空间作差而得的高斯差分尺度空间。

(2)、局部特征点检测

再构建完高斯差分(DOG)金字塔后，为了检测每层高斯差分金字塔的局部特征点，可以将每层的点和与相邻两层的18个邻接点及邻域点充分对比，如果要想使该点成为局部特征点，此点是必须是全局评比中的最大或最小值。

(3)、精确定位特征点

在我们获得上述目标点后，匹配算法为了提高特征点的噪声鲁棒性，使用拟合的三维二次函数确定尺度与位置信息，要去除对边缘响应敏感的特征点可以利用公式3-14的Hessian矩阵。

 (3-14)

其中，为高斯差分尺度空间对*x*坐标上的二阶偏导数。

设矩阵*He*的最大、最小特征值为，，则：

 (3-15)

 (3-16)

设，则有：

 (3-17)

由上述公式可知，当*r*越大时，说明两特征值比值越大（即为差值较大），所以两个方向的曲率差别也就越大。通常情况之下可以取*r*=10,当的值大于10时，我们就可以认为已经出现了边沿效应。换句话说，当*r*> 10时，为了消除边沿效应而将此位置特征点剔除。

(4)、特征点方向分配

因为必须要保证特征描述符的旋转不变性，所以我们要对特征点主方向进行分配，确定主方向时，以特征点为中心用高斯加权窗口取一邻域，然后以为单位，将邻域进行等分，如图3-6所示，利用直方图对邻域内梯度方向进行统计，与SURF相似的，可找出直方图中最大值，之后将该方向作为特征点主方向。在算法中，只选择一个方向在运算中万一无法查找或无法确定时，可能会造成模型不稳定，因此可以引入辅助方向。给每个特征点添加了一个或多个，这样就可以显著增强算法稳定性。得到主方向后，沿主方向对图像方向校正，为描述子做铺垫。

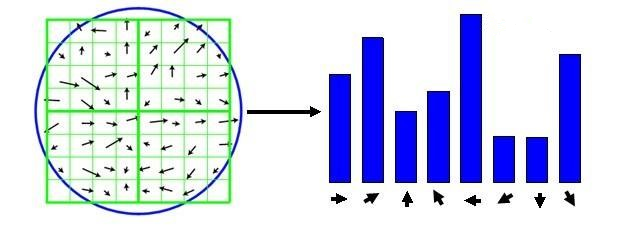


图3-6 确定特征点方向(第五列)

(5)、特征点的特征描述

为了获取特征点的描述子，首先要将目标特征点的邻域划分为44子区域，然后在各子区域内按上、右上、右、右下、下、左下、左、左上总共八个方向(类似于“米”字形)对上述直方图进行统计且排序梯度值，如图3-7所示，之后我们将会获得一个448的特征向量。最后，为了完成图像特征点匹配，每个特征点描述子都要利用欧氏距离特征点进行相似性度量，之后就可以实现特征点匹配。

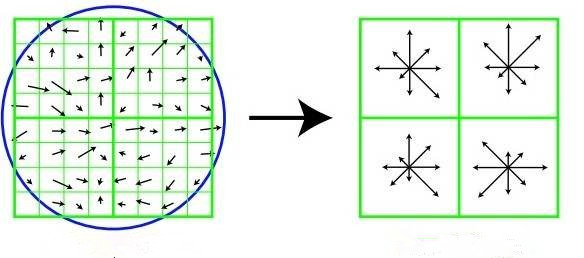


图3-7 图像梯度与关键点描述符生成的特征描述子

## 3.3三种方法的比较与特征点匹配算法的选择

SIFT算法、SURF算法与ORB算法都是机器视觉中常见的特征算法。SIFT算法、SURF算法与ORB之间存在本质上的差别。前者的本质上是由Harris斑点检测算子实现特征点匹配的，而后者是利用FAST角点检测算子实现特征点匹配的，因为角点与斑点表达形式有所区别，一个表达的是交叉点特征，另一个表达的是一片区域，两者的表达区域有所不同，斑点相对更大，因此斑点的抗干扰能力更强，同样的在斑点旋转与尺度不变性方面也优于角点。但FAST角点在计算效率这方面要远大于Harris斑点。SURF是基于SIFT改进得到的，与SIFT算法相比极大地降低了描述子维度，改进了算法的实时性能，提高了算法的计算效率。 ORB算法本身使用的FAST算法本身就具有速度优势，而且运算符描述特征点时，使计算机获得的描述符仅为0、1位字符串，节省了后续的特征点匹配时间，大大改善了操作效率。

总体上来说，SIFT算法鲁棒性最强、精度最高，但该方法运算时间较长。SURF算法是相对SIFT算法的“升级版”，提高了计算效率，但实时性变差。ORB在旋转与尺度不变性，定位精度与鲁棒性上比不上SIFT，但其采用了FAST算法与二进制描述子，因此在使用该方法进行运算时的检测速度大大提高，在特征提取与匹配的速度上相对于另外两种算法而言大幅领先。根据有关资料显示：ORB算法的匹配速度领先SURF算法两位数左右，比较SIFT算法领先三位数左右[10]。

总之，三种算法各有各的长处，同时各有所短，基于此本文采用的基于ORB和SIFT相结合的特征提取匹配算法，相比原本的两种算法而言，特征点提取的准确率有所提升，相比后者而言特征点提取时间大幅缩短，而且相比前者，提取特征点的精度也相对大幅提高，同时也解决了该算法中没有尺度不变性的问题。再实际应用中，相比原始的两种方法再尺度不变性、匹配效率等方面具有更高的效率。

## 3.4极线约束(RANSAC)筛选正确匹配点

为了解决在特征点匹配算法在匹配过程中可能存在的误匹配问题，本文使用了一种基于极线约束RANSAC 算法，对上文中的匹配算法匹配的结果进行优化，得到相对更优，正确的匹配点对。

最初由Fischler提出的用于模型估计的随机采样一致性算法(RANSAC)[17]，该算法既可以实现对图像中对上文中的匹配算法匹配的结果进行优化，也可以估计立体视觉中的基础矩阵。本文采用的ORB算法与SIFT算法结合得到的匹配特征点对中会存在错误匹配点对，所以我们使用一种基于极线约束RANSAC 算法来对上述结果进一步进行优化，得到相对更优的结果。

极线约束属于对极几何的范畴。如图3-8所示，为空间中某三维点，该点在双目相机中的成像位置分别为 *m*、。和  分别为双目相机中心，所以 、和共同构成了一个极平面，该平面与两像平面的交线就称为极线。比如在图中点*m*的极线就是直线，点的极线为是直线。极线与双目相机的中心连线交点和 是极点。

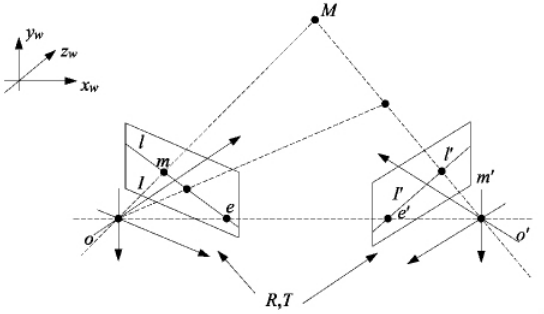


图3-8 极线约束图

RANSAC算法本身就允许数据中存在一部分误差相对较大的信息，通过尽量少的点对模型参数进行估计，然后该模型又会影响尽量多的点。建立参数模型后，在该模型误差范围内的点就称为局内点，反之是局外点。

RANSAC算法的基本流程如下：

(1)、随机采样*a*个点，将这些点拟合成参数模型。

(2)、计算其它点到模型的距离，如果小于阈值*t*就称为内点，之后利用一定的技术手段对内点个数进行统计。

(3)、自行设定一个算法重复次数，重复上述步骤*w*次，之后选择所有模型中内部点数最多的作为最优模型进行存储。

(4)、将上述最优模型的所有内点找出进行存储，用存储的数据对参数模型进行重新估计。

RANSAC算法对匹配结果进行优化，排除相对偏差较大的匹配点的算法核心环节是计算双目相机的基础矩阵*F*，基础矩阵就是像点与立体像对之间的一一对应关系。假设像素点坐标为，基础矩阵*F*定义如下:

 (3-18)

根据极线约束可知, 基本矩阵约束方程为：

 (3-19)

其中，为*m*像素点的坐标，为像素点的坐标。

若要计算F矩阵，至少需要8对特征匹配点，构成一个89的矩阵进行运算求解。但是我们在实际中已知点对数远远大于8对，所以使用而实际给定点数远远大于8对，因此使用RANSAC进行求解。步骤如下:

(1)、从已知匹配点中随机选取8对，以此作为输入参数。

(2)、使用8点法求解出基础矩阵*F*,使其满足奇异值约束条件。

(3)、计算模型到点之间的误差并判断内点，之后统计内点个数。

(4)、迭代*w*次，之后选取满足内点最多的矩阵*F*。

(5)、重新使用该模型中所有内点计算基础矩阵*F*。

使用RANSAC而得到的最后的最优模型中的点即为最优匹配集，之后将这些点对保存下来，其余点不错任何处理，保存完之后我们就相当于得到了正确且优化过的匹配集。

## 3.5奇异值分解求*R, t*矩阵

在点云自动拼接算法中，直接使用所有匹配点对来计算两点云的*R*，*t*矩阵计算量大，费时且复杂。因此我们需要对其进行分解，分解后分批次进行运算，在经过迭代找出最优值，根据相关资料记载，我们可以利用奇异值分解[18]对矩阵分解，奇异值分解(SVD)可用于计算配准旋转平移变换矩阵和点云法线求取等多个过程中。在需要对比较复杂的矩阵进行运算(比如最小二乘)的时候，该方法都有十分重要的应用，显著提高运算效率。

设存在源点云和目标点云，它们是由刚体变换关联的三维对应点，即，在已经被RANSAC筛选之后的匹配集就不存在数据误差时，这些对应点应该满足以下关系:

 (3-20)

设特征标志点集为和，令目标函数为：

 (3-21)

如果满足公式3-21的最小二乘解为，这样经进行变换，变换后的点集为，则与的质心*c*相同，即：

 (3-22)

其中，

 (3-23)

 (3-24)

 (3-25)

令：

 (3-26)

将上式带入目标函数中有：

 (3-27)

目标函数可以进行分解：

 (3-28)

令：

 (3-29)

其中，33的矩阵*H*定义为。

计算*H*的(奇异值分解)SVD分解可得：

*H=UVD* (3-30)

令，则是对称正定矩阵，则对任意33正交阵*B*，有:

 (3-31)

在所有33正交阵中，若其满足，那么*X*就是所求的*R*矩阵。

求得*R*矩阵后，根据公式3-22可知，经过如下转换关系可以进一步求出矩阵*t*:

 (3-32)

最后，因为第一次所得到的结果可能或存在误差，我们采用最小二乘法对其多次优化，得到旋转平移矩阵，之后在程序中使用*R, t*矩阵对点云进行坐标转换，就可以轻松实现三维点云自动拼接。

# 第四章 实验

前面的章节中分别叙述了点云读取算法、特征点提取算法比较，ORB-SIFT算法，极线约束RANSAC算法，SVD奇异性分解算法。本章将会开始对各种算法进行实验论证。文章中所列举的算法在计算机上均是采用C++语言编程实现的，硬件平台为双目相机、上位机、投影仪与笔记本电脑(软件开发平台是基于DELL灵越7000笔记本，inter i7-7200U处理器的Windows10系统上运行的，不同电脑运行速度可能有所差异)。编译开发环境为VS 2013。以及OpenCV计算机视觉库与实现点云通用算法的PCL(点云库)。

## 4.1点云数据的获取

本文中使用的多频外差与立体匹配的方法实现对目标物体的三维点云信息进行提取。点云图中的每个点都不是细腻的，而是代表了对应图像中一一对应的像素点三维坐标数据信息，将同一点的二维信息与三维信息一一对应起来存储于同一表格中。

以一个水杯采集对象，采集纸杯的特征点匹配二维原图(见图4-1)与三维点云信息图(见图4-2)，之后将二维点与三维点云数据一一对应，存储于.csv文件中，其中的部分数据展示于表4-1之中。

|  |  |
| --- | --- |
| 图4-1 待拼接物体采集图 | 图4-2 待测物体的点云图 |

表4-1 二维点与三维点对应数据表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 二维坐标*u* | 二维坐标*v* | 三维坐标*x* | 三维坐标*y* | 三维坐标*z* |
| 1 | 626 | 144 | 26.2534 | 9.17877 | -155.625 |
| 2 | 627 | 144 | 26.1118 | 9.15868 | -155.702 |
| 3 | 628 | 144 | 25.9287 | 9.17126 | -155.668 |
| 4 | 629 | 144 | 25.7831 | 9.1543 | -155.734 |
| 5 | 280 | 133 | 78.531 | 16.4011 | -134.575 |

表格中的数据已经二维点对与三维点对已经一一对应，我们只需要在二维点对中寻找特征点进行匹配，之后在表格中能够找到对应三维点即可。

## 4.2特征点匹配算法比较

通过查找相关文献，我们知道了三种常用算法，ORB算法，SURF算法，SIFT算法，并且了解了三种算法各自的优缺点，使用两张图片进行匹配，在运算速度，匹配效果等多方面进行比较。图4-3为SURF匹配示意图，图4-4为ORB匹配示意图，图4-5为SIFT匹配示意图。

|  |  |  |
| --- | --- | --- |
| 图4-3 SURF匹配示意图 | 图4-4 ORB匹配示意图 | 图4-5 SIFT匹配示意图 |

在实验中得出的规律汇总为表4-2：

表4-2特征点算法优缺点

|  |  |  |  |
| --- | --- | --- | --- |
|  | ORB算法 | SURF算法 | SIFT算法 |
| 特征点数量 | 少 | 一般 | 多 |
| 尺度变换鲁棒性 | 差 | 好 | 一般 |
| 计算速度 | 快 | 一般 | 慢 |

总之，三种算法各有各的优缺点，本文采用的基于ORB与SIFT相结合的特征提取算法，再实际应用中，相比原始的两种方法再尺度不变性、匹配效率等方面具有更高的效率[19]。

## 4.3极线约束RANSAC算法

在图4-6中，我们可以看出特征点匹配算法仅仅进行过一次距离约束时，多多少少存在一定的误匹配(见图4-6)，为了得到正确的匹配集，我们引入了极线约束RANSAC，首先，我们先用ORB-SIFT算法进行匹配，再将匹配之后的匹配集用RANSAC筛选剔除错误特征点对(见图4-7)。使用两张纸杯的图片进行匹配，观察得出结论。

|  |  |
| --- | --- |
| 图4-6 经过欧式距离约束后的匹配示意图 | 图4-7 经过RANSAC后的匹配示意图 |

从图4-6与图4-7中我们可以看出，使用RANSAC而得到了正确且最优的匹配集，同时，我们将最优特征点对找出来，通过二维点对在上表找到对应的三位点对，将数据存于.csv文件中，表格中的部分数据见表4-3中。

表4-3 正确三维匹配点对

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 序号 | 三维点  坐标*x* | 三维点  坐标*y* | 三维点  坐标*z* | 三维匹配  点坐标*x* | 三维匹配  点坐标*y* | 三维匹配  点坐标*z* |
| 1 | 64.070 | -22.309 | -84.920 | 65.270 | -6.482 | -89.693 |
| 2 | 54.328 | -17.584 | -93.454 | 55.910 | -2.128 | -96.624 |
| 3 | 60.196 | -19.210 | -89.785 | 61.544 | -3.638 | -93.525 |
| 4 | 59.730 | -21.421 | -88.871 | 60.906 | -5.668 | -92.703 |
| 5 | 63.569 | -16.465 | -88.969 | 65.042 | -0.820 | -93.294 |
| 6 | 58.859 | -17.549 | -91.560 | 60.340 | -1.966 | -95.092 |

## 4.4奇异值分解求*R, t*矩阵

将上式得到的最优的三维点对匹配集，经过奇异值分解求出旋转平移矩阵参数值，将所求的矩阵参数存于表格4-4中，之后在程序中实现对两幅点云的自动拼接见图4-8。

表4-4 SVD分解的*R, t*矩阵

|  |  |
| --- | --- |
| *R*矩阵 | *t*矩阵 |
|  |  |

|  |  |
| --- | --- |
|  |  |
|  |  |

图4-8 点云拼接图

图4-8中上面是两幅待拼接的点云，下面是拼接之后的点云不同方向的展示图。总共拼接6次左右，就可以完整呈现整个纸杯的三维点云信息。

拼接完成之后，计算点云之间的欧氏距离，计算得到拼接的平均距离为0.0385mm，拼接标准偏差为0.0394mm，通过比较可以得出，本文提出的基于特征点匹配的点云自动拼接算法精度较高，可以满足物体的拼接需求。

# 第五章 总结与展望

## 5.1总结

本文主要对当前主流的三维重建中的三维点云自动拼接技术进行了研究。经过阅读大量国内与国外权威文献，并且研究了目前主流特征匹配算法，经过大量实验论证发现目前的点云拼接算法还不够成熟，有较大的发展空间，特别是实时应用场合最为薄弱。在此基础上对本文的研究重点进行确立：为设计出一种实时应用的通过二维特征点匹配找到对应的三维关系来实现目标物体的三维拼接。本文分别从当前热门的粗配准和精配准分析入手，吸取各种特征点匹配算法的宝贵经验，扬长避短，从而设计出一种基于特征点匹配的三维点云拼接技术。本文通过结合2D特征检测和结构光3D扫描的高精度拼接方法，成功将奇异性分解的方法应用于点云拼接。最后通过编程实验验证了该设计算法的合理性，有效性。最终结果充分展现了算法的实用性。现对本文的主要工作做出如下总结:

(1)、本文首先给出了当前几种常见点三维云拼接方法，包括基于L型平面的转台标定法、特征点标记法、ICP(Iterative Closest Point)算法等，并分析了这些常用方法的长处与不足，提出一种基于二维特征点匹配得到匹配点对，并通过奇异性分解求出来两个点云之间的旋转平移矩阵，再利用一定技术手段完成点云自动拼接的算法，在此基础上，本文首先理论分析论证了算法的整个过程，最后通过实验进一步验证充实。

(2)、特征点提取匹配算法的研究。文章中将最常见的三种算法：SIFT算法、SURF算法与ORB算法进行了比较。对算法的优缺点进行深入分析，发现三种算法即各有千秋又各有所缺。所以本文采用的是基于ORB与SIFT相结合的特征提取算法，相比于原始的两种算法而言，特征点提取的精确度变高，相比后一种算法而言特征点提取速度大幅提高，相比前一种算法而言精度也大幅提高，同时结合之后的算法也解决了前一种算法的尺度变化性问题。再实际应用中，相比原始的两种方法具有更高的效率。

(3)、转换参数方法研究。在点云自动拼接技术中求解旋转矩阵与平移矩阵是十分重要的一步。本文充分分析了奇异值分解法的基本运算原理，最终选择了使用SVD(奇异值分解法)来求解所匹配点云间的旋转矩阵和平移矩阵，之后再使用一定的技术手段即可完成两幅点云之间的拼接，较完整的显示出目标物体三维信息。

## 5.2展望

本文以双目相机为基础，基于特征匹配而实现点云自动拼接，尽管文中的点云拼接算法性能较好，也能较好满足实时应用的需求，但其中依然存在部分问题，本文只是理论上分析了其可靠性和实用性。在工业环境，生活环境中还可能出现许许多多没有考虑到的问题，还存在很多改进之处和需要解决的问题，模型算法需要进一步的完善，下面提出部分建议[20-21]：

(1)、本算法是基于特征点匹配与奇异值分解求R，*t*矩阵的拼接算法，根据两幅图之间的运动矩阵进行拼接，因此会存在一定的误差，后续可以探索误差更小更优的点云拼接方案。

(2)、在待拼接物体是动态的情况下，双目相机采集的数据肯会失真，导致无法提取特征点或者提取的特征点有所偏差。

(3)、待测物体的图像特征是基于ORB和SIFT相结合的算法进行提取的，对于表面颜色和纹理同时单调的物体可能会不适用。

(4)、由于实际生活中的应用场合具有多样性，实验时能全部考虑到不现实，因此该算法仍要被更多实际场景检验，经过不断完善，成为一种通用实时算法。

参考文献

1. Crandall David J, Owens Andrew, Snavely Noah.*et al*. SFM with MRFs: discrete-continuous optimization for large-scale structure from motion. IEEE transactions on pattern analysis and machine intelligence,2013,35(12).
2. 雷禧生，肖昌炎，蒋仕龙.基于TOF相机的喷涂工件在线三维重建.电子测量与仪器学报，2017，31(12)：1991-1998.
3. Pyo Y, Tsuji T, Hashiguchi Y, *et al*. Immersive VR interface for informationally structured environment. IEEE, International Conference on Information and Automation 2015: 1766-1771.
4. 刘小通.面向文物复原的残差网络识别与多特征拼接技术研究：[硕士学位论文].陕西：西北大学，2019.
5. Ye Y, Song Z. An accurate 3D point cloud registration approach for the turntable-based 3D scanning system. IEEE International Conference on Information and Automation, 2015: 982-986.
6. 罗先波，钟约先，李仁举.三维扫描系统中的数据配准技术.清华大学学报(自然科学版)，2004(08)：1104-1106.
7. Rusu R B, Blodow N, Beetz M. Fast Point Feature Histograms (FPFH) for 3D registration, IEEE International Conference on Robotics and Automation, 2009: 1848-1853.
8. 许佩.结构光三维扫描系统多视角自动拼接算法研究：[硕士学位论文].广西：桂林电子科技大学，2019.
9. 翟志强，杜岳峰，朱忠祥等.基于Rank变换的农田场景三维重建方法.农业工程学报，2015，31(20)：157-164.
10. Rublee E, Rabaud V, Konolige K, *et al*. ORB: An efficient alternative to SIFT or SURF. International Conference on Computer Vision. 2012, 1(4): 5-8.
11. Jiao J, Zhao B, Wu S. A speed-up and robust image registration algorithm based on FAST[C], Computer Science and Automation Engineering (CSAE), IEEE, 2011:160-164.
12. Calonder M, Lepetit V, Strecha C, *et al*. BRIEF: Binary Robust Independent Elementary Features. Computer Vision, 2010:778-792.
13. 葛盼盼，陈强，顾一禾.基于Harris角点和SURF特征的遥感图像匹配算法.计算机应用研究，2014，31(07)：2205-2208.
14. 金斌英.基于改进的SURF图像配准算法研究.计算机测量与控制，2019，27(11)：228-232.
15. Divya S V, Sourabh Paul, Umesh Chandra Pati. Structure tensor-based SIFT algorithm for SAR image registration. 2020, 14(5):929-938.
16. 谢亚晋.基于ORB特征的电子稳像技术研究：[硕士学位论文].浙江：浙江大学，2018.
17. 张培耘，华希俊，夏乐春等.基于RANSAC算法的极线约束立体视觉匹配方法研究.组合机床与自动化加工技术，2013(11)：20-22.
18. Wei Xue, Yan Luo, Yue Yang, et al. Noise Suppression for GPR Data Based on SVD of Window-Length-Optimized Hankel Matrix. Sensors (Basel, Switzerland)，2019，19(17).
19. 葛山峰,于莲芝,谢振.基于ORB特征的目标跟踪算法.电子科技,2017(2) .
20. 苏志部.大尺寸测量的点云自动拼合技术研究：[硕士学位论文].湖北：华中科技大学,2009.
21. 方莹.三维点云自动拼接算法研究：[硕士学位论文].黑龙江：哈尔滨工程大学,2016.

# 附 录

英文文献翻译

## **Automatic reconstruction method for large scene based on multi-site point cloud stitching**

Haonan Xu, Lei Yu, Junyi Hou, Shumin Fei

School of Mechanical and Electric Engineering, Soochow University, Suzhou, China

School of Automation, Southeast University, Nanjing, China

## **Abstract**

At present, the three-dimensional (3D) reconstruction system is based on hand-held camera which causes certain problems, such as needing for a large number of human-computer interactions, demanding on high quality image data and insufﬁcient accuracy of 3D reconstruction. Aiming at these problems, a fully automatic reconstruction of large scene based on multi-site point cloud stitching is presented in this paper. The proposed method uses the Kinect sensor for image acquisition. Since, there are multiple sites in the room, each site of image data is processed using a separate model in order to get good 3D point cloud data. Then, these sites are used to constitute a local area network, and the method of bundle block adjustment is employed to stitch each site point cloud data. The proposed method achieves a high-degree automation and provides a high-precision 3D reconstruction which has two main advantages: (i) the reconstruction process is fully automatic, without any human-computer interaction; (ii) the automatic reconstruction is robust. Experimental results show that proposed automatic reconstruction method is convenient and practical, and can provide better 3D reconstruction model than commonly used methods. Moreover, it can be applied to virtual reality shopping malls, Virtual Reality (VR) and other ﬁelds.

**Keywords**: Large scene three-dimensional; reconstruction; Kinect sensor; Bundle adjustment; Point cloud registration; Automatic reconstruction method.

## **1. Introduction**

The three-dimensional reconstruction and display of scenes are widely used in many areas, such as design simulation, virtual reality and human-computer interaction. The existing three-dimensional reconstruction technology is mainly based on use of hand-held camera to obtain data for three-dimensional reconstruction. However, this method is not easy to operate, and quality of obtained image data and effect of three-dimensional reconstruction are not satisfactory. If a three-dimensional reconstruction model for photos and video is easy to obtain and works well, then that model could be applied to many areas, such as business website platform, virtual reality shopping malls and other shopping malls. Therefore, the achievement of a fully automated reconstruction and an ideal three-dimensional reconstruction model is very important goal of computer vision.

The current automated reconstruction system is mainly based on hand-held camera, and it provides good results in reconstruction of small scenes, but with slower and large human-computer interaction characteristics [1–8,23]. Henry et al. used the Kinect to implement an interactive 3D reconstruction system that selects only key frames for iterative closest point (ICP) cloud registration [2]. They designed a laboratory-wide three-dimensional model, but the model is rough and not complete enough, thus further improvement of accuracy and algorithm performance is needed. Shahram Izadi et al. constructed the Kinect-Fusion system, using Kinect for three-dimensional reconstruction of small-scale scenes and individual objects [3]. Compared to method presented by Henry et al., the results are more elaborate, but because almost all operations are done in the display and limited size of memory, the rebuild range and reconstruction accuracy cannot be good simultaneously, which makes reconstruction difficult. In view of the shortage of Kinect-Fusion system, Raul Mur-Artal et al. pro-posed a real-time reconstruction method named the simultaneous localization and mapping based on oriented brief feature recognition (ORB-SLAM) which uses parallel tracking and mapping (PTAM) architecture [4]. The ORB-SLAM adds the function of map initialization and closed-loop detection, and optimizes the method of key frame selection and map construction. It achieves good results in processing speed, tracking effect and map precision. However, ORB-SLAM still has some shortcomings [9–13]. The whole SLAM system is calculated by feature points; therefore, it is necessary to determine the ORB features for each image, which is time-consuming while the map construction is sparse [14,19]. The most of three-dimensional reconstruction methods for indoor scenes need a hand-held camera for shooting, moreover the operation is not convenient. The amount of human-computer interaction is large when a hand-held camera is used to acquire data. The quality of the data acquired by the camera will directly affect the results of the 3D reconstruction. The model obtained by the method of holding the camera for 3D reconstruction is not ideal.

As already mentioned, in current three-dimensional reconstruction of large scenes, the degree of automation is low and the reconstruction model is not ideal. This paper presents a fully automated reconstruction system for multi-site point cloud stitching intended for large scenes. The main feature of this work includes: (i) This paper presents a fully automated data acquisition platform, and uses advanced calibrated camera and rotating head for image acquisition. (ii) In the room with multiple sites, each site of image data is processed using an independent adjustment method to obtain a three-dimensional point cloud data. (iii) Each site constitutes a local area network and the bundle block adjustment method is used for each site point cloud data to get a three-dimensional reconstruction model and to achieve full automation.

The outline of this paper is as follows. In Section 2, the mathematical model of point cloud stitching and optimization is introduced, including mathematical model of single point cloud, multiple point cloud stitching and bundle block adjustment. Section 3 is devoted to the solution of mathematical model of cloud stitching and its optimization. Section 4 includes simulation results to illustrate feasibility and effectiveness of the proposed strategy. Finally, we give a conclusion in Section 5.

## **2. Mathematical model of point cloud stitching and optimization**

### 2.1 Mathematical model of single point cloud

In order to obtain the three-dimensional reconstruction model of a large scene, this paper uses a certain number of sites to form images, and then splices them into site point cloud [24–25]. In a single site point cloud, this paper uses the method of rotating head. The camera placed on the rotating head rotates with a step of to collect images, thus 36 pairs of color images and depth images are obtained. The data obtained by the scan are not in the same coordinate system, thus they need to be converted to the same coordinate system. In order to unify the data of each group, the space coordinate of the coordinate system is switched to the ﬁrst coordinate system, and it needs to go through (n-1) rotation and translation, but the excessive number of rotations and translations causes the accumulative error. In order to get a good model, the error should be evenly distributed, so the independent adjustment should be used [16–17,20–22].

The adjustment is based on determination of parameters of adjacent model, achievement of seven spatial similarities of parameters of three corner elements of exterior orientation (angle of rotation of the X, Y, and Z coordinate axes), namely, and K, three translations and a scale of the zoom:

 (1)

where, ,,,,,,,,andare the direction cosines of the function of corner elements, and K;,, andare the translations of the coordinate origin in three coordinate directions; andis the scaling factor. Then, we introduce the corrections of X, Y and Z, and coordinate each point of gravity. Therefore, the error equation is deﬁned by:

 (2)

where, is the correction of the observed value of;, ,,,,, andare the corrections of the myopic value of parameter to be determined;,, andare the coordinates of gravity center of spatial point in coordinate system to be rotated, and,, andare the constant terms of error equation.

 (3)

Then, (1) and (2) are combined to determine two adjacent images of transformation parameters. Using thepicture that corresponds to the camera coordinate system, we convert thepicture and thepicture to the point of the picture to the camera coordinate system that corresponds to thepicture.

### 2.2 Mathematical model of multiple point cloud stitching

Once we have obtained the 3D model data of one site, we can use the same method to get the data of the second site. The stitching of sites also uses coordinate transformation to unify the data, i.e. to translate the data into the same coordinate system. But due to the site increase, the error gradually accumulates, the reconstructed 3D model drifts and so on. In this paper, The idea of photogrammetry is applied. The site layout is the same as in the navigation area, and the model is optimized by constructing the bundle block adjustment method. The relationship between two adjacent sites based on the three-dimensional coordinate conversion model is as follows:

 (4)

where, ,andrepresent the translation parameters, which represent the three-dimensional coordinates of the scanning site translated into the reference coordinates of location, and they are the point cloud positioning parameters. And R matrix reflects the automatic orientation of the point cloud. According to the nature of Rodriguez matrix, the site is directed to its adjacent site h with an unknown number of adjustment equations. Based on (4), we get:

 (5)

where, I is the third order unit matrix, is the adjustment value of the antisymmetric matrix of thesite,,, andare the measured coordinates of the site h, ,, andare the coordinates of thesite scanning position,,, andrepresent the adjusted scan coordinates of thesite; and k is the scaling factor.

Since the above formula can be obtained for two adjacent sites point cloud stitching, then we can use the same method to get multi-point site cloud stitching, such that between multiple sites a local area network is constituted. And the bundle block adjustment is used to deal with each site point cloud stitching.

### 2.3 Mathematical model of bundle block adjustment

In the construction of bundle block adjustment for point cloud splicing optimization, thesite is used as a bundle adjustment of the base site. The bundle is adjusted as a single bundle with the radiation from the site adjacent to the site, and each bundle is used as the adjustment unit of the bundle block adjustment. The three-dimensional coordinate transformation model is used to establish a unified error equation in the block net, such that the bundle pointing to the same site is completely intersected, and then seven coordinate transformation parameters of the adjacent site transformation are determined. All bundles constitute block net, and the entire area of the block net can be classified as the public site, with one or more adjacent sites pointing to the public site. Here, the and the sites are directed to the public site *k*. The sites that point to the same public site can be used for public site k adjustment.

Then, if theand thesites point to the public site *k*, we can write the following:

 (6)

where, ,, andare the measured coordinates of thesite for the public site *k*; is the adjustment of the rotation matrix of thesite; ,, andare the scan coordinate adjustments of thesite to the public site *k*; , , andrepresent the site scanning position coordinates; is the adjustment of the rotation matrix of the site; ,, andare the scan coordinate adjustments of thesite to the public site *k*; andis the adjustment of the scaling parameter. Equation (6) can be rewritten as:

 (7)

Then, we get the point cloud stitching and its optimization mathematical model. And the solution of the mathematical model will be presented in the following.

## **3. Solution of mathematical model of cloud stitching and its optimization**

In order to facilitate the calculation, the subscripts *i* and *h* in (5) are omitted. Therefore, the linearization equation containing the unknowns of two adjacent sites is defined by:

 (8)

Then, we determine the positioning of the site and orientation parameter , parameter initialization, parameter correction value, coordinates of the public site center in the specified coordinate system , coordinates of the coordinate system, andcoefficient of the coordinate correction of a common site in the specified coordinate system.

Further:

 (9)

where, is the derivative matrix of function *F* for the specified coordinate system derivative where ,,and.The coordinate in coordinate system of the same public site corrects the coefficient as follows:

 (10)

And the orientation parameter is changed to the coefficient as follows:

 (11)

By combining Eqs. (10) and (11), we get:

 (12)

The close difference is :

 (13)

where.

The linearization equation (7) is defined by:

 (14)

where, andare the coefficients related to the coordinates coefficients,  andof the site and the site to the common site, , . andthe corrections of the and sites orientation parameters  and , respectively. The and coefficients are:

 (15)

The coefficient of the scaling parameter is:

 (16)

The correction factor of a is defined by:

 (17)

And the correction factor of b is defined by:

 (18)

Then, the correction factor of c is:

 (19)

 (20)

The linearization defined by (13) is the linearization of the mathematical model using the bundle block adjustment. Also, this linearization equation can be obtained by using the corresponding correction parameters.

## **4. Experiment results**

In the experiments, the Microsoft’s Kinect (RGB-D) camera was used to obtain data, and experimental platform consisted of Windows 10 (64-bit) system, Inter I7-7700HQ, graphic card Gtx980, and 12G memory. The amount of human-computer interaction is large when a hand-held camera is used to acquire image data for three-dimensional reconstruction. And 3D reconstruction model is not ideal. On the other hand, when the method described in this paper was used, and camera was placed on the rotating platform, and rotated with a step of, thus, 36 pairs of images were obtained per one full rotation. At each , the camera took only one pairs of pictures, thus we performed a week long scan to obtain 36 sets of data. After the data of the first site was obtained, we moved the device and obtained the data of the second site. The distance between two adjacent sites was 1.5 m–2 m. In the case of narrow aisle or house doors, the site layout needed to be encrypted. Due to the small number of pictures, a single site and two sites point cloud stitching were immediately available. Namely, if two sites point cloud stitching was intact, the subsequent site data were continued by device movement, otherwise the data of current site were retrieved. The above operation was repeated until all data within the site were obtained. Consequently, the ideal three-dimensional reconstruction model was obtained.

The three-dimensional reconstruction of laboratory scene obtained by commonly used hand-held camera is presented in Fig. 1. As it can be seen in Fig. 1, when the improved densely constructed ORB-SLAM system is used for 3D reconstruction [4], the 3D reconstruction model is incomplete because some indoor scenes have not been scanned. Therefore, this method is not convenient for human-computer interaction, and the effect of 3D reconstruction is not ideal. Consequently, it is more difficult to generate the 3D reconstruction model of a large scene.

The commonly used point cloud registration and stitching method reconstruction model of point cloud map is presented in Fig. 2a [15,18]. Namely, the internal and external parameters of camera are used to calculate the position of pixels in the world coordinate system to splice the point cloud model to form a single site. When the method described in Section II was used, the point cloud model shown in Fig. 2b was obtained which the independent adjustment method was used for each site point cloud stitching. Firstly, we determined the transformation parameters of two adjacent images,



Fig. 1. The hand-held camera for three-dimensional reconstruction of the point cloud map.

|  |  |
| --- | --- |
| Fig. 2a. The commonly used point cloud reconstruction model of point cloud map. | Fig. 2b. The independent adjustment method used to reconstruct the point cloud map. |

and then all coordinate system points were unified into one coordinate system using. By contrast, we can see that the independent adjustment method can be a good single point of the point cloud data, Fig. 2b. Ordinary methods can easily fail to stitch, so the model has only a flawed corner or a model stratified situation. Therefore, the independent adjustment method can be used to reconstruct a better single-site point cloud, and the bundle block adjustment correction processing can get better results.

The mathematical model of two-site splicing described in Section 2.2 was used and the splicing mathematical model was linearized by the method described in Section 3.1. The coefficient of coordinate corrections andof a common site in the specified coordinate system and the coefficient of the directional parameter Bwere obtained. Using the correction factor on two sites point cloud stitching we got the point cloud model shown in Fig. 3, where the positions ① and ② represent the location of the site.

With the increase of site, the cumulative error also increased. If the commonly used point cloud reconstruction method is used to match, that would easily lead to model confusion. As shown in Fig. 4a, in the results of commonly used point cloud stitching, there are two doors. At the upper right corner of the local enlarged view you can see two doors more clearly, and the right room is also moving to the right. When the traditional bundle adjustment method is used to optimize point cloud stitching for multiple sites, the resulting model is shown in Fig. 4b. It can be seen that the traditional bundle adjustment method can better solve the serious drift of the model. However, there are still some point cloud drifts in the reconstructed model. As shown in a partial enlarged view of the model reconstructed in the upper right corner of Fig. 4b, the point cloud of the cabinet still has a more serious drift. When the mathematical model of bundle adjustment described in Section 2.3 was used and the mathematical model was linearized, we got the correction parameters of the splicing model, and using the corrected parameters we carried out the point cloud model splicing. The resulting model is shown in Fig. 4c, wherein the position of the serial number represents the location of each map site. From the partial enlarged view of the model reconstructed in the upper right corner of Fig. 4c, it can be known that the model reconstructed in this paper is more accurate. In contrast, it can be found that when the beam splicing model was used

|  |  |
| --- | --- |
| Fig. 3. Two-site splicing point cloud map. | Fig. 4a. Results obtained without the bundle adjustment method. |
| Fig. 4b. Results obtained with the traditional bundle adjustment method. | Fig. 4c. Results obtained with the bundle block adjustment method. |

to construct the local area network, and the corrected parameters were used to solve the situation of the point cloud mismatch, the error during splicing process increased. Hence, the ideal three-dimensional reconstruction model was obtained.

Finally, 3D reconstruction model of all sites by using traditional bundle adjustment is obtained is shown in Fig. 5a. As shown in Fig. 5a, when using traditional bundle adjustment, the reconstructed model is severely drifting and difficult to identify. When all sites are reconstructed using the method proposed in this paper, the obtained three-dimensional model is shown in Fig. 5b. The location of serial number in Fig. 5b is the location of the site. In addition, using the fully automated reconstruction system of multi-site point cloud proposed in this paper, the reconstruction effect of a new house model is well obtained.

In this experiment, time between splicing, picture acquisition time and point cloud stitching between sites are shown in Table 1. After a large number of experiments, wherein the proposed method for three-dimensional reconstruction was used, the site point cloud three-dimensional reconstruction success rate was as shown in Table 2. According to the data in Table 2, we can use the independent adjustment method to carry on the stitching of a single site, and the bundle adjustment method can be used to obtain the ideal model. Moreover, the performances of a single site cloud stitching and final global point cloud reconstruction success rate are high.

|  |  |
| --- | --- |
| Fig. 5a. 3D reconstruction model of all sites by using traditional bundle adjustment. | Fig. 5b. The complete house model point cloud map by using our method. |
|  |  |

Table 1 The site reconstruction time.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Site | Splicing site (s) | Picture acquisition time (s) | Point cloud stitching calculation time (s) | Total time (s) |
| ① | / | 36 | 34.1 | 70.1 |
| ② | ① | 36 | 78.2 | 114.2 |
| ③ | ② | 36 | 77.6 | 113.6 |
| ④ | ③ | 36 | 78 | 114 |
| ⑤ | ① | 36 | 77.5 | 113.5 |
| ⑥ | ⑤ | 36 | 79.2 | 115.2 |
| ⑦ | ⑤ | 36 | 78.8 | 114.8 |
| ⑧ | ⑦ | 36 | 78.9 | 114.9 |
| ⑨ | ⑧ | 36 | 78 | 114 |
| ⑩ | ⑧ | 36 | 78.2 | 114.2 |
| ⑪ | ① | 36 | 78.1 | 114.1 |
| ⑫ | ⑪ | 36 | 78.3 | 114.3 |
| ⑬ | ⑫ | 36 | 78.3 | 114.3 |
| ⑭ | ⑪ | 36 | 77.9 | 113.9 |
| ⑮ | ⑭ | 36 | 78.6 | 114.6 |
| ⑯ | ⑮ | 36 | 78.8 | 114.8 |
| ⑰ | ⑯ | 36 | 77.6 | 113.6 |
| ⑱ | ⑰ | 36 | 78.3 | 114.3 |

Table 2 The three-dimensional reconstruction success rate.

|  |  |  |  |
| --- | --- | --- | --- |
| Point cloud | Total  number of  experiments | Number of  successful  experiments | Three-dimensional  reconstruction  success rate |
| Single site | 1000 | 985 | 98.5% |
| Two sites | 1000 | 982 | 98.2% |
| Global site | 1000 | 978 | 97.8% |

The experimental results show that the automatic image acquisition can be achieved by rotating head automatically, and the single point cloud can be obtained quickly. We conducted 1000 experiments on the Single site, Two sites, and Global site, respectively. We judge the success of 3D reconstruction by whether the model has a serious drift and does not meet the demand. Among them, the number of successes and success rates are shown in Table 2. The success rate is high, and the ideal single site point cloud is obtained. Each site constitutes a local area network and uses bundle adjustment method to deal with point cloud stitching, but the bundle adjustment method requires a good initial value. The error of independent adjustment used to get the single point cloud data is small, therefore the bundle adjustment method is used, which can effectively reduce the error transmission in order to achieve multi-site point cloud stitching and to reconstruct a large scene using the ideal three-dimensional reconstruction model.

## **5. Conclusions**

This paper presents a fully automatic reconstruction system based on a large-scene multi-site point cloud stitching, which is fast, robust and has high precision. The pre-calibrated camera and rotating head are used for image acquisition. By arranging multiple sites, the data of the whole scene are obtained. The single site point cloud is obtained by the independent adjustment that splices the processing, and the global site constitutes a local area network and adopts the bundle block adjustment method for global optimization. This system has simple operation, and can easily obtain the 3D reconstruction model of a large scene, whereas 3D reconstruction result is ideal. Therefore, the proposed fully automatic reconstruction system based on a large-scene multi-site point cloud stitching can be widely used in VR showings, business web site platform, virtual reality shopping malls and so on.

However, the proposed method has some limitations. During the shooting process the height of camera is fixed, so access to the panoramic information is relatively small. Therefore, the use of three cameras could present a better solution. Further, the calibration between three cameras and the calibration between camera and rotating head directly affect the reconstruction results. Accordingly, in order to get more ideal 3D reconstruction model of a large scene further study is needed.

## **Acknowledgements**

The work is supported by the National Natural Science Foundation of China (Nos. 61403268; 61873176). The Key Technology Pro-gram of Suzhou, China (SYG201639); Natural Science Fund for Colleges and Universities in Jiangsu Province (16KJB120005).

中文翻译：

## 基于多点云拼接的大场景自动重建方法

徐浩南，于雷，侯俊义，费树民

苏州大学机电工程学院，苏州

东南大学自动化学院，南京

## 摘 要

目前，三维(3D)重建系统是基于手持相机的，这会引起某些问题，例如需要大量的人机交互，对高质量图像数据的要求以及3D重建的准确性不足。针对这些问题，提出了一种基于多点云拼接的大型场景全自动重建方法。该方法利用运动传感器进行图像采集。由于房间里有多个站点，为了得到好的三维点云数据，每个站点的图像数据都是使用单独的模型进行处理的。然后，利用这些站点组成一个局域网，采用束块平差的方法对每个站点的点云数据进行拼接。该方法实现了高度自动化，提供了高精度的三维重建，具有两个主要优点：(i)重建过程是全自动的，不需要任何人机交互；(ii)自动重建是鲁棒的。实验结果表明，本文提出的自动重建方法方便实用，能提供比常用方法更好的三维重建模型。此外，它还可以应用于虚拟现实购物中心、虚拟现实(VR)等领域。

**关键词**: 大型场景三维重建； Kinect传感器；束调节；点云注册；自动重建方法

## 1. 简介

场景的三维重建和显示在设计仿真、虚拟现实、人机交互等领域得到了广泛的应用。现有的三维重建技术主要是利用手持摄像机获取数据进行三维重建。但该方法操作不方便，得到的图像数据质量和三维重建效果都不理想。如果一个照片和视频的三维重建模型容易获得并且工作良好，那么这个模型可以应用到很多领域，例如商业网站平台、虚拟现实商场等。因此，实现一个完全自动化的重建和一个理想的三维重建模型是计算机视觉非常重要的目标。

目前的自动重建系统主要基于手持式摄像机，对小场景的重建效果较好，但具有较慢和较大的人机交互特性[1-8,23]。Henry等人使用Kinect实现了一个交互式3D重建系统，该系统只选择关键点进行迭代最近点(ICP)云重建[2]。他们设计了一个实验室范围的三维模型，但是模型比较粗糙，不够完整，需要进一步提高准确性和算法性能。Shahram Izadi等人构建了Kinect- fusion系统，利用Kinect对小尺度场景和单个对象[3]进行三维重建。与Henry等人提出的方法相比，结果更加精细，但由于几乎所有的操作都是在显示和内存有限的情况下完成的，所以重建范围和重建精度不能同时很好，这使得重建困难。针对Kinect-Fusion系统的不足，Raul Mur-Arta等人提出了一种基于面向短特征识别(oriented brief feature recognition, ORB-SLAM)的实时重建方法，该方法采用并行跟踪与映射(parallel tracking and mapping, PTAM)架构[4]。ORB-SLAM增加了地图初始化和闭环检测功能，优化了关键的选择和地图构建方法。在处理速度、跟踪效果、地图精度等方面均取得了较好的效果。然而，ORB-SLAM仍然有一些缺点[9-13]。整个SLAM系统是通过特征点来计算的，因此需要确定每个图像的ORB特征，这是一个耗时的过程，而地图的构建是稀疏的[14,19]。大多数室内场景的三维重建方法都需要手持摄像机进行拍摄，而且操作不方便。当使用手持摄像机获取数据时，人机交互的数量是很大的。摄像机采集到的数据的质量将直接影响到三维重建的结果。通过手持相机进行三维重建得到的模型并不理想。

如前所述，目前大型场景的三维重构自动化程度较低，重构模型也不理想。提出了一种适用于大场景的多站点的点云拼接的全自动重建系统。本工作的主要特点包括:(1)介绍了一种全自动数据采集平台，采用先进的标定摄像机和旋转头进行图像采集。(2)在多站点房间中，对每个站点的图像数据采用独立的平差方法进行处理，得到三维点云数据。(3)每个站点构成一个局域网，对每个站点的点云数据采用束块调整法，得到三维重构模型，实现完全自动化。

本文概述如下。第二节介绍了点云拼接优化的数学模型，包括单点云、多点云拼接和束块平差的数学模型。第三节研究了点云拼接的数学模型求解及其优化。第四部分是仿真结果，说明了该策略的可行性和有效性。最后，我们在第5节给出结论。

## 2. 点云拼接优化的数学模型

### 2.1 单点云的数学模型

为了获得大场景的三维重建模型，本文利用一定数量的场地形成图像，然后将其拼接成场地点云[24-25]。在单点云中，本文采用旋转头部的方法。放置在旋转头上的相机以为一个步骤旋转以收集图像，从而获得36对彩色图像和深度图像。扫描得到的数据不在同一坐标系中，需要转换成同一坐标系。为了统一每组的数据，将坐标系的*n*次空间坐标转换为第一坐标系，需要进行(n-1)旋转和平移，但旋转和平移次数过多会造成累积误差。为了得到一个好的模型，误差应该均匀分布，所以应该使用独立的平差[16-17,20-22]。

调整是在确定相邻模型参数的基础上，实现外方位三角要素参数的七个空间相似性 (X、Y和Z坐标轴的旋转角度), 即, 和K，三个平移和缩放比例：

 (1)

其中,,,,,,,和是角元素函数的方向余弦，, 和K是坐标原点在三个坐标方向上的平移；,, 和是缩放因子。然后，我们引入X，Y和Z的修正，并对每个重力点进行坐标。因此，误差方程定义如下：

 (2)

其中，为观测值的修正；, ,,,,, 和为待测参数近视值的修正；,和为待旋转坐标系中空间点的重心坐标，,, 和为误差方程的常数项。

 (3)

然后，结合(1)和(2)确定两幅相邻图像的变换参数。使用与相机坐标系相对应的图片，我们将图片和图片转换为图片的点到与图片相对应的相机坐标系。

### **2.2** 多点云的数学模型

一旦得到了一个场地的三维模型数据，就可以用同样的方法得到第二个场地的数据。站点的拼接也使用坐标变换来统一数据，即将数据转换为相同的坐标系。但由于场地的增加，误差逐渐积累，重建的三维模型漂移等。本文应用摄影测量的思想。场地布置与航区相同，通过构造束块平差法对模型进行优化。基于三维坐标转换模型的两个相邻场地之间的关系如下：

 (4)

其中，,和表示平移参数，表示扫描点的三维坐标转换为位置的参考坐标，为点云定位参数。R矩阵反映了点云的自动定向。根据Rodriguez矩阵的性质，用一个未知的调整方程将点定向到其相邻的*h*点。根据(4)，我们得到：

 (5)

其中，I为三阶单位矩阵，为场地反对称矩阵的平差值，,, and为场地h的实测坐标，,, and为场地扫描位置的坐标，,, and表示场地的平差扫描坐标，*k*为标度系数。

由于上述公式可用于两个相邻站点的点云拼接，因此我们可以使用相同的方法得到多个站点的点云拼接，从而在多个站点之间构成一个局域网。并利用束块平差处理各站点的点云拼接。

### 2.3 束块平差的数学模型

在点云拼接优化的束块平差构造中，场地作为基准场地的束平差。利用邻近场地的辐射对束进行单束调整，每束作为束块调整的调整单元。利用三维坐标变换模型在块网中建立统一的误差方程，使指向同一点的束完全相交，进而确定相邻点变换的7个坐标变换参数。所有束线构成块网，块网的整个区域可以划分为公共站点，一个或多个相邻站点指向公共站点。这里，和站点指向公共站点*k*。指向同一公共站点的站点可以用于公共站点*k*的调整。

然后，如果和站点指向公共站点*k*，我们可以编写以下内容：

 (6)

其中，,和为公共场地*k*的场地测量坐标；为场地旋转矩阵的调整；,和为场地对公共场地*k*的扫描坐标调整；, 和为场地扫描位置坐标；为场地旋转矩阵的调整，,和是第*j*个站点到公共站点*k*的扫描坐标调整；是缩放参数的调整。公式(6)可重写为：

 (7)

然后，得到了点云拼接及其优化数学模型。数学模型的解将在下面给出。

## 3. 云缝数学模型的求解及优化

为了便于计算，省略(5)中的下标*i*和*h*。因此，包含两个相邻场地未知量的线性化方程定义如下：

 (8)

然后，确定场地的位置和方位参数、参数初始化、参数修正值、公共场地中心在指定坐标系中的坐标、坐标系的坐标、公共场地在指定坐标系中的坐标修正系数。

进一步:

 (9)

其中，为指定坐标系导数函数*F*的导数矩阵式,和中，且同一公共场所坐标系中的坐标修正系数如下：

 (10)

方向参数改为系数如下：

 (11)

通过组合等式。(10) 然后(11)，我们得到：

 (12)

最重要的区别是：

 (13)

其中，.

线性化方程(7)定义如下：

 (14)

其中，和是与坐标系数相关的系数，, ，和以及场地和场地到公共场地的系数。和位置和取向参数和的修正。和系数为：

 (15)

标度参数的系数为：

 (16)

*a*的修正系数定义为:

 (17)

*b*的修正系数定义为:

 (18)

之后，*c*的修正系数定义为:

 (19)

 (20)

由(13)定义的线性化是使用束块调整对数学模型进行线性化。同时，利用相应的校正参数可以得到线性化方程。

## 4. 实验结果

实验采用微软公司的Kinect（RGB-D）摄像机进行数据采集，实验平台由Windows 10（64位）系统、Inter I7-7700HQ、图形卡Gtx980和12G内存组成。手持式摄像机采集图像数据进行三维重建时，人机交互量大。三维重建模型也不理想。另一方面，当采用本文所述的方法时，将摄像机置于旋转平台上，以步为单位旋转，每旋转一次可获得36对图像。每台相机只拍摄一对照片，因此我们进行了为期一周的扫描以获得36组数据。在获得第一个站点的数据之后，我们移动设备并获得第二个站点的数据。两个相邻站点之间的距离为1.5 m–2 m。如果是窄通道或房门，则需要加密站点布局。由于图片数量较少，一个站点和两个站点的点云拼接立即可用。也就是说，如果两个站点的点云拼接完好无损，则后续的站点数据通过设备移动继续进行，否则将检索当前站点的数据。重复上述操作，直到获得现场所有数据。从而得到理想的三维重建模型。

图1显示了用常用的手持摄像机获得的实验室场景的三维重建。如图1所示，当使用改进的密集结构ORB-SLAM系统进行三维重建时[4]，由于一些室内场景尚未扫描，三维重建模型是不完整的。因此，该方法不便于人机交互，三维重建效果也不理想。因此，生成大型场景的三维重建模型变得更加困难。

点云地图常用的点云配准和拼接方法重建模型如图2a所示[15,18]。即利用摄像机的内外参数计算像素在世界坐标系中的位置，将点云模型拼接成单个场地。当使用第二节所述的方法时，得到图2b所示的点云模型，该模型对每个站点的点云拼接使用独立的调整方法。首先，我们确定了两幅相邻图像的变换参数，然



图1 用于点云地图三维重建的手持式摄像机

|  |  |
| --- | --- |
| 图2a 点云地图常用的点云重建模型 | 图2b 用于重建点云图的独立平差方法 |

后，使用将所有坐标系点统一为一个坐标系。相比之下，我们可以看到，独立平差方法可以很好的将单点的点云数据进行平差，图2b,普通方法很容易缝合失败，因此模型只有一个有缺陷的角点或模型分层的情况。因此，采用独立平差方法可以重建出较好的单点云，而束块平差校正处理可以取得较好的效果。

采用第2.2节所述的双位点拼接数学模型，并采用第3.1节所述的方法将拼接数学模型线性化。得到了指定坐标系下的坐标修正系数、同一场地的坐标修正系数和方向参数系数和。利用两站点点云拼接的校正因子，得到了如图3所示的点云模型，其中位置①和②表示站点的位置。

随着场地的增加，累积误差也随之增大。如果采用常用的点云重建方法进行匹配，很容易导致模型混乱。如图4a所示，在常用的点云拼接的结果中，有两个门。在局部放大视图的右上角，您可以更清楚地看到两扇门，并且右侧房间也正在向右移动。当使用传统的束平差方法对多个站点的点云拼接进行优化时，得到的模型如图4b所示，可见传统的束平差方法能较好地解决模型的严重漂移问题。然而，重建后的模型中仍然存在一些点云漂移。如图4b右上角重建的模型的局部放大图所示，机柜的点云仍然有更严重的漂移。当采用2.3节所述的束平差数学模型并将该数学模型线性化后，得到拼接模型的校正参数，并利用校正参数进行点云模型拼接。结果模型如图4c所示，其中序列号的位置表示每个地图站点的位置。从图4c右上角重建的模型的局部放大图可以看出，本文重建的模型更加精确。相比之下，当使用光束拼接模型时为了构建局域网，采用修正后的

|  |  |
| --- | --- |
| 图3 两站点拼接点云图 | 图4a 不使用束调整方法获得的结果 |
| 图4b 使用传统束调整方法获得的结果 | 图4c 使用束块调整方法获得的结果 |

参数来解决点云失配的情况，增加了拼接过程中的误差。从而得到理想的三维重建模型。

最后，利用传统的束平差得到了各场地的三维重建模型，如图5a所示，采用传统束平差时，重建模型漂移严重，难以识别。当使用本文提出的方法重建所有场地时，得到的三维模型如图5b所示。图5b中序列号的位置是场地的位置。此外，利用本文提出的多点云全自动重建系统，较好地获得了一个新的房屋模型的重建效果。

在本实验中，拼接间隔时间、图像采集时间和站点间的点云拼接如表1所示。经过大量实验，采用本文提出的三维重建方法，场地点云三维重建成功率如表2所示。根据表2中的数据，可以采用独立平差法对单个场地进行拼接，采用束平差法得到理想模型。此外，单点云拼接和最终全局点云重建的成功率都很高。

|  |  |
| --- | --- |
| 图5a 使用传统束平差的所有位置的三维重建模型 | 图5b 用我们的方法绘制的完整的房屋模型点云图 |

表1 场地重建时间

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 场地 | 拼接位置 (s) | 图像采集时间 (s) | 点云拼接计算时间(s) | 总时间 (s) |
| ① | / | 36 | 34.1 | 70.1 |
| ② | ① | 36 | 78.2 | 114.2 |
| ③ | ② | 36 | 77.6 | 113.6 |
| ④ | ③ | 36 | 78 | 114 |
| ⑤ | ① | 36 | 77.5 | 113.5 |
| ⑥ | ⑤ | 36 | 79.2 | 115.2 |
| ⑦ | ⑤ | 36 | 78.8 | 114.8 |
| ⑧ | ⑦ | 36 | 78.9 | 114.9 |
| ⑨ | ⑧ | 36 | 78 | 114 |
| ⑩ | ⑧ | 36 | 78.2 | 114.2 |
| ⑪ | ① | 36 | 78.1 | 114.1 |
| ⑫ | ⑪ | 36 | 78.3 | 114.3 |
| ⑬ | ⑫ | 36 | 78.3 | 114.3 |
| ⑭ | ⑪ | 36 | 77.9 | 113.9 |
| ⑮ | ⑭ | 36 | 78.6 | 114.6 |
| ⑯ | ⑮ | 36 | 78.8 | 114.8 |
| ⑰ | ⑯ | 36 | 77.6 | 113.6 |
| ⑱ | ⑰ | 36 | 78.3 | 114.3 |

表2 三维重建成功率

|  |  |  |  |
| --- | --- | --- | --- |
| 点云 | 实验总次数 | 成功试验数量 | 三维重建成功率 |
| 单场地 | 1000 | 985 | 98.5% |
| 双场地 | 1000 | 982 | 98.2% |
| 总场地 | 1000 | 978 | 97.8% |

实验结果表明，通过自动旋转头部可以实现图像的自动采集，快速获得单点云。我们分别在单站点、两站点和全局站点上进行了1000个实验。我们通过模型是否存在严重的漂移和不满足要求来判断三维重建的成功与否。其中，成功次数和成功率见表2。成功率高，得到了理想的单点云。每个站点组成一个局域网，采用束平差法进行点云拼接，但束平差法需要一个好的初值。单点云数据的独立平差误差较小，因此采用束平差方法，可以有效地减少误差传递，实现多点云拼接，利用理想的三维重建模型重建大场景。

## 5. 结论

本文提出了一种基于大场景多点云拼接的全自动重建系统，该系统速度快、鲁棒性强、精度高。图像采集采用预先标定的摄像机和旋转头。通过多个站点的排列，得到整个场景的数据。单站点的点云由拼接处理的独立平差得到，全局站点构成局域网，采用捆绑块平差方法进行全局优化。该系统操作简单，可以很容易地获得大场景的三维重建模型，而三维重建效果理想。因此，本文提出的基于大场景多点云拼接的全自动重建系统可以广泛应用于虚拟现实展示、商业网站平台、虚拟现实购物中心等领域。

但是，该方法存在一定的局限性。在拍摄过程中，摄像机的高度是固定的，因此获取全景信息的机会相对较小。因此，使用三个摄像头可以提供更好的解决方案。另外，三个摄像机之间的标定以及摄像机与旋转头之间的标定直接影响重建结果。因此，为了得到更理想的大场景三维重建模型还需要进一步的研究。

## 参考文献

1. A.J. Davison, I.D. Reid, N.D. Molton, *et al*., MonoSLAM: real-time single camera SLAM, IEEE Trans. Pattern Anal. Mach. Intell. 29 (6) (2007) 1052–1067.
2. P. Henry, M. Krainin, E. Herbst, *et al*. RGB-D mapping: using depth cameras for dense 3D modeling of indoor environments. In: Proceedings of the 12th International Sympsium on Experimental Robotics. Springer Derlin, India: IEEE, 2010.
3. Izadi S, Kim D, Hilliges O, Molyneaux D, et al. Kinect Fusion: real-time 3D reconstruction and interaction using a moving depth camera. In: Proceedings of the 2011 Annual ACM Symposium on User Interface Software and Technology. Santa Barbara, CA: ACM, 2011.559-568.
4. R. Mur-Artal, J.M.M. Montiel, J.D. Tardós, ORB-SLAM: A versatile and accurate monocular SLAM system, IEEE Trans. Rob. 31 (5) (2015) 1147–1163.
5. F. Endres, J. Hess, J. Sturm, et al., 3-D Mapping with an RGB-D Camera, IEEE Trans. Rob. 30 (1) (2014) 177–187.
6. Z. Xiao, L. Jin, D. Yu, et al., A cross-target-based accurate calibration method of binocular stereo systems with large-scale field-of-view, Measurement 43 (6) (2010) 747–754.
7. R. Hyunchul, J. Jinyong, C. Younggun, et al., Accurate mobile urban mapping via digital Map-Based SLAM, Sensors 16 (8) (2016) 1315.
8. P. Henry, M. Krainin, E. Herbst, et al., RGB-D mapping: Using Kinect-style depth cameras for dense 3D modeling of indoor environments, Int. J. Rob. Res. 31 (5) (2014) 647–663.
9. C.H. Jun, J. Kang, S. Yeon, et al., Towards a realistic indoor world reconstruction: preliminary results for an object-oriented 3D RGB-D mapping, Intell. Autom. Soft Compute. (2016) 1–12.
10. X. Gao, T. Zhang, Robust RGB-D simultaneous localization and mapping using planar point features, Rob. Auton. Syst. 72 (2015) 1–14.
11. A. Barmpoutis, Tensor Body: Real-Time reconstruction of the human body and avatar synthesis From RGB-D, IEEE Trans. Cybern. 43 (5) (2013) 1347.
12. C. Kim, R. Sakthivel, C.W. Kyun, Unscented Fast SLAM: a robust and efficient solution to the SLAM problem, IEEE Trans. Robotics 24 (4) (2008) 808–820.
13. J. Tong, J. Zhou, L. Liu, et al., Scanning 3D full human bodies using Kinects, IEEE Trans. Visual Comput. Graphics 18 (4) (2012) 643–650.
14. B. Triggs, P.F. Mclauchlan, R.I. Hartley, et al., Bundle adjustment- a modern synthesis. Vision Algorithms: Theory and Practice, Springer, Berlin Heidelberg, 1999, pp. 298–372.
15. K. Malowany, K. Magda, J. Rutkiewicz, et al., Measurements of geometry of a boiler drum by time-of-flight laser scanning, Measurement 72 (2015) 88–95.
16. K.H. Lee, J.N. Hwang, G. Okopal, et al., Ground-Moving-Platform-Based human tracking using visual SLAM and constrained multiple kernels, IEEE Trans. Intell. Transp. Syst. 17 (12) (2016) 3602–3612.
17. C.D. Herrera, J. Kannala, Heikkilä J. Joint, Depth and color camera calibration with distortion correction, IEEE Trans. Pattern Anal. Mach. Intell. 34 (10) (2012) 2058–2064.
18. M. Bueno, H. González-Jorge, J. Martínez-Sánchez, et al., Evaluation of point cloud registration using Monte Carlo method, Measurement 92 (2016) 264–270.
19. K.S. Arun, T.S. Huang, S.D. Blostein, Least-Squares fitting of two 3-D point sets, IEEE Trans. Pattern Anal. Mach. Intell. 9 (5) (1987) 698–700.
20. A.H. Ahmadabadian, R. Yazdan, A. Karami, et al., Clustering and selecting vantage images in a low-cost system for 3D reconstruction of texture-less objects, Measurement 99 (2017) 185–191.
21. G. Bresson, T. Féraud, R. Aufrère, et al., Real-Time monocular SLAM with low memory requirements, IEEETrans.Intell.Transp.Syst.16(4)(2015)1827–1839.
22. T. Zhang, J. Liu, S. Liu, et al., A 3D reconstruction method for pipeline inspection based on multi-vision, Measurement 98 (2017) 35–48.
23. R. Maier, J. Sturm, D. Cremers, Submap-Based bundle Adjustment for 3D Reconstruction from RGB-D Data, Lect. Notes compute. Sci. 8753 (2014) 54–65.
24. L. Zhou, X. Zhang, B. Guan, A flexible method for multi-view point clouds alignment of small-size object, Measurement 58 (2014) 115–129.
25. C. Shi, L. Zhang, H. Wei, et al., Attitude-sensor-aided in-process registration of multi-view surface measurement, Measurement 44 (4) (2011) 663–673.

## 致 谢

这项工作得到了国家自然科学基金（61403268；61873176）的资助。苏州市重点技术项目（SYG201639）；江苏省高校自然科学基金（16KJB120005）。

# 致　　谢

随着这篇本科毕业设计论文的完成，我的四年大学生活也即将划上一个圆满的句号。下一段征途即将揭晓，回想起刚刚入学时对大学生活的无限憧憬到课堂上对各位老师学术学识的深沉沉湎，一切中的一切仿佛历历在目，让人倍感留恋。

四年天津工业大学的生活将成为我人生中的一段重要旅程，四年来，我的学长、同学给予我的关怀与帮忙，使我终身收益。在此我向所有的帮助过我的人表达真挚的谢意。宋丽梅老师作为我的指导老师，她治学严谨，学识渊博，视野广阔，为我营造了一种良好的学术氛围。感谢我的导师宋丽梅与研究生黄学长的亲切关怀和耐心指导，黄学长对本课题做了不少工作，给予我不少的帮忙。感谢一起完成毕业设计论文的同学对我的帮助，感谢四年来各位老师教我知识、引我成长，也感谢百忙之中抽出时间进行论文审阅和答辩的老师们。

在论文即将完成之际，我的情绪无法平静，论文的完成标志着快要离开美丽的工大校园，内心中充满了不舍与感激。在这篇论文的写作中，深感自我学术水平十分有限，所写论文难免有不足之处，恳请各位老师和同学批评指正。