



浙江工业大学

# 硕士学位论文

基于伪随机编码结构光的三维重建系统设计与实现

作者姓名	蒋传鹏
指导教师	赵云波 教授
学科专业	控制工程
学位类型	工程硕士
培养类别	全日制专业学位硕士
所在学院	信息工程学院

提交日期：2020年01月

# Design and Implementation of 3D Reconstruction System Based on Pseudo-random Coded Structured Light

Dissertation Submitted to

**Zhejiang University of Technology**

in partial fulfillment of the requirement

for the degree of

**Master of Engineering**



by

**Chuan-Peng JIANG**

Dissertation Supervisor: Prof. Yun-Bo ZHAO

Jan., 2020

## 浙江工业大学学位论文原创性声明

本人郑重声明：所提交的学位论文是本人在导师的指导下，独立进行研究工作所取得的研究成果。除文中已经加以标注引用的内容外，本论文不包含其他个人或集体已经发表或撰写过的研究成果，也不含为获得浙江工业大学或其它教育机构的学位证书而使用过的材料。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。本人承担本声明的法律责任。

作者签名：蒋传鹏

日期：2019年12月

## 学位论文版权使用授权书

本学位论文作者完全了解学校有关保留、使用学位论文的规定，同意学校保留并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。本人授权浙江工业大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

本学位论文属于 1、保密□，在一年解密后适用本授权书。

2、保密□，在二年解密后适用本授权书。

3、保密□，在三年解密后适用本授权书。

4、不保密。

(请在以上相应方框内打“√”)

作者签名：蒋传鹏

日期：2019年12月

导师签名：赵云林

日期：2019年12月

中图分类号 TH741

学校代码 10337

UDC 681.2

密级 公开

研究生类别 全日制专业型硕士研究生



浙江工业大学

硕士学位论文

基于伪随机编码结构光的三维重建系统设计与实现

Design and Implementation of 3D Reconstruction System  
Based on Pseudo-random Coded Structured Light

作者姓名 蒋传鹏

第一导师 赵云波 教授

学位类型 工程硕士

学科专业 控制工程

培养单位 信息工程学院

研究方向 三维重建

答辩日期： 2019 年 11 月 26 日

## 基于伪随机编码结构光的三维重建系统设计与实现

### 摘 要

随着计算机视觉的不断发展，基于视觉的三维重建技术引起了学术界和产业界的广泛关注，目前在三维检测、智能机器人视觉、3D 打印等领域有着广泛的应用。基于伪随机编码结构光的三维重建在精度、实时性、鲁棒性等方面都具有一定的提升空间，因此对于该三维重建技术的研究具有重要意义。

本文针对近距离的物体进行三维重建，从算法、软件和硬件三个方面对基于伪随机编码结构光的三维重建系统进行分析与设计。该系统使用投影机主动投射伪随机编码图案，双目相机获取图像后对物体进行三维重建；该系统的最佳工作距离约为 350mm，帧率约为 15fps，三维重建的点间距的精细度范围是 0.2mm 到 3mm。本文的主要工作包括以下四个方面：

(1) 研究了伪随机编码的生成方法，设计并生成了符合本文要求的伪随机编码。本文设计的伪随机编码大小为 $320 \times 147$ ，由 7 个大小为 $320 \times 21$ 的伪随机编码拼接而成。在大小为 $320 \times 21$ 的伪随机编码区域中，任意 $5 \times 5$ 的区域具有唯一性。采用多个伪随机编码图案拼接的方案使得在极线搜索时保证图像块的灰度值在该极线上具有唯一性的同时，降低伪随机编码的难度。

(2) 制作了高精度的标定板，多次标定双目相机提高标定精度。将反光标记点贴在表面材质为漫反射的玻璃板上从而形成了高精度的标定板。基于 MFC 框架开发了相机内参比较器软件，可直观比较前后两次相机标定结果的差异。

(3) 设计了一套三维重建的软硬件系统。该硬件系统由两个工业相机、LED 补光灯和投影机等组成，具有轻巧、灵活、便携等特点。基于 MFC 框架开发的软件系统，实现了相机标定、硬件调试、实时显示三维点云等功能，方便了三维重建算法的验证。

(4) 本文的三维重建系统，改进了常规的三维重建算法，在精度、实时性和稳定性方面都有显著提高。在立体匹配的过程中使用了基于伪随机编码图案、图像金字塔和模板匹配等方法，提高了立体匹配时的精度和效率；利用双边滤波的图像预处理方法，提升了三维重建算法的稳定性；在算法的实现过程中，为了提升三维重建的实时性，采用了多线程和 GPU 加速方法。

**关键词：** 伪随机编码，结构光，相机标定，立体匹配，三维重建

# DESIGN AND IMPLEMENTATION OF 3D RECONSTRUCTION TECHNOLOGY BASED ON PSEUDO-RANDOM CODED STRUCTURED LIGHT

## ABSTRACT

With the continuous development of computer vision, vision-based 3D reconstruction technology has attracted extensive attention from academia and industry. It is widely used in 3D inspection, intelligent robot vision, 3D printing and other fields. The three-dimensional reconstruction of pseudo-random coded structured light has certain improvement space in terms of precision, real-time and robustness. Therefore, it is of great significance for the research and analysis of three-dimensional reconstruction technology based on pseudo-random coded structured light.

In this paper, three-dimensional reconstruction of close-range objects is carried out. The three-dimensional reconstruction system based on pseudo-random coded structured light is analyzed and designed from three aspects: algorithm, software and hardware. The system uses the projector to actively project the pseudo-random coding pattern, and the binocular camera acquires the image and reconstructs the object in three dimensions. The optimal working distance of the system is about 350mm, the frame rate is about 15fps, and the fineness range of the three-dimensional reconstruction point spacing. It is 0.2mm to 3mm. The main work of this paper includes the following four aspects:

(1) The generation method of pseudo-random coding is studied, and the pseudo-random coding conforming to the requirements of this paper is designed and generated. The pseudo-random code size designed in this paper is  $320 \times 147$ , which is composed of 7 pseudo-random codes with a size of  $320 \times 21$ . In a pseudo-random coding region of size  $320 \times 21$ , any  $5 \times 5$  region is unique. The scheme of splicing multiple pseudo-random coding patterns is combined to ensure that the gray value of the image block is unique in the pole line while the pole line search is performed, and the difficulty of the pseudo-random coding is reduced.

(2) A high-precision camera calibration plate was produced, and the binocular camera was calibrated multiple times to improve the calibration accuracy. The reflective markings are attached to a glass plate whose surface material is diffusely reflected to form a high-precision calibration plate. Based on the MFC framework, the

camera internal reference comparator software was developed to visually compare the difference between the two camera calibration results.

(3) A set of hardware and software systems for 3D reconstruction was designed. The hardware system consists of two industrial cameras, LED fill light and projector, which are light, flexible and portable. The software system developed based on MFC framework realizes camera calibration, hardware debugging, real-time display of 3D point cloud and other functions, which facilitates the verification of 3D reconstruction algorithm.

(4) The three-dimensional reconstruction system of this paper improves the conventional 3D reconstruction algorithm and has significant improvements in accuracy, real-time and stability. In the process of stereo matching, the methods based on pseudo-random coding pattern, image pyramid and template matching are used to improve the accuracy and efficiency of stereo matching. The image preprocessing method using bilateral filtering improves the stability of the 3D reconstruction algorithm. In the implementation process of the algorithm, in order to improve the real-time performance of 3D reconstruction, multi-threading and GPU acceleration methods are adopted.

**KEY WORDS:** pseudo-random coding, structured light, camera calibration, stereo matching, 3D reconstruction

## 目 录

摘 要.....	I
ABSTRACT.....	II
插图清单.....	VII
表格清单.....	VIII
第一章 绪论.....	1
1.1 课题的研究背景及意义.....	1
1.2 国内外研究现状.....	2
1.2.1 基于结构光的三维重建研究现状.....	2
1.2.2 立体视觉三维重建算法研究现状.....	4
1.3 本文内容安排.....	5
第二章 三维重建实验系统.....	7
2.1 相机的选型.....	7
2.1.1 相机的硬件结构.....	7
2.1.2 相机参数.....	7
2.1.3 相机型号.....	9
2.2 伪随机编码结构光的设计.....	10
2.2.1 伪随机序列的生成.....	10
2.2.2 伪随机阵列生成.....	11
2.3 LED 补光.....	13
2.4 硬件结构介绍.....	14
2.5 软件介绍.....	16
2.6 本章小结.....	17
第三章 基于伪随机编码结构光的三维重建.....	18
3.1 算法流程.....	18
3.2 相机标定.....	19
3.2.1 相机标定参数.....	19
3.2.2 相机模型.....	19
3.2.3 标定板制作.....	21
3.2.4 标定算法.....	22
3.2.5 标定实验.....	24
3.2.6 标定结果评价.....	25
3.3 立体匹配.....	27

3.4 三维重建.....	31
3.4.1 伪随机码结构光图案区域分割.....	31
3.4.2 对应点搜索.....	33
3.4.3 剔除误匹配.....	35
3.5 点云拼接.....	35
3.5.1 构建点云拓扑关系.....	36
3.5.2 点云配准.....	37
3.6 三角网格化.....	38
3.7 本章小结.....	38
<b>第四章 基于实时性与精度的研究.....</b>	<b>39</b>
4.1 图像预处理与增强.....	39
4.2 图像金字塔.....	40
4.2.1 模板匹配.....	40
4.2.2 亚像素优化.....	40
4.3 极线搜索优化.....	42
4.3.1 预计算确定平均深度.....	43
4.3.2 使用 ZNCC 快速算法.....	43
4.3.3 使用多线程方法加速.....	44
4.4 本章小结.....	44
<b>第五章 三维重建系统的精度分析与实验.....</b>	<b>45</b>
5.1 精度分析.....	45
5.1.1 理论分析.....	45
5.1.2 实验验证.....	45
5.2 不同场景下重建效果分析.....	48
5.2.1 平面物体的三维重建.....	49
5.2.2 类圆柱体的三维重建.....	50
5.3 本章小结.....	51
<b>第六章 总结与展望.....</b>	<b>52</b>
6.1 总结.....	52
6.2 展望.....	53
<b>参考文献.....</b>	<b>54</b>
<b>致 谢.....</b>	<b>57</b>
<b>作者简介.....</b>	<b>58</b>
1 作者简历.....	58
2 参与的科研项目及获奖情况.....	58
3 发明专利.....	58

学位论文数据集.....59

## 插图清单

图 1-1 普通二值码 .....	3
图 1-2 伪随机编码的设计方案 .....	4
图 2-1 MV-VDM130SC/SM 相机.....	9
图 2-2 线性反馈移位寄存器 .....	10
图 2-3 状态转移图 .....	11
图 2-4 伪随机编码图案 .....	13
图 2-5 伪随机编码图案 .....	13
图 2-6 LED 与镜头的位置关系示意图.....	14
图 2-7 三维重建系统的硬件结构示意图 .....	15
图 2-8 三维重建系统的硬件结构图 .....	15
图 2-9 软件界面 .....	16
图 3-1 算法流程 .....	18
图 3-2 针孔相机模型 .....	20
图 3-3 圆点标定板 .....	22
图 3-4 全局摄影测量仪 .....	22
图 3-5 标定原图像 .....	24
图 3-6 标定图像处理 .....	25
图 3-7 内参比较器 .....	27
图 3-8 相机图像 .....	28
图 3-9 对极约束 .....	28
图 3-10 人体雕像部分图像 .....	31
图 3-11 图像灰度值分布 .....	32
图 3-12 伪随机码区域分割 .....	33
图 3-13 左右图像对应点与极线 .....	33
图 3-14 ZNCC 相关性系数.....	34
图 3-15 人体雕像的三维点云 .....	34
图 3-16 去除误匹配的三维点云 .....	35
图 3-17 在不同位置获取图像 .....	36
图 3-18 八叉树示意图 .....	37
图 4-1 图像金字塔示意图 .....	40
图 4-2 伪随机码单块方格 .....	41
图 4-3 双线性插值 .....	42

图 4-4 采样点示意图 .....	43
图 5-1 墙面的三维重建结果 .....	45
图 5-2 不同距离下对墙面三维重建的深度值 .....	46
图 5-3 不同距离下对墙面三维重建的深度值的方差 .....	46
图 5-4 步距球规 .....	47
图 5-5 左相机获取步距球规图像 .....	47
图 5-6 步距球规的点云 .....	48
图 5-7 步距球规的三维重建结数据 .....	48
图 5-8 长城浮雕 .....	49
图 5-9 长城浮雕的三维重建结果 .....	49
图 5-10 死侍模型 .....	50
图 5-11 死侍模型的三维重建结果 .....	52

## 表格清单

表 2-1 MV-VDM130SC/SM 相机参数.....	9
表 2-2 移位寄存器的状态 .....	11
表 2-3 个人电脑配置 .....	16
表 3-1 左相机的两次标定数据 .....	25
表 3-2 左相机的两次标定结果的对比 .....	26
表 3-3 双目相机的内参 .....	26
表 3-4 图 3-9 中的符号释义 .....	29
表 5-1 步距球规参数 .....	47
表 5-2 长城浮雕参数 .....	49
表 5-3 死侍模型参数 .....	50

# 第一章 绪论

## 1.1 课题的研究背景及意义

随着社会的发展与科技的进步，越来越多的行业需要使用三维模型。有些三维模型可以通过 Solidworks 这样的软件设计并生成，然而这一方法在逆向工程中并不适用。例如在工业检测领域中，通过三维制作软件设计出了工件的模型，工件生产出来后需要对其进行分析，需要将该工件真实的三维模型与之前设计的三维模型进行对比，从而获得工件的偏差数据。利用三维重建技术，可以获得该工件的真实三维模型。在社会生活领域中，越来越多的人在电商平台上购买衣服，在挑选服饰的过程中只能看到该服装在模特身上的穿着效果。利用三维重建技术，获得身体的三维模型，将三维模型数据上传到电商平台后进行虚拟试衣，可以看到自己的穿着效果。还有很多行业和领域已经使用了三维重建技术，例如增强现实、虚拟现实、3D 打印、虚拟装配、文物复原等。

三维重建拥有多种实现方案。如果按照是否接触被测物体分类，可以分为接触式三维重建和非接触式三维重建。接触式三维重建方案有三坐标测量仪，非接触式三维重建方案有基于视觉的三维重建、基于激光雷达的三维重建、基于计算机断层扫描的三维重建等。基于视觉的三维重建技术越来越受到学术界和产业界的青睐，其具有结构简单、获取的信息丰富、硬件成本低廉等诸多优点。基于视觉的三维重建按照是否投射光源进行分类，可以分为主动式三维重建和被动式三维重建，主动式三维重建具有精度高、稳定性好、实时性好等诸多优点；当然其也有缺点，该方案的测量距离较短，不能对远距离的物体进行三维重建。基于视觉的主动式三维重建技术近年来之所以得到较大的发展，与其相关的硬件、算法的发展密不可分。首先是 CPU 计算能力大幅提升，因为基于视觉的三维重建，CPU 会实时接收并处理相机发送来的图像，CPU 拥有较强的计算能力才能满足实时三维重建的需求。其次是镜头的技术的成熟，再优秀的计算平台与算法，如果脱离了镜头，那也不能落到实处在各种场景下进行应用。最后是与三维重建相关算法的不断进步，三维重建中涉及到的相关算法有相机的标定、图像的认识与匹配、非线性优化算法等。

基于视觉的三维重建在研究与应用的过程中仍然有不少难题需要解决。随着各方面的技术不断进步，推动着三维重建技术向前发展。在发展的过程中，出现了多种基于视觉三维重建的解决方案，这些方案各有其优缺点与其适用的场景。从相机镜头的角度来说，有基于单目相机的三维重建<sup>[1]</sup>、基于双目相机

的三维重建<sup>[2]</sup>、基于深度相机的三维重建<sup>[3]</sup>。基于单目相机的三维重建虽然能够用最少的传感器实现了三维重建，但是单目相机本身尺度缺失，导致三维重建的结果没有尺度信息，无法获取被测物体的尺寸信息，所以不能用于测量领域。基于双目相机的三维重建，能够恢复出尺度信息，实现较为可靠的三维重建结果。但其受到被测物体的表面的灰度值、纹理等信息的影响较大，对于物体表面的纹理较弱甚至没有纹理的物体就无法进行三维重建。基于深度相机的三维重建，深度相机可以直接获取深度，但是其受到环境光的影响比较大，只能在室内使用，应用场景有一定的局限性。在上述的三种基于视觉的三维重建的解决方案中，都具有一定的局限性。只使用相机作为基于视觉三维重建的解决方案在很多情况下都受到限制，这里选择双目相机并加入伪随机编码结构光作为基于视觉的三维重建解决方案。向物体主动投射结构光，该结构光是提前通过编码生成并通过投影机投影到被测物体上，双目相机接受物体反射回来的光。该解决方案在精度和稳定性都有较高的优势，但其也有一些缺点，例如不可避免会发生遮挡问题，计算量较大等。

本文的重点是设计并搭建一套便携式的三维重建系统，该系统能够对近处的物体进行三维重建，在精度、实时性、稳定性都能满足一定的要求。考虑到本课题是对近处的物体进行三维重建，获得的三维模型可以用于 3D 打印，所以对精度有一定的要求，但又不是应用于工业检测领域，所以精度需要达到 0.2mm 左右即可。由于该三维重建系统设计成手持式，在使用过程中需要有一定的流畅度，所以该三维重建系统需要能够每秒钟处理 15 张左右的图像并将其重建出来，才能保证一定的实时性。被测物体与相机的距离大约在 350mm 左右。考虑到主要是对形状各异、大小各异的近处物体三维重建，所以该系统必须具有一定的灵活性，操作起来比较方便，因此设计成便携式。该三维重建系统的硬件系统由投影机、相机和补光灯等组成，通过投影机向被测物体主动投射伪随机编码结构光，双目相机采集图像，对物体进行三维重建。

## 1.2 国内外研究现状

基于伪随机编码结构光的立体视觉的三维重建的研究进展包括基于结构光的三维重建和基于立体视觉的三维重建算法两个方面，本文从这两个方面介绍近年来的研究进展。

### 1.2.1 基于结构光的三维重建研究现状

基于结构光<sup>[4]</sup>的三维重建属于主动式三维重建。被动式三维重建有着结构简单，实现起来比较容易。但被动式三维重建有相应的缺点，例如精度不高、容易受到场景的影响等。主动式三维重建向物体投射的光源有不同的形式。其分为时间调制和空间调制两类。时间调制通过相机接收到光的时间转换成距

离，从而可以获得深度信息，其中较为常见的是飞行时间法。空间调制通过相机接收到不同或者相同的光的形式，最终转换成深度信息，其中较为常见的有干涉法和结构光法。干涉法的精度高，但是其受环境的影响较大，而且只适用于测量微观表面形貌。结构光法精度高、抗干扰能力强、实时性好、成本低等优点，其也有相应的缺点，例如不可避免的会发生遮挡问题，这会导致图像匹配时无法找到正确的匹配点。

基于结构光法<sup>[5, 6]</sup>的三维重建在二十世纪七十年代被提出，经过不断的发展，这个方向的内容已经十分丰富。按照投射光源的面积大小可以分为点结构光<sup>[7]</sup>、线扫描结构光<sup>[8]</sup>、编码结构光<sup>[9]</sup>。点结构光法就是向被测物体投射一个点，相机获取图像后进行三维重建。这个缺点显而易见，速度极慢，因为每次只能重建出一个点。线扫描结构光的效率就提高了很多，之后又出现了多线结构光<sup>[10]</sup>。近年来许多优秀的公司已经将多线结构光落地应用，例如加拿大的 CREAFORM 公司的 HandySCAN 3D 扫描仪，中国的杭州思看科技有限公司的 KSCAN 扫描仪等，都是利用多线激光投射到被测物体，相机采集图像数据后进行三维重建。线结构光的效率与之前的点结构光有很多提升，依然需要拍摄多张图像才能进行三维重建，所以效率依然较低。

编码结构光法是向物体投射一幅图像或者多幅图像，相机获取从被测物体反射回来的光线从而记录下图像，然后进行三维重建。编码的方式分为时序编码与空间编码。时序编码中有二值码<sup>[11, 12]</sup>、n 值编码<sup>[13]</sup>、时间编码与相移法结合编码<sup>[14]</sup>和混合编码<sup>[15]</sup>。空间编码中有非正式编码<sup>[15]</sup>、基于 De Bruijn 序列编码<sup>[16]</sup>和 M-阵列编码<sup>[17, 18]</sup>。

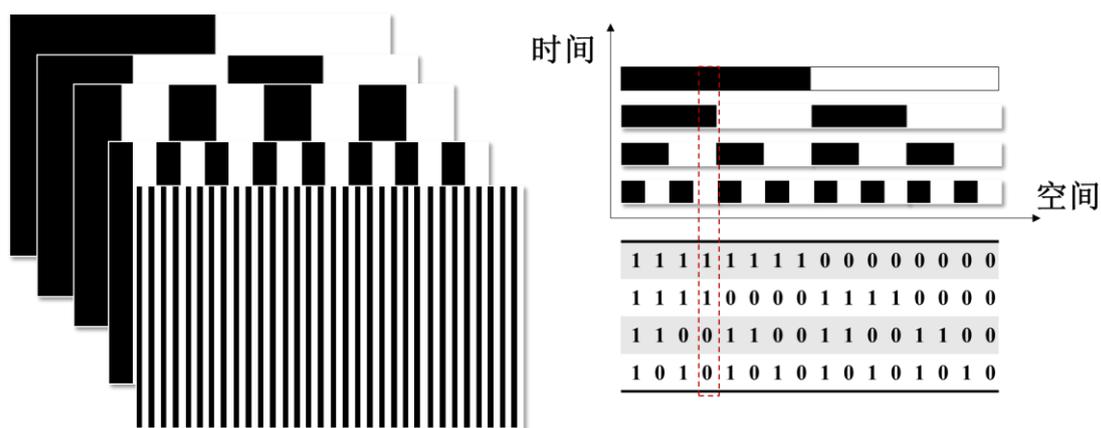


图 1-1 普通二值码

Figure 1-1. Normal binary code

时序编码是在一段时间内投射多张图案，每张图案投射到被测物体上相机都需要记录，这样在时间序列上，每个像素点的就有了编码。如图 1-1 中的左图为普通的二值码，只有黑白两种颜色，对应着二进制中的 1 和 0。图 1-1 右图中红色虚线框中所代表的编码为 0011，如果用二进制转换成十进制的方法，十进

制数为 3，也就是说在待检索的图像中搜索编码值为 3 的像素点即可。基于时序编码有很多改进，向抗干扰能力更强、唯一性更高、可识别度更高、匹配效果更好等方向发展。显然，使用该编码结构光需要投射不同的图案，因此相机需要拍摄大量的图片，并且只能对静态场景进行三维重建。

空间编码只向被测物体投射一张带有编码信息的图案，这张图案具有一定的编码信息，相机获取图像后进行三维重建。如图 1-2(a)所示，为 Spoelder<sup>[19]</sup> 等设计的  $65 \times 63$  大小的 M-阵列编码的一部分。其中的白格子表示 M 阵列码字，黑格子用来分割码字。如图 1-2(b)所示，为 Griffin<sup>[20]</sup> 等设计的编码  $18 \times 66$  大小的 M-阵列编码的一部分，使用 4 个符号生成编码，这使得编码具有更好的稳定性。如图 1-2(c)所示，为 Albitar<sup>[21]</sup> 等设计的 M-阵列编码，使用了 3 个元素，横线、圆圈、圆，该编码方式使得元素与元素之间的海明距离较大，使得其可靠性较高。

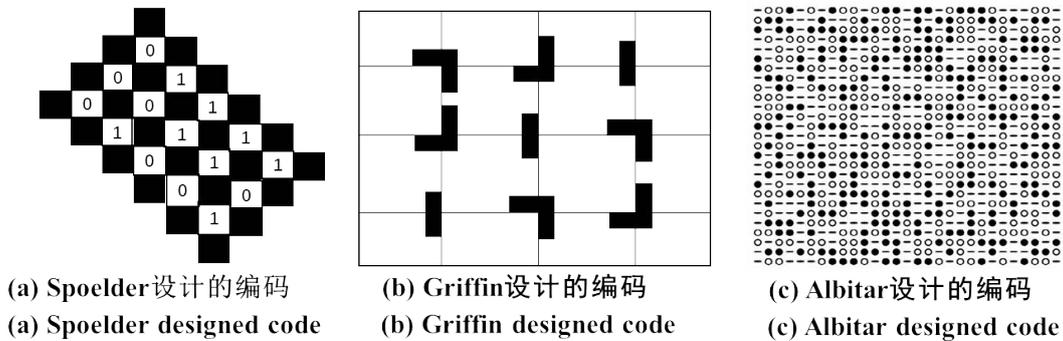


图 1-2 伪随机编码的设计方案

Figure 1-2. Pseudo-random coding design

国内在这方面也有很多科研人员在进行这方面的工作。张广军等<sup>[22]</sup>建立了结构光三维双视觉 RBF (Radial basis function) 神经网络模型，该模型不需要考虑外在因素的影响，使得该方法具有较高的精度。王宝光等<sup>[23]</sup>研究出了可以用来在生物医学、工业测量等领域的线结构光轮廓传感器。罗先波等<sup>[24]</sup>将线结构光传感器用于对运动的物体及腐蚀性的物体进行三维重建。樊强等<sup>[25]</sup>采用点结构光扫描重建复杂型面的刀具，测量精度达到  $\pm 10\mu\text{m}$ 。

### 1.2.2 立体视觉三维重建算法研究现状

立体视觉<sup>[26]</sup>在非接触式测量、自动驾驶、机器人视觉、增强现实、虚拟现实等领域有着广泛的应用。基于立体视觉的三维重建算法通常可以分为六个部分，分别为图像获取、相机标定、特征提取、立体匹配、深度感知、点云拼接，下面从这六个部分介绍近年来的研究现状。

图像获取是基于视觉三维重建的基础，图像获取的方式依据适用场景确定。例如，在自动驾驶<sup>[27]</sup>领域，相机装在车身上以较高的频率采集车身前后左右的情况；在工业检测领域，采用工业相机获取工件精准的图像数据；不管场

景如何变化，始终不变的要求只有一个：采集该场景下最好的图像数据，这时候需要考虑到该场景下的运动情况、光照、摄像机性能、视角与视野等因素。

相机标定<sup>[28]</sup>是基于视觉三维重建精度的保证。相机标定是获取相机的内外参数，建立三维世界中实际物体与其在图像平面上对应点之间的相对位置关系。相机标定的方法分为三类，传统相机标定方法<sup>[29]</sup>、相机自标定方法、基于主动视觉的相机的标定方法，传统标定方法中有直接线性变换法<sup>[30]</sup>、R.Tsai 的基于径向约束的两步法<sup>[31]</sup>、张正友平面标定法<sup>[32]</sup>。

特征提取<sup>[33]</sup>是立体匹配的前提。提取特征时按照某种约定将图像中的一些像素点作为兴趣点将其提取出来。提取的特征效果会影响到匹配的效果，如果该特征的可识别率不高，很容易匹配不上。所以对于特征点有 4 个要求，可重复性，相同的图像子区域可以在不同的图像中被找到；可区分性，不同的图像子区域不一样；高效率，提取特征点的速度必须要很快，并且在进行特征点匹配的时候也要很快；本地性，特征点仅与一小片图像子区域相关。常用的有角点特征、区域特征、边缘特征等。区域特征含有的信息丰富，适用于快速匹配，得到的是插图是稠密的，适用于三维重建与机器人导航，但是其特征提取和描述比较困难。角点特征中含有的信息较少，需要加入强有力的描述信息才能用于匹配，最终得到的视差图是稀疏的。例如 ORB<sup>[34]</sup>、SIFT<sup>[35]</sup>、SURF<sup>[36]</sup>特征就是一系列优秀的角点特征。

立体匹配<sup>[37, 38]</sup>的效果决定了三维重建的效果。立体视觉中的立体匹配是将同一场景下左右相机获取的两幅图像中的对应点一一对应起来。在灰度变化比较明显的场景下，找到同名对应点还是比较容易的，但是当场景中灰度变化不明显时，立体匹配就比较困难，需要从特定的场景出发加入特定的约束信息，增强匹配的效果。

深度感知<sup>[39]</sup>是将获取到的二维图像转换成三维信息，是立体视觉的重要目标。使用一个已经标定过的双目立体相机获取图像，立体匹配之后，使用三角测量或者透视变换原理，深度信息就可以得到了。深度感知之前的每一步都会对其精度造成影响。

点云拼接<sup>[40, 41]</sup>是三维重建的最后一步，将整个物体的点云进行拼。由于被测物体不能在一个视角下完全被相机观测到，需要在多个角度获取被测物体的形貌，将多个角度获取的图像生成点云后进行拼接。

### 1.3 本文内容安排

本文是基于伪随机码结构光的对物体进行三维重建，重点对三维重建系统的设计与制作、伪随机编码的设计、双目相机的标定、立体视觉的匹配、基于

三维重建系统的实时性与精度提升等内容进行研究。全文共分为六章，具体内容如下：

第一章：绪论。介绍了课题的研究背景与意义、国内外的研究现状。

第二章：三维重建实验系统。从相机的选型、伪随机编码的设计、软件构架的设计方面介绍本文搭建的三维重建系统。

第三章：基于伪随机码结构光的三维重建。从相机的标定、立体匹配、三维重建、点云拼接方面对三维重建的各部分算法进行实现。

第四章：基于实时性与精度的研究。在第三章的基础上，从三维重建的实时性与精度角度出发对其进行改进。

第五章：整体实验与过程分析。对本文设计的三维重建系统进行精度验证与实验，并对实验的数据进行分析。

第六章：结论与展望。总结全文的主要工作与展望未来发展的方向。

## 第二章 三维重建实验系统

本文基于伪随机编码结构光对近处物体进行三维重建，不仅需要对三维重建的算法进行研究，而且还需要对三维重建硬件和软件系统进行设计与实现。其中包括硬件系统的搭建、算法的实现和软件的开发。本章节阐述三维重建系统的设计，包括相机的选型、伪随机编码的设计、光源的选择、硬件结构和软件的简介。

### 2.1 相机的选型

相机是获取图像的设备，是三维重建的硬件系统的重要组成部分。相机成像的质量对三维重建的效果影响非常大，所以相机型号的选择很重要。首先需要了解相机的硬件结构，然后了解哪些相机参数对三维重建系统有影响，最后才可以确定本文采用哪种相机。我们这里所说的相机指的是工业相机，因为工业相机的图像质量稳定可靠，可用于图像处理，而民用相机不具备该特点。

#### 2.1.1 相机的硬件结构

相机的核心部件主要由镜头（lens）、图像传感器（image sensor）、数字信号处理器（digital signal processing）和滤光片（optical filter）等组成。成像质量主要受镜头、图像传感器和数字信号处理器这三个方面影响。

镜头主要是充当小孔成像模型中的小孔，不过镜头更能够聚集光线，接受更多的光线，在很短的时间内获得合适的曝光并有更为清晰的成像效果。一般来说镜头的构造越复杂，成像质量越高，价格就越贵。工业相机中通常组合塑胶透镜（plastic, P）与玻璃透镜（glass, G）构成镜头。玻璃透镜越多，成像效果越好，价格越高。常见的有 1G1P、1G2P、2G2P、4G 和 5G 等。

图像传感器是将通过镜头接收到的光线转换成电子信息的设备。目前常用的图像传感器有两种，分为感光耦合元件（charge-coupled device, CCD）和互补式金属氧化物半导体有源像素传感器（CMOS Active pixel sensor）。CCD 成像效果更好，但是成本高。CMOS 使用更为普遍，性价比高。

数字信号处理器是将信号以数字方式表示并处理的设备，是相机中的计算中心。其对图像传感器获取到的数据进行处理后上传到个人电脑中。

#### 2.1.2 相机参数

相机有很多参数，我们这里只考虑和三维重建紧密相关的一些参数。主要参数有快门类型、像素、帧率、景深、视场、成像颜色、焦距和相机传输接口。

快门是相机用来控制图像传感器有效曝光时间的机构。快门的类型目前分为卷帘快门和全局快门。卷帘快门对像素逐行扫描，而全局快门所有的像素同时曝光。所以当卷帘快门用于拍摄高速移动的物体的时候，会出现“果冻”现象，而全局快门不会出现这种情况。但是卷帘快门的价格便宜，体积小。本文是基于伪随机编码结构光对近处物体进行三维重建，会发生相对运动，如果选择卷帘快门会出现拖影的情况，所以我们这里选择全局快门。目前在计算机视觉检测领域，全局快门的应用较多。

相机的像素选择与图像大小、三维重建的精度密不可分。假设图像的宽和高分别为 $w$ 、 $h$ ，三维重建的精度为 $a$ ，像素为 $p$ ，那么像素的大小为：

$$p = \frac{w}{a} \times \frac{h}{a} = \frac{w \times h}{a^2} \quad (2-1)$$

像素越高，代表着图像质量越高，同时带来了较大的数据传输压力和较高的算力要求。三维重建的图像尺寸不能太小，否则对于较大的物体进行三维重建的时候会很慢，所以希望图像尺寸要尽可能大，这里选择 150\*150 (mm\*mm)，精度的范围是 0.1mm 到 1mm，那么图像的像素值的范围是 22.5 万像素到 225 万像素，所以我们选择 130 万像素的相机。

相机的帧率代表相机每秒钟拍摄照片的张数，帧率越高，代表相机每秒钟能够拍摄的照片数量越多。相机的帧率一定要大于你检测的速度。对于高速运动的场景来说，帧率比较高，这同时也带来了较大的数据传输压力和较高的算力要求。由于我们手持相机对近处的物体三维重建，手臂的移动速度在 300mm/s 左右，帧率设置为 15 帧时使用起来较为流畅，并能实现实时处理数据。在实际选择相机的过程中，帧率大于等于 15 帧就可以。

景深指的是焦点前后的清晰范围。本文主要是对近处的物体进行三维重建，最佳的距离设置在 350mm 左右，景深范围设置在 300mm 到 600mm 之内，则景深为 300mm。

视场 (field of view, FOV) 指的是相机的视野范围，视场又叫做视场角。视场角和焦距呈负相关，焦距越短视场角越宽，具体关系如下：

$$FOV = 2 \times \tan^{-1} \left( \frac{h}{2f} \right) \quad (2-2)$$

其中  $h$  是图像传感器的水平尺寸， $f$  为焦距。但是这两个条件并不知道，无法计算出视场角。有另外一种方法，通过相机与物体的距离和水平视场的大小可以得到视场角。这里我们用 600mm 作为工作距离，水平视场为 345mm，通过式(2-3)可以计算得到  $FOV = 60^\circ$

$$FOV = 2 \times \tan^{-1} \left( \frac{\text{水平视场}(mm)}{2 \times WD(mm)} \right) \quad (2-3)$$

我们通过相机最终获得的是一帧一帧的图片，如果处理的与图像颜色无关，那选择黑白相机，因为同样分辨率的彩色和黑白相机，黑白相机的精度比彩色相机高，尤其在看图像边缘的时候，黑白相机的效果更为突出。

焦距的选择，其他参数相同的情况下，焦距越短，物体的像就越大，所以，短焦距的镜头适合测量近处的物体，焦距较长的镜头适合拍摄距离较远的物体。我们这里是对一些物体进行三维重建，并非对一幢建筑物这样大型的物体进行三维重建，所以选择焦距较短的镜头就可以了。

传输接口，这个根据传输的距离、稳定性以及数据量确定，我们这里选择USB3.0类型的接口。

### 2.1.3 相机型号

根据本文对相机的一些需求，综合考虑后选择维视图像的 MV-VDM 小型 USB3.0 接口高速工业数字相机，相机如图 2-1 所示，该相机的主要参数如表 2-1 所示。



图 2-1 MV-VDM130SC/SM 相机

Figure 2-1. MV-VDM130SC/SM camera

表 2-1 MV-VDM130SC/SM 相机参数

Table 2-1. parameters of MV-VDM130SC/SM camera

参数名	数据
分辨率	1280 × 960
传感器类型	CMOS
快门类型	全局快门
电子快门	1/10000-1/4s
帧率	60fps
输出颜色	黑白
输出方式	USB3.0

## 2.2 伪随机编码结构光的设计

因为点、线结构光的效率问题，所以选择了面结构光。在第一章中的国内外研究现状也介绍了很多优秀科研人员设计的面结构光，但是很多的面结构光并不适合本文的要求。本文希望结构光在编码实现上能够简单，并且在三维重建的过程中需要保证一定的稳定性和精度，所以选择了伪随机编码结构光。而伪随机编码结构光的生成也有很多细节问题，接下来就具体设计并实现伪随机编码图案。

### 2.2.1 伪随机序列的生成

伪随机序列<sup>[42]</sup>同时具有随机序列的一些统计特性，又便于重复产生和处理。其拥有随机序列的优点，又避免了随机序列的缺点。伪随机序列最早出现于 20 世纪 60 年代，应用于通信领域。

伪随机序列通常由反馈移位寄存器组成的数字电路产生，反馈移位寄存器有分为线性和非线性。结合本文的需求，选择了线性移位寄存器，由线性移位寄存器产生的序列通常称为  $m$  序列，全称为最大长度线性反馈移位寄存器。移位寄存器由本原多项式  $h(x) = x^m + k_{m-1}x^{m-1} + \dots + k_2x^2 + k_1x + k_0$ ，生成的伪随机序列最大周期为  $T = q^m - 1$ ， $m$  为寄存器的个数， $q$  为存储器的状态数。本原多项式的系数是  $GF(q) = \{0, 1, A, A^2, \dots, A^{q-2}\}$  域中的元素， $A$  为本原。

在本文中， $m = 40, q = 2$ 。由于用  $m = 40$  生成伪随机序列的过程较为繁琐，下面用  $m = 4$  来说明伪随机序列的生成过程。则本原多项式为：

$$h(x) = x^4 + x + 1 \quad (2-4)$$

如图 2-2 所示，有 4 个移位寄存器，每个移位寄存器有 0 和 1 两种状态，最后输出的序列是不同时刻  $D_0$  的值，设为  $a_0a_1a_2a_3 \dots a_{q^m-2}$ 。

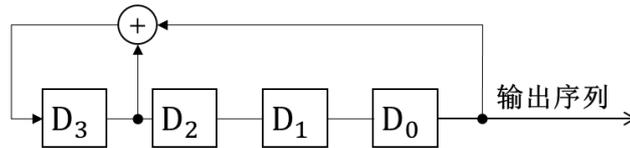


图 2-2 线性反馈移位寄存器

Figure 2-2. Linear feedback shift register

生成  $m$  序列需要移位寄存器的初始状态，假定初始状态为  $(D_3, D_2, D_1, D_0) = (1, 0, 0, 0)$ ，经过一次信号触发后，各个移位寄存器的状态为如表 2-2 所示。符号  $\oplus$  表示模 2 加运算，运算规则为  $1 \oplus 0 = 1, 1 \oplus 1 = 0 \oplus 0 = 0$ ，与二进制中的按位异或运算规则相同。如图 2-3 所示，经过不断的信号触发，得到了状态转移图，一共拥有 15 个状态，第 16 个状态和第 1 个状态相同，即周期为 15。显然，

初始状态不能全为 0，若某一状态为(0,0,0,)，下一状态为(0,0,0,0),并不是我们想要得到的结果。那图 2-2 输出的伪随机序列为：000111101011001。

表 2-2 移位寄存器的状态  
Table 2-2. state of shift register

移位寄存器	状态
$D_3'$	$D_3 \oplus D_0 = 1 \oplus 0 = 1$
$D_2'$	$D_3 = 1$
$D_1'$	$D_2 = 0$
$D_0'$	$D_1 = 0$

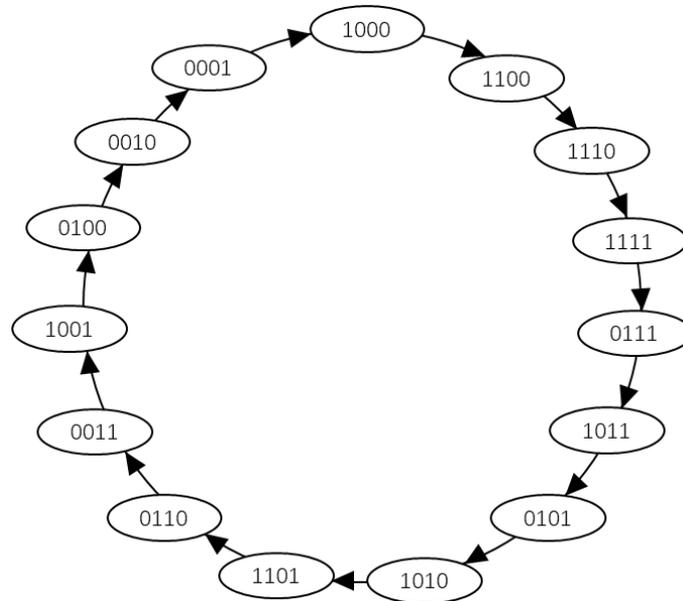


图 2-3 状态转移图

Figure 2-3. State transition diagram

### 2.2.2 伪随机阵列生成

将伪随机序列通过某种排列方式组成二维数组，该二维数组构成的阵列具有某种随机性，生成的二维数组叫做为随机阵列<sup>[43]</sup>。假设构成的阵列的维度为  $n_1 \times n_2$ ，其中维度为  $r \times v$  的子阵列在整个随机阵列中具有唯一性。

已知伪随机阵列的维度为  $n_1 \times n_2 = n$ , 则  $n_1$ 、 $n_2$  满足的关系为:

$$\begin{cases} n = 2^m - 1 = 2^{k_1 k_2} - 1 \\ n_1 = 2^{k_1} - 1 \\ n_2 = n_1 / n \end{cases}, n_1 \& n_2 > 1 \quad (2-5)$$

并且其中的  $n_1$ 、 $n_2$  互为素数。

当  $m=4$  时,  $n = 15 = 2^{2 \times 2} - 1$ , 则  $n_1 = 2^2 - 1 = 3$ ,  $n_2 = 5$ 。

本文选择的伪随机阵列的填充方式为从阵列的左上角开始, 沿着主对角线的方向填充, 当到达阵列的行边缘时, 就从下一列的第一行开始沿着平移注对角线方向继续填充; 当达到阵列的列边缘时, 就从下一行的第一列开始沿着平行主对角线方向继续填充。则当  $m = 4, q = 2$  时, 伪随机序列为  $a_0 a_1 a_2 a_3 \cdots a_{14}$ , 则伪随机阵列为:

$$\begin{bmatrix} a_0 & a_6 & a_{12} & a_3 & a_9 \\ a_{10} & a_1 & a_7 & a_{13} & a_4 \\ a_5 & a_{11} & a_2 & a_8 & a_{14} \end{bmatrix}$$

当寄存器的初始状态为 1000 时, 伪随机序列为: 000111101011001, 则伪随机阵列为:

$$\begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 \end{bmatrix}$$

由上面生成的阵列可以看出, 子窗口大为  $2 \times 2$  在伪随机阵列中具有唯一性。通用的伪随机阵列的表达方式为:

$$b = \begin{bmatrix} b_{0,0} & b_{0,1} & \cdots & b_{0,n_2-1} \\ b_{1,0} & b_{1,1} & \cdots & b_{1,n_2-1} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n_1-1,0} & b_{n_1-1,1} & \cdots & b_{n_1-1,n_2-1} \end{bmatrix} \quad (2-6)$$

伪随机阵列中的元素与伪随机序列的关系表示为:

$$\begin{cases} a_0 = b_{00} \\ a_1 = b_{11} \\ a_2 = b_{22} \\ \cdots \\ a_i = b_{ii} \end{cases} \quad (2-7)$$

$$\begin{cases} a_i = b_{i_1 i_2} \\ i \equiv i_1 \pmod{n_1}, 0 \leq i_1 \leq n_1 \\ i \equiv i_2 \pmod{n_2}, 0 \leq i_2 \leq n_2 \end{cases} \quad (2-8)$$

$$\begin{cases} n = q^m - 1 = q^{k_1 k_2} - 1 \\ n_1 = q^{k_1} - 1 \\ n_2 = n/n_1 \end{cases} \quad (2-9)$$

如果选择的  $m$  值很大, 使得整个伪随机码区域具有唯一性虽然可行, 但是会增加编码的难度。因为只要保证极线搜索的过程中, 每一条极线穿过的伪随机码区域中, 子区域具有唯一性即可。这条极线与下一条极线穿过的伪随机码如果相同, 对搜索匹配点没有影响。本文采用的用汇聚式样的投影模型, 相机

与相机之间的夹角为 60 度，极线的与像平面顶端的夹角不会太大，所以选择了  $m = 17, q = 2$ 。则本原多项式为：

$$h(x) = x^{13} + x^4 + x^3 + 1 \quad (2-10)$$

初始状态为：1000...000，可以生成最大周期为  $2^{13} - 1$  的伪随机序列。将伪随机序列编码成大小为  $320 \times 21$  伪随机阵列，编码结果如图 2-4 所示，其中任意  $5 \times 5$  大小的窗口在改编码图案中具有唯一性。



图 2-4 伪随机编码图案

Figure 2-4. Pseudo-random coded pattern

本文设计的三维重建系统，投影出的伪随机编码图案需要尽可能铺满两个相机的重合视场区域。最终生成的伪随机编码图案的大小为  $320 \times 147$ ，如图 2-5 所示，生成的伪随机编码图案。

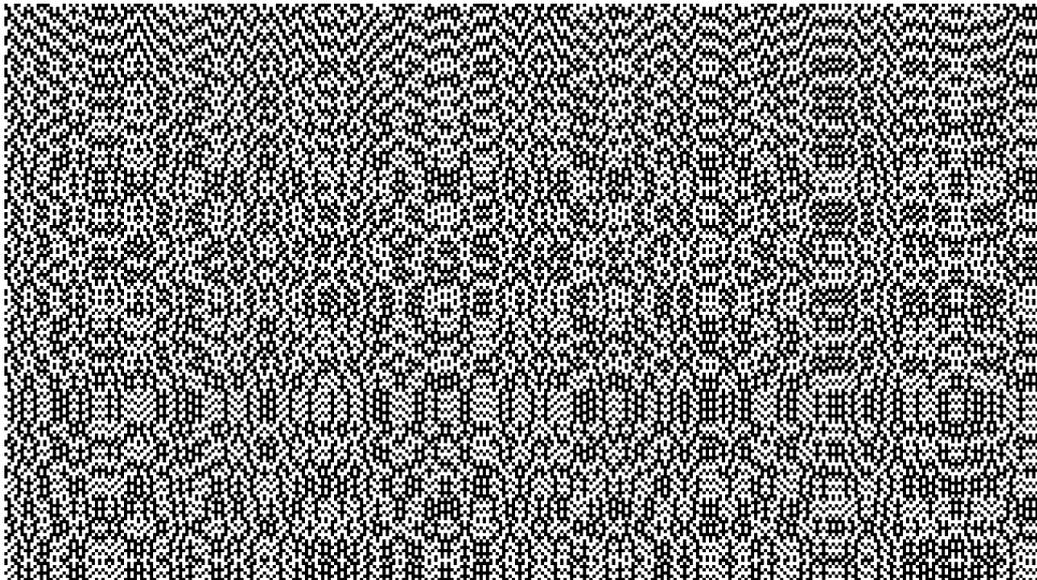


图 2-5 伪随机编码图案

Figure 2-5. Pseudo-random coded pattern

### 2.3 LED 补光

在本文搭建的三维重建系统中，采用 LED（发光二极管）进行补光提高三维重建的稳定性和鲁棒性。采用补光策略的优点有如下四点：1) 照亮被测物体，提高其亮度，与背景信息分离；2) 获得更好的成像效果，降低对算法的要求；3) 不受环境光的干扰，保证了图像质量；4) 提升系统的鲁棒性、精度等。

LED 具有寿命长、响应速度快、亮度可调节、形状自由度高、性价比高等优点，本文采用该光源进行补光。

LED 光源采用漫射光的照射方式。因为希望将被测物体照亮，若采取直射光的照射方式，只能照亮某一个点或者某几个点，并不符合本文的需求。采取漫射光的照射方式不仅能够将被测物体照亮，而且不会投射出阴影。

LED 光源的颜色采用白光。本文采用了伪随机编码结构光投影到被测物体，若采用其他颜色的光源会影响图像获取的质量，并且本文只需要将被测物体照亮，便于后期的图像处理。

采用明视场照明方式。明视场照明的方式是正面直射光照射的方式。

如图 2-6 所示，为 LED 与镜头的位置关系的示意图。中间的灰色部分代表了镜头，旁边的 4 个内嵌了黄色矩形的圆代表了 LED，这 4 个 LED 被固定在蓝色环状铁片上。

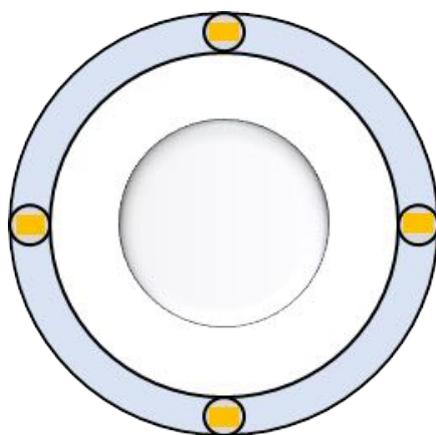


图 2-6 LED 与镜头的位置关系示意图

Figure 2-6. Schematic diagram of the position relationship between LED and lens

## 2.4 硬件结构介绍

基于伪随机编码结构光的三维重建系统的硬件结构主要由两个工业相机、投影机、两组 LED、个人电脑、型材铝、hub 和若干线缆组成。三维重建系统的硬件结构示意图如图 2-7 所示。

用两个工业相机和一个投影机使用螺丝固定在型材铝上，通过线缆连接到 hub 上，再连接到电脑上。这里的线缆指的是数据传输的线缆与供电的线缆。其中投影机用来投影出伪随机编码结构光，放置于左右相机的中间位置。示意图中的相机、型材铝等这些硬件是按照真实尺寸比例缩放绘制的。其中两个相机距离约为 200mm，相机与型材铝的夹角约为 30 度。实际的硬件结构图如图 2-8 所示。个人电脑的配置如表 2-3 所示。

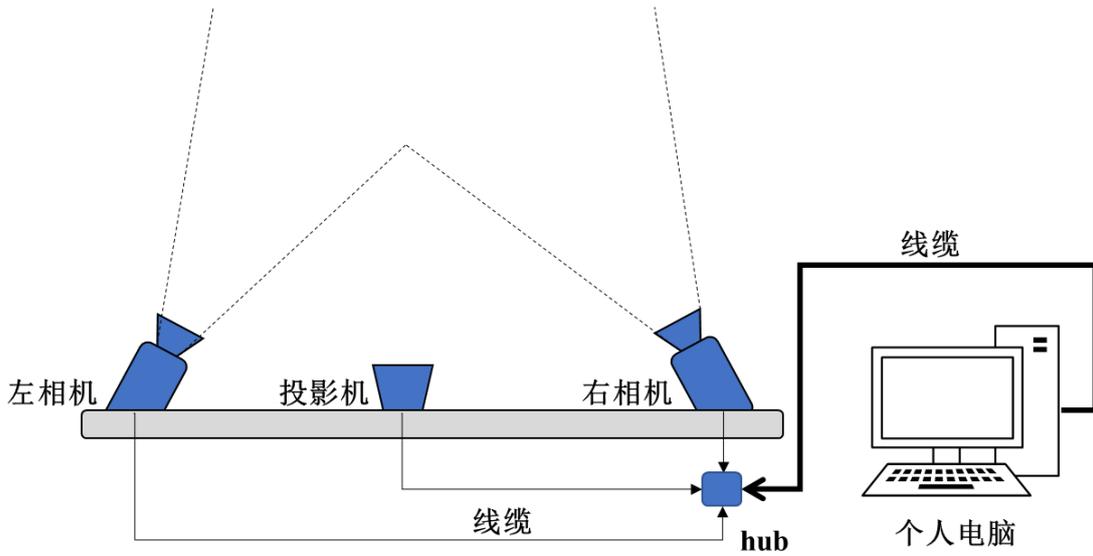


图 2-7 三维重建系统的硬件结构示意图

Figure 2-7. Hardware structure diagram of 3D reconstruction system



图 2-8 三维重建系统的硬件结构图

Figure 2-8. Hardware structure diagram of 3d reconstruction system

不同的计算平台的计算能力相差很大，例如嵌入式处理器、个人电脑 CPU、GPU 和服务端，它们都可以处理数据，但是它们的数据处理的能力相差巨大。我们这里选择的计算平台是个人电脑，首先因为个人电脑的计算能力适中，能够符合本文的计算要求；其次个人电脑比较常见，获取方便；最后使用个人电脑进行三维重建数据的显示、分析和处理也较为方便。

表 2-3 个人电脑配置  
Table 2-3. PC configuration

参数名		数值
处理器	型号	i7-8750H
	频率	2.21GHZ
显卡	型号	GTX1050
	容量	4G
内存	类型	DDR4 2400
	容量	32G
操作系统	型号	Windows 10

## 2.5 软件介绍

软件是基于 Microsoft Visual Studio 2017 版本中的 MFC 进行开发的。将一些功能与 MFC 中的 Button 控件进行关联，可以较为方便的进行重复操作；将一些输出信息与 MFC 中的 Edit 控件进行关联，可以输出一些信息，可以在使用软件时查看软件的运行状态。下面介绍一下软件界面中的一些功能按键，软件界面截图如图 2-9 所示。



图 2-9 软件界面

Figure 2-9. Software interface

软件界面中主要分为 6 块区域，分别为硬件测试、标定、相机参数设置、三维重建、信息提示和图像与点云显示。

硬件测试区域是对硬件部分进行调试，主要测试硬件和软件通讯是否正常、硬件能否正常工作。分别对硬件中的相机、投影机、LED 进行测试，对他们分别进行单次触发与连续触发。单次触发只触发一次，相机获取一次图像、投影机投影一次图案、LED 亮一次。连续触发后，相机连续获取图像、投影机一直投影图案、LED 一直补光。

标定区域是对相机进行标定。共有 4 个功能按键，分别为启动、抓取图像、关闭、计算标定参数。启动按键包含了连续触发相机与连续触发 LED。抓取图像按键代表在移动相机的过程中，当相机移动到合适的位置时，记录当前位置的图像。在标定之前需要选择标定类型，共有三种类型，分别为左相机内参、右相机内参、外参。

相机参数设置区域是对相机的参数进行设置。在三维重建的过程中，有时候需要根据被测物体的类型设置相机的一些参数。

三维重建区域是对物体进行三维重建时需要用到的一些功能按键。这个区域共有 5 个功能按键，分别为启动、结束、滤波、导出点云、抓取图片。启动按键包含启动相机、投影机、LED 三个硬件，启动后进行实时数据处理。结束后进行滤波，优化点云质量。获取图片按键是保存按下按钮这一刻左右相机获取到的图像。

信息提示和图像与点云显示区域是输出三维重建时的一些信息。在三维重建时需要看到图像的成像质量、点云的重建质量、每一帧点云的数量、相机的帧率、标记点识别的个数和一些其他提示信息。

## 2.6 本章小结

本章搭建了基于伪随机编码结构光的三维重建系统。该系统分为硬件部分和软件部分，主要是对相机的选型和伪随机编码两部分进行着重分析和设计。三维重建硬件系统的搭建为后期三维重建算法验证提供了平台。

### 第三章 基于伪随机编码结构光的三维重建

对物体进行三维重建需要经过相机标定、图像获取、提取特征、立体匹配、点云生成、点云拼接等多个步骤，每一个步骤都影响着三维重建的精度与稳定性。本章主要阐述了基于伪随机编码结构光的三维重建系统中的相机标定、立体匹配以及三维重建部分的实现。本文的三维重建算法更具有稳定性与鲁棒性，不受被测物体的颜色、形状以及环境光等因素的影响。

#### 3.1 算法流程

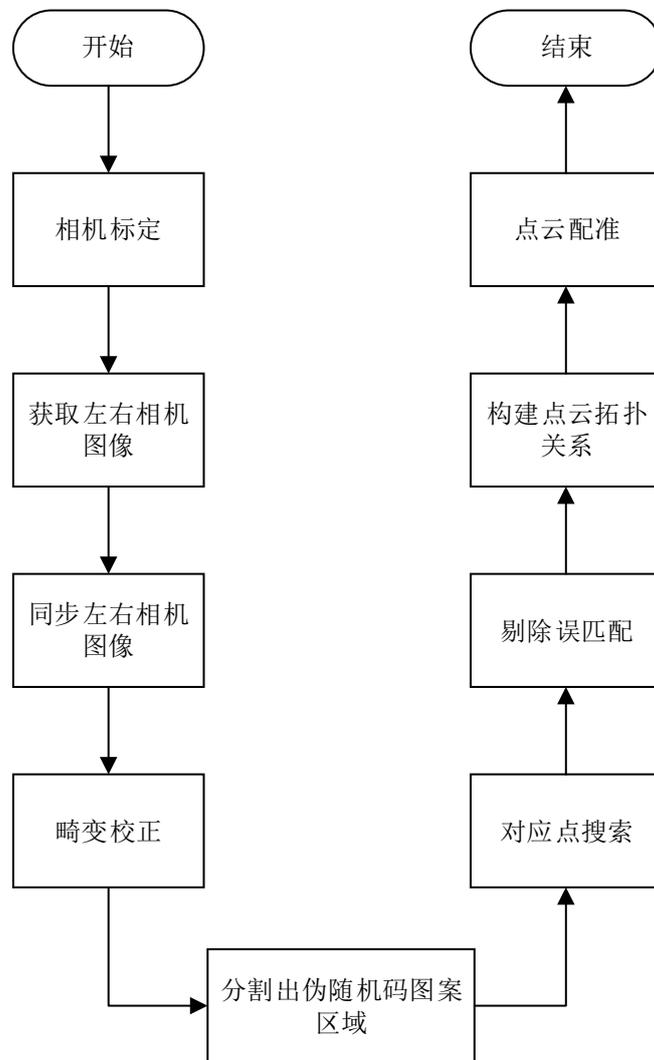


图 3-1 算法流程

Figure 3-1. Algorithm process

本文基于伪随机编码结构光对近处物体进行三维重建，设计了相应的算法流程。第一步是准备工作，即相机的标定。第二步正式进入到三维重建算法的过程中。先获取左右相机的图像，只对有伪随机码的区域进行三维重建，分割出伪随机码图案区域。然后，结合伪随机码与灰度相似在极线搜索寻找对应点，三维重建后剔除误匹配。为了方便点云的存储及后续的操作，建构点云的拓扑关系。最后将不同视角下生成的点云配准，完成整个物体的三维重建。三维重建算法流程如图 3-1 所示。

## 3.2 相机标定

### 3.2.1 相机标定参数

相机标定的参数有内参和外参，内参指的是焦距 ( $f$ )、缩放系数 ( $\alpha$ 、 $\beta$ )、平移系数 ( $c_x$ 、 $c_y$ )、畸变参数 ( $k$ 、 $p$ )；外参指相机相对于被测物体的相对位置关系 ( $\mathbf{R}$ 、 $\mathbf{t}$ )。 $\mathbf{R}$ 是 $3 \times 3$ 的旋转矩阵， $\mathbf{t}$ 是 $3 \times 1$ 的平移向量。焦距、缩放系数、平移系数构成透视矩阵 $\mathbf{K}$ ， $\mathbf{K}$ 矩阵的具体形式为：

$$\mathbf{K} = \begin{pmatrix} \alpha \cdot f & 0 & c_x \\ 0 & \beta \cdot f & c_y \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} \quad (3-1)$$

这里的畸变参数 $k$ 、 $p$ 总共选择了 8 个数，分别为 $k_1$ 、 $k_2$ 、 $k_3$ 、 $k_4$ 、 $p_1$ 、 $p_2$ 、 $p_3$ 、 $p_4$ 。

相机的内部参数分为两个部分，透视矩阵和畸变。内参描述三维世界中的物体转换到相机图像的变换关系。如果相机模型是标准的小孔成像模型，内参只需要透视矩阵。为了获得更好的成像效果，在相机成像平面的前方加入了透镜，光线经过镜头时发生折射，产生了径向畸变；由于镜头在装配的过程中，不能和相机成像平面完全平行，产生了切向畸变。相机的外部参数指的是相机坐标系和标定板的位姿关系。

相机的内部参数基本上不会发生变化的，只需要标定一次就可以使用。内参是三维重建的基础，内参的精度越高，三维重建的精度也越高。所以标定相机的重点就是求解相机内参，由于本实验系统采用的是双目相机，还需要标定双目相机之间的位置关系。

### 3.2.2 相机模型

相机模型按照是否忽略畸变参数，可以分为线性模型和非线性模型。线性模型可以将相机的模型建模成简单的针孔模型。如图 3-2 所示。

图中点 $p$ 是三维世界中的某一点，最终在物体成像平面得到点 $p'$ ，可以简化成一个相似三角形。假设点 $p$ 在相机坐标系下的坐标为 $[x, y, z]$ ，点 $p'$ 在物理成像平面坐标系下的坐标为 $[x', y', z']$ ，可以得到：

$$\begin{aligned} X' &= f \frac{X}{Z} \\ Y' &= f \frac{Y}{Z} \end{aligned} \quad (3-2)$$

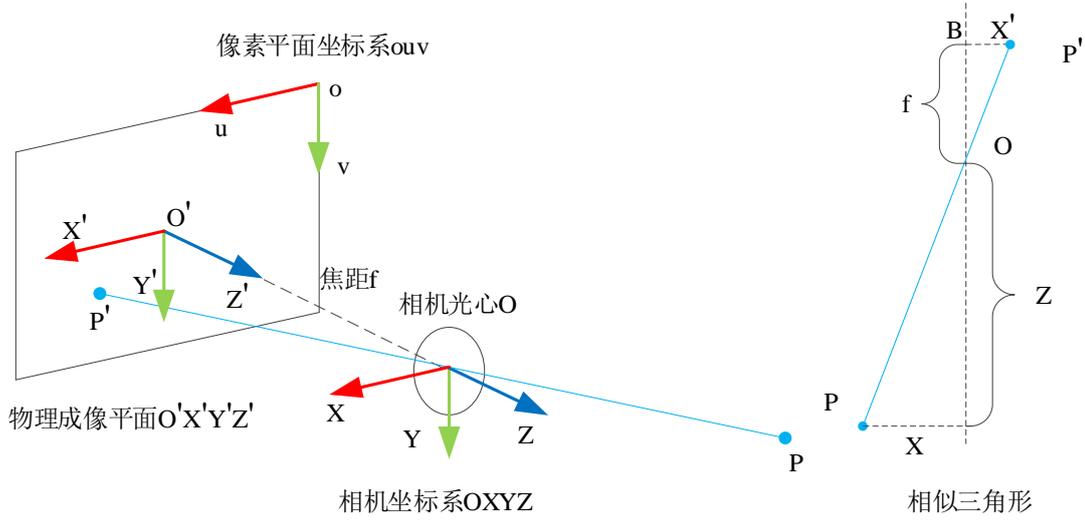


图 3-2 针孔相机模型  
Figure 3-2. Pin hole model

由于最后得到的像坐标是在像素平面坐标系下的坐标 $[u, v]^T$ ，所以假设像素坐标系和平面坐标系转换关系如下：像素坐标在 $u$ 轴上缩放 $\alpha$ ，平移 $c_x$ ；在 $v$ 轴上缩放 $\beta$ ，平移 $c_y$ 。所以，像素坐标 $[u, v]^T$ 和 $p'$ 的坐标关系如下：

$$\begin{cases} u = \alpha X' + c_x \\ v = \beta Y' + c_y \end{cases} \quad (3-3)$$

将式(3-2)代入到(3-3)得：

$$\begin{cases} u = \alpha f \frac{X}{Z} + c_x \\ v = \beta f \frac{Y}{Z} + c_y \end{cases} \quad (3-4)$$

令 $\alpha f = f_x$ ， $\beta f = f_y$ ，代入到(3-3)得：

$$\begin{cases} u = f_x \frac{X}{Z} + c_x \\ v = f_y \frac{Y}{Z} + c_y \end{cases} \quad (3-5)$$

将其写成矩阵的形式可得：

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \frac{1}{Z} \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \triangleq \frac{1}{Z} \mathbf{K} \mathbf{p} \quad (3-6)$$

所以这里的矩阵 $\mathbf{K}$ 就是标定需要求解的透视矩阵。在实际情况中，点 $\mathbf{p}$ 是三维世界中的真实点，已知的坐标值是在世界坐标系下得到的。世界坐标系下的坐标计作 $\mathbf{p}_w$ ，则有：

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \frac{1}{Z} \mathbf{K} \mathbf{p} = \frac{1}{Z} \mathbf{K} (\mathbf{R} \mathbf{p}_w + \mathbf{t}) = \frac{1}{Z} \mathbf{K} \mathbf{T} \mathbf{p}_w \quad (3-7)$$

将畸变参数考虑到相机的模型之中，可以得到：

$$\begin{pmatrix} u' \\ v' \\ 1 \end{pmatrix} = \frac{1}{Z} \mathbf{K} \mathbf{p} = \frac{1}{Z} \mathbf{K} (\mathbf{R} \mathbf{p}_w + \mathbf{t}) = \frac{1}{Z} \mathbf{K} \mathbf{T} \mathbf{p}_w \quad (3-8)$$

$$\begin{aligned} u_p &= u' \frac{(1 + k_1 r^2 + k_2 r^4 + k_3 r^6)}{(1 + k_4 r^2 + k_5 r^4 + k_6 r^6)} + 2p_1 u' v' + p_2 (r^2 + 2u'^2) \\ v_p &= v' \frac{(1 + k_1 r^2 + k_2 r^4 + k_3 r^6)}{(1 + k_4 r^2 + k_5 r^4 + k_6 r^6)} + p_1 (r^2 + 2v'^2) + 2p_2 u' v' \end{aligned} \quad (3-9)$$

其中， $(u_p, v_p)^T$ 是去畸变后的像素坐标。由于本次课题对三维重建的精度有一定的要求，所以畸变参数的选择了8个，所以本次标定的内参一共有12个参数，分别为： $f_x, f_y, c_x, c_y, k_1, k_2, p_1, p_2, k_3, k_4, k_5, k_6$ 。

### 3.2.3 标定板制作

相机标定的精度是三维重建精度的基础。标定的精度主要取决于两个方面，标定物和标定算法。是否需要标定物依赖于标定法的选择，我们这里选择传统的标定法。一般情况下，标定物的精度越高，制作该标定物的成本就越高。在二维标定物中，较为常见的是使用棋盘格标定板。本次毕业设计没有采用棋盘格标定板的原因有如下几点：第一，将棋盘格打印在普通的A4纸或者相片上，虽然操作方便、成本较低，但是标定物的精度无法保证；第二，采用简单打印的方式，相机在进行标定的过程中，棋盘格标定板会发生晃动，对标定的结果影响较大；第三，将棋盘格制作在光学玻璃上，操作麻烦，成本较高。为了同时兼顾精度与成本两个方面，本次标定物的制作方案是在漫反射玻璃面板上贴直径为6mm和8mm的圆形反光标记点。6mm标记点用来图像匹配时提取其中的圆心点，12mm标记点用来在匹配的过程中确定标定板的方向，以提高匹配的准确性、降低算法的复杂度。

如图3-3，该圆点标定板的尺寸是230mm × 350mm，标记点的个数是6\*7=42个，还有3个12mm的标记点。标定板的尺寸根据相机的视野选择，标定板的图像至少占相机视野的1/3。

使用圆点标定板，需要确定每个标记点的位姿（位置和姿态）。使用杭州思看科技有限公司的全局摄影测仪器（MSCAN），获取每个标记点圆心的坐标。

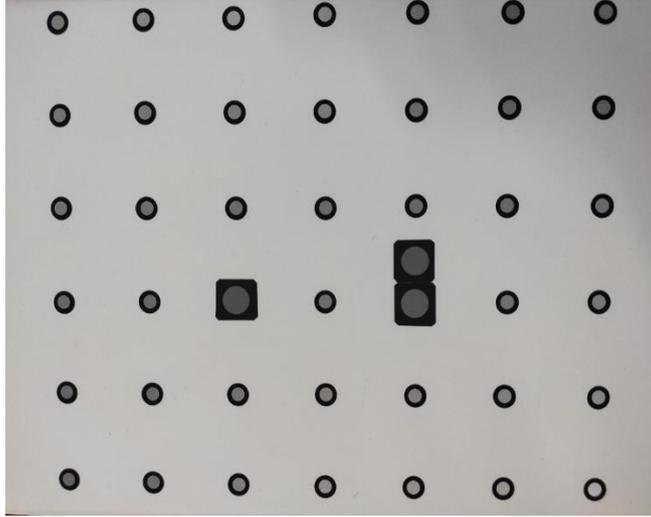


图 3-3 圆点标定板

Figure 3-3. Circle points calibration board



图 3-4 全局摄影测量仪

Figure 3-4. MSCAN

### 3.2.4 标定算法

假设相机没有畸变，先标定得到内参中的 $f_x, f_y, c_x, c_y$ ，以及外参的 $\mathbf{R}, \mathbf{t}$ ，再求解畸变参数。相机获取到标定板的图像，可以得到投影关系为：

$$\mathbf{p} = s\mathbf{H}\mathbf{P} \quad (3-10)$$

即：

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = s\mathbf{K}[\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{r}_3 \quad \mathbf{t}] \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} = s\mathbf{K}[\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{t}] \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \quad (3-11)$$

所以：

$$\mathbf{H} = s\mathbf{K}[\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{t}] \quad (3-12)$$

其中,  $p = [x, y]^T$  是像素坐标,  $P = [X, Y, Z]^T$  是世界坐标,  $s$  是尺度比例,  $H$  是  $3 \times 3$  的单应性矩阵,  $K = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$ ,  $r_1$ 、 $r_2$ 、 $r_3$  为旋转矩阵  $R$  分解出来的三个向量,  $t$  是平移向量。

将单应性矩阵写成列向量形式, 则有:

$$H = [h_1 \ h_2 \ h_3] = sK[r_1 \ r_2 \ t] \quad (3-13)$$

分解后, 得到:

$$\begin{cases} h_1 = sKr_1 & \text{or} & r_1 = \frac{1}{s}K^{-1}h_1 \\ h_2 = sKr_2 & \text{or} & r_2 = \frac{1}{s}K^{-1}h_2 \\ h_3 = sKt & \text{or} & t = \frac{1}{s}K^{-1}h_3 \end{cases} \quad (3-14)$$

因为  $r_1$ 、 $r_2$  相互正交, 所以可以得到:

$$\begin{cases} r_1^T r_2 = 0 \\ r_1^T r_1 = r_2^T r_2 \end{cases} \quad (3-15)$$

将式(3-14)代入式(3-15)可以得到:

$$\begin{cases} h_1^T K^{-T} K^{-1} h_2 = 0 \\ h_1^T K^{-T} K^{-1} h_1 = h_2^T K^{-T} K^{-1} h_2 \end{cases} \quad (3-16)$$

令  $B = K^{-T} K^{-1}$ , 展开后可以得到:

$$B = \begin{bmatrix} \frac{1}{f_x^2} & 0 & \frac{-c_x}{f_x^2} \\ 0 & \frac{1}{f_y^2} & \frac{-c_y}{f_y^2} \\ \frac{-c_x}{f_x^2} & \frac{-c_y}{f_y^2} & \frac{c_x^2}{f_x^2} + \frac{c_y^2}{f_y^2} + 1 \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{bmatrix} \quad (3-17)$$

将式(3-17)代入式(3-16), 通解为:  $h_i^T B h_j$ , 则有:

$$h_i^T B h_j = v_{ij}^T b = \begin{bmatrix} h_{i1} h_{j1} \\ h_{i1} h_{j2} + h_{i2} h_{j1} \\ h_{i2} h_{j2} \\ h_{i3} h_{j1} + h_{i1} h_{j3} \\ h_{i3} h_{j2} + h_{i2} h_{j3} \\ h_{i3} h_{j3} \end{bmatrix}^T \begin{bmatrix} B_{11} \\ B_{12} \\ B_{22} \\ B_{13} \\ B_{23} \\ B_{33} \end{bmatrix}^T \quad (3-18)$$

由  $v_{ij}^T$  的定义可得两个约束:

$$\begin{bmatrix} v_{12}^T \\ (v_{11} - v_{22})^T \end{bmatrix} b = 0 \quad (3-19)$$

获取  $n$  个不同位置的标定板图像, 那么矩阵  $V$  的维度是  $2n \times 6$ , 这里  $n$  取 20, 以保证标定板图像出现在相机视野的各个地方。则相机内参的代数解为:

$$\begin{aligned}
f_x &= \sqrt{1/sB_{11}} \\
f_y &= \sqrt{B_{11}/s(B_{11}B_{22} - B_{12}^2)} \\
c_x &= -sB_{13}f_x^2 \\
c_y &= (B_{12}B_{13} - B_{11}B_{23})/(B_{11}B_{22} - B_{12}^2)
\end{aligned} \tag{3-20}$$

其中：

$$s = B_{11}/[B_{33} \times B_{11} - B_{13}^2 - c_y(B_{12}B_{13} - B_{11}B_{23})] \tag{3-21}$$

相机外参为：

$$\begin{aligned}
\mathbf{r}_1 &= 1/(s\mathbf{K}^{-1}\mathbf{h}_1) \\
\mathbf{r}_2 &= 1/s\mathbf{K}^{-1}\mathbf{h}_2 \\
\mathbf{r}_3 &= \mathbf{r}_1 \times \mathbf{r}_2 \\
\mathbf{t} &= 1/s\mathbf{K}^{-1}\mathbf{h}_3
\end{aligned} \tag{3-22}$$

由式(3-5)可知，忽略相机畸变的时，世界坐标与像素坐标的变换关系。假设 $(u_p, v_p)$ 为点的位置， $(u, v)$ 为畸变位置，则有式(3-9)，畸变前后的像素坐标的变换关系，通过多个方程可以求解畸变参数。

### 3.2.5 标定实验

相机获取到图像后，关键的步骤是提取每个标记点的圆心。对图像先后进行滤波、增强、检测、定位，拟合圆、拟合圆心。图像处理的结果，如图 3-5 和图 3-6 所示。标记点的圆心数据得到之后，就可以构建式(3-19)组成的多个方程，进行方程的求解，即可得到标定的结果。

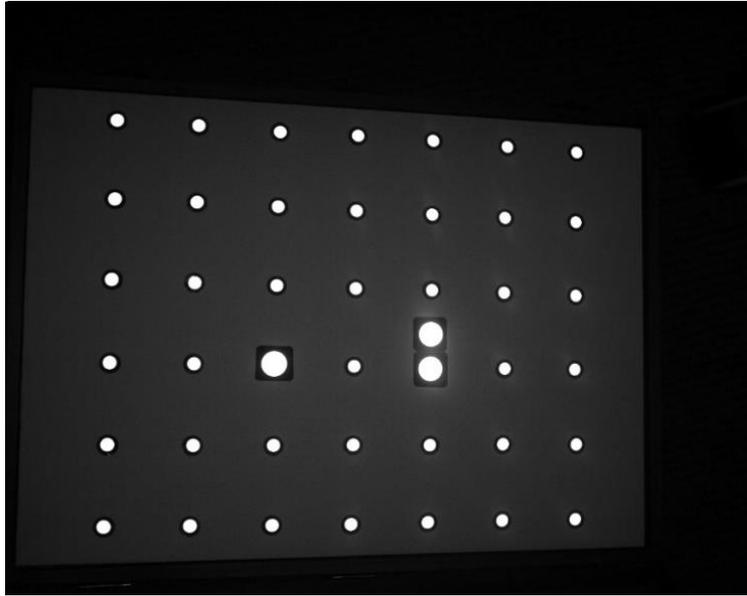


图 3-5 标定原图像

Figure 3-5. Raw image of calibration

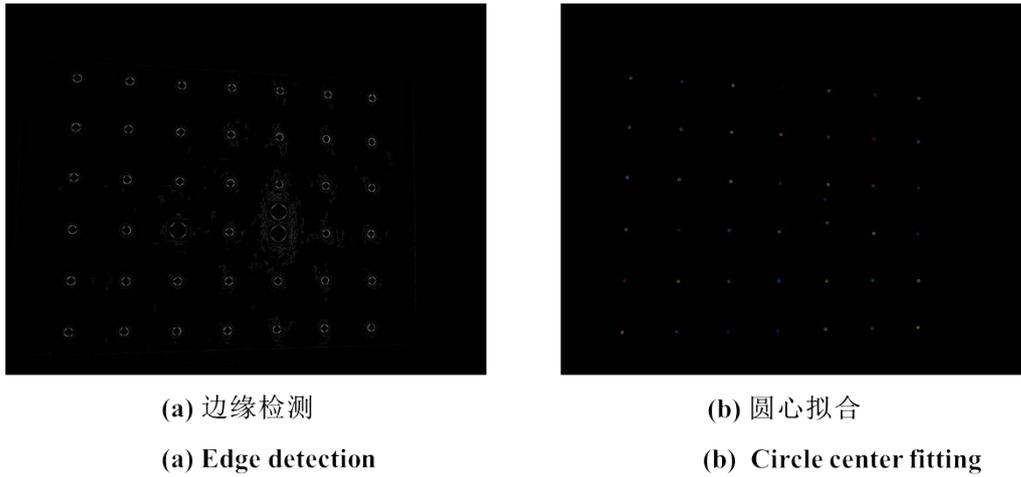


图 3-6 标定图像处理

Figure 3-6. Calibration image processing

### 3.2.6 标定结果评价

本文对相机进行了多次标定，发现每次标定的数据都有变化。每一组数据有 12 个数，无法直观的比较标定结果差异，于是采用可视化的方式显示出来。本文基于 Microsoft Visual Studio 2017 中的 MFC 开发了一个内参比较软件，采用色谱图呈现前后两次标定结果的差异。用鼠标点击显示色谱图的区域，可以显示当前这个像素点的坐标和误差值，软件界面如图 3-7 所示。

表 3-1 左相机的两次标定数据

Table 3-1. Two calibration results of the left camera

参数名	标定结果 1	标定结果 2
$f_x$	1713.66	1712.68
$f_y$	638.13	638.09
$c_x$	1713.63	1712.59
$c_y$	514.64	514.22
$k_1$	-0.014438	-0.017892
$k_2$	0.221141	0.252948
$k_3$	-0.003268	0.002389
$k_4$	-0.002013	-0.000207
$p_1$	0.377024	0.015804
$p_2$	0.019474	0.018005
$p_3$	0.081797	-0.061831
$p_4$	0.356338	-0.067625

表 3-2 左相机的两次标定结果的对比

Table 3-2. Comparison of two calibration results of the left camera

误差名称	误差值 (单位: 像素)
均值	0.058886
方差	0.001035
最大值	0.099981

如表 3-1 所示, 前后两次左相机的标定的数据, 每组 12 个数据。如表 3-2 所示, 前后两次左相机标定数据的对比, 其中最大的地方相差 0.099981 像素; 方差显示了整体标定结果的差异, 数据为 0.001035 像素, 表明整体的波动很小。如图 3-7 所示, 使用内参比较器软件, 直观地显示了前后两次左相机的标定数据差异。如图 3-7 所示, 以色谱图的方式显示各个像素点的标定误差比较, 这里的误差上限选择的是 0.1 像素。

色谱图是基于畸变公式(3-9)而来的, 最后转换成单个像素之间的比较。色谱图的颜色越红代表误差越大, 颜色越绿代表误差越小。这里的颜色深浅是相对值。四周的纯蓝色代表超出相机视野范围之外的图像, 这一块区域的标定误差不进行比较。

表 3-3 双目相机的内参

Table 3-3. Intrinsic of two cameras

参数名	左相机	右相机
$f_x$	1713.14	1712.45
$f_y$	639.84	686.32
$c_x$	1712.86	1711.31
$c_y$	517.88	481.31
$k_1$	-0.024631	-0.018253
$k_2$	0.044351	0.114411
$k_3$	-0.001273	0.001572
$k_4$	0.002398	-0.000189
$p_1$	-0.569129	0.392475
$p_2$	0.025461	0.018376
$p_3$	-0.159491	-0.002133
$p_4$	-0.409842	0.307753

通过多次标定相机的内外参数, 最终确定双目相机的内参, 以及左右相机的相对位置关系。

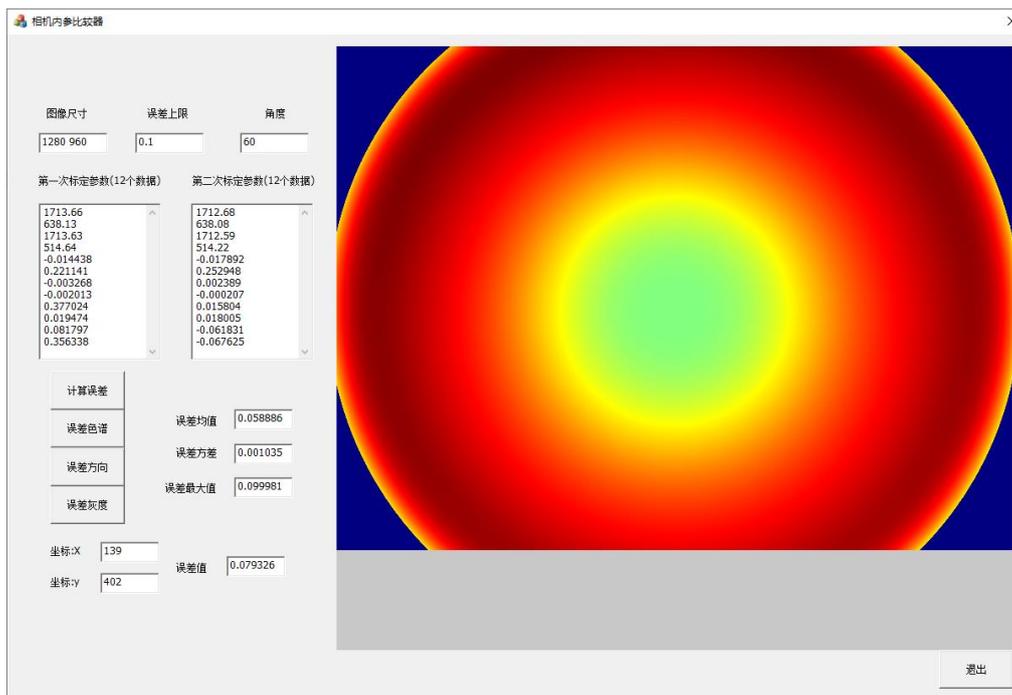


图 3-7 内参比较器

Figure 3-7. Intrinsic comparator

双目相机的内参如表 3-3 所示。双目相机的外参为：

$$\mathbf{R} = \begin{pmatrix} 0.888061 & -0.00125292 & 0.459724 \\ 0.00433001 & 0.999975 & -0.0056391 \\ -0.459705 & 0.00699847 & 0.888044 \end{pmatrix} \quad (3-23)$$

$$\mathbf{t} = (-204.265 \quad -0.252961 \quad 50.1539)^T \quad (3-24)$$

其中矩阵 $\mathbf{T}$ 指的是左相机到右相机的变换关系，即：

$$\mathbf{p}_r = \mathbf{R}\mathbf{p}_l + \mathbf{t} = \begin{pmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{pmatrix} \mathbf{p}'_l = \mathbf{T}\mathbf{p}'_l \quad (3-25)$$

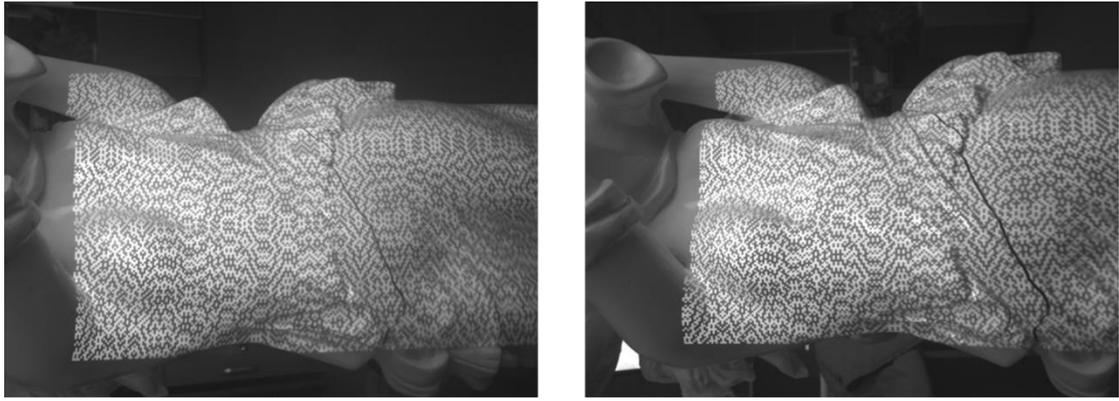
其中点的坐标形式为：

$$\mathbf{p}_r = \begin{pmatrix} x_r \\ y_r \\ z_r \end{pmatrix}, \mathbf{p}_l = \begin{pmatrix} x_l \\ y_l \\ z_l \end{pmatrix}, \mathbf{p}'_l = \begin{pmatrix} x_l \\ y_l \\ z_l \\ 1 \end{pmatrix} \quad (3-26)$$

双目相机的内外参数得到之后，计算相机的重投影误差以获取相机标定的精度。使用表 3-3 的内参和(3-23)与(3-24)的外参计算相机的重投影误差，重投影误差为：0.0283 像素。符合本文要求。

### 3.3 立体匹配

在三维重建的过程中，立体匹配<sup>[44]</sup>占有至关重要的地步，并且也是整个三维重建中难度较大的部分。立体匹配的目的是建立左右图像的像素点的一一对应关系。通过一定的匹配方法，实现这一目的。



(a) 左图像

(a) Left image

(b) 右图像

(b) Right image

图 3-8 相机图像

Figure 3-8. images of the cameras

如图 3-8 所示，是获取到的左右相机的图像。在立体视觉匹配之前需要确保获取到的左右相机图像是同一时刻的。对于人眼来说，立体匹配是一件非常容易的事情，因为人眼与人脑的高度配合，将大部分噪声去除，清晰的呈现这个物体的三维模型。但是这件事情对相机和计算机来说并非易事。因为光照不均匀、噪声干扰、相机视角差异、镜头畸变等因素，都会增加误匹配的概率。例如一个左、右相机都看到了一个圆，但是由于镜头畸变、相机视角差异，在左右镜头成像的时候变成了椭圆，并且两个椭圆也有差异。所以，准确无误的匹配较为困难。

由于本文侧重于近处物体的三维重建，希望左右相机的视野尽可能多的重合，所以采用了汇聚式视觉模型。如图 3-9 所示，为对极约束示意图。关于图中的参数说明如表 3-4 所示。

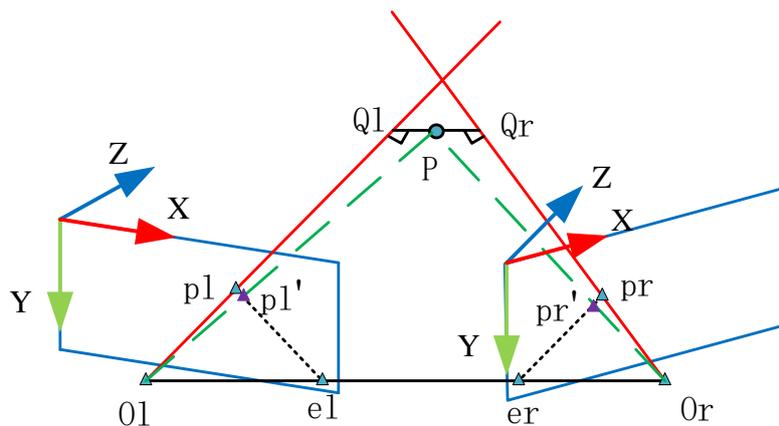


图 3-9 对极约束

Figure 3-9. Epipolar constraint

表 3-4 图 3-9 中的符号释义

Table 3-4. The symbolic interpretation in Figure 3-9

符号	释义
$O_l、O_r$	左右相机光心
$p_l、p_r$	点 $p$ 在左右相机的投影
$Q_lQ_r$	直线 $O_l p_l$ 与直线 $O_r p_r$ 公垂线
$P$	公垂线 $Q_lQ_r$ 的中点
$p_l'、p_r'$	重投影像素坐标

图像坐标系的建立如图 3-9 所示，以左上角为原点， $X$  轴向右， $Y$  轴向下， $Z$  轴向前。假设三维世界中的点  $P$  在左右相机的图像的像素点为  $p_l$  和  $p_r$ ，我们已知  $p_l$  的像素坐标，然后通过极线搜索的方法找到  $p_r$ ，建立以右相机中心为世界坐标系，坐标轴的方向与图像坐标系的坐标轴方向一致，通过异面直线公垂线<sup>[45]</sup>可得：

$$\begin{cases} Q_l Q_r \cdot l_1 = 0 \\ Q_l Q_r \cdot l_2 = 0 \end{cases} \quad (3-27)$$

即：

$$\begin{cases} a_1 = \frac{(\mathbf{p}_r - \mathbf{p}_l) \cdot \mathbf{l}_1 - ((\mathbf{p}_r - \mathbf{p}_l) \cdot \mathbf{l}_2) \cdot (\mathbf{l}_1 \cdot \mathbf{l}_2)}{1 - (\mathbf{l}_1 \cdot \mathbf{l}_2)^2} \\ a_2 = \frac{((\mathbf{p}_r - \mathbf{p}_l) \cdot \mathbf{l}_2) \cdot (\mathbf{l}_1 \cdot \mathbf{l}_2) - (\mathbf{p}_r - \mathbf{p}_l) \cdot \mathbf{l}_2}{1 - (\mathbf{l}_1 \cdot \mathbf{l}_2)^2} \end{cases} \quad (3-28)$$

则：

$$\mathbf{P} = \frac{\mathbf{Q}_l + \mathbf{Q}_r}{2} \quad (3-29)$$

其中  $\mathbf{l}_1、\mathbf{l}_2$  为直线  $O_l Q_l、O_r Q_r$  的单位向量； $\mathbf{Q}_l、\mathbf{Q}_r$  与  $\mathbf{p}_l、\mathbf{p}_r、a_1、a_2$  的关系为：

$$\begin{cases} \mathbf{Q}_l = \mathbf{p}_l + a_1 \mathbf{l}_1 \\ \mathbf{Q}_r = \mathbf{p}_r + a_2 \mathbf{l}_2 \end{cases} \quad (3-30)$$

通过上述方法，就可以得到点  $P$  在右相机坐标系下的坐标。点  $P$  坐标的准确性对后期的三维重建效果影响很大，而点  $P$  坐标的准确性依赖于像素坐标  $p_r$  的正确性。在实际的情况下，因为噪声、算法的鲁棒性等因素，都会影响  $p_r$  的正确性，所以误差基本上都会存在，所以点  $P$  重投影之后， $p_l'、p_r'$  与  $p_l、p_r$  不会完全重合，存在重投影误差。所以为了实现准确、可靠的匹配，需要在匹配的过程中加入各种已知的约束关系，常见的有极线约束、视差约束、唯一性约束、灰度相似性约束、几何相似性约束、光学测定学相容性约束、透视投影约束。

极线约束将计算复杂度从  $O(n^2)$  锐减到  $O(n)$ ，也就是从二维平面搜索范围变成了一维线搜索。如图 3-9 所示， $p_l$  对应的极线为  $e_r p_r$ 。左相机图中的每一个

点在右相机图中都存在唯一一条极线，对应点 $p_r$ 就在条极线上。所以，使用极线约束还有另外一个优点，提高了正确匹配的概率。

视差约束的作用是减少极线搜索的长度。点 $P$ 的深度范围在 0 到无穷远处。这个深度范围对应到极线上，就是极线的搜索长度。有的极线在图像上的很长，最长的是图像对角线那一根极线。本次实验的图像分辨率的大小为  $1280 \times 960$ ，那么对角线的那一条极线有 1639 个像素，如果不加任何约束，最坏的情况下会遍历 1639 个像素点。对应的深度本次课题针对的场景是对近处的物体进行三维重建，所以相机与镜头的距离在 350mm 左右，则加入的约束条件为 300mm 到 500mm。此时极线的搜索长度变成了 546 个像素，缩短了2/3。

唯一性约束指的左图像中的一个像素点在右图中有且只有一个对应的点。在进行匹配的时候，就可以选择相关性系数最大的那个。灰度相似性原则可以提高匹配的效率和正确性。虽然由于光照不均匀、图像噪声等因素会导致同一个点在左右相机的成像平面中的像素的灰度值不一致，但是也不会相差特别大，所以可以通过灰度值的预断剔除相似性极差的点，从而加快匹配的效率和正确性。

由于灰度值在左右相机中可能会出现不一样的结果，为了提高匹配的鲁棒性，选择以图像块的灰度值作为匹配的依据，降低立体匹配时受灰度值变化的影响，从而提高匹配的正确性。本文采用了零均值归一化互相关系数 (ZNCC) 作为匹配的依据，ZNCC 相关性计算公式如下：

$$C_{ZNCC}(x, y) = \frac{\sum_{j=-l}^l \sum_{i=-l}^l [(I_o^{i,j} - \bar{I}_o) (I_r^{i,j} - \bar{I}_r)]}{\sqrt{\sum_{j=-l}^l \sum_{i=-l}^l (I_o^{i,j} - \bar{I}_o)^2 \sum_{j=-l}^l \sum_{i=-l}^l (I_r^{i,j} - \bar{I}_r)^2}} \quad (3-31)$$

出于排版的整洁，将公式写在一行中，将其中的一些变量进行了简化，其中简化的变量如下：

$$I_o^{i,j} = I_{obj}(x + i, y + j) \quad (3-32)$$

$$I_r^{i,j} = I_{ref}(x + i, y + j) \quad (3-33)$$

$$\bar{I}_o = \overline{I_{obj}(x, y)} \quad (3-34)$$

$$\bar{I}_r = \overline{I_{ref}(x, y)} \quad (3-35)$$

其中  $I_{obj}(x + i, y + j)$  代表目标区域的中像素坐标为  $(x + i, y + j)$  的灰度值。 $\overline{I_{obj}(x, y)}$  代表目标区域中以  $(x, y)$  为中心的，大小为  $(2l + 1, 2l + 1)$  图像块的灰度值的平均数。

很多三维重建的方案中，采用了投影校正的方法，将汇聚式视觉模型校正为平行式视觉模型。本文并没有将这个办法加入三维重建算法中，有如下几个原因：1) 精度无法保证。因为本文需要实现较高精度的三维重建，投影校正采用逆向投影，然后通过像素插值计算得到投影校正之后的图像，用这个图像进

行三维重建。很显然在像素插值部分就已经丢失掉一部分精度，所以该方法不适合本文；2) 耗费额外的计算量。投影校正需要遍历图像中的每一个像素点，通过一定的校正方法得到校正之后的图像；当然这和方法也有一些优点，比如使用极线搜索的时候耗费的时间更少，也更靠谱。本文从三维重建的精度角度出发，没有采用投影校正的方法。

## 3.4 三维重建

### 3.4.1 伪随机码结构光图案区域分割

本文的三维重建基于伪随机码结构光，为了充分利用这个特点，需要将图像进行分割，只对图像中有伪随机码结构光图案区域进行三维重建。如图 3-10 所示，计算机获得左相机的图像，清晰可见左右图像中的伪随机码结构光图案区域。

立体匹配的精度和图像子区域的灰度值分布相关，子区域的与图像的其他区域的区分度越高，立体匹配的准确性和精度越高。本文设计的伪随机码图案，以其中某一个点为中心形成的方格图像在整个伪随机码图案的模板具有唯一性，与其他区域的相关性非常低，有助于提高匹配的精度。背景区域的灰度值分布是不可控制的，过分依赖于三维重建的物体会使得只能对灰度值变化较为明显的物体进行三维重建，这样算法的鲁棒性、实用性降低。如图 3-10 所示，是对一个雕像进行三维重建，这个雕像全身是乳白色，反应在图像中就是灰度值分布比较集中。

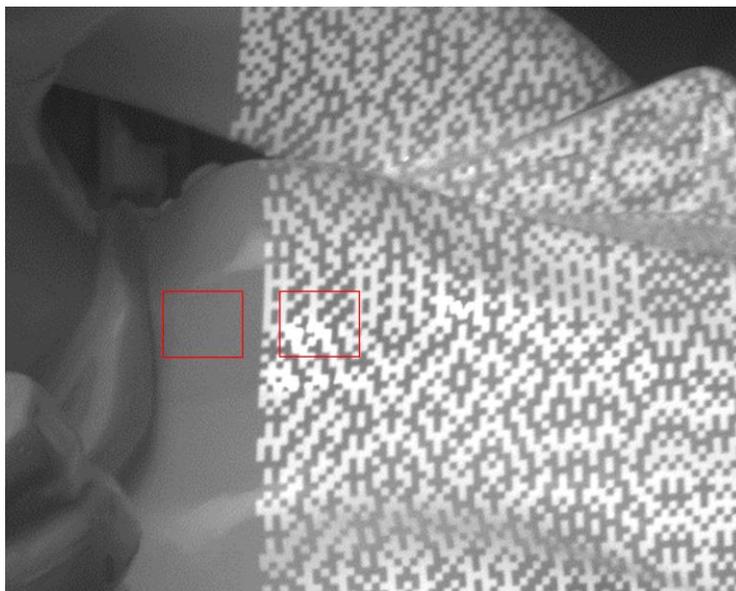


图 3-10 人体雕像部分图像

Figure 3-10. Partial image of a human statue

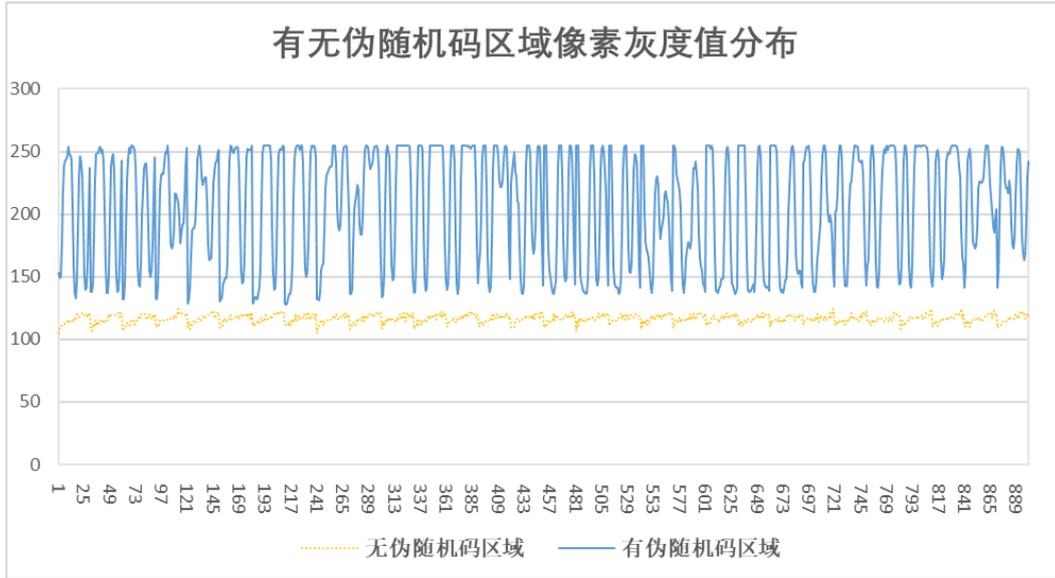


图 3-11 图像灰度值分布

Figure 3-11. Image gray value distribution

如图 3-10 所示，为相机获取的人体雕像的灰度图片的裁剪图片，伪随机码结构光图案打在该雕像上，有部分区域并未被伪随机码结构光覆盖到，各选取其中一块子区域输出它们的灰度值。需要注意的是，这两块子区域的选取一定要比较接近，如果相差的比较远那它们之间的灰度值对比就没有任何意义。选取了两块较为接近的区域，如图 3-10 中的红色框包围的两块区域，这两块区域大小都为  $30 \times 30$ ，左边的子区域的左上角的像素坐标为(110,435)，右下角的像素坐标为(140,465)；右边的子区域的左上角的像素坐标为(234,435)，右下角的像素坐标为(264,465)。这两块区域的灰度值分布如图 3-11 所示。

由图 3-10 可知，这两块区域靠的很近，又因为雕像本身是乳白色的，所以灰度值的分布应该相同。由图 3-11 可以看出，蓝色的直线是有伪随机码覆盖的区域的灰度值分布曲线，黄色的虚线是无伪随机码覆盖的区域的灰度值的分布曲线；横坐标是点的个数，这里一共选取了 900 个点，纵坐标是灰度值。可以看出，有伪随机码覆盖的区域的灰度值分布的方差更大，分布的范围是(128,255)，最大与最小值相差 127；无伪随机码覆盖的区域的灰度值分布的方差较小，分布的范围是(115,124)，最大与最小值相差仅为 9。可以看出，有伪随机码覆盖的区域计算其相关性系数准确率大大提高，没有伪随机码覆盖的区域很可能无法找到匹配点。

本文采用图像二值化方法分割出有伪随机码图案的区域，对于图像中的每一个像素点，用其邻域  $5 \times 5$  窗口大小的图像子区域简历平均差  $AD$ ：

$$AD = \frac{1}{25} \sum_{u=-2}^2 \sum_{v=-2}^2 \left| g(u,v) - \frac{1}{25} \sum_{u=-2}^2 \sum_{v=-2}^2 g(u,v) \right| \quad (3-36)$$

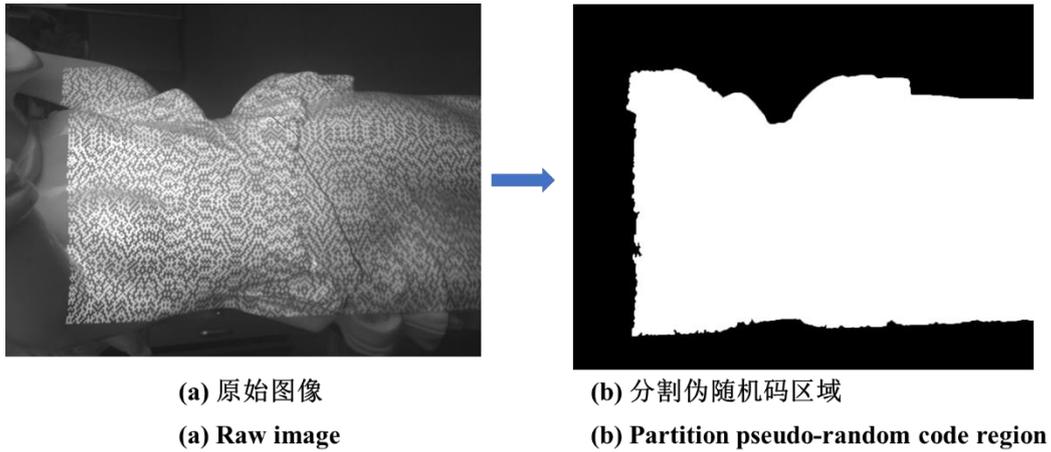


图 3-12 伪随机码区域分割

Figure 3-12. Pseudo-random coded region segmentation

假设给定分割阈值  $A$ ，如果当前像素点的  $AD$  值大于等于  $A$  时，记录下该像素点的坐标和灰度值，并将该像素点的灰度值设置成 255，即该像素点在伪随机码图案区域内；如果当前像素点的  $AD$  值小于  $A$  时，该像素点的不在伪随机码图案区域内。在实际的情况中，图像的灰度值会发生变化，分割阈值  $A$  不能对每种情况都表现出很好的分割效果，所以采用自适应确定阈值  $A$ ，以确保每次都能够实现很好的分割效果。如图 3-12 所示，分割出的伪随机码图案区域的效果。

### 3.4.2 对应点搜索

采用基于伪随机码结构光的相关方法搜索对应点，对于左图中的某一点，在右图中存在唯一一条极线与之对应。沿着这条极线计算每一个像素点的  $ZNCC^{[46]}$  相关系数， $ZNCC$  系数最大的那个像素点即为与左图对应的像素点。

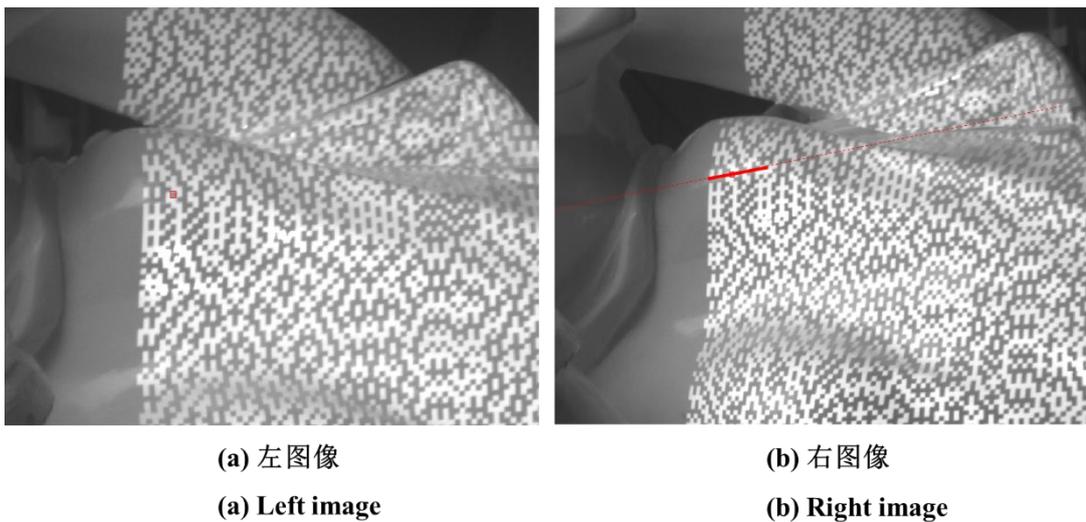


图 3-13 左右图像对应点与极线

Figure 3-13. Corresponding points and epipolar lines in the left and right images

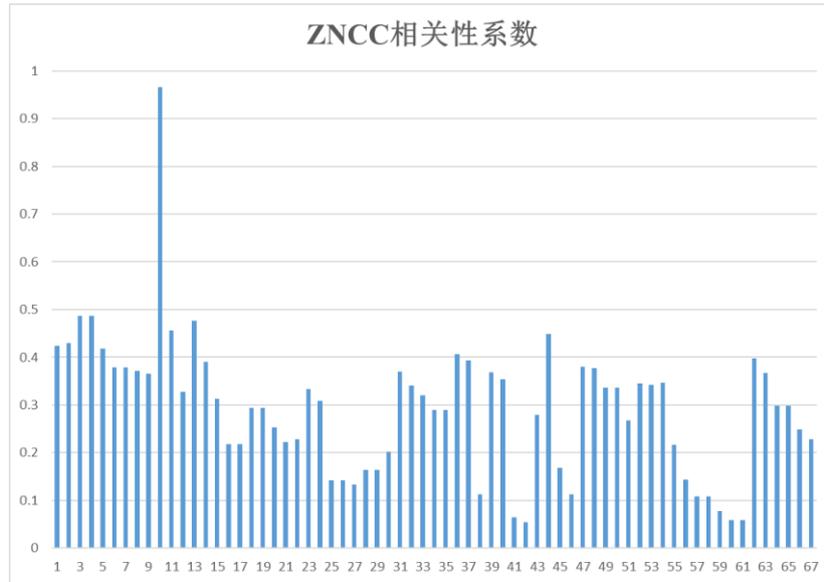


图 3-14 ZNCC 相关性系数

Figure 3-14. ZNCC correlation coefficient

如图 3-13 所示，选取了左右图像的对应点，极线的为图中红色虚线，红色实线是最终搜索的长度。在极线上搜索的并计算 ZNCC 相关性系数，ZNCC 相关性系数如图 3-14 所示。可以看出最高峰只有一个，并且与其它数值相差较大，最高峰对应的像素点就是要寻找的对应点。

通过对图像中的伪随机码图案区域进行对应点搜索，把重建出来的三维点坐标保存成文本格式后用 Geomagic Control 打开，如图 3-15 所示，雕像的三维点云。从图中可以看出，重建出了雕像大致形状，但是有一些误匹配的情况。虽然在立体匹配的过程中加入了极线约束、灰度约束、唯一性约束等约束条件，但是不能够完全消除误匹配的情况，需要进一步剔除误匹配以达到本文的要求。

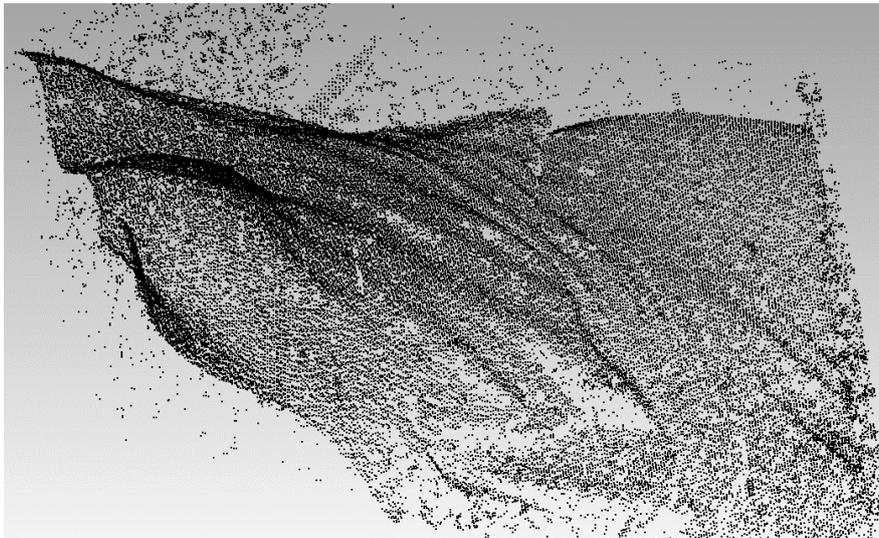


图 3-15 人体雕像的三维点云

Figure 3-15. Three-dimensional point clouds of human statues

### 3.4.3 剔除误匹配

基于视差连续性别除误匹配。三维空间中的物体在投影到相机成像平面的过程中，物体上相邻的点在图像上也是相邻的，所以重建出的三维点云中的点也是相邻的，即视差连续。如果这个点与周围的点视差不连续，说明在搜索匹配的点时，找到了错误的匹配点。通过这个特点，可以剔除点云中一部分误匹配的点。

基于离群点去除误匹配。点云中的点应该是连续的，并且每一个点的邻域内的点的数量不会相差太大。如果某个点的邻域内只有 1 个点，而其他点的邻域至少有 5 个点，那么可以推断出这个点是一个离群点，可以将其移除。在实践的过程中我们发现，半径设置为 5 个像素，邻域内的点数量为 8 可以达到较好的效果。

基于图像的空间邻近度和像素值相似度去除误匹配。同时考虑点云的空间信息与灰度相似性，使得很好的保存边缘，同时对点云起到平滑的效果。

去除误匹配的效果如图 3-16，由图可以看出，基本上达到了去除误匹配的要求。

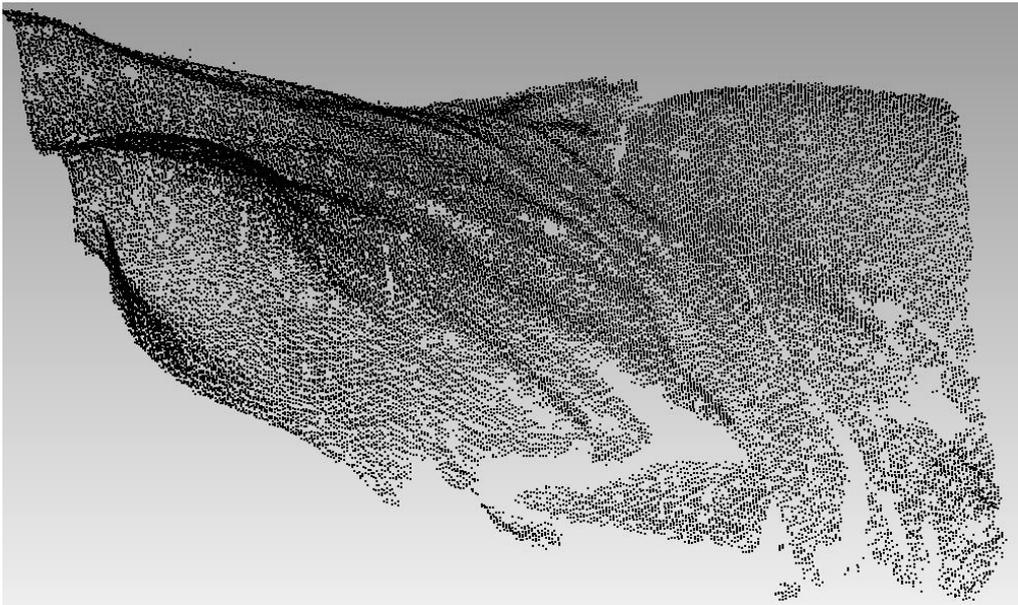


图 3-16 去除误匹配的三维点云

Figure 3-16. Remove mismatched 3D point clouds

## 3.5 点云拼接

到目前为止，我们只是获取了左右相机的各一幅图像，利用两幅图进行三维重建。这只能对物体的一面进行三维重建，我们需要对整个物体进行三维重建。连续获取图像，将不同视角下生成的点云通过一定的数据处理的方法整合

成整个物体点云。前面的章节已经实现了局部点云的生成，这一小节实现点云配准与点云融合。

在不同位置的获取图像的示意图，如图 3-17 所示。一个位置无法获得整个物体的全貌，需要在不同的位置获取图像。图中有被测物体、相机、反光标记点。

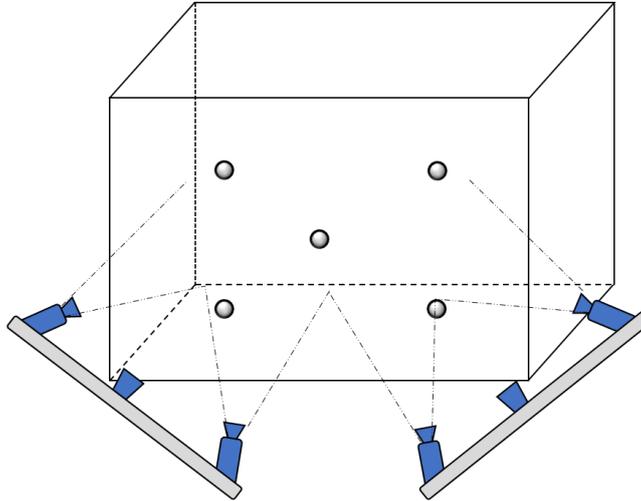


图 3-17 在不同位置获取图像

Figure 3-17. Obtain images of different positions

### 3.5.1 构建点云拓扑关系

构建点云的拓扑关系，方便存储和查询，特别是当点云的数据量很大时，高效可靠的拓扑关系决定了能否实时三维重建。常用的构建点云间的拓扑关系有八叉树法<sup>[47]</sup>、K-D 树法<sup>[48]</sup>、栅格化法、Voronoi 图法<sup>[49]</sup>、R 树法<sup>[50]</sup>等。每一种方法都有其适用的场景。例如 R 树法在高维空间搜索场景中表现突出，但是其动态扩展性不好。点云数据是不断生成的，需要不断向已有的点云数据中添加新的点云数据，需要有较好的动态扩展性，所以该方法不适合本文；K-D 树法在对点云数据进行构建 K-D 树后可以较快的进行搜索，但是对于大量的点云来说，遍历所有点云数据构建 K-D 树消耗的时间较长，所以该方法不适合用来构建大量点云数据的拓扑关系，所以不适合本文；八叉树法操作简单，根据事先是定的阈值分割点云数据，速度较快，虽然在过程中会产生空的方块，但是不会增加渐进复杂度。由于本文的点云数据量在 30 万左右，所以选择了八叉树法构建点云的拓扑关系。八叉树的构建过程如图 3-18 所示。

如图 3-18 中的左图所示，将立方体进行分割，每一次分割将方格分割成八个立方体，直到分割到符合分辨率的要求，停止分割。可以对应到图 3-18 中的右图，每一个节点有 8 个子节点，直到最后的叶节点。

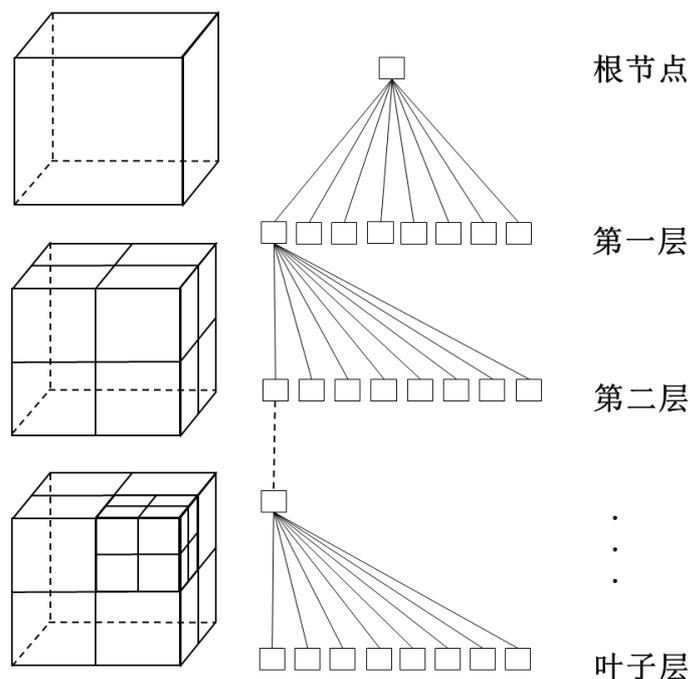


图 3-18 八叉树示意图

Figure 3-18. Octree diagram

### 3.5.2 点云配准

点云配准中较为常用的方法为迭代最近点（iterative closest point, ICP）算法，这个算法属于点云的精确配准。在精确配准之前需要一个初始值。初始值的好坏，对于 ICP 算法来说非常重要。因为 ICP 算法也属于非线性优化算法，初始值不理想，容易陷入局部最优。初值的获得需要进行点云的粗配准。

粗配准常用的方法有标签法、转台法、曲率特征法。本文采用了标签法和曲率特征法。转台法适用于物体较小的情况，并需要进行额外的硬件电路设计及通讯的方面的工作，与本文的契合度不高，故舍弃这一方法。标签法在物体表面贴标记点，通过标记点匹配完成粗配准。由于课题不是针对珍贵文物进行三维重建，此方法可以采用。曲率特征法是利用了刚体的表面几何特征不发生变化的特性，因此该方法适用于对表面几何特征较为明显的刚体进行三维重建的情况。故本文点云的粗配准方法融合标签法和曲率特征法。

假设两块点云进行配准，目标点云与源点云为  $p_t^i, p_s^i$ ，其中  $\{i = 1, 2, 3, \dots, n\}$ ，则粗配准获得的  $R$ 、 $t$  满足的关系式为：

$$p_t^i = R \cdot p_s^i + t \quad (3-37)$$

从源点云中选取  $k$  个采样点，采样点之间有一定的间隔，这里设置阈值为欧氏距离大于 10mm。遍历选取的  $k$  个点，计算每个点的 FPFH<sup>[51]</sup>特征，FPFH 的计算公式如下所示：

$$FPFH(p_q) = SPFH(p_q) + \frac{1}{k} \sum_{i=1}^k \frac{1}{\omega_k} \cdot SPFH(p_k) \quad (3-38)$$

在目标点云中找  $K$  个与源点云对应的点计算 FPFH 特征。精确配准是通过不断调整  $\mathbf{R}$ 、 $\mathbf{t}$  的值，使得式(3-39)的值最小。

$$f(\mathbf{R}, \mathbf{t}) = \frac{1}{n} \sum_{i=1}^n |\mathbf{p}_t^i - \mathbf{R} \cdot \mathbf{p}_s^i - \mathbf{t}|^2 \quad (3-39)$$

式(3-39)中的  $\mathbf{p}_t^i, \mathbf{p}_s^i$  为对应的点云对。通过式(3-37)通过  $\mathbf{R}$ 、 $\mathbf{t}$  将源点云变换到与目标点云相近的位置，因为这里的  $\mathbf{R}$ 、 $\mathbf{t}$  不够精确，所以两处点云基本上不会重合，但是对应的点的距离已经较小了。上一小节使用了八叉树法构建点云间的拓扑结构，基于这个特点可以较快的通过  $K$  近邻方法搜索到对应的点。此时计算对应点的 FPFH 特征。对比 FPFH，如果很接近说明找到的对应的点是正确匹配的，去除那些相差很大的点。有了对应点之后，重新估计  $\mathbf{R}$ 、 $\mathbf{t}$  的值并更新，代入式(3-39)获得该表达式的值进行迭代，如果式(3-39)的变化小于 0.05mm，则停止迭代，获得最终的  $\mathbf{R}$ 、 $\mathbf{t}$ 。点云配准的效果见第五章的实验部分。

### 3.6 三角网格化

基于前面的步骤，已经获得了整体的三维模型，这个三维模型依旧是一个个点组成，无法应用于三维打印等多个方面，需要将表面进行网格化。我们这里选用最常见的三角网格进行网格化。三角网格化有诸多的优点，例如稳定性好、网格简单、有助于恢复表面细节。

### 3.7 本章小结

本章对三维重建算法的各个部分进行了理论分析和实验，具体有相机的标定、立体匹配、三维重建、点云拼接和三角网格化这几个方面。总体上本文设计的三维重建算法能够实现对物体的三维重建，不过在精度和实时性等方面都有待提升，下一章节基于这两个方面对三维重建的算法进行提升。

## 第四章 基于实时性与精度的研究

在这一章节对三维重建的实时性和精度方面进行提升。在第三章部分，虽然已经对三维重建算法的各个部分进行了实现，但是仍然达不到实际应用的要求，这一章节加入一些策略，对三维重建算法的各个部分进行提升，以满足实际使用的需求。

### 4.1 图像预处理与增强

获取图像是基于视觉的三维重建算法的第一步，由于噪声的存在，直接对原始图像进行三维重建的效果并不理想，需要对图像进行预处理。图像预处理常见的方法有中值滤波、双边滤波、自适应中值滤波、自适应均值滤波等。每一种滤波方法尤其最佳的使用场景。在这里结合本文的需求，选择了双边滤波的预处理的方法。双边滤波对图像的边缘信息能够更好的保存，因此具有很高的保边去噪的效果。

双边滤波由两个函数组成，其中的两个滤波系数分别由几何空间和像素插值决定。双边滤波器输出结果为当前点的像素值，该值依赖于邻域像素的灰度值的加权组合，由式(4-1)给出：

$$gray(i, j) = \frac{\sum_{k,l} gray(k, l)w(i, j, k, l)}{\sum_{k,l} w(i, j, k, l)} \quad (4-1)$$

其中的 $w(i, j, k, l)$ 是权重系数因子，有定义域核和值域核的乘积共同决定。 $(i, j)$ 与 $(k, l)$ 代表像素点的坐标。定义域核为 $d(i, j, k, l)$ ，值域核为 $r(i, j, k, l)$ ，具体的表达式为：

$$d(i, j, k, l) = \exp\left(-\frac{(i-k)^2 + (j-l)^2}{2\sigma_d^2}\right) \quad (4-2)$$

$$r(i, j, k, l) = \exp\left(-\frac{\|gray(i, j) - gray(k, l)\|^2}{2\sigma_r^2}\right) \quad (4-3)$$

$$w(i, j, k, l) = d(i, j, k, l) \cdot r(i, j, k, l) \quad (4-4)$$

## 4.2 图像金字塔

### 4.2.1 模板匹配

图像处理中经常会用到图像金字塔<sup>[52]</sup>，原图像进行缩放后再进行图像处理。图像金字塔一般有两个用处，第一是提高精度，第二是加快速度。我们这里用来加快速度。直接在原图上极线搜索不仅搜索的距离较长，将原图缩小成 $1/4$ 后，图像中像素点的数量从原来的130万变成了32.5万。图像金字塔示意图，如图4-1所示。假设在 $1/4$ 倍原图中沿着极线搜索，通过模板匹配找到了相似区域是红色框包围起来的区域。映射到 $1/2$ 倍原图中，只需要在红色框包围的区域中沿着极线搜索，确定相似区域。最后将红色框区域映射到原图中，沿着极线搜索找到待匹配的点即可。这个方法不仅能加快速度而且能降低误匹配。

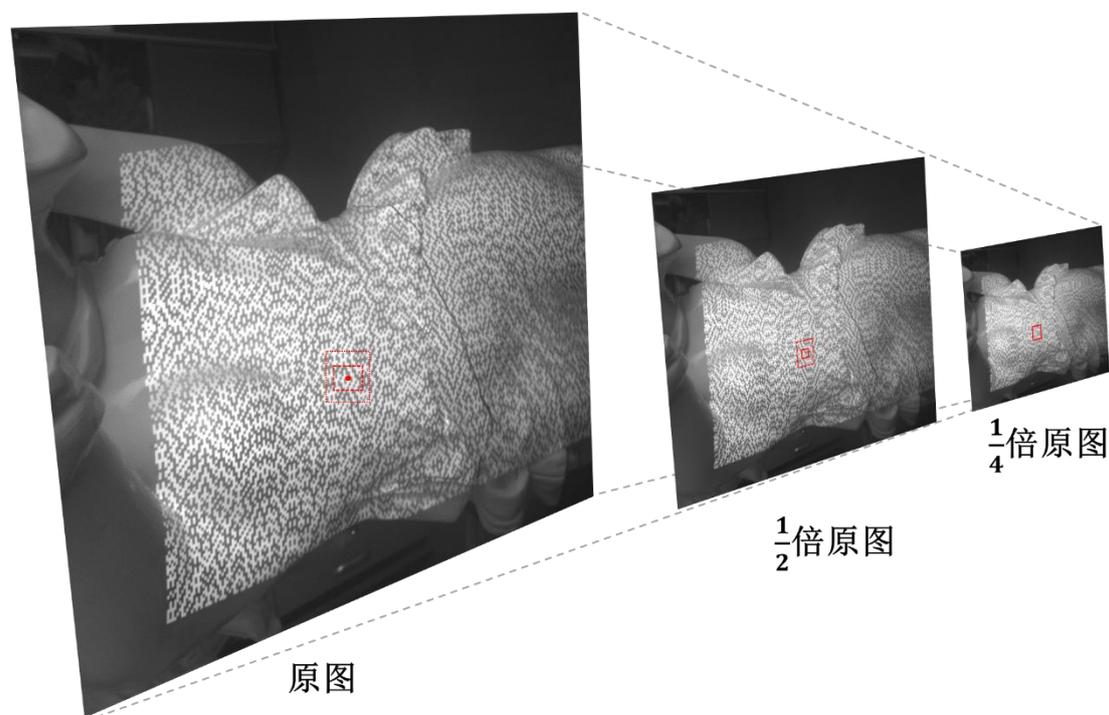


图 4-1 图像金字塔示意图

Figure 4-1. Image pyramid diagram

### 4.2.2 亚像素优化

出于三维重建精度的考虑，进行亚像素优化。在极线搜索的时候，假设参考点的坐标是整数，在目标区域中沿着极线搜索目标点时，找到了一个相关性系数最大的点，但不一定是正确的匹配点，因为目标点的坐标并不能保证是整数。进行亚像素搜索，有助于提高匹配的正确性，提高三维重建的精度。亚像素求解有很多常用的方法，比如线性插值法、双线行插值法和最近邻插值

法等。每个方法各有其使用的场景，具体使用哪种方法，根据具体的情况确定。

通常情况下伪随机码的一块黑色或者白色区域会占据不止一个像素大小。如图 4-2 所示，蓝色框包围起来的区域是伪随机码图案的一块黑色方块，占据了  $6 \times 6$  像素块，方块与方块之间的界限虽然已经变的很模糊了，但是有很明显的像素梯度。如图 4-2 中的蓝色框包围的像素点所示，其中一共包括了三个像素点，很明显的可以看出这三个像素点的灰度值存在很明显的像素梯度，这就为计算亚像素提供了方便，并且亚像素的变化不会太大。如果没有明显的像素梯度，再加上左右相机的视野获取的图像可能会发生变形后，亚像素甚至是像素的灰度值都可能发生很大变化，这在极线搜索找到正确匹配点时难度会大大增加。我们对比左右图中单个像素的灰度值可以看出，左图中每一个像素的灰度值都要比右图中对应的像素的灰度值要小，即左图要暗一点。这个是完全会发生的情况。本文模板匹配时采取的 ZNCC 相关性算法，这种情况完全不会影响该算法的鲁棒性和稳定性。

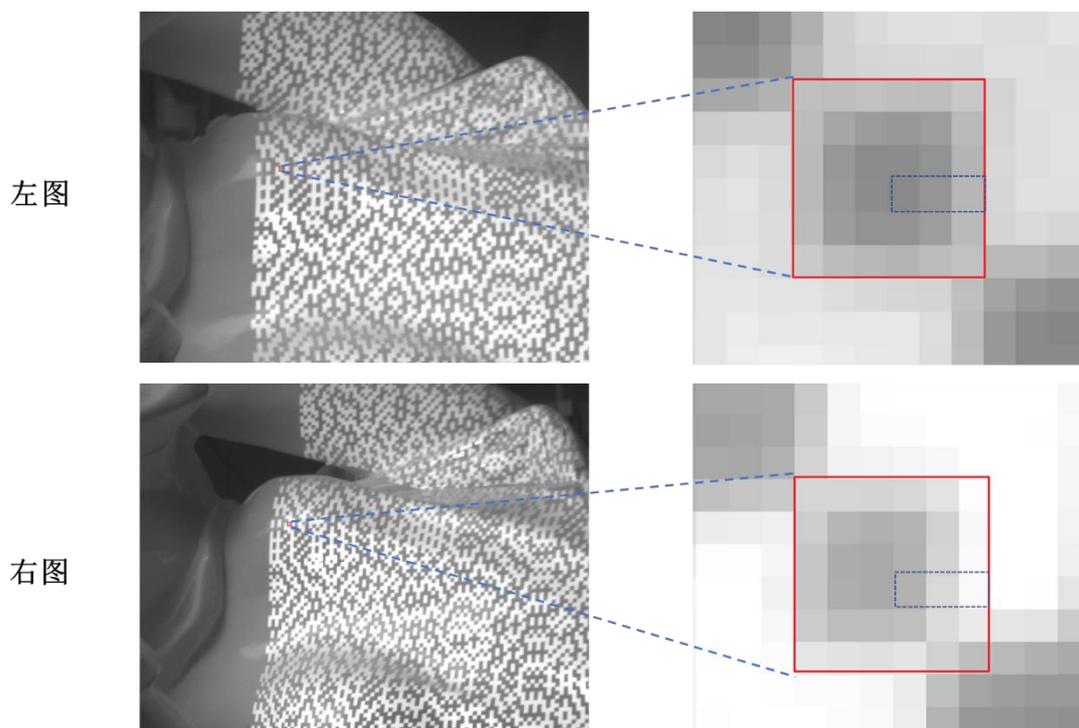


图 4-2 伪随机码单块方格

Figure 4-2. Pseudo-random code single square

采用双线性插值的方法计算亚像素。双线性插值的方法在实时性与准确性有较好的平衡。假设计算点  $p(x, y)$  的像素值，通过周围四个点的像素的灰度值计算得到。像素点的位置关系的示意图如图 4-3 所示。具体计算公式如下：

$$I_{R_1} \approx \frac{x_2 - x}{x_2 - x_1} I_{Q_{11}} + \frac{x - x_1}{x_2 - x_1} I_{Q_{21}} \quad (4-5)$$

$$I_{R_2} \approx \frac{x_2 - x}{x_2 - x_1} I_{Q_{12}} + \frac{x - x_1}{x_2 - x_1} I_{Q_{22}} \quad (4-6)$$

$$I_P \approx \frac{y_2 - y}{y_2 - y_1} I_{R_1} + \frac{y - y_1}{y_2 - y_1} I_{R_2} \quad (4-7)$$

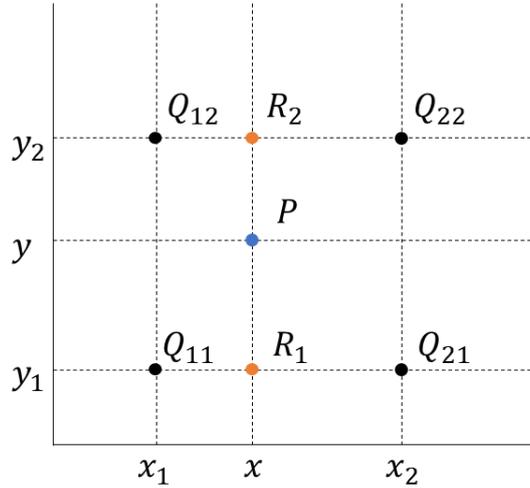


图 4-3 双线性插值

Figure 4-3. Bilinear interpolation

使用双线性插值的方法计算得到了亚像素的值，得到亚像素的值后进行相关性系数的计算。

使用伪随机码结构光比散斑图拥有更高的效率和更好的细节。伪随机码相对于散斑图有更好的细节效果。散斑是类似于小雪花片投影出去，相机获取到的图像也是类似的情况，里面的所有的小雪花片在整个图像中具有唯一性，在模板匹配的从理论上讲，具有很好的匹配效果，实际的使用并不是这样。投射散斑图像到物体上的时候，很多物体并不是一个平面，物体的表面是由多曲面拼接而成，曲面与曲面之间的地方会有陡峭的过渡，散斑图案在这里就会发生变形，左右相机获取的图像变化会很大，模板匹配的时候就会出错，即这种地方的重建效果就会比较差。使用伪随机码结构三维重建的时候光不会发生这种情况。通过图像金字塔确定了这目标点出现的区域后，进行亚像素求解的时候，根据上下左右 4 个点求取亚像素的值，变形后像素的梯度不会发生改变，因此具有较好的细节。

### 4.3 极线搜索优化

极线搜索在立体匹配的过程中起到了加速的作用，将搜索范围从整个图像平面缩小到一条线段上，从而将时间复杂度从  $O(n^2)$  减少到  $O(n)$ 。但是很多时候，极线还是很长，我们希望搜索的极线尽可能短，并且待匹配的点又在我们搜索的范围内。减少极线的长度不仅能缩减时间，而且能减少误匹配。

### 4.3.1 预计算确定平均深度

在上一章节中，我们简单的将待匹配点的深度范围限制在 300mm 到 500mm，虽然减少了2/3的搜索范围，还有 200mm 的搜索范围，可以进一步缩短。

先对图像中的点进行采样，用这部分点进行极线匹配，计算这部分点的深度均值和方差，如果方差超过阈值说明这些点的深度可能计算的不对，需要重新选取一部分点进行极线匹配。如果方差没有超过设定的阈值说明这些点的深度变化不大，将获得的平均深度作为一个深度基准，前后 50mm 作为搜索范围。与之前的 200mm 搜索范围相比，减少了1/2，即搜索时间减少了1/2。

本文中对图像中的点采样的步骤如下：将一张图像中的有伪随机编码图案分割出来，将分割出来的图像分为 36 块，每行和每列都进行 6 等分，如图 4-4 所示。图 4-4 中伪随机码区域是手工裁剪，作为示意，在实现的过程中，程序自动分割出伪随机码区域然后计算图像中部分点的三维坐标。

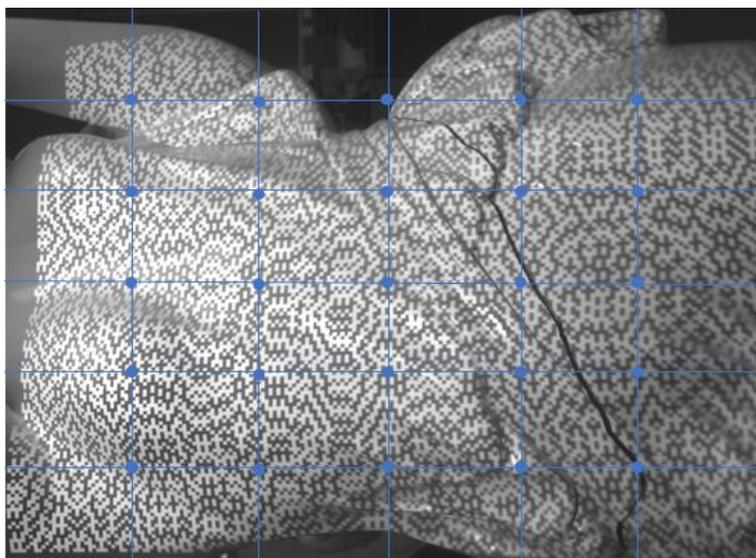


图 4-4 采样点示意图

Figure 4-4. Schematic diagram of sampling points

### 4.3.2 使用 ZNCC 快速算法

ZNCC 相关性系数的计算公式如式(3-31)，在每一次极线搜索的过程中都会使用到，而且在极线上的每一个点都会需要计算，因此降低 ZNCC 相关性系数的计算量会使得三维重建的实时性得到提高。

将 ZNCC 相关性系数的计算公式(3-31) 改写成如下形式：

$$C_{ZNCC}(x, y) = \frac{(N * S_{OR} - S_O * S_R)}{\sqrt{(N * S_{OO} - S_O^2) * (N * S_{RR} - S_R^2)}} \quad (4-8)$$

$$S_{OR}(x, y) = \sum_{j=-l}^l \sum_{i=-l}^l I_{obj}(x + i, y + j) * I_{ref}(x + i, y + j) \quad (4-9)$$

$$S_O(x, y) = \sum_{j=-l}^l \sum_{i=-l}^l I_{obj}(x + i, y + j) \quad (4-10)$$

$$S_R(x, y) = \sum_{j=-l}^l \sum_{i=-l}^l I_{ref}(x + i, y + j) \quad (4-11)$$

$$S_{OO}(x, y) = \sum_{j=-l}^l \sum_{i=-l}^l I_{obj}(x + i, y + j)^2 \quad (4-12)$$

$$S_{RR}(x, y) = \sum_{j=-l}^l \sum_{i=-l}^l I_{ref}(x + i, y + j)^2 \quad (4-13)$$

在一条极线上搜索匹配点时，重复计算的量有 $S_R$ 、 $S_{RR}$ 。在实现的过程中使用积分图的方法保存这两个值。

#### 4.3.3 使用多线程方法加速

在立体的匹配的过程中，遍历左图像中的点后极线搜索。在遍历的过程中，左图像中的每个像素点都是孤立存在，他们之间没有影响。基于这个特点，可以使用多线程进行加速运算。本文使用的 CPU 是 6 核 12 线程，可以同时开启 12 个线程并行计算，可以将计算时间缩减到 1/12。

#### 4.4 本章小结

本章节对三维重建的精度与实时性进行进一步的研究，将三维重建的效果进行提升。主要从立体匹配这一块进行了提升，使用了图像金字塔，亚像素优化、多线程等方法进行精度与速度的提升。

## 第五章 三维重建系统的精度分析与实验

### 5.1 精度分析

对物体进行三维重建后，需要进行 3D 打印或者其他用途时，需要获取该模型的精度。从理论分析和从实验分析该本文的三维重建系统的精度。

#### 5.1.1 理论分析

Thomas Luhmann<sup>[53]</sup>给出了简化版的精度分析，得出的结论是深度方向的测量误差与测量距离的平方成正比，x 与 y 方向的误差与距离成正比。其中采用的模型与本文不相同。本文采用的是汇聚式视觉模型,该模型采用的是平行式视觉模型。那么其中有两个假设在本文中不成立，两个相机的光轴平行与极线与光轴垂直。基于这个特点，结论依然适用，即深度方向的测量误差与测量距离的平方成正比，x 与 y 方向的误差与距离成正比。

#### 5.1.2 实验验证

实验验证分为两个部分，第一部分是在不同距离下对一个同一个墙面进行三维重建，从而获得最佳的重建距离；第二部分是在最佳的重建距离对步距球规三维重建的精度，从而获得该三维重建系统的精度。

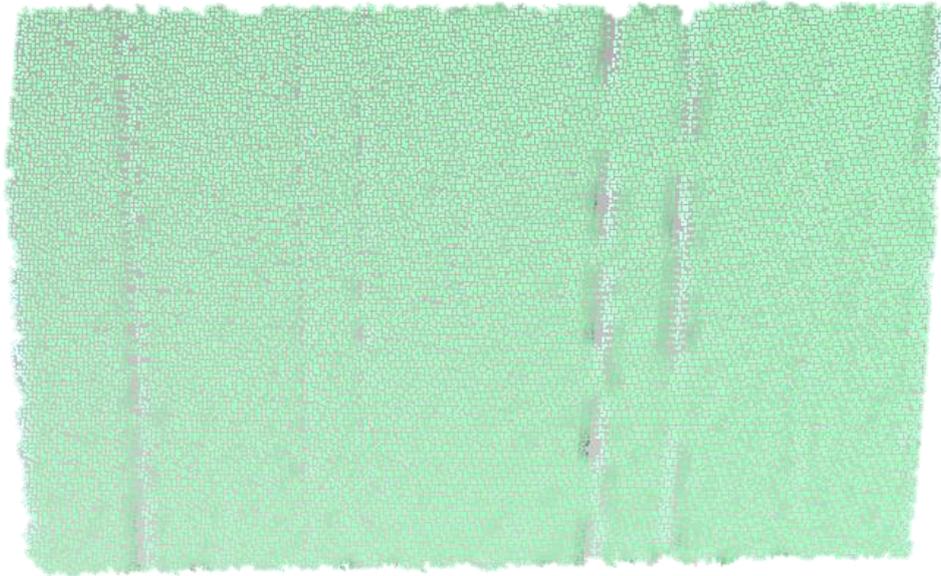


图 5-1 墙面的三维重建结果

Figure 5-1. 3D reconstruction of the wall surface

为了获得本文设计的三维重建系统最佳的重建距离，在不同距离下对同一个墙面进行三维重建，距离变化从330mm到650mm左右，一共采集五次图像。从理论上应该是对标准的平面进行三维重建，从而对比三维重建的结果，这里由于没有理想的平面，选取了白色墙面，可以大致代替理想平面。三维重建的结果如图5-1。对三维重建的结果进行数据分析，如图5-2所示。

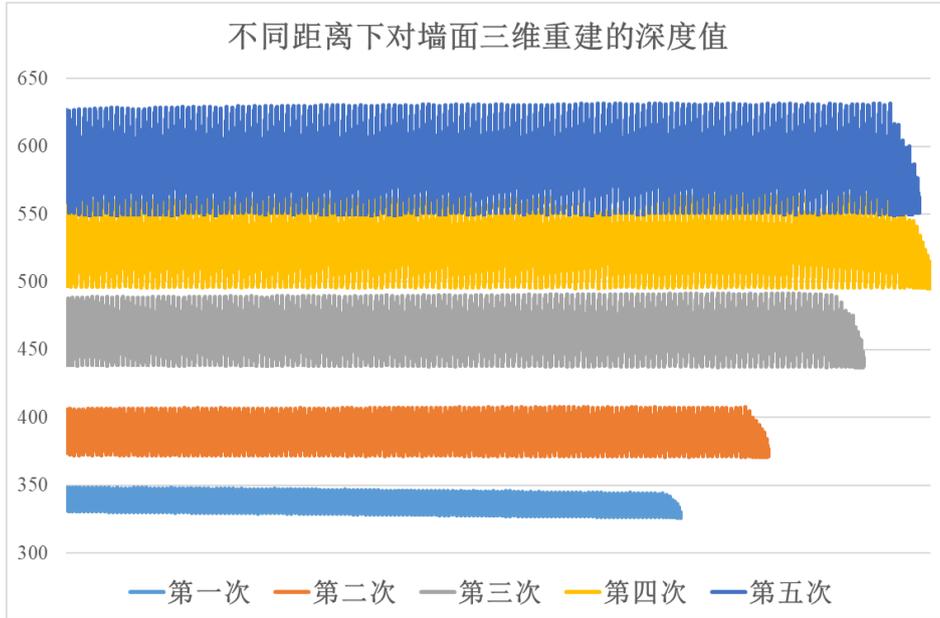


图 5-2 不同距离下对墙面三维重建的深度值

Figure 5-2. Depth values of 3D wall reconstruction at different distances

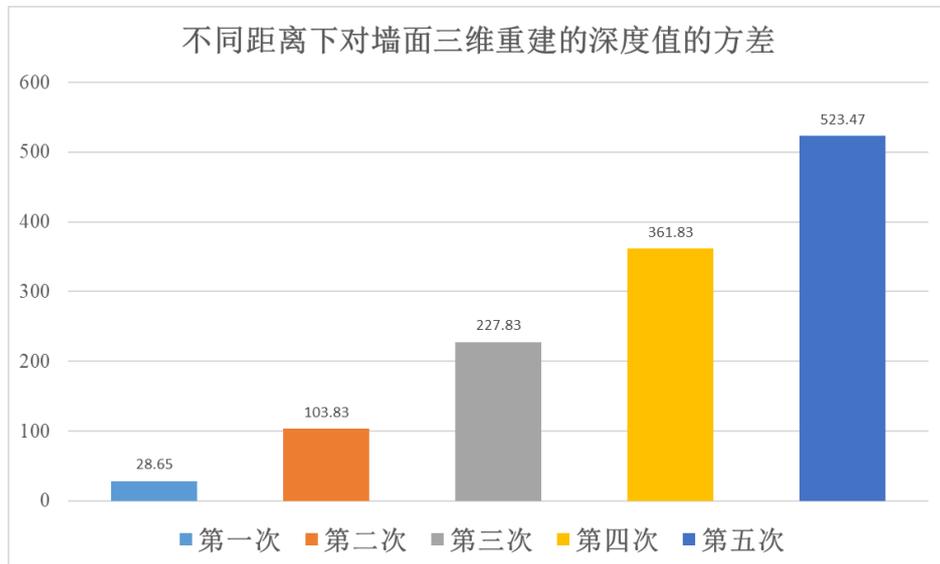


图 5-3 不同距离下对墙面三维重建的深度值的方差

Figure 5-3. Variance of depth values of 3D wall reconstruction at different distances

由图5-2和图5-3可以看出，随着重建距离的不断变大，深度的变化越来越大，即三维重建的精度越来越差。当距离在350mm左右时，深度变化的最小，

即该距离三维重建的效果最好。关于图中数据的的两点说明：其一是图中的数据是未经滤波处理过的数据，为了验证本文设计的三维重建系统的比较合适的重建距离。在后期进行三维重建的过程中，加入滤波和非线性优化方法优化三维重建的结果；其二是该平面并非标准的平面，平面本身就会有有很大的偏差，所以最小的方差有 28.65mm 也很正常。

为了得到本文设计的三维重建系统的精度，本文对如图 5-4 的步距球规进行三维重建，该步距球规的参数如表所示。该步距球规的参数由三坐标测量仪测量得到，精度可以保证。所以，将三维重建的结果与其参数进行比较即可。

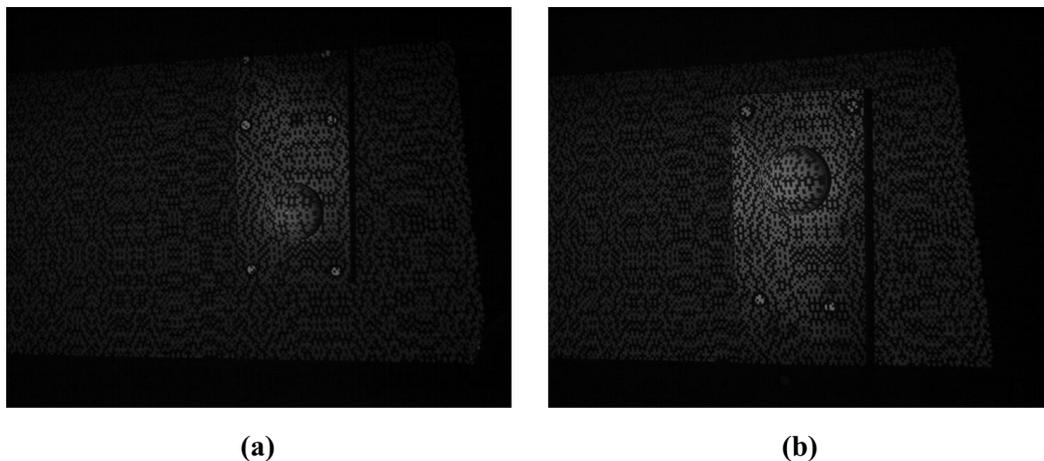


图 5-4 步距球规  
Figure 5-4. Step gauge

表 5-1 步距球规参数  
Table 5-1. Step gauge parameters

参数名	数据
球心距	300.00mm
球 1 直径	55.02mm
球 2 直径	55.01mm

当相机离步距球规约 350mm 时，左右相机各获取两张张图像后可以覆盖步距球规的全部区域，右相机获取的图像如图 5-5 所示。对其进行三维重建并去除非连接项，三维重建的点云结果如图 5-6 所示。



(a)

(b)

图 5-5 左相机获取步距球规图像  
Figure 5-5. Left camera captures the step gauge image

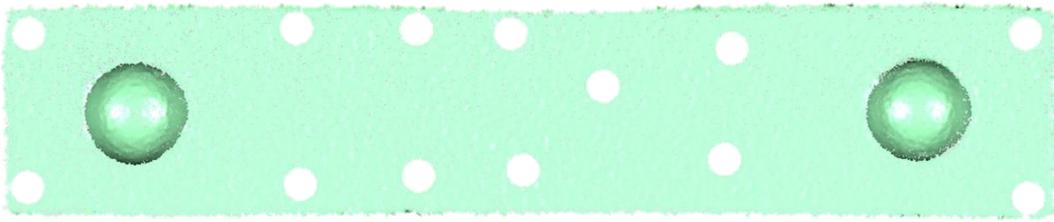


图 5-6 步距球规的点云

Figure 5-6. Point cloud of step gauge

将测量得到的数据使用 Geomagic Control 进行处理，将两个陶瓷球的点云数据拟合成球体，测量两个球心的距离。

通过对步距球规进行 10 次三维重建，将获取的数据制作成图表，如图 5-7 所示。从图中可以看出，精度的较为稳定，围绕标准值上下波动，精度与标准值最大相差 0.036mm。

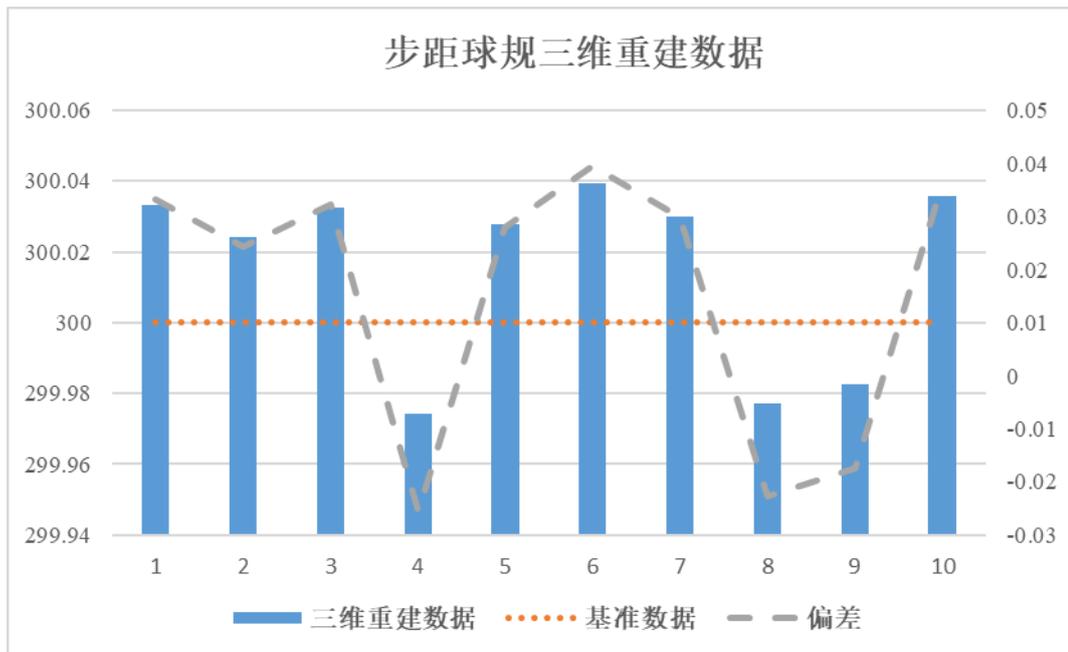


图 5-7 步距球规的三维重建数据

Figure 5-7. 3D reconstruction data of step gauge

## 5.2 不同场景下重建效果分析

本文设计的三维重建系统，目的是对于近处的物体进行三维重建，这一小节对两种不同的物体进行三维重建，一个是平面物体，测试该三维重建系统还原细节的能力；第二个是类圆柱体，测试该三维重建系统对于曲率较大的物体的重建效果。

### 5.2.1 平面物体的三维重建

在上一小节我们对白色墙面进行了三维重建并对其进行了分析。这个白色墙面并不带有任何的纹理与其他特征，并不能反映本文设计的三维重建系统对细节还原的效果。这里选择细节比较丰富的长城画像，如图 5-8 所示，该物体的具体参数如表 5-2 所示。如图 5-9 所示，长城浮雕的三维重建结果。



图 5-8 长城浮雕

Figure 5-8. The Great Wall relief

表 5-2 长城浮雕参数

Table 5-2. The Great Wall relief parameters

参数名	数值
长	573mm
宽	191mm
高	5mm

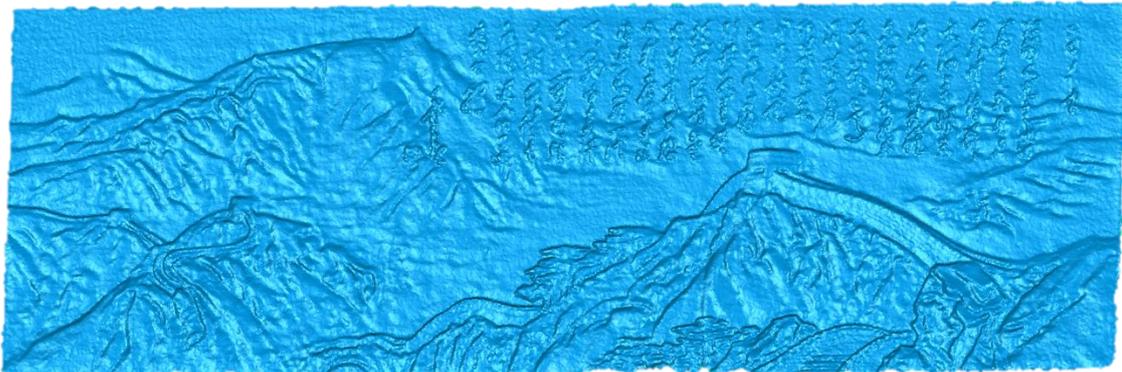


图 5-9 长城浮雕的三维重建结果

Figure 5-9. 3D reconstruction results of the Great Wall relief

从图 5-9 中可以看出，本文设计的三维重建系统对细节有一定的还原能力，可以清晰的看出长城、树木、诗句的轮廓，不过对于诗句而言，只能重建出大概的轮廓，并不能清晰的分辨出其中的字。

### 5.2.2 类圆柱体的三维重建

对类圆柱体的三维重建，可以验证本文设计的三维重建系统对曲率较大的物体三维重建的效果。本文选择的类圆柱体是电影《死侍》中的主角模型，模型图 5-10 所示。该物体的尺寸参数如表 5-3 所示，三维重建的结果如图 5-11 所示。



图 5-10 死侍模型

Figure 5-10. Deadpool model

表 5-3 死侍模型参数

Table 5-3. Deadpool model parameters

参数名	数值
长	110mm
宽	110mm
高	341mm



图 5-11 死侍模型的三维重建结果

Figure 5-11. 3D reconstruction results of Deadpool model

从图 5-11 中可以看出, 本文设计的三维重建系统对于类圆柱体的物体能够进行三维重建, 并且效果满足本文的要求。

### 5.3 本章小结

本章对三维重建的精度进行了分析, 从理论和实际验证两个角度进行了分析, 最终得到本文设计的三维重建系统的精度为 0.04mm。对细节丰富的平面物体和类圆柱体进行了三维重建实验, 实时处理数据的帧率为 15 帧, 重建出来的点云数据量在 30 万点左右。对其进行了分析, 实验结果表明, 本文设计的三维重建系统都能够得到较好的重建效果。

## 第六章 总结与展望

### 6.1 总结

目前三维检测、3D 打印、无人车、智能机器人视觉、虚拟现实和增强现实等领域对三维重建有着较高的需求，本文在实际需求的驱动下，从算法、硬件、软件三个方面对基于伪随机编码结构光的三维重建系统进行分析和设计，最终实现了精度高、实时性、稳定性好、鲁棒性强的三维重建系统。本文的主要工作总结如下：

(1) 开发了一套三维重建的硬件系统，该系统由两个工业相机、投影机、LED 补光灯、电路板、线缆组成。经过大量的测试，最终确定该系统所使用的硬件设备。该系统具有轻巧、灵活、便携性好等特点，能够对大型或者小型的物体进行实时三维重建，并实时显示三维重建的点云。

(2) 制作了高精度的标定板，多次标定双目相机提高相机的标定精度，并对标定的结果进行分析。基于 MFC 框架开发了一个相机内参比较器软件，该软件用于分析相机标定的结果，可以直观的比较前后两次相机标定结果的差异。

(3) 研究了伪随机编码的生成方法，设计并生成了符合本文要求的伪随机编码，本文设计的伪随机编码大小为 $320 \times 147$ ，由 7 个大小为 $320 \times 21$ 伪随机编码拼接而成。在大小为 $320 \times 21$ 的伪随机码区域中，任意 $5 \times 5$ 大小的区域具有唯一性。该编码方案具有编码简单，精度高等特点。

(4) 结合伪随机编码结构光、图像金字塔、模板匹配这三个方面实现高精度并快速的立体匹配，在细节和效率方面提升效果显著，从而每秒钟处理 15 张图像，达到实时重建的要求。

(5) 基于实时性与精度两个方面对三维重建的算法进行了研究，在原有的三维重建算法中加入一些策略，提高三维重建的效果，具体有图像预处理、图像金字塔、积分图保存相关性系数、多线程计算。

(6) 基于 MFC 框架开发三维重建软件，该软件用于基于伪随机编码结构光的三维重建，该软件将相机的标定、算法的实现、硬件的调试都包含在内，大大方便了三维重建的实验和后续的使用。

## 6.2 展望

基于伪随机编码结构的三维重建系统，日后有广阔的使用场景，还需要科研人员不断的进行改进与优化。本文虽然实现了对物体的三维重建，仍然有很多可以改进的地方。针对以后的设想和改进如下：

(1) 基于伪随机编码结构的三维重建系统，是立体视觉三维重建算法的一种妥协。不投射任何光线，采取被动式三维重建方案同样能够实现高精度的三维重建效果。

(2) 加入惯性测量单元（inertial measurement unit, IMU）传感器降低计算量，并提升三维重建效果。IMU 传感器属于内感受传感器，感知自身姿态的变化，相机属于外感受传感器，感知场景的变化。目前这两种传感器在智能飞行器与虚拟现实中的应用较多。

(3) 将计算平台更换成嵌入式。本文使用的计算平台是个人电脑，将所有的计算移到嵌入式芯片中，可以进一步提高三维重建系统的灵活性和实用性。

(4) 对点云拼接这一部分算法改进，使得点云拼接的效果鲁棒性与精度更好。

(5) 对非刚体进行三维重建，本文只针对刚体进行了三维重建的研究。

## 参考文献

- [1] AVIDAN S, SHASHUA A. Trajectory triangulation: 3D reconstruction of moving points from a monocular image sequence[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2000, 22(4): 348-357.
- [2] BRUNO F, BIANCO G, MUZZUPAPPA M, et al. Experimentation of structured light and stereo vision for underwater 3D reconstruction[J]. ISPRS Journal of Photogrammetry & Remote Sensing, 2011, 66(4): 508-518.
- [3] BAAK A, MILLER M, BHARAJ G, et al. A Data-driven approach for real-time full body pose reconstruction from a depth camera [M]. London: Springer, 2013:71-98.
- [4] GENG J. Structured-light 3D surface imaging: a tutorial[J]. Advances in Optics & Photonics, 2011, 3(2): 128-160.
- [5] VAN DER JEUGHT S, DIRCKX J J. Real-time structured light profilometry: a review[J]. Optics & Lasers in Engineering, 2016, 87: 18-31.
- [6] LIN H, NIE L, SONG Z. A single-shot structured light means by encoding both color and geometrical features[J]. Pattern Recognition, 2016, 54: 178-189.
- [7] RIOUX M. Laser range finder based on synchronized scanners[J]. Applied Optics, 1984, 23(21): 3837-3844.
- [8] OSHIMA M, SHIRAI Y. Object recognition using three-dimensional information[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 1983(4): 353-361.
- [9] SONG Z, TANG S, GU F, et al. DOE-based structured-light method for accurate 3D sensing[J]. Optics and Lasers in Engineering, 2019(120): 21-30.
- [10] KOWARSCHIK R M, KUEHMSTEDT P, GERBER J, et al. Adaptive optical 3-D-measurement with structured light[J]. Optical Engineering, 2000, 39(1): 150-158.
- [11] ZUO Y, ZHANG J, Zhang Y Y, et al. Weight threshold check coding for dimmable indoor visible light communication systems[J]. IEEE Photonics Journal, 2018, 10(3): 1-11.
- [12] HU X, ZHAO Q, YU P, et al. Dynamic shaping of orbital-angular-momentum beams for information encoding[J]. Optics Express, 2018, 26(2): 1796-1808.
- [13] CASPI D, KIRYATI N, SHAMIR J. Range imaging with adaptive color structured light[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 1998, 20(5): 470-480.
- [14] 于晓洋, 吴海滨, 尹丽萍, 等. 格雷码与相移结合的结构光三维测量技术[J]. 仪器仪表学报, 2007, 28(12): 2152-2157.
- [15] KONINCKX T P, LUC V G. Real-time range acquisition by adaptive structured light[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2006, 28(3): 432-445.
- [16] PAG S J, SALVI J, COLLEWET C, et al. Optimised De Bruijn patterns for one-shot shape acquisition[J]. Image & Vision Computing, 2005, 23(8): 707-720.
- [17] CHEN S, LI Y F, ZHANG J. Vision processing for realtime 3-D data acquisition based on coded structured light[J]. IEEE Transactions on Image Processing, 2008, 17(2): 167-176.
- [18] 陆军, 李积江, 黄春明. 符号 M 阵列结构光的解码[J]. 光学精密工程, 2013, 21(4): 1069-1078.
- [19] SPOELDER H J, VOS F M, PETRIU E M, et al. Some aspects of pseudo random binary array based surface characterization[C]. //Proceedings of the IEEE Instrumentation and Measurement Technology Conference. Piscataway: IEEE, 1998: 2-7.

- [20] GRIFFIN P M, NARASIMHAN L S, YEE S R. Generation of uniquely encoded light patterns for range data acquisition[J]. *Pattern Recognition*, 1992, 25(6): 609-616.
- [21] ALBITAR C, GRAEBLING P, DOIGNON C. Design of a monochromatic pattern for a robust structured light coding[C]. //Proceedings of the IEEE International Conference on Image Processing. Piscataway: IEEE, 2007: VI-529-VI-532.
- [22] 张广军, 李鑫, 魏振忠. 结构光三维双视觉检测方法研究[J]. *仪器仪表学报*, 2002, 23(6): 604-607.
- [23] 王宝光, 贺忠海, 陈林才, 等. 结构光传感器模型及特性分析[J]. *光学学报*, 2002, 22(4): 481-484.
- [24] 罗先波, 钟约先, 李仁举. 三维扫描系统中的数据配准技术[J]. *清华大学学报*, 2004, 44(8): 1104-1106.
- [25] 樊强, 陈大为, 习俊通. 高精度激光点扫描三维测量系统及应用[J]. *上海交通大学学报*, 2006(02): 227-230.
- [26] SUNG C, PAN Y K. 3D terrain reconstruction of construction sites using a stereo camera[J]. *Automation in Construction*, 2016, 64: 65-77.
- [27] GEIGER A, LENZ P, URTASUN R. Are we ready for autonomous driving? the kitti vision benchmark suite[C]. //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2012: 3354-3361.
- [28] AASEN H, BURKART A, BOLTEN A, et al. Generating 3D hyperspectral information with lightweight UAV snapshot cameras for vegetation monitoring: From camera calibration to quality assurance[J]. *ISPRS Journal of Photogrammetry & Remote Sensing*, 2015, 108(5): 245-259.
- [29] FAIG W. Calibration of close-range photogrammetric systems: Mathematical formulation[J]. *Photogrammetric Engineering & Remote Sensing*, 1975, 41(12): 1479-1486.
- [30] DAINIS A, JUBERTS M. Accurate remote measurement of robot trajectory motion[C]. //Proceedings of the IEEE International Conference on Robotics and Automation. Piscataway: IEEE, 1985: 92-99.
- [31] TSAI R. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses[J]. *IEEE Journal on Robotics and Automation*, 1987, 3(4): 323-344.
- [32] ZHANG Z. A flexible new technique for camera calibration[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2000, 22(11): 1330-1334.
- [33] WANG Y, LAI Z, HUANG G, et al. Anytime stereo image depth estimation on mobile devices[C]. //Proceedings of the International Conference on Robotics and Automation. Piscataway: IEEE, 2019: 5893-5900.
- [34] ROSTEN E, DRUMMOND T. Machine learning for high-speed corner detection[C]. //Proceedings of the European Conference on Computer Vision. Berlin: Springer: 430-443.
- [35] LOWE D G. Distinctive image features from scale-invariant keypoints[J]. *International Journal of Computer Vision*, 2004, 60(2): 91-110.
- [36] BAY H, TUYTELAARS T, VAN GOOL L. Surf: Speeded up robust features[C]. //Proceedings of the European Conference on Computer Vision. Berlin, Heidelberg, Germany: Springer, 2006: 404-417.
- [37] LIN Q, YANG R, ZHANG Z, et al. Robust stereo-match algorithm for infrared markers in image-guided optical tracking system[J]. *IEEE Access*, 2018(6): 52421-52433.

- [38] YIN Z, DARRELL T, YU F. Hierarchical Discrete Distribution Decomposition for Match Density Estimation[C]. //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2019: 6044-6053.
- [39] ZHOU Y, GALLEGO G, REBECQ H, et al. Semi-dense 3D reconstruction with a stereo event camera[C]. //Proceedings of the European Conference on Computer Vision. Berlin, Heidelberg, Germany: Springer, 2018: 235-251.
- [40] ZANG Y, YANG B, LI J, et al. An accurate TLS and UAV image point clouds registration method for deformation detection of chaotic hillside areas[J]. Remote Sensing, 2019, 11(6): 647-653.
- [41] YU J, LIN Y, WANG B, et al. An advanced outlier detected total least-squares algorithm for 3D point clouds registration[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(7): 4789-4798.
- [42] NESHVAD S, CHATZINOTAS S, SACHAU J. Wideband identification of power network parameters using pseudo-random binary sequences on power inverters[J]. IEEE Transactions on Smart Grid, 2015, 6(5): 2293-2301.
- [43] PAOLINI E, STEFANOVIC C, LIVA G, et al. Coded random access: applying codes on graphs to design random access protocols[J]. IEEE Communications Magazine, 2015, 53(6): 144-150.
- [44] NGUYEN T, SLAUGHTER D, MAX N, et al. Structured light-based 3D reconstruction system for plants[J]. Sensors, 2015, 15(8): 18587-18612.
- [45] 罗世民, 李茂西. 双目视觉测量中三维坐标的求取方法研究[J]. 计算机工程与设计, 2006, 27(19): 3622-3624.
- [46] LIN C, LI Y, XU G, et al. Optimizing ZNCC calculation in binocular stereo matching[J]. Signal Processing Image Communication, 2017, 52: 64-73.
- [47] BEŇO P, DUCHOŇ F, PAVELKA V, et al. Using octree maps and RGBD cameras to perform mapping and A\* navigation[C]. //Proceedings of the International Conference on Intelligent Networking & Collaborative Systems. Piscataway: IEEE, 2016: 66-72.
- [48] KLOKOV R, LEMPITSKY V. Escape from cells: Deep kd-networks for the recognition of 3d point cloud models[C]. //Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE, 2017: 863-872.
- [49] DADUSH D, BONIFAS N. Short paths on the Voronoi graph and closest vector problem with preprocessing[C]. //Proceedings of the Twenty-sixth Acm-SIAM Symposium on Discrete Algorithms. Pennsylvania: Society for Industrial and Applied Mathematics Philadelphia, 2015: 295-314.
- [50] OOSTEROM P V, MARTINEZ-RUBI O, IVANOVA M, et al. Massive point cloud data management: Design, implementation and execution of a point cloud benchmark[J]. Computers & Graphics, 2015, 49: 92-125.
- [51] RUSU R B, BLODOW N, BEETZ M. Fast point feature histograms (FPFH) for 3D registration[C]. //Proceedings of the IEEE International Conference on Robotics and Automation. Piscataway: IEEE, 2009: 3212-3217.
- [52] SCH PS T, SCH NBERGER J L, GALLIANI S, et al. A Multi-View Stereo Benchmark with High-Resolution Images and Multi-Camera Videos[C]. //Proceedings of the IEEE Conference on Computer Vision & Pattern Recognition. Piscataway: IEEE, 2017: 3260-3269.
- [53] LUHMANN T, ROBSON S, KYLE S, et al. Close-Range Photogrammetry and 3D Imaging[J]. Photogrammetric Engineering & Remote Sensing, 2015, 81(4): 273-274.

## 致 谢

时光荏苒，岁月如梭。光阴如流水一般，不曾停留，两年半的研究生生涯即将结束。在此，我要感谢这一路走来所有给予我帮助的老师、同学、家人及朋友。

首先，我要感谢浙江工业大学。感谢浙江工业大学在这两年半时间里的栽培，让自己有在知识、科研等多个方面都等到了提升。更要感谢我的研究生导师赵云波教授。非常有幸成为您的学生，在这两年半期间对我的学业及科研工作给予悉心的指导与帮助，才能得以攻克科研上的一个又一个难题。此外赵老师严谨的逻辑推理、一丝不苟的科研态度让我受益匪浅，将一直伴随着我走上社会。在此，真诚地感谢您。

非常感谢信息工程学院给予我指导的各位老师，是你们无私地将自己的知识输出，才有了我的不断成长，在这里，衷心地感谢你们。

同时，还要感谢我的同门师兄弟及给予我帮助的学长学姐们。他们是我在这两年半的研究生生涯中相处时间最多的人，他们不仅在学术上给予了我很多的帮助与指导，让我在这个过程中走更少的弯路，能够将主要的精力集中在关键的科研任务上。

我还要感谢我的家人，包括我的父母和两个姐姐，是他们在背后给我最大的支持与鼓励。父母总是给我无微不至的关心，让我在科研的道路上没有后顾之忧。在这里，由衷地感谢他们的养育之恩。

然后，我要感谢我的女友朱芳芳同学。感谢她一路的陪伴与鼓励。在科研的道路上，总会遇到学术难题，是她用笑容化解我内心的烦闷，让我再一次投入到科研任务中时，心如止水，勇敢向前。

最后，诚挚地感谢各位专家、评委能在百忙之中审阅本人的硕士毕业论文。

## 作者简介

### 1 作者简历

1994年2月出生于江苏省兴化市。

2013年9月——2017年6月，常州工学院电气与光电工程学院自动化专业学习，获得工学学士学位。

2017年9月——2020年1月，浙江工业大学信息工程学院控制工程专业学习，获得工程硕士学位。

### 2 参与的科研项目及获奖情况

[1] 基于资源调度和预测控制的无线网络化控制系统的联合设计. 国家自然科学基金面上项目(61673350).

### 3 发明专利

[1] 赵云波, 蒋传鹏, 倪洪杰, 等. 一种更鲁棒的视觉里程计的设计方法 [P]. CN109523589A, 2019-03-26.

## 学位论文数据集

密 级*	中图分类号*	UDC*	论文资助
公开	TH741	681.2	
学位授予单位名称	学位授予单位代码	学位类型*	学位级别*
浙江工业大学	10337	全日制工程硕士	硕士
论文题名*	基于伪随机编码结构光的三维重建系统设计与实现		
关键词*	伪随机编码, 结构光, 相机标定, 立体匹配, 三维重建		论文语种*
并列题名*	Design and Implementation of 3D Reconstruction System Based on Pseudo-random Coded Structured Light		中文
作者姓名*	蒋传鹏	学 号*	2111703386
培养单位名称*	培养单位代码*	培养单位地址	邮政编码
浙江工业大学 信息工程学院	10337	杭州市留和路 288 号	310023
学科专业*	研究方向*	学 制*	学位授予年*
控制工程	三维重建	2.5 年	2020 年
论文提交日期*	2020 年 1 月		
导师姓名*	赵云波	职 称*	教授
评阅人	答辩委员会主席*	答辩委员会成员	
盲评	石崇源	赵云波、宣琦、周晓、陈晋音	
电子版论文提交格式: 文本 ( ) 图像 ( ) 视频 ( ) 音频 ( ) 多媒体 ( ) 其他 ( )			
电子版论文出版(发布)者	电子版论文出版(发布)地		版权声明
论文总页数*	58 页		
注: 共 33 项, 其中带*为必填数据, 为 22 项。			