

西安邮电大学

毕业设计（论文）

题目： 基于机器视觉的深度相机设计与实现

学院： 计算机学院

专业： 计算机科学与技术

班级：

学生姓名：

学号：

导师姓名： 职称：

起止时间： 2018 年 11 月 19 日 至 2019 年 6 月 7 日

毕业设计（论文）声明书

本人所提交的毕业论文《基于机器视觉的深度相机设计与实现》是本人在指导教师指导下独立研究、写作的成果，论文中所引用他人的文献、数据、图件、资料均已明确标注；对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式注明并表示感谢。

本人完全理解《西安邮电大学本科毕业设计（论文）管理办法》的各项规定并自愿遵守。

本人深知本声明书的法律责任，违规后果由本人承担。

论文作者签名：

日期： 年 月 日

西安邮电大学本科毕业设计(论文)选题审批表

申报人			职 称		学 院	计算机学院	
题目名称	基于机器视觉的深度相机设计与实现						
题目来源	科研				教学		其它
题目类型	硬件设计	√	软件设计		论文		艺术作品
题目性质	应用研究		√		理论研究		
题目简述	<p>随着机器视觉、自动驾驶、机器人的火爆,采用深度相机采集环境的深度信息然后进行物体识别、环境建模等越来越普遍;相对于传统相机,深度相机增加了一维的深度,因而,能够更好的对真实世界进行描述;在许多领域如安防、监控、机器视觉、机器人等,拓展了更多的可能;如自动驾驶中的物体识别和障碍物检测,工业中散乱码放物体的识别、分拣、拆垛、码垛,物流场景中物体的货架抓取等等。</p>						
对学生知识与能力要求	<p>对学生知识与能力要求:</p> <ol style="list-style-type: none"> 1. 了解深度相机采集数据的工作原理; 2. 掌握 FPGA 硬件开发技术; 3. 掌握基于机器视觉的深度相机的实现方法。 						
具体任务以及预期目标	<p>设计一种基于空间编码进行三角测量原理的高精度深度相机(RGB-D 相机)设备;该设备具有抗环境光干扰、扫描精度高和硬件结构简单等优势;可用于过程控制、事件监测、物体与环境建模、设备交互和图像信息组织等工程方面应用。</p>						
时间进度	<p>2018.11.19 - 2018.12.16 理解和弄懂毕业设计任务和要求以及工作计划。根据毕业设计任务查找和搜寻相关参考资料在理解和研究相关资料的基础上编写完成毕业设计开题报告。</p> <p>2018.12.17 - 2019.3.12 学习和掌握 FPGA 开发工具和系统设计,并能熟练应用。分析理解机器视觉开发原理和实现方式,在分析和理解的基础上,开始构造和实现毕业设计所要求的模块化设计。</p> <p>2019.3.13 - 2019.5.19 使用开发工具和开发设计技术进行基于机器视觉得深度相机的设计开发与实现,同时完成对系统的测试和完善。</p> <p>2019.5.20 - 2019.5.31 整理毕业设计资料,设计总结,完成毕业设计论文的编写工作等。翻译一篇相关的英文文献,英文字符数不少于二万(中文 3000 字以上)。</p> <p>2019.6.1 - 2019.6.7 提交论文,完成毕业答辩。</p>						
系(教研室)主任 签字				主管院长 签字			
	年 月 日				年 月 日		

西安邮电大学本科毕业设计（论文）开题报告

学生姓名		学号		专业班级	
指导教师		题目	基于机器视觉的深度相机设计与实现		
<p>选题目的（为什么选该课题）</p> <p>图像传感器（Image Sensor）是一种将光学影像转换成电子信号的设备，长久以来作为计算机接收现实世界信息的重要端口设备；但随着时代的进步，技术的发展，单纯的二维图像数据已经不能满足 AR/VR、SLAM、自动驾驶、工业制造、工业流水线检测、自助机器人、三维重建、逆向工程、新式人机交互方式的需求；市场对于深度相机的需求逐年增长。使用深度相机采集周围环境的三维空间信息然后进行物体识别、环境建模等应用越来越普遍与重要；相对于传统相机，深度相机增加了距离这个对于现在的技术来说非常重要的信息，因而，深度相机采集到的深度图像可以更好的对真实世界进行描述；使用在在诸多领域如安防、监控、机器视觉、自主机器人等，使得单纯使用传统相机所不能实现的功能成为可能；如实现无人机或者无人小车的自动驾驶，使用深度摄像头进行视觉导航、识别外界的环境、规划路径、实现自动避障工作等，基于深度摄像头的视觉导航已经成为了公认的解决方案；实现工业机器人的自主运行以达到工业流水线的完全自动化；深度相机可以对工业产品进行三维重建，构建三维模型来进行工业流水线检测，相较于传统的机器视觉检测精度更高，更加准确；在物流行业，自助机器人将自动的将散乱码放物体的识别、分拣、拆垛、码垛等，甚至通过计算货物体积来实现最优化装载。</p>					
<p>前期基础（已学课程、掌握的工具，资料积累、软硬件条件等）</p> <p>1. 已学课程 《C 语言程序设计》、《工程制图与计算机制图》、《数字电路与逻辑设计》、《基于 Verilog 的 FPGA 设计基础》、《计算机图像处理》。</p> <p>2. 使用 Quartus、Modelsim、SolidWorks、Altium Designer 等设计工具，FPGA 开发硬件。</p>					
<p>要研究和解决的问题（做什么）</p> <p>设计与制作从通过 FPGA 硬件采集的各种信息后通过软件系统和 FPGA 加速处理，到得出深度图像为止全套的深度相机硬件以及配套的软件系统。结合市面上常见的深度相机实现方案的优势，弥补这些方案的固有缺陷，在保证精度与可靠性的前提下，降低深度相机的硬件成本，加强硬件的可定制化，使深度相机可以方便地嵌入各领域的生产过程以及最终产品中去。</p>					

工作思路和方案（怎么做）

初期方案考察与设计：初期研究已有的深度相机实现方案，如最为常见的双目、TOF 与结构光等三种方案，通过网络查找各方面的相关资料，了解其他深度相机方案的详细设计、优缺点等信息，最终确定使用以 FPGA 为核心的类结构光实现方案，这种方案结构简单、提升了精度与可靠性而且便于扩展。

深度相机硬件设计，分为三个主要模块进行设计：

1. 图像传感模块：由图像传感器与配套的驱动电路构成；负责记录光栅投射模块的标记周期并将图像数据传回主板做处理

2. 光栅投射模块：由无刷电机驱动模块、无线输电模块和标记周期脉冲模块等子模块构成，主要负责稳定的输出标记周期信号对空间进行标记与图像传感模块配合实现三角测量。

3. 主板：主板上集成了 FPGA 芯片以及其他通信模块，负责协调各模块之间的工作以及处理图像传感模块传回的图像数据，根据图像数据解算深度信息；向上位机输出深度数据以及接收上位机的控制指令等功能。

软件架构设计：软件需要进行深度解算以及误差校正等工作，主要进行畸变矫正，使用差分法消除干扰，图像边缘提取，深度解算的过程。图像传感模块在成像过程中，由于各种原因，相机模型不可能是一个完美的针孔成像模型，造成实际像与根据单透镜理论确定的理想像的偏离，这种误差造成的成像变形被称作相机畸变。相机畸变是光学透镜固有的透视失真的总称，也就是因为透视原因造成的失真，相机成像也不是一个完美的透射变换，这将会对三角测量造成严重影响。通过软件算法标定相机，通过算法得出相机模型的实际参数（焦距，焦点位置等信息）以及用来矫正畸变的参数，通过这些参数来修正变形的图像，再结合相机标定过程中得出的焦距，焦点位置等数据（比直接测量得出的数据更加精准），可以有效地减少三角测量的误差。

验证测试：验证测试与开发是同步进行的，在编写软件代码以及 FPGA 上运行的硬件代码的过程中建立测试环境，测试例程与开发代码互相完善。

指导教师意见

签字：

年 月 日

西安邮电大学毕业设计（论文）成绩评定表

学生姓名		性别		学号		专业 班级				
课题名称										
指导教师 意见	支撑指标点/赋分	3.1/10	3.2/20	3.4/10	4.3/10	8.2/20	10.1/10	12.1/10	12.2/10	合计
	得分									
	指导教师(签字): 年 月 日									
评阅 教师 意见	支撑指标点/赋分	3.1/10	3.2/30	3.4/10	4.3/20	10.1/10	12.1/10	12.2/10	合计	
	得分									
	评阅教师(签字): 年 月 日									
验收 小组 意见	支撑指标点/赋分	3.1/10	3.2/10	3.4/10	4.2/30	4.3/30	10.1/10	合计		
	得分									
	验收小组(签字): 年 月 日									
答辩 小组 意见	支撑指标点/赋分	3.1/10	3.2/10	3.4/10	4.2/10	4.3/10	10.1/50	合计		
	得分									
	答辩小组组长(签字): 年 月 日									
学生总评 成绩	评分比例	指导教师(20%)		评阅教师(30%)		验收小组(20%)		答辩小组(30%)		合计
	评分									
	毕业论文(设计)最终等级制成绩(优秀、良好、中等、及格、不及格)									
答辩委员会 意见	学院答辩委员会主任(签字、学院盖章): 年 月 日									

摘 要

双目立体视觉技术通过视差原理可以恢复目标场景的三维结构，具有结构简单、不容易受到光照干扰等优势，使得双目立体视觉技术有较大的研究价值；但由于双目立体视觉技术存在受目标场景的表面纹理、光照条件的影响较大、视差图的获取需要很复杂的计算等缺陷，导致双目立体视觉技术迟迟无法广泛应用。

本文提出了使用条带状图样标记辅助实现双目立体视觉技术的系统实现方案；通过结合结构光的设计思想，尝试解决传统双目立体视觉方案的固有缺陷。通过理论分析和搭建实验平台进行测试，验证了该系统实现方案确实提高了传统双目立体视觉方案的视差精度、降低了软件算法的复杂度，具有较高的研究价值。

但由于本文提出的系统实现方案是在传统双目立体视觉方案的硬件结构的基础上额外增加了标记模块，使得该系统实现方案不再具备传统双目立体视觉方案结构简单的优势、减慢了恢复目标场景三维结构的速度，还导致该系统实现方案对环境光的干扰更加敏感，降低了该系统实现方案的实用性。针对这些在实验测试过程中暴露出来的潜在问题，我们在文章末尾提出新的系统实现方案进行修正。

关键字：双目立体视觉；图像分割；线结构激光；三角测量

ABSTRACT

Binocular stereo vision technology can restore the three-dimensional structure of the target scene through the principle of parallax. It has the advantages of simple structure and not easy to be interfered by illumination, which makes the binocular stereo vision technology have great research value. However, binocular stereo vision technology suffers from The surface texture of the target scene and the influence of the illumination conditions are large, and the acquisition of the parallax map requires complicated calculations and the like, which leads to the delay in the application of the binocular stereo vision technology.

This paper proposes a system implementation scheme that realizes binocular stereo vision technology by using strip pattern labeling; by combining the design idea of structured light, it tries to solve the inherent defects of the traditional binocular stereo vision scheme. Through theoretical analysis and experimental platform construction, it is verified that the system implementation scheme has improved the parallax accuracy of the traditional binocular stereo vision scheme and reduced the complexity of the software algorithm, which has high research value.

However, because the system implementation scheme proposed in this paper is based on the hardware structure of the traditional binocular stereo vision scheme, the marking module is added, so that the system implementation scheme not only no longer has the advantages of the simple structure of the traditional binocular stereo vision scheme, but also slows down. The speed of restoring the three-dimensional structure of the target scene also makes the system implementation scheme more sensitive to ambient light interference, which reduces the practicability of the system implementation scheme. In response to these potential problems exposed during the experimental test, we propose a new system implementation at the end of the article to correct.

Keywords: Binocular stereo vision; Image segmentation; Line structure laser;
Triangulation-based measurement

目 录

第一章 绪论	1
1.1 研究背景、目的和意义	1
1.2 国内外主流实现方案	1
1.3 本文的主要研究内容和创新之处	3
1.4 本文的结构与内容安排	4
第二章 开发环境和理论原理概述	5
2.1 开发环境	5
2.1.1 OpenCV	5
2.1.2 OpenMP	5
2.1.3 SystemVerilog	5
2.1.4 FPGA	5
2.2 相机的透视投影模型	6
2.3 双视图几何模型	8
第三章 本系统实现方案的原理概述	11
3.1 本系统实现方案的系统模型	11
3.2 条带状图样中线提取算法	12
3.3 相差与场景扫描	14
第四章 本系统实现方案的实验装置设计	17
4.1 实验装置设计	17
4.1.1 实验装置的硬件设计	17
4.1.2 实验装置的软件设计	23
4.2 测试方案设计	26
4.2.1 标定测试	26
4.2.2 条带状图样中线提取算法测试	26
4.2.3 标记模块控制器测试	27
4.2.4 重建精度测试	27
第五章 实验结果	28
5.1 实验平台展示	28
5.2 测试结果展示	28
5.2.1 标定结果与分析	28
5.2.2 条带状图样中线提取算法结果与分析	30
5.2.3 标记模块控制器测试结果与分析	31
5.2.4 重建精度测试结果与分析	31
5.3 实验结果总结	33

5.3.1 目标场景视差图结果分析	33
5.3.2 潜在问题分析	34
5.3.3 解决方案探讨	35
结束语	36
致谢.....	37
参考文献	38

第一章 绪论

1.1 研究背景、目的和意义

相机作为我们生活中常见的硬件设备，它可以捕获目标场景的表面光学特征，将光电模拟信号量化得到数字信号并传输给计算机用于处理或者保存，是计算机观察现实世界的重要的端口设备，也是机器视觉技术主要也是非常重要的数据来源。

机器视觉技术经过几十年的发展，基于传统相机输出的二维数字图像（以下简称图像）的机器视觉技术和应用领域已经比较成熟和完善。传统的图像仅包含目标场景二维光强分布信息，但随着时代的进步和技术的发展，单纯的二维图像已经不能满足 AR/VR、实时定位与地图重建^[1]（Simultaneous Localization And Mapping 以下简称 SLAM）、自动驾驶、工业自动化制造、工业流水线检测、自主机器人、三维重建、逆向工程、新式人机交互方式的需求^[2]。机器视觉领域需要提供在不接触目标场景表面的情况下（非接触式）直接提供或者间接计算出目标场景的三维结构信息的解决方案。

深度图像不仅能提供传统图像所能提供的所有信息，还能额外提供目标场景与相机的相对距离信息。可以直接提供深度图像或者可以通过软件间接计算出深度图像的软硬件配套的系统被称作深度相机（本文主要讨论非接触式扫描）。深度图像的出现很好满足机器视觉的各个应用领域对目标场景三维结构信息的需求。

深度相机技术在机器视觉领域的推广，使得很多使用传统二维图像难以实现的技术得以应用，这些新技术的应用对我们生活带来的巨大的便利。深度相机可以用于智能家居、手势识别、目标跟踪、人脸识别、自动背景虚化等我们生活常见的领域；深度相机得到的深度图像可以建立周围环境详细的三维模型和实现自身的定位，这些优势使得深度相机技术在 3D 打印、自动化的缺陷检测和逆向工程等工业领域也有着极大的需求；深度相机与机器人控制技术、SLAM 技术相结合，可以实现机器人的自主导航和操作、生产流水线的全自动化、无人驾驶。新的技术的广泛应用将改变我们的生活。

1.2 国内外主流实现方案

实现深度相机的方案有很多种，我们根据深度相机自身是否向目标场景发射额外的能量而将深度相机的实现方案大致分为非接触式主动扫描和非接触式被动扫描两种。

主流的非接触式主动式扫描深度相机有时差测距（Time-of-Flight）、三角测

距 (Triangulation)、结构光源 (Structured Lighting) 和调变光 (Modulated Lighting) 四种。

时差测距方案又称作飞秒测距方案 (如图 1.1 所示), 通过主动的向场景发射高能光束脉冲, 测量光束脉冲照射在目标场景表面后返回的时间差以此计算出目标场景与相机的距离^[3]。这种深度相机的实现方案具有有效量程远、成像精度稳定 (时差测距方案测量精度只与控制芯片的运行频率有关, 距离的变化不会影响精度)、体积小 (不需要基线) 等优点, 故在专业领域有着广泛的应用; 但由于其控制芯片成本高昂、深度图像分辨率低等缺陷导致其未能广泛应用。

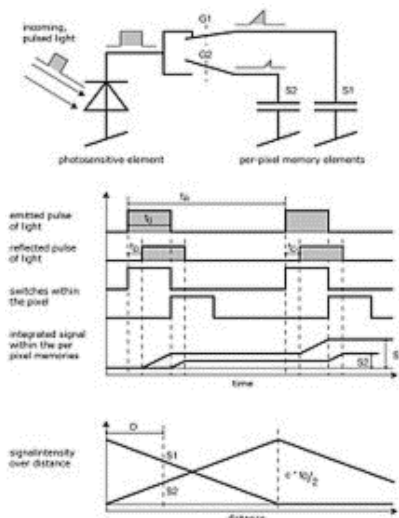


图 1.1 时差测距原理图

三角测距方案基于光沿直线传播的原理, 向目标场景发射一道光束 (如图 1.2 所示, 点光源或者线光源), 使用普通的相机拍着目标场景并搜寻该光点/线, 目标点/线与激光发射器、相机构成了三角形, 在这个三角形中, 我们已知相机和激光器的距离 (基线长度) 和激光器与基线的夹角, 通过光点/线的位置我们可以准确的构建三角形计算出目标点/线的距离。该方案具有控制系统简单、设备价格低廉、近距离测量精度高等优势, 但也存在随着距离的增加精度急剧下降、体积较大 (需要有一定的基线长度)、遮挡导致无法测量的缺陷。

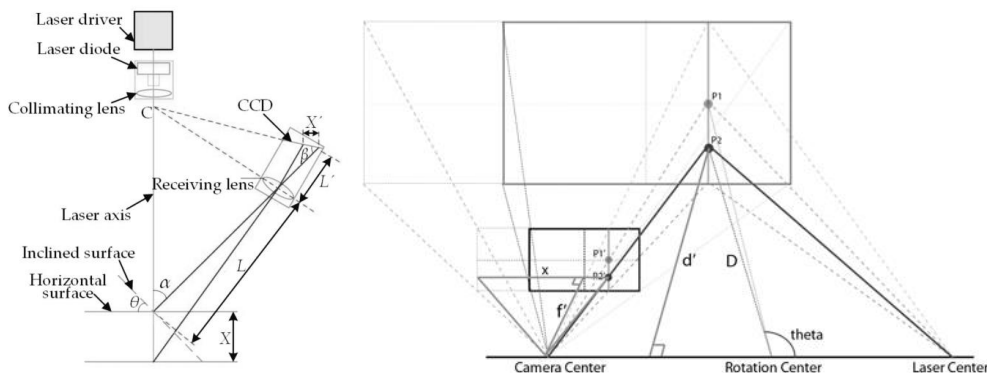


图 1.2 点光源三角测量 (图左) 和线光源三角测量 (图右)

结构光方案和调变光方案则是通过向目标场景投射特殊的图案，根据图案的变化计算距离。结构光方案根据目标场景表面的特殊图案形变来判断目标场景的距离信息，一次测量可以覆盖非常大的面积，可以实现非常快速的进行扫描，在动态测量场景具有得天独厚的优势；但由于特殊图案会受到环境光（日照）的影响，限制的结构光方案的适用范围，精度也较三角测量方案低。调变光方案则是向目标场景投射周期性变化的图案，如光强正弦变化的图案，通过计算每个像素亮度的变化与光的相差，可以达到极高的精度，然而这种方案是所有主动扫描深度相机中对噪声和干扰最敏感的设计。

主流的非接触式被动扫描有立体视觉（Stereoscopic）、立体光学（Photometric Stereo）等方案。我们主要对使用较为广泛的立体视觉方案进行分析。

立体视觉方案基于三角测量原理，通过配准立体图像计算视差来实现距离的计算；具有结构简单、成像精度较高、扫描速度较快、分辨率高等优势，但仍具有计算量大、受光照条件和目标场景表面纹理影响较大等缺点。本文从立体视觉方案出发，分析总结该方案的优缺点，提出系统实现方案并搭建测试平台验证系统实现方案的可行性。

1.3 本文的主要研究内容和创新之处

本文我们主要研究双目立体视觉方案（以下简称双目方案）的实现原理，分析双目方案优缺点，探讨双目方案可能存在的改进方向，我们重点对下述问题展开研究：

（1）针对如何减少双目方案计算量、加速处理速度的技术研究。传统的双目方案计算视差图需要对图像进行多次迭代计算，计算过程并行化程度低，浪费了计算机多核资源；本文提出通过限制配准范围，减少参与计算的图像并设计新的并行化配准算法，极大的加速每帧图像的处理速度。

（2）针对如何减少双目方案误差的技术研究。传统的双目方案对图像进行配准计算视差的算法大多属于全局优化类算法，这种算法对于图像噪声、环境遮挡、光照变化敏感，容易在视差图中造成大范围空洞区域；本文提出的系统实现方案结合三角测量方案的优势，通过标记模块主动向场景提供明显标志，利用三角测量的方法计算视差，可以明显地减少错误的配准。

（3）针对如何减少双目方案对目标场景光照条件和表面纹理敏感的技术研究。本文提出的系统实现方案在对图像进行处理的过程不需要考虑目标场景的光照条件和表面纹理，预处理算法只需要找出出现在目标场景明显的标记图案，匹配两幅图像中标记的位置即可计算视差，有效减少了图像噪声、目标光照条件变化、目标场景视线遮挡、纹理缺失的目标场景无法重建等问题。

在文章末尾我们设计和搭建了该系统实现方案的实验装置，并验证了该方案的可行性；执行测试过程、统计实验数据，通过对实验数据的分析和计算，总结并纠正每个环节出现的问题，分析该系统实现方案的优缺点和潜力，并针对该方案的缺陷提出了新的深度相机实现方案。

1.4 本文的结构与内容安排

本文主要研究了如何优化现有的深度相机实现方案。由双目方案出发，通过研究分析双目方案和三角测量方案的实现原理，结合两种方案的设计，提出改进的算法和设计方案。本文主要分为四章节进行讨论，各章节的内容如下：

第一章介绍了深度相机的功能以及它在各种领域的应用，分别对主流深度相机的实现方案进行介绍和优劣势分析，让我们对各类深度相机实现方案的原理有一个大概的认知；针对双目方案的缺陷提出研究的重点，最后介绍了本文的内容结构安排。

第二章介绍了机器视觉的相关技术并对双目深度相机实现方案的基本原理进行探讨，让我们对相机的几何模型，如何通过视差图计算目标场景的三维结构，双视图匹配算法如何得到视差图的过程有一个大致的了解，并在本章节结束展示了传统双目方案得到的结果。

第三章让我们对本文提出的系统实现方案的改进点和实验装置所依赖的基础技术原理有较为详细的了解。介绍了系统实现方案所依赖的算法的技术基础，包括该方案的技术模型、条带状图样中线提取算法和标记模块死区估计算法的设计等内容，为第四章设计实验装置打下坚实的理论基础。

第四章进行实验环境的介绍和构建；根据本系统实现方案的基本原理设计实验平台，介绍搭建实验平台所使用的所有元件设备，搭建平台和设计测试方案。

第五章我们使用实验装置根据测试方案的设计进行测试并统计实验数据，根据实验结果讨论和分析各个环节出现的问题，展示修正问题后各个测试方案的实验结果并进行总结。

第二章 开发环境和理论原理概述

2.1 开发环境

2.1.1 OpenCV

OpenCV 全称为 Open Source Computer Vision Library（开源计算机视觉库），主要接口由 C++ 语言编写并针对硬件进行了优化，以 BSD 协议发行，可以在研究领域免费试用；OpenCV 为我们实现的图像处理大多数的基础功能，极大的简化了开发图像处理软件的难度。

在本项目中我们主要使用 OpenCV 中图像处理和计算机视觉相关的模块，实现对摄像头的驱动、图像数据的存储和转移、相机标定等功能。

2.1.2 OpenMP

OpenMP 全称为 Open Multi-Processing（开放多线程处理），是一个可移植、可扩展的编程模型，OpenMP 用于共享内存并行系统的多线程程序设计的一套指导性注释，为程序员提供了一个简单而又灵活的开发平台，支持 C++ 语言和 Windows 平台。

在本项目中我们使用 OpenMP 来并发执行条带化图样中线提取算法，提高硬件利用效率，显著提升软件算法的执行速度。

2.1.3 SystemVerilog

SystemVerilog 是一种由 Verilog 发展、改善而来的硬件描述、硬件验证统一语言，主要在语言中加入了功能验证特性，很好地弥补了传统 Verilog 在芯片验证领域的缺陷，改善了代码可重用性，同时可以让验证工程师在比寄存器传输级更高的抽象级别，以事务而非单个信号作为监测对象，这些都大大提高了验证平台搭建的效率。

2.1.4 FPGA

FPGA 全称为 Field-Programmable Gate Array（现场可编程逻辑阵列），是专用集成电路领域的一种可半定制的电路系统，既解决了全定制电路灵活性的不足，又克服了原有可编程逻辑器件门电路数有限的缺点。

在本项目中我们使用 FPGA 构建标记模块的控制器，通过 Verilog HDL 语言构造硬件电路和使用 NIOS II 软核，实现对标记模块中触发元件输出信号的硬件滤波以及软件估计，计算旋转反光镜位置并向上位机传递死区范围的数据。

2.2 相机的透视投影模型

相机作为计算机了解现实世界重要的端口设备，主要的职能是捕获目标范围内场景的光学特征与二维物理结构并将其量化为计算机可以处理的数字信号，为计算机视觉、机器视觉以及更高层的软件应用提供重要的数据支持。针孔相机模型（以下简称相机模型）是在现实中使用最广泛、简单和通用的相机几何模型^[4]，该模型定义了目标场景范围内三维空间中的位置与其在相机的成像平面（指相机感光元件的感光平面）上对应的投影之间的几何关系（数学上称这种关系为透视投影）。使得现实中的相机具有通用、准确的几何模型与稳固的数学基础，为我们利用相机的几何模型的特性，通过计算恢复场景的三维结构提供了可靠的理论依据。

如图 2.1 所示的相机模型中，我们将从目标场景的表面出发、投向相机成像平面的所有光线相交的点（透视投影的中心点）称为该相机模型的光学中心或者相机中心；垂直于相机的成像平面并穿过相机中心的虚拟轴线称为主光轴；主光轴与相机成像平面的交点称为主点。

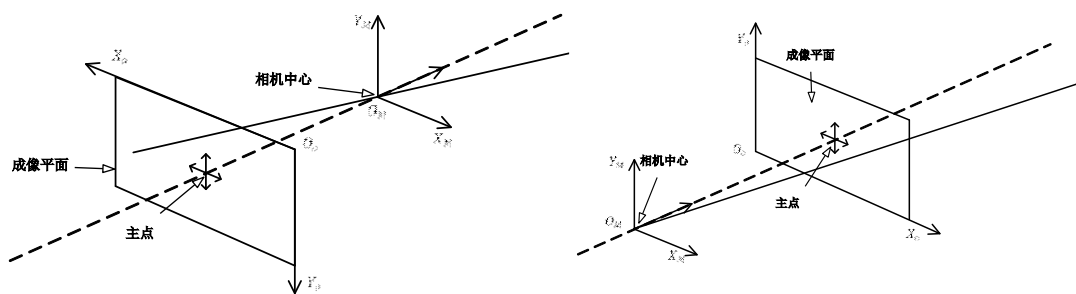


图 2.1 针孔相机模型

a.左图为实际相机模型，为了便于理解以及简化计算，我们将感光平面沿主光轴移动到相机中心的前面，如右图所示，像素坐标系翻转。

实际上现实中的相机大多采用如图 2.2 所示的光学透镜以增大进光量，提升相机感光元件的信噪比并减少如暗电流、宇宙辐射等随机噪声^[5]对相机图像质量的影响；但透镜的使用导致相机模型被引入了非线性变换，相机捕获到的图像会产生非线性失真；图像的失真造成理论上的相机几何模型与实际相机物理结构出现偏差，影响软件算法计算恢复目标场景三维结构的精度。对相机的几何模型进行标定并对图像进行修正可以减少这种现象带来的影响，在理论研究阶段我们讨论的相机模型均为理想模型，是指假设所有的相机均进行过相机标定使得相机的几何模型与物理模型贴近，不深入研究讨论相机标定的理论原理以及如何实现相机标定。

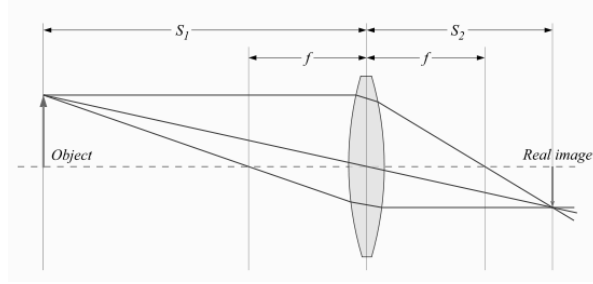


图 2.2 使用透镜后的相机光路图

如图 2.2 所示，以光心为中心，朝向成像平面的视角观察成像平面上的感光像素阵列，我们人为地规定像素阵列中左上角的像素为像素坐标系原点，以穿过像素坐标系原点并且与阵列较长的边平行的直线为 X 轴建立像素坐标系。

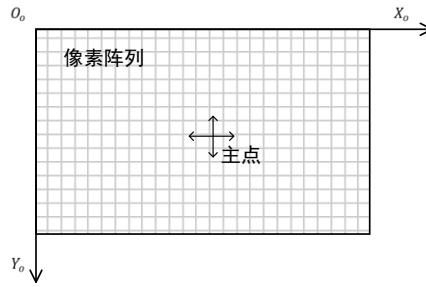


图 2.3 像素坐标系

如图 2.3 所示，以相机的中心作为计算机模型坐标系（以下简称模型坐标系）的原点，将穿过光心并平行于主光轴的直线作为模型坐标系的 Z 轴，主点向光心的方向作为模型坐标系 Z 轴的正方向；穿过模型坐标系原点，与成像平面较长的边平行的直线作为模型坐标系的 X 轴，主点向像素坐标系负方向为模型坐标系 X 轴的正方向，以此建立左手坐标系（见图 2.1 右）。相机主点与光心的距离为 f ；我们引入齐次坐标实现像素坐标系下的投影点到模型坐标系下的目标场景三维结构坐标的转换，得到模型坐标系中的任意位置 $(X, Y, Z)^T$ 与像素坐标系中的位置 $(u, v)^T$ 的投影方程：

$$\varphi \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} f & 0 & O_x & 0 \\ 0 & f & O_y & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad \text{公式 (2.1)}$$

式中 φ 作为齐次坐标比例因子，数值上等于 z 。由公式 (2.1) 我们可以得出，投影面（成像平面）上的任何一点的对应的投影坐标是无穷多的，仅凭一个视点的单张图像是无法还原出目标场景的三维结构。

2.3 双视图几何模型

为了解决单一视角的单张视图无法还原目标场景三维结构的问题，我们尝试使用两个视点不同的相机同时捕获目标场景来解决这个问题。视点不同的相机在捕获相同的目标场景的过程中，目标场景中的二维结构在两个相机得到的图像中出现的位置会出现差异；利用这个原理我们可以得到该场景的视差图，根据视差图以及相机的几何模型进行三角测量，实现对目标场景的三维结构的恢复。

视点不同的两个相机需要有一致的几何模型来描述这两个相机之间的物理关系，我们称这种几何模型为双视图模型^[6]。双视图模型的参数与每个单独的相机的几何模型参数、相机中心的相对位置、相机主光轴的方向都有着密切的联系。使用双视图相机标定我们可以较为准确的估计双视图模型的参数。

目标场景中的任意一点在两个视点不同的相机的成像平面上的投影点具有对应关系，描述这种对应关系的数学方法与工具被称作极线几何^[6]。极线几何将两个不同视点的相机的几何模型相结合并建立约束。极线几何主要解决了如下两个方面的问题：

一是点对应几何：如图 2.4 所示，给定视图中的任意点 P_l ，极线几何将对该点 P_r 在另一视图中出现的位置进行约束，我们称之为极线约束；目标表面的点 P 、投影点 P_l 、投影点 P_r 和两相机的中心五点共面，该面被称作极线平面；两相机的中心的连线被称为双视图模型的基线；极线平面与两相机的成像平面相交的交线被称作对极线。极线约束表明目标场景的任意一点 P 在两相机的投影点必定位于同一条对极线上，极线约束原理大大减少了双视图几何还原三维信息算法的设计难度与计算复杂度。

二是场景几何：给定投影点的对应和两个相机的极线几何，可以通过对这两个视点不同的图像的处理恢复目标场景的三维结构。也就意味着我们可以通过投影点的对应计算出视差图，再根据视差图和极线几何原理恢复目标场景的三维结构。

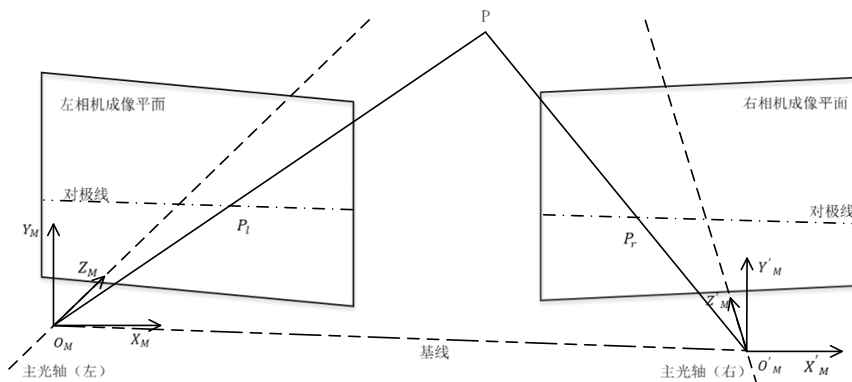


图 2.4 极线约束的示意图

如图 2.4 所示，由于装配误差的存在，双视图系统几乎不可能使双视图模型的基线与相机的成像平面完美平行，故对极线是很难与相机的成像平面上的像素阵列平行；不平行的对极线让图像上投影点的对应不在同一像素阵列行，使得软件算法增加了无谓的计算量。为了解决这个问题，我们将不同视点的相机捕获到的没有对齐和约束的视图进行转换合成（视图校正），视图校正将视点不同的相机捕获到的图像分别进行虚拟的旋转和移动，使极线与像素阵列平行。软件算法处理对齐后的图像仅仅只需要沿水平方向的阵列进行搜索匹配，提高匹配的准确度并降低了软件算法的时间复杂度。

经过标定后的双视图几何模型可以很精确的描述视点不同的相机的相对位置关系。如图 2.5 所示，我们取标定后的双视图几何模型的某一个极线平面来进行三维结构恢复算法的研究。经过校正的几何模型极线与基线平行；规定处在极线左侧的相机为主相机，称主相机的像素坐标系为主像素坐标系、主相机的模型坐标系为主模型坐标系。

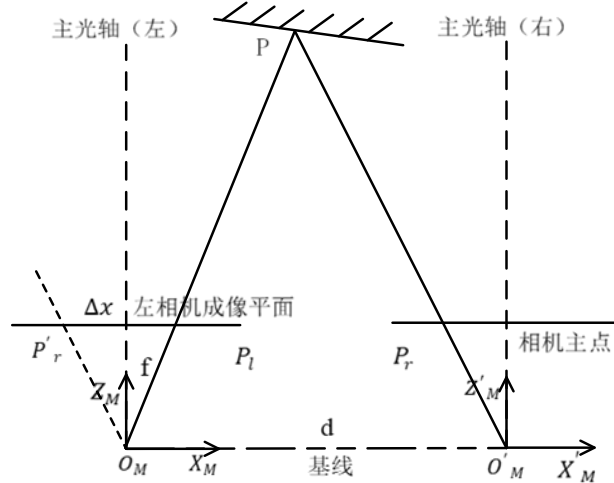


图 2.5 某一极线平面的图

主像素坐标系下主点坐标为 (u_0, v_0) ，像素阵列上点 P_l 的坐标为 (u, v) ， Δx 为该点处的视差（主相机图像上的某一点 P_l 与其在副相机对应的点 P_r 横坐标的差值），双视图模型的基线长度为 d ；我们结合三角测量原理和公式（2.1）进行推导得到公式（2.2）。根据公式（2.2）我们可以根据视差计算出主坐标系下 P_l 点对应的三维点 P 的坐标 $(X, Y, Z)^T$ 。

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} u - u_0 \\ v - v_0 \\ f \end{pmatrix} \begin{pmatrix} \Delta x \\ d \end{pmatrix} \quad \text{公式 (2.2)}$$

公式（2.2）向我们展示了影响恢复三维结构精度的两个因素：一是双视图几何模型对现实相机结构的还原是否准确（焦距 f 、基线长度 d 与物理模型是否吻合）；二是视差计算的准确度（对于主相机图像中的某一点能否正确的找出其在

副相机图像中的对应点)。

我们根据参与恢复目标场景三维结构计算过程的视差点对的多少将计算视差的算法大致分为两类：一类是视图中大部分的像素构成了视差点对，参与计算并恢复出目标场景较多的三维结构，如 SAD(Sum of absolute differences)算法、SGBM(Stereo Processing by Semiglobal Matching and Mutual Information)算法等，恢复出的三维结构构成了稠密的点云(如图 2.6)；另一类是使用特征点作为构成视差点对的匹配标准，如 ORB 算法；在图像中分别提取特征点，再将特征点进行匹配并计算视差，计算恢复三维结构得到较为稀疏的点云。

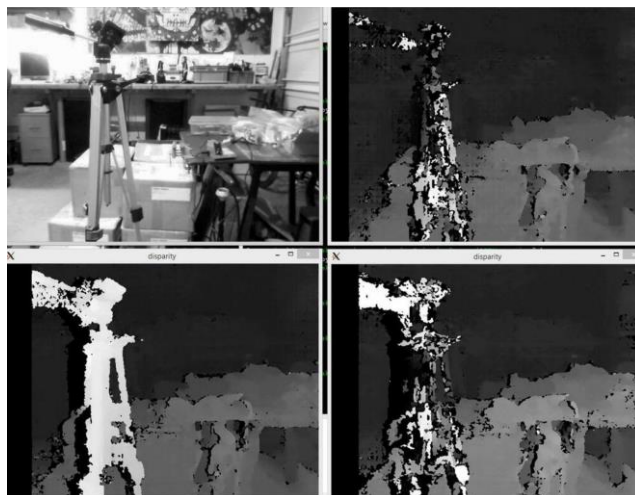


图 2.6 双目方案计算得到的视差图

计算视差的算法均需要图像具有明显的特征，这意味着在光照条件不足、特征不明显的场景以及重复纹理等目标场景中，计算视差的算法的准确度会受到极大的影响。为了减少特殊的场景对三维信息恢复精度的影响并降低计算视差算法的时间复杂度，我们将传统的双视图模型进行改进，在双视图模型中引入线结构激光主动对场景进行周期性的扫描，相机捕获激光照射在场景表面产生的明亮的条纹，软件算法对不同视图中出现的明亮条纹进行配对生成视差点对，恢复明亮条纹处的三维结构。这种方案减少了传统视差计算算法对场景的依赖，通过约束配对范围、增强特征的方法降低了算法的时间复杂度并且提高了系统的精度。

第三章 本系统实现方案的原理概述

3.1 本系统实现方案的系统模型

针对被动双视图模型的固有缺陷，我们引入线结构激光主动在场景中提供明显的标识（明亮的条纹）。如图 3.1 所示，反射镜搭载于旋转平台，将线结构激光束反射向目标场景以实现场景周期性标记扫描。

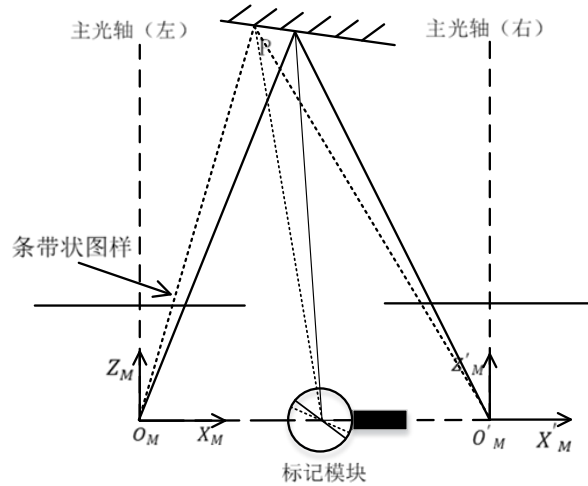


图 3.1 改进后的双视图几何模型

反光镜所在的旋转平台、线结构激光发生器和各种控制器等元件组成的模块（以下简称标记模块）可以设置在任意位置（原则上不遮挡相机对目标场景的捕获即可，图 3.1 中标记模块设置在基线中点附近的位置）；标记模块的位置并不会对视差的计算造成较大的影响，标记模块也不会影响双视图模型原本的几何约束，故也不会参与三维信息的恢复计算；标记模块仅负责对目标场景投射标记并使标记在目标场景中处于不断移动的扫描状态；这个过程会非常的迅速，导致在相机捕获目标场景的期间标记就移动（线结构激光的反射镜转动了一定的角度）了一段距离（如图 3.1 所示），导致相机捕获到的视图中的标记移动了一定距离并形成了条带状图样。

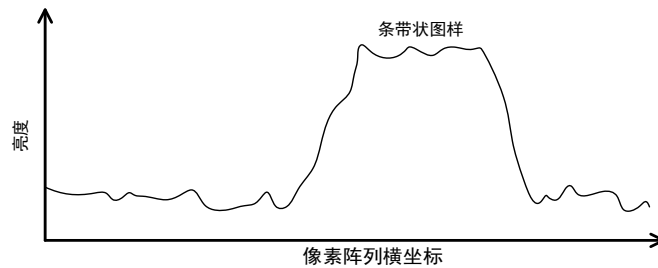


图 3.2 条带状图样的亮度分布

在目标场景表面材料的反射、散射、光源移动和线结构激光光束发散不均匀

等多重因素的影响下，相机捕获到的条带状图样具有边缘模糊、随着与目标场景的距离变化图样的宽度发生变化、图样各处亮度不平均等特点（图 3.2 向我们展示了条带状图样理论上的亮度分布图，由于线性激光镜头的缘故，像素阵列每行的条带状图样亮度峰值都有差异），为了提高软件算法计算视差的精度，我们需要设计算法处理条带状图样，提取出条带状图样的中线作为视差点对集。

3.2 条带状图样中线提取算法

为了实现对条带状图样进行预处理，我们设计了新的算法用于提取条带状图样的中线。条带状图样中线提取算法（以下简称中线提取算法）的设计思路借鉴 Mean Shift（以下简称均值漂移）算法并根据条带状图样的特点对其进行扩展，根据极线约束原理，我们将将中线提取算法的计算空间限制在一维并去除极大值区域漂移的过程，以此降低算法的时间复杂度。

给定 D 维空间中点集 $P \neq \emptyset$ ，对于点集中 P 中任意一点 X 来说该点上的漂移向量的基本形式为：

$$M_X = \frac{1}{K} \sum_{x_i \in R} (x_i - X) \quad \text{公式 (3.1)}$$

公式 (3.1) 中 M_X 是均值漂移算法求得的 D 维向量（漂移向量），指向在点 X 周围，点集 P 密度最大的区域； R 是一个非空点集，表示是以任一点 X 为球心，半径为 d 的一个范围（ $\|x_i - X\| < d$ ）内的所有点的集合，在该范围内的点参与漂移向量的累加计算； K 表示在点集 R 内的点的个数。

由公式 (3.1) 我们可以得出这样的结论：在区域 R 内的所有点对漂移向量 M_X 的贡献-权重是相等的，影响漂移向量的变化量只有点集中的点与中心点的距离和某个方向点的个数（邻域内点的密度）。由于相机捕获到的场景图像具有点集位置矩阵化、各点具有权重差异并且离散化（图像像素的亮度差异）等特点，我们需要将原本均值漂移算法中点与点之间的欧式距离提供对漂移向量贡献修改为点与点的亮度差异提供对漂移向量贡献。

条带状图样具有边缘模糊的特征；图像边缘是局部特征，图像边缘强度和该边缘邻域亮度空间分布、亮度极差有关。公式 (3.1) 所示，邻域内所有的点权重是相等的，但权重相等的平滑运算会使得离算子中心很远的像素点仍然有较大的贡献量，导致本是局部特征的图像边缘失真。为了解决这个问题，我们在中心点领域设置权重掩模，邻域内的点对漂移向量的贡献随着该点与中心点的距离递减。由于线结构激光的光源是高斯光源，条带状图样的边缘亮度的分布同样服从高斯分布，故我们在算法中使用径向基函数核（以下简称高斯核）构造权重掩模。

极线约束校正后的图像将可能匹配点对校正正在同一行（纵坐标相等），我们

将算法限定在一维（限制纵坐标相等，不考虑其他行），在图像中的每一行单独计算漂移向量。

综合以上因素，我们可以将均值漂移算法的基本形式（公式（3.1））变形修正：对于以点 X 为中心，与点 X 的欧式距离小于等于 d 范围内所有的像素集合 Q ，中心点 X 处的漂移向量 M_X 表示为：

$$M_X = \frac{1}{K} \sum_{x_i \in Q} G(\|x_i - X\|) (B_{x_i} - B_X) \quad \text{公式 (3.2)}$$

其中，中心点像素的亮度 B_X ， K 表示该点集 Q 中元素数量。

$$G(z) = G(\|x_i - X\|) = \exp\left(-\frac{\|x_i - X\|_1^2}{2\sigma^2}\right) \quad \text{公式 (3.3)}$$

但公式（3.2）和公式（3.3）构造的算法计算得到的只是关于条带状图样边缘强度分布的信息，我们需要求边缘强度的一阶导数以获得关于条带状图样边缘强度的方向分布，我们对公式（3.2）求一阶导数得公式（3.4）：

$$M'_{x_i} = \sum_{x_i \in Q} G'(x_i - X) (B_{x_i} - B_X) \quad \text{公式 (3.4)}$$

$$G'(x_i)_x = -\frac{1}{2\pi\sigma^4} (x_{ix} - X_x) \exp\left(-\frac{\|x_i - X\|_1^2}{2\sigma^2}\right) \quad \text{公式 (3.5)}$$

结合公式（3.4）和公式（3.5）构造滤波器（高通滤波器），在固定宽度的窗口进行滤波，以每个像素作为窗口中心对相机捕获到的图像进行预处理。预处理阶段我们得到了一个关于条带状图样位置的估计与边缘强度分布的、离散的一维离散向量场（以下简称向量场），这个向量场具有方向性（总指向领域中亮度最大的区域）、边缘性（漂移向量在条带状图样的边缘取得较大值，在条带状图样的内部或者没有条带出现较暗的区域取得较小值）和离散型（由于像素阵列本身就是离散分布的）等特性。

由于该向量场具有边缘性，我们将向量场中每个离散点处的向量根据其模长进行排序，并筛选出序列前 N 向量对应的离散点构成集合 H 。集合 H 中的任意一点 H_i 、 H_k 在向量场中与其对应的向量为 P_i 、 P_k ，将向量夹角为 π 的两个点将构成点对参与中点计算。权重均一化比例因子 Usf 将点对权重的线性归一化，使向量场的边缘向量构成的点对。

$$Usf = \sum_{P_i \cdot P_k = 0} (\|P_i\| \cdot \|P_k\|) \quad \text{公式 (3.6)}$$

$$Mid = \sum_{P_i \cdot P_k = 0} \frac{(H_i + H_k) \cdot (\|P_i\| \cdot \|P_k\|)}{2 \cdot Usf} \quad \text{公式 (3.7)}$$

根据公式 (3.6) 计算出权重均一化比例因子, 再使用公式 (3.7) 计算满足对应向量的夹角为 π 的点对连线的中点坐标乘以点对权重乘积线性均一化后求和, 计算得到集合 H 的中点。

在实际程序的编写设计中我们会引入一个参数限制点对连线的长度并限制点对的匹配方向, 减少图像噪点对条带状图样中线计算的干扰。

3.3 相差与场景扫描

使用条带状中线提取算法对相机捕获到的不同的场景视图分别进行处理, 得到条带状图样标记 (以下简称标记) 的位置集合, 我们规定在几何模型左侧的相机捕获场景视图并计算得到的标记集合为 G_L , 右侧相机捕获场景视图计算得到的标记集合为 G_R , 由于极线约束以及场景中标记唯一, 左侧相机标记对应的像差 K_{L_i} :

$$K_{L_i} = \|G_{L_i} - G_{R_i}\| \quad \text{公式 (3.8)}$$

根据公式 (3.8) 我们可以计算出标记处的相差, 接着我们可以通过双视图几何模型中推导出的公式 (2.2) 恢复条带状图样标记附近的三维结构。

通过控制相机的帧率 (曝光时间) 或者标记模块的标记周期, 使得相机捕获场景图像的间隔时间与标记模块中的反光镜的旋转周期不同步, 这种微小的时间差让条带状图样标记在场景中不断移动扫描; 通过控制标记在场景中的位置恢复局部三维结构, 不断地累积即可得到整个场景的三维结构。

如图 3.3 所示, 相机只能捕获特定角度范围内的场景, 我们称这个角度为视场 (FOV), 出现在两个重合的视场范围内的目标场景才能计算视差。由于标记模块采用了简单的平面反射镜, 使得场景内存在较长的时间区间不会出现条带状标识图样, 这个时间区间我们称之为扫描死区时间 (以下简称死区), 如果在死区期间相机对场景进行捕获, 软件算法无法在图像中计算出正确的条带状图样的位置, 进而导致软件算法无法恢复甚至恢复错误的目标场景三维结构。所以我们在标记模块中使用控制器对反光镜的实际位置进行估算, 并向软件提供大致的死区时间段参考以减少这种情况的发生。

我们使用如光电门或者霍尔传感器作为触发元件^[7], 当旋转平台通过某一位置 (角度) 时触发元件对标记模块的控制器发出一个脉冲信号 (或者只是反转了

电平)。高速旋转的旋转平台的每个回转周期内旋转平台的线速度我们可以认为是恒定的；控制器接收到信号并计算相邻的信号到达间隔并估算下一个回转周期可能的时间并计算出死区。

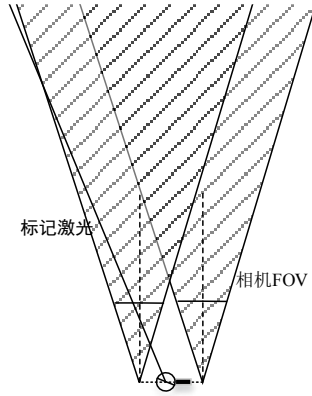


图 3.3 死区产生的原因

由于触发元件的模数转换电路存在较大的噪声干扰，导致控制器很难准确捕获到准确的触发脉冲的边沿。为了减少突发噪声对于回转周期估计的影响，我们采用了如下两种方案配合使用：一是使用 FPGA 搭建硬件滤波器对触发脉冲进行高频采样，使用多数投票策略决定触发脉冲边沿并减少亚稳态带来的影响；二是在控制器中使用软件或者硬件的方式实现卡尔曼滤波器^[8]，由于无刷电机驱动的旋转平台的重转周期恒定的，触发元件引入的噪声为高斯分布，触发脉冲是离散信号并且存在一定概率的随机的干扰发生，无控制离散型卡尔曼滤波器很适合对这样的系统进行估计。对触发脉冲边沿响应周期进行卡尔曼估计得出较为准确的重转周期，减少噪声的影响。

假设控制器对某一次触发脉冲时间间隔的测量值为 $M_{(i)}$ ，由于控制器主动保持反光镜的转速恒定，公式（3.9）中在该时刻的状态控制量 U_i 为 0，我们可以根据公式（3.9）对当前状态 $X_{(i|i-1)}$ 进行预测。

$$X_{(i|i-1)} = AX_{(i-1|i-1)} + BU_i \quad \text{公式（3.9）}$$

$P_{(i-1|i-1)}$ 代表上一时刻估算结果最优化后的协方差，由于标记模块为单一测量模型，公式（3.10）中 A 代表的系统参数；Q 为离散过程噪声的协方差，我们认为其不随时间而发生变化。

$$P_{(i|i-1)} = AP_{(i-1|i-1)}A^T + Q \quad \text{公式（3.10）}$$

完成协方差参数的更新后我们就可以通过公式（3.11）和公式（3.12）对系统当前最优值进行估计。将卡尔曼滤波器估计的最优值作为当前回转周期进行之后的死区计算。

$$Kg_{(i)} = \frac{P_{(i|j-1)} H^T}{HP_{(i|j-1)} + R} \quad \text{公式 (3.11)}$$

$$X_{(ij)} = X_{(ij-1)} + Kg_{(i)} (M_{(i)} - HX_{(ij-1)}) \quad \text{公式 (3.12)}$$

完成最优值估计后我们通过式公式 (3.13) 迭代更新均方差参数，使卡尔曼滤波器不断地运行。由于标记模块为单一测量的单模型，参数 $I = 1$ 。

$$P_{(ij)} = (I - Kg_{(i)} H) P_{(ij-1)} \quad \text{公式 (3.13)}$$

我们通过相机标定可以计算出相机的参数，从而计算出相机实际的视场；再根据双视图模型以及标记模块的位置与方向可以大致的估计出一个相对于触发元件的角度范围（死区）。由于电机高速旋转，我们可以认为电机在一个回转周期内的线速度是大致相等的；通过卡尔曼滤波器我们可以较为准确地估计到这一回合的回转周期，通过公式 (3.14) 我们可以计算出一个相对的延时区间，在该延时区间内曝光可以保证标记会出现在目标场景中。

$$\frac{\theta_{\text{offset}}}{2\pi} X_{(ij)} \leq T \leq \frac{\theta_{\text{offset}} + \theta_{\text{FOV}}}{2\pi} X_{(ij)} \quad \text{公式 (3.14)}$$

对于图 3.3 中的单反射镜标记模块，死区只有一个区间，公式 (3.14) 成立，当标记模块中的反射镜更换成双面反射镜甚至是棱角镜时，我们就需要对公式 (3.14) 进行修正，此处不再赘述。

第四章 本系统实现方案的实验装置设计

4.1 实验装置设计

4.1.1 实验装置的硬件设计

如图 4.1 所示的实验装置的硬件部分主要由两大部件构成：

一是目标场景捕获模块（相机模块）：该模块由两个全局快门相机构成，负责控制和配置图像传感元件的参数，接收到上位机发出的信号后捕获目标场景的图像，并通过 USB 接口向上位机传输数字图像数据。

二是标记模块：标记模块中的激光器为红外激光器，由于 808nm 红外激光不可见，大功率激光对人眼危害大^[9]，故我们采用 10mw 小功率激光器作为标记模块的发光源；标记模块使用无刷电机^[10]搭载反光镜片旋转向目标场景反射红外激光发生器发射的线结构激光，无刷电机由电子调速器（以下简称电调）控制调整转速；标记模块搭载霍尔元件用于记录无刷电机经过某一特定位置的时间，实现对反射线结构激光的角度的估计，减少由于目标场景捕获模块在死区时间内捕获场景图像，导致未能在场景中捕获到正确的条状标记图样，进而导致软件算法恢复错误的场景三维结构这种情况出现的概率。标记模块的控制器使用 FPGA 实现，在 FPGA 中构建硬件电路实现对霍尔元件发出的脉冲信号进行滤波，精确地计算触发周期；控制器也负责标记模块的电源控制，在暂时不需要运行的时间段内关闭标记模块以降低系统能耗。

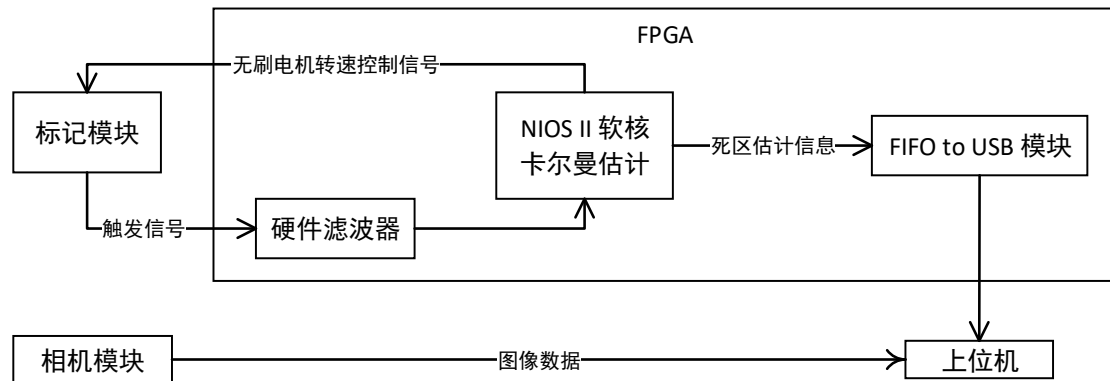


图 4.1 硬件结构

相机模块需要接受上位机的控制，在某个的时间点曝光固定的时长，然后传回捕获到的图像数据。由于曝光持续时间较短，负责标记的线性激光在标记模块的旋转反光镜的反射下导致条带状图样标记移动速度快（快速标记周期），为了避免普通相机感光元件的卷帘门快门的果冻效应，造成图像的变形，相机模块所

使用的图像传感器的曝光方式均需要支持全局曝光。

我们选用安森美半导体公司出品的型号是 MT9V034/D 的 CMOS 全局曝光灰度图像传感器芯片作为相机模块的感光元件（图 4.2），表 4.3 展示了该元件的关键参数。根据官方文档^[1]我们可以得知该芯片支持从模式，可以在外部控制下进行对目标场景的捕获以及如何配置该芯片的曝光时间长度和各种图像增益参数使得捕获到的图像质量最优。

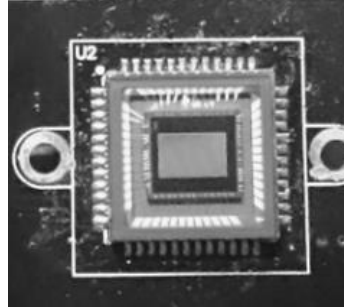


图 4.2 MT9V034 芯片图

表 4.3 MT9V034/D 芯片的关键参数

Parameter	Value
Optical Format	1/3-inch
Active Imager Size	4.51 mm (H) x 2.88 mm (V) 5.35 mm diagonal
Active Pixels	752 H x 480 V
Pixel Size	6.0 x 6.0 μm
Color Filter Array	Monochrome or color RGB Bayer
Shutter Type	Global Shutter
Maximum Data Rate Master Clock	27 Mp/s 27 MHz
Full Resolution	752 x 480
Frame Rate	60 fps (at full resolution)
ADC Resolution	10-bit column-parallel
Responsivity	4.8 V/lux-sec (550 nm)
Dynamic Range	>55 dB linear; >100 dB in HDR mode
Supply Voltage	3.3 V \pm 0.3 V (all supplies)
Power Consumption	<160 mW at maximum data rate (LVDS disabled); 120 μW standby power at 3.3 V
Operating Temperature	-30°C to + 70°C ambient
Packaging	48-pin CLCC

表 4.4 展示了配置寄存器的偏移量及其作用，我们将通过 Window Height 和 Window Width 两个寄存器将图像传感器传回的图像大小限制在宽度为 640 个像素、高度为 480 个像素的范围标准化图像尺寸。配置 Read Mode 寄存器使图像

传感元件进入从模式等待外部信号开始曝光。关闭自动分辨率控制和自动图像增益功能。

表 4.4 配置寄存器的偏移地址

Register Name	Register Number (Hex) For Context A	Register Number (Hex) for Context B
Column Start	0x01	0xC9
Row Start	0x02	0xCA
Window Height	0x03	0xCB
Window Width	0x04	0xCC
Horizontal Blanking	0x05	0xCD
Vertical Blanking	0x06	0xCE
Coarse Shutter Width 1	0x08	0xCF
Coarse Shutter Width 2	0x09	0xD0
Coarse Shutter Width Control	0x0A	0xD1
Coarse Shutter Width Total	0x0B	0xD2
Fine Shutter Width 1	0xD3	0xD6
Fine Shutter Width 2	0xD4	0xD7
Fine Shutter Width Total	0xD5	0xD8
Read Mode	0x0D [5:0]	0x0E [5:0]
High Dynamic Range enable	0x0F [0]	0x0F [8]
ADC Resolution Control	0x1C [1:0]	0x1C [9:8]
V1 Control – V4 Control	0x31 – 0x34	0x39 – 0x3C
Analog Gain Control	0x35	0x36
Row Noise Correction Control 1	0x70 [1:0]	0x70 [9:8]
Tiled Digital Gain	0x80 [3:0] – 0x98 [3:0]	0x80 [11:8] – 0x98 [11:8]
AEC/AGC Enable	0xAF [1:0]	0xAF [9:8]

图 4.5 为图像传感元件发出的关键时序信号，图像传感器传输的数字图像信号我们需要转换模块将 FIFO（First In First Out Buffer 先入先出缓冲，以下简称 FIFO）模式的数据通过 USB 接口发送给上位机，转换模块还需要接受从 USB 接口发送过来的控制信号去控制图像传感元件。我们选用市面上成熟的图像开发硬件模块，该模块提供的控制器可以实现我们上述的功能需求，控制器的信号连接如图 4.6。

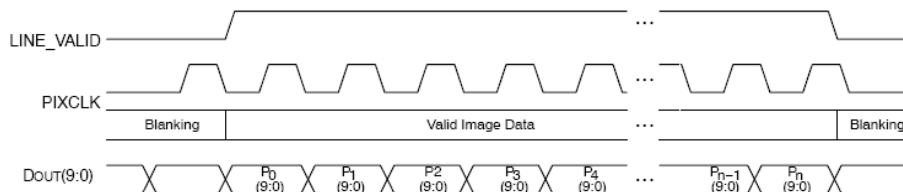


图 4.5 像素数据传输的时序信号

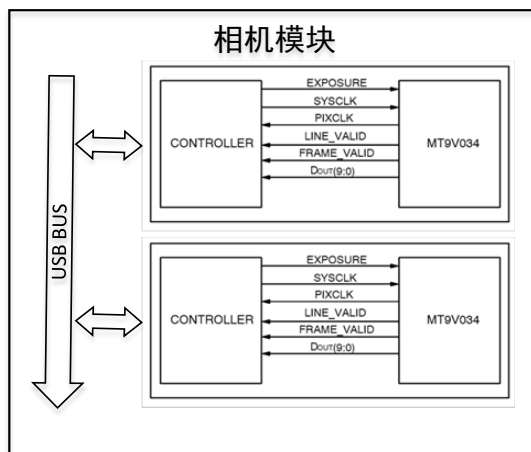


图 4.6 相机模块内部控制连接

标记模块由控制器、无刷电机驱动的旋转平台、触发元件以及其他控制器构成，我们选用市售的红外线性激光器作为标记模块的激光发生器（如图 4.7），该线性激光器的参数如表 4.8 所示，符合我们的需求。



图 4.7 红外线性激光器

表 4.8 红外线性激光器参数表

参数	值
电路模式	CW（连续波模式）
波长	808nm
镜头	玻璃/PMMA
输出功率	10mw
工作电压	2.8 - 3.2V
线宽	<300MM 处小于 2.5mm
激光等级	ClassIIIB
模组寿命	>8000 hrs

控制器作为标记模块的核心，需要由高速地完成滤波、标记周期估计等工作，考虑如果使用单片机作为控制器的实现，需要经常中断和滤波采集导致其他的任务执行缓慢和今后设备的扩展（将相机模块地控制器一起集成在标记模块内，通过硬件处理，不需要上位机参与计算）等因素，我们使用 FPGA 作为控制器的搭建平台（如图 4.9 所示），FPGA 芯片选用 Cyclone IV EP4CE6E22C8，足够满足我们目前的应用需求。滤波、通信等模块通过 FPGA 的硬件电路实现。在 FPGA

中搭建 NIOS II 软核运行标记周期估计算法，并通过通信模块将计算到的死区周期等数据传给上位机。

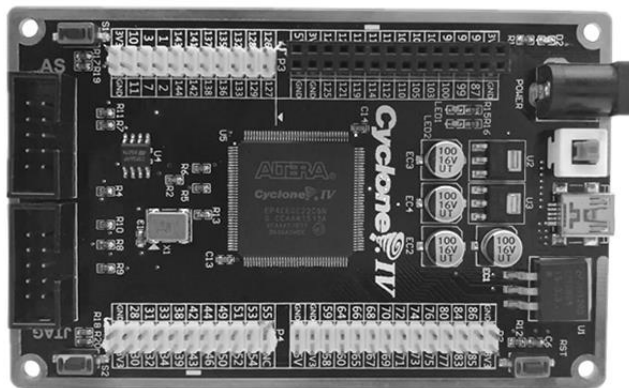


图 4.9 标记模块控制器主体

由于在 FPGA 中通过硬件电路搭建 USB 通信非常的复杂，我们使用已有的转换电路将 FIFO 通信转换为 USB 通信，实现控制器和上位机的通信。我们使用 FT245 芯片作为转换电路的核心，官方文档提供了该芯片外部连线的实例（如图 4.10 所示）。

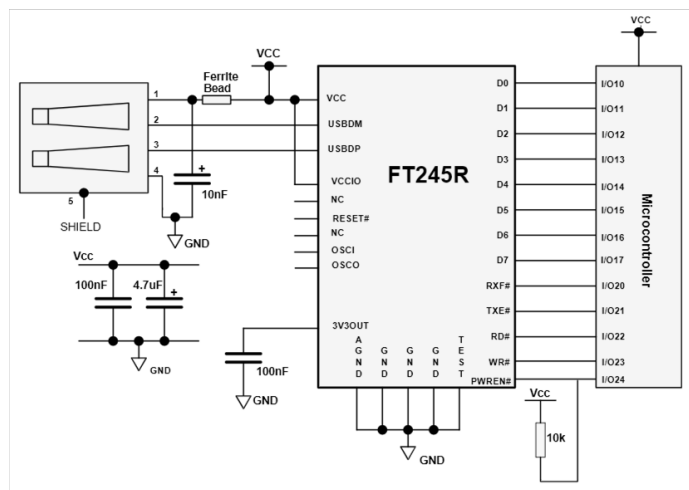


图 4.10 官方文档提供的实例电路图

通过图 4.10 我们可以很明显的看到 FT245 和控制器的通信接口由三态总线和异步握手信号实现。根据表 4.11 提供的接口信号定义和图 4.12 所示的信号时序图我们可以在 FPGA 中实现硬件驱动电路控制 FT245 的读写和实现 DMA 传输，减少 NIOS II 软核的读写压力。

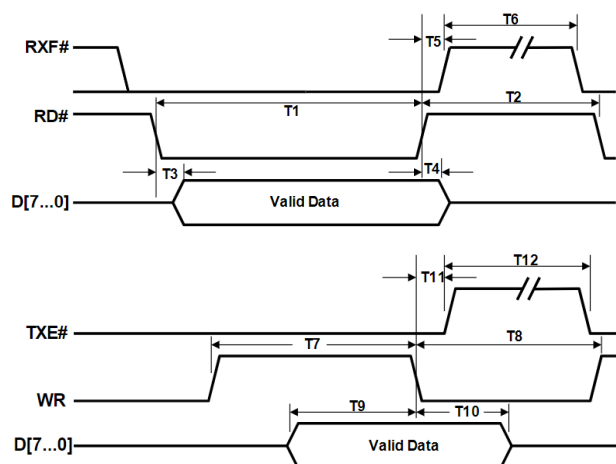


图 4.11 FT245 通信接口读过程握手信号（上图）和写过程握手信号（下图）

由于图像捕获模块的数据传输接口和 FT245 的数据传输接口类似，同为异步 FIFO，为了今后的扩展，我们将 FPGA 驱动电路设计模块化，提高模块代码复用率，如图 4.12 所示，FT245 模块的驱动模块，向 NIOS II 提供 DMA 通信方法，以便于今后传输图像数据。

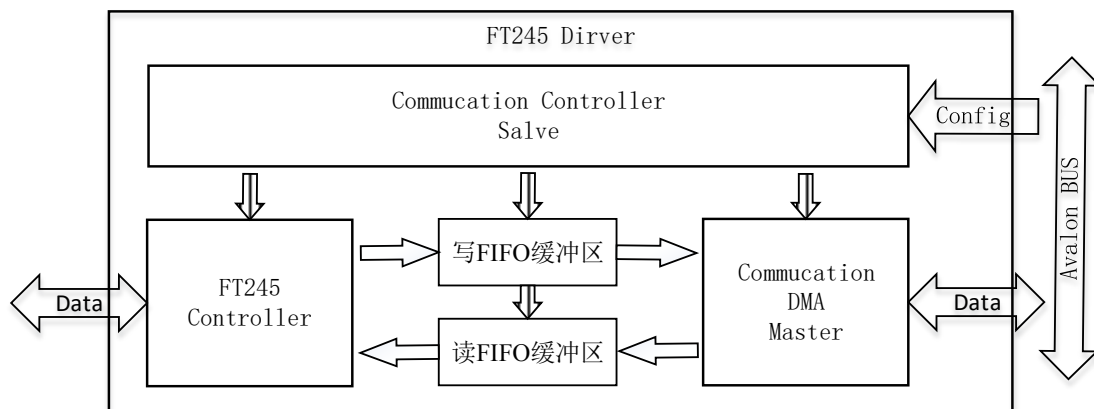


图 4.12 FPGA 中实现的 FT245 模块框图

FT245 驱动模块与 NIOS II 模块的通信通过 Avalon 总线^[12]进行，表 4.13 列出了我们使用的 Avalon 总线的信号表；Avalon 总线为我们将来扩展更多的模块提供了一个简单、高效的通信总线协议

表 4.13 Avalon 总线接口信号表

	信号名	宽度	注解
Avalon MM Master	Waitrequest	1	请求暂停信号，使能时设备不能对总线进行操作
	Address	1-32	发出读写目标地址
	Read	1	发出读请求信号
	Readdata	32	收到读取到的数据
	Readdatavalid	1	收到读取数据有效信号
	write	1	发出写请求信号
	writedata	1	发出写入的数据总线
	Burstcount	5	发出突发传输模式的突发长度
Avalon MM Slave	Waitrequest	1	发出暂停信号，暂停主设备的总线操作
	Address	1-32	收到读写目标地址
	Read	1	收到读请求信号
	Readdata	32	发出读的数据
	Write	1	收到写请求信号
	Writedata	32	收到写入数据
	Irq	1	向总线控制器发出中断请求

4.1.2 实验装置的软件设计

如图 4.14 所示，实验装置的软件部分主要由三个模块构成：

一是实现对场景捕获模块的相机进行标定，构建准确的双视图几何模型，为恢复目标场景的三维结构提供准确的模型支持。

二是软件需要在标记模块计算得到的合适捕获场景的时间区间内启动目标场景捕获模块，获得包含条状图样标识的图像并使用软件对两相机捕获到的不同视点的图像进行处理并计算出视差，最后根据双视图几何模型以及单目标定过程中计算出的焦距、主点位置等参数恢复场景标识附近的三维结构。

三是软件需要调整条状标记图样在场景中的位置，重复累积第二步计算得到的三维结构信息，实现对场景的扫描重建。

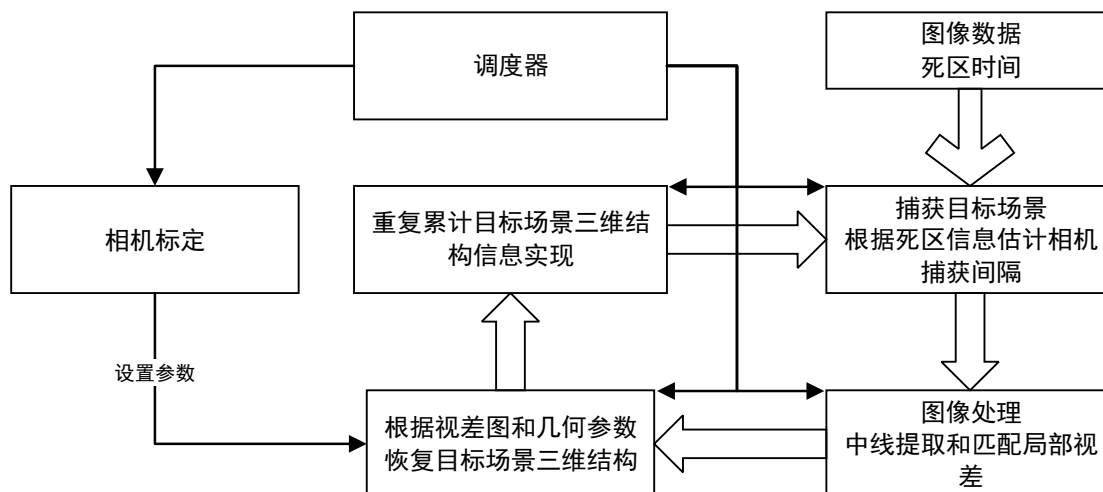


图 4.14 软件结构示意图

在系统初次运行或者相机模块结构有变动的情况，我们需要对相机模块进行标定工作。如图 4.15 所示的标定模块的工作流程，标定程序得到的重投影均方根误差在 0 - 1.0 范围内属于误差合理的情况，误差过高需要重新标定。

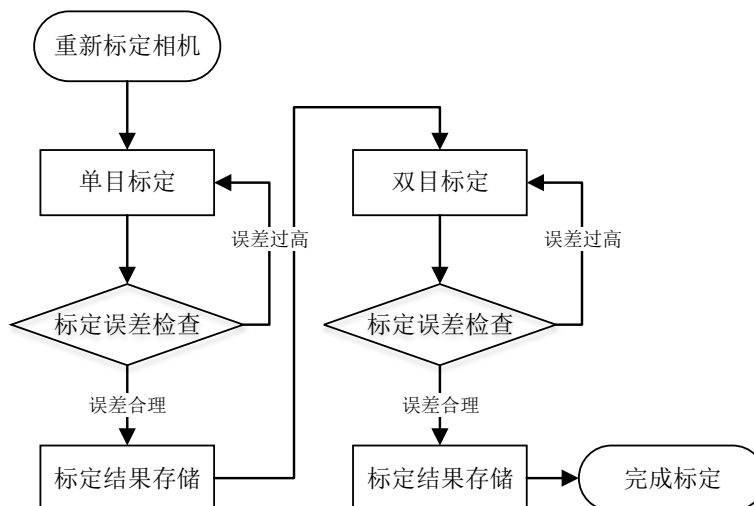


图 4.15 标定流程图

软件系统接收到相机模块传回的图像数据后首先进行校正，将左右两幅图像重投影，使其对极线平行，减少后续软件算法的计算量，为条带状图样中线提取算法的并行化奠定基础。

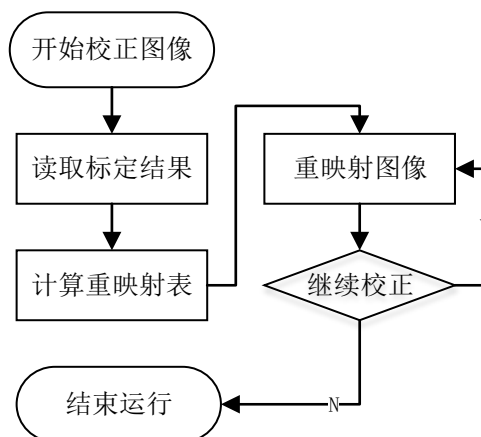


图 4.16 图像校正流程图

根据第三章设计的条带状图样中线提取算法的理论叙述，我们在程序中实现这样的算法流程，分别计算出左右图像条带状图样的中线，由于标定模块只是用一个反光镜，所以左右图像提取出的中线只有一条，极线约束原理使得我们只需要将纵坐标相同的中线点匹配为点对计算即可得到视差点对数据，将数据传给视差图累积模块进行处理。

如图 4.17 所示，视差图累积模块不断接受视差点对数据，累积到视差图中并加权平均降低视差图的误差，舍弃合理阈值外的视差点对，但对视图中的空洞不进行修补。

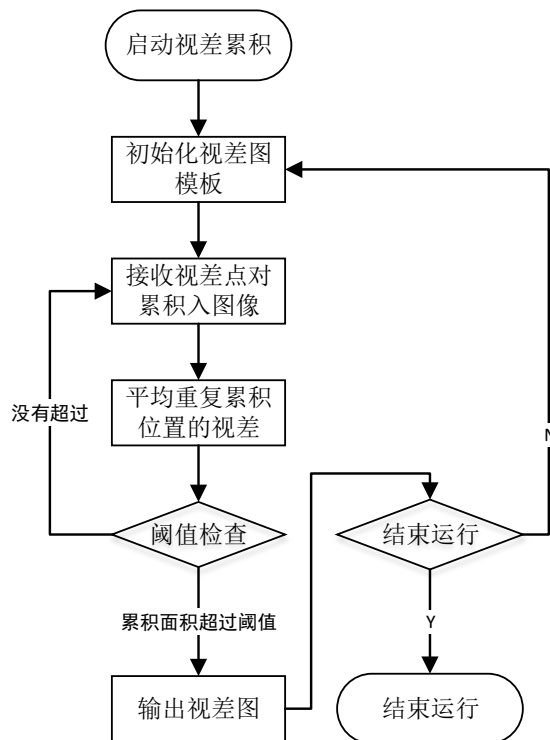


图 4.17 视差图累计流程图

当视差图像可恢复区域面积累积到超过阈值时，软件系统将会输出一幅视差

图像,这幅视差图是扫描期间目标场景的视差图像,如果目标场景有变化的情况,软件算法不能及时响应,该问题需要在后续方案修正设计中进行补充。

4.2 测试方案设计

4.2.1 标定测试

相机标定是求解双目方案的几何模型参数的过程,标定的结果是否准确直接关系到软件算法根据视差图恢复目标场景三维结构的精度。我们将通过一系列的测试验证标定模块给出的结果是否准确,具体过程如下:

标定准备: 我们将图 4.18 所示内角点为 7X5 的标定图样打印在 A4 纸上,并将打印好的标定图样粘贴在不易变形的平板上构成标定板。标定过程算法将会通过拍摄到的标定板搜索定位内角点的位置实现对相机几何参数的估计。

单目标定测试: 使用单个相机拍摄在场景中以不同角度和距离摆放的棋盘格标定板图像 13 张,确保图像中包含完整的标定板图样。角度之间的间隔大约 30° ,标定板和相机之间不易出现较大的角度,以免增加角点搜索的误差。输出标定后的参数和重投影均方根误差,通过观察校正后的图像是否准直,场景是否有明显变形验证校正的效果。

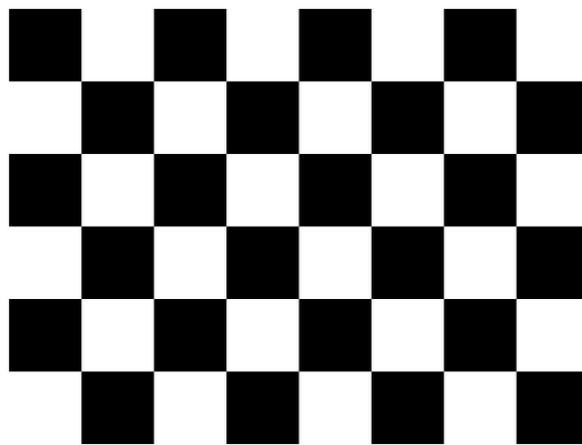


图 4.18 棋盘格标定板

双目标定测试: 完成单目标定测试后,我们分别得到的两个相机的内参和畸变参数。双目标定将同时使用两个相机拍摄标定板,计算内角点位置,软件算法通过两个相机的参数和拍摄到的标定板计算到的内角点位置计算双目几何模型的参数,输出标定后的参数和重投影均方根误差;查看校正后的两张图像是否配准,场景有无明显变形验证校正效果。

4.2.2 条带状图样中线提取算法测试

为了验证条带状图样中线提取算法的准确性,我们在较为黑暗的环境下进行

实验,尽可能的减少环境光对条带状图样的提取产生的干扰。将软件算法进入调试模式,输出条带状图样的标记

4.2.3 标记模块控制器测试

我们主要对标记模块的控制器进行测试,由于我们使用了卡尔曼滤波器的 IP 核减少我们的设计工作量,标记模块需要进行测试的模块现阶段只有 FT245 的驱动模块。

该驱动模块由 Verilog HDL 语言编写,使用 Modelsim 软件进行仿真。使用 System Verilog 搭建全自动仿真平台,使用 SVA (System Verilog Assertion) 描述设计期望行为的代码,构建系统断言基础属性,以保证调试能更加容易和快速地进行。通过随机化输入实现自动仿真平台的覆盖测试。

4.2.4 重建精度测试

重建精度测试方案主要是为了验证修正后的双视图模型相较于传统的双视图模型在精度上的提升。根据双视图几何模型的特性,我们决定在距离实验装置 1m~3m 的范围内设置测试平台,测试平台由位移度量平台和目标靶构成,目标靶的尺寸为 250mm X 250mm,目标靶选用不反光的材质减少场景表面材质对实验结果的影响。实验装置需要测量目标靶与相机中心的距离,实验装置的软件算法对目标靶的方向重复采集并限制了有效数据的窗口,减少可能存在的噪声和目标靶外的场景的干扰。

在测试范围内等距离选择五个基准位置作为五组测试数据的采集起点。每组测试数据的采集开始时首先将目标靶放置在该组的基准位置处,软件计算得到当前目标靶与相机的距离(测量值我们选用位于主相机视图中心处的窗口进行固定五点采样,取平均值作为传统双目方案的实验测量值),然后由基准位置开始向远离实验装置的方向平行移动一个标准移动位(20mm)并记录软件计算得到距离,再往靠近实验装置的方向移动同样的距离并记录,往复四次得到五个测量数据,计算测量数据差值作为测量位移数据记录。对标定靶重复进行上述测量方案完成五组原始测试数据的记录。

第五章 实验结果

5.1 实验平台展示

在第四章我们完成了实验装置的设计，根据设计搭建的实验装置的硬件结构如图 5.1 所示。完成实验平台的搭建后我们开始进行测试。

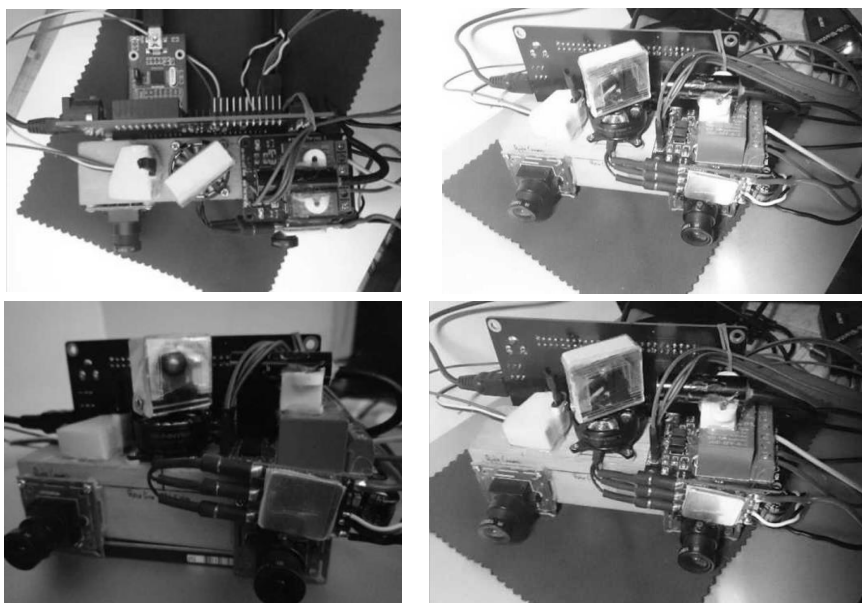


图 5.1 实验平台展示

5.2 测试结果展示

5.2.1 标定结果与分析

单目标定结果分析：

在实验装置的搭建完成后，我们需要对实验装置进行标定工作：首先我们对实验装置的每个相机单独进行单目标定，标定过程中使用需要标定的相机对摆放在场景中印有棋盘格图案的平面标定板捕获多张角度不同的图像，构成该相机的标定图像序列，再使用标定软件处理捕获到的图像序列，计算出相机的内部参数以及畸变模型的参数用于相机图像畸变的校正。图 5.2 我们展示了主相机（图左列）图像和副相机（图右列）图像的畸变校正结果。

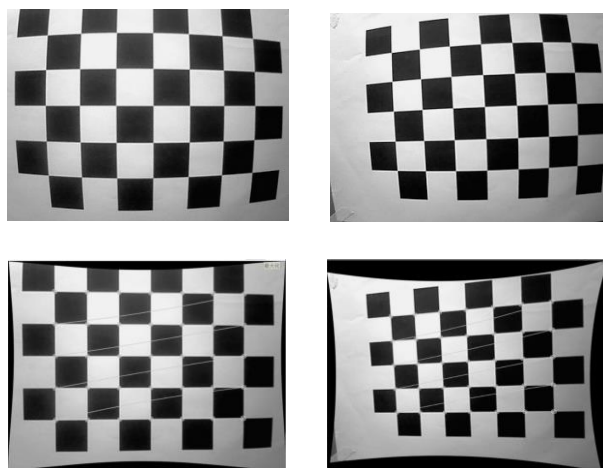


图 5.2 图像畸变校正结果展示

在标定的过程我们发现标定图像序列在修后仍存在明显的畸变的情况，标定算法返回的误差均方差值非常大。经过长时间的重复实验和在网络上查找相关资料，我们了解到标定过程中标定板不能与相机相距太远，尽量使标定板能覆盖整个相机视场，标定图像序列中的图像角度多样可以有效提高标定精度。

双目标定结果分析：

完成实验装置中所有相机的单目标定后，我们开始对图像捕获模块进行双目标定，确定双视图几何模型的各项参数。使用两个相机同时捕获多张摆放在场景中、角度不同的标定板图像（保证两个相机都能捕获到完整的标定板图像）构成双目标定的标定图像序列，再使用软件算法对双目标定图像序列进行处理，得到双视图几何模型的参数（基线位置与长度、相机旋转）信息。图 5.3 向我们展示了双目标定将图像旋转对齐后的结果。

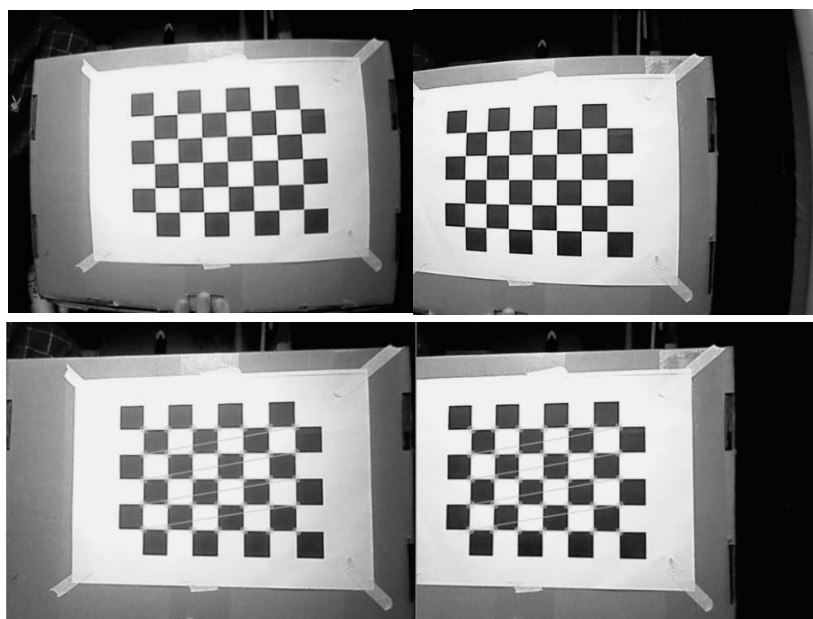


图 5.3 双目标定原图（上）和处理后的结果（下）

5.2.2 条带状图样中线提取算法结果与分析

为了减少目标场景干扰条带状图样的识别，我们在黑暗的环境进行条带状图样中线提取算法的测试，黑暗的环境使得相机捕获到的图像条带状图样部分亮度较高，便于我们进行测试。我们将软件算法进入调试状态，让软件程序输出条带状图样的中间处理结果便于我们分析。

图 5.4 为相机传回的图像数据通过软件 3D 化后转换成亮度高程图便于查看；通过图 5.4 我们可以明显的发现图像数据包含大量噪点，条带状图样由于激光发射器的透镜质量缘故导致光线亮度没有均匀的散布，实际的测试结果满足我们在三章对条带状图样的特点的亮度分布的预测。

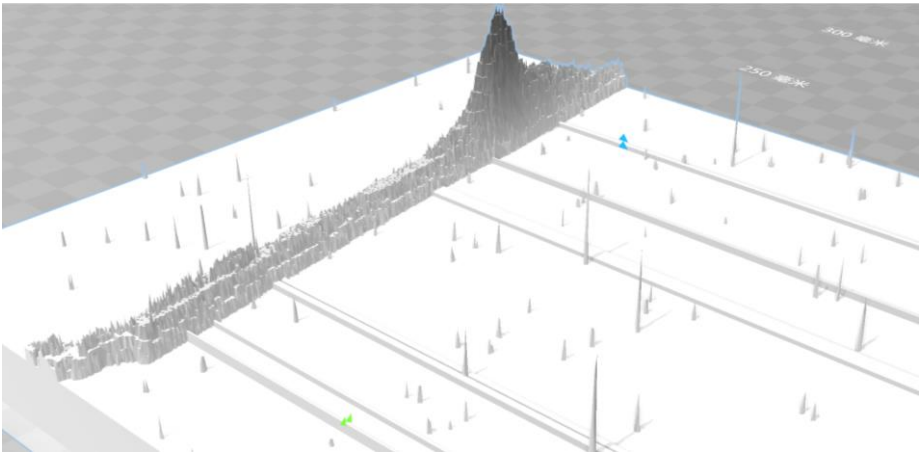


图 5.4 条带状图样亮度高程图

调试状态下我们让软件程序将计算得到的条带状图样的中线直接和输入图像融合输出便于我们查看，图 5.5 向我们展示了放大后条带状图样中线的位置。通过观察研究测试输出的数据（如图 5.6 所示），在没有对图像进行滤波的情况下，条带状图样中线提取算法较为准确的计算出了中线位置，并且对图像中出现的随机噪声具有一定的抑制效果。条带状图样中线提取算法的设计基本符合目标。

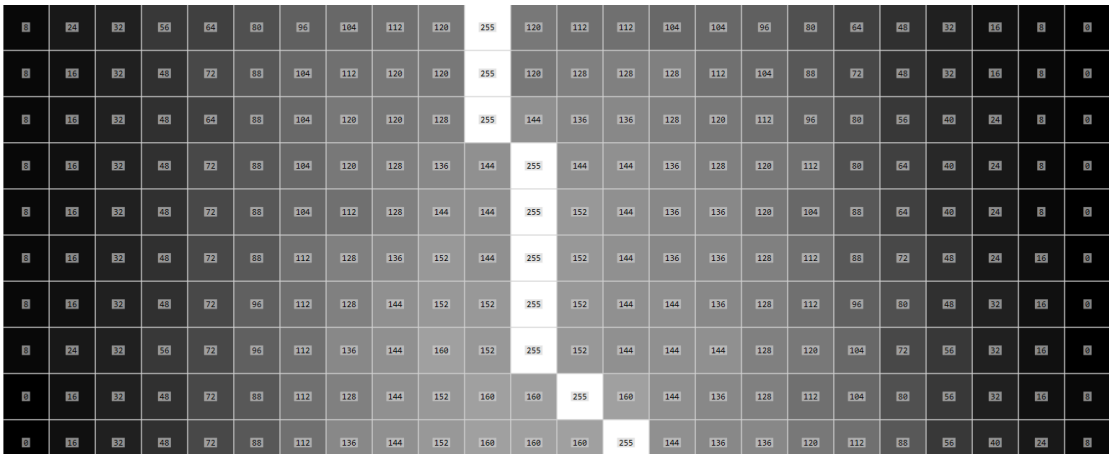


图 5.5 条带状图样和其中线的图像

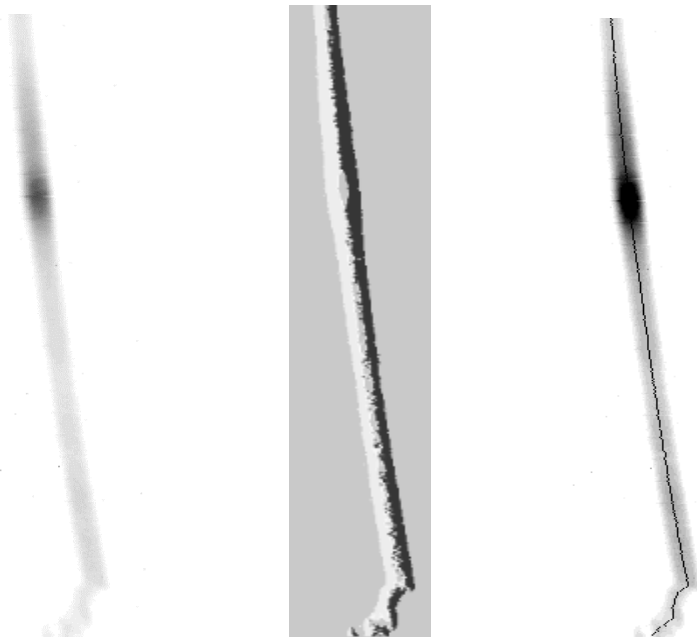


图 5.6 条带状图样原图（左）、概率密度计算结果（中）和中线位置（右）

5.2.3 标记模块控制器测试结果与分析

由于我们在自动化测试平台采用 SVA 建立了将要测试的模块的描述，任何在功能仿真过程中出现的不符合设计期望的电路行为都会导致断言失效。如图 5.7 所示，我们通过自动化测试实现 100% 的寄存器翻转和输入组合覆盖率并且不发生断言失效，验证了测试模块符合设计预期。

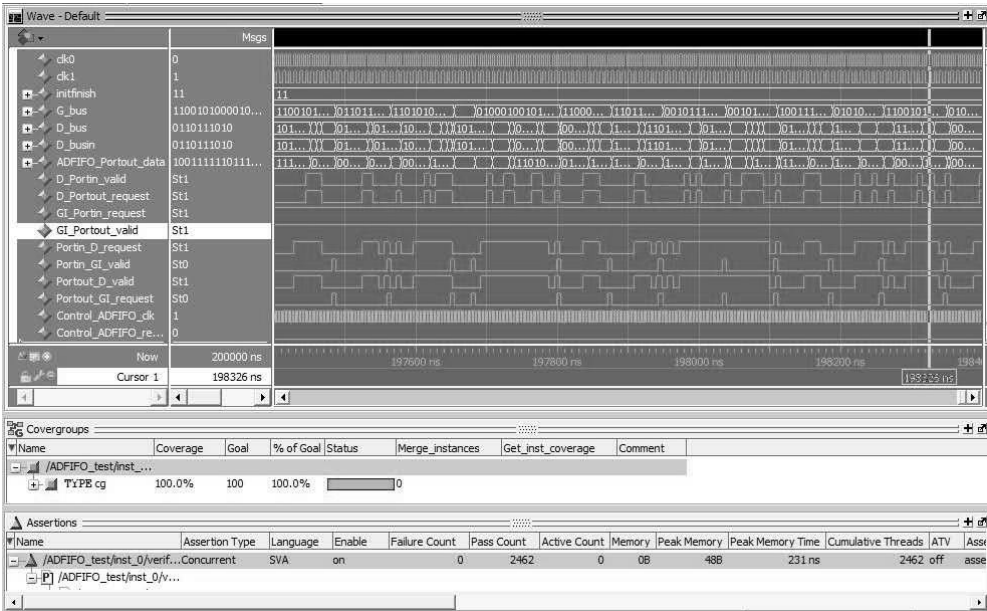


图 5.7 异步异带宽 FIFO 模块测试截图

5.2.4 重建精度测试结果与分析

如图 5.8 使用实验平台完成重建精度测试，记录数据于表 5.9 中

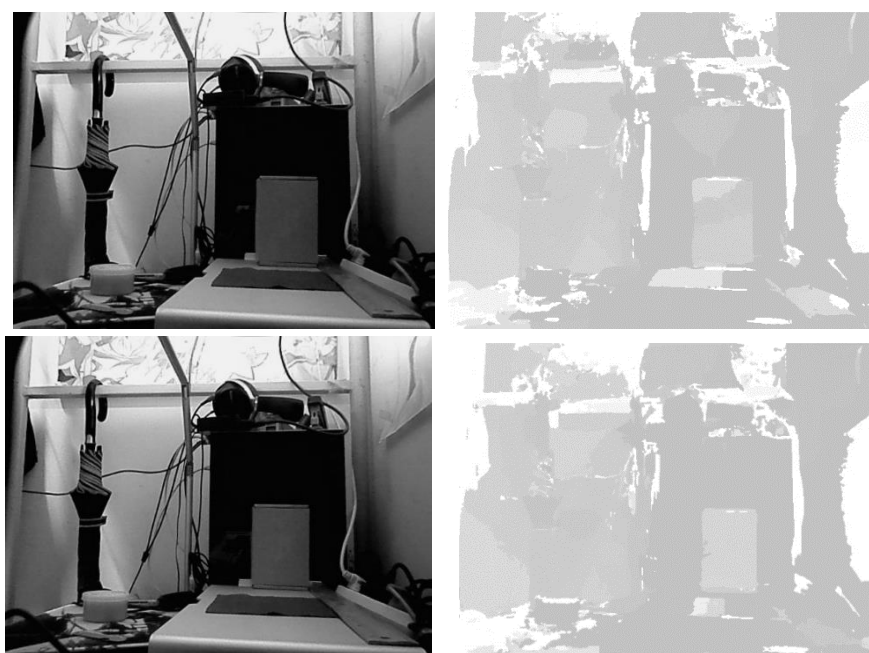


图 5.8 测试过程中拍摄到的目标靶图像

表 5.9 实验数据记录表

测试组别	实际位移/cm	原始数据的				
		测量位移 (部分)/cm	标准差/cm	平均位移/cm	绝对误差/%	计算距离/cm
1	2.0	2.002	0.055	2.033	1.67	98.4
		2.126				
		1.989				
		2.016				
2	2.0	1.966	0.068	2.079	3.95	125.5
		2.116				
		2.093				
		2.142				
3	2.0	2.190	0.114	2.233	11.65	157.7
		2.406				
		2.093				
		2.243				
4	2.0	2.229	0.137	2.162	8.11	180.2
		1.981				
		2.096				
		2.343				
5	2.0	2.376	0.077	2.255	12.73	214.3
		2.263				
		2.172				
		2.206				

根据测量到的原始数据（表 5.9 测量位移一行）计算出原始数据的标准差、平均位移并计算绝对误差并记录于表 5.9 中。原始数据中计算距离这一项为在视差图窗口中五点采样求平均值视差，以中间点为中心计算得到的目标靶与相机中心的距离信息，由于我们没有方法得到相机中心与标定平台起始位置的精确距离，此项仅用作参考分析。

通过对表 5.9 采集到的原始数据以及计算得到的各项数据分析，我们可以明显地看出，随着距离的增加，测量位移的标准差呈增大的趋势，说明随着距离的增加深度计算的精准度逐渐下降，采集到的深度数据逐渐的分散；而平均位移和相对误差反常地出现了降低的情况，我们猜测这是由于采样数据样本不充足造成的。虽然和成熟的结构光深度相机仍有较大的差距^[13]，但也成功的验证了本系统实现方案的可行性；通过后续对各环节可能仍潜在的问题进行推断总结，改善并提升该方案的实用性。

5.3 实验结果总结

5.3.1 目标场景视差图结果分析

由于条带状图样每帧只能还原条带状图样附近目标场景的三维结构，所以我们将通过控制每帧间隔大致调整条带状图样的位置，逐渐累积还原整个场景的三维结构。如图 5.10 所示，我们使用实验装置进行测试，目标场景为实验平台，设定有效视差面积阈值为 70%，通过软件算法处理并叠加重建结果得到目标场景视差图结果和计算到的场景距离。

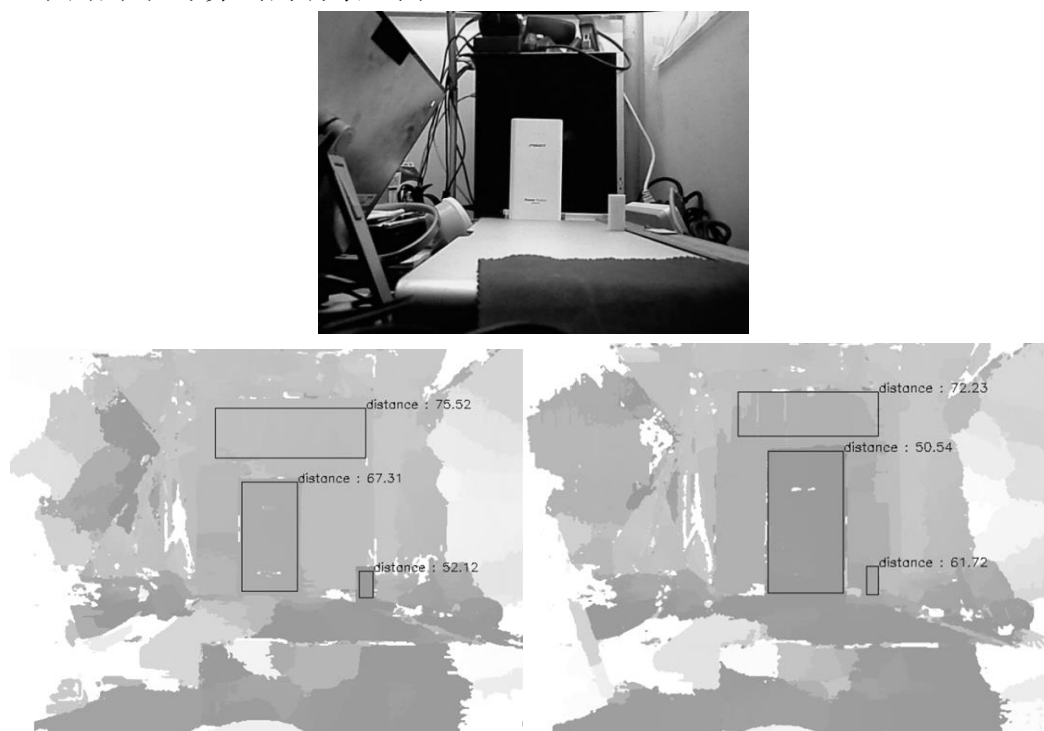


图 5.10 实验装置拍摄到的图像（上图）和测量不同距离的目标物体（下图）

5.3.2 潜在问题分析

软件算法和硬件电路完美配合，成功的实现了对目标场景的视差图计算，从实验角度上验证了本系统实现方案的可行性，重建精度测试环节也说明了优化算法对传统双目方案的补足还是有一定的效果。但该实现方案和其他成熟的深度相机实现方案相比得到视差图的速度慢、视差图完整性不足，我们将通过分析潜在的因素并提出修正方案，希望能提升本文提出的系统实现方案的可行性和实用性。

对重建结果分析我们发现使用累积法计算视差图的方法和其他方案在视差图的完整性和精确度上有一定的差距，通过观察我们发现视差图中出现了大量的空洞和没有正确还原到的区域，分析各环节仍存在以下潜在问题：

对于和条带状图样标记平行地目标场景物体，如果条带状图样地宽度超过物体的宽度，软件系统就不能计算出该物体表面的视差，甚至由于遮挡导致条带状图样分裂成两条，软件系统的设计时没有考虑到这种可能出现的情况的有效对策。

随着距离的增大，激光束扫过固定扇角形成的条带状图样标记地宽度也会随之增加，当前设计的条带状图样中线提取算法对较宽的图样中线提取地精度很差，后续需要对算法进行优化使其能适应较宽的条带状图样。原因分析如下：如图 5.11 所示，较宽的条带状图样整体亮度较低，收到背景光的干扰较强使得条带状图样出现不连续、和背景区分不明显的情况，导致中线提取误差很大。



图 5.11 固定变量测试时发现的潜在问题

较宽的标记图样会导致遮挡现象发生的更加频繁，非常影响得到的视差图的可靠性和精度。

5.3.3 解决方案探讨

针对在实验环节发现的潜在问题,在上一章节我们对这些潜在的问题进行了分析,并在本文提出的系统实现方案的基础上构思了新的深度相机实现方案,新的设计方案有如下四点改进方向:

1. 相机模块将使用一个相机,镜头部分加装滤光片以此减少可见光的干扰,修改相机模块的控制器,将使用 FPGA 替代控制器实现对相机模块的直接控制,通过改变曝光时间间接的控制条带状图样的宽度。

2. 相机模块与标记模块构建新的几何模型,该几何模型和三角测量模型相似,但存在较多的变量需要进行标定;为了解决这个问题,我们使用辅助相机,先实现对双目系统的标定,然后通过双目系统计算深度信息反向标定标记模块的参数,得到需要的参数后可以撤下辅助相机,仅用一个相机便可以实现三角测量。

3. 通过 FPGA 实现微秒级时间测量,精确控制相机模块的曝光时间,我们可以控制相邻两帧条带状图样存在微小的角度差,通过对这两帧图像进行差分计算,可以消除掉环境光的干扰并且得到两条极细的条带状图样,很好的改善了遮挡现象。

4. 设计新的算法对三角测量单次恢复的三维结构进行累积,以及在软件系统中实现深度图像的滤波,增强等后处理功能。

结束语

通过对重建结果环节的效果展示的分析,使用标记模块实现主动双目方案比传统的被动双目方案在恢复目标场景三维结构的精度上有一定的提升。本文提出的条带状图样中线提取算法,通过限定匹配范围和条带状图样方向,提取条带状图样中线并进行匹配计算出三维结构信息,耗时 12ms(图像分辨率为 640 X 480),和其他算法相比(如 Steger 算法^[14])在时间复杂度上有明显的降低,满足应用的实时性需求。该方案适用于对恢复三维场景结构精度要求较高、算力不充裕的应用场景。

虽然本系统实现方案在恢复三维场景的精度上有一定的提升,但引入标记模块使得本方案丧失了传统双目方案结构简单的优势,并且使相机拍摄的图像序列每帧只能恢复出标记附近的三维结构,极大的降低了目标场景三维结构的恢复速度,使得本方案在实用性方面大打折扣。

针对本方案的种种不足,总结分析在实验过程中发现的各种潜在问题以及没有考虑周全的部分设计,我们提出了新的深度相机设计方案,新提出的方案虽然没有改善恢复目标场景三维结构的速度缓慢的问题,但在保持恢复目标场景三维结构高精度的基础上,可以实现多机时分同步扫描相同的的目标场景(结构光方案无法实现的多机同时工作)、极强的抗干扰能力(差分图像减少了室外红外波段强光的干扰)、结构不受约束(单相机和标记模块的位置没有双目方案那么苛刻的要求,只需要满足一定的基线长度即可)、对平台的计算能力要求更低等目标,这些优势提高了新方案的实用性;未来我们将以设计、验证并实现新方案为目标继续努力。

致 谢

参考文献

- [1] 丁洁琼. 基于 RGB-D 的 SLAM 算法研究[D].西安电子科技大学,2014.
- [2] 佟帅,徐晓刚,易成涛,邵承永.基于视觉的三维重建技术综述[J].计算机应用研究,2011,28(07):2411-2417.
- [3] Lee, Seungkyu. Depth camera image processing and applications[P]. Image Processing (ICIP), 2012 19th IEEE International Conference on,2012.
- [4] 贾星伟,殷晨晖,隋国荣.单目摄像机线性模型可靠性分析与研究[J].光学仪器,2019,41(01):60-67.
- [5] 李轩. CMOS 图像传感器噪声抑制研究[D].天津大学,2010.
- [6] 阳鹏程. 基于双目结构光的三维测量技术研究[D].浙江大学,2014.
- [7] 王美丽.霍尔传感器的应用探讨[J].山西电子技术,2010(04):84-86.
- [8] 杨宏,李亚安,李国辉.一种改进扩展卡尔曼滤波新方法[J].计算机工程与应用,2010,46(19):18-2.
- [9] 潘秋娟,杨艳,常亮,王懿.400~1400nm 脉冲激光辐射人眼安全分析[J].激光与红外,2010,40(08):821-824.
- [10] 靖文,王影星.不同驱动方式对无刷直流电机性能的研究[J].电气传动,2019(05):3-6+24.
- [11] ON Semiconductor's, MT9V034C12STC-DP\ MT9V034C12STC-DP1: CMOS Image Sensor, Digital, Global Shutter, VGA, 1/3" Datasheet[January, 2017 – Rev. 7]
- [12] Intel,Avalon® Interface Specifications, Updated for Intel® Quartus® Prime Design Suite: 18.1, 2018.09.26
- [13] S. Zennaro et al., "Performance evaluation of the 1st and 2nd generation Kinect for multimedia applications," 2015 IEEE International Conference on Multimedia and Expo (ICME), Turin, 2015, pp. 1-6.
- [14] 刘正斌,朱丹,佟新鑫.基于 GPU 的结构光中心快速提取方法及实现[J].计算机应用与软件,2014,31(08):225-228.