

武汉大学大学生创新创业训练计 划项目科研总结

基于双目立体视觉的目标检测和距离测量 系统

院（系）名 称：测绘学院

专 业 名 称 ： 测绘工程

学 生 姓 名 ： 褚睿韬 孙章宇

胡昊杰 殷程昱 江洋

指 导 教 师 ： 贾剑钢 实验师

李明 副研究员

二〇一七年三月

郑 重 声 明

本项目组呈交的结题报告，是在导师的指导下，独立进行研究工作所取得的成果，所有数据、图片资料真实可靠。尽我们所知，除文中已经注明引用的内容外，本报告的研究成果不包含他人享有著作权的内容。对本报告所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确的方式标明。本报告的知识产权归属于培养单位。

项目组签名：_____ 日期：_____

摘 要

智能机器人、无人车、视觉导航正是当下研究的热点，双目视觉技术是计算机视觉技术的一个重要分支，它模拟人类视觉原理，使用计算机被动感知距离。从两个或者多个点观察一个物体，获取在不同视角下的图像，根据图像之间像素的匹配关系，通过三角测量原理计算出像素之间的偏移来获取物体的三维信息。得到了物体的景深信息，就可以计算出物体与相机之间的实际距离，物体 3 维大小，两点之间实际距离。目前也有很多研究机构进行 3 维物体识别，来解决 2D 算法无法处理遮挡，姿态变化的问题，提高物体的识别率。本项目研究的基于双目立体视觉的目标检测和距离测量系统，旨在实现仅利用双摄像头或立体相机进行高精度的距离测量和复杂环境下的动态目标检测，从而改进无人车、机器人、视觉导航中的感知精度等问题。本项目由褚睿韬、孙章宇、殷程昱、胡昊杰和江洋五位 2014 级导航工程的同学完成。项目过程中小组成员将计算机视觉知识和本专业的专业知识相结合，提高了项目开发能力，非常感谢项目过程中贾剑钢老师和李明老师在技术和实验环境上提供的支持，没有贾老师和李老师的鼓励和引导，我们团队无法顺利地进行系统的设计和开发。

一、项目背景

双目视觉技术是计算机视觉技术的一个重要分支，它通过模拟人的双眼用计算机实现人的视觉系统。获取空间三维场景的距离信息是计算机视觉研究中最基础的内容。获取距离信息的方法和技术很多。立体视觉是计算机被动测距方法中最重要的距离感知技术，它直接模拟了人类视觉处理景物的方式，可以在多种条件下灵活的测量景物的立体信息，其作用是其它计算机视觉方法所不能取代的。对它的研究，无论是从视觉心理的角度还是在工程应用中都具有十分重要的意义。双目立体视觉这一有着广阔应用前景的学科，随着光学，电子学以及计算机技术的发展，将不断进步，逐渐实用化，不仅将成为工业检测，生物医学，虚拟现实等领域。目前在国外，双目立体视觉技术已广泛应用于生产，生活中，而我国正处于初始阶段，尚需要广大科技工作者共同努力，为其发展做出贡献。

二、成员概况

（一）指导老师

贾剑钢：武汉大学测绘学院实验师，担任本项目的指导老师。贾老师对该项目任务进行细化并制定计划，一步步对我们进行指导，并督促我们按时按质完成计划，其过程不忘让我们学习额外的知识，让我们受益匪浅。

李明：武汉大学计算机学院副研究员，李老师多年来从事双目视觉的研究，为本项目的提供了重要的技术指导。

（二）小组成员

本组人员均为 14 级导航工程班的同学，对于摄影测量和数据处理有相关的专业知识基础，5 个人对于进行视觉识别和基于视觉识别的目标检测和距离测量有相关了解，并且兴趣浓厚。

褚睿韬：通过计算机三级考试，并且通过选拔进入测绘与遥感信息国家重点实验室无人车团队同时进行学习，参与过实际工程编程工作，能熟练使用 C++，java 等计算机常用编程语言，熟练使用 Linux 操作系统并可以进行常用的程序编写。

孙章宇：编写过导线测量和水准测量的平差程序，2015 年测绘技能大赛软件版块三等奖。熟悉 C, C++, 接触过 C#, JAVA 等编程语言，熟悉 PHP 语言，同时会 Linux 系统的操作，会写 shell 脚本语言。

胡昊杰：对编程感兴趣，通过了计算机二级 C 语言考试，获得过测绘技能大赛程序设计版块二等奖。

殷程昱：喜欢科研，关心世界上最新的科研进展，对编程有浓厚的兴趣，通过了全国计算机等级二级 C 语言考试。

江洋：计算机三级网络工程师，目前从事激光点云与图像处理相关研究和学习，研究方向为利用点云的栅格性质和图像的点阵性质综合处理点云或图像，在遥感院国家重点实验室李必军教授的无人车团队里学习。熟练掌握 android 移动开发, windows 开发, 网络编程, 网站网页设计, 一年内自学 C, C++, C#, JAVA, HTML, CSS, JavaScript, MATLAB, SVG 画图, autolisp, 汇编语言。

三、创新特色

智能机器人、无人车、视觉导航正是当下研究的热点，本项目研究的基于双目立体视觉的目标检测和距离测量系统，旨在实现仅利用双摄像头或立体相机进行高精度的距离测量和复杂环境下的动态目标检测，从而改进无人车、机器人、视觉导航中的感知精度等问题。研究涉及摄像机标定，立体匹配，基于视差图的距离测量，基于 V-视差图的目标检测，基于光流法、背景差分法、连续帧间差分法的动态目标检测技术等内容。研究内容充实，创新特色鲜明。

四、项目成果

本项目建立在双目视觉测距的基础上，通过单个相机获得物方点的平面坐标，通过获得两个摄像机的相对位置关系和两个摄像机拍摄到的影像的差异获得深度信息，重建拍摄时的三维情景。通过搭建自己的测试平台达成了基于视觉的距离测量和三维重建的目标。

（一）平台框架

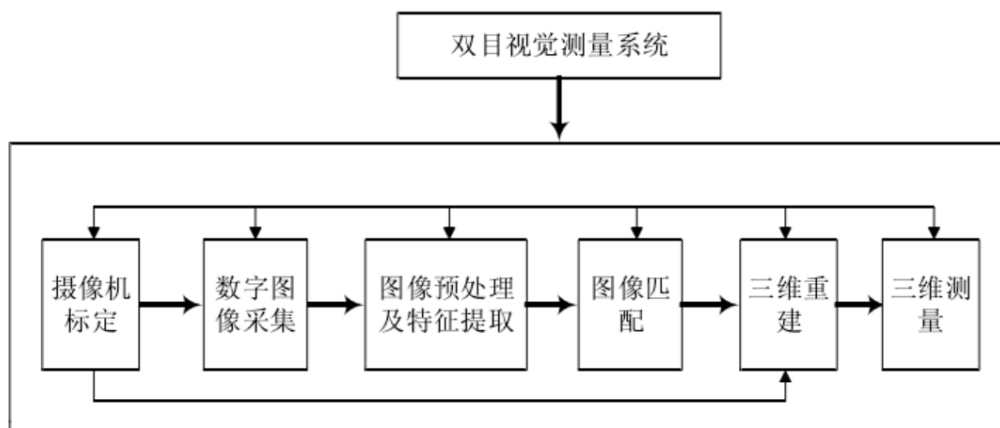


图 4.1 三维重建流程

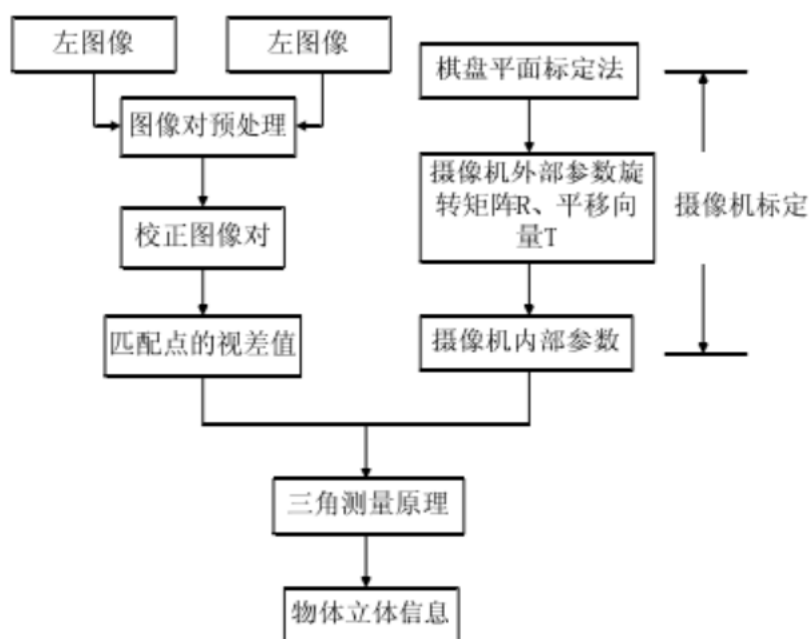


图 4.2 距离测量流程

项目平台依赖 OpenCV 开源视觉库，平台可视化框架建立在 Qt 之下。

程序窗口如下图，包含两个显示摄像头的窗口和一个显示视差图和三维重建效果的窗口，以及设置程序参数和控制的元件。

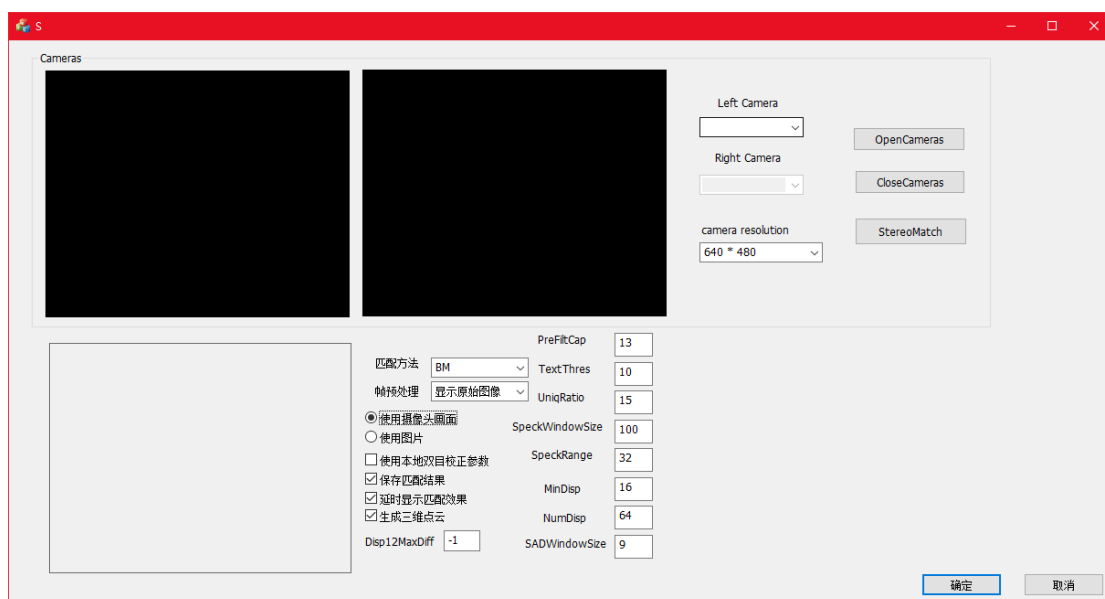


图 4.3 程序窗口

(二) 基于视觉的距离测量

1. 视差图生成

程序获得由摄像头拍摄的画面后将图像交由后台处理，但是由于计算机性能有限，无法在 CPU 上实现即时的处理与显示，在查找多方资料后决定使用 Intel 公司开发的 SSE 指令集以提高程序运行速度。使用 SSE 指令集直接操作 CPU 进行计算可以大大加快计算速度，减少了视频处理过程中的延迟。经过实验处理可将处理速度提到三倍以上。

在视差图计算过程中，小组成员对比后决定使用 SGBM 算法，而 SGBM 算法已经有和 SSE 指令集融合的高速算法 RSGM。SGBM 算法通过选取每个像素点的视差，组成一个视差图，设置一个和视差图相关的全局能量函数，使这个能量函数最小化，以达到求解每个像素最优视差的目的。

2. 程序运行结果

程序完成后选择一间会议室为场景进行测试，包括利用摄像头图像进行视差计算及计算结果的后处理，将视差按照大小进行上色，方便观察。



图 4.4 原图



图 4.5 视差图

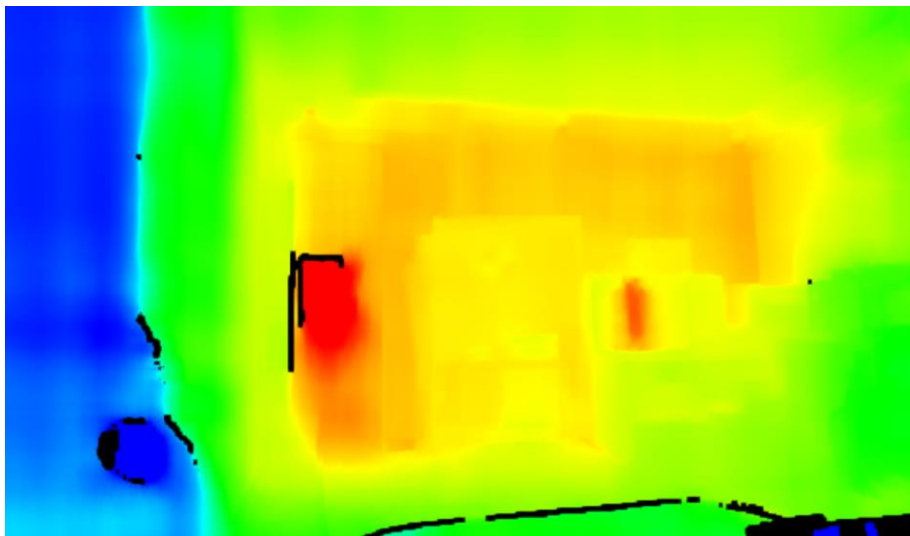


图 4.6 处理后的视差图

视差图除去进行距离测试外还可以进行物体与背景的剥离。通过视差图进行聚类后，可以将图像进行分层，分层后获得的结果进行上色后可以明显看出物体被从背景中剥离。依据这个原理可以动态检测摄像头视野中的障碍物。

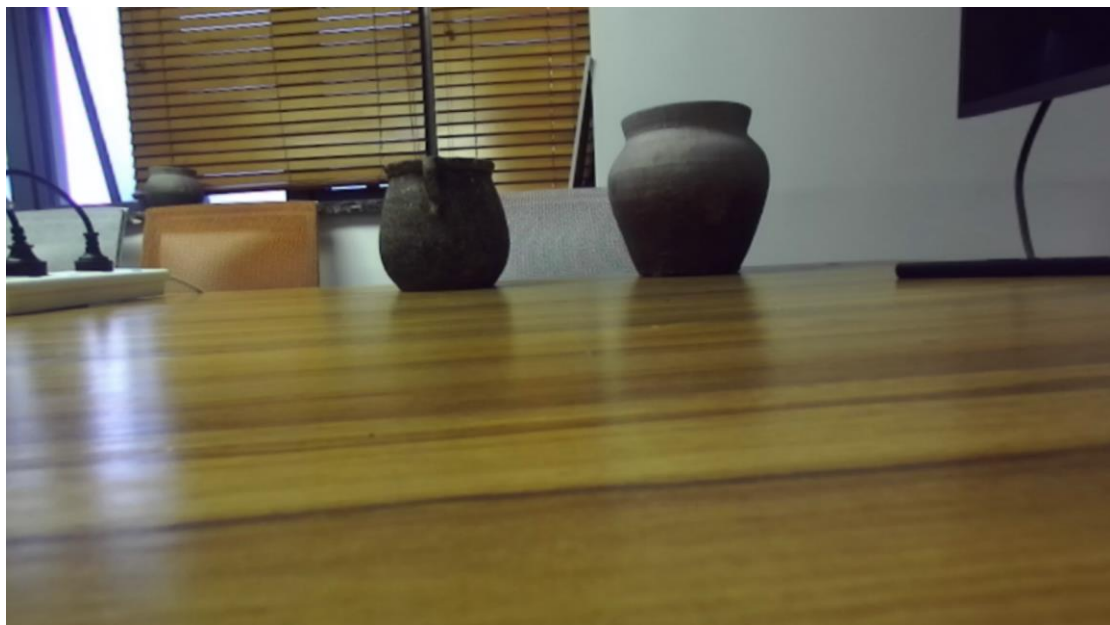


图 4.7 原图



图 4.8 分层后的视差图

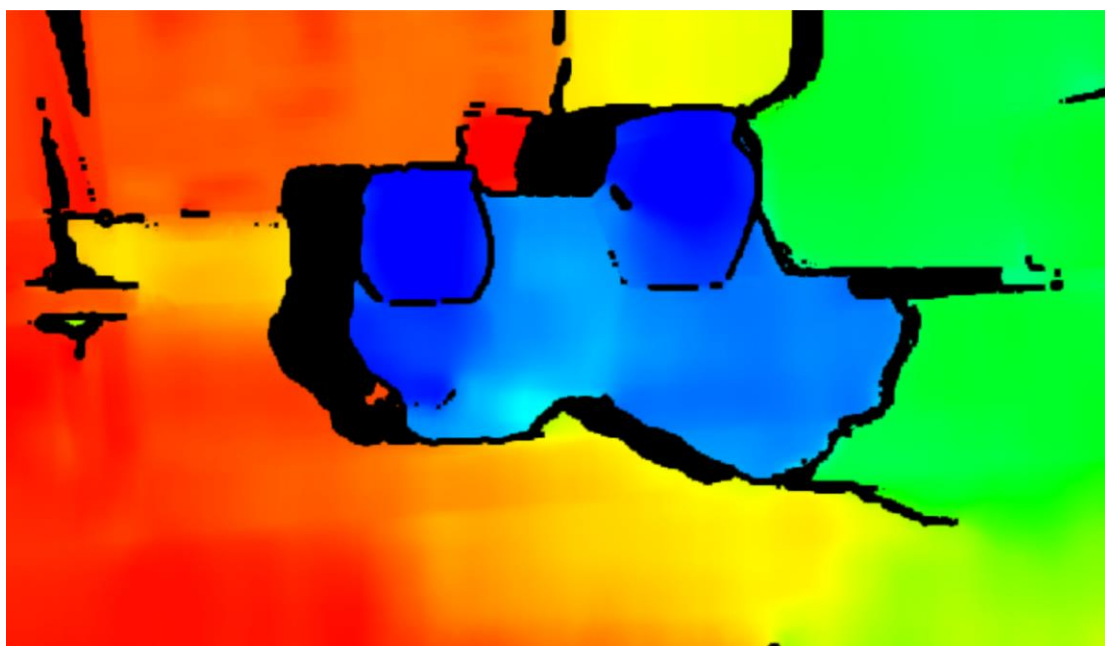


图 4.9 对分层图像进行上色

(三) 三维重建

1.程序架构

三维重建主要是用 OpenGL 实现的，根据立体匹配得到的视差图结果将不同远近的事物在三维空间里进行分层从而实现三维效果，同时提取左视图中的色彩构成纹理贴在三维重建好后的图形点云上，完成三维重建。

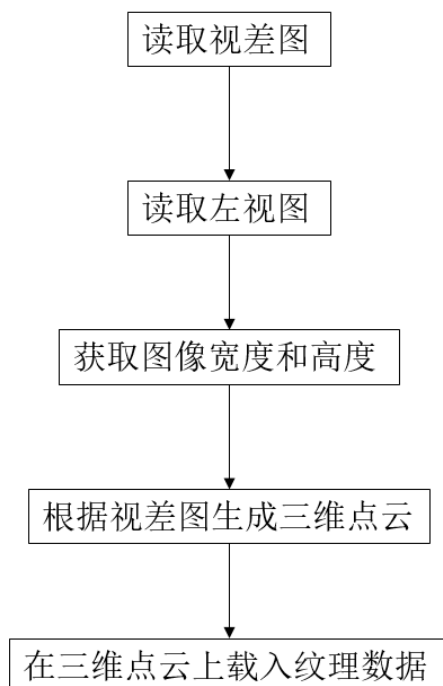


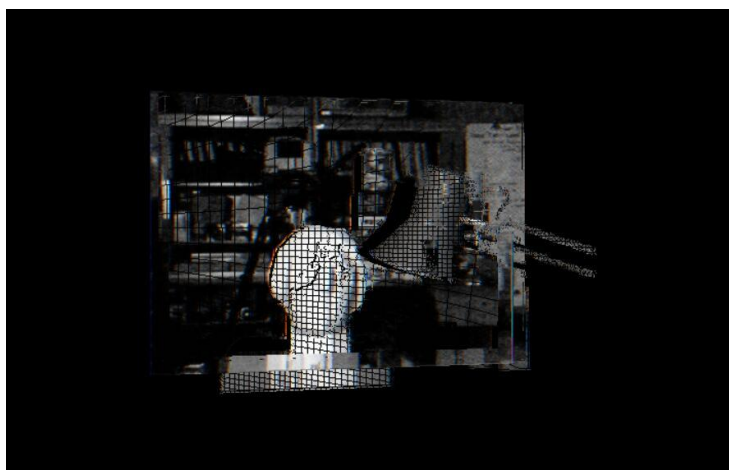
图 4.10 三维重建程序框图

程序中首先用 OpenCV 中的 `cvLoadImage` 读取视差图和左视图，视差图的信息存储在 `imgGrey` 变量中，左视图的信息存储在 `imgTexture` 变量中。然后获取图像宽度和高度，接着根据视差图生成三维点云，最后在三维点云上载入纹理数据。

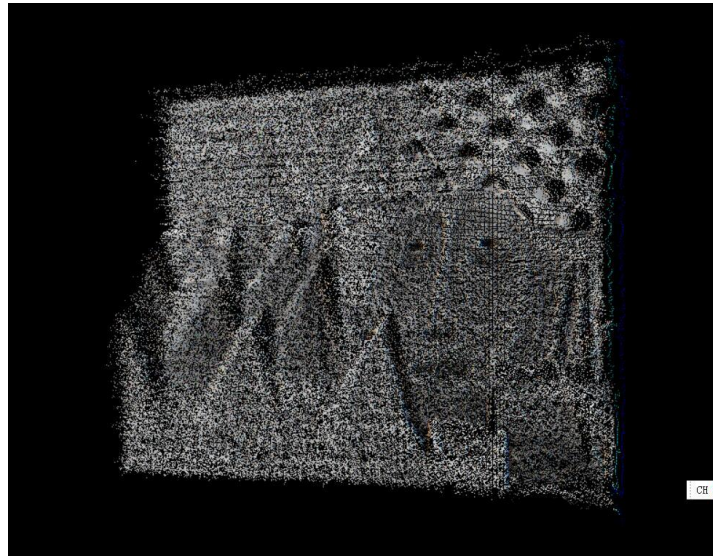
在进行三维重建生成三维点云时，我们并没有根据标定参数和视差图计算图像中每一个点的三维坐标再进行重建，而是用一种比较简单的办法呈现出三维效果。三维重建中，关键是每个物体的深度值如何获得，事实上，匹配得到的视差图中就已经包含了深度信息——视差图用灰度值深浅的方法标识深度信息，灰度值小的地方离摄像头越近，灰度值大的地方离摄像头越远。为此，我们在三维重建时将用灰度值最大值 255 减去该点的视差图中的灰度值作为该点的深度值。这样，在生成三维点云时，我们只用视差图就呈现出了三维效果，而免去了计算每个点三维坐标时庞大的计算，提高了运算速率。

2.重建效果

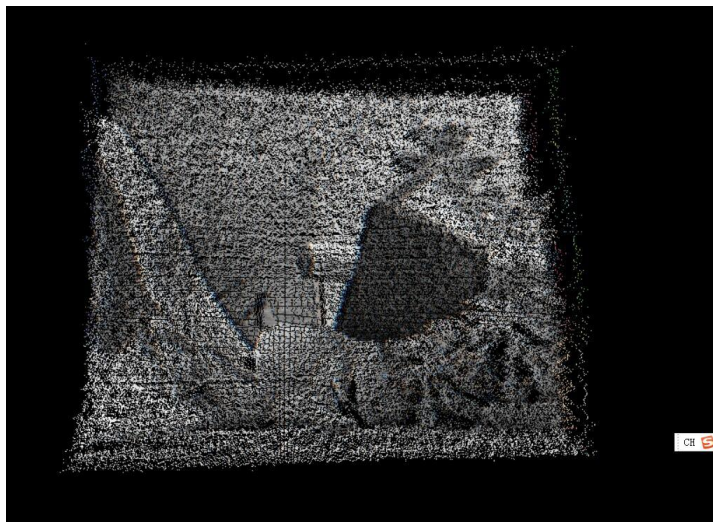
我们从网上还有别人的论文上找到了很多别人做出来的视差图 and 原视图来测试我们的三维重建程序，观察三维重建效果，以下是我们三维重建出来的效果（图 2）：



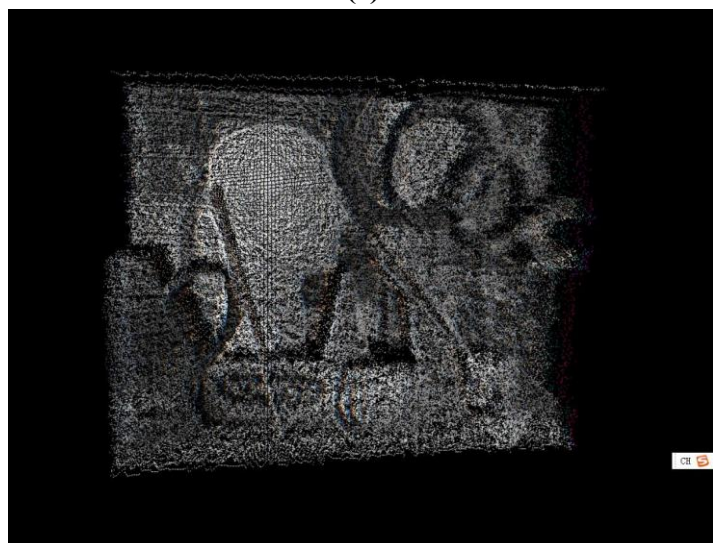
(a)



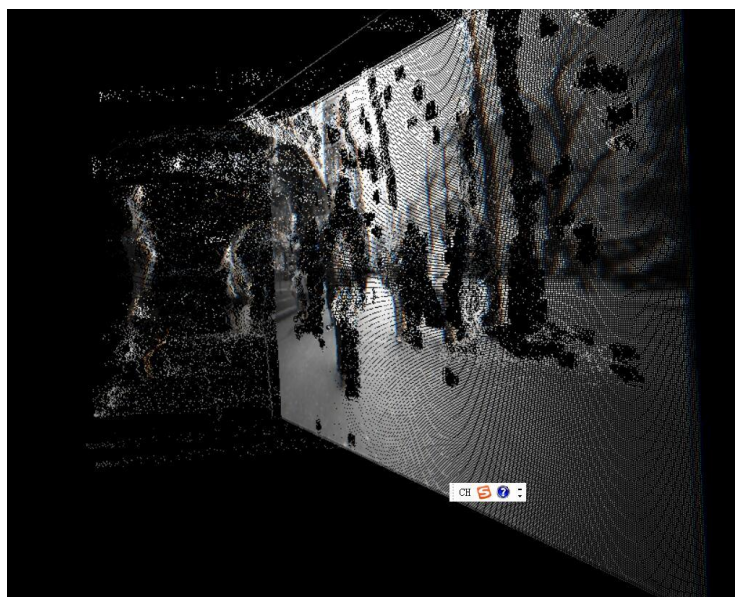
(b)



(c)



(d)



(e)

图 4.11 三维重建效果

从我们的三维重建图形中可以看出，虽然这些物体不是根据实际计算出来的三维坐标绘制出来的，但也已经出现了明显的三维效果，将远近事物进行了三维分层，基本实现了三维重建功能。

五、收获体会

二零一六年三月，我们团队五位同学在贾剑钢老师的指导下，对于本项目有了初步的认识和规划。迄今为止，我们也坎坎坷坷地走过了一年的时间，回顾这一年的科研历程，不及细细思量，颇多感触却已涌上心头。

一年前，当我们和贾老师对小组提出的科研项目进行讨论时，我们五位同学没有任何头绪，对于项目主题认识不清导致不应该从何下手。但是通过一遍一遍地写申请书，我们对于项目的印象在逐渐加强，对于项目的理解也越来越深刻，并且在此过程中，我们的文档写作能力也在逐渐提高，从最初会有段落的首行缩进等小问题到之后图表、图片的绘制与排版，我们写作的文档越来越规范。

虽然在项目初期对整个系统有进行完整的规划，但当进行深入的思考时却发

现有些功能我们目前根本无法实现，有些功能实用性较差，没有实现的必要。我们阅读了很多相关的资料，也比较详细的了解了与此有关的项目所做的事情，并且向学长们请教了很多，最后一步步地将自己将要做的项目的框架构建起来，再一步步完善细节。虽然我们不知道大概要做什么，但是，如何向别人明确表达，让其他人也能很容易明白我们将要做的事情，这是另一回事。这涉及到如何通过简洁明了的方式向别人表达我们的想法。我们尝试用不同的方式去呈现：文字、图表、视频，最后我们确定用制作相对简单，但又不会显得太过呆板的图表来表达。

项目程序上的窗口负责图像和视频的显示，窗口右边的参数控制可以在程序运行过程中进行调整，方便观察算法计算结果的变化。

开始构想程序框架时比较复杂，不仅要有完善的文件管理系统，还要有一个相对完整的内存处理系统和高效的计算方法，完全不知道该从何下手。经过长时间的考虑，最终一步步的理清了思路，并且将构想的功能逐个实现。首先我们的程序一定要有良好的界面方便调试程序，观察运行结果，因此我们可视化框架采用了可移植性强，相对成熟并且参考资料较多的 Qt 框架，在 windows 和 Linux 上运行都有较好的效果，无需对程序进行修改。然后是程序具体的功能和运行流程，我们主要提供了两方面的功能，一是展示视差图，可以观察算法的结果，初步对摄像机视野中的障碍物进行分离，为此我们查阅了大量资料同时和指导老师进行交流，在不断的尝试之后使用的 RSGM 开源库解决了程序运行过程中延时和计算效果差的问题，对于性能较差的设备也有较好的支持；另一个功能是对摄像机拍摄到的场景进行三维重建，三维重建后的数据进行上色后可以作为建模使用或对数据进行分析，获得更多的实时无法获取的场景信息。

这几个月，坎坎坷坷，有失败时的懊恼、有成功时的骄傲，在付出的同时，我们学到了很多。在此过程中，我们认识到了团队协作的重要性、提高了编程和文档写作的能力、掌握了自主学习的方法。在今后的学习和工作中，相信这次大学生科研一定会成为我们宝贵的经验和美好的回忆。