





# 基于平行双目视觉的环境特征提取方法研究

学位论文完成日期: 2010.5  
指导教师签字: [Signature]  
答辩委员会成员签字: [Signature]

[Signature]  
[Signature]  
[Signature]



## 独 创 声 明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的研究成果。据我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含未获得（注：如没有其他需要特别声明的，本栏可空）或其他教育机构的学位或证书使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

学位论文作者签名：王伟良 签字日期：2010 年 5 月 28 日

## 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，并同意以下事项：

1、学校有权保留并向国家有关部门或机构送交论文的复印件和磁盘，允许论文被查阅和借阅。

2、学校可以将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、汇编学位论文。同时授权清华大学“中国学术期刊(光盘版)电子杂志社”用于出版和编入 CNKI《中国知识资源总库》，授权中国科学技术信息研究所将本学位论文收录到《中国学位论文全文数据库》。

(保密的学位论文在解密后适用本授权书)

学位论文作者签名：王伟良

导师签字：田 斌

签字日期：2010 年 5 月 28 日

签字日期：2010 年 5 月 28 日

1000 1000 1000

1000

1000

1000

# 基于平行双目视觉的环境特征提取方法研究

## 摘 要

基于平行双目视觉的环境特征提取方法研究是移动机器人导航领域的重要研究内容。它利用左右两个视觉传感器采集的图像进行处理,提取其中的环境特征信息,为机器人导航的控制决策提供可靠的信息保障。

本文在传统的双目视觉理论研究的基础上,结合有关课题中移动机器人导航应用的实际要求,研究了平行双目视觉系统中的成像模型和摄像机内外参数的标定,并通过设计实验来达到实际应用的要求;接着研究了特征点的提取和立体匹配,视差图的生成,不规则障碍物的检测及其三维信息的计算,并且在各部分通过实验以验证算法的有效性。论文的主要创新点有以下几个方面:

(1) 研究了张正友摄像机标定方法,设计了符合实际应用的平行双目视觉系统,并通过设计实验确定摄像机内外参数,提出了实际应用中所应注意的问题并给出解决方法。

(2) 研究了特征点的提取和立体匹配,改进了一种基于点特征的双目视觉图像立体匹配方法。通过采用 Harris 角点的提取算子,首先对双目视觉图像中的特征点进行检测,并改进了 SIFT(Scale Invariant Feature Transform)算法中的特征描述方法在 Harris 特征点描述当中的应用;然后利用欧氏距离对特征点进行匹配;最后提出一种利用匹配点对的斜率来排除错误匹配的方法。实验表明该算法不仅可以达到较高的匹配精度,还可以有效缩短匹配时间,满足自主移动机器人双目视觉实时性的要求。

(3) 介绍了双目视觉在目标检测领域的研究进展,并根据实际项目中移动机器人的要求,提出了一种基于视差图的不规则障碍物检测方法。该方法利用视差图中匹配点的高度信息和匹配点连通区域的范围作为判别障碍物的首要依据,排除了形状和颜色所带来的判别障碍,并通过实验验证该法的有效性。

作者期待本文能够对基于平行双目视觉的环境特征提取具有一定的参考意义,从而对双目视觉和机器人导航的相关理论建设和实际应用做出贡献。

**关键词:** 双目视觉; 机器人导航; 特征点匹配; 障碍物监测



# Research on Extraction Technologies of Environmental Characteristics Based on Parallel Binocular Vision

## Abstract

Extraction technology of environmental characteristics based on binocular vision is an important research element in the robot navigation field. It uses two vision sensors to capture a great deal of environmental information, which other sensors can not compare with. The environmental information can guarantee a reliable decision for the control of robot navigation.

In this paper, the traditional theory of binocular vision is studied. Combining with the requirements in the practical application of mobile robot navigation, the paper carries out a research on the imaging models and calibration of camera parameters. Then the paper designs experiments to achieve the practical requirements. In the later chapters, extraction of the feature points and stereo matching, generation of disparity map, irregular obstacle detection, as well as the calculation of three-dimensional information are all studied. In the same time, experiments have been done to verify the effectiveness of the algorithm in all parts. The principal innovations of the paper are listed below:

(1) Camera calibration method raised by Zhengyou Zhang is studied in the first place. Then the paper designs the binocular vision system in line with the practical application and through the design of experiments, the paper calculates the parameters of the cameras. The paper puts forward the problems which should be paid attention to in the practical applications and gives appropriate solutions.

(2) The paper improves an image matching method of binocular stereo vision based on feature points. Firstly, the Harris corner detector is improved to detect feature points and the method to describe the features of the points with SIFT(Scale Invariant Feature Transform) is enhanced. Secondly, Euclidian distance is utilized to get exact matching in the points set above. Finally, a simple efficient way to eliminate wrong matches is given. This method is proved to meet the needs of real-time binocular vision by actual test, and it can be applied to practice.

(3) The paper describes the research progress in the target detection area of binocular vision, and based on the actual requirements of the project, the paper presents an irregular obstacle detection method based on disparity map. The method



considers the high information and continuous range of the matching points as the primary basis for distinguishing obstacles, excluding the difficulties by using shape and color. In the last part, the paper verifies the validity of the algorithm through experiments.

The author expects the paper may have some reference meaning in research on robot navigation based on binocular vision, so that it can make certain contribution to the theory construction and practical application of binocular vision and robot navigation.

**Key Words: binocular vision; robot navigation; feature point matching; obstacle detection**

# 目 录

1 绪论	1
1.1 课题背景及研究意义	1
1.2 国内外双目视觉发展现状	2
1.3 环境特征信息分析	4
1.4 双目视觉关键技术	5
1.5 论文的研究内容和创新点	8
1.5.1 论文的研究内容	8
1.5.2 论文的创新点	9
2 成像模型与摄像机标定	10
2.1 摄像机成像模型	10
2.1.1 四个参考坐标系	10
2.1.2 针孔模型下平行双目视觉坐标公式推导	12
2.2 摄像机标定方法	16
2.2.1 传统摄像机标定方法	16
2.2.1.1 透视变换法	17
2.2.1.2 利用最优化算法的标定方法	18
2.2.1.3 Tsai 两步法	19
2.2.1.4 双平面标定方法	19
2.2.2 摄像机自标定方法	20
2.2.3 张正友标定方法	20
2.3 摄像机参数标定实验	20
2.3.1 平行摄像机视觉系统内参数标定及实验	21
2.3.2 平行摄像机视觉系统外参数标定及实验	23
2.4 本章小结	25
3 特征点提取与基于点特征的立体匹配	26
3.1 特征点提取	26
3.1.1 Moravec 算子	27
3.1.2 SUSAN 算子	28
3.1.3 Trajkovic 算子	30
3.1.4 Harris 算子	31
3.2 特征点描述符	32
3.2.1 特征点描述符概述	33
3.2.2 改进的特征点描述符	33
3.3 基于点特征的立体匹配	35
3.3.1 基于点特征匹配方法的特征分析	36
3.3.2 改进的基于点特征的立体匹配方法	37
3.4 实验结果	39
3.4 本章小结	41
4 不规则障碍物检测与特征信息提取	43
4.1 双目视觉障碍物检测分析	43
4.2 本文采用的障碍物检测算法	44

4.2.1 视差图信息提取.....	44
4.2.2 障碍物检测算法.....	47
4.2.3 障碍物特征信息提取.....	51
4.4 系统整合与实验结果 .....	51
4.4.1 平行双目视觉系统整合.....	51
4.4.2 实验结果.....	52
4.5 本章小结 .....	54
<b>5 总结与展望 .....</b>	<b>55</b>
5.1 总结 .....	55
5.2 展望 .....	55
<b>参考文献 .....</b>	<b>57</b>
<b>致  谢 .....</b>	<b>59</b>
<b>个人简历、在学期间发表的学术论文 .....</b>	<b>61</b>

# 1 绪论

## 1.1 课题背景及研究意义

本课题来源于国家 863 项目：基于声纳和水下视觉的深海复杂环境下 AUV 组合导航系统关键技术研究。本论文旨在探索运用计算机双目立体视觉系统，实现周围环境的自动感知，包括运行环境当中障碍物的自动采集，实时立体匹配，不规则障碍物识别以及障碍物方位和角度的计算，以达到移动机器人后续利用这些信息实现自主导航的目的。

自主移动机器人的研究与发展水平体现了一个国家的高科技水平与工业自动化的程度。机器人技术是集合了计算机技术、自动控制技术、通讯技术、传感器技术和仿生学等多学科和应用的一门综合性研究技术，体现了当前高科技的发展前沿，也是科技研究的热门方向。自主移动机器人是机器人研究领域的一个重要分支。移动机器人按其控制方式和自主程度大致可分为三种，分别是：遥控式、半自主式和自主式。其中自主式移动机器人的设计目标是在没有人的干预并且无需对环境做任何规定和改变的条件下，机器人在行进过程中，不断感知周围的环境信息，自主地做出各种决策，有目的地移动和完成相应任务。因此，自主式移动机器人称得上是一种具有高度自规划、自组织、自适应能力，适合于在复杂的非结构化环境中工作的机器人。

在自主式移动机器人研究当中，自主导航始终是研究的核心问题，也是当下机器人研究的热点。机器人自主导航代表了一种更高层的智能，要求机器人在能够判别并且避开障碍物的同时，走向目标。自主导航需要移动机器人根据环境信息，按照先前给出的任务命令，做出全局路径规划，并且需要在行进过程中，不断地感知周围的局部环境信息，自主做出决策，并随时调整自身姿态，引导自身安全行驶达到目标位置。

由于人类 70% 以上的信息是通过视觉来获取的，因此对移动机器人来说视觉信息的获得是至关重要的。双目立体视觉是基于视差原理，由两幅图像获取物体

三维信息的方法。双目视觉测距方法在移动机器人导航中被大量采用,是因为与激光、雷达、超声相比,双目视觉导航具有以下几项优点:首先激光、雷达和超声的原理都是通过主动发射脉冲和接受脉冲来测距的,因此当多个机器人同时工作时,相互之间可能产生干扰,而双目视觉是被动测量,多个机器人相互之间的干扰可以减少到最小;其次是即使在丢弃了绝大部分的视觉信息后,所剩下的关于周围环境的信息仍然比激光、雷达和超声更多更精确。

随着双目视觉理论的发展和完善,基于双目视觉系统的环境特征感知技术正越来越多的应用于移动机器人的导航定位系统中。双目立体视觉技术是对移动机器人外部环境中同一目标在不同视角下的两幅图像进行提取、识别、匹配和三维重建的过程,它可以为移动机器人提供关于目标的位置描述,以支持移动机器人完成导航定位以及探测识别等特定任务,克服了利用声纳进行感知所带来的缺陷。基于双目视觉的导航技术不仅可以大大提高移动机器人对周围环境的考察精度,而且可以减少移动机器人对昂贵的高精度导航设备的依赖,降低运行成本,为许多移动机器人的用户提供低成本的移动机器人导航技术和装备。因而市场应用前景非常广阔。同时这必将推动移动机器人对周围环境感知技术的发展,也将对移动机器人的降低成本和产业化起到极大的推动作用。

## 1.2 国内外双目视觉发展现状

自 1977 年英国的 Marr 教授提出了自己的全新视觉理论以来,该理论是视觉研究中迄今为止最为完善,影响最为广泛的视觉理论<sup>[1-2]</sup>。但随着双目视觉技术研究的不断深入, Marr 视觉理论的局限性开始越来越多的显现出来<sup>[3]</sup>。进入 80 年代中后期,大量空间几何的方法以及物理知识被运用在双目视觉技术的研究中。这一时期由于引入了主动视觉的研究方法,并采用了距离传感器和信息融合等技术,可以直接取得深度图或通过移动获取深度图,从而使得 Marr 视觉理论中很多病态问题变成了良态<sup>[4-5]</sup>。在这种情况下,双目视觉系统才得以广泛应用。随着双目视觉理论的完善和视觉传感器技术的发展,基于双目视觉的环境特征提取技术正越来越多的应用于移动机器人的自主导航当中<sup>[6-8]</sup>。

在美国,华盛顿大学与微软公司展开合作,为火星卫星“探测器”号研制了宽基线立体视觉系统。该系统使用同一个摄像机在“探测器”的不同位置上拍摄

图像对,拍摄间距越大,基线越宽,能观测到越远的地貌。系统采用非线性优化得到两次拍摄图像时摄像机的相对准确的位置,利用鲁棒性强的最大似然概率法结合高效的立体搜索进行图像匹配,得到亚像素精度的视差,并根据此视差计算图像对中各点的三维坐标。相比传统的立体视觉系统,该系统能够更精确地绘制“探测器”号周围的地貌和以更高的精度观测到更远的地形,使“探测器”号能够在火星上对其即将跨越的几千米内的地形进行精确的定位导航<sup>[9]</sup>。另外,美国普渡大学(Purdue University)机器人视觉实验室对视觉导航移动机器人进行了研究,其研制的视觉导航移动机器人 Peter 采用了双目主动立体视觉系统,可以获得移动机器人的运行环境和障碍物的立体数据<sup>[10]</sup>。美国国家斯坦福研究所(Rlintematoinal)研制的移动机器人弗来克(Flkaye)能够实现全局路径规划和动态地局部路径规划<sup>[11]</sup>。

在日本,奈良科技大学信息科学学院提出了一种基于双目立体视觉的增强现实系统(AR)注册方法,通过动态修正特征点的位置达到提高注册精度的目的。该系统将单摄像机注册(MR)与立体视觉注册(SR)相结合,利用 MR 和三个标志点来算出特征点在每副图像上的二维坐标和误差,利用 SR 和图像对计算出特征点的三维位置总误差,反复修正特征点在图像对上的二维坐标,直至三维总误差小于某个阈值。该方法比仅使用 MR 或 SR 的方法大大提高了 AR 系统注册深度和精度。另外,大阪大学自适应机械系统研究院研制的一种自适应双目视觉伺服系统,利用双目立体视觉的原理,以每幅图像中相对静止的三个标志为参考,实时计算目标图像的雅可比矩阵,从而预测出目标下一步运动方向,实现了对运动方式未知的目标的自适应跟踪。该系统仅要求两幅图像中都有静止的参考标志,无需摄像机参数。而传统的视觉跟踪伺服系统需要事先知道摄像机的运动、光学等参数和目标的运动方式<sup>[12-13]</sup>。

其他国家,如法国的 IFREMER 研究开发了视频镶嵌技术,并在 2003 年 ROV “VICTOR6000” 的海底火山探测实验中得到应用;英国克兰菲尔大学研究视觉探测系统用于水下航行器的智能导航,采用距离选通技术结合激光三角法测距技术<sup>[14]</sup>;澳大利亚的 CSIRO ICT Centre 在 AUV “StarBug” 上嵌入了一个基于视觉的运动控制系统,装有朝向下和朝前向的 CMOS 摄像机,分别用于运动估计和避碰控制;澳大利亚国立大学信息科学与工程研究院开展了 AUV 视觉导航和目

标跟踪技术的研究,开发了用于水下机器人 Kambara 的视觉系统<sup>[15]</sup>。

我国的双目立体视觉环境特征提取研究起步较晚,但是目前我国的很多研究机构都开展了这方面的研究,并且有不少的原样机问世。虽然我国的大部分研究成果还没有转化为生产力,但已具备了开展视觉研究的人才和技术条件,如清华大学的智能技术与系统国家重点实验室等。同时我国各大院校也正在加大双目立体视觉研究的力度,陆续也取得了一些成果。

中国科学院自动化所自行设计、制造的全方位移动式机器人视觉导航系统 CASIA-I 可广泛应用于医院、图书馆、科技馆及展览馆等公共场所,以提供服务、作业、展示与娱乐等功能(如图1-1所示)。该移动机器人的基本结构包括传感器、控制器和运动机构。传感器由位于机器人底层的十六个触觉红外传感器、位于机器人中间两层的十六个超声传感器和十六个红外传感器、位于机器人顶部的摄像机等组成。再多传感器信息融合的基础上,完成感觉周围环境的情况变化等各项外界感觉功能,再由中央处理器将获取的各种信息综合起来进行处理,从而使其理解自己的状态和所处外部环境信息,实时做出运动控制决策,躲避障碍物和寻找最优路径,以实现自主移动和轨迹跟踪等<sup>[16]</sup>。

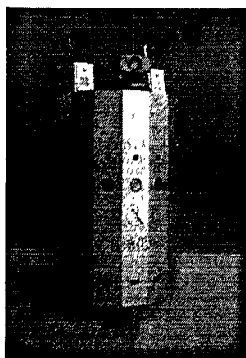


图1-1 中国科学院智能移动机器人CASIA-I

此外,哈工大采用异构双目视觉系统实现了全自主足球机器人的导航;中国科学院沈阳自动化研究所研制的自动引导车辆系统(AGV)和防爆机器人;吉林大学智能车辆课题组研制了一种新型视觉自动引导车,该车以地面上涂设的条带状路标作为路径标识符,运用计算机视觉快速识别路径。

### 1.3 环境特征信息分析

平行双目视觉系统在实际应用中,除了可以用图像自动记录下周围环境的信息外,一般用于移动机器人对周围环境的感知,这包括对采集到的图像进行处理和特征信息提取。本文针对移动机器人自主导航中需要用到环境特征信息进行分析。

自主导航移动机器人是一个集环境感知、动态决策和规划、行为控制与执行等多种功能于一体的综合系统。其在运行的过程中,不需要人工的任何干预,所有的运算都是靠自身的设计来完成。由此可以看出,环境特征感知是移动机器人后续决策等操作的基础,也是决定性条件。在移动机器人导航中,对环境的感知主要是对障碍物识别,并计算出障碍物与机器人的距离和角度,将此信息传达给后续的动态决策和路径规划模块,以完成机器人的自主导航。因此检测障碍物并获得障碍物的距离和角度成为移动机器人导航中环境信息的主要特征。

本文平行双目视觉系统的设计可以将前方视野范围 5 米之内的障碍物进行检测。针对基于点特征的 SLAM 导航算法的要求,本文对障碍物的特征进行了分析。在移动机器人导航中,对前方障碍物的信息要求能够满足控制系统进行判断的需求,距离和角度是必不可少的。本文研究的主要内容是对障碍物中心点的信息进行提取。这样做有以下好处。首先,本文针对的障碍物没有具体特征束缚,因此无法对整个障碍物的几何特征点进行确定。本文提出的障碍物中心点特指在视差图中组成障碍物联通域中心小方块的最接近小方块平均距离的点。其次,在移动机器人导航中,需要用到基于点特征的 SLAM 导航算法,该方法需要将障碍物视为一个点进行定位,本文将障碍物作为一个点来进行研究满足 SLAM 的要求。最后通过实验可以发现,将障碍物中心点的信息来作为整个障碍物信息,计算速度较快,可以快速传给 SLAM 以完成地图创建。

#### 1.4 双目视觉关键技术

双目立体视觉模仿人类的视觉立体感知过程,基本原理是从两个视点观察同一场景,以获取不同视角下的图像。通过计算两幅图像中像素点之间的对应关系,根据三角测量原理来获得物体的三维信息。一个完整的双目立体视觉系统一般包括图像的获取、摄像机标定、图像预处理与特征提取、立体匹配、深度信息提取以及三维重建等五个部分<sup>[17]</sup>。下面对平行双目视觉系统按照图像的处理顺序(如



图 1-2 所示) 做一个基本介绍。

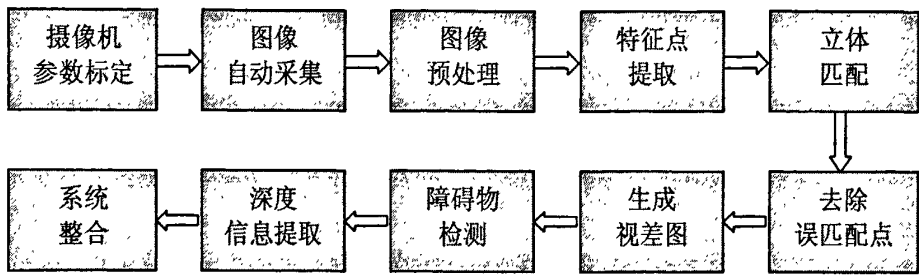


图 1-2 本文根据图像处理的流程图

图像的获取：获取数字图像是进行图像处理和实现计算机视觉的前提条件。在双目立体视觉当中，数字图像的获取所用的硬件设备一般是 CCD 摄像机和视频采集卡。图像采集不但要满足系统的应用要求，而且要考虑视点差异、光照条件、摄像机性能以及景物特点等因素的影响，以利于计算移动机器人导航所用的三维坐标和角度信息<sup>[17]</sup>。

摄像机标定：从二维图像中恢复出物体的三维信息，必须已知空间坐标系中的物体点与它在计算机图像平面上像点之间的对应关系，而这个对应关系是由成像模型和摄像机的位置、属性参数所决定的。确定这些摄像机参数的过程就称为摄像机标定。摄像机标定实质上就是根据成像模型确定由世界坐标系到计算机图像坐标系的变换矩阵。摄像机标定方法可分为两大类，第一类是直接估计摄像机位置、光轴方向、焦距等参数；第二类是通过最小二乘法拟合，确定三维空间点映射为二维图像点的变换矩阵。建立有效适用的摄像机模型，并且通过实验得到模型的精确参数和摄像机自身的参数，除了能够精确地恢复出空间景物的三维信息以外，还有利于提高立体匹配的精确性<sup>[17]</sup>。

图像预处理和特征提取：由光学成像系统生成的二维图像的过程中包含了各种各样的随机噪声和畸变，因此需要对原始图像进行预处理，以改善图像质量和突出有用信息、抑制无用信息。图像预处理的主要目的有两个：一是改善图像的视觉效果，提高图像的视觉清晰度；二是使图像变的更有利于计算机处理，便于在复杂的环境特征信息中提取障碍物特征。图像预处理技术包括图像随机噪声的去除、对比度的增强、边缘特征的加强等。特征提取是为了得到具有实际应用意义的图像匹配特征。由于目前尚没有一种普遍适用的理论可运用于图像特征的提

取,从而导致了立体视觉研究中匹配特征的多样性。良好的匹配特征应具有可区分性、不变性、稳定性以及有效解决歧义匹配的能力。目前,常用的匹配特征主要有区域特征、线特征和点特征等。一般来讲,小尺度特征的数目较多,其所含信息较少,因而在匹配时需要较强的约束准则和匹配策略,以克服歧义匹配和提高运算效率。而大尺度特征含有较丰富的图像信息,在图像中的数目较少,易于得到快速的匹配,但它们的定位精度差,特征提取与描述困难<sup>[17]</sup>。

**立体匹配:**该部分是双目视觉中最重要也是最困难的问题,要解决的是同一空间点在双目视觉两幅图像中像点的对应关系。当空间三维场景被投影为二维图像时,同一景物在不同视点下的计算机图像会有很大的不同,而且场景中的诸多因素,如景物几何形状和物理特性、光照条件、摄像机特性以及噪声干扰和畸变等,都被综合成单一的图像中的灰度值。因此,要准确地对包含了如此之多不利因素的图像进行无歧义的匹配,显然是十分困难的。由于立体匹配涉及的问题太多,至今仍未得到很好的解决,特别是在复杂场景中,如何提高算法去除歧义匹配和增加抗干扰能力,降低实现的复杂度和计算量,都需要更深入的探索和研究。对于任何一种立体匹配方法,其有效性有赖于几个问题的解决,即选择正确的匹配特征,寻找特征之间的本质属性及建立能正确匹配所选特征的稳定算法。立体匹配的研究都围绕着这三方面展开,并已经提出了大量各具特色的匹配方法<sup>[17]</sup>。

**障碍物检测:**机器人实现自主导航,双目视觉系统必须要对机器人移动过程当中遇到的障碍物进行检测。由于机器人是在未知的环境当中移动,单纯的依靠已知模型的障碍物进行检测是行不通的。国内在单纯已知模型方面的研究比较多,有基于形状的,有基于颜色的,但是在未知环境中对未知障碍物的研究目前还处于刚起步阶段,文献资料很少。本人尝试了一种对未知环境当中对障碍物进行识别的方法,实验效果可以满足实际移动机器人的应用,以实现机器人的自主导航。

**深度和角度信息提取:**基于平行双目视觉的环境特征信息提取最终要得到的信息是障碍物的深度和角度信息。已知立体成像模型、完成立体匹配并检测出障碍物之后,深度和角度信息的恢复是比较容易的。重要的是如何提高计算的精确度,其影响因素是多方面的,如摄像机参数标定、图像特征定位的精确程度和立体匹配的准确性等等,因此要提高计算的精度还需要更深入的研究。

## 1.5 论文的研究内容和创新点

### 1.5.1 论文的研究内容

本论文共分五章，章节内容如下安排：

第一章 绪论 主要介绍了基于平行双目视觉环境特征提取相关技术的国内外研究现状、主要技术难点以及论文的组织结构和研究内容。

第二章 成像模型与摄像机标定 介绍了双目视觉系统常用的成像模型、相机标定原理和常用的相机标定方法，并简单介绍了 Intel 公司的视觉开发库 OpenCV。最后给出本文所用的摄像机成像模型，首先通过推导，给出计算公式，并指出当中所要注意的关键问题，并通过设计实验，本章实现了摄像机内外参数的标定，并总结在今后的实际应用当中所应注意的问题以及所采取的方法。

第三章 特征点提取与基于点特征的立体匹配 介绍了常用的特征点提取方法和图像匹配方法，并针对以前的工作，改进了一种基于点特征的双目视觉图像匹配方法。通过采用 Harris 角点的提取算法，首先对双目视觉图像中的特征点进行检测，并改进了 SIFT(Scale Invariant Feature Transform)算法中的特征描述方法在 Harris 特征点描述当中的应用；然后利用欧氏距离对特征点进行匹配；最后提出一种利用匹配点对的斜率来排除错误匹配的方法。实验表明该算法不仅可以达到较高的匹配精度，还可以有效缩短匹配时间，满足自主机器人双目视觉实时性的要求。

第四章 不规则障碍物检测与特征信息提取 通过前一章的图像匹配计算，可以进一步生成视差图。首先介绍了视差图的生成原理，包括其对图像当中三维信息的隐藏；其次介绍了双目视觉中常用的目标检测方法以及这些方法在移动机器人导航当中的不足，并根据移动机器人导航的实际需要，本节提出了一种快速的不规则障碍物检测方法。最后，对整个双目视觉系统进行整合，对应用于机器人导航中的未知环境特征信息中障碍物深度和角度进行提取，在 VC++6.0 平台和实际环境当中进行实验。通过实验，验证了本文所设计的双目视觉系统满足移动机器人导航的应用。

第五章 结论和展望 总结本文的工作，并提出今后需要进一步研究的问题。

### 1.5.2 论文的创新点

本文对基于平行双目视觉的环境特征信息提取方法进行了研究,主要创新点有:

(1) 研究了成像模型和摄像机标定,并根据推导得出的三维坐标公式对平行双目视觉系统重新设计标定实验,根据实验结果确定平行双目视觉系统摄像机内外参数,并提出需要注意的问题以及给出解决办法;

(2) 研究了特征点的提取和匹配方法,采用了 Harris 角点的检测方法,改进了基于点特征的双目视觉立体匹配方法,并提出了一种利用匹配点的斜率来实现快速去除误匹配的算法;

(3) 研究了视差图的生成原理,并根据移动机器人对环境特征信息的要求提出了一种不规则障碍物的检测方法,并通过实际实验验证了该方法的有效性。

## 2 成像模型与摄像机标定

### 2.1 摄像机成像模型

在双目视觉中,摄像机成像几何模型解决的是客观世界场景中的点与视觉系统采集的计算机图像平面上的点之间的对应问题。几何模型中所涉及的参数包括摄像机内部参数和外部参数,内外参数的获得就是摄像机标定。简而言之,摄像机标定就是建立摄像机图像像素位置与场景点位置之间的对应关系,其途径是根据摄像机成像模型,由已知特征点的图像坐标和世界坐标求解摄像机成像模型的参数。因此确定摄像机成像模型是双目视觉当中至关重要的第一步,直接影响到摄像机标定等后续工作。

摄像机成像几何模型常用的有两种,平行投影和透视投影。平行投影是指将中心投影法的投射中心移至无穷远处,则各投射线成为相互平行的直线。透视投影是用中心投影法将形体投射到投影面上,从而获得的一种较为接近视觉效果单面投影图。透视投影符合人们心理习惯,即离视点近的物体大,离视点远的物体小,远到极点即为消失,成为灭点。它的视景体类似于一个顶部和底部都被切除掉的棱锥。本小节根据相关项目中对环境特征信息提取的实际要求,在相机的透视投影模型基础上展开研究,最后给出了本文所采用的双目视觉构造模型。

#### 2.1.1 四个参考坐标系

摄像机成像可分为三个过程:由世界坐标系到摄像机坐标系的转换、由摄像机坐标系到成像平面坐标系的转换和由成像平面坐标系到计算机图像坐标系的转换。为了方便后文定量叙述摄像机成像过程,在这里首先介绍一下这四个参考坐标系<sup>[18]</sup>:

(1) 计算机图像坐标系:摄像机采集的图像以标准电视信号的形式输入计算机,然后经过数模转换器转换为数字图像。每幅数字图像在计算机内以数组形式存储。在灰度图像的模式下,数组中的每一个元素(像素)的值就是图像点的亮度。如图 2-1 所示,在图像上定义直角坐标系  $O_0-uv$ , 原点  $O_0$  位于图像平面的

左上角，每一像素的坐标 $(u, v)$ 分别表示该像素在数组中的列数和行数。所以， $(u, v)$ 是以像素为单位的图像坐标系坐标。

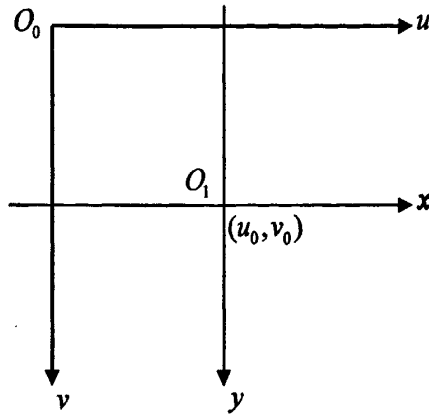


图 2-1 计算机图像坐标系与成像平面坐标系

(2) 成像平面坐标系：由于图像坐标系并没有用实际物理单位表示出该像素在图像中的位置，只表示像素位于图像数组中的列数和行数，因此还需要建立以物理单位(毫米)表示的图像坐标系 $O_1-xy$ ，我们称它为成像平面坐标系，如图 2-1 所示。在 $O_1-xy$ 坐标系中，原点 $O_1$ 位于摄像机光轴与图像平面的交点， $x, y$ 轴分别平行于 $u, v$ 轴。在本论文的以后章节中，如不加特别说明， $(u, v)$ 表示以像素为单位的计算机图像坐标系坐标， $(x, y)$ 表示以毫米为单位的成像平面坐标系坐标。

(3) 摄像机坐标系：以照相机的光心为摄像机坐标系原点 $O_c$ ， $X_c$ 轴和 $Y_c$ 轴分别平行于成像平面坐标系的 $x$ 轴和 $y$ 轴，相机的光轴为 $Z_c$ 轴，它与图像平面垂直，正方向指向正前方，构成了摄像机坐标系 $O_c-X_cY_cZ_c$ ，坐标值用 $(X_c, Y_c, Z_c)$ 表示。光心到成像平面的距离就是摄像机的有效焦距 $f$ 。

(4) 世界坐标系：由于双目摄像机和障碍物等可以放在现实环境中的任何位置，因此还需要在环境中选择一个基准坐标系来描述摄像机的位置，并用它来描述环境中障碍物目标的位置，这个坐标系就叫做世界坐标系 $O-X_wY_wZ_w$ 。在双目视觉中，通常以所观测物体的中心三维坐标系或摄像机坐标系来定义世界坐标系。这是一个假想的坐标系，坐标原点可人为设定，可以根据具体情况来选择，可仅用作一般的参考。

### 2.1.2 针孔模型下平行双目视觉坐标公式推导

摄像机的小孔成像模型如图 2-2 所示：在面 1 刺一个小孔，物体发射和反射的光线通过这个小孔在面 2 上形成一个倒立的像。其中，面 1 被称为焦平面，面 2 称为像平面，针孔为光学中心。根据移动机器人导航对环境信息提取的要求，可以不用考虑投影畸变所带来的细微误差，在这种情况下，该成像系统成为理想透视变换。

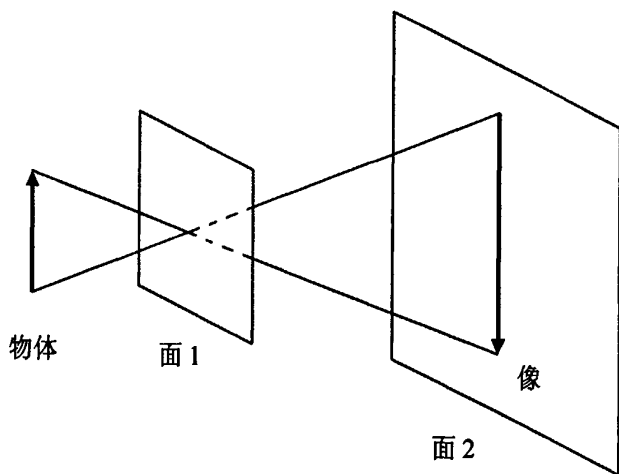


图 2-2 小孔成像模型

由于摄像机在随机器人的移动过程中对周围物体的观察距离有限，加上透视投影符合人类的观察习惯，本文选定透视投影作为研究对象，并以小孔成像的模型展开研究。平行双目视觉的特点是基线与水平轴重合，两个摄像机的光轴是平行的，型号也一样，极点移到无限远处，在采集到的图像中极线是平行的，如图 2-3 所示。这种模型计算简单并且可以满足实际要求。本小节将确定出平行双目视觉系统的系统搭建和从二维计算机图像坐标系坐标到世界坐标系坐标的计算公式，从而给出下一步需要进行标定的摄像机内外参数。

在双目视觉系统的左右摄像机架构部分，本文采用了平行双目视觉系统，这是因为在该系统下左右两摄像机之间的坐标转换比较容易计算，有利于摄像机的标定，而且有利于三维坐标的计算，符合移动机器人导航对环境特征信息的数据要求。在实验阶段，本文的摄像机型号采用的是陕西维视图像的 MV-VD030SC，通过 USB 连接笔记本电脑。摄像头型号采用 computar 的 M0814-MP，双目相机小支架采用的是陕西维视的 MV-BR28 型号。本文架设的平行双目立体视觉系

统实际效果如图 2-4 所示:

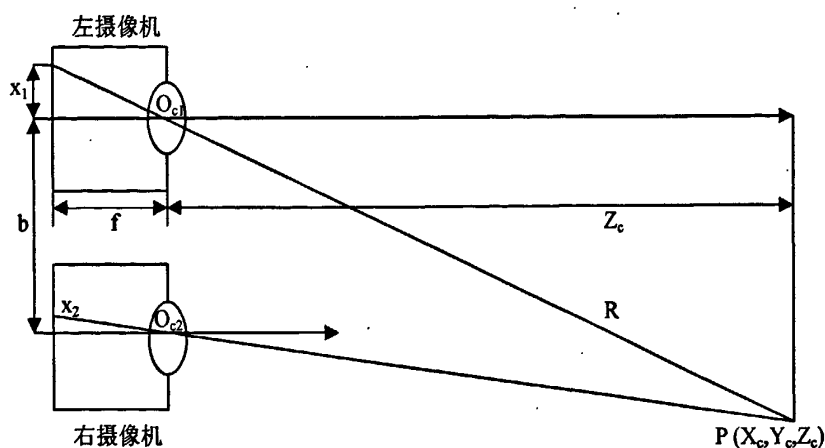


图 2-3 平行双目视觉系统成像模型图

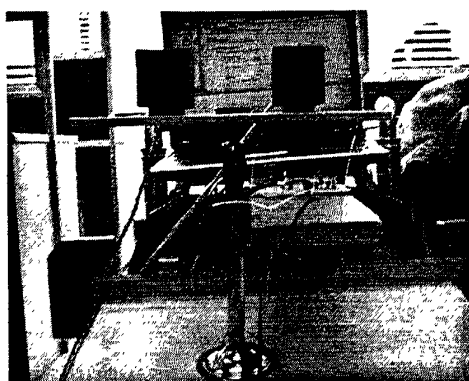


图 2-4 平行双目视觉系统

小孔成像模型下从二维计算机图像坐标系坐标到三维世界坐标系坐标的计算公式推导过程分为以下三步:

(1) 世界坐标系到摄像机坐标系的转换: 由于世界坐标系是一个假想的用来描述摄像机位置的坐标系, 所以我们可以取双目视觉左图像的摄像机坐标系为世界坐标系。在这种情况下, 平行双目视觉系统中世界坐标系到右摄像机坐标系的转换可表示如下:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad \text{式(2-1)}$$

其中  $R$  为旋转矩阵, 由于是平行双目视觉系统, 不存在旋转关系,  $R$  为单位矩阵,



$t$  为平移向量，模为基距  $b$ 。由于是向右平移，所以在计算时应取负号。 $0^T$  为元素为 0 的横向量。

(2) 摄像机坐标系到成像平面坐标系的转换：因为本文采用的摄像机成像模型为小孔成像模型，所以，摄像机坐标系到成像平面坐标系的转换过程符合透视投影，可用齐次坐标表示如下：

$$Z_c \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} -f & 0 & 0 & 0 \\ 0 & -f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \quad \text{式(2-2)}$$

其中， $f$  是摄像机的焦距， $(X_c, Y_c, Z_c)$  是摄像机坐标系空间中任一点  $P$  的三维坐标， $(x, y)$  是  $P$  点在成像平面坐标系中的二维坐标。

(3) 由成像平面坐标系到计算机图像坐标系的转换：

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} 1/d_x & 0 & u_0 \\ 0 & 1/d_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad \text{式(2-3)}$$

其中， $(u_0, v_0)$  是成像平面坐标系原点在计算机图像坐标系中的坐标，往往取为计算机图像的中心点坐标。 $(d_x, d_y)$  分别是成像平面坐标系在  $x$  方向和  $y$  方向相邻像素间的距离。

将式 2-1, 2-2 带入式 2-3 得：

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} -f/d_x & 0 & u_0 & 0 \\ 0 & -f/d_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad \text{式(2-4)}$$

设置等效变换如下：

$$N = \begin{bmatrix} -f/d_x & 0 & u_0 & 0 \\ 0 & -f/d_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad \text{式(2-5)}$$

$$M = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \quad \text{式(2-6)}$$

式 2-4 可简写为:

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = NM \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = Q \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad \text{式(2-7)}$$

其中,  $R$  为旋转矩阵,  $t$  为平移向量,  $N$  为摄像机的内参数矩阵,  $M$  为摄像机的外参数矩阵,  $Q = NM$  为摄像机的投影矩阵。在平行双目视觉系统中, 由于左右摄像机之间是平行移动关系, 而且已经确定左摄像机坐标系为世界坐标系, 因此旋转矩阵  $R$  为单位矩阵, 平移向量  $t$  的模为左右两摄像机之间的基距  $b$ , 因此针对左摄像机和右摄像机, 式 2-4 可以进一步简化如式 2-8 和式 2-9:

$$\begin{cases} z_w \cdot u_1 = \frac{-f \cdot x_w}{d_x} + u_0 \cdot z_w \\ z_w \cdot v_1 = \frac{-f \cdot y_w}{d_y} + v_0 \cdot z_w \end{cases} \quad \text{式(2-8)}$$

$$\begin{cases} z_w \cdot u_2 = \frac{-f \cdot x_w}{d_x} + \frac{f \cdot b}{d_x} + u_0 \cdot z_w \\ z_w \cdot v_2 = \frac{-f \cdot y_w}{d_y} + v_0 \cdot z_w \end{cases} \quad \text{式(2-9)}$$

整理以上两式可以得到世界坐标系和计算机图像坐标系之间的计算公式 2-10:

$$\begin{cases} x_w = \frac{-b \cdot (u_1 - u_0)}{u_2 - u_1} \\ y_w = \frac{-b \cdot d_y \cdot (v_1 - v_0)}{d_x \cdot (u_2 - u_1)} \\ z_w = \frac{b \cdot f}{d_x \cdot (u_2 - u_1)} \end{cases} \quad \text{式(2-10)}$$

由于在计算机显示图像时, 将小孔成像过程中的对应点经过 180 度的反转, 所以需要对最后的公式进行符号纠正, 即, 对  $z_w$  取相反数可得:

$$z_w = \frac{-b \cdot f}{d_x \cdot (u_2 - u_1)} \quad \text{式(2-11)}$$

由此可见,在本文所采用的平行双目视觉系统中,需要标定的摄像机内外参数有摄像机焦距  $f$ , 成像平面坐标系在  $x$  方向和  $y$  方向相邻像素间的距离  $d_x, d_y$ , 成像平面坐标系原点在计算机图像坐标系中的坐标  $(u_0, v_0)$  以及左右两平行摄像机之间的基距  $b$ 。

由于摄像机镜头的制造误差和安装误差等,实际摄像机成像过程中,摄像机镜头光轴与成像平面并不正交,或者计算机图像的  $u$  轴和  $v$  轴并不垂直。这样就需要考虑摄像机的线性参数模型,即在内参数矩阵  $N$  中加入畸变因子  $s$ 。考虑到移动机器人导航中对距离的计算可以保持在 5% 的误差之内,本文没有考虑对畸变因子  $s$  所带来的误差进行弥补。

## 2.2 摄像机标定方法

因为双目视觉可以应用在不同的场合中,所以对于不同的应用背景提出了不同的标定技术。如果视觉系统的任务是物体识别,则物体相对于参考坐标系的位置显得并不重要,重要的是物体特征点之间相对位置的精度。另一种情况下如果视觉系统的任务是物体的定位,则相对于某一参考坐标系的绝对定位精度就显得特别重要了。根据这些特点,在对摄像机进行标定时需要根据双目视觉系统的应用目的来进行实际操作。本文在分析了经典的标定方法,自标定方法后,介绍了张正友标定方法,并根据移动机器人导航对环境特征信息的要求,在下一节采取不同的方法分别对摄像机的内外参数通过实验进行了标定。

### 2.2.1 传统摄像机标定方法

传统的摄像机标定方法是在一定的摄像机成像模型和架构下,基于特定的条件,如形状尺寸已知的标定物,经过图像处理和一系列的数学变换计算方法来求取摄像机成像模型的内部参数和外部参数。从计算思路的角度来分,可以分为四类,即利用摄像机透视变换矩阵的标定方法,利用最优化算法的标定方法, Tsai 两步法和双平面标定法。

### 2.2.1.1 透视变换法

透视变换法采用线性针孔成像模型, 利用投影变换矩阵进行参数标定。如果忽略摄像机镜头的非线性畸变并且把透视变换矩阵中的元素作为未知数, 给定一组三维控制点和对应的图像点坐标, 就可以利用线性方法求解透视变换矩阵中的各个元素。该方法的优点是计算速度快, 能够实现摄像机参数的实时计算。缺点是不考虑镜头畸变, 标定精度受到影响, 并且线性方程中未知参数的个数大于求解的独立摄像机模型参数的个数, 线性方程中未知数不是相互独立的<sup>[19]</sup>。

在针孔成像模型中, 由式 2-12 可知投影矩阵为一  $3 \times 4$  的矩阵, 设为:

$$Q_{cam} = \begin{pmatrix} q_{11} & q_{12} & q_{13} & q_{14} \\ q_{21} & q_{22} & q_{23} & q_{24} \\ q_{31} & q_{32} & q_{33} & q_{34} \end{pmatrix} \quad \text{式(2-12)}$$

式 2-7 可写为:

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} q_{11} & q_{12} & q_{13} & q_{14} \\ q_{21} & q_{22} & q_{23} & q_{24} \\ q_{31} & q_{32} & q_{33} & q_{34} \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad \text{式(2-13)}$$

整理可得:

$$\begin{cases} Z_c u = q_{11} X_w + q_{12} Y_w + q_{13} Z_w + q_{14} \\ Z_c v = q_{21} X_w + q_{22} Y_w + q_{23} Z_w + q_{24} \\ Z_c = q_{31} X_w + q_{32} Y_w + q_{33} Z_w + q_{34} \end{cases} \quad \text{式(2-14)}$$

消去  $Z_c$  得:

$$\begin{cases} q_{11} X_w + q_{12} Y_w + q_{13} Z_w + q_{14} - q_{31} u X_w - q_{32} u Y_w - q_{33} u Z_w = q_{34} u \\ q_{21} X_w + q_{22} Y_w + q_{23} Z_w + q_{24} - q_{31} v X_w - q_{32} v Y_w - q_{33} v Z_w = q_{34} v \end{cases} \quad \text{式(2-15)}$$

取  $q_{34}=1$ , 则共有 11 个未知数, 求六个或以上的目标点可得到 12 个以上的方程组成的超定方程组, 求解该超定方程组, 得到投影矩阵  $Q_{cam}$ , 分解  $Q_{cam}$  获得相机的内外参数。利用透视变换矩阵的摄像机标定方法被广泛应用, 并取得了满意的结果<sup>[19-22]</sup>。

### 2.2.1.2 利用最优化算法的标定方法

利用最优化算法的摄像机标定方法可以假设摄像机的光学成像模型非常复杂,然而由此也带来了很多问题,如摄像机的初始给定值很重要,如果初始值给的不恰当,则很容易通过优化程序得到错误的标定结果;另外优化程序非常耗时,不适用于实时应用场合。本文对直接线性变换方法(DLT)<sup>[23]</sup>进行了研究。

直接线性变换方法是 Abdel-Aziz 和 Karara 首先于 1971 年提出的,具有两种含义,一种含义是直接通过求解一组线性方程得到摄像机的有关参数;另一种含义是求解的过程不排除使用非线性优化算法。直接线性变换方法是对摄影测量学中的传统方法的一种简化,以小孔模型为研究对象,忽略具体的中间成像过程,用一个  $3 \times 4$  阶矩阵表示空间物点与二维像点的直接对应关系。其线性变换矩阵是将像点和物点的成像几何关系在齐次坐标下写成透视投影矩阵的形式,如下式:

$$S \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = P_{3 \times 4} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad \text{式(2-16)}$$

式中各矩阵向量意义同透视变换方法。 $P$  为  $3 \times 4$  矩阵,  $S$  为未知尺度因子。消去  $S$ , 可得到如下方程组:

$$\begin{cases} p_{11}X_w + p_{12}Y_w + p_{13}Z_w + p_{14} - p_{31}uX_w - p_{32}uY_w - p_{33}uZ_w - p_{34}u = 0 \\ p_{21}X_w + p_{22}Y_w + p_{23}Z_w + p_{24} - p_{31}vX_w - p_{32}vY_w - p_{33}vZ_w - p_{34}v = 0 \end{cases} \quad \text{式(2-17)}$$

当已知  $N$  个空间点的三维坐标和对应图像上的点时,可以得到一个含有  $2N$  个方程的方程组:

$$AL = 0 \quad \text{式(2-18)}$$

其中  $A$  为  $2N \times 12$  的矩阵,  $L$  为透视投影矩阵元素组成的向量  $(p_{11}, p_{12}, p_{13}, p_{14}, p_{21}, p_{22}, p_{23}, p_{24}, p_{31}, p_{32}, p_{33}, p_{34})^T$ 。该标定方法的任务就是寻找合适的  $L$ , 使得  $\|AL\|$  为最小。由式 2-12 可见,  $P$  矩阵乘以任意不为 0 的常数并不影响  $(X_w, Y_w, Z_w, 1)$  与对应的图像点  $(u, v, 1)$  的关系, 因此, 给出约束  $p_{34} = 1$ , 用最小二乘法求解式 2-13 得  $L' = (C^T C)^{-1} C^T B$ , 其中  $L'$  为  $L$  的前 11 个元素组成的向量,

$C$  为  $A$  的前 11 列组成的矩阵,  $B$  为  $A$  第 12 列组成的向量<sup>[23]</sup>。

### 2.2.1.3 Tsai 两步法

透视变换法或者直接线性变换法可以利用线性方法求得摄像机的参数,但其缺点是没有考虑镜头的非线性畸变、精度不高。如果先利用透视变换方法或者直接线性变换方法求解相机参数,再以求得的参数为初始值,考虑畸变因素,并利用最优化算法进一步提高定标精度,这就形成了所谓的两步法<sup>[24]</sup>。80 年代中期 Tsai 提出的典型两步法中,CCD 阵列中的感光元横向间距和纵向间距被认为是摄像机厂家提供的已知值,他所假设的摄像机内部和外部参数分别是: $f$  为等效焦距; $k$  为镜头畸变参数; $S_x$  为非确定性尺度因子; $(C_x, C_y)$  为图像中心或主点; $R, T$  为三维空间坐标系与摄像机坐标系之间的旋转矩阵和平移向量<sup>[24]</sup>。

Tsai 两步法只考虑了径向畸变。第一步是利用最小二乘法求解超定线性方程,给出外部参数;第二步求解内部参数,如果相机无透镜畸变,则可由一个超定方程解出。如果存在一个以二次多项式近似的径向畸变,则可由一个三变量的优化搜索求解。Tsai 两步法的好处是它所使用的大部分方程是线性方程,从而降低了参数求解的复杂性,因此其标定过程快捷、准确。其算法校准分为两个阶段。

第一阶段:求解旋转矩阵  $R$ 、平移向量的  $T_x$ 、 $T_y$  以及图像尺度因子  $S_x$ 。

第二阶段:计算有效焦距  $f$ 、透镜畸变系数  $k$  和平移分量  $T_z$ 。

文献[25]同时考虑了径向畸变和切向畸变,是对基于径向排列约束的两步标定法的进一步发展。

### 2.2.1.4 双平面标定方法

在传统摄像机标定的另一方向,即寻找合理的摄像机模型,研究人员也做了深入的探讨: Martins 等首先提出了双平面模型<sup>[26]</sup>;马颂德和魏国庆在利用双平面模型标定摄像机参数方面做了大量的工作<sup>[27~29]</sup>。

双平面模型与小孔成像模型的基本区别在于,双平面模型不像小孔成像模型那样要求所有投射到面像平面上的光线必须经过光心。给定成像平面上的任意一个图像点,就能够求出两个标定平面上的相应点,从而确定了投射到成像平面上

产生该图像点的光线。该法的优点是可以利用线性方法就能解有关参数，缺点是要求解大量的未知参数，存在过分参数化的倾向。

### 2.2.2 摄像机自标定方法

与传统的摄像机标定方法相比，摄像机的自标定技术不需要已知参照物，而仅仅利用摄像机在运动过程中周围环境的图像及图像之间的对应关系进行标定。摄像机标定方法在很多情况下，由于存在经常性调整摄像机的要求，此时设置已知的标定参照物已经不现实，这时就需要一种不依赖标定参照物的所谓摄像机自标定方法，而这种标定方法成为目前摄像机标定研究的重点。

目前摄像机自标定技术主要有以下四种方法：分别是直接求解 Kruppa 方程的自标定方法；分层逐步标定方法；基于绝对二次曲线的自标定方法；Pollefeys 的模约束方法。摄像机自标定方法计算量比较大，因为大都需要解非线性方程组，另外一个问题是这些方法一般是基于点或者线的方法，匹配问题不容易实现。因此针对一般的摄像机自标定方法存在鲁棒性差、计算量大的缺点，Hartley 提出了一种新的摄像机自标定方法。在 Hartley 方法中的摄像机没有平移运动，而是只绕轴心旋转，在不同的方位拍摄三幅以上的图像，然后通过图像之间的点对应，除特殊情况外一般都可以顺利完成摄像机的自标定。这种方法极大地降低了算法的复杂性，同时标定的结果也满足要求。

### 2.2.3 张正友标定方法

张正友的平面标定方法是介于传统标定方法和自标定方法之间的一种方法，也是最近研究比较热门的标定算法之一。它既避免了传统标定方法中操作繁琐和对设备要求高等缺点，又较自标定方法精度较高。该方法只需要摄像机对同一标定板在不同方向进行多次拍照，而标定板的运动方式在标定过程中不必知道，然后通过对图像信息的提取计算来得到最终的标定结果。详细的算法过程请参照文献[30]，下一小节将运用张正友标定方法对平行双目视觉系统的摄像机内参数进行标定。

## 2.3 摄像机参数标定实验

由于本文在摄像机标定过程中用到了 OpenCV 开发库, 在这里首先简单介绍一下。OpenCV 由一系列 C 函数和少量 C++ 类构成, 是 Intel 开源计算机视觉库, 它实现了图像处理和计算机视觉方面的很多通用算法。在实际应用中通过在开发平台中进行配置, 可以直接调用其中的函数来完成开发需求。

### 2.3.1 平行摄像机视觉系统内参数标定及实验

在本文所论述的环境特征信息中, 所采用的双目视觉系统架构是平行双目视觉系统。该系统的特点左右两个摄像机的光轴处于平行状态, 相隔一定的距离。在这种情况下, 两个摄像机坐标系之间只存在简单的平移关系。如果将其中任意一个摄像机坐标系作为世界坐标系, 则世界坐标系和摄像机坐标系之间的转换关系也只存在简单的平移关系。这样做的好处在于可以简化坐标系之间的转换, 从而减少计算误差, 并且可以满足移动机器人对环境特征信息的要求, 即对障碍物三维坐标和角度的计算。

根据前面小孔成像模型的分析可知, 本文所涉及到的摄像机内部参数有以下五个, 分别是:

$(u_0, v_0)$ : 成像平面坐标系原点在计算机图像坐标系中的坐标, 其为摄像机光轴与摄像机成像平面的交点。在本文所采用摄像机中及在移动机器人导航允许的误差范围内, 可取计算机图像坐标系的中心点坐标来表示  $(u_0, v_0)$ 。

摄像机焦距  $f$ : 本文采用的是定焦摄像机, 摄像机的焦距  $f$  已经给出, 为 8mm。

$d_x, d_y$ : 为摄像机成像平面相邻像素点的距离, 同一型号的摄像机有时候也会因为制造工艺的差别, 导致此值不相同。一般摄像机不会提供相关数据, 本文所做的实验就是为了得到精确的相关数值。由于在计算距离的过程中,  $d_x, d_y$  对结果的影响很大, 所以该值显得特别重要。

本文根据张正友标定法制定了一张标定板, 并从不同的方向和距离分别进行拍摄。在张正友标定法中, 计算出的结果包括了摄像机内参数的方方面面, 由于我们已经得知摄像机的定焦距并且将计算机图像坐标系的中心点作为成像平面



坐标系的原点, 因此可以利用这两方面的约束来判定  $d_x, d_y$  的精确度。为了保证计算数据的质量, 本文在距离摄像机 1.3m 处和 3.6m 处分别拍摄了 23 副图像进行计算, 每幅图像的大小都是 640\*480。在计算结果当中, 对比两组图像所得的实验数据, 选取了计算误差较小的数据作为最后的标定值。为了节省篇幅, 本文仅列出所用每组图像当中的前 8 副为代表, 分别如图 2-5 和图 2-6 所示。

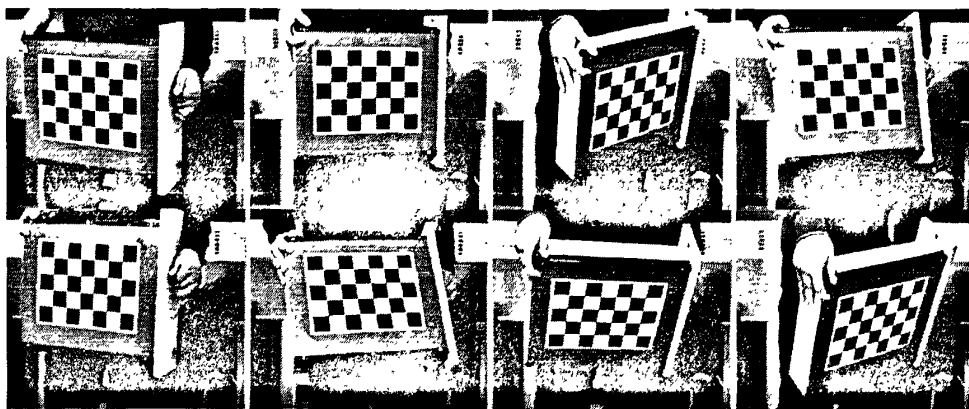


图 2-5 1.3m 处所测得的标定图像

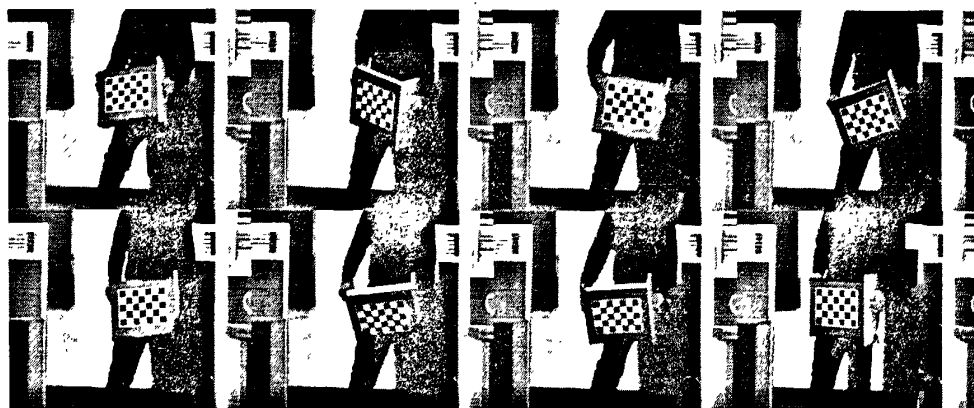


图 2-6 3.6m 处所测得的标定图像

根据张正友标定算法, 本文在 VC++ 6.0 中编写程序对拍摄的两组图像进行处理, 处理的数据如表 2-1 所示:

根据表 2-1 中的数据对比发现, 在 3.6m 处拍摄图像得出的数据的平均误差明显高于在 1.3m 处拍摄图像得出的数据的平均误差, 并且根据中心点的坐标位置可以发现, 在 1.3m 处的计算结果更接近 640\*480 图像中心点的坐标值。因此本文选用了在 1.3m 处拍摄图像所得出的计算结果作为  $d_x, d_y$  的实验值, 即  $d_x = d_y = 0.0055\text{mm}$ 。根据后面进一步实验所得的数据表明, 该实验数据符合实际应用的要求。

表 2-1 张正友标定法图像处理误差数据

1.3m 处图像	平均误差	3.6m 处图像	平均误差
1	0.218861 像素	1	0.26535 像素
2	0.35516 像素	2	3.51591 像素
3	0.266857 像素	3	0.216415 像素
4	0.276726 像素	4	0.187679 像素
5	0.295859 像素	5	0.253675 像素
6	0.313308 像素	6	0.235424 像素
7	0.388432 像素	7	0.217981 像素
8	0.326873 像素	8	0.301926 像素
9	0.217186 像素	9	0.316815 像素
10	0.350236 像素	10	0.171165 像素
11	0.32253 像素	11	0.497972 像素
12	0.404645 像素	12	0.509836 像素
13	0.398456 像素	13	0.486552 像素
14	0.26612 像素	14	0.34217 像素
15	0.375952 像素	15	0.527715 像素
16	0.139742 像素	16	3.17223 像素
17	0.264393 像素	17	0.261158 像素
18	0.359845 像素	18	0.297265 像素
19	0.238164 像素	19	0.30515 像素
20	0.228314 像素	20	0.293716 像素
21	0.284585 像素	21	0.189901 像素
22	0.211443 像素	22	0.315657 像素
23	0.293673 像素	23	0.26535 像素
总体平均误差	0.295807 像素	总体平均误差	0.571247 像素
中心点坐标	(315.235, 238.261)	中心点坐标	(528.439, 130.773)
$(d_x, d_y)$	(0.0055mm,0.0055mm)	$(d_x, d_y)$	(0.0095mm,0.0095mm)

### 2.3.2 平行摄像机视觉系统外参数标定及实验

由前文的讨论可知，针对平行双目视觉系统，摄像机外参数只有两摄像机之间的基距 $b$ ，本节主要目的就是通过设计实验来测定两摄像机之间的基距 $b$ 。

本小组通过实验发现,由不同型号的摄像机组成的平行视觉系统对基距 $b$ 有着不同的要求,并且采取不同的基距 $b$ 对实验结果的误差也会带来较大的影响,因此对基线 $b$ 的测量不是简单的测量两摄像机之间的距离,必须通过实验对比,得出误差最小的基距来作为实际应用。

本小节设计的实验步骤如下:

(1) 确定平行双目视觉系统:为了保证两摄像机的光轴尽量保持平行,本小组首先对 1500mm 位置的标定板进行测量,同时计算标定板中特征点离摄像机的距离和高度。在测量和计算的过程中,根据距离不断校正两摄像机之间的相对位置,根据高度值校正摄像机支架的位置最终使得测量值和计算值达到 10mm 的误差范围之内。最后固定两个摄像机的位置。这样做的目的是可以尽最大努力确定两摄像机光轴保持平行,并且使光轴所在的平面与地面保持平行。

(2) 由于不同的基距 $b$ 对双目视觉的测量结果会产生较大的影响,为了使基距 $b$ 带来的误差达到最小,本文在基距 $b$ 取不同的数值时分别进行了测量。由于需要对测量结果与实际结果进行对比,本文采取如下的方法:首先用直尺测量出从摄像机到被测物体(本文用一箱子代替,箱子表面有花纹)的真实距离,以毫米为单位;然后调用后续的特征点匹配方法,找出采集的图像对中的匹配点,并且计算出匹配点之间的视差值;最后根据得到的视差值计算该匹配点的实际三维坐标,从而计算其距离,用计算值和测量值进行比较得出误差。移动被测物体,分别在 1000mm, 2200mm, 5200mm 处进行测量。

(3) 变换基距 $b$ 的值。本文按照从小到大的顺序对 $b$ 分别取了 80mm, 100mm, 130mm, 150mm, 200mm, 分别进行试验。匹配点的三维坐标计算公式采用的是式 2-10 和式 2-11。在计算过程中,取 $f=8\text{mm}$ ,  $d_x=d_y=0.0055\text{mm}$ , 实验数据结果如表 2-2 所示:

由表 2-2 的数据对比发现,在基距 $b$ 取 130mm 的时候,测量距离从 1000mm 到 5200mm 变换的范围内,误差百分比始终维持在 5%内。本项目移动机器人导航要求双目视觉的测量范围维持在 5000mm 内,并且误差要达到 5%内。由此可见基距 $b=130\text{mm}$ 满足移动机器人导航对环境特征信息的要求。最终通过本次实验,确定了基距 $b=130\text{mm}$ 作为本文摄像机外参数标定的结果。

表 2-2 平行摄像机视觉系统外参数标定实验数据

基距 $b$	视差值 $\Delta x$	测量距离 $l_{\text{测}}$	计算距离 $l_{\text{计}}$	误差百分比
80mm	10	5200mm	11636mm	124%
80mm	49	2200mm	2378mm	8.1%
80mm	119	1000mm	978mm	2.2%
100mm	25	5200mm	5818mm	11.9%
100mm	62	2200mm	2346mm	6.6%
100mm	146	1000mm	996mm	0.4%
130mm	35	5200mm	5402mm	3.4%
130mm	84	2200mm	2251mm	2.3%
130mm	197	1000mm	957.4mm	4.2%
150mm	47	5200mm	4642mm	10.7%
150mm	106	2200mm	2058mm	6.5%
150mm	228	1000mm	956mm	4.4%
200mm	73	5200mm	3955mm	23.4%
200mm	149	2200mm	1952mm	11.3%
200mm	314	1000mm	926mm	7.4%

2.4 本章小结

要对障碍物进行三维坐标计算，首先需要对摄像机进行标定。在标定之前，需要确定双目视觉系统的架构。本文采用的是平行双目立体视觉架构和小孔成像模型对问题进行分析，包括四个坐标系之间的坐标转换和几种常用的标定方法。在前文论述的基础上，根据移动机器人导航对环境特征信息的要求，本文根据实际情况自行设计实验，对摄像机的内外参数进行了标定，包括 $(u_0, v_0)$ ，焦距 $f$ ，摄像机成像平面相邻像素点的横向纵向距离 $d_x, d_y$ 以及基距 $b$ 。通过对比实验结果，最终确定摄像机的内外参数，而且测试结果达到了应用要求。下一章将对双目视觉系统的立体匹配重要部分进行讨论。

### 3 特征点提取与基于点特征的立体匹配

基于图像特征的立体匹配是双目视觉应用中最常用的方法之一。图像特征多种多样,包括像素灰度特征、色彩特征、区域特征、纹理特征、角点特征等。由于图像的特征点比整幅图像的总像素点要少很多,因此可以大大减少匹配过程的计算量和时间;另外,特征点的匹配特征描述符对位置的变化比较敏感,因此可以提高匹配的精确度;最后特征点的提取可以减少各种噪声的影响,对灰度变化、遮挡以及图像形变都有很好的适应能力,因此本文在特征点的提取和基于特征点的匹配方面做了深入研究。

基于点特征的立体匹配是指寻找双目视觉左右两幅图像中的对应点问题。特征点提取方法的好坏直接影响到立体匹配的准确度,而立体匹配结果的准确率又直接决定了障碍物判别成功以及距离计算的准确性。在双目视觉导航中,障碍物判别以及障碍物的距离计算都要用到立体匹配的结果,因此立体匹配在整个系统中占据重要的地位。本文在分析了常用的立体匹配方法的实时性及结果的基础上,根据移动机器人导航对环境特征信息的实际要求,改进了文献[31]的匹配方法,并提出了一种快速的排除错误匹配点的方法。根据实验结果,改进后的匹配方法更加符合移动机器人导航的要求。

#### 3.1 特征点提取

基于特征点的立体匹配方法是建立在提取到正确的特征点的基础上的。也就是说,在双目立体视觉中,首先需要对左右两幅图像进行扫描,以便可以提取到能够进行匹配的特征点。特征点的提取可以大大减少匹配量,而对特征点采取恰当的特征描述符,可以大大提高匹配的正确率。综合两方面来研究,提取到合适的特征点数量并采取恰当的描述符,就可以提高后续立体匹配的实时性和正确率。特征点一般出现在物体图像中灰度急剧变化的地方,因此寻找灰度变化较大的点就能够找到特征点。常用的方法有一次微分、二次微分,另外还有一些其他方法,但实质一般都是对图像进行差分处理。

特征点提取的主要目的是确定特征点的位置和尺度等参数。Moravec 算子<sup>[32]</sup>

重点是在四个方向上, 选择了具有最大或者最小灰度方差的点作为特征点; SUSAN 算子<sup>[33]</sup>对角点的检测比对边缘检测的效果好, 适用于基于特征点匹配的图像配准, 并且该算子不需要梯度运算, 提高了算法的效率, 另外该算子具有积分特性, 提高了该算法的抗噪性和计算速度; Trajkovic 算子<sup>[34]</sup>的目标是只要最小的计算量就达到现有主流提取算子的重复性和定位准确性。最小计算量使得该算子非常适用于实时应用, 然而, 它并不是具有旋转不变性的, 对噪声敏感, 因此精确度比较低; Harris 算子<sup>[35]</sup>是一种基于信号的点特征提取算子。受信号处理中自相关函数的启发, 给出与自相关函数相联系的矩阵  $M$ 。 $M$  矩阵的特征值是自相关函数的一阶曲率, 如果两个曲率值都高, 那么就认为该点是特征点。

本节首先分析了以上几种常用算子的优缺点, 然后根据移动机器人导航对匹配点的数量, 位置和实时性的要求, 提出了本文所采用的特征点提取方法, 并在下一节给出了本文改进之后的特征点描述符。

### 3.1.1 Moravec 算子

Moravec 算子计算图像中各像素沿不同方向的灰度变化的平均值, 将特征点响应函数选定为灰度变化平均值的最小值, 最后定义特征点为在一定范围内具有最大角点响应的像素。该算子定义在各个方向上灰度值变化的点为特征点, 这是特征点的一个例子, 因此 Moravec 算子被认作是特征点提取算子。该算子的具体实现方法可参照文献[32]总结如下 4 步:

1. 针对每个像素点  $(x, y)$  灰度值, 计算出移动  $(u, v)$  后的灰度变化值:

$$V_{u,v}(x, y) = \sum_{\forall a, b \in \text{窗口}} (I(x+u+a, y+v+b) - I(x+a, y+b))^2,$$

$(1, 0), (1, 1), (0, 1), (-1, 1), (-1, 0), (-1, -1), (0, -1), (1, -1)$  为位移  $(u, v)$  的变化范围;

2. 通过计算每个像素点  $(x, y)$  的特征点化值  $C(x, y)$  构造特征点化映射

$$C(x, y) = \min(V_{u,v}(x, y));$$

3. 阈值映射设置所有  $C(x, y)$  在零到阈值  $T$  之间;

4. 执行抑制局部非发现局部最大值, 所有在特征点化映射值非零的像素点

都视为特征点。

在以上各式中, 本文用  $I(x, y)$  表示像素点  $(x, y)$  的灰度值。输入为灰度图像, 窗口的大小和阈值  $T$ ; 输出为提取特征点的位置。

Moravec 特征点的提取算子是一个相对简单的算法, 现在认为已经过时了。它的优点是计算相当有效, 因此在某些对精度要求低并且实时性要求比较高的场合有他的用武之地。缺点是它不具备旋转不变性, 因为它的反应是各向异性的。另外, 该算子容易把沿着边缘的点和孤立的点作为特征点加以提取, 所以该算子对噪音敏感<sup>[32]</sup>。

### 3.1.2 SUSAN 算子

SUSAN 算子是由 Smith 和 Brady 提出的一种与上述 Moravec 算子完全不同的特征点检测算法。该算子无需计算图像差分, 主要是基于亮度对比而实现的, 因此受噪声的影响较少。由于其计算速度比较快, 在实时性要求比较高的领域得到广泛的应用。

为了说明 SUSAN 算子的思想, 本文参考文献[33], 总结如下。如图 3-1 所示, a,b,c,d,e,f 为六个圆形模板。窗口的中心已经标出为模板的核, 边缘即为模板的边缘。六个字母分别代表了模板在整幅图像中的不同位置。在模板中所有与模板核具有相同或者相似灰度的像素区域称为 USAN(Univalue Segment Assimilating Nucleus)。根据图 3-1 可以看出, 在灰度平坦的区域, USAN 区域值越大, 越靠近边缘, USAN 区域就越小, USAN 区域包含了重要的图形结构信息。当模板核位于暗区边缘时, USAN 的区域为最大值的一半; 当模板靠近角点时, USAN 区域进一步减小为最大值的四分之一。因此可以通过 USAN 区域的大小来对图像中的一些重要二维信息做出判断。在图像中搜索 USAN 最小的点(小于一定的阈值), 即搜索最小化同化核分割相同值, 就是搜索图像边缘点或是图像的角点。该算子并没有用到微分运算, 因此这种算法具有较好的抗噪性<sup>[36]</sup>。具体的算法实现如下描述:

第一步需要构造一个圆形的模板。在数字图像中, 圆可以用点阵的形式来近似表示。在大多数情况下, 一个半径大小为三或四个像素的圆可以用 37 点点阵

的形式来近似代替,当然也可以用其它形状的模板,如矩形或者菱形等。37 点阵各行的像素点分别为{3, 5, 7, 7, 7, 5, 3}。

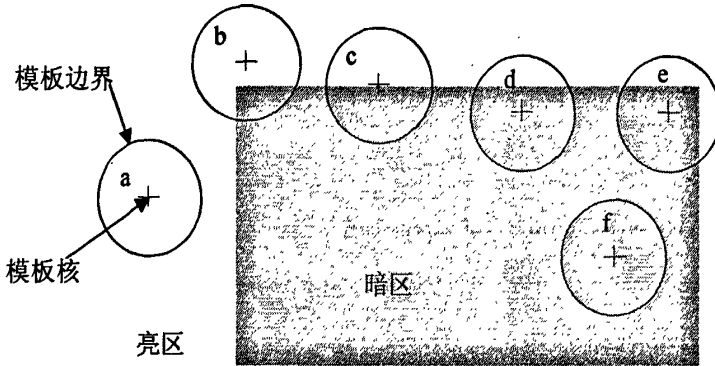


图 3-1 圆形模板在图像中的不同位置

构造了圆形模板之后,需要利用圆形模板来遍历图像中的每一个点。这时,可以采用式 3-1 来判断模板核灰度与模板所包含的区域内的点的相似程度。

$$C(\vec{r}, \vec{r}_0) = \begin{cases} 1 & \text{if } |I(\vec{r}) - I(\vec{r}_0)| \leq thr \\ 0 & \text{if } |I(\vec{r}) - I(\vec{r}_0)| > thr \end{cases} \quad \text{式(3-1)}$$

其中  $\vec{r}_0$  是模板核在灰度图像中的坐标,  $\vec{r}$  为模板区域内其它点的坐标。  $I(\vec{r}_0)$  为  $\vec{r}_0$  点的灰度值,  $I(\vec{r})$  为  $\vec{r}$  点的灰度值,  $thr$  为阈值,该值确定了两个相似点的最大差异。因此,模板区域 USAN 的值可以由式 3-2 计算得出,将下式计算得出的  $n$  与一固定阈值进行比较,就可以得出 SUSAN 算子对图像边缘的响应值,如式 3-3:

$$n(\vec{r}_0) = \sum_{\vec{r}} C(\vec{r}, \vec{r}_0) \quad \text{式(3-2)}$$

$$R(\vec{r}_0) = \begin{cases} g - n(\vec{r}_0) & \text{if } n(\vec{r}_0) < g \\ 0 & \text{otherwise} \end{cases} \quad \text{式(3-3)}$$

其中  $g = 3n_{\max}/4$  为模板区域中像素点的总个数减一。在一般情况下,模板区域的面积远远小于这个物体图像的面积,所以模板可以完全被包围在物体中,因此 USAN 值可以取得极大。为了取得更加稳定并且更加精确的结果,可以将式 3-1 改成如下式 3-4 形式,其中  $t$  为一般设为 75 的参数。这样就提高了稳定性,使阈值附近的值可以接近平缓的影响函数取值。



$$C(\vec{r}, \vec{r}_0) = \begin{cases} \exp\left\{-\left[\frac{I(\vec{r}) - I(\vec{r}_0)}{I(\vec{r}) + I(\vec{r}_0)}\right]^6\right\} & \text{if } |I(\vec{r}) - I(\vec{r}_0)| \leq thr \\ 0 & \text{if } |I(\vec{r}) - I(\vec{r}_0)| > thr \end{cases} \quad \text{式(3-4)}$$

提取图像的特征点时, 只需将 SUSAN 算法改变一个阈值, 即将阈值改为 8。需要考虑到噪声的影响时, 可以将其改写为  $g = n_{\max} / 2$ 。若边缘点在单像素宽度的直线上时, 该点的 USAN 值会很小, 很可能小于  $g = n_{\max} / 2$ 。要排除此类错误情况, 可以将阈值设定一个最小值。当 USAN 值满足在最大值和最小值之间时, 才将该点视为特征点<sup>[33]</sup>。

由以上分析可以得出, 因为 SUSAN 算子不涉及积分运算, 所以运算效率比较高, 对局部噪声也不敏感, 抗噪声能力也较强; 缺点在于定位不够准确, 在弱边缘处容易错误判别特征点。

### 3.1.3 Trajkovic 算子

Trajkovic 算子的目标是只需要最小的计算量就能够达到现有主流提取算子的重复性率和定位准确程度, 该算子是由 Trajkovic 和 Hedley 在 1998 年提出。经验表明, Trajkovic 算子比 SUSAN 算子的计算效率至少快了 3 倍。该算子定义特征点就是在所有方向上灰度值变化比较大的点。它定义角点化测量与 Moravec 算子一样, 是在所有方向上灰度变化的最小值。Trajkovic 在执行性能上超越了 Moravec 算子, 因为该算子采用了交互像素近似法来估计各个方向上的灰度变化。通过使用多栅格逼近, 该算子倾向于首先对初始图像的低分辨率版本进行特征点定位, 因此执行和计算需求都得到了改善。参照文献[34], 本文对该算子的特征点提取步骤归纳如下:

1. 针对低分辨率初始图像的每个像素点  $(x, y)$  计算简化的角点化测量值  $C_{\text{简化}}(x, y)$ , 并标记角点化测量值  $C_{\text{简化}}(x, y) \geq T_1$  的点  $(x, y)$  为潜在的特征点, 简化测量值可用式 3-5 计算得出:

$$C_{\text{简化}}(x, y) = \min(r_A, r_B) \quad \text{式(3-5)}$$

其中  $r_A = (I_A - I_c)^2 + (I_{A'} - I_c)^2$ ,  $r_B = (I_B - I_c)^2 + (I_{B'} - I_c)^2$ ;

2. 初始化角点化映射  $M$ ;

3. 对步骤 1 中找到的每一个潜在特征点, 计算出它在初始输入图像中的位

置  $(x', y')$  和简化的角点化测量值  $C_{\text{简化}}(x', y')$ 。如果  $C_{\text{简化}}(x', y') \leq T_2$ ，则此点不是特征点，角点化映射  $M$  设为 0。否则执行步骤 4；

4. 根据式 3-6 计算交互逼近角点化测量：

$$C_{\text{交互}}(x, y) = \begin{cases} C - \frac{B^2}{A} & B < 0, (A + B) > 0 \\ C_{\text{简化}}(x, y) & \text{其它} \end{cases} \quad \text{式(3-6)}$$

其中  $C = r_A, B = \min(B_1, B_2), A = r_B - r_A - 2B$ ,

$$B_1 = (I_B - I_A)(I_A - I_C) + (I_{B'} - I_{A'})(I_{A'} - I_C),$$

$B_2 = (I_B - I_{A'})(I_{A'} - I_C) + (I_{B'} - I_A)(I_A - I_C)$ ，如果  $C_{\text{交互}}(x', y') \leq T_2$ ，则角点化  $M$  设为 0，否则将  $C_{\text{交互}}(x', y')$  赋给  $M(x', y')$ ；

5. 对  $M$  执行非最大值抑制从而找出局部的最大值。最后所有在角点化映射  $M$  不为零的点均可认为是特征点。

最小计算量需求使得 Trajkovic 算子非常适用于实时应用，然而缺点是该算子不具有旋转不变性，对噪声敏感，对斜边缘的反应异常，检测效率会比其他算子较差<sup>[34]</sup>。

### 3.1.4 Harris 算子

Harris 算子是在 1988 年由 C.Harris 和 M.J.Setphens 提出的一种基于信号的特征点提取算子。该算子受到信号处理中自相关函数的启发，给出与自相关函数相联系的矩阵  $M$ 。 $M$  矩阵的特征值是自相关函数的一阶曲率，如果两个曲率值都高，那么就认为该点是特征点。该算子适合矩阵运算并且计算简单，具有提取的特征点均匀并且可以定量提取特征点等优点，适合在移动机器人导航中应用到整个视野的障碍物检测和三维坐标计算。根据文献[35]，本文对 Harris 算子的特征点检测步骤总结如下：Harris 算子可通过式 3-7 来确定：

$$M = \begin{bmatrix} \left( \frac{\partial I}{\partial x} \cdot \frac{\partial I}{\partial x} \right) \otimes e^{-(u^2+v^2)/2\sigma^2} & \left( \frac{\partial I}{\partial x} \cdot \frac{\partial I}{\partial y} \right) \otimes e^{-(u^2+v^2)/2\sigma^2} \\ \left( \frac{\partial I}{\partial x} \cdot \frac{\partial I}{\partial y} \right) \otimes e^{-(u^2+v^2)/2\sigma^2} & \left( \frac{\partial I}{\partial y} \cdot \frac{\partial I}{\partial y} \right) \otimes e^{-(u^2+v^2)/2\sigma^2} \end{bmatrix} \quad \text{式(3-7)}$$

其中， $I$  为图像， $\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}$  为图像在  $x$  和  $y$  方向的梯度图， $e^{-(u^2+v^2)/2\sigma^2}$  为高斯

滤波器。  $M$  为一个  $2 \times 2$  的对称矩阵，可以改下如式 3-8：

$$M = \begin{bmatrix} A & C \\ C & B \end{bmatrix} \quad \text{式(3-8)}$$

特征点与  $M$  矩阵密切相关。设定  $\alpha, \beta$  是  $M$  矩阵的两个特征值，则局部自相关函数的主要曲率与  $\alpha, \beta$  成比例，可以被用来描述  $M$  矩阵的旋转不变性。此时有以下三种情况需要考虑：

1. 假如两个特征值都是高的，导致局部的自相关函数是突变山峰形状，那么在任意方向的偏移都将增加该区域变化值，所以该点被认为是一个特征点；

2. 假如两个特征值都是小的，导致局部的自相关函数是平的，那么图像中的变化区域成为近似不变的亮度；

3. 假如一个特征值是低的，而另一个特征值是高的，从而导致局部自相关函数呈现出了山脊的形状，那么显然这是一条边。

为了加快计算速度，避免求矩阵的特征值，做出如下计算

$$Tr(M) = \alpha + \beta = A + B \quad \text{式(3-9)}$$

$$Det(M) = \alpha\beta = AB - C^2 \quad \text{式(3-10)}$$

$$R = Det(M) - kTr(M)^2 \quad \text{式(3-11)}$$

其中矩阵  $M$  的特征值为  $\alpha, \beta$ ， $Tr(M)$  为矩阵  $M$  的迹， $Det(M)$  为矩阵  $M$  的行列式， $k$  一般取为 0.04。考虑到三种情况，上式中的特征点判别准则  $R$  在不变的区域是一个很小的值，在特征点区域为一个正值，在边的区域为一个负值<sup>[35]</sup>。

在实际应用中，可以计算感兴趣区域中心点的  $R$  值，当大于某一个给定的门限值时，则这个点就判作特征点。可依次从以每个像素为中心的  $3 \times 3$  的窗口中提取  $R$  值，如果中心点像素的  $R$  值大于给定的门限值，则该点就被判作特征点。由于 Harris 算子求的是梯度运算，所以不受图像亮度改变的影响。再者 Harris 算子提取的特征点分布均匀而且合理，可以定量提取特征点，并且旋转后只是方向发生了改变，所以具有旋转不变性。由于具有以上特性，本文采取 Harris 算法提取特征点，使其提取的特征点个数和分布范围符合移动机器人导航在障碍物识别和三维坐标运算方面的要求。

### 3.2 特征点描述符

### 3.2.1 特征点描述符概述

提取适当的特征点之后,在对特征点进行匹配之前,需要对特征点进行特征描述。好的特征点描述符不但可以提高特征点匹配的速度,并且可以提高匹配正确率。特征描述符的关键在于要对不同区域具有较好的区分性,双目视觉中对这方面的研究一直是属于热门问题。

在对特征点进行特征描述的方法中,直接计算两个特征点邻域像素值的相似度来作比较是最直观的方法。但是由于描述符维数过高,直接影响该方法的计算速度,而且结果该方法很不稳定。基于区域梯度的方法,例如 Low 提出的 SIFT 描述符<sup>[37]</sup>,近年来受到广泛关注,理论和实验均表明该方法具有较好的性能。SIFT 描述符将特征点周围的 $16 \times 16$ 像素窗口分割成 16 个 $4 \times 4$ 的像素子窗口,然后统计每个像素子窗口的方向梯度直方图。将每个子窗口的梯度方向分成 8 个方向计算。这样就形成了 $8 \times 4 \times 4 = 128$ 维描述符:一共具有 $4 \times 4$ 个子窗口,并且每个窗口的描述是 8 位的,用来描述 8 个方向的梯度大小值。之所以要采用方向直方图来作为特征描述符,是因为直方图是由梯度计算得来的。基于梯度的特征描述方法会对明暗畸变具有较好的抵抗性,同时考虑了方向,因此具有旋转不变性。SIFT 的参数是由实验得出的。虽然理论上讲 SIFT 方法并不具备抗仿射变换,然而通过实验得出的 SIFT 对仿射变换的抵抗性并没有理论分析的那么差,因而该方法被认为是性能最好的描述符<sup>[38]</sup>。

### 3.2.2 改进的特征点描述符

移动机器人导航对双目视觉系统的要求之一是需要满足实时性,而双目视觉系统中图像匹配对实时性的影响最为明显。好的特征描述符不但可以加快立体图像匹配的速度,而且能够避免错误匹配的发生。由于移动机器人导航在对障碍物进行检测时,需要全面扫描整个视野,因此本文采用 Harris 角点检测算法,使特征点较均匀的分布于整幅图像,并设置适当的阈值,确保在有障碍物的情况下能够检测到满足要求的特征点个数。

在双目视觉系统中,左右两个摄像机对视频数据同时进行采集,视频分帧得到图像序列信息。由于两摄像头之间的基线长度对目标的大小影响较小,因此我

们可以假设它们在同一尺度上进行操作。对于左右图像内的特征点，利用公式(3-12)和公式(3-13)分别计算角点的梯度大小和梯度方向。

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad \text{式(3-12)}$$

$$\theta(x, y) = \tan^{-1}((L(x, y+1) - L(x, y-1)) / (L(x+1, y) - L(x-1, y))) \quad \text{式(3-13)}$$

因为单纯 Harris 角点的特征不足以表示出角点的所有特征，难以进行准确的立体匹配。根据 Harris 角点的特性，综合利用角点周围像素点的特征来描述一个 Harris 角点的特征方向。文献[31]已经给出计算 Harris 角点方向特征向量的方法，本文对其进行改进如下：

(1) 用公式(3-12)和公式(3-13)计算 Harris 角点周围 3x3 邻域内的像素梯度信息，这样得到了 9 个点的梯度方向和梯度大小信息。在双目视觉左右图像角点提取时，不存在图像边缘点，这样每个 Harris 角点周围的 9 个点的梯度方向和大小都可以求得。

(2) 在直方图中将 0 度到 360 度分成 36 个柱，如图 3-2 所示。将上一步 9 个点的梯度信息按方向投影到直方图当中，梯度方向在同一个柱内的梯度信息按梯度大小进行累加。

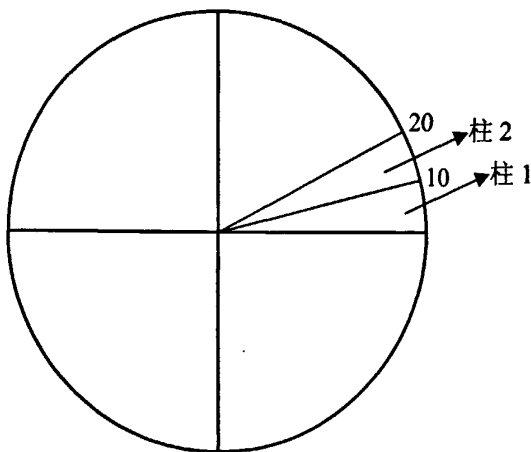


图 3-2 36 柱直方图

(3) 直方图中的峰值则代表了该特征点处邻域梯度的主方向，即作为该特征点的方向。在梯度方向直方图中，当存在另一个相当于主峰值 80% 能量的峰值时，则将这个方向认为是该特征点的辅方向，以增强特征点匹配的鲁棒性。

在描述特征点的特征向量时，文献[31]将 SIFT 算法<sup>[39]</sup>中 128 维向量引用到了 Harris 角点的特征描述中，本文对该算法进行了改进。

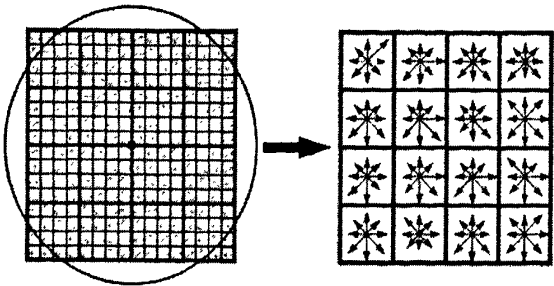


图 3-3 由 Harris 特征点邻域梯度信息生成特征向量

图 3-3 左部分的中央黑点为当前特征点的位置。每个小方格代表关键点邻域所在尺度空间的一个像素，箭头方向代表该像素的梯度方向，箭头长度代表梯度模值，绿色的圆圈代表高斯加权的范围（越靠近中央特征点的像素梯度方向信息贡献越大）。Harris 特征点的特征向量计算方法如下：

(1)以角点的特征方向为主方向建立坐标系，取角点周围 16x16 范围内的一个窗口区域，用公式(3-12)和公式(3-13)计算此矩阵内像素的梯度大小和方向，并将此 16x16 窗口区域分成 16 个 4x4 的小方块。

(2)将每个小方块中像素的特征向量投影到将 360 度平分为 8 个柱的直方图中。在投影的过程当中，考虑高斯加权的影响，越靠近中央特征点的像素梯度方向信息贡献越大，同时对同一个柱内的梯度投影进行累加，此时得到一个  $16 \times 8 = 128$  维的特征向量，此向量即为特征点的 SIFT 特征向量。对于有辅方向的特征点，可以进一步求得一个 128 维特征向量。对所求得的向量进行归一化处理，使其对旋转，亮度变化保持不变性。

在立体匹配部分，通过实验将会看到本节所改进的描述符对特征点的正确匹配起到了关键作用。

### 3.3 基于点特征的立体匹配

双目视觉中的立体匹配是指根据对所选的特征进行计算，建立特征之间的对应关系，将同一个空间物理点在左右摄像机拍摄图像中的映像点对应起来。双目视觉中的立体匹配与普通的图像配准并不同。立体像对之间的差异并不是由其它

如景物本身的运动、变化所引起的,而是由左右摄像机的观察点不同导致的。在双目视觉最后的障碍物判别以及距离计算中,都需要用到左右匹配点的视差值,因此立体匹配是双目立体视觉中最关键、困难的一步。双目视觉匹配算法根据匹配的基元不同可以分为三大类:区域匹配、特征匹配和相位匹配。这三种算法因匹配基元不同从而使得匹配基元的稳定性、致密性和歧义性程度直接决定了各类算法的基本特性,同时判断对应点匹配的理论依据也有所不同。另一方面,各类匹配算法中不乏一些共同的约束条件,比如 Marr 立体视觉计算理论中提出的唯一性、相容性和连续性约束。而对于不同的匹配基元,相似性测度的算法模型可以是通用的,比如都可以用欧式距离测度来进行匹配度量。

### 3.3.1 基于点特征匹配方法的特征分析

目前,针对双目视觉立体匹配方法的研究主要分为两大类,分别是基于窗口的匹配方法和基于特征的匹配方法。各种算法的主要差异在于相关度量标准以及搜索算法的选择不同。在图像匹配中最经常用的方法就是基于窗口的匹配方法,该法被看做是解决对应问题的最简单最直观的匹配方法。此方法思想如下,在一幅图像中选取一点作为中心选定一窗口,窗口是由待匹配点附近的像素灰度值组成的二维矩阵,在另一幅图像中寻找与该区域相关系数最大的区域,把找到的区域的中心认为是原来那窗口区域中心的对应点。最典型的如灰度互相关算法等,这是目前大多数匹配算法的基础。此类算法匹配精度较高,但时间复杂度也较高。Rosenholm 进行了改进,用连接窗在全局范围内进行匹配,而不使用固定大小的窗口,该方法对于处理重复纹理得到了较好的效果<sup>[40]</sup>。

目前大部分算法在另一幅图像中选取匹配像素对依据的是局部区域相关系数。基于区域的灰度相关算法实现简单,视差场密度大,但由于其易受噪声干扰并且计算量大,从而速度较慢。遗憾的是得到的相关值不能得到正确的匹配,因为这种方法依据局部区域相关系数来选取匹配像素对,只考虑到了图像的局部特征。

尽管人们提出了各种各样的改进方法,但是基于窗口的图像匹配仍存在着以下几点不足:首先由于基于窗口的图像匹配对于图像的旋转以及光强和对比度的变化等非常敏感,因为该方法是直接利用图像的像素灰度值进行匹配的;其次是

常常会引起匹配的混淆,给出错误的匹配结果,特别是当左右两幅图像中相关像素邻域内存在遮挡现象或存在重复结构的纹理特征时;再次这种方法的时空复杂性是很大的。虽然采用了内极线假设以及由粗到细的层次化结构等约束条件可以在一定程度上减少基于窗口的图像匹配的计算量,但互相关匹配的运算量仍然很大。

另一种双目视觉中常用的立体匹配方法是基于特征的匹配方法,这也是双目视觉应用中比较重要的一种匹配方法。基于特征的匹配方法不是直接利用图像灰度,而是通过灰度导出符号特征描述符来实现匹配。所以该法的第一步是提取感性的特征,并对特征值进行描述,如 D. G. Lowe 等提出用 SIFT 特征描述符<sup>[39]</sup>进行匹配。但可惜的是他使用的是不具有视觉特征而仅仅具有数学特征的兴趣点。该方法对于对比度和明显的光照变化等表现相对稳定。同时,可以通过对特征属性的简单比较而实现基于特征的匹配。因此,基于灰度的匹配要比该方法慢很多。目前在灰度图像特征匹配过程中常用的特征有边缘线段和特征点等。

基于特征的匹配最主要的是需要计算边缘或区域的特征。这些特征不但是图像内容的更加抽象的描述,而且在不同的照明条件和宽基线变换下是不变的。但是如果对这些特征直接进行计算,计算开销通常非常高,比较耗时,不能满足实时性的要求。相关研究者为了进一步提高速度已经做了长时间的深入研究。

### 3.3.2 改进的基于点特征的立体匹配方法

在上一节中,本文经过采用 Harris 算子,以使其能够提取出符合双目视觉导航要求的特征点。接着本文通过改进 Harris 特征点的 SIFT 描述符,以得到能够进行精确和快速匹配的特征向量。在本小节中,将利用得到的这一特征向量进行特征点匹配。本文之所以选取特征点作为匹配基元有以下几方面原因:

第一,基于点特征的立体匹配方法在速度上会比其他方法快很多。该方法主要利用了特征向量作为匹配中的首要元素,而在其他方法中,一般用图像本身的灰度值或亮度值来作为匹配元素,一方面受到匹配基元大小的影响,另一方面受到实时性的约束。

第二,基于点特征的立体匹配方法更加符合移动机器人导航的要求。移动机



机器人导航首先需要判别障碍物,依据特征点的特征对障碍物进行判断比利用灰度区域进行判断适用范围更加广泛。这是因为利用灰度进行判断只能对某一类别的障碍物做出识别。而在位置环境中,对障碍物的要求没有规定,此时基于灰度区域的判别方法达不到移动机器人导航的要求。其次移动机器人导航需要计算障碍物的距离和角度,在基于灰度区域的匹配方法中无法选定具有代表性的一点来作为计算距离和角度的基础。而在基于点特征的立体匹配中,由于本身已经提取出特征点,下一步要做的就是对障碍物中距离最近的特征点距离和角度进行计算,这样不但可以代表整个障碍物的三维信息,并且能够满足移动机器人导航的要求。

本文改进的基于点特征的立体匹配方法流程图如图 3-4 所示:

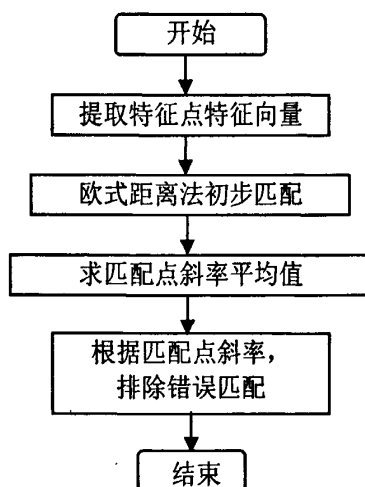


图 3-4 基于点特征的立体匹配方法流程图

当双目视觉采集到的两幅图像中特征点的特征向量生成之后,下一步本文采用的是特征向量的欧式距离来作为两幅图像当中特征点的相似性判断度量。取左图像中的某个特征点,在右图像当中找出与其特征向量欧氏距离最近的特征点。欧氏距离的计算公式如公式(3-14)所示:

$$D = \sqrt{(x_1 - x_1')^2 + (x_2 - x_2')^2 + \dots + (x_{128} - x_{128}')^2} \quad \text{式(3-14)}$$

其中,  $(x_1, x_2 \dots x_{128})$ ,  $(x_1', x_2' \dots x_{128}')$  为待匹配的左右图像两个特征点的特征向量。 $D$  越小,说明匹配程度越高。在匹配过程中,为了加快速度,需要考虑外极线约束。

得到初步匹配结果之后,下一步需要做的就是对错误匹配点进行处理。为了降低匹配时间的复杂度,本文提出一种根据初步匹配点对的斜率进行排除错误匹配的方法。在双目视觉的左右图像内,正确匹配点两坐标的斜率会局限在某一范围内,而只有错误的匹配点对的斜率才会超出这一个范围,根据这一特点,本文提出的排除错误匹配的方法如下:

(1)计算初步匹配得出的匹配点对的斜率,并计算平均值。

(2)设定阈值 $k$ ,删除匹配点对的斜率与平均值差的绝对值大于阈值的匹配点对。

阈值 $k$ 设定的越小,排除错误的匹配点越多,但误删除的情况也随之增加。本文设定 $k=5$ 。

### 3.4 实验结果

本文通过筑波大学提供的图像对进行测试,测试结果如图 3-5 所示。

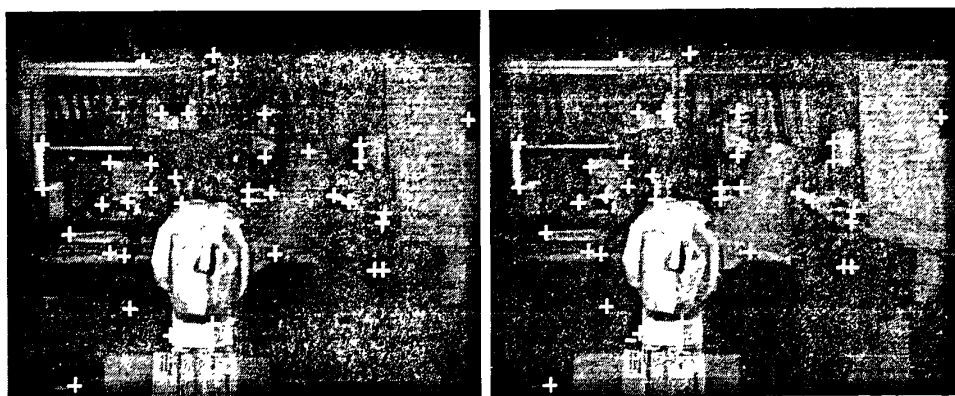


图 3-5 筑波大学测试图像对

本文算法首先采用 Harris 特征点提取算子对左右两幅图中的特征点进行提取,并采用改进了的 SIFT 描述符进行描述。通过本文改进的立体匹配方法,对这两幅图中的 36 对匹配点成功匹配(图中红色标记为最终所得到的 36 对匹配点)。在本算法所做的测试当中,对双目视觉当中障碍物图像的匹配正确率一直保持在 97%以上。阈值 $k$ 设定的越小,排除错误的匹配点越多,但误删除的情况也随之增加。本文设定 $k=5$ 。

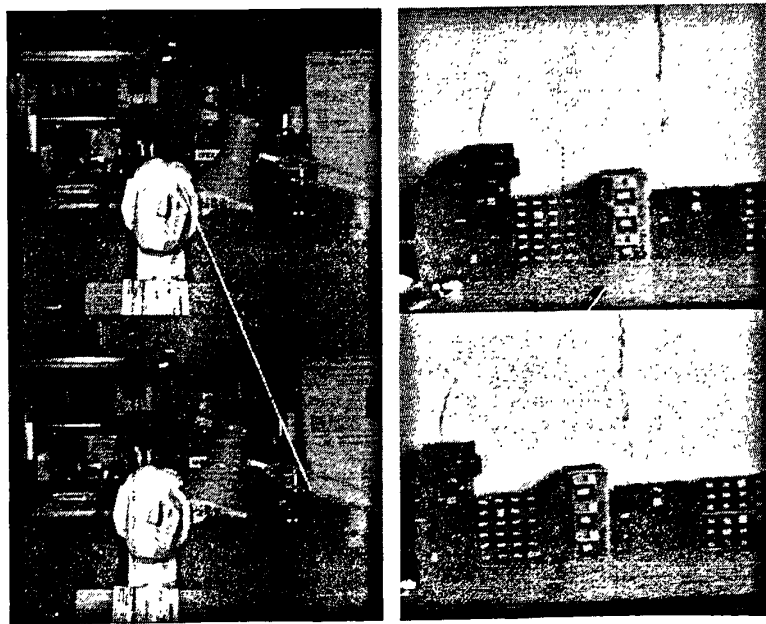
表 3-1 列出了实验所用的五对图片的匹配结果以及所用的时间。从列表中可

以看出,本文所提出的算法改进在时间复杂度上控制的比较好,对于一般的双目视觉障碍物来讲,可以满足实时的要求。本文提出的利用斜率进行排错的实验结果令人满意。

表 3-1 实验结果 (CPU:2.31GHz,Memory:1G)

图像组	本文算法匹配正确率	文献[31]方法匹配正确率	SIFT 算法匹配正确率	本文算法排除错误匹配对	匹配所用时间 ms
第 1 组	97.3%	95%	93%	1	447
第 2 组	97.5%	97%	96.9%	1	406
第 3 组	98%	95%	91.7%	3	125
第 4 组	98%	96%	95%	1	31
第 5 组	97.5%	94%	90%	4	594

在图 3-6 中,线段连接起来的匹配点对为本文提出的算法所排除的匹配点对。由于这些线段的斜率与匹配点对的平均斜率超过阈值(本文设为 5),这些错误匹配点均被成功排除。



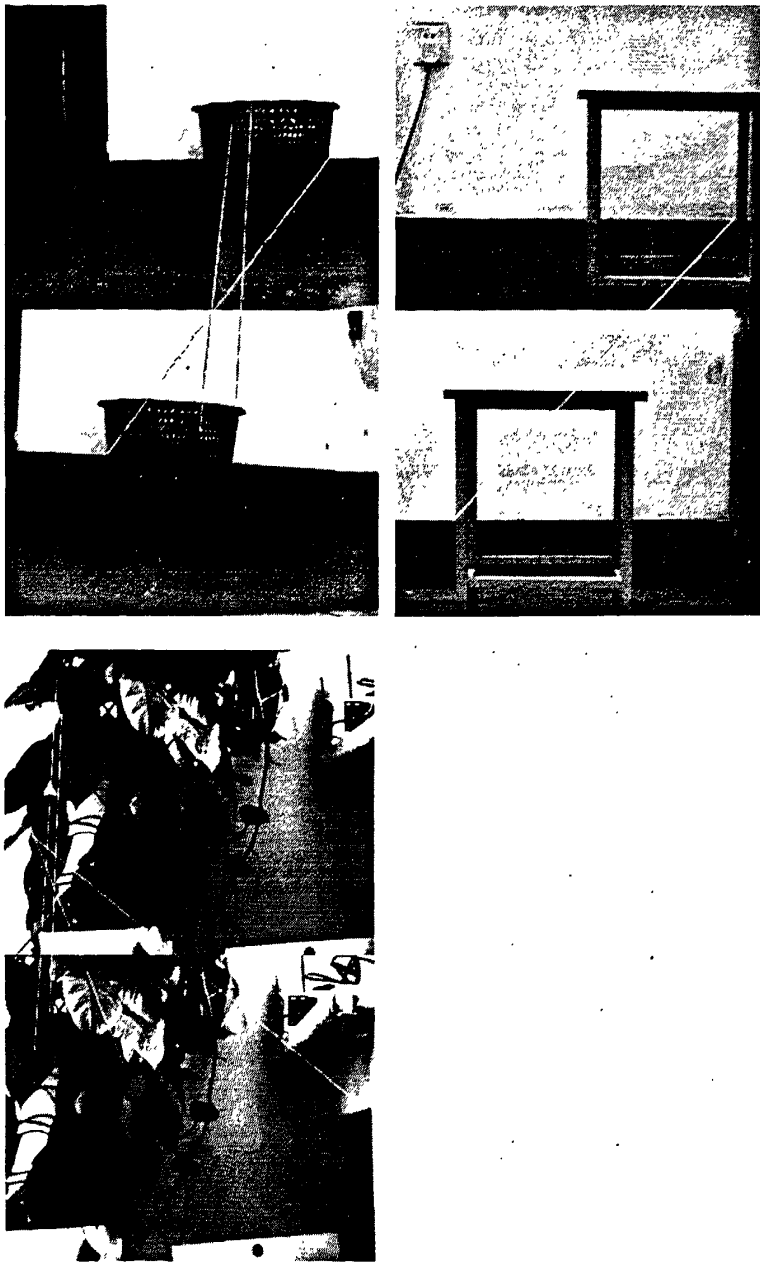


图 3-6 五组实验图的排除错误匹配点图

### 3.4 本章小结

本章首先分析了几种重要的特征点提取算子的优缺点,并根据移动机器人导航的实际要求,采用了 Harris 角点提取算法,保证在自然环境下双目视觉左右摄像机能够提取出符合要求的特征点。在第二节中分析特征点描述符对双目视觉立体匹配所带来的好处,并改进了对 SIFT 特征向量对特征点的描述。在第三节中,给出了本文所采用的立体图像匹配方法,提出了一种简单快速的去除错误匹配的

方法，最后通过实验证实了该方法实现了对匹配点较好的修正效果。

## 4 不规则障碍物检测与特征信息提取

### 4.1 双目视觉障碍物检测分析

障碍物检测作为移动机器人导航的关键,一直是图像理解领域研究的一个重要分支,也是基于模式识别、数据处理等理论,模仿人的视觉系统感知外部世界各种障碍物的一项技术。基于双目视觉的障碍物特征检测方法是目前障碍物检测中最常用的方法。

在双目视觉障碍物检测方法中,一般的做法首先使用左右两个摄像机同时获取场景图像。在标定过的平行双目视觉系统中,左右两摄像机之间不存在旋转,只有标定基距的平移关系,因此在拍摄的两幅图像中也是简单的平移关系。然后通过图像匹配发现障碍物并得到图像间的视差。由于在检测障碍物的过程中是基于图像匹配,因此需要对障碍物的特征进行选取,以便于匹配过程中找到符合要求的障碍物。最后根据障碍物在图像中的位置、视差以及摄像机的标定参数计算出障碍物距离摄像机的实际距离和角度,以应用到自主移动机器人的 SLAM 导航中。

目前,基于双目视觉的环境特征检测方法有多种。例如,在成熟果实的定位<sup>[41]</sup>应用中,首先根据成熟果实的颜色特征利用阈值自动设定的方法对图像进行分割,快速地识别成熟果实。然后采用圆形 Hough 变换算法拟合出目标图像的形心位置坐标。在该方法的判别过程中,因为果实具有一般的颜色特征和形状特征,可以通过这两个特征对果实进行识别,并且在实际的农业果实采摘中收到了很好的应用效果。但是该方法在移动机器人的导航中存在障碍物不规则障碍,无法对无规律障碍物进行识别。有的检测方法是利用了边缘提取方法提取双目图像对中的边缘,然后进行边缘匹配,最后通过场景重建提取目标。例如在水下声纳图像的目标检测与识别技术中<sup>[42]</sup>,基于形状相似度的概念,利用目标边缘轮廓特征,对水雷等目标形状进行分类识别。首先对检测到得水雷图像进行处理,然后利用模板更新的策略进行识别。另外有的交通障碍物检测方法是多幅已知双目图像中的颜色进行实验,找到草地、马路等非障碍物区的颜色大致范围作为经验值,

应用到未知的双目图像中。首先对目标区和非目标区进行划分,然后利用反透视原理,对双目图像对中一幅图像的几个特征点进行变换,并与另一幅图像内相同特征点比较,确定是否存在障碍物,但仅适用于同类型目标的识别。

在移动机器人导航中,需要对前进过程中的障碍物进行检测。在未知环境中,无法对障碍物的颜色,形状做一个统一的规定,因此应用到移动机器人目标跟踪方法中的策略不适合本文的自主移动机器人障碍物检测的要求。根据项目的实际要求,本文的实验小组通过实验提出了利用视差图中特征点的三维信息对障碍物进行检测。

## 4.2 本文采用的障碍物检测算法

在双目视觉系统中,视差图是以左右图像对中任意一幅图像为基准,即大小为该基准图像的大小,元素值为视差值的图像。在传统的视差图应用中,靠视差图中的三维信息来进行三维重建,并没有应用到障碍物的检测中。本实验小组充分运用了视差图中的三维信息,根据移动机器人 SLAM 的要求,提出了一种符合实际应用要求的障碍物检测算法。

### 4.2.1 视差图信息提取

在第三章中,本文采用了 Harris 特征点提取算子,并且改进了特征点的描述符。在初步提取特征点之后,提出了一种新的排除错误匹配的算法,得到很好的匹配结果,为本节视差图的生成提供了保障。

本节以双目视觉图像对的左图像作为基准图像,根据上一章中提取出的匹配点信息对视差图进行生成。首先按照左图像尺寸生成视差图大小,并以黑色填充整幅图像;其次根据匹配点的二维坐标信息  $(x, y)$  和视差值相关系数  $d_{\text{相关}}$  对图像进行着色,分一下几种情况:

第一,当视差值相关系数  $d_{\text{相关}} > 255$  时,在视差图的匹配点坐标  $(x, y)$  处涂以  $(255, 0, 0)$  的颜色;

第二,当视差值相关系数  $0 < d_{\text{相关}} < 255$  时,在视差图的匹配点坐标  $(x, y)$  处

涂以  $(d_{\text{相关}}, 0, 255 - d_{\text{相关}})$  的颜色;

第三, 当在左图像的  $(x, y)$  点处不存在匹配点时, 则保留生成图像时的黑色不变。

其中的颜色值是以三原色  $(r, g, b)$  的形式给出的。 $(r, g, b)$  中  $r$  代表红色,  $g$  代表绿色,  $b$  代表蓝色, 此三种颜色称为三原色, 可组成为计算机可以显示图像的所有颜色, 每个字母的取值范围为  $0 \sim 255$ , 数值越高, 表示此颜色所占的比重越大。 $d_{\text{相关}}$  可以通过式 4-1 计算:

$$d_{\text{相关}} = 3 \times \sqrt{(x - x')^2 + (y - y')^2} \quad \text{式(4-1)}$$

$(x, y)$  代表匹配点在左图中的坐标, 也即为在生成的视差图中的坐标,  $(x', y')$  为右图中相应匹配点的坐标。之所以需要  $d_{\text{相关}}$  是为了在视差图中使显示的效果更加明显, 更具有层次感。此时生成的视差图是由黑色的背景图像和由蓝色至红色过渡的匹配特征点点组成, 无法达到对障碍物进行检测的要求, 需要对匹配点进行区域自动增长才能生成色彩连续的视差图。本文的视差图生成流程图如下图 4-1 所示:

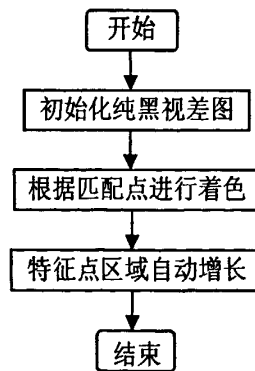


图 4-1 生成视差图流程图

由此得到的视差图包含了左图像中原点的坐标信息  $(x, y)$ , 图像由蓝到紫到红的变化反映了像素匹配时视差值的大小, 视差值越大, 红色越明显, 视差值越小, 蓝色越明显。本文生成的视差图效果如图 4-2:



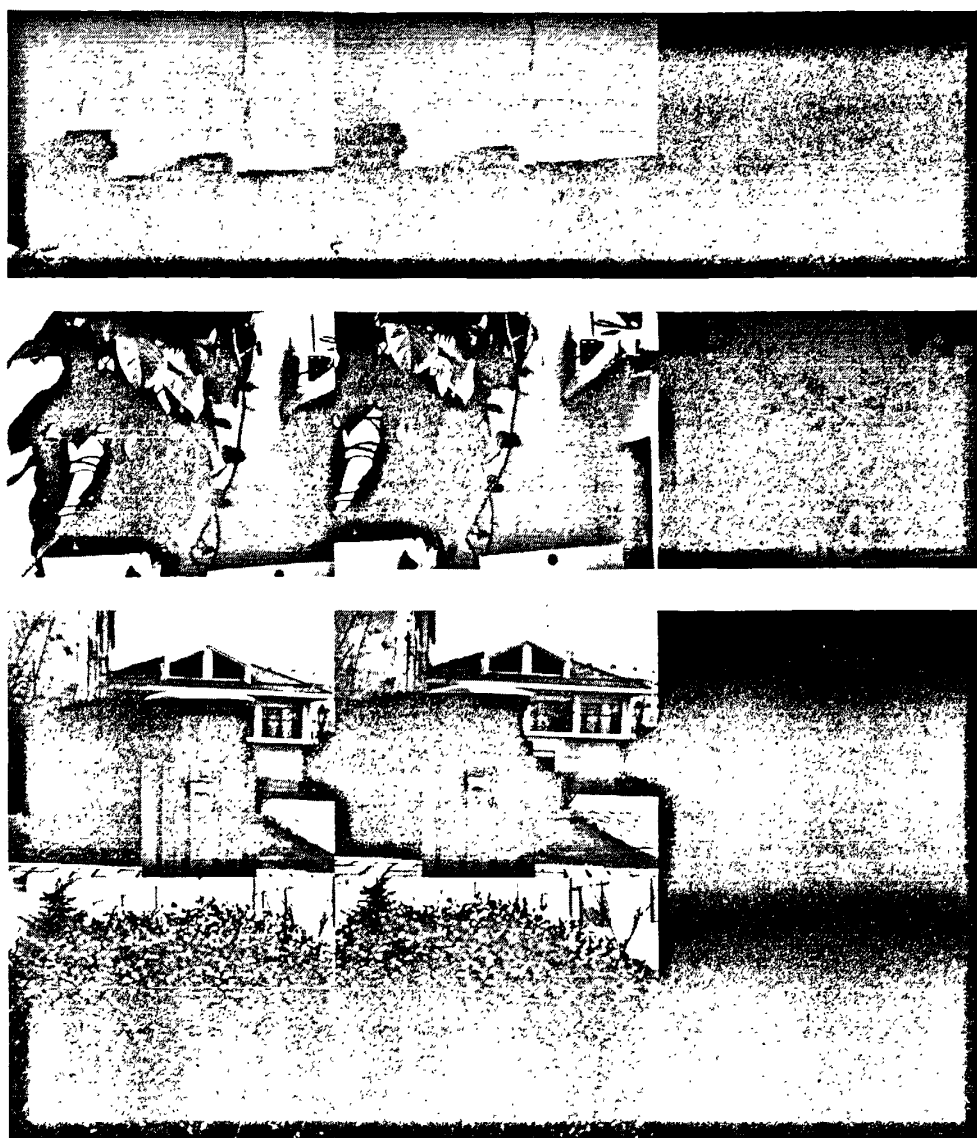


图 4-2 测试图像生成的视差图效果图

在图 4-2 中，左边一列是双目视觉系统左摄像机采集到的图像，中间一列是右摄像机采集到的图像，第三列为生成的相应视差图。在第一组图像中，本文拍摄的是实验室中的实验台。根据生成的视差图分析可知，实验台上的工具箱得到了很好的匹配，视差图的着色整体形状基本上与工具箱的排列整体形状相吻合，并且在彩色图像中，视差图中的红色（灰度图像显示偏白）很明显，表明工具箱离摄像机的距离很近。在第二组图像中，重点拍摄了实验室中的盆栽绿萝，绿萝和桌子一角都得到了很好的匹配，在颜色显示上两者均以红色（灰度图像偏白）为主，表明二者离摄像机的距离都比较近。在第三组图像中，本实验小组进行了外景实地拍摄。从视差图的显示效果来看，较近处的垃圾箱主要以紫色（灰度图

像显示偏白)显示为主,而较远处的楼房则显示成了蓝色(灰度图像显示偏白)。根据视差图的生成原理来看,楼房较远,得到匹配,但是视差值很小,相应的 $d_{\text{相关}}$ 也很小,所以显示的是蓝色,而近处的垃圾箱也得到匹配,但是由于距离摄像头比较近,相比远处得到匹配的楼房来看,视差图很大,相应的 $d_{\text{相关}}$ 也较大,所以显示了过渡颜色紫色。该副视差图层次感很明显,很好的体现了远近事物的不同。最后一组图像是双目视觉系统行进过程中遇到的一棵冬青。对生成的视差图分析可得,由于颜色是有红色向紫色渐渐过渡(灰度图像由灰白向灰黑过渡),因此可以表明,在中央部分冬青离摄像机较近,逐步向周围扩散,离摄像机越来越远。远处的松树一角也得到了很好的匹配,在视差图中以蓝色进行了显示。

由此可见,视差图中包含了环境事物中的深度信息,并且对匹配区域进行了很好的整合,可以从整体上对障碍物进行显示。因此可以通过视差图,提取障碍物的三维信息,并通过障碍物的高度和连通域来对障碍物进行检测。

#### 4.2.2 障碍物检测算法

移动机器人在移动过程中,首先要解决的问题就是避障,然后才能顺利按照路径规划的设计到达预设终点。而障碍物检测研究是双目视觉研究中的重点,也是难点。在很多场合中,可以设定障碍物的形状或者颜色,然后通过检测这些特征对障碍物进行检测或对目标进行跟踪。而在移动机器人运行过程中,周围环境未知,无法对障碍物进行预先特征标定,或者标定之后无法检测所有情况下的障碍物,因此预先设定特征的方法无法满足移动机器人导航的要求。

视差图在双目视觉系统中经常用作进一步进行三维重建,即在障碍物判别之后,根据视差图中信息进一步进行计算,视差图显示的不单单是匹配点的二维信息,高度信息也会通过计算得到,此时得到的视差图为三维的视差图,包含了特征点的三维信息。本文为满足移动机器人导航的要求,另辟蹊径,通过研究视差图可以发现,可以根据匹配点的高度信息来对障碍物进行判定,并且可以通过视差图中的相应连通域来判断障碍物的尺寸范围。本文为进一步提高运算速度,采用了 $320 \times 240$ 作为图像的基准大小,具体的算法描述如下:

1. 设定障碍物高度阈值。在本小组进行的实地试验中,选择以玩具车代替

机器人进行双目视觉系统障碍物检测测试。测量摄像机到地面的实际距离为 650mm, 在进行高度计算时, 需要在世界坐标系中对这一高度进行补偿, 以算出障碍物到水平面得高度。障碍物依据测试环境的不同而不同, 在信息学院门口公路上进行测试时, 本小组选定的高度阈值为 500mm。

2. 分割视差图。将  $320 \times 240$  的视差图进行分割, 每一小块为  $30px \times 30px$ , 边缘部分的剩余独立出来各成一小块;

3. 计算平均高度。针对视差图中的每一小块进行计算, 在进行特征点匹配时, 左右图像中的匹配的特征点坐标已经保存在相应的数组中。根据匹配时得到的坐标, 即可计算每个  $30px \times 30px$  小方块中特征点的视差值, 根据第二章中推导的公式便可计算该点对应的实际物点的高度。同时每一小方块中可计算高度特征点的个数可以进行计算, 包括在特征点区域自动生长时进行区域扩展的点。这样做的目的是为了提高运算的准确度, 以保障障碍物的范围可以全部包含在内。

4. 标记障碍物区域。在计算平均高度时, 如果小方块中点的平均高度大于给定的阈值, 并且像素的视差值  $d > 30$ , 则进行标记。这是因为本视觉系统的测定距离为 5 米之内, 在  $d = 30$  的情况下, 通过实验发现距离已经超过 5 米, 所以设定了视差值的取值范围。将该  $30px \times 30px$  小方块标记为绿色。在遍历完整个视差图后, 可以看到符合要求的区域已经被标记为绿色。

5. 判别障碍物。在上一步标记处的绿色区域, 有的地方是有很多小方块组成, 而有的地方是有零星的几个小方块, 甚至是单独的一个。本步要做的是对由 30 个小方块组成的连通区域进一步判别为障碍物, 而其他区域则不作为障碍物判断。

本文实验算法的流程如图 4-3 所示:

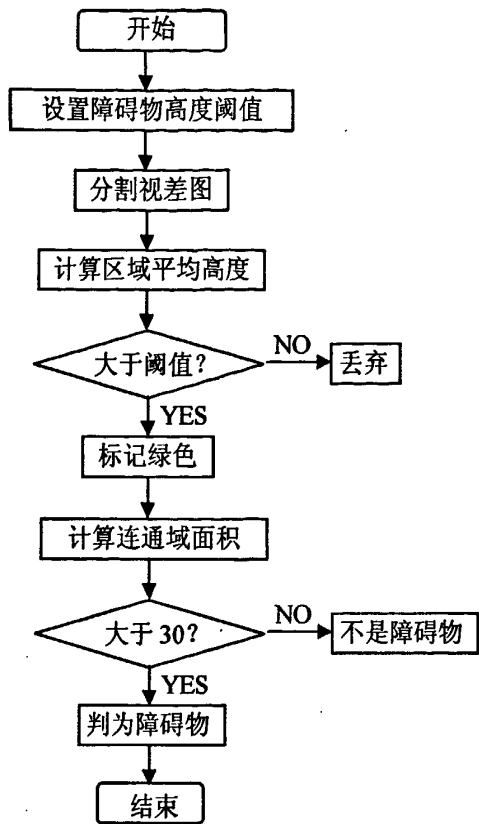


图 4-3 障碍物标记算法流程图

本文利用视差图对障碍物判别进行了尝试，并且在实验中收到了很好的效果。图 4-4，图 4-5 和图 4-6 为实验中标记出的障碍物区域，每组图像的上两幅为左右图像，左下图为视差图，右下图为障碍物识别图。



图 4-4 障碍物判别效果图 1

在图 4-4 中，视差图显示了大面积的蓝色和紫色（灰度图像显黑），表明视野范围距离摄像头都比较远，经过实地测量台阶和垃圾箱距离摄像头均在五米之外。在判别时，效果图和视差图一样，没有显示检测到障碍物，符合实际的要求。

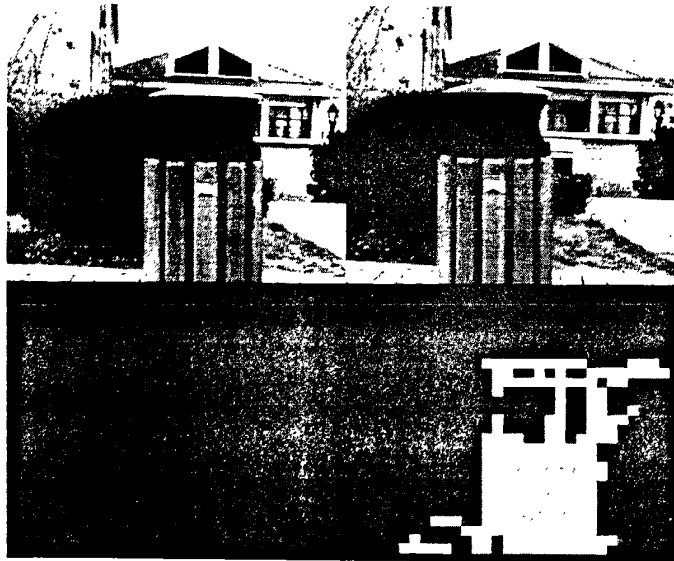


图 4-5 障碍物判别效果图 2

图 4-5 是在图 4-4 的基础上将摄像头向垃圾箱推进了一些距离，在 5 米范围之内。可以看出，在视差图中垃圾箱已经显示为暗红色（灰白图像显示灰黑），表明此时距离摄像头距离很近，而在障碍物判别效果图中，绿色覆盖区域已经基本涵盖了这跟个垃圾箱，并且连通的区域超过 30 个小方块，因此判为障碍物。

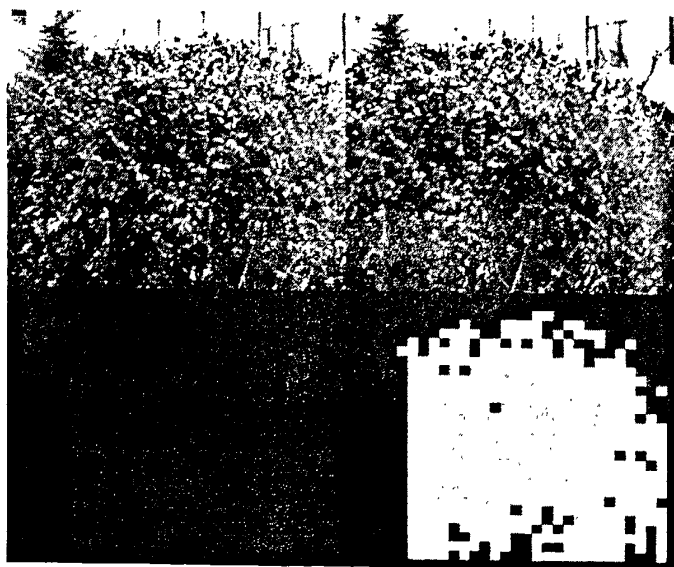


图 4-6 障碍物判别效果图 3

图 4-6 是测试时摄像头移动到冬青丛中拍摄的图像。在最后的障碍物判别效果图中可以看出，绿色区域已经覆盖整个冬青丛。

由实际实验可以看出，本节提出的算法在实际运行过程中收到了不错的效果：在前方有障碍物时，检测的障碍物标记区域能够准确覆盖障碍物尺寸；在前方无障碍物或者障碍物距离摄像头距离大于 5 米的时候，不会误判断障碍物。

#### 4.2.3 障碍物特征信息提取

基于点特征的 SLAM 导航算法需要将每一个障碍物视为一个点进行地图创建。根据这个要求，本实验小组对障碍物的连通域进行如下处理：

首先定位中心小方块所在位置：对障碍物的连通域进行扫描，并计算出中心小方块所在的像素范围。然后确定特征点的坐标：针对中心小方块，本实验小组首先求取该方块离摄像机的平均距离，然后对小方块进行扫描，求出与平均距离最相近的第一个点的坐标，该点即作为最后要输出的障碍物距离和角度的坐标。

对于具有其他应用的场合，在求得障碍物连通域范围以及视差图的情况下，对于障碍物的任何三维信息均可以通过计算给出，由于本项目现定于 SLAM 导航算法的要求，因此只给出其中一点的三维信息。

### 4.4 系统整合与实验结果

#### 4.4.1 平行双目视觉系统整合

根据基于点特征的 SLAM 导航算法对环境特征的要求，双目视觉系统需要对障碍物进行点化处理，并计算出该点的世界坐标系三维坐标 $(X_w, Y_w, Z_w)$ 以及与世界坐标系 $Z_w$ 轴在 $O_w - X_w Z_w$ 平面内的夹角。三维坐标公式在摄像机标定中已经推导得出，根据 SLAM 导航算法的要求，夹角的符号与 $X_w$ 的符号相反，式 4-2 为计算夹角的公式：

$$\theta = -\arctan(X_w / Z_w) \quad \text{式(4-2)}$$

本文平行双目视觉系统在 VC++ 6.0 平台下，并结合 Intel 公司的开放源代码 OpenCV 库进行程序实现。用于标定摄像机内外参数的程序不参与系统整合，对

其余的需要在实时环境中应用的部分进行系统整合。系统整合后的程序流程如图 4-7:

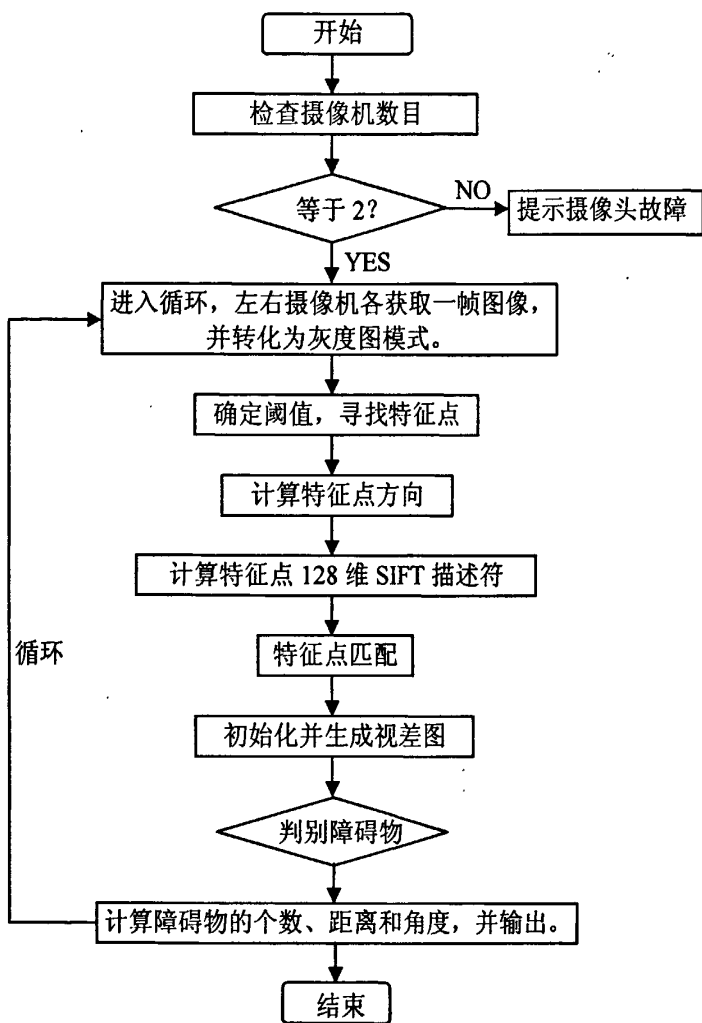


图 4-7 系统整合流程图

4.4.2 实验结果

本文对系统流程图中的各个部分分别进行了实现,对系统整合后,在实际环境中进行了测试。在前期工作准备充分的基础上,本小组进行的实验结果比较理想,不难能够在前方有障碍物时能够辨别出障碍物,在没有障碍物时能够不误判,并且通过计算公式计算出的距离和角度结果也在误差范围之内。图 4-8 分别给出了用玩具车搭载的实验系统在自然界场地中进行测试的场景以及本文在 VC++ 6.0 中进行实现的程序运行界面。

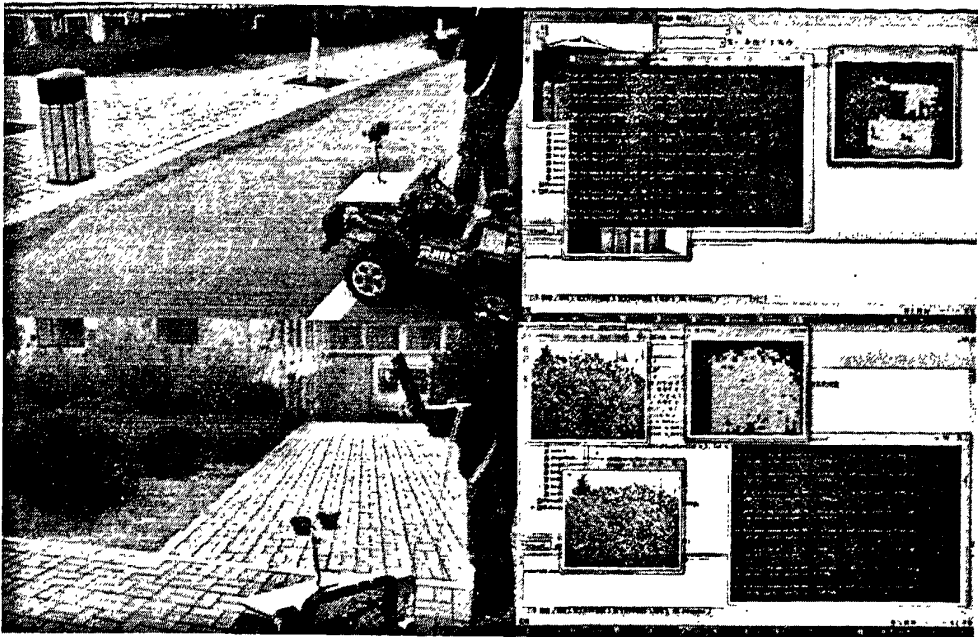


图 4-8 实验场景图及程序运行图

在实际测试中，本实验小组先后在走廊中对盆栽绿萝进行了测试，在实验室对标定板进行了测试，在外界场地中，对马路边的垃圾箱和冬青分别进行了测试，并在马路上进行了测试。在测试过程中，对障碍物的判别比较准确，由于光线的影响，输出的数据会有很小的变动，本文选取了中间值作为最后计算值。对测量值，本文采取的是最近测量点作为测量距离 $l_{测}$ ，并对测量值精确到厘米，所以会对误差百分比带来一定的影响。距离测试数据如表 4-1 所示：

表 4-1 系统运行后实际距离测试结果

障碍物	障碍物个数	测量距离 $l_{测}$	计算距离 $l_{计}$	误差百分比
盆栽绿萝	2	3210mm	3325mm	3.6%
		3210mm	3354mm	4.5%
标定板	1	2020mm	2056mm	1.9%
垃圾箱	1	1550mm	1607mm	3.7%
冬青	1	2630mm	2729mm	3.8%
马路	0			

由于在对盆栽绿萝的测试过程中，实验小组有意的摆放了两棵，根据实验结果来看，本文提出的障碍物判别方法准确的判定 2 个障碍物。在其他情形下，实验结果也较为理想，距离测试值达到了误差必须在 5% 以下的要求。

根据基于点特征的 SLAM 导航算法的要求，正前方的障碍物角度为 0，在左



前方的障碍物角度为正值，右前方的障碍物角度为负值，以弧度为单位。测试数据如表 4-2 所示：

表 4-2 系统运行后实际角度测试结果

障碍物	障碍物个数	测量角度 $\theta_{测}$	计算角度 $\theta_{计}$
盆栽绿萝	2	0.17	0.15
		-0.14	-0.16
标定板	1	0	0.02
垃圾箱	1	-0.09	-0.08
冬青	1	0.07	0.05

由于目前进行测试用的摄像机视野范围较窄，只能对正前方的障碍物进行测试，所以测试所得的数据都集中在 0 附近。对于角度的实际测量值，本实验小组测量的是障碍物最近点和左摄像机光心的连线与光轴之间的夹角。光心选为摄像机的中心点，光轴选为摄像机的正前方方向。

4.5 本章小结

本章首先在第三章特征点匹配的基础上介绍了视差图的生成过程，并紧接着给出了试验效果图；接着在分析了传统的障碍物判别方法之后，提出了本文所用的方法，同时进行了外景实地实验，根据实验效果来看，本文算法对障碍物有较好的判别效果。根据基于点特征的 SLAM 要求，对障碍物的特征信息进行提取，并进行系统整合和实验，实验结果达到要求。

## 5 总结与展望

### 5.1 总结

基于平行双目视觉系统的环境特征信息检测在当今社会中应用非常广泛。本文根据移动机器人中基于点特征的 SLAM 导航算法要求,对平行双目视觉的障碍物特征信息提取关键技术进行了研究。

本文研究了平行双目视觉系统的标定。首先确定了摄像机的成像模型和平行双目视觉系统的架设方案,根据自行推导的公式可以得出需要标定的内外参数,并自行设计实验来对参数进行标定。接着本文对双目视觉系统中最重要的特征点匹配环节进行了研究,改进了特征点的 SIFT 特征描述符,并提出了一种快速的去除错误匹配点的方法,同时通过编写程序进行实物测试实验,实验结果表明本文的改进和排错方案对整个匹配效果有较大改进。最后本文根据视差图提出了一种不规则障碍物判别方法,并分析项目要求,通过系统整合,对障碍物的特征信息进行输出。整个算法过程在 VC++ 6.0 的平台下配以 OpenCV 视觉开发库实现,实验效果较为理想。

### 5.2 展望

基于双目视觉的环境特征提取技术具有广阔的应用前景,本文以相关科研课题的基于双目视觉的机器人导航应用为背景展开研究。本文算法中有如下问题需要进一步研究:

1. 对整个双目视觉系统的计算速度需要进一步提高,或者通过其他导航传感器的数据进行数据补偿。
2. 本文主要针对自然环境的特征信息提取,因此对于一些人工平滑环境不能较好的提取特征点,会导致整个提取结果的质量下降。有待对人工环境中的特征匹配算法进行更加深入的研究。

由于时间原因,本文主要研究了摄像机标定技术和平行双目视觉系统的特征点描述方法,并且重点研究了自然环境下的不规则障碍物判别技术。基于点特

征的 SLAM 导航算法和多传感器信息融合等对移动机器人的自主导航的发展亦起着非常关键的作用。如何实现基于多传感器信息融合的移动机器人复杂环境下精确自主导航将是我们的一个长期的研究目标。

## 参考文献

- [1] Barnard S, et al. Computational stereo. ACM Computing Surveys, 1982, 14: 553-572
- [2] 李奇, 冯华君, 徐之海, 韩一石, 黄红强. 计算机立体视觉技术综述. 光学技术, 2007(5): 71-73
- [3] 项海兵. 计算机视觉发展中存在的问题. 新浪潮. 1996(4): 29-30
- [4] 张福学. 机器人学-智能机器人传感技术. 电子工业出版社, 北京, 1996
- [5] 吴祥瑞. 机器人技术及应用. 北京航空航天大学出版社, 北京, 1994
- [6] James C. Kinsey, Ryan M. Eustice, et al. A Survey of underwater vehicle navigation: Recent advances and new challenges, In the Proceedings of the IFAC Conference of Manoeuvring and Control of Marine Craft. Lisbon, Portugal, 2006. Invited Paper
- [7] 陈亮. 基于图像处理的水下目标识别方法研究, 硕士论文, 哈尔滨工程大学出版社, 2006
- [8] H. G. Nquyen, P. J. Heckman, et al. Real time pattern recognition for Guidance of an autonomous undersea submersible, in Proc. of IEEE conf. on Robotics and Automation, 1988: 1767-1770
- [9] Clark F. Slson, Habib Abi-Rached, Ming Ye, Jonathan P. Hendrich. Widebaselinestereo vision for mars rovers. Proceeding Of The 2003 IEEE/RSJ International Conference On Intelligent Robots and Systems, Las Vagas, Nevada, 2003: 1302-1307
- [10] 卢韶芳, 刘大维. 自主式移动机器人导航研究现状及其相关技术. 农业机械学报, 2002(3): 112-116
- [11] 包桂秋, 周兆英, 熊沈蜀, 叶维英. 图像导航技术的发展和应. 航空计测技术, 2003, 23(6): 1-4
- [12] Cheng-Jian Lin, Jun-Guo Wang, Chi-Yung Lee. Pattern recognition using neural-fuzzy networks based on improved particleswam optimization [J]. Expert System with Applications, 2008, 6: 1-9
- [13] 王鹏飞, 孙立宁, 黄博. 地面移动机器人系统的研究现状与关键技术, 机器设计. 2006, (7): 2-4
- [14] F. R. Dalglish, S. Tetlow, R. L. Allwood. Experiments in laser-assisted visual sensing for AUV navigation, Control Engineering Practice 12 (2004) 1561-1573
- [15] Ian Fitzgerald. A vision system for an autonomous underwater vehicle, Thesis, the Australian National University, 1999
- [16] 张广军. 机器视觉. 北京航天航空大学出版社, 2006. 9
- [17] Stanley K., Wu Q. M. J., Jerbi A., Gruver W. Neural network-based vision guided robotics. Robotics and Automation, 1999. Proceedings. 1999 IEEE International Conference on, May 1999, 1:281-286
- [18] 马颂德, 张正友, 计算机视觉-计算理论与算法基础, 科学出版社, 1998
- [19] Luh J Y, Klaasen J A. A three dimensional vision by off-shelf system with multi-cameras. IEEE Trans. Pam I, 1985, 7 (1)
- [20] Hall E L, TioM B K, M cPherson C A, Sadjadi F A. Curved surface measurement and recognition for robot vision. In: Proc. IEEE Workshop on Industrial Application of Machine Vision. 1982
- [21] Yakimovsky Y, Cunningham R. A system for extracting three dimensional measurements from a stereo pair of Tvcameras. Computer Graphics and Image Processing, 1978. 7:195-210

- [22] Itoh A M, Ozawa S. Distance measuring methods using only simple vision constructed for moving robots. In: Proc. ICPR' 84. 1: 192-197
- [23] Hartley R. Euclidean reconstruction and invariants from multiple images. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1994, 16(10): 1036-1041
- [24] Tsai R Y. An efficient and accurate camera calibration technique for 3D machine vision. In: Proc. CV PR' 86. 364-374
- [25] Weng J, Cohen P, Herniou M. Camera calibration with distortion models and accuracy evaluation. IEEE Trans. PAMI, 1992, 14 (10): 965-980
- [26] Martins H A, Birk J R, Kelley R B. Camera models based on data from two calibration planes. Computer Graphics and Imaging Processing. 1981, 17:173-180
- [27] Wei G, Ma S. Complete two-plane camera calibration and experimental comparisons. In: Proc. ICCV' 93. 439-446
- [28] Wei G, Ma S. Two-plane calibration: a unified model. In: Proc. CV PR' 91. 1991. 133-138
- [29] Wei G, Ma S. Implicit and explicit camera calibration: theory and experiments. IEEE Trans. PAMI, 1994, 16(5)
- [30] ZHANG Z. A Flexible new technique for camera calibration[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(11): 1330-1334
- [31] 魏志强, 黄磊, 纪筱鹏. 基于点特征的序列图像匹配方法研究. 中国图象图形学报 A, 2009, 3: 525-530
- [32] H. Moravec. Obstacle avoidance and navigation in the real world by a seeing robot rover. Technical Report CMU-RI-TR-3, Carnegie-Mellon University, Robotics Institute, 1980, p. 177-187
- [33] Smith. S. M, Brady. J. M. SUSAN-A new approach to low level image processing. International Journal of Computer Vision, 1997, 23 (1): 45-78
- [34] M. Trajkovic and M. Hedley. Fast Corner Detection. Image and Vision Computing, 1998, Vol. 16 (2), pp. 75-87
- [35] Harris C G, Stephens M J. A Combined Corner and Edge Detector. Proceedings Fourth Alvey Vision Conference, Manchester, 1988: 147-151
- [36] Perez, M. M. , Dennis, T. J. An adaptive implementation of the SUSAN method for image edge and feature detection [A]. Image Processing [C]. 1997, International Conference on Volume 2: 26-29, 394-397
- [37] D. Lowe. Object recognition from local scale-invariant features. Proceedings of the 1999 7<sup>th</sup> IEEE International Conference on Computer Vision (ICCV'99), Institute of Electrical and Electronics Engineers Inc, Piscataway, NJ, USA, 1999, p. 1150-1157
- [38] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 27, 2005, No. 10, p. 1615-1630
- [39] Lowe D. G. Distinctive Image Features from Scale-Invariant Keypoints[J]. International Journal of Computer Vision, 2004, 60 (2):91-110
- [40] Rosenholm D. Multi-Point matching using the least-squares technique for evaluation of three-dimensional models [J]. Photogrammetric Engineering & Remote Sensing, 1987, 53 (6): 621-626
- [41] 蔡健荣, 范军, 李玉良. 立体视觉系统标定及成熟果实定位. 农机化研究, 2007, 11 (11): 38-43
- [42] 陈立强, 刘忠, 田晓东. 基于边缘特征的水下目标分类识别算法. 自动化技术与应用, 2007, 26 (8): 77-79

## 致 谢

值此论文即将完成之际,在这里本生首先向导师魏志强教授致以最诚挚的谢意。每一次听魏志强对课题的讲解都使我多一回新的启迪和领悟,感谢魏志强在过去的两年多里对我的悉心栽培和孜孜教诲。魏志强对事业坚定、执着、永无止境的追求精神以及对自己与学生的高标准严要求,必将激励我在今后的人生道路上继续奋勇向前!

感谢殷波老师两年多来对我的指导和关心。在殷波老师的带领下,课题组才能严格按时间进度完成各部分的研究进展。殷波老师实事求是的研究精神,深厚的学术积累都给了我莫大的激励和鼓舞,与殷波老师每个周六的交流讨论解决了我做课题过程中遇到的一个又一个难题。

感谢我的父母,他们的理解、支持与鼓励使我有勇气面对碰到的每一个困难,他们全心的付出才使得我有时间和精力完成学业。他们的始终如一的爱给了我莫大的力量,让我不管遇到什么苦难,都能够鼓起勇气去克服,去战胜。

最后,感谢百忙之中抽出时间审阅论文的各位老师,感谢所有曾经关心和帮助过我的人们。



## 个人简历、在学期间发表的学术论文

### 个人简历

1984 年 12 月 4 日出生于山东省招远市。

2003 年 9 月考入中国海洋大学信息科学与工程学院通信工程专业, 2007 年 7 月本科毕业并获得工学学士学位。

2007 年 9 月保送中国海洋大学信息科学与工程学院通信与信息系统专业攻读硕士学位至今。

### 发表的学术论文

- [1] Weiliang Wang, Zhiqiang Wei, Bo Yin. Research on Binocular Stereo Matching Algorithm Based on Feature Points. The International Conference on Information Engineering and Computer Science (ICIECS2009). 2009, volume III:1598-1601
- [2] 董艳, 滕学志, 王伟良. 二维条码技术在超市领域的应用. 中国新技术新产品. 已录用



