

分类号: TP391

单位代码: 11407

密 级: 公开

学 号: 20157137



北方民族大学

硕士学位论文

基于 Webots 的家庭服务机器人 物体识别技术研究

Research on Object Recognition Technology
of Home Service Robot Based on Webots

学 位 申 请 人: 靳晓英

指 导 教 师: 马 行

申 请 的 学 位 门 类: 工学硕士

专 业 名 称: 电路与系统

研 究 方 向: 机器视觉与图像处理

所 在 学 院: 电气信息工程学院

论 文 完 成 日 期: 2018 年 3 月 16 日

独创性声明

本人声明所呈交的论文是我个人在导师指导下进行的研究工作及取得的科研成果。尽我所知，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，也不包含为获得北方民族大学或其它教育机构的学位或证书而使用过的材料。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了谢意。

研究生签名：靳晓英

时间：2018年5月28日

关于论文使用授权的说明

本人完全了解北方民族大学有关保留、使用学位论文的规定，即：学校有权保留送交论文的复印件和磁盘，允许论文被查阅和借阅，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。同意北方民族大学可以用不同方式在不同媒体上发表、传播学位论文的全部或部分内容。

(保密的学位论文在解密后应遵守此协议)

研究生签名：靳晓英

时间：2018年5月28日

导师签名：马红

时间：2018年5月28日

摘 要

在科技的发展, 人类需求的不断提升和社会老龄化等多方面的推动下, 服务机器人产业化势在必行。虽然许多的国内外学者对此进行了长期的研究, 但要实现机器人切实有效的帮助人类处理日常生活中面临的诸多问题还有一定的距离。物体识别便是其中的难点之一。主要原因在于基于机器视觉的物体识别要求识别精度和识别速度同时满足; 机器人面对的现实环境复杂多变, 干扰因素较多; 图像特征的选取会对识别的精度造成一定的影响。

本文以家庭环境作为机器人的识别场景, 针对上述难题进行研究设计, 主要完成了以下几方面的工作:

(1) 采用 SURF 算法和 SVM 分类器相结合的方法进行图像特征的提取和训练。在此过程中针对 SURF 算法提取的特征维度较高和数目不定的问题提出运用 PCA 和 Bag-of-visual-word 算法对其进行降维和聚类的处理。即提取出最具代表性的主成分后针对其特征数目不确定且 PCA 降维存在的缺少类别信息的特点运用 Bag-of-visual-word 算法对其进行聚类处理, 在弥补 PCA 算法不足的同时使 SURF 算法提取的特征数确定为一个固定的值, 从而便于后续 SVM 对特征的训练。

(2) 针对家庭场景的复杂性对机器人识别造成的影响, 本文提出将环境背景作为样本图进行训练的同时运用图像金字塔对测试图进行分割处理, 即对测试图进行分块识别, 以此来降低复杂背景的影响。

(3) 单一视角提取的 2D 图像不可避免的会造成信息的损失, 为进一步提升机器人识别的精度, 本文提出运用多视角的思想, 对目标物体多个角度提取图像以此来增加机器人对目标物体信息的获取。

(4) 本文主要运用 Webots 机器人仿真开发平台进行实验测试。在此平台上完成了对 OpenCV 的配置, 家庭场景和机器人的建模, 并运用 c++ 进行编程, 实现了在该平台上对机器人的运动控制和目标物体的识别。

实验结果表明, 本文的设计方案可以有效降低复杂环境和存在遮挡的情况对目标物体识别的影响, 在确保识别精度的同时能够保证一定的实时性。由于对家庭环

境的建模做到了尽量真实还原现实场景，因此，本文的研究设计具有一定的可行性和实用性。

关键词：PCA 降维，聚类，图像金字塔分割，多视角，Webots，SURF 算法

Abstract

With the development of science and technology, the continuous improvement of human needs and the ageing of society, the industrialization of service robots is imperative. Although many domestic and foreign scholars have conducted long-term research on this, there are still some distances to be achieved in order for robots to effectively and effectively help humans deal with the many problems that they face in daily life. Object recognition is one of the difficulties. The main reason is that Machine vision-based object recognition requires that the recognition accuracy and recognition speed be satisfied at the same time; The realistic environment faced by the robot is complex and changeable, with many interference factors. And the selection of image features will also have a certain impact on it.

This article takes the family environment as the recognition scene of the robot, carries on the research design to the above difficult problem, mainly accomplished the following several aspects of work:

(1) The SURF algorithm and SVM classifier are combined to extract and train image features. In this process, the problem of high and indeterminate feature dimensions extracted by the SURF algorithm is proposed. The PCA and Bag-of-visual-word algorithms are used to reduce the dimension and cluster. After extracting the most representative principal component, the characteristic of lacking category information whose feature number is uncertain and the PCA dimension is reduced is clustered by Bag-of-visual-word algorithm, which is used to make up for the deficiency of PCA algorithm. At the same time, the feature number extracted by the SURF algorithm is determined to be a fixed value, so as to facilitate the training of features by subsequent SVMs.

(2) In view of the influence of the complexity of family scenes on the robot recognition, this paper proposes to use the image of the environment as a sample map and use the image pyramid segmentation to segment the test chart. That is, the test chart is divided into blocks to reduce complexity. The effect of the background.

(3) The 2D image extracted from a single perspective will inevitably result in loss of information. In order to further improve the accuracy of robot recognition, this paper proposes to use multi-perspective ideas to extract images from multiple angles of the target object to increase the robot's acquisition of target object information.

(4) This article mainly uses Webots robot simulation development platform for experimental testing. On this platform, the configuration of OpenCV, family scenes and robots were modeled, and programming was done using c++. The robot's motion control and target object recognition were realized on this platform.

The experimental results show that the design of this paper can effectively reduce the impact of complex environments and occlusion situations on the recognition of target objects, and can ensure certain real-time performance while ensuring the recognition accuracy. Because the modeling of the home environment is done as realistic as possible to restore the real scene, therefore, the research design of this article has a certain degree of feasibility and practicality.

Keywords: PCA dimensionality reduction, clustering , image pyramid segmentation, multi-views , Webots , SURF

目 录

第一章 绪论.....	- 1 -
1.1 研究背景与意义.....	- 1 -
1.2 国外服务机器人研究现状与发展趋势.....	- 2 -
1.3 国内服务机器人研究现状与发展趋势.....	- 5 -
1.4 本文的主要研究内容.....	- 7 -
1.5 本文章节安排.....	- 7 -
第二章 物体识别相关内容.....	- 9 -
2.1 物体识别的应用范围.....	- 9 -
2.2 物体识别的制约因素.....	- 9 -
2.3 物体识别存在的问题.....	- 10 -
2.4 图像特征.....	- 11 -
2.5 物体识别的方法.....	- 12 -
2.6 本章小结.....	- 14 -
第三章 基于 Webots 的模型设计.....	- 16 -
3.1 Webots 开发平台概述.....	- 16 -
3.2 Webots 仿真建模.....	- 18 -
3.2.1 场景建模.....	- 18 -
3.2.2 机器人建模.....	- 19 -
3.3 超级机器人.....	- 22 -
3.4 本章小结.....	- 24 -
第四章 特征提取与处理.....	- 25 -
4.1 基于 SURF 算子提取特征.....	- 25 -
4.2 特征处理.....	- 29 -
4.2.1 基于 PCA 算法的特征降维.....	- 30 -

4.2.2 基于 Bag-of-Visual-Word 的特征聚类.....	- 32 -
4.3 基于 SVM 的特征训练.....	- 34 -
4.4 本章小结.....	- 35 -
第五章 实验设计与结果分析.....	- 36 -
5.1 搭建仿真环境.....	- 36 -
5.2 实验准备.....	- 37 -
5.2.1 样本图采集.....	- 37 -
5.2.2 机器人运动控制.....	- 37 -
5.2.3 PCA 降维处理.....	- 39 -
5.2.4 聚类处理.....	- 40 -
5.2.5 图像金字塔分割.....	- 41 -
5.3 实验结果测试与分析.....	- 41 -
5.4 本章小结.....	- 45 -
第六章 总结与展望.....	- 46 -
6.1 总结.....	- 46 -
6.2 展望.....	- 46 -
参考文献.....	- 48 -
致谢.....	- 52 -
个人简介.....	- 53 -

第一章 绪论

1.1 研究背景与意义

《机器人产业发展规划(2016-2020 年)》中指出^[1]：机器人作为高新制造业的关键的支撑装备之一，对改善人类的生活方式有着至关重要的作用。无论面向制造行业的工业型机器人，还是面向非制造业的服务型机器人，其相关的研究和实际的应用水平直接反映出一个国家的科技创新和高端制造的发展程度。对我国而言，重点开发机器人行业对推动我国工业的转型升级、提升中国制造的竞争优势、加快建设制造强国、改善人民的生活水平等方面均有重大而深远的意义。

而随着机器人智能领域相关技术的迅猛发展，机器人的应用领域也随之不断的扩展。机器人的研究早已不再仅仅满足于科学实验，而是立足于满足人们生产生活等方面的需求。机器人的适用环境也从静态的工场延伸至室外，海底，甚至是变化复杂，干扰因素较多的室内环境中^[2]。

由于社会人口老龄问题的不断加剧以及人们生活标准的逐步提升，服务型机器人特别是家用型服务机器人的需求更加与日俱增。中国产业信息网调查显示：在 2010-2016 年间，个人/家庭服务机器人的销量已从 215 万台增长到 690.2 万台^[3]。国际机器人联合会(International Federation of Roboyics, 简称 IFR)预计：在 2016-2019 年，全球的个人/家庭服务机器人将会迎来自身发展的高峰。2016-2019 年全球的个人/家庭服务机器人的累计销售量将增长至 4200 万台，其累计的销售额将达到约 223 亿美元^[3]。《机器人产业发展规划(2016-2020 年)》亦明确指出^[1]：家庭服务型机器人实现商品化是今后机器人产业发展的重要任务之一。今后将着力推动专业的服务型机器人系列化，个人/家用服务型机器人实现商品化。不仅如此，当前世界各国都致力于推进服务机器人产业的发展，日本和美国为其中的佼佼者。而相较于发达国家，我国的机器人产业的起步较晚，所以仍然存在一定的差距。但随着家庭服务型机器人产业时代的来临，对我国而言，与其说是挑战，不如说是机遇。

科技的发展使得机器人越来越智能，功能也更加的强大。但服务机器人要达到准确，实时的帮助人类解决实际的需求还面临着很多的问题。实现目标物体的准确识别便是其中之一，也是机器人达到能够切实解决日常生活中大小问题水平的不可缺少的环节之一。机器视觉作为人工智能研究的核心问题，其任务是让计算机达到和人类匹配的视觉能力^[4]。机器视觉的主要任务一般包括：识别、运动分析、场景重建和图像恢复四个方面^[5]。物体识别是机器视觉的主要任务之一，其目标是将图像和视频中包含的物体进行分类，属于核心问题，在其中发挥着至关重要

的作用。其原因主要是因为虽然人类可以轻而易举的识别出任意图像中的所有物体，即使物体的形态、观察视角、尺寸等发生变化，甚至是存在旋转或者局部存在遮挡的情况。但这对机器视觉而言是个很大的挑战。因此，它也是众多实际应用领域中难以突破的关键所在。目前，较为理想的方案也仅仅局限于设定某种特定情况下的识别，诸如：限定物体的形状、纹理、光线等的情况^[6]。广义下的物体识别，即在任何环境中识别出任意的物体是其最核心的问题，也是难以突破的难题之一^[7]。虽然全球诸如剑桥、麻省理工、斯坦福、耶鲁、普林斯顿等众多拥有先进技术和实验装备的研究机构对此已研究多年，广义下的物体识别依然远未得到理想的结果。其主要的原因在于根据人的视觉功能占用最少 60% 的人脑资源这一事实，机器视觉可能是等价于“让计算机拥有人类智能”的问题，即 AI-complete (Artificial Intelligence Complete, 人工智能完全)^[7]。由于物体识别是机器视觉的核心问题之一，因此，有人认为物体识别是等价于“人工智能领域最难的问题”的问题，即 AI-Hard (Artificial Intelligence Hard, 人工智能困难问题)^[7]。从机器学习的角度来说，物体识别的问题在一定程度上是可以实现的。准确来说，实现一个能够满足实际需要的识别系统是可行的。只要能够提取出适当的图像特征，准确的描述目标物体和适当的分类器对其进行训练。对家用服务机器人而言，由于家庭环境相对复杂，目前，物体识别仍然受到光照、拍摄角度、摆放姿势、复杂背景甚至是遮挡的影响。因此，机器人实现目标物体的准确识别仍然面临着较大的挑战。

仿真技术是机器人理论研究的一个重要组成部分。在机器人控制研究中具有非常重要的意义^[8]。主要在于通过仿真不仅可以有效快速的对现实环境进行复现和变换，而且可以很方便的对机器人的控制算法进行快速的改进。从而很好的避免测试和改进的过程中对机器人造成的不必要的耗损。兼具了成本低，效率高和灵活性好的特点的同时为发展智能机器人以及其它相关领域的课题提供了有效的理论依据和实践指导。

1.2 国外服务机器人研究现状与发展趋势

自 20 世纪 50 年代机器人诞生以来，迄今为止，机器人已从示教再现型的一代机器人，具备感知能力的二代机器人发展至智能型的三代机器人。服务机器人作为第三代机器人中的典型代表，其逐步向智能化与系统化的方向发展，应用领域也逐步向助老助残，家用服务，特种服务等方向扩展。

目前，世界上至少有 48 个国家致力于机器人领域的研究。其中，已有 25 个国家涉及服务类机器人的研究^[9]。其中，日本、北美和欧洲已将 7 类总计 40 多种服务类型的机器人投入实验和半商业化的应用当中。在处于发展前列的国家当中，西方以美国、德国和法国为代表，亚洲以日本和韩国为代表。

美国作为世界范围内最早涉足机器人领域的国家，尽管早期一直专注于理论的研究，但从

20 世纪 80 年代开始致力于应用型机器人的开发。立足市场需求，研制出众多具备先进技术和广泛用途的服务型机器人。2009 年 5 月，美国发布的《机器人技术路线图:从互联网到机器人》进一步拓宽了机器人的应用领域^[4]。

美国 iRobot 公司一直致力于实用机器人的研发，其研发的机器人种类众多，包括太空探测、搜救、国防安全和家用扫地机器人等各个领域。从 2011 年开始，研发了具有视频协作功能的高效的家庭清洁类机器人。实现了将导航、移动、操纵和人工智能领域的高新技术与家庭生活需求完美的融合。其 2015 年研发的 Roomba 980 扫地机器人采用了先进的导航系统，并与 VSLAM 视觉运算处理技术相结合。可以对家庭环境的地图进行绘制和全景规划导航，从而实现更智能和全面的清扫。在 2016 年推出的 Braava jet 240 喷水式擦地机器人，拥有干擦、湿抹、湿擦三种模式，适用于不同的清扫需求，拥有智能导航和多种传感器，可以规划记录清洁路径，实现全面智能的清洁。图 1-1 分别为 Roomba 980 扫地机器人和 Braava jet 240 喷水式擦地机器人。

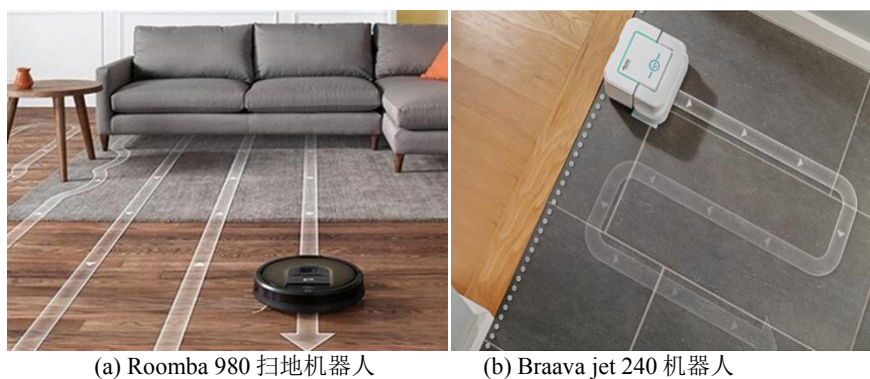


图 1-1 iRobot 清洁型机器人

德国和法国作为欧洲主要的发达国家之一，其对于机器人的研究始于 20 世纪 70 年代。

以严谨著称的德国，其服务机器人领域的研究和发展的水平获得了世界范围内的一致认可，处于公认的领先地位。其中机器人保姆 Care-O-Bot3，在正确识别日常生活用品的同时，能够避免对主人造成人身伤害，同时，它还对声音和手势的识别能力，能够接收语音和手势的命令。德国的独臂家政服务机器人，凭借其适当的型体，灵动的动作，高度的智能和良好的学习能力而成为了家庭当中的好帮手。

法国一直以来都十分重视机器人相关技术的研究，目前已形成了完整的一系列相关科学体系，无论在机器人数量和机器人技术水平和应用领域上都处于世界的领先地位。目前，致力于机器人应用领域的研究。

应用于养老院的 Zora 机器人是法国研制的 Nao 机器人系列中的一款拟人机器人，能够很好的服务老年人的生活所需，是法国首次将服务型的机器人正式应用于现实生活当中。在实现零突破的同时，更为满足老龄化社会的需求提供了可能。Care-O-Bot3 和 Zora 机器人如图 1-2 所示：

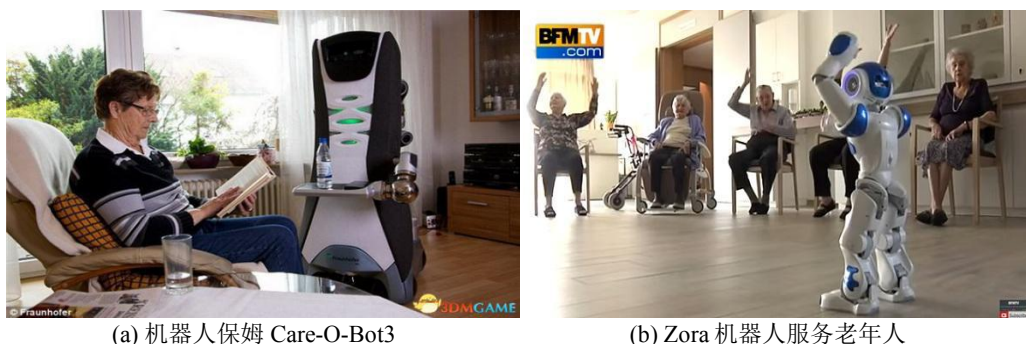


图 1-2 德国与法国研制的机器人

鉴于服务型机器人的应用领域的不断扩展，功能的不断提升和完善，欧洲许多其他的发达国家也踊跃的加入到此行业中来。

日本一直以来都十分的重视机器人产业的发展，其中以服务型机器人尤为突出，取得了许多先进的成果。无论是技术还是制造方面均在全球范围的相关领域产生了巨大的影响。而这离不开日本政府对产业的重视，更加离不开日本全社会人口的老龄化危机。相关资料显示，2005 年，日本的机器人产业的产值已在其国民经济中占有超过 7601 亿日元，而且在过去的 10 年间，产值一直处于上升的状态，截止 2015 年，机器人行业的总产值成功提升至 1 万亿日元。据相关人员估计，到 2020 年，以服务机器人为代表的日本机器人产业的产值将会达到 2.8 万亿日元。由上述资料不难看出，服务型机器人产业在日本的经济体系中占据非常重要的地位。

由日本本田肌研工业株式会社研发的 ASIMO 机器人是一款仿人类型的机器人。这款机器人通过精准模仿人类动作来实现帮助人类，特别是行动不便者的目的。配备视觉、听觉传感器，可以实现人脸、语音的识别，同时可以自主避开人群，避免碰撞发生。现在，ASIMO 不仅能自由行走，上下楼梯，而且还可以灵活的玩足球和倒水沏茶。目前，日本的 IBM 等企业已将其投入接待岗位。

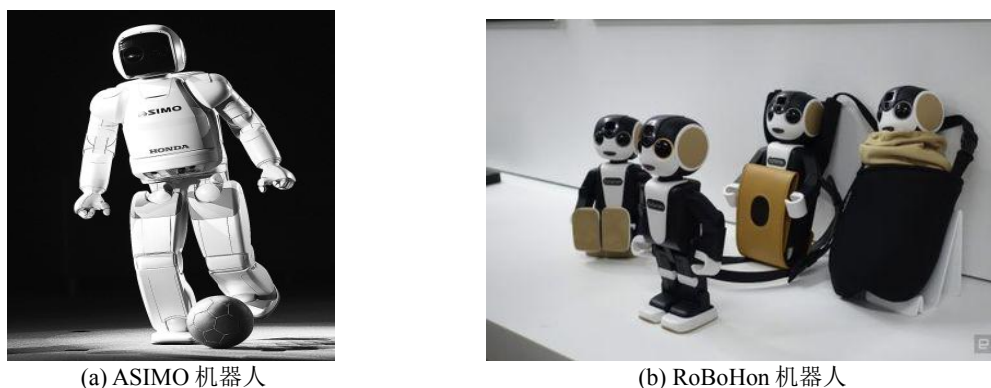


图 1-3 日本研发的机器人

RoBoHon 是夏普推出的一款拥有智能手机功能的移动型机器人。它是一款口袋型机器人，方便携带。可以完成双足行走和跳舞的动作。机身配置有摄像机和投影仪，可以实现拍照和投

影的功能。支持人脸和语音的识别,拥有打电话,读邮件和对登陆过的人进行识别的功能。ASIMO 和 RoBoHon 机器人如图 1-3 所示。

韩国将机器人产业作为其经济增长的重点发展项目,更将其列为未来重点发展的 10 大产业之一。

由韩国 LG 公司研发的 CLOi 机器人目前在韩国的仁川国际机场投入使用,可以作为为旅客寻找登机口的引路机器人还可以作为清扫地面的清洁机器人。机身上的显示屏可以为旅客提供航班信、机场地图、天气等信息,而且可以识别韩语、英语、汉语和日语四种语言。CLOi 机器人如图 1-4 所示:

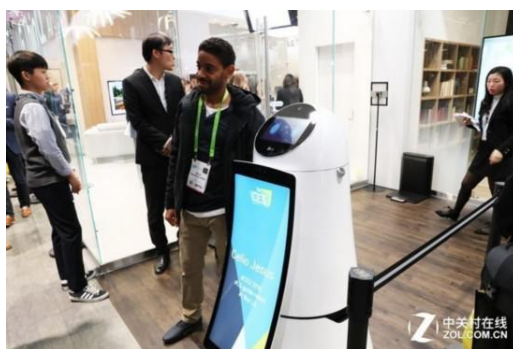


图 1-4 CLOi 机器人

1.3 国内服务机器人研究现状与发展趋势

我国的机器人研发开始于 20 世纪 70 年代。近年来,在相关政策的支持及社会需求的推动下,我国的机器人产业发展迅速^[10]。其在科学考察、医疗康复、教育娱乐和家庭服务等领域已获得了一系列的科研成果并投入了使用。当前,社会老龄化、医疗康复、救援预警、公共安全、教育娱乐等领域对服务型机器人的需求均呈现出迅速发展的趋向。

目前,中国的家用型服务机器人主要包括吸尘器、教育、娱乐和保安等多种类型的机器人。科沃斯是我国家庭服务类机器人的专业制造者。迄今为止创造了包括地面清洁机器人地宝、自动擦窗机器人窗宝、空气净化机器人沁宝、机器人管家亲宝在内的多种家用型机器人,专业从事家庭服务类机器人的研发、设计、制造和销售。其 2015 年研发的新品地宝朵朵通过专用的抹布和 wifi 智能控制功能来实现扫拖一体,从根本上解决了家庭地面清洁的问题。

2017 年 2 月,科沃斯生产的公共服务型机器人在杭州西湖区国地税联合自助办税服务厅正式投入使用^[11]。担任导税员,承担引导纳税人办税的工作,满足了纳税人对不同纳税服务的需求。成为指导纳税人办理业务的重要平台。

该机器人不仅可以依据纳税人所办理的业务指导纳税人去相应的业务办理区,还可以对纳税人进行自主办理时遇到的问题通过语音,视频等方式指导其运用自助终端来进行解决。此

外，该机器人还担当税宣大使，在等候区以向纳税人推送国税二维码和宣传视频的方式，达到服务和宣传的目的。



(a) 公共服务型机器人

(b) 扫地机器人 DJ35

图 1-5 科沃斯研发的机器人

近期，科沃斯推出升级版的扫地机器人 DJ35，在延续上一代优良清洁系统蓝鲸清洁系统 2.0 的同时，采用全新 Smart Eye 视觉导航技术，将 SLAM 算法与摄像头采集的 166° 环境图像相结合，使机器人实现精准定位和智能的导航移动，从而实现了更加全面和高效的清洁。公共服务型机器人和扫地机器人 DJ35 如图 1-5 所示。

海尔集团围绕“海尔智慧生活大脑”研发的 Ubot 机器人是全球第一款真正意义上的家用型智能机器人。集成了“听、说、嗅、知、动”五大拟人类的感知功能，且拥有强大的语音、物体、人脸的识别功能。Ubot 机器人如图 1-6 所示：



图 1-6 Ubot 机器人

Ubot 可以实时的感知家庭环境的温度，湿度，光照，亮度和安全设备的工作状态等的情况，并根据用户的命令作出相应的操作的同时，还可以通过学习和对大数据的分析来感知用户的习惯和喜好，从而为用户提供相应的服务。这是海尔自主研发的具备国际先进人工智能技术的新一代机器人，预示着其步入了自然交互感知的新时代^[12]。

艾富通科技最新研制的防盗类智能保安型机器人是一款拥有拟人类听，说，看功能的智能机器人。集聚了包括物联网、智能传感、云平台大数据在内的众多先进科技，支持实时报警、

视频监控、远程打击和隐私保护等多种功能，为“机器人”在防盗安全领域的实现提供了可能^[13]。智能保安型机器人如图 1-7 所示：



图 1-7 智能保安型机器人

当前，我国的服务机器人行业还处在初级阶段。大部分的产品目前仍处于研发实验的时期。相较发达国家而言，投放使用的产品依然很少，距离机器人的产业化依然有一段距离^[14]。

《机器人产业发展规划(2016-2020 年)》^[1]指出:从 2016 年起至 2020 年的发展目标是在 2020 年机器人产业规模持续增长；服务型机器人的年销售额至少达到 300 亿；持续提升相关技术水平，特别是医疗健康、家庭服务、反恐防暴、救灾救援、科学研究等领域的相关技术接近国际水平；扩展服务机器人的应用领域，围绕助老助残、家庭服务、医疗康复、救援救灾、能源安全、公共安全、重大科学研究等领域，培育智慧生活、现代服务、特殊作业等方面的需求^[1]。

1.4 本文的主要研究内容

本文针对目前机器人在家庭场景中对目标物体进行识别存在的问题进行研究和实验设计，并在 Webots 平台上进行仿真测试。主要完成以下工作：

1.在选择 SURF 算法提取图像特征和 SVM 分类器训练的基础上针对 SURF 算法的高维和数目不定的特点，提出运用 PCA 算法和 Bag-of-Visual-Word 算法对其进行特征处理，实现降维和聚类，便于 SVM 训练的同时，降低了计算的复杂程度，提升了计算效率。

2.在实验设计方面，将环境作为样本之一的同时提出运用图像金字塔分割和多视角的方法，以此降低复杂环境和遮挡对识别的影响。

3.在 Webots 平台上完成对场景和机器人模型的搭建和 OpenCV 的编译配置，使机器人在此平台上实现自主避障，图片采集和目标物体的识别。

1.5 本文章节安排

本文分为六个章节进行阐述，每章节的安排如下：

第一章 绪论。本章主要对服务机器人国内外发展现状和今后的发展趋势作了相关介绍，并简要概述了本文主要研究内容和章节安排。

第二章 物体识别相关内容。本章主要介绍物体识别相关方面的内容包括了物体识别目前存在的问题、识别的方法等多个方面。为后续特征的选取，处理和实验的设计提供了基础。

第三章 基于 Webots 的模型设计。本章首先对实验平台 Webots 及其使用范围和语言环境做简要概述。之后对场景模型和机器人模型建模设计的节点和方法进行阐述。完成实验场景和机器人模型的基本搭建。

第四章 特征提取与处理。本章详细介绍了 SURF 算法的原理，同时提出运用 PCA 和 Bag-of-Visual-Word 算法对特征进行降维和聚类的处理，并详细介绍了其处理的原理和参数选取的方法。在保证实验精度的同时简化了计算的复杂程度，提升了识别的效率。

第五章 实验设计与结果分析。本章详细阐述了实验的设计方案，并进行了相应的实验测试来验证该设计的可行性和实时性。

第六章 总结与展望。本章主要对本文的工作进行了总结，同时分析其不足之处，为后面进一步的研究工作进行展望。

第二章 物体识别相关内容

2.1 物体识别的应用范围

随着机器视觉的广泛应用，物体识别的应用领域也不断扩展^[15]，涵盖了以下几个方面：

1. 产品检测。可以实现对产品质量的检测，对存在瑕疵和裂隙的产品进行筛检。从而降低次品率。
2. 产品加工。可以实现对精密仪器的安装、生产和产品后期的包装和搬运，从而提升生产效率。
3. 航拍监测。对获取的航拍图和视频进行处理，获取所监测的目标事物和相关的环境，从而掌握实时信息。
4. 视觉导航。通过视觉系统获取的图像信息，对不同的参照物进行判断，从而感知周围环境，做出最佳的路径规划，实现自主高效的运动。
5. 智能家居。辅助机器人实现对特定物体的抓取，家庭环境的监测和家用电器的控制。

2.2 物体识别的制约因素

虽然，国内外的许多学者已对物体识别的相关问题进行了众多的研究^[16]，但其发展并没有达到非常成熟的地步，大规模的应用仍然尚未出现。究其原因是因为其相关的基础研究的制约，物体识别便是其中之一。主要原因在于：

1. 人类的视觉成像的原理和过程与机器视觉是不同的。人类是通过视网膜得到的二维成像来识别现实世界的三维物体，而机器视觉则是相反的过程。即从这种灰度阵列的平面图像中重新构造三维的现实世界^[17]。由于机器视觉获取的图像只能呈现真实世界的二维信息，因此，在采集图像的过程中不可避免的会造成信息的丢失。如何运用二维信息来理解现实世界的物体、是否能够还原原始三维信息仍未得到确切的答案。而且人类的视觉系统属于双目立体型视觉系统，而目前机器人的视觉系统主要有三种类型：单目视觉、立体视觉和全景视觉^[18]。三类系统均具有各自的弊端，不能很好的模拟人类的视觉系统。因此系统误差是不可避免的。

2. 与传统的计算机视觉相比，机器视觉在实现的目的上有很大的不同。机器视觉主要是让机器人在其相应的视觉系统的辅助下实现对其周围环境的感知。因此，该系统在要求高精度的同时，还需要满足一定的实时性的要求^[19]。

3. 现实世界很复杂，同样类型的物体会以不同的颜色、纹理、形状出现，且会受到光照、拍摄角度、甚至是非目标物体的干扰(遮挡)，从而使得目标物体的识别面临更多的挑战。

以上原因均使得物体识别的实现变的更加困难，因此，与其相关的任一方面的进步和突破，均会对机器人产业化和应用领域的进一步加深与扩大带来极为深远的意义^[6]。

2.3 物体识别存在的问题

由第一章的介绍可知，目前，家用机器人产业已经得到了很大的发展。但是对于实现家用机器人扮演人类生活伴侣这一目标而言依然还有一段距离。要实现这一目标需要机器人拥有更高级的智能，并以此来实现同人类共同学习、成长、甚至是共同体验生活^[20]。而如今，机器人严格说来只是一件减轻人类工作的工具，而非伙伴。因此，虽然，目前物体识别的研究已经获得了许多可喜得成果，但要进一步推进其产业化的程度还面临着许多的干扰和困难^[21]。由于家庭场景的复杂多变的特点，机器人要实现在家庭环境中对特定物体的识别还面临着以下几方面因素的干扰：

1. 光照

物体的成像颜色对光照敏感，不同的光照下可能会出现像素值的变化，也可能产生阴影。无论哪种都在一定程度上提升了目标识别的难度，并对目标识别系统产生影响。家庭场景中的光照情况主要受到自然光和灯光的影响，会随着时间和光照角度的变化而变化，从而影响到物体的识别。

2. 摄像机视角

摄像机在获取目标物体时，由于拍摄视角的变换和相机的抖动等干扰的存在，有可能会造成物体成像的畸变以及尺度，角度等的变化，从而影响到对目标物体的识别。

3. 目标物体姿势的改变

现实环境中，目标物体的位置和姿势不是一成不变的，而是存在随机性。机器人在识别时首先要找到目标物体的位置，在算法方面则要求具有克服同一物体不同姿态的影响的功能。

4. 复杂背景

复杂背景一直以来都是是干扰物体识别性能的最为严重的因素之一。目标所处环境的复杂程度、变化程度会直接影响到最终目标识别的效率。在现实情况中，目标物体很多时候都处于复杂的背景当中。之所以称其为最为严重干扰因素，其原因在于复杂的背景会极大影响目标物体特征的提取和表示。因此如何降低背景因素的影响并从中识别出目标物体是目前相关研究的重点之一。

5. 遮挡

目标被遮挡同复杂环境一样，是现实环境中经常会出现的状况。也是目前研究的热点。遮挡的存在会造成目标特征的缺失，进而直接影响到最终识别的效果。迄今为止，大面积的遮挡问题依然没有得到很好的解决。

6. 高维特征

图像特征很多都具有高维特性，但其对于目标的识别的重要性和贡献率是不一样的。图像本质特征的提取能够保证识别结果的准确性，反之不显著的特征对于识别只起到辅助作用。而且高维的特征会严重干扰到识别的实时性^[22]。

上述问题的存在，均在不同程度上对目标识别相关技术的研究和突破造成了较大的干扰，因此，本文主要针对上述几方面特别是最后三项进行相应的研究和实验设计。

2.4 图像特征

物体识别一般依赖于特征的提取。其实质就是在图像的实时特征和目标参考特征之间寻找一种特定的对应关系的过程。图像特征是图像信息的本质属性，每一幅图像均有其自身的表征。有些是视觉系统能直观感受到的，如：边缘、亮度、色彩等，有些需通过变换或测量才能得到，如频谱、直方图等^[23]。由于目标识别系统要实现实时、可靠地在复杂环境中对目标物体进行识别，取决于能否在大量复杂的图像信息中找出最具代表性的特征，因而特征的提取是目标识别的关键所在。由于在特征提取时会受到实际环境和技术的制约，因此，提取的特征会受到干扰，这就增加了特征的选择和提取的复杂程度。从某种程度上讲特征的选择比分类方法的选择更加重要，因为它是分类器事先正确分类的第一步。好的特征具有较高的分离性，可以有效的代表特定的物体，从而为后续识别奠定良好的基础，大幅提升后续识别的效率和性能^[24]。

图像特征一般可以分为全局特征和局部特征。由字面含义可知，全局特征代表的是一幅完整图像的特征，而局部特征代表的是图像中某一部分的特征^[25]。常见的全局特征有颜色、纹理和形状等。全局特征在应用上不如局部特征方便，易受到所处环境变化的干扰。全局特征存在计算量大，匹配速度慢的缺点。对于目标物体畸变或存在局部遮挡的情况时，无法保证其识别的精度。而局部特征具有抗干扰性强、稳定性高的优点。不仅可以更有效的对于图像进行分类，而且对图像存在的旋转、平移、缩放、甚至仿射变换等均保持不变的特性。即局部特征可以有效地克服存在尺度，平移，旋转，光照等干扰的情况，保证良好的识别性能。同时对于目标物体存在局部遮挡时依然能达到良好的识别效果^[26]。因此相较于局部特征来讲全局特征具有很大的局限性。不能很好的满足图像分类的需要。目前，国内外比较主流的局部特征主要包括角点特征、BLOB 斑点特征、区域特征等^[27]。

1. 角点特征

目前，对于角点存在不同的定义。常见的定义为图像中周围灰度发生剧烈变化的点或者图像边缘曲率极大的点。角点对应于物体的拐点、T 型节点等。角点邻域中也包含了较多的图像信息量，同时也具有一定的稳定性。角点检测是基于局部特征的一种常用方法，目前，角点检测算法根据其采用角点类型的不同可以分为三类：基于灰度图像的、基于二值图像的以及基于轮

廓曲线的角点检测。由于存在局部遮挡的目标物体额图像中依然包含许多的角点特征, 因此, 角点是也存在遮挡的识别依据的一类重要的特征。常用的角点检测算法包括 HansP.Moravec^[28]提出的 Moravec 算法、Chris Harris 等^[29]提出的 Harris 算法和 Smith 和 Brady^[30]提出的 SUSAN 算法等;

由于角点特征在对目标物体存在遮挡的情况进行检测时会产生较多的干扰点, 因此存在一定的误匹配的情况。

2. 斑点特征

斑点是与周围的颜色和灰度存在差异的区间。相较角点而言, 斑点特征代表的是区域, 因此它拥有良好的稳定性和较强的抗噪声的特点。常用的斑点特征包括 SIFT 算法^[31]、PCA-SIFT 算法^[32]和 SURF 算法^[33]。PCA-SIFT 是针对 SIFT 特征维数过高而提出的算法, 但削弱了特征的特异性。SURF 算法是一种具备快速稳健特性的局部特征, 与 SIFT 特征相比, 在大幅提升计算速度的同时, 完整的保留了原始特征的特异性。因此更加适用于实时性要求高的场合^[34]。

目前, 斑点特征在目标物体存在遮挡的检测和识别的应用方面取得了较好的效果, 不仅具有实时性, 而且能够保证较好的鲁棒性。

3. 区域特征

区域特征首先对图像当中拥有某类相同性质的相元进行分类, 之后将相同性质的相元连接起来, 从而形成一个个的区域, 并对其进行检测。区域特征包含 MSER^[35]、Saliency region detector^{[36][37]}、IBR^{[38][39]}等。

区域特征适用于存在尺度变换、仿射变换甚至是局部遮挡的情况。能够保证一定的鲁棒性。但计算复杂程度高, 而且由于区域特征是针对目标区域当中的某块区域来进行检测的, 因此不适合大面积遮挡的情况。

综上所述, 基于局部特征的目标识别能够在一定程度上克服光照、尺度、旋转和遮挡对目标物体识别的影响, 而 SURF 算法在众多局部特征中更适合本文的研究方向。因此本文选择 SURF 算法来进行特征提取。

2.5 物体识别的方法

物体识别不仅在机器视觉的研究领域占据重要的地位, 同时也是模式识别的一个重要分支。主要原因在于模式识别的任务是对表示事物的不同的模式进行分类, 以实现替代人类自主完成对事物的分类和识别。是人工智能的重要组成部分。而物体识别是以图像模式为前提的。图像的模式识别过程主要包括以下四个部分: 图像的获取、预处理、特征提取和决策分类^[40]。其系统框图如图 2-1 所示:



图 2-1 模式识别的系统框图

各部分的功能如下：

(1) 图像获取

运用各种信息输入设备如摄像机等将采集的信息转化为能够被计算机识别的数字图像信息，为之后其他环节的处理做好准备。

(2) 图像预处理

采集的图像信息中会包含一些干扰因素包括噪声，抖动等，会对后续的认可造成干扰。因此在此特征提取前会对图像进行预处理，包括图像的平滑，变换，增强和恢复等。

(3) 特征提取

运用不同的算法对图像中的特征进行提取，以此来代表图像的最原始和最具代表性的属性。之后，对提取的特征进行数字化的表示，方便计算机对其进行分类识别处理。

(4) 决策分类

该环节主要是寻找一种最佳的判别函数来对提取的特征进行分类和识别。该步骤的效果与特征提取的好坏有密切的关系。

物体识别方法的分类标准不是唯一的，同一种方法在不同的分类标准下可能会存在交叉。根据使用特征的不同可以将识别的方法大致分为两类:基于统计的和基于结构的识别。除此之外，目前运用比较广泛的还有模糊识别和神经网络识别的方法。

1. 统计模式识别

统计模式识别是以决策理论为基础建立的。该方法主要是依据对提取特征的分类训练来实现的。是发展较早也比较成熟的一种方法。主要选取能够反映图像本质的特征。之后对其采用数字化的处理方式，使原始特征变换为适用于计算机处理的数字化信息。其步骤主要包括训练和分类。该模式识别的系统框图如图 2-2 所示：



图 2-2 统计模式识别的系统框图

2. 结构模式识别

结构模式识别又被称为句法方法或语言学方法。其选取的特征是反映识别对象各部分之间的结构关系的特征。其基本的思想是将一个完整的模式运用树状结构进行表示。即可由子模块的组合来表示，而子模块又可以运用其相应的多个子模块表示。处于最底层的子模块是模式基元。通常其被要求在可以紧密的反应结构关系的同时需支持非句法结构抽取。模式以一组基元和它们的组合关系来描述，称为模式描述语句。因此，结构模式识别方法能够运用小而简单的

基元与语法规则相结合的方法来表达大而复杂的特征。相较于基于统计的模式识别方法，其侧重的方面是不同的，基于统计的方法主要是针对图像的性质进行识别，而基于结构(句法)的识别方法则是针对对于识别对象各部分之间的关系的识别。由于该方法忽略了噪声的干扰，因此，其元素和结构关系具有一定的随机性。且目前的图像处理算法对其相关的提取还存在一定的困难，因此该方法在实用性方面存在不足。

3. 模糊模式识别

模糊模式识别可以看做是对统计识别和结构识别的完善，即可以实现对模糊事物的识别和判断。为简化识别系统的结构，通过采用模糊技术对机器识别系统进行设计，同时根据人脑对事物识别的思维逻辑，对人脑思维过程的近似模拟，以此来提升其对客观事物的分类和识别的性能。

4. 神经网络模式识别

神经网络的模式识别源于对动物神经系统的研究，借助硬件或软件构建大量的处理单元，各单元之间通过一定的模式实现互相连接的方式形成拓扑结构。从而实现对人神经系统结构和功能的模拟。

对模式识别方法的选择主要是由问题的性质来决定的。若识别内容复杂且包含明显结构类信息时，通常采用结构识别。若识别的对象相对而言在内容和结构上较为简洁，则会采用统计的识别方法。

基于统计的模式识别的分类通常是依据采用的方法的不同来进行。根据训练样本是否具有类别信息可将其分为有监督学习和无监督学习两种。由于这两种方法同时属于基于学习的物体检测。因此，基于统计的模式识别与基于学习的物体检测之间拥有一定的联系。基于学习的物体检测可以实现目标物体的检测和识别的同步。同时，对存在复杂背景影响的情况仍然具备良好的识别性能。

无监督学习型的物体检测方法是运用提取的样本的空间特征或颜色、纹理等来建立主题模型(topic model)，并运用该模型来对物体进行识别和检测^[41]。有监督学习的物体检测方法主要是通过通过对不同物体，多种视角的特征进行学习来实现分类的目的。首先对图像特征的提取，训练和分类形成一定的样本信息，之后以一定的搜索策略对测试图进行分区域判断，从而得出分类的结果^[41]。该区域中包含目标物体，则分类器的判断结果为是。反之，则不包含目标物体。从而得到检测物体的位置信息。

本文采用有监督的学习分类方法对训练样本进行分类训练。

2.6 本章小结

本章对物体识别目前的现状进行了简单的介绍。并对机器人物体识别目前具有较好性能的

图像特征和识别的方法进行阐述，最终确定使用 SURF 算法实现本文中对图像特征的提取，从而克服光照，尺度和旋转对识别的影响；采用有监督的学习分类方法对训练样本进行训练和分类。

第三章 基于 Webots 的模型设计

3.1 Webots 开发平台概述

目前, 有关机器人开发常用软件主要有:

1. ADAMS 动力学仿真软件。其支持精细的机械系统的仿真, 可以对包含多组数据的实验方案进行分析, 给出优化的工作性能^[42]。

2. Robotbuilder 动力学仿真软件。由 Rodenbaugh 研发的适用于机器人开发的动力学仿真软件, 可适用于 windows 系统, 其包含构建、仿真和回放三种模式, 目前已被俄亥俄州立大学用于机器人教学课程^[43];

3. Microsoft Robotics Studio 机器人仿真平台。此软件由微软开发, 具备多种功能, 而且拥有很强的扩展性, 支持图形化的数据进行逻辑运算^[44]。

4. ROBCAD 仿真软件。主要用于图形的设计和运算, 并输出机器人任一零部件的尺寸^[45]。

5. Open Dynamics Engine(ODE)开源物理引擎。由 Russell Smith 研发, 包括刚体动力学和碰撞检测两部分, 以满足用户的不同的需求, 实现包含各种功能的物理环境^[46]。

6. Pro/Engineer 建模软件。其最早提出了参数化的设计理念, 主要实现三维造型设计。并通过使用单一数据库来避免特征相关性的问题。

7. MATLAB 软件。主要实现建模和数值分析。适用于多种研究领域, 包括数值分析, 图像处理和控制系统设计等。

以上均为机器人研究领域常涉及到的软件, 均有其不同的侧重点^[47]。但机器人领域的研究需要多种技术相结合, 因此需要各个不同的领域的理论与技术的支持。Webots 开发平台可以很好的解决上述问题。并解决了不同软件之间的兼容问题, 从而大幅提升了研发的效率。

Webots 是由瑞士联邦技术研究院(Swiss Federal Institute of Technology)研制的一款集建模, 编程和模拟功能为一体的专业机器人模拟器^[48]。并由 Cyberbotics 公司开发和支持。在全球有超过 1 200 家公司, 大学和研究中心在使用。它拥有开源的物体引擎 ODE 和原型环境, 支持用户创立包含质量, 摩擦系数, 力等物体特征在内的 3D 虚拟世界。同时, 通过添加相应的节点, 用户可以在此环境中添加机器人模型, 从而实现对机器人相关功能的开发。Webots 支持机器人以多种运动形态出现, 包括轮式, 双足和飞行。同时, 用户可以根据自己的需要为机器人添加节点来实现特定的功能。例如传感器, 定位系统, 管理者, 发射器, 照相机和接收器等。最终, 为每个机器人编写相应功能的程序以实现预期的行为。不仅如此, Webots 中还提供了许多的接口, 方便真实机器人程序的优化。对机器人的运动控制主要通过编程来实现。Webots 支持包括 C、C++、Java、Python 和 MATLAB 等在内的多种语言, 便于使用者不同的开发需求。

Webots 机器人开发平台的界面主要包含四个窗口^[49]：A.场景树窗口、B.仿真运行窗口、C.程序编辑窗口、D.控制台窗口。如图 3-1 所示。

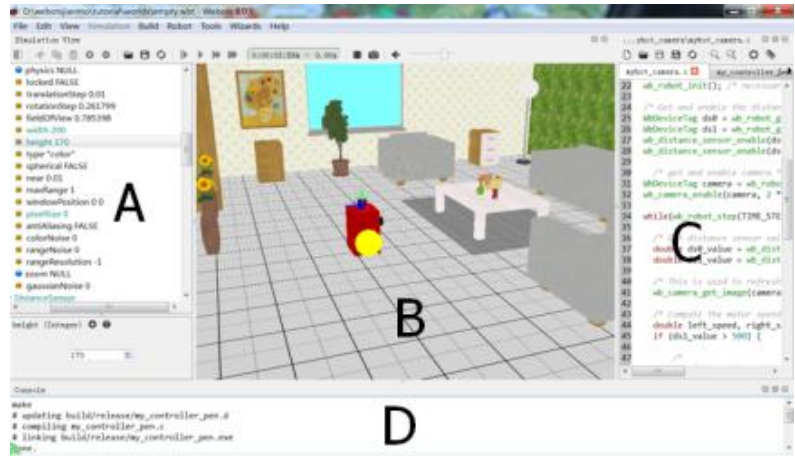


图 3-1 Webots 主窗口

在家庭环境下实现机器人的自主识别需要机器人具备三方面的功能：感知功能、运动控制功能和运算控制功能。由于 Webots 拥有完整的机器人开发环境，因此上述功能均可在该平台上得以实现。本文在 Webots 上主要实现家庭场景的建模，机器人的主体设计，各种设备的添加和参数的设置，之后进行控制程序的编写和实验的测试。主要的工作流程如图 3-2 所示。

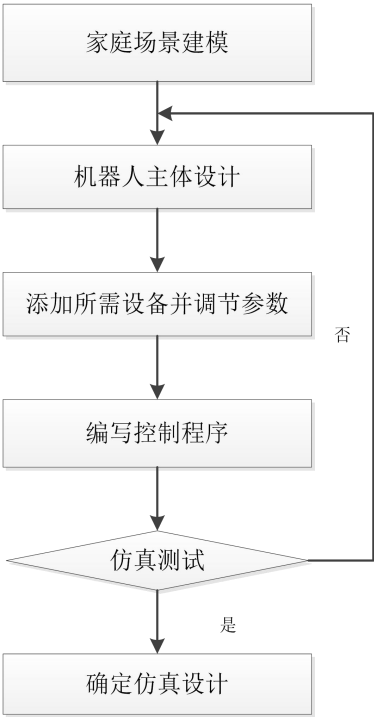


图 3-2 工作流程图

3.2 Webots 仿真建模

Webots 中模型的搭建主要通过节点选取及参数的设置和调节来实现。

3.2.1 场景建模

本文中家庭环境是机器人运行的主要环境。在仿真建模的过程中会考虑到现实环境中的各种影响因素，做到接近真实的模拟。因此家庭场景的模型选择的面积为 10m*10m。其中的家庭常见物品如墙，门，桌子，沙发等则按照现实中的大小进行选取。其模型如图 3-3 所示。



图 3-3 仿真世界中的模型

Webots 中提供 DirectionalLight 节点可以模拟无限远的光线(例如：太阳)。其定义的是定向光源，主要沿着平行于给定三维向量的光线照射。与 PointLight(点光源)不同，DirectionalLight 节点投射的光线不会随距离衰减。所以，模型中通过运用 DirectionalLight 节点和修改其中的位域值来模拟现实环境中的光源，同时显示阴影。光源模型如图 3-4^[50]所示：

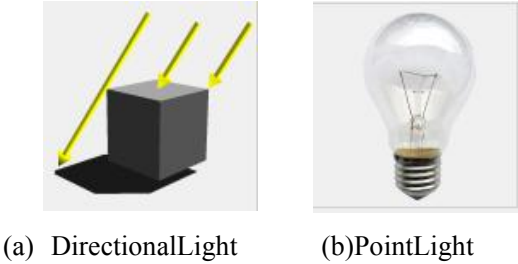


图 3-4 仿真世界中的光源模型

场景模型的仿真效果图如图 3-5 所示：



图 3-5 场景模型

3.2.2 机器人建模

本文中的机器人应具备在家庭环境中自主移动、避障、采图和识别的功能。因此，机器人模型应具备动力系统、感知系统、目标识别系统等。所以，本文中的机器人模型主要包括机身、差动轮、摄像机、距离传感器、GPS、笔。各器件的模型如图 3-6 所示。

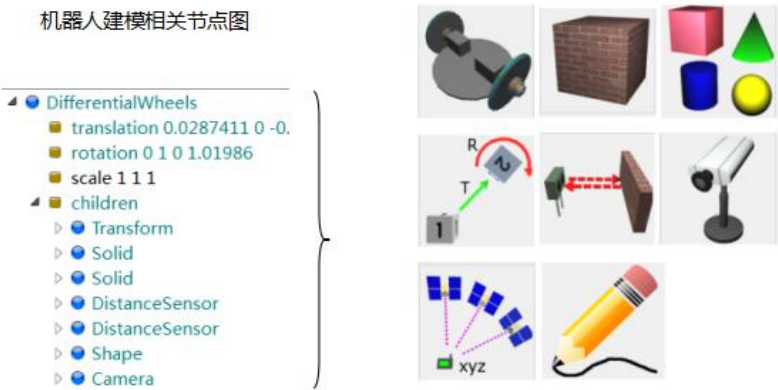


图 3-6 机器人相关节点模型

机器人模型如图 3-7 所示：

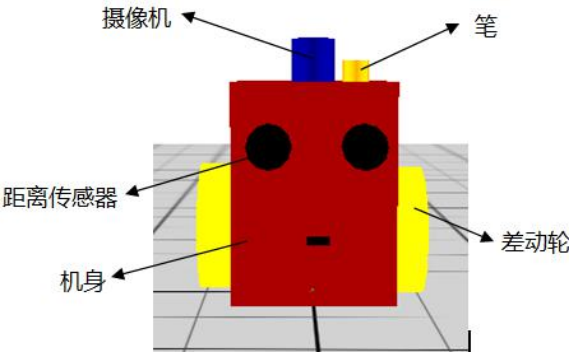


图 3-7 机器人模型

(1) 机身

对机器人机身的设计主要针对外观和形状两个方面来进行^[51]。外观的设计包含材质、纹理、颜色、形状结构等方面。Webots 的形状节点中支持正方形、长方形、胶囊体、圆柱体、圆锥体、球体等^[51]。对机身的设计主要通过对上述几个方面的节点参数进行设置和修改来实现。本文选取长方体作为机身的形状，颜色为红色。

(2) 差动轮

DifferentialWheels(差动轮)节点主要用来构造轮式机器人，为机器人提供差动转向功能。也可以作为子节点，此时机器人将实现自主运动。车轮必须是 Solid(实心)节点，且必须标注车轮的名字为“左车轮”和“右车轮”。Webots 中提供了 `wb_differential_wheels_set_speed()` 函数来实现对车轮旋转速度的设置和改变。

(3) 摄像机

Camera 节点用于实现对摄像机功能的模拟。Webots 中提供 "color", "range-finder" 和 "both" 三种类型的摄像机。"color" 相机允许从相机的 OpenGL 上下文中获取颜色信息。这个信息可以通过 `wb_camera_get_image()` 函数获得。本文主要采用此摄像机来完成对彩色图像的采集。相机节点 Camera 中包含的 `fieldofview` 节点表示的是照相机的水平视角，因此，若要调节相机的水平视角，只需改变 `fieldofview` 的值即可，从而找到最佳的拍摄尺寸。但需要注意的是，`fieldofview` 的取值在 $0 \sim \pi$ 弧度之间。同时，相机的水平视角，垂直视角以及相机的宽和高之间存在一定的对应关系。即相机的水平视角和垂直视角与相机的宽成反比，与相机的高成正比。所以，已知其中的三个值可以计算出另一个未知数。具体公式如公式 3-1 所示：

$$VerticalFOV = fieldofview * height / width \quad (3-1)$$

(4) 距离传感器

Webots 中提供 DistanceSensor 节点来模拟现实环境中的红外线传感器，声纳传感器，激光测距仪及普通传感器。通过获取障碍物与其所在环境中的所有传感器发出的射线之间的碰撞返回的参数值来实现其功能。其中，就距离传感器而言，其光线的分布呈 3D 锥形，开口的角度的调节可以通过改变 `aperture`(孔径)值来实现。图 3-8^[50]表示的是从一到十的光线分布图，其光线的分布要保证左右对称的同时，要确保其分布尽可能均匀。

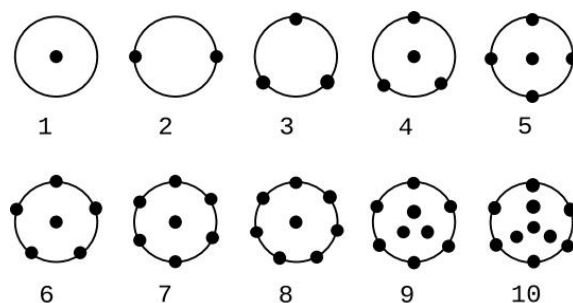


图 3-8 1 到 10 个传感器射线的预定义配置

其中，光线在数量上没有明确的限制，但在 Webots 中其性能会受到干扰。即其性能与光线数量成反比。

运用高斯分布求取每个传感器各个射线的权重的公式如公式 3-3 所示：

$$v_i = \exp\left(-\left(\frac{t_i}{a.g}\right)^2\right) \quad (3-2)$$

$$\omega_i = \frac{v_i}{\sum_{j=0}^n v_j} \quad (3-3)$$

其中 ω_i 表示的是第 i 条射线的权重， t_i 表示的是第 i 条射线与传感器轴线之间的夹角； a 代表的是传感器孔径角； g 代表的是高斯宽度， n 则是光线的数量。射线权重与射线数的关系图如图 3-9^[50]所示。

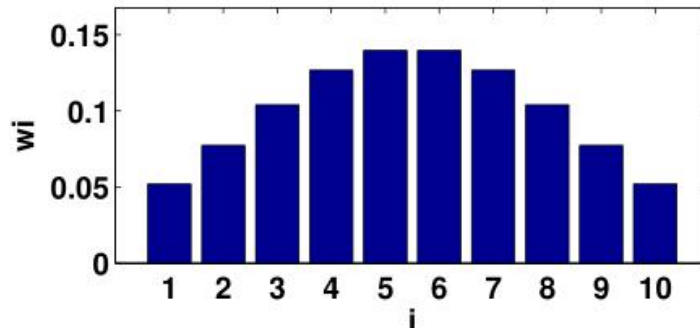


图 3-9 高斯宽度为 0.5 的 10 条射线的分布示例

由图可知，越靠近锥体中心的光线，其权重越大。若要改变其分布状态，可以通过调节 gaussian 值来实现。随着 gaussian Width 的增大，其分布会逐渐趋于平坦。

通过添加传感器的模型节点和其相应的节点值的调节设置，可以实现对其特定功能的模拟和控制。

DistanceSensor 节点中包含名为 lookuptable 的位域，用来指示传感器的返回值。即用于设置相关设备所需的响应曲线和噪声值，从而指示将 Webots 测量的射线交点距离映射至通过函数 `wb_distance_sensor_get_value()` 获取的响应值上的方法。具体的对应关系如图 3-10^[50]所示。

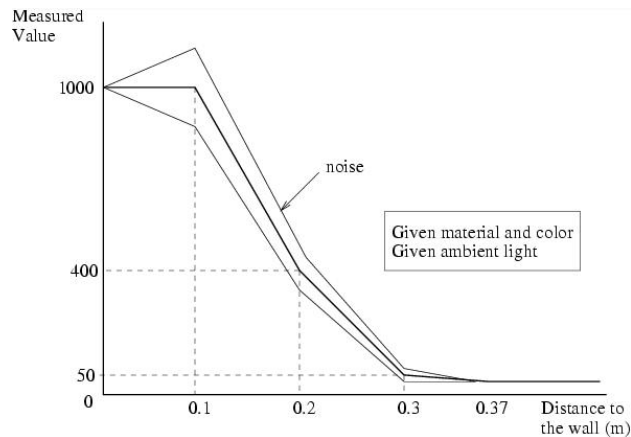


图 3-10 传感器和障碍物距离关系图

(5) GPS

GPS 节点用于模拟全球定位传感器(GPS)，该传感器可以从控制器程序获取关于其绝对位置的信息。该字段定义了像“卫星”或“激光”使用的 GPS 技术的类型。在 Webots 环境中，GPS 节点属于内部传感器，没有相应的实体模型。主要用来获取三个坐标轴方向的距离。在仿真运行时读取的 GPS 参数示意图如图 3-11 所示：

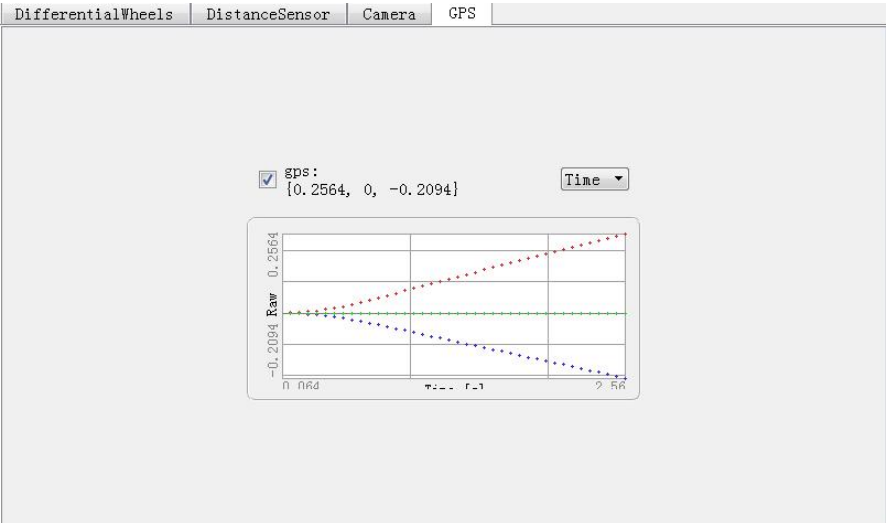


图 3-11 GPS 采集的数据

(6) 笔

用来显示机器人的运动轨迹，便于实验的观测。实验效果如图 3-12 所示：

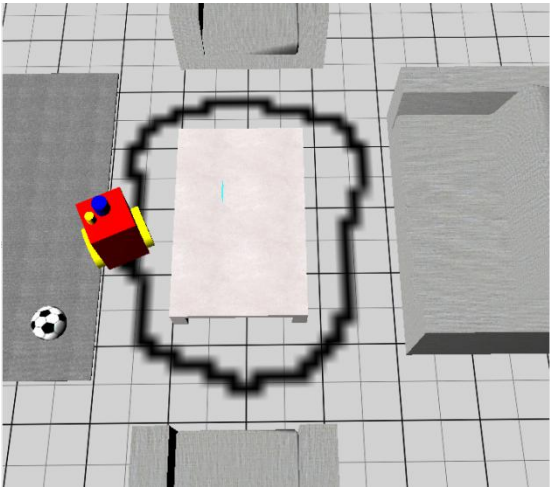


图 3-12 机器人运动轨迹

3.3 超级机器人

Webots 中含有一种特权类型机器人 Supervisor(超级机器人)。主要实现对仿真环境的管理。

由 Webots 中的 Robots(机器人)节点派生而来, 拥有 Robots 节点的一切属性^[52]。Supervisor 是一类拥有特权的机器人, 它可在仿真环境中实现人们在真实环境中的一系列动作, 例如追踪场景中物体的位置、改变物体或机器人的位置、拍照、录像等。不仅如此, Supervisor 可以修改仿真环境中包含的任意节点中的域值, 并且保证了实时性的操作, 即修改后的效果会立即在仿真环境中显现。本文中主要运用其实现对目标物体的坐标位置的获取, 并通过发射器和接收器来进行信息的传输和接收。其相关节点模型如图 3-13^[50]所示:

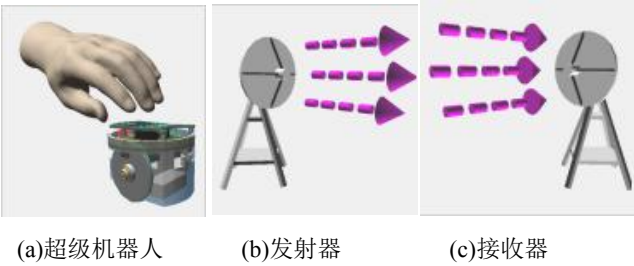
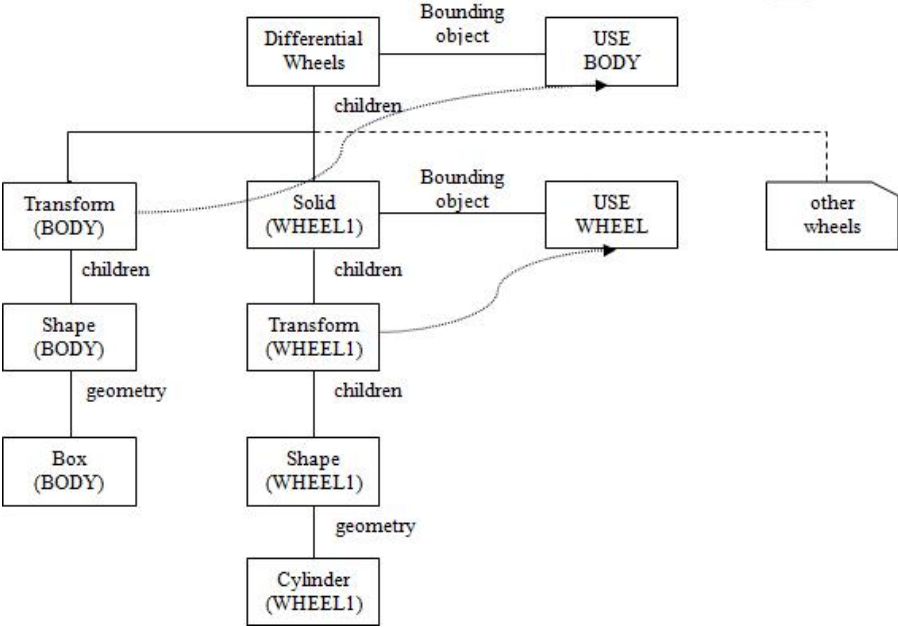


图 3-13 节点模型

机器人模型的整体构造主要是由 solid 节点连接而成的树形结构。在 Webots 中构建机器人模型首先要选择一个根节点作为起始节点, Webots 中提供了 DifferentialWheels 和 Robot 两类节点来搭建轮式机器人模型。一般情况下这颗树的节点应该是一个 Robot 节点, 通过 joint 节点将 solid 连接在一起。而 DifferentialWheels 节点是基于 Robot 节点而建立的。本文选取其作为机器人主体结构的根节点。部分主体树状结构如图 3-14 所示:



1.bounding object-边界对象 2.Solid-刚体 3.Transform-转换 4.geometry-形状 5.Box-矩形 6.Cylinder-圆柱
体 7.children-子节点 8.USE-调用 9.Shape-形状

图 3-14 部分主体树状结构图

3.4 本章小结

本章首先对目前运用较广泛的机器人研发软件和本文所要使用的 Webots 开发平台进行比较分析。之后对 Webots 中的有关场景和机器人建模的相关节点，原理和建模方法进行介绍，完成了对实验场景和机器人模型的构建。

第四章 特征提取与处理

物体识别在家用机器人领域面临的难点主要可以归结为对算法的鲁棒性、准确性和快速性三个方面的要求。因此，选择合适的方法来实现室内目标识别具有重要的意义^[53]。

4.1 基于 SURF 算子提取特征

SURF 算法是 Herbert Bay 等人在 SIFT 的基础上提出的基于盒状滤波器和积分图像理念的一种特征描述的算法，同时兼具了较高的计算效率和良好的鲁棒性。相比 SIFT 算法采用了不同的特征提取方式，降低了特征的维度，从而提升了算法的运行速度，并且仍保持了较高的特征的特异性，因此可以实现较好的识别效果。其不同主要体现在针对空间尺度的问题，SURF 算法提出将原始的高斯滤波器运用模板尺寸大小可随不同分辨率的图像进行调节的盒状滤波器来代替；针对计算效率的问题，SURF 算法提出运用积分图像的方法，提升卷积的速度。以此降低算法的复杂程度，提升计算速率，从而达到实时性的目的；针对特征点确定的问题，SURF 算法提出 Hessian 矩阵的思想。即通过找寻该矩阵的局部最大值来确定特征点。由于 SURF 算法在 SIFT 算法的基础做了如上的改进，因此，SURF 算法可以认为是对 SIFT 算法的进一步完善，是一种性能更为优良的特征提取类算法^[54]。其具体的步骤如下：

1. 求取积分图像。

积分图像是通过对原图像的左上角任意一点到图像原点对角线区域的灰度值求和得到的。即

$$ii(i,j) = \sum_{r \leq i, c \leq j} p(r,c) \quad (4-1)$$

式中， $ii(i,j)$ 表示的是积分图像中点 (i,j) 的值， $p(r,c)$ 表示图像中点 (r,c) 的灰度值。

求积分图像时，只需对原图像的所有像素扫描一遍即可。由于积分图像的计算时间不受矩形区域大小变化的影响，其计算的时间是恒定不变的。因此，SURF 算法的计算时间不受滤波器尺寸变化的影响，也基本保持在恒定的计算时间内完成卷积。

2. 构建 Hessian 矩阵。

SURF 算法对特征点的检测主要通过寻找 Hessian 矩阵的最大局部特征值的方法来实现的，以此来保证较高的计算精度和计算效率。其中 Hessian 矩阵的模是 SURF 检测的一个基础，主要采用 DOH(Determinant of Hessian)进行特征点检测。DOH 是一种常用的检测特征点的方法。Hessian 矩阵如公式 4-2 所示：

$$H(x,\sigma) = \begin{bmatrix} L_{xx}(x,\sigma) & L_{xy}(x,\sigma) \\ L_{xy}(x,\sigma) & L_{yy}(x,\sigma) \end{bmatrix} \quad (4-2)$$

其表示的是在尺度空间 σ 的 Hessian 矩阵。其中 $L_{xx}(x,\sigma)$ 表示的是 $\frac{\partial^2 g(\sigma)}{\partial x^2}$ 在图像点 x 处的卷积。 $L_{xy}(x,\sigma), L_{yy}(x,\sigma)$ 与此类似。Hessian 矩阵的模(判别式)如公式 4-3 所示:

$$det = \frac{\partial^2 f}{\partial x^2} \frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y} \right)^2 \quad (4-3)$$

高斯滤波在对图像尺度空间分析的方面虽然具备良好的性能,但由于其在实际的计算过程中的运算复杂且会产生不同程度的图像畸变。所以 SURF 算法运用盒状滤波器(box filter)来代替高斯滤波器进行特征提取,这种近似的运算可在积分图像中获得较快的运算速度,即可以忽略图像尺寸大小的变化对计算的影响,从而大幅缩短了特征提取的时间。其对比图如图 4-1^[42]所示。此图表示的是尺度空间因子为 $\sigma = 1.2$ 时的情况,其中,前两幅图像分别表示 y 和 xy 方向的二阶高斯滤波,后两幅图像则分别表示其相应方向的运用盒状滤波器对高斯滤波器进行近似的情况。从中可知,两者的性能相近。

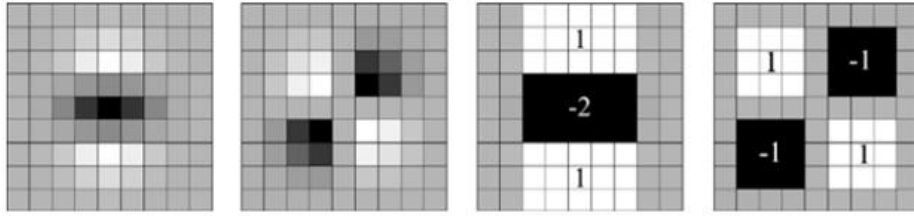


图 4-1 高斯滤波和盒状滤波

因此,可以运用 D_{xx}, D_{xy}, D_{yy} 对应高斯滤波中的 $L_{xx}(x,\sigma), L_{xy}(x,\sigma), L_{yy}(x,\sigma)$ 。得到公式 4-4:

$$\frac{|L_{xy}(1.2)|_F |D_{xy}(9)|_F}{|L_{xx}(1.2)|_F |D_{xy}(9)|_F} = 0.912 \approx 0.9 \quad (4-4)$$

其中 $|x|_F$ 为 Frobenius 范数,由此可以得到近似公式如公式 4-5 所示:

$$det(H_{approx}) = D_{xx}D_{yy} - (0.9D_{xy})^2 \quad (4-5)$$

实现替换之后对滤波后的结果依据不同大小的掩膜进行归一化的处理,确保其最终获得的 Frobenius 范数在任何尺寸的滤波器下均是一致的。。

3. 构建尺度空间。

通过增大滤波器模板的尺寸来求取 Hessian 的响应图像。一般而言,获取不同尺度的斑点需要通过图像的尺度空间金子塔来实现。由于 SURF 算法运用到盒状滤波和积分图像的理念,所以 SURF 算法只需改变模板的大小来实现,即采用积分图像与不同尺度 σ 的盒状滤波模板相结合的方法来求取 Hessian 矩阵行列式的响应图像。且不会增加运算的工作量。例如,首先利用 9×9 的模板对图像进行滤波处理并将其结果设为初始的尺度空间层(此时尺度值 s 为 1.2, 近似 $\sigma = 1.2$ 的高斯微分),之后逐渐扩大模板的尺寸,并进行相应的滤波处理。与 SIFT 算法相似,尺度空间需进行分组。每组分别代表了不同尺度下的模板对同一图像进行滤波处理后获得的响应图像

^[55]。且每个组均包含若干固定的层。每层之间的最小尺度变化量是盒状滤波模板尺寸的 1/3。即下一层的响应长度应至少要比上一层的响应长度多两个像素。因此模板的尺寸变化序列依次为：9×9，15×15，21×21，27×27 等。其过程如图 4-2^[42]所示，为保证中心像素始终唯一，黑白区域的像素的增加以偶数作为间隔。当图像的尺度变化较大时，其窗口变化的幅度和感兴趣点的采样间隔也要做出相应的变化。

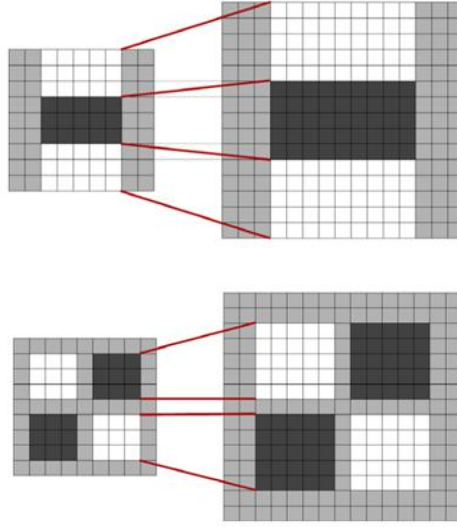


图 4-2 模板尺寸变换示意图

4. 精确定位特征点。

对提取的特征点及其周围相同尺度范围内的点构造一个 3×3×3 的立体邻域，对每个尺度下的每个像素点都与其三维邻域的 26 个邻点进行比较，26 个邻点包括周围的 8 个和上下层尺度上的各 9 个点。如图 4-3^[42]所示。在这个邻域内采用 3D 非最大值抑制法，寻到各尺度下的 Hessian 矩阵的局部最大值，其相关公式如式 4-6 所示：

$$H(x) = H + \frac{\partial H^T}{\partial x} + \frac{1}{2} x^T \frac{\partial^2 H}{\partial x^2} x \quad (4-6)$$

由公式可推出特征点位置为：

$$\hat{x} = \frac{\partial^2 H^{-1}}{\partial x^2} \frac{\partial H}{\partial x} \quad (4-7)$$

以此确定感兴趣点的最终位置。

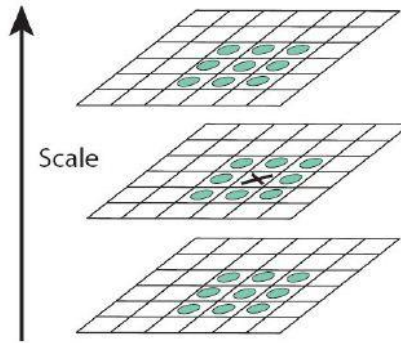


图 4-3 特征点的三维邻域示意图

5. 主方向确定

通过统计特征点邻域内的 Harr 小波特征，得到每一个特征点的主方向来确定。为保持算法的旋转不变性，SURF 算法会在感兴趣点周围的圆形区域内寻找一个主方向，主要通过对感兴趣点周围以 $6s$ 为半径的圆形区域内的点分别计算其 x 和 y 方向的 Haar 小波响应来实现。 s 为特征点所在的尺度值。首先，设置 Haar 小波的窗口尺寸为 $4s$ ，感兴趣点为圆心，以 60° 扇形区域进行旋转，遍历整个区域。同时计算各个区域中的水平方向和竖直方向的 Haar 小波的响应和，最大的响应和的方向即为该感兴趣点的主方向。图 4-4^[42]描述了这个过程。

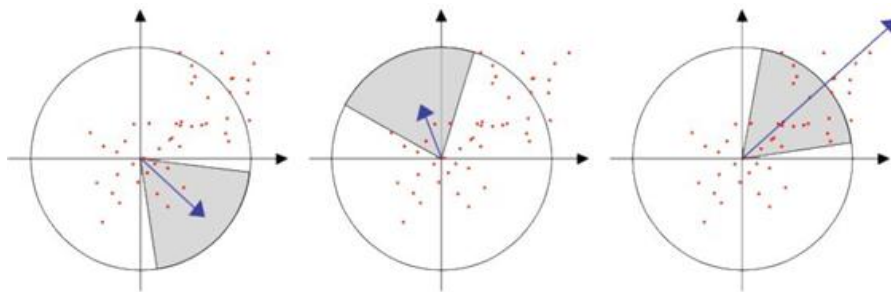


图 4-4 主方向确定过程示意图

6. 特征点描述子生成

感兴趣点的主方向确定之后，沿其划定一个长为 $20s$ 的方形区域，并将其分为 4×4 的子区域，对此以 5×5 的大小进行统计，之后分别计算各区域的水平方向和垂直方向的 Haar 小波响应。如图 4-5^[42]所示。为防止几何畸变和位置误差的影响，SURF 算法将小波响应的结果和高斯滤波进行相乘运算，得到一个四维向量： $\sum dx$ 、 $\sum |dx|$ 、 $\sum dy$ 、 $\sum |dy|$ ，分别代表水平方向，水平方向绝对值，垂直方向，垂直方向绝对值之和，最终在这个 4×4 区域中可以得到一个 64 维的值：

$$V_{64} = [\sum dx, \sum |dx|, \sum dy, \sum |dy|] \quad (4-8)$$

为降低光照和对比度的影响，对特征矢量进行归一化的处理得到一个特征点对应的描述符。

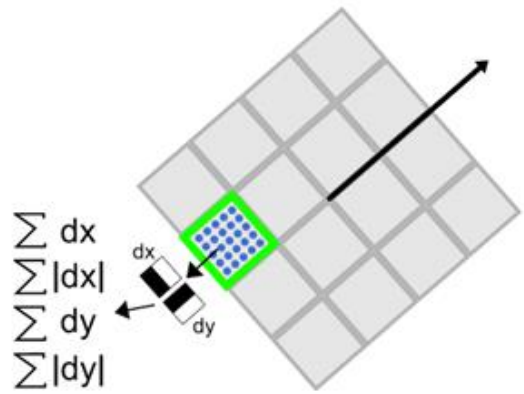


图 4-5 特征描述子生成表示

至此便实现了对一个特征点的特征描述，由上述步骤可知，SURF 不仅具有局部特征的优点，而且计算速度快，鲁棒性好。

本文运用 SURF 算子提取的特征示意图如图 4-6 所示：

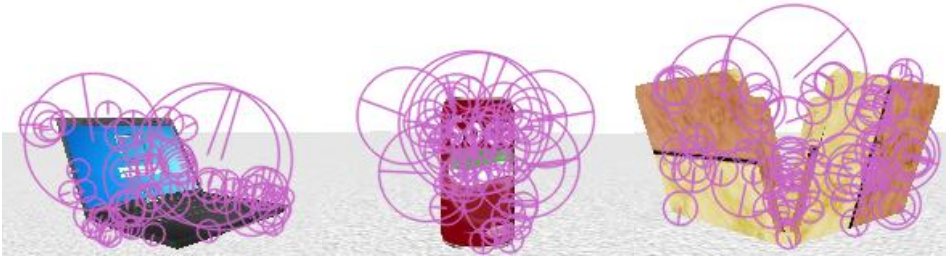


图 4-6 SURF 特征提取图

图中的圆圈代表的是特征点的位置和主方向。且圆圈的大小与特征的尺度成正比。

4.2 特征处理

降维在机器学习领域中主要是指通过某种映射方法改变原始数据的空间，使其实现从高维空间分布向低维空间的转变^[56]。

在机器学习中，随着数据维度的增多，数据在每个特征维度上的分布就会变得越稀疏，数据中包含的冗余也会越来越多。这对机器学习算法基本都是灾难性的。而且随着维数的增加，样本空间中的噪声会随之增加，这使得提取出的样本特征不能很精确的表示样本的特异性。从而在实际应用中会造成识别的误差。特别是样本数越少时，越应该采用少量的特征进行训练。数据降维直观地好处在于减少了数据存储所需的空间和计算所需的时间，提升了计算效率，同时去除噪声，提高了模型性能。而其深层次的意义在于实现了有效信息的提取，综合以及无用信息的摒弃，从而达到高效简化数据集的目的。虽然 SURF 算法相较 SIFT 算法而言实现了数据的降维，但获取的特征依然有 64 维而且是无序的。因此本文提出运用主成分分析(principal

component analysis,PCA)法对 SURF 算法提取出的特征进行降维处理。提取出其中的主成分。PCA 算法是机器视觉中常用的降维算法,也是目前应用最为广泛的方法^[57],该算法的主要思想是运用线性投影,将高维数据映射到低维空间。并寻找所投影维度上的最大方差,以此保证在使用较少数据维度的同时,保留尽可能多的原始数据信息。

PCA 算法通常对全体数据操作。经过变换后的数据仅最大程度地保留了原始空间的信息含量,但并未对提取的信息进行分类整理,即获取的特征属于无类别信息。因此该信息也并非全部有利于后续识别。同时,不同图像提取的 SURF 特征点的数目不同且数量较大,因此无法有效运用支持向量机(Support Vector Machine, SVM)对其进行训练。所以本文运用视觉词袋模型(Bag-of-Visual-Word)对降维后的数据进行进一步的聚类处理,以弥补上述不足。

4.2.1 基于 PCA 算法的特征降维

PCA 算法是由 K-L 变换发展而来的^[58],是一种在尽可能减少信息损失的情况下找到某种方式降低数据的维度的方法。是原始数据的重新组合。该算法的主要思想是将原来的一个坐标系映射到另外的一个空间坐标系,使降低维度后的特征形成某种非线性的关系,但是新的特征值(降低后的维度)能够最大程度地反映原始数据的特征,换言之,变换后的数据可以达到不影响原始的分类正确率,并且提高了运行速度。其具体的方法是通过线性变换,运用低维向量和特征向量矩阵对原始的高维向量进行重构,得到一组各维度线性无关的表示以此实现最大程度的反映出数据原有的特征信息。

新坐标系的选取主要依据相关的数据进行确定。其中,第一个坐标轴代表的是原始数据中方差最大的那个方向。而这对于数据来讲就是其自身最主要的方向。第二个坐标轴选取的是与第一个坐标轴正交的方向。对此过程进行重复,重复的次数为原始数据的特征数。而这些方向所表示的数据特征即为主成分。该过程的示意图如图 4-7^[59]所示。

图中的 B 线表示的是该数据中方差最大的方向,即其最主要的方向,为第一主成分。而 C 则是相对于 B 的正交坐标,即为第二主成分。

PCA 对于主成分的求取主要是通过求取其相应的协方差矩阵及该矩阵相应的特征值和特征向量来实现的。从而实现将样本从高维到低维的映射。

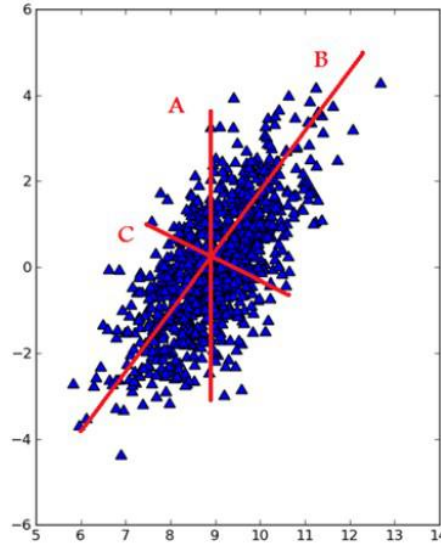


图 4-7 PCA 新坐标轴

本文中 PCA 算法与 SURF 算法相结合的方法是在保证 SURF 算法尺度不变性，高鲁棒性的同时提升了其计算效率。运用其对 SURF 算法提取的特征进行降维的主要步骤如下：

(1) 将 surf 算法提取的 n 个 64 维特征点存入矩阵。 x_i 表示第 i 个特征点的 64 维描述符，样本矩阵为 $[x_1, x_2, x_3 \dots x_n]^T$ 。

(2) 计算 n 个特征向量的平均值：

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i \quad (4-9)$$

(3) 对 n 个 x_i 进行零均值化。即计算特征向量与平均值的差，得到的差值向量：

$$D_i = x_i - \bar{X}, i = 1, 2, \dots, n. \quad (4-10)$$

(4) 求协方差矩阵：

$$C = \frac{1}{n} \sum_{i=1}^n D_i D_i^T \quad (4-11)$$

(5) 求协方差的 64 个特征值 λ_i 及其对应的特征向量 e_i 。

(6) 将所有的特征值和相应的特征向量依照从大到小的顺序进行排列。

(7) 取排列在前 m 个的特征向量作为最终的主成分。

(8) 构造 $64 * m$ 的矩阵 Y ， Y 即是降维后的数据。 y_i 表示第 i 个特征点的 m 维描述符，矩阵 $Y = [y_1, y_2, \dots, y_n]^T$ 。

其中 m 值的选取对后续匹配，训练的准确性至关重要。 M 值越大，其使用的特征维度越高，其相应的降维误差就越小，即其越完整的保留了原始数据的特征，更加接近真实的样本数据。反之，则会丢失一部分原始数据的信息。关于 m 值的确定，本文主要依据每个主成分的累积贡献率来进行确定。其计算公式如下：

$$\frac{\sum_{i=1}^m \lambda_i}{\sum_{i=1}^n \lambda_i} \geq 0.995 \quad (4-12)$$

其中 λ_i 为步骤(6)所求的特征值， $\sum_{i=1}^n \lambda_i$ 为所有特征值的和， $\sum_{i=1}^m \lambda_i$ 为前 m 个特征值的和，0.995 为贡献率的临界值，即当前 m 个特征值的贡献率大于等于 0.995 时，降维的个数为当前 m 值。

4.2.2 基于 Bag-of-Visual-Word 的特征聚类

视觉词袋模型(Bag-of-Visual-Word)近年来被广泛运用在计算机视觉中^[60]，由词袋模型(Bag-of-Word)扩展而来。BoW 模型是一种基于统计的方法。最初用于文本分类当中。将文本中的关键字作为词表，对所要分类的文本与生成的词表进行对照来实现分类。在机器视觉领域，将图像特征作为单词进行处理的方法，可以实现将词袋模型(BoW 模型)融入图像分类的应用当中。虽然该方法未考虑物体的一系列空间信息，但由于其构造简单的同时兼具良好的鲁棒性和较快的处理速度，因此可以满足机器人对实时性的要求。

Bag-of-Visual-Word 在对特征进行处理的过程中，其中心思想是运用 K-means 算法对图像特征进行聚类处理。K-means 是一种基于距离的聚类方法，其主要是依据欧式距离作为相似度测量的准则。算法采用误差平方和准则函数作为聚类准则。聚类时首先选取 K 个随机的特征向量作为初始的聚类簇中心，之后将其与其他的特征向量进行欧式距离求解，根据每个特征与各个簇中心距离的大小进行重新归类，以使其相似度最高。对上述步骤进行迭代直至新的簇中心与原中心相等或小于指定的阈值。K-means 的聚类示意图如图 4-8^[60]所示：

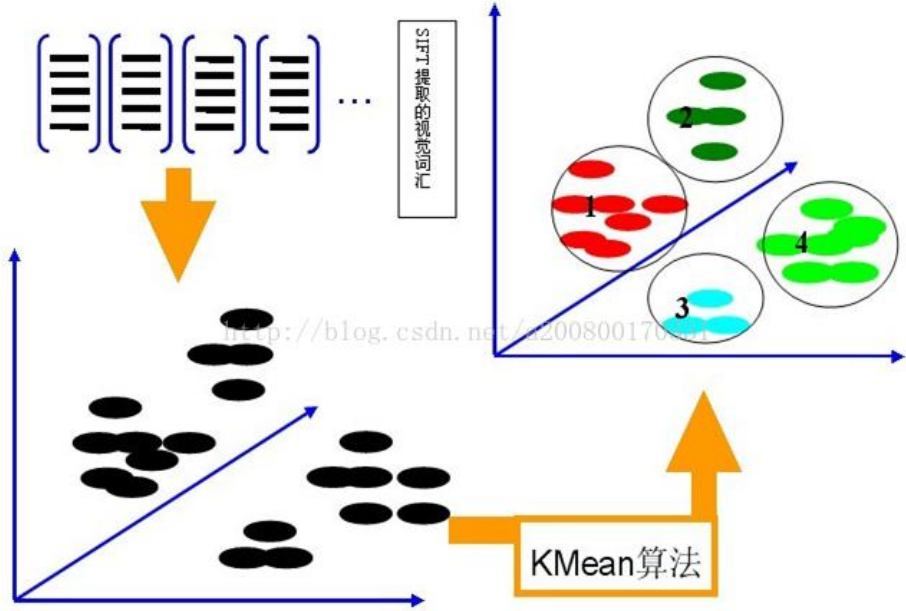


图 4-8 K-means 聚类示意图

聚类完成后，Bag-of-Visual-Word 会为每个簇分配一个标签，所有的标签构成一个词典。词典生成后，Bag-of-Visual-Word 对提取的每一个样本特征统计其在词典中的频数分布，得到的直方图向量即为该样本的视觉词袋模型。模型示意图如图 4-9^[61]所示：



图 4-9 视觉词袋模型示意图

由上述介绍可知，基于 Bag-of-Visual-Word 的特征聚类主要有以下三个步骤：

- (1) 将每幅图像的局部特征作为视觉词汇向量
- (2) 运用 K-Means 算法对所有的特征向量进行聚类，将相似的视觉词汇进行归并，构造一个具有 K 个词汇的词典。K 为聚类的个数。
- (3) 对词典中的每个单词在某一样本中出现的次数进行统计，之后将该类样本运用 k 维数值直方图进行表示。

运用 Bag-of-Visual-Word 对特征点进行聚类的示意图如图 4-10 所示：

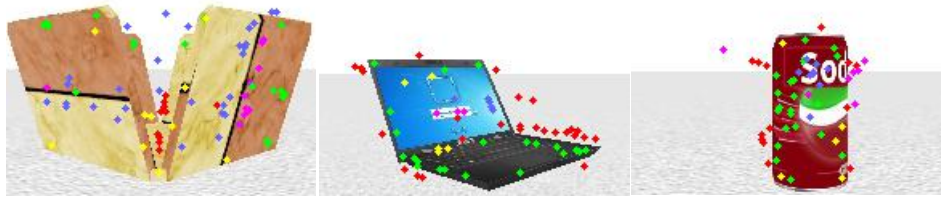


图 4-10 特征点聚类示意图

图中，相同颜色的点属于同一簇，不同颜色深浅的点代表不同的簇。

K 值的大小可以由几十到上千，若词典数目太小，将不利于对图像特征进行区分；若词典的数目太大，则会造成相同特征的误分离。以上情况均不利于物体的识别。所以，k 值的选择对于识别的精度至关重要。

聚类 K 值的选取问题，目前存在多种求解的方法。但单一的计算结果不一定准确，而且最终的 k 值要根据应用场景和实际情况来选取，而不能完全的依据评估参数，因此本文采用均方根法和 Elbow 法相结合的方法来初步确定 k 值，缩小取值范围，之后根据识别精度来确定最终的聚类数目。

首先运用均方根法来计算 k 值的大致取值，之后运用 Elbow 法来计算所得 k 值及其邻域值相应的畸变函数 J，并绘制折线图来观测斜率的变化。均方根法借助公式 4-13 进行计算：

$$K = \sqrt{m/2} \quad (4-13)$$

其中 m 代表样本数。畸变函数如公式 4-16 所示：

$$J(C^{(1)}, \dots, C^{(m)}, \mu_1, \dots, \mu_k) = \frac{1}{m} \sum_1^m (\|x^{(i)} - \mu_{c^{(i)}}\|) \quad (4-14)$$

其中 $x^{(i)}$ 表示第 i 个样本， $\mu_{c^{(i)}}$ 表示 $x^{(i)}$ 所属的聚类中心。 $c^{(i)}$ 表示最接近 $x^{(i)}$ 的聚类中心下标。

4.3 基于 SVM 的特征训练

支持向量机(Support Vector Machine, SVM)是常见的一种判别方法。有效的解决了传统分类器需要大量训练样本进行训练的缺陷。SVM 在机器学习领域中属于有监督的学习模型。多用于模式识别，回归分析以及分类等领域^[62]。适用于包括文字、语音、图像等领域的识别问题在内的多个领域。

由 Chapelle^[63]可得，图像的高维特征受到传统的分类方法的制约很难获得理想的结果，但 SVM 可以很好的解决上述的问题。

SVM 是 Vapnik 等人^[64]根据统计学习理论提出的一种新的线性分类器设计准则。其目的是寻找一个超平面来实现对样本的分类。它是一种基于结构风险最小化准则的学习方法。即在以训练误差为约束的情况下实现置信范围值最小化的目标。

SVM 可以解决线性可分的问题，也支持线性不可分的情况，其主要通过运用非线性映射算法实现低维输入空间线性不可分的样本向高维特征空间线性可分的转变，从而转化为解决线性可分的情况。

在分类问题中，SVM 可以看做是根据给定训练集寻找决策函数

$$f(x) = \text{sgn}(g(x)) \quad (4-15)$$

之后对测试数据对应的 y 值进行推断的过程。

本文中 SVM 的分类器构造和识别的流程如图 4-11 所示：

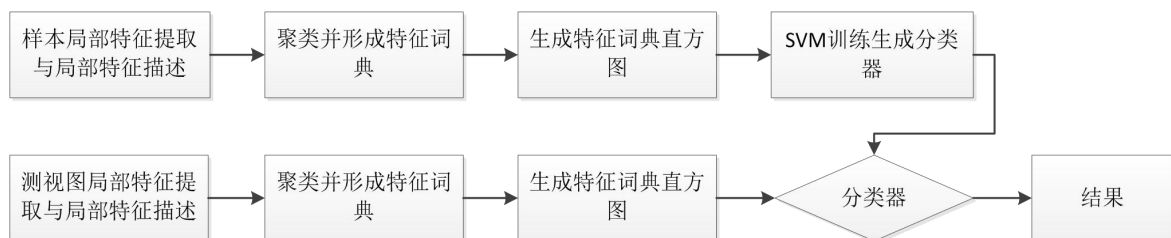


图 4-11 SVM 分类器构造和识别流程

4.4 本章小结

针对 SURF 算法提取的特征具有高维和特征点数不定的特点，运用 PCA 算法和 Bag-of-Visual-Word 算法进行降维和聚类处理，并具体介绍了算法的基本原理和特征处理方法，方便 SVM 的训练。

第五章 实验设计与结果分析

本文中的实验是根据机器人物体识别中存在的问题进行设计的，并在 Webots 平台上进行实验，验证其可行性和实用性。

5.1 搭建仿真环境

本文主要运用 Webots 和 OpenCV 来实现。运用 C++ 进行编程控制，同时将 OpenCV 库移植到 Webots 中，在 Webots 平台上实现机器人运动控制、样本采集、测试图片采集和特定物体识别。

为在 Webots 平台上完整实现上述功能需要对 OpenCV 进行移植和配置。具体过程如下所示：首先获取 OpenCV 安装包，并将 Webots 中 mingw 文件下的 bin 目录插入 Windows 环境变量 PATH 中，之后用 mingw32-make 对 OpenCV 的安装包进行配置和编译^[65]。

参数配置完成之后需要进一步对环境进行配置，具体的配置过程如下所示：

(1) 在相应的工程中的 Makefile 文件中添加所需的各个头文件相对应的目录，格式依据 INCLUDE = -I "C:\Users\YourName\XYZLib\include" 进行编写，其中 XYZLib 是所添加外部库的名字；

(2) 之后在 Makefile 文件中添加 Lib 库，格式依据 LIBRARIES = -L"C:\Users\YourName\XYZLib\lib" -lXYZLib 进行编写；

(3) 若工程中需要使用多项源文件，需要在其中添加相应的关键字。如果运用 C 语言则在 Makefile 文件中添加 "C_SOURCES" 关键字，格式依据 C_SOURCES = my_first.c my_second_file.c my_third_file.c 进行编写；如果运用 C++，则添加 "CXX_SOURCES" 关键字，格式依据 CXX_SOURCES = my_first.cpp my_second_file.cpp my_third_file.cc 进行编写；

(4) 若在 C++ 程序编写的过程中涉及到 C 语言相关的 API 函数时，则需要在 Makefile 文件添加 USE_C_API 关键字，格式为：USE_C_API= true。

完成上述配置后，只需在 Webots 平台上的机器人的控制器中添加相关的头文件，即可实现对 OpenCV 中的各项 API 函数的调用。本文编译完成的示意图如图 5-1 所示。

```

CXX_SOURCES =trainandtest.cpp
USE_C_API = true

INCLUDE = -I"D:\opencv2.4.9\webots_opencv\install\include"
INCLUDE += -I"D:\opencv2.4.9\webots_opencv\install\include\opencv"
INCLUDE += -I"D:\opencv2.4.9\webots_opencv\install\include\opencv2"

LIBRARIES = -L"D:\opencv2.4.9\webots_opencv\install\x64\mingw\bin"
-llibopencv_core249
LIBRARIES += -L"D:\opencv2.4.9\webots_opencv\install\x64\mingw\bin"
-llibopencv_features2d249
LIBRARIES += -L"D:\opencv2.4.9\webots_opencv\install\x64\mingw\bin"
-llibopencv_highgui249
LIBRARIES += -L"D:\opencv2.4.9\webots_opencv\install\x64\mingw\bin"
-llibopencv_nonfree249
LIBRARIES += -L"D:\opencv2.4.9\webots_opencv\install\x64\mingw\bin"
-llibopencv_legacy249
LIBRARIES += -L"D:\opencv2.4.9\webots_opencv\install\x64\mingw\bin"
-llibopencv_calib3d249
LIBRARIES += -L"D:\opencv2.4.9\webots_opencv\install\x64\mingw\bin"
-llibopencv_imgproc249
LIBRARIES += -L"D:\opencv2.4.9\webots_opencv\install\x64\mingw\bin"
-llibopencv_ml249

```

图 5-1 环境配置示意图

5.2 实验准备

5.2.1 样本图采集

样本图像运用机器人进行采集。尽可能降低背景的干扰，同时要多角度，多姿态，不同光线和阴影进行采集。本文主要实现三类物体的识别，分别是饮料，果盘和电脑。同时会将环境背景单独作为一类以识别无目标物体和存在复杂背景干扰的情况。每个类别包含 90 张样本图，每张图片的分辨率为 200*170。采集的样本图如图 5-2 所示：

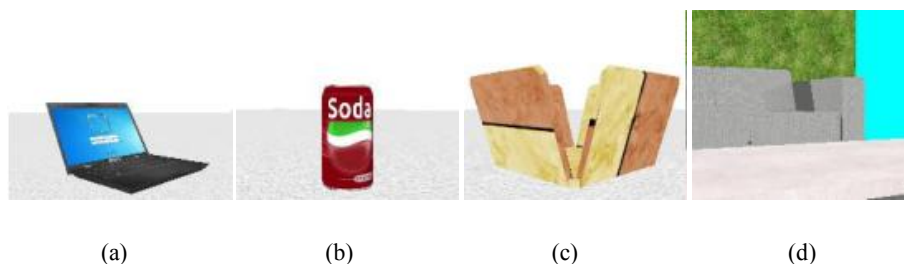


图 5-2 样本图

5.2.2 机器人运动控制

机器人的运动流程如图 5-3 所示：

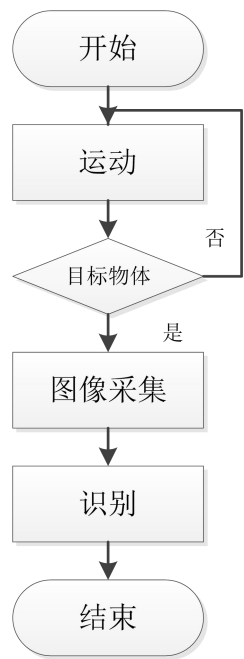


图 5-3 运动流程图

测试图像由机器人在搭建的家庭环境中进行采集。开启仿真后，机器人首先以初始速度进行匀速运动。之后调用 Webots 中的 `wb_robot_get_device()`，`wb_distance_sensor_enable()` 和 `wb_camera_enable()` 函数来调用和启动机器人的距离传感器和摄像机节点，之后运用 `wb_distance_sensor_get_value()` 和 `wb_camera_get_image()` 函数来返回距离传感器的值以及摄相机获取的图片，当两个距离传感器同时达到设定的阈值时，对 Supervisor 和 GPS 获取的障碍物和机器人的坐标值进行距离计算，判断是否靠近目标物体，如果是则运用 `wb_camera_save_image()` 进行图像采集，之后保持距离不变，围绕目标物体每 90° 拍摄一次照片。若只有一个值达到阈值则让机器人转向障碍物的一方后再进行上述计算来进行判断，从而提升机器人的工作效率。

机器人的运动轨迹如图 5-4 所示：

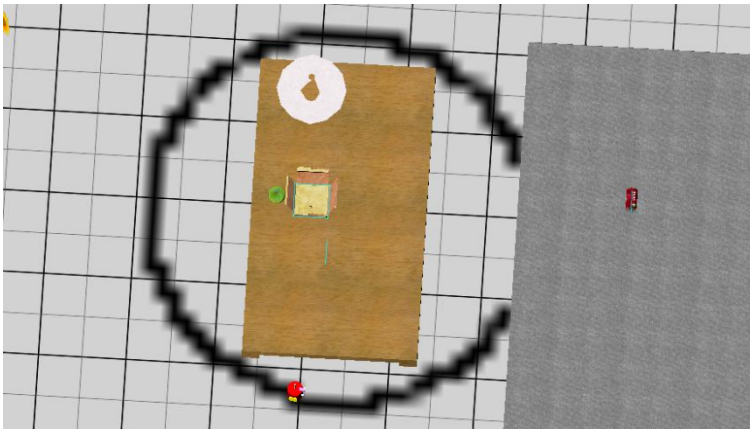


图 5-4 机器人运动轨迹示意图

根据之前对距离传感器的介绍，Webots 中提供的 `lookupTable` 表可以实现对距离传感器相关参数的设定，从而实现距离传感器阈值的选取。本文中对 `lookupTable` 节点参数的设置结果如表

5-1 所示：

表 5-1 lookupTable 参数设置表

$$lookupTable \begin{bmatrix} 0 & 1000 & 0 \\ 0.1 & 1000 & 0.1 \\ 0.15 & 700 & 0.1 \\ 0.2 & 400 & 0.1 \\ 0.3 & 50 & 0.1 \\ 0.37 & 30 & 0 \end{bmatrix}$$

其中第一列表示的是输入距离，第二列表示的是期望响应值，第三列表示的是期望的噪声标准偏差。返回值中的噪声主要依据高斯随机数分布进行计算，其大小表示为响应值的百分比（信号质量通常由标准偏差的两倍来表示）。但需要注意，查找表中的输入值始终为正数，并按递增顺序排序。图 5-5 为仿真过程中机器人与障碍物的距离与传感器返回值的示意图：

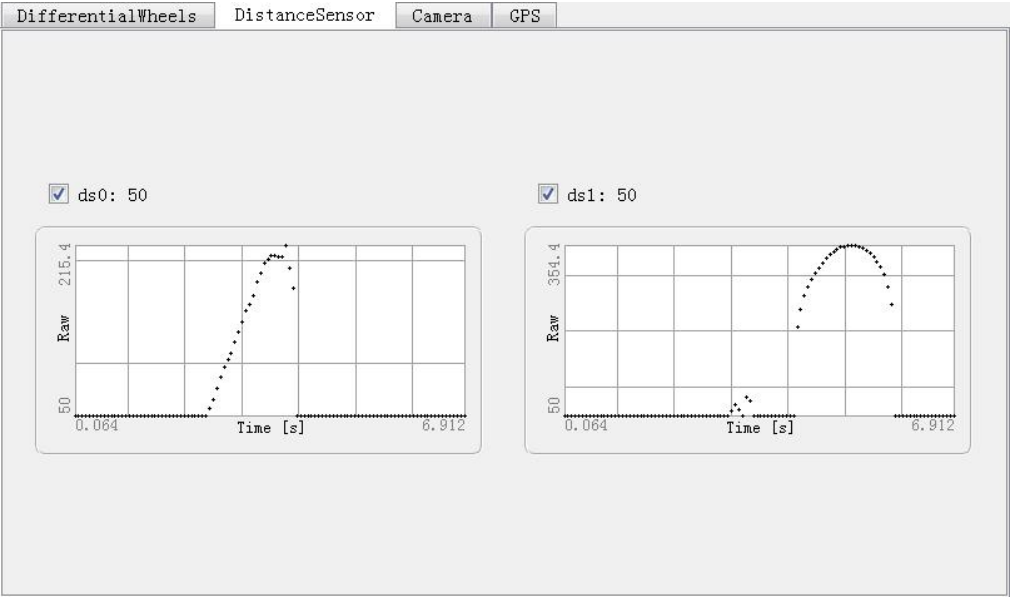


图 5-5 距离传感器的感应值

当与障碍物距离越近时，返回的值越大。

5.2.3 PCA 降维处理

首先分别对四类样本提取特征向量，对其进行降维处理获得相应的特征值并按从大到小的顺序进行排列，之后计算累积贡献率，四类样本各自的累积贡献率如图 5-6 所示。

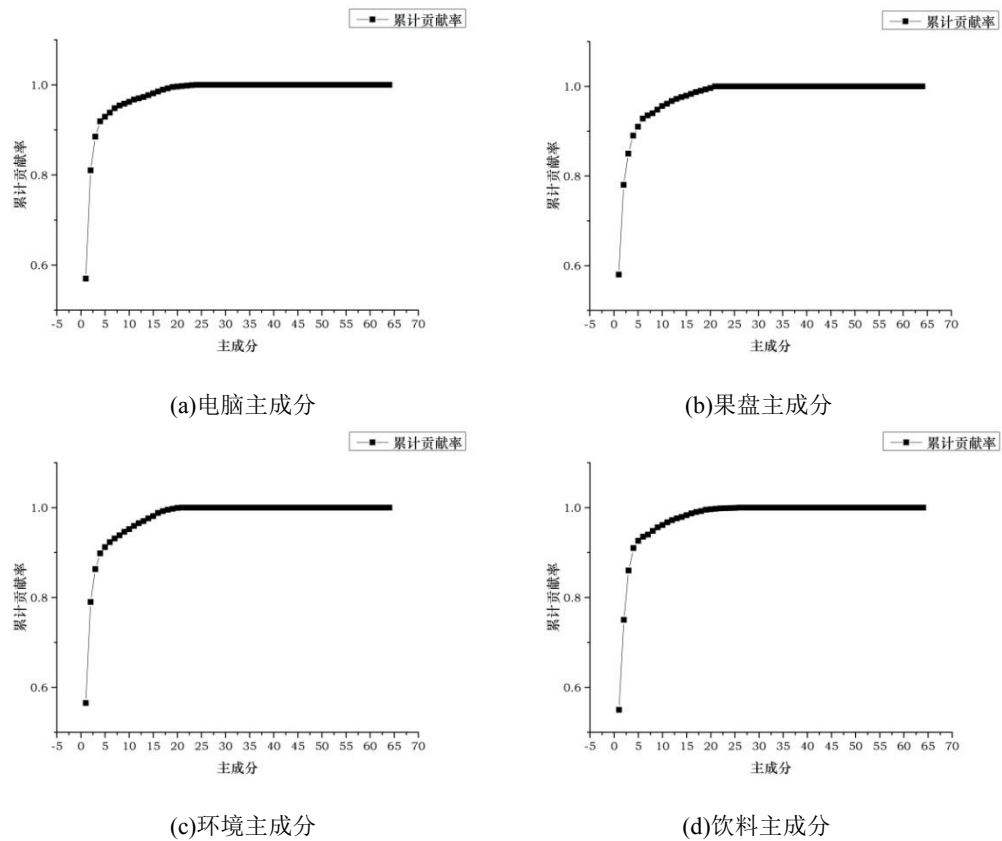


图 5-6 主成分图

由图 5-6 可得，计算机，果盘，环境，饮料的各主成分的累计贡献率分别在 23,20,20 和 25 维时达到了 0.995。所以，选择维度为 25，即 $m=25$ 作为最终的降维数。

5.2.4 聚类处理

对聚类值的确定同降维值的确定类似。对每一类样本的降维结果首先计算其各自的均方根，以此对 K 值进行初步确定。之后对 K 的邻域值运用 Elbow 方法计算对应的畸变函数 J 。计算机，果盘，环境，饮料均方根值分别为 52,58,70 和 60。其各自的 Elbow 变化曲线如图 5-7 所示。

由图 5-7 可知，当计算机的 k 值为 55 时，斜率变化最大；在 $k = 60$ 时，果盘的斜率变化最大；在 $k = 65$ 时，环境的斜率变化最大。饮料没有明显的斜率剧烈变化的点。这是合理的，并不是在任何情况下都会出现明显的拐点。在这种情况下，我们需要结合计算结果和实验结果来确定这个值。本文中，最终选取 $K=80$ 作为最终的聚类值，此时有较好的实验结果。

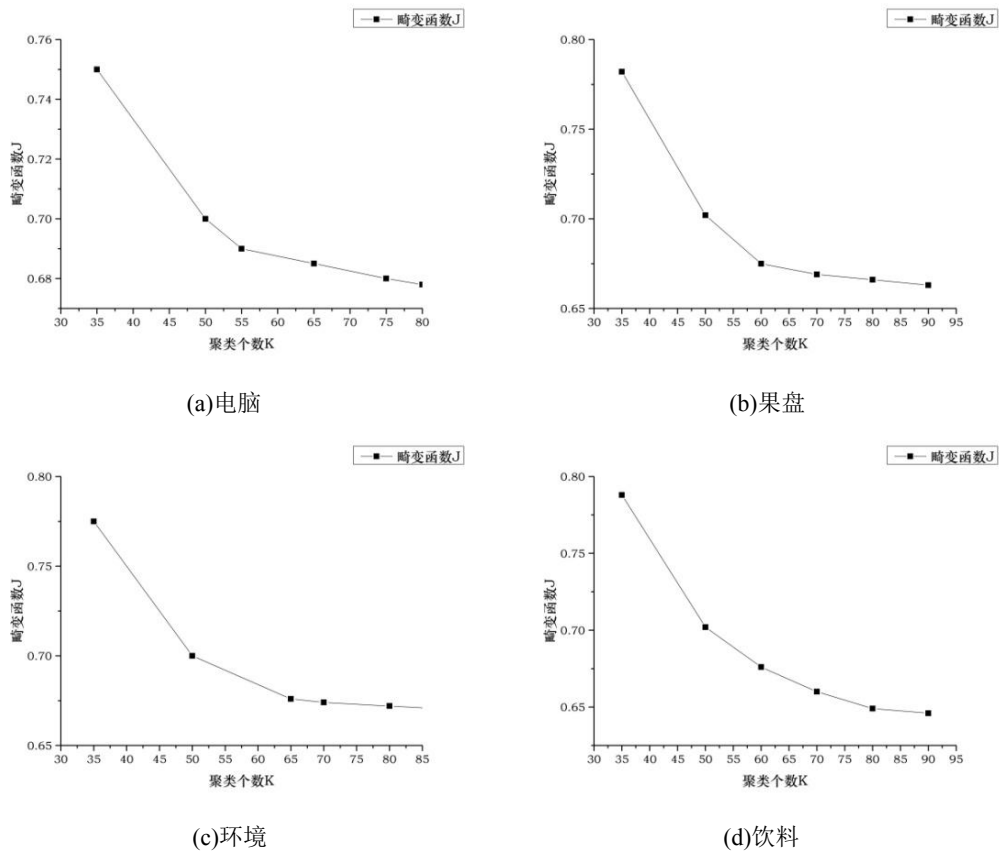


图 5-7 K 值与畸变函数 J 的关系图

5.2.5 图像金字塔分割

为解决复杂环境背景对目标物体识别的影响，进一步提升精确度，本文在图像测试环节运用图像金子塔对测试图像进行分割，然后对每一块图像都进行特征提取，词典描述。

图像金字塔中被广泛使用的类型有高斯金字塔和拉普拉斯金字塔。本文采用高斯图像金字塔进行分割。在分割的过程中涉及两个阈值。一个是建立连接的错误阈值，一个是分割簇的错误阈值。在本文中分别设置为 155 和 52。实验表明，此时的效果较好，可以分割出完整的目标物体。

5.3 实验结果测试与分析

1. 分类器性能统计

分类器的各方面性能如表 5-2 所示：

表 5-2 分类器性能表

样本数	SURF 特征提取时间	分类器构造时间	识别时间(单张)
60	9.964s	8.718s	0.413s
100	11.624s	14.256s	0.416s
200	13.130s	29.763s	0.420s
360	16.772s	40.291s	0.421s

2.环境背景作为样本对识别精度的影响

对环境背景是否作为样本进行训练对识别精度的影响进行统计。计算机选用 CORE i5 处理器，64 位操作系统。选择仿真环境中采集的 100 张图进行测试。测试图的示意图如图 5-8 所示，结果如图 5-9 所示：



图 5-8 测试图

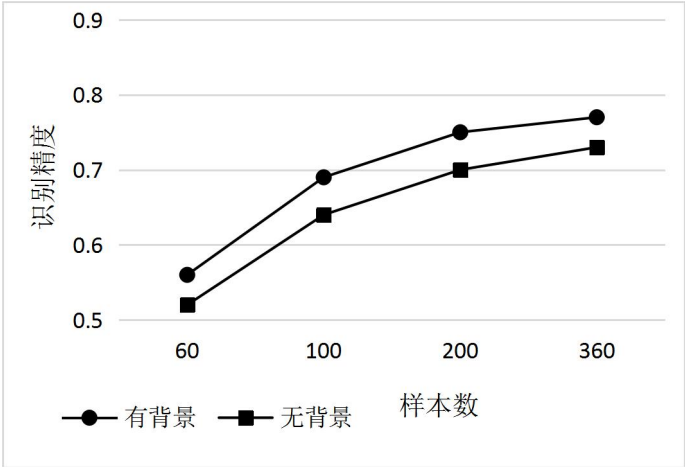


图 5-9 是否将环境背景作为样本对识别精度的影响

由图 5-9 可得，对环境样本进行训练可以在一定程度上提升识别的准确率。而且，随着样本数量的增加，识别的准确率将会提高。因此本文使用 360 张样本图进行训练。

3.图像金字塔分割对识别精度的影响。

对图像金字塔分割对识别精度的影响进行统计，结果如图 5-10 所示：

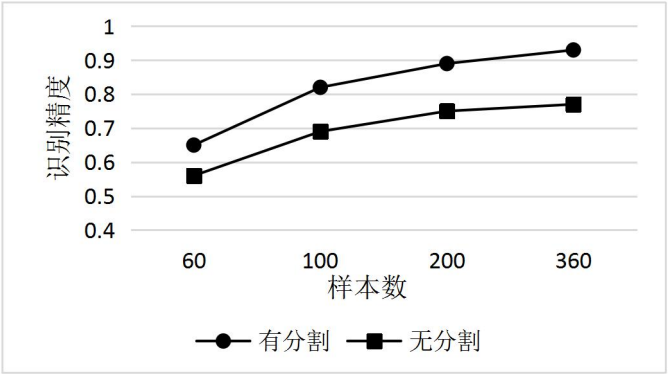


图 5-10 图像金字塔分割对识别精度的影响

由图 5-10 的对比结果可知，对测试图像进行图像金字塔分割后进行分块识别的处理可以在一定程度上提升识别的准确性。

4.不同干扰因素的影响

对不同干扰因素下本文的设计方法相应的识别精度进行统计。干扰因素包括无背景，不同光照，有背景，不同姿势和遮挡。统计结果如图 5-11 所示：

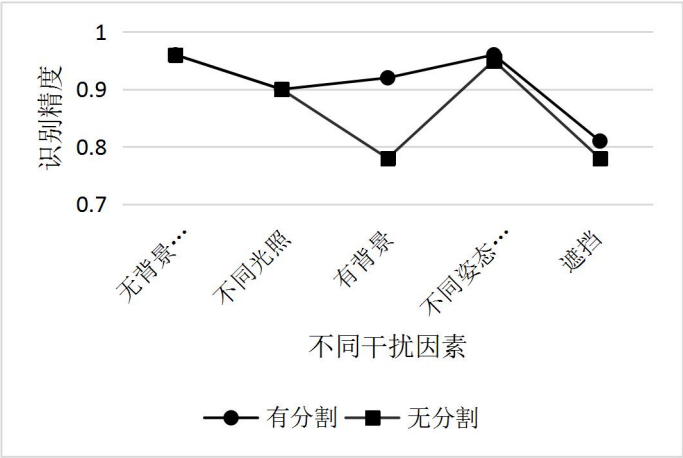


图 5-11 不同干扰因素对识别精度的影响

由图 5-11 可得，本文的算法对以上干扰因素的影响均保持了很高的识别准确率，尤其对复杂的环境有较好的识别结果。

5.多视角与单视角的识别精度对比

多视角相比单视角在很大程度上增加了识别信息的采集^[66]。它大大提高了识别的准确性^[67]。图为对同一目标物单视角与多视角图像信息的对比。



图 5-12 单视角与多视角对比图

由图 5-12 可得，多视角的图可以得到更多目标物体的信息，在目标物体存在遮挡时可以获得较完整的目标物体。因此可以在一定程度上克服遮挡对识别的影响。

因此，本文采取多视角的思路在测试图的采集阶段从目标物体的四个角度采集图像以获得更多的图像信息。其与单视角的对比结果如图 5-13 所示。

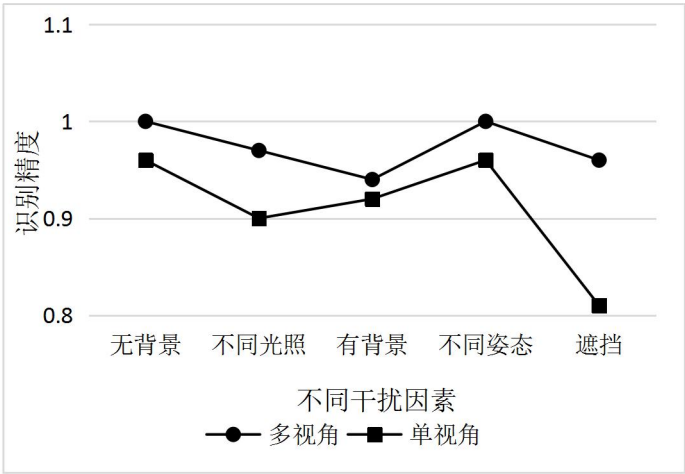


图 5-13 不同视角对识别精度的影响

由图 5-13 可得，本文的设计方案对上述干扰因素的影响均保持较好的识别效果，特别是对目标物体存在遮挡的情况，明显高于单视角图像的识别精度。

6.降维对识别速度的影响

对降维对识别速度的影响进行统计。结果如图 5-14 和图 5-15 所示。

由图 5-14 和图 5-15 可知，虽然 PCA 降维会增加总的识别时间，但可以有效的提升聚类速度。且其识别的总时间仍然满足实时性的要求。

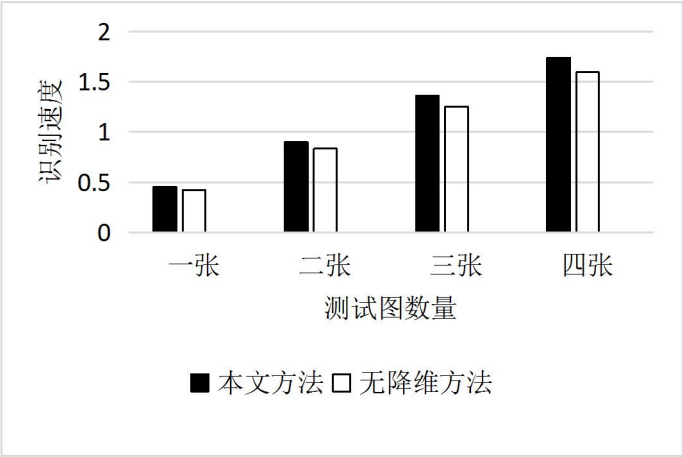


图 5-14 降维对识别总时间的影响

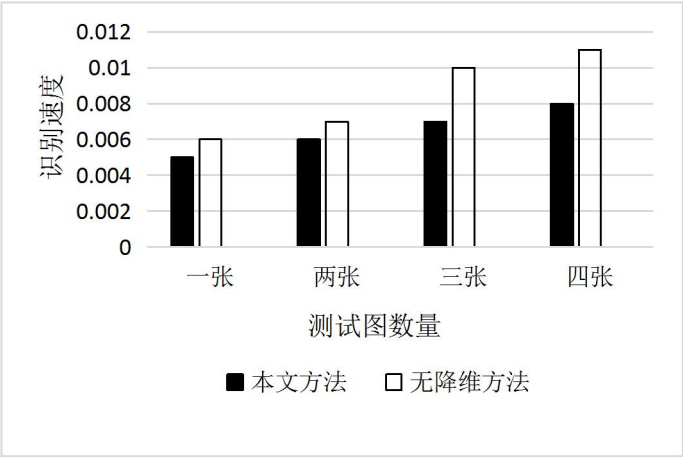


图 5-15 降维对聚类时间的影响

5.4 本章小结

针对家庭环境下机器人进行识别存在的问题进行设计，并在 Webots 仿真平台上进行试验，验证其可行性和实时性。结果表明，本文的设计在提高识别精度的同时能够保证较好的实时性。因此具有一定的实用价值。

第六章 总结与展望

6.1 总结

家庭服务机器人实现商品化作为今后机器人产业发展的主要任务之一，对其智能化的要求会越来越高，其应用的范围也会越来越广。但机器人实现为人类准确实时的解决问题仍然面临着诸多挑战。本文以家庭场景作为机器人的实验环境，针对目前机器人物体识别面临的问题进行实验设计，并在 Webots 机器人开发平台上实现机器人自主避障，图像采集和目标识别的功能。论文主要完成了以下工作：

1. 本文主要运用 Webots 机器人仿真开发平台进行实验测试。在此平台上完成了对 OpenCV 的配置。并对家庭场景和机器人的建模涉及的节点进行原理，功能的介绍和设计，并完成相关建模。之后运用 c++ 进行编程，实现了在该平台上对机器人的运动控制和目标物体的识别。

2. 采用 SURF 算法和 SVM 分类器相结合的方法进行图像特征的提取和训练的同时对 SURF 算法提取的图像特征运用 PCA 算法进行降维并运用 Bag-of-Visual-Word 对降维后的结果进行聚类处理，以此来弥补 PCA 算法处理后的特征缺乏类别信息的缺点。在确保识别精度，提升计算效率的同时，方便后续 SVM 的分类训练。

3. 针对复杂环境背景对物体识别的影响，本文将场景环境作为训练样本之一。同时，在测试环节将其与图像金字塔分割相结合来降低复杂背景的影响。

4. 针对单一视角提取的图像存在信息损失，且进一步提升机器人识别的精度，本文提出运用多视角的思想，即对目标物体多个角度提取图像，以此来增加机器人对目标物体信息的获取。从而在一定程度上克服目标物体存在遮挡对识别的影响。

实验结果表明，尽管 PCA 降维增加了总的识别时间，但它有效地提升了聚类处理的速度且识别时间仍然符合实时性要求。而且，使用图像金字塔分割和多视图方法在一定程度上达到了降低复杂背景和遮挡对识别的影响的目的。因此，本文的设计在确保实时性的同时具有良好的识别精度。由于本文试图在场景建模的过程中尽可能地恢复真实场景，因此，本文的方案具有一定的可行性和研究价值。

6.2 展望

虽然本文的设计取得了较好的实验结果，但依然存在许多的不足之处，下一步的研究工作主要包括：

1. 在现实场景中进行实验测试，验证设计方案，改进实验算法，进一步提升算法的可行性与实用性。

2. PCA 算法是一种线性降维的方法，属于无监督的算法。因此缺少类别信息，虽然本文运用 Bag-of-Visual-Word 算法弥补了上述的不足，但其仍未考虑空间关系，因此可以使用不同的聚类方法来进行研究。

3. 本文未从算法的角度对遮挡问题进行研究，而且，当目标物体存在大面积遮挡时仍无法取得较好的结果，所以，大面积遮挡的问题仍然是今后需要研究的重点。

4. 目前机器学习领域的降维算法还有很多，除了以 PCA 算法为代表的线性降维外，还有非线性的降维算法，如 LLE 等。也可以作为今后研究的方向。

参考文献

- [1] 机器人产业发展规划(2016-2022 年)[EB/OL].
http://www.ndrc.gov.cn/zcfb/zcfbghwb/201604/t20160427_799898.html
- [2] 贺俊耀. 单目视觉家庭机器人环境建模与导航问题研究[D].湖南:中南大学,2009.
- [3] 中国产业信息网[EB/OL].<http://www.chyxx.com/industry/201704/515800.html>,2018-2-15.
- [4] 于海旭.服务机器人的发展动态与趋势分析[J].机械设计与制造工程,2016,45(05):68-70.
- [5] 卞正岗.机器视觉技术的发展[J].中国仪器仪表,2015(06):40-42+65.
- [6] 王利明. 机器视觉中物体识别方法的研究与探讨[D].上海:复旦大学,2009.
- [7] 刘光灿. 基于机器学习的物体识别[D].上海:上海交通大学,2013.
- [8] Mon.Y.J,Lin,C.M.Image processing based obstacle avoidance control for mobile robot by recurrent fuzzy neuralnetwork[J].Journal of Intelligent and Fuzzy Systems,2014:2747-2754.
- [9] 余惠敏. 服务机器人: 广阔市场在招手[N]. 经济日报,2015-08-24(014).
- [10] 东方财富网[EB/OL].<http://finance.eastmoney.com/news/1355,20171207810365236.html>
- [11] 百度百科[EB/OL].<https://baike.baidu.com/item/科沃斯机器人/5657711>,2018-2-16.
- [12] 搜狐[EB/OL].http://www.sohu.com/a/60847855_114790,2018-2-16.
- [13] 中国网[EB/OL].<http://dz.china.com.cn/kj/2017-11-29/89510.html>
- [14] 中商情报[EB/OL].http://www.askci.com/news/chanye/20180131/114427117321_3.shtml,2018-2-16.
- [15] 綦学宇. 城市未来家居设计发展趋势研究[D].吉林:东北师范大学,2010.
- [16] 孙晋博,余隋怀,陈登凯.基于证据理论融合多特征的物体识别算法[J].计算机工程与应用,2015,51(09):147-151.
- [17] 胡小荣. 复杂场景中家庭服务机器人目标识别与人脸识别的研究和实现[D].安徽:合肥工业大学,2011.
- [18] 沈志华,赵英凯,王晓荣,严卉.全自主机器人垂直异构双目视觉研究[J].机械与电子,2005(08):46-48.
- [19] Lange.Sascha,Riedmiller,Martin.Evolution of computer vision subsystems in robot navigation and image classification tasks[J].Robo Cup 2004: Robot Soccer World Cup VIII,2005,3276:184-195.
- [20] EEPW[EB/OL].<http://www.eepw.com.cn/article/201801/374556.htm>,2018-2-18.
- [21] 张桂梅. 完整和部分遮挡物体的识别机理和算法研究[D].陕西:西北工业大学,2006.
- [22] 卢清薇,罗旌钰,王云峰.SURF 算法的降维研究[J].软件,2017,38(12):148-152.

- [23] 张驰. 部分遮挡目标识别算法研究[D].陕西:西安工业大学,2010.
- [24] 徐晓.计算机视觉中物体识别综述[J].电脑与信息技术,2013,21(05):4-6+41
- [25] 曹健. 基于局部特征的图像目标识别技术研究[D].北京:北京理工大学,2010.
- [26] 蔡星艳,赵和鹏,邱鹏,朱长仁.遮挡目标检测与识别技术研究[J].数字技术与应用,2013(09):73-75.
- [27] 李慧霞. 室内智能移动机器人规则物体识别与抓取[D].北京:北京交通大学,2016.
- [28] H.P.Moravec.Towards Automatic Visual Obstacle Avoidance[J].Proc.5th International Joint Conference on Artificial Intelligence.1997,584.
- [29] C.Harris,M.Stephens.A combined corner and edge detector[J].Proceedings in Alvey Conference.1988,189-192.
- [30] D.lowe.Distinctive Image Features from Scale-Invariant Keypoints[J].The International Journal of Computer Vision.2004,2(60),91-110.
- [31] DG.Lowe.Object recognition from local scale-invariant features[J]. IEEE International Conference on Computer Vision , 2002 , 2 :1150
- [32] Ke. Y, Sukthankar. R. PCA-SIFT: a more distinctive representation for local image descriptors[J]. Computer Vision and Pattern Recognition, 2004
- [33] Herbert.B, Tinne.T, Luc.V.G. SURF: Speeded Up Robust Features[J]. Computer Vision ECCV 2006, 2006 ,3951:404-417.
- [34] 柯翔,陈小平,靳国强,王锋,郭群.一种适用于室内服务机器人的实时物体识别系统[J].计算机系统应用,2013,22(10):84-89.
- [35] Matas.J,Chum.O,Urbán.M.et.al. Robust wide baseline stereo from maximally stable extremal region[J].BMVC,2002.
- [36] Timor. K, Michael.B.Saliency,scale and image description,Robotics research Laboratory[J], Department of Engineering Science,University of Oxford.2000.
- [37] Timor .K,Andrew .Z,Michael.B.An affine invariant salient region detector[J]. European Conference on Computer Vision-eccv , 2004 , 3021 :228-241
- [38] T.Tuytelaars,L.Van Gool.Wide baseline stereo matching based on local,affinely invariant Regions[J]. British Machine Vision Conference 2000.
- [39] Tinne.T,Luc.V.G.Matching widely separated views based on affine invariant regions[J]. International Journal of Computer Vision,2004,59(1):61-85.
- [40] 范立南,韩晓微,张广渊.图像处理 and 模式识别[M].北京:科学出版社,2007.
- [41] 高中杰. 基于快速鲁棒性特征的物体识别检测算法研究[D].天津:天津大学,2009.
- [42] 芮执元, 程林章. 基于 Pro/E 与 ADAMS 结合的虚拟样机动态仿真 [J] . 现代制造工程, 2005(1) : 56—58
- [43] 邓艳娇. 基于 Open GL 的四足机器人仿真系统研究与平台构建 [D] . 山东: 山东大学,

- 2015.
- [44] 李思齐. 基于 ODE 的四足机器人动力学仿真平台研究与实现 [D]. 山东: 山东大学, 2015.
- [45] 张继禹, 蔡鹤皋, 王树国, 等. 一个大型机器人仿真系统—ROB-CAD [J]. 哈尔滨工业大学学报, 1993, 25(3): 108—133.
- [46] Smith R. Open dynamics engine v0.5 user guide[J]. Computer Graphics, 2007, 176(2): 121-136.
- [47] 赵小川, 罗庆生, 韩宝玲. 基于 Webots 仿真软件的仿生六足机器人机构设计与步态规划[J]. 系统仿真学报, 2009, 21(11): 3241-3245+3267.
- [48] 辛亚先, 李彬, 洪真. 基于 Webots 仿真软件的四足机器人建模和运动控制[J]. 齐鲁工业大学学报(自然科学版), 2016, 30(02): 45-51.
- [49] 高越. 计算机辅助吊装选型及碰撞检测研究[D]. 宁夏: 北方民族大学, 2016.
- [50] Webots 手册 [EB/OL]: file:///E:/webots8.0.5/Webots/doc/reference/section3.24.html, 2018-3-16.
- [51] 李柳群. 基于 Webots 汽车起重机工作场景的感知与路径规划研究[D]. 宁夏: 北方民族大学, 2017.
- [52] 王铭昌. 基于 Webots 的服务机器人故障仿真平台研究与设计[D]. 山东: 山东大学, 2016.
- [53] 陈佳洲. 室内场景物体同时识别与建模[D]. 广东: 广东工业大学, 2016.
- [54] 马朋飞. 面向移动机器人的全方位视觉系统关键技术研究[D]. 宁夏: 北方民族大学, 2017.
- [55] 方露. 视频中运动目标检测与追踪算法研究[D]. 安徽: 安徽工程大学, 2017.
- [56] 刘瑞华, 耿国华, 周明全, 冯亚. 一种新的等距映射下的降维方法[J]. 计算机应用与软件, 2008(06): 21-22.
- [57] Kazor, K. Holloway, R. Cath. T. et al. Comparison of linear and nonlinear dimension reduction techniques for automated process monitoring of a decentralized wastewater treatment facility[J]. Stochastic Environmental Research and Risk Assessment, 2016, 1527-1544.
- [58] 刘达. 基于 KINECT 的室内目标识别技术研究[D]. 江苏: 江苏科技大学, 2016.
- [59] 降维和 PCA [EB/OL]: <https://www.cnblogs.com/mooba/p/5530577.html>, 2018-3-16.
- [60] Aldavert, D. Rusiol, M. Toledo, R. et al. A study of Bag-of-Visual-Words representations for handwritten keyword spotting[J]. International Journal on Document Analysis and Recognition, 2015, 223-234.
- [61] bag-of-word 模型介绍 [EB/OL]: <https://blog.csdn.net/wsj998689aa/article/details/47089153>, 2018-3-16.
- [62] Phu, V. N., Chau, V. T. N., Tran, V. T. N. SVM for English semantic classification in parallel Environment[J]. International Journal of Speech Technology, 2017, 487-508.
- [63] Chapelle, O., Haffner, P., Vapnik, V. N. Support vector machines for histogram-based image

- classification[J]. Neural Networks, IEEE Transactions on, 1999, 10(5): 1055-1064
- [64] Cortes. C, Vapnik. V. Support Vector Networks[J]. Machine Learning, 1995.
- [65] 王晓龙.基于 ISO26262 的嵌入式车道偏离预警系统仿真与实现[D].宁夏: 北方民族大学, 2016.
- [66] Potthast.C, Breitenmoser.A, Sha.F. et al. Active multi-view object recognition: A unifying view on online feature selection and view planning[J].Robotics andAutonomous Systems,2016,31-47.
- [67] Guo.H,Wang.J,Gao.Y,Lu.H.Multi-View 3D Object Retrieval With Deep Embedding Network[J].IEEE Trans Image Process,2016,5526-5537.

致谢

光阴似箭，岁月如梭。一转眼，三年的研究生生活即将画上句点。回首这三年的生活，虽然喜忧参半，但心中仍然感激北方民族大学给予我宝贵的学习机会。

在学习过程中，遇到过许许多多的难题，庆幸的是得到了来自老师和同学的关心和帮助，因此，在这里我要感谢所有给予我帮助的人。

首先要感谢的是我的导师马行和穆春阳老师。两位老师不仅在学习上给了我很大的帮助，而且在平时的生活上也给予我很大的支持。从两位老师身上我不仅学习到了相关领域的知识，而且提升了自己自主学习和研究的能力。

其次，要感谢我的同门师兄妹们。虽然我们来自祖国的大江南北，但我们相处融洽，每当我在学习和生活中遇到困难时，他们都会毫不犹豫的给予我帮助。在学习期间，我们建立了深厚的友谊，为我的研究生生活增添了别样的乐趣和色彩。

最后，我要感谢我的家人。在我求学的过程中给予了我无条件的支持和关爱，是我最坚强的后盾。

研究生三年的学习过程中，我获益匪浅，收获良多。希望自己今后能成为社会的有用之才，以此来回报大家曾对我的帮助和支持！

靳晓英

2018 年 3 月 15 日，电气信息工程学院

个人简介

攻读硕士期间参与的项目与发表的论文情况如下：

(一)科研项目：

1. 校级创新项目《基于 Webots 的家庭环境下物体识别的设计与实现》，主要对机器人在家庭环境下的物体识别存在的问题进行研究。2017.3-2018.6

(二)论文发表：

1. 《Design and Implementation of Family Service Robots'Object Recognition Based on Webots》，ICDIP 第十届数字图像处理国际会议，2018.3 录用，第一作者