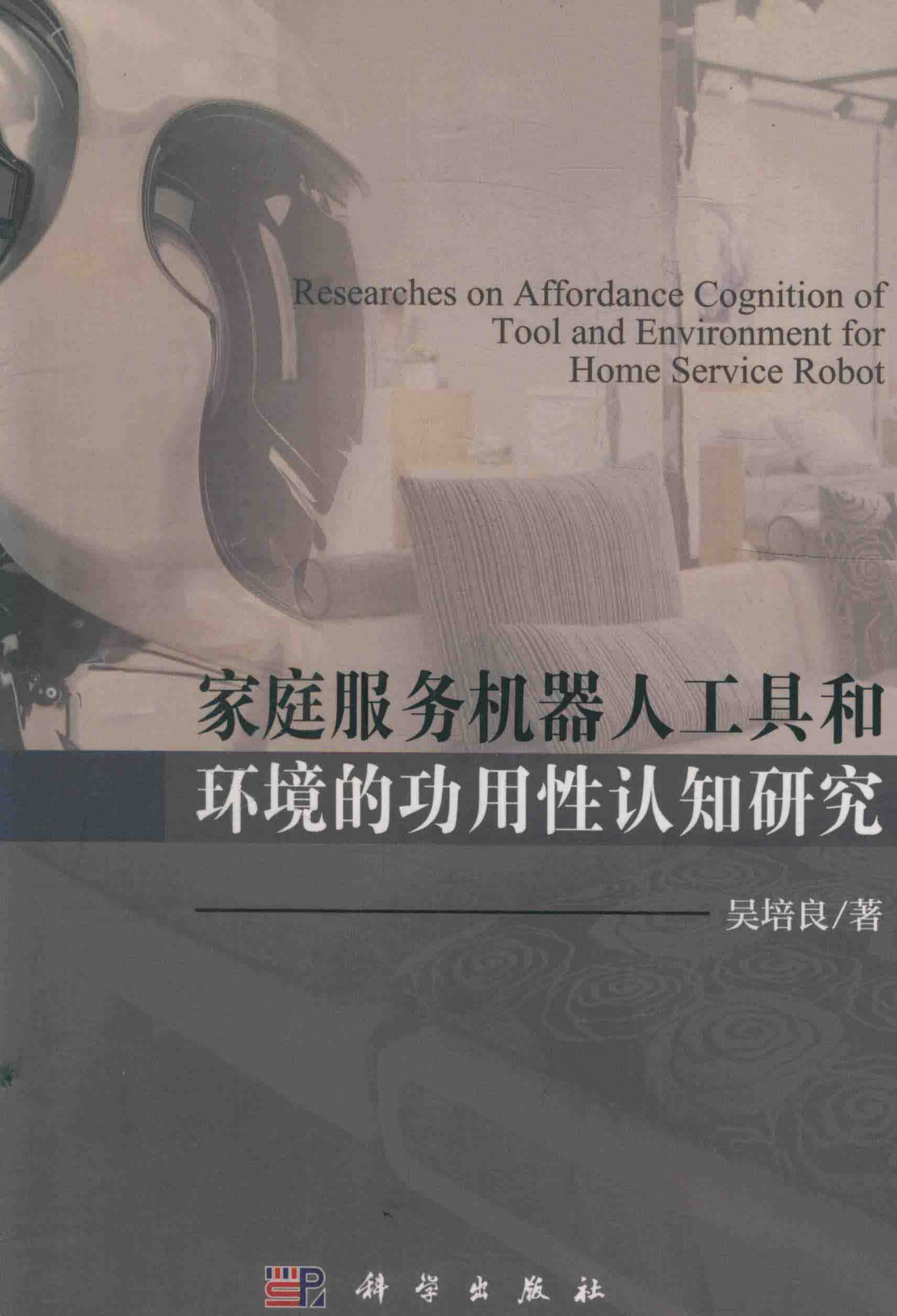




Researches on Affordance Cognition of
Tool and Environment for
Home Service Robot

家庭服务机器人工具和 环境的功用性认知研究

吴培良/著



科学出版社

家庭服务机器人工具和 环境的功用性认知研究

Researches on Affordance Cognition of Tool and
Environment for Home Service Robot

吴培良 著



科学出版社

北京

内 容 简 介

本书系统介绍了家庭服务机器人工具与环境认知研究的最新成果。全书分为工具功能认知、环境功能认知、环境建图三大部分。其中,工具功能认知着重阐述工具功用性部件及工具整体的建模与检测;环境功能认知着重阐述室内功能区建模与分类;环境建图着重阐述家庭全息地图表示与构建,以及物联网机器人系统同时定位、标定与建图。

全书注重系统性、严谨性、理论性和可读性,可以作为高等院校计算机、智能科学与技术、模式识别等专业本科生及研究生的教学用书,也可作为相关专业科研工作者的参考辅导工具书。

图书在版编目(CIP)数据

家庭服务机器人工具和环境的功用性认知研究/吴培良著. —北京:科学出版社, 2018. 5

ISBN 978-7-03-057258-5

I. ①家… II. ①吴… III. ①家政服务-智能机器人-研究 IV. ①TP242.6

中国版本图书馆 CIP 数据核字 (2018) 第 081794 号

责任编辑:周 涵/责任校对:杨 然
责任印制:张 伟/封面设计:无极书装

科学出版社 出版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

北京教图印刷有限公司 印刷

科学出版社发行 各地新华书店经销

*

2018 年 5 月第 一 版 开本: 720 × 1000 B5

2018 年 5 月第一次印刷 印张: 8 1/2 彩插: 4

字数: 181 000

定价: 68.00 元

(如有印装质量问题, 我社负责调换)

前 言

2017年7月8日,国务院印发了《新一代人工智能发展规划》,其开篇指出:人工智能的迅速发展将深刻改变人类社会生活、改变世界。

在机器智能研究领域,2007年比尔·盖茨提出了“机器人进入千家万户”的宏伟设想,十年后的今天,扫地机器人、教育机器人等具备基本服务能力的机器人产品进入普通家庭,极大地方便和改善了人们的日常生活。在人工智能与机器人结合的前沿领域,能够与作业环境、人和其他机器人自然交互,自适应复杂动态环境并协同工作的共融机器人概念已被提出,其研究框架雏形正在逐步形成。

作为共融机器人的基础技术之一,功用性认知研究旨在使机器人像人一样认识日常生活中工具及环境的功能用途。近几年,斯坦福大学的李飞飞研究团队和康奈尔大学的 Ashutosh Saxena 研究团队不约而同地指出,功用性是物品及场景的重点语义特征,这一语义特征在人类认知和活动中起到巨大作用,频繁出现在人们的日常用语中,并且深刻地影响着人们的理解和交互。从某种程度上讲,功用性是一种比名称和空间位置更有意义的语义描述。

事实上,纵观人类文明史,社会每一次进步几乎都与使用工具息息相关;在人的成长过程中,学习使用工具也是必须具备的能力。依此推理,假如拥有了学习和使用工具的能力,毫无疑问,机器人也将迈向更高层次的文明。

而人工智能的发展,为机器人搭建好了阶梯!

受比尔·盖茨先生“机器人进入千家万户”设想的鼓舞,本课题组于2008年明确了家庭服务机器人的研究方向,起初的研究侧重适应机器人服务需要的包含工具等信息的环境地图,我们称之为全息地图。“全息”含义有二:一为空间,二为时间。继而,为了真正从时空维度完整掌握服务环境的空间分布和动态变化,我们将物联网理念引入家庭环境,与机器人有机构成家庭智能空间,研究了家庭智能空间服务机器人全息环境地图相关理论与技术,以及网络机器人系统协同定位、标定与建图问题解耦及算法实现,构建了表征空间分布与实时变化的家庭环境全息地图,有助于机器人全局环境认知与服务任务规划。

但上述研究本质上仍属于感知问题研究,在研究过程中,我们越来越深刻地认识到,要想提高机器智能和共融式宜人化服务水平,仅研究环境地图远远不够,还需要深入认知环境及工具的功能、用途。基于该思想,我们发现并开始关注工具和环境的功用性认知问题,我们认为,作为环境全息地图的有效补充,功用性是连通现实世界与任务世界的桥梁,是由被动认知到主动认知的关键转折。

时至今日,智能算法已经在部件功用性、工具功用性、环境功用性、全息建图等方面崭露头角,这几个方面由点到面,互相关联、互为补充,构成了本书的主体内容。全书由吴培良负责统稿,具体章节撰写分工如下:

第1章由吴培良撰写;第2章由付卫兴、吴培良撰写;第3章由隰晓珺、吴培良撰写;第4章由李亚南、吴培良撰写;第5章由何犇、吴培良、付卫兴撰写;第6章由何犇、吴培良撰写;第7章由李亚南、吴培良撰写;第8章至第10章由吴培良撰写。书稿的校核工作由金昱旻、杨芳负责。

回首全息地图及功用性认知的研究之路,过程绝非一帆风顺。在每次迷茫的时候总能有幸得到我的博士生导师孔令富先生和博士后导师侯增广先生的耐心指导,使我得以坚持至今,感激之情非言语所能表达;感谢张世辉师兄和赵逢达师兄的悉心关怀和点拨;在研究过程中与师弟孔亮博士、段亮亮博士、景荣博士的深入交流至今受益匪浅;还需要感谢的是一路陪我走来的研究生团队,在人工智能研究处于低谷的时候,他们可以顶住压力随我一起潜心研究,并为本书的写作贡献了重要力量。

在本书的撰写过程中,我们参考了许多同类学术论文,吸收了许多观点,并在参考文献中列出,在此由衷地表示感谢。

本书的出版得到了国家自然科学基金项目“家庭智能空间服务机器人全息环境地图相关理论与技术研究”(项目编号:60975062)和“网络机器人系统协同定位、标定与建图问题解耦及算法实现”(项目编号:61305113)、河北省自然科学基金项目“家庭物联网机器人系统同时定位标定与建图问题研究”(项目编号:F2012203199)和“学习型家庭服务机器人免标签环境含功用性认知方法研究”(项目编号:F2016203358),以及中国博士后科学基金项目“家庭服务机器人日常工具功用性认知研究”(项目编号:2018M631620)的资助,在此一并表示感谢。

由于作者水平有限,书中难免存在不足之处,敬请读者批评指正。

吴培良

燕山大学

2018年5月

目 录

前言

第 1 章 基础知识	1
1.1 研究背景	1
1.2 国内外研究现状	2
1.3 本书所用特征	5
1.3.1 深度几何特征	5
1.3.2 SIFT 特征	6
1.3.3 SURF 特征	8
1.4 本书所用算法	8
1.4.1 结构随机森林	8
1.4.2 K 均值聚类	9
1.4.3 SVM	10
1.4.4 RBPF-SLAM	11
1.5 本章小结	12
参考文献	12
第 2 章 基于 SRF 的工具功用性部件建模与检测	15
2.1 系统框架	15
2.2 模型离线训练	16
2.2.1 功用性边缘检测器构建	16
2.2.2 工具部件功用性检测器构建	18
2.2.3 由粗到精阈值选取	19
2.3 工具功用性部件在线检测	19
2.4 实验	20
2.4.1 实验数据集	20
2.4.2 评价方法	20
2.4.3 实验结果分析	22
2.5 本章小结	25
参考文献	25
第 3 章 基于联合学习的家庭日常工具功用性部件检测	27
3.1 系统框架	27

3.2	问题描述与公式化表示	28
3.3	模型优化与解耦求解	29
3.3.1	模型优化	30
3.3.2	CRF 权重求解	31
3.3.3	字典求解	31
3.4	算法实现	32
3.4.1	几何特征表示与提取	32
3.4.2	基于联合学习的模型构建算法	32
3.4.3	功用性部件在线检测算法	33
3.5	实验	33
3.5.1	实验数据集	33
3.5.2	实验条件	34
3.5.3	实验结果分析	34
3.6	本章小结	37
	参考文献	37
第 4 章	基于特征优选和部件组合的家庭日常工具分类	39
4.1	系统框架	39
4.2	特征描述与最优组合选取	40
4.2.1	特征描述	40
4.2.2	ReliefF 最优特征组合选取	41
4.3	基于 BOW 的服务机器人家庭日常工具分类模型构建	42
4.3.1	模型离线构建	43
4.3.2	在线分类检测	45
4.4	实验	46
4.4.1	实验数据集	46
4.4.2	实验结果及分析	46
4.5	本章小结	50
	参考文献	50
第 5 章	部件功用性语义组合的家庭日常工具分类	51
5.1	系统框架	51
5.2	模型离线构建	52
5.2.1	功用性部件边缘特征描述	52
5.2.2	功用性边缘检测器构建	52
5.2.3	基于功用性部件组合聚类的工具字典构建	54
5.3	基于功用性工具字典的家庭日常工具分类	55

5.4	实验	56
5.4.1	实验数据集	56
5.4.2	实验结果与分析	57
5.5	本章小结	60
	参考文献	60
第 6 章	基于空间金字塔池化的工具识别	62
6.1	系统框架	62
6.2	构建工具特征图的空间池化模型	63
6.2.1	工具几何特征描述	63
6.2.2	基于 cciPCA 的多尺度特征块提取	63
6.2.3	空间池化金字塔的构建	64
6.3	多分类的线性分类模型构建	65
6.4	模型选择及实验结果	66
6.4.1	实验数据集	66
6.4.2	实验结果及分析	67
6.5	本章小结	69
	参考文献	69
第 7 章	基于 CLM 模型的服务机器人室内功能区分类方法	71
7.1	系统框架	71
7.2	基于 CLM 算法的功能区图像模型构建	72
7.2.1	提取图像特征描述符	72
7.2.2	图像高斯模型表示与匹配	73
7.2.3	图像高斯模型的优化	74
7.2.4	学习改进的 SVM 分类器	74
7.2.5	室内功能区建模算法描述	75
7.3	在线检测算法	76
7.4	实验	76
7.4.1	实验数据集	76
7.4.2	实验结果及分析	77
7.5	本章小结	78
	参考文献	79
第 8 章	服务机器人家庭全息地图表示与构建	80
8.1	家庭全息地图的表示	80
8.2	家庭全息地图的构建	81
8.2.1	机器人观测模型	81

8.2.2	局部几何特征地图的构建	83
8.2.3	全局拓扑地图的构建	87
8.2.4	几何-拓扑地图的存储结构	89
8.3	实验	90
8.3.1	实验描述	90
8.3.2	家庭环境地图的构建	92
8.3.3	实验分析	92
8.4	本章小结	93
	参考文献	94
第 9 章	物联网机器人同时标定与定位算法	95
9.1	系统模型	95
9.1.1	系统构成	95
9.1.2	机器人运动模型	95
9.1.3	摄像机对机器人的观测模型	96
9.1.4	摄像机投影矩阵的标定	98
9.2	基于 RBPF 的物联网机器人定位	98
9.2.1	基本思想	98
9.2.2	构建位姿粒子的提议分布	99
9.2.3	计算粒子权重及重采样	100
9.2.4	估计投影矩阵	102
9.3	实验	103
9.4	本章小结	105
	参考文献	105
第 10 章	物联网机器人系统同时定位、标定与建图方法	106
10.1	系统描述	106
10.1.1	系统构成	106
10.1.2	问题简化	107
10.1.3	模型建立	107
10.2	物联网机器人系统同时定位、标定与建图	109
10.2.1	基本思想	109
10.2.2	传感器网络辅助机器人位姿估计	110
10.2.3	粒子权值计算及重采样	113
10.2.4	传感器网络标定	115
10.2.5	环境特征地图构建	117
10.2.6	同时定位、标定与建图的完整算法	119

第1章 基础知识

1.1 研究背景

随着社会进步和机器人技术的不断发展,特别是在比尔·盖茨提出“机器人进入千家万户”的战略设想后,家庭服务机器人逐渐受到人们关注,此外,在人口老龄化日益严重的今天,智能服务机器人进入家庭为老人、残障人士提供服务的需求越来越迫切.在这样的背景下,众多发达国家都对家庭服务机器人的研究高度重视,相继制订了相应的战略发展规划.我国科学技术部也组织编制了《服务机器人科技发展“十二五”专项规划》,将服务机器人研究作为“十二五”期间机器人技术研究的重要部分.国家自然科学基金委员会也启动了“共融机器人基础理论与关键技术研究”重大研究计划,明确提出了对能够与作业环境、人和其他机器人自然交互、自主适应复杂动态环境并协同作业的共融机器人的概念及其研究方向.在这些政策导向的推动下,国内外很多著名高校和学者先后投身其中,家庭服务机器人的研究逐渐进入了高速发展时期.

尽管已有众多的高校和科研机构对服务机器人展开了研究,但是目前服务机器人的研究仍处于一个较低的水平,相比人们期望的机器人进入日常家庭,提供宜人化服务还有一定差距,其原因主要在于机器人自身智能程度的限制以及家庭环境的复杂性.在家庭环境下,服务机器人要完成一项服务任务(如获取和操作工具),往往涉及“用到哪些工具?”“工具是什么样的?”“工具在哪里?”“我在哪里?”和“我怎样到达工具处?”等问题,其中,前两个问题属于工具知识表示与推理的范畴,后三个问题则可通过有效的地图模型寻求解决.事实上,在服务机器人研究领域,有效的工具知识表示和环境地图模型是机器人理解并适应环境,继而完成路径规划和定位导航,并最终提供自然交互服务的基础.

服务机器人对工具和环境的理解研究经历了两个阶段,即感知阶段(面向导航)和认知阶段(面向自然交互).在感知阶段,工具表示主要基于颜色、纹理等底层表观特征,环境也主要以几何、栅格、拓扑等地图形式描述,该阶段缺乏对物品功用性等语义的描述,且地图信息也过度简化而仅适用于路径规划和定位导航,对机器人环境理解与人自然交互帮助甚少.发展到认知阶段,工具的中层和高层语义信息得到重视,并且通过在传统感知地图中加入这些语义信息,初步建立了与人类认知相兼容的环境描述,即认知地图.

在人类认知过程中,对家庭工具的认知通常通过有参照地学习(聚类和监督式

学习的思想) 同类工具的表现和语义来构建该类工具的模型, 对家庭场景的认知则根据其中所含标志性工具将场景赋予特定语义信息。总体来看, 这些语义信息在人类活动中起着至关重要的作用, 频繁出现在人们的日常用语中, 并且深刻地影响着人们的理解和交互。传统感知型机器人正是由于没有考虑工具及场景的语义, 因此无法以与人相同或相似的方式认知服务环境, 故而难以与人形成自然交互而做到人性化服务。

目前, 认知环境及其中功能区场景、工具的研究已经初步展开, 但问题还远未到解决的程度, 通用的、实用化的工具与功能区知识表示及认知地图表示方法仍需要进一步深入研究。在认知层面上, 当前研究主要聚焦在工具名称及空间关系语义上。但根据李飞飞等研究团队的最新调查研究, 功用性这一语义特征在人类认知中起到更大作用, 某种程度上讲, 功用性是一种比名称及空间关系更有意义的语义描述, 将其融入环境认知中将极大地提升机器人认知水平。此外, 目前可认知环境中工具及功能区的语义信息大多依靠所粘贴的语义标签来提供, 机器人通过阅读标签获取语义, 属于一种被动式认知方式, 且标签标记存在负担繁重、部分工具标记困难的问题。随着图像处理和机器学习理论的发展, 模拟人类的环境认知与学习机制, 使机器人自主识别工具与场景已成为服务机器人领域的一个重要研究方向。

此外, 对于工作在家庭这种动态复杂环境的机器人, 受自身感知范围所限, 机器人难以实时地获得整个家庭的动态信息, 不利于其提供快捷高效的服务。物联网技术将日常物品连成网络并进行智能化识别、定位、跟踪、监控与管理。通过将物联网与服务机器人有效结合, 可实现两者的优势互补: 一方面, 物联网为机器人提供全局感知, 弥补机器人全局感知能力弱的缺陷; 另一方面, 机器人可视为物联网的执行机构, 从而使物联网具备主动服务能力。可见, 物联网机器人系统是机器人进入家庭提供智能服务的可行发展方向。

1.2 国内外研究现状

本书围绕基于学习的家庭物联网机器人系统工作环境深层认知关键理论与技术展开探讨, 所研究的问题属于机器学习、状态估计等理论及其在机器人环境理解与认知方面的应用。由于家庭环境认知的复杂性, 目前国内外研究主要集中在工具识别、室内场景识别, 以及环境地图表示与构建三个方面独立展开。

1.2.0.1 工具识别研究

工具识别是服务机器人应具备的基本功能。目前, 工具识别算法大致可以分为三个类别: 第一类以 Dalal 和 Triggs 提出的基于支持向量机分类器的方向梯度直

方图 (histogram of oriented gradient, HOG) 特征刚性检测器为代表^[1], 作为刚性模型的扩展, 变形部件模型^[2] 在很多基准测试中取得了不错的效果. 第二类建立在 Viola 和 Jones^[3] 工作基础之上, 使用了 Boosting 方法和各种不同的特征通道. 此类检测器在处理工具轮廓方面不够灵活, 但是在行人数据集下具有很好的识别精度, 并可达到在线识别. 第三类为基于投票机制的工具中心预测器^[4]. 上述方法均在 RGB 图像上进行处理, 未考虑场景图像的深度信息. 目前随着可同时获取颜色和深度数据的低成本 RGB-D 相机的出现, 基于 RGB-D 进行工具识别逐渐引起学者的关注. 与此同时, 受特征学习研究的推动, 机器人领域的学者们开始对其展开研究并应用于 RGB-D 工具识别. Bo 等^[5] 通过将核观点引入深度图和 3D 点云提出了 RGB-D 核描述符, 并在 RGB-D 工具数据集下得到了更高的识别精度. Blum 等^[6] 将 K 均值 (K -means) 聚类特征学习用于 RGB-D 工具识别. 最近, 卷积神经网络 (convolutional neural networks, CNN) 被用来训练并提取图像的 CNN 特征, 以进行场景中工具的识别^[7], 实验结果表明, 基于 CNN 特征的工具识别算法准确率较其他常规算法明显提高, 该研究还指出, 在 RGB 图像上提取 CNN 特征的方法同样适用于深度图像, 且将 RGB 和深度信息结合可生成更加丰富的特征. 上述方法虽然在识别准确率上有不断提高, 但大部分都基于滑动窗口机制, 由于 RGB-D 图像搜索空间巨大, 因此很难达到在线识别. 将视觉显著性计算、随机森林、 K 均值等方法引入工具检测和识别算法中, 为实现在线识别提供了一种有效的解决策略, 在该方面本课题组已经展开初步研究^[8-10].

此外, 受自然交互和宜人服务任务驱动, 机器人已能识别工具的一些基础语义信息, 如工具名称、种类等. 近年来, 功用性 (affordance) 被提出来^[11,12], 其因连通现实世界与任务世界而被认为是语义的重要组成部分. 借鉴人类认知方式, 使机器人具备工具部件功用性认知能力, 对机器人主动智能提升具有重要意义. 目前, 机器人主要通过读取语义标签方式被动获取工具功用性等语义, 基于推理学习的主动认知方法研究刚刚出现^[9-13].

1.2.0.2 室内场景识别研究

室内场景识别在机器人和计算机视觉领域是一个非常具有挑战性的开放问题. Lazebnik 等^[14] 在该方面做出了开创性的工作, 随后场景识别主流方法致力于利用表观特征或空间信息来构建场景图像的全局表示, 并在室外场景识别时取得了很好的效果, 但针对室内场景表现不佳, 其原因有二: 第一, 室内工具存在形式的多样性, 造成同类场景呈现较大差异; 第二, 人造环境中不同场景类间存在相似性. 为了弥补低层特征表示方法的不足, 一些中层和高层语义信息被用来进行图像的表达. 文献 [15] 提出了一种考虑不同视觉单词间的空间关系的空间金字塔匹配模型. 文献 [16] 建立了一种语义框架, 通过基于隐藏工具的语义表示方法来预测图

像类别。Wan 等^[17]提出了一种更加有效的高层特征表示方式,不仅考虑到场景图像中出现的工具,还将各个工具之间的空间位置关系考虑其中,包括前、后、上、下、左、右、包含、被包含、近、远。

传统视觉感知模型通过识别场景的工具推断场景类型,然而目前的研究成果表明人类对图像的局部细节并不敏感。心理学家则预测功用性是实现人类感知外部环境的关键^[18]。将所含工具功用性、布局(位置)及空间关系等语义信息融合起来,是提高场景识别准确率的行之有效的研究方向,本课题组提出一种规避码本的室内功能区表示与建模方法^[19]。

1.2.0.3 环境地图表示与构建研究

地图是服务机器人理解所处环境,实现自主导航以及与人进行交互的基础。环境地图的精度和内容的完备性将直接影响机器人的服务质量。从地图形式和内容看,对环境的描述主要分为感知类地图和认知类地图两大类。感知地图研究较早,发展到现在主要有几何地图、拓扑地图、混合地图等。认知地图^[20,21]也称语义地图^[22,23]、全息地图^[24,25]等,是一种集成了工具空间分布及自身属性,同时包括各种上下文语义及其推理关系的功能强健的新型环境模型。文献^[26]提出了空间混合级联地图模型,文献^[22]通过识别粘贴在大工具上的 QR Code 标签,构建了含大工具功能属性和归属关系描述的三维栅格语义地图。

在环境地图构建方面,基于概率方法的机器人同时定位与建图(simultaneous localization and mapping, SLAM)仍是目前的主流研究方法^[27,28]。Rao-Blackwellized 粒子滤波(Rao-Blackwellized particle filtering, RBPF)同时具备扩展卡尔曼滤波(extended Kalman filter, EKF)和粒子滤波的优势,已被成功应用于 SLAM 中^[28]。尽管国内外学者对 SLAM 问题已经进行了深入研究,定位与建图的精度得到了大幅提升,然而,此过程中传感器机载,机器人观测误差与运动误差相耦合,导致定位和建图误差会随机器人运动距离发生不可避免的扩散。

综上所述,现有认知地图更多地关注表现特征和工具空间关系等语义信息,而忽略了功用性这一关键语义。此外,通过人工标签获取工具语义的方式存在标注工作繁重、适用性有限等缺点,且因破坏环境原貌、不符合日常习惯而显得不够自然、和谐。

1.2.0.4 发展动态分析

家庭物联网机器人系统环境认知是当前服务机器人研究的重要热点问题,但该问题的解决还有很长的路要走。目前研究仍处于初级阶段,存在以下难点问题。

(1) 完整有效的环境知识表示模型尚待研究建立。必须明确,仅依靠认知地图并不能达到自然交互、宜人服务级别的环境认知。尽管认知地图能够描述家庭环境

中场景与工具间的包含关系及空间上下文语义,使机器人具备了初步的服务空间概念,但这对达到自然交互的要求还远远不够.受人类认知启发,分别针对工具及功能区构建包含更深层次语义信息的知识库,并研究其与认知地图的关联关系,将是一种提升服务质量的有效途径.

(2) 机器人环境认知的主动学习能力尚需提高.当前,机器人对环境及其组成成分的表现特征已可通过各类传感器获取并分析,但对语义层面的特征则需要借助语义标签方式由人告知.这种被动式获取方式,一方面限制了机器人在自然家庭环境中的应用,另一方面,对机器人智能增长亦无推动.因此,对服务机器人环境认知的主动学习方法亟待展开研究.

(3) 工具、功能区的语义内涵有待扩展.面向自然交互,机器人需要与人类认知方式相兼容,因此,现有工具及功能区的语义描述需要扩展.近几年斯坦福大学李飞飞研究团队^[11]和康奈尔大学 Ashutosh Saxena 研究团队^[12]不约而同地指出,功用性是工具及功能区的重点语义特征.此外,对功能区功用性的认知还需要参考其中工具功用性、布局及空间关系等诸多因素联合确定.

综上所述,服务机器人对服务环境的理解已由感知层面深化到认知层面,相关研究也已由对物理空间和表现特征空间的建模转移到语义空间下的建模,但该方面研究刚刚开始,亟待展开进一步深入研究.综合研究现状以及作者本人和所在课题组在环境认知方面已展开的研究工作,本项目组提出在免标签自然家庭环境下,基于学习的家庭物联网机器人系统自主环境认知研究课题,对提升机器人环境认知深度、广度及自主性方面所涉及的关键理论和方法展开综合性的研究,针对工具及功能区的知识表示、知识库构建与更新、与知识库关联的认知地图分层描述与物联网机器人联合建图等问题提出较为系统的解决方案.

1.3 本书所用特征

1.3.1 深度几何特征

本书根据深度图像计算功用性边缘检测模型对应的几何特征,其中平均曲率为微分几何中反映曲面弯曲程度的内蕴几何量,记为 f_{MC} ,主曲率为 (k_1, k_2) , $k_1 > k_2$, 则 $f_{MC} = (k_1 + k_2)/2$. 梯度幅值和方向梯度直方图特征是用来进行物体边缘检测的有效特征描述子. 形状指数 (SI) 和曲度 (CV) 表征表面在不同方向的弯曲,体现人对形状的感知. 形状指数 SI 和曲度 CV 的计算公式如下:

$$SI = -\frac{2}{\pi} \arctan\left(\frac{k_1 + k_2}{k_1 - k_2}\right), \quad CV = \sqrt{\frac{k_1^2 + k_2^2}{2}} \quad (1.1)$$

表面法向量是几何体表面的重要属性. 本书从深度数据恢复 3D 点云,再从中

估计出 3D 表面法向量, 并去除样本块均值, 使得视角变化时表面法向量特征鲁棒性更强。

高斯曲率 (Gaussian curvatures) 同平均曲率一样是曲面论中重要的内蕴几何量, 记为 f_{GC} , 则 $f_{GC} = k_1 k_2$ (k_1 和 k_2 为曲面上一个点的两个主曲率)。联合高斯曲率和平均曲率可以确定 8 种曲面类型: 峰、脊、鞍形脊、最小面、平面、阱、谷和鞍形谷, 有助于识别不同功用性的内部结构。

1.3.2 SIFT 特征

尺度空间极值检测 首先在图像尺度空间通过高斯差分 (difference of Gaussian, DoG) 函数查找潜在的兴趣点, 初步确定这些点的位置和尺度。DoG 定义为两个不同尺度的高斯滤波器的差分, 即

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (1.2)$$

式中, $G(x, y, k\sigma)$ 为二维高斯函数, σ 为高斯正态分布的方差, $I(x, y)$ 表示原始图像, $*$ 表示卷积运算。

如图 1.1 所示, 在检测尺度空间极值时, 图中标记为叉号的像素为当前处理的像素点, 它需要跟包括同一尺度的周围邻域 8 像素和相邻尺度对应位置的周围邻域 9×2 像素共 26 像素进行比较, 如果该点为局部最小值或最大值点, 则该点为兴趣点。最底部和最顶部的 DoG 尺度图像, 由于没有下一级图像和上一级图像, 因此不再进行极值点检测。

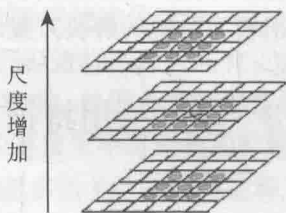


图 1.1 DoG 尺度空间局部极值检测

兴趣点初选择 以上获得兴趣点是候选兴趣点, 需要对这些兴趣点进行检验以保证可靠性。首先删除对噪声敏感的低对比度或在边缘定位差、不稳定的兴趣点, 以增强匹配稳定性、提高抗噪声能力。为了进一步提高定位的稳定性, Lowe 利用一个 3D 二次函数拟合兴趣点周围的采样点来寻找插值的最大值^[29], 不再简单地将兴趣点位置选择为中心采样点的位置。实验表明, 精确确定候选兴趣点的位置对兴趣点的匹配也有极大的改善。

确定兴趣点方向向量 排除掉不稳定兴趣点和精确定位兴趣点后, 利用兴趣

点邻域像素的梯度方向分布特性为每个兴趣点指定方向参数, 使特征具备旋转不变性. 兴趣点 $[x \ y]$ 处的邻域内梯度大小 m 和梯度方向 θ 的计算公式如下:

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (1.3)$$

$$\theta(x, y) = \arctan 2((L(x, y+1) - L(x, y-1)) / (L(x+1, y) - L(x-1, y))) \quad (1.4)$$

实际计算时, 在以兴趣点为中心的邻域窗口内采样, 并用直方图统计邻域像素的梯度方向. 方向梯度直方图的范围是 $0^\circ \sim 360^\circ$, 其中每 10° 为一个组 (bin), 总共 36 个 bin. 直方图的峰值则代表了该兴趣点处邻域梯度的主方向, 即作为该兴趣点的方向. 为了增强匹配的鲁棒性, 在方向梯度直方图中, 当存在另一个相当于主峰值 80% 能量的峰值时, 则认为这个方向是该兴趣点的辅方向. 一个兴趣点可能会被指定具有多个方向 (一个主方向, 一个以上辅方向).

至此, 图像的兴趣点已检测完毕, 每个兴趣点的位置、所处尺度和方向已经确定, 下一步是生成对应每个兴趣点的特征描述符.

生成尺度不变特征变换 (scale-invariant feature transform, SIFT) 兴趣点特征描述符 首先将坐标轴旋转为兴趣点的方向, 以确保旋转不变性. 然后以兴趣点为中心取 8×8 的窗口. 图 1.2(a) 的中央黑点为当前兴趣点的位置, 每个小格代表兴趣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, 箭头长度代表梯度大小, 图 1.2(a) 中圆圈代表 0.5 倍高斯函数加权范围 (越靠近兴趣点的像素梯度方向信息贡献越大). 最后在每 4×4 的小块上计算 $0^\circ, 45^\circ, \dots, 315^\circ$ 共 8 个方向的方向梯度直方图, 绘制各梯度方向的累加值形成种子点, 如图 1.2(b) 所示.

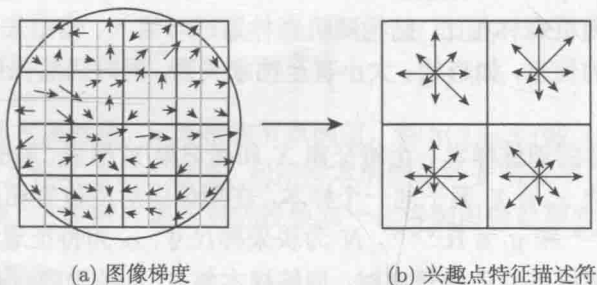


图 1.2 SIFT 特征描述符的生成

图 1.2 中一个兴趣点由 2×2 共 4 个种子点组成, 每个种子点有 8 个方向向量. 这种邻域方向性信息联合的思想增强了算法抗噪声的能力, 同时对于含有定位误差的特征匹配提供了较好的容错性.

此时 SIFT 特征向量已经去除了尺度变化、旋转等几何变形因素的影响, 再继续将特征向量的长度归一化, 则可去除光照变化的影响, 使其对亮度变化不敏感.