

内 容 提 要

自然环境中视觉导航的研究涉及到计算机视觉的许多方面,是一个极有意义但又颇具难度的综合性课题。视觉导航的基本要求是机器人能够自主进行全局定位、道路跟踪和障碍物检测,其中的每一部分都要以环境建模为基础。本论文提出了面向任务的多尺度全视野环境综合建模方法,将路标定位的全景图像建模、道路识别的全方位图像建模和障碍物检测的双目注视图像建模相结合,改进了依赖局部、单一信息的视觉导航方法。本论文研究了灵巧传感器设计、鲁棒视觉算法和多层次模型表示等问题,主要贡献有:基于图像稳定和全景外极面图像分析的3D环境建模方法,全方位图像特征和神经网络相结合的道路建模方法,以及基于重投影变换的双目障碍物检测方法。论文给出了大量室外景物建模的实验结果。

图书在版编目(CIP)数据

全视野时空视觉导航:真实景物的成像、建模与表示/
朱志刚. —北京:高等教育出版社, 2001. 12

研究生教材

ISBN 7-04-010417-2

I. 全... II. 朱... III. 计算机视觉—导航—研究
生—教材 IV. TP302.7

中国版本图书馆 CIP 数据核字(2001)第 079248 号

全视野时空视觉导航——真实景物的成像、建模与表示
朱志刚

出版发行 高等教育出版社

社 址 北京市东城区沙滩后街 55 号 邮政编码 100009

电 话 010-64054588 传 真 010-64014048

网 址 <http://www.hep.edu.cn>

<http://www.hep.com.cn>

经 销 新华书店北京发行所

排 版 高等教育出版社照排中心

印 刷

开 本 850×1168 1/32

印 张 7.625

字 数 190 000

插 页 1

版 次 年 月 第 版

印 次 年 月 第 次印刷

定 价 12.60 元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究

责任编辑 康兆华
封面设计 张楠
责任绘图 郝林
版式设计 史新薇
责任校对 朱惠芳
责任印制

前 言

自然环境中视觉导航的研究涉及到计算机视觉中的各个主要方面,是一个极有意义但又颇具难度的综合性课题。机器人视觉导航的基本要求是机器人能够自主进行全局定位、道路跟踪和障碍物检测,其中每一部分都要以环境建模为基础。而真实景物的环境建模又离不开三个主要部分:灵巧视觉成像、鲁棒视觉算法和多层次视觉表示。本书是笔者1997年在清华大学所做博士论文的修订版。在论文中笔者系统、深入地研究了视觉导航中的自然环境建模问题,主要内容包括以下几个方面。

1. 提出了面向任务的多尺度全视野的环境综合建模方法。将路标定位的全景图像建模、道路识别的全方位图像建模和目标(障碍物)检测的双目注视图像建模相结合,改进了依赖局部、单一信息的视觉导航方法。研究了面向任务的环境建模方法中的传感器设计、视觉算法和模型表示等问题。

2. 提出了由非精确摄像机运动下的图像序列建立3D环境全景模型的两步法,即基于运动滤波的图像稳定和基于时-空-频域遮挡模型的全景外极面图像分析。本文推广了全景图像方法和外极面图像方法,使之能适用于具有抖动特征的图像序列分析,并避免了一般运动视觉方法的不适定问题、特征对应问题,以及基于空域约束的迭代方法的局部最小化等问题。

3. 提出了全方位图像特征和神经网络相结合的道路建模方法,较好地解决了机器人依赖局部视野信息导致迷路、视觉算法依赖特定环境特征导致推广性差等问题。本文提出了主分量分析和 Fourier 变换相结合的全方位图像特征空间数据压缩和旋转不变特征提取方法,设计了基于识别道路类型进行道路方向估计的组合神经网络,从而提供了一种使机器人自动适应不同道路类型的有效方法。

4. 提出了基于重投影变换的障碍物检测方法,超越了传统立体视觉方法中特征抽取、匹配和三维恢复的模式。设计了无特征提取和对应的立体视觉新算法,并给出了克服摄像机俯仰影响的动态重投影变换算法,增强了系统在颠簸道路上运行的自适应性。利用重投影变换后的双目图像对中的路面特征零视差的特性,通过障碍物判断、区域确定和三维测量的“分步渐进”过程,实现了对路面障碍物高效可靠的实时检测。

5. 系统实现 (1) 实现了全景建模中的图像稳定、遮挡恢复和深度分层,从而为机器人全局定位的自然路标提取和真实环境再现的图像合成建立了基础 (2) 设计和实现了适用于室外自然环境的全方位成像系统和单摄像机双目立体成像系统 (3) 实验验证了全方位道路图像神经网络学习方法的有效性 (4) 实现了实时障碍物检测系统,通过大量的室外道路环境下障碍物检测的试验,说明该系统具有很强的实用性。

由于写作时间有限,书中若有疏漏或不妥之处,敬请读者不吝指正。

朱志刚

2001 年 5 月

Preface

Visual navigation of a mobile robot in natural environment has always been a very interesting but challenging problem. It involves almost every aspect of computer vision issues : from visual sensors , robust algorithms to visual representation. The fundamental tasks of visual navigation include global localization , road following and obstacle detection. Visual environment modeling is the foundation of visual navigation , and to some extent , most of the real world problems in computer vision. Three key components , smart visual sensing , robust visual algorithms and hierarchical visual representations , are equally important in real scene modeling. This book is based on the author 's Ph. D. thesis at Tsinghua University. The book presents a systems approach to visual modeling of natural scene for robot navigation and in large to scene understanding and representation. There are four main contributions in the thesis.

1. A purposive , multi-scale and full-view visual scene modeling approach is proposed for visual navigation in natural environment. It combines three novel modules into an integrated system , the POST (Panoramic vision for scene modeling , Omnidirectional vision for road understanding and STereo vision for obstacle detection). This

approach tries to overcome the drawbacks of traditional visual navigation methods that have mostly depended on local and/or single view visual information. However, the proposed approach is not just a simple combination of the three novel sensors and methods, but rather a systematic integration under the strategy of purposive vision (“the right way for the right work”), and under the philosophy of systems approach which emphasizes that “the whole is more than sum of its components”. Thus, right sensor design, adequate level of scene representation and corresponding robust algorithms are specifically explored for each given task while the interconnection among the vision sub-systems are taken into consideration under the overall goal of autonomous navigation.

2. A two-stage method is presented for 3D panoramic scene modeling from image sequences captured by a video camera subject to vibration on a common road surface. Firstly, a 3D image stabilization method is proposed which eliminates vibration from vehicle’s motion so that “seamless” epipolar plane images (EPIs) and panoramic view images (PVI) can be generated. Secondly, an efficient panoramic EPI analysis method is proposed to combine the advantages of both PVI and EPIs efficiently in three important steps: locus orientation detection, motion boundary localization, and occlusion/resolution recovery. The two-stage method not only combines Zheng Tsuji’s PVI method with Baker Bolle’s EPI analysis, resulting in the so-called panoramic EPI method, but also generalizes them to handle image sequences subject to unpredictable camera vibration. Finally, a compact layered representation for a large-scale scene has been proposed, which can be used in both visual navigation and image-based rendering. Since camera calibration, image segmentation, feature extraction and matching are avoided, all the proposed algo-

rithms are fully automatic and rather general.

3. A new road following approach , the Road Omni-View Image Neural Networks (ROVINN) , has been proposed. It combines the omnidirectional image sensing technique with neural networks in such a manner that the robot is capable to learn recognition and steering knowledge from the omnidirectional road images which in turn guarantee the robot never to get lost. The ROVINN approach brings Yagi 's COPIS (Conic Omnidirectional Projection Image Sensor) method to the outdoor road scene and provides an alternative solution different from the CMU 's ALVINN. Compact and rotation-invariant image features are extracted by integrating the Principle Component Analysis (PCA) and the Discrete Fourier Transform (DFT). The modular neural networks of the ROVINN estimate road orientations more robustly and efficiently by classifying the roads as the first step , which enables the robot to adapt to various road types automatically.

4. A novel method called Image Reprojection Transformation (IRT) is presented for stereo-vision-based road obstacle detection. Obstacle detection is modeled as a reflexive behavior of detecting anything that is different from the planar road surface. Dynamic re-projection transformation algorithms are developed so that the algorithms can work in rough road surface. The novelty of the (dynamic) re-projection transformation method , which assembles the gaze control of the human vision , lies in the fact that it brings the road surface to zero disparity so that the feature extraction and matching procedures of the traditional stereo vision are avoided in the proposed obstacle detection algorithms. The progressive processing strategy from yes/no verification , focus of attention , to 3D measurement based on the re-projection transformation makes the obstacle detec-

tion efficient , fast and robust.

5. To validate the proposed strategies and methods , we have implemented the following algorithms and systems.

(1) 3D Scene modeling system.

In the 3D panoramic scene modeling , the algorithms of motion filtering and image stabilization , kinetic occlusion detection and depth layering , have been developed , and the novel depth layered setting(DLS) models have been constructed for many image sequences. These efforts found a ground base for landmark selection of global localization and image synthesis of photo-realistic image-based rendering.

(2) Designs of novel sensors.

An omni-view image sensor is designed and realized , and its properties for the outdoor road understanding are thoroughly studied. A patented single camera binocular vision system is also designed and constructed , and has been put into real use of obstacle detection. An integrated full view smart sensor POST (Panoramic , Omnidirectional and STereo vision) is proposed , which integrates 360-degree omnidirectional view , binocular forward view as well as both left and right side views by using a single camera.

(3) Real scene experiments.

Experimental results of training and testing the ROVINN using real road images have shown that the proposed method for road following is quite promising. A real – time visual obstacle detection system has been set up and tested on the outdoor road scenes extensively.

致 谢

我的博士学位论文是在博士生导师石纯一教授和博士生导师徐光祐教授的指导下完成的。石纯一教授严谨的治学态度和富有启发性的指导使我受益匪浅。徐光祐教授引导我进入视觉环境建模的研究领域,本论文的许多思想方法和视觉导航体系的建立与他的指导是密不可分的。他对问题深刻的洞察力和超乎寻常的远见使我学到了超越专业研究范围的东西,并将使我终身受益。林学言教授在视觉算法的深入研究方面给予我许多启迪和帮助,并共同提出了重投影变换的原理,他那追求创新的精神对我永远是巨大的鞭策。石定机教授启发我设计新型的视觉传感器,并且共同研制了双目立体成像系统和全方位成像系统,他使我时刻牢记:研究工作必须与工程和应用相结合。在论文的撰写过程中也得到吴伯新高级工程师、陈永康高级工程师等老师的帮助,在此深表感谢。

在做论文的过程中还得到清华大学计算机科学与技术系视觉实验室研究生、工作人员和我家人的协助和关心。杨波先生帮助我实现了实时障碍物检测系统,图像运动检测和图像稳定的算法是由杨雨东先生、张强先生帮助实现的。尤其要感谢我的妻子李晓彦,她是我十几年如一日坚持研究工作的强大后盾。

本书的出版应该感谢我的博士学位论文评审专家及答辩委员会,中国首届优秀博士学位论文评审委员会和高等教育出版社,他们

卓有成效的工作和极具价值的评审意见促使我的论文能够入选中国首届优秀博士论文,并在高等教育出版社正式出版。在本书成稿的过程中,正值笔者在美国安城麻省大学(UMASS-Amherst)作访问教授、资深研究员,UMASS 计算机系视觉实验室的联合主任 Edward M. Riseman 教授和 Allen R. Hanson 教授不仅给予诸多便利的研究条件,而且还提出许多极有价值的意见和建议。在此一并表示衷心的感谢。

目 录

第 1 章 绪论	1
1.1 问题的提出	1
1.2 研究历史与现状	3
1.2.1 Marr 视觉理论	3
1.2.2 DARPA 视觉导航	5
1.2.3 障碍物检测	10
1.2.4 全方位视觉和全景视觉	11
1.2.5 图像镶嵌和运动分层表示	15
1.3 论文工作	19
1.3.1 概述	19
1.3.2 论文的主要贡献	21
1.3.3 论文的结构	22
第 2 章 场景的全景图像建模	24
2.1 引言	24
2.2 车体运动模型	26
2.3 运动滤波和图像稳定	28
2.3.1 运动矢量估计和运动参数计算	29
2.3.2 帧间运动分解和图像预校正	29
2.3.3 运动滤波和图像稳定	32
2.3.4 全景图像和外极面图像生成	37

2.4	全景外极面图像分析方法	41
2.4.1	运动纹理模型和运动遮挡模型.....	42
2.4.2	大窗口方向检测算子 GFOD	46
2.4.3	全景外极面图像深度估计方法.....	50
2.5	全景图像建模过程和表示	59
2.5.1	图像初始校正和图像稳定.....	60
2.5.2	运动纹理方向估计和运动边界定位.....	62
2.5.3	运动和形状的融合.....	62
2.5.4	深度估计结果分析.....	64
2.5.5	分层表示.....	67
2.6	3D全景模型的应用	71
2.6.1	任意视角图像的合成.....	71
2.6.2	选择机器人定位的显著路标.....	73
2.7	总结和讨论	73
2.7.1	全景建模方法的优点.....	73
2.7.2	方法和算法的意义.....	74
2.7.3	进一步的工作.....	75
第3章	道路的全方位图像建模	76
3.1	引言	76
3.2	基于全方位图像的特征空间表示方法	78
3.2.1	问题的描述.....	79
3.2.2	特征图像表示.....	80
3.2.3	全方位视图特征空间表示的问题.....	82
3.3	实时全方位成像	83
3.3.1	原理系统几何.....	83
3.3.2	实际成像系统.....	85
3.3.3	系统标定.....	88
3.3.4	地面投影图像.....	88
3.3.5	图像校正.....	90
3.3.6	极坐标图像的路面特征分析.....	91
3.4	全方位图像的旋转不变性特征空间表示	94

3.4.1	沿径向的主分量分析.....	95
3.4.2	沿角度方向的 Fourier 变换	96
3.4.3	旋转角度估计.....	98
3.5	基于全方位图像的神经网络模型	100
3.5.1	问题的描述	100
3.5.2	模型表示	101
3.5.3	具体问题	102
3.6	实验结果和分析	104
3.6.1	实验系统结构	104
3.6.2	数据采集	107
3.6.3	训练和测试	110
3.7	总结和讨论	113
第 4 章	障碍物的双目注视图像建模	115
4.1	引言	115
4.2	平面重投影变换原理	116
4.2.1	原理	117
4.2.2	性质	119
4.2.3	意义	121
4.3	单摄像机双目成像系统	126
4.3.1	左右图像双目成像装置	126
4.3.2	上下图像双目成像装置	132
4.4	实时双目障碍物检测算法	134
4.4.1	目标和基本假设	135
4.4.2	算法原理分析	136
4.4.3	基本算法	142
4.4.4	算法实现及性能分析	143
4.5	动态重投影变换	146
4.5.1	迭代的方法	147
4.5.2	图像稳定的方法	148
4.5.3	广义 HOUGH 变换的方法	149
4.6	总结和讨论	153

第 5 章 多尺度全视野视觉集成	155
5.1 “分解 – 综合”的视觉导航策略	155
5.2 集成视觉传感器 POST	160
5.3 数据融合	162
5.4 导航工作模式的初步讨论	166
结束语	168
附录	175
附录 1 运动模型的扩展	175
附录 2 图像稳定实验分析	179
附录 3 运动遮挡的能量模型的证明	191
附录 4 GFOD 快速算法	192
附录 5 全方位成像几何	193
附录 6 柱面投影分析	194
附录 7 式(4.8)的重投影变换几何	195
附录 8 式(4.9)和(4.10)重投影视差关系的证明 ...	196
附录 9 差值和视差的关系	196
附录 10 障碍物检测性能分析	197
博士论文期间及后续工作中的研究成果和发表的学术论文 ...	205
参考文献	215

第1章 绪论

移动机器人视觉导航和物体识别是三维计算机视觉的两大应用领域,而研究自然环境中的机器人视觉导航对视觉的研究具有重要的意义。从20世纪80年代开始,世界各国对此进行了广泛研究。多年的三维视觉研究取得了不少进展,也提出许多新的思路和框架,如主动视觉、有目的视觉等。但视觉研究还缺少理论基础。在视觉表示方面,缺少描述复杂真实世界景物更好的模型;在视觉算法方面,有待提出更鲁棒的方法,如在深度、运动边界无法使用平滑性约束的情况下解决病态问题的有效方法;在应用方面,也远未出现能像人一样能够在自然环境中应付自如的机器人视觉系统。使计算机(或机器人)视觉系统应用于实际自然环境中,针对所处环境的特点,进行切实有效的景物描述,提出鲁棒性强、行之有效的算法,并将三维计算机视觉推向更广泛的应用,应是当前计算机视觉研究的主要方向。

1.1 问题的提出

理解环境和适应环境的能力应是智能机器人的主要特点。视

觉导航是指机器人通过视觉等感觉器官来认识环境,并根据任务的需要作出决策的过程。环境建模,即对环境的学习、记忆和表示,是视觉导航的基础。视觉环境建模是指通过视感觉,在机器内部记忆环境,并给出适当表示的过程,它包括传感方法、建模算法和环境表示三个部分。换句话说,对环境信息的获取、处理和表示就是环境建模。环境建模不一定是完全的三维景物恢复,它根据具体任务的要求有选择性地建立有关的视觉环境模型,环境建模不只是景物表示,它还包括对景物的感知和学习过程,环境建模可能需要获取三维信息,但并不只是三维恢复。

在更广泛的意义上,视觉导航,不论是自主、半自主或是遥控的,都是为了完成一定的任务,行为者(主体)和环境(客体)进行交互的过程。视觉导航的研究所关心的三个问题是:主体(actor, agent)是人还是机器人?所处的环境(scene, environment)是室内还是室外,太空还是海底?视觉导航的任务(task, purpose)是什么?针对这三个方面对象的不同,视觉系统既具有一般性,又不可避免地具有与任务、行为者和环境有关的特殊性。

在室外自然环境下,机器人导航的基本任务(task)有如下3个:

(1) 基于环境理解的全局定位:通过对环境中景物的理解,识别人造或自然的路标,以完成对机器人的定位。

(2) 道路识别和跟踪:包括对自然环境中道路的理解和可行通路方向的判别。

(3) 目标识别和障碍物检测:包括对路面上障碍物的检测和对特定目标的识别。

视觉导航的工作环境(scene)是人们所处的三维物理世界。随着环境状况的不同,视觉处理的工作内容可能会大不相同。本文研究的导航环境为室外自然道路环境。

在对自主式智能移动机器人的研究出现巨大困难的时候,人们转向了半自主和遥控机器人的研究(如 Weisbin92),导航的主体