



普通高等教育“十一五”国家级规划教材

第1版获全国普通高等学校优秀教材一等奖

清华大学信息科学技术学院教材——信息与通信工程系列

图像工程(下册)

图像理解

(第2版)

章毓晋 编著
Zhang Yujin

清华大学出版社



普通高等教育“十一五”国家级规划教材

TP391.41
634
:3
2005-2007

第1版获全国普通高等学校优秀教材一等奖

清华大学信息科学技术学院教材——信息与通信工程系列

图像工程(下册)

图像理解

(第2版)

Image Engineering (III)
Image Understanding
(Second Edition)



章毓晋 编著
Zhang Yujin

清华大学出版社
北京

内 容 简 介

本书为《图像工程》的下册,主要介绍图像工程的第三层次——图像理解的基本概念、基本原理、典型方法、实用技术以及国际上有关研究的新成果。

本书主要内容归纳在四个单元中。第一个单元(包含第1,2,3,4章)主要介绍图像工程的整体发展状况和图像理解与其他相关学科的联系,基本的视觉感知原理和过程,高维图像采集以及3-D目标表达方法等。这些也为进一步学习后面单元的内容打下了基础。第二个单元(包含第5,6,7,8章)论述景物恢复(重建)的各种典型技术,对应图像理解的较低层次。这里主要涉及立体视觉技术(包括双目和多目),以及由单目图像恢复深度信息的技术(包括立体光度学、从运动求取结构、从阴影恢复形状、从纹理变化确定表面朝向等)。第三个单元(包含第9,10,11,12章)论述场景解释的概念和原理,对应图像理解的较高层次。这里论述知识和表达基础及常用方法、广义匹配的多种技术,以及图像模式识别的基础工具、图像理解理论的内容发展和图像信息系统的概况比较。第四个单元(包含附录A,B,C)分别介绍了三个典型图像理解技术的应用领域:多传感器图像信息融合、人脸和表情识别、基于内容的图像和视频检索。书中还提供了大量例题、思考题和练习题,并对近半数练习题提供了解答或解题思路。

本书可作为信号与信息处理、通信与信息系统、电子与通信工程、模式识别与智能系统、计算机视觉等学科大学本科和研究生专业基础课教材,也可供信息与通信工程、电子科学与技术、计算机科学与技术、测控技术与仪器、机器人自动化、生物医学工程、光学、电子医疗设备研制、遥感、测绘和军事侦察等领域的科技工作者参考。

版权所有,侵权必究。侵权举报电话:010-62782989 13501256678 13801310933

图书在版编目(CIP)数据

图像工程(下册)——图像理解/章毓晋编著. —2版. —北京:清华大学出版社,2007.2
(清华大学信息科学技术学院教材. 信息与通信工程系列)
ISBN 978-7-302-14220-1

I. 图… II. 章… III. 计算机应用—图像处理—高等学校—教材 IV. TP391.4

中国版本图书馆CIP数据核字(2006)第147500号

责任编辑:陈国新 张占奎

责任校对:白 蕾

责任印制:何 芊

出版发行:清华大学出版社 地 址:北京清华大学学研大厦A座

<http://www.tup.com.cn> 邮 编:100084

c-service@tup.tsinghua.edu.cn

社总机:010-62770175 邮购热线:010-62786544

投稿咨询:010-62772015 客户服务:010-62776969

印刷者:北京市昌平环球印刷厂

装订者:三河市新茂装订有限公司

经 销:全国新华书店

开 本:185×230 印 张:29 字 数:595千字

版 次:2007年2月第2版 印 次:2007年2月第1次印刷

印 数:1~4000

定 价:39.00元

本书如存在文字不清、漏印、缺页、倒页、脱页等印装质量问题,请与清华大学出版社出版部联系调换。联系电话:(010)62770177 转 3103 产品编号:014684-01

《清华大学信息科学技术学院教材》

编 委 会

(以姓氏拼音为序)

主 任： 郑大钟

副 主 任： 蔡鸿程 邓丽曼 胡事民 任 勇 章 征
王希勤 王 雄 余志平

编 委： 高文焕 华成英 陆文娟 王诗宓 温冬婵
萧德云 谢世钟 殷人昆 应根裕 郑君里
郑纬民 周立柱 周润德 朱雪龙

秘 书： 王 娜

责任编辑： 马瑛珺 王一玲 陈国新 邹开颜

出版说明

本套教材是针对清华大学信息科学技术学院所属电子工程系、计算机科学与技术系、自动化系、微电子研究所、软件学院的现行本科培养方案和研究生培养计划的课程设置而组织编写的。这些培养方案和培养计划是基于清华大学对研究型大学的定位和对研究型教学的强调,吸纳多年来在教学改革与实践中所取得的成果和形成的共识,历经多届试用和不断修订而形成的。贯穿于其中的“本科教育的通识性、培养模式的宽口径、教学方式的研究型、专业课程的前沿性”的相关思想是我们组编本套教材所力求体现的基本指导原则。

本套教材以本科教材为主并适量包括研究生教材。定位上,属于信息学科大类中各个基本方向的基本理论和前沿技术的一套高等院校教材。层次上,覆盖学院公共基础课程、专业技术基础课程、专业课程、研究生课程。领域上,涉及6个系列14个领域,即学院公共基础课程系列,信息与通信工程系列(含通信、信息处理等领域),微电子光电子系列(含微电子、光电子等领域),计算机科学与技术系列(含计算机科学、计算机网络与安全、计算机应用、软件工程、网络计算等领域),自动化系列(含控制理论与控制工程、模式识别与智能控制、检测与电子技术、系统工程、现代集成制造等领域),实验实践系列。类型上,以文字教材为主并适量包括多媒体教材,以主教材为主并适量包括习题集、教师手册等辅助教材,以基本理论和工程技术教材为主并适量包括实验和实践课程教材。列入这套教材中的著作,大多是清华大学信息科学技术学院所属系所院开设的课程中经过较长教学实践而形成的,既有多年教学经验和教学改革基础上新编著的教材,也有部分已出版教材的更新和修订版本。教材在总体上突出求新与求实的风格,力求反映所属领域的基本理

论和新进展,力求做到学科先进性和教学适用性的统一。

本套教材的主要读者对象为电子科学与技术、信息与通信工程、计算机科学与技术、控制科学与工程、系统科学、电气工程、机械工程、化学与技术工程、核能工程等相关理工专业的大学生和研究生,以及相应领域和部门的科学工作者和工程技术人员。我们希望,这套教材既能为在校大学生和研究生的学习提供内容先进、论述系统和适于教学的教材或参考书,也能为广大科学工作者与工程技术人员的知识更新与继续学习提供适合的和有价值的进修或自学读物。我们同时要感谢使用本系列教材的广大教师、学生和科技工作者的热情支持,并热忱欢迎提出批评和意见。

《清华大学信息科学技术学院教材》编委会

2003年10月

全套书第2版前言

Preface

本套书为《图像工程》第2版,共分3册,分别为《图像工程(上册)——图像处理》、《图像工程(中册)——图像分析》和《图像工程(下册)——图像理解》。整套书全面介绍图像工程的基本概念、基本原理、典型方法、实用技术以及国际上有关研究的新成果。

本套书的第1版也分3册,其中《图像工程(上册)——图像处理和分析》1999年出版,《图像工程(下册)——图像理解和计算机视觉》2000年出版,《图像工程(附册)——教学参考与习题解答》2002年出版。至今各册已一共重印近30次,总计达10多万册。根据在网上查询的结果,已有约上百所学校院所采用该套书作为教材、教学参考书或考研指定书/参考书。从读者向作者发来的上千封电子邮件中,也得知许多人在利用该套书进行自学。

本套书第1版自出版后也得到相关领域研究人员的广泛使用。根据利用“中国期刊网”中的“中国期刊全文数据库”和“万方数据库”中的“万方数字化期刊”所做的检索,可知本套书第1版自出版至今,已被600多篇在国内重要学术期刊上发表的研究文献所引用。又根据利用“中国期刊网”中的“中国优秀博硕士论文全文数据库”所做的检索,可知本套书第1版自出版以来,已被1000多篇博士和硕士论文所引用。

该套书的多次重印和多次被引用,表明作者一直倡导的,为了对各种图像技术进行综合研究、集成应用而建立的整体框架——图像工程,作为一门系统地研究各种图像理论、技术和应用的新的交叉学科得到了广泛的认可。同时,随着研究的深入和技术的发展,编写新版的工作也逐渐提到议事日程上来。

第2版的编写开始于2003年,是年作者在学术休假期间,作为新加

坡南洋理工大学的访问教授,一方面在那里开出一门新课“图像分析”(Advance in Image Analysis),并参与了其他几门课的教学工作;另一方面,静心构思了整套书的新框架,并开始了第2版的编写。为使新版有更多新的内容,在那里收集了10多本相关书籍,还借阅了20多本书,并作了15万字的笔记。这些书籍中,既包括多本典型的图像处理、图像分析和图像理解教材,也包括如成像技术、数字电视、小波变换、二值图像处理、非线性图像处理、三维图像处理、图像分割和组合、纹理分析、形状分析、立体图像分析、体视学、多视几何、图像代数、视觉表达和识别等专著。这为新版的编写打下了坚实的基础。第2版在内容的深度和广度选择上,仍希望读者能通过本书的学习,独立地和全面地了解该领域的基本理论、技术、应用和发展。

第2版的编写仍如第1版那样,在编写的方针上,力求具有理论性、实用性、系统性、实时性;在内容叙述上,力求理论概念严谨,论证简明扼要。考虑到图像技术的飞速发展,第2版既保留了第1版有代表性的经典内容,也削减了一部分近年使用不多的陈旧内容,同时还充实了近年许多最新的研究成果和得到广泛使用的典型技术。这些新内容既参考了许多有关文献,也结合了作者的一些研究工作和成果以及这些年来教学教案。第2版在整个篇幅上比第1版有相当的增加。

与第1版相比,第2版有如下几个特点:

第一,从总体框架上,第2版按图像工程的三个层次分为对应的三册。2004年,作者在清华大学也开出了“图像分析”课程,这样加上原有的两门课程(“图像处理”和“图像理解”),三门课程和三册教材就互相对应起来了。为此,第2版每册书都增加了教学使用建议,以帮助使用者。

第二,从教学的角度出发,以章为单位,对内容进行了均衡;同时对内容较多的技术,划分成基础部分和深入部分,以有利于不同专业背景使用者的选择。

第三,从学习的角度出发,结合教学实践和读者对第1版书的反馈意见,对一些重要和基本的章节概念,增加了解释和示例,以帮助读者理解和自学。

第四,从深入学习和开展科研的角度出发,在每章结尾处增加了各节小结以及为读者扩大视野和进一步研究而提供的一些相关参考文献的介绍,以帮助读者有的放矢地查阅。

第五,从使用方便的角度出发,将习题和对应的解答都集中在同一册书中。

本套书的第1版从1996年开始编写,至今已近10年。这期间,作者与许多读者(包括教师、学生、自学者)有过各种形式的讨论和交流,有的是通过面谈,有的是通过电话或电子邮件。这些讨论和交流使作者获得了许多宝贵的意见和建议,为完善第1版和编写第2版起到了不可或缺的作用。值得指出的是,书中汇集了几年来不少听课学生的贡献,许多例题和练习题是在历届学生作业和课堂讨论的基础上提炼出来的,一些图片还直接由学生帮助制作,在选材上也从学生的反馈中受到许多启发。借此机会一并表示衷心的感谢。

本套书中有相当内容基于作者和他人共同研究的成果,特别是历年的学生边辉、陈挺、陈伟、戴声扬、傅卓、高永英、葛菁华、胡浩基、黄英、黄翔宇、贾波、贾慧星、姜帆、李娟、李乐、李勍、李睿、刘青棣、刘惟锦、刘晓旻、刘忠伟、陆海斌、罗惠韬、罗沅、钱宇飞、秦暄、谭华春、王志明、吴高洪、吴玮、徐丹、徐枫、徐寅、薛景浩、严严、杨劲波、杨翔英、姚玉荣、俞天利、张宁和朱小青,博士后高立志以及进修教师崔京守、石俊生、杨卫平、张贵仓等,第1版和第2版采用的图片除作者自己制作的外,也包括他们在研究工作中收集和实验得到的。本套书应该是多人合作成果的体现。

全书在编写过程中得到陈国新编辑认真的审阅和精心的修改,这既使全书增色不少,也使作者受益匪浅。

最后,作者感谢妻子何芸,女儿章荷铭以及父母家人在各方面的理解和支持!

章毓晋

2006年夏于清华园

通信:北京清华大学电子工程系,100084

电话:(010) 62781430

传真:(010) 62770317

电子邮件:zhangyj@ee.tsinghua.edu.cn

个人主页:www.ee.tsinghua.edu.cn/~zhangyujin/

研究室网:image.ee.tsinghua.edu.cn

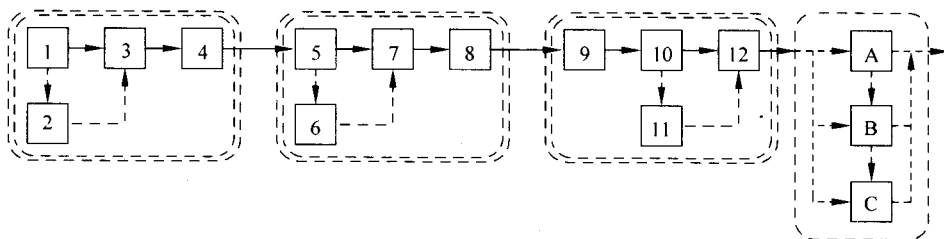
本册书概况和使用建议

本册书为《图像工程》的下册,主要介绍图像工程的第三层次——图像理解的基本概念、基本原理、典型方法、实用技术以及国际上有关研究的新成果。

本册书主要内容可划分在四个单元中。第一个单元(包含第 1,2,3,4 章)主要介绍图像工程的整体发展状况和图像理解与其他相关学科的联系,基本的视觉感知原理和过程,高维图像采集以及 3-D 目标表达方法等。这些也为进一步学习后面单元的内容打下了基础。第二个单元(包含第 5,6,7,8 章)论述景物恢复(重建)的各种典型技术,对应图像理解的较低层次。这里主要涉及立体视觉技术(包括双目和多目),以及由单目图像恢复深度信息的技术(包括立体光度学、从运动求取结构、从阴影恢复形状、从纹理变化确定表面朝向等)。第三个单元(包含第 9,10,11,12 章)论述场景解释的概念和原理,对应图像理解的较高层次。这里论述知识和表达基础及常用方法、广义匹配的多种技术,以及图像模式识别的基础工具、图像理解理论的内容发展和图像信息系统的概况比较。第四个单元(包含附录 A,B,C)分别介绍了三个典型图像理解技术的应用领域:多传感器图像信息融合、人脸和表情识别、基于内容的图像和视频检索。

本册书共有 12 章正文、3 个附录以及习题解答和参考文献。在这 17 个一级标题下共有 71 个二级标题(节),再往下还有 130 个三级标题(小节)。全册书共有文字(也包括图片、绘图、表格、公式等)约 60 万字。本册书共有编了号的图 298 个、表格 34 个、公式 654 个。为便于教学和理解,本册书共给出各类例题 73 个。为便于检查教学和学习效果,各章后均有思考题和练习题,本册书共有思考题和练习题 204 个,对其中的 94 个练习题(题前标“*”号)提供了参考答案。另外,书末列出了直接引用和提供参考文献的 300 多篇文献的目录。

关于本册书的使用可借助下图解释。



本册书主要包括 12 章,根据内容可结合成为三个单元,每个单元包括的章如三个双虚线框所示。每个单元选了三章作为基本内容,即第 1 章、第 3 章、第 4 章,第 5 章、第 7

章、第8章、第9章、第10章、第12章(可按此次序学习)。在包括了这9章基本内容的基础上,可考虑根据专业和学时选择其他内容。如果需要了解人类视觉感知方面的内容,可在第1章后学习第2章。第6章是第5章的扩展,如果有较好的数学基础且比较关注立体视觉的技术,可在第5章后学习。第11章涉及模式识别的内容,如没有学习过相关的课程可在第10章后学习。最后本册书还包括三个附录(在图中单虚线框内),均为涉及有关图像理解(图像工程高层)技术的应用内容,可根据专业和科研方向在第12章后选择学习。

本册书各章内容基本平衡,根据学生的基础和背景,每章可用2~4课堂学时。各章后面均有较多的思考题和练习题(包括若干涉及编程实现算法、编写例题或习题、进行开放式讨论等),每章平均可安排1~3课外学时来完成,有些可作为专题讨论、大作业或项目训练。

本册书主要介绍图像理解的内容,相比图像处理和图像分析更专业化一些,需在学习完本套书上册(或至少学习了其中的第一个单元和第二个单元)后再学习(特别是对于自学的读者)。如果已经学习了本套书中册(或至少其中的第二个单元和第三个单元的基本内容),则对学习本册书更为有利。

目 录

Contents

1 绪论	1
1.1 图像工程的发展	1
1.2 图像理解概述	6
1.2.1 图像理解	6
1.2.2 计算机视觉	7
1.2.3 其他相关学科	9
1.2.4 图像理解的应用领域	10
1.3 主要内容和安排	11
总结和复习	14
2 视感觉和视知觉	16
2.1 从感觉到知觉	16
2.2 视觉特性	19
2.2.1 视觉的时间特性	19
2.2.2 视觉的空间特性	21
2.3 形状知觉	23
2.3.1 形状的感知	23
2.3.2 轮廓	24
2.3.3 图形和背景	26
2.3.4 几何图形视错觉	30
2.4 空间知觉	34
2.4.1 非视觉性深度线索	35
2.4.2 双目深度线索	35

2.4.3	单目深度线索	38
2.5	运动知觉	40
	总结和复习	42
3	高维图像采集	44
3.1	高维图像	44
3.1.1	高维图像种类	44
3.1.2	本征图像和非本征图像	45
3.2	成像变换和摄像机模型	47
3.2.1	一般摄像机模型	47
3.2.2	近似投影模式	50
3.2.3	通用成像模型	52
3.3	摄像机标定	54
3.3.1	标定程序和参数	54
3.3.2	两级标定法	57
3.4	深度图像采集	60
3.4.1	飞行时间法	60
3.4.2	结构光法	63
3.4.3	莫尔等高条纹法	64
3.4.4	深度和亮度图像同时采集	67
3.5	显微镜 3-D 分层成像	68
3.5.1	显微镜 3-D 成像	68
3.5.2	共聚焦显微镜 3-D 成像	70
	总结和复习	72
4	3-D 目标表达	75
4.1	曲线和曲面的局部特征	76
4.1.1	曲线局部特征	76
4.1.2	曲面局部特征	79
4.2	3-D 表面表达	84
4.2.1	参数表达	84
4.2.2	表面朝向表达	86
4.3	等值面的构造和表达	88
4.3.1	行进立方体算法	88

4.3.2 覆盖算法	91
4.4 从并行轮廓插值 3-D 表面	92
4.5 3-D 实体表达	97
4.5.1 基本表达方案	97
4.5.2 广义圆柱体表达	100
总结和复习	101
5 立体视觉：双目	103
5.1 立体视觉	104
5.1.1 立体成像方式	104
5.1.2 立体视觉模块	105
5.2 双目成像和视差	107
5.2.1 双目横向模式	107
5.2.2 双目横向会聚模式	110
5.2.3 双目纵向模式	111
5.3 基于区域的双目立体匹配	112
5.3.1 模板匹配	112
5.3.2 双目立体匹配	114
5.4 基于特征的双目立体匹配	117
5.4.1 基本方法	117
5.4.2 动态规划匹配	119
5.5 视差图误差检测与校正	121
总结和复习	124
6 立体视觉：多目	127
6.1 水平多目立体匹配	127
6.1.1 水平多目图像	128
6.1.2 倒距离	129
6.2 正交三目立体匹配	132
6.2.1 基本原理	132
6.2.2 基于梯度分类的正交匹配	136
6.3 多目立体匹配	140
6.3.1 任意排列三目立体匹配	140
6.3.2 正交多目立体匹配	142

6.4	亚像素级视差计算	142
	总结和复习	147
7	景物恢复：多图像	149
7.1	单目景物恢复	149
7.2	光度立体学	151
7.2.1	景物亮度和图像亮度	151
7.2.2	表面反射特性和亮度	154
7.2.3	目标表面朝向	156
7.2.4	反射图和亮度约束方程	157
7.2.5	光度立体学求解	159
7.3	从运动求取结构	162
7.3.1	光流和运动场	162
7.3.2	光流方程求解	164
7.3.3	光流与表面取向	170
	总结和复习	172
8	景物恢复：单图像	174
8.1	从阴影恢复形状	174
8.1.1	阴影与形状	174
8.1.2	利用单目图像求解照度方程	178
8.2	纹理与表面朝向	183
8.2.1	单目成像和畸变	183
8.2.2	由纹理变化恢复朝向	185
8.2.3	线段纹理消失点的确定	192
8.3	由焦距确定深度	194
8.4	根据三点透视估计位姿	196
	总结和复习	198
9	知识和表达	201
9.1	知识分类和表达	202
9.2	场景知识	203
9.2.1	模型	203
9.2.2	属性超图	205

9.2.3 基于知识的建模	206
9.3 过程知识	208
9.4 知识表达基础	210
9.4.1 对知识表达的要求	210
9.4.2 知识表达类型	211
9.4.3 图像理解系统中的知识模块	212
9.4.4 图像理解中的知识表达	213
9.5 逻辑系统	215
9.5.1 谓词演算规则	215
9.5.2 利用定理证明来推理	218
9.6 语义网络	221
9.7 产生式系统	224
总结和复习	227
10 广义匹配	229
10.1 匹配基础	230
10.1.1 匹配策略和类别	230
10.1.2 匹配和配准	231
10.2 目标匹配	232
10.2.1 匹配的度量	233
10.2.2 字符串匹配	235
10.2.3 惯量等效椭圆匹配	235
10.3 动态模式匹配	236
10.4 关系匹配	239
10.5 图同构	243
10.5.1 图论简介	243
10.5.2 图同构和匹配	245
10.6 线条图标记	247
总结和复习	251
11 图像模式识别	254
11.1 模式和分类	254
11.2 统计模式识别	258
11.2.1 最小距离分类器	258

11.2.2	最优统计分类器	260
11.3	感知机和支持向量机	264
11.3.1	感知机	264
11.3.2	支持向量机	268
11.4	结构模式识别	272
11.4.1	字符串结构识别	272
11.4.2	树结构识别	275
11.4.3	学习和推理	278
	总结和复习	281
12	图像理解理论和系统	284
12.1	从感知到理解	285
12.2	图像理解理论框架	287
12.2.1	马尔视觉计算理论	287
12.2.2	对马尔理论框架的改进	293
12.2.3	关于马尔重建理论的讨论	294
12.2.4	新理论框架的研究	296
12.3	图像理解系统模型	298
12.3.1	系统模型结构	298
12.3.2	多层次串行结构	299
12.3.3	以知识库为中心的辐射结构	300
12.3.4	以知识库为根的树结构	301
12.3.5	多模块交叉配合结构	302
12.4	具体系统分析	303
12.4.1	VISIONS 系统	303
12.4.2	ACRONYM 系统	304
12.4.3	KB Vision 系统	306
12.5	典型系统比较	306
12.6	讨论和展望	308
	总结和复习	312
附录 A	多传感器图像信息融合	315
A.1	信息融合概述	315
A.1.1	多信息融合	316