



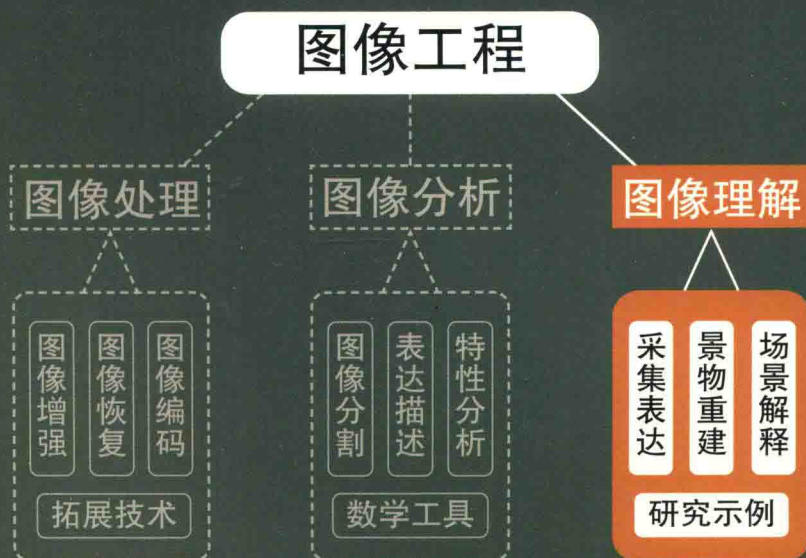
“十二五”普通高等教育本科国家级规划教材

第1版获全国普通高等学校优秀教材一等奖
第2版评为全国普通高等教育“十一五”精品教材
第3版评为2013年北京高等教育精品教材

图像工程（下册） 图像理解 （第4版）

章毓晋 (ZHANG Yu-Jin) 编著

IMAGE ENGINEERING (III)
IMAGE UNDERSTANDING
(Fourth Edition)



清华大学出版社





“十二五”普通高等教育本科国家级规划教材

图像工程 (下册)

图像理解

(第4版)

章毓晋 (ZHANG Yu-Jin) 编著

IMAGE ENGINEERING (III)

IMAGE UNDERSTANDING

(Fourth Edition)

清华大学出版社

北京

内 容 简 介

本书为《图像工程》第4版的下册,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。

本书主要分为4个单元。第1单元(包含第2~5章)介绍图像采集表达技术,其中第2章介绍摄像机成像模型和标定技术,第3章介绍压缩感知理论及其在成像中的应用,第4章介绍采集含深度信息图像的方法,第5章介绍各种表达3-D景物的技术。第2单元(包含第6~9章)介绍景物重建技术,其中第6章介绍双目立体视觉方法,第7章介绍多目立体视觉方法,第8章介绍从多幅图像恢复景物的技术,第9章介绍从单幅图像恢复景物的技术。第3单元(包含第10~12章)介绍场景解释技术,其中第10章介绍知识表达和推理方法,第11章介绍目标和符号匹配技术,第12章介绍场景分析和语义解释的内容。第4单元(包含第13~15章)介绍三个研究方向的示例,其中第13章介绍多传感器图像信息融合方法,第14章介绍基于内容的图像和视频检索技术,第15章介绍时空行为理解的内容。书中的附录介绍了有关视觉和视知觉的一些知识,与各章都有一些联系。书中还提供大量例题、思考题和练习题,并对部分练习题提供了解答。书末还给出了主题索引。

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、计算机视觉等学科研究生专业基础或专业课教材,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。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

图像工程.下册,图像理解/章毓晋编著. —4版. —北京:清华大学出版社,2018
ISBN 978-7-302-50361-3

I. ①图… II. ①章… III. ①计算机应用—图像处理 IV. ①TP391.41

中国版本图书馆CIP数据核字(2018)第117718号

责任编辑:文怡

封面设计:李召霞

责任校对:时翠兰

责任印制:沈露

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦A座

邮 编:100084

社总机:010-62770175

邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

课件下载: <http://www.tup.com.cn>, 010-62795954

印装者:三河市国英印务有限公司

经 销:全国新华书店

开 本:185mm×260mm 印 张:28.5

字 数:692千字

版 次:2000年8月第1版 2018年8月第4版

印 次:2018年8月第1次印刷

定 价:89.00元

产品编号:077216-01

全套书第4版前言

这是《图像工程》第4版,全套书仍分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》。它们全面介绍图像工程的基础概念、基本原理、典型方法、实用技术以及国际上相关内容研究的新成果。

《图像工程》第3版也分3册,名称相同。上、中、下册均在2012年出版,而2013年出版了《图像工程》第3版的3册合订本。第3版至今已重印13次,总计3万多册。

《图像工程》第2版也分3册,名称相同。上、中、下册分别在2006年、2005年和2007年出版,2007年还出版了《图像工程》第2版的3册合订本。第2版共重印18次,总计近7万册。

《图像工程》第1版也分3册,名称分别为《图像工程(上册)——图像处理和分析》《图像工程(下册)——图像理解和计算机视觉》和《图像工程(附册)——教学参考及习题解答》。这3册分别在1999年、2000年和2002年出版。第1版共重印27次,总计约11万册。

《图像工程》的多次重印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,也在教学中得到大量使用。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4版的编写开始于2016年,是年暑假静心构思了全套书的整体框架。其后,根据框架陆续收集了一些最新的相关书籍和文献(包括印刷版和电子版),并仔细进行了阅读和做了笔记。这为新版的编写打下了一个坚实的基础。期间,还结合以往课堂教学和学生反馈,对一些具体内容(包括习题)进行了整理和调整。第4版内容具有一定的深度和广度,希望读者通过本套书的学习,能够独立和全面地了解该领域的基本理论、技术、应用和发展。

第4版在编写的方针上,仍如前3版那样力求具有理论性、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。在内容方面,第4版基本保留了第3版中有代表性的经典内容,同时考虑图像技术的飞速发展,还认真选取了近年的一些最新研究成果和得到广泛使用的典型技术进行充实。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的教学教案。除每册书增加一章全新内容外,还各增加了多个节和小节,并特别增加了许多例题,其中有些介绍新的选学内容,有些则从其他角度来补充解释已有的概念和方法。这些例题可根据课时安排、学生基础等选择使用,比较灵活。总体来说,第4版的内容覆盖面更广,介绍更全面细致,整体篇幅比第3版有约20%的增加。

第4版在具体结构和章节安排方面仍然保留了第3版的特点:

第一,各册书均从第2章就开始介绍正式内容,更快进入主题。先修或预备内容分别安排在需要先修部分的同一章前部,从教学角度来说,更加实用,也突出了主线内容。

第二,除第1章绪论外,各册书的正式内容仍都结合成4个主题相关的单元(并画在封

面上),每个单元都有具体说明,帮助选择学习。全书有较强的系统性和结构性,也有利于复习考核。

第三,各章中的习题均只有少部分给出了解答,使教师可以更灵活地选择布置。更多的习题和其余的习题解答将会放在出版社网站上,便于补充、改进,网址为 www.tup.com.cn。

第四,各册书后均仍有主题索引(并给出了英文),这样既方便在书中查找有关内容,又方便在网上查找有关文献和解释。

第4版还增加了一项新的举措。书中的彩色图片印刷后均为黑白的,但可以通过手机扫描图片旁的二维码,调出存放在出版社网站上的对应彩色图片,获得更多的信息和更好的观察效果。

从1996年开始编写《图像工程》第1版以来至今已20多年。期间,作者与许多读者(包括教师、学生、自学者等)有过各种形式的讨论和交流,除了与一些同行面谈外,许多人打来电话或发来电子邮件。这些讨论和交流使作者获得了许多宝贵的意见和建议,在编写这4版中都起到了不可或缺的作用,特别是在解释和描述的详略方面都结合读者反馈意见进行了调整,从而更加容易理解和学习。值得指出的是,书中还汇集了多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。借此机会对他们一并表示衷心的感谢。

书中有相当内容基于作者和他人共同研究的成果,特别是历年研究室的学生(按姓名拼音排序):卜莎莎、边辉、蔡伟、陈权崎、陈挺、陈伟、陈正华、崔崑、程正东、戴声扬、段菲、方慕园、冯上平、傅卓、高永英、葛菁华、侯乐天、胡浩基、黄英、黄翔宇、黄小明、贾波、贾超、贾慧星、姜帆、李佳童、李娟、李乐、李品一、李勍、李睿、李硕、李闻天、李相贤(LEE Sang Hyun)、李小鹏、李雪、梁含悦、刘宝弟、刘晨阳、刘峰、刘锴、刘青棣、刘惟锦、刘晓旻、刘忠伟、陆海斌、陆志云、罗惠韬、罗沅、朴寅奎(PARK In Kyu)、钱宇飞、秦暄、秦垠峰、阮孟贵(NGUYEN Manh Quy)、赛义(Saeid BAGHERI)、沈斌、谭华春、汤达、王树徽、王宇雄、王志国、王志明、王钟绪、温宇豪、文熙安(Tristan VINCENT)、吴高洪、吴纬、夏尔雷(Charley PAULUS)、向振、徐丹、徐枫、徐洁、徐培、徐寅、许翔宇、薛菲、薛景浩、严严、杨劲波、杨翔英、杨忠良、姚玉荣、游钱皓喆、鱼荣珍(EY Young Jin)、俞天利、于信男、袁静、负亮、张宁、赵雪梅、郑胤、周丹、朱施展、朱小青、朱云峰、博士后高立志、王怀颖以及进修教师崔京守(CHOI Jeong Swu)、郭红伟、石俊生、杨卫平、曾萍萍、张贵仓等。第1版、第2版、第3版和第4版采用的图表除作者本人制作的外,也包括他们在研究工作中收集和实验得到的。该书应该说是多人合作成果的体现。

最后,感谢妻子何芸、女儿章荷铭在各方面的理解和支持!

章毓晋

2018年元旦于书房

通信:北京,清华大学电子工程系,100084

办公:清华大学,罗姆楼,6层305室

电话:(010) 62798540

传真:(010) 62770317

电邮:zhang-yj@tsinghua.edu.cn

主页:oa.ee.tsinghua.edu.cn/~zhangyujin/

下册书概况和使用建议

本书为《图像工程》第4版的下册,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。

本书第1章是绪论,介绍图像理解基础并概述全书。本书主要内容分别在4个单元中介绍。第1单元(包含第2~5章)介绍图像采集表达技术,其中第2章介绍摄像机成像模型和标定技术,第3章介绍压缩感知理论及其在成像中的应用,第4章介绍采集含深度信息图像的方法,第5章介绍各种表达3-D景物的技术。第2单元(包含第6~9章)介绍景物重建技术,其中第6章介绍双目立体视觉方法,第7章介绍多目立体视觉方法,第8章介绍从多幅图像恢复景物的技术,第9章介绍从单幅图像恢复景物的技术。第3单元(包含第10~12章)介绍场景解释技术,其中第10章介绍知识表达和推理方法,第11章介绍目标和符号匹配技术,第12章介绍场景分析和语义解释的内容。第4单元(包含第13~15章)介绍三个研究方向的示例,其中第13章介绍多传感器图像信息融合方法,第14章介绍基于内容的图像和视频检索技术,第15章介绍时空行为理解的内容。书中的附录为有关视觉和视知觉的一些知识,与各章都有一些联系。

本书包括15章正文,1个附录,以及“部分习题解答”“参考文献”和“索引”。在这19个一级标题下共有91个二级标题(节),再下还有144个三级标题(小节)。全书共有文字(也包括图片、绘图、表格、公式等)共60多万字。本书共有编了号的图349个(包括297幅图片)、表格37个、公式775个。为便于教学和理解,本书共给出各类例题99个。为便于检查教学和学习效果,各章后均有12个思考题和练习题,全书共有180个,对其中的30个(每章2个)提供了参考答案(更多思考题和练习题的解答将考虑另行提供)。另外,参考文献中统一列出了直接引用和提供参考文献的370多篇文献的目录。最后,书末还给出了700多个主题索引(及英译)。

本书各章主要内容和讲授长度基本平衡,根据学生的基础和背景,每章可用3~4个课堂学时讲授,另外可能还需平均2~3个课外学时练习和复习。本书电子教案可在清华大学出版社网站(<http://www.tup.com.cn>)或作者主页(<http://oa.ee.tsinghua.edu.cn/~zhangyujin>)下载。

本书主要介绍图像理解的内容,需要有一定的图像处理和图像分析基础,最好作为学习图像技术的第三本书来学习(如果自学,最好在学习完上册的前两个单元和中册的前两个单元后进行)。当学时比较紧张时,可先学习完本书的前两个单元,再从后两个单元中选取部分内容。

目 录

第 1 章 绪论	1
1.1 图像工程的发展	1
1.1.1 基本概念和定义概括	1
1.1.2 图像技术发展情况回顾	3
1.2 图像理解及相关学科	7
1.2.1 图像理解	7
1.2.2 计算机视觉	8
1.2.3 其他相关学科	10
1.2.4 图像理解的应用领域	11
1.3 图像理解理论框架	12
1.3.1 马尔视觉计算理论	12
1.3.2 对马尔理论框架的改进	17
1.3.3 关于马尔重建理论的讨论	18
1.3.4 新理论框架的研究	20
1.4 内容框架和特点	22
总结和复习	25

第 1 单元 采集表达

第 2 章 摄像机成像	29
2.1 视觉过程	29
2.2 摄像机成像模型	31
2.2.1 基本摄像机模型	32
2.2.2 近似投影模式	36
2.2.3 一般摄像机模型	39
2.2.4 通用成像模型	41
2.3 摄像机标定	43
2.3.1 标定程序和参数	43
2.3.2 两级标定法	46
2.4 亮度成像	48
2.4.1 光度学和光源	49

2.4.2 从亮度到照度	50
总结和复习	53
第3章 压缩感知与成像	56
3.1 压缩感知概述	56
3.2 稀疏表达	58
3.3 测量矩阵及特性	60
3.3.1 采样/测量模型	60
3.3.2 测量矩阵特性	61
3.4 解码重构	63
3.4.1 重构原理	63
3.4.2 测量矩阵的校准	65
3.4.3 典型重构算法	66
3.5 稀疏编码与字典学习	71
3.5.1 字典学习与矩阵分解	72
3.5.2 非负矩阵分解	74
3.5.3 端元提取	75
3.5.4 稀疏编码	76
3.6 压缩感知的成像应用	78
3.6.1 单像素相机	78
3.6.2 压缩感知磁共振成像	79
总结和复习	80
第4章 深度信息采集	83
4.1 高维图像和成像方式	83
4.1.1 高维图像种类	83
4.1.2 本征图像和非本征图像	84
4.1.3 深度成像方式	85
4.2 双目成像模式	86
4.2.1 双目横向模式	86
4.2.2 双目会聚横向模式	90
4.2.3 双目轴向模式	92
4.3 深度图像直接采集	94
4.3.1 飞行时间法	94
4.3.2 结构光法	96
4.3.3 莫尔等高条纹法	98
4.3.4 深度和亮度图像同时采集	100
4.4 显微镜 3-D 分层成像	101
4.4.1 景深和焦距	101
4.4.2 显微镜 3-D 成像	102
4.4.3 共聚焦显微镜 3-D 成像	104

总结和复习	106
第 5 章 3-D 景物表达	108
5.1 曲线和曲面的局部特征	108
5.1.1 曲线局部特征	108
5.1.2 曲面局部特征	111
5.2 3-D 表面表达	114
5.2.1 参数表达	114
5.2.2 表面朝向表达	115
5.3 等值面的构造和表达	117
5.3.1 行进立方体算法	118
5.3.2 覆盖算法	120
5.4 从并行轮廓插值 3-D 表面	121
5.5 3-D 实体表达	126
5.5.1 基本表达方案	126
5.5.2 广义圆柱体表达	128
总结和复习	129

第 2 单元 景物重建

第 6 章 立体视觉：双目	133
6.1 立体视觉模块	133
6.2 基于区域的双目立体匹配	135
6.2.1 模板匹配	135
6.2.2 立体匹配	138
6.3 基于特征的双目立体匹配	143
6.3.1 基本步骤	144
6.3.2 尺度不变特征变换	146
6.3.3 加速鲁棒性特征	148
6.3.4 动态规划匹配	153
6.4 视差图误差检测与校正	154
总结和复习	157
第 7 章 立体视觉：多目	159
7.1 水平多目立体匹配	159
7.1.1 水平多目图像	159
7.1.2 倒距离	160
7.2 正交三目立体匹配	163
7.2.1 基本原理	163
7.2.2 基于梯度分类的正交匹配	167
7.3 多目立体匹配	170

7.3.1	任意排列三目立体匹配	170
7.3.2	正交多目立体匹配	173
7.4	亚像素级视差计算	175
	总结和复习	178
第8章	景物恢复：多图像	181
8.1	单目景物恢复	181
8.2	光度立体学	182
8.2.1	景物亮度和图像亮度	182
8.2.2	表面反射特性和亮度	185
8.2.3	景物表面朝向	186
8.2.4	反射图和亮度约束方程	187
8.2.5	光度立体学求解	190
8.3	从运动求取结构	192
8.3.1	光流和运动场	193
8.3.2	光流方程求解	195
8.3.3	光流与表面取向	199
8.3.4	光流与相对深度	201
	总结和复习	202
第9章	景物恢复：单图像	204
9.1	从影调恢复形状	204
9.1.1	影调与形状	204
9.1.2	亮度方程求解	207
9.2	纹理与表面朝向	212
9.2.1	单目成像和畸变	212
9.2.2	由纹理变化恢复朝向	214
9.2.3	检测线段纹理消失点	219
9.2.4	确定图像外消失点	220
9.3	由焦距确定深度	222
9.4	根据三点透视估计位姿	224
	总结和复习	226

第3单元 场景解释

第10章	知识表达和推理	231
10.1	知识基础	231
10.2	场景知识	232
10.2.1	模型	233
10.2.2	属性超图	234
10.2.3	基于知识的建模	235

10.3	过程知识	237
10.4	知识表达	238
10.4.1	知识表达要求	239
10.4.2	知识表达类型	239
10.4.3	图像理解系统中的知识模块	240
10.4.4	基本知识表达方案	241
10.5	逻辑系统	243
10.5.1	谓词演算规则	243
10.5.2	利用定理证明来推理	245
10.6	语义网	248
10.7	产生式系统	251
	总结和复习	253
第 11 章	广义匹配	255
11.1	匹配概述	255
11.1.1	匹配策略和类别	256
11.1.2	匹配和配准	257
11.1.3	匹配评价	258
11.2	目标匹配	259
11.2.1	匹配的度量	259
11.2.2	对应点匹配	261
11.2.3	字符串匹配	262
11.2.4	惯量等效椭圆匹配	262
11.2.5	形状矩阵匹配	263
11.3	动态模式匹配	264
11.4	关系匹配	266
11.5	图同构匹配	269
11.5.1	图论简介	269
11.5.2	图同构和匹配	271
11.6	线条图标记和匹配	273
	总结和复习	277
第 12 章	场景分析和语义解释	279
12.1	场景理解概述	279
12.2	模糊推理	281
12.2.1	模糊集和模糊运算	281
12.2.2	模糊推理方法	283
12.3	遗传算法图像解释	285
12.3.1	遗传算法原理	285
12.3.2	语义分割和解释	286
12.4	场景目标标记	289

12.5 场景分类	292
12.5.1 词袋/特征包模型	292
12.5.2 pLSA 模型	294
12.5.3 LDA 模型	298
总结和复习	300

第4单元 研究示例

第13章 多传感器图像信息融合	305
13.1 信息融合概述	305
13.2 图像融合	307
13.2.1 图像融合的主要步骤	307
13.2.2 图像融合的三个层次	309
13.2.3 图像融合效果评价	310
13.3 像素级融合方法	313
13.3.1 基本融合方法	313
13.3.2 融合方法的结合	315
13.3.3 小波融合时的最佳分解层数	318
13.3.4 压缩感知图像融合	319
13.3.5 像素级融合示例	319
13.4 特征级和决策级融合方法	322
13.4.1 贝叶斯法	322
13.4.2 证据推理法	324
13.4.3 粗糙集理论法	326
总结和复习	329
第14章 基于内容的图像和视频检索	331
14.1 图像和视频检索原理	331
14.2 视觉特征的匹配和检索	335
14.2.1 颜色特征匹配	335
14.2.2 纹理特征计算	336
14.2.3 多尺度形状特征	337
14.2.4 综合特征检索	338
14.3 基于运动特征的视频检索	340
14.3.1 全局运动特征	340
14.3.2 局部运动特征	341
14.4 视频节目分析和索引	342
14.4.1 新闻视频结构化	343
14.4.2 体育比赛视频排序	345
14.4.3 家庭录像视频组织	350

14.5	语义分类检索	354
14.5.1	基于视觉关键词的图像分类	354
14.5.2	高层语义与气氛	355
	总结和复习	357
第 15 章	时空行为理解	360
15.1	时空技术	360
15.2	时空兴趣点	362
15.3	动态轨迹学习和分析	364
15.3.1	自动场景建模	364
15.3.2	学习路径	366
15.3.3	自动活动分析	368
15.4	动作分类和识别	370
15.4.1	动作分类	370
15.4.2	动作识别	371
15.5	活动和行为建模	375
15.5.1	动作建模	375
15.5.2	活动建模和识别	379
15.6	主体与动作联合建模	382
15.6.1	单标签主体-动作识别	383
15.6.2	多标签主体-动作识别	383
15.6.3	主体-动作语义分割	384
	总结和复习	387
附录A	视觉和视知觉	389
A.1	视知觉概述	389
A.2	视觉特性	391
A.2.1	视觉的空间特性	391
A.2.2	视觉的时间特性	393
A.2.3	视觉的亮度特性	394
A.3	形状知觉	396
A.3.1	轮廓	396
A.3.2	图形和背景	397
A.3.3	几何图形错觉	401
A.4	空间知觉	404
A.4.1	非视觉性深度线索	404
A.4.2	双目深度线索	405
A.4.3	单目深度线索	406
A.5	运动知觉	408
	部分思考题和练习题解答	413
	参考文献	418
	主题索引	433

第1章 绪 论

本书为《图像工程》整套书的下册,是提高和总结的一册。

本章对全书内容要点和布局结构进行概括介绍,安排如下:

1.1 节先对图像工程的发展情况加以回顾和概述,并列举两个相关文献综述的一些统计数据,展示图像工程的发展情况。

1.2 节概括介绍图像理解的研究内容和在图像工程中的位置,讨论图像理解与计算机视觉,以及与一些相关学科的联系和区别,并介绍图像理解的一些应用领域。

1.3 节全面介绍对图像理解有重要基础作用的马尔视觉计算理论的各个要点以及对马尔理论框架的改进工作,并对马尔重建理论的不足进行了一些讨论,既帮助总体把握整个领域,也促进开展进一步的深入研究。

1.4 节概括介绍本书主要内容、框架结构、编写特点以及先修知识要求。

1.1 图像工程的发展

先对图像工程的发展情况给一个概述(更多细节可参见上册和中册)。

1.1.1 基本概念和定义概括

图像是用各种观测系统以不同形式和手段观测客观世界而获得的,可以直接或间接作用于人眼,并进而产生视知觉的实体[章 1996a]。人的视觉系统是具有这样能力的典型系统。视觉是人类观察世界、认知世界的重要功能手段。人类从外界获得的信息约有 75% 来自视觉系统,这既说明视觉信息量巨大,也表明人类对视觉信息有较高的利用率。正因为如此,人们还制造了许多利用各种辐射对客观场景成像的系统,以利用视觉信息观察世界。图像是表达视觉信息的一种物理形式。对场景采集数字图像的最终结果常是某种能量的样本阵列,所以常用矩阵或数组来表示,其中每个元素的坐标对应场景点的位置,而元素的值对应场景点的某个物理量。

人们用各种技术方式和手段对图像进行加工,以获得需要的信息。从广义的角度,图像技术可看作各种图像加工技术的总称。它覆盖了利用计算机和其他电子设备进行和完成的一系列工作,例如图像的采集、获取、编码、存储和传输,图像的合成和产生,图像的显示、绘制和输出,图像的变换、增强、恢复(复原)、修补和重建,图像的分割,各种特征的提取和测量,目标的检测、定位、表达和描述,序列图像的校正和配准,图像数据库的建立、索引、查询和抽取,图像的分类、表示和识别,3-D 景物的重建复原,图像模型的建立,图像信息的融合,图像知识的利用和匹配,图像和场景的解释和理解,以及基于它们的推理、学习、判断、决策

和行为规划(如何推断出应实现的目标及构造实现目标的操作序列)等。另外,图像技术还可包括为完成上述功能而进行的硬件和系统设计及制作等方面的技术。上述许多具体技术已在本套书上册和中册进行了介绍。

对图像技术的综合研究和集成应用可在**图像工程**这个整体框架下进行[章 1996a]。众所周知,工程是指将自然科学的原理应用到工业部门而形成的各学科的总称。图像工程学科则是利用数学、光学等基础科学的原理,结合电子技术、计算机技术及在图像应用中积累的技术经验而发展起来的一个对整个图像领域进行研究应用的新学科。事实上,图像技术多年来的发展和积累为图像工程学科的建立打下了坚实的基础,而各类图像应用也对图像工程学科的建立提出了迫切的需要([章 2000c]、[章 2002c]、[章 2009c]、[Zhang 2009f]、[Zhang 2015f]、[Zhang 2018a])。

图像工程的内容非常丰富,应用也非常广泛,根据抽象程度、研究方法、操作对象和数据量等的不同可分为 3 个层次,见图 1.1.1: 图像处理、图像分析和图像理解(可分别见 [Zhang 2017a]、[Zhang 2017b]、[Zhang 2017c])。图像处理是比较低层的操作,主要在图像像素级上进行处理,处理的数据量非常大。图像分析则处于中层,通过分割和特征提取把原来以像素描述的图像转变成比较简洁的非图形式的描述。图像理解主要指高层的操作,基本上根据较抽象的描述进行解析、判断、决策(符号运算),其处理过程和方法与人类的思维推理有许多类似之处。在这里,随着抽象程度的提高,数据量是逐渐减少的。具体说来,原始图像数据经过一系列的处理过程,逐步转化为更有组织和用途的信息。在这个过程中,语义信息不断引入,操作对象也逐步发生变化。另外,高层操作对低层操作有指导作用,能提高低层操作的效能,完成复杂的任务。

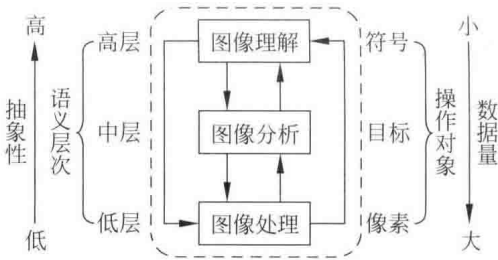


图 1.1.1 图像工程 3 层次示意图

概括地说,图像工程是既有联系又有区别的图像处理、图像分析及图像理解三者的有机结合,并包括对它们的工程应用。从概念上讲,图像工程既能较好地许多相近学科兼容并蓄,也更强调图像技术的应用,所以这里选用图像工程来概括整个图像领域的研究和应用,也使图像处理、图像分析及图像理解三者的关系更紧密。

图像工程是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。从研究方法来看,它与数学、物理学、生物学、生理学(特别是神经生理学)、心理学、电子学、计算机科学等许多学科可以相互借鉴;从研究范围来看,它与模式识别、计算机视觉、计算机图形学等多个专业又相互交叉。另外,图像工程的研究进展与人工智能、神经网络,遗传算法、模糊逻辑等理论和技术都有密切的联系,它的发展应用与生物医学、材料、遥感、通信、交通管理、军事侦察、文档处理和工业自动化等许多领域也是不可分割的。

图像工程是全面系统地研究图像理论方法,阐述图像技术原理,推广图像技术应用及总

结生产实践经验的新学科。结合本册书的内容和重点,图像工程的主要构成可用图 1.1.2 所示的整体框架来表示,其中虚线框内为图像工程的基本模块。这里要用到各种图像技术以帮助人们从场景获得信息。首先要进行的就是利用各种方式从场景获得图像。接下来对图像的低层处理主要是为了改善图像的视觉效果或在保持视觉效果的基础上减少图像的数据量,处理的结果主要是给用户观看的。对图像的中层分析主要是对图像里感兴趣的目标进行检测、提取和测量。分析的结果能为用户提供描述图像目标特点和性质的数据。最后对图像的高层理解是要通过对图像里各目标的性质和它们之间相互关系的研究,了解把握图像内容并解释原来的客观场景。理解的结果能为用户提供客观世界的信息,从而指导和规划行动。这些从低层到高层所用的图像技术都得到了包括人工智能、神经网络、遗传算法、模糊理论、图像代数、机器学习、深度学习等新理论、新工具、新技术的有力支持。为完成这些工作还要采取合适的策略进行控制。

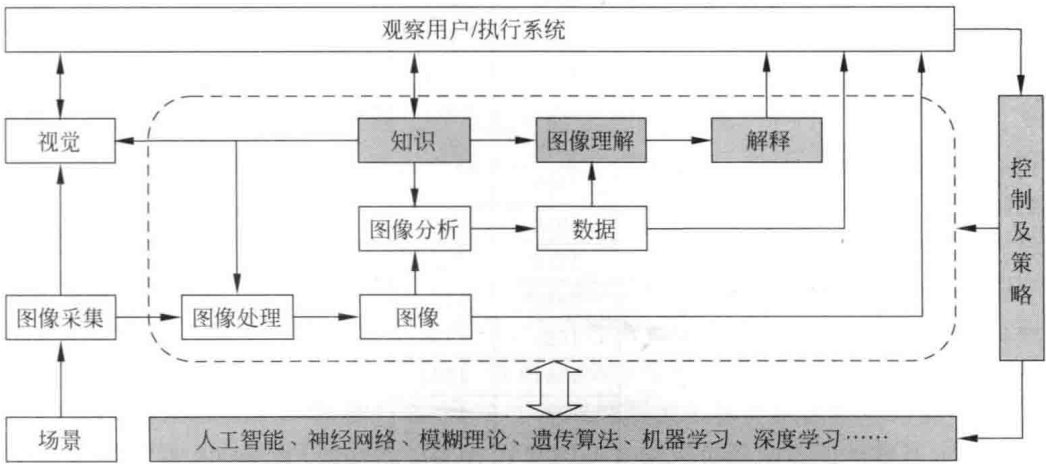


图 1.1.2 图像工程整体框架

本套书上册对低层图像处理的基本原理和技术进行了详细的介绍,本套书中册对中层图像分析的基本原理和技术进行了详细的介绍,本册书的内容将主要涉及高层图像理解的基本原理和技术,包括(在处理和 analysis 基础上)对 3-D 客观场景信息的获取和表达、景物的重建、场景的解释等,以及上述过程中的相关知识及应用,为完成这些工作所采用的控制和策略等。对高层理解的研究现已成为图像技术研究发展的一个重点,对不同层次图像技术的综合应用推动了图像事业的快速发展。

1.1.2 图像技术发展情况回顾

一个领域的研究发展可借助相关文献的发表情况分析,因为文献的发表是研究成果的一种体现。下面将借助两个综述系列及对它们的数据统计来概括地介绍一些图像技术的发展情况。

1. 一个已结束的图像技术综述系列

在 20 世纪最后约 30 年的时间内曾有一个由 30 篇论文组成的关于图像技术方面的综述系列。这个系列中综述论文的题目从开始直到 1986 年均称为“图像处理”,而从 1987 年开始称为“图像分析和计算机视觉”。该综述系列已于 2000 年由作者结束。作者声称,由于

在线信息获取的进展,该综述系列已不必要再延续下去了[Rosenfeld 2000]。该综述系列 30 年的总体情况,包括各综述论文题目、所综述的文献年份(引用年)和数量、刊登论文的刊物名称和论文的发表年份概括在表 1.1.1 中[Zhang 2002a]。由表中数据可见,文献数量有比较明显的“大小年”现象,这是由于综述考虑了一些两年一次的会议,导致有会年份的文献数量有较大的增加。

表 1.1.1 一个已结束综述系列的概况

序号	综述论文题目	引用年	数量/篇	刊登综述论文的刊物	发表年
1	Picture Processing by Computer	—1969	408	ACM Computing Surveys	1969
2	Progress in Picture Processing:1969—1971	1969—1971	580		1972
3	Picture Processing: 19xx (1972—1986)	1972	350	Computer Graphics and Image Processing (CGIP)	1973
4		1973	245		1974
5		1974	341		1975
6		1975	354		1976
7		1976	461		1977
8		1977	609		1978
9		1978	819		1979
10		1979	700		1980
11		1980	897		1981
12		1981	982		1982
13		1982	1185	Computer Vision, Graphics and Image Processing (CVGIP)	1983
14		1983	1138		1984
15		1984	1252		1985
16		1985	1063		1986
17		1986	1436		1987
18		1987	1412		1988
19		1988	1635		1989
20		1989	1187		1990
21		1990	1611		1991
22		1991	1178		1992
23	Image Analysis and Computer Vision: 19xx (1987—1999)	1992	1897	CVGIP: Image Understanding	1993
24		1993	1281		1994
25		1994	1911	Computer Vision and Image Understanding (CVIU)	1995
26		1995	1561		1996
27		1996	2148		1997
28		1997	1691		1998
29		1998	2268		1999
30		1999	1693		2000

这个综述系列在 30 年里所引用的共 34293 篇文献分别来自 40 多个刊物(大部分选自美国或国际刊物)和 10 多个大型国际会议。由表 1.1.1 可见,除了开始的两篇综述论文,其余在当前年发表的综述论文都是对上一年有关文献的统计(提供了文献清单)。这个系列的前两篇综述论文发表在 *ACM Computing Surveys* 上,其后的 28 篇都发表在 *Computer Graphics and Image Processing* (CGIP)和其更名后的刊物上。