

国外计算机科学教材系列

计算机视觉

—— 一种现代方法

Computer Vision: A Modern Approach

Computer
Vision

A Modern Approach

FORSYTH ■ PONCE



[美] David A. Forsyth 著
Jean Ponce

林学闾 王宏 等译

PEARSON
Prentice
Hall



電子工業出版社
Publishing House of Electronics Industry
<http://www.phei.com.cn>

国外计算机科学教材系列

计算机视觉

——一种现代方法

Computer Vision: A Modern Approach

[美] David A. Forsyth 著
Jean Ponce

林学间 王 宏, 等译

电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 提 要

计算机视觉的应用日趋成熟,学习与掌握计算机视觉基本原理及计算方法是从事计算几何、计算机图形学、图像处理、机器人学等专门人才的需要。由于计算机视觉集多种学科,如数字图像处理、数字信号处理、光学、物理学、几何学、应用数学、模式识别及人工智能等知识于一体,因而编写一本适用教材十分困难。

本书作者力求该书既做到系统条理,又能各章相对独立,便于学生通读或选择部分阅读。本书既讲清基本原理,又密切联系应用,使学生既能掌握基本原理又能与实际应用联系起来。既不乏经典理论,又侧重近年来的新鲜成果,使学生既了解计算机视觉的发展历史,又能把主要精力放在被实践证明为有效的近年研究的新成果上。本书将必要的数学知识融入各相关章节中,具有深入浅出的效果。本书作者为不同需要的学生设计了若干种不同的教学计划,该书网络简版已在美国若干学校试用,反映颇好。这是近年较成功的一本计算机视觉教材,也适合有兴趣的专业人才自学。

Simplified Chinese edition Copyright © 2004 by PEARSON EDUCATION ASIA LIMITED and Publishing House of Electronics Industry.

Computer Vision: A Modern Approach, ISBN: 0130851981 by David A. Forsyth and Jean Ponce. Copyright © 2003. All Rights Reserved.

Published by arrangement with the original publisher, Pearson Education, Inc., publishing as Prentice Hall.

This edition is authorized for sale only in the People's Republic of China (excluding the Special Administrative Region of Hong Kong and Macau).

本书中文简体字翻译版由电子工业出版社和 Pearson Education 培生教育出版亚洲有限公司合作出版。未经出版者预先书面许可,不得以任何方式复制或抄袭本书的任何部分。

本书封面贴有 Pearson Education 培生教育出版集团激光防伪标签,无标签者不得销售。

版权贸易合同登记号:图字:01-2003-0586

图书在版编目(CIP)数据

计算机视觉:一种现代方法/(美)福赛思(Forsyth, D. A.)等著;林学闯等译. —北京:电子工业出版社, 2004.6

(国外计算机科学教材系列)

书名原文:Computer Vision: A Modern Approach

ISBN 7-120-00086-1

I. 计... II. ①福...②林... III. 计算机视觉—教材 IV. TP302.7

中国版本图书馆 ICP 数据核字(2004)第 050522 号

责任编辑: 窦 昊 特约编辑: 詹文军

印刷者: 北京兴华印刷厂

出版发行: 电子工业出版社 <http://www.phei.com.cn>

北京市海淀区万寿路 173 信箱 邮编: 100036

经 销: 各地新华书店

开 本: 787×1092 1/16 印张: 37.5 字数: 966 千字

版 次: 2004 年 6 月第 1 次印刷

定 价: 55.00 元

凡购买电子工业出版社的图书,如有缺损问题,请向购买书店调换;若书店售缺,请与本社发行部联系。联系电话:(010) 68279077。质量投诉请发邮件至 zltz@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

译者序

人们利用自己的眼、耳、鼻、嘴以及其他器官从周围的环境中获取信息,从而在纷繁的世界中生存与发展。因此,一旦产生了计算机,人们也希望计算机能够替代他们与周围的环境交换信息,这就导致了计算机视觉这门学科的产生与发展。

计算机视觉成为一门独立的学科,至少可以从美国麻省理工学院 Marr 教授这一代人所做的奠基工作开始,因此计算机视觉这门学科至少已有 20~30 年的历史。在此期间,由于计算机视觉具有的潜在应用十分广泛,所涉及的学科知识极其繁多,研究的问题又极富挑战性,因此它一直是计算机学科中的一门热门学科,并吸引了许多从事心理学、神经科学、生理学、生物物理学、数学与计算机等各种学科研究人员的关注,从而把图像处理、模式识别、人工智能、数学、认知科学、机器学习、计算机图形学等各方面的研究成果融会进来。

要从事计算机视觉的研究,就要先学好这门学科,掌握这门学科的基础理论与治学方法,掌握最必要的数学工具。因此选择一本合适的教材与参考书是十分关键的,我们翻译的这本书是近年来一本关于计算机视觉的新书,与先前的各种介绍计算机视觉原理的教科书与著作相比较,作者从他们多年从事研究的经历出发,敏锐地感到计算机视觉这门学科已经日趋成熟,已经能在实际应用中发挥极其重要的作用,他们抓住服务于应用这条线,对计算机视觉发展过程中积累的理论成果进行了认真的挑选,从而编写了这样一本颇具特色的教科书:

- 作者力求该书既做到系统条理,又能各章相对独立,便于学生通读或选择部分内容阅读;
- 既讲清基本原理,又密切联系应用,使学生既能掌握基本原理又能与实际应用联系起来;
- 既不乏经典理论,又侧重近年的新鲜成果,使学生既了解计算机视觉的发展历史,又能把主要精力放在被实践证明为有效的近年研究的新成果上;
- 该书将必要的数学知识融入各相关章节中,具有深入浅出的效果;
- 作者为不同需求的学生设计了若干种不同的教学计划。该书的网络简版已在美国若干学校试用,反映颇好。

总之,这是近年来较成功的一本计算机视觉教材,也非常适合有兴趣的专业人才自学。

本书的主要译著者从事计算机视觉的研究与教学工作已有 20 年的历史,我们感到把这本书翻译成中文,对我国学习与研究计算机视觉的研究人员与学生会有帮助。

本书的第 1~5 章、第 12 章、第 13 章、第 19 章、第 20 章、第 22~26 章由林学闾及其学生马庚宇、杜杨洲、李旸和高华翻译,林学闾校阅;第 6~11 章、第 14~18 章、第 21 章由王宏及其学生董斌、于骞、安东、蔡文超、何英华、钱文杰、杨阳、叶明江翻译,王宏校阅。全书由林学闾校订。

在翻译过程中,我们力求忠实、准确地把握原著,同时保留原著风格。但是由于翻译时间仓促,原著涉及内容又十分广泛,难免有翻译不当乃至错误之处,殷切希望读者在学习过程中发现问题并随时与译著者联系,以便今后重印或再版时更正改进。

林学闾
王宏
清华大学计算机系

前 言

计算机视觉是一个处于知识前沿的领域。与其他前沿领域一样,它既激动人心又显得头绪繁多;在该领域经常可以看到缺乏权威性的现象;许多有用的做法并没有理论基础,而一些理论在实际应用中又毫无用处;虽然许多方面的研究已见成效,但是经常可以看到它们之间缺乏联系。尽管如此,我们还是力图在这本书中对这个领域给出一个相当有条理的分析。

我们认为计算机视觉,或简称为“视觉”,是一项事业,它与研究人类或动物的视觉是不同的。它借助于几何、物理和学习技术来构筑模型,从而用统计的方法来处理数据。因此从我们的角度看,在透彻理解摄像机性能与物理成像过程的基础上(这是本书第一篇的内容),视觉对每个像素值进行简单的推理(第二篇),将在多幅图像中可能得到的信息合成和谐的整体(第三篇),确定像素集之间的联系以便将它们彼此分割开,或推断一些形状信息(第四篇),使用几何信息(第五篇)或概率统计技术(第六篇)来识别物体。计算机视觉具有很广泛的应用,有早期研究较多的应用(如移动机器人导航、工业检测、军事侦察),也有较新的应用(如人机交互、数字图书馆中用的图像检索、医学图像处理以及计算机图形学中合成场景的逼真绘制)。我们将在第七篇中讨论一些具体应用。

为什么要研究视觉

计算机视觉主要的诀窍在于从图像或图像序列中获取对世界的描述。毫无疑问,这是很有使用价值的。摄取图像通常不带破坏性,是安全的。它又是一件不费力的事,现在的成本也不昂贵。用户希望从图像中获取的描述对不同的应用可能相差很大。例如,一种称为从运动求取结构(Structure from motion)的技术可以从图像序列中获取所见物体的描述以及摄像机的运动规律。娱乐产业中的人们利用这种技术来构造建筑物的三维模型,此时人们关注结构而忽略运动信息。这些模型可以应用到实际建筑物无法使用的场合,如火灾、爆炸等场合。只要利用数量很少的一组照片就可以构造出好的、简单准确的、令人信服的模型。而用这种技术来控制移动机器人时,人们一般关注运动,而将结构舍弃。这是因为,一般仅知道机器人工作的区域的某些方面,而并不知道机器人在这个区域的确切位置。可以从固定在机器人上的摄像机的运动信息中确定机器人的位置。

计算机视觉还有许多其他方面的重要应用,其中之一是医学图像处理与理解。人们可以设计软件系统来增强图像,或鉴别重要的现象或事件,或通过成像获得可视化信息。另一种应用是:检验人们对物体拍摄的图像,以便确定它们是否符合规定。第三种应用是卫星图像的理解。这既可用于军事目的,如编制程序来确定近来有没有与军事有关的现象在给定的地区内发生,或估计轰炸所引起的损害,也可以服务于民用的目的。今年的玉米收成会怎样,有多少雨林被保存下来?第四个应用是对收集的图片加以组织与结构化。我们知道如何去搜索与浏览图书馆的文本(尽管这仍然是具有未解决难题的课题),但确实不知道对于图片或视频图书馆应该如何做。

计算机视觉正处于一个自身发展的关键时期。从 20 世纪 60 年代起,人们就想利用计算机视觉的原理构造出有用的计算机系统,但这只是在最近才成为可能。这种繁荣的局面是受多方面的因素驱动的;计算机和图像系统的价格已经很便宜。在不太长时间之前,得到好的数字彩色图像要花费上万美元,而现在至多几百美元就够了。在不太久之前,彩色打印机是很难看到的,如果有的话也往往在研究实验室中才能见到,而现在它们出现在许多家庭中。这就意味着搞研究工作变得容易了。这也意味着有许多人产生了一些需要用计算机视觉方法来解决的问题。例如人们希望将收集的图片组织起来,为他们所在的周围世界构造三维模型以及管理与编辑收集的视频。我们对视觉中的基本几何学和物理学的理解以及如何运用它这一点已经得到了极大改善。我们正开始有可能解决许多人关注的问题,但是还没有一个困难的问题已得到解决,并且还有许多较容易的问题仍然没有解决(打算解决困难问题时要保持清醒的头脑)。现在正是研究这个主题的时候!

本书的内容

这本书包含的内容,我们认为从事计算机视觉研究的人们都应该知道。然而这本书是面对更广大的读者的。我们希望包括从事计算几何、计算机图形学、图像处理(总的说来与图像有关的)以及机器人学者会感到这是一本有益的参考书。我们试图做到让这本书对视觉这一课题有中等兴趣的本科高年级学生或研究生一年级的学生感到能接受。每一章覆盖这个题目不同的部分,而各章之间是相对独立的,这从表 1 就可瞥见。这就意味着读者不仅可以阅读整本书,也可着重于某一部分。一般说来,我们已努力做到使每一章从容易的内容开始,把深奥难懂的内容放在最后。每一章末尾都有一个小结,包含历史性资料与相应的观点。我们努力使这本书叙述有用的概念,或者在今后或许会有用。我们把重点放在理解成像的基本几何学和物理学知识上,但是也力图把它们与实际应用联系起来。总的说来,这本书反映了几何学以及多种形式的应用统计学在近年来对计算机视觉的多方面的影响。

如果读者从头到尾阅读本书,是会有收获的。当然也会很累,这本书包含的内容在一个学期学习是太多了。当然,未来的(或现正从事)计算机视觉专业人员应该逐字阅读、做每一个练习、报告所发现的问题,以便在第二版中更正。尽管学习计算机视觉并不要求学生具有很深的数学知识,但它的确要求学生熟悉多种不同的数学概念。我们已经努力使得该书能自成体系,这是指具有工科高年级数学理科水平的读者,对本书的内容会感到合适而不需要去参考其他课本。我们也注意到将数学降低到必不可少的程度,因为这本书是关于计算机视觉的,而不是应用数学方面的。我们还选择了将所保留的数学内容穿插在主要的章节内而不是单独的附录中的做法。

总而言之,我们已经努力减少各章之间的联系,以免只对某些专门话题感兴趣的读者在整本书中漫游。但做到各章完全自成体系是不可能的,各章之间的联系在表 1 中列出。

表 1 各章之间的关联程度：“必读章节”一列给出的章节必须很好地理解，“必读章节”一列给出的章节必须很好地理解，而“有助章节”一列给出的章节对学习该章是有帮助的

部 分	章 号	必 读 章 节	有 助 章 节
第一部分	1. 摄像机		
	2. 摄像机的几何模型	1	
	3. 摄像机的几何标定	2	
	4. 辐射学——光亮度度量		
	5. 光源、阴影与影调		4, 1
	6. 颜色		5
第二部分	7. 线性滤波		
	8. 边缘检测	7	
	9. 纹理	7	8
第三部分	10. 多视角几何学	3	
	11. 立体视觉	10	
	12. 从运动估计仿射模型	10	
	13. 从运动估计投影模型	12	
第四部分	14. 基于聚类的分割方法		9, 6, 5
	15. 基于模型拟合的分割		14
	16. 使用随机方法的分割与拟合		15, 10
	17. 基于线性动态模型的跟踪		
第五部分	18. 基于模型的视觉	3	
	19. 平滑曲面及其轮廓	2	
	20. 外观图	19	
	21. 距离数据		20, 19, 3
第六部分	22. 利用分类器建立模板		9, 8, 7, 6, 5
	23. 基于模板间关系的识别		9, 8, 7, 6, 5
	24. 基于空间关系的几何模板	2, 1	16, 15, 14
第七部分	25. 应用：在数字化收藏库中查找		16, 15, 14, 6
	26. 应用：基于图像的绘制	1	13, 12, 11, 6, 5, 3, 2, 1

本书未包含的内容

计算机视觉的参考文献数量是十分巨大的，因此要写出一本书能让普通人感兴趣不是一件容易的事。为此我们不得不删掉一些素材，去掉一些主题等。在最后一刻，删掉了两个完整的章节：一个是对概率论和推理的介绍，另一个是用非线性动力学跟踪物体的方法的叙述。这些章出现在本书的网页 <http://www.cs.berkeley.edu/~daf/book.html> 上。

去掉某些主题取决于我们个人的判断，或者是由于我们已感到筋疲力尽而停止了撰写某些方面，或者是由于我们得知它们太晚了以至于无法将它们加进去，也可能由于我们不得不缩

短某些章,或许还有其他种种理由。我们有意忽略那些主要与历史有关的细节,而将历史的评论放在每章的末尾。我们不认为自己是个成功的知识考古学家,这意味着一些概念可能比我们所表达的有更深远的历史意义。我们并没有回避撰写变形模板与拼图这两个具有相当重要实用意义的话题,我们将试图把它们放到第二版中去。

致谢

在筹备这本书的过程中,我们已经欠下了一大笔债。许多不知其名的评阅者已经读过这本书的几种草稿版本,并且对这本书做出了非常有用的贡献。感谢他们所花费的时间与精力。编辑 Alan Apt 在 Jake Warde 的帮助下组织了评阅,对他们表示感谢。Integre Technical Publishing 的 Leslie Galen, Joe Albrecht 和 Dianne Parish 帮助解决了校对和说明方面的许多问题。书中使用的一些图像是从 IMSI's Master Photos Collection 公司得到的。在筹集参考文献的过程中,我们大量使用了 Keith Price 提供的有关计算机视觉方面极好的参考文献,这可以从网上查到,地址为 <http://iris.usc.edu/vision-notes/bibliography/contents.html>。

一些同事评阅了整本书的内容以及若干章节,他们对这些章节的修订提出了宝贵且详细的建议。我们要感谢 Kobus Barnard, Margaret Fleck, David Kriegman, Jitendra Malik 和 Andrew Zisserman。我们的许多学生也在提建议、图示创意、校对评论以及其他方面做出了贡献。我们要感谢 Okan Arikan, Sébastien Blind, Martha Cepeda, Stephen Chenney, Frank Cho, Yakup Genc, John Haddon, Sergey Ioffe, Svetlana Lazebnik, Cathy Lec, Sung-il Pae, David Parks, Fred Rothganger, Attawith Sudsang 以及在加州大学伯克利分校 UIUC 上视觉课的一些学生们的贡献。幸运的是,一些大学的同事们在视觉课中用了这本书的草稿版本。使用过这本书初稿版本的学校有 Carnegie-Mellon 大学、Stanford 大学、在 Madison 的 Wisconsin 大学、在 Santa Barbare 的 California 大学以及 Southern California 大学,也可能有一些我们不知道的其他学校。我们对所有使用这本书的读者所提的建设性意见表示感谢,特别要感谢 Chris Bregler, Chuck Dyer, Martial Hebert, David Kriegman, B. S. Manjunath 和 Ram Nevatia, 他们提供了许多详尽的、非常有帮助的评论与改正意见。这本书还受益于 Aydin Alaylioglu, Srinivas Akella, Marie Banich, Serge Belongie, Ajit M. Chaudhari, Navneet Dalal, Richard Hartley, Glen Healey, Mike Heath, Hayley Iben, Stéphanie Jonqnières, Tony Lewis, Benson Limketkai, Simon Maskell, Brian Milch, Tamara Miller, Cordelia Schmid, Brigitte 和 Gerry Serlin, Ilan Shimshoni, Eric de Sturler, Camillo J. Taylor, Jeff Thompson, Claire Vallant, Daniel S. Wilkerson, Jinghan Yu, Hao Zhang 和 Zhengyou Zhang。如果读者发现了明显的印刷排版错误,请给 DAF 发电子邮件(daf@cs.berkeley.edu),使用短语“book typo”告诉我们细节,我们将在第二版中感谢每个错误的第一个发现者。

还要感谢 P. Besl, B. Boufama, J. Costeira, P. Debevec, O. Fangeras, Y. Genc, M. Hebert, D. Huber, K. Ikenchi, A. E. Johnson, T. Kanade, K. Kutulakos, M. Levoy, S. Mahamud, R. Mohr, H. Moravec, H. Murase, Y. Ohta, M. Okutami, M. Pollefeys, H. Saito, C. Schmid, S. Sullivan, C. Tomasi 以及 M. Turk, 感谢他们为本书的某些插图提供了原件。

对教学大纲的建议

整本书从第一页开始讲起,可以分为两个学期讲,内容还是比较紧凑的。可以安排应用中的一章(如关于“基于图像绘制”这一章)放在第一学期,而将有关应用的另一章放在第二学期。但是估计有的系不会需要有如此详细的序列课。我们已在编排这本书时考虑到教师可以按他们的喜好选择其中的内容。表 2 到表 6 列出一些用于一学期 15 周课的教学大纲案例,它们是根据我们的设想来安排的。我们鼓励(并希望)教师们按自己的兴趣来重新安排。

表 2 列出的教学大纲,是为计算机科学、电气工程或其他工程与自然科学学科的本科高年级学生或一年级研究生设计的计算机视觉导论课,历时一个学期。学生们可以学到这个领域的多个方面,包括在数字图书馆以及基于图像绘制等方面的应用。尽管最难的理论部分被去掉了,但是成像的基本几何和物理学知识是较深入的。我们假设学生具有较广的背景知识,并建议在第二周或第三周读一下概率论的背景知识(网上的这本书有关于其的一章)。我们将应用章节放到书的末尾,但许多人可能会选择在第 10 周讲第 20 章,在第 6 周讲第 21 章。

表 2 对计算机科学、电子工程或其他工科或理科的一年级研究生或本科高年级学生开一个学期计算机视觉的导论课的内容

周次	章号	节号	主要内容
1	1,4	1.1,4(只限小结)	针孔摄像机,辐射学术语
2	5	5.1~5.5	局部影调模型,点、线和面光源,光度学体视
3	6	全部	颜色
4	7,8	7.1~7.5,8.1~8.3	线性滤波器,平滑抑制噪声,边缘检测
5	9	全部	纹理,用滤波器输出表示其统计量,纹理合成;由纹理推断形状
6	10,11	10.1,11	基本的多视角几何,立体视觉
7	14	全部	用聚类实现分割
8	15	15.1~15.4	拟合直线与曲线,用最大似然率进行拟合,鲁棒性
9	16	16.1,16.2	隐变量与 EM 算法
10	17	全部	用卡尔曼滤波来跟踪,数据相关
11	2,3	2.1,2.2,第 3 章全部	摄像机标定
12	18	全部	使用特征对应和摄像机标定的基于模型的视觉
13	22	全部	使用分类器的模板匹配
14	23	全部	基于关系的匹配
15	25,26	全部	在数字图书馆中检索图像,基于图像的绘制

表 3 所列教案是为计算机图形学的学生设计的。他们想知道与他们的课题有关的视觉基础知识。我们在此强调了能从图像提供信息恢复物体模型的方法,了解这些内容需要了解摄像机和滤波器的运作机理。跟踪在图形学领域变得很有用,其对运动分析十分重要。我们认

为学生已具有很广的背景知识,并且对概率论有一定的了解。

表3 适用于计算机图形学学生用的教学大纲,他们想知道视觉与他们的课题有关的一些内容

周次	章号	节号	主要内容
1	1,4	1.1,4(只限小结)	针孔摄像机,辐射学术语
2	5	5.1~5.5	局部影调模型,点、线和面光源,光度学电视
3	6.1~6.4	全部	颜色
4	7,8	7.1~7.5,8.1~8.3	线性滤波器,平滑抑制噪声,边缘检测
5	9	9.1~9.3	纹理,用滤波器输出表示其统计量,纹理合成
6	2,3	2.1,2.2,第3章全部	摄像机标定
7	10,11	10.1,11	基本的多视角几何,立体视觉
8	12	全部	由运动恢复仿射结构
9	13	全部	由运动恢复射影结构
10	26	全部	基于图像的绘制
11	15	全部	拟合,鲁棒性,RANSAC
12	16	全部	隐变量与EM算法
13	19	全部	表面与轮廓
14	21	全部	距离数据
15	17	全部	跟踪,卡尔曼滤波与数据相关

表4中的教学大纲主要是为对计算机视觉应用感兴趣的学生制定的。该教学大纲覆盖了与应用直接有关的内容,我们假定这些学生已具有相当广泛的背景知识,也可以在第2~3周时安排一个关于概率论的背景阅读。

表4 对计算机应用感兴趣的学生的教学大纲

周次	章号	节号	主要内容
1	1,4	1.1,4(只限小结)	针孔摄像机,辐射学术语
2	5,6	5.1,5.3,5.4,5.5,6.1~6.4	局部影调模型;点、线和面光源;光度学电视;颜色-物理学,人的感知,颜色空间
3	2,3	全部	摄像机模型及其标定
4	7,9	第7章全部,9.1~9.3	线性滤波器,用滤波器输出表示纹理的统计量,纹理合成
5	10,11	全部	多视角几何,以立体视觉为例
6	12,13	全部	由运动推断仿射结构,由运动推断射影结构
7	13,26	全部	由运动推断射影结构,基于图像的绘制
8	14	全部	由聚类方法进行分割,重点在镜头边界检测与背景差分
9	15	全部	拟合直线、曲线,鲁棒性,RANSAC
10	16	全部	隐变量与EM算法
11	25	全部	在数字图书馆中检索图像

(续表)

周次	章号	节号	主要内容
12	17	全部	跟踪,卡尔曼滤波与数据相关
13	18	全部	基于模型的视觉
14	22	全部	使用分类器检测模板
15	20	全部	距离数据

表5的教案是为认知科学或人工智能学科的学生设计的,他们需要对计算机视觉重要概念的基本梗概有所了解。这个教案显得不那么步步紧逼,对学生在数学方面的要求也较少。学生需要在第2周或第3周学一些概率论的内容。

表5 对认知科学或人工智能学生的大纲,他们希望对计算机视觉的重要概念有一个基本的了解

周次	章号	节号	主要内容
1	1,4	1,4(只限小结)	针孔摄像机,镜头,摄像机与人眼,辐射度学术语
2	5	全部	局部影调模型;点、线和面光源;光度学电视,互反射,光亮度计算
3	6	全部	颜色:物理,人的感知,空间,图像模型颜色恒常性
4	7	7.1~7.5,7.7	线性滤波器;采样;尺度
5	8	全部	边缘检测
6	9	全部	纹理;表达式,合成,由纹理推断形状
7	10.1,10.2	全部	基本的多视角几何
8	11	全部	立体视觉
9	14	全部	用聚类方法实现分割
10	15	全部	拟合直线、曲线,鲁棒性,RANSAC
11	16	全部	隐变量与EM算法
12	18	全部	基于模型的视觉
13	22	全部	用分类器检测模板
14	23	全部	基于模板间关系进行识别
15	24	全部	用空间关系建立几何模板

表6的教案是为那些对应用数学、电气工程或物理学有强烈兴趣的学生设计的。这个教案使一学期的内容很紧凑,进展很快,并且假设学生能够适应许多教学内容。我们假设学生具有宽广的背景知识,并且能够在第2周或第3周给他们指定阅读一些概率论的内容。我们在这样一个相当概括和要求甚高的教案中,安插了对数字图书馆的简短综述作为缓冲,也可以用基于图像绘制这一章或距离数据来代替。

表 6 适用于对应用数学、电子工程或物理学有浓厚兴趣的学生的大纲

周次	章号	节号	主要内容
1	1,4	全部	摄像机,辐射度学
2	5	全部	影调模型;点、线和面光源;光度学体视,互反射以及影调基元
3	6	全部	颜色-物理学,人的感知,空间,颜色恒常性
4	2,3	全部	摄像机参数与标定
5	7,8	全部	线性滤波器与边缘检测
6	8,9	全部	边缘检测;纹理;表达式,合成,推断形状
7	10,11	全部	多视角几何,以立体视觉为例
8	12,13	全部	由运动推断结构
9	14,15	全部	用聚类方法进行分割;拟合直线、曲线;鲁棒性;RANSAC
10	15,16	全部	拟合;隐变量与 EM 算法
11	17,25	全部	跟踪;卡尔曼滤波器,数据相关,数字图书馆中检索图像
12	18	全部	基于模型的视觉
13	19	全部	表面与它们的轮廓
14	20	全部	外观图
15	22	全部	模板匹配

编程作业和源程序

本书中给出的编程作业经常需要数值线性代数、奇异值分解以及线性与非线性最小二乘的程序。这些程序的较完整集合,可以在 MATLAB 以及一些公共图书馆中得到,例如 LINPACK、LAPACK 和 MINPACK,它们可以从 Netlib 库中下载(<http://www.netlib.org/>)。我们在本书的网页 <http://www.cs.berkeley.edu/~daf/book.html> 上也给出了一些至其他软件的链接。也可在此处找到编程作业的数据集或到数据集的链接。

目 录

第一部分 图像生成与图像模型

第 1 章 摄像机	2
1.1 针孔照相机	2
1.2 带镜头的摄像机	6
1.3 人的眼睛	10
1.4 信号感应	12
1.5 注释	15
习题	16
第 2 章 摄像机的几何模型	17
2.1 欧几里得解析几何基础	17
2.2 摄像机参数和透视投影	23
2.3 仿射摄像机和仿射投影方程	26
2.4 注释	29
习题	30
第 3 章 摄像机的几何标定	32
3.1 最小二乘法的参数估计	32
3.2 使用线性方法进行摄像机标定	37
3.3 径向畸变	40
3.4 分析摄影地形测量法	42
3.5 应用: 机器人定位	43
3.6 注释	44
习题	45
第 4 章 辐射学——光亮度度量	46
4.1 空间中的光	46
4.2 到达表面的光	50
4.3 重要的特殊情况	53
4.4 注释	56
习题	57
第 5 章 光源、阴影与影调	59
5.1 定性辐射学	59
5.2 光源及其产生的效果	60

5.3	局部影调模型	65
5.4	应用:光度学体视	68
5.5	互反射:全局影调模型	74
5.6	注释	79
	习题	81
第6章	颜色	83
6.1	物理学中的颜色	83
6.2	人类的颜色感知	87
6.3	颜色表示	90
6.4	图像颜色的一个模型	97
6.5	从图像颜色中找到表面颜色	103
6.6	注释	109
	习题	111

第二部分 低层视觉:使用一幅图像

第7章	线性滤波	114
7.1	线性滤波和卷积	114
7.2	移不变线性系统	118
7.3	空间频率和傅里叶变换	123
7.4	采样和折叠失真	126
7.5	滤波器与模板	132
7.6	技术:归一化相关和检测模式	133
7.7	技术:尺度和图像金字塔	135
7.8	注释	137
	习题	138
第8章	边缘检测	140
8.1	噪声	140
8.2	导数估计	143
8.3	对边缘进行检测	148
8.4	注释	157
	习题	159
第9章	纹理	161
9.1	纹理表示	161
9.2	使用有方向性金字塔的分析(和合成)	167
9.3	应用:合成纹理来绘制	174
9.4	由纹理得到形状	176
9.5	注释	179
	习题	180

第三部分 低层视觉:使用多幅图像

第 10 章 多视角几何学	182
10.1 双视角	182
10.2 三视图	188
10.3 更多的视图	192
10.4 注释	195
习题	197
第 11 章 立体视觉	199
11.1 重建	200
11.2 人类的立体视觉过程	201
11.3 双目融合	204
11.4 使用多个摄像机	209
11.5 注释	211
习题	213
第 12 章 从运动估计仿射模型	214
12.1 仿射几何基础	215
12.2 仿射结构和两幅图之间的运动	220
12.3 从多幅图像估计仿射结构和运动	224
12.4 从仿射到欧氏图像	226
12.5 仿射运动分割	229
12.6 注释	231
习题	232
第 13 章 从运动估计投影模型	234
13.1 投影几何基础	234
13.2 从双目对应估计运动和投影结构	243
13.3 多线性约束估计投影运动	246
13.4 多幅图像恢复运动和投影结构	247
13.5 从投影图像到欧氏图像	250
13.6 注释	252
习题	252

第四部分 中层视觉

第 14 章 基于聚类的分割方法	256
14.1 什么是分割	256
14.2 人类视觉:分类和格式塔原理	258
14.3 应用:镜头的边界检测和背景差分	263
14.4 基于像素点聚类的图像分割	266

14.5	基于图论的聚类分割	269
14.6	注释	277
	习题	278
第 15 章	基于模型拟合的分割	280
15.1	哈夫变换	280
15.2	直线拟合	283
15.3	拟合曲线	286
15.4	作为概率问题的拟合	290
15.5	鲁棒性	291
15.6	举例:用 RANSAC 来拟合基础矩阵	296
15.7	注释	299
	习题	300
第 16 章	使用随机方法的分割与拟合	302
16.1	丢失数据问题、拟合和分割	302
16.2	EM 算法的应用	306
16.3	模型选择:哪个模型拟合得最好	315
16.4	注释	317
	习题	318
第 17 章	基于线性动态模型的跟踪	320
17.1	把跟踪作为一个抽象的推理问题	320
17.2	线性动态模型	322
17.3	卡尔曼滤波	326
17.4	数据相关	333
17.5	应用和例子	336
17.6	注释	340
	习题	340

第五部分 高层视觉几何方法

第 18 章	基于模型的视觉	344
18.1	初始假设	344
18.2	通过位姿一致性获取假设	345
18.3	位姿聚类获得假设	349
18.4	采用不变量获得假设	351
18.5	校验	357
18.6	应用:医学图像系统的对准	359
18.7	曲面与对准	363
18.8	注释	363
	习题	365

第 19 章 平滑表面及其轮廓	367
19.1 微分几何的基本要点	368
19.2 表面轮廓几何学	375
19.3 注释	379
习题	379
第 20 章 外观图	381
20.1 视觉事件:微分几何的补充	383
20.2 计算外观图	391
20.3 外观图与物体定位	395
20.4 注释	398
习题	399
第 21 章 距离数据	401
21.1 主动距离传感器	401
21.2 距离数据的分割	402
21.3 距离图像的匹配和模型获取	410
21.4 物体识别	414
21.5 注释	419
习题	421

第六部分 高层视觉:基于概率和推理的方法

第 22 章 利用分类器建立模板	424
22.1 分类器	424
22.2 基于类直方图创建分类器	431
22.3 特征选择	434
22.4 神经网络	443
22.5 支持向量机	451
22.6 注释	455
习题	457
22.7 附录 I:向后传播算法	457
22.8 附录 II:线性不可分数据集上的支持向量机	460
22.9 附录 III:非线性支持向量机	461
第 23 章 基于模板间关系的识别	463
23.1 通过对模板间关系投票检测物体	463
23.2 利用概率模型及搜索的关系推理	468
23.3 利用分类器简化搜索	471
23.4 隐马尔可夫模型	474
23.5 应用:基于隐马尔可夫模型的手语理解	483
23.6 应用:基于隐马尔可夫模型的人体检测	486