

普通高等教育
“九五”国家级重点教材



全国高技术重点图书
自动化技术领域



中国科学院研究生教学丛书

计算机视觉

——计算理论与算法基础

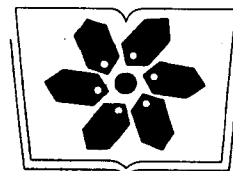
马颂德 张正友 著

科学出版社



139.4
M22

410749



中国科学院科学出版基金资助出版

计算机视觉

——计算理论与算法基础

马颂德 张正友 著



科学出版社

1998

JS142/B
内 容 简 介

计算机视觉是在图像处理的基础上发展起来的新兴学科。计算机视觉从信息处理的层次研究视觉信息的认知过程,研究视觉信息处理的计算理论、表达与计算方法。本书系统地介绍了计算机视觉的重要理论与算法,包括图像特征提取,摄像机定标,立体视觉,运动视觉(或称序列图像分析),由图像灰度恢复三维物体形状的方法,物体建模与识别方法以及距离图像分析方法等。

本书是在作者十多年来从事计算机视觉的研究和研究生教育的基础上编写而成的,书中不仅包含了初次接触本学科的读者所需要的基础知识,也介绍了近年来国内外计算机视觉研究的重要理论研究成果。本书的大多数内容已在中国科学院北京研究生院教授多年。本书附有图像实验数据与参考实验结果,可供研究生或研究者进行实验研究。

本书可作为信息处理、计算机、机器人、人工智能、遥感图像处理、认知神经科学等有关专业的高年级学生或研究生的教学用书,也可供以上领域的研究工作者参考。

图书在版编目(CIP)数据

计算机视觉——计算理论与算法基础/马颂德,张正友著.

北京:科学出版社,1997

ISBN 7-03-006070-9

I. 计… I. ①马… ②张… III. 计算机视觉
N. TP302.7

中国版本图书馆 CIP 数据核字(97)第 07413 号

科学出版社出版

北京东黄城根北街 16 号

邮政编码: 100717

中国科学院印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

*

1998 年 1 月第 一 版 开本: 787×1092 1/16

1998 年 1 月第一次印刷 印张: 18 3/4

印数: 1—2 000 字数: 418 000

定价: 32.00 元

《中国科学院研究生教学丛书》总编委会

主任：

路甬祥

常务副主任：

白春礼

副主任：

李云玲 师昌绪 杨 乐 汪尔康 沈允钢

黄荣辉 叶朝辉 李 佩

委员：

赵保恒 匡廷云 冯克勤 冯玉琳 朱清时

王 水 刘政凯 龚 立 侯建勤 颜基义

黄凤宝

《中国科学院研究生教学丛书》技术学科编委会

主 编：

师昌绪

副主编：

冯玉琳

编 委：

刘政凯 徐至展 陈先霖 王占国 马颂德

吴承康 史忠植

《中国科学院研究生教学丛书》序

在 21 世纪曙光初露,中国科技、教育面临重大改革和蓬勃发展之际,《中国科学院研究生教学丛书》——这套凝聚了中国科学院新老科学家、研究生导师们多年心血的研究生教材面世了。相信这套丛书的出版,会在一定程度上缓解研究生教材不足的困难,对提高研究生教育质量起着积极的推动作用。

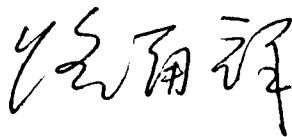
21 世纪将是科学技术日新月异,迅猛发展的新世纪,科学技术将成为经济发展的最重要的资源和不竭的动力,成为经济和社会发展的首要推动力量。世界各国之间综合国力的竞争,实质上是科技实力的竞争。而一个国家科技实力的决定因素是它所拥有的科技人才的数量和质量。我国要想在 21 世纪顺利地实施“科教兴国”和“可持续发展”战略,实现小平同志规划的三步战略目标——把我国建设成中等发达国家,关键在于培养造就一支数量宏大、素质优良、结构合理,有能力参与国际竞争与合作的科技大军。这是摆在我国高等教育面前的一项十分繁重而光荣的战略任务。

中国科学院作为我国自然科学与高新技术的综合研究与发展中心,在建院之初就明确了出成果出人才并举的办院宗旨,长期坚持走科研与教育相结合的道路,发挥了高级科技专家多,科研条件好,科研水平高的优势,结合科研工作,积极培养研究生;在出成果的同时,为国家培养了数以万计的研究生。当前,中国科学院正在按照江泽民同志关于中国科学院要努力建设好“三个基地”的指示,在建设具有国际先进水平的科学研究基地和促进高新技术产业发展基地的同时,加强研究生教育,努力建设好高级人才培养基地,在肩负起发展我国科学技术及促进高新技术产业发展重任的同时,为国家源源不断地培养输送大批高级科技人才。

质量是研究生教育的生命,全面提高研究生培养质量是当前我国研究生教育的首要任务。研究生教材建设是提高研究生培养质量的一项重要的基础性工作。由于各种原因,目前我国研究生教材的建设滞后于研究生教育的发展。为了改革这种情况,中国科学院组织了一批在科学前沿工作,同时又具有相当教学经验的科学家撰写研究生教材,并以专项资金资助优秀的研究生教材的出版。希望通过数年努力,出版一套面向 21 世纪科技发展,体现中国科学院特色的高水平的研究生教学丛书。本丛书内容力求具有科学性、系统性和基础性,同时也兼顾前沿性,使阅读者不仅能获得相关学科的比较系统的科

学基础知识,也能被引导进入当代科学研究的前沿。这套研究生教学丛书,不仅适合于在校研究生学习使用,也可以作为高校教师和专业研究人员工作和学习的参考书。

“桃李不言,下自成蹊。”我相信,通过中国科学院一批科学家的辛勤耕耘,《中国科学院研究生教学丛书》将成为我国研究生教育园地的一丛鲜花,也将似润物春雨,滋养莘莘学子的心田,把他们引向科学的殿堂,不仅为科学院,也为全国研究生教育的发展作出重要贡献。

A handwritten signature in black ink, reading 'Guo Yongqiang' (郭永强) in a cursive style.

致 谢

特别感谢中国科学院自动化研究所模式识别国家重点实验室的同事和学生们,实验室良好的研究环境为本书的写作创造了条件.谨将本书献给我的妻子、父母和家人,感谢他们在我繁忙的工作中所给予的理解与支持.

马颂德

特别感谢法国国家信息与自动化研究院(INRIA)为我的研究工作提供了便利,使得本书的写作成为可能;感谢 Faugeras 教授多年的关怀与指导,和为实验室创造了活跃的富有成效的科研环境.谨将本书献给我的父母和家人,感谢他们对我工作的理解与支持.

张正友

前 言

人类通过视觉识别文字、图片和周围的环境；利用听觉识别与理解语言。视觉与听觉是人类智能的重要组成部分。从信息处理的角度搞清它们的机理，研究它们的计算理论与算法，以使计算机实现人的视觉、听觉等模式识别能力，是人类在基础研究与应用研究中面临的最重大的挑战之一。如果机器不能自动感知与识别周围环境，机器智能也就无从谈起。因此，用计算机实现模式的自动识别，是开发智能机器的一个最关键的突破口。它的成功应用将大大推动人工智能系统的发展，拓广计算机与各种自动机器的应用范围。

人类是通过眼睛与大脑来获取、处理与理解视觉信息的。周围环境中的物体在可见光的照射下，在人眼的视网膜上形成图像，由感光细胞转换成神经脉冲信号，经神经纤维传入大脑皮层进行处理与理解。视觉，不仅指对光信号的感受，它包括了对视觉信息的获取、传输、处理、存储与理解的全过程。信号处理理论与计算机出现以后，人们试图用摄像机获取环境图像并转换成数字信号，用计算机实现对视觉信息处理的全过程，这样，就形成了一门新兴的学科——计算机视觉。

计算机视觉的研究目标是使计算机具有通过一幅或多幅图像认知周围环境信息的能力。这种能力将不仅使机器能感知环境中物体的几何信息，包括它的形状、位置、姿态、运动等，而且能对它们进行描述、存储、识别与理解。计算机视觉发展得益于神经生理学、心理学与认知科学对生物视觉系统的研究，但多年来的研究实践表明，人类对自身具有的认知能力的认识还远远不够。从事心理学、神经科学、生理学、生物物理学、数学与计算机科学的研究人员越来越迫切地感到，需要联合起来对人脑的认知过程进行从宏观到微观的深入研究。虽然由于人脑的高度复杂性，这种跨学科的研究还远远不够深入，但从事计算机视觉的研究者们已发展起一套独立的计算理论与算法，从而能用计算机对视觉信息（或者说，对图像）进行分析与处理。计算机视觉，作为一门独立的学科正在受到广泛的重视，计算机视觉已应用于遥感图像分析，文字识别，医学图像处理，多媒体技术，图像数据库，工业检测与军事等方面。我们相信，随着跨学科基础研究的不断深入，随着计算机性能的快速提高，计算机视觉将广泛地应用于更复杂的应用场合。

本书是作者在中国科学院自动化研究所模式识别国家重点实验室与法国国家信息与自动化研究院机器人视觉实验室从事十多年的研究工作的基础上编写的。书中较全面地介绍了计算机视觉的专业基础知识与视觉信息处理各阶段的主要计算理论与计算方法。作者在国内外学术刊物与国际学术会议上发表了 200 多篇论文，书中介绍了作者在这些论文中报道的部分研究成果。本书的大部分内容还是作者多年来在中国科学院北京研究生院和法国尼斯大学讲授计算机视觉课程的教材内容，有些还是作者 1994 年在香港浸会大学计算机系的讲课内容。本书第三章的一部分与第六至第十章由张正友编写，其余各章由马颂德编写，张正友还为本书加上了附录软盘，其中包含计算机视觉的实验数据及参考结果，以便于读者作实验。需要数据者，请与作者联系。

本书的出版得到了中国科学院科学出版基金的资助，本书作者在研究工作中还获得

多方面的研究基金与研究计划的资助,其中有国家攀登计划、国家自然科学基金、国家863计划、中国科学院重点研究项目、欧洲 ESPRIT 计划、中法高级研究合作计划、法国航天局(CNES)及欧洲航天局(ESA)等研究项目的资助。作者对组织这些研究计划与基金的有关部门,如国家科委基础研究与高技术司,国家自然科学基金委,中国科学院基础局等表示感谢;作者还特别感谢国家计委科技司,十多年前,由于国家计委的远见卓识,在国内建立了一批国家重点实验室,本书第一作者所在的模式识别国家重点实验室就是其中之一,国家重点实验室较好的研究条件与对基础研究的重视,为作者创造了良好的研究环境。作者还要感谢香港浸会大学计算机系与香港 Croucher 基金会,在该基金的支持下,本书第一作者在香港浸会大学计算机系作为访问教授工作了5个月,使作者有时间完成了本书的大部分编写工作。本书另一作者感谢法国国家信息与自动化研究院(INRIA)为他提供了良好的科研条件及各种便利,使得本书的写作得以顺利完成。

作者十分感谢他们的同事与学生,他们是中国科学院自动化研究所模式识别国家重点实验室的罗曼丽、胡占义、蒋田仔、卢汉清、田英利、华军、魏国庆、李炳成、管伟光、丛舸、杨智勇等,法国国家信息与自动化研究院的 O. Faugeras, N. Ayache, R. Deriche, O. Monga, ……。本书介绍的许多工作是他们的研究工作或是作者与他们在合作中完成的,在与他们的讨论中,作者获益匪浅。

作者还要特别感谢娄文利与赵微,本书的大部分初稿都是她们打印的,她们为本书的出版付出了艰辛与繁琐的劳动;作者特别感谢本书的责任编辑,科学出版社的李淑兰女士,她为本书的文字润色与校对作了大量的工作,没有她的鼓励与协作,本书的出版是不可能的。

计算机视觉的内容十分广泛,与之有关的学科也很多,由于作者学识有限以及书的篇幅限制,书中错误与片面之处在所难免,敬请读者不吝批评指正。

马颂德 张正友

1997年

目 录

第一章 绪论	(1)
1.1 生物视觉通路简介	(2)
1.1.1 生物视觉通路	(2)
1.1.2 感受野的分层等级假设	(2)
1.1.3 视觉信息处理的多通道、多任务并行处理性质	(4)
1.2 Marr 的计算视觉理论框架	(5)
1.2.1 视觉系统研究的三个层次	(5)
1.2.2 视觉信息处理的三个阶段	(6)
1.3 本书各章内容简介	(7)
1.4 计算机视觉的现状与阅读本书需注意的问题	(8)
思考题	(9)
参考文献	(9)
第二章 边缘检测	(11)
2.1 边缘检测与微分滤波器	(12)
2.1.1 边缘检测与微分运算	(12)
2.1.2 微分滤波器	(14)
2.1.3 离散信号的差分滤波器	(16)
2.2 边缘检测与正则化方法	(19)
2.2.1 病态问题	(19)
2.3 多尺度滤波器与过零点定理	(21)
2.4 最优边缘检测滤波器	(25)
2.5 边缘检测快速算法	(28)
2.5.1 滤波器的递推算法	(29)
2.5.2 二维可分离滤波器	(32)
2.6 图像低层次处理的其他问题	(33)
思考题	(34)
参考文献	(34)
第三章 射影几何与几何元素表达	(36)
3.1 仿射变换与射影变换的几何表达	(36)
3.2 仿射坐标系与射影坐标系	(38)
3.3 仿射变换与射影变换的代数表达	(40)
3.4 不变量	(42)

3.5	由对应点求射影变换	(43)
3.6	点	(44)
3.7	指向和方向	(44)
3.7.1	二维	(44)
3.7.2	三维	(44)
3.8	平面直线及点线对偶关系	(45)
3.9	空间平面及点面对偶关系	(46)
3.10	空间直线	(48)
3.11	二次曲线与二次曲面	(49)
	思考题	(49)
	参考文献	(51)
第四章	摄像机定标	(52)
4.1	线性模型摄像机定标	(52)
4.1.1	图像坐标系、摄像机坐标系与世界坐标系	(52)
4.1.2	线性摄像机模型(针孔模型)	(54)
4.1.3	线性模型摄像机定标	(56)
4.2	非线性模型摄像机定标	(59)
4.3	立体视觉摄像机定标	(60)
4.4	机器人手眼定标	(63)
4.5	摄像机自定标技术	(67)
	思考题	(70)
	参考文献	(71)
第五章	立体视觉	(72)
5.1	立体视觉与三维重建	(72)
5.1.1	空间点重建	(72)
5.1.2	空间直线重建	(75)
5.1.3	空间二次曲线重建	(76)
5.2	极线约束	(78)
5.3	对应基元匹配	(80)
5.4	射影几何意义下的三维重建	(87)
	思考题	(93)
	参考文献	(93)
第六章	运动与不确定性表达	(95)
6.1	欧氏平面上的刚体运动	(95)
6.2	欧氏空间中的刚体运动	(95)
6.2.1	定义	(96)
6.2.2	旋转向量表示法	(97)

6.2.3 四元数表示法	(98)
6.3 不确定性的描述	(99)
6.3.1 引言	(99)
6.3.2 几何概率论	(99)
6.4 不确定性的运算	(101)
6.4.1 一个不确定向量是另一个不确定向量的函数	(102)
6.4.2 通过优化方法得到一个不确定向量	(102)
6.4.3 Mahalanobis 距离	(103)
6.5 不确定性运算的几个例子	(105)
6.5.1 边缘点的不确定性	(105)
6.5.2 从边缘点到二维直线的不确定性	(106)
6.5.3 三维重建的点的确定性	(107)
6.6 三维直线段的不确定性	(110)
6.6.1 直线段与无限长直线	(110)
6.6.2 直线段的方向	(111)
6.6.3 直线段的位置	(112)
6.7 不确定性的显示	(113)
思考题	(116)
参考文献	(116)

第七章 基于光流场的运动分析

7.1 光流场和运动场	(117)
7.1.1 运动场	(117)
7.1.2 光流	(118)
7.2 光流的约束方程	(119)
7.2.1 开孔问题	(119)
7.2.2 光流的约束方程	(120)
7.3 微分技术	(121)
7.3.1 正则化方法	(121)
7.3.2 局部调整法	(122)
7.3.3 鲁棒法	(123)
7.3.4 高阶导数法	(124)
7.4 其他方法	(124)
7.4.1 区域匹配法	(125)
7.4.2 频域法	(125)
7.5 基于光流场的定性运动解释	(127)
7.5.1 光流的分类及其与三维运动的关系	(127)
7.5.2 汇集点	(130)
思考题	(131)
参考文献	(132)

第八章 长序列运动图像特征跟踪	(133)
8.1 引论	(133)
8.2 参数估计理论初步	(135)
8.2.1 状态空间表示法	(135)
8.2.2 参数估计	(135)
8.2.3 卡尔曼滤波	(137)
8.3 特征运动模型	(138)
8.3.1 多项式运动学	(139)
8.3.2 刚体运动学	(140)
8.4 特征跟踪的阐述	(142)
8.4.1 二维特征跟踪	(142)
8.4.2 三维特征跟踪	(143)
8.5 匹配	(145)
8.5.1 Mahalanobis 距离	(146)
8.5.2 匹配冲突的解决	(146)
8.6 实际应用中需要考虑的问题	(147)
8.6.1 运动模型的不适当	(147)
8.6.2 特征继续存在的支持	(148)
思考题.....	(149)
参考文献.....	(150)
第九章 基于二维特征对应的运动分析	(151)
9.1 极线方程和本质矩阵	(152)
9.1.1 极线方程	(152)
9.1.2 本质矩阵	(152)
9.2 基于点匹配的运动计算	(155)
9.2.1 求本质矩阵的算法	(155)
9.2.2 从本质矩阵求运动参数	(157)
9.2.3 非线性方法	(157)
9.2.4 鲁棒法	(159)
9.2.5 结构再投影法	(162)
9.2.6 运动多义性和临界曲面	(162)
9.3 当图像是一个空间平面的投影时的运动计算	(165)
9.4 基于直线匹配的运动计算	(169)
9.4.1 两幅图像不足以确定运动	(169)
9.4.2 从三幅图像求运动	(170)
9.4.3 几何解释	(171)
9.5 基本矩阵的估计	(171)
思考题.....	(171)
参考文献.....	(172)

第十章 基于三维特征对应的运动分析	(173)
10.1 由三维点匹配估计运动.....	(173)
10.1.1 问题陈述	(173)
10.1.2 直接利用旋转矩阵 R	(173)
10.1.3 利用四元数	(178)
10.1.4 利用欧拉角	(179)
10.1.5 利用旋转向量	(180)
10.1.6 三个点匹配时的特殊方法	(181)
10.1.7 鲁棒法	(183)
10.2 不需显式匹配的方法.....	(184)
10.2.1 共面点的运动	(184)
10.2.2 一般点群的运动	(184)
10.3 从三维直线匹配估计运动.....	(185)
10.3.1 三维直线的变换	(185)
10.3.2 解析解	(186)
10.3.3 数值方法	(187)
10.4 从平面匹配估计运动.....	(190)
10.4.1 问题的描述	(190)
10.4.2 解析解	(191)
10.4.3 卡尔曼滤波	(191)
10.5 二维-三维的物体定位	(191)
思考题.....	(192)
参考文献.....	(192)
第十一章 由图像灰度恢复三维物体形状	(194)
11.1 辐射度学与光度学.....	(194)
11.1.1 辐射度学	(194)
11.1.2 光度学	(196)
11.1.3 图像灰度与物体表面光辐射度的关系	(198)
11.2 光照模型.....	(199)
11.3 由多幅图像恢复三维物体形状.....	(203)
11.4 由单幅图像恢复三维物体形状.....	(205)
思考题.....	(207)
参考文献.....	(207)
第十二章 建模与识别	(209)
12.1 CAD 系统中的三维模型表达	(210)
12.2 曲线与曲面的表达.....	(211)
12.2.1 三次曲线与双三次曲面	(212)
12.2.2 广义圆柱面	(216)
12.2.3 超二次曲面	(217)

12.3	三维世界的多层次模型	(218)
12.4	由二维图像建模	(219)
12.4.1	矩不变量	(219)
12.4.2	变换群与不变量	(221)
12.5	识别的一般原则——问题与策略	(225)
12.6	特征关系图匹配	(227)
12.7	“假设检验”识别方法	(229)
12.7.1	识别的系统结构	(229)
12.7.2	三维空间中的坐标转换	(231)
12.7.3	由单幅图像中的几何元素进行物体定位	(231)
12.7.4	假设检验	(233)
	思考题	(234)
	参考文献	(235)
第十三章	距离图像获取与处理	(237)
13.1	距离传感器	(237)
13.1.1	基于雷达原理的距离传感器	(238)
13.1.2	基于三角原理的距离传感器	(240)
13.2	数据预处理	(242)
13.3	深度图分割	(243)
13.3.1	高斯球分割法	(244)
13.3.2	根据曲率特征分割	(245)
13.3.3	区域分割方法	(248)
	思考题	(249)
	参考文献	(249)
第十四章	计算机视觉系统体系结构讨论与展望	(250)
14.1	计算机视觉系统的基本体系结构	(250)
14.2	视觉系统体系结构讨论	(250)
14.2.1	视觉信息处理系统的“任务”	(250)
14.2.2	关于模块化问题	(251)
14.2.3	“注视”与“注意”(attention)问题	(252)
14.2.4	关于定性与定量分析	(252)
14.2.5	多种方法的融合	(253)
14.2.6	图像分割与知觉组织	(253)
14.2.7	局部特征与全局特征	(254)
14.2.8	物体建模	(255)
14.2.9	非线性、自适应、自学习与人工神经网络	(255)
14.2.10	交互作用问题	(256)
14.3	主动视觉	(257)

14.4 计算机视觉的应用展望.....	(258)
参考文献.....	(259)
附录 A 实验数据及参考结果.....	(261)
A.1 图像的格式	(261)
A.2 摄像机定标	(265)
A.3 立体视觉	(266)
A.4 基于光流场的运动分析	(269)
A.5 长序列运动图像特征跟踪	(271)
A.6 基于二维特征对应的运动分析	(276)
A.7 基于三维特征对应的运动分析	(276)

第一章 绪 论

人类通过视觉感知外界信息. 俗话说“百闻不如一见”, 就是说视觉感知环境信息的效率很高. 人类感知外界信息, 80%以上是通过视觉得到的. 让计算机或机器人具有视觉, 是人类多年以来的梦想. 虽然, 目前我们还不能让计算机也具有像生物那样高效、灵活的视觉, 但这种希望正在逐步实现.

人类是通过眼睛与大脑来获取、处理与理解视觉信息的. 周围环境中的物体在可见光的照射下, 在人眼的视网膜上形成图像, 由感光细胞转换成神经脉冲信号, 经神经纤维传入大脑皮层进行处理与理解. 视觉, 不仅指对光信号的感受, 它包括了对视觉信息的获取、传输、处理、存储与理解的全过程. 信号处理理论与计算机出现以后, 人们试图用摄像机获取环境图像并将其转换成数字信号, 用计算机实现对视觉信息处理的全过程, 这样, 就形成了一门新兴的学科——计算机视觉.

计算机视觉的研究目标是使计算机具有通过二维图像认知三维环境信息的能力. 这种能力将不仅使机器能感知三维环境中物体的几何信息, 包括它的形状、位置、姿态、运动等, 而且能对它们进行描述、存储、识别与理解. 计算机视觉发展得益于神经生理学、心理学与认知科学对动物视觉系统的研究, 但计算机视觉已发展起一套独立的计算理论与算法, 它并不刻意去“仿真”生物视觉系统, 其原因有三:

(1) 几十年来, 虽然神经生理学、心理学与认知科学对生物视觉系统从解剖学、电生理过程与信息处理等不同角度进行了大量卓越的研究, 但对视觉认知过程的认识还远远不够, 尤其是对大脑皮层中高层次视觉信息处理的认识还是十分肤浅的.

(2) 研究表明, 大脑神经细胞构成了一个极其复杂的, 巨大的互连网络, 这种神经网络在体系结构上与当前计算机体系结构有很大的差别. 虽然, 近年来人工神经网络的研究得到很大的重视(即用计算机或专用芯片去局部仿真神经网络), 但这种仿真也只能是近似的、局部的, 要实现大量人工神经元的互连网络在技术上目前还是很困难的. 另外, 不少计算机视觉工作者认为, 只要从信息转换的角度真正理解了视觉信息处理过程并发展出一套信息处理的计算理论, 用哪种体系结构去实现它是无关紧要的. 本书要讲述的内容就是多年来计算机视觉工作者对实现各种视觉功能发展出的一整套计算理论与算法. 近年来对神经网络的研究表明, 其中许多算法既可以用通用的串行的计算机体系结构实现, 也可以用神经网络实现.

(3) 在许多工业应用场合, 视觉系统要观察的环境常常比较简单. 60年代第一次做的计算机视觉系统的实验[Roberts 1965], 环境被限制在所谓“积木世界”, 即周围的物体都是由多面体组成的, 需要识别的物体可以用简单的点、直线、平面的组合表示. 计算机视觉的研究工作就是从对这些简单的物体识别与定位开始的. 另一方面, 用计算机视觉的方法还可以处理与识别更广泛的图像, 例如遥感图像、医学图像等, 这些图像的处理和识别与一般摄像机的图像有较大的不同, 处理与识别的方法也不同.