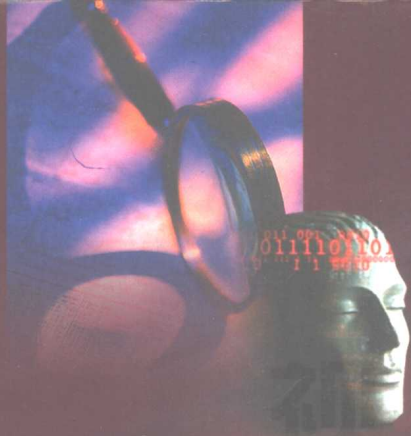


图象图形科学丛书



贾云得 编著

# 机器视觉

科学出版社



图象图形科学丛书

# 机 器 视 觉

贾云得 编著

科 学 出 版 社

2000

FF30/A

2000.11.15

王云得

## 内 容 简 介

本书旨在提供一本机器视觉(或计算机视觉)导论性教材,主要介绍机器视觉的基础理论、基本方法和实用算法,内容包括:二值图象分析、图象预处理、边缘检测、图象分割、纹理分析、明暗分析、彩色感知、深度图象与立体视觉、视觉系统标定、物体表示、二维运动分析、三维运动分析、物体识别等,另外还简要介绍了人类视觉的基本原理和 Marr 视觉计算理论。

与目前同类著作不同,本书摒弃最新研究现状和成果的展示以及复杂的数学推导,注重基本概念、基本原理的阐述,以及视觉基本算法的详尽介绍。本书内容几乎涉及到视觉的所有研究方向,但取材精炼,重点突出。

本书要求学生具有大学理工科微积分等数学知识,包括矩阵和线性空间;也要求学生具备基本概率理论、计算机编程等知识。本书适合作为大学高年级本科生和研究生的教程,也适合作为从事视觉领域研究和应用的科技工作者的导论性读物。

### 图书在版编目(CIP)数据

机器视觉/贾云得编著. -北京:科学出版社,2000

(图象图形科学丛书)

ISBN 7-03-007232-4

I. 机… II. 贾… III. 计算机视觉 IV. TP302.7

中国版本图书馆 CIP 数据核字(1999)第 34611 号

科学出版社 出版

北京东黄城根北街 16 号

邮政编码:100717

北京双青印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

\*

2000 年 4 月第 一 版 开本: 787×1092 1/16

2000 年 4 月第一次印刷 印张: 20 3/4

印数: 1—3 800 字数: 460 000

定价: 32.00 元

(如有印装质量问题, 我社负责调换(环伟))

# 图象图形科学丛书编委会名单

## 主任委员

潘云鹤

## 副主任委员

(按姓氏笔划为序)

王宝兴 王淑兰 罗志安 章毓晋

## 委 员

(按姓氏笔划为序)

王宝兴 王淑兰 刘健勤 朱迷龙 江 早  
石教英 何江华 鲍虎军 罗志安 张永生  
章毓晋 崔 屹 潘云鹤 潘志庚

## 丛书序言

图象图形是人类相互交流和认识客观世界的主要媒体。科学研究和统计表明，视觉系统帮助人类从外界获得 3/4 以上的信息，而图象图形带给我们的正是视觉世界中的所有信息。视觉信息所获得的客观作用是其他信息不能替代的，百闻不如一见就是一个非常形象的例子。图象图形是现代信息化社会的重要支柱。

图象图形科学是一门理论与现代高科技相结合来系统地研究各种视觉原理、技术和应用的综合性很强交叉学科。图象图形技术在广义上是各种与视觉有关技术的总称。人类基于视觉的活动，是一个广阔、复杂、富有挑战性的研究领域。图象图形科学和技术是这个领域的有力工具。该学科包括利用计算机和其他电子设备观察世界而获得的数据及按产生数据处理并且显示这些数据的理论和技术的研究。

图象图形科学具有涉及面广，内容丰富，跨行业、跨学科的特点。从它的研究方法来看，它与数学、物理学、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；从它的研究范围来看，它与人工智能、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；它的发展应用与医学、遥感、通信、影视、文档处理和工业自动化等许多领域也是不可分割的。在科学史上，它代表了最活跃和令人振奋的边缘学科之一。

图象图形技术已经迅速渗透到人类生活和社会发展的各个方面。图象图形技术无论在科学研究、工业生产、文化娱乐、管理部门都得到越来越多的重视。图象图形技术在工业检测、高空侦察、制导、文件处理、动画、虚拟现实、生物医学、人体科学、艺术、远程教育、科学可视化、计算机辅助设计、遥感、航天等方面都得到越来越多的应用。进入 21 世纪后，其发展将更加迅速。

“图象图形科学丛书”正是在这种形势下组织出版的。中国图象图形学会和科学出版社为该套书的出版付出了很多的努力。这套丛书比较全面地覆盖了图象图形科学的各个分支，是广泛了解图象图形领域基本理论、技术应用和发展动态的最好读物；也是从事图象图形领域研究、技术开发和实际应用人员的工具书。

“图象图形科学丛书”由我国该领域的专家编写，这些专家既对图象图形领域的发展有全面的把握，又分别在其中的某一方向上有深入的研究和独到的见解，充分反映了当前图象图形科技研究的前沿、进展和水平。希望该套书能为发展图象图形科学技术，活跃学术气氛，交流研究成果，促进科技发展，为迎接信息技术的挑战，为我国图象图形事业做出应有贡献。

潘雪鹤

2000 年 2 月 1 日

## 前 言

本书的编写可以追溯到 1996 年,当时我正在美国卡内基-梅隆大学(Carnegie Mellon University, CMU)计算机科学学院([www.cs.cmu.edu](http://www.cs.cmu.edu))进行客座研究. 我有幸得到 K. Ikeuchi 教授、M. Hebert 博士、E. Krothov 博士在 CMU 共同主讲的“计算机视觉”课程讲义,同时多次观摩了该课的教学过程,受益匪浅. 他们在教学大纲、授课内容、授课方式、课程作业、计算机实践等方面给我留下了非常深刻的印象,许多方面值得我们学习和借鉴,这也促使我决心为我国大学高年级本科生和研究生编写一本计算机视觉的导论性教材. 为了配合本教材的使用,我们将建立“机器视觉”课程教学网页,包括:测试图象库,图象函数程序库,计算机实践项目,视觉系统设计库等等.

本书的编写大纲和部分内容是在美国完成的. 为了使本书尽力体现计算机视觉研究领域的基本、全面、先进和实用的内容,除了参阅 CMU“计算机视觉”讲义外,还参阅了 Ballard 和 Brown 1982 年著《Computer Vision》,Horn 1986 年著《Robot Vision》,Faugeras 1993 年著《Three Dimensional Computer Vision》,Jian, Kasturi 和 Schunck 1995 年著《Computer Vision》,Tekalp 1995 年著《Digital Video Processing》,Castleman 1996 年著《Digital Image Processing》等几十部国内外有关计算机视觉的专著,以及近十几年来发表的 200 多篇优秀论文.

在本书的编写过程中,得到了许多人的大力支持和帮助. 我首先要感谢北京理工大学机电工程与控制国家级重点实验室给我提供的良好办公环境和研究环境,特别要感谢我国著名机电工程专家、北京理工大学教授马宝华先生对我的关心和支持.

我们计算机视觉研究组的刘万春和朱玉文两位老师,以及吴小平、吕宏静、李莉、冷树林、裴洪安、徐一华等十几位博士生、硕士生为本书提供了大量的编译材料、计算机程序和实验结果,并完成了本书的大部分绘图、文字录入、校对等工作. 可以说,本书是我们研究组集体的劳动结晶,在此我对他们的无私帮助表示衷心的感谢.

我还要感谢在美国期间的导师 K. Ikeuchi 教授和 M. Hebert 博士,以及 Fred Solomon, Harry Shum(沈向洋), Mark Wheeler, Yoichi Sato's, George Paul, Larry Zitnick, Richard Volpe, Yanbin Jia 等,还有当时任职 CMU 的徐阳生教授、杨杰博士. 他们在工作 and 生活上给予我很大的关心和帮助,许多学术构思也得益于与他们的讨论,以及他们给予的帮助.

我热切希望能得到各位读者对本书提出的宝贵意见.

我的 E-Mail 地址是: [yjjar @ bit.edu.cn](mailto:yjjar@bit.edu.cn).

贾云得

1999 年 8 月于北京

# 目 录

<b>第一章 引论</b> .....	1
1.1 机器视觉 .....	1
1.2 Marr 视觉计算理论 .....	2
1.2.1 三个层次 .....	2
1.2.2 视觉表示框架 .....	3
1.3 机器视觉的应用 .....	4
1.4 机器视觉的研究内容与面临的困难 .....	6
1.4.1 机器视觉的研究内容 .....	6
1.4.2 机器视觉面临的困难 .....	7
1.5 机器视觉与其它学科领域的关系 .....	9
1.6 成象几何基础 .....	10
1.6.1 透视投影 .....	10
1.6.2 正交投影 .....	11
1.6.3 视觉系统坐标系 .....	12
1.7 本书内容向导 .....	13
思考题 .....	14
<b>第二章 人类视觉</b> .....	15
2.1 人类视觉简介 .....	15
2.1.1 人眼的构造 .....	16
2.1.2 视觉通路 .....	17
2.1.3 视觉信息处理过程 .....	18
2.2 感受野 .....	20
2.3 视觉信息的多层次并行处理 .....	22
2.3.1 视觉信息的多层串行处理机制 .....	22
2.3.2 视觉信息的并行处理 .....	23
2.4 视觉信息的集成和反馈 .....	24
思考题 .....	25
<b>第三章 二值图象分析</b> .....	26
3.1 阈值 .....	27
3.2 几何特性 .....	28
3.2.1 尺寸和位置 .....	28
3.2.2 方向 .....	29
3.2.3 密集度和体态比 .....	30
3.3 投影 .....	31

3.4	游程长度编码 .....	32
3.5	二值图象算法 .....	33
3.5.1	定义 .....	33
3.5.2	连通成分标记 .....	36
3.5.3	欧拉数 .....	37
3.5.4	区域边界 .....	38
3.5.5	距离测量 .....	39
3.5.6	中轴 .....	39
3.5.7	细化 .....	40
3.5.8	扩展与收缩 .....	42
3.6	形态算子 .....	44
	思考题 .....	47
	计算机练习题 .....	47
<b>第四章</b>	<b>区域分析 .....</b>	<b>48</b>
4.1	区域和边缘 .....	48
4.2	分割 .....	49
4.2.1	自动阈值化法 .....	49
4.2.2	直方图方法的局限性 .....	53
4.3	区域表示 .....	53
4.3.1	阵列表示 .....	53
4.3.2	层级表示 .....	54
4.3.3	基于特征的区域表示 .....	54
4.3.4	图象分割数据结构 .....	56
4.4	分裂和合并 .....	58
4.4.1	区域合并 .....	58
4.4.2	区域分裂 .....	60
4.4.3	分裂和合并 .....	61
4.5	区域增长 .....	61
	思考题 .....	63
	计算机练习题 .....	64
<b>第五章</b>	<b>图象预处理 .....</b>	<b>65</b>
5.1	直方图修正 .....	65
5.2	图象线性运算 .....	67
5.2.1	线性系统 .....	67
5.2.2	傅里叶变换 .....	69
5.3	线性滤波器 .....	69
5.3.1	均值滤波器 .....	70
5.3.2	高斯平滑滤波器 .....	71
5.4	非线性滤波器 .....	78
5.4.1	中值滤波器 .....	78



5.4.2 边缘保持滤波器 .....	79
思考题 .....	81
计算机练习题 .....	81
<b>第六章 边缘检测</b> .....	<b>82</b>
6.1 梯度 .....	84
6.2 边缘检测算法 .....	85
6.2.1 边缘检测算法的基本步骤 .....	85
6.2.2 Roberts 算子 .....	85
6.2.3 Sobel 算子 .....	86
6.2.4 Prewitt 算子 .....	86
6.2.5 各种算法比较 .....	86
6.3 二阶微分算子 .....	89
6.3.1 拉普拉斯算子 .....	90
6.3.2 二阶方向导数 .....	91
6.4 LoG 算法 .....	92
6.5 图象逼近 .....	94
6.6 Canny 边缘检测器 .....	97
6.7 子像素级位置估计 .....	100
6.8 边缘检测器性能 .....	101
6.8.1 性能评价方法 .....	102
6.8.2 品质因数 .....	102
6.9 线条检测 .....	103
思考题 .....	103
计算机练习题 .....	104
<b>第七章 轮廓表示</b> .....	<b>105</b>
7.1 数字曲线及其表示 .....	105
7.1.1 链码 .....	106
7.1.2 斜率表示法 .....	107
7.2 曲线拟合 .....	108
7.2.1 多直线段 .....	109
7.2.2 二次曲线 .....	112
7.3 样条曲线 .....	115
7.3.1 三次样条曲线 .....	115
7.3.2 B 样条曲线 .....	117
7.4 曲线回归逼近 .....	118
7.4.1 全回归方法 .....	118
7.4.2 角点估计 .....	119
7.4.3 鲁棒回归法 .....	120
7.5 Hough 变换 .....	122
7.6 傅里叶描述子 .....	126

思考题 .....	126
计算机练习题 .....	127
<b>第八章 纹理 .....</b>	<b>129</b>
8.1 概述 .....	129
8.2 纹理分析统计方法 .....	130
8.2.1 灰度级同现矩阵 .....	130
8.2.2 自相关法 .....	132
8.3 有序纹理的结构分析 .....	133
8.4 基于模型的纹理分析 .....	134
8.5 用分形理论分析纹理 .....	134
8.6 从纹理恢复形状 .....	135
思考题 .....	138
计算机练习题 .....	138
<b>第九章 明暗分析 .....</b>	<b>139</b>
9.1 图象辐射度 .....	139
9.1.1 照明 .....	140
9.1.2 反射 .....	141
9.2 表面方向 .....	143
9.3 反射图 .....	144
9.4 从图象明暗恢复形状 .....	145
9.5 光度立体 .....	146
思考题 .....	149
计算机练习题 .....	149
<b>第十章 彩色感知 .....</b>	<b>150</b>
10.1 三色原理 .....	150
10.2 颜色模型 .....	151
10.2.1 RGB 和 CMY 颜色模型 .....	151
10.2.2 CIE - XYZ 颜色模型 .....	153
10.2.3 NTSC - YIQ 彩色模型 .....	154
10.3 颜色的视觉处理 .....	155
10.3.1 颜色的三个基本属性 .....	155
10.3.2 HSV 颜色感知模型 .....	155
10.4 彩色不变性 .....	157
10.5 讨论 .....	158
思考题 .....	159
计算机练习题 .....	159
<b>第十一章 深度图 .....</b>	<b>160</b>
11.1 立体成象 .....	160
11.2 立体匹配 .....	162

11.2.1	基本约束 .....	163
11.2.2	边缘匹配 .....	164
11.2.3	区域相关性 .....	164
11.2.4	讨论 .....	166
11.3	多基线立体成象 .....	166
11.4	从 X 恢复形状的方法 .....	170
11.5	测距成象 .....	171
11.5.1	结构光测距 .....	172
11.5.2	激光测距雷达 .....	174
11.5.3	变焦测距 .....	176
11.5.4	激光阵列成象 .....	178
11.5.5	各种方法比较 .....	178
11.6	主动视觉 .....	179
	思考题 .....	180
	计算机练习题 .....	180
<b>第十二章</b>	<b>标定 .....</b>	<b>181</b>
12.1	刚体变换 .....	181
12.1.1	旋转矩阵 .....	182
12.1.2	旋转轴 .....	183
12.1.3	四元数 .....	183
12.2	绝对定位 .....	184
12.3	相对定位 .....	187
12.4	校正 .....	188
12.5	双目立体深度测量 .....	190
12.6	含有比例因子的绝对定位 .....	191
12.7	外部定位 .....	192
12.8	内部定位 .....	194
12.9	摄像机标定 .....	197
12.9.1	摄像机标定的基本方法 .....	197
12.9.2	仿射变换法 .....	200
12.9.3	非线性方法 .....	201
12.10	双目立体标定 .....	202
12.11	主动三角测距标定 .....	203
12.12	鲁棒方法 .....	204
12.13	讨论 .....	205
	思考题 .....	205
	计算机练习题 .....	206
<b>第十三章</b>	<b>三维场景表示 .....</b>	<b>207</b>
13.1	三维空间曲线 .....	207
13.1.1	三次样条曲线 .....	207

13.1.2 三维 $B$ 样条曲线 .....	208
13.2 三维空间曲面的表示 .....	208
13.2.1 多边形网面 .....	208
13.2.2 曲面片 .....	211
13.2.3 张量积曲面 .....	211
13.2.4 超二次曲面 .....	212
13.3 曲面插值 .....	213
13.3.1 三角形面插值 .....	213
13.3.2 二元线性插值 .....	214
13.3.3 鲁棒插值 .....	215
13.4 曲面逼近 .....	215
13.4.1 回归样条 .....	216
13.4.2 加权样条逼近 .....	220
13.5 曲面分割 .....	221
13.5.1 初始分割 .....	221
13.5.2 曲面片增长 .....	222
13.6 曲面配准 .....	223
思考题 .....	225
计算机练习题 .....	225
<b>第十四章 二维运动估计</b> .....	<b>226</b>
14.1 图象运动特征检测 .....	226
14.1.1 差分图象 .....	226
14.1.2 时变边缘检测 .....	230
14.1.3 运动对应性 .....	230
14.2 光流法 .....	235
14.2.1 运动场与光流 .....	235
14.2.2 光流约束方程 .....	237
14.3 光流计算 .....	238
14.3.1 Horn - Schunck 方法 .....	238
14.3.2 Lucas - Kanade 方法 .....	239
14.3.3 Nagel 方法 .....	240
14.3.4 鲁棒计算方法 .....	242
14.4 基于块的运动分析 .....	244
14.4.1 块运动模型 .....	244
14.4.2 傅里叶方法 .....	245
14.4.3 相位相关法 .....	246
14.4.4 块匹配方法 .....	247
14.4.5 广义图象运动估计 .....	249
思考题 .....	251
计算机练习题 .....	251
<b>第十五章 三维运动估计</b> .....	<b>252</b>
15.1 基于成像模型的对应点估计 .....	252
15.1.1 正交投影模型 .....	252

15.1.2	基于正交投影的运动估计 .....	253
15.1.3	透视投影模型 .....	254
15.1.4	外极线方程和基本矩阵 .....	255
15.1.5	从基本矩阵估计运动参数 .....	257
15.1.6	从外极线方程估计运动参数 .....	258
15.2	三维运动估计光流法 .....	260
15.2.1	速度场正交投影模型与仿射流 .....	260
15.2.2	速度场透视投影模型与二次流 .....	261
15.2.3	延伸焦点 .....	262
15.2.4	代数法 .....	264
15.2.5	优化法 .....	266
15.2.6	几种方法比较 .....	266
15.3	光流分割 .....	268
15.3.1	Hough 变换分割方法 .....	268
15.3.2	贝叶斯分割方法 .....	269
	思考题 .....	272
	计算机练习题 .....	272
<b>第十六章</b>	<b>物体识别</b> .....	<b>273</b>
16.1	识别系统的基本组成 .....	273
16.2	物体识别的复杂度 .....	275
16.3	图象矩不变量特征表示 .....	276
16.4	三维物体模型表示 .....	277
16.4.1	多视图表示 .....	278
16.4.2	结构立体几何表示 .....	278
16.4.3	体积表示 .....	279
16.4.4	扫掠表示 .....	280
16.4.5	函数表示 .....	281
16.4.6	三角形网面表示 .....	281
16.5	特征检测与识别策略 .....	282
16.5.1	特征检测 .....	282
16.5.2	特征分类 .....	284
16.5.3	特征匹配 .....	286
16.5.4	特征标记 .....	287
16.6	验证 .....	288
16.6.1	模板匹配 .....	288
16.6.2	形态方法 .....	289
16.6.3	符号 .....	290
16.6.4	类比法 .....	291
16.7	物体定位 .....	291
16.7.1	三维 - 三维物体定位 .....	292
16.7.2	二维 - 二维物体定位 .....	295
	思考题 .....	295
	计算机练习题 .....	296
	参考文献 .....	297
	英汉词汇对照表 .....	309

# 第一章 引 论

## 1.1 机器视觉

人类在征服自然、改造自然和推动社会进步的过程中,面临着自身能力、能量的局限性,因而发明和创造了许多机器来辅助或代替人类完成任务.智能机器,包括智能机器人,是这种机器最理想的形式,也是人类科学研究中所面临的最大挑战之一.智能机器是指这样一种系统,它能模拟人类的功能,能感知外部世界并有效地解决人所能解决的问题.人类感知外部世界主要是通过视觉、触觉、听觉和嗅觉等感觉器官,其中约 80%的信息是由视觉获取的.因此,对于智能机器来说,赋予机器以人类视觉功能对发展智能机器是极其重要的,也由此形成了一门新的学科——机器视觉(也称计算机视觉或图象分析与理解等).机器视觉的发展不仅将大大推动智能系统的发展,也将拓宽计算机与各种智能机器的研究范围和应用领域.

机器视觉是研究用计算机来模拟生物外显或宏观视觉功能的科学和技术.机器视觉系统的首要目标是用图象创建或恢复现实世界模型,然后认知现实世界.机器视觉系统获取的场景图象一般是灰度图象,即三维场景在二维平面上的投影.此时,场景三维信息只能通过灰度图象或灰度图象序列来恢复处理,这种恢复需要进行多点对一点的映射逆变换.在信息恢复过程中,还需要有关的场景知识和投影几何知识.

机器视觉是一个相当新且发展十分迅速的研究领域,并成为计算机科学的重要研究领域之一.机器视觉是在 20 世纪 50 年代从统计模式识别开始的,当时的工作主要集中在二维图象分析和识别上,如光学字符识别、工件表面、显微图片和航空图片的分析和解释等.60 年代,Roberts(1965)通过计算机程序从数字图象中提取出诸如立方体、楔形体、棱柱体等多面体的三维结构,并对物体形状及物体的空间关系进行描述[Roberts 1965].Roberts 的研究工作开创了以理解三维场景为目的的三维机器视觉的研究. Roberts 对积木世界的创造性研究给人们以极大的启发,许多人相信,一旦由白色积木玩具组成的三维世界可以被理解,则可以推广到理解更复杂的三维场景.于是,人们对积木世界进行了深入的研究,研究的范围从边缘、角点等特征提取,到线条、平面、曲面等几何要素分析,一直到图象明暗、纹理、运动以及成象几何等,并建立了各种数据结构和推理规则.到了 70 年代,已经出现了一些视觉应用系统[Guzman 1969, Mackworth 1973].

70 年代中期,麻省理工学院(MIT)人工智能(AI)实验室正式开设“机器视觉”(Machine Vision)课程,由国际著名学者 B. K. P. Horn 教授讲授.同时,MIT AI 实验室吸引了国际上许多知名学者参与机器视觉的理论、算法、系统设计的研究,David Marr 教授就是其中的一位.他于 1973 年应邀在 MIT AI 实验室领导一个以博士生为主体的研究小组,1977 年提出了不同于“积木世界”分析方法的计算视觉(computational vision)理论,该理论在 80 年代成为机器视觉研究领域中的一个十分重要的理论框架[Marr 1982].

可以说,对机器视觉的全球性研究热潮是从 20 世纪 80 年代开始的,到了 80 年代中期,机器视觉获得了蓬勃发展,新概念、新方法、新理论不断涌现,比如,基于感知特征群的物体识别理论框架、主动视觉理论框架、视觉集成理论框架等。

到目前为止,机器视觉仍然是一个非常活跃的研究领域.许多会议论文集都反映了该领域的最新进展,比如,国际计算机视觉与模式识别会议(International Conference on Computer Vision and Pattern Recognition, CVPR);国际计算机视觉会议(International Conference on Computer Vision, ICCV);国际模式识别会议(International Conference on Pattern Recognition, ICPR);国际机器人学与自动化会议(International Conference on Robotics and Automation, ICRA);计算机视觉研讨会(Workshop on Computer Vision, WCV)以及许多 SPIE 的会议.还有许多学术期刊也包含了这一领域的最新研究成果,如,IEEE Transaction on Pattern Analysis and Machine Intelligence(PAMI); Computer Vision, Graphics, and Image Processing(CVGIP); IEEE Transaction on Image Processing; IEEE Transaction on Systems, Man, and Cybernetics (SMC); Machine Vision and Applications; International Journal on Computer Vision(IJCV); Image and Vision Computing; Pattern Recognition. 另外,每年还出版许多研究专集、学术著作、技术报告,专题讨论会文集等.所有这些都是研究机器视觉及其应用的很好信息来源。

## 1.2 Marr 视觉计算理论

Marr 视觉计算理论立足于计算机科学,系统地概括了心理生理学、神经生理学等方面业已取得的所有重要成果,是视觉研究中迄今为止最为完善的视觉理论.Marr 建立的视觉计算理论,使计算机视觉研究有了一个比较明确的体系,并大大推动了计算机视觉研究的发展.人们普遍认为,计算机视觉这门学科的形成与Marr的视觉理论有着密切的关系.事实上,尽管 20 世纪 70 年代初期就有人使用计算机视觉这个名词[Binford 1971],但正是 Marr 于 70 年代末建立的视觉理论促使计算机视觉这一名词的流行.下面简要地介绍 Marr 视觉理论的基本思想及其理论框架。

### 1.2.1 三个层次

Marr 认为,视觉是一个信息处理系统,对此系统研究应分为三个层次:计算理论层次,表示(representation)与算法层次,硬件实现层次,如表 1-1 所示。

表 1-1

计算理论	表示和算法	硬件实现
计算的目的是什么	如何实现这个计算理论	在物理上如何实现
为什么这一计算是合适的	输入、输出的表示是什么?	这些表示和算法?
执行计算的策略是什么	表示与表示之间的变换是什么?	

按照 Marr 的理论,计算视觉理论要回答视觉系统的计算目的和策略是什么,或视觉系统的输入和输出是什么,如何由系统的输入求出系统的输出.在这个层次上,信息系统的特征是将一种信息(输入)映射为另一种信息(输出).比如,系统输入是二维灰度图象,

输出则是三维物体的形状、位置和姿态,视觉系统的任务就是研究如何建立输入输出之间的关系和约束,如何由二维灰度图象恢复物体的三维信息。表示与算法层次是要进一步回答如何表示输入和输出信息,如何实现计算理论所对应的功能的算法,以及如何由一种表示变换成另一种表示,比如创建数据结构和符号。一般来说,不同的输入、输出和计算理论,对应不同的表示,而同一种输入、输出或计算理论可能对应若干种表示。在解决了理论问题和表示问题后,最后一个层次是解决用硬件实现上述表示和算法的问题,比如计算机体系结构及具体的计算装置及其细节。从信息处理的观点来看,至关重要的乃是最高层次,即计算理论层次。这是因为构成知觉的计算本质,取决于解决计算问题本身,而不取决于用来解决计算问题的特殊硬件。换句话说,通过正确理解待解决问题的本质,将有助于理解并创造算法。如果考虑解决问题的机制和物理实现,则对理解算法往往无济于事。

上述三个层次之间存在着逻辑的因果关系,但它们之间的联系不是十分紧密,因此,某些现象只能在其中一个或两个层次上进行解释。比如神经解剖学原则上与第三层次(即物理实现)联系在一起。突触机制、动作电位、抑制性相互作用都在第三个层次上。心理物理学与第二层次(即表示与算法)有着更直接的联系。更一般地说,不同的现象必须在不同的层次上进行解释,这会有助于人们把握正确的研究方向。例如,人们常说,人脑完全不同于计算机,因为前者是并行加工的,后者是串行的。对于这个问题,应该这样回答:并行加工和串行加工是在算法这个层次上的区别,而不是根本性的区别,因为任何一个并行的计算程序都可以写成串行的程序。因此,这种并行与串行的区别并不支持这种观点,即人脑的运行与计算机的运算是不同的,因而人脑所完成的任务是不可能通过编制程序用计算机来完成。

### 1.2.2 视觉表示框架

视觉过程划分为三个阶段,如表 1-2 所示。第一阶段(也称为早期阶段)是将输入的原始图象进行处理,抽取图象中诸如角点、边缘、纹理、线条、边界等基本特征,这些特征的集合称为基元图(primitivesketch);第二阶段(中期阶段)是指在以观测者为中心的坐标系

表 1-2 由图象恢复形状信息的表示框架

名 称	目 的	基 元
图象	亮度表示	图象中每一点的亮度值
基元图	表示二维图象中的重要信息,主要是图象中的亮度变化位置及其几何分布和组织结构	零交叉,斑点,端点和不连续点,边缘,有效线段,组合群,曲线组织,边界
2.5 维图	在以观测者为中心的坐标系中,表示可见表面的方向、深度值和不连续的轮廓	局部表面朝向("针"基元) 离观测者的距离 深度上的不连续点 表面朝向的不连续点
3 维模型表示	在以物体为中心的坐标系中,用由体积基元和面积基元构成的模块化多层次表示,描述形状及其空间组织形式。	分层次组成若干三维模型,每个三维模型都是在几个轴线空间的基础上构成的,所有体积基元或面积形状基元都附着在轴线上。



中,由输入图象和基元图恢复场景可见部分的深度、法线方向、轮廓等,这些信息包含了深度信息,但不是真正的物体三维表示,因此,称为二维半图(2.5 dimensional sketch);在以物体为中心的坐标系中,由输入图象、基元图、二维半图来恢复、表示和识别三维物体的过程称为视觉的第三阶段(后期阶段)。

Marr 理论是计算机视觉研究领域的划时代成就,但该理论不是十分完善的,许多方面还有争议.比如,该理论所建立的视觉处理框架基本上是自下而上,没有反馈.还有,该理论没有足够地重视知识的应用.尽管如此,Marr 理论给了我们研究计算机视觉许多珍贵的哲学思想和研究方法,同时也给计算机视觉研究领域创造了许多研究起点.

### 1.3 机器视觉的应用

机器视觉技术正广泛地应用于各个方面,从医学图象到遥感图象,从工业检测到文件处理,从毫微米技术到多媒体数据库,不一而足.可以说,需要人类视觉的场合几乎都需要机器视觉.应该指出的是,许多人类视觉无法感知的场合,如精确定量感知、危险场景感知、不可见物体感知等,机器视觉更突显其优越性.下面是一些机器视觉的典型应用.

#### 1.1 零件识别与定位

由于工业环境的结构、照明等因素可以得到严格的控制,因此,机器视觉在工业生产和装配中得到了成功的应用.图 1.1 是一个具有简单视觉的工业机器人系统示意图,其视觉系统由一个摄像机和相关的视觉信息处理系统组成.摄像机位于零件传输带上方,对于不同的零件,可以选择不同颜色的传输带,比如,明亮的物体,选择黑色传输带,暗色的零件,选择白色的背景,这样有利于视觉系统将零件从传输带上分离出来,并进行识别和定位,识别的目的是为机器人提供是否操作或进行何种操作的信息,定位的目的是导引机器

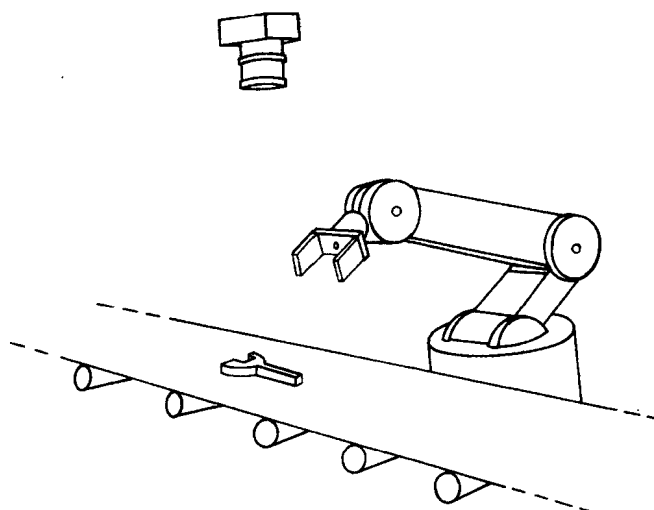


图 1.1 用于生产线上具有简单视觉系统的工业机器人系统示意图