

机器视觉理论及应用

赵 鹏◎著



电子工业出版社

PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

<http://www.phei.com.cn>

机器视觉理论及应用

赵 鹏 著

電子工業出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

本书涉及了机器视觉中的主要研究内容,各章节按照由视觉模型到视觉系统、由视觉算法到视觉应用编排,体现从简到繁、从浅到深、从理论到实际的特点,力求具有层次性、系统性、先进性和实用性。本书的前4章主要介绍了机器视觉的基本概念、基本几何变换、边缘检测及摄像机标定,这4章内容是从事机器视觉研究应该具备的基础知识。第5章与第6章侧重于机器视觉计算领域,涉及了图像融合和视觉模型融合与跟踪这两个热点领域。第7章与第8章侧重于机器视觉精密测量领域,具体讨论了基于视觉的几何量精密测量和运动物体速度精密测量这两个实际工程应用问题。此外,本书第9章和第10章还介绍了机器视觉在林业工程及农业工程领域的两个典型成功的应用范例。第11章介绍了光笔式3D坐标测量系统及其应用实例。

本书可作为光学工程、自动化及仪器仪表、电子信息工程等专业的低年级本科生和研究生的教材及参考书,也可以供从事机器视觉相关研究工作的研究人员和工程技术人员参考。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有,侵权必究。

图书在版编目(CIP)数据

机器视觉理论及应用 / 赵鹏著. —北京:电子工业出版社,2011.12
ISBN 978-7-121-15312-9

I. ①机… II. ①赵… III. ①计算机视觉 IV. ①TP302.7

中国版本图书馆CIP数据核字(2011)第244068号

策划编辑:秦绪军 赵娜

责任编辑:徐云鹏 文字编辑:韩奇桅

印刷:北京天宇星印刷厂

装订:三河市皇庄路通装订厂

出版发行:电子工业出版社

北京市海淀区万寿路173信箱 邮编 100036

开本:787×1092 1/16 印张:15.75 字数:403.2千字 彩插:2

印次:2011年12月第1次印刷

印数:2000册 定价:49.80元



凡所购买电子工业出版社图书有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系,联系及邮购电话:(010) 88254888。

质量投诉请发邮件至 zlt@s-phci.com.cn, 盗版侵权举报请发邮件至 dbqq@phci.com.cn。

服务热线:(010) 88258888。

前 言

随着信号处理技术、计算机应用技术、精密测量技术的快速发展，机器视觉得到了广泛的重视与应用。它具有非接触、实时性、可视化好、自动化和智能性高等优点，在科学研究和工程实践中具有广泛应用前景，逐渐受到研究人员的高度重视。

机器视觉是学科交叉研究领域，研究内容和研究方法涉及光学工程、图像处理、模式识别、信号处理、人工智能、仪器仪表、机器人学及自动化等诸多学科领域。本书按照机器视觉所涉及的基础理论及最新应用成果来安排章节，比较系统地阐述了机器视觉的基础理论、关键技术及最新应用实例。与同类书籍相比，本书既比较系统地论述了机器视觉理论及算法，又详细介绍了机器视觉最新的工程应用实例。

本书涉及了机器视觉中的主要研究内容，各章节按照由视觉模型到视觉系统、由视觉算法到视觉应用、由单传感器视觉到多传感器视觉的顺序来安排，体现从简到繁、从浅到深、从理论到实际的特点，力求具有层次性、系统性、先进性和实用性。本书的前4章主要介绍了机器视觉的基本概念、基本几何变换、边缘检测及摄像机标定，这4章内容是从事机器视觉研究应该具备的基础知识。鉴于国内外已经有一些关于机器视觉的专著，并且这些专著一般都对这些基础内容进行了论述，本书未对这些基础内容进行详细论述，因此读者可以参考有关专著或教材详细学习相关基础内容。第5章~第11章是本书的重点内容，这几章包括了作者本人近几年的全面系统的研究成果，它们反映了近10年来机器视觉领域的一些最新发展动向和热点领域。其中的第5章与第6章侧重于机器视觉计算领域，涉及了图像融合和视觉模型融合与跟踪这两个热点领域。第7章与第8章侧重于机器视觉精密测量领域，具体讨论了基于视觉的几何量精密测量和运动物体速度精密测量这两个实际工程应用问题。第9章和第10章介绍了机器视觉在农业工程和林业工程中的两个典型应用范例，即杂草分类识别和木材材种分类识别。第11章介绍了光笔式3D坐标视觉测量系统及其在几何量测量中的应用。

作者近年来在这些领域进行了深入系统的研究，取得了一些重要研究成果。在相应的章节中，作者对这些研究成果都进行了详细论述。因此，和机器视觉同类书籍比较，本书具有较好的新颖性、先进性和实用性。这些研究得到了作者主持的科研项目的资助，具体包括1项哈尔滨市青年科技创新项目、1项中央高校基本业务费专项基金、1项科技部北京理工大学973计划项目横向课题、1项黑龙江省自然科学基金面上项目、1项东北林业大学青年拔尖人才计划基金及1项黑龙江省博士后科研启动金。作者对批准这些研究基金和研究计划的相关行政部门表示真诚感谢。本书在编写过程中，得到了有关专家和同事的帮助，这里作者表示感谢。由于作者本人研究水平有限，经验不足，敬请读者批评指正。

目 录

第 1 章 引论	1
1.1 机器视觉的发展及系统构成	1
1.1.1 机器视觉的发展	1
1.1.2 机器视觉系统构成	2
1.2 Marr 的视觉理论框架	4
1.2.1 视觉系统研究的三个层次	4
1.2.2 视觉信息处理的三个阶段	5
1.3 机器视觉的应用领域及面临的问题	7
1.3.1 机器视觉的应用领域	7
1.3.2 机器视觉面临的问题	8
本章参考文献	9
第 2 章 空间几何变换与摄像机模型	11
2.1 空间几何变换	11
2.1.1 齐次坐标	11
2.1.2 射影变换	12
2.1.3 仿射变换	13
2.1.4 比例变换	14
2.1.5 欧氏变换	14
2.2 几何变换的不变量	15
2.2.1 简比与交比	15
2.2.2 不变量	16
2.3 欧氏空间的刚体变换	17
2.3.1 刚体变换过程	18
2.3.2 旋转矩阵的表示形式	18
2.4 摄像机透视投影模型	20
2.4.1 图像坐标系、摄像机坐标系与世界坐标系	20
2.4.2 针孔成像模型	22
2.4.3 非线性模型	23
2.5 摄像机透视投影近似模型	23
2.5.1 正投影	23
2.5.2 弱透视	24

2.5.3 平行透视	25
2.5.4 仿射摄像机	26
本章参考文献	27
第3章 视觉图像特征信息提取	28
3.1 图像边缘与图像平滑	28
3.1.1 图像边缘	28
3.1.2 图像平滑滤波	29
3.2 一阶微分边缘检测算子	31
3.2.1 梯度	32
3.2.2 边缘算子	32
3.3 二阶微分边缘检测算子	36
3.3.1 拉普拉斯算子	36
3.3.2 LoG 算子	37
3.4 子像素级边缘检测	39
3.4.1 Hessian 矩阵法	40
3.4.2 曲面拟合法	42
3.5 角点探测器	43
3.5.1 K-R (Kitchen - Rosenfeld) 法	43
3.5.2 图像灰度法	44
3.5.3 简单算法	45
3.6 形状特征分析	46
3.6.1 矩形度	46
3.6.2 球状性	46
3.6.3 圆形性	47
3.6.4 中心矩	47
3.6.5 长轴	48
3.7 椭圆孔图像中心的提取	48
3.7.1 重心法	49
3.7.2 椭圆拟合法	49
3.8 给定形状曲线的检测——Hough 变换	50
3.8.1 简单形状曲线的检测	50
3.8.2 复杂形状曲线的检测	52
本章参考文献	53
第4章 摄像机标定与双目立体视觉	55
4.1 非线性优化方法	55
4.1.1 非线性优化目标函数	55
4.1.2 最小二乘法	56

4.1.3 Levenberg - Marquardt 算法	60
4.1.4 罚函数法	61
4.2 基于3D立体靶标的摄像机标定	61
4.2.1 线性模型摄像机标定	61
4.2.2 非线性模型摄像机标定	64
4.3 双目立体视觉原理	64
4.3.1 双目立体视觉三维测量原理	64
4.3.2 双目立体视觉数学模型	65
4.4 双目立体视觉中的对应点匹配	68
4.4.1 图像匹配的常用方法	68
4.4.2 已知极线几何的对应点匹配方法	69
4.4.3 未知极线几何的对应点匹配方法	70
本章参考文献	75
第5章 视觉计算与融合理论	76
5.1 融合处理中需考虑的问题	76
5.2 融合方法概论	78
5.2.1 信号级融合方法	78
5.2.2 像素级融合方法	79
5.2.3 特征级融合方法	86
5.2.4 决策级融合方法	87
5.3 模式识别与信息融合	88
5.3.1 视觉信息融合	89
5.3.2 视觉与触觉融合	89
5.3.3 视觉与红外信息融合	90
5.3.4 自动目标识别	90
5.3.5 移动机器人的感知信息融合	91
本章参考文献	92
第6章 基于图像融合的变形轮廓线	95
6.1 变形轮廓线概述	95
6.1.1 主动轮廓线	95
6.1.2 动态轮廓线	96
6.2 B样条动态轮廓线	96
6.3 基于图像融合和微分耦合机制的动态轮廓线	103
6.3.1 基于Lagrangian动力学原理的微分耦合的动态轮廓线	103
6.3.2 微分耦合机制的局限性	106
6.3.3 基于图像融合和微分耦合的动态轮廓线	106
6.3.4 实验结果	107

6.4	基于图像融合和 B 样条曲线范数极小化的动态轮廓线	109
6.4.1	B 样条形状空间	110
6.4.2	自适应 Kalman 滤波	110
6.4.3	基于 B 样条曲线范数极小化的图像融合	111
6.4.4	实验结果	112
6.5	基于图像融合的运动目标轮廓提取	114
6.5.1	运动目标分割	114
6.5.2	特征级融合	116
6.5.3	实验结果	116
6.6	基于图像融合的多分辨率动态轮廓线	117
6.6.1	多分辨率图像像素级融合	118
6.6.2	基于多分辨率对比度分解的图像融合	120
6.6.3	基于图像融合的多分辨率动态轮廓线	122
6.6.4	动态轮廓线与目标运动速度的关系	123
6.6.5	实验结果	125
6.7	基于图像融合和形状约束机制的主动轮廓线	128
6.7.1	基于能量极小化原理的参数型主动轮廓线	128
6.7.2	形状约束主动轮廓线及其局限性	128
6.7.3	基于图像融合和形状约束的主动轮廓线	129
6.7.4	实验结果	129
	本章参考文献	130
第 7 章	基于变形轮廓线的微小物体表面积周长精密测量	133
7.1	二维图像测量机研究现状	133
7.2	二维图像测量机	134
7.2.1	总体结构与原理	134
7.2.2	图像式自动调焦瞄准系统	134
7.2.3	调焦评价函数的确定	135
7.2.4	二维图像测量机的工作方式	137
7.3	基于动态轮廓线的物体表面积周长测量	137
7.3.1	图像目标的像素面积计算概述	137
7.3.2	基于动态轮廓线的物体表面积计算方法	138
7.3.3	基于动态轮廓线的物体周长计算方法	139
7.4	像素尺寸当量的标定	139
7.4.1	概述	139
7.4.2	形心自标定技术	140
7.5	基于多分辨率动态轮廓线的物体表面积周长测量	141
7.5.1	多分辨率动态轮廓线	141
7.5.2	实验结果	142

7.6 基于短程线主动轮廓线的多物体面积周长并行测量研究	144
7.6.1 短程线主动轮廓线概述	144
7.6.2 基于短程线主动轮廓线的像素面积及周长的计算方法	145
7.6.3 实验结果	146
本章参考文献	148
第8章 运动模糊图像恢复及其在运动物体速度测量中的应用	150
8.1 运动模糊图像恢复综述	150
8.2 应用区域划分法进行空间可变运动模糊图像恢复研究	151
8.2.1 奇偶场图像提取及运动偏移量的计算	153
8.2.2 区域空间不变运动模糊图像恢复	153
8.2.3 实验结果	154
8.3 采用运动模糊图像信息进行物体速度精密测量	156
8.3.1 基于运动模糊的物体旋转速度测量	156
8.3.2 基于车载摄像机采集运动模糊图像的车辆平移速度测量	157
8.3.3 基于公路两侧安装摄像机采集运动模糊图像的车辆平移速度测量	159
8.3.4 考虑车辆运动模糊与摄像机离焦模糊耦合时的车辆速度测量	162
8.3.5 采集单幅隔行扫描 CCD 图像进行车辆速度测量	164
本章参考文献	165
第9章 林业工程应用范例——板材材种显微细胞图像分类识别	167
9.1 概述	167
9.1.1 国内研究概述	167
9.1.2 国外研究概述	168
9.1.3 板材材种分类识别	168
9.2 系统硬件及软件构成	170
9.2.1 系统硬件组成与配置	170
9.2.2 系统软件设计	171
9.3 板材细胞形状特征提取与分类识别	173
9.3.1 细胞图像预处理	173
9.3.2 板材细胞图像分割	176
9.3.3 细胞外轮廓定型	181
9.3.4 基准细胞模拟	190
9.4 细胞纹理特征提取与分类识别	193
9.4.1 纹理图像研究方法	194
9.4.2 Contourlet 变换理论	196
9.4.3 板材细胞图像分类识别	202
本章参考文献	206

第 10 章 农业工程应用范例——农作物杂草分类识别	208
10.1 杂草识别	209
10.2 利用杂草位置信息进行识别	211
10.3 利用杂草颜色特征进行识别	212
10.4 利用杂草形状特征进行识别	214
10.5 利用杂草纹理特征进行识别	215
10.6 利用杂草光谱特征进行识别	216
10.7 利用杂草多特征融合进行识别	217
10.8 利用杂草模糊信息进行识别	218
本章参考文献	221
第 11 章 应用实例——光笔式 3D 坐标视觉测量系统及其应用	226
11.1 系统建模	226
11.1.1 系统坐标系的建立	226
11.1.2 共线 3 点透视问题的求解	227
11.1.3 被测点三维坐标的求解	228
11.1.4 系统模型的唯一性证明	228
11.1.5 实验	230
11.2 光笔式坐标测量系统中控制点光斑图像的识别	230
11.2.1 光笔上发光二极管的成像特点	230
11.2.2 发光二极管椭圆形光斑图像的识别	231
11.3 光笔式坐标测量系统中控制点光斑图像中心的并行定位算法	233
11.4 采用光笔式坐标测量系统进行曲面物体边界周长的精密测量	236
11.5 采用光笔式坐标测量系统进行曲面物体表面积的精密测量	237
11.5.1 B 样条变形曲面	238
11.5.2 面积计算	238
本章参考文献	239

第 1 章 引 论

人类是通过眼睛和大脑来获取、处理与理解视觉信息的。周围环境中的物体在可见光照射下，在人眼的视网膜上形成图像，由感光细胞将其转换成神经脉冲信号，并经神经纤维传入大脑皮层进行处理与理解。所以说，视觉不仅是指对光信号的感受，还包括对视觉信息的获取、传输、处理与理解的全过程。

随着信号处理理论和计算机技术的发展，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一幅或多幅图像认知周围环境信息的能力。这使计算机不仅能模拟人眼的功能，更重要的是使计算机完成人眼所不能胜任的工作。机器视觉则是建立在计算机视觉理论基础之上，侧重于计算机视觉技术工程化应用。与计算机视觉研究的视觉模式识别、视觉理解等内容不同，机器视觉重点在于感知环境中物体的形状、位置、姿态、运动等几何信息。

本章首先讨论机器视觉的发展及系统构成，然后介绍 Marr 的视觉理论框架，最后对机器视觉的应用领域与面临的问题，以及本书各章内容进行介绍。

1.1 机器视觉的发展及系统构成

视觉是人类观察世界和认知世界的重要手段。据统计，人类从外部世界获得的信息约有 80% 是由视觉获取的。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，同时又体现了人类视觉功能的重要性。随着信息技术的发展，给计算机、机器人或其他智能机器赋予人类视觉功能，成为人类多年以来的梦想。虽然目前还不能够使计算机、机器人或其他智能机器也具有像人类等生物那样高效、灵活和通用的视觉，但自 20 世纪 50 年代以来，视觉理论和技术得到了迅速发展，这使得人类的梦想正在逐步实现。



1.1.1 机器视觉的发展

计算机视觉是用计算机实现人的视觉功能——对客观世界的三维场景的感知、识别和理解。计算机视觉是在 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像分析、识别和理解上，如光学字符识别、工件表面、显微图片和航空照片的分析和解释等。20 世纪 60 年代，Roberts 将环境限制在所谓的“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。通过计算机程序从数字图像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。到 20 世纪 70 年代，已经出现了一些视觉应用系统 [Guzman 1969, Mackworth 1973]。

1973 年，英国的 Marr 教授应邀在麻省理工学院（MIT）的人工智能实验室创建并领导一个以博士生为主体的研究小组，从事视觉理论方面的研究。1977 年，Marr 提出了不同于“积木世界”分析方法的计算视觉理论——Marr 视觉理论，该理论在 20 世纪 80 年代成为计算机视觉研究领域中的一个十分重要的理论框架 [Marr 1982]。到了 20 世纪 80 年代中期，计算机视觉获得了迅速发展，主动视觉理论框架、基于感知特征群的物体识别理论框架等新概念、新方法、新理论不断涌现。而到了 20 世纪 90 年代，计算机视觉在工业环境中得到广泛应用 [Mouaddib 1997, Zhang 1997, Tech 1995]，同时基于多视几何的视觉理论得到迅速发展 [Xu 1996]。

在上面的讨论中，计算机视觉（Computer Vision）和机器视觉（Machine Vision）两个术语是不加以区分的，在很多文献中也是如此，但其实这两个术语是既有区别又有联系的。计算机视觉是采用图像处理、模式识别、人工智能技术相结合的手段，着重于一幅或多幅图像的计算机分析。图像可以由单个或多个传感器获取，也可以是单个传感器在不同时刻获取的图像序列。分析是对目标物体的识别，确定目标物体的位置和姿态，对三维景物进行符号描述和解释。在计算机视觉研究中，经常使用几何模型、复杂的知识表达，采用基于模型的匹配和搜索技术，搜索的策略常使用自底向上、自顶向下、分层和启发式控制策略。机器视觉则侧重于计算机视觉技术工程化，能够自动获取和分析特定的图像，以控制相应的行为。具体地说，计算机视觉为机器视觉提供图像和景物分析的理论及算法基础，机器视觉为计算机视觉的实现提供传感器模型、系统构造和实现手段。因此可以认为，一个机器视觉系统就是一个能自动获取一幅或多幅目标物体图像，对所获取图像的各种特征量进行处理、分析和测量，并对测量结果做出定性分析和定量解释，从而得到有关目标物体的某种认识并做出相应决策的系统。机器视觉系统的功能包括物体定位、特征检测、缺陷判断、目标识别、计数和运动跟踪。



1.1.2 机器视觉系统构成

机器视觉系统一般以计算机为中心，主要由视觉传感器、高速图像采集系统及专用图像处理系统等模块组成，如图 1.1 所示。

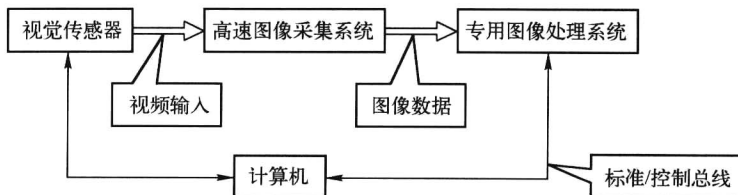


图 1.1 机器视觉系统的基本组成框架

视觉传感器是整个机器视觉系统信息的直接来源，主要由一个或两个图像传感器组成，有时还要配以光投射器及其他辅助设备。它的主要功能是获取足够的机器视觉系统要处理的最原始图像。图像传感器可以是激光扫描器、线阵和面阵 CCD 摄像机或 TV 摄像机，也可以

是最新出现的数字摄像机等。尤其是线阵和面阵 CCD 摄像机，它们在计算机视觉的发展和应用中起着至关重要的作用。随着半导体集成技术和超大规模微细加工技术的发展，面阵 CCD 摄像机不仅商品化，而且具有高分辨率和高工作速度。另外，它所具有的二维特性、灵敏度高、可靠性好、几何畸变小、无图像滞后和图像漂移等优点使其成为计算机视觉中非常适合的图像传感器。光投射器可以作为普通照明光源、半导体激光器或红外激光器等，它的功能主要是参与形成被分析的物体图像的特征。其他辅助设备为传感器提供电源和控制接口等功能。

进入 20 世纪 90 年代，为满足对小型化、低功耗和低成本成像系统消费需求的增加，出现了几种新的固体图像传感技术，其中最引人注目且最有发展潜力的是采用标准 CMOS 半导体工艺生产的图像传感器，即 CMOS 图像传感器。可以预计，CMOS 图像传感器以其独特的优点在计算机视觉系统中将具有广泛的应用前景。

高速图像采集系统是由专用视频解码器、图像缓冲器和控制接口电路组成的。它的主要功能是实时地将视觉传感器获取的模拟视频信号转换为数字图像信号，并将图像直接传送给计算机进行显示和处理，或将数字图像传送给专用图像处理系统进行视觉信号的实时前端处理。随着专用视频解码器芯片和现场可编程逻辑门阵列（FPGA）芯片的出现，现在的大多数高速图像采集系统由少数几个芯片就可以完成。图像采集系统与计算机的接口采用工业标准总线，如 ISA 总线、VME 总线或 PCI 总线等，使图像采集系统到计算机的实时图像数据传输成为可能。

专用图像处理系统是计算机的辅助处理器，主要采用专用集成芯片（ASIC）、数字信号处理器（DSP）或 FPGA 等设计的全硬件处理器。它可以实时高速完成各种低级图像处理算法，减轻计算机的处理负荷，提高整个视觉系统的速度。专用图像处理系统与计算机之间的通信可以采用标准总线接口、串行通信总线接口或网络通信等方式。各种硬件处理系统的出现，如基于 FPGA 的超级计算机和实时低级图像处理系统等，为机器视觉系统实时实现提供了有利的条件。

计算机是整个机器视觉系统的核心，它除了控制整个系统的各个模块的正常运行外，还承担着视觉系统的最后结果运算和输出。由图像采集系统输出的数字图像可以直接传送到计算机，由计算机采用纯软件方式完成所有的图像处理和其他运算。如果纯软件处理能够满足视觉系统的要求，专用硬件处理系统就不出现在机器视觉系统中。这样，一个实用机器视觉系统的结构、性能、处理时间和价格等都可以根据具体应用而定，因此比较灵活。

针对有些机器视觉系统，还需配有相应的工件传输和定位系统，以使待监测的工件通过特定的传送系统安放到预定的空间内，必要的时候，加以定位限制。

为适应现代工业发展的需要，在各种小型机、微型机，特别是在功能强大的 IBM-PC 上开发各种专用微型视觉组件变得更为重要。越来越多的公司投入大量人力、物力研究视觉组件产品。单就美国而言，早在 1983 年年底就有一百多家公司跻身于计算机视觉系统的市场，而到现在，经过近 20 年的发展，投入到这个领域的公司不计其数。随着微处理器和超大规模集成电路技术日益成熟，从而能生产出更小、更先进、更灵活可靠、更耐用的视觉组件产品，并使它们走出实验室进入实际工作现场。

在二维视觉处理中，随着机器视觉的飞速发展，已从二值视觉系统发展为灰度视觉系统，并达到实用。二值视觉系统仅通过像素由 0 到 1 或由 1 到 0 的变化提取图像边缘点，它

需要高对比度图像。灰度视觉系统具有检测复杂场景的能力,如复杂工件识别和表面特征(纹理、阴影、模式等)分析。采用一定的算法,系统精度受照明变化的影响很小。灰度是图像辐射或亮度的量化测量。该信息是通过视频 A/D 转换器存储在帧存体中获得的,灰度分辨率随机器视觉系统的不同而不同,但数值通常是 2 的乘方:4、16、64 和 256。灰度分辨率将确定视觉系统检测区域亮度值的最小变化。灰度分辨率结合“子像素”能力在机器视觉系统中起着重要作用。

在三维视觉信息获取上,近年来也取得了巨大的进步,由于实现思想和条件不同,产生了相应的诸多方法。例如,根据照明方式可分为主动测距法和被动测距法。前者需要利用特别的光源所提供的结构信息,而后者获取深度信息是在自然光下完成的。被动测距法适合于受环境限制和需保密的场合。而主动测距法可应用的领域非常广泛,且具有测量精度高、抗干扰性能好和实时性强等优点。总之,三维视觉的引入进一步扩大了机器视觉的应用领域。

此外,机器视觉系统的界面是开放性的,用户可根据应用需要进行计算机编程,以改善系统的功能。为进一步使机器视觉系统与工业自动化现场相适应,由菜单驱动系统,用鼠标或光笔在屏幕上选择功能项即可,易于操作,无须专门培训操作员。

机器视觉系统具有高度的智能和普遍的适用性,随着它的不断完善,已完全能用于工业现场,满足现代生产过程的要求。

1.2 Marr 的视觉理论框架

20 世纪 80 年代初, Marr (1982) 首次从信息处理的角度综合了图像处理、心理物理学、神经生理学及临床神经病学等方面已取得的重要研究成果,提出了第一个较为完善的视觉系统框架,使计算机视觉研究有了一个比较明确的体系。虽然这个理论还需要通过研究不断改进和完善,但 Marr 的视觉计算理论是首次提出的阐述视觉机理的系统理论,并且对人类视觉和计算机视觉的研究都产生了深远的推动作用。下面简要介绍 Marr 视觉理论的基本思想及理论框架 [Marr 1982, 吴健康 1993]。



1.2.1 视觉系统研究的三个层次

Marr 从信息处理系统的角度出发,认为对视觉系统的研究应分为三个层次,即计算理论层次、表达与算法层次和硬件实现层次。

计算理论层次要回答视觉系统的计算目的与计算策略是什么或视觉系统的输入/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,视觉系统输入是二维图像,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,如何由二维灰度图像恢复物体的三维信息。表达与算法层次是要进一步回答如何表达输入和输出信息,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转换成另一种表示。一般来说,不同的表达方式,完成同一计算的算法会不同,但 Marr 算法与表达是比计算理论低一层次的问题,不同的表达与算法,在计算理论层次上可以是相同的。最后一个硬件实现层次,是解决用硬件实现上述表达和算法的问题,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，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，则往往对理解算法无济于事。

区分以上三个不同层次，对于深刻理解计算机视觉与生物视觉系统及它们的关系都是有益的，例如，人的视觉系统与目前的计算机视觉系统在“硬件实现”层次上是完全不同的，前者是极为复杂的神经网络，而后者是目前使用的计算机，但它们可能在计算理论层次上完成相同的功能。

视觉系统研究的三个层次的含义如表 1.1 所示。

表 1.1 视觉系统研究的三个层次的含义

要素	名称	含义和所解决的问题
1	计算理论	什么是计算目的，为什么要这样计算
2	表达和算法	怎样实现计算理论，什么是输入/输出表达，用什么算法实现表达间的转换
3	硬件实现	怎样在物理上实现表达和算法，什么是计算结构的具体细节



1.2.2 视觉信息处理的三个阶段

Marr 从视觉计算理论出发，将系统分为自下而上的三个阶段，即视觉信息从最初的原始数据（二维图像数据）到最终对三维环境的表达经历的三个阶段的处理，如图 1.2 及表 1.2 所示。第一阶段（也称为早期阶段）构成所谓“要素图”或“基元图”（Primary Sketch），基元图由二维图像中的边缘点、直线段、曲线、顶点、纹理等基本几何元素或特征组成；第二阶段（中期阶段），Marr 称其为对环境的 2.5 维描述，2.5 维描述是一种形象的说法，意即部分的、不完整的三维信息描述，用“计算”的语言来讲，就是重建三维物体在以观察者为中心的坐标系下的三维形状与位置。当人眼或摄像机观察周围环境物体时，观察者对三维物体最初是以自身的坐标系来描述的，另外，我们只能观察到物体的一部分（另一部分是物体的背面或被其他物体遮挡的部分）。这样，重建的结果是以观察者为中心的坐标系下描述的部分三维物体形状，称为 2.5 维描述。这一阶段中存在许多并行的相对独立的模块，如立体视觉、运动分析、由灰度恢复表面形状等不同处理单元。事实上，从各种不同角度去观察物体，观察到的形状都是不完整的，不能设想，人脑中存有同一物体从所有可能的观察角度看到的物体形象，以用来与所谓的物体的 2.5 维描述进行匹配与比较。因此，2.5 维描述必须进一步处理以得到物体的完整三维描述，而且必须是物体本身在某一固定坐标系下的描述，这一阶段称为第三阶段（后期阶段）。视觉信息处理的三个阶段如图 1.2 所示。

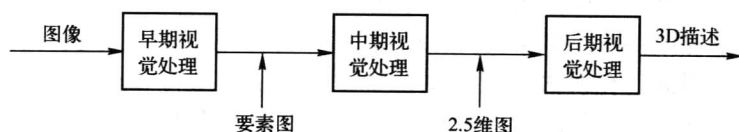


图 1.2 Marr 框架的视觉信息处理的三个阶段

表 1.2 由图像恢复形状信息的表达框架

名称	目的	基元
图像	亮度表示	图像中每一点的亮度值
基元图	表示二维图像中的重要信息, 主要是图像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 维基图	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
三维模型表示	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

Marr 理论是计算机视觉研究领域的划时代成就, 积极推动了这一领域的研究, 多年来对图像理解和计算机视觉研究的发展起了重要的作用。但 Marr 理论也有其不足之处, 其中有 4 个有关整体框架 (如图 1.2 所示) 的问题 [吴立德 1993]:

- (1) 框架中输入是被动的, 给什么图像, 系统就处理什么图像。
- (2) 框架中加工目的不变, 总是恢复场景中物体的位置和形状等。
- (3) 框架缺乏或者说未足够重视高层知识的指导作用。
- (4) 整个框架中信息加工过程基本自下而上, 单向流动, 没有反馈。

针对上述问题, 近年来人们提出了一系列改进思路, 对应图 1.2 所示的框架, 可将其改进并融入新的模块得到如图 1.3 所示的框架, 具体改进如下:

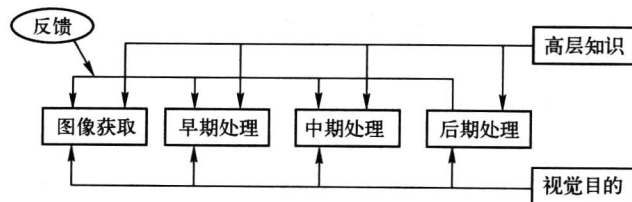


图 1.3 改进的 Marr 框架

(1) 人类视觉是主动的, 会根据需要改变视角, 以帮助识别。主动视觉指视觉系统可以根据已有的分析结果和视觉的当前要求, 决定摄像机的运动, 以从合适的视角获取相应的图像。人类的视觉又是有选择的, 可以注视 (以较高分辨率观察感兴趣区域), 也可以对场景中某些部分视而不见。选择性视觉指视觉系统可以根据已有的分析结果和视觉的当前要求, 决定摄像机的注意点以获取相应的图像。考虑到这些因素, 在改进框架中增加了图像获取模块。该模块要根据视觉目的选择采集方式。

(2) 人类的视觉可以根据不同的目的进行调整。有目的视觉 (也称为定性视觉) 指视觉系统根据视觉目的进行决策, 例如, 是完整地恢复场景中物体的位置和形状等还是仅仅检测场景中是否有某物体存在。事实上, 有相当多的场合只需定性结果就可以, 并不需要复杂性高的定量结果。因此, 在改进框架中增加了视觉目的模块 [Aloimonos 1992], 但定性分析还缺乏完备的数学工具。顺便指出, 有一种相关的观点认为 Marr 关于对场景先重建后解释的思路可以简化视觉任务, 但与人的视觉功能并不完全吻合。事实上重建和解释不总是串行的。

(3) 人类可在仅从图像获取部分信息的情况下完全解决视觉问题,原因是隐含地使用了各种知识。例如,借助CAD设计资料获取物体形状信息(使用物体模型库),可帮助解决由单幅图恢复物体形状困难。利用高层知识可解决低层信息不足的问题,所以在改进框架中增加了高层知识模块 [Huang 1993]。

(4) 人类视觉中前后处理之间是有交互作用的,尽管对这种交互作用的机理了解得还不充分,但高层知识和后期处理的反馈信息对早期处理的作用是重要的。从这个角度出发,在改进框架中增加了反馈控制流向。

最后需要指出,限于历史等因素, Marr 没有研究如何用数学方法严格地描述视觉信息的问题,虽然较充分地研究了早期视觉,但基本没有论及对视觉知识的表达、使用和基于视觉知识的识别等。近年来,有许多试图建立计算机视觉理论框架的工作,其中 Grossberg 宣称建立了一个新的视觉理论:表观动态几何学 (Dynamic Geometry of Surface Form and Appearance)。它指出感知的表面形状是分布在多个空间尺度上多种处理动作的总结果,因此 2.5 维图并不存在,向 Marr 的理论提出了挑战 [吴健康 1993]。但 Marr 的理论使得人们对视觉信息的研究有了明确的内容和较完整的基本体系,仍被看做是研究的主流。现在新提出的理论框架均包含它的基本成分,多数被看做它的补充和发展。尽管 Marr 理论在许多方面还存在争议,但至今它仍是广大计算机视觉工作者所普遍接受的计算机视觉理论基本框架。

1.3 机器视觉的应用领域及面临的问题

机器视觉在国民经济、科学研究及国防建设等领域都有着广泛的应用。视觉的最大优点是与被观测的对象无接触,因此对观测与被观测者都不会产生任何损伤,这是其他感觉方式无法比拟的。另外,视觉方式所能检测的对象十分广泛,人眼观察不到的范围,机器视觉也可以观察,例如,红外线、微波、超声波等人类就观察不到,而机器视觉则可以利用这方面的敏感器件形成红外线、微波、超声波等图像。因此可以说是扩展了人类的视觉范围。另外,人无法长时间地观察对象,机器视觉则不知疲劳,始终如一地观测,所以机器视觉可以广泛地用于长时间恶劣的工作环境。不过机器视觉技术仍处于一个完善和发展的阶段,其发展还远远落后于人们所希望的发展水平。



1.3.1 机器视觉的应用领域

机器视觉应用系统将在第7章~第11章中介绍,下面只列举一些机器视觉的主要应用领域。

(1) 工业自动化生产线应用。产品检测、工业探伤、自动流水线生产和装配、自动焊接、PCB印制板检查,以及各种危险场合工作的机器人等。将图像和视觉技术用于生产自动化,可以加快生产速度,保证质量的一致性,还可以避免人的疲劳、注意力不集中等带来的误判。

(2) 各类检验和监视应用。标签文字标记检查,邮政自动化,计算机辅助外科手术,显微医学操作,石油、煤矿等钻探中数据流自动监测和滤波,在纺织、印染业进行自动分色、配色,重要场所门廊自动巡视,自动跟踪报警等。