

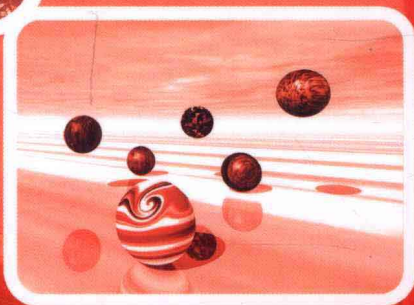
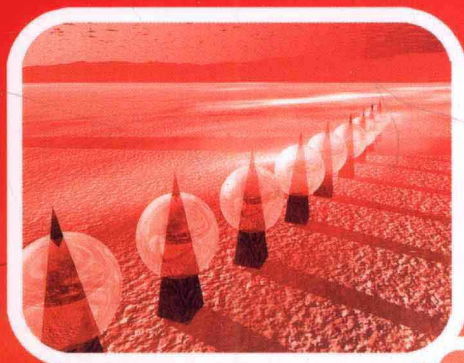


人工心理与数字人技术丛书

视觉测量技术

○ 迟健男 主编

○ 王先梅 胡四泉 曹洋 魏世丞 副主编



机械工业出版社
CHINA MACHINE PRESS



人工心理与数字人技术丛书

视觉测量技术

迟健男 主 编

王先梅 胡四泉 副主编

曹 洋 魏世丞



机械工业出版社

视觉测量技术的研究内容主要包括视觉测量系统的组成、理论模型和图像算法等。本书分别介绍视觉测量技术的历史与发展、系统的硬件组成、图像与处理、图像分割、摄像机标定、单目视觉测量与双目视觉测量的相关技术与方法。

本书适合从事计算机、自动化、模式识别、智能科学、人机交互技术的科技人员阅读,也可以作为高等院校相关专业的学生、研究生的教学参考书。

图书在版编目(CIP)数据

视觉测量技术/迟健男主编. —北京:机械工业出版社, 2011.6
(人工心理与数字人技术丛书)
ISBN 978-7-111-34687-6

I. ①视… II. ①迟… III. ①计算机视觉—测量 IV. ①TP391.41

中国版本图书馆CIP数据核字(2011)第088474号

机械工业出版社(北京市百万庄大街22号 邮政编码100037)

策划编辑:王欢 责任编辑:王欢

版式设计:霍永明 责任校对:李秋荣

封面设计:陈沛 责任印制:乔宇

北京机工印刷厂印刷(三河市南杨庄国丰装订厂装订)

2011年6月第1版第1次印刷

184mm×260mm·16.5印张·407千字

0 001—3 000册

标准书号:ISBN 978-7-111-34687-6

定价:48.00元

凡购本书,如有缺页、倒页、脱页,由本社发行部调换

电话服务

网络服务

社服务中心:(010)88361066 门户网:<http://www.cmpbook.com>

销售一部:(010)68326294 教材网:<http://www.cmpedu.com>

销售二部:(010)88379649

读者购书热线:(010)88379203 封面无防伪标均为盗版

丛书序言

在2006年1月26日公布的《国家中长期科学和技术发展规划纲要（2006—2020年）》中，指出了信息技术将继续向高性能、低成本、普适计算和智能化等主要方向发展，寻求新的计算与处理方式和物理实现是未来信息技术领域面临的重大挑战。纳米科技、生物技术与认知科学等多学科的交叉融合，将促进基于生物特征的、以图像和自然语言理解为基础的“以人为中心”的信息技术发展，作为信息学科前沿技术要研究个性化的智能机器人和人机交互系统。在国家科技发展纲要中如此描述信息技术的发展方向，这表明新世纪科学发展的特征是对人的研究、对人与自然和谐相处的研究；其表现是纳米-生物-信息-认知（Nano-Bio-Info-Cognition, NBIC）多学科交叉、多技术融合研究的突起；是数字人（物理机器人与虚拟软件人的总称，数字人与数字社会的关系研究）技术的趋于热门化和普遍化。

以人为本，以人与自然的和谐相处为研究目标；多学科交叉结合为研究手段，以人工科学为主要研究领域，将是未来若干年科学技术研究的主要特征。

为了响应执行国家《科技发展纲要》，顺应科学技术潮流，我们萌生了编写“人工心理与数字人技术丛书”的想法，希望能够对国家的科技进步有所贡献。

“人工心理与数字人技术丛书”选题主要包括以下范围（不局限于这些范围）：

- 1) NBIC 技术；
- 2) 广义人工智能技术；
- 3) 生物特征识别技术；
- 4) 虚拟现实技术；
- 5) 机器人技术；
- 6) 虚拟人技术；
- 7) 人机交互技术；
- 8) 普适计算。

“丛书”的选题是开放的，我们殷切希望国内外同行专家学者一起来撰写此领域的学术著作，为中华民族的科学技术事业共同努力。

本丛书有如下特色：

- 1) 本丛书主要是前沿技术专题论著，选题内容新颖；
- 2) 选题主要是前沿技术，重点在于紧跟世界科技发展新趋势；
- 3) 内容深入浅出，便于自学。

本丛书以科研人员及大专院校师生为主要读者，也可供工程技术人员学习前沿技术时作为参考。

丛书主编 王志良

前 言

视觉测量技术是以现代光学为基础，融合计算机技术、激光技术、图像处理与分析技术等现代科学技术为一体，组成光机电一体化综合测量系统。基于视觉测量技术的检测仪器设备能够实现智能化、数字化、小型化、网络化和多功能化，具有精度高、非接触、在线检测、实时分析与控制、连续工作等特点，能够适应多种危险的应用场合，广泛应用于军事、工业、农林业、医学、航空航天、科学研究等领域。

作者长期从事视觉测量技术的研究及教学工作，全书以作者多年的教学经验与研究成果为主要内容，对视觉测量的各项关键技术进行了详细的阐述。全书共分七章。第1章“视觉测量技术概述”，主要介绍了视觉测量技术的基本概念、研究历史与发展趋势。第2章“视觉测量系统”，主要介绍了系统的各个组成部分及相关概念。第3章“图像预处理”，主要介绍了图像与处理各个阶段的各种模型算法。第4章“图像分割”，主要介绍了图像分割的各种检测与处理方法。第5章“摄像机标定”，主要介绍了摄像机模型及几种主要的标定方法。第6章“单目视觉测量”，主要介绍了单目视觉测量的几何测量方法，虚拟立体视觉的原理、模型、参数计算、标定方法，以及光笔式三维坐标测量方法和几何光学法等。第7章“双目视觉测量”，主要介绍了双目立体视觉的原理、模型、结构参数、测量精度，以及摄像机标定、匹配等方法。

全书由迟健男副教授撰写和统稿，参与编写工作的还有胡四泉、王先梅、曹洋、魏世丞、黄伟明、崔丽波、赵大海、李明、杨林、杨靖、朱文翔、王开然等。

本书深入浅出，内容系统，适合从事视觉测量行业的工程师、技术人员作为参考书使用，也适合高等院校相关专业的本科生、研究生及师生作为教材及相关参考书籍使用。本书编写工作得到了国家自然科学基金、国家高技术研究发展计划（863）的支持，在此表示感谢。由于作者的水平有限，书中难免有错误和不足，请读者批评指正。

迟健男

2011年5月29日于北京科技大学

目 录

丛书序言

前言

第1章 视觉测量技术概述 1

- 1.1 计算机视觉 1
 - 1.1.1 计算机视觉的概念 1
 - 1.1.2 计算机视觉的实现方法 1
 - 1.1.3 计算机视觉技术的研究目标 2
 - 1.1.4 计算机视觉技术的发展历程 2
 - 1.1.5 Marr 视觉理论简介 3
 - 1.1.6 计算机视觉技术的主要研究内容 6
 - 1.1.7 计算机视觉系统的构成 8
 - 1.1.8 计算机视觉技术的应用领域 10
- 1.2 计算机视觉检测技术 16
 - 1.2.1 计算机视觉检测技术分类 16
 - 1.2.2 计算机视觉测量技术 17
 - 1.2.3 计算机视觉测量系统构成 18
 - 1.2.4 计算机视觉测量技术的性能指标 20
 - 1.2.5 计算机视觉测量系统的关键技术 20
 - 1.2.6 计算机视觉测量技术的优点 22
 - 1.2.7 影响测量系统精度的因素与提高方法 22
 - 1.2.8 计算机视觉测量技术的主要应用领域 23
 - 1.2.9 计算机视觉测量技术的发展趋势 27
 - 1.2.10 计算机视觉测量的流程 29

第2章 计算机视觉测量系统 30

- 2.1 光源照明系统 30
 - 2.1.1 光源系统 31
 - 2.1.2 常用可见光光源 32
 - 2.1.3 光源类型 37
 - 2.1.4 照明方式 38
- 2.2 固态摄像机 47
 - 2.2.1 固态摄像机简介 47
 - 2.2.2 摄影光学系统 48
 - 2.2.3 光学镜头的一般知识介绍 51
 - 2.2.4 分色滤光片 54

2.3 图像传感器 55

- 2.3.1 CCD 图像传感器 55
- 2.3.2 CMOS 图像传感器 61
- 2.3.3 CCD 与 CMOS 图像传感器的差异 62
- 2.4 图像采集卡 63
- 2.5 人工标志及被测物体表面处理 65
 - 2.5.1 一般人工标志 66
 - 2.5.2 人工标志的设计原则 66
 - 2.5.3 人工标志的分类 67
 - 2.5.4 回光反射标志 67
 - 2.5.5 常用人工标记 69
 - 2.5.6 被测物体的表面处理 71
- 2.6 计算机 71

第3章 图像预处理 73

- 3.1 概述 73
- 3.2 图像滤波 74
 - 3.2.1 图像噪声类型 74
 - 3.2.2 空间域滤波 76
 - 3.2.3 变换域滤波 81
- 3.3 频域滤波 81
 - 3.3.1 一维离散傅里叶变换 81
 - 3.3.2 二维离散傅里叶变换 83
 - 3.3.3 二维离散傅里叶变换的性质 84
 - 3.3.4 频域滤波的基本步骤 85
 - 3.3.5 频域滤波和空间域滤波的关系 86
 - 3.3.6 频域低通滤波 86
- 3.4 图像增强 89
 - 3.4.1 空间域增强 89
 - 3.4.2 直方图均衡化 92
 - 3.4.3 直方图规定化方法 99
 - 3.4.4 频域增强 100

第4章 图像分割 105

- 4.1 图像分割概述 105
- 4.2 间断检测 108
- 4.3 边缘检测 112
 - 4.3.1 边缘点检测 113
 - 4.3.2 边缘连接 125
- 4.4 阈值分割 132
 - 4.4.1 阈值分割的基本原理 132

4.4.2	双峰法	134	6.5	光笔式三维坐标测量方法	203
4.4.3	p 参数法	135	6.5.1	系统构成	205
4.4.4	最小类内方差法	136	6.5.2	测量原理	205
4.4.5	最大类间距离法	136	6.5.3	数学模型	206
4.4.6	最大类间类内方差比法	137	6.5.4	关键技术	209
4.4.7	最大熵法	137	6.5.5	典型测量系统	210
4.4.8	自动迭代法	138	6.6	几何光学法	212
4.4.9	直方图变换法	138	6.6.1	散焦测距的基本原理	214
4.4.10	连通区域标记	139	6.6.2	散焦测距的 Pentland 算法	215
4.5	区域生长法	142	6.6.3	Subbarao 的深度恢复算法	218
4.5.1	基于像素相似性的简单区域生长法	143	第7章 双目视觉测量	221	
4.5.2	基于像素与区域相似性的质心型区域生长法	145	7.1	双目视觉测量系统	221
4.5.3	混合型区域增长	149	7.2	双目视觉三维测量原理	222
4.5.4	基于区域形状的区域增长	149	7.3	双目视觉测量数学模型	224
4.6	区域分裂-合并方法	150	7.4	双目视觉测量系统精度分析	225
第5章 摄像机标定	152		7.4.1	测量系统结构参数对测量精度的影响	225
5.1	摄像机模型	153	7.4.2	摄像机焦距对测量精度的影响	231
5.1.1	坐标变换	153	7.5	双目视觉系统结构设计	232
5.1.2	成像变换	157	7.5.1	基于两台摄像机的双目视觉系统结构	232
5.1.3	摄像机的镜头畸变	158	7.5.2	基于单台摄像机的双目视觉系统结构	232
5.1.4	摄像机模型	160	7.5.3	同一物镜法	233
5.2	摄像机标定	166	7.6	运动物体的同步摄影方法	234
5.2.1	摄像机标定的透视变换法	166	7.7	双目视觉测量系统标定	234
5.2.2	摄像机标定两步法	174	7.8	双目视觉测量系统立体匹配方法概述	236
5.2.3	摄像机标定两步法的标定过程	177	7.8.1	立体匹配技术	236
5.2.4	摄像机标定的张正友法	180	7.8.2	稠密匹配方法剖析	236
5.2.5	摄像机标定双平面法	187	7.8.3	特征匹配方法剖析	237
5.2.6	畸变校正	189	7.9	双目视觉测量系统常用约束条件	238
第6章 单目视觉测量	191		7.9.1	约束规则	238
6.1	几何相似法测量	191	7.9.2	外极约束	239
6.2	几何形状约束法测量	192	7.9.3	其他约束条件	241
6.3	基于单摄像机的虚拟立体视觉	193	7.10	双目视觉测量特征匹配算法介绍	243
6.3.1	虚拟立体视觉简介	193	7.10.1	基于兴趣点的匹配算法	244
6.3.2	虚拟立体视觉测量原理	194	7.10.2	基于固定滑窗区域的匹配算法	244
6.3.3	单摄像机虚拟立体视觉模型	194	7.10.3	基于形状模板的匹配算法	246
6.3.4	系统相关参数计算	195	7.10.4	基于边缘特征的匹配算法	247
6.4	结构光视觉测量法	196	参考文献	248	
6.4.1	系统构成	196			
6.4.2	数学模型	196			
6.4.3	投影模式	198			
6.4.4	标定方法	202			

第1章 视觉测量技术概述

基于计算机视觉的视觉测量技术是近年来测量领域中迅速发展起来的崭新技术。它是现代光学为基础，融合计算机技术、激光技术、图像处理与分析技术等现代科学技术为一体，组成光机电一体化综合测量系统。基于视觉测量技术的检测仪器设备能够实现智能化、数字化、小型化、网络化和多功能化，具有精度高、非接触、在线检测、实时分析与控制、连续工作等特点，能够适应多种危险的应用场合，广泛应用于军事、工业、农林业、医学、航空航天、科学研究等领域。视觉测量的理论基础是计算机视觉技术，因此本章首先介绍计算机视觉相关知识，在此基础上对视觉测量技术进行概述和介绍。

1.1 计算机视觉

1.1.1 计算机视觉的概念

计算机视觉也称为机器视觉，是指利用计算机对采集的图像或者视频进行处理，从而代替人眼的视觉功能，实现对客观世界的三维场景的感知、识别和理解的技术。也就是说，计算机视觉是一门研究如何使机器实现“看”功能的科学，它不仅能模拟人眼的功能，更重要的是完成人眼所不能胜任的工作。计算机视觉是工程和科学领域中的一个富有挑战性研究领域。计算机视觉的研究范畴涉及物理学、数学、光学、计算机科学、信号处理学、生物学、神经生理学和认知科学等，是一门综合性的前沿学科，具有广阔的应用前景。

计算机视觉使用计算机及相关设备对生物视觉进行模拟。其主要任务是通过采集的图像或视频进行处理，以获得相应场景的三维信息。

计算机视觉使用的理论方法主要是基于几何、概率、运动学计算和三维重构等视觉计算理论，其基础包括射影几何学、刚体运动学、概率论与随机过程、人工智能等理论。

1.1.2 计算机视觉的实现方法

根据计算机视觉的技术机理和相关应用，可将其实现方法分为两类：

1. 仿生学方法

仿生学方法主要是参照人类视觉系统的结构原理，建立视觉过程的计算模型和相应的处理模块，实现与人类视觉系统类似的或者全部的功能。例如，仿生眼、主动视觉系统等都是典型的基于仿生学原理的计算机视觉系统。

2. 工程学方法

工程学方法并不刻意模拟人类视觉系统的内部结构和工作过程，它主要是从分析人类视觉过程的功能着手，仅考虑系统的输入和输出，并采用任何现有的、可行的手段实现所需的系统功能。例如，目前广泛应用的双目视觉测距系统、双目和多目视觉空间三维坐标测量系统等都是典型的计算机视觉工程应用系统。

1.1.3 计算机视觉技术的研究目标

借鉴人类视觉的原理，计算机视觉技术研究的长远目标是建立通用的视觉系统。即利用计算机对三维场景的理解，实现人类视觉的功能。其本质问题是利用二维图像重构三维物体的可视部分。计算机视觉技术的优势不仅在于能模拟人眼的功能，还能完成人眼所不能胜任的工作。

根据计算机视觉技术的应用领域，目前计算机视觉技术的研究目标主要体现在以下两个方面：

1. 建成各种专用视觉系统，完成各种实际场合提出的专门视觉任务

其技术原理是，计算机借助各种图像传感器（电荷耦合器件（CCD）、互补金属氧化物半导体（CMOS）摄像器件等）获取场景图像，感知和恢复 3D 环境中物体的几何结构、位置姿态、运动信息等，并对客观场景进行识别、描述、解释，进而做出决策。

2. 建立人类视觉的功能模型和计算模型

即从生物学角度进一步探索人脑和视觉系统的工作机理，加深对人脑及视觉系统的掌握和理解（如计算神经科学），其研究内容主要是生物学机理。长期以来，对于人脑视觉系统已从生理、心理、神经、认知等方面进行了大量的研究，但还远没有揭开视觉过程的全部奥秘，对视觉机理的研究和了解远落后于对视觉信息处理的研究和掌握。因此，对人脑视觉的充分理解将促进对计算机视觉更深入的研究，并对计算机视觉系统提供坚实的生物理论基础。

具体而言，计算机视觉至少要达到如下几个目标：

1) 根据一幅或多幅二维图像，计算出目标物体的几何参数，如物体的高度、宽度、远近距离及姿态等。

2) 根据一幅或多幅二维图像，计算出目标物体的运动参数，如物体的运动速度、加速度及运动方向等。

3) 根据一幅或多幅二维图像，计算出目标物体的表面物理特性，如光滑与粗糙及纹理等。

4) 根据多幅二维图像，恢复出更大空间区域的图像，如根据不同角度拍摄的一组相互间存在重叠部分的目标图像，恢复出包含各图像序列信息的、完整的、高清晰的整个场景图像。图像拼接是该目标的主要技术基础。

1.1.4 计算机视觉技术的发展历程

计算机视觉技术起源于在 20 世纪 50 年代，经过几十年的发展，各种研究理论与研究方法层出不穷，研究内容已经从最初的二维图像分析扩展到当前的三维复杂场景理解，并广泛应用于工业、农业、军事、医学、教育与娱乐等多个领域。

1. 20 世纪 50 年代

统计模式识别的兴起，标志着计算机视觉技术研究的开始。当时的研究工作主要集中在二维图像的分析、识别和理解，如字符识别，工件表面、显微图片和航空图片的分析和解释等。

2. 20 世纪 60 年代

Robert 的工作开创了以理解三维场景为目标的三维计算机视觉的研究。Robert 将环境限制在所谓的“积木世界”，即周围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。

3. 20 世纪 70 年代

出现了计算机视觉应用系统，并首次提出了较为完整的计算机视觉理论。1973 年英国的 Marr 教授应邀在麻省理工学院（Massachusetts Institute of Technology, MIT）的人工智能实验室创建并领导了一个以博士生为主的研究小组，从事视觉理论方面的研究。1977 年，Marr 教授提出了不同于“积木世界”分析方法的计算机视觉理论——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十分重要的理论框架。

4. 20 世纪 80 年代中期

计算机视觉技术获得迅速发展，主动视觉理论框架、基于感知特征群的物体识别理论框架、视觉集成理论框架等新概念、新方法和新理论不断涌现。

5. 从 20 世纪 90 年代至今

进入 20 世纪 90 年代以后，计算机视觉理论的发展十分迅速，已逐渐发展成为一门由计算机技术、控制理论、模式识别、人工智能和生物技术等众多领域交叉综合的新学科。它受到了学术界和工业界的极大重视，在机器人、工业检测、空间遥感、医学图像分析、军事导航和交通管理等诸多领域得到了广泛应用，而且应用领域正在日益扩大。

1.1.5 Marr 视觉理论简介

1977 年，Marr 教授首次从信息处理的角度综合了图像处理、神经生理学及临床神经病学等方面已经取得的重要研究成果，提出了第一个较为完善的视觉理论框架。虽然这个理论还需要不断改进和完善，但 Marr 视觉理论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和框架。

1. Marr 视觉系统研究的三个层次

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

(1) 计算理论层次

这一层次研究的是视觉系统的计算目的与计算策略是什么，或者视觉系统的输入、输出是什么，如何由系统的输入求出系统的输出。在这一层次上，视觉系统的输入是二维图像，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和输出之间的关系和约束，如何由二维图像恢复成物体的三维信息。

(2) 表达与算法层次

这一层次进一步研究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算法，以及如何由一种表示变换成另一种表示。

(3) 硬件实现层次

这一层次解决用硬件实现上述表达与算法的问题，如视觉系统的计算机体系结构和具体的计算装置及其细节。

视觉系统研究的三个层次见表 1.1。

表 1.1 视觉系统的三个层次

层次	名称	含义和所研究的问题
1	计算理论层次	什么是计算目的，为什么要这样计算
2	表达与算法层次	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变换
3	硬件实现层次	怎样在物理上实现表达和算法，什么是视觉系统计算结构的具体细节

2. Marr 视觉信息处理的三个阶段

Marr 从视觉计算理论出发，将视觉系统分为自下而上的三个阶段，即视觉信息从最初的二维图像原始数据到最终三维环境表达经历了三个处理阶段，如图 1.1 所示。

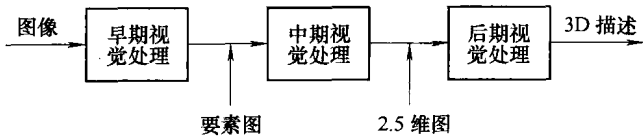


图 1.1 Marr 视觉信息处理的三个阶段

第一阶段是早期视觉处理，其目的是从输入的原始图像中抽取观察者周围景物表面的物理特性，如距离、表面方向、材料特性（反射、颜色、纹理）等，并构成要素图或基元图。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。这一阶段包括边缘检测、双目立体匹配、由阴影确定形状、由纹理确定形状、光流计算等。第一阶段也称为低层视觉阶段。

第二阶段是中期视觉处理，它在以观测者为中心的坐标系中，由输入图像和基元图恢复场景可见部分的深度、法线方向、轮廓等。根据这些描述，可以重建物体边界，并按表面和体积分割景物，但在以观察者为中心的坐标系中只能得到可见表面的描述，无法得到遮挡表面的描述，这还不是真正的物体三维表示。因此 Marr 称之为对环境的 2.5 维描述，即对环境部分的、不完整的三维信息描述，是以观察者为中心的坐标系中描述的部分三维物体形状和位置。第二阶段也称为中层视觉阶段。

第三阶段是后期视觉处理，用 2.5 维图中的表面信息建立适用于视觉识别的三维形状描述。这个描述与观察者的视角无关。也就是说，在以物体为中心的坐标系中，用各种符号和几何结构来描述物体的三维结构和空间关系。第三阶段也称为高层视觉阶段。表 1.2 简单总结了视觉处理三个阶段中的目的和特点。

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	灰度表示	图像中每一像素灰度值
基元图	表示二维图像中的重要信息，主要是图像中的灰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	在以观察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、离观测者距离、深度上的不连续点、表面朝向的不连续点
三维模型	在以物体为中心的坐标系内，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，所有体积基元和面积形状基元都附着在轴线上

3. Marr 视觉理论的意义及不足

Marr 视觉理论是计算机视觉研究领域的划时代成就。几十年来，它对图像理解和计算机视觉的研究发展产生了深远的影响。但它还不是一个完善的理论，因为它没有反映人类视觉的某些重要的本质和特征，即人类视觉的选择性和整体性，而且在实践中也遇到了严重困难。Marr 视觉理论的不足之处主要表现为以下几点：

- 1) 框架中的输入是被动的，给什么图像，系统就处理什么图像；
- 2) 框架中的加工目的是不变的，只能是恢复场景中物体的位置和形状等；
- 3) 框架缺乏或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上是自下而上、单向流动、没有反馈的。研究发现，在生物视觉系统中有许多从高层传送信息的神经纤维，甚至视网膜上也有许多来自中枢的神经，给予支配信息。虽然上述反馈神经的确切作用尚不明确，但是可以证明视觉系统应该有反馈存在。

4. Marr 视觉理论框架的改进

针对 Marr 视觉理论的不足，人们提出了一系列改进思路。

(1) 人类视觉是主动的

人类视觉可根据需要改变视角，以帮助识别。主动视觉是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。

(2) 人类视觉是有选择的

人类视觉可以注视（以较高的分辨率观察感兴趣的区域），也可以对场景中的某些部分视而不见。

选择性视觉是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。该模块要根据视觉的目的选择采集方式。

(3) 人类视觉可以根据不同目的进行调整

有目的的视觉也称为定性视觉，是指视觉系统根据视觉目的进行决策。例如辨别完整的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。在很多情况下，只需要定性结果就可以了，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。

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，重建和解释并不总是串行的。

(4) 人类视觉由高层知识指导

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，其原因是隐含地使用了各种知识。例如，借助计算机辅助设计（CAD）物体模型库资料，可获取物体形状信息，可帮助解决由单幅图像恢复物体形状的困难。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。

(5) 人类视觉有反馈

人类视觉中前后处理是有交互作用的，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模块。

改进后的 Marr 视觉理论框架如图 1.2 所示。

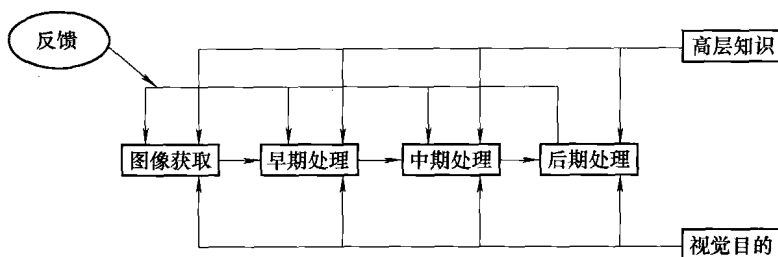


图 1.2 改进后的 Marr 视觉理论框架

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。虽然他较充分地研究了早期视觉，但基本上没有论及对视觉知识的表达、使用和识别等。

近年来，有许多试图建立计算机视觉理论框架方面的研究。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的、多种处理动作的总的结果，因此 2.5 维信息是不存在的。从而向 Marr 视觉理论提出了挑战。

但 Marr 视觉理论使得人们对信息的研究有了明确的内容和较完整的基本体系，仍被看作是研究的主流。现在新提出的理论框架均包含它的基本成分，多数被看作是它的补充和发展。尽管 Marr 视觉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普遍接受的计算机视觉理论基本框架。

1.1.6 计算机视觉技术的主要研究内容

1. 计算机视觉技术的研究内容

下面从图像输入设备、低层视觉、中层视觉、高层视觉和体系结构五个方面介绍计算机视觉技术的主要研究内容。

(1) 图像输入设备（Input Device）

输入设备包括成像设备和数字化设备。成像设备是指通过光学系统或红外、激光、超声、X 射线等技术对周围场景或物体进行探测成像。之后通过数字化设备得到关于场景或物体的二维或三维数字化图像的装置。

(2) 低层视觉（Low Level Vision）

低层视觉主要对输入的原始图像进行处理。这一过程借用了大量的图像处理技术和算法，如图像滤波、图像增强、图像边缘检测、图像纹理检测等，以便从图像中抽取诸如角点、边缘、线条、边界及色彩等关于场景的基本特征。这一过程还包括了各种图像变换（如校正）、图像压缩、图像运动检测等。

(3) 中层视觉（Intermediate Level Vision）

中层视觉的主要任务是恢复场景的深度、表面法线方向、轮廓等有关场景的 2.5 维信息。实现的途径有立体视觉（Stereo Vision）、测距成像（Range Finder）、运动估计（Motion Estimation）等。系统标定等研究内容一般也在这个层次上进行。

(4) 高层视觉（High Level Vision）

高层视觉的主要任务是在以物体为中心的坐标系中，在原始输入图像、图像基本特征、

2.5 维图的基础上, 恢复出物体的完整三维图, 建立物体三维描述, 识别三维物体, 并确定物体的位置和方向等。

(5) 体系结构 (System Architecture)

体系结构这一术语最通常的含义是指在高度抽象的层次上, 根据系统模型而不是根据事先设计的具体例子来研究系统的结构。为了说明这一点, 可以考虑建筑设计中某一时期的建筑风格 (如清朝时期) 和根据这一风格设计出来的具体建筑之间的区别。体系结构研究涉及一系列相关的课题, 包括并行结构、分层结构、信息流结构、拓扑结构, 以及从设计到实现的途径。

2. 图像工程

(1) 概念

图像工程是有关图像的理论和技术, 是对图像数据的分析、管理及各种应用的总称。

1) 多年来, 图像技术有了迅猛的发展, 受到了广泛的重视和关注, 出现了许多新理论、新方法、新技术、新手段、新设备。它要求图像技术工作在一个整体的框架下进行。

2) 图像技术多年来的发展和积累, 为图像工程学科打下了坚实的基础, 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。图像工程成为整个图像领域进行研究和应用的新学科。

(2) 图像工程的内容

图像工程按照抽象程度和研究方法可分为三个层次: 图像处理、图像分析和图像理解。图像工程是三者既有联系又有区别的有机结合体。

1) 图像处理 广义图像处理泛指各种图像技术, 实际上着重强调在图像之间的变换。狭义图像处理是指图像进行各种加工, 以改善图像的视觉效果, 并为自动识别打下基础; 或对图像进行压缩编码, 以减少对所需存储空间和传输时间、传输通路的要求。图像处理研究内容包括图像采集和获取、图像重建、图像变换、图像滤波、图像增强、图像恢复或复原等。

2) 图像分析 对图像中比较感兴趣的目标进行检测和测量, 以获得它们的客观信息, 从而建立对图像的描述。图像分析的特点是一个从图像到数据的过程。数据可以是对目标特征测量的结果, 或是基于测量的符号表示。它描述了图像中目标的特点和性质。图像分析的研究内容包括图像分割、图像的目标表达和描述、图像检测与测量、图像的目标形状、图像的纹理、图像的运动分析、图像的目标识别等。

3) 图像理解 在图像分析的基础上, 进一步研究图像中各目标的性质和它们之间的相互联系, 并得出对图像内容含义的理解及对客观场景的解释, 从而指导和规划行动。图像理解的研究内容包括图像匹配、图像三维 (3D) 建模、图像客观场景恢复、图像解释和推理等。

(3) 小结

图像处理、图像分析和图像理解是处在三个抽象程度和数据量各有特点的三个不同层次上。

图像处理是底层操作, 它主要在图像的像素级上进行处理, 处理数据量大。图像处理是一个从图像到图像的过程。

图像分析是中层操作, 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

的非图像形式的描述。

图像理解主要是高层操作，基本上是对从描述抽象出来的符号进行运算，其处理方法与人类的思维推理有许多类似之处。

图 1.3 所示为图像工程的三层次示意。

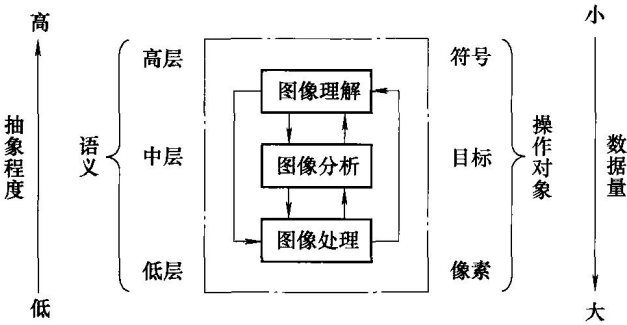


图 1.3 图像工程的三层次示意

1.1.7 计算机视觉系统的构成

如图 1.4 所示，计算机视觉系统以计算机为中心，一般由光源、图像传感器（包括光学成像系统、光电转换装置）、图像采集装置、图像处理系统（包括专用图像处理系统与计算机图像处理系统）组成。随着网络技术的发展，网络化的视觉系统已经应用到诸如大型制造业、军事、空间技术等领域，这里的信息传输系统（包括各种有线形式与无线形式）也是视觉系统的重要组成部分。在实际应用中，根据不同的任务和应用目标，各类视觉系统在构成方面会略有差异。

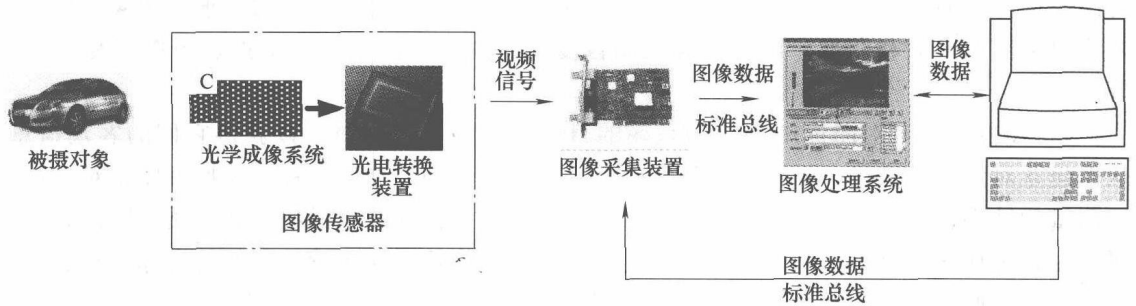


图 1.4 计算机视觉系统构成示意

1. 图像传感器

图像传感器是整个视觉检测系统的感觉部分，主要包括光学成像系统、光电转换装置和摄像机控制系统三个部分。

光学成像系统将真实物体的景象投影到摄像机的像平面上，其主要参数有焦距、视角、放大倍率、景深和线性度等，对检测系统的性能有重要影响。

摄像机控制系统接收反馈控制系统的信号，主要用来实现摄像机的精确运动控制和参数的调整。

光电转换装置将接收到的光信息转成电信号，是构成计算机视觉系统的核心部件。光电转换装置主要有电荷耦合器件（Charge Coupled Device, CCD）和互补金属氧化物半导体

(Complementary Metal-Oxide Semiconductor, CMOS) 器件。

CCD 是目前计算机视觉系统中最为常用的摄像器件。它集光电转换及电荷存储、电荷转移、信号读取于一体,是典型的固体成像器件。不同于其他器件以电流或者电压为信号,CCD 的突出特点是以电荷为信号。这类摄像器件通过光电转换形成电荷包,之后在驱动脉冲的作用下转移、放大输出图像信号。典型的 CCD 摄像机由光学镜头、时序及同步信号发生器、A-D 信号处理电路组成。CCD 作为一种功能器件在计算机视觉的发展和应用中起着至关重要的作用。

CMOS 器件的开发最早出现在 20 世纪 70 年代初。20 世纪 90 年代初期随着超大规模集成 (Very Large Scale Integrated, VLSI) 电路制造工艺技术的发展,CMOS 器件得到迅速发展。与 CCD 相比,CMOS 器件的优点可概括如下:可以实现窗口、子样和随机像素存取;无需专用驱动电路;CMOS 器件将光敏单元阵列、图像信号放大器、信号读取电路、A-D 转换电路、图像信号处理器及控制器集成在一块芯片上,因此易与信号转换和处理电路实现单片集成;功耗极低,且无需制冷;抗辐射性能好;动态范围宽;CMOS 芯片成本只相当于同类 CCD 芯片的 10% ~ 30%。可以预计,以其独特的优点,CMOS 器件在计算机视觉系统中将有广阔的应用前景。

2. 图像采集系统

图像采集系统由专用视频解码器、图像缓冲器和接口控制电路构成。它的主要功能是将摄像系统生成的视频信号实时地转换成便于计算机处理的数字图像信号。随着专用视频解码芯片的出现和现场可编程门阵列 (Field Programmable Gate Array, FPGA) 器件的普遍应用,使得图像采集装置只需要很少的几个芯片就可以完成图像数据的实时高速转换。所以,有些图像采集装置集成了显示功能,可以实现高质量图像的实时显示。一些高档图像采集装置还带有数字信号处理器 (Digital Signal Processor, DSP) 处理功能,在对图像数据进行预处理,如快速傅里叶变换 (Fast Fourier Transform, FFT) 后,再传输到计算机,使得后续的图像处理工作的运算速度得到大幅度提高。在以个人计算机 (Personal Computer, PC) 为主的视觉检测系统中,图像采集装置一般以板卡的形式出现,因而有时称其为图像采集卡。图像采集卡一般通过 64 位的外围设备接口 (Peripheral Component Interconnect, PCI) 总线实现与计算机的通信。随着标准通用接口数据传送速率的提高,近年来出现了基于通用串行总线 (Universal Serial Bus, USB) 接口的图像采集卡,具备了即插即用能力,为计算机视觉检测技术的实验研究创造了极为方便的条件。在组建视觉检测系统过程中,要注意图像采集卡允许的输入信号与摄像机的输出 (如视频信号或彩色 RGB 信号等信号类型) 的一致性,同时还要考虑采样速度、分辨率等参数。合理的性能选择与参数配置对提高视觉检测系统整体性价比具有较大的益处。

3. 专用图像处理系统

专用图像处理系统是计算机的辅助处理器,主要采用专用集成电路 (Application Specific Integrated Circuit, ASIC) 芯片、数字信号处理器 (DSP) 或者 FPGA 器件等设计的全硬件处理器。其主要功能是实时高速地完成各种底层计算机图像处理功能,例如图像的缩放、滤波、增强、边缘提取等。其目的是提高整个计算机视觉系统的处理速度。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者网络通信等形式。

4. 计算机及其图像处理系统

计算机作为整个视觉测量系统的核心，不仅要控制整个系统的各个模块的正常运行，还承担着视觉系统最后结果的运算和输出。由图像采集系统输出的数字图像可以直接传送到计算机，由计算机采用软件方式完成所有图像的处理和其他计算。如果软件处理不能够满足视觉系统的要求，则需要使用专用硬件处理系统，实时高速地完成各种低层次的图像处理算法，从而减轻计算机的处理负荷，提高整个计算机视觉测量系统的速度。

随着技术的发展，摄像系统、图像采集系统已经可以实现高度集成，例如各种数字摄像机的出现就是一个实例。而且对于一个实用的计算机视觉系统而言，其结构、性能、处理时间和价格等均应该根据具体应用情况进行灵活的选择。

1.1.8 计算机视觉技术的应用领域

随着技术的发展，计算机视觉技术已广泛应用于工业自动化生产线、各类检验和监视、视觉导航、图像自动解释、人机交互与虚拟现实等各个领域。

1. 在产品的分类与检测

计算机视觉中的应用技术已成功地应用于工、农、林等各个领域的产品分类检测，不但大幅提高了产品的质量和可靠性，而且提高了生产效率。

在工业生产中，计算机视觉技术可以自动检测各种机械零件的几何尺寸，测量其精度范围，目前已经用于产品外形检验、表面缺陷检验等。例如，对滑块及滑槽外形的检验及装配后位置的检验，可以决定它们能否装配在一起，并且准确无误地完成装配任务；对发动机内壁麻点、刻痕等缺陷的检查，可以决定产品的质量；通过 X 射线照相或超声探测获取物体内部的图像，可以实现内部缺陷检验，如钢梁内部裂纹和气孔等缺陷检验。另外，计算机视觉还可对零部件文字标识进行检测与识别。序列号由数字、字母和汉字等组成，以数字和字母居多，作为零部件的惟一身份标志，通常打印在电子元器件、集成电路板、大型工件成品的外表面上。这些序列标号或编码，如采用人工方式用眼睛对这些字符进行观察识别，将不能满足高效率、自动化生产的要求，并且字符在工件上的位置很多时候不便于人眼观察，人眼疲劳因素也会使检测结果不能从根本上得到保证。采用计算机视觉、模式识别、智能控制等技术可以有效地克服上述缺点，完成文字标识的自动检测与识别，提高生产效率，实现生产过程自动化。

在纺织、印染业中，计算机视觉技术可以完成自动分色、配色和产品成分与瑕疵检验等工作。例如，利用计算机视觉来进行羊毛的细度和羊绒羊毛的纤维含量分析，或对织造布、非织造布、丝绸的质量和工艺作出分析。

计算机视觉在农林业生产中的应用也日益广泛，主要表现在：

1) 利用计算机视觉技术根据检测到的农产品的颜色、纹理等信息对其进行分级、检验和分类。例如，根据检测到的粒型、至白度等参数，将稻米分为不同等级；根据检测到的农作物的叶片信息，可以对不同的农作物进行分类；根据检测到的颜色、形状等信息，对苹果、柑橘等水果进行分类。

2) 利用计算机视觉技术进行病虫害的监测。例如，根据检测到的棉花叶片的图像信息，用轮廓跟踪的方法确定棉花叶片上的虫蚀痕迹，对棉花生长过程中棉铃虫虫害的发病程度进行诊断。

3) 利用计算机视觉技术对育苗、栽培、管理和收获等各个环节进行自动化管理等。例