

# 视觉测量

张广军 著

科学出版社

北 京

## 内 容 简 介

从计算机视觉概念和方法出发,将计算机视觉应用于空间几何尺寸的精确测量和定位,从而产生了一种新的计算机视觉应用概念——视觉测量。本书是作者研究组 15 年来从事视觉测量研究工作的总结和提炼,系统地介绍了视觉测量的基础原理、测量方法、关键技术与实用算法,并给出了视觉测量系统实例。本书涉及视觉测量中的主要研究内容,包括空间几何变换与摄像机模型,视觉图像特征信息提取,典型算法硬件 IP 核设计,摄像机标定,双目立体视觉测量,结构光三维视觉测量,多传感器三维视觉测量,流动式三维视觉测量,以及三个典型视觉测量系统。

本书可作为测量与控制、自动化、计算机、机器人及人工智能等专业的高年级本科生和研究生的教材,也可供从事计算机视觉和相关专业研究工作的技术人员参考。

### 图书在版编目(CIP)数据

视觉测量/张广军著. —北京:科学出版社,2008

ISBN 978-7-03-020923-8

I. 视… II. 张… III. 计算机视觉-测量 IV. TP302.7

中国版本图书馆 CIP 数据核字(2008)第 012034 号

责任编辑:马长芳 潘继敏 / 责任校对:陈丽珠  
责任印制:张克忠 / 封面设计:耕者设计工作室

科学出版社 出版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

中国科学院印刷厂印刷

科学出版社发行 各地新华书店经销

\*

2008 年 3 月第 一 版 开本:B5(720×1000)

2008 年 3 月第一次印刷 印张:21 3/4

印数:1—4 000 字数:410 000

定价:38.00 元

(如有印装质量问题,我社负责调换(科印))

## 序

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,通过一幅或多幅图像认知周围环境信息。从计算机视觉概念和方法出发,将计算机视觉应用于空间几何尺寸的精确测量和定位,从而产生了一种新的计算机视觉应用概念——视觉测量。它以其非接触性、测量精度高及响应速度快等特点,在非接触在线测量、质量监控与运动分析中具有广阔的应用前景,其应用正逐渐渗透到航空航天、军事装备、生物医药、目标识别、装备制造、工业测量和自动化控制等诸多领域。美国三大汽车公司相继与美国 Michigan 大学和 Perceptron 公司合作,研制出用于轿车车身、分总成和总成产品尺寸的在线自动测量的视觉测量系统,还用于生产线上机器手的定位与瞄准,提高了汽车产品质量和生产效率。天津大学精密测试技术及仪器国家重点实验室在国内率先开展了视觉测量与应用研究,将机器视觉与精密测量有机结合,先后研制成功车身三维尺寸自动检测系统和机器人柔性测量系统,为国内多个汽车厂家所采用。

视觉测量是以计算机视觉理论为基础,涉及光学、光电子学、图像处理、模式识别、信号与数据处理及计算机技术等诸多学科。目前,信号处理理论与超大规模集成电路(VLSI)技术的快速发展为视觉测量信号的处理提供了新的手段,为实时在线视觉测量系统的构建奠定了基础。

多年来,作者带领其研究组一直致力于钻研视觉测量,在理论方法、关键技术、系统装置及工程应用等方面开展了系统深入地研究,取得了多项创新性研究成果。本书是作者及其研究组多年来从事视觉测量研究工作的总结和提炼,是其研究成果的结晶,对国内视觉测量及相关专业研究具有重要参考价值。

中国工程院院士



2007 年 10 月于天津大学

# 前 言

视觉是人类观察世界和认知世界的重要手段,使计算机或机器人具有视觉是人类多年以来的梦想。随着信号处理理论和电子、计算机等相关技术的发展,计算机视觉得到了迅速发展。从计算机视觉概念和方法出发,将计算机视觉应用于空间几何尺寸的精确测量和定位,从而产生了一种新的计算机视觉应用概念——视觉测量。视觉测量作为当今高新技术之一,在电子学、光电探测、图像处理和计算机技术不断成熟和完善的基础上得到了突飞猛进的发展,并有着广泛的应用,如应用于产品在线质量监控、微电子器件(IC 芯片、PCB 板、BGA 封装)的自动检测、各种模具三维形状的测量及生产线中机械手的定位与瞄准等,所以很多发达国家都竞相发展这项高新技术。

视觉测量涉及光学、光电子学、图像处理、模式识别、信号与数据处理及计算机技术等诸多学科领域,内容广泛。本书在作者 2005 年 6 月出版的研究生教材用书《机器视觉》基础上,融入了作者研究组近年来在视觉测量研究方面的最新研究成果,是作者研究组 15 年来从事视觉测量研究工作的总结和提炼,并获得国防科工委国防专著出版基金资助,以专著形式由科学出版社出版。

本书涉及视觉测量中的主要研究内容,各章节按照由视觉测量模型到视觉测量系统、由视觉测量算法到视觉测量应用、由单传感器视觉测量到多传感器视觉测量、由视觉测量关键技术到视觉测量系统集成的顺序来安排撰写,体现了从简到繁、由浅入深、从理论到实际、从技术到系统的特点,力求具有层次性、系统性、先进性和实用性。全书共分 12 章,第 1 章为引论,第 2 章介绍了空间几何变换与摄像机模型,第 3 章介绍了视觉图像特征信息提取,第 4 章介绍了典型算法硬件 IP 核设计,第 5 章介绍了摄像机标定,第 6 章介绍了双目立体视觉测量,第 7 章介绍了结构光三维视觉测量,第 8 章介绍了多传感器三维视觉测量,第 9 章介绍了流动式三维视觉测量,第 10 章介绍了微小构件内表面测量系统,第 11 章介绍了轮胎综合几何参数测量系统,第 12 章介绍了高扇翅频昆虫运动参数测量系统。

15 年来,作者研究组在视觉测量研究工作中获得了多项研究基金与计划资助,包括国家杰出青年科学基金、国家自然科学基金、国家 863 计划、民用航天重大研究项目、国防基础科研项目、教育部跨(新)世纪优秀人才培养计划、北京市科技新星人才培养计划、航空科学基金及企业委托项目等 20 余项研究项目。作者就对这些研究基金和研究计划给予支持和资助的国家自然科学基金委员会、科技部 863 计划联合办公室、国防科工委科技质量司、教育部科学技术司、北京市科学技

术委员会、航空科学基金办公室等有关部门表示感谢。

本书是作者研究组多年来从事视觉测量的研究成果的结晶,作者十分感谢研究组的周富强教授、魏振忠副教授、江洁副教授、孙军华博士、魏新国博士,及贺俊吉、王颖、陈大志、李秀智、李海超、樊巧云、李苏祺、雷鸣等同学,本书介绍的许多工作是由他们具体完成的。

在此书撰写完成之际,作者要特别感谢自己的恩师——天津大学叶声华院士,感谢他对作者 20 年来一如既往的关心、指导和帮助。作者也十分感谢天津大学精密测试技术及仪器国家重点实验室的师兄师弟,在与他们的学术交流中受益匪浅。

本书参考和引用的参考文献与研究成果已在文中列出或说明,对相关内容感兴趣的读者可直接查阅。

视觉测量内容较新,且十分广泛,涉及诸多学科领域。由于作者水平有限,经验不足,书中不妥之处在所难免,敬请广大读者、同行与专家批评指正。

张广军

2007 年 8 月于北京航空航天大学

# 目 录

第 1 章 引论	1
1.1 计算机视觉的发展与系统构成	1
1.2 生物视觉简介	4
1.3 Marr 视觉理论框架	9
1.4 计算机视觉应用领域及面临问题	12
1.5 视觉测量系统与关键技术	15
1.6 本书各章内容简介	20
参考文献	21
第 2 章 空间几何变换与摄像机模型	23
2.1 空间几何变换	23
2.2 几何变换的不变量	27
2.3 欧氏空间的刚体变换	30
2.4 摄像机透视投影模型	33
2.5 摄像机透视投影近似模型	36
参考文献	41
第 3 章 视觉图像特征信息提取	42
3.1 图像边缘与图像平滑	42
3.2 改进的 Steger 图像边缘检测算法	46
3.3 Harris 角点探测器	52
3.4 X 型角点子像素级提取	55
3.5 栅格型角点子像素级提取	61
3.6 椭圆形图像中心的提取	66
3.7 空间椭圆中心图像位置的提取	70
参考文献	76
第 4 章 典型算法硬件 IP 核设计	78
4.1 高斯滤波 IP 核设计	79
4.2 角点探测器 IP 核设计	83
4.3 光斑图像中心提取 IP 核设计	90
4.4 改进的 Steger 算法 IP 核设计	93

参考文献	101
<b>第 5 章 摄像机标定</b>	102
5.1 基于 3D 立体靶标的摄像机标定	102
5.2 基于径向约束的摄像机标定	106
5.3 基于 2D 平面靶标的摄像机标定	110
5.4 基于交比不变的摄像机标定	116
5.5 基于卡尔曼滤波的摄像机标定	120
5.6 机器人手眼标定	125
5.7 机器人足目标定	128
参考文献	132
<b>第 6 章 双目立体视觉测量</b>	134
6.1 测量原理与数学模型	134
6.2 测量系统精度分析	137
6.3 测量系统结构设计	140
6.4 极线几何与基本矩阵	144
6.5 两幅图像对应点匹配	148
6.6 基于角点引导的边缘匹配	159
6.7 测量系统标定方法	163
6.8 光笔式三坐标测量机	166
参考文献	173
<b>第 7 章 结构光三维视觉测量</b>	175
7.1 测量原理与数学模型	175
7.2 结构光投射器	185
7.3 测量系统常规标定方法	189
7.4 基于双重交比不变的标定方法	193
7.5 基于自由移动平面靶标的标定方法	197
7.6 基于神经网络的标定方法	201
参考文献	205
<b>第 8 章 多传感器三维视觉测量</b>	207
8.1 测量系统概述	207
8.2 全局标定方法	213
8.3 基于双经纬仪的全局标定	220
8.4 基于单经纬仪的全局标定	227
8.5 全局标定精度分析与评价	231

---

参考文献	237
<b>第 9 章 流动式三维视觉测量</b>	238
9.1 光栅条纹编码识别	238
9.2 基于平面基线靶标的测量数据拼接	242
9.3 基于平面靶标三维点的测量数据拼接	253
9.4 拼接实验与精度分析	257
参考文献	263
<b>第 10 章 微小构件内表面测量系统</b>	264
10.1 概述	264
10.2 系统硬件与软件构成	268
10.3 图像特征分析与提取	273
10.4 系统标定与测量应用	280
参考文献	285
<b>第 11 章 轮胎综合几何参数测量系统</b>	286
11.1 概述	286
11.2 轮胎光条图像特征信息提取	289
11.3 基于单传感器的静态测量系统	292
11.4 基于双传感器的动态测量系统	297
参考文献	302
<b>第 12 章 高扇翅频昆虫运动参数测量系统</b>	303
12.1 概述	303
12.2 系统结构设计 with 构成	309
12.3 系统标定与精度评价	317
12.4 昆虫运动参数测量	322
参考文献	335



# 第 1 章 引 论

视觉是人类观察世界和认知世界的重要手段,人类通过眼睛和大脑来获取、处理与理解视觉信息。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,由感光细胞转换成神经脉冲信号,经神经纤维传送入大脑皮层进行处理与理解。所以,视觉不仅指对光信号的感受,还包括对视觉信息的获取、传输、处理与理解的全过程。

使计算机或机器人具有视觉是人类多年以来的梦想。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(computer vision)。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。它不仅在于模拟人眼能完成的功能,更重要的是它能完成人眼所不能胜任的工作。

从计算机视觉概念和方法出发,将计算机视觉应用于空间几何尺寸的精确测量和定位,从而产生了一种新的计算机视觉应用概念——视觉测量(vision measurement)。视觉测量作为当今高新技术之一,在电子学、光电探测、图像处理和计算机技术不断成熟和完善的基础上得到了突飞猛进的发展,并有着广泛的应用,如应用于产品在线质量监控、微电子器件(IC 芯片、PCB 板、BGA 封装)的自动检测、各种模具三维形状的测量及生产线中机械手的定位与瞄准等,所以各先进国家都竞相发展这项高新技术。另外,由多个视觉传感器可以组建一个柔性的空间三坐标测量站(或称为多传感器视觉测量系统),以完成对大型物体的三维空间尺寸的全自动实时测量。

本章首先讨论计算机视觉的发展过程及系统构成,简介生物视觉,然后介绍 Marr 视觉理论框架,并阐述计算机视觉应用领域及其面临问题,最后对视觉测量系统与关键技术,以及本书各章内容进行介绍。

## 1.1 计算机视觉的发展与系统构成

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的巨大,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。随着技术的发展,将人类视觉功能赋予计算机、机器人或其他智能机器是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,使得人类的梦想正在逐步实现。

### 1.1.1 计算机视觉的发展

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,如字符识别及工件表面、显微图片和航空照片的分析和解释等。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<sup>[1]</sup>。Roberts 的研究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。到 70 年代,已经出现了一些计算机视觉应用系统<sup>[2,3]</sup>。

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<sup>[4]</sup>。到了 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 90 年代,计算机视觉在工业环境中得到广泛应用<sup>[5~7]</sup>,同时基于多视几何的视觉理论也得到迅速发展<sup>[8]</sup>。

### 1.1.2 计算机视觉系统构成

计算机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构成,如图 1.1 所示。

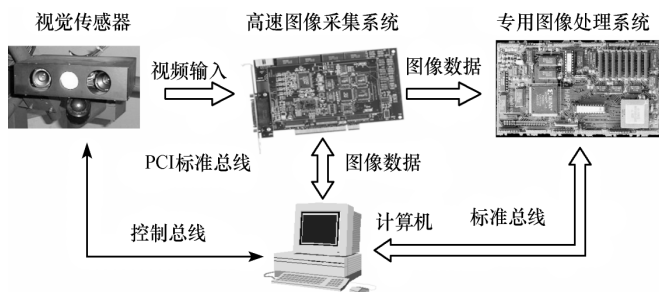


图 1.1 计算机视觉系统基本组成模块

视觉传感器是整个计算机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计算机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长,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。与 CCD 相比,CMOS 图像传感器的优点可以概括如下:①可以实现窗口、子样和随机像素存取;②无需专用驱动电路;③易与信号转换和处理电路实现单片集成;④功耗极低,且无需制冷;⑤抗辐射性能好;⑥动态范围宽;⑦芯片成本只相当于同类 CCD 芯片的 10%~30%。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、图像缓冲器以及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,或者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只需由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 总线或者 PCI 总线等,使得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)或者 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,减轻后端计算机的处理负荷,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者网络通信等方式。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,为计算机视觉系统实时实现提供了有利的条件。

计算机是整个计算机视觉系统的核心,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,就不需专用硬件处理系统出现在计算机视觉系统。这样,一个实用计算机视觉系统的结构、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。

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 机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 年底就有 100 多家公司跻身于计算机视觉系统的经营市场,而到现在 20 多年的发展,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技术日益成熟,越来越多的公司能生产出更小、更先进、更灵活可靠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。

随着计算机视觉的飞速发展,二维视觉处理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 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,而且通过采用一定的算法,视觉系统精度受照明变化的影响很小。灰度是图像辐射度或亮度的量化测量。该信息是通过视频 A/D 转换器存储在帧存体中获得的,灰度分辨率随计算机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”能力在计算机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照明视觉和被动照明视觉。前者需要利用特别的光源所提供的结构信息,如结构光视觉;而后者是在自然光下完成的,如双目立体视觉。被动照明视觉适合于受环境限制和自然光下的场合。而主动照明视觉可应用的领域非常广泛,且抗干扰性能好、实时性强。总之,三维视觉的引入进一步扩大了计算机视觉的应用领域。

## 1.2 生物视觉简介

了解生物视觉的构成、信息处理过程与特性,对研究计算机视觉研究人员来讲是非常有启发性和吸引力的。因此,下面在参考相关文献<sup>[9,10]</sup>的基础上对生物视觉通路、感受野及生物视觉信息处理特性作一简要介绍。

### 1. 生物视觉通路

图 1.2 示意了生物视觉通路。物体在可见光照射下经眼的光学系统在眼底视

网膜上形成物像,由杆体和锥体感光细胞转换为神经信号,并经视网膜中的神经节细胞(ganglion cell,GC)加工后传出视网膜。经神经节细胞加工的神经信号,经过视交叉部分地交换神经纤维后,再形成视束,传到中枢神经的许多部分,包括丘脑的外膝体(lateral geniculate nucleus, LGN)、上丘(superior colliculus)和视皮层(visual cortex)。上丘只与眼动等视觉反射有关,外膝体和视皮层直接与视知觉有关。神经节细胞轴突在外膝体换神经元后,由外膝体神经元直接经视放线到视皮层,这是视束的大部分纤维去向,称为视觉第一视通路。视束的一小部分纤维走向内方,经上丘臂到达上丘和顶盖前区。上丘浅层神经元透射到丘脑枕换元后,再透射到视皮层,上丘还有纤维直接透射到视皮层。由于这条通路不经过外膝体,故称为视觉第二视通路。神经信号主要是通过视觉第一视通路到达视皮层。因此,第一视觉视通路在视觉过程中起着主要作用,而第二视觉视通路的作用极其微小。

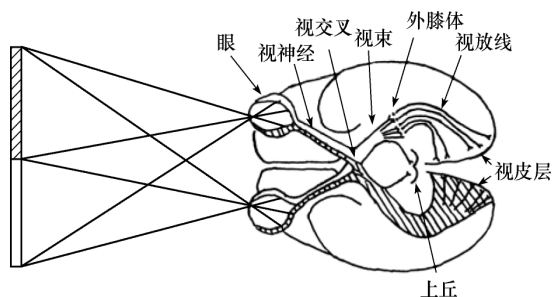


图 1.2 生物视觉视通路示意图

由视觉视通路可以看到,眼、外膝体与视皮层构成了对视觉信息处理的三个基本层次。进一步分析表明,外膝体与视皮层,尤其是视皮层还有更为复杂的分块结构。分块表明了视觉信息处理的并行性,不同区域的神经细胞具有不同的功能;分层表明了视觉信息处理的串行性。因此,生物视觉系统是一个串行与并行处理相结合的复杂系统。

## 2. 感受野的分层等级假设

视觉通路上各层次的神经细胞,由简单到复杂,它们所处理的信息,分别对应于视网膜上的一个局部区域,层次越深入,该区域就越大,这就是著名的感受野(receptive field)与感受野等级假设。感受野是支持视觉信息分层串行处理的最重要的生理学证据。

以信息处理的第一级为例,视网膜上的神经节细胞将感光细胞上接收到的光信号转换成电信号再由它的轴突传出,但每一个 GC 细胞只能接收视网膜上一个局部区域的信号,该区域就是 GC 的感受野。研究表明,GC 感受野及其对光信号

的转换作用可划分为以下几种：

(1) 对空间亮度变化敏感的感受野,形状可用两个同心圆表示。如图 1.3 所示,这种同心圆形状的感受野按其对光信号的转换作用又可以分为中心兴奋区、周边抑制区组成的 on 型,及中心抑制区、周边兴奋区组成的 off 型。图中还画出了这两种感受野对光信号的径向截面响应曲线,该曲线近似于两个方差不同的高斯函数的差。具有这种感受野的 GC 细胞对于在同心圆区域受到均匀光照的光刺激的反应,为响应曲线的积分。一般,GC 的输出为响应曲线与光信号乘积的积分。图 1.4 表示当视网膜上光信号为一边亮一边暗的具有一定对比度的信号时,感受野位于不同空间位置的 GC 的输出。只有当亮暗边缘线过同心圆中心时,GC 的输出与感受野受到均匀光照时一样,设为  $E$ ,而当边缘线位于同心圆其他位置时,输出分别高于或低于该平均输出  $E$ 。如将输出看作实际输出减去平均输出  $E$ ,则当亮暗边缘线过感受野同心圆中心时,输出为零。可见,由 GC 的输出与感受野的位置可以检测亮暗边缘线。

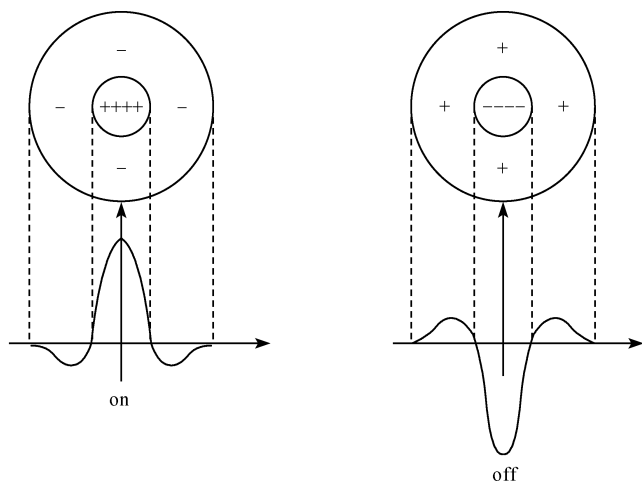


图 1.3 on 型与 off 型感受野

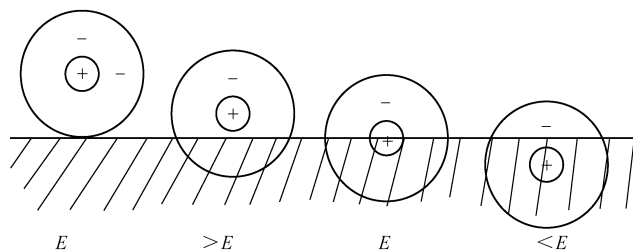


图 1.4 不同位置感受野的 GC 细胞对具有黑白对比度的光信号的响应

(2) 除具有上述响应特征及感受野的 GC 细胞外(这种 GC 细胞也称为 x 型细胞),还有一种 Y 型细胞。它的响应不对亮度的空间变化敏感,而是对时间变化敏感,即当感受野的圆形区域内的亮度随时间变化时,GC 的输出会增大或减少。y 细胞也可划分为 on 型或 off 型,这种局部亮度随时间变化敏感的性质是物体运动分析的基础。

以上介绍的是视网膜 GC 细胞的感受野,这是视觉信息处理第一层次的感受野。对于更高层次,包括外膝体细胞与初级视皮层(即视皮层的前几层),也发现了类似的性质,即每一个单个细胞只接受视网膜上的局部信息,但层次越高感受野越大,即信息处理是从局部到更大的区域的。例如外膝体上的一个细胞,可以接受来自空间感受野相邻的多个 GC 细胞的信息,由于每个 GC 细胞有一个感受野,多个 GC 细胞的感受野就组成了一个更大的区域,而且该区域也具有同心圆的形状。这种感受野从局部到整体的层次结构就是著名的感受野层次等级假设<sup>[1]</sup>。

### 3. 视觉信息的并行处理

感受野的等级假设与局部性质主要支持视觉信息处理的自下而上的分层次串行处理性质。视网膜、外膝体和视皮层构成对视觉信息处理的多级串行处理,特别是视皮层,表现出更为复杂的多级分层。然而视觉系统的任务不是单一的,它要识别物体的形状和颜色,要得到三维物体的深度信息,要检测物体的方位和运动参数。另外,物体的空间和时间频率性质也有很大的差异,例如,较细的表面纹理表现了物体表面较高的空间频率特征;较快的运动,表现为图像较高的时间频率特征。神经生理学的研究表明,视觉通路的各个层次上存在着基本互相独立的并行通道,分别完成不同的视觉任务,下面简要介绍几个不同的并行通道:

(1) X、Y 和 W 通道。在视网膜神经节细胞中,一类神经节细胞的感受野的兴奋和抑制作用可以线性相加,称为 X 细胞,另一类神经节细胞的空间总和性质是非线性的,称为 Y 细胞。外膝体的神经元也可按其空间总和性质划分为 X 和 Y 细胞,并且在传递信息过程中,X 型神经节细胞总是与 X 型外膝体神经元发生联系,Y 型神经节细胞总是与 Y 型外膝体神经元发生联系。X 型外膝体神经元多数传至视皮层简单细胞和超级复杂细胞,而多数 Y 型外膝体神经元只传给视皮层复杂细胞。由此可见,视觉系统内存在一个 X 和 Y 通道,它们在功能上表现出 X 细胞的感受野可能与空间信息的检测与传递有关,而 Y 细胞的感受野可能与时间信息的检测与传递有关。除了 X 和 Y 细胞外,人们在猴、猫视网膜上还发现了一种称为 W 型的神经节细胞,其感受野与 Y 细胞大小相仿,但轴突直径特别细,因而动作电位在其轴突上传导速度也最慢。W 细胞的轴突主要传至中脑上丘部,是控制眼球运动的。

(2) on 型和 off 型通道。在视网膜上,on 型和 off 型细胞是一种均匀镶嵌式

的排列,其总数基本相等,而在外膝体,它们开始呈现一定程度的分离。实验证明, on 型通道和 off 型通道在外膝体到视皮层是充分地平行分离的。例如,在猴视网膜水平细胞、双极细胞处用药物选择性地阻断 on 型通道,可以取消神经节细胞、外膝体和视皮层的 on 型反应,但对 off 型细胞的反应和视皮层细胞方位、方向选择却毫无影响。

(3) 空间与时间频率通道。X、Y 型细胞分别对具有空间频率或时间频率的信息敏感。心理物理实验证明,外膝体上的某一个神经细胞并不对所有频率的信息敏感,而是对某一频段的信息呈现较强的反应。在视皮层也发现了类似的证据,即视皮层的某一部分只对某一定频率的信号敏感。这些实验证据表明,在视觉通路中存在处理不同频道的信号的独立通道。

(4) 颜色信息通道。在视网膜上就存在不同的感光细胞,即锥状细胞与杆状细胞,其中杆状细胞对颜色不敏感,而锥状细胞又分为不同光谱(红、绿、蓝)敏感的三种细胞。对外膝体与视皮层 17 区的分析也表明,它们都有专门的区域从事颜色信息的处理与识别。

(5) 左右眼信息通道与立体视觉。由图 1.2 的视觉通道示意图中可见,眼睛、外膝体与视皮层都有左右两侧。分析表明,左右两侧的神经细胞分别处理由左右两眼的半侧来的信息,也就是说,每一个眼睛的左右侧视野的信息是交叉地分别投射到左右侧外膝体与视皮层的。视皮层处理来自两眼的信息得到双眼视差信息是立体视觉的基础,也就是说,只有比较来自双眼的信息,才能使我们有深度感。最近的研究表明,一直到视皮层的 17,18,19 区,还存在独立处理不同视差的并行通道。

(6) 空间方位信息通道。早期对视觉通路信息处理的研究认为,视皮层细胞对空间几何元素(如直线)的方位敏感。相关研究表明,这种方位敏感性也存在于外膝体细胞,即具有相似最优方位敏感的细胞在外膝体层次已经聚集在一起,且具有与视网膜神经元细胞相似的、向心的最优方位分布规律。可见在视皮层处理之前,外膝体已经对方位信息进行了组织,在视觉通路中存在着处理方位信息的通道。因此,视网膜、外膝体和视皮层形成了一个空间方位信息通道。

(7) 视皮层对形状、颜色、运动与深度信息的并行处理。对视皮层 17 区至更高层次的研究表明,对物体的形状、颜色、运动、深度等不同视觉信息处理已明显分离开,对视皮层不同区域分别处理什么信息已有大量的研究,最近发展的功能核磁共振(FMRI)与高分辨率脑地形图的分析视皮层各部分的功能提供了技术手段。

以上介绍表明,视觉信息处理是一个串行与并行相结合的复杂信息处理过程,目前被人们所认知的仅仅是极小的一部分。从信息处理的角度看,我们对大多数处理单元的知识还非常有限,只知道这个单元对某种信息“敏感”,而信息是如何表征的、如何变换的,则仍不清楚,尤其是较高层信息的处理。关于串行与并行处理



进入高级皮层后,信息是如何综合的,更是很有争论的问题。还有一些认知心理学实验,观察到与上述有些论述完全不相容的现象。

## 1.3 Marr 视觉理论框架

20 世纪 80 年代初, Marr 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<sup>[4]</sup>。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<sup>[4,12]</sup>。

### 1.3.1 视觉系统研究的三个层次

Marr 从信息处理系统的角度出发,认为对此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、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变换成另一种表示。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。最后一个硬件层次,是解决用硬件实现上述表达和算法的问题,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。

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以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可归纳为表 1.1。

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

1.3.2 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了三个阶段的处理,如图 1.5 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表 1.2 所示。

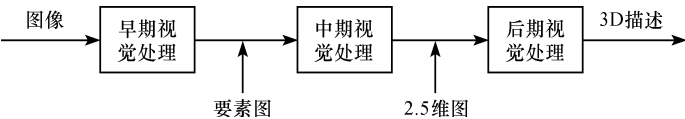


图 1.5 Marr 框架的视觉三阶段

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

续表

名 称	目 的	基 元
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

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的研究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(见图 1.5)的问题<sup>[13]</sup>:

- 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 1.5 的框架,可将其改进并融入新的模块得到图 1.6 的框架,具体改进如下。

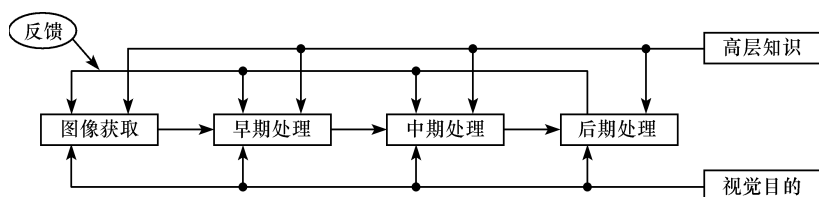


图 1.6 改进的 Marr 框架

(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<sup>[14]</sup>,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<sup>[15]</sup>。

(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<sup>[12]</sup>。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,多数被看作它的补充和发展。尽管 Marr 的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。

## 1.4 计算机视觉应用领域及面临的问题

计算机视觉在国民经济、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,视觉方式所能观测的对象十分广泛,人眼观察不到的范围,计算机视觉也可以观察,例如,红外线、微波、超声波等人类就观察不到,而计算机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,计算机视觉则不知疲劳,始终如一地观测,所以计算机视觉可以广泛地用于长时间恶劣的工作环境。不过,计算机视觉技术仍处于十分不成熟的阶段,其发展还远远落后于人们所寄予的发展水平。

### 1.4.1 计算机视觉应用领域

计算机视觉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、人机交互及虚拟现实等领域。

(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。

(3) 视觉导航应用。巡航导弹制导、无人驾驶飞机飞行、自动行驶车辆、移动机器人、精确制导及自动巡航捕获目标和确定距离等,既可避免人的参与及由此带来的危险,也可提高精度和速度。

(4) 图像自动解释应用。对放射图像、显微图像、医学图像、遥感多波段图像、合成孔径雷达图像、航天航测图像等的自动判读理解。由于近年来技术的发展,图像的种类和数量飞速增长,图像的自动理解已成为解决信息膨胀问题的重要手段。

(5) 人机交互应用。人脸识别、智能代理等。同时让计算机可借助人的手势动作(手语)、嘴唇动作(唇读)、躯干运动(步态)、表情测定等了解人的愿望要求而执行指令,这既符合人类的交互习惯,也可增加交互方便性和临场感等。

(6) 虚拟现实应用。飞机驾驶员训练、医学手术模拟、场景建模、战场环境表示等,它可帮助人们超越人的生理极限,“亲临其境”,提高工作效率。

计算机视觉的应用是多方面的,它已经取得并将继续取得越来越广泛的应用。

### 1.4.2 计算机视觉面临的问题

对于人的视觉来说,由于人的大脑和神经的高度发展其目标识别能力很强。但是人的视觉也同样存在障碍,例如,即使具有一双敏锐视觉和极为高度发达头脑的人一旦置于某种特殊环境(即使曾经具备一定的先验知识),其目标识别能力也会急剧下降。事实上人们在这种环境下在对简单物体时,仍然可以有效而简便地识别,而在对复杂目标或特殊背景时则在视觉功能上发生障碍,两者共同的结果是导致目标识别的有效性和可靠性的大幅度下降。将人的视觉引入计算机视觉中,计算机视觉也存在着这样的障碍。它主要表现在三个方面<sup>[16]</sup>:一是如何准确、高速(实时)地识别出目标;二是如何有效地增大存储容量,以便容纳下足够细节的目标图像;三是如何有效地构造和组织出可靠的识别算法,并且顺利地实现。前两者相当于人的大脑这样的物质基础,这期待着高速的阵列处理单元,以及算法(如神

经网络、分维算法、小波变换等算法)的新突破,用极少的计算量以及高度地并行性实现功能。

另外,由于当前对人类视觉系统的机理、人脑心理和生理的研究还不够,目前人们所建立的各种视觉系统极大多数是只适用于某一特定环境或应用场合的专用系统,而要建立一个可与人类的视觉系统相比拟的通用视觉系统是非常困难的。主要原因有以下几点:

(1) 图像对景物的约束不充分。首先是图像本身不能提供足够的信息来恢复景物,其次是当把三维景物投影成二维图像时丧失了深度信息。因此,需要附加的约束才能解决从图像恢复景物时的多义性。

(2) 多种因素在图像中相互混淆。物体的外表受材料的性质、空气条件、光源角度、背景光照、摄像机角度和特性等因素的影响。所有这些因素都归结到一个单一的测量,即像素的灰度。要确定各种因素对像素灰度的作用大小是很困难的。

(3) 理解自然景物要求大量知识。例如,要用到阴影、纹理、立体视觉、物体大小的知识;关于物体的专门知识或通用知识,可能还有关于物体间关系的知识等,由于所需的知识量极大,难以简单地用人工进行输入,可能要求通过自动知识获取方法来建立。

(4) 人类虽然自己就是视觉的专家,但它又不同于人的问题求解过程,难以说出自己是如何看见事物,从而给计算机视觉的研究提供直接的指导。

人类视觉系统具有如下特点:具有高分辨率,有立体观察、优越的识别能力和灵活的推理能力,可灵活地根据各种视觉线索进行推理。为了便于理解,现将人的视觉与计算机视觉对比列于表 1.3、表 1.4。视觉机理的复杂深奥使有些学者不禁感叹道:如果不是因为有人的视觉系统作为通用视觉系统的实例存在的话,甚至都怀疑能不能找到建立通用视觉系统的途径。正因如此,赋予机器以人类视觉功能是几十年来人们不懈追求的奋斗目标。

表 1.3 计算机视觉与人的视觉能力比较

能 力	计算机视觉	人的视觉
测距	能力有限	定量估计
定方向	定量计算	定量估计
运动分析	定量分析,但受限制	定量分析
检测边界区域	对噪声比较敏感	定量、定性分析
图像形状	受分割、噪声制约	高度发达
图像机构	需要专用软件,能力有限	高度发达
阴影	初级水平	高度发达
二维解释	对分割完善的目标能较好解释	高度发达
三维解释	较为低级	高度发达
总的能力	最适合于结构环境的定量测量	最适合于复杂的、非结构化环境的定量解释

表 1.4 计算机视觉与人的视觉性能标准比较

性能标准	计算机视觉	人的视觉
分辨率	能力有限	定量估计
处理速度	零点几秒/每帧图像	定量估计
处理方式	串行处理,部分并行处理	每只眼睛每秒处理(实时) $10^{10}$ 空间数据
视觉功能	二维、三维立体视觉有限	自然形式三维立体视觉
感光范围	紫外线、红外线、可见光	可见光

从长远来看,建立人类视觉的计算理论,并进而建成可与人类视觉系统相比拟的通用视觉系统是计算机视觉研究的最终目标。人类视觉的研究涉及神经生理学、心理物理学、心理学等多方面,对其机理的了解可为建立视觉的计算理论提供有益的启示。与此同时,视觉计算理论的研究又促进了在上述领域中引入计算机,这又推动了这些学科自身的发展。强调计算机视觉研究与人类视觉研究之间的密切关系,并不意味着计算机视觉系统要完全模仿人类视觉系统,而是通过对人类视觉系统的研究发现是什么因素使人类视觉系统的性能如此之好,并且把它结合到计算机视觉系统中去。

## 1.5 视觉测量系统与关键技术

视觉测量是以计算机视觉为理论基础,采用先进的高密度、低噪声和畸变小的图像传感器,通过高速实时图像采集系统,专用图像硬件处理系统以及高性能计算机完成对二值或灰度图像的有效处理的先进系统。视觉测量任务对生产线上的工件或装配件进行测量以确保它们与预置标准量的相符程度,具有非接触、动态响应快、大量程、高效率、全自动等优点,广泛应用于制造加工工业中。

### 1.5.1 视觉测量系统

根据测量对象和测量任务的要求,视觉测量系统集成到加工制造生产线或者生产作业设备中。在实际生产线和作业线上,首先实现视觉测量系统的现场标定和量值传递,在此基础上完成被测物体的综合几何参数测量。尽管视觉测量系统的具体结构不尽相同,但其原理框图如图 1.7 所示。它主要由视觉传感器、高速视觉图像采集系统、高速视觉图像专用硬件处理系统、高级算法硬件处理系统、现场校准装置及软件、测量与控制软件及机械结构与电气控制系统等组成。视觉传感器获取被测物体表面特征图像,经高速视觉图像采集系统转换为数字信号,由高速视觉图像专用硬件处理系统完成视觉数字图像的高速底层处理,并提取出特征信息的图像坐标。根据视觉测量数学模型,可由高级算法硬件处理系统直接快速计算出被测物体的特征空间坐标,最终得到被测物体空间几何参数和位置姿态等参

数,也可由上位计算机实现被测物体空间几何参数和位置姿态等参数的快速计算,并完成相关的系统控制。

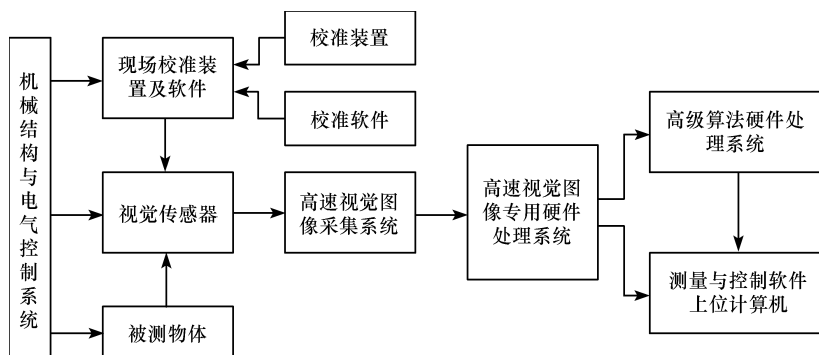


图 1.7 视觉测量系统原理框图

工作时,被测工件处于生产线上,视觉测量系统安装在生产线的固定工位;或者被监测工件静止,而视觉测量系统集成在处于运动中的生产作业设备中。

在视觉测量系统工作之前,由现场校准装置及软件对视觉测量系统进行现场校准,使得视觉测量的工作状态和校准状态完全一致,保证了现场测量精度。

针对测量对象的不同,视觉测量系统的规模和复杂程度也不尽相同。如测量小型零部件或者大型物体的一部分,只需要单个视觉传感器或者为数不多的视觉传感器,相应的控制、电气等系统也比较简单。而像轿车白车身总成视觉测量系统,则由于被测物体庞大,需测量区域较多,且要求在线测量,因此整个系统结构相当复杂。基本组成包括多种类型的视觉传感器阵列、传感器安装及调整架、电气系统控制台、机械定位系统、传输系统以及视觉信息采集和处理系统,及其他辅助部件。这类型测量系统的最主要特点是系统具有多种类型的视觉传感器阵列,因此,我们称这种系统为多传感器视觉测量系统。由于不同传感器获取的坐标必须统一到被测物体本身的坐标系去,所得结果才具有意义,因此视觉测量系统还包括传感器标定以及整个系统全局标定所需要的各种硬件,如经纬仪、标定靶标等。

多传感器视觉测量系统是在 20 世纪 80 年代末才逐步发展起来的工业测量系统,由于它能完成工业生产中一些复杂、庞大工件的 100% 在线测量,现已经成为视觉测量发展的一支主流。国外一些著名公司、大企业都竞相研究与开发多传感器视觉测量应用技术,以满足生产中不同测量目的的需要。图 1.8 是美国 Perceptron 公司研制的用于轿车车身生产的 Perceptron 1000 型多传感器视觉测量系统。



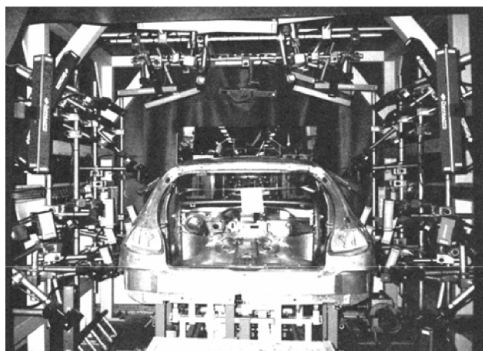


图 1.8 Perceptron 1000 型多传感器视觉测量系统

除了多传感器视觉测量系统,由单个视觉传感器的移动也可以实现多传感器视觉测量系统的功能,有时称这种方式为流动式三维扫描。在这种方式下,根据测量任务要求,单个视觉传感器可随意放置到(或称为流动到)大型工件周围的任何位置,通过在工件表面上设置适当的标记点,单个视觉传感器在不同位置的局部测量数据就可以自动拼接,从而完成对大型工件的整体测量。德国 GOM 公司的 ATOS 便携式三维扫描仪是这类系统的典型代表,如图 1.9 所示。

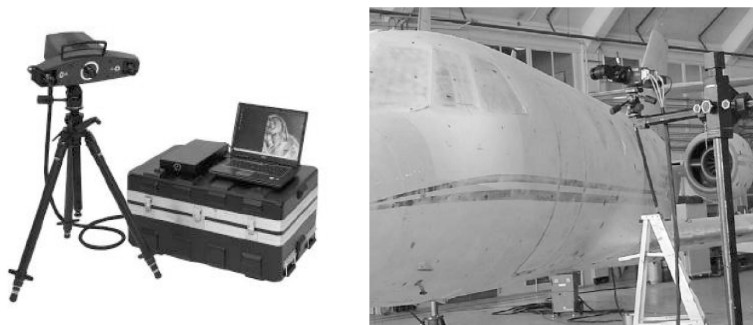


图 1.9 ATOS 便携式三维扫描仪

英国 3D Scanners 公司和美国 Faro 公司运用线结构激光扫描技术配合关节式测量手臂,可以对被测物进行多角度测量。图 1.10 和图 1.11 分别为英国 3D Scanners 公司和美国 Faro 公司的产品。

### 1.5.2 视觉测量关键技术

视觉测量系统主要性能指标包括测量精度、测量范围、测量速度、自动化智能程度、易维护性等方面。下面围绕这些指标来讨论视觉测量的共性关键技术。



图 1.10 英国 3D Scanners 公司产品



图 1.11 美国 Faro 公司产品

### 1. 视觉测量模型

视觉测量模型是实现视觉测量的基础,在此基础上通过对视觉图像中的各种特征信息进行处理、分析和计算,才可以实现被测物体三维几何尺寸、形貌及位置的测量。一般来讲,视觉图像是二维图像,如何从二维图像恢复出三维场景,这需要视觉测量模型给以描述,即视觉测量模型表征了从二维图像集合向三维场景集合的映射关系。

视觉测量模型对视觉测量系统的测量精度、测量范围及测量速度有着直接的影响。在视觉测量模型建立中,不但要求模型准确、简洁,而且应通用性好、适用性强。

### 2. 视觉传感器结构优化

一般来讲,三维视觉传感器主要包括双摄像机构成的双目视觉传感器以及单摄像机与光投射器构成的结构光视觉传感器,这两种传感器在三维视觉测量系统中起着重要作用。为了获得较高的测量精度,要求传感器的基线距,即两个摄像机之间或者单摄像机与光学投射器之间的距离尽可能大,这必然导致传感器的体积增大,重量增加。因此,解决传感器结构与传感器精度要求之间的矛盾成为传感器设计的主要内容。

### 3. 视觉视频信号实时采集

视觉传感器获取的通常是模拟视频信号,必须转换为数字图像信号才能作进一步的处理。通常的商用图像采集卡,可以实时地将图像信号采集到计算机内存,

能够满足大多数场合的应用要求,然而正是因为它的通用性,很难满足多传感器视觉测量系统的图像采集特殊要求,它不能向专用硬件前端处理器提供合理的有效接口。研究一种专用实时图像采集系统,使之能够向计算机和专用硬件处理系统同时提供处理信号,并保证图像采集和图像处理能够同时进行,是实现三维在线视觉测量的前提条件之一。

#### 4. 视觉视频信号前端实时处理

视觉图像信息数据量大,对于在线测量或多视觉传感器测量,能够完成视觉图像特征提取与处理的视觉视频信号前端实时处理技术尤为重要。因此,研制视觉视频信号前端实时处理系统,最大程度地减少计算机处理的数据量,使计算机只完成主要的高级处理任务,是提高整个系统测量速度的最有效途径。

低级图像处理的速度一直是图像界的热门话题之一,也是具有挑战性课题之一。由于图像数据量大,数据传输率高,同时低级图像处理算法涉及的数据领域性很强,这些因素促使人们去寻找新的提高图像处理速度的策略,构造新的处理体系结构。

#### 5. 摄像机内部参数标定

采用共面标定参照物的传统摄像机的内部参数标定,必须由摄像机获取多幅相互位置已知的二维标定参照物图像,才能求得摄像机内部参数。这种标定方法只能在传感器固定在测量架之前进行,要求高精度标定参照物以及精密移动导轨。首先应将标定参照物与导轨的垂直度调整到要求的范围内,然后每一个摄像机的标定都至少要求标定参照物精确移动到三个空间位置,因此标定劳动强度大、效率低。而且在安装标定好的传感器时,需要特别小心以保证传感器在固定前和固定后的结构不发生变化,因此不易维护。如何能够减轻标定劳动强度,并且保持摄像机的标定状态与使用状态完全一致,降低对标定设备的要求,成为当今摄像机标定技术的主要研究方向。

#### 6. 视觉传感器结构参数现场标定

标定视觉传感器的结构参数,即两摄像机之间的平移和旋转参数或单摄像机和光学投射器的位置关系,是视觉传感器能够进行三维测量的必要前提。同样,研究视觉传感器的现场标定方法,是保持传感器标定状态和使用状态完全一致的有效途径,同时也为降低劳动强度,降低对标定设备的要求提供了可能。

#### 7. 多传感器视觉测量系统现场全局标定

多传感器视觉测量系统是由两个以上(含)视觉传感器构成视觉三维测量系

统。多传感器视觉测量系统可以克服单视觉传感器测量中存在的盲区问题,同时又能够组成视觉测量站,实现大型物体尺寸和形位参数的测量,解决大范围空间尺寸三坐标的现场测量。

多传感器视觉测量系统现场全局标定是建立多传感器视觉测量系统的核心技术,其作用是把各个视觉传感器的测量数据统一到一个总体世界坐标系中,也就是确定各个视觉传感器坐标系相对总体世界坐标系的旋转矩阵和平移矢量。

## 8. 其他关键技术

除了以上关键技术外,视觉测量的关键技术还包括被测物体的硬定位和软定位、视觉测量系统小型化与可靠性设计、系统电气网络控制、机械结构设计、测量软件编制及测量数据分析与质量预测等。

另外,建立通用视觉测量系统及研究开发基于彩色图像与多谱图像的视觉测量也是视觉测量的关键技术与发展方向。总之,视觉测量作为一种新兴的测量技术,现代工业为其提供了巨大的需求空间。随着计算机视觉自身的成熟和发展,视觉测量必将在现代和未来工业中得到越来越广泛的应用。

# 1.6 本书各章内容简介

本书是作者研究组 15 年来从事视觉测量研究工作的总结,系统地介绍了视觉测量的基础原理、测量方法、关键技术与实用算法,并给出了视觉测量系统实例。本书涉及了视觉测量中的主要研究内容,各章节按照由视觉测量模型到视觉测量系统、由视觉测量算法到视觉测量应用、由单传感器视觉测量到多传感器视觉测量、由视觉测量关键技术到视觉测量系统集成的顺序来安排,体现了从简到繁、从浅到深、从理论到实际、从技术到系统的特点,力求具有层次性、系统性、先进性和实用性。

第 1 章介绍了计算机视觉的发展过程及系统构成,简介了生物视觉和 Marr 视觉理论框架,阐述了计算机视觉应用领域及面临问题,最后讨论了视觉测量系统与关键技术。生物视觉对研究计算机视觉非常有启发性,而 Marr 视觉计算理论对计算机视觉研究又产生了深远的作用,因此,建议有兴趣的读者可查阅相关文献作为参考。第 2 章介绍了射影变换、仿射变换、比例变换和欧氏变换等空间几何变换的一些常用结论,以及各种变换的不变量性质。另外基于中心射影变换,还介绍了摄像机的线性模型和非线性模型,并对摄像机透视投影的各种线性近似模型和应用条件进行讨论。视觉测量所涉及的数学知识较多,除本章介绍的空间几何外,随机过程、矩阵理论与最优化方法也是研究视觉测量的重要数学工具。第 3 章首先介绍了图像边缘与图像平滑,在此基础上讨论了改进的 Steger 图像边缘检测

算、各种角点探测器、椭圆形图像中心的提取及空间椭圆中心图像位置的提取。另外,许多图像处理的专著和文献中很多算法可直接应用于视觉图像特征信息的提取,由于篇幅有限,本章不再介绍,有兴趣的读者可查阅相关文献以作为补充。第4章介绍了视觉测量中典型算法硬件 IP 核设计,包括高斯滤波 IP 核设计、角点探测器 IP 核设计、光斑图像中心提取 IP 核设计及改进的 Steger 算法 IP 核设计。本章内容为视觉图像特征信息的快速准确提取奠定了基础。第5章介绍了摄像机标定的现有典型方法,如基于3D立体靶标的摄像机标定、基于径向约束的摄像机标定、基于2D平面靶标的摄像机标定及机器人手眼标定,并较为详细地讨论了基于交比不变的摄像机标定、基于卡尔曼滤波的摄像机标定及机器人足目标定。这些方法各有特点,适合于不同应用场合的摄像机内外参数的标定。第6章和第7章分别介绍了两种典型的三维视觉测量方法——双目立体视觉测量和结构光三维视觉测量,详细讨论了其视觉测量原理与数学模型、三维重构约束条件、结构设计及系统构成、标定方法及精度分析。第8章介绍了由多传感器构成的视觉测量系统,包括全局标定方法、全局标定系统以及全局标定精度分析与评价。本章内容为实现批量大型物体尺寸和形位参数的测量提供了技术途径。第9章介绍了流动式三维视觉测量原理、光栅条纹编码识别、测量数据拼接,及拼接实验与精度分析。本章内容为实现单件和小批量大型物体尺寸和形位参数的测量提供了技术途径。视觉测量在机械制造加工、工业自动化、运动监控及科学研究等领域有着广泛的应用,第10~12章分别介绍了微小构件内表面测量系统、轮胎综合几何参数测量系统及高扇翅频昆虫运动参数测量系统,对其关键技术进行了较为详细地阐述。

## 参 考 文 献

- [1] Roberts L G. Machine perception of three-dimensional solids. Tippet J T. In Optical and Electro-optical Information Processing. Cambridge; MIT Press, 1965, 159~197
- [2] Guzman A. Decomposition of a visual scene into three-dimensional bodies//Grasseli A. In Automation Interpretation and Classification of Images. New York; Academic Press, 1969
- [3] MacKworth A K. Interpreting pictures of polyhedral scenes. Artificial Intelligence, 1973, 4(2): 121~137
- [4] Marr D. Vision; A computational investigation into the human representation and processing of visual information. San Francisco; W. H. Freeman and Company, 1982
- [5] Mouaddib E, Batle J, Salvi J. Recent progress in structured light in order to solve the correspondence problem in stereo vision. Proceedings of the 1997 IEEE International Conference on Robotics and Automation, 1997, 130~136
- [6] Kovacevic Y R. Real-time sensing of sag geometry during GTA welding. Journal of Manufacturing Science and Engineering, 1997, 119(2): 151~160
- [7] Tech E K, Mital D P. A transputer-based automated visual inspection system for electronic

- devices and PCBs. Optics and Lasers in Engineering, 1995, 22: 161~180
- [8] Xu G, Zhang Z. Epipolar Geometry in Stereo. Motion and Object Recognition. The Netherlands: Kluwer Academic Publishers, 1996, 79~204
- [9] 寿天德. 视觉信息处理的脑机制. 上海: 上海科技教育出版社, 1997
- [10] 马颂德, 张正友. 计算机视觉. 北京: 科学出版社, 1998
- [11] Hubel H H, Wiesel T N. The Brain: A Scientific American Book. New York: W. H. Freeman, 1979
- [12] 吴健康, 肖锦玉. 计算机视觉基本理论和方法. 合肥: 中国科学技术大学出版社, 1993
- [13] 吴立德. 计算机视觉. 上海: 复旦大学出版社, 1993
- [14] Aloimonos Y. Special issue on purposive, qualitative, active vision. CVGIP-IU, 1992, 56(1): 1~29
- [15] Huang T, Stucki P. Special section on 3-D modeling in image analysis and synthesis. IEEE Trans. Pattern Analysis and Machine Intelligence, 1993, 15(6): 529~616
- [16] 王斌, 罗志勇, 刘栋玉等. 机器视觉心理与分辨锥体的应用. 华中理工大学学报, 1996, 24(6): 59~62