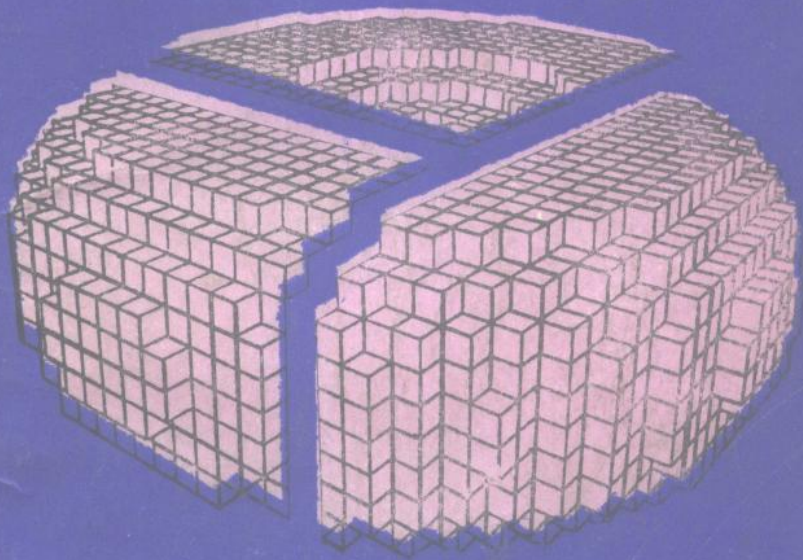


电信技术普及丛书



计算机视觉

罗宇华 编著

人民邮电出版社

电信技术普及丛书

计 算 机 视 觉

罗宇华 编著

人 民 邮 电 出 版 社

此成果得到中国国家自然科学基金资助

内 容 提 要

本书为“电信技术普及丛书”之一，系统地介绍了计算机视觉的基本概念、系统组成和应用，以及计算机视觉中常用的各种方法和具体应用实例。明确地阐述了目前计算机视觉系统所能解决的问题范围以及与人的视觉间尚存在的差距。

本书着重讲清物理概念，免去了过于抽象和繁琐的数学推导，因而通俗易懂，适合于该专业及其它科技人员和学生阅读参考。

电信技术普及丛书

计算机视觉

罗宇华 编著

*

人民邮电出版社出版

北京东长安街27号

中国铁道出版社印刷厂印刷

新华书店北京发行所发行

各地新华书店经售

*

开本：787×1092 $\frac{1}{32}$ 1990年8月 第一版

印张：5 $\frac{28}{32}$ 页数：94 1990年8月北京第1次印刷

字数：134千字 印数：1—2000册

ISBN7-115-04223-3/TP·051

定价：2.20元

出版者的话

为了普及电信技术知识,特别是电信新技术知识,为我国的通信现代化服务,我们组织编写了一套“电信技术普及丛书”,陆续出版。这套丛书的主要读者对象是具有中学以上文化水平、有一些电信基本知识的工人、管理干部和有关技术人员。在编写中,力求做到内容正确,概念清楚,深入浅出,通俗易懂;使读者读过一书后,能对某项技术的基本原理和主要情况有一个概括的了解,作为进一步学习的入门向导。我们殷切希望广大读者对这套丛书提出意见和建议,帮助我们做好这一工作。

前 言

视觉是人类最重要的感觉功能。它产生出对外部世界有效的描述。计算机视觉就是利用计算机来达到这一目的。计算机视觉是七十年代以来发展起来的一门多学科交叉的边沿学科。它在国民经济、科学技术及国防等各个方面都有广泛的应用。计算机视觉从理论上涉及人工智能、计算机科学、图象处理、模式识别、神经生物学、心理物理学等多个学术领域。

用简明、浅显的语言来阐述这门理论上较为艰深而实践范围又较广阔的新兴学科的基本内容,是本书努力达到的目的。全书力图用最少的篇幅着力介绍计算机视觉学科中重要的基本概念、原理和方法,并试图介绍该学科的一些新成果。

全书共分为六章。第一、二章为对整个学科概貌性的了解。三、四、五章则按计算机视觉中对视觉信息处理的最主要的层次顺序作介绍:如何对视觉信息进行早期处理、如何从图象中提取它所代表的场景特征以及怎样对场景进行描述。第六章则把这些内容统一起来介绍计算机视觉系统。

由于计算机视觉学科还在发展之中,本书的写法是一种尝试,欢迎广大读者对本书提出宝贵意见。

承清华大学葛成辉教授、林行刚副教授对全书作了深入细致地审阅,在此深表谢意。

本书的出版得到中国国家自然科学基金的资助,在此表示感谢。

广州暨南大学 罗宇华

1989年7月

目 录

第一章 概述	1
1.1 计算机视觉	1
1.2 与人的视觉比较	3
1.3 计算机视觉系统的基本组成	8
1.4 计算机视觉的应用	9
第二章 视觉信息的多层表示形式	12
2.1 什么是视觉的表示形式	12
2.2 视觉信息的多层表示形式	13
2.3 高级视觉过程与低级视觉过程	17
第三章 早期处理	20
3.1 视觉信号的获得	20
3.2 本征图象	28
3.3 图象滤波	33
3.4 局部边沿提取	42
第四章 场景特征的提取	47
4.1 边界	48
4.2 区域	62
4.3 纹理	78
4.4 运动图象与光流	86

9110038

第五章 对场景的更高级描述	97
5.1 表面	98
5.2 立体	115
5.3 物体	120
5.4 相互关系结构	130
第六章 计算机视觉系统	141
6.1 专用视觉系统举例	142
1. CONSIGHT-I 专用视觉系统	150
2. 检验印刷电路板专用视觉系统	157
6.2 视觉系统的控制方式	161
6.3 通用视觉系统的实验与展望	170

第一章 概 述

1.1 计算机视觉

视觉,是人类最重要的感觉功能。视,就是看;觉,就是感觉、感知。通过看来感知外部世界丰富多彩的信息,是我们每个人每日每时都在亲身体会的经验。“百闻不如一见”,这句话生动地说明了视觉对获得客观世界信息的重要性。据认为,从视觉获得的信息约占人所获得信息总量的 90%。

我们眼睛所看到的一切称为场景。我们对场景进行观察时,视网膜内就会形成一个象。人脑的视觉部分对这个象进行加工处理,就得到对客观世界场景的描述,最终达到对场景的理解和感知。可见视觉是一种复杂的信息处理任务,它对场景的象进行加工处理,最终产生出对观察者有意义的描述。

自从计算机出现以来,其发展日新月异,它已深入到国民经济的各个领域和人们的日常生活中,显示出对信息处理的巨大威力。既然视觉是一种复杂的信息处理任务,是否可以利用计算机来达到视觉的目的呢?回答是肯定的。计算机视觉就是使计算机具备“看”的功能,使计算机能认识和看懂所要看的东西,能找到它所见范围内目标的位置等等。利用计算机来达到这一目的的系统,称为计算机视觉系统。在计算机视觉系统中,如果说计算机好比人脑的视觉部分,那么成象器就好比人的眼睛,许多计算机视觉系统是用摄象机来作成象器的,成象器所形成的象

称为图象。

既然计算机视觉是用以完成视觉任务的,那么视觉任务又是什么呢?其实,视觉任务对我们每个视觉正常的人来说实在是太熟悉了。给出一幅照片,要说明照片上拍的是什么,就是一个视觉任务。又如要横过一条马路,也包含了视觉任务,因为你必须看清路上的情况,来往车辆的位置、速度,什么时候、从什么地方过去最为安全等等。再如给医生一幅病人胸部的 X 光照片,要求判断是否正常,如果不正常是什么毛病,在胸部的什么部位等等,这些都是视觉任务。

计算机在其他领域的成功鼓励人们进入这样一个新领域,也就是希望计算机能代替人完成上述所列的这一类视觉任务,而计算机视觉,正是为达到这一目的而形成的新学科。研究这门新学科的目的归纳起来有以下三点:

第一:建成一些专用的计算机视觉系统,完成在各种实际应用场景提出来的专门视觉任务。这些专用视觉系统在特定的场合下代替人的视觉功能,以便提高自动化的程度。这些专用系统对不同视觉任务和服务对象会有不同的结构,它们的结构与所完成的任务关系很大,例如工业生产线上的视觉机器人就是这一类系统。

第二:把对计算机视觉的研究作为探索人脑视觉部分工作机理的重要手段。人的视觉系统是最成功的视觉系统的范例,人可以非常轻松自如地完成非常复杂的视觉任务,但是,人对自己极强的视觉能力仅仅处于只能意会、不能言传的不自觉阶段。为什么人有如此完善的视觉功能?视觉究竟是一个怎样的过程?要揭开视觉的奥秘,除了从医学生物学、神经解剖学等角度去研究之外,计算机视觉是研究的重要途径。视觉本身很可能就是一个复杂的计算过程,计算机视觉着重从计算理论方面去研究视觉,

很可能正抓住了视觉过程最本质的东西。试想如果人们能用计算机重现与人视觉一样的效果,其主要处理过程不就是对生物视觉过程的重大启发吗?

第三:建成象人的视觉系统那样的通用计算机视觉系统。这样的系统可以适合于范围较广的视觉目标,甚至有可能超过人的视觉能力。当然,要达到这样的目标是长期艰巨的任务,然而这正是计算机视觉这门学科的最终目标。

至于以图象作为主要研究对象的学科还有一些,计算机视觉与它们有非常密切的关系和不同程度的交叉。但是各门学科都有自己的发展历史和特点,着重点亦有所不同,例如图象处理就是着重图象之间的变换,即从一种形式的图象变换到另一种形式的图象。输出的图象可能有更好的信噪比,或更适合于人的观察,或更加便于存储和传输,它一般不着重对图象的描述。模式识别则把模式,包括图象的模式进行分类,在对图象的模式分类过程中,需要首先从图象中提取一些预定的特征,并在特征空间中利用一系列判决准则进行分类,最终把图象模式归入预先定义的不同类别之中,例如文字识别就是模式识别的典型课题。

1.2 与人的视觉比较

人的视觉系统可以说是现有的视觉系统中最为成功的例子。根据光的强度不同、波长不同、方向位置不同,人眼可以感受到明暗、颜色、方向、形状和运动。人有区分物体和背景的能力,还有把所看到的模式聚合成有意义的物体的能力等等。

通过观察图 1.1 中的不同图案就可以显示出人对形状的感知能力。对图(a),人眼会不顾每一个小图案位置的分散而感觉到中间三角形的存在。对图(c),总会在中间加一条线,理解成为

一页折过的纸片。而对图(b)，总是觉得中间有一个圆，实际上只有几段直线段，并没有画出圆来。对图(d)，如果取不同部分作背景，就可以看出是一个花瓶或两个人脸部的侧面轮廓。人的视觉器官总是尽量把图形组合成有意义的个体。

人的视觉系统对不少景物的观察能力更是令人惊叹。想想我们识别人的面孔的能力吧！我们可以区分出多少不相同的面孔？

男女老少，熟悉的、不熟悉的，甚至对方脸部呈现的细微的喜怒哀乐的变化都逃不过人的眼睛。如果每一种形态都需要一定的存储单元来记忆，这该需要多少存储单元啊！我们的视觉系统对图象的存储能力之强真是令人难以想象。然而不知道人们有没

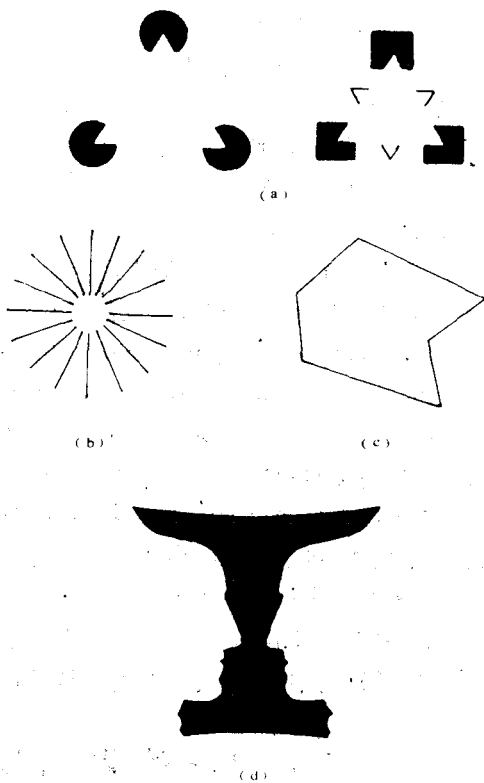


图 1.1 人眼对形状的感知

有注意到一个事实,就是我们口头上能对图象的定量描述却少得可怜,人类高超的视觉能力似乎完全是不自觉的,似乎与我们直接能控制的部分并不相连。视觉器官通过亿万年的生物进化已达到非常完美的地步,然而我们对它的认识却非常肤浅。

人的视觉系统是一部如此完善的自动机器,它是高效率的、实时的、可以并行处理的。计算机视觉正是要用一般计算机那有限的数字硬件产生出同样的效果来,其难度是可想而知的。总结起来说,困难主要表现在如下一些方面:

(1) 计算机视觉的唯一输入是图象,而从图象中所能得到的是一个二维的数字阵列。阵列中的每个元素 $f(x,y)$ 表示出该位置 (x,y) 处接收到的景物亮度,或称强度,要从强度值中求得形状本身就是一件极困难的事。图 1.2 左上角是一幅人脸右上部的图象,毫无问题,人们可以从图象中看到脸这一部分的形状。而下部的立体图形则是强度值本身的形象表示,图形中每一点的高度与强度的数值成正比,人们很难把二者直接联系起来。想从下图的强度值领会出左上角脸部的形状,真是难而又难的事,计算机视觉正是要用计算机来完成这样一个困难的任務。

(2) 由三维场景到二维图象是一个投影过程,在这一过程中失去了不少信息,如第三维的景深信息就基本丧失了。另外,引起强度变化的因素很多,诸如光照条件、物体表面反射率、物体的形状、摄像的角度、距离等。但在图象中每一个位置的强度值却只有一个,所有因素对该点的共同作用只由单一的一个强度值反映出来,要想从这一个已知值倒推影响它的多个因素,本身的条件就不够。

(3) 要恢复上述失去的信息以及产生出对图象景物有意义的描述,都需要有知识。视觉的知识一般由各种视觉模型表达出来,这些模型有关于客观世界物体的物理性质、形态的模型、成

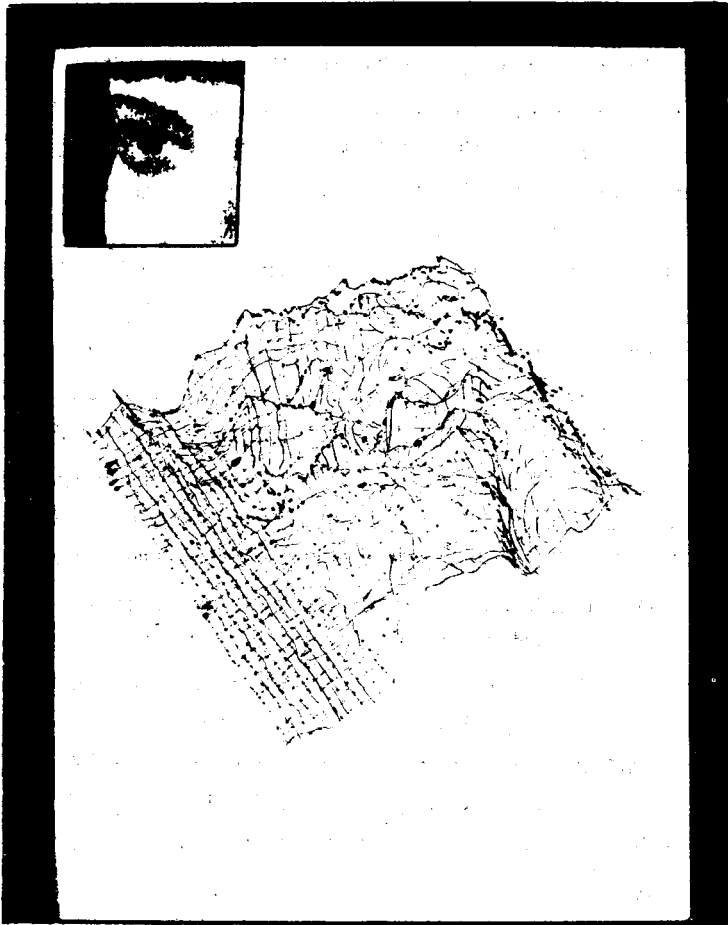


图 1.2 图象强度与形状

象器成象几何模型、物体的空间关系模型……,等等。在计算机视觉中如何表示模型是视觉系统的一个至关重要的问题,另外,一个成功的视觉系统往往不能缺少视觉目的的指导。我们在看东西时如果知道要找的是什,就比较容易找到,这点从图 1.3 中非常明显地表示出来,读者不妨自己试试,看看图中画的是什



图 1.3 视觉目的的指导作用

么?如果不告诉你要在图中找什么,大概都只会看到是一堆杂乱无章的点。如果告诉你要在图中找一只狗,就很容易找到,因为这是在有目的指导的情况下进行观察的,而且这里也体现了知识的作用,因为一般人对狗这种常见动物都会有一定的知识,利用这些知识有目的地去观察,效果就明显得多。可见计算机视觉涉及的不仅仅是图象中的问题,还有对视觉知识如何表达、如何利用的问题,这些都是计算机视觉这门新学科必须要解决的。

(4)实际技术上的问题。视觉信息比起语音信息数据量要大得多,计算机视觉系统必须解决视觉信息量大的问题,例如一幅典型的航空照片有 3000×3000 个数字,如果每个数字用一个 8 比特的字节表示,每幅图象就有多达 9M 个字节。如果对每个数进行 10 次简单的运算,就必须进行 9 千万次运算!因此,寻找高效率的算法和有效的表示形式是计算机视觉所要解决的基本技术问题。

1.3 计算机视觉系统的基本组成

图 1.4 表示出计算机视觉系统最基本的组成部分,系统由成象器、计算机、装有各种模型的视觉知识库以及输出设备组成,每一部分的详细作用可见图 1.5。

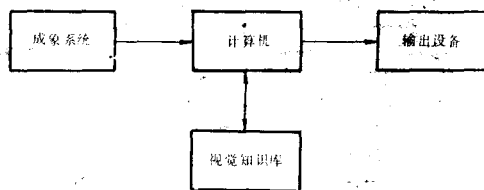


图 1.4 计算机视觉系统的基本组成部分

成象系统把三维场景的象记录下来,

形成强度的二维阵列——图象。图象有可能是黑白的,称为

灰度图象。同时也可能是彩色图象,成象器得到的是代表红、绿、蓝三个基色分量大小的三个数字阵列。另外,成象系统还可能得到运动物体在不同时刻的多幅图象,这称为图象序列。在图象序列中,每一幅依然是一个二维数字阵列。计算机则利用它内部知识库中的各种视觉模型对输入图象进行分析 and 解释,输出得到的是对图象三维场景的描述,这些描述随视觉目的的不同而不同,一般包括对物体的描述以及对物体相互关系的描述,如图 1.5 中所示。同

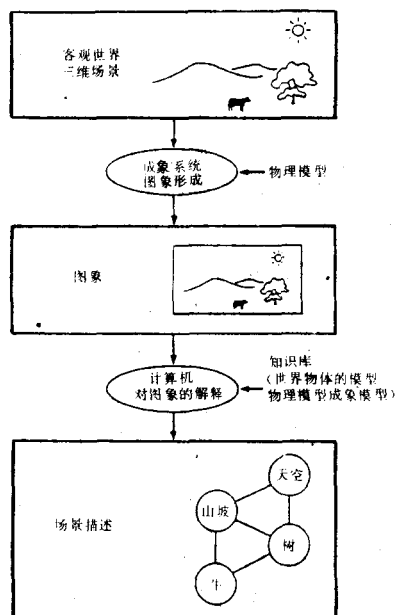


图 1.5 视觉系统各部分的作用

时也有可能对物体表面的结构、形状、纹理、颜色、质地、影子位置、光源位置等进行描述,输出设备可以是计算机的常规输出,如打印机、绘图仪等,更多的是专用的图象显示器以及机器人和其他自动机器的随动机构。

1.4 计算机视觉的应用

计算机视觉被称为自动化的眼睛,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,因此对观测和被观测者都不会产生任何损伤,十分安全可靠,这是其他感觉方式所不能比拟的。另外,视觉方式所能检测的对象十分广泛,可以说是对对象不加选择,理论上,凡是人眼能观察的、能判别、决策的对象范围都可以应用计算机视觉,甚至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,而计算机视觉可以利用这方面的敏感器件形成红外线、微波、超声波等图象,因此可以说它扩展了人类的视觉范围。人工观察者观察时间长了就会产生疲劳,这就会降低精度、引起误差增加。计算机视觉机器则不知疲劳,可以在长时间内保持始终如一的精度,因此,计算机视觉系统可以广泛用于长时间的、恶劣的工作环境。

由于有以上种种优点,计算机视觉有着广阔的发展前途,被认为是第五代新型智能计算机不可缺少的输入方式和组成部分,计算机视觉还在迅速发展之中。下面仅就其已取得的应用成果举例说明如下。

工业是计算机视觉广泛应用的重要方面。具有视觉的机器人可以直接参加生产,例如在生产线上进行部件安装、自动焊接、作切割加工。带有视觉的机器在大规模集成电路生产线上自动连接引线;对准芯片和封装;在石油、煤矿等地质钻探中对数据流自动监视和滤波;在纺织、印染业,进行自动配色、辅助设计,大大减轻人的视力劳动。

在各类检验、监视工作中,计算机视觉更能大显身手。这方面的应用有很多,如检查印刷底板的裂痕、短路及不合格的连接部;检查铸件的杂质和断口;对产品样品进行常规检查;检查标签的文字标记;检查包装与内装物是否相符;检查玻璃产品的裂缝和气泡等。