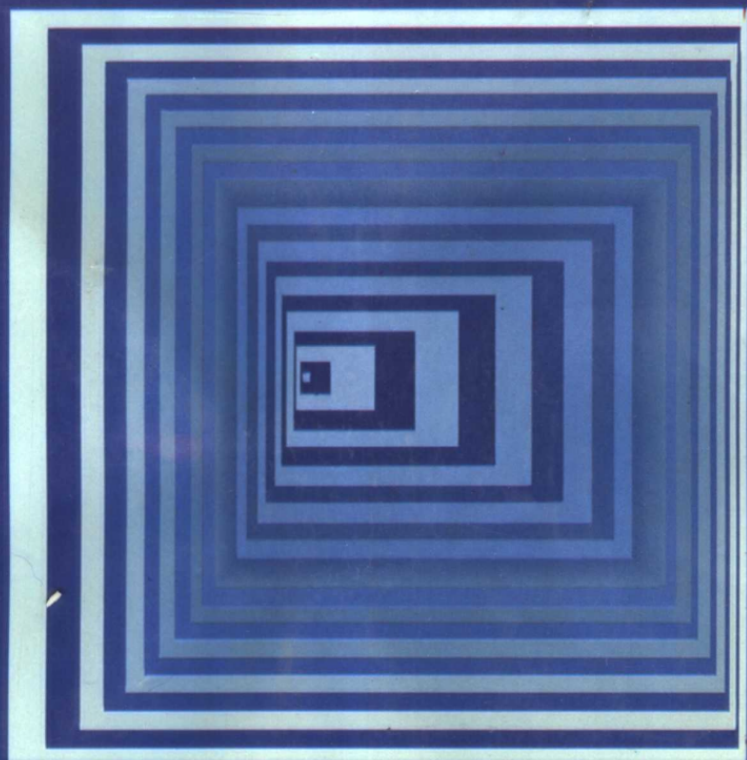


COMPUTER
VISION
计算机视觉

朱淼良 著



浙江大学出版社

计算机视觉

朱森良 著

浙江大学出版社

内容介绍

本书从人的视觉的机理出发,对机器模仿视觉功能的原理和方法作了探讨,指出了视觉世界的本质和计算机视觉信息处理的特点,并系统地叙述了视觉过程从低层、中层到高层的三层次模型。

全书分四部分共十二章,分别介绍了视觉的本质及计算机视觉的特点;特征抽取和区域分割;二维、三维信息的表达方式和空间建模方法;模式识别、模型匹配、知识推理、规划与导航。这些章节的安排,基本上对应于计算机视觉系统特别是知识基视觉系统的结构组成。通过本书,读者可以从中了解计算机视觉系统的基本构造和一些典型的算法描述。

本书可作为从事有关计算机科学和工程的专业人员的专业书,也可作为对计算机图像、图形及视觉有兴趣者的参考书。

计算机视觉

朱森良 著

责任编辑 洪保平

* * *

浙江大学出版社出版

(杭州玉古路 20 号 邮政编码 310027)

浙江大学出版社电脑排版中心排版

德清第二印刷厂印刷

浙江省新华书店经销

* * *

850×1168 32 开 11.5 印张 309 千字

1997 年 3 月第 1 版

1997 年 3 月第 1 次印刷

印数 0001—1000

定价:13.00 元

ISBN 7-308-01884-9/TP·159

前 言

让计算机“看见”世界，这无疑是一个令人兴奋的想法。由于计算机技术的飞速发展及在各个领域中无所不能地代替人类发挥神奇的作用，人们自然地相信计算机能像人那样具有视觉功能。确实，计算机视觉早已成为一大批科学家孜孜以求的目标，而且为此作了巨大的努力。

早先的有关计算机视觉的研究着眼于“看”的功能。人们认为只要把周围光学世界进入计算机之中，凭借计算机非凡的计算处理能力不难实现视觉目标。随着电视摄像及高速大规模集成电路的发展，光学图像能够方便自如地进出计算机。但是人们发现，所有有关图像操作的努力仅仅起到增强特征、改善图像质量的作用。人们意识到，视觉是一个从形象到抽象的识别理解的信息处理过程。于是具有高度抽象概括能力的数学工具被用来实现这个目标，出现了一大批有关视觉计算的理论、算法和系统。它们的观点是：所有真实世界的空间信息均可以采用“数学+算法”的方式从二维图像中得到恢复和重构。但是，人们很快感觉到，依靠严格的数学方法来规范五彩缤纷、千变万化的视觉世界实在是力不从心。许多精心设计的方法一旦对象稍有变化就难干奏效，因而失去应用价值。人们意识到：人的视觉系统能方便自如地应付各种复杂环境，是与人类的宝贵财富——知识处理能力分不开的。视觉处理是一个心理感知过程，其中知识起到关键的作用。利用认知知识模型来指导视觉——即知识基视觉系统是一个可取的研究途径。事实上，早在重构学派之前，基于心理模式的模式识别技术已经成功地解决了许多实际问题，取得令人瞩目的成就。它的特点是利用视觉特征，通过学习训练建立目标对象的共同模式。利用模式可进行目标的识别和分类，但是，把特定的模式推广到

更为一般的知识,从而使系统更加接近人的功能,这是模式识别技术所难于解决的。人工智能技术的发展给知识基视觉系统带来更大的希望和促进。在80年代中期,出现了视觉专家系统,出色地解决了一些复杂环境下的视觉任务,特别是景物的描述和理解。但是和其他学派一样,它也受到知识系统、特别是专家系统的脆弱性的影响。人工神经网络的勃起,使得大多数专家们转而开发这块新兴的“矿藏”。神经网络以它特有的结构仿真和神秘的工作方式,特别是它的高度平行处理能力,强烈地冲击着计算机领域。人们不仅认为它将代替传统的人工智能技术,而且期望它成为新的计算机结构方向。人工神经网络兴起至今已有十年了,在早期出现的一些令人兴奋的冲击波之后,由于机理的不明及种种实际原因,发展的步调很快地慢了下来。在视觉方面,人们发现神经网络仅擅长于模式识别的一些功能。相反地由于大规模集成电路产业的大步前进,高性能、低价位的计算机日新月异的发展,基于知识、算法的视觉系统得到稳定、持久地研究开发,并且进入许多实用系统。特别是多媒体信息时代的到来,作为其核心——视觉媒体的识别、理解和转换等高级处理必然成为未来多媒体技术的主流。因此,计算机视觉将会再次成为人们研究的热点。这就是撰写本书的原因。

本书力求包容上述计算机视觉发展过程中的各种有用的技术(除人工神经网络之外),但主要阐明下列几个观点:

一、视觉世界表面上是一个光学世界,实质上物体空间特性是其主要特征。视觉系统的目标是探求视觉世界的空间特征及其运动规律。

二、视觉过程是一个从客观实体到抽象感知的信息处理过程,其中先验知识起到主导作用。

三、计算机视觉的目的是用机器模仿人类视觉功能,是一个从具体的图像数据到抽象的符号表达的信息处理过程。可通过低层的特征抽取,中层的模式表达和建模,及高层的知识推理、匹配、判断三个层次来实现。

本书分为四部分共 12 章。第一部分包括第一章和第二章,分别就什么是视觉、什么是计算机视觉,及有关视觉的生理学、心理学、物理学的基本知识作了阐述和介绍,表明全书的基本观点。第二部分主要讲述以图像为基础的一些概念和操作,其中第三章讨论了图像的属性,第四、五章介绍了图像的线性特征的抽取及区域的分割。第三部分为中层处理,其中第六章主要介绍了有关二维图像特征的数学表达方式,第七章为三维空间物体模型表达方式,第八章则从知识角度讨论视觉世界的建模方式。第四部分讨论高层信息处理的几种操作,其中第九章对基于特征模型的模式识别进行简单介绍,第十章讨论了基于空间模型的匹配方法,第十一章讲述了基于知识模型的视觉专家系统(这是本书的重点),第十二章则对视觉系统重要的应用领域——空间规划与导航技术作了探讨。以上这些章节各自独立,分属专项技术,但从总体上构成了一个完整的计算机视觉系统。本书的观点是基于人工智能知识系统的。

本书为多年研究生教学的总结。材料一方面取自于各种学术著作,另一方面是本人多年从事国家自然科学基金、863 高技术计划和国防科研项目及国外访问研究的结果。考虑到读者的情况,本书力求避免繁复的数学推导和深奥的理论证明,一些理论和方法的介绍尽量深入浅出,偏重于概念性的描述。

计算机视觉是一个不断深化发展的学科,本书难免有错误和不完善之处,望专家和读者批评指正。

作者

1996 年秋

目 录

第一部分 导论

第一章 视觉与计算机视觉系统	3
1.1 视觉与视觉世界	3
1.1.1 视觉世界	4
1.1.2 视觉的本质	5
1.2 计算机视觉	9
1.2.1 计算机视觉的目标与任务	9
1.2.2 计算机视觉系统的基本结构	12
1.2.3 计算机视觉的关键技术	16
1.2.4 计算机视觉的相关学科	18
第二章 视觉的基本知识	23
2.1 视觉世界的物理特性	23
2.1.1 光度学	24
2.1.2 彩色视觉与色度学基础	26
2.2 生理视觉系统的基础知识	32
2.2.1 视觉系统的解剖结构	32
2.2.2 神经元及视觉神经结构	37
2.2.3 图像特性与视觉生理结构关系	41
2.3 视觉心理学知识	43
2.3.1 视觉的适应和对比特性	44
2.3.2 盖式塔组合(Gestalt grouping)规律	47
2.3.3 三维感受	49

第二部分 图像低层处理

第三章 图像与图像属性	56
3.1 成像原理	58
3.2 图像的属性	65
3.3 灰度图像的统计特性	68
3.4 纹理特性	74
3.5 图像的运动特性	83
第四章 边缘特征的抽取	93
4.1 边缘检测	94
4.1.1 梯度法边缘算子	94
4.1.2 边缘模板匹配	98
4.1.3 区域拟合法	101
4.1.4 边缘检测的几个问题	105
4.2 边界曲线的提取	106
4.2.1 直线段检测	106
4.2.2 曲线的跟踪和连接	108
4.3 Hough 变换法	112
4.3.1 直线 Hough 变换	113
4.3.2 Hough 变换的推广	116
第五章 图像的区域分割	119
5.1 基于象元分类的图像分割法	120
5.1.1 灰度直方图分割法	122
5.1.2 多特性直方图分割法	125
5.2 区域增长法	130
5.2.1 区域增长法的基本原则	130
5.2.2 区域增长算法	131
5.3 图像划分	135

5.3.1	图像划分的基本原理和方法	135
5.3.2	分割—归并算法	137
5.4	图像分割中的松弛算法	141
5.4.1	松弛算法的基本方法	141
5.4.2	图像分割中的松弛算法	144

第三部分 中层处理——空间表达与建模

第六章	二维特征的数学表达	150
6.1	曲线与边界	150
6.1.1	细化算法与边界跟踪算法	151
6.1.2	曲线的拟合法	156
6.1.3	曲线的链码	161
6.1.4	曲线的其他表达方法	167
6.2	区域的表示	170
6.2.1	行程码	171
6.2.2	中轴变换 MAT	173
6.2.3	四叉树	178
6.3	特性的量度	188
6.3.1	几何特性	189
6.3.2	拓扑特性	199
6.3.3	光学灰度特性	201
第七章	三维物体的建模	205
7.1	曲面表示法	206
7.1.1	曲面拟合法	206
7.1.2	翼形边——一种多面体的表示系统	209
7.1.3	通用锥表示法	211
7.2	体素表达法	215
7.2.1	体素堆积法	215

7.2.2	结构体素几何	217
7.3	几个有关三维立体信息的模型	224
7.3.1	本征图像	224
7.3.2	三维空间透视变换模型	227
7.3.3	运动物体的三维坐标	230
7.3.4	激光(Range)图像成像模型	233
第八章	视觉世界的知识模型	237
8.1	视觉世界的知识模型及知识库	238
8.1.1	空间知识表达系统	238
8.1.2	世界模型和知识库	241
8.2	视觉系统中的知识表达方式	245
8.2.1	框架系统	245
8.2.2	语义网络	248
8.2.3	面向对象的编程设计	253

第四部分 高层分析

第九章	基于特征模式的识别系统	261
9.1	图像识别系统的基本原理	261
9.2	判决理论的模式识别方法	263
9.2.1	一般方法	263
9.2.2	统计模式识别	265
9.3	句法模式识别	267
9.3.1	形式语言和句法识别系统	268
9.3.2	模式表达	269
9.3.3	识别和文法分析	273
第十章	基于模型匹配的视觉系统	278
10.1	模板匹配法	279
10.2	模型的装配技术	283

10.2.1	模型装配中的优化方法	284
10.2.2	Shape from x 问题	286
10.3	图结构模型的匹配	294
10.3.1	图论中的匹配问题	295
10.3.2	景物的解释	299
第十一章	知识基视觉专家系统	306
11.1	人工智能技术及视觉专家系统	306
11.1.1	专家系统的一般概念	306
11.1.2	知识库及知识模型	310
11.1.3	推理及控制	314
11.2	视觉专家系统的建立	317
11.2.1	视觉对象知识的实时获取	317
11.2.2	视觉专家系统的设计	320
11.3	实例 AARDS——航片道路网检测专家系统	324
第十二章	规划与导航	329
12.1	状态空间与地形信息系统	330
12.2	路径规划	334
12.2.1	全局规划	334
12.2.2	局部路径规划	339
12.3	导航定位	342
12.3.1	地标定位法	342
12.3.2	地形匹配定位	345
12.4	自主式视觉导航系统	349
附录	专用术语及其英汉对照	353

第一部分 导 论

什么是视觉,这显然是计算机视觉首先要搞清楚的问题,只有对视觉的本质、对象、目标有了正确的了解,才能说得上用计算机模拟视觉。同时,从视觉探讨过程中,还可以了解到视觉处理的特点,这无疑对计算机视觉有很大的启发意义。这是第一章前半部分探讨的主要内容,而后半部分则相应地讨论了计算机视觉的目标、任务、对象和方法,同时对当前的计算机系统结构和特点也作了简单的介绍。

任何一门技术学科均直接或间接地受到相关学科启发影响而诞生的,同时受到多个学科的渗透而发展起来的。计算机视觉既然是模仿人的视觉功能,那么有关人的视觉的学科——生理学、心理学是必不可少的。本书第二章开始将对神经生理学、认知心理学中有关视觉研究的结果加以介绍。它们分别从生理解剖、实验及心理测试多个方面揭示人的视觉系统的一些规律,对计算机视觉的研究发展起到重要的启示作用。另外第一章中明确地提出视觉世界是一个以空间特性为主的三维物理世界,光学特性则是它的表面特征。视觉的目标是从表面的现象去揭示空间特征的本质及运动规律,因此在第二章中对研究物体的光学色彩等物理学科——光度学、色度学作了简洁的介绍。它们是视觉计算中不可缺少的理论,是视觉低层处理中图像特征抽取的主要依据。

第一章

视觉与计算机视觉系统

1.1 视觉与视觉世界

视觉是人类和一些动物的基本功能。大家都知道,人们用眼睛来观察周围的客观世界,通过眼睛识别环境物体,辨明物体的方位,确定周围事物的组成及相互之间的关系。人们也可以通过视觉来识别文字、资料,理解其中的含义。人们把这种特有的功能称为视觉功能,这是人们认识世界、了解世界的重要途径之一。那么,什么是视觉呢?人类的上述功能是建立在人的眼睛器官对于周围光学世界的感知而得到的。人们常说失去了眼睛就失去了视觉,那么,是否可以说视觉是人和动物眼睛特有的一种感知外部光学世界的一种属性呢?要回答这个问题,首先看下面的一些实例:

第一类实例是有关不依靠眼睛的视觉系统。大家都知道,蝙蝠的眼睛是瞎的,它依靠对超声波的传感,同样能够感知周围物体的存在,识别跟踪猎物,避开障碍。响尾蛇利用对红外线的传感,能感知一定范围内物体的组成,特别是对运动动物的闯入,确定其方位,从而迅速出击。这两个例子说明,不通过眼睛,同样可以认识周围世界的空间结构及存在,达到眼睛器官所产生的同样的功能。

第二类例子是一些非光学视觉世界。一个明显的例子是x射线。人们通过x射线而产生的非光学图像,同样能够了解光线不能进入的内部物体的存在和结构。随着科学的发展,人们借助于各种波段的

电磁波能够探索原子结构、金属内部结构,或夜间室内外景物等。从这方面来看,客观世界的光学特性不是唯一的视觉系统研究的对象。

第三类例子是“视”与“觉”的关系。一个摄像机能够把周围世界的物体无一遗漏地摄入照片中,但是不能说摄像机有视觉功能。因为它并不达到对周围世界的理解。同样,一个出生不久的小孩,它有正常的眼睛系统能够对周围的光线产生反应,但是,因为它对周围世界一无所知,因此无法了解眼睛所看到的景物,从而不能了解周围的世界。

上述的三类例子,围绕着什么是视觉进一步提出了几个重要的问题:什么是视觉的真正对象,视觉的目的是什么,视觉的本质是什么和视觉有什么特点,下面分别加以说明。

1.1.1 视觉世界

既然视觉是人类认识客观世界的一种重要感知手段,而客观世界具有各种方面特性和征状,那么不同的感知系统,其结果会得到不同的世界概念。例如听觉得到的是各种发声源组成的声音世界。按日常的体会,大家都认为视觉是通过客观世界的光学特性来认识世界的。光学特性仅仅是感知物体存在的一种媒介,而视觉真正的目的是为了识别物体,了解周围环境的物体存在及相互关系,从而进一步理解周围环境的现状和意义。因此,我们把视觉系统所要处理的对象称为视觉世界。视觉世界主要是由占有空间的物体组成,人们通过视觉世界具有的光学或非光学特性来了解视觉世界的存在。所以这里说的视觉世界是一个以客观物体的空间特性为主要特征的客观世界,而它的光学和某些非光学特性是它的一些重要属性。视觉系统的研究目的是感知视觉世界的空间存在,了解周围视觉世界的空间结构、特点、组成以及它们的空间运动变化规律。

根据对视觉世界的讨论,可以说,蝙蝠是通过超声波的感知来了解它周围空间物体的存在、分布及猎物的空间运动变化,这就是蝙蝠的“视觉”。事实上,仿生科学中研究蝙蝠的功能,已经成为仿生视觉

技术的一大成就。响尾蛇通过红外线感知器官来了解具有红外线辐射物体的空间位置和运动信息,它也是具有一定的“视觉”功能的。从广义的角度来说,不管采用什么传感系统,只要目的是探索了解客观物体的空间存在结构和空间运动规律的,均可视为视觉系统。

1.1.2 视觉的本质

为了说明视觉的本质是什么,不妨根据我们日常视觉感知的经验,来讨论一下典型的景物分析的过程。

在我们眼前是一个典型的乡村景物,由竹林、房屋、农田、池塘、小路以及天空组成,如图 1.1-1(a)。该景物映射到我们的眼睛视网膜上,成为一平面型的图像。它与照相机相片类似,所有景物的色彩,纹理基本上保持原来的特点。当然由于透视关系,远的物体变小,如图 1.1-1(b)所示。根据人们习惯的方法,根据表面的色彩,纹理不同,把图像划分成几个不同的区域。在同一区域中它们的色彩、纹理基本相同或相类似,区域的边界勾划出了该类物体的空间轮廓,如图 1.1-1(c)所示。

从人们日常的经验出发,根据每个区域的轮廓形状和色彩等特征,很快地识别出各个区域的名称。此时,该景物在心目中不再是复杂的图像,而是抽象成由几种能说出其类型的物体概念组成,如图 1.1-1(d)所示。

人们还可以用语言来描述所见到的景物:“蓝天下面是一片树林,树林中有一座房子,房子前有池塘、道路、田地……”这时,在人们的心目中便抽象成为简单的拓扑关系的概念结构。如图 1.1-1(e)所示。

人们最后可以把整个景物归纳为“乡村景物”这一结构,如图 1.1-1(f)所示。

上述的过程说明了视觉感知过程是一个由客观空间的物理存在到主观概念确定的不断概括抽象的过程,是一个从具体的多彩、繁复的空间景物到简练的语言式概念描述的过程。这就是视觉系统的实

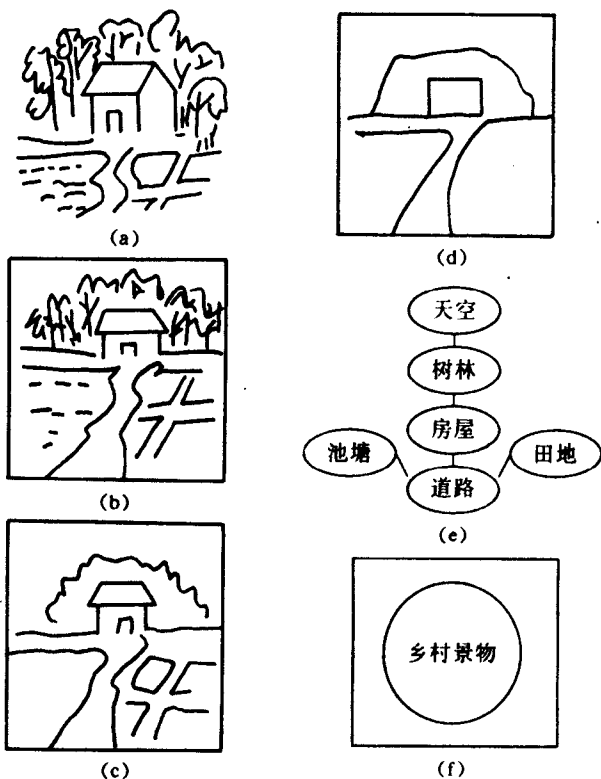


图1.1-1 人的视觉系统工作过程

质,是人类独特的认识世界的本领。它的形成是基于两个实用的原因。一是为了记忆。世界景物千变万化,人的头脑记忆能力毕竟有限,因此,必须把大量相同类型的景物抽象成少量的具有一定特征的“模式”或“框架”。记住这些“模式”便可以掌握一大类景物的形象,在具体使用时可通过经验知识加上空间想像,基本上可以复现原来的形象。二是为了思想的交流。当一个人需要把它们看到的景物告诉另一个人时,它可以把上述抽象的结果用简练的语言进行交流。例如,当一个人告诉另一个人说他看到了一个乡村景物时,另一个人凭他