

计算机视觉

D.H.巴拉德 C.M.布朗 著



科学出版社

1.4

2/1

计 算 机 视 觉

D. H. 巴拉德 C. M. 布朗 著

王东泉 徐心平 赵经伦 译

潘裕焕 校

科 学 出 版 社

1 9 8 7

内 容 简 介

本书在总结有关计算机视觉领域的最新研究成果的基础上,力图对这一新学科建立一个完整的概念和系统化的理论基础,重点讨论图象理解过程中输入图象信息和有关客观世界的先验知识间交互作用的各种表示法和原理,它综合了人工智能、心理学、计算机图示学和图象处理等有关学科的部分知识,构成其自己的独立体系。本书按描述景物抽象程度的不同而分成四部分:(1)广义图象;(2)分割图象;(3)几何结构;(4)关系结构。全书共分十三章和两个附录,除第一章外,每章均有练习题。

本书内容新颖、材料丰富,并附有大量的图片、算法和参考文献,本书可作为高等学校人工智能、模式识别、图象处理以及其它有关专业的教科书或教学参考书,可供以上专业的学生、研究生、教师及从事计算机视觉研究的广大工程技术人员和科学研究工作者参考。

D. H. Ballard C. M. Brown
COMPUTER VISION
Prentice-Hall, Inc., 1982

计 算 机 视 觉

D. H. 巴拉德 C. M. 布朗 著

王东泉 徐心平 赵经伦 译

潘裕焕 校

责任编辑 刘兴民 袁放尧

科学出版社出版

北京朝阳门内大街137号

轻工业出版社印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

1987年9月第 一 版 开本: 787×1092 1/16

1987年9月第一次印刷 印张: 28¹/₂ 插页: 1

印数: 0001--3,000 字数: 658,000

统一书号: 15031·866

本社书号: 5165·15-8

定价: 6.70元

译者前言

计算机视觉就是用机器代替人的眼睛和大脑,对客观世界进行视觉感知和解释。众所周知,视觉对于人类来说是最有用的,但也是最复杂的一种感知周围环境的手段。随着科学技术的日益发展,计算机视觉的应用范围正在不断扩大,例如具有视觉反馈的机器人可以代替人完成各种复杂的任务,包括工业产品的自动装配、焊接和检验;生物医学中X光照片的自动诊断;遥感照片的自动解释;各种自治车辆(如火星车)的导航;以及各种军事应用。计算机视觉是实现机器智能及第五代计算机的关键因素之一。

图象理解或景物分析是非常复杂的过程。人类感知和解释一幅图象似乎毫不费力,即使对图象中的某些物体并不熟悉,但我们可以从中抽取最有用的特征,并利用我们头脑中已有的有关客观世界的先验知识或模型来解释。因此机器实现图象理解也必须将输入图象与预先存入的有关物体结构和环境约束的知识进行交互作用。计算机视觉的研究近十年取得了非常迅速的发展,已成为人工智能的一个独立分支。作为计算机视觉的重要组成部分,有关图象处理、模式识别的研究已较为成熟,国内外有关的专著已不下数十种。而本书则是系统地从人工智能的角度讨论图象理解的原理和方法的第一本完整的著作。

本书第一章简单介绍计算机视觉的基本概念,第二章和第三章讨论图象的生成和前处理,包括图象的基本性质,滤波处理,以及产生边缘、距离、表面方向和光流等本征图象的方法。第四章和第五章讨论图象分割的两个基本方法:边界检测和区域生长。纹理和运动是两个更为重要的、包含大量信息的图象特征,因此是目前受到特别重视的两个研究领域,分别在第六章和第七章中讨论。第八章和第九章分别讨论描述物体形状的二维和三维结构,这对于物体识别和景物分析具有重要的意义。作为计算机视觉的高级能力处理则要利用人工智能中的知识表达、语义网、符号匹配、推理和规划等基本原理,如何将这些高级智能控制方法用于图象理解则分别在第十、十一、十二和十三章中讨论。

本书内容新颖、材料丰富,并附有大量的图片、算法和参考文献,反映了计算机视觉领域的最新研究成果和发展趋势,对于我国正在蓬勃兴起的计算机视觉、人工智能以及其它有关学科的研究必将有较大的参考价值。

本书第二、三、十、十一、十二、十三各章及附录由王东泉翻译,第一、四、五、六、七各章及索引由徐心平翻译,第八、九两章由赵经伦翻译。潘裕焕同志在百忙中为全书进行了精心审核,在此表示衷心感谢。

本书在译出过程中对原作者已指出的错误进行了改正,对我们认为错误的地方也做了订正。由于本书内容包含的学科面很广,许多新的名词、术语尚无统一译法,限于译者水平,误解或翻译不当之处在所难免,衷心希望读者予以批评指正。

译者

1985年12月9日

序 言

人类对智能自动机的幻想虽然可以追溯到遥远的过去,但直到1950年左右才在Turing的数字计算机论著中第一次较广泛地阐述清楚。其后,继续促使其实现的,主要是从事人工智能研究的学者,他们的目标是赋予计算机类似生物体所具有的信息处理功能。因此,从一开始,人工智能研究的目标之一就是给机器配备能处理传感器输入信息的功能。

计算机视觉是由图象对客观物体建立明确而有意义的描述。图象理解与图象处理截然不同。后者只研究从图象到图象的变换方法,并不明确地描述图象的构成。描述是物体识别、处理和评价的前提。

我们用许多不变的特性来接受和认识相关的三维物体世界。客观地说,输入的视觉数据并不会显示出相应的相关性或不变性;却含有大量无关的或甚至使人误解的偏差。倒是我们的视觉系统,从视网膜到认识的各个阶段,能以某种方式设法理解这些杂乱无章的视觉输入数据,或整理清楚这些数据。这可以通过从输入数据中可靠地抽取得到的固有信息来实现,也可通过视觉处理各个阶段用到的假设和知识来完成。

计算机视觉的研究课题是很明显的。确切地说,这些问题是,只用物理学和光学中非常基本的假设,从一幅图象中能提取景物的什么信息?必须完成哪些计算?在哪个阶段必须向理解过程引入有关客观世界域相关的先验知识?如何表示和利用客观世界模型及知识?本书主要阐述有关图象理解中促使图象信息与先验知识相互作用的那些表示方法和机理。

计算机视觉是一个相当新、而且发展很快的领域。最初的试验是在五十年代后期开始进行的,而许多基本概念,则是最近五年发展起来的。随着计算机视觉的迅速发展,致使在诸如人工智能、心理学、计算机图示学以及图象处理这些根本不同的领域里,出现了许多具有决定意义的新思想。我们选编这些材料的意图在于供高等院校的高年级学生和研究生作为教科书,或给建造视觉系统的专业人员作为得力的参考书。本书人工智能学科味很浓,我们希望这将促进研究人员思考问题。我们相信,在颇为成功的视觉系统中,固有的图象信息和客观世界的内在模型两者均具有重要意义。

本书基于对物体的四个不同抽象程度的描述而分成四大部分:

1. 广义图象——图象和与图象类似的实体。
2. 分割图象——组成子图象的各种图象,子图象可能对应于“感兴趣的物体”。
3. 几何结构——图象和客观世界结构的定量模型。
4. 关系结构——图象和客观世界结构的复杂的符号描述。

上述各部分按抽象程度递增的顺序排列。虽然按理说应该依次阅读上述四个部分,但它们之间的联系不很紧密。第一部分是第二部分的基础,而第三、四两部分则可以独立地阅读。

本书各部分中假定读者已有若干数学和计算技术的基础知识(微积分、线性代数、数据结构、数值方法);但就全书而言,数学的严密性同概念相比处于次要地位。我们

的意图是为尽可能广泛的读者提供关于这个新兴领域的一套完整的概念。

用一本书不可能非常完善地评述过去在计算机视觉方面所做工作的广度和深度。况且我们认识到，在这一迅速发展的领域里还将会不断地涌现出新思想和新概念。我们只希望本书能引起读者深入地阅读、思考和提出有益的批评，更希望迅速超越和摆脱本书的局限。

D.H.巴拉德 C.M.布朗

目 录

译者前言

序言

第一章 计算机视觉	(1)
1.1 简单视觉目标的获取	(1)
1.2 高级和低级能力	(3)
1.3 表示法系列	(5)
1.4 计算机的作用	(8)
1.5 计算机视觉的研究及应用	(9)

第一部分 广义图象

第二章 图象生成	(12)
2.1 图象	(12)
2.2 图象模型	(12)
2.2.1 图象函数	(12)
2.2.2 成象几何	(13)
2.2.3 反射	(16)
2.2.4 空间特性	(18)
2.2.5 彩色	(24)
2.2.6 数字图象	(27)
2.3 计算机视觉的成象设备	(34)
2.3.1 照相成象	(34)
2.3.2 距离检测	(42)
2.3.3 重构成象	(46)
第三章 前处理	(51)
3.1 恢复本征结构	(51)
3.2 图象的滤波	(52)
3.2.1 模板匹配	(53)
3.2.2 直方图变换	(56)
3.2.3 背景消去法	(57)
3.2.4 滤波模型和反射模型	(59)
3.3 局部边缘的检测	(60)
3.3.1 边缘算子的类型	(62)
3.3.2 边缘阈值化策略	(64)
3.3.3 三维边缘算子	(66)
3.3.4 边缘算子的性能如何	(67)

3.3.5	边缘松弛法	(69)
3.4	由几何确定距离信息	(72)
3.4.1	立体视觉和三角形法	(72)
3.4.2	立体观察松弛算法	(73)
3.5	由反射模型确定表面方向	(76)
3.5.1	反射率函数	(76)
3.5.2	表面梯度	(77)
3.5.3	光度立体观察	(80)
3.5.4	用松弛法由明暗确定形状	(81)
3.6	光流	(83)
3.6.1	光流的基本约束条件	(84)
3.6.2	用松弛法计算光流	(84)
3.7	分辨率锥体	(86)
3.7.1	灰度级合并	(87)
3.7.2	相关运算中的锥体结构	(87)
3.7.3	边缘检测中的锥体结构	(89)

第二部分 分割图象

第四章	边界检测	(94)
4.1	边缘元素的联合	(94)
4.2	近似位置附近搜索	(95)
4.2.1	先验边界的修正	(95)
4.2.2	非线性相关用于边缘空间	(96)
4.2.3	分割-占据法边界检测	(96)
4.3	Hough曲线检测法	(96)
4.3.1	梯度的使用	(98)
4.3.2	实例	(99)
4.3.3	参数空间计算和图象空间计算的替换	(100)
4.3.4	Hough变换的推广	(101)
4.4	图搜索法边缘追踪	(103)
4.4.1	品质估价函数	(105)
4.4.2	查找整个边界	(105)
4.4.3	A算法的替代	(107)
4.5	动态规划法边缘追踪	(108)
4.5.1	动态规划	(108)
4.5.2	动态规划用于图象	(110)
4.5.3	低分辨率估价函数	(111)
4.5.4	动态规划的理论问题	(113)

4.6 轮廓追踪	(114)
4.6.1 推广到灰度图象	(114)
4.6.2 推广到高维图象数据	(115)
第五章 区域生长	(116)
5.1 区域	(116)
5.2 局部法: 团点着色法	(117)
5.3 整体法: 阈值法区域生长	(118)
5.3.1 多维空间的阈值处理	(119)
5.3.2 多级精化	(122)
5.4 分裂与合并	(122)
5.4.1 状态空间法用于区域生长	(123)
5.4.2 低级边界数据结构	(123)
5.4.3 根据图形建立的区域结构	(124)
5.5 语义的结合	(126)
第六章 纹理	(130)
6.1 什么是纹理	(130)
6.2 纹理基元	(132)
6.3 纹理元排列的结构模型	(133)
6.3.1 文法模型	(135)
6.3.2 形文法	(135)
6.3.3 树文法	(137)
6.3.4 阵文法	(140)
6.4 纹理模式识别问题	(142)
6.4.1 纹理能量	(145)
6.4.2 空间灰度级的相依性	(148)
6.4.3 区域纹理元	(149)
6.5 纹理梯度	(151)
第七章 运动	(155)
7.1 运动理解	(155)
7.1.1 与域无关的理解	(155)
7.1.2 与域相关的理解	(156)
7.2 光流理解	(158)
7.2.1 延伸焦点	(158)
7.2.2 毗邻、深度和碰撞	(160)
7.2.3 表面方向和边缘检测	(161)
7.2.4 观察者本身运动	(164)
7.3 图象序列理解	(165)
7.3.1 由离散图象计算光流	(165)

7.3.2	运动刚体	(168)
7.3.3	运动光显示的解釋——一种与域无关的方法	(171)
7.3.4	人体运动理解——一种模型指导方法	(174)
7.3.5	分割图象	(175)

第三部分 几何结构

第八章	二维几何结构表示	(182)
8.1	二维几何结构	(182)
8.2	边界表示	(183)
8.2.1	折线	(183)
8.2.2	链码	(185)
8.2.3	ψ - s 曲线	(187)
8.2.4	Fourier 描述	(188)
8.2.5	二次曲线	(188)
8.2.6	B 样条	(189)
8.2.7	带树	(193)
8.3	区域表示	(197)
8.3.1	空间占有阵列	(197)
8.3.2	y 轴	(197)
8.3.3	四叉树	(198)
8.3.4	中轴变换	(200)
8.3.5	复杂区域的分解	(202)
8.4	简单形状性质	(203)
8.4.1	面积	(203)
8.4.2	偏心率	(203)
8.4.3	Euler数	(203)
8.4.4	紧凑性	(204)
8.4.5	斜率密度函数	(204)
8.4.6	标记	(205)
8.4.7	凹面形树	(206)
8.4.8	形状数	(207)
第九章	三维结构的表示	(210)
9.1	立体及其表示	(210)
9.2	表面表示	(210)
9.2.1	由面构成的表面	(211)
9.2.2	以样条为基础的表面	(214)
9.2.3	作为球面函数的表面	(215)
9.3	广义圆柱体表示	(219)

9.3.1	广义圆柱体坐标系统及其特性	(220)
9.3.2	抽取广义圆柱体	(222)
9.3.3	骨架的一种离散体积形式	(223)
9.4	体积表示法	(224)
9.4.1	空间占有阵列	(224)
9.4.2	单元分解	(225)
9.4.3	结构立体几何	(225)
9.4.4	立体表示的算法	(227)
9.5	线条画理解	(233)
9.5.1	线条画与三维基元的匹配	(235)
9.5.2	将区域分组聚合成为物体	(237)
9.5.3	标记线	(238)
9.5.4	关于平面的推理	(243)

第四部分 关系结构

第十章	知识表达和使用	(253)
10.1	表达	(253)
10.1.1	知识库——模型和过程	(253)
10.1.2	类比和命题的表达	(255)
10.1.3	过程知识	(256)
10.1.4	计算机实现	(257)
10.2	语义网	(258)
10.2.1	语义网基础	(258)
10.2.2	推理语义网	(261)
10.3	语义网例子	(267)
10.3.1	框架的实现	(267)
10.3.2	定位网络	(268)
10.4	复杂视觉系统的控制问题	(272)
10.4.1	并行计算和串行计算	(272)
10.4.2	分层和非分层控制	(273)
10.4.3	信念维护和目标实现	(277)
第十一章	匹配	(279)
11.1	匹配的概况	(279)
11.1.1	解释、构造、匹配和标记	(279)
11.1.2	图象、几何和关系诸结构的匹配	(280)
11.2	图论算法	(283)
11.2.1	算法	(283)
11.2.2	复杂性	(284)

11.3 图论算法的实现.....	(285)
11.3.1 匹配量度.....	(285)
11.3.2 回溯搜索.....	(288)
11.3.3 结合图技术.....	(290)
11.4 实践中的匹配.....	(294)
11.4.1 决策树.....	(294)
11.4.2 决策树和子图同构.....	(298)
11.4.3 信息特征的分类.....	(299)
11.4.4 一个复杂的匹配程序.....	(302)
第十二章 推理.....	(305)
12.1 一阶谓词演算.....	(305)
12.1.1 子句型句法 (非形式的)	(306)
12.1.2 非子句型句法和逻辑语义 (非形式的)	(306)
12.1.3 非子句形式转变成子句形式.....	(308)
12.1.4 定理证明.....	(309)
12.1.5 谓词演算和语义网络.....	(311)
12.1.6 谓词演算和知识表达.....	(313)
12.2 计算机推理.....	(315)
12.3 产生式系统.....	(316)
12.3.1 产生式系统的详细说明.....	(317)
12.3.2 模式匹配.....	(318)
12.3.3 一个例子.....	(319)
12.3.4 产生式系统的优点和缺点.....	(324)
12.4 景物标记和约束松弛法.....	(326)
12.4.1 一致和最优标记法.....	(326)
12.4.2 离散标记算法.....	(329)
12.4.3 线性松弛算子和线条标记的例子.....	(332)
12.4.4 一种非线性算子.....	(335)
12.4.5 用作线性规划的松弛法.....	(338)
12.5 活动知识.....	(345)
12.5.1 假设.....	(346)
12.5.2 HOW-TO和SO-WHAT进程.....	(346)
12.5.3 原始控制语句.....	(346)
12.5.4 活动知识的特点.....	(347)
第十三章 目标实现.....	(350)
13.1 符号规划.....	(350)
13.1.1 现实世界的表达.....	(351)
13.1.2 动作的表达.....	(352)

13.1.3	堆叠积木块	(353)
13.1.4	框架问题	(355)
13.2	有代价的规划	(356)
13.2.1	规划、评分及其相互作用	(356)
13.2.2	简单规划的评分	(357)
13.2.3	增强型规划的评分	(361)
13.2.4	实用上的简化	(362)
13.2.5	基于规划的一个视觉系统	(362)
附录一	一些数学工具	(371)
A1.1	坐标系	(371)
A1.1.1	笛卡尔坐标系	(371)
A1.1.2	极坐标系和空间极坐标系	(371)
A1.1.3	球面坐标系和柱面坐标系	(372)
A1.1.4	齐次坐标	(373)
A1.2	三角学	(373)
A1.2.1	平面三角学	(373)
A1.2.2	球面三角学	(374)
A1.3	向量	(374)
A1.4	矩阵	(376)
A1.5	直线	(378)
A1.5.1	两点	(378)
A1.5.2	一点和方向	(378)
A1.5.3	斜率和截距	(378)
A1.5.4	比率	(378)
A1.5.5	法向和到原点的距离(直线方程)	(378)
A1.5.6	参数式	(379)
A1.6	平面	(379)
A1.7	几何变换	(380)
A1.7.1	旋转	(380)
A1.7.2	尺度变换	(381)
A1.7.3	扭斜	(381)
A1.7.4	平移	(382)
A1.7.5	透视	(382)
A1.7.6	直线和平面的变换	(383)
A1.7.7	小结	(383)
A1.8	摄像机校准和逆透视	(384)
A1.8.1	摄像机校准	(384)
A1.8.2	逆透视	(385)

A1.9	最小二乘方误差拟合	(387)
A1.9.1	伪逆法	(387)
A1.9.2	主轴法	(388)
A1.9.3	利用伪逆法的曲线拟合	(389)
A1.10	二次曲线	(390)
A1.11	插值	(391)
A1.11.1	一维	(391)
A1.11.2	二维	(391)
A1.12	快速 Fourier 变换	(392)
A1.13	二十面体	(394)
A1.14	求根	(395)
附录二	高级控制结构	(398)
A2.1	标准控制结构	(398)
A2.1.1	递归	(398)
A2.1.2	协同子程序	(399)
A2.2	固有序列结构	(399)
A2.2.1	自动回溯	(399)
A2.2.2	上下文转换	(400)
A2.3	序列结构或并行结构	(400)
A2.3.1	模块和信息	(401)
A2.3.2	优先级任务队列	(402)
A2.3.3	有向模式调用	(404)
A2.3.4	“黑板”系统	(405)
参考文献		(407)
参考文献中引用的会议录及专集名称的缩写符		(426)
人名索引		(427)
名词索引		(430)

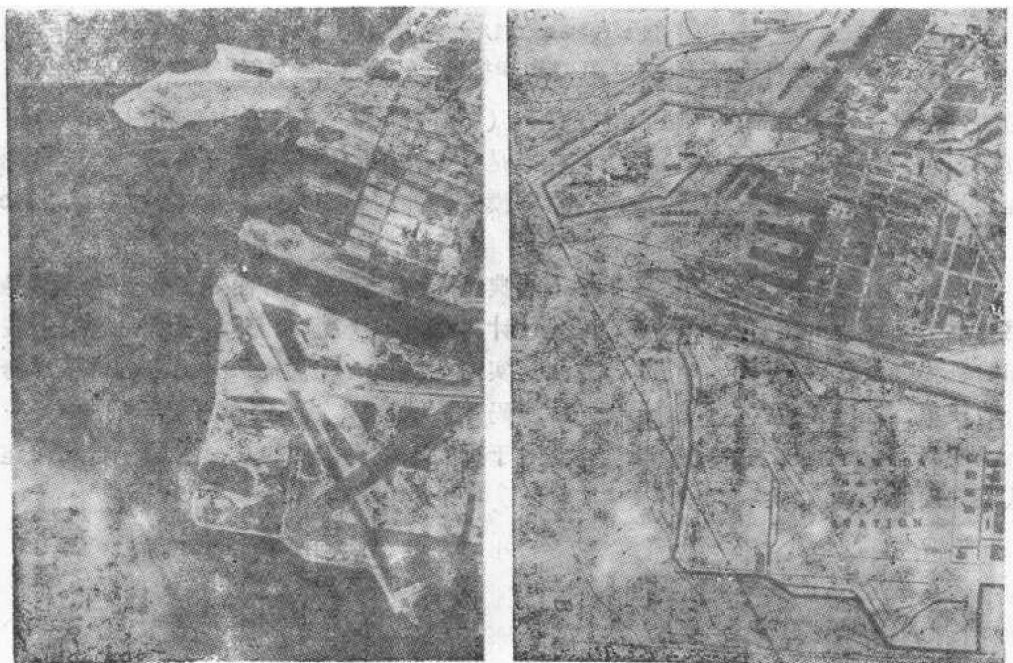
第一章 计算机视觉

计算机视觉问题

1.1 简单视觉目标的获取

假定给你一张如图1.1(a)所示的航空照片,请你确定船在照片中的位置。以前你可能从未在航空照片中见过海军舰艇,但你会毫无困难地对船的样子作出一般性的预测。你可能推断说,在陆地上绝不会找到船,所以你就把注意力转向海洋。你可能被水面上耀眼的阳光一时弄得无所适从,但一旦你意识到这是由于太阳光的反射引起的时候,你会发觉海洋是一块平坦而连绵不断的广阔天地。在无边无际的洋面上,船只很容易被发现(如果你已经从空中看见过船,你会懂得如何在水面上寻找船迹)。近岸的图象则比较混乱,但你却知道,紧靠岸边的不是停泊的船只,就是进入船坞里的船只。如果你有一张地图(图1.1(b)),这张地图可以帮助你确定船坞的位置(图1.1(c));对于一个质量较差的照片,地图还可以帮助你辨识海岸线。按此方式在地图和图象间建立起对应关系则会收到节省时间的效果。沿平行于船坞区的海岸线进行搜索,就可以发现几条船只(图1.1(d))。

另一个例子是,假设给你一组显示人体腹部“切片”的计算机辅助层析(CAT)扫描图片(图1.2(a))。这些图象是用一种高级技术得到的,使用常规的X射线方法得不到这种图象。你的工作是由这些断面图来重新构造肾脏的三维形状。这个任务似乎比



(a)

(b)

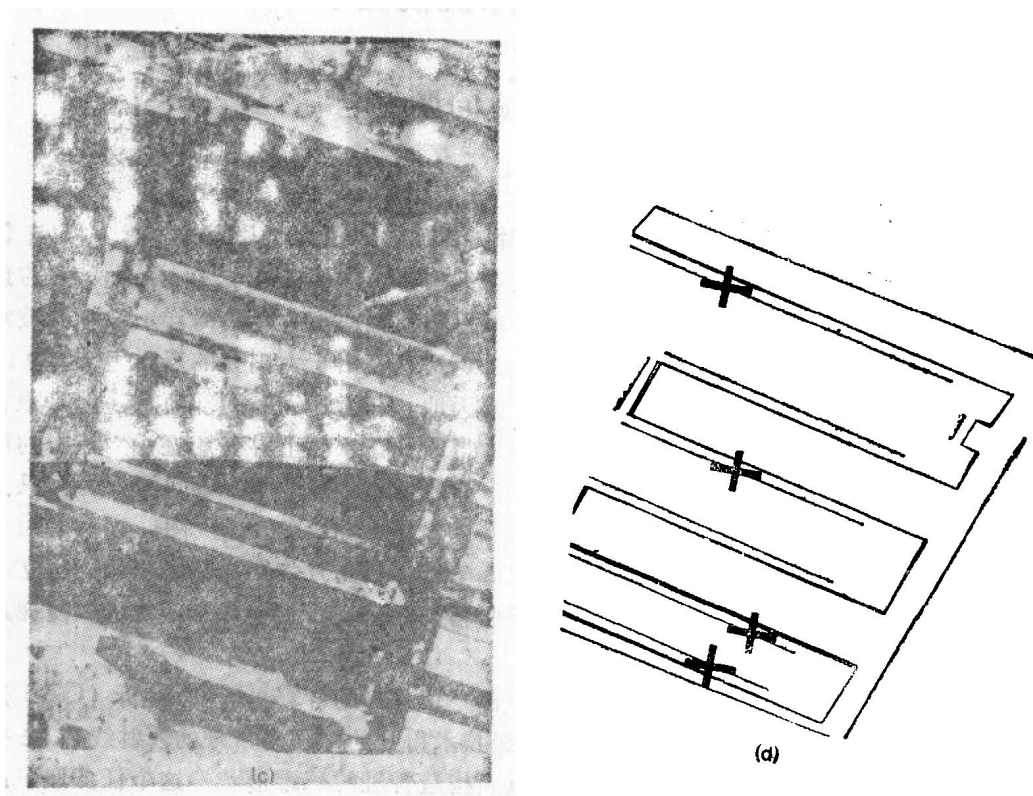


图 1.1 从航空照片中寻找船只

(a) 照片. (b) 对应的地图. (c) 照片中的船坞区. (d) 配准的地图和图象以及船的位置

找船要困难些. 你首先需要知道: 寻找什么 (图1.2(b)), 在CAT扫描图片中的什么地方去找它, 以及在这种扫描图片中它应当是什么样子. 你必须能在心里把这些图片“堆积”起来, 并形成一个肾脏形状的固有模型, 如同其切片所显示的那样(图1.2(c), (d)).

本书讨论计算机视觉. 上述两个例子是典型的计算机视觉方面的工作. 二者都是用本书各章节中所提到的各种知识和方法通过计算机来解决的. 计算机视觉是将用于视觉感知的各种处理和表示综合在一起, 并使其实现自动化的一种技术. 它包含许多本身已实用的技术, 如图象处理 (图象的变换、编码及传输) 统计模式分类 (用于一般模式、视觉或其它方面的统计决策理论). 对于我们来说更重要的是, 它还包含几何建模和认识处理的一些方法.

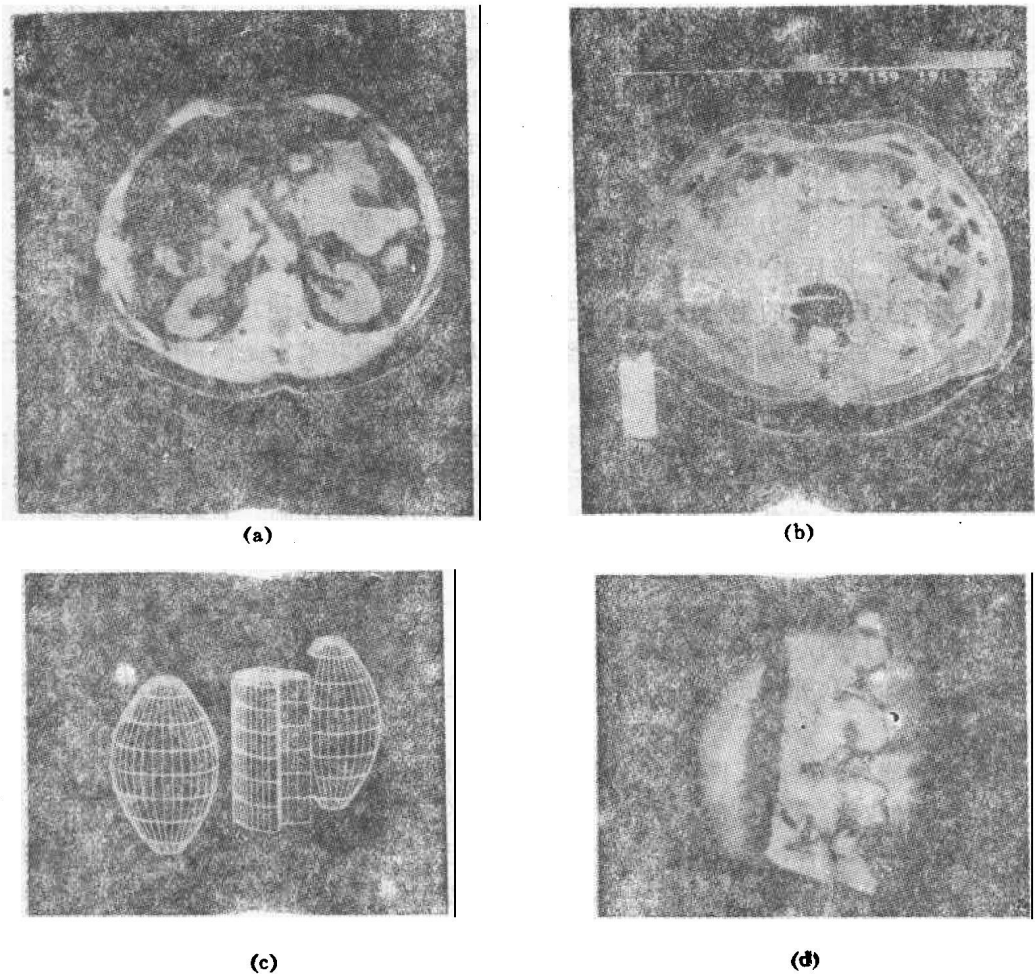


图 1.2 从计算机辅助层析图象中查找肾脏。

(a) 一个扫描数据切片。 (b) 肾脏模型样本。 (c) 模型拟合。 (d) 最终得到的肾脏和脊髓

1.2 高级和低级能力

1.1节的例子表明视觉需要使用认知处理、几何模型、目标和规划。这些高级处理很重要，我们的例子仅简略地表示了它们的能力和范围。找船肯定有一个总的目标，可能还有一些附属的信息，如在港口内有潜艇、驳船或小船等等。CAT扫描图片的使用还要考虑某些诊断目标和可供利用的有关医疗史。目标和知识属于指导视觉活动的高级能力，因此视觉系统应能利用这些能力。

即使这样复杂的任务也还只是很特殊的情形，比平常的视觉感知（如抱起一个婴儿、横穿繁华的街道、或去参加一个晚会时，能迅速地“发现”你所认识的人，主人屋子装饰的格调以及晚会进行了多久等等）要简单。所有这些任务都需要进行判断，并需要大量的有关实际物体是什么样子，以及如何行动等方面的知识。这些高级能力被完美