



智能

科/学/技/术/著/作/丛/书

序列图像中运动目标 跟踪方法

李培华 著



科学出版社

www.sciencep.com

序列图像中运动目标跟踪方法

李培华 著

科学出版社

北 京

内 容 简 介

本书主要内容包括目标跟踪的介绍(第1章)、基于滤波理论的目标跟踪方法(第2~4章)、基于 Mean Shift 的目标跟踪方法(第5~9章)和基于偏微分方程的目标跟踪方法(第10章),介绍了相关跟踪方法的研究背景、理论基础、算法描述、计算复杂性分析,并给出了相应的实验结果。本书是序列图像中目标跟踪方面的专著,反映了作者近年来在这一领域的主要研究成果。

本书内容新颖、结构清晰、语言简练,可作为大专院校及科研院所图像处理、视频处理和计算机视觉等领域的高年级本科生、研究生的教材和参考书,也可作为相关领域的教师、科研人员以及从事安防、视频监控行业的工程技术人员参考书。

图书在版编目(CIP)数据

序列图像中运动目标跟踪方法 / 李培华编. —北京:科学出版社,2010
ISBN 978-7-03-027242-3

I. ①序… II. ①李… III. ①计算机视觉-目标识别 IV. ①TP391.41

中国版本图书馆 CIP 数据核字(2010)第 069164 号

责任编辑:张海娜 / 责任校对:赵桂芬

责任印制:赵 博 / 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>

源海印刷有限责任公司印刷

科学出版社发行 各地新华书店经销

*

2010年4月第 一 版 开本:B5(720×1000)

2010年4月第一次印刷 印张:11 3/4

印数:1—3 000 字数:219 000

定价:35.00 元

(如有印装质量问题,我社负责调换)

本书获国家自然科学基金(60505006、60673110、60973080)
和黑龙江大学重点学科出版经费的资助

前 言

20 世纪 70 年代中期, Hogg 等率先开展了序列图像中目标跟踪方面的探索, 致力于跟踪场景中的行人并试图进行识别。在目标跟踪的研究历史中具有里程碑意义的文章是 Kass 等在 80 年代末发表的关于主动轮廓线模型的论文。从 90 年代初期到现在的近二十年中, 由于计算机和摄像机性价比的大幅提高, 以及目标跟踪在智能人机交互、医疗诊断、智能机器人、视频监控、视频编码和军事等领域不断增长的广泛应用价值, 目标跟踪逐渐成为图像处理和计算机视觉领域的研究热点。

在实际应用中, 由于摄像机和目标之间的相对运动和应用场景的复杂多变, 采集的视频图像一般具有光照变化明显、图像中杂物和噪声显著、目标的姿态变化很大等特点, 因此目标跟踪是图像处理和计算机视觉中一个公认的具有挑战性的问题。它的研究需要图像处理、模式识别、概率论及随机过程、泛函及优化和偏微分方程等多学科的知识。

序列图像中目标跟踪是指根据视频图像中的时空相关信息在每一帧图像中确定目标的位置、大小或形状信息等。目标跟踪方法大致可以分成三大类: 基于滤波理论的目标跟踪方法、基于 Mean Shift 的目标跟踪方法和基于偏微分方程的目标跟踪方法。

本书的第 1 章是绪论, 主要介绍了目标跟踪的应用和分类。

本书的第 2~4 章主要研究基于滤波理论的目标跟踪方法。基于滤波理论的目标跟踪方法将传感器噪声和模型误差等进行建模, 将目标跟踪问题表达为系统状态的后验概率密度估计问题。当后验概率密度是高斯分布时, 卡尔曼或扩展的卡尔曼滤波器可以对系统的状态进行准确估计, 而粒子滤波器能够解决更一般的非线性非高斯的目标跟踪问题。

本书的第 5~9 章主要研究基于 Mean Shift 的目标跟踪方法。基于 Mean Shift 的目标跟踪方法首先建模目标特征的概率密度分布, 然后采用概率密度之间的相似性测度度量目标和候选目标之间的相似性。通过将梯度下降算法应用于相似性函数, 推导出 Mean Shift 迭代方程, 从而对目标进行跟踪。

第 10 章主要研究一种基于偏微分方程的目标跟踪方法。在基于偏微分

方程的目标跟踪方法中,目标跟踪问题被建模为关于曲线能量的泛函。通过变分法或形状导数法等推导出描述轮廓曲线进化的偏微分方程,通过求解偏微分方程得到泛函的极值,从而获得目标的轮廓。

感谢作者的导师张田文教授将作者引入图像处理及目标跟踪这一富有生机、蓬勃发展的研究领域,本书的部分研究成果是作者攻读博士学位期间完成的。感谢作者的硕士研究生肖莉娟和赵楠楠,她们在攻读硕士学位期间对书中提出的新方法做了大量实验。作者的硕士研究生宋琦、赵楠楠和马宏伟帮助完成了部分文字的录入工作,在此表示感谢。

本书的研究工作得到多项国家自然科学基金(60505006、60673110、60973080)和黑龙江大学重点学科出版经费的资助。没有这些资助,本书的研究工作就不可能完成。在本书出版之际,作者向支持自己的机构致以诚挚的谢意。

由于作者水平有限,书中不足之处在所难免,希望读者批评指正。

目 录

前言

第 1 章 绪论	1
1.1 目标跟踪的应用	1
1.1.1 在智能人机交互中的应用	1
1.1.2 在医疗诊断中的应用	4
1.1.3 在智能机器人中的应用	5
1.1.4 在视频监控中的应用	7
1.2 目标跟踪的分类	8
1.2.1 基于滤波理论的目标跟踪方法	9
1.2.2 基于 Mean Shift 的目标跟踪方法	11
1.2.3 基于偏微分方程的目标跟踪方法	13
参考文献	16
第 2 章 基于 Unscented 卡尔曼滤波器的目标跟踪	23
2.1 引言	23
2.2 目标轮廓的 B 样条曲线表达	24
2.2.1 样条函数空间和 B 样条基函数	24
2.2.2 B 样条曲线空间、样条向量空间和形状空间及其内积和范数	27
2.3 运动模型的学习	29
2.3.1 运动模型的最大似然学习	29
2.3.2 运动模型的最大期望学习	30
2.4 卡尔曼滤波器跟踪算法	32
2.5 Unscented 卡尔曼滤波器跟踪算法	32
2.5.1 Unscented 变换	33
2.5.2 非线性观测模型	34
2.5.3 跟踪算法	35
2.5.4 实验	37
2.6 本章小结	41
参考文献	41

第 3 章 基于粒子滤波器的目标跟踪算法	43
3.1 引言	43
3.2 动态模型、形状模型及观测模型	44
3.2.1 动态模型	44
3.2.2 形状模型和观测模型	45
3.3 标准的粒子滤波器和条件概率密度传播算法	45
3.3.1 一般的目标跟踪问题	45
3.3.2 一般的粒子滤波器	46
3.3.3 条件概率密度传播算法	47
3.4 基于卡尔曼粒子滤波器和 Unscented 粒子滤波器的目标跟踪算法	48
3.4.1 基于卡尔曼粒子滤波器的跟踪算法	48
3.4.2 基于 Unscented 粒子滤波器的跟踪算法	50
3.4.3 算法性能分析和时间分析	52
3.4.4 实验	53
3.5 本章小结	57
参考文献	57
第 4 章 基于积分图像的快速粒子滤波器跟踪	60
4.1 引言	60
4.2 基于颜色的粒子滤波器算法	61
4.2.1 自适应颜色模型	61
4.2.2 基于积分图像的颜色模型计算	62
4.2.3 计算积分图像的并行算法	64
4.3 跟踪算法及实验	67
4.3.1 跟踪算法	67
4.3.2 实验	67
4.4 本章小结	75
参考文献	76
第 5 章 利用二阶信息的 Mean Shift 目标跟踪算法	78
5.1 引言	78
5.2 传统 Mean Shift 跟踪算法	78
5.2.1 非参量核概率密度估计和 Mean Shift 迭代	78
5.2.2 模型表达及跟踪算法	80

5.3 牛顿法	82
5.3.1 步长选择模型及终止条件	84
5.3.2 步长选择算法	85
5.4 信赖域方法	87
5.4.1 信赖域算法	88
5.4.2 信赖域方向	89
5.5 实验	90
5.6 本章小结	94
参考文献	94
第 6 章 基于聚类的颜色模型及 Mean Shift 跟踪	96
6.1 引言	96
6.2 基于聚类的颜色模型	98
6.2.1 颜色空间的自适应剖分	98
6.2.2 颜色模型及相似性度量	99
6.3 Mean Shift 跟踪算法	100
6.3.1 算法的推导过程	100
6.3.2 算法的复杂性分析	101
6.3.3 实验	102
6.4 本章小结	108
参考文献	108
第 7 章 基于 Mean Shift 的仿射变换和相似性变换目标跟踪	110
7.1 引言	110
7.2 基于 Mean Shift 算法的仿射运动目标跟踪	111
7.2.1 平移向量的估计	113
7.2.2 旋转角度的估计	113
7.2.3 跟踪算法	114
7.2.4 实验	115
7.3 基于 Mean Shift 算法的相似性变换目标跟踪	120
7.3.1 平移向量的估计	121
7.3.2 旋转角度的估计	121
7.3.3 跟踪算法	122
7.3.4 实验	122
7.4 计算复杂性分析	125

7.5 本章小结	125
参考文献	126
第 8 章 并行 Mean Shift 目标跟踪算法	128
8.1 引言	128
8.2 CUDA 介绍	129
8.3 基于多元高斯分布的颜色空间剖分	131
8.4 基于 GPU 的并行 Mean Shift 跟踪算法	133
8.5 实验	136
8.6 本章小结	141
参考文献	141
第 9 章 基于 EMD 的 Mean Shift 跟踪算法	142
9.1 引言	142
9.2 颜色概率密度表达及 EMD 定义	143
9.2.1 颜色概率密度表达	143
9.2.2 EMD 定义	146
9.3 单纯形法计算 EMD	147
9.4 基于 EMD 的 Mean Shift 算法	152
9.4.1 相似性度量函数最优化	152
9.4.2 基于 EMD 的 Mean Shift 算法	153
9.5 实验	153
9.5.1 走廊序列	153
9.5.2 行人序列	155
9.6 本章小结	156
参考文献	156
第 10 章 基于偏微分方程的目标跟踪方法	157
10.1 引言	157
10.2 基于 Jensen-Shannon 散度的区域泛函	157
10.3 偏微分方程的推导	158
10.3.1 形状导数理论简介	158
10.3.2 基于形状导数理论的偏微分方程推导	161
10.3.3 基于变分理论的偏微分方程推导	162
10.4 目标跟踪的水平集方法	162
10.4.1 水平集函数的推导	162

10.4.2 偏微分方程的时空离散化	163
10.4.3 数值算法中的重新初始化	166
10.4.4 实验	170
10.5 本章小结	173
参考文献	173

第 1 章 绪 论

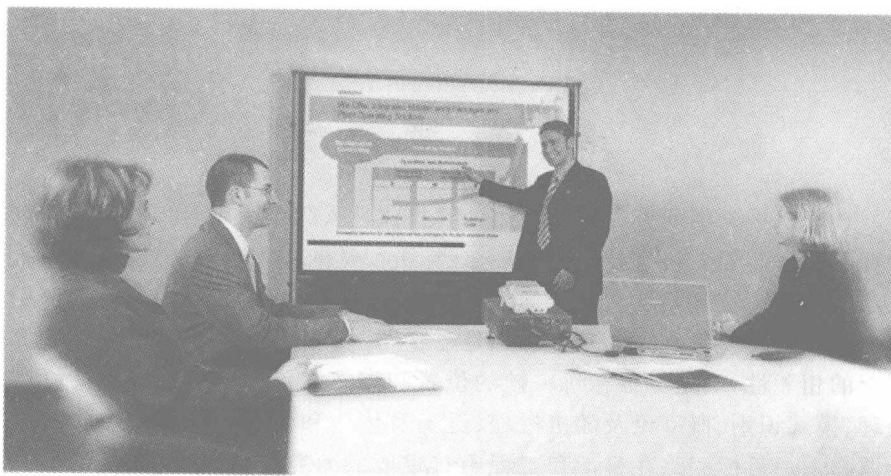
计算机视觉的研究目的是使用计算机代替人眼及大脑,对环境进行感知、解释和理解。随着科学技术的发展,计算机视觉的应用范围日益扩大,对人类生产生活的影 响日益增大。目标跟踪作为计算机视觉的重要研究方向吸引了越来越多的研究者。它的基本思想是在序列图像中根据视频信息在空间和时间上的相关性,确定目标在每一帧的位置和姿态。目标跟踪是一项融合图像处理、模式识别、概率论及随机过程、泛函及优化和偏微分方程等多学科的复杂课题^[1]。目标跟踪在科学和工程中有重要的研究价值,在生产生活中有广泛的应用前景。它已经在智能人机交互、医疗诊断、智能机器人、视频监控等许多领域获得广泛应用。

1.1 目标跟踪的应用

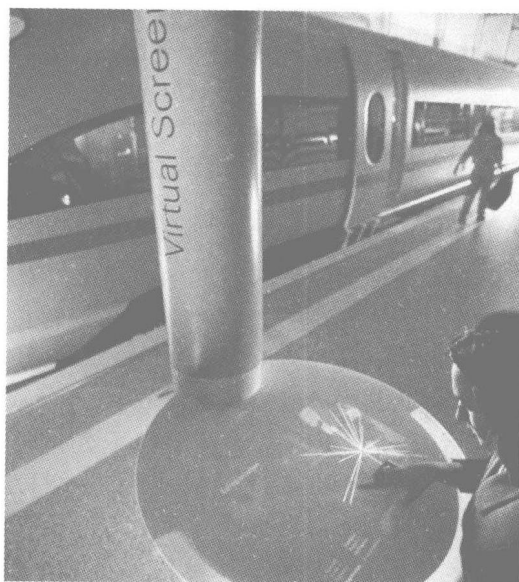
1.1.1 在智能人机交互中的应用

计算机计算能力的提高遵循摩尔定律,呈指数倍增长。然而目前的人机交互仍旧停留在以键盘和鼠标为主。人们期望计算机具有智能,能够以自然的方式与人类进行交流^[2~4]。在虚拟现实中得到广泛研究的设备如数据手套、红外头盔等是接触式的,但使用者穿戴起来比较笨重、烦琐。跟踪技术能够提供一种灵活的、方便的、非接触的输入方式。

在人机交互领域内最为成功的发明是西门子公司开发的虚拟触摸屏 SIVIT^[5,6]。该系统包括投影仪、红外发射器、摄像机和手部跟踪识别模块。SIVIT 将监视器的内容投影到演讲白板或桌面上,其操作通过将索引手指置于光束中并指向目标来实现。SIVIT 在摄像机记录的序列图像中连续跟踪手的轮廓,将索引手指的运动理解为鼠标的移动,手指的短暂停顿理解为按动鼠标。图 1.1(a)说明了西门子公司 Folex SIVIT Walkboard 演示系统。当演讲者进行演示时,把图像投影仪与个人计算机连接起来,将图像投影到反射墙或幻灯幕上,使用索引手指来操作。图 1.1(b)演示了安装在火车站的公共信息查询终端,用户接触的只有一个普通的桌面,用来投影显示器的内容。



(a) Folex SIVIT Walkboard 演示系统



(b) 公共信息查询终端

图 1.1 西门子虚拟触摸屏技术的应用

剑桥大学工程系信息工程实验室开发了基于指向手势的人机接口^[7]。摄像机在约 1.6m 远的距离上拍摄机械手、工作空间及操作者的指向手势。操

作者指向目标,机械手抓取该目标。该系统在立体摄像机拍摄的两个序列图像中分别实时跟踪指向手势,通过三维-二维的摄影变换来确定指向目标,然后指示机械手抓取该目标。图 1.2 说明了这一人机接口系统,其中图 1.2(a)和图 1.2(b)说明了未确定目标的立体摄像机、桌面及机器人的放置,图 1.2(c)和图 1.2(d)是立体摄像机对这一场景采集的一对图像。

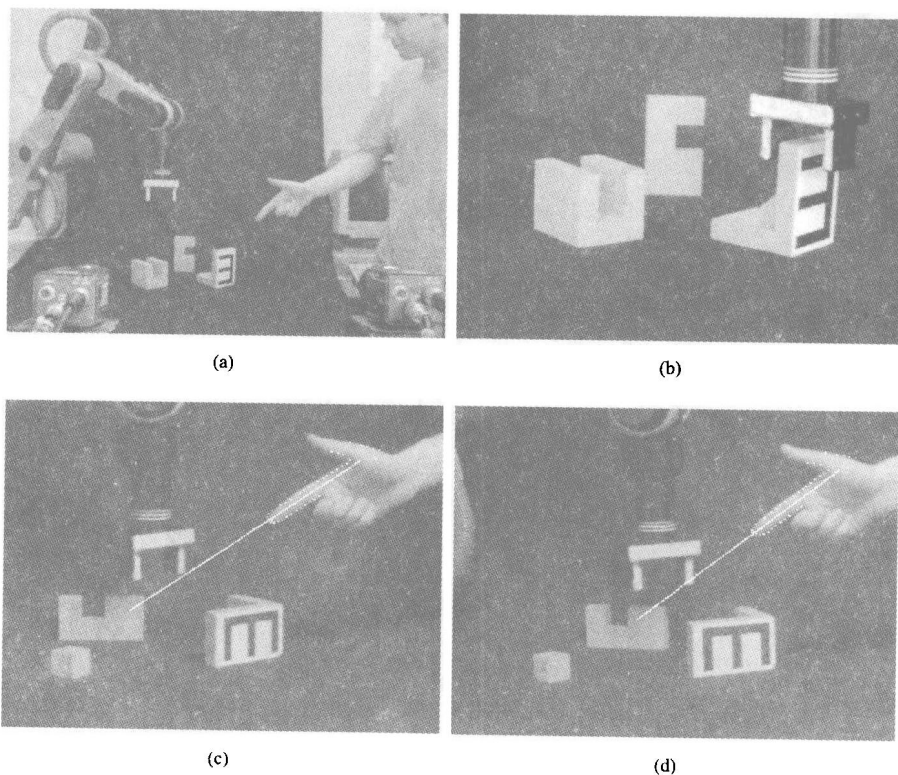


图 1.2 基于指向手势的人机接口

牛津大学工程科学系机器人研究组开发了基于视觉的绘制系统^[8,9]。该系统仅仅使用计算机、彩色摄像机和图像采集卡,不使用其他任何设备,能够实现自动初始化以及实时、鲁棒而准确的绘制。图 1.3 演示了这一系统。

语音识别在低噪声环境中取得了很大成功,然而在存在较大背景噪声的环境中,特别是当存在其他人交叉谈话的情况下,识别率会大大降低。研究表明,基于视觉的唇读对于提高语音识别的鲁棒性起到重要作用^[10~13],如图 1.4所示。唇读是通过在获得语音信号的同时同步地跟踪唇部的内外轮廓

得以实现的。通过这一辅助视觉信息,自动语音识别的性能得到改善和增强。

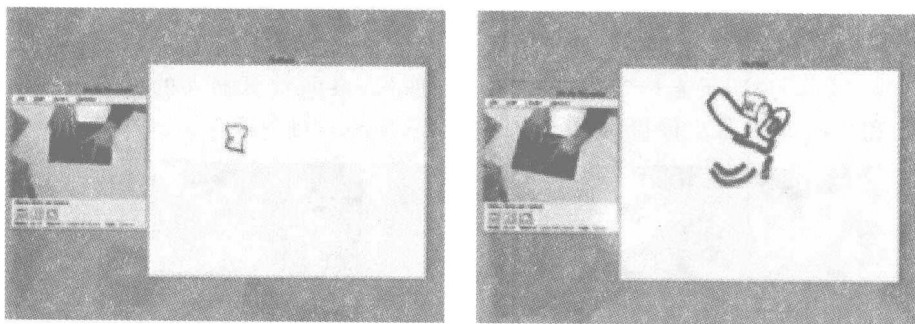


图 1.3 基于视觉的绘制系统

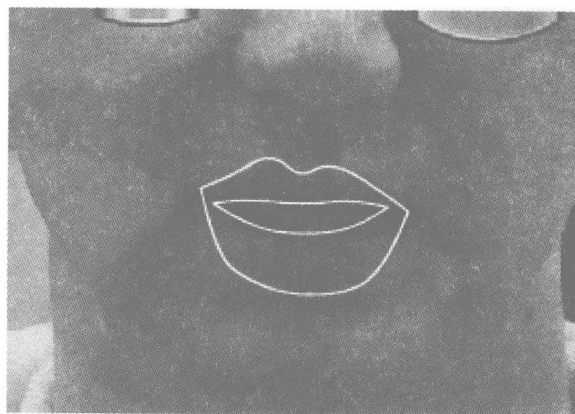


图 1.4 自动语音识别的性能通过唇读得到改善和增强

1.1.2 在医疗诊断中的应用

运动目标的跟踪和分割在医学图像处理中应用十分广泛^[14~19]。超声波扫描器(ultrasound scanner)由于成本低廉而成为广泛应用的医疗诊断成像设备,它们能够产生实时的视频序列。由于超声波图像中噪声的存在使单帧图像中有用的信息几乎被淹没,使得静态分析十分困难,而如果利用序列图像中目标的几何上的连续性和时间上的相关性进行动态分析,得到的结果将更加准确、更有价值。Jacob 和 Ayache 等使用主动轮廓线跟踪超声波序列图像中心室的跳动,能够为诊断心脏病变提供准确可靠的信息,如图 1.5 所示。磁

共振成像(MRI)由于对人体无害和成像质量高也很受欢迎。Marais 等跟踪图像序列中每一帧的脑半球的断层扫描图像,最后将跟踪结果应用于脑半球的三维重建,如图 1.6 所示。

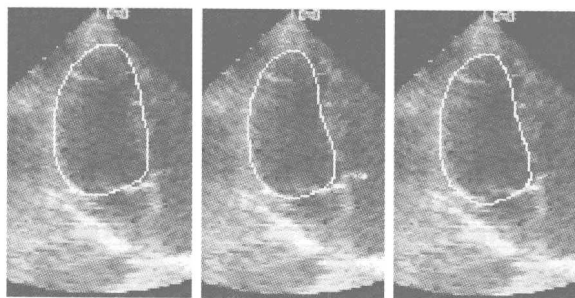


图 1.5 跟踪心室跳动

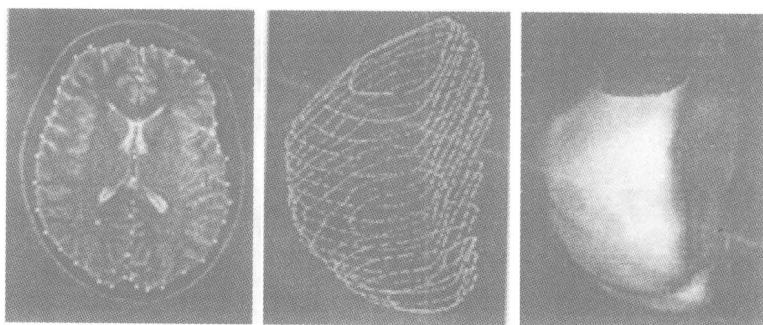


图 1.6 脑半球的 MRI 成像

1.1.3 在智能机器人中的应用

视觉传感器是智能机器人的一种十分重要的信息源。为了能够自主运动,智能机器人需要能够跟踪和识别环境中的物体^[20~24]。一个非常重要的具有应用前景的自主机器人的例子是无人驾驶自动车辆,它们必须能够跟踪道路的特征和其他移动车辆^[25~27]。在机器人手眼应用中,安装在机械手上的单目或多目摄像机可以运动,从不同的视角采集目标的图像,用来恢复物体的三维结构或计算最优的抓取方式^[28~30]。基于视觉跟踪的机器人能够进行农作物的自动收割、自动识别和肥料、药剂的自动喷洒^[31,32]。图 1.7 显示了一个自主的农业机器人系统,其中包括移动车辆、车载计算机、摄像机和喷嘴等。

在车辆运动过程中,摄像机拍摄喷嘴下的植物的序列图像,计算机跟踪植物的运动和形状,然后指示喷嘴按照需要将化学肥料喷洒到植物的旁边,或远离它们。在类机器人的研究中,图 1.8(a)显示了本田公司研发的类机器人阿西莫(ASIMO),图 1.8(b)显示了阿西莫与人握手的情景。阿西莫的头顶安装有两架视频摄像机,用来检测目标、感知运动、识别人脸和理解手势等。目标跟踪算法在其中起着重要的作用。例如,通过对移动目标的连续跟踪,阿西莫能够估计出目标的距离和运动方向,从而能够选择合适的行为(避开或停止运动),在对手势的连续跟踪基础上进一步对人的手势进行识别等。

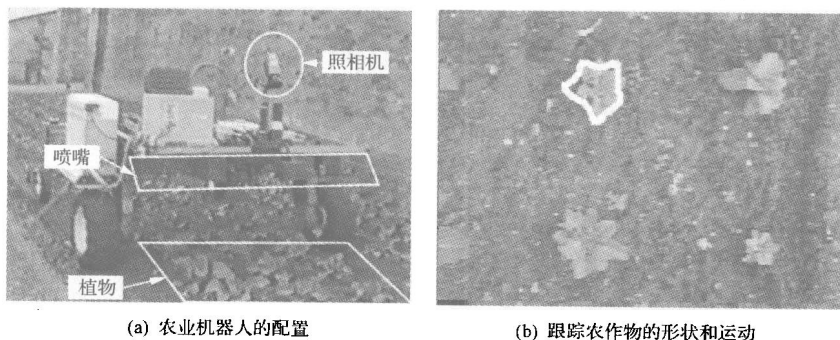


图 1.7 农业机器人

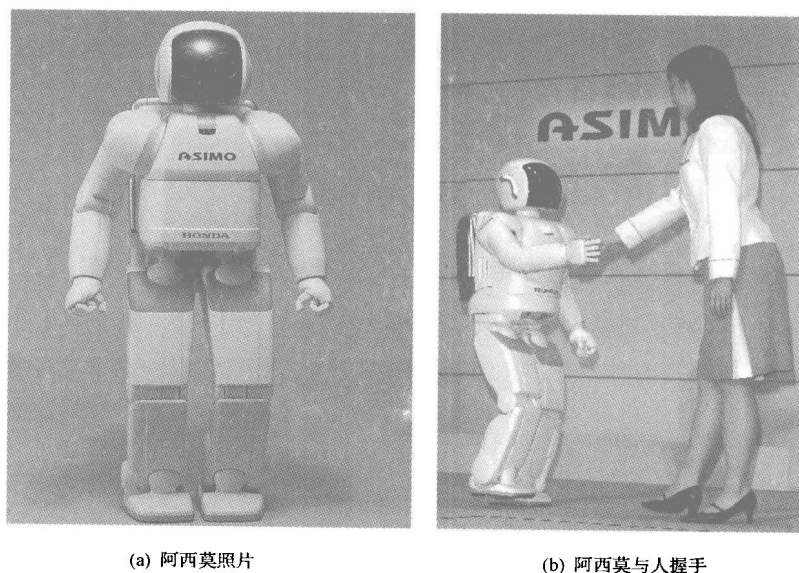


图 1.8 本田公司研发的类机器人阿西莫