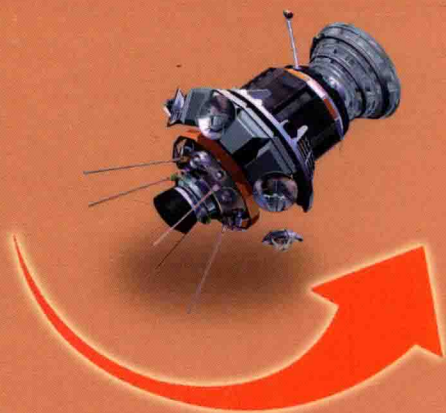


# 计算机视觉中的目标特征模型 和视觉注意模型

刘 玮 魏龙生 / 编 著



华中科技大学出版社  
<http://www.hustp.com>

# 计算机视觉中的目标特征 模型和视觉注意模型

刘 玮 魏龙生 编著

华中科技大学出版社  
中国·武汉

## 内 容 简 介

本书选择了计算机视觉研究领域目前最核心同时也是最热门的目标特征模型和视觉注意模型两大部分作为研究内容。第一部分,目标特征模型。根据目标特征由简单到复杂,其包含的语义信息由单一到复合,介绍了目标在不同层面上的特征模型建立及其在目标识别检测中的应用。目标特征模型包括目标全局特征模型、目标局部特征模型、目标结构特征模型和目标生物视觉特征模型。第二部分,视觉注意模型。根据视觉信息处理中注意机制计算模型中存在的若干问题,以视觉心理学和生理学的最新研究成果为依据,介绍了基于目标背景对比度的视觉注意计算模型、基于目标自身特性的视觉注意计算模型、基于动态和静态的视觉注意计算模型及基于视觉注意区域可变分辨率的图像压缩模型。

### 图书在版编目(CIP)数据

计算机视觉中的目标特征模型和视觉注意模型/刘玮,魏龙生编著. —武汉:华中科技大学出版社,2016.9

ISBN 978-7-5680-2189-0

I. ①计… II. ①刘… ②魏… III. ①计算机视觉-研究 IV. ①TP302.7

中国版本图书馆 CIP 数据核字(2016)第 213968 号

### 计算机视觉中的目标特征模型和视觉注意模型

刘 玮

Jisuanji Shijue Zhong De Mubiao Tezheng Moxing He Shijue Zhuyi Moxing

魏龙生

编著

策划编辑:范 莹

责任编辑:陈元玉

封面设计:原色设计

责任校对:张 琳

责任监印:朱 玟

出版发行:华中科技大学出版社(中国·武汉)

武汉市东湖新技术开发区 邮编:430223 电话:(027)81321913

录 排:武汉市洪山区佳年华文印部

印 刷:虎彩印艺股份有限公司

开 本:710mm×1000mm 1/16

印 张:13

字 数:260千字

版 次:2016年9月第1版第1次印刷

定 价:32.00元



华中出版

本书若有印装质量问题,请向出版社营销中心调换  
全国免费服务热线:400-6679-118 竭诚为您服务  
版权所有 侵权必究

# 前 言

对计算机视觉领域的研究历史悠久,但是对其在工业界应用的探索还处于刚起步阶段,随着每年图像和视频数据以数十倍的速度增长,同时随着计算机技术、网络技术和传感器技术的发展,计算机视觉领域显示出巨大的发展空间。计算机视觉领域包含很多研究内容,其中自动目标识别是计算机视觉中最主要也是应用最广泛的研究内容,而视觉注意机制的研究则是计算机视觉近些年最新和最热门的研究内容,同时两部分内容可以互为交叉和补充。本书以这两部分研究内容为基础,主要阐述了自动目标识别中的目标特征模型和视觉注意机制中的视觉注意模型的分析与建模。

本书可供从事计算机视觉研究的科研人员了解一般的目标特征模型和主要的视觉注意模型,同时提供合理的建模思路和方法。

本书共分为 10 章,第 1 章介绍了计算机视觉的范畴、应用和关键研究问题。基于识别问题提出了目标特征模型和视觉注意模型所研究的内容。

第 2 章讲述了目标特征模型的层级结构。根据目标特征所包含的信息从底层到高层,从少到多,建立了特征模型。模型分为两大层,包含目标的基本特征和目标的高级特征。其中基本特征又分为全局特征和局部特征,高级特征又分为结构特征和视觉特征。

第 3 章基于目标的初级特征对目标进行建模。本章以目标的全局特征为基础建立机场目标模型,并将该模型应用于前下视红外机场自动识别中。这里以目标的两种全局特征建立不同的目标特征模型并采用不同的分类识别方法。

第 4 章基于目标的局部特征建模,根据本书分析统计模型提出了图像目标语义概率模型,并且将这种概率模型应用于目标识别和复杂场景下的地物分析。首先将图像表示成多个特征局部区域的集合,然后根据目标语义概率模型得到图像、特征局部和目标语义之间的概率关系,通过计算后验概率可以实现目标语义类别的识别。

第 5 章在局部特征的基础上引入了空间特征,结合两种特征得到了整体视觉模型。整体视觉模型不但考虑了目标的局部特征,而且考虑了目标局部之间的空间位置关系。本书将这一模型应用于遥感图像的典型地物识别之中,不但能够识别单目标,还可以识别多目标,识别率较高,且定位较准确。

第6章基于大脑皮层视觉特征对目标进行建模,并基于该模型实现目标的识别和检测算法。脑颞叶皮层视觉认知系统的标准量化模型是一种前馈式层级模型,其很好地仿效了脑皮层视觉从简单认知元到复杂认知元的识别机理,由量化层级模型获得高级视觉特征。基于量化模型特征矢量建立了高性能的目标识别系统。

第7章介绍了基于目标背景对比度的视觉注意模型,该模型将所有的训练目标融合成目标类,所有的训练背景融合成背景类,对于每一个特征,目标类与背景类的显著性均值的比值得到权重。在待注意场景中,特征图通过权重联合生成自顶向下的显著性图,与自底向上的显著性图融合生成全局显著性图。

第8章介绍了基于目标自身特性的视觉注意计算模型,模型根据目标自身而不依靠背景信息提取特征,并表示成均值和标准差。通过比较训练图和待注意图得到相似性图,所有相似性图被联合生成自顶向下的显著性图,与自底向上的显著性图乘积得到全局显著性图。

第9章介绍了动态和静态的视觉注意计算模型,对于视频的每一帧,提取颜色特征、亮度特征、方位特征和纹理特征,计算特征图中的熵值,这些熵值逐步融合形成动态显著性图;动态显著性图和静态显著性图融合生成最终显著性图。最后,计算出每个显著性区域的尺寸。

第10章介绍了基于视觉注意区域可变分辨率的图像压缩模型,显著的区域采用较高的分辨率,不显著的区域采用较低的分辨率,这样达到一个可变分辨率的效果,该模型不仅可以对整个图像达到一个高压缩率的效果,还可以保持显著性区域的高分辨率。由于该模型在设计 and 实现过程中没有任何先验的知识,所以可被用于任何场景中。这为图像压缩提供了一种新的方法。

本书作者刘玮博士及魏龙生博士所带领的团队长期从事计算机视觉的相关研究,在自动目标识别、目标检测跟踪和视觉注意等方面取得了较多研究成果。该团队主持2项国家自然科学基金、1项省部级项目、2项校级优秀青年项目,参与多项自然科学基金、973及863等重大项目。在国内外权威学术期刊及会议上发表了多篇与计算机视觉相关的论文,其中SCI论文6篇、EI检索论文23篇;获发明专利2项。

书稿在两位作者的博士论文基础上编写完成,但是由于水平有限,时间紧张,书中难免出现疏漏,希望读者提出宝贵意见,以便再版时修改和完善。

编者

2016年4月

# 目 录

<b>第 1 章 绪论</b> .....	(1)
1.1 计算机视觉 .....	(1)
1.2 与计算机视觉相关的学科 .....	(3)
1.3 计算机视觉的应用 .....	(4)
1.4 计算机视觉的经典问题 .....	(6)
1.5 目标特征模型 .....	(7)
1.6 视觉注意模型.....	(13)
<b>第 2 章 目标特征模型</b> .....	(19)
2.1 目标的基本特征.....	(19)
2.2 目标的高层特征.....	(27)
2.3 本章小结.....	(30)
<b>第 3 章 基于全局特征的目标模型</b> .....	(31)
3.1 引言.....	(31)
3.2 目标全局特征建模.....	(32)
3.3 目标建模预处理.....	(33)
3.4 基于不变矩的机场建模及识别算法.....	(35)
3.5 基于形状特征的机场建模及识别算法.....	(46)
3.6 本章小结.....	(59)
<b>第 4 章 基于局部特征的目标模型</b> .....	(61)
4.1 引言.....	(61)
4.2 基于局部特征的目标模型.....	(62)
4.3 图像目标建模方法.....	(63)
4.4 目标局部模型在目标识别和场景分析中的应用.....	(67)
4.5 本章小结.....	(87)
<b>第 5 章 基于结构特征的目标模型</b> .....	(88)
5.1 引言.....	(88)
5.2 目标的结构化模型.....	(89)
5.3 目标的整体视觉模型.....	(90)
5.4 典型目标结构模型的实现.....	(96)
5.5 结构模型在目标检测中的实现及应用.....	(97)

5.6	本章小结 .....	(109)
<b>第 6 章</b>	<b>基于视觉特征的目标模型</b> .....	(110)
6.1	引言 .....	(110)
6.2	大脑皮层视觉认知的前馈式层级模型 .....	(111)
6.3	目标视觉认知模型及其应用 .....	(114)
6.4	本章小结 .....	(120)
<b>第 7 章</b>	<b>基于目标背景对比度的视觉注意计算模型</b> .....	(121)
7.1	引言 .....	(121)
7.2	自底向上显著性图 .....	(123)
7.3	自顶向下显著性图 .....	(125)
7.4	全局显著性图及显著性区域的尺寸 .....	(128)
7.5	视觉转移 .....	(129)
7.6	实验结果与分析 .....	(129)
7.7	本章小结 .....	(132)
<b>第 8 章</b>	<b>基于目标自身特性的视觉注意计算模型</b> .....	(133)
8.1	引言 .....	(133)
8.2	目标表示 .....	(135)
8.3	自顶向下的显著性图 .....	(137)
8.4	全局显著性图 .....	(137)
8.5	实验结果与分析 .....	(138)
8.6	目标背景对比度模型与目标自身特性模型的比较 .....	(145)
8.7	本章小结 .....	(145)
<b>第 9 章</b>	<b>基于动态和静态的视觉注意计算模型</b> .....	(147)
9.1	引言 .....	(147)
9.2	动态显著性图 .....	(148)
9.3	静态显著性图和时空显著性图 .....	(152)
9.4	实验结果与分析 .....	(153)
9.5	本章小结 .....	(161)
<b>第 10 章</b>	<b>基于视觉注意区域可变分辨率的图像压缩模型</b> .....	(162)
10.1	引言 .....	(162)
10.2	图像压缩方法 .....	(163)
10.3	可变分辨率的图像压缩模型 .....	(165)
10.4	实验结果与分析 .....	(166)
10.5	本章小结 .....	(171)
<b>参考文献</b>	.....	(172)

### 1.1 计算机视觉

Google 的自动驾驶小萌车可在街道上自由穿行(见图 1-1(a)),波士顿动力公司的机器人 Atlas 能在雪地行走(见图 1-1(c)),大疆推出的多旋翼无人机 Phantom4 可以低空贴地任意自主智能飞行(见图 1-1(b)),微软的 KINECT 技术让电影《少数派报告》中手势操作人机界面场景成为现实(见图 1-1(d)),最酷的可穿戴设备 Google 眼镜将虚拟现实带入现实生活(见图 1-1(e))。所有这些最新的科技产品的背后都离不开一种最关键的技术,即计算机视觉技术。

人类获取的外界信息,主要是通过视觉、听觉、触觉、嗅觉和味觉等感觉器官得到的,其中 80% 的信息获取都来自视觉,而且视觉获取的信息也是最丰富、最复杂的。人的生理构造决定了我们能够看清楚并理解身边的场景,但是要让计算机看懂这个世界,却是一件非常困难的事情。即使在很多人看来,现在的计算机技术已经足够先进了,但是要达到看懂并自主分析各种复杂信息的程度,还有很长的一段路要走,这就是计算机视觉这门学科要做的事情。

什么是计算机视觉? 维基百科给出了相对严谨的概念,计算机视觉是一门研究如何使机器“看”的学科,更进一步地说,就是指用摄影机和计算机代替人眼对目标进行识别、跟踪和测量等机器视觉,并进一步进行图像处理,使计算机处理成为更适合人眼观察或传送给仪器检测的图像。俗话说,计算机视觉也称机器视觉,就是让计算机或机器像人类一样能够感知世界,即看到世界和分析世界的一门学科。作为一门科学学科,计算机视觉研究相关的理论和技术,试图建立能够从图像或者多维数据中获取“信息”的人工智能系统。这里所指的信息指 Shannon 定义的,可以用来帮助做

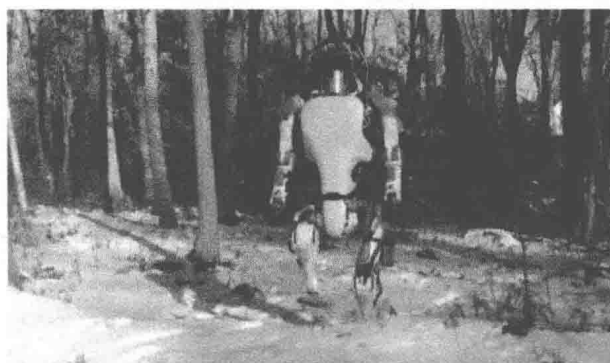




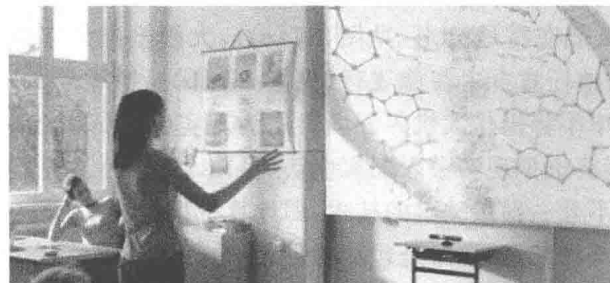
(a) Google自动驾驶小萌车



(b) 大疆无人机智能跟拍



(c) 机器人Atlas在雪地行走



(d) 体感人机交互



(e) Google智能眼镜

图 1-1 计算机视觉的应用

一个“决定”的信息。因为感知可以看成是从感官信号中提取信息,所以计算机视觉也可以看成是研究如何使人工系统从图像或多维数据中“感知”的学科。

## 1.2 与计算机视觉相关的学科

计算机视觉既属于工程领域,也是科学领域中的一个极富有挑战性的重要研究领域。计算机视觉是一门综合性的学科,它吸引了来自各个学科的研究者参与到对它的研究之中,其中包括计算机科学和工程、信号处理、物理学、应用数学和统计学、神经生理学和认知科学等。

### 1. 计算机视觉是计算机科学的重要分支

计算机视觉的研究最早开始于 20 世纪 70 年代,一些从事人工智能和机器人研究的先驱者们试图引入“视觉输入”来解决如高层次推理和规划中的困难问题。人工智能所研究的一个主要问题是如何让系统具备规划和决策能力,从而使之完成特定的技术动作(例如,在某种特定环境中)。移动一个机器人这一问题与计算机视觉问题息息相关,在这里,计算机视觉系统作为一个感知器,为决策提供信息。另外一些研究方向包括模式识别和机器学习(隶属于人工智能领域,都与计算机视觉有着重要联系),因此计算机视觉通常被视为计算机科学的一个重要分支。

### 2. 计算机视觉与物理有着重要联系

计算机视觉关注的目标在于充分理解电磁波(主要是可见光与红外线部分)遇到物体表面被反射所形成的图像,而这一过程是基于光学物理和固态物理的。一些尖端的图像感知系统甚至会应用到量子力学理论来解析影像所表示的真实世界。同时,物理学中的很多测量难题也可以通过计算机视觉得到解决,例如流体运动。因此,计算机视觉可以被看成是物理学的拓展。

### 3. 计算机视觉与神经生物学有着千丝万缕的联系

在整个 20 世纪,人类对各种动物的眼睛、神经元以及与视觉刺激相关的脑部组织都进行了广泛研究,这些研究得出了一些关于生物视觉系统如何运作的描述,这也形成了计算机视觉中的一个子领域。人们试图建立人工系统,使之在不同的复杂程度上模拟生物的视觉运作。同时在计算机视觉领域中,一些基于机器学习的方法也有参考部分生物机制。比如近几年最新的深度学习方法,仿照人的神经系统,提出的深度卷积神经网络目前应用于很多大数据相关的领域。而计算机视

觉反过来又可以作为研究探索人脑视觉工作机理的手段,进一步加深对人脑视觉机理的掌握和理解。

#### 4. 计算机视觉与信号处理密不可分

很多有关单元变量信号的处理方法,尤其是对时变信号的处理,都可以很自然地扩展为计算机视觉中对二元变量信号或者多元变量信号的处理方法。但由于图像数据的特有属性,很多计算机视觉中发展起来的方法,在单元信号的处理方法中却找不到对应版本。这类方法的一个主要特征便是它们的非线性以及图像信息的多维性。以上两点作为计算机视觉的一部分,在信号处理学中形成了一个特殊的研究方向。

除了上面提到的领域,很多研究课题同样可被当成纯粹的数学问题。例如,计算机视觉中的很多问题,其理论基础便是统计学、最优化理论以及几何学。

### 1.3 计算机视觉的应用

随着计算机视觉技术的不断发展,越来越多的研究和新技术、越来越多的全新用户体验方式正在强烈地冲击着人们传统的工作和生活方式,从日常超市条形码检测、上下班指纹考勤、人脸识别考勤、高速路车牌识别、航空遥感测控地形地貌、电影特效制作、工业生产自动化检测、医学影像检测,到航空天文领域等。基于计算机视觉的应用广泛遍布以下领域。



图 1-2 日本一家公司研发的用于核辐射和酷暑环境的机器人

随着计算机视觉技术的不断发展,越来越多的研究和新技术、越来越多的全新用户体验方式正在强烈地冲击着人们传统的工作和生活方式,从日常超市条形码检测、上下班指纹考勤、人脸识别考勤、高速路车牌识别、航空遥感测控地形地貌、电影特效制作、工业生产自动化检测、医学影像检测,到航空天文领域等。基于计算机视觉的应用广泛遍布以下领域。

(1) 工业视觉。将工业机器人应用到过程自动控制中,如自动生产流水线、邮政自动化、邮件分拣等;在制造业中实现自动检测,如无损探测、印刷版质量检验、精细印刷品缺陷检测等;在各种危险场合工作环境中应用机器人,如核辐射环境下的工作机器人,如图 1-2 所示。

(2) 视觉导航。无人驾驶汽车通过车载摄像头检测道路区域、识别道路交通标志实现导航驾驶;移动机器人,如火星探测器勇气号(见图 1-3)、机遇号以及我国的玉兔号都是基于双目摄像头实现实时路径规划;无人机可以实现

避障和智能识别等从而保障低空飞行的安全性;巡航导弹精确制导及自动巡航捕获目标和确定距离。

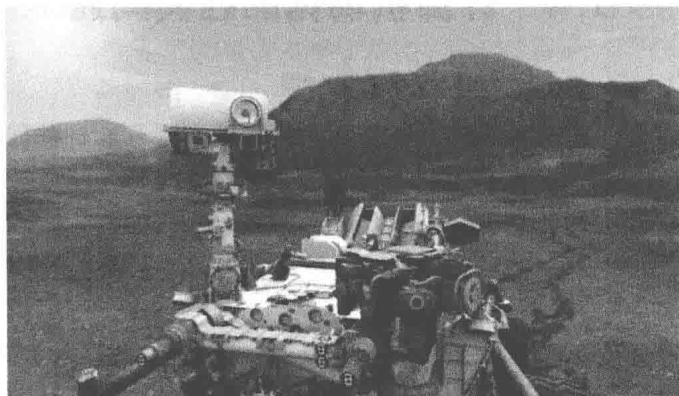


图 1-3 美国勇气号机器人探索火星

(3) 检测事件。视频智能监控、安全监控,在安防摄像头中融合视频分析、识别跟踪等技术,实现如看守所、监狱、博物馆、厂区及地铁站等特定场景的无人看守监控;在某些重要关口,如火车站,海关口岸实现人群计数;智能交通,能够实现车流量计数等。

(4) 组织信息。图像序列数据库或者视频数据的检索,如 Google 的图像搜索引擎能够实现基于内容的图像检索等。

(5) 物体或环境的建模。如生物医学图像分析、地形学建模、虚拟现实或增强现实、遥感测绘等。

(6) 人机交互。计算机或机器人等借助人的手势动作(手语)、嘴唇动作(唇读)、躯干动作(步态)、人脸表情测定等了解人的愿望要求而执行命令或进行互动(见图 1-4)。



图 1-4 人机交互游戏

## 1.4 计算机视觉的经典问题

这些计算机视觉技术精彩应用的背后都要解决一系列相同的问题。我们将这些相同的问题归纳成为 4 个计算机视觉的经典问题:识别、运动分析、场景重建和图像复原。

### 1. 识别问题

识别是计算机视觉最经典的问题之一,即让计算机“知道”某个物体“是什么”。其典型的应用有:图像搜索引擎、安保中的指纹或人脸识别、高速路摄像头自动检测超速车辆车牌号、书面文本的识别等。

识别问题是指判定一组图像数据中是否包含某个特定的物体,识别其图像特征或运动状态。我们希望通过计算机或机器自动解决以上问题,但是到目前为止,还没有某个单一的方法能够广泛地对各种情况进行判定,即在任意环境中识别任意物体。目前的技术只能很好地解决特定目标的识别,比如简单几何图形识别、人脸识别、印刷或手写文件识别或者车辆识别,而且这些识别需要在特定的环境中,比如满足特定的光照、背景和姿态要求。

### 2. 运动分析

运动分析通过对图像序列(视频)的处理,可以对其中运动物体的速度、方位等作出预估,即可实现运动物体的跟踪。

### 3. 场景重建

场景重建是指通过输入图像或视频对场景进行 3D 建模。

### 4. 图像复原

图像复原是指用最简单的办法,一般是使用低通滤波器或中值滤波器等去除图像的噪声(传感器噪声、运动模糊等),但复杂一点的基于 CV 技术的图像复原,便是同各国对图像中点、线、边缘等的分析检测得到的信息,能得到更好的噪声去除效果。

本书主要讨论了计算机视觉技术识别问题中的两个关键研究内容:目标模型建立和视觉注意特征模型建立。在识别问题中,尤其是在人脸识别、自动目标识别等问题中,识别效果的好坏主要取决于待识别目标特征模型的好坏。现在识别问

题发展的趋势是模拟生物的视觉系统来实现精确的识别和定位,而生物视觉系统的注意机制则可以实现大量背景信息的过滤,同时实现显著性目标的定位。

## 1.5 目标特征模型

本书以目标自动识别为背景讨论目标特征模型,但是目标特征模型可以应用于计算机视觉的各种相关应用中。

### 1.5.1 目标自动识别中面临的挑战

目标识别领域的算法研究面临三个主要的挑战,它们分别是目标描述(如何为目标建立模型)、学习(如何通过训练样本得到模型)和检测识别(如何在图像中检测出我们所感兴趣的目标或类,并且检测识别结果不受目标位置、视角和光线强弱的影响)。众多领域内的文献显示,这些问题还远远没有得到很好的解决。一个比较成功的自动目标识别算法应该具有以下几种情况的应对能力:① 视角的变化:同一个目标在不同的视角下可能有很大的不同。② 观察点的变化:观察点的不同可能导致目标的位置变化、尺度变化、旋转变化和仿射变化等。③ 光照的不同会导致图像中目标像素灰度值的变化。④ 背景复杂度的变化:一般现实中目标很难处于一个相对单一的背景环境,其背景环境可能是多变而复杂的。⑤ 目标遮挡:在图像中待识别的目标可能部分被遮挡,导致目标外形等发生变化。⑥ 类内目标差异性(类目标识别):属于同一个类别的目标也可能有很大的外形或者灰度的差异。

本书对目标的各种特征进行了研究,根据特征从底层特征到高层特征建立了特征模型。以特征模型的结构为基础,依次研究了各层次的目标特征,分别以这些特征为基础建立目标模型,将目标模型用于目标识别或者场景分析中,并且分析这些目标模型对亮度变化、视角变化、背景复杂度以及其他变化的适应性。

这里将目标识别问题分为单个特定物体的识别<sup>[1-4]</sup>和某一类具有共性物体的识别<sup>[5-15]</sup>,本书着重阐述了某类目标识别问题的研究成果。特定目标的识别相对来说比较容易,其识别算法相对来说已经获得了比较好的效果,不但效率高<sup>[3]</sup>,而且光照和视角变化具有不变性<sup>[3,4,16,17]</sup>。对类目标的识别来说相对困难些,因为类目标具有普遍性,所以需要更复杂的表达,因此其学习过程也相对复杂。以往的文献显示在类目标的学习中需要数以千计的训练样本,本书的研究集中在如何用尽量少的样本实现对类目标模型参数的学习。

本书的识别除了对常用目标的一些识别,还着重于阐述对遥感图像中类目标的识别问题。近年来的研究主要集中在某几种类目标的识别上,包括人脸<sup>[12,14]</sup>、

行人<sup>[18]</sup>、手写字<sup>[8]</sup>和车辆识别<sup>[7,12]</sup>。而文献中对遥感图像中类目标的识别相对比较少,已有的一些传统方法都是对图像进行分割后再引入目标的一些特征进行判断。而我们希望能够实现更加普遍的遥感图像类目标识别技术,引入计算机视觉中的高层信息,使其更加接近于人类识别的能力。

另一个关注的方面是,如何尽量减小监督的作用。也就是在模型学习的时候,对样本不需要太多的注释,因为注释大多都是靠手工标注的,这是需要工作量非常大的一件事情。因此,本书使用的统计模型和概率检测技术会在后面章节详细讲解。

目标识别又是一个多面问题,在不同的应用中可能有不同的理解。对于给定的一幅图片,我们感兴趣的可能是:图像中是否存在某类目标?这个目标在图像中的位置在哪里?又或者我们可能会问:图像存在多类中的哪类目标?目标识别可能有多层的含义,这里对几个名词进行定义。

(1) 二分类(classification):图像二分类的任务是,决定图像中是否存在某类目标,即存在或者不存在某类目标。

(2) 分类(categorization):图像分类的任务是,图像中存在多类目标中的哪种目标,属于多类分类问题。

(3) 定位(localization):定位的任务是,当图像中存在某类目标时,找到此目标在图像中的位置,即本文用矩形框标出目标所在位置。

(4) 检测(detection):与定位的含义相同。

(5) 目标识别(recognition):包含分类、检测和定位,目标识别可以是其中任何一个概念。

## 1.5.2 识别中特征模型的发展

最早的目标识别研究始于1960年,所以领域内有非常多的文献。过去50年间识别算法的进步在于所用图像越来越贴近现实,比如早期的目标识别仅指单一特定物体的识别,而某一类目标的识别随后才出现。在1970年才出现的距离数据(range data)给出了3D数据,而1980年的识别目标多放置于一个单一背景或者恒定背景中,使得目标的分割非常容易。随后再发展,目标识别技术可以慢慢开始处理自然图片,但是仍然局限于单一特定物体的识别,开始出现对于多视角的识别。1990年集中于对自然背景中目标的识别,并且可以实现类目标的识别,但是目标的视角被限定。

要了解目标识别领域的历史,就要从单一目标的识别开始,然后延续到类目标的识别。

### 1. 特定目标的识别

对于特定目标的识别,通常分析所识别目标的特征,采用目标的某一个或者几

个全局特征建立目标的特征模型。

20世纪70年代的识别技术仅用于对具体单一目标的识别,所以不存在类内差异性。目标的可变性由5种因素决定,分别为视角变化、亮度变化、目标遮挡、背景复杂度和图像噪声(通常比较小)。这段时间的目标识别算法采用的都是几何特征,主要提取的是目标的内轮廓或者外边缘。使用几何特征的优点是对光照的不变性以及对2D和3D形态表达的直观性。这段时期产生了各种各样的几何特征识别方法,主要分为匹配(alignment techniques)识别方法和几何不变量(geometric invariants)识别方法。

Lowe提出的匹配识别系统SCERPO<sup>[19]</sup>,用于提取图像中的直线,再根据共线与平行的原则分类直线,与模板进行匹配后找到识别目标;Huttenlocher和Ullman<sup>[20]</sup>根据已有的CAD模型,识别3D目标,实验表明他们的算法能够识别复杂背景中有遮挡的目标;Ullman和Basri<sup>[21]</sup>提出了一个3D和2D混合模型,这引起了领域内对于3D模型或者混合2D模型哪个效果更好的争论;人类视觉科学家在人脑视觉建模上也做了很多的工作,比如Biederman<sup>[22]</sup>、Riesenhuber和Poggio<sup>[23]</sup>;Mundy和Heller<sup>[24]</sup>将3D模型应用于遥感图像的目标识别中。文献<sup>[25]</sup>为匹配最优搜索提供了快速的算法。

几何不变量识别方法中,非常重要的一个贡献是散列法/哈希法(hashing)的引入,Lamdan和Wolfson<sup>[26]</sup>以及Rigoutsos和Hummel<sup>[27]</sup>都提到了这一算法。这种算法的目标模型由感兴趣点集组成,其中为了实现对仿射变换的适应性,以三个点组成的三角形为一个组成单元。在学习阶段,分析每个目标兴趣点组成的所有可能的三角形,将最后保留的三角形对应的兴趣点存放在哈希表中;在识别阶段,分析待识别目标的兴趣点并对应到哈希表,投票给映射的目标模型,得票最多者即为识别出来的目标。由于这种表示方法的冗余性,可以实现对被遮挡目标的识别,但是在噪声影响较大或者复杂背景的情况下,误识率(FPR)迅速上升,于是Rigoutsos和Hummel介绍了一种概率投票的方法来解决这个问题。Rothwell等<sup>[28,29]</sup>针对平面目标的识别提出了投影不变量,结合兴趣点、直线和曲线等特征得到各目标的不变量投影函数,不同目标的投影函数大不相同。

上述几何特征识别方法都有对仿射的不变性,比如几何哈希法不但具有对仿射的不变性,而且识别速度快。但是这种方法仅依靠目标的几何轮廓,对于轮廓内的纹理特征可以忽略不计,这种方法有利于对某类外形一致的目标的识别,比如酒瓶,但是对于人脸识别这种具体识别就完全不适用了。这些方法最大的问题在于,其假设目标外边缘已经准确被提取,但是实际上目标的边缘并不容易准确提取,尤其在噪声干扰以及复杂背景的情况下,所以这类方法的实用性受到了限制。

1970—1980年的技术突破在于3D信息的距离数据(range data)被广泛引



入<sup>[30-34]</sup>,取代灰度信息,保留了像素点的深度信息,这样便于提取出 3D 目标的形状。比如高斯曲线<sup>[34]</sup>的距离数据可以方便地从图像中提取边缘,然后建立目标模型与 3D 模型做匹配。

除了几何特征的识别方法以外,还有一些其他的方法值得一提。Schiele 和 Crowley<sup>[35]</sup>采用目标局部表面的统计直方图,并且结合局部描述子的空间位置;之后 Swain 和 Bollard<sup>[36]</sup>在此算法的基础上衍生出其他的算法;Murase 和 Nayar<sup>[37]</sup>采用 3D 特征空间的方法实现对 COIL 数据库<sup>[38]</sup>中的车辆识别,并且估计汽车的 3D 姿态。模板库中记录了目标的各种视角的姿态,然后将其映射到主成分空间中;Pontil 和 Verri<sup>[39]</sup>采用支持向量机(SVM)的方法识别 COIL 数据库中的目标,将目标对应到点集空间中,然后采用 SVM 寻找点集间的最佳划分超平面。但是这种算法对遮挡很敏感,并且对复杂背景中的目标识别效果也不佳。

以上方法都采用 COIL 数据库为实验对象,但是 COIL 数据库的最大缺点是其图像大部分都是人为制作出来的,比如玩具等都被置于黑色的背景中,这种目标的分割比自然场景中的目标分割要简单得多。因此,上述方法虽然适用于 COIL 数据库,但是却不适用于日常应用中。

上面所有的算法都是基于目标的整体特性的,比如整体目标轮廓、目标灰度等,这种算法的优点是实现简单,但是对背景的变化和遮挡都非常敏感。针对这些缺点,基于局部特征的算法开始发展起来,比如局部纹理特征的引入。最早由 Schmid 和 Mohr<sup>[40]</sup>提出了基于局部区域的识别方法,用于特定目标的识别。特征兴趣点的提取采用 Harris 角点提取方法,每个点取其局部区域的多尺度灰度特征,此算法在目标部分缺失的情况下依然能够识别,但是对背景变化的适应性不强。这种算法被应用于 COIL 数据库和遥感图像中。基于局部特性识别最著名的算法当数 Lowe<sup>[41]</sup>提出的算法了,其提出的 SIFT 描述子之后被广泛应用于各种目标的识别中,此方法具有对仿射的不变性,并且对遮挡目标也有很好的识别率。之后在此方法的基础上衍生出大量的改进识别算法<sup>[42-47]</sup>。

综上所述,基于局部区域的算法要优于基于几何等特征的算法,由于其在复杂背景和目标遮挡识别上的优势。但是对于某些变化造成目标的局部纹理变化的情况,几何特征的算法要远远优于局部纹理的算法。所以上述两种算法各有其适应的应用。

## 2. 类目标的识别

最早的某类目标的识别研究起始于 1980 年,与某一特定目标的识别相比,某类目标的识别更加困难,因为类内的物体并不完全相同而是彼此存在差异性,所以识别算法需要类内物体差异性具有不变性,即建立的目标模型具有不变性或者选择具有稳健性的匹配算法。截至目前,大部分文献中的类识别都集中在特定目标