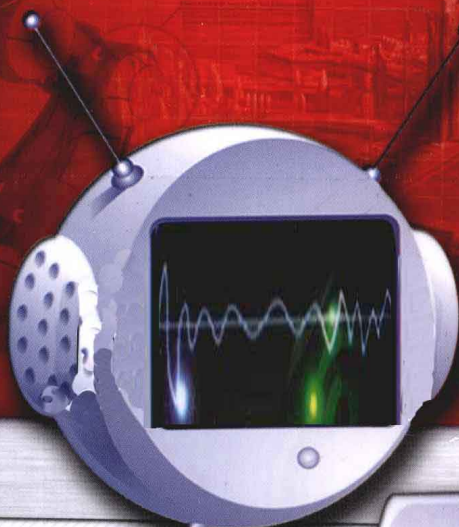


图像目标的

表示与识别

曹健 著

REPRESENTATION AND RECOGNITION OF THE IMAGE TARGET

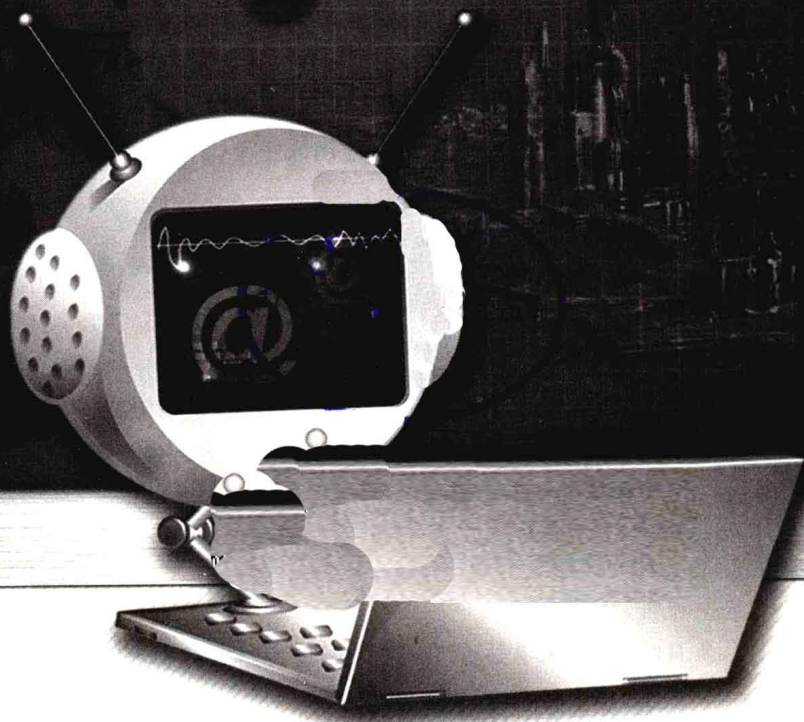


 机械工业出版社
CHINA MACHINE PRESS



图像目标的 表示与识别

曹健 著



 机械工业出版社
CHINA MACHINE PRESS

本书较为全面地介绍了图像目标识别的相关概念、原理和技术方法。主要内容包括图像目标的特征提取、图像目标的表示与描述、图像目标匹配和图像目标分类等。本书紧跟上述内容的国内外发展现状和最新成果，阐述了作者对图像目标识别的理解和认识，尤其针对局部特征在图像目标识别中的应用，进行了深入的探讨、分析和实例验证。

本书可以作为从事图像理解、模式识别、机器视觉等相关专业研究人员的参考书，对于计算机科学与技术、信息与通信工程、电子科学与技术等专业的研究生和高年级本科生也有一定的参考价值。

图书在版编目 (CIP) 数据

图像目标的表示与识别/曹健著. —北京: 机械工业出版社, 2012. 7
ISBN 978-7-111-38182-2

I. ①图… II. ①曹… III. ①图像处理 IV. ①TN911. 73

中国版本图书馆 CIP 数据核字 (2012) 第 081984 号

机械工业出版社 (北京市百万庄大街 22 号 邮政编码 100037)

策划编辑: 吕 潇 责任编辑: 吕 潇

版式设计: 霍永明 责任校对: 薛 娜

封面设计: 赵颖喆 责任印制: 杨 曦

北京中兴印刷有限公司印刷

2012 年 6 月第 1 版第 1 次印刷

169mm × 239mm · 13.25 印张 · 235 千字

0 001—2 500 册

标准书号: ISBN 978-7-111-38182-2

定价: 39.80 元

凡购本书, 如有缺页、倒页、脱页, 由本社发行部调换

电话服务

网络服务

社服务中心: (010)88361066

门户网: <http://www.cmpbook.com>

销售一部: (010)68326294

销售二部: (010)88379649

教材网: <http://www.cmpedu.com>

读者购书热线: (010)88379203

封面无防伪标均为盗版

前 言

图像目标的表示与识别作为图像处理与模式识别领域的一个重要的研究方向，在安全监控、军事侦察、产品检验、人机交互、医疗诊断等方面得到了越来越广泛的应用。但目前尚未形成一个成熟统一的技术方案，往往需要针对特定的任务，甚至针对特定的图像，选用一种或几种不同的方法。而相关领域的数学算法和具体技术林林总总各不相同，甚至从思路上就已经大相径庭，这更需要我们下工夫进行梳理和提炼。

本书围绕着图像目标的表示与识别这一主题，全面系统地介绍了相关的概念、原理和技术方法。针对可见光图像和刚性目标，学习并借鉴了图像工程、模式识别、机器视觉和人工智能学科中一些先进技术，探讨了复杂背景下的目标识别以及局部遮挡物体的识别中的关键问题，为增强现有图像识别系统的自动化程度和信息处理能力提供理论支持和技术帮助。

本书分为7章，内容安排如下：

第1章概述了图像目标识别的基础理论和研究思路，介绍了图像目标识别常用的图像库，指出了图像目标识别的主要难点和发展趋势；第2章讨论了图像分割和目标分割的关系，介绍了提取目标整体特征的相关技术；第3章介绍了目标匹配和目标分类的基本理论，详细论述了常用的图像目标分类器的设计和训练方法；第4章回顾了局部特征的研究现状，给出了几种典型区域检测算子和特征描述子的具体算法和改进方法；第5章针对局部特征匹配在目标图像拼接和图像检索中应用的不足，提出了基于多分辨率技术的航拍图像拼接方法，以及基于原型匹配的图像检索方法；第6章阐述了视觉单词的理论依据以及视觉单词库特征库构造方法，结合信息论的相关技术进行特征选择，提出了一种基于局部特征的目标分类方法；第7章结合主分量法和 Hausdorff 距离，提出了一种在视点变化下目标匹配识别方法和基于角点标记图的 BP 网络分类方法。

本书的研究成果首先要感谢北京理工大学计算机学院的刘玉树教授等多位老师给予作者的长期指导和教诲，还要感谢众多师兄弟和一些硕士研究生在作者攻读博士学位期间给予的启发与激励，更要感谢北京工商大学计算机与信息工程学院的领导和同事们不遗余力的关怀和帮助，尤其感谢国家自然科学基金

和北京市自然科学基金项目（编号：4123095，4112016）对本书相关课题研究的支持。本书的出版也得到了北京工商大学青年教师科研启动基金资助项目（编号：QNJJ2011-38）以及实验室课题组项目（科技创新平台，编号：19005118053）经费的支持，在此一并致谢。

由于图像处理与模式识别领域的相关技术仍处于不断发展和完善阶段，加之作者水平有限，书中难免存在一些不足之处，敬请读者批评指正。

曹 健

2012年5月

目 录

前言

第1章 绪论	1
1.1 引言	1
1.2 图像目标识别的意义	2
1.3 图像目标识别的框架与思路	5
1.3.1 图像目标识别问题的分类	5
1.3.2 图像目标识别的基本框架	7
1.3.3 图像目标识别的两种思路	8
1.4 图像目标识别的数据集	10
1.5 图像目标识别的开发环境	15
1.6 主要难点与发展趋势	18
1.7 研究内容与结构安排	21
1.7.1 本书的研究内容	21
1.7.2 本书的结构安排	22
第2章 图像目标的整体特征提取	25
2.1 引言	25
2.2 图像目标分割	29
2.2.1 图像目标分割概述	29
2.2.2 图像目标分割现状	30
2.2.3 图像目标分割技术	33
2.3 目标的表示与描述	42
2.3.1 光谱特征	42
2.3.2 纹理特征	44
2.3.3 形状特征	46
2.4 特征空间的优化	48
2.4.1 特征选择	48
2.4.2 特征变换	50
2.5 本章小结	52

第3章 基于整体特征的目标识别	55
3.1 引言	55
3.2 模式识别方法概述	56
3.3 目标匹配的研究现状	58
3.3.1 两种目标匹配方式	58
3.3.2 匹配的相似度度量	59
3.4 目标分类的研究现状	61
3.4.1 分类器设计技术	62
3.4.2 性能评估方法	64
3.5 典型的图像目标分类器	66
3.5.1 基于聚类分析的分类器	66
3.5.2 基于朴素贝叶斯的分类器	69
3.5.3 基于BP神经网络的分类器	71
3.5.4 基于支持向量机的分类器	73
3.6 本章小结	76
第4章 图像目标的局部特征提取	77
4.1 引言	77
4.2 特征区域的稀疏选取算法	78
4.2.1 特征区域检测的研究现状	78
4.2.2 高斯差分检测算子	80
4.2.3 边缘点检测算子	83
4.3 局部特征的定量描述	85
4.3.1 特征区域描述的研究现状	85
4.3.2 基于梯度分布的描述子	87
4.3.3 线矩特征描述子	89
4.4 角点的检测算法	90
4.4.1 直线投影检测算法	91
4.4.2 SUSAN算法的自适应阈值改进	92
4.5 实验结果与分析	94
4.6 本章小结	97
第5章 基于局部特征的目标匹配	99
5.1 引言	99
5.2 结合 NNDR 与霍夫变换的匹配方法	100
5.2.1 基于 NNDR 的匹配策略	100
5.2.2 邻近特征点的搜索算法	101

5.2.3	基于霍夫变换的目标检测	103
5.3	基于局部特征和多分辨率技术的图像拼接	105
5.3.1	图像拼接技术的研究现状	105
5.3.2	多分辨率下的图像配准	107
5.3.3	渐入渐出的图像融合算法	112
5.4	基于局部特征和原型匹配的图像检索	114
5.4.1	CBIR 的研究现状和发展趋势	114
5.4.2	基于模板匹配的检索方法	117
5.4.3	基于原型匹配的反馈技术	118
5.5	实验结果与分析	119
5.6	本章小结	124
第 6 章	基于局部特征的目标分类	127
6.1	引言	127
6.2	目标的向量空间模型表示	129
6.3	构造视觉单词库	130
6.3.1	视觉单词的生成方法	131
6.3.2	基于 RNN 的层次聚类算法	132
6.4	基于信息论的特征选择方法	134
6.4.1	信息论的相关概念	135
6.4.2	基于信息增益法的特征选择	136
6.4.3	基于 CHI 统计量的特征选择	137
6.4.4	基于互信息法的特征选择	138
6.5	视觉单词的权重计算	139
6.6	实验结果与分析	141
6.7	本章小结	146
第 7 章	基于角点特征与视面模型的目标识别	147
7.1	引言	147
7.2	三维物体的视面模型表示	150
7.3	基于角点特征的目标匹配	152
7.3.1	利用基准角点进行目标匹配	152
7.3.2	基于主分量与 Hausdorff 距离的匹配算法	154
7.4	基于角点标记图的目标分类	157
7.4.1	角点特征的优化技术	157
7.4.2	角点标记图的生成方法	159
7.5	实验结果与分析	160

7.6 本章小结	164
附录 A 图像处理的一些相关理论	167
A.1 数字图像的基本概念	167
A.2 数字图像的信息内容	168
A.3 图像处理的技术门类	169
附录 B 模式组合的一些基本概念	173
B.1 图	173
B.2 树	173
B.3 符号串	174
附录 C 概率统计的一些预备知识	177
C.1 概率	177
C.2 最大似然估计	177
C.3 条件概率	177
C.4 贝叶斯公式	178
C.5 随机变量	179
C.6 二项式分布	179
C.7 联合概率分布和条件概率分布	179
C.8 贝叶斯决策理论	180
C.9 期望和方差	181
附录 D 信息检索的一些基础模型	183
D.1 布尔模型	183
D.2 向量空间模型	183
D.3 概率模型	184
D.4 语言模型	185
附录 E 名词术语解释	187
参考文献	192

第 1 章 绪 论

我们只能向前看到很短的距离，但是我们能够看到仍然有很多事情要做。

——阿兰·麦席森·图灵（1912—1954）

1.1 引言

视觉是人类获取信息、感知世界，进而改造世界的一个重要途径。有资料显示，人类接受到的外界信息中约有 60% 以上来自于视觉，而听觉、味觉、触觉、嗅觉总共占不到 40%。但是从技术发展来看，图像信息的处理远远滞后于语音信息，随着计算能力的不断提高，如何使计算机具有和生物类似的视觉感知功能成为目前计算机领域中的一个研究热点。

图像目标的表示与识别，又称关于视觉图像的模式识别，旨在利用图像处理与模式识别等领域的理论和方法，确定图像中是否存在感兴趣的目标，如果存在则为目标赋予合理的解释，必要时还要确定其位置^[1]。虽然国内外科研工作者就如何在复杂环境下检测、辨识和准确跟踪目标进行了理论分析和实践探索，但目前尚未形成一个成熟统一的技术方案，往往需要针对特定的任务，甚至针对特定的图像，选用一种或几种不同的方法。而相关领域的数学算法和具

体技术林林总总各不相同，甚至从思路上已经大相径庭，这更需要我们下工夫进行梳理和提炼。

在这里，识别 (Recognition)、分类 (Classification)、检测 (Detection)、定位 (Location) 和鉴别 (Identification) 几个概念需要简要说明一下。从上面的定义可以看出，识别的内涵最为宽泛，分类、检测、定位、鉴别都能看做是识别的子任务之一；分类的定义比较清晰，即对图像目标按照类别标签进行划分；检测和定位的目的是相似的，一般是确定图像中某个目标的具体位置；鉴别往往指同类目标间的区分，如对人物张三和人物李四进行辨认。虽然这几个概念在不同的文献中稍有差异，本书中对它们的解释也并不唯一，然而把握好它们在具体问题中的界定，还是有助于加深对图像识别领域中实际问题的理解。

1.2 图像目标识别的意义

近年来，许多重要的国际期刊 (IEEE Transactions on Pattern Analysis and Machine Intelligence、IEEE Transactions on Image Processing、IEEE Transactions on Medical Imaging、IEEE Transactions on Vehicular Technology、International Journal of Computer Vision、Computer Vision and Image Understanding、Image and Vision Computing、Pattern Recognition、Pattern Recognition Letters、Machine Vision and Application 等) 以及重要的国内期刊 (计算机学报、软件学报、自动化学报、机器人、模式识别与人工智能、计算机研究与发展等) 都发表了大量关于图像模式识别方面的论文。在国外召开的顶级国际会议，如 IEEE 国际计算机视觉与模式识别 (Computer Vision and Pattern Recognition, CVPR) 会议、欧洲计算机视觉会议 (European Conference on Computer Vision, ECCV)、国际信息处理会议 (International Conference on Information Processing, ICIP) 等，也收录了许多知名学者在相关领域的学术成果。这几年，国内学术界积极开展了一系列的学术交流活动，比如 2005 年在北京举办的国际计算机可视化会议 (International Conference on Computer Vision, ICCV)、2006 年在香港特别行政区举办的第 18 届模式识别会议 (International Conference on Pattern Recognition, ICPR)、2008 年全国模式识别学术会议、2009 年在西安举办的第 9 届亚洲计算机可视化会议 (Asian Conference on Computer Vision, ACCV) 等。

图像目标的表示与识别之所以备受关注，是由于它能够广泛应用于国防和民用的许多领域，其中包括安全监控、军事侦察、产品检验、人机交互和医学

应用等多个方面。

1. 安全监控

图像目标识别在安全领域的应用范围非常广泛，大城市很多地方，如民宅、停车场、银行等，都装有闭路电视监控系统（Close Circuit TV），以便能够对可疑的物品和人员进行有效的监控。而随着各种新的 DNA 分析、分型技术方法的建立，借助多模态的生物特征辨识系统，法医 DNA 分析技术可将从犯罪现场提取的 DNA 轮廓（手掌的纹理、指纹、脸的几何形状）与疑犯的 DNA 信息进行更准确、快速、自动的匹配。2004 年，根据市场研究公司——国际生物测定组织的分析，在人们首选的在线银行认证方法中，选择生物特征辨识的占了 50%，是智能卡、密码、身份证号码等方式的总和。

在交通系统中除了视频摄像外，还需要大量的识别监视跟踪系统。例如，目前的车牌识别技术已经非常成熟，这对道路上异常车辆的监控和交通事故的事后处理都具有非常重要的意义。西门子公司交通监控性能非常优越，不仅能探测隧道中慢行或停止的汽车，还可探测处于 U 形转弯处的违规汽车，以及自动检测可疑的行李。智能车辆的最终目的是实现车辆的自动驾驶，目前主要是利用车上安装的摄像机、雷达等传感器设备进行道路检测并识别前方的障碍物（如车辆、行人），以保证车辆的安全行驶。

2. 军事侦察

相对而言，军事领域的识别与监测要求就非常苛刻了，主要是因为战场环境要比一般的民用环境更为复杂。例如，检测有遮挡和伪装的机动目标就十分困难，由于假设的局限性，在民用上已经比较成熟的算法在军事上往往效果很不理想。美国洛克希德·马丁公司开发的数字式侦察图像处理系统已安装到尼米兹级航空母舰上，成为美国海军联合部队图像处理系统（JSIPS-N）的战术组成部分，它能接受和处理来自多个传感器平台（U-2、“全球鹰”无人机、F/A-18 共享侦查吊舱等）的图像，极大增强了美国海军识别和打击关键目标的能力。2006 年 6 月以色列 IAI 公司在巴黎展示了其一元化的战争指挥室，其中实时图像情报中心（EL/S-8894RT-RiCENT）具有对战场全天候一体化的监视和侦察能力。

3. 产品检验

由于工业环境的结构、照明等因素可以得到严格的控制，图像目标识别在工业生产和装配中得到了成功的应用。一个具有简单视觉感知功能的自动化生产线包含一个摄像机和相关的信息处理系统，通过摄像机对零件进行识别和定位，为机器人提供是否操作或进行何种操作的信息，并引导机器人手臂实时准

确地夹取零件；此外，图像识别技术已经应用在集成电路设计、图形设计和电视电影制作中；通过多源图像融合，可以进行产品外形检验、表面缺陷检验，加强对产品质量的严格把关。

对多个摄像机的图像同步识别处理，利用某一时刻关于某个目标的不同角度的图像可以恢复场景的三维信息，并依据三维信息做出决断，实现即时规划、自主导航、与周围环境实时交互作用等。这是生产控制的进一步发展，让机器人不仅仅停留在简单的自动化生产线上，而且能够代替人类进入危险的环境进行生产活动，例如，在核辐射区或火灾现场抢修设备，远程控制的无人开采矿藏、星际探测设备的自主导航等。

4. 人机交互

对包含文字和符号的图像进行识别可以让人与计算机的交互更加便捷。目前这方面的技术大量应用于信函分拣、稿件输入、支票查对、期刊阅读和自动排版中，而超市的条码阅读器更是对销售管理的一场革命。现在美国和日本的客户已经能够通过把他们的手机指向汉堡包的包装纸，获得其营养信息并显示在屏幕上，也可以通过这种方式获得商品报价。例如，在日本东京的一座建筑物上粘贴的超高速识读条码（Quick Response Code, QRcode）就含有很多信息，通过带有摄像头的可正确编译的手机就能方便地读取。

面部表情传达了一种非口头性的暗示，对其进行自动识别是人机接口的重要元素，也被用于行为科学和临床实践中。比如，具有微笑探测和眨眼探测的两个功能独立的数字照相机可以在恰当的时机（用户微笑的时候或眨眼之后）捕捉到主体，并提示用户，进行抓拍。手势识别也称手语识别，是机器视觉领域中比较前沿的研究领域。当用户做出一个手势，摄像机（一般为双目或三目）将图像传送到计算机，然后由特定软件结合视差来提取手臂、手指等三维特征，完成这些特征的进一步识别，最后对这个手势做出响应^[2]。

5. 医学应用

如今，计算机图像分析逐步融入到了医疗诊断的过程中，这就促生了计算机辅助诊断（Computer-Assisted Diagnosis, CAD）技术。利用该技术，可进行核磁共振成像（主要用于医疗成像来可视化人体结构和功能，提供任何平面内身体的细节图像），癌细胞、白细胞、染色体检查，修复手术控制设计等。

通过一组切片图像进行人体器官的三维重构，可以为医疗诊断和病理分析提供重要和直观的帮助。同样，可以根据图像序列中的信息对普通目标进行三维重构，无论观察点在何处，都能利用其三维信息进行识别，这也为解决视点

变化下的目标识别提供了一个思路。

除了以上几个方面，图像目标识别在生产生活中还有很多应用。对目标描述信息的分析处理，可以用在天气预报、森林火灾及地质灾害监测、空气污染预报等领域。人脸检测（Facial Detection）技术可以将画面及时地锁定在讲话人身上，这样就很大程度地降低了远程电视会议的图像传输比率^[3]。在虚拟现实、计算机动画、视频评注等应用领域，目标识别技术同样也发挥着不可替代的作用。所以，开展图像目标识别研究意义重大，其研究成果具有非常广阔的应用前景。

1.3 图像目标识别的框架与思路

1.3.1 图像目标识别问题的分类

针对图像领域中的各种具体问题，目标识别所采用的研究方法和技术方案都有所区别，甚至迥然不同。所以需要将目标识别问题按照一定的标准进行分类，对具体问题进行分析。

1. 按照获取图像的传感器的种类

按照获取图像的传感器的种类，可以将图像目标识别分为可见光图像目标识别，红外图像目标识别和合成孔径雷达（Synthetic Aperture Radar, SAR）图像目标识别。这三种传感器的成像原理不同，对拍摄时间、天气情况、地理环境、光照的要求也不一样。通常条件下，可见光图像比较清晰、直观、费用低，有利于实时传输，但可见光传感器只敏感于目标场景的可见光反射，容易受到各种场地因素的干扰；红外图像^[4]适合夜间使用，具有特殊的识别伪装的能力，但图像清晰度低，且大气红外辐射和吸收作用对图像质量影响很大；合成孔径雷达图像^[5,6]易于判读线性地物、表面光滑的面状地物、森林、草地、水体等，具有很强的穿透力，但雷达视向对目标的表达色调和形状影响很大。目前，国外先进的无人侦察平台都采用多种传感器成像技术，并通过图像融合得到了信息更为丰富的图像。

2. 按照图像背景的复杂程度

按照图像背景的复杂程度，可以分为简单背景下的目标识别和复杂背景下的目标识别。简单背景下的目标识别，如文字识别、符号识别和人脸识别等，目标和背景的对比度非常大，一般的图像处理和分割算法就能准确完整地提取出目标。此类研究侧重于如何辨识出更加细微的区别，或者对目标的不同姿态进行识别。而在复杂背景下进行图像目标识别受到噪声的影响非常大，目标的

检测效果往往差强人意，要想提取出完整的目标更是困难，一般需要在先验知识的指导下进行目标的检测和图像的分割。

3. 按照相关图像的性质

按照相关图像的性质，可以分为静态图像识别和动态图像识别。静态图像，也称静止图像，指的是关于目标的单幅图像，我们一般的图像检索和图像分类大都是针对这类图像的。而动态图像为我们提供了比静态图像更为丰富的信息，通过对多帧动态图像（图像序列）的分析，可以检测出目标的运动信息，识别与跟踪运动目标和估计三维运动及结构参数。动态图像识别面临的首要挑战是，如何从图像序列中实现有效的图像分割和图像对应。图像分割在静态图像识别领域也尚未得到有效解决，图像对应问题则是与模式识别和人工智能紧密相连的难题。

4. 按照图像中目标的数目

按照图像中目标的数目，可以分为单目标识别和多目标识别。单目标的图像，顾名思义，就是只有一个感兴趣的目標，其余属于背景，这就相当于提供了一个重要的前提。在这个前提下，我们更多关注的是如何利用各种图像处理技术抑制背景，完整准确地检测和提取出这一个目标。而多目标识别要比单目标识别困难得多，因为多个目标同时出现在一幅图像中，不光有复杂背景的干扰，还必须考虑到目标之间会相互遮挡（Occlusion）、合并（Merge）、分离（Split）等种种情况。这更需要通过知识来指导信息的选择和整合，并进行反复的假设验证（Hypothesis Verification）和复杂的反馈处理。

5. 按照图像中目标的类型

按照图像中目标的类型，可以分为刚性目标识别和非刚性目标识别。刚性（Rigid）目标一般指具有刚性结构、不易变形的物体，如飞机、车辆、建筑物等人工物体，它们的共同特点是结构比较规范，适合用几何模型进行描述，一般采用基于形状特征的方法进行识别。而非刚性（Non-rigid）是指外形能够变化的物体，如细胞、动物、人体等。对这类目标可以采用光谱特征、纹理特征以及变形模板（Deformable Template）技术等进行识别。

6. 按照对图像语义的理解程度

按照对图像语义的理解程度，可以分为图像分类、目标检测以及目标识别。图像分类只是根据低层图像特征和相似度度量，将内容类似的图像归为一类，并不需要对图像中的对象进行分割和定位，如基于内容的图像检索；目标检测不仅要确定图像中是否存在感兴趣的目標，还要在必要时确定其位置，以便于

进一步提取目标进行处理，如车牌提取、人脸检测等；具体目标的识别需要对图像信息进行深入分析，例如在视觉跟踪中，不仅要检测出感兴趣的目標，还要与周围其他目标进行区分，避免产生混淆。

1.3.2 图像目标识别的基本框架

一个典型的图像目标识别系统如图 1-1 所示，主要由图像增强与变换（图像预处理）、图像分割、图像描述、分类决策四部分构成。这四个部分关系非常密切，在看做一个有机整体的同时，也可以看成三个层次的计算处理——低层、中层和高层处理。

低层是对图像数据进行预处理，如对有噪声的图像要进行滤波去噪，对信息微弱的图像要进行对比度增强，对失真图像要进行几何校正等，以达到改善图像质量、突出兴趣区域的目的。其鲜明的特点是输入和输出的都是图像。

中层处理涉及分割（把图像分为不同区域或目标物），将给定图像或已分割的图像区域用更为简单明确的数值、符号或图来表征（特征描述），以使其更适合计算机处理及对不同目标的分类（识别）^[7]。中层处理输入为图像，但输出的是从这些图像中提取的特征。

高层处理一般是基于知识进行推理和证实的，涉及图像或图像区域的理解，以及执行与视觉相关的识别函数^[8]。也可以简单地认为是对图像或图像区域进行分类和估计。其输入是向量、串或树等形式的特征描述，输出则是图像或物体的类别。

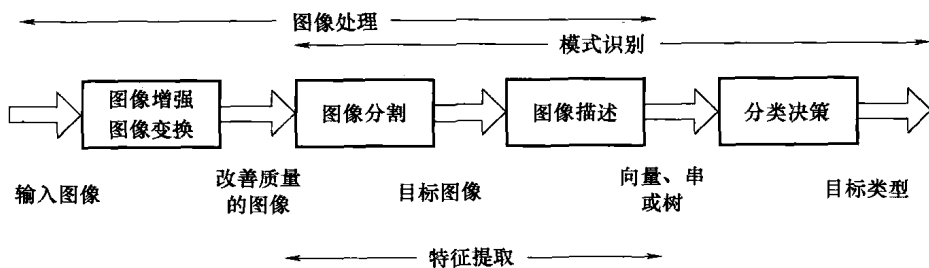


图 1-1 图像目标识别系统的基本框架图

图像目标识别技术是延伸和扩展人的视觉功能的方式和方法，其实信息技术整体都可以认为是扩展人的信息器官功能的技术。人的信息器官主要包括感觉器官、传导器官、思维器官和效应器官四大类型，其功能主要是信息获取、信息传输、信息处理和信息应用，因此感测技术、通信技术、智能技术与控制

技术被认为是信息技术的四基元，其他信息技术通常被看做是这四种基本技术的高阶逻辑综合或分解衍生^[9]。

如表 1-1 所示，我们把图像识别看作图像处理 and 模式识别的交叉，而这两门学科分别属于信息处理和计算智能两个大的学科门类，甚至还涉及信息传输的一些内容，从这个意义上，也看出设计和执行算法来模仿人类对物体的视觉识别能力是一项有趣而富有挑战性的任务，因此，这门学科不断吸引了许多来自不同领域的科研人员钻研和探讨，也不断涌现新的理论和方法。

表 1-1 图像识别在信息学科中的位置

信息器官 (人)	器官的作用	相应技术	学科门类	研究方向 (举例)
感觉器官	信息获取	感测技术	信息处理	图像处理 信号分析
传导器官	信息传输	通信技术	信息传输	信息编码 信息安全
思维器官	信息加工	智能技术	计算智能	人工智能 模式识别
效应器官	信息应用	控制技术	自动控制	集中控制 分散控制 现场控制

1.3.3 图像目标识别的两种思路

人类认知过程可以用图 1-2 描述^[10]。不同视觉基本特征，如方位、方向、空间频率、眼优势、空间拓扑和颜色等在不同层次视觉皮层具有一定的空间组织形式，多种基本特征功能柱共存于一片皮层空间，实现多种特征表达的最优化；特异性反应细胞在高级与初级视觉皮层上进行自下而上的前馈和自上而下的反馈，完成视觉表征自下而上地逐级抽象，以及在整合后自上而下地反馈、对初级水平的调控；大脑自动建立基于皮层自组织的计算视觉模型^[11,12]。

对于图像目标识别问题的研究，也是遵循着人的认知形式，总体上讲有两种思路，一种是自下而上的加工 (Bottom-up Process)，另一种是自上而下的加工 (Top-down Process)^[13]。这两类处理方法有着各自的优点和缺点，将它们结合起来各取所长，就有可能实现更为理想的识别。

1. 自下而上的加工

也被称为数据驱动 (Data-driven) 的加工，其核心观点是系统工作是单向