

**B** 大数据丛书  
BIG DATA SERIES

 Springer

# Pattern Recognition and Classification: An Introduction

## 模式识别与分类导论

---

【美】 杰夫·多尔蒂 著  
GEOFF DOUGHERTY

殷娟娟 答凯艳 程国建 译

 机械工业出版社  
CHINA MACHINE PRESS

大数据丛书

# 模式识别与分类导论

[美] 杰夫·多尔蒂 (Geoff Dougherty) 著  
殷娟娟 答凯艳 程国建 译



机械工业出版社

模式识别与分类的使用是当今许多自动化电子系统的基础。尽管该领域已出版了许多名著，但该主题仍然非常具有挑战性，特别是对于初学者而言。

本书全面介绍了自动模式识别领域中所涉及的核心概念。本书的构思适用于具有不同背景的新手，对图像、信号处理分析以及计算机视觉方面的研究人员和专业人员也非常有用。对于监督分类与非监督分类的基本概念，本书以非公理化、非形式化的方式呈现出来，便于读者快速地获得将基本概念应用于解决实际问题的必要背景知识。更高级的主题，如半监督分类、聚类算法与相关反馈的结合，在本书的后半部分中进行讨论。

本书适合作为高等院校本科生和研究生“模式识别与机器学习”课程的教学参考书或自学指导书。

Translation from English language edition;

Pattern Recognition and Classification: An Introduction

by Geoff Dougherty

Copyright © 2013 Springer New York

Springer New York is a part of Springer Science + Business Media

All Rights Reserved

本书中文简体字版由 Springer 授权机械工业出版社独家出版，未经出版者书面允许，本书的任何部分不得以任何方式复制或抄袭。

版权所有，翻印必究。

北京市版权局著作权合同登记 图字：01-2015-0414 号。

## 图书在版编目 (CIP) 数据

模式识别与分类导论/(美)杰夫·多尔蒂 (Geoff Dougherty) 著; 殷娟娟, 答凯艳, 程国建译. —北京: 机械工业出版社, 2019. 2

(大数据丛书)

书名原文: Pattern Recognition and Classification: An Introduction

ISBN 978-7-111-62072-3

I. ①模… II. ①杰… ②殷… ③答… ④程… III. ①模式识别—研究  
IV. ①TP391. 4

中国版本图书馆 CIP 数据核字 (2019) 第 033409 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: 王 康 责任编辑: 王 康 任正一

责任校对: 李 杉 封面设计: 陈 沛

责任印制: 张 博

北京铭成印刷有限公司印刷

2019 年 4 月第 1 版第 1 次印刷

169mm × 239mm · 10.5 印张 · 192 千字

标准书号: ISBN 978-7-111-62072-3

定价: 45.90 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

服务咨询热线: 010-88379833

机工官网: [www.cmpbook.com](http://www.cmpbook.com)

读者购书热线: 010-68326294

机工官博: [weibo.com/cmp1952](http://weibo.com/cmp1952)

教育服务网: [www.cmpedu.com](http://www.cmpedu.com)

封面无防伪标均为盗版

金 书 网: [www.golden-book.com](http://www.golden-book.com)

# 致 谢

---

我想感谢我的同事 Matthew Wiers，我们进行了许多次深入的对话，他帮助处理了与本书捆绑在一起的多个 Excel 文件。感谢我以前的所有学生对课程的反馈，并最终催生了这本书。特别要感谢的是 Brandon Ausmus，Elisabeth Perkins，Michelle Moeller，Charles Walden，Shawn Richardson，Ray Alfano。

我感谢 Springer 的 Chris Coughlin 在写作本书的过程中给予的支持和鼓励，并感谢批评稿件的各位匿名评论家，并在课堂上试用这些评论。特别感谢我的妻子 Hajjah 和家人 Danel，Adeline 和 Nadia 的耐心和支持，还有我的父母 Mand 和 Harry（2009 年去世），没有他们，这本书将难以面世。

# 译 者 序

人们在观察外界事物或现象的时候，常常要寻找它与其他事物或现象的不同之处，并根据特定的目标把各个相似的但又不完全相同的事物或现象组成一类，正所谓“物以类聚”，这就是模式识别与分类。人工智能中的模式识别就是通过计算机用算法来研究模式或样本数据的自动处理和分析判断，进而对表征事物或现象的信息进行描述、辨认、分类和解释的过程。模式识别技术是人工智能与信息科学的基础技术。21世纪是信息化、计算化、网络化和智能化的世纪，在这个以数字计算为特征的世界里，作为人工智能与信息科学与技术基础学科的模式识别技术，必将获得巨大的发展空间。

模式识别及其分类与统计学、心理学、语言学、计算机科学、生物学、控制论等都有密切的关系，它与人工智能、图像处理、机器视觉的研究又有交叉关系。在国际上，各大权威研究机构及各大前沿技术公司（如谷歌、IBM、微软、百度、腾讯、阿里巴巴等）纷纷开始将模式识别技术作为其战略研发重点。业界流行的大数据分析、机器视觉、深度学习等技术模式识别及分类有着密不可分的关系。

本书是模式识别与分类的基础教程，共包含10章内容：引言、分类、非度量方法、统计模式识别、监督学习、非参数学习、特征提取及选择、非监督式学习、评估和比较分类器、项目等。

本书既适合个人自学，也适合作为物理学、计算机科学与技术、自动化、机器人工程、生物医学工程、智能科学与技术，以及数学与应用数学等专业高年级本科生和研究生的模式识别和机器学习课程的教材。

本书的翻译出版得到西安培华学院学术基金的支持，在此表示感谢。

程国建

# 前 言

模式识别和分类的使用是当今许多自动化电子系统的基础。其应用范围遍及军事防御、医疗诊断、生物识别、机器学习、生物信息、家庭娱乐，等等。然而，尽管在这个领域有一些著名的书籍，但这个题目仍然是非常具有挑战性的，特别是对初学者来说。

我们发现目前的教科书无论是对以计算机科学为主的学生，还是对数学和物理学的学生以及其他工科专业的学生来说，都不能令人满意。他们的数学和计算机背景已相当多样，他们都希望以最少的时间投入理解和吸收核心概念，以便能够解决他们所从事领域的问题。而那些具有广泛的数学或统计学先决条件的文本令人望而生畏，对他们没有吸引力。

我们的学生抱怨“只见森林不见树木”，这对于模式识别教材来说颇为讽刺。对于这个领域的新人来说，必须以有序的合乎逻辑的方式介绍基本的概念，使他们能够欣赏“大局”。以先前的知识为基础，可以逐渐处理更多的细节，而不会被压倒。我们的学生经常尝试从各种教科书中学习不同的方法，但最终却被不同的术语弄糊涂了。

我们注意到，大多数学生对视觉学习非常感兴趣，由于他们的入学知识往往有限，需侧重于通过实例和练习来学习关键概念。我们相信，更多的视觉呈现和工作实例可以更好地帮助学生理解相关知识，提高他们的洞察力。

本书最初是作为加州州立大学海峡群岛（CSUCI）分校的高级本科课程和模式识别研究生课程的笔记和讲义开始的。随着时间的推移，它逐渐成长并接近目前的形式，这已经在 CSUCI 上经过了几年的等级测试。它适用于高年级本科生或研究生。它只要求学生在统计学和数学方面有适度的背景，由于在文本中加入了必要的附加材料，本书基本上是独立的。

本书既适用于物理学、计算机科学与技术、自动化、机器人工程、生物医学工程、智能科学与技术、数学与应用数学等专业的本科生和研究生的模式识别和机器学习课程，也适合于个人学习。它全面介绍了必须理解的核心概念，以便为该领域做出贡献。虽然它旨在为来自不同背景的新手提供便利，但对于图像和信号处理和分析以及计算机视觉领域的研究人员和专业人员也是有用的。我们的目标是以非正式而不是公理化的方式呈现监督和无监督分类的基本概念，以便读者能够快速将概念应用于实际问题。最后一章指

出了一些有用和可以进行的项目。

我们在图像探索和分析的早期阶段使用 ImageJ (<http://rsbweb.nih.gov/ij/>) 和相关的发行版, Fiji (<http://fiji.sc/wiki/index.php/Fiji>) 也具有直观的界面和易用性。然后, 我们倾向于转向 MATLAB, 因为它在处理矩阵及其图像处理和统计工具箱方面具有丰富的功能。我们推荐使用名为 DipImage (来自 <http://www.diplib.org/download>) 的有吸引力的 GUI 来避免在处理图像时进行大量的命令行输入。

还有一些可用于 MATLAB 的分类工具箱, 例如分类工具箱 Toolbox (<http://www.wiley.com/WileyCDA/Section/id-105036.html>, 其需要从关联的计算机手册获得密码) 和 PRTools (<http://www.prtools.org/download.html>)。我们在第 8 章使用了分类工具箱, 展示了其直观的图形用户界面。一些学生已经探索了 Weka, 这是一组机器学习算法, 用于解决在 Java 中实现的数据挖掘问题, 并且是开源的 ([http://www.cs.waikato.ac.nz/ml/weka/index\\_downloading.html](http://www.cs.waikato.ac.nz/ml/weka/index_downloading.html))。

还有一些额外的资源, 可以从 <http://extras.springer.com/> 的本书的配套网站上下载, 其中包括一些有用的 Excel 文件和数据文件。

尽管我们在校对方面做出了最大的努力, 但仍然有可能存在一些差错和缺漏。如果您发现任何问题, 请通知我们。

我非常喜欢写这本书。我希望你喜欢读它!

杰夫·多尔蒂

# 目 录

## 致 谢 译者序 前 言

第1章 引言 .....	1
1.1 概述 .....	1
1.2 分类 .....	3
1.3 本书的组织结构 .....	5
1.4 练习 .....	5
参考文献 .....	6
第2章 分类 .....	7
2.1 分类过程 .....	7
2.2 特征 .....	9
2.3 训练与学习 .....	13
2.4 监督学习与算法选择 .....	14
2.5 分类方法 .....	15
2.6 实例 .....	18
2.6.1 按形状分类 .....	18
2.6.2 按大小分类 .....	19
2.6.3 更多的实例 .....	20
2.6.4 字母的分类 .....	21
2.7 练习 .....	21
参考文献 .....	22
第3章 非度量方法 .....	23
3.1 引言 .....	23
3.2 决策树分类器 .....	23
3.2.1 信息、熵、扭曲度 .....	24
3.2.2 信息增益 .....	26
3.2.3 决策树存在的问题 .....	30
3.2.4 优缺点 .....	32
3.3 基于规则的分类器 .....	33
3.4 其他方法 .....	33
3.5 练习 .....	34
参考文献 .....	35
第4章 统计模式识别 .....	36
4.1 测量数据与测量误差 .....	36
4.2 概率论 .....	37
4.2.1 简单概率论 .....	37
4.2.2 条件概率和贝叶斯规则 .....	38
4.2.3 朴素贝叶斯分类器 .....	44
4.3 连续随机变量 .....	45
4.3.1 多变量高斯分布 .....	48
4.3.2 协方差矩阵 .....	49
4.3.3 马氏距离 .....	57
4.4 练习 .....	60
参考文献 .....	62
第5章 监督学习 .....	63
5.1 参数与非参数学习 .....	63
5.2 参数学习 .....	63
5.2.1 贝叶斯决策理论 .....	63
5.2.2 判别函数与决策边界 .....	72
5.2.3 MAP (最大后验) 估计量 .....	80
5.3 练习 .....	81
参考文献 .....	82
第6章 非参数学习 .....	83
6.1 直方图估计与 Parzen 窗口 .....	83
6.2 k-NN (k-最近邻域) 分类 .....	85
6.3 人工神经网络 .....	88
6.4 内核机 .....	99
6.5 练习 .....	101
参考文献 .....	101



<b>第7章 特征提取及选择</b> .....	103	9.1.1 偏差和方差 .....	133
7.1 降维 .....	103	9.2 交叉验证和重采样方法 .....	134
7.1.1 预处理 .....	104	9.2.1 Holdout 方法 .....	134
7.2 特征选择 .....	104	9.2.2 $k$ 重交叉验证 .....	135
7.2.1 类内/类间距离 .....	104	9.2.3 自助法 .....	136
7.2.2 子集选择 .....	106	9.3 测量分类器的性能 .....	137
7.3 特征提取 .....	106	9.4 比较分类器 .....	141
7.3.1 主成分分析法 .....	106	9.4.1 ROC 曲线 .....	141
7.3.2 线性判别分析 .....	113	9.4.2 McNemar 检验 .....	143
7.4 练习 .....	118	9.4.3 其他统计检验 .....	143
参考文献 .....	119	9.4.4 分类工具箱 .....	143
<b>第8章 非监督式学习</b> .....	120	9.5 组合分类器 .....	146
8.1 聚类 .....	120	参考文献 .....	149
8.2 $k$ -Means 聚类 .....	121	<b>第10章 项目</b> .....	150
8.2.1 模糊 $c$ -Means 聚类 .....	125	10.1 视网膜弯曲度作为疾病的	
8.3 (聚合)层次聚类 .....	125	指示器 .....	150
8.4 练习 .....	129	10.2 纹理分割 .....	153
参考文献 .....	130	10.3 生物特征识别系统 .....	156
<b>第9章 评估和比较分类器</b> .....	131	10.3.1 指纹识别 .....	157
9.1 比较分类器和“没有免费午餐		10.3.2 面相识别 .....	158
定理” .....	131	参考文献 .....	159

### 1.1 概述

人类善于识别物体（通用术语称为模式），并且理所当然地认为我们具备这种能力，但却发现在分析的过程中每个步骤又很困难。通常，区分人的声音和小提琴的声音、手写数字“3”和“8”、洋葱和玫瑰的气味是很容易的。每天，我们都在识别周围的人脸，但这完全是无意识的，因为不能解释这一专长，而且我们发现很难编写计算机程序来做同样的事情。每个人的脸都是由特定的结构（眼睛、鼻子、嘴巴等）组成的一个模式，并且分布在脸的某个位置。通过分析人脸样本图像，一个程序应该能捕获人脸的模式，并将其作为脸（作为我们已知的某一范畴或者某一类中的一个成员）来识别（或辨认），这就是模式识别。这里可能有许多范畴（或类别），我们必须将特定的脸分类到某个范畴（或类别）中，这就是分类。要注意的是：“模式识别”这个术语得到了广泛地解释，并不需要重复的暗示，它用于包括我们想分类的所有对象，如苹果（或橘子），语言的波形和指纹。

类属是相似对象的集合，对象只需要相似不需要完全相同，不同的类属之间是有区别的。图 1.1 表明了我们预先知道的类和在检查对象后所创建的类别之间的不同。

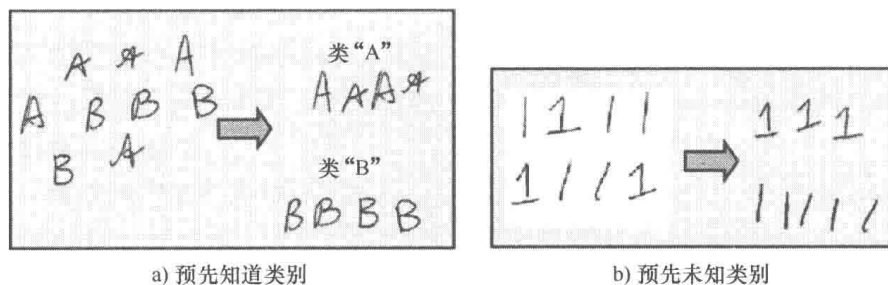


图 1.1 分类问题图解

由于新兴的应用，人们对模式识别和分类的研究不断增加和深入，这不仅具有挑战性，而且对计算技术要求很高。这些应用包括：

- 数据挖掘（通过筛选大量的数据来提取少量相关的和有用的信息，如欺诈监测、金融预测和信用评分）

- 生物统计学（基于面部、虹膜和指纹等的物理属性的个人身份识别）
- 机器视觉（例如，一条流水线上的自动视觉检查）
- 字符识别（例如，根据邮政编码自动分拣邮件、自动取款机上的自动检查扫描器）
- 文档识别（例如，通过邮件的标题和内容识别邮件是否是垃圾邮件）
- 计算机辅助诊断（例如，基于对医学数据的解释，如乳腺摄影图像、超声波图像、心电图（ECGs）和脑电图（EECs），帮助医生做出诊断决定）
- 医学图像（例如，基于核磁共振图像（MRI）扫描的结果可将细胞分类为恶性或良性，或者将功能 MRI 中的大脑活动图像分类为不同的情绪和认知状态）
- 语言识别（例如，帮助残疾人控制机器）
- 生物信息学（例如，通过对 DNA 序列的分析来检测与基因相关的特定疾病）
- 遥感（例如，土地使用和作物产量）
- 天文学（根据其形状对星系进行分类；例如像外星球智能调查（SETI）中的自动调查，通过分析无线电望远镜的数据试图去定位可能源于人造的信号）

这些实用的方法已经在各个领域独立发展。在统计学中，从特定观察到一般描述的过程称之为推论，学习 [即，使用例证（训练）数据] 称之为估计，我们所知的分类方法称之为判别分析（McLachlan 1992）。在工程中，分类叫作模式识别，其方法是无参数的，更多的是凭经验。其他的方法起源于机器学习（Alpaydin 2010）、人工智能（Russell & Norvig 2002）、人工神经网络（Bishop 2006）和数据挖掘（Han 和 Kamber 2006）。我们将结合这些技术从不同的侧重点，给出一个更加统一的分类说明（见图 1.2）。

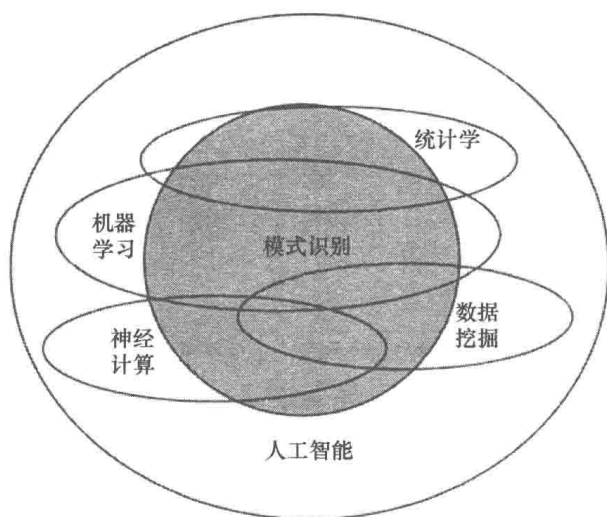


图 1.2 模式识别及其相关学科领域

## 1.2 分类

分类通常是一般过程的最后一步（见图 1.3），它的任务是将对象分类到单独的类属中。以图像为例，对获取的图像加以分割以从背景中分离出不同的对象，不同的对象贴上不同的标签。典型的模式识别系统包括一个感知器、预处理机制（预先分割）、特征提取机制、一系列已经分类（后期处理）实例（训练数据）和分类算法。特征提取这一步通过测量标签对象的确切的特征属性或特征（如大小、形状和结构）来减少数据量。然后将这些特征（确切地说是这些特征的值）传递给分类器，分类器对所提交的特征进行评估计算，将物体分配给某个类别，其分类准则一般是判断特征的值落在类容差范围以内还是以外。例如，通过这个过程可用于分类身体损伤是良性还是恶性。

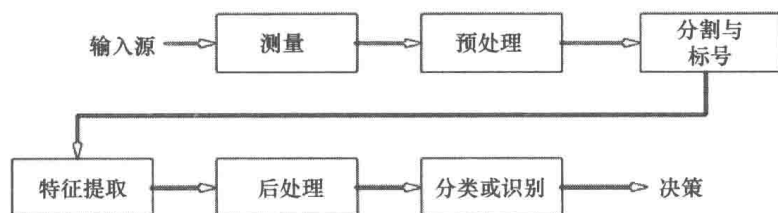


图 1.3 通用分类系统流程

获取图像的质量取决于图像系统的分辨率、敏感度、带宽和信噪比。如图像增强（如亮度调整、对比增强、图像拉平、频域滤波、边缘增强）和图像复原（如照片标准校正、逆滤波、维纳滤波）的预处理步骤需要在分割前进行，这通常是一个有挑战性的过程。典型情况下，图像增强先于图像复原。通常情况下，这些过程都是按顺序执行的，但在较复杂的任务中需要循环处理，就是说，后续的处理步骤有时需将参数返回到先前的步骤中，这样图像处理过程就包括大量的迭代循环。

特征的质量与它们能够区分来自不同类属的能力有关。来自相同类属的实例应该有相似的特性值，而不同类属的例子应该有不同的特性值，也就是说，好的特征应该有小的类内变化和大的类间变化（见图 1.4）。在加载给分类器前，测量特征会转换或映射到另一个可选择的特征空间，以产生更好的有区分力的特征。

假定特征可以是连续的（定量的），也可以是类属的或非测量的（定性的），这在数据挖掘应用中比较常见。分类的特征可以是名义上的（即无序的，例如邮政编码、员工 ID、性别）或是依照次序的（即有序的，例如街道

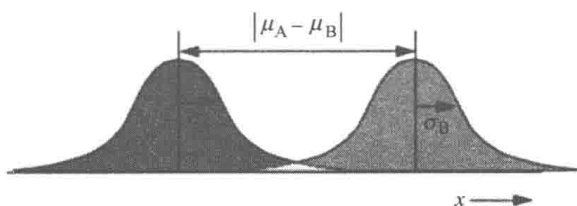


图 1.4 对两个不同的类别（蓝色和红色）进行度量的  
一个好的特征  $x$  应该有小的类内变化和大的类间变化

号、年级、满意的级别可以是非常不满意、不满意、还可以、好、非常好)。有些能力可以将数据从一种类型转换为另一种类型，例如，连续的数据可以离散化为序列数据，序列数据可以指派为整型数据（尽管它们缺少很多真实数字的属性，但应该可以当作符号对待）。首选的特征总是包含大量的信息（因此，最具有区分力）。如果可以选择，科学应用一般更喜欢连续数据，因为可以利用它们（例如，均值和标准差）做更多的研究。对于属类数据，人们可能会怀疑是否所有相关的类别已经做出了解释，或者它们可能随时间而变化。

人类擅长于使用尺寸、形状、颜色和其他视觉线索在图像内辨别对象。尽管对象会在不同的视角和不同的光照条件下呈现，有不同的大小，或者会加以旋转，甚至它们在视觉上被部分遮挡时（见图 1.5），我们仍能够辨认出它们。但是通常这些任务对于机器识别系统来说很具有挑战性。



图 1.5 人脸识别需要能够处理不同的表情、光照和遮挡

分类器的目标是将新数据（测试数据）分类到某一个类别，这些类别是被决策域定义的，决策域之间的边界称之为决策线（见图 1.6）。

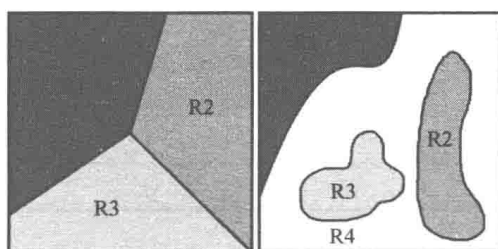


图 1.6 分类可映射为带有边界线的决策域

分类技术可以分为两个宽泛的领域：统计技术和结构（或称句法）技术。另外还有组合使用二者的第三个领域，称为认知方法，其中包括神经网络和遗传算法。第一个领域处理的对象或模式，对它们的产生有潜在的、可计量的统计基础，可以通过数量的特性来描述，如长度、面积和结构。第二个领域处理的对象由定性特征可进行最好地描述，描述对象中固有的结构或句法关系。统计分类方法比结构方法使用更为广泛。认知方法已在过去数十年间受到欢迎。这些方法并不是独立的，涉及多种分类器的混合系统日益普遍（Fu1983）。

### 1.3 本书的组织结构

第2章介绍分类过程的细节和不同的方法以及分类任务的一些实例。第3章介绍非度量方法（如决策树等）。第4章将讨论概率论，导入到贝叶斯规则和统计模式识别的起源。第5章讲解监督学习以及参数学习的例子，我们将看到用不同的方法去评估分类器的性能。第6章讲解非参数学习。第7章讲解维度灾难以及如何将特征数量保持到一个有效的最小值。第8章讲解非监督式学习的技术。第9章讲解评估各种分类器的性能的方法。第10章讨论一些有意义的分类问题。

如果有选择地回避一些细节，本书可以用于一个学期的学习。如果进行精讲，并附带一些适量的练习或编程及一些项目工作，则可能需要两个学期。另外，对于自学者，可以根据自己的节奏来学习这些材料，通过几个月的时间尽情享受这些内容，快乐地学习吧！

### 1.4 练习

1. 列举一些文章中未提到的分类应用程序。
2. 根据四个成年人的数据，说明了他们的体重和他们的健康状况。设计一个简单的分类器可以合理地分类所有四个模式。

体重 (kg)	类别标签
50	不健康
60	健康
70	健康
80	不健康

使用这个分类器，如何分类体重为 76kg 的第五个成年人？

3. 超市购买的以下物品和它们的一些属性如下：

项目编号	价格 ( \$ )	体积 ( $\text{cm}^3$ )	颜色	类别标签
1	20	6	蓝色	便宜
2	50	8	蓝色	便宜
3	90	10	蓝色	便宜
4	100	20	红色	昂贵
5	160	25	红色	昂贵
6	180	30	红色	昂贵

三个特征（价格、体积和颜色）中哪个是最好的分类属性？

4. 考虑将对象分类为圆和椭圆的问题，该如何分类这些对象？

## 参考文献

- [1] Alpaydin, E.: Introduction to Machine learning, 2nd edn. MIT Press, Cambridge (2010)
- [2] Bishop, C. M.: Neural Networks for Pattern Recognition. Oxford University Press, Oxford (2006)
- [3] Duda, R. O., Hart, P. E., Stork, D. G.: Pattern Classification, 2nd edn. Wiley, New York (2001)
- [4] Fu, K. S.: A step towards unification of syntactic and statistical pattern recognition. IEEE Trans. Pattern Anal. Mach. Intell. 5, 200–205 (1983)
- [5] Han, J., Kamber, M.: Data Mining: Concepts and Techniques, 2nd edn. Morgan Kaufmann, San Francisco (2006)
- [6] McLachlan, G. J.: Discriminant Analysis and Statistical Pattern Recognition. Wiley, New York (1992)
- [7] Russell, S., Norvig, P.: Artificial Intelligence: A Modern Approach, 2nd edn. Prentice Hall, New York (2002)

## 2.1 分类过程

常见的分类过程是阶段之间无反馈的分类系统，如图 2.1 所示。

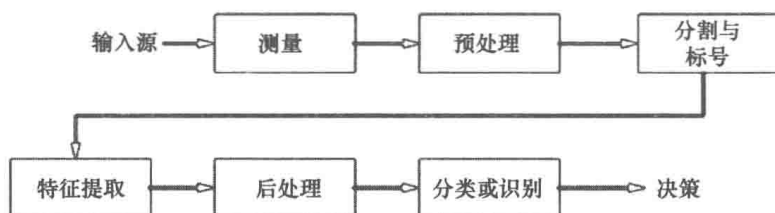


图 2.1 通用分类系统流程

感知/采集阶段使用诸如照相机或麦克风这样的变换器。所获取的信号（如一幅图像）必须具有足够的品质特性，这样所区分的“特征”能加以正确测量。这取决于变换器的特性，例如对于照相机性能应包括以下几个方面：分辨率、动态范围、灵敏度、失真、信噪比、聚焦好坏等。

预处理通常用于图像分割的情况。例如，平滑图像（比如，通过高斯模板卷积）通过阈值来减轻分割时噪声引起的混杂效应（因为包含噪声的随机波动可以导致像素跨越阈值移动以及错误分类）。用中值滤波器进行图像预处理可以有效去除散射噪声（例如椒盐噪声）。去除可变背景亮度和直方图均衡化经常用来确保均匀的亮度。

根据不同情况，我们可能要处理丢失的数据（Batista 和 Monard, 2003），并检测和处理异常数据（Hodge 和 Austin, 2004）。

分割，即将一幅图像分割成对特定任务有意义的几个区域，比如前景（包括感兴趣对象）和背景（包括其他所有景象）。有两种主要的方法：一是基于区域的方法，可以检测相似性；二是基于边界的方法，可以检测到间断性（边缘）并连接形成连续的边界和区域。

基于区域的方法可以根据像素之间的相似性找出连通区域。定义区域最基本的特征是图像的灰度级别和亮度，而其他一些特征如颜色或纹理也可以使用。



然而，如果要求一个区域中的像素非常相似，我们可能会过度分割图像；如果允许区域中的像素有很大的差异性，我们可能会将原来独立的目标对象融合起来。要达到的目标就是找到目标对象相对应的区域，就像人类看到它们一样，而这个目标是难以实现的。

基于区域的方法包括阈值方法（使用全局或局部自适应阈值法，最优阈值法（如 Otsu 方法、isodata 方法或最大熵阈值法））。如果这样导致目标对象重叠，那么使用图像的距离转换阈值或分水岭算法可有助于将它们分开。基于区域的其他方法包括区域生长法（使用“种子”像素的自下而上方法）和分裂合并法（自上而下的基于四叉树的方法）。

基于边界的方法趋向于使用边缘检测器（如 Canny 检测器）或者边缘连接，用于连接边缘的所有断裂处，或通过边界跟踪形成连续的边界。或者使用活动轮廓（或蛇形轮廓）方法，这是一个可以弹性控制的连续轮廓，可以通过对它边缘的锁定来封闭目标对象。

分割提供了一个简化的二元图像，它把感兴趣的对象（前景）从背景中分离，同时为以后的测量保留它们的形状和大小。前景像素设置为“1”（白色），背景像素设置为“0”（黑色）。通常，我们更愿意用离散的数值给图像中的对象做标记。连通组件标记算法扫描分割的二元图像，然后根据像素的连通性将像素分为不同的组件，也就是说，在一个连通组件里的所有像素具有相同的像素值，并且在某种程度上互相连接。一旦分组确定，根据它分配的组件，每个像素都标记为一个数值（1, 2, 3, ...），这些数值可以看作图像显示的灰度级或颜色（见图 2.2）。

标记的一个明显结果是图像中的对象可以很容易地加以计算。更一般的是，标记的二进制对象可以用来掩盖原始图像从而分离出每个（灰度）对象，但仍保留其原始的像素值，这样可以单独测量其属性或特征。掩模可以通过几种不同的方式进行设计。二进制掩模可用在显示硬件中的覆盖通道或 alpha 通道中，用来屏蔽像素的显示。还可以使用掩码来修改存储的图像。这可以通过灰度图像乘以二进制掩模来实现或通过位操作将原始图像和二进制掩模进行“与”运算来实现。可以独立测量的隔离特征，是显著性区域（RoI）处理的基础。

图像分割的后期处理可以为特征提取做准备。例如，局部对象可以从图像周边加以移除（见图 2.2e），非连通对象可以进行融合，小于或大于特定限制的对象可以加以移除，对象或背景中的孔可以通过形态学中的开运算或闭运算得到填充。