

数字图像分析与处理技术

# Visual C++

## 数字图像模式识别技术 及工程实践

◆ 求是科技 张宏林 编著

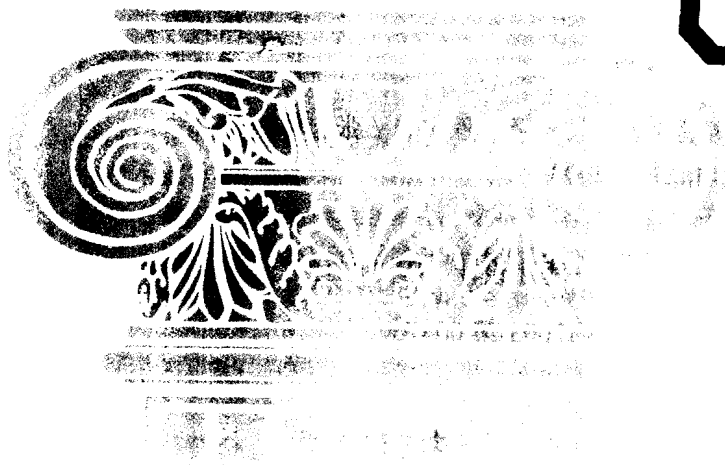
人民邮电出版社  
POSTS & TELECOMMUNICATIONS PRESS



附光盘  
CD-ROM

数字图像分析与处理技术

TP312  
1049D



# Visual C++

## 数字图像模式识别技术 及工程实践

◆ 求是科技 张宏林 编著

人民邮电出版社

## 图书在版编目(CIP)数据

Visual C++数字图像模式识别技术及工程实践 / 张宏林编著.

—北京: 人民邮电出版社, 2003.2

ISBN 7-115-10951-6

I. V... II. 张... III. C 语言—程序设计 IV. TP312

中国版本图书馆 CIP 数据核字(2003)第 004766 号

## 内容提要

本书介绍了模式识别和人工智能中的一些基本理论以及一些相关的模型, 包括贝叶斯决策、线性判别函数、神经网络理论、隐马尔可夫模型、聚类技术等, 同时结合模式识别中的一些经典问题, 比如说字符识别、笔迹鉴定、人脸检测、车牌识别、印章识别以及遥感图片、医学图片处理等内容, 从多种不同的角度, 介绍了这些问题的解决思路。

本书的最大的特色在于对于其中的大多数问题, 给出了基于 C/VC++6.0 的实现代码, 且具有一定的扩展性。有的实例还给出了不同方法的实现, 以供读者选择。

本书可作为读者学习模式识别与人工智能时的参考书。

数字图像分析与处理技术

Visual C++数字图像模式识别技术及工程实践

◆ 编 著 求是科技 张宏林  
责任编辑 张立科

◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街 14 号  
邮编 100061 电子函件 315@ptpress.com.cn  
网址 <http://www.ptpress.com.cn>  
读者热线 010-67132692  
北京汉魂图文设计有限公司制作  
北京密云春雷印刷厂印刷  
新华书店总店北京发行所经销

◆ 开本: 787×1092 1/16

印张: 30.5

字数: 946 千字

印数: 1-5 000 册

2003 年 2 月第 1 版

2003 年 2 月北京第 1 次印刷

ISBN7-115-10951-6/TP·3270

定价: 54.00 元 (附光盘)

本书如有印装质量问题, 请与本社联系 电话: (010) 67129223

# 前言

我们在日常生活中经常进行模式识别的活动。比如说，我们能够分辨出桌子、椅子，很小的时候就能够分出自己的父母，能够听出是谁的声音，能够进行正常的阅读，这些都是我们习以为常的能力，在计算机出现以前，没有人对此表现出惊奇，也没有人注意到人类的模式识别能力是一个值得研究的课题。当计算机出现以后，人工智能开始发展，模式识别也随之成为一个热门课题，当科学家发现用机器实现人类的模式识别能力是如此麻烦时，人们才意识到这个问题的难度。

“模式”(Pattern)这个词的原意是指供模仿用的，完美无缺的标本，这是一个相当含蓄的定义，并且触及了一些相当深奥的论题。它使我们想起了 Plato 的观点：客观世界本身正是完美理想的不尽完美的复制品。在心理学中，模式识别是作为一个过程来定义的，通过这一过程，到达感觉器官的外界信号被转换成有意义的感性经验。然而，要精确地定义什么是“有意义的感性经验”，本身就是一件很麻烦的事。

模式识别和人工智能，在大学计算机、自动化、电子等专业中都是比较重要的课程，作者在学习这些课的时候，深深体会到，有些算法，理解起来容易，但真正用程序实现，却又未必是很简单的事情，对于初学者而言，有时候面对一些问题，甚至根本无法下手。

近几年中，作者看了不少这方面的书，也查阅了大量的关于机器学习、文字识别、签名鉴定以及人脸识别方面的论文，当然了，也得益于很多网络上的文章和示例代码，自己动手实现了一些算法，写了一些应用程序。所以如果能将这些程序及收获整理出来，也许对初学模式识别和人工智能的朋友会有所帮助。

书中介绍的一些算法，大部分是散见在各种书籍和论文中的，也有部分来自网络上的资料，这些思想都属于原作者所有。由于中间持续的时间比较长，有些算法的出处已经不是太清楚，如果没有在书末的参考文献中列举出来，还望原作者原谅。书中的程序大部分都是作者自己实现的，有少数是同学提供的或者是由网上一些公开的算法改变的（未见这些算法有版权申明），这些程序仅供读者学习使用，请勿用于其他目的。

另外，由于本书不是 VC++ 的入门书籍，所以除了个别程序外，大部分程序没有给出项目如何创建的细节，相信这些对了解 VC++ 的读者不是难事。

本书由求是科技策划，张宏林编写，此外参加资料收集和整理工作的人有肖洪伟、李廷文、张增强、王洪涛、吴继刚、周学明、李闽溟、黄沙、宣小平、但正刚、张文毅、张小磊、胡昱、范国平、陈晓鹏、王凯封、潘邦传、王锐、闫卫东、赵明华、许福、施新刚、郑刚、李现勇、谭思量、邹超群、郭瑞军、宋文超、彭珂珂、忤浩、齐燕杰、赵苏琦、徐建军、胡伟、刘江、王茹、闫海荣、刘理、谭春华、张益贞、刘韬、刘鹏、黄良大、周金萍、黄超、王非、冉光志、金海、赵振、孙越、纪慧波、刘子锐、刘成、沈兆普、陈雍、章丽萍、王钦、许主红、陈志华、王朝龙等。如果本书能够对读者有所帮助，作者将感到万分荣幸。另外，如果读者在阅读本书时发现了什么错误或者在使用程序方面有什么疑问，以及有什么好的建议，交流想法等，欢迎写信联系：zhangkile@ptpress.com.cn。

最后深深感谢曾经提供过源代码的朋友以及其他一些在作者学习过程中给予帮助的朋友！

作者

2003 年 1 月

# 目 录

|                     |    |
|---------------------|----|
| 第1章 绪论              | 1  |
| 1.1 模式和模式识别的概念      | 1  |
| 1.2 模式空间、特征空间和类型空间  | 2  |
| 1.3 模式识别系统的构成       | 3  |
| 1.3.1 信息获取          | 3  |
| 1.3.2 预处理           | 4  |
| 1.3.3 特征提取和选择       | 4  |
| 1.3.4 分类决策          | 4  |
| 1.4 物体的结构表示         | 4  |
| 1.5 图片识别问题          | 5  |
| 1.6 关于本书的内容安排和程序    | 5  |
| 第2章 模式识别中的基本决策方法    | 6  |
| 2.1 基于最小错误率的贝叶斯决策   | 6  |
| 2.2 分类器设计           | 9  |
| 2.2.1 多类情况          | 10 |
| 2.2.2 两类情况          | 11 |
| 2.3 关于分类器的错误率       | 11 |
| 2.4 关于贝叶斯决策一些讨论     | 11 |
| 2.5 线性判别函数的基本概念     | 12 |
| 2.6 设计线性分类器的主要步骤    | 13 |
| 2.7 Fisher 线性判别     | 13 |
| 2.8 解决多类问题决策树       | 15 |
| 2.8.1 决策树的基本概念      | 16 |
| 2.8.2 决策树设计的基本考虑    | 16 |
| 第3章 常用的模型和算法介绍      | 17 |
| 3.1 人工神经网络的发展简史     | 17 |
| 3.2 人工神经元           | 18 |
| 3.2.1 神经元模型         | 18 |
| 3.2.2 几种常用的作用函数     | 19 |
| 3.3 人工神经网络构成        | 22 |
| 3.3.1 基本模型          | 22 |
| 3.3.2 前向网络          | 22 |
| 3.4 人工神经网络的学习规则     | 23 |
| 3.4.1 Hebb 学习规则     | 23 |
| 3.4.2 $\delta$ 学习规则 | 23 |
| 3.5 BP 神经网络         | 24 |
| 3.5.1 BP 网络模型       | 24 |
| 3.5.2 输入输出关系        | 24 |
| 3.5.3 网络学习训练        | 24 |
| 3.5.4 BP 网络的设计问题    | 25 |
| 3.5.5 BP 网络的限制与不足   | 27 |

|                                    |           |
|------------------------------------|-----------|
| 3.5.6 BP 算法的改进 .....               | 27        |
| 3.6 BP 算法的 C 语言实现及使用方法 .....       | 29        |
| 3.7 用 BP 网络解决异或问题 .....            | 40        |
| 3.8 标量量化 .....                     | 43        |
| 3.8.1 基本概念 .....                   | 43        |
| 3.8.2 均匀量化 .....                   | 44        |
| 3.8.3 非均匀量化 .....                  | 45        |
| 3.9 矢量量化 .....                     | 46        |
| 3.9.1 基本原理 .....                   | 46        |
| 3.9.2 失真测度 .....                   | 48        |
| 3.9.3 设计码本 .....                   | 48        |
| 3.10 矢量量化算法的 C 语言实现 .....          | 50        |
| 3.11 HMM 基本思想 .....                | 58        |
| 3.11.1 Markov 链 .....              | 58        |
| 3.11.2 HMM 的概念 .....               | 59        |
| 3.12 HMM 基本算法 .....                | 60        |
| 3.12.1 前向后向算法 .....                | 61        |
| 3.12.2 Viterbi 算法 .....            | 63        |
| 3.12.3 Baum-Welch 算法 .....         | 63        |
| 3.13 基本 HMM 模型的 C 语言实现 .....       | 64        |
| 3.13.1 数据结构和函数定义 .....             | 65        |
| 3.13.2 一些基本工具 .....                | 66        |
| 3.13.3 HMM 结构的操作函数 .....           | 74        |
| 3.13.4 前向后向算法 .....                | 78        |
| 3.13.5 Viterbi 算法 .....            | 82        |
| 3.13.6 Baum-Welch 算法 .....         | 85        |
| 3.13.7 随机数生成函数 .....               | 89        |
| 3.13.8 序列操作函数 .....                | 89        |
| <b>第 4 章 常用搜索算法 .....</b>          | <b>93</b> |
| 4.1 状态空间法 .....                    | 93        |
| 4.1.1 状态 (State) .....             | 93        |
| 4.1.2 问题的状态空间 (State Space) .....  | 93        |
| 4.2 盲目搜索算法 .....                   | 93        |
| 4.2.1 宽度优先搜索 .....                 | 94        |
| 4.2.2 深度优先搜索 .....                 | 94        |
| 4.3 启发式搜索算法 .....                  | 96        |
| 4.3.1 搜索深度、启发函数和评价函数 .....         | 96        |
| 4.3.2 A 算法和 A* 算法 .....            | 96        |
| 4.4 A* 算法类的实现 .....                | 98        |
| 4.5 8 数码游戏 (Eight-Puzzle) 简介 ..... | 105       |
| 4.6 关于 8 数码游戏解的存在性讨论 .....         | 106       |
| 4.6.1 问题的表达 .....                  | 106       |
| 4.6.2 问题的转化与证明 .....               | 106       |
| 4.7 算法设计 .....                     | 107       |

|                                  |     |
|----------------------------------|-----|
| 4.8 程序实现 .....                   | 108 |
| 4.8.1 程序创建步骤 .....               | 109 |
| 4.8.2 数据结构和函数定义 .....            | 110 |
| 4.8.3 各种算法的实现 .....              | 111 |
| 4.8.4 可视化的实现 .....               | 122 |
| 4.9 黑白棋简介 .....                  | 131 |
| 4.9.1 黑白棋规则 .....                | 131 |
| 4.9.2 黑白棋基本战术 .....              | 131 |
| 4.10 算法设计 .....                  | 133 |
| 4.10.1 博弈算法基础 .....              | 133 |
| 4.10.2 Alpha-Beta 剪枝 .....       | 135 |
| 4.10.3 估值函数 .....                | 139 |
| 4.10.4 开局及终局 .....               | 145 |
| 4.11 程序实现 .....                  | 146 |
| 4.11.1 程序创建步骤 .....              | 146 |
| 4.11.2 程序代码 .....                | 147 |
| <b>第 5 章 联机字符识别</b> .....        | 167 |
| 5.1 汉字识别的历史和现状 .....             | 167 |
| 5.2 联机字符识别原理框图 .....             | 168 |
| 5.2.1 统计决策方法 .....               | 169 |
| 5.2.2 句法结构方法 .....               | 170 |
| 5.3 基于笔划及笔划特征二级分类的联机汉字识别 .....   | 170 |
| 5.3.1 笔划的分类 .....                | 170 |
| 5.3.2 笔划识别前的噪声处理 .....           | 171 |
| 5.3.3 笔划方向码合并处理及笔划识别 .....       | 172 |
| 5.3.4 笔划间特征量的定义及识别 .....         | 173 |
| 5.3.5 整字匹配的距离准则 .....            | 174 |
| 5.4 基于活动模板引导的子结构的识别 .....        | 175 |
| 5.4.1 系统模型 .....                 | 176 |
| 5.4.2 活动模板子结构的构造 .....           | 176 |
| 5.4.3 子结构引导的结构匹配 .....           | 178 |
| 5.5 实例之联机手写数字识别 .....            | 180 |
| 5.5.1 难点及特征的选取 .....             | 180 |
| 5.5.2 相应的预处理及模板的建立 .....         | 181 |
| 5.5.3 程序的实现 .....                | 182 |
| 5.6 实例之联机手写数字、英文字符及汉字识别 .....    | 196 |
| <b>第 6 章 脱机字符识别</b> .....        | 219 |
| 6.1 印刷体汉字的识别 .....               | 219 |
| 6.1.1 印刷体汉字的基本知识 .....           | 219 |
| 6.1.2 汉字的行切割和字切割 .....           | 219 |
| 6.1.3 文字的归一化 .....               | 220 |
| 6.1.4 基于统计量的特征 .....             | 221 |
| 6.2 基于置信度分析和多信息融合的手写数字识别方法 ..... | 225 |
| 6.2.1 多种特征和多种分类器 .....           | 225 |

|       |                             |     |
|-------|-----------------------------|-----|
| 6.2.2 | 集成方法 .....                  | 227 |
| 6.2.3 | 预处理 .....                   | 229 |
| 6.3   | 其他手写数字识别方法简介 .....          | 231 |
| 6.3.1 | 基于支持向量机 (SVM) 的方法 .....     | 231 |
| 6.3.2 | 伪二阶隐马尔可夫模型应用于手写数字识别 .....   | 232 |
| 6.3.3 | 基于骨架特征顺序编码的识别方法 .....       | 232 |
| 6.4   | 手写数字识别实例之模板匹配法 .....        | 233 |
| 6.4.1 | 位图的读写 .....                 | 233 |
| 6.4.2 | 细化算法 .....                  | 235 |
| 6.4.3 | 特征提取与识别 .....               | 251 |
| 6.4.4 | 程序实现 .....                  | 251 |
| 6.5   | 手写数字识别实例之 Fisher 线性判别 ..... | 271 |
| 6.5.1 | USPS 数据库 .....              | 271 |
| 6.5.2 | Fisher 判别程序 .....           | 274 |
| 6.6   | 数字识别实例之神经网络法 .....          | 286 |
| 第 7 章 | 在线签名鉴定 .....                | 299 |
| 7.1   | 时间弯折算法 .....                | 299 |
| 7.1.1 | 时间弯折的概念 .....               | 299 |
| 7.1.2 | 时间弯折的限制 .....               | 300 |
| 7.1.3 | 时间弯折的 DP 方法 .....           | 300 |
| 7.1.4 | DTW 方法的扩充和变形 .....          | 302 |
| 7.1.5 | 模板的建立 .....                 | 302 |
| 7.1.6 | 算法的实现 .....                 | 302 |
| 7.2   | 签名分段算法 .....                | 312 |
| 7.3   | 自回归分析 .....                 | 313 |
| 7.4   | 联机签名可资利用的特征 .....           | 320 |
| 7.5   | 基于特征函数法的联机签名鉴定 .....        | 321 |
| 7.5.1 | 系统框图 .....                  | 321 |
| 7.5.2 | 预处理 .....                   | 321 |
| 7.5.3 | 特征提取 .....                  | 322 |
| 7.5.4 | 特征匹配 .....                  | 323 |
| 7.6   | 在线签名鉴定系统实例 .....            | 324 |
| 7.6.1 | 签名数据的输入 .....               | 324 |
| 7.6.2 | 一些结构的定义 .....               | 324 |
| 7.6.3 | 方向分布的计算 .....               | 326 |
| 7.6.4 | 文件数据的读取 .....               | 331 |
| 7.6.5 | 预处理函数 .....                 | 333 |
| 7.6.6 | 识别算法 .....                  | 342 |
| 7.6.7 | 保存和打开模板 .....               | 345 |
| 第 8 章 | 离线签名鉴定 .....                | 347 |
| 8.1   | 离线签名鉴定的一些基于统计特征的方法 .....    | 347 |
| 8.1.1 | 距离匹配变换 .....                | 347 |
| 8.1.2 | 形状特征 .....                  | 347 |
| 8.1.3 | 纹理特征 .....                  | 353 |

|               |                          |            |
|---------------|--------------------------|------------|
| 8.2           | 伪动态特征 .....              | 361        |
| 8.3           | 总结 .....                 | 362        |
| <b>第 9 章</b>  | <b>人脸的检测与定位 .....</b>    | <b>363</b> |
| 9.1           | 人脸检测方法综述 .....           | 363        |
| 9.1.1         | 基于知识的自顶向下的方法 .....       | 363        |
| 9.1.2         | 基于人脸特征的自底向上的方法 .....     | 363        |
| 9.1.3         | 模板匹配的方法 .....            | 363        |
| 9.1.4         | 基于人脸外观的方法 .....          | 364        |
| 9.2           | 基于肤色的人脸检测算法 .....        | 364        |
| 9.2.1         | 色彩空间与色彩空间的聚类 .....       | 364        |
| 9.2.2         | 肤色模型 .....               | 365        |
| 9.2.3         | 人脸区域分割 .....             | 370        |
| 9.2.4         | 肤色模型在人脸检测的后期验证中的应用 ..... | 373        |
| 9.3           | 人脸特征的检测 .....            | 374        |
| 9.3.1         | 候选特征的提取 .....            | 375        |
| 9.3.2         | 双眼和嘴巴的定位 .....           | 375        |
| 9.3.3         | 双眼和嘴巴的轮廓提取 .....         | 375        |
| 9.4           | 人脸检测与定位实例 .....          | 377        |
| 9.4.1         | 人脸区域的检测 .....            | 377        |
| 9.4.2         | 眼睛的标定 .....              | 398        |
| 9.4.3         | 鼻子的确定 .....              | 407        |
| 9.4.4         | 嘴的确定 .....               | 410        |
| 9.4.5         | 主程序的其他一些代码 .....         | 413        |
| <b>第 10 章</b> | <b>车牌识别技术 .....</b>      | <b>422</b> |
| 10.1          | 系统简介 .....               | 422        |
| 10.1.1        | 车牌定位技术综述 .....           | 423        |
| 10.1.2        | 车牌字符识别技术综述 .....         | 423        |
| 10.2          | 车牌图像定位与分割算法 .....        | 424        |
| 10.2.1        | 车牌图像的特点及识别难点 .....       | 424        |
| 10.2.2        | 边缘提取算法 .....             | 426        |
| 10.2.3        | Hough 变换提取直线 .....       | 434        |
| 10.2.4        | 车牌检测的要点 .....            | 438        |
| 10.2.5        | 算法流程 .....               | 440        |
| 10.3          | 车牌字符的识别 .....            | 441        |
| 10.3.1        | 二值化 .....                | 441        |
| 10.3.2        | 倾斜度的校正 .....             | 442        |
| 10.3.3        | 大小归一化 .....              | 442        |
| 10.3.4        | 匹配识别字符 .....             | 442        |
| <b>第 11 章</b> | <b>印章识别 .....</b>        | <b>444</b> |
| 11.1          | 伪印章的制作及人工防伪技术 .....      | 444        |
| 11.1.1        | 常用伪造印章的方法及其特征 .....      | 444        |
| 11.1.2        | 真假印章印文的检验 .....          | 446        |
| 11.2          | 自动印章识别系统的设计 .....        | 446        |
| 11.2.1        | 预处理 .....                | 447        |

|                               |            |
|-------------------------------|------------|
| 11.2.2 特征的提取 .....            | 447        |
| 11.3 算法实现 .....               | 447        |
| <b>第 12 章 图像的纹理分析方法 .....</b> | <b>452</b> |
| 12.1 纹理分析概念 .....             | 452        |
| 12.2 空间灰度层共现矩阵 .....          | 453        |
| 12.3 纹理能量测量 .....             | 456        |
| 12.4 纹理的结构分析方法和纹理剃度 .....     | 458        |
| 12.4.1 纹理的结构分析方法 .....        | 458        |
| 12.4.2 纹理梯度 .....             | 458        |
| 12.5 遥感图像的纹理分析 .....          | 458        |
| 12.5.1 云类的自动识别 .....          | 458        |
| 12.5.2 台风的自动识别 .....          | 460        |
| 12.6 细胞图像的彩色纹理分析 .....        | 463        |
| 12.6.1 纹理的彩色分布特征描述 .....      | 463        |
| 12.6.2 纹理彩色特征 .....           | 464        |
| 12.6.3 细胞图像处理 .....           | 464        |
| 12.7 VC 应用实例 .....            | 465        |
| 12.7.1 灰度共现矩阵类 .....          | 466        |
| 12.7.2 几个响应函数 .....           | 473        |
| <b>参考文献 .....</b>             | <b>476</b> |

# 第 1 章 绪论

模式识别诞生于 20 世纪 20 年代，随着 40 年代计算机的出现，50 年代人工智能的兴起，模式识别在 20 世纪 60 年代迅速发展成为一门学科。它所研究的理论和方法在很多科学和技术领域中得到了广泛的重视，推动了人工智能系统的发展，扩大了计算机应用的可能性。几十年来，模式识别研究取得了大量的成果，也实现了很多成功的应用，但由于实际系统中涉及到很多复杂的问题，面对这些问题，现有的理论和方法就显得有些不足了。如何将这些方法结合实际问题加以应用，是绝大多数涉足这一领域的读者所必须考虑的问题。

## 1.1 模式和模式识别的概念

我们在日常生活中经常进行模式识别的活动。比如说，我们能够分辨出桌子、椅子，很小的时候就能够分出自己的父母，能够听出是谁的声音，能够进行正常的阅读，这些都是我们习以为常的能力，在计算机出现以前，没有人对此表现出惊奇，也没有人注意到人类的模式识别能力是一个值得研究的课题。当计算机出现以后，人工智能开始发展，模式识别也随之成为一个热门课题，当科学家发现用机器实现人类的模式识别能力是如此麻烦时，人们才意识到这个问题的难度。

“模式”(Pattern)这个词的原意是指供模仿用的，完美无缺的标本，这是一个相当含蓄的定义，并且触及了一些相当深奥的论题。它使我们想起了 Plato 的观点：客观世界本身正是完美理想的不尽完美的复制品。在心理学中，模式识别是作为一个过程来定义的，通过这一过程，到达感觉器官的外界信号被转换成有意义的感性经验。然而，要精确地定义什么是“有意义的感性经验”，本身就是一件很麻烦的事。

一个主要的问题是模式这一概念是如何形成的。它可以由演绎或者归纳过程而得到。首先，我们假定模式这一概念是观察者本身所固有的，或者假定观察者是通过许多不完全例子的观察而抽象出这一概念的。当这些被观察的例子被标以一种或几种已给定的模式时，这一过程可以称之为“有导师”的学习。如果没有这样的标记也同样行之有效的过程，称之为“无导师”的学习。

完成这一工作的精神过程显然很复杂，且还不太为我们所知晓。模式识别是一个不仅被人类也被动物所履行的过程。并且，在进化的意义上，这个过程有着确定的生存价值。“抽象化”或者“理想化”也确实允许一个生物以一种相似于由以前的经验所证实行之有效的方法去应付新的同类型情况。较低等的生物所具备的抽象概念也许只限于危险、食物和交配，而较高级生物所具有的显然更为丰富。模式识别在生物学意义上的重要性表明，人类和动物的神经系统可能已经发展了行之有效的回路。我们也许能概括这一观察，而说神经系统对处理具备生存价值的任务要比没有生存价值的任务远远的有效，这样，我们就不会对人类很容易完成一项复杂的模式识别任务，而对做多位数乘法却感到相当困难这一点感到惊奇。

30 年前，数字计算机出现了，它的强有力的数值计算能力，使得人们期望着它也能像人脑一样，具备非数值计算能力，然而，人们早期的乐观期望并没有成为现实，比起先进的计算机来，人脑的机制更为复杂，计算机在数值计算能力方面超过人类并不代表它的通用计算能力也很强，这种非数值计算能力如何在计算机上实现，恰恰成为人工智能和机器模式识别的重要任务。

简单一点说，模式被理解成取自世界有限部分的单一样本的被测量值的综合；模式识别就是试图确定一个样本的类别属性，即把某一样本归属于多个类型中的某一个类型。

我们这里要研究的是通过机器的自动识别。这就需要把人们的知识和经验教给机器，为机器制定一

些规则和方法，并且让机器具有综合分析和自动分类的判断能力，以使使机器能够完成自动识别的任务。当机器学会自动识别后，可以完成人们自己难以完成的许多工作。因此，模式识别技术被广泛地应用于人工智能、计算机工程、机器入学、神经生物学、医学、侦探学以及高能物理、考古学、地质勘探、宇航科学和武器技术等许多重要的领域，随着高科技的迅速发展，模式识别技术必将获得更广泛的应用，其基本理论和方法也会愈益丰富

1.2 模式空间、特征空间和类型空间

从模式识别技术途径来说，由模式空间经过特征空间到类型空间是模式识别所经历的过程。为了说明这些概念，首先解释“物理上可以觉察到的世界”在模式识别范畴内，在客观世界里存在这样的一些物体和事件，它们各自都能被适当选择的和足够多的函数来描述，或者说它们在物理上都是可以测量的，它们的可测数据的集合就称为物理上可以觉察到的世界。显然，这些可测数据，或者说这个世界维数是无限多的。

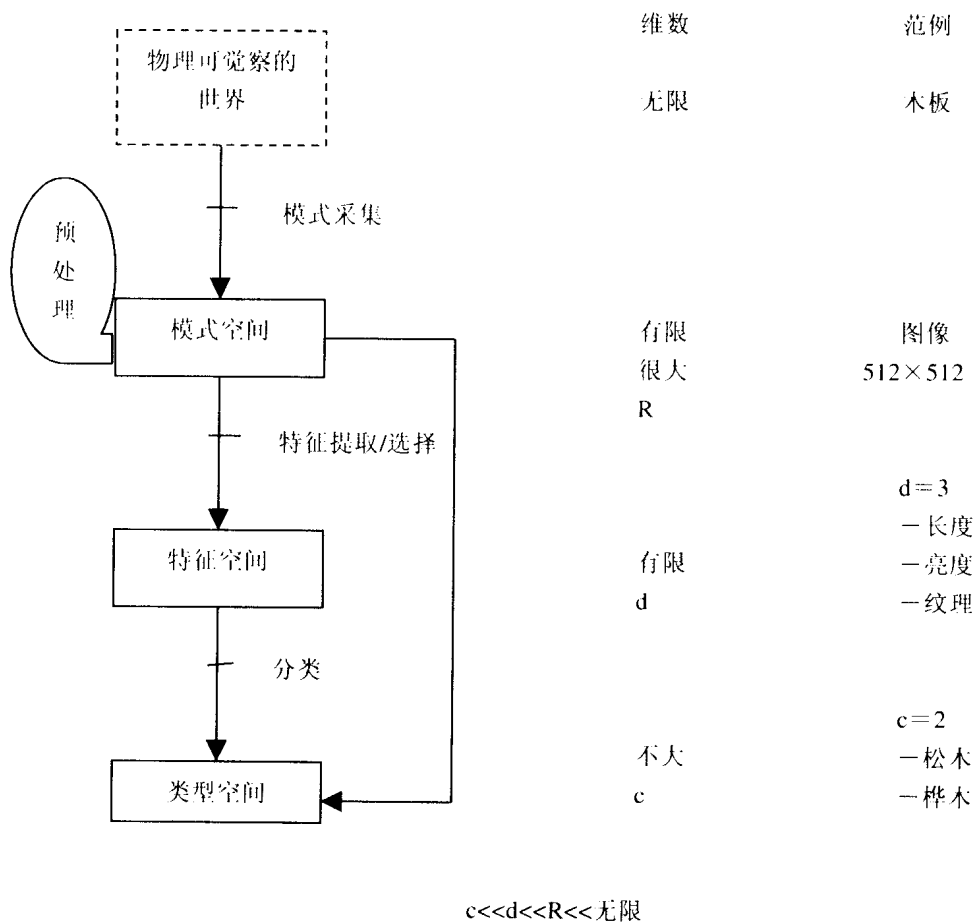


图 1-1 模式识别过程

在物理上可以觉察到的世界里，适当地选择某些物体和事件，我们把它们称为样本，对它们分别进行观测。如前所述，每个样本的观测数据的综合都构成模式，所有的样本观测数据则构成模式空间。显然，模式空间的维数与所选择的样本和测量方法有关，也与特定的应用有关，一般说来是很大的，但是

有限值。在模式空间里，每个模式样本都是一个点，点的位置由该模式在各维上的测量数据来确定。由物理上可以觉察到的世界到模式空间经历的过程称为模式采集。

模式空间的维数虽然是有限的，还是非常多的，其中有些并不能有效地提示样本的实质。正像人们对事物进行判断之前要进行综合分析一样，机器在做出判断之前也要对模式空间里的各坐标元素进行综合分析，获取最能揭示样本属性的观测量作为主要特征，这些主要特征就构成特征空间。显然，特征空间的维数大大压缩了。特征空间的每个坐标都是样本的主要特征，简称特征。每个样本在特征空间里也是一个点，点的位置由该样本的各特征值来确定。由模式空间到特征空间所需要的综合分析，往往包含适当的变换和选择，称为特征提取和特征选择。关于特征提取和特征选择的内容，我们在后面还会研究。

由某些知识和经验可以确定分类准则，称之为判决规则。根据适当的判决规则，把特征空间里的样本区分成不同的类型，从而把特征空间塑成了类型空间。类型空间里不同类型之间的分界面，常称为决策面。类型空间的维数与类型的数目相等，一般地说，小于特征空间的维数。由特征空间到类型空间所需要的操作就是分类判决。

这个完整的过程可以由图 1-1 形象地表示。

## 1.3 模式识别系统的构成

有两种基本的模式识别方法，即统计模式识别方法和结构（句法）模式识别方法，与此相应的模式识别系统都是有二个过程（设计与实现）所组成。“设计”是指用一定数量的样本（训练集/学习集）进行分类器的设计。“实现”是指用所设计的分类器对待识别的样本进行分类决策。本书的例子主要是用统计模式识别方法，在用到结构模式识别的方法时，我们会对其再加以介绍。基于统计模式识别方法的系统主要由以下几个部分组成：信息获取、预处理、特征提取和选择、分类决策，如图 1-2 所示。

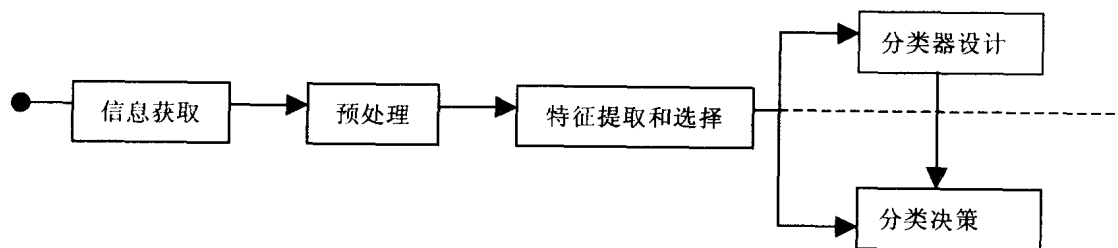


图 1-2 模式识别系统的基本构成

下面，我们对这几个部分作一下简化的说明。

### 1.3.1 信息获取

为了使计算机能够对各种现象进行分类识别，要用计算机可以运算的符号来表示所研究的对象，通常输入对象的信息有下列 3 种类型，即

#### 1. 二维图像

如文字、指纹、地图、照片这类对象。

#### 2. 一维波形

如脑电图、L1 电图、机械震动波形等。

### 3. 物理参量和逻辑值

前者如在疾病诊断中病人的体温及各种化验数据等；后者如对某参量正常与否的判断或对症状有无的描述，如疼与不疼，可用逻辑值即 0 和 1 表示。在引入模糊逻辑的系统中，这些值还可以包括模糊逻辑值，比如很大、大、比较大等。

通过测量、采样和量化，可以用矩阵或向量表示二维图像或一维波形。这就是信息获取的过程。

## 1.3.2 预处理

预处理的目的是去除噪声，加强有用的信息，并对输入测量仪器或其他因素所造成的退化现象进行复原。

## 1.3.3 特征提取和选择

由图像或波形所获得的数据量是相当大的。例如，一个文字图像可以有几千个数据，一个心电图波形也可能有几千个数据，一个卫星遥感图像的数据量就更大。为了有效地实现分类识别，就要对原始数据进行变换，得到最能反映分类本质的特征。这就是特征提取和选择的过程。

## 1.3.4 分类决策

分类决策就是在特征空间中用统计方法把被识别对象归为某一类别。基本作法是在样本训练集基础上确定某个判决规则，使按这种判决规则对被识别对象进行分类所造成的错误识别率最小或引起的损失最小。

## 1.4 物体的结构表示

为了让计算机处理任何类型的信息，首先必须使其量化，即用数字表示之。这往往会产生大量的数据。例如，一个图像可以用格子划分为一些充分小的区域，在这些区域上，就所用的测量仪器而言，亮度和彩色强度均一致。从而它可以由二个矩阵来描述。第一个矩阵的元素是相应于各个格子上彩色的代码值，第二个矩阵的元素是相应于各个格子上所测得的光强度。一个心电图按离散时间间隔取样，且表示为一个在这些采样时刻上的波形振幅值。这样一个数组是量化了的波形。一些机械设备的工作状况可以用描述温度、油压等时间函数的一组量化波形及另外一些单个测量来表现。病人的状况可以用一组波形(心电图、脑电图等)、数值(温度、血球数等)和逻辑值(对症状有无的描述)来描述。诚然这样的描述保存了全部有用信息，却太详细了以至不能直接应用。

为把握这些未加预处理的原始信息费力甚大，所以用更紧凑的方法来表示所考虑的物体是必要的，但如可能不要损失任何信息。在某种意义下，这是个编码问题且是所有模式识别算法的第一步。显然，表示很强烈地依赖于数据类型。例如，波形可以通过级数展开的系数来表达。对图片(如亲笔签名)，固然亦可如法炮制，但也存在其他方法。图片可被划分成最大的有相同色彩和亮度的区域。进而用这些区域的边界来描述图片。这两种表示法的任何一种都不一定会有信息丢失。如有任何信息丢失，也是可以控制的。例如，一个包含 512 个取样点的波形可以由 512 个 Fourier 级数展开式来表达(256 个正弦和 256 个余弦)。如不考虑低振幅谐波可以忽略，以进一步减少表示的大小。在这一情况下，由此引起的误差仍可保持在严格的界限之下。此外，在图片亮度函数的分段常数逼近时，可以通过允许各区域上亮度与其平均值间有更大的变异而扩大区域。

## 1.5 图片识别问题

在一些最棘手的模式识别问题中,数据以图片形式出现。由于人类视觉的物理基础尚未被很好认识,这一课题对模式识别提出了特殊的挑战。光学错觉的存在提示我们:我们正着手于一个高级的复杂过程,并且如不考虑更高级的精神过程视觉信息的机器处理也许是不妥当的。概括的图片模式识别将涉及下列各步:

- 图片的取样与量化:一幅景物或一张照片被转换为一个数组,以适合计算机处理。
- 图片分割:按亮度、彩色或纹理的一致与否确定区域。
- 景物分析:由分割获得的区域被合并或修改,使之可被定为物体。
- 形状描述:物体被编码为反映它们形状的定量的结构。
- 物体描述:可能是简单分类(例如,所见物体被分入字母A类);或者是一段语言的拙述(例如所见物体由二个被水平线连接的圆盘所组成)。

上述第一步一般属于信号处理的范围,涉及到硬件,我们将不予讨论。在讨论图片时总假设有一个图片文件(.bmp)存在,且可以通过函数将之读取到二维数组中。图片分割位于图像处理和模式识别的分界线,所以本书予以讨论。有一点是显而易见的,即至少在人进行模式识别时,上述各步是相互作用的。如果初始分割结果得出了“陌生”的形式,那么可能进行一个新的分割。用技术的术语来说,在分割时当我们对该物体没有先验知识时,应用自底向上方法分割。而在有这类知识的情况下,常用自顶向下方法。

## 1.6 关于本书的内容安排和程序

本书介绍了模式识别中的一些基本理论,给出了一些应用实例,具体内容安排如下:

第2章介绍了模式识别中的一些基本决策方法,包括贝叶斯决策,线性判别函数等;第3章介绍了常用的模型和算法,包括人工神经网络、矢量量化、隐马尔可夫模型;第4章介绍了常用的一些搜索算法,给出了8数码问题和黑白棋这两个实例;第5~12章介绍了模式识别应用中的一些实例,包括字符识别、签名鉴定、人脸的检测定位、车牌识别、印章识别以及图像的纹理分析等。对于每个实例,给出了详细的分析过程,有的提供了多种方法,并给出了实现代码,且有一定的扩展性。

## 第2章 模式识别中的基本决策方法

贝叶斯方法是现有模式识别文献中的经典方法，也一直被认为是统计模式识别的基础方法。近几年来，对贝叶斯模式识别方法有一些批评意见。当然，这些批评不是针对其中的模式识别部分，而是针对贝叶斯方法的本质。有一类批评意见是怀疑用先验概率和条件类别概率表达后验概率时，贝氏关系是否有效和具有实际意义。批评者认为，由于很少可能知道先验概率，因而这种表达关系的实用性是很有限的；另外，由于它没有提供可信度的主观度量手段，所以，这些表达关系是不充分的。

尽管有这些批评意见，并不能否认贝叶斯方法的重要地位，至少，它可以让我们对模式识别系统有一个直观而具体的理解

用贝叶斯方法进行分类时要求：

- (1) 各类别总体的概率分布是已知的；
- (2) 要决策分类的类别数是一定的。

考虑连续的情况，不妨假设要识别的物理对象有  $d$  种特征观察量  $x_1, x_2, \dots, x_d$ ，这些特征的所有可能的取值范围构成了  $d$  维特征空间，这里，称  $x = [x_1, x_2, \dots, x_d]^T$  为  $d$  维特征向量

另外，我们还假设要研究的问题有  $c$  个类别，每个类别的状态用  $\omega_i$  来表示， $i = 1, 2, \dots, c$ ；对应于各个类别出现的先验概率  $P(\omega_i)$  是已知的。如果在特征空间已观察到某一向量  $x$ ， $x = [x_1, x_2, \dots, x_d]^T$  就是  $d$  维空间上的某一个点，本章要研究的问题就是怎样合理的将  $x$  归到某一类里面去。由于本书不是理论书籍，所以只介绍最基本的决策准则。

### 2.1 基于最小错误率的贝叶斯决策

在模式分类问题中，人们往往希望尽量减少分类的错误，从这样的要求出发，利用概率论中的贝叶斯公式，就能得出使错误率为最小的分类规则，称之为基于最小错误率的贝叶斯决策。

在讨论一般理论之前，先举一个癌细胞的识别例子来说明解决问题的过程。假设每个要识别的细胞已作过预处理，抽取出  $d$  个表示细胞基本特性的特征，成为一个  $d$  维空间的向量  $x$ ，识别的目的是要将  $x$  分类为正常细胞或者异常细胞。用决策论的术语来讲就是将  $x$  归类于两种可能的自然状态之一，如果用  $\omega$  表示状态，则：

$\omega = \omega_1$  表示正常；       $\omega = \omega_2$  表示异常；

类别的状态是一个随机变量，而某种状态出现的概率是可以估计的。例如，根据医院细胞病理检查的大量统计资料可以对某一地区正常细胞和异常细胞出现的比例作出估计，这就相当于在识别前已知正常状态出现的概率  $P(\omega_1)$  和异常状态出现的概率  $P(\omega_2)$ 。这种由先验知识在识别前就得到的概率  $P(\omega_1)$  和  $P(\omega_2)$  成为状态的先验概率。在两类识别问题中显然有  $P(\omega_1) + P(\omega_2) = 1$ 。如果不做细胞特征的仔细观测，只依靠先验概率  $P(\omega_1)$  和  $P(\omega_2)$  趋做决策，那么合理的决策就应该为：若  $P(\omega_1) > P(\omega_2)$ ，则作出

$\omega = \omega_1$  的决策; 反之, 则作出  $\omega = \omega_2$  的决策。显然这是不合理的, 因为在这个例子中, 由于  $P(\omega_1) > P(\omega_2)$ , 如果仅仅按照先验概率来决策, 就会把所有的细胞都归于正常细胞类别中。

为此, 我们做进一步分析, 这里, 为简单起见, 我们假定特征向量是一维的 (即只用一个特征),  $d = 1$ 。根据本章开始时提的要求, 在自然状态下观察的类别条件概率分布应为已知, 如图 2-1 所示。

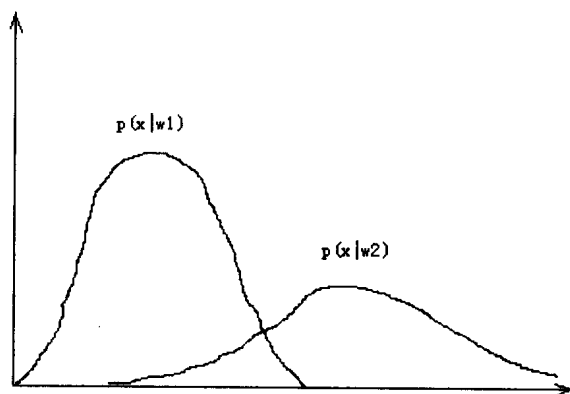


图 2-1 类条件概率密度

$p(x|\omega_1)$  是正常状态下细胞特征观察  $x$  的类条件概率密度;

$p(x|\omega_2)$  是异常状态下细胞特征观察  $x$  的类条件概率密度。

设  $P(\omega_i|x)$  为后验概率, 则有贝叶斯公式

$$P(\omega_i|x) = \frac{p(x|\omega_i)P(\omega_i)}{\sum_{j=1}^2 p(x|\omega_j)P(\omega_j)}$$

可见, 贝叶斯公式实际上是通过观察  $x$  (记被识别细胞特征的测量) 把状态的先验概率转化为状态的后验概率  $P(\omega_i|x)$ , 见图 2-2。

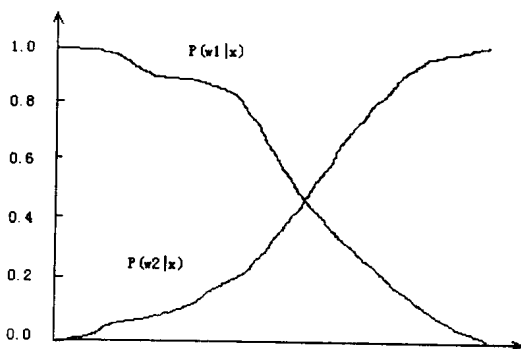


图 2-2 后验概率

这样, 基于最小错误概率的贝叶斯决策规则为: