



高等学校计算机科学与技术教材

- ① 原理与技术的完美结合
- ① 教学与科研的最新成果
- ① 语言精炼，实例丰富
- ① 可操作性强，实用性突出

图像模式识别 ——VC++技术实现

□ 杨淑莹 编著

□ 李兰友 主审



配套光盘提供

遗传算法、神经网络、
贝叶斯、模糊聚类等
先进算法的源代码，
VC++可编译执行。



清华大学出版社

● 北京交通大学出版社

高等学校计算机科学与技术教材

图像模式识别

——VC++ 技术实现

杨淑莹 编著

李兰友 主审

清华大学出版社

北京交通大学出版社

·北京·

内 容 简 介

本书介绍图像模式识别的各种算法及其编程实现。全书共分为 10 章, 内容包括: 模式识别的基本概念、分类器设计, 模板匹配分类器, 基于概率统计的 Bayes 分类器, 几何分类器, 神经网络分类器, 图像分割与特征提取, 聚类分析, 模糊聚类分析, 遗传算法聚类分析。

本书实用性强, 选材新颖, 包括了神经网络、模糊集理论、遗传算法等新技术, 针对每一种模式识别技术, 书中分为理论基础、实现步骤、编程代码三部分, 所有算法都用 VC++ 编程实现, 程序结构简单, 代码简洁, 便于初学者很快掌握模式识别技术。

本书可作为高等院校计算机工程、信息工程、生物医学工程、智能机器人学、工业自动化、模式识别等学科本科生、研究生的教材或教学参考书, 亦可供有关工程技术人员参考。

版权所有, 翻印必究。举报电话: 010-62782989 13501256678 13801310933

本书封面贴有清华大学出版社防伪标签, 无标签者不得销售。

本书防伪标签采用特殊防伪技术, 用户可通过在图案表面涂抹清水, 图案消失, 水干后图案复现; 或将表面膜揭下, 放在白纸上用彩笔涂抹, 图案在白纸上再现的方法识别真伪。

图书在版编目(CIP)数据

图像模式识别: VC++ 技术实现/杨淑莹编著. —北京: 清华大学出版社; 北京交通大学出版社, 2005.7

(高等学校计算机科学与技术教材)

ISBN 7-81082-481-3

I. 图… II. 杨… III. ① 计算机应用-图像识别-高等学校-教材 ② C语言-程序设计-高等学校-教材 IV. ① TP391.4 ② TP312

中国版本图书馆 CIP 数据核字 (2005) 第 061742 号

责任编辑: 韩 乐 特邀编辑: 郑宏云

出版者: 清华大学出版社 邮编: 100084 电话: 010-62776969 <http://www.tup.com.cn>

北京交通大学出版社 邮编: 100044 电话: 010-51686414 <http://press.bjtu.edu.cn>

印刷者: 北京瑞达方舟印务有限公司

发行者: 新华书店总店北京发行所

开 本: 185×260 印张: 17.75 字数: 443 千字 附光盘 1 张

版 次: 2005 年 7 月第 1 版 2005 年 7 月第 1 次印刷

书 号: ISBN 7-81082-481-3/TP·207

印 数: 1~5000 册 定价: 36.00 元

本书如有质量问题, 请向北京交通大学出版社质监组反映。对您的意见和批评, 我们表示欢迎和感谢。

投诉电话: 010-51686043, 51686008; 传真: 010-62225406; E-mail: press@center.bjtu.edu.cn。

前 言

模式识别属于当代高科技研究的重要领域,已经发展为一门独立的新学科。模式识别技术迅速扩展,几乎遍及各个学科,广泛应用于人工智能、机器人、系统控制、遥感数据分析、生物医学工程、军事目标识别等领域,在国民经济、国防建设和社会发展等方面得到广泛应用,产生了深远的影响。

国内外论述模式识别技术的书为数不少,这一领域涉及深奥的数学理论,往往使实际工作者理解起来感到困难,大部分书是罗列模式识别的各种算法,见不到算法的实际效果和各种算法对比的结果,而这正是学习者和实际工作者所需要了解和掌握的内容。目前确实缺少一本关于模式识别技术在实际应用方面具有系统性、可比性和实用性的书籍。

本书内容基本涵盖了目前模式识别重要的理论和方法,但并没有简单地将各种理论方法堆砌起来,而是将作者自身的研究成果和实践经验传授给读者,在介绍理论方法的同时,将各种算法应用于实际实例中讲解。书中含有需要应用模式识别技术解决的实际问题,有模式识别理论的讲解和推理,有将理论转化为编程的步骤,有计算机能够运行的源代码,有计算机运行模式识别算法程序后的效果,有不同算法应用于同一个问题的效果对比,使读者不至于面对如此丰富的理论和方法无所适从,而是有所学就会有所用。

本书主要介绍统计模式识别,由于篇幅有限,没有讨论句法模式识别。第1章介绍图像模式识别的基本概念,位图的基础知识;第2章至第6章介绍分类问题,其中第2章介绍分类器设计,第3章介绍模板匹配分类器,第4章介绍基于概率统计的 Bayes 分类器,第5章介绍几何分类器,第6章介绍神经网络分类器;第7章至第10章介绍聚类问题,其中第7章介绍图像分割与特征提取,第8章介绍聚类分析,第9章介绍模糊聚类分析,第10章介绍遗传算法聚类分析。

本书的特点有以下几点。

1. 选用新技术。在介绍许多重要的经典内容基础上,书中还详细介绍了最近十几年来才发展起来的并被实践证明有用的新技术、新理论,比如神经网络、模糊集理论、遗传算法等,并将这些新技术应用于模式识别当中,提供这些新技术的源代码。

2. 实用性强。采用具体实例介绍理论和技术,使理论和实践相结合,避免了空洞的理论说教。书中实例取材于手写数字模式识别,由于数字识别属于多类问题,在实际应用中具有广泛的代表性,读者对程序稍加改进,就可以应用到不同的场合,例如文字识别、字符识别、图形识别等。

3. 针对每一种模式识别技术,书中分为理论基础、实现步骤、编程代码三部分,在掌握了基本理论之后,按照实现步骤的指导,可以了解算法的实现思路和方法,再进一步体会短小精悍的核心代码,学习者可以很快掌握模式识别技术,经过应用本书提供的实例程序,立刻会见到算法的实际效果。所有算法都用 VC++ 编程实现,便于读者学习和扩展应用。

4. 考虑到手写数字每次有所不同,本书配套光盘中提供了两个综合性的实例(分类实例和聚类实例),每个实例包含了多种模式识别算法,这样针对同一次手写的数字,可以应用各种算法进行识别,读者可以比较各种算法的识别效果。

5. 当读者运行本书配套光盘实例时,首先在左视图随机手写 0~9 的数字,手写数字的图

像大小固定,背景是白色,笔是黑色,并可擦除修改。然后单击右视图(激活右视图),调用各种模式识别算法,输出识别结果。书中省略了滤去干扰、噪声等有关图像处理的内容,请参考有关书籍。

本书可作为高等院校计算机工程、信息工程、生物医学工程、智能机器人学、工业自动化、模式识别等学科本科生、研究生的教材或教学参考书,亦可供有关工程技术人员参考。

参加本书编写工作的还有章慎锋、王厚雪、洪俊、任翠池等,由天津工业大学软件学院李兰友教授主审,正是在他们的帮助下本书得以顺利完成,作者表示衷心的感谢。由于编者业务水平和实践经验有限,书中缺点与错误在所难免,欢迎读者予以指正!

电子邮箱:ysying1262@126.com。

杨淑莹
2005年5月于天津

目 录

第 1 章 模式识别的基本概念	1
1.1 模式识别的基本概念	1
1.2 图像识别	5
1.3 位图基础	7
1.3.1 数字图像的基本概念	7
1.3.2 BMP 文件结构	9
1.3.3 CDib 类库的建立	11
小结	18
习题	18
第 2 章 分类器设计	19
2.1 特征空间优化设计问题	19
2.2 分类器设计准则	21
2.3 分类器设计基本方法	24
2.4 判别函数	25
2.5 分类器的选择	27
2.6 训练与学习	28
小结	29
习题	29
第 3 章 模板匹配分类器	30
3.1 特征类设计	30
3.2 待测样品特征提取	35
3.3 训练集特征库的建立	37
3.4 模板匹配分类法	39
小结	42
习题	42
第 4 章 基于概率统计的 Bayes 分类器	43
4.1 Bayes 决策的基本概念	43
4.1.1 Bayes 决策所讨论的问题	43
4.1.2 Bayes 公式	44
4.2 基于最小错误率的 Bayes 决策	46
4.3 基于最小风险的 Bayes 决策	49
4.4 Bayes 决策比较	51
4.5 基于二值数据的 Bayes 分类实现	52
4.6 基于最小错误率的 Bayes 分类实现	55
4.7 基于最小风险的 Bayes 分类实现	59
小结	64

习题	64
第 5 章 几何分类器	65
5.1 几何分类器的基本概念	65
5.2 线性判别函数	66
5.3 线性判别函数的实现	70
5.4 感知器算法	70
5.5 增量校正算法	78
5.6 LMSE 验证可分性	85
5.7 LMSE 分类算法	91
5.8 Fisher 分类	95
5.9 线性分类器实现分类的局限性	101
5.10 非线性判别函数	103
5.11 分段线性判别函数	105
5.12 势函数法	108
小结	114
习题	114
第 6 章 神经网络分类器	115
6.1 人工神经网络的基本原理	115
6.1.1 人工神经元	115
6.1.2 人工神经网络模型	118
6.1.3 神经网络的学习过程	118
6.1.4 人工神经网络在模式识别问题上的优势	119
6.2 BP 网络设计	120
6.2.1 三层 BP 网络学习算法	121
6.2.2 BP 网络设计需要考虑的问题	123
6.3 神经网络分类器设计	125
小结	134
习题	134
第 7 章 图像分割与特征提取	135
7.1 聚类简介	135
7.2 图像阈值分割	138
7.2.1 直方图门限选择阈值	139
7.2.2 半阈值选择分割	141
7.3 图像的标识及特征提取	143
7.4 图像的轮廓提取	151
7.5 图像的测量	156
7.5.1 二值图像的区域面积测量	156
7.5.2 二值图像的周长测量	157
小结	158
习题	158

第 8 章 聚类分析	159
8.1 聚类的设计	159
8.2 模式相似性测量	161
8.3 基于试探的未知类别聚类算法	167
8.3.1 最临近规则的试探法	167
8.3.2 最大最小距离算法	172
8.4 层次聚类算法	175
8.4.1 最短距离法	176
8.4.2 最长距离法	180
8.4.3 中间距离法	184
8.4.4 重心法	189
8.4.5 类平均距离法	193
8.5 动态聚类算法	201
8.5.1 K 均值算法	202
8.5.2 迭代自组织的数据分析算法 (ISODATA)	205
小结	213
习题	213
第 9 章 模糊聚类分析	214
9.1 模糊集的基本概念	214
9.2 模糊集运算	217
9.2.1 模糊子集运算	217
9.2.2 模糊集运算性质	218
9.3 模糊关系	218
9.4 模糊集在模式识别中的应用	223
9.5 模糊聚类分析	224
小结	240
习题	240
第 10 章 遗传算法聚类分析	241
10.1 遗传算法基本概念	241
10.2 遗传算法的构成要素	243
10.2.1 染色体的编码	243
10.2.2 适应度函数	244
10.2.3 遗传算子	245
10.3 控制参数的选择	247
10.4 基于遗传算法的聚类分析	247
小结	266
习题	266
附录 A 几种主要矩阵运算的程序代码	267
参考文献	275

第 1 章 模式识别的基本概念

本章要点:

- 模式识别的基本概念
- 图像识别
- 位图基础

1.1 模式识别的基本概念

模式识别(Pattern Recognition)就是机器识别、计算机识别或机器自动识别,目的在于让机器自动识别事物。例如手写数字的识别,就是将手写的数字分到具体的数字类别中;智能交通管理系统的识别,就是判断是否有汽车闯红灯,闯红灯的汽车车牌号码;还有文字识别,语音识别,图像中物体识别,等等。该学科研究的目的是使机器能做人类能做的事,具备人所具有的对各种事物与现象进行分析、描述与判断的部分能力。模式识别是直观的、无所不在的。实际上人类在日常生活的每个环节,都从事着模式识别的活动。人和动物的模式识别能力是极其平常的,但对计算机来说却是非常困难。让机器能识别、分类,需要研究识别的方法,这就是这门学科的任务。

模式识别研究的内容是利用计算机对客观物体进行分类,在错误概率最小的条件下,使识别的结果尽量与客观物体相符合。让机器辨别事物的最基本方法是计算,原则上讲是对计算机要分析的事物与标准模板的相似程度进行计算。例如要识别一个手写的数字,就要将它与 0 到 9 的模板做比较,看跟哪个模板最相似或最接近。因此首先要能度量,从中看出不同事物之间的差异,才能分辨当前要识别的事物。因此找到度量不同类事物差异的有效方法最为关键。

在模式识别学科中,就“模式”与“模式类”而言,模式类是一类事物的代表,而“模式”则是某一事物的具体体现,如数字 0,1,2,3,4,5,6,7,8,9 是模式类,而用户任意手写的一个数字或任意一个印刷数字则是“模式”,是数字的具体化。

1. 模式的描述方法

在模式识别技术中,被观测的每个对象称为样品,例如,在手写数字识别中每个手写数字可以作为一个样品,如果共写了 N 个数字,我们把这 N 个数字叫做 N 个样品 $(X_1, X_2, \dots, X_j, \dots, X_N)$,其中数字 0 有 N_0 个样品,数字 1 有 N_1 个样品,数字 2 有 N_2 个样品,数字 3 有 N_3 个样品……,一共有 $\omega_1, \omega_2, \dots, \omega_M (M=10)$ 个不同的类别。

对于一个样品来说,必须确定一些与识别有关的因素作为研究的根据,每一个因素称为一个特征。模式就是样品所具有的特征描述。模式的特征集由处于同一个特征空间的特征向量表示,特征向量的每个元素称为特征,该向量也因此称为特征向量。一般地用小写英文字母

x, y, z 来表示特征。如果一个样品 X 有 n 个特征, 则可把 X 看做一个 n 维列向量, 该向量 X 称为特征向量, 记作:

$$\mathbf{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = (x_1, x_2, \dots, x_n)^T$$

抽取图像特征的目的是为了进行分类, 识别图像。也就是说把图像变成 n 维空间的一个向量, 实际上就是看成 n 维空间中的一个点, 这样有利于从几何上考虑问题, 计算时比较方便。

若有一批样品共有 N 个, 每个样品有 n 个特征, 这些数值可以构成一个 n 行 N 列的矩阵, 称为原始资料矩阵, 见表 1-1。

表 1-1 原始资料矩阵

特征 \ 样品	X_1	X_2	...	X_j	...	X_N
x_1	x_{11}	x_{12}	...	x_{1j}	...	x_{1N}
x_2	x_{21}	x_{22}	...	x_{2j}	...	x_{2N}
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
x_i	x_{i1}	x_{i2}	...	x_{ij}	...	x_{iN}
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
x_n	x_{n1}	x_{n2}	...	x_{nj}	...	x_{nN}

模式识别问题就是根据 X 的 n 个特征来判别模式 X 属于 $\omega_1, \omega_2, \dots, \omega_M$ 类中的哪一类。待识别的不同类模式都在同一特征空间中考察, 不同类物体由于性质上的不同, 它们在各自特征取值范围上有所不同, 因而在特征空间的不同区域中出现。要记住向量的运算是建立在各个分量基础之上的。因此, 模式识别系统的目标是在特征空间和解释空间之间找到一种映射关系。特征空间是由从模式得到的对分类有用的度量、属性或基元构成的空间。解释空间由 M 个所属类别的集合构成。

如果一个对象的特征观察值为 $\{x_1, x_2, \dots, x_n\}$, 它可构成一个 n 维的特征向量值 X , 即 $X = (x_1, x_2, \dots, x_n)^T$, 式中, x_1, x_2, \dots, x_n 为特征向量 X 的各个分量。一个特征可以看作 n 维空间中的向量或点, 此空间称为模式的特征空间 R_n 。在模式识别过程中, 要对许多具体对象进行测量, 以获得许多观测值, 其中有均值、方差、协方差与协方差矩阵。

1) 均值

N 个样品的均值可表示为:

$$\bar{X} = \frac{1}{N} \sum_{j=1}^N X_j = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)^T \quad (1-1)$$

其中 \bar{x}_i 是第 i 个特征的平均值, $\bar{x}_i = \frac{1}{N} \sum_{j=1}^N x_{ij}$ 。

2) 方差

方差用来描述一批数的分散程度, 第 i 个特征的 N 个数的方差公式是:

$$s_i^2 = \frac{1}{N-1} \sum_{j=1}^N (x_{ij} - \bar{x}_i)^2 \quad (1-2)$$

方差的平方根称为均方差, $\sigma_i = \sqrt{s_i^2} = \sqrt{\frac{1}{N-1} \sum_{j=1}^N (x_{ij} - \bar{x}_i)^2}$ 。

3) 协方差与协方差矩阵

在 N 个样品中, 第 i 个特征和第 j 个特征之间的协方差定义为:

$$s_{ij} = \frac{1}{N-1} \sum_{k=1}^N (x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j) \quad (1-3)$$

对于同一批样品来说, 很明显有 $s_{ij} = s_{ji}$ 。

如果一批样品有 n 个特征 x_1, x_2, \dots, x_n 。求出每两个特征的协方差, 总共得到 n^2 个值, 将这 n^2 个值排列成以下的 n 维方阵, 称为协方差矩阵:

$$S = \begin{pmatrix} s_{11} & s_{12} & \cdots & s_{1n} \\ s_{21} & s_{22} & \cdots & s_{2n} \\ \vdots & \vdots & & \vdots \\ s_{n1} & s_{n2} & \cdots & s_{nn} \end{pmatrix}$$

协方差矩阵是对称矩阵, 而且主对角线元素 s_{ii} 就是特征 x_i 的方差 $s_i^2, i=1, 2, \dots, n$ 。

例如有 10 个学生, 其中 5 个男生, 5 个女生。对每个学生取身高、体重两项指标作为特征, 测得的数据如表 1-2 所示。

表 1-2 学生数据

样品(学生) 特征(指标)	男 生					女 生				
	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}
x_1 (身高/m)	1.70	1.75	1.65	1.80	1.78	1.60	1.55	1.60	1.65	1.70
x_2 (体重/kg)	65	70	60	65	70	60	45	45	50	55

10 个样品的均值为

$$\bar{X} = \frac{1}{10} \sum_{i=1}^{10} X_i = (1.678, 58.5)^T$$

男生和女生样品点的均值为

$$\bar{X}^{(1)} = \frac{1}{5} \sum_{i=1}^5 X_i = (1.736, 66.0)^T$$

$$\bar{X}^{(2)} = \frac{1}{5} \sum_{i=6}^{10} X_i = (1.62, 51.0)^T$$

特征 x_1 对于全体样品的方差为

$$s_1^2 = \frac{1}{10-1} [(1.70-1.678)^2 + (1.75-1.678)^2 + \cdots + (1.70-1.678)^2] = 0.00684$$

特征 x_1 对于男生和女生样品的方差分别为

$$s_1^2(1) = \frac{1}{5-1} [(1.70-1.736)^2 + (1.75-1.736)^2 + \cdots + (1.78-1.736)^2] = 0.00373$$

$$s_1^2(2) = \frac{1}{5-1} [(1.60-1.62)^2 + (1.55-1.62)^2 + \cdots + (1.70-1.62)^2] = 0.00325$$

全体样品点中特征 x_1 与 x_1 的协方差 s_{11} 即 s_1^2 , x_1 与 x_2 的协方差为

$$s_{12} = \frac{1}{10-1} [(1.70-1.678)(65-58.5) + (1.75-1.678)(70-58.5) + \dots + (1.70-1.678)(55-58.5)] = 0.636$$

在男生和女生样品点中分别有

$$s_{12}(1) = \frac{1}{5-1} [(1.70-1.736)(65-66) + (1.75-1.736)(70-66) + \dots + (1.78-1.736)(55-66)] = 0.18$$

$$s_{12}(2) = \frac{1}{5-1} [(1.60-1.62)(65-51.0) + (1.55-1.62)(45-51.0) + \dots + (1.70-1.62)(55-51.0)] = 0.163$$

很明显, 无论对于哪一批样品点总有 $s_{21} = s_{12}$ 。

2. 模式识别系统

一个典型的模式识别系统如图 1-1 所示, 由数据获取、预处理、特征提取、分类决策及分类器设计五部分组成。一般分为上下两部分。上半部分完成未知类别模式的分类; 下半部分属于设计分类器的训练过程, 利用样品进行训练, 确定分类器的具体参数, 完成分类器的设计。而分类决策在识别过程中起作用, 对待识别的样品进行分类决策。

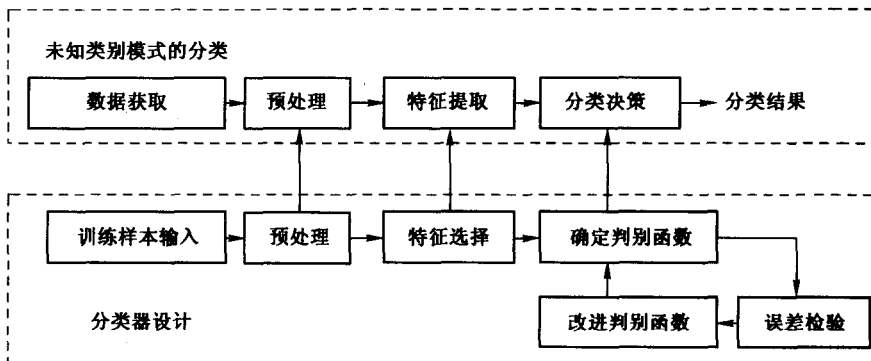


图 1-1 模式识别的过程

模式识别系统组成单元功能如下。

① 数据获取。用计算机可以运算的符号来表示所研究的对象, 一般获取的数据类型有以下几种。

- 二维图像: 文字、指纹、地图、照片等。
- 一维波形: 脑电图、心电图、季节震动波形等。
- 物理参量和逻辑值: 体温、化验数据、参量正常与否的描述。

② 预处理。对输入测量仪器或其他因素所造成的退化现象进行复原, 去噪声, 提取有用信息。

③ 特征提取。对原始数据进行变换, 得到最能反映分类本质的特征。将维数较高的测量空间(原始数据组成的空间)转变为维数较低的特征空间(分类识别赖以进行的空间)。

④ 分类决策。在特征空间中用模式识别方法把被识别对象归为某一类别。

⑤ 分类器设计。基本做法是收集样品训练集, 在此基础上确定判别函数, 改进判别函数

和误差检验。

模式识别的关键是解决如何用计算机对样品进行分类。执行模式识别的计算机系统称为模式识别系统。设计人员按需要设计模式识别系统,而该系统被用来执行模式分类的具体任务。

3. 统计模式识别主要研究问题

统计模式识别主要研究问题有:特征的选择与优化、分类判别、聚类。

1) 特征的选择与优化

如何确定合适的特征空间是设计模式识别系统一个十分重要的问题,对特征空间进行优化有两种基本方法。一是特征选择,如果所选用的特征空间能使同类物体分布具有紧致性,将为分类器设计成功提供良好的基础;反之,如果不同类别的样品在该特征空间中混杂在一起,再好的设计方法也无法提高分类器的准确性。另一种是特征的组合优化,通过一种映射变换改造原特征空间,构造一个新的精简的特征空间。

2) 分类判别

分类判别的前提是已知若干个样品的类别以及每个样品的特征,在此基础上才能对待测样品进行分类判别。例如,手写阿拉伯数字的判别是一个具有 10 类的分类问题,由于同一个数字,不同的人有不同的写法,甚至同一个人对同一个数字也有多种写法,机器首先要知道大多数手写数字的形状特征,它们属于哪一类。因此对分类问题需要建立样品库。根据这些样品库建立判别分类函数,这一过程是由机器来实现的,称为学习过程,然后对一个未知的新对象分析它的特征,决定它属于哪一类。这是一种监督分类的方法。

3) 聚类

聚类分析前提是已知若干对象和它们的特征,但是不知道每个对象属于哪一类,而且事先并不知道究竟分成多少类,在此基础上用某种相似性度量的方法,把特征相似的归为一类,即“人以类聚,物以群分”。例如,手写了若干个阿拉伯数字,把相同的数字归为一类。这是一种非监督学习的方法。

机器识别也往往借鉴人的思维活动,像人类一样找出待识别物的外形或颜色等特征,进行分析、判断,然后加以分门别类,即识别它们。模式识别的方法很多,很难将其全部概括,也很难说哪种方法最佳,常常需要根据实际的情况运用多种方法进行实验,然后选择最佳的分类方法。

1.2 图像识别

模式识别是人工智能领域应用的基础,它利用计算机和光学系统来识别计算机“看到”的图像信息,模拟人的视觉;用计算机和声音传感器来识别计算机接收到的声波信息,模拟人的听觉;用计算机通过压力、温度、气体、液体等传感器来识别计算机获得的各种特征信息,模拟人的触觉和嗅觉等功能。在视觉、听觉和触觉的识别中,基于视觉图像识别具有特别重大的意义。从信息论的角度来看,“图像”所包含的信息量最大,内容极为广泛,具体来说可以是各种物体的黑白或彩色图画、手写字符、遥感图片、声波信号、X 射线透视胶片、指纹图案、空间物体,等等。这些对象可以概括为两个类型,一是有直觉形象的——图片、相片、图案、文字图样等;一是无直觉形象而只有数据或信号的波形——语言、声音、心电图、地震波等。图像识别目

的在于用计算机自动处理某些信息系统,以代替人去完成图像分类及辨识的任务,对图像识别来说,无论是数据、信号或平面图形,甚至物体,都是除掉与它们各不相同的物理内容,考虑对它们进行“分类”这一共性来研究的。针对这一共性,以统一的观点把同一种共性者归为一类,另一种共性者归为另一类。例如,手写阿拉伯数字就需分为 10 类,手写英文字母就要分成 26 类。

1. 图像识别过程

1) 图像预处理

为了研究图像内容的识别,首先要对获得的图像信息进行预处理,滤去干扰、噪声,当图像信息微弱无法辨识时,还须对图像进行增强处理,几何调整,颜色校正等,以便人、机分析。

对于文字信息的处理,在定位和分离之后还需要采用平滑处理、细化处理和边界跟踪。平滑的目的是去掉笔道中孤立的黑点,以及笔道边缘小的凹凸处。细化处理的目的是获得文字的骨架或轮廓,减少信息量,即将字符的笔道宽度减少为一个模式元素的宽度,这样剩下的是原字符的骨架。

例如,汽车牌照的识别、邮政信件的分拣都需要滤去背景、干扰、噪声等信息,定位分离出不同的数字,还需要平滑、细化和边缘跟踪检测,这些都属于图像处理的范畴。这一过程输入是图像,输出也是图像。

2) 图像分割

为了从图像中找到需要识别的物体,还要对图像进行分割,也就是定位和分离出不同的待识别物体。这一过程输入是整幅图像,输出是像元图像。

3) 图像特征抽取

在需要识别的物体被分割出来的基础上,提取需要的特征,并对某些参数进行计算、测量,根据测量结果进行分类。例如,车牌照、邮政信件的数字定位分割之后,对每个数字进行特征提取。这一过程输入是庞大的信息图像,输出则是少量的特征信息,而且这些特征信息仅仅代表物体,无法还原回原物体。

4) 图像分类

根据提取的特征值,利用模式识别的方法进行分类,确定类别名称,以便对图像的重要信息得到一种理解和解释。这一过程输入是特征信息,输出是类别名称。

图像识别过程如图 1-2 所示。

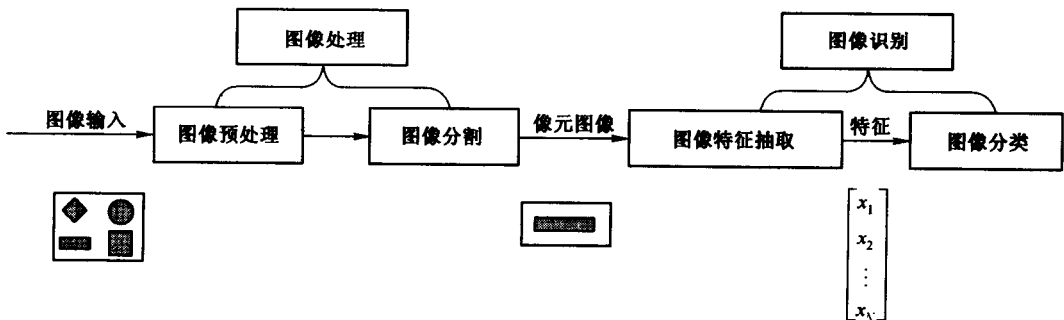


图 1-2 图像识别四个主要步骤

一个汽车车牌识别的全过程如图 1-3 所示。

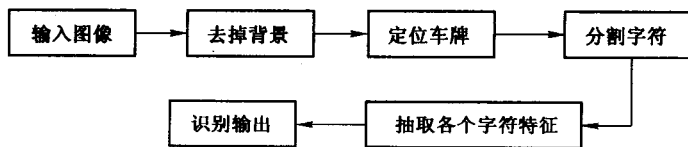


图 1-3 汽车车牌识别过程

由此可见,如果模式识别的任务来源是图像,这样的识别系统称为图像识别。

2. 图像识别的应用

图像识别广泛应用于机械、冶金、勘探、农业、造林、渔业、天文气象、医务、邮电、交通、公安、财务等部门,以及许多工矿企业中。到目前为止,图像识别受到广泛的重视,世界各国已经研制成多种多样的自动识别机和有视觉的机器人。这里例举出图像识别在以下方面的应用。

① 机器视觉识别:通过相机捕捉图像,分析图像的描述信息,作为自动化视觉检验或装配线的自动化。

② 文字和字符识别:信函分拣,文件处理,卡片输入,稿件输入,支票查对,期刊阅读,自动排版等。

③ 图形识别:遥感和航空照片分析,指纹、唇纹和面貌辨认,X光,显微图像,热成象及超声图像检查等。

④ 声音识别:语音识别和鉴定,侦听和机器故障判断等。

⑤ 生物医学应用:疾病诊断,癌细胞、白血球、染色体检查,修复手术控制设计等。

⑥ 工业应用:产品质量检验,集成电路设计,自动键合,图形设计等。

⑦ 预报问题:天气预报,工业烟雾预报,地震预报,经济预报等。

1.3 位图基础

1.3.1 数字图像的基本概念

计算机内的数字图像通常由采样点的值所组成的矩阵来表示:

$$\begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0,M-1) \\ f(1,0) & f(1,1) & \cdots & f(1,M-1) \\ \vdots & \vdots & & \vdots \\ f(N-1,0) & f(N-1,1) & \cdots & f(N-1,M-1) \end{bmatrix}$$

每一个采样单元叫做一个像素(pixel),上式中, M, N 分别为数字图像在横、纵方向上的像素总数。图像文件按其数字图像格式的不同一般具有不同的扩展名。最常见的图像格式是位图格式,其文件名以BMP为扩展名。

数字图像的颜色深度表示每一像素的颜色值所占的二进制位数。颜色深度越大则能表示的颜色数目越多。颜色深度的不同,就产生不同种类的图像文件,在计算机中常使用如下类型的图像文件。

1. 单色图像

单色图像中每个像素点仅占一位,其值只有“0”或“1”,“0”代表黑,“1”代表白或相反。

因为图像中的每个像素仅需 1 位信息,常把单色图像称为是 1 位图像。单色图像具有比较简单的格式,一般由黑色区域和白色区域组成。

2. 灰度图像

灰度图像具有如下特征。

① 灰度图像的存储文件带有图像颜色表,此颜色表共有 256 项,图像颜色表中每一表项由红、绿、蓝颜色分量组成,并且红、绿、蓝颜色分量值都相等,即

$$\text{fred}(x, y) = \text{fgreen}(x, y) = \text{fblue}(x, y)$$

② 每个像素由 8 位组成,其值范围从 0~255,表示 256 种不同的灰度级。每个像素的像素值 $f(x, y)$ 是图像颜色表的表项入口地址。例如,某灰度图像(16×6)对应的数值矩阵如下:

```
125,153,158,157,127,70,103,120,129,144,144,150,150,147,150,160,
133,154,158,100,116,120,97,74,54,74,118,146,148,150,145,157,
155,163,95,112,123,101,137,108,81,71,63,81,137,142,146,152,
167,69,85,59,65,43,85,34,69,78,104,101,117,132,134,149,
54,46,38,44,38,36,44,36,25,48,115,113,114,124,135,152,
58,30,44,35,28,69,144,147,57,60,93,106,119,124,131,144
```

3. 伪彩色图像

伪彩色图像与灰度图像相似,其存储文件中也带有图像颜色表,伪彩色图像具有如下特征。

① 图像颜色表中的红、绿、蓝颜色分量值不全相等,即

$$\text{fred}(x, y) \neq \text{fgreen}(x, y) \neq \text{fblue}(x, y)$$

② 整幅图像仅有 256 种颜色,要表示 256 种不同的颜色,像素必须由 8 位组成,每个像素值不是由每个基色分量的数值直接决定,而是把像素值当作图像颜色表的表项入口地址。把具有 256 色的图像称为 8 位彩色图像。256 色图像有照片效果,比较真实。

4. 24 位真彩色图像

具有全彩色照片表达能力的图像为 24 位彩色图像,24 位真彩色图像存储文件中不带有图像颜色表,其具有如下特征。

① 图像中每一像素由 R,G,B 三个分量组成,每个分量各占 8 位,每个像素需 24 位。

② $\text{fred}(x, y)$, $\text{fgreen}(x, y)$, $\text{fblue}(x, y)$ 取值范围为 0~255。

例如,某彩色图像(6×8)对应的数值矩阵如下:

```
(207,137,130) (220,179,163) (215,169,161) (210,179,172) (210,179,172) (207,154,146)
(217,124,121) (215,169,161) (216,179,170) (216,179,170) (207,137,120) (159,51,71)
(213,142,135) (216,179,170) (221,184,170) (190,89,89) (204,109,113) (204,115,118)
(216,179,170) (220,188,176) (190,77,84) (206,95,97) (217,113,113) (189,85,97)
(222,192,179) (150,54,71) (177,65,73) (145,39,65) (150,47,67) (112,20,56)
(136,38,65) (112,20,56) (112,20,56) (109,30,65) (112,20,56) (95,19,64)
(136,38,65) (91,11,56) (113,25,60) (103,19,59) (81,12,59) (126,62,94)
(138,46,71) (103,19,59) (158,65,83) (124,40,70) (145,62,79) (130,46,73)
```


由于24位真彩色图像所需的存储空间很大,处理速度较慢,当需要存储空间不大,并且要求实时快速处理图像时,一般都要利用相应的8位位图对其进行近似处理,因此8位位图是图像技术中涉及范围比较广泛的一种图像表示方法。

为了使实例中的代码简洁,本书中打开处理的位图文件都是针对8位位图。

1.3.2 BMP 文件结构

1. BMP 文件组成

8位位图(BMP)文件由文件头、位图信息头、颜色信息和图像数据四部分组成,结构如图1-4所示。

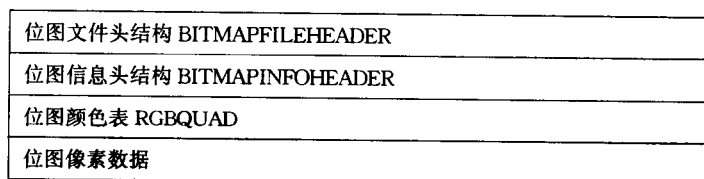


图 1-4 位图结构

2. 位图文件头

位图文件头结构含有 BMP 文件的类型、文件大小和位图起始位置等信息。其结构定义如下。

```
typedef struct tagBITMAPFILEHEADER
{
    WORD bfType;           // 位图文件的类型,必须为 BMP
    DWORD bfSize;         // 位图文件的大小,以字节为单位
    WORD bfReserved1;     // 位图文件保留字,必须为 0
    WORD bfReserved2;     // 位图文件保留字,必须为 0
    DWORD bfOffBits;      // 位图数据的起始位置,以相对于位图文件头的偏移量表示,
                          // 以字节为单位
} BITMAPFILEHEADER;
```

3. 位图信息头

BMP 位图信息头结构用于说明位图的尺寸等信息。其结构定义如下。

```
typedef struct tagBITMAPINFOHEADER
{
    DWORD biSize;         // 本结构所占字节数
    LONG biWidth;         // 位图的宽度,以像素为单位
    LONG biHeight;        // 位图的高度,以像素为单位
    WORD biPlanes;        // 目标设备的级别,必须为 1
    WORD biBitCount;      // 每个像素所需的位数,必须是 1(双色),4(16 色),8(256 色)
                          // 或 24(真彩色)之一
    DWORD biCompression; // 位图压缩类型,必须是 0(不压缩),1(BI_RLE8 压缩类型)或
                          // 2(BI_RLE4 压缩类型)之一
```