

# 模式识别导论

沈清汤霖 编著

国防科技大学出版社

## 内 容 简 介

本书兼顾计算机模式识别的基础与发展两个侧面,从统计模式识别、结构模式识别、模糊模式识别与智能模式识别这样四个方面,系统地有机地组织了全书的内容,形成了有别于以往同类教科书的独特体系结构。书中一方面展开了诸如线性分类器、Bayes决策、串文法、自动机等模式识别基础知识的介绍,并力图在数学推导的同时,尽可能引入其物理内涵以加深理解与记忆,另一方面又引入了模糊隶属度函数的建立、基于推理的知识表示与目标搜索、人工神经网络中的BP算法与Hopfield算法等内容,以及它们在模式识别中的应用,以使读者从中得到启迪,并为继续探索作好准备。

本书是大专院校计算机科学与工程专业学生、研究生教材,亦适合于计算机应用单位的广大科技人员参考。书中有较丰富的例题与图表,有助于阅读,各章之后附有习题,可供自我测试。

JS44/126

## 模 式 识 别 导 论

沈清 汤霖 编著

责任编辑 张序君 何晋

封面设计 陆荣斌

责任校对 朱淑娥

\*

国防科技大学出版社出版发行

新华书店总店科技发行所经销

国防科技大学印刷厂印装

\*

开本: 787×1092 1/16 印张: 18.625 字数: 430千

1991年5月第1版第1次印刷 印数: 1—4 100册

ISBN 7-81024-150-8

TP·29 定价: 4.80元

## 序

模式识别是人工智能的一个重要分支，它已发展成为一门独立的新学科，成为当代高科技研究和应用的重要领域之一。在过去三十年中，模式识别研究获得了重大的进展，在国民经济、国防建设、社会发展和人民生活的各个方面得到广泛应用，产生并正在继续产生十分深远的影响。正如国务委员兼国家科委主任宋健教授三年前给我的来信中指出：“以人工智能和模式识别为带头的这门新学科，将为人类迈进智能化时期作出奠基性贡献。”

沈清和汤霖编著的《模式识别导论》一书，涉及模式识别学科的各个主要研究问题，反映了模式识别研究和应用的最新进展，适应了进一步发展我国模式识别技术和相关课程教学的需要。本人有幸先读该书初稿，受益匪浅。读后觉得本书具有下列几个特点：

一、综合性 该书内容比较系统和全面，不仅包含了统计模式识别和结构模式识别等“古典”部分，而且涉及了模糊模式识别和智能模式识别等最新研究方面。这样，读者能够对模式识别的内容有个系统的和全面的了解。

二、实践性 本书取材既着重比较系统的理论阐述，又注意技术方法和应用实例的介绍，并用实例来说明基本原理。这就能够使有关科技人员和学生会如何处理工业、医学、军事和其它部门提出的实际问题。

三、新颖性 早在1983—1985年我首次赴美国从事人工智能研究时，美国国家工程科学院院士、国际模式识别和人工智能的杰出奠基者和先驱、国际模式识别学会首届主席、我的导师普渡大学(Purdue University)教授傅京孙(King-Sun Fu)博士就曾反复强调过：“模式识别必须与人工智能相结合，才能有更大的发展。”在过去几年中，模式识别所取得的许多新的重要进展，证明了傅先生这一论断的正确性。本书以清晰和较易理解的方法向读者介绍了这些新的进展成果及其实质，它不仅综合了模糊模式识别和逻辑推理法智能模式识别，而且发展了神经网络法智能模式识别。这一新内容，是迄今出版过的模式识别教材中所未曾见到的。

四、可读性 本书的可读性较好。作者尽力将数学推导与物理内涵有机地结合起来，既保持了必要的数学推演，又加强了物理意义的论述，便于读者学习与理解。

由于本书具有上述诸特点，因此，我相信，她的正式出版和发行，必将为广大读者提供一本最新的和有价值的模式识别读物和教材，并为活跃学术空气和促进我国模式识别学科的发展作出应有贡献。

联合国专家，中国人工智能学会理事  
中国计算机视觉与智能控制学会理事

蔡自兴

1991年4月15日

于中南工业大学

## 前 言

计算机硬件的迅速发展, 计算机应用领域的不断开拓, 急切地要求计算机能更有效地感知诸如声音、文字、图象、温度、震动等人类赖以发展自身、改造环境且能运用自如的信息资料。但就一般意义来说, 目前计算机却无法直接感知它们。计算机对外部世界感知能力的低下, 成为开拓计算机应用的狭窄瓶颈。这一问题已日益引起计算机科学工作者的密切关注, 并投入了大量人力、物力从事计算机模式识别的理论与应用的研究。其中在图象目标的检测与识别, 语音、图形、文字的识别, 地质探矿, 指纹鉴别, 机器人等领域取得了可喜的成果。

不久以前, 用于模式识别的方法, 还局限于统计模式识别与结构模式识别这两大类。近年来, 迅速发展的模糊数学已深入到模式识别的许多环节, 并出现了模糊模式识别这样的提法。不过在本书中, 作者还首次提出了智能模式识别这一名词, 并从逻辑推理与人工神经网络这两方面来探讨其必要性与实现途径, 力图使它与上述三种方法, 共同构成支持模式识别大厦的四根支柱。智能模式识别是嫁接于人工智能这一母株上的新枝, 作者深信随着其母株的不断成长, 不断输送营养, 必将结出丰硕的成果。

作者从事模式识别的教学与研究多年, 深感讲授这一课程的难度, 因为一方面它涉及数学与计算机科学, 如概率论、线性代数、模糊数学、离散数学、图象处理、形式语言与专家系统等学科, 另一方面它又是一门迅速发展中的学科, 新的研究成果不断充实着它的内容。鉴于此, 我们试图从基础与发展两方面展开, 即一方面是那些经典的内容, 如线性分类器、Bayes 决策、串文法、自动机等必不可少的基础知识, 并力图在数学推导的同时, 尽可能地引入其物理内涵。另一方面是那些大有希望的发展方向, 即前述模糊模式识别与智能模式识别方面的内容, 希望读者能从中得到启迪并继续探索。

本书第一章介绍模式识别的概况并在模式采集与预处理两方面涉及较多。第二、三、四章是统计模式识别的传统内容。第二章主要是相似性度量与几何分类法。第三章是以概率为基础的统计分类法与非参数技术。第四章则是聚类。第五章是介绍模糊数学及其在模式识别上的应用, 重点是隶属度函数的设计方法。第六章是结构模式识别。第七、八两章则是迄今为止的模式识别教材中均未涉及的, 人工智能技术与模式识别的结合。附录中的霍夫曼编码表则是图形数据转换中的常用标准。

本书第一章语音识别部分及第二章由汤霖编写, 其余部分由沈清编写。王兵山教授与胡运发副教授在百忙中, 分别审阅了本书第六、七两章。刘江宁、戴葵、卢春兰、胡晓敏等同志在本书的形成中, 提供了宝贵的帮助, 我们在此一并致以深切的谢意。蔡自兴、李树祥、胡守仁、谷士文, 王润生等五位教授、专家, 他们对本书作了客观的评审, 指出了本书的长处与不足, 使本书得以出版与提高, 其中, 蔡自兴教授还在百忙中为本书作序。对于他们的帮助与鼓励, 我们表示衷心的感谢。

由于我们的水平有限, 书中错误及不当之处, 在所难免, 敬请读者指正。

编著者

1990年11月

# 目 录

## 第一章 引言

§1.1 模式识别与模式识别的目的	1
§1.2 模式识别的全过程	3
§1.3 模式信息的采集	5
§1.3.1 地面景物图象的获取	5
§1.3.2 图形、文字信息的获取	6
§1.3.3 语音信息的获取	10
§1.4 模式识别前的预处理	12
§1.4.1 遥感图象的预处理	12
§1.4.2 图形、文字的预处理	19
§1.4.3 语音识别的预处理	23

## 第二章 统计模式识别(一)——几何分类法

§2.1 统计分类的基本思想	26
§2.2 模板匹配法及其数学描述	27
§2.2.1 光学模板匹配	27
§2.2.2 电子模板匹配(模拟灰度)	28
§2.2.3 电子模板匹配(数字灰度)	29
§2.2.4 模式匹配的实现	30
§2.3 模式的相似性度量及距离分类法	30
§2.3.1 相似性度量	30
§2.3.2 距离函数	31
§2.3.3 标准样本的距离分类器	33
§2.3.4 分散样本的距离分类器	34
§2.4 几何分类法(线性可分时)	36
§2.4.1 线性判别函数	36
§2.4.2 广义线性判别函数	40
§2.4.3 感知器算法	41
§2.4.4 LMSE 算法	46
§2.4.5 线性分类器用于多类问题	49
§2.5 几何分类法(非线性可分时)	51
§2.5.1 势函数的选择	51
§2.5.2 势函数算法的迭代训练	52

## 第三章 统计模式识别(二)——概率分类法

§3.1 统计分类的判别准则	56
§3.1.1 Bayes 法则	56
§3.1.2 Bayes 风险	57
§3.1.3 基于 Bayes 法则的分类器	58

§ 3.1.4	最小最大决策	60
§ 3.1.5	Neyman-person决策	62
§ 3.1.6	离散情况下的BaYes判决	63
§ 3.1.7	复合BaYes决策理论	64
§ 3.2	正态密度及其判别函数	65
§ 3.2.1	正态密度函数	65
§ 3.2.2	正态分布样品的判决函数	86
§ 3.3	密度函数的估计	70
§ 3.3.1	最大似然估计	71
§ 3.3.2	BaYes估计	74
§ 3.4	统计量的充分性	77
§ 3.4.1	因子化定理	78
§ 3.4.2	充分统计量与指数族	79
§ 3.5	非参数方法	80
§ 3.5.1	概率密度的估计	82
§ 3.5.2	Parzen窗估计法	83
§ 3.5.3	近邻估计	88
§ 3.6	分类错误率问题	93
§ 3.6.1	特征维数与错误率的关系	93
§ 3.6.2	协方差矩阵的估计	94
§ 3.6.3	一个分割平面的能力	95
§ 3.6.4	平均错误率问题	96
§ 3.6.5	错误率的估计	98
§ 3.7	降低特征的维数	99
<b>第四章 聚类分析</b>		
§ 4.1	聚类的基础	106
§ 4.2	基于试探的聚类算法	109
§ 4.2.1	基于最邻近规则的试探法	109
§ 4.2.2	最大最小距离算法	109
§ 4.3	层次聚类算法	110
§ 4.4	动态聚类法	112
§ 4.4.1	K均值算法	112
§ 4.4.2	迭代自组织的数据分析算法	113
§ 4.5	概念合取聚类法	118
§ 4.5.1	模式类的表示	118
§ 4.5.2	最佳分类的评判	119
§ 4.5.3	算法实现	119
§ 4.6	最小张树分类法	124
§ 4.6.1	基本概念	124
§ 4.6.2	分类法	124
<b>第五章 模糊模式识别</b>		
§ 5.1	综述	126

§5.2 模糊子集	127
§5.2.1 模糊子集的定义	127
§5.2.2 隶属函数的确定	128
§5.2.3 模糊子集的运算	136
§5.3 模糊关系	141
§5.3.1 模糊关系的性质及其建立	141
§5.3.2 模糊矩阵的运算	144
§5.3.3 模式分类的隶属原则与择近原则	149
§5.3.4 基于模糊等价关系的模式分类	152
§5.3.5 基于模糊相似关系的分类	154

## 第六章 结构模式识别

§6.1 概述	157
§6.2 结构模式识别系统	158
§6.3 模式基元的选择与抽取	159
§6.3.1 强调边界或骨架的基元选择	161
§6.3.2 强调区域的基元选择	162
§6.3.3 模式基元的抽取	163
§6.4 模式文法	166
§6.4.1 串文法	166
§6.4.2 扩展的串文法	172
§6.4.3 阵列文法	174
§6.4.4 树文法	175
§6.4.5 网文法	177
§6.4.6 图文法	178
§6.5 串的识别与分析	182
§6.5.1 有限自动机	183
§6.5.2 非确定的下推自动机	188
§6.5.3 导出树	192
§6.5.4 CYK 算法	193
§6.5.5 Early 算法	195
§6.5.6 转移图法	199

## 第七章 智能模式识别 (一) —— 逻辑推理法

§7.1 研究智能模式识别的必要性	202
§7.2 知识表示方法	204
§7.2.1 谓词逻辑表示法	205
§7.2.2 产生式表示法	209
§7.2.3 语义网络表示法	213
§7.2.4 框架表示法	215
§7.2.5 过程表示法	217
§7.2.6 知识表示的发展动向	220
§7.3 基于知识的推理	220
§7.3.1 求解问题的重要手段——推理	221

§ 7.3.2	问题的变换与分解	223
§ 7.3.3	基本搜索方法	225
§ 7.3.4	基于规则的推理	229
§ 7.3.5	语义网络推理	232
§ 7.4	知识的获取	234
§ 7.4.1	人工地获取知识	235
§ 7.4.2	自动地获取知识	237
§ 7.5	智能模式识别实例	239
§ 7.5.1	汉字的知识性表示	239
§ 7.5.2	获取待识汉字的符号性表述	242
§ 7.5.3	基于框架结构的推理	247
§ 7.5.4	规则的添加	248
§ 7.5.5	小结	249
<b>第八章 智能模式识别(二) ——神经网络法</b>		
§ 8.1	大脑神经元的构成及其机理	251
§ 8.2	人工神经网络概述	254
§ 8.2.1	基本模型	254
§ 8.2.2	与传统模式分类器的对比	255
§ 8.3	BP模型及其在模式识别上的应用	267
§ 8.3.1	BP模型的背景	267
§ 8.3.2	BP模型的算法及其特点	269
§ 8.3.3	应用实例	271
§ 8.4	Hopfield模型及其在模式识别上的应用	273
§ 8.4.1	Hopfield网络模型及算法	273
§ 8.4.2	应用	276
§ 8.5	其它神经网络模型及其在模式识别中的应用	278
§ 8.5.1	高阶关联神经网络	279
§ 8.5.2	其它几种神经网络	281
附录 A	改进的霍夫曼编码	285
附录 B	逻辑运算公式	286
主要参考文献		288



# 第一章 引言

## § 1.1 模式识别与模式识别的目的

按照广义的定义，模式是一些供模仿用的、完美无缺的标本。模式识别就是识别出特定客体所模仿的标本。识别能力是人类和其它生物的一种基本属性，根据被识别的客体的性质可以将识别活动分为两个主要类型：具体的客体与抽象的客体。字符、图画、音乐……是具体的客体，它们通过对感官的刺激而被识别；论点、思想、信仰……则是非物质的客体，对它们的研究主要属于哲学、政治学的范畴。我们主要是研究具体客体的识别，而且仅局限于研究用机器完成与识别任务有关的基本理论与实用技术，这一类课题属于工程学、计算机科学、应用数学的范畴。而研究人类和其它生物的识别能力，虽然也属于模式识别这一大课题，但却属于生物学、心理学、生理学的范畴。针对本书将要讨论的内容，我们对模式、模式识别作如下狭义的定义：模式是对某些感兴趣的客体的定量的或结构的描述，模式类是具有某些共同特性的模式的集合。模式识别是研究一种自动技术，依靠这种技术，机器将自动地（或人尽量少地干涉）把待识模式分配到各自的模式类中去。

模式识别是伴随着计算机的研究、应用日益发展起来的。从第一台计算机诞生到现在，已有四十多年了，在这期间，计算机不断更新换代，运算速度不断加快，内外存容量不断加大，各种外部设备不断增添，计算对象已由单纯数值计算发展到许多非数值领域，各种大型的管理系统、控制系统、检索咨询系统……纷纷投入运行，而且其应用领域还在不断扩大。人们亲切地将计算机称之为“电脑”，几乎所有本来由人脑实现的功能，都试图用“电脑”来完成，并不断取得令人振奋的成就。但比起人脑来，“电脑”毕竟是小巫见大巫，不仅仅在于人脑约有 $10^{11} \sim 10^{12}$ 个脑细胞（被称为神经元），更在于每个神经元约有 $10^3 \sim 10^4$ 个突触，亦即一神经元可通过突触与8000个别的神经元交换信息，当生物电流通过某一突触，甲神经元就将信息传送到乙神经元。所以人脑有极丰富的联想能力。可以超越时空，任意跳跃，传说中的孙悟空一个跟头可达十万八千里，已经很了不起，但人的思维可在一瞬间从地球跳跃到外银河系，又何止十万八千里？！不仅在联想、判断、推理能力等许多方面电脑远远不及人脑，特别是在对外界信息的感知能力方面，更远远不如人脑。作为群体的人类，生活在自然界中，作为个体的人，生活在自然界与社会中，每日每时不断感知外界的信息，并作出相应的反应。一只老虎与一只小猫外形颇为相似，所做的动作也颇相似，但前者使人惧怕，本能地设法自卫；后者使人感到亲切，想去抚摸一番。一声严厉的呵斥，一声亲切的呼唤，可能频谱与强度都相差无几，但却引起你截然不同的感受。在从事模式识别工作的人看来，客观事物（景物、图象、图形、语言、文字、电压、光强……）全都以不同的模式向人传送

不同的信息，正是区分了这些不同的模式，人才感知了这许许多多不同的信息。在陌生人看来似乎是一模一样的孪生兄弟，但他们的父母必然能区分谁是哥哥，谁是弟弟。但目前，“可怜的”电脑，连老虎和猫都区分不开，更别说是孪生兄弟中的哥哥与弟弟了。在一次人工智能国际会议上，日本学者曾展出了一个脸谱识别器，表演时能把几个日本人的脸谱识别出来，叫出他们的名字。有些外国代表也想试一试，但计算机识别不出来，显示“不是人”，引起哄堂大笑。当然从那次会到现在已过去好多年了，尽管可以将“不是人”改为“不认识”，但至今尚无质的飞跃。识别脸谱、指纹、签名等都是很有用的课题，但还需要作很大的努力。我们模式识别工作者就是要努力教会电脑正确识别自然界，人类社会中亿万种不同的模式，不仅要教会电脑识别老虎和猫，而且最好能迅捷、准确地区分孪生兄弟中的哥哥与弟弟。电脑的识别能力越强，它就越能更好地为人类服务。

而且目前计算机的内部处理能力与对外界的感知能力之间，也日益表现出尖锐的矛盾。这当然不是说计算机的处理能力已尽善尽美了，但相比之下，计算机与外界的信息交换能力却太可怜了。例如作为计算机的大容量外存储设备：磁盘、光盘已有了长足的发展，整台磁盘机存储容量达数百~数千兆字节，薄薄的一张光盘（直径5.25英寸）即可存储600兆字节信息。一个汉字只需2字节的存储容量，故可存300兆汉字，即可存40万字的书7500本，但这些文字怎样存入计算机呢？以目前的汉字编码键入速度计算（8小时键入25000字）约需50人年。这就需要开展汉字识别的研究来解决输入速度的问题。又如，由于电声技术的发展，人类的语言已能以相当高的保真度记录下来，但机器仅是如实的记录、再现而已，对于其含义并不十分理解，尽管世界上许多学者集中了巨大的精力，从事各国语言的识别，但至今仍未达到广泛应用的程度，其原因之一就是语音信息不仅包含声音的特征，而且包含着文法、语义，甚至感情、心理等因素。再如，大量统计表明：人的70%以上的信息是从眼睛中得来的。人的视觉机理是很复杂的，人看东西并不是把所有的信息都送到大脑去的，在眼睛周围就有许多“微型计算机”，把看到的東西进行分类、抽取、压缩，仅仅把重要的信息与大脑交换。在大脑中，对形状、运动、色彩的识别，都是由不同部分完成的。用现在计算机的术语来说，是一个分级控制的分布式的信息处理系统。特别重要的是：人是用先验知识，通过联想、比拟、推断才能对看到的東西作出决策的。但对于计算机来说，尤其是建立在诺依曼体系基础上的计算机来说，这实在是太难了。所以日本提出的第五代计算机是以突破诺依曼体系为前提的，即所谓智能型的计算机。诚然，日本五代机计划的结局，似乎并不理想，这也许是因为人们对此计划的期望过高。但五代机计划的实施，毕竟促成了人们的共识，即必须毫不懈怠地以更大的努力，继续展开人工智能各个领域（包括模式识别）的研究，以使计算机更好地为人类服务。

总而言之，研究和发展模式识别的目的，在于提高计算机的感知能力，从而大大开拓计算机的应用；而计算机感知能力的真正提高，不仅与模式识别这一学科本身有关，而且与数学、工程技术学以及计算机本身的体系结构，软硬件性能有关，我们应及时吸取兄弟学科的新成就，把握计算机科学的新发展，为我所用，为人类造福。

## § 1.2 模式识别的全过程

图 1.2.1 简要地表示了模式识别的全过程，但可以毫不夸大地说，其中每一阶段设计的好坏都会对全盘的工作产生严重的影响，所以每一阶段都应争取尽可能完美的效果。

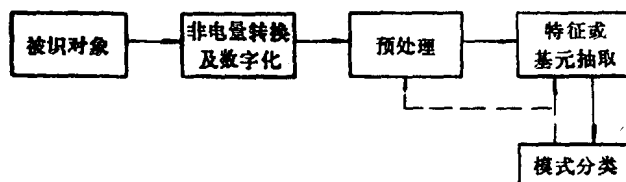


图 1.2.1 模式识别的全过程

由于被识别的对象多半是具有不同特征的非电量，如灰度、色彩、声音、压力、温度等等，所以第一步就要将它们转变为电信号。（若被识对象本身就是某种形式的电信号，则自然可省略这一步骤），然后经A/D变换，将它们转换为能由计算机处理的数字量。当然模式识别亦可不用计算机（例如光学的办法），这就不属本书讨论的范围了。

数字化后的电信号需经预处理，以滤除样品采集过程中参入的干扰、噪声，并人为地突出有用信号，以得到良好的识别效果，这将在后面予以讨论。

经改善后的有用信号，还要作特征抽取或基元抽取（视分类方法而定），才能对其分类。由于特征（或基元）的抽取与待识模式的类别密切相关，很难有某种泛泛的规律可依循。而且可以相信，通过以后几章的介绍，读者定能相应地了解到对特征（或基元）抽取的要求，再结合具体的待识对象，便可有针对性地完成此阶段的工作。但随着工作的深入，不断修改与完善此阶段的工作，不仅是可能的，从某种意义上说简直是必需的。这就是图1.2.1中虚线回溯的含义。

模式分类就是在前几步准备工作的基础上，把被识对象归并分类，确认其为哪种模式的过程。这是模式识别“出成果”的阶段，直接以其分类结果表明本次识别的结束。显然这是举足轻重的阶段，弄得不好，会使前几阶段的工作付之东流。

模式分类按其方法，大致可分为统计决策法、句法结构法、模糊判决法与人工智能法这样四大类。前两类方法有悠久的历史，发展得较成熟，在解决相应领域中的模式识别问题时，均有明显的效果，是模式分类的经典性与基础性技术。在这两类方法中引入模糊数学的研究成果，则往往大大改善分类的效果。当然，独立地运用模糊判决方法，也不乏模式分类的成功实例。特别值得指出的是：随着新一代计算机研制工作的进行，对人类逻辑思维能力模拟的巨大成功，专家系统与 PROLOG、LISP 语言的广泛应用，在模式识别中将传统的分类方法与基于逻辑推理的（狭义的）人工智能的各项研究成果密切结合，无疑会将这一工作推向新的高度，具有无比诱人的前景，我们对此应给予足够的重视。80年代再度活跃起来的人工神经网络的研究，更以崭新的姿态，以其全局相关的特色，在一系列模式识别领域中，取得了许多用传统方法所难以达到的，令人瞩目的成就，大大拓广了模式识别工作者的思路。但一般说来，人工神经网络的研究，仍属于对人类智能的模拟，亦即属于（广义的）人工智能的研究领域。尽管它更倾向于人类形象

表 1.2.1 几种识别方法的比较

识别方法 比较项目	统计决策	句法结构	模糊判决	逻辑推理	神经网络
主要理论 支柱	概率论、数理统计	形式语言、自动机技术	模糊数学	演绎逻辑、布尔代数	神经生理学, 心理学
模式描述 方法	特征向量: $X = (x_1, x_2, \dots, x_n)$	符号串、树、图	模糊集合 $A = \{(\mu_a, a), (\mu_b, b), \dots, (\mu_n, n)\}$	字符串表示的事实	以不同活跃度表示的 输入结点集
模式类判 定	是一个聚合类。用条件概率分布 $P(X/\omega_i)$ 表示, $m$ 类就有 $m$ 个分布, 然后判定未知模式属于那一个分布。	是一种语言。用一个文法表示一类, $m$ 类就有 $m$ 个文法, 然后判定未知模式遵循那个文法。	是一种集合运算。用隶属度将模糊集合划分为若干子集, $m$ 类就有 $m$ 个子集, 然后根据择近原则分类。	是一种布尔演算。从事实出发运用一系列规则, 推理得到不同结果, $m$ 个类就有 $m$ 个结果。	是一个非线性动态系统。通过对样本的学习建立起记忆, 然后将未知模式判决为其最接近的记忆。
主要方法	几何分类: 线性分类 非线性分类 统计分类: Bayes 决策 无教师的分类: 聚类分析	自动机技术 CYK 剖析算法 Earley 算法 转移图法	隶属度函数的设计: 模糊统计法 二元对比排序法 推理法 模糊集运算规则 模糊矩阵	产生式推理 语义网推理 框架推理	BP模型 HOP模型 高阶网
主要优缺点	优点: 比较成熟, 能考虑干扰、噪声等影响, 识别模式基元能力强。 缺点: 对结构复杂的模式抽取特征困难。不能反映模式的结构特征, 难以描述模式的性质。难以从整体角度考虑识别问题。	优点: 识别方便, 可从简单的基元开始, 由简至繁。能反映模式的结构特性, 能描述模式的性质, 对图象畸变的抗干扰能力较强。 缺点: 当存在干扰及噪声时, 抽取基元困难, 且易失误。	优点: 由于用隶属度函数作为样品与模板间相似程度的度量, 故往往能反映它们整体的与主要的特性, 从而允许样品有相当程度的干扰与畸变。 缺点: 准确合理的隶属度函数往往难以建立, 故限制了它的应用。	优点: 已建立了关于知识表示及组织, 目标搜索及匹配的完整体系。对需通过众多规则的推理达到识别目标确认的问题, 有很好的效果。 缺点: 当样品有缺损, 背景不清晰, 规则不明确甚至有歧义时, 效果不好。	优点: 可处理一些环境信息十分复杂, 背景知识不清楚, 推理规则不明确的问题。允许样品有较大的缺损、畸变。 缺点: 模型在不断丰富与完善中, 目前能识别的模式类还不够多。

思维能力的模拟，而不局限于逻辑思维能力，但它毕竟仍属于人工智能的范畴，故我们仍将它归类为解决模式分类的人工智能方法。

所有这些方法均将在后面的章节中深入介绍。此处我们先以表格的方式，对它们作一比较。见表1.2.1。

### § 1.3 模式信息的采集

在对模式识别的全过程作初步介绍后，本节介绍几种模式信息的采集方法。实际上这就是图1.2.1中的非电量转换及数字化。

#### §1.3.1 地面景物图象的获取

接受太阳能的各种地面物质，都在连续地反射各种电磁波。各种不同的物质，它所反射的电磁波强度也不一样。图 1.3.1 中以雪、湿地、小麦为例，给出了它们的反射率。

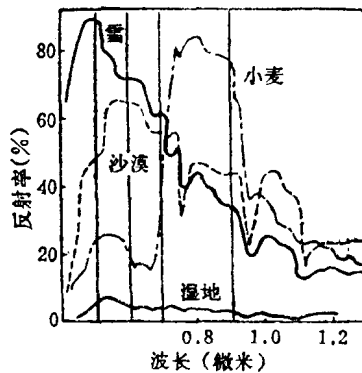


图 1.3.1 雪、湿地、小麦等物质的反射率

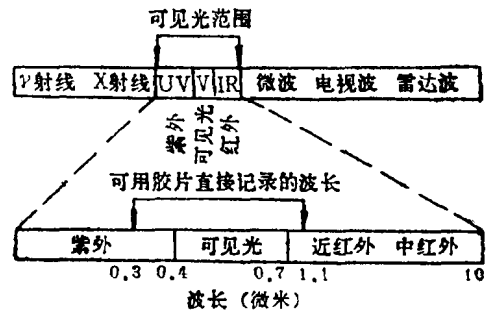


图 1.3.2 电磁波的反射光谱

采用全色胶片摄影时，在可见光区包含整个波长范围的反射能量积分值，是以黑白色调的差别来记录的。彩色摄影分为蓝绿红三个不同波段合成的感光层，用这种方法获得的图象，各波段合成后可以用自然色调表现出来。多光谱摄影则是以不同波长的反射能量来记录的。根据需要，选择其中某一波段的图象(或几个波段图象进行彩色合成)，并进而用假色彩技术表现它们，则更易于得到不同物质物体的轮廓。如图 1.3.1 所示的三种物质，若分别在0.5微米和0.8微米波段采录，就会得到很好的效果。

除了地面摄影外，还可在飞机，卫星上从高空对地球表面摄影，以获得解决不同问题所需要的信息：如地形测量、资源开发、植物保护……这些都是近年涌现出来的热门课题。采用航空摄影，除了获得地表的几何信息外，还可获得自可见光区至近红外区各个波段的光谱辐射强度。

使用单个摄影机，选用适当的滤光法和胶片，可对一个特定波长范围的目标的辐射能量进行记录。使用多台摄影机，对同一目标采用多波段摄影，称为多光谱摄影，现一般是对蓝、绿、红、红外四个波段同时摄影，再分别合成，以期达到特定的分辨效果。

采用摄影图象方式可记录的波长范围，受限于胶片的性能，故而较窄，见图1.3.2。

但从可见光区到近红外区的图象判读，有悠久的历史，已成为比较完整的技术系统。各种型式的红外胶片、彩色胶片、红外假彩色胶片、以及二色胶片等不断问世，不断扩展着图象可记录的波段范围。

除摄影图象外，还有非摄影图象方式，这种方式是基于以下目的而研制的：

- 图象信息的实时传输；
- 覆盖宽的波长区域；
- 后继的自动分类。

这就是说希望不用胶片作为传送图象的媒介，因为胶片只能事后回收、显影，再用光扫描鼓、飞点扫描管，……等转换成电信号，处理很麻烦。

非摄影图象方式的选择之一，是采用电视摄像管（其构造见下节），采录到的电视信息，不仅能远距离传输，由计算机即时处理，且能在显象管上显示出来，实时地观察或合成，用录象磁带进行的记录与还原也很方便。非摄影图象方式的另一种选择是用旋转扫描镜以固定的现场对目标扫描，然后将接收来的电磁波聚焦在接收元件上。例如红外成象装置就是一例。由旋转扫描镜接收的红外辐射能量，通过由抛物镜组成的光学系统，聚焦在红外接收器的阵列面上，经过接收元件，转换为电信号，此种电信号可由长余辉显象管显示，以录象带记录或远距离传输。改变接收元件的光电特性，可得到从紫外到红外各波段的图象。

还有一种多光谱图象扫描仪（MSS），它把由旋转式扫描镜接收到的电磁波，通过棱镜或光栅等分光元件进行分光，再以各种不同波长的接收元件同时取出电信号，分别记录到磁带上。如地球资源技术卫星2号上装有五个波段（0.5~0.6，0.6~0.7、0.7~0.8，0.8~1.1，10.4~12.6微米）的多光谱扫描仪，连续不断地向地面发送各种图象信息。

上述成象方式，均是探测自然界的物体所辐射的能量强度，是一种被动的方式。相对而言各种成象雷达的出现及其所获得图象的判读，则不仅扩展了探测的波段，而且进入了一种主动的方式，即所接收的是由人工发展的能源，经被测物体反射回来的能量强度，它对阴晴雨雪等因素也很不敏感。其中显示出很大威力和受到很高评价的是侧视雷达（SLAR）。

侧视雷达是从飞机装载的天线上，向飞行方向一侧的斜方发射出扇形的波束，再接收从地表反射回来的微波，故信号的发射与接收是反复交替式进行的，它随着飞机的飞行，沿其轨迹覆盖了一相当宽阔的区域。从地表反射回来的微波能量，是以时间与振幅的相关而获得的。经过阴极射线管的变换，再依次地记录在同步传输的胶片上。根据飞行高度与探测装置的性能，在飞行的侧视方向可达到从1公里到100公里的范围，可实现速度快，探视面积广的要求。根据水平和垂直两个方向的偏振，可获得不同信息的图象。已有产品的工作波长为0.86厘米、3厘米。这类波长范围的微波不受云雨等方面影响，故是全天候的。

### §1.3.2 图形、文字信息的获取

本节所介绍的系统，可以将图形、文字这样的非电量（包括显影后的胶片）转换为电信号，经A/D变换后即可由计算机处理。

## 一、飞点扫描系统

图1.3.3示出一个飞点扫描系统。飞点扫描管是一种特殊的阴极射线管，工作原理和显像管一样，只是余辉时间极短，光点亮度很强，电子枪的聚焦质量好，光点细，分辨率高。

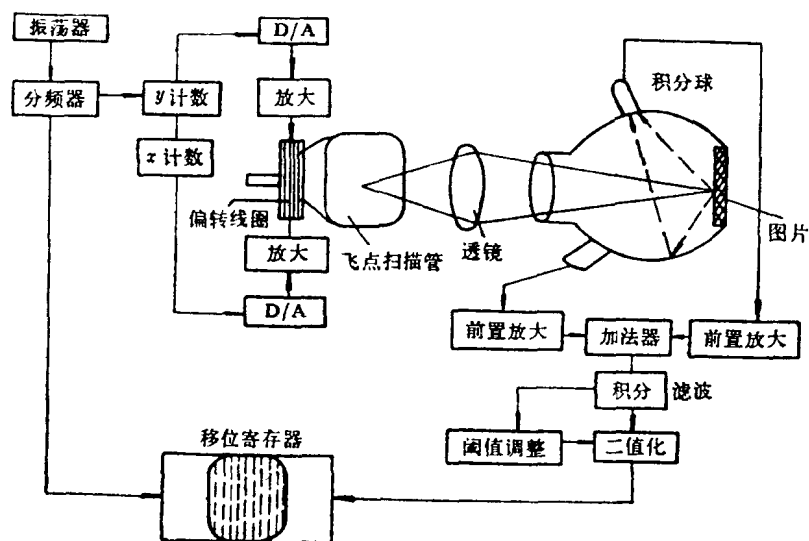


图 1.3.3 飞点扫描系统

振荡器的输出经分频器降低频率后，送往Y计数器和移位寄存器，Y计数器各位输出经D/A转换并放大后，供给飞点管Y方向的偏转电流。Y计数器每计满一列光点后，进位到X计数器进行计数，X计数器的各位输出经D/A转换并放大后，供给飞点管X方向的偏转电流。X、Y两方向的偏转电流波形，如图1.3.4所示。

光点上下左右扫描形成一点阵，并经过透镜在被识别对象（例如图片）上成象。图片漫反射的光由积分球收集起来。积分球是一个金属圆球，避光密封，内壁反光极好，它可将图片上漫反射的光线，经过一次或多次反射，射入光电倍增管接收，转换为模拟电信号。经前置放大后在加法器迭加，再经积分滤波网络，对随机噪声进行最佳滤波。去除干扰噪声后的模拟视频信号输入二值化部件，与阈值调整系统计算出来的阈值相比较，得到二值信号方阵，送至移位寄存器，移位脉冲也由振荡器提供，以保证与扫描光点同步。

飞点扫描应用上稳定可靠，扫描速度快，控制灵活，在反馈跟踪系统的控制下，有可能使光点沿图形或字符的边缘移动，即螺旋状扫描，这是其它采样设备所不具备的。

其缺点是使用不够方便，因为它是利用飞点管发出的光源射于图片上，所以不可使外界光源射到图片上。为此要求有能对图片全自动馈送的密闭机构，故而成本较高。

## 二、电视摄象系统

这是一种被广泛使用的采样设备，不仅用于图形、文字的采录，而且广泛用于图象、景物的采录。它利用摄象管的光电转换特性，将由外界进入摄象管的光通

量，转换成变化的电信号，再经A/D变换后送入计算机。

摄像管的简单原理如下（参见图1.3.5）：

摄像管主要由电子枪、线圈及光电靶三部分组成。电子枪由灯丝、阴极、控制栅极、加速极和聚焦极所组成。线圈有偏转、校正、聚焦线圈三部分。由于电子枪与线圈均与一般阴极射线管相似，故此处从简。

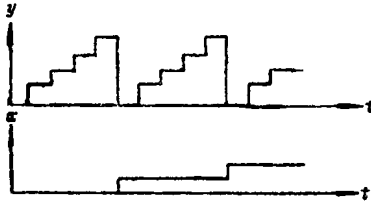


图 1.3.4 X、Y两方向偏转电流波形

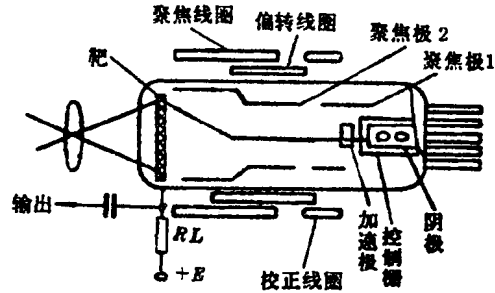


图 1.3.5 摄像管原理图

光电靶很像一块“夹心饼干”，中间是信号极，它是喷涂在玻璃板上的一层透明金属导电层，其特点是透明度高、导电性能好，其左侧（向着被摄取对象一侧）是玻璃板本身，其右侧是光电阴极，它由蒸镀在信号极上，具有内光电效应的一层材料所构成，这种材料当受光照射后亦发出光电子，并立即成为参加内部导电的自由电子，因而降低了材料的电阻率，降低量的大小与光通量成正比。该材料的量子输出可远大于1，所以光电转换效率高，但惰性较大。光电子然后通过信号极输出。

一个实际的电视摄像采样系统见图1.3.6所示。摄像管靶极输出的电信号经放大滤波后，并联地送入16级比较级左侧的输入端1，而比较级左侧的输入端2则加上参考电压经分压器分压后的电压，当输入端1的电压高于输入端2的电压时，在比较级右侧输出高电平，反之输出低电平。当比较级输出全为低电平时，编码器输出为零，当比较级输出全为高电平时，编码器输出为16进制F，这样，主机就能得到比例于图片（或景物）灰度的16级不同灰度值。一个准确的振荡器，控制着接口的输出，使之与电视摄像的行、帧信号同步，并抹去电视回扫时的输出。

### 三、电荷耦合器件

随着MOS工艺的提高，电荷耦合器件（CCD）已广泛用于图文的采录。图文采录设备成为计算机办公自动化的标准外设。不同档次的图文扫描仪和传真机由于采用了CCD后大大提高了质量，降低了成本。

CCD是一种MOS集成电路，目前已达到每毫米十余象素的分辨率。它是在P型单晶硅的衬底上生成一层 $\text{SiO}_2$ ，再在其上蒸发一层间距排列很小的铝电极，电极上加有适当的正偏压，该电压形成的电场穿过 $\text{SiO}_2$ 层，并排斥P型硅中的空穴，从而在电极下形成电荷耗尽区，而在 $\text{SiO}_2$ -Si的界面上得到一贮存电子的势阱。当硅晶体受光照而产生出电子空穴对时，电子就被吸入势阱中，投射光越强，势阱中捕获的电荷量越多，这



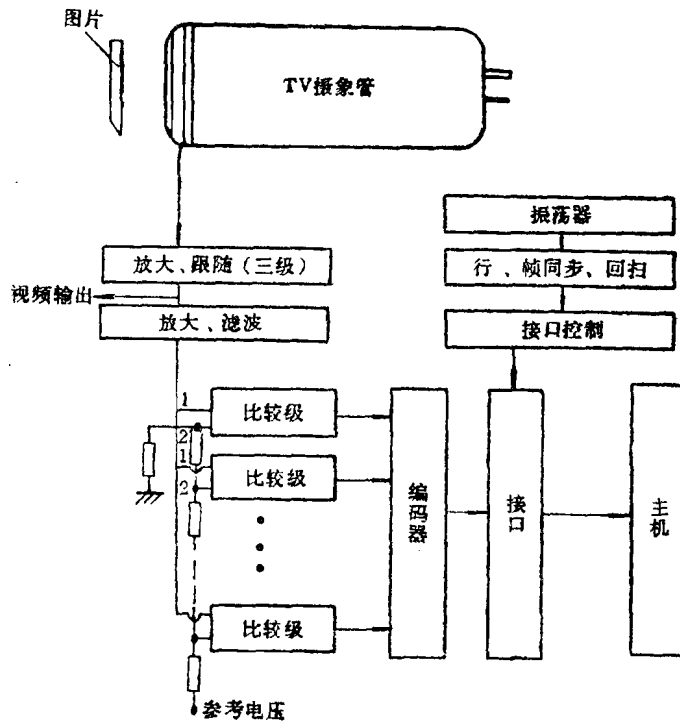


图 1.3.6 一个电视摄像采样系统

样就把物体的光像变为由积累电荷所描绘的电子图象，完成了光电转换与储存信息的过程。

为了读出CCD中储存的电子图象，可用移位寄存器的办法，即在减小本电极正偏压的同时，提高相邻电极的正偏压，使原来存储在势阱中的电荷，向相邻电极转移，实现一次移位。经多次移位后，便可将CCD中的电子图象移入计算机。

图1.3.7为电极上所加的电压波形及电极下电子的转移情况。

设某一时刻 $t_1$ ，这时 $v_1$ 为高电位， $v_2$ 、 $v_3$ 为低电位，则在 $v_1$ 电位下的电极形成一电场，当硅晶体受到光投射而产生电子空穴对时，电子在电场作用下被聚集在势阱中，投射光越强，聚集的电子越多。当 $t_2$ 、 $t_3$ 时刻到来， $v_2$ 、 $v_3$ 依次出现高电位，势阱中的电子依次向右迁移，故首先是右侧第1个像素的电子图象输出，到 $t_3 + T$ 后又有第2个像素的电子图象输出，经历 $2048T$ 后，全部2048个像素依次串行输出。恰在这时，柱状透镜旋转一小角度，将另一行图象投入CCD器件，如上述过程又输出另一行光电转换后的电子图象。

CCD器件有线式的（仅一行）和面式的（若干行构成一方阵）。已有2048单元线式CCD商品。这类器件要求每一组电极的功能都是合格的，故需严格筛选，也使面式CCD难以商品化。