

计算机模式识别

· 结构方法 ·



计 算 机 模 式 识 别

· 结构方法 ·

周冠雄 编 著
邹海明 主 审

华中工学院出版社

计算机模式识别 (结构方法)

周冠雄 编 著

责任编辑 唐元瑜

华中工学院出版社出版发行

(武昌喻家山)

新华书店湖北发行所经销

华中工学院出版社沔阳印刷厂印刷

开本: 787×1092 1/16 印张: 16.25 字数: 388,000

1987年4月第1版 1987年4月第1次印刷

印数: 1—3 000

ISBN 7-5609-0028-3/TP·5

统一书号: 15255-091 定价: 2.75元

前 言

本书是《计算机模式识别(统计方法)》(华中工学院出版社1986年出版)的姐妹篇。两本书一起构成计算机模式识别学科的一个基本体系。

全书共十章。第一章介绍有关的基本概念及基本思想方法;第二章论及文法及语言的知识,目的是使那些对形式语言理论并不了解的读者也能掌握这些知识,并建立起它们与结构模式识别学科之间的有机联系;第三章介绍模式描述语言,它是形式语言在模式识别应用中的进一步深化与具体化;第四章讨论基本模式元的选择与识别方法,虽然这部分内容属于数字图象处理及统计模式识别学科的范畴,但它又是实现结构模式识别系统的基本任务,因此,有必要在本书中也作概括的介绍,以便为读者说明解决该问题的思路及可能采取的方法;第五章讨论句法分析,介绍实现结构识别的基本分析方法;第六章论及文法推断,集中介绍构造式的文法推断技术;第七章介绍模式的字符串描述及树状描述的相似性度量及其误差校正分析;第八章专门讨论模式的图描述及图变换方法;第九章介绍如何综合应用模式的统计特征与结构特征,以实现一个混合分层的模式识别系统;第十章讨论结构模式识别方法的应用,考虑到结构方法在外形分析及波形分析中的应用已在参考文献〔4〕中有了较全面的讨论,故本章将集中介绍这一方法在纹理分析中的应用。

由于结构模式识别的内容十分庞杂,因此,在一个规定的有限篇幅内,实际上是无法作包罗万象的讨论的。所以,本书在内容的取舍上,力图突出那些最基本而且是行之有效的方法;在论述上力争通俗易懂,注重方法的完整性,使读者能加以应用;而对于那些未能完整地加以论述的重要结果,则列出相应的参考文献,以备读者查阅。

本书的第五章与第七章的初稿由秦友淑编写,其余各章由周冠雄编写。统编时,周冠雄对第五章与第七章作了若干内容的增删及文字的修改。邹海明教授主审全书。

华中工学院计算机系领导及该系模式识别与人工智能研究室的钟春香副教授、胡昌赤讲师及其他同志,对该书的编写工作给予了很大的帮助;外语系德法教研室的李梅讲师为该书的编写提供了许多德文文献的中译稿;华中工学院出版社在该书的编辑与出版过程中作了大量辛勤、细致的工作;在此一并致谢。

由于本人水平有限,本书错误与不妥之处在所难免,敬请同行专家及读者不吝赐教。

周 冠 雄

一九八六年三月

04-04-87

目 录

第一章 导 言

1.1 基本概念	(1)
1.2 关于结构方法与统计方法的比较	(5)
1.3 结构方法的基本思想及一般原则	(6)
1.4 结构模式识别系统	(9)
1.5 模式的结构描述方法举例	(10)

第二章 文法和语言

2.1 链文法	(13)
2.2 程序文法	(15)
2.3 随机文法	(16)
2.4 属性文法	(18)
2.5 带控制集的文法	(20)
2.6 散射前后文文法	(21)
2.7 随机前后文文法	(22)
2.8 标识文法	(23)
2.9 阵列文法	(25)
2.10 树文法	(27)
2.11 图文法	(30)

第三章 模式描述语言

3.1 一般模型	(34)
3.2 一个用于模式分类的图象描述语言	(36)
3.3 前后文无关的图象描述语言 (PDL)	(37)
3.4 PDL 应用于粒子物理图片的句法结构描述	(41)
3.5 一个化学结构描述语言 (CSL)	(44)
3.6 丛状图象描述语言 (PLEX 语言)	(47)
3.7 图象处理语言	(50)

第四章 基本模式元的选择及识别

4.1 一般方法及要点	(54)
4.2 基本模式元的选择对链文法的影响	(56)
4.3 基本模式元的抽取	(57)
4.3.1 一般方法	(57)
4.3.2 逐段线性近似方法	(60)
4.4 基本模式元的识别	(62)
4.4.1 最优统计分类器	(62)
4.4.2 线性分类器	(65)
4.4.3 可调线性分类器	(66)
4.4.4 非参数分类器	(68)

第五章 句法分析

5.1 关于句法分析问题的一般论述	(71)
5.2 自上而下的分析	(72)
5.2.1 带回溯的自上而下的分析算法	(72)
5.2.2 自上而下分析算法中的问题及其解决办法	(76)
5.2.3 不带回溯的自上而下的分析算法	(79)
5.3 自下而上的分析	(80)
5.3.1 归约与句柄	(80)
5.3.2 删剪语法树	(82)
5.4 CYK分析算法	(83)
5.5 Early分析算法	(86)
5.6 LR(k)文法	(89)
5.7 前后文无关程序语言的句法分析	(93)

第六章 文法推断

6.1 关于文法推断技术的一般论述	(96)
6.1.1 文法推断方法的基本要点	(96)
6.1.2 文法推断技术的历史沿革	(97)
6.2 链文法的推断	(98)
6.2.1 样本集及推断问题	(98)
6.2.2 正则文法的推断	(99)
6.2.3 前后文无关文法的推断	(106)

第七章 结构模式的相似性度量及分类

7.1 字符串的Levenshtein相似性度量	(124)
7.1.1 Levenshtein距离与加权Levenshtein距离	(124)
7.1.2 实现Levenshtein距离的动态规划方法	(126)
7.2 最小距离误差校正分析	(128)
7.3 字符串的Findler相似性度量及最小距离分类	(130)
7.4 树的相似性度量	(133)
7.4.1 树状模式的误差变换	(133)
7.4.2 两棵树之间的距离	(134)
7.5 树的误差校正分析	(137)
7.5.1 结构保持的误差校正树自动机(SPECTA)	(137)
7.5.2 广义误差校正树自动机(GECTA)	(143)
7.6 树的聚类分析	(148)
7.6.1 计算两棵树之间距离的一种算法	(148)
7.6.2 分级算法在树的聚类分析中的应用	(150)

第八章 结构模式识别中的图变换方法

8.1 图与模式	(153)
8.2 模式变形与图的变形	(154)
8.3 误差校正的图变换	(157)
8.4 确定最优图变换的研究式搜索方法	(160)
8.4.1 状态描述	(160)
8.4.2 运算符	(161)

8.4.3 损失	(162)
8.5 加速搜索时信息的利用 (一个确定最优图变换的具体搜索算法)	(164)
8.6 应用举例——几何图形的识别	(172)
第九章 统计与结构特征混合使用的识别系统	
9.1 混合分层的外形识别系统	(176)
9.1.1 系统描述	(177)
9.1.2 客体轮廓线的跟踪	(177)
9.1.3 利用方向编码计算Fourier展式系数特征	(178)
9.1.4 方向编码的跑长编码	(179)
9.1.5 近邻分类器的应用	(180)
9.1.6 结构分类器的实现	(180)
9.1.7 最大相似性度量	(181)
9.2 混合分层的数字符号识别系统	(181)
9.2.1 识别系统的构成	(182)
9.2.2 模式依据正交归一化函数系的展开及统计分类的实现(第一层)	(183)
9.2.3 基于模式拓扑结构分析的分类与分类假设的产生(第二层)	(184)
9.2.4 精简类别假设的方向元素特征检测方法(第三层)	(198)
9.2.5 最终有效分类的实现(第四层)	(208)
第十章 结构模式识别方法在纹理分析中的应用	
10.1 纹理、纹理图象与纹理分析	(222)
10.1.1 概念与定义	(222)
10.1.2 纹理分析的统计途径与结构途径	(223)
10.1.3 纹理分析系统的设计	(226)
10.2 纹理分析中的纯结构方法	(227)
10.2.1 以单个象素作为模式基元的方法	(228)
10.2.2 以波峰作为基元的方法	(230)
10.3 纹理分析的结构统计方法	(233)
10.3.1 以区域为基元的方法	(233)
10.3.2 铺瓦方法	(236)
10.3.3 纹理匹配模型方法	(239)
10.4 自上而下的纹理分析方法	(241)
10.4.1 纹理周期性向量的确定	(241)
10.4.2 纹理元素的抽取	(242)
参考文献	(244)

第一章 导 言

按照现代的观点,电子计算机已经不再单纯是数字计算的工具,而同时又是信息处理的工具,今后还将发展成为具有智能处理能力的系统。正是由于计算机技术的迅速发展,才使得模式识别在最近二十年中,迅速地成长为一门相对独立的,几乎在所有领域中都被广泛应用的学科。

目前,该学科在其三个分支(统计模式识别、结构模式识别及模糊模式识别)上,奠定理论基础的工作已基本完成,但作为学科的理论体系,其归纳整理的工作还在进行⁽¹⁾。按其历史发展的逻辑,首先获得充分发展的是统计模式识别的理论方法,因为它建筑在经典的决策理论的基础上,许多经典的数学方法也为它的发展提供了强有力的工具。它与计算机科学相结合,已经产生了自己的一整套关于模式的预处理、模式特征抽取、模式分类、学习以及识别系统的数学问题等专门的理论方法⁽²⁾⁽³⁾。结构模式识别的理论方法相对来说发展得较晚,因为它在形式语言理论充分发展的基础上才能获得自身的发展。但是,其发展速度,在以往的十多年中是惊人的。傅京荪教授作为这一学科的创始人,他的研究工作及一系列专著,不仅标志着该学科的发展里程,而且为该学科的理论方法作出了经典性的阐述⁽⁴⁾。模糊模式识别植根于模糊数学理论的基础上,是近几年来异常活跃的领域,目前已积累了大量的研究成果和应用经验。系统地总结这些经验和成果,完成建立学科体系的工作,现在已经是水到渠成的时候了。

本书将集中论述结构模式识别的基本理论和方法。考虑到这方面的内容十分庞杂,因而仅限于论述那些对于该体系来说是最为基本的知识。这些知识在模式识别的实际应用中也是最有价值的。在论述结构模式识别之前,将首先介绍有关模式识别的基本概念。

1.1 基本概念

所谓“模式识别”,是指借助于电子计算机,就人类对外部世界某一特定环境中的客体、过程及现象的识别功能进行自动模拟的一门学科⁽⁵⁾。人对于外部世界环境中的客体、现象及过程具有识别能力,已经是一个众所周知的事实。但是,其识别机理却是至今未明的。为了弄清楚人的识别能力的机理,科学家们正在各个领域(诸如生理学、心理学、生物控制论、人工智能学与模式识别等等)中进行着顽强的探索。这种探索在基本方向上可以归纳为两种:一种是研究人的识别能力的数学生物学机理,另一种是研究人的识别能力的数学技术机理。模式识别在其目前的发展阶段,采取的是后一种。它主要是借助计算机来模拟识别的功能,而不是模拟识别过程中大脑运动规律的本身。所以,概而言之,模式识别是研究人的识别能力的数学技术机理的科学。所谓“数学”,是指提出与识别功能相适应的数学模型,以便描述这种识别功能;所谓“技术”,是指通过数学模型的程序化,以便实现这种识别功能的计算机模拟。了解这一点,对于正确地掌握模式识别的研究方法,无疑具有重要的意义。

一般地,由于人们识别的对象均处于外部世界的环境之中,因此,有关模式识别的基本概念可以通过对“环境”及“环境中的对象”的抽象化描述来逐步给出。

可以形式化地把“环境”表述为一个可测物理量（或函数） $\overline{b^p(x)}$ 的集合

$$U = \{\overline{b^p(x)} | p = 1, 2, \dots\} \quad (1.1)$$

而环境 U 的一个子集

$$\Omega = \{\overline{f^p(x)} | p = 1, 2, \dots\} \quad (1.2)$$

称为研究的“问题范围”，它给一个模式识别问题规定了确定的研究对象。

在式(1.1)中，每一 $\overline{b^p(x)}$ 的分量数目都是待定的，都可以是不同的。它表示，每一环境中的对象都可以通过足够多经过适当选择的函数来描述。而在式(1.2)中，每一 $\overline{f^p(x)}$ 的分量的数目都是确定的，且是相同的，从而可以保证对确定的问题范围定量地加以研究。例如，对于写阿拉伯数字符号的识别，是模式识别的一个“问题范围” Ω 的具体例子。此时，在 Ω 中包含了一切可能出现的阿拉伯数字符号，而每一数字符号均作为一个函数，且以同等数目的分量来加以描述，如图1.1所示。在问题范围 Ω 中，每一元素均要求通过确定的接收设备或

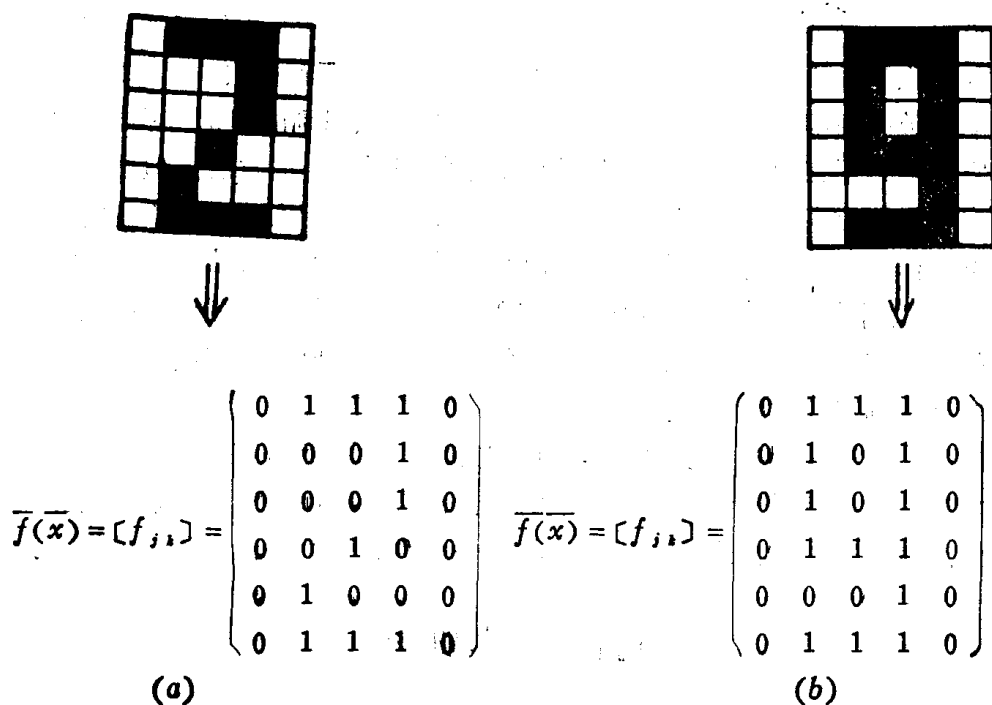


图1.1

数学模型来提供。反之，通过选择接收设备或数学模型，一个可测量的集合 Ω 将完全确定。接收设备的选择对模式识别问题的其后解决具有基本的意义，因为它将决定问题范围的对象性质，从而亦决定着其后的处理方法。例如，如果测量设备是多光谱图象接收装置，则识别任务必须借助于多光谱图象的处理与判读方法来完成；如果接收设备是光学扫描装置，则可应用灰度图象的处理与识别方法来解决，等等。

集合 Ω 中的元素称为“模式”。一个模式乃是一个函数

$$\overline{f(x)} = \begin{pmatrix} f_1(x_1, x_2, \dots, x_n) \\ f_2(x_1, x_2, \dots, x_n) \\ \vdots \\ f_m(x_1, x_2, \dots, x_n) \end{pmatrix} \quad (1.3)$$

对于已确定的模式识别的问题范围 Ω 来说， $\overline{f}$ 与 $\overline{x}$ 的分量的数目是常数，而下标 m, n 对所有

的 $\overline{f(x)} \in \Omega$ 是恒定不变的。例如，一张黑白照片可以用一个二维模式 $f(x, y)$ 来描述，即 $m=1, n=2$ ；在机器故障的诊断中，一个时序模型的输出向量是一个一维模式 $f(t)$ ，即 $m=n=1$ ，等等。

将集合 Ω 分解为 k 个（或 $k+1$ 个）子集 $\Omega_k \subset \Omega$ ($k=1, 2, \dots, k$ 或 $k=0, 1, 2, \dots, k$) 的运算，称为“模式分类”。前者称为“无拒识别要求的分类”；后者称为“带拒识别要求的分类”，并称类 Ω_0 为“拒识别类”，它表示 Ω 中的那些类别属性尚无充分理由确定的元素的集合。每一子集 Ω_k 称为“模式类”，它们应满足下述条件

$$\begin{cases} \Omega_k \neq \phi & k=1, 2, \dots, k \\ \Omega_k \cap \Omega_\lambda = \phi & \lambda \neq k \\ \bigcup_{k=1}^k \Omega_k = \Omega \text{ 或 } \bigcup_{k=0}^k \Omega_k = \Omega \end{cases} \quad (1.4)$$

集合 Ω 分解的一种特殊形式是包含在式(1.4)中的分层分解形式，即将式(1.4)理解为第一层分解，从而组成部分集合 Ω_k ；在第二层分解中，每一部分集合 Ω_k 的本身继续分解为更小的部分集合 $\Omega_{k\lambda}$ 。这一过程可以根据问题范围 Ω 的需要继续执行下去。

在模式分类中，可将模式视为一个整体，并将其映射到 k 个可能类别中的唯一的一个类别中去。一个模式分类系统，如果借助于计算机来实现，一般地具有一个分层的结构。如图1.2(a)所示，第 i 个模块的输出视为第 $i+1$ 个模块的输入，而其具体的实现形式则由图1.2

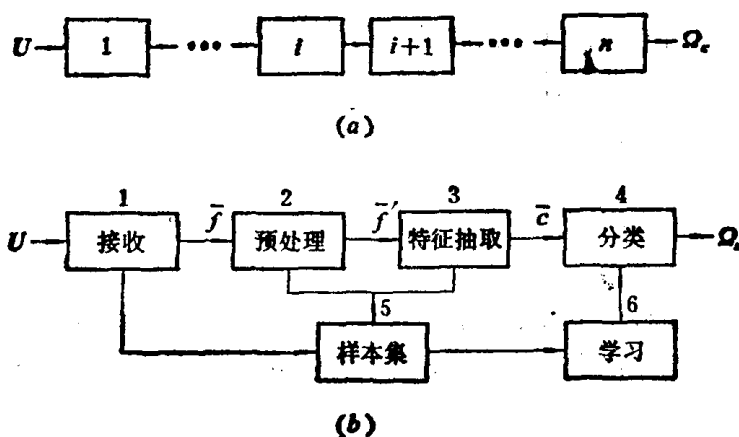


图1.2

(b)所给出。图中：框1实现环境 U 中问题范围 Ω 内的模式的接收，作量化（离散化）处理，并实现计算机的内部描述和存贮。它可以借助传感器、扫描器等装置来实现，其输出一般是模式的离散的灰度图象，但在某些情况下亦可能是实数向量。例如，在系统（或机器）的故障分析和诊断中，经常运用时序分析方法。此时，该框可理解为通过某一确定设备测量到的与时间有关的数据所建立起来的时序模型，其输出模式 $\bar{f}$ 为一实数向量。框2是对 $\bar{f}$ 的预处理，它实现一系列关于模式 $\bar{f}$ （数据模式或图象模式）的预处理措施。例如，在数据模式的情况下，实现平滑、规格化、加权等算法，以便为后续处理提供一个较为理想的数据集合，其输出仍为数据模式；在图象模式的情况下，实现平滑处理、增强反差、正规化处理、阈值处理、边缘检测等预处理措施，以便为后续处理提供一个质量较好的、更有利于达到识别目的的图象，其输出仍为图象模式。框3是对已经过预处理的模式抽取反映其类别属性的特征，这些特征综合为向量 $\bar{c}$ 。 $\bar{c}$ 称为模式的“特征向量”，是特征抽取模块的输出，并作为

识别分类器的输入。框4表示一个识别分类器，即一个分类程序，它实现未知类别的模式分类，其输出是类别号 Ω_k 。框5存贮所研究的问题范围 Ω 中的一个已知的样本集 ω ，它实际上是关于 Ω 的一个已知的先验知识的集合。框6则是通过 ω 来训练一个分类器，使得该分类器对于 Ω 中规定的分类任务来说是有效的（所谓有效，是指符合使用者的要求）。因此，模式分类亦可理解为由若干子任务组成的识别任务的总和，该总和意味着实现一个变换

$$\bar{f} \rightarrow \kappa \in \{0, 1, 2, \dots, k\} \text{ 或 } \{1, 2, \dots, k\} \quad (1.5)$$

实现上述形式的一个模式分类系统，是统计模式识别领域的主要研究课题。与此有关的基本理论和基本方法，亦成为结构模式识别的学科基础。

对模式识别的研究，除了上述模式分类的任务外，还有一个异常重要的分支，即模式分析。所谓“模式分析”，就是对每一模式 $\bar{f}(x) \in \Omega$ ，给出一个对应的符号化描述。而所谓“描述”，则是指把一个模式分解为模式的简单的组成部分，并分析它们之间的相互关系。

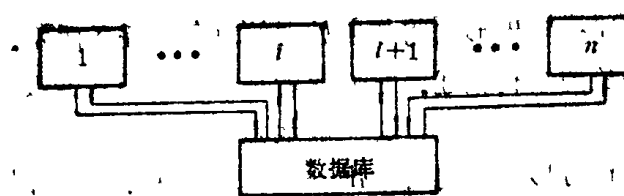
显然，模式分析是用于处理较为复杂的模式识别任务的。对于一个较为复杂的模式来说，往往包含着环境中的不同对象、客体、过程和现象，它们之间存在着丰富的时空结构信息。此时，若单纯将模式作为一个整体来分类是不可能的，或者是无意义的。即使这种分类是可能的或有意义的，但对于使用者来说，单纯给出一个类别号往往是不足够的，因为这种输出为使用者提供的信息太少了。

例如，对连续语言的自动理解，对电子线路图的自动分析，都是模式分析的例子。前者不仅要求识别组成语言句子的每一个词语（当然随之而来的是要求识别组成词语的字母），而且要求识别词语间的结构关系（句法的及语义的）；后者不仅要求识别组成线路图的各种元件，而且应正确地理解这些元件之间的接线关系，并在一个相应的数据结构中存贮起来，直至能够重现该电子线路图。

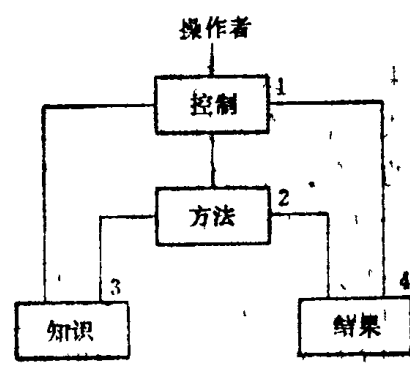
一般地可以认为，统计模式识别主要是研究模式分类问题，而结构模式识别既研究模式分析问题，也研究模式分类问题。而两者在研究模式分类问题时各自的特点与相互联系，将在1.2节及1.3节中论述。至于一个复杂模式的分析，则既需要综合应用统计与结构的识别方法，又需要应用人工智能科学中关于控制、搜索、知识的描述与运用等多种方法。前者将在第九章中论述，而后者已超出本书的范围，只在第九章中稍有涉及。就本书的任务而言，主要是介绍如何利用模式的结构信息，实现模式的自动描述及识别的基本理论与方法。但是，统计模式识别、结构模式识别与人工智能学一起，可构成模式分析的理论与技术方法的有机整体。

一个模式分析系统，如果借助于计算机来实现，一般地具有一个无顺序的分层结构（模块结构）。如图1.3(a)所示，模块间的运行顺序是由一个数据库来加以引导的，它是根据每一当前处理结果及数据库中存贮的数据来决定分析系统的后续处理步骤的。实验证明，在一个复杂的模式分析问题中，要想按一个规定的顺序无一例外地执行所有模块中存贮的运算是不经济且是无效率的，有时甚至还会适得其反。图1.3(b)给出了一个分析系统的具体形式——人机交互系统。在这个系统中，模块1与2是程序模块，其中模块2存贮全部作用于模式的处理程序；模块3与4是数据模块，其中，模块3是知识库模块，实现知识的内部描述与存贮，模块4是图象库模块，存贮分析过程的中间结果。控制模块用来实现系统之间的信息交换，并依据当前处理结果及已有知识选择后续处理方法和控制系统运行。

无论是设计一个模式分类系统或者是设计一个模式分析系统，都要求一个基本的前提，即要有一个有代表性的样本集



(a)



(b)

图1.3

$$\omega = \{f^1(\bar{x}), f^2(\bar{x}), \dots, f^k(\bar{x})\} \subset \Omega \quad (1.6)$$

可供利用，以便收集有关问题范围 Ω 的先验信息。此外，如果运用统计识别方法来设计分类系统，则还要求每一 $f^k(\bar{x}) \in \omega_k$ 具备其所属类别 Ω_k 的类属性特征 $\bar{c}_k$ ，所有 $\bar{c}_k$ 组成类别特征子空间 $C_{k,k}$ ，而所有 $C_{k,k} (k=1, 2, \dots, k)$ 组成“特征空间” $C_k = \bigcup_{k=1}^k C_{k,k}$ 。在 C_k 中，各个 $C_{k,k}$ 应在某种程度上相互分离。如果运用结构模式识别方法来设计分类或分析系统，则还要求模式 $f^k(\bar{x}) \in \omega_k$ 应能分解为简单的组成部分，且它们之间应具有确定的关系，即具有一种确定的结构。

在两种情况下，都要求能够对两个模式的“相似性”，给出一个可比较的度量。通过上述内容，读者已对模式识别的有关概念及其相互之间的关系有了概括的了解，下面将回到结构模式识别这一相对来说是较为狭窄的具体领域，以便进一步说明结构模式识别研究的基本方法及其特点。

1.2 关于结构方法与统计方法的比较

模式识别统计方法的特点是，将经过预处理的模式，通过研究式的或者分析的方法，抽取模式的特征，并把它整理到一个特征向量中去，作为分类器的输入，以实行决策判别，其输出是待识别模式的类别号。也就是说，首先将一个给定的模式变换为一个实数串，这一实数串确定地描述了该模式，然后对此实数串执行判决识别。用这种方法实现的模式识别系统，基本上只是处理待识别模式的统计特征的。但是，在一类模式识别问题中，模式的统计特征看起来并不起重要的作用，而起决定性作用的是模式的结构特征。所谓模式识别的结构方法，就是将经过预处理的模式，通过与识别对象相适应的不同方法，抽取模式的基本结构特征及这些特征之间的相互关系，组成基本模式元及其相互关系的一个字符串—棵树或一个图，并以此作为结构分类器或语法分析器的输入，实行决策分类或语法分析，其输出或者为

类别号, 或者为模式的描述; 输出形式将依所处理的对象和任务而定。在这种情况下, 一个二维的模式, 具有一个抽象的结构描述。它们或者是由基本模式元及其相互关系组成的一个字符串(即“句子”), 或者是一棵树, 更一般地是一个图^[5]。

例如, 为了描述一个正方形, 人们将指出, 该模式由四条具有相同长度的边所组成, 每两条边在其端点处相接并构成一个直角。在这里, 模式特征即“基本模式元”是“长度相等的边”, 其“相互关系”则是“两条边构成一个直角”, 这种相互关系确定了两个特征之间的相互几何位置。

用基本模式元及其相互关系来描述模式, 其作用是有可能使用比较少的基本模式元来描述一个较为复杂的模式, 且将模式的结构信息较为充分地显示出来。这一方法, 实际上是基于人们对自己所使用的自然语言的了解。在任何一种自然语言中, 词语总是应用字母来描述的, 每一个词均由字母组成的一个确定的字符串来表示, 而为数相对地极少的字母却能给出成千上万的不同的词。于是, 可以将一个相对地比较复杂的模式, 分解为一些简单的模式元; 或相反地, 可以借助于这些简单的基本模式元, 重新构造出一个比较复杂的模式。

把上面所介绍的结构模式识别方法的基本思想, 应用到具体的模式识别问题, 其区别于统计方法的特殊点在于, 寻找一个合适的基本模式元的集合; 对模式分解为基本模式元及由基本模式元重建该模式的机理, 加以形式化的描述和理解。

在统计模式识别方法中, 一个模式是作为一个整体来理解的, 并最终使其对应于 k 个可能类别中的一个。与此相反, 在结构模式识别方法中, 可以形式地描述一个模式, 这种形式的描述可能导致对模式实行分类、也可能导致对模式实行分析。也就是说, 作为结果而输出的是模式的描述, 或者是简单模式的分类, 或者是复杂模式的分析。但是, 这两者之间并不存在着截然可分的界限。

一般地, 结构模式识别方法相对于统计模式识别方法来说, 更多地是用于复杂模式的分析。由于复杂模式的分析所要求的运算时间相当长, 且寻找合适的基本模式元集合及寻找分析算法(包括描述基本模式元及其相互关系的文法)亦颇为困难, 故使得结构模式识别方法在实际应用中还处于发展的阶段。因此, 该领域目前正在吸引着大量科学工作者的注意力, 以力图解决结构模式识别方法应用于实际的问题。最近, 吴立德教授在研究描述直线的方向链码时, 得出了一个颇为有趣的结论: 直线的方向链码, 作为 $\{0, 1, 2, 3, 4, 5, 6, 7\}$ 上的一个集合, 不可能是一个前后文无关的语言。鉴于非前后文无关的语言在分析上的困难, 故这一结论起码提示了, 单纯的结构方法在模式识别应用中的局限性。因此, 在一个实际的模式识别问题中, 往往必须考虑同时运用统计识别与结构识别方法。它们是相互联系、相互补充的, 即使是纯粹的结构方法, 确定基本模式元的集合、亦依赖于统计分类方法的应用。所以, 统计模式识别方法对于结构模式识别方法来说, 是必不可少的基本理论方法^{[3][5]}。

1.3 结构方法的基本思想及一般原则

如前节所述, 结构模式识别方法是建立在形式语言的基础上的, 它借助于反映基本模式元及其相互关系的一个字符串来描述一个模式, 并通过语法分析来实现模式的分类或分析。

图1.4(a)给出了一个数学公式的例子, 它可视为一个复杂模式。显然, 实现对组成该公式的符号的识别及其排列顺序的判读, 对于该公式的识别来说是不充分的, 因为在一个这样的数学公式中, 存在着一个确定的三维有序关系。这种有序关系, 即前后文关系, 给出了公

式符号的正确意义,对于图1.4(b)所示的例子也是如此。这个模式的各不同组成部分可孤立地仅作为“线段”、“三角形”或“矩形”被描述。但是,一个矩形可能是一个窗口,而另一个矩形则可能是一扇门。同样地,依据量的关系,左边的模式可能是一间房子,而右边的模式则可能是一个猪舍。对于上述要说明的结论,人眼很容易判别。然而,电子计算机却无法在对问题不加形式化描述的条件下进行判别和分析。因此,借助计算机进行模式识别总是要求把如图1.4所示的模式转换为一个基本模式元的串,而该串中两个基本模式元的关系,则应以一个唯一明确的方式表达出来。

图1.5表示,对于一个数学公式来说,可对应着一个依据高级程序语言FORTRAN所改写的字符串。对于该公式来说,两种不同的书写方式,完全具有同样的意义和计算值。原始公式是一个符号的二维关系,而在其字符串的描述中,仅具有一维的顺序关系。很显然,后一种描述方式,直接地适用于计算机的后续处理。上述例子表明,把一个模式转换为一个字符串,对于人来说是无困难的。问题在于,这个过程应该加以形式化,以便使得机器能自动地加以处理,即必须发展一种转换的算法,使得一个模式到一个字符串的变换能被机器执行。

图1.5的例子还包含着某些其他关系。例如,字母所组成的串“EXP”的意义区别于串“TAU”,前者是函数名,后者是一个变量。若计算机正确地把“TAU”或者“OMEGA”

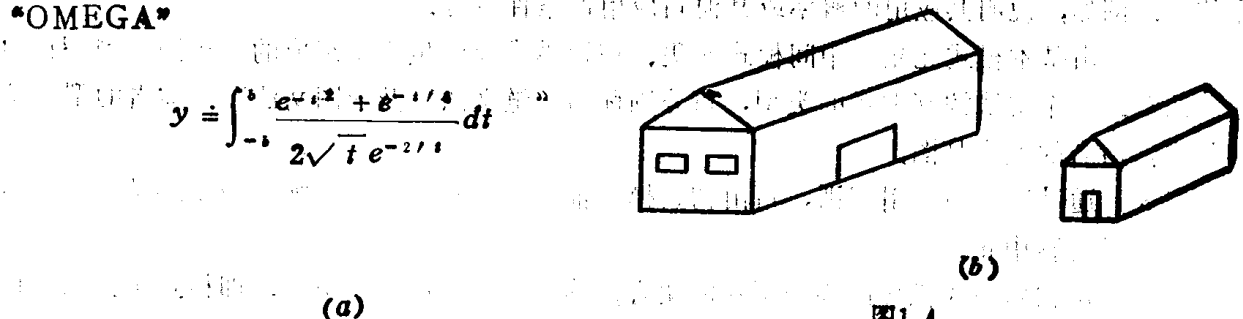


图1.4

$$y = \frac{ae^{-t/\tau} \sin \omega t}{\sqrt{1 + (b\omega)^2}}$$

$$Y = A \cdot \text{EXP}(-T/\text{TAU}) \cdot \text{SIN}(\text{OMEGA} \cdot T) / \text{SQRT}(1 + (B \cdot \text{OMEGA})^2)$$

图1.5

理解为“变量”,把“EXP”理解为“函数名”,则就意味着执行了一个分类。当然,这里除了对基本符号进行研究外,还必须对其相互关系进行研究。这种情况,正如对自然语言中的一个句子进行分析时的情况一样。一个词被称为“主语”或者“宾语”,这主要依赖于该词与句子中的其他成分(组成部分)的关系。一个句子的分析规则存在于一个自然语言的“语法”之中。如果熟知了该语言的语法,则对该语言中的句子进行分析就是可能的了。于是,可以想象,如果一个模式被一个字符串(即句子)所描述,且相应的语法是已知的,则对该模式进行分类或分析也是可能的了。反之,应用基本模式元及相应的语法规则,即可构造新模式,而这样构造的新模式,在语法所规定的意义上,是合乎规范的。

上述分析表明,在结构模式识别中主要解决下面三个问题:第一个问题是确定基本模式元的集合;第二个问题是形式地给出语法及与之相对应的语言;第三个问题是给出句法分析方法,使得对确定地限制了的模式识别的问题范围 Ω ,在一个可接受的计算时间内,求得一个解,即判别一个由基本模式元组成的抽象模式确定地来自于一种由相应语法生成的语言。属于同一语言的模式将视为同类的模式。

应用语法及句法分析来实现的结构模式识别方法，亦称为句法模式识别方法。这种方法的基本点可以归纳如下：

(1) 存在着一个有穷的“模式元”的集合，这个集合分解为“基本模式元”和“中间模式元”两个子集。如果把模式理解为一个字符串，则相应地亦称为“终结符集”及“非终结符集”。

(2) 模式元的连结构成“型式”，而最简单的型式是字符串。在字符串中，模式元是顺序相连的。一般地，一个型式是一个“图”(graph)。在图中，模式元是网络连接的。

(3) 如果一个型式仅由基本模式元所组成，则这个型式称为“终结的”。否则，称为“非终结的”。

(4) 存在一个非终结的“初始型式”，亦即待识别的初始模式。

(5) 型式之间的转换由一个有穷的规则集合所规定，称为“产生式集”。借助于产生式集，一个非终结的型式可以进行改变。

(6) 每一产生式均由一个“左边”(称为条件)及一个“右边”(称为结果)所构成。左边是一个非终结型式，右边是一个非终结型式或终结型式。如果在一个给定的型式中，出现一个与产生式的左边相同的子型式，则可由产生式的右边替代。产生式亦可包含有一些特殊信息，例如，说明其应用的概率或其他可应用性条件等等。

(7) 由基本模式元集、中间模式元集、初始模式及产生式集组成的一个四元组称为“文法”。一个文法按其型式的类型，可分别称为“链文法”及“图文法”。前者的型式是字符串，后者的型式是图。

(8) 通过反复地应用规则，可由初始模式推导出另外的型式。如果一个终结型式被给出，则推导过程中止。

(9) 由初始模式推导的终结型式的集合，称为文法所产生的语言，即模式类。为了便于模式的分类，选择文法时，应使得所产生的语言包含所有相同类模式的全部型式，其中当然也计及由该模式所得到的型式中可能出现的干扰。

(10) 由初始模式产生一个型式的反过程称为句法分析。通过句法分析确定，来自初始模式的一个预先给定的型式是否可以产生；应用哪些产生式可以产生。同样地，也应该考虑到在这些型式中可能产生的干扰。

(11) 分析可以由下而上(bottom-up)进行，即从一个给定的终结型式出发，通过产生式的相反的作用，试图回复到初始模式。实行这种分析过程的基本前提是，模式必须预先变换到一个由基本模式元组成的型式。

(12) 另一种分析方法是由上而下的分析方法(top-down)，即从给定的模式出发，通过产生式的运用，产生出给定的终结型式。因为初始模式是非终结的，借助于产生式，给定的终结型式的基本模式元将仅在分析过程中产生，因此，模式无须预先变换成基本模式元的描述。此时，模式的预处理和特征抽取也可以在文法及句法分析程序的控制下进行。

概而言之，句法识别方法要求确定合适的基本模式元集，设计对于一个确定的问题范围来说是合适的产生式集合，并给出一个句法分析程序。前两者统称为文法规则，它在大多数情况下是由识别系统的设计者来确定的。此外，亦存在着某些文法的自动设计方法。至于句法分析，原则上都归结为某些标准的分析算法。而在必要的情况下，也必须考虑到干扰的校正。当然，也存在着不依赖于文法及句法分析来实现模式分类的结构方法。这种方法建立在两个字符串，两棵树，或是两个图的“距离”的概念的基础上。这种结构分类器原则上相当

于统计方法的距离测量分类器。

上述两种结构模式识别方法，将在其后各章中分别详细地加以介绍。

1.4 结构模式识别系统

一般地，一个结构模式识别系统，或称为模式的分析和描述系统，可如图 1.6 所示。其任务是把一个模式分解为比较简单的模式元，而实现一个简单模式元的分割又相当于执行一个分类的任务。分解的困难在于，要决定应确定怎样的简单模式元（包括基本模式元和中间模式元）。在模式的自动处理中，信息的压缩和选择，总是与模式的描述和分析相联系的，而运用哪一些简单模式元来描述一个模式，又总是依赖于使用者的要求。通常，用于描述的简单模式元可直接通过一个分类算法来确定，而简单模式元又可进一步分解为更为简单的模式元。一个最为简单的来自自然语言的例子是，“句子”分解为“词语”，“词语”又分解为“字母”，“字母”又进一步分解为“外形元素”。另一个例子如图 1.7 所示。由图也可以看出，为了实现一个总的系统，这种分解应达到何种程度，一般是不能预先确定的。因为基本模式元的选择总是视使用者的要求而定。

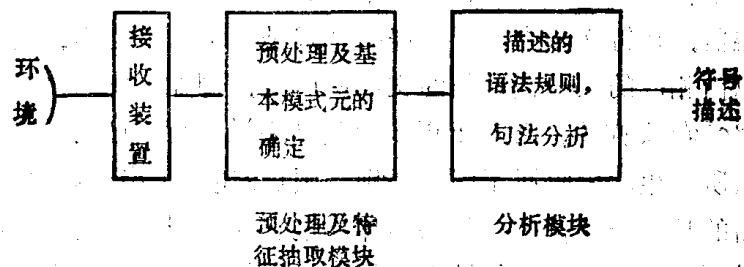


图1.6

5	场景	房子群
4	形体关系	 (岗亭)  (房子)  (塔) ...
3	几何图形关系	 (三角形) "在...上" {  (四边形) "在...中"  (四边形) }
2	线条关系	    ...
1	线条元素关系	/ "在...旁" / "在...旁" / ...

图1.7

为了实现这样的一个分析系统，如同在统计模式识别方法中所面临的问题一样，应收集一个在确定的问题范围内的模式的样本集，并找到一个对分析来说是合适的产生式集合，即

一种形式化的描述方法。一个模式的描述一般地并不是唯一的。例如，它依赖于人们是从图象的左上方还是从图象的右下方开始执行这种描述，还依赖于一个图象在平移、相似及旋转变换下的形态。因此，作为模式的特征描述应尽可能注意消除这些变换所造成的影响。当然，如果模式在相似和旋转变换下的形态应判为属于不同的类别时，情况是一个例外。此时，应保留模式变换时所导致的特征描述上的差别。同时，应考虑到，实现自动处理的方法不应导致不可接受的计算量，因为在一个复杂模式的分析系统中，处理的时间相对于简单模式的分类致来说要长得多。因此，计算时间的考虑具有特别重要的意义。

1.5 模式的结构描述方法举例

图1.8及图1.9给出了某些模式及其由基本模式元组成的型式并加以描述的例子，它们可用来揭示模式结构描述的若干基本方法，以便为后面第二章及第三章的形式化处理给读者提供一个直观的印象，并提示这些描述方法的基本点。

一维模式（即一维的函数过程）的描述仅要求一个基本模式元的顺序排列，即对应于一个字符串，其中每一字符都有一个左边字符及一个右边字符（仅第一个及最后一个例外）。实现这样的字符串的识别，依赖于一个链文法，或一个依据确定的“距离”准则而执行决策的结构分类器。在链文法中，所允许的型式是一维的字符串。所谓识别，就是判决一个字符串是否可以由相应的某一模式类的文法所产生。例如，图1.8中的一个正方形，可通过一个字符串 $a^5b^5c^5d^5$ 来描述。而对于正方形类别来说，可以应用链文法判别出不同正方形的型式均由该文法所产生，因而属于同一类别。图1.8还显示出，如果客体通过它的外形曲线来描述，则给出的是一个字符串；如果客体通过它的面积来描述，则给出的是一个字符场。在后一种情况下，如果采取4邻域定义规则，则每一字符有4个近邻字符；如果采取8邻域定义规则，则每一字符有8个近邻字符（仅边缘符号例外）。此时，所允许的模式的型式为二维的符号场。这样的一个二维符号场的集合将通过一个阵列文法来表征。而这样的二维甚至多维的关系也可以通过线性化而用符号串来描述。然而，人们当然还期待，通过对二维连结型式的直接理解将导致出现一个简单的、一目了然的关系。

图1.8还说明，一个电路图可以有选择地通过不同的形式加以描述，以便使模式元或者作为“场”的元素，或者作为“树”的元素。一个基本模式元在一个“场”中仅有4个或8个邻近的模式元，而在一棵“树”中，则可以有任意多个。在后者的情况下，模式元对应着树的结点，此时，所允许的型式是符号树，而符号树的一个集合将通过树文法来表征。树文法对于复杂模式的分析来说是合适的工具，因为模式及模式部分之间的分层关系在树文法中很容易理解，而且有成熟的分析算法。

模式的二维描述的最一般形式是图。普遍的观点认为，图文法在复杂模式的分析中将是合适的工具，因为模式及模式部分之间的位置关系在图文法中很容易理解。模式的图描述及图文法，是目前结构模式识别领域中最活跃的理论研究领域之一，它在模式识别中的应用尚在发展中。借助于模式的图描述，还可以通过图变换的方法来确定图的“距离”及其相似性度量，以便实现识别的目的。

至此，可以对结构模式识别的基本任务概括如下：

- (1) 确定基本模式元的集合（基本模式元的抽取与识别）；
- (2) 设计用以描述模式的文法（文法推断）；