

结构模式识别

〔美〕帕夫利迪斯 著

上海科学技术文献出版社

结构模式识别

[美] T·帕夫利迪斯 著

张寿萱 徐立明 译

吴立德



上海科学技术文献出版社

1981

结构模式识别

[美] T·帕夫利迪斯 著
张寿萱 徐立明 译 吴立德 校

*

上海科学技术文献出版社出版
(上海高安路六弄一号)

书店上海发行所发行
上海商务印刷厂印刷

*

开本 787×1092 1/32 印张 11 字数 262,000

1981年7月第1版 1981年7月第1次印刷

印数: 1—3,800

书号: 15192·147 定价: 1.40 元

《科技新书目》198-132

内 容 简 介

本书主要阐述结构模式识别的基本原理和一些具体的计算机算法。全书共分十章,包括:引论、曲线拟合的数学技术、图与格子、图片分割初步、高级分割技术、景物分析、区域边界的分析描述、区域边界和其它曲线的句法分析、用区域分析进行形状描述以及分类、描述和句法分析。可供与计算机科学有关的大专院校学生、研究生、教师及工程技术人员阅读参考。

Structural Pattern Recognition

[美] T. Pavlidis

Springer-Verlag Berlin Heidelberg 1977

作者为中译本写的前言

我感到十分荣幸, 我的书“结构模式识别”已被选出译成中文。中国人民的科学技术成就是众所周知的, 因此对中国科学家们对我的工作感到兴趣, 我觉得是一种荣誉, 也为此感到十分自豪。我希望本书的翻译将有助于增进美中两国之间的科学和技术交流。

我很幸运, 本书的译者之一张寿萱先生和校阅者吴立德教授我是认识的。我确信, 他们出色地完成了译校工作。我还要指出, 他们在翻译和校阅的过程中, 发现了原文中的一些错误, 承蒙他们作了校正。此外, 别的读者发现的差错他们也一一作了校正, 因此可以期望, 书中的中译本将会比包括第二次印刷在内的英文本更好些。

T. Pavlidis

普林斯顿

1980. 3

原 序

本书是以作者在普林斯顿大学为计算机科学和电气工程系高年级学生和研究生讲授一学期教程的讲稿为基础写成的。在讲授中作者着重强调了这样一个过程,即把图片数据转换成更适合于运用统计学技术或句法分析技术作进一步处理的数学结构。本书对题材的取舍正是反映了作者这一意图。作者也曾力求阐明通常所说的“特征生成”能以某种系统的方法加以实现,而不必过分依赖于启发和直观。

模式识别还是一个新兴领域,在方法上尚未定型,所以应把哪些内容选入一本一学期教程这样篇幅有限的书里是比较困难的。作者在本书中仅粗略地讨论了统计模式识别,这是由于这一主题似乎已趋于成熟,且已为许多好的教本和专著所叙述的缘故。另一方面,作者也试图全面地介绍形状分析的种种方法,其中包括,中轴对称变换,矩法, Fourier 描述器,链码,多边形逼近,分解法,积分投影等等。关于这些论题确实存在大量文献,但它们却往往没有被收入教科书或专著。在本书中,图片分割是按一般性的方法加以讨论,强调了数据的结构:区域相邻图,“金字塔”等等。其相应算法通过逻辑谓词来描述,从而允许算法适应于不同判别准则的分割。这在纹理分析等研究领域可能是十分有用。在这些领域中,不存在可接受的普遍准则,人们往往必须试验各种不同技术。作者也尽可能详细地列出了形形色色的算法,以便于它们可以不十分困难地以相当高级的语言来实现。作者希望这将使本书不仅对学生有用,也能对研究工作者有所裨益。

在写作过程中,作者荣幸地获得许多人的帮助:作者要感谢这部书的编者, K. S. Fu 教授,以往在结构模式识别的研究和在本书的写作过程中,我们都进行了许多次的讨论。C. M. Bjorklund 和 T. Boek 仔细阅读了各个阶段的手稿。承蒙他们的建议,我们作了多处形式上和内容上的改

进。关于松弛法的处理, 作者从与 S. W. Zucker 和 S. L. Davis 的讨论中颇获收益。许多例子利用了 S. L. Morowitz 和 F. Ali 所编写的软件。特别感谢 A. W. Evans 的卓越的绘图工作。衷心感谢 Springer-Verlag 出版社 H. Lotsch 博士在编辑事务方面的合作。本书的原稿是在电气工程和计算机科学系 PDP11/45 计算机 Unix 软件支持下准备的。当此机不能利用时, 又使用了统计系的计算机。在此也向国家科学基金会的支持致以诚挚谢意, 特别是对以 N. Caplan 为首的生物工程和自动化小组致谢。最后的但并非最不重要的是要感谢我的妻子 Marion, 孩子 Paul, Karen, Harry 在我沉湎于写作时所表现的忍耐和支持。

T. Pavlidis

于新泽西州普林斯顿

1977 年 9 月

校译者前言

本书是以作者在美国普林斯顿大学电机工程和计算机科学系为研究生和高年级大学生多年讲授模式识别这一课程的讲义为基础而写成的,适合于国内计算机科学系,自动化系,电子工程系,电机工程系中诸有关专业为研究生和高年级大学生开设这一课程时作为教材或基本参考书。

模式识别是近 30 年来随着计算机的迅速发展而发展起来的一门综合性很强的,既有深刻理论意义,又有广泛实际应用的新学科,它必将在我国的四化建设中发挥应有的作用。

本书以较少的篇幅,介绍了结构模式识别的基本内容,并把重点放在系统地介绍模式识别中最困难的特征抽取问题,即如何将输入的图片资料变换成便于进一步处理的各种数学结构。这在同类书籍中是少见的。同时,书中的许多方法都是作者首创并加以系统化的。因此,在一定意义上,本书也是一本专著。正因为如此,本书于 1977 年出版后,很受读者欢迎,并在 1979 年又作了第二次印刷。

本书作者 T. Pavlidis 教授是国际上知名的模式识别专家,美国电机和电子工程师协会的荣誉会员(Fellow),先后共发表了 100 多篇有关论文。在第四届国际模式识别会议(1978 年)上,作者应邀作了有关形状分析的综述报告,并被推选为第五届国际模式识别会议(1980 年)的总主席。作者也是多种杂志的编辑。

校、译者中的两位有机会与作者一起工作,并承作者好意为中译本写了前言和将第二次印刷时所作的校正告诉了我们,致使中译本有可能以更好的形式与读者见面。与作者一样,我们也希望本书中译本的出版能有助于促进中美两国的科学和技术的交流。

限于校、译者的水平,译文中错误和缺点在所难免,欢迎读者批评指正。

吴立德 张寿萱 徐立明

1980 年 3 月

目 录

1. 引 论	1
1.1 什么是我们所指的模式识别	1
1.2 研究模式识别的种种方法	3
1.3 物体的结构表示法	5
1.4 图片模式识别	7
1.5 模式识别的前景展望	8
1.6 算法书写格式的说明	9
2. 曲线拟合的数学技术	12
2.1 图片数据的数学表示	12
2.2 曲线拟合的基本知识	13
2.3 积累平方误差逼近	16
2.4 Karhounen-Loeve 展开及其变种	18
2.5 一致逼近	21
2.6 变换和空间域技术	26
2.7 分段逼近和样条	28
2.8 带固定结点的逼近样条	31
2.9 带可变接点的分段逼近	32
I. 一般性质	32
2.10 带可变接点的分段逼近	35
II. L_2 情况	35
2.11 带可变接点的分段逼近	39

III. 简单的例子	59
2.12 带可变接点分段逼近的算法	42
3. 图与格子	46
3.1 引言	46
3.2 图论初步	47
3.3 关于图的一些基本算法	52
3.4 离散几何初步	59
3.5 二值图的连通性和拓扑等价	62
3.6 多色图片的离散几何	65
4. 图片分割初步	68
4.1 引言	68
4.2 分割的形式定义	72
4.3 用于分割的数据结构	75
4.4 分割方法	79
4.5 取阈值和直方图分割法	82
4.6 边缘检测技术	85
4.7 分割作为函数逼近	92
5. 高级分割技术	97
5.1 引言	97
5.2 分裂和合并算法概述	98
5.3 图片编辑和 <i>RAG</i> 的一些性质	101
5.4 二维分裂和合并算法	112
5.5 由分割矩阵和区域相邻图构造区域边界表	124
5.6 三个区域相交处	127
5.7 逐行扫描分析和行相邻图	128
5.8 逐行扫描分析的扩充	134

6. 景物分析	137
6.1 引言	137
6.2 景物分析作为图的匹配问题	140
6.3 景物分析的局部化技术	145
6.4 景物分析的滤波和松弛技术	147
6.5 松弛法应用举例	151
6.6 非线性松弛技术	156
7. 区域边界的分析描述	161
7.1 引言	161
7.2 投影	164
7.3 边界标量变换技术	168
7.4 Fourier 形状描述器的性质和应用	171
7.5 空间域中曲线和边界的描述	175
7.6 曲率和角的重要性	178
7.7 用于多边形逼近的算法: 合并方案	
7.8 多边形逼近算法: 分裂和混合方案	
7.9 用于多边形逼近的分裂和合并算法	
8. 区域边界和其它曲线的句法分析	201
8.1 引言	201
8.2 边界的句法分析	202
8.2.1 正方形识别	203
8.2.2 缺口识别	205
8.3 局部特征检测的句法分析手段	208
8.4 噪声边界的分析	213
8.5 正则表达式	217
8.6 一种紧凑的边界描述	218

8.7	滤波和松弛匹配	229
8.8	波形	230
9.	用区域分析进行形状描述	235
9.1	引言	235
9.2	掩模和矩	237
9.3	细化算法和中轴变换	242
9.4	由骨架以及细线影象形成图和树	250
9.5	分解为初等凸子集	256
9.6	初等图	262
9.7	在凹角处分解	267
9.8	进一步分解和形状描述	272
10.	分类、描述和句法分析	280
10.1	引言	280
10.2	向量的统计模式分类	282
10.3	行的模式分类	286
10.4	树文法	288
10.5	图语言与图文法	294
10.6	一阶图文法	297
10.7	二阶图文法	300
10.8	CFGG 生成图的举例	304
10.9	线性图文法 (LGG)	306
	术语索引	309
	参考文献	321

1. 引 论

“眼见未必为实” M. Luckl Esh^[1.9]

1.1 什么是我们所指的模式识别

“模式”(Pattern)这个词与保护神(Patron)这个词来自同一根词。它的原意是指这样一些东西,它们作为供模仿用的、完美无缺的标本。于是,模式识别就是指识别出给定物体所模仿的标本。这是一个相当含蓄的定义,并且触及了一些非常深奥的论题。其一,它使我们回想起 Plato 的观点:客观世界本身正是完美理想的不尽完满的复制品。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”^[1.1]。但这个定义也相当繁复。因为要精确地定义什么是“有意义的感性经验”这件事本身就是困难的。

一个主要的问题是理想,或者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,我们假定理想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代表一种或几种已给定的理想时,这一过程可以称为“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(生物学意义上)。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。

三十年前，数字计算机的出现，有效地使一台机器以远远超过人的能力去进行数值计算。然而，一个自然的问题是是否这一机器也能被证明具备非数值运算能力，特别，能具有如模式识别这样的精神功能。这后一类问题一般被汇集在人工智能这一术语之下，虽然，另一个更恰当的名称也许应该是机器智能。

今天，也许我们可以说早先的一些乐观想法并未被证实。回想起来这不奇怪，特别是用生物学观点看尤其如此。比起先进的计算机来，人脑有着远为复杂的机制。“人脑计算机”不太适应于数值计算这一事实并非意外，因为，直至上一世纪，人进行这种运算的能力(在生物学意义上)对适应环境并没有什么好处。计算机在这些方面胜过人类这一点丝毫不意味着计算机的通用“计算”能力超过了人类。假如这一情况发生，那末也许只不过是惊人的巧合。

如果不是因为机器模式识别有重大实用价值这一事实，那末，上述讨论只有纯粹智力价值。然而，有许多“单调、乏味”工作，它们的自动化却都涉及到模式识别：邮件分类和把数据输入计算机，两者都需要既识别印刷体字符，也识别手写体字符；血球计数；质量控制的抽样检查和机器零件的装配；需要识别一般

的形状。类似的需要也还可以在航空和卫星照片，核物理气泡室照片，自动扫描射线照片等领域的分析中找到。在本书写作时，这些问题中的一些已获解决。商业上已出现利用光学设备可阅读印刷体或多种字体的打印文本的机器。血球计数器近来已用在现场检验中。然而，也仍然有许多其他问题尚待解决。

在继续叙述前，我们先给出一些有用的定义。模式识别显然是与分类有关，此时一个物体必须被判定为属于若干类中的一类。但这两个过程又不是等价的，我们可以涉及其中一个而不涉及另一个。在此，我们用“分类”这个术语表示指定类别。这些类别的每一个都对应于一个模式。我们还应对模式和性质加以区分。后者更广泛，且可以作为一种特殊情况而包含模式。模式这个词传统上被用来描述复杂性质的全体。模式识别的心理学方面的系统讨论已超出本书范围。这方面，读者可以参阅由 LINDSAY 和 NORMAN 所写的卓越的导论性的书^[1,11]。

1.2 研究模式识别的种种方法

由于其实用价值，模式识别已成为一个非常活跃的领域。有必要作一简单的历史回顾。许多早期的工作本质上都利用“gestalt”方法。无论输入的是什么，它都被表示为一个数组，它们是对原始数据进行各种测量所得的结果。这些数可以解释为向量空间中点的坐标，这也是几何学的这一术语经常被用来描述这种方法的原因。一旦完成了适当的测量，就可以发现源自同一模式的物体，就几何距离而论，其对应点总是互接近的。于是模式识别问题就变成寻找空间中一些区域的问题，这些区域上的点来源于某个单一模式。按照我们先前的讨论，这也称为分类。由于解决这类问题的方法基于概率统计理论，所以也使用术语统计模式识别。

在这一领域中有相当多的文献，且由于这个原因，在 10.2

节中,我们将只粗略地论及这个问题,大多给出适宜的参考文献的指示。此刻,我们要指出几何方法的主要优缺点。从好的方面说,人们也已经有了许多有效算法,形成了好的文件和得到详尽的分析。假如能做到“正确”测量,它们就能被成功地应用。然而,这是一个非常强的“假如”。一般地,这个方法不提供对测量的指导。有许多手段可以从一个大的候选集中找出某个最适合的子集合(这一过程往往称为特征选取),但是人们必须先找到这样一个大的候选集合。进行特征生成必须基于其它考虑,可能是物体的物理模型。

这一困难早在 1960 年就被认识,从而把许多研究者引导到一个不同的方法上。其基本思想是一个复杂模式可以由一些简单模式递归地描述。例如,字母表中的字母可以由笔划描述^[1,2,3]。这种递归描述与课文中短语通过单词来描述十分相似,由于这一原因,许多形式语言中的方法都可以被应用。因此,语言模式识别和句法模式识别这些术语就被用来描述基于这一思想的种种方法。NARASIMHAN^[1,4]和 LEDLEY^[1,5]属于最早发展的一些模式识别文法分析器。后来,SHAW 发展了图片描述语言并应用于气泡室照片分析^[1,6]。随后的年代,这一方向上的研究异常活跃。读者可以从 FU 的教本^[1,7]中看到关于这一论题的透彻处理。

形式语言理论的应用仅是这类方法的一个方面。它有希望的主要原因可能是模式的直观或物理的定义能被用来产生为计算机处理所适合的描述这样一个事实。以物体的结构作为对它们的描述,这是本质的。而句法分析技术的应用虽然重要但却不是下定义的出发点。于是,在 60 年代后期,结构模式识别这个术语开始被应用^[1,8]。

结构模式识别的目标是使用适合于下一步处理的数学模型

对物体描述作出推导。这下一步处理过程应该是相对简单的。如果运用句法分析工具,则相应的文法应易于分析。如果使用几何学描述,则同一类物体的相应点,应按照某一简单的距离紧密聚集在一起。注意这最后一句话意味着结构描述并不必须是递归的。诚然,大多数有趣的方法是与递归情况相关的。

1.3 物体的结构表示法

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,一个图象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。在某种意义上,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。这两种表示法的任一种都不一定会有信息丢失。如有任何信息丢