

神经网络模式识别及其实现

Pattern Recognition with Neural Networks in C++

[美]Abhijit S. Pandya Robert B. Macy 著

徐勇 荆涛 等译

電子工業出版社

Publishing House of Electronics Industry

北京·BEIJING

- 尽管有许多有关统计模式识别的著作,但它们并不包括神经网络方法。反过来,有关神经网络的教科书也极少讨论模式识别问题。

本书试图在同一框架下讨论模式分类和神经网络方法,并使之适合于专业人员。认为神经网络是由复杂机制控制、能提供令我们惊奇或失望结果的黑盒子是不应该的。只要有了对基本理论和实际例子的理解,专业人员就能够更好地在设计中作出正确选择,使神经网络的应用更加具有预见性和更为有效。

对于每一个网络示例,我们努力为每一个方法给出直观的解释,并在合适的地方用严格的数学方法加以论证和扩充。为了丰富和更加明确地描述各种概念、模型和算法,给出了大多数所讨论的神经网络模型的实际 C++ 实现程序,提供了论题的解释图表和方法的描述。我们还提供了数学模型必要的推导过程,以便聪明的读者能够随着本学科领域新的发展,将自己新的思想加到程序中。

对于每一个方法,我们试图简单明了地阐述已知的理论成果、发展趋势,以及对如何从该方法中获得最好结果的建议。一些人也许对这些建议持有异议,但至少专业人员可以从中受益和获得实验经验。我们力图使几乎没有或不清楚神经网络背景知识的工程师们,容易理解这些方法。然而,由于所提供的材料有足够的深度,那些预先具有背景知识的读者会发现本书对他们也是有裨益的。

本书所讨论的内容,适用于计算机科学和大多数工科专业的高年级大学生和研究生作为学习神经网络方面的课程。通过学习模式识别的知识,可以很容易地观察和理解神经网络的原理和应用。此外,在个别情况下,如果需要重点学习神经方法或与其它介绍统计方法的教科书结合起来学习时,本书也适合于用作模式识别的教程。

那些具有某种神经网络范例的实践经验的读者,也许对于其中的网络结构之一已经进行过编程应用,那么他应该能够顺利地应用本书的材料。我们希望本教科书能够对加强学术和实践的基本原理研究,提供有价值的帮助。

译 者 的 话

神经网络是人们在模仿人脑处理问题的过程中发展起来的一种新型智能信息处理理论,它通过大量的称为神经元的简单处理单元构成非线性动力学系统,对人脑的形象思维、联想记忆等进行模拟和抽象,实现与人脑相似的学习、识别、记忆等信息处理能力。神经网络在经历了 40 多年的曲折发展之后,在信息科学领域等许多应用方面已显示出巨大潜力和广阔的应用前景。

神经网络模式识别方法是近几年兴起的模式识别领域的一个新的研究方向。由于神经网络的高速并行处理、分布存贮信息等特性符合人类视觉系统的基本工作原则,具有很强的自学习性、自组织性、容错性、高度非线性、高的鲁棒性、联想记忆功能和推理意识功能等,能够实现目前基于计算理论层次上的模式识别理论所无法完成的模式信息处理工作,所以,采用神经网络进行模式识别,突破了传统模式识别技术的束缚,开辟了模式识别发展的新途径。同时,神经网络模式识别也成为神经网络最成功和最有前途的应用领域之一。研究神经网络模式识别系统,无论对于神经网络理论的发展,还是对于模式识别技术的实际应用,都具有特别重要的意义。

为适应从事神经网络和模式识别领域学习、研究的各种读者的需要,我们翻译了美国 CRC 出版公司出版的专著“Pattern Recognition with Neural Networks in C++”。本书将神经网络理论与模式识别技术有机地结合起来,重点论述神经网络在模式识别系统中的应用问题,给出了大量的用于模式识别的神经网络模型的 C++ 实现程序。其突出特点是内容新颖,概念清晰,既有深入浅出的理论分析,又有实用易行的计算机实现方法。我们将本书翻译过来介绍给国内读者,希望能为高等院校有关专业的研究生和高年级大学生提供相应的教材或教学参考书,并对我国从事神经网络和模式识别领域工作的研究人员和工程技术人员有所帮助,对该领域的研究和应用起到一定的促进作用。

承蒙电子工业出版社购买了原著的中文本版权,并给予大力支持,本书才得以顺利出版。在本书的翻译过程中,赵晓晖帮助校阅了译文的部分章节,王伟、赵岩、邓小英、魏静参加了部分翻译工作。在此一并表示衷心的感谢。

对于原书中的一些印刷错误尽可能作了订正。由于译者水平有限,翻译中错误和不妥之处在所难免,欢迎读者批评指正。

译 者

序 言

为什么在已经有许多优秀著作的情况下,我们还感觉有必要写一部专著,来论述模式识别这个传统的论题?答案在于,我们是要深入全面地讨论作为自然模式分类器或聚类器的神经网络。人工神经网络计算是一个极为活跃的研究领域,它的一个重要应用就是处理模式化信息以及含有潜在模式的信息。神经网络研究无论过去还是现在的迅速进展,促进了模式识别其它相关方法的发展。较之传统的统计模式识别的范畴,模式识别已经成为更为广泛意义上的方法学。将人工神经网络计算引入传统模式识别中,则使模式识别这一方法学变得更加具有新意和有力,这正是本书要向这方面的专业人员介绍的内容。

模式识别系统是基于复杂模式或对象的可测属性,或由可测属性得到的特征,对这些模式进行自动分类或聚类的系统。根据这个观点,神经网络可以被看作为一种识别模式的系统。然而,确切地说,神经网络的优势是完成寻求潜在规则方面的任务。所以,在这个意义上,神经网络的研究和模式识别的研究是相互交叉覆盖的,任何一方缺了对方,它的学科就不是真正完整的。在一些通常明确不属于模式识别的领域(例如控制)中,神经网络得到了许多有效的应用。这些领域利用的也是神经网络探测和识别输入空间潜在规则的能力,从这个意义上说,它们也可以被认为是属于模式识别的范畴。

神经网络应该体现出多少生物系统的性能,一直是一个有争议的问题。本书主要讨论工程方法,不涉及任何生物学上似乎真实的论述。尽管如此,我们也认识到,设计一个系统时,观察一下能很好完成预期功能的其它系统是很有用的。生物系统是超级模式识别器。人们观察鸟类,分析其飞行特性(例如,翅膀形状、质量/体积比,等等)并进行模仿,对设计第一架飞机起了很重要的作用。但是,在模仿过程中,人们必须避免跟着生物系统亦步亦趋,否则,就不可能出现能进行超音速飞行的航天飞机。所以,我们可以相信,在从生物系统获得灵感的同时,人工神经网络最终一定能使其本身更加真正有效。

我们的目的是引导人们实际应用人工神经网络。尽管已有许多神经网络方面的教科书,并且可以作为很好的参考资料,但是以我们所见还不能完全满足该领域应用的需要,其中一些理由如下:

- 大多数有关人工神经系统的论著,是由从事理论研究的学者撰写的,对一般理论进行了很好的论述,但通常忽略了各种网络结构和学习算法应用的实用信息。它们经常给出的是某段程序代码和个别算法的计算机实现结果,而很少提供能有助于实验设计的完整程序代码。
- 最近出版的神经网络论著中,绝大部分是将个人或几个人的研究论文编辑在一起,阅读起来必须需要预先从事相关专业领域的经历。对于一些神经网络算法,即使加以详尽论述,网络所完成的计算细节还是不容易实现,而且它们的结果也不容易跟踪。
- 目前出版了几本有关神经网络方案的详细说明书,其中包含有简化公式和实现相应方法的 C++ 程序,但是这些书是为外行编写的。

- 尽管有许多有关统计模式识别的著作,但它们并不包括神经网络方法。反过来,有关神经网络的教科书也极少讨论模式识别问题。

本书试图在同一框架下讨论模式分类和神经网络方法,并使之适合于专业人员。认为神经网络是由复杂机制控制、能提供令我们惊奇或失望结果的黑盒子是不应该的。只要有了对基本理论和实际例子的理解,专业人员就能够更好地在设计中作出正确选择,使神经网络的应用更加具有预见性和更为有效。

对于每一个网络示例,我们努力为每一个方法给出直观的解释,并在合适的地方用严格的数学方法加以论证和扩充。为了丰富和更加明确地描述各种概念、模型和算法,给出了大多数所讨论的神经网络模型的实际 C++ 实现程序,提供了论题的解释图表和方法的描述。我们还提供了数学模型必要的推导过程,以便聪明的读者能够随着本学科领域新的发展,将自己新的思想加到程序中。

对于每一个方法,我们试图简单明了地阐述已知的理论成果、发展趋势,以及对如何从该方法中获得最好结果的建议。一些人也许对这些建议持有异议,但至少专业人员可以从中受益和获得实验经验。我们力图使几乎没有或不清楚神经网络背景知识的工程师们,容易理解这些方法。然而,由于所提供的材料有足够的深度,那些预先具有背景知识的读者会发现本书对他们也是有裨益的。

本书所讨论的内容,适用于计算机科学和大多数工科专业的高年级大学生和研究生作为学习神经网络方面的课程。通过学习模式识别的知识,可以很容易地观察和理解神经网络的原理和应用。此外,在个别情况下,如果需要重点学习神经方法或与其它介绍统计方法的教科书结合起来学习时,本书也适合于用作模式识别的教程。

那些具有某种神经网络范例的实践经验的读者,也许对于其中的网络结构之一已经进行过编程应用,那么他应该能够顺利地应用本书的材料。我们希望本教科书能够对加强学术和实践的基本原理研究,提供有价值的帮助。

目 录

第一章 引言	(1)
1.1 模式识别系统	(1)
1.2 人工神经网络方法的产生	(4)
1.3 模式识别序言	(7)
1.4 统计模式识别	(8)
1.5 按句法规则的模式识别	(11)
1.6 字符识别问题	(13)
1.7 题目的组织	(15)
参考书与文献	(15)
第二章 神经网络概述	(17)
2.1 生物神经网络概述	(17)
2.2 背景	(17)
2.3 生物神经网络	(18)
2.4 大脑中的分层组织	(21)
2.5 历史背景	(25)
2.6 人工神经网络	(30)
参考书与文献	(32)
第三章 预处理	(35)
3.1 概述	(35)
3.2 扫描图像的处理	(35)
3.3 图像压缩	(36)
3.3.1 图像压缩的例子	(37)
3.4 边缘检测	(38)
3.5 骨架处理	(44)
3.5.1 细化的例子	(46)
3.6 处理手写板输入	(51)
3.7 图像的分割	(54)
参考书与文献	(55)
第四章 有监督学习的前馈网络	(57)
4.1 前馈多层感知器结构	(57)
4.2 用 C++ 实现前馈多层感知器	(59)
4.3 利用 B-P 算法进行网络训练	(68)
4.3.1 用 C++ 实现 B-P 算法	(74)
4.4 一个基本例子	(84)
4.5 训练策略和避免局部最小	(90)
4.6 梯度下降中的变量	(92)
4.6.1 块适应和数据适应梯度下降方法的比较	(92)

4.6.2 一阶和二阶梯度下降方法的比较	(93)
4.7 拓扑	(93)
4.8 ACON 和 OCON 的比较	(93)
4.9 过训练和推广	(95)
4.10 训练集合和网络大小	(98)
4.11 共轭梯度方法	(98)
4.12 ALOPEX	(100)
参考书与文献	(112)
第五章 其它类型的神经网络	(115)
5.1 概述	(115)
5.2 径向基函数网络	(115)
5.2.1 网络结构	(115)
5.2.2 RBF 训练	(116)
5.2.3 RBF 网络的应用	(118)
5.3 高阶神经网络	(119)
5.3.1 引言	(119)
5.3.2 结构	(120)
5.3.3 几何变换的不变性	(122)
5.3.4 范例	(123)
5.3.5 实际应用	(124)
参考书与文献	(125)
第六章 特征提取 I:几何特征和变换	(129)
6.1 概述	(129)
6.2 几何特征(环、交叉点、端点)	(129)
6.2.1 交叉点和端点	(129)
6.2.2 环	(132)
6.3 特征映射	(141)
6.4 基于几何特征的一个网络例子	(143)
6.5 利用变换进行特征提取	(144)
6.6 傅立叶描述符(FD)	(144)
6.7 Gabor 变换和子波	(146)
参考书与文献	(151)
第七章 特征提取 II:主分量分析	(153)
7.1 降维	(153)
7.2 主分量	(154)
7.2.1 PCA 示例	(155)
7.3 KARHUNEN-LOEVE(K-L)变换	(156)
7.3.1 变换示例	(157)
7.4 主分量神经网络	(159)
7.5 应用	(160)
参考书与文献	(163)
第八章 Kohonen 网络和学习矢量量化	(165)
8.1 概述	(165)

8.2 K-均值算法	(166)
8.2.1 K-均值算法举例	(171)
8.3 Kohonen 模型介绍	(180)
8.3.1 Kohonen 网络示例	(187)
8.4 侧反馈规则	(189)
8.5 Kohonen 自组织特征映射	(191)
8.5.1 SOFM 举例	(200)
8.6 学习矢量量化	(204)
8.6.1 LVQ 举例	(211)
8.7 LVQ 的改进	(221)
8.7.1 LVQ2	(222)
8.7.2 LVQ2.1	(222)
8.7.3 LVQ3	(223)
8.7.4 LVQ 的最后变形	(223)
参考书与文献	(224)
第九章 神经联想记忆和 Hopfield 网络	(225)
9.1 概述	(225)
9.2 线性联想记忆(LAM)	(226)
9.2.1 一个自联想 LAM 例子	(227)
9.3 Hopfield 网络	(236)
9.4 Hopfield 网络的一个范例	(241)
9.5 讨论	(242)
9.6 位图范例	(244)
9.7 BAM 网络	(250)
9.8 一个 BAM 网络范例	(252)
参考书与文献	(255)
第十章 自适应共振理论(ART)	(257)
10.1 概述	(257)
10.2 寻求聚类结构	(257)
10.3 矢量量化	(257)
10.3.1 VQ 举例 1	(264)
10.3.2 VQ 举例 2	(269)
10.3.3 VQ 举例 3	(274)
10.4 ART 基本原理	(279)
10.5 稳定性和可塑性两难问题	(280)
10.6 ART1:基本工作方式	(281)
10.7 ART1:算法	(287)
10.8 增益控制机制	(288)
10.8.1 增益控制举例 1	(294)
10.8.2 增益控制举例 2	(298)
10.9 ART2 模型	(303)
10.10 讨论	(304)
10.11 应用	(307)
参考书与文献	(309)

第十一章 神经认知机 (311)

11.1 引言 (311)

11.2 网络的结构 (311)

11.3 神经认知机的一个例子 (316)

参考书与文献 (321)

第十二章 多分类器系统 (323)

12.1 综述 (323)

12.2 多种识别器组合成的系统结构 (325)

12.3 投票方案 (327)

12.4 混淆矩阵 (329)

12.5 可靠性 (331)

12.6 一些经验方法 (332)

参考书与文献 (333)

第一章 引言

什么是模式?模式实际上可以说是存在有某种基本结构组织的排列或有序化。我们可以把整个世界看作是由模式构成的。Watanabe[1985]把模式定义成一个能够给出名字,但不能明确定义的一个实体。

模式也可称为对某一事物或其它一些感兴趣项目的定量或结构上的描述。一组具有公共特性的模式可以看作为一个模式类。通过机器进行模式识别的主要问题,就是如何采用更好的处理技术自动地、尽可能少由人介入地把模式分到各自的类中。例如:在邮局根据用五位数字的邮区代码自动分拣邮件的机器,就要识别各位数字。在这种情况下有10个模式类,每个模式类中有10个数字。邮区代码识别机的功能就是识别有一定几何图形的模式(每个模式代表一个输入数字),即识别出它是可用的某个模式类中的哪一个。

模式可以用一个由标准激励或用取自标准激励和它们相互关系的属性组成的矢量来表示。通常,一个模式的特性是通过组成它的各元素的顺序来表示,而不是通过这些元素的内在性质来表示。广泛地说,模式识别就是把测量结果、激励或输入模式,划分或分配到有意义的类别中。自然地,它包含从来自不相关细节的背景数据提取重要属性。语音识别是从代表语音的波形中映射变换出话音。在字符识别中,模式识别就是从像素(或笔划)矩阵中识别出字符和词。模式识别的其它例子包括:核实签名、从像素图中识别人的面孔以及敌友识别。同样地,一个接收声纳数据的系统可以通过输入数据判定目标是一艘潜艇还是一条鱼,这样的系统就是一个模式识别系统。

1.1 模式识别系统

对于一个典型的模式识别系统,类的判别仅仅是整个系统任务的一个方面。总的来说,模式识别系统以“原始”测量结果的形式接收数据,这些原始数据集合形成一个激励矢量。寻求存在于激励矢量内部特征中的相关属性,是这种系统一个典型的基本要求(在一些情况下可能是全部要求)。这些能够更真实、更清晰地表达模式基本结构相关属性的一个有序集合构成一个特征矢量。

类别仅仅是根据问题的性质,可以或不可以必须确定的属性之一。属性可以是离散的值、布尔实体、按句法规则的句子或模拟值。在模式识别中,学习就相当于模式特征和属性之间相互关系规则的判定。

通常来说,实际的图像识别系统除了识别机器本身外通常还包括几级。在集中研究神经网络识别方法之前,我们先简要描述一下某种典型的识别系统[Chen,1973]。

图1.1中示出了一个典型模式识别任务的所有方面,它们包括:

预处理:把图像划分成独立的识别对象(即字符等)。

此外,它还能按比例变换图像以便将任务集中于对识别对象的处理上。

特征提取:提取各个模式的高层信息使其容易识别。

分类器:分类器识别出模式所属的类别或者大体上确定给定模式的属性。

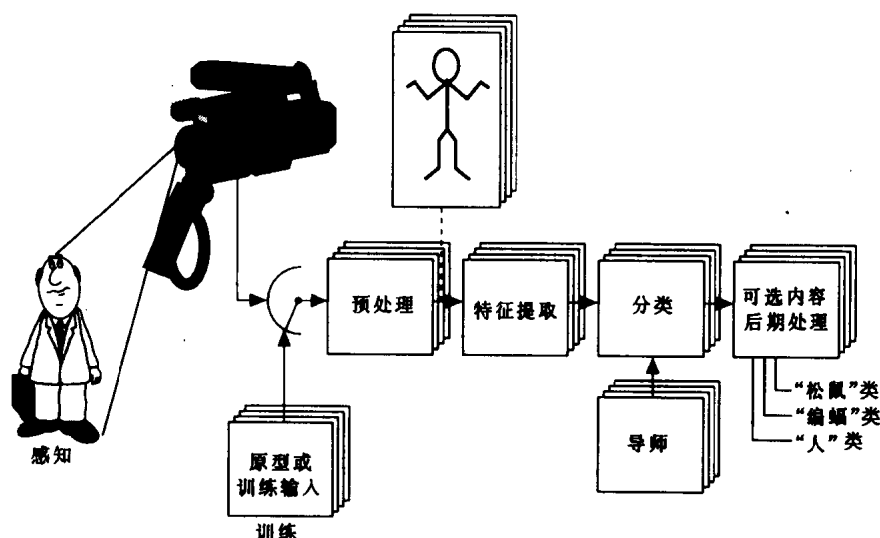


图 1.1 模式识别系统的构成

相关处理:相关处理器通过提供关于识别对象周围环境的相关信息来提高识别精度。例如,在字符识别的情况下,它可以是字典或语言模型支持。

图 1.2 给出了一个典型模式识别系统的设计步骤。如何选择合适的传感器、预处理技术和判决算法,由这个领域问题的特性来确定。不像专家系统,在设计中,特定领域的知识是隐含的并且不能通过一个独立的模块来代表。

模式分类系统能完成:

- (1)有监督分类:对一个给定模式可以识别它是一个已知的或已定义的类别的那一类;
- (2)无监督分类或聚类:对一个模式需要把它分到一个到目前为止还未知的模式类别中。

模式识别可以是静态的,也可以是动态的。在异步系统的情况下,时间或相继顺序不起任何作用。这种情况可以使用静态模式识别。图像识别/理解属于这一类。对动态模式识别,相对定时是很重要的。输入和输出之间时间关系具有重要作用,学习过程必须决定控制这些时间关系的准则。这包括诸如使用人工神经网络进行控制,或使用神经网络来预测等的一些应用。例如,在识别手写字符的情况中,在一个数字化的便笺簿上,笔划出现的顺序可提供很多对识别处理有用的信息。

当各类别间具有重叠区时(参见图 1.3),模式识别的任务就更复杂了。在这种情况下,识别系统必须设法使分类误差最小化。分类误差很明显地受到训练集中样本数目的影响。一些研究者(例如 Jain 和 Chandrasekaran[1982],Fukunaga 和 Hayes[1989],Foley[1972])已经讨论过这个问题。

设计一个模式识别的三种主要方法为:(1)统计方法;(2)按句法规则的或结构化方法;以及(3)人工神经网络方法。统计模式识别技术使用统计信息和估计理论的结果,去获得从表达空间到解释空间的映射。这些结果依靠一个合适的特征值组合来确定,这些特征值提供两种类别之间差别的测度。然而在某些情况下,特征本身并不是重要的。而关于模式类别或模式属性的关键信息,包含在特征的结构化关系之中。有关形象化的模式(由可识别的形状表示其

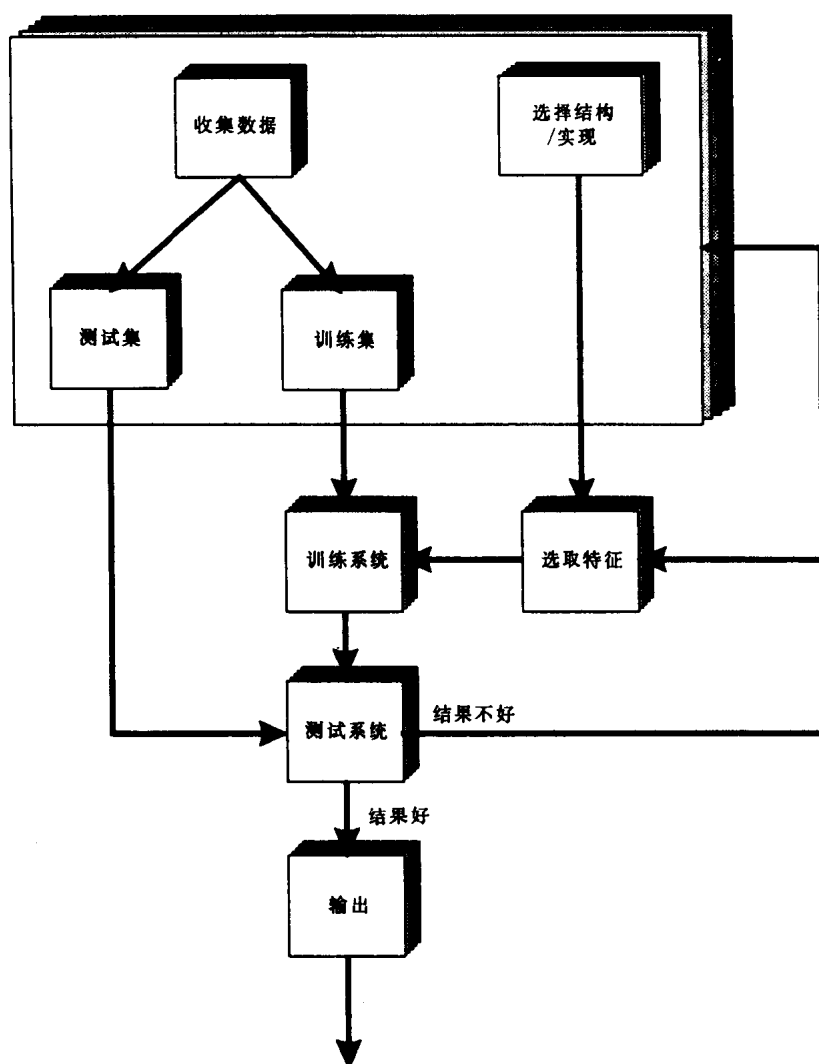


图 1.2 设计一个用于模式识别的学习机的过程流程图

特征)识别的应用,诸如字识别、染色体鉴别、基本粒子碰撞照片等都属于这一类。既然按句法规则的模式识别具有统计模式识别方法所缺乏的结构处理能力,那么它的目标就是处理这方面的问题。在这个领域的许多技术,都源自于数学语言学的早期工作和计算机语言学的研究成果。在这个领域,像 Watanabe[1972],Fu[1974,1977],Gonzalez 和 Thomason[1978]等已有大量文献来论述这些问题。

尽管存在着许多好的统计的、按句法规则的(以语法为基础)和图解的方法可用于模式识别,我们这本书还是仅限于讨论各种基于人工神经网络的方法。然而,统计的方法和神经网络技术之间的关系非常密切,所以也讨论了适当的统计方法。除此之外,不应该忽视的是神经识别器能够并且已经用于和其它类型诸如缓冲模式匹配器等识别机器的组合之中。

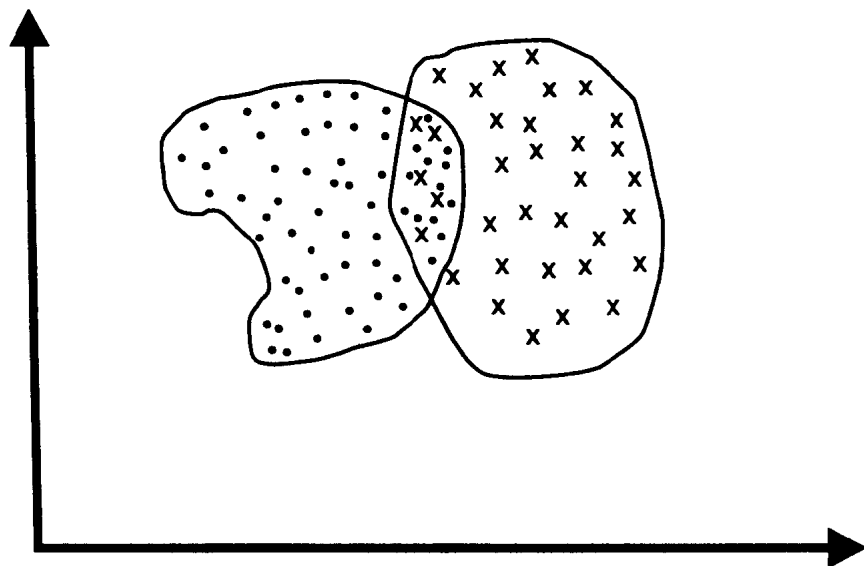


图 1.3 模式空间中的两类模式,可以看到在重叠区域中的模式属于两个模式类

1.2 人工神经网络方法的产生

计算机作为一个不仅仅只是用于计算的机器,它的发展标志着模式识别领域的诞生。我们已经看到,关于使用机器来执行通常和人的行为相关的智能任务的研究兴趣的提高。模式识别技术就是在机器智能领域已经使用的重要工具之一。识别终究可以看作是生物组织的一种基本属性。生物系统(包括人类)模式识别能力的研究,属于诸如心理学、生理学、生物学和神经科学等学科的范围。对于用机器实现一个给定任务的识别的实际技术和用于设计这些系统的必要数学构架研究,属于工程、计算机科学和应用数学领域。随着神经网络技术的出现,在从事生物系统领域研究的工程师和学生以及心理学家、生理学家、语言学家等等之间,建立了共同的研究领域。可以指出,在模式识别和神经网络理论中使用的数学运算,通常形式上相似甚至相同。所以,从数学的角度,有很好的理由将这两个领域的内容放在一起讲授。

识别模式(并以识别为基础产生动作)是所有生物系统共有的主要活动。总的来说,生物系统,特别是人类,是已知的最灵活、有效和万能的模式识别器;而且他们的行为提供了用于研究模式识别问题的足够数据。例如,我们能够不费力地识别出手写字符,尽管这个字符写得有失真、遗漏或较大的变化。在语音识别中,也可以发现同样的能力。人也具有在只有模式的一部分存在时,根据模式的相互联系来获得信息的能力。以鸡尾酒会中出现的现象为例,在酒会大厅任一方向,当人们的谈话中提到你的名字,你都会听到,即使这时因为噪音大多数谈话都听不到。类似地,你能在远处的人群中辨别出朋友的身影,即使这时大多数身影你都看不到。

人类做决定的过程通常与模式识别有关。人类善于寻找模式的相关关系,并根据它们提取规律性。这种观察力使人们能按预期想法行动,这样可以缩短反应时间并给出反应行为的界限。人们通常把机器设计成以某些特定事件的出现作出的反应为基础进行工作。这样使它们在诸如控制等应用上的速度减慢了。被识别模式的特性可以是感觉识别或理性识别。前者包括用诸如视觉或听觉刺激等感觉信息对具体实体作出识别。物体、字符、音乐、语音、签名等的识别可以看作是感觉识别的例子。另一方面,理性识别包括诸如对解决一个问题或一个旧

观点的识别。它包括抽象实体而且在这种情况下不需要外界刺激。在这本书中,我们将只涉及具体实体的识别。

然而,模式识别的真正问题是产生一种理论。这种理论通过一种机器能稳健识别物体的方式来确定其特性。生物系统工作方式的研究,使人们对如何解决这个问题有了深入了解。图 1.4 中的图像表示了我们所讨论问题的复杂性。在图 1.4(a)中,图像为像素间具有明显边界的人脸图像。所以,一种图像理解/模式识别算法,如果用具有不同明暗度的区域标识不同部分的表面,在识别人脸这种模式时就会有困难。另一方面,一个观察者对如图 1.4(b)所示的像素之间边界模糊不清图像,则更容易识别成熟熟悉人的面孔。这种能力可以归因于,在人的视觉系统中存在着的高低空间频率通道的相互作用。



图 1.4(a) 低分辨率具有可见像素网格的面部图像

工程和人工智能界的一个很大的目标,就是创造出行为如人的“智能”系统。这种智能行为能使人类在一种更为自然的方式下进行人/机交互。那就是,我们将提供感觉的和认识的能力,使计算机能以一种自然的、直观的方式与我们交流信息。设计具有判决能力的机器是众多

目标之一。为了达到这个目的,这样的机器应具有和人相同的模式信息处理能力。



图 1.4(b) 象素之间边界模糊不清的同一面部图像

早期的建立模式识别系统的一些工作,实际上是出于生物学上的动机。最常见的历史例子就是所谓的感知机和自适应线性组合器(ADALINE)设备。这些研究的目的,就是研制出一个识别系统,它的结构和工作方式同人的识别系统相仿。在 Minsky 和 Pappert[1969]的书中表现出的认为感知机能力有限的悲观结论,使这个领域的研究大大地减少了。后来,随着更有力的神经技术的出现,神经网络研究领域再一次充满活力。在人工神经网络和联接主义学说领域中,目前发生的重大事情使人联想到神经计算机研究兴起的早年时期。

早年对感知机方法的悲观失望,使许多研究人员把注意力集中到格式化模式信息处理方面的数学或计算机科学领域。例如:研究重点转移到了统计模式识别和具有按句法规则结构的模式分类领域。神经网络或联接主义学说,提供了通向真正拥有智能能力的计算机系统这一充满希望的道路。最近十年在人工神经网络领域中的进展,已经使我们离创造出能够像人一样工作的系统这一目标更近了。

Jain 和 Mao[1994]很好地讨论了人工神经网络方法和统计模式识别方法之间的共同之处。

总之,神经网络是有效的和能满足预期需要特性的天然分类器。这些特性包括:

1. 噪声抑制能力。
2. 对失真图像或模式的容错能力(再生能力)。
3. 对具有丢失信息或低质量图像的超级识别能力。
4. 并行处理潜力。

1.3 模式识别序言

一个模式可以由一组 n 个数字的测量值一个 n 维模式或测量值矢量 Z 来表示:

$$\mathbf{Z} = z_1, z_2, \dots, z_{N_m} \quad (1-1)$$

然后,一个特征矢量 $\mathbf{X}$ 可以取自模式矢量:

$$\mathbf{X} = x_1, x_2, \dots, x_N \quad (1-2)$$

从而,一个模式可以看作是一个在 N_m 维测量超空间,或者是在 N 维特征超空间中的一个点。典型的特征空间的维数,要选得比相应的测量空间维数低。模式分类,即把一个模式正确地由特征/测量空间映射到一个类空间中。所以,模式分类的判决过程可以归纳如下。

考虑一个由 n 维特征矢量来表示的模式:

$$\mathbf{X} = (x_1, x_2, \dots, x_n)^T \quad (1-3)$$

其中 T 表示转置。

模式识别的任务,就是判别出一个模式属于 K 个类别 $C_1, C_2, \dots, C_K$ 中的哪一个。注意到测量矢量代表检测到的数据,其中 N_m 是测量值的数目。例如,如果一幅图像由 $m \times m$ 具有 16 个灰度级的象素阵列来表示,那么可以确定模式矢量的维数, $n = m^2$ 。假设矢量 $\mathbf{Z}$ 的每个元素 z_i 是来自于 16 个可能灰度值中的一个相应的灰度值。

在语音模式识别问题中,声信号是时间的函数。感兴趣的实体是变量 t 的连续函数,而不像前面的例子中它是离散的灰度值。为了完成这种类型的分类,我们必须首先测量可观察到的抽样值特性。在这种情况下,这些抽样包括观测一段时间内的语音波形。一个模式矢量可以通过以离散的时间间隔 $t_1, t_2, \dots, t_n$ 等对这些函数进行抽样来形成。图 1.5 示出了对波形进行抽样的值 $z(t_1), z(t_2), \dots, z(t_n)$ 。

用于语音识别的特征矢量,可以由捕获波形的前 N 个傅立叶系数构成。

模式识别系统的设计也包括选择一种合适的模式描述方法,这种描述应考虑到是机器可以接受的形式。这种决定也受到识别系统所应用问题范围的影响。例如,在人脸识别问题中,图像可以被转化成一个象素阵列,象素具有通过光敏矩阵设备(或具有帧抓取功能的摄像机)来表示的灰度。在彩色编码应用中,使用红、蓝、绿(RBG)信号的强度来描述是更合适的。

所以,特征提取器首先应设计成可以找出用于表示输入模式的合适特征,以便在特征空间中增大来自不同类的模式之间的差别。在定义了特征集和选择好适当的特征提取算法之后,典型的识别过程包括两个阶段:训练和预测。一旦特征空间的映射建立起来,训练过程就可以开始。在这个问题相关范围内有代表性的训练数据是必须要获得的。调节识别机器以便它能将特征矢量(取自训练数据)映射到分类误差最小的类别中。在第二个阶段(预测阶段)中,训

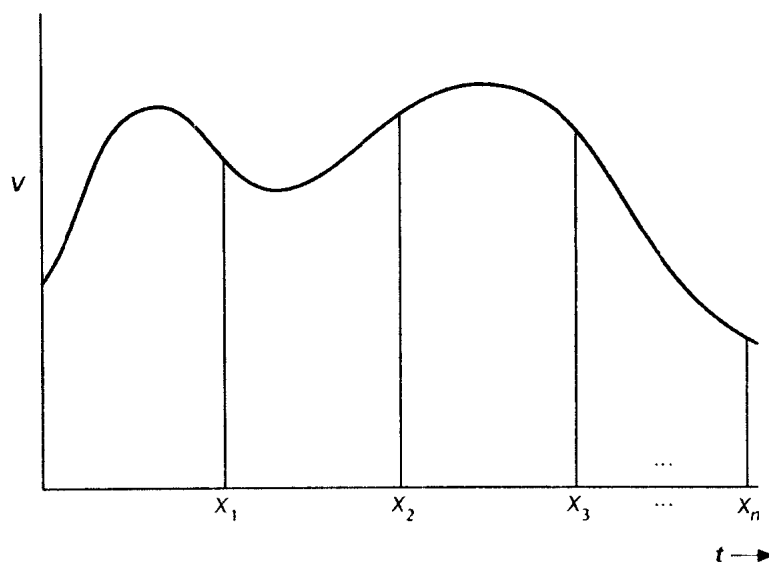


图 1.5 以离散时间间隔波形的抽样

练过的分类器把未知输入模式根据提取出的特征矢量分到某一类别中。如果预测结果不可接受,则需要重复处理,特征选择需要重做或者在不同参数下的训练需要重新执行。

两个原始数据表示(比特图或在字识别情况中的笔划)都不特别适合于直接输入到神经识别器。就像将会看到的,对变化的范围上讲“坏”的程度与讨论中的识别器特性是不同的。在上面的情况下,使用原始数据输入形式作为直接输入送到神经识别器中,存在如下一些问题:

1. 它们是非正交的。
2. 它们不可能表示出被识别模式的显著特征。
3. 它们是冗长的。多余的大输入矢量会导致需要更大的网络,使得在训练和识别过程中网络性能下降。
4. 它们对图像中的微小变化,即各种字体或笔划的变化很敏感。
5. 它们可能包含许多无关的或不相干信息,从而需要提供给识别器过多的训练。
6. 对于旋转变换,它们不会是恒定的,等等。

对于给定的这些问题,必须注意采用其它更好的数据表示方法把数据提供给识别器。

1.4 统计模式识别

在这个领域,模式分类问题可以被系统地阐述成统计判决问题。统计模式识别是一种相对成熟的理论,并且已经基于这种方法设计了许多商用的识别系统。关于这个领域,几本早期的书仍是可借鉴的,如:Tou 和 Gonzalez, 1974; Duda 和 Hart, 1973; Fu, 1977; Fukunaga, 1990。Pao[1989]的著作是一本相当好的书,它给出了从实际工程应用前景来寻找最合适技术这一思想。这些书把模式识别问题表示成多维空间中密度函数的估计问题,并且把这个超空间分成多个类别或分类区域。在这种情况下,判决是使用合适的判别函数来完成的。从而数学统计成为这个领域的基础。

由于这个学科利用判别函数来划分模式空间,那么它也可称为决策理论方法。这些函数也可称为判决函数,它是模式 X 的标量函数。以被判别函数所规定的边界包围的模式空间中