

编者按:模式识别可以使人们在影响因素很多的情况下仍能对众多信息进行方便的处理,进而对物群进行分类。而在工业中使用模式识别,则可以节约原材料、提高产品质量和产量、降低单产能耗,解决企业和工程技术人员所关心的问题。从本期开始,“模式识别”专栏将陆续邀请国内外知名专家向广大读者介绍模式识别的发展和具体应用,希望使更多读者认识和了解模式识别对自动化技术的推动。



模式识别简述

A Brief Introduction to Pattern Recognition



严红平

女,博士后,中国地质大学(北京)信息工程学院副教授,主要研究方向为模式识别、计算机图形学、图像处理。

(中国地质大学(北京)信息工程学院,北京 100083) 严红平
(中国科学院自动化所模式识别国家重点实验室,北京 100080) 潘春洪

1 序言

人们在观察事物或现象的时候,常常要根据一定需求寻找观察目标与其他事物或现象的相同或不同之处,并在此特定需求下将具有相同或相似之处的事物或现象组成一类。例如字母“A”、“B”、“a”、“b”,如果从大小写上来分,会将“A”、“B”划分为一类,“a”、“b”划分为另一类;但是如果从英文字母发音上来分,则又将“A”、“a”划分为一类,而“B”、“b”则为另一类。另外,不同人写的“A”、“B”、“a”、“b”都不同,但即使人们从未见过某个人写的“A”、“B”、“a”、“b”,或者这些字符出现在混乱的背景里,或部分被遮盖,人们也可以正确地分出它们,并根据需要将它们进行准确归类,当然,前提条件是人们对“A”、“B”、“a”、“b”一般的书写格式、发音方式等有所了解。人脑的这种思维能力就构成了“模式识别”的概念。那么,什么是模式?什么是模式识别呢?

2 模式和模式识别

从以上的例子可以看出,对字符的准确识别首先需要在头脑中对相应字符有个准确的认识。当人们看到某物或现象时,人们首先会收集该物体或现象的所有信息,然后将其行为特征与头脑中已有的相关信息相比较,如果找到一个相同或相似的匹配,人们就可以将该物体或现象识别出来。因此,某物体或现象的相关信息,如空间信息、时间信息等,就构成了该物体或现象的模式。Watanabe^[16]定义模式“与混沌相对立,是一个可以命名的模糊定义的实体”。比如,一个模式可以是指纹图像、手写草字、人脸、或语言符号等。“广义的说,存在于时间和空间中可观察的事物,如果我们可以区别他们是否相同或相似,都可以称之为模式”^[16]。而将观察目标

与已有模式相比较、配准,判断其类属的过程就是模式识别。模式以及模式识别是和类别(集合)的概念分不开的,只要认识某类事物或现象中的几个,人们就可以识别该类中的许多事物或现象。为了强调能从具体的事物或现象中推断出总体,我们把通过对具体的个别事物进行观测所得到的具有时间和空间分布的信息称为模式,而把模式所属的类别或同一类模式的总体称为模式类(我们下面进行的模式识别的讨论都是基于该定义的)。也有人习惯上把模式类称为模式,把个别具体的模式称为样本^[6]。如“字符”、“植物”、“动物”等等都是模式,而“A”、“松树”、“狗”则是相应模式中的一个样本。在此意义上,人们可以认为把具体的样本归类到某一个模式,就叫做模式识别,或模式分类。

人类具有很强的模式识别能力。通过视觉信息识别文字、图片和周围的环境,通过听觉信息识别与理解语言等。模式识别是人类的一种基本认知能力或智能,是人类智能的重要组成部分,在各种人类活动中都有着重要作用。在现实生活中,几乎每个人都会在不经意间轻而易举地完成模式识别的过程。但是,如果要让机器做同样的事情,恐怕决非这么轻松。文中,笔者将从人工智能的角度,更深层次地分析什么是模式识别,以及如何用机器进行模式识别。

要让机器具有人的模式识别能力,人们首先需要研究人类的识别能力,因此模式识别是研究人类识别能力的数学模型,并借助于计算机技术让计算机模拟人类识别行为的科学。换言之,模式识别是研究如何让机器观察周围环境,学会从背景中识别感兴趣的模式,并对该模式的类属作出准确合理的判断。模式识别研究主要集中在两方面,即研究生物体(包括人)如何感知对象,以及研究在给定的任务下,如何用计算机实现模式识别的理论和方法。前者属于认知科学的范畴,是生理学家、心理学家、生物学家和神经生理学家的研究内容,后者属于信息科学的范畴,是数学家、信息学专家和计算机科学工作者的研究内容。识别行为可以分为两大类:识别具体事物和识别抽象事物。具体事物的识别涉及到时空信息的识别。空间信息的例子,如指纹、气象图和照片等;时间信息的例子,如波形、信

号等。抽象事物的识别涉及到某一问题解决办法的识别、一个古老的话题或论点等。换言之,抽象事物的识别是识别那些不以物质形式存在的现象,属于概念识别研究的范畴。笔者所指的模式识别主要是对具体事物的识别,如语音波形、地震波、心电图、脑电图、图片、文字、符号、三位物体和景物以及各种可以用物理的、化学的、生物传感器进行测量的具体模式等。要识别的数据有:一维数据,如语音、心电图、地震数据等;二维数据,如文字图片、医学图像、卫星图像等;三维数据,如图像序列、结晶学或X像断层摄影术等。

3 模式识别系统

一个完整的模式识别系统基本上由三大部分组成,即数据采集、数据处理和分类决策或模型匹配,如图1所示。在设计模式识别系统时,需要注意模式类的定义、应用场合、模式表示、特征提取和选择、聚类分析、分类器的设计和学习、训练和测试样本的选取、性能评价等。针对不同的应用目的,模式识别系统三部分的内容可以有很大的差异,特别是在数据处理和模式分类这两部分,为了提高识别结果的可靠性往往需要加入知识库(规则)以对可能产生的错误进行修正,或通过引入限制条件大大缩小待识别模式在模型库中的搜索空间,以减少匹配计算量。在某些具体应用中,如机器视觉,除了要给出被识别对象是什么物体外,还要求出该物体所处的位置和姿态以引导机器人的工作。下面笔者分别简单介绍模式识别系统这三部分的工作原理。



图1 模式识别系统框架图

(1) 数据采集

数据采集是指利用各种传感器把被研究对象的各种信息转换为计算机可以接受的数值或符号(串)集合。习惯上,称这种数值或符号(串)所组成的空间为模式空间。这一步的关键是传感器的选取。为了从这些数字或符号(串)中抽取对识别有效的信息,必须进行数据处理,包括数字滤波和特征提取。

(2) 数据处理

数字滤波是为了消除输入数据或信息中的噪声,排除不相干的信号,只留下与被研究对象的性质和采用的识别方法密切相关的特征(如表征物体的形状、周长、面积等等)。举例来说,在进行指纹识别时,指纹扫描设备每次输出的指纹图像会随着图像的对比度、亮度或背景等的不同而不同,有时可能还会产生变形,而人们感兴趣的仅仅是图像中的指纹线、指纹分叉点、端点等,而不需要指纹的其它部分或背景。因此,需要采用合适的滤波算法,如基于块方图的方向滤波、二值滤波等,过滤掉指纹图像中这些不必要的部分。

特征提取是指从滤波数据中衍生出有用的信息,从许多特征中找出最有效的特征,以降低后续处理过程的难度。比如,图像识别时,提取的特征有灰度变化、纹理、形状等。我们对滤波后的这些特征进行必要的计算(比如进行快速傅里叶变换以得到信号功率谱)后,通过特征选择和提取或基元

选择形成模式的特征空间。那么,如何判别什么特征是最有效的呢?人类很容易获取的特征对于机器来说就很难获取了,这就是模式识别中的特征选择与提取的问题。特征选择和提取是模式识别的一个关键问题。一般情况下,候选特征种类越多,得到的结果应该越好。但是,由此可能会引发维数灾害,即特征维数过高,计算机难以求解。因此,数据处理阶段的关键是滤波算法和特征提取方法的选取。不同的应用场合,采用的滤波算法和特征提取方法以及提取出来的特征也会不同。

(3) 分类决策或模型匹配

基于数据处理生成的模式特征空间,人们就可以进行模式识别的最后部分:模式分类或模型匹配。该阶段最后输出的可能是对象所属的类型,也可能是模型数据库中与对象最相似的模式编号。模式分类或描述通常是基于已经得到分类或描述的模式集合而进行的。人们称这个模式集合为训练集,由此产生的学习策略称为监督学习。学习也可以是非监督性学习,在此意义下产生的系统不需要提供模式类的先验知识,而是基于模式的统计规律或模式的相似性学习判断模式的类别。模式分类或模式匹配的方法有很多,主要是基于以下思想设计的:

成员表:即模板匹配。基于该思想,分类系统中会预先存储属于同一模式类的模式集,然后将输入的未知模式与系统中已有的模式相比较,具有相同或相似匹配的模式类即为该未知模式的所属类型。

一般特征:这里模式的一般特征被存储在一个分类系统中,当有一个未知模式进入该系统时,系统会将其一般特征与系统中现有类的一般特征相比较,并将其归入到与其有相似特征的类中。

聚类:文中笔者用实数向量来表示目标类的模式,这样,利用其聚类特性,可以轻易地将未知模式进行分类。如果目标向量在几何位置上相距很远,就容易确定未知模式的类别。但是如果目标向量相距较近,或甚至有重叠,人们就需要采用比较复杂的算法来确定未知模式的类别。最小距离分类法就是一个基于聚类概念的简单算法。该算法通过计算未知模式与希望的已知模式集之间的距离,来决定哪一个已知模式与该未知模式最近,并最终将该未知模式归入到与其相距最短的已知模式类中。该算法对于目标向量在几何位置上相距很远的模式分类很有效。

神经元:上面的模式分类思想都是基于机器的直接计算,而直接计算则是基于数学相关的技术。仿生学是指将生物学知识应用到电子机器中。神经网络方法就是将生物知识应用于机器中来模式识别,从而引进了人工神经网络。

一个神经网络是一个信息处理系统,由大量简单的数据处理单元组成,这些单元互相连接,协同工作,从而实现大规模并行分布处理。神经网络的设计和功能是模仿了生物的脑部和神经系统的功能而设计的。神经网络具有自适应学习、自组织和容错力等优点。由于神经网络的这些突出特点,人们可以应用神经网络进行模式识别。一些最好的神经网络模型是后向传播网络、高阶网络、时延和周期性网络。

通常,人们利用前向传播网络进行模式识别。前向传播也就是没有回到输入端的反馈信息。与人类从错误中得到教训

相似,神经网络也能通过向输入端反馈信息,从其错误中得到教训。通过反馈可以重建输入模式,避免产生错误,从而提高神经网络的性能。当然,构造这样的神经网络非常复杂。这类神经网络要用到后向传播算法(BP)。后向传播算法的主要问题之一是局部极小问题。另外,神经网络在学习速度、结构选择、特征表示、模块性、缩放性等方面也都存在一些问题。虽然神经网络存在这样那样的问题和困难,但是其发展潜力还是巨大的。

基于上面的思想,形成了几种常见的模式识别方法:模板匹配、统计(决策论)模式识别、句法(或结构)模式识别、模糊模式识别和神经网络模式识别。

4 模式识别方法

(1) 模板匹配

模板匹配是最早出现,也是最简单的模式识别方法之一。匹配是模式识别的一种分类操作,主要是判断同一类的两个实体(如点、曲线、形状等)之间的相似性。要进行模板匹配,首先需要存储一些已知模板,然后考虑所有可能的变化,将待识别模板与已知模板相比较,从而得出二者之间的相似性度量。已知模板一般是通过训练得到的。模板匹配方法在字符识别、人脸识别等领域有广泛的应用,但是该方法计算量非常大,而且该方法的识别率严重依赖于已知模板。如果已知模板产生变形,会导致错误的识别,由此产生了可变形模板匹配方法。

(2) 统计模式识别

自上世纪60年代用统计决策理论求解模式识别问题以来,统计模式识别方法得到了迅速的发展,70年代前后出版了一系列反映统计模式识别理论和方法的专著^{[7][9][10][12][20]}。到目前为止,统计模式识别的理论体系已经相当完善。统计模式识别,又称决策理论识别方法,该方法根据模式的统计特征,用一个 n 维特征空间(特征集)来描述每个模式,然后基于概率论、数理统计以及矩阵理论和向量代数的知识,利用合适的判别函数(每个模式类的特征值分布函数),将这个 n 维特征空间划分为 m 个区域,即类别。特征值分布函数可以通过指定或学习得到。比如,字符识别器确定一个模式的类别为“a”到“z”26类中的一个。同样地,在进行签名的有效性验证时,人们将某一签名确定为“真实”或“伪造”。统计模式识别技术对于解决分类问题非常有用。在统计模式识别中,贝叶斯决策规则^[2]从理论上解决了最优分类器的设计问题,但其实施却必须首先解决更困难的概率密度估计问题。

(3) 句法(结构)模式识别

1962年,R.Narasimahan提出了一种基于基元关系的句法模式识别方法,傅京孙在这个领域进行了卓有成效的工作,形成了句法模式识别的系统理论。句法(结构)模式识别主要是基于特征的结构相关性进行内部模式结构的描述。比如,图像分析常常涉及到图像的描述而不仅仅是分类。一个描述包括图像基元的信息以及这些信息之间的关系。句法模式识别(结构模式识别的一种)利用句法、句法分析和自动推理机理论来描述和分析一个模式的结构,是相对较成熟的模式识别方法。一

维字符串的语法分析可以在许多方面进行拓宽,从而应用于二维和三维模式的识别。错误信息和不确定信息的句法处理是目前的研究热点。统计模式识别和句法模式识别是模式识别领域的两大主流研究方向。

(4) 模糊模式识别

1965年L.A.Zadeh的《模糊集合论》(《Fuzzy Sets》)宣告了模糊数学的诞生,从那以来,有关模糊信息处理的理论和应用取得了重大进展,并由此产生了模糊模式识别方法。模糊模式识别是基于模糊数学的模式识别方法。现实世界中存在许多界限不分明、难以精确描述的事物或现象,而模糊数学则可以用数学的方法研究和处理这类具有“模糊性”的事物或现象。模糊数学的出现使得人们可以模拟人类神经系统的活动,描述模式属于某类的程度,因此,模糊数学在模式识别中得到了很好的应用。模糊聚类分析是非监督模式识别的重要分支。1994年,Randas利用模糊聚类从原始数据中直接提取特征,并对提取出来的特征进行优选和降维操作,以免造成维数灾害。马少平在汉字的方向像素特征进行了模糊化描述,是特征变化比较平滑,提高了汉字特征描述的准确度。在模糊模式识别中,隶属函数的选取是关键。

(5) 人工神经网络模式识别

上世纪50年代末,F.Rosenblatt^[16]提出了一种简化的模拟人脑进行识别的数学模型—感知机,初步实现了通过给定类别的各个样本对识别系统进行训练,使系统在学习完毕后具有对其他未知类别的模式进行正确分类的能力。80年代,J.Hopfield深刻揭示出人工神经网络所具有的联想存储和计算能力,为模式识别技术提出了一种新的途径,短短几年在很多方面就取得了显著成果,从而形成了人工神经网络模式识别方法。神经元模式识别利用神经网络中出现的神经计算模式进行。大部分神经网络都有某种训练规则,如基于现有模式调节连接权重。换句话说,神经网络直接对例子进行学习,得出其结构特征进行推广,就像孩子从狗的例子中认识狗一样。

人工神经网络可以超越传统基于计算机的模式识别系统的能力。人们可以利用计算机或神经网络进行模式识别。计算机利用传统的数学算法来检测给定的模式是否跟现有模式相匹配。这是一个简单易懂的方法。但是,该方法只能进行是或非的判断,且不允许模式有噪声。另一方面,神经网络允许模式可以有噪声,而且如果训练得当,神经网络会对未知模式的类别做出正确的响应。虽然神经网络不能创造奇迹,但是如果采用合适的结构,对好的数据进行正确的训练,不仅在模式识别领域,而且在其他科学或商业应用中,神经网络都可以给出令人惊异的结果。比如,BP神经网络直接从观测数据(训练样本)学习,非常简便有效,因而获得了广泛应用,但它是一种启发式技术,缺乏指定工程实践的坚实理论基础。

模糊模式识别和神经网络模式识别是新近发展起来的模式识别方法,是信息科学和人工智能的重要组成部分。在过去的几十年里,人们对模糊数学、人工智能和基于规则的专家系统的兴趣高涨。在这些研究领域里,模式识别起着重要作用。

实际上,现在的专家系统和模式分析之间存在着许多交叉。而模式识别的核心,包括“学习技术”和“推理”在人工智能中也起着非常重要作用。模式识别中的视觉理解、情景分析、图像理解对于机器人视觉也是绝对必要的。另一方面,人工智能中的方法,如知识表示、语义网络和启发式搜索算法,也能被用在许多模式识别问题中来改善模式描述和匹配,从而产生“聪明的”模式识别。此外,像语音或图像这样的感觉数据总被认为是人工智能领域的重要分支,同时它们也是模式识别的研究热点。

表1对上面介绍的这几种模式识别方法进行了简单的归纳总结。实际上,上述模式识别方法并非完全独立、互不相干的,这些方法互相渗透、互相补充。在许多新兴的应用领域,没有唯一最优的方法,必须同时使用几种不同的模式识别方法。人们已经尝试设计融合了许多识别方法的模式识别系统。

表1 常见的模式识别方法

| 方法 | 表示模式 | 识别函数 | 判别准则 |
|----------|----------|-----------|------|
| 模板匹配 | 样本、像素、曲线 | 相关性、距离度量 | 分类误差 |
| 统计模式识别 | 特征 | 类属判别函数 | 分类误差 |
| 结构模式识别 | 基元 | 规则、语法 | 接受误差 |
| 模糊模式识别 | 特征 | 隶属函数 | 隶属度 |
| 神经网络模式识别 | 样本、像素、特征 | 非线性信号处理函数 | 均方误差 |

5 结语

随着计算机软硬件技术的快速发展,模式识别得到越来越多的关注,模式识别技术也越来越完善,并在越来越多的领域得到了成功应用,如数据挖掘、文献分类、财政预测、多媒体数据库的组织和检索、生物(比如根据人的物理特征,如人脸、指纹等识别人)、医学(医学图像分析)、地质、能源、气象(天气预报)、化工、冶金、航空(卫星航空图片解释)、工业产品检测等领域。近年来发展最快的模式识别领域应属于计算机视觉与听觉领域,如手写文字识别,生物特征识别(包括:指纹识别、虹膜识别、视网膜识别、掌纹识别、人脸识别、手掌静脉分布识别等),印刷品版面分析识别,互联网有害信息检测,语音识别(语音输入系统,语音应答系统)等。Picard^[13]提出了模式识别的又一应用—情感计算,使得计算机可以像人类一样识别和表达感情、对人类的情感做出睿智的响应,并可以利用情感机制进行决策。美国微软公司主席比尔·盖茨认为人类计算的未来就是要让计算机看、会听、会说、会思考;美国总统信息技术顾问委员会PITAC(President's Information Technology Advisory Committee)在其提交给联邦政府关于“面向未来的信息技术研究”的咨询报告中,强调模式识别和感知等方面的基础研究将极大地改善人类与机器之间的交互。如果有一天,模式识别发展到了可以普遍应用阶段,将给人们的生活、工作带来很大的便利。模式识别不仅是一个有重大科学意义的研究领域,也是网络数字化时代诸多实际应用中需要突破的一项重大关键技术。

模式识别是一个快速发展的学科,因此很难对该领域的最新研究进展做一个很全面详尽的总结。模式识别从上世纪20年代发展至今,已经有了许多行之有效的模式识别方法来解决不同问题,但是还没有发展成统一的、有效的可应用于所有问题的模式识别理论。模式识别研究的目的是要开发出一

的不依赖于应用领域的数据分析技术,使得机器可以像人类一样分析解决各种问题。这是一个难以实现的目标,人们当前要做的工作是结合具体问题提出新的模式识别方法,或者根据各种现有模式识别方法的特性,将不同的模式识别方法融合在一起解决同一问题的不同方面,互相取长补短,开创模式识别研究和应用的新局面。

其他作者:潘春洪,男,博士后,现就职于中科院自动化所模式识别国家重点实验室工作,主要研究方向为计算机视觉、图像处理和模式识别。

参考文献

- [1] Arkadev A.G. and Braverman E.M., Teaching Computers to Recognize Patterns, Academic Press, London, 1966.
- [2] Bernardo J. and Smith A., Bayesian Theory, John Wiley & Sons, 1994.
- [3] Bezdek J.C., Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, New York, 1982.
- [4] Bezdek J.C. and Pals, Eds. S.K., Fuzzy Models for Pattern Recognition: Methods that Search for Structures in Data, IEEE Press, 1992.
- [5] 边肇祺, 张学工, 等编著. 模式识别(第二版)[M]. 北京:清华大学出版社, 2000.
- [6] Bishop C.M., Neural Networks for Pattern Recognition, Clarendon Press, Oxford, 1995.
- [7] Devijver P.A. and Kittler J., Pattern Recognition, a Statistical Approach, Prentice Hall, Englewood Cliffs, London, 1982.
- [8] Duda R.O., Hart P.E. and Stork D.G., Pattern Classification (2nd ed.), John Wiley and Sons, 2001.
- [9] Friedman M. and Kandel A., Introduction to Pattern Recognition, statistical, structural, neural and fuzzy logic approaches, World Scientific, Singapore, 1999.
- [10] Fukunaga K., Introduction to Statistical Pattern Recognition (Second Edition), [11] Gonzalez R.C. and Thomason M.G., Syntactic pattern recognition - An introduction, Addison-Wesley, Reading, 1982.
- [12] Jain A. K., Duin R. P. W., Jianchang Mao, Statistical pattern recognition: a review, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000; 22(1): 4 ~ 37.
- [13] Picard R., Affective Computing, MIT press, 1997.2.
- [14] Rosenblatt, F., Principles of Neurodynamics: Perceptrons and the theory of brain mechanisms, Spartan Books, Washington, D.C., 1962.
- [15] Stephen I. Gallant, Neural Network Learning and Expert systems, Massachusetts Inst. of Technology, Cambridge, Massachusetts, 1993.
- [16] Watanabe S., Pattern Recognition: Human and Mechanical, Wiley, New York, 1985.
- [17] Watanabe S., Frontiers of Pattern Recognition, Academic Press, New York, 1972.
- [18] 王碧泉, 陈祖荫. 模式识别理论、方法和应用[M]. 北京:地震出版社, 1989.
- [19] 王守觉. 仿生模式识别(拓扑模式识别)——一种模式识别新模型的理论与应用[N]. 电子学报, 2002, 30(10): 1417 ~ 1420.
- [20] Webb A.M., Statistical pattern recognition, Oxford University Press Inc., New York, 1999.

编号:060104