

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

王守觉

(中国科学院半导体研究所神经网络实验室, 北京 100083)

摘 要: 本文提出了一种模式识别理论的新模型, 它是基于“认识”事物而不是基于“区分”事物为目的, 与传统以“最佳划分”为目标的统计模式识别相比, 它更接近于人类“认识”事物的特性, 故称为“仿生模式识别”. 它的数学方法在于研究特征空间中样本集合的拓扑性质, 故亦称作“拓扑模式识别”. “拓扑模式识别”的理论基点在于它确认了特征空间中同类样本的连续性(不能分裂成两个彼此不邻接的部分)特性. 文中用“仿生模式识别”理论及其“高维空间复杂几何形体覆盖神经网络”识别方法, 对地平面刚体目标全方位识别问题作了实验. 对各种形状相像的动物及车辆模型作全方位 8800 次识别, 结果正确识别率为 99.75%, 错误识别率与拒识率分别为 0 与 0.25%.

关键词: 模式识别; 神经网络; 仿生学; 高维几何

中图分类号: O235 **文献标识码:** A **文章编号:** 0372-2112 (2002) 10-1417-04

Bionic (Topological) Pattern Recognition — A New Model of Pattern Recognition Theory and Its Applications

WANG Shou-jue

(Lab of Artificial Neural Networks, Institute of Semiconductors, CAS, Beijing 100083, China)

Abstract: A new model of pattern recognition principles, which is based on “matter cognition” instead of “matter classification” in traditional statistical pattern recognition, has been proposed. This new model is better closer to the function of human being, rather than traditional statistical pattern recognition using “optimal separating” as its main principle. So the new model of pattern recognition is called the Bionic Pattern Recognition. Its mathematical basis are topological analysis of sample set in the high dimensional feature space, therefore it is also called the Topological Pattern Recognition. The basic idea of this model is based on the fact of the continuity in the feature space of any one of the certain kinds of samples. We did experiments on recognition of omnidirectionally oriented rigid objects on the same level, with the Bionic Pattern Recognition using neural networks, which acts by the method of covering the high dimensional geometrical distribution of the sample set in the feature space. Many animal and vehicle models (even with rather similar shapes) were recognized omnidirectionally thousands of times. For total 8800 tests, the correct recognition rate is 99.75%, the error rate and the rejection rate are 0 and 0.25 respectively.

Key words: pattern recognition; neural networks; bionics; high dimensional geometry

1 引言

模式识别的发展已有几十年的历史. 早在 20 世纪 30 年代, Fisher 提出的判别分析就是利用已知的两类向量的概率分布函数来设计将两类向量分开的决策规则的问题^[1]. 随即在统计模式识别发展以后, 模式识别就被认为是用来选择使平均风险达到最小的模式类别. 因而, 划分类别问题, 与统计判决理论完全相同^[2].

近年受到人们关注的支持向量机(SVM)的最早提出者 Vapnik 在 1974 年提出了“最优分类超平面”的概念^[3]并在此基础上后来发展构造了支持向量机(SVM)^[4]. 综上所述, 数十年来人们在研究模式识别问题中, 考虑的出发点都是在若干类别的最佳分类划分上. 其根本原因或许在于用这样的数学描述与处理方法最具有—般性、通用性. 但众所周知, 即使基于目前最先进的模式识别理论基础上的识别机, 其实际效

果却仍然远不能令人满意. 虽然 Vapnik 在一本非常出色的书中是这样批评人工智能学者的^[5]. (为了忠于原文, 把 Vapnik 书中(中译本)原话一段引在下面, 这对说明本文的目的与出发点是必要的.) Vapnik 写道: “在 80 年代, 人工智能学者在计算学习领域中扮演了主要的角色. 在这些人工智能研究者中间, 一些较极端的学者有很大的影响(正是他们强调了‘复杂的理论是没有用的, 有用的是简单的算法’). 这些人工智能学者对于处理学习问题有丰富的经验, 善于对一些理论上非常复杂的问题构造‘简单的算法’. 60 年代末, 人们认为在几年之内就可以完成计算机自然语言翻译器(但时至今日我们尚离解决这个极其复杂的问题很远); 在此之后的另一个计划是构造通用问题求解器; 再之后又是建立大系统自动控制机的计划, 等等. 所有这些科研计划都没有取得成功. 接下来要研究的问题就是如何建立一种计算学习技术. 这些学者首先做

的是改变了所用的术语.特别地,感知器被改称为神经网络.然后这些研究被称作是与生理学家共同进行的,对学习问题的研究减少了一般性,增加了主观色彩.在 60 和 70 年代学习问题研究的主要目标是寻找从小数量样本出发进行归纳推理的最好途径,而到了 80 年代,目标变成了构造利用大脑来推广的模型.(原文注:当然研究人是如何学习的是很有意义的,但是这并不一定是建立人工学习机器的最佳途径,正如人们对鸟类如何飞行的研究实际上对建造飞机没有多少帮助一样).”

本文对 Vapnik 这段话并非有所评议,只想说明,本文题目用“仿生模式识别”中的“仿生”二字,并非 Vapnik 所批评的“目标变成了构造利用大脑来推广的模型”,而只是为了从功能上更接近于人类对事物的识别过程,而从数学模型上加入了实际“被识别对象”的某些普遍存在的“特殊性”.这个所谓“特殊性”是针对目前传统的模式识别数学描述方法的“一般性与通用性”而言的,它在实际的模式识别问题中却仍然具有一般性与通用性.

本文的基本出发点就在于把模式识别问题看成模式的“认识”,而不是分类划分,不是模式分类;是一类一类样本的“认识”,而不是多类样本的划分.因而,其基本数学模型当然会与传统模式识别的“最优分类”界面的概念大不相同.

2 人类的识别功能与传统模式识别方法的差异

设想农村中的一孩子,能认识牛、羊、马、狗、树林、房子等等农村中所有的一切,但没见过汽车;如果有一辆汽车突然出现在他面前时,试问他会怎样想呢?他会把汽车与以前见过的牛、羊、马、狗……等一件件比较,看与哪件最相像吗?(传统模式识别正是这样的)不,他绝不是这样想的,而是认为“我没见过这种东西”“我

不认识这种东西”.这正是人类认识事物与传统模式识别的差别.人类侧重于“认识”,只有在细小之处才重视“区别”(例如要区分狼与狗或马与驴等等),而传统模式识别则只注意“区分”,没有重视“认识”的概念.本文提出的“仿生模式识别”就是从“认识”模式的角度进行模式识别;既然它的基本概念不同于已经发展数十年的传统模式识别,为强调与传统模式识别在概念上的不同,用了“仿生模式识别”这一名词;其中“仿生”二字的含义只是在模式识别的功能和数学模型上强调了从“认识”的概念出发,更接近于人类的识别;而从实现的途径上看,还没有“仿生”的含意.下面把仿生模式识别与传统模式识别的比较与差异以图表方式表示在图 1 中.

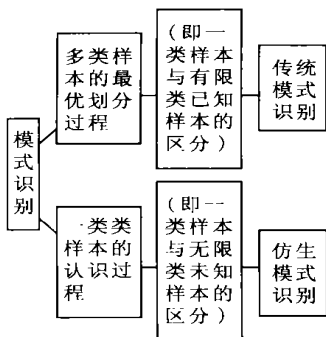


图 1

3 从基本数学模型看仿生模式识别与传统模式识别的本质差别

传统模式识别其基本数学模型的基点在于如文献[5]中第 12 页所述的:《所有可用的信息都包含在训练集(1-1)式中的情况下》(注:原文(1-1)式即训练集坐标 $(x_1, y_1), \dots, (x_1, y_1)$)这就是说在同类样本点相互之间没有任何关系的先验知识存在,因此,一切都只能从特征空间中不同类样本的划分出发了.然而在自然界实际规律中并非如此;本文所提出的仿生模式识别就在于引入了同类样本间存在的某些普遍的规律性而建立起来的“多维空间中非超球复杂几何形体覆盖”的识别原理.下节将重点讨论同类样本间普遍存在的规律性(亦即普遍存在的先验知识)

4 仿生模式识别的基点——特征空间中同类样本全体的连续性规律

人们在对事物的识别过程中,存在两种不同的归类过程.一种是“同源”的归类过程,例如同为简体汉字的手写体识别.另一种是不同源的按逻辑知识归类过程,例如手写体中简体汉字与繁体汉字的归并.对学过简体汉字而未学过繁体汉字的人,对于简体汉字的印刷体、手写楷书和行书等往往都能认识,而对于繁体汉字却不认识.因为繁体汉字与简体汉字并不源于同一个字形,而是被人为地定义为同一个字的,靠事先学习的逻辑知识把二类归并成一类.我们把“同源”的称为同一类,而在识别以后再按逻辑归类知识把需要归在一起的类别合并.因而以下所称同类样本指的都是“同源”的一类样本.以下讨论“同源”同类样本的连续性规律:

自然界任何欲被认识的事物(包括事物、图像、声音、语言、状态等等)若存在两个“同源”同类而不完全相等的事物,而这两个事物的差别是可以渐变的或非量子化的(宏观世界的绝大部分事物正是如此,即使像“苹果产量”那样不可能细化到一只苹果以下,但以重量计的“苹果产量”是考虑为可以渐变的)则这两个同类事物之间必至少存在一个渐变过程,在这个渐变过程中间的各事物都是属于同一类的.以数学公式描述为:

特征空间 R^n 中设所有属于 A 类事物的全体所做成的点集为 A ,若集合 A 中存在任意两个元素 x 与 y ,则对 ϵ 为任意大于零的值时,必定存在集合 B ,使

$$B = \{x_1, x_2, x_3, \dots, x_n \mid x_1 = x, x_n = y, n \in N, \rho(x_m, x_{m+1}) < \epsilon, \epsilon > 0, n-1 \geq m \geq 1, m \in N\}, B \subset A$$

在特征空间 R^n 中同类样本点之间所存在的这个连续性规律是超出了传统模式识别与学习理论的基本假定的,该假定认为“可用的信息都包含在训练集中”.但这个连续性规律却是客观世界中人类直观认识范围的客观存在的规律;因而也是仿生模式识别中用来作为样本点分布的“先验知识”,从而来提高对事物的认识能力.

传统模式识别中把不同类样本在特征空间中的最佳划分作为目标,而仿生模式识别则以一类样本在特征空间的分布的最佳覆盖作为目标,以二维空间的情况示意如图 2.

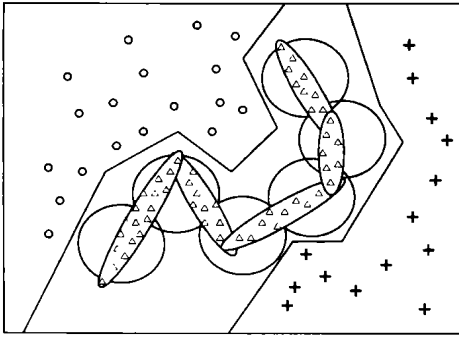


图 2

图中,三角形为要识别的样本,圆圈和十字形为与三角形不同类的两类样本,折线为传统 BP 网络模式识别的划分方式,大圆为 RBF 网络的划分方式(等同于以模板匹配的认识方式)细长椭圆形构成的曲线代表仿生模式识别的“认识”方式.因而,仿生模式识别把分析特征空间训练样本点的关系作为基点,而特征空间中样本分布的连续性规律为此提供了可能性.

5 仿生模式识别理论的数学工具问题

在仿生模式识别中引入了特征空间中同类样本的连续性规律后,对一类事物的“认识”,实质上就是对这类事物的全体在特征空间中形成的无穷点集合的“形状”的分析和“认识”.正如前苏联多名院士的数学名著^[6]中所指出的“任何对象的每一个集合,只要在有着关于连续性、关于无限接近的自然概念,就是一个拓扑空间.拓扑空间的概念是非常普遍的,而关于这种空间的学问——抽象拓扑学——不是别的,正是关于连续的最普遍的数学分支”.而本文所提出的仿生模式识别的理论分析数学工具正是点集拓扑学中对高维流形的研究问题.这同传统的以数理统计为基础的模式识别显然在数学工具的基础上有根本的差别.与传统的统计模式识别相别,把仿生模式识别也称为拓扑模式识别.

在仿生模式识别(拓扑模式识别)中,特征空间 R^n 中任何一类事物(如 A 类)全体在 R^n 中映射(必须是连续映射)的“象”所做成的点集被视为都是一个闭集,这个闭集我们称它为 A ,集合 A 根据仿生模式识别的具体应用对象可以是不同维数的“流形”.关于一般点集的维数分析在前苏联数学家乌利逊(1898~1924)所建立的“一般维数论”中已有讨论.但模式识别是一门实际应用的科学.在实际工程中,无论模式识别欲解决的具体问题是什么,在采集样本和识别对象时必然会带来一定的随机“噪声”.因而,在实用的仿生模式识别中,对 A 类事物“认识”的判别覆盖集合应该用集合 P_a 取代集合 A ,

$$P_a = \{x | \rho(x, y) \leq k, y \in A, x \in R^n\}, k \text{ 为选定的距离常数}.$$

集合 P_a 是 n 维的,因而仿生模式识别的任务,就是判别“被识别事物”映射在特征空间 R^n 中的“象”是否属于集合 P_a .

6 神经网络高维空间复杂几何形体覆盖识别方法及其应用实例

在实际的仿生模式识别中为了判别是否属于集合 P_a ,必须用软件或硬件为手段,在特征空间 R^n 中构筑一个能覆盖集合 P_a 的 n 维空间几何形体.近似于覆盖集合 P_a 的 n 维空间几何形体是以不同维数的“流形”(集合 A)中,无穷多的点作球心,以常数 k 作为半径的无穷多个 n 维超球体的并,即集合 A 与 n 维超球体的拓扑乘积.

根据维数理论^[7],要把 n 维空间分成两部分,其界面必须是一个 $n-1$ 维的超平面或超曲面,而人工神经网络的一个神经元正是在 n 维空间中作一个 $n-1$ 维的超平面或超曲面,把 R^n 分成两个部分.一个神经元,也可以是多种多样的复杂的封闭超曲面^[8].因而,人工神经网络是实现仿生模式识别的十分合适的手段.

为了方便发展神经网络仿生模式识别,我们在前一篇论文中^[9]引入了神经网络高维空间几何分析方法,用来作为发展仿生模式识别的一种实用性工具.该文中对 n 维空间的点、直线、平面、超平面、圆、球面、超球面间的关系作了叙述,但未对非球超曲面进行讨论.以下,将介绍和讨论一个应用非球超曲面的仿生模式识别的实例.

仿生模式识别应用实例的要求是在海面上或地平面上对不同方向观察的目标(如舰艇、坦克、汽车、牛、马、羊等)的认识.样本的采集是从不同方向观察所采集到的 bmp 文件,进行前处理(连续映射)后压缩成 256 维特征空间样本点.由于观察方向都是水平的,可以说方向的改变只有一个变量,因而,特征空间中样本点的分布应近似于呈一维流形分布.加以其他方向存在的微弱变动,可以考虑某类对象在特征空间中的覆盖形状应是一个与圆环同胚的一维流形与 256 维超球的拓扑乘积.用语言描述也就是在 256 维特征空间中,离开一条头尾相接的空间曲线的最小距离小于某定值 k 的所有点的集合 P_a ,而该空间曲线包含所有采集的样本点集合 S ,即

$$S = \{x | x = S_i (i=1, 2, \dots, \text{采集样本总数})\}$$

$$P_a = \{x | \rho(x, y) \leq k, y \in A, x \in R^n\}$$

$$A = \{x | x = x_i, i = (1, 2, \dots, n), n \leq N, \rho(x_m, x_{m+1}) < \epsilon,$$

$$\rho(x_1, x_n) < \epsilon, \epsilon > 0, n-1 \geq m \geq 1, m \leq N\}, A \subset R^n, S \subset A$$

为了实际用神经网络中若干神经元来实现近似覆盖 P_a ,我们用若干空间直线段来近似空间曲线 A ,每个神经元覆盖的是一个直线段与 n 维超球的拓扑乘积.

设采集原始样本集合为 S ,在其中取样本 j 个组成新的样本集 S' ,使

$$S' = \{x | x = S'_i (i=0, 1, 2, \dots, j), \rho(s'_i, s'_{i+1}) \leq d <$$

$$\rho(s'_{i-1}, s'_{i+1}), d \text{ 为设定常数}, s'_0 = s'_j\}, S' \subset S$$

用 j 个神经元来近似地覆盖 P_a ,第 i 个神经元覆盖范围 P_i 为:

$$P_i = \{x | \rho(x, y) \leq k, y \in B_i, x \in R^n\}$$

$$B_i = \{x | x = \alpha S'_i + (1 - \alpha) S'_{i+1}, \alpha = [0, 1]\}$$

$$\text{全部 } j \text{ 个神经元覆盖范围为: } P'_a = \bigcup_{i=0}^j P_i$$

7 实验与结果

以狮子、老虎、坦克、小汽车等八件实物模型(见图 3),各以旋转 360° 一圈采集 400 个样本,称为第一样本集八类共 3200 个样本,再在不同时间下另行重复采集一次,得第二样本集八类 3200 个样本。

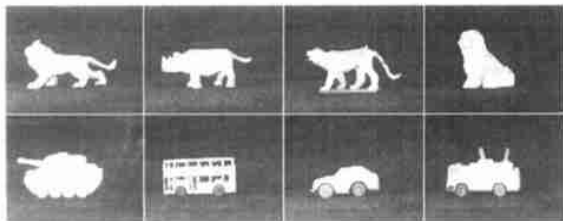


图 3

再用猫、小狗、斑马等六件实物模型(见图 4),同样各采集 400 个样本,称为第三样本集六类共 2400 个样本。试验步骤与结果如下:

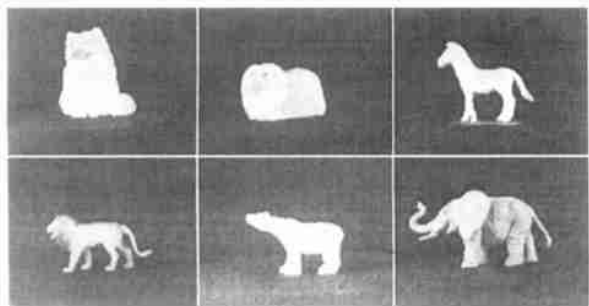


图 4

(1)首先在第一样本集 3200 个样本中,按同类样本间的距离大小每类分别选择 26 至 50 个不等数量的样本作为训练样本,共八类 338 个样本作为训练集。

(2)以统一的距离参数 k ,分别用神经网络构造 P_a, P_b, \dots 等八类对象拓扑流形的识别网络,其神经元数量与训练集样本数量相同。

(3)把第一、第二两样本集共 6400 个样本作为正确识别率的测试样本,结果正确识别的样本总数为 6384 个,正确识别率为 99.75%,未正确识别的 16 个样本全部属拒识,拒识率为 0.25%。

(4)把第一、第二、第三样本集全部 8800 个样本作为错误识别率的测试样本,结果无一样本被误识。

由于参数 k 的选择与 P_a, P_b, \dots 等八类识别网络的建立等工作的描述与讨论需占用太大的篇幅,将在另文发表。

8 讨论与结论

由以上实验结果说明:

(1)仿生模式识别对于未经过训练的任意对象决不会误识(无一次误识,即使动物模型的后影是十分相似的)这正是人们识别功能的特点,也是传统模式识别的致命弱点。

(2)仿生模式识别是一类一类分别训练“认识”的,对新增加样本的训练不会影响原有识别知识,这是一大优点。

(3)正确识别率 99.75%,剩下的 0.25% 全部是拒识(而不是误识),这在实际应用中往往没有影响,只要重复一次即可,不会造成误差。

(4)该实验实例是拓扑模式识别中最简单的“一维流形实例”,即每一类样本的全体在特征空间中呈与圆环同胚的“一维流形”状分布(亦即空间闭合曲线),它与 n 维超球的拓扑乘积 P_a 易于用神经网络覆盖。

(5)上述情况说明研究具体的仿生模式识别首先要分析研究被识别事物在特征空间中分布状况的拓扑属性,再考虑用神经网络覆盖方法。

(6)在上述实验实例中,首先保证了预处理(即特征提取)过程是连续映射,无跳变现象,这是拓扑模式识别的基本要求,否则将会影响效果。

(7)上述实验中用的识别对象实物模型,为消除光线色度的影响和反光的影响(正如一般摄影技术中要避免反光一样),这些实物都采用无色毛面,而摄像采样只用了黑白灰度。

(8)综上所述可以说明,仿生模式识别开辟了一个效果较好有前途的新方向,正有待于从理论到实践上研究发展众多具体针对性的应用领域。

参考文献:

- [1] Fisher R. A. Contributions to Mathematical Statistics [M]. New York: J. Wiley, 1952.
- [2] 陈季镛(美)著,邱炳章,邱华译.统计模式识别[M].北京:北京邮电学院出版社,1989.
- [3] Vapnik V. N and Chervonenkis A. Ja. Theory of Pattern Recognition [M]. Nauka, Moscow, 1974.
- [4] Boser B, Guyon I and Vapnik V. N. A training algorithm for optimal margin classifiers [A]. Fifth Annual Workshop on Computational Learning Theory [C]. Pittsburgh: ACM, 1992. 144—152.
- [5] Vladimir N Vapnik 著,张学工译.统计学习理论的本质[M].北京:清华大学出版社,2000.
- [6] A D 亚历山大洛夫等著,王元等译.数学——它的内容、方法和意义[M].北京:科学出版社,2001.
- [7] Ryszard Engelking. Dimension Theory [M]. PWN-Polish Scientific Publishers-Warszawa, 1978.
- [8] 王守觉,李兆洲,陈向东,王柏南.通用神经网络硬件中神经元基本数学模型的讨论[J].电子学报,2001,29(5):577—580.
- [9] 王守觉,王柏南.人工神经网络的多维空间几何分析及其理论[J].电子学报,2002,30(1):1—4.

作者简介:

王守觉 男,1925 年生于上海,早年就读于西南联大和同济大学,毕业后在北平研究院镭学研究所从事氧化亚铜研究,解放后改为中国科学院应用物理所结晶学室,1960 年成立半导体所后历任器件室主任、副所长、所长等职务,1980 年当选为学部委员(院士),现为中国电子学会副理事长,《电子学报》主编,中国半导体学科奠基人之一,现从事半导体超高速电路与人工神经网络算法、模型、硬件和应用的研究。