

多层自适应模式识别系统模型

杨国庆 陈松灿 吕 军

(南京航空学院计算机系, 南京, 210016)

A MULTI-LAYER ADAPTIVE PATLERN RECOGNITION SYSTEM MODEL

Yang Guo-qing, Chen Song-can, Lu Jun

(Department of Computer Nanjing Aeronautical Insfitute, Nanjing, 210016)

摘 要 在英国 WISARD 单层模式识别系统的基础上, 借助 P.Kanerva 的稀疏分布存贮 (SDM) 的概念, 提出了一种新的多层自适应模式识别系统模型 (MAPR)。并就其工作过程和主要特点作了较详细的叙述, 还列出了多体印刷体汉字识别的初步试验结果。MAPR 用稀疏 RAM 阵列代替 WISARD 的常规 RAM 阵列, 用对 n 元模式计频的训练策略代替了原系统的直接置位策略。使系统除了保持原系统的重要优点外, 在大维数或非确定性模式数据识别方面, 其性能有了明显改善。

关键词 模式识别, 稀疏分布存贮, 并行分布处理, 神经网络

Abstract Based-on a single-layer pattern recognition system (WISARD) and referred to the concept of a sparse distributed memory (SDM) by P.Kanerva, this paper presents a novel multi-layer adaptive pattern recognition system model and describes in detail its processing procedures and main features. The results of primitive multi-font printed Chinese character recognition experiment are given.

Unlike WISARD, MAPR replaces the normal RAM array with sparse RAM array, and the directly setting bit strategy with counting frequency training for n -tuple pattern, so that not only MAPR maintains the main advantages of WISARD, but its performance is obviously improved as well in the aspect of recognizing large-dimensional or non-deterministic pattern data.

Key words pattern recognition system, sparse distrobuted memory

传统的模式识别方法几乎都是基于算法、在冯诺依曼机器上串行实现的。其通用性、适应性、实时性都受到很大的限制。因此, 探索和模拟人的模式识别机制、研究非算法极度并行分布式模式识别系统的结构, 已是感兴趣的研究内容。英国 I.Aleksander 等人发明的 WISARD 系统是这种系统的雏型^[1,2], 是一种非算法、用 RAM 网络实现的极度并行系统, 能以每秒 25 帧的速度处理电视分辨率的实际图象, 具有良好的识别特性和应用前景。然而, 这种系统也有其固有的某些限制, 如硬件成本、非确定模式数据处理能力等, 影响了系统的性能, 限制了使用范围。P.Knerva 提出的稀疏分布存贮 (SDM) 模型可以认为是推广的 Hopfield 神经网络模式, 实现模式的联想^[3,4]。本文在 WISARD 系统的 n 元模式识别原理的基础上, 结合 SDM 的概念, 提出了一种新的多层自适应模

1990年6月2日收到, 1991年5月31日收到修改稿

863 高科技资助课题

式识别系统模型, 并通过简例, 详细地介绍了这样系统的训练和分类过程, 简单地分析了这种系统性能上的改善, 最后还给出了多体印刷体汉字识别的初步结果。

1 WISARD 系统模型

WISARD 系统逻辑结构如图 1 所示。

输入模式一般是 $m_1 \times m_2$ 的二值数据阵列。这可以是一幅图象、一个字符或其它类型的模式数据。随机地或确定地分成 R 个子模式, 每个子模式为 n 维大小, 故称 n 元子模式。每个 n 元子模式分别连接到具有 2^n 个存贮单元的 RAM 地址端。这样, $k(k = m_1 \times m_2 / n)$ 个内容寻址的 RAM 部件构成了一个分类器的存贮空间。对于 C 类问题, 一般需要 C 个分类器与之对应。各分类器均按相同方式与输入模式连接。它们的训练单独进行, 而分类则同时进行。当未知模式输入系统时, 各分类器的每个 n 元所对应 RAM 的输出经求和逻辑形成分类器的输出响应。然后由决策逻辑根据响应对未知模式进行分类决策。具体工作过程如下

在训练前, 系统各分类器的 RAM 置为 0。训练阶段, 被训练的分类器所有 RAM 部件置 '写' 状态。当训练样本的 n 元模式数据寻址 RAM 时, 被选单元被置 1。一个模式类数据训练完毕时, 该分类器的 RAM 恢复成 '读' 状态。以相同的方法重复进行其它分类器的训练。分类阶段, 各分类器的 RAM 置为 '读' 状态。当未知模式输入系统时, 各 n 元子模式同时寻址各分类器的 RAM, 相应单元的内容被读出。每个分类器所有 RAM 的输出之和形成了该分类器对未知输入模式的响应。其最大响应为 K 。决策逻辑根据各分类器的响应对未知模式进行分类, 其分类决策为最大响应决策原理。决策输出包括最大响应、最大类别号(决策 D)和置信度(最大响应和次大响应之差)等。

和传统的模式识别方法相比, 这种系统的主要特点有:(1) 自适应性、通用性强;(2) 并行处理能力强;(3) 训练样本集的大小和 RAM 容量无关。 n 元子模式可以看作作为特征, 它的大小对分类性能影响较大, 并主要决定 RAM 容量, 但这些子模式自身并不保存在 RAM 中;(4) 松散地基于神经元模型。从信息的分布式存贮方式、大规模并行的处理能力、RAM 部件同时具备的信息存贮和处理功能和系统的自适应能力看, 系统可以认为是一种松散基于 McCulloch-pitts 模型的神经网络, 称之为逻辑神经网络。

有关这种系统的工作原理、理论行为、应用结果见文献[1,2,5,6]

2 稀疏分布存贮模型

联想记忆是神经网络的重要研究内容之一。所谓联想是指对已记忆(或存贮)的一组模式 $(X^{(\alpha)}, Y^{(\alpha)})_{\alpha=1}^M$, 当用地址 $X^{(\alpha)}$ 或与 $X^{(\alpha)}$ 相似的地址 X 检索时, 产生对应的 $Y^{(\alpha)}$ 或与 $Y^{(\alpha)}$ 相

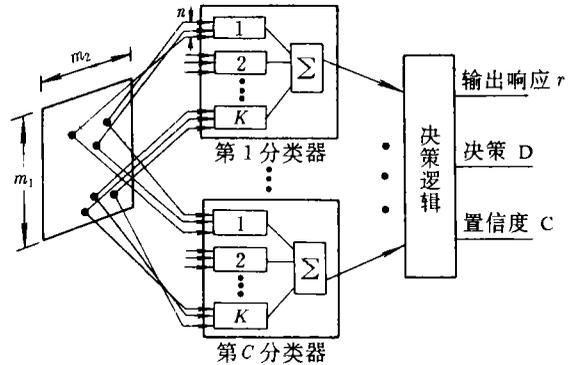


图 1 WISARD 系统模型

似的数据 Y 。对于一般的计算机存储器, $X^{(a)}$ 与地址对应, $Y^{(a)}$ 与 $X^{(a)}$ 寻址的存贮单元内容相对应。当地址长度为 n 位时, 具有 2^n 个不同的地址状态, 每个地址状态寻址一个确定的存贮单元。随着 n 的增长, 可能的地址状态呈指数增长。例如, 当 $n=1\ 000$ 时, 其地址数已是天文数字了。显然不可能在物理存贮单元与全部的、甚至是部分的地址状态之间建立联想关系。为此, P.Kanerva 提出了稀疏分布式存贮模型, 解决了大地址空间的联想问题^[3,4]。在 SDM 中, 从 2^n 个地址状态中, 随机地选取 m 个状态与确定的 m 个物理存贮单元对应。因为 $m \ll 2^n$, 所以这些被选的地址状态稀疏分布在地址空间中, 如图 2 所示。

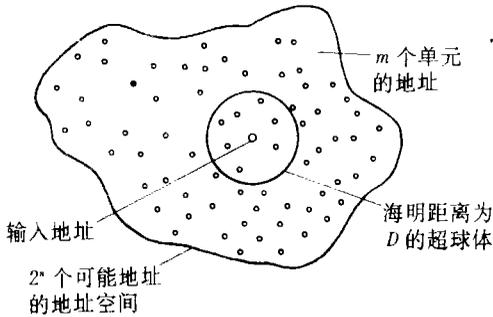


图 2 SDM 地址选择示意图

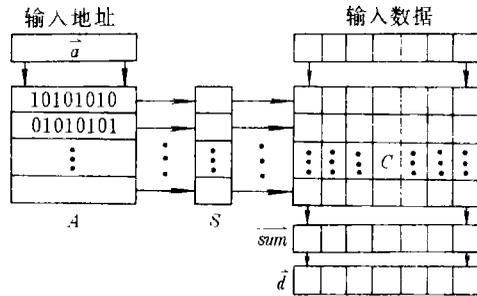


图 3 稀疏分布存贮模型

例如, $n=1\ 000$, 可以选取 $m=2^{20}$ 代表 $2^{1\ 000}$ 个地址的空间。在写操作过程中, 给定地址, 相应数据被写入与该地址相近的一些单元。这些单元的选择办法是: 在地址空间中, 以给定地址为中心, 海明距离 D 为半径作超球体, 落在超球体的地址状态点(指上述 m 个地址中的少数几个)所对应的单元被选中, 数据同时写入这些单元, 也即数据信息分布存贮在这些单元中。新数据的写入不是代替旧数据, 而是相加后再写入。在读操作过程中, 给定期望数据的地址, 以该地址为中心, D 为半径的超球体内的所有 SDM 的存贮单元被选中, 其内容被读出, 相加后取阈值产生输出数据。只要没有太多的其它数据写入存贮器, 该输出数据和原数据是相近的。由此实现了自联想。这种模型的结构如图 3 所示。

A 矩阵保存着预置的随机选择的 m 个地址。 C 矩阵由 m 个存贮单元构成, 用以保存数据, S 为 $m \times 1$ 矩阵。当输入给定地址向量 \vec{a} 时, A 中落在以 \vec{a} 为中心, D 为半径的超球体内的地址矢量置位 S 的相应位, 其余位为 0。 S 的这些为 1 的位寻址矩阵 C 的相应单元, 在写方式时, 输入数据写入相应的单元。在读方式时, 相应单元的数据经求和形成 \vec{sum} , 再经阈值运算后形成数据 \vec{d} 。如前所述, 只要没有太多的其它数据被写入这些单元, \vec{d} 和原写入数据是充分相似的。

3 多层自适应模式识别系统模型

3.1 体系结构

根据 WISARD 的 N 元原理, 借助 SDM 的概念, 提出了一种新的双层自适应模式识别系统模型, 其结构如图 4 所示。

和 WISARD 一样, 输入模式被映射为 K 个子模式, 每个子模式大小为 n 维, 这些 n 元子模式分别与 A 矩阵的各个子矩阵 $A_i(i=1,2,\dots,k)$ 相连, 每个 A_i 和 SDM 中的 A 矩阵结构相同, m 个单元, n 位字长。 A_i 的输出连接 S_i , K 个 S_i 构成矩阵 S , 其大小为 $Km \times 1$ 位。 C 为 $Kmb \times c$ 矩阵, 每个分类器有 Km 个单元, 字长为 b 位(b 的大小取决于训练集的大小, 一般不超过一个字节), c 为分类器数目。各个分类器以相同的方式与 S 并行连接。 C 的每个分类器输出经求和和运算后形成矢量 \overline{sum} , 再经阈值运算后形成输出决策矢量 \vec{d} 。在这种结构中, A 和 S 是各分类器的公共部分, 作为系统的第一层, 其中输入模式到 A 矩阵的映射基于 WISARD 的 N 元原理。 A 矩阵和 S 矩阵的构造遵从 SDM 概念。第二层的识别网络采用 SDM 的稀疏 RAM 阵列, 其分类策略遵从 N 元识别原理。当 $m=2^n$, $D=0$ 时, MAPR 则是 WISARD 模型。因此, WISARD 可以认为是 MAPR 的特殊情形。

3.2 工作过程

为叙述方便, 以简化两类分类器为例, 说明 MAPR 的工作过程。图 5(a)是这种分类器的结构。其中, 模式大小为 3×3 位, $n=3, m=4$ 。 n 元模式的形式为 $n_i = A_i B_i C_i, i=1,2,3$ 。 A 矩阵的每个子矩阵预置的内容皆为 000, 001, 011, 111。 C 矩阵仅是图 4 C 矩阵的一列(一个分类器)。 C 的所有单元输出连接到求和逻辑, 其输出 Y 经阈值运算形成分类决策。下面以图 5(b)的样本集为例说明工作过程(取海明半径 $\theta=1$)。

在训练阶段, C 的 RAM 为计算状态。取 x_1, x_2, x_3 为训练样本。当训练 x_1 时, $n_1=111$,在以 111 为中心, 1 为半径的球内, A_1 的 011、111 包括在内, S_1 被置成

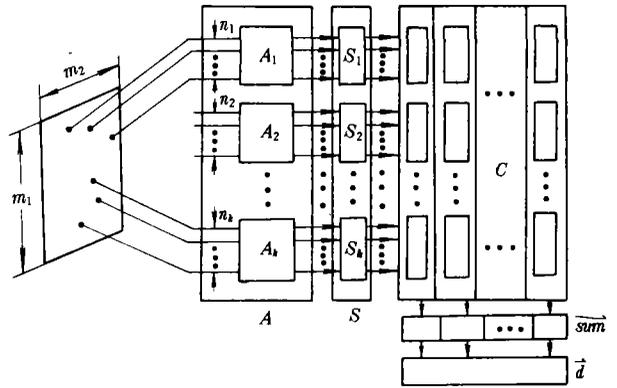


图 4 多层自适应模式识别系统模型

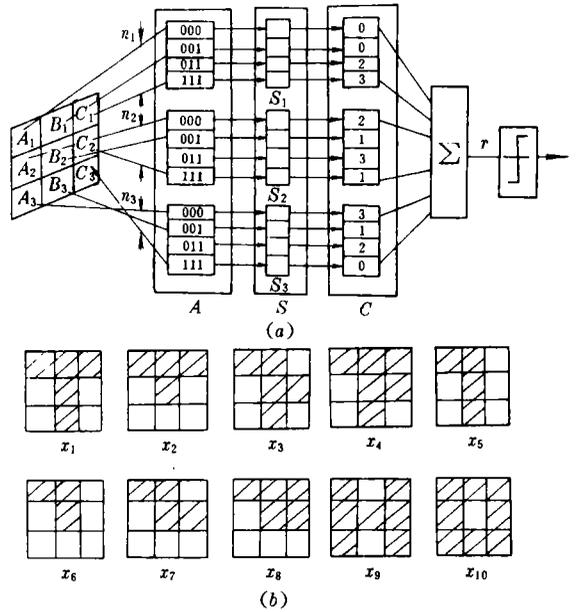


图 5 (a)简化两类分类器 (b)训练和测试样本

$[0,0,1,1]^T$, C_1 被相应地写成 $[0,0,1,1]^T$ 。同样, C_2 被写成 $[1,0,1,0]^T$, C_3 被写成 $[1,0,1,0]^T$ 。当训练 x_2 时, S_1 仍置成 $[0,0,1,1]^T$ 。 C_1 用新的内容 $[0,0,1,1]^T$ 和原内容相加得 $[0,0,2,2]^T$ 。同样, C_2 为 $[2,0,2,0]^T$, C_3 为 $[2,1,1,0]^T$ 。当训练完毕时, $C_1 \sim C_3$ 分别为 $[0,0,2,3]^T$ 、 $[2,1,3,1]^T$ 、 $[3,1,2,0]^T$ 。在测试阶段, C 为读状态。当输入模式为 x_1 时, S_1 为 $[0,0,1,1]^T$, C_2 的第 2、3 单元被寻址(对应 S_1 中为 1 的单元), 相应的内容被读出。同样, C_2 、 C_3 第 0.2 单元被读出。这些读出数据经求和逻辑输出响应 $r_{v_1} = 15$ 。图 5(b)的所有样本的响应如表 1 所示。

表 1 图 5(b)样本集测试响应结果

样本	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
响应	15	14	13	15	13	12	12	14	9	9
分类	ω_1							ω_2		

设分类法则为

$$x \in \begin{cases} w_1 & r_v \geq 10 \\ w_2 & r_v < 10 \end{cases} \quad (1)$$

则有表 1 的分类结果。这一结果是合理的, x_1, x_2, x_3 是训练集中字符 'T' 和它的变形, 而 $x_4 \sim x_8$ 是 $x_1 \sim x_3$ 中各 n 元子模式组合而成的, 可以把它们认为是训练集模式 $x_1 \sim x_3$ 的推广模式, 也是 'T' 的变形, 应该属于 ω_1 。在 WISARD 中, 这模式的响应都是最大⁽⁵⁾。而 x_9 是字符 'H', x_{10} 是字符 '0', 不应属于 w_1 。

上述简化模式仅是为了说明系统的工作过程。实际的 MAPR 将工作在模式维数较大的情况下, 如高分辨率图象识别等。才能真正体现这种系统的特点。

3.3 初步试验结果

一个包括 WISARD 和 MAPR 两种模型的仿真软件包已开发完成。这个功能较为完备的通用开发工具, 用于 MAPR 的结构研究及两种模型的应用研究。根据现有条件, 初步试验选取多种印刷体汉字识别问题。取 20 种汉字(20 类), 每种选宋、仿宋、楷、黑 4 种字体的样本各一个。共 80 个样本。样本大小为 24×24 阵列, 取 $n=14, m=64, D=5$ 。当采用重复置换法检验时("重复置换法"指样本集同时用于训练和测试^[7]), 其正确率为 100%。当采用"留一个出来"方法检验时("留一个出来"方法指 N 个样本中, $N-1$ 个样本用于训练分类器, 1 个样本用于测试, 共进行 N 次训练和测试), 其正确率为 92%。如果增加和优化训练样本, 其正确率下限还可以进一步提高。这是因为 WISARD 模型受成本限制, 只能选取 $n=2 \sim 8$, 而初步试验表明, $n=14$ 其性能最佳。

3.4 性能比较

关于 MAPR 系统和 WISARD 及其它模式识别模型的理论比较有待进一步研究。本文仅就该系统与 WISARD 系统作一些初步的比较。

3.4.1 成本 WISARD 系统需要的物理存储器容量为

$$V_w = m_1 \times m_2 / n \cdot 2^n \cdot c \quad (2)$$

MAPR 系统需要的物理存储器容量为

$$V_w = m_1 \times m_2 / n[m(n+1) + bc] \quad (3)$$

其中 $m_1 \times m_2$ 为模式大小, n 为 n 元子模式大小, $m_1 \times m_2 / n = k$, c 为分类器的个数, m 为稀疏地址矩阵的单元数, b 为 C 矩阵中每个分类器 RAM 的字长。

当分类器数目 c 很大时, (3)式可近似为

$$V_M \approx m_1 \times m_2 / nmbc \quad (4)$$

比较(2)式和(4)式, WISARD 系统随 n 呈指数增长, 而 MAPR 的成本随 n 呈线性增长。由于 $m \ll 2^n$, 则对于较大的 n , MAPR 系统的成本将远远低于 WISARD 系统。例如, 取 $m_1 \times m_2 = 512 \times 512$, $n = 16$, $m = 64$, $c = 16$, $b = 8$ 时, $V_W = 2^{34}$ 位, $V_M = 2^{27}$ 位, V_M 仅为 V_W 的 $1/128$ 。

3.4.2 推广特性 分类器的推广特性是指系统对训练中未出现的模式的正确识别能力。WISARD 系统的推广模式集指使分类器能作出最大响应的模式集合, 其大小为

$$\#G = \prod_{i=1}^K m_i \quad (5)$$

其中 m_i 是第 i 个 n 元的不同训练子模式个数, K 为 n 元个数。若用图 5(b) 的 x_1, x_2, x_3 训练 WISARD n 元大小。映射方法和图 5(a) 相同, 则 $\#G = 8$, 即对于图 5(b) 中的 $x_1 \sim x_8$, WISARD 都能得出最大响应 ($r = 3$)。在理想情况下, 一类的推广集应恰好包括了该类的全部模式数据, 这体现了系统良好的推广特性(智能特性)。但事实上, 推广集的大小与模式的差异、训练集的大小及 n 元的大小关系密切。对于非确定模式数据, 要求训练集较大, 这很容易形成存贮器的“饱和”, 也即 RAM 中多数单元被置 1, 当任一模式输入时, 各分类器的响应都很高, 从而降低了分类能力。

MAPR 的推广特性取决于 n 元子模式在 A 矩阵中的相关地址模式(A 矩阵中预置的地址模式)的多少及这些相关模式在训练中出现的频率。在非确定数据的分类中, 由于噪声模式的出现频率较低, 在训练阶段, 与之对应的相关地址模式的频率也较低。因此, 只要适当选择输出阈值, 就会有效地抑制这些模式。

3.4.3 自适应性和通用性 如前所叙述的, WISARD 具有较强的通用性和自适应性。由于结构上的密切相关, MAPR 也同样具有这种通用性和自适性。对于 WISARD 系统性能的优化, 主要取决于 n 元大小、映射方式、训练集大小和模式的性质。而 MAPR 系统, 在训练时可调节的参数除了上述内容外, 还有 m 的大小、 A 矩阵中预置地址模式的分布方式, 海明半径等。因此, 它比 WISARD 具有更强的自适应能力。

此外, 就硬件结构而言, WISARD 结构比较简单、规整, 实现比较容易。而 MAPR 的结构比较复杂, 实现比较困难。

4 结 论

在文献[8]中, 曾以“稀疏分布存贮概念和模式计算机”为小标题介绍了 SDM 概念的模式计算机, 并援引了美国计算机学会前主席 PD. Enning 的预言: 这种非数字、非字符处理的模式计算机在模式识别等领域有着广阔的发展前景。因为它在一定程度上模拟了人脑的模式识别过程。本文介绍的 MAPR 系统是在对 WISARD 系统多年研究的基础上, 利用 SDM 概念提出的一种模型。它使这种模式计算的设想更加具体化。和 WISARD 系统相

比,这种系统在大模式维数的情况下,如高分辨率图象识别,优越性更为突出。它的通用性、实时性和自适应性的特点,使这种系统在工业视觉、印刷体和手写体汉字识别、安全监视等多种应用领域具有很好的应用前景。有关这种系统的理论行为及应用的研究还有待进一步深入,可望不久将有更多的结果与读者见面。

研究生刘川为本文提供了多体印刷体汉字识别的试验结果,特此致谢。

参 考 文 献

- 1 Aleksander I, Thomas W V, Bowden P A. WISARD a radical step forward in Image Recognition Sensor Review, July, 1984. 120~124
- 2 Aleksander I, Stonham T J. A guide to pattern recognition using radom access memories IEEJ. Computer and Digital Tech. 1979,2(1):29~40
- 3 Kanerva P. Sparse distributed memory. Cambridge. MA: MIT Press, 1988.53~77
- 4 Keeler D J. Comparision between Kanerva's SDM and Hopfield-type Neural Networks Cognitive Science 12. 1988.229~329
- 5 杨国庆,夏振华,梅宏. 自适应数字学习网络. 微型计算机, 1988,(4):1~4
- 6 杨国庆,陈松灿,徐涛. 一种基于单层逻辑神经网络的通用自适应模式识别系统. 微电子学与计算机, 1990, 7(8):1~3
- 7 王碧泉. 模式识别—理论、方法和应用,北京:地震出版社, 1989. 192~199
- 8 李国杰. 面向知识表示的智能机. 模式识别与人工智能, 1988,2. 1~10