

混沌神经网络在分离叠加模式和多对多联想记忆中的应用^{*}

刘光远 段书凯

(西南师范大学电子与信息工程系 重庆400715)

A Chaotic Neural Network and its Applications in Separation of Superimposed Pattern and Many-to-Many Associative Memory

LIU Guang-Yuan DUAN Shu-Kai

(Dept. of Electronic & Information Engineering, Southwest China Normal University, Chongqing 400715)

Abstract In this paper, we propose a modified chaotic associative memory neural network(MCAM). It has two important features: it can recall stored patterns from superimposed input; (2) it can deal with many-to-many associative memory. The computer simulations show the effectiveness of the proposed model.

Keywords Chaotic associative memory neural network, Superimposed pattern, Many-to-many associative memory

混沌是不含外加随机因素的完全确定性的系统表现出来的介于规则和随机之间的内秉随机行为。脑神经系统是由神经细胞组成的网络。类似于人脑思维的人工神经网络与冯·诺依曼计算机相比,在信息处理方面有很大的优越性。混沌和神经网络相互融合的研究是从90年代开始的,其主要的目标是通过分析大脑的混沌现象,建立含有混沌动力学的神经网络模型(即混沌神经网络模型),将混沌的遍历性、对初始值敏感等特点与神经网络的非线性、自适应、并行处理优势相结合,用于联想记忆、模式识别和组合优化等,提高信息处理的效率和适应性。

用于联想记忆的混沌神经网络模型主要有两类:(1)基于耦合的混沌神经网络模型:1990年,Kaneko提出了全局耦合的网络模型(GCM)^[1],网络单元演化有聚类的特性。1991年,Inoue^[2]等提出的混沌神经元由若干个处理单元构成,每个处理单元由两个彼此耦合的混沌振荡子组成。(2)基于时空的总和的混沌神经网络模型:由E. R. Caianiello提出的Caianiello模型,南云(Nagumo)和佐藤(Sato)、Aihara^[3]、Y. Osana^[4]等对此模型作了进一步改进,建立了新的混沌神经网络模型。但是,已有的模型分离叠加模式、多对多联想速度慢,信息处理效率较低。本文提出一种改进的混沌神经网络模型(MCAM),该模型能分离叠加模式、实现多对多联想。仿真实验结果证明MCAM模型比以前的模型在分离叠加模式的速度和多对多联想效率方面有了明显的提高。本文分析了网络性能提高的原因,这有助于推进混沌神经网络在信息处理中的应用。

1 改进的混沌神经网络

1.1 Y. Osana 提出的模型

在历史上,最重要的神经元数学模型有Macculloch-Pitts模型(1943年)和Caianiello模型(1961年)。1971年,日本人南云(Nagumo)和佐藤(Sato)对此模型进行了修改。1990年,Aihara为了模拟真实神经元,在前人推导和实验基础上,提出了一种混沌神经模型,并以神经元为基本单元建立了相应的神经网络,即混沌神经网络模型。由 M 个混沌神经元构成的神经网络的第 i 个混沌神经元动力学方程为:

$$x_i(t+1) = f_i \left(\sum_{j=1}^M W_{ij} \sum_{r=0}^t k^r h_j(x_j(t-r)) + \sum_{j=1}^N V_{ij} \sum_{r=0}^t k^r I_j(t-r) - \alpha \sum_{r=0}^t k^r g_i(x_i(t-r)) - \theta_i \right) \quad (1)$$

其中, $x_i(t+1)$ 是第 i 个神经元在 $t+1$ 时刻的输出; W_{ij} 和 V_{ij} 分别表示神经元之间和外部输入与神经元之间的权值; h_j 是内部反馈函数; N 是外部输入数; $I_j(t-r)$ 是外部输入强度; f_i 是连续输出函数; g 是神经元输出与不应度之间的关系函数。 θ 是阈值, α 为分支参数, k^r 表示不应性衰减因子。1998年,Y. Osana和M. Hagiwara对此模型作了进一步改进,在多对多联想记忆模型(MAM)^[5]的基础上,用混沌神经元代替常规神经网络的常规神经元,用连续的输入代替初始化输入,建立了新的混沌神经网络模型。该模型具有分离叠加模式和实现多对多联想功能^[4],但分离的速度较慢。

1.2 改进的混沌神经网络

在分析Aihara模型的基础上,我们提出了一种改进的混沌神经网络模型(即MCAM),使联想过程更接近于人脑的思维过程和认知心理学原理。在Aihara等人提出的模型中,考虑了所有时间、空间的总和,即从 $t=0$ 到当前所有时间内的不应性、层内反馈、层间影响等因素,回忆成功率较低。认知心理学的研究成果表明,人脑记忆包括长时记忆和短时记忆,由于短时记忆存储的项目较少,信息提取速度较快,Seternberg的经典研究结果证实了这个规律^[6]。

在MCAM模型中,我们用部分的时空总和代替Y. Osana模型中的全部时空总和。当被记忆的模式作为神经网络外部输入时,网络围绕输入模式进行搜索,混沌神经元通过混沌改变它们的状态,从而实现分离叠加模式和实现多对多联想。本文所提出MCAM模型是一个类似Hopfield网络的全互连网络。其中第 α 层第 i 个神经元在 $(t+1)$ 的动力学方程为:

(1)当 $t > t_0$ 时,

$$x_i(t+1) = \Psi \left[v \sum_{d=t-t_0}^t k_d^{\alpha} I_i^{(\alpha)}(t-d) + \sum_{\beta=1}^L \sum_{j=1}^{N(\beta)} \omega_{ij}^{(\beta\alpha)} \sum_{d=t-t_0}^t k_d^{\beta} x_j^{(\beta)}(t-d) + \sum_{d=t-t_0}^t k_d^{\alpha} x_i^{(\alpha)}(t-d) - \sum_{j=1}^{N(\alpha)} \omega_{ji}^{(\alpha\alpha)} \sum_{d=t-t_0}^t k_d^{\alpha} x_j^{(\alpha)}(t-d) - \theta_i \right]$$

^{*} 本文受重庆市科委应用基础项目(项目号:99—5909)的资助。刘光远 博士,教授,研究方向为神经网络、计算智能等。段书凯 讲师,硕士,研究方向为神经网络。

$$\gamma \sum_{d=t-t_0}^t k_r^d x_i^{(a)} (t-d) - \theta_i^{(a)}]$$

(2)当 $t \leq t_0$ 时,将上式中 $d=t-t_0$ 改为 $d=0$ 即可。

其中,函数 $\psi(\cdot)$ 为具有陡度参数 ϵ 的 Logistic 函数 $\psi(\cdot)=2/(1+e^{-y/\epsilon})$, $\psi(\cdot)$ 的第一项表示外部输入,第二项表示不同层之间的连接(异联想),第三项表示同层中神经元之间的连接(自联想),第四项是神经元的不应性影响,最后一项是神经元的阈值。

1.3 分离叠加模式和全部记忆信息的恢复

在混沌神经网络中用被记忆的模式作为输入时,混沌神经网络能围绕这个模式进行搜索。当我们把训练模式1、2、3记忆到网络中以后,如果我们以模式1作为网络的输入时,网络围绕输入模式进行搜索,模式1能被回忆出来。当以模式1和模式2的叠加模式作为输入时,网络围绕模式1和模式2进行搜索,因为神经元能通过混沌改变其自身的状态,所以模式1和模式2能在不同的时刻分离回忆出来。

1.4 多对多联想

利用多目标联想记忆(MAM)模型不实现多对多联想。考虑训练集 $\{(X_1,Y_1,Z_1),(X_1,Y_2,Z_2),(X_3,Y_3,Z_3)\}$,其中包含共同项 X_1 。如将 X_1 作为第一层的初始化输入,第二层神经元的内部状态由以下公式决定:

$$I^{\beta}=X_1\omega_{(\alpha\beta)}=X_1X_1^TY_1+X_1X_1^TY_2+X_1X_1^TY_3=X_1X_1^T\cdot(Y_1+Y_2)+X_1X_1^TY_3$$

$\omega^{(\alpha\beta)}$ 是第一层与第二层之间的连接矩阵,在上式中 $X_1X_1^TY_3$ 是噪声项, (Y_1+Y_2) 是共同项 X_1 引起的叠加模式,同样在第三层会出现叠加模式 (Z_1+Z_2) 。因为 MAM 不能分离叠加模式,所以模式 (X_1,Y_1,Z_1) 和 (X_1,Y_2,Z_2) 不能分离出来。在本文提出的改进型混沌联想记忆模型 MCAM 中,网络通过混沌搜索,在第二层分离出 Y_1 和 Y_2 ,在第三层分离出 Z_1 和 Z_2 ,从而回忆出完整的模式 (X_1,Y_1,Z_1) 和 (X_1,Y_2,Z_2) ,实现多对多联想。

2. 仿真实验及讨论

本文做了相应的计算机仿真实验,实验结果证明本文提出的模型能分离叠加模式和恢复全部信息。实验设计的参数为:不应性衰减因子 $K_s=0.95$, $K_d=0.10$, $K_m=0.90$,分支参数 $\alpha=9.8$, $\nu=100$, Logistic 函数的陡度参数 $\epsilon=0.015$ 。模式信息用双极性代码1和-1表示。

2.1 分离叠加模式

我们设计了一个用49个神经元组成的联想记忆混沌神经网络,分别记忆字母、数字和汉字。当训练集(见图1)被记忆到网络中后,以叠加模式“(1”+“3”)作为外部输入,输出步数为30(即 $t=30$)。图2、3、4、5分别表示考虑不同范围内时空总和的输出情况。



图1 训练集1



图2 $t_0=0$ 时网络输出



图3 $t_0=3$ 时网络输出(MCAM 模型)



图4 $t_0=15$ 时网络的输出(MCAM 模型)



图5 $t_0=30$ 时网络输出(Y. Osana 模型)

由图2、3、4、5可以看出,如果只考虑当前状态的输出(即 $t_0=0$),网络没有完成分离,见图2所示;当考虑最近3、15次时空总和(即 $t_0=3、15$)时的输出(即 MCAM 模型),分离效果较好,如图3、4所示,其中 $t_0=3$ 时(即考虑最近三次的时空总和),效果最好;图5表示考虑全部时空总和时的输出(如 Y. Osana 模型),能分离模式“1”和“3”,但效果较差。

我们分析了在相同的输出步数($t=30$)时,分离效果与考虑时空总和范围的关系,如表1所示。从表1中可以看出,本文提出的模型 MCAM 在 $t_0=3$ 时在30次输出中分离成功次数为17次,分离成功率为56.7%,比 Y. Osana 模型的分离成功率提高了23.3%。

表1 分离效果与考虑时空总和范围的关系

实验序号	考虑时空总和范围(t_0)	分离成功次数	分离成功率
1	0	0	0
2	3	17	56.7%
3	15	13	43.3%
4	30	10	33.3%

2.2 多对多联想

本文设计了一个三层的改进型混沌神经网络 MCAM,每层包含81个神经元。本文将包含了共同信息“A”的模式(A,a,1)、模式2(A,b,2)、模式3(C,c,3)记忆到三层网络中,如图6所示。



图6 训练集2

图7、8、9分别表示当输出次数为50时,考虑不同范围内时空总和的多对多联想效果。当“A”作为第一层的输入时,分别联想模式1(A,a,1)和模式2(A,b,2)。其中,当 $t_0=0$ 时,没有实现多对多联想(如图7所示);当 $t_0=4$ 时,模式1出现8次,模

式2出现16次,多对多联想效果较好(如图8所示);当 $t_0=50$ 时(即 Y. Osana 模型),成功次数为4次,明显没有本文提出的 MCAM 模型效率高(如图9所示)。

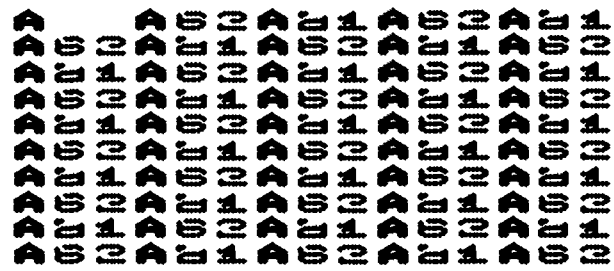


图7 $t_0=0$ 时网络的多对多联想输出

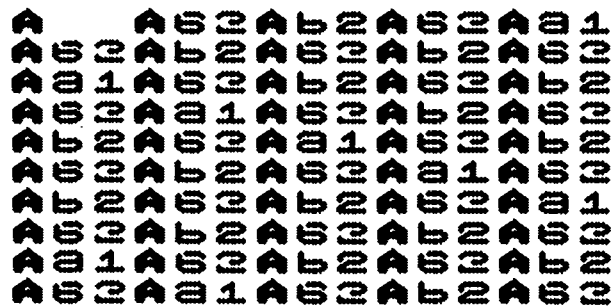


图8 $t_0=4$ 时网络的多对多联想输出(MCAM 模型)

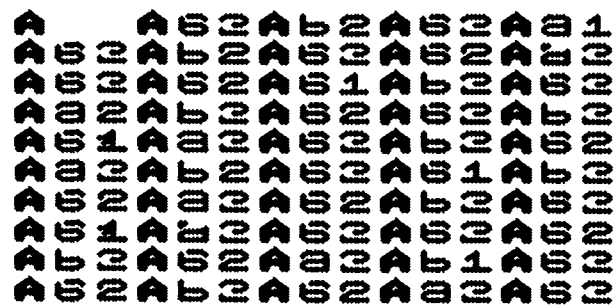


图9 $t_0=50$ 时网络的多对多联想输出(Y. Osana 模型)

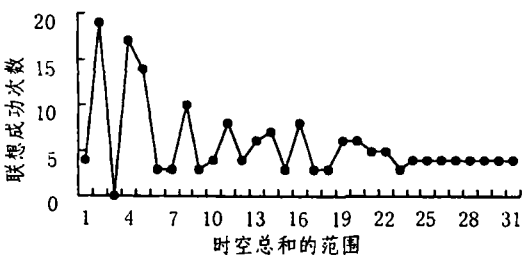


图10 时空总和范围与多对多联想成功次数关系图

为了比较模型中时空总和范围大小对多对多联想效率的影响,我们完成了30次输出中考虑 $t_0=0\sim30$ 所有情况的仿真实验.时空总和范围(即 t_0 的大小)与多对多联想成功次数关系图如图10所示。

图中数据显示:当时空总和范围选择恰当时,多对多联想成功机会就会增加(如 $t_0=1$ 时,联想成功19次; $t_0=3$ 时,联想成功17次; $t_0=4$ 时,联想成功14次)。当时空总和范围进一步增加时,联想成功机会减小,联想成功机会维持在4次左右。从图中还可以看出,当 $t_0=2$ 时,没有联想成功,这说明 t_0 的选择对联想效率有很大的影响。

我们进一步分析了对于不同输出步数多对多联想效果的变化,如表2所示。从表中可以看出,随着输出步数的增加,多对多联想成功的次数相应增加。可见本文提出的 MCAM 模型实现多对多联想的性能比较稳定。

表2 多对多联想效果与输出步数的关系

实验序号	输出步数	多对多联想成功次数	
		$t_0=3$ 时	$t_0=4$ 时
1	30	17	14
2	50	28	24
3	80	35	39

讨论 由以上的实验可以看出,用本文改进的混沌神经网络模型(MCAM)分离叠加模式时,输入的叠加模式都能在很快的时间里正确地分离出来。模式“3”能在 $t_0=3$ 时,模式“1”能在 $t_0=4$ 时从输入的叠加模式中完整地分离出来,与 Y. Osana 模型模型相比,分离效率有了很大的提高。当改进的模型用于多对多联想记忆时,联想效率提高了50%。实验证明了本文改进的模型在分离叠加模式和多对多联想记忆中的有效性。这是因为在我们改进的网络模型中使联想过程更接近于人脑的思维过程,用部分的时空总和代替 Y. Osana 模型中的全部时空总和,故本文改进的网络模型在叠加分离模式和多对多联想记忆方面有更强的能力。

参考文献

1 Kaneko K. Clustering, coding, switching, hierarchical ordering and control in a network of chaotic elements: [Physics D]. 1990. 137~172

2 Inoue M, Nagayoshi A. A chaos neuro-computer. Physics A, 1991, 158: 373~376

3 Aihara K, Takabe T, Toyoda M. Chaotic neural networks. Physics Letters A, 1990, 6, 7: 333~340

4 Osana Y, Hagiwara M. Separation of superimposed pattern and many-to-many associations by chaotic neural networks. IJCNN, Anchorage, 1998, 1: 514~519

5 Hagiwara M. Multidirectional associative memory. IJCNN, Washington, D. C. 1990, 1: 3~6

6 王 蓀, 汪安圣. 认知心理学. 北京大学出版社, 1992. 154