

# 用分维神经网络实现短期记忆<sup>1</sup>

李霞 谢维信

(深圳大学信息工程学院 深圳 518060)

**摘要** 神经网络中的突触连接矩阵储存有限的信息。当神经网络在线使用时, 基于 Hebb 规则的网络模型将不再适用。本文提出用分维神经网络实现短期记忆, 要求旧的模式随时间而遗忘, 最近存储的模式可以可靠记忆。实验结果表明, 该网络模型适用于在线工作, 且优于对比方法。

**关键词** 分维神经网络, 短期记忆, 突触连接矩阵

**中图分类号** TN-052

## 1 引言

记忆是一种过程, 它可分成三种模式: 感觉信息存储, 短期记忆和长期记忆。我们常有这样的体验, 当我们看到或听人叙述某一事项时, 往往只记住了最后几个内容, 而对前面的内容则有些模糊不清, 较难记住。可见, 短期记忆 (Short Term Memory) 具有如下重要特征: 容量极为有限, 随时间的经过产生忘却。生理数据显示短期记忆存在于自然界的生命中, 对短期记忆的深入研究有望应用于计算机的高速数据缓存器中。

近年来, 将神经网络用于联想记忆取得了极大的进展。由 Hopfield 提出的全连接无自反馈网络模型被广泛地研究。理论分析和实验结果均表明 Hopfield 网络只具有有限的存储能力<sup>[1,2]</sup>。当新的模式逐渐加入时, 网络最终会导致失控状态, 即没有一个模式可以被可靠地记忆。就短期记忆而言, 常要求在线使用。Morris-Wong 模型<sup>[3]</sup>是基于全连接 Hopfield 网络的短期记忆模型, 它需要大量的计算时间, 在线使用时效率低。事实上, 实际的生理系统常是稀疏连接的, 本文基于 Y. Baram 提出的具有自相似连通性的分维神经网络模型<sup>[4]</sup>来实现短期记忆。实验结果表明, 该网络模型适用于在线工作, 且优于 Morris-Wong 法。

## 2 分维神经网络模型

分维神经网络<sup>[4]</sup>具有层状结构, 其信息的存储是以子模式的形式完成的。分维神经网络由  $N$  个神经元组成, 它可分成每个子网具有  $p$  个神经元的  $L$  个子网, 每个子网内是全连接的。高层中的神经元只与低层中某一子网的神经元相连, 不同层的子网通常含有相同的结构, 亦即子网具有自相似连通性。假设  $K$  为分维神经网络的分层数, 则有

$$N = \frac{(p-1)^{K+1} - 1}{p-2}, \quad L = \frac{(p-1)^K - 1}{p-2}. \quad (1)$$

图 1 中所示为  $N=85, p=5, K=3$  时的分维神经网络结构。

设  $x_k$  表示第  $k$  个神经元的状态, 并取二值  $\pm 1$ , 网络状态由列向量  $X = \{x_1, x_2, \dots, x_N\}^T$  表示。  $X_{(j)}^{(i)}$  同样是一个  $N$  维列向量, 其第  $j$  个子网所对应的神经元的状态由第  $i$  个子模式决定, 其余元素为 0。对应于 Hebb 存储规则的分维网络的突触连接矩阵可表示为

$$T = \sum_{i=1}^{M_j} \sum_{j=1}^L X_{(j)}^{(i)} X_{(j)}^{(i)T}, \quad (2)$$

<sup>1</sup> 1998-02-09 收到, 1998-10-25 定稿

式中  $M_j$  为第  $j$  个子网中存储的子模式的数目。显而易见,  $T$  是对称的和高度稀疏的。在任一时刻, 各神经元依下述 McCulloch-Pitts 规则选择其状态:

$$x_i(t+1) = \text{sgn}\left(\sum_{j=1}^N T_{i,j}x_j(t) - \theta_i\right), i = 1, 2, \dots, N, \quad (3)$$

式中  $\text{sgn}(x) = x/|x|$ ,  $\theta_i$  为神经元阈值, 不妨设为零。突触连接矩阵  $T$  的对称性决定该网络具有全局稳定性。可恢复的存储子模式必须是网络的平衡点<sup>[1]</sup>。

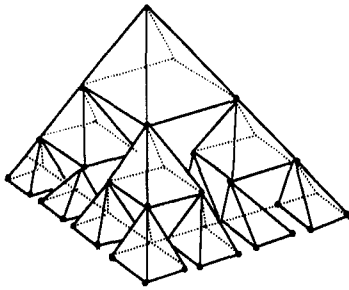


图 1  $N=85, p=5, K=3$  时的分维神经网络结构示意图

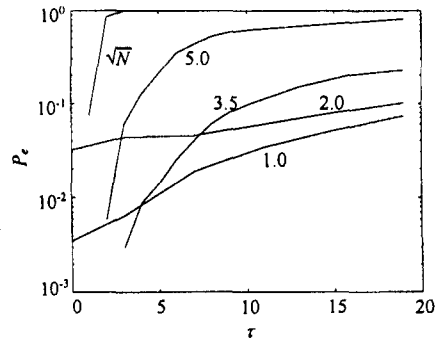


图 2  $\epsilon$  不同时记忆误差率随时间延迟变化的曲线

### 3 短期记忆学习算法

分维神经网络实现短期记忆的学习算法是在改进的临界学习算法<sup>[5]</sup>的基础上提出来的。定义  $T_{n,i}$  为一个  $p \times p$  维的矩阵, 它表示输入  $n$  个模式后第  $i$  个子网所对应的突触连接矩阵,  $X_{n,i}$  是一个  $p$  维列向量, 其元素值等于第  $n$  个输入模式对应于第  $i$  个子网的子模式值。子网突触连接矩阵  $T_{n,i}$  按下述规则进行修正:

$$T_{n,i} = \frac{1}{\sqrt{1 + \epsilon^2/N}} T_{n-1,i} + \frac{\epsilon/\sqrt{N}}{\sqrt{1 + \epsilon^2/N}} (I_p - rT_{n-1,i}) X_{n,i} X_{n,i}^T, \quad (4)$$

式中  $I_p$  为  $p \times p$  维单位矩阵,  $\epsilon$  和  $r$  为系统参数。当  $\epsilon$  较小时, 网络随时间忘却能力加强; 反之, 当  $\epsilon$  较大时, 网络趋向于记住较多的输入模式。显然, 随着记忆模式的增加, 新输入的模式总是对旧的模式产生排挤作用, 旧的模式渐渐变弱, 经过一段时间就被忘却, 因此想起的正确率降低。突触连接矩阵  $T$  的修正以串行的方式由低层向高层逐层进行。由于每个神经元最多与两个子网相连, 且子网可以做到很小, 这大大节省了计算时间。

根据 (3) 和 (4) 式, 可得到如下定理:

**定理**  $X_n$  是系统平衡状态的充要条件为: 对所有元素  $x_i$ ,  $i = 1, 2, \dots, N$ , 有

$$\left(\frac{\sqrt{N}}{\epsilon} - pr\right) \left| \sum_{j \neq k} T_{n-1, \nu_i}(k, j) X_{n,k}(j) \right| < p + \left(\frac{\sqrt{N}}{\epsilon} - pr\right) T_{n-1, \nu_i}(k, k), \quad j = 1, 2, \dots, p, \quad (5)$$

其中  $1 \leq \nu_i \leq L$ ,  $k$  表示第  $i$  个神经元所在子网  $\nu_i$  所对应的神经元的序号,  $1 \leq k \leq p$ 。

为满足 (5) 式,  $\epsilon$  必须有一个下限  $\epsilon_{\text{inf}}$ 。它可由下述不等式粗略估算<sup>[6]</sup>:

$$\left(\frac{\sqrt{N}}{\varepsilon} - pr\right)(p-2) \frac{\varepsilon}{\sqrt{N+\varepsilon^2} - \sqrt{N} + \varepsilon} < p. \quad (6)$$

#### 4 实验结果和讨论

以下对 Morris-Wong 模型<sup>[3]</sup>和本文提出的方法进行仿真比较。假设两种模型具有相同的神经元数目, 实验中需存储的模式采用相互独立且均匀分布的符号随机序列, 该序列以等时间间隔产生并顺序送入网络。以最近存入模式  $X_n$  的时刻为研究起点, 它对应的时间延迟  $\tau = 0$ , 则第  $(n-k)$  个模式所对应的时间延迟  $\tau = k$ 。记忆误差率  $P_e$  表示错误恢复模式数与存储模式总数的比值。

先讨论 (4) 式中  $\varepsilon$  的作用。采用图 1 所示的网络结构, 选择  $r = 1/\sqrt{N}$ 。图 2 表示当  $\varepsilon$  取不同值时, 记忆误差率随时间延迟变化的曲线。显然,  $\varepsilon$  较大时, 网络可以可靠存储最后的几个模式, 延迟越大, 记忆误差率也越大; 另一方面, 当  $\varepsilon$  较小时, 曲线变得平坦, 如图 2 中  $\varepsilon=1.0, 2.0$  时的情形。实验中发现  $\varepsilon=3.5$  是一个关键值, 当  $\varepsilon$  的取值小于 3.5 时, 没有模式能可靠记忆。它与由 (6) 式推出的理论估算值  $\varepsilon_{\text{inf}} = 3.64$  非常接近。以下的实验结果均假设  $\varepsilon=3.5$ 。

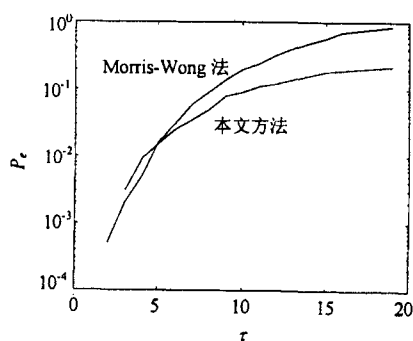


图 3 暂态记忆误差率曲线

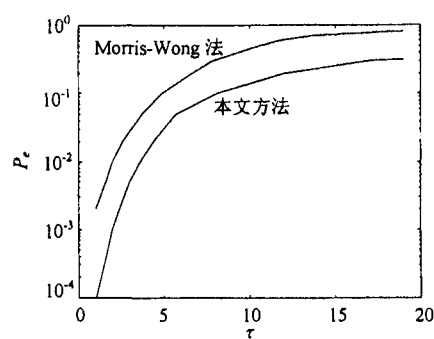


图 4 稳态记忆误差率曲线

图 3 表示网络离线使用时, 输入  $M$  个模式后暂态记忆误差率 (样本平均) 随时间延迟  $\tau$ ,  $\tau = 0, 1, \dots, M-1$  变化的曲线。可以看到, 记忆误差率  $P_e$  是时间延迟  $\tau$  的单调递增函数, 这与短期记忆的特性相吻合。Morris-Wong 模型和本文的方法比较, 多数情况下后者优于前者。由图中还可看出, 前者的正确记忆 ( $P_e = 0$ ) 模式数为 3, 而后者为 4, 这暗示后者更适合于在线使用。

图 4 表示稳态时 ( $n \rightarrow \infty$ ), 对应两种比较模型的记忆误差率随时间延迟变化的曲线。实验结果为 50,000 个输入模式在扣除了最初 200 个后的时间平均。由图中曲线可见, 两种方法对最新存入模式均能可靠记忆, 且记忆误差率随时间延迟单调递增。定义短期记忆模型的性能指标  $C_p$  表征允许记忆误差率范围内的模式数目, 设  $P_e \leq 5\%$  则对应于 Morris-Wong 模型,  $C_p = 4$ , 而本文的方法可得到  $C_p = 6$ 。显然本文的方法优于对比方法。

#### 5 结束语

本文提出了一种基于分维神经网络的短期记忆模型。该模型神经元间的连接高度稀疏, 突触连接矩阵可以快速修正, 节省大量的存储空间与运算时间。实验结果表明, 该模型在网络结构、运算速度以及性能等方面都优于对比方法, 适于在线使用。

## 参 考 文 献

- [1] Hopfield J J. Neural networks and physical systems with emergent computational abilities, Proc. Nat. Acad. Sci., 1982, 79: 2554-2558.
- [2] Ho C Y, Sasase I, Mori S. On the capacity of the Hopfield associative memory, Proc. of IEEE/INNS IJCNN'92, Baltimore: 1993, 196-201.
- [3] Morris Robert J T, Wong Wing Shing. A short-term neural network memory, SIAM J. Comput., 1988, 17(6): 1103-1118.
- [4] Baram Y. Associative memory in fractal neural networks, IEEE Trans. on SMC, 1989, SMC-19(5): 1133-1141.
- [5] Mezard M, Nadal J P, Toulouse G. Solvable models of working memories, J. Physique, 1996, 47(7): 1457-1462.
- [6] Li Xia, Wong Wing Shing. Fractal neural networks for short term memory, Proc. of 1994 ICNN of the IEEE WCCL, Orlando: 1995, 2454-2459.

## FRACTAL NEURAL NETWORKS FOR SHORT TERM MEMORY

Li Xia    Xie Weixin

*(Faculty of Information Engineering, Shenzhen University, Shenzhen 518060)*

**Abstract** Synaptic connection matrix encodes limited information. It is well known that neural network memory with storage prescriptions based on Hebb's rule will collapse as more patterns are stored. By requiring that old patterns are automatically forgotten and the memory recalls only the most recent ones, a new short-term neural network memory based on Y.Baram's fractal neural network is proposed. Comparison is made with Morris and Wong's method and the experimental results are shown to be rather satisfactory and encouraging.

**Key words** Fractal neural network, Short-term memory, Synaptic connection matrix

李霞: 女, 1968年生, 博士, 主要从事信号处理、神经网络、信息与控制等方面的研究.

谢维信: 男, 1941年生, 教授, 博士生导师, 主要从事智能信息处理、模糊信息处理、信号处理、模式识别和图象处理等方面的研究.