

联想记忆神经网络的一个有效学习算法

梁学斌 吴立德 俞俊

(复旦大学计算机科学系 上海 200433)

摘要 提出一种新的联想记忆网络模型的有效学习算法,它具有下述特点:(1)可以全部存储任意给定的训练模式集,即对于训练模式的数目和它们之间相关性的强弱没有限制;(2)最小的训练模式吸引域达到最大;(3)在(2)的基础上,每个训练模式具有尽可能大的吸引域;(4)联想记忆神经网络是全局稳定的.大量的计算机仿真实验结果充分说明所提出的学习算法比已有算法具有更强的存储能力和联想容错能力.

关键词 联想记忆网络,训练模式集,全部存储,最大吸引域,全局稳定性.

1 引言

联想记忆神经网络可分为有自反馈和无自反馈两种模型,它具有信息记忆和信息联想的功能,能够从部分信息或有适当畸变的信息联想出相应的存储在联想记忆神经网络中的完整的记忆信息^[1,2].其性能主要是由具体的学习算法来决定.

至今,已经提出了不少关于联想记忆神经网络的学习算法,主要有 Hebbian 学习算法^[3]、投影学习算法^[4]、Gardner 学习算法^[5]、最小重叠学习算法^[6]、Ho-Kashyap 学习算法^[7]、神经元或训练模式加权学习算法^[8]和优化学习算法^[9]等.其中只有优化学习算法严格考虑了如何提高联想记忆神经网络的存储能力和联想容错能力.

基于文献[9]的思想,本文提出了设计联想记忆网络的极大极小准则,它要求设计出的对称连接权阵应使得网络最小的记忆模式吸引域达到最大.并进一步发展了综合联想记忆神经网络的一个有效学习算法,它具有如下特点:(1)可以全部存储任意给定的训练模式集;(2)最小的训练模式吸引域达到最大;(3)在(2)的基础上,每个训练模式具有尽可能大的吸引域;(4)网络连接权阵是主对角元为 1 的对称阵,因此所设计出的联想记忆神经网络是全局稳定的^[11].

2 无自反馈网络模型训练式集的基本约束

联想记忆神经网络模型是由 N 个互联神经元组成的非线性动力系统.网络状态 $\mathbf{v} = (v_1, \dots, v_N)^t$, 其中 $v_i (i=1, 2, \dots, N)$ 表示第 i 个神经元的状态,取值空间是 $\{-1, 1\}$;网络连接权阵 $W = (W_{ij})_{N \times N}$, 其中 W_{ij} 表示从第 j 个神经元到第 i 个神经元的连接权.联想记忆神经网络模型可表述为^[1]

$$v_i' = \operatorname{sgn}\left(\sum_{j=1}^N W_{ij} v_j\right), \quad i = 1, 2, \dots, N. \quad (1)$$

其中 $v_i' = (v_1', \dots, v_N')^t$ 表示下一时刻的状态向量, 非线性函数 $\operatorname{sgn}(x)$ 定义为当 $x \geq 0$ 时为 1, 而当 $x < 0$ 时为 -1.

若网络连接权阵 W 满足 $W_{ij} = W_{ji}$ ($i, j = 1, 2, \dots, N$), 且 $W_{ii} = 0$ ($i = 1, 2, \dots, N$), 即连接权阵是一个具有零对角的实对称阵, 则称为 Hopfield 网络^[1].

设有 M 个不同的训练模式, 即 x^1, x^2, \dots, x^M , 其中 $x^u = (x_1^u, \dots, x_N^u)^t$, $u = 1, 2, \dots, M$. 它们成为系统(1)的稳定吸引子等价于对所有 $u = 1, 2, \dots, M$, 都成立

$$x_i^u = \operatorname{sgn}\left(\sum_{j=1}^N W_{ij} x_j^u\right), \quad i = 1, 2, \dots, N. \quad (2)$$

文献[11]已证明, Hopfield 网络的存储容量不超过 N , 即对于任何正整数 M , 若 $M > N$, 则一定存在 M 个训练模式, 它们不可能同时是 Hopfield 网络的稳定吸引子.

本节的结果是, 即使 $M \leq N$, 若 M 个不同训练模式同时成为 Hopfield 网络的稳定吸引子, 则必须满足如下基本约束(Fundamental Constraint, FC): 任意两个不同的训练模式至少有两个分量不同. 若不然, 则存在某两个不同训练模式刚好有且只有一个分量不同. 由于 Hopfield 网络无自反馈, 故这两个训练模式必然是同一个训练模式, 这就出现了矛盾.

由于有些训练模式间的分类仅靠某个模式分量, 如汉字“王”和“玉”, “已”和“己”等. 因此, 上述基本约束(FC)反映了 Hopfield 网络在存储能力方面的局限性.

以下不妨称连接权阵是零对角实对称阵的网络为无自反馈的 Hopfield 网络, 称连接权阵是具有非零对角元的实对称阵的网络为有自反馈的 Hopfield 网络. 有自反馈的 Hopfield 网络可以存储任意给定的训练模式集^[4], 问题的实质在于如何使每个训练模式具有尽量大的吸引域.

3 稳定吸引子和联想容错性分析

Hopfield 网络的连接权阵可由 Hebbian 学习准则确定, 即

$$W_{ij} = \begin{cases} \sum_{u=1}^M x_i^u x_j^u, & i \neq j, \\ 0, & i = j. \end{cases} \quad (3)$$

Hebbian 学习准则主要是为了模拟生物神经网络工作原理而提出的^[3], 并不是从联想记忆人工神经网络的工程设计准则出发. 它通常只能存储相关性较弱的训练模式集.

定理 1. M 个不同训练模式 x^1, x^2, \dots, x^M 成为系统(1)的稳定吸引子的充分条件是对所有 $u = 1, 2, \dots, M$ 都成立

$$F_i^u = \sum_{j=1}^N (x_i^u x_j^u) W_{ij} > 0, \quad i = 1, 2, \dots, N. \quad (4)$$

定理 2. 设 x^u ($u \in \{1, 2, \dots, M\}$) 是系统(1)的一个记忆模式, x 是一个畸变模式, $H(x, x^u)$ 表示 x 和 x^u 之间的 Hamming 距离. 若 $H(x, x^u) < F_i^u / (2W_{mi})$, $i \in \{1, 2, \dots, N\}$, 则 x 一步迭代联想出 x_i^u , 其中 $W_{mi} = \max_{1 \leq j \leq N} |W_{ij}|$.

证明. 即要证明 $x_i' = x_i''$. 该等式成立的一个充分条件是

$$\sum_{j=1}^N (x_i'' x_j) W_{ij} > 0. \quad (5)$$

注意到
$$\sum_{j=1}^N (x_i'' x_j) W_{ij} = \sum_{j=1}^N (x_i'' x_j'') W_{ij} + \sum_{j=1}^N x_i'' (x_j - x_j'') W_{ij}, \quad (6)$$

由于
$$\sum_{j=1}^N x_i'' (x_j - x_j'') W_{ij} \geq -2H(\mathbf{x}, \mathbf{x}'') W_{mi},$$

故由(6)式得
$$\sum_{j=1}^N (x_i'' x_j) W_{ij} \geq F_i'' - 2H(\mathbf{x}, \mathbf{x}'') W_{mi} > 0,$$

从而不等式(5)成立.

证毕.

为了书写和结论简洁,且不影响结论的实质,文后均假定对于所有 $i=1,2,\dots,N, W_{ij}$ ($j=1,2,\dots,N$)不全为0,且 $W_{mi}=1$. 这样, F_i'' 值越大,则 \mathbf{x}'' 的第 i 个分量的联想容错性就越强. 类似的定性讨论也可参考文献[6,12].

令
$$F^u = \min_{1 \leq i \leq N} F_i'', \quad u = 1, 2, \dots, M. \quad (7)$$

推论 1. 设 $\mathbf{x}'' (u \in \{1, 2, \dots, M\})$ 是系统(1)的一个记忆模式, \mathbf{x} 是一个畸变模式. 若 $H(\mathbf{x}, \mathbf{x}'') < F^u/2$, 则一步迭代联想起 \mathbf{x}'' .

由推论 1 可知, $F^u/2$ 是记忆模式 \mathbf{x}'' 的吸引域半径的一个下界.

4 极大极小设计准则和两个学习算法

设
$$F = \min_{1 \leq u \leq M} F^u, \quad (8)$$

则 $F/2$ 是最小的记忆模式吸引域的半径的一个下界.

从数学上,极大极小设计准则可表示为下列有约束不可微优化问题

$$W^* = \arg \max(F) = \arg \max_{1 \leq u \leq M} \min F^u,$$

即
$$W^* = \arg \max_{1 \leq u \leq M} \min_{1 \leq i \leq N} F_i'', \quad (9)$$

其中 $\arg \max$ 的约束条件是 $W_{ij} = W_{ji}$, 且 $|W_{ij}| \leq 1; i, j = 1, 2, \dots, N$.

将(9)式中的 $W_{ij} (i > j; i, j = 1, 2, \dots, N)$ 都变化成 W_{ji} , 则这个约束不可微优化问题含有 $N(N+1)/2$ 个自由变量, 即 W_{ij} , 而 $\arg \max$ 的约束条件是 $|W_{ij}| \leq 1$. 因此, 有约束不可微优化问题(9)等价于下列线性规划问题:

$$(P) \max(z), \quad (10)$$

并带有约束条件

$$(C_1) F_i'' \geq z, \quad i = 1, 2, \dots, N; \quad u = 1, 2, \dots, M; \quad (11)$$

$$(C_2) |W_{ij}| \leq 1, \quad i \leq j; \quad i, j = 1, 2, \dots, N. \quad (12)$$

线性规划问题(10)–(12)可以用单纯形方法来求解. 由于单纯形方法相对较复杂,特别是在 N 和 M 都比较大时. 为此, 本文将提出更有效的简单方法.

现对(9)式作分析

$$\max_{(C_2)} \min_{1 \leq u \leq M} \min_{1 \leq i \leq N} F_i'' = \max_{(C_2)} \min_{1 \leq i \leq N} \min_{1 \leq u \leq M} F_i'' = \min_{1 \leq i \leq N} \max_{(C_3)} \min_{1 \leq u \leq M} F_i'', \quad (13)$$

其中第二个等号利用了 F_i'' 只与 $W_{ij} (j=1, 2, \dots, N)$ 有关的事实, $\max_{(C_2)}$ 和 $\max_{(C_3)}$ 分别表示在约

束(C2)和(C3)下求最大值,(C3)定义为

$$(C3) \quad |W_{ij}| \leq 1, \quad j = 1, 2, \dots, N. \quad (14)$$

由于 $\min(a+b) \geq \min(a) + \min(b)$ (其中 a 和 b 是变量), 故由(13)式得

$$\max_{(C2)} \min_{1 \leq u \leq M} \min_{1 \leq i \leq N} F_i^u \geq \min_{1 \leq i \leq N} \max_{(C3)} \sum_{j=1}^N \min_{1 \leq u \leq M} (x_i^u x_j^u) W_{ij}. \quad (15)$$

易知约束优化问题

$$\max_{(C3)} \sum_{j=1}^N \min_{1 \leq u \leq M} (x_i^u x_j^u) W_{ij} \quad (16)$$

的解是

$$W_{ij} = \begin{cases} 1, & \text{若 } x_i^u x_j^u = 1, \quad u = 1, 2, \dots, M, \\ -1, & \text{若 } x_i^u x_j^u = -1, \quad u = 1, 2, \dots, M, \\ 0, & \text{其它情形.} \end{cases} \quad (17)$$

由不等式(15)可知,(17)式是不可微优化问题(9)式的一个次优解.

Hopfield 联想记忆网络的连接权可以由(17)式来确定,该规则就称为快速学习算法.它具有下述特点:

- 1) 可以存储任意给定的训练模式集,且具有一定的联想容错能力^[13];
- 2) 可以用作其它约束优化迭代算法的初值;
- 3) 设计出的连接权值为 $-1, 0$ 或 1 ,特别是,主对角权元素全为 1 ,故网络易于硬件实现和光学实现.

设 f 是线性规划问题(10)—(12)的解.线性规划问题(10)—(12)等价于在约束条件

$$(C4) \quad |W_{ij}| \leq 1 \quad (i < j; \quad i, j = 1, 2, \dots, N) \quad (18)$$

$$\text{下求解线性不等式组 } F_i^u \geq f \quad (i=1, 2, \dots, N; \quad u=1, 2, \dots, M); \quad (19)$$

而在无约束条件下求解线性不等式组(19)可用感知器算法^[14].

设训练模式的学习序列是 $x^1, x^2, \dots, x^M, \dots, x^1, x^2, \dots, x^M, \dots$, 即是一个循环学习序列.感知器算法按此循环顺序接受训练模式 $x^u (u \in \{1, 2, \dots, M\})$, 且按下列规则来修改连接权: 对于 $i=1, 2, \dots, N$, 有

$$W_{ij}(t+1) = \begin{cases} W_{ij}(t), & \text{若 } F_i^u \geq f, \\ W_{ij}(t) + qx_i^u x_j^u, & \text{若 } F_i^u < f, \text{ 且 } i < j. \end{cases} \quad (20)$$

其中 t 是算法迭代次数, $q > 0$ 是任意正数, W_{ij} 的迭代初值可以任意选定.

为使得每个训练模式具有一定大小的吸引域,在感知器算法(20)中须加入约束条件(C4).同时,将快速学习算法(17)的结果作为连接权阵的迭代初值,并且只对初值为 0 的非对角权元素进行优化迭代.在此基础上,再使得每个训练模式具有尽量大的吸引域.

本文的约束感知器优化学习算法可总结如下:

(1) 初始化. $W_{ij}(0) (i < j; \quad i, j = 1, 2, \dots, N)$ 由快速学习算法(17)确定, $f(0) = 1$, 并选取学习因子 q , 迭代次数 δ_1 和 δ_2 ;

(2) 学习. 设从循环学习序列 $x^1, x^2, \dots, x^M, \dots, x^1, x^2, \dots, x^M, \dots$, 接受到训练模式 $x^u (u \in \{1, 2, \dots, M\})$, 则对于 $i=1, 2, \dots, N$, 按下式来修改 $W_{ij}(t) (i < j; \quad j=1, 2, \dots, N)$

$$W_{ij}(t+1) = \begin{cases} W_{ij}(t), & \text{若 } F_i^u \geq f, \\ \Phi(W_{ij}(t) + qx_i^u x_j^u), & \text{若 } F_i^u < f, \text{ 且 } W_{ij}(0) = 0. \end{cases} \quad (21)$$

其中函数 $\Phi(x)$ 定义为当 $x > 1$ 时为 1, 当 $x < -1$ 时为 -1, 否则为 x . 每完成一个循环序列 $\{x^1, x^2, \dots, x^M\}$, 就称为迭代一次;

(3) 判断. 若算法(21)收敛, 即 $F_i^u \geq f$ ($i=1, 2, \dots, N; u=1, 2, \dots, M$) 时的迭代次数小于 δ_1 , 则 $f = f + 1$, 转至步骤(2); 若算法(21)迭代次数超过 δ_1 后仍不收敛, 则 W_{ij} 返回算法(21)式迭代之前的值, 并转至步骤 4;

(4) 增大每个训练模式的吸引域. 设 $f^u = f, u=1, 2, \dots, M$. 设从循环学习序列 $x^1, x^2, \dots, x^M, \dots, x^1, x^2, \dots, x^M, \dots$, 接受到训练模式 $x^u, u \in \{1, 2, \dots, M\}$, 则对于 $i=1, 2, \dots, N$, 按下式来修改 $W_{ij}(t)$ ($i < j; j=1, 2, \dots, N$)

$$W_{ij}(t+1) = \begin{cases} W_{ij}(t), & \text{若 } F_i^u \geq f^u, \\ \Phi(W_{ij}(t) + qx_i^u x_j^u), & \text{若 } F_i^u < f^u, \text{ 且 } W_{ij}(0) = 0. \end{cases} \quad (22)$$

每完成一个循环序列 $\{x^1, x^2, \dots, x^M\}$, 就称为迭代一次.

每迭代一次, 且迭代次数小于 δ_2 时, 对于 $u=1, 2, \dots, M$, 按如下规则递增 f^u : 若 $F_i^u \geq f^u$ ($i=1, 2, \dots, N$), 且 $F_i^{u'} \geq f^{u'} - 1$ ($u' \neq u, u'=1, 2, \dots, M; i=1, 2, \dots, N$), 则 $f^u = f^u + 1$, 并且迭代次数重新从 0 开始, 返回步骤(4);

若算法(22)式在迭代次数超过 δ_2 后, 所有 f^u ($u=1, 2, \dots, M$) 都不能再作递增, 则 W_{ij} 返回算法(22)式迭代之前的值, 并结束算法.

5 计算机实验结果

实验中取 $q=1/N, \delta_1=\delta_2=40$. 设

$$F_{\max} = \max_{1 \leq i \leq N} \max_{1 \leq u \leq M} F_i^u, F_{\min} = \min_{1 \leq i \leq N} \min_{1 \leq u \leq M} F_i^u, F_{\text{ave}} = \left(\sum_{i=1}^N \sum_{u=1}^M F_i^u \right) / (MN), \quad (23)$$

F_{\max}, F_{\min} 和 F_{ave} 可用作统计量来定量客观地评价 Hopfield 联想记忆学习算法的联想容错能力. F_{\max}, F_{\min} 和 F_{ave} 越大, 则说明算法的联想容错能力越强.

第一组实验的训练模式集是汉字“己”、“巳”和“巳”, 如图 1 所示. 用快速学习算法可以记忆该训练模式集. Hebbian 学习规则和文献[9]的优化学习算法都不能存储这个训练模式集. 这是因为, “己”和“巳”以及“己”和“巳”的 Hamming 距离均为 1. 实验结果表明 Hebbian 学习规则的联想结果是 3 个训练模式都联想到模式“己”. 表 1 是快速学习算法和本文优化学习算法的联想容错能力比较. 由表 1 可知, 快速学习算法和本文优化学习算法在此特例下, 具有相同的联想容错能力.

表 1 两种算法的联想容错能力比较

	F_{\max}	F_{\min}	F_{ave}
快速学习算法	98.00	1.00	96.06
本文优化学习算法	98.00	1.00	96.06

第二组实验的训练模式集是英文字母“A”—“H”, 如图 2 所示, 对于 8 组训练模式集 $\{A\}, \{A, B\}, \{A, B, C\}, \{A, B, C, D\}, \{A, B, C, D, E\}, \{A, B, C, D, E, F\}, \{A, B, C, D, E, F, G\}$ 和 $\{A, B, C, D, E, F, G, H\}$ 分别用本文优化学习算法和文献[9]的优化学习算法进

行学习. 表 2 是两种优化学习算法的容错能力比较, 其中 M 表示上述 8 组训练模式集分别所含的模式个数. 由表 1 和 2 的结果可知, 本文优化学习算法比文献[9]的优化学习算法具有更强的联想容错能力, 同时说明快速学习算法的结果作为迭代算法的初值可大大提高迭代算法的联想容错性.

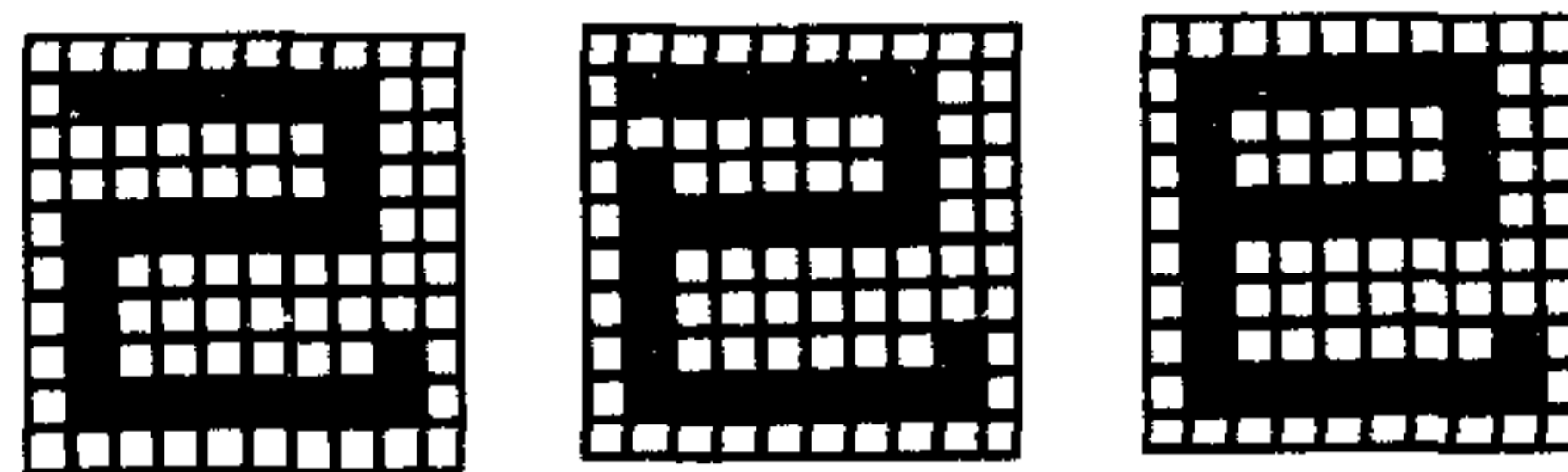


图 1 实验 1 训练模式集

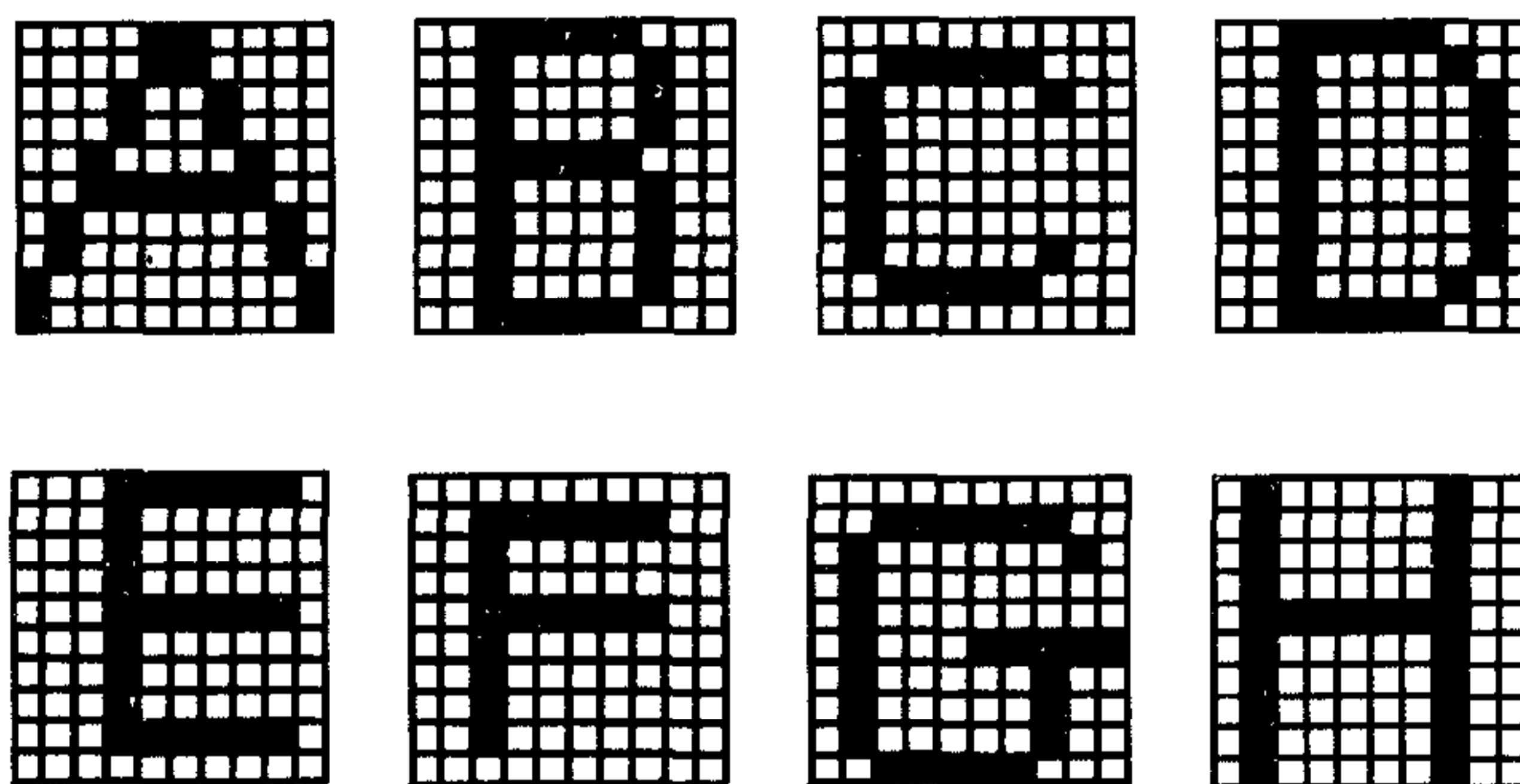


图 2 实验 2 训练模式集

表 2 两种优化学习算法的联想容错能力比较

M	1	2	3	4	5	6	7	8
F_{\max} (本文)	100	57.00	49.78	40.47	36.06	36.28	33.62	30.38
F_{\max} (文献[9])	99	42.56	34.63	14.70	14.83	14.83	14.41	17.88
F_{\max} (快速算法)	100	57.00	42.00	38.00	33.00	33.00	31.00	27.00
F_{\min} (本文)	100	43.00	33.34	14.08	14.19	13.83	11.08	12.02
F_{\min} (文献[9])	99	42.00	32.03	13.04	13.08	11.00	10.01	10.01
F_{\min} (快速算法)	100	43.00	15.00	1.00	1.00	1.00	1.00	1.00
F_{ave} (本文)	100	50.98	40.51	25.78	22.22	21.84	19.63	17.66
F_{ave} (文献[9])	99	42.32	32.77	13.75	13.81	11.86	10.94	11.00
F_{ave} (快速算法)	100	50.98	29.26	22.58	15.38	14.12	12.02	9.56

参 考 文 献

- [1] Hopfield J J. Neural networks and physical systems with emergent collective computational abilities, In: Proc. Natl. Acad. Sci., U. S. A., 1982, 79(4): 2554—2558.
- [2] Michel A N, Farrell J A. Associative memories via artificial neural networks. *IEEE Control Syst. Mag.*, 1990, 10(4): 6—17.
- [3] Hebbian D O. The organization of behavior. New York: Wiley, 1949.
- [4] Personnaz L, Guyon I, Dreyfus G. Collective computational properties of neural networks; new learning mecha-

- nisms. *Phys. Rev. A*, 1986, **34**(5): 4217—4228.
- [5] Gardner E. The space of interactions in neural network models. *J. Phys. A: Math. Gen.*, 1988, **21**: 257—270.
- [6] Krauth W, Mezard M. Learning Algorithms with Optimal Stability in Neural Networks. *J. Phys. A: Math. Gen.*, 1987, **20**: 745—752.
- [7] Hassoun M H, Youssef A M. High performance recording algorithm for hopfeild model associative memories. *Optical Eng.*, 1989, **28**(1): 46—54.
- [8] Wang T, Zhuang X H, Xing X L. A neuron-weighted algorithm and its hardware implementation in associative memories. *IEEE Trans. Computer*, 1993, **42**(5): 636—640.
- [9] Wang T. Learning of Hopfield associative memory by global minimization. *Int. J. Pattern Recog. Artif. Intell.*, 1993, **7**(3): 559—567.
- [10] Atiya A, Abu-Mostafa Y S. An analog feedback associative memory. *IEEE Trans. Neural Networks*, 1993, (4): 117—126.
- [11] Goles E, Fogelman F, Pellegrin D. Decreasing energy functions as a tool for studying threshold networks. *Discrete Appl. Math.*, 1985, **12**: 261—277.
- [12] Abbott L F, Kepler T B. Optimal learning in neural network memories. *J. Phys. A: Math. Gen.*, 1989, **22**: 711—717.
- [13] 梁学斌, 吴立德. 综合联想记忆神经网络的外积取等准则. *通信学报*, 1995, **16**(2): 1—6.
- [14] Tou J T, Gonzalez R C. *Pattern recognition principles*. Massachusetts: Addison-Wesley, 1977. 165—168.

AN EFFICIENT LEARNING ALGORITHM FOR ASSOCIATIVE MEMORY NEURAL NETWORK

LIANG XUEBIN WU LIDE YU JUN

(Dept. of Computer Science, Fudan University, Shanghai 200433)

Abstract A new and efficient learning algorithm of asociative memory neural network is proposed, with the following characteristics: (1) it can store any given training pattern set no matter how much and what correlation among them may be; (2) the smallest domain of attraction of training patterns is maximized; (3) each domain of attraction of training patterns is guaranteed to be as large as possible; (4) the designed associative memory network is globally stable. A large number of computer experimental results confirm that our algorithm possesses more powerful storage ability and more fault-tolerance capability than existing ones.

Key words Associative memory, training pattern set, total storage ability, maximized domain of attraction, global stability.

梁学斌 1997年1月获复旦大学计算机软件专业博士学位. 现从事计算机智能领域的研究工作, 在国内外已发表论文30余篇.

吴立德 复旦大学计算机科学系教授, 博士生导师. 从事计算机视觉和计算机语言学方面的研究, 已发表论文100余篇, 著作多部.

俞俊 复旦大学计算机科学系研究生, 主要学习计算机视觉和图象编码.