

用于数字式细胞神经网络设计的学习算法¹

乔长阁

(清华大学计算机系 北京 100084)

摘 要 本文针对数字式细胞神经网络 (DCNN) 的应用, 给出了 DCNN 模板的设计方法, 即不等式构造法, 并提出一个基于松弛法的 DCNN 有教师学习算法, 为 DCNN 的设计提供了理论根据。在连通片检测等应用中的模拟表明了算法的有效性和正确性。

关键词 数字式细胞神经网络, 邻接, 模板, 松弛算法

中图分类号 TN-052

1 引言

细胞神经网络^[1]的应用主要取决于其模板的设计。以往的细胞神经网络应用中, 一般采用试探法或者根据图象处理中的各种算子来设计模板系数^[2]。对我们所提出的数字式细胞神经网络^[3]DCNN(Digital Cellular Neural Network)而言, 模板的设计也同样决定了网络的应用范围。为了在图象处理和模式识别等领域中发挥 DCNN 的作用, 本文探讨了 DCNN 模板的设计方法, 即不等式构造法。同时, 基于解线性不等式的松弛法, 提出一个基于松弛法的学习算法。

2 数字式细胞神经网络

DCNN 是这样—个数字式系统, 其基本单元是一个细胞, 而每个细胞仅仅与它的近郊细胞相连接, 只有近郊间的细胞才能直接地相互作用。从结构上讲, 它与 CNN^[1]相同。一个典型的二维 DCNN 如图 1 所示。记第 i 行第 j 列的细胞为细胞 $C(i, j)$ 。下面是 $C(i, j)$ 邻接细胞的定义:

定义 1 r -邻接 在一个 DCNN 中, 细胞 $C(i, j)$ 的 r -邻接定义如下:

$$N_r(i, j) = \{C(m, n) | \max\{|m - i|, |n - j|\} \leq r, 1 \leq m \leq M, 1 \leq n \leq N\}, \quad (1)$$

这里, r 是一个正整数。

定义 2 数字式细胞神经网络 DCNN 可以用一个无向加权图来描述。图的每条边都附有一个权值, 每个节点 [神经元] 有一个阈值。设 DCNN 是一个 $M \times N$ 阶网络, 则 DCNN 由 $[A, \Theta]$ 唯一地确定。其中, (1) A 是一个权值 (或称为模板系数) 矩阵, 也称为模板, 它有 $(2r+1) \times (2r+1)$ 个元素, A_{ijmn} 为附给 $C(i, j)$ 和 $C(m, n)$ ($C(m, n) \in N_r(i, j)$) 两个细胞间的权值。(2) Θ 是一个 $M \times N$ 维向量, Θ_{ij} 表示细胞 $C(i, j)$ 的阈值。

定义 3 DCNN 状态变化 对上面所定义的 DCNN, 每个神经元可以处于两个可能状态之一, 即 1 或 0。若我们用 $Y_{ij}(t)$ 表示 t 时刻细胞 $C(i, j)$ 的输出, $X_{ij}(t)$ 表示神经元 $C(i, j)$ 在

¹ 1995-08-30 收到, 1996-06-12 定稿

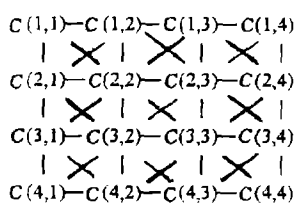


图 1 一个二维 DCNN 实例

t 时刻的状态, 则 $C(i, j)$ 的次态由以下规则来描述:

$$Y_{ij}(t) = f(X_{ij}(t)) = \begin{cases} 1, & \text{当 } X_{ij}(t) < 0; \\ Y_{ij}(t), & \text{当 } X_{ij}(t) = 0; \\ 0, & \text{当 } X_{ij}(t) > 0; \end{cases} \quad (2)$$

$$X_{ij}(t) = \sum_{C(m,n) \in N_r(i,j)} [A_{ijmn} Y_{mn}(t)] - \theta_{ij}. \quad (3)$$

整个网络的输出向量 $\mathbf{Y}(t) \in \{0, 1\}^{M \times N}$. DCNN 网络的权值由 r -邻接来定义, 只有近邻间才有权值, 较远邻接权值为零.

3 DCNN 的不等式构造法和应用

由于 DCNN 采用了离散形式, 因此可以很容易地用线性不等式来描述状态变化关系. 我们就利用不等式构造法来给出次态与现态、输入和阈值间的关系, 通过解这个线性不等式集来得到所要求的模板. 下面结合在连通片检测中的应用实例给出不等式构造过程.

连通片检测的目的是检测一条含 N 个像素的线段 $\{0, 1\}^N$ 中连通片的数目^[4]. 如图 2(a) 所示为一条线段, 假设 $\blacksquare = 1$, $\square = 0$, 检测结果应如图 2(b) 所示, 结果态中黑块像素的数目等于线段初始态中连通片的数目. 这是一个一维问题, 重写 DCNN 状态变化如下:

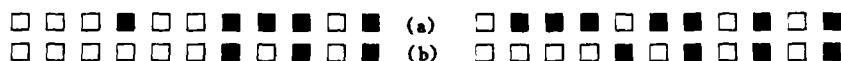


图 2 连通片检测实例

$$Y_i(t) = f(X_j(t)) = \begin{cases} 1, & \text{当 } X_j(t) < 0; \\ Y_i(t), & \text{当 } X_i(t) = 0; \\ 0, & \text{当 } X_j(t) > 0; \end{cases} \quad (4)$$

$$X_j(t) = \sum_{C(j) \in N_r(i)} [A_{ij} Y_j(t)] - \theta_j. \quad (5)$$

如果考虑 $r=1$ 邻域, 则对应这一邻域有 $2^{2r+1} = 8$ 种状态和输出结果. 若假设 $A = (a \ b \ c)$, $-\theta_i = \theta$ (即所有神经元的阈值相同), 则 (4) 式将变为: $Y_i(t+1) = f[X_i(t)] = f[aY_{i-1}(t) + bY_i(t) + cY_{i+1}(t) + \theta]$, 从而根据状态的变化可对应得到构造的不等式. 表 1 给出对应的 8 种转变态. 例如, 对状态 4, 若连通片的当前状态为 $\{0, 1, 1\}$, 则中间神经元的次态应变为 0 (因为后面的 1 可继续向后传播). 与表 1 相对应, 根据 (4), (5) 式列不等式如表中最下一行所示. 对这 8 个不等式, 可以给出 1 组解为: $a=-2$, $b=-2$, $c=2$, $\theta=1$, 即 $\theta_i = -1$.

这种模板设计的不等式构造方法同样可推广到其他应用中. 实际应用时, 首先针对一维或二维问题设计一种模板形式, 给出受影响的邻接细胞, 然后根据周边和当前点状态的预测变化来构造不等式集, 通过解这个不等式集即可求出最后的模板. 为了解这些不等式集, 需要给出求解的方法. 下面我们将介绍解线性不等式的松弛法, 并提出我们的 DCNN 模板学习算法.

4 求解线性不等式的松弛法

表 1 连通片检测的 8 种状态

状态	1	2	3	4	5	6	7	8
当前	Y_{-1}	0	0	0	1	1	1	1
前	Y_0	0	0	1	1	0	1	1
态	Y_1	0	1	0	1	0	0	1
次态	Y_0	0	0	1	0	1	1	1
不等式	$\theta \geq 0$	$c + \theta \geq 0$	$b + \theta \leq 0$	$b + c + \theta > 0$	$a + \theta < 0$	$a + c + \theta \geq 0$	$a + b + \theta \leq 0$	$a + b + c + \theta \leq 0$

松弛法是迭代解一组线性不等式集的算法^[5]。考虑 m 个线性不等式集

$$\sum_{j=1}^n a_{ij}x_j + b_i \geq 0, \quad (i = 1, 2, \dots, m) \quad (6)$$

这里 $a_{ij} \in R$, $b_i \in R$ 。问题是设计一个数值过程, 可以找到满足不等式 (6) 式的解 $(x_1, x_2, \dots, x_n)$ 。

定义 4 逐点接近 假设 p 与 p_1 为欧几里得空间的点, 若 $|p - a| > |p_1 - a|$ 对每一个 $a \in A$ 成立, 则称 p_1 比 p 更逐点接近于集合 A , 如图 3 所示。

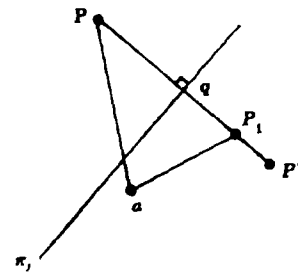


图 3 逐点接近点构造

假设每个线性不等式定义的闭环半空间为 H_i , 而解空间为这些半空间的交集, 即 $A = H_1 \cap H_2 \cap \dots \cap H_m$ 。若给定 $p \notin A$, 则可以用下面的方法构造一个比 p 更逐点接近于 A 的点 p_1 : 设 q 为 p 在超平面 π_j 上的投影, 若 $p \in A$, 即其坐标满足 (6) 式, 则寻找终止; 若 $p \notin A$, 定义 $p_1 = p + \lambda(q - p)$, 其中 λ 是一个常数, $0 < \lambda \leq 2$ 。为方便起见, 简写为 $p_1 = F_\lambda(p)$ 。若 $p_1 \in A$, 则过程终止, 若 $p_1 \notin A$, 则继续迭代得到 $p_2 = F_\lambda(p_1), \dots$, 最后得到一个序列 $p = p_0, p_1, p_2, \dots$, 且有如下连接关系: $p_{i+1} = F_\lambda(p_i)$ 。Motzkin^[5] 证明了, 序列 $\{p_i\}$ 总是可以终止的, 即算法收敛。

5 基于松弛法的 DCNN 学习算法

为了得到 DCNN 学习算法, 在松弛法中取 $\lambda = 2$, 即 $p_1 = p + 2(q - p) = 2q - p$, 则 p_1 与 p 对称。对给定点 $p \notin H_i$, 可以求得 p_1 点坐标 (即违约点 $\sum_{j=1}^n a_{ij}x_j^k + b_i < 0$ 对于超平面 $\sum_{j=1}^n a_{ij}x_j^k + b_i = 0$ 的映象) 为 $X^{k+1} = X^k - 2(A_i^T X^k + b_i) \cdot A_i / (A_i^T \cdot A_i)$, 其中,

$$A_i = \begin{bmatrix} a_{i1} \\ a_{i2} \\ \vdots \\ a_{in} \end{bmatrix}, \quad A_i^T = [a_{i1} \ a_{i2} \ \dots \ a_{in}], \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

我们给出 DCNN 基于松弛法的学习算法如下:

第 1 步 根据学习实例从 (2),(3) 式构造不等式集形如: $\sum_{j=1}^n a_{ij}x_j + b_i \leq 0, (i = 1, 2, \dots, m)$ 。

第 2 步 任意选择初始值 X^0 , $X^0 \in R$ 。

第3步 对所有违约不等式, 即 $\sum_{j=1}^n a_{ij}x_j + b_i > 0$, 取 $X^{k+1} = X^k - 2(A_i^T X^k + b_i) \cdot A_i / (A_i^T \cdot A_i)$ 。

第4步 若 X^{k+1} 满足所有不等式, 则算法终止; 否则转至第3步循环执行, 直至找到这样的 $(x_1, x_2, \dots, x_n)$ 为止。

实际上, 为了得到 DCNN 中这些不等式表达式, 必须先给出输入和输出样本才可进行这种神经网络的训练。在用样本训练完网络后, 用松弛法得到模板权值和细胞阈值, 就可将它们用于以后的应用中而不必再训练和学习。同时, 为了得到最后变量集的可靠性, 用于学习过程的输入和初始样本应有图象的状态变化, 而且不同约束数应较大。对一个绝对可靠的结果, 必须训练初始和输入的所有组合。

6 结论

本文基于 DCNN 的特点, 给出 DCNN 模板的不等式构造法, 并在连通片检测问题中给出具体的应用。在此基础上, 根据解线性不等式系统的松弛法, 提出了一个基于松弛法的有教师学习算法。模拟表明, 不等式构造算法和基于松弛法的学习算法对 DCNN 设计非常有效。我们已将它们应用于连通片检测、方向投影、边缘检测和迷宫通路求解等问题中, 均得到了正确的结果^[3]。同样, 算法也可用于 Hopfield 离散网络的学习中。

参 考 文 献

- [1] Chua L O, Yang L. Cellular neural networks : Theory and applications. IEEE Trans. on CAS, 1988, 35(10): 1257-1290.
- [2] Chua L O, Thiran P. An Analytic method for designing simple CNNs. IEEE Trans. on CAS, 1991, 38(11): 1332-1341.
- [3] 乔长阔. 数字式细胞神经网络理论与应用研究: [博士论文]. 西安: 西北工业大学, 1993.
- [4] Matsumoto T, et al. CNN cloning template : Connected component detector. IEEE Trans. on CAS, 1990, 37(5): 633-635.
- [5] Motzkin T S, Schoenberg L J. The Relaxation method for linear inequalities. Canada Journal of Mathematics, 1954, 6(3): 393-404.

LEARNING ALGORITHM FOR THE DESIGN OF DIGITAL CELLULAR NEURAL NETWORK

Qiao Changge

(Department of Computer Science and Technology, Tsinghua University, Beijing 100084)

Abstract The application of cellular neural networks is determined by its templates. In accordance with the application of digital cellular neural network the author proposed before, a method for the design of its templates, i.e. inequalities construction method, is given and a supervised learning algorithm is proposed based on the relaxation method. The learning algorithm provides theoretical basis for the design of DCNN. Simulation results on examples such as connected component detection shows the effectiveness and feasibility of the algorithm.

Key words Digital cellular neural network, Neighborhood, Template, Relaxation algorithm

乔长阔: 男, 1967年生, 博士, 助理研究员, 主要的研究领域包括: 超大规模集成电路计算机辅助设计和 EDA 系统, 图象处理, 细胞神经网络, 并行算法等。