异步四元数神经网络模型的能量函数研究 *

陈振湘 帅建伟 刘瑞堂 吴伯僖

(厦门大学物理系，厦门， 361005)

摘 要

本文证明了异步离散四元数神经网络模型在其动力学演化过程中，网络的能量函数单调递减，网络最终将稳定在一个平衡态上，存贮图像在能量函数空间中对应一能量极小点，为网络的稳定吸引子。

关键词： 神经网络，四元数，图像识别

1 引 言

近年来，神经网络的研究在世界范围内受到高度重视，它不仅对理解生物神经信息处理的一些基本方面有帮助，而且还给新一代智能计算机的研究带来巨大影响，并将推动整个人工智能领域的发展。该领域中一个重要课题即是多值图像自动识别，对于多值图像的识别，有两大思路的工作：一类是直接利用二值 Hopfield 神经网络模型 [1]，用多个二值神经来表示一个多态像点。Q 值灰度或彩色图像表示每一像点取值具有 Q 种可能，从物理上可理解为外部空间的每一个像点还存在着一个内部空间，该内部空间具有 Q 种状态。因此另一类工作是用一个多值神经元表示一个 Q 值像点的状态，从而提出一些多值神经网络模型 [2–5] 。我们把四元数 [6] 引入神经网络中，建立了 16 态四元数神经网络模型 [5]，文 [7] 引入量子力学的 Dirac 算符符号表示该模型，自然地在连接矩阵中引入了四元数共轭空间，并研究了它的存贮容量。该模型具有以下的特点：为双极 16 态模型，可用现在成熟的集成电路兑现；可应用于 16 值灰度或彩色图像识别中，其图像编码方式与计算机 16 值图像编码一一对应；其存贮容量与 Hopfield 神经网络模型相同。本文对该网络定义了一个能量函数，通过对能量函数的考察，进一步分析了异步离散四元数神经网络模型的存贮图像性质和动力学演化规律。

2 四元数神经网络模型简介

四元数也称为 Hamilton 数，具有一个实部和三个虚部。令四元数系

$$Q(R) = \{\alpha : \alpha = a + b\vec{i} + c\vec{j} + d\vec{k}, a, b, c, d \epsilon R\}$$

其中 $\vec{i}, \vec{j}, \vec{k}$ 表示虚部的三个基矢元，各基矢元间的乘法关系为：

$$\vec{i}^2 = \vec{j}^2 = \vec{k}^2 = -1, \vec{i}\vec{j} = -\vec{j}\vec{i} = \vec{k}$$

$$\vec{j}\vec{k} = -\vec{k}\vec{j} = \vec{i}, \vec{k}\vec{i} = -\vec{i}\vec{k} = \vec{j} \tag{1}$$

* 国家自然科学基金重点项目资助课题。

收稿日期： 1995-01-17

由上可知，四元数虽不满足乘法交换律，但满足结合律，即有 $\alpha\beta \neq \beta\alpha$，但 $(\alpha\beta)\gamma = \alpha(\beta\gamma)$。因此，对于四元数间的乘法运算，其前后排列顺序是很重要的，不可随意调动，但其运算秩序则可改变。对任一四元数 $\alpha = a+b\vec{i}+c\vec{j}+d\vec{k}$，定义其四元共轭 $\alpha^* = a-b\vec{i}-c\vec{j}-d\vec{k}$，故有 $\alpha\alpha^* = a^2+b^2+c^2+d^2 = |\alpha|^2$，其中 $|\alpha|$ 表示为 α 的模，可证有关系式

$$(\alpha\beta)^* = \beta^*\alpha^* \tag{2}$$

该式在下面的运算中会常出现。

我们把四元数引入 Hopfield 模型中，建立了四元数神经网络模型，四元数神经元具有 16 种状态 $(\pm 1 \pm \vec{i} \pm \vec{j} \pm \vec{k})$。设网络存贮模式即基矢为 $|\mu\rangle(\mu = 1, \cdots, M)$，连接矩阵表示为总投影算符：

$$J = \sum_{\mu=1}^{M} |\mu\rangle\langle\mu| \tag{3}$$

$\langle\mu|$ 表示为 $|\mu\rangle$ 的共轭基矢。连接矩阵 J 具有性质：

$$J = J^+$$

J^+ 表示 J 的共轭转置矩阵。故对任一态矢 $|S\rangle$，其有效局域场矢量为

$$|H\rangle = J|S\rangle \tag{4}$$

于是动力学方程为：

$$|t+1\rangle = \Theta|H(t)\rangle = \Theta J|t\rangle \tag{5}$$

其中 Θ 为推广的四元数阶跃函数。$\Theta(x)$ 取值规则如下：当 x 的实部以及某一虚部不小于 0 时，$\Theta(x)$ 的对应部分取为正单位元；当 x 的实部以及某一虚部小于 0 时，$\Theta(x)$ 的对应部分取为负单位元。

对于四元数神经网络模型，我们可以定义对态矢 $|S\rangle$ 的能量函数为：

$$E(S) = \langle S|J|S\rangle = -\frac{1}{2}\sum_{i,j=1}^{N} S_i^* J_{ij} S_j \tag{6}$$

我们可证明对于任意具有 $J = J^+$ 性质的连接矩阵网络，能量函数定义式为一实数：能量函数的共轭函数为

$$\begin{aligned} E(S)^* &= -\tfrac{1}{2}, \sum_{i,j=1}^{N} (S_i^* J_{ij} S_j)^* \\ &= -\tfrac{1}{2}, \sum_{i,j=1}^{N} S_j^* J_{ji} S_i \\ &= -\tfrac{1}{2}, \sum_{i,j=1}^{N} S_j^* J_{ji} S_i = E(S) \end{aligned}$$

也即能量函数定义式为一实数，它说明该定义是合理的。

3 存贮图像的能量函数研究

对随机图像存贮的 Hopfield 模型，当网络的神经元个数 N 远大于存贮图像数 M 时，在能量函数空间，存贮图像为 Hopfield 网络的能量极小值。对于四元数神经网络模型，同样可证明存贮图像在能量函数空间中对应一能量极小点。

对任一存贮图像 $|\nu\rangle$, 其能量为：

$$\begin{aligned} E(S^\nu) &= -\frac{1}{2}\langle\nu|J|\nu\rangle \\ &= -\frac{1}{2}\sum_{\mu=1}^{N}\langle\nu|\mu\rangle\langle\mu|\nu\rangle \\ &= -2N^2 - \frac{1}{2}\sum_{\mu\neq\nu}\langle\nu|\mu\rangle\langle\mu|\nu\rangle \\ &\approx -2N^2 + A/2 \end{aligned} \tag{7}$$

当 $N \gg M$ 时，第二项很小，可看作一微扰项。式 $\langle\nu|\mu\rangle\langle\mu|\nu\rangle$ 展开时，为 4^4N 项和，其中有 3/4 项为虚部而相互抵消，仅剩 4^3N 项实数 $+1,-1$ 的和。由于存贮图像的随机性，可得 A 为 $4^3N(M-1)$ 项随机实数 $+1,-1$ 的和，故其统计平均值为 0, 均方差为 $\sigma = 2^3\sqrt{N(M-1)}$, 且具有正态分布：

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}}\exp(-\frac{x^2}{2\sigma^2}) \tag{8}$$

设一输入图像 $|S\rangle$ 与存贮图像 $|\nu\rangle$ 只有第一个四元数神经元中的一位不同，即

$$\Delta S_1 = S_1 - S_1^\nu \neq 0$$

$$S_i = S_i^\nu \qquad (i > 1)$$

如 $S_1^\nu \doteq 1+\vec{i}+\vec{j}+\vec{k}$, 而 $S_1 = 1-\vec{i}+\vec{j}+\vec{k}$, 则易知有关系式：

$$S_1^{\nu *}\Delta S + \Delta S^* S_1^\nu = -4$$

$$\Delta S^*\Delta S = 4$$

为求图像 S 的能量函数值，我们先分析下式：

$$\begin{aligned} &\langle S|\nu\rangle\langle\nu|S\rangle \\ =\ & \sum_{i,j=2}^{N} S_i^* S_i^\nu S_j^{\nu *} S_j + \sum_{i=2}^{N} S_i^* S_i^\nu S_1^{\nu *}(S_1^\nu + \Delta S) \\ &+2\sum_{j=2}^{N}(S_1^\nu + \Delta S)^* S_1^\nu S_j^{\nu *} S_j + (S_1^\nu + \Delta S) * S_1^{\nu *} S_1^\nu (S_1^\nu + \Delta S) \\ =\ & 4N^2 + N(S_1^{\nu *}\Delta S + \Delta S^* S_1^\nu) + \Delta S^*\Delta S \\ =\ & 4N^2 - 4(N-1) \end{aligned}$$

则

$$\begin{aligned} E(s) &= -\frac{1}{2}\sum_{\mu=1}^{N}\langle S|\mu\rangle\langle\mu|S\rangle \\ &= -2N^2+2(N-1)-\frac{1}{2}\sum_{\mu\neq\nu}\langle S|\mu\rangle\langle\mu|S\rangle \end{aligned}$$

由于存贮图像及 S 图像的随机性，第二项仍可看作一微扰项，且同样可写为表达式 $A/2$, 其中 A 为具有平均值为 0, 均方差为 $\sigma=2^3\sqrt{N(M-1)}$ 的正态分布统计规律，故

$$\begin{aligned} E(S) &\approx -2N^2+2(N-1)+A/2 \\ &= E(S^\nu)+2(N-1) \\ &> E(S^\nu) \end{aligned} \tag{9}$$

所以，对于该四元数神经网络模型，存贮图像在能量函数空间中对应一能量极小点。

4　异步离散复相角神经网络模型的能量函数研究

与 Hopfield 模型类似，该模型同样可工作在完全异步方式下，或完全同步方式下，也可工作在部分同步方式下。下面我们考察该网络工作在完全异步方式下，且不考虑自反馈连接作用，即不考虑自能情况时，网络能量函数的变化情况。异步工作方式是指在任一时刻 t, 只有某一按某种顺序轮换或完全随机选择的神经元 i 的状态发生变化，而其他神经元的状态不变。

不失一般性，设在时刻 t, 只有神经元 i 的状态发生变化:

$$\Delta S=S_1'-S_1$$

则能量变化量为:

$$\begin{aligned} \Delta E &= E'-E \\ &= -\frac{1}{2}\sum_{\mu=1}^{M}(\langle S'|\mu\rangle\langle\mu|S\prime\rangle-\langle S|\mu\rangle\langle\mu|S\rangle) \\ &= -\frac{1}{2}\sum_{j=2}^{N}\sum_{\mu=1}^{M}[S_j^{\mu*}S_j^{\mu}S_1^{\mu*}(S_1'-S_1)+(S_1'-S_1)^*S_1^{\mu}S_j^{\mu*}S_j^{\mu}] \\ &= -\frac{1}{2}\sum_{j=2}^{N}\sum_{\mu=1}^{M}(S_1^{\mu}S_j^{\mu*}S_j^{\mu})^*(S_1'-S_1)+(S_1'-S_1)+(S_1'-S_1)^*\sum_{j=2}^{N}\sum_{\mu=1}^{M}S_1^{\mu}S_j^{\mu*}S_j^{\mu} \end{aligned}$$

若不考虑自能作用，$J_{ii}=0$, 因此

$$H_1=-\frac{1}{2}\sum_{j\neq1}\sum_{\mu=1}^{N}(S_1^{\mu}S_j^{\mu*}S_j) \tag{10}$$

并对四元数模型，设

$$H_1=H_{10}+\vec{i}H_{11}+\vec{j}H12+\vec{k}H_{13}$$

$$\Delta S=\Delta S_0+\vec{i}\Delta S_1+\vec{j}\Delta S_2+\vec{k}\Delta S_3$$

所以得能量差为：

$$\begin{aligned}\Delta E &= -\frac{1}{2}[H_1^*\Delta S + \Delta S^* H_1] \\ &= -\sum_{i=0}^{3} H_{1i}\Delta S_i \end{aligned} \tag{11}$$

从而由 $\Theta(x)$ 的取值规则知：当 H_{1i} 不小于 0 时，ΘH_{1i} 取为正单位元，即 ΔS_i 非负；当 H_{1i} 小于 0 时，ΘH_{1i} 取为负单位元，即 ΔS_i 非正。因此总有

$$\Delta E \leq 0 \tag{12}$$

也就是说，任意一个神经元，当其输出发生变化时，网络能量函数都将减少，也即在异步方式下，且不考虑自能情况时，网络能量函数 E 随时间单调下降。由于能量有界，系统必趋于一个稳定状态，并对应于能量函数状态空间的局域极小值。

上节我们得出对于四元数神经网络模型，当网络的神经元个数 N 远大于存贮图像数 M 时，存贮图像在能量函数空间中对应一能量极小点，因此存贮图像为网络的稳定吸引子，一定噪声范围的输入图像最终将稳定在该稳定吸引子上。

对于工作在完全同步方式或部分同步方式下的四元数神经网络，由于在一次迭代中常不只一个神经元发生变化，其网络能量函数 E 随时间的变化关系将不满足式 (14)，情况会变得较复杂，有待进一步研究。

5 结论

本文对四元数网络定义了一个能量函数，通过对能量函数的分析，证明了异步离散四元数神经网络模型在其动力学演化过程中，网络的能量函数单调递减，网络最终将稳定在一个平衡态上。网络的存贮图像在能量函数空间中对应一能量极小点，因此存贮图像为网络的稳定吸引子。

参考文献

1 J.J.Hopfield.*Proc.Natl Acad.Sci.U.S.A.*, 1982, 79:2554-2558

2 A.J.Noest.*Phys.Rev.A*,1988, 38:2196-2199

3 I.Kanter.*Phys.Rev.A*,1988, 37: 2739-2742

4 H.Rieger.*J.Phys.A*,1990, 23: L1273-L1279

5 J.W.Shuai,Z.X.Chen,R.T.Liu,B.X.Wu.*SICOESE'94* (第二届国际光电子科学及工程学术会议),p. Ⅱ -26

6 E.U.Condon,*et al*,*Handbook of Physics*.New York:McGraw-Hill Second Edition,1967,1:22

7 J.W.Shuai,Z.X.Chen,R.T.Liu,B.X.Wu.*The Hamilton neural network model*,*Physica A*,1995,216:20-31

An Investigation of the Energy Function of the Sequential Hamilton Number Neural Network

Chen Zhenxiang　*Shuai Jianwei*　*Liu Ruitang*　*Wu Boxi*

(Physical Department,Xiamen University,Xiamen 361005)

Abstract

In this paper, it is demonstrated that the energy of the random sequential Hamilton number neural network model decreases monotonously with time and so the network must end up in a state of equilibrium, the stored patterns are local minima of the energy function and are stable attractors of the network.

Key words: neural network, the Hamilton number, pattern recognition

陈振湘　男，副教授。1964 年毕业于厦门大学物理系，留系工作至今。一直从事发光物理研究，近期从事神经网络研究，主持电致发光神经网络自动纠错样机的设计及研制工作，发表论文多篇。