第 34 卷 第 6 期　厦门大学学报(自然科学版)　Vol. 34 No. 6
1995 年 11 月　Journal of Xiamen University (Natural Science)　Nov. 1995

异步复相角神经网络模型的能量函数研究①

陈振湘　帅建伟　刘瑞堂
(物　理　学　系)

摘要　证明了具有零对角的厄米联结矩阵的异步离散复相角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不动点吸引子上;当网络的神经元个数远大于存贮图象数时,随机存贮图象在能量函数空间中对应一能量极小点,因此存贮图象为网络的不动点吸引子.

关键词　神经网络,复相角模型,图象识别
中国图书分类号　O 411.3

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象自动识别. 对于多值图象的识别,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],用多个二值神经元来表示一个多态象点;另一类工作是用一个多值神经元表示一个 Q 值象点的状态,从而提出一些多值神经网络模型[2~6].

Noest 在文[2,6]中,提出了用复平面单位圆上等距的 Q 个离散点表示 Q 值象点,利用复数的转动特性建立了复相角神经网络模型,并认为研究复相角神经网络模型有以下几个生物的或技术上的理由:在神经系统中存在着相互作用的有限循环振荡,大脑的一些部分,如在嗅觉的根部和简单生物的运动图象的生成中,似乎用了局域反馈神经回路来作为结构单元网络,该振荡现象可能暗示了复相角神经网络模型具有生物学上的存在依据;复相角神经网络模型可用来处理具有循环值域的信号,如图象中的边缘定向或光流方向,或探测器阵列所探测的波相位图;从光神经网络的技术实现来看,利用光束的相角来对信息进行编码,也是具有吸引力的,因为基于振幅的器件的光开关速度,较基于饱和光放大器的速度慢. 由于该模型具有以上所提及的各种特点,值得我们深入的研究. 本文先简要讨论了复相角神经网络模型及其几个特例,并通过对网络能量函数的考察,较详细地分析了异步离散复相角神经网络模型的存贮图象性质和动力学演化规律.

1　复相角神经网络模型简述

在平面上绕原点转动 θ,可以用 $R(\theta)=e^{i\theta}=\cos\theta+i\sin\theta$ 表示,令转角 $0\leqslant\theta<2\pi$,而 $i=$

① 本文 1995-01-10 收到;　国家自然科学基金重点资助项目

$\sqrt{-1}$. 对于正转动 $R(\theta)$，其共轭反转动为 $R(-\theta)$. 这里有归一化等式: $|R(\theta)|=1$.

把平面转动引进神经网络，每个神经元表示一个转动操作，令神经元为

$$S=\cos(n\alpha+\beta)+i\sin(n\alpha+\beta)=e^{i(n\alpha+\beta)}\equiv[n] \tag{1}$$

其中 $\alpha=2\pi/Q$,Q 为正整数,$n=0,1\cdots,Q-1$. 也即神经元的取值为 Q 个平面离散转动态而 β 为最小转动角. 该复相角神经元状态可简记为$[n]$. 由定义式(1)，可以看出神经元的状态实际上由 $n=0,1\cdots,Q-1$ 来描述，也即神经元具有 Q 个态.

设网络有 N 个复相角神经元 S，存储 M 个 Q 态记忆模式 $S^{\mu}(\mu=1,2,\cdots M)$，则连接矩阵为

$$J_{ij}=\sum_{\mu=1}^{N}S_i^{\mu}(S_j^{\mu})^* \tag{2}$$

$(S_i^{\mu})^*$ 表示取 S_i^{μ} 的共轭. 该连接矩阵 J 具有性质

$$J=J^* \tag{3}$$

J^* 表示 J 的共轭转置矩阵，也即 J 为厄米矩阵，故其有效局域场为

$$H_i=\sum J_{ij}S_j \tag{4}$$

有效局域场实际上为一合成转动，可表达为 $H_i=|H_i|e^{i(x_i\alpha+\beta)}$，这里 x_i 为实数，$x_i\alpha+\beta$ 表示合成转动角的大小. 把它转换为归一化局域场

$$h_i=\frac{\sum J_{ij}S_j}{|H_i|}=[x_i] \tag{5}$$

对任一输入态，网络的动力学方程为

$$S_i(t+1)=\Theta H_i(t)=\Theta h_i(t)=[m_i] \tag{6}$$

当且仅当 $-1/2<m_i-x_i\leqslant 1/2$.

其几何意义是:神经元把各分转动操作合成为一总转动操作，即得有效局域场值，再把该总转动操作经输出函数作用后，映射到 Q 个神经元离散转动态的最近邻态上. 也即在复平面的单位圆上，均匀分布着 Q 个点，以这 Q 个点为中心，把单位元等分成 Q 个扇形，输出函数则把落在每个扇形内的旋转操作映射到其中心的离散点上.

对于复相角神经网络模型，我们可以引入一能量函数定义[6]

$$E(S)=-\frac{1}{2}\sum_{i=1}^{N}H_iS_i^*=-\frac{1}{2}\sum_{i,j=1}^{N}J_{ij}S_i^*S_j \tag{7}$$

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(7)为一实数，该能量函数的共轭函数为

$$E^*(S)=-\frac{1}{2}\sum_{i,j=1}^{N}(J_{ij}S_i^*S_j)^*=-\frac{1}{2}\sum_{i,j=1}^{N}J_{ij}^*S_iS_j^*=-\frac{1}{2}\sum_{i,j=1}^{N}J_{ij}S_i^*S_j=E(S)$$

也即能量函数定义式为一实数.

2　复相角神经网络模型的两个特例

对于复相角模型，若令 $Q=2,\beta=0$ 则神经元状态为实数 $S=\pm1$，若演化动力学方程中的两个映射扇形以虚轴为界，即网络的动力学方程为:当其局域场不小于零时，神经元为$+1$;反之则为-1. 实际上得到的是双极 Hopfield 模型[7]. 当该模型退化为双极 Hopfield 模型时，式(7)也变成 Hopfield 模型的能量表达式，它说明能量定义式(7)是合理的.

对于复相角模型,若令 $Q=4$,$\beta=\pi/4$,则神经元状态为 $S=\sqrt{2}/2(\pm 1\pm i)$,为四象限的角平分线. 我们可取动力学方程中的四个映射扇形分别为四象限. 由于转动模型中求的是角度,实际上神经元是否归一化,也即是否 $|S|=1$,对网络的动力学演化并无影响,只要取统一的模值即可,这样可统一令神经元取值为 $(\pm 1\pm i)$,则网络的上述动力学方程有一简单的表述形式,它可避免先换算为角度后再映射的过程,具体形式如下:

$$S_i(t+1)=\Psi\{\sum J_{ij}S_j(t)\} \tag{8}$$

函数 $\Psi\{x\}$ 的取值规则如下:当 x 的实部及虚部不小于 0 时,$\Psi\{x\}$ 的对应部分取为 $+1$ 或 $+i$,当 x 的实部及虚部小于 0 时,$\Psi\{x\}$ 的对应部分取为 -1 或 $-i$. 该模型即是文献[8]所提出的四态复数神经网络模型,也是文[5]中当 $n=1$ 时的复数神经网络模型.

3 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,该模型同样可以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. 下面我们证明网络在异步工作条件下,当联接矩阵 $J=J^*$ 且 $J_{ii}=0$ 时,网络的能量函数随时间必定单调下降,因而系统必稳定在一个能量局域极小的不动点吸引子上. 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,而其它神经元的状态不变.

不失一般性,设在时刻 t,只有神经元 k 的状态发生变化

$$S'_k=\Theta(H_k)=[m], \qquad S_k=[n]$$

则能量变化量 ΔE 为

$$\begin{aligned}\Delta E=E'-E&=-\frac{1}{2}\sum_{i,j=1}^{N}(J_{ij}S'^*_iS'_j-J_{ij}S_i^*S_j)\\&=-\frac{1}{2}\sum_{j\neq k}^{N}(J_{jk}S_j^{\mu *}(S'_k-S_k)+J_{kj}(S'_k-S_k)^*S_j^{\mu}\\&=(-\frac{1}{2}\sum_{j\neq k}^{N}J_{kj}S_j^{\mu})^*(S'_k-S_k)+(-\frac{1}{2}\sum_{j\neq k}^{N}J_{kj}S_j^{\mu})(S'_k-S_k)^*\end{aligned}$$

若不考虑自能作用,$J_{ii}=0$,因此有效局域场为

$$H_i=\sum_{j\neq k}J_{ij}S_j=|H_i|h_i \tag{9}$$

所以得能量差为

$$\Delta E=-\frac{1}{2}|H_k|[h_k^*(S_k'-S_k)+h_k(S_k'-S_k)^*]$$

设归一化局域场 $h_k=[x]$,则可得

$$\begin{aligned}\Delta E&=-\frac{1}{2}|H_k|\{[-x][m]-[-x][n]+[x][-m]-[x][-n]\}\\&=-\frac{1}{2}|H_k|\{[m-x]+[x-m]-[n-x]-[x-n]\}\\&=-|H_k|[\cos(m-x)\alpha-\cos(n-x)\alpha]\\&=2|H_i|\sin\frac{m+n-2x}{2}\alpha\sin\frac{m-n}{2}\alpha\end{aligned} \tag{10}$$

可证该式为一非正值:若 $m>n$,则有 $x>n$,且 $x-n>1/2$,又因 $-1/2<m-x\leqslant 1/2$,$m\alpha<2\pi$,故得

$$-\pi<\frac{m+n-2x}{2}\alpha\leqslant 0,\qquad 0\leqslant\frac{m-n}{2}\alpha<\pi$$

而若 $m<n$,则有 $x<n$,且 $n-x>1/2$,$n\alpha<2\pi$,可得

$$-\pi<\frac{m-n}{2}\alpha\leqslant 0,\qquad 0\leqslant\frac{m+n-2x}{2}\alpha<\pi$$

也即式(14)中的两个正弦角必有一个在上半轴,而另一个在下半轴,从而有

$$\Delta E\leqslant 0 \tag{11}$$

也就是说,任意一个神经元,当其输出发生变化时,网络能量函都将减少. 上面的推证中关于连接矩阵 J 只用了关系式(3)和 $J_{ii}=0$,它说明对任意具有零对角的厄米联结矩阵网络,在异步方式下网络能量函数 E 随时间必定单调下降. 由于能量有界,系统必稳定在一个不动点吸引子上,并对应于能量函数状态空间的局域极小值.

对于工作在完全同步方式或部分同步方式下复相角神经网络,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,与 Hopfield 模型类似,网络不一定稳定在一个不动点吸引子上,有可能稳定在一周期吸引子上,有待进一步具体分析.

4　存贮图象的能量函数研究

对随机图象存贮的 Hopfield 模型,当网络的神经元个数 N 远大于存贮图象数 M 时,在能量函数空间,存贮图象为 Hopfield 网络的能量极小值. 对于复相角神经网络模型,同样可证当网络的神经元个数远大于存贮图象数时,且联接矩阵 $J_{ij}=\sum S_i^{\mu}(S_j^{\mu})^*$ 时,存贮图象在能量函数空间中对应一能量极小点.

对任一存贮图象 S^{v},其能量为

$$\begin{aligned}E(S^{v})&=-\frac{1}{2}\sum_{i,j=1}^{N}\sum_{\mu=1}^{N}S_i^{\mu}S_j^{\mu *}S_i^{v *}S_j^{v}\\&=-\frac{1}{2}N^2-\frac{1}{2}\sum_{i,j=1}^{N}\sum_{\mu\neq v}S_i^{\mu}S_j^{\mu *}S_i^{v *}S_j^{v}\\&\approx-\frac{1}{2}N^2+A\end{aligned}\tag{12}$$

第二项 A,实际上为一噪声项,具体展开为实部和虚部时,为$2^4N^2(M-1)$项之和,其中一半为虚部,相互抵消. 而余下$2^3N^2(M-1)$项实数和,且每项又为4个正余弦值的积,由于存贮图象的随机性,可得 A 的统计平均值为0,均方差为 $\sigma=\sqrt{N^2(M-1)}$,且具有正态分布

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{x^2}{2\sigma^2}\right)\tag{13}$$

当 $N\gg M$ 时,噪声项 A 很小,可看作一微扰项.

设一输入图象 S 与存贮图象 S^{v} 只有某一神经元的取值不同,不失一般性,设

$$\Delta S=S_1-S_1^{v}\equiv[n]-[m]\neq 0$$

$$S_i=S_i^{v}\qquad(i>1)$$

为求图象 S 的能量函数值,我们先分析下式:

$$\begin{aligned}&\sum_{i,j=1}^{N}S_iS_j^{*}S_i^{v *}S_j^{v}\\&=\sum_{i,j=2}^{N}S_iS_j^{*}S_i^{v *}S_j^{v}+\sum_{j=2}^{N}(S_i^{v}+\Delta S)S_j^{*}S_i^{v *}S_j^{v}\end{aligned}$$

$$
\begin{aligned}
&+\sum_{i=2}^{N} S_i(S_i^v+\Delta S)^* S_i^{v*} S_i^v+(S_i^v+\Delta S)(S_i^v+\Delta S)^* S_i^{v*} S_i^v \\
&=N^2+N(S_i^{v*}\Delta S+S_i^v\Delta S^*)+\Delta S\Delta S^* \\
&=N^2+N\{[-m]([n]-[m])+[m]([-n]-[-m])+([n]-[m])([-n]-[-m]\} \\
&=N^2+N([n-m]-1+[m-n]-1)+2-[n-m]-[m-n] \\
&=N^2-2[N-1][1-\cos(n-m)\alpha]
\end{aligned}
$$

上式最后一步用了下列表达式

$$
2-[n-m]-[m-n]=2[1-\cos(n-m)\alpha] \tag{14}
$$

因此，图象 S 的能量值为

$$
\begin{aligned}
E(S)&=-\frac{1}{2}\sum_{i,j=1}^{N} J_{ij}S_i^* S_j \\
&=-\frac{1}{2}N^2+(N-1)[1-\cos(n-m)\alpha]-\frac{1}{2}\sum_{i,j=1}^{N}\sum_{\mu\neq v} S_i^\mu S_j^{\mu *} S_i^* S_j
\end{aligned}
$$

由于存贮图象及 S 图象的随机性，第二项仍可看作一微扰项，同样具有平均值为0，均方差为 $\sigma=\sqrt{N^2(M-1)}$ 的正态分布统计规律. 故

$$
\begin{aligned}
E(S)&\approx-\frac{1}{2}N^2+(N-1)[1-\cos(n-m)\alpha]+A \\
&=E(S^v)+(N-1)[1-\cos(n-m)\alpha] \\
&>E(S^v)
\end{aligned} \tag{15}
$$

所以，对于由矩阵式(2)连接构成的复相角神经网络模型，当 $N\gg M$ 时，存贮图象在能量函数空间中对应一能量极小点.

结合上一节所推出的结论，我们可进一步得到如下的结果：当网络的神经元个数 N 远大于存贮图象 M 时，对于由矩阵式(2)连接构成的复相角神经网络模型，存贮图象不仅在能量函数空间中对应一能量极小点，而且在异步工作方式情况下且不考虑自反馈作用时，存贮图象为网络的不动点吸引子. 因此在一定噪声范围的输入图象最终将稳定在该吸引子上.

5 结　论

本文先简要论述了复相角神经网络模型，指出了双极 Hopfield 模型和四态复数神经网络模型为该模型的两个特例，证明了具有零对角的厄米联结矩阵的异步离散复相角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不动点稳定吸引子上. 当网络的神经元个数远大于存贮图象数时，随机存贮图象在能量函数空间中对应一能量极小点，因此存贮图象为网络的稳定不动点吸引子.

实际上，复相角神经元为 Q 态神经元，网络对图象的识别是按照一种整体联想记忆装置进行工作的，在图象识别应用中，复数间的运算关系与象素间的变换关系并不一定要求存在着数学或物理意义上的一一对应特征，重要的是把该复相角神经元的每一个状态与多值象点的不同状态对应起来，所以我们认为该模型不仅可用来处理具有循环值域的信号，而且可处理任意的 Q 态图象，如 Q 态灰度或彩色图象.

参 考 文 献

1 Hopfield J J. Neural networks and physical systems with emergent collective computertational abilities. *Proc, Natl Acad, Sci U. S. A.* ,1982, 79: 2 554～2 558

2 Noest A J. Discrete-state phasor neural networks. *Phys, Rev. A*, 1988, 38: 2 196～2 199

3 Kanter I. Potts-glass models of neural networks. *Phys, Rev. A*, 1988, 37: 2 739～2 742

4 Rieger H. Storing an extensive number of grey-toned patterns in a neuralnetwork using multistate neurons. *J. Phys. A* ,1990, 23: L1 273～L1 279

5 帅建伟，陈振湘，刘瑞堂等. 2^n 元数神经网络模型：存贮容量分析. 厦门大学学报(自然科学版)，1995，34：353～358

6 Noest A J. Associative memory in sparse phasor neural networks. *Europhys. Lett.* ,1988, 6: 469～474

7 Amit D J, Gurfreund H , Sompolinsky H. Storinginfinite numbers of patterns in a spin-glass model of neural networks. *Phys. Rev. Lett.* ,1985, 55: 1 530～1 533

8 Zhou C H, Liu L R. Thecomplex Hopfield model. *Optics Comm.* ,1993, 103: 29～32

An Investigation of the Energy Function of the Sequential Complex Phasor Neural Network Model

Chen Zhenxiang　　Shuai Jianwei　　Liu Ruitang

(Dept. of Phys.)

Abstract In this paper, it is demonstrated that the energy of the random sequential complex phasor neural network model with zero-diagonal ermy connection matrix decreases monotonically with time and so the network must end up in a fixed spot attractor. When the number ofstored patterns is much smaller than the total number of neurons in the network, the storedpatterns are local minima of the energy function and so are the fixed spot attractors of the network.

Key words Neural network, Complex phasor neural network, Pattern recognition