

一种基于能量人工神经元模型的自生长、自组织神经网络

班晓娟¹ 刘浩¹ 徐卓然¹

摘要 本文结合近年生物学中神经科学的发展, 针对神经胶质细胞对生物神经网络的生长提供能量支持的特性, 将神经胶质细胞的能量模型引入到人工神经元的概念模型中, 提出了能量人工神经元 (Energy artificial neuron, EAN) 的概念模型, 并给出了数学表述. 同时, 在能量人工神经元模型的基础上, 实现了一种新型自生长、自组织人工神经网络 (EAN based self-growing and self-organizing neural network, ESGSONN), ESGSONN 将神经元中的能量、网络的熵增量及样本与神经元权值的相似度的竞争作为生长的条件, 并对最优生长点中的获胜神经元进行单位步长调整. ESGSONN 实现了快速生长、精确的样本数据分布密度保持、死神经元少的特性. 本文使用经典的 16 种动物实验 (Ritter and Kohonen, 1989) 验证了 ESGSONN 的正确性, 并通过同 SOFM、GCS 等自组织网络的对比实验验证 ESGSONN 网络的特性. 最后, 本文对 ESGSONN 在高维空间中的本质进行了讨论.

关键词 自组织, 生长网络, 能量人工神经元, 无监督学习, 高维空间, 熵, 聚类

DOI 10.3724/SP.J.1004.2011.00615

An Energy Artificial Neuron Model Based Self-growing and Self-organizing Neural Network

BAN Xiao-Juan¹ LIU Hao¹ XU Zhuo-Ran¹

Abstract In this paper, we established a new artificial neuron model called EAN (energy artificial neuron) based on the energy concept from the glial cells according to the recent achievements in the neuroscience field. We suggested a way to demonstrate EAN model in mathematics. In addition, we realized a self-growing and self-organizing neural network based on the EAN model called ESGSONN (EAN based self-growing and self-organizing neural network). ESGSONN considers the energy in EAN, the entropy productions in the network and the similarity (between the sample and neurons' weights) as its conditions of growing and competitions. Its main features are described as below: rapid growing, probability density preserving and few superfluous neurons. A classical experiment of 16-kind animals (after Ritter and Kohonen, 1989) proved ESGSONN can work correctly. We showed the new features of ESGSONN by comparing it with the traditional self-organizing networks such as SOFM and GCS. Finally, we argued about the essence of this network in the high-dimensional space.

Key words Self-organizing, growing network, energy artificial neuron (EAN), unsupervised learning, high-dimensional space, entropy, clustering

人脑的形态通常呈现出高度自组织的特征, 而早在 20 世纪 70 年代, 就有学者开始研究能够模拟人脑生物学规律的人工神经网络和算法^[1]. 目前按照网络的结构特征可以将自组织神经网络分成两大类: 一类是经典的具有固定结构的网络,

如 1982 年 Kohonen 提出的自组织特征映射网络 (Self-organizing feature maps, SOFM)^[2-3]; 另一类是具有动态增长特性的网络, 如 Fritzke 提出的增长细胞结构 (Growing cell structure, GCS)^[4]、生长网格网络 (Growing grid, GG)^[5] 和生长神经气网络 (Growing neural gas, GNG)^[6]. GCS 是基于 SOFM 的神经网络, 其用神经元间联结构成的三角形网络代替 SOFM 的二维平面, 每个三角形都分布在输入空间中非零概率密度的位置上. 算法以一个随机三角形开始, 用启发式方法增加神经元和联结, 生成三角形网络, GNG 是基于 NG 的神经网络, 初始时拥有两个神经元, 通过训练建立及移除节点, 直到神经元生长到指定数目, Choi 等提出的自创造、自组织神经网络 (Self-creating and organizing neural networks, SCONN)^[7], Martinetz 等提出的神经气网络 (Neural gas, NG)^[8], 其神经元个数固

收稿日期 2010-08-03 录用日期 2010-12-27
Manuscript received August 3, 2010; accepted December 27, 2010

国家高技术研究发展计划 (863 计划) (2009AA062800, 2009AA04Z163), 国家自然科学基金 (60973063), 北京自然科学基金 (4092028), 中央高校基本科研业务费专项资金 (FRF-TP-09-016B) 资助

Supported by National High Technology Research and Development Program of China (863 Program) (2009AA062800, 2009AA04Z163), National Natural Science Foundation of China (60973063), Beijing Natural Science Foundation (4092028), and Fundamental Research Funds for the Central Universities (FRF-TP-09-016B)

1. 北京科技大学信息工程学院 北京 100083
1. Information School of Engineering, University of Science and Technology Beijing, Beijing 100083

定,但初始时神经元间无联结,随着训练数据的输入,神经元对数据进行适应,在获胜神经元和其邻近竞争者间产生联结,Blackmore 的增量式生长神经网络 (Incremental grid growing, IGG)^[9],可根据输入数据改变神经网络,其以几个神经元开始,通过启发式生长法在网络边缘生长神经元,当内部节点权值偏差量低于某一阈值时增加联结,反之,当节点权值偏差量增加时移除联结,Chow 等提出的细胞分裂神经网络 (Cell-splitting grid, CSG)^[10],其模拟细胞的分裂机制,在细胞到达某一年龄时会进行分裂,分裂后的部分会保持原细胞部分属性;此外还有 Alahakoon 等基于 IGG 提出的动态自组织映射网络 (Dynamic self-organizing map, DSOM)^[11],Shen 的自组织增量式神经网络 (Self-organizing incremental neural network, SOINN)^[12-14] 和加强的自组织增量式神经网络 (Enhanced self-organizing incremental neural network, ESOINN)^[15] 等. 由于自组织神经网络具有无监督学习,不需要先验信息的特点,这种方法逐步被运用到模式识别、聚类、图像处理、计算 NP 完全问题的近似解等诸多领域中.

但是,传统的自组织网络模型通常具有如下的特性: 1) 网络规模不可调整,容易产生死神经元,或神经元生长的条件由局部误差确定,使得网络生长缓慢; 2) 使用平均量化误差 (Mean quantization error, MQE) 作为网络收敛的判断,导致需要样本反复迭代,训练周期长; 3) S & P 二难问题.

本文结合近年生物学中神经科学的发展,将神经胶质细胞的能量模型引入到人工神经元的概念模型中,提出了能量人工神经元的概念模型,并给出了数学表述. 同时,在能量人工神经元模型的基础上,实现了一种新型的自生长、自组织人工神经网络—ESGSONN, ESGSONN 将神经元的能量、网络的熵增量及相似度竞争作为生长的条件,并对最优生长点中的获胜神经元进行单位步长调整. 实验表明, ESGSONN 实现了快速生长、较精确的样本数据密度保持、死神经元少的特性.

1 能量人工神经元 (EAN) 模型

1.1 生物学中神经胶质细胞的重要作用

组成大脑的细胞主要有两大类: 神经元和神经胶质细胞. 在过去的生物学研究中,一直把神经元看作是大脑活动的主要对象,而神经胶质细胞被认为仅仅起到神经元的支撑骨架、提供营养成分等从属作用. 1997 年, Pfrieger 与 Barres 在 *Science* 上发表文章,首次指出神经胶质细胞能强烈促进神经元间的突触联系后^[16],神经生物学研究学者相继在

Nature 与 *Science* 上发表论文, Ullian 等通过定量实验验证了神经胶质细胞能够显著提高突触数量及增强突触传递效率,同时证明了神经胶质细胞是维持突触的必要条件^[17], Song 等发现神经胶质细胞能够明显提高神经发生率^[18], Slezak 等证实突触周围的神经胶质细胞是突触的构成要素^[19]. 其中,神经胶质细胞对神经元的能量供给是神经胶质细胞的重要功能之一,这主要体现在神经胶质细胞通过氧化代谢可以提高葡萄糖的利用率并且释放乳酸从而产生能量^[20-21] (以三磷酸腺苷形式) 对其周围的神经元产生调节作用,这种对神经元的调节可以表现为激活或抑制^[22-23],并以此为基础完成其他复杂的功能,图 1 中为两个神经元之间的星状神经胶质细胞 (Astrocyte).

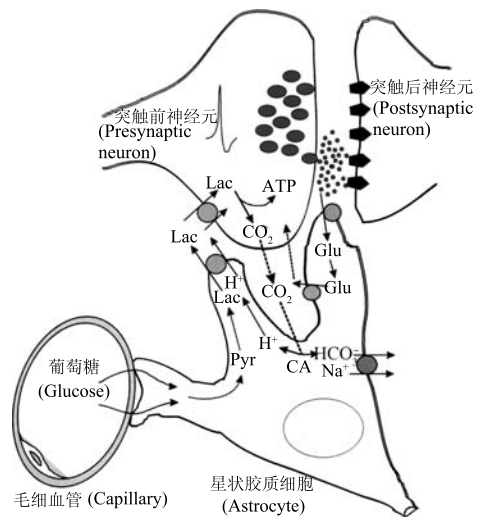


图 1 两个神经元之间的星状胶质细胞^[20]
Fig. 1 An astrocyte between two neurons^[20]

这些生物学研究表明,神经胶质细胞在神经系统的高级活动中具有复杂的功能,对神经系统的发生、突触的形成以及学习与记忆具有重要意义. 随着生物学研究的发展,大脑应该被看作是神经元和神经胶质细胞构成的整合网络^[24].

1.2 能量人工神经元 (EAN) 的概念模型

为了表征生物学中神经胶质细胞可为神经元的生长提供所需能量,本文在传统 M-P 神经元模型^[25]的基础上,引入神经胶质细胞中能量对神经元的调节作用,从而构成一个有能量特征的人工神经元 (Energy artificial neuron, EAN) 概念模型. EAN 模型使得神经元模型具有能量的特性,从而为自生长、自组织神经网络提供生长条件. 为了避免引入人工神经胶质模型而使系统的复杂度增加,本文对生物学中神经胶质细胞的功能特性进行了简化,仅保留其对神经元的能量支持的概念,以此表达神经

元之间建立连接的条件.

定义 1. 一个能量人工神经元 (EAN) 的概念模型可由如下一个七元组来表示:

$$EAN = \langle \mathbf{X}, \mathbf{O}, M, I, F, E_{AN}, L_{AN} \rangle$$

其中:

\mathbf{X} 表示输入向量, $\mathbf{X} = [x_1, \dots, x_n]^T$;

\mathbf{O} 表示神经元计算后的输出, $\mathbf{O} = \{o(t)\}$;

M 表示局域记忆单元, $M = M_W \cup M_L$, 其中, $M_W = \{u\}$ 为工作记忆, $M_L = \{w = \{w_i\}_{i=1}^n, b\}$ 为长时记忆;

I 表示整合运算, 如 $I: u(t) = I(\mathbf{w}, \mathbf{x}) = \mathbf{w}^T \mathbf{x}(t)$;

F 表示激发函数, 如 $F: o(t) = F(u(t) - \theta)$;

E_{AN} 表示当前神经元的总能量;

L_{AN} 表示广义 Hebb 学习律.

改进后的能量人工神经元 (EAN) 模型结构如图 2 所示.

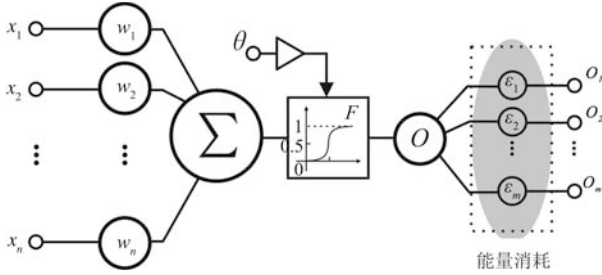


图 2 能量人工神经元 (EAN) 模型基本结构

Fig. 2 The structure of the EAN model

定义 2. 一个神经元拥有的总能量为:

$$E_{AN} = E_{AN}^{\text{consumption}} \cup E_{AN}^{\text{overplus}}$$

其中, $E_{AN}^{\text{consumption}} = \{e_{ij} | i, j = 1, 2, \dots, m\}$ 表示当前神经元用于与其他神经元连接的能量, e_{ij} 表示在 t 时刻, 当前神经元的输出分量 i 与其周围的神经元 j 在 $t+1$ 时刻产生连接时所消耗的能量, E_{AN}^{overplus} 表示当前神经元能量的剩余.

定义 3. 在定义 2 中的 $E_{AN}^{\text{consumption}}$ 中, 一次连接建立时产生的能量消耗为: $e_{ij} = \varepsilon_{ij} \times \bar{e}$, 其中, \bar{e} 表示两个神经元间建立连接时消耗能量的期望值, ε_{ij} 为当前神经元的输出分量 i 与其拓扑相邻的神经元 j 连接时消耗能量的系数, 记为这两个神经元

$$\text{之间的能量权值 } \varepsilon_{ij}, W_e = \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \cdots & \varepsilon_{1j} \\ \varepsilon_{21} & \varepsilon_{22} & \cdots & \varepsilon_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{i1} & \varepsilon_{i2} & \cdots & \varepsilon_{ij} \end{bmatrix}$$

为神经元的能量权值矩阵. E_{AN}^{overplus} 与前同义. 由此, 一个神经元的能量可通过式 (1) 来计算:

$$E_{AN} = \bar{e} \times \left(\sum_{i,j=1}^n \varepsilon_{ij} + \sum_{i,j=1}^m \varepsilon'_{ij} \right) \quad (1)$$

性质 1. 对 $\forall \varepsilon_{ij}$ 都满足 $\varepsilon_{ij} \times \bar{e} \leq E_{AN}^{\text{overplus}}$, 换言之, 若 $\exists \varepsilon_{ij}$ 使得 $E_{AN}^{\text{overplus}} < \varepsilon_{ij} \times \bar{e}$, 则神经元的输出分量 i 不能与神经元 j 建立连接.

性质 2. 神经元的能量仅在神经元之间建立连接时消耗, 连接一旦建立, 神经元之间的连接强度由其连接权值决定, 与神经元的能量权值无关.

2 ESGSONN: 基于能量人工神经元模型的自生长、自组织神经网络

2.1 ESGSONN 的相关概念

ESGSONN 是以能量人工神经元模型 (以下简称能量神经元) 为基础构建的神经网络. ESGSONN 将时刻 t 能量神经元中的能量、网络的熵增量及与样本的相似程度作为最优生长点的判决条件, 通过神经元的生长、删除、合并形成聚类决策域, 然后在网络的输出层生长出神经元, 神经元与决策域一一对应. 当训练结束后, 输出层中的神经元的个数代表了聚类的个数, 而其决策域本质上是对应的第二层拓扑相邻的神经元映射在高维空间中超球的拓扑乘积, 代表了同一类样本在高维空间中的分布. 下面对 ESGSONN 中的相关概念进行定义:

定义 4. 神经元单元. 神经元单元是网络生长的基本单位. 若由六个神经元在二维空间中按照正六边形的拓扑结构排序形成图 G_{AN}^{unit} , 则 G_{AN}^{unit} 与连接权值矩阵 W_G 及在 $t+1$ 时刻的熵增量 $\Delta H_G^{\text{unit}}(t+1)$ 构成一个人工神经元单元, 即:

$$EAN^{\text{unit}} = \langle G_{AN}^{\text{unit}}, W_G, \Delta H_G^{\text{unit}}(t+1) \rangle$$

其中, $G_{AN}^{\text{unit}} = \langle V_c, E_c \rangle$, $V_c = \{EAN_i | i = 1, 2, \dots, 6\}$ 是能量神经元 EAN 的集合, $E_c = \{e_{ij} | i, j = 1, 2, \dots, 6\}$, e_{ij} 为神经元由 EAN_i 起终止于 EAN_j 的边, W_G 为 6 个神经元连接的权值矩阵, 1 个能量神经元的输出端分别与其拓扑相连的 2 个能量神经元相连, 构成六边形结构. 此外, 规定每个能量神经元能量总数为 4, 能量消耗的期望为 1 且在生长时任意两个神经元之间建立连接或充当公用神经元时所消耗的能量都相同, 即对 $\forall EAN_i \in EAN^{\text{unit}}$ 满足:

$$E_{AN} = E_{AN}^{\text{consumption}(i)} \cup E_{AN}^{\text{overplus}(i)} = 4 \quad \text{且}$$

$$W_e = \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix}$$

EAN^{unit} 中的熵增量 $\Delta H_G^{\text{unit}}(t+1)$ 可根据 EAN 中的能量约束进行香农熵的计算, 由式 (2) 给出:

$$\Delta H_G^{\text{unit}}(t+1) = - \sum_{R \in S(R)} p(R_{t+1}) \log_2 p(R_{t+1}) \quad (2)$$

其中, $p(R_{t+1})$ 可由时刻 t 的能量剩余决定. 一个神经元单元的拓扑结构如图 3 所示.

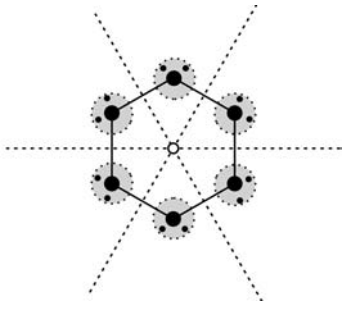


图 3 神经元单元的拓扑结构

Fig. 3 Neuron unit topological graph

定义 5. 可生长点. 可生长点是网络进行生长的地方, 可由一个二元组来表述:

$$G' = \langle V_G, E_G \rangle$$

其中, $V_G = \{EAN_i, EAN_j, \dots, EAN_k\}$ 是能量神经元 EAN_i 及与 EAN_i 拓扑相连的神经元的集合, $E_G = \{e_{ij} | i, j = 1, 2, \dots, k\}$ 是边的集合, e_{ij} 为神经元由 EAN_i 起终止于 EAN_j 的边. 若 G' 中任意一个神经元在生长中的能量消耗小于能量剩余, 且 G' 的拓扑结构图是一个神经元单元的一个真子图并外凸生长, 则称 G' 是一个可生长点. 即: $\forall EAN_i \in G'$ 都满足:

$$\sum_{i=1}^n \bar{e} \times \varepsilon_{ij} < E_{AN}^{\text{overplus}} \quad \text{且}$$

$$G' \subset G_t \wedge G_{AN(t+1)}^{\text{unit}} \subset G_{AN}^{\text{unit}}$$

其中, G_t 为当前网络图, $G_{AN(t+1)}^{\text{unit}}$ 为时刻 $t+1$ 新增的神经元单元拓扑.

性质 3. 一个神经元单元只能在可生长点生长新的神经元单元结构, 新的神经元单元与原神经元单元共用可生长点中的神经元. 每次生长过程中, 可生长点 G' 中的神经元单元的能量减少量为

$$\Delta E_{AN}^{\text{unit}} = \sum_{i=1}^n \Delta E_{AN}^{\text{consumption}(i)} = \sum_{i,j=1}^n \bar{e} \times \varepsilon_{ij}$$

其中, $\Delta E_{AN}^{\text{consumption}(i)}$ 为每个神经元能量减少量, n 为 G' 中神经元的个数. 结合定义 5 可知

$$\Delta E_{AN}^{\text{unit}} = \sum_{i=1}^n \Delta E_{AN}^{\text{consumption}(i)} = n$$

定义 6. 最佳生长点 G'^{optimal} . 在时刻 t 的网络中有 n 个可生长点, 若一个可生长点 G' 使得网络在 $t+1$ 时刻的熵增量 $\Delta H_G^{\text{unit}}(t+1)$ 最小, 且其处于 G' 拓扑中心的神经元的权值 W_c 与目标样本 X 的相似度最高, 则称该可生长点为最佳生长点. 即 G'^{optimal} 满足:

$$\Delta H_{G'^{\text{optimal}}}^{\text{unit}}(t+1) \leq \Delta H_{G'_i}^{\text{unit}}(t+1) \quad \text{且}$$

$$\|X - W_{C^{\text{optimal}}}\| < \|X - W_{C_i}\|, \quad \forall G'_i \in \bigcup_{i=1}^n G'_j$$

其中, G'_i 为 t 时刻网络中的一个可生长点.

定义 7. 三神经元单元. 若 EAN^{unit} 是一个神经元单元, 且 3 个神经元按图 4 拓扑结构排列, 则构成一个三神经元单元, 可以由如下三元组表示:

$$EAN^{\text{unit}3} = \langle G_{AN}^{\text{unit}3}, W_G, \Delta H_G^{\text{unit}3}(t+1) \rangle$$

$EAN^{\text{unit}3}$ 中共有 13 个能量神经元, 这种结构可以看作是由神经元单元通过两次生长后得到的, 如图 4 所示.

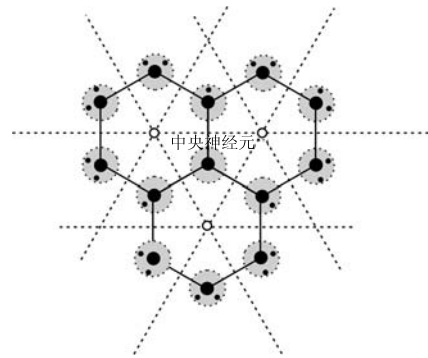


图 4 三神经元单元的拓扑结构

Fig. 4 Treble neuron unit topological graph

定义 8. 中央神经元. 若一个能量神经元 EAN_c 处于一个三神经元单元 $EAN^{\text{unit}3}$ 拓扑结构的几何中心位置, 则称 EAN_c 是三神经元单元 $EAN^{\text{unit}3}$ 的中央神经元. 结合性质 1、定义 4 和定义 7 可知, 中央神经元不能再与其他神经元相连接, 也不能作为当前三神经元单元外的公用神经元.

定义 9. 覆盖神经元. 一个神经元若其权值与样本的距离小于一个期望值, 则称此神经元为一个覆盖神经元. EAN 对应的权值满足:

$$\|W - X_i\| \leq \xi$$

其中, ξ 是经验选择常数. 覆盖神经元本质上代表了一个神经元从高维空间中以 W 为球心, ξ 为半径的超球体积域到神经元在二维平面三角形面积域的映射.

定理 1. 若在时刻 t , 网络中存在 $G'_{\text{type}1}$, $G'_{\text{type}2}$ 两种类型的生长点, 其中 $G'_{\text{type}1}$ 满足: $G'_{\text{type}1}$ 是可生长点且至少含有一个公用神经元; $G'_{\text{type}2}$ 满足: $G'_{\text{type}2}$ 是可生长点且不含有公用神经元, 如图 5 所示, 那么, 在时刻 $t+1$ 时, 沿 $G'_{\text{type}1}$ 生长的系统熵增量比沿 $G'_{\text{type}2}$ 的系统熵增量小, 即:

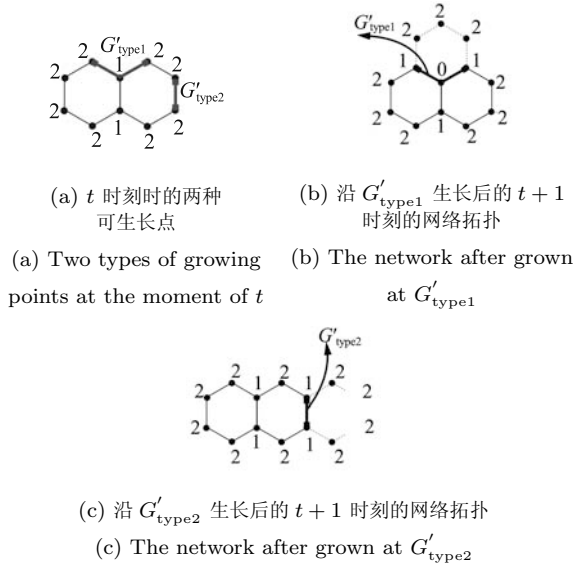


图 5 t 时刻网络中存在的两种可生长点及在 $t+1$ 时刻沿这两种可生长点生长后的网络拓扑结构

Fig. 5 Two types of growing point at the moment of t and their topological graphs after growing at the moment of $t+1$

$$\Delta H_{G'_{type1}}(t+1) < \Delta H_{G'_{type2}}(t+1)$$

证明. 设 m, n 分别为 G'_{type1} 和 G'_{type2} 在时刻 t 的生长形态总数, 则有:

$$\Delta H_{G'_{type1}}(t+1) = - \left(\sum_{i \in type1} \frac{X_i}{\sum_m X_m} \log_2 \frac{X_i}{\sum_m X_m} + \sum_{j \in type2} \frac{X_j}{\sum_m X_m} \log_2 \frac{X_j}{\sum_m X_m} \right)$$

$$\Delta H_{G'_{type2}}(t+1) = - \left(\sum_{i \in type1} \frac{X_i}{\sum_n X_n} \log_2 \frac{X_i}{\sum_n X_n} + \sum_{j \in type2} \frac{X_j}{\sum_n X_n} \log_2 \frac{X_j}{\sum_n X_n} \right)$$

假设每一种生长形态出现的概率相等, 则香农熵可退化为玻尔兹曼熵的形式:

$$\Delta H_{G'_{type2}}(t+1) - \Delta H_{G'_{type1}}(t+1) = \sum_{i \in m} \frac{X_i}{\sum_m X_m} \log_2 \frac{X_i}{\sum_m X_m} - \sum_{j \in n} \frac{X_j}{\sum_n X_n} \log_2 \frac{X_j}{\sum_n X_n} =$$

$$\sum^m \frac{1}{m} \log_2 \frac{1}{m} - \sum^n \frac{1}{n} \log_2 \frac{1}{n} = \log_2 n - \log_2 m$$

又因 G'_{type1} 中至少有一个公用神经元, 因此, 沿 G'_{type1} 生长要比沿 G'_{type2} 生长至少少增加一个神经元, 结合其拓扑结构可知, 沿 G'_{type1} 生长的形态数 m 要比沿 G'_{type2} 生长的形态数 n 少, 即 $m < n$, 则

$$\Delta H_{G'_{type1}}(t+1) < \Delta H_{G'_{type2}}(t+1)$$

□

定理 1 表明, 网络总是尽可能地向着最小熵增量的方向生长, 而这些最小熵增点往往出现在学习次数较多的地方, 学习过的神经元对类似的样本输入具有较高的响应, 即只要能量充足, 学习过的神经元在时刻 $t+1$ 具有优先学习的优势.

2.2 ESGSONN 网络的训练算法

ESGSONN 算法共有 3 个基本环节: 1) 网络的初始化; 2) 网络中神经元的产生、学习和删除; 3) 决策域的计算及生成输出神经元. 图 6 为网络生长的过程.

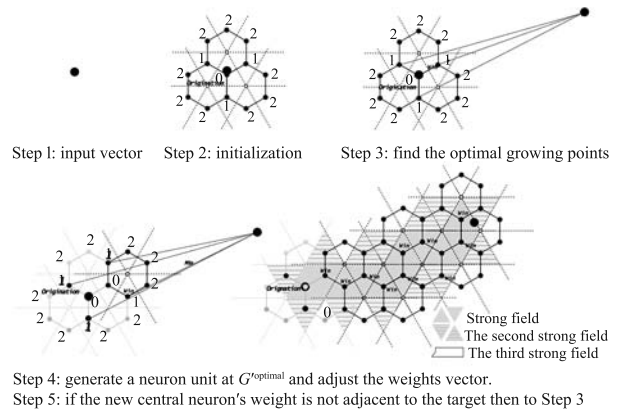


图 6 网络的生长过程

Fig. 6 The growing procedure of ESGSONN algorithm

步骤 1. 从样本空间分布 \mathbf{R}^n 中随机输入一个样本 \mathbf{X}_i 作为初始化样本, 步骤 2 之后的样本可以顺序输入或随机输入.

步骤 2. 初始化生长. 在二维平面上生长出第一个三神经元单元 EAN^{unit3} , 使满足: 若 $\forall EAN_i \in EAN^{unit3}$, 则 $\mathbf{W}_i = \mathbf{X}_i$.

步骤 3. 根据定义 6 获得网络中的最优生长点 $G^{optimal}$.

步骤 4. 网络的生长条件及神经元的学习规则: 若一个样本的学习次数大于 σ 或已经学习到一个新样本, 则在最佳生长点 $G^{optimal}$ 的二维外凸空间生长出一个神经元单元 EAN_{new}^{unit} , 新增节点 V_i 的权值为

$$\mathbf{W}_{\text{new}} = \begin{cases} \mathbf{W}_{\text{topo}}, & V_i \text{ 与 } G^{\text{optimal}} \text{ 拓扑相连} \\ \mathbf{W}_c, & V_i \text{ 与 } V_c \text{ 在 } EAN_{\text{new}}^{\text{unit}} \text{ 中成对角} \end{cases}$$

其中,若在最佳生长点 G^{optimal} 的二维外凸空间生长出一个神经元单元 $EAN_{\text{new}}^{\text{unit}}$,使得 t 时刻 G^{optimal} 中处于中间位置的神经元 V_t 在 $t+1$ 时刻成为三神经元的中央神经元 V_c :对 G^{optimal} 中的神经元权值按式 (3) 进行调整.

$$\mathbf{W}_{(t+1)} = \begin{cases} \mathbf{W}_{(t)} + \eta_{\text{max}} \times \Delta \hat{\mathbf{W}}_{it}, & V_t = V_c \\ \mathbf{W}_{(t)} + \eta_{\text{topo1}} \times \Delta \hat{\mathbf{W}}_{it}, & V_t = V_{\text{topo1}} \\ \mathbf{W}_{(t)} + \eta_{\text{topo2}} \times \Delta \hat{\mathbf{W}}_{it}, & V_t = V_{\text{topo2}} \end{cases} \quad (3)$$

其中, t 为时间, η 为学习效率, $V_{\text{topo1}}, V_{\text{topo2}}$ 是与 V_c 直接拓扑相邻的两个神经元, $\eta_{\text{topo1}}, \eta_{\text{topo2}}$ 和 η_{max} 是属于 $[0,1]$ 的常数,且保证 η_{max} 最大. $\Delta \hat{\mathbf{W}}_{it}$ 为时刻 t 神经元与样本 \mathbf{X}_i 距离的归一化向量,由式 (4) 计算:

$$\Delta \hat{\mathbf{W}}_{it} = \frac{\mathbf{X}_i - \mathbf{W}_t}{\|\mathbf{X}_i - \mathbf{W}_t\|} = \left(\frac{\mathbf{X}_i^1 - \mathbf{W}_t^1}{\sqrt{\sum_{j=1}^n (\mathbf{X}_i^j - \mathbf{W}_t^j)^2}}, \dots, \frac{\mathbf{X}_i^k - \mathbf{W}_t^k}{\sqrt{\sum_{j=1}^n (\mathbf{X}_i^j - \mathbf{W}_t^j)^2}} \right) \quad (4)$$

步骤 5. 停止判断. 若在新增的神经元单元中,存在一个神经元单元,满足定义 9,则将此中央神经元设为出发点进行下一个样本学习,直到全部样本输入完毕,否则回到步骤 3.

步骤 6. 神经元的删除. 当全部学习结束后,若存在中央神经元及其拓扑相邻神经元皆为非覆盖神经元,则删除该神经元及其拓扑相邻的非覆盖神经元.

步骤 7. 计算决策域,生成输出神经元. 计算拓扑相邻的神经元对应超球的拓扑乘积,其结果作为这一类型样本的决策域:

$$D_c = \bigcup_{i=1}^n Vol_{EAN_i}$$

其中, Vol_{EAN_i} 代表一类样本 c 中的一个神经元在高维空间中的超球体积. 最后,在网络的第三层增加与决策域数量一致的神经元,使每个神经元对应一个决策域.

2.3 实验结果与分析

实验 1. 使用 Ritter 与 Kohonen 的 16 种动物的实验数据^[26] 进行聚类实验,同 SOFM 和 GCS 进行对比.

实验结果如图 7 所示. 可以看出, ESGSONN 算法经过 53 步学习 ($\eta_{\text{max}} = 0.3, \eta_{\text{max}} = 0.25, \eta_{\text{max}} = 0.24, \xi = 0.25, \sigma = 5$), 聚类的运行结果和 GCS 的运行结果一致 (分为 4 类), 同 SOFM 的运行结果基本一致.

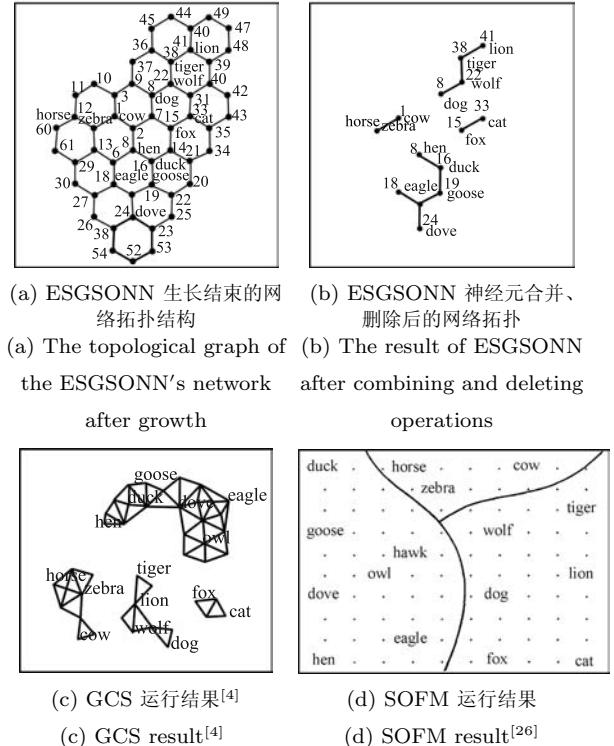


图 7 实验 1 中 ESGSONN 与 SOFM、GCS 的对比结果
Fig. 7 The results of ESGSONN compared with SOFM and GCS in Experiment 1

实验 2. 使用两组具有特定分布的数据,来验证 ESGSONN 算法对处理死神经元、边缘效应及复杂结构数据的能力. 实验中 SOFM 采用 20×20 的规模, $\alpha = 0.05$, iterations = 10 000, Gaussian 邻域函数. ESGSONN 中参数 $\eta_{\text{max}} = 0.3, \eta_{\text{max}} = 0.25, \eta_{\text{max}} = 0.24, \xi = 0.25, \sigma = 5$. 第一组数据共 3 243 个样本,每个样本维数为 2,呈圆环分布,如图 8(a) 所示.

采用 SOFM 和 ESGSONN 算法进行对比,结果如图 9(a) 和图 9(b) 所示. 可以看出, SOFM 在这种情况下,有很多神经元的权值处于数据分布以外的区域,并集中在圆环的几何中心和每一小组数据的缝隙位置,这些神经元很难被激活,成为死神经元. 同时在数据分布的边缘很少有神经元能够覆盖到,通常情况下, SOFM 需要更多的迭代才能消除这种边缘现象. 而 ESGSONN 算法则可以很好地反映原数据的密度分布. 由于按照单位长度更新权值,则不会在圆环的几何中心和每小组数据的缝隙中生长神经元,从而尽量避免了死神经元的产生.

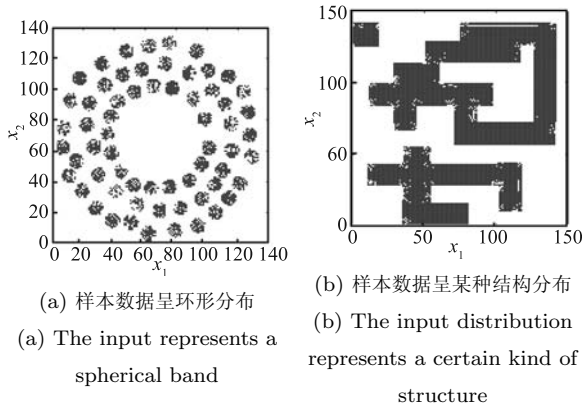


图 8 实验 2 中的样本数据分布

Fig. 8 Distributions of the samples in Experiment 2

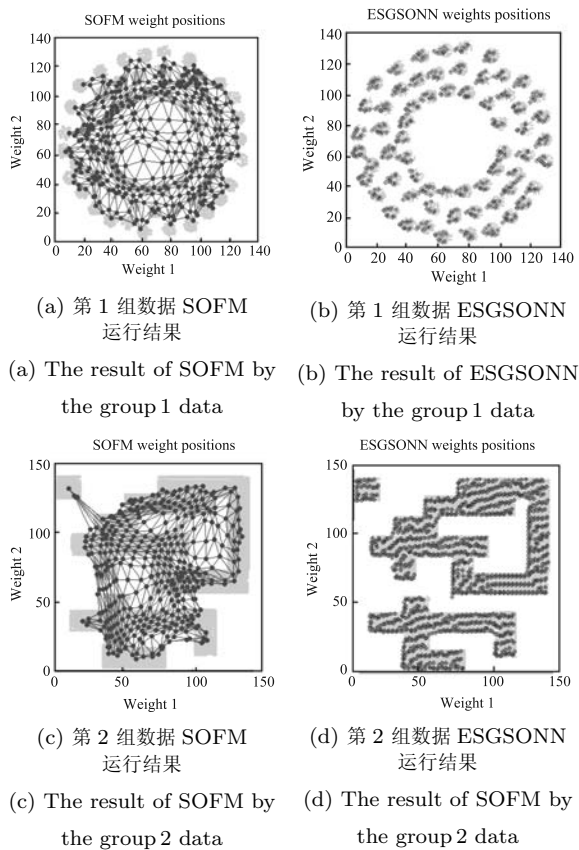


图 9 实验 2 中 ESGSONN 与 SOFM 的对比结果
Fig. 9 The results of ESGSONN compared with SOFM in Experiment 2

SOFM in Experiment 2

第 2 组数据使用 9 173 个样本, 每个样本维数为 2, 用来测试 ESGSONN 算法对复杂结构的数据的稳定性. 输入数据的分布如图 8 (b) 所示. SOFM 和 ESGSONN 运行结果分别如图 9 (c) 和图 9 (d) 所示. 可以看出, SOFM 出现了结构上的不稳定, 随着学习邻域半径的不断缩小, 网络的震荡逐渐停止, 而在这个过程中非 0 密度区的神经元不断对 0 密度区域的神经元产生影响, 从而导致训练结束后, 存在大量

的神经元处于 0 密度区域. 而 ESGSONN 则较好地反映了样本数据的分布.

3 ESGSONN 在高维几何空间的本质分析

实际上, 表征样本特征的模式可以用高维欧氏几何空间中的一个点来表示, 可以假设同类样本间的相似度远远高于不同样本之间的相似度, 反映在空间分布中为同类样本分布密度较大, 而不同样本之间会有明显的 0 密度空间. 因而, 如果能找到每一类样本在高维空间中的分布区域, 并用一个几何形状封闭起来, 则可以达到聚类的目的. 这样, 空间聚类问题可以转化为高维空间中寻找每一类样本的分布并用几何形状加以封闭的问题.

在 ESGSONN 中, 神经元的权值和样本具有同样的维度, 网络中任意一个 EAN 可以看作是在高维空间中以权值 W_i 为球心、参数 ξ 为半径的超球, 这样就在高维空间中获得了一个与该超球体积大小相当的封闭几何区域. 在学习过程中, ESGSONN 生成、删除超球及调整这些超球在高维空间中的位置, 使全部的样本都位于超球内部, 最终超球的拓扑乘积构成封闭的几何形状. 神经元的权值代表相应超球的球心, 实际上是一小部分样本的聚类中心.

在对样本的学习过程中, 一共存在着两种学习的方式, 即: 对不同种类样本的学习的过程和对同类样本的学习过程. 对于不同种类样本的学习过程, 用“过程的神经元”表示每一个学习的步骤, “过程的神经元”的产生过程代表了一类样本的模式到另一类样本的模式演变过程; 对于同类样本的学习过程, 用“过渡的神经元”和中央神经元共同表示, 这两种神经元对应超球的拓扑乘积即为当前类别的封闭几何形状. 引用“过渡的神经元”是为了填补同一类样本之间在高维几何空间中分布的“缝隙”, 增加网络的泛化能力. 在学习结束后, “过程的神经元”会被删除, 以便形成不同的类, “过渡的神经元”得以保留. 学习过程中不用刻意区分这两种类型的神经元, 实际上, 这两种类型的神经元都可以用 EAN 模型来表示, 只是与周围神经元所组成的网络的拓扑结构不同.

4 结论

人工神经元是神经计算科学的基础, 本文结合生物学中神经科学的发展, 将神经胶质细胞的能量模型引入到人工神经元的概念模型中, 提出了能量人工神经元模型, 并给出了数学表述. 此外, 在能量人工神经元模型的基础上, 实现了一种新型自生长、自组织人工神经网络 — ESGSONN, 通过实验表明, ESGSONN 具有如下的特性:

- 1) ESGSONN 通过学习、生长、合并、删除等

操作, 实现了一个高维输入空间 \mathbf{R}^n 到一个二维空间的拓扑结构 \mathbf{H} 的映射. \mathbf{H} 中神经元的权值密度分布与样本空间 \mathbf{R}^n 中的样本具有类似的密度分布. 同时, 两个神经元在二维拓扑中距离越近, 代表对应的样本相似度越高.

2) ESGSONN 的算法简单, 网络生长以构建新的神经元单元为目标, 而且学习过程只对可生长点中的神经元权值进行调整, 每次权值调整的神经元数量非常少. 学习效率是常数, 不需要样本的反复迭代, 全部样本一次输入后立即完成对网络的训练.

3) 训练后的网络很少出现死神经元及边缘效应.

References

- 1 Willshaw D J, Von D M C. How patterned neural connections can be set up by self-organization. *Proceedings of the Royal Society of London, Series B, Biological Sciences*, 1976, **194**(1117): 431–445
- 2 Kohonen T. Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 1982, **43**(1): 59–69
- 3 Kangas J A, Kohonen T K, Laaksonen J T. Variants of self-organizing maps. *IEEE Transactions on Neural Networks*, 1990, **1**(1): 93–99
- 4 Fritzke B. Growing cell structures — a self-organizing network for unsupervised and supervised learning. *Neural Networks*, 1994, **7**(9): 1441–1460
- 5 Fritzke B. Growing grid — a self-organizing network with constant neighborhood range and adaption strength. *Neural Processing Letters*, 1995, **2**(5): 9–13
- 6 Fritzke B. A growing neural gas network learns topologies. *Advances in Neural Information Processing System*. Cambridge: The MIT Press, 1995. 625–632
- 7 Choi D I, Park S H. Self-creating and organizing neural networks. *IEEE Transactions on Neural Networks*, 1994, **5**(4): 561–575
- 8 Martinetz T M, Schulten K J. A “neural-gas” network learns topologies. *Artificial Neural Networks*, 1991, **1**: 397–402
- 9 Blackmore J, Mikkulainen R. Incremental grid growing: encoding high-dimensional structure into a two-dimensional feature map. In: *Proceedings of the IEEE International Conference on Neural Networks*. San Francisco, USA: IEEE, 1993. 450–455
- 10 Chow T W S, Wu S. Cell-splitting grid: a self-creating and self-organizing neural network. *Neurocomputing*, 2004, **57**: 373–387
- 11 Alahakoon D, Halgamuge S K, Srinivasan B. Dynamic self-organizing maps with controlled growth for knowledge discovery. *IEEE Transactions on Neural Networks*, 2000, **11**(3): 601–614
- 12 Shen F. An Algorithm for Incremental Unsupervised Learning and Topology Representation [Ph.D. dissertation], Tokyo Institute of Technology, Japan, 2006
- 13 Shen F, Hasegawa O. An incremental network for on-line unsupervised classification and topology learning. *Neural Networks*, 2006, **19**(1): 90–106
- 14 Shen F, Hasegawa O. An adaptive incremental LBG for vector quantization. *Neural Networks*, 2006, **19**(5): 694–704
- 15 Shen F, Oqura T, Hasegawa O. An enhanced self-organizing incremental neural network for online unsupervised learning. *Neural Networks*, 2007, **20**(8): 893–903
- 16 Pfrieger F W, Barres B A. Synaptic efficacy enhanced by glial cells in vitro. *Science*, 1997, **277**(5332): 1684–1687
- 17 Ullian E M, Sapperstein S K, Christopherson K S, Barres B A. Control of synapse number by glia. *Science*, 2001, **291**(5504): 657–661
- 18 Song H, Stevens C F, Gage F H. Astroglia induce neurogenesis from adult neural stem cells. *Nature*, 2002, **417**(6884): 39–44
- 19 Slezak M, Pfrieger F W, Soltys Z. Synaptic plasticity, astrocytes and morphological homeostasis. *Journal of Physiology-Paris*, 2006, **99**(2–3): 84–91
- 20 Deitmer J W. Strategies for metabolic exchange between glial cells and neurons. *Respiration Physiology*, 2001, **129**(1–2): 71–81
- 21 Deitmer J W. Glial strategy for metabolic shuttling and neuronal function. *BioEssays*, 2000, **22**(8): 747–752
- 22 Newman E A. Glial cell inhibition of neurons by release of ATP. *The Journal of Neuroscience*, 2003, **23**(5): 1659–1666
- 23 Deitmer J W, Brockhaus J, Casel D. Modulation of synaptic activity in Purkinje neurons by ATP. *The Cerebellum*, 2006, **5**(1): 49–54
- 24 Bezzi P, Volterra A. A neuron-glia signaling network in the active brain. *Current Opinion in Neurobiology*, 2001, **11**(3): 387–394
- 25 McCulloch W S, Pitts W. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biology*, 1943, **5**(4): 115–133
- 26 Ritter H, Kohonen T. Self-organizing semantic maps. *Biological Cybernetics*, 1989, **61**(4): 241–254



班晓娟 北京科技大学教授. 主要研究方向为人工智能, 人工生命和计算机动画. E-mail: banxj@ustb.edu.cn
(**BAN Xiao-Juan** Professor at University of Science and Technology Beijing. Her research interest covers artificial intelligence, artificial life, and computer animation.)



刘浩 北京科技大学计算机科学与技术系硕士研究生. 主要研究方向为人工智能和神经网络. 本文通信作者. E-mail: liuhao520@gmail.com
(**LIU Hao** Master student at the Department of Computer Science and Technology, University of Science and Technology Beijing. His research interest covers artificial intelligence and neural network. Corresponding author of this paper.)



徐卓然 北京科技大学计算机科学与技术系硕士研究生. 主要研究方向为人工智能和神经网络. E-mail: xuzhuoran0106@gmail.com
(**XU Zhuo-Ran** Master student at the Department of Computer Science and Technology, University of Science and Technology Beijing. His research interest covers artificial intelligence and neural network.)