

# 一种改进的 ART 型神经网络学习算法

杨 兴<sup>1</sup>, 朱大奇<sup>1</sup>, 桑庆兵<sup>2</sup>

(1. 江南大学 控制科学与工程研究中心, 江苏 无锡 214122;

2. 江南大学 信息工程学院, 江苏 无锡 214122)

**摘 要:** 指出常规 ART1 型神经网络的不足, 提出了一种改进的相似度计算方法。它同时考虑两向量对应位子值, 避免了 ART1 网络中两个向量由于输入顺序的不同而得到不同的相似度的结果; 针对 ART1 网络模式识别的漂移问题, 提出了少数服从多数原则来减少这种问题的出现。改进了 ART1 型神经网络的应用效果。

**关键词:** 神经网络; 自适应共振理论; 模式分类; 模式漂移

中图分类号: TP183

文献标识码: A

文章编号: 1673-629X(2006)09-0027-03

## An Improved ART Neural Network Learning Algorithm

YANG Xing<sup>1</sup>, ZHU Da-qi<sup>1</sup>, SANG Qing-bing<sup>2</sup>

(1. Research Centre of Control Science and Engineering, Southern Yangtze University, Wuxi 214122, China;

2. School of Information Engineering, Southern Yangtze University, Wuxi 214122, China)

**Abstract:** The shortage of traditional ART1 neural network is indicated, and an improved calculating method of comparability is presented. The corresponding place value of two vectors is considered in this method. The method avoids the different result of ART1 neural network because of inputting different sequence. In order to solve pattern-drafting problem of ART1 neural network, the principle of minority subordinate to majority is proposed to reduce the appeared problem. They improve the applicative effect of ART1 neural network.

**Key words:** neural network; adaptive resonance theory; pattern recognition; pattern drafting

## 0 引 言

自适应共振理论(Adaptive Resonance Theory, 简称 ART)<sup>[1]</sup>是美国波士顿大学的 S. Grossberg 和 G. A. Carpenter 于 1976 年提出的一种神经网络模型。ART 是用生物神经细胞自兴奋与侧抑制的动力学原理指导学习, 让输入模式通过网络双向连接权的识别和比较达到共振来完成自身的记忆, 并以同样的方式实现联想。ART 网络是一种无教师学习的反馈人工神经网络, 可以在线学习, 适用于复杂系统的分类。ART 网络分为 3 种类型: ART1, ART2 和 ART3。其中 ART1 型网络应用广泛, 是一个闭环的自适应系统, 具有自组织、自稳定和自学习能力, 表现了良好的自适应能力和可塑性<sup>[2,3]</sup>。

## 1 ART1 网络的工作原理

ART1 型神经网络<sup>[4,5]</sup>由输入层、比较层和识别层构成。网络运行时接受来自外部环境的输入模式, 并检测输

入模式与识别层所有模式之间的匹配程度。对于相似程度最高的模式类, 网络要继续考察该模式的典型向量与当前输入模式的相似度。如果大于相似度阈值, 就归为该模式类。如果不大于相似度阈值, 就在剩下的节点中再次匹配相似。如果已有节点都不能匹配相似, 就自适应地增加一种新的模式来对应此次输入的模式。

①初始化内外星权向量的权值: 设  $t_{ij}$  表示外星权向量, 也即识别层第  $j$  个节点到比较层第  $i$  个节点的向量, 令  $t_{ij}(0) = 1$ ; 设  $b_{ij}$  表示内星权向量, 也即比较层第  $i$  个节点到识别层第  $j$  个节点的向量, 令  $b_{ij}(0) = 1/(1+n)$ 。其中,  $i = 1, 2, \dots, n; j = 1, 2, \dots, m$ 。

②输入外部信号  $X$ :  $X = \{x_i\}^n = (x_1, x_2, \dots, x_n)$ ,  $i = 1, 2, \dots, n$ , 其中  $x_i$  取 0 或 1。

③计算最大匹配度  $\text{net}_j^*$ :  $\text{net}_j = \sum_{i=1}^n b_{ij}x_i$ ,  $j = 1, 2, \dots, m$ ,  $\text{net}_j^* = \max\{\text{net}_j\}$ , 其中  $j^*$  为获胜节点。

④相似度的检测:  $N_0 = \sum_{i=1}^n t_{ij^*}x_i$ ,  $N_1 = \sum_{i=1}^n x_i$ ,  $i = 1, 2, \dots, n$ ,  $j = 1, 2, \dots, m$ 。

⑤相似度的比较: 设  $\rho$  为相似度阈值。如果  $N_0/N_1 \geq \rho$ , 转到 ⑦; 如果  $N_0/N_1 < \rho$ , 且所有节点都匹配过, 转到 ⑥; 如果  $N_0/N_1 < \rho$ , 所有节点未全部匹配过, 转到 ③, 寻找下一个最大匹配度节点。

收稿日期: 2005-12-19

基金项目: 江苏省自然科学基金资助项目(BK2004021); 教育部科学技术研究重点项目(105088)

作者简介: 杨 兴(1981-), 男, 江苏镇江人, 硕士研究生, 研究方向为故障诊断; 朱大奇, 博士, 教授, 研究方向为智能故障诊断与容错控制、人工神经网络及应用。

⑥ 增加一个节点, 即此时有  $m+1$  个输出节点。  $t_{i, m+1}$  与  $b_{i, m+1}$  的权值为初始化的值,  $i = 1, 2, \dots, n$ 。

⑦ 权值调整:

$$t_{ij}*(t+1) = t_{ij}*(t)x_i$$
$$b_{ij}*(t+1) = \frac{t_{ij}*(t+1)}{0.5 + \sum_{i=1}^n t_{ij}*(t+1)} \quad (1)$$

$i = 1, 2, \dots, n$

⑧ 接受新的输入, 转到②。

## 2 ART1 型神经网络存在的问题与改进

### 2.1 常规相似度判定的不足与改进

常规 ART1 的相似度是比较获胜节点  $j^*$  外星权向量  $T_j^*$  与输入向量  $X$  对应位子中为“1”的个数, 其表达如下:

$$\rho_1 = \frac{N_0}{N_1} = \frac{\sum_{i=1}^n t_{ij^*} x_i}{\sum_{i=0}^n x_i}, i = 1, 2, \dots, n \quad (2)$$

这种常规的相似度的判据并未考虑外星权向量中“0”的作用, 只考虑“1”的作用。在某些只有“1”表示的状态需要考虑时, 这种方法是可行的。但是在实际运用中, 0, 1 分别表示两种状态, 在判断中都是有用的信息, 是同等重要的。因此, 这种方法有明显的不足之处。这里以英文字母的识别的例子来进行比较。

例 1: 图 1 是英文字母的模式图。其中, (a) 为标准“E”模式信号, (b)、(c)、(d) 为“E”的噪声信号, (e) 为标准的“F”模式信号。计算相似度时, 以涂黑的部分为“1”, 空白的部分为“0”。对应 (a)、(b)、(c)、(d) 和 (e) 图信号进行编码, 分别用符号  $X_1, X_2, X_3, X_4$  和  $X_5$  表示。各信号的编码如下:

$X_1 = 11111100001000011110100001000011111,$   
 $X_2 = 11111100001000011110100001000011011,$   
 $X_3 = 11111100001000011110100001000010111,$   
 $X_4 = 11111100001000011110100001000011110,$   
 $X_5 = 11111100001000011110100001000010000.$

网络第一次输入信号(a) 时( $X = X_1$ ), 信号(a) 与第一个节点匹配相似, 外星权向量的权值进行调整, 由式(1) 计算第一个节点调整的内外星权向量。

第二次输入信号(b) 时( $X = X_2$ ), 计算  $X_2$  与竞争层各个节点的匹配相似度, 可知在第一个节点处最大(大于其它未使用的节点), 按公式(2), 它的相似度公式得到的值  $\rho_1$  为 1.000, 这说明(b) 与(a) 是完全相同的信号, 从图中可以清楚看到(b) 与(a) 是有差异的。而将(a)、(b) 的输入次序调换一下, 根据公式(2) 得到的相似度  $\rho_1$  为  $17/18 = 0.944$ 。就(a) 与(b) 而言, 在不同的输入次序下得到不同的相似度。这是因为在比较相似度的时候, 只要  $t_{ij^*}$  为 0, 在其对应的位置上  $x_i$  也为 0, 此时不论  $t_{ij^*}$  为 1 的对应位置上, 输入  $x_i$  为 1 还是为 0, 所得相似度一定为 1。

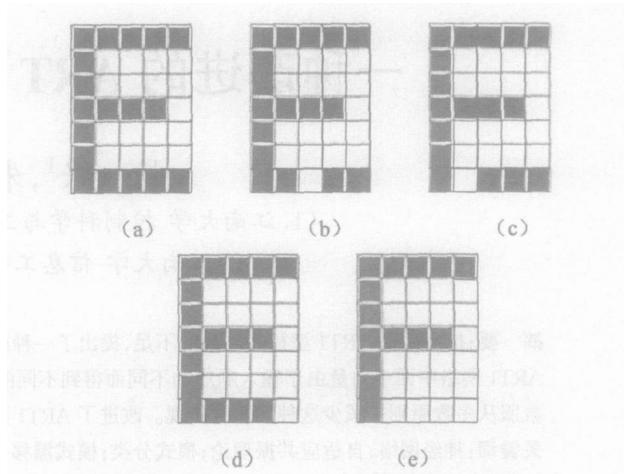


图 1 英文字母的模式图

为进一步说明其不足, 此处给出另一实例。

例 2: 取  $n = 10$ , 外星权的权值为  $T_j^* = 1111111110$ , 而输入  $X = 1000000000$ , 根据公式(2) 得到的相似度  $\rho_1$  为“1”, 但这两个矢量明显不是同类, 这就给实际的模式分类带来错误判断。

为解决这种不足, 文中设计了一种简单快捷且准确的相似度计算方法, 此处所用的思想就是同时考虑“0”和“1”的状态, 看  $t_{ij^*}$  和  $x_i$  对应位置上相同状态的个数, 即  $t_{ij^*}$  与  $x_i$  “同或”值的个数与输入矢量维数  $n$  的比值。此相似度的计算方法如公式(3)。这种方法能够很准确地比较的两个向量, 而不存在孰轻孰重的问题。

$$\rho_2 = \frac{\sum_{i=1}^n t_{ij^*} x_i + \sum_{i=1}^n (1 - t_{ij^*})(1 - x_i)}{n} \quad (3)$$

$i = 1, 2, \dots, n$

用这种方法时, 网络第一次输入信号(a) 的结果与常规 ART1 型网络一样。在第二次输入信号(b) 时( $X = X_2$ ), 它匹配度仍是在第一节节点处最大(大于其它未使用的节点), 根据公式(3) 得它的相似度的值  $\rho_2$  为:  $\rho_2 = 34/35 = 0.971$ , 说明(b) 与(a) 是有差异的。

对例 2 来说, 外星权向量  $T_j^* = 1111111110$  与输入向量  $X = 1000000000$  明显不相似, 按(3) 式可得相似度  $\rho_2 = (1+1)/10 = 0.2$ , 而按公式(2) 得其相似度却为“1”。

这种改进的相似度判据是针对输入模式与已有匹配模式进行的。由于网络初始化时的  $t_{ij}$  都为 1, 对于未使用的节点在第一次使用时, 直接将它的外星权向量调整为输入的向量, 即  $t_{ij^*} = x_i, i = 1, 2, \dots, n$ 。

### 2.2 常规外星权向量调整的不足与改进

常规 ART1 型神经网络的外星权向量调整如式(1) 所示, 是外星权  $t_{ij^*}$  和输入  $x_i$  的“与”。这在网络模式识别时容易产生渐进模式漂移<sup>[9]</sup>。即对某一获胜节点来说, 其外星权向量随着输入模式的逐渐变化, 越来越远离最初的模式矢量。这种情形是输入的模式与某个节点匹配相似, 并且输入矢量与原有模式有很小的变化, 这样每次这个节

点的权值都会被改变,这使得识别层每个节点代表的模式是飘忽不定的,有时就会漂移到别的模式中去,使网络在模式分类中出现错误。

ART1 型神经网络明显的特点就是只有 1 和 0 两种状态,要想从根本上消除渐进模式漂移问题是十分困难的,只有削弱这种情况的出现。常规的 ART1 型神经网络是一旦输入的模式与已有节点匹配相似就调整该节点外星权向量的权值,无论是朝我们期望的方向还是背离的期望。针对此种情况,在外星权向量调整的时间上有所改变,即不是输入模式与某个节点匹配相似后就调整该外星权的权值。所采取的方法就是对某个节点有多组输入模式与其匹配相似时,把这几组模式对应的信息存储起来,等到达到一定数量时,综合这几种外星权向量的知识,对其进行调整。例如,取 3 组向量来进行综合知识的调整,即采用 2/3 原则。比较每组向量对应位子的 3 个数,取相同数多的作为外星权向量调整后的值。当 3 个数都为“1”或“0”时,调整后的数就为“1”或“0”;当两个为“1”,一个为“0”时,调整后的值还为“1”;当两个为“0”,一个为“1”时,调整的值就为“0”了。

2.3 结果讨论

为了更好地提高 ART1 神经网络的识别性能,此处进一步讨论将改进相似度方法与外星权向量调整算法相结合的 ART1 网络识别效果。在上述的英文字母识别的例子中,取相似度的阈值  $\rho_0$  为 0.9,图 1 中(e) 与(a) 的相似度按公式(2) 得  $\rho_1 = 14/14 = 1 > \rho_0$ ;按式(3) 得相似度  $\rho_2 = 31/35 = 0.886 < \rho_0$ ,通过改进的 ART1 型相似度判据可以将两组模式区分开来。但是,如果将信号按(a) (b) (c) (d) 的顺序依次输入,每次在匹配相似后进行内外星权权值调整,再输入(e) 信号后,系统会将(e) 信号归为标准“E”模式。这是因为在权值调整中,表征模式特征的外星权向量每次都朝不期望的方向移动,当这种移动达到一定程度时,就漂移到了别的模式中去了。如果用改进相似度判据的方法,同样有漂移问题。

① 常规外星权向量权值调整的结果。

如表 1 所示,将信号按(a) (b) (c) (d) (e) 的顺序依次输入。相似度阈值  $\rho_0$  取 0.9。当第一次输入信号(a) 时,与“空”模式的匹配相似度为“1”。第二次输入信号(b) 时,由于信号(b) 与信号(a) 相比,只是在(a) 即  $X_1$  为“1”的位子上有一个变为“0”,用常规的方法得到的相似度  $\rho_1$  一定为 1,与第一个节点匹配,此时对第一个节点外星权向量的权值进行调整。同样第三次输入信号(c) 时,常规的方法得到(c) 的相似度  $\rho_1$  就不为 1 了,而是 0.941,再经过外星权向量的权值调整;在输入信号(d) 得到常规的相似度  $\rho_1$  为 0.880,小于相似度阈值  $\rho_0$ ,被归为另一类了。而在输入信号(e) 时,它又被归为标准“E”类了,且相似度  $\rho_1$  为 1,完全相似。显然所得结果与实际情况不符。

用改进相似度的计算方法,可以将(d) 归为标准“E”类,但模式漂移仍存在,信号(e) 仍然被归为标准“E”类。

表 1 常规与改进外星权权值调整结果的比较

相似度	常规外星权向量权值调整的结果		改进外星权向量权值调整的结果	
	$\rho_1$	$\rho_2$	$\rho_1$	$\rho_2$
(a)	1.000	1.000	1.000	1.000
(b)	17/17 = 1.000	34/35 = 0.971	17/17 = 1.000	34/35 = 0.971
(c)	16/17 = 0.941	33/35 = 0.943	17/17 = 1.000	34/35 = 0.971
(d)	15/17 = 0.880	32/35 = 0.914	17/17 = 1.000	34/35 = 0.971
(e)	14/14 = 1.000	35/35 = 1.000	14/14 = 1.000	31/35 = 0.886

② 改进外星权向量权值调整的结果。

改进外星权向量权值调整的方法就是少数服从多数的原则。在常规的外星权向量权值调整中,  $t_{ij}^*$  只能从“1”变为“0”或是不变;而采用少数服从多数原则后,  $t_{ij}^*$  可从“1”变为“0”,或者从“0”变为“1”。在用改进相似度计算方法的基础上,用改进外星权向量权值调整的方法,解决了上述模式漂移问题。如表 1 所示,所得到的  $\rho_2$  的值,明显将(b) (c) (d) 信号归为标准“E”类,将(e) 信号归为另一类。

对于某个已建立的模式,用改进的外星权向量权值调整方法在以后的学习过程中,是有选择地对网络进行学习,只有那些经常出现的重要的信息才被网络学习进去,那些经常变化的不确定的信息将不被网络学习。这种方法并未从本质中消除渐进模式漂移,但大大削减了漂移问题出现的情况。以 2/3 原则为例,当输入上述例子时是不会出现模式漂移的,但每组向量依次输入,至少连续输入两次后才出现上述渐进模式漂移。如果在实际运用中出现这种情况,可用更高的原则,如 3/5 原则、4/7 原则等。

3 结 论

文中提出的改进的 ART1 型神经网络是在实际运用中出现的问题。提出了两个向量“同或”的相似度的判据方法,很好解决了在输入次序先后所产生的问题。但是仅仅对相似度的改进,并未解决网络存在的模式漂移问题,基于这种情况提出了少数服从多数的原则来改进权值调整方式。模式漂移的问题虽然没有能够完全消除,但改进的方法十分有效,大大削弱了这种漂移问题的出现。

参考文献:

[1] Carpenter G A, Grossberg S. A Massively Parallel Architecture for a Self-Organizing Neural Pattern Recognition Machine[J]. Trans. IEEE on Computer Vision, Graphics and Image Processing, 1987, 37(1): 54-115.

[2] Grossberg S. Competitive Learning, From Interactive Activation to Adaptive Resonance[J]. Cognitive Science, 1987, 11(1): 23-63.

[3] 全朝阳,石教英. 一种面向模式分类的修正的 ART1 型神经网络[J]. 计算机学报, 1995, 18(9): 671-677.

[4] 朱大奇. 电子设备故障诊断原理与实践[M]. 北京: 电子工业出版社, 2004. 145-150.

[5] 韩力群. 人工神经网络的理论、设计及应用[M]. 北京: 化学工业出版社, 2002. 81-87.

[6] Laura I, Burke. Clustering Characterization of Adaptive Resonance[J]. Neural Networks, 1991, 16(4): 485-491.