维普资讯 http://www.cqvip.com

FUZZY ARTMAP 神经网络综述*)

A Survey of Fuzzy Adaptive Resonance Theory Map

钟金宏　杨善林

(合肥工业大学计算机网络系统研究所　合肥230009)

Abstract　Adaptive resonance theory offers a network solution of the stability-plasticity dilemma that exists in conventional artificial neural networks by simulating perception and cognition of humans. Fuzzy adaptive resonance theory map synthesizes fuzzy set theory and adaptive resonance theory, holds a number of desirable properties. This paper surveys its development course, principle, application area and the newest results, forecasts the directions of the future development.

Keywords　Neural networks, Adaptive resonance theory, Fuzzy adaptive resonance theory map

1. 引言

神经网络模拟人脑神经功能，用大量简单关系连接来表示复杂的函数关系，具有很多特殊优点，已在系统建模、模式识别、图像处理、知识发现和控制等领域得到了广泛的应用。但当使用前向神经网络，如多层感知器(MLP)和径向基函数(RBF)网络时，需要依赖一些试探法去选择最优的网络尺寸和参数[1]，此外这些网络在学习后通常是静止的，不具备增量学习能力(在线学习)、对新模式学习时，会破坏网络已记忆的模式[2,3]。

相反地，自适应谐振理论(ART)[4]神经网络家族克服了所谓的可塑性-稳定性二难问题，能够在变化的环境中学习新信息，而不会腐蚀以前已学习的信息。目前已有一系列ART网络的模型和结构[4~9]，其中模糊自适应谐振理论网络(FUZZY ARTMAP)[9]摒弃了ARTMAP的集合观点，采用隶属函数的方法表示输入特性，从而扩展了对模糊属性的学习能力。FUZZY ARTMAP神经网络在国外有许多研究报道，本文从发展历程、基本原理、主要应用领域和当前研究进展等四个方面对其进行了综述。

2. 发展历程

在ART家族中，最早期的自适应谐振理论只适用于二值输入模式，称为ART-1网络[4]，为解决这个问题，S. Grossberg 和 G. Gartpenter 又提出了一个适合处理模拟值输入模式的自适应谐振理论，称为ART-2网络[5]，两者的基本原理是相同的，学习过程都由一个警戒参数控制，通过选择适当的警戒参数就可以生成粗细适当的分类；且都具有增量学习能力，当给网络增加新模式时不用重构原网络。区别是ART-2既可处理二值输入模式，又能处理模拟值输入模式，故ART-2的适用性比ART-1更广泛，ART-3[6]是Cohen和Grossberg等人在做了ART的同形级联结构体系研究的基础上发展出来的，它允许任意数量结构相同的ART-2模型级联，通过一个嵌入网络层间的化学突触模型(发射器)解决了ART系统的记忆搜索问题。FUZZY ART[7]是模糊集理论和ART相结合的产物，用模糊集运算取代ART中的集合运算，使网络具有更强大的性能，不但能处理二值和模拟值输入模式，而且能表示和处理模糊信息。以上都属于无导师学习，存在着诸如分类精度固定等多方面问题。ARTMAP[8]算法通过实现有导师增量学习解决了这些问题，该算法通过一个映射域将两个ART模型连接起来，映射域在对ART的分类结果进行连接处理的同时，还执行"匹配跟踪"规则，从而使匹配过程中出现的错误不重复出现。但ARTMAP是由两个ART1模型和映射域构成，因此只能处理二值输入模式，这限制了它的使用。FUZZY ARTMAP[9]算法是在ARTMAP算法基础上，用FUZZY ART模型取代ARTMAP中的ART-1模型，从而能处理二进制值和连续值，并扩展了对模糊属性的学习，使算法具有更高的性能和更广泛的适用性。

*)国家自然科学基金资助项目(编号　79970058)，安徽省自然科学基金资助项目(编号　99043645)。钟金宏　博士生，主要研究方向为模式识别和人工智能。杨善林　教授，博导，主要研究方向为模式识别、人工智能、知识发现和决策支持系统等。

维普资讯 http://www.cqvip.com

3. FUZZY ARTMAP 基本原理

FUZZY ARTMAP 神经网络是监督 ARTMAP 网络的一个变种，将模糊集理论并入了 ARTMAP 动力系统，从而可实现任意维模拟输入、输出空间间的任意映射。其通过匹配跟踪过程实现了一种新的最小最大学习规则，即最小化预测误差，最大化编码压缩率（泛化）。FUZZY ARTMAP 是由一个内部映射域 F^{ab} 和两个 FUZZY ART 模型（ART_a 和 ART_b）构成（图1）。每个 ART 模型为三层结构，F_0层的节点代表现在的输入向量，F_1层接受来自 F_0的下一上的输入和来自 F_2层的上一下的输入，F_2层的节点代表激活的编码（类别），F_0和 F_1层的节点数与输入有关，F_2层的节点数是变化的，与每个 F_2层类别节点相连的是权矢量，称为自适应权重或 LTM 跟踪，包容了 ART-1网络的下一上和上一下权矢量。

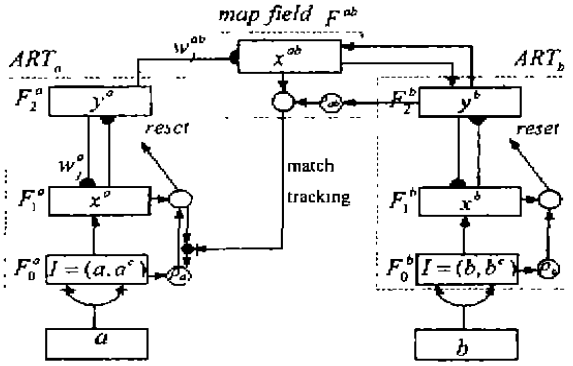

图1 FUZZY ARTMAP 神经网络结构图

在监督学习时，ART_a 接受输入模式，进行自组织学习，最终形成输入向量的自动分类，ART_b 接受目标模式，进行非监督学习，输出导师信号。在 ART_a 和 ART_b 中，为避免类别增生问题，在传给 F_2层前输入模式 a 在 F_1层进行补码编码，也就是 $I=(a,a^c)=(a,1-a)$，F_2层是一个胜者为王（winner－take－all）的竞争层，一个类别选择函数被用来衡量输入向量 I 与 LTM 跟踪（第 j 个节点的 w_j）间的匹配程度：

$$T_j=\frac{|I\wedge w_j|}{\alpha+|w_j|} \tag{1}$$

这里 α 为选择参数，一般取 $\alpha\approx 0$。模糊与操作[10]定义为：$(p\wedge q)\equiv\min(p_i,q_i)$，模 $|\cdot|$ 的定义为：$|p|\equiv\sum_{i=1}^{M}|p_i|$，式(1)反映了权向量 w_j 是输入向量 I 的模糊子集的程度，特殊地，$|I\wedge w_j|/|w_j|=1$表示 w_j 是 I 的模糊子集。获胜神经元由(2)式确定。在有多个 T_j 最大时，选择具有最小索引的节点。

$$T_J=\max\{T_j:j=1,\cdots,N\} \tag{2}$$

当 F_2层节点未被激活时，F_1层的激活向量为 $x=I$，否则，$x=I\wedge w_J$。获胜神经元接着送它的权向量（原形模式）到 F_1层。一个警戒测试被用来检查原形模式和输入模式间的相似性：

$$\frac{|I\wedge w_J|}{|I|}\geqslant\rho \tag{3}$$

这里 ρ 为警戒参数，取值在[0,1]间。如果(3)式满足，表示 x 和 I 的差别不超过警戒参数，获胜神经元正确地代表了输入 I 的类别，共振发生。与获胜神经元相关的权值按(4)式进行自适应，使其与输入模式更相似。

$$w_J^{(new)}=\beta(I\wedge w_J^{(old)})+(1-\beta)w_J^{(old)} \tag{4}$$

这里 β 为学习率，$\beta=1$为快速学习，$\beta<1$为慢速学习。当不满足(3)式时，即 $|x|=|I\wedge w_J|<\rho|I|$时，表示 x 和 I 的差别超过警戒参数，获胜神经元不能代表输入 I 的恰当分类，获胜节点被禁止，输入模式被重新传给 F_2层去搜索满足(3)式的新节点，如无这样的节点存在，则增加一新节点，

映射域 F^{ab}的作用是在 ART_a 的类别与 ART_b 中导师信号间建立联系，执行匹配跟踪规则。映射域的输出向量 x^{ab}遵循(5)式。模式对通过它们的压缩代表（类别节点）联系，关于内部模型（F_2层间的连接）的 LTM 信息储存在映射域的权矩阵中，其分配一权向量给每一已与 ART_b 节点连接的 ART_a 节点。映射域学习规则为：初始时，$w_{jk}^{ab}(0)=1$，在共振期间第 J 个 ART_a 节点被激活时，$w_{jk}^{ab}=x^{ab}$。在快速学习时，一旦连接建立将是永久的。

$$x^{ab}=\begin{cases} y^b\wedge w_j^{ab} & \text{第 } J \text{ 个 } F_2^a \text{节点和 } F_2^b \text{被激活时，}\\ w_j^{ab} & \text{第 } J \text{ 个 } F_2^a \text{节点被激活和 } F_2^b \text{未被激活时，}\\ y^b & F_2^a \text{未被激活而 } F_2^b \text{被激活时，}\\ 0 & F_2^a \text{和 } F_2^b \text{都未被激活时}\end{cases} \tag{5}$$

匹配跟踪：当映射域出现不匹配时，其内部的控制机制使 ART_a 的警戒值增加更正预测误差所需的最小值，即牺牲最小量的泛化能力（信息压缩率）去更正由原来较高的信息压缩率所造成的预测误差。映射域匹配函数 $|y^b\wedge w_J^{ab}|/|y^b|$不满足(7)式时，表示映射域输出向量与 ART_b 输出向量差别超过警戒参数，匹配跟踪发生，增加 ART_a 的警戒值使 $|x^a|=|I\wedge w_J^a|<\rho_a|I|$，这样导致 ART_a 重新搜索，寻找另一满足(6)和(7)式的节点。式(7)中 ρ_{ab}为映射域警戒参数。匹配跟踪循环结束的条件为：一个匹配的 ART_a 类别从已储存的类别中被选择；在训练期间一个新类别被创立；ART_a 的警戒参数值大于1。

$$|x^a|=|I\wedge w_J^a|\geqslant\rho_a|I| \tag{6}$$

维普资讯 http://www.cqvip.com

$$|x^{a2}| = |y^a \wedge w_j^a| \geqslant \rho_{a1}|y^a| \quad (7)$$

综上所述，FUZZY ARTMAP 的学习算法如下：

1）网络初始化，包括网络连接权及相关属性的初始化。

2）提供一目标模式给 ART_1，确定 ART_a 中的获胜神经元，并进行学习。

3）提供与目标模式相对应的输入模式给 ART_a，确定 ART_a 中的获胜神经元。

4）确定映射域输出向量并与 ART_1 的输出向量进行匹配。

5）如两者相匹配，映射域权向量学习，ART_a 权向量学习，转7）。

6）如不匹配，执行匹配跟踪，转3）。

7）如还有未学习的模式，则返回2），否则结束。

4. 主要应用领域

FUZZY ARTMAP 神经网络有许多应用研究报道，如心电图波形识别[11]、卫星遥感图像分析[12]、机器人导航[13]和数字电路设计[14]等等。总体上可归纳为三个方面：作为有导师分类器[9]；作为概率估计器[15]；用于知识发现[16]。

在映射域快速学习模式下（$\rho_{ab}=1$），FUZZY ARTMAP 神经网络用作有导师分类器，训练阶段建立输入模式和目标模式间的映射，测试阶段根据输入模式给出网络输出，实现有导师分类。在映射域慢速学习模式下（$\rho_{ab}<1$），映射域权向量提供了输入向量 a 属于类别 c 的后验贝叶斯概率 p(c|a) 估计，故 FUZZY ARTMAP 神经网络可用作概率估计器，且概率估计是非参数性，不需对潜在的概率分布作先验的假定。

FUZZY ARTMAP 神经网络在监督学习时，首先进行自组织学习，形成输入样本的不同类别（F_2^a 层的不同神经元），然后将每一类别与各自的预测（F_2^b 中的神经元）联系起来，这种联系可以看作是以输入特征为前提以分类结果为结论的 IF－THEN 规则。因此，在 FUZZY ARTMAP 中，F_2^a 层中每个已被编码的神经元对应着一条规则，且与该类别神经元相应的权向量可翻译为该条规则前提的语言描述。规则抽取一般分为两个阶段：网络剪枝，移去可信度低于选择阈值的识别节点；连续权向量的量化，以使最后的系统状态能被翻译成可用的规则集。

5. 最近研究进展

FUZZY ARTMAP 神经网络近期研究进展主要体现在：针对实际应用中存在的问题，对网络进行的改进，即牺牲网络的一般性来换取在具体问题上性能的提高，下面进行具体说明。

FUZZY ARTMAP 神经网络性能受训练阶段模式提供顺序影响，常规的做法是通过投票策略来削弱其影响，文[17]中给出了另一种方法，通过一个基于最大－最小（max－min）聚类算法的排序过程实现对训练模式的排序，从而以一个固定的训练模式次序提供给网络。搜索时间上，在最坏的情况下存在着搜索时间与已学习的模板数成正比的问题，文[18]将搜索和共振转化为一个优化问题来解决这个问题。FUZZY ARTMAP 倾向于创立比实际需要更多的类别，这往往引起对训练集的过拟合，使网络在测试集的性能并不随循环次数和所创立的类别数增加而单调增加，为此一个基于可信度的网络剪枝方法在文[19]中给出。

在解决有噪音映射时，噪音连接可能被看作新连接，这将导致过学习和网络泛化性能的降低，在分类问题上允许子类被分解的匹配跟踪机制易引起类别增生。PROBART[20]修正了这一点，通过使用层间节点连接的概率信息去估计输出，ART－EMAP[21]是通过空间和时间证据积累来扩展 FUZZY ARTMAP 神经网络性能，合并了监督、非监督学习和中期记忆处理去完成在噪音输入环境下稳定的模式类别识别。FUZZY ARTMAP 神经网络在区分相似输入序列时，需要较高的警戒参数，但这将使系统对噪音敏感，阻止系统正确分类有噪音的输入，ARTMAP-DS[22]通过折扣相似性，即通过抑制类别表示共同拥有的特征，从而突出它们间的不同，较好地解决了这个问题。

在快速学习时，FUZZY ARTMAP 神经网络不能编码不相容事件（相同的输入得出不同的结果），慢速学习通过在事件中进行平均能提供更可能的估计，但牺牲了网络编码稀有事件的能力。ARTMAP－IC[23]通过修改其搜索算法使网络能编码不相容事件。传统的匹配跟踪算法（MT＋）在搜索期间警戒参数值不允许减小，仿佛搜索循环可无限制地快，新算法（MT－）允许有限快速搜索，搜索期间警戒参数值在下一个节点被测试违反匹配标准前可轻度减小，从而允许相同的输入预测不同的输出，建立截然不同的类别。

胜者为王（WTA）编码使网络能保持稳定的记忆，但这种类型编码在快速学习和训练集有噪音时可能引起类别增生，多层感知器（MLP）具有好的编码压缩和噪音容忍性能，但在快速学习时会出现灾难性的遗忘。分布自适应谐振理论模型（dART）[24]合并了两者的优点，允许任意分布编码表示，并用动态权向量取代传统的神经网络通道权向量。

此外，亦有 FUZZY ARTMAP 神经网络与其他神经网络联合使用的研究报道，如与概率神经网络 PNN 混合利用[25]，在学习阶段 FUZZY ARTMAP 被用作潜在的聚类算法，将输入模式分类成不同类别，在预测阶段，PNN 被用来给出一个输出的贝叶斯概率估计。

维普资讯 http://www.cqvip.com

结束语 自适应谐振理论(ART)通过对人类认知活动的模拟,解决了传统神经网络存在的可塑性-稳定性二难问题。模糊自适应谐振理论网络(FUZZY ARTMAP)综合了模糊集理论和自适应谐振理论,具有很多优良的性能,其未来的发展方向我们认为主要在以下三个方面:(1)针对具体问题进行改进;(2)随神经科学和脑功能研究的深入而发展,如 Coupled-ART[26]就是对人类两大脑半球结构和功能的模拟;(3)与其他理论结合,如 FTART[27]是域理论和自适应谐振理论的结合,FUZZY ART[7]是模糊理论与自适应谐振理论的结合。

参 考 文 献

1 Lippmann R P. An introduction to computing with neural nets. IEEE ASSP Magazine, 1987. 4~22

2 Kramer M A, Leonard J A. Diagnosis using back-propagation neural networks: analysis and criticism. Computer Chem. Engin., 1990, 14. 1328~1338

3 Ripley B D Neural networks and related methods for classification. J. R. Statist. Soc. B., 1994, 56(3): 409~456

4 Carpenter G A, Grossberg S. A massively parallel architecture for a self-organizing neural pattern recognition machine. Computer Vision, Graphics and Image Processing, 1987, 37: 54~115

5 Carpenter G A, Grossberg S. Stable self-organization of pattern recognition codes for analog input patterns. Applied Optics, 1987, 26(23): 4919~4930

6 Carpenter G A, Grossberg S. ART3: Hierarchical search using chemical transmitters in self-organizing pattern recognition architectures. Neural Networks, 1990, 3: 129~152

7 Carpenter G A, Grossberg S, Rosen D. Fuzzy ART. Fast stable learning and categorization of analog patterns by an adaptiver resonance system. Neural Networks, 1991, 4. 759~771

8 Carpenter G A, Grossberg S, Reznolds J H. ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network. Neural Networks, 1991, 4: 565~588

9 Carpenter G A, Grossberg S. Fuzzy ARTMAP: A neual network architecture for incremental supervised learning of analog multidimensional maps. IEEE Transactions on Neural Networks, 1992, 3(5): 698~713

10 Zadeh L. Fuzzy sets, Inform. Contr, 1965, 8: 338~353

11 Ham F M, Han S. Classification of cardiac arrhythmia using fuzzy ARTMAP. IEEE Transactions on Biomedical Engineering, 1996, 43: 425~430

12 Carpenter G A, Gjaja M N, Gopal S, Woodcock C E. ART neural networks for remote sensing: Vegetation classification from Landsat TM and terrain data. IEEE Transactions on Geoscience and Remote Sensing, 1997, 35: 308~325

13 Racz J, Dubrawski A. Artificial neural networks for mobile robot topological localization. Robotics and Autonomous Systems, 1995, 16: 73~80

14 Kalkunte S S, Kumar J M, Patnaik L M A network approach for high resolution fault diagnosis in digital circuits. Proc. of the Intl Joint Conf. Neural Networks, 1992. 83~88

15 Carpenter G A, Grossberg S, Reynolds J H. A Fuzzy ARTMAP nonparametric probability estimator for nonstationary pattern recognition problems. IEEE Transaction on Neural Networks, 1995, 6(6): 1330~1336

16 Carpenter G A, Tan A H. Rule extraction: From neural architecture to symbolic representation. Connection Science, 1995, 7: 3~27

17 Dagher I, et al. An ordering algorithm for pattern presentation in Fuzzy ARTMAP that tends to improve generalization performance. IEEE Transaction on Neural Networks, 1999, 10(4): 768~778

18 Wu P S, Li Ming. Supervised and unsupervised fuzzy-adaptive Hamming net. Pattern Recognition, 1999, 32. 1801~1816

19 Carpenter G A, Tan A H. Rule extraction, Fuzzy ARTMAP and medical databases. Proc. of World Congress on Neural Networks, 1993, 1. 501~506

20 Marriott S, Harrison R F A modified fuzzy ARTMAP architecture for the approximation of noisy mappings Neural Networks, 1995, 8(4). 619~641

21 Carpenter G A, Ross W D. ART-EMAP: A neural network architecture for object recognition by evidence accumulation. IEEE Transactions on Neural Networks, 1995, 6(4): 805~818

22 Carpenter G A, Wilson F D M. ARTMAP-DS. Pattern discrimination by discounting similarities. In: Proc. of the Intl. Conf. Artificial Neural Networks, 1997. 607~612

23 Carpenter G A, Markuzon N. ARTMAP-IC and medical diagnosis: instance counting and inconsistent cases. Neural Networks, 1998, 11: 323~336

24 Carpenter G A. Distributed learning, recognition, and prediction by ART and ARTMAP neural networks. Neural Networks, 1997, 10(8): 1473~1494

25 Lim C P, Harrison R F. Probabilistic fuzzy ARTMAP: An autonomous neural network architecture for bayesian probability estimation. Proc. of Fourth Intl. Conf. on Artificial Neural Networks, 1995, 6: 148~153

26 Lee N I. A coupled-ART neural network capable of modularized categorization of patterns. Pattern Recognition Letters, 1999, 20: 131~140

27 陈兆乾,周戎,等. 一种新的自适应谐振算法 FTART. 软件学报, 1996, 7(8): 458~465