

油水层快速识别改进的 CP 网络方法初探

陆万雨¹, 管志宁¹, 宋子齐²

(1. 中国地质大学 工程技术学院 北京 100083; 2. 西安石油学院 陕西 西安 710065)

摘要: CP 网络具有分类快速、准确等优点。文中详细介绍了 CP 网络的算法, 并对其进行了部分改进, 增加了误差检验和网络测试功能, 使用 Visual Basic 6.0 开发了相应仿真软件, 并将其用于某油田的油、水层快速识别。实际应用表明, 改进的 CP 网络训练周期短、识别准确率高, 完全可以用于油、水层等的快速识别, 具有广阔的应用前景。

关键词: CP 网络; CP 算法; 误差检验; 油水识别

中图分类号: P631.1; TP31

文献标识码: A

文章编号: 1000-8918(2002)01-0053-04

用神经网络解决“识别”问题已得到人们的认可^[1~4], 但目前使用最为广泛的 BP(back-propagation)网络有学习速度慢等缺陷^[1,2], 满足不了快速识别的需要。CP 网络即对向传播(counter-propagation)神经网络, 是由美国神经计算机专家 Robert Hecht-Nielsen 于 1987 年提出的^[1,2]。CP 网络算法简练, 在数据压缩等领域已得到有效应用^[1], 但目前尚未见到在油田测井资料处理方面的相关应用报导。为了解决油、水层快速识别问题, 本文首次引入了具有分类快速、准确等优点的 CP 网络, 并结合某油田的实际资料对 CP 网络进行了应用研究, 这些研究工作同时也是对 CP 网络在测井领域应用的一种尝试。

1 CP 网络及其改进

1.1 CP 网络及其算法

CP 网络是 Kohonen 特征映射网络(又叫自组织特征映射网络—self-organizing feature map, 简称 SOM)与 Grossberg 基本竞争型网络的巧妙结合。如图 1 所示, 网络由输入、竞争、输出 3 层组成。由输入层至竞争层, 网络按 SOM 的学习规则产生竞争层的获胜神经元, 并按这一规则调整相应的输入层至竞争层的连接权; 由竞争层至输出层, 网络按 Grossberg 基本竞争型网络学习规则得到各输出神经元的实际输出值, 并按有教师示教的误差校正方法, 调整由竞争层至输出层的连接权。

CP 网络的具体算法^[1]如下: 设网络输入层有 N 个神经元, 竞争层有 Q 个神经元, 输出层有 M 个神经元, 输入层与竞争层之间的连接权为 w_{ji} ($j=1, 2, \dots, Q; i=1, 2, \dots, N$), 竞争层与输出层之间的连接权为 v_{li} ($l=1, 2, \dots, M; i=1, 2, \dots, Q$), 共有 P 对学习模式, 输

入模式 $A_k = (a_1^k, a_2^k, \dots, a_N^k)$, 希望输出模式 $C_k = (c_1^k, c_2^k, \dots, c_M^k)$, 网络实际输出为 $C'_k = (c'_1, c'_2, \dots, c'_M)$, $k=1, 2, \dots, P$ 。

网络的学习(训练)规则如下: ①将连接权 w_{ji} 和 v_{li} 赋予 $[0, 1]$ 区间内的随机值, 令学习次数记录变量 $t=1$, 将输入模式 A_k 按

$$a_i^k = a_i^k / \|A_k\|, \quad (1)$$

$$\|A_k\| = \sqrt{\sum_{i=1}^N a_i^2} \quad (2)$$

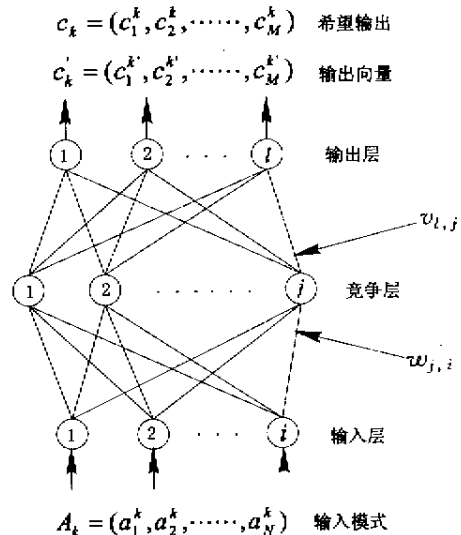


图 1 CP 网络结构

收稿日期 2001-06-06

进行归一化处理; ②将输入模式 A_k 提供给网络的输入层; ③将连接权 w_{ji} 按

$$w_{j,i} = w_{j,i} / \|w_{j,i}\|, \quad (3)$$

$$\|w_{j,i}\| = \sqrt{\sum_{i=1}^N w_{j,i}^2} \quad (4)$$

做归一化处理, ④求每个竞争层神经元加权输入和

$$s_j = \sum_{i=1}^N (a_i^k w_{j,i}); \quad (5)$$

⑤求连接权 $w_{j,i}$ 中与 A_k 距离最近的向量

$$w_g = \max_{j=1}^Q \sum_{i=1}^N (a_i^k w_{j,i}) = \max_{j=1}^Q s_j, \quad (6)$$

将神经元 g 的输出 b_g 置 1, 其余竞争层神经元输出置为 0, 即

$$b_j = \begin{cases} 1 & j = g \\ 0 & j \neq g \end{cases}; \quad (7)$$

⑥将连接权向量 w_g 按

$$w_{g,i}(t+1) = w_{g,i}(t) + \alpha \cdot [a_i^k - w_{g,i}(t)] \quad (8)$$

进行调整, 式中 α 为学习率 ($0 < \alpha < 1$), 然后将连接权向量 w_g 按式 (3) 重新进行归一化, ⑦按

$$v_{l,j}(t+1) = v_{l,j}(t) + \beta \cdot b_j \cdot (c_l^k - c_l^{k'}) \quad (9)$$

调整由竞争层至输出层的连接权 $v_{l,j}$ (9) 式简化为

$$v_{lg}(t+1) = v_{lg}(t) + \beta \cdot (c_l^k - c_l^{k'}), \quad (10)$$

式中 β 为学习率 ($0 < \beta < 1$); ⑧求输出层各神经元的加权输入, 并将其作为输出神经元的实际输出

$$c_l^{k'} = \sum_{j=1}^Q (b_j \cdot v_{lj}) = v_{lg}; \quad (11)$$

⑨返回步骤②, 直至 P 个输入模式全部提供给网络;

⑩令 $t = t + 1$, 返回步骤②, 直至 $t = T$ (一般情况下, $500 \leq T \leq 10\,000$, T 为总学习次数)。

网络学习结束后, 按以下步骤进行网络回想(预测): ①将输入模式 A 提供给网络的输入层, ②按

$$b_g = \max_{j=1}^Q \left(\sum_{i=1}^N w_{j,i} a_i \right) \quad (12)$$

求出竞争层的获胜神经元 g ; ③令 $b_g = 1, b_j = 0 (j \neq g)$, 按

$$c_l^{k'} = v_{lg} \cdot b_g \quad (13)$$

求得输出层各神经元的输出值。由此产生输出模式 $C' = (c_1', c_2', \dots, c_M')$, 从而得到输入模式 A 的分类识别结果。

1.2 CP 网络的改进

从 CP 网络的具体算法可以看出, 网络的学习(训练)是以指定的学习次数进行的, 学习样本按固定的顺序进入网络输入层, 这对 CP 网络在数据压缩、图像处理等领域的应用来说是必要的^[2]。但对于分类识别方面的应用, 我们在实用中发现, 以这种方式进行网络学习, 有 2 个较明显的缺点: ①容易造

成网络的“过学习”; ②网络的输出“飘忽不定”。究其原因, 一方面是在学习过程中没有进行误差计算和分析, 对输出精度也没有进行监测, 网络的输出精度取决于最后一次训练的输出精度, 而这不一定是整个训练过程中输出精度最高的; 另一方面是网络竞争层神经元个数的选取往往根据经验而定, 而它又对网络性能有较大影响, 当神经元个数太多时不仅会造成网络的冗余, 增加网络学习时间, 而且极易造成“过学习”现象, 而太少时又会使网络输出精度降低甚至无法工作。

为此, 我们进行了改进研究, 采取了以下几个措施。

1. 打乱学习样本的输入顺序, 让计算机随机选取。

2. 增加误差处理过程, 并对误差进行监测, 每一次学习结束后都进行误差计算, 一旦达到误差要求就立即停止训练(即学习过程结束)。误差计算公式如下:

$$E = \frac{1}{2} \sum_{k=1}^P \sum_{l=1}^M (c_l^k - c_l^{k'})^2. \quad (14)$$

3. 增设网络测试功能, 训练好的网络在应用前先进行测试, 符合要求后才可用于工作, 否则重新进行训练或调整网络结构。测试样本一般从学习样本中提取(要加上适当噪声)。

对于竞争层神经元个数的选取, 我们经过反复试验, 发现当竞争层神经元个数稍多于 2 倍学习样本数时, 网络的性能较好, 但这仅为经验取值, 仍有待以后进一步的探讨和研究。

我们以 Visual Basic 6.0 为工具开发出了相应实现软件^[5](图 2), 并以异或问题为例进行试验, 对比

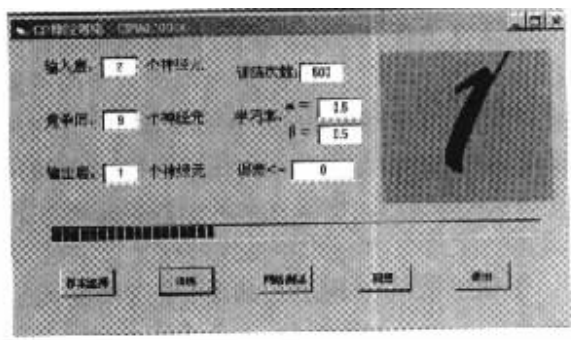


图 2 CP 网络界面

表 1 异或问题样本

学习样本			预测样本		
输入		希望输出	输入		理论输出
0	0	0	0.1	0.2	0
0	1	1	0.05	0.8	1
1	0	1	0.9	0.2	1
1	1	0	0.75	0.95	0

表 2 CP 网络参数及取值

网络参数	参数值
网络结构	2-9-1
预设学习次数	500
误差要求	0
α	0.5
β	0.5

表 3 CP 网络改进前后试验结果

试验次数	改进前			改进后		
	学习次数	全局误差	正确识别率/%	学习次数	全局误差	正确识别率/%
1	500	0	100	149	0	100
2	500	0.000 10	100	151	0	100
3	500	0.555 56	0	167	0	100
4	500	0.537 78	50	89	0	100
5	500	0.231 40	100	143	0	100
平均	500	0.264 97	70	140	0	100

改进前后的 CP 网络性能(表 1~3)。从试习速度和输出精度上有了大幅度的提高,而且输出结果稳定,消除了改进前 CP 网络输出结果“飘忽不定”的现象。因此,改进后的 CP 网络可以适用于解决快速识别问题。

2 应用实例

我们将改进的 CP 网络应用于解决某油田一研究区的油、水层快速识别问题。该区共有 6 口井,部分井段进行了试油,其测井资料共有 5 条测井曲线:

感应电导率(σ_i)、自然伽马(N)、电阻率(ρ)、密度(d)、自然电位(U)。根据实际情况,我们选用了 5-35-2 的网络结构:5 个输入节点对应 5 种测井值,2 个输出节点对应油、水,竞争节点数取 35 个(稍多于 2 倍学习样本数),网络的学习系数(学习率)根据经验取值,在此取 $\alpha = \beta = 0.5$,预设网络训练次数为 500。根据试油资料提取了 10 组学习样本,另外从可靠的测井解释中补充了 6 组学习样本(主要为试油资料验证结果(表 3)可以看出,改进后的 CP 网络不仅在学缺少的水层及干层样本),网络的学习样本及

表 4 油水快速识别学习样本

样本序号	输入参数					希望输出		网络输出对应的综合解释结论
	$\sigma_i/(\text{ms}\cdot\text{m}^{-1})$	$\rho/(\Omega\cdot\text{m})$	U/mV	N/API	$d/(\text{g}\cdot\text{cm}^{-3})$	Y1	Y2	
1	44.1	22.7	-43.8	33.1	2.29	1	0	油
2	100.0	10.0	-41.1	59.3	2.24	1	1	油水(同层)
3	42.3	23.6	-39.1	48.4	2.15	1	0	油
4	87.8	11.4	-39.4	60.0	2.16	1	0	油
5	57.0	17.5	-30.9	61.7	2.35	1	1	油水(同层)
6	34.8	28.7	-32.0	41.5	2.17	1	1	油水(同层)
7	23.6	42.3	-51.0	26.5	2.22	1	0	油
8	54.7	18.3	-54.8	43.7	2.26	1	1	油水(同层)
9	75.6	13.2	-40.1	41.3	2.33	1	0	油
10	50.4	19.9	-37.8	26.0	2.30	1	0	油
11	221.7	4.5	-42.2	90.0	1.91	0	1	水
12	265.8	3.8	-60.9	59.6	2.14	0	1	水
13	281.7	3.5	-57.8	57.1	2.33	0	1	水
14	248.8	4.0	-48.9	69.7	2.52	0	1	水
15	91.7	10.9	-14.6	70.2	2.15	0	0	干层
16	129.2	7.7	-19.8	69.7	2.15	0	0	干层

输出对应的情况如表 4 所示。

采用以上学习样本对网络进行了训练和测试, 万方数据

然后应用训练和测试好的网络对 25 个重点层位进行了油、水层预测,并与用交会图法、统计分析法得

表 5 CP 网络工作情况及结果

试验 次数	学习情况			测试情况		预测情况 (与交会图法、统计分析法解释结论对比)				与最新试油结论对比	
	学习 样本数	学习次数	误差	测试 样本数	符合率 %	待识别 样本数	完全符合 样本数	基本符合 样本数	不符合 样本数	试油层数	预测符合 层数
1	16	194	0	4	100	25	14	9	2	4	4
2	16	192	0	4	100	25	13	10	2	4	4
3	16	190	0	4	100	25	14	9	2	4	4
4	16	194	0	4	100	25	13	10	2	4	4
5	16	193	0	4	100	25	13	10	2	4	4
平均	16	193	0	4	100	25	13	10	2	4	4

注:基本符合是指解释结论接近(如油层与油水同层)且与岩芯、试油等资料不矛盾的情况。

出的结论进行了对比(如表 5 所示),从中可以看出:CP 网络的预测结果与交会图法、统计分析法的综合解释结论比较接近(25 个待预测样本仅有 2 个样本结论不一致,基本符合的平均占 92%),特别是与最新试油层位的结论完全一致,准确率达到 100%;而 CP 网络的学习周期很短(平均学习次数才 193 次),用时不足 1 min,实现了油、水层的快速识别。

3 结 语

从整体上看,CP 网络属于有教师示教型网络,而由输入层和竞争层构成的 SOM 网络又是一种典型的无教师示教型网络,因此,这一网络既具有无教师示教型网络分类灵活、算法简练的优点,又兼备有教师示教型网络分类精细、准确的长处,是 2 种不同类型网络的有机结合。实践表明,经适当改进后的

CP 网络学习周期短、识别准确率高,完全可以用于油、水层的快速识别,具有广阔的应用前景。

本文完成过程中得到了侯俊胜副教授、蔚远江博士、陈程博士的大力帮助,在此表示衷心感谢!

参考文献:

[1] 王伟.神经网络原理—入门与应用[M].北京:北京航空航天大学出版社,1995.
[2] 焦李成.神经网络系统理论[M].西安:西安电子科技大学出版社,1992.
[3] 肖慈珂,姜建立,谭世君.神经网络技术用于测井解释的评述[J].测井技术,1999,23(5):389-392.
[4] 李云省,李士伦.BP 网络快速识别油气层初探[J].天然气工业,1999,19(4):40-42.
[5] 希望图书创作室.中文 Visual Basic 6.0 教程[M].北京:宇航出版社,北京希望电子出版社,2000.

THE TENTATIVE APPLICATION OF THE IMPROVED CP-NETWORK
TO RAPID IDENTIFICATION OF OIL-WATER LAYERS

LU Wan-yu¹, GUAN Zhi-ning¹, SONG Zi-qi²

(1. China University of Geosciences, Beijing 100083, China; 2. Xi'an Petroleum College, Xi'an 710065, China)

Abstract: The Counter-Propagation network is excellent in solving the sorted problems. It can sort effectively, and its algorithm is concise. This paper describes in detail the algorithm of CP-network and its improvement-the addition of error processing and network testing functions. The authors developed its simulated software by using the Visual Basic 6.0 and utilized it to identify the oil-water layers. Practice shows that the training period of the improved CP-network was short and its identification ratio was very high. It can realize the quick identification of oil-water layers and will hence have a very wide application prospect.

Key words: CP-network, CP-algorithm, error processing, identification of oil-water layers.

作者简介:陆万雨(1972-),男,广西人,壮族,1999年毕业于中国地质大学(北京)工程技术学院,获硕士学位,现为在读博士,地球探测与信息技术专业。