

· 综合研究 ·

# 动态概率神经网络及油气概率分布预测

徐旺林\* 庞雄奇

吕淑英

Michael R Berthold

(石油大学盆地与油藏研究中心·北京) (海顿新技术有限公司)

(Konstanz 大学·德国)

## 摘 要

徐旺林, 庞雄奇, 吕淑英, Michael R Berthold. 动态概率神经网络及油气概率分布预测. 石油地球物理勘探, 2005, 40(1): 65~70

传统概率神经网络(PNN)采用了前向径向基函数神经网络的局部化的、高斯型作用函数,虽然具有较好的分类能力,但也存在以下局限性:①如果学习样本增多,模式层的神经元个数也增多,导致运算矩阵的增大而使之几乎丧失处理大数据量实际资料的能力;②模式层到求和层的权值固定为常数,必然要求学习集合中各类样本的数量相等,从而会影响其接受实际数据的能力。为此文中引入了动态概率神经网络(DPNN),它与传统PNN在结构上的区别在于:①模式层到求和层不以等权连接,权值由学习样本集的概率分布所确定;②学习样本集中各类样本的数量(各类在模式层的神经元个数)可以不相等。文中还介绍了DPNN算法。理论数据测试充分展示了DPNN结构具有动态调整、学习收敛快速、分类识别能力强大等特点。通过选取G油田实际资料的22种属性作为网络的输入向量,用DPNN进行分类识别得到了含油气概率分布图,可为预测有利油气圈闭及油水分布规律提供依据。

**关键词** 动态概率神经网络 地震属性 模式识别

## 1 引言

近年来,利用先进、快速的工作站技术,提取地震资料中丰富的地震属性信息,用于分析储集层的物性甚至含油(气)性是油气概率分布预测领域的研究方向。对地震属性进行综合分析或模式分类、识别的方法很多,其中以人工智能神经网络的应用尤为广泛。在国内,朱广生、杜世通等人于1993年将人工神经网络用于地球物理勘探<sup>[1~3]</sup>,随后,有人将人工神经网络用于油气预测和反演处理<sup>[4~6]</sup>,也取得了成功。

地球物理勘探中广泛使用的神经网络主要为BP网络,SOM、Hopfield、ART等网络也得到了不同程度的应用,这些网络大多存在学习收敛速度慢、容易陷入局部极小、训练过程中有振荡现象等不足之处。Specht等人于1990年将贝叶斯概率理论与神经网络相结合<sup>[7]</sup>,借助径向基核函数提出了概率神经网络(PNN),在一定程度上避免了BP网络的缺点,并且使神经网络在分类识别时以概率方式来

表述。Daniel等人利用概率神经网络在地震属性与测井曲线之间建立映射关系,由数据驱动的策略来预测测井曲线特征属性<sup>[8]</sup>。早期的概率神经网络相对简单,其网络结构在学习训练过程中固定不变,对数据的适应能力也较差。

本文在早期概率神经网络基础上,引入了高效的动态概率神经网络(Dynamic Probabilistic Neural Network,简称DPNN),可依据学习数据动态调整网络结构,根据学习数据的概率分布动态确定网络节点之间的连接权,网络的分类识别结果也以概率值来表示,网络的学习收敛速度也较快(一般迭代次数小于5次,网络即可收敛)。DPNN的最大特点是分类识别能力比早期的PNN高。

## 2 概率神经网络与分类(PNN)

Specht在1990年提出了概率神经网络(Probabilistic Neural Networks,简称PNN)的概念<sup>[7]</sup>,它基于前向径向基函数神经网络(Radial Basis Function,简称RBF)发展而来,所以从理论体系上来说,

\* 北京昌平石油大学盆地与油藏研究中心,102249

本文于2003年12月29日收到,修改稿于2004年10月12日收到。

PNN 可以归入径向基函数网络。由于它以概率密度函数(Probabilistic Density Function, 简称 PDF)的形式解释了所建立的神经网络的结构, 而且性能优于其他分类器, 因此在应用中受到广泛关注。

在用神经网络进行分类识别时, 网络输出的一个有效解释是样本归入各类的概率。在这种情况下, 网络的作用实际上是估计学习样本的概率密度函数。如果网络的输出是输入空间中给定点的期望输出值, 则形成一种映射关系, 可以把分类问题看成回归拟合问题。

在常规统计中, 人们可以根据给定的概率模型来得到某一数值变量的概率, 而采用贝叶斯统计则可依据给定的数据来估计概率模型的有效性, 更主要的是由贝叶斯统计可以根据给定的数据来估计概率密度模型。

在分类问题中, 可以构造所有可能类别的概率密度函数的估计。对于待判别数据, 可以比较其归入各类的概率, 以选择概率最大的类别。因此用概率神经网络来进行分类, 实质上就是用给定的学习数据来估计并建立概率密度函数。

构造概率密度函数的传统方法是根据数据来进行估计, 即首先假定一个概率密度函数的模型(如正态分布), 然后利用解析方法估计模型参数(如正态分布的均值与标准差); 另一种估计概率密度函数的方法是采用基于内核的估计。可以认为, 某一事件(输入空间中的一个点)的发生表明了在该事件处的概率密度。也就是说, 一组相互接近的事件指示了一个高概率密度区, 某一事件与这一组事件越接近, 则其概率密度值越高; 反之, 当某一事件远离这一组事件时, 其概率密度值就会下降。在基于内核的估计方法中, 在每一个能够得到的事件(学习样本)处都有一个密度函数(内核), 所有事件的密度函数合起来便可以得到总的概率密度函数。一般来说, 典型的内核函数是一系列钟形的高斯型函数, 如果有足够多的学习数据, 就会得到一个对真实概率密度函数任意精度的近似。这种基于多个内核的概率密度函数的逼近过程事实上是径向基函数网络最佳逼近特性的延伸, Specht 称之为概率神经网络(PNN)<sup>[7]</sup>。概率神经网络的一个主要应用是进行分类模式识别, 它利用贝叶斯概率理论, 可对模式分类问题提供一个统一的解决方案。在学习模式下, 概率神经网络由有监督训练集在模式层(即 RBF 的隐层)获得分

布函数; 在识别模式下, 概率神经网络利用分布函数来估计输入特征向量归入某类的似真性。若假定所有类在是否接受某一输入特征向量时是等概率的, 这时类别的确定仅依赖于输入特征向量与某一类分布函数的接近程度。

图 1 为概率神经网络结构示意图, 概率神经网络由输入层、模式层、求和层、输出层(未绘出)共 4 层组成。输入层将特征向量传递给网络; 模式层计算输入特征向量与训练集中各个模式的匹配关系, 以其距离送入高斯型激活函数得到模式层的输出; 模式层到求和层以等权连接来确定匹配概率。每一类只有一个求和层单元, 求和层单元只与属于自己类的模式层单元相连接, 而与模式层中的其他单元没有连接, 因此求和层单元只接受属于自己类的模式层单元的输出, 而与属于其他类别的模式层单元的输出无关。求和层单元的输出与各类基于内核的概率密度函数的估计成比例, 通过输出层的处理, 便得到各类的概率估计。

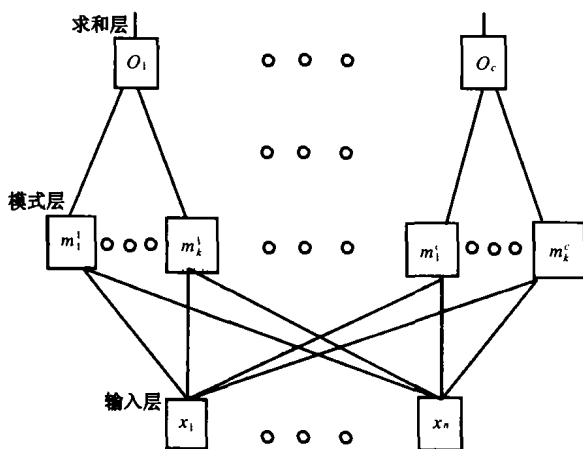


图 1 概率神经网络结构示意图

模式层到求和层以等权连接; 模式层的神经元个数与学习样本数目一致

### 3 动态概率神经网络(DPNN)原理

由于传统的概率神经网络采用了 RBF 神经元的局部化的、高斯型作用函数, 因而具有较好的分类能力, 但是传统的概率神经网络也存在一定的局限性。首先, 它没有解决 RBF 网络中关于如何确定模式层节点中心和模式层各单元概率密度函数的标准差问题, 只是简单地将各个学习样本作为相应模式层节点的中心, 并将所有模式层节点的标准差置为常数, 这势必会影响模式识别的效果。其次, 传统的

概率神经网络简单地将模式层到求和层的权值固定为常数,而且模式层的神经元个数也与学习样本数目一致,这会导致以下两个问题:①如果学习样本增多,模式层的神经元个数也增多,导致运算矩阵的增大而使之几乎丧失了处理大数据量实际资料的能力;②模式层到求和层的权值固定为常数,必然要求学习集合中各类样本的数量相等,这也会影响其接受实际数据的能力。

与传统概率神经网络相比,动态概率神经网络(DPNN)在结构上具有模式层的节点数是动态可调整的突出特点。由图2可以看出,DPNN与传统的PNN在结构上有较大区别。首先,模式层到求和层不再以等权连接,实际上权值是由学习样本集的概率分布所确定的;其次,没有要求学习样本集中各类样本的数量相等,也意味着各类在模式层的神经元个数可以不相等。

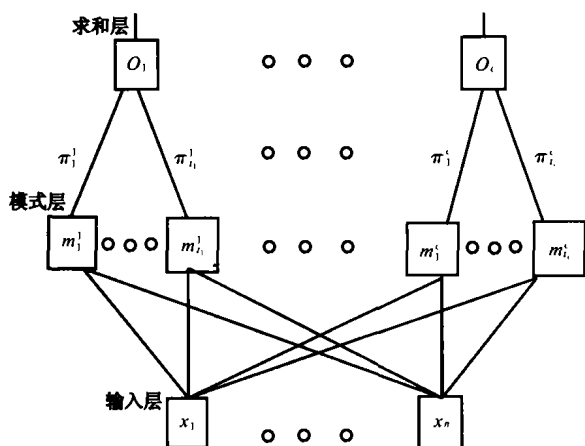


图2 动态概率神经网络(DPNN)结构图

模式层到求和层不以等权连接;各类模式层的神经元个数可以不等,第1类~第c类模式层的神经元个数为 $t_1 \sim t_c$

DPNN以 $n$ 维特征向量 $\mathbf{x} = (x_1, x_2, \dots, x_n)$ 作为输入,输入向量作用到输入神经元并传递给模式层。在模式层中,对于第 $k(1 \leq k \leq c)$ 类,共有 $t_k$ 个 $n$ 维高斯分布 $N(\mu_j^k, \sum_j^k)(j = 1, \dots, t_k)$ 需要计算,其中第 $j$ 个 $n$ 维高斯分布的密度为

$$p_j^k(x) = \frac{1}{(2\pi)^{n/2} |\sum_j^k|^{-1/2}} \times \exp \left\{ -\frac{1}{2} [\mathbf{x} - \mu_j^k]^T (\sum_j^k)^{-1} [\mathbf{x} - \mu_j^k] \right\}$$

式中: $\mu_j^k$ 表示第 $k$ 类中第 $j$ 个神经元的高斯分布的中心; $\sum_j^k$ 为相应的协方差矩阵。一般情况下, $\sum_j^k$ 为正定矩阵。对于第 $k$ 类,共有 $t_k$ 个 $n$ 维高斯分布。

输出层通过组合这些多变量的高斯分布来计算每个类的概率函数值(这是DPNN与传统PNN的主要差别),即

$$O_k(x) = \sum_{j=1}^{t_k} \pi_j^k p_j^k(x)$$

这里 $\pi_j^k$ 表示类内的组合比例因子, $\pi_j^k$ 为非负且满足

$$\sum_{j=1}^{t_k} \pi_j^k = 1 \quad k = 1, \dots, c$$

## 4 动态概率神经网络(DPNN)学习算法

DPNN采用了在学习过程中依据学习数据动态调整网络结构和动态调整标准差的算法,通过匹配和排斥邻近神经元达到分类目的。DPNN的算法在学习过程中依据学习数据动态调整网络结构,从而避免了预定义网络结构无法处理大规模实际资料的缺点。DPNN算法通过在学习过程中动态调整 $\alpha$ 和 $\beta$ 两个参数求取每个神经元的标准差,即利用 $\alpha$ 为合适类别的学习样本确定最小的正确概率,而利用 $\beta$ 来避免错误分类(对每个学习样本而言,它属于某个不正确类别的概率,小于 $\beta$ )。图3为DPNN算法的概率密度图,由图中可见,对学习样本来说,由于 $\alpha$ 和 $\beta$ 的限制,形成了一个既没有匹配、又没有排斥的区域,称之为冲突区。

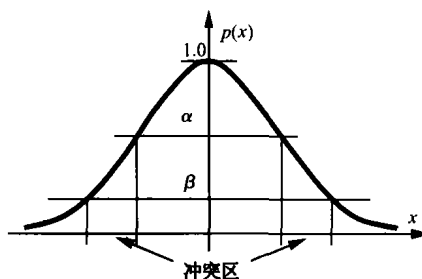


图3 DPNN的算法概率密度图

通过动态调整阈值 $\alpha$ 和 $\beta$ 所确定的冲突区

DPNN算法分为学习过程和分类过程。在学习过程中,对于难以确定类别的学习样本向量,要么自发建立新的RBF神经元,要么调整RBF神经元的冲突半径。学习结束以后,对所有的学习数据集 $(x, k)$ ,要求必须满足两个条件:

(1)第 $k$ 类中,至少有一个神经元的响应大于 $\alpha$ ,即至少存在一个 $i$ 使下式成立

$$p_i^k(x) \geq \alpha$$

(2) 所有与  $x$  不相容的类的响应总是小于  $\beta$ , 也就是对于任意第  $f$  类, 它的所有  $t_f$  个神经元的响应满足

$$\forall f \neq k, 1 \leq j \leq t_f: p_j^f(x) \leq \beta$$

DPNN 网络的学习算法如下:

(1) 在一个学习循环开始前, 将所有的权置为零, 以避免学习样本重复信息累积, 即

$$W_j^k = 0.0$$

(2) 把所有的学习样本送入网络, 如果新的模式能被正确分类, 则相应的权值增加; 否则建立新的 RBF 神经元

$$t_k \leftarrow t_k + 1; \mu_{t_k}^k = x; W_{t_k}^k = 1.0$$

以新的样本向量作为其中心, 初始权为 1, 初始标准差在与其它已有类别不冲突的情况下尽可能大。

(3) 对所有 RBF 神经元来说, 若对当前具体的学习样本响应太高, 则需收缩其标准差, 即

$$\sigma_j^k = \min \left\{ \sigma_j^k, \sqrt{-\frac{\|x_j - \mu_{t_k}^k\|^2}{\ln \beta}} \right\}$$

经过几次迭代以后, 网络结构确定, 误差收敛, 表明学习训练已结束。由于学习训练的迭代性, 最后可能要删除一些权值为零的冗余神经元, 这些神经元在学习早期被引进, 后来被能覆盖更大特征空间的更为优化的神经元所代替。学习结束后, 按下式依据权值计算归一化的类内组合比例因子

$$\pi_i^k = \frac{W_i^k}{\sum_{j=1}^{t_k} W_j^k}$$

其中:  $1 \leq k \leq c; 1 \leq i \leq t_k$ 。

网络的实现采用了 JAVA 语言的动态数组—链表复合数据结构, 这样既发挥了 JAVA 语言本身安全、面向对象等优越特性, 又采用了动态数据结构, 使网络的规模能自动适应数据的规模, 并充分利用了计算机的存储空间, 可处理大数据量的实际资料。

## 5 数值计算

为了检验算法的分类性能, 对传统的概率神经网络(PNN)和动态概率神经网络(DPNN)分别采用经典的双螺旋曲线数据进行分类测试。

双螺旋曲线学习数据如图 4 所示, 采用该图的 196 个数据(分为两类)为样本对 PNN 和 DPNN 分别进行测试, 测试的分类结果如图 5、图 6 所示。从

图 5 中可以看出, PNN 能区分开测试区域内的绝大部分数据点, 展示了其良好的分类能力, 但是并不能区分开测试区域中心点间距离较小的少量样本数据点, 这表明 PNN 具有一定的局限性。由图 6 可以看出, 图中可以完全区分在图 5 中不能区分开的数据点, 展示了良好的分类能力, 表明 DPNN 分类性能优于 PNN。从图 6 中还可以看出, DPNN 的分类结果的两类数据样本之间的过渡带是渐变的, 这表明 DPNN 对模式的识别尚具有一定的模糊性, 这一点与 PNN 只是简单地将特征空间分为两类的结果不同。

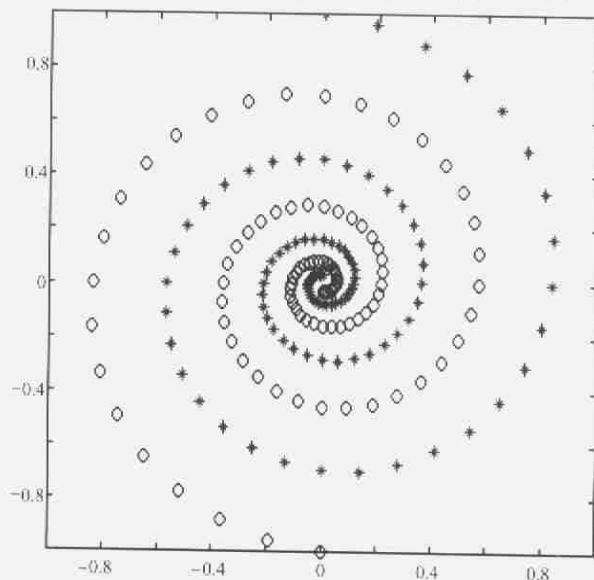


图 4 双螺旋曲线型的学习数据(特征空间)

以图中的 196 个数据为样本(分为两类), 对 PNN 和 DPNN 分别进行学习训练

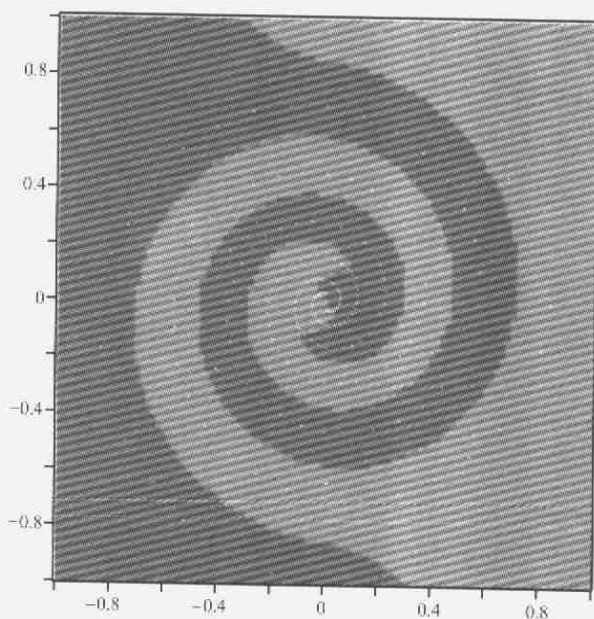


图 5 PNN 对特征空间的分类结果

取图 4 中 10201 个间距为 0.02 的等距离均匀分布在横坐标区间  $[-1, 1]$ , 纵坐标区间  $[-1, 1]$  上的数据点作为测试集, 图 6 同

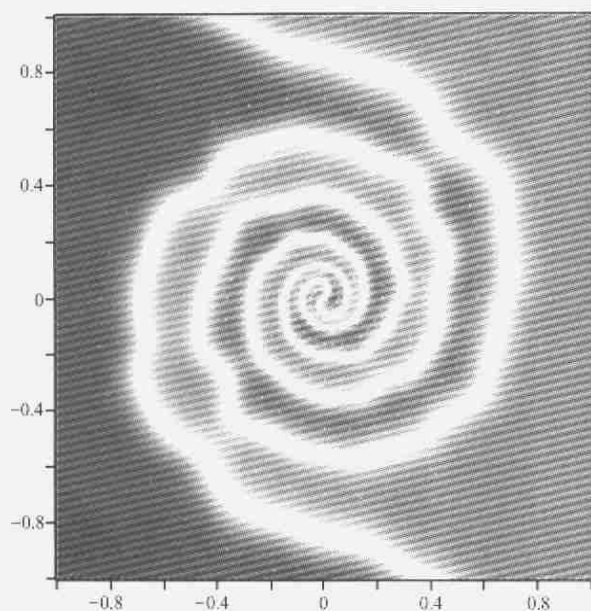


图6 DPNN对特征空间的分类结果

## 6 实际资料处理

我国G油田的勘探目标位于凹陷西南斜坡上,为一鼻状隆起。工区面积约为 $12\text{km}^2$ ,收集到的资料有7口早期探井资料和2000年采集的主频达 $80\text{Hz}$ 的三维高分辨率地震资料,其中1口井为产油井(G25),其余为产水井。研究的目的是寻找有利的岩性圈闭并预测含油气范围。在研究过程中,除了采用波阻抗反演技术预测目的层砂岩岩性圈闭的厚度之外,还提取了沿层的地震特征属性。特征属性提取工作的主要步骤为:①制作精细的合成记录,进行目的层时深标定;②根据标定结果进行目的层顶、底精细追踪解释;③以追踪解释的层位为约束,提取沿层特征属性参数;④利用解释系统在平面上对计算结果进行分析研究,以筛选合理的参数。在上述地震特征属性分析的基础上,共计算了40种特征属性参数(包括反演储层的波阻抗数据体),从中筛选出22种特征属性参数(储层参数反演类1个、振幅类9个、复数道统计类3个、频谱统计类4个、层序统计类5个)构成DPNN的输入特征向量空间<sup>[11]</sup>。在得到地震特征属性参数之后,在井点附近抽取地震道特征属性,按产油和非产油分为两类,构成学习样本集。将学习样本集送入DPNN进行学习训练,迭代5次,网络达到收敛。然后用训练好的网络对整个工区的数据进行识别,从而得到含油概率分布图。

图7为研究区目的层砂岩等厚图,图8为研究

区含油概率分布图。由图7、图8可知,最高含油概率位于G25处东南边(指示最好的含油区)。另外,在工区的西边角上,也有一片含油有利区,是下一步重点勘探的对象。工区中部厚砂岩发育带虽然处在构造较高部位,但预测结果为含水砂岩,究其可能是该砂体与下倾部位的含油砂体之间存在隔夹层,致使低部位的油(气)不能向上运移而在较低部位的圈闭中赋存,推测该砂体为典型的透镜体砂岩隐蔽油气藏。此例成功地解释了以前难以解释的油水分分布规律,为对该区透镜体砂岩隐蔽性油气藏的深入认识提供了有力证据。图7、图8还表明,在研究区内不仅要预测与研究有利的储集圈闭,而且要加强油气运聚规律的深入认识。

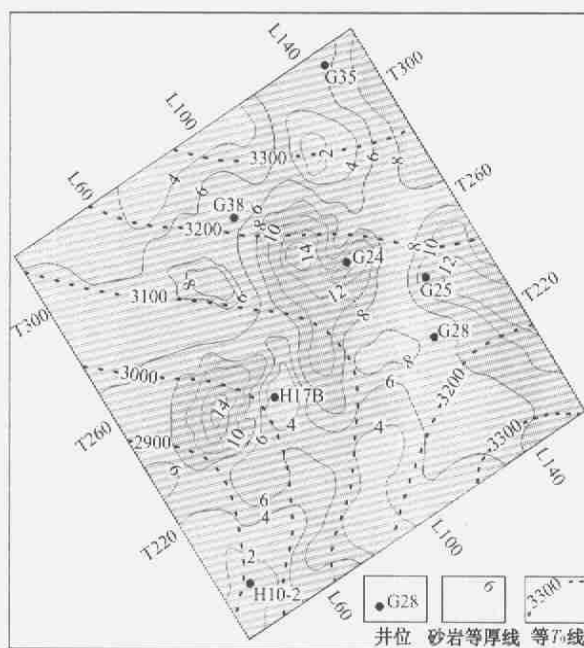
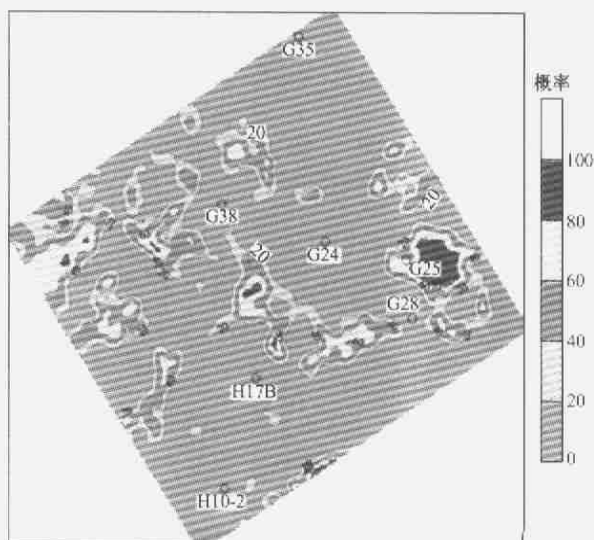
图7 研究区目的层砂岩等厚图(背景为等 $T_0$ 线)

图8 DPNN模式识别得到的含油概率平面图

## 7 结论

理论分析和数值计算结果表明,动态概率神经网络(DPNN)与传统的概率神经网络(PNN)相比,有其独特的优点:①依据学习数据动态调整网络结构;②根据学习数据的概率分布动态确定网络节点间的连接权;③网络的学习收敛速度较快,一般迭代次数不到 5 次网络即可收敛;④DPNN 的分类识别能力较 PNN 有较大提高。尚需指出,本文仅仅涉及 DPNN 在分类识别中应用,下一步的研究方向应是建立动态概率映射神经网络,以对描述同一对象且相互关联的数据进行不同测量域之间的连续映射。从现有的应用效果看,DPNN 展示了良好的应用前景。

### 参考文献

- [1] 朱广生编译. 人工神经网络——勘探地球物理学家的新工具. 国外油气勘探, 1993, 5(2): 155~173
- [2] 杜世通等. 油藏地球物理异常目标的预测——“七五”国家重点科技项目成果报告. 石油大学(华东), 1990
- [3] 杜世通. 地震属性分析与应用技术. 见: 地球物理勘探新方法新技术. 中国石油勘探开发研究院西北分院, 2000
- [4] 杨凤丽, 印兴耀. 神经网络在埕岛北区馆上段油藏描

- 述中的应用. 胜利石油管理局地质科学研究所, 1993
- [5] 许建华, 蔡瑞. 有监督 SOM 神经网络在油气预测中的应用. 石油物探, 1998, 37(1)
- [6] 印兴耀等. 用神经网络和广义逆联合反演波阻抗. 国际地球物理研讨会论文集(北京), 1998
- [7] Spetch D F. Probabilistic Neural Networks. *Neural Networks*, 1990, 3: 109~118
- [8] Daniel P H et al. Use of multiattribute transforms to predict log properties from seismic data. *Geophysics*, 2001, 66(1): 220~236
- [9] Michael R B. Constructive training of probabilistic Neural Networks. *Neurocomputing*, 1997, 19: 167~183
- [10] Xu W L et al. Application of comprehensive reservoir prediction technique in areas with complicated lithologies. *Bulletin of Chinese Geophysical Society*, 2000, 4
- [11] 徐旺林等. 沿层属性分析在储层研究中的应用. 石油地球物理勘探, 2002, 37(增刊): 83~87
- [12] 陈永波等. 陆相地层渗透性砂层预测方法. 石油地球物理勘探, 1999, 34(1): 93~98, 118
- [13] Paul J K. Learning probabilistic Networks. Philips Research Laboratories United Kingdom
- [14] 王士同, 於东军. 径向神经网络的构造理论. 模糊技术与神经网络技术选编
- [15] Wey S H and Juyang W. Hierarchical discriminant regression. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, 22(11)

(本文编辑: 刘勇)

(上接第 64 页)

以取得理想的去噪处理效果。

(3) 利用上述随钻地震波场的时空传播关系, 可以对观测到的随钻地震记录中的各种波场作出合理的解释。

在随钻单炮地震记录的成像处理过程中, 得到了东方地球物理公司研究院严又生先生的帮助, 在此表示感谢。

### 参考文献

- [1] Rector III J W and Hardage B A. Radiation pattern and seismic waves generated by a working roller-cone drill bit. *Geophysics*, 1992, 57: 1319~1333
- [2] Rector III J W and Marion B P. The use of drill-bit energy as a downhole seismic source. *Geophysics*, 1991, 56: 628~634
- [3] Haldorsen J B U et al. Walk-away VSP using drill noise as a source. *Geophysics*, 1995, 60: 978~997

- [4] Poletto F and Dordolo G. A new approach to offshore drill-bit reverse vertical seismic profiling (RVSP). *Geophysics*, 2002, 67: 1071~1075
- [5] Aleotti L et al. Seismic while-drilling technology: use and analysis of the drill-bit seismic source in a cross-hole survey. *Geophysical Prospecting*, 1999, 47: 25~39
- [6] Batini F et al. Seismic while drilling measurements in Italian deep geothermal wells. *Proceedings World Geothermal Congress*, 2000
- [7] Petronio L and Poletto F. Seismic-while-drilling by using tunnel boring machine noise. *Geophysics*, 2002, 67: 1798~1809
- [8] Petronio L et al. Geology prediction ahead of the excavation front by Tunnel Seismic while Drilling (TSWD) method. SEG International Exposition and 73rd Annual Meeting, 2003
- [9] 长春地质学院、成都地质学院、武汉地质学院合编. 地震勘探——原理和方法. 北京: 地质出版社出版, 1980, 96~102

(本文编辑: 朱汉东)