

· 综合研究 ·

# 二维叠前模式识别方法研究

郭淑文\*

(中国地质大学(北京)地球物理与信息技术学院)

郭淑文. 二维叠前模式识别方法研究. 石油地球物理勘探, 2008, 43(3): 313~317

**摘要** 现有的模式识别技术都是针对叠后数据提取属性,而叠加过程恰好丢失了属性随炮检距的变化信息。为此提出针对叠前地震数据提取各种属性,然后再对每个CDP所有地震道属性求取它们相对于炮检距的变化梯度或平均值,再对提取的地震属性采用聚类分析和分类判别的方法进行模式识别。在做分类判别时,提出了先对井旁道目的层提取属性,再应用聚类和测井资料指定样本类别的新方法。经实际资料检验表明,此法能够分辨出储层纵向、横向的变化。

**关键词** 模式识别 叠加 属性提取 聚类分析 分类判别 神经网络

## 1 引言

模式识别可分为聚类分析和分类判别。聚类的任务是把各个样品点间具有相似特性的点划分成一些类,即每个类内的各点在某种意义上彼此相似,而属于不同类的点则具有不同的性质。分类判别与聚类不同的是类数已经确定,然后利用它们去判断每个还不知道类别的点应该属于哪一类<sup>[1]</sup>。

从20世纪80年代初,模式识别技术用于地震储层预测以来,引起了广泛关注<sup>[2~6]</sup>。Bios(1980, 1981, 1983)研究了用聚类技术来确定油藏边界。Sinvhal等(1983, 1984)用线性判别分析方法对某些砂岩—页岩层序进行分类。钱绍新(1992)应用Fisher判别进行储层预测。Huang和Fu(1985a, 1986b)采用句法模式识别技术来辨别亮点。Quine和Kubichek则将模式识别方法结合基于规则的专家系统进行储层分析。

二维模式识别技术现已被成功地应用于油藏描述的许多方面,如储层参数预测、储层含油气性预测、构造识别、岩性识别等。但它也存在着许多弊端:①由于它所利用的属性数据都是从叠后地震资料中提取出来的,因而不能识别出储层的细微变化;②以往的二维模式识别系统大多是沿着某一层开窗抽取地震属性,得到的结果是测线上该解释层位

的含油气性概率预测值;③二维模式识别的样本拾取方式都是先从深度域的测井资料中获得各个解释层位,然后转换成时间域,再投影到相应的地震资料中,并选取模式样本,由于测井资料和地震资料分辨率的差异,造成测井资料深时转换的误差较大,势必影响到样本拾取的准确性。为此,本文提出一种叠前模式识别技术,根据此模式识别的结果可对储层在纵向和横向上的分布进行预测,并用于岩性解释中。

## 2 属性提取

地震属性指的是那些由叠前或叠后的地震数据,经过数学变换而导出的有关地震波的几何形态、运动学特征、动力学特征和统计学特征。不论使用聚类分析或是分类判别,基于叠前地震资料进行模式识别必须遵循如下步骤:首先要进行CDP叠加;然后根据地质资料中的分层数据,综合地震数据的特征,在叠加后的地震剖面上做层位解释;最后将叠后地震剖面上的解释层位投影到相应的叠前地震道集上。

在此基础上进行叠前数据各种属性提取。先选择任意两个解释层位(或具有一定厚度的时间段),并对这两个解释层位之间的每个深度点提取多个属性;然后对每个CDP的所有地震道的相同反射点提取出来的地震属性求取相对于炮检距的变化梯度,或对整个CDP取平均值,这样对于每个CDP就提

\*北京市海淀区学院路中国地质大学(北京)地球物理与信息技术学院,100083  
本文于2007年7月5日收到。

取出来一个随深度变化的属性矩阵。这样一个属性矩阵从矩阵的维数上来说就相当于从叠后地震剖面中提取出来的,所以它在实际应用方面和运算速度方面都便于应用模式识别。这种属性提取方法较之以往的叠后地震属性更能反映储层的细微变化。

我们的目的是预测油气储层的纵、横向连续性,所提取的地震属性分别为:

(1)瞬时地震属性 包括瞬时振幅  $f(t)$ 、瞬时积分振幅  $q(t)$ 、瞬时相位  $\gamma(t)$ 、瞬时频率  $\psi(\omega)$ 、反射强度  $A(t)$ 及振幅加权频率;

(2)单道时窗地震属性 包括最大峰值振幅、最大谷值振幅、复合绝对振幅、第一/二/三峰值谱频率、设定频宽能量、几何分维数、模糊分维数、模糊自相似从属度;

(3)多道时窗地震属性 包括相关长度、平均相关、加权相关、最小相关、最大相关、似相关系数。

### 3 聚类分析算法

鉴于聚类方法的结果比较稳定,所以适合于无井或少井的勘探区和井类型不全、不具代表性的开发区。同时聚类方法不但可以用于分类判别时分类数的确定,其结果还可以用于沉积相分析和岩性解释中。在提取地震属性后,本文采用了两种不同的聚类算法对属性数据分别聚类和比较。

#### 3.1 VQ 算法

VQ算法是一种无监督学习的非神经的矢量量化方法。矢量量化方法开始时并没有分配类别(第一个模式会强迫创建一个类以支持该模式),在遇到每一个新的输入模式时,系统将计算它和任何一个已分配的类别之间的欧氏距离,并选择一个距离最小的类别分配。一旦一个模式确定属于某一类别,最新修改的类别的聚类中心就必须调整。新的聚类中心是通过求得所有成员矢量的平均值形成的<sup>[3]</sup>,即

$$C_k = \frac{1}{N_x} \sum_{x \in S_k} X \quad (1)$$

式中: $k$ 表示现在输入模式所属的聚类中心; $C_k$ 是聚类中心对应的矢量; $N_x$ 表示 $N$ 是 $X$ 的函数,随着矢量 $X$ 的变化,第 $k$ 类的矢量个数 $N$ 是变化的; $S_k$ 为第 $k$ 类的矢量模式集。

#### 3.2 自组织神经网络聚类算法

本文采用了 Kohonen 自组织神经网络算法进

行聚类分析。该网络结构由输入层和 Kohonen 层(即输出层)构成。这两层是全连接的,每个输入层神经元与每个输出层神经元有一前馈连接。假设输入是标准化的(即  $\|x\| = 1$ ),Kohonen 层的输出可用下式

$$I_j = \sum_{i=1}^n (W_{ij} x_i) \quad (2)$$

计算,应用胜者取全部的原则,获胜的输出层神经元将是最大  $I_j$  的神经元。该方法选择一个获胜神经元,该神经元权值矢量与输入矢量的夹角要小于其他所有神经元所对应的点积(图 1)。

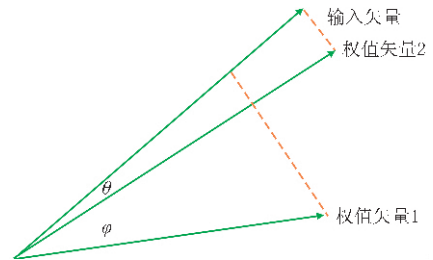


图 1 模式矢量与权值矢量点积

注意模式矢量与最相似的权值矢量之间的夹角最小

### 4 分类判别算法

如果测井资料比较齐全,而且需要确切知道分类类别时,就需要采用分类判别方法。分类判别方法就是除原始资料矩阵之外还知道一批已分类的点,并从这些已知条件出发去寻求某种判别函数或判别标准,然后判断还不知道类别的点应该属于哪一类。因此要达到这一目的,首先要从测井资料中拾取已知样本,并用已知样本学习(即训练识别的概念)得到分类判别函数,然后利用学习的结果去判断大量未知类型的样品属于哪一类。

#### 4.1 样本提取

将深度域的测井资料转换成时间域的地震资料,其过程产生的误差总是很大。如果直接利用这个转换后的测井资料拾取样本,必然要将误差传递给下面的计算过程。为解决上述问题,本文提出了一种新的样本提取方法。该方法充分利用了聚类结果稳定的优势,首先拾取所有井的井旁道,对井旁道的有一定厚度的目的层提取各种属性,再利用聚类方法直接聚类,这样就得到了一组并不确定类别的样本以及分类数。然后利用测井资料指定每个样本的类别,也可以利用声波测井曲线或“拟声波测井曲

线”合成的 VSP 资料修改每个样本值。此法利用了测井资料或地震资料上已确定的目的层顶、底面的深度,同时避免了从测井曲线上读取中间层位深度点再进行深时转换引起的误差。

#### 4.2 模式训练

在模式训练中引入了神经网络技术。采用的神经网络为多层感知网络,学习算法为 BP 算法和 ALOPEX 算法。当学习任务变得复杂时,BP 算法表现得不尽人意,梯度下降法的缺欠被反向传播算法采纳。ALOPEX 是把神经网络学习过程看作最优化问题的随机并行算法。对于用于训练神经网络的 ALOPEX 算法,第  $i$  神经元的网络输入是

$$\text{net}_i = \sum_j \omega_{ij} \text{out}_j + \theta_i \quad (3)$$

其中  $\omega_{ij}$  表示第  $i$  个和第  $j$  个神经元之间的连接权值。假设 sigmoid 激励函数第  $i$  神经元的输出是

$$\text{out}_i = \frac{1}{(1 + \exp(-\text{net}_i/Q_0))} \quad (4)$$

由 ALOPEX 算法确定的网络权值在第  $n$  次递归期间的新值为

$$\omega_{ij}(n) = \omega_{ij}(n-1) + \delta(n) \quad (5)$$

$$\delta_{ij}(n) = \begin{cases} -\delta & \text{概率为 } P_{ij}(n) \\ +\delta & \end{cases} \quad (6)$$

式中  $\delta$  为随机行走步长,是一个小正常数。概率  $P_{ij}(n)$  为

$$P_{ij}(n) = \frac{1}{1 + \exp(\Delta_{ij}(n)/T)} \quad (7)$$

$$\Delta_{ij}(n) = \Delta \omega_{ij}(n) \Delta E(n) \quad (8)$$

其中

$$\Delta \omega_{ij} = [\omega_{ij}(n-1) - \omega_{ij}(n-2)] \quad (9)$$

$$\Delta E(n) = [E(n-1) - E(n-2)] \quad (10)$$

如果  $e_k$  是输出神经元  $k$  的误差,那么总的取自所有

输出神经元的误差  $E$  为

$$E = \sum_k e_k \quad (11)$$

此算法沿着减少误差  $E$  的方向上采取有偏随机行走。步长  $\delta$  是一常量,温度  $T$  决定随机行走的效果。建议  $T$  的初始值选得大一些,然后,随训练过程的进行下降到相关平均值  $\Delta_{ij}$ 。概率的误差总是向总误差  $E$  降低的方向倾斜。

## 5 实例分析

为了检验本文提出的聚类分析和分类判别算法的实际效果,针对不同模式识别的方法对实际的资料分别进行了测试,通过对比不同算法的应用效果,确定本文方法的有效性。

### 5.1 聚类分析实例

本文应用聚类分析的理论和算法对委内瑞拉 Maporal 油田实际数据进行了试算。试验中选择了三个方向的测线(图 2),从图中可看出,测线 3 与断层接近平行,测线 1 和测线 2 与断层相交。

图 3 分别是上述三条测线经动校正后的部分叠前 CDP 道集。由图 3 可以看出,目的层的反射随炮检距的不同有明显的变化,而且三条线的变化规律也不尽相同,其中测线 3 上的振幅变化最为复杂。

对以上三条测线提取地震属性并采用不同的聚类方法计算,得到了不同的聚类结果。首先在计算地震属性时,对每个 CDP 道集的同—反射深度取随炮检距的变化梯度,分别用矢量量化方法和神经网络方法进行聚类,也可以对每个 CDP 道集的地震属性取平均后,用神经网络方法聚类。图 4 就是对测线 1 提取梯度属性后用神经网络的方法聚类得到的聚类结果剖面。将所有聚类结果进行对比分析可看

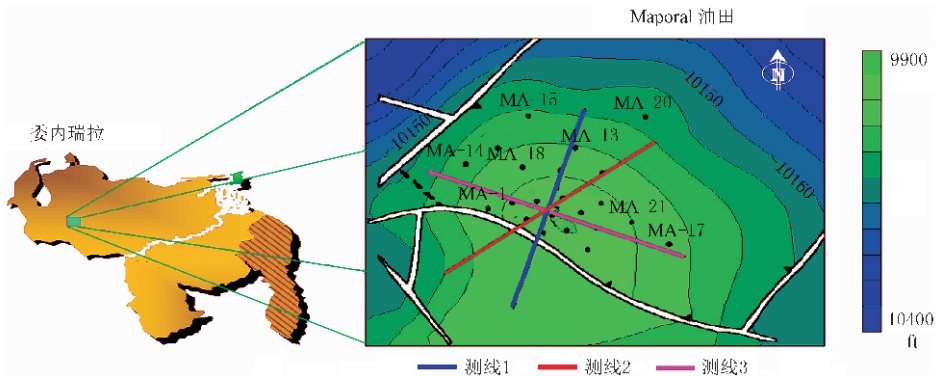


图 2 委内瑞拉的 Maporal 油田位置图

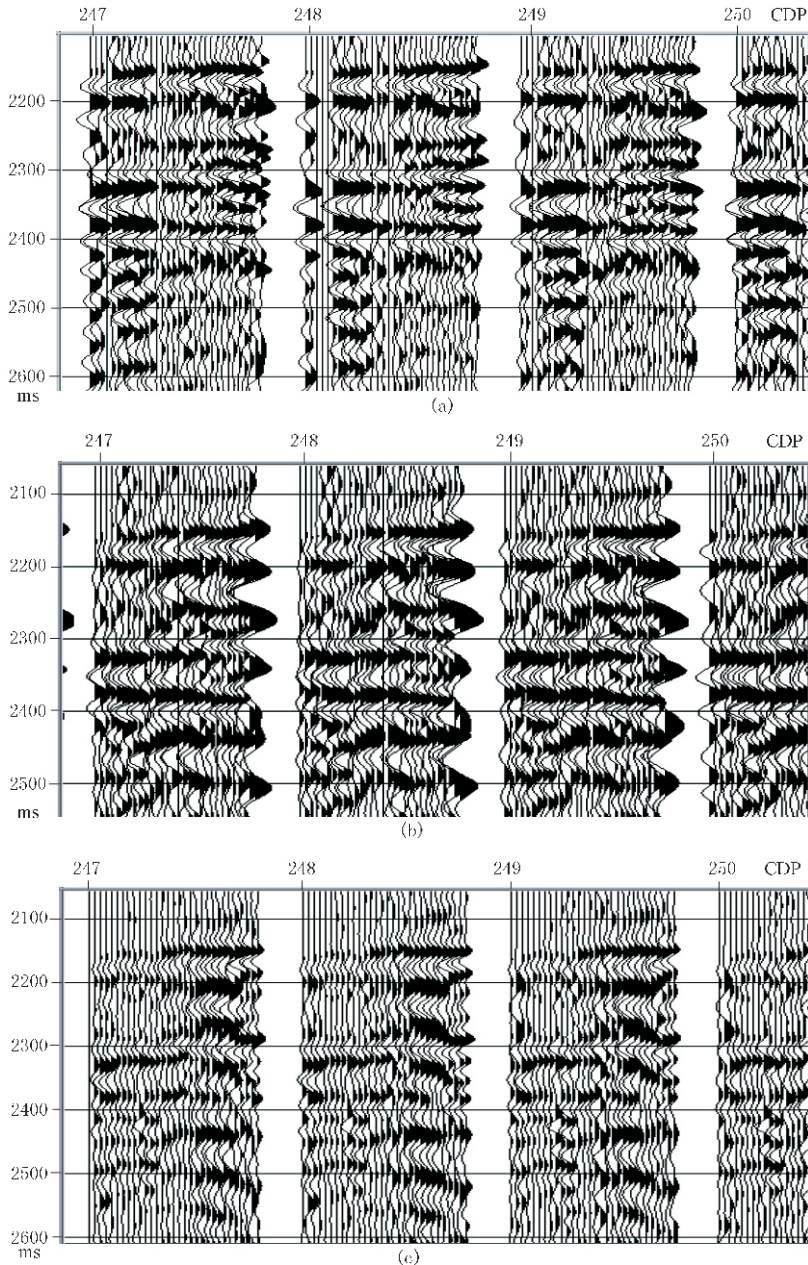


图3 Maporal 油田三条测线的部分叠前 CDP 道集  
(a)测线 1; (b)测线 2; (c)测线 3

出:①用神经网络方法得到的聚类结果在反映储层纵向变化方面要比矢量量化方法的结果理想;②利用 CDP 道集平均属性进行的聚类效果不如将属性对整个 CDP 道集取其随炮检距的变化梯度后再进行聚类,后者更能反映储层在横向上的分布情况;③神经网络方法比矢量量化方法的计算速度要快、更容易分类。对于测井资料缺乏或井资料不好的工区,可以直接应用聚类分析,以克服因测井资料过少而造成的模式识别的局限性。

### 5.2 分类判别实例

针对同一数据,应用分类判别的理论和算法检验本文提出的样本拾取方法以及模式训练算法的实际应用效果。首先选择井旁道并拾取样本,然后根据测井资料修改样本值,再选择合适的神经网络进行模式训练得到判别标准,最后对整条测线进行分类。图 5 就是对测线 3 采用了以上各个应用过程后得出的识别结果剖面图,由图可看出,该模式识别结果预测了砂体的分布和连续性情况。

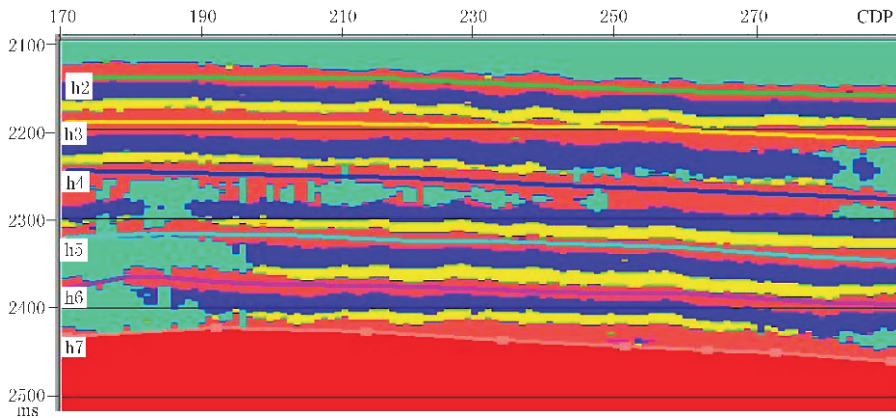


图 4 用神经网络方法对测线 1 的聚类结果剖面

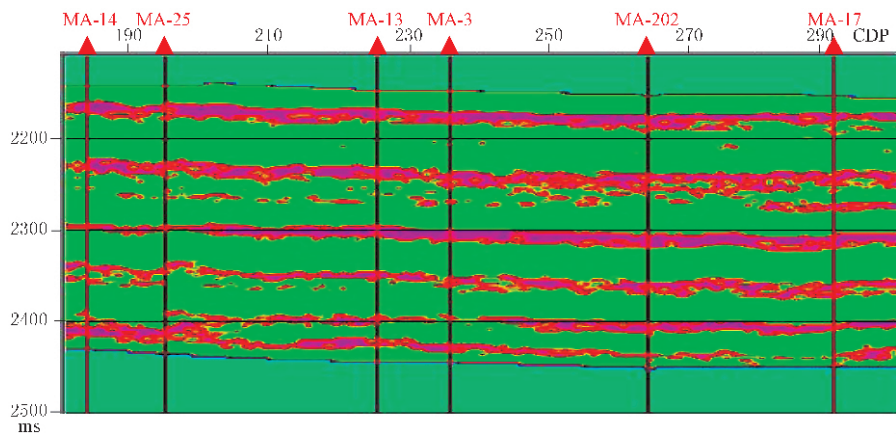


图 5 应用 BP 网络对测线 3 的分类判别结果

红色代表物性好的砂岩,黄色代表物性差的砂岩,绿色表示泥岩

从分类判别方法的应用效果来看,对砂体分布的预测结果比较成功,这是因为加入了井约束条件,而在利用测井资料修改样本值时,忽略了储层岩性在纵向上的微小变化,只考虑比较大的岩性变化(将许多聚类样本合并了)。

### 6 结论

(1)本文所采用的属性提取方法与传统的叠后属性相比,前者对储层岩性的纵、横向变化的分辨率更高;

(2)利用本文提出的聚类分析和分类判别算法,通过实例分析预测储层纵、横向分布,取得了较好的效果,但也存在局限和不足;

(3)由于储层的横向变化特征不明显,需要进一

步寻求更多、更能反映储层连续性的地震属性;

(4)应充分发挥聚类分析和分类判别成果的应用价值,例如将二维叠前模式识别结果应用于沉积相分析等,使之能更好地为地质分析服务。

### 参 考 文 献

- [1] 王碧泉,陈祖荫. 模式识别理论、方法与应用. 北京:地震出版社,1989
- [2] 边肇祺. 模式识别. 北京:清华大学出版社,1988
- [3] Abhijit S et al; 徐勇,荆涛译. 神经网络模式识别及其实现. 北京:电子工业出版社,1999
- [4] 马在田等. 计算地球物理学概论. 上海:同济大学出版社,1997
- [5] 刘雯林. 油气田开发地震技术. 北京:石油工业出版社,1996
- [6] 焦李成. 神经网络的应用与实现. 西安:西安电子科技大学出版社,1993

(本文编辑:张亚中)