

文章编号: 1004 — 5589 (2007) 03 — 0381 — 04

利用模式识别技术识别致密砂岩含气层

程洪亮, 范晓敏, 董建刚

吉林大学 地球探测科学与技术学院, 长春 130061

摘要: 准确识别致密砂岩含气层一直是测井解释工作所面临的一个难题, 为准确识别致密砂岩含气储层, 根据各种测井的响应特征, 运用计算机模式识别技术进行气层综合识别。通过对实际资料处理与气层测试对比, 气层识别的准确率可达到 91%, 储层划分的纵向误差均 $< 0.450\text{ m}$ 。

关键词: 模式识别; 特征提取; 致密砂岩

中图分类号: P618.13; TP317

文献标识码: A

Distinguishing gas-bearing formation from compact sandstone with pattern recognition

CHENG Hong-liang, FAN Xiao-min, DONG Jian-gang

College of Ge exploration Science & Technology, Jilin University, Changchun 130026, China

Abstract: It is a difficult to solve the problem in log data interpretation to distinguish accurately gas-bearing formation. Based on the character of log data response, the authors used the computer pattern recognition technique to identify it synthetically. By comparison with processing of actual data and the gas-bearing test, it can approach to 91% in accuracy rate for distinguishing gas-bearing formation, with the lengthwise error of reservoir zonation less than 0.450 m .

Key words: pattern recognition; feature extraction; compact sandstone

0 引言

致密砂岩储层, 埋藏深度较大, 在压实作用下储集性变差。致密砂岩储层一般具有孔隙度低 ($< 10\%$)、渗透率低 ($< 0.5 \times 10^{-3}\text{ }\mu\text{m}^2$)、含气饱和度低而含水饱和度高、毛管压力高、天然气在其中流动速度较为缓慢、储层在测井资料上的响应不明显的点^[1]。这类储层是中国常见的含气储层, 目前在中国新探明的天然气储量中近一半产自这种储层^[2]。常规砂岩气层在测井资料上有明显的反映, 与含水及含油地层相比, 气层的声波孔隙度、密度孔隙度增大而中子孔隙度减小, 识别效果好。因为致密砂岩储层的孔、渗等性能非常差, 所以气

层识别较常规砂岩气层困难, 许多方法的识别效果都不理想^[3]。为此, 需要采用新的气层识别方法来提高致密砂岩气层识别的准确率。笔者根据自然伽玛、中子、声波、密度、自然电位和双侧向等常规测井资料对气层的响应特征, 采用模式识别技术综合识别致密砂岩含气层。

随着人工智能的兴起, 模式识别技术在 20 世纪 60 年代初迅速发展成一门学科, 并在很多领域得到了成功的应用。80 年代起人工智能中的模式识别技术和人工神经网络技术开始大量运用到地质勘探与开发中, 并取得了显著的成果^[4]。人工神经网络技术已经被广泛运用到储层识别中, 但在综合识别中要求输入的样本维数高, 网络结构复杂,

为使网络具有推广性必需有大量的样本数据,而且网络计算时间长^[5]。基于综合识别数据高维的特点,采用人工智能中的有监督模式识别技术进行致密砂岩气层的综合识别,将会取得显著效果。

1 模式识别原理

模式识别技术主要有两种基本方法,统计模式识别方法和结构(句法)模式识别方法,与此相对应的模式识别系统都由设计和实现两个过程组成。设计是用一定数量的样本(学习集)进行分类器的设计,其基本思想:给定某个判别函数类,利用样本集确定出判别函数中的未知参数。实现是用所设计的分类器对待识别的样本进行分类决策^[6]。本文采用有监督统计模式识别方法进行致密砂岩气层的识别(图1)。

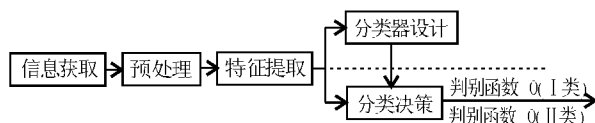


图1 模式识别方法原理图

Fig 1 Schematic figure of pattern recognition

通过预处理和特征提取,从每个样本得到一个维特征矢量,可将每一个样本看成是 n 维空间的一个点。如果此特征矢量能反映致密砂岩含气与否,则反映含气与不含气的特征矢量点在 n 维空间中应分布在不同的位置上。设计分类器是要找到一个超平面,使含气类矢量点和非含气类矢量点分别位于该超平面的两侧,对未知样本分类是计算其特征矢量相对这个超平面的位置。在分类器的实际设计过程中根据测井资料样本数据的特点,采用广义线性分类器,即广义线性判别函数。

2 广义线性判别函数

从每个样本(测井资料)中提出的维特征矢量,记为 $X = (x_1, x_2, x_3, \dots, x_n)$,第 i 个样本记为 $X_i = (x_{1i}, x_{2i}, x_{3i}, \dots, x_{ni})$ 。广义线性判别函数 $g(X_i)$ 是 X_i 的线性函数,即

$$g(X_i) = w_0 + w_1 x_{1i} + w_2 x_{2i} + \dots + w_n x_{ni} \quad (1)$$

式中: $w_0, w_1 \dots w_n$ 为每项权系数,由设计分类器确定。表示成矩阵的形式 $g(X_i) = W^T X_i$

其中:

$$W = \begin{pmatrix} w_0 \\ w_1 \\ \vdots \\ w_n \end{pmatrix}, \quad X_i = \begin{pmatrix} 1 \\ x_{1i} \\ \vdots \\ x_{ni} \end{pmatrix} \quad (2)$$

设有已知类别的 m 个样本 X_1, X_2, \dots, X_m , $g(X)$ 是期望的判别函数,则对 m 个已知样本,都有 $g(X_i) = W^T X_i = g_i (i = 1, 2, \dots, m)$,矩阵形式为 $X^T W = B$,其中 X 和 B 分别为

$$X = \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_{11} & x_{12} & \dots & x_{1m} \\ \vdots & \vdots & \dots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{pmatrix}, \quad B = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_m \end{pmatrix} \quad (3)$$

组成 B 矢量的各分量是期望的已知样本的判别函数值。反映含气和非含气样本分别位于判别超平面的两侧,即有不同的符号。令

$$g_i = \begin{cases} 1 & \text{样本 } X_i \text{ 是 I 类} \\ -1 & \text{样本 } X_i \text{ 是 II 类} \end{cases} \quad (4)$$

通常样本数大于特征矢量维数,因此矩阵 X^T 是一个列满秩阵,方程 $X^T W = B$ 没有精确解,只有最优解。定义一个平方误差准则函数 $J_s(W)$ 为

$$J_s(W) = \|X^T W - B\|^2 \quad (5)$$

求解一个使 $J_s(W)$ 极小化的 W 作为问题的解,这是矛盾方程组的最小二乘近似解,称为伪解或 MSE 解,记作 W^* ^[7]。解析法求出的 MSE 解为

$$W^* = (X X^T)^{-1} X^T B \quad (6)$$

3 测井资料特征提取

综合气层识别选择的各测井资料对致密砂岩含气层均有反映,但各种测井资料对含气层的反映程度不同,即各种测井资料对气层识别的贡献大小不同。特征提取是根据测井资料对气层识别贡献的大小,对原始数据进行变换,得到最能反映含气层识别本质的一个或几个综合本质特征,来实现有效的含气层识别。测井资料对气层识别贡献大小就是分类识别过程中的类别可分离性,数据变换要使变

换后的综合本质特征数据可分离性最大化。在模式识别中有多种准则（判据）来衡量数据的可分离性，如离散度、概率分布和熵函数等。致密砂岩含气层识别的特征提取基于离散度判据，期望变换后的数据类间离散度尽量大，而类内离散度尽量小，定义判据函数

$$J = tr(S_w^{-1}S_b)$$
 (7)

式中： S_b 为类间离散度矩阵， S_w 为类内离散度矩阵^[8]。

用离散度判据进行测井数据特征提取的步骤为：

气层识别采用的原始特征数据： $y = (GR, CNL, AC, DEN, SP, LLD, LLS)^T$ ，通过变换映射为 n ($n < 7$) 个特征： $X = (X_1, x_2, \dots, x_n)^T$ ，其变换关系为

$$X = A^T y$$
 (8)

令 S_b, S_w 为原始特征数据离散度矩阵； S_b^*, S_w^* 为映射后特征数据离散度矩阵，则

$$S_b^* = A^T S_b A, S_w^* = A^T S_w A$$
 (9)

经变换后的判据函数 J 变为：

$$J(A) = tr((A^T S_w A)^{-1} A^T S_b A)$$
 (10)

特征提取问题转化为求取使 $J(A)$ 取得最大值的变换 A 。经过计算求出使判据函数 J 达到最大的变换 A 为：

设矩阵 $S_w^{-1}S_b$ 的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_7$ ，按大小排列为： $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_7$ ，对应的特征向量为 u_1, u_2, \dots, u_7 ，则选前 n 个特征值对应的特征向量作为 A ，即 $A = (u_1, u_2, \dots, u_n)$ 。

图 2 为 A 井 1 340. 000~ 1 380. 000 m 段测井资料特征提取成果图，为与原始测井资料进行对比仅提取出 1 个 ($n = 1$) 综合本质特征。通过测井曲线与提取特征曲线对比，发现图中提取特征曲线的储层识别能力较任何单一测井资料识别能力都强，在测井曲线上有含气响应的井段，提取特征曲线的突变加强，在其他井段曲线起伏减小。但要定量识别出含气层，必须通过广义线性判别函数定量计算。

4 实际应用及效果

测井数据标准化是使已知样本数据与待识别的测井数据在研究区域内有一个统一的刻度标准，使可能影响原始测井数据的各种系统误差减至最小，以消除非地质因素造成的测井响应的井间差异；通过归一化，可以克服不同测井数据以及非测井数据由于量纲不同而造成的差异。还能在很大程度上减小非地质因素导致的井间差异的影响。这是使用模式识别技术处理气层识别实际问题的关键。用标准化与归一化后的已知样本数据进行分类器设计。对标准化与归一化后的测井资料作气层识别。

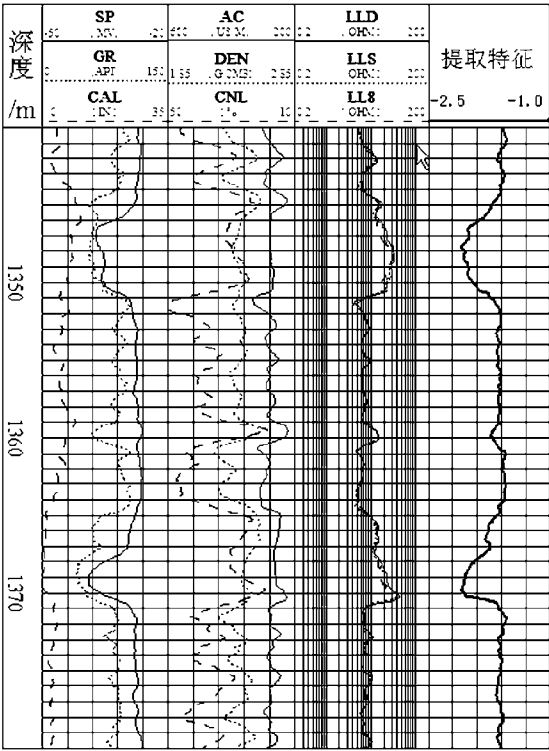


图 2 A 井测井资料特征提取成果图

Fig 2 Result of feature extraction from log data of well A

利用研究区域两已测试井的测井资料设计分类器，对该区域的多口井目的层作气层识别与测井解释、测试结论对比分析，其吻合率较常规识别方法有显著的提高。图 3 为 B 井 1 310. 000~ 1 350. 000 m 段直接利用原始测井资料综合识别的气层结论，图中右边第一道为测井解释和测试的气层结论，5 号层 1 318. 000~ 1 322. 000 m，6 号层 1 345. 500~ 1. 350. 500 m，两层均为气层。图中右边第二道为测

井资料综合气层识别结论, 5号层 1 317.750~1 322.000 m, 6号层 1 345.875~1 350.625 m均与测井解释和测试相符。通过对5口井的对比结果统计, 表明利用模式识别技术作气层识别完全能满足快速直观气层识别的要求。5口井综合测井解和测试划分了34个含气层, 利用测井资料作气层识别, 其结果为31个含气层符合率为91%, 纵向误差均<0.450 m。常用的气层识别方法一般采用一种或几种测井资料的相对幅度大小或计算的含气指数的相对幅度来划分含气层, 而运用模式识别技术的综合气层识别方法实现了气层识别的定量化。如图3所示零值线作为气层临界值, 当广义线性判别函数计算的结果大于临界值时在处理过程中程序将自动识别为气层, 不需要根据全井段的相对幅度变化来识别。

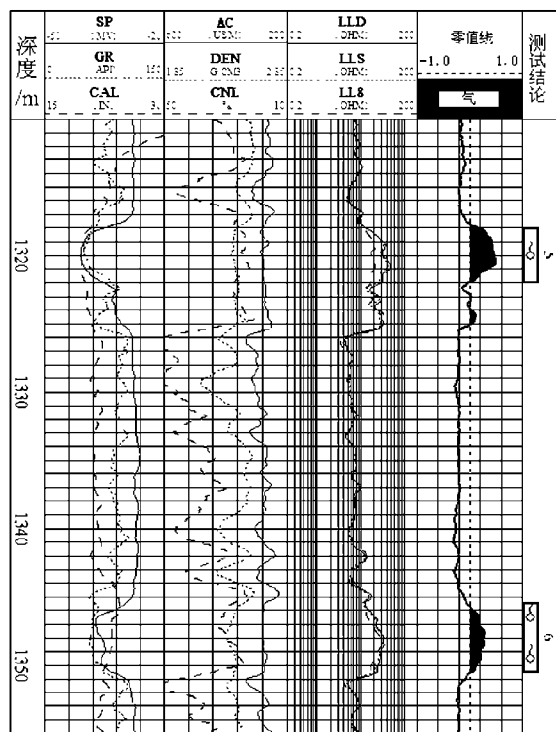


图3 B井气层识别成果图

Fig 3 Result for distinguishing gas formation of well B

5 结论

(1) 模式识别技术在致密砂岩气层识别方面的运用克服了常规多参数识别的缺陷, 不用选择参数, 消除了参数误差引起的识别错误, 而且气层识别的灵敏度有明显提高、准确率可达91%;

(2) 模式识别技术在致密砂岩气层识别方面的

运用, 实现了气层识别的定量化, 使储层划分的纵向误差均<0.450 m;

(3) 测井资料的标准化与归一化是提高气层识别准确率的关键。

参考文献:

- [1] 董瑞霞, 范晓敏. 致密砂岩气层综合识别方法的改进[J]. 世界地质, 2003, 22 (3): 266-270.
DONG Rui xia, FAN Xiaomin. Improvement of method for distinguishing formation containing gas from compact sandstone [J]. *Global Geology*, 2003, 22 (3): 266-270.
- [2] 赵彦超, 吴春萍, 吴东平. 致密砂岩气层的测井评价[J]. 地质科技情报, 2003, 22 (4): 65-70.
ZHAO Yan chao, WU Chun ping, WU Dong ping. Logging evaluation to tight gas sandstone [J]. *Geological Science and Technology Information*, 2003, 22 (4): 65-70.
- [3] 李云省, 曾渊奇, 田建波. 致密砂岩气层识别方法研究[J]. 西南石油学院学报, 2003, 25 (1): 26-28.
LI Yun sheng, ZENG Yuan qi, TIAN Jian bo. A study on the recognition method of tight sandstone gas layers [J]. *Journal of Southwest Petroleum Institute*, 2003, 25 (1): 26-28.
- [4] 罗利, 姚声贤, 任兴国. 神经网络及模式识别技术在测井解释中的应用[J]. 测井技术, 2002, 26 (5): 364-368.
LUO Li, YAO Sheng xian, REN Xing guo. Application of numerical calculation and model identification technology based on neural network in log interpretation [J]. *Well Logging Technology*, 2002, 26 (5): 364-368.
- [5] Hagan M T, Demuth H B, Beale M H. 神经网络设计[M]. 戴葵, 译. 北京: 机械工业出版社, 2002: 227-335.
Hagan M T, Demuth H B, Beale M H. Neural network design [M]. DAI Kui, translated. Beijing: China Machine Press, 2002: 227-335.
- [6] 边肇祺, 张学工. 模式识别[M]. 北京: 清华大学出版社, 2000: 2-3.
BIAN Zhao qi, ZHANG Xue gong. Pattern recognition [M]. Beijing: Tsinghua University Press, 2000: 2-3.
- [7] 王新民, 董小刚. 工程中的数值方法[M]. 长春: 吉林人民出版社, 2001: 48-53.
WANG Xin min, DONG Xiao gang. Numerical method of engineering [M]. Changchun: Jilin People's Press, 2001: 48-53.
- [8] 鞠正卫, 陈文苗. 概率与数理统计[M]. 哈尔滨: 黑龙江科技出版社, 1983: 31-55.
JU Zheng wei, CHEN Wen miao. Probability and numerical statistics [M]. Harbin: Heilongjiang Science and Technology Press, 1983: 31-55.