

# 基于粗糙集和 BP 神经网络的天然裂缝识别

张 龙, 吴 江, 张德同

(西北大学 信息科学与技术学院, 陕西 西安 710127)

**摘 要:**研究粗糙集理论和 BP 神经网络算法, 以及如何结合两者构建天然裂缝智能识别的应用。选取与天然裂缝发育有关的测井曲线作为样本集, 先用粗糙集理论约简样本信息, 然后采用 BP 神经网络识别天然裂缝。给出了基于粗糙集和 BP 神经网络的天然裂缝智能识别的设计方案, 通过长庆油田实际油井的应用, 结果表明这种识别方法效果显著。粗糙集有效地约简样本信息, BP 神经网络很好地识别天然裂缝, 故基于粗糙集和神经网络的智能识别方法切实可行。

**关键词:**天然裂缝识别; 粗糙集; BP 神经网络; 智能识别

**中图分类号:** TP183

**文献标识码:** A

**文章编号:** 1673-629X(2008)11-0041-03

## A Method for Natural Fractures Identification Based on Rough Sets Theory and BP Neural Network

ZHANG Long, WU Jiang, ZHANG De-tong

(College of Information Science and Technology, Northwest University, Xi'an 710127, China)

**Abstract:** Investigate the rough sets theory and BP neural network algorithm, and how to combine the two that for constructing a natural fractures intelligence identification application. Choose the logging curve related to natural fractures development to collect as the sample, first, the rough sets theory is applied in sample-information reduction, then the BP neural network is applied to identify the natural fractures. Puts forward a design plan that the natural fractures intelligence identification based on rough set and BP neural network. Actual application example in the Changqing oil-field shows the method is practicable, and the interpretation effect is satisfactory. Rough set is simple to invite simple sample information, natural fractures of identification that BP neural network can be very good. The method of intelligence identification based on rough set and BP neural network is feasible.

**Key words:** natural fractures identification; rough sets; BP neural network; intelligent identification

## 0 引 言

天然岩石裂缝普遍存在于各类岩体中, 了解裂缝的性质对于地质、地球物理、采矿、石油工程、水文地质以及核废料隔离都具有重要意义<sup>[1]</sup>。但是常规的各种识别裂缝的方法均存在局限性, 从而造成识别裂缝的主观不确定性及多解性。成像高新测井技术优势明显, 但毕竟其代价过高, 只能在少数井中测量, 并且测量时具有一定的局限性, 其中最突出的一点是大多数成像测井仪探测深度都较浅, 受井壁状况影响较大, 当探测深度较大时, 分辨率大大降低, 图像变的模糊<sup>[2]</sup>。

测量天然裂缝的数据, 主要为岩芯和各类井下成

像测井。当缺少岩芯和井下成像数据时, 也可使用常规测井曲线数据<sup>[2]</sup>。天然裂缝识别的关键是选取最能反映裂缝发育的测井曲线参数值, 为此文中根据常规测井曲线数据, 采用粗糙集和人工神经网络相结合的方法识别天然裂缝。此方法的特点是避开事先建立的测井解释模型, 通过其本身算法的结果来自动确定裂缝带与各测井响应或特征参数之间的复杂关系。为了降低神经网络构成系统的复杂性以及减少网络的训练时间, 采用粗糙集方法优选输入神经网络学习的样本信息。该方法在成像测井资料及岩芯样品有限的情况下, 对裂缝识别评价具有极其重要的意义。

## 1 基于粗糙集和 BP 神经网络的智能识别方法

采用基于粗糙集与神经网络的智能处理方法进行天然裂缝识别的创新点是: 采用单隐层的 BP 神经网络作为信息识别系统; 使用粗糙集方法对神经网络的学

收稿日期: 2008-02-24

基金项目: 陕西省重大科技创新项目(2005ZKC(二)02-11)

作者简介: 张 龙(1983-), 男, 硕士研究生, 研究方向为计算智能、数据挖掘系统; 吴 江, 副教授, 硕士生导师, 研究方向为知识库与数据库、人工智能和基于网络应用系统; 张德同, 副教授, 硕士生导师, 研究方向为智能信息系统、算法分析与设计。

习训练样本集进行信息约简,优选出能全面反映天然裂缝的各种属性。这种信息约简方法是其它方法无法比拟的。

### 1.1 信息约简的粗糙集方法

粗糙集理论是 Zdzislaw Pawlak 于 20 世纪 80 年代提出的关于模糊和不确定关系的数学工具<sup>[3,4]</sup>。它把无法确认归属于某个确定的类的对象都归属于边界线区域,定义为上近似集和下近似集之差集,能有效地分析和处理不精确、不一致、不完整等各种不完备信息,并从中发现隐含的知识,揭示潜在的规律。它不仅能对知识系统进行数据约简,从决策表中导出决策规则,而且能分析属性间的依赖关系,并可对导出的规则进行评价,是处理不确定、不精确、不完善系统的一种有效方法,已被广泛用于专家系统、决策支持系统、机器学习、归纳推理、模式识别等领域。其中信息约简是一个重要的研究内容。

信息约简原理<sup>[4]</sup>:设  $K = (U, R)$  为一信息系统,其中  $U = \{u_1, u_2, \dots, u_n\}$  为有限非空的个体全集,  $R = \{r_1, r_2, \dots, r_m\}$  为有限的属性集。

设  $r \in R$  为  $U$  上的一个属性,全集  $U$  上的元素关于属性  $r$  的等价类可简记为  $\{x\}$ , 其中  $x \in U$ 。

设  $P \subseteq R, P \neq \emptyset$  (空集),  $P = \{r_{i1}, r_{i2}, \dots, r_{ik}\}$ , 则  $\cap P$  是一等价关系,简记为  $\text{ind}(P)$ , 即不可分辨关系。

设  $r \in R$ , 如果  $\text{ind}(R) = \text{ind}(R - \{r\})$ , 则称  $r$  是  $R$  中的一个可约简的属性,否则称是不可约简的。

如对  $r \in R$  都是不可约简的,则称集合  $P$  是独立的。若  $Q \subseteq P$  是独立的,且  $\text{ind}(Q) = \text{ind}(P)$ , 则称  $Q$  是  $P$  的一个简化。 $P$  中所有不可省属性的集合称为  $P$  的核,记成  $\text{core}(P)$ 。

归纳起来,信息约简的基本步骤如下:

(1) 信息表中属性的离散化和泛化,即对具有连续性属性的离散化和针对名词性属性的面向属性的归纳,连续属性的离散化,可以由领域专家根据经验给出相应的区间,也可以由系统根据某种原则自动地对输入空间进行划分;

(2) 从信息表中去除冗余对象,即值约简;

(3) 从信息表中去除冗余属性,求得独立属性的最小子集,得到约简的属性集。

### 1.2 BP 神经网络原理

BP(back propagation)算法全称“误差逆传播算法”,是一种具有二层或三层以上的阶层型神经网络。上、下层之间各神经元实现全连,BP 下层的每一个单元与上层每一个单元都实现连接,而每层的神经元之间无连接。

BP 算法的基本思想是<sup>[5,6]</sup>,学习过程由信号的正向传播与误差的反向传播两个过程组成。正向传播时,输入样本从输入层传入,经各隐层逐层处理后,传向输出层。若输出层的实际输出与期望输出(教师信号)不符,则转入误差的反向传播阶段。这种信号正向传播与误差反向传播的各层权值调整过程一直进行到网络输出的误差减少到设定阈值,或到设定的学习次数为止。图 1 所示的单隐层网络的应用最为普遍。

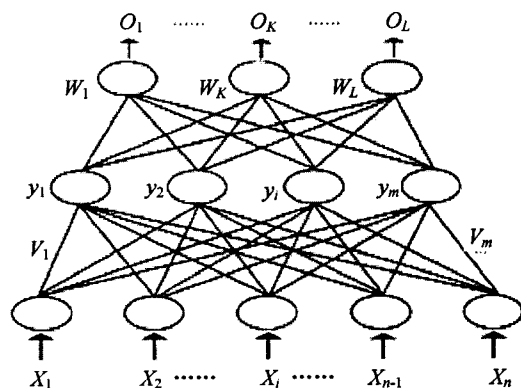


图 1 三层 BP 网络结构

BP 网络的数学模型:

(1) 输入层节点  $i$  的输出为  $X'_i = f(x_i)$ , 其中  $x_i$  为模型的输入。为了使 BP 模型隐层的输入规范化, 函数  $f$  为一规整函数, 即  $f = (x - x^{(1)}) / (x^{(2)} - x^{(1)})$ 。其中,  $x^{(1)}$  为节点输入的下限,  $x^{(2)}$  为节点输入的上限。

(2) 隐层节点  $j$  的输出为  $y_j = f(I_j)$ , 其中  $I_j = \sum w_{ij} X'_i - \theta_j$

其中,  $I_j$  为隐层节点  $j$  的输入加权, 且  $w_{ij}$  为输入层节点  $i$  到隐层节点  $j$  间的权值, 在算法实现由系统生成随机数, 目前有人研究由神经网络学习生成较为科学的权值。 $\theta_j$  为隐层节点  $j$  的阈值, 在本方法中由系统生成随机数。在 BP 网络中函数通常取拟合精度较高的 S 型函数。即  $y_j = 1 / (1 + e^{-\sum w_{ij} X'_i - \theta_j})$ 。

(3) 输出层节点  $k$  的输出为:

$$O_k = \sum w_{jk} y_j - \theta_k$$

其中,  $w_{jk}$  为隐层节点  $j$  到输出层节点  $k$  之间的权值,  $\theta_k$  为输出层节点的阈值。

为了提高学习速度, 使用了改进的加入动态学习因子的 BP 算法。

### 1.3 天然裂缝识别系统流程

基于粗糙集与神经网络的天然裂缝识别系统流程图如图 2 所示。

(1) 样本信息的选取与预处理: 天然裂缝在大多数常规测井曲线上都有反映, 可使用几种与井径变形、泥浆漏失有关的测井曲线做为样本集。这些曲线包括纵波、横波、伽玛、密度、井径变形、浅层电阻率、深层电阻

率。样本信息的预处理包括数据校正及归一化处理。

(2)粗糙集约简样本信息:将输入的样本信息表中去除冗余对象,再去除冗余属性,求得最简决策表。从而得到反映天然裂缝特性的条件属性。

(3)网络学习训练:将通过信息约简后所得的样本集进行学习训练,直到满足精度要求或循环次数为止,得出权值和阈值,并存储之。

(4)神经网络识别与输出:根据(2)中求得的反映天然裂缝的条件属性,相应地优选和计算待识别的岩层信息有关属性,其中包括信息的预处理和计算有关特性参数。再使用训练好的神经网络对之进行识别或分类,并输出结果。

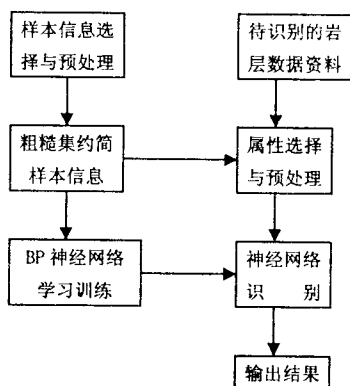


图 2 天然裂缝识别系统流程

## 2 应用实例

文中采用长庆油田的某井作为研究对象,选择 30 个典型岩层的数据做样本集,条件属性为  $C = \{r_1, r_2, r_3, r_4, r_5, r_6, r_7\}$ ,  $r_1$  至  $r_7$  分别对应纵波、横波、伽玛、密度、井径变形、浅层电阻率、深层电阻率。决策属性为  $D = \{\text{裂缝发育程度}\}$ 。这样就构成了一个决策信息表。

首先要将决策表中的属性进行离散化和泛化,即决策属性用数表示和条件属性量化。令决策属性  $D = \{d\}$ ,  $d = 0$  表示裂缝不发育,  $d = 1$  表示发育。测井曲线数据的离散化由裂缝领域专家根据经验给出相应的区间,利用此离散区间对条件属性  $C$  离散成对应决策层 0、1 的数值。

消除相同行的样本,即进行值约简后得到 14 个样本,如表 1 所示。

按照如下步骤对处理后的决策表进行约简:

- 1) 删除表中相同的规则;
- 2) 消除表中多余的条件属性;
- 3) 简化决策规则,即删除分类规则中多余的属性,只保留必要的属性;
- 4) 如有必要,则进一步进行决策算法最小化,从

简化的决策表中去除冗余决策项;

5) 最后利用得到的决策规则进行决策活动。

表 1 值约简后的样本信息决策表

样本号	$r_1$	$r_2$	$r_3$	$r_4$	$r_5$	$r_6$	$r_7$	$D$
1	0	0	0	0	1	0	1	1
2	0	0	0	0	0	1	1	0
3	0	0	0	0	0	0	0	1
4	1	1	1	2	0	1	1	1
5	1	0	1	2	0	1	2	1
6	0	0	0	2	0	0	1	0
7	3	0	0	1	0	0	2	1
8	3	1	0	1	2	0	2	0
9	0	3	2	1	2	0	3	1
10	3	0	1	1	0	1	2	1
11	1	2	2	1	1	1	3	0
12	0	3	3	0	2	1	3	0
13	2	0	1	2	1	2	2	0
14	1	1	2	2	1	1	0	1

由于表中已不存在相同的规则,因此可从第二步开始。从表中去掉一个条件属性,观察决策表中相同的规则是否会导致不同的决策,如相同规则对应着不同的决策,决策表不协调,则该属性不可删除。

在表 1 中删除条件属性  $r_1$ ,可见表仍是协调的,即不存在相互矛盾的决策规则。同理,删除属性  $r_2, r_3$  后,决策表依旧保持协调。

因此,  $\{r_4, r_5, r_6, r_7\}$  是最简约简,即决策表中的条件属性可由原来的 7 个变为 4 个。至此,采用粗糙集方法对样本集进行了包括值和条件属性的信息约简,得到了由密度、井径变形、浅层电阻率和深层电阻率这 4 个属性的 14 个层位样点组成的样本集,即最简样本集。

以密度、井径变形、浅层电阻率和深层电阻率作为输入,以裂缝发育程度作为输出,构建 BP 神经网络模型,对于隐含层的节点数目的确定是利用检验数据通过 BP 算法及作用函数来实现,其实质是通过误差的反向传递和连续的修正完成训练,直至期望值与实际值之间的误差最小;一般无严格要求下可取输入层神经元数的 1~1.5 倍,文中取 6。用此神经网络结构对 12 个样本信息进行学习训练,再用训练好的网络对另外 13 个典型层位进行裂缝识别,具体识别结果如表 2 所示,其中成像显示作为裂缝发育评价标准。

由表 2 可知,13 个验证样本经 BP 网络预测后,有 2 个样本类型与给定类型不一致,准确率达到了 84.62%。这说明基于粗糙集和神经网络的智能识别方法的应用效果理想,可以用于该油田其他井裂缝的识别。

(下转第 46 页)

同的操作的。窗体有两种类型,即 WindowType-Common 和 WindowType- Main, 对于 WindowType- Main 型的窗体,会在窗体初始化的时候,为该系列 window 创建一个 eventbuffer。而对于 WindowType-Common 型的窗体,它与同 environment 的主窗体共用一个 eventbuffer。一个应用可能有多个窗体,其中只有一个主窗体,只要这些窗体具有相同的 enviroment,所有相关的事件统一进入主窗口创建的 eventbuffer。

### 3.4 窗口绘制的实现

Window 对应的内存可绘制区域是由 surface 管理的,通过类 buffer 主要用来记录一些层的信息,用于 surface 的绘制。在 layer 初始化的时候,会创建 windowstack,并且在 windowstack 中记录了 primary-surface,primarysurface 是记录显存地址的一块特殊的 surface。当窗口内容改变时,引擎修改窗口的 surface 相应内存地址的内容,经过窗口的一系列处理,最终通过 Blit 操作使更新后的内容显示在屏幕上。

上层的一些绘图操作实际上是将颜色信息写入 window 对应的 surface buffer,这只是在内存中进行操作,没有反映在屏幕上。经过一系列判断和处理之后,此时内存中已经是应该显示的结果,只需要将内存中记录的颜色信息拷贝到显存中即可。primarysurface 的 buffer 中记录的就是当前 layer 在显存中的首址,Enable Overlay1 时需要修改 primarysurface 的 buffer 中的值,将 buffer 中记录的首址改为驱动为 Overlay1 分

配的显存首址。确定了显存的地址后,开始拷贝操作,这里涉及到格式转换的问题。不同格式之间的拷贝要注意格式转换。

## 4 结束语

基于此图形引擎的 Vinci 系统实现了具有高效的事件处理机制和面向构件的体系结构,这样大大提高了图形程序与用户交互效率,同时它为用户灵活地实现多种图形界面提供了保障。

### 参考文献:

- [1] Koretide Website[EB/OL]. 2008-03. <http://www.koretide.com.cn/>.
- [2] 陈 榕,刘艺平.基于构件、中间件的因特网操作系统及跨操作系统的构件与中间件运行平台[R].上海:上海科泰世纪科技有限公司,2003.
- [3] 上海科泰世纪科技有限公司. CAR 构件大全[EB/OL]. 2008-01-01. <http://www.elastos.com.cn/>.
- [4] Elastos, Inc. Elastos 2.1 Information Repository[EB/OL]. 2008-03. <http://www.koretide.com/download.php>.
- [5] Chen Rong. The Application of Middleware Technology in Embedded OS[C]//Workshop on Embedded System, in Conjunction with the ICYCS(6th). Hangzhou: [s. n.]. 2001: 1-3.
- [6] Yuan Feng. Windows 图形编程[M]. 北京:机械工业出版社,2002.

(上接第 43 页)

表 2 某井天然裂缝识别结果

序号	深度(米)	神经网络输出结果	成像显示	文中方法
1	1507.25	0.9981	1	1
2	1513.00	0.0020	0	0
3	1525.12	0.9976	1	1
4	1551.37	0.3540	1	0
5	1560.5	0.9876	1	1
6	1594.37	0.0040	0	0
7	1818.87	0.0120	0	0
8	1842.75	0.9021	1	1
9	1869.37	0.0325	0	0
10	1876.87	0.9640	1	1
11	1880.25	0.5637	1	0
12	2010.56	0.0011	0	0
13	2204.32	0.0010	0	0

## 3 结束语

采用粗糙集方法能够挖掘出反映天然裂缝特征的有用的常规测井曲线属性,这样可以得到最简样本集;

再采用具有容错性、抗干扰性和自适应性等优点的神经网络作为后置的信息识别系统能够很好地识别天然裂缝。实例表明,文中提出的基于粗糙集和神经网络的智能识别方法切实可行,在天然裂缝的识别中效果显著。

### 参考文献:

- [1] 张仕强,焦 棣,罗平亚,等.天然岩石裂缝表面形态描述[J].西南石油学院学报,1998,20(2):45-49.
- [2] 郑雷清,刘东付,左新玉,等.测井资料在砂岩裂缝识别中的应用[J].吐哈油气,2003,8(3):79-82.
- [3] Pawlak Z. Rough sets[J]. International Journal of Information and Computer Science, 1982,11(5):341-356.
- [4] 曾黄麟.粗糙集理论及其应用[M].重庆:重庆大学出版社,1998.
- [5] 韩力群.人工神经网络理论、设计及应用[M].北京:化学工业出版社,2002.
- [6] 刘光中,李晓峰.人工神经网络 BP 算法的改进和结构的自调整[J].运筹学学报,2001,5(1):81-88.