



(12)发明专利

(10)授权公告号 CN 106121622 B

(45)授权公告日 2019.10.25

(21)申请号 201610597512.X

(22)申请日 2016.07.27

(65)同一申请的已公布的文献号

申请公布号 CN 106121622 A

(43)申请公布日 2016.11.16

(73)专利权人 渤海大学

地址 121013 辽宁省锦州市松山新区科技
路19号渤海大学

(72)发明人 李琨 韩莹 余东生 魏泽飞

杨一柳 于震

(74)专利代理机构 西安铭泽知识产权代理事务

所(普通合伙) 61223

代理人 俞晓明

(51)Int.Cl.

E21B 44/00(2006.01)

E21B 47/00(2012.01)

G06F 17/50(2006.01)

G06F 17/10(2006.01)

(56)对比文件

CN 102094626 A,2011.06.15,

CN 102011576 A,2011.04.13,

CN 104481496 A,2015.04.01,

CN 105257277 A,2016.01.20,

CN 105672988 A,2016.06.15,

CN 204347530 U,2015.05.20,

李琨等.有杆泵抽油系统井下工况诊断的物
元分析方法.《东北大学学报(自然科学版)》
.2013,第34卷(第5期),

审查员 江定国

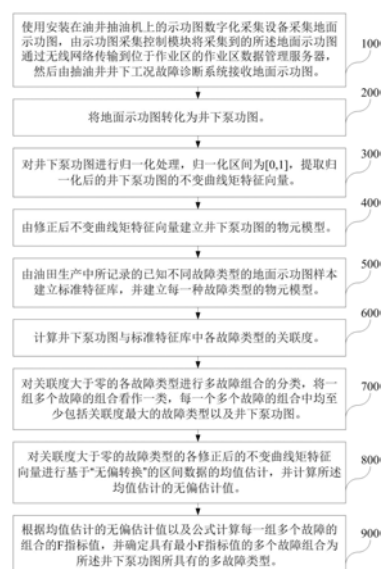
权利要求书4页 说明书13页 附图2页

(54)发明名称

一种基于示功图的有杆泵抽油井的多故障
诊断方法

(57)摘要

本发明公开了一种基于示功图的有杆泵抽油井的多故障诊断方法,涉及石油生产技术领域。将采集的地面示功图转换为井下泵功图,根据油田生产中已有的示功图数据建立各故障类型的标准特征库,对于每一个不变曲线矩特征向量的特征值,由区间数据的形式进行表示,然后计算所采集样本与标准特征库中各故障类型的关联度,判断所采集样本可能具有的故障类型,提高了故障诊断的可信性。最后将所采集样本可能具有的故障类型进行各种多个故障的组合,通过计算各个组合的F指标值来确定所具有的多故障类型。该方法原理简单,计算复杂性小,容易实现,诊断的准确度高。



1. 一种基于示功图的有杆泵抽油井的多故障诊断方法,其特征在于,包括:

采集地面示功图,并将所述地面示功图转化为井下泵功图;

对所述井下泵功图进行归一化处理,归一化区间为 $[0, 1]$,提取归一化后的井下泵功图的不变曲线矩特征向量,并对提取的不变曲线矩特征向量进行修正;

由修正后的不变曲线矩特征向量建立所述井下泵功图的物元模型,如下:

$$M = \begin{bmatrix} \text{井下工况} & \delta_1 & v_1 \\ & \delta_2 & v_2 \\ & \vdots & \vdots \\ & \delta_7 & v_7 \end{bmatrix} \quad (1)$$

其中, $\delta_1, \delta_2, \dots, \delta_7$ 分别为7个所述修正后的不变曲线矩特征向量, $v_1, v_2, \dots, v_7$ 分别为7个所述修正后的不变曲线矩特征向量的取值;

由油田生产中所记录的已知不同故障类型的地面示功图样本建立标准特征库,并建立每一种故障类型的物元模型,如下:

$$M_j = \begin{bmatrix} W_j & \delta_1 & [v_{j1a}, v_{j1b}] \\ & \delta_2 & [v_{j2a}, v_{j2b}] \\ & \vdots & \vdots \\ & \delta_7 & [v_{j7a}, v_{j7b}] \end{bmatrix} \quad (2)$$

其中, $j=1, 2, \dots, T$, T 表示已知故障类型的数量, W_j 表示第 j 种故障类型; $[v_{j1a}, v_{j1b}]$ 表示所述标准特征库中第 j 种故障类型下的第1个修正后的不变曲线矩特征向量的取值区间, $[v_{j2a}, v_{j2b}]$ 表示所述标准特征库中第 j 种故障类型下的第2个修正后的不变曲线矩特征向量的取值区间;依次类推, $[v_{j7a}, v_{j7b}]$ 表示所述标准特征库中第 j 种故障类型下的第7个修正后的不变曲线矩特征向量的取值区间;

计算所述井下泵功图与所述标准特征库中各故障类型的关联度,计算公式如下:

$$\varepsilon(v_k, v_{jk}) = \left| v_k - \frac{v_{jka} + v_{jkb}}{2} \right| - \frac{v_{jkb} - v_{jka}}{2} \quad (3)$$

$$\xi_{jk}(v_k) = \begin{cases} \frac{\varepsilon(v_k, v_{jk})}{\varepsilon(v_k, X) - \varepsilon(v_k, v_{jk})}, & v_k \notin v_{jk} \\ -\frac{\varepsilon(v_k, v_{jk})}{|v_{jk}|}, & v_k \in v_{jk} \end{cases} \quad (4)$$

$$\lambda(W_j) = \frac{1}{7} \sum_{k=1}^7 \xi_{jk}(v_k) \quad (5)$$

其中, $k=1, 2, \dots, 7$, v_k 表示所述井下泵功图第 k 个修正后的不变曲线矩特征向量的取值, v_{jk} 表示所述标准特征库中第 j 种故障类型下第 k 个修正后的不变曲线矩特征向量的取值, v_{jka} 和 v_{jkb} 分别表示所述标准特征库中第 j 种故障类型下第 k 个修正后的不变曲线矩特征向量取值的最小值和最大值, $|v_{jk}|$ 表示所述标准特征库中第 j 种故障类型下第 k 个修正后的不变曲线矩特征向量取值范围的大小;节域 $X=[0, 20]$; $\lambda(W_j)$ 表示所述井下泵功图与所述

标准特征库中第j种故障类型的关联度, $\varepsilon(v_k, v_{jk})$ 表示所述井下泵功图第k个修正后的不变曲线矩特征向量与所述标准特征库中第j种故障类型下第k个修正后的不变曲线矩特征向量的距离, $\xi_{jk}(v_k)$ 表示第k个修正后的不变曲线矩特征向量下所述井下泵功图与所述标准特征库中第j种故障类型的关联函数, $\varepsilon(v_k, X)$ 表示所述井下泵功图第k个修正后的不变曲线矩特征向量与节域的距离;

对 $\lambda(W_j) \geq 0$ 的各故障类型进行多故障组合的分类, 将一组多个故障的组合看作一类, 每一个所述多个故障的组合中均至少包括 $\lambda(W_j)$ 最大的故障类型以及所述井下泵功图;

对 $\lambda(W_j) \geq 0$ 的故障类型的各修正后的不变曲线矩特征向量进行基于“无偏转换”的区间数据的均值估计, 并计算所述均值估计的无偏估计值;

根据所述均值估计的无偏估计值以及以下公式计算每一组多个故障的组合的F指标值, 并确定具有最小F指标值的多个故障组合为所述井下泵功图所具有的多故障类型:

$$F = \frac{Sum_1 / (num-1)}{Sum_2 / (expt-1)} = \frac{\sum_{k=1}^{num} expt_k \left\| \overline{\alpha_k^{**}} - \overline{\alpha_{all}^{**}} \right\|^2 \cdot (expt - num)}{\sum_{k=1}^{num} \sum_{g=1}^{expt_k} \left\| \alpha_{kg}^{**} - \overline{\alpha_k^{**}} \right\|^2 \cdot (num-1)} \quad (6)$$

其中, $k=1, 2, \dots, num$, num 表示每一组多个故障组合中故障类型的个数; $expt_k$ 表示第k个故障类型下进行实验的个数, 即为第k个故障类型下修正后的不变曲线矩特征向量的个数, $expt_k=7$, $expt$ 为不同故障类型下所有实验的个数, 有: $expt = \sum_{k=1}^{num} expt_k$; α_{kg}^{**} 表示第k个故障类型下第g个修正后的不变曲线矩特征向量的区间数据的均值估计的无偏估计值, $g=1, 2, \dots, expt_k$; $\overline{\alpha_k^{**}}$ 表示第k个故障类型下所有修正后的不变曲线矩特征向量的区间数据的均值估计的无偏估计值的平均值, 有: $\overline{\alpha_k^{**}} = \frac{1}{expt_k} \sum_{g=1}^{expt_k} \alpha_{kg}^{**}$; $\overline{\alpha_{all}^{**}}$ 表示所有故障类型下的所有修正后的不变曲线矩特征向量的区间数据的均值估计的无偏估计值的总平均值, 有:

$\overline{\alpha_{all}^{**}} = \frac{1}{expt} \sum_{k=1}^{num} \sum_{g=1}^{expt_k} \alpha_{kg}^{**}$; sum_1 表示 $\overline{\alpha_k^{**}}$ 与 $\overline{\alpha_{all}^{**}}$ 的差异之和, 有: $Sum_1 = \sum_{k=1}^{num} expt_k \left\| \overline{\alpha_k^{**}} - \overline{\alpha_{all}^{**}} \right\|$; Sum_2 表示 α_{kg}^{**} 与 $\overline{\alpha_k^{**}}$ 的差异之和, 有: $Sum_2 = \sum_{k=1}^{num} \sum_{g=1}^{expt_k} \left\| \alpha_{kg}^{**} - \overline{\alpha_k^{**}} \right\|$;

步骤对 $\lambda(W_j) \geq 0$ 的故障类型的各修正后的不变曲线矩特征向量进行基于“无偏转换”的区间数据的均值估计, 并计算所述均值估计的无偏估计值具体包括:

假设所述区间数据的均值估计的值为 α , 是一个非负随机变量, 由 $[v_a, v_b]$ 表示修正后的不变曲线矩特征向量的取值的区间形式, $[v_a, v_b]$ 是与 α 独立的随机向量, 假设 $[v_a, v_b]$ 满足某种连续数据分布, 由 $g(\cdot, \cdot)$ 表示, 认为 $g(\cdot, \cdot)$ 满足均匀分布, 修正后的不变曲线矩特征向量取值区间 $[v_a, v_b]$ 的均匀分布由如下公式表示:

$$g(v_a, v_b) = \begin{cases} \frac{1}{v_b - v_a}, & 0 < v_a \leq v_b \\ 0, & \text{其它} \end{cases} \quad (7)$$

假设 $\{(v_{as}, v_{bs})\}_{s=1}^r$ 和 $\{\alpha_s\}_{s=1}^r$ 分别是 (v_a, v_b) 和 α 的独立同分布样本, 假设在实验中观测到

了 $(v_{as}, v_{bs}, \tau_{1s}, \tau_{2s}, \tau_{3s})$, 其中: $\tau_{1s} = I_{\alpha_s < v_{as}}$, $\tau_{2s} = I_{v_{as} < \alpha_s < v_{bs}}$, $\tau_{3s} = I_{\alpha_s > v_{bs}}$, I 表示示性函数, r 表示独立同分布样本数量, 令:

$$\tau_s = \begin{cases} 0, & \alpha_s < v_{as} \\ 1, & v_{as} \leq \alpha_s \leq v_{bs} \\ 2, & \alpha_s > v_{bs} \end{cases} \quad (8)$$

其中, τ_s 表示第 s 个独立同分布样本的观测值;

令:

$$\alpha_s^* = \theta_1(v_{as}, v_{bs}) \cdot \tau_{1s} + \theta_2(v_{as}, v_{bs}) \cdot \tau_{2s} + \theta_3(v_{as}, v_{bs}) \cdot \tau_{3s} \quad (9)$$

其中, α_s^* 表示第 s 个独立同分布样本的无偏估计值, $\theta_1(\cdot, \cdot)$ 、 $\theta_2(\cdot, \cdot)$ 和 $\theta_3(\cdot, \cdot)$ 是与 α 的分布无关的连续函数, 且 $\theta_1(\cdot, \cdot)$ 、 $\theta_2(\cdot, \cdot)$ 和 $\theta_3(\cdot, \cdot)$ 满足如下条件:

$$\begin{cases} \int_{v_b=0}^{\infty} \int_{v_a=0}^{v_b} \theta_1(v_a, v_b) g(v_a, v_b) dv_a dv_b = 0 \\ \int_{\alpha}^{\infty} [\theta_2(\alpha, v_b) - \theta_1(\alpha, v_b)] g(\alpha, v_b) dv_b \\ + \int_0^{\alpha} [\theta_3(v_a, \alpha) - \theta_2(v_a, \alpha)] g(v_a, \alpha) dv_a = 1 \end{cases} \quad (10)$$

计算区间数据 $[v_a, v_b]$ 的均值估计的无偏估计值如下:

$$\alpha^{**} = \frac{1}{r} \sum_{s=1}^r \alpha_s^* \quad (11)。$$

2. 如权利要求1所述的方法, 其特征在于, 步骤提取归一化后的井下泵功图的不变曲线矩特征向量, 并对提取的不变曲线矩特征向量进行修正具体包括:

归一化后的井下泵功图是由 N 个离散数据点 (x_i, y_i) , $(i=1, 2, \dots, N)$ 组成的曲线, 定义其 $p+q$ 阶曲线矩 c_{pq} 为:

$$c_{pq} = \sum_{i=1}^N x_i^p y_i^q \Delta l_i \quad (12)$$

其中, $\Delta l_i = \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2}$, $(p, q=0, 1, 2, 3)$ 为相邻两个离散数据点之间的直线距离;

相应的 $p+q$ 阶中心矩定义为:

$$\phi_{pq} = \sum_{i=1}^N (x_i - \bar{x})^p (y_i - \bar{y})^q \Delta l_i \quad (13)$$

其中, $\bar{x} = c_{10} / c_{00}$, $\bar{y} = c_{01} / c_{00}$, 点 $(\bar{x}, \bar{y})$ 为归一化后的井下泵功图曲线的重心坐标;

计算各阶中心矩如下:

$$\begin{aligned} \phi_{00} &= c_{00}, \phi_{10} = 0, \phi_{01} = 0, \phi_{11} = c_{11} - \bar{x}c_{01} = c_{11} - \bar{y}c_{10}, \phi_{20} = c_{20} - \bar{x}c_{10}, \phi_{02} = c_{02} - \bar{x}c_{01}, \phi_{30} = c_{30} - 3\bar{x}c_{20} + 2c_{10}\bar{x}^2, \\ \phi_{03} &= c_{03} - 3\bar{y}c_{02} + 2c_{01}\bar{y}^2, \phi_{12} = c_{12} - 2\bar{y}c_{11} - \bar{x}c_{02} + 2\bar{y}^2 c_{10}, \phi_{21} = c_{21} - 2\bar{x}c_{11} - \bar{y}c_{02} + 2\bar{x}^2 c_{01}; \end{aligned}$$

对得到的各阶中心矩进行规范化处理, 采用如下规范化公式:

$$\varphi_{pq} = \phi_{pq} / (\phi_{00})^{p+q+1} \quad (14)$$

计算7个所述不变曲线矩特征向量如下：

$$\gamma_1 = \varphi_{20} + \varphi_{02} \quad (15)$$

$$\gamma_2 = (\varphi_{20} - \varphi_{02})^2 + 4\varphi_{11}^2 \quad (16)$$

$$\gamma_3 = (\varphi_{30} - 3\varphi_{12})^2 + (3\varphi_{21} - \varphi_{03})^2 \quad (17)$$

$$\gamma_4 = (\varphi_{30} + \varphi_{12})^2 + (\varphi_{21} + \varphi_{03})^2 \quad (18)$$

$$\gamma_5 = (\varphi_{30} - \varphi_{12})(\varphi_{30} + \varphi_{12})[(\varphi_{30} + \varphi_{12})^2 - 3(\varphi_{21} + \varphi_{03})^2] + (3\varphi_{21} + \varphi_{03})(\varphi_{21} + \varphi_{03})[3(\varphi_{30} + \varphi_{12})^2 - (\varphi_{21} + \varphi_{03})^2] \quad (19)$$

$$\gamma_6 = (\varphi_{20} - \varphi_{02})[(\varphi_{30} + \varphi_{12})^2 - (\varphi_{21} + \varphi_{03})^2] + 4\varphi_{11}(\varphi_{30} + \varphi_{12})(\varphi_{21} + \varphi_{03})^2 \quad (20)$$

$$\gamma_7 = (3\varphi_{21} - \varphi_{03})(\varphi_{30} + \varphi_{12})[(\varphi_{30} + \varphi_{12})^2 - 3(\varphi_{21} + \varphi_{03})^2] + (3\varphi_{12} - \varphi_{30})(\varphi_{21} + \varphi_{03})[(\varphi_{30} + \varphi_{12})^2 - 3(\varphi_{21} + \varphi_{03})^2] \quad (21)$$

对7个所述不变曲线矩特征向量进行修正,修正公式如下：

$$\delta_i = |\lg |\gamma_i|| \quad (22)$$

其中, $i=1, 2, \dots, 7$, 得到7个修正后的不变曲线矩特征向量。

一种基于示功图的有杆泵抽油井的多故障诊断方法

技术领域

[0001] 本发明涉及石油生产技术领域,特别涉及一种基于示功图的有杆泵抽油井的多故障诊断方法。

背景技术

[0002] 有杆泵抽油井是国内外油田的主要生产方式,实现其连续、稳定和高效的运行,是提高油井生产效率、降低生产成本和减少生产安全隐患的重要措施。由于抽油泵工作在数千米的地下,工作环境恶劣,使得井下工况极其复杂,故障时有发生,轻则停井影响生产,重则发生安全事故,因此,对有杆泵抽油井井下工况的准确故障诊断是油田企业一直关注的问题。

[0003] 随着计算机信息技术、通信技术、电子技术等的不断发展,利用计算机实现抽油井井下工况的故障诊断已越来越受到重视。计算机诊断方式代替人工诊断,可以高效率处理大规模数据,在较短的时间内得到诊断结论,为管理部门根据实时工况制定合理的生产措施提供便利。

[0004] 但是,目前已有的诊断方法一般只考虑一种故障工况发生的情况,而实际生产中多种故障同时发生的情况是普遍存在的,因此,当井下多个故障工况同时出现时,现有的方法无法做出准确分析以提供有效的技术指导。

发明内容

[0005] 本发明实施例提供了一种基于示功图的有杆泵抽油井的多故障诊断方法,用以解决现有技术中存在的问题。

[0006] 一种基于示功图的有杆泵抽油井的多故障诊断方法,包括:

[0007] 采集地面示功图,并将所述地面示功图转化为井下泵功图;

[0008] 对所述井下泵功图进行归一化处理,归一化区间为 $[0,1]$,提取归一化后的井下泵功图的不变曲线矩特征向量,并对提取的不变曲线矩特征向量进行修正;

[0009] 由修正后的不变曲线矩特征向量建立所述井下泵功图的物元模型,如下:

$$[0010] \quad M = \begin{bmatrix} \text{井下工况} & \delta_1 & v_1 \\ & \delta_2 & v_2 \\ & \vdots & \vdots \\ & \delta_7 & v_7 \end{bmatrix} \quad (1)$$

[0011] 其中, $\delta_1, \delta_2, \dots, \delta_7$ 分别为7个所述修正后的不变曲线矩特征向量, $v_1, v_2, \dots, v_7$ 分别为7个所述修正后的不变曲线矩特征向量的取值;

[0012] 由油田生产中所记录的已知不同故障类型的地面示功图样本建立标准特征库,并建立每一种故障类型的物元模型,如下:

$$[0013] \quad M_j = \begin{bmatrix} W_j & \delta_1 & [v_{j1a}, v_{j1b}] \\ & \delta_2 & [v_{j2a}, v_{j2b}] \\ & \vdots & \vdots \\ & \delta_7 & [v_{j7a}, v_{j7b}] \end{bmatrix} \quad (2)$$

[0014] 其中, $j=1, 2, \dots, T$, T 表示已知故障类型的数量, W_j 表示第 j 种故障类型; $[v_{j1a}, v_{j1b}]$ 表示所述标准特征库中第 j 种故障类型下的第1个修正后的不变曲线矩特征向量的取值区间, $[v_{j2a}, v_{j2b}]$ 表示所述标准特征库中第 j 种故障类型下的第2个修正后的不变曲线矩特征向量的取值区间; 依次类推, $[v_{j7a}, v_{j7b}]$ 表示所述标准特征库中第 j 种故障类型下的第7个修正后的不变曲线矩特征向量的取值区间;

[0015] 计算所述井下泵功图与所述标准特征库中各故障类型的关联度, 计算公式如下:

$$[0016] \quad \varepsilon(v_k, v_{jk}) = \left| v_k - \frac{v_{jka} + v_{jkb}}{2} \right| - \frac{v_{jkb} - v_{jka}}{2} \quad (3)$$

$$[0017] \quad \xi_{jk}(v_k) = \begin{cases} \frac{\varepsilon(v_k, v_{jk})}{\varepsilon(v_k, X) - \varepsilon(v_k, v_{jk})}, & v_k \notin v_{jk} \\ -\frac{\varepsilon(v_k, v_{jk})}{|v_{jk}|}, & v_k \in v_{jk} \end{cases} \quad (4)$$

$$[0018] \quad \lambda(W_j) = \frac{1}{7} \sum_{k=1}^7 \xi_{jk}(v_k) \quad (5)$$

[0019] 其中, $k=1, 2, \dots, 7$, v_k 表示所述井下泵功图第 k 个修正后的不变曲线矩特征向量的取值, v_{jk} 表示所述标准特征库中第 j 种故障类型下第 k 个修正后的不变曲线矩特征向量的取值, v_{jka} 和 v_{jkb} 分别表示所述标准特征库中第 j 种故障类型下第 k 个修正后的不变曲线矩特征向量取值的最小值和最大值, $|v_{jk}|$ 表示所述标准特征库中第 j 种故障类型下第 k 个修正后的不变曲线矩特征向量取值范围的大小; 节域 $X=[0, 20]$; $\lambda(W_j)$ 表示所述井下泵功图与所述标准特征库中第 j 种故障类型的关联度, $\varepsilon(v_k, v_{jk})$ 表示所述井下泵功图第 k 个修正后的不变曲线矩特征向量与所述标准特征库中第 j 种故障类型下第 k 个修正后的不变曲线矩特征向量的距离, $\xi_{jk}(v_k)$ 表示第 k 个修正后的不变曲线矩特征向量下所述井下泵功图与所述标准特征库中第 j 种故障类型的关联函数, $\varepsilon(v_k, X)$ 表示所述井下泵功图第 k 个修正后的不变曲线矩特征向量与节域的距离;

[0020] 对 $\lambda(W_j) \geq 0$ 的各故障类型进行多故障组合的分类, 将一组多个故障的组合看作一类, 每一个所述多个故障的组合中均至少包括 $\lambda(W_j)$ 最大的故障类型以及所述井下泵功图;

[0021] 对 $\lambda(W_j) \geq 0$ 的故障类型的各修正后的不变曲线矩特征向量进行基于“无偏转换”的区间数据的均值估计, 并计算所述均值估计的无偏估计值;

[0022] 根据所述均值估计的无偏估计值以及以下公式计算每一组多个故障的组合的F指标值, 并确定具有最小F指标值的多个故障组合为所述井下泵功图所具有的多故障类型:

$$[0023] \quad F = \frac{Sum_1 / (num-1)}{Sum_2 / (expt-1)} = \frac{\sum_{k=1}^{num} expt_k \left\| \overline{\alpha_k^{**}} - \overline{\alpha_{all}^{**}} \right\|^2 \cdot (expt - num)}{\sum_{k=1}^{num} \sum_{g=1}^{expt_k} \left\| \alpha_{kg}^{**} - \overline{\alpha_k^{**}} \right\|^2 \cdot (num-1)} \quad (6)$$

[0024] 其中, $k=1, 2, \dots, num$, num 表示每一组多个故障组合中故障类型的个数; $expt_k$ 表示第 k 个故障类型下进行实验的个数, 即为第 k 个故障类型下修正后的不变曲线矩特征向量的

的个数, $expt_k=7$, $expt$ 为不同故障类型下所有实验的个数, 有: $expt = \sum_{k=1}^{num} expt_k$; α_{kg}^{**}

表示第 k 个故障类型下第 g 个修正后的不变曲线矩特征向量的区间数据的均值估计的无偏估计值, $g=1, 2, \dots, expt_k$; $\overline{\alpha_k^{**}}$ 表示第 k 个故障类型下所有修正后的不变曲线矩特征向

量的区间数据的均值估计的无偏估计值的平均值, 有: $\overline{\alpha_k^{**}} = \frac{1}{expt_k} \sum_{g=1}^{expt_k} \alpha_{kg}^{**}$; $\overline{\alpha_{all}^{**}}$ 表示

所有故障类型下的所有修正后的不变曲线矩特征向量的区间数据的均值估计的无偏估计

值的总平均值, 有: $\overline{\alpha_{all}^{**}} = \frac{1}{expt} \sum_{k=1}^{num} \sum_{g=1}^{expt_k} \alpha_{kg}^{**}$; sum_1 表示 $\overline{\alpha_k^{**}}$ 与 $\overline{\alpha_{all}^{**}}$ 的差异之和, 有:

$Sum_1 = \sum_{k=1}^{num} expt_k \left\| \overline{\alpha_k^{**}} - \overline{\alpha_{all}^{**}} \right\|$; Sum_2 表示 α_{kg}^{**} 与 $\overline{\alpha_k^{**}}$ 的差异之和, 有:

$$Sum_2 = \sum_{k=1}^{num} \sum_{g=1}^{expt_k} \left\| \alpha_{kg}^{**} - \overline{\alpha_k^{**}} \right\|。$$

[0025] 优选地, 步骤提取归一化后的井下泵功图的不变曲线矩特征向量, 并对提取的不变曲线矩特征向量进行修正具体包括:

[0026] 归一化后的井下泵功图是由 N 个离散数据点 (x_i, y_i) , $(i=1, 2, \dots, N)$ 组成的曲线, 定义其 $p+q$ 阶曲线矩 c_{pq} 为:

$$[0027] \quad c_{pq} = \sum_{i=1}^N x_i^p y_i^q \Delta l_i \quad (7)$$

[0028] 其中, $\Delta l_i = \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2}$, $(p, q=0, 1, 2, 3)$ 为相邻两个离散数据点之间的直线距离;

[0029] 相应的 $p+q$ 阶中心矩定义为:

$$[0030] \quad \phi_{pq} = \sum_{i=1}^N (x_i - \bar{x})^p (y_i - \bar{y})^q \Delta l_i \quad (8)$$

[0031] 其中, $\bar{x} = c_{10} / c_{00}$, $\bar{y} = c_{01} / c_{00}$, 点 $(\bar{x}, \bar{y})$ 为归一化后的井下泵功图曲线的重心坐标;

[0032] 计算各阶中心矩如下：

$$\begin{aligned} [0033] \quad \phi_{00} &= c_{00}, \phi_{10} = 0, \phi_{01} = 0, \phi_{11} = c_{11} - \bar{x}c_{01} = c_{11} - \bar{y}c_{10}, \quad \phi_{20} = c_{20} - \bar{x}c_{10}, \quad \phi_{02} = c_{02} - \bar{x}c_{01}, \\ \phi_{30} &= c_{30} - 3\bar{x}c_{20} + 2c_{10}\bar{x}^2, \quad \phi_{03} = c_{03} - 3\bar{y}c_{02} + 2c_{01}\bar{y}^2, \quad \phi_{12} = c_{12} - 2\bar{y}c_{11} - \bar{x}c_{02} + 2\bar{y}^2c_{10}, \\ \phi_{21} &= c_{21} - 2\bar{x}c_{11} - \bar{y}c_{02} + 2\bar{x}^2c_{01}; \end{aligned}$$

[0034] 对得到的各阶中心矩进行规范化处理,采用如下规范化公式：

$$[0035] \quad \varphi_{pq} = \phi_{pq} / (\phi_{00})^{p+q+1} \quad (9)$$

[0036] 计算7个所述不变曲线矩特征向量如下：

$$[0037] \quad \gamma_1 = \varphi_{20} + \varphi_{02} \quad (10)$$

$$[0038] \quad \gamma_2 = (\varphi_{20} - \varphi_{02})^2 + 4\varphi_{11}^2 \quad (11)$$

$$[0039] \quad \gamma_3 = (\varphi_{30} - 3\varphi_{12})^2 + (3\varphi_{21} - \varphi_{03})^2 \quad (12)$$

$$[0040] \quad \gamma_4 = (\varphi_{30} + \varphi_{12})^2 + (\varphi_{21} + \varphi_{03})^2 \quad (13)$$

$$\begin{aligned} [0041] \quad \gamma_5 &= (\varphi_{30} - \varphi_{12})(\varphi_{30} + \varphi_{12})[(\varphi_{30} + \varphi_{12})^2 - 3(\varphi_{21} + \varphi_{03})^2] + \\ &\quad (3\varphi_{21} + \varphi_{03})(\varphi_{21} + \varphi_{03})[3(\varphi_{30} + \varphi_{12})^2 - (\varphi_{21} + \varphi_{03})^2] \end{aligned} \quad (14)$$

$$\begin{aligned} [0042] \quad \gamma_6 &= (\varphi_{20} - \varphi_{02})[(\varphi_{30} + \varphi_{12})^2 - (\varphi_{21} + \varphi_{03})^2] + \\ &\quad 4\varphi_{11}(\varphi_{30} + \varphi_{12})(\varphi_{21} + \varphi_{03})^2 \end{aligned} \quad (15)$$

$$\begin{aligned} [0043] \quad \gamma_7 &= (3\varphi_{21} - \varphi_{03})(\varphi_{30} + \varphi_{12})[(\varphi_{30} + \varphi_{12})^2 - 3(\varphi_{21} + \varphi_{03})^2] + \\ &\quad (3\varphi_{12} - \varphi_{30})(\varphi_{21} + \varphi_{03})[(\varphi_{30} + \varphi_{12})^2 - 3(\varphi_{21} + \varphi_{03})^2] \end{aligned} \quad (16)$$

[0044] 对7个所述不变曲线矩特征向量进行修正,修正公式如下：

$$[0045] \quad \delta_i = |\lg |\gamma_i|| \quad (17)$$

[0046] 其中, $i=1, 2, \dots, 7$, 得到7个修正后的不变曲线矩特征向量。

[0047] 优选地,步骤对 $\lambda(W_j) \geq 0$ 的故障类型的各修正后的不变曲线矩特征向量进行基于“无偏转换”的区间数据的均值估计,并计算所述均值估计的无偏估计值具体包括：

[0048] 假设所述区间数据的均值估计的值为 α , 是一个非负随机变量,由 $[v_a, v_b]$ 表示修正后的不变曲线矩特征向量的取值的区间形式, $[v_a, v_b]$ 是与 α 独立的随机向量,假设 $[v_a, v_b]$ 满足某种连续数据分布,由 $g(\cdot, \cdot)$ 表示,认为 $g(\cdot, \cdot)$ 满足均匀分布,修正后的不变曲线矩特征向量取值区间 $[v_a, v_b]$ 的均匀分布由如下公式表示：

$$[0049] \quad g(v_a, v_b) = \begin{cases} \frac{1}{v_b - v_a}, & 0 < v_a \leq v_b \\ 0, & \text{其它} \end{cases} \quad (18)$$

[0050] 假设 $\{(v_{as}, v_{bs})\}_{s=1}^r$ 和 $\{\alpha_s\}_{s=1}^r$ 分别是 (v_a, v_b) 和 α 的独立同分布样本,假设在

实验中观测到了 $(v_{as}, v_{bs}, \tau_{1s}, \tau_{2s}, \tau_{3s})$, 其中: $\tau_{1s} = I_{\alpha_s < v_{as}}$,

$\tau_{2s} = I_{v_{as} < \alpha_s < v_{bs}}$, $\tau_{3s} = I_{\alpha_s > v_{bs}}$, I 表示示性函数, r 表示独立同分布样本数量, 令:

$$[0051] \quad \tau_s = \begin{cases} 0, & \alpha_s < v_{as} \\ 1, & v_{as} \leq \alpha_s \leq v_{bs} \\ 2, & \alpha_s > v_{bs} \end{cases} \quad (19)$$

[0052] 其中, τ_s 表示第 s 个独立同分布样本的观测值;

[0053] 令:

$$[0054] \quad \alpha_s^* = \theta_1(v_{as}, v_{bs}) \cdot \tau_{1s} + \theta_2(v_{as}, v_{bs}) \cdot \tau_{2s} + \theta_3(v_{as}, v_{bs}) \cdot \tau_{3s} \quad (20)$$

[0055] 其中, α_s^* 表示第 s 个独立同分布样本的无偏估计值, $\theta_1(\cdot, \cdot)$ 、 $\theta_2(\cdot, \cdot)$ 和 $\theta_3(\cdot, \cdot)$ 是与 α 的分布无关的连续函数, 且 $\theta_1(\cdot, \cdot)$ 、 $\theta_2(\cdot, \cdot)$ 和 $\theta_3(\cdot, \cdot)$ 满足如下条件:

$$[0056] \quad \begin{cases} \int_{v_b=0}^{\infty} \int_{v_a=0}^{v_b} \theta_1(v_a, v_b) g(v_a, v_b) dv_a dv_b = 0 \\ \int_{\alpha}^{\infty} [\theta_2(\alpha, v_b) - \theta_1(\alpha, v_b)] g(\alpha, v_b) dv_b \\ + \int_0^{\alpha} [\theta_3(v_a, \alpha) - \theta_2(v_a, \alpha)] g(v_a, \alpha) dv_a = 1 \end{cases} \quad (21)$$

[0057] 计算区间数据 $[v_a, v_b]$ 的均值估计的无偏估计值如下:

$$[0058] \quad \alpha^{**} = \frac{1}{r} \sum_{s=1}^r \alpha_s^* \quad (22)。$$

[0059] 本发明的有益效果是:

[0060] 1、泵功图的图形特征可以反映有杆泵抽油井的井下工况, 提取能够准确描述图形特征的特征向量是进行故障诊断的有效依据。本发明采用不变矩理论所提取的7个不变曲线矩特征向量, 其中的各阶矩都具有特定的物理意义, ϕ_{00} 表示泵功图曲线的长度; ϕ_{10} 和 ϕ_{01} 可以用来确定泵功图曲线的灰度重心; ϕ_{20} 、 ϕ_{11} 和 ϕ_{02} 用来衡量泵功图曲线的大小和方向; ϕ_{30} 和 ϕ_{03} 表示泵功图图形的不对称性, ϕ_{30} 表示泵功图曲线关于垂直轴线不对称性度量, ϕ_{03} 表示泵功图曲线关于水平轴线的不对称度量。

[0061] 2、本发明根据油田生产中已有的示功图数据建立各故障类型的标准特征库, 对于每一个不变曲线矩特征向量的特征值, 由区间数据的形式进行表示, 解决传统方法只考虑特征向量的单值形式, 而不能定量化体现数据变化差异的问题, 然后计算所采集样本与标准特征库中各故障类型的关联度, 通过量化分析的方式来判断所采集样本可能具有的故障类型, 提高了故障诊断的可信性。

[0062] 3、本发明所建立的有杆泵抽油井井下工况的多故障诊断方法, 将所采集样本可能具有的故障类型进行各种多个故障的组合, 通过计算各个组合的F指标值来确定所具有的多故障类型, F指标值越小说明该组合中各故障类型之间的差异最小, 各个故障类型同时发

生的可能性最高。该方法原理简单,计算复杂性小,容易实现,诊断的准确度高。

附图说明

[0063] 为了更清楚地说明本发明发明实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本发明发明的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0064] 图1为本发明实施例提供的一种基于示功图的有杆泵抽油井的多故障诊断方法的步骤流程图;

[0065] 图2为由有杆泵抽油井地面示功图无线远程采集系统采集到的地面示功图;

[0066] 图3为图2中地面示功图转化得到的井下泵功图。

具体实施方式

[0067] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0068] 参照图1,本发明实施例提供了一种基于示功图的有杆泵抽油井的多故障诊断方法,所述方法包括:

[0069] 步骤100,使用安装在油井抽油机上的示功图数字化采集设备采集地面示功图,由示功图采集控制模块将采集到的所述地面示功图通过无线网络传输到位于作业区的作业区数据管理服务器,然后由抽油井井下工况故障诊断系统接收所述地面示功图。

[0070] 步骤200,将所述地面示功图转化为井下泵功图。

[0071] 步骤300,对所述井下泵功图进行归一化处理,归一化区间为 $[0,1]$,然后对归一化后的井下泵功图进行特征向量提取,具体提取步骤为:

[0072] 子步骤310,归一化后的井下泵功图是由 N 个离散数据点 (x_i, y_i) , $(i=1, 2, \dots, N)$ 组成的曲线,定义其 $p+q$ 阶曲线矩 c_{pq} 为:

$$[0073] \quad c_{pq} = \sum_{i=1}^N x_i^p y_i^q \Delta l_i \quad (1)$$

[0074] 其中, $\Delta l_i = \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2}$, $(p, q=0, 1, 2, 3)$ 为相邻两个离散数据点之间的直线距离。

[0075] 相应的 $p+q$ 阶中心矩定义为:

$$[0076] \quad \phi_{pq} = \sum_{i=1}^N (x_i - \bar{x})^p (y_i - \bar{y})^q \Delta l_i \quad (2)$$

[0077] 其中, $\bar{x} = c_{10} / c_{00}$, $\bar{y} = c_{01} / c_{00}$, 点 $(\bar{x}, \bar{y})$ 为归一化后的井下泵功图曲线的重心坐标。

[0078] 计算各阶中心矩如下:

[0079] $\phi_{00} = c_{00}, \phi_{10} = 0, \phi_{01} = 0, \phi_{11} = c_{11} - \bar{x}c_{01} = c_{11} - \bar{y}c_{10}, \phi_{20} = c_{20} - \bar{x}c_{10}, \phi_{02} = c_{02} - \bar{x}c_{01},$
 $\phi_{30} = c_{30} - 3\bar{x}c_{20} + 2c_{10}\bar{x}^2, \phi_{03} = c_{03} - 3\bar{y}c_{02} + 2c_{01}\bar{y}^2, \phi_{12} = c_{12} - 2\bar{y}c_{11} - \bar{x}c_{02} + 2\bar{y}^2c_{10},$
 $\phi_{21} = c_{21} - 2\bar{x}c_{11} - \bar{y}c_{02} + 2\bar{x}^2c_{01}。$

[0080] 子步骤320,对得到的各阶中心矩进行规范化处理,采用如下规范化公式:

$$[0081] \quad \varphi_{pq} = \phi_{pq} / (\phi_{00})^{p+q+1} \quad (3)$$

[0082] 子步骤330,计算7个不变曲线矩特征向量如下:

$$[0083] \quad \gamma_1 = \varphi_{20} + \varphi_{02} \quad (4)$$

$$[0084] \quad \gamma_2 = (\varphi_{20} - \varphi_{02})^2 + 4\varphi_{11}^2 \quad (5)$$

$$[0085] \quad \gamma_3 = (\varphi_{30} - 3\varphi_{12})^2 + (3\varphi_{21} - \varphi_{03})^2 \quad (6)$$

$$[0086] \quad \gamma_4 = (\varphi_{30} + \varphi_{12})^2 + (\varphi_{21} + \varphi_{03})^2 \quad (7)$$

$$[0087] \quad \gamma_5 = (\varphi_{30} - \varphi_{12})(\varphi_{30} + \varphi_{12})[(\varphi_{30} + \varphi_{12})^2 - 3(\varphi_{21} + \varphi_{03})^2] + \\ (3\varphi_{21} + \varphi_{03})(\varphi_{21} + \varphi_{03})[3(\varphi_{30} + \varphi_{12})^2 - (\varphi_{21} + \varphi_{03})^2] \quad (8)$$

$$[0088] \quad \gamma_6 = (\varphi_{20} - \varphi_{02})[(\varphi_{30} + \varphi_{12})^2 - (\varphi_{21} + \varphi_{03})^2] + \\ 4\varphi_{11}(\varphi_{30} + \varphi_{12})(\varphi_{21} + \varphi_{03})^2 \quad (9)$$

$$[0089] \quad \gamma_7 = (3\varphi_{21} - \varphi_{03})(\varphi_{30} + \varphi_{12})[(\varphi_{30} + \varphi_{12})^2 - 3(\varphi_{21} + \varphi_{03})^2] + \\ (3\varphi_{12} - \varphi_{30})(\varphi_{21} + \varphi_{03})[(\varphi_{30} + \varphi_{12})^2 - 3(\varphi_{21} + \varphi_{03})^2] \quad (10)$$

[0090] 子步骤340,对7个不变曲线矩特征向量进行修正,修正公式如下:

$$[0091] \quad \delta_i = |\lg |\gamma_i|| \quad (11)$$

[0092] 其中, $i=1,2,\dots,7$,得到7个修正后的不变曲线矩特征向量。

[0093] 步骤400,由计算得到的修正后的不变曲线矩特征向量建立所述井下泵功图的物元模型,如下:

$$[0094] \quad M = \begin{bmatrix} \text{井下工况} & \delta_1 & v_1 \\ & \delta_2 & v_2 \\ & \vdots & \vdots \\ & \delta_7 & v_7 \end{bmatrix} \quad (12)$$

[0095] 其中, $v_1, v_2, \dots, v_7$ 分别为7个所述修正后的不变曲线矩特征向量的取值。

[0096] 步骤500,由油田生产中所记录的已知不同故障类型的地面示功图样本建立标准特征库,根据步骤200~400得到每一种故障类型的物元模型如下:

$$[0097] \quad M_j = \begin{bmatrix} W_j & \delta_1 & [v_{j1a}, v_{j1b}] \\ & \delta_2 & [v_{j2a}, v_{j2b}] \\ & \vdots & \vdots \\ & \delta_7 & [v_{j7a}, v_{j7b}] \end{bmatrix} \quad (13)$$

[0098] 其中, $j=1, 2, \dots, T$, T 表示已知故障类型的数量, W_j 表示第 j 种故障类型; $[v_{j1a}, v_{j1b}]$ 表示所述标准特征库中第 j 种故障类型下的第1个修正后的不变曲线矩特征向量的取值区间, $[v_{j2a}, v_{j2b}]$ 表示所述标准特征库中第 j 种故障类型下的第2个修正后的不变曲线矩特征向量的取值区间;依次类推, $[v_{j7a}, v_{j7b}]$ 表示所述标准特征库中第 j 种故障类型下的第7个修正后的不变曲线矩特征向量的取值区间。

[0099] 步骤600, 计算所述井下泵功图与标准特征库中各故障类型的关联度, 计算公式如下:

$$[0100] \quad \varepsilon(v_k, v_{jk}) = \left| v_k - \frac{v_{jka} + v_{jkb}}{2} \right| - \frac{v_{jkb} - v_{jka}}{2} \quad (14)$$

$$[0101] \quad \xi_{jk}(v_k) = \begin{cases} \frac{\varepsilon(v_k, v_{jk})}{\varepsilon(v_k, X) - \varepsilon(v_k, v_{jk})}, & v_k \notin v_{jk} \\ -\frac{\varepsilon(v_k, v_{jk})}{|v_{jk}|}, & v_k \in v_{jk} \end{cases} \quad (15)$$

$$[0102] \quad \lambda(W_j) = \frac{1}{7} \sum_{k=1}^7 \xi_{jk}(v_k) \quad (16)$$

[0103] 其中, $k=1, 2, \dots, 7$, v_k 表示所述井下泵功图第 k 个修正后的不变曲线矩特征向量的取值, v_{jk} 表示标准特征库中第 j 种故障类型下第 k 个修正后的不变曲线矩特征向量的取值, v_{jka} 和 v_{jkb} 分别表示标准特征库中第 j 种故障类型下第 k 个修正后的不变曲线矩特征向量取值的最小值和最大值, $|v_{jk}|$ 表示标准特征库中第 j 种故障类型下第 k 个修正后的不变曲线矩特征向量取值范围的大小; 节域 $X=[0, 20]$; $\lambda(W_j)$ 表示所采集的井下泵功图与标准特征库中第 j 种故障类型的关联度, $\varepsilon(v_k, v_{jk})$ 表示所述井下泵功图第 k 个修正后的不变曲线矩特征向量与所述标准特征库中第 j 种故障类型下第 k 个修正后的不变曲线矩特征向量的距离, $\xi_{jk}(v_k)$ 表示第 k 个修正后的不变曲线矩特征向量下所述井下泵功图与所述标准特征库中第 j 种故障类型的关联函数, $\varepsilon(v_k, X)$ 表示所述井下泵功图第 k 个修正后的不变曲线矩特征向量与节域的距离。

[0104] 当 $\lambda(W_j) < 0$ 时, 表示所述井下泵功图没有发生该类型的故障; 当 $\lambda(W_j) \geq 0$ 时, 表示所述井下泵功图可能发生该类型的故障, 数值越大, 发生的程度就越高。

[0105] 步骤700, 对 $\lambda(W_j) \geq 0$ 的各故障类型进行多故障组合的分类, 认为 $\lambda(W_j)$ 最大的该类型的故障是必然发生的, 将一组多个故障的组合看作一类, 每一个分组中均至少包括所述 $\lambda(W_j)$ 最大的故障类型以及所述井下泵功图。

[0106] 步骤800, 由于标准特征库中每一种故障类型的修正后的不变曲线矩特征向量的

取值都为区间值,因此对步骤700中所得到的 $\lambda(W_j) \geq 0$ 的故障类型的各修正后的不变曲线矩特征向量进行基于“无偏转换”的区间数据的均值估计,并计算所述均值估计的无偏估计值,具体步骤如下:

[0107] 子步骤810,假设区间数据的均值估计的值为 α ,是一个非负随机变量,由 $[v_a, v_b]$ 表示修正后的不变曲线矩特征向量的取值的区间形式, $[v_a, v_b]$ 是与 α 独立的随机向量, α 可能在区间 $[v_a, v_b]$ 内,也可能在区间 $[v_a, v_b]$ 的左侧,或在区间 $[v_a, v_b]$ 的右侧。假设 $[v_a, v_b]$ 满足某种连续数据分布,由 $g(\cdot, \cdot)$ 表示,本实施例中认为 $g(\cdot, \cdot)$ 满足均匀分布,修正后的不变曲线矩特征向量取值区间 $[v_a, v_b]$ 的均匀分布由如下公式表示:

$$[0108] \quad g(v_a, v_b) = \begin{cases} \frac{1}{v_b - v_a}, & 0 < v_a \leq v_b \\ 0, & \text{其它} \end{cases} \quad (17)$$

[0109] 子步骤820,假设 $\{(v_{as}, v_{bs})\}_{s=1}^r$ 和 $\{\alpha_s\}_{s=1}^r$ 分别是 (v_a, v_b) 和 α 的独立同分布样本,假设在实验中观测到了 $(v_{as}, v_{bs}, \tau_{1s}, \tau_{2s}, \tau_{3s})$,其中: $\tau_{1s} = I_{\alpha_s < v_{as}}$,
 $\tau_{2s} = I_{v_{as} < \alpha_s < v_{bs}}$, $\tau_{3s} = I_{\alpha_s > v_{bs}}$, I 表示示性函数, r 表示独立同分布样本数量,令:

$$[0110] \quad \tau_s = \begin{cases} 0, & \alpha_s < v_{as} \\ 1, & v_{as} \leq \alpha_s \leq v_{bs} \\ 2, & \alpha_s > v_{bs} \end{cases} \quad (18)$$

[0111] 其中, τ_s 表示第 s 个独立同分布样本的观测值。

[0112] 子步骤830,令:

[0113]

$$\alpha_s^* = \theta_1(v_{as}, v_{bs}) \cdot \tau_{1s} + \theta_2(v_{as}, v_{bs}) \cdot \tau_{2s} + \theta_3(v_{as}, v_{bs}) \cdot \tau_{3s} \quad (19)$$

[0114] 其中, α_s^* 表示第 s 个独立同分布样本的无偏估计值, $\theta_1(\cdot, \cdot)$ 、 $\theta_2(\cdot, \cdot)$ 和 $\theta_3(\cdot, \cdot)$ 是与 α 的分布无关的连续函数,且 $\theta_1(\cdot, \cdot)$ 、 $\theta_2(\cdot, \cdot)$ 和 $\theta_3(\cdot, \cdot)$ 满足如下条件:

$$[0115] \quad \begin{cases} \int_{v_b=0}^{\infty} \int_{v_a=0}^{v_b} \theta_1(v_a, v_b) g(v_a, v_b) dv_a dv_b = 0 \\ \int_{\alpha}^{\infty} [\theta_2(\alpha, v_b) - \theta_1(\alpha, v_b)] g(\alpha, v_b) dv_b \\ + \int_0^{\alpha} [\theta_3(v_a, \alpha) - \theta_2(v_a, \alpha)] g(v_a, \alpha) dv_a = 1 \end{cases} \quad (20)$$

[0116] 子步骤840,那么计算区间数据 $[v_a, v_b]$ 的均值估计的无偏估计值如下:

$$[0117] \quad \alpha^{**} = \frac{1}{r} \sum_{s=1}^r \alpha_s^* \quad (21)$$

[0118] 步骤900,计算每一组多个故障组合的F指标值,并具有最小F指标值的多个故障组合为所采集的井下泵功图所具有的多故障类型。具体包括以下方法:

[0119] 子步骤910,采用如下公式计算每一组多个故障组合的F指标值,

$$[0120] \quad F = \frac{Sum_1 / (num-1)}{Sum_2 / (expt-1)} = \frac{\sum_{k=1}^{num} expt_k \left\| \overline{\alpha_k^{**}} - \overline{\alpha_{all}^{**}} \right\|^2 \cdot (expt - num)}{\sum_{k=1}^{num} \sum_{g=1}^{expt_k} \left\| \alpha_{kg}^{**} - \overline{\alpha_k^{**}} \right\|^2 \cdot (num-1)} \quad (22)$$

[0121] 其中, $k=1, 2, \dots, num$, num 表示每一组多个故障组合中故障类型的个数; $expt_k$ 表示第 k 个故障类型下进行实验的个数,即为第 k 个故障类型下修正后的不变曲线矩特征向量的个数,在实施例 $expt_k=7$, $expt$ 为不同故障类型下所有实验的个数,有:

$$expt = \sum_{k=1}^{num} expt_k; \alpha_{kg}^{**} \text{表示第} k \text{个故障类型下第} g \text{个修正后的不变曲线矩特征向量的区}$$

间数据的均值估计的无偏估计值, $g=1, 2, \dots, expt_k$; $\overline{\alpha_k^{**}}$ 表示第 k 个故障类型下所有修正后的不变曲线矩特征向量的区间数据的均值估计的无偏估计值的平均值,有:

$$\overline{\alpha_k^{**}} = \frac{1}{expt_k} \sum_{g=1}^{expt_k} \alpha_{kg}^{**}; \overline{\alpha_{all}^{**}} \text{表示所有故障类型下的所有修正后的不变曲线矩特征向}$$

量的区间数据的均值估计的无偏估计值的总平均值,有: $\overline{\alpha_{all}^{**}} = \frac{1}{expt} \sum_{k=1}^{num} \sum_{g=1}^{expt_k} \alpha_{kg}^{**}$; Sum_1

表示 $\overline{\alpha_k^{**}}$ 与 $\overline{\alpha_{all}^{**}}$ 的差异之和,有: $Sum_1 = \sum_{k=1}^{num} expt_k \left\| \overline{\alpha_k^{**}} - \overline{\alpha_{all}^{**}} \right\|$; Sum_2 表示 α_{kg}^{**} 与 $\overline{\alpha_k^{**}}$ 的

差异之和,有: $Sum_2 = \sum_{k=1}^{num} \sum_{g=1}^{expt_k} \left\| \alpha_{kg}^{**} - \overline{\alpha_k^{**}} \right\|$ 。

[0122] 子步骤920,对每一组多个故障组合的F指标值进行分析。F指标值越小,说明该组多个故障类型组合中的各故障类型之间的距离越小,各故障类型之间的差异小,确定具有最小F指标值的多个故障组合为所采集的井下泵功图所具有的多故障类型。

[0123] 下面结合一个具体实例对本发明进行具体说明:

[0124] 1、由有杆泵抽油井地面示功图无线远程采集系统采集地面示功图,如图2所示。

[0125] 2、将采集到的地面示功图转化为井下泵功图,如图3所示。

[0126] 3、对转化后的井下泵功图进行归一化处理,归一化到区间 $[0, 1]$,然后进行特征提取,根据公式(1)-(11)提取其7个不变曲线矩特征向量,如表1所示。

[0127] 表1所采集井下泵功图的7个不变曲线矩特征值

[0128]

不变曲线矩 特征向量	δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7
	0.36	2.07	1.51	3.38	2.30	2.25	15.98

[0129] 4、根据公式 (12)，由计算得到的不变曲线矩特征向量建立所采集的井下泵功图的物元模型，如下：

$$[0130] \quad M = \begin{bmatrix} \text{井下工况} & \delta_1 & 0.36 \\ & \delta_2 & 2.07 \\ & \vdots & \vdots \\ & \delta_7 & 15.98 \end{bmatrix}$$

[0131] 5、由油田生产中所记录的已知不同故障类型的地面示功图样本建立标准特征库，得到每一种故障类型的物元模型，如表2所示，其中：所建立的标准特征库包含11种故障类型，分别为：“正常”、“气体影响”、“供液不足”、“抽油杆断落”、“油稠”、“游动阀漏失”、“泵上碰”、“泵下碰”、“固定阀漏失”、“油井出砂”、“柱塞脱出工作筒”，由 W_1 – W_{11} 表示。

[0132] 表2标准特征库各故障类型的不变曲线矩特征向量取值区间

[0133]

不变曲线矩特征向量	δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7
W_1	[0.05,0.45]	[0.87,2.45]	[0.65,1.75]	[1.31,3.56]	[0.51,2.35]	[1.31,2.11]	[4.33,5.81]
W_2	[0.27,0.39]	[0.7,0.91]	[1.52,1.76]	[3.75,5.19]	[0.48,0.82]	[1.69,1.99]	[4.5,5.06]
W_3	[0.22,0.36]	[0.79,2.88]	[1.31,1.73]	[2.8,5.58]	[0.61,2.32]	[1.35,5.93]	[3.31,5.98]
W_4	[0.06,0.58]	[1.13,1.44]	[0.52,0.96]	[1.96,3.87]	[0.56,0.9]	[1.5,2.42]	[1.99,2.49]
W_5	[0.03,0.05]	[0.67,2.56]	[0.65,1.15]	[1.31,2.31]	[0.53,2.55]	[1.31,2.24]	[4.67,5.67]
W_6	[0.05,0.68]	[1.14,3.02]	[0.59,1.08]	[1.26,2.2]	[1.5,2.67]	[1.27,4.37]	[2.93,4.94]
W_7	[0.02,0.1]	[0.94,3.98]	[0.67,1.92]	[1.44,4.42]	[1.38,2.77]	[1.3,2.33]	[4.28,5.65]
W_8	[0,0.02]	[1.27,4.09]	[0.71,1.28]	[1.47,2.56]	[0.73,1.99]	[1.15,3.32]	[3.17,4.33]
W_9	[0.06,0.64]	[1.16,3.08]	[0.63,1.35]	[1.27,3.33]	[0.46,1.47]	[2.57,5.21]	[3.82,6.37]
W_{10}	[0.08,0.45]	[0.29,1.14]	[0.33,2.17]	[0.68,4.53]	[0.53,1.92]	[0.81,3.57]	[4.77,5.92]
W_{11}	[0.05,0.15]	[0.84,1.94]	[0.66,1.61]	[1.33,3.25]	[0.35,2.47]	[1.1,1.92]	[3.88,5.66]

[0134] 6、根据公式 (14) – (16) 计算所采集的井下泵功图与标准特征库中各故障类型的关联度，如表3所示。

[0135] 表3所采集井下泵功图与标准特征库中各故障类型的关联度

[0136]

	$\lambda(W_1)$	$\lambda(W_2)$	$\lambda(W_3)$	$\lambda(W_4)$	$\lambda(W_5)$	$\lambda(W_6)$	$\lambda(W_7)$	$\lambda(W_8)$	$\lambda(W_9)$	$\lambda(W_{10})$	$\lambda(W_{11})$
关 联 度	-0.258	0.316	0.643	0.151	-0.274	0.206	-0.244	-0.119	-0.103	-0.296	-0.073

[0137] 根据表3,所采集的井下泵功图与第 W_2 、 W_3 、 W_4 和 W_6 种故障类型的关联度大于等于0,说明所采集的井下泵功图有可能具有的故障类型是“气体影响”、“供液不足”、“抽油杆断落”和“游动阀漏失”。

[0138] 7、对 $\lambda(W_j) \geq 0$ 的各类型故障进行多故障组合的分类,根据表3结论,有: $\lambda(W_3) > \lambda(W_2) > \lambda(W_6) > \lambda(W_4) > 0$,认为具有最大值的 $\lambda(W_3)$ 所对应的该类型的故障是必然发生的,该类型故障为:“供液不足”。将多个故障进行组合分类,分别为: $\{W_3, \beta\}$, $\{W_3, W_2, \beta\}$, $\{W_3, W_4, \beta\}$, $\{W_3, W_6, \beta\}$, $\{W_3, W_2, W_4, \beta\}$, $\{W_3, W_2, W_6, \beta\}$, $\{W_3, W_4, W_6, \beta\}$, $\{W_3, W_2, W_4, W_6, \beta\}$, 其中, β 为所采集的井下泵功图样本。将一组多个故障的组合看作一类,分别由 Z_1 - Z_8 表示。

[0139] 8、根据公式(17)-(22),计算 Z_1 - Z_8 的F指标值,如表4所示,其中: $\theta_1(v_a, v_b) = 0$ 、 $\theta_2(v_a, v_b) = (v_b - v_a) \cdot e^{-v_b}$ 、 $\theta_3(v_a, v_b) = (v_b - v_a) \cdot (2 - v_b) \cdot e^{-a}$ 。

[0140] 表4 Z_1 - Z_8 各组合分类的F指标值

[0141]

	Z_1	Z_2	Z_3	Z_4	Z_5	Z_6	Z_7	Z_8
F指标值	22.5	14.2	19.7	22.6	21.3	25.2	23.4	20.8

[0142] Z_2 分类组合 $\{W_3, W_2, \beta\}$ 的F指标值最小,说明所采集的井下泵功图与标准特征库中的第 W_2 和 W_3 种故障类型之间的差异小,确定所采集的井下泵功图具有“供液不足”和“气体影响”两种故障。

[0143] 本领域内的技术人员应明白,本发明的实施例可提供为方法、系统、或计算机程序产品。因此,本发明可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且,本发明可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0144] 本发明是参照根据本发明实施例的方法、设备(系统)、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器,使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

[0145] 这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中,使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品,该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。

[0146] 这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上,使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理,从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

[0147] 尽管已描述了本发明的优选实施例,但本领域内的技术人员一旦得知了基本创造性概念,则可对这些实施例作出另外的变更和修改。所以,所附权利要求意欲解释为包括优选实施例以及落入本发明范围的所有变更和修改。

[0148] 显然,本领域的技术人员可以对本发明进行各种改动和变型而不脱离本发明的精神和范围。这样,倘若本发明的这些修改和变型属于本发明权利要求及其等同技术的范围之内,则本发明也意图包含这些改动和变型在内。

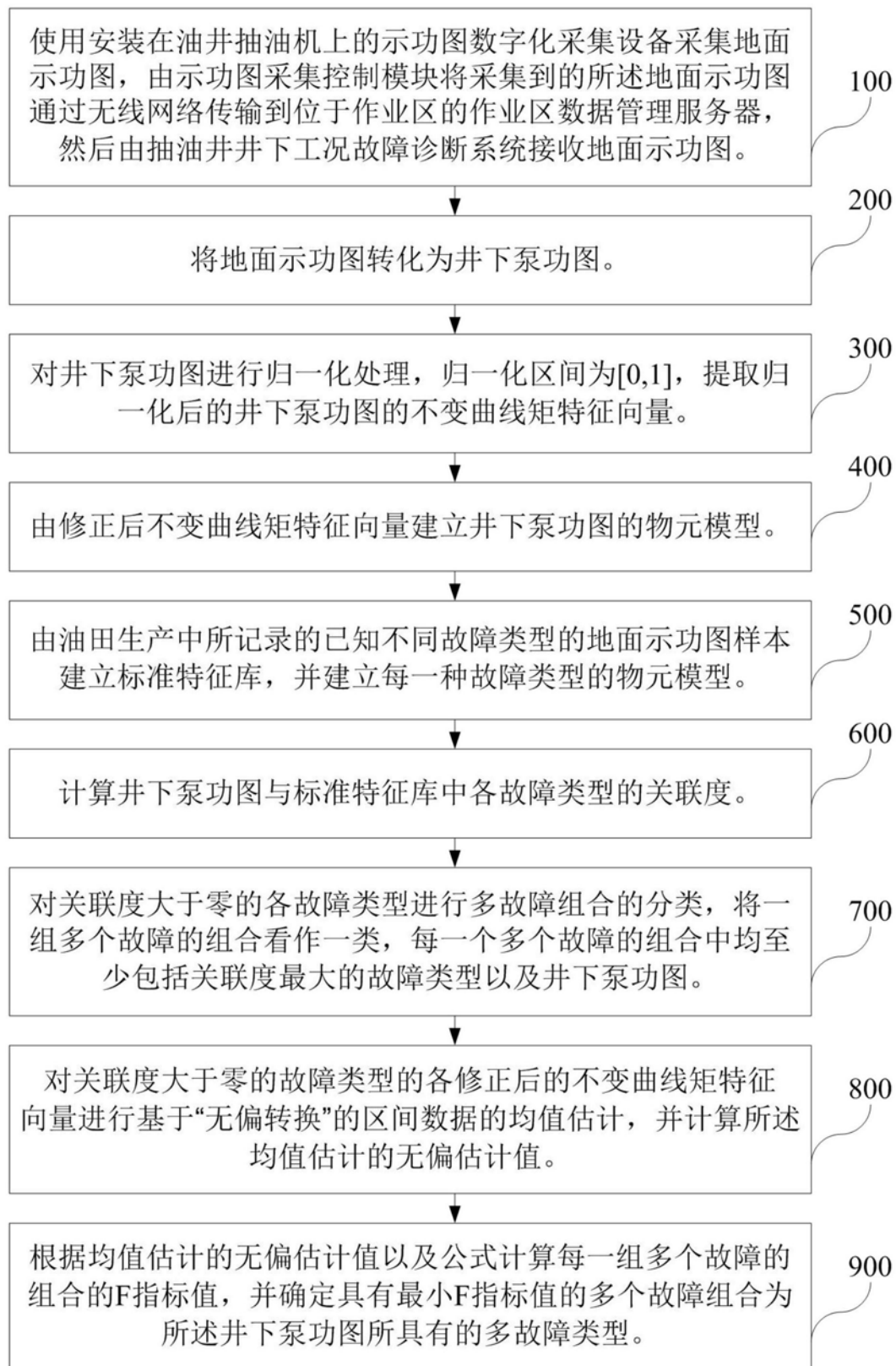


图1

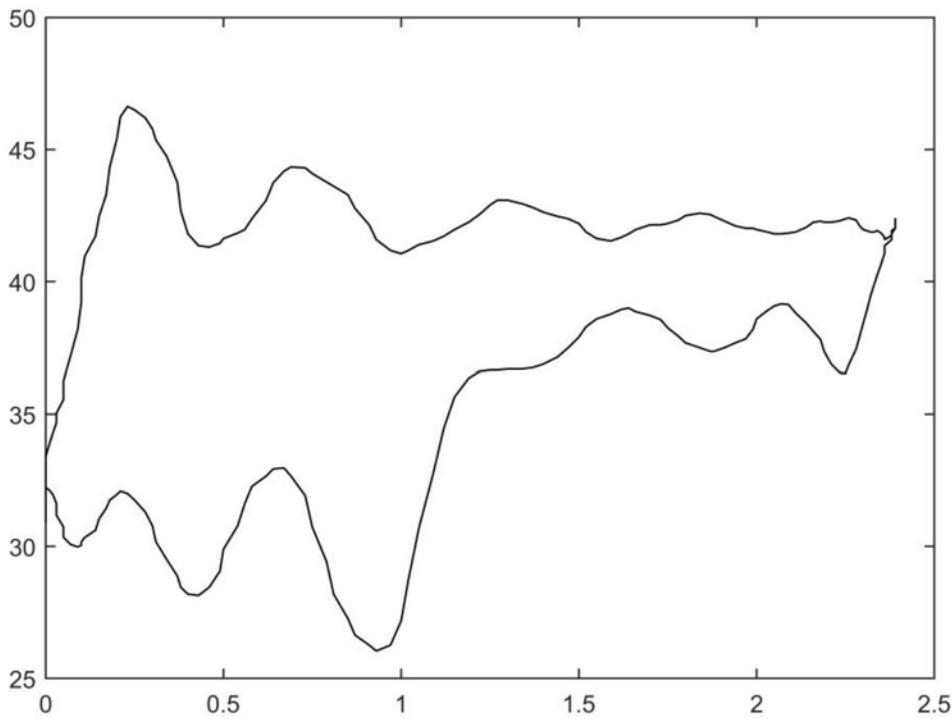


图2

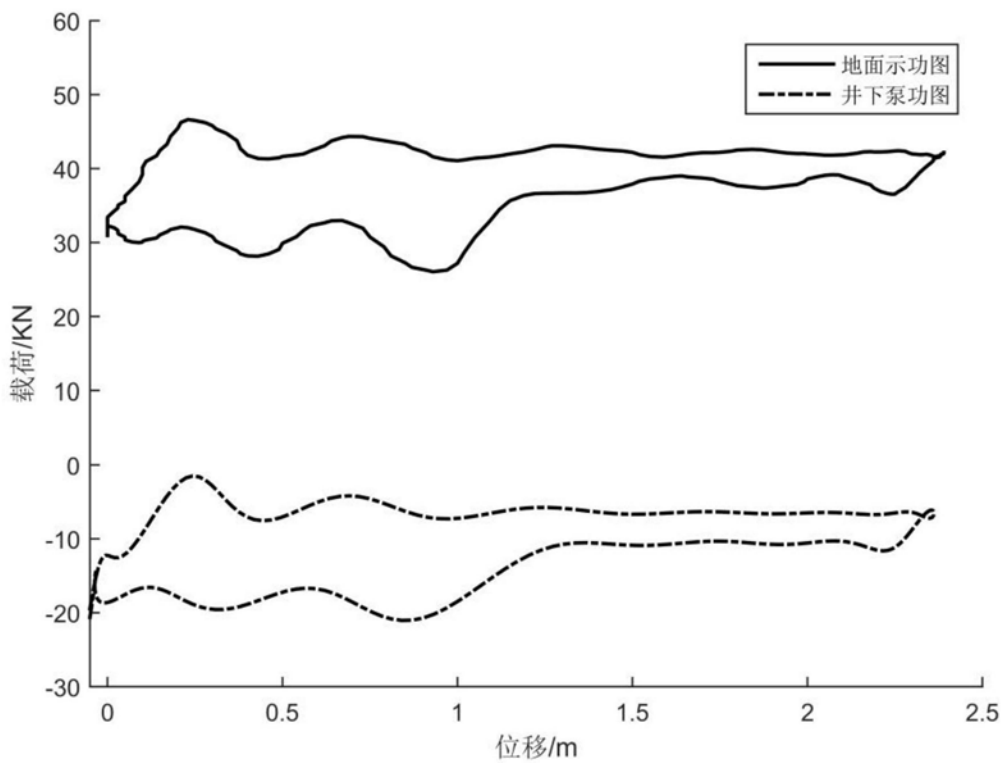


图3