

基于抽油机工况的特征提取与建立和全域故障识别

任泰珠 樊军 蒋夏新

(新疆大学机械工程学院, 乌鲁木齐 830017)

摘要: 针对目前抽油机井下工况故障特征分类任务难以解决,使得所建立诊断模型适应性差且识别率低的问题,通过对抽油机阀门和抽油杆运动状态的分析,首先将示功图进行数学形态学预处理;然后提出阀门开闭点获取和载荷变化特征获取的两种方法,提取到抽油机全域故障的54个全新特征,建立了抽油机工况的特征库;最后运用决策树、Logistic回归和支持向量机算法,验证了在不同工况下,该特征库均具有较好的分类效果,评估了不同故障的工况特征指标,得到各工况私有规则库。研究表明,提取的特征能够有效识别出抽油机全域故障,并且具有较高的识别精度。

关键词: 抽油机; 示功图; 特征提取; 故障识别; 阀门

中图分类号: TE933 **DOI:** 10.16579/j.issn.1001.9669.2025.01.002

0 引言

石油是我国重要的战略资源,抽油机井在石油开采中具有重要地位。准确了解油井采油状况对于抽油机工况诊断和油田整体经济效益的提升具有重要意义^[1-4]。示功图是判断抽油机工作状况的主要依据,如若不能准确及时识别抽油机故障类型,就会影响油田的安全高效生产,甚至造成生产安全事故^[5]。因此,亟需对抽油机工况进行特征提取和全域故障识别。

不同抽油机故障导致不同形状的示功图,而每个故障的示功图都具有独有的特征。具体表现为阀门开闭点位置变化、油杆的运动状况、载荷的变化等。对于示功图的诊断识别方法研究,主要有人工诊断识别方法和计算机智能诊断识别方法两种。人工诊断识别方法是根据油田专业技术人员多年的经验,通过对实测示功图的判断和识别,从而判别抽油机的故障类型。计算机智能诊断识别方法^[6]主要是通过提取图形特征作为分类器的分类依据来判断工况类别。示功图特征提取的方法主要包括傅里叶变换^{[7]343-355}^[8]、差分曲线法^{[9]96-98}、网格法^{[10]265-1268}、Freeman链码法^{[11]4788-4796}^[12-13]、神经网络算法^{[14]2559-2573}^[15-18]等。LI等^{[7]343-355}将示功图的傅里叶描述子作为特征与图形特征进行融合,增强了诊断模型的有效性;于泳滢^{[9]96-98}提出了一种结合几何参数法、差分曲线法和生产参数的方法,用于抽油机井故障的诊断;许广繁等^{[10]265-1268}将网格法和灰色理论相结合,建立了示功图的6个特征值用来识别抽油机

的工况种类;周斌等^{[11]4788-4796}使用八方向链码来提取示功图的轮廓特征,精确识别了阀门的开闭点;YE等^{[14]2559-2573}将卷积神经网络算法改进和简化,使得示功图判断及时准确。以上方法都是从图形的角度出发,没有与抽油机的工作原理相结合,未能直接反映出在抽油机工作过程中阀门、抽油泵、抽油杆等之间的关系,因此只能对某一部分故障有效识别,对全域故障没有普适性。

为解决上述问题,在示功图上获取能够真实反映工况的特征。本文首次提出通过机制方式提取特征,先对抽油杆进行工作原理和受力分析,并加入增载、卸载和上下冲程阶段阀门开合对载荷的影响,确定了阀门开合过程中载荷特征和变化趋势;然后对图形处理获取阀门开合位置和各阶段载荷变化,获取到全新的54个特征,建立全域故障特征库;最后利用决策树、Logistic回归和支持向量机验证提取特征的准确性和有效性,并评估不同故障的工况特征指标。

1 机制描述

1.1 抽油机工作原理

抽油机抽油泵的工作过程如图1所示。在上冲程阶段(A→C段),柱塞和液柱在抽油杆的作用下向上运动,抽油杆受力沿轴线方向向上,抽油杆弹性伸长,载荷增加,固定阀打开,到达上死点后关闭,对应图2的C点(上死点),AB段为抽油杆弹性伸长阶段,BC段为柱塞上行程阶段;在下冲程阶段(C→A段),除了抽油

收稿日期: 2023-06-03 修回日期: 2023-07-07

基金项目: 国家自然科学基金项目(11462021)

作者简介: 任泰珠,女,2000年生,陕西渭南人,硕士研究生;主要研究方向为图像处理和故障识别;E-mail:lolly0312@163.com。

樊军(通信作者),男,1965年生,山东青岛人,副教授,硕士研究生导师;主要研究方向为采油装备及工艺、模式识别和计算机图形学等;E-mail:xj_fanjun@163.com。

引用格式: 任泰珠,樊军,蒋夏新.基于抽油机工况的特征提取与建立和全域故障识别[J].机械强度,2025,47(1):12-19.

REN Taizhu, FAN Jun, JIANG Xiixin. Feature extraction and establishment based on pumping unit working conditions and global fault identification[J]. Journal of Mechanical Strength, 2025, 47(1): 12-19.

杆的自重之外,其他力都朝着轴线方向向上。因此,上部分的杆柱受拉力,而下部分的杆柱受压力。在这个过程中,抽油杆会发生弹性收缩,从而导致载荷降低。同时,游动阀会打开,然后在达到下死点后关闭,对应图 2 的 A 点(下死点),CD 段为抽油杆弹性收缩阶段,DA 段为柱塞下行程阶段,如此周期运行。示功图是了解油井工作状况的主要方法,通过定量分析可以促进抽油机的故障诊断和特征分类识别。由图 2 可以看出,示功图的图形表现为一标准的平行四边形,但实际现场工作过程中,由于作业地理环境等的因素十分复杂,固定阀、游动阀受到水、蜡、砂、气和惯性、振动、冲击载荷等的影响,其打开和关闭的时间会出现差错、程度也会受到影响,此时 B、D 点的位置就会发生变化,加之抽油杆在上下运动中又会受到摩擦力的作用,就会出现各种各样的异常情况,BC 段和 DA 段会出现不同程度的波动。最终,杆的影响因素、泵故障以及原油因素等多种影响因素耦合起来联合反馈到悬点载荷中。

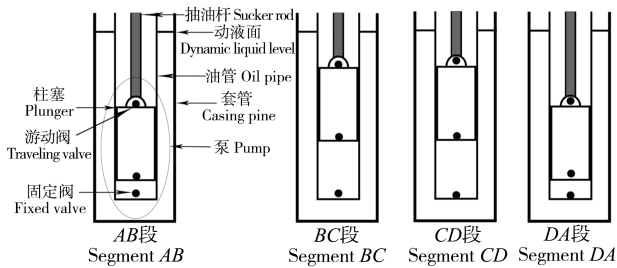


图 1 抽油泵工作原理

Fig. 1 Working principle of the oil well pump

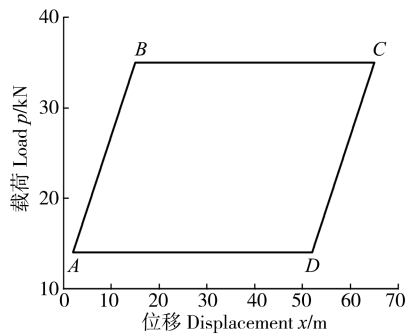


图 2 理论示功图

Fig. 2 Theoretical dynamometer card

1.2 抽油杆受力及运动分析

杆管液耦合动力学模型中的抽油杆运动方程^[19]为

$$\rho H \frac{\partial^2 u}{\partial t^2} = EH \frac{\partial^2 u}{\partial s^2} + \rho Hg - F_n - F_r \quad (1)$$

式中, ρ 为抽油杆材料密度, kg/m^3 ; H 为抽油杆横截面积, m^2 ; u 为抽油杆单元节点位移, m ; t 为时间, s ; s 为抽油杆节点位置, m ; g 为自由落体加速度, m/s^2 ; F_n 为油管对单位长度抽油杆摩阻力, N ; F_r 为井液对单位长度抽油杆摩阻力, N ; E 为抽油杆弹性模量, Pa 。

泵阀运行特性函数 Φ ^[20]为

$$\Phi = v_p(t) - v_b(t) \quad (2)$$

式中, v_p 为柱塞速度, m/s ; v_b 为泵筒速度, m/s 。

杆管环空井液底端流速函数 Ψ ^[21]为

$$v_f(L_p, t) = \begin{cases} \frac{(H_u - H_p)v_b(t) + (H_p - H_m)v_p(t)}{H_m}, f_d = 0 \\ \frac{H_u v_b(t) - H_m v_p(t)}{H_m}, f_d = 1 \end{cases} \quad (3)$$

式中, L_p 为抽油杆柱的总长度, m ; H_u 为油管横截面内圆面积, m^2 ; H_m 为柱塞横截面积, m^2 ; H_p 为最后一级抽油杆横截面积, m^2 ; F_d 为最后一级杆管环空井液的横截面积, m^2 ; f_d 为游动阀的状态,0表示阀关闭,1表示阀打开。

式(1)~式(3)根据抽油杆、油管、液面的受力、石油进出泵筒速度及泵筒内井液压力的变化规律建立了抽油杆运动方程、泵阀运行函数和井液底端流速函数,表明了抽油杆载荷不但会受到阻力、流速的影响,还会受到阀门状态和底部流量的影响。同时,GIBBS^[22]通过波动方程和抽油杆运动边界条件描述了抽油杆运动、受力、抖动和井下泵的运动条件,石油负载会影响在抽油杆柱内传播力波的形状,并可能导致施加在抽油杆上的最大和最小载荷发生显著变化,石油载荷的影响和抽油杆的惯性效应被集中到所谓的动态载荷中耦合起来反馈到示功图中。

2 特征的提取与建立

2.1 载荷波动影响的去除

原始示功图曲线由 120 个点组成,将其分为上、下冲程两部分,为了去除抽油杆振动和冲击载荷波动的影响,分别对上下冲程的 60 个点进行数学形态学操作。式(4)~式(7)为用结构元素 B 对图形 A 进行处理^[23],腐蚀操作、膨胀操作、开运算和闭运算分别定义如下:

$$A \ominus B = \{x, y | (B)xy \subseteq A\} \quad (4)$$

$$A \oplus B = \{x, y | (B)xy \cap A\} \neq \emptyset \quad (5)$$

$$A \circ S = (A \ominus B) \oplus B \quad (6)$$

$$A \bullet S = (A \oplus B) \ominus B \quad (7)$$

开运算可以去除示功图中由于载荷波动而形成的孤立小点和毛刺,闭运算可以填补示功图中的小孔并弥合裂缝。本文选用形状分别为 3×3 、 5×5 和 7×7 的正方形作为结构元素,以正方形的中点作为结构元素的中心,用 3 种不同大小的结构元素分别对示功图进行开运算和闭运算操作。利用式(6)、式(7),以上冲程为例,使用 3 种不同大小的结构元素进行开运算和闭运算处理,得到最终处理后的图形,处理结果可分别去除阻力和流速、波动以及振动影响,如图 3 所示。

2.2 阀门属性特征的获取

阀门属性特征的获取首先要找出 B、D 点,在进行 B、D 点获取之前要给出理论密度和实际密度的定义。

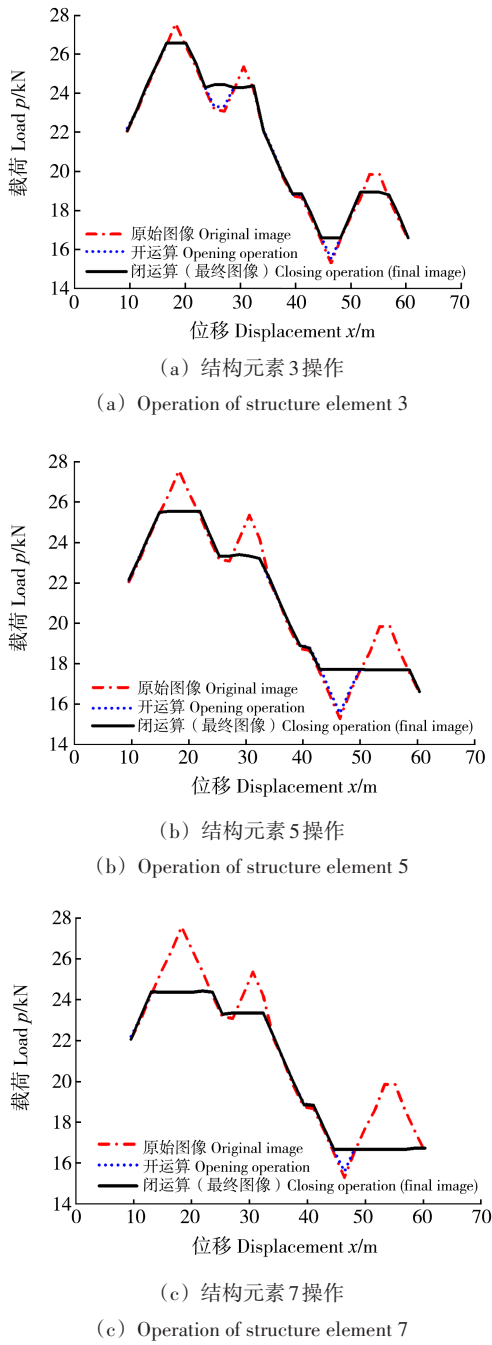


图3 不同结构元素处理图形

Fig. 3 Handle graphics of different structural elements

如图4所示,以最左点和最右点形成的直线作为最上层,最低点所在的纵坐标为最下层,得出理论密度;以原始图形的曲线作为最上层,最低点所在的纵坐标为最下层,得出实际密度。

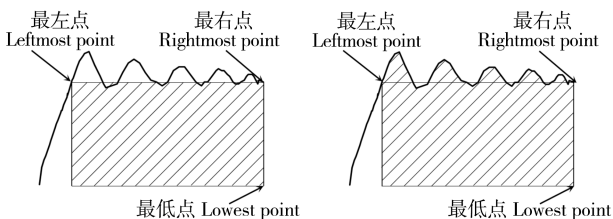


图4 理论密度和实际密度

Fig. 4 Theoretical density and actual density

分别对3种不同结构元素消除载荷影响后的图形进行如下操作,如图5所示(以上冲程为例)。

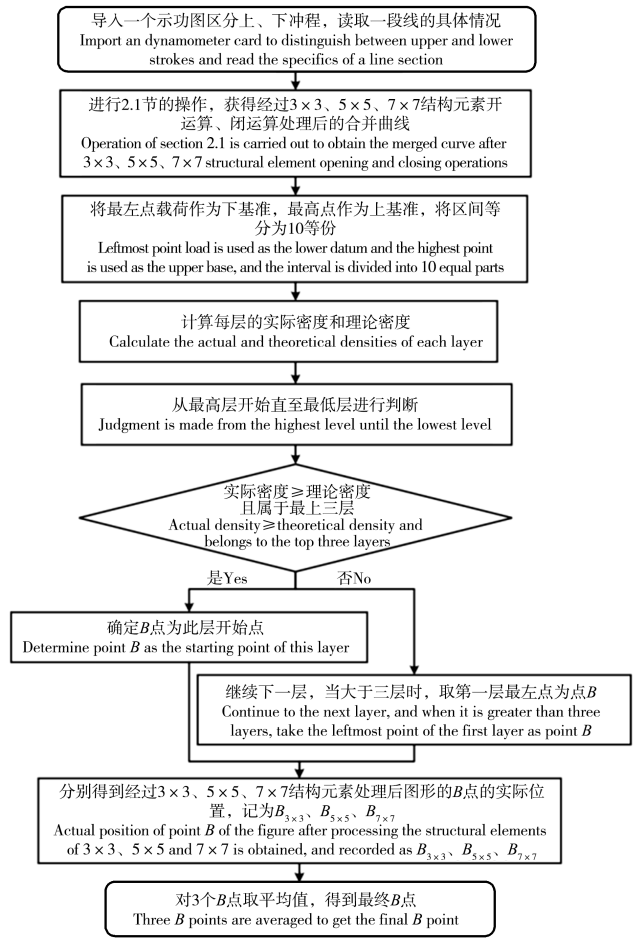


图5 B点的获得

Fig. 5 Acquisition of point B

通过图5的方式将最右点载荷作为最上基准,最低点作为最下基准对下冲程的图形进行处理,可以得到D点。A点和C点机制上对应示功图最左端和最右端的点,最终我们可以获得A、B、C、D点的具体数据信息。图6是得到A、B、C、D四点的示功图图形(以3种类型示功图为例)。

2.3 载荷变化属性特征的获取

针对AB、BC、CD、DA段之间的线段进行微小载荷波动过滤处理,步骤与第2.1节相同,对获取的线段进行如下处理:

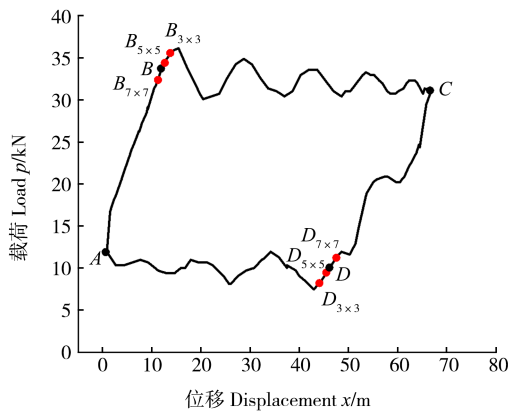
步骤一:对A、B、C、D形成的4段线中每段按点的顺序依次计算相邻两点之间的斜率[式(8)]和斜率的增量[式(9)]。

两点之间的斜率为

$$k_i = \frac{y_{i+1} - y_i}{x_{i+1} - x_i} \quad (8)$$

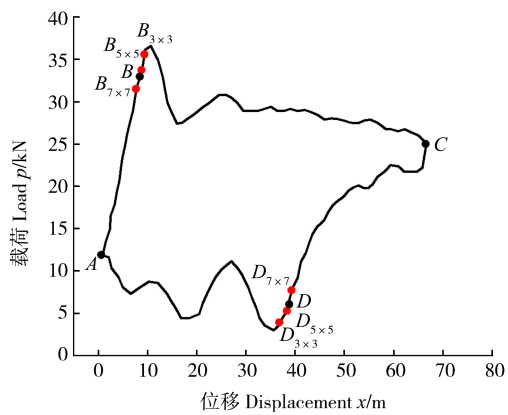
斜率的增量为

$$\Delta k = k_{i+1} - k_i \quad (9)$$



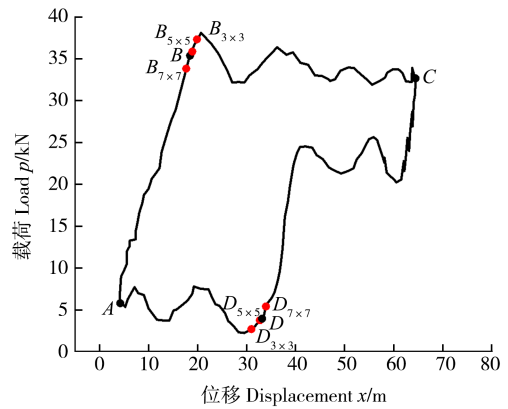
(a) 正常示功图

(a) Normal dynamometer card



(b) 气体影响

(b) Influence the gas



(c) 供液不足

(c) Insufficient liquid supply

图 6 不同故障示功图中 B、D 点的获取

Fig. 6 Acquisition of B, D points in different fault dynamometer cards

步骤二: 根据相邻点斜率和相邻斜率增量的各自比较, 以 15° 为界, 对线段进行分类和比较, 定义出 9 种曲线类型(表 1)。

步骤三: 每段线内从第一个数据开始, 按照步骤一的方法计算斜率和斜率增量, 按照步骤二的方法判断线型, 并且记录每段线的长度、起始点、方向以及线型。

步骤四: 对每段线的数据进行如下操作。

1) 过滤。将所有长高比大于 5 的曲线重新定义为直线。

2) 合并。若直线与直线斜率之差小于 0.3, 重新合并为一条直线; 若直线长度与曲线长度之差小于 0.1, 则重新合并为一条曲线。

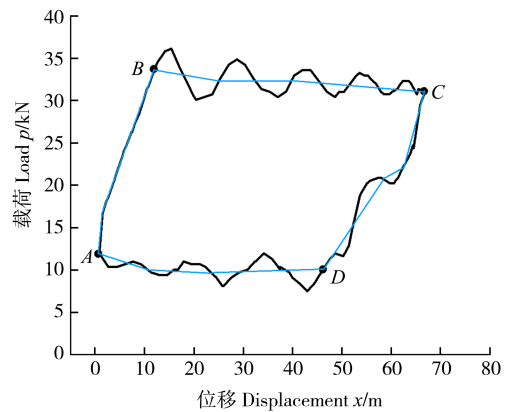
3) 重复。将相邻线依次进行合并, 直至每段的最后一条线。

表 1 线型

Tab. 1 Line type

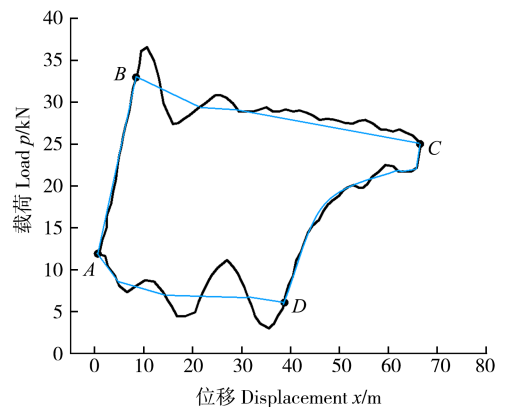
线型 Line type	判断条件 Judgment conditions
向上凹弧 Upward concave arc	$k_i > 0$ $k_{i+1} > 0$ $\Delta k > 0$
向上凸弧 Upward convex arc	$k_i > 0$ $k_{i+1} > 0$ $\Delta k < 0$
向上直线 Straight up line	$k_i > 0$ $k_{i+1} > 0$ $\Delta k = 0$
向下凹弧 Downward concave arc	$k_i < 0$ $k_{i+1} < 0$ $-\Delta k < 0$
向下凸弧 Downward convex arc	$k_i < 0$ $k_{i+1} < 0$ $-\Delta k > 0$
向下直线 Straight down line	$k_i < 0$ $k_{i+1} < 0$ $\Delta k = 0$
水平凹弧 Horizontal concave arc	$k_i = 0$ $k_{i+1, i+2, \dots} > 0$ $\Delta k_{i+1} > 0$
水平凸弧 Horizontal convex arc	$k_i = 0$ $k_{i+1, i+2, \dots} < 0$ $-\Delta k_{i+1} > 0$
水平线 Horizontal line	$k_i = 0$ $k_{i+1} = 0$ $\Delta k = 0$

步骤五: 将 AB、BC、CD、DA 每段线总的线条数合并为 3 个, 若线型的数量超过 3 个, 将最短那条线取消, 直至数量为 3; 如不够 3 条, 在此段线的最后补数据均为 0 的线型, 直至补齐 3 个。图 7 所示为经过线条合并后所得到的图形(以 3 种类型示功图为例)。



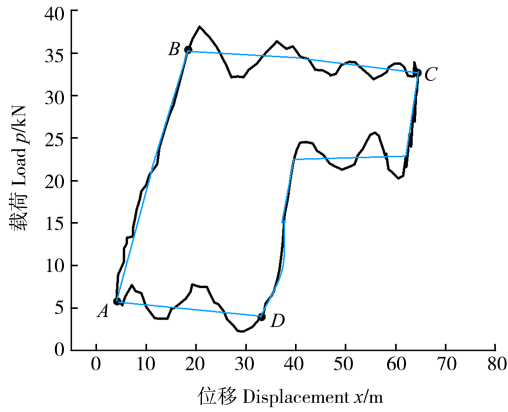
(a) 正常示功图线型

(a) Line type of the normal dynamometer card



(b) 气体影响线型

(b) Line type influenced by the gas



(c) 供液不足线型
(c) Line type influenced by the insufficient liquid supply

图7 不同故障示功图线型
Fig. 7 Line type of different fault dynamometer cards

2.4 抽油机工况特征库的建立

由于抽油机在工作过程中作业环境的复杂,阀门打开和关闭的时间会受到影响,B、D点的位置也会发生变化,加之抽油杆在上下运动中又会受到摩擦力的作用,就会出现各种各样的异常情况,这些都会反映在位移-载荷图像中。依据第2.1~2.3节中所得到的A、B、C、D四点在机制上对应载荷的最大点和最小点,阀门打开和关闭的点,这四点不同的位置可以形成不同的四边形,并且在相同的四边形中边长、角度也会有所不同,可以得到不同的信息;相邻的每3个点可构成对应边长、角度大小不一的角,不同故障所形成角的大小、边长都有所不同,AB反映增载情况,CD反映卸载情况,BC、DA反映抽油杆在运动过程中的力学特征对应轨迹,角度的大小反映了阀门的打开程度,不同故障对应的角度具有不同的属性;相邻每两点之间形成的线型

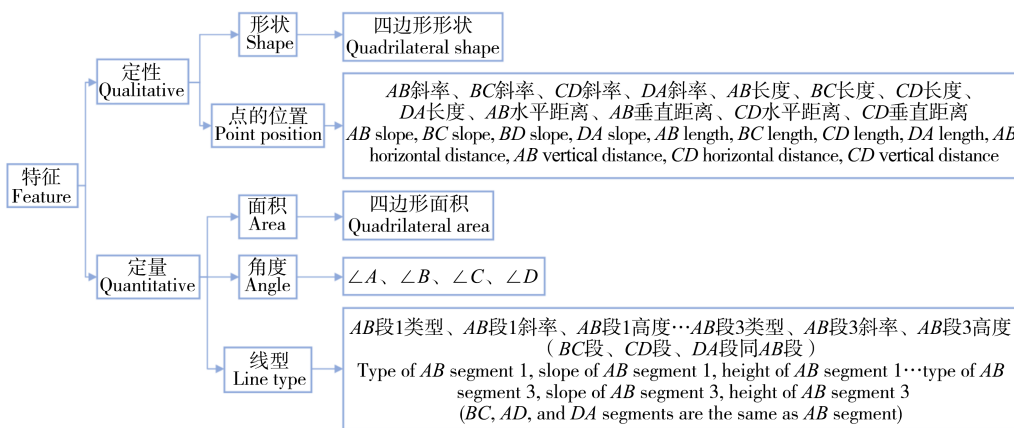


图9 全域特征库

Fig. 9 Global feature library

针对上述11种工况,通过Logistic回归、决策树和支持向量机对每种工况进行二分类,验证本文建立属性特征的全局意义和有效性,并评判每种工况不同特征在故障识别过程中的权重。将数据的70%作为训练集样本,得到二分法模型,并利用训练得到的结果对30%测试集样本进行二分类,验证特征的有效性以

包含了抽油杆波动的影响,经过第2.1节的操作可以获得最真实的加载和卸载特征,这3类信息可以描述出抽油机采油过程中所有变化的属性特征(图8)。

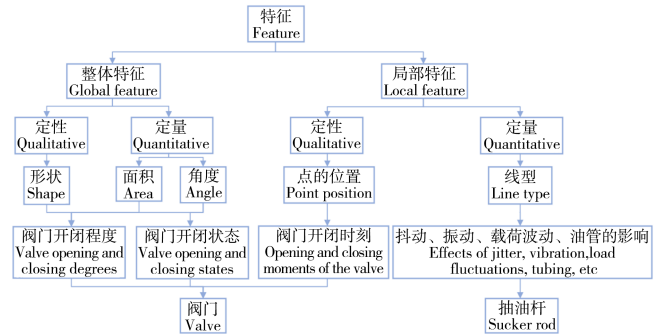


图8 属性特征

Fig. 8 Attribute feature

根据属性特征,从定性和定量两个角度分析,最终建立全域故障的所有特征(54个全新特征),如图9所示。

3 属性特征的有效性验证

3.1 属性数据获取

本文使用来自新疆克拉玛依油田的两个采油区块约120口井2019年3月底到2020年3月底的示功图以及由示功图分析所对应工况类型作为数据来源,其包含的工况主要有正常示功图、上碰泵、供液不足、下碰泵、气体影响、固定凡尔漏失、出砂、游动凡尔漏失、油稠、油井结蜡,双凡尔漏失等11种。原始示功图成图数据按日期存放于单独的文件中,每条数据均由120个坐标点构成。每条数据都经过第2.1、2.2和2.3节算法的处理,最终形成图9所对应特征库里的每条特征数值。

及分类的精确度。具体流程如图10所示。

3.2 实验验证及结果

3.2.1 分类模型

1)将连续数据采用ordinal编码离散化,用分位数把所表现的特征进行四等分,每个特征中每个箱内的样本数量都相同。

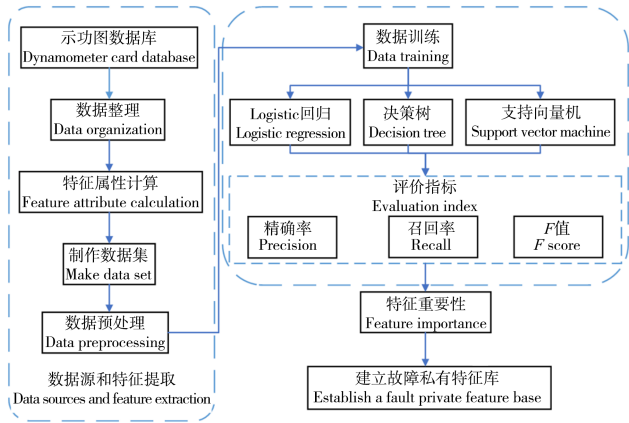


图 10 故障诊断流程图

Fig. 10 Flow chart of the fault diagnosis

2) 决策树开始时根节点包含数据集中的所有样例数据,计算每个特征信息熵,计算数据集中每个特征的信息增益,确定模型的根节点。根据根节点的划分,再依次计算每个决策属性的信息增益,选择最佳节点和最佳的分枝方法,对每一个分枝进行进一步的划分。根据对数据集各特征属性信息熵的计算,生成故障二分类识别的决策树模型。

3) Logistic 回归由线性回归衍生而来,可以解决离

散数据的分类问题,其中引入了 Sigmoid 激活函数将所有计算结果映射到 $[0, 1]$,以 0.5 为界,实现对故障的二分类。

4) 支持向量机将特征映射到一个更高维的空间里找到样本数据之间最优的划分超平面,完成对故障的二分类。

3.2.2 评估指标

本文采用准确率 $P_{\text{Precision}}$ (Precision rate)、召回率 R_{Recall} (Recall rate) 和平衡 F_1 分数 $F_{1\text{-score}}$ (Balance F_1 score) 3 个指标对模型的性能进行评价,为

$$P_{\text{precision}} = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \quad (10)$$

$$R_{\text{recall}} = \frac{T_P}{T_P + F_N} \quad (11)$$

$$F_{1\text{-score}} = 2 \times \frac{P_{\text{precision}} \times R_{\text{recall}}}{P_{\text{precision}} + R_{\text{recall}}} \quad (12)$$

式中, T_P 、 T_N 、 F_P 和 F_N 分别为真阳性、真阴性、假阳性、假阴性。

3.2.3 不同故障模型性能评估

利用第 3.1 节所得到的属性数据集在 3 种算法模型上的评估结果如表 2 所示。

表 2 实验结果

Tab. 2 Experimental results

故障类型 Fault type	决策树 Decision tree			Logistic 回归 Logistic regression			支持向量机 Support vector machine		
	准确率 Precision rate $P_{\text{precision}} / \%$	召回率 Recall rate $R_{\text{recall}} / \%$	平衡 F_1 分数 Balance F_1 score $F_{1\text{-score}} / \%$	准确率 Precision rate $P_{\text{precision}} / \%$	召回率 Recall rate $R_{\text{recall}} / \%$	平衡 F_1 分数 Balance F_1 score $F_{1\text{-score}} / \%$	准确率 Precision rate $P_{\text{precision}} / \%$	召回率 Recall rate $R_{\text{recall}} / \%$	平衡 F_1 分数 Balance F_1 score $F_{1\text{-score}} / \%$
	正常示功图 Normal dynamometer card	96.8	96.9	95.79	94.64	95.9	93.29	95.84	96.31
供液不足 Insufficient liquid supply	95.75	95.75	93.92	94.53	95.46	93.32	95.04	95.67	93.75
气体影响 Gas influence	95.87	95.87	96.23	94.74	94.89	95.74	95.16	95.09	95.56
固定凡尔漏失 Suction valve leakage	90.75	91.68	91.63	92.63	92.83	93.72	91.37	92.26	92.63
游动凡尔漏失 Delivery valve leakage	95.68	95.92	96.32	94.37	94.69	95.13	95.16	95.27	95.79
双凡尔漏失 Double valve leakage	95.25	95.33	95.89	93.54	93.62	93.85	94.33	94.58	94.85
油稠 Oil consistency	90.85	91.27	90.49	90.28	92.57	90.94	93.11	94.97	93.13
油井结蜡 Oil well waxing	95.2	95.54	95.64	93.34	95.93	94.65	94.87	95.78	94.82
出砂 Sand production	93.32	94.23	93.98	92.27	92.63	93.37	93.12	93.17	93.65
上碰泵 Top pump bumping	92.35	93.34	92.83	93.78	94.95	93.62	92.98	93.93	92.73
下碰泵 Bottom pump bumping	91.3	93.27	93.42	90.24	91.98	92.77	91.28	90.88	91.78

由表 2 可知,在分别对 11 种不同的故障数据集进行测试时,所确定出的特征属性在 3 种算法运行后都有较高精度指数,实验结果证明,本文建立的全域特

征库具有合理性和正确性。

3.2.4 工况私有规则库的建立

通过 3 种算法对本文第 2 章所述方法建立的全域

特征库进行特征重要性计算及特征选择,分别建立属于上述 11 种故障的私有规则库。以供液不足故障为例,经过多次实验分析,从全域特征库 54 个特征中选

择了 14 个特征建立供液不足私有规则库的特征公共集,具体特征重要性如图 11 所示。

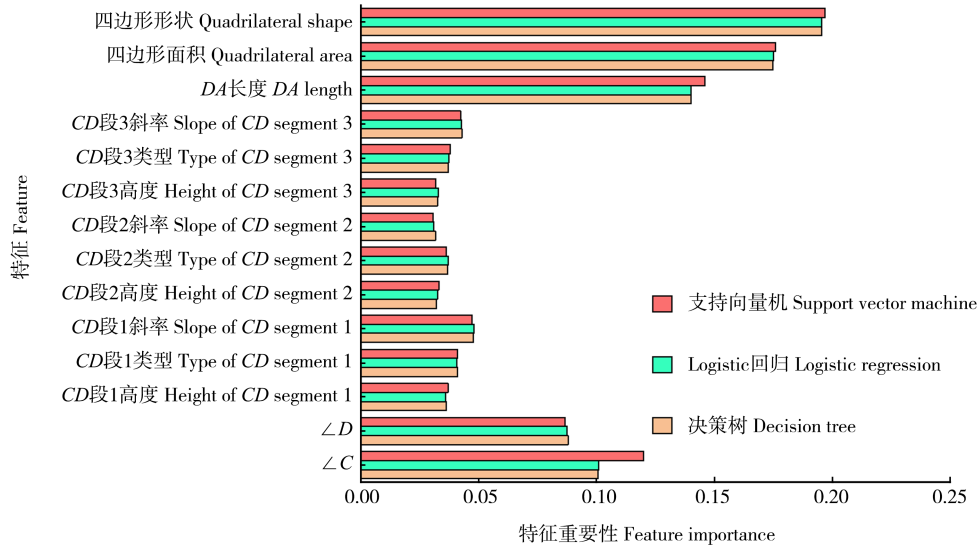


图 11 供液不足特征重要性

Fig. 11 Feature importance of insufficient liquid supply

4 结论

针对抽油机故障特征分类难且故障识别率低的问题,基于抽油机工况建立了示功图全域故障识别特征集,通过应用机器学习算法,对不同特征在不同故障类型中的影响差异进行了深入的对比分析,得出以下结论:

1) 抽油机的每一种故障对全域特征集中的某些特征具有较高敏感度,通过对特征重要性的展示和特征的选择,可以建立针对每种故障的私有规则库。

2) 通过抽油机工况建立的示功图全域故障识别特征集,可为实际生产提供通用的抽油机故障识别工具,且具有较高的准确性。

3) 该特征集可有效地为抽油机各类故障识别提供有效特征,为抽油机生产运行和维护提供了重要的参考依据。

参考文献 (References)

- [1] 冯立涛. 油田抽油机常见故障分析及处理[J]. 化学工程与装备, 2022(1):64-65.
FENG Litao. Analysis and treatment of common faults of oilfield pumping machines [J]. Chemical Engineering & Equipment, 2022 (1):64-65. (In Chinese)
- [2] 杜娟, 刘志刚, 宋考平, 等. 基于卷积神经网络的抽油机故障诊断[J]. 电子科技大学学报, 2020, 49(5):751-757.
DU Juan, LIU Zhigang, SONG Kaoping, et al. Fault diagnosis of pumping unit based on convolutional neural network [J]. Journal of University of Electronic Science and Technology of China, 2020, 49 (5):751-757. (In Chinese)
- [3] 王晓菡. 用于工况诊断的示功图特征提取方法研究[D]. 北京: 中国石油大学, 2011:1-4.

- WANG Xiaohan. Study on indicator card feature extraction method for working condition diagnosis [D]. Beijing: China University of Petroleum, 2011: 1-4. (In Chinese)
- [4] KRASNOV A N, PRAKHOVA M Y. Predictive state analytics of gas pumping units [J]. Journal of Physics: Conference Series, 2022, 2388 (1):012078.
- [5] 陈紫琴. 基于DBN的游梁式抽油机生产工况识别方法研究[D]. 沈阳: 沈阳工业大学, 2022:1-6.
CHEN Ziqin. Research on the identification method of production working condition to swimming beam type pumping machine based on DBN [D]. Shenyang: Shenyang University of Technology, 2022: 1-6. (In Chinese)
- [6] 杨亚勋. 基于多特征融合的示功图诊断方法研究[D]. 大庆: 东北石油大学, 2016:3-4.
YANG Yaxun. Fault diagnosis of indicator diagram based on feature fusion [D]. Daqing: Northeast Petroleum University, 2016: 3-4. (In Chinese)
- [7] LI J N, SHAO J, WANG W, et al. An evolutionary deep learning method based on multi-feature fusion for fault diagnosis in sucker rod pumping system [J]. Alexandria Engineering Journal, 2023, 66: 343-355.
- [8] LUKASZ L, LECH J S. The Fourier transform as a new approach of evaluating the internal combustion engine indicator diagram [J]. Journal of KONES, 2017, 24(4):173-180.
- [9] 于泳澎. 抽油机井示功图自动诊断方法的改进[J]. 化学工程与装备, 2017, 7:96-98.
YU Yongpeng. Improvement of automatic diagnosis method of pumping well dynamometer card [J]. Chemical Engineering & Equipment, 2017, 7:96-98. (In Chinese)
- [10] 许广繁, 杨善林. 灰色理论在有杆抽油井故障诊断中的应用[J]. 合肥工业大学学报(自然科学版), 2013, 36(10):1265-1268.
XU Guangfan, YANG Shanlin. Application of gray theory in fault diagnosis of rod-pumped well [J]. Journal of Hefei University of Technology (Edition of Natural Science), 2013, 36 (10) : 1265-

1268. (In Chinese)
- [11] 周斌,邢林林,张敏. 基于八链码形状匹配的抽油机井地面示功图凡尔开闭点识别[J]. 科学技术与工程, 2022, 22(12): 4788-4796.
ZHOU Bin, XING Linlin, ZHANG Min. Valve opening and closing points recognition in surface dynamometer cards from sucker-rod pumping wells based on 8-direction chain code and shape matching [J]. Science Technology and Engineering, 2022, 22(12): 4788-4796. (In Chinese)
- [12] HAN Y, LI K, GE F W, et al. Online fault diagnosis for sucker rod pumping well by optimized density peak clustering[J]. ISA Transactions, 2022, 120: 222-234.
- [13] 肖维民,梁波. 基于PNN网络和Freeman链码的抽油机井工况诊断[J]. 计算机测量与控制, 2015, 23(6): 1885-1887.
XIAO Weimin, LIANG Bo. Working condition diagnosis of pumping unit based on PNN and Freeman chain code[J]. Computer Measurement & Control, 2015, 23(6): 1885-1887. (In Chinese)
- [14] YE Z W, YI Q J. Working-condition diagnosis of a beam pumping unit based on a deep-learning convolutional neural network[J]. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 2022, 236(5): 2559-2573.
- [15] JIANG J X, LI X F. Identification of indicator diagram type in the oil well by BP neural network[J]. IOP Conference Series: Earth and Environmental Science, 2021, 781(2): 022057.
- [16] SHI J F, MA W M, SONG P. Transform electric power curve into dynamometer diagram image using deep recurrent neural network [J]. IEICE Transactions on Information and Systems, 2018, E101. D(8): 2154-2158.
- [17] 张强,朱刘涛,王彦琦. 基于改进胶囊神经网络的示功图诊断模型[J]. 电子科技大学学报, 2022, 51(1): 117-122.
ZHANG Qiang, ZHU Liutao, WANG Yanqi. Indicator diagram diagnosis model based on improved capsule neural network [J]. Journal of University of Electronic Science and Technology of China, 2022, 51(1): 117-122. (In Chinese)
- [18] 吴定海,任国全,王怀光,等. 基于卷积神经网络的机械故障诊断方法综述[J]. 机械强度, 2020, 42(5): 1024-1032.
WU Dinghai, REN Guoquan, WANG Huaiguang, et al. The review of mechanical fault diagnosis methods based on convolutional neural network [J]. Journal of Mechanical Strength, 2020, 42(5): 1024-1032. (In Chinese)
- [19] XING M M. Response analysis of longitudinal vibration of sucker rod string considering rod buckling [J]. Advances in Engineering Software, 2016, 99: 49-58.
- [20] 薛军,陈春坤,杨振迎,等. 抽油杆三维力学模型的建立与应用[J]. 石油化工应用, 2021, 40(11): 75-79.
XUE Jun, CHEN Chunjun, YANG Zhenying, et al. Establishment and application of three dimensional mechanical model of sucker rod [J]. Petrochemical Industry Application, 2021, 40(11): 75-79. (In Chinese)
- [21] 吕孝孝,王早祥,刘延鑫,等. 故障工况下有杆泵采油系统运行特性及示功图研究[J]. 中国石油大学学报(自然科学版), 2020, 44(2): 117-126.
LÜ Xiaoxiao, WANG Hanxiang, LIU Yanxin, et al. Study on operating characteristics and indicator diagram of rod pump production system under fault condition [J]. Journal of China University of Petroleum (Edition of Natural Science), 2020, 44(2): 117-126. (In Chinese)
- [22] GIBBS S G. Predicting the behavior of sucker-rod pumping systems [J]. Journal of Petroleum Technology, 1963, 15(7): 769-778.
- [23] 肖引昌. 数学形态学在数字图像处理中的应用[J]. 集成电路应用, 2023, 40(2): 297-299.
XIAO Yinchang. Application of mathematical morphology in digital image processing [J]. Application of IC, 2023, 40(2): 297-299. (In Chinese)

Feature extraction and establishment based on pumping unit working conditions and global fault identification

REN Taizhu FAN Jun JIANG Xiaxin

(School of Mechanical Engineering, Xinjiang University, Urumqi 830017, China)

Abstract: In response to the challenging task of fault feature classification in the current working conditions of pumping units, which results in poor adaptability and low recognition rate of the established diagnosis model. The dynamometer card was pretreated by mathematical morphology through the analysis of the motion state of the pumping unit valve and the sucker rod. Then, two methods of obtaining valve opening and closing points and load variation characteristics were proposed, and 54 new features of global faults of pumping units were extracted, and the characteristic database of working conditions of the pumping unit was established. Finally, the algorithm of decision tree, logistic regression and support vector machine was used to verify that the feature database has good classification effect under different working conditions. The characteristic indexes of different fault conditions were evaluated, and the private rule database of each working condition was obtained. The research results demonstrate that the proposed features in this study are capable of effectively identifying comprehensive faults in pumping units, exhibiting a high level of recognition accuracy.

Key words: Oil pumping machine; Dynamometer card; Feature extraction; Fault identification; Valve

Corresponding author: FAN Jun, E-mail: xj_fanjun@163.com

Fund: National Natural Science Foundation of China (11462021)

Received: 2023-06-03 Revised: 2023-07-07