

基于大数据分析的多算法融合滤波方法

邵艳鸣, 秦绪祥

(上汽大众汽车有限公司, 上海 201805)

摘要: 某涂装车间调漆间实时采集的废溶剂回收缸数据存在周期性异常波动问题。从异常检测算法、常规滤波算法以及多算法融合方法等测试了不同算法对于该液位曲线的滤波效果。最终的测试结果显示,一种基于大数据分析的多算法融合方法整合了各滤波算法的优点,有效解决了废溶剂液位数据失真问题,为后续的溶剂降耗研究提供了良好的数据基础。同时,该方法可推广至车间类似设备,依托大数据分析处理方法处理液位扰动问题,为后续大数据分析的研究提供高质量数据。

关键词: 大数据分析; 液位滤波; 水性废溶剂

中图分类号: TQ639 **文献标志码:** B **文章编号:** 1007-9548(2025)01-0033-04

Multi-algorithm Fusion Filtering Method Based on Big Data Analysis

SHAO Yan-ming, QIN Xu-xiang

(SAIC-Volkswagen Co., Ltd., Shanghai 201805, China)

Abstract: In response to the periodic abnormal fluctuations in the waste solvent recovery tank data collected in real-time from a paint mixing room in a certain paint workshop, this paper tested different algorithms, including anomaly detection algorithms, conventional filtering algorithms, and multi-algorithm fusion methods, to filter the liquid level curve. The final test results showed that a self-designed multi-algorithm fusion method integrated the advantages of various filtering algorithms and effectively solved the problem of distortion in the water-based waste solvent level data, providing a good data foundation for subsequent research on reducing water consumption in waste solvents. At the same time, this method can be applied to related equipment in various paint workshops, relying on big data analysis methods to deal with liquid level disturbance problems and providing high-quality data for subsequent big data analysis research.

Key words: big data analysis; liquid level filtering; water-based waste solvents

0 引言

某涂装车间调漆间水性废溶剂回收缸的液位数据存在周期性异常波动问题,导致采集的液位数据无法直接用于数据分析。原始数据如不进行处理则无法转化为符合建模要求的过程数据,也就无法应用到后续的开发环节中。具体而言,该车间调漆间的废溶剂缸采用静压式液位计实时测量液位变化数据。静压式液位

计的工作原理是基于应变感应原理将压力信号转化为电信号获得实时的液位数据。其存在的问题是一旦出现非液位变化引起的压力信号改变,系统反馈的液位信号将剧烈扰动,其液位数据也会失真。

图1为水性废溶剂回收缸管路,液位计(10.2)被安装在缸底部的抽料管位置。一旦废溶剂缸的液位达到MAX位置,工人手动启动隔膜泵(4.4)将缸内的废溶剂抽至废料桶内。由于隔膜泵工作时抽料管内流体被迅速抽取,导致管内瞬时出现抽真空状态而影响管内流体的真实压力,这种变化将对液位计的应变片受力产生较大的冲击。实际液位曲线显示在隔膜泵启动的瞬间液位直接从5 000 L掉至1 000 L左右。这部分

收稿日期: 2023-11-28

作者简介: 邵艳鸣(1985—),男,硕士,工程师,主要从事喷涂机器人规划及涂装车间数字化应用等相关工作。E-mail: worijue@163.com。

值检测得分, 图中最上方的蓝色曲线代表实际液位, 中间绿色曲线代表异常得分, 橙色的线段代表异常判定, 1 代表数据正常, -1 代表异常值。

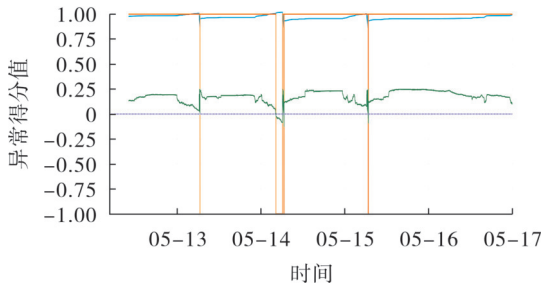


图4 液位曲线及异常值检测得分

该算法实现了大部分异常区域的检测, 但一部分的数据在异常划分上存在明显的错误。提取其中一段液位数据(见图5), 图中异常检测判定为-1的液位曲线实际均为正常液位曲线, 而液位曲线中间部位有几个抖动的波峰位置则全部被判定为正常值。即正常液位数据被算法判定为异常值, 而异常液位数据则刚好被判定为正常值。因此, 该算法无法通过找到异常值来定位液位突变区间。究其原因, 异常检测算法是研究所有数据在一个区间内数据分布概率, 只要数据处于分布概率低的区间, 就会存在一定的几率被判定为异常数据。由此一部分正常的高液位和低液位(数据分布少)将有可能被误判为异常数据, 而一部分异常数据因为存在于正常液位范围内则有可能被误判为正常数据。但该算法对于突变点的判断较为准确。

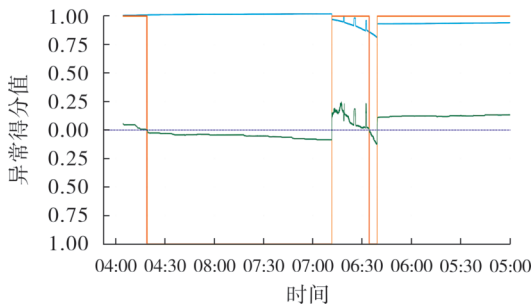


图5 孤立森林算法检测异常示意

3 基于多算法融合的滤波方法

常见的滤波算法包括限幅滤波法、算数平均滤波法和中位值滤波法等。限幅滤波法的思路是根据经验判断两次采样允许的最大偏差值(设为A), 每次检测到新采样值时进行判断: 如果本次新采样值与上次滤波结果之差小于A, 则本次采样值有效, 将本次滤波结果设为新采样值; 反之, 本次采样值无效, 放弃本次值, 将本次滤波结果设为上次滤波结果。然而, 由于突变失

真数据存在较大范围的情况, 显然该方法不适用于当前场景。

3.1 算数平均滤波法和中位值滤波法

算数平均滤波法是指连续取N个采样值进行算数平均运算, 当N值较大时, 信号平滑度较高, 但灵敏度较低; 当N值较小时, 信号平滑度较低, 但灵敏度较高。而中位值滤波法是指连续采样N次(N为奇数), 将N次采样值按大小排列, 取中间值作为本次有效值。中位值滤波法能有效克服因偶然因素引起的波动干扰, 对于变化缓慢的被测参数, 如温度、液位等, 具有良好的滤波效果。利用Python将采集到的液位数据清洗后处理为时间序列数据, 通过滑动窗口函数分别求取了算数平均值和中位值数据, 处理后的液位曲线如下, 其中图6表示经过算数平均滤波处理的液位曲线, 图7表示经过中位数滤波处理的液位曲线。

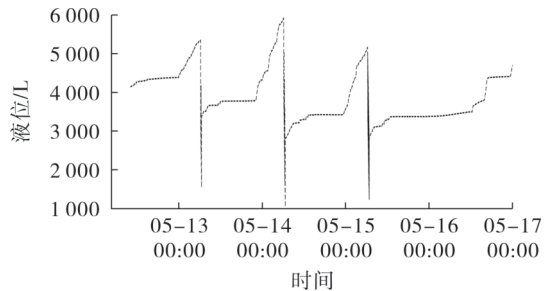


图6 算数平均滤波曲线

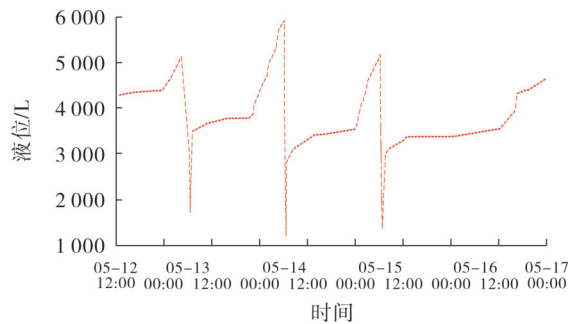


图7 中位数滤波曲线

从图可知, 算法平均滤波和中位数滤波均不能很好解决液位失真问题, 经过算数平均过滤法处理的数据和原始数据的变化不大, 其抖动部分的数据依然存在。而中位数过滤法虽然过滤掉了抖动的液位数据, 但却将一部分的正常数据也剔除了, 导致部分液位数据丢失而无法准确计算实际的耗量。因此, 这两种算法无法处理液位失真问题。

3.2 基于多算法融合的液位滤波法

水性废溶剂的液位曲线滤波处理关键是要把突变的液位数据剔除掉。如何判断突变点, 最简单的逻辑是

液位突然发生了巨大变化,变化前后的两个液位数据的差值必然超过了正常的波动范围,往往这种变化只有在液位计管路内被抽真空才会发生,因为正常的液位曲线是一种缓慢上升的变化过程。因此,找到突变发生前的最后一个真实数据及对应时间,以及突变恢复后第一个真实数据及对应时间,将这个时间区间的数据剔除即完成滤波目的。

提出了一种融合了限幅思想和中位值滤波的算法,按时间顺利从前到后遍历所有液位数据找到限幅后的突变点并删除其中的失真数据,在上述数据处理过程中将删除一部分后期正常液位数据,接着再按时间顺序由后到前再次遍历所有数据,同样这个过程将会删除一部分前期数据,最后将这个两组数据合并后可减少真实数据的丢失数量。完成该遍历操作后再通过中位值滤波的方法清洗掉剩余的抖动值即完成了液位数据滤波过程。图8为原始的液位曲线,图9为经过处理后的滤波曲线。

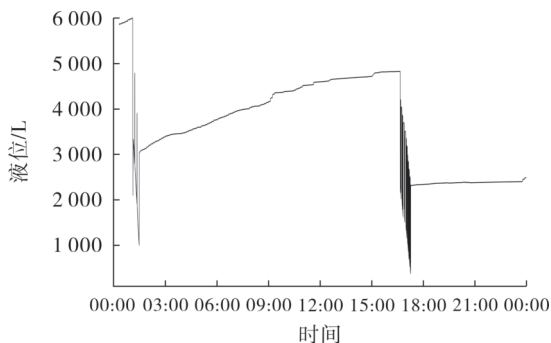


图8 原始液位曲线

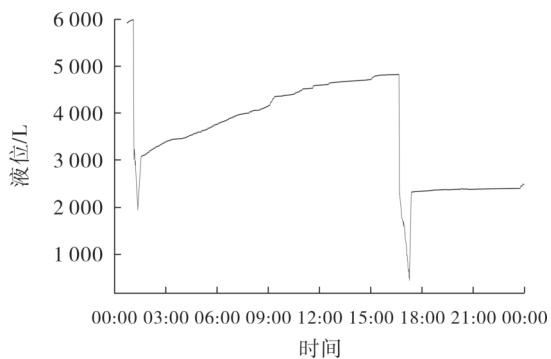


图9 多算法融合滤波液位曲线

4 结语

基于多算法融合滤波算法很好地解决了水性废溶剂液位失真问题,经过该算法过滤后的液位数据能够准确反映液位的真实走势,剔除了抖动部分的数据将作为目标值的一部分为下一步的建模做数据准备。由于水性废溶剂的液位在失真抖动过程中存在不断增加的情况,剔除的数据也会不可避免地将该部分增加的液位数据

一并处理,实际的水废耗量会比滤波后的液位数据略高一些。具体的差异还需要在实际生产中不断跟踪观察,这也是本项目后续工作的一个重要课题。◆

(上接第16页)定位基准来调整改造点的一致性。

4.1.2 机器人仿形调试

1)离线仿形:按照喷漆全工艺流,从鸵鸟毛→BC1→BC2内喷→BC2外喷→CC内喷→CC外喷,开展调试制作;2)离线导入:导入后按照全流程进行对点调试;3)带车过线:按照流程先后顺序,开展实车过线就不同站点进行逐一确认;4)导入参数实车外观调试。

4.1.3 通过性确认

需从焊装上线点至涂装下线转接点全流程进行确认,确认点不仅要考虑输送机运设备、转接点、吊具、滑橇等,同步要关注工艺设备的可达性,以免造成清洗不净或喷涂受限,从而影响产品品质。

4.2 工艺调试

4.2.1 调试资源

1)车身资源(如:1台/通过性、3台/每个颜色、2台/防腐剖解);2)工装器具(电泳、面漆工装各25套,量产300套);3)人力资源(前处理电泳开线人员、面漆线开线人员、机器人及机运操作人员等);4)能源准备(纯水、电、压缩空气、燃气、热水);5)材料(油漆、胶、蜡投罐、随车及一次性堵件、沥青板等)。

4.2.2 工艺参数确认

电泳及面漆工艺调试前,均需确认好各项工艺参数指标在正常工艺范围内,不仅要确认材料参数(如:泳透力、pH、黏度等),同步要确认工艺环境(如:温度、湿度、风速等)。

5 管理优化

为提升后期同类工作开展效率及问题规避,经小组共同复盘研讨并形成标准流程:项目指令下发→现场通过性调研→风险评估→组织架构签批→工作界面确认→框架协议签订→技术任务书签批→商务定标→车身资源协调→改造样件验证→详细方案设计→图纸会签→预制件跟踪→入场准备→批量改造→仿形调试→车身资源协调→工艺调试→小批量试制→SOP→终验收→总结和复盘。

6 结语

通过此次过桥生产改造案例推进,解决了产能溢出问题。同时在改造过程中从立项到项目SOP也遇到了工作中部分不足,诸如改造过程中的工作界面分歧、配合度等问题。同步也形成了转产过桥生产的宝贵经验,为以后过桥生产事宜提供珍贵的理论知识及数据支持。◆