

控制与决策

Control and Decision

面向厚板板形质量监控与异常诊断的大数据可视分析

丁进良, 关凯丰, 张同康

引用本文:

丁进良, 关凯丰, 张同康. 面向厚板板形质量监控与异常诊断的大数据可视分析[J]. *控制与决策*, 2023, 38(8): 2192–2202.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2023.0661>

您可能感兴趣的其他文章

Articles you may be interested in

[基于云模型的煤矿安全大数据多粒度表示方法及应用](#)

Multi-granularity representation method of big data in coal mine safety based on cloud model and its application

控制与决策. 2021, 36(10): 2359–2368 <https://doi.org/10.13195/j.kzyjc.2020.0325>

[基于核典型相关性-熵成分分析的工业过程质量监测方法](#)

A quality monitoring method for industrial process based on kernel canonical correlation-entropy component analysis

控制与决策. 2021, 36(12): 2999–3006 <https://doi.org/10.13195/j.kzyjc.2020.0548>

[数据驱动的综合能源系统运行优化方法研究综述](#)

Review of research of data-driven methods on operational optimization of integrated energy systems

控制与决策. 2021, 36(2): 283–294 <https://doi.org/10.13195/j.kzyjc.2020.0725>

[基于局部信息增量与MPLS的质量相关故障检测方法](#)

Quality-related fault detection method based on local information increment and MPLS

控制与决策. 2021, 36(7): 1647–1654 <https://doi.org/10.13195/j.kzyjc.2019.1402>

[大数据服务商参与下供应链联合减排的动态协调策略](#)

Dynamic coordination strategy of joint emission reduction in supply chain involving big data service provider

控制与决策. 2021, 36(8): 2013–2022 <https://doi.org/10.13195/j.kzyjc.2019.1560>

面向厚板板形质量监控与异常诊断的大数据可视分析

丁进良[†], 关凯丰, 张同康

(东北大学 流程工业综合自动化国家重点实验室, 沈阳 110819)

摘要: 随着信息技术和传感技术的发展与应用,工业制造已迈入“大数据”时代,分析工业大数据和提取价值信息已成为大数据研究的重要挑战. 板形作为厚板生产的重要质量指标,其控制复杂度高,与生产工艺参数密切相关. 由于厚板生产具有多品种小批量特点,且厚板大数据具有规模大、多源异构等特性,现有单一的大数据分析方法难以直接应用. 对此,采用工业大数据可视分析方法,以数据可视化窗口,融合人的感知、分析和推理能力以及数据分析算法,提出“可视+交互+算法”融合模式的厚板板形质量监控与异常诊断方法,并研发厚板生产大数据可视分析系统,解决面向厚板板形质量监控的大数据综合分析利用问题. 实验验证结果表明了所提出方法与系统的有效性.

关键词: 工业大数据; 可视分析; 人机交互; 厚板板形; 生产监控; 异常诊断

中图分类号: TP393.0

文献标志码: A

DOI: 10.13195/j.kzyjc.2023.0661

引用格式: 丁进良,关凯丰,张同康. 面向厚板板形质量监控与异常诊断的大数据可视分析[J]. 控制与决策, 2023, 38(8): 2192-2202.

Big data visual analytics for heavy plate shape quality monitoring and anomaly diagnosis

DING Jin-liang[†], GUAN Kai-feng, ZHANG Tong-kang

(State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang 110819, China)

Abstract: With the development and application of information technology and sensing technology, industrial manufacturing has entered the era of ‘big data’. Analyzing industrial big data and extracting valuable information has become a significant challenge in big data research. Plate shape, as a critical quality indicator in heavy plate production, is intricately linked to production process parameters and presents a high level of complexity in control. Due to the characteristics of multi-variety and small-batch production in heavy plate manufacturing, as well as the large scale and heterogeneous nature of heavy plate big data, existing single data analysis methods are inadequate for direct application. In this study, an industrial big data visual analytics approach is adopted, using data visualization as a window to integrate human perception, analysis and reasoning abilities, as well as data analysis algorithms. A hybrid model combining ‘visualization + interaction + algorithms’ is proposed for heavy plate shape quality monitoring and anomaly diagnosis. A big data visual analytics system is developed to address the challenges of comprehensive analysis and utilization of big data in heavy plate shape quality monitoring. Experimental verification results demonstrate the effectiveness of the proposed approach and system.

Keywords: industrial big data; visual analytics; human-computer interaction; heavy plate shape; production monitoring; anomaly diagnosis

0 引言

随着“中国制造2025”计划的实施以及工业信息技术的革新,我国的工业制造已迈入“大数据”时代^[1],工业企业在生产线上布置大量传感器和应用计算机控制系统,采集并存储跨越多个层级的工业大数

据,建立了大规模工厂级数据中心. 工业大数据存在规模大、结构复杂、维度不统一、内容多样等特点^[2],因此,分析复杂工业大数据和提取数据价值信息已成为工业大数据研究的重要任务和关键挑战^[3].

目前,利用工业大数据实现工业过程分析的理论

收稿日期: 2023-05-16; 录用日期: 2023-05-25.

基金项目: 国家重点研发计划项目(2022YFB3304700).

责任编辑: 杨光红.

[†]通讯作者. E-mail: jlding@mail.neu.edu.cn.

和方法研究主要从两个维度展开:一是从机器或计算机角度出发的数据挖掘和智能建模算法;二是以人和需求为主体,强调人机融合的大数据可视分析法^[4]。数据挖掘与智能建模方法应用于多种实际工业场景的分析任务中,包括多种数据挖掘算法^[5]、关联规则分析模型^[6]、分类与回归模型^[7-8]、机器学习与深度学习^[9-10]等。这些方法虽然可以对工业大数据进行有效分析,但只能针对单一分析任务和业务场景,对数据规模、变量特征和数据特性等提出要求;机器学习和深度学习模型效果更好,但对数据适应性较差,可解释性弱,难以适应工业产品规格多变的特点。大数据可视分析法通过人机交互的方式,将人的认知规律与智能算法结合,把人所具备的、机器不擅长的认知能力融合进分析过程^[11]。人从外界获得的信息约有80%以上来自视觉系统^[12],通过将复杂工业数据转化为视觉编码,并有效结合人的视觉认知和机器的计算能力,可以极大增加分析的灵活性、可解释性和准确率^[13]。因此,可视分析是实现工业大数据分析的有效手段。

厚板是广泛应用于国防军工、基础设施建设、能源工程等多个领域的重要钢铁产品^[14]。不同领域对厚板质量的要求不尽相同,其板形是各应用领域都关注的重要指标^[15]。板形与生产过程密切相关,且控制复杂度高,因此受到广泛关注^[16]。板形异常通常在最终质量检查工序进行判断,因而呈现批量特点。因此,利用生产后的数据进行异常分析,对厚板板形质量进行监控与异常诊断具有重要意义。

本文采用工业大数据可视分析方法,将数据可视

化、人类认知推理能力与数据分析算法有机结合,发挥人机协同的优势,实现厚板板形质量监控与异常诊断。首先介绍大数据可视分析的基本方法及其研究现状;然后针对厚板生产工艺与厚板大数据特点,提出面向板形监控与异常诊断的可视分析框架和方法;最后介绍系统实现并进行实验验证。

1 大数据可视分析概述

数据可视化是指利用常见图像、图标以及动画等视觉表现进行直观展示数据信息的技术,以实现信息记录、对信息进行推理和分析、信息传播与协同等功能^[17]。数据可视化发展史与人类现代文明启蒙和科技发展一脉相承,早在16世纪至18世纪,人类已开始采用图表和图形来表示信息,记录数据的演变规律,并发明了折线图、柱状图、流图等如今常见的图形,如普莱费尔发明的折线图用于贸易平衡时序分析(见图1(a))。随着工艺设计的完善,19世纪产生了概念制图的新思维,人们将数据可视化工具应用到经济、社会、医学、规划等多个工程和统计领域中解决实际问题,如近代护理事业的创始人南丁格尔创作堆叠饼图(玫瑰图)分析克里米亚战争中士兵的死亡原因(见图1(b))。自20世纪至今,随着计算机技术与信息科学的发展,人类迎来大数据时代,复杂信息展示与探索式数据分析蓬勃发展,数据可视化发展成包括科学可视化、信息可视化、可视分析等研究分支的数据科学^[18]。探索式交互技术可以辅助人们动态地查看复杂的大数据,快速挖掘有用的信息,如鱼眼方法模拟鱼眼视觉效果对重要细节提供专注(见图1(c))。

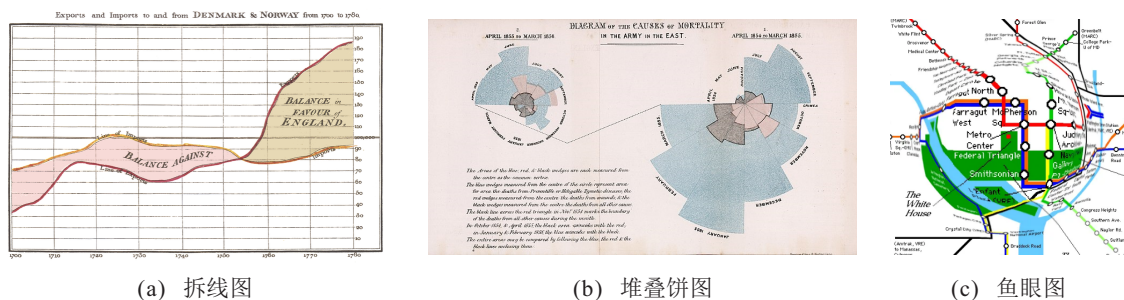


图1 数据可视化发展历史中的经典可视化图形

科学可视化研究如何对物理、化学、医学、生物等学科的数据和模型进行解释、操作与处理,寻找其中的模式、特点、关系以及异常情况,所涉及的数据往往包含真实世界中存在的几何结构;信息可视化则更关注抽象、高维的非结构化数据,结合统计图形学、信息图形、视觉设计等技术,通常在二维空间中以直观的方式传达大量抽象信息,这些数据本身没

有看得见摸得着的几何结构^[19];可视分析是以可视交互界面为基础的分析推理科学^[20],通过可视交互界面将图形学、数据挖掘和人机交互技术结合起来,把人的感知、认知和推理能力以可视的方式融入数据分析过程,提升信息交流效果,完成有效的分析推理和决策。可视分析更关注系统中的分析含量^[19],更多地集成了数据挖掘等自动算法对数据进行综合分

析. 图2^[11]展示了可视分析与人、机器之间的关系.

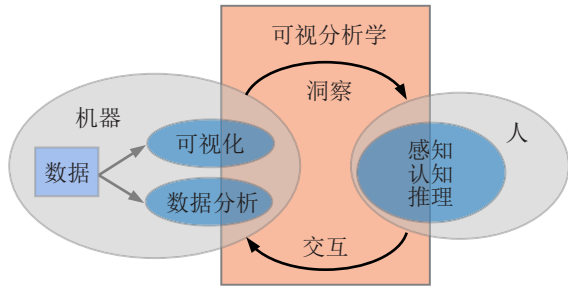


图2 可视分析与人、机器之间的关系

数据可视化视图种类丰富、设计灵活,能很好地应对结构复杂、内容多样的工业大数据信息表达问题,结合人的智慧和机器算力,可以从全新的角度实现工业大数据的分析,提高分析结果的灵活性、可解释性和准确率. 大数据可视分析在工业过程应用也逐步展开. 比如, LiveGantt 系统^[21]支持操作员从不同角度探索离散制造的大规模并发调度; Sun 等^[22]提出 PlanningVis 支持生产计划的探索和比较,通过交互式视觉探索与自动规划算法结合促进生产计划高效优化; Kostoláni 等^[23]用增强现实技术实现对工厂的实时过程监控; Xu 等^[24]设计了 ViDX,采用视觉聚合技术扩展 Marey 图,实现对装配线上故障的有效排除和实时的数据探索; Huettenberger 等^[25]通过可视化方法实现汽车零件上故障点的自动检测,并支持交互操作,帮助专家确定检测到的故障特征的严重程

度.

综上所述,大数据可视分析能够为工业生产过程大数据分析提供一个有效的综合分析方法,根据特定的领域需求和工业数据特点进行可视化设计,展示复杂的工业大数据,集成多种分析算法对数据展开多维度的分析,融合交互技术使人的感知和推理能力参与分析过程,挖掘大数据中有用的信息.

2 板形监控与异常诊断问题描述

2.1 工艺流程描述

厚板生产的工序流程如图3所示,包括加热、轧制、冷却和矫直等主要工序. 首先,钢坯在加热炉中经过预热段、加热段和均热段充分加热,使得钢坯温度均匀且达到所需温度,以降低变形抗力,提高钢坯的塑性. 除磷后钢坯进入轧制工序,由粗轧机完成成形轧制和展宽轧制,接着由精轧机完成精轧轧制,获得尺寸和板形等质量符合要求的钢板. 随后进入冷却工序,通过加速冷却装置迅速降低钢板温度或进行淬火处理,控制其相变与碳化物析出,以确保韧性和硬度达到要求. 由于前道工序加工时无法完全避免各种不良生产因素和生产环境对板形的影响,在冷却后需要通过热矫直机对冷却后钢板的瓢曲、浪形等板形问题进行矫正. 厚板生产结束后,在质量检查工序中完成厚板板形检测. 常见板形异常如图4所示,包括边浪、凸起和凹陷等.

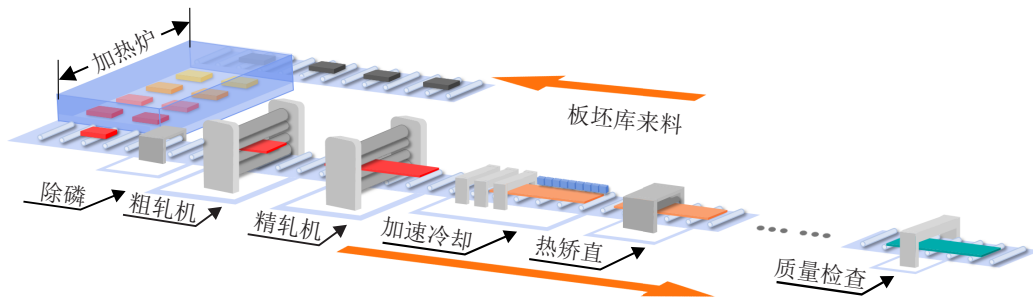


图3 厚板生产流程

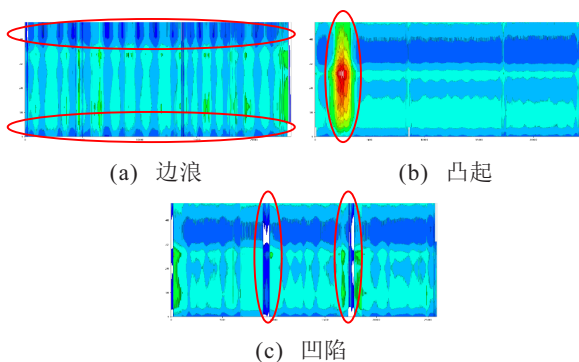


图4 厚板板形异常

2.2 数据特征分析

厚板板形本质上指的是厚板内部残余应力的分布^[26],当残余内应力足够大时,会引起翘曲、浪形、凹陷等厚板板形异常,在厚板生产过程中,加热、轧制、冷却等多个工序均可能对板形质量造成影响,因此,分析导致出现板形问题的原因需要深入厚板工艺及各个加工环节的过程参数.

加热环节需要根据坯料的钢种和断面尺寸等因素采取正确的加热工艺制度,对加热温度、加热速度、加热时间等条件严格设定,保证板坯在加热炉中受热均匀且升温曲线合理. 轧制过程需要确保轧制条件

的控制与设备状态的稳定,以防止轧辊对钢板的内应力造成异常影响.冷却工序要完成整个厚板的快速均匀冷却,通过控制冷却温度、冷却速率、冷却水流遮蔽因子等使得厚板温度均匀下降.根据生产过程对板形影响因素分析,从厚板生产过程数据中选择关键变量,以实现板形质量监控与异常诊断,选取相关数据包括:厚板规格数据,加热炉温度数据、钢坯温度均匀性、加热时间等,轧制工序每道次的轧制力、弯辊力、钢板尺寸变化、轧辊凸度、窜动位置、磨损度等数据,开始和结束冷却温度、温度均匀性、冷却速度、遮蔽因子等数据;在质量检查工序由板形仪通过激光三角法测量得到一个二维数值矩阵,衡量厚板表面的平整度和形状.

对所选取的与板形质量相关的生产过程数据进行分析可以发现,该数据涉及厚板生产过程中的多个工序和多个设备,涵盖厚板从钢坯到形成最终厚板产品的生产周期.由于采集设备和数据表现方式的差异,所选取的数据呈现出典型的多源异构性,数据维度不统一且长短不相同,如:各工序的过程变量形成一维数据,各工序温度测量点组成的钢板二维温度场、板形仪采集的板形二维数值矩阵,轧制时厚度、轧制力等测量数据增加道次维度使得数据存在三维结构;此外,厚板生产过程是一个典型的流程型生产过程,前一个工序的输出是后一个工序的输入,工序间变量存在耦合性与非线性关系,且厚板的生产规格多变,存在典型的多规格、小批量特性.

2.3 问题描述

钢铁企业亟需分析厚板生产过程数据和设备状态信息,总结导致板形质量异常的因素,进而改进生产过程和工艺参数,改善厚板板形问题.由此,钢铁企业突破了跨层级的多频度异构数据的采集、处理和存储技术,建立工厂级厚板数据中心^[27],以实现厚板大数据智能应用.但有效分析厚板大数据具有较大挑战:一是数据具有规模大、结构复杂、维度不统一等多源异构特征,不能直接应用于建立监控分析模型;二是厚板生产是典型流程工业生产过程,具有工艺复杂、产品品种多、批量小、规格多的特点,难以灵活地筛选分析数据;三是难以考虑全工序过程数据实现板形质量监控与异常诊断分析.

综上,本文将工业大数据可视分析、人机交互和模型算法相结合,提出面向厚板板形监控与异常诊断的可视分析框架和方法,旨在解决厚板大数据难以分析与信息挖掘问题,实现板形异常监控,分析和诊断导致板形异常的关键过程变量,进而辅助现场改进生产过程.

3 面向板形监控与异常诊断的可视分析

本文提出的面向板形监控与异常诊断的可视分析方法的总体结构如图5所示,包括数据处理、异常诊断和辅助决策3部分.数据处理模块首先展示钢板板形异常率日历分布,可刷选异常日期;然后,对选定日期的钢板进行流形学习降维,展示工艺数据的二维投影分布和异常数量的日期分布,提供目标区域刷选

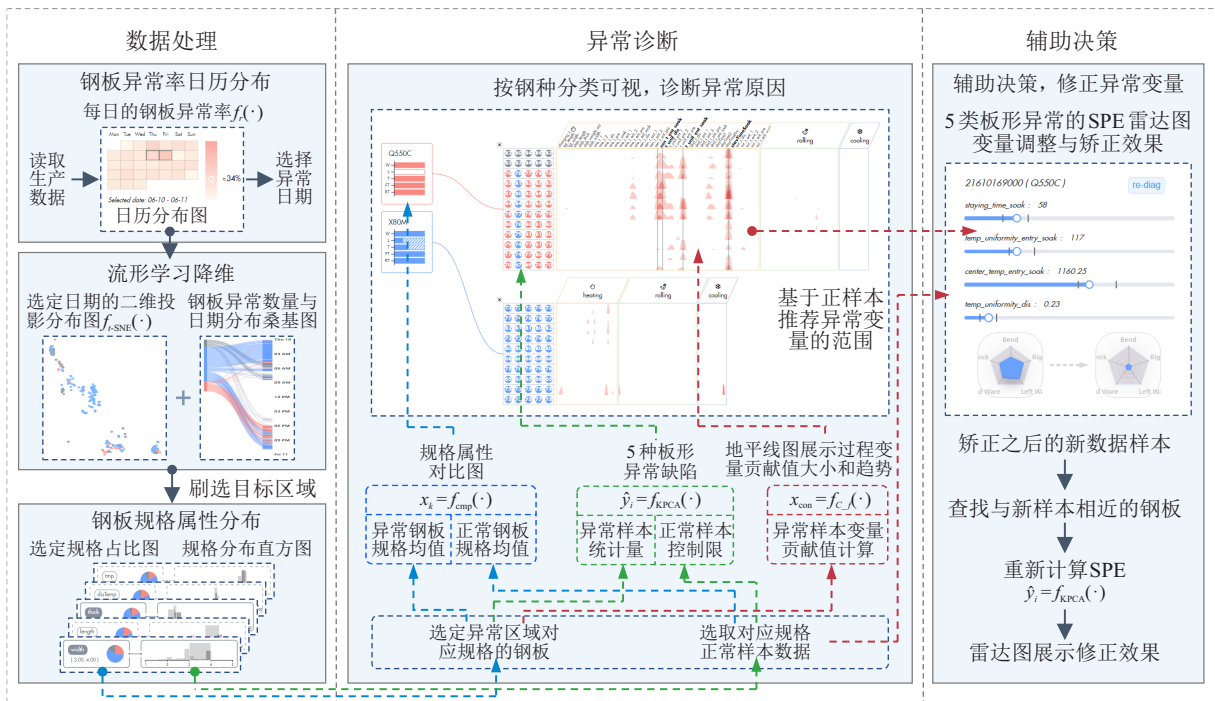


图5 面向板形监控与异常诊断的可视分析总体结构

交互;之后,钢板规格属性分布展示目标区域的选定规格占比以及分布直方图,用于选定异常区域对应的钢板和选取对应规格正常样本数据.异常处理模块按钢种进行分类可视并诊断异常原因,规格属性对比图展示其波动程度,板形异常图展示5类板形异常监控结果,地平线图展示过程变量贡献值大小和趋势.辅助决策模块修正异常变量,实现辅助决策,5类板形异常的雷达图展示变量调整与矫正效果.

3.1 数据处理

3.1.1 钢板异常率日历分布

板形异常率使用阈值分割法对板形矩阵进行分割,将表面分成几个区域,根据不同区域的高度值范围可以将板形矩阵数据转化为向量 $[y_i] = [y_{i,1}, y_{i,2}, y_{i,3}, y_{i,4}, y_{i,5}]$, $y_{i,j}$ 的取值为1或0,表示是否存在头尾翘曲、厚度异常、中浪、左边浪、右边浪这5种板形异常.每天的钢板异常率计算公式为

$$R_f = f_r(Y) = \frac{n}{N} \% \quad (1)$$

其中: n 为当天出现异常的钢板数量, N 为当天生产的钢板总数量.

钢板板形异常率日历分布如图6所示,是对选定的一个月的钢板历史生产数据的展示.通过日历热力图表示每日的板形异常率的高低,右侧颜色映射编码根据颜色深浅表征异常率的高低.工艺专家可通过在该视图交互选择异常日期来快速确定异常率较高的生产区间,进而选定日期来进行后续分析.系统自动设置平均异常率最高的两天作为默认选择的日期.

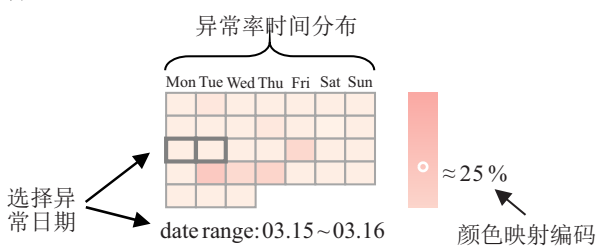


图6 钢板异常率日历分布图

3.1.2 流形学习降维

选择拟分析的异常日期后,对该异常日期范围内生产的全部钢板的工艺相关数据进行流形学习降维和时间分布展示,表示生产工艺和生产计划与板形质量之间的关系.

在众多对高维数据进行降维并可可视化的方法中, t -SNE (t -distributed stochastic neighbor embedding) 算法产生的可视化效果明显较优^[28]. t -SNE 方法是随机邻域嵌入的变体,通过最小化高维空间和低维空间的概率分布之间的KL散度来优化映射结果.提取包

括各加热段的温度、各加热段停留时间、目标尺寸规格、轧制完成温度、冷却相关温度数据和冷却速度等变量的工艺相关数据矩阵,记为 X ,采用 t -SNE 降维的数学表达式如下:

$$X_{t\text{-SNE}} = f_{t\text{-SNE}}(X), \quad (2)$$

其中 $f_{t\text{-SNE}}(\cdot)$ 是 t -SNE 降维方法,具体实现过程参见文献[28]. $X_{t\text{-SNE}}$ 是钢板工艺相关数据降为二维的数据矩阵,每个向量表示在直角坐标系中每个钢板数据点的坐标.数据点的颜色表示钢板板形标签的状态,红色表示出现板形异常,蓝色表示正常板形,灰色表示无法进行板形检测的特厚钢板.厚板数据降维分析视图如图7(a)所示,该视图表示异常钢板的聚类分布模式与生产工艺的相关性.

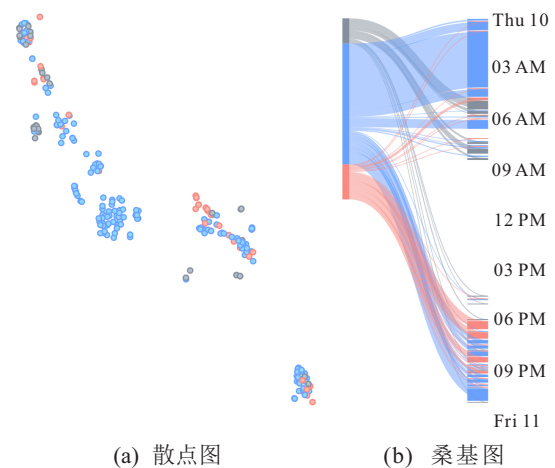


图7 厚板数据降维分析视图

采用图7(b)所示的桑基图表示板形异常状况标签,颜色表示同散点图所述.桑基图左侧表示板形异常状况标签的数量,右侧是分钟级的时间轴,表示钢板的生产时间,桑基图的节点间连接线的宽度表示流量大小.该视图表示不同板形异常状况的钢板数量及其生产时间分布,探索其板形异常的批次性和时间聚集性.

流形学习降维模块旨在帮助工艺专家探索发现钢板的生产工艺数据与板形标签的聚集关系,或总结是否集中出现板形异常,工艺专家可通过鼠标来刷选散点图或桑基图中显示的异常集中的钢板样本区域,进一步选定拟分析出现异常钢板的目标区域.

3.1.3 钢板规格属性分布

钢板规格属性分布模块展示异常日期区间以及拟分析目标区域的钢板规格信息,包括宽度设定值(width)、长度设定值(length)、厚度设定值(thick)、出钢温度设定值(disTemp)、轧后完成滚动温度设定值(tmp)共5个指标.该模块部分指标的可视化如图8所示.图8(a)展示了规格指标名称及其上下限范围,饼

图表示该上下限范围内钢板板形标签状态的占比情况,红色表示出现板形异常,蓝色表示正常板形,灰色表示无法进行板形检测的特厚钢板;图8(b)展示了异常日期区间内所生产钢板的规格属性直方图,反映规格指标的数据分布情况。

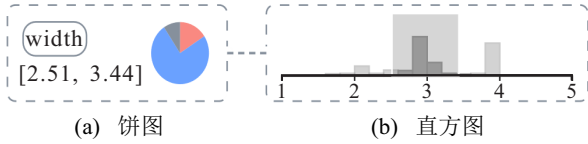


图8 钢板规格属性分布视图

该视图提供动态交互功能,即当工艺专家在图7刷选异常集中的目标区域后,系统会在直方图中初始给出目标区域内钢板的规格属性的上下限范围,饼图根据规格范围的变化动态展示目标区域钢板中符合该范围的板形状态占比情况.工艺专家根据流形学习降维分析结果并结合厚板生产小批量多规格的特点,通过人机交互方式可以灵活地动态调整同规格的上下限范围,筛选出对应规格范围内的异常钢板,并从工业数据库中查找对应规格范围内的正常板形的数据样本,以进行深入可视分析和诊断异常原因。

3.2 板形监控与异常原因诊断

通过对前述模块的数据可视化和探索分析,选定异常区域对应规格的钢板数据集合 Z ,并选取对应规格的正常样本数据集合 V ,正常样本数据即无板形异常的同规格数据,可以作为板形监控与异常原因诊断的参考数据.板形监控与异常原因诊断模块如图9所示,对异常钢板按钢种分类分析,视图包括规格属性对比、5类板形异常监控结果和过程变量贡献值显示。

规格属性对比图展示了按钢种分类后钢板的规格属性波动情况,波动程度是相对同规格的正常样本数据而言的,其计算数学表达式如下:

$$x_k = f_{cmp}(z_{:,k}, v_{:,k}) = \frac{\bar{z}_{:,k}}{\bar{v}_{:,k}} \quad (3)$$

其中: x_k 为第 k 个规格属性的对比结果,一般为大于或小于1的值,表示规格属性增大或减小; $\bar{z}_{:,k}$ 和 $\bar{v}_{:,k}$ 分别为异常钢板集合和正常样本集合中第 k 个规格属性的平均值.图9(a)中的横向柱状图用不同的渲染形态表示这种规格属性波动状态,当规格减小时,柱

状图的空白区域宽度表示减少的程度;当规格增大时,柱状图的阴影区域宽度表示增大的程度;规格属性不出现剧烈波动,则应无其他异常区域.规格属性对比图较好地展示了同类钢种的不同量纲下的规格属性波动情况。

板形监控与异常原因诊断采用基于数据驱动的多变量统计过程监控方法,其中主元分析(principal component analysis, PCA)和偏最小二乘法(partial least square, PLS)等方法本质上都是线性变换,无法有效提取生产过程数据中的非线性信息,因此板形监控与异常原因诊断主要基于核主元分析(kernel principal component analysis, KPCA)模型来实现。

基于KPCA方法监控生产过程是否出现板形异常有3个步骤:首先基于正常样本计算该类板形异常的控制限;然后计算异常样本的统计量;最后比较统计量是否超过控制限来判断是否出现该类板形异常.统计量包括 T^2 和SPE,这里选择监控准确率较高的SPE作为控制限和统计量的计算.KPCA监控板形异常的计算过程如下:

$$K = f_{\phi}(V), \quad (4)$$

$$Q_{lim} = f_{CL}(\alpha^m, K), \quad (5)$$

$$Q_i = f_{SPE}(\alpha^m, z_{i,:}). \quad (6)$$

其中: V 是正常样本数据集合, α^m 是非线性变换后的前 m 个主成分, $f_{\phi}(\cdot)$ 表示非线性变换, $f_{CL}(\cdot)$ 表示计算控制限, $f_{SPE}(\cdot)$ 表示计算SPE统计量,具体的数学表达式见文献[29].

分别针对5类板形异常进行上述监控计算,得到每个钢板是否出现5类板形异常的监控结果,对其可视化如图9(b)所示.图中一列代表一块钢板的板形异常类型,从左到右分别为头尾翘曲、厚度异常、中浪、左边浪、右边浪,红色表示存在该类板形异常,蓝色表示不存在该类板形异常。

异常原因诊断需要计算过程变量在所有主成分中的重要程度,这在基于线性变换的方法中比较容易求得,但无法直接应用于KPCA等非线性监控方法中.针对这个问题,Deng等^[30]为贡献图计算提供了一个新的统一框架:

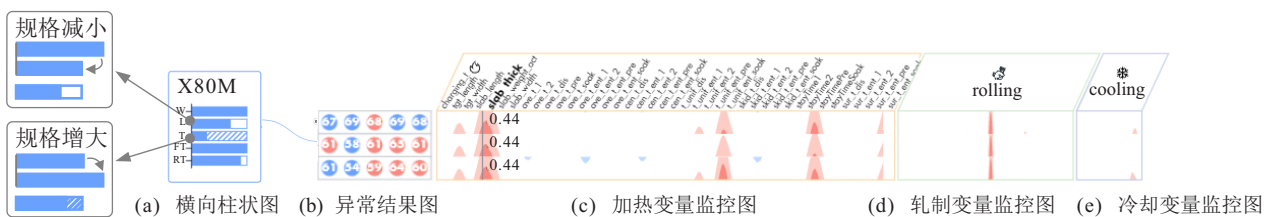


图9 板形监控与异常原因诊断视图

$$z_j^c = f_{C-f}(z_j, \theta), 0 < \theta < 1. \quad (7)$$

其中:异常样本的第 j 个变量的贡献值 z_j^c 由 $f_{C-f}(\cdot)$ 计算, θ 可根据正样本集 V 近似估计.

由式(7)可以计算得到每个异常钢板的生产过程变量贡献值,贡献值越高,表明该变量对板形异常的释程度越大,因此可实现异常原因诊断.此时得到生产批次内多块钢板的多个变量的诊断结果,对于此类数据可以采用多重折线图、热力图、地平线图等方式(见图10).多重折线图线条过多,颜色和趋势变化信息难以获取,而单一折线图由于聚合度太高而无法在有效空间中表达信息;热力图方案颜色清晰,有助于发现变量的分布,但图元过于简单失去了趋势性;地平线图能够有效结合二者优势,实现视觉平衡^[31].因此,采用地平线图方案.

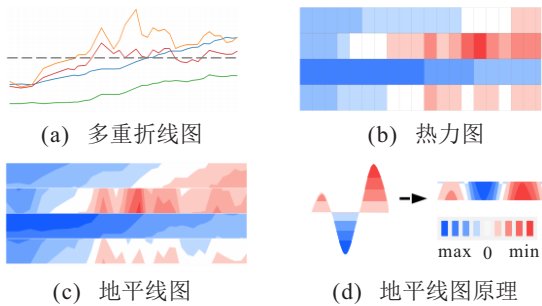


图10 多样本多变量贡献值可视化方案

地平线图原理如图10(d)所示,将多个变量绘制成面积图,并根据数值大小对面积图进行切割和压缩,在有限的视觉空间内保留数值信息.基于地平线图展示生产批次内多块钢板的多个变量的诊断结果如图9(c)~(e)所示,横向表示每个工序中按生产进度排序的过程变量,纵向表示当前生产批次内按时间排序的每块钢板.通过生产批次钢板的过程变量地平线图可以发现异常变量的贡献值大小和出现规律,同时可以总结出各个工序的生产进度中异常变量的数值趋势.点击工序图标可以将该工序的变量折叠,可以更聚焦地分析异常变量较多的工序.工艺专家可以点选钢板和对应异常的变量,进行后续分析和决策辅助.

基于KPCA的板形监控与异常诊断方法基于生产过程大数据进行计算分析,根据贡献值分析挖掘出导致板形缺陷的异常变量,进而由工艺专家分析异常原因,如同规格钢板的工艺参数波动、换规格生产导致关键工序的过程变量设置不合理等.

3.3 辅助决策

异常变量的数值对5类板形异常的SPE结果通过图11(a)所示的雷达图展示,其中每个异常变量滑

动条上的两条标记线之间的区域表示同规格钢板的当前变量推荐值范围.根据正常样本数据集 V 中该变量的上下分位点计算得到

$$v_j^p = f_Q(v_{:,j}, p). \quad (8)$$

其中: v_j^p 表示第 j 个异常变量的 p 分位点,每个变量的 p 分别取0.25和0.75.

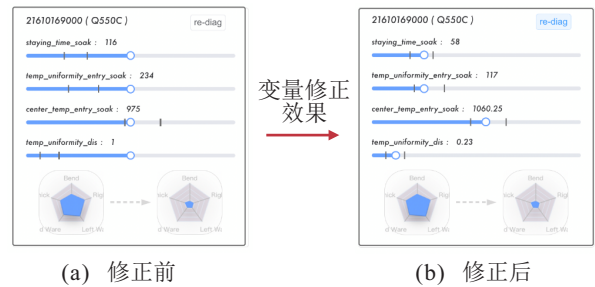


图11 异常变量调整与矫正效果

工艺专家可以根据前述模块对钢板数据的感知和推理结果以及异常变量的推荐范围,拖动滑块来调整异常变量的数值,之后点击视图右上方的按钮,在数据库中检索与调整后的新样本相近的钢板,并按式(6)重新计算5类板形异常的SPE值.绘制变量修正后的SPE雷达图如图11(b)所示,期望调整后的雷达图面积更小,且各类板形缺陷的SPE统计量能有效降低,此异常变量修正效果用于辅助决策并优化生产过程.

4 系统实现与验证

4.1 系统实现

为了验证本文所述可视分析方法的有效性,本文实现了面向板形监控与异常诊断的大数据可视分析系统,采用3层架构模式,将软件分为数据层、业务逻辑层和表现层.系统架构设计如图12所示.

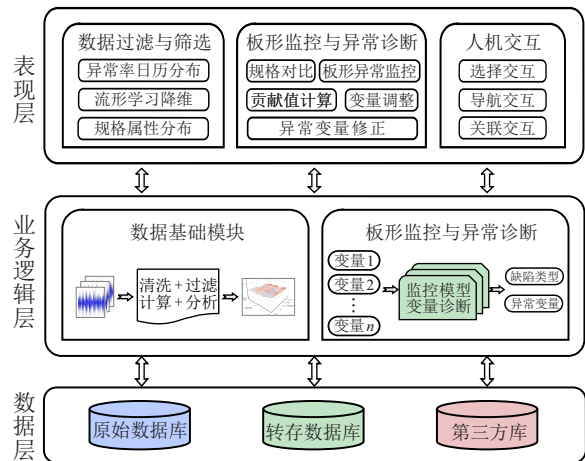


图12 系统架构模式示意图

数据层是系统体系架构的底层,包括原始数据库、转存数据库和第三方库,主要负责接收业务逻辑

层的数据处理指令并将结果返回给服务器.原始数据库储存的是从实际厚板生产现场采集的多层级异构大数据;转存数据库主要保存预处理耗时较长的临时数据,以提高系统的响应效率;第三方库中保存的是常用的程序应用.

业务逻辑层是系统的中间层,位于数据层和表现层之间,根据表现层业务逻辑的需要,查询数据层中所需的数据,调用相关任务函数进行计算和分析,并返回计算结果.在本系统中,业务逻辑层需要实现系统所需的数据基础模块和板形监控与异常原因诊断,包括数据处理函数和算法实现,支撑表现层的功能.

表现层作为3层结构的最上层,是最接近用户的

一层,给用户提供包含丰富交互功能的数据可视化界面,通过Web页面实现数据可视化和板形质量监控,向用户展示信息,同时根据人机交互功能接收用户交互输入,调用业务逻辑层中的接口函数,动态返回结果,实时更新页面信息.

本文基于此系统架构模式,采用JavaScript语言并结合数据可视化框架D3完成表现层页面的开发,采用Python语言和Flask框架实现业务逻辑层的数据处理模块和算法模块,数据存储层利用PostgreSQL数据库存储相关数据.实现面向厚板板形质量监控与异常诊断的大数据可视分析系统运行界面如图13所示.

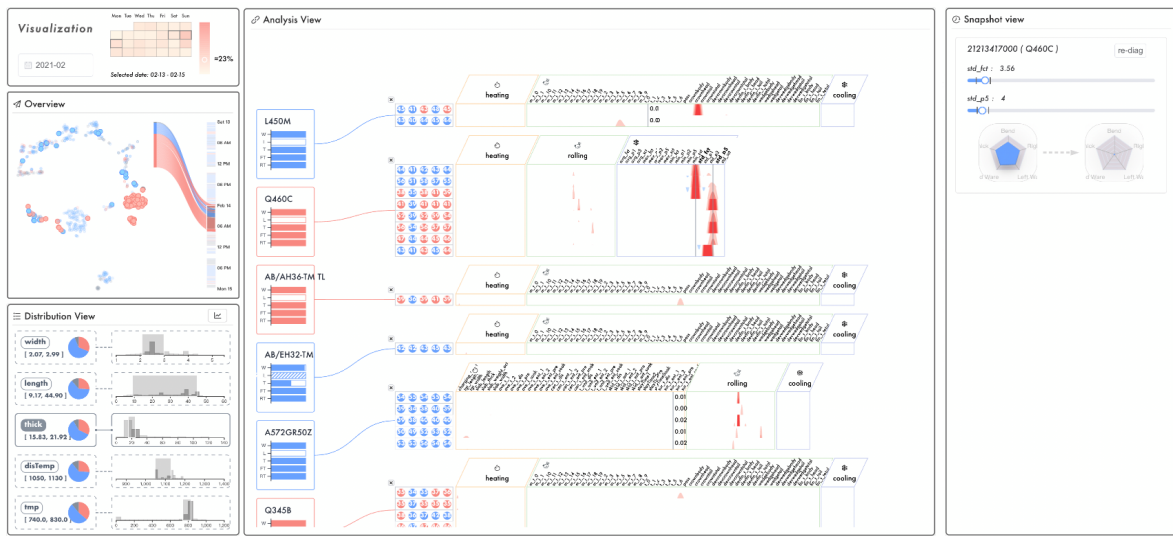


图 13 系统界面

4.2 实验验证与结果分析

使用国内某钢铁企业厚板生产线上实际采集的工业大数据,对本文所述的大数据可视分析方法和系统功能进行验证.

4.2.1 案例 1

案例1的数据范围包括2021年2月生产的全部厚板生产过程数据,验证与分析步骤见图14.

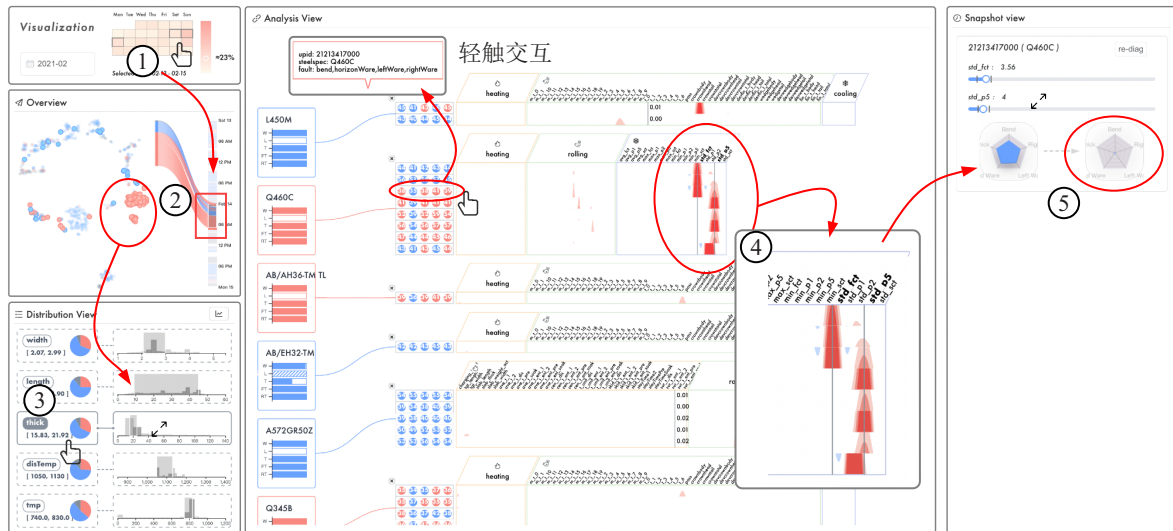


图 14 大数据可视分析系统案例1分析

由钢板异常率日历分布视图可以看出,当生产进行到2月13日至15日时,监控到异常率升高,经过交互探索发现异常率达到23%,通过步骤①选择交互选择异常日期进行深入分析.

观察散点图和桑基图发现,存在异常集中的日期,通过步骤②刷选该异常时间区间,发现在散点图中集中出现,轻触交互发现这类异常钢板的钢种包括“Q460C”“AB/AH36-TMTL”“Q345B”等,其中数量较多的是“Q460C”.由经过步骤②后的规格属性分布图可以看出异常钢板的规格分布范围较大,因此该时间段生产的异常钢板的规格是存在差异的.对于异常数量较多的“Q460C”钢种,工艺专家调整其同规格属性,设置为轧制目标厚度范围在0.015 83 m至

0.021 92 m.

由步骤③筛选出同规格范围内的异常钢板,并从工业数据库中查找对应规格范围内的正常板形数据样本后,进行板形监控与异常原因诊断分析.由异常诊断视图可知,“Q460C”钢种出现了头尾翘曲、中浪、左右边浪等板形异常.由变量贡献值地平线图可以看出,存在2个异常变量:最终冷却温度的标准差、冷却出口温度标准差.通过步骤④点选钢板和对应的异常变量,进行变量分析和决策辅助.

4.2.2 案例2

案例2的数据范围包括2021年6月生产的全部厚板生产过程数据,验证与分析步骤见图15.

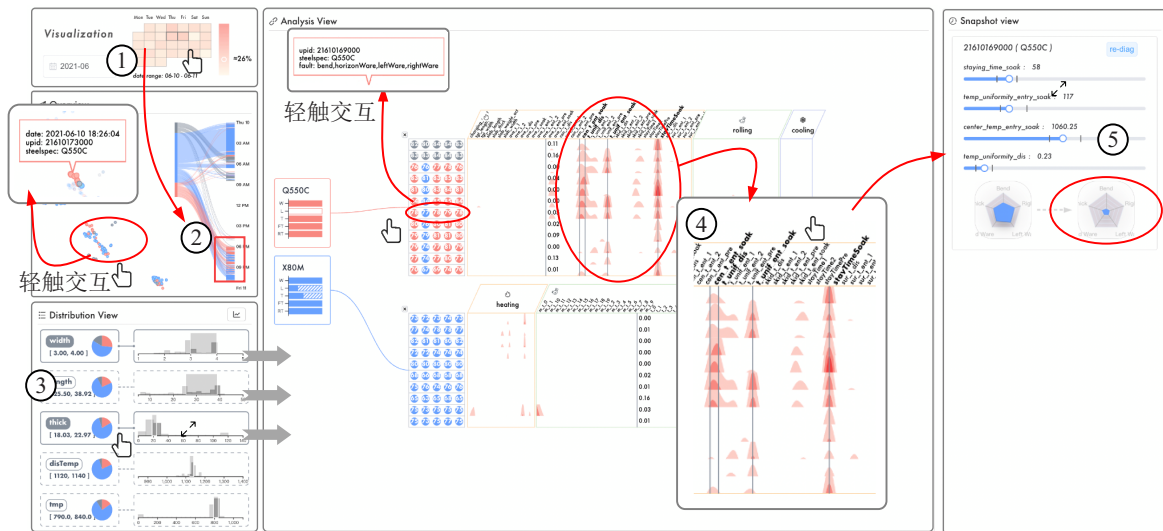


图15 大数据可视分析系统案例2分析

由钢板异常率日历分布视图可以看出,当厚板生产进行到6月10日至11日时,监控到异常率升高,经过交互探索发现异常率达到26%,通过步骤①选择异常日期进行深入分析,此时系统会对异常日期钢板进行流形学习降维和规格属性分布计算.

由滑动条上的标记线可以看出系统根据同规格正常样本计算得到的变量推荐值范围,工艺专家根据步骤⑤拖动滑块将变量值调整至合适的范围,系统根据新样本数值在数据库中查找相近的钢板重新计算SPE雷达图,前后雷达图对比可以看出异常变量修正效果.根据本案例的结果得知,降低最终冷却温度和冷却出口温度的标准差,使得冷却时厚板温度的均匀性更好,可以有效避免同规格厚板出现头尾翘曲、中浪、左右边浪等复合板形异常.

对异常日期内生产的所有钢板的工艺数据采用流形学习降维,鼠标轻触异常钢板聚集区域,并结合日期分布桑基图查看异常集中的日期,发现钢种为

“Q550C”的钢板出现板形异常且在投影视图中聚集,其生产时间也具有连续性,主要分布在10日18时至21时.工艺专家通过步骤②刷选板形异常集中的时间区间,选定拟分析的目标区域,系统计算目标区域内钢板的规格属性的上下限并在其分布视图中初始化,工艺专家可以根据饼图中异常钢板数量占比和钢种特点来动态调整规格属性的数值范围,并通过点击交互激活该规格属性作为同规格筛选条件.在本次验证分析中,同规格钢板条件为轧制目标宽度范围是3 m至4 m,轧制目标厚度是0.018 03 m至0.022 97 m.

由步骤③筛选出同规格范围内的异常钢板,并从工业数据库中查找对应规格范围内的正常板形数据样本后,进行板形监控与异常原因诊断分析.由按钢种分类的分析结果可视化可知,“Q550C”钢种出现了批量板形异常,具体的板形缺陷监控结果为头尾翘曲、中浪、左右边浪.由变量贡献值地平线图可以看出,存在4个异常变量:均热段的停留时间、均热段

入口处的温度均匀性、均热段入口处的中心温度、出炉处的温度均匀性。通过步骤④点选钢板和对应的异常变量,进行变量分析和决策辅助。

由滑动条上的标记线可以看出,系统根据同规格正常样本计算得到的变量推荐值范围,工艺专家根据步骤⑤拖动滑块将变量值调整至合适的范围,系统根据新样本数值在数据库中查找相近的钢板重新计算SPE雷达图,由前后雷达图对比可以看出异常变量修正效果。根据本案例的结果得知,减少均热段的停留时间至58s,增大均热段入口处的中心温度至1082℃,并优化均热段和入口处的温度均匀性,使得钢坯温度更稳定,可以有效避免同规格厚板出现头尾翘曲、中浪、左右边浪等复合板形异常。

5 结论

本文针对厚板生产过程的多品种、小批量、多规格特性以及工业大数据中存在数据规模大、结构复杂等特点,提出了面向厚板板形质量监控与异常原因诊断的可视分析方法,并基于该方法研发了工业大数据可视分析系统。工艺专家以数据可视化视图为窗口,通过人机交互融合自身感知和推理能力,结合系统提供的多视角数据展示和分析能力,总结钢板规格属性波动规律,探索板形异常原因并修正异常变量,实现厚板板形监控与异常变量修正。系统验证分析表明,本文所提的可视分析方法能够在多个维度展示厚板生产过程大数据,以及有机融合“可视+交互+计算”方法的辅助决策能力,说明大数据可视分析方法应用于厚板板形监控的有效性。

参考文献(References)

- [1] 罗立辉. 工业大数据的应用和实践研究[J]. 信息记录材料, 2022, 23(2): 167-169.
(Luo L H. Research on the application and practice of industrial big data[J]. Information Recording Materials, 2022, 23(2): 167-169.)
- [2] 王建民. 工业大数据技术综述[J]. 大数据, 2017, 3(6): 3-14.
(Wang J M. Survey on industrial big data[J]. Big Data Research, 2017, 3(6): 3-14.)
- [3] 何文韬, 邵诚. 工业大数据分析技术的发展及其面临的挑战[J]. 信息与控制, 2018, 47(4): 398-410.
(He W T, Shao C. The development and challenges of industrial big data analysis technology[J]. Information and Control, 2018, 47(4): 398-410.)
- [4] 任磊, 杜一, 马帅, 等. 大数据可视分析综述[J]. 软件学报, 2014, 25(9): 1909-1936.
(Ren L, Du Y, Ma S, et al. Visual analytics towards big data[J]. Journal of Software, 2014, 25(9): 1909-1936.)
- [5] 曹建国, 江军, 赵秋芳, 等. 基于数据挖掘的宽厚板板凸度控制[J]. 中南大学学报: 自然科学版, 2019, 50(11): 2743-2752.
(Cao J G, Jiang J, Zhao Q F, et al. Wide and heavy plate crown control based on data mining[J]. Journal of Central South University: Science and Technology, 2019, 50(11): 2743-2752.)
- [6] Caiado R G G, Scavarda L F, Gavião L O, et al. A fuzzy rule-based industry 4.0 maturity model for operations and supply chain management[J]. International Journal of Production Economics, 2021, 231: 107883.
- [7] Chai Z, Zhao C H. Enhanced random forest with concurrent analysis of static and dynamic nodes for industrial fault classification[J]. IEEE Transactions on Industrial Informatics, 2020, 16(1): 54-66.
- [8] 刘学娟. 基于随机系数回归模型的退化过程及维修策略[J]. 控制与决策, 2021, 36(3): 754-760.
(Liu X J. Degradation process and maintenance planning based on random coefficient regression model[J]. Control and Decision, 2021, 36(3): 754-760.)
- [9] Wu X F, Liu Y H, Zhou X L, et al. Automatic identification of tool wear based on convolutional neural network in face milling process[J]. Sensors, 2019, 19(18): 3817.
- [10] Zhou X K, Hu Y Y, Liang W, et al. Variational LSTM enhanced anomaly detection for industrial big data[J]. IEEE Transactions on Industrial Informatics, 2020, 17(5): 3469-3477.
- [11] Cui W Q. Visual analytics: A comprehensive overview[J]. IEEE Access, 2019, 7: 81555-81573.
- [12] Card S K, Mackinlay J D, Shneiderman B. Readings in information visualization: Using vision to think[M]. San Francisco: Morgan Kaufmann Publishers, 1999.
- [13] Zhou F F, Lin X R, Liu C, et al. A survey of visualization for smart manufacturing[J]. Journal of Visualization, 2019, 22(2): 419-435.
- [14] Xie B S, Cai Q W, Yun Y, et al. Development of high strength ultra-heavy plate processed with gradient temperature rolling, intercritical quenching and tempering[J]. Materials Science and Engineering: A, 2017, 680: 454-468.
- [15] Ma X B, Liu H M, Sun J L, et al. Impact of main drive system of 5m wide and heavy plate mill on screw-down load deviation[J]. Engineering Failure Analysis, 2017, 79: 913-927.
- [16] Wang G D. Status and prospects of research and development of key common technologies of high-quality heavy and medium plate production[J]. Steel Rolling, 2019, 36(1): 1-8.
- [17] 陈为, 沈则潜, 陶煜波. 数据可视化[M]. 北京: 电子工业出版社, 2013.
(Chen W, Shen Z Q, Tao Y B. Data visualization[M]. Beijing: Publishing House of Electronics Industry, 2013.)
- [18] Ward M, Grinstein G G, Keim D. foundations, techniques, and applications[M]. Natick: A K Peters/CRC Press, 2010.
- [19] 屈华民. 大数据时代的可视化与协同创新[J]. 新美术,

- 2013, 34(11): 21-27.
(Qu H M. Visualization and collaborative innovation in the era of big data[J]. New Arts, 2013, 34(11): 21-27.)
- [20] Cook K A, Thomas J J. Illuminating the path: The research and development agenda for visual analytics[R]. Richland: Pacific Northwest National Lab, 2005.
- [21] Jo J, Huh J, Park J, et al. LiveGantt: Interactively visualizing a large manufacturing schedule[J]. IEEE Transactions on Visualization and Computer Graphics, 2014, 20(12): 2329-2338.
- [22] Sun D, Huang R F, Chen Y Z, et al. PlanningVis: A visual analytics approach to production planning in smart factories[J]. IEEE Transactions on Visualization and Computer Graphics, 2020, 26(1): 579-589.
- [23] Kostoláni M, Murín J, Kozák Š. Intelligent predictive maintenance control using augmented reality[C]. 2019 22nd International Conference on Process Control. Strbske Pleso, 2019: 131-135.
- [24] Xu P P, Mei H H, Ren L, et al. ViDX: Visual diagnostics of assembly line performance in smart factories[J]. IEEE Transactions on Visualization and Computer Graphics, 2017, 23(1): 291-300.
- [25] Huettnerberger L, Feige N, Ebert A, et al. Application of Pareto sets in quality control of series production in car manufacturing[C]. 2015 IEEE Pacific Visualization Symposium (PacificVis). Hangzhou, 2015: 135-139.
- [26] 肖桂林, 刘晔. 基于PIDAS平台厚板质量管理的探索与实践[J]. 宝钢技术, 2021(1): 19-23.
(Xiao G L, Liu Y. Research and practice of heavy plate quality management based on PIDAS platform[J]. Bao-Steel Technology, 2021(1): 19-23.)
- [27] 焦四海, 丁建华, 闫博, 等. 厚板数据中心和智能制造的实践与探索[J]. 宝钢技术, 2020(6): 8-15.
(Jiao S H, Ding J H, Yan B, et al. Practice and exploration on data center and intelligent manufacture of plate mill[J]. Bao-Steel Technology, 2020(6): 8-15.)
- [28] Maaten L, Hinton G. Visualizing Data using t-SNE[J]. Journal of Machine Learning Research, 2008, 9: 2579-2605.
- [29] Lee J M, Yoo C, Choi S W, et al. Nonlinear process monitoring using kernel principal component analysis[J]. Chemical Engineering Science, 2004, 59(1): 223-234.
- [30] Deng X G, Tian X M. A new fault isolation method based on unified contribution plots[C]. Proceedings of the 30th Chinese Control Conference. Yantai, 2011: 4280-4285.
- [31] Xu K, Wang Y, Yang L N, et al. CloudDet: Interactive visual analysis of anomalous performances in cloud computing systems[J]. IEEE Transactions on Visualization and Computer Graphics, 2020, 26(1): 1107-1117.

作者简介

丁进良(1976—), 男, 教授, 博士生导师, 从事复杂工业过程智能建模与优化控制、生产全流程运行优化等研究, E-mail: jlding@mail.neu.edu.cn;

关凯丰(1997—), 男, 硕士生, 从事工业大数据可视分析的研究, E-mail: kaifeng_guan@163.com;

张同康(1993—), 男, 博士生, 从事工业大数据可视化与可视分析、复杂工业过程数据驱动建模等研究, E-mail: tkzhang@stumail.neu.edu.cn.



特邀专家 丁进良, 2001年获东北大学自动化专业学士学位, 2004年和2012年分别获控制理论与控制工程专业硕士和博士学位. 现任东北大学二级教授、博士生导师, 流程工业综合自动化国家重点实验室副主任, 东北大学第八届学术委员会委员. 长江学者特聘教授、国家杰出青年科学基金获得者. 长期从事复杂工业过程智能建模、智能优化与控制、生产全流程运行优化及其应用等研究. 先后主持与完成国家重点研发计划项目、国家自然科学基金重大项目等科研项目30余项. 发表论文200余篇, 获得发明专利50余项(国际专利4项)、计算机软件著作权20余项, 获第十四届中国青年科技奖、国家技术发明二等奖、省部级一等奖3项. 享受国务院特殊津贴, 入选“万人计划”科技创新领军人才、国家百千万人才工程和辽宁省兴辽英才计划科技创新团队. 目前担任IFAC MMM和LSC两个技术委员会委员、IEEE高级会员、中国自动化学会常务理事、过程控制专委会副主任、国际期刊IEEE TCAS II和《控制工程》副主编等.

专家寄语 百年树人, 英才辈出! 愿母校在新的征程里, 再铸辉煌!