

引文格式:

刘书含, 翁荻, 巫英才. 工业数据可视分析综述 [J]. 集成技术, 2021, 10(6): 3-19.

Liu SH, Weng D, Wu YC. A review on industrial data visual analytics [J]. Journal of Integration Technology, 2021, 11(6): 3-19.

工业数据可视分析综述

刘书含 翁 荻 巫英才*

(浙江大学计算机辅助设计与图形学国家重点实验室 杭州 310058)

摘 要 自 2013 年工业 4.0 的概念被提出以来, 全世界的工业进程都飞速奔向智能制造时代。而数据感知技术的发展进一步帮助收集海量工业数据, 工业信息化革新正是机遇。但是, 工业数据具有规模大、维度高、结构多变和内容复杂的特性, 对其进行分析是一项严峻的挑战。多变的应用场景又导致分析的灵活度要求提高, 往往需要专家参与分析循环, 因此可视化在工业数据分析中有了更广泛的应用。该综述首先按生产阶段及属性分别总结了工业场景下常用的数据类型; 其次, 根据数据属性, 按时间、空间、时空结合分类, 介绍了对应的可视化方法; 再次, 总结了可视分析在工业场景下的应用, 并讨论如何合理地集成自动化分析方法以提高分析能力; 最后, 展望了工业数据可视分析的发展前景, 并提出未来的研究方向。

关键词 工业 4.0; 智能制造; 可视化; 可视分析

中图分类号 TP 319 **文献标志码** A **doi**: 10.12146/j.issn.2095-3135.20210630003

A Review on Industrial Data Visual Analytics

LIU Shuhan WENG Di WU Yingcai*

(State Key Lab of CAD&CG, Zhejiang University, Hangzhou 310058, China)

*Corresponding Author: ycwu.zju.edu.cn.

Abstract Since industry 4.0 was introduced in 2013, industries across the globe have been rushing towards the era of intelligent manufacturing. The advances in data sensing technologies have further helped the collection of massive industrial data, providing an excellent opportunity for innovations in industrial informatization. However, it remains a major challenge to analyze industrial data because of its large scale, high dimensionality, heterogeneity and complexity. Constantly changing application scenarios also lead to strict requirements in the flexibility of analyses, which demand placing domain experts in the analysis

收稿日期: 2021-06-30 修回日期: 2021-09-29

基金项目: 国家自然科学基金委员会面上项目 (62072400); 浙江省杰出青年自然科学基金项目 (LR18F020001); 之江实验室经济运行态势分析及模拟推演平台 (2021KE0AC02)

作者简介: 刘书含, 博士研究生, 研究方向为信息可视化、可视分析和人机交互; 翁荻, 博士研究生, 研究方向为信息可视化、可视分析和人机交互; 巫英才(通讯作者), 教授, 研究方向为信息可视化、可视分析和人机交互, E-mail: ycwu.zju.edu.cn.

loop. Therefore, visualization has been widely applied to analyzing industrial data. This review article first summarizes the data types commonly used in the industrial scenarios based on the production stages and properties. Then, based on the data properties, this paper introduces visualization methods for the temporal, spatial and spatio-temporal types. Further, this paper overviews the applications of visual analytics in the industrial scenarios and discusses the integration of automated analysis methods in the visual analytics systems. Finally, this paper prospects the development of industrial data visual analytics and possible research directions for the future.

Keywords industry 4.0; intelligent manufacturing; visualization; visual analytics

Funding The work is supported by National Natural Science Foundation of China (62072400), Zhejiang Provincial Natural Science Foundation (LR18F020001), (in part) by Zhejiang Lab (2021KE0AC02)

1 引言

自 2013 年工业 4.0^[1] (或互联工业、工业互联网^[2-3]) 的概念被提出以来, 全世界的工业进程都飞速奔向智能制造时代。工厂中越来越多的传感器、3D 扫描仪等各类新设备以及蓝牙网络等各种连接方式为人类带来了海量的工业数据资源, 从而开启了工业信息化革新的大门^[4]。但是, 对工业数据进行分析仍然是一项严峻的挑战。一方面是因为工业数据规模大、维度高、结构多变、内容复杂; 另一方面是因为工业数据具有多变的应用场景, 难以直接应用全自动化的分析模型。

数据可视化使用视觉通道对数据属性进行编码, 并通过直观的图表传递信息^[5]。种类丰富的可视化形式满足了对大量复杂的异构数据量身定制图表的需求。再者, 可视分析结合数据可视化方法将人类感知融入数据分析的过程中^[6], 同时利用人类智慧和机器算力, 极大地增加了分析的灵活性、可解释性和准确度。因此, 可视分析被广泛地应用于工业数据, 是工业 4.0 必不可少的关键技术^[7-8]。

本文收集了近十年来在可视化、可视分析、人机交互会议与期刊中发表的有关工业场景的研

究, 并在网络上采用关键词搜索的方式提取了相关研究。本综述的结构包括: 第二章总结了工业场景下常用的数据类型, 并按照生产阶段(设计规划、生产制造、运维服务)及属性(时间、空间、时空结合)分别归类介绍; 第三章根据数据的共有属性, 按时间、空间、时空结合分类介绍了对应的可视化方法; 第四章总结了可视分析在不同工业场景下的应用, 并讨论如何合理地把自动化分析方法应用到可视分析中; 第五章展望了工业数据可视分析的发展前景, 并提出未来的研究方向。

2 工业数据

本文收集并统计了近十年来常见的工业数据类型, 按照生产阶段分类介绍其用途、采集方式、形式, 并按照数据属性进行归纳总结, 如表 1 所示。

2.1 按生产阶段分类

经过文献调研^[9], 并结合工厂实际生产的先后顺序, 本文将数据的产生时间划分为设计规划、生产制造和产品运维等 3 个阶段。设计规划阶段主要包括工厂初期设计建造、改造以及生产规划等过程, 生产制造阶段涵盖工厂实际制造产

表 1 常用的工业数据

Table 1 Commonly used industrial data

数据类别	数据属性	数据类型			数据使用示例
		数值	类别	文字	
设计规划域	抽象结点	位置	√		[10-13,29]
		设备编号		√	
	工厂布局	位置	√		
		形状及面积	√		
		设备编号		√	
	立体点云	位置	√		
		设备编号		√	
	工序设计	时间	√		[10,15]
		工序名称		√	
		优先级	√		
		使用设备		√	
	日程规划	时间	√		[17-18]
		工作负载量	√		
		优先级	√		
生产制造域	工况状态	时间	√		[19-22,30-31]
		状态	√	√	
	控制信息	时间	√		
		控制参数	√	√	
	污染情况	时间	√		[24]
		废气	排放量	√	
			方向	√	
		时间	√		
		废水	排放量	√	
			方向	√	
产品运维域	产品质检	时间	√		[25]
		产品编号		√	
		质检结果	√	√	
	物流服务	货物编号		√	[5,27-28]
		时间	√		
		位置	√		

品的过程, 产品运维阶段包含产品离开工厂装配线后的维护过程。

2.1.1 设计规划数据域

设计规划阶段产生的数据主要包括工厂布局数据、工序设计数据和日程规划数据等。

(1) 工厂布局数据标明工厂中每个机器设备的位置信息和空间物理信息。由于布局能够影响

生产效率, 收集并分析工厂布局数据尤为重要。如, 在自动化程度高的车间中, 不合理的布局会增加产品在工序间的传递时间; 而在自动化程度低的场景下, 糟糕的布局会影响工人之间的可见性, 从而降低沟通效率^[10]。利用 3D 扫描仪从实际场景采集^[10-12]和从设计的 CAD 文件中分析获取^[13]等都是工厂布局数据的常见收集途径。数据

的形式有点云^[14]、坐标等。

(2) 工序设计数据记录了工厂中装配线或流水线的顺序信息和时间配置,包括但不限于某阶段某设备的计划装配时长、设备与设备之间的装配顺序等^[15]。在工厂中,设备之间实际上可能存在多个拓扑序。如,先安装左前轮还是右前轮对车辆装配过程不会造成太大影响,但如果左轮和右轮的传送流水线速度差异过大,就可能导致装配线堆积阻塞。该种数据可从工厂系统日志、设计文件等内容中分析提取,又被称为材料流^[16]。

(3) 日程规划数据表示工厂的生产计划。对部分工厂而言,同样的装配线可以通过简单的改动完成多种产品的生产^[17]。所以如何安排每种产品生产的时间、如何调整装配线、何时进行调整,都是日程规划需要考虑的问题。日程规划数据可从历史系统日志中提取,也可以是分析系统的输出^[18]。

2.1.2 生产制造数据域

生产制造阶段产生的数据主要包括工况状态数据、控制信息数据和污染情况数据等。

(1) 工况状态数据记录了工厂车间或设备每个时刻的工作状态。工况状态中有大量的数据采集自传感器,反映了工厂中某些设备的物理或化学属性,如温度、压强、速度、流量等^[19-20]。在这些数据的基础上进行简单组合和计算,还可以进一步得到聚合信息,如工作效率、危险指标等^[21-22],这也是工况状态数据的一种类型。

(2) 控制信息数据记录了工厂中各个时刻可直接调控的信息。控制信息既包含车间设备机器的配置数据(如传送带速度、电流频率、输水流量、送风温度等),还包含车间的人力资源分配情况^[13](如工种安排、工人数量、工人调度等),特别是在建筑业中^[23],合理的人力资源分配意味着更高的安全性以及更高的效率。机器的参数既可采集自传感器,又可人工记录自仪表盘。人力资源信息则一般来自于刷卡信息或系统调度记录。

(3) 污染情况数据记录了工厂的污染情况,包括废气(时间、排放量、方向),废水(时间、排放量、方向),废渣(时间、污染物总量、其他信息)等。污染数据的采集方式众多,既可以收集自传感器,又可以通过定期的采样化验得到^[24]。

2.1.3 产品运维数据域

产品运维阶段产生的数据主要包括产品质检数据和物流服务数据等。

(1) 产品质检数据记录了工厂所生产产品的质量情况。对以制造业为代表的有实际产品的工业类型而言,不同的产品有各自的质检标准,如检测车辆前保险杠的密度、强度、边缘切割情况等^[25]。至于以燃煤发电为代表的重工业类型产业^[26],因为不生产实际产品,所以一般无产品质检数据。

(2) 物流服务数据记录了产品离开装配线之后的物流及服务维护信息。一方面,它包括产品离开装配线后的物流配送信息^[27],如物流位置、配送时间等;另一方面,它也涵盖了产品的维护及服务数据,如保修时间、退换情况等。特别地,物流服务数据有别于物流仓库数据^[28]:前者以产品为主体进行记录,后者还需考虑储物、工人等信息。物流仓库实际可视为一个独立的工厂。

2.2 按数据属性分类

不同生产阶段的数据有共同的属性,按照属性可以进一步归纳总结各工业生产阶段的共性。归纳数据属性便于在可视化设计时,用不同的视觉通道对不同的数据属性进行合适的编码。本节在表 1 归纳的基础上,主要介绍其中使用最普遍、意义最重要的两个维度:时间与空间。

2.2.1 时间序列

工厂装配的先后逻辑使其天然带有时间的维度,所以在工厂中许多数据都是时间序列数据。准确地说,工厂采集到的是一系列有序的离散点,各自记录了该时间点上的某数据值,时间点

按先后排布就得到了时间序列数据。

时间维度信息对工厂分析而言至关重要,特别是级联因果的分析。举例来说,假设工厂流水线上发生了事件 A(传送带 M 堵塞)和事件 B(传送带 N 堵塞),传送带 M 和 N 相邻。在未知时间信息的情况下,无法进行两个事件的关联分析。但若已知事件 B 紧接在事件 A 后出现,则事件 A 较有可能导致了事件 B。

工序设计、日程规划、工况状态、控制信息、污染情况、物流服务都具有时间属性,属于时间序列数据。

2.2.2 空间位置

工业数据中的空间信息虽然大多数都可以表示为离散的坐标位置,但根据抽象到具体的程度,坐标又可以分为三级。第一级是用一个抽象的坐标结点表示一个设备,设备与设备之间的关联(如路径、可见方向)就是结点之间的连线。第二级是用二维平面中的一个区域表示一个设备,比原始的位置信息增加了面积属性。随着数据采集技术的不断发展,越来越多的新型数据采集设备被采用(如 3D 扫描仪),于是点云逐渐成为记

录工厂空间信息的第三级重要的数据形式。相比于结点或区域,点云更加准确地记录了设备的 3D 信息,这意味着更容易进行还原。工厂布局数据形式众多,三级皆有。其中,物流服务、工况状态大都使用第一级坐标结点。

3 工业数据可视化

时间、空间是工业数据普遍具有的重要维度属性,本节将逐一介绍在可视化时间、空间、时空相结合的数据时常用的形式,以及近年来各种可视化形式在工业数据上的实际应用情况,如表 2。

3.1 时间序列数据的可视化

Aigner 等^[39]根据时间数据的性质将其分为 3 个正交的维度:线性时间或周期时间、时间点或时间区间、有序时间或分支时间。3 个维度的不同组合有不一样的可视化设计。Brehmer 等^[40]根据时间序列数据的特征,探索了时序数据的设计空间^[40]。线性时间通常用从左到右的一根轴编码^[41-42],象征着时间的流逝,周期时间通常用环形或螺旋状编码^[43-44];时间点和时间区间的粒度

表 2 常用的工业数据可视化形式

Table 2 Commonly used industrial data visualization

维度	维度分类	可视化形式	适用场景举例	应用示例
时间	时间点	折线图	分析变化趋势(如工作效率、产量、排污量)	[20,24,32-33]
		堆叠图	分析总量和分量的关系(如总工作负载和各机器负载),观察累积特征(如最高产量点)	[12,32]
		面积图	结合折线图和直方图,既呈现累积量(如产量)又需呈现变化趋势(如生产效率)	[12,17,30]
		山图	比较具有不同单位的事件序列数据	[34]
	时间区间	甘特图	生产规划、分析机器占用效率、工序优化	[17-18,33]
空间	时间区间	嵌套环形图	周期性时间序列的生产规划、故障分析、工序优化	[30]
		日历图	周期性时间序列,分析特定日期的关联、生产日程安排	[19,29-30]
		点线图	材料流、工序图、工人轨迹的呈现、设备的关联分析	[12,24]
	结点坐标	区域图	占地面积规划,设备布局规划(如建筑图)	[13]
	平面区域	3D 还原	工人培训、工厂的设计与改造、增强现实下的工厂监控	[21,35-36]
时空结合	时空结合	时间曲线	布局优化(如在区域面积图上分析工人轨迹耗时)	[37-38]
		马雷图	异常分析,可以利用时空上不同的变化斜率突出故障部分及时间区间	[30]

不同,一个常用点表示,另一个常用区域表示;有序时间常用有向图表明顺序,而分支时间的一个典型代表就是故事线^[45-46]。此外,线性时间上面还存在一些离散的区间,这些区间构成了事件的序列^[47-48]。在工业数据中,线性时间点及时间区间、周期时间点及时间区间的可视化较为常见。

3.1.1 时间点

线性时间点最普遍的可视化形式是可以用于反应趋势的折线图以及可以用于反映数量的柱状图或面积图。Castor 等^[24]使用图 1 的折线图表示工厂排放的废水中有害化学物质随时间变化的含量,而 Post 等^[12]使用图 2 的面积图表示单一机器某时刻的工作负载。

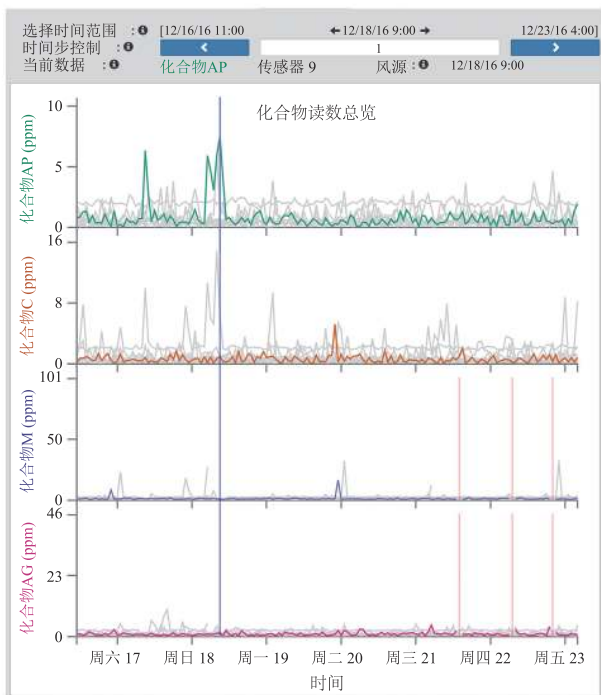


图 1 工厂排放废水中所含化学物质的折线图^[24]

Fig. 1 Line chart of chemicals contained in waste water^[24]

线性时间点还有一种常见的形式是堆叠图, ThemeRiver^[49]就是一个非常典型的案例,如图 3 所示。Post 等^[12]在模拟、分析和评估工厂设备位置变化的影响时,也采用了堆叠图来展示总的工作负载,如图 4 所示。

目前,工业时间序列数据普遍使用的可视化

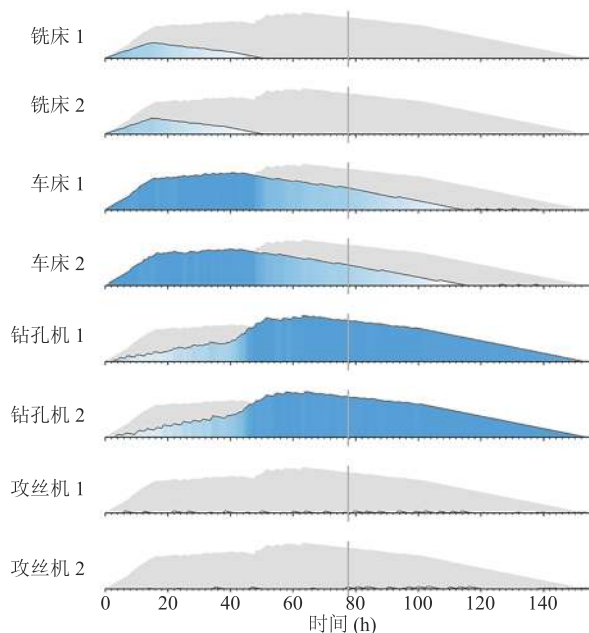


图 2 单一机器某时刻的工作负载^[12]

Fig. 2 Workload of each machine^[12]

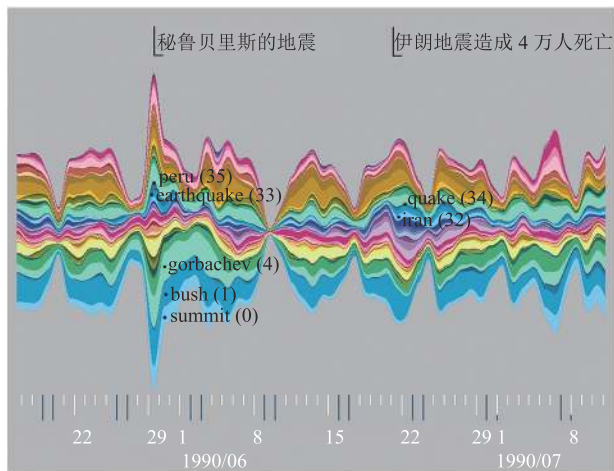
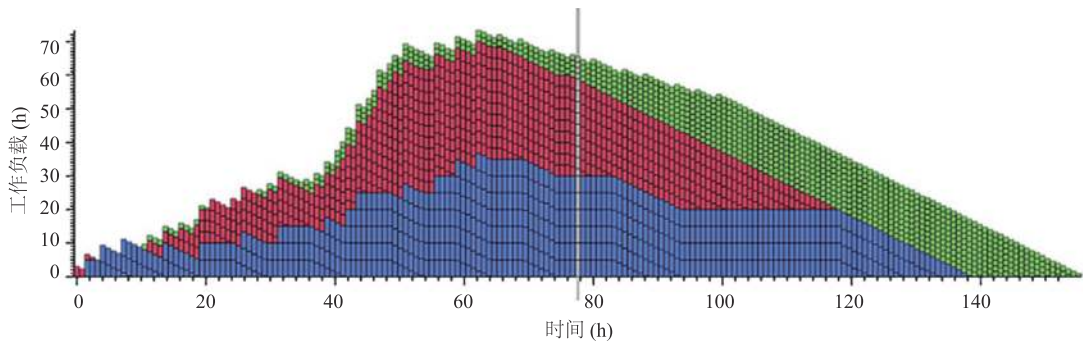


图 3 堆叠图^[49]

Fig. 3 Stacked chart^[49]

形式中,折线图长于呈现变化趋势;柱状图则牺牲了一部分连续性来兼容区间总量统计的功能;面积图能综合前两者的优点,但其编码较为复杂,带来了理解成本;而堆叠图在需要同时呈现整体和部分的关系时更具优势。

此外,还有许多线性时间点的新颖可视化形式未被应用到工业数据中。如 Time curves^[37]可以通过连线在散点图中叠加时间信息,这和材

图 4 用堆叠图展示总工作负载^[12]Fig. 4 Using stacked chart to show workload^[12]

料流较为相似。但材料流主要强调拓扑序和关联性, 而 Time curves^[37]还能编码更加细粒度的信息, 特别适合用户在细节视图上根据需求自定义使用。还有山图^[34]通过面积图的重叠排布来减小所占空间, 且因为不共用 y 轴, 能够同时分析比较具有不同单位的数据。

3.1.2 时间区间

在可视化线性时间区间时, 甘特图^[50]常被用在生产计划、流水线占用情况等数据上。甘特图由多个横条组成, 每个横条的左侧表示开始时间、右侧表示结束时间, 横条的长度就是所需时长。这样一个横条就能表示一个产品在某机器上的生产占用时间。

如图 5 所示, LiveGantt^[17]使用甘特图辅助分析了制造进程中产生的数据。以任务 A 为例, 其从 0 点至 6 点占用了 DA001。图中一共有 13 个任务 4 个资源, 而用户可以很快速地发现任务 B、C 之间有一段时间重叠且分别在两个同类资源

源上, 说明甘特图长于简洁明了地呈现多任务多资源的信息, 这和工业场景非常契合。而任务 C 和任务 D 之间的 Ch 代表 12 点后对 DA002 进行了配置调整, 意味着甘特图也可应用于辅助插入临时任务。

周期时间区间常使用基于角度编码时间的嵌套环形图(如 24 小时、周一至周末、12 个月、四季等), 或日历图。Herr 等^[19]在尝试从工业流水线数据中挖掘事件特征时采用了日历图。其中, 颜色的深浅表明异常的程度, 如图 6 所示。

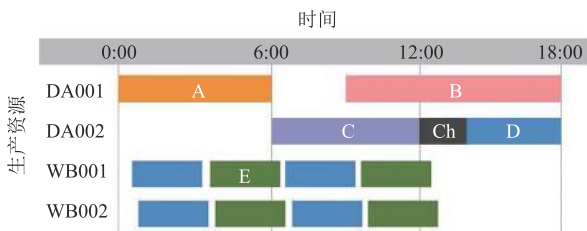
线性时间区间与周期时间区间的可视化具有许多相似之处, 事实上嵌套的环形图类似于卷成圆的甘特图, 在表现机器繁忙程度、规避生产冲突、进行生产规划时都能起到优秀的作用。日历图的独特之处在于可以展示与特定日期的联系, 如节假日、产品的淡季和旺季。

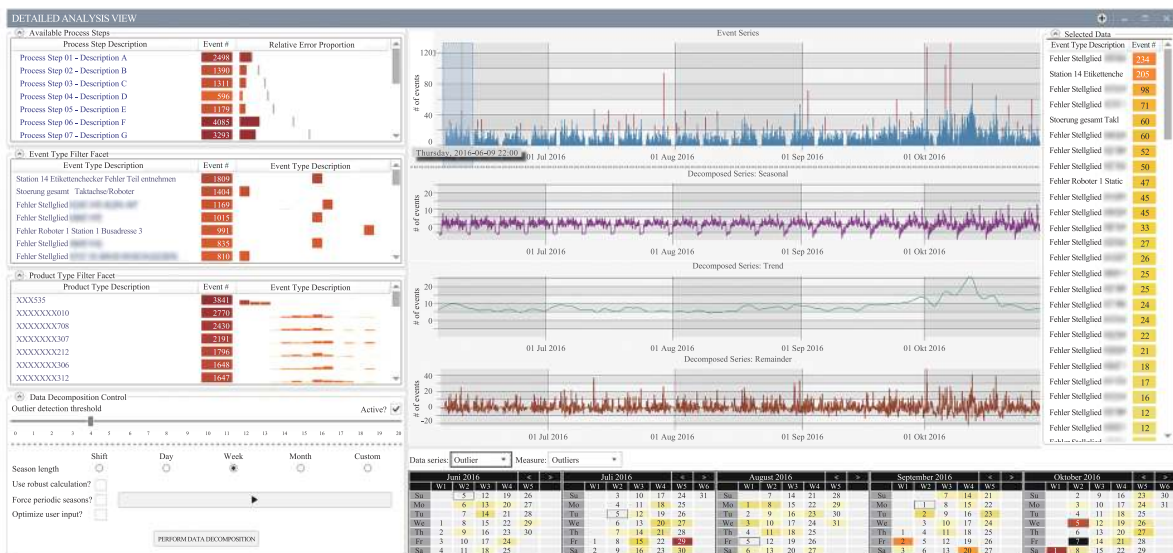
3.2 空间位置数据的可视化

工业数据中的空间信息按照抽象到具体的程度分为结点坐标、平面区域、立体点云三级, 本节将分别介绍这三级数据的常用可视化设计及其在工业数据上的实际应用情况。

3.2.1 结点坐标

用一个单独结点抽象表示位置是最简洁直观的一种方法。在工业数据的背景下, 这种表示形式虽然以较粗的粒度描述了单个设备信息, 但是能帮助用户快速概览工厂的整体布局情况以及设

图 5 甘特图^[17]Fig. 5 Gantt chart^[17]

图 6 日历图^[19]Fig. 6 Calendar^[19]

备之间的关联情况。

Post 等^[12]使用结点抽象表示工厂中的设备,如图 7 所示,首先用不同颜色流线编码材料类型,然后由结点和流线共同组成的聚合视图呈现了材料流的整体情况。这一视图的优点在于简洁的抽象结点既完成了表示设备的任务又为其他信息留出了空间——流线的粗细和密集程度能快速反映工作负载,而且流线的方向明确了产品在各设备间的流动信息。

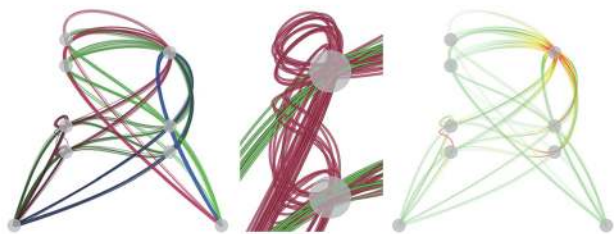
图 7 用抽象结点表示坐标的示例^[12]

Fig. 7 Example of using abstract nodes to represent locations^[12]

3.2.2 平面区域

比起结点坐标,平面区域能更加直观地反映设备或车间的占地面积。其优点是在设计或改造工厂布局时,能更加快速地判断能否完成设备的移动、评估车间的拥挤程度等。

BlueCollar 系统中^[13]使用不同大小的平面区域来表示各类设备,区域间的连线就是工人在工厂车间中的轨迹。如图 8 所示,图中 A 设备的面积明显大于 B 设备。不仅如此,还能基本判断出 B 和 C 属于同一类设备,且设备类别在工厂布局时也是很有意义的信息。

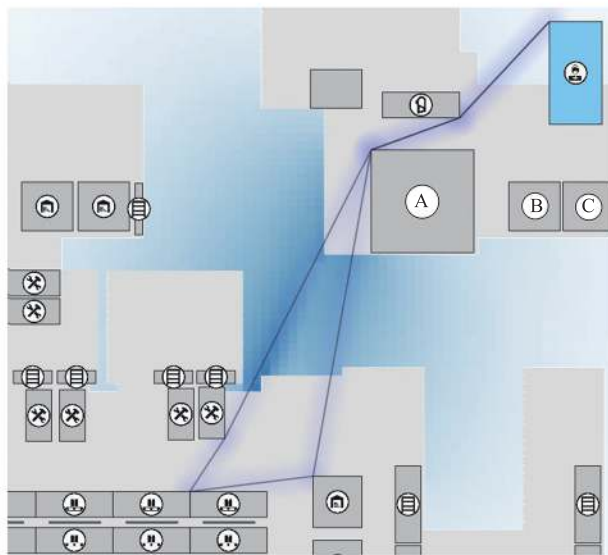
图 8 用平面区域表示位置的示例^[13]

Fig. 8 Example of using a flat area to indicate a location^[13]

3.2.3 立体点云

点云是通过 3D 扫描仪得到的物体外观表面

点数据集合, 根据点的密集程度可以分为稀疏点云和密集点云。传统方法常通过 CAD 等软件或工具将点云还原到 3D 图中进行可视化。近年来, 随着沉浸式技术的发展, 点云的可视化形式越来越丰富。特别是在增强现实 (Augmented Reality, AR)、虚拟现实 (Virtual Reality, VR) 的场景下, 得益于点云信息的详细度, 用户能够在更加真实的还原场景上作进一步的分析。

特别地, 已经有许多综述介绍增强现实^[38,53]、虚拟现实^[52]在工业 4.0 时代的应用及前景, 其中也不乏与可视化密切相关的系统和工作。如 FlapAssist^[10]利用 VR 技术可视化工厂进行设计并

评估设计。在该工作中, 点云被用于还原设备来评估其间的可见度, 如图 9(a) 所示。而图 9(b) 中也显示了 Lindskog 等^[11]尝试用点云记录多种类型的设备, 便于在 AR 及 VR 环境下共同合作设计工厂。

3.3 时空结合的可视化

在某些案例中, 时空数据维度的组合能呈现更多信息。在 ViDX 中^[30], 研究者使用马雷图可视觉化工厂的装配线。如图 10 所示, 横轴表示时间, 纵轴表示机组 (即空间位置), 一条线表示一个产品从头至尾的装配流程。从图中可以发现, 12 点附近和别处密度差异极大, 并且连续一系列



图 9 用点云表示位置的示例^[10-11]

Fig. 9 Example of using points cloud to indicate a location^[10-11]

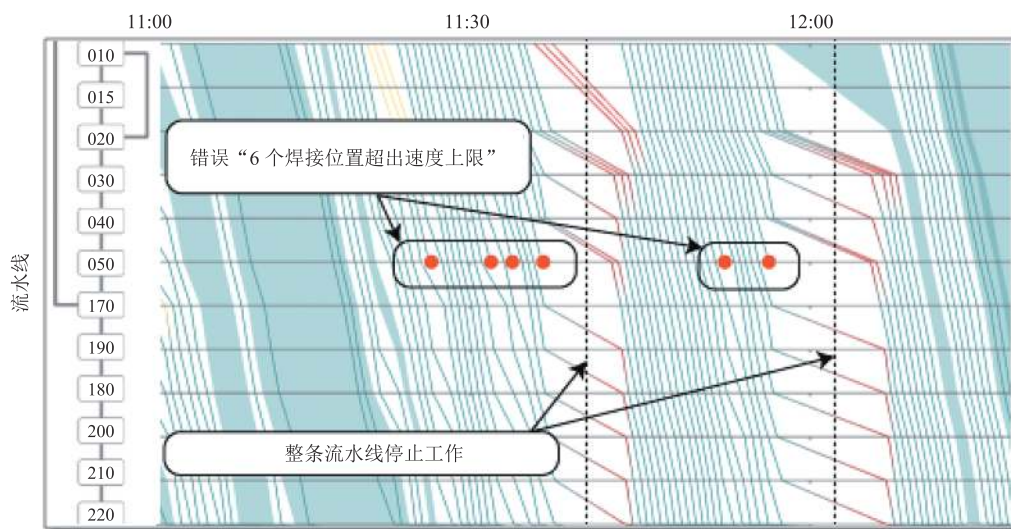


图 10 马雷图^[30]

Fig. 10 Marey chart^[30]

的产品在各个设备上都出现了问题。这也证明了时间和空间结合能够更好地反映级联问题。

4 工业数据的智能可视分析方法

4.1 工业数据的可视分析及应用

目前, 已经有许多可视分析系统被设计出来, 并服务于工业分析的各个方面(如布局改造、日程规划、运行监控、污染处理等)。

布局改造方面, KnowIME^[29]以知识图谱的形式可视化了与某一设备相近的所有设备, 便于调整或选择相近的备选项。BlueCollar^[13]系统对工人走动路线进行可视分析, 并利用结果调整工厂布局, 借此提高工厂效率。如图 11 所示, 概览视图是工厂当前的布局, 机器间连线是历史数据中的工人路径, 机器外框颜色编码了算法得出的建议调整程度, 颜色越深说明越需要调整。当用户选中某个机器后, 热力图会提示用户将该机

器调到各个位置的提效概率。Gebhardt 等^[10]的工作和 Lindskog 等^[11]的工作都利用了 VR 技术, 其中一个用于辅助合作设计工厂布局, 另一个用于评估。Yang 等^[35]利用增强现实方法优化点云配准算法, 辅助还原工厂。Büttner 等^[53]的工作探究了 AR 技术在训练工人装配上的意义。

日程规划方面, Wörner 等^[31]的工作结合甘特图、3D 还原场景、工件详细信息面板实现

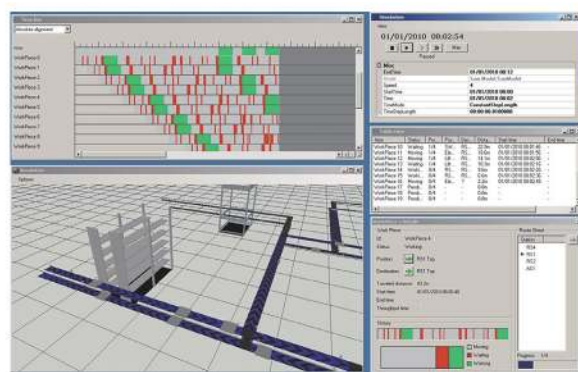


图 12 实时日程规划系统^[31]

Fig. 12 Real-time schedule planning system^[31]

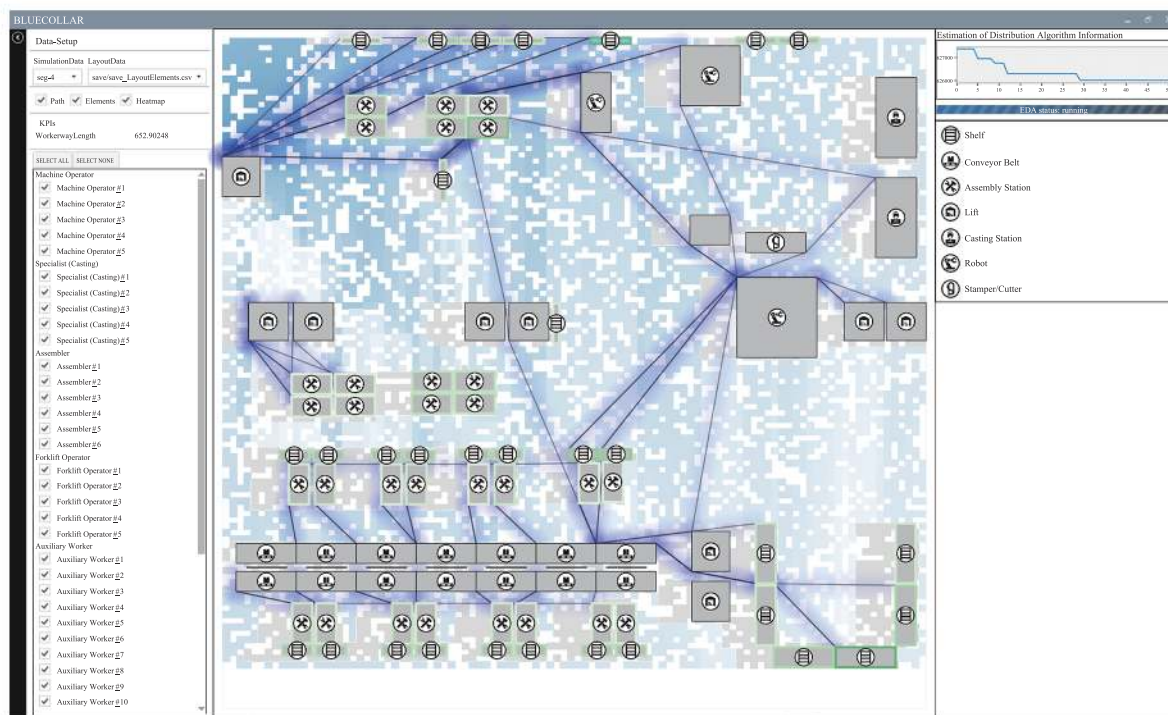


图 11 BlueCollar 的系统示意图^[13]

Fig. 11 BlueCollar's overview^[13]

了对实时日程规划或突发生产需求的支持。如图 12 所示, 左上角视图呈现了各生产任务对资源的占用情况, 左下角视图监控了工厂中实际生产的流水线, 右侧是更加细节的任务信息。根据左上角可以快速发现机器的工作情况, 并结合生产需求找到闲瑕区间插入突发任务或进行后续日程规划。LiveGantt 系统^[17]中有 3 个视图, 如图 13 所示, (a) 记录了探索的视图序列; (b) 是主要视图, 横轴编码时间, 每个横条代表一个任务, 黑色粗线表示用户当前关心的时间, 高亮最近的任務, 当用户点击关心的任务后, 高亮其

后的任务序列; (c) 记录了随时间变化的工作负载, 橙色表示增加、蓝色表示减少, 饱和度越高斜率越大即变化程度越大。相较于早期的甘特图, LiveGantt 一方面智能地对任务进行了排序, 使多个占用同类资源的任务更加靠近, 另一方面对占用同类资源的任务进行了可视化上的聚合, 图 14(a~d) 表明, 这种可视化的聚合进一步清楚明了地呈现了生产任务。PlanningVis^[33]提供了更加细节的视图来对比多种生产计划——在选中等待优化的生产计划之后, 备选的计划会在效率、关联性等各个方面分别用折线图、平行坐标

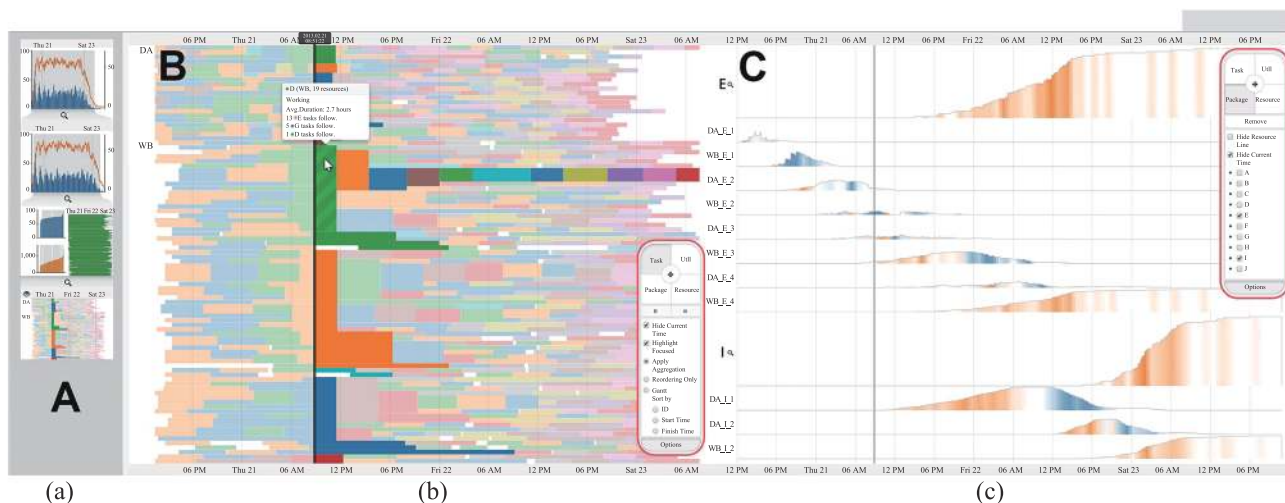


图 13 LiveGantt 的系统示意图^[17]

Fig. 13 LiveGantt's overview^[17]

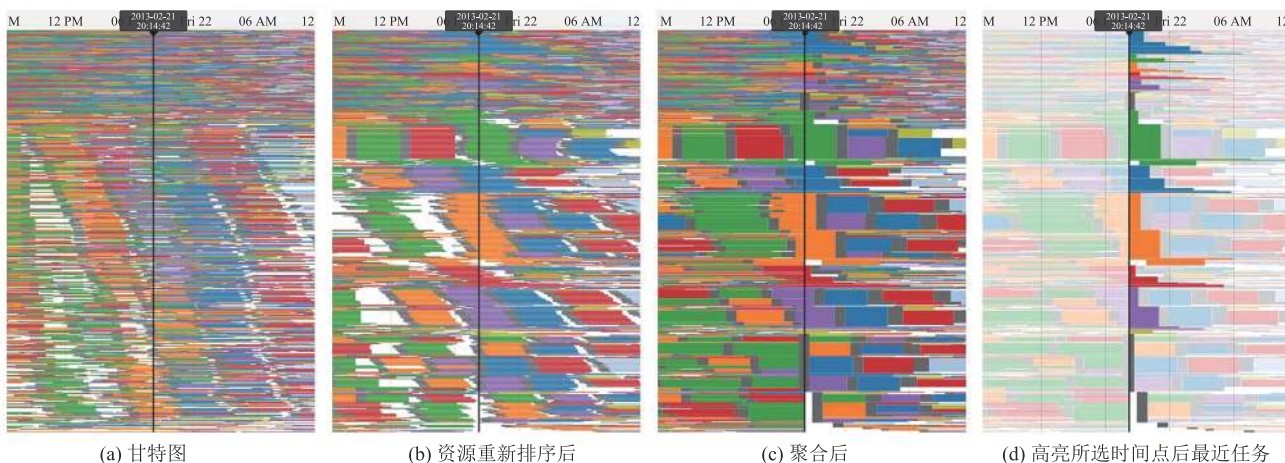
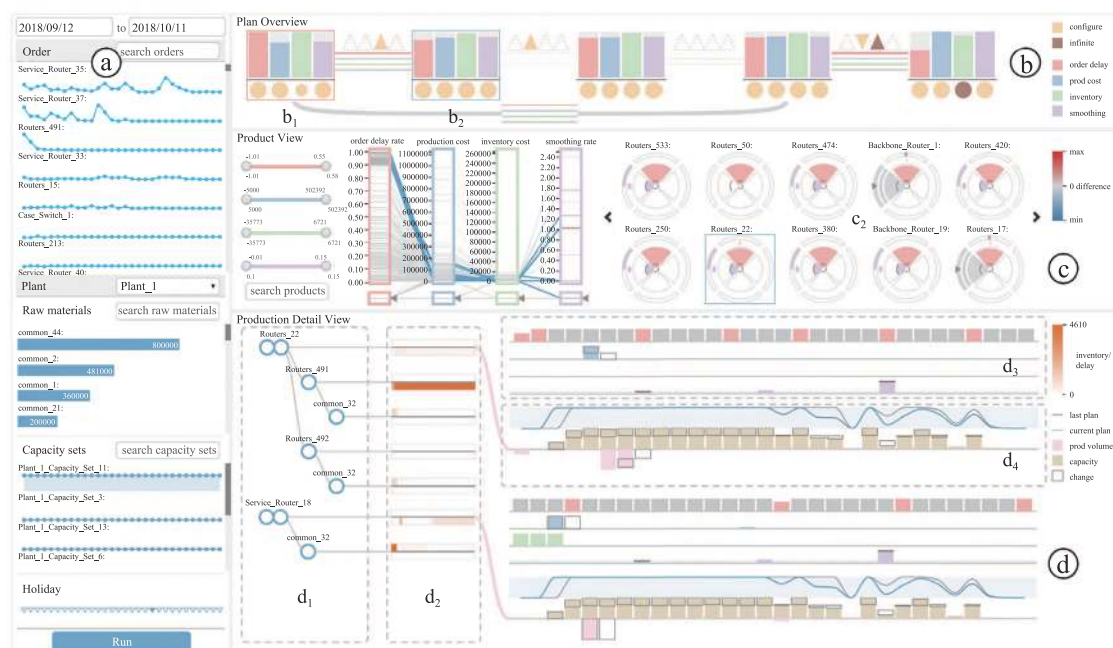


图 14 LiveGantt 的优化示意图^[17]

Fig. 14 The optimization process of LiveGantt^[17]

图 15 PlanningVis 的系统示意图^[33]Fig. 15 PlanningVis's overview^[33]图 16 增强现实下的仪表盘^[21]Fig. 16 Dashboard in augmented reality^[21]

轴等进行更加细粒度的对比，而且系统高效的处理性能支持实时的计划调整，如图 15 所示。

运行监控方面实际上包括数据监控和异常分析。Kostoláni 等^[21]采用 AR 技术，实现对工厂的监控，使人行走在工厂中就能看到数字仪表盘和信息图表，如图 16 所示。Mahmoodpour 等^[22]为不同用户角色设计了不同的可视化面板来实现监控和调整。图 17 中的 ViDX^[30]是一个非常典型的可视分析系统，利用马雷图、环形图和 3D 设备

还原来分析装配线上的异常。

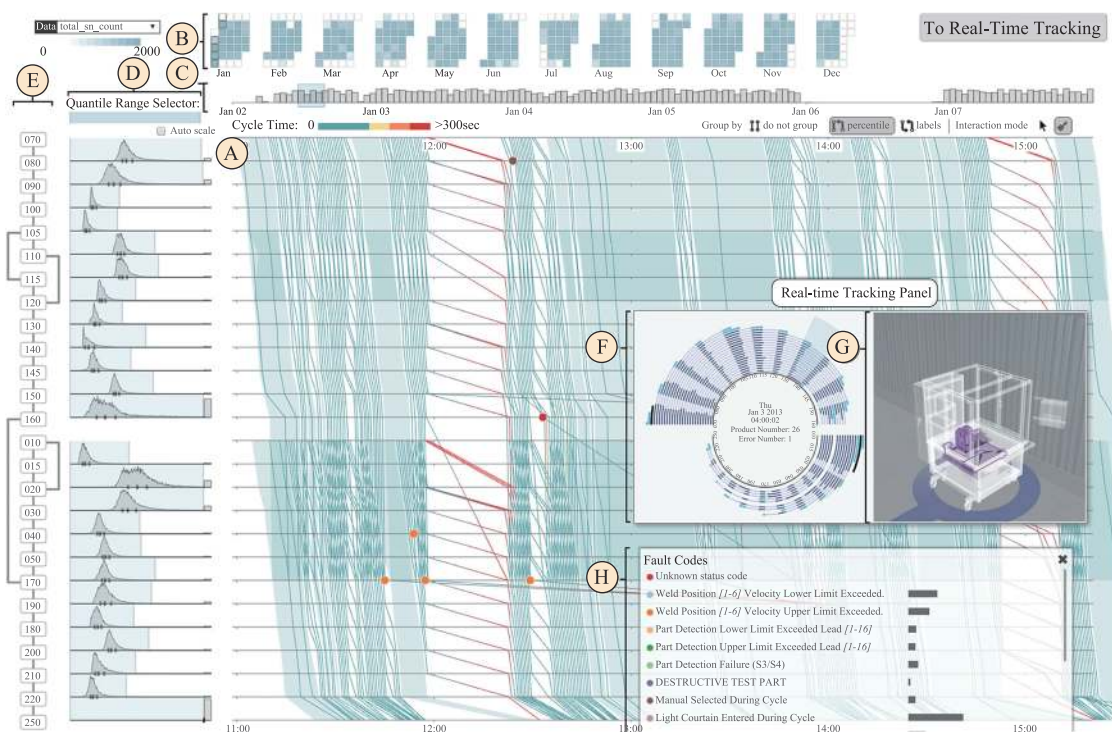
产品运维方面，Castor 等^[24]结合折线图和流程图对废水废气进行可视分析，Huettnerberger 等^[25]利用可视化方法对车辆进行质量检测。产品运维阶段的可视分析应用相较于其他阶段还不多，但物流数据、售后数据等都有着丰富的信息价值和极大的探索空间，今后可能会出现更多相关的应用。

4.2 与自动分析方法的结合

在可视分析系统中，合理地集成智能的机器学习算法也尤为重要，一方面人类智慧可以进一步提升算法的准确度，另一方面机器算力可以帮助人类提升计算效率。包括异常检测、优化决策在内的问题都有许多已经很成熟的方法。

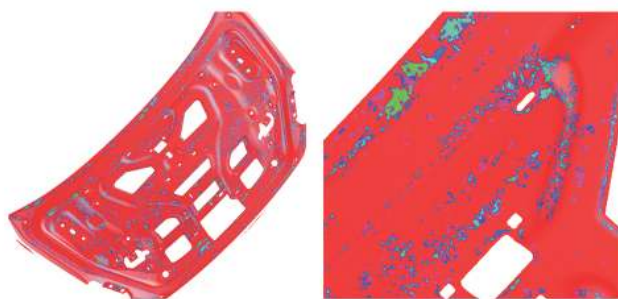
4.2.1 异常检测

异常检测包括无监督、半监督和全监督算法。无监督算法中基于聚类的 K-means^[54]和 EM 算法^[55]非常具有代表性，可以通过探索离群点来找到异常。半监督和全监督算法中常用的则有分类树、遗传算法神经网络等。

图 17 ViDX 的系统示意图^[30]Fig. 17 ViDX's overview^[30]

在工业数据的应用场景下, 需要根据数据的特征、可视化的形式, 并结合实际需求对异常进行合适的定义。一方面, 数据不一定如聚类图中的离群点一样具有显著的异常特征; 另一方面, 在不同的可视化形式下, 异常的呈现形式也会随之变化。如 BlueCollar^[13]基于分布估计算法对布局中的每一个设备都进行合理程度的评估, 通过量化布局的优劣获得分数来定义该场景下的异常。Huettnerberger 等^[25]在三角网格上采用帕累托前沿的思想, 对产品可能存在问题的部分通过加权计算进行突出可视化, 如图 18 所示, 在连续的空间上用不同颜色编码异常(颜色越绿意味着异常越严重)。

异常检测的自动分析方法已有较为成熟的体系, 在可视分析中对异常的编码也有许多不同的尝试, 如位置(散点图中的离群点)、颜色(如图 18)、标签(文字标注)等, 这些工作的可视化重点在于“展示”异常。然而, ViDX 等工作是

图 18 可视化异常检测^[25]Fig. 18 Visual anomaly detection^[25]

基于可视化对异常进行“检测”和“探索”, 这类可视分析系统在结合领域知识、找寻未知异常方面有自动化算法不可替代的启迪作用。

4.2.2 优化决策

工业场景中的优化决策就是对工业资源进行调配控制, 即得到合理的配置方式。许多配置决策问题都可以抽象为求解多目标优化问题, 也就是通过搜索解空间找到帕累托最优, 整数规划可以完成这项任务。PlanningVis^[33]生成备选生产计

划时就考虑采用整数规划,但考虑到耗时较长,所以该工作采用了一种结合线性规划^[56]和启发式算法^[57]的混合优化算法^[58]。

强化学习的方法也常被用于优化决策,Zhan等^[59]利用强化学习对火力发电厂进行了效率优化。此外,还有诸如贝叶斯模型、遗传算法等机器学习方法^[32]也被用于优化决策。

5 总 结

进入工业 4.0 时代,智能制造既是机遇又是挑战。工业数据可视化在解决大规模、高维度、结构多变、内容复杂的工业数据导致的问题上依然发挥极为重要的作用。本文概览了智能制造背景下的可视化技术、可视分析系统及其在工业场景的应用,按生产阶段和数据属性归纳并总结了常用的工业数据,又根据数据属性的共性按时间、空间、时空组合属性呈现了可视化方法,并介绍了现有的可视分析系统及工具,最后讨论了与自动分析算法的结合。

随着数据采集方式的丰富、可视化形式的创新、机器算力的腾飞,工业 4.0 时代的可视分析仍存在巨大发展空间。

从数据来源的角度,此前工业 4.0 曾被概括为三大主题:“智能工厂”、“智能生产”、“智能物流”。本文所收集文章大都围绕制造业展开研究,也就是“智能生产”,而在重工业、建筑业、物流业方面存在不少空白。特别是物流服务发展飞速的今天,工业 4.0 下一阶段的信息化或许会转向运维和服务。

从可视化形式的角度,虽然诸如折线图、直方图的传统图表已经足以支持大量的基础可视化,然而对复杂数据的探索不会止步于了解简单图表即可呈现的统计信息。未来可能的创新方向包括:(1)通过创新的算法为传统图表添加智慧,如 LiveGantt^[17];(2)基于新的数据设计新

的可视化形式;(3)基于新的交互为可视分析提效,随着人机交互的不断发展,如增强现实、虚拟现实等新的交互形式必将为可视分析注入更多新鲜血液。

从智能分析方法的角度,许多算法发展已经非常成熟,如异常检测、优化决策、模式挖掘,但仍需要进一步地探索和调整其在可视分析系统中的应用。如在考虑与深度学习的结合时如何更好地兼顾可视分析的可解释性和深度学习的优秀效果。可视化作为“完整解决方案的粘合剂”在工业数据分析问题上可以更好地与机器合作,将来有望进一步探索自动分析方法和可视分析的组合。

参 考 文 献

- [1] Drath R, Horch A. Industrie 4.0: hit or hype? [Industry Forum] [J]. IEEE Industrial Electronics Magazine, 2014, 8(2): 56-58.
- [2] Annunziata M, Evans PC. The industrial internet@work [EB/OL]. (2013.01.06)[2021.09.29]. https://www.ge.com/jp/sites/www.ge.com.jp/files/Japan_IndustrialInternetatWork_0106s.pdf.
- [3] Evans PC, Annunziata M. Industrial internet: pushing the boundaries [J]. General Electric Reports, 2012: 488-508.
- [4] Auschitzky E, Hammer M, Rajagopaul A. How big data can improve manufacturing [J]. McKinsey & Company, 2014: 822.
- [5] Liu SX, Cui WW, Wu YC, et al. A survey on information visualization: recent advances and challenges [J]. The Visual Computer, 2014, 30(12): 1373-1393.
- [6] Keim D, Kohlhammer J, Ellis G, et al. Mastering the Information Age: Solving Problems with Visual Analytics, Eurographics Association [M]. Goslar: Eurographics Association, 2010.
- [7] Posada J, Toro C, Barandiaran I, et al. Visual computing as a key enabling technology for industrie 4.0 and industrial internet [J]. IEEE Computer Graphics and Applications, 2015, 35(2): 26-40.

- [8] Zhou F, Lin X, Liu C, et al. A survey of visualization for smart manufacturing [J]. *Journal of Visualization*, 2019, 22(2): 419-435.
- [9] 布轩. 《工业数据分类分级指南(试行)》解读 [N]. 人民邮电报, 2020-03-11(004).
Bu X. Interpretation of “Guidelines for Classification and Classification of Industrial Data (Trial)” [N]. *People’s Posts and Telecommunications News*, 2020-03-11(004).
- [10] Gebhardt S, Pick S, Voet H, et al. FlapAssist: how the integration of VR and visualization tools fosters the factory planning process [C] // 2015 IEEE Virtual Reality (VR), 2015: 181-182.
- [11] Lindskog E, Berglund J, Vallhagen J, et al. Visualization support for virtual redesign of manufacturing systems [J]. *Procedia CIRP*, 2013, 7: 419-424.
- [12] Post T, Ilsen R, Hamann B, et al. User-guided visual analysis of cyber-physical production systems [J]. *Journal of Computing and Information Science in Engineering*, 2017, 17(2): 021005.
- [13] Herr D, Grund S, Ertl T. BlueCollar: optimizing worker paths on factory shop floors with visual analytics [C] // *Proceedings of the Hawaii International Conference on System Sciences*, 2019.
- [14] Girardeau-Montaut D, Roux M, Marc R, et al. Change detection on points cloud data acquired with a ground laser scanner [J]. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2005, 36(3): W19.
- [15] Wells LJ, Megahed FM, Camelio JA, et al. A framework for variation visualization and understanding in complex manufacturing systems [J]. *Journal of Intelligent Manufacturing*, 2012, 23(5): 2025-2036.
- [16] Bringezu S, Moriguchi Y. *Green Accounting* [M]. Routledge, 2018: 149-166.
- [17] 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.
- [18] Wörner M, Ertl T, Miksch S, et al. Visual analysis of advanced manufacturing simulations advanced manufacturing simulations [C] // *EuroVA*, 2011.
- [19] Herr D, Beck F, Ertl T. Visual analytics for decomposing temporal event series of production lines [C] // *Proceedings of the 22nd International Conference Information Visualisation*, 2018: 251-259.
- [20] Zhou FF, Lin XR, Luo XB, et al. Visually enhanced situation awareness for complex manufacturing facility monitoring in smart factories [J]. *Journal of Visual Languages & Computing*, 2018, 44: 58-69.
- [21] Kostoláni M, Murín J, Kozák Š. Intelligent predictive maintenance control using augmented reality [C] // *Proceedings of the 22nd International Conference on Process Control*, 2019: 131-135.
- [22] Mahmoodpour M, Lobov A, Lanz M, et al. Role-based visualization of industrial IoT-based systems [C] // *Proceedings of the 14th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications*, 2018: 1-8.
- [23] Guo HL, Yu YT, Skitmore M. Visualization technology-based construction safety management: a review [J]. *Automation in Construction*, 2017, 73: 135-144.
- [24] Castor J, Borowicz J, Burks A, et al. MC2-Mining factory pollution data through a Spatial-Nonspatial flow approach (honorable mention for clarity in visual communication) [C] // 2017 IEEE Conference on Visual Analytics Science and Technology, 2017: 203-204.
- [25] Huettnerberger L, Feige N, Ebert A, et al. Application of Pareto sets in quality control of series production in car manufacturing [C] // *Proceedings of the IEEE Pacific Visualization Symposium (PacificVis)*, 2015: 135-139.
- [26] Lei ZC, Zhou H, Hu WS, et al. Towards a web-based digital twin thermal power plant [J]. *IEEE Transactions on Industrial Informatics*, 2021: 1-10.
- [27] Qabbaah H, Sammour G, Vanhoof K. Using k-means clustering and data visualization for monetizing logistics data [C] // 2019 2nd International Conference on new Trends in Computing Sciences, 2019: 1-6.

- [28] Tang JX, Zhou YH, Tang T, et al. A visualization approach for monitoring order processing in e-commerce warehouse [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2021.
- [29] Yan HH, Yang J, Wan JF. KnowIME: a system to construct a knowledge graph for intelligent manufacturing equipment [J]. *IEEE Access*, 2020, 8: 41805-41813.
- [30] Xu PP, Mei HH, Ren L, et al. ViDX: visual diagnostics of assembly line performance in smart factories [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2016, 23(1): 291-300.
- [31] Li M, Zhou Q. Industrial big data visualization: a case study using flight data recordings to discover the factors affecting the airplane fuel efficiency [C] // *IEEE International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*, 2017: 853-858.
- [32] Wu WC, Zheng YX, Chen KY, et al. A visual analytics approach for equipment condition monitoring in smart factories of process industry [C] // *Proceedings of the IEEE Pacific Visualization Symposium*, 2018: 140-149.
- [33] Sun D, Huang RF, Chen YZ, et al. PlanningVis: a visual analytics approach to production planning in smart factories [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2019, 26(1): 579-589.
- [34] Mountains Out of Molehills Interactive [EB/OL]. (2019.10.14)[2021.09.29]. <https://informationisbeautiful.net/visualizations/mountains-out-of-molehills/>.
- [35] Yang X, Yang JF, He HW, et al. A hybrid 3D registration method of augmented reality for intelligent manufacturing [J]. *IEEE Access*, 2019, 7: 181867-181883.
- [36] Paelke V. Augmented reality in the smart factory: supporting workers in an industry 4.0 environment [C] // *Proceedings of the IEEE Emerging Technology and Factory Automation*, 2014: 1-4.
- [37] Bach B, Shi CL, Heulot N, et al. Time curves: folding time to visualize patterns of temporal evolution in data [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2015, 22(1): 559-568.
- [38] Stopar L, Skraba P, Grobelnik M, et al. Streamstory: exploring multivariate time series on multiple scales [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2018, 25(4): 1788-1802.
- [39] Aigner W, Miksch S, Schumann H, et al. *Visualization of time-oriented data* [M]. Springer Science & Business Media, 2011.
- [40] Brehmer M, Lee B, Bach B, et al. Timelines revisited: a design space and considerations for expressive storytelling [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2016, 23(9): 2151-2164.
- [41] Fulda J, Brehmer M, Munzner T. TimeLineCurator: interactive authoring of visual timelines from unstructured text [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2015, 22(1): 300-309.
- [42] André P, Wilson ML, Russell A, et al. Continuum: designing timelines for hierarchies, relationships and scale [C] // *Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology*, 2007: 101-110.
- [43] Tominski C, Schumann H. Enhanced interactive spiral display [C] // *The Annual SIGRAD Conference Special Theme*, 2008: 53-56.
- [44] Wang CC, Zhu LG, Zhang D. Application of fisheye view in temporal spiral visualization [C] // *International Conference on Applied Computing and Information Technology/International Conference on Computational Science and Intelligence (ACIT-CSI)*, 2016: 389-393.
- [45] Tanahashi Y, Ma K-L. Design considerations for optimizing storyline visualizations [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2012, 18 (12): 2679-2688.
- [46] Tang T, Li RZ, Wu XK, et al. Plotthread: creating expressive storyline visualizations using reinforcement learning [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2020, 27(2): 294-303.
- [47] Chen YZ, Xu PP, Ren L. Sequence synopsis:

- optimize visual summary of temporal event data [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2017, 24(1): 45-55.
- [48] Wongsuphasawat K, Gotz D. Exploring flow, factors, and outcomes of temporal event sequences with the outflow visualization [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2012, 18(12): 2659-2668.
- [49] Havre S, Hetzler B, Nowell L. ThemeRiver: visualizing theme changes over time [C] // *Proceedings of the IEEE Symposium on Information Visualization*, 2000: 115-123.
- [50] Wilson JM. Gantt charts: a centenary appreciation [J]. *European Journal of Operational Research*, 2003, 149(2): 430-437.
- [51] Gattullo M, Evangelista A, Uva AE, et al. What, how, and why are visual assets used in industrial augmented reality? A systematic review and classification in maintenance, assembly, and training (from 1997 to 2019) [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2020: 1-15.
- [52] Hamid NSS, Aziz FA, Azizi A. Virtual reality applications in manufacturing system [C] // *Proceedings of the Science and Information Conference*, 2014: 1034-1037.
- [53] Büttner S, Prilla M, Röcker C. Augmented reality training for industrial assembly work—are projection-based AR assistive systems an appropriate tool for assembly training? [C] // *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020: 1-12.
- [54] Krishna K, Murty MN. Genetic K-means algorithm [J]. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, 1999, 29(3): 433-439.
- [55] McLachlan GJ, Krishnan T. The EM algorithm and extensions [M]. John Wiley & Sons, 2007.
- [56] Shapiro JF. Mathematical programming models and methods for production planning and scheduling [J]. *Handbooks in Operations Research and Management Science*, 1993, 4: 371-443.
- [57] Li YJ, Chen J, Cai XQ. Heuristic genetic algorithm for capacitated production planning problems with batch processing and remanufacturing [J]. *International Journal of Production Economics*, 2007, 105(2): 301-317.
- [58] Sahling F, Buschkühl L, Tempelmeier H, et al. Solving a multi-level capacitated lot sizing problem with multi-period setup carry-over via a fix-and-optimize heuristic [J]. *Computers & Operations Research*, 2009, 36(9): 2546-2553.
- [59] Zhan XY, Xu HR, Zhang Y, et al. DeepThermal: combustion optimization for thermal power generating units using offline reinforcement learning [Z/OL]. *arXiv Preprint arXiv: 2102.11492*, 2021.