

数据驱动构造流程工业系统状态特征谱的方法

陈倩¹, 孙锴^{1,2+}

(1. 西安建筑科技大学 机电工程学院, 陕西 西安 710055;

2. 中国科学院 微电子研究所, 北京 100029)

摘要: 流程工业系统是一种典型的分布式复杂机电系统, 其分布式控制系统(DCS)实时记录着整个系统所有设备的运行信息。为了挖掘出蕴含在 DCS 监测数据集中的系统运行健康状态信息, 全面掌握系统总体运行健康状态, 快速准确地辨识和定位故障, 提出一种名为状态特征谱的方法, 将蕴含在海量高维数据中的系统运行健康状态信息以图像的形式展现在二维平面上。通过观察和分析状态特征谱上像素点的位置和颜色的分布, 在企业层面全面掌握系统运行健康状态。分别以化工领域通用的田纳西仿真数据和某化工企业 DCS 数据集为实例, 构造了对应的状态特征谱, 验证了所提方法的先进性、创新性和有效性。

关键词: 流程工业; 分布式控制系统; 数据分析; 多变量时序数据; 状态特征谱

中图分类号: N945.17; TH86 **文献标识码:** A

Forming state feature spectrum of process industry based on its DCS data set

CHEN Qian¹, SUN Kai^{1,2+}

(1. School of Mechanical and Electrical Engineering, Xi'an University of
Architecture and Technology, Xi'an 710055, China;

2. Institute of Microelectronic of Chinese Academy of Science, Beijing 100029, China)

Abstract: Production system of process industry is a typical distributed complex electromechanical system, whose Distributed Control System (DCS) records operational information of all the equipment of the whole system in real time. For the purpose of mining system health status information contained with its DCS data set, fully grasping overall operating health status, quickly and accurately identifying and locating faults, a method named state feature spectrum was proposed, which displayed operating health status concealing inside DCS data set on a two-dimensional plane by the form of digital images. By observing and analyzing position and color distribution of pixels on the spectrum, the plant-wide operating health status could be fully comprehended. DCS data set of Tennessee-Eastman process and a real chemical plant were both used to form their state feature spectrum correspondingly verifying the advancement, innovation and effectiveness of the method.

Keywords: process industry; distributed control system; data analysis; multivariable time series; state feature spectrum

0 引言

以能源重化工企业为代表的流程工业系统在我国国民经济中起着极其重要的作用, 是我国制造业不可或缺的组成部分。流程工业系统由一系列复杂

的机电设备组成, 各设备种类和数量庞大, 是一类典型的分布式复杂机电系统, 其生产对象多数属于高温、高压、易燃易爆等物质, 发生火灾、爆炸、污染物泄露和人员中毒等安全事故的概率远高于其他生产企业。在流程工业系统中, 生产设备的安全管理与

收稿日期: 2020-08-13; 修订日期: 2020-10-19。Received 13 Aug. 2020; accepted 19 Oct. 2020.

基金项目: 国家自然科学基金青年科学基金资助项目(51705393)。Foundation item: Project supported by the National Natural Science Foundation, China (No. 51705393).

运行健康状态监测是永恒的主题^[1]。

流程工业系统内部结构较为复杂,各设备之间联系紧密,具有高度耦合性,导致需要监测和控制的回路数数以千计,并且传感器需将监测到的数据实时地上传给分布式控制系统,因此其分布式控制系统实时记录着整个系统的运行数据,保存了所有的系统特征信息。系统在连续生产过程中产生的物质流、能量流和信息流三者之间相互交换,导致系统监测数据日益剧增,具有海量性、耦合性和非线性关系,很难建立数学解析表达式^[2-4]。从系统层面直观地展示系统的运行健康状态及动态演化过程,对于分析复杂机电系统的运行状态至关重要。

分布式控制系统(Distributed Control System, DCS)数据集是由数以千记的时间序列构成,本质上是一个多元时间序列。传统的DCS数据分析方法广泛采用多元统计分析方法^[5-7],具有代表性的有主元分析法(Principal Component Analysis, PCA)^[8]、独立分量分析法(Independent Component Analysis, ICA)^[9]、偏最小二乘算法(Partial Least Squares, PLS)^[10]、核主元分析(Kernel Principal Component Analysis, KPCA)^[11-13]、神经网络-PCA分析(Neural Network-PCA, NN-PCA)^[14]和核独立分量分析(Kernel Independent Component Analysis, KICA)^[15-17]等。其主要思路是采用多变量映射方法对数据进行降维,提取少量“关键”因素数据进行分析,忽略掉大多数“非关键”信息。从历次安全事故的事后分析来看,被高度关注的“关键”因素一般来说都运行正常,而造成事故的正是所谓的“非关键”因素。这是由于流程工业系统本身所特有的设备间高度的关联性和耦合性,导致从系统工程理论的角度来看不存在“非关键”数据。因此,不降维分析DCS的所有数据是准确提取企业级系统运行健康状态信息的重要前提,也是数据分析领域面临的重大挑战。

为了避免数据降维,解决数据之间高度耦合的问题,数据可视化的方法得到了广泛的使用^[18-20]。鲁文波等^[21]通过近场声全息重新构建机械声压场和声源的可视化声像图,从声像图中提取出灰度共生的矩阵特征,采用支持向量机(Support Vector Machine, SVM)模式识别达到故障诊断的效果;LU等^[22]提出了声全息图方法,通过对比正常状况和异常状况时的灰度图像,实现了机械设备非接触式的故障模式识别;杜党党等^[23]提出一种基于故障

彩色图谱的发动机故障模式识别方法;郇绍强^[24]通过构造采油生产知识图谱及建立预警规则方法,研制了一套智能预警系统;刘鑫等^[25]用灰度图表征时域振动信号,提出一种基于灰度图像纹理分析的二维故障特征提取模型。孙锴等^[3,26]将数字图像处理技术引入到DCS数据分析中,定义了色彩相空间,将DCS监测数据集转化为数字图像,制定了构造系统图谱,以图像的形式反映系统运行状态特征的方法,有效地展现了多变量数据间的内在关系,从宏观层面将多变量数据间的高耦合性、非线性关系直观地呈现出来。其中,系统故障图谱将监测数据分为正常数据和异常数据两类,分别以白色和黑色对数据进行着色,以黑白二值图像反映了系统中异常数据的分布特征。然而,仅将监测数据划分为正常和异常两类过于简单,降低了分析的准确性,无法满足企业生产要求。对于流程工业来说,一次停机或大修的成本代价很高,企业工程师需要了解数据的异常程度作为维修设备的依据,而系统故障图谱简化了数据异常程度信息。

为了解决上述问题,本文提出了构造系统状态特征谱的方法,在系统故障图谱的数据分类基础上对DCS监测数据集中所有数据按照偏离最优值的程度进行了细化分类和着色,最大限度地保留了数据异常程度信息。通过对系统状态图谱中像素点的位置以及色彩的分布规律进行分析,可以更加准确地对系统的整体运行健康状态做出判断。

1 数据矩阵 X

DCS监控数据集可以被看作是一个多变量时间序列集,其中包含了 n 个监测变量的时间序列。每经过一个采样时间间隔,对各个监测变量参数进行一次采样,每个变量在一个监测时间间隔内可以截取 m 个采样值。将采样监测变量作为行向量,采样时间序列作为列向量进行排布,可以构造出一个 $m \times nm \times n$ 的二维矩阵 X 。 n 个监测变量构成 n 维Hausdorff拓扑空间 $S \in R^n$, m 个采样周期构成 m 维Hausdorff拓扑空间 $D \in R^m$,系统监测数据集由它们的笛卡尔乘积空间 $S \times D \in R^{m \times n}$ 构成,这是一个 $m \times n$ 维的Hausdorff拓扑空间。定义一个二维Hausdorff拓扑空间 $X \in R^2$,建立映射关系 $f: S \times D \rightarrow X$,将蕴含复杂机电系统运行健康状态的DCS监测数据集投影到二维数据空间中。

定义1 数据矩阵 X 。本文将既能反映复杂机

电系统的时间和空间分布,又能反映系统运行健康状态的矩阵,称为分布式复杂机电系统的数据矩阵 X ,

$$X = \begin{matrix} & \text{传感器} \\ \begin{matrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{matrix} & \begin{matrix} \\ \\ \\ \end{matrix} & \begin{matrix} \\ \\ \\ \end{matrix} & \text{时间。} \end{matrix} \quad (1)$$

式中 x_{ij} 表示第 j 个监测变量在第 i 个采样周期的数值。数据矩阵 X 的行向量 $X_i \in R^m$ 表示系统的 n 个监测变量在时刻 i 的一个采样值,代表着某个特定时间点的系统动态特性。数据矩阵 X 的列向量 $X_j \in R^n$ 表示系统的第 j 个监测变量的时间序列。

2 分类矩阵 Q

根据企业生产的实际工况,挑选出系统处于运行状态最佳时段下所监测到 DCS 数据集,并将该数据集中的数据作为评判系统运行状态的基准,即 DCS 标准集。通过 DCS 标准集中各列向量的平均值和方差,得到最优监测值向量 E 和异常偏离度 G 作为待测数据集的分类的基准向量。

2.1 数据分类器 T

定义 2 最优监测值向量 E 。对 DCS 标准集所构成的数据矩阵中的每一列求平均值,可以得到其

均值行向量,将该均值行向量称为最优监测值向量 E :

$$E = [\mu_1 \quad \mu_2 \quad \cdots \quad \mu_j \quad \cdots \quad \mu_n]_{1 \times n} \quad (2)$$

$$\text{式中 } \mu_j = \frac{1}{m} \sum_{i=1}^m x_{ij}。$$

定义 3 异常偏离度 G 向量。对 DCS 标准集所构成的数据矩阵中的每一列求标准差,所得到的系统标准差行向量称为异常偏离度向量 G ,

$$G = [\sigma_1 \quad \sigma_2 \quad \cdots \quad \sigma_j \quad \cdots \quad \sigma_n]_{1 \times n} \quad (3)$$

$$\text{式中 } \sigma_j = \sqrt{\frac{1}{m} \sum_{i=1}^m (x_{ij} - \mu_j)^2}。$$

定义 4 标准质量区间。根据异常偏离度对 DCS 监测数据进行分类的基准区间。当变量 j 的监测数据在标准区间中时,表明系统处于最佳运行状态,如式(4)所示:

$$\mu_j - \sigma_j \leq x_j \leq \mu_j + \sigma_j \quad (4)$$

根据定义 2 和定义 3,将传感器变量 j 的监测数据划分到不同的质量间隔中,如图 1 所示。图中 h 表示偏离标准值的异常偏离值,偏离标准区间的值越高,表示系统的运行故障程度越高。

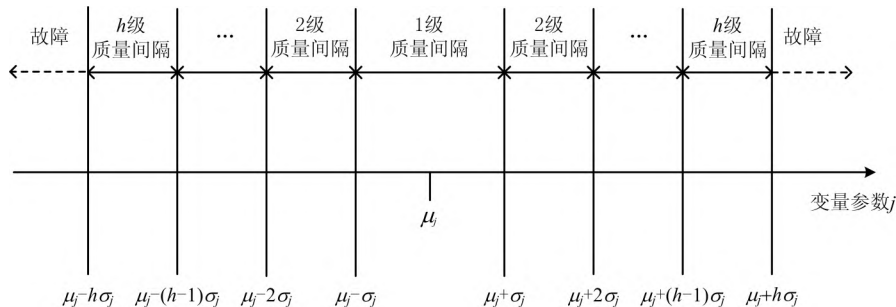


图1 变量 j 的时间变换序列

根据标准质量区间,将每个参数变量在一段采样周期内监测的采样值通过数据分类器 T 进行分类,将各变量的监测值按偏离标准值的不同程度分类到相应的作用域内,构成数据分类矩阵 Q 。

定义 5 数据分类器 T 。由系统各变量的标准值和各等级异常偏离度值所构成的矩阵,称其为数据分类器 T ,

$$T = \begin{bmatrix} \mu_1 + (h+1)\sigma_1 & \mu_2 + (h+1)\sigma_2 & \cdots & \mu_j + (h+1)\sigma_j & \cdots & \mu_N + (h+1)\sigma_N \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mu_1 + \sigma_1 & \mu_2 + \sigma_2 & \cdots & \mu_j + \sigma_j & \cdots & \mu_N + \sigma_N \\ \mu_1 - \sigma_1 & \mu_2 - \sigma_2 & \cdots & \mu_j - \sigma_j & \cdots & \mu_N - \sigma_N \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mu_1 - (h+1)\sigma_1 & \mu_2 - (h+1)\sigma_2 & \cdots & \mu_j - (h+1)\sigma_j & \cdots & \mu_N - (h+1)\sigma_N \end{bmatrix} \quad (5)$$

该数据分类器的主要特点是将难以辨识的浮点型数据以简单的整型数值代替并进行分类,对变量

j 监测数据的分类过程如式(6)所示:

$$T_j = \begin{cases} 0 & \mu_j - \sigma_j \leq x_{*j} \leq \mu_j + \sigma_j \\ 1 & \mu_j - 2\sigma_j \leq x_{*j} \leq \mu_j - \sigma_j \cup \mu_j + \sigma_j \leq x_{*j} \leq \mu_j + 2\sigma_j \\ 2 & \mu_j - 3\sigma_j \leq x_{*j} \leq \mu_j - 2\sigma_j \cup \mu_j + 2\sigma_j \leq x_{*j} \leq \mu_j + 3\sigma_j \\ \vdots & \vdots \\ h & \mu_j - (h+1)\sigma_j \leq x_{*j} \leq \mu_j - h\sigma_j \cup \mu_j + h\sigma_j \leq x_{*j} \leq \mu_j + (h+1)\sigma_j \end{cases} \quad (6)$$

2.2 分类矩阵 Q

定义6 分类矩阵 Q 。将待测的监测数据通过数据分类器 T 进行分类后,所得到的整数型矩阵称为数据分类矩阵 Q ,

$$Q = \begin{matrix} & \begin{matrix} \text{传感器} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{matrix} \\ \begin{matrix} \cdots & 0 & 2 & 0 & 6 & 1 & 5 & \cdots \\ \cdots & 5 & 0 & 4 & 4 & 3 & 2 & \cdots \\ \vdots & 5 & 0 & \vdots & 0 & 1 & 2 & \vdots \\ \vdots & 5 & 4 & 2 & \vdots & 0 & 2 & \vdots \\ \cdots & 4 & 4 & 5 & 6 & 1 & 0 & \cdots \\ \cdots & 0 & 0 & 1 & 2 & 5 & 4 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{matrix} & \begin{matrix} \\ \\ \\ \text{时间。} \\ \\ \\ \end{matrix} \end{matrix} \quad m \times n \quad (7)$$

分类矩阵中不同元素代表不同的系统运行状态。根据三西格玛质量控制原理,当异常值在 $0 \sim 2$ 之间时,说明系统处于安全运行状态范围内。当异常值高于正常运行范围时,异常值越高,则系统异常程度越高。

3 系统状态特征谱

虽然人眼难以直观地辨识出整型数据的变化规律,并从中找出异常数据,但是人眼对色彩变化的敏感度远远高于对数字变化的敏感度。通过制定着色规则,对分类矩阵 Q 中的元素进行着色,构造系统状态特征谱。

3.1 状态特征谱的着色规则

根据 RGB 色彩规则,对 DCS 监控数据集中所有数据按照偏离最优值的程度进行细分。为了人眼有效地区分色彩,三原色 RGB 的各个分量只能在全 0 和全 1 中取,其中 RGB 三个分量取 0 和 1 为色彩的两个极端情况,本文选择 $2^3 = 8$ 种最极端的颜色表示系统运行健康状态。孙锴等^[3]在构造系统故障图谱研究中,用黑色表示故障点,白色表示无故障数据。对数据矩阵 X 着色后,以二值数字图像的形式展示系统运行健康状态。为了与黑白二值系统故障

图谱相区别,同时由于在彩色图像中,人眼对红色的辨识度高于黑色,视觉感官更强,本文用红色代替黑色表示系统变量观测值偏离最优值最远的区间,用绿色代表观测值处于最优区间,中间过渡色根据偏离最优值距离区间从近到远依次着色为黄、白、蓝、青、紫。将蕴含复杂机电系统运行健康状态的数据矩阵 X 从海量高维 Hausdorff 空间投影到二维色彩相空间中,构造出系统状态特征谱。

定义7 状态特征谱 P 。对数据分类矩阵 Q 进行染色,其中 $pixel_0$ 表示为绿色、 $pixel_1$ 表示为黄色、 $pixel_2$ 表示为白色、 $pixel_3$ 表示为蓝色、 $pixel_4$ 表示为青色、 $pixel_5$ 表示为紫色、 $pixel_6$ 表示为红色,在二维平面中可以得到彩色图像,称为状态特征谱 P ,如式(8)所示:

$$P = \{pixel_i\} = \begin{cases} \text{Green} & i = 0 \\ \text{Yellow} & i = 1 \\ \text{White} & i = 2 \\ \text{Blue} & i = 3 \\ \text{Cyan} & i = 4 \\ \text{Purple} & i = 5 \\ \text{Red} & i = 6 \end{cases} \quad (8)$$

依据式(8)所定义的着色规则,对数据分类矩阵中的数据着色,系统整体的运行健康状态通过色彩的分布和变化被呈现在一幅二维平面数字彩色图像上。

3.2 构造流程

系统状态特征谱的构造流程如图2所示。根据现场监控记录,挑选出系统运行状态最佳的一天所对应的 DCS 数据集作为运行状态的分析标准。将待测 DCS 数据集以同样的采样周期为单位划分为的独立数据集。通过构建数据分类器 T 、分类矩阵 Q ,构造系统特征谱 P ,分析系统运行健康状态。具体步骤如下:

步骤1 根据现场监控待测的 DCS 数据中,人为选取某段系统运行状态时段最佳的 DCS 监测数据集,作为标准集。

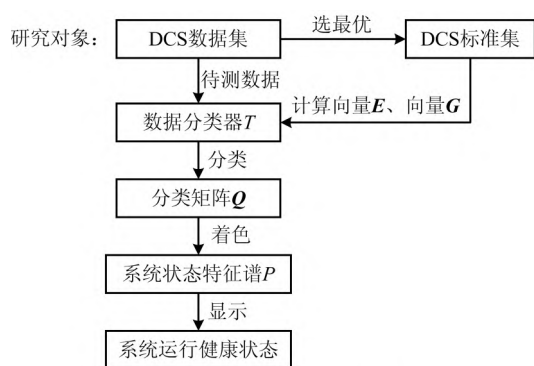


图2 系统运行健康状态分析流程图

步骤 2 对 DCS 标准数据集求取最优监测值向量和异常偏离度向量,构造数据分类器 T 。

步骤 3 将待测的 DCS 数据集通过数据分类器 T 进行分类,根据不同程度偏离最优标准值进行分类后,构建出分类矩阵 Q 。

步骤 4 按照数据着色规则对分类矩阵 Q 进行着色,构造系统状态特征谱 P 。

步骤 5 根据分析所构造的系统状态特征谱,分析系统运行健康状态。

4 算法的先进性和有效性

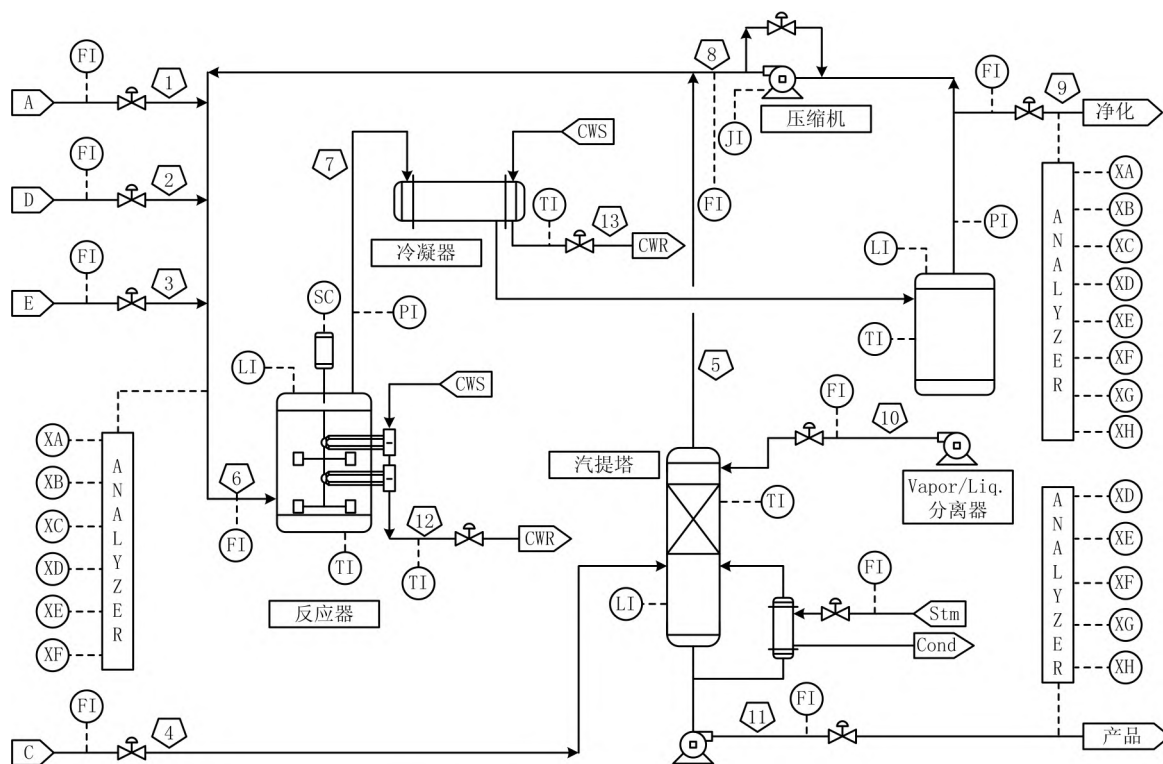
田纳西-伊斯曼过程仿真过程 (Tennessee-

Eastman Process, TEP) 是由伊斯曼化工公司提出的一种通用化学过程模拟器,是一个用微分方程模拟化工生产过程的仿真系统,其主要目的是为了评估监测方法和工程控制提供一个现实的工业流程仿真。

田纳西-伊斯曼过程主要由反应器、压缩机、冷凝器、分离器和汽提塔 5 个单元组成,包含了 A、B、...、H8 种成分,其主要工艺流程如图 3 所示。图 3 中 XA, XB, ..., XH 为输出;1、2、...、13 为流号;CWR 为冷却水回流;CWS 为水煤浆;LI、JI、TI、FI 和 PI 为变量控制器。反应器的物质流经过冷凝器冷却后,送入分离器中进行分离,将分离出来的蒸汽经过压缩机压缩后再循环送回反应器中。

TEP 包含可以模拟正常运行状态的 52 个变量,其中 XMEAS(1)~XMEAS(41)为监测变量, XMV(1)~XMV(12)为控制变量,所有变量同步采样周期为 1 min。并且 TEP 测试数据集包含 1 个无故障数据包和 21 个典型类型的故障数据包,每个数据包包括 480×52 个数据点。

传统的数据驱动故障辨识方法主要包含 PCA, 规范变量分析(Canonical Variable Analysis, CVA)和动态主元分析(Dynamic Principal Component A-

图3 田纳西-伊斯曼仿真工艺流程^[3]

analysis, DPCA), 这些方法的主要思想是基于数据降维。利用本文所提出的方法与传统数据驱动故障辨识的方法对田纳西仿真系统的 21 种典型故障的数据进行分析, 各方法的故障辨识精确度情况对比如表 1 所示。

表 1 状态特征谱方法与传统方法的对比

	分析方法			系统状态特征谱
	PCA	DPCA	CVA	
故障辨识精确度/%	52.38	52.38	71.43	100

由表 1 可知, PCA 和 DPCA 方法对于 21 种典型故障模式的故障识别精确度为 52.38%, CVA 方法的故障识别精确度为 71.43%, 而本文提出系统状态特征谱方法的故障识别精确度为 100%。通过对比, 本文所提出的系统状态特征谱的方法对故障的识别精确度更高。

根据定义 1 可知, 将 52 个监测变量所构成的时间序列作为列向量, 按照 XMEAS(1)~XMEAS(41)、XMV(1)~XMV(11)的顺序排列, 构造了 480×52 的二维矩阵。根据图 2 的系统状态特征谱的流程, 构造出 TEP 有无故障状态特征谱, 对比无故障系统故障图谱和有故障系统状态特征谱, 如图 4 所示。

从图 4a 可以看出, 无故障状态下的系统故障图谱是一副纯白色的图像, 只能反映系统设备都正常运行, 更多的信息无法被呈现出来。而如图 4b 所示, 本文所提出构造系统状态特征谱的方法在无故障状态下呈现出绿、黄、白色彩相间的一副图像, 显示了系统中各个设备在安全范围内波动的情况。这虽然表明监测数据还在正常运行的范围内, 但偏离最优值程度各有不同, 被呈现出来的信息更加具体丰富。根据本文提出的分类矩阵的概念, 将 TEP 无故障数据构成的分类矩阵中的各列进行求和, 各参数的异常程度值如图 5 所示。将监测值异常程度由高到低进行排序, 需要着重观察并予以关注的变量序列一目了然, 然而这些信息在图 4a 的故障图谱中均被抛弃掉了, 导致无法具体地呈现系统运行健康状态。

以 TEP 故障模式 4 为例, 对比系统故障图谱^[3]与系统状态特征谱绘制出的系统有故障图谱, 如图 6 所示。

从图 6a 中可以看到, 变量 51 处有一条纵穿图谱的黑色条纹和其他零星分布的黑色斑点, 这表明

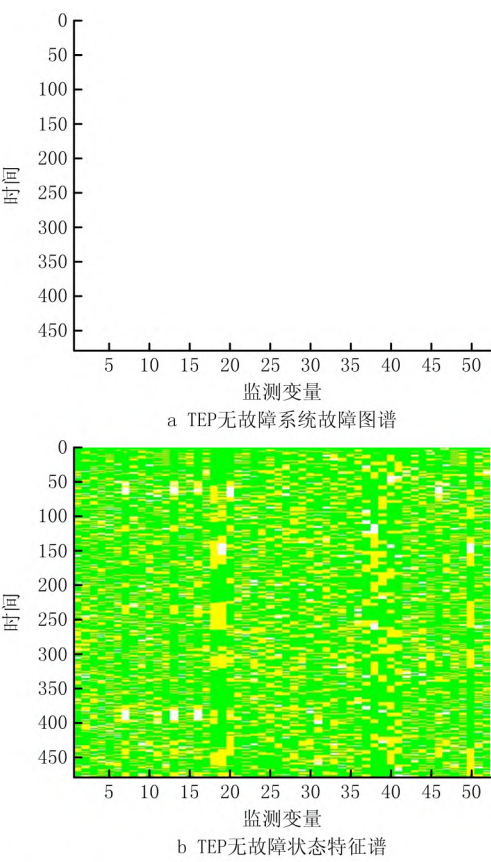


图4 无故障状态下的系统故障图谱与状态特征谱对比

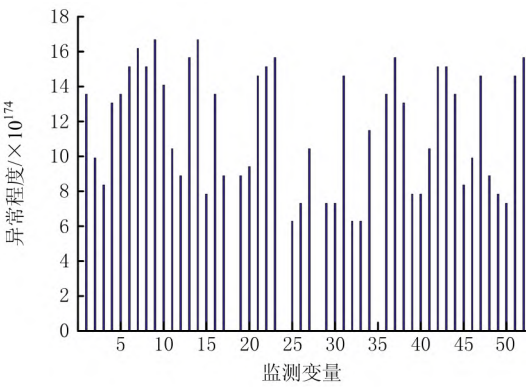


图5 TEP无故障数据各参数的总偏离度

变量 51 在整个采样周期中都出现了异常, 其他变量的监测值零星出现异常, 但很快消失。由于整个图谱上的异常值都用黑色表示, 无法区分变量 51 出现的异常值和其它变量显示的异常值的异常程度及区别, 无法判定出着重观察的区域。

由图 6b 可知, TEP 状态特征谱是由不同颜色所构成的。图 6b 中, 变量 51 是一条贯穿整个监测时序的纵向红色条纹, 红色像素代表了系统目前已经发出了严重故障的预警。而在其他位置出现部分

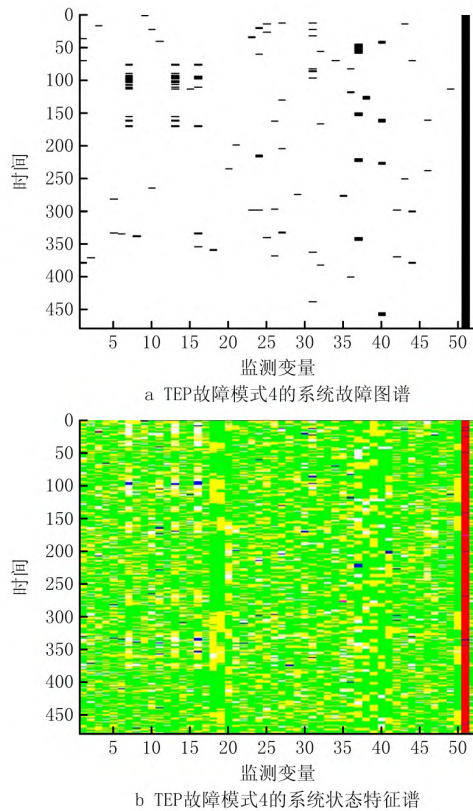


图6 故障模式4的系统故障图谱与状态特征谱对比

蓝色、青色、紫色斑点,这些不同的色彩代表着不同

的故障程度,但它们的异常严重程度远远小于变量 51“反应器冷却水流量”出现的严重故障。根据图 6b 给出的信息,工程师可以判断出其他系统变量造成的故障程度处于安全范围内,仅予以观察,但不采取措施,而是专注解决变量 51 造成的故障,这一信息无法从图 6a 中获取到。

相比于系统故障图谱而言,本文提出构造系统状态特征谱的方法不仅能够表达出系统是否处于异常状态的信息,同时通过细化系统故障的映射关系,不同的颜色代表系统不同的异常程度,根据像素点的位置来判断系统异常情况并采取相应措施,从全局上更准确地把握了系统的运行状态。

5 某化工企业空气压缩机组实例应用

某化工厂空气压缩机组的设备连接如图 7 所示。该压缩机组由汽轮机、空压机、增压机、变速箱等设备组成,设备与设备之间通过连接管进行连接,各设备间相互协调,达到空气压缩的功能,是一个典型的复杂机电系统。该化工厂空气压缩机组总共包括 250 个传感器,种类繁多,如温度、流量、压力、转速、功率等传感器。这些传感器对系统进行实时监控,实时将监测数据上传给控制系统——DCS。

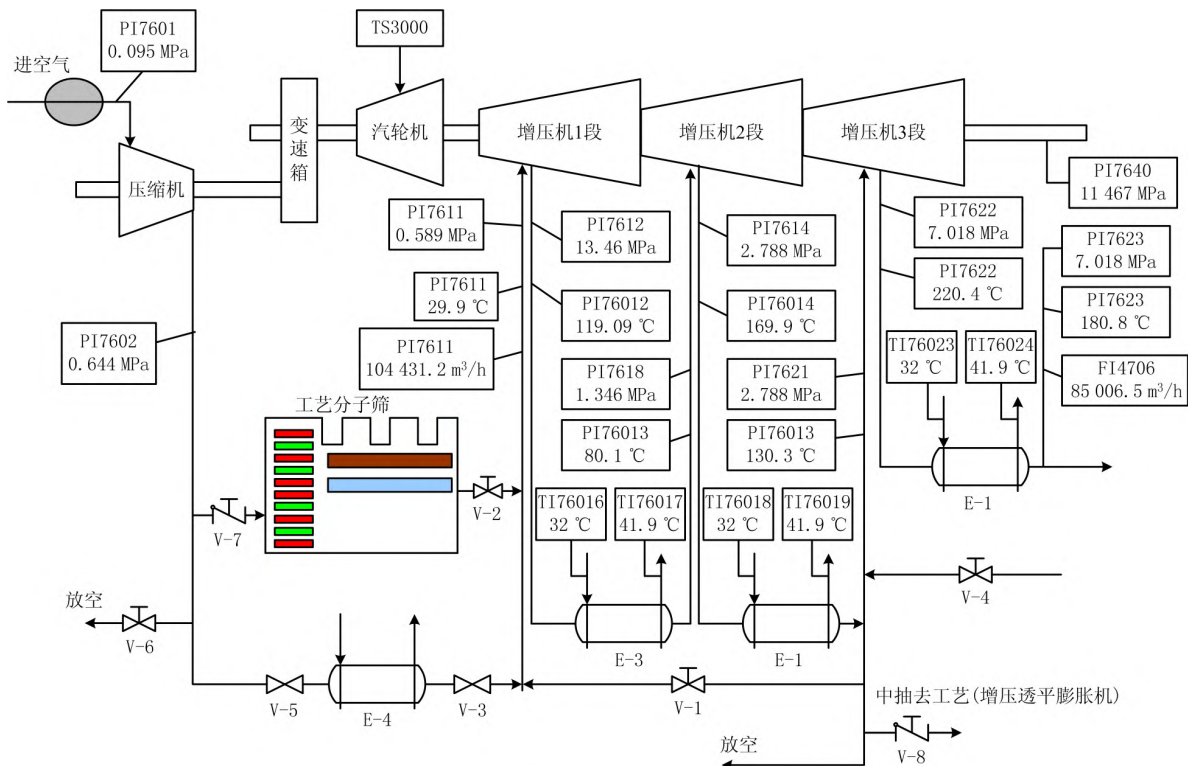


图7 空气压缩机设备连接图

对该空气压缩机组运行半年时间的监测数据进行分析。根据定义 2~定义 7,将该空气压缩机组 2013 年上半年所有的 DCS 监测数据构造出反映半年内系统运行健康状态的状态特征谱,如图 8 所示。其中横轴表示 250 个系统采样参数变量,纵轴表示系统的采样时序为 721,监测数据的测试样本数为 32 805 500 个数据点。

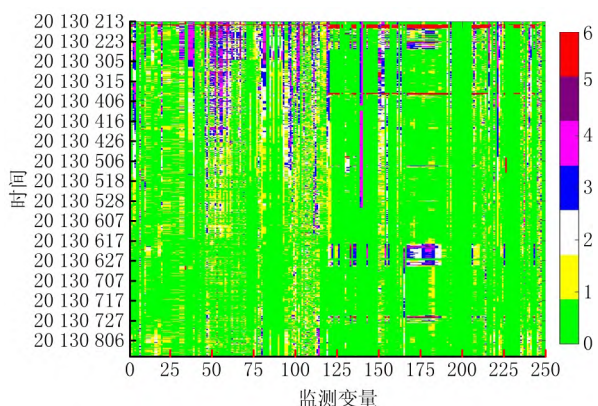


图8 某化工企业2013年上半年生产系统状态特征谱

由图 8 可知,该化工企业 2013 年上半年的系统状态图谱上分布着不同色彩的像素点。根据本文所提出的数据分类算法以及着色规则可知,图 8 中呈现的绿、黄、白 3 种颜色代表系统处于正常运行状态。其中绿色为最佳的系统运行状态,黄色和白色代表在系统允许的波动范围内正常运行,但实际上已经说明系统处于亚健康状态了,应给予留意和观察;而绿、黄、白以外的其他颜色分别代表着不同程度的故障运行状态,其中红色为严重故障状态。如图 8 所示,有两条较为醒目的红色条纹横穿整幅图谱,它们分别代表 2013 年 02 月 14 日、2013 年 03 月 25 日这两天的系统运行状态。2 月 14 日的红色条纹是横向贯穿整幅状态图谱的,说明当天所有的传感器监测的设备都在发出严重故障预警;03 月 25 日的红色条纹代表传感器 120~250 所监测的设备都在发出严重预警,此时必须要采取相应的检修措施,整个系统应该立即停车大修。因此,本文对图 8 中较为明显的两条红色条纹及前后时段作为重点进行放大分析,如图 9 和图 10 所示。

图 9 中,在 2013 年 2 月 13 日中午 12:00 左右至 14 日零点左右,整个系统状态特征谱的色彩呈红色,且基本贯穿了 250 个系统监测变量,这说明整个系统都在发出严重故障预警。2 月 14 日零点~2 月 16 日下午 13:00 左右,系统变量为 177~195、变量

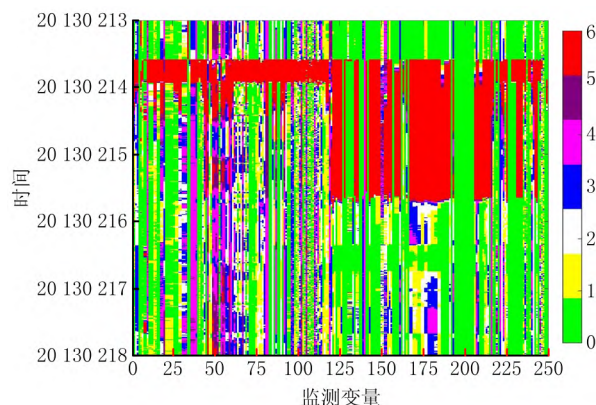


图9 某化工厂生产系统有故障部分 I 状态特征谱

210~248 左右的采样点仍持续发出严重故障预警。变量为 35~68 在系统状态特征谱的整个时序上贯穿着紫色和蓝色,代表着不同程度的故障发生,但这些异常程度不高,可将这部分时段呈现出来的系统运行状态与实际工况相结合,对此时段的系统运行状况予以关注或采取相应措施。

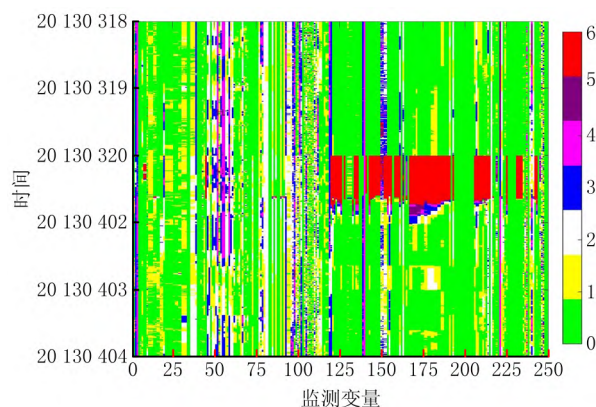


图10 某化工厂生产系统有故障部分 II 状态特征谱

如图 10 所示,传感器 120~190、210~240 在 2013 年 3 月 20 日~4 月 1 日范围内呈现着红色斑点,代表在这段时序范围内,这些系统变量发出严重故障预警。并且图 10 中的部分变量在一定的时序范围内呈现紫色像素和蓝色像素,而这些色彩同样也反映着各系统变量不同程度的故障状态,这些部分将作为重点观察区域,观察系统故障运行状态是否有向严重故障运行状态的趋势。因此,本文提出的方法是通过观察系统状态特征谱颜色的变化及分布规律,全面地把握系统运行健康状态,在重大事故发生之前准确快速地作出预警,避免不必要的损失。

6 结束语

本文提出一种基于 DCS 监测数据集构造系统状态特征谱分析系统整体运行健康状态的方法。该方法改进了之前提出的系统故障图谱无法精确反映故障程度的缺点,通过设计数据分类器和制定着色规则,对系统监测数据集中的数据进行分类和着色,将所有的数据呈现在二维图像上,构造出信息丰富细致的系统状态特征谱。通过分析像素点的位置以及颜色的分布范围,实现具体全面的掌握系统的运行健康状态,提高了分析准确度,为之后的量化分析奠定了基础。下一步的研究将利用数字图像处理技术中丰富的图像分析算法,更加充分地挖掘潜藏在系统状态特征谱中的系统健康运行状态,实现复杂机电系统运行健康状态的量化评级和快速、准确溯源。

参考文献:

- [1] CHAI Taiyou, DING Jinliang. Process industry intelligent optimized manufacturing[J]. Chinese Engineering Science, 2018, 20(4): 51-58(in Chinese). [柴天佑, 丁进良. 流程工业智能优化制造[J]. 中国工程科学, 2018, 20(4): 51-58.]
- [2] SUN Kai, GAO Jianming, GAO Zhiyong. Analysis of the health status of modern industrial system based on data-driven system color map[J]. Journal of Mechanical Engineering, 2012, 48(18): 186-191(in Chinese). [孙 锴, 高建民, 高智勇. 基于数据驱动的系统彩色图谱分析现代工业系统健康状态[J]. 机械工程学报, 2012, 48(18): 186-191.]
- [3] SUN Kai, GAO Jianming, GAO Zhiyong, et al. An enterprise-level fault pattern recognition method based on fault map[J]. Computer Integrated Manufacturing Systems, 2015, 21(2): 519-527. [孙 锴, 高建民, 高智勇, 等. 基于故障图谱的企业级故障模式识别方法[J]. 计算机集成制造系统, 2015, 21(2): 519-527.]
- [4] SUN Kai. Status analysis method of complex electromechanical system based on system spectrum[M]. Xi'an: Northwestern Polytechnical University Press, 2016(in Chinese). [孙 锴. 基于系统图谱的复杂机电系统状态分析方法[M]. 西安: 西北工业大学出版社, 2016.]
- [5] SÜRMELİ B G, TÜMER M B. Multivariate time series clustering and its application in industrial systems[J]. Cybernetics and Systems, 2020, 51(3): 315-334.
- [6] NIU Yuguang, WANG Shilin, LIN Zhongwei, et al. Boiler process fault detection based on multivariate statistical process monitor[J]. Journal of Power Engineering, 2017, 37(10): 829-836(in Chinese). [牛玉广, 王世林, 林忠伟, 等. 基于多元统计过程监控的锅炉过程故障检测[J]. 动力工程学报, 2017, 37(10): 829-836.]
- [7] WANG Qingning, YANG Xin, DAI Yiyang. Application of PPA based multivariate statistical analysis method in process fault diagnosis[J]. Computer and Applied Chemistry, 2016, 35(10): 821-832(in Chinese). [汪庆宁, 杨 鑫, 戴一阳. 基于 PPA 的多元统计分析方法在过程故障诊断的应用[J]. 计算机与应用化学, 2016, 35(10): 821-832.]
- [8] DONG Guosheng, REN Wei. Study on real-time fault monitoring system for dynamic chemical process based on PCA[J]. Computer and applied chemistry, 2010, 27(10): 1357-1360(in Chinese). [董国胜, 任 伟. 基于 PCA 的动态化工过程实时故障监测系统研究[J]. 计算机与应用化学, 2010, 27(10): 1357-1360.]
- [9] RADOJIĆ U, NORDHAUSEN K, OJA H. Notion of information and independent component analysis[J]. Application of Mathematics, 2020, 65(3): 311-330.
- [10] KUMAR K, MISHRA A K. Application of partial least square (PLS) analysis on fluorescence data of 8-anilinoanthralene-1-sulfonic acid, a polarity dye, for monitoring water adulteration in ethanol fuel[J]. Journal of Fluorescence, 2015, 25(4): 1055-1061.
- [11] PENG Hongxing, PAN Lili, ZHAO Hongtu. Abnormal diagnosis method of monitoring system sensor based on KPCA[J]. Sensors & Microsystems, 2016, 35(12): 33-36(in Chinese). [彭红星, 潘梨莉, 赵鸿图. 基于 KPCA 的监控系统传感器异常诊断方法[J]. 传感器与微系统, 2016, 35(12): 33-36.]
- [12] LI Zhijun, NIU Xiaoxu, LI Yueheng. TE process fault prediction based on KPCA-ARIMA model[J]. Industrial Control Computer, 2019, 32(4): 45-48(in Chinese). [李志军, 牛晓旭, 李月恒. 基于 KPCA-ARIMA 模型的 TE 过程故障预测[J]. 工业控制计算机, 2019, 32(4): 45-48.]
- [13] SU Huaizhi, WEN Zhiping, REN Jie. A kernel principal component analysis-based approach for determining the spatial warning domain of dam safety[J]. Soft Computing, 2020, 24: 14921-14931.
- [14] CHEN Yaqin, LIN Ling, LI Gang, et al. Accuracy of PCA-NN in non-invasively determining tissue optical properties from spatially resolved diffuse reflectance[C]// Proceedings of SPIE—The International Society for Optical Engineering. Bellingham, Wash., USA: SPIE, 2004.
- [15] ZHANG Ruicheng, PEI ran. Fault diagnosis of nonlinear industrial process based on nuclear independent element analysis[J]. Science, Technology and Engineering, 2020, 20(17): 6944-6949(in Chinese). [张瑞成, 裴 然. 基于核独立元分析的非线性工业过程故障诊断[J]. 科学技术与工程, 2020, 20(17): 6944-6949.]
- [16] SUN D D, GONG X F, CHEN Y L. Integrating canonical variate analysis and kernel independent component analysis for tennessee eastman process monitoring[J]. Journal of Chemical Engineers of Japan, 2020, 53(3): 126-133.
- [17] LIU Yan, WANG Fuli, CHANG Yuqing, et al. Perform-

- ance-relevant kernel independent component analysis based operating performance assessment for nonlinear and non-Gaussian industrial processes[J]. Chemical Engineering Science, 2019, 209: 115167. DOI: 10.1016/j.ces.2019.115167.
- [18] GAO Huihui, XU Yuan, ZHU Qunxiong. Visualization monitoring technology and application of process industry alarm system[J]. Journal of Chemical Industry, 2015, 66(1): 215-221. [高慧慧, 徐圆, 朱群雄. 过程工业报警系统可视化监控技术及应用[J]. 化工学报, 2015, 66(1): 215-221.]
- [19] XU Yongshun, PAN Wei. A brief analysis of information visualization[J]. Modern Industrial Economy and Information Technology, 2016, 6(10): 93-94(in Chinese). [徐永顺, 潘伟. 简析信息可视化[J]. 现代工业经济和信息化, 2016, 6(10): 93-94.]
- [20] LIU Hongyu. Research on information visualization technology and application[J]. Software, 2019, 40(11): 156-157, 181(in Chinese). [刘鸿宇. 信息可视化技术及应用研究[J]. 软件, 2019, 40(11): 156-157, 181.]
- [21] LU Wenbo, JIANG Weikang. Fault diagnosis method and system implementation based on near field acoustic holography[J]. Vibration and Shock, 2013, 32(5): 48-51(in Chinese). [鲁文波, 蒋伟康. 基于近场声全息故障诊断方法及系统实现[J]. 振动与冲击, 2013, 32(5): 48-51.]
- [22] LU Wenbo, Jiang Weikang, Wu Haijun, et al. A fault diagnosis scheme of rolling element bearing based on near-field acoustic holography and gray level co-occurrence matrix[J]. Journal of Sound and Vibration, 2012, 331(15): 3663-3674.
- [23] DU Dangdang, JIA Xiaoliang, WANG Zhuo. Fault pattern recognition of aero engine based on fault color picture[J]. Computer Integrated Manufacturing Systems, 2012, 24(9): 2297-2305(in Chinese). [杜党党, 贾晓亮, 王卓. 基于故障彩色图谱的航空发动机故障模式识别[J]. 计算机集成制造系统, 2012, 24(9): 2297-2305.]
- [24] BING Shaoqiang. Research and application of intelligent early warning system for oil recovery production based on knowledge map[J]. Petrochemical Technology, 2020, 27(1): 30-32(in Chinese). [邵绍强. 基于知识图谱的采油生产智能预警系统研究与应用[J]. 石化技术, 2020, 27(1): 30-32.]
- [25] LIU Xin, JIA Yunxian, SU Xiaobo, et al. Feature extraction of diesel engine misfire based on gray-scale image texture analysis[J]. Vibration and Impact, 2019, 38(2): 140-145(in Chinese). [刘鑫, 贾云献, 苏小波, 等. 基于灰度图像纹理分析的柴油机失火故障特征提取[J]. 振动与冲击, 2019, 38(2): 140-145.]
- [26] SUN Kai, GAO Jianmin, GAO Zhiyong, et al. Plant-wide quantitative assessment of a process industry system's operating state based on color-spectrum[J]. Mechanical Systems and Signal Processing, 2015, 60-61: 644-655.

作者简介:

- 陈倩(1996—),女,陕西西安人,硕士研究生,研究方向:机电系统测试、诊断与控制,E-mail:chenqian_0510@163.com;
- 孙锴(1977—),女,北京人,中国科学院微电子研究所副研究员、西安建筑科技大学硕士生导师,研究方向:机电系统测试、诊断与控制、复杂产品可靠性质量、安全与可靠性工程,通讯作者,E-mail:clara_sun96@163.com。