

基于局部 – 整体相关特征的多单元化工过程分层监测

姜庆超^{1,2} 颜学峰^{1,2}

摘 要 针对一类多单元化工过程的监测问题, 提出基于局部-整体相关特征的分层故障检测与故障定位方法, 通过表征单元内部变量相关性、单元与单元间相关性、局部单元与过程整体相关性, 对过程运行状态进行判断, 以提升过程监测的准确性与可靠性. 首先, 采用典型相关分析, 通过引入邻域单元相关变量提取每个单元的独有特征和外部相关特征; 其次, 对每个单元的独有特征和所有单元的外部相关特征建立统计模型实现分层故障检测; 然后, 建立单元-变量分层贡献图, 对故障单元以及故障变量实现分层定位. 通过在 Tennessee Eastman 仿真过程和一个实验室级甘油精馏过程中的应用说明所提分层监测方法的有效性.

关键词 局部-整体相关特征, 多单元化工过程, 分层过程监测, 故障检测, 故障定位

引用格式 姜庆超, 颜学峰. 基于局部 – 整体相关特征的多单元化工过程分层监测. 自动化学报, 2020, 46(9): 1770–1782

DOI 10.16383/j.aas.c190671

Hierarchical Monitoring for Multi-unit Chemical Processes Based on Local-global Correlation Features

JIANG Qing-Chao^{1,2} YAN Xue-Feng^{1,2}

Abstract A hierarchical process monitoring method based on local-global correlation features is proposed for a class of multi-unit chemical processes. The process operation status is identified by characterizing the correlation within a local unit, between units, and between the local unit and the whole process, through which the monitoring reliability is enhanced. First, based on canonical correlation analysis, individual characteristics and external correlation characteristics of each unit are extracted by introducing correlated variables from neighboring units; Second, multivariate statistical monitoring models are established for individual characteristics of each unit and external correlation characteristics of all units; Then, the unit-variable hierarchical contribution plot is established to locate the fault units and fault variables hierarchically. The effectiveness of the proposed hierarchical monitoring is demonstrated through applications to the Tennessee Eastman process and a laboratory distillation process.

Key words Local-global correlation feature, multi-unit chemical process, hierarchical process monitoring, fault detection, fault location

Citation Jiang Qing-Chao, Yan Xue-Feng. Hierarchical monitoring for multi-unit chemical processes based on local-global correlation features. *Acta Automatica Sinica*, 2020, 46(9): 1770–1782

过程监测作为流程工业过程综合自动化的重要组成部分, 对于保障过程安全、平稳、高效运行与保证产品品质具有重要意义, 是过程控制与过程系统工程领域的研究热点之一^[1-3]. 同时, 伴随着工业数

据采集、传输、存储设备的发展, 过程信息以数据形式得以保存^[4-5]. 基于数据驱动的过程监测不依赖过程机理知识, 在复杂过程优化控制中有着巨大应用前景^[6-10].

由于产品从原料到成品生产工序的日益复杂, 现代流程工业生产过程通常具有规模大、单元多、流程长等特点^[11-14]. 对这类大规模多单元过程的监测, 不仅需要关注过程的整体运行状态, 也需要关注局部单元状态以及单元间的彼此关联影响^[7, 15]. 例如, 水合法制乙二醇生产过程主要包括乙烯氧化制环氧乙烷过程单元、环氧乙烷吸收和解吸过程单元、环氧乙烷精制过程单元、环氧乙烷水合反应过程单元, 以及乙二醇分离与精制过程单元等^[16]. 各个单元既具有独立结构或功能, 又彼此关联, 相互影响, 协同运行. 一个局部的微小故障可能引发多

收稿日期 2019-09-23 录用日期 2020-01-17
Manuscript received September 23, 2019; accepted January 17, 2020

国家自然科学基金 (61973119, 61603138, 21878081) 资助
Supported by National Natural Science Foundation of China (61973119, 61603138, 21878081)

本文责任编辑 王卓
Recommended by Associate Editor WANG Zhuo
1. 华东理工大学化工过程先进控制与优化技术教育部重点实验室 上海 200237 2. 同济大学上海智能科学与技术研究院 上海 200092

1. Key Laboratory of Advanced Control and Optimization for Chemical Processes of Ministry of Education, East China University of Science and Technology, Shanghai 200237 2. Shanghai Institute of Intelligent Science and Technology, Tongji University, Shanghai 200092

个装置或单元的连锁反应, 导致整个过程运行状态恶化, 甚至造成严重事故^[17-18]. 因此, 对多单元过程进行精细化监测, 关注重要单元及过程整体运行状态, 及时发现过程局部或整体异常, 对异常报警原因进行分析识别, 具有重要的理论意义和工程应用价值.

多元统计过程监测 (Multivariate statistical process monitoring, MSPM) 方法, 如主成分分析 (Principal components analysis, PCA)、偏最小二乘分析 (Partial least squares, PLS) 以及典型相关分析 (Canonical correlation analysis, CCA) 等, 在工业过程生产中取得了广泛应用^[10, 19-22]. 直接采用这些经典 MSPM 方法对大规模多单元过程进行监测存在以下几个方面的问题^[8, 15, 23]: 首先, 大规模多单元过程监测变量众多, 造成监测模型复杂度高, 解释性差; 其次, 建立单一整体监测模型往往会忽略过程局部信息, 降低对局部故障的监测性能; 第三, 整体单一监测模型不利于局部故障的定位及诊断; 第四, 对每个单元单独建模往往会忽略单元之间、单元与过程整体的相关关系.

针对大规模多单元工业过程, 文献提出的监测方法主要有以下四种: 多块监测方法^[24-25]、多层监测方法^[26]、联合-独立特征监测方法^[27-28]、以及分布式监测方法^[15, 23, 29]. 分布式监测方法保持变量间关系, 通过局部单元间信息交流增强故障检测性能, 取得了显著成效与快速发展. Ge 等提出基于 PCA 载荷分块的分布式监测方法^[23]; Jiang 等从几何意义和统计意义出发分析了分布式过程监测的必要性与可行性, 建立了分布式过程监测的设计框架^[15, 30]. 最近, Jiang 等提出了基于分布式 CCA 的过程监测方法, 并在多元统计假设检验的框架下分析了分布式 CCA 过程监测对于局部故障检测的优越性^[27]. 然而, 现有分布式 CCA 监测方法有以下问题需要进一步探讨: 首先, 现有分布式 CCA 过程监测方法未对过程的局部单元特征以及过程整体特征进行区分; 其次, 当多个单元受故障影响产生协同变化时, 经典 CCA 产生的残差对于故障检测并不是最优的; 最后, 现有分布式监测方法尚未对过程故障变量的定位问题进行深入探讨, 不利于对过程及故障状态进行进一步解析.

此外, 贡献图方法广泛应用于故障变量定位^[31-33]. 然而, 经典的贡献图方法被指出具有“涂抹 (Smearing)”效应影响, 即单个变量引起的故障会造成多个相关变量的贡献率升高, 进而影响故障定位的效果^[33-34]. 尤其在测量变量众多、变量相关关系复杂的大规模多单元过程中, 变量间以及单元间的相互影响会使得“Smearing”效应更为明显, 为实现故障变量的准确定位带来困难.

本文在之前工作的基础上, 提出基于局部-整体相关特征的多单元化工过程分层监测方法. 该分层监测方法既关注过程的整体运行状态, 也关注局部单元状态以及单元间的彼此关联影响. 首先, 采用 CCA 方法, 通过引入邻域单元相关变量提取局部单元的局部特征和外部相关特征, 建立单元内、局部单元与邻域单元、局部单元与过程整体故障检测模型, 实现对过程局部-整体运行状态更为清晰的描述; 其次, 构建单元-变量分层贡献图, 描述单元对故障统计量、变量对故障统计量的贡献, 减少不同单元变量之间的“Smearing”效应影响, 实现故障变量的准确定位.

1 基于 CCA 的过程故障检测

CCA 是一种经典的多变量统计分析方法. 对于两组随机变量 $\mathbf{u} \in \mathbf{R}^{p \times 1}$ 和 $\mathbf{y} \in \mathbf{R}^{q \times 1}$ (p 和 q 为变量个数), 典型相关分析寻找规范向量 \mathbf{J} 和 \mathbf{L} , 使得经过线性变换之后的典型相关特征 $\mathbf{J}^T \mathbf{u}$ 和 $\mathbf{L}^T \mathbf{y}$ 之间的相关性最大. 该优化问题可以表示为^[19, 27]

$$(\mathbf{J}^*, \mathbf{L}^*) = \arg \max_{(\mathbf{J}, \mathbf{L})} \frac{\mathbf{J}^T \boldsymbol{\Sigma}_{uy} \mathbf{L}}{(\mathbf{J}^T \boldsymbol{\Sigma}_u \mathbf{J})^{\frac{1}{2}} (\mathbf{L}^T \boldsymbol{\Sigma}_y \mathbf{L})^{\frac{1}{2}}} \quad (1)$$

其中, $\boldsymbol{\Sigma}_u$, $\boldsymbol{\Sigma}_y$ 以及 $\boldsymbol{\Sigma}_{uy}$ 表示数据方差和协方差矩阵. 该优化问题的解可以通过构造矩阵 \mathbf{G} 并对其进行奇异值分解得到, 即

$$\mathbf{G} = \boldsymbol{\Sigma}_u^{-\frac{1}{2}} \boldsymbol{\Sigma}_{uy} \boldsymbol{\Sigma}_y^{-\frac{1}{2}} = \mathbf{R} \boldsymbol{\Sigma} \mathbf{V}^T \quad (2)$$

其中, $\boldsymbol{\Sigma} = \begin{bmatrix} \text{diag}\{\rho_1, \dots, \rho_l\} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \in \mathbf{R}^{p \times q}$ 为相关关系矩阵, ρ_i 为线性相关系数, $l = \text{rank}(\boldsymbol{\Sigma}_{uy})$. 两组变量的线性相关程度可以用 $\Upsilon = \sum_{i=1}^l \rho_i$ 来衡量. 两组规范向量可以通过式 (3) 和式 (4) 获得

$$\mathbf{J} = \boldsymbol{\Sigma}_u^{-\frac{1}{2}} \mathbf{R} \in \mathbf{R}^{p \times p} \quad (3)$$

$$\mathbf{L} = \boldsymbol{\Sigma}_y^{-\frac{1}{2}} \mathbf{V} \in \mathbf{R}^{q \times q} \quad (4)$$

通过累积相关系数贡献率, 即 $Cum(k) = (\sum_{i=1}^k \rho_i) / (\sum_{i=1}^l \rho_i) \geq 90\%$, 可以定义最相关的前 k 个典型相关特征. 将 CCA 用于过程故障检测, 需要生成以下残差向量

$$\mathbf{r}_u = \mathbf{J}^T \mathbf{u} - \boldsymbol{\Sigma} \mathbf{L}^T \mathbf{y} \quad (5)$$

$$\mathbf{r}_y = \mathbf{L}^T \mathbf{y} - \boldsymbol{\Sigma}^T \mathbf{J}^T \mathbf{u} \quad (6)$$

已有研究表明, 这些残差具有最小的方差, 对于故障检测是最优的^[19, 27]. 以下两个 T^2 统计量可以用于判别过程运行状态

$$T_u^2 = \mathbf{r}_u^T \boldsymbol{\Sigma}_{ru}^{-1} \mathbf{r}_u \quad (7)$$

$$T_y^2 = \mathbf{r}_y^T \boldsymbol{\Sigma}_{ry}^{-1} \mathbf{r}_y \quad (8)$$

其中, $\boldsymbol{\Sigma}_{ru}$ 和 $\boldsymbol{\Sigma}_{ry}$ 为残差 \mathbf{r}_u 和 \mathbf{r}_y 的协方差矩阵 (具体计算见附录 A)^[19, 27]. T_u^2 将变量组 \mathbf{u} 与变量组 \mathbf{y} 的相关性考虑进来, 用以对变量组 \mathbf{u} 的状态进行监测; T_y^2 将变量组 \mathbf{u} 与变量组 \mathbf{y} 的相关性考虑进来, 用以对变量组 \mathbf{y} 的状态进行监测. 已有文献指出, 当故障只影响变量组 \mathbf{u} 或变量组 \mathbf{y} 时, T_u^2 和 T_y^2 具有最优的监测性能^[19, 27, 35]. 当假设过程数据服从高斯分布时, 统计量的控制限可以通过 χ^2 或 F 分布来获得, 这里不再赘述.

鉴于 CCA 在过程故障检测方面的优越性, 众多 CCA 拓展形式被提出来, 以处理诸如非线性、非高斯性、动态性、多模态等复杂过程特性^[8, 35-38]. 然而, 将 CCA 应用于多单元化工过程监测中, 还有以下几个方面需要进一步探讨:

1) 经典 CCA 方法基于残差构建统计量对过程进行监测, 对于一个只影响变量组 \mathbf{u} 或变量组 \mathbf{y} 的故障, 具有最优的监测性能. 然而, 当故障对变量组 \mathbf{u} 和变量组 \mathbf{y} 造成共同影响时, 故障对 CCA 残差产生的影响可能会被抵消或削弱, 造成监测性能下降.

2) 厂级化工过程变量众多, 单元内变量可能维数高、相关性强. 经典 CCA 方法在计算规范向量时, 如式 (2) 所示, 需要计算每个协方差矩阵的逆, 这使得经典 CCA 方法处理单元内高维相关变量时存在困难.

3) 为实现对过程状态的准确描述, 既要对整个运行状态进行判别, 也要对局部过程运行状态进行判别. 对大规模、多单元厂级化工过程局部-整体特征的分析还需要进一步研究.

4) 现有的基于 CCA 的过程监测方法研究大多集中于故障检测, 对故障定位的研究尚不充分. 尤其是, 在大规模、多单元厂级化工过程监测中, 单元整体对故障的贡献尚未得到充分探讨.

2 基于局部-整体相关特征的多单元化工过程分层监测

基于局部-整体相关特征的多单元过程分层监测方法包括局部单元监测器设计、过程全局监测器设计、以及分层故障定位三部分, 具体介绍如下.

2.1 局部单元监测器设计

对于多单元厂级化工过程, 需要建立局部单元监测模型, 对局部单元的运行状态进行监测与分析. 假设厂级化工过程共有 B 个典型操作单元. 对于第 b , $b = 1, 2, \dots, B$, 个单元, 单元内变量组记为 \mathbf{u}_b ,

对应的历史数据记为 $\mathbf{U}_b \in \mathbf{R}^{m_b \times N}$, 其中, m_b 为第 b 个单元中变量个数, N 为样本个数. 对变量组 \mathbf{u}_b 的协方差矩阵 $\boldsymbol{\Sigma}_u$ 进行奇异值分解可得

$$\boldsymbol{\Sigma}_u = \frac{\mathbf{U}_b \mathbf{U}_b^T}{N-1} = \begin{bmatrix} \mathbf{P}_{U_{pc}} & \mathbf{P}_{U_{res}} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Lambda}_{U_{pc}} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Lambda}_{U_{res}} \end{bmatrix} \begin{bmatrix} \mathbf{P}_{U_{pc}}^T \\ \mathbf{P}_{U_{res}}^T \end{bmatrix} \quad (9)$$

其中, 投影矩阵 $\mathbf{P}_{U_{pc}} \in \mathbf{R}^{m_b \times k_{b,u}}$ 将 \mathbf{u}_b 投影到 $k_{b,u}$ 维主元空间, $\boldsymbol{\Lambda}_{U_{pc}}$ 为包含特征向量 $\lambda_1, \dots, \lambda_{k_{b,u}}$ 的对角矩阵, $k_{b,u}$ 为保留的主元个数. 主元得分 \mathbf{t}_u 可以计算为

$$\mathbf{t}_u = \boldsymbol{\Lambda}_{U_{pc}}^{-\frac{1}{2}} \mathbf{P}_{U_{pc}}^T \mathbf{u}_b \in \mathbf{R}^{k_{b,u} \times 1} \quad (10)$$

相应可得与主元空间互补的残差 \mathbf{e}_u 为

$$\mathbf{e}_u = \mathbf{u}_b - \mathbf{P}_{U_{pc}} \mathbf{P}_{U_{pc}}^T \mathbf{u}_b \quad (11)$$

因残差对应的协方差矩阵求逆可能存在困难, 可构建 Q 统计量对残差空间的变异进行监测, 即

$$Q_b = \mathbf{e}_u^T \mathbf{e}_u = \mathbf{u}_b^T (\mathbf{I}_{m_b \times m_b} - \mathbf{P}_{U_{pc}} \mathbf{P}_{U_{pc}}^T) \mathbf{u}_b \quad (12)$$

其中, $\mathbf{I}_{m_b \times m_b}$ 为 m_b 阶单位矩阵. 记该单元的邻域单元变量组或者第 b 个单元以外的其他所有变量组为 $\mathbf{y}_b \in \mathbf{R}^{m_{\bar{b}} \times 1}$, $m_{\bar{b}}$ 为变量个数. 第 b 个单元的邻域变量组是指和第 b 个单元中变量存在相关关系的变量组. 因为 \mathbf{y}_b 中变量也可能存在较强相关性, 对变量组 \mathbf{y}_b 的协方差矩阵 $\boldsymbol{\Sigma}_y$ 进行同样处理, 可得

$$\boldsymbol{\Sigma}_y = \frac{\mathbf{Y}_b \mathbf{Y}_b^T}{N-1} = \begin{bmatrix} \mathbf{P}_{Y_{pc}} & \mathbf{P}_{Y_{res}} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Lambda}_{Y_{pc}} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Lambda}_{Y_{res}} \end{bmatrix} \begin{bmatrix} \mathbf{P}_{Y_{pc}}^T \\ \mathbf{P}_{Y_{res}}^T \end{bmatrix} \quad (13)$$

$$\mathbf{t}_y = \boldsymbol{\Lambda}_{Y_{pc}}^{-\frac{1}{2}} \mathbf{P}_{Y_{pc}}^T \mathbf{y}_b \in \mathbf{R}^{k_{b,y} \times 1} \quad (14)$$

其中, $\mathbf{Y}_b \in \mathbf{R}^{m_{\bar{b}} \times N}$ 为 \mathbf{y}_b 对应的历史数据, $k_{b,y}$ 为对 \mathbf{y}_b 进行投影时保留的主元个数. 对历史过程数据 \mathbf{U}_b 和 \mathbf{Y}_b 进行投影, 可得 \mathbf{t}_u 和 \mathbf{t}_y 对应的投影数据 $\mathbf{T}_u \in \mathbf{R}^{k_{b,u} \times N}$ 和 $\mathbf{T}_y \in \mathbf{R}^{k_{b,y} \times N}$. 对 \mathbf{T}_u 和 \mathbf{T}_y 进行典型相关分析, 可以得到典型相关向量 \mathbf{J}_T 和 \mathbf{L}_T , 以及相关系数矩阵 $\boldsymbol{\Xi}_T$. 然后, 对于局部单元变量组 \mathbf{u}_b , 可以生成以下残差向量

$$\mathbf{r}_u = \mathbf{J}_T^T \mathbf{t}_u - \boldsymbol{\Xi}_T \mathbf{L}_T^T \mathbf{t}_y = \mathbf{J}_T^T \boldsymbol{\Lambda}_{U_{pc}}^{-\frac{1}{2}} \mathbf{P}_{U_{pc}}^T \mathbf{u}_b - \boldsymbol{\Xi}_T \mathbf{L}_T^T \boldsymbol{\Lambda}_{Y_{pc}}^{-\frac{1}{2}} \mathbf{P}_{Y_{pc}}^T \mathbf{y}_b \quad (15)$$

经过推导 (详细过程见附录 B), 可以得到

$$\mathbf{r}_u = \mathbf{J}_T^T (\mathbf{t}_u - \boldsymbol{\Sigma}_{T_{uy}} \boldsymbol{\Sigma}_{T_y}^{-1} \mathbf{t}_y) \quad (16)$$

其中, $\boldsymbol{\Sigma}_{T_{uy}}$ 为 \mathbf{T}_u 和 \mathbf{T}_y 的协方差矩阵, $\boldsymbol{\Sigma}_{T_y}$ 为 \mathbf{T}_y

对应的协方差矩阵. 记 $\hat{t}_u = \Sigma_{T_{uy}} \Sigma_{T_y}^{-1} t_y$, 可以看出 \hat{t}_u 为基于 t_y 对 t_u 的最小二乘估计. 因此, 基于典型相关分析得到的残差具有最小的方差. 从最小方差的角度来看, 该残差对于故障检测是最优的^[10, 27, 35]. 为增强对监测结果的解释性和对故障定位的准确性, 进一步对残差进行分解. 基于相关系数矩阵及累计相关系数贡献率, 保留前 $k_{b,out}$ 个典型相关变量, 可以将典型相关向量划分为

$$\mathbf{J}_T = [\mathbf{J}_{T,out}, \mathbf{J}_{T,in}] \quad (17)$$

外部相关典型相关变量 \mathbf{h}_b 可以记为

$$\mathbf{h}_b = \mathbf{J}_{T,out}^T \mathbf{t}_u = \mathbf{J}_{T,out}^T \mathbf{A}_{U_{pc}}^{-\frac{1}{2}} \mathbf{P}_{U_{pc}}^T \mathbf{u}_b \quad (18)$$

残差形式可以写为

$$\mathbf{r}_u = \begin{bmatrix} \mathbf{J}_{T,out}^T \\ \mathbf{J}_{T,in}^T \end{bmatrix} \mathbf{t}_u - \begin{bmatrix} \Xi_{T,out} & \mathbf{0} \\ \mathbf{0} & \Xi_{T,in} \end{bmatrix} \mathbf{L}_T^T \mathbf{t}_y \quad (19)$$

单元外部相关残差可以构建为

$$\begin{aligned} \mathbf{r}_{b,out} &= \mathbf{J}_{T,out}^T \mathbf{t}_u - [\Xi_{T,out} \quad \mathbf{0}] \mathbf{L}_T^T \mathbf{t}_y = \\ & \mathbf{J}_{T,out}^T \mathbf{A}_{U_{pc}}^{-\frac{1}{2}} \mathbf{P}_{U_{pc}}^T \mathbf{u}_b - \\ & [\Xi_{T,out} \quad \mathbf{0}] \mathbf{L}_T^T \mathbf{A}_{Y_{pc}}^{-\frac{1}{2}} \mathbf{P}_{Y_{pc}}^T \mathbf{y}_b \end{aligned} \quad (20)$$

单元内部残差可以计算为

$$\begin{aligned} \mathbf{r}_{b,in} &= \mathbf{J}_{T,in}^T \mathbf{t}_u - [\mathbf{0} \quad \Xi_{T,in}] \mathbf{L}_T^T \mathbf{t}_y = \\ & \mathbf{J}_{T,in}^T \mathbf{A}_{U_{pc}}^{-\frac{1}{2}} \mathbf{P}_{U_{pc}}^T \mathbf{u}_b - \\ & [\mathbf{0} \quad \Xi_{T,in}] \mathbf{L}_T^T \mathbf{A}_{Y_{pc}}^{-\frac{1}{2}} \mathbf{P}_{Y_{pc}}^T \mathbf{y}_b \end{aligned} \quad (21)$$

假设 \mathbf{u}_b 和 \mathbf{y}_b 服从高斯分布并经过均值-方差归一化处理. 基于外部相关残差 $\mathbf{r}_{b,out}$ 和单元内部残差 $\mathbf{r}_{b,in}$, 可以构建外部相关统计量 $T_{b,out}^2$ 以及内部特征统计量 $T_{b,in}^2$ 为

$$T_{b,out}^2 = \mathbf{r}_{b,out}^T \Sigma_{b,out}^{-1} \mathbf{r}_{b,out} \sim \chi^2(k_{b,out}) \quad (22)$$

$$T_{b,in}^2 = \mathbf{r}_{b,in}^T \Sigma_{b,in}^{-1} \mathbf{r}_{b,in} \sim \chi^2(k_{b,in}) \quad (23)$$

其中, $\Sigma_{b,out}$ 和 $\Sigma_{b,in}$ 为 $\mathbf{r}_{b,out}$ 和 $\mathbf{r}_{b,in}$ 的协方差矩阵, $k_{b,in} = k_{b,u} - k_{b,out}$. 关于式 (20)~(23) 的说明详见文献^[19, 27]. 局部监测器设计实现了以下目的:

1) 构建了 $T_{b,out}^2$, $T_{b,in}^2$ 和 Q_b 统计量, 用于对过程局部单元运行状态进行监测. Q_b 统计量描述局部单元内部变量的相关性信息. $T_{b,in}^2$ 对局部单元的独有特征进行监测, 当只有 $T_{b,in}^2$ 检测出故障时, 该故障通常情况下为局部过程故障. $T_{b,out}^2$ 统计量将局部单元与外部变量的相关性考虑进来, 实现局部单元的监测. 当 $T_{b,out}^2$ 检测出故障时, 该故障通常与单元外部其他变量相关.

2) 提取了每个单元的外部相关特征 \mathbf{h}_b , 这些外部相关特征表征了每个单元与外部其他单元的关联信息, 可以用于对过程全局的监测.

2.2 过程全局监测器设计

一个局部单元的外部相关特征包含该单元和其他单元的所有相关信息, 因而可以用来建立全局过程监测模型, 实现对过程全局运行状态的描述. 记外部相关特征组合 $\mathbf{z} = [\mathbf{h}_1^T, \mathbf{h}_2^T, \dots, \mathbf{h}_B^T]^T \in \mathbf{R}^{m_z \times 1}$ (m_z 为 \mathbf{z} 中变量个数), 则基于历史训练数据得到的外部相关特征数据可以记为 $\mathbf{Z} \in \mathbf{R}^{m_z \times N}$. 因为不同单元间的外部相关特征可能存在较强的相关性, 对其直接建立 T^2 统计量用以监测过程整体运行状态存在困难. 因此, 对外部相关特征的协方差矩阵进行奇异值分解, 将外部相关特征投影到两个互补空间进行监测. 对协方差矩阵 Σ_z 进行奇异值分解可以得到

$$\begin{aligned} \Sigma_z &= \frac{\mathbf{Z}\mathbf{Z}^T}{N-1} = \\ & [\mathbf{P}_{Z_{pc}} \quad \mathbf{P}_{Z_{res}}] \begin{bmatrix} \Lambda_{Z_{pc}} & \mathbf{0} \\ \mathbf{0} & \Lambda_{Z_{res}} \end{bmatrix} \begin{bmatrix} \mathbf{P}_{Z_{pc}}^T \\ \mathbf{P}_{Z_{res}}^T \end{bmatrix} \end{aligned} \quad (24)$$

其中, $\Lambda_{Z_{pc}}$ 包含前 k_z 个方差最大主元的方差信息. 构建 T^2 统计量用以监测主元空间可以得到

$$T_z^2 = \mathbf{z}^T \mathbf{P}_{Z_{pc}} \Lambda_{Z_{pc}}^{-1} \mathbf{P}_{Z_{pc}}^T \mathbf{z} \quad (25)$$

残差向量 \mathbf{e}_z 可以计算为

$$\mathbf{e}_z = \mathbf{z} - \mathbf{P}_{Z_{pc}} \mathbf{P}_{Z_{pc}}^T \mathbf{z} \quad (26)$$

构建 Q 统计量以监测残差空间可得

$$Q_z = \mathbf{e}_z^T \mathbf{e}_z = \mathbf{z}^T (\mathbf{I}_{m_z \times m_z} - \mathbf{P}_{Z_{pc}} \mathbf{P}_{Z_{pc}}^T) \mathbf{z} \quad (27)$$

其中, $\mathbf{I}_{m_z \times m_z}$ 为 m_z 阶单位矩阵. 通过全局监测器设计, 构建了 T_z^2 和 Q_z 统计量, 实现了对单元与单元间、局部单元与过程整体间相关关系的描述.

2.3 分层故障定位

对于多单元过程, 分层故障定位首先定位故障发生单元, 其次定位单元内故障相关变量. 在定位多单元过程发生故障的单元时, 针对不同的故障有两种途径: 1) 当全局过程故障发生时, 需要衡量每个单元外部相关特征对故障统计量的贡献; 2) 当局部故障发生时, 找到局部监测统计量超限的单元.

对于过程整体, 有 T_z^2 和 Q_z 两个统计量对过程运行状态进行监测; 对于每个局部单元, 有 $T_{b,out}^2$ 、 $T_{b,in}^2$ 和 Q_b 三个统计量对过程状态进行监测. 当 T_z^2 或 Q_z 超出控制限时, 需要计算每个单元的外部相关特征对统计量的贡献. 某个外部相关特征 z_j 对于 T_z^2 统计量的贡献可计算为^[32-33]

$$CONT(z_j) = \sum_{i=1}^{k_z} (cont^{j,i}) = \sum_{i=1}^{k_z} \left(\frac{t_i}{\lambda_i} P_{Z_{pc}}^{j,i} z_j \right) \quad (28)$$

其中, $p_{Zpc}^{j,i}$ 为投影矩阵 P_{Zpc} 中第 (j, i) 个元素. 某个外部相关特征 z_j 对于 Q_z 统计量的贡献可计算为

$$CONQ(z_j) = (z_j - \hat{z}_j)^2 = \left(z_j - (P_{Zpc} P_{Zpc}^T z) \right)_j^2 \quad (29)$$

其中, $(P_{Zpc} P_{Zpc}^T z)_j$ 表示 $P_{Zpc} P_{Zpc}^T z$ 中第 j 个元素.

单元外部相关特征贡献图描述了单元外部相关特征对统计量 T_z^2 和 Q_z 的影响. 当影响多个单元或过程全局特性的故障发生时, T_z^2 或 Q_z 统计量会超出控制限, 通过外部相关特征贡献图可以找到对统计量贡献大的单元, 实现故障单元的定位.

当找到故障单元后, 需要进一步定位单元内的故障变量. 对于第 b 个过程单元, 第 i 个变量 u_b^i 对 Q_b 统计量的贡献可计算为

$$CONQ_b^i = (u_b^i - \hat{u}_b^i)^2 = \left(u_b^i - (P_{Upc} P_{Upc}^T u_b) \right)_i^2 \quad (30)$$

某个变量对 $T_{b,out}^2$ 或 $T_{b,in}^2$ 统计量的贡献需要先计算所提取主元对残差的贡献. 对于贡献大的主元, 再计算变量对主元的贡献. 记第 b 个单元变量组 u_b 对应的第 j 个主元得分为 t_u^j . 该主元对第 i 个残差 r_u^i 的贡献为

$$CONR_U^{j,i} = (J^{j,i} (t_u^j - \hat{t}_u^j))^2 = \left(J^{j,i} \left(t_u^j - (\Sigma_{Tuy} \Sigma_{T_y}^{-1} t_y) \right)_j \right)^2 \quad (31)$$

则 t_u^j 对 $T_{b,out}^2$ 的贡献可以计算为

$$CONR_{U,out}^j = \sum_{i=1}^{k_{b,out}} \frac{CONR_U^{j,i}}{\Sigma_{r_u}^{i,i}} \quad (32)$$

其中, $\Sigma_{r_u}^{i,i}$ 为残差向量 r_u 对应的协方差矩阵 Σ_{r_u} 的第 (i, i) 个元素. 相应的, t_u^j 对 $T_{b,in}^2$ 的贡献可以计算为

$$CONR_{U,in}^j = \sum_{i=k_{b,out}+1}^{k_{b,u}} \frac{CONR_U^{j,i}}{\Sigma_{r_u}^{i,i}} \quad (33)$$

当 $T_{b,out}^2$ 发生故障时, 需计算变量 u_b^i 对 $T_{b,out}^2$ 中贡献大的主元的贡献. u_b^i 对前 k_1 个对 $T_{b,out}^2$ 贡献大的主元的贡献可以计算为

$$CONU_{out}^i = \sum_{j=1}^{k_1} (cont^{i,j}) = \sum_{j=1}^{k_1} \frac{(p_{Zpc}^{i,j} u_b^i)^2}{\lambda_j} \quad (34)$$

u_b^i 对前 k_2 个对 $T_{b,in}^2$ 贡献大的主元贡献为

$$CONU_{in}^i = \sum_{j=1}^{k_2} (cont^{i,j}) = \sum_{j=1}^{k_2} \frac{(p_{Zpc}^{i,j} u_b^i)^2}{\lambda_j} \quad (35)$$

在变量贡献计算中, 应只计算那些变量对于超出控制限的统计量的贡献, 因此, 引入一个加权值 w 将变量贡献校正为

$$CON = w \times CONTR \quad (36)$$

其中, $CONTR$ 为原始计算得到的变量贡献, $CONTR$ 为 $CONQ_b^i$, $CONU_{out}^i$ 或者 $CONU_{in}^i$. 当对应统计量超出控制限时, $w = 1$; 当对应统计量未超出控制限时, $w = 0$.

基于局部-整体相关特征的分层监测设计框架如图 1 所示. 对于第 b 个操作单元, 首先对单元内变量组 u_b 和邻域变量组 y_b 进行主成分提取, 得到单元残差子空间、单元主元子空间、以及邻域主元子空间. 其次, 基于 CCA 将单元主元子空间分解为

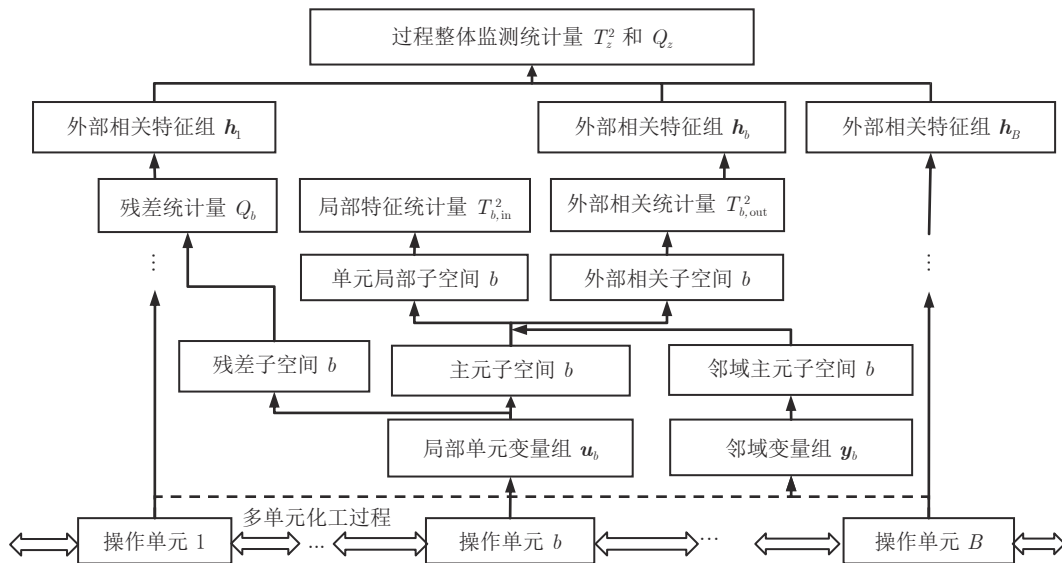


图 1 基于局部-整体相关特征的分层监测设计框架

Fig.1 Framework of the local-global correlation feature-based hierarchical monitoring

单元局部子空间和外部相关子空间. 然后建立残差统计量 Q_b 、局部特征统计量 $T_{b,in}^2$ 、以及外部相关统计量 $T_{b,out}^2$, 实现对局部单元的监测. 最后, 采用所有单元的外部相关特征构建整体监测统计量 T_z^2 和 Q_z , 实现对过程整体运行状态的监测. 当检出故障时, 采用分层定位方法对故障单元及故障变量实现定位.

值得指出的是, 本文工作的出发点是对多单元过程进行精细化监测, 既关注局部单元运行状态, 也关注过程全局运行状态, 及时发现过程局部或全局异常, 对异常报警原因进行分析识别. 分层监测方法对各个单元及过程全局分别构建统计量, 不可避免地增加了统计量个数. 统计量个数的增加一方面增强了故障检测和定位的准确性以及对故障状态的解释性, 另一方面不可避免地增加了整体监测故障误报率. 然而, 分层监测方法是建立在统计分析框架内, 对于每个局部单元或是过程全局监测的故障误报率都可以通过调整统计量控制限中置信度来控制, 以满足实际应用需求.

3 仿真与实验研究

3.1 在 Tennessee Eastman 测试平台的应用

Tennessee Eastman (TE) 过程是通过模拟真实的工业过程构建的^[39-40]. 如图 2 所示^[41], 该过程

由 5 个典型操作单元组成, 分别为进料单元、反应器单元、分离器单元、汽提塔单元和压缩机单元. 相应的变量如表 1 所示^[40-41]. 首先, 基于过程正常运行数据建立分层监测模型. 然后, 对于 21 个故障测试集, 计算故障检出率来检验算法性能, 如表 2 所示.

表 1 TE 过程的典型操作单元和对应变量
Table 1 Operation units and corresponding variables in the TE process

单元	变量描述	变量名称	符号
进料	A 进料 (流1)	XMEAS(1)	①
	D 进料 (流2)	XMEAS(2)	②
	E 进料 (流3)	XMEAS(3)	③
	A 和 C 进料	XMEAS(4)	④
	D 进料	XMV(1)	⑫
	A 进料流量	XMV(3)	⑬
	E 进料流量	XMV(2)	⑭
	A 和 C 进料流量	XMV(4)	⑮
反应器	反应器进料量	XMEAS(6)	⑥
	反应器压力	XMEAS(7)	⑦
	反应器液位	XMEAS(8)	⑧
	反应器温度	XMEAS(9)	⑨
	反应器水温	XMEAS(21)	⑳
	反应器冷却水流量	XMV(10)	⑩
	冷凝器冷却水流量	XMV(11)	㉑
	冷凝器冷却水流量	XMV(11)	㉒
	冷凝器冷却水流量	XMV(11)	㉓
	冷凝器冷却水流量	XMV(11)	㉔
	冷凝器冷却水流量	XMV(11)	㉕

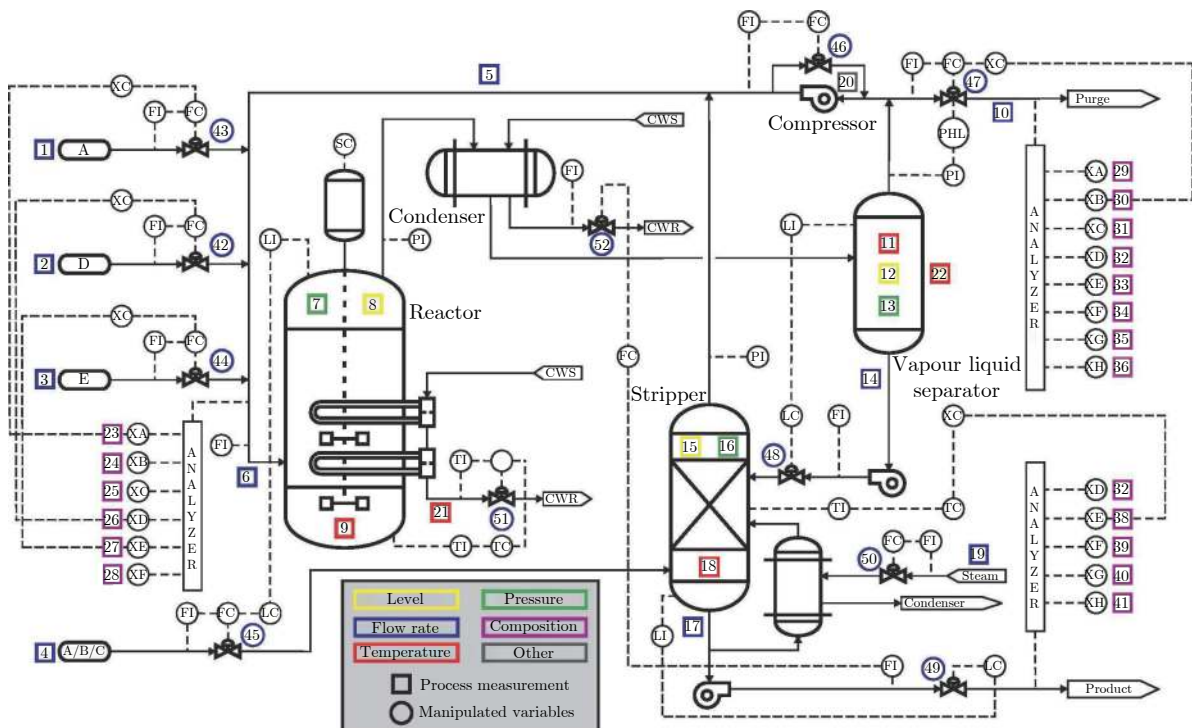


图 2 TE 过程流程图^[41]

Fig. 2 Flowchart of the TE process^[41]

表 1 TE 过程的典型操作单元和对应变量 (续)

Table 1 Operation units and corresponding variables in the TE process (continued)

单元	变量描述	变量名称	符号
分离器	分离器温度	XMEAS(11)	⑪
	分离器液位	XMEAS(12)	⑫
	分离器压力	XMEAS(13)	⑬
	分离器底物流量	XMEAS(14)	⑭
	分离器水温度	XMEAS(22)	⑳
	分离器液流量	XMV(7)	㉓
汽提塔	汽提塔液位	XMEAS(15)	⑮
	汽提塔压力	XMEAS(16)	⑯
	汽提塔底物流量	XMEAS(17)	⑰
	汽提塔温度	XMEAS(18)	⑱
	汽提塔蒸汽流量	XMEAS(19)	㉑
	汽提塔产物流量	XMV(8)	㉒
压缩	汽提塔蒸汽阀开度	XMV(9)	㉔
	再循环流量	XMEAS(5)	⑤
	排放速度	XMEAS(10)	⑩
	压缩机功率	XMEAS(20)	㉒
	压缩机再循环阀	XMV(5)	㉕
	排放阀	XMV(6)	㉖

特别地, 我们详细分析分层监测方法对于故障 4 和故障 5 的监测效果.

故障 4 是影响反应器冷却水入口温度的局部故障, 导致反应器冷却水流量测量值发生阶跃变化 (变量 51, XMV(10)), 其他变量受到轻微影响^[39-40], 图 3(a) 显示经典 CCA 全局监测方法 (即 11 个操作变量作为变量组 u , 22 个测量变量作为变量组 y) 的监测结果; 图 3(b) 和 3(c) 显示分层监测的整体监测结果和局部统计量监测结果. 对比图 3(a) 和 3(b), 可以发现经典 CCA 和分层监测的整体监测效果有显著不同. 图 3(a) 持续指示过程有故障发生, 但并不显示故障性质信息. 从图 3(b) 和 3(c) 中可以看出, 分层监测方法的整体监测统计量很少受到影响, 而反应器对应的第二个操作单元局部监测统计量受到显著影响, 表明该故障是局部故障. 并且, 该故障仅影响反应器单元, 与其他操作单元无关. 分层故障检测结果与实际故障设置情况相符^[40]. 因为全局监测统计量很少超出控制限, 我们直接定位到局部监测器的统计量, 查看检测出故障的 161 到 165 点的平局贡献图, 如图 4 所示.

图 4(a) 给出变量对于各单元外部相关统计量

表 2 分层监测对于 21 个故障测试集的监测效果

Table 2 Hierarchical monitoring results for the 21 faults in TE process

编码	单元及过程 故障描述/统计量	进料单元			反应器单元			分离器单元			汽提塔单元			压缩单元			过程整体	
		$T_{1,out}^2$	$T_{1,in}^2$	Q_1	$T_{2,out}^2$	$T_{2,in}^2$	Q_2	$T_{3,out}^2$	$T_{3,in}^2$	Q_3	$T_{4,out}^2$	$T_{4,in}^2$	Q_4	$T_{5,out}^2$	$T_{5,in}^2$	Q_5	T_z^2	Q_z
1	A/C 进料比率, B 成分不变 (阶跃)	0.99	0.31	0.04	0.77	0.26	0.06	0.44	0.04	0.07	1	0.06	0.98	0.17	0.02	0.23	1	1
2	B 成分, A/C 进料比率不变 (阶跃)	0.92	0.02	0.27	0.95	0.22	0.03	0.92	0.14	0.06	0.99	0.06	0.89	0.99	0.01	0.42	0.98	0.98
3	D 的进料温度 (阶跃)	0.01	0.01	0.01	0.32	0.02	0.00	0.14	0.01	0.00	0.20	0	0.01	0	0.01	0.09	0.01	0.02
4	反应器冷却水入口温度 (阶跃)	0.02	0.01	0.02	0.25	0.75	1	0.12	0.00	0.01	0.21	0	0.01	0	0.00	0.01	0.03	0.06
5	冷凝器冷却水入口温度 (阶跃)	0.16	0.03	0.04	0.99	0.09	0.03	0.23	0.01	0.02	1	0.00	0.19	0.07	0.00	0.13	0.22	0.18
6	A 进料损失 (阶跃)	0.99	0.91	1	1	0.98	0.96	0.98	0.82	0.98	0.99	0.96	0.97	0.99	0.92	0.99	0.99	0.99
7	C 存在压力损失 (阶跃)	0.98	1	0.87	0.98	0.22	0.09	0.38	0.03	0.04	0.76	0.01	0.22	0.24	0.01	0.27	1	0.98
8	A、B、C 进料成分 (随机)	0.78	0.10	0.16	0.97	0.48	0.13	0.90	0.03	0.35	0.94	0.11	0.68	0.87	0.02	0.61	0.97	0.89
9	D 的进料温度 (随机)	0.00	0.01	0.01	0.27	0.02	0.01	0.14	0.01	0.01	0.16	0	0.01	0	0.00	0.02	0.01	0.02
10	C 的进料温度 (随机)	0.08	0.02	0.02	0.43	0.04	0.02	0.33	0.01	0.00	0.46	0.00	0.81	0.07	0.00	0.10	0.29	0.13
11	反应器冷却水入口温度 (随机)	0.11	0.01	0.01	0.39	0.61	0.70	0.17	0.01	0.01	0.42	0.00	0.05	0	0.01	0.02	0.20	0.27
12	冷凝器冷却水入口温度 (随机)	0.74	0.22	0.25	0.95	0.60	0.29	0.94	0.28	0.65	0.96	0.06	0.89	0.34	0.03	0.83	0.96	0.91
13	反应动态 (慢偏移)	0.77	0.19	0.30	0.92	0.72	0.39	0.89	0.10	0.48	0.95	0.24	0.86	0.85	0.03	0.89	0.94	0.95
14	反应器冷却水阀门 (粘滞)	0.75	0.01	0.00	1	0.97	0.12	0.36	0.07	0.01	0.88	0.01	0.01	0.04	0.01	0.01	1	1
15	冷凝器冷却水阀门 (粘滞)	0.01	0.01	0.01	0.29	0.02	0.01	0.18	0.00	0.01	0.23	0	0.03	0.00	0.00	0.03	0.02	0.06
16	未知	0.03	0.02	0.01	0.37	0.03	0.01	0.26	0.00	0.00	0.40	0	0.85	0.03	0.01	0.05	0.15	0.11
17	未知	0.64	0.02	0.01	0.95	0.94	0.44	0.35	0.04	0.02	0.76	0.00	0.22	0.03	0.01	0.03	0.84	0.85
18	未知	0.88	0.82	0.83	0.92	0.87	0.79	0.91	0.12	0.87	0.90	0.75	0.88	0.80	0.71	0.85	0.88	0.88
19	未知	0.01	0.02	0.01	0.24	0.08	0.01	0.14	0.01	0.01	0.16	0	0.12	0.01	0.32	0.66	0.01	0.03
20	未知	0.03	0.02	0.01	0.61	0.02	0.01	0.38	0.05	0.39	0.81	0.00	0.23	0.47	0.01	0.89	0.32	0.43
21	流 4 的阀门固定在稳态位置	0.01	0.00	0.00	0.66	0.44	0.01	0.87	0.00	0.01	0.73	0.00	0.45	0.31	0.00	0.02	0.41	0.84

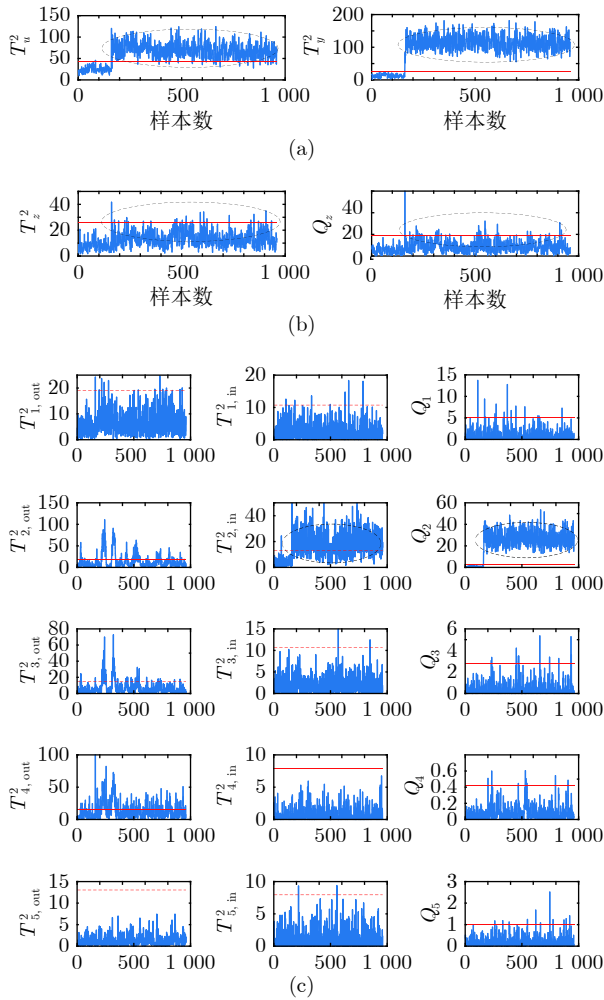


图 3 TE 过程故障 4 的监测效果 ((a) 经典 CCA 监测结果; (b) 分层监测整体监测结果; (c) 分层监测局部监测效果)

Fig.3 Monitoring results for the TE fault 4 ((a) Conventional CCA method; (b) Global monitoring using hierarchical method; (c) Local monitoring using hierarchical method)

的贡献图, 图 4(b) 给出变量对于单元内部特征统计量的贡献图, 图 4(c) 给出了变量对 Q_b 统计量的贡献图. 从图中可以看出, 第二个操作单元, 即反应器单元内的变量对故障统计量贡献最大. 进一步分析可得, 反应器温度和反应器冷却水流量两个变量对于描述单元内部特征状态的 $T_{2,in}^2$ 和 Q_2 贡献最大, 因此这两个变量被识别为与故障相关的变量. 这与故障设置和实际过程运行情况相符.

故障 5 影响冷凝器冷却水进口温度. 在故障开始时, 多个变量受到影响, 故障对各个单元以及过程整体运行状态影响显著. 但是, 由于控制回路的存在, 除冷凝器冷却水流量外, 在第 500 点以后的大多数变量中故障影响得以消除^[40-41]. 图 5(a) 显示了经典 CCA 全局监测方法对于故障 5 的监测结果;

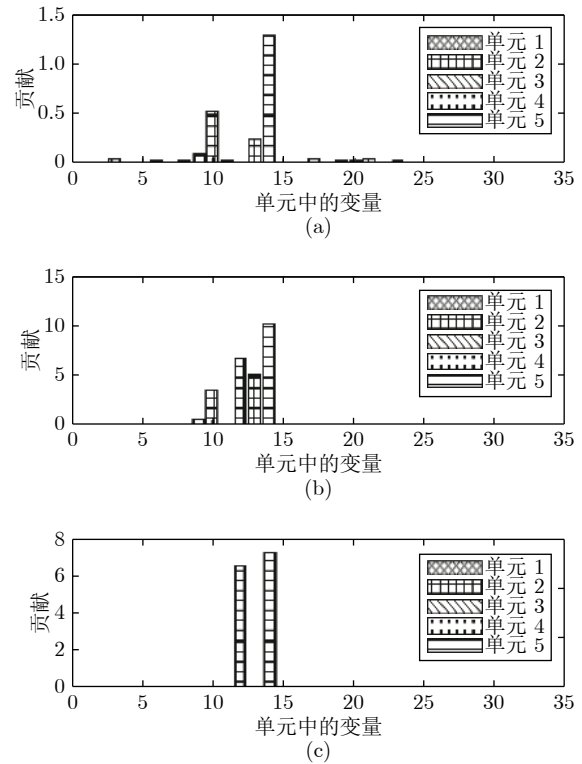


图 4 TE 过程故障 4 的分层贡献图 ((a) $T_{b,out}^2$; (b) $T_{b,in}^2$; (c) Q_b)

Fig.4 Contribution plots for the TE fault 4 ((a) $T_{b,out}^2$; (b) $T_{b,in}^2$; (c) Q_b)

图 5(b) 和 5(c) 提供了分层监测对于故障 5 的整体和局部监测结果. 从图 5(a) 中可以看出, 经典 CCA 全局监测方法持续显示故障, 但并不区分故障性质. 从图 5(b) 和 5(c) 中可以看出, 在故障开始时全局监测统计量检测到故障, 但在 500 点后恢复正常. 第二个单元和第四个单元的外部相关统计量持续指示故障, 而这两个单元的内部监测统计量也在 500 点之后恢复正常. 这表示由于控制回路的补偿, 局部单元的状态恢复正常. 但是, 不同单元之间的相关性发生了变化, 这种变化无法得到补偿.

最后, 故障定位贡献图如图 6 所示. 图 6(a) 显示了 161 点到第 165 点 (故障刚开始时) 的全局监测贡献图. 从图中可以看出, 单元 2~4 中的外部相关特征对于全局监测统计量有较大贡献, 其中反应器单元和分离器单元中的变量对 T_z^2 的贡献最大. 图 6(b)~(d) 显示了故障刚开始时变量对局部监测统计量 $T_{b,out}^2$ 、 $T_{b,in}^2$ 以及 Q_b 的贡献图. 从图中可以看出, 单元 3 中的变量对统计量有最大的贡献. 图 6(e) 显示了第 701 至 705 点变量对外部相关监测统计量的平均贡献图 (控制回路补偿之后). 可见, 反应器中的变量贡献率相对较高, 其中冷凝器冷却水流量

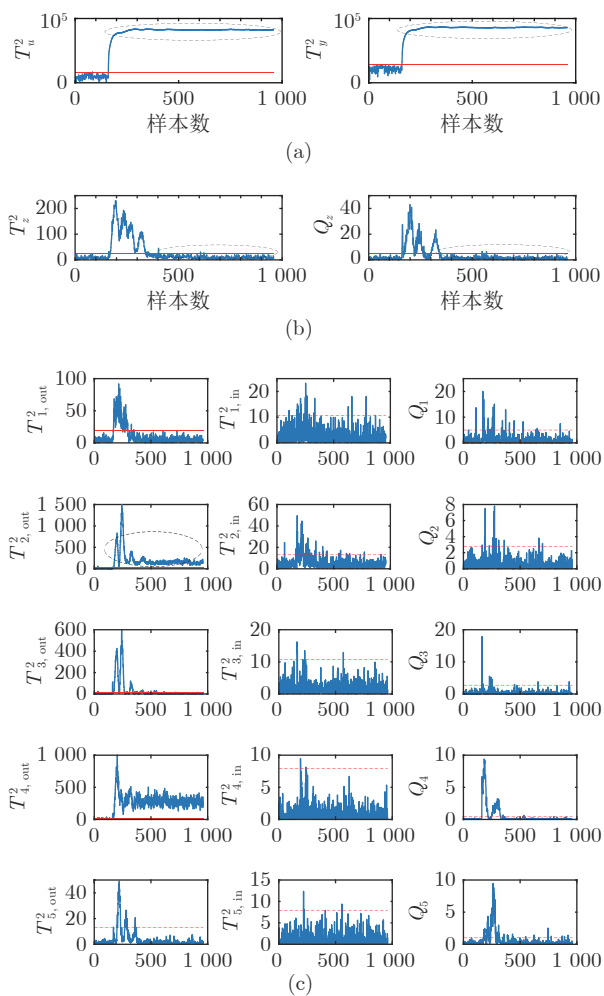


图 5 TE 过程故障 5 的监测效果 ((a) 经典 CCA 监测结果; (b) 分层监测整体监测结果; (c) 分层监测局部监测效果)

Fig.5 Fault detection results for the TE fault 5 ((a) Conventional CCA method; (b) Global monitoring using hierarchical method; (c) Local monitoring using hierarchical method)

的贡献最大, 故障定位结果和实际设置相符^[41]. 以上结果表明分层监测方法准确实现了故障检测与故障定位.

3.2 在实验室精馏过程中的应用

甘油是一种用途广泛的化工原料, 甘油蒸馏装置直接影响甘油产品质量, 在甘油生产中占有重要地位. 按照与实际过程 8 : 1 的比例, 实验室建立了甘油蒸馏过程的实验装置, 如图 7 所示. 过程进料液是水和甘油的混合物. 精馏塔设备的下端与蒸发器相连. 热油经导热油循环泵送至精馏塔和预热罐. 热油加热系统由循环泵强制循环, 热油输送至供热器以满足过程的热量需求, 然后返回加热炉对下面的循环泵进行加热. 液相导热油加热系统的工作压

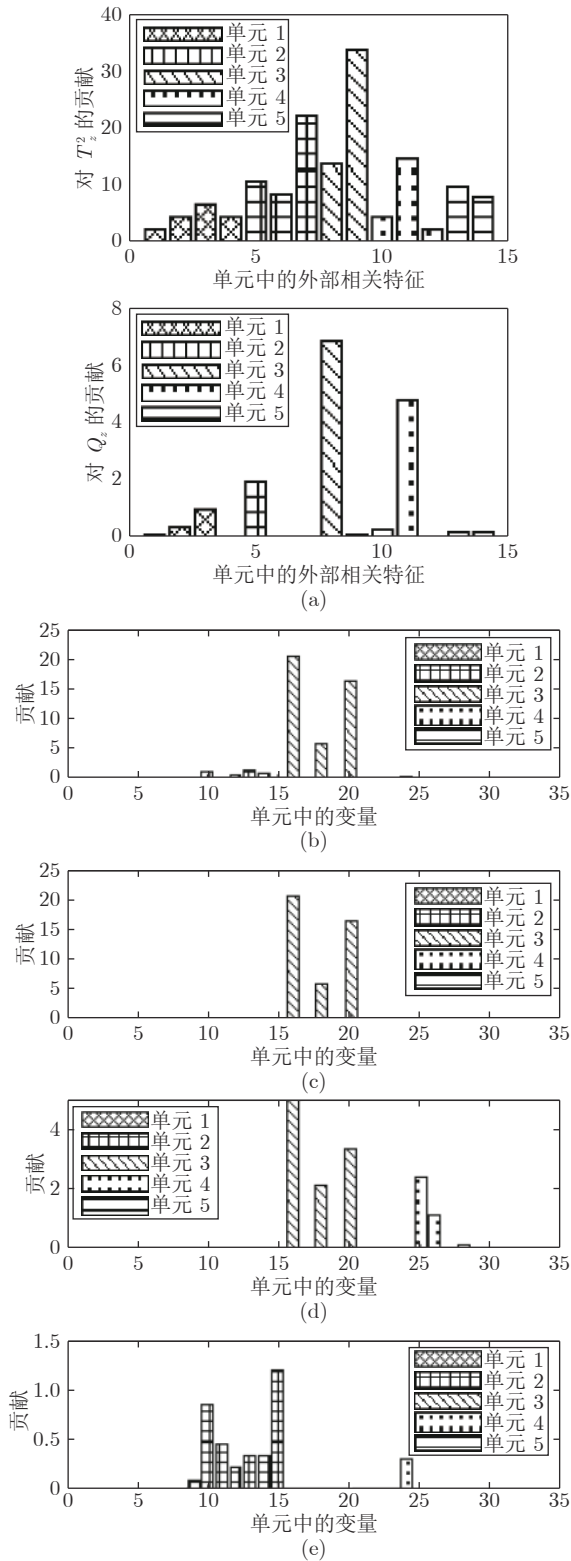


图 6 TE 过程故障 5 的分层监测贡献图 ((a) T_z^2 和 Q_z ; (b) $T_{b,out}^2$; (c) $T_{b,in}^2$; (d) Q_b ; (e) 控制补偿后 $T_{b,out}^2$)

Fig.6 Contribution plots for the TE fault 5 ((a) T_z^2 and Q_z ; (b) $T_{b,out}^2$; (c) $T_{b,in}^2$; (d) Q_b ; (e) $T_{b,out}^2$ after compensation)

力也由循环泵决定. 蒸发器加热液体产生蒸汽, 蒸

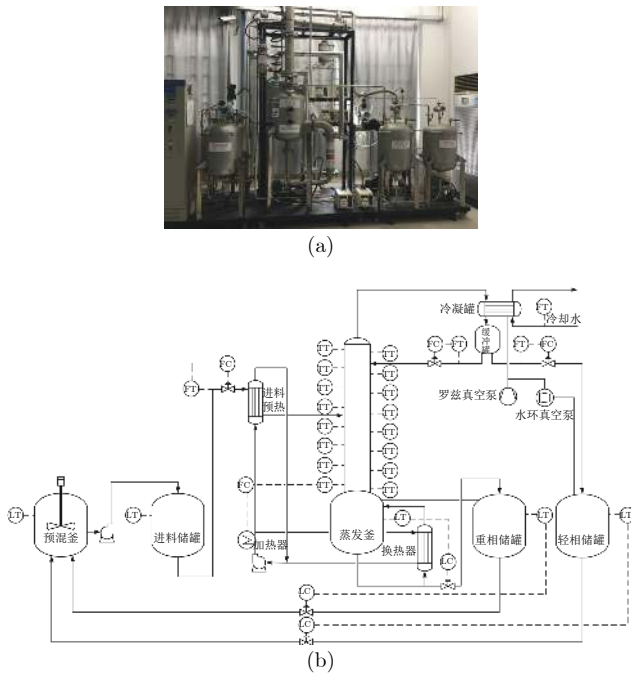


图 7 实验室甘油精馏装置与流程图
(a) 设备图; (b) 流程图

Fig. 7 Lab-scale glycerin distillation process ((a) Equipment diagram; (b) Simplified flowchart)

汽沿着塔板上升, 并与下降的液体逆流接触, 重组分流入蒸发器。塔顶设有冷凝器, 通过管道与塔体相连。顶部馏出物通过冷凝器进入缓冲罐, 一些馏出物作为轻组分流入轻组分罐, 其余的则回流到蒸馏塔进行蒸馏和提纯。我们将整个过程的监测变量分为两个单元, 各单元变量如表 3 所示。在过程正常运行情况下, 采集 500 个样本作为训练数据。为了测试分层监测的性能, 在第 130 个样本之后, 引入了以下两个故障: 1) 进料流量阶跃变化; 2) 温度传感器 4 缓慢漂移。

故障 1 是一个会对两个单元运行状态都造成影响故障。分层监测对于故障 1 的故障检测效果如图 8 所示。图 8(a) 显示全局监测统计量监测效果, 图 8(b) 显示局部单元监测统计量监测效果。从图 8 中可以看出, 两个局部单元以及过程全局特性均受

表 3 甘油精馏过程中的监测变量

Table 3 Measured variables in the distillation process

单元 1	变量名称	单元 2	变量名称
1	进料流量	1	进料储罐液位
2	灵敏板温度	2~13	塔板温度 1~12
3	塔底液位	14	冷却水流量
4	塔顶回流	15	重相储罐液位
5	塔顶产品流	16	轻相储罐液位

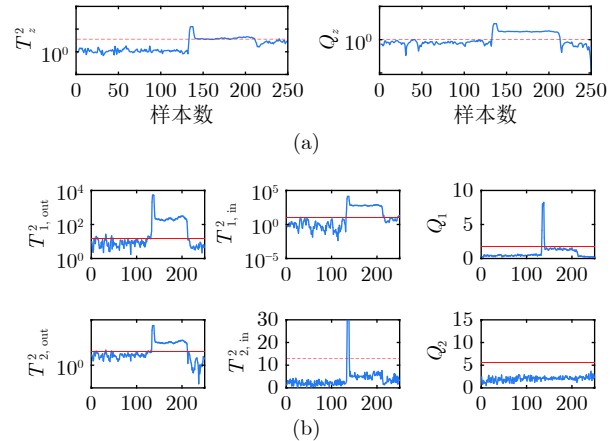


图 8 精馏过程故障 1 的分层故障检测效果 ((a) 整体监测统计量; (b) 局部监测统计量)

Fig. 8 Fault detection results for the distillation process fault 1 ((a) Global monitoring statistics; (b) Local monitoring statistics)

到故障影响。过程整体和局部监测贡献图如图 9 所示。图 9(a) 显示两个单元的外部相关特征均对过程全局故障有贡献。变量对各个单元 $T_{b,out}^2$ 贡献图如图 9(b) 所示, 变量对各单元 $T_{b,in}^2$ 贡献图如图 9(c) 所示。从图中可以看出, 单元 1 内变量对故障贡献最大, 故障被定位在单元 1 内。在单元 1 内, 进料量对故障贡献最大, 这与实际情况相符。

故障 2 是一个传感器故障, 是发生在单元 2 内的一个局部故障, 没有对过程全局造成影响。分层监测对于故障 2 的监测效果如图 10 所示。图 10(a) 显示全局监测统计量监测效果; 图 10(b) 显示局部监测统计量对局部单元的监测效果。从图 10 中可以看出, 故障最先被 $T_{2,in}^2$ 检出, 指示该故障为影响单元 2 局部特征的一个局部故障。随着故障幅度的逐渐增大, 其他统计量开始受到影响。变量贡献图如图 11 所示, 从图 11 中可以看出, 第二个单元的变量对 $T_{2,in}^2$ 有较大贡献, 尤其是温度传感器 4 对故障贡献最大, 这与实际情况相符。以上实验与监测结果表明了所提分层监测方法的有效性。

4 总结与展望

本文提出一种基于局部-整体相关特征的分层过程监测方法。首先, 通过引入每个单元的邻域变量, 提取每个局部单元的外部相关特征与局部特征, 实现对局部单元的外部相关特征与局部特征的监测; 其次, 通过每个单元的外部相关特征构建全局监测模型, 实现对过程整体运行状态的监测; 最后, 构建单元-变量贡献图, 实现从故障单元到单元内故障变量的准确定位。通过在 TE 标准测试数据和一个实验室精馏过程的应用验证了所提分层监测方

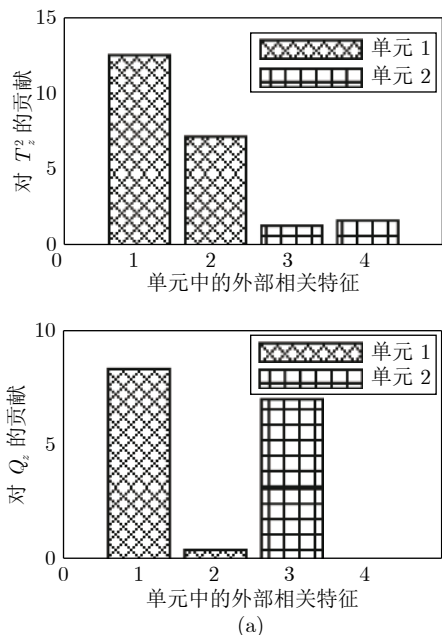


图 9 精馏过程故障 1 分层监测贡献图
(a) T_z^2 和 Q_z ; (b) $T_{b,out}^2$; (c) $T_{b,in}^2$

Fig.9 Contribution plots for the distillation fault 1
(a) T_z^2 and Q_z ; (b) $T_{b,out}^2$; (c) $T_{b,in}^2$

法的有效性. 分层监测方法考虑单元内变量相关性、单元间相关性、以及单元与过程整体相关性. 相比于经典分块监测方法, 所提分层监测方法既识别故障状态, 又识别故障位置, 在处理多单元过程监测中显示出优秀监测性能.

本文提出的分层监测方法面向线性高斯过程, 但是在该设计框架内, 可以融合动态、非线性、多模态等监测方法, 实现具有复杂特性的多单元化工过程的分层监测. 同时, 近年来基于稀疏回归变量选择实现故障变量定位的方法也受到越来越多的关注^[13, 34], 如何在分层监测中融合变量选择来

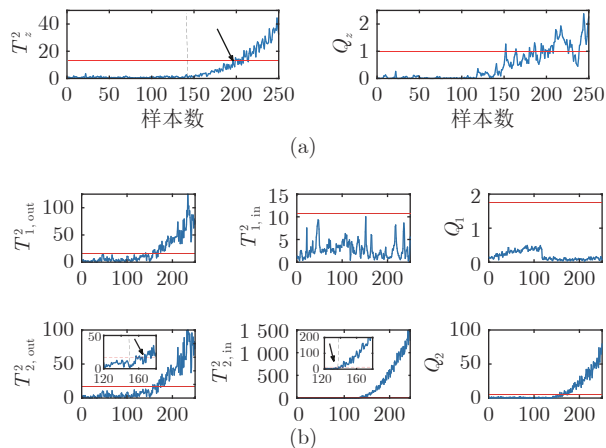


图 10 精馏过程故障 2 的分层故障检测效果
(a) 整体监测统计量; (b) 局部监测统计量

Fig.10 Fault detection results for the distillation process fault 2 ((a) Global monitoring statistics; (b) Local monitoring statistics)

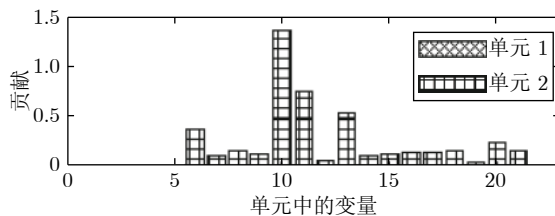
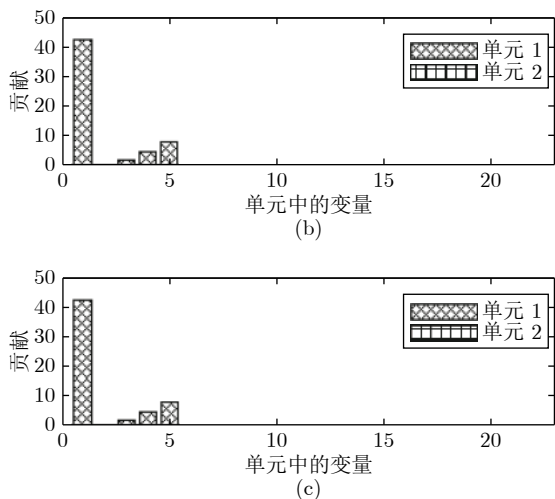


图 11 精馏过程故障 2 分层监测贡献图

Fig.11 Contribution plots for the distillation process fault 2

提升故障定位的准确性值得进一步研究. 另外, 随着数据资源的日益丰富, 如何融合多源数据信息实现过程运行状态的准确识别, 也是值得研究的课题.

附录 A CCA 故障检测中残差的统计性质^[19, 27]

假设变量组 \mathbf{u} 和 \mathbf{y} 经过均值方差归一化处理, 可以得到残差的统计性质为

$$E(\mathbf{r}_u) = \mathbf{J}^T E(\mathbf{u}) - \Xi \mathbf{L}^T E(\mathbf{y}) = \mathbf{0} \quad (A1)$$

$$E(\mathbf{r}_y) = \mathbf{L}^T E(\mathbf{y}) - \Xi \mathbf{J}^T E(\mathbf{u}) = \mathbf{0} \quad (A2)$$

$$\Sigma_{r_u} = \mathbf{I}_p - \Xi \Xi^T = \begin{bmatrix} \text{diag}\{(1 - \rho_1^2), \dots, (1 - \rho_l^2)\} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{(p-l) \times (p-l)} \end{bmatrix} \quad (A3)$$

$$\Sigma_{r_y} = \mathbf{I}_q - \Xi^T \Xi = \begin{bmatrix} \text{diag}\{(1 - \rho_1^2), \dots, (1 - \rho_l^2)\} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{(q-l) \times (q-l)} \end{bmatrix} \quad (A4)$$

其中, $\mathbf{I}_p, \mathbf{I}_q, \mathbf{I}_{(p-l) \times (p-l)}$ 以及 $\mathbf{I}_{(q-l) \times (q-l)}$ 为对应阶次的单位矩阵。

附录 B 式 (16) 中的残差向量推导过程

对 \mathbf{T}_u 和 \mathbf{T}_y 进行典型相关分析, 可构造矩阵 \mathbf{G}_T 并对其进行奇异值分解得

$$\mathbf{G}_T = \Sigma_{T_u}^{-\frac{1}{2}} \Sigma_{T_{uy}} \Sigma_{T_y}^{-\frac{1}{2}} = \mathbf{R}_T \mathbf{\Xi}_T \mathbf{V}_T^T \quad (\text{B1})$$

其中, $\Sigma_{T_u}, \Sigma_{T_{uy}}, \Sigma_{T_y}$ 为协方差矩阵。两组规范变量可计算为

$$\mathbf{J}_T = \Sigma_{T_u}^{-\frac{1}{2}} \mathbf{R}_T \quad (\text{B2})$$

$$\mathbf{L}_T = \Sigma_{T_y}^{-\frac{1}{2}} \mathbf{V}_T \quad (\text{B3})$$

由式 (B1) 和式 (B3) 可得

$$\mathbf{\Xi}_T \mathbf{V}_T^T = \mathbf{R}_T^T \Sigma_{T_u}^{-\frac{1}{2}} \Sigma_{T_{uy}} \Sigma_{T_y}^{-\frac{1}{2}} \quad (\text{B4})$$

$$\mathbf{\Xi}_T \mathbf{L}_T^T \mathbf{t}_y = \mathbf{\Xi}_T \mathbf{V}_T^T \Sigma_{T_y}^{-\frac{1}{2}} \mathbf{t}_y = \mathbf{R}_T^T \Sigma_{T_u}^{-\frac{1}{2}} \Sigma_{T_{uy}} \Sigma_{T_y}^{-1} \mathbf{t}_y \quad (\text{B5})$$

进而可得故障检测残差为

$$\mathbf{r}_u = \mathbf{J}_T^T \mathbf{t}_u - \mathbf{\Xi}_T \mathbf{L}_T^T \mathbf{t}_y = \mathbf{J}_T^T \left(\mathbf{t}_u - \Sigma_{T_{uy}} \Sigma_{T_y}^{-1} \mathbf{t}_y \right) \quad (\text{B6})$$

因此, 式 (16) 成立。式 (16) 中 $\Sigma_{T_{uy}} \Sigma_{T_y}^{-1} \mathbf{t}_y$ 一项是 \mathbf{t}_y 对 \mathbf{t}_u 的最小二乘估计。因此, \mathbf{r}_u 具有最小方差。

References

- Chai Tian-You, Ding Jin-Liang. Smart and optimal manufacturing for process industry. *Strategic Study of CAE*, 2018, **20**(4): 51-58
(柴天佑, 丁进良. 流程工业智能优化制造. 中国工程科学, 2018, **20**(4): 51-58)
- Liu Qiang, Zhuo Jie, Lang Zi-Qiang, Qin S Joe. Perspectives on data-driven operation monitoring and self-optimization of industrial processes. *Acta Automatica Sinica*, 2018, **44**(11): 1944-1956
(刘强, 卓洁, 郎自强, 秦泗钊. 数据驱动的工业过程运行监控与自优化研究展望. 自动化学报, 2018, **44**(11): 1944-1956)
- Qian Feng, Du Wen-Li, Zhong Wei-Min, Tang Yang. Problems and challenges of smart optimization manufacturing in petrochemical industries. *Acta Automatica Sinica*, 2017, **43**(6): 893-901
(钱锋, 杜文莉, 钟伟民, 唐漾. 石油和化工行业智能优化制造若干问题及挑战. 自动化学报, 2017, **43**(6): 893-901)
- Qin S J. Process data analytics in the era of big data. *AIChE Journal*, 2014, **60**(9): 3092-3100
- Zhao Chun-Hui, Yu Wan-Ke, Gao Fu-Rong. Data analytics and condition monitoring methods for nonstationary batch processes—current status and future. *Acta Automatica Sinica*, 2020, DOI: 10.16383/j.aas.c190586
(赵春晖, 余万科, 高福荣. 非平稳间歇过程数据解析与状态监控—回顾与展望, 自动化学报, 2020, DOI: 10.16383/j.aas.c190586)
- Wang Y Q, Si Y B, Huang B, and Lou Z J. Survey on the theoretical research and engineering applications of multivariate statistics process monitoring algorithms: 2008-2017. *The Canadian Journal of Chemical Engineering*, 2018, **96**(10): 2073-2085
- Ge Z Q. Review on data-driven modeling and monitoring for plant-wide industrial processes. *Chemometrics and Intelligent Laboratory Systems*, 2017, **171**: 16-25
- Jiang Q C, Yan X F, Huang B. Review and perspectives of data-driven distributed monitoring for industrial plant-wide processes. *Industrial and Engineering Chemistry Research*, 2019, **58**(29): 12899-12912
- Guo Xiao-Ping, Liu Shi-Yang, Li Yuan. Fault detection of multi-mode processes employing sparse residual distance. *Acta Automatica Sinica*, 2019, **45**(3): 617-625
(郭小萍, 刘诗洋, 李元. 基于稀疏残差距离的多工况过程故障检测方法研究. 自动化学报, 2019, **45**(3): 617-625)
- Ding S X. *Data-driven Design of Fault Diagnosis and Fault-tolerant Control Systems*: Springer London, 2014.
- Zhang Y W, Zhou H, Qin S J. Decentralized fault diagnosis of large-scale processes using multiblock kernel principal component analysis. *Acta Automatica Sinica*, 2010, **36**(4): 593-597
- Zhao C H, Sun H. Dynamic distributed monitoring strategy for large-scale nonstationary processes subject to frequently varying conditions under closed-loop control. *IEEE Transactions on Industrial Electronics*, 2019, **66**(6): 4749-4758
- Shang C, Ji H Q, Huang X L, Yang F, Huang D X. Generalized grouped contributions for hierarchical fault diagnosis with group Lasso. *Control Engineering Practice*, 2019, **93**: 104193
- Zou Xiao-Yu, Wang Fu-Li, Chang Yu-Qing, Wang Min, Cai Qing-Hong. Plant-wide process operating performance assessment and non-optimal cause identification based on hierarchical multi-block structure. *Acta Automatica Sinica*, 2019, **45**(2): 315-324
(邹筱瑜, 王福利, 常玉清, 王敏, 蔡庆宏. 基于分层分块结构的流程工业过程运行状态评价及非优原因追溯. 自动化学报, 2019, **45**(2): 315-324)
- Jiang Q C, Huang B. Distributed monitoring for large-scale processes based on multivariate statistical analysis and Bayesian method. *Journal of Process Control*, 2016, **46**: 75-83
- Luo Na, Jiang Yong, Ye Zhen-Cheng, Du Wen-Li, Qian Feng. Simulation of ethylene glycol process based on integrated modeling method. *Chinese Journal of Chemical Engineering*, 2009, **60**(1): 151-156
(罗娜, 蒋勇, 叶贞成, 杜文莉, 钱锋. 基于集成建模方法的乙二醇全流程模拟. 化工学报, 2009, **60**(1): 151-156)
- Wen Cheng-Lin, Lv Fei-Ya, Bao Zhe-Jing, Liu Mei-Qin. A review of data driven-based incipient fault diagnosis. *Acta Automatica Sinica*, 2016, **42**(9): 1285-1299
(文成林, 吕菲亚, 包哲静, 刘妹琴. 基于数据驱动的微小故障诊断方法综述. 自动化学报, 2016, **42**(9): 1285-1299)
- Yang Hao, Jiang Bin, Zhou Dong-Hua. Review and perspectives on fault tolerant control for interconnected systems. *Acta Automatica Sinica*, 2017, **43**(1): 9-19
(杨浩, 姜斌, 周东华. 互联系统容错控制的研究回顾与展望. 自动化学报, 2017, **43**(1): 9-19)
- Chen Z W, Ding S X, Zhang K, Li Z B, Hu Z K. Canonical correlation analysis-based fault detection methods with application to alumina evaporation process. *Control Engineering Practice*, 2016, **46**: 51-58
- Cao Yu-Ping, Huang Lin-Zhe, Tian Xue-Min. A process monitoring method using dynamic input-output canonical variate analysis. *Acta Automatica Sinica*, 2015, **41**(12): 2072-2080
(曹玉苹, 黄琳哲, 田学民. 一种基于 DIQVA 的过程监控方法. 自动化学报, 2015, **41**(12): 2072-2080)
- Shi L K, Tong C D, Lan T, Shi X H. Statistical process monitoring based on ensemble structure analysis. *IEEE/CAA Journal of Automatica Sinica*, 2018, DOI: 10.1109/JAS. 2017.7510877
- Peng Kai-Xiang, Ma Liang, Zhang Kai. Review of quality-related fault detection and diagnosis techniques for complex industrial processes. *Acta Automatica Sinica*, 2017, **43**(3): 349-365
(彭开香, 马亮, 张凯. 复杂工业过程质量相关的故障检测与诊断技术综述. 自动化学报, 2017, **43**(3): 349-365)

- 23 Ge Z Q, Song Z H. Distributed PCA model for plant-wide process monitoring. *Industrial and Engineering Chemistry Research*, 2013, **52**(5): 1947–1957
- 24 Liu Q, Qin S J, Chai T Y. Multiblock concurrent PLS for decentralized monitoring of continuous annealing processes. *IEEE Transactions on Industrial Electronics*, 2014, **61**(61): 6429–6437
- 25 Qin S J, Valle S, Piovoso M J. On unifying multiblock analysis with application to decentralized process monitoring. *Journal of Chemometrics*, 2001, **15**(9): 715–742
- 26 Ge Z Q. Quality prediction and analysis for large-scale processes based on multi-level principal component modeling strategy. *Control Engineering Practice*, 2014, **31**(1): 9–23
- 27 Jiang Q C, Ding S X, Wang Y, Yan X F. Data-driven distributed local fault detection for large-scale processes based on the GA-regularized canonical correlation analysis. *IEEE Transactions on Industrial Electronics*, 2017, **64**(10): 8148–8157
- 28 Wang Y, Jiang Q C, Yan X F, Fu J Q. Joint-individual monitoring of large-scale chemical processes with multiple interconnected operation units incorporating multiset CCA. *Chemometrics and Intelligent Laboratory Systems*, 2017, **166**: 14–22
- 29 Zhu J L, Ge Z Q, Song Z H. Distributed parallel PCA for modeling and monitoring of large-scale plant-wide processes with big data. *IEEE Transactions on Industrial Informatics*, 2017, **13**(4): 1877–1885
- 30 Jiang Q C, Yan X F, Huang B. Performance-driven distributed PCA process monitoring based on fault-relevant variable selection and Bayesian inference. *IEEE Transactions on Industrial Electronics*, 2015, **63**(1): 377–386
- 31 Jiang B B, Huang D X, Zhu X X, Yang F, Richard D Braatz. Canonical variate analysis-based contributions for fault identification. *Journal of Process Control*, 2015, **26**: 17–25
- 32 Westerhuis J A, Gurden S P, Smilde A K. Generalized contribution plots in multivariate statistical process monitoring. *Chemometrics and Intelligent Laboratory Systems*, 2000, **51**(1): 95–114
- 33 Kerkhof P, Vanlaer J, Gins G, Impe J. Analysis of smearing-out in contribution plot based fault isolation for statistical process control. *Chemical Engineering Science*, 2013, **104**: 285–293
- 34 Yan Z B, Yao Y. Variable selection method for fault isolation using least absolute shrinkage and selection operator (LASSO). *Chemometrics and Intelligent Laboratory Systems*, 2015, **146**: 136–146
- 35 Chen Z W, Ding S X, Peng T, Yang C H, Gui W H. Fault detection for non-Gaussian processes using generalized canonical correlation analysis and randomized algorithms. *IEEE Transactions on Industrial Electronics*, 2018, **65**(2): 1559–1567
- 36 Jiang Q C, Gao F R, Yi H, Yan X F. Multivariate statistical monitoring of key operation units of batch processes based on time-slice CCA. *IEEE Transactions on Control Systems Technology*, 2019, **27**(3): 1368–1375
- 37 Jiang Q C, Yan X F. Learning deep correlated representations for nonlinear process monitoring. *IEEE Transactions on Industrial Informatics*, 2019, **15**(12): 6200–6209
- 38 Jiang Q C, Yan X F. Multimode process monitoring using variational Bayesian inference and canonical correlation analysis. *IEEE Transactions on Automation Science and Engineering*, 2019, **16**(4): 1814–1824
- 39 Yin S, Luo H, Ding S X. Real-time implementation of fault-tolerant control systems with performance optimization. *IEEE Transactions on Industrial Electronics*, 2014, **61**(5): 2402–2411
- 40 Downs J J, Vogel E F. A plant-wide industrial process control problem. *Computers and Chemical Engineering*, 1993, **17**(3): 245–255
- 41 Chiang L H, Russell E L, Braatz R D, *Fault detection and diagnosis in industrial systems*: Springer Science and Business Media, London, 2000.



姜庆超 华东理工大学自动化系副研究员。2010年和2015年分别获得华东理工大学学士和博士学位，之后分别于阿尔伯塔大学、杜伊斯堡-艾森大学、中国香港科技大学、以及京都大学从事研究工作。主要研究方向为机器学习与工业应用，工业大数据解析，过程监测与故障诊断。

E-mail: qchjiang@ecust.edu.cn

(JIANG Qing-Chao Associate professor in the Department of Automation, East China University of Science and Technology (ECUST). He received his bachelor and Ph.D. degrees from ECUST, Shanghai, China, in 2010 and 2015, respectively. He had been a researcher with University of Alberta, Canada, University of Duisburg-Essen, Germany, the Hong Kong University of Science and Technology, China, and Kyoto University, Japan. His research interest covers machine learning and industrial application, industrial big data analysis, process monitoring, and fault diagnosis.)



颜学峰 华东理工大学自动化系教授。1995年和2002年分别获得浙江大学学士和博士学位。主要研究方向为复杂化工过程建模、优化与控制，过程监测与故障诊断，智能信息处理。本文通信作者

E-mail: xfyan@ecust.edu.cn

(YAN Xue-Feng Professor in the Department of Automation, East China University of Science and Technology. He received his bachelor and Ph.D. degrees from Zhejiang University, Hangzhou, China, in 1995 and 2002, respectively. His research interest covers complex chemical process modeling, optimizing and controlling, process monitoring, fault diagnosis, and intelligent information processing.)