



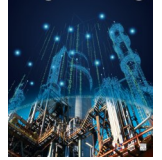
ELSEVIER

Contents lists available at ScienceDirect

Engineering

journal homepage: www.elsevier.com/locate/eng

Engineering



Research
AI Energizes Process Manufacturing—Article

一种用于工业过程监测的鲁棒迁移字典学习算法

阳春华, 梁慧平, 黄科科*, 李勇刚, 桂卫华

School of Automation, Central South University, Changsha 410083, China

ARTICLE INFO

Article history:

Received 13 May 2020

Revised 27 June 2020

Accepted 25 August 2020

Available online 4 August 2021

关键词

过程监测
多模态过程
字典学习
迁移学习

摘要

由于数据驱动的过程监测方法具有普遍性,且不依赖反应机理,其已经成为复杂工业系统过程监测的主流。然而,大多数数据驱动的过程监测方法均假设历史训练数据和在线测试数据遵循相同的分布。事实上,由于工业系统恶劣的环境,从实际工业过程中收集的数据总是受到许多因素的影响,如多变的操作环境、原材料的变化和生产指标的修改。这些因素通常会使得在线监测数据和历史训练数据分布不同,从而导致过程监测任务中的模型失配。因此,当从训练数据中学习的模型应用于实际的在线监测时,很难实现精确的过程监测。为了解决操作环境变化导致的历史训练数据和在线测试数据之间的分布差异问题,提出了一种鲁棒的迁移字典学习(RTDL)算法用于工业过程监测。RTDL是表示学习和域自适应迁移学习的协同方法。该方法将历史训练数据和在线测试数据分别作为迁移学习问题的源域和目标域。然后将最大均值差异正则化和线性判别分析正则化引入字典学习框架,可以减少源域和目标域之间的分布差异。这样,即使源域和目标域的特征在实际变化的操作环境的干扰下明显不同,仍可以学习鲁棒的字典。这样的字典可以有效地提高过程监测和模态识别的性能。通过数值仿真和两个工业系统的实验验证了该方法的有效性和优越性。

© 2021 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. 引言

过程监测对工业系统来说是必要的和有意义的,并吸引了工业界和学术界相当多的关注[1-4]。一般来说,过程监测方法分为三类:基于模型的方法、基于知识的方法和驱动驱动的方法[5-8]。基于模型的方法使用系统的数学表示,将对系统的物理理解纳入监测方案。基于知识的方法使用图形模型,如Petri网、多信号流图和贝叶斯网络(BN),用于系统监测和故障排除。这种方法特别适合于耦合系统的预测[9]。数据驱动的监测方法不需要反应

机理或过程的机理知识。近年来,通过利用智能传感器、数据分析和深度学习技术的快速进步,数据驱动方法已经被开发出来,以提高诊断的有效性和性能[10]。该领域的进展包括玻尔兹曼机、支持向量机(SVM)、卷积神经网络(CNN)等[11-13]。近年来,数据驱动的方法已经成为复杂工业过程监测的主流。

然而,目前大多数数据驱动的方法均假定历史训练数据和在线监测数据遵循相同的分布[14-16]。事实上,从实际工业过程中收集的数据总是受到许多因素的影响,如多变的操作环境、原材料的变化和生产指标的修改[17]。

* Corresponding author.

E-mail address: huangkeke@csu.edu.cn (K. Huang).

当基于训练数据学习的模型被应用于实际在线监测时, 这些因素常常导致模型失配等问题。这样的问题使其难以实现精确的过程监测。为了解决历史训练数据和在线测试数据遵循不同分布的问题, 研究者已经做了很多的工作。Hou 等[18]提出了一个用于时变过程监测的增量主成分分析 (PCA) 在线模型。在这个模型中, 当获得一个新的样本时, PCA 模型利用原来的 PCA 模型和新的样本进行更新, 预测误差平方 (SPE) 和 T^2 的控制限也被更新。Jiang 等[19]提出了一种基于当前数据和历史数据的多模型判别部分最小二乘法 (DPLS) 来诊断数据。Zhang 等[20]建立了一个深度信念网络 (DBN), 该网络具有很强的泛化能力, 可以在线监测焊接状态。为了适应新数据样本的增加, Zeng 等[21]提出了增量局部保存投影 (LPP) 算法, 该算法通过使用拉普拉斯矩阵和原始样本的投影值来更新。然而, 这些算法不能应对两个域之间的巨大差异, 例如, 被监测的工业过程处于一个完全不同的操作环境。Ge 和 Song [22]提出了在线批处理独立成分分析-主成分分析 (ICA-PCA) 方法, 在该方法中, 使用从模型库中选择合适的 ICA-PCA 模型来监测新数据。然而, 这种方法需要构建多个模型, 而且数据的不平衡性会降低监测性能。

字典学习通常学习一个过完备的字典, 然后, 原始数据可以通过字典和稀疏矩阵进行重构。真实的原始数据通常具有结构冗余的特点。通过字典学习, 原始数据被映射到一个较低的维度空间, 从而消除了原始数据的结构冗余, 同时保留了最简洁的信息。Peng 等[23]提出了一种通过 LPP 计算出一个映射字典的方法, 以保留原始数据的几何结构。Chen 等[24]提出了一种利用字典正则化从少量目标域数据中学习字典的方法, 同时保障这个字典与从源域数据中学习的字典相似。Zhang 等[25]提出了一种方法, 在该方法中, 为了实现跨域分类, 学习了一个公共字典、一个源域字典和一个目标域字典。在 Jie 等[26]提出的方法中, 学习了源域字典和目标域字典之间的若干个子空间字典。Long 等[27]提出了迁移稀疏代码 (TSC), 该方法引入了图的正则化并减少了分布距离, 以实现知识迁移。这些字典学习方法在信号重建、信号降噪、图像识别、图像校正等方面取得了很大的成就, 引起了学术界和工业界的高度重视。此外, 最近的研究表明, 字典学习在过程监测中具有非凡的优势。Huang 等[28]提出了一种核字典学习方法用于实现非线性过程监测, 以及一种分布式字典学习方法用于实现高维过程监测。

尽管上述方法在工业系统中表现出卓越的过程监测性能, 但它们没有考虑历史数据和在线监测数据的分布差异。迁移学习是不同域间知识迁移的一般框架, 近年来得

到了广泛的研究, 特别是在人工智能、图像识别和计算机视觉界。受字典学习强大的表示能力和迁移学习的跨域知识迁移能力的启发, 本文提出了一种鲁棒迁移字典学习 (RTDL) 方法, 用于处理现实工业过程监测的数据分布畸变问题。本方法是联合表示学习和域自适应性迁移学习的一种协同方法。简而言之, 历史训练数据和在线测试数据被分别作为迁移学习问题的源域和目标域。对于一个工业系统来说, 虽然系统总是在不同的运行环境中运行, 会不可避免地导致不同的数据分布, 但其内部的基本信息, 如机理, 往往是相同或相似的。换句话说, 工业系统的高维观测数据往往在不同域间有不变的子空间。因此, 将源域和目标域映射到一个公共的子空间是可行的, 在该子空间中源域和目标域之间的分布差异被消除。之后, 在该子空间中进行过程监测。在实际应用中, 提出了一种判别性的字典学习方法, 从多模态工业数据中提取特征。接下来, 提出了最大平均差异 (MMD) 正则化[29]作为非参数距离度量来表达分布距离, 减少了源域数据和目标域数据之间的分布畸变。此外, 为了减少模内数据的距离, 引入了线性判别分析 (LDA) 正则化。因此, 即使源域和目标域受到不同操作环境的严重影响, 本方法也能学习到一个鲁棒的公共字典。因此, 该方法能有效地提高过程监测和模态识别的性能。

本文的主要贡献如下。第一, 提出了一种 RTDL 方法, 以减少工业系统多变的运行环境所带来的负面影响。通过减少域间差异, 本方法可减少由多变的运行环境引起的过程监测和模态识别的性能下降。第二, 给出了关于带约束不可微分字典学习问题的详细优化步骤, 可以有效实现 RTDL 的求解。第三, 通过大量的实验验证了本方法, 包括数值仿真实验和实际工业实验。实验结果表明, 该方法在精度上优于一些先进的方法。因此, 该方法适用于工业系统的过程监测任务。

本文的其余部分组织如下。第2节简要介绍了域自适应迁移学习、字典学习和本文的动机。第3节提出了 RTDL 模型并给出了其有效的优化步骤。第4节进行了广泛的实验, 包括数值仿真、连续搅拌釜加热器 (CSTH) 基准和风机系统实验, 以验证 RTDL 方法的有效性。第5节为结论和总结意见。

2. 相关知识

2.1. 域自适应迁移

假设有一个拥有大量数据的源域 ($\mathbf{X}_s = [x_{s1}, x_{s2}, \dots$,

$x_{n_s}]$, 其中 n_s 是源域数据 \mathbf{X}_s 的数量) 和一个拥有少量数据的目标域 ($\mathbf{X}_t=[x_{t1}, x_{t2}, \dots, x_{t_{n_t}}]$, 其中 n_t 是目标域数据 \mathbf{X}_t 的数量)。这里, 源域和目标域是相关的, 但不相同。以风机系统为例, 将风机系统在冬季的过程数据设为源域, 在夏季的过程数据设为目标域。这两个域是相关的, 因为每个域的数据均是在相同的机理下从同一个风机系统中采集的。然而, 由于外部运行环境的不同, 两个季节的观测数据往往具有不同的分布。在数学上, 各域的输入特征空间是相同的, 但各域的边缘分布和条件分布是不同的, 即 $\mathbf{X}_s \in \mathcal{X}$ 、 $\mathbf{X}_t \in \mathcal{X}$ 、 $P_s(x) \neq P_t(x)$ 、 $P_s(y|x) \neq P_t(y|x)$ 。其中 P_s 和 P_t 分别代表源域和目标域的概率分布; \mathcal{X} 代表数据空间; x 代表数据样本; y 代表数据 x 的标签。

为了实现目标域的过程监测, 需要消除源域和目标域之间的分布差异, 即条件分布差异和边缘分布差异。当源域和目标域的分布畸变消除后, 源域数据可以表现出与目标域数据相同的过程信息, 因此该方法可以利用丰富的源域数据辅助训练模型, 从而达到积极的知识迁移效果。

2.2. 字典学习

字典学习的思想是通过学习由一系列原子组成的字典和一个稀疏矩阵来最小化数据重构误差。假设 $\mathbf{X}_N=[x_1, x_2, \dots, x_N] \in \mathbf{R}^{m \times N}$ 是一组原始样本, 其中 x_N 表示维数为 m 的第 N 个样本, \mathbf{R} 是向量空间, N 表示数据量。 $\mathbf{D}_K=[d_1, d_2, \dots, d_K] \in \mathbf{R}^{m \times K}$ 是由 K 个原子组成的字典, 其中 d_K 是第 K 个原子, K 是原子数。 $\mathbf{S}_N \in \mathbf{R}^{K \times N}$ 代表稀疏矩阵。字典学习的问题可以表示如下:

$$(\mathbf{D}_K, \mathbf{S}_N) = \arg \min_{\mathbf{D}_K, \mathbf{S}_N} \left\| \mathbf{X}_N - \mathbf{D}_K \mathbf{S}_N \right\|_F^2 + \alpha \left\| \mathbf{S}_N \right\|_0 \quad (1)$$

式中, α ($\alpha > 0$) 是控制 \mathbf{S}_N 稀疏性的参数; $\|\cdot\|_F$ 表示矩阵 \mathbf{X}_N 的 F 范数; $\|\cdot\|_0$ 表示矩阵 \mathbf{S}_N 的 L_0 范数。

2.3. 本文动机

如前所述, 尽管一个工业系统在不同的运行环境下(如不同的地点、时间、天气和人工操作等外部干扰)运行时, 会不可避免地导致数据分布的畸变, 但基本的内部机理往往是相同或相似的。也就是说, 不同域下的工业系统的高维观测数据往往具有不变的子空间。因此, 有必要提取不变的知识或子空间, 以消除外在干扰, 从而进一步提高工业过程监测的性能。为了生动地展示源域数据和目标域数据之间可迁移特征的效果, 图1显示了仅受不同环境因素影响而导致不同分布的数据散点图。源域和目标域的边缘分布和条件分布明显不同。传统的数据驱动方法有

两种常见的策略。如图1(a)所示, 第一种策略是忽略源域数据, 只用目标域数据作为输入数据, 这样可以满足训练数据的分布与测试数据的分布相同的假设。然而, 由于目标域的训练数据非常少, 最终的模型很容易出现过拟合。如图1(b)所示, 第二种策略是忽略源域和目标域的不同特征。这种策略直接使用大量的历史数据和少量的新数据来进行模型训练任务, 因此, 最终的模型会混淆域间差异信息和异常信息。此外, 该模型很容易被拥有大量样本的源域所支配。相比之下, 图1(c)是本文提出的RTDL模型。该方法试图找到一个映射关系函数 $\Phi(\cdot)$ 。通过这个映射关系, 原始数据被映射到一个子空间。在这个子空间中, 源域的边缘分布和条件分布与目标域的分布相同; 也就是说 $P_s(\Phi(x))=P_t(\Phi(x))$, $P_s(y|\Phi(x))=P_t(y|\Phi(x))$, 其中 $\Phi(x)$ 是关于 x 的映射。我们认为如果 $\Phi(\cdot)$ 能克服外在环境的干扰, 只保留最简明的内部机理信息, 就能将知识从源域迁移到目标域。也就是说, 通过在字典学习目标函数中加入MMD和LDA-like的正则化, RTDL方法可以利用丰富的源域数据来辅助训练模型, 达到迁移的效果, 以提高工业过程监测的效果。

3. 方法

在详细讨论该方法之前, 这里先引入一个假设。这个假设是合理的, 且通常情况下工业系统满足这个假设。

假设: 一个复杂的工业过程常常以不同的模态运行, 以满足不同的现实需求。在不同的模态下, 观察到的变量的特征是不同的。为了清楚地描述不同的观察结果, 历史训练数据和在线测试数据分别被看作是源域和目标域。

一般来说, 进行多模态数据过程监测有两种方式。第一种方式是单独处理多模态数据, 然后单独完成过程监测任务。第二种方式是全局处理多模态数据, 然后使用单一模型完成过程监测任务。当在每个单独模态下收集的数据足够时, 第一种方式是更好的选择。然而, 对于实际的工业过程监测任务来说, 目标域的数据总是严重不足, 单独的方法容易造成过拟合, 因此, 最好是全局处理多模态数据。此外, 观察到的变量不仅由工业过程的内部机理决定, 还受到外在环境(如人工操作、不确定性、参数测量的不连续性、噪声等)的影响。在线测试数据的外在环境与历史训练数据的外在环境不同, 所以会出现域畸变。为了获得准确的过程监测结果, 一个明智的选择是通过使用域自适应迁移学习的方法来消除无关的外在干扰。

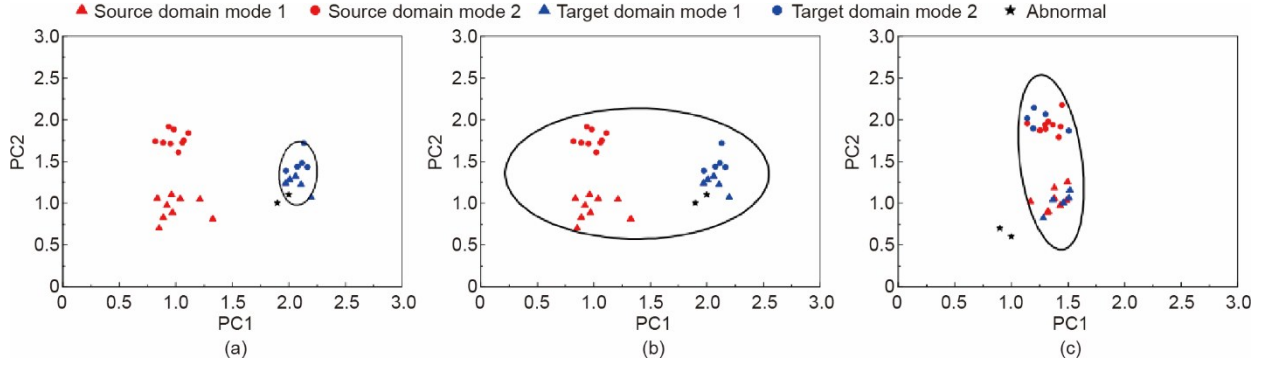


图1. 源域数据量大而目标域数据量小的情况。(a) 传统策略1: 忽略源域, 仅利用目标域数据进行模型训练。(b) 传统策略2: 忽略源域和目标域的特征差异, 直接利用所有数据进行模型训练。(c) RTDL模型将源域和目标域的数据进行正则化约束, 消除域间差异。该模型是三种模型中最合理的。PC1和PC2表示数据的两个主要成分。

3.1. 判别性字典

传统的字典学习已经被广泛地引入过程监测。此外, 最近的研究表明, 学习判别性字典可以使字典具备模态识别能力[30-33]。因此, 对于过程监测任务来说, 迫切需要一种判别性的字典学习方法。这里, 判别式字典被表示为 $\mathbf{D}=[\mathbf{D}_1, \mathbf{D}_2, \dots, \mathbf{D}_C] \in \mathbf{R}^{m \times Ck}$, 其中 C 代表模态的数量; k 代表每个模态的原子数量; \mathbf{D}_C 是一个由 k 个原子组成的子字典, 用来代表第 C 个模态的特征。原始样本矩阵 (\mathbf{X}) 关于字典 \mathbf{D} 的稀疏矩阵为 $\mathbf{S}=[\mathbf{S}_1, \mathbf{S}_2, \dots, \mathbf{S}_C] \in \mathbf{R}^{Ck \times (n_s + n_t)}$, 其中 $\mathbf{S}_C=[\mathbf{S}_{s_C}, \mathbf{S}_{t_C}]$; \mathbf{S}_{s_C} 代表第 C 个模态源域数据的稀疏编码; \mathbf{S}_{t_C} 代表第 C 个模态目标域数据的稀疏编码。为了简单起见, 我们将 \mathbf{X}_i 关于字典 \mathbf{D} 的稀疏编码记为 $\mathbf{S}_i=[\mathbf{S}_i^1; \mathbf{S}_i^2; \dots; \mathbf{S}_i^C]$, 其中 \mathbf{S}_i^c 是 \mathbf{X}_i 关于子字典 \mathbf{D}_C 的编码系数; \mathbf{X}_i 是由第 i 个模态样本组成的矩阵 ($i \in \{1, 2, \dots, C\}$), 是 \mathbf{X} 的子矩阵。

为了提高多模态数据的表示能力, 应该在字典学习中加入先验约束。第一, 数据应该被字典和相应的稀疏矩阵很好地重构, 即 $\mathbf{X} \approx \mathbf{D}\mathbf{S}$ 。第二, 数据应该被它自己的子字典和子稀疏矩阵很好地表示出来; 即 $\mathbf{X}_i \approx \mathbf{D}_i \mathbf{S}_i^i \forall i \in \{1, 2, \dots, C\}$ 。第三, 由于数据可以由它自己的子字典和子稀疏矩阵很好地表示, 因此 $\mathbf{S}_i^i (i' \neq i)$ 项应该尽可能地接近于零, 公式 (1) 可以转化为以下形式[31]。

$$\begin{aligned} (\mathbf{D}, \mathbf{S}) = \arg \min_{\mathbf{D}, \mathbf{S}} & \|\mathbf{X} - \mathbf{D}\mathbf{S}\|_F^2 + \|\mathbf{X} - \mathbf{D}\mathbf{S}_{\text{in}}\|_F^2 \\ & + \|\mathbf{D}\mathbf{S}_{\text{out}}\|_F^2 + \alpha \|\mathbf{S}\|_0 \\ \text{s.t. } & \forall \|d_b\|_2 \leq 1 \end{aligned} \quad (2)$$

式中, d_b 代表字典 \mathbf{D} 中的第 b 个原子; \mathbf{S}_{in} 和 \mathbf{S}_{out} 是关于 \mathbf{S} 的表达式, 如下所示。

$$\mathbf{S}_{\text{in}}[a, b] = \begin{cases} \mathbf{S}[a, b] & x_a, d_b \in \text{same mode} \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

$$\mathbf{S}_{\text{out}}[a, b] = \begin{cases} \mathbf{S}[a, b] & x_a, d_b \notin \text{same mode} \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

式中, x_a 是 \mathbf{X} 的第 a 个样本。

特别地, $\|\mathbf{X} - \mathbf{D}\mathbf{S}_{\text{in}}\|_F^2 = \sum_{i=1}^C \|\mathbf{X}_i - \mathbf{D}_i \mathbf{S}_i^i\|_F^2$ 意味着数据 \mathbf{X}_i 应该由它自己的子字典 \mathbf{D}_i 和子稀疏矩阵 \mathbf{S}_i^i 来很好地表示。此外, $\|\mathbf{D}\mathbf{S}_{\text{out}}\|_F^2 = \sum_{i=1}^C \sum_{i' \neq i} \|\mathbf{D}_i \mathbf{S}_i^{i'}\|_F^2$, $\mathbf{S}_i^{i'} (i' \neq i)$ 项应尽可能地接近于零。

3.2. 正则化

由于工业系统的源域和目标域受到不同环境因素的影响, 所以数据分布是不同的。为了确保学习到的字典能够捕捉到源域和目标域中潜在的公共机理信息, 而不是无关的外在干扰, 一个直接的方法是通过最小化一些预定义的度量来减少分布差异。MMD 正则化被认为是一种表达域间分布差异的非参数指标[27], 它可以使稀疏矩阵 \mathbf{S}_{s_i} 和 \mathbf{S}_{t_i} 的中心接近。在数学上, MMD 正则化表示如下。

$$\begin{aligned} c_1 &= \sum_{i=1}^C \left\| \frac{1}{n_{s_i}} \sum_{s_j \in S_{s_i}} s_j - \frac{1}{n_{t_i}} \sum_{s_j \in S_{t_i}} s_j \right\|_F^2 \\ &= \sum_{i=1}^C \text{tr}(\mathbf{S}_i \mathbf{M}^i \mathbf{S}_i^T) = \text{tr}(\mathbf{S} \mathbf{M} \mathbf{S}^T) \end{aligned} \quad (5)$$

式中, c_1 代表源域和目标域之间的 MMD 正则化; n_{s_i} 和 n_{t_i} 分别代表源域和目标域中第 i 个模态的样本数; s_j 是稀疏矩阵 \mathbf{S} 中的一个稀疏代码; \mathbf{M}^i 代表 MMD 矩阵 \mathbf{M} 的第 i 个模态部分, 可以计算如下。

$$\mathbf{M}_{aa}^i = \begin{cases} 1/n_{s_i}^2 & x_a, x_a \in \text{source domain} \\ 1/n_{t_i}^2 & x_a, x_a \in \text{target domain} \\ -1/n_{s_i} n_{t_i} & \text{otherwise} \end{cases} \quad (6)$$

$$\mathbf{M} = \text{diag}(\mathbf{M}^1, \mathbf{M}^2, \dots, \mathbf{M}^C)$$

式中, M_{aa}^i 代表 \mathbf{M}^i 的第 a 行和第 a 列中的元素; x_a 是 \mathbf{X} 的

第 \bar{a} 个样本; $x_{\bar{a}}$ 是 \mathbf{X} 的第 \bar{a} 个样本。

众所周知, 同一模态下的观测数据的分布应该是相同的。然而, 由于工作环境的不确定性, 同一模态下的数据也表现出一定的差异。为了消除不确定环境的干扰, 无论数据来自哪个域, 最好使同模态数据的稀疏码彼此更加接近。也就是说, $\mathbf{S}_i = [\mathbf{S}_{s_i}, \mathbf{S}_{t_i}] \forall i \in \{1, 2, \dots, C\}$ 的每一个列向量均应该相互接近。因此, $\sum_{s_j \in S_i} (s_j - \bar{s}_i)^T (s_j - \bar{s}_i)$ 的值应该尽可能小, 其中 \bar{s}_i 是 S_i 的中心。因此, 应该引入以下约束条件。

$$c_2 = \sum_{i=1}^C \sum_{s_j \in S_i} (s_j - \bar{s}_i)^T (s_j - \bar{s}_i) = \text{tr}(\mathbf{S}(\mathbf{I} - \mathbf{H})\mathbf{S}^T) \quad (7)$$

式中, c_2 是所有模态的模内距离; \mathbf{I} 是单位矩阵。LDA-like 矩阵 \mathbf{H} 可以通过以下方式得到。

$$\mathbf{H} = \text{diag}(\mathbf{H}_1, \mathbf{H}_2, \dots, \mathbf{H}_C), \mathbf{H}_i = \frac{1}{n_i} \mathbf{1}_{n_i} \mathbf{1}_{n_i}^T \quad (8)$$

式中, $\mathbf{1}_{n_i}$ 是一个长度为 n_i 及所有元素均等于 1 的向量; n_i 是第 i 个模态在源域和目标域的总数据数。公式 (7) 的正则化形式上与 LDA 正则化相似[34], 所以它被称为 LDA-like 正则化。

总之, MMD 和 LDA-like 正则化可以被认为是一种渐进的关系。MMD 正则化使每个模态的源域中心接近每个模态的目标域中心。作为补充, LDA-like 正则化使同一模态的数据相互靠近, 以减少模态内的距离。基于这两种正则化的协同作用, 我们将公式 (2)、(3) 和 (7) 连接起来, 形成了本文中所提出的方法, 该方法的目标函数可以表示如下。

$$\begin{aligned} (\mathbf{D}, \mathbf{S}) = \arg \min_{\mathbf{D}, \mathbf{S}} & \|\mathbf{X} - \mathbf{D}\mathbf{S}\|_F^2 + \|\mathbf{X} - \mathbf{D}\mathbf{S}_{\text{in}}\|_F^2 \\ & + \|\mathbf{D}\mathbf{S}_{\text{out}}\|_F^2 + \alpha \|\mathbf{S}\|_0 + \beta_1 \text{tr}(\mathbf{S}\mathbf{M}\mathbf{S}^T) \\ & + \beta_2 \text{tr}(\mathbf{S}(\mathbf{I} - \mathbf{H})\mathbf{S}^T) \end{aligned} \quad (9)$$

$$\text{s.t. } \forall d_b \leq 1$$

式中, β_1 和 β_2 ($\beta_1, \beta_2 > 0$) 分别是 MMD 正则化和 LDA-like 正则化的超参数, 这些超参数可以平衡重建误差和正则化约束之间的权重。由于本方法减少了源域和目标域之间的分布畸变, 因此可以达到消除外在环境干扰的效果。

3.3. 优化

上述目标函数的优化变量是 \mathbf{D} 和 \mathbf{S} 。由于优化问题不是两个变量的联合凸问题, 而是分别凸于 \mathbf{D} (保持 \mathbf{S} 固定的情况下) 和凸于 \mathbf{S} (保持 \mathbf{D} 固定的情况下), 因此引入迭代优化方法来计算 \mathbf{D} 和 \mathbf{S} 的最优值[35]。

3.3.1. 更新 \mathbf{D}

当通过固定 \mathbf{S} 更新字典 \mathbf{D} 时, 目标函数可以简化为如下形式:

$$\begin{aligned} \mathbf{D} = \arg \min_{\mathbf{D}} & \|\mathbf{X} - \mathbf{D}\mathbf{S}\|_F^2 + \|\mathbf{X} - \mathbf{D}\mathbf{S}_{\text{in}}\|_F^2 + \|\mathbf{D}\mathbf{S}_{\text{out}}\|_F^2 \\ \text{s.t. } & \forall \|d_b\|_2 \leq 1 \end{aligned} \quad (10)$$

构建了一个新的数据矩阵 $\mathbf{X}_{\text{new}} = [\mathbf{X}, \mathbf{X}, \mathbf{O}]$ 和一个新的稀疏矩阵 $\mathbf{S}_{\text{new}} = [\mathbf{S}, \mathbf{S}_{\text{in}}, \mathbf{S}_{\text{out}}]$, 其中 \mathbf{O} 是与矩阵 \mathbf{X} 维度相同的零矩阵, 目标函数[公式 (10)]变为如下形式:

$$\mathbf{D} = \arg \min_{\mathbf{D}} \|\mathbf{X}_{\text{new}} - \mathbf{D}\mathbf{S}_{\text{new}}\|_F^2, \forall \|d_b\|_2 \leq 1 \quad (11)$$

该目标函数可以通过使用拉格朗日对偶法[35]来有效解决。首先, 考虑拉格朗日函数:

$$\begin{aligned} L(\mathbf{D}, \vec{\lambda}) = & \text{trace}((\mathbf{X}_{\text{new}} - \mathbf{D}\mathbf{S}_{\text{new}})^T (\mathbf{X}_{\text{new}} - \mathbf{D}\mathbf{S}_{\text{new}})) \\ & + \lambda_j \left(\sum_{i=1}^m \mathbf{D}[i, j]^2 - 1 \right), \forall \lambda_j \geq 0 \end{aligned} \quad (12)$$

式中, L 是拉格朗日函数; $\vec{\lambda} = [\lambda_1, \lambda_2, \dots, \lambda_{Ck}]$ 是引入的非负参数。

通过最小化拉格朗日函数[公式 (12)]得到拉格朗日对偶问题, 如下所示。

$$\begin{aligned} B(\vec{\lambda}) = & \min_{\mathbf{D}} L(\mathbf{D}, \vec{\lambda}) \\ = & \text{trace}(\mathbf{X}_{\text{new}}^T \mathbf{X}_{\text{new}} - \mathbf{A} - \mathbf{X}_{\text{new}} \mathbf{S}_{\text{new}}^T \\ & (\mathbf{S}_{\text{new}} \mathbf{S}_{\text{new}}^T + \mathbf{A})^{-1} (\mathbf{X}_{\text{new}} \mathbf{S}_{\text{new}}^T)^T) \end{aligned} \quad (13)$$

式中, B 是拉格朗日对偶公式; \mathbf{A} 是由 $\vec{\lambda}$ 组成的对角矩阵, $\mathbf{A} = \text{diag}(\vec{\lambda})$ 。拉格朗日对偶问题[公式 (13)]可以用牛顿法或共轭梯度法来优化。在最大化 $B(\vec{\lambda})$ 后, 我们得到最优字典 \mathbf{D} 如下。

$$\mathbf{D}^T = (\mathbf{S}_{\text{new}} \mathbf{S}_{\text{new}}^T + \mathbf{A})^{-1} (\mathbf{X}_{\text{new}} \mathbf{S}_{\text{new}}^T)^T \quad (14)$$

3.3.2. 更新 \mathbf{S}

矩阵 \mathbf{S} 将被逐列进行更新。当更新其中一列 s_j 时, 目标函数可以表示如下。

$$\begin{aligned} s_j = \arg \min_{s_j} & f(s_j) \\ = \arg \min_{s_j} & \|x_j - \mathbf{D}s_j\|_F^2 + \|x_j - \mathbf{D}\mathbf{Q}s_j\|_F^2 + \|\mathbf{D}\mathbf{P}s_j\|_F^2 \\ & + \alpha \sum_{p=1}^{Ck} |s_j^{(p)}| + (\beta_1 \mathbf{M}_{jj} + \beta_2 (\mathbf{I} - \mathbf{H})_{jj}) s_j^T s_j + s_j^T h_j \end{aligned} \quad (15)$$

式中, \mathbf{Q} 、 \mathbf{P} 和 h_j 是目标函数简化过程中的中间变量。 $\mathbf{Q} = \text{diag}(0, 0, \dots, 1, 1, 1, \dots, 0, 0, 0) \in \mathbf{R}^{Ck \times Ck}$; $\mathbf{P} = \mathbf{I} - \mathbf{Q}$; $h_j = 2 \sum_{j \neq j} (\beta_1 \mathbf{M}_{jj} + \beta_2 (\mathbf{I} - \mathbf{H})_{jj}) s_j$ 。 \mathbf{Q} 是一个对角线矩阵, 元素 1 只存在于 s_j 所属模态的相应的 k 个位置上; x_j 是 \mathbf{X} 中的第 j 个样

本； $s_j^{(p)}$ 是 s_j 中的第 p 个元素。我们通过特征符号搜索算法[27,35]来优化公式(15)。定义 $g(s_j)$ 如下：

$$g(s_j) = \|x_j - \mathbf{D}s_j\|_F^2 + \|x_j - \mathbf{DQ}s_j\|_F^2 + \|\mathbf{DP}s_j\|_F^2 + (\beta_1 \mathbf{M}_{jj} + \beta_2 (\mathbf{I} - \mathbf{H})_{jj}) s_j^T s_j + s_j^T h_j \quad (16)$$

式中， $g(s_j)$ 是公式(15)的可微分部分。为了实现公式(15)中的特征符号搜索算法，引入以下定理。

定理 1: 定义一个关于 x 的连续函数为 $F(x) = G(x) + \lambda \|x\|_1$ ； $x = \arg \min_x F(x)$ 的最优必要条件为

$$\begin{cases} \nabla^p G(x) + \lambda \text{sign}(x^p) = 0, & \text{if } |x^p| \neq 0 \\ \|\nabla^p G(x)\| \leq \lambda, & \text{if } |x^p| = 0 \end{cases} \quad \forall x^p \quad (17)$$

式中， $G(x)$ 是关于向量 x 的连续可微函数[35-36]； x^p 是向量 x 的第 p 个元素； $\nabla^p G(x)$ 是 $G(x)$ 关于 x^p 的偏导。

证明: 我们通过反证法来提供一个简短的证明。假设最优解 x 中有一个元素 x^p 不符合上述最优必要条件。首先，对于 $|x^p| \neq 0$ ， $\nabla^p G(x) + \lambda \text{sign}(x^p) \neq 0$ ，则很明显可知 $\nabla^p F(x) = \nabla^p G(x) + \lambda \text{sign}(x^p) \neq 0$ 。因此，我们可以找到另一个值 x^{*p} 来代替 x^p ，使 $F(x^*)$ 值更小。这与 x 为最优解的假设是矛盾的。其次，对于 $|x^p| = 0$ ， $\nabla^p G(x) > \lambda$ ，由于 $G(x)$ 是一个连续可微的函数，我们可以找到一个 $x^{*p} < 0$ 来代替 x^p ，并且满足 $G(x^*) - G(x) < \lambda x^{*p}$ 。因此 $F(x^*) = G(x^*) + \lambda \|x^*\|_1 = G(x^*) + \lambda \|x\|_1 - \lambda x^{*p} < G(x) + \lambda \|x\|_1 = F(x)$ ，这也是与假设相矛盾的。对于 $|x^p| = 0$ ， $\nabla^p G(x) < -\lambda$ ，可以用同样的方法来证明该假设不成立。综上所述，命题得证。

根据定理 1，公式(15)中的必要条件可以描述如下。

$$\begin{cases} \nabla^{(p)} g(s_j) + a \text{sign}(s_j^{(p)}) = 0, & \text{if } |s_j^{(p)}| \neq 0 \\ \|\nabla^{(p)} g(s_j)\| \leq a, & \text{if } |s_j^{(p)}| = 0 \end{cases} \quad (18)$$

当第一个条件被违反时，因为其符号是已知的，公式(15)中的目标函数是可微的，它成为一个无约束的优化问题(QP)。当第二个条件被违反时，假设 $\nabla^{(p)} g(s_j) > a$ ，由于 $\nabla^{(p)} f(s_j)$ 一定大于零，为了最小化 $f(s_j)$ 的值， $s_j^{(p)}$ 必须减少。由于 $s_j^{(p)}$ 从零开始，对其进行的任何无限小的调整均会使其符号为负，因此，我们直接让 $s_j^{(p)}$ 为-1。那么， $f(s_j)$ 关于 $s_j^{(p)}$ 同样是可微的，问题可以简单解决。如果 $\nabla^{(p)} g(s_j) < -a$ ， $s_j^{(p)}$ 可以用同样的方法更新。

据此，稀疏矩阵优化算法的完整过程被总结为算法1。

3.4. 在线过程监测

在获得通用字典 \mathbf{D} 后，利用该字典可以进行过程监测

和模态识别任务。

3.4.1. 过程监测

字典 \mathbf{D} 和正交匹配追求(OMP)算法[37]被用来计算训练集中目标域数据的稀疏表示，训练集中目标域数据的重构残差(RES)可以根据公式(19)得到。

$$r = \|x - \mathbf{D}s\|_2 \quad (19)$$

式中， $\|\cdot\|_2$ 代表向量的 L_2 范数； r 是样本 x 的重建误差。

接下来，用核密度估计法(KDE)[38]计算目标域数据的残差分布区间，以检测新的测试数据是否正常。当新的测试数据 x 来自目标域时，使用同样的OMP算法得到稀疏代码 s ；然后使用公式(19)可以得到测试数据的RES。当RES属于上述分布区间时，它是正常数据，否则，就是故障数据。

3.4.2. 模态识别

测试数据被检测为正常数据后，进行模态识别，通过公式(20)来识别数据 x 。

$$\text{mode} = \arg \min_i \|x - \mathbf{D}_i s^i\|_F^2 \quad (20)$$

式中， s^i 是 s 的子稀疏表示。

综上所述，RTDL的完整流程总结于算法2。

4. 验证实验

在本节中，我们在数值仿真、CSTH基准和风机系统上进行了广泛的实验，以证明RTDL方法的有效性和优越性。为了性能的可视化和参数的敏感性分析，在数值仿真中分别引入MMD正则化和LDA-like正则化，以便能够直观地观察样本的分布。对于CSTH基准和风机系统，将提出的方法与一些最先进的方法进行了性能比较。

4.1. 性能可视化和参数敏感性分析

4.1.1. 数据集

本小节在数值仿真上进行了实验，以验证本文动机并直观地评价性能。数值仿真的数据生成模型如下。

$$\bar{\mathbf{x}} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \\ \bar{x}_4 \\ \bar{x}_5 \end{bmatrix} = \begin{bmatrix} 0.5768 & 0.3766 \\ 0.7382 & 0.0566 \\ 0.8291 & 0.4009 \\ 0.6519 & 0.2070 \\ 0.3972 & 0.8045 \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix} \quad (21)$$

式中， $\bar{\mathbf{x}} = [\bar{x}_1, \bar{x}_2, \bar{x}_3, \bar{x}_4, \bar{x}_5]^T$ 是过程监测的观测变量； t_1 和 t_2 是数据仿真模型的两个独立输入变量； e_1, e_2, \dots, e_5 是由于当前环境干扰造成的观测变量的噪声。为简洁起见，我们

Algorithm 1. Sparse matrix optimizing algorithm.**Input:** Data matrix X , MMD matrix M , LDA-like matrix H , parameters α , β_1 , and β_2 **Output:** optimal sparse matrix S

Begin optimizing algorithm

For each s_j in S do**Step initialization:** $s_j = \mathbf{0}$, $\theta = \mathbf{0}$, active set $A = \{ \}$, where θ is a vector that represents a symbol of s_j , $\theta_p \in \{-1, 0, 1\}$ denotes $\text{sign}(s_j^{(p)})$ **Step activate:**From zero coefficients of s_j , select $p := \arg \max_p |\nabla^{(p)} g(s_j)|$. Activate $s_j^{(p)}$ (add p to A) only if it locally improves Eq. (15); that is, if $\nabla^{(p)} g(s_j) > \alpha$, then set $\theta_p = -1$, $A = \{p\} \cup A$; otherwise, if $\nabla^{(p)} g(s_j) < -\alpha$, then set $\theta_p = 1$, $A = \{p\} \cup A$ **Step feature-sign search:**Let \hat{D} , \widehat{DQ} , and \widehat{DP} be a submatrix of D , DQ , and DP that contains only the columns in A ; let \hat{s}_j , \hat{h}_j , and $\hat{\theta}_j$ be sub-vectors of s_j , h_j , and θ in A , respectively

Compute the solution to the resulting unconstrained quadratic optimization problem (QP):

$$\min \hat{f}(\hat{s}_j) = \|x_j - \hat{D}\hat{s}_j\|_F^2 + \|x_j - \widehat{DQ}\hat{s}_j\|_F^2 + \|\widehat{DP}\hat{s}_j\|_F^2 + (\beta_1 M_{jj} + \beta_2 (\mathbf{I} - H)_{jj}) \hat{s}_j^T \hat{s}_j + \hat{s}_j^T \hat{h}_j + \alpha \hat{\theta}_j^T \hat{s}_j$$

If we let $\partial \hat{f}(\hat{s}_j) / \partial \hat{s}_j = 0$, we can obtain the optimal value of s_j under the current A :

$$\hat{s}_j^{\text{new}} = \left\{ \hat{D}^T \hat{D} + \widehat{DQ}^T \widehat{DQ} + \widehat{DP}^T \widehat{DP} + \left(\beta_1 M_{jj} + \beta_2 (\mathbf{I} - H)_{jj} \right) \mathbf{I} \right\}^{-1} \left(\hat{D}^T x_j + \widehat{DP}^T x_j - (\alpha \hat{\theta}_j + \hat{h}_j) / 2 \right)$$

Perform a discrete line search on the closed line segment from \hat{s}_j to \hat{s}_j^{new} Check the objective value at \hat{s}_j^{new} and all points where any coefficient changes signUpdate \hat{s}_j (and the corresponding entries in s) to the point with the lowest objective valueRemove zero coefficients of \hat{s}_j from the active set A and update $\theta := \text{sign}(s_j)$ **Step check optimality conditions:**(i) Optimality condition for nonzero coefficients: $\nabla^{(p)} g(s_j) + \alpha \text{sign}(s_j^{(p)}) = 0$, $\forall s_j^{(p)} \neq 0$

If condition (i) is not satisfied, go to the step “feature-sign search” (without any new activation); otherwise, check condition (ii)

(ii) Optimality condition for zero coefficients: $|\nabla^{(p)} g(s_j)| \leq \alpha$, $\forall s_j^{(p)} = 0$ If condition (ii) is not satisfied, go to the step “activate;” otherwise, return s_j as the optimal solution**End for****End optimal algorithm**

假设在源域和目标域有两种模态。模态状态由独立的输入变量 t_1 和 t_2 决定，而环境干扰 (e_1, e_2, \dots, e_s) 导致域的差异。数据的详细描述见表1，其中 $N(u, \sigma^2)$ 表示均值为 u 、方差为 σ^2 的高斯分布； $U(\underline{\omega}, \bar{\omega})$ 表示范围从 $\underline{\omega}$ 到 $\bar{\omega}$ 的均匀分布； $E(\lambda)$ 表示参数为 λ 的指数分布； $\bar{F}(\gamma, \rho)$ 表示自由度为 γ 和 ρ 的 F 分布。异常数据从目标域的模态2中收集，在 x_2 维度上发生了加性故障。

对于训练数据，我们在源域的每个模态收集了100条正常数据，在目标域的每个模态收集了10条正常数据。对于测试数据，我们在每个模态收集50条正常数据，在

目标域收集300条异常数据。

4.1.2. 性能可视化实验

如前文所述，MMD正则化和LDA-like正则化可以从不同角度消除分布畸变。为了验证两者均有消除分布畸变的能力，我们分别引入其中一个，设置另一个正则化参数为零，直观地观察样本的分布。训练数据被用来学习一个公共的字典。然后通过公共字典和OMP算法得到源域数据、目标域数据和异常数据的稀疏代码。我们通过PCA将稀疏代码可视化。图2显示了可视化的情况。

图2 (a) 显示了原始数据的分布。由于环境干扰，

Algorithm 2. RTDL.

Input: Source domain data matrix $X_s = [x_{s1}, x_{s2}, \dots, x_{sn}]$, target domain data matrix $X_t = [x_{t1}, x_{t2}, \dots, x_{tn}]$, parameters α , β_1 , and β_2

Begin RTDL algorithm**Step 1.** Initialization

Structure data matrix $X = [X_{s1}, X_{t1}, \dots, X_{sc}, X_{tc}, \dots, X_{sc}, X_{tc}]$, MMD matrix M by Eq. (4) and LDA-like matrix H by Eq. (8). D and S are initialized as random values, where X_{sc} and X_{tc} represent the data sets of C th mode of the source domain and target domain, respectively

Step 2. Off-line training

Obtain D and S iteratively

To fix S , update D using Eq. (14)

To fix D , update S using **Algorithm 1**

Step 3. Online testing

(i) Process monitoring

Get the RES limit by D and the OMP and KDE algorithms

Conduct process monitoring by comparing the RES statistic and RES limit

(ii) Mode classification

After the testing data is detected as normal data, conduct mode classification using Eq. (20)

End RTDL algorithm

表1 关于仿真数据的细节描述

Mode	Variable		
	t_1	t_2	(e_1, e_2, \dots, e_5)
Mode 1 in source domain	$N(5, 0.8)$	$N(15, 1.3)$	$3E(0.5)$
Mode 2 in source domain	$U(2, 3)$	$N(15, 1.7)$	$3E(0.5)$
Mode 1 in target domain	$N(5, 0.8)$	$N(15, 1.3)$	$0.1\bar{F}(5, 10)$
Mode 2 in target domain	$U(2, 3)$	$N(15, 1.7)$	$0.1\bar{F}(5, 10)$

域间数据出现了分布畸变，且异常的信息是不明显的。图2 (b)、(c) 显示，在单独使用其中一个正则化后，分布畸变被部分消除，异常信息通过字典学习后开始显示出来。图2 (d) 显示，源数据和目标数据完全混合，通过使用两种正则化的方法，异常信息很容易被识别。据此，实验结果验证了本文的动机（第2.3节）。

4.1.3. 参数敏感性分析实验

参数 α 是字典学习中一个重要的调整参数，它控制着 S 的稀疏程度。一般来说， α 可以通过观察 S 的稀疏程度来选择， S 的稀疏程度表示矩阵中非零元素所占的比例。如图3所示，当稀疏率（SPR）为20%~50%时，可以得到满意的过程监测结果；据此，对SPR的参数敏感性分析验证了所提方法的鲁棒性。

接下来，对修改后的 β_1 和 β_2 参数的过程监测性能进行分析。令 β_1 从 10^1 变化到 10^5 ，而 β_2 从10变化到70。错误报警率（FAR）和故障检测率（FDR）被认为是过程监

测的两个评价指标。结果显示在图4和图5中。可以看出，当 β_1 和 β_2 改变时，FAR总是低于3%，FDR总是大于80%。尽管 β_1 和 β_2 的变化范围很大，但过程监测的这种性能仍是满意的。当 β_1 和 β_2 的值过小时，学习的字典可以很好地重建训练数据。然而，训练数据中不相关的环境干扰也会被字典学习。当 β_1 和 β_2 的值过大时，目标函数只是减少域之间的分布差异，学习到的字典不能很好地重建过程数据，而且潜在的语义信息，如数据中的机理信息将被丢失。这两种情况均会降低字典捕捉过程数据中常见的潜在信息的能力，导致过程监测的性能不佳。为了给这些超参数选择一个合适的值，可以选择网格搜索这样的优化方法。

4.2. 性能比较实验

4.2.1. 数据集

由于RTDL在数值仿真中的性能是比较满意的，我们现在考虑其在现实工业场景中的性能。本实验准备了两个数据集。

(1) **CSTH**. CSTH过程是一个非线性的真实平台，已被广泛用作评估不同过程监测方法的基准[39]。CSTH的原理如图6所示。在CSTH过程中，有两个物理平衡：质量平衡和热平衡。冷水和热水同时流入水槽进行搅拌，同时被蒸汽加热[40]。假设这个过程有两种模态，其中模

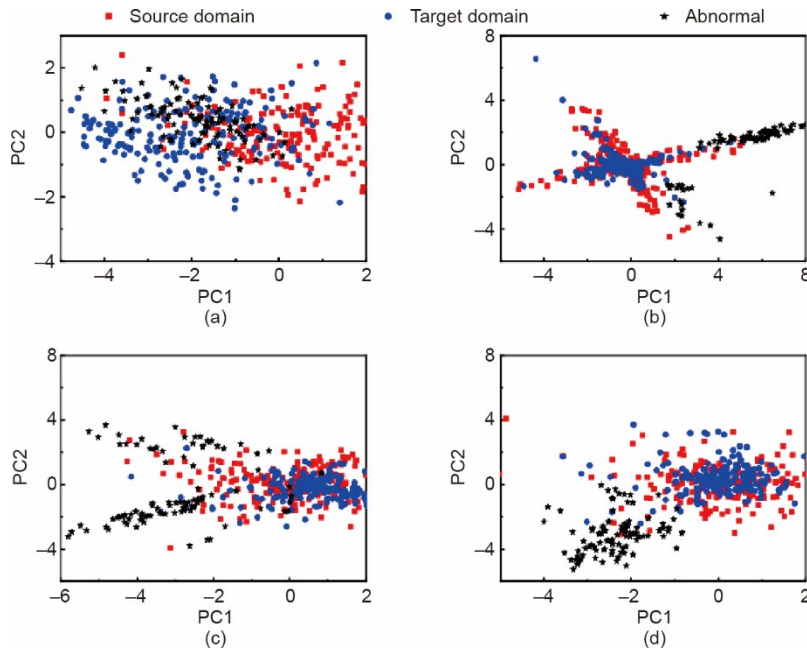


图2. 样本分布散点图。(a) 原始数据；(b) 通过带 LDA-like 正则化的字典学习得到的稀疏表示；(c) 通过带 MMD 正则化的字典学习得到的稀疏表示；(d) 通过同时带 LDA-like 正则化和 MMD 正则化的字典学习得到的稀疏表示。

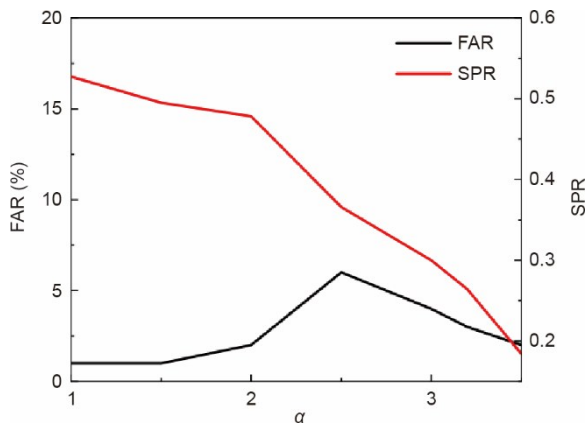


图3. 参数 α 对指标 FAR 和 SPR 的灵敏度分析。FAR: 错误报警率。

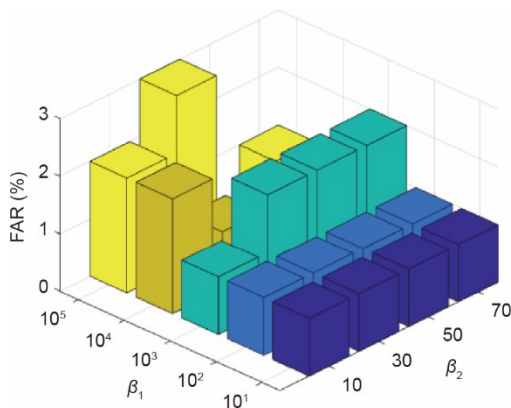


图4. 参数 β_1 和 β_2 对指标 FAR 的灵敏度分析。

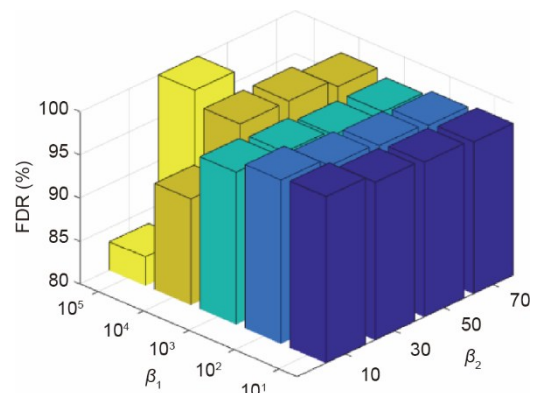


图5. 参数 β_1 和 β_2 对指标 FDR 的灵敏度分析。

态状态由液面设置、温度设置和热水阀位置设置决定。详细情况见表2。为了模拟环境干扰，我们在所有观测变量

中加入一个指数分布变量，以形成源域，在所有观测变量中加入一个 F 分布变量，以形成目标域。在观察到的流量变量上施加一个加性故障，以产生异常数据。我们在源域的每个模态收集了100个数据，在目标域的每个模态收集了30个正常数据，以形成训练数据。测试数据包括每个模态的50个数据和目标域的200个异常数据。

(2) **风机系统。** 风机系统的数据来自于北京的一家风力发电公司。从2011年1月1日至2011年11月11日，8台风力发电机每分钟采样一次，表3所示的15维数据被用于过程监测和模态识别。每台风机的数据中存在着不同的流形结构，可以认为是风机系统的运行模态。夏季的温度和风力与冬季不同，导致风力发电机组的工作温度和工作功率不同，从而导致数据分布畸变。风力涡轮机系统在不同季节会受到不同的环境干扰。我们假设冬季为源域，而夏季为

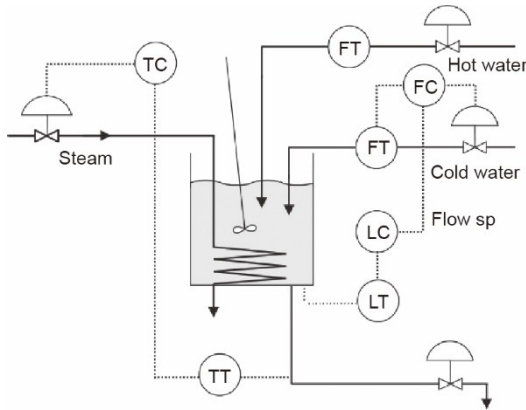


图6. CSTH原理图。TC: 温度控制器; FC: 流量控制器; LC: 液位控制器; TT: 温度传感器; FT: 流量传感器; LT: 液位传感器; sp: 设定值。摘自参考文献[40], 经Elsevier许可, ©2008。

表2 CSTH的传感器电信号设置参数

Mode	Liquid level sensor current (mA)	Temperature sensor current (mA)	Water valve position sensor current (mA)
Mode 1	12	10.5	5.5
Mode 2	12	10.5	5.0

目标域。详细说明见表4。训练数据由源域中每种模态的350个正常数据和目标域中每种模态的50个正常数据组成。每个模态的50个正常数据和目标域的300个异常数据构成测试数据。由于风机系统数据的尺寸差异很大, 所有的数据在实验前均进行了标准化处理。

4.2.2. 过程检测的比较实验

为了定量评估RTDL方法, 我们使用了另外两种新颖的字典学习方法和一种自适应监测方法进行比较实验。此外, 第2.3节中提到的两种数据处理策略也被用于对比。比较方法包括: label consistent K-singular value decompo-

表3 风机系统实验中使用的特征

Serial number	Feature	Serial number	Feature	Serial number	Feature
001	Active power	006	First phase current	007	Second phase current
008	Third phase current	012	Motor speed	026	Rotor speed
035	First phase voltage	036	Second phase voltage	038	Third phase voltage
038	Average wind speed	061	Gear shaft temperature	065	Gear box temperature
068	Engine room temperature	069	External ambient temperature	098	The stator temperature

表4 关于风机系统数据的详细描述

Domain	Model		Abnormal data
	Mode 1	Mode 2	
Source domain	Normal data in January 2011 in wind turbine 20	Normal data in January 2011 in wind turbine 70	No
Target domain	Normal data in July and August 2011 in wind turbine 20	Normal data in July and August 2011 in wind turbine 70	Abnormal data in July and August 2011 in wind turbine 70

sition (LC-KSVD) (S+T)、LC-KSVD (T)、Fisher discrimination dictionary learning (FDDL) (S+T)、FDDL (T)、moving window PCA (MWPCA) 和RTDL。LC-KSVD [41]、FDDL [31]和MWPCA [42]方法是用于过程监测的三种先进的方法。LC-KSVD和FDDL同时具有模态识别的能力, 而MWPCA是另一种用于过程监测的自适应方法。这里, LC-KSVD (S+T) 和FDDL (S+T) 分别指LC-KSVD方法和FDDL方法, 其直接使用所有源域训练数据和目标域训练数据作为输入数据, 不考虑域之间的不同特征。LC-KSVD (T) 和FDDL (T) 分别指LC-KSVD方法和FDDL方法, 其只使用目标域训练数据作为输入数据。MWPCA方法直接使用所有源域训练数据和目标域训练数据作为输入数据, 不考虑域之间的不同特征。RTDL将源域的训练数据和目标域的训练数据有区别地作为输入数据。为了公平, 字典的大小和其他参数被设定为相同。为了比较各模型的性能, 我们参考了两个指标: FAR 和FDR。

结果显示在图7和图8中。如图所示, 在这两个实际数据集中, 本方法在准确性方面均优于对比方法。有一些有趣的结果值得注意, 在CSTH中, LC-KSVD (S+T) 的FDR接近于LC-KSVD (T) 的FDR, 但是LC-KSVD (S+T) 的FAR明显大于LC-KSVD (T) 的FAR。这一结果与我们的观察相吻合, 即如果忽略了域的分布畸变, 模型可能更容易将域间差异信息与异常信息混淆, 导致过程监测结果更差。

4.2.3. 模态识别比较实验

本方法可以处理多模态数据, 当检测到数据为正常数据时, 可以进行模态识别。因此, 在本节中, 通过与对比

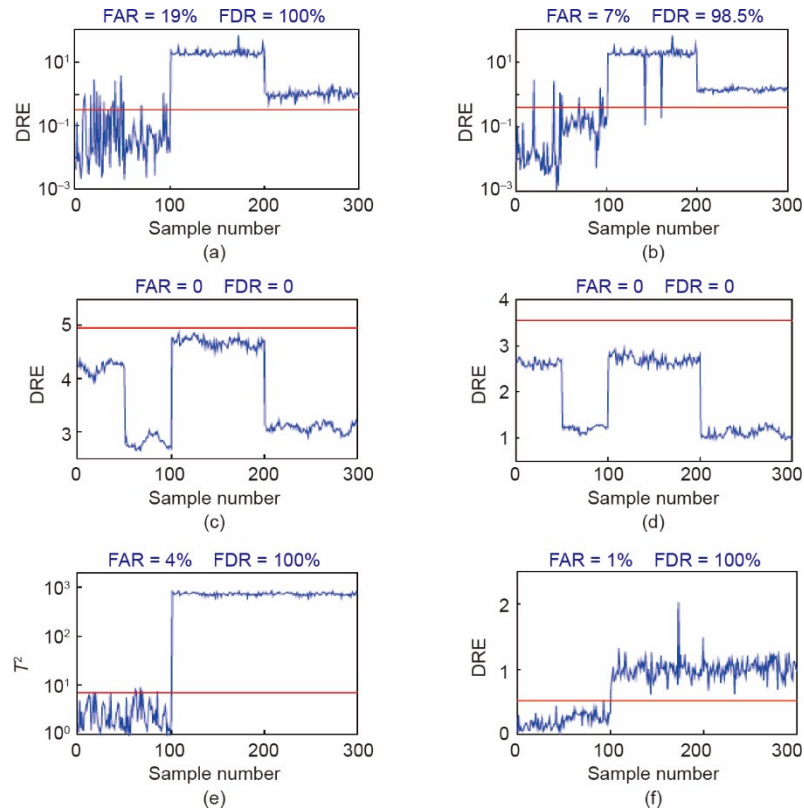


图7. 本方法与对比方法在CSTD上的监测结果。(a) LC-KSVD (S+T)的统计量DRE; (b) LC-KSVD (T)的统计量DRE; (c) FDDL (S+T)的统计量DRE; (d) FDDL (T)的统计量DRE; (e) MWPCA的统计量 T^2 ; (f) RTDL的统计量DRE。DRE: 字典重构误差。

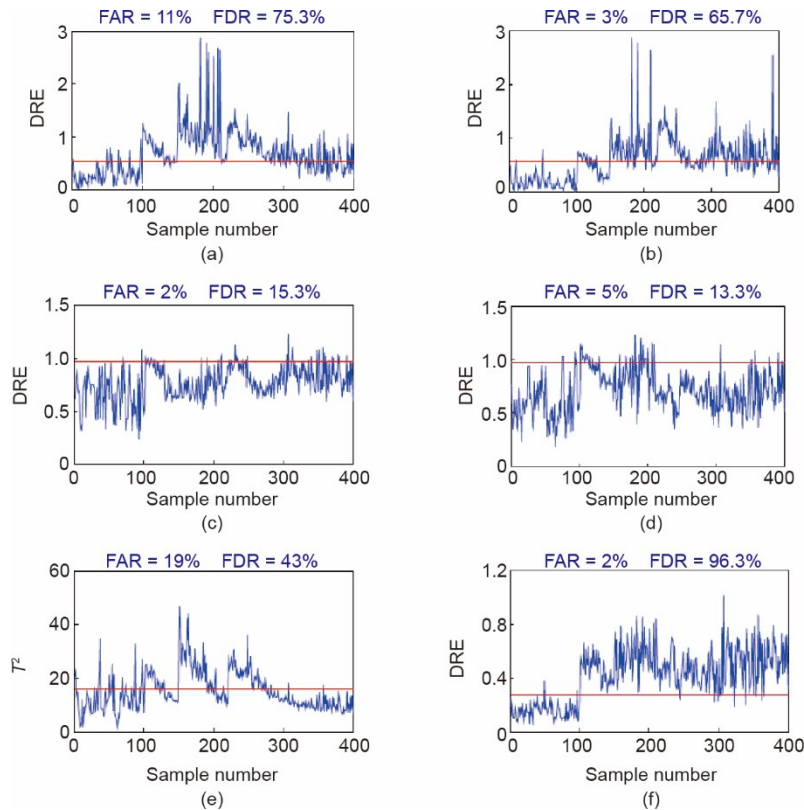


图8. 本方法与对比方法在风机系统上的监测结果。(a) LC-KSVD (S+T)的统计量DRE; (b) LC-KSVD (T)的统计量DRE; (c) FDDL (S+T)的统计量DRE; (d) FDDL (T)的统计量DRE; (e) MWPCA的统计量 T^2 ; (f) RTDL的统计量DRE。

表5 模态识别结果

Method	CSTH		Wind turbine system	
	Mode 1 accuracy	Mode 2 accuracy	Mode 1 accuracy	Mode 2 accuracy
LC-KSVD(S+T)	96	98	92	74
LC-KSVD(T)	0	64	76	84
FDDL(S+T)	0	100	18	88
FDDL(T)	0	100	48	42
RTDL	98	100	94	88

方法的比较来评估模态识别的有效性。对比方法的参数设置与过程检测比较实验一样。值得注意的是，MWPCA方法不能完成模态识别任务，所以该方法不用于本实验。为了比较模型的性能，我们参考了两个指标：模态1准确率和模态2准确率。结果显示在表5中。可以看出，本方法在模态识别中的表现优于对比方法的表现，这进一步验证了本方法的有效性。

5. 结论

由于工业过程经常受到多变的操作环境的影响，在线监测数据和历史训练数据并不总是遵循相同的分布。因此，基于历史训练数据的过程监测模型不能准确地执行监测在线流数据的任务。为此本文提出了一种RTDL方法。该方法是一个表示学习和域自适应迁移学习的协同框架。也就是说，首先使用字典学习方法，将原始数据投射到一个子空间，学习一个公共的字典来表示源域数据和目标域数据。在减少子空间中的域间分布距离和模内距离后，消除了环境干扰引起的分布畸变，从而提高了字典表示内部语义信息的能力，如机理信息。通过大量的实验，包括数值仿真、CSTH基准和真实的风机系统，证明了所提方法在域迁移问题上的优越性。因此，可以得出结论，本方法可以将知识从单一源域迁移到单一目标域。由于工业过程通常会遇到多种操作环境，未来的工作将集中在实现从多个源域到多个目标域的知识迁移。

致谢

本研究由国家自然科学基金(61988101)和国家重点研发计划(2018YFB1701100)项目资助。

Compliance with ethics guidelines

Chunhua Yang, Huiping Liang, Keke Huang, Yong-

gang Li, and Weihua Gui declare that they have no conflict of interest or financial conflicts to disclose.

References

- [1] Qian F, Zhong W, Du W. Fundamental theories and key technologies for smart and optimal manufacturing in the process industry. *Engineering* 2017; 3(2): 154–60.
- [2] Mao S, Wang B, Tang Y, Qian F. Opportunities and challenges of artificial intelligence for green manufacturing in the process industry. *Engineering* 2019; 5(6):995–1002.
- [3] Zhao C, Sun Y. Multispace total projection to latent structures and its application to online process monitoring. *IEEE Trans Control Syst Technol* 2014;22(3):868–83.
- [4] Liu Q, Qin SJ, Chai T. Multiblock concurrent PLS for decentralized monitoring of continuous annealing processes. *IEEE Trans Ind Electron* 2014; 61(11): 6429–37.
- [5] Jiang Q, Yan X, Huang B. Review and perspectives of data-driven distributed monitoring for industrial plant-wide processes. *Ind Eng Chem Res* 2019;58(29): 12899–912.
- [6] Shang C, You F. Data analytics and machine learning for smart process manufacturing: recent advances and perspectives in the big data era. *Engineering* 2019;5(6):1010–6.
- [7] Zhang W, Yang D, Wang H. Data-driven methods for predictive maintenance of industrial equipment: a survey. *IEEE Syst J* 2019;13(3):2213–27.
- [8] Zhou P, Zhang R, Xie J, Liu J, Wang H, Chai T. Data-driven monitoring and diagnosing of abnormal furnace conditions in blast furnace ironmaking: an integrated pca-ica method. *IEEE Trans Ind Electron* 2020;68(1):622–31.
- [9] Sankavaram C, Pattipati BR, Kodali A, Pattipati K, Azam M, Kumar S, et al. Model-based and data-driven prognosis of automotive and electronic systems. In: *Proceedings of 2009 IEEE International Conference on Automation Science and Engineering*; 2009 Aug 22–25; Bangalore, India. New York City: Curran Associates; 2009. p. 22–5.
- [10] Feng L, Zhao C. Fault description based attribute transfer for zero-sample industrial fault diagnosis. *IEEE Trans Ind Inform* 2020;17(3):1852–62.
- [11] Liang YC, Wang S, Li WD, Lu X. Data-driven anomaly diagnosis for machining processes. *Engineering* 2019;5(4):646–52.
- [12] Jiang Q, Yan X, Huang B. Deep discriminative representation learning for nonlinear process fault detection. *IEEE Trans Autom Sci Eng* 2020; 17(3): 1410–9.
- [13] Jiang Q, Yan S, Yan X, Yi H, Gao F. Data-driven two-dimensional deep correlated representation learning for nonlinear batch process monitoring. *IEEE Trans Ind Inform* 2020;16(4):2839–48.
- [14] Pan SJ, Yang Q. A survey on transfer learning. *IEEE Trans Knowl Data Eng* 2010;22(10):1345–59.
- [15] Chai Z, Zhao C. A fine-grained adversarial network method for cross-domain industrial fault diagnosis. *IEEE Trans Autom Sci Eng* 2020;17(3):1432–42.
- [16] Huang K, Wen H, Zhou C, Yang C, Gui W. Transfer dictionary learning method for cross-domain multimode process monitoring and fault isolation. *IEEE Trans Instrum Meas* 2020;69(11):8713–24.
- [17] Wang J, Zhao C. Mode-cloud data analytics based transfer learning for soft sensor of manufacturing industry with incremental learning ability. *Control Eng Pract* 2020;98:104392.
- [18] Hou R, Wang H, Xiao Y, Xu W. Incremental PCA based online model updating for multivariate process monitoring. In: *Proceedings of the 10th World*

- Congress on Intelligent Control and Automation; 2012 Jul 6–8; Beijing, China. New York City: Curran Associates; 2012.
- [19] Jiang LY, Xie L, Wang SQ, Wang N. Monitoring batch processes using multi-model discriminant partial least squares. In: Proceedings of 2005 International Conference on Machine Learning and Cybernetics; 2005 Aug 18–21; Guangzhou, China. New York City: Curran Associates; 2005.
- [20] Zhang Y, You D, Gao X, Katayama S. Online monitoring of welding status based on a DBN model during laser welding. *Engineering* 2019;5(4):671–8.
- [21] Zeng J, Gao C, Luo S, Li Q. Online process monitoring based on incremental LPP. In: Proceedings of Online process monitoring based on incremental LPP; 2011 Jul 22 – 24; Yantai, China. New York City: Curran Associates; 2011.
- [22] Ge Z, Song Z. Online batch process monitoring based on multi-model ICA–PCA method. In: Proceedings of the 2008 7th World Congress on Intelligent Control and Automation; 2008 Jun 25–27; Chongqing, China. New York City: Curran Associates; 2008.
- [23] Peng X, Tang Y, Du W, Qian F. Multimode process monitoring and fault detection: a sparse modeling and dictionary learning method. *IEEE Trans Ind Electron* 2017;64(6):4866–75.
- [24] Chen G, Xiong C, Corso JJ. Dictionary transfer for image denoising via domain adaptation. In: Proceedings of the 19th IEEE International Conference on Image Processing; 2012 Sep 30–Oct 3; Orlando, FL, USA. New York City: Curran Associates; 2013.
- [25] Zhang K, Yuan M, Xiong Y, Qu L. Common dictionary and domain-specific dictionary based cross-domain image classification. In: Proceedings of 2017 Chinese Automation Congress (CAC); 2017 Oct 20–22; Jinan, China. New York City: Curran Associates; 2017. p. 2824–9.
- [26] Jie N, Qiang Q, Chellappa R. Subspace interpolation via dictionary learning for unsupervised domain adaptation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2013 Jun 23–28; Portland, OR, USA. New York City: Curran Associates; 2013. p. 692–9.
- [27] Long M, Ding G, Wang J, Sun J, Guo Y, Yu PS. Transfer sparse coding for robust image representation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2013 Jun 23–28; Portland, OR, USA. New York City: Curran Associates; 2013. p. 407–14.
- [28] Huang K, Wen H, Ji H, Cen L, Chen X, Yang C. Nonlinear process monitoring using kernel dictionary learning with application to aluminum electrolysis process. *Control Eng Pract* 2019;89:94–102.
- [29] Gretton A, Borgwardt K, Rasch MJ, Scholkopf B, Smola AJ. A kernel method for the two-sample problem. 2008. arXiv: 0805.2368.
- [30] Ramirez I, Sprechmann P, Sapiro G. Classification and clustering via dictionary learning with structured incoherence and shared features. In: Proceedings of 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition; 2010 Jun 13–18; San Francisco, CA, USA. New York City: Curran Associates; 2010.
- [31] Meng Y, Lei Z, Feng X, Zhang D. Fisher discrimination dictionary learning for sparse representation. In: Proceedings of 2011 International Conference on Computer Vision; 2011 Nov 6–13; Barcelona, Spain. New York City: Curran Associates; 2011.
- [32] Shekhar S, Patel VM, Nguyen HV, Chellappa R. Generalized domain-adaptive dictionaries. In: Proceedings of the IEEE conference on Computer Vision and Pattern Recognition; 2013 Jun 23–28; Portland, OR, USA. New York City: Curran Associates; 2013. p. 361–68.
- [33] Huang K, Wu Y, Wen H, Liu Y, Yang C, Gui W. Distributed dictionary learning for high-dimensional process monitoring. *Control Eng Pract* 2020;98:104386.
- [34] Kokiopoulou E, Chen J, Saad Y. Trace optimization and eigenproblems in dimension reduction methods. *Numer Linear Algebra Appl* 2011; 18(3): 565–602.
- [35] Schölkopf B, Platt J, Hofmann T. Efficient sparse coding algorithms; Stanford: MIT Press; 2007.
- [36] Fletcher R. Practical methods of optimization. 2nd ed. New York City: Wiley; 1987.
- [37] Rebollo-Neira L, Lowe D. Optimized orthogonal matching pursuit approach. *IEEE Signal Process Lett* 2002;9(4):137–40.
- [38] Chen Q, Wynne R, Goulding P, Sandoz D. The application of principal component analysis and kernel density estimation to enhance process monitoring. *Control Eng Pract* 2000;8(5):531–43.
- [39] Huang K, Wu Y, Yang C, Peng G, Shen W. Structure dictionary learning-based multimode process monitoring and its application to aluminum electrolysis process. *IEEE Trans Autom Sci Eng* 2020;17(4):1989–2003.
- [40] Thornhill NF, Patwardhan SC, Shah SL. A continuous stirred tank heater simulation model with applications. *J Process Control* 2008, 8(3–4):347–60.
- [41] Jiang Z, Lin Z, Davis LS. Label consistent K-SVD: learning a discriminative dictionary for recognition. *IEEE Trans Pattern Anal Mach Intell* 2013;35(11): 2651–64.
- [42] Jeng JC. Adaptive process monitoring using efficient recursive PCA and moving window PCA algorithms. *J Taiwan Inst Chem Eng* 2010;41(4):475–81.