

基于动态时间规整和自适应顺序形态变换分类的故障诊断方法^①李 晗^② 萧德云^③

(清华大学自动化系 北京 100084)

摘 要 研究了将动态时间规整(DTW)用于故障诊断的技术,使具有相同趋势但时序上并不完美一致的两序列之间距离最小化,发挥其在序列匹配中的重要作用。在前期研究提出的一种采用欧式距离度量两序列间的匹配程度的基于自适应顺序形态变换分类的故障诊断方法的基础上,由采用欧式距离变为采用 DTW 匹配来度量信号间误差,提出了三种用于故障诊断的新的自适应顺序形态变换分类方法。通过连续搅拌釜式加热器故障诊断问题的研究验证了所提出的三种新方法的有效性,并对比了它们与前期研究提出的方法的故障诊断效果。

关键词 动态时间规整(DTW), 自适应顺序形态滤波, 故障诊断, 连续搅拌釜式加热器(CSTH)

0 引 言

各类模式匹配技术在故障诊断中有着广泛应用。基于模式分类的故障诊断方法的主要思想是将由过程数据变换得到的模式和已有表示过程各种运行或故障状态的标准模式一一进行匹配,定义匹配程度,选取匹配程度最高的标准模式实现未知模式的识别分类。但是由于发生时刻和故障幅值(例如阶跃型故障)不同^[1]以及仪器不精准将引起测量噪声、垂直方向平移、振幅缩放和时间轴不一致^[2]等原因,都会使得在同一时间内观察到的相同故障的现象不完全一样,这将直接影响各种定性趋势分析技术提取特征和降低后续模式匹配的效果。通常,采用信号模式直接匹配时,对信号进行标称化(零均值和单位方差化)或使用滤波可能会降低幅值差别和噪声扰动的影响^[3],但对于两个发生趋势相同但波形扭曲的信号来说,例如其中一个时间轴压缩或扩展,采用欧式距离作为模式间的不匹配程度描述会遇到困难^[4]。

文献[5]利用自适应顺序形态滤波器构建了一种分类方法,该方法可使检测信号在不同的监督信

号下利用自适应顺序形态变换得到波形形态不同的变换信号,通过变换信号和监督信号的匹配程度,完成对检测信号的识别和分类,本文称其为第 I 种基于自适应顺序形态变换分类的故障诊断法,简称为第 I 种分类。此方法采用欧式距离描述信号间误差来实现两个信号的匹配度量。而动态时间规整(dynamic time warping, DTW)技术是语音识别领域的一项重要模式匹配技术^[6,7],运用动态规划使两个模式中的重要特征排列一致,从而使总距离最小化,得到 DTW 距离。这突破了欧式距离中两个信号任意时刻的元素都要严格对应的限制,可使不同发生时刻的同一类故障进行更好的匹配。而第 I 种分类同 DTW 技术一样,也以整段信号为特征计算模式间的匹配程度,采用欧式距离描述信号间误差,变换实现了检测信号变换后有与监督信号重合的趋势,即实现两者具有相同特征的时标靠拢,使变换后较变换前提高了信号的匹配程度,可实现模式分类和故障诊断。本文结合 DTW 匹配和第 I 种分类,采用 DTW 距离分别描述自适应顺序形态变换后和变换中间过程的信号误差,提出了三种新分类方法,最后采用连续搅拌釜式加热器故障诊断问题验证了这三种新分类方法的有效性和较第 I 种分类的优势。

① 国家自然科学基金(60736026, 60904044)和清华大学信息科学与技术国家实验室(筹)学科交叉基金资助项目。

② 男,1986年生,博士生;研究方向:基于数据的故障诊断;E-mail: lihan07@mails.tsinghua.edu.cn

③ 通讯作者,E-mail: xiaody@mail.tsinghua.edu.cn

(收稿日期:2012-09-13)

1 DTW 技术基础

当两个需要进行匹配的时间序列趋势相同但时间长度不一致,或时间长度虽一致但重要特征有时间上的错位时,依然采用时标一一对应的方式求出两信号间的距离将偏大,可能得出两序列不同类的判断,这时常并不是所期望得到的结果。而 DTW 技术是处理这类问题的有效途径,它能在一定限制下,将序列元素错位对应,并最终确定所有元素最优错位对应的组合,以此实现两个序列的错位对应。在此过程中,一个错位对应将产生所对应元素的差距,称为局部距离,局部距离的组合将形成总体距离。如前所述,此组合可以有多种组合形式,而产生

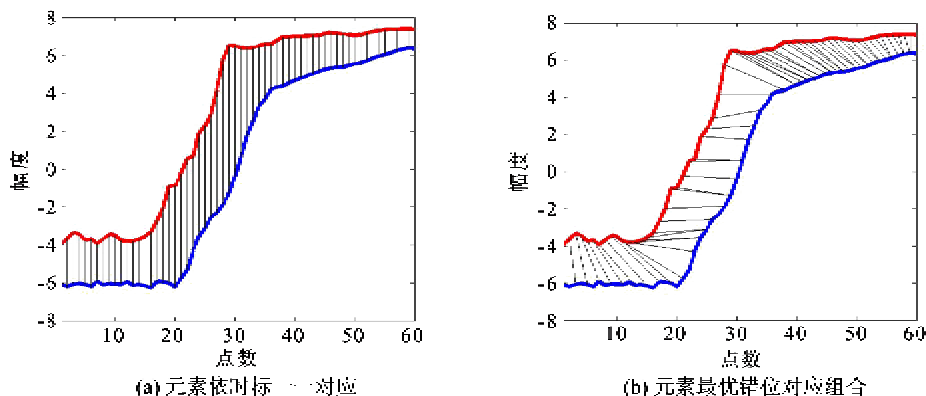


图1 两个具有相同趋势的序列匹配方式

号模式最佳匹配,而当前模式比历史数据库长度小很多,以此实现化工过程状态转变中的故障诊断和状态辨识。

2 基于 DTW 和自适应顺序形态变换分类的三种故障诊断法

第 I 种基于自适应顺序形态变换分类的故障诊断法^[5]在特征构建和特征分类两个阶段都需要采用序列匹配技术确定两序列间的差距,图 2 给出了方法的基本架构,以便能够直观地认识到序列匹配在整个架构中的位置、作用及与其它部分的联系。

图 2 中 x 是检测信号,以 $d_1, d_2, d_3, \dots, d_{N_p}$ 共 N_p 个模板信号(统一标记为 d)分别作为监督信号;得到的自适应顺序形态变换的变换信号 $y_1, y_2, y_3, \dots, y_{N_p}$ (统一标记为 y)以及 $y_{i1}, y_{i2}, \dots, y_{iN_p}$ (统一标记为 y_i),是自适应顺序形态变换在误差收敛过程中的中间过程变换信号。在该架构中,共有三种场合需要进行信号匹配和计算信号间误差,分别是:

总体距离最小的组合就是所有元素最优错位对应组合,图 1 说明了两序列元素依时间一一对应和最优错位对应的情况。可以看到,在错位对应情况下,可通过改变、增加和丢弃一些原有元素对应关系实现最佳匹配。Kassidas 等人^[1]用 DTW 技术来诊断田纳西伊斯曼流程(Tennessee Eastman process, TEP)的确定性故障。Colomer 等人^[8]先将信号表示成一系列连续特征(episodes),再对特征序列采用 DTW 匹配。Zhou 和 Wong^[9]指出,在基于相似度查询中,传统的以逐点方式将时间轴伸缩变换可能会引起较大的波形失真,因此将逐段方式时间轴伸缩和 DTW 结合起来形成逐段时间规整技术,但文献[2]指出,STW 的时间复杂度非常高。Srinivasan 和 Qian^[10]采用动态规划在历史数据库中在线搜寻和当前多元信

计算监督信号和变换信号间距离 e_1 得到分类依据;计算监督信号和中间过程变换信号间距离 e_2 且当该值收敛时完成自适应顺序形态变换;计算监督信号和中间过程变换信号间距离或局部距离 e_3 并将其用于计算参数更新量。其中 e_1 和 e_2 是信号整体匹配得到的距离, e_3 可以是信号整体匹配的距离也可以是部分信号段匹配的局部距离。

在第 I 种分类中, e_1 和 e_2 采用了欧式距离度量信号间误差。由于每个采样点都可能具有不同的顺序形态变换参数, e_3 采用了监督信号和中间过程变换信号各采样点间的直接差值,而非信号整体匹配的距离^[5]。以上三种情况中都采用了以欧式距离为基础的度量原则,但如果检测信号经过变换后得到的变换信号和监督信号相似的波形形态特征在时序上仍然有一定差距,此时若采用欧式距离度量会得到很大距离值,降低了匹配程度,造成错误诊断。

若借助 DTW 技术的柔性匹配功能,对具有相同波形或趋势的两信号时序上错位对应找到最小距

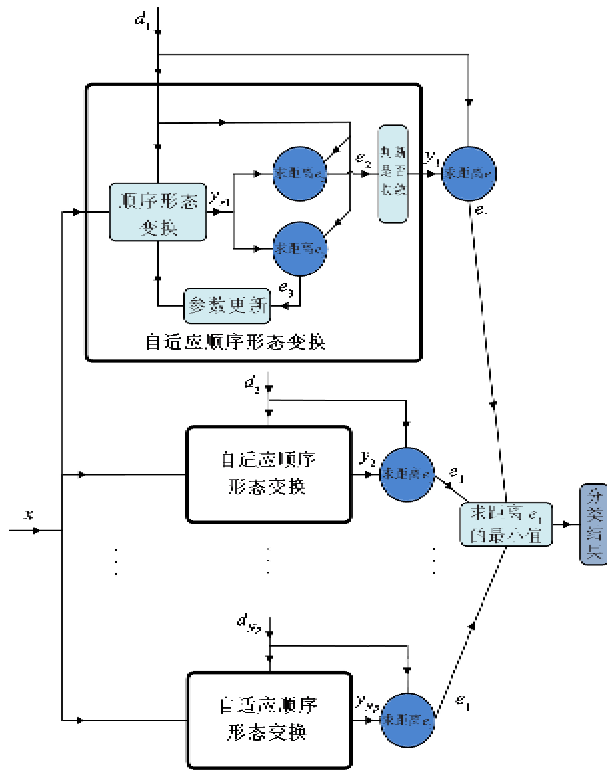


图2 基于自适应顺序形态变换分类的故障诊断法架构

离,在极大降低同类信号间匹配距离的同时,又能使不同类信号间匹配距离和采用欧式距离相比降低不大,这样的特点将帮助第I种分类减少误诊断。而另一方面,由于需要防止匹配时,时标不发生过度的非线性扩张和压缩,DTW一般对两序列时标错位对应关系有约束条件,因此对于存在相同特征但时标差距较大的两个信号,可以先采用自适应顺序形态变换,使得两信号波形往更加重合的趋势上进行变换,再采用DTW匹配,以此有效扩展DTW匹配可处理信号错位对应的范围。

综上所述,结合自适应顺序形态变换和DTW匹配的特点,可将第I种分类中序列匹配的三种场合由欧式距离度量改进为DTW距离度量,可有三种具体改进形式:(1)采用DTW距离度量变换信号 y_i 和监督信号 d 的误差 e_1 ,构建基于DTW距离的自适应顺序形态变换分类法;(2)采用DTW距离度量中间过程变换信号 y_i 和 d 的误差 e_2 (意味着在自适应顺序形态变换参数训练过程中以DTW距离收敛代替欧式距离收敛),构建基于DTW距离的收敛自适应顺序形态变换分类法;(3)采用DTW距离度量信号对应各采样点的局部距离 e_3 反馈参与各点的顺序形态变换参数更新,构建基于DTW距离反馈自适应顺序形态变换分类法。以上三种改进分类法

统称为第II种基于自适应顺序形态变换分类的故障诊断方法,简称为第II种分类,分别介绍如下:

(1)基于DTW距离的自适应顺序形态变换分类法

基于DTW距离的自适应顺序形态变换分类法将第I种分类中的 e_1 由采用欧式距离度量变为采用DTW距离度量, e_2 和 e_3 计算方法仍然采用欧式距离度量。变换信号取迭代过程中对应产生 e_2 最小值的那个中间过程变换信号。

(2)基于DTW距离的收敛自适应顺序形态变换分类法

如果采用DTW距离度量中间过程变换信号 y_i 和监督信号 d 的误差 e_2 ,这样 e_1 也将以DTW距离度量,取迭代过程中产生 e_2 最小值的那个中间过程变换信号作为变换信号,参数更新方法仍然同基于第I种分类,由此形成的分类方法称为基于DTW距离的收敛自适应形态变换分类法。

(3)基于DTW距离的反馈自适应顺序形态变换分类法

将第I种分类中 e_3 度量中间过程变换信号 y_i 和监督信号 d 各采样点间的差值变为与各采样点相联系的一段信号的DTW距离,作为下一次该采样点的参数更新信息,由此构建了基于DTW距离的反馈自适应顺序形态变换分类法。

由于参数更新采用了同最小均方算法相似的结构,所以只要参数更新步长的选择满足一定条件(小于某值)^[5],第II种分类三种形式的自适应顺序形态变换参数都将收敛或进入稳态振荡,因此变换信号的波形形态和误差 e_2 最终也将保持稳定或进入稳态振荡。当给出的监督信号与检测信号波形形态相似时,即属同类信号时,采用第II种分类得到的 e_2 稳态值相对于采用I种分类得到 e_2 稳态值的下降幅度一般要比监督信号和检测信号属不同类时得到的相应下降幅度更多。这样第II种分类就能在很大程度上解决当采用第I种分类时,监督信号和检测信号属不同类时得到的 e_2 稳态值比属于同类时得到的 e_2 稳态值更小的情况,也即避免根据更小的 e_2 稳态值将造成的误分类。因此采用第II种分类可能比采用第I种分类更能降低故障误诊断风险。但注意到采用第II种分类方法3的参数更新规则不以两个信号各采样点重合为目标,变换信号波形可能并不稳定,进而又可能影响信号分类的效果,因此第II种分类方法1和方法2结合了第I种分类和DTW各自的优势,将最有可能改进第I种分

类的故障诊断效果。

3 连续搅拌釜式加热器故障诊断

连续搅拌釜式加热器 (continuous stirred tank heater, CSTH) 是一个将冷热水混合, 通过加热管盘利用蒸汽加热并由长管将水导出的实验装置, 如图 3 所示。模型由 Thornhill 和 Patwardhan 等人给出^[11], 仿真模型提供于网址: <http://www.ps.ic.ac.uk/~nina/CSTHSimulation/index.htm> (2012 年 7 月 7 日链接有效)。CSTH 的输入是蒸汽和入流冷水的阀门调节信号, 采用串级控制; 输出是实际温度、水位和冷水入流量。设定的 CSTH 各类故障及用于生成模板信号的故障训练数据参数列于表 1。

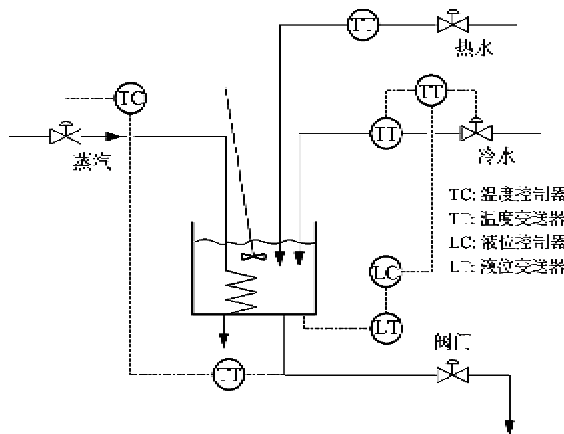


图 3 连续搅拌釜式加热器结构图

代表系统故障状态的特征信号产生方法如下: 选取三个输出量, 即液位、入流冷水速率和温度在一段时间内的采样信号经过小波去噪、减去首值然后首尾相连组成长信号, 再对长信号进行标称化, 形成特征信号。设置 Simulink 仿真时间 101s, 得到每个输出量信号长度超过 1000, 组合成的特征信号长度大于 3000, 进一步进行间隔采样, 构成长度 60 的特征信号。表 2 列出了 18 个测试故障的生成参数, 其中前 12 个对应不同的故障类型, 后 6 个和前 12 个某些故障是同类故障, 但生成参数不一致, 有一些更接近训练数据生成参数, 有一些则更远离训练数据生成参数, 这样设计的目的是为了对比第 I 种分类和第 II 种分类三种方法的故障诊断效果。得到的诊断结果如表 3 所示, 统计结果列于表 4。

表 1 CSTH 训练数据数据产生状态参数

故障标号	故障描述	训练数据参数
N	正常运行	-
LSPS	液位变送器正阶跃故障	跳变值 +2, 发生时刻 51s
LSPR	液位变送器正斜坡故障	斜率 +0.02, 发生时刻 51s
FSNS	入流冷水流量变送器负阶跃故障	跳变值 -2, 发生时刻 51s
FSPR	入流冷水流量变送器正斜坡故障	斜率 +0.02, 发生时刻 51s
LAPS	液位调节阀正阶跃故障	跳变值 +2, 发生时刻 51s
LANR	液位调节阀负斜坡故障	斜率 -0.02, 发生时刻 51s
FANS	入流冷水调节阀负阶跃故障	跳变值 -2, 发生时刻 51s
FANR	入流冷水调节阀负斜坡故障	斜率 -0.02, 发生时刻 51s
TSNS	温度变送器负阶跃故障	跳变值 -2, 发生时刻 51s
TSPF	温度变送器正斜坡故障	斜率 +0.02, 发生时刻 51s
TAPS	温度调节器正阶跃故障	跳变值 +2, 发生时刻 51s
TANR	温度调节器负斜坡故障	斜率 -0.02, 发生时刻 51s

表 2 CSTH 测试故障数据产生状态参数

测试故障标号	测试故障种类	测试数据参数
1	LSPS	跳变值 +1.2, 发生时刻 20s
2	LSPR	斜率 +0.01, 发生时刻 10s
3	FSNS	跳变值 -0.7, 发生时刻 80s
4	FSPR	斜率 +0.005, 发生时刻 30s
5	LAPS	跳变值 +4, 发生时刻 5s
6	LANR	斜率 -0.02, 发生时刻 95s
7	FANS	跳变值 -1, 发生时刻 65s
8	FANR	斜率 -0.01, 发生时刻 1s
9	TSNS	跳变值 -0.1, 发生时刻 92s
10	TSPF	斜率 +0.07, 发生时刻 5s
11	TAPS	跳变值 +1.3, 发生时刻 20s
12	TANR	斜率 -0.005, 发生时刻 80s
13	FANS	跳变值 -0.5, 发生时刻 65s
14	TAPS	跳变值 +1.3, 发生时刻 13s
15	FSNS	跳变值 -0.7, 发生时刻 60s
16	TANR	斜率 -0.005, 发生时刻 68s
17	FSNS	跳变值 -0.85, 发生时刻 60s
18	TANR	斜率 -0.005, 发生时刻 40s

表3 各种方法对 Csth 故障的诊断结果(采用具有 Itakura 局部约束的 DTW,√表示正确诊断)

故障编号 采用方法	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
E				√						√	√			√				
D				√	√		√			√	√			√	√		√	
EA		√		√	√		√			√	√		√	√				
DA		√		√	√		√			√			√				√	
DC	√	√		√	√		√			√			√	√	√		√	√
DF				√	√					√					√		√	

表4 各种方法对 Csth18 种故障正确诊断的统计

采用的方法	结合 DTW 匹配 具有 Itakura 局部约束
E	4/18
D	8/18
EA	8/18
DA	7/18
DC	11/18
DF	5/18

从表3中可以看到,基于DTW距离收敛自适应顺序形态变换分类(DC)的正确诊断故障次数最多;对于有些与训练数据相比,故障幅度减小的测试故障,DC更易于诊断出,例如故障15、17和18,而第I种分类(EA)没有正确诊断出,表明依据DTW距离收敛可避免依据欧式距离收敛带来的过度训练问题,这是对EA方法的改进。表3还显示,直接采用DTW匹配(D)、基于DTW距离自适应顺序形态变换分类(DA)在总体上和EA得到的正确诊断次数都相同,两者都不比直接采用DTW匹配效果差,且都明显好于直接采用计算欧式距离匹配(E),而DTW反馈自适应顺序形态变换分类(DF)则出现诊断效果下降,即由于此种方法可能会造成输出信号波形不稳定,进而影响故障诊断效果。

4 结 论

本文分析了文献[5]提出的第I种基于自适应顺序形态变换分类的故障诊断方法(简称第I种分类)基本架构中的信号间误差度量三种场合,结合动态时间规整(DTW)模式间柔性匹配的特点,提出了三种新分类方法,即方法1——基于DTW距离的自适应顺序形态变换分类法,方法2——基于DTW距离的收敛自适应顺序形态变换分类法,方法3——基于DTW距离的反馈自适应顺序形态变换分类法,统称第II种基于自适应顺序形态变换分类

的故障诊断方法(简称第II种分类)。方法1采用DTW距离度量自适应顺序形态变换信号与监督信号间的误差;方法2采用DTW距离度量变换中间过程变换信号与监督信号间的误差,并以该距离收敛为变换结束的判断;方法3将中间过程变换信号与监督信号的误差以DTW距离度量并参与参数更新过程。

采用连续搅拌釜式加热器故障诊断仿真分析,验证了第II种分类三种方法进行故障诊断的有效性,其中第II种分类方法2可以使检测信号根据监督信号发生相应的波形改变得得到变换信号,又能在避免第I种分类由于要求欧式距离收敛反而可能产生检测信号波形过度改变的现象,这一性质是自适应顺序形态变换和DTW柔性匹配相结合的优势,因此第II种分类方法2能比第I种分类在一定程度上减少故障误诊断。第II种分类方法1和第I种分类相比正确诊断故障的效果相当;第II种分类方法3计算量较大,且变换信号波形不稳定,误差可能处于振荡状态,和第I种分类相比,诊断效果可能会下降。

参考文献

- [1] Kassidas A K, Taylor P A, MacGregor J F. Off-line diagnosis of deterministic faults in continuous dynamic multivariable processes using speech recognition methods. *Journal of Process Control*, 1998, 8(5-6):381-393
- [2] Gameroa F I, Meléndez J, Colomera J. QSSI: A new similarity index for qualitative time series application to classify voltage sags. *Applied Artificial Intelligence*, 2011, 25(2): 141-162
- [3] Akbaryan F, Bishnoi P R. Fault diagnosis of multivariate systems using pattern recognition and multisensor data analysis technique. *Computers & Chemical Engineering*, 2001, 25(9-10):1313-1339
- [4] Perng C S, Wang H, Zhang S R, et al. Landmarks: a new model for similarity-based pattern querying in time

- series databases. In: Proceedings of 16th International conference on data engineering, San Diego, USA, 2000. 33-42
- [5] Li H, Xiao D Y. Fault diagnosis using pattern classification based on one-dimensional adaptive rank-order morphological filter. *Journal of Process Control*, 2012, 22 (2): 436-449
- [6] Itakura F. Minimum prediction residual principle applied to speech recognition. *IEEE Trans Acoustics, Speech, and Signal Processing*, 1975, 23(1):67-72
- [7] Sakoe H, Chiba S. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Trans Acoustics, Speech and Signal Processing*, 1978, 26(1): 43-49
- [8] Colomer J, Melendez J, Gamero F I. Qualitative representation of process trends for situation assessment based on cases. In: Proceedings of 15th IFAC World Congress, Barcelona, Spain, 2002
- [9] Zhou M, Wong M H. A segment-wise time warping method for time scaling searching. *Information Sciences*, 2005, 173(1-3): 227-254
- [10] Srinivasan R, Qian M S. Online fault diagnosis and state identification during process transitions using dynamic locus analysis. *Chemical Engineering Science*, 61 (18): 6109-6132
- [11] Thornhill N F, Patwardhan S C, Shah S L. A continuous stirred tank heater simulation model with applications. *Journal of Process Control*, 2008, 18(3-4):347-360

Fault diagnosis based on dynamic time warping and classification of adaptive rank-order morphological transform

Li Han, Xiao Deyun

(Department of Automation, Tsinghua University, Beijing 100084)

Abstract

The techniques of applying dynamic time warping (DTW) to fault diagnosis were studied to raise DTW's important role in sequence matching by minimizing the distance between two sequences which display the same trends but do not perfectly align with each other. Based on the fault diagnosis method proposed in the earlier stage of the study that uses the classification of adaptive rank-order morphological transform and the Euclidean distance for measuring the matching degree between two sequences, the DTW matching was substituted for the Euclidean distance metric to measure the matching degree, and three new classification approaches for fault diagnosis were presented. The effectiveness of the three methods in diagnosis of faults was verified by the test of a continuous stirred tank heater, and their performances were compared with that of method proposed in the earlier study.

Key words: dynamic time warping (DTW), adaptive rank-order morphological filter, fault diagnosis, continuous stirred tank heater (CSTH)