

密级:_____



中国科学院大学
University of Chinese Academy of Sciences

硕士学位论文

基于动态特性与稀疏表示的故障诊断方法研究

作者姓名: 武利军

指导教师: 焦建彬 教授 中国科学院大学

学位类别: 工学硕士

学科专业: 计算机应用技术

研究所: 中国科学院大学电子电气与通信工程学院

二零一三年五月

Fault Diagnosis Based on Dynamic Characteristics
And Sparse Representation

By
Lijun Wu

A Thesis Submitted to
The University of Chinese Academy of Sciences
In partial fulfillment of the requirement
For the degree of
Master of Computer Application

School of Electronic, Electrical and Communication Engineering

May, 2013

中国科学院大学直属院系
研究生学位论文原创性声明

本人郑重声明：所呈交的学位论文是本人在导师的指导下独立进行研究工作所取得的成果。尽我所知，除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的研究成果。对论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明或致谢。

作者签名：

日 期：

中国科学院大学直属院系
研究生学位论文原创性声明

本人完全了解并同意遵守中国科学院有关保存和使用学位论文的规定，即中国科学院有权保留送交学位论文的副本，允许该论文被查阅，可以公布该论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存、汇编本学位论文。

涉密的学位论文在解密后适用本声明。

作者签名：

日 期：

导师签名：

日 期：

摘 要

近年来，随着现代科学技术的发展和进步，现代化工生产过程日趋复杂化、大型化。任何系统内某一点故障的发生，如若不能及时发现处理，都会带来巨大的经济损失和人员伤亡。因此，准确、及时的故障诊断已经成为保证产品质量和工厂稳定运行的关键因素。智能化技术的发展，促成了生产过程中的大量实时数据被采集保存起来。这些数据隐含了大量信息却并没有被充分利用。为此本文依托于这些数据，采用模式识别的思想提出了一种新的故障诊断方法，以提高生产系统的安全性。论文主要研究的是基于动态特性与稀疏表示的故障方法，主要工作有以下几个方面：

(1) 采用基于稀疏表示的故障诊断方法。目前多变量统计分析方法在进行故障诊断时，将故障检测与确定故障类型作为两个独立的问题，采用不同的方法分步进行解决，相对比较繁琐复杂。模式识别方法中用于故障诊断的分类器大部分是针对于两类问题进行设计的，因此在对多类故障进行分类时需要训练多个分类器加以组合才能确定故障类型。针对以上问题，本文提出了一种基于稀疏表示的故障诊断方法。

(2) 引入了动态特性信息作为新特征改善诊断效果。当前各种故障诊断方法都是假设每维数据在时间轴上是独立的，然而这种假设在大多数的工业过程中都是不现实的，测量变量并不是序列无关的，具有一定的动态特性。而且不同的故障产生时，同一变量所具有的动态特性也是不一样的。因此，故障数据的动态特性有效提取和表示是提高故障诊断效果的有效途径之一。同时在模式识别领域，有效的特征表示是成功分类的关键因素。本文基于以上考虑，引入了故障动态特性作为分类特征的补充。

(3) 通过字典优化提高诊断水平。稀疏表示的本质就是利用过完备字典中的基元对测试样本进行线性组合。过完备字典的好坏直接决定了稀疏表示的效果，以至于决定了最后的分类效果，因此构建一个优异的冗余字典是非常重要的。本文采用 Feature-sign 算法和 language dual 算法对字典进行迭代优化提高诊断效果。

关键词：故障诊断，动态特性，稀疏表示，字典学习

Abstract

Lijun Wu(Computer Application)

Directed by: **Jianbin Jiao**(Professor)

In recent years, with the development of modern science and technology, the modern chemical production processes become more complex and more large-scale. Once an accident happens in a complex system, it can cause great property damage and casualties. Therefore, a correct and real-time fault diagnosis method has become the key factor to guarantee the high quality of the product and the stable operation of the factory. A large number of real-time monitoring data are collected and stored due to the development of intelligent technology development, but these data are not effectively applied to fault diagnosis. Consequently, relying on these data, a new fault diagnosis method based on pattern recognition is proposed in to improve the security of the production system. This thesis mainly makes a study about fault diagnosis based on dynamic characteristic and sparse representation. The main research contents are as follows:

(1) Fault diagnosis based on sparse representation. Multivariate statistical analysis methods treating fault detection and fault diagnosis separately are cumbersome. Meanwhile, because the classifiers currently used for fault diagnosis are usually designed for two classification problems, a large number of classifiers need to be trained and combined to identify the fault type when the number of fault is large. To solve the above problems, this thesis proposes a new fault diagnosis method based on sparse representation.

(2) Extracting dynamic characteristics information as new features for sparse representation classification. Now all kinds of fault diagnosis methods usually assume that each dimension of data on the time axis is independent. However, this assumption is unrealistic in most industrial processes. The measured variables have dynamic characteristics, which are different from the fault type. A good dynamic characteristics extracting way is an effective way to improve fault diagnosis. Moreover, a good

feature representation is very important for classification. In view of these, we propose a simple fault dynamic characteristic extraction method for fault diagnosis.

(3) Dictionary learning for improving fault diagnosis. The essence of the sparse representation is that the test sample is linearly combined using atoms of the overcomplete dictionary and classification results depend on the quality of the dictionary. Therefore, a good overcomplete dictionary is very important for sparse representation. The algorithms feature-sign and language dual are used for dictionary learning to improve fault diagnosis in this thesis.

KEY WORDS: Fault Diagnosis, Dynamic Characteristics, Sparse Representation, Dictionary Learning

目 录

摘 要	I
Abstract.....	II
目 录	IV
图目录	VI
表目录	VII
第一章 绪论	1
1.1 课题研究背景及意义.....	1
1.2 故障诊断研究发展及方法分类	2
1.2.1 定性分析的方法.....	3
1.2.2 定量分析的方法.....	5
1.3 故障诊断国内外研究现状及研究难点.....	6
1.4 本文研究内容	8
1.5 本文的组织结构	9
第二章 压缩感知和信号的稀疏表示理论	11
2.1 压缩感知的提出与内容.....	11
2.2 稀疏表示基本理论	12
2.3 稀疏表示的实现算法.....	14
2.3.1 全局优化算法.....	14
2.3.2 贪婪算法	15
2.4 稀疏表示的相关应用.....	17
2.5 本章小结.....	17
第三章 TEP 过程介绍	19
3.1 TEP 过程概述.....	19
3.2 TEP 过程变量和故障描述.....	20
3.3 实验仿真数据描述	22
3.4 本章小结.....	24
第四章 基于动态特性与稀疏表示的故障诊断	25
4.1 基于稀疏表示的故障诊断方法	25
4.1.1 故障样本字典的构造	25
4.1.2 基于稀疏表示的故障诊断	26
4.2 基于稀疏表示的故障诊断方法实验.....	29
4.3 基于动态特性的故障表示方法	34

4.4 字典的学习优化	38
4.5 本章小结	42
总结与展望	43
参考文献	45
作者简历与论文发表	51
附录 1: 主要符号对照表	53
致 谢	55

图目录

图 1-1 开县井喷事故和辽宁大连中石油输油管道爆炸	2
图 1-2 故障诊断方法分类示意图	3
图 2-1 稀疏表示模型	13
图 3-1 TEP 过程工艺流程图	19
图 3-2 反应器压力	23
图 3-3 反应器温度	24
图 4-1 基于稀疏表示的故障诊断方法研究框架	26
图 4-2 故障样本数据集内部结构示意图	26
图 4-3 测试样本利用故障样本数据集稀疏表示示意图	27
图 4-4 实验 1 中若干测试样本的重构系数和重构误差分布图	32
图 4-5 实验 2 中若干测试样本的重构系数和重构误差分布图	33
图 4-6 不同 K 值下实验 1 诊断效果图	34
图 4-7 不同 K 值下实验 2 诊断效果图	34
图 4-8 故障 4、9 和 11 训练集第 51 个观测变量随时间上变化情况	35
图 4-9 故障 4、9 和 11 训练集第 51 个观测变量方差随时间上变化情况 ..	36
图 4-10 故障 4、9 和 11 训练集第 51 个观测变量均值随时间上变化情况 ..	37

表目录

表 3-1 TEP 过程操作模式参数.....	20
表 3-2 连续测量变量	21
表 3-3 成分变量.....	22
表 3-4 故障描述.....	23
表 4-1 三组实验描述	29
表 4-2 实验 1 实验混淆矩阵	29
表 4-3 实验 2 实验混淆矩阵	29
表 4-4 20 类故障诊断效果比较.....	30
表 4-5 SRC 方法与 DSRC 实验效果对比	38
表 4-6 SRC、SRC_L、DSRC 和 DSRC_L 实验效果对比	41

第一章 绪论

近年来,随着现代科学技术的飞速发展,化工生产过程日趋复杂化、大型化,自动化装置的规模越来越大,复杂度越来越高,各个系统之间的内部联系越来越密切。任何系统内某一点故障的发生,如若不能及时发现处理,都将引发整个生产系统发生故障,最终导致生产过程停止,可谓“牵一发而动全身”。这些故障的发生不仅带来巨大的经济损失,更严重者会造成环境污染,危及人身安全。因此,作为提高系统可靠性和安全性的重要途径,故障诊断越来越被受到重视,成为当今社会的研究热点问题^[1-3]。

1.1 课题研究背景及意义

作为国民经济发展中的重要支柱产业,化工工业在工业总产值和工业增加值中占着很大比重,它为经济的发展和社会的进步做出了重大贡献。然而,化工过程具有生产工艺复杂,产品危险性高和生产环境恶劣等特点。生产过程中如操作不当,疏于检测或因不可抗拒的自然因素,都会造成重大事故。危险化学事故具有突发性、复杂性、激变性、群体性的特点,不仅能导致生产中断、爆炸、泄露毒气、火灾、中毒窒息、灼伤等危险,还有可能引起多米诺式灾害链的发生,导致重大的人员伤亡事故。

据有关材料证明,美国石油化工业每年由于系统故障引起的损失在 100 亿元至 200 亿美元之间,设备维修需要花费 100 亿美元。我国冶金业每年的设备维修费用也达到就达到了 250 亿元。除了巨额的经济损失外,危险化学品生产安全形势日益严峻,爆炸、火灾、中毒等各类安全事故频频发生,人员伤亡总数逐年上升。2003 年 12 月 23 日 21 时 55 分,四川石油管理局川东钻探公司川钻 12 队对气井起钻时,突然发生井喷,来势特别猛烈,富含硫化氢的气体从钻具水眼喷涌达 30 米高程,硫化氢浓度达到 100ppm 以上,失控的有毒气体随空气迅速扩散,导致在短时间内发生大面积灾害,人民群众的生命财产遭受了巨大损失。据统计,井喷事故发生后,离气井较近的 4 个乡镇,30 个村,9.3 万余人受灾,6.5 万余人被迫疏散转移,累计门诊治疗 27011 人,住院治疗 2142 人,243 位无辜人员

遇难，直接经济损失达 8200 余万元。2010 年 1 月 7 日 17 时 30 分左右，中国石油天然气集团公司兰州石化公司 303 厂 316 烃类罐区一裂解碳四储罐阀门处突然发生泄漏，现场可燃气体浓度达到极限，在当班操作人员紧急处理时发生爆炸，爆炸事故造成了 6 人遇难，1 人重伤，5 人轻伤。同年 7 月 16 日晚间 18 时左右，大连新港附近中石油一条输油管道起火爆炸。经过 2000 多名消防官兵彻夜奋斗，截至 17 日上午，火势才得以扑灭。事故造成大连附近海域至少 50 平方公里海面被原油污染。以上事故的发生给我们以沉痛的教训，时刻提醒我们安全生产至关重要。



图 1-1 开县井喷事故和辽宁大连中石油输油管道爆炸

因此，针对现代生产具有规模大、复杂性高、变量多，并在闭环控制下运行种种挑战，设计出具有实时性、准确性故障诊断方法是意义重大的。这不仅可以减少停产时间，增加系统运行的安全性，减少制造成本，还可以使企业避免人员和财产的巨大损失，给企业带来可观的经济效益。

本论文受到国家重大基础研究计划“973”项目课题资助。课题信息：“事故致灾过程和事故致因理论”（课题编号：2011CB706901）。2011.01-2015.12，在研；项目研究重大危险化学品事故数据源及其设备完整性理论。其科学目标是获取有效性、完整性的事故数据源，为科学研究和事故应急提供平台基础。同时通过数据挖掘与修补技术，探索其与设备完整性的科学联系，为化工机械和监测设备提供研究开发的科学依据。

1.2 故障诊断研究发展及方法分类

故障诊断指的是通过利用被诊断系统中的一定数量测量设备观测到的数据信、被诊断系统动力学模型、系统结构知识等各种已有知识以及过程异常变化的征兆与过程系统故障之间的内在联系，对系统的运行状态进行分析和判断，查明

故障发生的时间、位置和故障模式^[4]。

故障诊断技术涉及控制理论、模式识别、计算机科学、统计学、最优化理论等多个学科。整个故障诊断系统包括：故障的检测、故障隔离与诊断和故障的评价决策。目前故障诊断研究主要集中于故障检测与诊断两方面。1990年，国际故障诊断领域的理论权威 P.M.Frank 教授认为所有的故障诊断方法可以分为三类：基于解析模型的方法、基于信号处理的方法和基于知识的方法^[5]。2003年，Venkatasubramania 将故障诊断方法分为基于定量模型的方法、基于定性模型的方法和基于过程历史数据的方法^[6-7]。然而，近年来随着故障诊断的理论研究的不断深入，各类新颖方法的不断出现，使得原来的分类方法已经不再适用。2009年清华大学自动化系周东华教授提出了一种全新的分类框架，将故障诊断方法整体分为两大类：定性分析的方法和定量分析的方法，其中定量分析的方法又分为基于解析模型的方法和数据驱动的方法^[8]，如图 1-2 所示。

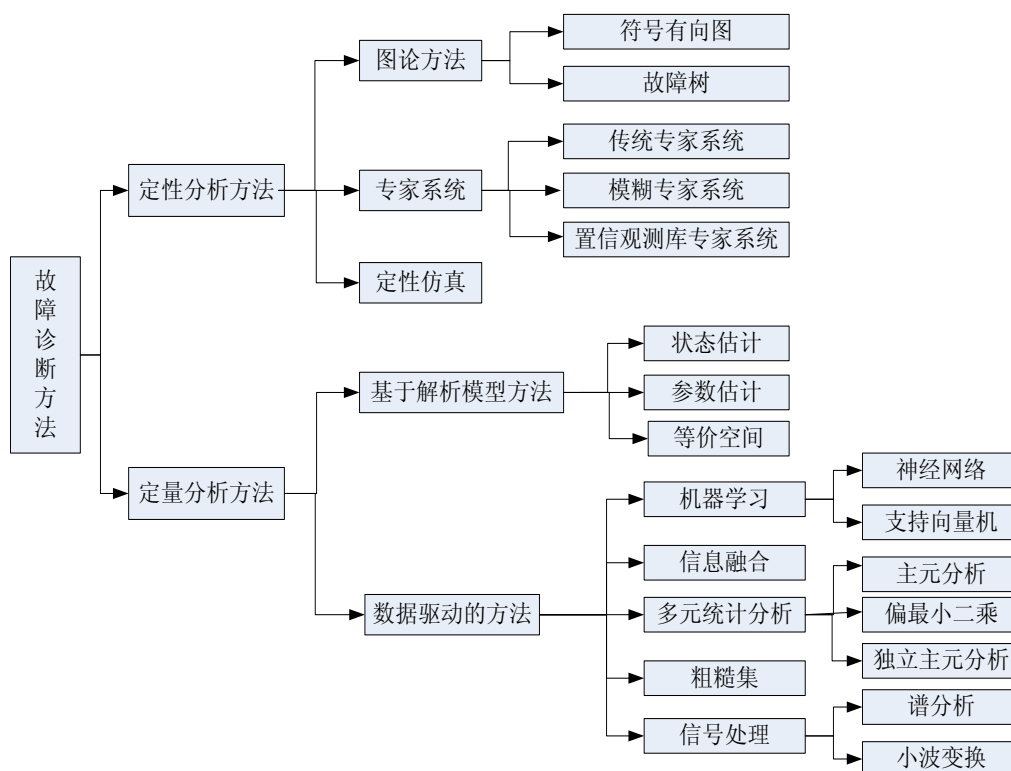


图 1-2 故障诊断方法分类示意图

1.2.1 定性分析的方法

所谓定性的故障诊断方法是采用定性的方法对系统结构和功能进行描述，对系统行为进行预测推理，通过和实际系统行为进行对比，检测系统是否发生故障。

常见的定性分析方法有图论方法、专家系统和定性仿真等。

■ 图论方法

有向图（Signed Directed Graph, SDG）方法和故障树方法是常用的两种基于图论的故障诊断方法。在 SDG 中，节点代表事件或者变量，变量之间的因果关系从原因节点指向结果节点的有方向的边表示。系统正常情况下，SDG 中的节点处于正常值范围内，一旦发生故障相应的故障节点将会偏离正常值并发生报警。根据图中节点之间的因果关系，并结合一定的搜索策略就可以分析出故障发生的原因及其发展的演变过程^[9]。

基于故障树的诊断方法最早是由 Caceres^[10]等人提出的，故障树是基于图论方法的一种具有语义的逻辑图。该类方法采用由果到因的分析过程，根据系统的故障现象出发，依据逻辑图逐级进行推理分析，最终确定故障发生的原因、影响程度和发生概率。

基于图论的方法建模简单，结果容易理解，应用范围广，但是当系统比较复杂时，这类方法的搜索过程会变得非常复杂，且准确率不高。

■ 专家系统

专家系统（Expert System）主要由知识库、推理机、综合数据库、人机接口及解释模块等部分构成，其中知识库和推理机是专家系统核心。基于专家系统的故障诊断方法就是利用领域专家在长期实践中积累起来的经验建立知识库，并设计出一套计算机程序控制模拟人类专家的推理和决策过程进行故障诊断^[11-12]。

基于专家系统的故障诊断方法利用专家丰富的经验知识，具有广泛的应用，但是该方法知识的获取比较困难，诊断的准确程度依赖于知识库中专家经验的丰富程度和知识水平的高低，推理规则较多时推理效率低下。

■ 定性仿真

定性仿真（Qualitative Simulation, QS）是以非数字手段处理信息输入、建模、行为分析和结果输出等仿真环节，通过定性建模推导系统的定性行为描述。定性仿真得到的系统在正常和各种故障情况下的定性行为描述可以作为系统知识用于故障诊断^[13]。该方法将系统描述成一个代表物理参数的符号集合以及反映这些物理参数之间相互关系的约束方程集合，然后从系统的初始状态出发，生成可能的后继状态，并用约束方程过滤掉那些不合理的状态，重复此过程直到没有新的状态出现为止。

1.2.2 定量分析的方法

■ 基于解析模型的方法

基于解析模型的方法是最早发展起来的,此方法需要建立被诊断对象的比较准确的数学模型。进一步,它又可以分为参数估计方法、状态估计方法和等价空间方法。把残差作为特征数据是基于解析模型方法的主流,残差是对象观测值和数学模型之间一致性比较的结果。一般如果故障发生了,残差会较大,如果是噪声、干扰、模型误差等,残差值会较小,所以可以定义适当的阈值来检测故障是否发生了。因此,基于解析模型的方法就是将系统的数学模型产生的信号与系统实际测得的信号进行比较,比较的结果就是表示这两种信号差别的残差量,然后对残差进行分析处理,进而实现故障诊断^[5]。

实际生产过程中,由于各种未知扰动、噪声等因素的影响,很难得到精确的数学模型,因此,基于解析模型的故障诊断方法的鲁棒性问题具有极大研究价值。由于建模不确定和故障实际上都是不期望的动态特性,如果不能很好处理建模不确定的影响,原有的故障诊断系统将会受到极大影响,甚至会完全失效。

基于解析模型的故障诊断利用了对系统内部的深层认识,具有很好诊断效果,但是这类方法需要对系统描述精确的数学模型。实际生产系统的数据模型往往难以建立。

■ 基于数据驱动的方法

在近 20 年里,随着机械、化工、冶金等大型工业向复杂化、智能化的发展,这些大型复杂工业过程有着两个显著的特点:一是日趋复杂的生产系统难以建立精确的数学模型,二是由于计算机技术和海量数据存储技术的快速发展,生产过程中产生的大量的反映过程运行机理和状态的数据,然而这些数据未能得到有效利用。基于以上两个因素,数据驱动的故障诊断方法得以重大发展。基于数据驱动的故障诊断方法对过程运行数据进行分析处理,主要分为机器学习类方法、多元统计分析类方法、信号处理类方法、粗糙集和信息融合类方法等。

机器学习类方法利用系统在正常和各种故障情况下历史数据训练神经网络或支持向量机等分类器算法用于故障诊断^[14]。本文采取的算法就是基于这种原理;多元统计分析方法根据过程变量的历史数据,利用多元投影方法将多变量样本空间分解成由主元变量构成的低维空间和一个残差子空间,并分别在这两个子空间中构造能够反映系统状态变化的统计量,然后将观测向量分别向两个子空间进行投影并计算相应的统计量指标用于故障诊断^[15-17]。目前常用的监控统计量有

Hotelling's T^2 和平方预测误差 (Square Predicted Error, SPE 或 Q); 信号处理方法是利用各种信号处理的方法对观测变量进行分析处理, 提取与故障相关的信号的时域或频域特征用于故障诊断^[18-19], 主要包括谱分析方法和小波变换方法; 信息融合方法通过获取被监控系统的多源信息, 对这些信息加以综合和自动分析以获得比单源信息更为可靠的结论。信息融合按照融合时信息抽象层次可分为数据层融合、特征层融合和决策层融合^[20-21]。

综上所述, 基于数据驱动的故障诊断方法完全从系统的历史数据出发, 不需要建立精确的解析模型, 因此在实际系统中具有广泛的应用。但是, 由于该方法没有内部系统结构和机理的信息, 因此对于故障的分析和解释比较困难。

1.3 故障诊断国内外研究现状及研究难点

如上面所述, 基于解析模型优点在于不需要大量的先验数据, 易于分析和诊断, 但是它计算量大, 需要深入系统本质机理建立准确的解析模型^[22-23]。但实际应用中复杂的动态往往难以建立精确的数学模型, 因此该类方法的适用范围和效果受到大大限制, 进展缓慢。

随着科技的发展, 现代化工过程中安置了大量传感器, 每天都会有数以百计的变量数据被采集存储下来, 这些数据背后中隐藏着故障相关信息。另外 MacGregor^[24] 指出化工工业故障诊断中的过程变量具有十分重要的地位, 将过程变量纳入监控范围, 能够有效地加强过程的诊断能力。基于以上两点原因, 基于数据驱动的故障诊断方法已成为当前过程控制领域研究热点。基于数据驱动的方法原理是通过各种数据分析处理方法对这些数据进行处理, 挖掘数据中隐含的有用信息, 获取正常操作和故障的特征模式, 实现故障诊断^[25-26]。这种方法正成为过程诊断领域内的研究热点。目前常用的有多变量统计方法、支持向量机等。

多变量统计分析方法的本质是采用各种数据降维方法来对历史数据降维, 建立一个低维的数据模型, 使得过程的故障诊断可以在这个低维空间内进行。降维的主要目的在于设法找到主元变量或者隐变量, 以使用更少的维数来描述整个过程的主要特征。目前常用的多变量统计方法有: 主元分析 (Principal Components Analysis, PCA)、Fisher 判别分析 (Fisher Discriminant Analysis, FDA)、独立分量分析 (Independent Component Analysis, ICA)、偏最小二乘 (Partial Least Square, PLS) 等。

早在 1980 年, Jackson 等人提出利用 PCA 算法地域主元信息进行提取, 并采用 Hotelling's T^2 和 SPE 两种统计量进行故障诊断^[27]。MacGregor 等在 1994 年进一步完善了 PCA 理论内容^[28-29]。1998 年 Qin 等人分析了 PCA 算法在故障诊断方面的应用, 利用子空间分析法分析了残差子空间的故障检测性、重构性、区别性^[30]。Kano 以主元子空间的差异来检测和区分故障的产生^[31]。然而, 传统的 PCA 算法是在过程数据都服从相同的高斯分布假设下进行的, 实际工业工程中的数据往往具有非正态、动态、非线性、非平稳等特性。因此近年来 PCA 改进算法陆续出现。Nomikos, P. & MacGregor 等人在 1994 年的时候提出了 multi-way PCA^[28]。Ku 等人在 1995 年提出 dynamic PCA^[32]。Cho 等人在 2005 年提出了 KPCA (Kernel PCA) 来对非线性过程进行故障监控^[33]。PCA 算法有效地解决了数据间的相关性, 并未对独立性进行分析。针对这一问题, 基于高阶统计量的信号分析方法 ICA 被引入到故障诊断中。2004 年 Lee 等人把 ICA 引入到工业过程监控和故障诊断中, 并对化工过程进行故障诊断^[34], 07 年进一步提出了核独立成分分析法^[35]。蒋丽英等^[36]提出了多向判别部分最小二乘 (MDPLS) 混合模型与多向主元分析 (MPCA) 相结合的故障诊断方法。Balla 等^[37]提出了一种多尺度 FDA 法, 与小波分析结合对故障进行分离。

基于模式识别的方法是将对已经存在的故障数据进行分类, 故障诊断问题就转化为把系统现有的工作状态归入哪一类的问题。从而可以故障诊断看作是模式分类和识别问题。Chiang 等^[38]提出了一种 FDA 和支持向量机 (Support Vector Machine, SVM) 相结合的故障诊断方法。该方法首先通过 FDA 降维方法进行特征提取, 剔除了不相关特征, 然后通过特征扩展将系统的动态特性加入其中, 最后采用 SVM 分类器进行故障诊断。Zhu 等人^[39]在 2011 年提出采用 KFDA (Kernel FDA) 对故障数据进行特征提取, 然后采用高斯混合模型 (Gaussian Mixture Model, GMM) 和 K 近邻 (K Nearest Neighbor, KNN) 两种分类器对故障数据进行分类, 该方法在 TEP (Tennessee Eastman Process) 数据集上对正常数据、故障 4、故障 8 和故障 14 数据集上得到了验证。Monroy 等人^[40]在 2010 年提出了一种半监督的故障诊断方法。该方法首先利用 ICA 进行特征提取, 之后利用具有 Bayesian Information Criterion (BIC) 信息的 GMM 和 SVM 进行组合用于最终的故障识别。

作为故障诊断领域的研究热点, 基于数据驱动的方法近年来取得了较好发展, 尤其是在多变量统计分析方法和机器学习中模式识别的方法。然而目前仍有

一些需要解决的问题。多变量统计分析方法中用于故障诊断的主要思路首先利用 PCA 等降维方法将原始特征空间投影到一个低维的主元空间和一个残差空间，之后利用 Hotelling's T^2 和平方预测误差 SPE 两种统计量进行故障检测，如果没有检测到故障的产生，便认为是正常状态，否则就要采用另外的方法进行故障诊断，判断出哪种类型故障的产生。这个过程需要数据降维，检测与诊断分开进行，相对比较繁琐，不能直接准确地定位何种故障的产生。故障诊断在作为模式识别问题解决时是被看作一个多类问题进行解决。当前的机器学习模式识别方法大都是针对两类问题，在处理多类问题时基本采用两种策略：one-against-one 和 one-against-all^[41]。若采用 one-against-one 需要训练 $n(n-1)/2$ 个分类器（假设 n 为类别个数），采用 one-against-all，需要训练 n 个分类器。当 n 较大时候，需要建立很多分类器，训练起来相对麻烦，分类时需经过多个分类器。采用 one-against-all 时还会出现训练数据不平衡的问题导致训练出来的分类器不准确^[42]。除此之外故障诊断领域还存在一些其他问题，都是亟需解决的。

1.4 本文研究内容

结合前文提到的当前故障诊断方法中遇到的研究难点问题，本文主要研究的内容是基于动态特性和稀疏表示的故障诊断方法，分别从以下三个方面进行相关研究，以便提高故障诊断效果。

基于稀疏表示的故障诊断方法研究。近年来，稀疏表示（Sparse Representation）在信号处理、机器学习和模式识别等领域中取得了迅速的进展，并在人脸识别、目标跟踪和异常路径行为检测等方面得到了很好的应用。本文采用基于稀疏表示方法用于解决多类分类问题时具有以下优势：(1). 模型统一，对于多类分类问题，目前的主要策略是利用两类分类器采用一对一或者一对多的分类器构造方法，这种分类策略在分类数目较少时可行，但是当分类的数目大于 3 时，就会变得很繁琐，当分类数目为 n 时，需要构建 $n(n-1)/2$ 分类器；而基于稀疏表示的分类方法对于多类问题只需一种分类模型即可解决多类问题。(2). 节省训练时间，基于稀疏表示的分类方法不需要进行分类器训练，不需要调试各种参数，只需要分类的各类选出一定的故障训练数据组成重构样本字典用于测试样本的重构，获得重构系数和重构残差进行故障诊断。(3). 不需要过多的训练样本数据，实验证明基于稀疏表示的分类效果并不与构成重构样本字典样本数目成

正比,在组成字典的重构样本较少时所获得的分类效果既已达到较好效果。这大大解决了现实生活中故障训练样本数据较少的难题。(4).不需要进行特征提取。基于稀疏表示的分类方法不需要采用数据降维等方法进行特征提取即可取得较好的分类效果。

引入了**动态特性信息作为分类特征的补充**。众所周知,在模式识别领域中有有效的特征描述对于得到良好的分类效果起着至关重要的作用。因此如何寻找一种有效的故障数据描述方法是当前研究的一大热点问题。在当前流行故障诊断方法中,用于分类的特征空间大都是由化工过程中相关传感器的测量变量组成的,对于同一个传感器取得的数据在时间轴上都是假设独立的,然而这在大多数的工业过程中都是不现实的。测量变量并不是序列无关的^[43],当前时刻的测量变量与过去若干时刻的变量都有关系,都是具有一定的动态特性的,并且不同的故障产生时,同一变量所具有的动态特性也是不一样的。因此,如何对故障数据的动态特性进行提取并有效地加以表示是提高故障诊断效果的有效途径之一。本文通过引入了动态特性信息作为分类特征的补充,利用添加该特性之后的特征空间进行故障诊断时发现效果大有改善。

字典学习优化。稀疏表示的本质就是利用过完备字典中的基元对测试样本进行线性组合。过完备字典的好坏直接决定了稀疏表示的效果,以至于决定了最后的分类效果,因此构建一个优异的冗余字典是非常重要的。本文通过 Feature-sign 算法和 language dual 算法对字典进行迭代优化,提高诊断效果。

1.5 本文的组织结构

第一章,绪论。主要论述了故障诊断的研究背景和意义,总结归纳了当前故障诊断的分类及方法,分析了国内外的行为故障诊断方法研究现状,并分析指出了当前面临的研究问题,列出了本文的主要研究目的和研究内容。

第二章,压缩感知与信号的稀疏表示理论。主要论述了压缩感知基本理论、稀疏表示的基础知识,对稀疏表示目前的实现方法进行了分类和介绍,为下文方法的提出做了铺垫。最后介绍了稀疏表示理论为在各个领域的应用。

第三章,TEP 过程介绍。主要介绍了本文的实验平台 TEP,从过程的工艺流程,反应物、生成物、过程变量、生产模式、过程故障和过程仿真数据等方面进行了描述,为后续实验部分奠定了基础。TEP 过程是一个典型的化工过程,具有

现代化工过程各方面特点，是故障诊断领域内公认的实验验证平台，易于在 MATLAB 上仿真获取实验数据。

第四章，基于动态特性与稀疏表示的故障诊断方法研究。首先论述了基于稀疏表示的故障诊断方法的基本原理，进行了验证实验。然后，介绍了动态特性引发展和本文引入动态特性的方法。最后论述了字典优化的学习方法。同时多组对比实验在实验平台 TEP 仿真数据上进行了测试，实验结果有效地验证了本文提出的算法性能。

最后是结束语，总结了本文的主要工作，展望了未来工作的方向，以及对如何更好地提高基于动态特性与稀疏表示的故障诊断方法的鲁棒性、准确性等研究热点和难点的探讨。

第二章 压缩感知和信号的稀疏表示理论

2.1 压缩感知的提出与内容

压缩感知 (Compressive Sensing, Compressed Sampling) 理论是近几年来信号处理领域诞生的一种新的信号处理理论, 由 D.Donoho、E.Candes 以及华裔科学家 T.Tao 等人提出^[44-48], 自诞生之日便极大地吸引了相关人员的关注。

压缩感知理论认为, 只要信号具有稀疏性, 那么可以把高维空间中的信号通过测量矩阵投影到一个低维空间, 然后通过求解一个最优化问题就能够精确重构出原始信号。压缩感知过程中采样和压缩编码是同时进行的, 它利用信号的稀疏性对信号实现非自适应的测量编码, 测量值并不是信号本身, 而是从高维空间到低维空间的投影值。假设信号 $x(n)$ 为 N 维信号, 基向量为 $\psi_i (i=1,2,3,\dots,N)$, 通过变换后原始信号可以表示为:

$$x = \sum_{i=1}^N \psi_i \alpha_i \quad \text{或} \quad x = \psi \alpha \quad (2-1)$$

显然这种变换是信号在不同域的代表形式。信号是否具有稀疏性或者近似稀疏性是运用压缩感知理论的关键问题, 若式 (2-1) 中的 α 只有 K 个非零值 ($N \gg K$); 或者说 α 经排序后按指数级衰减并趋近于零, 可认为信号是稀疏的^[49]。

若 $x \in R^N$ 在某空间可以稀疏表示, 那么由式 (2-1) 我们可以得到 $\alpha = \psi^T x$, 其中 α 可以看作是 ψ 的等价或者逼近的稀疏表示。然后我们设计出一个与变换基 α 不相关的观测矩阵对 α 进行观测, 得到观测值的集合 $Y = \phi \alpha = \phi \psi^T x (A^{cs} = \phi \psi^T)$ 这个过程也可以看作是原始信号 x 通过矩阵 A^{cs} 进行非自适应观测。最后, 通过求解一个最优化问题得到 x 的精确或者近似的解 \hat{x} :

$$\begin{aligned} \min_{\alpha} & \|\psi^T x\|_{l_0}, \\ \text{s.t.} & A^{cs} x = \phi \psi^T x = Y \end{aligned} \quad (2-2)$$

得到的解 \hat{x} 在 ψ 上基本则是最稀疏的。

通过上文的分析我们可以看出, 压缩感知的本质与传统的先压缩和采样的思想不同, 它是通过对少量的观测值求解最优化问题就可以得到原始信号 x , 是对

那奎斯特采样定理的突破。由于它具有使用少量的非线性观测值就能对信号精确的重构的优势，在图像处理、数据压缩、模式识别及集成电路等方面有着广泛的应用^[50]。

2.2 稀疏表示基本理论

稀疏表示的研究兴起于二十世纪九十年代，近几年来得到了蓬勃的发展，压缩感知的提出更是为其提供了工程应用的土壤^[51]。广义上，稀疏是指集合元素大部分为 0，这里的集合可以是矩阵、向量等。信号稀疏表示是指，信号在一个过完备基上表示是稀疏的。采用稀疏逼近原始的数据的方法进行稀疏表示，可以从本质上降低数据处理的成本，提高压缩效率。

假设有一个实值、长度有限的一维离散时间信号 x ，它是在 R^N 上的 $N*1$ 的列向量，向量元素为 $x(n)$, $n=1,2,\dots,N$. 基向量 $\{\psi_i\}_{i=1}^N$ 可以表示 R^N 上的任何 $N*1$ 维的信号。我们可以假设所选基是标准正交的。使用 $N*N$ 的基 $\Psi = [\psi_1 | \psi_2 | \dots | \psi_n]$ ，其中向量 $\{\psi_i\}$ 是列向量。那么原始信号 x 的稀疏表示为：

$$x = \sum_{i=1}^N s_i \psi_i \quad \text{或者} \quad x = \Psi s \quad (2-3)$$

其中 s 称为权重系数，是一个 $N*1$ 的列向量， $s_i = \langle x, \psi_i \rangle = \psi_i^T x$ 。简单的说， s 是原始信号 x 在不同域的等价表示， x 是时域或空间域的， s 属于 Ψ 域。

如果在上式 (2-3) 中求解的权重系数 s 中只有 K 个是非零的， $(N-K)$ 个系数为零，即信号 x 为 K 个基向量的线性组合，那么就说 x 是 K 稀疏的。我们感兴趣的是 $K \ll N$ 的情况，因为只有满足这个条件信号才具有稀疏性，也就是式

(2-3) 中只有少量的大系数，绝大部分系数都接近零，信号 x 才是稀疏的，或者是可压缩的^[52]。

所以，信号的稀疏表示就是求解稀疏系数 s ，因此稀疏表示可以描述为：在误差条件允许的前提下，在所有可能的基元组合中，选出 s 最为稀疏的一个解。即：

$$\arg \min \|s\|_0 \quad \text{subject to} \quad x = \Psi s \quad (2-4)$$

其中 $\|\cdot\|_0$ 为 l_0 范数，代表的含义为系数非零项的个数。

当 Ψ 是 Hilbert 空间的一个正交基时，上式最优化问题的解是很容易求得的。但是对于一个超完备基来说，该最优化问题是一个 NP-hard 问题。为了得到稀疏表示的最优化的解，一般采用稀疏逼近的方法来求解。

所谓稀疏逼近指的是：存在某一集合 $D = \{g_k; k = 1, 2, \dots, K\}$ ，其中元素 $g_k \in R^N$ 为单位矢量，原子库中基元个数 $K \gg N$ ，集合 D 称为原子库，它是过完备的。则对于任何信号 $f \in H$ ，在 D 中能自动选取 m 个基元对信号 f 实施 m 项线性组合^[48]：

$$f_m = \sum_{\gamma \in I_m} c_\gamma g_\gamma, \quad \text{card}(I_m) = m \quad (2-5)$$

其中， c_γ 为展开系数，由于 $m \ll N$ ，因而被称为稀疏逼近。由于原子库的冗余性，基元矢量 g_k 不再是线性无关的。因此上式有不同的线性组合，不同组合中基元和基元个数也不尽相同。模型示意图如下：

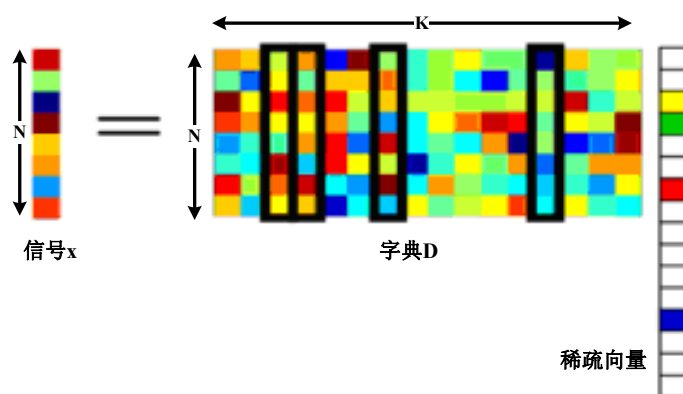


图 2-1 稀疏表示模型

2004 年，在前人的基础上，Elad 和 Bruckstein^[53]对相关理论定理进行了证明：

定理 1：设 D 是一个相干系数为 μ 的原子库，如果一个信号 $s \in R^N$ 可以表示为：

$$s = \sum_{i=0}^m c_i g_i \quad (2-6)$$

并且 $\|c_0\| < 1/\mu$ ，则式子就是信号 s 在 D 中最稀疏的表示。

定理 1 中相干系数 μ 定义如下：

$$\mu = \sup_{\substack{i,j \\ i \neq j}} |\langle g_i, g_j \rangle| \quad (2-7)$$

μ 的大小与原子库的关联性成正比。若 $\mu = 1$ ，表明原子库中至少两个基元一样；当 μ 较小时，则称原子库是非相干的。

定理 1 结果表明，如果 s 在 D 中有非常稀疏的表示，那么通过基追踪算法 (Basis Pursuit, BP) 对原信号进行精确稀疏表示。

定理 2：如果信号 s 在原子库 D 中存在一个稀疏表示，而且满足下式

$$\|c_0\| < \frac{\sqrt{2}-0.5}{\mu} \quad (2-8)$$

则 l_1 的最小化问题有一个唯一解，它就是 l_0 最小化问题的解。这就意味着我们可以通过 BP 算法求解 l_1 问题，取代原来的 l_0 问题。

综上所述，在满足上述条件下， l_0 范数的优化问题与 l_1 范数的优化问题结果一致，即求最小化 l_0 范数的问题就可以转化为求最小化 l_1 范数的问题。

$$\arg \min_s \|s\|_1 \quad \text{subject to } x = \psi s \quad (2-9)$$

2.3 稀疏表示的实现算法

当前，稀疏表示的主流求解方法分为全局优化方法和贪婪算法。基追踪算法 (Basis Pursuit, BP)^[53] 是全局优化方法中的代表算法。贪婪算法主要有匹配追踪算法 (Matching Pursuit, MP)^[54]、正交匹配追踪算法 (Orthogonal Matching Pursuit, OMP)^[55]、正则化正交匹配追踪算法 (Regularized Orthogonal Matching Pursuit, ROMP)^[56]，最优正交匹配追踪方法 (Optimized Orthogonal Matching Pursuit, OOMP)^[57] 等。

2.3.1 全局优化算法

全局优化方法是指在满足一定的约束条件下，使得某个特殊的目标函数最小。典型的目标函数是凸函数，并且任何局部最小值也是全局最小值。

■ 基追踪算法

基追踪算法用最小化 l_1 范式代替为最小化 l_0 范式问题，这是一个凸优化的问题，可以转化为线性规划问题加以求解。一般实际应用中也要考虑重构误差，具体问题描述如下：

$$\arg \min_s \|s\|_1 \quad \text{subject to } x = \psi s \quad (2-10)$$

常用的基追踪求解方法由内点法、梯度投影法以及同伦算法。其中内点法精度高、稳定性好，较为常用，但是内点法计算复杂度高，速度慢。

基追踪方法并不适用于实际信号，Donoho 等人提出了基追踪去噪算法，式子改为：

$$\arg \min_s \left\{ \frac{1}{2} \|x - \psi s\|_2^2 + \lambda \|s\|_1 \right\} \quad (2-11)$$

2.3.2 贪婪算法

贪婪算法基本思想是每次迭代过程中，从字典中选择与信号最相匹配的基元来构建稀疏逼近，一定迭代次数后，信号就可以用一些基元的线性组合表示。

■ 匹配追踪算法

假设给定的过完备的字典 $D = \{g_k, k=1, 2, \dots, K\}$ ，其中所有的基元都进行了归一化，即 $\|g_k\|_2 = 1$ 。初始化时，将 f 投影到一个向量 $g_{\gamma_0} \in D$ 上，就可以得到残差 R_f ：

$$f = \langle f, g_{\gamma_0} \rangle g_{\gamma_0} + R_f \quad (2-12)$$

由于 R_f 和 g_{γ_0} 正交，所以有：

$$\|f\|^2 = \|\langle f, g_{\gamma_0} \rangle\|^2 + \|R_f\|^2 \quad (2-13)$$

如果要极小化 $\|R_f\|^2$ ，那么就要使得 $|\langle f, g_{\gamma_0} \rangle|$ 为极大的 $g_{\gamma_0} \in D$ 。为了有效地减少计算量，也可以通过调整参数来寻找向量 g_{γ_0} 比如引入最佳因子 $\alpha \in (0, 1]$ ，那么就可以得到：

$$|\langle f, g_{\gamma_0} \rangle| > \alpha \sup |\langle f, g_{\gamma_0} \rangle| \quad (2-14)$$

然后将 R_f^m 投影到 g_{γ^m} 上，那么可以得到：

$$R_f^m = \langle R_f^m, g_{\gamma^m} \rangle g_{\gamma^m} + R_f^{m+1} \quad (2-15)$$

由于 R_f^m 与 g_{γ^m} 正交，那么有：

$$\|R_f^m\|^2 = \left\| \left\langle R_f^m, g_{\gamma^m} \right\rangle \right\|^2 + \|R_f^{m+1}\|^2 \quad (2-16)$$

分别对上面的两式的 m 从 0 到 $M-1$ 求和，可以得到：

$$f = \sum_0^{M-1} \left\langle R_f^m, g_{\gamma^m} \right\rangle g_{\gamma^m} + R_f^M \quad (2-17)$$

$$\|f\|^2 = \sum_0^{M-1} \left\| \left\langle R_f^m, g_{\gamma^m} \right\rangle \right\|^2 + \|R_f^M\|^2 \quad (2-18)$$

研究表明，当 m 趋向于无穷的时候 $\|R_f^m\|$ 将会按指数级趋向于 0。

该算法一个明显的缺点是在已知基元组成的子空间上，信号的拓展上不是最好的，因为它不是一个正交投影。虽然已经证明匹配追踪在有限维的情况下是指数级收敛的，但是在无穷维的情况下并不收敛。正交匹配追踪递归对所选基元的集合进行正交化，解决了这个问题。

■ 正交匹配追踪算法

OMP 算法是在 MP 算法基础上的一种改进算法。它与 MP 都是从过完备字典中找出与待分解信号最为匹配的基元，不同的是 OMP 算法需要将所选基元利用 Gram-Schmidt 正交化方法进行正交化处理，再将信号在这些正交基元构成的空间上进行投影^[58]。从而可以得到信号在已选定基元上的分量和剩余残差分量，最后使用 MP 相同的方法分解残差分量。

OMP 采用 MP 算法相同的方法选取基元：

$$g_{\gamma_m} \text{ subject to } \left| \left\langle R_m, g_{\gamma_m} \right\rangle \right| = \max_{\gamma} \left| \left\langle R_m, g_{\gamma} \right\rangle \right| \quad (2-19)$$

基元选定之后，使用 Gram-Schmidt 进行正交化处理：

$$u_m = g_{\gamma_m} - \sum_{p=0}^{m-1} \frac{\left\langle g_{\gamma_m}, u_p \right\rangle}{\|u_p\|^1} u_p \quad (2-20)$$

此时残差 R_m 投影到 u_m 上，而不是 g_{γ_m} 上：

$$R_m = \frac{\left\langle u_m, R_m \right\rangle}{\|u_m\|^2} u_m + R_{m+1} \quad (2-21)$$

信号 f 的 M 项表示为

$$f = \sum_{m=0}^{M-1} \frac{\langle u_m, R_m \rangle}{\|u_m\|^2} u_m + R_M \quad (2-22)$$

M 小于信号 f 的空间维数。

从收敛速度上来看，OMP 收敛速度比 MP 更快，因此，在精度相同的情况下，OMP 表示的信号更为稀疏，但是精度上远不如最小 l_1 范数，且对于字典的要求更加严格。

研究者在 OMP 算法的基础上提出了更多变形的算法，比如正则正交匹配追踪算法、压缩采样匹配追踪算法、最优正交匹配追踪算法等。

2.4 稀疏表示的相关应用

近年来，大量稀疏表示研究论文出现在模式识别与计算机视觉领域的各种会议期刊上，稀疏表示与分析的专题报告被开设。基于稀疏表示理论得到了广泛研究和应用^[59-61]，在图像处理方面最为突出，主要有人脸识别、图像恢复、目标跟踪、行为检测等。

2009 年，John Wright 和 Yi Ma^[62] 等人提出基于稀疏表示分类（Sparse Representation Classification, SRC）的人脸识别方法，该算法自动找出被遮挡的部分，并对该部分进行重建，有效的解决了人脸部分遮挡问题。Mairal 等^[63] 使用稀疏表示对图像去噪，同时对较小的区域进行修复，在图像恢复上取得了良好的效果。Han 等^[64] 将 l_1 范数最小化的稀疏重构应用到多目标跟踪中，用搜索区域内的跟踪样本来线性重构需要跟踪的目标，已达到目标跟踪的目的。Li 等^[65] 采用三次 B 样条方法作为行为轨迹的表示方法，通过稀疏表示方法对行为轨迹进行稀疏表示，以实现异常行为检测的目的。除此之外，在图像的其他领域如图像融合、图像目标检测、图像修复等也提出了令人瞩目的基于稀疏表示的算法。

2.5 本章小结

压缩感知理论的发展积极带动了各个领域的研究，本章主要介绍了压缩感知理论的基本框架和大体过程，并着重阐述了本章故障诊断的核心思想稀疏表示理论。本章对稀疏表示理论做了详细的数学描述，给出了稀疏表示问题是如何转化为最优化问题的数学推导过程。之后介绍了当前该问题的几种代表性求解方法。

最后介绍了稀疏表示理论近年来在各个领域的实际应用。本章内容为接下来章节算法的提出做好了背景知识的铺垫。

第三章 TEP 过程介绍

3.1 TEP 过程概述

田纳西-伊斯曼过程 (Tennessee Eastman process, TEP) 最早提出于 1990 年在芝加哥举行的美国化工学会 (AIChE) 年会上, 该过程是由美国伊斯曼化学公司的过程控制小组创建的, 基于实际工业过程的过程控制实例, 测试过程是基于一个真实工业过程的仿真^[66]。建立该过程的目的在于为评价过程控制和监控方法建立一个现实的工业平台。大量文献引用它作为数据源来进行控制、优化、过程监控、故障诊断等方面的研究。TEP 可以在 MATLAB 上进行仿真。

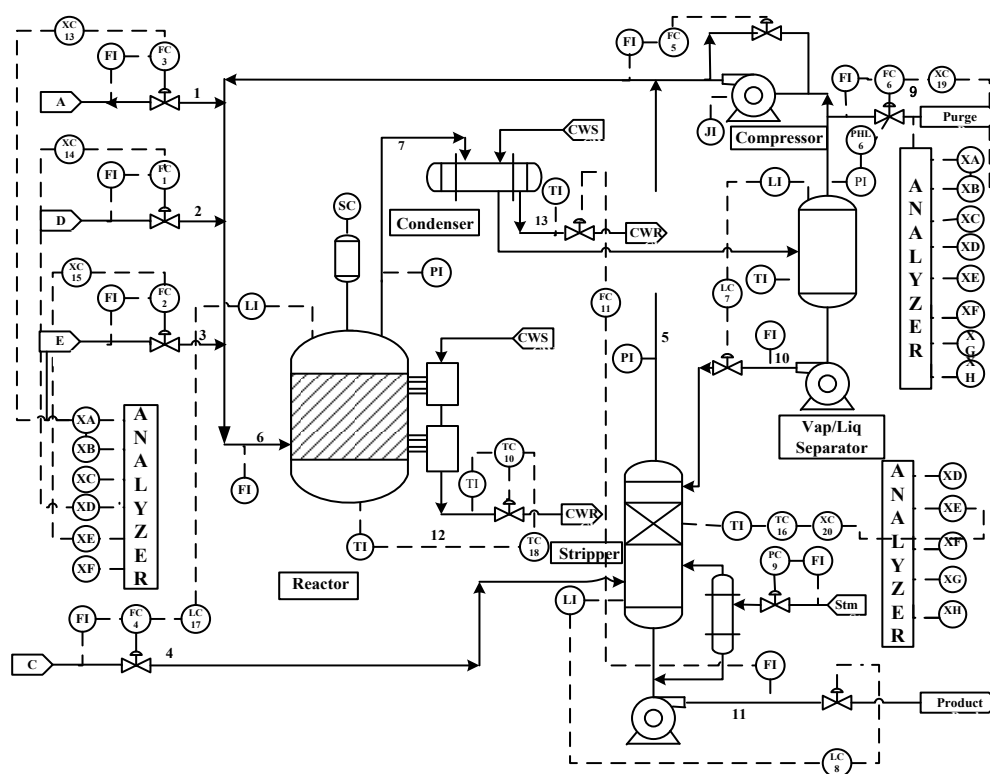
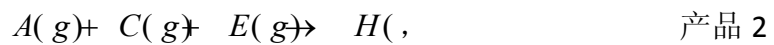
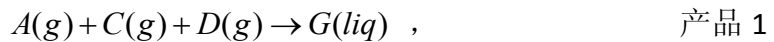


图 3-1 TEP 过程工艺流程图

TEP 过程流程图如图 3-1 所示, 该过程由五个反应单元组成: 反应器 (reactor)、冷凝器(Condenser)、分离器(Separator)、汽提塔(Stripper)和压缩机(Compressor); 共有八种成分: A、B、C、D、E、F、G 和 H, 其中 A、C、D、E 为四种气态反应物, B 为催化剂, F 为副产物, G 和 H 为最终产物。在反应器

中进行的化学反应如下：



反应物在反应器中经过催化剂催化作用生成液态产物。之后产品之后进入冷凝器，进行冷却，紧接着送入到冷/液分离器。从分离器出来的蒸汽通过压缩机再循环进入反应器。为了防止过程中惰性成份和反应副产物产品的积聚，必须排放一部分再循环流。来自分流器的冷凝成份（流 10）被泵送入汽提塔。流 4 用于流 10 中的剩余反应物，这些剩余反应物通过流 5 与再循环流结合。从汽提塔底部出来的产品 G 和 H 被送到下游过程。

根据 G/H 质量比的不同，TEP 过程具有六个操作模式，本文采用的数据为基本工况下产生的。见表 3-1：

表 3-1 TEP 过程操作模式参数

模式	G/H 质量比	生产率
1	50/50	G:7038 kgh-1; H: 7038 kgh-1 (基本工况)
2	10/90	G:1408 kgh-1; H: 12669 kgh-1 (优化基工况)
3	90/10	G:10000 kgh-1; H: 1111 kgh-1 (优化操作)
4	50/50	最大生产率
5	10/90	最大生产率
6	90/10	最大生产率

3.2 TEP 过程变量和故障描述

TEP 过程包括 12 个控制变量和 41 个测量变量，测量变量包含 19 个非连续测量组分变量和 22 个连续变量，其中 19 个非连续测量成份变量由 3 个成份分析仪获得。表 3-2 为 22 个连续测量值描述表，用 XMEAS 进行标示定义。12 个控制变量用 XMV 来定义，取值范围为 0~100，初始值为基础值，如表 3-3。表 3-4 中 XMEAS (23) ~XMEAS (41) 为 19 个成分测量值。本文实验采用的训练集和测试集中的数据包含了 41 个测量变量和 11 个控制变量。除了反应器中搅拌器的搅拌速度外，共有过程变量 52 个。

表 3-2 连续测量变量

变量名	变量符号	基础值	单位
A 物料流量	XMEAS1	0.25052	kscmh
D 物料流量	XMEAS2	3664.0	Kgh-1
E 物料流量	XMEAS3	4509.3	Kgh-1
A、C 混合物料流量	XMEAS4	9.3477	kscmh
回收流量	XMEAS5	26.902	kscmh
反应器进料率	XMEAS6	42.339	kscmh
反应器压力	XMEAS7	2705.0	kPa gauge
反应器液位	XMEAS8	75.00	%
反应器温度	XMEAS9	120.40	oC
放空率	XMEAS10	0.3372	kscmh
产品分离器温度	XMEAS11	80.109	°C
产品分离器液位	XMEAS12	50.000	%
产品分离器压力	XMEAS13	2633.7	kPa gauge
产品分离器出口流量	XMEAS14	25.160	m3h-1
汽提塔液位	XMEAS15	50.000	%
汽提塔压力	XMEAS16	3102.2	kPa gauge
汽提塔出口流量	XMEAS17	22.949	m3h-1
汽提塔温度	XMEAS18	65.731	°C
汽提塔蒸汽流量	XMEAS19	230.31	kg h-1
压缩机工作功率	XMEAS20	341.43	kw
反应器冷却水出口温度	XMEAS21	94.599	°C
分离器冷却水出口温度	XMEAS22	77.297	°C

TEP 仿真模型一共包括 21 个预设故障，其中 16 个是已知的，5 个是未知的^[67]。故障类型主要分为五种：故障 1~7 为阶跃型故障，指的是由过程变量的阶跃变化的故障，如冷却水的入口温度或者进料成分的变化；故障 8~12 为随机变量型，这是因为一些过程变量的可变性增大有原因；故障 13 是反应动力学中的缓慢漂移；故障 14、15 和 21 是与粘滞阀有关的；故障 16~20 是未知的。

表 3-3 成分变量

组分	变量名	基础值	单位
A	XMEAS23	32.188	Mol%
B	XMEAS24	8.8933	Mol%
C	XMEAS25	26.383	Mol%
D	XMEAS26	6.8820	Mol%
E	XMEAS27	18.776	Mol%
F	XMEAS28	1.6567	Mol%
A	XMEAS29	32.958	Mol%
B	XMEAS30	13.823	Mol%
sC	XMEAS31	23.978	Mol%
D	XMEAS32	1.2565	Mol%
E	XMEAS33	18.579	Mol%
F	XMEAS34	2.2633	Mol%
G	XMEAS35	4.8436	Mol%
H	XMEAS36	2.2986	Mol%
D	XMEAS37	0.01787	Mol%
E	XMEAS38	0.83570	Mol%
F	XMEAS39	0.09858	Mol%
G	XMEAS40	53.724	Mol%
H	XMEAS41	43.828	Mol%

3.3 实验仿真数据描述

本文实验中采用的数据集由 MATLAB 仿真程序获得，训练集和测试集的数据包含 41 个测量变量和 11 个控制变量，采样间隔取 3 分钟，包括正常数据和前 20 类故障数据。训练集中正常模式数据中有 500 个样本数据，其余 20 类故障模式具有 480 个样本数据；测试集中每种模式的样本数据个数为 960，正常模式所有样本为正常数据。故障模式中前 160 个样本数据为正常数据，后 800 个数据为故障数据，这是因为仿真程序进行仿真时，各种故障是在 8 小时后引入的。正常状态下某些测量值时间变化情况如下图 3-2 和图 3-3。

表 3-4 故障描述

编号	故障描述	分类
1	A/C 进料流量比变化, 组分 B 含量不变 (流 4)	阶跃
2	组分 B 含量发生变化, A/C 进料流量比不变 (流 4)	阶跃
3	物料 D 的温度发生变化 (流 2)	阶跃
4	反应器冷却水入口温度发生变化	阶跃
5	冷凝器冷却水入口温度发生变化	阶跃
6	物料 A 损失 (流 1)	阶跃
7	物料 C 压力损失 (流 4)	阶跃
8	物料 A,B,C 的组成发生变化 (流 4)	随机变量
9	物料 D 的温度发生变化 (流 2)	随机变量
10	物料 C 的温度发生变化 (流 2)	随机变量
11	反应器冷却水入口温度发生变化	随机变量
12	冷凝器冷却水入口温度发生变化	随机变量
13	反应动力学特性发生变化	慢漂移
14	反应器冷却水阀门	黏住
15	冷凝器冷却水阀门	黏住
16	未知	未知
17	未知	未知
18	未知 s	未知
19	未知	未知
20	未知	未知
21	流 4 的阀门固定在稳态位置	恒定位置

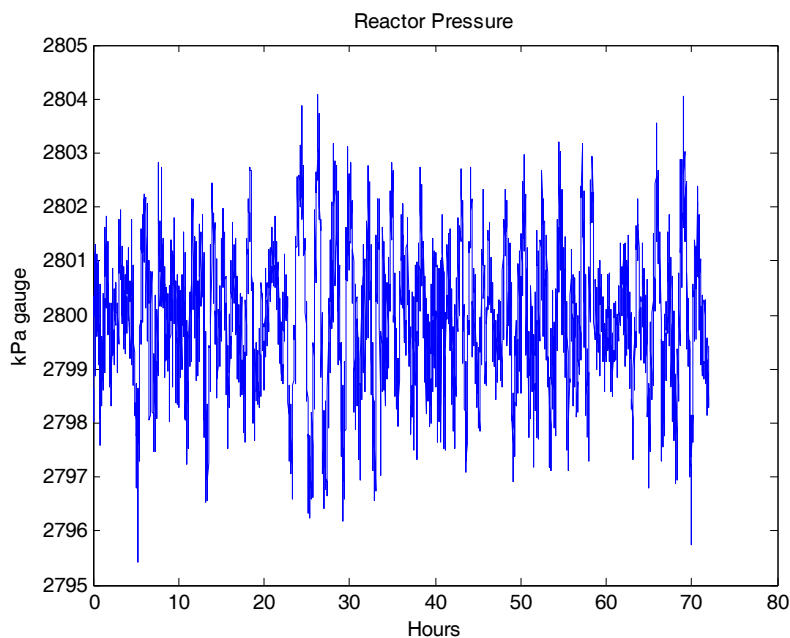


图 3-2 反应器压力

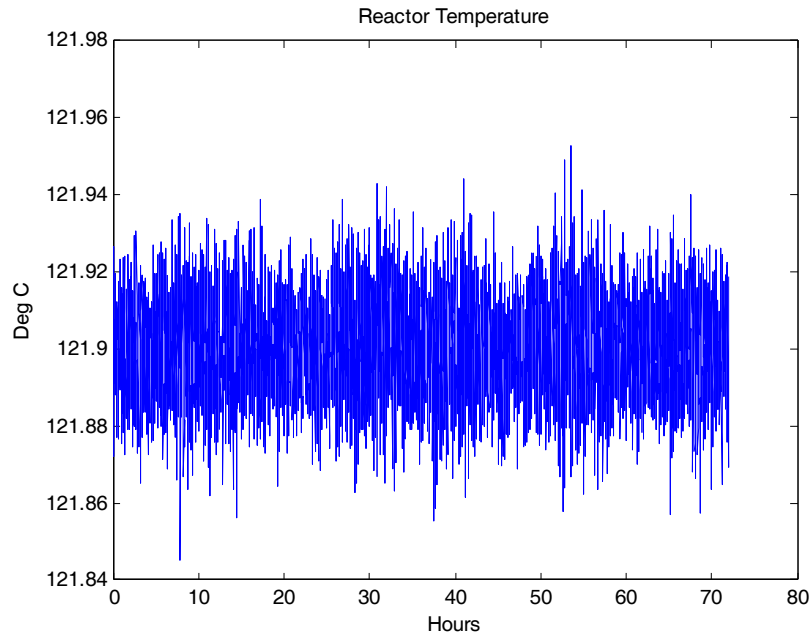


图 3-3 反应器温度

3.4 本章小结

本章对本文算法验证平台 TEP 化工过程从各个方面做了详细的介绍，包括过程的工艺流程，反应物、生成物、过程变量、生产模式、过程故障和过程仿真数据等，为后续实验部分奠定了基础。TEP 过程是一个典型的化工过程，具有现代化工过程各方面特点，是故障诊断领域内公认的实验验证平台。过程故障类型具有代表性，而且易于在 MATLAB 上仿真获取数据。利用它作为我们实验的验证平台能全面的检验算法的可靠性。

第四章 基于动态特性与稀疏表示的故障诊断

4.1 基于稀疏表示的故障诊断方法

如前文所说，稀疏表示在信号领域取得了重大发展，表现出了种种优点，也被广泛的用于图像处理、人脸识别、目标跟踪等各个领域。本文采用稀疏表示的方法用于故障诊断，同样取得了较好的效果。

稀疏表示的基本假设是：某一类的未知样本可以被该类的若干样本在某一子空间有效地进行线性表示。因此基于稀疏表示的故障诊断方法，实质就是将化工过程中故障检测与诊断理解为每一时刻测试样本与预先标定的能代表化工故障典型的训练样本进行匹配分类进而检测判断的过程。在现实化工过程中由于系统相对比较复杂，故障的类型往往很多，导致使用故障样本数据集进行新样本稀疏表示的系数会很稀疏，故障样本数据集中的多数类别的样本对应重构系数趋近于零，因此在本文中提出了基于稀疏表示分析获取样本稀疏重构系数和重构误差的方法。这种方法在获得故障样本数据集的基础上，对样本集中的所有样本通过重构系数和重构误差进行评估，从而选择出最具有代表性的一些样本来重构新样本。这些样本揭示着新样本数据的主要特征和内在结构，通过对新样本的重构，从而使得故障检测与诊断具有较强的鲁棒性。基于稀疏表示的故障诊断方法研究内容分为以下两个方面：故障样本字典的构造和基于稀疏表示的故障诊断。基于稀疏表示的故障诊断方法研究框架如图 4-1 所示。

4.1.1 故障样本字典的构造

基于故障诊断方法故障字典构造，就是建立化工过程中各类故障表示的完备集合。本文在进行实验时所选的数据集为 TEP 数据集，该数据集包含正常数据和 20 类故障数据共 21 种类别样本数据，每个样本数据有 52 维属性，每种类别包含有 480 个训练样本数据和 800 个测试样本数据。因此，本文从每种类别训练数据集中选取一定数目的样本数据用来构建故障样本字典。其中，每一个样本数据的表示序列是故障样本字典中的一个列向量。

具体地，利用化工过程中的每种典型故障样本数据建立样本集合，

$$A_j = \{a_j^1, a_j^2, \dots, a_j^K\} \quad (4-1)$$

其中， K 表示第 j 类故障样本个数。将所有故障类型的样本数据进行合并，形成故障样本字典，如下形式：

$$B = \cup\{A_j\} = \{a_1^1, a_1^2, \dots, a_1^K, a_2^1, a_2^2, \dots, a_j^K\} \quad (4-2)$$

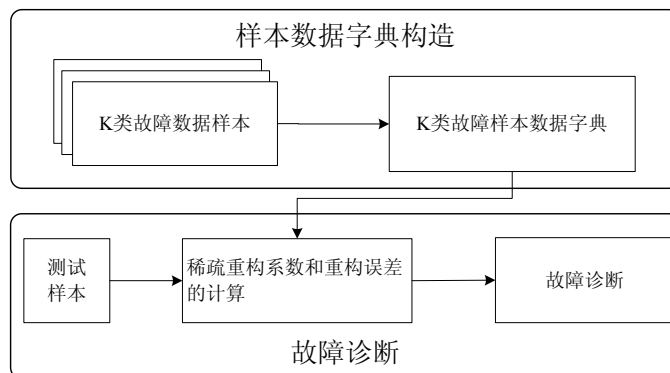


图 4-1 基于稀疏表示的故障诊断方法研究框架

其中， $j=1\dots J$ 。 J 表示该化工过程中典型正常类型与各种故障种类数目之和。因此，我们可以得到由 $K \times J$ 个故障样本组成的故障样本字典 B ，如图 4-2 所示。

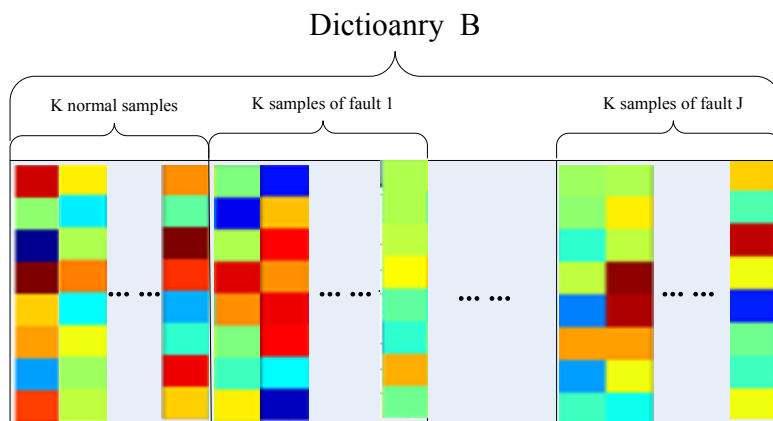


图 4-2 故障样本数据集内部结构示意图

4.1.2 基于稀疏表示的故障诊断

由于现实化工过程中往往含有包括正常模式在内的很多故障模式，所以构造的故障样本字典也很大，这便会导致使用故障样本字典进行新样本稀疏表示的系数会很稀疏，样本集中的多数类别样本对应的重构系数趋近于 0。这样的稀疏表示系数在实际中保证了故障样本集中只有少数和测试样本相似的故障样本会用于新样本的重构。

从压缩感知理论出发，故障诊断的本质在于使用样本选择或分类的方法（分类器、稀疏表示等）从故障样本字典中获取新的样本的故障分类结果，所以假设新的故障样本数据可以通过以上所构造的故障样本字典来线性重构得到，如公式 4-3:

$$B\psi \approx F_{unknown} \quad (4-3)$$

其中， $F_{unknown}$ 是新故障样本表示， $\psi = \{\psi_j^k\}, j=1, \dots, J, k=1, \dots, K$ 是对应 B 中样本的系数向量， ψ_j^k 就是第 j 种故障第 k 个样本的重构系数。

在实际实验中， $\psi = \{\psi_j^k\}, j=1, \dots, J, k=1, \dots, K$ 中只有 l 个是非零的系数 ($l \ll K$)。在数学中，我们将这样的矩阵称为样本的 l 稀疏表示。非零系数的个数可以用 $\|\psi\|_0$ 来表示，所以我们可以通过最小化 $\|\psi\|_0$ 获得样本的稀疏重构。这个 0-范数的问题是 NP-hard 的问题。根据上文介绍的求解方法，可以通过 1-范数最小化可以用来解决稀疏表示问题:

$$\begin{aligned} \arg \min \|\psi\|_1 \\ s.t. \|B\psi - F_{unknown}\|_2 \leq \varepsilon \end{aligned} \quad (4-4)$$

其中， $\|\cdot\|_1$ 表示 1-范数， $\varepsilon > 0$ 表示一个较小的数值。通过故障样本字典 B 和计算得到的稀疏系数向量 ψ ，新样本就可以通过重构来得到，重构公式为:

$$F_{unknown} \approx B\psi = \sum_{k=1}^K B^k \psi^k \quad (4-5)$$

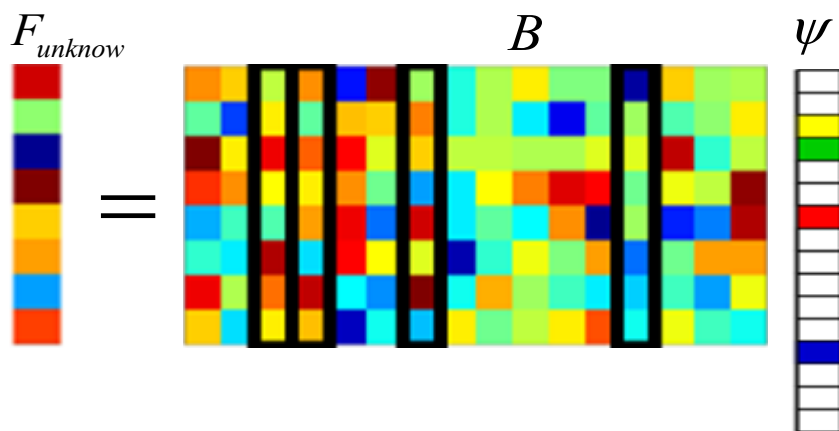


图 4-3 测试样本利用故障样本数据集稀疏表示示意图

基于 1-范数最小化的稀疏重构，保证了该重构是很紧致的^[68-69]，也就是说，

我们从故障样本数据集中挑选出最有代表的样本来稀疏表示新样本即样本重构，这样可以有效地通过稀疏重构系数和重构误差来揭示新故障样本的主要特征和内在结构与故障样本数据集中的哪些样本相似一致，使得稀疏重构系数和稀疏重构误差用作故障诊断时具有较强的鲁棒性。

当测试样本稀疏重构完成，采用公式(4-4)可以计算求得其对所有故障样本数据集 B 的稀疏系数 ψ 。随后，对于每一种故障模式 j ，我们定义其对应的特征函数 δ_j 。其中 δ_j 定义为只保留故障样本数据集 B 中第 j 种故障模式所对应的稀疏系数，同时将其其他故障模式样本所对应的稀疏系数都赋值为 0。基于 δ_j 的定义与集合 B ，我们利用重构误差来判断新样本属于哪种故障模式或正常模式，过程如下公式：

$$r_j(F_{unknown}) = \|F_{unknown} - B\delta_j(\psi)\|_2, j = 1 \dots J \quad (4-6)$$

其中， $r_j(F_{unknown})$ 表示利用第 j 种故障模式样本数据对新样本进行重构时的误差，通过此方法我们可以依次求出利用每种故障样本数据集对新样本进行重构的误差。基于某一类的未知样本可以被该类的若干样本在某一子空间有效地进行线性表示这一假设，我们可以认为采用同类故障样本数据对新样本进行稀疏重构时所获得重构误差是最小的。因此，我们还可以通过公式(4-7)将新样本 $F_{unknown}$ 进行故障诊断分类：

$$Classify(F_{unknown}) = \arg \min_j r_j(F_{unknown}) \quad (4-7)$$

这种首先采用数据字典进行稀疏表示，之后利用重构系数与重构误差进行分类的方法我们称之为稀疏表示分类方法（Sparse Representation Classification, SRC）。

通过上述基本原理的介绍，基于稀疏表示的分类器在故障诊断方面具有三个特点：一是模型统一，SRC 对于多类问题只需一种分类模型即可解决多类问题。二是减少了分类器训练环节，SRC 不需要进行分类器训练，只需要分类的各类选出一定的故障训练数据组成重构样本字典用于测试样本的重构，获得重构系数和重构残差进行故障诊断。三是不需要进行特征提取。SRC 不需要进行利用 PCA 等降维方法进行特征提取，在原始特征空间上进行即可。

4.2 基于稀疏表示的故障诊断方法实验

为了多方面的验证本论文提出的算法，首先我们设置了三组实验，前两组实验的故障种类数目为 4，相对较小，其中还包括了正常模式，这样做的目的在于体现故障检测与诊断两个问题的统一性；第三组我们则将进行 TEP 前 20 种故障模式分类效果实验。具体的实验描述如表 4-1。

表 4-1 三组实验描述

实验	模式	故障描述	类型	训练数据	测试数据
1	正常	-	-	480	800
	故障 1	A/C 进料流量比变化，组分 B 含量不变	阶跃	480	800
	故障 6	物料 A 损失	阶跃	480	800
	故障 18	未知	未知	480	800
2	正常	-	-	480	800
	故障 2	组分 B 含量变化，A/C 进料流量比不变	阶跃	480	800
	故障 5	冷凝器冷却水入口温度发生变化	阶跃	480	800
	故障 7	物料 C 压力损失	阶跃	480	800
3	20 类故障		-----	-----	-----

表 4-2 实验 1 实验混淆矩阵

模式	正常	故障 1	故障 6	故障 18
正常	764	6	2	100
故障 1	2	794	0	0
故障 6	0	0	798	0
故障 18	34	0	0	700
总数	800	800	800	800

表 4-3 实验 2 实验混淆矩阵

模式	正常	故障 2	故障 5	故障 7
正常	800	24	1	0
故障 2	0	776	0	0
故障 5	0	0	799	0
故障 7	0	0	0	800
总数	800	800	800	800

表 4-2 与表 4-3 为表示前两组实验效果的混淆矩阵，表 4-2 显示在第一组实

验中，800 个正常模式测试数据被正确分为正常模式的样本 764，被误分为故障 1 的样本有 2 个，被误分为故障 18 的样本有 34 个；800 个故障 6 的测试样本中，798 个样本被正确分为故障 6，仅有 2 个测试样本被分为正常数据。在第二组的实验中，正常模式和故障 7 的测试数据全都被正确分类；故障 5 测试样本中仅有 1 个样本被误分为正常数据；故障 2 中 776 个样本被正确分类。由此我们可以看出，当选定某些故障类别且故障类别较小时，采用基于稀疏表示的故障诊断方法分类效果较好，误分率可以控制在 5%以内。同时由于两组实验中都包括了正常模式，所以采用本文提出的方法可以确定是否有故障产生的同时，确定何种类型的故障产生，即将故障检测与故障诊断合为一个问题进行解决，不需要分布进行解决。同时可以说明，对于多类分类问题，SRC 不需要训练多个分类器加以组合实现多类分类，只需要一个统一的模型即可解决多类问题。

表 4-4 20 类故障诊断效果比较

F1 index(%)	<i>ref[40]</i>		<i>Ref[72]</i>		SRC
	ml	ML	Original space	Squared standardization	
	1	98.1	96.1	95.6	
2	98.0	98.0	100	98.8	93.3
3	0.0	0.0	7.3	0	7.7
4	88.9	89.1	8.4	55.3	33.8
5	0.0	0.0	8	53.1	75.3
6	100	100	100	95.6	100
7	99.8	99.5	100	100	94.7
8	25.9	24.8	10	9.5	41.3
9	0.0	0.0	9.5	16.5	9.3
10	0.0	0.0	9.3	27.8	51.7
11	0.0	0.0	8.4	80.2	12.1
12	9.2	0.0	9.8	65.3	42.4
13	26.3	22.1	13.5	63.1	35.9
14	0.0	0.0	5	92.8	56.9
15	0.0	0.0	6.3	0	8.0
16	11.8	11.5	10.4	0	27.4
17	97.4	97.5	99.4	8.8	69.9
18	89.9	90.5	17.4	42.7	84.6
19	0.0	0.0	7.7	83.1	48.8
20	89.3	84.9	41.2	6.2	64.2
Average	41.7	40.7	33.4	49.9	52.6

为了更好的将本文提出的方法与其他方法进行一个比较。在此我们也采用 F1 index^{[40][70]}作为实验的评估指标。F1 index 是一种组合正确率和召回率的评价指标，它的具体定义如公式（4-10）：

$$Prec(f) = \frac{tp}{tp + fp} \quad (4-8)$$

$$Rec(f) = \frac{tp}{tp + fn} \quad (4-9)$$

$$F1(f) = \frac{2 * Prec(f) * Rec(f)}{Prec(f) + Rec(f)} \quad (4-10)$$

其中， tp 表示的被正确分类的正例数， fp 为被误分为正例的反例数， fn 表示的是被误分为反例的正例数， $Prec(f)$ 和 $Rec(f)$ 分别表示准确率和召回率。

表 4-4 是本文提出的方法与其他方法的一个对照比较表，从表中我们可以看出某些故障的 F1 得到了提高，如故障 3、5、8、10 等，某些故障则降低了，如故障 1、2、4 等。但是从总体平均水平来看，21 类故障 F1 index 平均值为 52.6% 相对于其他方法有所提高，效果不是特别的明显。

为了更好地展现稀疏表示算法的过程，我们分别在前两组实验中选取了若干测试样本实例对其重构系数和重构误差进行分析。如图 4-4 和 4-5 所示。在图 4-7 中，a(1)和 a(2)为实验 1 中选取的某个正常测试样本的重构系数和重构误差分布图，同样的 b(1)和 b(2)，c(1)和 c(2)，d(1)和 d(2)是分别为来自故障 1、6、18 测试样本。可以看出，当某类测试样本利用故障数据字典进行稀疏表示时，我们发现其重构系数主要集中在该测试样本所属类别的字典样本上面，其他类别字典样本上的重构系数大部分为 0 或者相对比较小。这与前文中所说的某一类的未知样本可以被该类的若干样本在某一子空间有效地进行线性表示的假设。而且由同类样本进行稀疏表示时，重构误差远远小于其他类别。图 4-5 为实验 2 中所选的若干测试样本，同样可以体现上述效果。

另外，为了确定组成故障样本数据集的每类样本数目 K 是否对分类效果有影响，我们也在前两组实验的基础上做了相关实验。我们依次从每个相关的故障训练数据集中选取 50、100、150、200、250、300、350、400、450 和 480 个样本构建故障样本数据集，在这些建立的故障样本数据集上我们进行了故障诊断实验，实现效果如图 4-6 和图 4-7 所示。从图上我们可以看出随着 K 值的变化，故

故障诊断的准确率基本没有太大波动，维持较高的准确率。这说明当我们构建故障样本数据集时不需要太多训练样本即可获得较好故障诊断效果，这在现实生活中很具有意义，因为故障数据并不是那么容易获得。

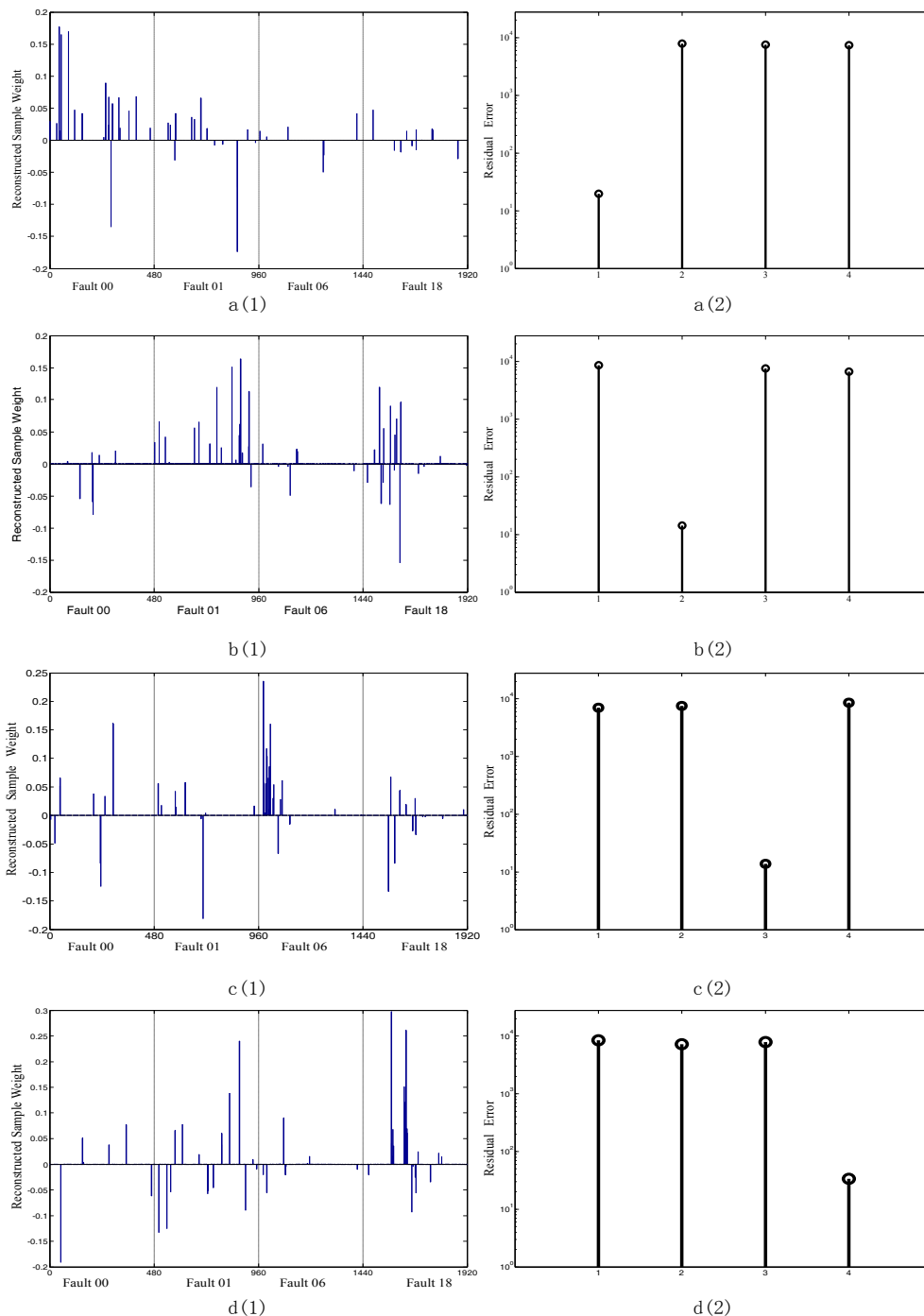


图 4-4 实验 1 中若干测试样本的重构系数和重构误差分布图

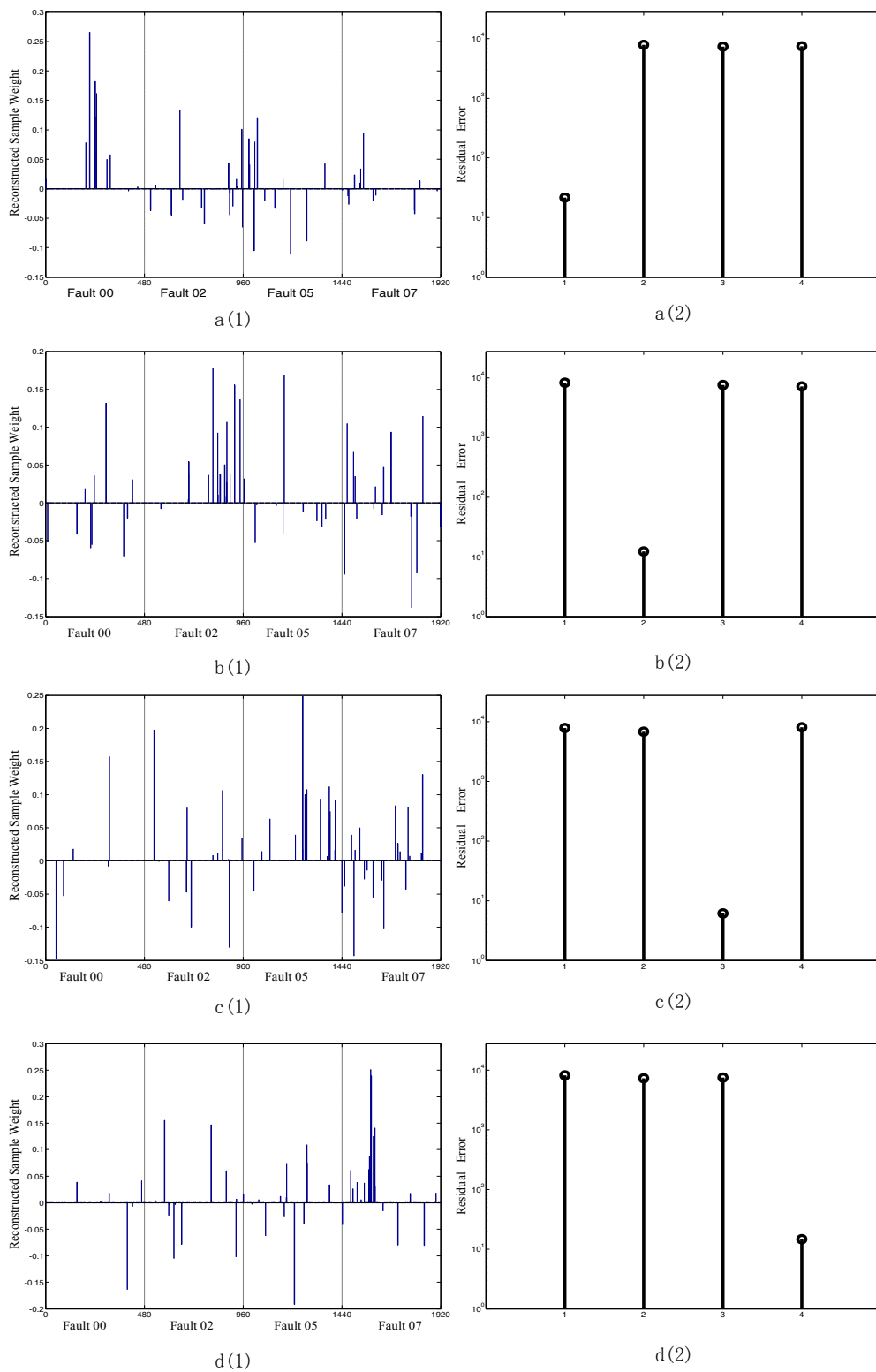


图 4-5 实验 2 中若干测试样本的重构系数和重构误差分布图

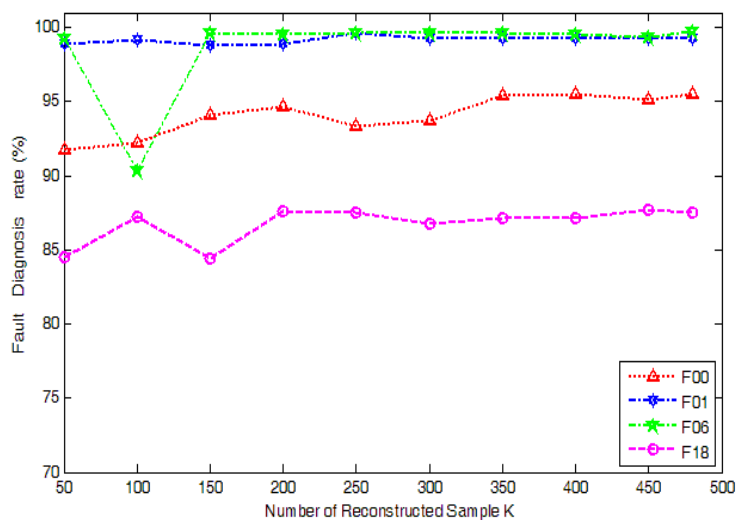


图 4-6 不同 K 值下实验 1 诊断效果图

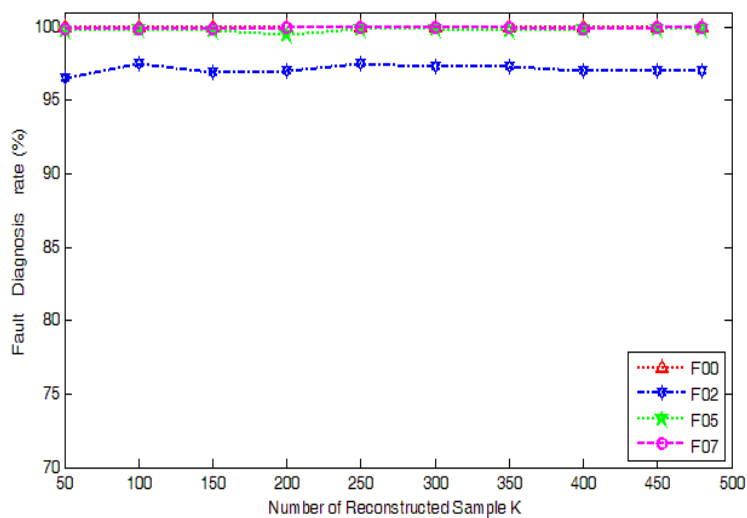


图 4-7 不同 K 值下实验 2 诊断效果图

4.3 基于动态特性的故障表示方法

现代企业中，高度自动化的流程生产系统通常由多个相互连接的设备组成，各设备之间的生产流程及相互影响关系复杂。为保证设备的无故障运行，需要利用各种传感器对设备的各种参数信息进行实时监控，由此积累了大量带有时间标

记的检测数据。如何从这些时间序列数据库中挖掘出故障所特有的动态特性信息，对于复杂系统的诊断与生产预警都是非常必要的。

另一方面在模式识别领域中，有效的特征表达对于分类效果起着至关重要的作用。在当前流行的故障诊断方法中，数据特征空间基本都是仅仅由相关传感器的测量值所组成的，对于每维数据在时间轴上都是假设独立的，然而这种假设在大多数的工业过程中都是不现实的，测量变量并不是序列无关的，当前时刻的测量变量与过去若干时刻的变量都有关系，具有一定的动态特性的。不同的故障产生时，同一变量所具有动态特性也是不一样的。因此，故障数据的动态特性有效提取和表示是提高故障诊断效果的有效途径之一^[38,40]。

在故障诊断领域对于动态特性的信息研究一直都在进行。早在 1995 年，Ku 等人将动态行为与 PCA 算法相结合，提出了 DPCA 算法^[32]。Li 和 Qin 在此基础上提出了一种基于 EVI-SMI(error-in-variables subspace model identification)的一致 DPCA 方法^[71]。2003 年 Lee 等人结合 sensor validity index(SVI)和 DPCA 提出了一种用于故障检测的 SVI^[72]。

前人在故障特性方面的研究表明设备检测数据动态特性信息的提取对于改善故障诊断效果起到促进作用。因此如何建立一种动态特性信息的特征表达方式是重要的。前文中所采用的动态信息加入的方法通常是在原始特征空间的基础上添加若干连续样本数据，这样做的直接后果是显著地提高了特征空间维数，加大数据量，增大了计算复杂度。本文引入了一种精简的动态特性表达方法。

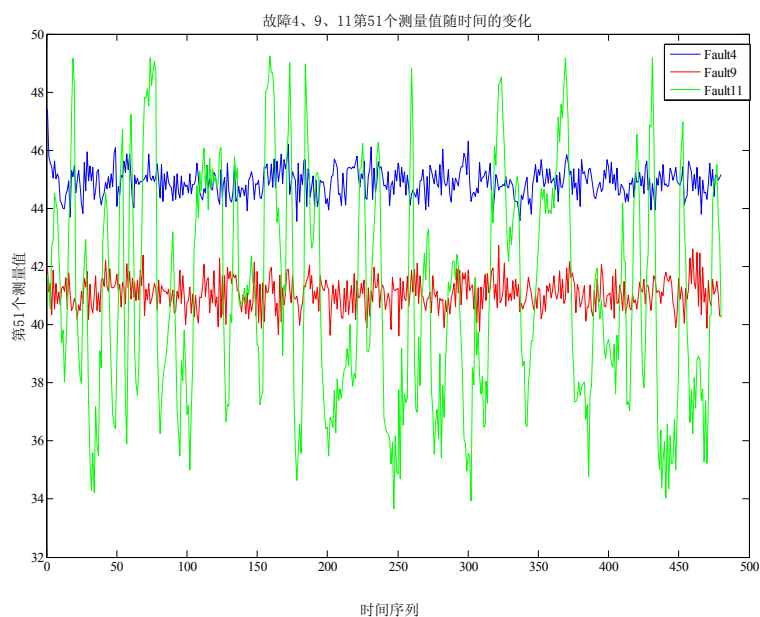


图 4-8 故障 4、9 和 11 训练集第 51 个观测变量随时间上变化情况

在对 TEP 进行过程仿真时，我们发现在对不同故障进行仿真时，设备某些指标检测数据的变化情况相差比较悬殊。其中故障 4、9 和 11 第 51 个观测变量在时间上的变化情就很具有代表性，如图 4-8 所示。

通过图 4-8 我们发现，当故障 11 发生时，第 51 个观测变量波动幅度相对于故障 4 和故障 9 发生时来说波动幅度比较大，故障 4 和 9 发生时第 51 测量值的波动幅度较小。这就说明不同故障发生时某些测量变量的波动情况会因故障而异。同时我们还可以看出，不同故障发生时，某些测量值的波动范围也不同。故障 4 发生时，第 51 个观测变量值基本在 45 上下小范围波动而故障 9 则是在 41 上下波动，两者不存在重叠。

考虑到前面介绍的样本添加方法所带来的问题，并结合上文的分析，本文基于以上的发现引入了称为动态特性时间窗统计方法用于动态信息的提取。该方法的主要原理是在一定时间窗口内，根据窗口内测量值信息统计出该时刻测量值在其维度上的方差信息和均值信息，利用这些统计信息对该样本点原始特征空间进行动态特性的加入。提取方差和均值的意义在于体现测量值在故障发生时所表现出的波动幅度信息和平均取值信息。这样做的好处有：一方面方差信息很好地表示出某一维度数据在时间轴上的波动信息；另一方面均值信息结合方差信息较好地反映测量值波动大致范围。图 4-9 和图 4-10 为按照时间窗统计方法所提取的故障 4、9 和 11 第 51 个测量值的方差信息和均值信息。

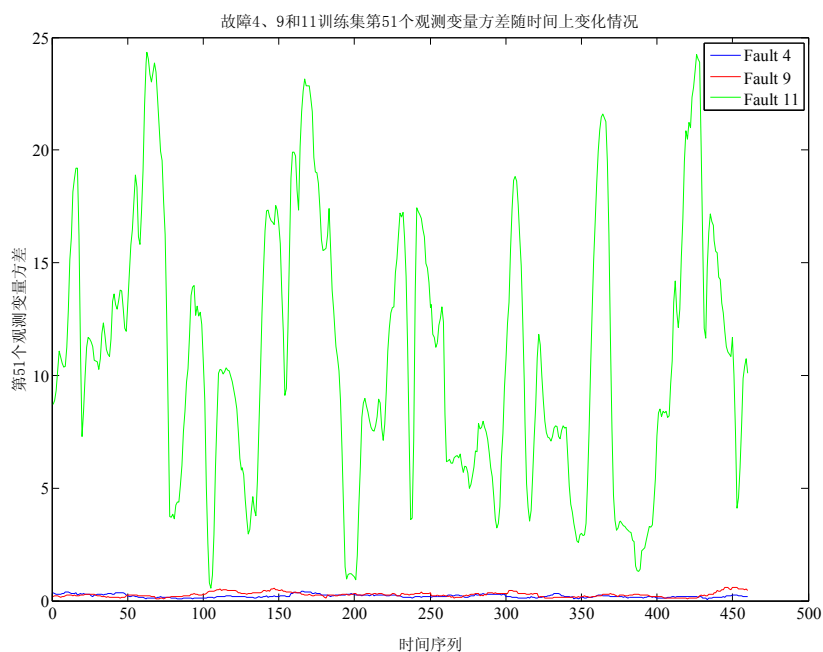


图 4-9 故障 4、9 和 11 训练集第 51 个观测变量方差随时间上变化情况

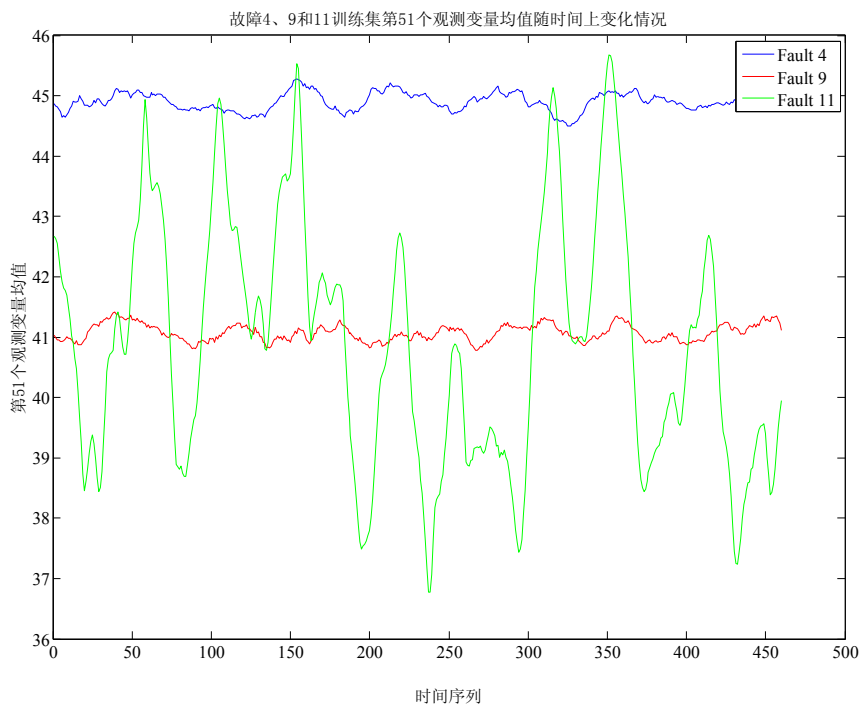


图 4-10 故障 4、9 和 11 训练集第 51 个观测变量均值随时间上变化情况

上图所设时间窗口大小为 10，可以清楚地故障 11 样本的第 51 个测量值大部分方差比较大，而且变化比较大；故障 4 和 9 的方差值则明显较小，基本接近于 0。均值变化图更能区别故障 4 和 9。可以看出这两个统计特征结合之后能够很好地区别不同故障的产生。因此，有效的提取这种区别性很大的动态信息对于故障诊断具有很大影响。将此方法提取的动态特性作为新的特征补充到原始特征空间，为故障诊断提供了更多的可判别信息，可以提高分类器的效果。

在此为了验证引入动态特性作为特征补充的诊断效果，我们对实验仿真训练集和测试集进行了动态特性信息加入，对 20 类故障进行了实验 4 与前文没有加入动态信息的实验进行对比。由于该方法是在原有特征空间上加入了动态特性信息，因为我们定义该方法为 DSRC (Dynamic SRC, DSRC)。结果如表 4-5 所示。

通过表 4-5 的实验结果可以看出，在引入了动态特性之后，基于稀疏表示的故障诊断效果得到较大的提高。平均故障诊断效果 F1 index 值由原来的 52.6% 提高到了 68.9%，效果比较显著。同类故障进行比较可以发现除了故障 6 有所下降外，剩余的故障诊断效果都有所改善，其中故障 14 和故障 19 相比之前 F1 index 值提高了 40% 以上。由此可以得出结论：引入动态特性作为分类特征的补充对于

基于稀疏表示的故障诊断方法有很大促进作用。

表 4-5 SRC 方法与 DSRC 实验效果对比

F1 index(%)	SRC	DSRC
1	94.3	95.6
2	93.3	99.1
3	7.7	22.7
4	33.8	68.4
5	75.3	84.0
6	100	79.3
7	94.7	100
8	41.3	69.8
9	9.3	13.4
10	51.7	54.1
11	12.1	37.3
12	42.4	76.3
13	35.9	38.7
14	56.9	98.7
15	8.0	17.5
16	27.4	48.8
17	69.9	95.9
18	84.6	87.6
19	48.8	98.0
20	64.2	93.5
Average	52.6	68.9

4.4 字典的学习优化

稀疏表示的本质就是利用超完备字典中的基元对测试样本进行线性组合。过完备字典的好坏直接决定了稀疏表示的效果，以至于决定了最后的分类效果，因此构建一个优异的冗余字典是非常重要的^[73]。前文实验所构造的字典数据直接来源于原始数据集，没有进行相关的优化工作，导致基于稀疏表示的分类优势未能完全体现出来。字典学习优化就是一种寻找稀疏表示下最优基的构造问题。不仅可以满足稀疏表示唯一性条件的约束同时也可获得更稀疏和更精确的表示。转化为目标函数为式子 4-11 的形式：

$$\arg \min_{\{D,W\}} \frac{1}{2} \|x - DW\|_F^2 + \lambda \|W\|_1 \quad (4-11)$$

由于上式目标函数不能同时对 D 和 W 进行凸优化，因此可以采用交替优化 D 和 W 的方式进行优化，该问题可以转化为以下两个子优化问题^[74]：

(1) 固定字典 \hat{D} 不变， W 凸优化问题。

$$W = \arg \min_W \frac{1}{2} \|x - \hat{D}W\|_F^2 + \lambda \|W\|_1 \quad (4-12)$$

(2) 固定 \hat{W} 不变， D 凸优化问题。

$$\begin{aligned} D &= \arg \min_D \frac{1}{2} \|x - D\hat{W}\|_F^2 \\ \text{subject to } \sum_i D_{ij}^2 &\leq c, \quad \forall j=1,2,\dots,K \end{aligned} \quad (4-13)$$

其中问题 (1)，本文采用 Feature-sign 查找法，Feature-sign 算法是一系列的符号猜测和更新的过程。每一个子过程中，猜测符号并赋值。详细的过程如算法 4-1 所示。

对于问题 (2)，我们采用 language dual 法，首先令 $\lambda = [\lambda_1, \dots, \lambda_K]$ ，其中 λ_i 为拉格朗日乘子，对应于不等式约束为 $\|D_i\|^2 - c \leq 0$ ，公式 4-14 的 language dual 函数为：

$$\begin{aligned} G(\lambda) &= \max_D L(D, \lambda) \\ &= \max_D \left(\|X - DW\|_F^2 + \sum_{i=1}^K \lambda_i (\|D_i\|^2 - c) \right) \end{aligned} \quad (4-14)$$

令 $\Lambda \in R^{K \times K}$ ，对角线元素 $\Lambda_{ii} = \lambda_i$ 。那么 $L(D, \lambda)$ 可以表示如下：

$$\begin{aligned} L(D, \lambda) &= \text{trace}(\|X - DW\|_F^2) + \text{trace}(D^T D \Lambda) - c * \text{trace}(\Lambda) \\ &= \text{trace}(X^T X) - 2 * \text{trace}(D^T X W^T) + \text{trace}(W^T D^T D W) + \\ &\quad \text{trace}(D^T D \Lambda) - c * \text{trace}(\Lambda) \end{aligned} \quad (4-15)$$

可以通过令上式的一阶导数为零，求得最优解 D^* 。

$$D W W^T - X W^T + D \Lambda = 0 \quad (4-16)$$

那么可得：

$$D^* = X W^T (W W^T + \Lambda)^{-1} \quad (4-17)$$

将上式 4-17 代入上式 4-15，language dual 函数变为：

$$\begin{aligned}
 g(\lambda) &= \text{trace}(X^T X) - 2 * \text{trace}(XW^T (WW^T + \Lambda)^{-1} WX^T) + \\
 &\quad \text{trace}(W^T D^T DW) + \text{trace}((WW^T + \Lambda)^{-1} WX^T XW^T) - c * \text{trace}(\Lambda) \quad (4-18) \\
 &= \text{trace}(X^T X) - \text{trace}(XW^T (WW^T + \Lambda)^{-1} WX^T) - c * \text{trace}(\Lambda)
 \end{aligned}$$

进一步可以得到以下 language dual 函数为：

$$\begin{aligned}
 \min_{\Lambda} &\text{trace}(XW^T (WW^T + \Lambda)^{-1} WX^T) - c * \text{trace}(\Lambda) \\
 \text{subject to } &\lambda_i \geq 0, i = 1, \dots, K.
 \end{aligned} \quad (4-19)$$

上式可以通过共轭梯度法或牛顿法求解。如果 Λ^* 为最优解，那么： $D^* = XW^T (WW^T + \Lambda)^{-1}$ 。值得注意的是 $WW^T + \Lambda$ 不能够保证可逆，因此实际应用中伪逆代替直接求逆运算。

算法 4-1 Feature-sign 查找算法

输入：输入信号 $x \in R^d$ ，超完备字典 $D \in R^{d \times K}$ ，稀疏参数 λ

输出：稀疏系数 $w \in R^K$

1. 初始化： $w = 0$ （或者提出初始值）

2. 外循环 while

2.1 选择 $i = \arg \max_i |D_i^T D_i w_i - D_i^T x|$, where $w_i = 0$

如果， $D_i^T D_i w_i - D_i^T x > \lambda$ ，那么， $w_i = \frac{\lambda - (D_i^T D_i w_i - D_i^T x)}{D_i^T D_i}$

如果， $D_i^T D_i w_i - D_i^T x < -\lambda$ ，那么， $w_i = \frac{-\lambda - (D_i^T D_i w_i - D_i^T x)}{D_i^T D_i}$

如果， $|D_i^T D_i w_i - D_i^T x| \leq \lambda$ ，那么，退出外循环

2.2 内循环 while

1) 获取当前非零系数集合 $p \in R^m$, where $m \in [1, K]$, $w_{pj} \neq 0, j = 1, 2, \dots, m$

2) 获取非零系数对应的 $w_a = w(p), x_a = x(p), D_a = D(p)$ （列为主）

3) 求解解析解 $w_{new} = (D_a^T D_a)^{-1} (D_a^T x_a - \lambda \theta_a)$

4) 判断， w_a 与 w_{new} 之间是否有符号不一样的元素？

a. 没有，更新 $w_{p(j)} = w_{newj}$, where $j = 1, \dots, m$ ，然后退出内循环。

b. 有, 执行闭合线段 (从 w_a 到 w_{new}) 离散查找过程

2.3 直到满足条件 a, 结束内循环

3. 直到满足条件 $|D_i^T D_i w_i - D_i^T x| \leq \lambda$ where $\forall w_i = 0$, 结束外循环

通过以上两步不断迭代, 即可实现字典的优化学习。为了验证字典优化的有效性, 我们在原有的实验基础之上对构造的字典进行了学习, 并对训练集和测试集进行了格式转换。具体的实验结果如表 4-6 所示。

表 4-6 SRC、SRC_L、DSRC 和 DSRC_L 实验效果对比

F1 index(%)				
	SRC	SRC_L	DSRC	DSRC_L
1	94.3	94.1	95.6	93.8
2	93.3	95.4	99.1	97.5
3	7.7	8.1	22.7	20.0
4	33.8	40.0	68.4	66.1
5	75.3	56.4	84.0	81.5
6	100	100	79.3	99.5
7	94.7	100	100	99.5
8	41.3	40.0	69.8	61.2
9	9.3	30.5	13.4	18.1
10	51.7	40.7	54.1	55.6
11	12.1	20.8	37.3	36.1
12	42.4	45.2	76.3	80.9
13	35.9	41.3	38.7	41.7
14	56.9	61.4	98.7	98.0
15	8.0	12.6	17.5	17.4
16	27.4	28.8	48.8	50.0
17	69.9	84.5	95.9	94.4
18	84.6	81.9	87.6	87.1
19	48.8	57.3	98.0	97.5
20	64.2	66.2	93.5	90.2
Average	52.6	53.9	68.9	69.2

在表 4-6 中, SRC_L 表示原始数据字典经过学习优化后的新方法, DSRC_L 则表示加入动态特性的数据字典经过学习优化后的新方法。可以看出故障数据字典经过优化之后再用于故障诊断时, 诊断效果都有了一定的提高。但是两者提高的效果并不很明显。分析原因是由于在进行学习过程中所设优化参数比较单一, 没有对所设参数进行调整, 如果参数设置合适, 字典优化学习之后的诊断效果一

定会有较大提高。因此参数的设置问题也是下一步的研究重点。

4.5 本章小结

本章主要详细介绍了所提出的基于动态特性与稀疏表示故障诊断方法。首先通过若干公式和示意图对基于稀疏表示的故障诊断方法原理进行了详细地阐述。之后以 TEP 为实验平台，在仿真数据集上进行了部分实验工作对基于稀疏表示的故障诊断方法进行验证。然后简单介绍了引入动态特性特征信息的方法，同时进行了对比实验。最后，对字典学习优化问题的方法原理进行的数学描述和实验验证。

总结与展望

本文首先简述了本课题的研究背景及意义、国内外研究现状和目前研究的热点问题。针对当前故障诊断领域内存在的若干问题，本文提出了一种基于动态特性与稀疏表示的故障诊断方法，主要做了三个方面的工作。一是故障诊断的方式进行了改进。本文采用的基于稀疏表示的故障诊断方法，既不同于多变量统计分析方法的统计变量监控，也不同于一般模式分类方法。基于稀疏表示的分类器在故障诊断方面具有以下优点：1.模型统一，基于稀疏表示的分类方法对于多类问题只需一种分类模型即可解决多类问题。2.节省训练时间，基于稀疏表示的分类方法不需要进行分类器训练，不需要调试各种参数，只需要分类的各类选出一定的故障训练数据组成重构样本字典用于测试样本的重构，获得重构系数和重构残差进行故障诊断。3.不需要过多的训练样本数据，实验证明基于稀疏表示的分类效果并不与构成重构样本字典样本数目成正比，在组成字典的重构样本较少时所获得的分类效果既已达到较好效果。这大大解决了现实生活中故障训练样本数据较少的难题。4.不需要进行特征提取。二是在基于稀疏表示的故障诊断方法上充分考虑了化工过程中故障测量数据动态差异性，通过引入动态特征作为分类特征补充有效的改善了诊断效果。三是对字典的学习优化问题进行了研究，并采用 Feature-sign 查找法和 language dual 算法对字典进行了优化。本文所提方法在 TEP 仿真数据集合上得到了有效验证。

虽然本文所提出的方法有效地改善了故障诊断效果。但是基于稀疏表示的方法往往需要通过一个超完备基来求得测试样本的稀疏表示，当数据量比较大时，算法的计算机复杂度成为限制其优良性能的瓶颈，导致 SRC 无法进行实时识别，这对于故障诊断是致命的。因此，如何提出一张简便有效的改进算法，大大减小计算机复杂度成为下一步的研究重点。所引入的动态特性方法比较简单，还存在一些缺陷不足以表示全面的动态，更加完善的动态特性信息表示方法也是亟待解决问题。字典优化有助于 SRC 的分类，但是参数的设置问题需要不断探索。

参考文献

- [1] 刘强, 柴天佑, 秦泗钊, 赵立杰. 基于数据和知识的工业过程监视及故障诊断综述[J]. 控制与决策, 2010,25(6):801-807.
- [2] 李钢, 秦泗钊, 周东华. 基于多尺度 PCA 的工业过程故障预测[J]. 华中科技大学学报, 2009,37(Sup I):32-35.
- [3] 胡玉成. 基于统计特征提取的故障诊断方法研究[D]. 杭州电子科技大学硕士学位论文, 2011.
- [4] 王新明. 基于 TE 的化工过程故障诊断算法研究[D]. 兰州理工大学硕士学位论文, 2010.
- [5] Frank P M. Fault diagnosis in dynamics systems using analytical and knowledge based redundancy: a survey and some new results [J]. Automatic, 1990,26(3):459-474.
- [6] Venkatasubramanian V, Rengaswamy R and Kavuri S N. A review of process fault detection and diagnosis Part II: qualitative models and search strategies[J]. Computers and Chemical Engineering.2003, 27(3):313-326.
- [7] Venkatasubramanian V, Rengaswamy R and Kavuri S N, Yin K. A review of process fault detection and diagnosis Part III: process history based methods [J]. Computers and Chemical Engineering.2003,27(3):327-346
- [8] 周东华, 胡艳艳. 动态系统的故障诊断技术[J]. 自动化学报,2009, 35(6): 748-758.
- [9] Iri M, Aoki K, O'Shima E and Matsuyama H. An algorithm for diagnosis of system failures in the chemical process[J]. Computers and Chemical Engineering, 1979, 3(1): 489-493.
- [10] Caceres S, Henley E J. Process failure analysis by block diagrams and fault trees. Industrial and Engineering Chemistry, Fundamentals, 1976, 15(2): 128-134
- [11] Leung D, Romagnoli J. Dynamic probabilistic model based expert system for fault diagnosis[J], Computers and Chemical Engineering,2000,24(11):2473-2492.
- [12] Scenna N J. Some aspects of fault diagnosis in batch processes [J]. Reliability Engineering and System Safety, 2000, 71(1):95-110.

- [13] Kulpers B. Qualitative simulation as causal explanation[J]. IEEE Transactions on Systems, Man and Cybernetics, 1987, 17(3):432-444.
- [14] Thukaram D, Khincha H P, Vijaynarasimha H P. Artificial neural network and support vector machine approach for locating faults in radial distribution systems[J]. Power Delivery, IEEE Transactions on, 2005, 20(2): 710-721.
- [15] Wise B M, Ricker N L, Veltkamp D F, Kowalski B R. A theoretical basis for the use of principal component models for monitoring multivariate processes[J]. Process Control and Quality, 1990, 1(1): 41-51.
- [16] Li W, Yue H H, Valle-Cervantes S, et al. Recursive PCA for adaptive process monitoring[J]. Journal of process control, 2000, 10(5): 471-486.
- [17] Kano M, Tanaka S, Hasebe S, et al. Monitoring independent components for fault detection[J]. AIChE Journal, 2003, 49(4): 969-976.
- [18] Drago R J. Incipient failure detection[J]. Power Transmission Design, 1979, 21(2): 40-45.
- [19] Sun Q, Tang Y. Singularity analysis using continuous wavelet transform for bearing fault diagnosis[J]. Mechanical Systems and Signal Processing, 2002, 16(6): 1025-1041.
- [20] Sun Wei Xiang, Chen Jin, Wu Xing, Dong Guang Ming, Ning Zuo Gui, Wang Dong Sheng. Early loosening fault diagnosis of clamping support based on information fusion[J]. Journal of Shanghai Jiaotong University, 2006, 40(2):239-242.
- [21] Lu F, Du H L, Diao Z J, et al. Rotor fault diagnosis based on fusion estimation of multi-circuit model of induction motor[C]//Machine Learning and Cybernetics, 2004. Proceedings of 2004 International Conference on. IEEE, 2004, 4: 2157-2161.
- [22] 程宇, 王武, 崔福军. 基于模型的故障诊断方法研究综述[C]. In: Proceedings of the 27th Chinese Control Conference, 2008, Kunming, China: 17-20.
- [23] 张强. 基于系统输入/输出数据的预测控制器性能监控[D]. 上海:上海交通大学, 2004.
- [24] MacGregor J F. Using On - Line Process Data to Improve Quality: Challenges for Statisticians*[J]. International Statistical Review, 1997, 65(3): 309-323.

- [25] Chiang L H, Russell E H and Braatz R D. Fault diagnosis in dynamic systems[M]. Prentice Hall, Englewood Cliffs, NJ, 1989.
- [26] 张杰, 阳宪惠. 多变量统计过程控制[M]. 北京: 化学工业出版社, 2000.
- [27] Jackson J E, Mudholkar G S. Control procedures for residuals associated with principal component analysis[J]. *Technometrics*, 1979, 21(3): 341-349.
- [28] Nomikos P, MacGregor J F. Monitoring batch processes using multiway principal component analysis[J]. *AIChE Journal*, 1994, 40(8): 1361-1375.
- [29] Nomikos P, MacGregor J F. Multivariate SPC charts for monitoring batch processes[J]. *Technometrics*, 1995, 37(1): 41-59.
- [30] Dunia R, Joe Qin S. Subspace approach to multidimensional fault identification and reconstruction[J]. *AIChE Journal*, 1998, 44(8): 1813-1831.
- [31] Kano M, Hasebe S, Hashimoto I, et al. A new multivariate statistical process monitoring method using principal component analysis[J]. *Computers & chemical engineering*, 2001, 25(7): 1103-1113.
- [32] Ku W, Storer R H, Georgakis C. Disturbance detection and isolation by dynamic principal component analysis[J]. *Chemometrics and intelligent laboratory systems*, 1995, 30(1): 179-196.
- [33] Cho J H, Lee J M, Wook Choi S, et al. Fault identification for process monitoring using kernel principal component analysis[J]. *Chemical engineering science*, 2005, 60(1): 279-288.
- [34] Lee J M, Yoo C K, Lee I B. Statistical monitoring of dynamic processes based on dynamic independent component analysis[J]. *Chemical Engineering Science*, 2004, 59(14): 2995-3006.
- [35] Lee J M, Qin S J, Lee I B. Fault Detection of Non - Linear Processes Using Kernel Independent Component Analysis[J]. *The Canadian Journal of Chemical Engineering*, 2007, 85(4): 526-536.
- [36] 蒋丽英, 王树青. 基于MPCA-MDPLS 的间歇过程的故障诊断[J]. *化工学报*, 2005, 56(3): 482-486.
- [37] Balla R, Misra M. Multi-scale Fisher discriminant analysis[C]//2005AIChE Annual Meeting and Fall Showcase, New York. 2005.
- [38] Chiang L H, Kotanchek M E, Kordon A K. Fault diagnosis based on Fisher

- discriminant analysis and support vector machines[J]. Computers & chemical engineering, 2004, 28(8): 1389-1401.
- [39] Zhu Z B, Song Z H. A novel fault diagnosis system using pattern classification on kernel FDA subspace[J]. Expert Systems with Applications, 2011, 38(6): 6895-6905.
- [40] Monroy I, Benitez R, Escudero G, et al. A semi-supervised approach to fault diagnosis for chemical processes[J]. Computers & Chemical Engineering, 2010, 34(5): 631-642.
- [41] Wu L, Chen X, Peng Y, et al. Fault Detection and Diagnosis based on Sparse Representation Classification (SRC)[C]. Robotics and Biomimetics, ROBIO 2012, IEEE International Conference on. IEEE, 2012: 926-931.
- [42] Zhu Z B, Song Z H. Fault diagnosis based on imbalance modified kernel Fisher discriminant analysis[J]. Chemical Engineering Research and Design, 2010, 88(8): 936-951.
- [43] 郭明. 基于数据驱动的流程工业性能监控与故障诊断研究 [D]. 浙江大学博士学位论文, 2004.
- [44] Candès E J, Romberg J, Tao T. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information[J]. Information Theory, IEEE Transactions on, 2006, 52(2): 489-509.
- [45] Donoho D L. Compressed sensing[J]. Information Theory, IEEE Transactions on, 2006, 52(4): 1289-1306.
- [46] Baraniuk R G. Compressive sensing [lecture notes][J]. Signal Processing Magazine, IEEE, 2007, 24(4): 118-121.
- [47] DeVore R A. Deterministic constructions of compressed sensing matrices[J]. Journal of Complexity, 2007, 23(4): 918-925.
- [48] Tsaig Y, Donoho D L. Extensions of compressed sensing[J]. Signal processing, 2006, 86(3): 549-571.
- [49] 李卓凡, 闫敬文. 压缩感知及应用[J]. 微计算机应用, 2010, 31(3): 12-16.
- [50] 韩红平. 压缩感知中信号重构算法的研究[D]. 南京邮电大学硕士学位论文, 2012.
- [51] 李立. 基于稀疏表示的人脸图像识别方法研究[D]. 南京理工大学硕士学位

- 论文, 2012.
- [52] Paredes J L, Arce G R, Wang Z. Ultra-wideband compressed sensing: channel estimation[J]. *Selected Topics in Signal Processing, IEEE Journal of*, 2007, 1(3): 383-395.
- [53] Elad M, Bruckstein A M. A generalized uncertainty principle and sparse representation in pairs of bases[J]. *Information Theory, IEEE Transactions on*, 2002, 48(9): 2558-2567.
- [54] Mallat S G, Zhang Z. Matching pursuits with time-frequency dictionaries[J]. *Signal Processing, IEEE Transactions on*, 1993, 41(12): 3397-3415.
- [55] Tropp J A, Gilbert A C. Signal recovery from random measurements via orthogonal matching pursuit[J]. *Information Theory, IEEE Transactions on*, 2007, 53(12): 4655-4666.
- [56] Needell D, Vershynin R. Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit[J]. *Selected Topics in Signal Processing, IEEE Journal of*, 2010, 4(2): 310-316.
- [57] Rebollo-Neira L, Lowe D. Optimized orthogonal matching pursuit approach[J]. *Signal Processing Letters, IEEE*, 2002, 9(4): 137-140.
- [58] Friedman J H, Stuetzle W. Projection pursuit regression[J]. *Journal of the American statistical Association*, 1981, 76(376): 817-823.
- [59] Ma J. Single-pixel remote sensing[J]. *Geoscience and Remote Sensing Letters, IEEE*, 2009, 6(2): 199-203.
- [60] Taubock G, Hlawatsch F. A compressed sensing technique for OFDM channel estimation in mobile environments: Exploiting channel sparsity for reducing pilots[C]//*Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*. IEEE, 2008: 2885-2888.
- [61] Wakin M B, Laska J N, Duarte M F, et al. An architecture for compressive imaging[C]//*Image Processing, 2006 IEEE International Conference on*. IEEE, 2006: 1273-1276.
- [62] A.Y. Yang, J. Wright, Y. Ma, and S. S. Sastry, Robust Face Recognition via Sparse Representation [J]. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 31, No. 2, 2009.

- [63] Mairal J, Elad M, Sapiro G. Sparse representation for color image restoration[J]. Image Processing, IEEE Transactions on, 2008, 17(1): 53-69.
- [64] Han Z, Jiao J, Zhang B, et al. Visual object tracking via sample-based adaptive sparse representation (AdaSR)[J]. Pattern Recognition, 2011, 44(9): 2170-2183.
- [65] Li C, Han Z, Ye Q, et al. Abnormal Behavior Detection via Sparse Reconstruction Analysis of Trajectory[C]//Image and Graphics (ICIG), 2011 Sixth International Conference on. IEEE, 2011: 807-810.
- [66] Ricker and Lee. Tennessee Eastman Plant-wide Industrial Process Challenge Problem[J]. Computers and Chemical Engineering, 1995, 19(9): 983-1005
- [67] Laukonen E G, Passino K M, Krishnaswami V, et al. Fault detection and isolation for an experimental internal combustion engine via fuzzy identification[J]. Control Systems Technology, IEEE Transactions on, 1995, 3(3): 347-355.
- [68] Candes E J, Romberg J K, Tao T. Stable signal recovery from incomplete and inaccurate measurements[J]. Communications on pure and applied mathematics, 2006, 59(8): 1207-1223.
- [69] Kašin B S. Diameters of some finite-dimensional sets and classes of smooth functions[J]. Mathematics of the USSR-Izvestiya, 1977, 11(2): 317.
- [70] Yélamos I, Graells M, Puigjaner L, et al. Simultaneous fault diagnosis in chemical plants using a multilabel approach[J]. AIChE Journal, 2007, 53(11): 2871-2884.
- [71] W. Li, S.J. Qin, Consistent dynamic PCA based on error-in-variables subspace identification,[J]. Process Control 11 (2001) 661-678
- [72] Lee C, Choi S W, Lee I B. Sensor fault identification based on time-lagged PCA in dynamic processes[J]. Chemometrics and intelligent laboratory systems, 2004, 70(2): 165-178.
- [73] 苏杭. 过完备字典下的稀疏信号重构研究[D]. 武汉理工大学博士论文, 2012.
- [74] Lee H, Battle A, Raina R, et al. Efficient sparse coding algorithms[J]. Advances in neural information processing systems, 2007, 19: 801.

作者简介与论文发表

姓名：武利军 性别：男 出生日期：1987.11.11 籍贯：山东曲阜

2010.9-2013.7 中国科学院大学 计算机应用技术 硕士

2006.9-2010.7 中国矿业大学（北京） 计算机科学与技术 学士

已发表文章目录

- **Lijun Wu**, Xiaogang Chen, Yi Peng, Qixiang Ye, Jianbin Jiao. Fault Detection and Diagnosis based on Sparse Representation Classification [C]. *Proceedings of IEEE International Conference on Robotics and Biomimetics, 2012 (ROBIO)*, (EI)

- Yi Peng, Qixiang Ye, Jianbin Jiao, Xiaogang Chen and **Lijun Wu**, "Fault Diagnosis via Structural Support Vector Machines," Proc.of IEEE International Conference on Mechatronics and Automation(ICMA), 2012.

附录 1: 主要符号对照表

SPE	平方预测误差 (Square Predicted Error)
PCA	主元分析 (Principal Components Analysis)
FDA	判别分析 (Fisher Discriminant Analysis)
ICA	独立分量分析 (Independent Component Analysis)
PLS	偏最小二乘 (Partial Least Square)
SR	稀疏表示 (Sparse Representation)
BP	基追踪 (Basis Pursuit)
MP	匹配追踪 (Matching Pursuit)
OMP	正交匹配追踪算法 (Orthogonal Matching Pursuit)
ROMP	正则化正交匹配追踪 (Regularized Orthogonal Matching Pursuit)
OOMP	最优正交匹配追踪 (Optimized Orthogonal Matching Pursuit)
SRC	稀疏表示分类 (Sparse Representation Classification)
TEP	田纳西-伊斯曼过程 (Tennessee Eastman process)
DPCA	动态主元分析 (Dynamic Principal Components Analysis)

致 谢

在中国科学院大学攻读硕士学位三年的学习生活中，我经历了诸多坎坷，也付出了艰辛的努力，同时也得到了很大的收获。在毕业论文完成之际，由衷地感谢这三年来曾经给予我无数帮助的老师、同学、朋友和家人。

首先，我要感谢导师焦建彬教授在我硕士期间每一次的指导和鼓励，感谢他在论文撰写和修改中倾注的心血。三年来，焦老师对我的学习给予了悉心的指导，为我的研究提供了良好的环境和机会。他严谨的治学态度和平易近人的学者风范深深地影响着我，他诲人不倦的精神让我感动，这些言传身教将使我受益终生。

其次，感谢叶齐祥老师、韩振军老师和陈孝罡师兄。他们在我的理论学习和实验研究中给予了耐心的指导和帮助，并在论文的撰写和审稿中付出了辛苦的劳动。他们具有渊博的专业知识，自强不息的学习精神和扎实的动手能力，我从他们的谆谆教导中学到了一些令我受益匪浅的可贵知识。

衷心感谢我的师兄师姐和同学们，感谢我们一起学习生活中结下的深厚友谊，特别是张立国师兄、彭艺师姐、李策师姐，好友张晓丹、杨威、高山、梁吉祥和邢宇龙。三年中，他们曾和我一起面对成长中的喜悦和困惑，一起书写这段美好回忆。在即将毕业之际，我想感谢他们一路上的相伴，真心希望我们不会因为毕业的分离而隔断了友情。

特别要感谢我的家人，感谢他们多年来一直给我最无私的关爱与支持，永远是我最坚强的后盾和避风的港湾。感谢他们为我所做的一切，愿他们能够为我而骄傲，祝福他们健康长寿。亲人们对我的支持和理解指引我奋发拼搏、积极向上，我将终生铭记，时刻激励自己不断进步。

感谢参加开题及中期评阅的各位老师和专家们，他们曾利用丰富的经验来无私地帮助我把握论文方向和研究进度。

最后，感谢参加论文评审和答辩的各位老师。

武利军

2013年5月

