

基于压缩感知的设备多源信息传输与分类算法

赵小虎^{1,2}, 王刚^{1,2}, 宋泊明³, 于嘉成¹

(1. 矿山互联网应用技术国家地方联合工程实验室, 徐州 江苏 221000;
2. 中国矿业大学物联网(感知矿山)研究中心, 徐州 江苏 221000; 3. 西澳大学电气电子与计算机工程系, 珀斯 6009)

摘 要: 针对选煤厂设备种类繁多、监测点分散的特点, 提出了一种基于压缩感知的设备多源信息无线传输与分类算法。通过构建一种多跳信息传输模型, 将信息传输问题转换为多路测量信号的压缩感知问题, 将测量矩阵获取问题转化为多跳信息传输模型的路由问题。针对所获得的测量矩阵存在较大相干性、影响信号重构效果问题, 将随机路由的思想引入路由构建当中, 提出了一种随机动态自组织路由算法。为了解决重构后的信号时域特征难以对故障类型进行精确分类的问题, 针对重构信号, 引入了一种新的时域特征——振动信号的全变分, 并通过补偿距离评估算法(CDET)验证了引入指标的优越性。通过选煤厂实测数据分析表明, 所提多源信息传输与分类算法在提高监测数据实时传输效率情况下, 能够有效提高故障识别精度。

关键词: 压缩感知; 多跳网络; 随机路由; 全变分; 故障分类

中图分类号: TP18

文献标识码: A

doi: 10.11959/j.issn.1000-436x.2020040

Multi-source information transmission and classification algorithm for equipment based on compressed sensing

ZHAO Xiaohu^{1,2}, WANG Gang^{1,2}, SONG Boming³, YU Jiacheng¹

1. The National Joint Engineering Laboratory of Internet Applied Technology of Mines, Xuzhou 221000, China

2. IoT Perception Mine Research Center, China University of Mining and Technology, Xuzhou 221000, China

3. Department of Electrical, Electronic and Computing Engineering, The University of Western Australia, Perth 6009, Australia

Abstract: Aiming at the characteristics of various types of equipment in coal preparation plant and the dispersion of monitoring points, a multi-source information wireless transmission and classification algorithm for equipment based on compressed sensing was proposed. By constructing a multi-hop information transmission model, the information transmission problem was transformed into the compressed sensing problem of multi-path measurement signals, thereby the measurement matrix acquisition was transformed into the routing problem of the multi-hop information transmission model. Aiming at the large coherence of the obtained measurement matrix and affecting the signal reconstruction effect, the idea of random routing was introduced into the routing construction, and a random dynamic self-organizing routing algorithm was proposed. In order to solve the problem that the time domain features of the reconstructed signal were difficult to accurately classify the fault type, a new time domain feature, the total variation (TV) of the vibration signal, was introduced for the reconstructed signal, and the compensation distance estimation algorithm was adopted to verify the superiority of the introduction of indicators. The analysis of the measured data of the coal preparation plant shows that the proposed multi-source information transmission and classification algorithm can effectively improve the fault recognition accuracy under the condition of improving the real-time transmission efficiency of the monitoring data.

Key words: compressed sensing, multi-hop network, random routing, total variation, fault classification

收稿日期: 2019-10-28; 修回日期: 2019-12-12

通信作者: 王刚, wgcumt@yeah.net

基金项目: 国家重点研发计划基金资助项目 (No.2017YFC0804404)

Foundation Item: The National Key Research and Development Project of China (No.2017YFC0804404)

1 引言

由于选煤厂设备种类繁多, 监测点分散, 建立传统的有线监测系统存在较大的困难, 基于自组和多跳路由的无线传感器网络为选煤厂设备状态监测提供了灵活的解决思路。设备振动信号由于包含丰富的设备运行信息, 常用于设备故障诊断。然而, 设备故障诊断中振动信号频带较宽, 依据传统香农采样定理进行数据采集时, 会得到大量振动数据, 给实时设备监测中的存储、传输和处理带来很大困难。压缩感知 (CS, compressed sensing)^[1] 由于利用了信号的稀疏性, 能以远低于奈奎斯特采样速率对信号进行非自适应测量编码, 突破了香农采样定理的瓶颈, 使高分辨率信号的采集成为可能。因此, 本文研究适用于选煤厂设备振动信号的压缩感知与分类算法。

Fazel 等^[2]提出了一种用于大规模传感器网络中长期数据收集的随机压缩感知方案, 该方案能够延长传感器网络的寿命。针对分布式压缩与重构问题, 田子建等^[3]研究了一种基于分布式压缩感知构造指纹数据库的方法, 减少了数据采集工作量, 提高了工作效率。Wang 等^[4]结合压缩感知和多跳网络用于移动瓦斯测量, 实现了瓦斯测量数据的重构。然而, 利用压缩感知理论采集选煤厂设备振动数据时, 由于监测点分散, 监测点向无线接入点数据传输采用多跳通信方式, 测量矩阵的构成依赖于多跳传输时所选择的路由技术。由于每次随机选择 M 条不同链路传输数据, 而不同路由策略形成不同测量矩阵, 从而影响不同重构效果, 因此路由选择策略是选煤厂振动信号压缩感知实际使用时迫切需要解决的问题。

为了解决拥塞、时延、能耗等问题, Wang 等^[5]提出了一种基于蚁群算法的服务质量 (QoS, quality of service) 路由协议。兰巨龙等^[6]提出了基于深度强化学习的软件定义网络 QoS 优化算法, 将网络资源和状态信息统一到网络模型中, 通过长短期记忆网络提升算法的流量感知能力, 最后基于深度强化学习生成满足 QoS 目标的动态流量调度策略。Gelenbe 等^[7]将随机路由算法用于基于无线网络的事件监测, 为监测事件提供更短的时延和更好的 QoS。

结合选煤厂设备的监测需求, 本文利用自组织特征映射 (SOM, self-organizing map) 神经网络将链路通信质量汇聚到路由构建中, 保证了信息传输

质量, 同时, 提出了一种基于服务质量的随机动态自组织路由技术, 解决了压缩感知中测量矩阵存在较大相关性的问题, 提高了信号重构精度。

对压缩重构的选煤厂设备运行信号进行状态分类时, 传统故障诊断方法是构建设备振动信号混合域特征向量集, 然而高维特征向量存在的冗余信息会对后续模式分析造成维数灾难, 影响故障分类精度。Yang 等^[8]提出将距离评估技术应用于特征降维和旋转机械设备故障分类。Lei 等^[9]将距离评估技术改进为一种降维性能更优越的补偿距离评估技术 (CDET, compensation distance evaluation technique)。王前等^[10]将噪声信号梅尔倒谱与 CDET 结合, 用于滚动轴承故障诊断, 取得了良好的聚类结果。以上研究成果表明, CDET 能够反映不同特征的敏感性特征权重。采用敏感特征组成的特征集能够提高故障模式识别的准确精度。

在分类模型选择上, 基于数据驱动的机器学习算法取得了更好的诊断效果。人工神经网络 (ANN, artificial neural network)^[11-12]和支持向量机 (SVM, support vector machine)^[13]应用于旋转机械设备的故障诊断均可获得较高的故障分类精度。而 SVM 作为小样本的机器学习方法, 分类模型训练时间远少于 ANN, 特别适用于工业应用场景, 并且对旋转机械设备故障类型识别率更高^[14]。

因此, 为解决混合域特征集内特征维数较高、特征间相关性严重的问题, 本文以压缩重构的多通道振动信号时域特征为基础, 引入振动信号的全变分, 将补偿距离评估算法与支持向量机结合, 提出了一种基于全变分的滚动轴承数据驱动特征提取与故障分类算法, 并从以下两方面验证了该算法的优越性: 1) CDET 算法优于传统的降维算法, 如主成分分析 (PCA, principal component analysis)、线性判别式分析 (LDA, linear discriminant analysis)、非线性主元分析 (KPCA, kernel PCA) 等, 在特征筛选中能优先选出全变分特征, 表明全变分具有描述轴承振动信号时域特征的优良特征; 2) 引入全变分后的特征集能够提高故障的识别精度。

2 多跳通信与压缩感知

压缩感知主要包括信号稀疏表示、测量矩阵设计与重构算法三部分。信号稀疏表示是压缩感知的先决条件。长度为 N 的 K 稀疏信号 \mathbf{x} , 测量矩阵 $\Phi \in R^{M \times N} (M \ll N)$, 当 \mathbf{x} 稀疏时, 测量值 $\mathbf{y} (\mathbf{y} \in R^M)$

可由 $y = \Phi x$ 获得。当 x 非稀疏时，若 x 在基底 $\Psi(N \times N)$ 下是稀疏的，稀疏信号为 s ，则测量值 $y = \Phi x = \Phi \Psi s$ 。 Φ 的每一行可以看作一个传感器阵列，它与信号相乘，拾取了信号的一部分信息。

振动监测信号从一个无线采集点传输到另一个无线采集点，为了延长生存时间，常采用多跳通信方式，如图 1 所示。

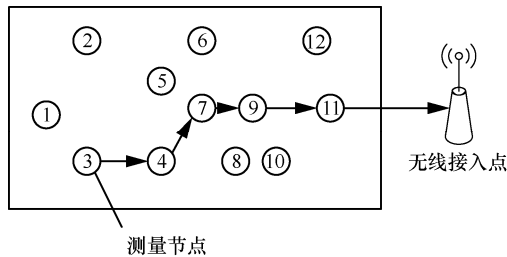


图 1 多跳无线网络

数据采集采用时间独立的 M 个具有不同起始节点的多跳通信链路，对于每一个多跳网络，测量节点随机选择一个初始节点，并计算振动测量值与自身产生的随机比例系数乘积，根据选定的路由将该乘积传送给下一个测量节点。而下一个测量节点将接收到的信号与本地产生的乘积信号相加，再传送给下一个测量节点，以此类推。这样，无线接入点接收到的第 i 路多跳信号的测量值为^[15-16]

$$y_i = \sum_{j \in P_i} \Phi_{i,j} x_j \quad (1)$$

其中， P_i 是在第 i 条多跳通信链路中经过的测量节点序号集合， $\Phi_{i,j}$ 是第 j 个测量节点产生的随机系数， x_j 是第 j 个测量节点测量到的振动数据，则在无线接入点接收到的 M 路信号向量 $y = [y_1, \dots, y_M]$ 可表示为

$$y = \Phi \Psi s \quad (2)$$

这样，数据多跳传输就转变为压缩感知问题。其中， Φ 是 $M \times N$ 阶测量矩阵，当 $j \in P_i$ 时，第 (i, j) 元素的值等于 $\Phi_{i,j}$ ；否则为 0。显然测量矩阵 Φ 与监测信息传输的路由技术相关。

3 随机动态自组织路由

对于由 N 个测量节点构成的无线传感器网络，网络中每个测量节点有若干条通信链路同其他节点相连，网络可以用图的方式描述为 $G(V, E)$ 。其中， $V = \{v_1, v_2, \dots, v_N\}$ 是网络测量节点集，它对应于图中

每一个顶点； $E = \{w_{ij}\}$ 是网络链路的集合，每条链路对应一条弧 (v_i, v_j) ， w_{ij} 为该弧的权值或距离。测量网络结构如图 2 所示。在无线传感网络里，假设所有链路均是对称的，即 $w_{ij} = w_{ji}$ 。

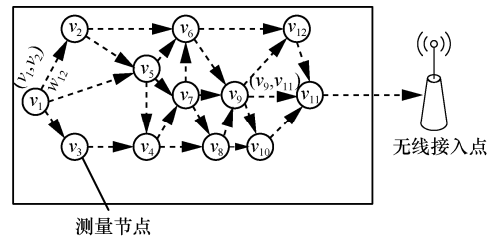


图 2 测量网络结构

由于测量矩阵由网络中采用的路由算法确定，当采用最短路径树、贪婪路由算法测量数据流传输时，测量矩阵会存在较大的相干性，影响信号重构效果。因此，引入随机路由的思想。

考虑到选煤厂通信环境复杂，在构造最小路径树时，可将 QoS 定义为通信链路的权值。影响服务质量的因素包括时延、数据传输率、误码率、带宽等，采用 SOM 神经网络将以上因素映射为链路 QoS，即链路距离。SOM 神经网络由两层神经元组成，第一层为输入层，作为传感器感知信息的输入端，这里是描述传感器邻居链路间 QoS 优劣的指标；第二层为竞争层。运行过程分为训练和执行 2 个阶段。由于训练阶段需要较高的计算开销，因此需要在上位机上完成，称为离线处理；相反，执行阶段不需要高计算开销，可在每个传感器节点上在线运行，因此可以把 SOM 神经网络应用到传感器网络节点。最后，利用最小路径树构造节点传输路径，称为动态自组织路由。随机动态自组织路由是在最小路径树基础上增加路由随机扩展。由于原有路由在构造路径时每次都选择节点到汇聚节点的最小距离，使节点路由唯一，出现测量节点只存在一个前向节点或后向节点的情况，不能满足压缩感知对测量矩阵的要求。因此，在最小路径树基础上，通过扩展路由增加前向节点，并对每一个前向节点赋予一定选择概率，使路由选择在保证 QoS 的基础上又具有一定随机性，解决了测量矩阵存在较大相关性的问题。具体实现方法如下。

定义 $d(v_i)$ 为节点 v_i 到汇聚节点的距离， $\Gamma(v_i)$ 为节点 v_i 单跳可到达的邻域节点集合， r 为汇聚节点， w_{ij} 为节点 i 与 j 的距离，其值由链路 QoS 确定。假设所有节点链路间 QoS 已通过 SOM 神经网络汇

聚计算得到。

1) 节点距离初始化

当节点距离超过通信范围时,将距离 d 设为 ∞ 。

$$\text{令 } d(r)=0, \quad d(v_i)=\begin{cases} w_{ri}, & v_i \in \Gamma(r) \\ \infty, & v_i \notin \Gamma(r) \end{cases}$$

2) 最小路径树构建

在集合 T 中,寻找到汇聚节点距离最小的节点 v_j ,更新 T 中元素, $T=T-\{v_j\}$ 。对任意 $v_i \in T \cap \Gamma(v_j)$,计算 $d_i=d(v_j)+QoS$ 。若 $d_i < d(v_i)$, $d(v_i)=d_i$ 。

3) 如果 $|T| > 0$,重复步骤 2); 如果 $|T|=0$,进入步骤 4)。

4) 路由扩展

对最小路径树中的每一节点 v_i ,若只有一个前向节点,则在其 $\Gamma(v_i)$ 任意选择一个除已有前向节点 v_f 和后向节点 v_h 之外任一 w_{ij} 最小的节点 v_j ,进行路由扩展,并令节点 v_i 的下一跳选择节点 v_f 和 v_j 的概率分别为 p_1 和 p_2 , $p_1 > p_2$,且 $p_1 + p_2 = 1$ 。若节点 v_i 已有 2 个或以上前向节点,则不再进行路径的前向扩展,并平均分配每条支路一个选择概率。

获得了信号传输采用的路由策略,即获得了多跳传输时压缩感知的测量矩阵,最后根据接收到的信息即可通过测量矩阵重构出原始测量信号。

4 信号重构与时域处理

4.1 信号重构

信号重构算法是指由长度为 M 的测量向量 y 重构长度为 N 的稀疏信号 x 的过程。信号的重构问题可表示为 l_0 范数的最小优化问题,如式(3)所示。

$$\begin{aligned} \hat{s} &= \operatorname{argmin} \|s\|_0 \\ \text{s.t. } y &= \Phi x = \Phi \Psi s \end{aligned} \quad (3)$$

其中, $\|\cdot\|_0$ 为向量的 l_0 范数。最小 l_0 范数问题是一个 NP-hard 问题。鉴于此,研究人员提出了一系列求次最优解的算法,主要包括正交匹配追踪 (OMP, orthogonal matching pursuit) 算法^[17]、基追踪 (BP, basis pursuit) 算法^[18]、广义正交匹配追踪 (gOMP, generalized OMP) 算法^[19]、稀疏度自适应匹配追踪 (SAMP, sparsity adaptive MP) 算法^[20]、分段弱正交匹配追踪 (SWOMP, stagewise weak OMP) 算法^[21] 和子空间追踪 (SP, subspace pursuit) 算法^[22] 等。

4.2 振动信号特征提取

对重构的每一路测试信号进行时频分析,获取信号特征。传统的信号处理方法包括时域分析、频域分析、小波分组变换等。为了反映信号变化情况,本节参考文献[23-24]提取信号的均值、标准差、均方根值、歪度指标、峭度指标、波形指标、裕度指标、脉冲指标、峰值指标等传统时域特征,并构建了一种新的时域特征——振动信号的全变分。

设 x 为轴承的振动信号,由于正常运行下机器的振动或噪声一般是大量的、无规则的、大小接近的随机扰动的结果;而对于不正常运行状态下的振动信号,通常是在随机信号中出现有规则的周期性脉冲,即满足 $x_i \approx x_{i+1}$ 。因此,定义振动信号 $x \in R^N$ 的全变分如式(4)所示。

$$\|Dx\|_1 = \sum_{i=1}^{N-1} |D_i x| = \sum_{i=1}^{N-1} |x_{i+1} - x_i| \quad (4)$$

其中, $D \in R^{(N-1) \times N}$ 为双对角矩阵, D_i 为双对角矩阵 D 的第 i 行元素。

$$D = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & -1 & 1 \end{bmatrix} \quad (5)$$

全变分反映了一个采样周期内信号时域相邻采样点的总体变化情况。在正常运行下,振动信号的全变分近似为微小量,而对于不正常运行状态下的振动信号,其全变分是一个较大的量,因此全变分函数对于快速变化的 x 给予大的值,更能反映有变化的故障信号的统计时域特征。这在信号的稀疏采样中也获得了验证^[19]。

作为对比,本文选择了一般文献中常用的 9 个信号时域特征,组成了本文轴承振动信号所选择的 10 个时域特征及其排列顺序,作为后续特征筛选的源特征,如表 1 所示,其中 μ_x 为均值, σ_x 为标准差。

4.3 补偿距离评估算法

设含有 C 个模式类 $\omega_1, \omega_2, \dots, \omega_C$ 的特征集为 $\{p_{c,m,k}\}$, $c=1,2,\dots,C$, $m=1,2,\dots,M_c$, $k=1,2,\dots,K$ 。其中, $p_{c,m,k}$ 为状态 c 下第 m 个样本的第 k 个特征, M_c 为状态 c 下样本总数, K 为每个样本特征数目。因此总共可得 $\sum_{c=1}^C M_c$ 个样本集,共计 $K \sum_{c=1}^C M_c$ 个特征。

表1 振动信号的10个时域特征及其表达式

序号	特征名称	计算式
1	均值	$\sqrt{\frac{1}{N} \sum_{i=1}^N x_i}$
2	标准差	$\sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu_x)^2}$
3	均方根值	$\sqrt{\frac{1}{N} \sum_{i=1}^N (x_i)^2}$
4	歪度指标	$\frac{\sum_{i=1}^N (x_i - \mu_x)^3}{N\sigma_x^3}$
5	峭度指标	$\frac{\sum_{i=1}^N (x_i - \mu_x)^4}{N\sigma_x^4}$
6	波形指标	$\frac{\sqrt{\frac{1}{N} \sum_{i=1}^N (x_i)^2}}{\frac{1}{N} \sum_{i=1}^N x_i }$
7	裕度指标	$\frac{\max x_i }{\left(\frac{1}{N} \sum_{i=1}^N \sqrt{ x_i }\right)^2}$
8	脉冲指标	$\frac{\max x_i }{\frac{1}{N} \sum_{i=1}^N x_i }$
9	峰值指标	$\frac{\max x_i }{\sqrt{\frac{1}{N} \sum_{i=1}^N (x_i)^2}}$
10	全变分	$\sum_{i=1}^N x_{i+1} - x_i $

补偿距离评估算法实现特征降维的具体步骤如下。

1) 计算 $\omega_c (c=1,2,\dots,C)$ 类中所有特征向量的平均距离，如式(6)所示。

$$d_{c,k} = \frac{1}{M_c(M_c - 1)} \sum_{i=1}^{M_c} \sum_{j=1}^{M_c} |p_{c,i,k} - p_{c,j,k}| \quad (6)$$

对 $d_{c,k} (c=1,2,\dots,C)$ 求平均后得到平均类内距离，如式(7)所示。

$$d_k^\omega = \frac{1}{C} \sum_{c=1}^C d_{c,k} \quad (7)$$

2) 定义并计算 d_k^ω 的方差因子，如式(8)所示。

$$v_k^\omega = \frac{\max(d_{c,k})}{\min(d_{c,k})} \quad (8)$$

3) 计算 C 个模式类的类间距离，如式(9)所示。

$$d_k^b = \frac{1}{C(C-1)} \sum_{c=1}^C \sum_{e=1}^C |\mu_{c,k} - \mu_{e,k}| \quad (9)$$

其中， $\mu_{c,k} = \frac{1}{M} \sum_{m=1}^{M_c} p_{c,m,k}$ 为 ω_c 类中所有的 k 个特征

均值。

4) 定义并计算 d_k^b 的方差因子，如式(10)所示。

$$v_k^b = \frac{\max |\mu_{e,k} - \mu_{c,k}|}{\min |\mu_{e,k} - \mu_{c,k}|} (c, e = 1, 2, \dots, c \text{ 且 } c \neq e) \quad (10)$$

5) 定义并计算补偿因子，如式(11)所示。

$$\gamma_k = \frac{1}{\frac{v_k^\omega}{\max(v_k^\omega)} + \frac{v_k^b}{\max(v_k^b)}} \quad (11)$$

6) 计算类间距离 d_k^b 与类内距离 d_k^ω 的比值，得到距离评估指标 α_k 。

$$\alpha_k = \gamma_k \frac{d_k^b}{d_k^\omega} \quad (12)$$

对 α_k 进行归一化处理，得到归一化的距离评估指标为

$$\bar{\alpha}_k = \frac{\alpha_k}{\max(\alpha_k)} \quad (13)$$

7) 设阈值 $\varphi (\varphi \in [0,1])$ ，从特征集 $\{p_{c,m,k}\}$ 中选择距离评估指标 $\bar{\alpha}_k \geq \varphi$ 特征为灵敏特征；或设目标维度为 L ，选取距离评估指标最高的 L 个特征为敏感特征，将 K 维降为 L 维。

5 算法流程

基于压缩感知的选煤厂设备多源监测信息传输与分类算法流程如图3所示，具体步骤如下。

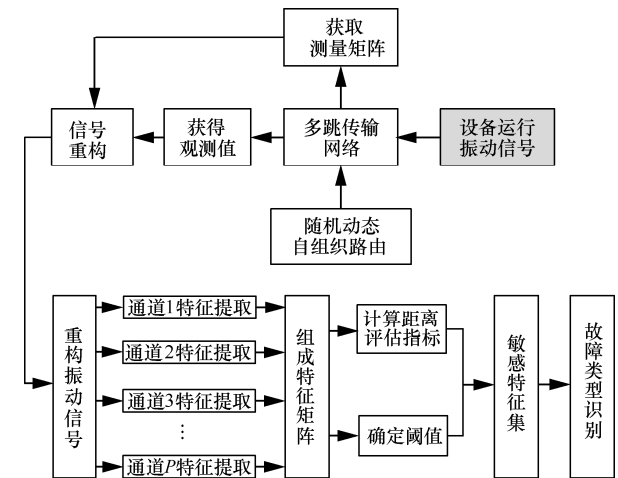


图3 算法流程

1) 利用随机动态自组织路由部署选煤厂设备多元监测信息多跳传输网络，获取观测值 y 。

2) 根据多跳传输网络获取压缩感知中的测量矩阵 Φ 。

3) 信号重构, 获取重构测量信号 \mathbf{x} 。

4) 根据重构获得的 C 个模式类型的 P 个通道振动信号, 将每种故障类型振动信号分别按时间序列分帧处理组成 M_c 个样本。

5) 引入全变分, 提取不同故障类型样本的每个通道时域特征。其中每个通道提取 10 个时域特征, 10 个特征排序按表 1 顺序排列。 P 个通道共提取 K 个特征, 即 $K = 10P$ 。

6) 样本特征按列排列, 组成特征集 $\{p_{c,m,k}\}$, 表示为

$$\{p_{c,m,k}\} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1N} \\ a_{21} & a_{22} & \dots & a_{2N} \\ \dots & \dots & \dots & \dots \\ a_{K1} & a_{K2} & \dots & a_{KN} \end{bmatrix} \quad (14)$$

其中, N 为 C 个模式类的样本数总和, 特征集元素 a_{ij} 表示第 j 个样本第 p 个通道的第 k 个特征值, $i = (k-1)P + p, j = 1, \dots, N, p = 1, \dots, P, k = 1, \dots, 10$ 。

7) 采用 CDET 算法, 根据式(13)计算时域特征集补偿距离指标 $\overline{\alpha}_k$, 并确定相应阈值 φ , 选取敏感特征。

8) 确定敏感特征集, 通过“一对一”的多分类 SVM 对设备故障类型进行分类, 确定其具体故障类型。

从信号重构开始, 各步骤均在上位机中完成。

6 实验测试

6.1 随机动态自组织路由

轴承是机械设备易损件之一, 因此在设备状态监测时, 对选煤厂每台设备采用 2 个三轴振动加速度传感器, 分别测取 X、Y、Z 这 3 个轴向的振动加速度信号, 传感器的具体位置如图 4 所示。其中, 驱动端的 3 个轴向分别用 T_1 、 T_2 和 T_3 表示, 非驱动端的 3 个轴向分别用 T_4 、 T_5 和 T_6 表示。

图 5 描绘了根据图 2 场景, 采用不同路由由协议时测量数据传输路径树的构造结果。图 5(a)为根据 SOM 神经网络计算出节点间链路质量的带 QoS 的网络带权图, 箭头上数字表示链路质量 QoS; 图 5(b)为基于 QoS 的多权值动态自组织路由技术, 每条路由保证节点到汇聚节点累积 QoS 最小, 即链路质量最高; 图 5(c)为随机动态自组织路由, 在图 5(b)基础

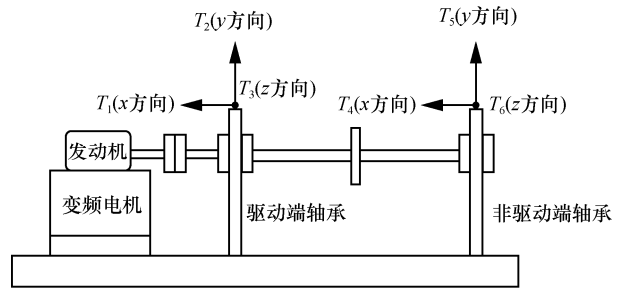
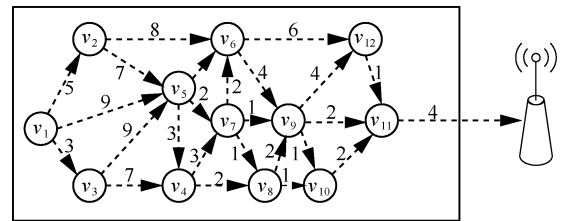
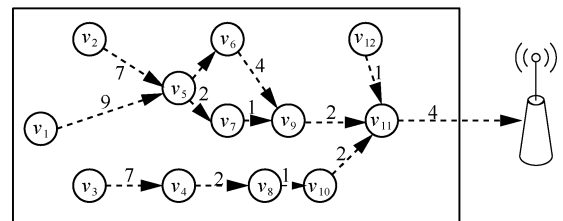


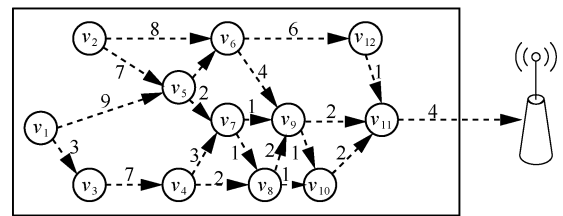
图 4 传感器的具体位置



(a) 网络带权图



(b) 多权值动态路由



(c) 随机动态自组织路由

图 5 几种路由方式

上进行了路由扩展, 对只有一个前向节点的路由节点, 在可能的情况下扩展了一个前向节点, 并对多个前向节点进行了概率分配, 从而使节点到汇聚节点的路径树不唯一, 增加了路径选择的随机性。

为了测试路由效果, 随机产生了 12 路测试信号, 每一路信号稀疏度为 2, 恢复算法采用 BP 算法。图 6 所示为分别采用多权值动态路由和随机动态自组织路由由技术构建测量矩阵的重建效果。其中, 横轴为采样点数; 纵轴为加速度幅值, 单位为 g 。由图 6 可知, 采用多权值动态路由时, 由于测量矩阵的确定性, 12 路信号中只有不到一半的路数信号可以精确重构; 而采用随机动态自组织路由时, 由于测量矩阵的随机性, 数据重建效果变好。

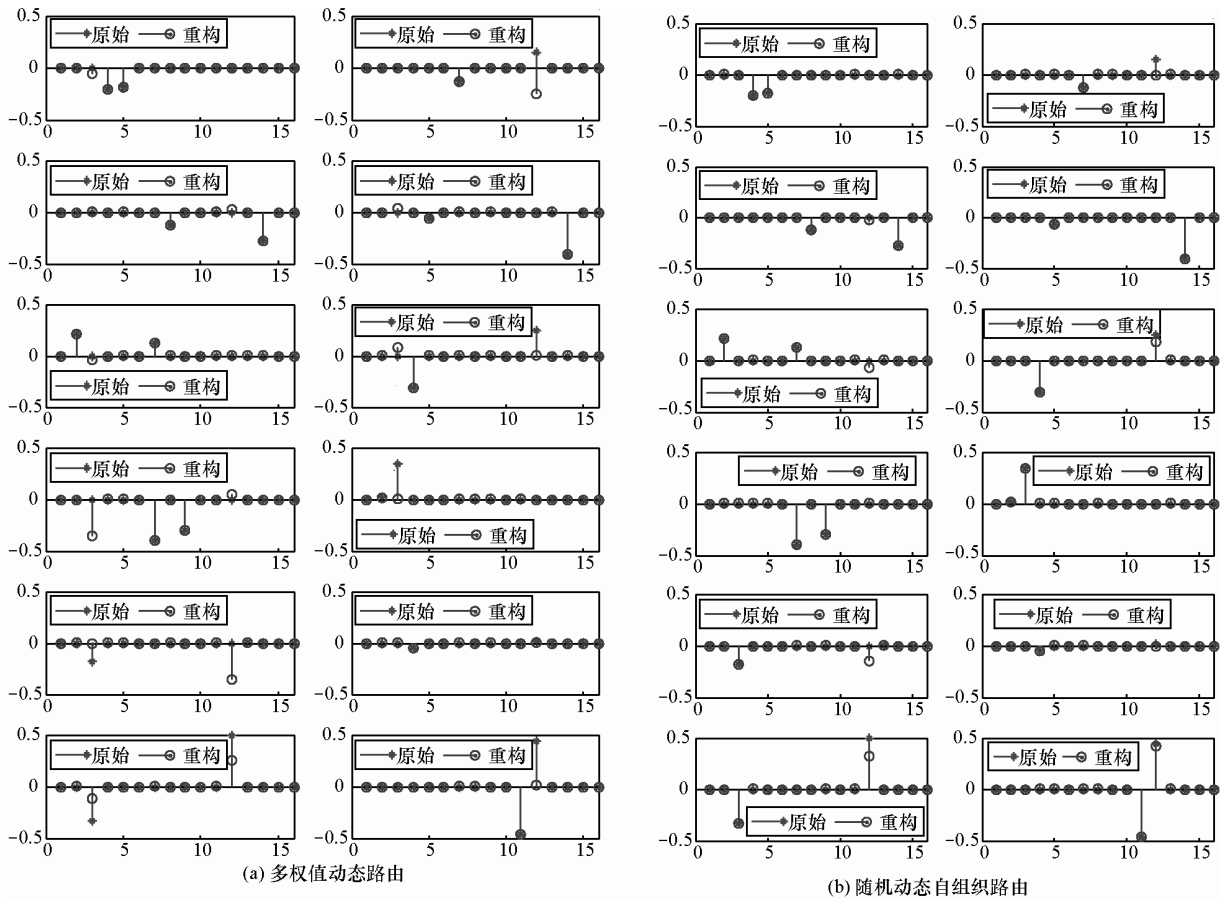


图 6 2 种路由技术构建测量矩阵的重建效果

6.2 重构的测量信号

具体工业实验时，采集 1 800 r/min 转速下 5 种轴承工作状态（正常、内圈故障、外圈故障、滚动

体故障、组合故障）实验数据。不同设备轴承工作状态采样周期都为 1 min，采集得到 6 个轴向振动传感器的 5 种运行状态数据集。图 7 所示为轴 T_1

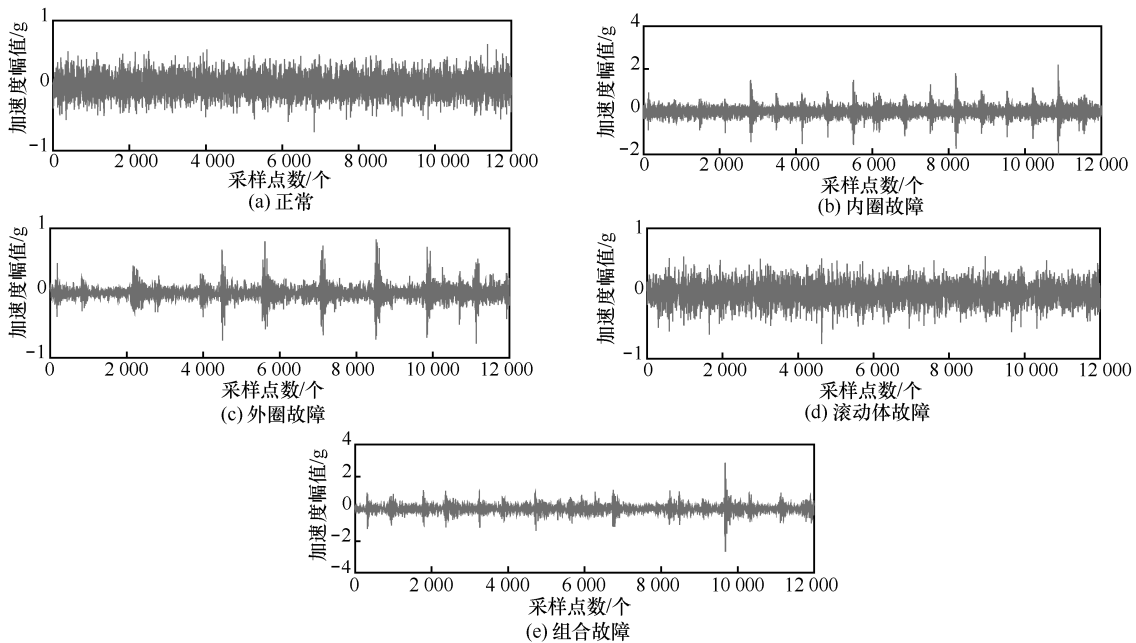


图 7 轴 1 方向 5 种轴承运行状态

方向 5 种轴承运行状态，其中横轴为采样点数，纵轴为加速度幅值。

图 8 显示了随机动态自组织路由下，采用不同重构算法压缩比与重构误差结果。由图 8 可见，在相同压缩比时，OMP、gOMP 以及 SAMP 算法整体性能相近。

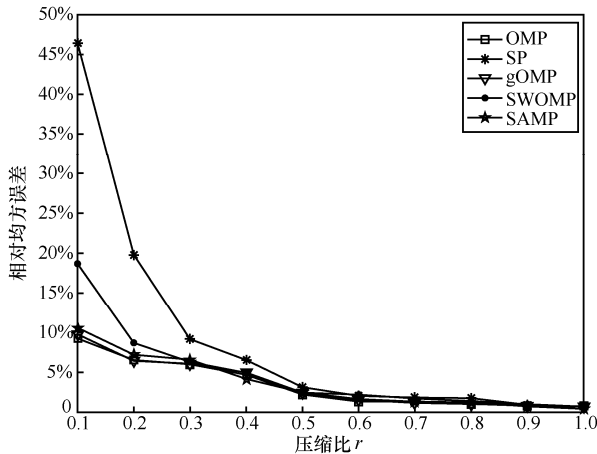


图 8 不同重构算法压缩比与重构误差关系

以正常状态信号为例，图 9 为 $N=256$ 长轴承信号在压缩比为 0.3 时利用随机动态自组织路由在 OMP、gOMP、SAMP 和 SP 重构算法下的时域波形对比。由于重构的测量信号主要应用于后续设备工

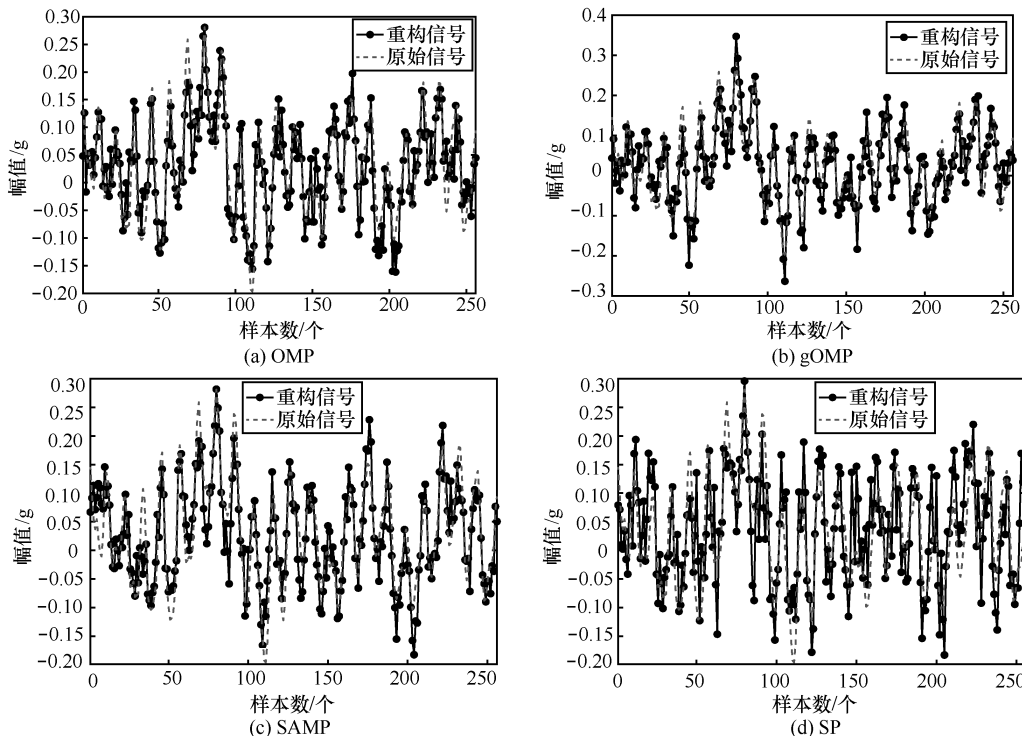


图 9 正常状态信号在不同重构算法下重构效果

作状态分类，因此不需要过于在意重构时域误差。

6.3 特征提取

利用随机动态自组织路由作为测量矩阵，采用 OMP 算法对多源监测信号进行重构。重构信号的数据处理针对每个方向不同工作状态提取 200 组数据为训练集，提取 150 组数据为测试集，每组 3 000 个数据点。提取每组数据均值、标准差、均方根值、歪度指标、峭度指标、波形指标、裕度指标、脉冲指标、峰值指标、全变分共 10 个特征构成测试特征集，6 轴向传感器共构成 60 个特征集。

6.4 无全变分特征后特征筛选与分类情况

1) 基于 CDET 筛选出的时域特征

通过补偿距离评估算法得到没有振动信号全变分特征时的 54 个特征距离评估指标 $\bar{\alpha}_k$ ($k=1,2,\dots,54$)，为了使目标维数为 3，本文将阈值 ϕ 设为 0.6，即 $\bar{\alpha}_k > 0.6$ 的指标被选为特征指标，用于下一步 SVM 故障类型分类。图 10 为特征补偿距离评估，选择的 3 个特征分别为轴 4 方向振动信号的均方根值，以及轴 1 和轴 3 方向的振动信号平均值。

2) 不同降维算法分类情况比较

根据补偿距离评估算法得到的三维特征，与 PCA、LDA、KPCA 降维得到的三维特征分别应用

SVM 进行故障分类。图 11 为各降维算法降维后的聚类效果。表 2 为无全变分各降维算法对应的分类精度。

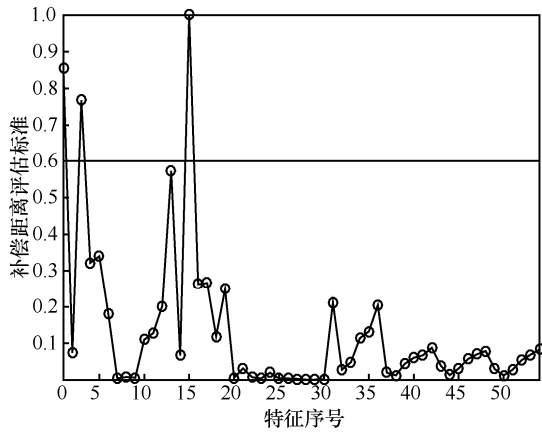


图 10 特征补偿距离评估

表 2 无全变分各降维算法对应的分类精度

降维算法	SVM 分类精度
LDA	94%
PCA	96%
KPCA	97.73%
CDET	99.33%

由表 2 可知，PCA 降维后的数据集分类精度约

为 96%；KPCA 降维方法处理非线性数据集降维效果较好^[16]，分类精度较高，但其故障分类结果要比补偿距离评估算法降维后分类精度低 1.6%；LDA 分类精度最低约为 94%；而 CDET 的分类准确度达到 99.33%。由此可见，当目标维数为 3 时，采用补偿距离评估算法得到的三维特征已可达到较高的分类精度，说明目标维数为三或阈值 ϕ 为 0.6 满足实际应用需求。

6.5 考虑全变分特征后特征筛选与分类情况

1) 基于 CDET 筛选出的时域特征

当引入本文振动信号全变分特征后，对 60 维数据集进行降维，计算 60 个特征距离评估指标 $\bar{\alpha}_k$ ($k=1,2,\dots,60$)，设阈值 $\phi=0.6$ ，即 $\bar{\alpha}_k > 0.6$ 的指标被选为特征指标，图 12 为引入全变分特征集补偿距离评估，选择的 3 个特征分别为轴 4 和轴 5 方向的振动信号全变分，以及轴 4 方向振动信号的均方根值。

根据表 1 以及 K 个样本特征排列顺序，由图 12 可见，CDET 筛选出的 3 个时域特征中 2 个均为全变分特征，而且轴 4 和轴 5 方向的振动信号全变分补偿距离评估值均大于未引入全变分的所有特征的补偿距离评估值，表明全变分具有描述轴承振动信号时域特征的优良特征。

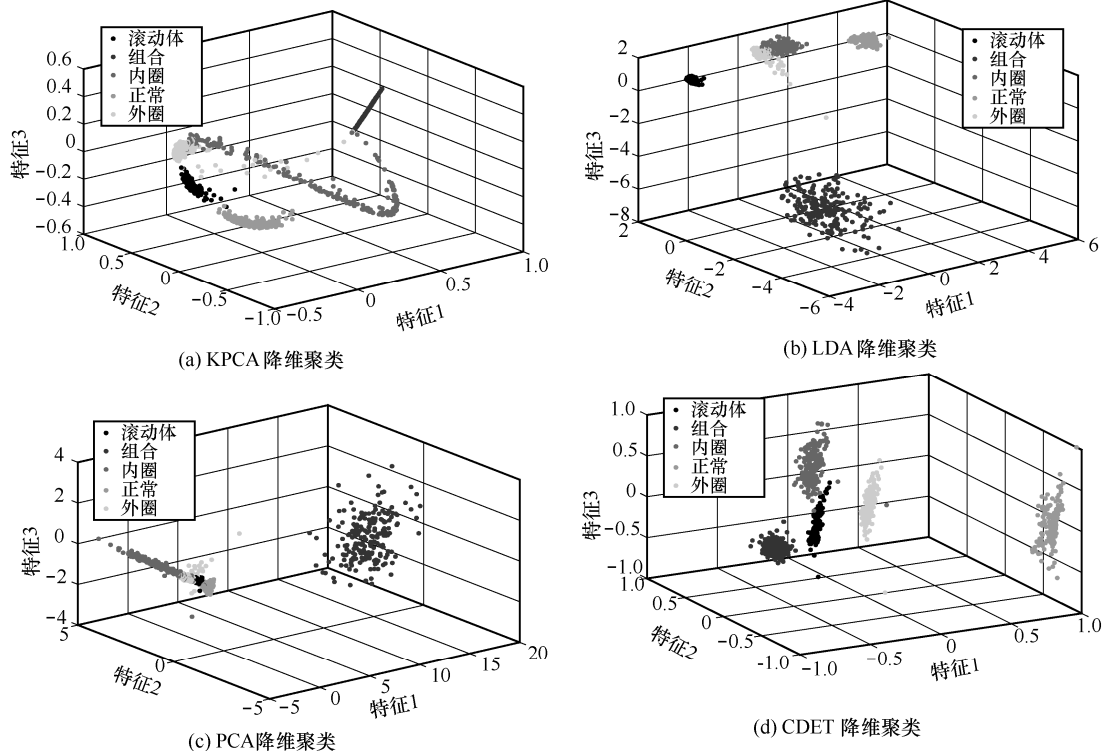


图 11 各降维算法降维后的聚类效果

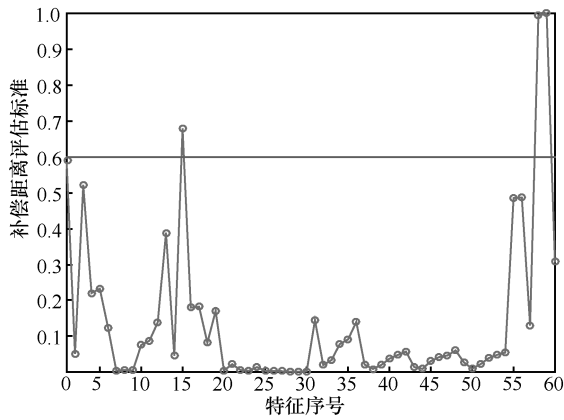


图 12 引入全变分特征集补偿距离评估

2) 不同降维算法分类情况比较

采用 4 种降维算法得到低维特征集合，通过 SVM 故障分类算法的识别精度验证全变分特征对分类效果的影响。图 13 为引入振动信号全变分特征的聚类效果。与图 11 对比可以看出，各种降维算法的聚类效果得到提高。表 3 为 4 种降维算法分类结果。

表 3 4 种降维算法分类结果

降维算法	SVM 分类精度
LDA	99.46%
PCA	99.46%
KPCA	99.47%
CDET	99.87%

对比表 2 和表 3 可知，引入全变分指标后，PCA 和 LDA 故障分类精度分别提高了 5.46%和 3.46%；KPCA 分类精度为 99.47%，相较于未引入全变分特征提高了 1.74%；CDET 降维后分类精度最高约为 99.87%。

图 14 所示为引入振动信号全变分特征前后故障分类精度比较。由图 14 可知，引入全变分特征后相较于未引入全变分特征 4 种降维算法故障分类精度都得到明显提高。而且对本文 2 个时域特征集降维，CDET 降维算法的分类结果都优于传统的降维算法。上述实验结果表明，引入全变分后的特征集能够提高故障的识别精度。

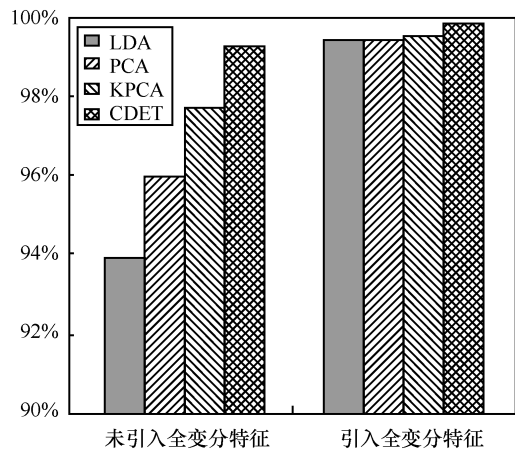
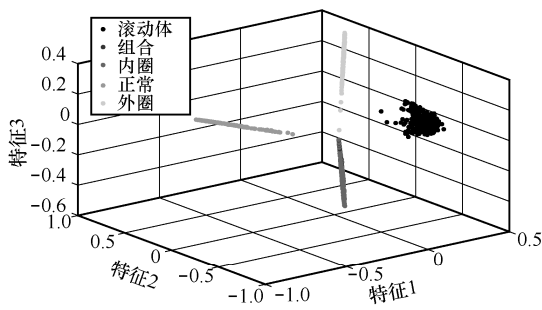
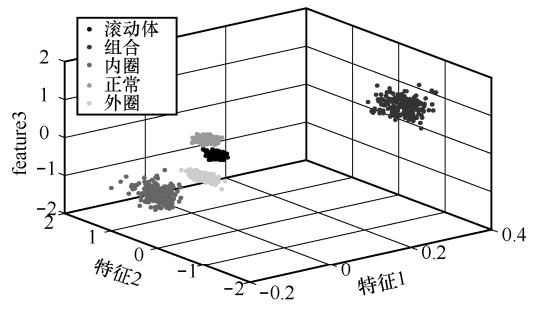


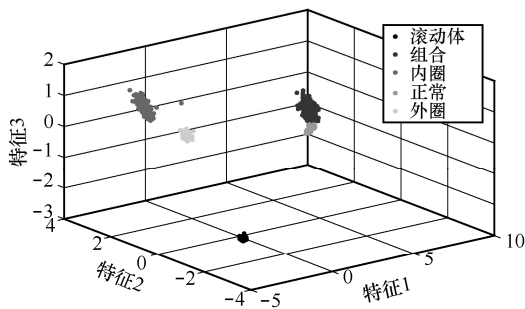
图 14 引入全变分特征前后故障分类精度比较



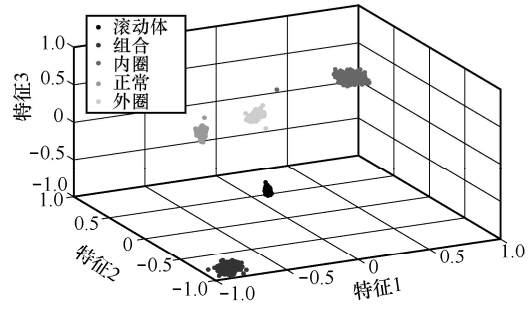
(a) KPCA降维聚类



(b) LDA降维聚类



(c) PCA降维聚类



(d) CDET降维聚类

图 13 引入振动信号全变分特征的聚类效果

7 结束语

针对选煤厂设备种类繁多和监测点分散的特点, 本文提出了一种基于压缩感知的选煤厂设备多源信息无线传输与分类算法。通过构建一种多跳信息传输模型, 将信息传输问题转换为多路测量信号的压缩传感问题, 并将测量矩阵获取转化为多跳信息传输模型的路由问题。针对所获得的测量矩阵存在较大相干性, 影响信号重构效果问题, 将随机路由的思想引入路由构建当中, 提出了一种随机动态自组织路由算法, 该技术在最小路径树的基础上, 通过扩展路由, 增加前向节点, 并对每一个前向节点赋予一定的选择概率, 使路由的选择在保证 QoS 的基础上具有一定的随机性, 解决了测量矩阵存在较大相关性的问题。为了解决重构后的信号时域特征难以对故障类型进行精确分类的问题, 针对重构信号, 提出了一种基于全变分与补偿距离评估相结合的支持向量机分类算法对轴承故障类型分类, 并从两方面论述了该算法的优越性: 1) 补偿距离评估算法降维能够有效地得到特征主元, 而且其降维后聚类结果优于传统的降维算法, CDET 算法在特征筛选中能优先选出全变分特征, 表明全变分具有描述轴承振动信号时域特征的优良特征; 2) 当引入振动信号全变分特征后, 4 种降维方法的故障类型识别精度都得到了提高, 其中补偿距离评估算法降维后数据集相较于传统降维算法 PCA、LDA 与 KPCA, 分类精度最高, 表明引入全变分后的特征集能够提高故障的识别精度。因此, 基于全变分的滚动轴承数据驱动特征提取与故障分类算法, 可以更有效地确定滚动轴承的故障类型。所提出的多源信息传输与分类算法在提高监测数据实时传输效率情况下, 能够有效提高故障识别精度, 特别适用于监测点分散、有线监测系统布置困难的应用场景。

参考文献:

- [1] 温江涛, 闫常弘, 孙洁娣, 等. 基于压缩采集与深度学习的轴承故障诊断方法[J]. 仪器仪表学报, 2018, 39(1):171-179.
WEN J T, YAN C H, SUN J D, et al. Bearing fault diagnosis method based on compressed acquisition and deep learning [J]. Chinese Journal of Scientific Instrument, 2018, 39(1):171-179.
- [2] FAZEL F, FAZEL M, STOJANOVIC M. Random access compressed sensing for energy-efficient underwater sensor networks[J]. IEEE Journal on Selected Areas in Communications, 2011, 29(8):1660-1670.
- [3] 田子建, 贺方圆. 一种基于分布式压缩感知的矿井目标指纹数据库建立方法[J]. 电子与信息学报, 2019, 41(10):2450-2456.
TIAN Z J, HE F Y. Method of establishing mine target fingerprint database based on distributed compressed sensing[J]. Journal of Electronics & Information Technology, 2019, 41(10):2450-2456.
- [4] WANG G, ZHAO Z, NING Y. Design of compressed sensing algorithm for coal mine IoT moving measurement data based on a multi-hop network and total variation[J]. Sensors, 2018, 18(6):1732.
- [5] WANG Z, LIU X. An QoS routing protocol based on ant colony algorithm in WSN[J]. Chinese Journal of Sensors & Actuators, 2011, 2(11):1007-1009.
- [6] 兰巨龙, 张学帅, 胡宇翔, 等. 基于深度强化学习的软件定义网络 QoS 优化[J]. 通信学报, 2019, 40(12):60-67.
LAN J L, ZHANG X S, HU Y X, et al. Software-defined networking QoS optimization based on deep reinforcement learning[J]. Journal on Communications, 2019, 40(12):60-67.
- [7] GELENBE E, NGAI E. Adaptive random re-routing for differentiated QoS in sensor networks[J]. The Computer Journal, 2018, 53(7):1052-1061.
- [8] YANG B S, HAN T, AN J L. ART-Kohonen neural network for fault diagnosis of rotating machinery[J]. Mechanical Systems & Signal Processing, 2004, 18(3):645-657.
- [9] LEI Y, HE Z, ZI Y, et al. New clustering algorithm-based fault diagnosis using compensation distance evaluation technique[J]. Mechanical Systems & Signal Processing, 2008, 22(2):419-435.
- [10] 王前, 王刚, 蒋晗略, 等. 基于 MFCC 与 CDET 的滚动轴承故障诊断方法研究[J]. 控制工程, 2019, 26(9):1682-1686.
WANG Q, WANG G, JIANG H H, et al. Study on fault diagnosis of rolling bearing based on MFCC and CDET[J]. Control Engineering of China, 2019, 26(9):1682-1686.
- [11] LONG Y, LI Z D, XIE M H, et al. Neural network based preprocessing and dictionary methods for switched current circuit fault diagnosis[J]. Journal of Computational & Theoretical Nanoscience, 2017, 14(4):1914-1924.
- [12] GUO S, YANG T, GAO W, et al. A novel fault diagnosis method for rotating machinery based on a convolutional neural network[J]. Sensors, 2018, 18(5):1429.
- [13] YAO C Y, LAI B W, et al. Fault diagnosis method based on med-vmd and optimized SVM for rolling bearings[J]. China Mechanical Engineering, 2017, 28(24):3001-3012.
- [14] WIDODO A, YANG B. Support vector machine in machine condition monitoring and fault diagnosis[J]. Mechanical Systems and Signal Processing, 2007, 21(6):2560-2574.
- [15] HAYASHI K, NAGAHARA M, TANAKA T. A user's guide to compressed sensing for communications systems[J]. IEICE Transactions

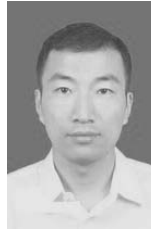
on Communications, 2013, 96(3): 685-712.

- [16] HAUPT J, BAJWA W U, RABBAT M, et al. Compressed sensing for networked data[J]. IEEE Signal Processing Magazine, 2008, 25(2): 92-101.
- [17] LI Q, SHI J, LI Z, et al. Acoustic sound speed profile inversion based on orthogonal matching pursuit[J]. Acta Oceanologica Sinica, 2019, 38(11): 149-157.
- [18] YANG B Y, LIU R N, CHEN X F. Sparse time-frequency representation for incipient fault diagnosis of wind turbine drive train[J]. IEEE Transactions on Instrumentation & Measurement, 2018, 67(11): 2616-2627.
- [19] BO L, YI S, RAJAN S, et al. Theoretical results for sparse signal recovery with noises using generalized OMP algorithm[J]. Signal Processing, 2015, 117(C):270-278.
- [20] DO T T, GAN L, NGUYEN N, et al. Sparsity adaptive matching pursuit algorithm for practical compressed sensing[C]// Asilomar Conference on Signals, Systems and Computers. IEEE, 2008: 581-587.
- [21] ZHANG Y, SUN G. Stagewise arithmetic orthogonal matching pursuit[J]. International Journal of Wireless Information Networks, 2018, 25(2):221-228.
- [22] DAI W, MILENKOVIC O. Subspace pursuit for compressive sensing signal reconstruction[J]. IEEE Transactions on Information Theory, 2009, 55(5): 2230-2249.
- [23] LEI Y, HE Z, ZI Y, et al. New clustering algorithm-based fault diagnosis using compensation distance evaluation technique[J]. Mechanical Systems and Signal Processing, 2008, 22(2): 419-435.
- [24] LU S, WANG X, HE Q, et al. Fault diagnosis of motor bearing with speed fluctuation via angular resampling of transient sound signals[J]. Journal of Sound and Vibration, 2016, 385(12): 16-32.

[作者简介]



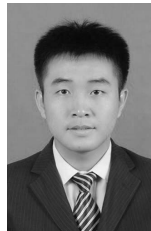
赵小虎（1976- ），男，江苏徐州人，博士，中国矿业大学教授、博士生导师，主要研究方向为物联网、机器学习。



王刚（1977- ），男，江苏徐州人，博士，中国矿业大学副教授、硕士生导师，主要研究方向为机器学习、设备故障诊断、压缩感知。



宋泊明（1989- ），男，江苏徐州人，西澳大学博士生，主要研究方向为机器学习、基于微波信号的雨衰估计、层析重构算法等。



于嘉成（1994- ），男，河北沧州人，中国矿业大学硕士生，主要研究方向为机器学习、设备故障诊断等。