

文章编号: 2095-2163(2019)06-0032-10

中图分类号: U216

文献标志码: A

基于压缩感知的轨道结构故障模式识别研究

许汪歆, 袁天辰, 杨 俭

(上海工程技术大学 城市轨道交通学院, 上海 201620)

摘要: 本文提出一种基于压缩感知的轨道结构故障模式识别方法。该方法通过设计重构信号的稀疏基与测量矩阵,将原始振动信号稀疏重构,解决了轨道振动信号在设计分类器时会出现“过拟合”的问题。通过构造 11 个特征征兆指标,研究重构后的数据特征征兆指标分布规律,解决了数据集维度高的问题。将特征征兆指标两两组合,结合密度聚类算法,成功区分轨道结构正常工况、道床板结工况、道床翻浆工况和轨枕空吊工况。利用归一化互信息(NMI)指标评价密度聚类结果有效性。算例表明,该方法实现了大量样本下轨道基础结构不同故障类型的特征征兆指标提取与故障模式识别。本文所提方法能够有效识别轨道结构故障,为轨道结构故障智能诊断与剩余寿命预测奠定基础。

关键词: 振动信号; 压缩感知; 轨道结构故障; 特征征兆指标; 密度聚类

Research on fault pattern recognition of track structure based on compressed sensing

XU Wangxin, YUAN Tianchen, YANG Jian

(School of Urban Rail Transportation, Shanghai University of Engineering Science, Shanghai 201620, China)

[Abstract] This paper proposes a fault pattern recognition method based on compressed sensing for track structures. The method refines the original vibration signal by designing the sparse basis and measurement matrix of the reconstructed signal. It solves the problem that the track vibration signal will appear “over-fitting” when designing the classifier. By constructing 11 characteristic symptom indicators, the distribution law of the characteristic features of the reconstructed data features is studied. It solves the problem that the data set has a high dimension and a long calculation time. Combining the two groups of characteristic indicators and density clustering algorithm, the paper successfully distinguishes the normal working conditions of track structure, ballast slab working conditions, ballast boiling working conditions and sleeper empty crane working conditions, and the validity of the density clustering results was evaluated using the normalized mutual information (NMI) indicator. A numerical example shows that; This method realizes the feature symptom extraction and fault pattern recognition of different fault types of track infrastructure under a large number of samples. The method proposed in this paper can effectively identify the track structure failure and lay the foundation for the intelligent diagnosis of the track structure fault and the remaining life prediction.

[Key words] vibration signal; compressed sensing; track structure failure; characteristic sign; DBSCAN

0 引言

随着铁路运营里程的增加,轨道结构病害时有发生,严重威胁着铁路线路的运营安全。轨道结构在列车动载荷和环境温度等因素作用下,空吊等病害日益凸显。目前主要依赖于铁路工务人员在天窗时间人工目视检查或手动探伤。这些方法不仅效率低下,而且有漏检的可能,如果不及时检测出故障,将严重威胁列车的行车安全。因此,如何对轨道结构进行监测、故障诊断和预警是保证轨道结构稳定的重要研究方向。

近年来,随着无线通讯技术的发展,机械设备的检测与故障诊断正朝着网络化、无损化方向发展^[1]。轨道结构振动过程中产生的数据集存在样本类标签未知、各数据集间分布不平衡、非线性、多工况、瞬态等诸多问题。目前,研究人员结合机械结构振动状态与压缩感知理论进行了一些探索。文献[2]采用压缩感知与信号稀疏表示,结合机械振动信号自身特点,实现了压缩信号的高精度重构。文献[3]针对稀疏字典较难重构的问题,提出优化分类的方法,将信号进行分块,得到与机械振动信号相适应的稀疏基。现有的故障诊断经验方法通常是根

基金项目: 国家自然科学基金(高铁联合基金)(U1834201);国家自然科学基金(青年基金)(11802170);国家自然科学基金(面上项目)(51575334);上海工程技术大学研究生创新项目基金(18KY1002);上海市自然科学基金(19ZR1421700)。

作者简介: 许汪歆(1992-),女,硕士研究生,主要研究方向:基于振动响应的轨道结构故障诊断;袁天辰(1988-),男,博士,讲师,主要研究方向:非线性振动能量采集;杨 俭(1962-),男,博士,教授,主要研究方向:城市轨道交通车辆运行节能技术。

收稿日期: 2019-09-20

哈尔滨工业大学主办 ◆ 学术研究与应用

据已知故障数据集类别进行监督学习^[4], 无监督聚类方法, 因其实现不需要任何训练样本, 能较好地根据数据结构将相似度较高的样本正确分类, 在故障诊断领域^[5-6]已有应用。上述研究多侧重于对信号进行分解与重构, 以及通过改进算法进而提高信号压缩重构概率, 并没有将压缩重构恢复的信号与故障诊断结合。密度聚类^[7]作为无监督学习, 在数据分析、模式识别中根据数据结构自动划分不同的类别, 为面向数据分析和多源数据融合的轨道结构服役状态监测及故障诊断提供新的方法。

故障诊断与模式识别的首要问题是对数据进行训练。虽然原始数据包含了样本最丰富的信息, 能最完整表达样本特征, 然而在数据传输过程中, 采用原数据传输会增加传感器耗电, 同时在设计分类器时会出现“过拟合”问题。

针对上述问题, 本文提出一种新的轨道结构故障智能诊断方法。采用压缩感知原理, 将轨道结构振动数据集经过稀疏分解与压缩重构, 实现从数据的采集到预处理, 减少数据“冗余性”; 选用时域和频域特征参数作为特征征兆指标, 实现降低数据维度, 解决了“过拟合”问题; 特征征兆指标通过在密度聚类算法下两两组合, 实现在不同特征指标组合下识别轨道结构状态, 为轨道结构故障智能诊断提供参考。

1 车辆—轨道耦合动力学计算模型

车辆采用四轴二系悬挂的整车模型, 轨道系统模拟成三层连续弹性点支撑梁体系, 各支撑点以轨枕间距隔开, 钢轨采用 Euler 梁模型, 以减少模型的复杂程度; 轨枕、钢轨与道床间采用弹簧和阻尼连接; 考虑到钢轨、轨枕及道床垂向振动特性, 道床采用锥台参振模型。文献[8]的计算结果表明, 计算长度为 100 m 的轨道长度满足仿真计算要求。本文中轨道长度为 $l = 120\text{ m}$, 取位于轨道结构中心 $i = 97 \sim 103$ 共 7 个轨枕等截面单元, 轨枕间距 $l_s = 0.6\text{ m}$ 。翟婉明院士在《车辆—轨道耦合动力学》^[9]中指明, 轨下基础支承刚度或阻尼沿纵向不均匀变化时, 可以对模型中各支点刚度和阻尼元件逐一赋值。轨下基础支承引起的轨道刚度或阻尼突变, 缺陷通常为轨枕空吊、道床翻浆和道床板结。某处轨枕完全丧失正常的工作能力可以相应地在模型中设为 $K_{bi} = C_{bi} = 0$; 道床板结或翻浆等缺陷, 相应地在模型中设为 $K_{bi}' = \eta_k K_{bi}, C_{bi}' = \eta_c C_{bi}$, η_k, η_c 分别为道床刚度、阻尼变化系数, 对于不同的情况, η_k, η_c 可在 0.1 ~ 10 范围内取值。图 1 给出了车辆—轨道耦合模型的振动截面图。其中, $B_1 - B_1, B_2 - B_2$ 和 $B_3 - B_3$ 为故障可能出现的断面。

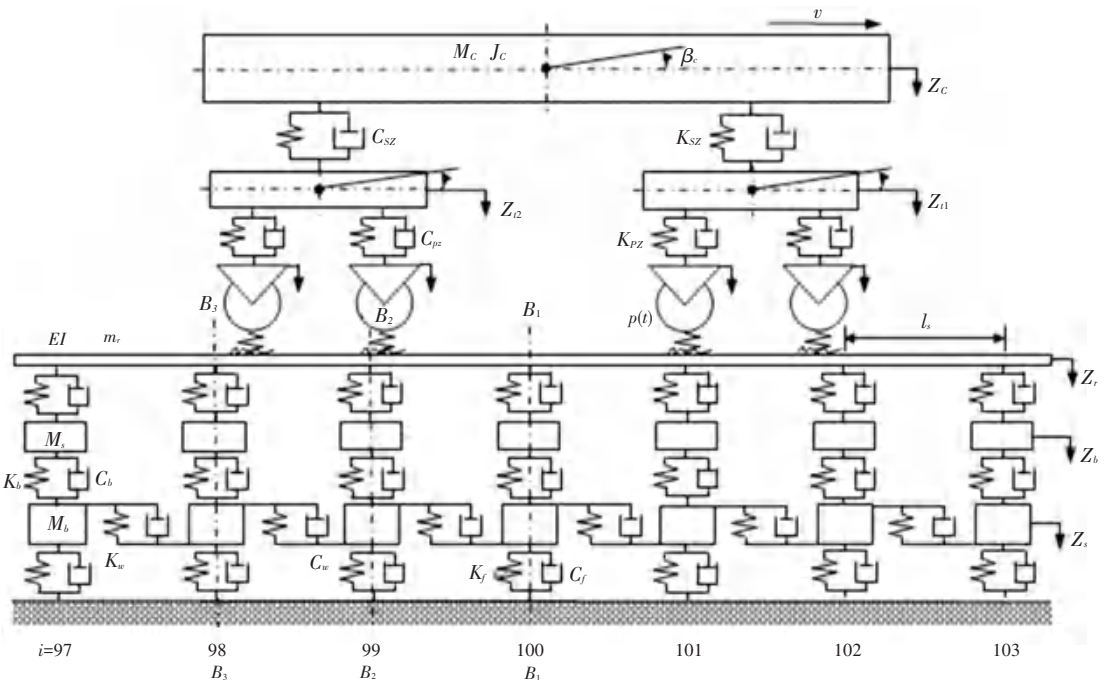


图 1 车辆—轨道垂向振动结构断面模型

Fig. 1 Vehicle-track vertical vibration structure section model

1.1 仿真计算

运用达朗贝尔原理,可以建立轨道系统动力学方程如式(1):

$$M\ddot{Z} + C\dot{Z} + KZ = Q. \quad (1)$$

其中, Z 为位移向量; Q 为力向量; M 为质量矩阵, C 为阻尼矩阵; K 为刚度矩阵。运用 Newmark 方法,通过理论分析和数值计算,可以得到轨道系统动力学响应的时间历程。通过改变轨道结构的质量、刚度和阻尼矩阵以及列车运行速度等参数,模拟不同工况下轨道结构振动响应的变化规律。在分析车辆通过轨道故障结构路段的耦合振动时,选取轨枕振动加速度作为动力学响应指标。模拟轨道结构振动响应工况,以正常工况下 99 号轨枕为例。设定采样频率为 10 kHz,行车速度 200 km/h,如图 2 所示。采样点数为 $N = 8\,000$,在 0.5 s ~ 1.3 s 内振动响应时间历程满足振动响应分析要求。

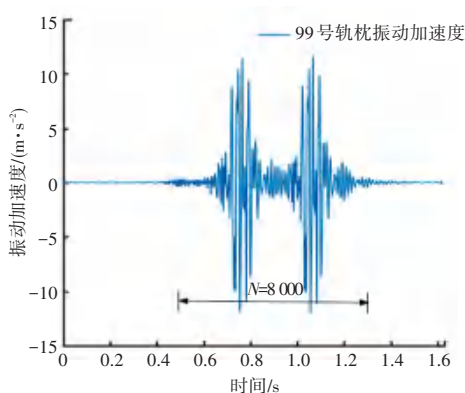


图2 99号轨枕振动响应时程曲线

Fig. 2 Time history curve of vibration response of No. 99 sleeper

1.2 数值计算结果及仿真分析

仿真计算中,车辆运行速度取 80 km/h、120 km/h、160 km/h、200 km/h。本文对正常工况、轨枕空吊、道床翻浆和道床板结的工况分别作了仿真计算,具体计算方案见表 1。

根据上述计算工况表,不同工况下的轨枕加速度都发生变化。当轮轨冲击作用力沿着钢轨传递到轨枕,空吊工况下轨枕的加速度也产生较大变化。轨枕空吊引起的轨枕加速度增加表明轨道结构故障区段的轨枕间距和轨道结构支承刚度已发生变化。

由于车身的分布质量和轮轨作用力对轨道结构振动响应有影响,并且轨道结构服役周期长,巡线检测获取数据量大,单纯从时域分析角度无法较为全面地识别轨道结构故障。因此基于振动信号处理和特征征兆指标提取的故障模式识别方法是预测轨道结构状态的重要方法之一。

表1 轨道结构出现不同病害时轨枕数值模拟计算工况

Tab. 1 Sleeper numerical simulation of track structure with different diseases

| 工况编号 | 故障 | 速度/ (km · h ⁻¹) | 故障处(99号轨枕)加速度峰值 /(m · s ⁻²) | 故障毗邻处(98号轨枕)加速度峰值/(m · s ⁻²) |
|----------------|------|--------------------------------|--|--|
| C ₁ | 轨枕空吊 | 200 | 58.666 2 | 11.842 7 |
| C ₂ | | 160 | 46.739 0 | 7.874 6 |
| C ₃ | | 120 | 41.067 7 | 6.703 8 |
| C ₄ | | 80 | 12.792 0 | 1.843 3 |
| C ₁ | 道床翻浆 | 200 | 34.701 0 | 11.417 3 |
| C ₂ | | 160 | 31.733 8 | 7.568 8 |
| C ₃ | | 120 | 21.079 7 | 6.609 6 |
| C ₄ | | 80 | 12.792 0 | 2.140 4 |
| C ₁ | 道床板结 | 200 | 11.360 7 | 12.968 1 |
| C ₂ | | 160 | 10.687 0 | 9.937 3 |
| C ₃ | | 120 | 7.370 8 | 6.094 9 |
| C ₄ | | 80 | 3.060 9 | 2.887 5 |

2 轨道结构振动信号压缩感知

原振动信号 $x(t) \in 8\,000 \times 1$ 虽然维度高,而实际的有效信息集中在低维空间中。压缩感知是基于线性降维的思想,采用远低于传统的奈奎斯特采样定理要求的采样方式重建原信号。通过信号的稀疏性表示、观测矩阵的不相关性设计以及通过某种算法对原信号的非线性重构,实现信号的稀疏分解与重构^[10]。本文采用压缩感知与信号稀疏重构的过程如图 3 所示。

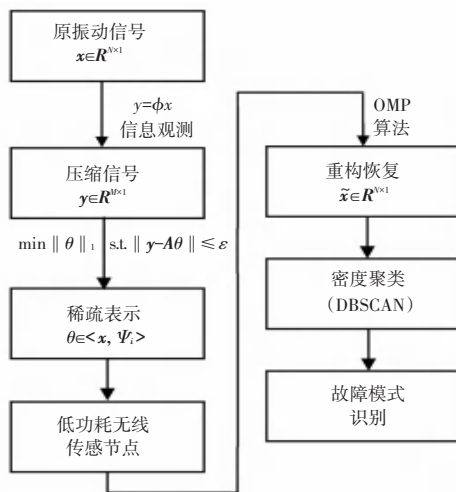


图3 压缩感知框架

Fig. 3 Compressed sensing framework

2.1 轨道结构振动信号的稀疏表示与压缩测量

振动信号的稀疏表示是一个将原信号降维的过程。通过选择合适的稀疏基矩阵 ψ , 求出观测向量 y

在稀疏基矩阵上的稀疏表示 S 。取正常工况的轨枕振动加速度时程曲线作为原信号 $x(t) \in \mathbf{R}^N$ 。如果轨枕振动加速度时程曲线可以稀疏表示, 那么原信号 $x(t) = [x_1, x_2, \dots, x_N]$ 表示为:

$$x = \Psi S = \sum_{i=1}^N \psi_i s_i, \quad (2)$$

式中, $\Psi = [\psi_1, \psi_2, \dots, \psi_N]$ 为稀疏基, $S = [s_1, s_2, \dots, s_N]$ 是稀疏系数。当 $\|S\|_0 = k$ ($k \ll N$) 时, 称原信号 x 是 k -稀疏信号, 其中 $\|\cdot\|_0$ 表示 l_0 范数。事实上, 大多数信号在时域上并不稀疏, 通过选择合适的变换域, 将时域信号转换为稀疏域信号, 实现轨枕振动信号的稀疏表示。一般情况下, 振动信号稀疏域的选择可以是离散傅里叶变换 (Discrete Fourier Transform, DFT)、离散小波变换 (Discrete

$$C = \begin{bmatrix} \sqrt{\frac{1}{2}} & \sqrt{\frac{1}{2}} & \dots & \sqrt{\frac{1}{2}} \\ \cos \frac{\pi}{2N} & \cos \frac{3\pi}{2N} & \dots & \cos \frac{(2N-1)\pi}{2N} \\ \vdots & \vdots & \vdots & \vdots \\ \cos \frac{(N-1)\pi}{2N} & \cos \frac{3(N-1)\pi}{2N} & \dots & \cos \frac{(2N-1)(N-1)\pi}{2N} \end{bmatrix}_{N \times N}, \quad (4)$$

根据图 2 的振动时程曲线, 取采样点长度 $N = 8\,000$, 进行 DCT 域稀疏表示, 仿真结果如图 4 所示。

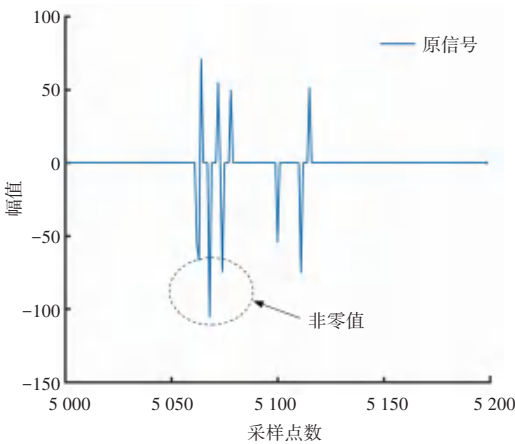


图 4 轨枕振动原信号的稀疏表示系数 S

Fig. 4 Sparse representation coefficient S of original sleeper vibration original signal

根据图 4 的仿真结果, 在 DCT 变换域中, 已将原信号采样点 $N = 8\,000$ 稀疏表示为 $N' = 200$ 稀疏采样点, 并且仅有少量非零值 (如图 4 虚线圈所示)。因此, 99 号轨枕振动信号是可压缩的, 可以对该信号进行随机亚采样。信号在采样阶段, 通过观

Wavelet Transform, DWT)、离散余弦变换 (discrete cosine transform, DCT)^[11]。离散余弦变换的变换矩阵能较好地描述时变信号的相关特征。因此, 本文选取 DCT 变换作为轨枕振动信号的稀疏基矩阵^[12-13]。对于 99 号轨枕振动加速度时程曲线, 其一维离散余弦变换可以表示为,

$$F(0) = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} f(t), \quad \mu = 0$$

$$F(\mu) = \sqrt{\frac{2}{N}} \sum_{t=0}^{N-1} f(t) \cos \frac{\pi}{2N} (2t+1)\mu, \quad \mu = 1, 2, \dots, N-1 \quad (3)$$

(3) 式的矩阵形式为 $F = Cf$, 其中, 变换矩阵 C 为:

测矩阵 $\Phi \in \mathbf{R}^{M \times N}$ 将原信号 $x(t)$ 投影到低维空间 $y \in \mathbf{R}^M$, 观测值与原信号之间的表示为:

$$y = \Phi x, \quad (5)$$

将(2)式代入到(5)式, 得到:

$$y = \Phi \Psi S, \quad (6)$$

令 $\Theta = \Phi \Psi$, 则有:

$$y = \Theta S, \quad (7)$$

由于观测值的维数远小于原信号维数, 式(5)是欠定方程组, 即原方程没有唯一解, 因此不能直接通过(4)式求解出原信号 x 。要想从 M 个观测值中重构恢复出 N 个原信号, 这就要求矩阵

$$\Phi = \begin{bmatrix} \varphi_{11} & \varphi_{12} & \dots & \varphi_{1K} \\ \varphi_{21} & \varphi_{22} & \dots & \varphi_{2K} \\ \vdots & \vdots & \dots & \vdots \\ \varphi_{m1} & \varphi_{m2} & \dots & \varphi_{mK} \end{bmatrix}_{M \times N}$$

与矩阵

$$\Psi = \begin{bmatrix} \psi_{11} & \psi_{12} & \dots & \psi_{1K} \\ \psi_{21} & \psi_{22} & \dots & \psi_{2K} \\ \vdots & \vdots & \dots & \vdots \\ \psi_{m1} & \psi_{m2} & \dots & \psi_{mK} \end{bmatrix}_{N \times N}$$

不相关。则:

$$\mu(\Phi, \Psi) = \sqrt{n} \max_{1 \leq k, j \leq n} |\langle \varphi_k, \psi_j \rangle|, \quad (8)$$

其中, μ 的范围: $\mu(\Phi, \Psi) \in [1, \sqrt{n}]$, μ 越小, Φ 和 Ψ 越不相关。Candes 等人^[14]于 2005 年给出了更为准确的要求: 观测矩阵 Φ 应满足约束等距性条件 (Restricted Isometry Property, 简称 RIP)。即对于任意常数有:

$$(1 - \delta_k) \|x\|_2^2 \leq \|\Phi x\|_2^2 \leq (1 + \delta_k) \|x\|_2^2, \quad (9)$$

观测矩阵与稀疏基矩阵的乘积满足 RIP 性质。满足上述条件后, 重构振动信号 x 的过程可以转化为求解如下最优化问题:

$$P_0: \operatorname{argmin} \operatorname{norm}(x, 0) \text{ s.t. } y = \Phi x, \quad (10)$$

式中, $\operatorname{norm}(x, 0)$ 是正则项, 虽然能够保证式 (5) 的解具有唯一性, 但仍不能求解出该问题。文献[10]已经证明, 正则化的 l_0 范数可以使用 l_1 范数替代, 所以优化问题转换成一个凸优化问题。如式

(11) 表达式:

$$P_1: \operatorname{argmin} \operatorname{norm}(x, 1) \text{ s.t. } y = \Phi x. \quad (11)$$

凸优化问题可以通过线性优化来解决。近几年, 研究者们对于求解优化问题做出了不少研究, 主要包括贪婪算法求解、lasso 模型以及组合算法求解。

2.2 重构算法的理论框架

重构算法主要包括匹配追踪 (Matching Pursuit, MP) 算法^[15]、正交匹配追踪 (Orthogonal Matching Pursuit, OMP) 算法^[16]及 l_1 范数 (l_1 Norm) 算法^[17]。OMP 算法可以实现将原轨枕振动信号转化为较多 0 稀疏, 同时不降低信号精确度, 从而产生稀疏解, 尤其适用于故障诊断的振动信号数据传输^[16]。本文采用正交匹配追踪 (Orthogonal Matching Pursuit) 算法重构步骤见表 2。

表 2 OMP 重构算法

Tab. 2 OMP reconstruction algorithm

核心算法步骤如下:

输入: 观测矩阵 Φ , 观测向量 y ;

输出: 原信号 x 的 K -稀疏的逼近 \hat{x} ;

初始化: 残差 $r_0 = y$, 索引集合 $A_0 = \emptyset, t = 1$;

循环执行步骤 1-5:

步骤 1: 找出残差 r 和观测矩阵 φ_j 乘积最大值对应脚标 λ , 即 $\lambda_t = \operatorname{argmax}_{j=1, \dots, N} |\langle r_{t-1}, \varphi_j \rangle|$;

步骤 2: 更新索引集 $A_t = A_{t-1} \cup \{\lambda_t\}$ 以及观测矩阵中的重建集合 $\Phi_t = [\Phi_{t-1}, \varphi_{\lambda_t}]$

步骤 3: 令 $P_t = X(c_t)(X(c_t)'X(c_t))^{-1}X(c_t)'$, 由最小二乘得到 $\hat{x}_t = \operatorname{argmin}_2 \|y - \Phi_t \hat{x}_t\|_2$;

步骤 4: 更新残差 $r_t = y - \Phi_t \hat{x}_t, t = t + 1$;

步骤 5: 判断终止条件 $\|\hat{x} - x\|_2 \leq \varepsilon$, 若满足, 则停止迭代; 若不满足, 则执行步骤 1。

2.3 轨道结构振动信号算例分析

结合轨道结构振动响应, 根据矩阵的 RIP 性, 稀疏变换基 Ψ 为 $8\,000 \times 8\,000$ 的离散余弦变换 (DCT) 矩阵, 观测矩阵 Φ 为高斯矩阵。文献[18]已证明, 观测矩阵在满足观测次数

$$M \leq c \times S \times \log(N/S), \quad (12)$$

就能够以较低的采样率高概率恢复原始信号。式中, S 是信号稀疏度; N 是信号的长度; $S \ll N$; c 是依赖于 (9) 式的 RIP 常数 δ_t 。将振动信号稀疏分解后, 得到图 5 所示稀疏表示, 该稀疏表示由两部分组成。在第一部分即采样点数为 200 时, 原信号稀疏分解的幅值较明显; 稀疏表示在第二部分幅值趋于 0。经过稀疏分解的 200 个信号, 通过 OMP 重构算法得到如图 6 所示的重构曲线。每一条曲线代表了一种工况的重构恢复率, 尽管不同工况需要的稀疏采样点数不同, 但四种工况都在采样点数达到 200

时, 重构恢复率达到 100%。

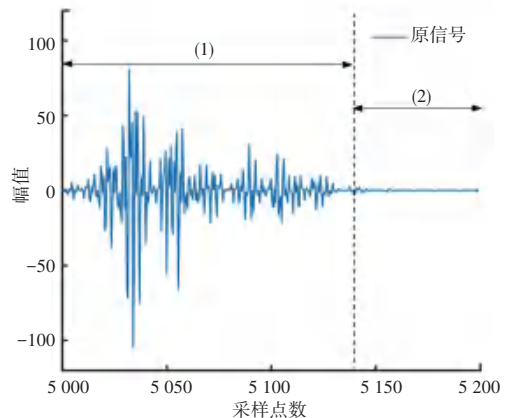


图 5 99 号轨枕振动加速度在 DCT 域下的稀疏表示

Fig. 5 Sparse representation of vibration acceleration of No. 99 sleeper in DCT domain

以稀疏表示的振动信号作为输入, 根据表 2, 得

到轨枕振动加速度的信号重构对比仿真图。尽管有峰值失真(如图 7 所示),但整体重构信号没有明显的误差。图 7 的仿真结果表明,可以采用 OMP 算法重构轨枕振动加速度信号。

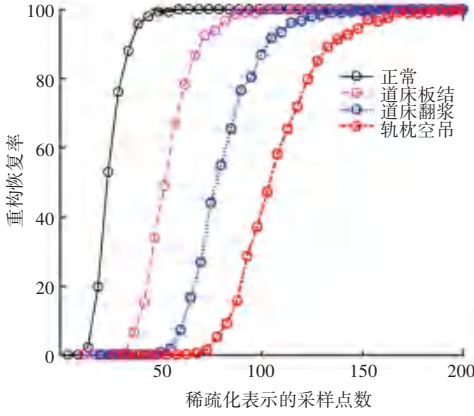


图 6 99 号轨枕的四种工况振动信号重构恢复精度

Fig. 6 Vibration compensation and reconstruction accuracy of four working conditions of No. 99 sleeper

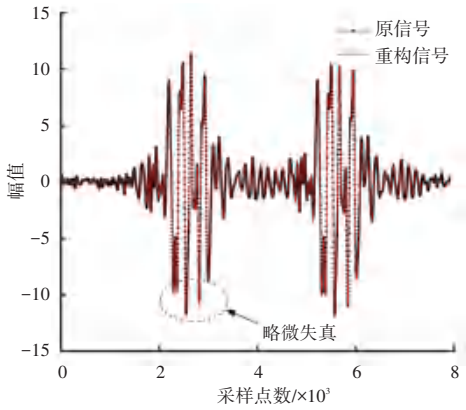


图 7 采用 OMP 算法对 99 号轨枕振动加速度信号重构

Fig. 7 Reconstruction of vibration acceleration signal of No. 99 sleeper by OMP algorithms

2.4 轨道结构振动信号重构及迭代残差分析

采用压缩率衡量轨道振动信号的压缩程度,表示为:CR 值越大,信号越稀疏。

$$CR = \frac{\varphi}{e} - \frac{M}{N} \times 100\%, \quad (13)$$

为了定量衡量 OMP 算法的重构精度,本文选取均方误差 (Mean Squared Error, MSE) 计算原信号的重构误差 $MSE = \frac{1}{N} \sum_{i=1}^N (x - y)^2$, 式中, x 为原信号, y 代表重构信号。MSE 值越小,说明原信号恢复的准确度越高。

OMP 算法重构的残差表示为:

$$re = y - \Theta S = y - \Theta (\Theta^T \Theta)^{-1} \Theta^T y. \quad (14)$$

正常工况和故障工况(道床板结、道床翻浆和轨枕空吊)的稀疏信号在该算法的迭代残差结果如图 8 所示。当迭代次数达到稀疏采样点数 $N' = 200$ 时(图 8 第(1)部分),残差值已经趋于 0,以上结果进一步证明 OMP 算法能较好地重构原振动信号。

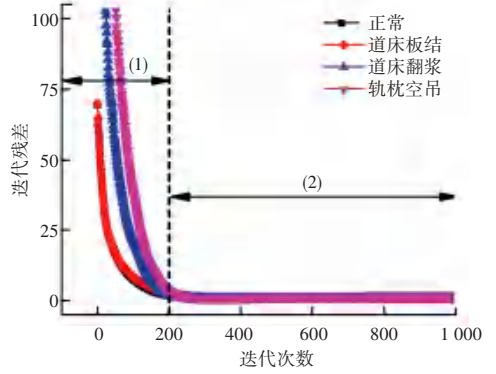


图 8 99 号轨枕的四种工况残差迭代次数

Fig. 8 Iterative number of residual errors in four working conditions of No.99 sleeper

2.5 轨道振动信号特征征兆指标提取

特征征兆指标的提取是对轨道基础结构进行监测、故障诊断以及故障预测的基础^[19],振动加速度能较好地反映轨道结构状态。结合文献[20],采用表 3 定义的各个时域和频域特征参数对轨道结构振动数据进行特征征兆指标提取。信号进行特征征兆指标提取后是高维的特征向量,本文采用基于密度聚类算法 (Density - Based Spatial Clustering of Applications with Noise, DBSCAN),将高维特征向量通过两两组合达到降维效果,实现在不影响轨道结构故障模式识别的基础上,减少特征向量的输入。

3 基于密度聚类算法的轨道结构的故障模式识别

类别可分离性反映不同的类在特征征兆指标组成的平面中的离散情况。通过特征形成、特征提取、特征选择和故障模式识别,识别出不同的轨道结构故障,表征轨道结构服役状态,以实现轨道结构故障在线监测。

3.1 基于空间密度的聚类方法 (DBSCAN)

密度聚类算法 (DBSCAN) 是由 Martin Ester^[21] 等人提出,该算法只需输入一个参数 ϵ (邻域半径),并且可以根据一定的启发式规则选取参数。该算法无需预先指定聚类数目,可以对大规模无规则形状的数据点进行有效的聚类。通过 DBSCAN 算法,可以实现将数据集 D 分为 M 个不同的类。

该方法基于数据点的密度可达和密度相连的概念。从某个选定的核心点出发,不断向密度可达的区域扩张,从而得到一个包含核心点和边界点的最大区

域,区域中任意两点密度相连,将划分结果定义为密度相连对象的集合,类 C 是数据集 D 的非空子集:

表3 轨道结构振动加速度时域频域特征参数统计表

Tab. 3 Statistical table of time-domain and frequency-domain characteristic parameters of vibration acceleration of track structures

| 特征征兆指标 | 表达式 | 特征征兆指标 | 表达式 |
|--------|--|--------|--|
| 波形指标 | $W_s = \frac{X_{rms}}{\bar{X}}$ | 绝对值 | $Vab = x_i(t) $ |
| 脉冲指标 | $I = \frac{X_p}{ \bar{X} }$ | 方差 | $S^2 = \frac{1}{N} \sum_{i=1}^N (x_i(t) - \bar{X})^2$ |
| 峭度指标 | $C_q = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^4}{X_{rms}^4}$ | 均方根值 | $X_{rms} = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2(t)}$ |
| 方根幅值 | $x_r = \left[\frac{1}{T} \int_{-T}^{+T} \sqrt{ X(t)dt } \right]^2$ | 偏斜度指标 | $C_w = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^3}{X_{rms}^3}$ |
| 峰-峰值 | $Vpp = \max(x_i(t)) - \min(x_i(t))$ | 平均幅值 | $Taa = \frac{1}{N} \sum_{i=1}^N (x_i(t))$ |
| — | — | 裕度指标 | $L = \frac{X_p}{X_r}$ |

表3中, $X_p = \max(\max(x_i(t)) - \min(x_i(t)))$, $\bar{X} = \frac{1}{N} \sum_{i=1}^N x_i(t)$.

(1) ε -邻域:某点 p 的 ε -邻域用 $N_\varepsilon(p)$ 表示,定义为集合 $N_\varepsilon(p) = \{q \in D \mid dist(p, q) \leq \varepsilon\}$,由此设置合适的阈值 $MinPts$,通过判断 $|N_\varepsilon(p)|$ 是否小于 $MinPts$,将数据点分为边界点和核心点。

(2) 如果 $p \in N_\varepsilon(p)$ 并且 $|N_\varepsilon(p)| \geq MinPts$,那么,点 p 从点 q 关于 ε 和 $MinPts$ 直接密度可达。

(3) 密度可达: 如果存在数据链 $p_1, p_2, \dots, p_n, p_1 = q, p_n = p$, 并且满足点 p_{i+1} 从 p_i 直接密度可达,那么点 p 从点 q 关于 ε 和 $MinPts$ 密度可达。

(4) 密度相连: 如果存在核心点 $o \in D$, 且边界点 p, q 分别从核心点 o 关于 ε 和 $MinPts$ 密度可达,那么边界点 p 与边界点 q 密度相连。该算法的流程如图9所示。

3.2 基于 DBSCAN 的轨道结构故障模式识别

DBSCAN 算法为了获取一个簇,形成具有相同密度数据之间的聚类,在扫描输入的数据集时,该算法会判断当前数据 q 是否为核心点,进而判断是否需要形成新的簇类^[22]。以1.1节中的工况为例,根据表3计算出有效的特征征兆子集作为输入数据集,得到结果见表4。

根据图9的算法流程,对轨道结构特征征兆指标数据集聚类,识别不同的故障模式。仿真结果如图10所示。双向箭头表示不同类别间的距离。一

个特征征兆指标代表一个维度,不同特征征兆指标之间两两组合,可以从不同特征空间对轨道结构故障识别。

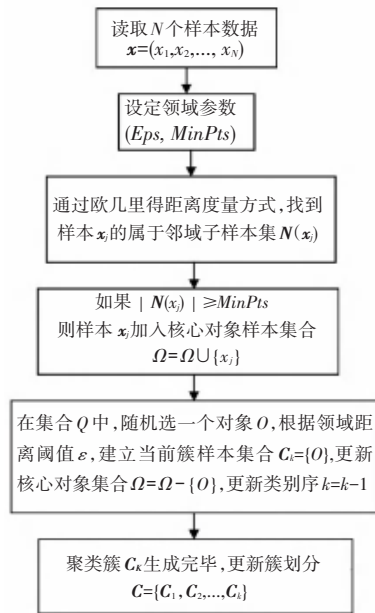


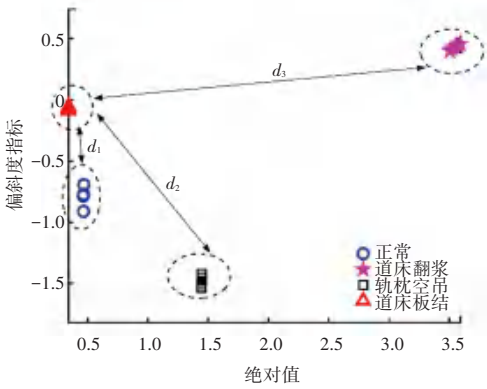
图9 DBSCAN 算法流程

Fig. 9 Flow chart of DBSCAN algorithm

表 4 200 km/h 轨枕不同故障对应的特征征兆指标数据集

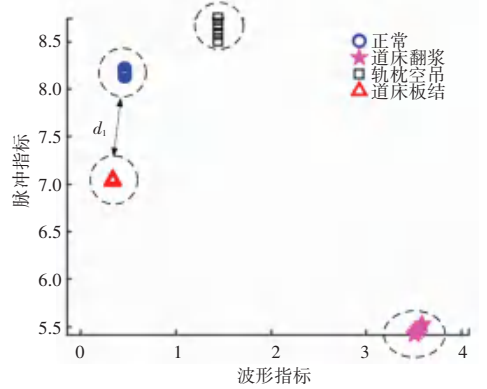
Tab. 4 Data Set of characteristic indicators for different faults of 200 km/h sleepers

| 200 km/h | 波形指标 | 脉冲指标 | 峭度指标 | 裕度指标 | 峰-峰值 | 绝对值 | 方差 | 均方根值 | 偏斜度指标 | 平均幅值 | 方根幅值 |
|----------|---------|---------|---------|----------|----------|---------|----------|---------|----------|---------|---------|
| 正常 | 1.761 2 | 8.095 4 | 8.209 6 | 12.584 1 | 6.601 1 | 0.464 9 | 0.670 5 | 0.818 8 | -0.786 9 | 0.464 9 | 0.299 1 |
| 道床板结 | 1.682 6 | 6.902 0 | 7.050 6 | 10.325 8 | 4.617 1 | 0.341 0 | 0.329 2 | 0.573 8 | -0.151 5 | 0.341 0 | 0.227 9 |
| 道床翻浆 | 1.788 6 | 7.784 7 | 8.738 9 | 11.943 1 | 18.127 0 | 1.443 2 | 6.663 8 | 2.581 3 | -1.536 6 | 1.443 1 | 0.940 7 |
| 轨枕空吊 | 1.583 5 | 6.252 1 | 5.414 5 | 9.228 7 | 37.351 4 | 3.506 6 | 30.834 4 | 5.552 5 | 0.400 1 | 3.506 6 | 2.375 6 |



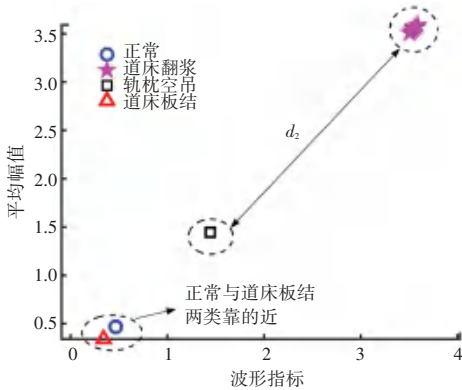
(a) 绝对值与偏斜度指标聚类

(a) Absolute value and skewness index clustering



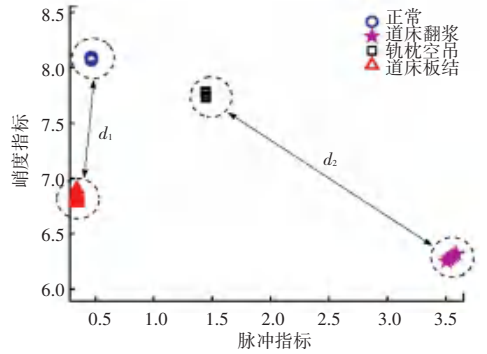
(b) 波形指标与脉冲指标聚类

(b) Waveform index and pulse index clustering



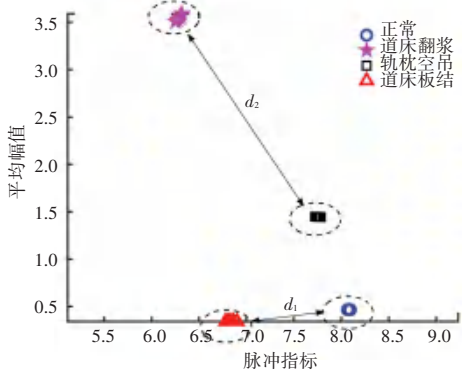
(c) 波形指标与平均幅值聚类

(c) Waveform index and average amplitude clustering



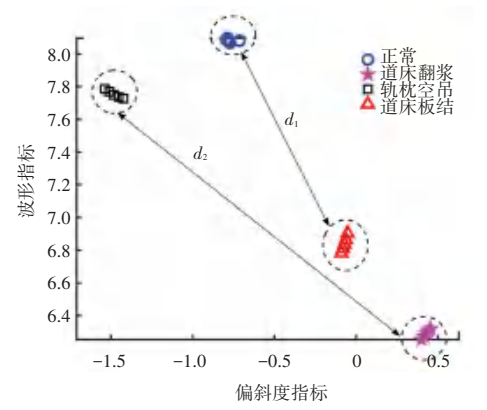
(d) 脉冲指标与峭度指标聚类

(d) Pulse index and kurtosis index clustering



(e) 脉冲指标与平均幅值聚类

(e) Pulse index and average amplitude clustering



(f) 偏斜度指标与波形指标聚类

(f) Tilt index and waveform index clustering

图 10 不同特征征兆指标之间的聚类

Fig. 10 Clustering between different feature indications

将图 10 的(a)与(d)比较来看,道床板结工况与正常工况在脉冲指标和峭度指标这两个特征征兆指标下,类间离散程度较大,而在图 10(a)中没有这样的聚类效果。图 10 的(b)与(c)是波形指标和脉冲指标、波形指标和平均幅值的对比图。从波形指标与平均幅值的组合来看,密度聚类算法不能对正常工况与道床板结这两个类别间进行区分。事实上,道床板结虽然降低了整个轨道系统的弹性功能,但对轨道结构垂向振动影响并不大,单纯从轨道动力学角度不容易识别,而这一点在密度聚类算法中(图 10 的(d)、(e)、(f))已能较好地区分。因此,轨道动力学结合振动信号处理和特征征兆指标提取,能够实现轨道结构服役状态故障模式识别,为轨道结构在线监测奠定基础。

3.3 聚类有效性评价指标

由于在密度聚类识别中,数据集的结构及分布是未知的,因此需要验证密度聚类结果的合理性和有效性。可以通过建立一个评价函数,评价聚类的有效性,聚类评价指标实现了度量聚类算法结果是否有效的性能。

由于轨道结构振动时域数据存在冗余性、数据结构分布不明确等特点,从紧致性、分离性和重叠度等方面评价类别划分结果。采用归一化互信息(NMI)指标描述密度聚类结果的有效性。NMI 用于衡量两个聚类结果 c_k 和 c_l 的吻合程度。其表达式如下:

$$NMI = \frac{-2 \sum_{i=1}^{C_K} \sum_{j=1}^{C_L} C_{ij} \cdot \log \frac{C_{ij} \cdot N}{C_i \cdot C_j}}{\sum_{i=1}^{C_K} C_i \cdot \log \frac{C_i \cdot N}{C_i \cdot N} + \sum_{j=1}^{C_L} C_j \cdot \log \frac{C_j \cdot N}{C_j \cdot N}} \quad (15)$$

其中, N 是数据个数; C 是混淆矩阵; 矩阵中的元素 C_{ij} 表示既在 K 类中也在 L 类; $C_K(C_L)$ 表示 $K(L)$ 类的个数; $C_i(C_j)$ 是矩阵 C 中所有元素的总和。NMI 的值越大, K 和 L 的划分越相似, 当 NMI 的值为 1 时, 表明 K 和 L 是属于同一个类。

基于密度聚类的轨道结构聚类的有效性评价指标函数,道床板结与正常工况这两个类别在波形指标与平均幅值中, NMI 值接近于 1, 这表明道床板结与正常工况这两类几乎是同一个类别, 同时也表明其重叠度明显。对于道床翻浆与轨枕空吊, 这两个类别无论在哪个特征征兆指标对应的维度下, 类间分离性较大, 类内具有高度紧致性。

4 结束语

随着无线通讯与物联网技术的发展, 轨道服役状态的智能监测与预警问题受到广泛关注。然而轨道结构的病害检测目前依赖于人工检修与轨检车巡线检测。针对这一问题, 提出一种基于压缩感知的轨道结构故障模式识别的方法, 结论如下:

(1) 利用轨枕振动信号的可稀疏化表示, 提出采用离散余弦矩阵作为稀疏基, 解决了轨枕时变信号冗余特性带来的无法构造正交基矩阵的问题。

(2) 基于 OMP 贪婪算法对轨枕的振动信号进行重构, 比较了正常工况、翻浆工况、板结工况和空吊工况下的重构恢复率, 结果表明上述工况在 OMP 算法下都能精确重构。

(3) 利用时域和频域特征参数表达式将压缩重构的轨枕振动信号转换成 11 个特征征兆指标, 结合密度聚类算法, 成功区分出轨道结构的正常工况和三种故障工况(翻浆工况、板结工况和空吊工况), 实现了轨道结构无损化监测。

参考文献

- [1] 曾超, 汤宝平, 肖鑫, 等. 低功耗机械振动无线传感器网络节点结构设计[J]. 振动与冲击, 2017, 36(14): 33-37, 65.
- [2] 温江涛, 孙洁娣, 于洋, 等. 基于小波包字典优化的旋转机械振动信号压缩感知重构方法[J]. 振动与冲击, 2018, 37(22): 164-172.
- [3] 王强, 张培林, 王怀光, 等. 基于优化分类的机械振动信号压缩感知[J]. 振动与冲击, 2018, 37(14): 86-93.
- [4] 梁潇, 王海峰, 郭进, 等. 基于贝叶斯网络的列控车载设备故障诊断方法[J]. 铁道学报, 2017, 39(8): 93-100.
- [5] 王瑞峰, 陈旺斌. 基于灰色神经网络的 S700K 转辙机故障诊断方法研究[J]. 铁道学报, 2016, 38(6): 68-72.
- [6] 杨帅杰, 马跃, 张旭, 等. 一种 ELMD 模糊熵和 GK 聚类的轴承故障诊断方法[J]. 机械设计与制造, 2018(6): 118-121.
- [7] 郭昆, 彭胜波, 陈羽中, 等. 基于密度聚类的增量动态社区发现算法[J]. 模式识别与人工智能, 2018, 31(11): 965-978.
- [8] 史红梅. 基于车辆动态响应的轨道不平顺智能感知算法研究[D]. 北京: 北京交通大学, 2012.
- [9] 翟婉明. 车辆-轨道耦合动力学[M]. 4 版. 北京: 科学出版社, 2007.
- [10] 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.
- [11] 张新鹏, 胡莴庆, 程哲, 等. 基于压缩感知的振动数据修复方法[J]. 物理学报, 2014, 63(20): 200506(1)-200506(10)
- [12] PRAGATI D, PRIYANKA J. Realization of Second-Order Structure of Recursive Algorithm for Discrete Cosine Transform [J]. Circuits, Systems, and Signal Processing, 2019, 38(2): 791-804.