

基于变分贝叶斯的轴承故障诊断方法

王 岩 罗 倩 邓 辉

(北京信息科技大学信息与通信工程学院 北京 100101)

摘 要 滚动轴承是旋转机械结构中常用的零件,如果发生故障,会造成极大的危害。随着大数据时代的到来,现代智能诊断方法已被广泛应用到轴承故障诊断中。针对目前智能诊断方法存在的问题,将统计模型引入轴承故障诊断中,提出了基于变分贝叶斯的轴承故障诊断方法。该方法对轴承振动信号进行局部特征尺度分解,得到若干个内禀尺度分量,并分别提取时域特征组成特征集,使用特征集训练产生基于变分贝叶斯的混合多维高斯分布模型,通过计算不同轴承故障的概率实现故障诊断。实验结果表明,所提方法的诊断正确率达到 99.6%,与基于支持向量机的轴承诊断方法相比,在所组成的特征集上诊断正确率最高提升了 39.6%。文中提出的方法能够全面且有效地诊断滚动轴承故障,对高维复杂的故障数据也有很好的诊断效果。

关键词 变分贝叶斯,轴承故障诊断,局部特征尺度分解,高斯混合模型

中图法分类号 TP206+.3 **文献标识码** A **DOI** 10.11896/jsjcx.180901719

Bearing Fault Diagnosis Method Based on Variational Bayes

WANG Yan LUO Qian DENG Hui

(College of Information and Communication Engineering, Beijing Information Science and Technology University, Beijing 100101, China)

Abstract Rolling bearings are common parts in rotating mechanical structures and can cause significant damage if they fail. With the advent of the era of big data, modern intelligent diagnostic methods are widely used in bearing fault diagnosis. Aiming at the problems existing in the intelligent diagnosis method, this paper introduced the statistical model into the bearing fault diagnosis, and proposed a fault diagnosis method based on the variational Bayesian. The method performs local feature scale decomposition on the bearing vibration signal to obtain several intrinsic scale components and extracts the time domain feature composition feature set. The feature set training is used to generate the mixed multidimensional Gaussian distribution model based on variational Bayes, and the different bearings are calculated. The probability of failure is to achieve fault diagnosis. The experimental results show that the diagnostic accuracy rate is 99.6%. Compared with the bearing diagnosis method based on support vector machine, diagnostic accuracy rate is up to 39.6%. The proposed method can comprehensively and effectively diagnose rolling bearing faults, and has a good diagnostic effect on high-dimensional complex fault data.

Keywords Variational Bayes, Bearing fault diagnosis, Local characteristic scale decomposition, Gaussian mixture model

1 引言

滚动轴承故障诊断的理论和应用研究一直是机械故障诊断领域的一个重点。随着大数据和人工智能时代到来,越来越多的机器学习算法被用于轴承故障诊断中。Zhang 等^[1]提出了一种基于神经网络的故障诊断方法,神经网络模型从大量源数据中学习特征并相应地训练参数。该方法可以提高分类精度,缩短训练时间。孙珊珊等^[2]提出基于 SVM 的故障诊断方法,使用双树复小波包变换分解原始振动信号,得到不同频带的分量,并从各个频带分量中提取能量特征来训练支持向量机模型。Fu 等^[3]提出基于 Elman-AdaBoost 的故障诊断方法。原始振动信号经验模态分解为若干个固有模式函

数,使用固有模式函数重建信号,从每个重建信号中提取均方根值和功率谱中心作为二维特征向量,使用特征向量训练 Elman-AdaBoost 诊断模型。

现有的这些诊断方法都基于机器学习中的判别式模型,而判别式模型本身存在一定的局限性,比如支持向量机中针对每一种轴承故障都设置一个判别函数进行诊断,造成输入空间中存在无法分类的区域,即诊断片面性的问题^[4];并且在工况不理想的情况下,特征集合中的部分特征质量较差,对不同故障类别的分辨度较低,判别式模型的诊断正确率会出现明显降低。本文通过引入基于变分贝叶斯的混合高斯分布生成式模型^[5]即轴承故障诊断的统计模型来解决此问题,从一个新的角度实现了滚动轴承故障的智能诊断。

到稿日期:2018-09-13 返修日期:2019-01-10 本文受促进高校内涵发展-信息与通信工程一级学科(5121911025),北京市教委科研计划项目(201811232009)资助。

王 岩(1993—),男,硕士生,主要研究方向为信号处理、大数据处理,E-mail: wyleo7@qq.com;罗 倩(1965—),女,博士,副教授,主要研究方向为信号处理、大数据处理,E-mail: luoqian@bistu.edu.cn(通信作者);邓 辉(1994—),男,硕士生,主要研究方向为信号处理、大数据处理。

本文提出一种新的基于变分贝叶斯的轴承故障诊断方法。轴承故障诊断的统计模型通常是高维复杂的^[6]。使用变分贝叶斯方法推断模型参数的总体思路为:首先对轴承时域振动信号进行局部特征尺度分解(Local Characteristic scale Decomposition, LCD),得到信号的内禀尺度分量(Intrinsic Scale Component, ISC)^[7];再对轴承故障振动信号和若干个 ISC 分别提取时域特征组成特征集。对于所有轴承故障类型的特征集,使用基于变分贝叶斯的方法训练产生混合多维高斯分布模型。对于待诊断的样本,使用统计模型分别计算该样本属于每一种轴承故障的概率,并选取最大概率判定为该故障类别,从而达到故障诊断的目的,解决了判别式模型在轴承故障诊断中存在的问题。

2 数据预处理与特征提取

2.1 数据 LCD 预处理

轴承故障信号时域特征通常是非平稳的,包含周期性的冲击信号成分,冲击信号是滚动体通过故障点时撞击故障点而产生的。冲击的重复频率与滚动体通过频率完全一致。滚动体通过故障点时激起轴承体系的结构共振,随着滚动体通过故障点,这种共振会迅速衰减。为了提高故障诊断质量,本文使用程军圣等^[7]提出的局部特征尺度分解方法对原始轴承振动信号进行预处理。LCD 是一种基于极值点的局部尺度参数自适应地非平稳信号分解方法,该方法以任意相邻的两个极值点为跨度,通过分段的形式对信号进行线性变换来构造基线函数,从而将原始信号自适应地分解为若干个 ISC 分量。其具体步骤为:通过两个循环将原始信号分解为具有物理意义的内禀尺度分量 ISC 和一个剩余项。第一个循环是寻找基线信号,并经循环迭代求出 ISC 分量的过程,ISC 分量应满足 ISC 分量条件;第二个循环是将原始信号减掉已获得的 ISC 分量后的剩余信号作为原始信号再次进行循环迭代得到 ISC 分量。重复上述步骤,直至所得剩余项为一个单调函数为止^[8]。对于任一时间序列,通过 LCD 分解可以得到若干个 ISC 和一个剩余项。LCD 分解的本质是将非平稳信号进行平稳化处理,将信号中不同尺度的波动或趋势逐级分解。

2.2 特征提取

LCD 预处理之后,对原始振动信号和若干个 ISC 分量分别提取时域特征。为了能够更加全面、准确地诊断故障,本文对轴承原始振动信号和若干个 ISC 分别提取了 5 个无量纲特征:偏斜度、峰值、波形指标、脉冲、裕度;以及 5 个有量纲特征:有效值、均值、标准差、方根幅值、最大值。无量纲指标对剥落、压痕等故障有很好的识别效果;有量纲指标对磨损类故障的识别效果比较好。

3 基于变分贝叶斯的轴承故障诊断

本文只涉及到轴承运行的 4 种状态:内圈故障状态、外圈故障状态、滚动体故障状态和正常状态。基于变分贝叶斯的轴承故障诊断的基本思路是:从原始轴承振动数据和若干个 ISC 中分别提取时域特征组成特征集,分析所有特征的分布,对 4 种轴承运行状态的特征集建立混合多维分布模型,使用变分贝叶斯算法对模型进行参数推断,最终得到混合多维概率分布模型。对于待诊断样本,使用混合多维概率模型分别

计算出其属于每一种轴承状态的概率,选取的最大概率判定样本为该类别。

3.1 特征值分布

建立特征的多维概率分布,首先需要分析每个特征在不同状态下的一维分布,然后构建出合适的多维概率分布。生成轴承不同状态下各特征的分布直方图,通过直观观察发现绝大多数直方图服从高斯分布。进一步,使用 Quantile-Quantile 正态检验方法^[9]和 Michael 拟合优度检验方法^[10]检验各特征分布,发现这些特征都落在 90% 接受区间内,可以认为这些特征都服从高斯分布。因此,针对轴承的每一个状态,可以建立特征集的多维高斯分布。

3.2 变分贝叶斯推断

变分贝叶斯方法的原理为:数据集 $\mathbf{X}$ 和待求解参数 Θ 的后验分布为 $p(\Theta|\mathbf{X})$,其中模型 Θ 包含模型参数 θ 和隐藏变量 $\mathbf{Z}$ 。在高维数据中,考虑到后验分布的复杂性,很难根据贝叶斯理论求得后验概率分布的精确解。因此,用一种简单的分布 $q(\Theta)$ 作为后验分布的近似^[11]。为了描述 $q(\Theta)$ 对 $p(\Theta|\mathbf{X})$ 的逼近程度,引入 KL 散度(Kullback-Leibler divergence):

$$KL(q \parallel p) = - \int q(\Theta) \ln \{ \frac{p(\Theta|\mathbf{X})}{q(\Theta)} \} d\Theta \tag{1}$$

令:

$$L(q(\Theta)) = \int q(\Theta) \ln \{ \frac{p(\Theta, \mathbf{X})}{q(\Theta)} \} d\Theta \tag{2}$$

$KL(q(\Theta) \parallel p(\Theta|\mathbf{X})) \geq 0$,当 $q(\Theta) = p(\Theta|\mathbf{X})$ 时, KL 取 0; $L(q(\Theta))$ 为 ELOB(Evidence Lower Bound)。将式(1)与式(2)相加,可以得到:

$$\ln p(\mathbf{X}) = L(q(\Theta)) + KL(q(\Theta) \parallel p(\Theta|\mathbf{X})) \tag{3}$$

目标是最小化 $KL(q(\Theta) \parallel p(\Theta|\mathbf{X}))$,同时 $\ln p(\mathbf{X})$ 是一个不依赖于 Θ 的常数^[12],因此最小化 $KL(q(\Theta) \parallel p(\Theta|\mathbf{X}))$ 等价于最大化 $L(q(\Theta))$ 。变分贝叶斯通过 $q(\Theta)$ 的迭代更新来实现 $L(q(\Theta))$ 取得最大值^[13]。引入平均场假设,可以将 $q(\Theta)$ 分解:

$$q(\Theta) = \prod_{i=1}^M q_i(\Theta_i) \tag{4}$$

通过对每个因子 $q_i(\Theta_i)$ 进行最优化来完成整体的最优化。根据泛函理论^[13],分别对每一个 $q_i(\Theta_i)$ 求偏导,可得到相应的通解表达式,将 $q_i(\Theta_i)$ 记作 q_i 。

$$q_i(\Theta_i) = \frac{\exp(E_{j \neq i}[\ln p(\Theta, \mathbf{X})])}{\int \exp(E_{j \neq i}[\ln p(\Theta, \mathbf{X})]) d\Theta_i} \tag{5}$$

E 表示期望,对式(5)两侧取对数可得:

$$\ln q_i(\Theta_i) = E_{j \neq i}[\ln p(\Theta, \mathbf{X})] + \text{常数} \tag{6}$$

本文中常数定义为与函数自变量无关的项。其中:

$$E_{j \neq i}[\ln p(\Theta, \mathbf{X})] = \int \ln p(\Theta, \mathbf{X}) \prod_{j \neq i} q_j d\Theta_j \tag{7}$$

变分贝叶斯最优化涉及到在两个阶段之间的循环更新。第一阶段计算每个隐藏变量 $\mathbf{Z}$ 的期望 $E[z_{nk}]$,第二阶段更新模型参数 θ 使得 $E[z_{nk}]$ 取最大值,每次迭代检查变分下界是否达到阈值^[14]。

变分贝叶斯轴承故障诊断方法的思想为:通过特征值分布分析,对轴承运行的 4 种状态建立混合多维高斯分布模型,使用变分贝叶斯进行模型参数的推断。轴承故障诊断模型的混合多维高斯分布为:

$$p(\mathbf{X}|\mathbf{Z},\boldsymbol{\pi},\boldsymbol{\mu},\mathbf{A})=\prod_{n=1}^N\prod_{k=1}^K(\pi_k\mathcal{N}(\mathbf{x}_n|\boldsymbol{\mu}_k,\mathbf{A}_k^{-1}))^{z_{nk}} \quad (8)$$

针对轴承的 4 种状态,各取相同数量的特征输入向量 $\mathbf{x}_n$ 组成训练集 $\mathbf{X}$; π_k 表示混合系数,每种轴承状态特征集取相同数量的特征输入向量,因此 $\pi_k=0.25$; $\boldsymbol{\mu}_k, \mathbf{A}_k$ 分别表示第 k 个多维高斯分布的均值和精度。 $\mathbf{X}$ 由 N 个特征输入样本组成,每个输入向量的维度 D 表示提取的特征总数,隐藏变量 $\mathbf{Z}$ 表示每个输入向量 $\mathbf{x}_n$ 属于某个特定的多维高斯分布,因此 K 个混合数量的混合高斯分布中隐藏变量 $\mathbf{Z}$ 表示为 $N \times K$ 矩阵。每个元素 z_{nk} 取 0 或 1。 $z_{nk}=1$ 表示输入向量 $\mathbf{x}_n$ 属于第 k 个高斯分量,每个输入向量仅能对应一个高斯分量,因此隐藏变量 $\mathbf{Z}$ 的条件概率分布可以表示为:

$$p(\mathbf{Z}|\boldsymbol{\pi})=\prod_{n=1}^N\prod_{k=1}^K\pi_k^{z_{nk}} \quad (9)$$

在共轭指数域内,先验分布与后验分布具有一致性的保证,使得通过在共轭指数域内选取分布函数能够得到闭合的解析表达式,并通过循环迭代方法求出最优的超参数解^[15]。引入参数 $\boldsymbol{\pi}, \boldsymbol{\mu}$ 和 $\mathbf{A}$ 的共轭先验概率分布。混合系数 $\boldsymbol{\pi}$ 的先验分布为狄利克雷分布,每种轴承状态的特征输入向量数量相同,因此初始化时每个分类选择了同样的参数 α_0 , $C(\alpha_0)$ 是狄利克雷分布归一化常数:

$$p(\boldsymbol{\pi})=Dir(\boldsymbol{\pi}|\alpha_0)=C(\alpha_0)\prod_{k=1}^K\pi_k^{\alpha_0-1} \quad (10)$$

$\boldsymbol{\mu}$ 和 $\mathbf{A}$ 的先验分布为高斯-Wishart 分布^[16]:

$$p(\boldsymbol{\mu},\mathbf{A})=p(\boldsymbol{\mu}|\mathbf{A})p(\mathbf{A}) \\ =\prod_{k=1}^K\mathcal{N}(\boldsymbol{\mu}_k|\mathbf{m}_0,(\beta_0\mathbf{A}_k)^{-1})\mathcal{W}(\mathbf{A}_k|\mathbf{W}_0,\nu_0) \quad (11)$$

其中, $\mathcal{W}$ 表示 Wishart 分布; $\mathbf{W}_0$ 为初始尺度矩阵 (Scale matrix); ν_0 为自由度参数。根据变分贝叶斯原理,引入形式简单分布 $q(\boldsymbol{\Theta})$ 近似后验分布 $p(\mathbf{Z},\boldsymbol{\pi},\boldsymbol{\mu},\mathbf{A}|\mathbf{X})$ 。令 $q(\boldsymbol{\Theta})=q(\mathbf{Z},\boldsymbol{\pi},\boldsymbol{\mu},\mathbf{A})$ 为变分近似的联合密度函数,根据平均场假设进行因子化,在隐藏变量和模型参数之间进行分解:

$$q(\mathbf{Z},\boldsymbol{\pi},\boldsymbol{\mu},\mathbf{A})=q(\mathbf{Z})q(\boldsymbol{\pi})\prod_{k=1}^Kq(\boldsymbol{\mu}_k,\mathbf{A}_k) \quad (12)$$

根据式(6), $q(\boldsymbol{\Theta})$ 为:

$$\ln q^*(\mathbf{Z})=E_{\boldsymbol{\pi},\boldsymbol{\mu},\mathbf{A}}[\ln p(\mathbf{X},\mathbf{Z},\boldsymbol{\pi},\boldsymbol{\mu},\mathbf{A})]+\text{常数} \\ =E_{\boldsymbol{\pi}}[\ln p(\mathbf{Z}|\boldsymbol{\pi})]+E_{\boldsymbol{\mu},\mathbf{A}}[\ln p(\mathbf{X}|\mathbf{Z},\boldsymbol{\pi},\boldsymbol{\mu},\mathbf{A})]+\text{常数} \\ =\sum_{n=1}^N\sum_{k=1}^Kz_{nk}\ln \rho_{nk}+\text{常数} \quad (13)$$

其中, ρ_{nk} 定义为:

$$\ln \rho_{nk}=E[\ln \pi_{nk}]+\frac{1}{2}E[\ln |\mathbf{A}_k|]-\frac{D}{2}\ln(2\pi)- \\ \frac{1}{2}E_{\boldsymbol{\mu}_k,\mathbf{A}_k}[(\mathbf{x}_n-\boldsymbol{\mu}_k)^T\mathbf{A}_k(\mathbf{x}_n-\boldsymbol{\mu}_k)] \quad (14)$$

式(13)中 * 表示最优因子; D 为多维高斯分布的维数。对式(13)两侧取指数并归一化:

$$q^*(\mathbf{Z})=\prod_{n=1}^N\prod_{k=1}^K\pi_k^{z_{nk}} \quad (15)$$

其中, r_{nk} 表示第 n 个输入向量属于第 k 类轴承状态的概率:

$$E_{z_{nk}}[z_{nk}]=\frac{\rho_{nk}}{\sum_{j=1}^K\rho_{nj}}=r_{nk} \quad (16)$$

初始化先验分布中的超参数后,迭代如下步骤。

步骤 1 使用 r_{nk} 更新模型的相关参数。

$$N_k=\sum_{n=1}^Nr_{nk} \quad (17)$$

$$\bar{\mathbf{x}}_k=\frac{1}{N_k}\sum_{n=1}^Nr_{nk}\mathbf{x}_n \quad (18)$$

$$\mathbf{S}_k=\frac{1}{N_k}\sum_{n=1}^Nr_{nk}(\mathbf{x}_n-\bar{\mathbf{x}}_k)(\mathbf{x}_n-\bar{\mathbf{x}}_k)^T \quad (19)$$

使用式(6)求解因子 $q(\boldsymbol{\pi},\boldsymbol{\mu},\mathbf{A})$ 为:

$$\ln q^*(\boldsymbol{\pi},\boldsymbol{\mu},\mathbf{A})=\sum_{k=1}^K\sum_{n=1}^Nr_{nk}\ln \mathcal{N}(\mathbf{x}_n|\boldsymbol{\mu}_k,\mathbf{A}_k^{-1})+\text{常数}+ \\ E_{\mathbf{Z}}[\ln p(\mathbf{Z}|\boldsymbol{\pi})]+\ln p(\boldsymbol{\pi})+ \\ \sum_{k=1}^K\ln p(\boldsymbol{\mu}_k,\mathbf{A}_k) \quad (20)$$

分离式(20)中与 $\boldsymbol{\pi}$ 有关的项,可得:

$$\ln q^*(\boldsymbol{\pi})=(\alpha_0-1)\sum_{k=1}^K\ln \pi_k+\sum_{n=1}^N\sum_{k=1}^Kr_{nk}\ln \pi_k+\text{常数} \quad (21)$$

对两侧取指数,得到狄利克雷分布:

$$q^*(\boldsymbol{\pi})=Dir(\boldsymbol{\pi}|\boldsymbol{\alpha}) \quad (22)$$

其中, $\boldsymbol{\alpha}$ 为向量,每一个元素为 α_k ,初始化时都设置为 α_0 ,迭代时为 $\alpha_k=\alpha_0+N_k$, N_k 表示训练集中属于第 K 类轴承状态的特征输入向量个数,初始化时为 $N/4$ 。从式(20)中分离出与 $\boldsymbol{\mu}_k$ 和 $\mathbf{A}_k$ 有关的项并取指数,得到高斯-Wishart 分布,形式为:

$$q^*(\boldsymbol{\mu}_k,\mathbf{A}_k)=q^*(\boldsymbol{\mu}_k|\mathbf{A}_k)q^*(\mathbf{A}_k) \\ =\mathcal{N}(\boldsymbol{\mu}_k|\mathbf{m}_k,(\beta_k\mathbf{A}_k)^{-1})\mathcal{W}(\mathbf{A}_k|\mathbf{W}_k,\nu_k) \quad (23)$$

其中,参数 $\mathbf{W}_k, \beta_k, \mathbf{m}_k, \nu_k, \bar{\mathbf{x}}_k, \mathbf{S}_k, N_k$ 的计算公式为:

$$\mathbf{W}_k^{-1}=\mathbf{W}_0^{-1}+N_k\mathbf{S}_k+\frac{\beta_0N_k}{\beta_0+N_k}(\bar{\mathbf{x}}_k-\mathbf{m}_0)(\bar{\mathbf{x}}_k-\mathbf{m}_0)^T \quad (24)$$

$$\beta_k=\beta_0+N_k \quad (25)$$

$$\mathbf{m}_k=\frac{1}{\beta_k}(\beta_0\mathbf{m}_0+N_k\bar{\boldsymbol{\mu}}_k) \quad (26)$$

$$\nu_k=\nu_0+N_k \quad (27)$$

步骤 2 为了完成变分推断的迭代过程,需要更新参数 r_{nk} ,其计算式为:

$$r_{nk}\propto\pi_k\bar{\mathbf{A}}_k^{1/2}\exp\{-\frac{D}{2\beta_k}-\frac{\nu_k}{2}(\mathbf{x}_n-\mathbf{m}_k)^T\mathbf{W}_k(\mathbf{x}_n-\mathbf{m}_k)\} \quad (28)$$

其中:

$$\ln \tilde{\mathbf{A}}_k=\sum_{d=1}^D\psi(\frac{\nu_k+1-d}{2})+D\ln 2+\ln |\mathbf{W}_k| \quad (29)$$

$$\ln \tilde{\pi}_k=\psi(\alpha_k)-\psi(\alpha) \quad (30)$$

其中, $\psi(\cdot)$ 是 Digamma 函数, $\alpha=\sum_k\alpha_k$ 。混合高斯分布的变分下界为:

$$\mathcal{L}=L(q(\boldsymbol{\Theta})) \\ =\sum_{\mathbf{Z}}\iiint q(\mathbf{Z},\boldsymbol{\pi},\boldsymbol{\mu},\mathbf{A})\ln\{\frac{p(\mathbf{X},\mathbf{Z},\boldsymbol{\pi},\boldsymbol{\mu},\mathbf{A})}{q(\mathbf{Z},\boldsymbol{\pi},\boldsymbol{\mu},\mathbf{A})}\}d\boldsymbol{\pi}d\boldsymbol{\mu}d\mathbf{A} \\ =E[\ln p(\mathbf{X},\mathbf{Z},\boldsymbol{\pi},\boldsymbol{\mu},\mathbf{A})]-E[\ln q(\mathbf{Z},\boldsymbol{\pi},\boldsymbol{\mu},\mathbf{A})] \quad (31)$$

更新 r_{nk} 值后,使用更新后的模型参数计算变分下界 $\mathcal{L}^{(n)}$ 及与上一次变分下界的差值 $\Delta^{(n)}=\mathcal{L}^{(n)}-\mathcal{L}^{(n-1)}$,上标为第 n 次迭代,若 $\Delta^{(n)}$ 小于预设阈值,说明收敛,迭代完成并输出模型,否则 $\Delta^{(n)}$ 大于预设阈值,再进行下一次迭代^[17]。

3.3 轴承故障诊断

针对轴承的每一种状态,选择 D 个敏感特征组成特征集,分别对轴承振动信号时间序列和若干个 ISC 时间序列提取 D 个特征,所有特征点组成一个特征输入向量。对轴承的 4 种状态在提取相同数量的特征输入向量,并训练模型。对于第 n 个待诊断的轴承诊断数据,提取同样的特征输入向量,分别在不同轴承状态下计算式(14)、式(16)得到 r_{nk} ,即输入向量属于不同轴承状态的概率 r_{n1}, r_{n2}, r_{n3} 和 r_{n4} ,选择最大概率诊断为该状态。

4 实验分析

4.1 故障数据仿真

使用凯斯西储大学的滚动轴承振动数据进行仿真。振动信号由驱动端加速度传感器采集,轴承型号为 SKF6205,电机转速为 1 772 r/s,使用电火花加工技术在驱动端轴承上布置了单点故障。内圈、滚动体及外圈故障的直径都为 0.007 英寸,外圈故障位于 6 点钟位置,采样频率为 12 kHz,实际故障频率:内圈为 159.93 Hz,滚动体为 139.20 Hz,外圈为 105.88 Hz。训练模型时,将轴承振动信号 200 个采样点作为一组,从凯斯西储大学原始数据中共提取 240 组特征输入向量组成训练样本,其中轴承的 4 种状态各占 60 组。测试模型时,从原始数据中另外提取 240 组组成测试数据,其中轴承的 4 种状态也各占 60 组。训练数据按组分别进行 LCD 预处理,按照 LCD 分解级数选取不同数量的 ISC 和原始数据分别提取时域特征组成特征输入向量和训练集。表 1 是轴承内圈故障的第 i 个 ISC 与原始信号的互相关系数表,从中可知,前 5 个 ISC 有相对更大的互相关系数($c_i \geq \max(c_i)/10$),因此本文 LCD 的最大分解级数为 6。

表 1 内圈故障信号与 ISC 之间的互相关系数

Table 1 Cross-correlation coefcient between ith ISC and inner ring fault signal								
ISC	ISC1	ISC2	ISC3	ISC4	ISC5	ISC6	ISC7	ISC8
$corr_i$	0.8412	0.5533	0.3448	0.2312	0.0895	0.0502	0.0239	0.0086

注: $corr_i$ 表示第 i 个 ISC 与原始信号之间的互相关系数

使用训练集进行变分贝叶斯模型训练,因为 LCD 分解级数不同可得到不同的模型。将模型参数初始化值设置为:随机生成 $r_{nk}, \alpha_0=1, \beta_0=1, \mathbf{m}_0$ 为所有特征的均值组成的向量, $\mathbf{W}_0=0.0001 * \text{diag}(1,1 \cdots 1), \mathbf{v}_0$ 为特征个数。对测试数据按照训练集所使用的 LCD 分解级数提取同样的特征,并分别计算在不同模型中的诊断正确率。变分下界的迭代收敛过程如图 1 所示,迭代预设阈值为 1×10^{-8} 。

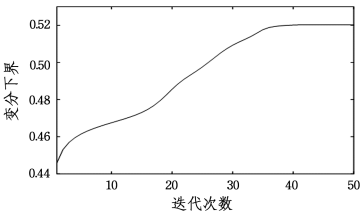


图 1 变分下界的迭代收敛过程

统计 240 组测试数据中轴承运行状态诊断正确的组数,诊断正确组数的比例即为诊断正确率。本文所提方法的诊断正确率如表 2 所列,原始信号提取 10 个时域特征组成训练集时,正确率最高能达到 99.6%。使用支持向量机作为分类器进行对比,支持向量机的损失函数参数为 1.2;使用高斯核函数,训练不同模型时变分贝叶斯的超参数、支持向量机的参数和模型设置保持不变。对原始信号和前 6 个 ISC 提取 70 个特征组成特征集训练得到的模型诊断结果中,采用支持向量机作为分类器进行诊断的正确率仅为 59.6%;而本文所提算法的诊断正确率为 99.2%,诊断正确率提升了 39.6%。

表 2 不同 ISC 训练集的诊断正确率

Table 2 Diagnostic accuracy of different ISC training sets			
训练集	特征个数	变分贝叶斯的正确率/%	支持向量机的正确率/%
原始信号	10	99.6	97.9
原始信号 ISC1	20	99.6	95.8
原始信号 ISC1 ISC2	30	99.6	89.1
原始信号 ISC1 ISC2 ISC3	40	99.6	79.6
原始信号 ISC1 ISC2 ISC3 ISC4	50	99.6	70.8
原始信号 ISC1 ISC2 ISC3 ISC4 ISC5	60	99.6	62.1
原始信号 ISC1 ISC2 ISC3 ISC4 ISC5 ISC6	70	99.2	59.6

对 240 组测试集数据的原始信号提取 10 个特征作为特征集并训练模型,使用多维尺度法 (Multidimensional Scaling, MDS)^[18] 将诊断结果的 10 维数据可视化在二维空间中,如图 2 所示。从图 2 可以看出,4 种状态分布具有明显的区分度,说明本文提出的方法可以很好地判别 4 种不同的轴承状态。

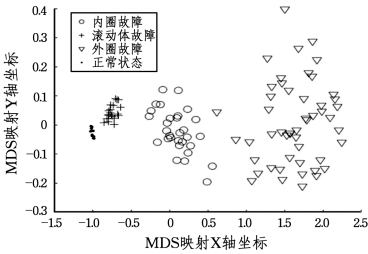


图 2 高维测试集数据的二维映射

Fig. 2 Two dimensional mapping of high dimensional test set data

凯斯西储大学的数据是在实验室理想条件下得到的,为了进一步证明本文方法的适用性,使用该方法对中国铁路某局机车滚动轴承实测数据进行了诊断。该数据为中国铁路某局 2014 年某月的测量数据,只有轴承故障和正常两个状态,并没有对故障进行细分。该实测数据是经过共振解调处理过的^[19],图 3 是其振动信号 200 个采样点的时间序列波形。

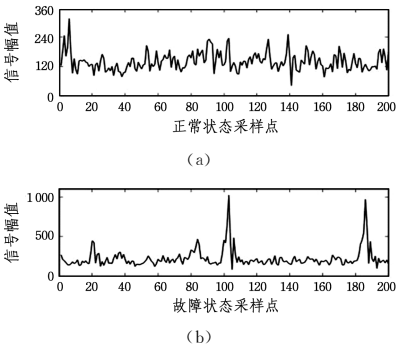


图 3 轴承振动信号的时间序列

Fig. 3 Bearing vibration signal time series

使用 60 组故障数据和 60 组正常数据作为训练样本;另外 60 组故障数据和 60 组正常数据作为测试数据。表 3 是该局使用本文算法和支持向量机在不同数量特征集下的诊断正确率,训练不同模型时变分贝叶斯的超参数、支持向量机的参数和模型设置保持不变。由表 3 可知,使用原始信号和 ISC1 取 20 个特征组成特征集时的诊断正确率最高为 91.7%;而在同样的特征和参数值下,支持向量机算法的诊断正确率最

高仅为 75%。在实测数据上的仿真结果表明,本方法依然有较高的故障诊断正确率。

表 3 不同 ISC 训练集的诊断正确率

训练集	特征 个数	变分贝叶斯的 正确率/%	支持向量机的 正确率/%
原始信号	10	72.5	75.0
原始信号 ISC1	20	91.7	73.7
原始信号 ISC1 ISC2	30	91.6	50
原始信号 ISC1 ISC2 ISC3	40	91.6	50
原始信号 ISC1 ISC2 ISC3 ISC4	50	90.0	50
原始信号 ISC1 ISC2 ISC3 ISC4 ISC5	60	77.5	50
原始信号 ISC1 ISC2 ISC3 ISC4 ISC5 ISC6	70	74.7	50

4.2 结果分析

本文对凯斯西储大学的滚动轴承振动数据进行仿真实验,当特征数为 70 时,与基于支持向量机的方法相比,基于变分贝叶斯的方法的诊断正确率提升了 39.6%。当特征数为 70 个时,使用 MDS 多维标度法将训练集的 70 维数据可视化在二维空间中,发现某些特征对于 4 种轴承运行状态的辨识度降低,不同类型的轴承运行状态特征点出现了重叠分布,并且随着 LCD 分解级数的增多,ISC 与原始轴承振动信号的相关性降低,故 ISC5 和 ISC6 提取的特征分辨率会降低,特征质量较差。此时,70 个特征分辨率的情况为:部分特征对 4 类轴承状态具有较高的分辨率,部分特征分辨率较低且不同类型的轴承状态特征点出现重叠。在这种情况下,基于变分贝叶斯的方法依然有较高的诊断正确率,说明文中所提方法对特征集中存在分辨率较差特征的情况也有良好的诊断效果,具有更好的鲁棒性。

结束语 本文针对现有智能诊断方法存在的问题,提出了一种基于变分贝叶斯的轴承故障诊断方法,从一个新的角度实现了滚动轴承故障的智能诊断。本方法可以有效地对轴承故障进行诊断,对实际数据也有良好的诊断效果,在工况不理想的情况下也可以实现准确的轴承故障诊断。本文未对分辨率较差的特征作进一步分析,有必要从统计分布角度提出一种特征质量评估的方法。

参 考 文 献

[1] ZHANG R,TAO H,WU L,et al. Transfer Learning with Neural Networks for Bearing Fault Diagnosis in Changing Working Conditions[J]. IEEE Access,2017,5:14347-14357.

[2] SUN S S,HE G H,CUI J. Fault diagnosis method of rolling bearing based on dual-tree rational-dilation complex wavelet packet transform and SVM[J]. Computer Science,2015(S2): 131-134. (in Chinese)

孙珊珊,何光辉,崔建. 基于有理双树复小波和 svm 的滚动轴承故诊断方法[J]. 计算机科学,2015(S2):131-134.

[3] FU Q,JING B,HE P,et al. Fault feature selection and diagnosis of rolling bearings based on EEMD and optimized Elman_Ada-Boost algorithm[J]. IEEE Sensors Journal,2018,18(12):5024-5034.

[4] GHAHRAMANI Z. Probabilistic machine learning and artificial intelligence[J]. Nature,2015,521(7553):452-459.

[5] KONG H H,SHI H B. A Comparative Study of Generative and Discriminative Classification Methods [J]. Sci-Tech Information Development & Economy,2010,20(8):78-81. (in Chinese)

孔环环,石洪波. 产生式与判别式分类方法比较研究[J]. 图书情报导刊,2010,20(8):78-81.

[6] QIU L,GAO S,CAO C G. Improved Estimation of Distribution Algorithms Based on Normal Distribution [J]. Computer Science,2015,42(8):32-35. (in Chinese)

邱玲,高尚,曹存根. 改进的正态分布的分布估计算法[J]. 计算机科学,2015,42(8):32-35.

[7] WU Z T,CHENG J S,LI B Q,et al. The method of generalized local characteristic-scale decomposition and its application[J]. Journal of Vibration Engineering,2016,29(2):331-339. (in Chinese)

吴占涛,程军圣,李宝庆,等. 广义局部特征尺度分解方法及其应用[J]. 振动工程学报,2016,29(2):331-339.

[8] LUO S,CHENG J. VPMCD based novelty detection method on and its application to fault identification for local characteristic-scale decomposition[J]. Cluster Computing,2017,20(4):2955-2965.

[9] CHAMBERS J M. Graphical Methods for Data Analysis [M]. Chapman and Hall/CRC,2017.

[10] PAOLELLA M S. New graphical methods and test statistics for testing composite normality[J]. Econometrics,2015,3(3):532-560.

[11] MA Y,ZHAO S,HUANG B. Multiple-Model State Estimation Based on Variational Bayesian Inference[J]. IEEE Transactions on Automatic Control,2018,64(4):1679-1685.

[12] CHEN Y,GUO Q,SUN H,et al. A Distributionally Robust Optimization Model for Unit Commitment Based on Kullback-Leibler Divergence[J]. IEEE Transactions on Power Systems,2018,33(5):5147-5160.

[13] LIM K L,WANG H. MAP approximation to the variational Bayes Gaussian mixture model and application[J]. Soft Computing,2017(3):1-13.

[14] WU G. Fast and scalable variational Bayes estimation of spatial econometric models for Gaussian data [J]. Spatial Statistics,2018,24:32-53.

[15] WU W,NAGARAJAN S,CHEN Z. Bayesian Machine Learning:EEG/MEG signal processing measurements[J]. IEEE Signal Processing Magazine,2016,33(1):14-36.

[16] KANGIN D,MARKARIAN G. Multi-Bernoulli filter for group object tracking and its Gaussian-Wishart implementation[C]// International Joint Conference on Neural Networks. IEEE,2017:3161-3168.

[17] KAMIL A,VINCENT E. Variational Bayesian Inference for Source Separation and Robust Feature Extraction[J]. IEEE/ACM Transactions on Audio Speech & Language Processing,2016,24(10):1746-1758.

[18] STURROCK K,ROCHA J. A Multidimensional Scaling Stress Evaluation Table:[J]. Field Methods,2016,12(1):49-60.

[19] LI J R,YUE J H. Fault diagnosis algorithm for the axle box bearing of walking unit in EMU based on HHT and resonance demodulation method[J]. Journal of Beijing Jiaotong University,2017,41(4):85-90. (in Chinese)

李佳睿,岳建海. 基于 HHT 及共振解调方法的动车组走行部轴箱轴承故障诊断算法[J]. 北京交通大学学报,2017,41(4):85-90.