

# 电特征分析与机器学习助力数控铣床异常检测

马志勇, 程 旺, 郑海波

(山东法因数控机械设备有限公司, 山东 济南 250100)

**摘 要** 电特征分析是一种强大的工具, 它利用机器的电压和电流信号来推断其健康状况。电特征分析可以作为预测性维护工具, 用于在早期检测数控铣床常见故障, 从而防止灾难性故障和生产中断, 并延长设备使用寿命。本研究提出了一种将电特征分析和机器学习应用于数控铣床工作状态监测和健康状况评估的新方法, 实验结果表明了该方法的有效性。

**关键词** 异常检测; 状态监测; 健康状况评估; 机器学习; 数控铣床

**中图分类号**: TP3; TG54

**文献标志码**: A

**DOI**: 10.3969/j.issn.2097-3365.2025.22.002

## 0 引言

在工业系统中, 设备突然发生故障可能引发诸多严重问题。一方面, 会导致昂贵的停机时间成本, 打乱正常的生产节奏, 造成巨大的经济损失; 另一方面, 故障还可能对周围设备造成损坏, 甚至对操作人员构成安全威胁<sup>[1]</sup>。因此, 对工业系统的可靠性和可用性的要求与日俱增。

为确保设备有效运行、避免过早出现故障, 持续监测机器的健康状况至关重要。因为一旦设备发生故障, 不仅会扰乱生产线, 还会带来高昂的维修成本。在电气设备故障诊断领域, 振动分析和声发射是较为常用的方法, 然而, 这两种方法存在明显弊端, 它们在操作复杂性以及所使用传感器的成本方面都相对较高。

电特征分析是一种极具优势的故障检测手段, 它通过捕捉设备的电源信号, 如电流、电压等, 并对这些信号进行分析, 从而实现对设备故障或早期故障的有效检测。电特征分析可作为预测性维护工具, 在设备出现常见故障的早期阶段就及时发现问题, 进而预防可能出现的昂贵灾难性故障和生产中断情况, 延长设备的使用寿命<sup>[2]</sup>。与传统故障诊断方法相比, 电特征分析只需监测电流和电压信号, 大大减少了所需安装的传感器数量, 而且仅需接入电源线路, 无需直接接触设备, 就能在不干扰设备正常运行的前提下完成诊断和预测分析工作。因此, 电特征分析是一种可实现在线甚至远程监测的、真正意义上的无侵入式检测方法。本文重点介绍经验模态分解在计算机数字控制铣床状态监测和健康状况评估方面的创新性应用。

## 1 实验设备与数据采集系统

案例研究机器对象是 3 轴的哈斯迷你铣床, 该机器主要规格数据见表 1 所示。

表 1 机器主要规格数据

规格	值	规格	值
最大主轴转速	6 000 Rpm	Z—轴行程	254 mm
最大切削进给速度	12.7 m/min	表大小	305*914 mm
轴功率	5.6 kW	最大重量	227 kg
X—轴行程	406 mm	输入最低电压	220 V
Y—轴行程	305 mm	输入最高电压	440 V

由于哈斯迷你铣床使用的伺服电机控制器频率限制在 30 kHz, 根据奈奎斯特准则建议以信号最高频率的两倍收集数据<sup>[3]</sup>, 确定 60 kHz 为可能的频率下限。由于数据来自 7 个通道, 因此需要 420 kHz 的最大串行频率来处理多路数据采集系统。考虑到这一点, 选择 70 kHz 的采样频率。数据采集系统使用 Raspberry Pi 4B 数据处理平台, 选择它是因为它有两个内置的 USB3.0 端口、数字 IO、四核和高频 CPU, 选择 USB-1608GX 进行数据采集, 1608GX 的显著特性包括其最大串行测量频率为 500 kHz, 电压扫描范围为  $\pm 10$  V, 以及通过专有的函数库进行通信。更高的频率允许所有 7 个通道以两倍于铣床估计的最高频率进行采样。该工厂的机械图由 4 个电流输入和 4 个电压输入构成。

模式识别一种基于过程历史的方法被用于进行数据分析: 一个数学模型, 它解释了测量输入和输出之间的关系, 是在大量的历史过程数据的基础上进行训练的<sup>[4]</sup>。这个模型必须是描述性的, 从数据中提取知识, 并且必须是预测性的, 以便进行预测。在工程领域, 基于过程历史的方法是指模式识别方法。这种方法涉及将由可测量表征的模式分类为有限数量的类别称为标签, 按照以下步骤进行: (1) 通过适当的测量设置

进行数据收集；(2) 从收集的数据开始进行特征提取；(3) 特征选择：对特征进行筛选以识别出最重要的特征；(4) 分类器设计：分类器的训练；(5) 对所开发的分类器的系统评估。

## 2 数据分析与模型构建流程

### 2.1 数据采集策略

这台机器被编程以三个方向执行一系列动作：(1) 沿 X 轴从坐标 (0, 0, 0) 移动到坐标 (10, 0, 0)，进刀量为 300 英寸 / 分钟 (指令 X10)；(2) 沿 Y 轴从坐标 (0, 0, 0) 移动到坐标 (0, 10, 0)，进刀量为 300 英寸 / 分钟 (指令 Y10)；(3) 以 300 英寸 / 分钟的进刀量沿 Z 轴从坐标 (0, 0, 0) 移动到坐标 (0, 0, 10) (指令 Z10)；(4) 沿 X 轴从坐标 (0, 0, 0) 移动到坐标 (-10, 0, 0)，进刀量为 300 英寸 / 分钟 (指令 X-10)；(5) 沿 Y 轴从坐标 (0, 0, 0) 移动到坐标 (0, -10, 0)，进刀量为 300 英寸 / 分钟 (指令 Y-10)；(6) 沿 Z 轴从坐标 (0, 0, 0) 移动到坐标 (0, 0, -10)，进刀量为 300 英寸 / 分钟 (指令 Z-10)；(7) 主轴以 7000 转 / 分钟的转速旋转 (S7000M03)；(8) 主轴以 6000 转 / 分钟的转速旋转 (S6000M04) 持续五秒。

通过采集系统，收集了每次操作中与机器电流相关的数据，以表征其电特征。这些数据包括三相电流和中性电流。对于每个动作，总共记录了 500 次重复 ( $r=500$ )。此外，还在机器闲置时进行了数据采集，以表征基线噪声。

### 2.2 特征提取方法

处理在实验活动期间收集的数据，以提取和综合信号中的信息内容，信息内容通过显著特征的生成来表示。在执行时间序列特征提取的整合技术中，检测两个时间序列相似性的技术被广泛使用。这些技术包括计算时间序列与参考信号在希尔伯特空间中的距离作为相似性的度量<sup>[5]</sup>。为了消除不太重要的频段，进行了初步的数据处理。信号滤波的方法包括识别最重要的信号频段，并创建带通滤波器来消除最不重要的频段，实际上这种技术提供了关于信号的频率内容以及如何随时间变化的信息。

所有信号的频谱图显示，最有趣的频段是 0 至 15 千赫兹之间的频段，因此设计了一个数字低通滤波器来截断高于 15 千赫兹的频率。在分析的特征提取技术中，选择了多维动态时间规整，因为它允许对不同长度的非相位信号进行比较。参考向量是通过每个动作的获取向量的均值生成的，分析结果总结在一个矩阵中，其中  $p=8$  是通过多维动态时间规整提取的相似性

指标的数量， $n=r*p$ 。这些为每个获取向量提取的相似性指标被用作特征来发展分析。另外，添加了一列以提供每行的标签。该标签是一个索引，表示与每行相关联的操作。

### 2.3 特征筛选优化

提取的特征经过处理，以选择更相关的特征，或者一个候选特征子集。通过这种方式，可以加快数据挖掘算法，提高预测准确性，并增加可理解性。相对于这个阶段，可以使用两种方法：一种是生成新特征作为已存在特征的组合，例如主成分分析和线性判别分析；另一种是通过优化目标函数来识别最佳特征子集，例如顺序前向选择和顺序后向消除。

通过使用最小冗余最大相关性算法对这两种技术的结果进行了比较，该方法通过最大化与目标变量的相关性，同时最小化所选变量之间的冗余来迭代扩展所选解释变量的集合。需要注意的是，已对每个特征的数据进行了初步标准化。

原始数据需要 4 个特征来解释，而通过主成分分析或线性判别分析处理后的数据分别只需要 2 个和 3 个特征。鉴于此，继续对通过主成分分析处理后的数据进行分析至关重要。详细来说，通过对数据应用主成分分析，在保证 92% 的信息含量的同时，将特征数量从 8 个减少到 2 个成为可能。

### 2.4 分类器设计与训练

处理后的数据用于训练一个能够确定向量标签的分类器，而向量标签是事先未知的。监督学习被用于完成这一任务。它指的是两种方法：参数化方法，涉及对分布形状的假设的制定，并从训练集中估计关键参数，然后使用贝叶斯规则；非参数化方法，利用训练集中包含的信息进行分类：这种方法基于相似输入将产生相似输出的假设。属于第一种方法的技术更简单，不需要大型数据集；属于第二种方法的技术需要大型数据集，但能确保更高的灵活性和性能。由于在案例研究中，有大量的数据可用，但没有先验知识，因此选择了参数化方法。基于非参数方法的主要技术包括：支持向量机 (SVM)、人工神经网络 (ANN)、k 近邻 (k-NN)、随机森林 (RF)、集成方法。通过对数据矩阵实现 (k-NN)、线性支持向量机 (L-SVM)、二次支持向量机 (Q-SVM)、三次支持向量机 (C-SVM)、随机森林和人工神经网络来训练分类器。对每种技术的性能进行了测试，在训练阶段使用了交叉验证方法以消除过拟合的风险。

用于比较模型的性能指标是预测的准确性，其计

算方式为正确预测的数量与总预测数量的比率。表2总结了第一次比较不同分类器预测的准确性。只有那些在此次评估中报告准确率为100%的模型才被纳入系统评估阶段。

表2 第一次比较不同分类器预测的准确性

模型	精度
数据矩阵实现	100%
二次支持向量机	100%
三次支持向量机	100%
随机森林	98.61%
人工神经网络	97.22%
线性支持向量机	92.12%

## 2.5 系统综合评估

系统评估是通过另一项实验活动进行的,其中为每个动作收集了100次重复数据。为了与第一个数据集对比,这个新数据集也经过了标准化和主成分分析。第二次比较的结果总结在表3中。基于支持向量机的两种分类器结果证明比基于k近邻的分类器具有更高的准确性。由于监督机器学习模型具有度数更低的核函数,结果证明其在其他因素相同的情况下,计算时间更短,因此监督机器学习模型被选为模型。

表3 第二次比较不同分类器预测的准确性

模型	精度
数据矩阵实现	98.61%
二次支持向量机	97.22%
三次支持向量机	92.12%

## 3 分类器在工业应用中的关键环节

### 3.1 机器状态实时检测

为了在工业环境中应用该分类器,它需要能够实时进行信号分析。这需要一个算法,当所检测的电流发生改变时,能够激活分类器。这种改变将由一系列异常值表示,这些异常值与机器的正常工作条件和空闲状态相关的工作条件相比是不正常的。

为了开发激活算法,决定使用机器学习来创建异常检测模型。这些异常数据块是实时信号的一部分,作为输入提供给先前训练好的分类器,以确定它们的标签。这些算法应用于与空闲机器相关的采集数据。结果是一个模型的训练,该模型能够区分空闲和异常行为。为了减少误报,创建了异常数据识别模型,以便它们能够检测到持续时间至少半秒的异常数据块。

测试是通过将模型应用于先前获取的每个动作的

数据来进行的。用于比较模型的性能指标是检测的准确率,计算为正确检测数量与总检测数量的比率。LOF的准确率为91%,OCSVM为88%,而IF为89%。基于这些结果,LOF是被选择的模型。

### 3.2 设备健康状况评估

在开发了用于识别机器实时状态的分类器和模型之后,就可以对信号中的异常进行识别了。机器状态检测功能识别待分类信号部分并触发模式识别;模式识别使识别出的信号部分能够确定其参考动作;最后,健康状况评估功能确定机器在正常运行条件下是否存在异常。

在案例研究中,健康状况评估包括两个阶段:第一阶段是相似性检测,由多维动态时间规整进行,用于检测获取的向量与特定动作相关的参考向量之间的相似性;第二阶段是异常值检测,基于图基的栅栏法。模型验证包括向获取信号的一个子集中引入人工异常,以测试系统准确检测偏差的能力。用于比较模型的性能指标是异常值检测的准确率,计算为正确检测数量与总检测数量的比率。结果显示准确率为97%。

## 4 结束语

电特征分析和机器学习应用于数控铣床工作状态监测和健康状况评估,机械状态检测功能的估计准确率为91%,模式识别功能的估计准确率为98%,健康状况评估功能的估计准确率为97%。即使在工业环境中以如此简单的形式应用这样一个系统,也可以通过减少故障的发生频率和严重程度以及提高机器的可用性来节省大量成本。所提出的方法的主要优势之一是,即使是较旧的设备也能轻松配备所需的数据采集设备,从而实现具有成本效益的诊断和预测系统,而无需昂贵和复杂的硬件基础设施。

## 参考文献:

- [1] 余国瑞,曾学淑,崔欢欢.人工智能在数控加工程序优化中的应用分析[J].机电产品开发与创新,2024,37(05):124-126,130.
- [2] 朱翔.基于多种机器学习算法的铣刀磨损故障监测[D].合肥:安徽大学,2020.
- [3] 余捷.哈斯科技扎根PPM[J].制造技术与机床,2013(09):13-14.
- [4] 李康举.数控铣床切削稳定性研究[D].沈阳:东北大学,2011.
- [5] 徐博,唐浩,严家霖,等.基于希尔伯特半张量压缩感知的亚采样率采集技术[J].信号处理,2024,40(10):1846-1854.