DOI: 10.3901/JME.2019.14.001

基于压缩感知和加噪堆栈稀疏自编码器的 铣刀磨损程度识别方法研究*

李宏坤 郝佰田 代月帮 杨蕊
(大连理工大学机械工程学院 大连 116024)

摘要: 数控机床在加工过程中, 刀具磨损会对被加工零件的表面质量、尺寸精度产生巨大影响, 而传统依靠切削力系数来分析刀具磨损的方法, 需要在工作台上安装额外测力装置, 这将干扰机床正常加工, 限制被加工零件尺寸, 引起加工质量降低等问题, 限制了其在实际工业环境中的应用。针对上述问题, 提出利用主轴电流结合深度学习网络识别铣刀磨损状态的监测方法。首先, 理论论证利用主轴电流代替切削力识别刀具磨损的可行性; 然后, 利用压缩感知对电流信号的频域数据进行数据压缩, 其中为提高网络的鲁棒性, 对观测信号添加高斯白噪音; 最后, 将压缩后的数据输入堆栈稀疏自编码网络, 利用有监督学习与无监督学习相结合的方法, 提取刀具磨损所引起的特征信息, 用于表征刀具磨损程度。试验结果表明, 该方法可以有效对铣刀磨损程度进行识别。

关键词: 主轴电流; 铣刀磨损; 压缩感知; 堆栈稀疏自编码器

中图分类号: TG156

Wear Status Recognition for Milling Cutter Based on Compressed Sensing and Noise Stacking Sparse Auto-encoder

LI Hongkun HAO Baitian DAI Yuebang YANG Rui

(School of Mechanical Engineering, Dalian University of Technology, Dalian 116024)

Abstract: In the machining process of CNC machine tools, the state of tool wear has a great influence on the surface quality and dimensional accuracy of the parts being machined. However, the traditional method of relying on cutting forces to analyze tool wear requires the installation of additional force measuring devices on the workbench. This will interfere with the normal machining of the machine tool, limit the size of the part being machined, cause the processing quality to be reduced and other issues, and limit its application in the actual industrial environment. Aiming at the above problems, a monitoring method for identifying the wear state of the milling cutter by using the spindle current combined with the deep learning network is proposed. Firstly, the feasibility of using spindle current instead of cutting force to identify tool wear is demonstrated. Then, the frequency domain data of the current signal is compressed by compressed sensing, and Gaussian white noise is added to the observation signal to improve the robustness of the network. Finally, the compressed data is input into a stack sparse auto-encoder, which uses a combination of supervised learning and unsupervised learning to extract feature information caused by tool wear and to characterize tool wear. The experimental results show that this method can effectively monitor the wear state of the milling cutter.

Key words: spindle current; milling cutter wear; compressed sensing; stack sparse auto-encoder

0 前言

近年来, 深度学习在特征提取与模式识别方面显示出独特的优势与潜力, 将深度学习应用于生产制造过程, 实现生产过程的状态检测的研究已初现

端倪, 智能制造、高精度制造日益成为制造行业的主流趋势^[1]。数控机床加工过程中最基本的要素为机床、夹具、工件和刀具, 而刀具作为与工件直接接触的零件, 在加工过程中非常容易损坏。刀具损坏最常见的形式是刀具磨损, 它是一种在机械、热、化学、磨粒等共同作用于刀刃而产生的不可避免的刀具退化现象^[2]。数控机床刀具磨损不仅制约加工精度的提高和自动化的发展, 而且如果不能及时发现, 极有可能造成工件报废, 严重限制着企业生

* 国家自然科学基金资助项目(51575075, U1808214)。20180702 收到初稿, 20190425 收到修改稿

产率及其利润的提高。因此,将深度学习方法在特征提取方面的优势应用到刀具磨损程度识别中去,逐渐成为该方向的研究热点。

针对数控机床刀具磨损状态的监测方法,国内外学者已经进行了大量的研究,有直接测量法和间接监测法两类。直接测量法虽然可以精确测得刀具磨损程度,但需要暂停生产、装夹刀具,在实际生产中不实用。间接监测法主要通过采集力信号、声发射信号、振动信号及电流信号等参数,对这些信号进行分析,实现监测目的。传统分析方法主要是使用时域、频域、时频域分析手段提取磨损特征,实现刀具磨损监测,其中基于力信号的铣刀磨损状态监测研究取得了许多的成果。如 ALTINTAS^[3]通过比较平均每齿合力的一阶差分和二阶差分的大小,实现了铣刀破损的实时监测;李宏坤等^[4]提取特征频带小波包能量,结合 logistic 回归模型评估刀具运行可靠性,很好的识别出了刀具的磨损状态。但依靠切削力信号来分析刀具磨损的方法,需要在工作台上额外增加测力装置,这将干扰机床正常加工,限制了其在实际工业环境中的应用。STEIN 等^[5-8]针对机床进给系统进行分析,建立了控制模型,提出利用进给电流参数对切削力的预测和刀具状态监测,验证了进给电流代替切削力监测刀具状态的可行性。

传统分析方法用到大量信号处理技术,人为选定信号特征参数^[9],需要具备专业知识的技术人员才能辨识出刀具磨损状况。因此,张存吉等^[10]提出对加工时的振动信号进行小波包变换,得到对应能量频谱图,然后采用卷积神经网络对预处理后的数据进行训练学习,对刀具磨损程度的识别正确率达到了 98%。林扬^[11]提出利用小波包变换之后的切削力信号,输入堆栈稀疏自编码器,识别精度识别正确率达到了 93%。以上所提出的方法尽管解决了人为选择信号特征参数的缺点,但由于加工过程中的切削力信号和主轴振动信号的采集,需要在数控机床加工端安装额外传感器,一定程度上限制了机床的加工。

针对以上不足,本文提出一种既方便数据实时采集,又能去除人为选择特征参数的刀具磨损监测方法,即利用压缩感知技术和加噪处理对主轴电流信号进行预处理,将处理后数据输入深度学习网络,实现铣刀磨损状态的监测。

1 机床主轴电流监测原理

刀具磨损后会影响到切削力变化,如能建立主轴

电机电流信号与切削力之间的关系,便可以证明切削电流可以反映刀具磨损状态。

铣削力会随着刀具旋转周期变化,并通过刀柄、主轴等传动装置传递至主轴电机,引起电机电流随切削力周期性变化。由于径向铣削力和轴向铣削力方向穿过刀具旋转中心,故不产生扭矩。因此,作用在刀具上的切向铣削力与主轴电机电流存在如下关系模型

$$K_t \cdot I_{rms} = T_m = J \cdot \frac{d\omega}{dt} + T_f + F_t \cdot R \quad (1)$$

式中, K_t 为电机扭矩常量; I_{rms} 为电机三相电流有效值; T_m 为作用在电机轴上扭矩; J 为主轴传动系统等效转动惯量; ω 为主轴角频率; T_f 为电机轴克服的摩擦扭矩; F_t 为切向铣削力; R 为铣刀半径。

本文中机床采用三相交流异步电动机作为主轴系统的动力源,假设电机的三相电流分别为 i_u , i_v , i_w , 则 I_{rms} 可以表示为

$$I_{rms} = \sqrt{(i_u^2 + i_v^2 + i_w^2) / 3} \quad (2)$$

由于式(1)中存在等效转动惯量 J 和摩擦扭矩 T_f 两未知项的存在,使测得的主轴电流不能直接代替铣削力进行计算。实际加工时,一次走刀中主轴转速一般恒定不变,主轴角加速度为零,即 $d\omega/dt$ 为零,故不需要考虑等效转动惯量。当主轴空转时,铣削力 F_t 为零,故有

$$K_t \cdot I_{rms0} = T_f \quad (3)$$

式中, I_{rms0} 为某转速下空切电流。

将式(4)代入式(1)中有

$$\begin{aligned} \Delta I &= F_t \cdot R / K_t \\ \Delta I &= I_{rms} - I_{rms0} \end{aligned} \quad (4)$$

式中, ΔI 为铣削电流增量。显然,切削电流增量与切向铣削力成正比,因此,在理论上可以使用切削电流增量代替切向铣削力反应刀具磨损状态。

2 数据预处理

2.1 压缩感知

压缩感知(Compressed sensing, CS)是利用一组特定波形的测量矩阵去感知信号,即将信号在给定波形上进行投影,得到一组压缩数据^[12]。压缩感知理论指出,如果信号在某个变换域内是稀疏的(非零数目较少),那么就可以利用一个与变换基不相关的测量矩阵将高维信号投影到低维空间,并可以通过求解最优问题,利用这些低维观测信号重构出原始信号,即表明观测信号包含了原始信号的足够信息^[13]。

图 1 为压缩感知原理示意图, ϕ 为 $M \times N$ 维测量矩阵, $\mathbf{x}$ 为 N 维可稀疏化的原始数据, $\mathbf{c}$ 为 M 维观测数据(压缩所得数据), $\bar{\mathbf{x}}$ 为 N 维重后数据, 为达到数据压缩目的, 应保证 $M < N$ 。当原始数据 $\mathbf{x}$ 本身不是可稀疏数据时, 总可以找到一个稀疏基方阵 ψ , 使得 $\mathbf{x}$ 在该稀疏基 ψ 上的变换系数 $\mathbf{s}$ 是稀疏的, 因此压缩感知方程可为

$$\mathbf{c} = \phi \mathbf{x} = \phi \psi \mathbf{s} \quad (5)$$

定义信号 $\mathbf{x}$ 中非零值个数为 S , 可以证明当测量长度 $M \geq S \log_{10}(N/S)$ 时, 有很大的概率从长度为 M 观测值 $\mathbf{c}$ 中恢复出信号 $\mathbf{x}$ ^[14]。对于不同的原始信号, 常用的测量矩阵也会不同, 选择的标准是保证压缩得到的观测信号能够尽可能地重构出原始信号, 即不破坏原始信号 $\mathbf{x}$ 中所包含的信息。对一维原始信号而言, 使用高斯测量矩阵可以在很大程度上保证了其基矢量和任意稀疏基 ψ 不相关, 与之类似的测量矩阵还有 Bernoulli 矩阵等^[14]。

图 1 压缩感知示意图

2.2 加噪处理

深度学习网络在训练过程中经常会出现过拟合现象, 这将降低网络的鲁棒性, 当系统受到干扰时, 网络对新采集到的样本识别准确率会降低^[15]。为提高网络的鲁棒性, 在第一层稀疏自编码器前端, 对输入训练数据集添加 M 种程度的高斯白噪声

$$n_i(t) = (k^2 i + k) \frac{\sigma}{M} n_{st}(t) \quad (6)$$

式中, k 是加噪强度参数, $i=1,2,\dots,M$ 用于控制加噪的不同强度, σ 为训练样本的标准方差。加噪后, 每一个组训练样本数据扩展为 $M+1$ 组样本 $\{c(t), c_1^*(t), \dots, c_M^*(t)\}$ 。

$$c_i^*(t) = c(t) + n_i(t) \quad (7)$$

很明显, 在对样本加入噪音之后, 训练样本数目也会随之增加, 进而增加网络训练时长, 因此应当合理选择加噪个数 M 和加噪强度 k 。

3 深度学习网络

在多种深度学习网络中, 堆栈稀疏自编码器(Stack sparse auto-encoder, SSAE)通过将有监督学习和无监督学习相结合的方法, 能够很好地逐层提取数据深层特征, 具有挖掘同类数据之间关联信息和

不同数据之间区别的能力^[8]。在堆栈稀疏自编码器提取特征之后, 连接 Softmax, 便可实现模式识别。堆栈稀疏自编码器由多层稀疏自编码器(Sparse auto-encoder, SAE)连接而成, 网络训练过程主要分为预训练和全局微调。预训练阶段是一个无监督学习的过程。首先, 数据输入第一层网络, 构建网络损失函数, 采用梯度下降法寻找损失函数局部最优值, 此时的权重和偏置作为该层网络参数的初始值。然后将上一层稀疏自编码器的隐藏层神经元作为下一层网络的输入, 逐层训练, 直至得到每一层网络的初始权重和偏置。全局微调阶段是一个有监督学习的过程, 以原始数据标签作为监督信号, 构建整个网络的损失函数, 通过梯度下降法调节各层稀疏自编码器的网络权重和偏置。

3.1 稀疏自编码器

单层自动编码器是一种无监督学习网络, 不需要事先对原始数据添加标签, 基础网络结构如图 2 所示。 $\mathbf{x} = (x_1, x_2, \dots, x_n)$ 为网络的 n 个输入层节点, $\mathbf{a} = (a_1, a_2, \dots, a_s)$ 为隐含层节点, $\hat{\mathbf{x}} = (\hat{x}_1, \hat{x}_2, \dots, \hat{x}_n)$ 为网络输出。自动编码器包含两个部分: 编码器和解码器。编码器负责在输入数据中提取特征信息, 解码器负责利用特征数据恢复输入数据。

图 2 单层自动编码器结构图

编码过程

$$\mathbf{a} = f_{\theta}(\mathbf{x}) = S_f(\mathbf{W}\mathbf{x} + \mathbf{b}) \quad (8)$$

式中, $\mathbf{a}$ 代表编码后提取的 s 维特征参数; S_f 代表编码用激活函数, 通常使用 Sigmoid 函数; $\mathbf{W}$ 表示输入层和隐含层之间的权重, 维数为 $s \times n$, 其初始值一般随机初始化为很小的值; $\mathbf{b}$ 表示隐含层偏置, 维数为 s , 同样对其进行随机初始化; 用 f_{θ} 代表编码过程, $\theta = (\mathbf{W}, \mathbf{b})$ 表示编码部分的参数。

解码过程

$$\hat{\mathbf{x}} = g_{\theta'}(\mathbf{a}) = S_g(\mathbf{W}'\mathbf{a} + \mathbf{b}') \quad (9)$$

式中, S_g 代表解码用的激活函数, 通常也使用 Sigmoid 函数; $\mathbf{W}' = \mathbf{W}^T$ 表示隐含层与输出层之间的权重, 维数为 $n \times s$; $\mathbf{b}'$ 表示输出层偏置, 维数为 n ; 用 g_θ 代表编码过程, $\theta' = (\mathbf{W}', \mathbf{b}')$ 表示编码部分的参数。

很明显, 自动编码器是一个对称的网络结构, 为尽可能减小解码出数据与输入数据之间的差异, 自动编码器损失函数被定义为

$$J(\mathbf{W}, \mathbf{b}) = \frac{1}{m} \sum_{i=1}^m \frac{1}{2} \|x^{(i)} - \hat{x}^{(i)}\|^2 + \frac{\lambda}{2} \sum_{l=1}^{n_l-1} \sum_{i=1}^{s_l} \sum_{j=1}^{s_{l+1}} (w_{ji}^{(l)})^2 \quad (10)$$

式中, m 表示训练样本个数; $x^{(i)}$ 、 $\hat{x}^{(i)}$ 分别表示第 i 个样本数据和网络输出数据; λ 表示权重衰减参数, 用于控制权重衰减项与网络重构误差之间的相对重要性; n_l 表示网络层数, 自动编码器中 $n_l = 3$; s_l 、 s_{l+1} 分别表示第 l 层和第 $l+1$ 层网络神经元节点个数; $w_{ji}^{(l)}$ 表示第 l 层的第 j 个神经元与 $l+1$ 层的第 i 个神经元之间的权重。

公式前一项是网络重构误差, 后一项是权重衰减项, 其目的是减小权重幅度, 防止过拟合。

为提高网络训练效率, 充分挖掘原始数据中的特征信息, 增加特征数据的可分性, 稀疏自编码器在自动编码器的基础上对网络隐含层神经元添加强制性稀疏限制, 迫使只有少数神经元被激活^[16]。引入一个接近于 0 的参数——稀疏性参数 ρ , 加入限制性条件。

$$\hat{\rho}_j = \rho \quad (11)$$

式中, $\hat{\rho}_j$ 表示在 m 个样本输入下, 第 j 个隐藏层神经元的平均活跃度。

$$\hat{\rho}_j = \frac{1}{m} \sum_{i=1}^m [a_j^{(2)}(x^{(i)})] \quad (12)$$

式中, $a_j^{(2)}(x^{(i)})$ 表示在第 i 个样本 $x^{(i)}$ 输入时, 第 j 个隐含层神经元的激活度。

在增加稀疏性限制之后, 对损失函数做出修改, 增加稀疏惩罚因子

$$\sum_{j=1}^{s_2} KL(\rho \| \hat{\rho}_j) = \sum_{j=1}^{s_2} \left(\rho \log_{10} \frac{\rho}{\hat{\rho}_j} + (1-\rho) \log_{10} \frac{1-\rho}{1-\hat{\rho}_j} \right) \quad (13)$$

式中, s_2 为隐藏层神经元个数。由公式可知, 稀疏惩罚因子基于相对熵而来, 用于表示稀疏性参数 ρ 与平均活跃度 $\hat{\rho}_j$ 之间的差异。当 $\hat{\rho}_j$ 与 ρ 相等时, 其值等于 0, 当 ρ 与 $\hat{\rho}_j$ 之间差异的增大时, 其值随之增大。

因此总体损失函数改为

$$J_{\text{sparse}}(\mathbf{W}, \mathbf{b}) = J(\mathbf{W}, \mathbf{b}) + \beta \sum_{j=1}^{s_2} KL(\rho \| \hat{\rho}_j) \quad (14)$$

式中, 前一项与之前定义相同; 后一项为稀疏惩罚因子项, β 为稀疏惩罚因子的权重系数。

为使得稀疏自编码器损失函数 $J_{\text{sparse}}(\mathbf{W}, \mathbf{b})$ 最小, 利用梯度下降法, 优化权重 $\mathbf{W}$ 和偏置 $\mathbf{b}$, 每次迭代对两个参数做出更新。

$$\begin{cases} \mathbf{W}_{ji}^{(l)} = \mathbf{W}_{ji}^{(l)} - \alpha \frac{\partial}{\partial \mathbf{W}_{ji}^{(l)}} J_{\text{sparse}}(\mathbf{W}, \mathbf{b}) \\ \mathbf{b}_i^{(l)} = \mathbf{b}_i^{(l)} - \frac{\partial}{\partial \mathbf{b}_i^{(l)}} J_{\text{sparse}}(\mathbf{W}, \mathbf{b}) \end{cases} \quad (15)$$

由此可见, 网络训练过程中, 权重更新的关键在于求解损失函数对权重 $\mathbf{W}$ 和偏置 $\mathbf{b}$ 的偏导数, 下面利用反向传播算法求解。

3.2 反向传播算法

由前面可知网络传播过程可知, 第 l 层神经元的第 j 个节点的输入 $z_j^{(l)}$ 与第 $l-1$ 层神经元节点输出 $a_i^{(l-1)}$ 之间的关系。

$$\begin{cases} z_j^{(l)} = \mathbf{W}_{ji}^{(l-1)} a_i^{(l-1)} + b_j^{(l-1)} \\ a_i^{(l)} = f_\theta(z_j^{(l)}) \end{cases} \quad (16)$$

假设给定样本 (x, y) , 对网络各参数随机初始化, 经过前向传播运算后, 得出网络各节点激活值 $f_\theta(z_j^{(l)})$ 。为表示第 l 层神经元的第 j 个节点对输出误差的影响, 定义误差项 $\delta_j^{(l)}$, 对于网络输出层(用 n_l 表示), 该误差项为

$$\begin{aligned} \delta_j^{(n_l)} &= \frac{\partial}{\partial z_j^{(n_l)}} \frac{1}{2} \|y - a^{(n_l)}\|^2 = \frac{\partial}{\partial z_j^{(n_l)}} \frac{1}{2} \sum_{i=1}^{s_l} \|y_i - a_i^{(n_l)}\|^2 = - \\ &= (y_j - a_j^{(n_l)}) \frac{\partial}{\partial z_j^{(n_l)}} a_j^{(n_l)} = -(y_j - a_j^{(n_l)}) f'_\theta(z_j^{(n_l)}) \end{aligned} \quad (17)$$

对于 n_l-1 层, 第 j 个节点的误差项

$$\begin{aligned} \delta_j^{(n_l-1)} &= \frac{\partial}{\partial z_j^{(n_l-1)}} \frac{1}{2} \|y - a^{(n_l)}\|^2 = \\ &= \frac{\partial}{\partial z_j^{(n_l-1)}} \frac{1}{2} \sum_{i=1}^{s_l} \|y_i - a_i^{(n_l)}\|^2 = \\ &= \left(\sum_{i=1}^{s_l} \delta_i^{(n_l)} \mathbf{W}_{ji}^{(n_l-1)} \right) f'_\theta(z_j^{(n_l-1)}) \end{aligned} \quad (18)$$

利用递推关系, 得出任意 l (在 $n_l, n_l-1, \dots, 2$ 取值)层的误差项

$$\delta_j^{(l)} = \left(\sum_{i=1}^{s_{l+1}} \delta_i^{(l+1)} \mathbf{W}_{ji}^{(l)} \right) f'_\theta(z_j^{(l)}) \quad (19)$$

通过上面推导可知, 计算第 l 层网络误差项时, 需要提前知道第 $l+1$ 层的误差项, 即误差从后向前求导, 因此称之为反向传播算法。

稀疏自编码器中损失函数对权重 $\mathbf{W}$ 的偏导数为

$$\frac{\partial}{\partial \mathbf{W}_{ji}^{(l)}} J_{sparse}(\mathbf{W}, \mathbf{b}) = \left[\frac{1}{m} \sum_{i=1}^m \sum_{j=1}^{S_l} \delta_i^{(l+1)} \mathbf{W}_{ji}^{(l)} f'_\theta(z_j^{(l)}) a_j^{(l)} \right] + \lambda \sum_{i=1}^{n_l-1} \sum_{j=1}^{S_l} \sum_{j=1}^{S_{l+1}} (\mathbf{W}_{ji}^{(l)}) + \frac{\beta}{m} \sum_{i=1}^m \sum_{j=1}^{S_l} \left(-\frac{\rho}{\hat{\rho}_i} + \frac{1-\rho}{1-\hat{\rho}_i} \right) f'_\theta(z_j^{(l)}) a_j^{(l)} \quad (20)$$

稀疏自编码器中损失函数对权重 $\mathbf{b}$ 的偏导数

$$\frac{\partial}{\partial b_i^{(l)}} J_{sparse}(\mathbf{W}, \mathbf{b}) = \frac{1}{m} \sum_{i=1}^m \sum_{j=1}^{S_l} \delta_i^{(l+1)} \mathbf{W}_{ji}^{(l)} f'_\theta(z_j^{(l)}) + \frac{\beta}{m} \sum_{i=1}^m \sum_{j=1}^{S_l} \left(-\frac{\rho}{\hat{\rho}_i} + \frac{1-\rho}{1-\hat{\rho}_i} \right) f'_\theta(z_j^{(l)}) \quad (21)$$

对于样本 $\mathbf{x}$ 输入时, 第三层网络中第 i 个神经元的误差项 $\delta_i^{(3)}$ 为

$$\delta_i^{(3)} = -(x_i - \hat{x}_i) f'_\theta(z_i^{(3)}) \quad (22)$$

将式(19)代入式(17)、(18), 便可求出损失函数对隐藏层神经元权重 $\mathbf{W}$ 和偏置 $\mathbf{b}$ 的偏导数, 实现网络参数的更新。

3.3 堆栈稀疏自编码器

由前面介绍可知, 稀疏自编码器中隐藏层是一个稀疏层, 且在经过训练之后, 该层神经元可以很好地恢复输入数据, 因此该层神经元实质是对输入层数据提取的特征数据。当把稀疏自编码器输出层去掉, 仅保留 Encoder 部分, 并将提取的特征数据作为下一个稀疏自编码器的输入, 同样方法连接多个稀疏自编码器, 便形成堆栈稀疏自编码器。图 3 所示为由 3 层 SAE 组成的堆栈稀疏自编码器。

图 3 堆栈稀疏自编码器网络结构

3.4 Softmax 分类器

Softmax 分类器常被连接在深度学习网络的最后一层, 用于解决多分类问题。假设有训练集 $\{(x^{(1)}, y^{(1)}), \dots, (x^{(m)}, y^{(m)})\}$, $x^{(i)}$ 为样本数据, $y^{(i)}$ 为样本标签, 取值范围在 1 到 k 。那么对于给定测试样本 x , 可利用假设函数估计出该样本属于任意第 j 类的概率 $p(y=j|x)$, 即假设函数将输出 k 维数据分别表示该样本属于这 k 类的概率。假设函数表达

式为

$$\mathbf{h}_\theta(x^{(i)}) = \begin{pmatrix} p(y^{(i)}=1|x^{(i)}; \theta) \\ p(y^{(i)}=2|x^{(i)}; \theta) \\ p(y^{(i)}=3|x^{(i)}; \theta) \end{pmatrix} = \frac{1}{\sum_{j=1}^k e^{\theta_j^T x^{(i)}}} \begin{pmatrix} e^{\theta_1^T x^{(i)}} \\ e^{\theta_2^T x^{(i)}} \\ \vdots \\ e^{\theta_k^T x^{(i)}} \end{pmatrix} \quad (23)$$

式中, $\theta_1, \theta_2, \dots, \theta_k \in \mathbb{R}^{n+1}$ 是模型参数。

添加权重衰减项的损失函数为

$$J(\theta) = -\frac{1}{m} \left[\sum_{i=1}^m \sum_{j=1}^k 1\{y^{(i)}=j\} \log_{10} \frac{e^{\theta_j^T x^{(i)}}}{\sum_{l=1}^k e^{\theta_l^T x^{(i)}}} \right] + \frac{\lambda}{2} \sum_{i=1}^m \sum_{j=0}^n \theta_{ij}^2 \quad (24)$$

式中, 当 $y^{(i)}=1$ 为真时, $1\{y^{(i)}=1\}=1$; 否则 $1\{y^{(i)}=1\}=0$ 。

通过最后最优化算法, 寻找 $J(\theta)$ 最小值对应的参数, 便可以得到 Softmax 分类模型。

3.5 有监督微调

网络的有监督微调过程是在确定堆栈稀疏自编码器和 Softmax 分类器的初始权重和偏置之后, 构建整个网络的损失函数, 利用训练样本标签和训练结果对整个网络的参数做出调整的过程。在寻找整个网络损失函数最优解的过程, 使用之前介绍的梯度下降法对网络中各参数更新, 使用反向传播算法求解各层神经元的偏导。具体过程如下。

(1) 将样本 $x^{(i)}$ 输入经过预训练的深度学习网络, 经过数据前向传播, 得到各层神经元的激活值。

(2) 对最后的 Softmax 层输出, 求残差。

$$\delta_i^{(n)} = -\theta^T (I - P) f'_\theta(z_i^{(n)}) \quad (25)$$

式中, P 为网络得到样本属于各类别的概率; I 为样本标签经过变换后得到的矩阵。

(3) 利用式(19)求前 l (在 $n_l, n_l-1, \dots, 2$ 取值) 层的残差。

(4) 利用式(20)、式(21)求出损失函数对各层神经元权重与偏置的偏导。并利用式(15)完成权重更新。

经过微调后的深度学习网络, 各参数得到进一步优化, 整个网络提取特征的能力得到提高, 测试结果也更加精确。

4 深度学习的刀具磨损状态监测方法

为实现数控机床刀具磨损状态监测, 本文提出了一种更加方便有效的监测方法, 该方法诊断流程图, 如图 4 所示。

图 4 刀具磨损诊断流程图

诊断流程如下。

(1) 样本采集和数据预处理。该诊断方法将采集数控机床的三相电流信号；然后将其融合成电流有效值信号，计算信号频谱；因频域数据具有稀疏性，因此利用压缩感知对频域数据压缩；之后，将观测信号 $c(t)$ 随机取一半作为训练样本，另一半作为测试样本，并对训练样本做加噪处理；最后将所有样本归一化到 $[0, 1]$ 范围之内。

(2) 深度学习网络训练。该诊断方法采用两层 SAE 加 Softmax 的深度学习网络结构，训练过程主要分为两部分：网络预训练和有监督微调。网络预训练时需要先确定网络参数：各层神经元个数、训练步数、隐藏层稀疏性参数 ρ 等；然后逐层训练，得到每层的最优权重和偏置作为有监督微调的初始参数值。有监督微调将构建整个网络的损失函数，利用样本标签再次训练网络，对权重和偏置微调，使整个网络达到最优。

(3) 将测试样本输入训练好的网络，得出训练结果。

5 试验验证

5.1 数据采集

为验证所提出的方法可以用于数控机床刀具磨

损程度的状态监测，在台湾东昱精机公司生产的 CMV-850A 型立式加工中心上进行磨损试验，该机床主轴电机型号是由 FANUC 公司生产的 $\beta 12/8000i$ 型三相异步电动机。利用三个瑞士 LEM 公司生产的 LT 108-S7 型闭环霍尔效应电流传感器采集主轴电流信号，通过丹麦 B&K 公司生产的 3053B 型 12 通道数据采集模块及其软件系统采集存储。刀具为台湾 Di 牌 2 刃 40Cr 硬质合金平头立铣刀，其直径为 12 mm，螺旋角为 30° 。试验现场图如图 5 所示。

图 5 试验现场图

对刀具主切削刃加工出 6 种不同程度的磨损状态，其中第 1 种状态为刀具无磨损状态，第 6 种状态为刀具严重磨损状态。在每种磨损状态进行 9 组加工参数的切削试验，切削参数见表 1，共设置三种切削速度 n 、三种进给量 f_z 、三种纵向切深 a_p 、三种径向切深 a_r 。

表 1 加工参数

序号	切削速度 $n/(r/min)$	进给量 f_z/mm	纵向切深 a_p/mm	径向切深 a_r/mm
1	1 200	0.15	2.0	1.50
2	1 200	0.15	1.5	1.50
3	900	0.15	1.0	1.50
4	1 200	0.15	1.0	1.50
5	1 500	0.15	1.0	1.50
6	1 200	0.10	1.0	1.50
7	1 200	0.20	1.0	1.50
8	1 200	0.15	1.0	1.75
9	1 200	0.15	1.0	2.00

试验所用采样频率为 4 kHz，共得到 54 组主轴电流数据，每组数据包含电机三相电流信号。磨损状态 2 下的铣刀在工况 1 下进行切削时，采集到单相电流的时域和频域信号如图 6 所示。从图中可以

图 6 主轴电机单相电流信号及其频谱

看出单相电流信号时域图比较光滑, 十分接近正弦信号, 在工况 1 下, 最大幅值处正好对应转频 2 倍, 即转频乘铣刀刃数。

5.2 参数确定

对数控机床电机的三相电流信号按照式(2)融合成电流有效值信号, 并按照磨损状态的不同, 分为 6 类, 以每 4 096 点为一个样本, 共得到 656 组原始数据样本。计算每个样本的频谱, 并选取前一半的有效频段, 图 7 给出了工况 1 下的 6 种不同磨损程度的电流有效值信号及其频谱。

图 7 6 种磨损情况下的原始数据及其频谱

从图 7 中可以看出主轴电流信号频域数据具有稀疏性质, 因此可直接利用压缩感知压缩数据, 根据信号重建误差趋势图(图 8), 取测量长度 $M=500$ 点。

图 8 不同测量长度下的测量误差平方和

所有样本经过压缩感知之后, 每个样本长度压缩至 500 点, 随机选取 328 组作为训练样本, 另外 328 组作为测试样本, 对训练样本添加 4 种不同噪音, 加噪强度系数 $k=0.1$, 样本个数扩展到 1 640 组。

同时, 经过反复试验确定稀疏自编码中的稀疏参数 $\rho=0.3$, 稀疏惩罚因子的权重系数 $\beta=2$ 。

5.3 试验结果分析

为验证本文提出方法具备在电流信号中提取刀具磨损特征信息的优异能力, 将用本文提出的方法与人工提取特征方法、SSAE、文献[14]提出的方法进行对比分析, 诊断结果如表 2 所示。其中计算机配置为 Intel(R) Core(TM) i5-4460 处理器, 16 GB 内存, 2T 机械硬盘, 处理软件 Matlab R2016a。

与人工提取特征方法对比。人工提取特征方法共提取 25 个特征, 其中 5 个时域特征: 均值、方差、峰峰值、均方根值和峭度值, 取 4 个频带内的 3 个特征: 平均能量、频域方差和频域均方根值, 同时通过 3 层小波包分解得到 8 组小波包能量, 将这 25 组特征输入支持向量机(Support vector machine, SVM), 进行分类识别。从诊断结果可以看出, 人工提取特征的方法和本文提出的方法对刀具磨损的第 6 种磨损状态都有很好的识别效果, 这是因为刀具严重磨损时, 主轴电机电流信号发生明显的变化, 这种变化可以通过人工提取信号特征的方式表现出来, 因此利用 SVM 识别精度可以达到 100%。但当磨损区别不大时, 如磨损状态 1、2、3 和 4, 主轴电机电流信号仅发生微弱变化, 同时由于加工参数的变化以及噪声信号的存在, 无法通过传统特征提取的方法区分出这种微弱变化, 因此利用 SVM 识别效果较差, 而利用本文提出的方法识别效果最差的第 4 种磨损状态识别准确率依旧达到 95.1%, 说明该方法能够适应加工参数的变化, 充分提取表征信号特点的特征信息。从总体来看, 人工提取特征的方法总的识别准确率仅有 81.7%, 却用时 294 s; 而本文提出的方法在将识别准确率提高至 98.5%的同时, 训练用时减少到 130 s, 充分体现了本文提出方法的优势。

与堆栈稀疏自编码器对比。利用电流有效值的时域和频域信号作为堆栈稀疏自编码器的输入, 进行刀具磨损状态识别, 其中时域信号取三种不同的样本点数: 4 096、1 024、512, 并做傅里叶变换形成三种不同点数的频谱, 所用堆栈稀疏自编码器的参数与之前介绍的相同, 各层网络神经元个数, 在表 2 中给出。从识别结果可以看出, 直接用时域信号作为堆栈稀疏自编码器的输入时, 平均识别准确率低于 70%。第 6 种刀具磨损状态的识别效果明显, 识别准确率高达 95%。而其他几种刀具磨损区别不大的状态识别准确率较低, 这是由于时域信号变化微弱, 加上随机噪声的存在, 使得堆栈稀疏自编码器无法有效的对时域信号提取特征。直接利用频域信号作为堆栈稀疏自编码器的输入时, 相对时域输入,

识别精度有所提高,但由于频域内存能力,使得仅到达 84.8%,且准确率随做傅里叶变换样本点数的减小而降低,而点数增加又会增加训练时间,无法像

本文提出的方法那样兼顾训练精度与训练效率。而从单一磨损程度来看,对磨损比较严重的第 6 种刀具磨损状态的识别准确率同样达到 99.7%~100%。

表 2 诊断精度与用时

序号	方法	磨损 1(%)	磨损 2(%)	磨损 3(%)	磨损 4(%)	磨损 5(%)	磨损 6(%)	平均	用时/s
1	人工提取特征方法+SVM	71.9	61.5	74.5	81.6	98.3	100	81.7	294
2	时域(4 096+SSAE hid1=500, hid2=200	55	82.5	54.6	47.3	83.3	98.1	67.0	493
3	时域(1 024)+SSAE hid1=500, hid2=200	57	66.2	52.4	54.8	82.9	96.0	68.3	153
4	时域(512)+SSAE hid1=200, hid2=100	49	62.6	46.1	49.2	87.0	94.2	65.4	40
5	频域(4 096)+SSAE hid1=500, hid2=200	79	80.4	59.6	85.9	100	100	84.8	245
6	频域(1 024)+SSAE hid1=200, hid2=100	75.5	80.3	64.4	81.9	97.9	100	83.6	116
7	频域(512)+SSAE hid1=200, hid2=100	59.2	59.7	49.9	75.5	89.4	99.7	72.6	24
8	文献[13]方法 hid1=600, hid2=1 000	50.2	44.3	32.7	35.4	62.2	98.8	61.5	170
9	本文方法 hid1=200, hid2=100	98.0	100	98.2	95.1	100	100	98.5	130

本文与文献[13]的不同之处在于,文献[13]利用 CS 压缩时域信号,形成完备特征,然后利用堆栈稀疏自编码器逐层扩展提取特征的个数,然后用于状态识别。大量试验表明,长度为 1 024 点的时域数据用 CS 压缩到 400 点后,可以很好地恢复出原信号。从试验结果可以看出,这种方法对时域信号区别不明显的第 1、2、3 和 4 种磨损的识别准确率较低,对时域信号区别明显的第 6 种刀具磨损状态识别准确率达到 98.8%。这同样是因为加工过程中的随机噪声淹没了信号中的微弱差异,降低了堆栈稀疏自编码器提取特征的能力,而本文提出的方法将时域信号转换到频域,能消除时域内随机噪声的干扰,可有效提取信号中微弱差异的特征信息。

为进一步验证本文提出方法的特征提取能力,利用主成分分析(Principal component analysis, PCA)对人工方法、直接使用 SSAE 和“CS+SSAE”提取得到的特征进一步降维,取前三个主元成分,进行可视化,如图 9 所示。从图中可以明显看出利用人工提取信号特征的方法和直接使用频域信息输入堆栈稀疏自编码器提取特征的方法,很难将 6 种磨损程度区分开,尤其是对前四种磨损状态,提取得到的特征重叠交叉到一起,无法识别。而利用本文提出方法提取得到的特征,将每类磨损状态都聚集到了一起,不同磨损状态之间的边界分明,充分体现了该方法对信号的深度表征能力。

(a) 人工提取特征

(b) 频域+SSAE

(c) 本文提出方法

图 9 特征提取散点图

6 结论

(1) 针对数控机床加工过程中刀具磨损程度识别问题, 本文提出了利用融合后的主轴电机三相电流信号作为刀具磨损的诊断信号, 使用压缩感知加稀疏自编码器的方法提取频域内观测信号的特征信息, 从而有效的实现对刀具磨损程度的识别。

(2) 本文提出的方法相对于传统特征提取方法更加智能, 可适用于不同的加工参数, 避免了人为提取特征时加工参数的影响, 同时避免了必须利用先验知识才能对故障特征进行识别的弊端。

(3) 本文提出的方法相对直接使用堆栈稀疏自编码器的方法, 可以更适当的表征信号中隐藏的特征信息, 更加准确的识别出不同类别信号之间的差异特征。

参 考 文 献

- [1] 任浩, 屈剑锋, 柴毅, 等. 深度学习在故障诊断领域中的研究现状与挑战[J]. 控制与决策, 2017, 32(8): 1345-1358.
REN Hao, QU Jianfeng, et al. Deep learning for fault diagnosis: The state of the art and challenge[J]. Control and Decision, 2017, 32(8): 1345-1358.
- [2] Sandvik Coromant (Firm). Modern metal cutting: A practical handbook[M]. Sandviken: Sandvik Coromant, 1994.
- [3] ALTINTAS Y. Prediction of cutting forces and tool breakage in milling from feed drive current measurements[J]. Asme Journal of Engineering for Industry, 1992, 114(4): 386-392.
- [4] 李宏坤, 汪寅虎, 杨树华, 等. 基于声发射信号的铣刀运行可靠性评估[J]. 大连理工大学学报, 2014(4): 418-423.
LI Hongkun, WANG Yinhu, YANG Shuhua, et al. Estimation of operational reliability for milling cutter based in acoustic emission signal[J]. Journal of Dalian University of Technology, 2014(4): 418-423.
- [5] STEIN J L, COLVIN D, CLEVER G, et al. Evaluation of DC servo machine tool feed drives as force sensors[J]. Transactions of Asme Journal of Dynamic Systems Measurement & Control, 1986, 108(4): 279-288.
- [6] ALTINTAS Y, YELLOWLEY I. In-Process detection of tool failure in milling using cutting force models[J]. Journal of Engineering for Industry, 1989, 111(2): 149-157.
- [7] JEONG Y H, CHO D W. Estimating cutting force from rotating and stationary feed motor currents on a milling machine[J]. International Journal of Machine Tools & Manufacture, 2002, 42(14): 1559-1566.
- [8] LI X, TSO S K, WANG J. Real-time tool condition monitoring using wavelet transforms and fuzzy techniques[J]. IEEE Transactions on Systems Man & Cybernetics Part C, 2000, 30(3): 352-357.
- [9] 雷亚国, 贾峰, 周昕, 等. 基于深度学习理论的机械装备大数据健康监测方法[J]. 机械工程学报, 2015, 51(21): 49-56.
LEI Yaguo, JIA Feng, ZHOU Xin, et al. A deep learning-based method for machinery health monitoring with big data[J]. Journal of Mechanical Engineering, 2015, 51(21): 49-56.
- [10] 张存吉, 姚锡凡, 张剑铭, 等. 基于深度学习的刀具磨损监测方法[J]. 计算机集成制造系统, 2017, 23(10): 2146-2155.
ZHANG Cunji, YAO Xifan, ZHANG Jianming, et al. Tool wear monitoring based on deep learning[J]. Computer Integrated Manufacturing Systems, 2017, 23(10): 2146-2155.
- [11] 林杨. 基于深度学习的刀具磨损状态监测技术的研究[D]. 合肥: 中国科学技术大学, 2017.
LIN Yang. Research on tool wear state monitoring technology of high speed milling based on deep learning[D]. Hefei: Hefei University of Science and Technology of China, 2017.
- [12] 郭俊峰, 石斌, 雷春丽, 等. 基于双稀疏字典模型机械振动信号压缩感知方法[J]. 机械工程学报, 2018, 54(6): 118-127.
GUO Junfeng, SHI Bin, LEI Chunli, et al. Method of compressed sensing for mechanical vibration signals based on double sparse dictionary model[J]. Journal of Mechanical Engineering, 2018, 54(6): 118-127.
- [13] AHMED H O A, WONG M L D, NANDI A K. Intelligent condition monitoring method for bearing faults from highly compressed measurements using sparse over-complete features[J]. Mechanical Systems & Signal Processing, 2018, 99: 459-477.
- [14] 石光明, 刘丹华, 高大化, 等. 压缩感知理论及其研究进展[J]. 电子学报, 2009, 37(5): 1070-1081.
SHI Guangming, LIU Danhua, GAO Dahua, et al. Advances in theory and application of compressed sensing[J]. ACTA Electronica Sinica, 2009, 37(5): 1070-1081.

- [15] CHEN R, CHEN S, HE M, et al. Rolling bearing fault severity identification using deep sparse auto-encoder network with noise added sample expansion[J]. Journal of Risk & Reliability, 2017, 231(6): 1748006X-1772645X.
- [16] WANG Fengtao, DENG Gang, LIU Chenxi, et al. A deep feature extraction method for bearing fault diagnosis

based on empirical mode decomposition and kernel function [J]. Advances in Mechanical Engineering, 2018, 10(9): 1-12.

作者简介: 李宏坤, 男, 1974 年出生, 教授, 博士研究生导师。主要研究方向为机械设备动态分析与故障诊断, 信号处理。
E-mail: lihk@dlut.edu.cn