

# 基于声学特征的微电机故障诊断方法

刘艳杰,陈炳发,丁力平

(南京航空航天大学 机电学院,江苏 南京 210016)

**摘要:**针对微电机质量检测水平低、故障识别困难等问题,设计一种基于声学特征的微电机故障诊断方法。通过声音采集装置获得微电机转动时的正常声音信号和三种故障信号;从声音信号中提取 39 维梅尔频率倒谱系数和短时能量,搭建一维卷积神经网络模型进行识别。将声音信号转化成语谱图,建立二维卷积神经网络模型并识别。利用多模型融合技术中的加权平均算法将两个模型融合,融合后模型的准确率为 93.58%,比单个模型平均提高 2.43%。

**关键词:**声学特征;模型融合;卷积神经网络;微电机;故障诊断

**中图分类号:**TP391 **文献标志码:**B **文章编号:**1671-5276(2022)02-0190-05

## A Fault Diagnosis Method of Micro-motor Based on Acoustic Feature

LIU Yanjie, CHEN Bingfa, DING Liping

(College of Mechanical and Electrical Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

**Abstract:** A fault diagnosis method based on acoustic characteristics is designed to solve the problems of low quality detection and difficulty in fault identification of needle micro-motor. The normal sound signal and three kinds of fault signals are obtained by sound acquisition device. 39 dimensions Mel frequency cepstrum coefficient and short-time energy are extracted from sound signals to built one dimensional convolutional neural network for identification, and the sound signals are transformed into speech spectrum diagram to establish two dimensional convolutional neural network model for identification. The weighted average algorithm in the multi-model fusion technology is applied to fuse the two models with the accuracy of the fused models up to 93.58%, which is 2.43% higher than the single model on average.

**Keywords:** acoustic feature; model fusion; convolutional neural network; micro-motor; fault diagnosis

## 0 引言

微电机,全称是微型电动机,是指直径 $<160\text{ mm}$ 或者额定功率 $<750\text{ mW}$ 的电机,经常被用于控制系统中实现机电信号的解析运算和转换等功能。微电机装配过程已经实现了自动化,但是在最终质量检测环节中,却要依赖有经验的工人听微电机的转动声音来判断是否有噪声故障。虽然人耳是个极为精密的器官,但人耳对微电机装配质量的判定受诸多因素的影响,不同人对同一微电机、甚至同一个人对同一台微电机的多次判定结果均可能不一致,因此识别率难以得到保证。随着数字信号处理<sup>[1]</sup>和机器学习<sup>[2]</sup>等技术的发展,将机器学习技术用于声音信号处理,以实现生产线中微电机装配质量的智能化识别,有利于提高装配质量的检测精度和效率。

基于声学特征的故障诊断主要包含特征提取和故障分类两个部分。电机在运行的过程中,会由自身的振动产生噪声,噪声信号中包含着电机的状态信息。当发生故障时,其噪声的振幅、频域都会发生变化<sup>[3]</sup>。声音信号的特征参数一般可以分为时域特征、频域特征和倒谱域特征<sup>[4]</sup>。信号处理的方法一般为时频分析方法,主要有小波变换<sup>[5]</sup>和经验模式分解<sup>[6]</sup>。胡文明等<sup>[7]</sup>对微电机声音

信号进行小波包能量谱分析,将信号分成不同频段,通过每个频段内能量分布的不同来识别故障类型。HUANG N E 等<sup>[8]</sup>提出了经验模态分解方法(EMD),这种分解方法是自适应的,适用于非线性和非平稳过程。

梅尔频率倒谱系数(mel frequency cepstrum coefficient, MFCC)<sup>[9]</sup>将人耳的听觉感知特性和语音的产生机制相结合,广泛适用于音频信号处理。语谱图是语音频谱图,它表达了语音信号不同频段的信号强度,在语音情感识别中广泛应用此特征<sup>[10]</sup>。卷积神经网络(CNN)<sup>[11]</sup>是一种深度监督学习下的机器学习模型,具有很强的适应性,便于挖掘数据的局部特征,在语音识别、图像处理<sup>[12]</sup>等领域都取得了很好的成果。

本文提出了一种基于声学特征的微电机故障诊断方法,以融合 CNN 模型作为声音识别模型。首先利用声音采集装置获得微电机转动声音信号,然后提取信号的 MFCC、短时能量和语谱图等声学特征,最后将提取的特征数据输入到模型中进行故障识别。本文对该方法的设计与实现的过程进行阐述。

## 1 微电机故障诊断的整体流程

本文提出的微电机故障诊断方法的整体流程如图 1

所示,主要包含以下 5 步。

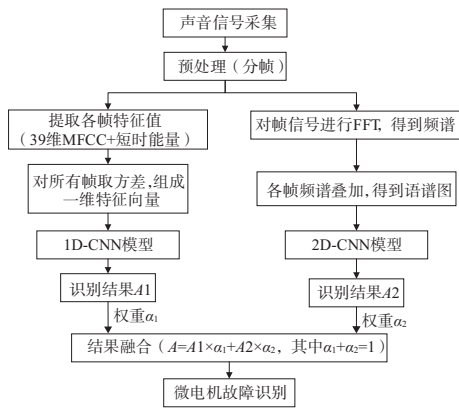


图 1 微电机故障诊断的整体流程

1) 声音信号采集: 使用声音传感器采集微电机声音信号, 并进行预处理;

2) 特征提取: 提取声音信号的 39 维 MFCC、短时能量和语谱图特征, 用于后面的分析识别;

3) 模型识别: 将 MFCC 和短时能量输入 1D-CNN(一维卷积神经网络)模型, 语谱图输入 2D-CNN(二维卷积神经网络)模型中, 分别得到各自识别的结果;

4) 模型融合: 给不同模型识别的结果赋予不同的权重, 以提高识别的准确率;

5) 故障识别: 输出微电机的故障类型。

## 2 微电机声音信号采集及特征提取

### 2.1 声音信号采集

本文研究的微电机型号是 24BYJ46, 微电机声音信号使用隔音装置采集, 没有环境噪声。首先将微电机接通电源, 将声音传感器靠近微电机进行采集, 采样频率是 16 kHz, 采样位数是 16 位, 采样点是 30 000, 信号的保存格式是无损的 WAV 格式。采集的声音信号有 4 种, 正常微电机转动声音、齿轮啮合不良故障、擦铁异音故障、内有垃圾故障。每种信号有 480 个样本, 总共 1 920 个样本。4 种微电机声音信号的时域波形图如图 2 所示。

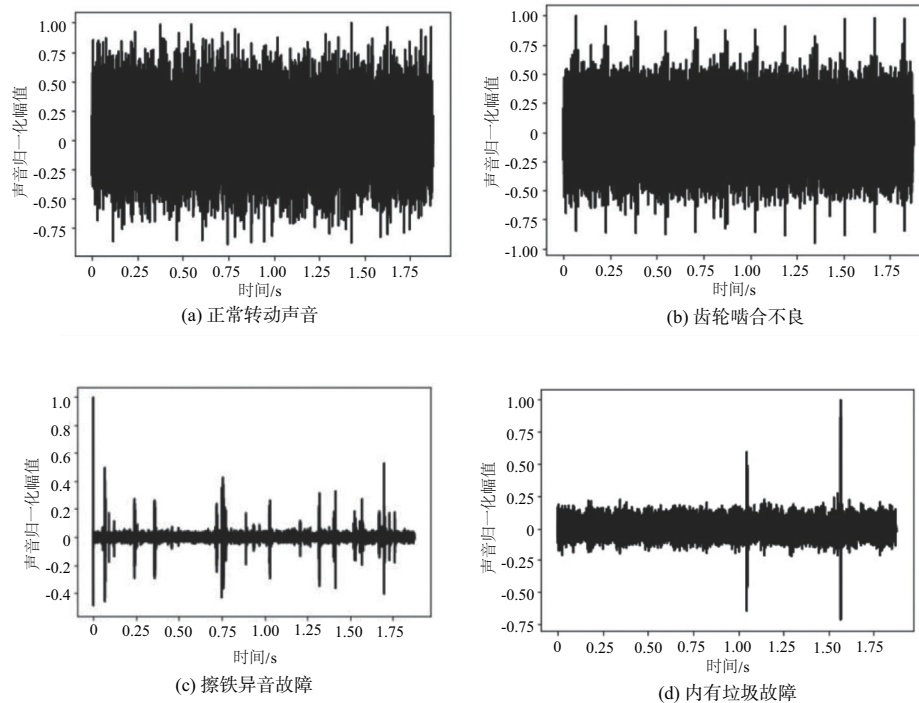


图 2 4 种微电机声音信号的时域波形图

### 2.2 特征提取

#### a) 梅尔频率倒谱系数

梅尔频率倒谱系数是基于人耳的听觉特性提出来的, 它与频率成非线性对应关系。MFCC 是一种常用的音频特征, 广泛应用于自动语音和说话人识别中。在实际应用中, MFCC 特征提取的过程<sup>[13]</sup>如图 3 所示, 具体计算过程如下。

1) 预加重。将采集的声音信号  $s(n)$  通过一个高通滤波器, 以增强声音信号中的高频部分。高通滤波器的传递

函数为

$$X(n) - s(n) - a \times s(n-1) \quad (1)$$

式中  $a$  为预加重系数, 值介于 0.9~1.0 之间。

2) 分帧。声音信号具有短时平稳性, 分帧的过程就是将一段声音信号分成  $N$  段, 每一段为一帧, 通常情况下每帧涵盖的时间约为 20~30 ms, 为了避免相邻两帧之间变化过大, 因此会让两帧之间有一段重叠的区域, 称为帧移, 一般为帧长的一半左右。

3) 加窗。加窗可以增加帧两端的连续性, 减少频谱泄漏。常用的窗函数是汉明窗 (Hamming), 表达式如下为

$$w(n) = \begin{cases} 0.54 - 0.46 \cos\left(\frac{2\pi n}{N-1}\right), & 0 \leq n \leq N-1 \\ 0, & \text{其他} \end{cases} \quad (2)$$

式中  $N$  为窗长,即帧长。

4) DFT。将声音信号进行离散傅里叶变换后得到频谱,将频谱取模平方得到信号的能量谱。声音信号的频谱为

$$X(k) = \text{DFT}(x'(n)) = \sum_{n=0}^{N-1} x'(n) e^{-j\frac{2\pi nk}{N}}, 0 \leq k \leq N \quad (3)$$

5) Mel 滤波器组。求出能量谱后,用  $M$  个 Mel 带通滤波器组进行滤波,将每个滤波频带内的能量进行叠加,这时第  $k$  个滤波器输出功率谱  $x'(k)$ 。

6) 对数运算。对  $m$  个滤波器组的输出取对数运算,得到相应频带的对数功率谱。计算公式为

$$s(m) = \ln\left(\sum_{k=0}^{N-1} |X(k)|^2 H_m(k)\right), 0 \leq m \leq M \quad (4)$$

式中:  $s(m)$  是对数功率谱;  $H_m(k)$  是滤波器组。

7) DCT。将对数功率谱进行 DCT(离散余弦变换),得到  $L$  个 MFCC 系数,计算公式为

$$c(n) = \sum_{m=0}^{M-1} s(m) \cos\left[\frac{\pi n(m-0.5)}{M}\right], n=1,2,\dots,L \quad (5)$$

式中  $c(n)$  是第  $n$  个 MFCC 系数。

8) 动态差分参数的提取。上面直接得到的 MFCC 特征只反映了信号的静态特征,对这种静态特征做一阶、二阶差分,可得到相应的动态特征。差分参数的计算公式为

$$d(t) = \begin{cases} c(t-1) - c(t), & t < K \\ \sum_{k=1}^K k [c(t+k) - c(t-k)], & \text{其他} \\ c(t) - c(t-1), & t \geq Q-K \end{cases} \quad (6)$$

式中:  $d(t)$  表示第  $t$  个一阶差分;  $c(t)$  表示第  $t$  个倒谱系数,  $Q$  表示倒谱系数的阶数;  $K$  表示一阶导数的时间差,可取 1 或 2。将上式结果再代入就能得到二阶差分的参数。

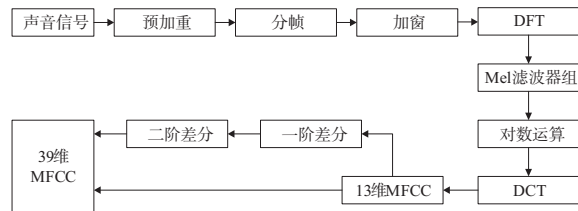


图3 MFCC 的提取过程

b) 短时能量

声音信号的能量随时间变化比较明显,其短时能量可以分析这些幅度的变化。对于信号  $x(n)$ , 短时能量的定义如下:

$$E_n = \sum_{m=-\infty}^{\infty} [s(m)w(n-m)]^2 = s^2(n) \times h(n) \quad (7)$$

式中  $h(n) = w^2(n)$ 。

4 种微电机声音信号的短时能量如图 4 所示。

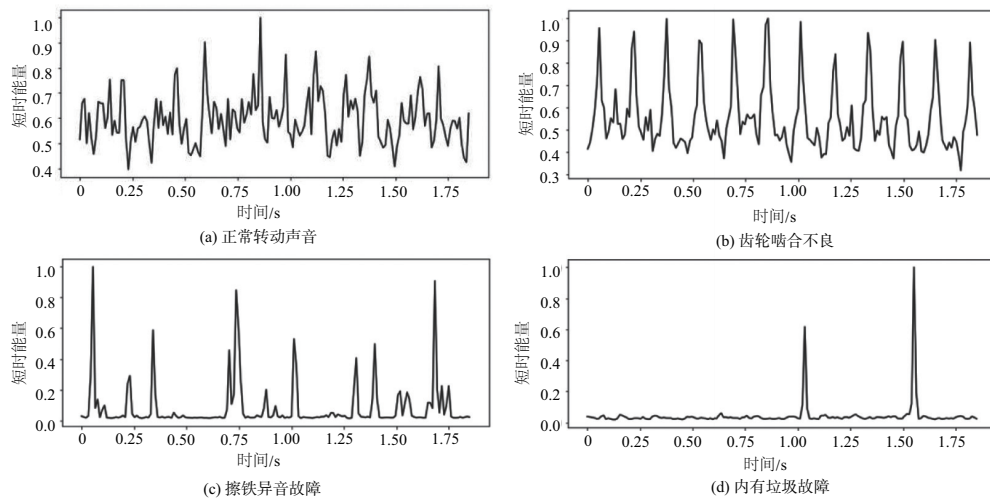


图4 4种微电机信号的短时能量图

c) 语谱图

语谱图是语音频谱图,它的横坐标是时间,纵坐标是频率,坐标点值为语音数据能量。由于语谱图采用二维平面表达三维信息,所以能量值的大小是通过颜色来表示的,颜色越深表示该点的语音能量越强。语谱图表达了语音信号不同频段的信号强度,便于观察频率随时间的变化情况。通过语音信号获得语谱图的过程如下:

- 1) 获取音频数据;
- 2) 将声音信号分帧加窗;
- 3) 对帧信号进行 FFT,得到帧信号的频谱;
- 4) 将各帧频谱叠加,得到整个声音信号的语谱图。

4 种微电机声音信号的语谱图如图 5 所示。

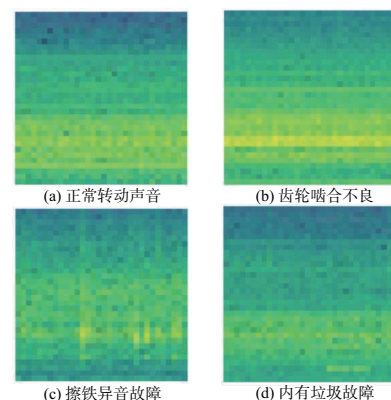


图5 4种微电机信号的语谱图

### 3 微电机故障声音信号识别及结果分析

#### 3.1 融合 CNN 模型

本文训练了两种模型:1D-CNN 和 2D-CNN,同时对

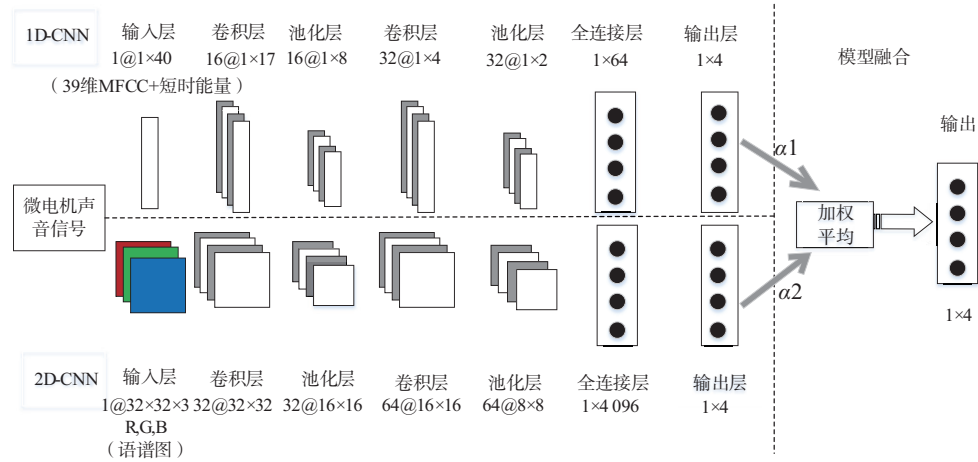


图 6 1D-CNN 与 2D-CNN 模型融合过程

1D-CNN 的网络结构包括输入层、2 个卷积层、2 个池化层、全连接层和输出层。该网络的输入是大小为 40(39 维 MFCC 和短时能量)的特征向量。第一个卷积层卷积核的数目为 16,卷积核大小为 8,步长为 2,卷积后使用 relu 激活函数引入非线性因素,卷积后得到的特征向量用 0 填充边缘部分;第一个卷积层后连接第一个池化层对卷积后的特征向量进行压缩,以简化网络计算复杂度,采用大小为 2 的卷积核进行最大池化,得到 16 个大小为 1x8 的特征向量;第二个卷积层卷积核的数目为 32,卷积核大小为 8,步长为 2,激活函数是 relu,边缘部分用 0 填充;第二个池化层采用大小为 2 的卷积核进行最大池化,得到 32 个大小为 1x2 的特征向量;第二个池化层后是全连接层,用来连接所有的特征,将输出值送给 softmax 分类器,经全连接层后得到的 64 维特征向量;最后一个是输出层,大小为 4。

2D-CNN 的网络结构包括输入层、2 个卷积层、2 个池化层、全连接层和输出层。该网络的输入是大小为 32 像素x32 像素x3 通道(RGB)的语谱图。第一个卷积层卷积核的数目为 32,卷积核大小为 3x3,卷积后使用 relu 激活函数引入非线性因素,卷积后得到的特征图用 0 填充边缘部分,使其大小保持不变;第一个卷积层后连接第一个池化层对卷积后的特征图进行压缩,以简化网络计算复杂度,采用大小为 2x2 的卷积核进行最大池化,得到 32 个大小为 16x16 的特征图;第二个卷积层卷积核的数目为 64,卷积核大小为 3x3,激活函数是 relu,边缘部分用 0 填充;第二个池化层采用大小为 2x2 的卷积核进行最大池化,得到 64 个大小为 8x8 的特征图;第二个池化层后是全连接层,用来连接所有的特征,将输出值送给 softmax 分类器,经全连接层后得到的 4 096 维特征向量;最后一个是输出层大小为 4。

本文采用加权平均法对 2 个模型的输出结果进行融

合,加权平均的公式如下:

$$H(x) = \frac{1}{T} \sum_{i=1}^T \alpha_i \cdot h_i(x) \quad (8)$$

式中  $\alpha_i$  表示权重,  $\alpha_i \geq 0$ ;  $\sum_{i=1}^T \alpha_i = 1$ 。

#### 3.2 试验验证与结果分析

##### 1) 试验设计

从样本中随机选择 70% 的样本作为卷积神经网络的输入训练样本,剩下 30% 的样本作为测试样本。将训练样本提取特征值并分别输入到一维 CNN 和二维 CNN 模型中,获得训练模型;再将测试样本提取特征值并输入到训练好的一维 CNN 和二维 CNN 模型中并输出,然后将 2 个模型的结果融合后再输出,对比 3 种模型的输出结果。

##### 2) 状态编码

微电机状态对应编码表如表 1 所示。

表 1 微电机状态对应编码表

类别号	微电机状态	输出向量
1	正常	[1 0 0 0]
2	齿轮啮合不良	[0 1 0 0]
3	擦铁异音	[0 0 1 0]
4	内有垃圾	[0 0 0 1]

编码后的向量便是 CNN 的输出向量。

##### 3) 结果分析

本文分析了 3 个模型:1D-CNN、2D-CNN 和融合 CNN。在融合 CNN 模型中,1D-CNN 的权重为  $\alpha_1$ ,2D-CNN 的权重为  $\alpha_2$ ,且  $\alpha_2 = 1 - \alpha_1$ 。融合 CNN 模型中每个模型的权重和识别准确率的变化如表 2 所示。当  $\alpha_1 = 0$  时,2D-CNN 模型的准确率为 91.67%;当  $\alpha_1 = 1$  时,1D-CNN 模型的准确率为 90.63%。当  $\alpha_1 = 0.4$  并  $\alpha_2 = 0.6$  时,准确

率最高,为93.58%。故本文融合 CNN 模型中,取1D-CNN 的权重为0.4,2D-CNN 的权重为0.6。不同深度学习模型的故障识别准确率对比如表3所示。

表2 融合模型识别准确率

权重		准确率
$\alpha_1$	$\alpha_2$	
0	1	91.67
0.1	0.9	91.84
0.2	0.8	92.53
0.3	0.7	92.71
0.4	0.6	93.58
0.5	0.5	92.01
0.6	0.4	91.49
0.7	0.3	91.32
0.8	0.2	91.15
0.9	0.1	90.97
1	0	90.63

表3 不同深度学习模型的故障识别准确率对比

方法	特征	准确率/%
1D-CNN	MFCC+短时能量	90.63
2D-CNN	语谱图	91.67
融合 CNN	MFCC+短时能量+语谱图	93.58

由表3可知,1D-CNN 模型主要分析特征为时域特征和倒谱域特征,准确率为90.63%;2D-CNN 模型主要分析特征为频域特征,准确率为91.67%,融合 CNN 模型集合了两个模型的优势,识别准确率为93.58%,比1D-CNN 和2D-CNN 的准确率各提高了2.95%和1.91%,平均提高了2.43%。

## 4 结语

本文提出了一种基于声学特征和融合 CNN 模型的微电机故障诊断方法。该方法提取了微电机声音信号的 MFCC 特征、短时能量和语谱图特征,并建立了一维 CNN

和二维 CNN 分类模型来进行识别,然后利用模型融合技术将两个模型训练的结果进行加权平均,充分发挥了不同模型的优势,提高了准确率。该方法具有很好的准确度,在实际的工程中有良好的应用前景。

### 参考文献:

- [1] 吴伟. 相位噪声测试系统中信号处理关键算法研究及实现[D]. 成都:电子科技大学,2020.
- [2] 沈飞,陈超,严如强. 奇异值分解与迁移学习在电机故障诊断中的应用[J]. 振动工程学报,2017,30(1):118-126.
- [3] 刘旭南,赵丽娟,付东波,等. 采煤机截割部传动系统故障信号小波包分解方法研究[J]. 振动与冲击,2019,38(14):169-175,253.
- [4] 胡耀文. 音频信号特征提取及其分类研究[D]. 昆明:昆明理工大学,2018.
- [5] 陈东华,张凤雏. 基于小波变换的同步交流发电机旋转整流器故障诊断[J]. 机械制造与自动化,2018,47(6):224-227.
- [6] 卓仁雄,肖金凤. 基于改进的集合经验模态分解的电动机滚动轴承故障诊断研究[J]. 机械制造与自动化,2019,48(1):36-39.
- [7] 胡文明,陈胜义. 基于小波包能量谱的微电机故障检测方法研究[J]. 时代农机,2017,44(8):107-108,110.
- [8] HUANG N E, SHEN Z, LONG S R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis [J]. Proceedings of the Royal Society of London Series A: Mathematical, Physical and Engineering Sciences, 1998, 454(1971):903-995.
- [9] 王彪. 一种改进的 MFCC 参数提取方法[J]. 计算机与数字工程, 2012, 40(4):19-21.
- [10] 陶华伟,查诚,梁瑞宇,等. 面向语音情感识别的语谱图特征提取算法[J]. 东南大学学报(自然科学版),2015,45(5):817-821.
- [11] 李时奇,赵东标,申珂楠,等. 基于 CNN 特征提取及模型融合的飞机液压系统故障诊断[J]. 机械制造与自动化,2020,49(5):192-195,199.
- [12] HAO X, ZHANG G G, MA S. Deep learning[J]. International Journal of Semantic Computing, 2016, 10(3):417-439.
- [13] 孙汉文,李喆,盛戈峰,等. 基于机器学习与卷积神经网络的放电声音识别研究[J]. 高压电器,2020,56(9):107-113.

收稿日期:2021-01-26

(上接第189页)

## 4 结语

本文以一种常规固定翼无人机为研究对象,采用传统控制策略对无人机自主着陆的短距问题进行了详细分析。通过选择合适的连续襟翼控制策略的基准构型提高了无人机减速能力,并设计了连续襟翼控制策略以直接升力控制,提高了无人机在低速下的抗风能力以及轨迹控制能力,形成了以抵御阵风作用下襟翼下沉率的 PI 控制结构;升降舵以高度控制为外回路,俯仰角回路为内回路的控制结构。

通过仿真试验证明本文设计的控制策略的有效性,提高了无人机风干扰下的着陆精度以及短距性能。

### 参考文献:

- [1] 朱雯雯. 轮式起降无人机全包线控制律设计技术研究[D].

南京:南京航空航天大学,2012.

- [2] 宋辉. 复杂条件下无人机自动着陆控制技术[D]. 南京:南京航空航天大学,2011.
- [3] RUDOLPH P K C. High-lift systems on commercial subsonic airliners[J]. NASA Contractor Report 4746, 1996(9):1-154.
- [4] BRINE G. Flap system for short takeoff and landing aircraft: USA, US4784355 A[P]. 1988-11-15.
- [5] 吴森堂,费玉华. 飞行控制系统[M]. 北京:北京航空航天大学出版社,2005.
- [6] 郭艳艳,陈澜,杨常伟. 无人机着陆抗风鲁棒  $H_\infty$  控制器设计研究[J]. 计算机测量与控制,2010,18(2):348-350.
- [7] 章卫国,王新民,刘长林. 舰载飞机纵向自动着陆控制系统研究[J]. 西北工业大学学报,1996,14(4):549-552.
- [8] 贾家宁. 大载荷无人机自主着陆纵向控制策略研究[D]. 南京:南京航空航天大学,2019.

收稿日期:2020-12-29