

基于 LSTM 的辅助动力装置系统辨识与仿真

杜维仲, 王 硕

(深圳航空有限责任公司 沈阳分公司维修工程部, 沈阳 110000)

摘要: 针对民航飞机辅助动力装置的复杂非线性时序特性, 依据长短时记忆网络 (LSTM) 的长时序记忆、非线性关系表达能力, 提出一种基于 LSTM 的辅助动力装置系统参数辨识模型; 同时建模过程中考虑作用于辨识参数的多种影响因素, 进一步提出基于 LSTM 的多变量时间序列预测参数辨识模型, 解决了传统时序模型难以解决的多变量或者多输入问题; 最后利用辅助动力装置试车样本, 建立了 APU 启动阶段的参数辨识模型; 仿真结果表明, 发动机排气温度 EGT 温度均方根误差小于 4℃, 发动机转速百分比 N 均方根误差小于 1, 满足辅助动力装置仿真需求。

关键词: LSTM; 系统辨识; 非线性系统; APU; 深度学习

System Identification and Simulation of Auxiliary Power Plant Based on LSTM

Du Weizhong, Wang Shuo

(Maintenance Engineering Department, Shenyang Branch of Shenzhen Airlines, Shenyang 110000, China)

Abstract: Aiming at the complex nonlinear timing characteristics of civil aircraft auxiliary power plant, based on the long-term memory and nonlinear relationship expression of long and short time memory network (LSTM), a parameter identification model of auxiliary power plant system based on LSTM is proposed. At the same time, various influence factors of the identification parameters are considered in the modeling process, and the LSTM-based multivariate time series prediction parameter identification model is further proposed, which solves the multivariable or multi-input problem that is difficult to solve by the traditional time series model. Finally, using the auxiliary power device test sample, the parameter identification model of the APU startup phase is established. The simulation results show that the engine exhaust gas temperature root mean square error is less than 4℃, and the engine speed percentage N root mean square error is less than 1, which satisfies the auxiliary power plant simulation requirements.

Keywords: LSTM; system identification; nonlinear system; APU; deep learning

0 引言

辅助动力装置 (APU) 是飞机上的一个小型的燃气涡轮发动机。其功能是在地面或者飞行中为飞机的空调系统及用电设备提供引气和电力供应, 是飞机不可缺少的子系统。对 APU 进行参数辨识对飞机仿真建模具有重要研究意义^[1]。

系统辨识是根据系统的输入输出函数来确定描述系统行为的数学模型, 系统辨识的目的是通过表征系统行为的重要参数, 建立一个能模仿真实系统行为的数学模型^[2]。近年来, 已经有大量科研人员对发动机系统进行系统辨识, 常用的辨识方法有基于最小二乘法、支持向量机、BP 神经网络等静态辨识方法, 以及 ARIMA、NARX 等基于时间序列模型的动态辨识方法^[3-6]。APU 是一个复杂的非线性时序系统, 静态辨识模型忽略了系统的时序信息, 而传统的时序模型对非线性系统表征能力不足^[7-8]。随着深度学习的兴

起, 深度学习模型逐渐被引入到系统辨识的研究中, 循环神经网络 (RNN) 将时序信息嵌入到神经网络结构中, 使得神经网络具备了动态辨识能力。长短期记忆 (LSTM) 模型作为 RNN 的变体, 解决了 RNN 梯度消失或爆炸、长时记忆能力较差等问题, 使其兼具静态模型的非线性表征能力与时序模型的长时间存储能力。在系统辨识中表现出更强的适应性及泛化性^[9-10]。

APU 工作过程较为复杂, 大致可分为启动、电提取功率、主发动机启动、环控引气及停车 5 个阶段。本文选取 APU 的启动阶段, 结合发动机试车数据, 针对表征 APU 工作状态的排气温度 EGT、转速 N, 提出了一种基于 LSTM 的辅助动力装置参数辨识模型。在建模过程中, 充分考虑影响辨识参数的其他静态特征, 进一步提出了基于 LSTM 的多变量时间序列辨识模型。将所建模型与最小二乘法、神经网络等传统辨识方法进行对比, 验证了该模型的优越性。

1 长短时记忆网络

1.1 网络构型

长短期记忆神经网络 (long short-term memory, LSTM) 是一种时间递归神经网络 (RNN)。传统的神经网

收稿日期: 2019-06-29; 修回日期: 2019-07-25。

基金项目: 国家自然科学基金重点项目 (60832011); 天津市科技攻关计划重点项目 (06YFGZGX00700)。

作者简介: 杜维仲 (1990-), 男, 黑龙江鸡西人, 大学本科, 中级工程师, 主要从事航空器维修方向的研究。

络模型分为输入层、隐藏层、输出层，层与层之间是全连接的，但是每层之间的节点是无连接的，此网络结构对于序列问题建模效果较差^[11-12]。RNN 的隐藏层节点具有连接性与记忆性，会处理信息并将信息传递给当前节点，对序列问题建模有较好的效果，如图 1 所示。

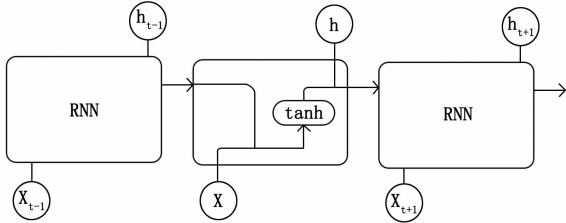


图 1 RNN 网络结构

RNN 输出两层单元，其中隐藏层单元的输出 h_t 以及状态的单元输出 y_t 表达式：

$$h_t = f_a(W_{sh}x_t + W_{hh}h_{t-1} + b_h) \quad (1)$$

$$y_t = W_{hy}h_t + b_y \quad (2)$$

其中： W_{sh} 为输入层到隐藏层的权重系数矩阵； W_{hh} 为隐藏层到隐藏输出的权重系数矩阵； b_h 为隐藏层的偏置向量； f_a 为激活函数； W_{hy} 为隐藏层到状态输出层的权重系数矩阵； b_y 为状态输出层的偏置向量； t 表示时刻。

RNN 虽然解决了序列建模的问题，但是序列长度比较长时，回传时残差指数会下降，进而降低网络权重更新速率，导致长期记忆效果较差，但 LSTM 可以存储记忆因此被提出解决问题。

LSTM 同样是循环结构，但是重复的单元有不同的网络结构。LSTM 采用门结构方式来改变信息至目前细胞状态。门采用选择式的通过方式，其中包含 sigmoid 神经网络以及 pointwise 乘法操作。LSTM 通过遗忘门、输入门、输出门三种门方式来控制细胞状态^[13-15]，网络结构如图 2 所示。

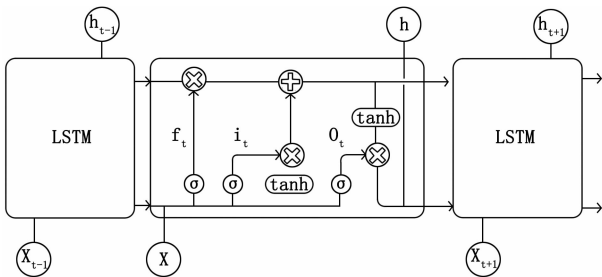


图 2 LSTM 网络结构

遗忘门 f_t 负责控制上一个时刻的输入，选择性地忘记过去某些无用信息，减小记忆负担。

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (3)$$

输入门 i_t 控制当前时刻的输入，同样是过滤无用信息，选择性记忆当前重要信息。

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (4)$$

C_t 为当前细胞状态，此时合并需要记忆的信息，更新细胞状态。

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (5)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (6)$$

输出门 o_t 将过去信息与当前信息合并，输出新的状态信息。

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (7)$$

$$h_t = o_t * \tanh(C_t) \quad (8)$$

1.2 网络训练

LSTM 采用前向传播和反向传播两种训练过程，前向传播通过式 (3) ~ (8) 计算出神经元输出值，反向传播则是 BPTT 算法，从时间和网络两个方面反向计算每个神经元的误差项，根据计算的误差结果，采用不同的优化算法，优化神经元权重。

梯度下降法是优化神经网络中常用算法，例如 SGD、Adagrad、RMSprop、Adam 等梯度下降法。本文采用自适应矩估计 (adaptive moment estimation, Adam) 优化算法，Adam 可优化每一个参数的自适应学习率。拥有 Adadelta 和 RMSprop 算法存储一个指数衰减的历史平方梯度平均值功能的同时还可保存一个历史梯度的指数衰减均值。实际应用表明，Adam 具有算法优势。

2 基于 LSTM 的多变量时间序列辨识模型

根据 APU 的工作原理，表征其工作状态的参数主要有以下两个：排气温度 EGT、发动机转速 N，即本文的辨识目标。传统的时间序列建模只考虑当前时刻与之前时刻的数据关系。这类只考虑时序信息的建模方法有明显不足，时序信息以外的干扰因素同样有着非常重要的作用，比如预测 APU 的发动机排气温度 (EGT)，影响当前 EGT 的因素除了上一时刻的 EGT，其他重要因素还包括发动机点火状态、起动机状态等因素。在 APU 系统辨识建模的过程中，需要将时间以外的其他因素同时考虑进去。

APU 起动阶段，采用时序开环控制，既不引气也不提取电功率；负载压气机出口管道喘振阀 SCV 全开，引气阀 LCV 关闭，负载压气机出口气流通过 SCV 排到 APU 尾喷管，进口导叶 IGV 处于角度最小状态，以提高 APU 起动速度。当启动按钮打开时，APU 起动机通电打开，当 APU 转速达到 7% 时，燃油活门打开，APU 点火启动，当 APU 转速达到 50% 时，APU 起动机断电，当 APU 转速达到 95% 时，点火断开，APU 启动完成。在 APU 起动阶段，SCV、LCV、IGV 三个阀门状态固定，为非变量因素，不考虑其对辨识系统的影响，此阶段系统辨识的影响因素见表 1。

表 1 影响因素

输入信号	信号类型
点火器	布尔型
起动机	布尔型
供油量	连续型

在 APU 工作其他阶段，同样也可以根据工作原理，总

结出其与辨识参数相关的各个特征, 与时间序列特征拼接, 构建基于多变量的时间序列辨识模型。

2.1 模型架构

基于多变量时间序列特征输入的考虑, 本文提出了一种基于 LSTM 的多变量时间序列预测模型, 充分考虑影响辨识参数的其他特征因素, 模型架构如图 3 所示。

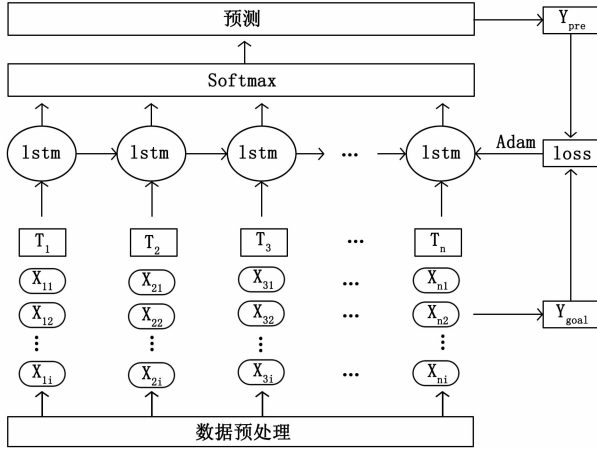


图 3 多变量时间序列辨识模型

在数据准备阶段, 需要对收集的发动机试车数据进行数据预处理, 统一数据类型和量纲, 然后拼接时间序列数据和多变量特征数据, 将拼接的数据送入模型。在网络训练阶段, 前向传播部分通过 softmax 进行回归预测, 反向传播部分通过 Adam 梯度下降算法更新网络参数。最后通过参数寻优、模型评估等方法, 优化网络参数, 提高模型性能。以下分别做了详细介绍。

2.2 数据预处理

通过 APU 发动机试车采集的原始数据, 并不能直接送给模型训练。主要存在以下两个方面问题: 一方面数据量纲不一致, EGT 的温度范围是 $0 \sim 1038 \text{ }^\circ\text{C}$, N 的转速百分比范围是 $0 \sim 110$, 直接送入模型训练, 会导致模型对量纲小的数据不敏感; 另一方面数据类型不同, 影响系统辨识的主要参数数据类型多样, 包括布尔型、离散型、连续型, 不同类型数据含义不一, 直接送入模型会导致模型学习误差, 因此需要对数据进行以下预处理。

1) 布尔型特征和离散性特征向量化, 主要用到的方法有 One-hot 编码、哈夫曼编码等。

2) 连续型特征归一化, 主要解决量纲不一致的情况, 采取数据归一化的方法, 一般会用 maxminmap 算法, 公式为:

$$x_{scale} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (9)$$

其中: x 为待归一数据, x_{min} 为预处理数据最小值, x_{max} 为预处理数据最大值, x_{scale} 即为归一数据。

2.3 多变量时间序列数据输入

传统的时间序列预测的模型, 其输入信息只考虑前一个或者多个时刻对辨识结果的影响。在时刻 k , 传统时序辨

识模型表达式为:

$$EGT_k = f_{EGT}(EGT_{k-1}, EGT_{k-2}, \dots, EGT_{k-n}) \quad (10)$$

$$N_k = f_N(N_{k-1}, N_{k-2}, \dots, N_{k-n}) \quad (11)$$

其中: n 为网络步长。 f_{EGT} 和 f_N 分别为排气温度 EGT 和发动机转速 N 的传统辨识模型表达式。

在 APU 启动阶段, 作用于 k 时刻的辨识参数, 不仅与前 n 个时刻的时间序列有关, 同时作用于辨识参数的影响因素还包括点火器 FIR、起动机 STR、供油量 OIL。此阶段多变量时间序列辨识模型表达式为:

$$EGT'_k = f'_{EGT}(EGT'_{k-1}, EGT'_{k-2}, \dots, EGT'_{k-n}) \quad (12)$$

$$N'_k = f'_N(N'_{k-1}, N'_{k-2}, \dots, N'_{k-n}) \quad (13)$$

$$EGT'_{k-n} = EGT_{k-n} + FIR_{k-n} + STR_{k-n} + OIL_{k-n} \quad (14)$$

$$N'_{k-n} = N_{k-n} + FIR_{k-n} + STR_{k-n} + OIL_{k-n} \quad (15)$$

其中: f'_{EGT} 和 f'_N 分别为排气温度 EGT 和发动机转速 N 的多变量输出辨识模型。参数 k 和 n 的意义同上。

2.4 基于 Grid-Search 的参数寻优

网格搜索算法 (Grid-Search) 是一种通过遍历给定的参数组合来优化模型表现的方法。在模型训练的过程中, 需要调试的参数主要包括网络步长、隐藏层个数、批尺寸、学习率等。在模型训练过程中, 为了确定最优的参数, Grid-Search 采取了一种穷举搜索的调参手段, 在所有候选的参数中, 通过循环遍历, 尝试每一种组合可能, 性能最好的就是最终的辨识模型网络参数。

以隐藏层单元个数为例, 在训练过程中, 首先会根据经验设置参数为 200 个, 以经验参数为中心, 其他参数保持不动的情况下, 模型分别遍历 100、150、200、250、300。测试效果最好的参数, 即作为最终的系统辨识模型隐藏层单元个数。

2.5 基于 K 折交叉验证的模型评估

通过发动机试车得到的数据, 相对来说数量偏小, 划分训练集、验证集和测试集之后, 用于训练的数据更少, 在这种情况下, 分割后的数据分布很有可能和原始数据集的分布情况不一致, 容易造成过拟合或者欠拟合的情况。

为了更加有效地评估模型的泛化性能, 采取 K 折交叉验证的方式评估模型, 此方法对 K 个不同分组数据进行训练, 并将结果进行平均, 进而减少方差, 降低数据敏感性, 提升模型性能, 具体步骤为:

1) 将原始数据不重复抽样并随机分 k 份。

2) 每次挑选其中一份作为测试集, 其余的 $k-1$ 份作为训练集用于模型训练。

3) 重复第二步 k 次, 每个子集都有一次机会作为测试集, 其余机会作为训练集。

4) 在每个训练集上训练后得到一个模型。

5) 计算模型在相应的测试集上的结果, 并保存模型的评估指标。

6) 保存训练模型的评估指标。

7) 计算 k 组测试结果的平均值, 作为 K 折交叉验证下模型的性能指标。

对于模型的评估指标, 本文选择均方根误差 (Root Mean Square Error, RMSE) 作为度量标准。其公式为:

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (\hat{y}_t - y_t)^2}{n}} \quad (16)$$

其中: \hat{y}_t 为时间 t 的回归的预测值, y_t 为时间 t 的观测值, n 为样本点个数。本文通过对测试集的 RMSE 计算, 衡量各个模型的预测精度。

3 模型仿真

3.1 仿真平台

实验所采用的计算机配置如下: 处理器采用英特尔 i7-8550U; 内存为 8.00 GB; 操作系统为 Win10 (64 位); 集成开发环境为 Anaconda 3; 程序设计语言为 Python 3.6.5 (64 位); 采用机器学习包 scikit-learn、谷歌开源框架 tensorflow 1.4.0 进行实验。

3.2 实验结果

数据来源于采集的某型发动机地面起动阶段试车样本, 采集样本包括供油量、点火状态、起动机运行状态。APU 发动起启动时长大约为 12 s 因此以该时长作为试车样本采样周期。采集样本的时间间隔为 0.2 s, 单个样本点为 60 个。

3.2.1 参数寻优

模型训练过程中, 需要确定的参数包括, 网络步长、隐藏层个数、学习率分别采用 Grid-Search 策略进行参数寻优。实验过程记录了在不同参数条件下, 取 RMSE 作为辨识参数 EGT 和 N 辨识结果的评判标准。实验结果如图 4 所示。

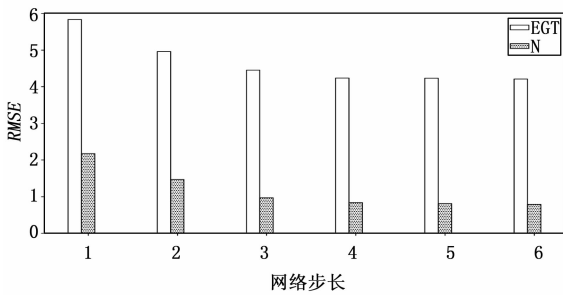


图 4 网络步长参数寻优

根据图 4 的实验结果, 网络步长越大, 其 RMSE 越小, 在步长超过 4 时, 其 RMSE 变化不大, 综合模型效率与准确率考虑, 取网络步长为 4。同时根据图 5 和图 6 的结果可以判断, 转速 N 的参数辨识模型其学习率参数选取为 0.01, 隐藏层个数选取为 250, 排气温度 EGT 的参数辨识模型其学习率参数选取为 0.005, 隐藏层个数选取为 200, 模型辨识效果最佳。

根据网格搜索算法得出的最优参数, 分别训练出 APU 的 EGT 和 N 的参数辨识模型, 其中 EGT 的 RMSE 为 3.64, N 的 RMSE 为 0.86。辨识结果如图 7、图 8 所示。

3.2.2 模型对比

为了验证模型效果, 分别和传统的系统辨识模型进行

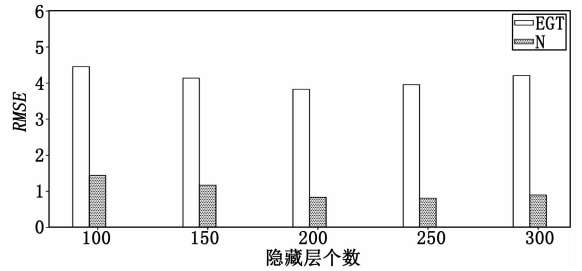


图 5 隐藏层参数寻优

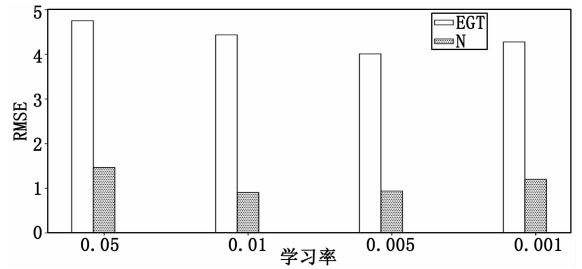


图 6 学习率参数寻优

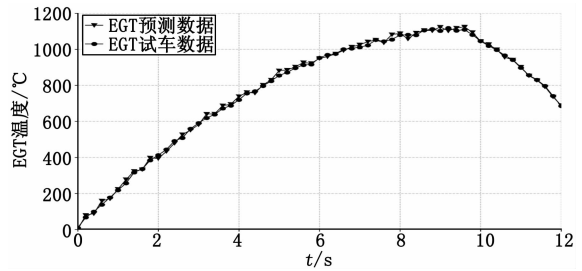


图 7 EGT 辨识模型拟合曲线

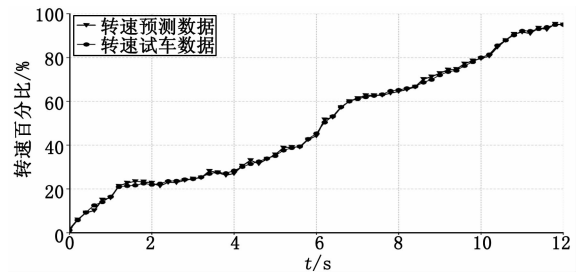


图 8 N 辨识模型拟合曲线

对比, 对比模型包括最小二乘法 (OLS)、支持向量回归 (SVR)、自回归积分滑动平均模型 (ARIMA)、神经网络 (MLP)。实验结果如下:

表 2 模型对比

模型	RMSE/EGT	RMSE/N
OLS	4.82	1.43
SVR	4.89	1.37
ARIMA	4.31	1.17
MLP	4.22	1.08
LSTM	3.54	0.66

实验结果表明, 在 EGT 和 N 的系统辨识任务中, LSTM 因为其长时记忆的特性, 其结果均优于其他同类模型。

4 结束语

本文针对表征 APU 工作状态的 EGT 和 N 两个系统参数, 提出一种基于 LSTM 辅助动力装置系统辨识模型。建模过程中, 结合 APU 发动机工作过程, 分析影响辨识参数的影响因素, 进一步提出了基于 LSTM 的多变量时间序列系统参数辨识模型。最后利用 APU 发动机试车数据, 经过模型对比, 实验结果证实了基于 LSTM 的多变量时间序列参数辨识模型的辨识性能。

基于多变量 LSTM 的时间序列模型在系统参数辨识上具有较大优势, 但是文中模型仍有不足之处, 目前辨识模型还是单步滚动预测, 多步预测还有待探索, 另外目前只进行了 APU 启动阶段的系统参数辨识, 后续阶段的辨识任务还没有进行, 下一步工作将围绕这些任务继续展开。

参考文献:

- [1] 王立文, 王浩. 飞行模拟机辅助动力装置仿真研究 [J]. 计算机工程, 2010, 36 (8): 233 - 235.
- [2] Wang Y. A new concept using LSTM neural networks for dynamic system identification [A]. American Control Conference (ACC) [C]. 2017: 5324 - 5329.
- [3] 荀倩, 王培良, 李祖欣, 等. 基于递推最小二乘法的永磁伺服系统参数辨识 [J]. 电工技术学报, 2016, 31 (17): 161 - 169.
- [4] 熊剑, 刘建业, 赖际舟, 等. 基于高斯粒子滤波的陀螺 ARI-MA 模型辨识方法 [J]. 中国惯性技术学报, 2010, 18 (4):

493 - 497.

- [5] Laurain V, Tóth R, Piga D, et al. An instrumental least squares support vector machine for nonlinear system identification [J]. Automatica, 2015, 54: 340 - 347.
- [6] 江群, 王道波, 李猛. 基于 PEM 的辅助动力装置系统辨识与仿真 [J]. 燃气涡轮试验与研究, 2010, 23 (3): 46 - 48.
- [7] 常博博, 苏三买, 刘铁庚, 等. 辅助动力装置建模及数值仿真 [J]. 航空动力学报, 2011, 26 (9): 2122 - 2127.
- [8] 杨娟, 任仁良, 韩勇. 飞机辅助动力装置电起动系统模型设计及仿真 [J]. 计算机仿真, 2018 (1): 61 - 65.
- [9] 王鑫, 吴际, 刘超, 等. 基于 LSTM 循环神经网络的故障时间序列预测 [J]. 北京航空航天大学学报, 2018, 44 (4): 772 - 784.
- [10] 王晓辉, 朱永利, 王艳, 等. 基于深度学习的电容器介损角在线辨识 [J]. 电工技术学报, 2017, 32 (15): 145 - 152.
- [11] Karim F, Majumdar S, Darabi H, et al. LSTM Fully Convolutional Networks for Time Series Classification [J]. IEEE Access, 2018, 6 (99): 1662 - 1669.
- [12] Ghaeini R, Hasan S A, Datla V, et al. DR-BiLSTM: Dependent Reading Bidirectional LSTM for Natural Language Inference [J]. 2018.
- [13] 罗龙, 李两桓, 王成阳, 等. 基于 ARIMA-LSTM 的绝缘子状态数据挖掘方法 [J]. 电力科学与技术学报, 2017 (4): 38 - 43.
- [14] 杨煜, 张炜. TensorFlow 平台上基于 LSTM 神经网络的人体动作分类 [J]. 智能计算机与应用, 2017, 7 (5): 41 - 45.
- [15] 毛景慧. 基于 LSTM 深度神经网络的股市时间序列预测精度的影响因素研究 [D]. 广州: 暨南大学, 2017.

(上接第 125 页)

- [5] 丁鹏, 田学锋, 葛如海, 等. 农用车空气悬架的混沌振动特性 [J]. 江苏农业科学, 2017, 45 (8): 202 - 206.
- [6] 马克, 米林, 谭伟, 等. 主动悬架非脆弱 H_∞ 控制器设计 [J]. 重庆理工大学学报, 2017 (12).
- [7] 梁军, 庞辉, 陈嘉楠, 等. 车辆半主动悬架分数阶天棚阻尼控制研究 [J]. 机械科学与技术, 2017, 36 (12): 1949 - 1955.
- [8] 周创辉, 文桂林. 基于改进型天棚阻尼控制算法的馈能式半主动油气悬架系统 [J]. 振动与冲击, 2018, 37 (14): 168 - 174.
- [9] Nabawy A E, Abdelrahman A A, Abdalla W S, et al. Analysis of the dynamic behavior of the double wishbone suspension system [J]. International Journal of Applied Mechanics, 2019, 11 (34).
- [10] Kavitha C, Shankar S A, Karthika K, et al. Active camber and toe control strategy for the double wishbone suspension system [J]. Journal of King Saud University - Engineering Sciences, 2018: S1018363917302799.

- [11] 陈涛, 巍郎. 人一车一路互动模式虚拟仿真系统 [J]. 长安大学学报: 自然科学版, 2017, 27 (1): 67 - 71.
- [12] 赵成, 胡增荣, 陈大跃. 半主动悬架的滑模变结构控制 [J]. 中国公路学报, 2007, 20 (3): 109 - 114.
- [13] 刘金琨, 孙富春. 滑模变结构控制理论及其算法研究与进展 [J]. 控制理论与应用, 2016, 24 (3): 407 - 418.
- [14] 余强, 马建. 半主动悬架系统对汽车侧翻稳定性的改善 [J]. 交通运输工程学报, 2005, 5 (1): 53 - 56.
- [15] 郭建辉, 邹金校, 高恩壮. 汽车半主动悬架减振器阻尼匹配设计 [J]. 长安大学汽车学院, 2017: 1677 - 7988.
- [16] 周长城. 车辆悬架设计及理论 [M]. 北京: 北京大学出版社, 2011.
- [17] 周长城. 汽车液压筒式减振器设计及理论 [M]. 北京: 北京大学出版社, 2012.
- [18] 彭莫, 刁增祥, 党潇正. 汽车悬架构件的设计计算 [M]. 北京: 机械工业出版社, 2012.
- [19] 丁鹏, 葛如海, 李智, 等. 汽车自适应可调尾翼系统的研发 [J]. 汽车工程, 2017, 39 (8): 895 - 899.