



ISSN 1001-5965  
CODEN BHHDE8

# 北京航空航天大学 学报

JOURNAL OF BEIJING UNIVERSITY OF  
AERONAUTICS AND ASTRONAUTICS



**2022-06**

**Vol.48 No.6**

目 次

直升机前飞状态旋翼结冰风洞试验研究…………… 黄明其, 王亮权, 袁红刚, 彭先敏, 章贵川 (929)

近距空战训练中的智能虚拟对手决策与导引方法……………  
…………… 孟光磊, 刘德见, 周铭哲, 朴海音, 陈耀飞 (937)

基于损伤力学的增材制造金属材料疲劳寿命预测…………… 洪海铭, 詹志新, 王佳莹 (950)

基于朝向约束和重识别特征的目标轨迹关联方法…………… 艾明晶, 单国志, 刘鹏高, 杨澎钢 (957)

基于  $\mu$ CT 图像的 PVC 泡沫微观试样几何模型生成算法…………… 周勇, 薛斌, 郭昀鑫, 王人鹏 (968)

金属蜗杆与塑料斜齿轮不等齿距啮合方法…………… 任继华, 石照耀, 王得峰, 余石鹏 (979)

卷积神经网络卫星信号自动调制识别算法…………… 崔天舒, 崔凯, 黄永辉, 赵文杰, 安军社 (986)

产品协同设计中异构模型数据融合的有效策略……………  
…………… 薛俊杰, 周军华, 施国强, 宋晓, 蒋炎红, 全红艳 (995)

一种基于自动分区的海量科学数据计算框架…………… 田杨, 晏海华 (1004)

移动机器人自主动态避障方法…………… 张贺, 缪存孝, 唐友军, 闫晓强, 时彦洋, 余远金 (1013)

基于 VMD-MPE 的开关磁阻电机功率变换器故障诊断…………… 张竞文, 熊立新, 马宏昌, 边敦新 (1022)

基于多任务辅助推理的近眼视线估计方法…………… 王小东, 谢良, 闫慧炯, 闫野, 印二威, 李卫国 (1030)

输入受限下多无人机三维协同路径跟踪控制…………… 刘荣华, 刘树光, 张博洋, 王欢, 李伟 (1038)

一种改进 PSO-ARMA 半球谐振陀螺温度误差建模方法…………… 吴宗收, 汪立新, 李新三, 李灿 (1050)

基于 AK-IS 法的航空齿轮泵滑动轴承可靠性分析…………… 高宁, 李华聪, 洪林雄, 曹瑞, 符江锋 (1057)

吸气流动控制对翼身融合飞机气动特性的影响…………… 贾媛, 曹翔, 吴江浩 (1065)

基于 KF-LESO-PID 洛伦兹惯性稳定平台控制…………… 熊颖, 刘强, 任元, 樊亚洪, 孙津济 (1072)

基于 ESO 的电液位置伺服系统反步滑模控制…………… 张震阳, 汪成文, 郭新平, 陈帅 (1082)

多四旋翼无人机系统分布式分层编队合围控制…………… 郑伟铭, 徐扬, 罗德林 (1091)

采用改进导重法的拓扑结构灰度单元过滤技术…………… 高翔, 王林军, 杜义贤, 付君健 (1106)

# JOURNAL OF BEIJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS

Vol. 48 No. 6 (Sum 352) June 2022

---

---

## CONTENTS

- Icing wind tunnel investigation of helicopter rotor model in forward flight state  
..... HUANG Mingqi, WANG Liangquan, YUAN Honggang, PENG Xianmin, ZHANG Guichuan (929)
- Intelligent virtual opponent decision making and guidance method in short-range air combat training  
..... MENG Guanglei, LIU Dejian, ZHOU Mingzhe, PIAO Haiyin, CHEN Yaofei (937)
- Damage mechanics-based fatigue life prediction for additive manufacturing metal materials  
..... HONG Haiming, ZHAN Zhixin, WANG Jiaying (950)
- Target trajectory association method based on orientation constraint and re-identification feature  
..... AI Mingjing, SHAN Guozhi, LIU Penggao, YANG Penggang (957)
- An algorithm for generating geometric models of microscopic specimens of PVC foam based on  $\mu$ CT images  
..... ZHOU Yong, XUE Bin, GUO Yunxin, WANG Rempeng (968)
- Method of unequal pitch meshing of metal worm and plastic helical gear  
..... REN Jihua, SHI Zhaoyao, WANG Defeng, YU Shipeng (979)
- Convolutional neural network based algorithm for automatic modulation recognition of satellite signals  
..... CUI Tianshu, CUI Kai, HUANG Yonghui, ZHAO Wenjie, AN Junshe (986)
- Effective strategy of heterogeneous model data fusion in product collaborative design  
..... XUE Junjie, ZHOU Junhua, SHI Guoqiang, SONG Xiao, JIANG Yanhong, QUAN Hongyan (995)
- A computing framework for massive scientific data based on auto-partitioning algorithm  
..... TIAN Yang, YAN Haihua (1004)
- Dynamic obstacle avoidance method for mobile robots  
..... ZHANG He, MIAO Cunxiao, TANG Youjun, YAN Xiaoqiang, SHI Yanyang, YU Yuanjin (1013)
- Fault diagnosis of switched reluctance motor power converter based on VMD-MPE  
..... ZHANG Jingwen, XIONG Lixin, MA Hongchang, BIAN Dunxin (1022)
- Near-eye gaze estimation based on multitasking auxiliary reasoning  
..... WANG Xiaodong, XIE Liang, YAN Huijiong, YAN Ye, YIN Erwei, LI Weiguo (1030)
- 3D cooperative path following control of multi-UAVs with input saturation  
..... LIU Ronghua, LIU Shuguang, ZHANG Boyang, WANG Huan, LI Wei (1038)
- An improved PSO-ARMA method for temperature error modeling of hemispherical resonator gyroscope  
..... WU Zongshou, WANG Lixin, LI Xinsan, LI Can (1050)
- Reliability analysis of journal bearings inside aero-gear pump based on AK-IS method  
..... GAO Ning, LI Huacong, HONG Linxiong, CAO Rui, FU Jiangfeng (1057)
- Influence of suction flow control on aerodynamic characteristics of blended-wing-body aircraft  
..... JIA Yuan, CAO Xiang, WU Jianghao (1065)
- Lorentz inertial stability platform control based on KF-LESO-PID  
..... XIONG Ying, LIU Qiang, REN Yuan, FAN Yahong, SUN Jinji (1072)
- Backstepping sliding mode control of electro-hydraulic position servo system based on ESO  
..... ZHANG Zhenyang, WANG Chengwen, GUO Xinping, CHEN Shuai (1082)
- Distributed hierarchical formation-containment control of multiple quadrotor UAV systems  
..... ZHENG Weiming, XU Yang, LUO Delin (1091)
- Gray element filtering technology of topology structure based on improved guide-weight method  
..... GAO Xiang, WANG Linjun, DU Yixian, FU Junjian (1106)

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2020.0703

# 直升机前飞状态旋翼结冰风洞试验研究

黄明其, 王亮权\*, 袁红刚, 彭先敏, 章贵川

(中国空气动力研究与发展中心 旋翼空气动力学重点实验室, 绵阳 621000)

**摘 要:** 为研究直升机旋翼在不同前飞状态的结冰情况, 研制了直升机旋翼模型结冰试验台, 在中国空气动力研究与发展中心结冰风洞 4.8 m × 3.2 m 试验段中进行了旋翼结冰风洞试验。以某型直升机 2 m 直径旋翼缩比模型为试验对象, 分析了旋翼转速、初始拉力系数对旋翼结冰的影响。结冰试验过程中保持旋翼操纵恒定, 利用天平测量了旋翼拉力和扭矩性能的动态变化, 并采用二维冰形切割及三维冰形扫描的方式分别测量了桨叶展向典型剖面的翼型及桨叶的整体结冰形态。试验得到了旋翼结冰关键数据, 分析结果表明: 旋翼桨叶结冰主要集中在桨叶前缘和下表面, 结冰会在降低旋翼升力的同时增大旋翼扭矩和功率; 小拉杆杆端轴承的积冰可能造成卡塞, 变距拉杆上的积冰可能造成杆端轴承卡塞, 从而使旋翼操纵失效。

**关键词:** 直升机; 旋翼; 前飞状态; 结冰风洞; 气动性能

**中图分类号:** V211.52

**文献标志码:** A

**文章编号:** 1001-5965(2022)06-0929-08

当环境温度接近或低于冰点、湿度等其他条件达到一定程度时, 飞行器表面会出现结冰问题, 严重的结冰现象是造成飞行事故的一项重要因素。对直升机而言, 飞行过程中旋翼、发动机进气口、驾驶舱风挡等是相对容易结冰的部位, 其中旋翼结冰会直接降低直升机飞行性能, 危害直升机操纵品质和飞行安全<sup>[1]</sup>。中国民用航空规章第 27 部<sup>[2]</sup>和第 29 部<sup>[3]</sup>都对直升机的结冰适航要求做出了明确规定。

为了促进旋翼防除冰系统的研制<sup>[4-7]</sup>, 并获得结冰环境下安全飞行的适航合格审定, 有必要开展直升机旋翼结冰研究。旋翼结冰的主要研究方法包括数值模拟、飞行试验和地面试验 3 种<sup>[8]</sup>。已有学者进行了直升机旋翼结冰三维数值模拟的尝试<sup>[9-12]</sup>, 但由于旋翼桨叶是具有大展弦比的细长体外形, 积冰会明显改变桨叶的质量及刚度分布, 准确的旋翼结冰数值模拟除了需要进行气动/热力学耦合之外, 还需考虑桨叶动力学问题<sup>[13]</sup>。

此外, 高速旋转的旋翼桨尖线速度很高, 离心力效应显著, 可能出现的积冰脱落也给数值模拟带来了很大的困难。目前, 旋翼结冰三维数值模拟技术尚不够成熟。为减小数值计算量, 还有学者提出了旋翼结冰工程简化计算模型<sup>[14-15]</sup>。

飞行试验需要使直升机在自然结冰环境下飞行<sup>[16-17]</sup>, 试验气象条件和结冰环境参数无法控制, 另外由于直升机航程偏短, 无法长时间留在空中, 单次飞行有效试验时间很短。这些因素导致直升机结冰飞行试验周期十分漫长, 以法国 AS332 超级美洲豹直升机的结冰飞行试验为例, 试验持续了 9 年才获得最终认证, 时间及经济成本高昂, 且直升机结冰飞行试验安全风险较高。

地面试验是飞行试验良好的补充及替代手段<sup>[18]</sup>。在地面试验手段中, 结冰风洞能够提供稳定的结冰试验环境, 相对结冰飞行试验具有很高的试验效率, 试验风险和成本显著低于飞行试验。关于二维翼型<sup>[19-21]</sup>及固定翼飞机<sup>[22-24]</sup>的结冰风

收稿日期: 2020-12-21; 录用日期: 2021-02-20; 网络出版时间: 2021-03-09 19:24

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20210309.0905.001.html](http://kns.cnki.net/kcms/detail/11.2625.V.20210309.0905.001.html)

基金项目: 国家自然科学基金(11902335); 中国空气动力研究与发展中心基础和前沿技术研究基金(PJD20180146)

\* 通信作者。E-mail: wangliangquan@cardc.cn

**引用格式:** 黄明其, 王亮权, 袁红刚, 等. 直升机前飞状态旋翼结冰风洞试验研究[J]. 北京航空航天大学学报, 2022, 48(6): 929-936. HUANG M Q, WANG L Q, YUAN H G, et al. Icing wind tunnel investigation of helicopter rotor model in forward flight state [J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(6): 929-936 (in Chinese).

洞试验进行得较多,国外开展了一些围绕对直升机翼型<sup>[25]</sup>和旋翼的结冰试验<sup>[26-27]</sup>,这些试验主要研究结冰环境参数(如液态水含量、水滴平均粒径及静温等)对结冰特性的影响。国内宋建宇等<sup>[28]</sup>在2020年进行了关于直升机涡轴发动机结冰的风洞试验,但有关旋翼结冰风洞试验的研究仍较为稀缺。

开展直升机旋翼结冰风洞试验需要克服以下难题:①为了测量旋翼结冰过程中拉力和扭矩的动态变化,需要专门研制能适应低温环境、具有温度补偿功能的天平,旋翼主轴倾斜系统、操纵系统等部件也均需要能够在低温环境下正常工作;②与飞机机翼、发动机进气口等固定的试验模型不同,直升机旋翼是高速旋转部件,强离心力场可能使桨叶上尤其是桨尖的积冰脱落,桨叶模型较为细长的外形使得精确测量桨叶冰形较为困难;③旋翼上各片桨叶的结冰形态并不完全一致,由于离心力导致积冰的随机脱落可能造成旋翼动力学不平衡,从某一片桨叶上脱落的冰甚至会直接撞击到后续高速旋转的桨叶上,轻则带来试验台异常振动,重则有损坏试验台的风险,需要建立完善的试验信息监控系统,确保试验台的运行安全。

本文利用研制的直升机旋翼模型结冰试验台及建立的配套试验技术,在中国空气动力研究与发展中心(China Aerodynamics Research and Development Center, CARDC)结冰风洞中开展了直升机旋翼结冰风洞试验。试验研究了旋翼转速和初始拉力系数对旋翼结冰特性的影响,获取了结冰过程中旋翼拉力和扭矩性能的动态变化数据,测量了桨叶展向典型剖面的二维翼型冰形和三维桨叶的整体冰形,分析了桨叶结冰现象与旋翼性能变化之间的关系。试验结果有助于建立对直升机旋翼结冰问题的直观认识,也可为旋翼结冰数值模拟研究提供参考。

## 1 试验设备和试验模型

本文试验于2020年在CARDC结冰风洞4.8 m×3.2 m试验段开展<sup>[29]</sup>,该风洞可以模拟低至-40℃的环境温度,位于试验段上游的喷雾系统提供试验所需的结冰云雾,云雾中的液态水含量(LWC)可在0.2~3.0 g/m<sup>3</sup>、平均粒径(MVD)可在10~300 μm范围内调节。

安装在结冰风洞中的直升机旋翼结冰试验台如图1所示。使用某型直升机旋翼2 m直径缩比模型,旋翼旋转方向为俯视逆时针,桨毂型式为无铰式,安装有5片桨叶,在桨叶展向特定位置处进

行标记以便于翼型冰形取样,桨叶弦长约0.06 m,桨叶尖部存在抛物线尖削,桨叶模型三视图如图2所示。缩比的旋翼模型试验时与全尺寸旋翼拉力系数一致,桨尖及桨叶特征剖面的线速度相近,以求较好地反映全尺寸旋翼的结冰特性。

使用能在低温环境下工作的专用旋翼天平 and 扭矩天平,分别测量旋翼结冰过程中拉力和扭矩的动态变化,天平包裹在旋翼下方的整流罩中。驱动旋翼的电机及减速器安装在风洞地板下方,电机最大输出功率为90 kW。通过一根垂直安装的传动轴将电机的输出功率传递到旋翼上。



图1 安装在CARDC结冰风洞中的旋翼试验台

Fig. 1 Rotor test rig installed in CARDC icing wind tunnel

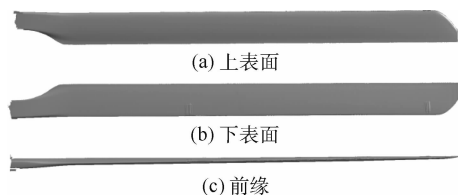


图2 桨叶模型三视图

Fig. 2 Three-view drawing of rotor blade model

## 2 试验状态和试验方法

以Case01状态为基准试验状态,该状态旋翼转速为1 800 r/min,旋翼旋转桨尖马赫数为0.554 4,拉力系数为0.008。研究变旋翼转速(Case02、Case03)、拉力系数(Case04、Case05)对旋翼结冰的影响,试验状态如表1所示。所有试验状态保持风洞来流速度和结冰环境参数一致,来流速度为37.7 m/s(对应基准试验状态旋翼前进比为0.2),结冰云雾播撒时间为180 s,其中LWC=0.70 g/m<sup>3</sup>,过冷水滴MVD=20 μm,来流静温为-15℃。

进行直升机旋翼结冰试验时,采用如下试验流程:①通过结冰风洞制冷系统将洞体内空气温度降至指定温度;②试验台动力电机启动,驱动旋

表1 旋翼结冰试验状态

Table 1 Icing conditions for rotor test

试验编号	来流速度/ ( $m \cdot s^{-1}$ )	旋翼转速/ ( $r \cdot min^{-1}$ )	拉力系数
Case01	37.7	1 800	0.008
Case02	37.7	1 080	0.008
Case03	37.7	1 440	0.008
Case04	37.7	1 800	0.006
Case05	37.7	1 800	0.010

翼模型;③风洞起风,使风速达到试验指定值;④根据风洞起风过程中气流速度的变化,对旋翼进行动态配平,随后开始采集旋翼拉力和扭矩信息;⑤喷雾系统启动,向风洞试验段流场中注入结冰云雾;⑥待喷雾结束后,风洞停止吹风,试验台停车;⑦进行桨叶翼型冰形测量和三维冰形扫描;⑧对试验台及旋翼模型除冰后进行检查,并准备开展下一状态的试验。按照上述步骤完成单个试验状态需耗时1~2 h。

笔者团队基于模糊控制理论,开发了旋翼风洞试验自动配平软件<sup>[30]</sup>。进行动态配平时,固定旋翼轴倾角,调整旋翼总距和周期变距,使旋翼拉力达到预定值,并使旋翼滚转力矩和偏航力矩为零。完成旋翼配平后,不再调整旋翼的操纵量。待试验台停止运行后,试验人员需穿着防寒服进入结冰风洞试验段,取下冰形保持较为完整的桨叶,采用描迹法测量桨叶展向0.3R、0.6R和0.8R(R为旋翼半径)3个翼型剖面处的冰形<sup>[31]</sup>,使用三维扫描设备测量整片桨叶的结冰情况。需要指出的是,由于离心力的影响,单片桨叶难以在展向不同的翼型剖面均保留完整的结冰形态,因此二维翼型冰形测量的结果可能来自不同的桨叶。对于整片桨叶的三维冰形扫描,则只能选取一片冰形保持相对最完整的桨叶。此外,使用数码相机对试验台及旋翼模型的局部结冰进行拍照。

### 3 试验结果及讨论

#### 3.1 旋翼结冰典型试验结果

图3和图4分别为旋翼模型、桨毂及变距机构上的典型结冰情况。可以看出,旋翼结冰主要集中于桨叶前缘,桨叶外段至桨尖有明显的结冰脱落痕迹。在180 s的结冰时间内,自动倾斜器及变距拉杆上积了一层质地坚硬的冰,而小拉杆杆端轴承的结冰已经在一定程度上限制了轴承的活动范围。若结冰云雾播撒时间进一步加长,轴承将非常容易出现卡塞现象,造成旋翼操纵失灵,影响试验台的安全。

图5为基准试验状态旋翼的拉力和扭矩时间历程(用拉力系数 $C_t$ 和扭矩系数 $C_q$ 表示)。在 $t=15$  s喷雾系统启动后至 $t=40$  s,旋翼拉力快速下降,扭矩则急剧增大。 $t=40$  s后旋翼拉力和扭矩围绕某一平均值波动,说明该状态下旋翼结冰和冰脱落速率基本一致,一个积冰生成-脱落周期持续时间约为15~20 s。试验过程中旋翼拉力最大下降了32.9%,而扭矩的增加超过4倍。图6为基准试验状态桨叶前缘局部积冰特写,该状态下桨叶前缘积冰呈晶莹透明状,由于离心力的作



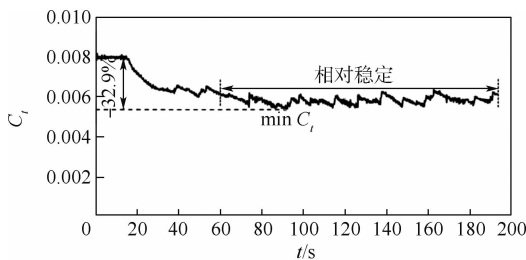
图3 桨叶上的结冰情况

Fig. 3 Ice accretion on rotor blade

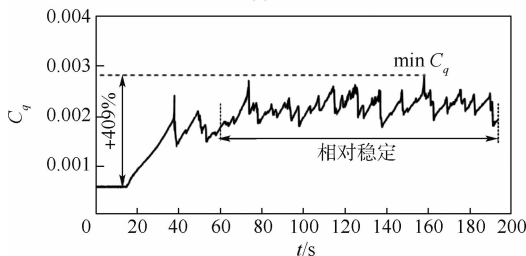


图4 桨毂及变距机构上的结冰情况

Fig. 4 Ice accretion on hub and pitch link



(a) 旋翼拉力



(b) 旋翼扭矩

图5 基准试验状态(Case01)旋翼性能的变化  
Fig. 5 Time history of rotor performance for baseline case (Case01)

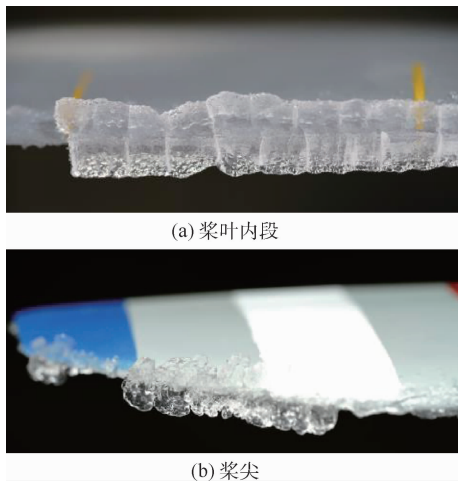


图6 基准试验状态 (Case01) 桨叶前缘的积冰特写  
Fig. 6 Close-up view of ice accretion on blade for baseline case (Case01)

用,可以在图6(a)中看到数条冰裂纹,而在图6(b)中桨尖位置处,可以看到仅残存了少量的明冰。

### 3.2 转速对旋翼结冰的影响

图7为3个不同转速下旋翼拉力和扭矩时间历程对比曲线。1 080 r/min 的低转速试验状态 (Case02)下,旋翼拉力下降幅值最小,拉力先下降了18.4%,然后又恢复了13.5%,该状态旋翼扭矩增长也较慢,在  $t = 120$  s 之后才达到峰值,但扭矩峰值很高,这可能与积冰在桨叶上的附着位置和附着形态有关。1 440 r/min 试验状态 (Case03)的旋翼拉力变化趋势与基准试验状态 (Case01)基本一致,第1个扭矩峰值出现在喷雾开始后36.6 s,比基准试验状态第1个扭矩峰值的出现滞后了13.5 s。从  $t = 40$  s 后的扭矩曲线上看,1 440 r/min 试验状态 (Case03)的一个波动周期约为30 s,但该状态的扭矩波动幅值显著高于1 800 r/min 基准试验状态 (Case01)。综合而言,旋翼转速越低,单位时间内桨叶表面尤其是桨叶前缘受到结冰云雾中的过冷液滴冲击次数越少,因此,积冰导致的旋翼拉力和扭矩性能的变化也相对较慢。

图8为旋翼转速对桨叶展向3个剖面的翼型积冰冰形的影响。可以看出,尽管低转速试验状态 (Case02)桨叶积冰速度较慢,但试验结束后桨叶上存留的积冰反而更多,这在0.6R和0.8R两个剖面表现得尤为明显,这2个剖面的积冰冰形超出桨叶前缘分别约0.14c和0.19c(c为翼型弦长),冰形厚度超过了翼型最大相对厚度,而另外2个转速试验状态 (Case01、Case03)的积冰冰形超出桨叶前缘不到0.06c。

图9为1 080 r/min 转速试验状态 (Case02)

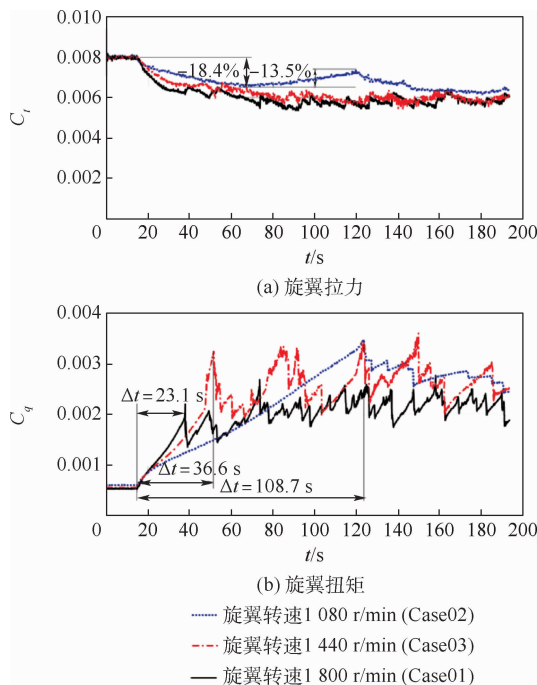


图7 旋翼转速对旋翼性能的影响  
Fig. 7 Effect of rotational speed on rotor performance

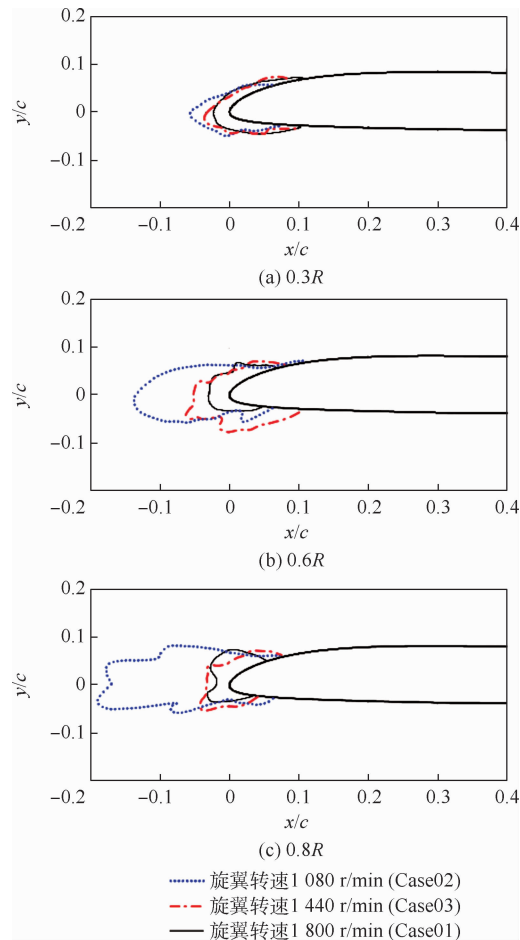


图8 旋翼转速对翼型剖面冰形的影响

Fig. 8 Effect of rotational speed on airfoil ice accretion  
桨叶的积冰照片。从桨叶局部特写照片可以看出,前缘积冰呈霜冰和明冰结合形成的混合冰形态,桨叶上表面除前缘外较为光洁,而下表面则分

散凝结了较多的细小颗粒状霜冰。由于离心力的影响,图中2片桨叶尖部的冰基本脱落。图10为1440 r/min 转速试验状态 (Case03) 桨叶的积冰照片。照片中桨叶前缘的积冰保持较为完整,积冰呈不透明状,且在抛物线桨尖前缘形成了独特的贝壳状冰。对比图9和图10可以看出,1080 r/min低转速试验状态桨叶前缘的积冰尺寸更大。

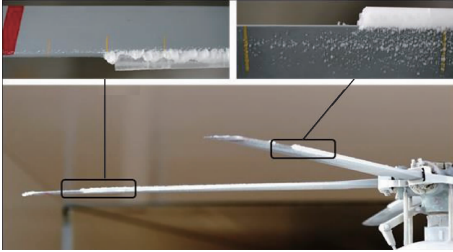


图9 Case02 试验状态桨叶积冰照片

Fig.9 Picture of ice accretion on blade for test Case02

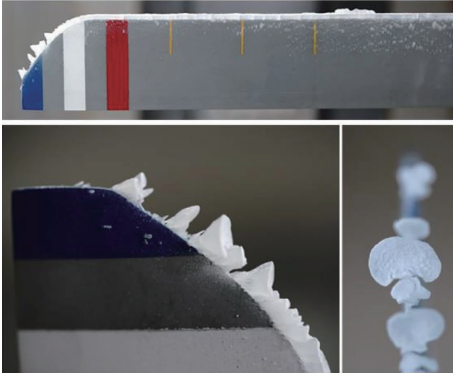


图10 Case03 试验状态桨叶积冰照片

Fig.10 Picture of ice accretion on blade for test Case03

### 3.3 初始拉力系数对旋翼结冰的影响

图11为初始拉力系数  $C_{t0}$  对旋翼性能的影响,3种试验状态下旋翼初始拉力系数分别为0.006 (Case04)、0.008 (Case01) 和0.010 (Case05)。可以看出,3种试验状态下旋翼拉力和扭矩变化趋势基本一致,旋翼经过一段时间的拉力快速下降和扭矩快速增加后,性能变化达到一种动态平衡。初始拉力系数为0.010的大拉力试验状态 (Case05) 旋翼拉力和扭矩的波动略高于另外2种试验状态,而基准试验状态 (初始拉力系数  $C_{t0} = 0.008$ ) 的旋翼性能波动幅值相对较小。

图12为旋翼初始拉力系数对桨叶展向3个剖面的翼型积冰冰形的影响。与旋翼性能的波动情况相对应,初始拉力系数为0.008的基准试验状态 (Case01) 桨叶前缘的积冰范围最小,在0.6R桨叶剖面,初始拉力系数0.006和0.010两种试验状态的积冰范围均较大,约超出前缘0.1c,冰形厚度同样超过了翼型最大厚度,其中  $C_{t0} =$

0.010试验状态的积冰范围略大。综合而言,未发现旋翼初始拉力系数对桨叶结冰特性有明显的影响规律。

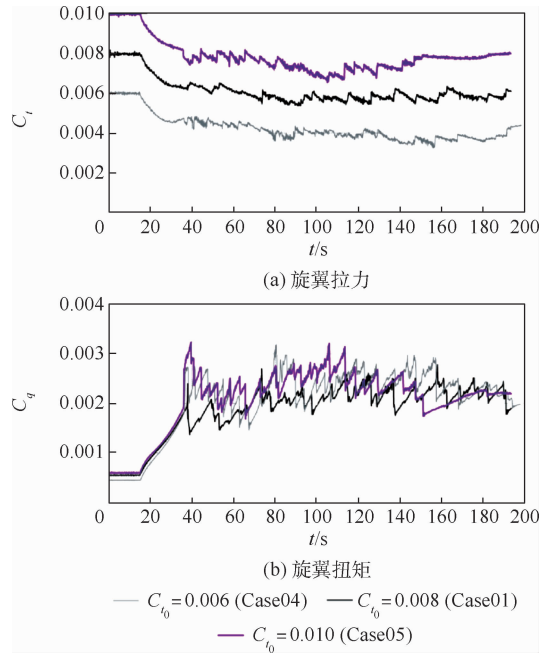


图11 初始拉力系数对旋翼性能的影响

Fig.11 Effect of initial thrust coefficient on rotor performance

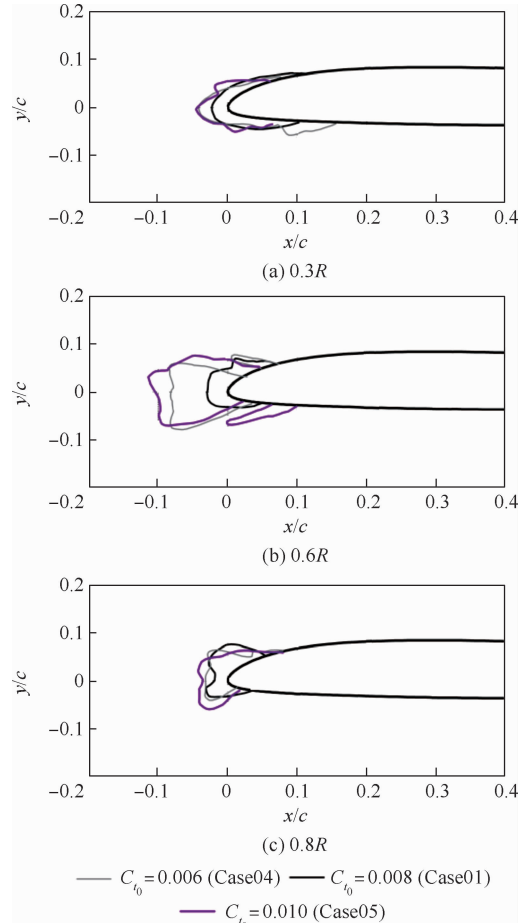


图12 初始拉力系数对翼型剖面冰形的影响

Fig.12 Effect of initial thrust coefficient on airfoil ice accretion

图 13 和图 14 分别给出了  $C_{t0} = 0.006$  (Case04) 和  $0.010$  (Case05) 两种试验状态的桨叶冰形三维扫描结果。可以看出, 桨叶冰形都保持较为完好。桨叶前缘的积冰使得前缘气动外形不再光滑, 同时也加大了桨叶的气动弦长。在桨叶下表面内段区域, 附着了大量粗糙的冰粒。图 15 为  $C_{t0} = 0.006$  (Case04) 试验状态桨叶前缘的积冰及下表面的粗糙冰粒特写。图 15(a) 中桨叶前缘的积冰上有一道较深的沟槽, 从图 15(b) 中可以看出桨叶下表面冰粒的强大附着能力。

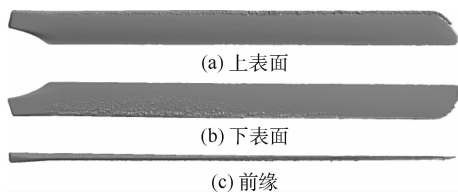


图 13 Case04 试验状态的桨叶积冰扫描结果

Fig. 13 Scan results of ice accreted on blade for test Case04

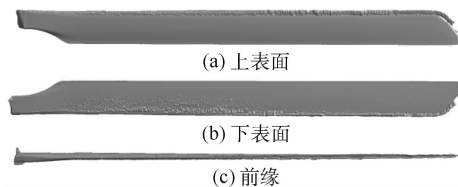
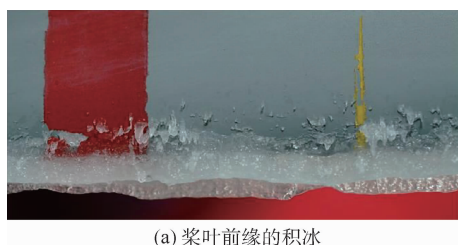


图 14 Case05 试验状态的桨叶积冰扫描结果

Fig. 14 Scan results of ice accreted on blade for test Case05



(a) 桨叶前缘的积冰



(b) 附着在桨叶下表面的粗糙冰粒

图 15 桨叶局部的积冰特写 (Case04)

Fig. 15 Picture of ice accretion on blade for test Case04

## 4 结论

基于所研制的直升机旋翼结冰试验台, 使用某型直升机 2 m 直径缩比旋翼模型, 在中国空气动力研究与发展中心结冰风洞 4.8 m × 3.2 m 试验段中开展了旋翼结冰试验, 研究了旋翼转速和初始拉力系数对旋翼结冰外形及性能的影响, 主

要结论如下:

1) 直升机旋翼结冰问题值得引起更多的关注。桨叶积冰会导致旋翼拉力迅速下降的同时, 扭矩大大增加。基准试验状态旋翼拉力最大下降了 32.9%, 扭矩最大增加了 4 倍。旋翼拉力下降意味着真实飞行时直升机会迅速丢失飞行高度, 扭矩增加则意味着发动机功率消耗急剧增大。为了保持旋翼输出拉力的稳定, 旋翼扭矩和功率增加量将会更大, 容易造成飞行事故。

2) 转速对桨叶积冰和旋翼性能的影响较大。低转速试验状态 (1 080 r/min) 桨叶上的积冰速度显著低于 1 400 r/min 和 1 800 r/min 试验状态, 180 s 的结冰喷雾过程中旋翼拉力的损失不超过 20%, 但该转速状态试验结束后桨叶上存留的积冰反而更多。旋翼初始拉力系数对桨叶结冰的影响程度不大, 未发现明显的变化规律。

3) 除了要注意桨叶表面结冰对旋翼气动特性的影响外, 还需关注桨毂及变距机构上的积冰对结构的影响。其中变距拉杆杆端轴承由于需要传递操纵旋翼的铰链力矩, 结构较为脆弱, 轴承缝隙处的积冰是较大的安全隐患, 积冰过量可能造成轴承运动受阻并导致旋翼操纵失灵。

本文是在旋翼操纵恒定的前提下, 研究了旋翼结冰过程中的拉力和扭矩性能变化, 未来可以研究在结冰风洞试验环境下, 保持一定的旋翼拉力, 需要的旋翼操纵量及扭矩 (功率) 的变化, 确定不同结冰条件下的安全飞行边界。

## 参考文献 (References)

- [1] CAO Y, CHEN K. Helicopter icing[J]. The Aeronautical Journal, 2010, 114(1152): 83-90.
- [2] 中华人民共和国交通运输部. 中国民用航空规章第 27 部 正常类旋翼航空器适航规定: CCAR-27-R2[S]. 北京: 中华人民共和国交通运输部, 2017.  
Ministry of Transport of the People's Republic of China. China civil aviation regulation No. 27: Airworthiness regulations for normal rotorcraft: CCAR-27-R2[S]. Beijing: Ministry of Transport of the People's Republic of China, 2017 (in Chinese).
- [3] 中华人民共和国交通运输部. 中国民用航空规章第 29 部 运输类旋翼航空器适航规定: CCAR-29-R2[S]. 北京: 中华人民共和国交通运输部, 2017.  
Ministry of Transport of the People's Republic of China. China civil aviation regulation No. 29: Airworthiness regulations for transport rotorcraft: CCAR-29-R2[S]. Beijing: Ministry of Transport of the People's Republic of China, 2017 (in Chinese).
- [4] TARQUINI S, ANTONINI C, AMIRFAZLI A, et al. Investigation of ice shedding properties of superhydrophobic coatings on helicopter blades[J]. Cold Regions Science and Technology,

- 2014,100:50-58.
- [ 5 ] PALACIOS J, WOLFE D, BAILEY M, et al. Ice testing of a centrifugally powered pneumatic deicing system for helicopter rotor blades [ J ]. *Journal of the American Helicopter Society*, 2015, 60 ( 3 ): 1-12.
- [ 6 ] 常士楠,刘达经,袁修干. 直升机旋翼桨叶防/除冰系统防护范围研究 [ J ]. *航空动力学报*, 2007, 22 ( 3 ): 360-364.
- CHANG S N, LIU D J, YUAN X G. Research on protected range of the anti-icing/deicing system for helicopter rotor [ J ]. *Journal of Aerospace Power*, 2007, 22 ( 3 ): 360-364 ( in Chinese ).
- [ 7 ] 洪海华,刘伟光,艾剑波,等. 直升机的防除冰系统 [ J ]. *直升机技术*, 2010 ( 1 ): 52-56.
- HONG H H, LIU W G, AI J B, et al. Anti-icing/deicing system of helicopter rotor [ J ]. *Helicopter Technique*, 2010 ( 1 ): 52-56 ( in Chinese ).
- [ 8 ] SHAW R, RICHTER G. The UH-1H helicopter icing flight test program-An overview [ C ] // 23rd Aerospace Sciences Meeting. Reston: AIAA, 1985: 338.
- [ 9 ] 钟国,曹义华,赵明. 直升机旋翼积冰的数值模拟 [ J ]. *北京航空航天大学学报*, 2012, 38 ( 3 ): 330-334.
- ZHONG G, CAO Y H, ZHAO M. Numerical simulation of ice accretion on helicopter rotor [ J ]. *Journal of Beijing University of Aeronautics and Astronautics*, 2012, 38 ( 3 ): 330-334 ( in Chinese ).
- [ 10 ] KELLY D, HABASHI W G, QUARANTA G, et al. Ice accretion effects on helicopter rotor performance, via multibody and CFD approaches [ J ]. *Journal of Aircraft*, 2018, 55 ( 3 ): 1165-1176.
- [ 11 ] WANG Z Z, ZHAO N, ZHU C L. Numerical simulation for three-dimensional rotor icing in forward flight [ J ]. *Advances in Mechanical Engineering*, 2018, 10 ( 4 ): 168781401877240.
- [ 12 ] CHEN X, ZHAO Q J, BARAKOS G. Numerical analysis of aerodynamic characteristics of iced rotor in forward flight [ J ]. *AIAA Journal*, 2018, 57 ( 4 ): 1523-1537.
- [ 13 ] WANG L Q, XU G H, SHI Y J. Rotor airload and acoustics prediction based on CFD/CSD coupling method [ J ]. *Transactions of Nanjing University of Aeronautics and Astronautics*, 2018 ( 2 ): 343-352.
- [ 14 ] 李国知,曹义华. 旋翼结冰对直升机飞行动力学特性的影响 [ J ]. *航空学报*, 2011, 32 ( 2 ): 187-194.
- LI G Z, CAO Y H. Effect of rotor icing on helicopter flight dynamic characteristics [ J ]. *Acta Aeronautica et Astronautica Sinica*, 2011, 32 ( 2 ): 187-194 ( in Chinese ).
- [ 15 ] CHEN N L, JI H H, CAO G Z, et al. A three-dimensional mathematical model for simulating ice accretion on helicopter rotors [ J ]. *Physics of Fluids*, 2018, 30 ( 8 ): 83602.
- [ 16 ] FLEMMING R J, HANKS K W, HANKS M L. US army UH-60M helicopter main rotor ice protection system [ C ] // 2007 SAE Aircraft and Engine Icing International Conference, 2007: 1-16.
- [ 17 ] HOFF S C, VORST J, FLEMMING R J, et al. Icing certification of Korean utility helicopter KUH-1: NLR-TP-2017-098 [ R ]. Amsterdam: NLR, 2017: 1-18.
- [ 18 ] FLEMMING R, ALLDRIDGE P, DOEPPNER R. Artificial icing tests of the S-92A helicopter in the McKinley Climatic Laboratory: AIAA 2004-737 [ R ]. Reston: AIAA, 2004.
- [ 19 ] BRAGG M B, BROEREN A P, BLUMENTHAL L A. Iced-airfoil aerodynamics [ J ]. *Progress in Aerospace Sciences*, 2005, 41 ( 5 ): 323-362.
- [ 20 ] LYNCH F T, KHODADOUST A. Effects of ice accretions on aircraft aerodynamics [ J ]. *Progress in Aerospace Sciences*, 2001, 37 ( 8 ): 669-767.
- [ 21 ] 张大林,陈维建. 飞机机翼表面霜状冰结冰过程的数值模拟 [ J ]. *航空动力学报*, 2004, 19 ( 1 ): 137-141.
- ZHANG D L, CHEN W J. Numerical simulation of rime ice accretion process on airfoil [ J ]. *Journal of Aerospace Power*, 2004, 19 ( 1 ): 137-141 ( in Chinese ).
- [ 22 ] CEBECI T, KAFYEKE F. Aircraft icing [ J ]. *Annual Review of Fluid Mechanics*, 2003, 35: 11-21.
- [ 23 ] 桂业伟,周志宏,李颖晖,等. 关于飞机结冰的多重安全边界问题 [ J ]. *航空学报*, 2017, 38 ( 2 ): 520734.
- GUI Y W, ZHOU Z H, LI Y H, et al. Multiple safety boundaries protection on aircraft icing [ J ]. *Acta Aeronautica et Astronautica Sinica*, 2017, 38 ( 2 ): 520734 ( in Chinese ).
- [ 24 ] 易贤. 飞机积冰的数值计算与积冰试验相似准则研究 [ D ]. 绵阳:中国空气动力研究与发展中心, 2007: 17-19.
- YI X. Numerical computation of aircraft icing and study on icing test scaling law [ D ]. Mianyang: China Aerodynamics Research and Development Center, 2007: 17-19 ( in Chinese ).
- [ 25 ] FLEMMING R J, LEDNICER D A. Correlation of icing relationships with airfoil and rotorcraft icing data [ J ]. *Journal of Aircraft*, 1986, 23 ( 10 ): 737-743.
- [ 26 ] MILLER T L, BOND T H. Icing research tunnel test of a model helicopter rotor: NASA-TM-101978 [ R ]. Washington, D. C. : NASA, 1989: 1-12.
- [ 27 ] BRITTON R, BOND T, FLEMMING R. An overview of a model rotor icing test in the NASA Lewis icing research tunnel: AIAA-94-0716 [ R ]. Reston: AIAA, 1994.
- [ 28 ] 宋建宇,吴昊峰,邱长波,等. 民用涡轴发动机进气系统结冰试验 [ J ]. *航空动力学报*, 2020, 35 ( 5 ): 1089-1098.
- SONG J Y, WU J F, QIU C B, et al. Civil turboshaft engine induction system icing test [ J ]. *Journal of Aerospace Power*, 2020, 35 ( 5 ): 1089-1098 ( in Chinese ).
- [ 29 ] 倪章松,刘森云,王桥,等. 3 m × 2 m 结冰风洞试验技术研究进展 [ J ]. *实验流体力学*, 2019, 33 ( 6 ): 46-53.
- NI Z S, LIU S Y, WANG Q, et al. Research progress of test technologies for 3 m × 2 m icing wind tunnel [ J ]. *Journal of Experiments in Fluid Mechanics*, 2019, 33 ( 6 ): 46-53 ( in Chinese ).
- [ 30 ] 章贵川,彭先敏,车兵辉,等. 共轴刚性旋翼试验自动配平技术研究 [ J ]. *南京航空航天大学学报*, 2019, 51 ( 2 ): 226-231.
- ZHANG G C, PENG X M, CHE B H, et al. Research on automatic trim technology of coaxial rigid rotor test [ J ]. *Journal of Nanjing University of Aeronautics & Astronautics*, 2019, 51 ( 2 ): 226-231 ( in Chinese ).
- [ 31 ] 易贤,王斌,李伟斌,等. 飞机结冰冰形测量方法研究进展 [ J ]. *航空学报*, 2017, 38 ( 2 ): 520711.
- YI X, WANG B, LI W B, et al. Research progress on ice shape measurement approaches for aircraft icing [ J ]. *Acta Aeronautica et Astronautica Sinica*, 2017, 38 ( 2 ): 520711 ( in Chinese ).

## Icing wind tunnel investigation of helicopter rotor model in forward flight state

HUANG Mingqi, WANG Liangquan<sup>\*</sup>, YUAN Honggang, PENG Xianmin, ZHANG Guichuan

(Key Laboratory of Rotor Aerodynamics, China Aerodynamics Research and Development Center, Mianyang 621000, China)

**Abstract:** To investigate the effect of ice accretion on helicopter rotors in forward flight in different flight states, a helicopter rotor icing test system was developed. A rotor model icing test was conducted in the 4.8 m × 3.2 m test section of the China Aerodynamics Research and Development Center icing wind tunnel. The effect of the rotational speed and initial thrust on the performance of a 2 m-diameter rotor model was examined. During the icing test, the control angle of the rotor was maintained constant, and the dynamic variation of the rotor thrust and torque was measured by a balance. The ice shape of the airfoil at typical span stations and of the entire blade was measured by two-dimensional ice cutting and three-dimensional scanning system, respectively. Test results showed that ice accretion occurred primarily on the blade leading edge and lower surface, and that icing increased rotor torque and power while decreasing rotor thrust. The rod end bearing of the pitch link could be stuck by accreted ice, resulting in a loss of rotor control.

**Keywords:** helicopter; rotor; forward flight state; icing wind tunnel; aerodynamic performance

**Received:** 2020-12-21; **Accepted:** 2021-02-20; **Published online:** 2021-03-09 19:24

**URL:** [kns.cnki.net/kcms/detail/11.2625.V.20210309.0905.001.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210309.0905.001.html)

**Foundation items:** National Natural Science Foundation of China (11902335); CARDC Fundamental and Frontier Technology Research Fund (PJD20180146)

\* **Corresponding author.** E-mail: wangliangquan@cardc.cn

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2020.0687

# 近距离空战训练中的智能虚拟对手 决策与导引方法

孟光磊<sup>1,\*</sup>, 刘德见<sup>1</sup>, 周铭哲<sup>1</sup>, 朴海音<sup>2</sup>, 陈耀飞<sup>1</sup>

(1. 沈阳航空航天大学自动化学院, 沈阳 110136; 2. 航空工业沈阳飞机设计研究所, 沈阳 110035)

**摘 要:** 针对近距离空战训练中智能虚拟对手攻防博弈的自主决策与占位导引问题, 提出了基于动态贝叶斯网络(DBN)和约束梯度法的智能虚拟对手决策和导引一体化方法。结合空间占位态势、火控攻击区和机动动作识别结果等信息, 建立近距离空战决策动态贝叶斯网络模型, 实现根据战场动态环境变化的占位导引指标决策。针对在线识别的各类目标机动动作, 建立轨迹预测模型, 实现目标轨迹的实时预测。根据占位导引指标和目标预测轨迹, 考虑飞行性能约束, 采用约束梯度法计算智能虚拟对手的优化占位导引量。实现了近距离空战智能虚拟对手空间占位决策与导引量计算的无缝结合。近距离空战仿真实验结果表明: 所提方法能够实现智能虚拟对手的合理化自主决策和占位导引, 克服了传统方法实现机动动作方式固化的问题, 具备较好的实时性和优化性。

**关键词:** 空战训练; 智能虚拟对手; 占位导引; 机动识别; 轨迹预测; 动态贝叶斯网络(DBN)

**中图分类号:** V221<sup>+</sup>.3; TB553

**文献标志码:** A **文章编号:** 1001-5965(2022)06-0937-13

随着人工智能技术的发展, 智能决策系统在军事领域的应用越来越深入。基于人工智能技术开发具有自主决策能力的智能虚拟对手协助飞行员进行空战训练, 受到各航空强国的重视<sup>[1-2]</sup>。采用智能虚拟对手协助飞行员完成对抗训练, 既可以在缺少飞行员陪练的情况下保障训练正常进行, 又能够在很大程度上降低训练成本。近距离空战过程中, 战场态势变化剧烈, 对于飞行员的决策和操作水平要求很高, 如果采用智能虚拟对手模拟飞行员的决策, 则需要其具备实时给出适应态势变化的决策和占位导引指令的能力。因此, 空战智能决策和占位导引技术是智能虚拟对手建模的核心技术, 直接影响智能虚拟对手对于飞行员

思维拟合的逼真性, 从而影响受训人员在空战对抗模拟训练中的真实感体验。

学者们提出了多种飞机空战智能决策和占位导引方法, 按照实现方式主要可以归纳为2类。第1类方法的共有特征是总结出战机可以执行的机动动作集合, 设计空战智能决策方法, 实现适应战场态势变化的机动动作智能调用, 按照机动动作的预定轨迹导引战机实现有利于己方的空间占位。采用的智能决策方法包括专家系统、矩阵对策、影响图、神经网络、近似动态规划、共生生物搜索等。文献[3]根据专家知识设计决策规则, 构建了飞机机动动作库, 实现了典型态势下的机动动作智能决策; 文献[4]采用矩阵对策法进行空

收稿日期: 2020-12-08; 录用日期: 2021-02-10; 网络出版时间: 2021-03-01 15:17

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20210301.1328.001.html](http://kns.cnki.net/kcms/detail/11.2625.V.20210301.1328.001.html)

基金项目: 国家自然科学基金(61503255); 沈阳市中青年科技创新人才支持计划(RC180174)

\*通信作者. E-mail: mengguanglei@yeah.net

**引用格式:** 孟光磊, 刘德见, 周铭哲, 等. 近距离空战训练中的智能虚拟对手决策与导引方法[J]. 北京航空航天大学学报, 2022, 48(6): 937-949. MENG G L, LIU D J, ZHOU M Z, et al. Intelligent virtual opponent decision making and guidance method in short-range air combat training[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(6): 937-949 (in Chinese).

战决策,得到的机动策略序列满足分段优化条件;文献[5-9]基于影响图理论建立空战智能决策模型,以逻辑可释的方式实现了空战机动决策,较好地实现了在典型空战态势下对于飞行员决策思维的拟合;文献[10-11]采用神经网络方法进行空战机动决策,该方法需要获得大量包含各种态势的空战样本数据来进行模型训练;文献[12-15]提出了空战机动决策的近似动态规划算法,有效提高了空战决策的实时性;文献[16]用共生生物搜索方法进行飞机空战机动决策,收敛速度与精度较传统智能算法均有改善。上述方法均以机动动作库为基础,采用人工智能方法实现不同态势下的机动决策。机动动作的固化使得导引轨迹缺乏随着态势微变而及时调整的灵活性,成为影响决策逼真性和占位导引优化性的关键因素。第2类方法针对进攻和逃逸态势分别设计导引指标函数,采用最优控制方法进行导引量解算,从而满足空战占位导引的优化性需求。文献[17-19]开展了基于微分对策的追逃博弈导引方法研究,但在求解过程中不能保证鞍点一定存在;文献[20-21]提出应用滚动时域来预测未来一段时间内空战态势的变化,采用最优控制方法求解战机导引量,实现了优化占位导引。该类方法虽然能够克服机动动作固化导致的空间占位导引灵活性和优化性不足的问题,但是在如何根据战场态势自主进行攻击和逃逸决策,以及将决策和导引进行一体化设计方面,没有给出系统的解决方案。

本文针对近距空战模拟训练过程中对于智能虚拟对手决策的逼真性和占位导引的优化性需求,开展智能虚拟对手决策和导引一体化方法研究。在有人机空战过程中,飞行员通过分析各类传感器获取的战场信息,得到当前空战态势,并据此经过层级化推理做出有利于己方的机动决策。本文模拟飞行员的逻辑思维方式,综合考虑影响

近距空战决策的态势因素,基于动态贝叶斯网络(dynamic Bayesian network, DBN)建立智能虚拟对手自主占位决策模型,实现了适应战场态势变化的占位导引优化指标智能决策。根据目标预测轨迹,采用约束梯度法实时解算导引指标泛函,得到优化占位导引量,使智能虚拟对手能够做出有利于己方的优化机动。本文方法克服了传统方法实现机动动作固化、对抗表现缺乏灵活性的问题,实现了决策和导引的无缝结合。仿真结果表明,本文方法能够使智能虚拟对手实现合理决策和占位导引,快速达到攻击条件或者规避威胁,具备较好的实时性和优化性,有较好的工程应用价值。

## 1 智能虚拟对手自主占位决策模型

在空战仿真对抗训练过程中,智能虚拟对手自主占位决策模型的作用是替代飞行员给飞行模拟器的导引解算系统下达占位决策指令。战场环境中存在电子干扰,或受探测设备性能影响,因此,飞行员常常需要针对不确定性信息进行决策。如何真实模拟飞行员根据战场信息进行态势分析和决策推理,是提高智能虚拟对手自主占位决策模型逼真性的关键。

DBN 适用于表达和分析前后时刻具有关联关系的不确定性和概率性事件,可以从不完全或不确定的知识或信息中做出推理<sup>[22]</sup>。本文借鉴飞行员的层级化推理过程,根据领域专家经验,采用 DBN 建立近距空战自主占位决策模型。首先,综合分析影响空战决策的态势因素,明确各因素之间的因果关系,从而设计整个模型的网络结构,如图 1 所示。然后,根据领域先验知识,确定网络参数,包括网络模型有向弧概率、状态转移概率和网络节点先验分布。最后,完成基于 DBN 的近距空战自主占位决策模型构建。

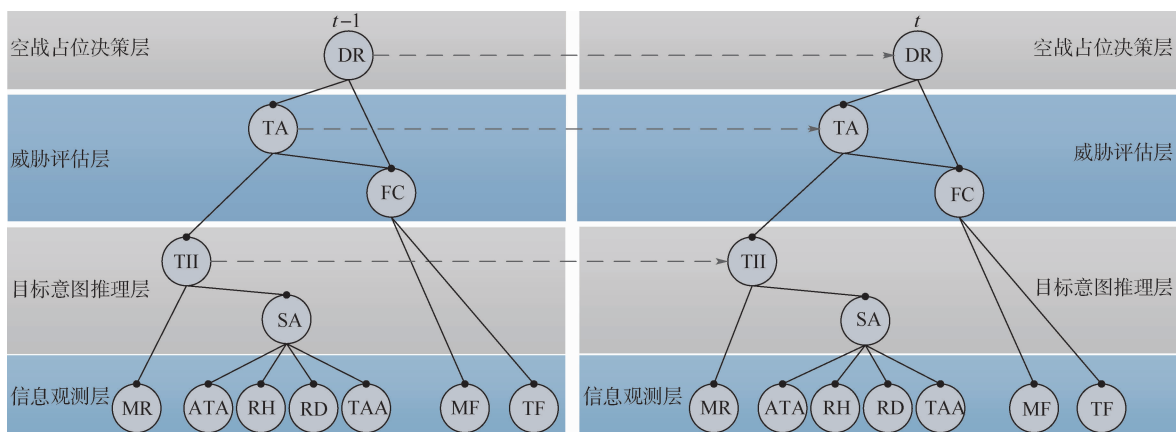


图 1 近距空战自主占位决策 DBN 模型

Fig. 1 Short-range air combat autonomous occupancy decision DBN model

DBN 模型共有 4 层,底层为信息观测层,采集目标机动动作、目标方位角、目标进入角、相对高度、目标距离、双方火控攻击区等观测信息,为层次化的空战决策过程提供战场态势信息输入。其中,MR 为目标机动动作识别网络模型的输出节点,识别网络的构建参照文献[23]的方法,此处不做详述,其识别结果为目标意图推理提供了决策依据。目标意图推理层先对敌我双方的空间相对位置观测信息进行评估,得到目标相对于我机空间占位的综合判断,再结合目标机动识别结果推断目标在对抗过程中采用的策略。威胁评估层先根据火控攻击区信息得到我机对于目标的攻击和规避条件判断,再结合目标意图推理结果评估目标对于我机的威胁程度。

战机火控攻击区域如图 2 所示。目标相对火控攻击区域的位置关系包括 3 种情况:攻击区域内、不可逃逸区和攻击区域外。信息观测层、目标意图推理层和威胁评估层的各节点定义及状态集划分如表 1 所示。空战占位决策层是该模型的顶层,综合威胁评估结果和攻击/规避条件评估结果推理我方应采用的优化占位导引指标。

在近距空战过程中,飞行员根据先验知识对获取的目标感知信息进行综合评判,得到关于己方是否处于目标火控攻击区域的判断结果。智能虚拟对手决策模型中,TF(目标火控信

息)节点的状态值是通过拟合飞行员的决策思维,根据先验知识设定攻击区域、不可逃逸区的边界条件阈值,结合获取的目标方位角、目标进入角和目标距离等信息,进行基于规则的条件推理得到的。

在近距空战中,进行尾后攻击是飞行员或战机的理想攻击状态,而机尾遭到目标攻击则是最糟糕的情况。在同有人机飞行员进行空战仿真对抗过程中,对智能虚拟对手实施导引的目标是通过调节智能虚拟对手与目标之间的角度和距离,在有利态势下,使机载武器进入可发射区,在不利态势下,破坏目标武器的发射条件。近距空战自

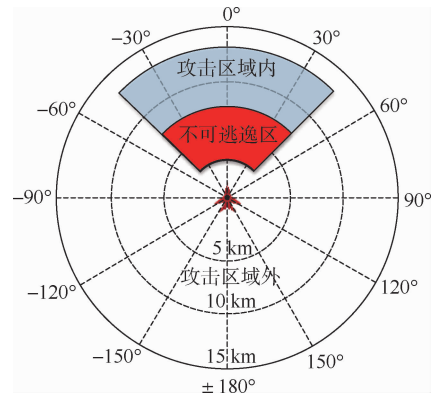


图 2 近距空战火控攻击区域包络  
Fig. 2 Fire control attack area envelope in short-range air combat

表 1 近距空战自主占位决策模型节点及状态集定义

Table 1 Description of node and state set of short-range air combat autonomous occupancy decision model

节点	状态集
ATA(目标方位角)	左前方(30°~90°)、右前方(-90°~-30°)、前方(-30°~30°)、后方(150°~180°、-180°~-150°)、左后方(90°~150°)、右后方(-150°~-90°)
TAA(目标进入角)	左前方(30°~90°)、右前方(-90°~-30°)、前方(-30°~30°)、后方(150°~180°、-180°~-150°)、左后方(90°~150°)、右后方(-150°~-90°)
RH(相对高度)	高于目标、基准面、低于目标
RD(目标距离)	近距(<10 km)、中距(10~20 km)、远距(>20 km)
MF(我方火控信息)	目标在攻击区域内、目标在不可逃逸区、目标在攻击区域外,角度条件不满足、目标在攻击区域外,距离条件不满足、目标在攻击区域外,角度和距离条件均不满足
TF(目标火控信息)	我机在攻击区域内、我机在不可逃逸区、我机在攻击区域外,角度条件不满足、我机在攻击区域外,距离条件不满足、我机在攻击区域外,角度和距离条件均不满足
MR(目标机动识别)	直线飞行、跃升、俯冲、筋斗、半筋斗翻转、蛇形机动、左盘旋、右盘旋、左战斗转弯、右战斗转弯
SA(目标空间占位)	$\{(\alpha, t, h, d)   \alpha \in \{ATA\}, t \in \{TAA\}, h \in \{RH\}, d \in \{RD\}\}$ 。其中, $\alpha$ 为目标方位角, $t$ 为目标进入角, $h$ 为相对高度, $d$ 为目标距离
TII(目标意图推理)	左后方攻击、右后方攻击、左侧方绕飞攻击、右侧方绕飞攻击、左前方逃逸、右前方逃逸、左后方分离、右后方分离
TA(威胁评估)	进攻、规避
FC(攻击和规避条件评估)	攻击角度与攻击距离均满足、攻击角度满足但攻击距离需调整、攻击距离满足但攻击角度需调整、攻击角度与攻击距离均需调整、规避角度与规避距离均满足、规避角度满足但规避距离需调整、规避距离满足但规避角度需调整、规避角度与规避距离均需调整

主占位决策 DBN 模型的决策节点输出的是不同态势下智能虚拟体应采用的占位导引指标泛函,是综合考虑空间占位、目标意图、威胁评估及火控信息等因素后,得到的优化决策结果。

网络输出的决策节点状态集定义如表 2 所示。ATK1 ~ ATK3 为目标处于劣势条件下,智能

虚拟对手采取进攻策略时选择的占位导引指标,目的是导引智能虚拟对手达到武器攻击条件;AVD1 ~ AVD3 为目标处于优势条件下,智能虚拟对手采取规避策略时选择的占位导引指标,目的是导引智能虚拟对手通过调整距离和角度,规避敌方攻击,破坏敌机武器攻击条件。

表 2 决策节点 DR 的定义

Table 2 Decision node specification of DR

节点	状态集	备注
DR (决策结果)	ATK1(距离优先攻击指标) $J_{\min} = \int_{t_0}^{t_f} \frac{ \mathbf{r}_{AB} }{D_{\max}} dt$	智能虚拟对手具有攻击优势且攻击角度条件满足,选择距离优先攻击指标,调整两机距离达到尾后攻击条件
	ATK2(角度优先攻击指标) $J_{\min} = \int_{t_0}^{t_f} (\cos(\pi - \alpha) - \cos \beta) dt$	智能虚拟对手具有攻击优势且攻击距离条件满足,选择角度优先攻击指标,调整目标方位角 $\alpha$ 和进入角达到尾后攻击条件
	ATK3(绕飞攻击指标) $J_{\min} = \int_{t_0}^{t_f} \left[ M \left( \frac{ \mathbf{r}_{AB}  - D_{\exp}}{D_{\max}} \right)^2 + N(\cos(\pi - \alpha) - \cos \beta) \right] dt$	双方均势且攻击角度、距离条件均不满足,采用绕飞攻击指标,同时调整角度和距离达到尾后攻击条件
	AVD1(角度优先规避指标) $J_{\min} = \int_{t_0}^{t_f} (1 - \cos \beta) dt$	目标具有攻击优势且满足攻击角度条件,破坏角度条件,调整目标进入角以躲避目标攻击
	AVD2(距离优先规避指标) $J_{\min} = \int_{t_0}^{t_f} \left( \frac{ \mathbf{r}_{AB}  - D_{\exp}}{D_{\max}} \right)^2 dt$	目标具有攻击优势且满足攻击距离条件,破坏距离条件,增大距离以躲避目标攻击
	AVD3(距离和角度同时规避指标) $J_{\min} = \int_{t_0}^{t_f} \frac{E}{ \mathbf{r}_{AB} } + F(1 - \cos \beta) dt$	目标具有攻击优势但攻击角度、距离条件均不满足,同时破坏角度、距离条件,调整距离及目标进入角来躲避目标攻击

表 2 中:  $J_{\min}$  为所建立的优化导引指标泛函,用于求解满足期望条件时的优化导引量;  $D_{\exp}$  为智能虚拟对手导引飞行过程中期望同目标保持的距离;  $D_{\max}$  为对于双机博弈占位有影响的两机最大距离阈值,可以根据领域专家经验,参考最大攻击区距离进行设定;  $M$ 、 $N$ 、 $E$ 、 $F$  为权重系数,用于选择调节角度或距离的优先级,且  $M + N = 1$ ,  $E + F = 1$ ;  $\beta$  为目标进入角的补角,  $\cos \beta$  的计算如下:

$$\cos \beta = \frac{\mathbf{r}_{AB} \cdot \mathbf{V}_B}{|\mathbf{r}_{AB}| \cdot |\mathbf{V}_B|} \quad (1)$$

式中:  $\mathbf{V}_B$  为目标速度矢量;  $\mathbf{r}_{AB}$  为两机距离矢量。

根据  $D_{\max}$  对  $\mathbf{r}_{AB}$  进行限值处理得到  $\bar{\mathbf{r}}_{AB}$  为

$$\bar{\mathbf{r}}_{AB} = \begin{cases} |\mathbf{r}_{AB}| & |\mathbf{r}_{AB}| < D_{\max} \\ D_{\max} & |\mathbf{r}_{AB}| \geq D_{\max} \end{cases} \quad (2)$$

红方为目标机,蓝方为智能虚拟对手,变量具体含义如图 3 所示。

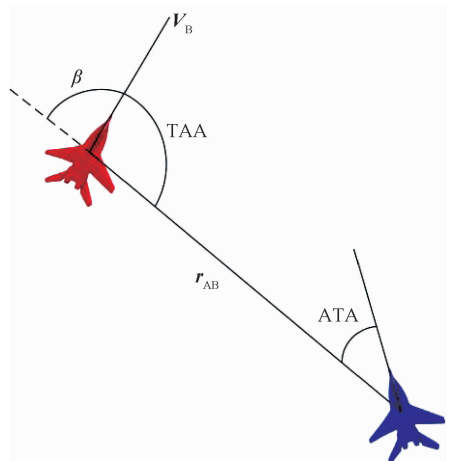


图 3 空间态势变量定义

Fig. 3 Space situation variable specification

DBN 推理算法的实现步骤如下:

**步骤 1** 计算当前时刻的性能指标决策概率。由当前时刻观测节点的特征提取结果和各条

链路的条件概率完成各观测节点到根节点的推理,再将所有链路推理结果相乘得到当前时刻的性能指标决策概率分布。

$$p(\text{DR}_t | e_t) = p(\text{DR}_t | \text{MR}_t, \text{ATA}_t, \text{RH}_t, \text{RD}_t, \text{TAA}_t, \text{MF}_t, \text{TF}_t) = p(\text{DR}_t | \text{MR}_t) p(\text{DR}_t | \text{ATA}_t) \cdot p(\text{DR}_t | \text{RH}_t) p(\text{DR}_t | \text{RD}_t) p(\text{DR}_t | \text{TAA}_t) \cdot p(\text{DR}_t | \text{MF}_t) p(\text{DR}_t | \text{TF}_t) \quad (3)$$

式中:  $e_t$  为动态贝叶斯网模型与时间相关联。

**步骤2** 计算当前时刻决策概率。由于当前时刻的最终决策结果受上一时刻的决策结果影响,根据贝叶斯公式,可求解条件概率分布。

$$p(e_t | \text{DR}_t) = \frac{p(\text{DR}_t | e_t) p(\text{MR}_t, \text{ATA}_t, \text{RH}_t, \text{RD}_t, \text{TAA}_t, \text{MF}_t, \text{TF}_t)}{p(\text{DR}_t)} \quad (4)$$

根据式(3)和式(4)计算结果,可求解各性能指标在综合历史证据信息下的决策概率。

$$p(\text{DR}_t | e_{1:t}) = \frac{p(e_t | \text{DR}_t) \sum_{\text{DR}_{t-1}} p(\text{DR}_t | \text{DR}_{t-1}) p(\text{DR}_{t-1} | e_{1:t-1})}{\sum_{\text{DR}_t} p(e_t | \text{DR}_t) \sum_{\text{DR}_{t-1}} p(\text{DR}_t | \text{DR}_{t-1}) p(\text{DR}_{t-1} | e_{1:t-1})} \quad (5)$$

**步骤3** 判断收敛性。当某性能指标的决策概率达到90%,且观测节点特征提取结果不变时,判断网络处于收敛状态。

**步骤4** 输出决策结果。最终得到的决策结果为推理概率最大值对应的导引性能指标。若接收到仿真结束指令,则停止计算,否则返回步骤1

进行下一解算周期推理。

## 2 目标轨迹预测模型

在近距空战智能虚拟对手导引过程中,考虑目标轨迹变化对战场态势演变的影响。采用文献[23]方法进行目标机动动作在线识别,针对每类机动动作,建立飞行轨迹预测模型。在每个预测周期起始时刻,采用从空战对抗仿真系统获取的飞行状态信息,对目标位置、速度、航向、高度、转弯角速度等信息进行初始化。

### 2.1 直线类机动

直线类机动包括跃升、俯冲、水平直线飞行等机动。根据目标当前时刻的状态信息预测目标未来一段时间之内的运动轨迹,选择状态向量  $\mathbf{X}(k) = [x(k), y(k), z(k), \dot{x}(k), \dot{y}(k), \dot{z}(k)]^T$ 。则直线类机动轨迹预测模型状态方程为

$$\mathbf{X}(k+1) = \mathbf{F}_{\text{CA}}(k) \mathbf{X}(k) \quad (6)$$

式中:  $x(k)$ 、 $y(k)$  和  $z(k)$  为目标在地面坐标系的三轴位置信息;  $\dot{x}(k)$ 、 $\dot{y}(k)$  和  $\dot{z}(k)$  为目标三轴速度信息;  $\mathbf{F}_{\text{CA}} = \{\mathbf{A}, \mathbf{A}\}$ ,  $\mathbf{A}_{3 \times 3}$  为牛顿矩阵。

### 2.2 转弯类机动

转弯类机动包括左/右盘旋机动、筋斗机动、战斗转弯等。该模型将飞机运动状态描述为平面的匀速圆周运动,通过飞行角速度和飞行速度之间的运动关系来描述飞机的运动状态。选取状态向量  $\mathbf{X}(k) = [x(k), y(k), z(k), \dot{x}(k), \dot{y}(k), \dot{z}(k)]^T$ , 转弯机动过程模型可表示为

$$\begin{bmatrix} x(k+1) \\ y(k+1) \\ z(k+1) \\ \dot{x}(k+1) \\ \dot{y}(k+1) \\ \dot{z}(k+1) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \frac{\sin(\omega_1 T)}{\omega_1} + \frac{\sin(\omega_2 T)}{\omega_2} & -\frac{1 - \cos(\omega_1 T)}{\omega_1} & -\frac{1 - \cos(\omega_2 T)}{\omega_2} \\ 0 & 1 & 0 & \frac{1 - \cos(\omega_1 T)}{\omega_1} & \frac{\sin(\omega_1 T)}{\omega_1} + \frac{\sin(\omega_3 T)}{\omega_3} & -\frac{1 - \cos(\omega_3 T)}{\omega_3} \\ 0 & 0 & 1 & \frac{1 - \cos(\omega_2 T)}{\omega_2} & \frac{1 - \cos(\omega_3 T)}{\omega_3} & \frac{\sin(\omega_2 T)}{\omega_2} + \frac{\sin(\omega_3 T)}{\omega_3} \\ 0 & 0 & 0 & \cos(\omega_1 T) + \cos(\omega_2 T) & -\sin(\omega_1 T) & -\sin(\omega_2 T) \\ 0 & 0 & 0 & \sin(\omega_1 T) & \cos(\omega_1 T) + \cos(\omega_3 T) & -\sin(\omega_3 T) \\ 0 & 0 & 0 & \sin(\omega_2 T) & \sin(\omega_3 T) & \cos(\omega_2 T) + \cos(\omega_3 T) \end{bmatrix} \begin{bmatrix} x(k) \\ y(k) \\ z(k) \\ \dot{x}(k) \\ \dot{y}(k) \\ \dot{z}(k) \end{bmatrix} \quad (7)$$

式中:  $\omega_1$ 、 $\omega_2$ 、 $\omega_3$  分别为飞机绕机体  $z$ 、 $y$ 、 $x$  轴运动的角速度,当  $\omega_1 > 0$ 、 $\omega_2 > 0$ 、 $\omega_3 > 0$  时,飞机分别绕对应轴做顺时针转动,当  $\omega_1 < 0$ 、 $\omega_2 < 0$ 、 $\omega_3 < 0$  时,飞机分别绕对应轴做逆时针转动;  $T$  为采样周期。

### 2.3 蛇形机动

蛇形机动的轨迹可描述为飞机在地面坐标系下按正弦轨迹运动,如图4所示。

根据蛇形机动动作的特性,选取状态向量  $\mathbf{X}(k) = [x(k), y(k), z(k), \dot{x}(k), \dot{y}(k), \dot{z}(k)]^T$ ,

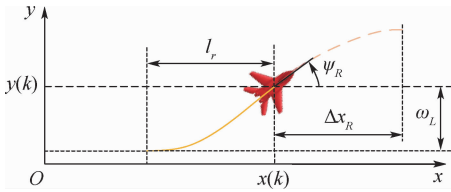


图4 蛇形机动轨迹预测模型

Fig. 4 Snaking maneuver trajectory prediction model

则蛇形机动轨迹预测模型的状态方程可表示为

$$\begin{cases} x(k+1) = x(k) + \dot{x}(k)T \\ y(k+1) = \frac{\omega_L}{2} \sin\left(\frac{\pi}{l_r} \Delta x_R - \frac{\pi}{2}\right) + y(k) \\ z(k+1) = z(k) \\ \dot{x}(k+1) = \sin(\psi_R T) \dot{x}(k) \\ \dot{y}(k+1) = \cos(\psi_R T) \dot{z}(k) \\ \dot{z}(k+1) = 0 \end{cases} \quad (8)$$

式中:  $\Delta x_R$  为蛇形机动未来时刻横向偏移量;  $\omega_L$  和  $l_r$  分别为蛇形机动识别周期起始至结束时刻的纵向偏移量和横向偏移量;  $\psi_R$  为飞机航向角。

## 2.4 目标轨迹合成模型

综合目标轨迹预测子模型, 选取状态向量  $\mathbf{X}(k) = [x(k), y(k), z(k), \dot{x}(k), \dot{y}(k), \dot{z}(k)]^T$ , 对应  $j \in \{1, 2, \dots, r\}$  类机动动作, 结合目标机动动作识别概率  $p(j)$ , 得到目标轨迹预测综合模型为

$$\mathbf{X}(k+1) = \sum_{j=1}^r \mathbf{X}_j(k+1) p(j) \quad (9)$$

## 2.5 目标轨迹预测误差

为验证轨迹预测模型的准确性, 将每个时间点的轨迹预测值与实际飞行轨迹进行误差计算。对预测轨迹进行等间隔采样, 外推时域为  $T_{re}$ , 则采样数量为  $L = \lceil T_{re}/T \rceil$ , 得到每个采样点的目标预测轨迹与实际飞行轨迹的均方根误差为

$$RMSE = \frac{\sum_{i=1}^L \sqrt{(x'_i - x)^2 + (y'_i - y)^2 + (z'_i - z)^2}}{L} \quad (10)$$

式中:  $(x, y, z)$  为采样点真实位置;  $(x'_i, y'_i, z'_i)$  为采样点预测位置信息。

## 3 基于约束梯度法的优化占位导引量解算

针对智能虚拟对手自主占位决策模型得到的占位导引指标泛函, 采用约束梯度法解算优化占位导引量。核心思想是: 对于微分方程形式的飞机运动学模型, 给定初始状态、横截条件和所能达到的性能边界状态, 计算在某一段时间内, 能够使性能指标泛函达到极小值的导引量。

## 3.1 飞机运动学模型

采用飞机运动模型如下<sup>[17]</sup>:

$$\dot{x} = V \cos \gamma \cos \chi \quad (11)$$

$$\dot{y} = V \cos \gamma \sin \chi \quad (12)$$

$$\dot{z} = V \sin \gamma \quad (13)$$

$$\dot{V} = \frac{\eta P - D}{m} - g \sin \gamma \quad (14)$$

$$\dot{\gamma} = \frac{g}{V} (n_z \cos \mu - \cos \gamma) \quad (15)$$

$$\dot{\chi} = \frac{g}{V \cos \gamma} n_z \sin \mu \quad (16)$$

式中:  $\dot{x}, \dot{y}, \dot{z}$  为飞机在地面坐标系下各轴向的位置变化率;  $\dot{\gamma}, \dot{\chi}$  分别为航迹倾角和航迹偏角的变化率;  $V$  为速度大小;  $P$  为飞机所能产生的最大推力;  $\eta$  为推力系数, 取值在  $0 \sim 1$  之间;  $D$  为空气阻力;  $\mu$  为滚转角;  $m$  为机体质量;  $g$  为当地重力加速度;  $n_z$  为法向过载, 纵向过载  $U_1 = n_z \cos \mu$ , 横向过载  $U_2 = n_z \sin \mu$ 。选取状态向量  $\mathbf{X} = [x, y, z, V, \gamma, \chi]^T$ , 控制向量  $\mathbf{U} = [\eta, U_1, U_2]^T$ 。

## 3.2 飞行性能约束算子

为了在空战训练过程中, 逼真模拟飞机的飞行性能, 需要针对飞机运动学模型, 设计飞行性能约束算子, 包括速度大小  $V$ 、速度变化率  $\Delta V$ 、推力系数  $\eta$ 、法向过载  $n_z$  及滚转角变化率  $\Delta \mu$ , 具体参数取值范围设定如下:

$$\begin{cases} V_{\min} \leq V \leq V_{\max} \\ |\Delta V| \leq \Delta V_{\max} \\ 0 \leq \eta \leq 1 \\ |n_z| \leq n_{z \max} \\ |\Delta \mu| \leq \Delta \mu_{\max} \end{cases} \quad (17)$$

## 3.3 优化占位导引量解算

根据飞机运动学模型和目标轨迹预测结果, 求解能够使占位导引指标泛函最小的飞机导引量, 属于最优控制问题, 计算量较大, 为满足实时性需求, 本文采用约束梯度法<sup>[24]</sup>进行解算, 具体流程如图5所示。

首先, 根据状态方程建立哈密顿函数, 求解协态方程。然后, 计算状态变量在  $t_0 \sim t_f$  内的积分, 由目标的状态预测信息求出协态方程在  $t_0 \sim t_f$  内的反向积分。目标状态信息可由第2节中的轨迹预测模型预测得出, 但由于智能虚拟对手状态方程较为复杂, 直接积分解算会有很大困难。本文采用龙格库塔法来解算状态方程顺向积分和协态方程反向积分的数值解。求出梯度向量, 对控制向量进行修正。应用约束梯度法解算过程中, 在每个迭代周期, 通过约束算子将求解出的状态量

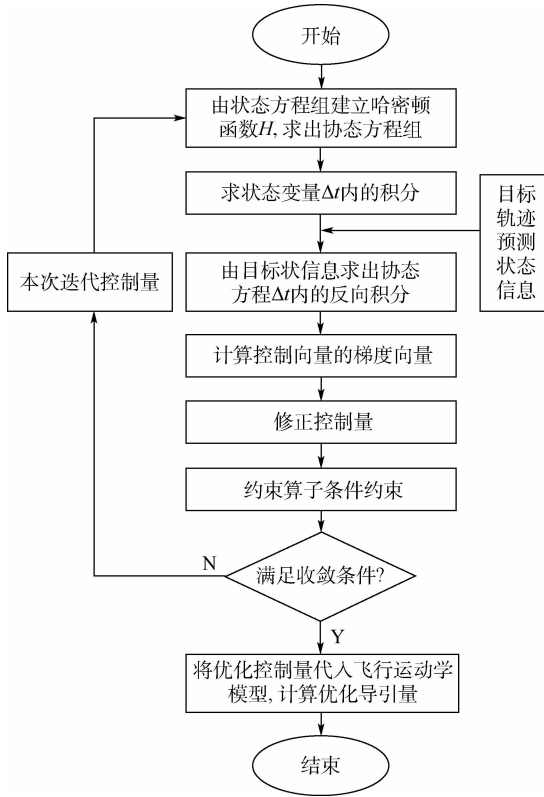


图 5 约束梯度法解算流程

Fig. 5 Solution flow of constrained gradient method

和控制向量限制到合理区域内。重复上述过程进行迭代计算, 收敛条件为本次输出的控制向量与上次控制向量之差的绝对值小于 0.01 或已迭代 100 次。在满足收敛条件后, 将控制向量代入飞机运动学模型计算优化导引量, 包括航迹偏角、航迹倾角和速度, 实现智能虚拟对手的优化占位导引。

### 3.4 优化占位导引量解算步骤

**步骤 1** 建立哈密顿函数。联合无人机的状态方程和优化指标泛函, 建立对应该指标的哈密顿函数  $H$ :

$$H = J_{\min} + \lambda_1 V \cos \gamma \cos \chi + \lambda_2 V \cos \gamma \sin \chi + \lambda_3 V \sin \gamma + \lambda_4 \left( \frac{P - D}{m} - g \sin \gamma \right) + \lambda_5 \frac{g}{V} (n_z \cos \mu - \cos \gamma) + \lambda_6 \frac{g}{V \cos \gamma} n_z \sin \mu \quad (18)$$

**步骤 2** 积分状态方程。给定一个初始控制向量  $\mathbf{U}^k(t_0) = [\eta^k(t_0), U_1^k(t_0), U_2^k(t_0)]^T$  和初始状态向量信息  $\mathbf{X}(t_0) = [x(t_0), y(t_0), z(t_0), V(t_0), \gamma(t_0), \chi(t_0)]^T$ , 通过四阶龙格库塔法在待求时间  $t_0 \sim t_f$  内顺向积分状态方程, 求出状态向量  $\mathbf{X}(t_f)$ 。

**步骤 3** 求解协态方程。结合哈密顿函数  $H$ , 对每个状态变量分别求偏导, 得到每个状态变量所对应的协态方程:

$$\begin{cases} \dot{\lambda}_1 = -\frac{\partial H}{\partial x} \\ \dot{\lambda}_2 = -\frac{\partial H}{\partial y} \\ \dot{\lambda}_3 = -\frac{\partial H}{\partial z} \\ \vdots \\ \dot{\lambda}_6 = -\frac{\partial H}{\partial \chi} \end{cases} \quad (19)$$

**步骤 4** 求解协态向量。将横截条件求得的终端值  $\boldsymbol{\lambda}(t_f)$ 、状态向量  $\mathbf{X}(t_f)$  和控制向量  $\mathbf{U}^k(t_f)$ , 通过四阶龙格库塔法从待求时间  $t_f$  到  $t_0$  逆向积分协态方程, 求出协态向量  $\boldsymbol{\lambda}(t)$ , 其中协态方程目标位置信息由轨迹预测给出。

**步骤 5** 求解梯度向量。计算哈密顿函数  $H$  对控制向量  $\mathbf{U}$  的梯度向量  $\mathbf{g}^k$ 。

$$\mathbf{g}^k \triangleq \left( \frac{\partial H}{\partial \mathbf{U}} \right)_k \quad (20)$$

**步骤 6** 修正控制向量。用一维寻优法 (0.618 法) 确定步长因子  $\tau$ , 修正控制向量:

$$\mathbf{U}^{k+1}(t) = \mathbf{U}^k(t) - \tau \mathbf{g}^k \quad (21)$$

**步骤 7** 循环输出控制向量, 得到飞机导引量。令  $\varepsilon$  小于某个极小值, 计算控制向量  $\mathbf{U}^{k+1}(t)$  是否满足下列输出条件, 满足则输出, 否则跳回步骤 2, 使用控制向量  $\mathbf{U}^{k+1}(t)$  作为下一次解算的初始控制向量, 以此类推。

$$\left| \frac{J(\mathbf{U}^{k+1}) - J(\mathbf{U}^k)}{J(\mathbf{U}^k)} \right| < \varepsilon \quad (22)$$

$$\|\mathbf{g}^k\| < \varepsilon \quad (23)$$

将控制向量代入飞机运动学模型得到导引量, 包括飞机速度大小  $V$ 、飞机航迹倾角  $\gamma$ 、飞机航迹偏角  $\chi$ 。

## 4 仿真分析

开展红蓝双方的近距离自由空战仿真, 对本文方法进行验证。红蓝双方采用相同的动力学模型, 具有相同的机动能力。蓝方智能虚拟对手采用本文设计的算法进行智能决策和优化占位导引; 红方为有人机, 在操纵杆的控制下机动飞行。

### 4.1 双机侧方态势下的近距离空战仿真

#### 4.1.1 仿真参数设置

飞机运动学模型参数设置为: 质量为 1 000 kg, 最大推力为 10 000 N, 最大速度为 320 m/s, 最大法向过载与最大纵向过载均为 10  $g$ 。红蓝双方的初始状态信息如表 3 所示。仿真系统硬件配置为处理器 Intel(R) Core(TM) i7-6700 CPU@ 3.4 GHz, 内存 (RAM) 16.0 GB。

表3 红蓝双方战机初始状态信息

Table 3 Initial situation information of red and blue fighters

战机	x 坐标/ km	y 坐标/ km	高度/ km	空速/ ( $m \cdot s^{-1}$ )	航向/ ( $^{\circ}$ )
红方	10	0	6	250	0
蓝方	10	-10.2	7	200	0

4.1.2 侧方态势下双机对抗轨迹

完成仿真参数配置后,启动仿真实验。在自由空战仿真过程中,红蓝双方不断进行机动变换以占据有利攻击位置,侧方态势下双机对抗轨迹如图6所示。飞行轨迹反映出了对抗过程中双方运动状态的变化情况,蓝方智能虚拟对手在智能决策引导下,能快速接近红方并持续占据红方后方有利攻击位置,构成攻击条件。

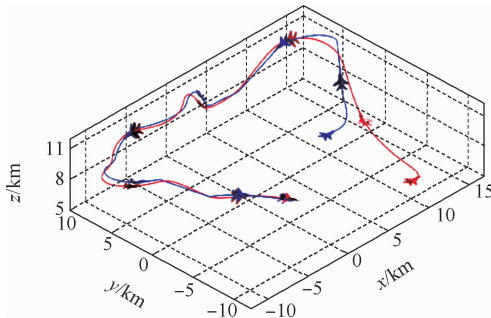


图6 侧方态势下近距离格斗对抗轨迹  
Fig. 6 Short-range combat confrontation trajectory under lateral status

4.1.3 目标机动识别结果及轨迹预测误差分析

空战对抗过程中,智能虚拟对手根据目标机动动作识别结果进行目标意图推理,进而做出适

应战场态势变化的占位决策。参照文献[23]方法构建目标机动动作识别模型,在每个仿真周期输出识别概率最大的机动动作作为识别结果,识别概率随时间变化的过程如图7所示。机动动作编号1~10分别对应水平直线飞行、跃升、俯冲、左战斗转弯、右战斗转弯、左盘旋、右盘旋、筋斗、半筋斗、蛇形机动。为便于观察被识别机动动作的切换变化过程,时间坐标轴采用变比例缩放,省略了机动动作识别概率持续保持大于0.9的部分。对目标机动动作识别结果进行分析可知,当红方做出战术机动,飞行参数的特征提取结果发生变化时,机动识别网络能够迅速收敛到正确的识别结果。

蓝方智能虚拟对手对红方机动动作识别的时序图如图8所示,输出结果为每个机动动作识别周期概率最大的机动动作。该图描述了整个空战过程中,智能虚拟对手识别的红方机动变化,为空战占位决策提供了可靠依据。

根据机动动作识别概率结果,蓝方智能虚拟对手对红方进行轨迹预测,预测时间外推5 s,采样间隔为50 ms,根据机动动作识别概率结果,按照式(9)进行目标预测轨迹合成。采用2.5节方法计算轨迹预测均方根误差,计算结果如图9所示。当红方做大机动时,轨迹预测精度会有所下降,但在整个空战过程中,平均轨迹预测均方根误差为21.65 m,可为智能虚拟对手优化占位导引量解算提供较为准确的目标预测轨迹。

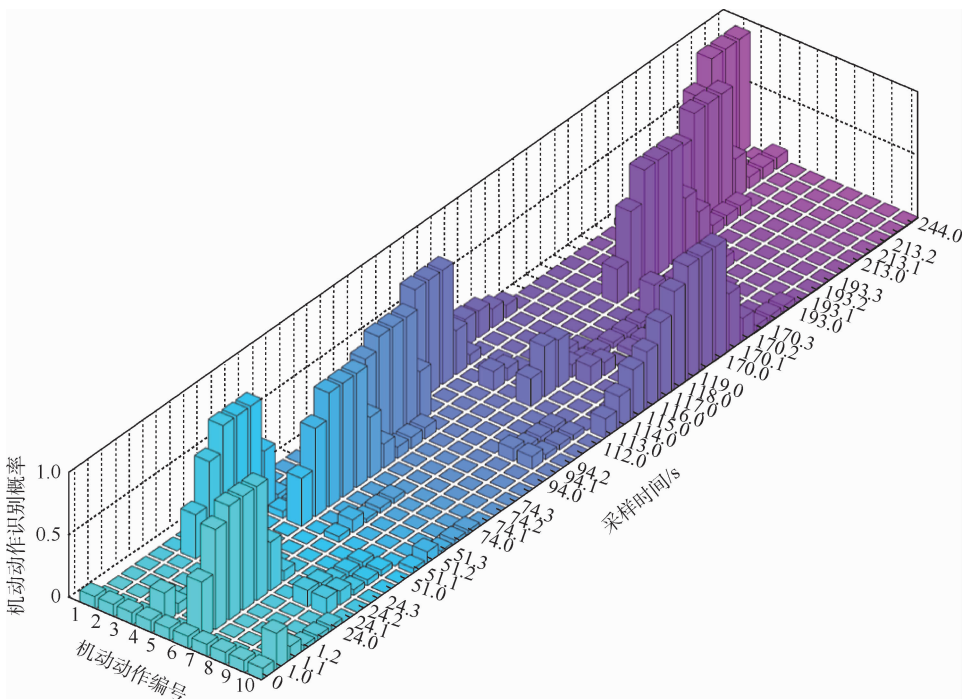


图7 目标识别过程机动动作概率

Fig. 7 Probability of maneuver in target recognition process

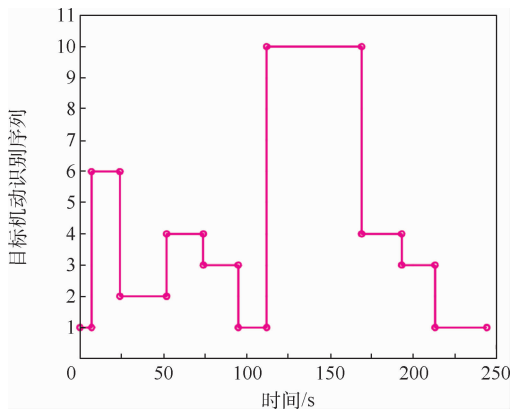


图 8 目标机动识别时序图

Fig. 8 Sequence diagram of target maneuver recognition

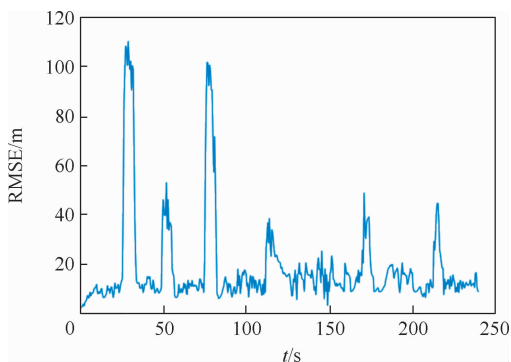


图 9 目标预测轨迹均方根误差

Fig. 9 Root mean square error of target prediction trajectory

#### 4.1.4 空战决策过程及态势分析

在空战仿真对抗训练过程中,智能虚拟对手自主占位决策模型根据空战态势变化,实时决策出己方应该采用的优化占位导引指标泛函,并采

用约束梯度法计算对应的优化占位导引量,实现了空战占位决策与优化占位导引的无缝衔接。在本例中,由近距离空战自主占位决策模型得到的优化导引指标切换 4 次,概率分布如图 10 所示。

为便于分析占位导引的优化程度,对空战过程中智能虚拟对手的占位导引指标值进行规范化处理,使其变化范围为  $[-1, 1]$ , 指标值收敛情况如图 11 所示。空战初始阶段(0 ~ 23 s),红方采用直线飞行,蓝方攻击距离与角度均不满足,此时决策模型迭代输出优化指标泛函 ATK3 来解算导引量,同时调整角度和距离,快速接近红方机尾后,指标值的变化趋势如图 11 中的红色线所示,从中可以看到指标值逐渐收敛。在 24 ~ 110 s 时间内,红方依次做出跃升机动、左战斗转弯机动,此时蓝方攻击距离条件不满足,智能虚拟对手输出 ATK1 进行优化占位导引,为调整距离达到尾后攻击条件,蓝方快速爬升,缩小双方距离,重新获得攻击占位优势,指标值逐渐减小。在 111 ~ 170 s 内,红方进行蛇形机动、左战斗转弯等机动以摆脱蓝方追击,蓝方攻击角度条件不满足,智能虚拟对手采用 ATK2 指标泛函解算导引量,优先调整攻击角度,由于红方在蛇形机动过程中不断大范围改变飞行角度,在追击过程中,指标值有所波动,但总体仍呈现出收敛趋势。在近距离空战后期的 171 ~ 237 s 内,红方又进行了左战斗转弯、俯冲等机动以摆脱蓝方的尾追态势,智能虚拟对手采用 ATK1 牢牢追击红方,取得了攻击优势。

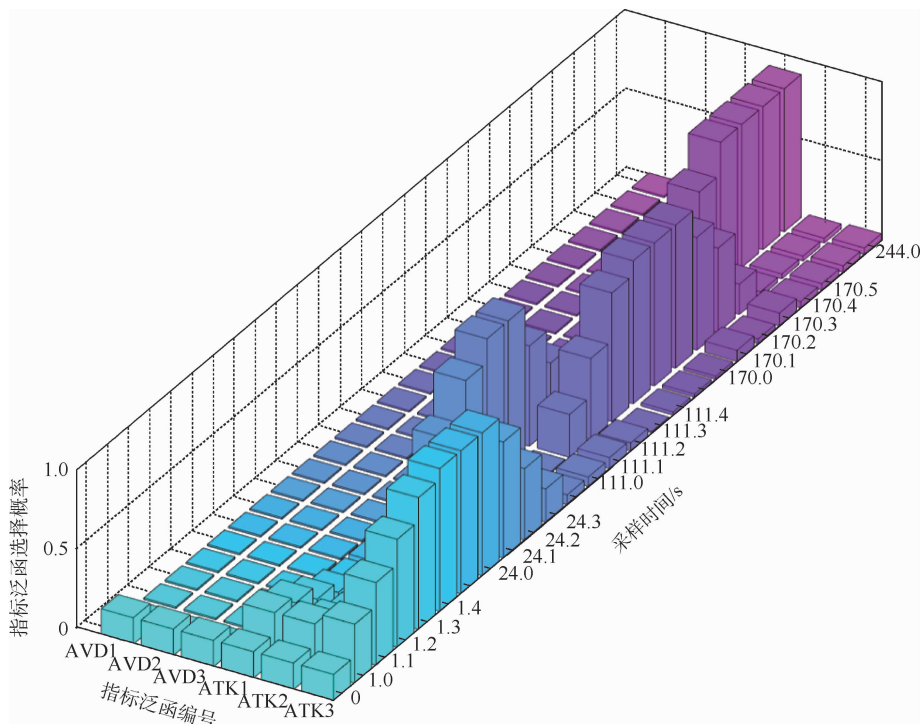


图 10 指标选择过程指标泛函概率

Fig. 10 Index selection process index functional probability

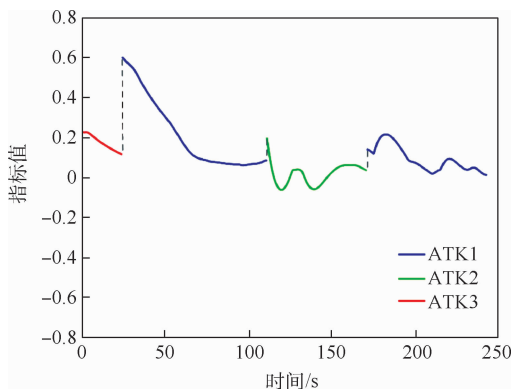


图 11 不同时间段内指标的收敛情况

Fig. 11 Convergence of indicators in different time periods

### 4.1.5 智能虚拟对手控制向量和导引量分析

蓝方智能虚拟对手初始控制向量  $U = [0.6, 0, 0]^T$ , 在空战开始阶段, 蓝方与目标距离较远, 通过接收到的目标状态信息, 由决策网络输出优化指标, 解算优化导引量, 推力系数迅速增至最大; 飞行 24 s 后, 目标做跃升机动, 蓝方纵向过载增大, 进行加速爬升机动, 以获得更好的攻击占位优势; 在空中缠斗阶段, 蓝方预测红方轨迹, 通过改变推力系数大小来调整速度保持追击态势, 改变纵向过载和横向过载大小来控制蓝方机动的航迹倾角和航迹偏角, 使其始终紧紧跟随红方并占据攻击优势。整个近距离格斗空战中的蓝方智能虚拟对手控制向量参数变化趋势如图 12 所示。

在空战仿真过程中的 0 ~ 23 s、24 ~ 110 s、111 ~ 170 s、171 ~ 237 s 各时间段内, 由近距离空战自主占位决策模型迭代输出指标泛函分别为 ATK3、ATK1、ATK2、ATK1, 将输出的指标泛函代入优化占位导引量解算流程得到对应导引量变化趋势, 如图 13 所示。

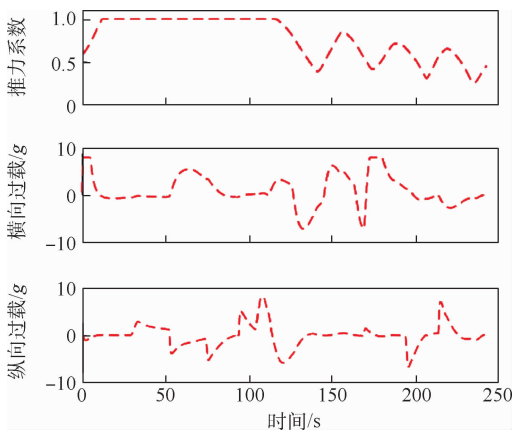


图 12 智能虚拟对手控制向量变化趋势

Fig. 12 Variation trend of intelligent virtual adversary control quantity

### 4.1.6 红蓝双方空战占位分析

在空战仿真过程中, 双方实时距离变化与红

方实时方位角、进入角变化分别如图 14 和图 15 所示。仿真结果表明, 蓝方智能虚拟对手能够有效跟踪红方机动状态变化, 当红方快速逃逸时, 蓝方在决策网络的导引下能快速迭代输出有效的机动优化指标泛函, 通过约束梯度法在短时间内计算优化导引量, 使蓝方迅速减小双方距离, 占据有利空间态势。当红方做出大过载机动时, 蓝方也能快速调整导引指标, 使红方方位角收敛至  $0^\circ$  附近, 进入角保持在正负  $180^\circ$  左右, 保持攻击优势。

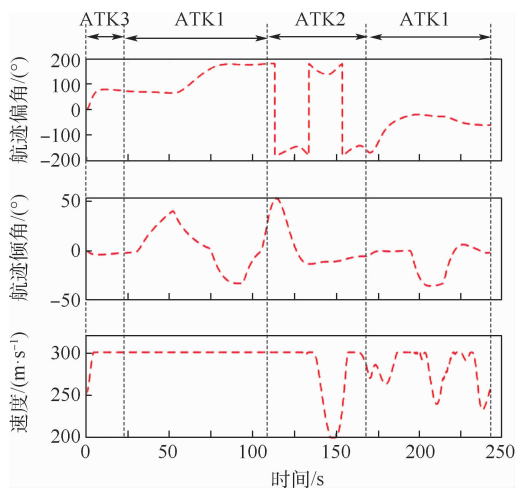


图 13 智能虚拟对手导引量变化趋势

Fig. 13 Variation trend of intelligent virtual adversary guidance quantity

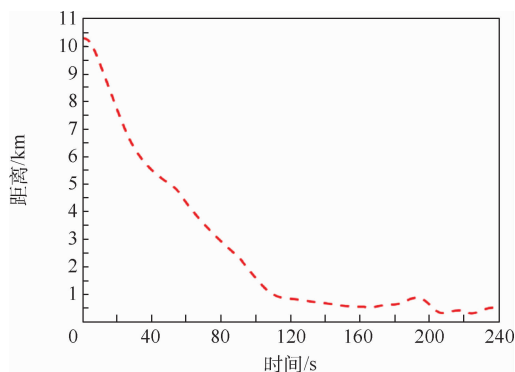


图 14 双机实时距离

Fig. 14 Real time distance of two fighters

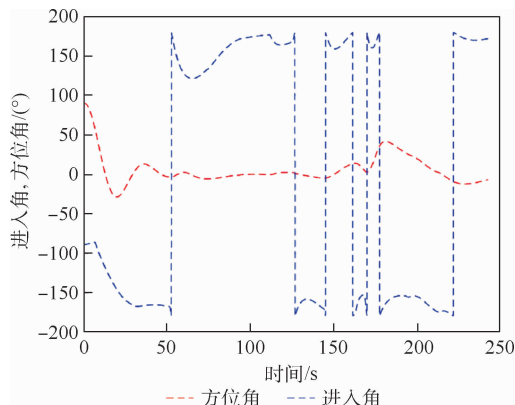


图 15 红方进入角、方位角

Fig. 15 Angle of approach and azimuth of red

### 4.1.7 算法实际应用性能分析

空战仿真过程中,在处理器为 Intel(R) Core(TM) i7-6700 CPU@ 3.4 GHz、内存(RAM)为 16.0 GB、64 位操作系统配置下,开展 DBN 算法和约束梯度法的耗时仿真实验,如图 16 所示。

实验结果表明,整个空战过程中 DBN 算法和约束梯度法在每一个解算周期的耗时均不超过 10 ms,符合工程实际要求。

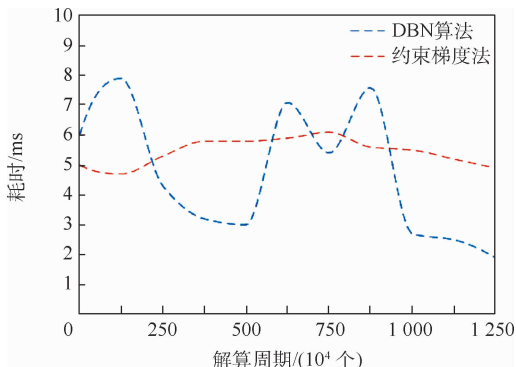


图 16 DBN 算法和约束梯度法耗时

Fig. 16 Elapsed time of DBN and constrained gradient algorithm

## 4.2 双机迎头态势下的近距离空战仿真

### 4.2.1 仿真参数设置

为验证智能虚拟对手算法的合理性与通用性,本文在空战中不同起始态势下进行了多次仿真实验。以迎头态势为例,设置飞机运动学模型参数为:质量为 1 000 kg,最大推力为 10 000 N,最大速度为 320 m/s,红方目标机最大法向过载与最大纵向过载均为 9 g,蓝方智能虚拟对手最大法向过载与最大纵向过载均为 7 g。

### 4.2.2 迎头态势下双机对抗轨迹

飞行轨迹反映出了对抗过程中双方运动状态的变化情况,蓝方智能虚拟对手在智能决策导引下,能快速接近红方并保持占位优势,最终占据红方后方有利攻击位置,构成攻击条件。迎头态势下,基于 DBN 与约束梯度法的双机近距离格斗对抗轨迹如图 17 所示。

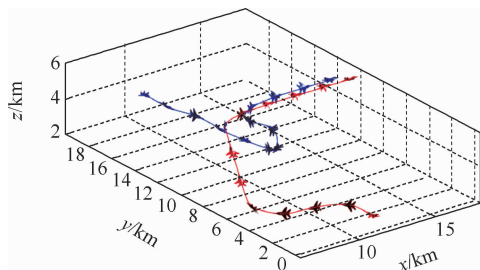


图 17 迎头态势下近距离格斗对抗轨迹

Fig. 17 Dog-fight combat trajectory in head-on position

### 4.2.3 相同条件下的对比实验

在空战仿真参数、双机初始空间占位信息及红方目标机飞行轨迹相同的情况下进行对比实验,基于影响图对策方法<sup>[8]</sup>的双方对抗轨迹如图 18 所示。

整个空战过程中,基于 2 种方法进行智能虚拟对手飞行导引后的双机相对态势信息如图 19 所示。就双方相对距离而言,2 种方法的相对距离变化趋势一致,且本文方法的相对距离缩小趋势更为明显,优化占位导引达到的效果更好。就目标进入角与目标方位角而言,2 种方法最终达到的结果相似,但在整个空战过程中,基于本文方法进行飞行导引后的相对态势更为有利,即目标方位角一直控制在 ±30°之内,目标进入角最终达到 120°左右,满足攻击条件。

实验结果表明,采用本文方法构建的智能虚拟对手模型决策结果符合飞行员逻辑思维,导引轨迹具有较好的优化性,能够满足近距离空战对抗模拟训练要求。

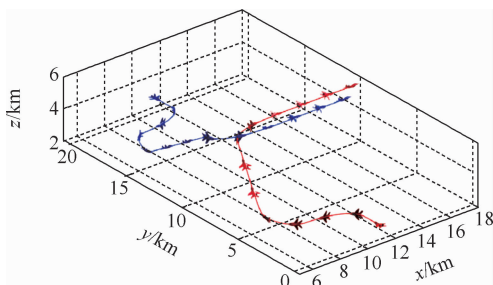


图 18 基于影响图对策方法的双方对抗轨迹

Fig. 18 Trajectories of two opposing parties based on influence diagram game method

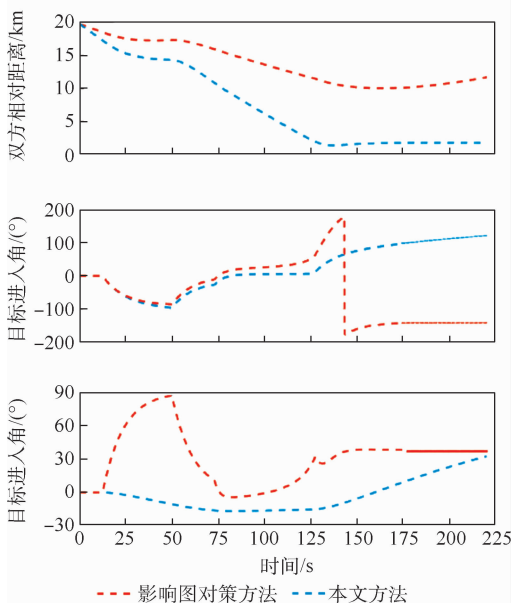


图 19 蓝方智能虚拟对手与红方有人机的相对态势信息

Fig. 19 Relative situation information between blue intelligent virtual adversary and red manned aircraft

## 5 结 论

本文对智能虚拟对手的决策与优化占位导引问题进行了深入研究,仿真实验结果表明,采用本文方法可以实现智能虚拟对手的合理化自主决策和占位导引,为实现智能化的空战模拟训练奠定了基础。

1) 模拟飞行员在近距离空战过程中的决策思维,建立了基于 DBN 的智能虚拟对手自主占位决策模型,实时适应战场态势变化,智能化选取占位导引指标,克服了传统方法机动动作调用方式固化的问题,提高了智能虚拟对手建模的逼真度。

2) 针对在线识别的各类目标机动动作,建立了轨迹预测模型,根据目标轨迹预测结果,采用约束梯度法对占位导引指标进行解算,导引智能虚拟对手在优势条件下能够达到攻击占位,在劣势条件下能够破坏敌方优势条件,规避敌方进攻,具备较好的优化性和实时性,能够满足空战对抗仿真训练的需求。

3) 智能虚拟对手自主占位决策模型中的决策节点间因果关系明晰,具备决策逻辑可释性。由于决策结果为占位导引指标,采用约束梯度法进行解算,可实现高动态战场环境下的优化占位导引。通过上述方式实现了空战决策与占位导引的无缝结合,该方法具备较好的泛化能力,能够满足实际工程应用中的实时性要求,既可以应用于空战模拟训练,也可推广至智能空战决策领域。

### 参考文献 (References)

- [ 1 ] 耿振余,刘思彤,李德龙. 嵌入式空战训练中虚拟智能对手的生成研究[J]. 现代防御技术,2014,42(3):172-177.  
GENG Z Y, LIU S T, LI D L. Generating virtual intelligent adversary in embedded training of air combat counterwork[J]. Modern Defence Technology, 2014, 42(3):172-177 (in Chinese).
- [ 2 ] 陈斌,王江,王阳. 战斗机嵌入式训练系统中的智能虚拟陪练[J]. 航空学报,2020,41(6):366-380.  
CHEN B, WANG J, WANG Y. Intelligent virtual training partner in embedded training system of fighter[J]. Acta Aeronauticae Astronautica Sinica, 2020, 41(6):366-380 (in Chinese).
- [ 3 ] RODIN E Y, LIROV Y, MITNIK S, et al. Artificial intelligence in air combat games[J]. Computers and Mathematics with Applications, 1987, 3(1):261-274.
- [ 4 ] 徐光达,吕超,王光辉,等. 基于双矩阵对策的UCAV空战自主机动决策研究[J]. 舰船电子工程,2017,37(11):24-28.  
XU G D, LV C, WANG G H, et al. Research on UCAV autonomous air combat maneuvering decision-making based on bi-matrix game[J]. Ship Electronic Engineering, 2017, 37(11):24-28 (in Chinese).
- [ 5 ] SCHREIER M, WILLERT V, ADAMY J. An integrated approach to maneuver-based trajectory prediction and criticality assessment in arbitrary road environments[J]. IEEE Transactions on Intelligent Transportation Systems, 2016, 17(10):2751-2766.
- [ 6 ] ZHONG L, TONG M A, ZHONG W. Sequential maneuvering decisions based on multi-stage influence diagram in air combat[J]. Journal of Systems Engineering and Electronics, 2007, 18(3):551-555.
- [ 7 ] VIRTANEN K, KARELAHTI J, RAIVIO T. Modeling air combat by a moving horizon influence diagram game[J]. Journal of Guidance, Control, and Dynamics, 2006, 29(5):1080-1091.
- [ 8 ] 钟麟,佟明安,钟卫. 影响图对策在多机协同空战中的应用[J]. 北京航空航天大学学报,2007,33(4):450-453.  
ZHONG L, TONG M A, ZHONG W. Application of multistage influence diagram game theory for multiple cooperative air combat[J]. Journal of Beijing University of Aeronautics and Astronautics, 2007, 33(4):450-453 (in Chinese).
- [ 9 ] VIRTANEN K, RAIVIO T, RAIMO P H. Modeling pilot's sequential maneuvering decisions by a multistage influence diagram[J]. Journal of Guidance, Control, and Dynamics, 2004, 27(4):665-677.
- [ 10 ] 常勇,姜长生,陈志伟. 基于模糊神经网络的多机多目标空战决策[J]. 电光与控制,2011,18(4):13-17.  
CHANG Y, JIANG C S, CHEN Z W. Decision-making based on fuzzy neural network for air combat of multi-aircraft against multi-target[J]. Electronics Optics & Control, 2011, 18(4):13-17 (in Chinese).
- [ 11 ] 钟麟,佟明安,张圣云. 粗糙集-神经网络集成在编队空战中的应用[J]. 系统工程与电子技术,2006,28(6):881-884.  
ZHONG L, TONG M A, ZHANG S Y. Cooperative team air combat decision based on integration of rough sets and neural networks[J]. Systems Engineering and Electronics, 2006, 28(6):881-884 (in Chinese).
- [ 12 ] MCGREW J S, HOW J P, WILLIAMS B. Air-combat strategy using approximate dynamic programming[J]. Journal of Guidance, Control, and Dynamics, 2010, 33(5):1641-1654.
- [ 13 ] MA Y F, MA X L, SONG X. A case study on air combat decision using approximated dynamic programming[J]. Mathematical Problems in Engineering, 2014, 10(4):115-122.
- [ 14 ] 李惠峰,易文峰,程晓明. 基于近似动态规划的目标追踪控制算法[J]. 北京航空航天大学学报,2019,45(3):597-605.  
LI H F, YI W F, CHENG X M. Target tracking control algorithm based on approximate dynamic programming[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(3):597-605 (in Chinese).
- [ 15 ] 黄长强,赵克新,韩邦杰,等. 一种近似动态规划的无人机机动决策方法[J]. 电子与信息学报,2018,40(10):2447-2452.  
HUANG C Q, ZHAO K X, HAN B J, et al. Maneuvering decision-making method of UAV based on approximate dynamic programming[J]. Journal of Electronics & Information Technology, 2018, 40(10):2447-2452 (in Chinese).
- [ 16 ] 高阳阳,余敏建,韩其松,等. 基于改进共生生物搜索算法的空战机动决策[J]. 北京航空航天大学学报,2019,45(3):429-436.  
GAO Y Y, YU M J, HAN Q S, et al. Air combat maneuver decision-making based on improved symbiotic organisms search algorithm[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(3):429-436 (in Chinese).
- [ 17 ] IMADO F, KURODA T. Family of local solutions in a missile-aircraft differential game[J]. Journal of Guidance, Control, and

- Dynamics, 2011, 34(2): 583-591.
- [18] PARK H, LEE B Y, TAHK M J, et al. Differential game based air combat maneuver generation using scoring function matrix [J]. International Journal of Aeronautical and Space Sciences, 2016, 17(2): 204-213.
- [19] 傅莉, 王晓光. 无人战机近距离空战微分对策建模研究[J]. 兵工学报, 2012, 33(10): 1210-1216.  
FU L, WANG X G. Research on close air combat modeling of differential games for unmanned combat air vehicles [J]. Acta Armamentarii, 2012, 33(10): 1210-1216 (in Chinese).
- [20] HUANG C Q, DONG K S, HUANG H Q, et al. Autonomous air combat maneuver decision using Bayesian inference and moving horizon optimization [J]. Journal of Systems Engineering and Electronics, 2018, 29(1): 86-97.
- [21] 付昭旺, 李战武, 强晓明, 等. 基于滚动时域控制的战斗机空战机动决策[J]. 电光与控制, 2013, 20(3): 20-24.  
FU Z W, LI Z W, QIANG X M, et al. Tactical decision-making method based on receding horizon control for air combat [J]. Electronics Optics & Control, 2013, 20(3): 20-24 (in Chinese).
- [22] 曲婉嘉, 周兴旺, 刘颖. 基于动态贝叶斯网络的防空雷达毁伤效果评估研究[J]. 电子设计工程, 2016, 24(6): 121-124.  
QU W J, ZHOU X W, LIU Y. Research on damage evaluation of air defense radar based on Bayesian network [J]. Electronic Design Engineering, 2016, 24(6): 121-124 (in Chinese).
- [23] 孟光磊, 张慧敏, 朴海音, 等. 自动化飞行训练评估中的战机动作识别[J]. 北京航空航天大学学报, 2020, 46(7): 1267-1274.  
MENG G L, ZHANG H M, PIAO H Y, et al. Recognition of fighter maneuver in automatic flight training evaluation [J]. Journal of Beijing University of Aeronautics and Astronautics, 2020, 46(7): 1267-1274 (in Chinese).
- [24] 张洪钺, 王青. 最优控制理论与应用[M]. 北京: 高等教育出版社, 2006: 127-135.  
ZHANG H Y, WANG Q. Optimal control theory and application [M]. Beijing: Higher Education Press, 2006: 127-135 (in Chinese).

## Intelligent virtual opponent decision making and guidance method in short-range air combat training

MENG Guanglei<sup>1,\*</sup>, LIU Dejian<sup>1</sup>, ZHOU Mingzhe<sup>1</sup>, PIAO Haiyin<sup>2</sup>, CHEN Yaofei<sup>1</sup>

(1. School of Automation, Shenyang Aerospace University, Shenyang 110136, China;

2. AVIC Shenyang Aircraft Design and Research Institute, Shenyang 110035, China)

**Abstract:** To train pilots' short-range air combat skills, the traditional way based on flight simulation technology is to have multiple pilots operate multiple fighter simulators at the same time. If an intelligent virtual opponent is used to assist pilots in confrontation training, not only could the normal training process without other pilots be guaranteed, but the training cost could also be reduced to a great extent. In this paper, an integrated method based on dynamic Bayesian network (DBN) and constrained gradient algorithm is proposed to realize autonomous decision making and space occupancy guidance for intelligent virtual opponents in the attack and defense game during short-range air combat training. A dynamic Bayesian network model for short-range air combat decision making is established in combination with the space occupying situation, the fire control attack area and the identification results of maneuvering actions. This model realizes an intelligent selection of occupancy guidance index in accordance with the dynamic battlefield environment. A target trajectory prediction model is built for each type of maneuvers identified online to obtain the real-time prediction of the target trajectory. With the occupancy guidance index, target trajectory prediction, and the flight performance constraints in consideration, a constraint gradient method is used to calculate the optimal occupancy guidance quantity of the intelligent virtual opponent. Thus, a seamless combination of space occupancy decision and guidance quantity calculation for intelligent virtual opponent is achieved. The simulation results of short-range air combat show that the proposed method can realize rational autonomous decision making and space occupancy guidance for intelligent virtual opponent, overcome the problem of solidifying the maneuver mode in traditional methods, and thus have better real time and optimization performance.

**Keywords:** air-combat training; intelligent virtual opponent; occupancy guidance; maneuver recognition; trajectory prediction; dynamic Bayesian network (DBN)

**Received:** 2020-12-08; **Accepted:** 2021-02-10; **Published online:** 2021-03-01 15:17

**URL:** [kns.cnki.net/kcms/detail/11.2625.V.20210301.1328.001.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210301.1328.001.html)

**Foundation items:** National Natural Science Foundation of China (61503255); Shenyang Support Program for Young and Middle-aged Scientific and Technological Innovation Talents (RC180174)

\* **Corresponding author.** E-mail: mengguanglei@yeah.net

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2020.0722

# 基于损伤力学的增材制造金属材料 疲劳寿命预测

洪海铭<sup>1</sup>, 詹志新<sup>2,\*</sup>, 王佳莹<sup>1</sup>

(1. 航空工业沈阳飞机设计研究所 综合强度部, 沈阳 110035; 2. 北京航空航天大学 航空科学与工程学院, 北京 100083)

**摘 要:** 增材制造(AM)技术发展迅速,广泛应用于航空航天领域合金构件的加工制造,而很多增材制造合金构件承受循环载荷,疲劳失效破坏十分普遍。通过建立考虑增材制造过程影响的疲劳损伤模型,计算了增材制造金属材料的疲劳寿命。给出了弹塑性本构模型和考虑增材制造过程参数的疲劳损伤模型,进而给出了疲劳寿命计算的有限元数值方法;对增材制造金属材料进行了疲劳寿命预测,预测值与试验值基本吻合,并从疲劳数据的分散性及增材制造金属材料内部的孔隙率2个方面分析了计算误差;讨论了体积能量密度比对增材制造金属材料疲劳性能的影响,并对结果进行了分析,为增材制造金属材料的疲劳损伤评定提供一种有效的方法。

**关键词:** 增材制造(AM); 金属材料; 疲劳寿命; 预估方法; 体积能量密度比

**中图分类号:** O346.2; V215.5

**文献标志码:** A

**文章编号:** 1001-5965(2022)06-0950-07

增材制造(additive manufacturing, AM)技术是一种先进的近净成形技术,通过激光或电子束的逐层扫描并熔化凝固,形成高性能结构件。增材制造技术具有低成本、快速、高效等优势<sup>[1]</sup>,近年来,大量的航空航天金属构件采用增材制造技术<sup>[2]</sup>生产制造。增材制造金属材料往往承受循环载荷,造成其明显的疲劳破坏。因此,有必要研究增材制造金属材料的疲劳分析方法。

研究表明,增材制造金属构件的疲劳分析包括试验、理论及数值计算3个方面。Zhang等<sup>[3-5]</sup>研究了增材制造的加工参数对316L不锈钢的疲劳、断裂特性的影响,提出了预测疲劳性能的模式。Spierings等<sup>[6]</sup>对增材制造的SS316L和15-5PH金属构件开展了疲劳试验研究,结果表明,增材制造的这2种材料的拉伸和疲劳特性与传统加工工艺制造出的这2种材料的特性差别不大,但是该研究仅局限在试验方面,并没有开展数值

模拟工作。Kahlin等<sup>[7]</sup>对不同增材制造工艺生产的Ti6Al4V开展了缺口疲劳特性的研究。Akita等<sup>[8]</sup>研究了选择性激光熔化630不锈钢的疲劳行为,发现加工工艺会带来材料内部的缺陷,而疲劳裂纹往往会在这些缺陷处萌生。Edwards和Ramulu<sup>[9]</sup>开展的试验工作表明,材料内部的气孔会对增材制造Ti6Al4V的疲劳寿命产生显著影响。

在金属增材制造过程中,由于激光能量的不稳定(不足或过大),容易在层间或相邻扫描线之间产生缺陷<sup>[10]</sup>,其典型类型主要有气孔、未熔合孔洞和裂纹。增材制造工艺参数(包括激光能量密度、扫描策略及制粉方式等)对缺陷的尺寸、形状和空间特征有很大的影响。已有研究表明,内部缺陷一般会随着输入能量密度的增大而减少,但当能量密度超过一定值后,熔池发生严重变形,成形件内部会产生大量缺陷<sup>[11]</sup>。在扫描策略方面,正交扫描和“岛型”扫描方式有利于激光能量

收稿日期: 2020-12-30; 录用日期: 2021-04-11; 网络出版时间: 2021-04-14 17:24

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20210414.1717.003.html

\* 通信作者。E-mail: zzxupe@163.com

**引用格式:** 洪海铭, 詹志新, 王佳莹. 基于损伤力学的增材制造金属材料疲劳寿命预测[J]. 北京航空航天大学学报, 2022, 48(6): 950-956. HONG H M, ZHAN Z X, WANG J Y. Damage mechanics-based fatigue life prediction for additive manufacturing metal materials[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(6): 950-956 (in Chinese).

的均衡输入,使得金属粉末充分熔化,从而减少未熔合缺陷<sup>[12]</sup>。制粉方式采用金属雾化法制得球形度较好的粉末,且粒度分布集中,成形件的密度相对较高,缺陷含量较少<sup>[13]</sup>。此外,球形粉末颗粒及尺寸较小的粉末具有更好的流动性,有利于减少粉末飞溅,其成形件的内部缺陷较少<sup>[14]</sup>。国内外学者对增材制造金属粉末熔化数值模拟开展了相关的研究工作<sup>[15]</sup>,通过建立热分析模型(包括热流耦合、热传导等)及相应的计算方法,分析了内部缺陷的形成机理及演化规律。内部缺陷会导致增材制造金属构件的局部应力集中,往往成为疲劳裂纹源,显著影响增材制造金属材料的疲劳损伤累积、疲劳寿命及裂纹萌生位置。

目前,研究疲劳问题的方法大多基于试验和以统计为基础的经验公式,但是该方法需要依赖大量的构件疲劳试验<sup>[16]</sup>,较为耗费资源。因此,很多学者对疲劳累积损伤理论进行了研究,并提出了多种疲劳损伤累积理论<sup>[17-18]</sup>。而损伤力学是通过引入损伤变量来研究结构材料的力学性能退化及破坏的规律<sup>[19-20]</sup>,其优点在于物理意义较明确,相应的损伤累积理论是预估构件疲劳寿命的最有效方法之一。Chaboche<sup>[21]</sup>、Lemaitre<sup>[22]</sup>等开展了大量的基于连续损伤力学理论的疲劳分析工作。

为了预测航空航天常用合金材料的疲劳寿命,本文建立了一种考虑增材制造工艺过程影响的疲劳损伤模型,选取的增材制造工艺为激光熔化技术,其增材制造工艺参数包括激光功率、扫描速度、激光间距和粉末层厚度,这4个因素可以用体积能量密度来表示。本文研究内容主要为:①推导考虑增材制造过程影响的疲劳损伤模型,并给出理论模型的有限元数值方法;②对增材制造工艺制造的合金构件的疲劳寿命进行预测,并通过试验结果验证疲劳损伤模型的正确性;③分析增材制造构件的疲劳性能随体积能量密度比的变化趋势,并分析其影响规律。

## 1 理论模型

### 1.1 弹塑性本构模型

在连续损伤力学的理论框架内,对于各向同性材料,在循环载荷作用下,一般通过弹性模量的衰减来表征材料力学性能的劣化,用损伤变量  $D$  来表示,其表达式为

$$D = (E - E_D) / E \quad (1)$$

式中: $E$ 为无损伤时材料的弹性模量; $E_D$ 为有损伤时的等效弹性模量。

对于含损伤材料,其线弹性本构方程为

$$\varepsilon_{ij}^e = \frac{1+\nu}{E} \left( \frac{\sigma_{ij}}{1-D} \right) - \frac{\nu}{E} \left( \frac{\sigma_{kk}}{1-D} \right) \delta_{ij} \quad (2)$$

式中: $\sigma_{ij}$ 为应力张量; $\sigma_{kk}$ 为正应力分量; $\varepsilon_{ij}^e$ 为弹性应变张量; $\nu$ 为泊松比; $\delta_{ij}$ 为克罗内克符号。

在小变形条件下,总应变  $\varepsilon_{ij}$  为弹性分量  $\varepsilon_{ij}^e$  和塑性分量  $\varepsilon_{ij}^p$  之和:

$$\varepsilon_{ij} = \varepsilon_{ij}^e + \varepsilon_{ij}^p \quad (3)$$

假设材料符合 von Mises 屈服准则,则屈服方程  $f$  表示为

$$f = \sqrt{\frac{3}{2} \left( \frac{S_{ij}}{1-D} - X_{ij} \right) \left( \frac{S_{ij}}{1-D} - X_{ij} \right)} - Q \quad (4)$$

式中: $S_{ij}$ 为应力偏量部分; $Q$ 为各向同性硬化屈服面半径; $X_{ij}$ 为背应力分量。

塑性应变分量的演化关系为

$$\dot{\varepsilon}^p = \frac{3}{2} \cdot \frac{\dot{\lambda}}{1-D} \cdot \frac{S_{ij}/(1-D) - X_{ij}}{(S_{ij}/(1-D) - X_{ij})_{eq}} \quad (5)$$

$$\dot{p} = \sqrt{\frac{2}{3} \dot{\varepsilon}_{ij}^p \dot{\varepsilon}_{ij}^p} = \frac{\dot{\lambda}}{1-D} \quad (6)$$

$$\dot{X}_{ij}^{(k)} = (1-D) \left( \frac{2}{3} C_k \dot{\varepsilon}_{ij}^p - \gamma_k X_{ij}^{(k)} \dot{p} \right) \quad (7)$$

式中: $\lambda$ 为塑性乘子; $p$ 为累积塑性应变; $\gamma_k$ 、 $C_k$ 为材料参数。

### 1.2 疲劳损伤模型

在单轴循环载荷作用下,疲劳损伤演化方程为<sup>[23]</sup>

$$\frac{dD}{dN} = [1 - (1-D)^{\beta+1}]^\alpha \left[ \frac{\sigma_a}{m(1-n\sigma_m)(1-D)} \right]^\beta \quad (8)$$

式中: $N$ 为循环次数; $\sigma_a$ 和 $\sigma_m$ 分别为应力幅值和平均应力; $\alpha$ 、 $\beta$ 、 $m$ 、 $n$ 为材料常数。

在多轴循环载荷作用下,疲劳损伤演化方程可以表示为

$$\frac{dD}{dN} = [1 - (1-D)^{\beta+1}]^\alpha \left[ \frac{A_{11}}{m(1-n\sigma_{H,m})(1-D)} \right]^\beta \quad (9)$$

式中: $A_{11}$ 为八面体剪切应力幅值; $\sigma_{H,m}$ 为平均静水应力。

由式(8)和式(9)可见,在传统的单轴和多轴疲劳损伤演化方程中,并没有与增材制造工艺参数相关的项,因此,这2个方程不适合直接用来预测增材制造合金构件的疲劳寿命。文献[3-5]表明,增材制造工艺参数(包括激光功率 $p_1$ 、扫描速度 $v$ 、激光间距 $h$ 和层厚 $t$ )与材料的孔隙率、制造缺陷和微观结构密切相关,这些参数对增材制造合金构件的疲劳寿命有显著影响。考虑到上述4

个增材制造工艺参数能够由体积能量密度来表示,因此,本文提出修正的疲劳损伤演化方程。

对于增材制造金属材料,在单轴循环载荷下,疲劳损伤演化方程为

$$\frac{dD}{dN} = [1 - (1 - D)^{\beta+1}]^{\alpha} \left\{ \frac{\sigma_a(1 + |1 - R_E|)}{m(1 - n\sigma_m)(1 - D)} \right\}^{\beta} \quad (10)$$

对于多轴疲劳,疲劳损伤演化方程为

$$\frac{dD}{dN} = [1 - (1 - D)^{\beta+1}]^{\alpha} \left\{ \frac{A_{II}(1 + |1 - R_E|)}{m(1 - n\sigma_{II,m})(1 - D)} \right\}^{\beta} \quad (11)$$

式中: $R_E = E_d/E_{d0}$ 为体积能量密度比, $E_d$ 为体积能量密度, $E_{d0}$ 为参考体积能量密度。参考体积能量密度用于描述增材制造过程状态<sup>[3]</sup>,没有明确的物理意义。在激光增材过程中,当激光体积能量密度达到某一个值时,增材制造材料/结构的孔隙率最小,并且其表面光洁度和硬度都较好,那么这个值即被指定为该种材料的参考体积能量密度,其实用价值在于对增材工艺参数的选择及优化提供指导。参考体积能量密度的值可以由以下方式得到:根据经验,选取一组增材工艺参数制造加工金属材料,并测量其孔隙率及硬度等;改变增材工艺参数(如激光功率、扫描速度、激光间距和粉末层层厚等),并测量不同增材制造工艺参数下金属材料的孔隙率及硬度等;通过一系列对比研究,将具有最小孔隙率及较好硬度的金属材料所对应的那一组参数,即对应的激光体积能量密度,作为参考体积能量密度。

在恒定应力条件下,疲劳寿命 $N_f$ 可以通过对 $D$ 从0到1积分计算得到。对于单轴情况:

$$N_f = A \left\{ \frac{\sigma_a(1 + |1 - R_E|)}{(1 - C\sigma_m)} \right\}^{-B} \quad (12)$$

对于多轴情况:

$$N_f = A \left\{ \frac{A_{II}(1 + |1 - R_E|)}{(1 - C\sigma_{II,m})} \right\}^{-B} \quad (13)$$

式中: $A = m^{\beta}/[(1 - \alpha)(1 + \beta)]$ ;  $B = \beta$ ;  $C = n$ 。

### 1.3 理论模型中材料的参数标定方法

#### 1.3.1 弹塑性本构模型中的参数标定方法

由单轴拉伸静力试验可以直接确定材料的弹性模量、屈服应力和强度极限。各向同性硬化参数需要单轴拉伸-压缩循环试验确定,假设材料服从随动硬化,采用3个指数形式的背应力分量来描述材料的非线性硬化行为:

$$X_k = \frac{C_k}{\gamma_k} (1 - e^{-\gamma_k \varepsilon^p}) \quad (14)$$

式中: $\varepsilon^p$ 为塑性应变。

则单轴应力应变曲线可以表示为

$$\sigma = \sigma_y + \sum_{k=1}^3 \frac{C_k}{\gamma_k} (1 - e^{-\gamma_k \varepsilon^p}) \quad (15)$$

式中: $\sigma_y$ 为初始屈服应力。

#### 1.3.2 疲劳损伤模型中的参数标定方法

在弹性损伤演化方程中,一共有4个参数需要标定。根据标准光滑试件在不同应力比下的疲劳试验,通过使用最小二乘法,可以标定得到参数,材料参数随着增材制造工艺参数的变化而改变。表1中列出了2种标定的材料参数结果。

表1 标定的2种材料参数

Table 1 Two kinds of calibrated material parameters

材料	$\alpha$	$\beta$	$m$	$n$
SS316L	0.84	2.46	52 164	0.001 8
Ti6Al4V	0.9	2.89	70 405	0.004 26

## 2 疲劳寿命预测的有限元数值方法

基于损伤力学的增材制造金属材料疲劳寿命预测的计算方法如下:

**步骤1** 初始化所有参数,将增材制造金属材料的初始损伤度设为0。

**步骤2** 利用有限元软件计算在循环载荷作用下的循环应力和应变。对于弹性损伤累积的情况,逐个计算疲劳损伤很耗时,因此采用循环块的方法计算疲劳损伤增量,以及对应于循环次数的总损伤度。

**步骤3** 如果任何单元的损伤累积达到1,则假定在该单元处疲劳裂纹萌生,相应的循环次数就是疲劳裂纹萌生寿命;否则,重新计算每个单元的材料特性,以及应力场和损伤场,直至裂纹萌生。

## 3 增材制造金属材料的疲劳寿命预测结果

### 3.1 SS316L合金的疲劳寿命预测结果

根据第1节和第2节理论模型和有限元数值方法,在不同的载荷条件下,预测增材制造SS316L合金的疲劳寿命。对于SS316L合金构件,利用MTS试验系统对其进行疲劳试验<sup>[3-5]</sup>。其中,应力比 $R = 0.1$ ,循环载荷中的最大应力从200 MPa变化到660 MPa。体积能量密度比 $R_E$ 的变化范围为0.3~1。全部的预测结果与试验数据的对比如图1所示。可以看到,计算结果与试验结果基本一致,一共有93组试验数据,只有10组数据超出了三倍误差带。当体积能量密度比 $R_E$ 较小时,本文提出的疲劳损伤模型不能非

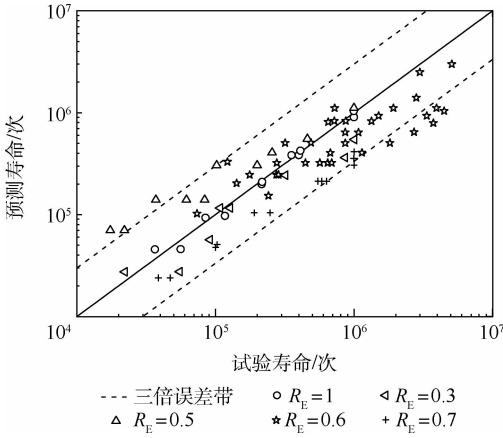


图 1 SS316L 合金的疲劳寿命预测结果

Fig. 1 Predicted fatigue lives for SS316L

常准确地预测 SS316L 合金的疲劳寿命。可能有以下 2 个原因导致这些试验疲劳数据超出了误差带的范围:①疲劳数据存在分散性,当体积能量密度比  $R_E$  较小时,分散性尤为明显;②体积能量密度比  $R_E$  较小时,增材制造材料内部的孔隙率变得很高,而且增材制造构件内部会有大量缺陷。

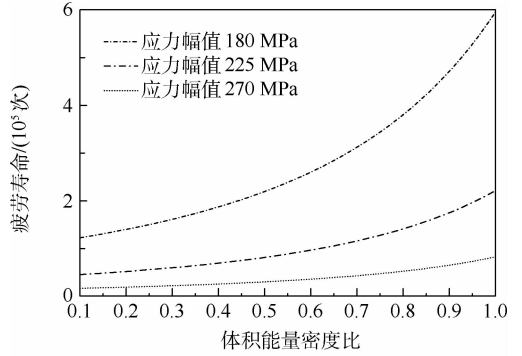
基于疲劳损伤模型分析 SS316L 合金部件疲劳损伤的 2 个重要特性:①不同应力幅值下,疲劳寿命随体积能量密度比的变化趋势;②不同应力比下,疲劳寿命随体积能量密度比的变化趋势。

图 2(a)给出了在不同的应力幅值,应力比均为 0.1 的情况下,疲劳寿命随体积能量密度比  $R_E$  的变化曲线。可以看到:①疲劳寿命随  $R_E$  的增大而增加;②应力幅值越小,疲劳寿命的增长趋势越快,应力幅值越大,疲劳寿命的增长趋势越不明显;③当  $R_E$  达到 1 时,疲劳寿命达到最大值,意味着 SS316L 具有最小的孔隙率。图 2(b)给出了在不同的应力比、最大应力均为 500 MPa 的情况下,疲劳寿命随体积能量密度比  $R_E$  的变化曲线。可以看到:①疲劳寿命随  $R_E$  的增大而迅速增加;②在相同的体积能量密度比下,应力比越大,疲劳寿命越长;③对于不同的应力比,当  $R_E$  达到 1 时,SS316L 的疲劳寿命达到最大值。

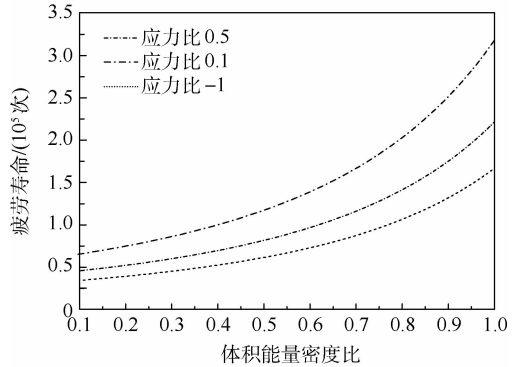
### 3.2 Ti6Al4V 合金的疲劳寿命预测结果

基于 1.2 节的疲劳损伤模型,预测了增材制造 Ti6Al4V 合金的疲劳寿命,并与文献[24-26]中的试验结果进行了对比。

在 Ti6Al4V 合金部件的疲劳试验中,应力比不同,应力范围为 90 ~ 800 MPa,体积能量密度比  $R_E$  介于 0.6 ~ 1.8 之间。图 3 给出了预测结果与试验数据的对比。结果表明,疲劳寿命预测值与试验值基本吻合,一共有 75 组试验数据,只有 9 组数据位于三倍误差带之外。有 2 个可能的原



(a) 不同应力幅值下, SS316L 合金疲劳寿命随体积能量密度比的变化曲线



(b) 不同应力比下, SS316L 合金疲劳寿命随体积能量密度比的变化曲线

图 2 SS316L 合金疲劳寿命的变化曲线

Fig. 2 Variation curves of fatigue life for SS316L

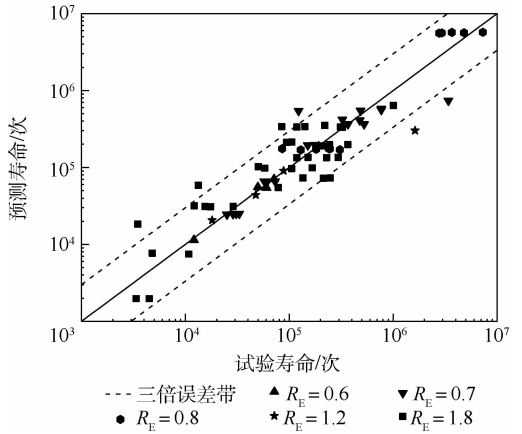


图 3 Ti6Al4V 合金的疲劳寿命预测结果

Fig. 3 Predicted fatigue lives for Ti6Al4V

因使得这些数据超出了误差带的范围:①由疲劳数据的分散性直接造成;②当体积能量密度比  $R_E$  超过 1 时,有可能是激光功率  $p_1$  过大,或者激光扫描速度  $v$  过低导致,此时增材制造材料内部会有一些未熔化的颗粒,也会对疲劳寿命带来很大影响。因此,相应的预测疲劳寿命不准确。

利用本文提出的疲劳损伤模型研究讨论 Ti6Al4V 合金部件疲劳损伤的 2 个重要特性:①不同应力幅值下,疲劳寿命随体积能量密度比的变化趋势;②不同应力比下,疲劳寿命随体积能量密度比的变化趋势。

图4(a)给出了在不同的应力幅值、应力比均为0.2的情况下,疲劳寿命随体积能量密度比 $R_E$ 的变化曲线,可以看到:①当体积能量密度比 $R_E$ 小于1时,疲劳寿命随 $R_E$ 的增大而增加;②应力幅值越小,在 $R_E$ 小于1时,疲劳寿命的增长趋势越快,应力幅值越大,疲劳寿命的增长趋势越不明显;③当 $R_E$ 超过1时,疲劳寿命随 $R_E$ 的增大而降低。图4(b)给出了在不同的应力比、最大应力均为600 MPa的情况下,疲劳寿命随体积能量密度比 $R_E$ 的变化曲线,可以看到:①在 $R_E$ 小于1时,疲劳寿命随 $R_E$ 的增大而迅速增加;②在相同的体积能量密度比下,应力比越大,疲劳寿命越长;③对于不同的应力比,当 $R_E$ 达到1时,Ti6Al4V的疲劳寿命达到最大值。

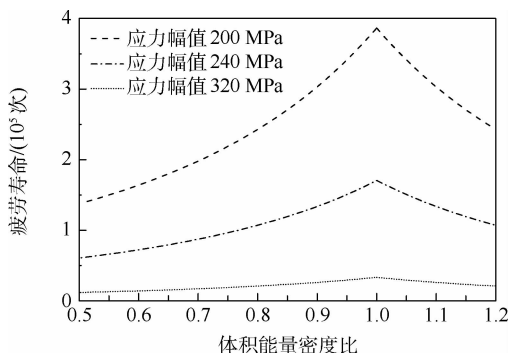
模型,并给出了增材制造金属材料疲劳分析的有限元数值方法。

2) 预测了增材制造 SS316L 和 Ti6Al4V 合金材料的疲劳寿命,预测值与试验值基本吻合,并从疲劳数据的分散性及增材制造材料内部的孔隙率 2 个方面分析了计算误差。

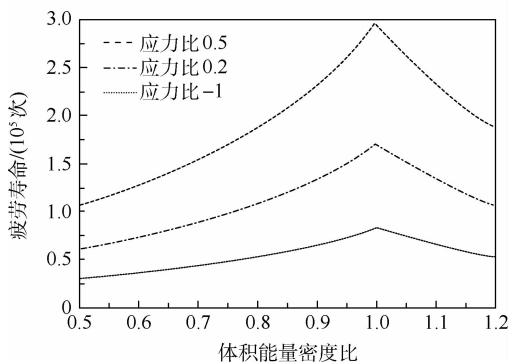
3) 增材制造合金部件的疲劳特性与体积能量密度比 $R_E$ 密切相关,当 $R_E$ 小于1时,疲劳寿命随 $R_E$ 的增加而迅速增加。当 $R_E$ 大于1时,疲劳寿命则随 $R_E$ 的增加而略有降低。当 $R_E$ 等于1时,SS316L 和 Ti6Al4V 合金材料具有较小的孔隙率,其疲劳性能达到最好。

### 参考文献 (References)

- [1] NGO T D, KASHANI A, IMBALZANO G, et al. Additive manufacturing (3D printing): A review of materials, methods, applications and challenges [J]. *Composites Part B: Engineering*, 2018, 143: 172-196.
- [2] BLAKEY-MILNER B, GRADL P, SNEDDEN G, et al. Metal additive manufacturing in aerospace: A review [J]. *Materials & Design*, 2021, 209: 110008.
- [3] ZHANG M, SUN C N, ZHANG X, et al. Fatigue and fracture behaviour of laser powder bed fusion stainless steel 316L; Influence of processing parameters [J]. *Materials Science and Engineering A*, 2017, 703(4): 251-261.
- [4] ZHANG M, SUN C N, ZHANG X, et al. Elucidating the relations between monotonic and fatigue properties of laser powder bed fusion stainless steel 316L [J]. *The Journal of the Minerals, Metals & Materials Society*, 2018, 70(3): 390-395.
- [5] ZHANG M, SUN C N, ZHANG X, et al. Predictive models for fatigue property of laser powder bed fusion stainless steel 316L [J]. *Materials & Design*, 2018, 145: 42-54.
- [6] SPIERINGS A B, STARR T L, WEGENER K. Fatigue performance of additive manufactured metallic parts [J]. *Rapid Prototyping Journal*, 2013, 19(2): 88-94.
- [7] KAHLIN M, ANSELL H, MOVERARE J J. Fatigue behaviour of notched additive manufactured Ti6Al4V with as-built surfaces [J]. *International Journal of Fatigue*, 2017, 101: 51-60.
- [8] AKITA M, UEMATSU Y, KAKIUCHI T, et al. Defect-dominated fatigue behavior in type 630 stainless steel fabricated by selective laser melting [J]. *Materials Science and Engineering A*, 2016, 666(6): 19-26.
- [9] EDWARDS P, RAMULU M. Fatigue performance evaluation of selective laser melted Ti-6Al-4V [J]. *Materials Science and Engineering A*, 2014, 598(3): 327-337.
- [10] SANAEI N, FATEMI A. Defects in additive manufactured metals and their effect on fatigue performance: A state-of-the-art review [J]. *Progress in Materials Science*, 2020, 117(4): 100724.
- [11] KASPEROVICH G, HAUBRICH J, GUSSONE J, et al. Correlation between porosity and processing parameters in TiAl6V4 produced by selective laser melting [J]. *Materials & Design*, 2016, 105: 160-170.



(a) 不同应力幅值下, Ti6Al4V合金疲劳寿命随体积能量密度比的变化曲线



(b) 不同应力比下, Ti6Al4V合金疲劳寿命随体积能量密度比的变化曲线

图4 Ti6Al4V合金疲劳寿命的变化曲线

Fig. 4 Variation curves of fatigue life for Ti6Al4V

## 4 结论

本文建立了一种考虑增材制造过程参数的疲劳分析模型,计算了增材制造 SS316L 和 Ti6Al4V 的疲劳寿命,与试验结果进行了对比,并研究了疲劳损伤特性随体积能量密度比的变化规律。主要结论如下:

1) 基于连续损伤力学理论,推导了损伤耦合的本构模型和考虑增材制造过程参数的疲劳损伤

- [12] ZHANG B, LI Y, BAI Q. Defect formation mechanisms in selective laser melting: A review [J]. Chinese Journal of Mechanical Engineering, 2017, 30(3): 515-527.
- [13] 陈迪, 王焯, 高海燕, 等. 3D 打印钛合金内部孔洞的研究进展 [J]. 应用激光, 2019, 39(1): 72-78.  
CHEN D, WANG L, GAO H Y, et al. Research progress on 3D printing internal cavity of titanium alloy [J]. Applied Laser, 2019, 39(1): 72-78 (in Chinese).
- [14] TONELLI L, LIVERANI E, VALLI G, et al. Effects of powders and process parameters on density and hardness of A357 aluminum alloy fabricated by selective laser melting [J]. The International Journal of Advanced Manufacturing Technology, 2020, 106(1-2): 371-383.
- [15] YAN W, GE W, QIAN Y, et al. Multi-physics modeling of single/multiple-track defect mechanisms in electron beam selective melting [J]. Acta Materialia, 2017, 134(8): 324-333.
- [16] 姚卫星. 结构疲劳寿命分析 [M]. 北京: 国防工业出版社, 2003: 20-25.  
YAO W X. Fatigue life analysis of structures [M]. Beijing: National Defense Industry Press, 2003: 20-25 (in Chinese).
- [17] ZHU S P, LIAO D, LIU Q, et al. Nonlinear fatigue damage accumulation: Isodamage curve-based model and life prediction aspects [J]. International Journal of Fatigue, 2019, 128: 105185.
- [18] BENKABOUCHE S, GUECHICHI H, AMROUCHE A, et al. A modified nonlinear fatigue damage accumulation model under multiaxial variable amplitude loading [J]. International Journal of Mechanical Sciences, 2015, 100: 180-194.
- [19] YUAN H, ZHANG L, MA S. Damage evolution and characterization for sintered powder metals with the varying porosity [J]. Engineering Fracture Mechanics, 2019, 207: 86-98.
- [20] 郑战光, 蔡敢为, 李兆军. 一种新的疲劳损伤演化模型 [J]. 工程力学, 2010, 27(2): 37-40.  
ZHENG Z G, CAI G W, LI Z J. A new model of fatigue damage evolution [J]. Engineering Mechanics, 2010, 27(2): 37-40 (in Chinese).
- [21] CHABOCHE J L, GALLERNEAU F. An overview of the damage approach of durability modelling at elevated temperature [J]. Fatigue & Fracture of Engineering Materials & Structures, 2001, 24(6): 405-418.
- [22] LEMAITRE J. A course on damage mechanics [M]. Berlin: Springer, 2012: 1-20.
- [23] SHEN F, HU W, MENG Q. A non-local approach based on the hypothesis of damage dissipation potential equivalence to the effect of stress gradient in fretting fatigue [J]. International Journal of Fatigue, 2016, 90: 125-138.
- [24] LI P, WARNER D H, FATEMI A, et al. Critical assessment of the fatigue performance of additively manufactured Ti-6Al-4V and perspective for future research [J]. International Journal of Fatigue, 2016, 85: 130-143.
- [25] LIU Q C, ELAMBASSERIL J, SUN S J, et al. The effect of manufacturing defects on the fatigue behaviour of Ti-6Al-4V specimens fabricated using selective laser melting [J]. Advanced Materials Research, 2014, 891(3): 1519-1524.
- [26] VAN HOOREWEDER B, BOONEN R, MOENS D, et al. On the determination of fatigue properties of Ti6Al4V produced by selective laser melting: AIAA 2012-1733 [R]. Reston: AIAA, 2012.

# Damage mechanics-based fatigue life prediction for additive manufacturing metal materials

HONG Haiming<sup>1</sup>, ZHAN Zhixin<sup>2,\*</sup>, WANG Jiaying<sup>1</sup>

(1. Comprehensive Strength Department, AVIC Shenyang Aircraft Design and Research Institute, Shenyang 110035, China;

2. School of Aeronautic Science and Engineering, Beihang University, Beijing 100083, China)

**Abstract:** The additive manufacturing (AM) technology develops rapidly and it is widely employed in the fabrication of aerospace alloy components. Many additive manufacturing alloy components bear cyclic loadings, and the fatigue failure is very common. The fatigue damage model considering the influence of additive manufacturing process is established and the fatigue lives of additive manufacturing alloy materials are predicted. The elastic-plastic constitutive model and the fatigue damage model considering the additive manufacturing process parameters are presented, and the finite element numerical method is presented for the fatigue life computation. The fatigue lives of additive manufacturing metal materials are predicted, which are basically consistent with the experimental results, and the computed errors are analyzed from two aspects including the scatter of fatigue data and the porosity in the additive manufacturing materials. The influence of the volume energy density ratio on the fatigue properties of additive manufacturing metal materials is discussed, and the results are analyzed. This research provides an effective method for the fatigue damage evaluation of additive manufacturing metal materials.

**Keywords:** additive manufacturing (AM); metal materials; fatigue life; prediction method; volume energy density ratio

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2021.0089

# 基于朝向约束和重识别特征的目标轨迹关联方法

艾明晶\*, 单国志, 刘鹏高, 杨澎钢

(北京航空航天大学 虚拟现实技术与系统国家重点实验室, 北京 100083)

**摘 要:** 基于检测关联和深度学习的目标轨迹关联方法是计算机视觉领域的研究热点之一,但现有方法设计中缺乏有效的时空约束,且目标表现特征泛化能力不足,在目标朝向差异明显的情况下会发生识别错误,在目标轨迹关联时会导致频繁的目标ID切换和错误关联。针对该问题,提出了一种基于朝向约束和重识别特征的目标轨迹关联方法。将行人朝向判别引入行人重识别中,提出了一种具有朝向约束力的行人重识别网络模型,提升了目标特征的代表能力。结合目标朝向、卡尔曼滤波得到的位置信息、重叠面积等时空特征,提出一种基于朝向约束的分层轨迹关联模型,得到单相机内的目标轨迹。在跨相机场景中,通过引入一种简单有效的双向竞争匹配机制,实现了目标轨迹的有效关联。实验结果表明:所提方法在MOT数据集上度量指标优于多种方法,能够减少频繁的目标ID交换,有效解决了相似目标相向而行时的错误关联;帧率达到19.6帧/s,能够满足近实时场景下的使用要求。

**关键词:** 轨迹关联; 行人重识别; 多目标跟踪; 朝向约束; 时空约束

**中图分类号:** TP391.4; TP391

**文献标志码:** A **文章编号:** 1001-5965(2022)06-0957-11

多目标跟踪(multi-object tracking, MOT)和轨迹关联是计算机视觉最重要的研究方向之一。目标轨迹关联主要解决如何根据时间自动生成目标完整运动轨迹的问题。跨相机完整轨迹的生成需要通过多目标跟踪获取单相机内的目标轨迹,而所获得跟踪轨迹的准确性主要依赖于目标检测的准确性、表现特征区分度及时空关联约束有效性等方面。

基于检测和数据关联的方法(TBD)是目标轨迹关联中最主要的研究方法,主要利用检测器获得视频序列帧中目标可能存在的区域,通过关联模型将属于同一运动目标的检测包围框关联到一起。对于这种关联模型,如何提取和得到有效的行人特征是关键问题。

因此,行人重识别(pedestrian re-identification, ReID)与轨迹关联技术紧密相关,相辅相成。自2006年开始,行人重识别技术逐渐从目标跟踪中分离出来,成为了一个独立的研究课题<sup>[1]</sup>。行人重识别能够依靠简单的外观特征,实现跨相机场景下的行人目标检索;同时,其重点研究的特征提取和度量方法可以为视频目标之间的数据关联提供强有力的支撑,是实现多目标跟踪和轨迹关联的重要研究基础,特别是在跨相机目标轨迹关联场景中,重识别技术更能发挥优势。

轨迹关联方法的难点主要体现在2个方面:①如何刻画和表达每帧内目标的特征,主要涉及外观、运动和时空等建模;②如何基于得到的特征来判断不同帧内的目标是否属于同一行人,主要

收稿日期: 2021-02-23; 录用日期: 2021-03-26; 网络出版时间: 2021-06-22 15:52

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20210622.1225.001.html](http://kns.cnki.net/kcms/detail/11.2625.V.20210622.1225.001.html)

基金项目: 国家自然科学基金(61872024); 国家重点研发计划(2018YFB2100601)

\* 通信作者. E-mail: amj@buaa.edu.cn

**引用格式:** 艾明晶, 单国志, 刘鹏高, 等. 基于朝向约束和重识别特征的目标轨迹关联方法[J]. 北京航空航天大学学报, 2022, 48(6): 957-967. AI M J, SHAN G Z, LIU P G, et al. Target trajectory association method based on orientation constraint and re-identification feature[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(6): 957-967 (in Chinese).

涉及数据关联模型。

然而,受限于现有重识别特征的泛化能力,并且由于缺乏有效的时空约束,现有轨迹关联方法在目标朝向差异明显的情况下会导致频繁>ID切换和错误关联,在大多数场景中难以保证轨迹关联的准确性。

朝向作为运动目标的一种固有属性,其所蕴含的目标前进方向是一种有效的时空约束。有效的朝向先验信息不仅可以提升目标识别的准确率,还能够克服行人相向而行时的错误关联,从而增强目标轨迹关联的置信度。因此,本文研究主要以行人目标为例,并以准确的行人朝向判别为基础,提出一种基于朝向约束的行人重识别和轨迹关联方法,并通过实验详细证明了该方法的优越性和有效性。

随着深度学习算法在目标检测、语义分割等研究领域的广泛应用,特别是 Faster R-CNN<sup>[2]</sup>、YOLO<sup>[3]</sup>等基于深度学习的检测算法的出现和行人重识别技术的不断研究,基于深度学习的轨迹生成和关联方法逐渐成为研究热点。其中,基于检测和数据关联的方法是当下最主要的研究框架,主要利用检测器获得视频序列每帧中目标可能存在的区域,通过关联模型将属于同一运动目标的检测包围框关联到一起。

随着行人重识别的不断研究,图像的特征表达能力和可区分度逐渐提升,越来越多的研究将行人重识别技术引入到轨迹关联模型的设计中,将行人重识别得到的外观特征作为数据关联和目标跟踪的相似性判别依据,并获得了显著的性能提升。

2016年,Ristani等<sup>[4]</sup>发布了 DukeMTMC 数据集,并提出了相应的外观特征提取方法,极大推动了目标重识别在轨迹关联技术中的应用。同年,Leal-Taixe等<sup>[5]</sup>基于行人重识别的研究经验,直接使用 Siamese 网络提取重识别特征用于生成目标运动轨迹。

2017年,Wojke等<sup>[6]</sup>在原有 SORT 算法<sup>[7]</sup>的基础上引入行人重识别特征,提出了 DeepSORT 方法,通过卡尔曼滤波预测运动信息,利用重识别网络得到的全局特征向量作为外观信息,将运动和外观信息加权融合后作为数据关联依据,并使用匈牙利算法将预测后得到的轨迹和当前帧中的检测框进行数据匹配。

2018年,Chen等<sup>[8]</sup>提出了 MOTDT 方法,将目标检测和跟踪算法结合起来,相互补充,并使用区域选择网络 R-FCN 和轨迹评分机制选择候选

目标,利用行人重识别得到的行人特征实现轨迹关联。同年,Ristani和Tomasi<sup>[9]</sup>研究了行人重识别任务在轨迹关联中的适配方案,提出了 DeepCC 方法,但不同之处在于该方法重点关注训练策略,提出了一种改进的自适应加权三元组,并探究了行人重识别和轨迹关联的关系。

2020年,Wang等<sup>[10]</sup>基于 YOLO 检测算法和 MOTDT 方法,提出了融合检测和行人重识别的 JDE 方法,将重识别特征的提取和检测网络同时实现,在检测网络的特征图中提取了一个嵌入向量,用三元组损失进行训练,从而提升了轨迹关联的速度。Bergmann等<sup>[11]</sup>认为检测和跟踪同时产生的结果很难进行评判,因此直接采用检测框的回归和重叠面积代替了跟踪器的预测。

2020年,Zhan等<sup>[12]</sup>认为行人检测和行人重识别使用2个单独的模型难以共享特征,实时性不高,进一步提出了一种 FairMOT 方法,在目标检测算法的基础上,增加了一个分支来提取行人重识别特征,同时获取目标中心,从而将目标检测和重识别很好地统一起来。同年,Hou等<sup>[13]</sup>深入分析了跟踪问题和行人重识别问题的区别,认为跟踪只需要局部匹配,直接用行人重识别模型得到的全局匹配特征并不能得到最好的结果,提出了一种局部邻域内目标外观的度量方法。

综上所述,行人重识别技术与目标轨迹关联密不可分,二者虽然是独立研究的课题,但相辅相成。将行人重识别应用到轨迹关联中已经是目前研究中最主流的方法。因此,本文提出一种基于朝向约束和重识别特征的轨迹关联方法,以更好地将二者适配起来,平衡其准确率和实时性。

## 1 朝向约束行人重识别方法

行人重识别技术最直接的应用就是多目标跟踪和跨相机轨迹关联。目标的表现特征是跟踪过程中关联和匹配的主要依据。

因此,本文提出一种具有朝向约束力的行人重识别方法,得到的行人特征表示是后续进行目标轨迹关联的重要基础,也是保证目标关联准确性的前提,其主要技术路线如图1所示。

如图2所示,为了适应轨迹关联中行人朝向变化对表现特征的影响,本文设计了一种基于朝向约束的行人重识别网络。该网络是基于卷积神经网络的双分支网络结构,可以同时将1个样本映射到2个不同的特征空间,每个特征空间对应1个网络分支。其中,每个分支设计了不同的混

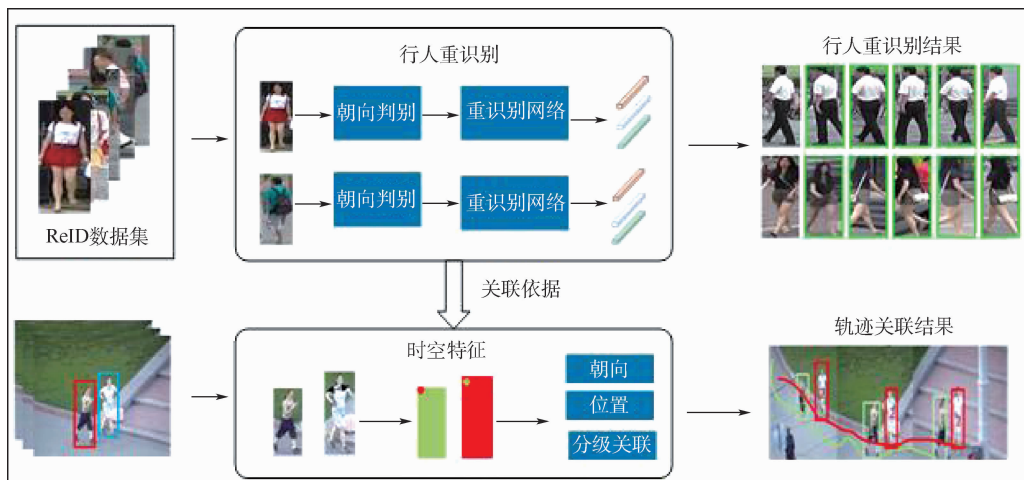


图 1 整体研究思路与技术路线

Fig. 1 Overall research ideas and technical route

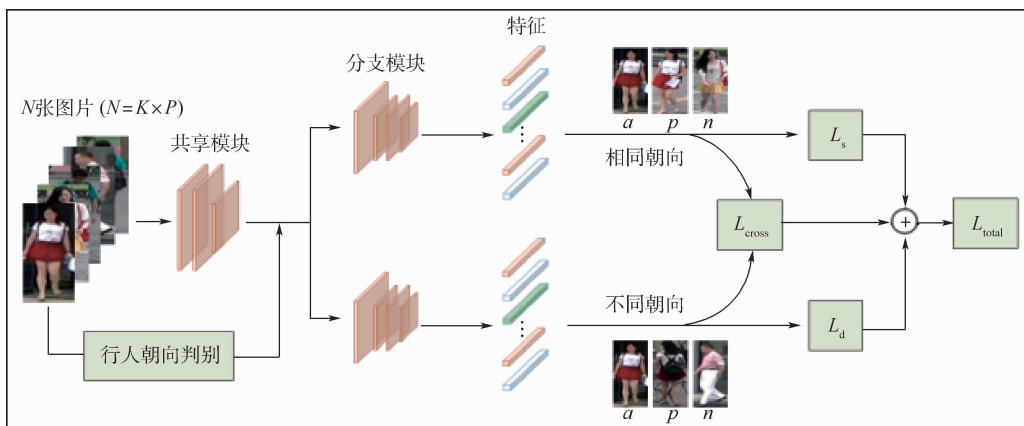


图 2 朝向约束行人重识别网络结构示意图

Fig. 2 Structure diagram of pedestrian re-identification network with orientation constraint

合损失函数,这样的网络设计使得第 1 分支(称为相同朝向分支)主要关注相同朝向的样本,而第 2 分支(称为不同朝向分支)更加适应不同朝向样本的变化情况。

在训练时,采用小批量样本采样策略,从数据集中随机挑选  $P$  个行人,并根据朝向两两组合,为每个行人选择  $K$  张图像。由此得到的每个 batch 先通过共享模块提取浅层特征,再分别使用 2 个分支模块将其映射到高维子空间中。

2 个分支的不同之处在于损失函数的选择。三元组损失<sup>[14-15]</sup>(triplet loss)可以考虑正负样本之间的类间距,在行人重识别任务中具有很好的表现。其中,在相同朝向分支中,主要采用相同朝向的样本组成三元组,损失函数  $L_s$  为

$$L_s = \frac{1}{P \times K} \sum_{a \in \text{batch}} [\max d(a, p_s) - \min d(a, n_s) + \alpha]_+ \quad (1)$$

式中: $P$  为行人数目; $K$  为图像个数; $d$  为样本之间的距离; $a$  为训练样本 anchor; $p_s$  为与 anchor 具

有相同朝向的正样本; $n_s$  为与 anchor 朝向不同的负样本; $\alpha$  为三元组阈值。因此,该损失主要关注的是相同朝向样本对之间的距离关系。

在不同朝向分支中,主要从 batch 中选择不同朝向的样本组成三元组,考虑不同朝向样本的训练,从而弥补相同朝向分支的不足,损失函数的数学表达为

$$L_d = \frac{1}{P \times K} \sum_{a \in \text{batch}} [\max d(a, p_d) - \min d(a, n_d) + \beta]_+ \quad (2)$$

式中: $p_d$  为与 anchor 朝向不同的正样本; $n_d$  为与 anchor 朝向不同的负样本; $\beta$  为距离阈值。为了综合协调整个训练过程,在两分支的基础上添加交叉约束和类内约束三元组。其中,交叉约束用于保证正样本对间的距离小于负样本对,类内约束保证同一行人内部不同朝向之间的样本可分。

经过多损失函数的联合训练,可以得到具有朝向约束力的行人重识别模型。在测试时,先通过行人朝向判别得到图像中行人的朝向标签,再

根据朝向关系选择不同分支的行人特征计算距离,从而转化为行人相似度。

因此,朝向判别方法在网络中的作用至关重要。具体实现时,行人朝向判别采取基于 Res-Net50 的网络进行四分类,将行人图像划分为正、背、左、右 4 个类别。该朝向判别模型在 RAP 数据集<sup>[16]</sup>上训练得到,实验证明,其判别准确率达到 89%,在行人重识别模型训练过程中,该判别模型参数保持不变。

## 2 单相机轨迹生成方法

多目标跟踪主要用于单相机内目标轨迹的生成。对于视频来说,由同一个人同一时刻只能出现在一个位置的客观规律可知,帧号所表达的时间信息及检测框位置所体现的空间信息是关联目标的主要依据。朝向作为最直观的空间信息之一,反映的目标行进方向对前后帧之间的目标匹配具有重要价值,因为前后帧之间目标的朝向几乎不会突变,并且准确的朝向信息可以有效解决相似目标相向而行时的错误关联。

本节主要针对轨迹关联中时空约束缺失的问题,提出一种基于朝向的单相机轨迹生成方法,以朝向约束行人重识别作为表观特征关联依据,结合朝向、位置等时空运动特征,实现一种基于朝向的分层轨迹关联模型,用以生成单相机内的目标轨迹。

### 2.1 目标特征提取与匹配

对于某一个目标在单相机内的运动轨迹,用式(3)所示的九元组来描述轨迹在某一时刻的状态。

$$\text{Track} = \{u, v, \gamma, h, o, x', y', \gamma', h'\} \quad (3)$$

式中: $(u, v)$ 为目标矩形框的中心坐标; $\gamma$ 为矩形框的高宽比; $h$ 为矩形框的高度; $o$ 为该目标当前的朝向; $x', y', \gamma', h'$ 为目标在图像中对应的坐标速度。

基于以上轨迹状态,分别对目标进行运动特

征和表观特征的提取。

#### 2.1.1 运动特征

对于视频中的每一帧,先通过 YOLOv4 对其中的行人目标进行检测,同时使用基于线性观测模型的卡尔曼滤波器提取行人运动特征。预测结果用 $(u, v, \gamma, h)$ 表示,代表目标可能的位置,因此直接通过马氏距离计算其与检测目标之间的标准差,如下:

$$d_1(i, j) = (d_j - y_i)^T S_i^{-1} (d_j - y_i) \quad (4)$$

式中: $d_1$ 为马氏距离; $d_j$ 为第  $j$  个检测位置; $y_i$ 为第  $i$  个追踪器对目标的预测值; $S_i$ 为检测与预测之间的协方差矩阵。

在基于运动特征进行目标匹配时,如果某次关联的马氏距离小于指定阈值,则关联成功,如下:

$$b_1(i, j) = 1 \quad d_1(i, j) \leq t_1 \quad (5)$$

式中: $b_1$ 为关联结果; $t_1$ 为关联阈值。

马氏距离实际代表了对目标位置预测的置信度,是对目标运动特征的估计。

#### 2.1.2 表观特征

仅靠运动特征难以保证长时间的目标跟踪。基于第 1 节中提出的朝向约束行人重识别方法,对于每一帧中检测出的目标矩形框  $d_j$ ,可以先通过朝向判别得到其朝向标签,再将其输入到网络中,得到每个样本从上下 2 个分支输出的 2 个不同的特征向量  $F_s$  和  $F_d$  (特征的长度定义为 1 024 维)。其中,  $F_s$  用于相同朝向样本之间的距离计算,  $F_d$  用于不同朝向样本之间的距离计算。

在匹配当前检测框时,根据已关联轨迹的最后一个特征  $c$  及最后一次出现的朝向,与当前帧中的检测结果  $d$  两两计算欧氏距离。如图 3 所示,相同朝向的距离和不同朝向的距离根据朝向判定结果融合生成一个最终的距离矩阵  $d_2(i, j)$ 。

对于得到的距离矩阵,若其距离小于阈值  $t_2$ ,则说明检测目标与该轨迹具有较高的相似度,视为关联成功,如下:

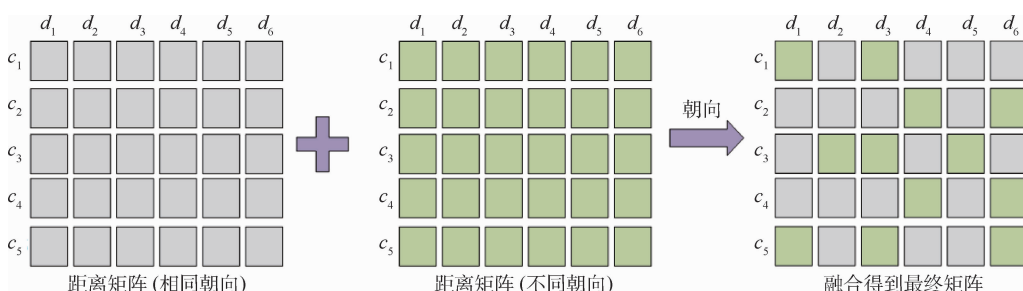


图 3 基于朝向的距离矩阵计算示意图

Fig. 3 Schematic diagram of distance matrix calculation based on orientation

$$b_2(i, j) = 1 \quad d_2(i, j) \leq t_2 \quad (6)$$

综上所述,对从视频帧中得到的每一个目标,本文用其运动特征和表观特征 2 种方法进行描述。

### 2.2 分层轨迹关联模型

本节基于提取的目标特征,提出一种具体的行人轨迹关联模型。在单相机内部,将目标从出现到消失的整个轨迹生命周期定义为 4 种状态,如图 4 所示。图中:New 代表新目标出现;Confirmed 代表轨迹确认;Lost 代表轨迹消失;Deleted 代表轨迹删除。基于以上轨迹状态,可以发现在每一次进行目标关联时,需要考虑处于 New、Confirmed 和 Lost 三种状态的轨迹。但这 3 种状态的轨迹并不是同等重要,处于 Confirmed 状态的轨迹状态最为稳定,在当前帧中持续出现的概率更大,而处于 Lost 和 New 状态的轨迹,很可能是已经消失或由误检引起的错误轨迹。因此,每种状态的轨迹应具有不同的优先级。

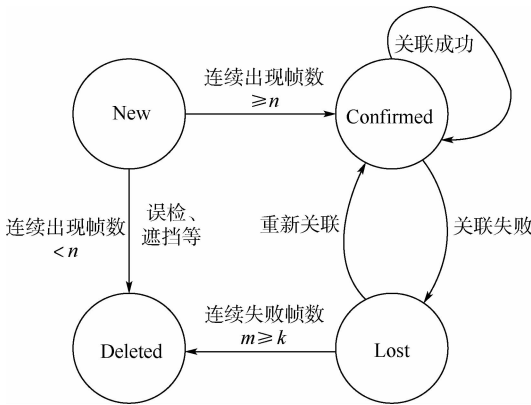


图 4 轨迹生命周期状态转换

Fig. 4 Trajectory life cycle state transition

算法 1 描述了基于轨迹状态进行分层关联的整体流程。

#### 算法 1 基于轨迹状态的分层关联模型。

输入:轨迹集合  $T$ (长度  $N$ )、当前帧检测目标集合  $D$ (长度  $M$ )、帧号  $num$ 、间隔参数  $g$ 。

输出:轨迹与目标检测框的关联结果,即  $matches$ 、 $unmatched\_tracks$ 、 $unmatched\_detections$ 。

procedure  $match(T, D, num, g)$

- 1: 初始化已关联集合  $R \leftarrow \Phi$
- 2: 初始化未关联的轨迹  $Q \leftarrow T$
- 3: 初始化未关联的检测框  $H \leftarrow D$
- 4: //根据轨迹当前的状态划分为 3 个不同的集合
- 5: 划分未关联的轨迹段  $Q$  为  $Q_{confirmed}$ 、 $Q_{unconfirmed} \leftarrow$

- 6: //计算关联矩阵,短期使用运动特征,每隔  $g$  帧使用表观特征进行校正
- 7: if  $num \% g == 0$  then
- 8:     使用式(4)计算距离矩阵  $C = [d_1(i, j)]$
- 9: else
- 10:    按图 3 所示计算距离矩阵  $C = [d_2(i, j)]$
- 11: end if
- 12: //处于 Confirmed 状态的轨迹基于朝向级联匹配
- 13:  $(Y, G, U) \leftarrow match\_orientation(Q_{confirmed}, H, C)$
- 14: update  $R \leftarrow R \cup Y$  //已正确关联的轨迹
- 15: update  $H \leftarrow U, Q_{other} \leftarrow G$  //未关联
- 16: //处于其他状态的轨迹,基于朝向级联匹配
- 17:  $(Y, G, U) \leftarrow match\_orientation(Q_{unconfirmed}, H, C)$
- 18: update  $R \leftarrow R \cup Y, H \leftarrow U, Q \leftarrow Q \cup G$
- 19: //未关联的轨迹,基于重叠面积尝试关联
- 20: 基于  $Q_{other}$  和  $H$  的重叠面积计算矩阵  $C$
- 21:  $(Y, G, U) \leftarrow min\_iou\_matching(Q_{other}, H, C)$
- 22: update  $R \leftarrow R \cup Y, H \leftarrow U, Q \leftarrow Q \cup G$
- 23: return  $(R, Q, H)$

end procedure

算法 1 融合运动特征和表观特征的优势,在短期跟踪中使用基于卡尔曼滤波得到的预测位置,同时为了防止长时间遮挡等问题引起的轨迹偏移,每隔  $g$  帧基于重识别得到的表观特征进行一次关联校正,兼顾算法的实时性和准确性,主要基于不同的代价矩阵来实现(见算法 1 中第 6 ~ 11 行)。在具体的分层关联中,根据优先级依次为处于 Confirmed 和其他状态的轨迹寻找正确的关联目标,而在具体关联方式上,采用基于算法 2 的朝向级联策略。对于处于 Confirmed 状态但未能关联的轨迹,通过检测框和预测位置之间的重叠面积进行一次关联尝试,防止目标被遮挡而引起的关联错误。

算法 2 描述了基于朝向进行关联的具体方式。

#### 算法 2 基于朝向的分层轨迹关联模型。

输入:未关联的轨迹集合  $T$ ,未关联的目标集合  $H$ ,关联矩阵  $C$ 。

输出:轨迹与目标检测框的关联结果,即  $matches$ 、 $unmatched\_tracks$ 、 $unmatched\_detections$ 。

procedure  $match\_orientation(T, H, C)$

- 1: 初始化已关联集合  $R \leftarrow \emptyset$
- 2: 根据朝向分类器获取集合  $T$  和集合  $G$  中每个行人的朝向
- 3: //相同朝向优先关联,匈牙利算法
- 4: 根据朝向是否相同更新矩阵  $C$ , 朝向不同的目标矩阵中的值无限大
- 5:  $[t_1, d_1] \leftarrow \text{min\_cost\_matching}(C_{\text{same}}, T, H)$
- 6:  $\text{update } T \leftarrow T - t_1, H \leftarrow H - d_1, R \leftarrow R \cup [t_1, d_1]$
- 7: //相近朝向组合进行关联(正和左、正和右、背和左、背和右)
- 8: 根据相近的朝向更新矩阵  $C$ , 并删除已关联上的轨迹
- 9:  $[t_2, d_2] \leftarrow \text{min\_cost\_matching}(C_{\text{close}}, T, H)$
- 10:  $\text{update } T \leftarrow T - t_2, H \leftarrow H - d_2, R \leftarrow R \cup [t_2, d_2]$
- 11: //相反朝向组合进行关联(正和背、左和右)
- 12: 根据相反朝向更新矩阵  $C$ , 并删除已关联上的轨迹
- 13:  $[t_3, d_3] \leftarrow \text{min\_cost\_matching}(C_{\text{opposite}}, T, H)$
- 14:  $\text{update } T \leftarrow T - t_3, H \leftarrow H - d_3, R \leftarrow R \cup [t_3, d_3]$
- 15:  $\text{return}(R, T, H)$

end procedure

算法2是算法1的内层嵌套,用于相同状态轨迹的关联。在分别为 New、Confirmed 和 Lost 三种状态的轨迹寻找匹配目标时,考虑到相邻帧之间目标朝向几乎不会突变的规律,将轨迹与检测框之间的朝向划分为相同、相近(正和左、正和右、背和左、背和右)、相反(正和背、左和右)3种不同的组合。朝向相同的轨迹和检测框在匹配时应当赋予优先权,而相近和相反的组合优先级逐渐降低。在每一种朝向组合进行匹配时,通过代价矩阵基于匈牙利算法实现二分图之间的最大匹配,找到全局最优解。

综上所述,本节提出的分层轨迹关联模型核心思想是:在关联模型中优先考虑最近出现过的确定轨迹,优先考虑相同或相近朝向的关联组合。该模型共包括2层:第1层考虑被关联轨迹当前所处的状态,优先对已确认轨迹进行匹配,未能与确认轨迹关联的目标检测框再与 New 和 Lost 状态的轨迹尝试匹配,对于只在当前帧未能关联的轨迹,则用卡尔曼滤波预测的位置代替;第2层考虑被关联轨迹与目标检测框之间的朝向关系,通过对不同朝向组合赋予不同的优先级,使得模型优先考虑朝向相近的目标,从而克服相似目标相向而行时的目标交换,这是一种简单但非常有效的时空约束。

## 2.3 单相机轨迹生成方法

利用分层轨迹关联模型,可以有效实现单相机内多个行人的实时跟踪,得到单相机内的目标完整轨迹。方法整体流程如图5所示。

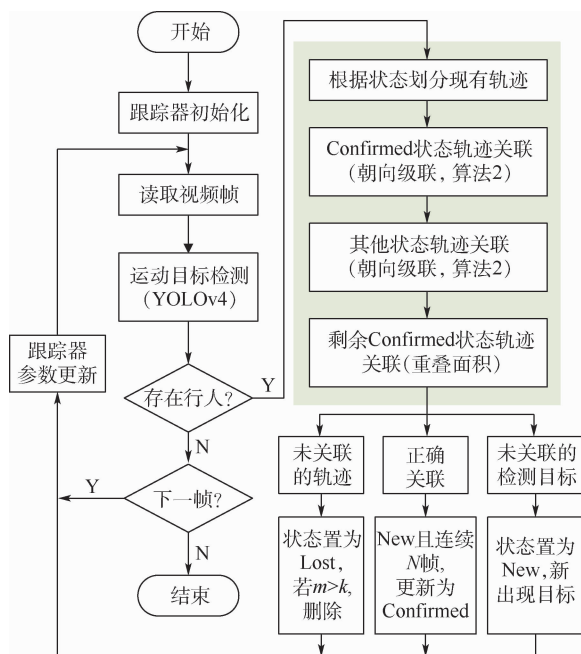


图5 单相机轨迹生成方法流程

Fig. 5 Process of single camera trajectory generation

方法过程如下:

1) 通过多目标检测算法获得每帧图像中所有行人目标的矩形包围框,本文使用同时兼顾检测准确率和实时性的YOLOv4检测算法。

2) 利用卡尔曼滤波预测下一帧中行人的位置,得到运动特征,并通过行人重识别网络得到每一个检测框的表观特征表示。

3) 对于检测到的每一个行人,对其进行朝向判别,得到朝向标签。

4) 通过分层轨迹关联模型(算法1)将现有轨迹和检测框之间进行关联(流程图中的阴影区域)。

5) 根据关联结果,设置每一条轨迹的当前状态,更新相应参数。

6) 提取下一帧,重复以上步骤进行关联,直到没有下一帧,结束。

## 3 基于重识别特征的双向匹配轨迹关联方法

基于单相机轨迹生成方法,可以得到每个行人目标在单个相机内的运动轨迹,为跨相机间的目标轨迹关联提供基础输入。为将不同相机下的行人轨迹联系起来,本节进一步提出一种基于重识别特征的双向匹配轨迹关联方法。其基本思想

是:基于单相机目标轨迹生成过程中记录的每个行人目标的重识别特征(由于跨相机之间的朝向不确定性,采用不同朝向分支得到的行人特征向量),取所有特征向量的平均值作为整条轨迹的表示。在具体的 2 个相机之间进行轨迹关联时,通过引入前向后向竞争的双向匹配机制,增强轨迹关联的可靠性。

假设在相机 1 中得到的行人轨迹集合为  $T = \{t(1), t(2), \dots, t(n)\}$ ,  $t(i)$  表示第  $i$  时刻的轨迹,  $i = 1 \sim n$ ; 而在相机 2 中得到的行人轨迹集合为  $G = \{g(1), g(2), \dots, g(m)\}$ ,  $g(j)$  表示第  $j$  时刻的轨迹,  $j = 1 \sim m$ 。则每个相机内的轨迹集合可以看做是行人重识别问题模型中的被检索库 gal-

lery。因此,可以将跨相机轨迹关联转化为行人重识别的检索问题。本文主要基于前向后向竞争的思想,提出一种双向匹配的轨迹关联方法。

双向指先将集合  $T$  中的每一条轨迹  $t(i)$  看作检索对象 query, 在集合  $G$  中查找与其最相似的轨迹  $g(j)$ ; 然后反过来, 对于集合  $G$  中的每一条轨迹  $g(j)$ , 将其看作检索对象 query, 在集合  $T$  中查找与其最相似的轨迹  $t(k)$ 。在 2 次重识别过程中, 如果得到的  $t(k)$  与  $t(i)$  是同一个目标轨迹, 则认为 2 条轨迹是完全匹配的, 标记为已匹配; 反之, 如果前向后向匹配得到的结果不同, 则认为匹配可能性较低, 视为单独的轨迹, 如图 6 所示。

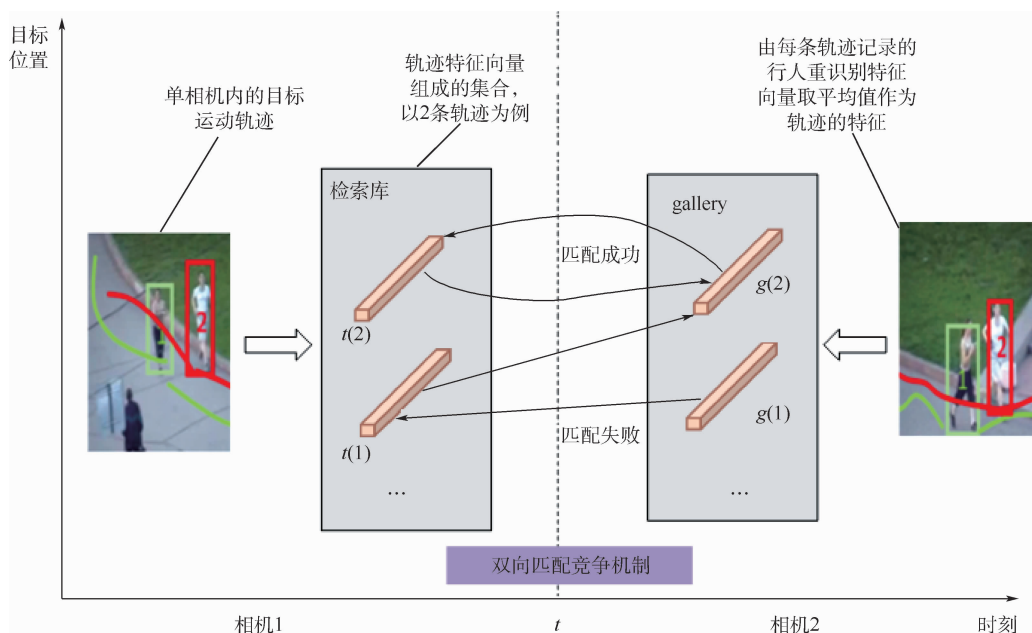


图 6 跨相机目标轨迹关联示意图

Fig. 6 Schematic diagram of cross camera target trajectory association

在具体实现上,可以通过将集合  $T$  与集合  $G$  的平均值特征向量计算欧氏距离,得到一个  $n \times m$  的距离矩阵,前向匹配过程相当于矩阵中的一行,后向匹配过程相当于矩阵中的一列。当该距离矩阵的行最小值等于所在列的列最小值时,说明双向匹配是成功的,否则视为匹配失败。

综上所述,基于重识别特征的双向匹配轨迹关联方法,实际上将跨相机目标轨迹关联的问题转化为了 2 次行人重识别的检索过程,其原理和实现相对简单,但关联准确性依赖于行人重识别的特征可信度及单相机轨迹的正确性。实验证明,对于长期跟踪得到的稳定轨迹,其平均重识别特征是比较可靠的,因为平均特征减弱了运动过程中的短暂遮挡所引起的特征变化,能够较好地

## 4 实验结果及分析

### 4.1 行人重识别实验结果

行人重识别有效的特征表示和泛化性能是实现目标轨迹有效关联的基础。本文基于 market-1501 数据集<sup>[16-17]</sup>,使用 Rank- $k$  和 mAP 两种评价指标对朝向约束行人重识别方法进行了定量实验。

在实验中,以 ResNet-50 作为骨干网,conv2\_x 及之前的所有层作为共享模块,其余网络层作为分支模块(不共享权重)。对于所有的输入数据,将图像大小重新设置为  $256 \times 128$ ,将每一批次的 batch size 设置为 128,其中包括 32 个行人和每个行人的 4 张图像样本( $N = 128, P = 32, K = 4$ )。网络共训练 120 代(epoch),初始学习率为  $3.5 \times$

$10^{-5}$ ,采用学习率预热(warmup)策略在前10代将学习率逐渐增加到 $3.5 \times 10^{-4}$ ,随后在第35、75、95代时分别将学习率降低为原来的0.1倍,最终得到的特征向量为2048维。

在market-1501数据集上得到的测试结果及与其他方法的对比如表1所示。可知,本文方法在market-1501数据集上的Rank-1和Rank-5准确率分别达到了94.21%和97.96%,mAP为83.41%,优于对比方法。

图7为采用BaseLine和本文方法的行人重识别效果。最左侧代表查询图像query,随后依次为其前5张检索结果,相似度从高到低排列,绿色边框和红色边框分别代表识别正确、识别错误。可以看出,相比于BaseLine实验,本文方法在行人重识别上准确率更高;同时由于朝向约束的影响,匹配时能够优先识别具有相同朝向的行人,对实

现目标轨迹关联非常有益。

表1 行人重识别实验结果对比

Table 1 Comparison of experimental results of pedestrian re-identification %

方法	Rank-1	Rank-5	mAP
PCB <sup>[18]</sup>	92.3	97.2	77.4
AlignedReID <sup>[19]</sup>	91.8	97.1	79.3
PIE <sup>[20]</sup>	87.33	95.56	69.25
GLAD <sup>[21]</sup>	89.9		73.9
Spindle <sup>[22]</sup>	76.9	91.5	
HA-CNN <sup>[23]</sup>	91.2		75.7
TriHard <sup>[15]</sup>	86.67	93.38	81.07
HPM <sup>[24]</sup>	94.2	97.5	82.7
PGR <sup>[25]</sup>	93.87	97.74	77.21
OSCNN <sup>[26]</sup>	83.9		73.5
本文方法	94.21	97.96	83.41



图7 行人重识别效果直观展示

Fig. 7 Visual display of pedestrian re-identification effect

#### 4.2 目标轨迹关联实验结果

MOT系列数据集是在多目标跟踪和轨迹关联相关研究中使用最为广泛的数据集,最常用的是MOT16数据集和MOT17数据集<sup>[27]</sup>,其由14个不同角度、不同自然条件下拍摄的視頻组成,训练集和测试集各7个,其中的行人数量和各种因素互不相同,是比较有挑战性的数据集。

本文选用了比较有代表性的5个评价指标:MOTA、MOTP、IDP、IDR和IDF1。

针对2.2节提出的分层轨迹关联模型,在实

验中通过多次测试,得到其主要的参数分配如下:基于卡尔曼滤波得到的不确定性估计参数 $t_1 = 9.4877$ ,表观特征距离阈值 $t_2 = 0.5$ ,判定新目标出现的连续观察帧数 $n = 5$ ,判定轨迹消失的最大连续帧数 $k = 100$ ,使用表观特征进行校正的间隔参数 $g = 3$ 。

通过以上参数,使用MOT数据集中给定的目标检测框进行实验,得到的实验结果与常见的在线多目标跟踪算法SORT<sup>[7]</sup>、DeepSORT<sup>[6]</sup>、MOTDT<sup>[8]</sup>、EAMTT<sup>[28]</sup>、Tractor<sup>[11]</sup>、FairMOT<sup>[12]</sup>等进行比较,结果如表2和表3所示。

表2 轨迹关联方法在 MOT16 数据集上的性能比较

Table 2 Performance comparison of trajectory association methods on MOT16 dataset %

方法	MOTA	MOTP	IDP	IDR	IDF1
EAMTT <sup>[28]</sup>	52.5	78.8		31.5	42.4
JDE <sup>[10]</sup>	64.4				55.8
MOTDT <sup>[8]</sup>	47.6	50.9		40.3	50.9
SORT <sup>[7]</sup>	33.4	72.1	54.2	29.6	26.5
DeepSORT <sup>[6]</sup>	60.3	79.1	76.9	31.8	37.6
本文方法	60.8	78.6	78.2	45.6	54.7

表3 轨迹关联方法在 MOT17 数据集上的性能比较

Table 3 Performance comparison of trajectory association methods on MOT17 dataset %

方法	MOTA	MOTP	IDP	IDR	IDF1
FairMOT <sup>[12]</sup>	67.5	84.7	75.9	45.3	69.8
Tractor <sup>[11]</sup>	53.5				52.3
SORT <sup>[7]</sup>	31.6	69.8	45.2	37.5	43.2
DeepSORT <sup>[6]</sup>	44.1	83.2	66.6	43.0	52.3
本文方法	57.6	79.7	75.8	47.2	53.5

从表2可以看出,在参与实验的各种方法中,本文提出的目标轨迹关联方法取得了较好的效果,指标 MOTA、MOTP、IDF1 均是次优, IDP 和 IDR 为最优。

从表3可以看出,本文方法在关联 MOTA 准确率和 MOTP 召回率上分别为 57.6% 和 79.7%, 略低于 FairMOT 方法,但 IDR 指标优于 FairMOT 方法,说明本文方法能够减少频繁 ID 交换和错误关联。

在 MOT17 数据集的视频数据上统计了方法处理的帧率,本文方法的帧率达到了 19.6 帧/s, 满足近实时性的要求,在大多数场景中可用。

图8为本文方法与 DeepSORT 方法在经典场景中的轨迹关联效果对比。可以看出,本文方法的 ID 增长速度较慢,在很大程度上减少了频繁 ID 变换,说明轨迹之间的关联特征是比较有区分度的,能够在目标发生误检和丢失的情况下进行重新关联,大大降低了轨迹碎片的产生。在该场景的第 337 帧,2 个行人目标相向而行,在第 351 帧时,2 个行人之间由于互相遮挡,导致检测器未能将被遮挡行人检测出来,发生了漏检现象。此时,对于 DeepSORT 方法,如图 8(a)、图 8(c) 和图 8(e) 所示,当 2 个行人从遮挡状态逐渐分离之后,发现其 ID 发生了互换,轨迹匹配错误;而如图 8(b)、图 8(d) 和图 8(f) 所示,本文方法在 2 个行人逐渐分离之后,其 ID 仍旧能够正确关联。这一方面是由于表观特征的泛化性能有所增强,另一方面是因为在分层轨迹关联模型的约束下,由于相同朝向的行人在匹配时具有较高的优先级,从而避免了 ID 之间的错误关联。

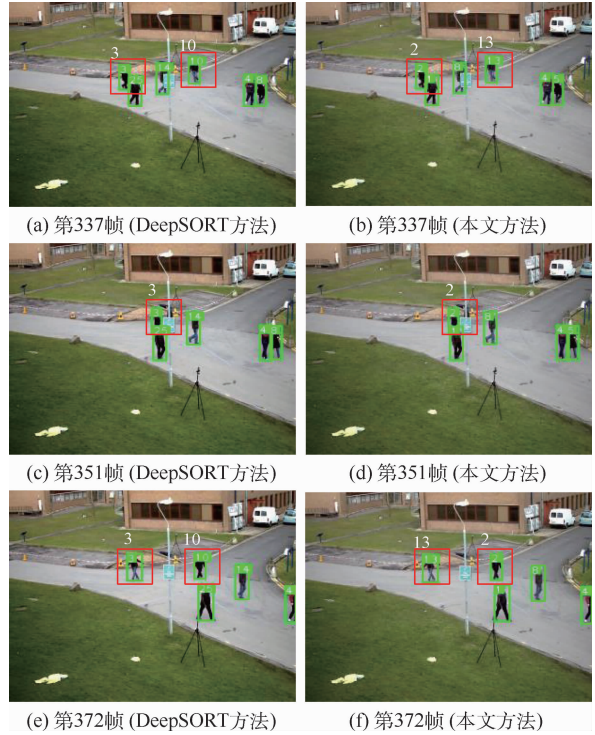


图8 轨迹关联效果对比

Fig. 8 Comparison of trajectory association effect

## 5 结论

本文针对轨迹关联中时空约束缺失和表观特征泛化能力不足的问题,提出了一种基于朝向约束和重识别特征的目标完整轨迹生成方法。其中,所提出的行人重识别方法在 market-1501 数据集上达到了 94% 的 Rank-1 准确率,减少了朝向不同带来的识别误差。同时,所提出的轨迹关联方法在 MOT 数据集上 ID 相关指标优于多种方法,说明在朝向约束的作用下能够有效减少频繁的 ID 交换和错误关联。

## 参考文献 (References)

- [1] GHEISSARI N, SEBASTIAN T B, HARTLEY R. Person reidentification using spatiotemporal appearance [C] // 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2006: 1528-1535.
- [2] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [3] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 779-788.
- [4] RISTANI E, SOLERA F, ZOU R, et al. Performance measures and a data set for multi-target, multi-camera tracking [C] // European Conference on Computer Vision. Berlin: Springer, 2016:

- 17-35.
- [ 5 ] LEAL-TAIXE L, CANTON-FERRER C, SCHINDLER K. Learning by tracking: Siamese CNN for robust target association [ C ] // 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops. Piscataway: IEEE Press, 2016: 418-425.
- [ 6 ] WOJKE N, BEWLEY A, PAULUS D. Simple online and real-time tracking with a deep association metric [ C ] // 2017 IEEE International Conference on Image Processing. Piscataway: IEEE Press, 2017: 3645-3649.
- [ 7 ] BEWLEY A, GE Z Y, OTT L, et al. Simple online and realtime tracking [ C ] // 2016 IEEE International Conference on Image Processing. Piscataway: IEEE Press, 2016: 3464-3468.
- [ 8 ] CHEN L, AI H Z, ZHUANG Z J, et al. Real-time multiple people tracking with deeply learned candidate selection and person re-identification [ C ] // 2018 IEEE International Conference on Multimedia and Expo. Piscataway: IEEE Press, 2018: 1-6.
- [ 9 ] RISTANI E, TOMASI C. Features for multi-target multi-camera tracking and re-identification [ C ] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 6036-6046.
- [ 10 ] WANG Z D, ZHENG L, LIU Y X, et al. Towards real-time multi-object tracking [ C ] // European Conference on Computer Vision. Berlin: Springer, 2020: 107-122.
- [ 11 ] BERGMANN P, MEINHARDT T, LEAL-TAIXÉ L. Tracking without bells and whistles [ C ] // 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE Press, 2019: 941-951.
- [ 12 ] ZHAN Y, WANG C, WANG X, et al. A simple baseline for multi-object tracking [ C ] // 2020 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020.
- [ 13 ] HOU Y, ZHENG L, WANG Z, et al. Locally aware appearance metric for multi-target multi-camera tracking [ C ] // 2020 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020.
- [ 14 ] SCHROFF F, KALENICHENKO D, PHILBIN J. FaceNet: A unified embedding for face recognition and clustering [ C ] // 2015 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2015: 815-823.
- [ 15 ] HERMANS A, BEYER L, LEIBE B. In defense of the triplet loss for person re-identification [ EB/OL ]. ( 2017-03-22 ) [ 2021-02-01 ]. <https://arxiv.org/abs/1703.07737>.
- [ 16 ] LI D W, ZHANG Z, CHEN X T, et al. A richly annotated dataset for pedestrian attribute recognition [ EB/OL ]. ( 2016-03-23 ) [ 2021-02-01 ]. <https://arxiv.org/abs/1603.07054>.
- [ 17 ] ZHENG L, SHEN L Y, TIAN L, et al. Scalable person re-identification: A benchmark [ C ] // 2015 IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2015: 1116-1124.
- [ 18 ] SUN Y F, ZHENG L, YANG Y, et al. Beyond part models: Person retrieval with refined part pooling ( and a strong convolutional baseline ) [ C ] // European Conference on Computer Vision. Berlin: Springer, 2018: 480-496.
- [ 19 ] ZHANG X, LUO H, FAN X, et al. AlignedReID: Surpassing human-level performance in person re-identification [ EB/OL ]. ( 2017-11-22 ) [ 2021-02-01 ]. <https://arxiv.org/abs/1711.08184>.
- [ 20 ] ZHENG L, HUANG Y J, LU H C, et al. Pose-invariant embedding for deep person re-identification [ J ]. IEEE Transactions on Image Processing, 2019, 28 ( 9 ): 4500-4509.
- [ 21 ] WEI L H, ZHANG S L, YAO H T, et al. GLAD: Global-local-alignment descriptor for scalable person re-identification [ J ]. IEEE Transactions on Multimedia, 2019, 21 ( 4 ): 986-999.
- [ 22 ] ZHAO H Y, TIAN M Q, SUN S Y, et al. Spindle Net: Person re-identification with human body region guided feature decomposition and fusion [ C ] // 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 907-915.
- [ 23 ] LI W, ZHU X T, GONG S G. Harmonious attention network for person re-identification [ C ] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 2285-2294.
- [ 24 ] FU Y, WEI Y C, ZHOU Y Q, et al. Horizontal pyramid matching for person re-identification [ EB/OL ]. ( 2018-11-10 ) [ 2021-02-01 ]. <https://arxiv.org/abs/1804.05275v3>.
- [ 25 ] ZHAO L M, LI X, ZHUANG Y T, et al. Deeply-learned part-aligned representations for person re-identification [ C ] // 2017 IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2017: 3239-3248.
- [ 26 ] ZHENG L, YANG Y, HAUPTMANN A G. Person re-identification: Past, present and future [ EB/OL ]. ( 2016-10-10 ) [ 2021-02-01 ]. <https://arxiv.org/abs/1610.02984>.
- [ 27 ] MILAN A, LEAL-TAIXE L, REID I, et al. MOT16: A benchmark for multi-object tracking [ EB/OL ]. ( 2016-03-02 ) [ 2021-02-01 ]. <https://arxiv.org/abs/1603.00831>.
- [ 28 ] SANCHEZ-MATILLA R, POIESI F, CAVALLARO A. Online multi-target tracking with strong and weak detections [ C ] // European Conference on Computer Vision. Berlin: Springer, 2016: 84-99.

## Target trajectory association method based on orientation constraint and re-identification feature

AI Mingjing<sup>\*</sup>, SHAN Guozhi, LIU Penggao, YANG Penggang

(State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing 100083, China)

**Abstract:** Target trajectory association method based on detection association and deep learning is one of the research hotspots in the field of computer vision. However, due to the lack of effective space-time constraints in the design of existing methods, and the insufficiency of generalization ability of target apparent features, recognition errors will occur in the case of obvious differences in target orientation, which will lead to frequent ID switching and error association. To solve this problem, we propose a target trajectory association method based on orientation constraint and re-identification feature. This paper introduces pedestrian orientation discrimination into pedestrian re-identification, and presents a pedestrian re-identification network model with orientation constraint, which improves the representation ability of target features. Combining the spatial and temporal characteristics of target orientation, position information from Kalman filter and overlap area, a hierarchical trajectory association model based on orientation constraint is proposed to obtain the target trajectory in a single camera. A simple and effective bi-directional competitive matching mechanism is introduced to implement effective association of target trajectories in the cross camera scene. Experimental results show that the proposed method achieves a competitive level on MOT datasets. It can reduce frequent ID exchange, and can effectively solve the problem of error association when similar objects are moving towards each other. Meanwhile, with a frame rate of 19.6 frame/s, it can satisfy the requirements of near real-time scene.

**Keywords:** trajectory association; pedestrian re-identification; multi-target tracking; orientation constraint; spatiotemporal constraint

**Received:** 2021-02-23; **Accepted:** 2021-03-26; **Published online:** 2021-06-22 15:52

**URL:** [kns.cnki.net/kcms/detail/11.2625.V.20210622.1225.001.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210622.1225.001.html)

**Foundation items:** National Natural Science Foundation of China (61872024); National Key R & D Program of China (2018YFB2100601)

**\* Corresponding author.** E-mail: amj@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2020.0726

# 基于 $\mu$ CT 图像的 PVC 泡沫微观试样 几何模型生成算法

周勇, 薛斌, 郭昀鑫, 王人鹏\*

(同济大学 土木工程学院, 上海 200092)

**摘 要:** 在泡沫微观结构的数值模拟中, 泡沫空腔的几何特征和排列状态对计算效率及计算结果有着重要的影响, 基于前进面搜索几何构造算法和 Laguerre 划分算法, 提出了一种新的 PVC 泡沫微观试样几何模型生成算法。从  $\mu$ CT 扫描图像重构泡沫的真实几何模型, 测量泡沫空腔的几何特征及体积分布规律, 将测量得到的泡沫空腔体积转化为球体, 并通过前进面搜索几何构造算法投入空间, 通过 Laguerre 划分算法将空间球体进行区域划分, 赋予壁厚参数, 构建出闭孔 PVC 泡沫的微观几何模型。所建立的模型在微观几何特征上与实际材料符合较好。

**关键词:** CT 扫描; 前进面法; Laguerre 划分; 闭孔泡沫; 微观模型

**中图分类号:** V254.1<sup>+</sup>9; TB332

**文献标志码:** A

**文章编号:** 1001-5965(2022)06-0968-11

泡沫材料由于质量轻、冲击能量吸收、隔热隔音等优势正越来越多地被用于汽车、航空航天、船舶等领域。随着泡沫的大量使用, 其机械性能引起了广泛的关注。泡沫的机械性能主要取决于制备材料的性能、泡沫相对密度及泡沫的微观结构, 如空腔尺寸、空腔几何形状、空腔壁厚等<sup>[1]</sup>。许多学者通过建立空间填充多面体模型对泡沫进行了研究。Gibson 和 Ashby<sup>[2]</sup>建立了立方晶胞模型, Mills 和 Zhu<sup>[3]</sup>建立了开尔文十四面体模型, Weaire 和 Phelan<sup>[4]</sup>建立了 Weaire-Phelan 模型, 由于这些模型具备简单统一的几何形状, 可以采用经典的梁壳理论进行分析, 但是其过于简单, 对于空腔的排列过于理想化, 忽略了泡沫排列的随机性, 无法代表真实的泡沫微观结构。随着科技的发展, CT 技术已经被用来获取泡沫内部的真实形状并重构出闭孔泡沫的真实模型<sup>[5-14]</sup>。通过图像建立泡沫模型的主要优点在于: 这些模型与扫描的泡沫样本具有相同的几何形状, 适合于探究在固定形状下的泡沫破坏机理。但是除了计算量大

之外, 通过图像直接重构三维模型还具有一定的局限性, 即无法生成任意方式的微观结构, 因此不能综合考虑每个空腔几何特征的影响, 且扫描所得到的参数都是固定的, 在进行诸如壁厚、空腔体积等参数分析时难度较大, 为此许多学者基于 Voronoi 模型进行闭孔泡沫的微观结构生成<sup>[15-19]</sup>。Voronoi 模型类似于泡沫形成的实际过程, 气泡从种子点向外匀速生长, 当气泡与其他气泡接触时生长停止, 通过 Voronoi 模型得到的泡沫微观结构在每个单元的平均面数、每个面的平均边数等几何特征上与实际结构较为符合<sup>[20]</sup>。Voronoi 模型虽然简单, 但是不能构造出遵循指定大小分布的空腔特征, 为使空腔体积大小的分布更加符合实际, 学者们提出了利用 Laguerre 划分进行泡沫微观模型的重构<sup>[1, 21-22]</sup>。Laguerre 划分通过对 Voronoi 划分赋予一定的权重, 即使得气泡从种子点向外以一定的速率比进行生长, 假设以一组随机紧密堆积的球体中心作为种子点, 并将这些球体的半径作为权重, 则构造出的 Laguerre 划分模

收稿日期: 2020-12-31; 录用日期: 2021-03-07; 网络出版时间: 2021-03-25 16:35

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20210324.1349.002.html](http://kns.cnki.net/kcms/detail/11.2625.V.20210324.1349.002.html)

基金项目: 国家重点研发计划(2016YFB1200601-16)

\*通信作者. E-mail: renpengwang@126.com

**引用格式:** 周勇, 薛斌, 郭昀鑫, 等. 基于  $\mu$ CT 图像的 PVC 泡沫微观试样几何模型生成算法[J]. 北京航空航天大学学报, 2022, 48(6): 968-978. ZHOU Y, XUE B, GUO Y X, et al. An algorithm for generating geometric models of microscopic specimens of PVC foam based on  $\mu$ CT images[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(6): 968-978 (in Chinese).

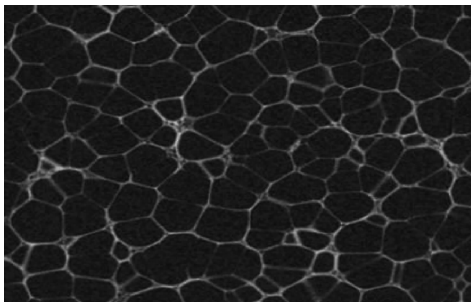
型拥有近似于原始球体大小分布的特征<sup>[1]</sup>。

本文基于  $\mu$ CT 扫描图像,利用 Dragonfly 软件进行了泡沫微观结构的三维重构,统计出实际的空腔体积、空腔表面积、空腔壁厚等几何参数,将空腔体积通过一系列处理等效为球体后,采用前进面搜索几何构造算法进行球体的随机紧密堆积,并利用 Laguerre 划分算法重构出聚氯乙烯(PVC)泡沫微观结构几何模型,所建立的闭孔 PVC 泡沫数值模型的几何特征接近于实际材料微观结构。

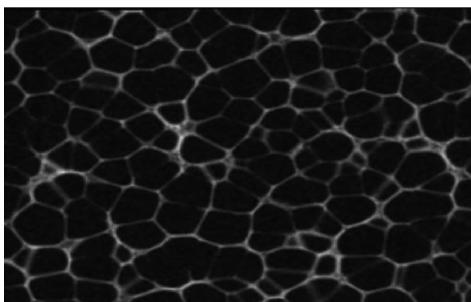
### 1 实验材料与方法

#### 1.1 分析方法

本文研究的样本为 Diab 的 Divinycell H100 泡沫,该泡沫现今被广泛应用于航天航空、船舶等行业。H100 是由 PVC 制成的密度为  $100 \text{ kg/m}^3$  的闭孔泡沫。 $\mu$ CT 扫描图像由 X 射线显微镜(Xradia 520 Versa)获取。实验所采用的射线源电压为 70 kV、电流为  $86 \text{ }\mu\text{A}$ ,光学放大倍率为 0.4。原始试样尺寸为  $10 \text{ mm} \times 10 \text{ mm} \times 10 \text{ mm}$ ,图片大小为  $1\ 004 \times 1\ 024$ ,像素大小为  $9.899\ 3 \text{ }\mu\text{m}$ 。扫描试样在厚度方向上共生成了 1 016 个二维断层图像,为了减小计算量同时确保统计效果,提取边长为  $6\ 345 \text{ }\mu\text{m} \times 5\ 137 \text{ }\mu\text{m} \times 4\ 563 \text{ }\mu\text{m}$  的长方体进行处理。为了消除  $\mu$ CT 扫描中原有噪点的影响,采取高斯滤波对噪点进行消除,图 1 为滤波前后的扫描图像。



(a) 原始图像



(b) 滤波后图像

图 1 滤波前后二维断层图像

Fig. 1 Two-dimensional tomographic images before and after filtering

对于滤波处理后的图像,通过简单的阈值划分可以划分出泡沫的孔壁与空腔,此时虽然没有大量噪点的影响,但是仍然会存在孔壁缺失的情况,不利于之后的连通性分析及单个空腔的提取。为此,对所得到的黑白图像进行欧几里得变换,将到孔壁的距离分配给每个像素,理想情况下在细胞中心会产生局部最大值。利用 Sobel 滤波建立空腔的边界,从局部最大值处向孔壁扩张,扩张到 Sobel 滤波形成的边界时停止,从而形成独立的空腔,图 2 为划分后的图像。

对处理后的图像模型进行连通性分析,分别提取出空腔与孔壁。其中,空腔用于测量空腔体积率、空腔体积大小、表面积,孔壁用于测量壁厚大小。重构后的空腔与孔壁模型如图 3 所示。

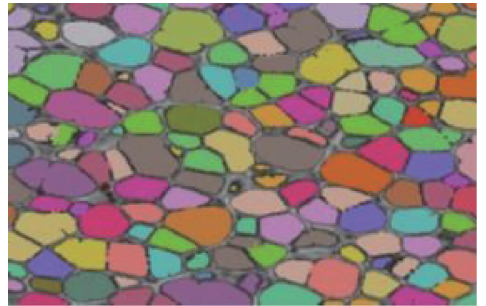
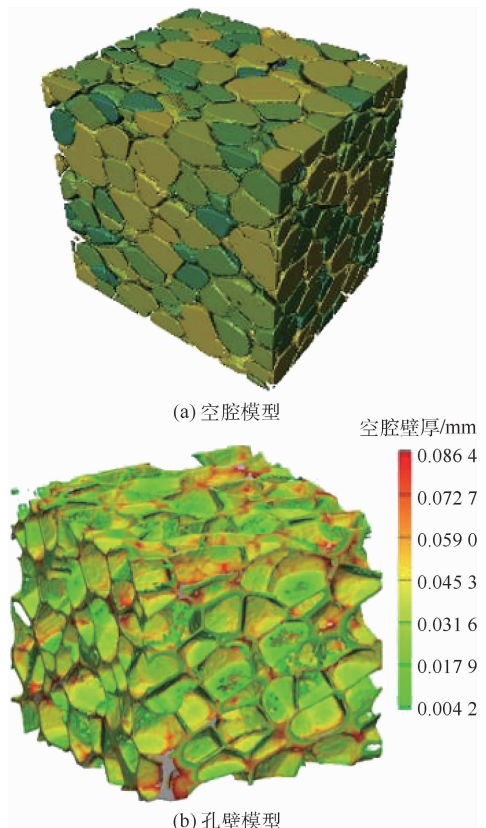


图 2 划分后空腔

Fig. 2 Cavity after division



(a) 空腔模型

(b) 孔壁模型

图 3 三维重构模型

Fig. 3 Three-dimensional reconstruction model

## 1.2 几何特征分析

对处理后的三维空腔图像进行体积表面积等几何特征统计时,考虑到边缘地带泡沫空腔被切割后变得不完整的情况,将边缘地区的泡沫剔除后进行统计。在进行泡沫空腔率统计时,考虑到空腔排列的周期性特征,对整体进行分析。测得泡沫的空腔率为 72.03%。

经统计,样品中共存在空腔 3 130 个,对这些空腔进行统计分析发现,空腔体积最大值为  $2.040 3 \times 10^8 \mu\text{m}^3$ ,最小值为  $2.207 9 \times 10^6 \mu\text{m}^3$ ,平均值为  $3.410 1 \times 10^7 \mu\text{m}^3$ ,空腔体积大小分布服从  $\alpha = 9.702 6 \times 10^7$ 、 $c = 1.507 9$ 、 $k = 4.958 0$  的 Burr 分布;空腔表面积最大值为  $3.232 3 \times 10^6 \mu\text{m}^2$ ,最小值为  $1.030 1 \times 10^5 \mu\text{m}^2$ ,平均值为  $6.234 2 \times 10^5 \mu\text{m}^2$ ,空腔表面积大小分布服从  $\alpha = 6.644 8 \times 10^5$ 、 $c = 2.663 1$ 、 $k = 1.503 8$  的 Burr 分布;空腔体积/表面积最大值为  $85.388 6 \mu\text{m}$ ,最小值为  $14.049 9 \mu\text{m}$ ,平均值为  $48.747 4 \mu\text{m}$ ,空腔体积/表面积大小分布大致服从  $\mu = 48.747 4$ 、 $\sigma = 13.502 6$  的正态分布。对于各种参数的分布均采用概率密度函数(PDF)及累积分布函数(CDF)来进行描述。具体的分布数据及参数拟合数据如图 4 所示。

将重构后的空腔壁模型导入到 3-matic Research 中,对其进行壁厚分析。对空腔壁厚进行分析时,不对缺失的表面进行补面等措施,防止造成孔壁厚度与实际不符的情况,从而影响统计结果。统计后发现,空腔壁厚大小分布大致服从  $\mu = 36.297 8$ 、 $\sigma = 13.637 1$  的正态分布。空腔壁厚最大值为  $86.4 \mu\text{m}$ ,最小值为  $4.7 \mu\text{m}$ ,平均值为  $36.3 \mu\text{m}$ 。具体的空腔壁厚分布数据和参数拟合数据如图 5 所示。

由于本文泡沫微观结构的重构方法是先在指定空间内进行球体的紧密堆积,再采用 Laguerre 划分算法进行以半径为权重的划分,划分后的空腔率会大于原始球体的填充率。为使划分后的空腔率与真实的空腔率更加符合,需要在球体紧密堆积时缩小球体的半径。Matzke<sup>[23]</sup>发现每个空腔的面数为 11 ~ 17,平均值为 13.7,且超过 2/3 的细胞面是五边形,为了计算的简便,假设每个空腔都是正多面体,与之最相近的正多面体为正十二面体,正十二面体的面均是正五边形。正十二面体的内切球体积占正十二面体体积的体积分数为 74.79%,即在生成填充球体时,生成球体体积为原始空腔体积大小的 74.79%。对紧密堆积的球体进行 Laguerre

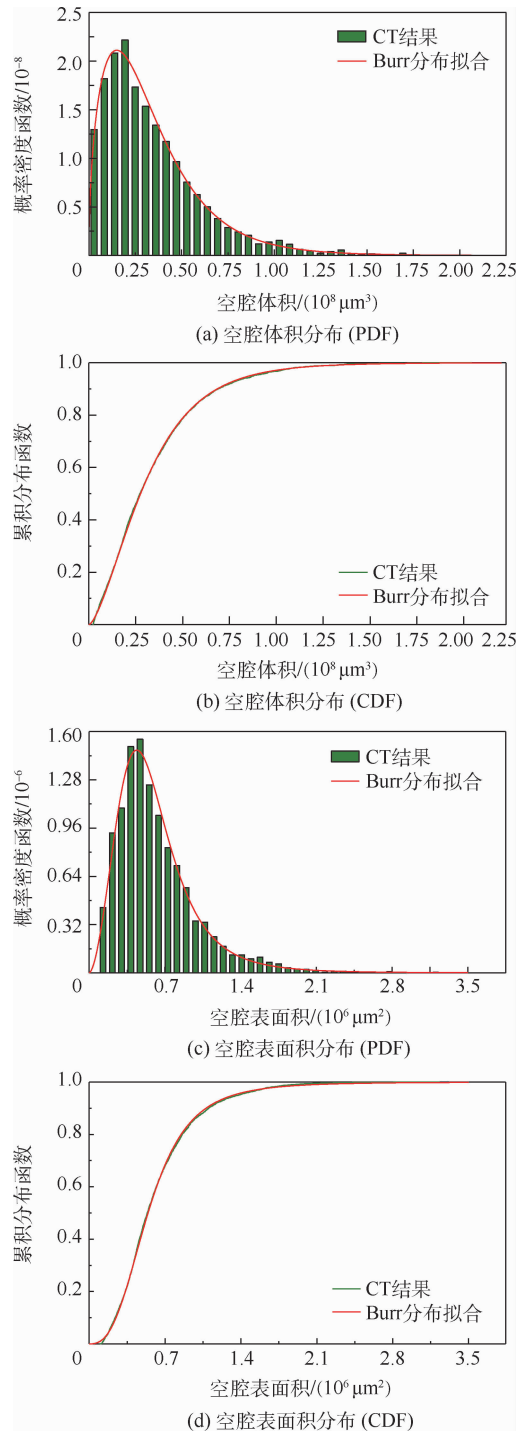


图4 空腔几何特征统计

Fig.4 Statistics of cavity geometry

划分后,所得到的面均为 0 厚度,因此为了从一开始就将厚度的影响考虑进去,将用于堆积的球体半径均增加 0.5 倍壁厚后视为填充球体半径。对填充球体半径进行统计发现,填充球体半径大小分布大致服从  $\mu = 188.549 5$ 、 $\sigma = 46.581 5$  的正态分布。填充球体半径最大值为  $349.50 \mu\text{m}$ ,最小值为  $91.32 \mu\text{m}$ ,平均值为  $188.55 \mu\text{m}$ 。具体的填充球体分布数据和参数拟合数据如图 6 所示。

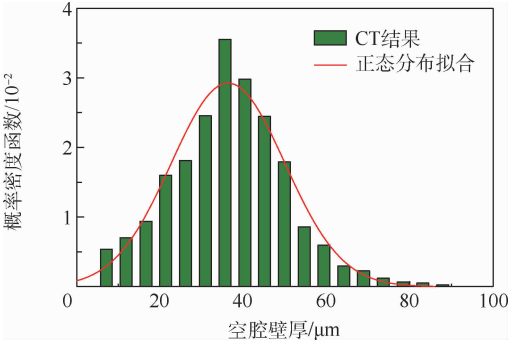
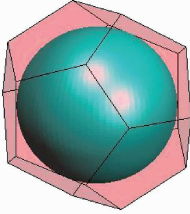
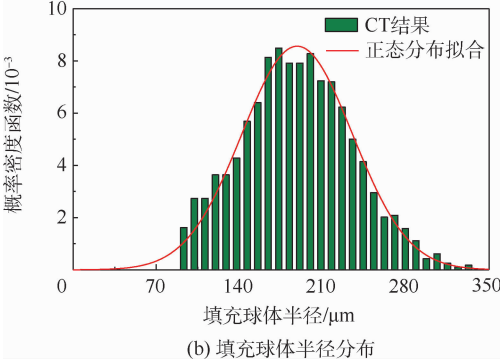


图 5 空腔壁厚统计(PDF)

Fig. 5 Cavity wall thickness statistics(PDF)



(a) 内切于正十二面体的球



(b) 填充球体半径分布

图 6 填充球体生成示意及分布

Fig. 6 Filled sphere generation and distribution

## 2 微观试样生成算法

### 2.1 前进面搜索几何构造算法

前进面搜索几何构造算法能够实现球体的紧密堆积,并能够允许球体符合一定的粒径分布规律<sup>[24-25]</sup>。前进面搜索几何构造算法通过 3 个两两相切的球体构成前进面,在前进面上放置球体,为此需要在空间内预先生成 3 个两两相切的球体。根据生成的填充半径,建立半径集合  $R$ ,所有用于填充的球体半径均从此集合中选取。通过前进面搜索几何构造算法,随机选择球体放置于前进面上,使得新球体不与周围球体发生重叠,每放置成功一个半径为  $r$  的球体即从集合  $R$  中删除此半径。在实际情况下,会出现球体与周围球体重叠的情况,此时说明新生成球体的半径太大,出于填充率的考虑,需要对半径进行重新选取。李勇俊和季顺迎<sup>[24]</sup>提出了球体重叠时的放置方法,本文对其进行延伸。例如,构成前进面三球为  $S_1$ 、

$S_2$ 、 $S_3$ ,先选取的需要放置的球体为  $S_{old}$ ,结果与周围球体发生了重叠(见图 7(a)),并且球体  $S_i$  是与  $S_{old}$  球心距离最近的一个球,则球  $S_{old}$  不符合排列要求。将  $S_i$  加入到  $S_1$ 、 $S_2$ 、 $S_3$  组中,可以得到方程组为

$$\begin{cases} (x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2 = (r + r_1)^2 \\ (x - x_2)^2 + (y - y_2)^2 + (z - z_2)^2 = (r + r_2)^2 \\ (x - x_3)^2 + (y - y_3)^2 + (z - z_3)^2 = (r + r_3)^2 \\ (x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2 = (r + r_i)^2 \end{cases} \quad (1)$$

通过求解方程组可以得到半径  $r$  及相应的球体  $S$ ,此时的  $S$  与  $S_1$ 、 $S_2$ 、 $S_3$ 、 $S_i$  均相切(见图 7(b)),但是获得的  $r$  不一定在集合  $R$  中,为此需要对求解到的半径进行判断。若  $r$  位于集合  $R$  中,则此半径有效,从  $R$  中删除使其不参与后面的半径选取,若此半径不在集合  $R$  中,则从集合  $R$  中寻找比  $r$  小且最接近  $r$  的半径  $r_{new}$  构成新球  $S_{new}$ 。若不改变  $S$  球心直接用  $S_{new}$  替换  $S$ (见图 7(c)),此时的球体  $S_{new}$  必不与周围球体相交,但是也不相切,不相切会造成前进面的缺失,为此需要将  $S_{new}$  放置到原来的前进面  $S_1S_2S_3$  上(见图 7(d))。若集合  $R$  中不存在比  $r_{new}$  更小的半径,则说明此前前进面不能够放置符合要求的球体,转而寻找下一个前进面。图 7 说明了发生重叠时球体的放置方法。

此外,可以通过划分空间网格并给球体赋予网格属性的方式来提高前进面的搜索效率<sup>[24]</sup>,图 8 以流程图的方式作进一步说明。

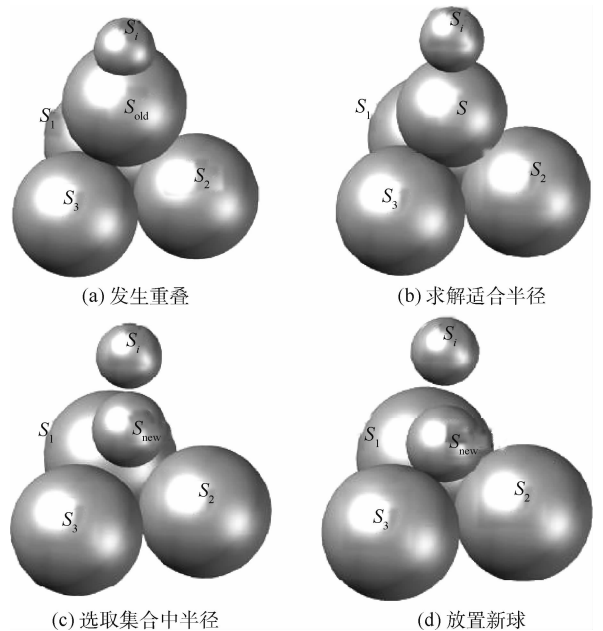


图 7 重叠时球体的放置演化方法

Fig. 7 Placement evolution method of spheres when overlapping

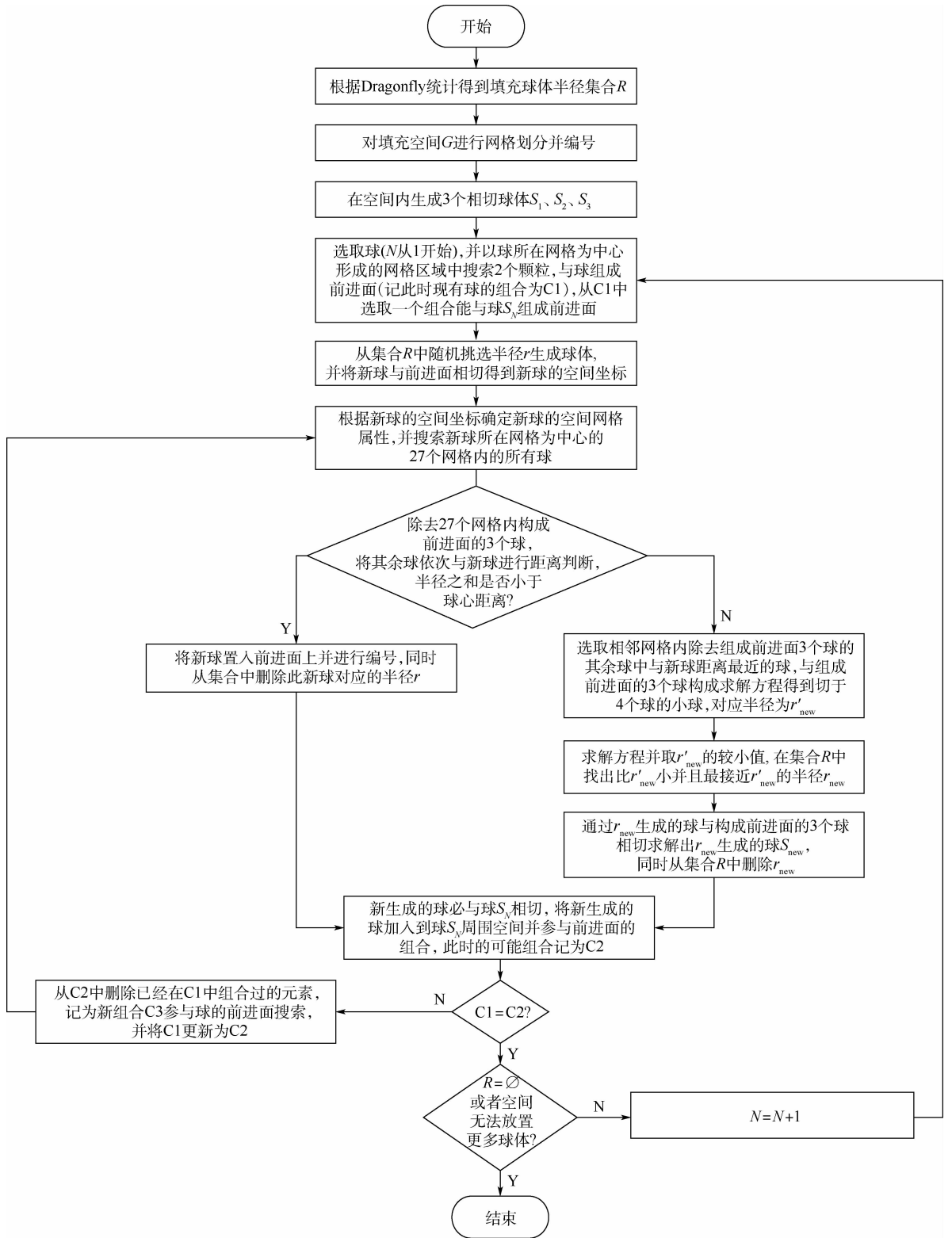


图8 前进面搜索几何构造算法流程

Fig. 8 Flow chart of advancing surface search geometric construction algorithm

2.2 Laguerre 划分算法

假设  $D = \{P_1, P_2, \dots, P_n\}$  为空间  $G$  中的一组种子点, 对于  $G$  中的任意点  $Q$ , 用  $d_v(Q, P_i)$  表示  $Q$  和  $P_i$  之间的欧几里得距离。用  $V(P_i)$  表示  $P_i$  所包含的 Voronoi 单元, 则  $V(P_i)$  可以表示为空间中满足条件的任意点  $Q$  的集合:

$$V(P_i) = \{Q \mid Q \in R, d_v(Q, P_i) < d_v(Q, P_j), j \neq i\} \quad (2)$$

从几何特征上看,  $V(P_i)$  由空间  $G$  中相较于其他种子点更为接近  $P_i$  的点构成,  $G$  中的每个种子点  $P_i$  都控制着一个 Voronoi 单元, 这些单元将空间  $G$  划分为各种凸出的多面体, 进而形成 Voronoi 划分。

Laguerre 划分是加权 Voronoi 划分的一种,若给每个种子点  $P_i$  赋予一个权重  $r_i$ ,如此就有一个与  $D$  同维度的权重集合  $U = \{r_1, r_2, \dots, r_n\}$ ,根据此集合重新定义空间  $G$  中任意点  $Q$  与种子点  $P_i$  之间的欧几里得距离:

$$d_L(Q, P_i) = (d_v(Q, P_i))^2 - r_i^2 \quad (3)$$

此距离下  $P_i$  所包含的单元  $V(P_i)$  可以表示为  $V(P_i) = \{Q | Q \in R, d_L(Q, P_i) < d_L(Q, P_j), j \neq i\}$  (4)

此时  $V(P_i)$  称为  $P_i$  的 Laguerre 划分单元。 $G$  中的每个种子点  $P_i$  都以权重  $r_i$  控制着空间大小,这些单元将空间  $G$  划分为各种凸出的多面体,进而形成 Laguerre 划分。

### 2.3 生成数值试样

在经由前进面搜索几何构造算法生成大量球体后,通过切割出中间部分来消除边界处可能出现的排列不充分的情况,切割时会出现完整球体被切割,恰巧对应了试样边缘空腔不完整的情况。对切割生成的试样进行 Laguerre 划分,由于球体与其相邻的球体大都处于相切的状态,以各自半径为权重的 Laguerre 划分,划分所得的面为与相邻两球同时相切的切平面,一个球体周围所划分的空间近似为其外接多面体,由于每个 Laguerre 单元都紧密地围绕着一个球体, Laguerre 划分的单元大小分布接近于密排球体的体积分布。图 9

显示了切割出的球体堆积试样及 Laguerre 划分后的空间试样。

## 3 结果与分析

采用 Voro++ 来实行 Laguerre 划分并统计划分后的诸多几何特征参数<sup>[26]</sup>。

### 3.1 生成试样几何特征

对于生成后的试样进行空腔体积、空腔表面积、单元平面数及泡沫空腔率统计。由于此时所生成的试样由 0 厚度的面所构成,在进行体积统计时,需用生成试样的空腔体积减去空腔周围面总面积乘以厚度的 1/2,在进行表面积统计时,忽略给 0 厚度面赋予厚度时所造成的边棱处重合带来的影响。对于泡沫空腔率,通过赋予面平均厚度后进行统计。选取一次典型的模拟结果进行几何特征统计分析,统计结果如图 10 所示。

列举 5 次模拟所得的结果与  $\mu$ CT 扫描结果对比,模拟数据与  $\mu$ CT 扫描结果的对比如图 11 所示。

本节采用的空腔率为经过高斯滤波后的空腔率,高斯滤波对整个图像进行加权平均,使得每一个点位的灰度值都由其自身及周围空间的点位经过加权平均得到,这种操作很好地消除了图像中的噪点,但是也导致实测壁厚比实际壁厚增大,从而空腔率低于实际的空腔率。高斯滤波后的平均壁厚为 36.3  $\mu\text{m}$ ,但实际壁厚应该比此值小, Liu 等<sup>[27]</sup>通过 SEM 扫描测得了 H100 实际壁厚平均值为 11.9  $\mu\text{m}$ 。Gibson 和 Ashby<sup>[2]</sup>通过测量 PVC 泡沫的原始参数,得到了 PVC 原材料的密度为  $\rho_s = 1400 \text{ kg/m}^3$ 。样品 H100 的密度  $\rho_{s1} = 100 \text{ kg/m}^3$ ,通过等质量计算得到泡沫的实际空腔率为  $1 - \rho_{s1}/\rho_s = 92.86\%$ 。

对模拟模型分别赋予实验测得的壁厚及实际壁厚,分别与实验测得及实际空腔率进行对比,如图 12 所示。

赋予实验壁厚的模拟结果空腔率平均值为 74.712%,与实验结果相差 3.72%。赋予实际壁厚的模拟结果空腔率为 91.572%,与实际结果相差 1.39%。

为了验证本文算法在空腔率方面能够适合不同密度的泡沫,选取“Divinycell H”系列的 H45、H80、H100、H130、H200 泡沫,这 5 组泡沫的名义密度分别为 45、80、100、130、200  $\text{kg/m}^3$ 。Liu 等<sup>[27]</sup>测量了以上不同密度的泡沫壁厚,分别将这些壁厚赋予算法所得到的数值模型,并与采用名义密度计算得到的实际空腔率对比,结果如表 1 所示。

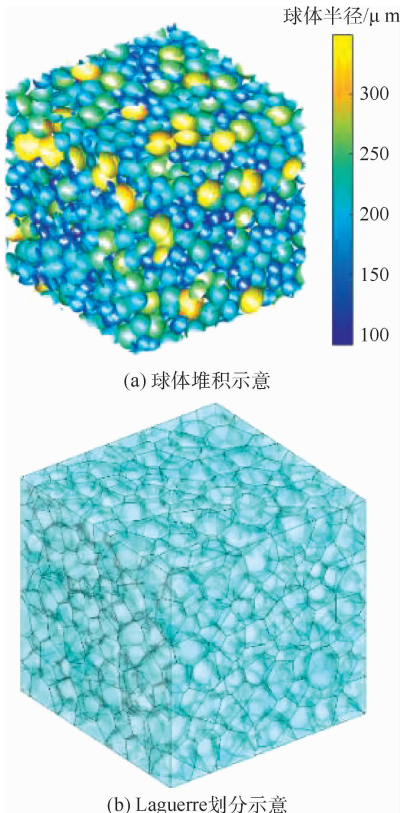


图 9 数值试样生成

Fig. 9 Numerical sample generation

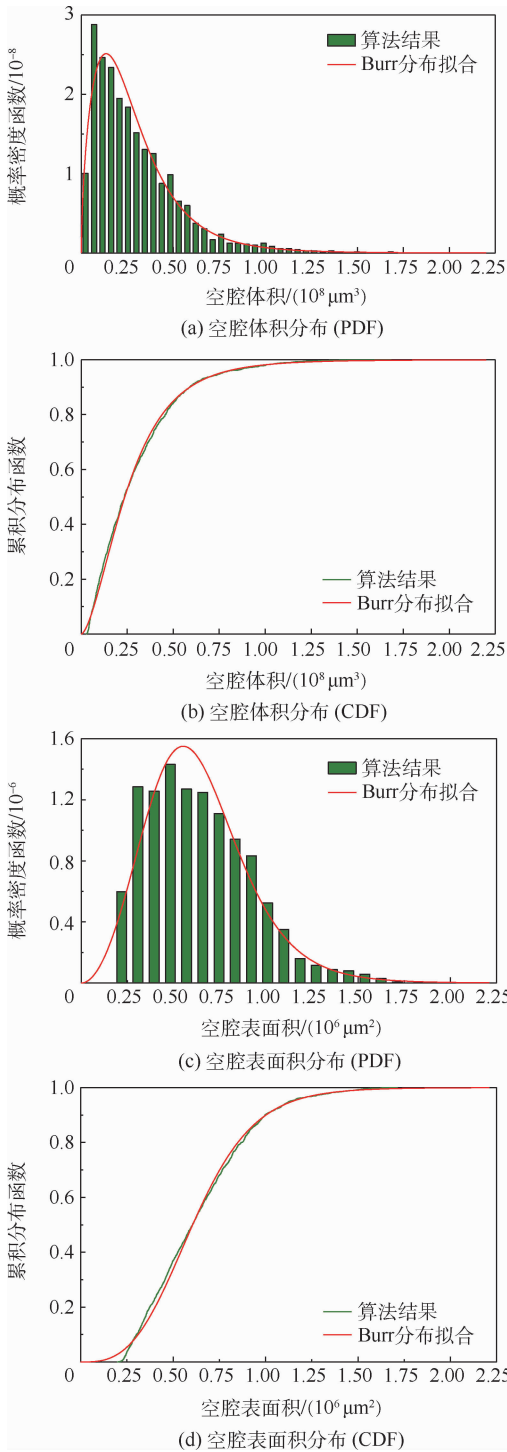
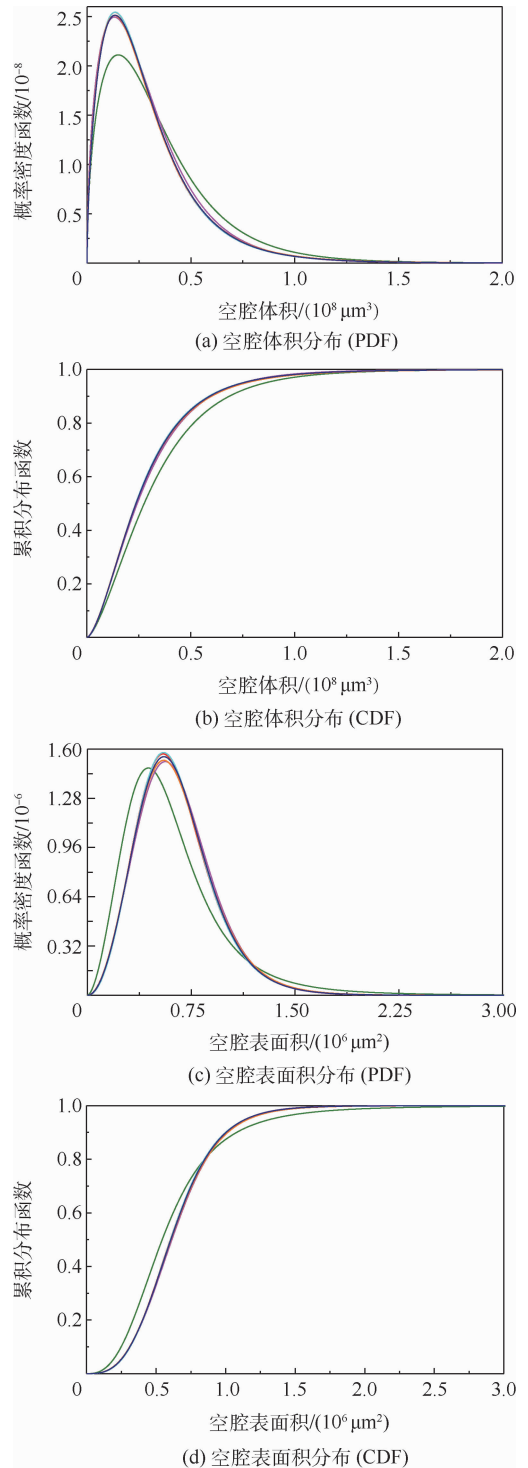


图 10 数值模型几何特征统计

Fig. 10 Numerical model geometric feature statistics

Matzke<sup>[23]</sup>观测到的空腔面数的平均值为13.7,模拟所得的平均值为12.99,相差5.2%。模型体积最大值为 $1.7428 \times 10^8 \mu\text{m}^3$ ,实验所得体积最大值为 $2.0403 \times 10^8 \mu\text{m}^3$ ,相差14.58%,模型体积最小值为 $2.7178 \times 10^6 \mu\text{m}^3$ ,实验所得体积最小值为 $2.2079 \times 10^6 \mu\text{m}^3$ ,相差23.09%,模型体积平均值为 $2.9508 \times 10^7 \mu\text{m}^3$ ,实验所得体积平均值为 $3.4101 \times 10^7 \mu\text{m}^3$ ,相差13.47%。



— CT结果 — 模拟结果1 — 模拟结果2  
— 模拟结果3 — 模拟结果4 — 模拟结果5

图 11 数值模型结果与实验结果对比  
Fig. 11 Comparison of numerical model results with experimental results

### 3.2 误差分析

根据上述几何特征的分析对比可以发现,本文算法具有很好的稳定性,每次的结果波动较小,并且与实验的几何特征在空腔面数、体积表面积空间分布等参数上符合较好,但在体积最大最小值相差较大,分析其误差产生原因:

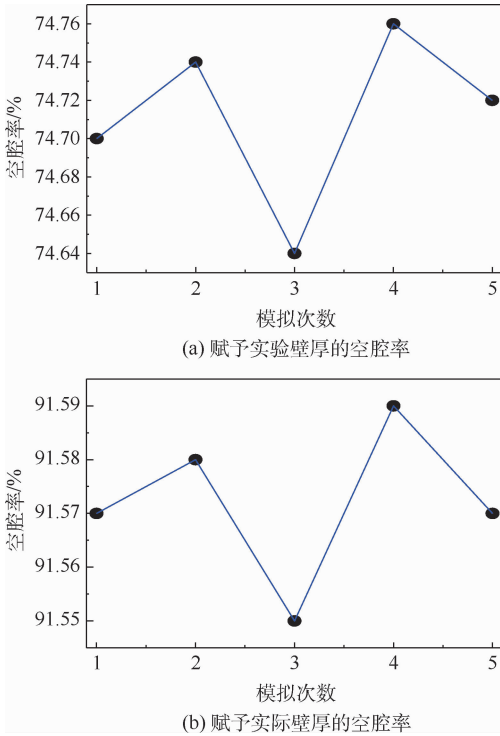


图 12 数值模型空腔率

Fig. 12 Numerical model cavity ratio

表 1 不同密度泡沫空腔率的数值模型与实验对比

Table 1 Numerical model and experimental comparison of foam cavity ratio of different densities

泡沫类型	空腔壁厚/ $\mu\text{m}$	真实空腔率/%	数值模型平均空腔率/%	误差/%
H45	7.44	96.79	94.77	2.09
H80	8.84	94.29	93.79	0.53
H100	11.9	92.86	91.57	1.39
H130	13.86	90.71	90.26	0.50
H200	20.41	85.71	85.67	0.05

始体积相差为 22.85%。若最小半径球体周围排布的都是最大半径的球体(见图 13(c)),则最多可以排布 4 个球体,中心球体区域经 Laguerre 划分后的体积为  $4.7003 \times 10^6 \mu\text{m}^3$ ,与原始体积相差为 112.89%。由此可见,采用十二面体内切球来确定填充球体大小对于小空间区域影响更大。

### 3.3 优化算法

本文算法产生的体积分布曲线(CDF)总体位于原始曲线上方,说明算法产生的小空间数量比原始小空间数量多,而大空间数量比原始大空间数量少,对此提出一种优化算法。由于小球经划分后体积受周围球体影响较大,而大球受自身大小影响较大,以空间体积最大值及空间体积平均值作为体积修正的依据。

在整个空间体积大小取三点(空间体积最大值、最小值、平均值),将空间体积分为 2 部分:第

基于正十二面体内接球假设的前提下,球体相邻的球体都是与其等大小的,而在实际情况下这种情形是不存在的,当中心球体为最大球,而周边的球体比最大球半径小,就会造成与中心球相切的球体数量增多,从而导致进行 Laguerre 划分时单元的面数增多,在同为内切球的情况下,面数多的比面数少的内切球所占体积比更大,因此导致 Laguerre 划分后的单元体积大小小于原始体积大小。同理,小球周围的球体由于比其大的原因,会导致 Laguerre 划分后的单元体积大小大于原始体积大小。

采用等半径的球体进行前进面搜索几何构造算法的排列并采用 Laguerre 划分,可以发现以此方法形成的中心球体周围恰巧是 12 个相同大小的球体(见图 13(a)),以最大半径为例,中心球体区域经 Laguerre 划分后的体积为  $2.0261 \times 10^8 \mu\text{m}^3$ ,与原始体积相差仅为 0.21%。若最大半径球体周围排布的都是最小半径的球体(见图 13(b)),则可以排布最多 74 个小球,中心球体区域经 Laguerre 划分后的体积为  $1.5741 \times 10^8 \mu\text{m}^3$ ,与原

1 部分为最小值至平均值所包含的空间体积大小,第 2 部分为平均值至最大值所包含的空间体积大小,两段区域分别以边界值中较大值作为修正的标准。首先,以之前存在的数值模型作为对比数据,与实验数据进行对比,得到体积最大值的比值及体积平均值的比值。然后,以平均值的比值作为第 1 部分的修正标准,以最大值的比值作为第 2 部分的修正标准,将体积比折算为填充球体半径比,进行填充球体半径的缩放调整,以调整后的填充球体半径集合进行前进面搜索几何构造算法的填充,将生成的新微观试样重复以上操作,循环直至最大值及平均值的误差均在预设范围之内,得到最终的微观试样。图 14 以流程图形式阐述了这一操作。其中采用  $e$  作为误差限值,取 2%,即数值模型体积最大值与实验体积最大值、数值模型体积平均值与实验体积平均值误差均在 2% 以内时,停止迭代。

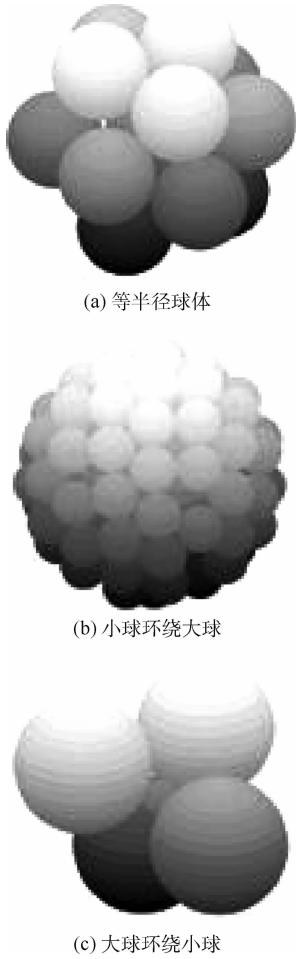


图 13 不同种类球体堆积

Fig. 13 Different kinds of spheres stacked

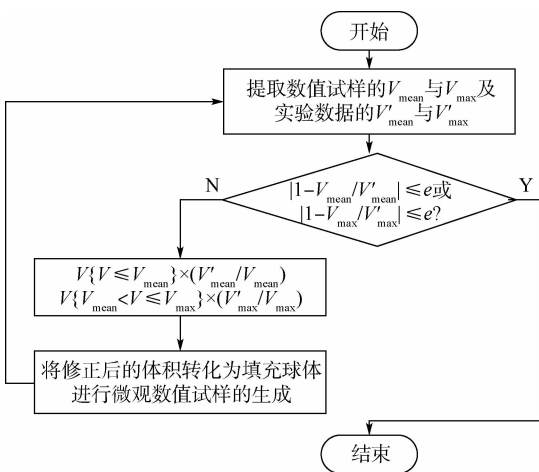


图 14 优化算法流程

Fig. 14 Flow chart of optimization algorithm

### 3.4 优化后试样几何特征

应用优化算法后,3次迭代后流程结束。将3次迭代后模型、初始模型与实验几何特征进行对比,如图15所示。相较于初始模型,经过3次迭代后的模型在体积分布上与实验数据具有更好的贴合度。

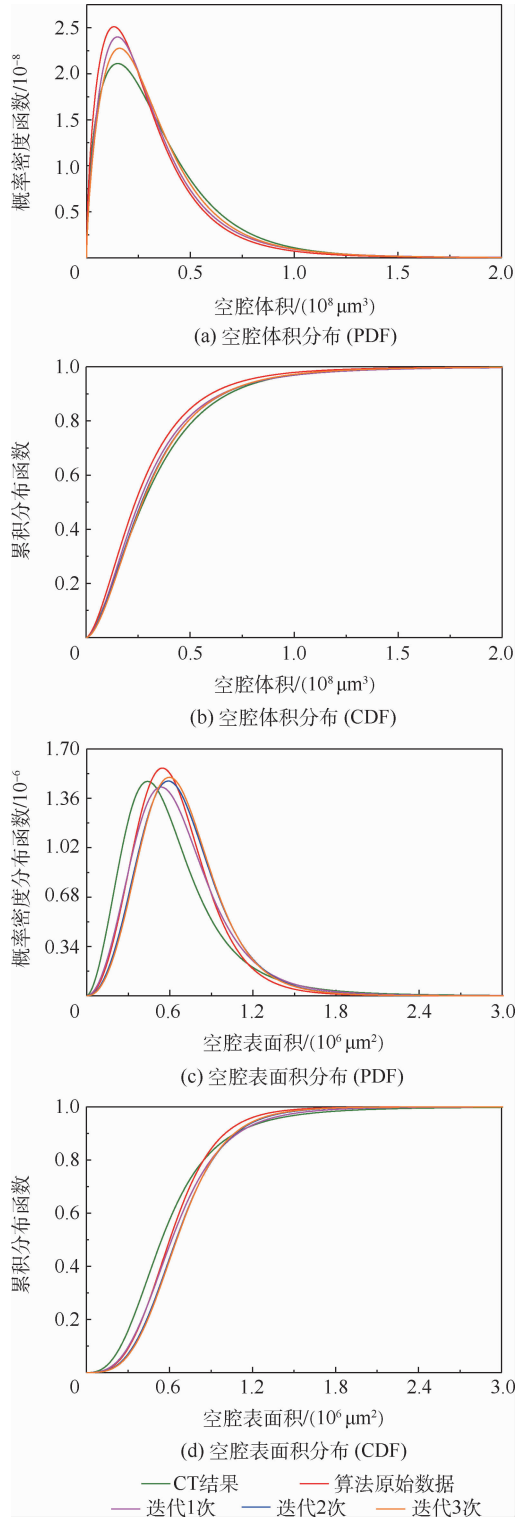


图 15 优化试样几何特征对比

Fig. 15 Comparison geometric characteristics of optimized sample

## 4 结论

1) 基于  $\mu$ CT 扫描数据,得到了 PVC 闭孔泡沫几何特征,发现空腔体积服从 Burr 分布,空腔表面积服从 Burr 分布,空腔体积/表面积服从正态分布,空腔壁厚服从正态分布。

2) 基于前进面搜索几何构造算法和 Laguerre 划分算法编写了考虑壁厚大小的 PVC 闭孔泡沫微观试样的几何模型生成算法。

3) 将生成的 PVC 闭孔泡沫几何特征与  $\mu$ CT 扫描数据对比,分析了误差产生的原因,提出了改进算法,并将改进算法生成的模型与  $\mu$ CT 扫描数据对比,验证了本文算法生成的几何模型的可用性。

### 参考文献 (References)

- [ 1 ] CHEN Y M, DAS R, BATTLE M. Effects of cell size and cell wall thickness variations on the stiffness of closed-cell foams [ J ]. *International Journal of Solids and Structures*, 2015, 52: 150-164.
- [ 2 ] GIBSON L J, ASHBY M F. *Cellular solids* [ M ]. Cambridge: Cambridge University Press, 1997.
- [ 3 ] MILLS N J, ZHU H X. The high strain compression of closed-cell polymer foams [ J ]. *Journal of the Mechanics and Physics of Solids*, 1999, 47(3): 669-695.
- [ 4 ] WEAIRE D, PHELAN R. A counter-example to Kelvin's conjecture on minimal surfaces [ J ]. *Philosophical Magazine Letters*, 1994, 69(2): 107-110.
- [ 5 ] WISMANS J G F, GOVAERT L E, VAN DOMMELEN J A W. X-ray computed tomography-based modeling of polymeric foams: The effect of finite element model size on the large strain response [ J ]. *Journal of Polymer Science Part B: Polymer Physics*, 2010, 48(13): 1526-1534.
- [ 6 ] VESENJAK M, VEYHL C, FIEDLER T. Analysis of anisotropy and strain rate sensitivity of open-cell metal foam [ J ]. *Materials Science and Engineering: A*, 2012, 541: 105-109.
- [ 7 ] SUN Y L, LOWE T, MCDONALD S A, et al. In situ investigation and image-based modelling of aluminium foam compression using micro X-ray computed tomography [ M ] // LETA F R. *Visual computing*. Berlin: Springer, 2014: 189-197.
- [ 8 ] SUN Y L, LI Q M, LOWE T, et al. Investigation of strain-rate effect on the compressive behaviour of closed-cell aluminium foam by 3D image-based modelling [ J ]. *Materials & Design*, 2016, 89: 215-224.
- [ 9 ] KIM S H, CHUNG H J, RHEE K Y. Numerical analysis on the compressive behaviors of aluminum foam material using computed tomography imaging [ J ]. *Advanced Materials Research*, 2010, 123-125: 567-570.
- [ 10 ] JEON I, ASAHINA T, KANG K J, et al. Finite element simulation of the plastic collapse of closed-cell aluminum foams with X-ray computed tomography [ J ]. *Mechanics of Materials*, 2010, 42(3): 227-236.
- [ 11 ] HUANG R X, LI P F, LIU T. X-ray microtomography and finite element modelling of compressive failure mechanism in cenosphere epoxy syntactic foams [ J ]. *Composite Structures*, 2016, 140: 157-165.
- [ 12 ] DAPHALAPURKAR N P, HANAN J C, PHELPS N B, et al. Tomography and simulation of microstructure evolution of a closed-cell polymer foam in compression [ J ]. *Mechanics of Advanced Materials and Structures*, 2008, 15(8): 594-611.
- [ 13 ] CATY O, MAIRE E, YOUSSEF S, et al. Modeling the properties of closed-cell cellular materials from tomography images using finite shell elements [ J ]. *Acta Materialia*, 2008, 56(19): 5524-5534.
- [ 14 ] 李侯贞强, 张亚栋, 张锦华, 等. 基于 CT 的泡沫铝三维细观模型重建及应用 [ J ]. *北京航空航天大学学报*, 2018, 44(1): 160-168.
- LI H Z Q, ZHANG Y D, ZHANG J H, et al. Reconstruction and application of three-dimensional mesoscopic model of aluminum foam based on CT [ J ]. *Journal of Beijing University of Aeronautics and Astronautics*, 2018, 44(1): 160-168 (in Chinese).
- [ 15 ] LI K, GAO X L, SUBHASH G. Effects of cell shape and strut cross-sectional area variations on the elastic properties of three-dimensional open-cell foams [ J ]. *Journal of the Mechanics and Physics of Solids*, 2006, 54(4): 783-806.
- [ 16 ] SONG Y Z, WANG Z H, ZHAO L M, et al. Dynamic crushing behavior of 3D closed-cell foams based on Voronoi random model [ J ]. *Materials & Design*, 2010, 31(9): 4281-4289.
- [ 17 ] ZHU H X, HOBDELL J R, WINDLE A H. Effects of cell irregularity on the elastic properties of open-cell foams [ J ]. *Acta Materialia*, 2000, 48(20): 4893-4900.
- [ 18 ] ZHU H X, WINDLE A H. Effects of cell irregularity on the high strain compression of open-cell foams [ J ]. *Acta Materialia*, 2002, 50(5): 1041-1052.
- [ 19 ] ZHU W Q, BLAL N, CUNSOLO S, et al. Effective elastic behavior of irregular closed-cell foams [ J ]. *Materials*, 2018, 11(11): 2100.
- [ 20 ] RIBEIRO-AYEH S. *Finite element modelling of the mechanics of solid foam materials* [ D ]. Karlstad: Karlstad University, 2005.
- [ 21 ] REDENBACH C, SHKLYAR I, ANDRÄ H. Laguerre tessellations for elastic stiffness simulations of closed foams with strongly varying cell sizes [ J ]. *International Journal of Engineering Science*, 2012, 50(1): 70-78.
- [ 22 ] GHAZI A, BERKE P, KAMEL K E M, et al. Multiscale computational modelling of closed cell metallic foams with detailed microstructural morphological control [ J ]. *International Journal of Engineering Science*, 2019, 143: 92-114.
- [ 23 ] MATZKE E B. The three-dimensional shape of bubbles in foam—An analysis of the rôle of surface forces in three-dimensional cell shape determination [ J ]. *American Journal of Botany*, 1946, 33(1): 58-80.
- [ 24 ] 李勇俊, 季顺迎. 基于球形颗粒几何排列的离散元试样高效生成方法 [ J ]. *应用力学学报*, 2020, 37(2): 469-476.
- LI Y J, JI S Y. Construction approach of DEM samples with high efficiency based on geometrical packing of spherical particles [ J ]. *Chinese Journal of Applied Mechanics*, 2020, 37(2): 469-476 (in Chinese).
- [ 25 ] VALERA R R, MORALES I P, VANMAERCKE S, et al. Modified algorithm for generating high volume fraction sphere packings [ J ]. *Computational Particle Mechanics*, 2015, 2(2): 161-172.
- [ 26 ] RYCROFT C H. *Voro++: A three-dimensional voronoi cell library in C++* [ J ]. *Chaos*, 2009, 19(4): 041111.
- [ 27 ] LIU Y, RAHIMIDEHGOLAN F, ALTENHOF W. Anisotropic compressive behavior of rigid PVC foam at strain rates up to  $200 \text{ s}^{-1}$  [ J ]. *Polymer Testing*, 2020, 91: 106836.

# An algorithm for generating geometric models of microscopic specimens of PVC foam based on $\mu$ CT images

ZHOU Yong, XUE Bin, GUO Yunxin, WANG Renpeng\*

(College of Civil Engineering, Tongji University, Shanghai 200092, China)

**Abstract:** In numerical simulation of foam microstructure, the geometric characteristics and arrangement of foam cavities have an important influence on calculation efficiency and calculation results. We propose a new algorithm, based on the advancing surface search geometric construction algorithm and the Laguerre partition algorithm, to generate the geometric model of the PVC foam microscopic specimen. First, reconstruct the authentic geometric model of the foam from  $\mu$ CT scan image, and measure the geometric characteristics of the foam cavity and the volume distribution pattern. Then, convert the measured foam cavity volume into a sphere, and put it into the space through the advancing surface search geometric construction algorithm. Finally, divide the space sphere into regions by means of Laguerre division, and assign wall thickness parameters to form a geometric model of the microstructure. The established model is in good agreement with the actual material in terms of micro-geometric characteristics.

**Keywords:** CT scan; advancing surface method; Laguerre division; closed cell foam; microstructural model

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2020.0690

# 金属蜗杆与塑料斜齿轮不等齿距啮合方法

任继华<sup>1</sup>, 石照耀<sup>1,\*</sup>, 王得峰<sup>2</sup>, 余石鹏<sup>2</sup>

(1. 北京工业大学 机械工程与应用电子技术学院, 北京 100124; 2. 东莞市星火齿轮有限公司, 东莞 523723)

**摘 要:** 为了进一步提高金属蜗杆与塑料斜齿轮传动中塑料齿轮的承载能力, 研究了传统等齿距蜗杆与斜齿轮啮合传动时受力的特点, 提出了蜗杆与斜齿轮不等齿距啮合方法。基于梁弯曲理论和轮齿变形理论, 得到了齿面载荷、变形及接触刚度的关系, 并以轮齿齿根弯曲变形率相等为前提, 推导了不等齿距啮合的设计方法, 得到了不等齿距啮合时蜗杆的齿距调整量, 通过静态强度实验进行了验证。实验结果表明: 不等齿距设计可以使塑料斜齿轮的承载能力提高 13.69%。

**关键词:** 不等齿距; 金属蜗杆; 塑料斜齿轮; 啮合理论; 齿根弯曲应力

**中图分类号:** TH132.4; TH124; TH112.3

**文献标志码:** A

**文章编号:** 1001-5965(2022)06-0979-07

蜗轮蜗杆机构是一种常见的空间交错轴传动机构, 空间中两轴一般为  $90^\circ$ , 具有传动比大、传动平稳及可自锁等特点, 广泛应用于各种机械产品中。但传统的金属蜗轮设计复杂, 需要特定的蜗轮滚刀, 且加工精度不高, 对安装精度要求较高, 成本往往较高, 因此, 用斜齿轮代替传统蜗轮的蜗杆斜齿轮传动的方式便应运而生。

基于等齿距蜗杆斜齿轮齿面载荷及变形规律, 目前国内学者对此进行了深入研究。云永琥等<sup>[1]</sup>分析了金属蜗杆与塑料斜齿轮的啮合规律, 并提出了一种计算齿轮齿面载荷和接触点处变形的的方法。石照耀和欧阳志喜<sup>[2]</sup>通过采用刘易斯方程式和赫兹应力系数面压分别对塑料直齿轮和斜齿轮进行强度设计计算, 在齿轮的齿顶处受力, 计算齿根处的弯曲应力小于许用应力, 可以很好地保证塑料齿轮工作时的稳定性。文献[3]提出了一种高效计算塑料圆柱直齿轮齿面载荷分配规律的方法, 并分析了齿面载荷分配规律对齿轮传

动误差及啮合刚度的影响。文献[4]针对双圆弧齿轮提出了一种快速计算齿轮齿面接触力的方法, 根据材料力学和弹性力学计算了齿轮的弯曲刚度, 从而求得接触力。文献[5]利用 MATLAB-APDL 对直齿圆柱齿轮进行了精确建模, 并利用有限元软件分别从静态/准静态载荷和动态载荷 2 个方面分析了齿轮的强度和载荷分布规律, 验证了齿端刚度效应、应力集中效应和弹性体力场的连续性。

为了进一步提高塑料斜齿轮的承载能力, 本文根据梁弯曲理论定义了齿根弯曲率, 并分析了等齿距啮合时各啮合齿的齿根弯曲率, 在使各啮合齿的齿根弯曲变形率相等的前提下, 基于传统等齿距蜗杆斜齿轮啮合理论, 得到了此时塑料齿轮齿面的变形情况, 推导了两者不等齿距啮合的方法, 并通过齿轮静态强度实验对 2 种模型的承载能力进行比较验证。不等齿距蜗杆斜齿轮啮合方法为进一步提高塑料齿轮的强度提供了理论指导。

收稿日期: 2020-12-10; 录用日期: 2020-12-31; 网络出版时间: 2021-03-02 12:28

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20210302.0946.002.html](http://kns.cnki.net/kcms/detail/11.2625.V.20210302.0946.002.html)

基金项目: 国家自然科学基金重点项目(51635001)

\*通信作者. E-mail: shizhaoyao@bjut.edu.cn

**引用格式:** 任继华, 石照耀, 王得峰, 等. 金属蜗杆与塑料斜齿轮不等齿距啮合方法[J]. 北京航空航天大学学报, 2022, 48(6): 979-985. REN J H, SHI Z Y, WANG D F, et al. Method of unequal pitch meshing of metal worm and plastic helical gear[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(6): 979-985 (in Chinese).

## 1 等齿距金属蜗杆与塑料斜齿轮变形分析

在钢制蜗杆与塑料斜齿轮传动中,由于金属蜗杆的强度远远大于塑料斜齿轮,蜗杆斜齿轮传动的破坏往往发生在斜齿轮上。塑料斜齿轮齿根弯曲断裂是一种常见的失效形式,究其原因主要有2种:①由于斜齿轮承受的载荷超过其许用载荷<sup>[6]</sup>引起的过载折断;②由于蜗杆斜齿轮长时间的运行,导致在塑料斜齿轮齿面产生点蚀、疲劳裂纹等,并进一步发展为较大裂纹,最终引起斜齿轮齿根弯曲疲劳折断。因此,在设计蜗杆斜齿轮时,有必要对齿根处的强度进行校核计算。根据文献[7-8]中的梁弯曲理论知识,进行如下定义:在蜗杆斜齿轮啮合的任意位置,斜齿轮齿根处的弯曲变形率 $\varepsilon_i$ 等于接触点处的综合接触变形 $\delta_i$ 与该接触到斜齿轮齿根圆的距离 $h_i$ 的比值,即

$$\varepsilon_i = \frac{\delta_i}{h_i} \quad (1)$$

等齿距蜗杆斜齿轮在某时刻下啮合的位置如图1所示。设啮合点 $a$ 、 $b$ 、 $c$ 到斜齿轮齿根圆的距离分别为 $h_1$ 、 $h_2$ 、 $h_3$ ,则由图1可知:

$$h_1 > h_2 > h_3 \quad (2)$$

假设不考虑变位系数、齿廓间隙、轮齿修形等引起的基节误差,由等齿距蜗杆斜齿轮啮合的特点可知,各啮合节点处的综合接触变形相等<sup>[1]</sup>,即

$$\delta_1 = \delta_2 = \delta_3 \quad (3)$$

式中: $\delta_1$ 、 $\delta_2$ 、 $\delta_3$ 分别为啮合点 $a$ 、 $b$ 、 $c$ 处的综合接触变形。

由式(1)可得,该啮合时刻下啮合点 $a$ 、 $b$ 、 $c$ 在斜齿轮齿根处的弯曲变形率为 $\varepsilon_1$ 、 $\varepsilon_2$ 、 $\varepsilon_3$ ,显然:

$$\varepsilon_1 < \varepsilon_2 < \varepsilon_3 \quad (4)$$

与之对应,该时刻下啮合点 $a$ 所承受的载荷最小,啮合点 $c$ 所承受的载荷最大,斜齿轮更容易在3齿齿根处发生断裂,因此,该齿轮的承载能力主要由 $c$ 点承受的载荷决定,这种承载方式对斜齿轮是非常不利的。

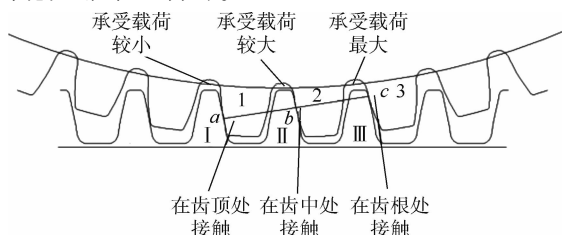


图1 等齿距啮合

## 2 金属蜗杆与塑料斜齿轮不等齿距理论的提出与设计

由分析可知,斜齿轮的承载能力主要受限于啮合点 $c$ ,本文的设计思想是:通过调整蜗杆与斜齿轮的齿距,使蜗杆与斜齿轮在受载时各啮合点处的变形量发生不同的改变,如图2所示,则各啮合点的啮合齿面综合接触变形均会发生改变,通过调整两者齿距,使斜齿轮单齿承受最大载荷时,各啮合齿根处的弯曲变形率近似相等,理论上可以大大提高齿轮的承载能力。

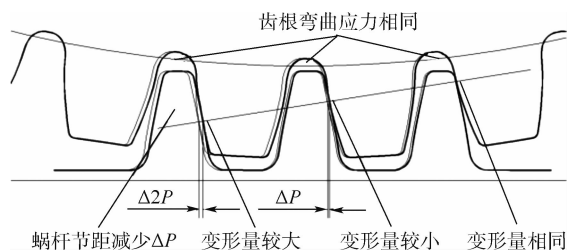


图2 不等齿距啮合

Fig. 2 Unequal pitch meshing

### 2.1 金属蜗杆与塑料斜齿轮齿面接触刚度分析

蜗杆斜齿轮齿面接触刚度 $k$ 是指两者在啮合接触时抵抗变形的能力,其与接触点处的几何形状及位置有关,定义如下:

$$k_i = \frac{1}{\delta_i} \quad (5)$$

式中: $\delta_i$ 为各啮合点处的总变形; $k_i$ 为各啮合点为单位作用力下的接触刚度; $i$ 为参与啮合的齿对数。

蜗杆斜齿轮在啮合过程中,齿面接触情况非常复杂,在任一啮合位置下,啮合接触点处的总变形 $\delta_i$ 包含4个方面,分别为轮齿接触点弹性弯曲变形 $\delta_{bi}$ 、轮齿剪切变形 $\delta_{si}$ 、轮齿基体变形 $\delta_{gi}$ 和赫兹接触压力产生的变形 $\delta_{pi}$ <sup>[9-12]</sup>。齿轮参数如图3所示,即

$$\delta_i = \delta_{bi} + \delta_{si} + \delta_{gi} + \delta_{pi} \quad (6)$$

式中:

$$\delta_{bi} = \frac{12F_i z \cos^2 \alpha}{EBS_F^3} \left( S_m^2 + \frac{z^2}{3} - S_m z \right) + \frac{6F_i \cos^2 \alpha (w - z)^3}{EBS_F^3} \quad (7)$$

$$\delta_{si} = \frac{2(1 + \nu)F_i \cos^2 \alpha}{EBS_F} \left[ z + (w - z) \cdot \ln \left( \frac{w - z}{w - S_m} \right) \right] \quad (8)$$

$$\text{其中: } w = \frac{nS_F - zS_k}{S_F - S_k} \circ$$

$$\delta_{gi} = \frac{24F_i S_m^2 \cos^2 \alpha}{\pi EBS_F^2} \quad (9)$$

$$\delta_{pi} = \frac{4F_i(1 - \nu^2)}{\pi EB} \cdot \frac{\rho_2}{\rho_1 + \rho_2} \quad (10)$$

式(7)~式(10)中符号及意义如表1所示。

由式(5)~式(10)可分别解得蜗杆与斜齿轮的接触刚度,设参与啮合齿中蜗杆的接触刚度为  $k_{wi}$ ,斜齿轮的接触刚度为  $k_{gi}$ ,则蜗杆斜齿轮在啮合时啮合齿对的组合刚度  $k_{eqi}$  为

$$k_{eqi} = \frac{k_{wi}k_{gi}}{k_{wi} + k_{gi}} \quad (11)$$

式中: $i$ 为总的啮合齿对数。

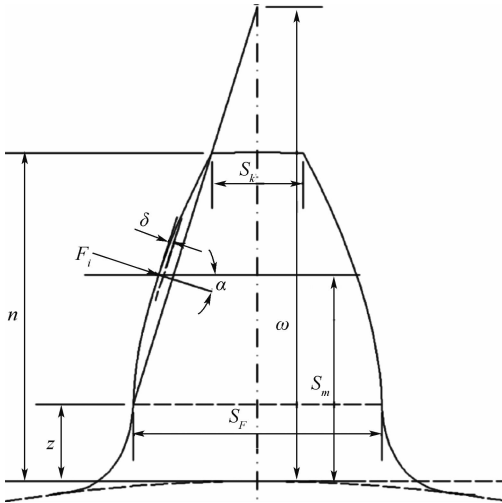


图3 齿轮轮齿参数

Fig. 3 Gear tooth parameters

表1 式(7)~式(10)中符号及意义

Table 1 Symbols and meanings in formula (7) - (10)

参数	含义
$B/\text{mm}$	齿宽
$E/\text{Pa}$	弹性模量
$F_i/\text{N}$	每齿法向载荷
$S_k/\text{mm}$	齿顶厚
$w/\text{mm}$	悬臂梁长度
$S_m/\text{mm}$	载荷作用线与齿根圆之间的距离
$S_F/\text{mm}$	危险截面宽度
$z/\text{mm}$	危险截面与齿根圆之间的距离
$n/\text{mm}$	全齿高
$\alpha/(\text{°})$	加载角
$\delta_B/\text{mm}$	轮齿接触点弹性弯曲挠度
$\delta_S/\text{mm}$	轮齿剪切变形
$\delta_g/\text{mm}$	轮齿基体偏移量
$\delta_p/\text{mm}$	赫兹接触压力产生的变形
$\delta/\text{mm}$	轮齿总挠度
$\nu$	泊松比
$\rho_1, \rho_2/\text{mm}$	蜗杆斜齿轮的曲率半径

## 2.2 不等齿距金属蜗杆与斜齿轮的节距计算

蜗杆与斜齿轮传动中,蜗杆相当于一个大螺旋角的斜齿轮<sup>[13]</sup>,设计时蜗杆与斜齿轮均以法向参数作为标准值<sup>[14-16]</sup>,即蜗杆与斜齿轮的法向模数和法向压力角相等且蜗杆导程角等于斜齿轮的螺旋角,按照正常等齿距蜗杆斜齿轮进行设计。

得到齿轮参数后,由于所有啮合点均位于啮

合线上,基于重合度计算公式和各啮合点的位置关系(相邻两啮合点间隔一个节距),可以确定所有啮合点的位置,并定义啮合点到齿轮圆心的距离为啮合半径  $r$ 。通过计算一个齿从啮入到啮出的齿面载荷及其对应的位置,得到斜齿轮单齿承受最大载荷时的综合接触变形和对应的位置。由于齿距改变前后对啮合点位置的影响非常小,为了简化计算,假设齿距改变前后单齿承受最大载荷的位置不变。由齿轮参数计算端面重合度为  $\varepsilon_\alpha$ ,根据齿轮表面啮合力的变化规律<sup>[17-18]</sup>,斜齿轮单齿承受最大载荷位于少齿区域,此时参与啮合的总齿数  $N$  为端面重合度的整数部分,即

$$N = [\varepsilon_\alpha] \quad (12)$$

则设斜齿轮单齿承受最大载荷时各啮合点的啮合半径为  $r_1, r_2, \dots, r_N$ ,啮合点到斜齿轮齿根圆的距离  $h_i$  为

$$h_i = r_i - \frac{d_2}{2} \quad (13)$$

式中: $r_i$ 为各啮合点的啮合半径; $d_2$ 为斜齿轮分度圆; $i$ 为对应啮合齿上啮合点的序号。

由前面的分析可知,不等齿距设计的最终目的是为了在最大载荷处使斜齿轮齿根处的弯曲变形率相同,即

$$\varepsilon_1 = \varepsilon_2 = \dots = \varepsilon_N \quad (14)$$

齿根处的弯曲变形率反映了各个齿在啮合时所承受的法向载荷,则在最大载荷处,不等齿距蜗杆斜齿轮啮合时斜齿轮上啮合点处的法向载荷有如下关系:

$$F_1 = F_2 = \dots = F_N \quad (15)$$

由齿轮传动的工作原理可知,任意啮合时刻下,所有参与啮合的齿受到的法向载荷之和等于齿轮轮齿的总法向载荷  $F_n$ <sup>[19]</sup>,即

$$\sum_{i=1}^N F_i = F_n \quad (16)$$

式中:总法向载荷  $F_n$  为

$$F_n = \frac{2T}{d_2 \cos \beta \cos \alpha_n} \quad (17)$$

其中: $T$ 为扭矩; $\beta$ 为斜齿轮螺旋角; $\alpha_n$ 为法向压力角。

则在最大载荷处各啮合点处的法向载荷为

$$F_1 = F_2 = \dots = F_N = \frac{2T}{Nd_2 \cos \beta \cos \alpha_n} \quad (18)$$

假设不考虑由于齿侧间隙的调整、轮齿修形及齿轮变位系数等引起的基节误差,蜗杆斜齿轮的组合刚度为  $k_{eqi}$ ,则斜齿轮各啮合齿在接触点处的综合接触变形  $\delta'_i$  为

$$\delta'_i = \frac{F_i}{k_{eqi}} \quad (19)$$

设得到相邻两啮合点处的综合接触变形为 $\delta'_1, \delta'_2, \dots, \delta'_N$ , 则相邻两啮合点之间的齿距差为

$$\begin{cases} \Delta P_1 = |\delta'_1 - \delta'_2| \\ \Delta P_2 = |\delta'_2 - \delta'_3| \\ \vdots \\ \Delta P_{N-1} = |\delta'_{N-1} - \delta'_N| \end{cases} \quad (20)$$

则对蜗杆齿距的平均调整量  $\Delta P$  为

$$\Delta P = \frac{\Delta P_1 + \Delta P_2 + \dots + \Delta P_{N-1}}{N - 1} \quad (21)$$

正常情况下, 一对啮合的齿轮之间应该留有一定的侧隙, 这是为了防止齿轮在啮合时轮齿因摩擦发热膨胀卡死, 另外, 便于在两啮合齿廓之间形成油膜, 进一步改善啮合的流畅性<sup>[20-22]</sup>。为了避免蜗杆齿距调整后, 在啮合的过程中蜗杆与斜齿轮产生干涉, 需要对齿距的调整量进行校核, 即校核齿距的调整量和端面重合度与斜齿轮的齿侧间隙之间满足一定关系。设斜齿轮的齿侧间隙为 $j_n$ , 应满足以下关系:

$$\Delta P \cdot \varepsilon \alpha \leq j_n \quad (22)$$

根据齿轮基本参数定义, 由齿距与齿轮模数的关系, 则改变齿距后蜗杆的新模数 $m'_i$ 为

$$m'_i = \frac{P_i - \Delta P}{\pi} \quad (23)$$

式中: $p_i$ 为端面齿距。

转换为标准参数下的模数 $m'_n$ 为

$$m'_n = \frac{P_i - \Delta P}{\pi} \cos \alpha \quad (24)$$

不等齿距设计除对蜗杆模数进行调整外, 其他设计参数如齿数、螺旋角、压力角等均不改变, 其他计算参数如齿顶圆、分度圆和齿根圆等按新的模数进行计算。采用模数改变的蜗杆与不改变模数的斜齿轮进行相配, 两者的安装中心距按改变模数前的距离进行安装。由式(24)得到, 在新的模数下, 在单齿斜齿轮承受最大载荷的位置, 相邻两啮合齿的综合接触变形分别为 $\delta'_1, \delta'_2, \dots, \delta'_N$ , 啮合点位置在齿距改变前后不变, 啮合半径依然为 $r_1, r_2, \dots, r_N$ , 啮合点与斜齿轮齿根圆之间的距离为 $h_1, h_2, \dots, h_N$ , 由式(1)得到齿距改变后的斜齿轮齿根弯曲变形率 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_N$ , 对其进行校验。

### 3 实验验证

#### 3.1 实验准备

根据不等齿距理论的设计计算, 加工得到2支不同齿距的蜗杆与同一塑料齿轮配合, 一对蜗杆副是传统的等齿距的金属蜗杆与斜齿轮, 一对是不等齿距的金属蜗杆与塑料斜齿轮啮合, 蜗杆的齿距小于塑料斜齿轮的齿距。参数如表2所

示, 样板如图4和图5所示, 蜗杆齿距测试如图6和图7所示。

表2 实验样板参数

Table 2 Experimental model parameters			
参数	等齿距蜗杆	不等齿距蜗杆	塑料斜齿轮
法向模数	1	0.985	1
法向压力角/(°)	12	12	12
螺旋角/旋向	9°9'44"(左)	9°9'44"(左)	9°9'44"(左)
齿数	1	1	63
轴向齿距/mm	3.182 2	3.134 5	3.182 2
齿顶圆直径/mm	7.9	7.9	65.35
齿根圆直径/mm	4.36	4.36	61.9

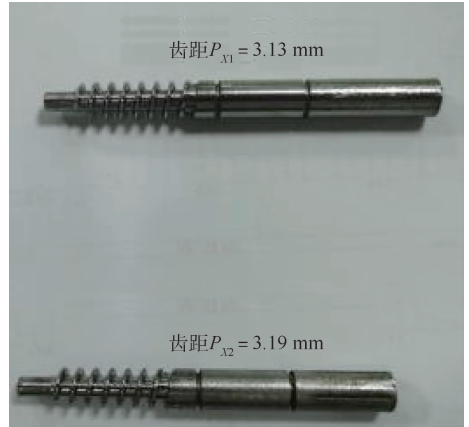


图4 两种齿距蜗杆样板

Fig. 4 Worm model with two pitches

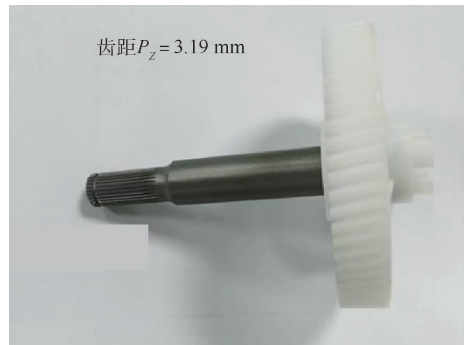


图5 正常齿距塑料齿轮样板

Fig. 5 Plastic gears model with normal pitch



图6 蜗杆齿距测量(齿距为3.19 mm)

Fig. 6 Worm pitch measurement (tooth pitch is 3.19 mm)



图 7 蜗杆齿距测量(齿距为 3.13 mm)

Fig. 7 Worm pitch measurement (tooth pitch is 3.13 mm)

将蜗杆与塑料斜齿轮固定在实验台上,实验台示意图如图 8 所示。保持相等的中心距啮合,在蜗杆端部通过力学性能仪匀速加力(见图 9),直至塑料齿轮破坏,记录数据。

### 3.2 实验测试结果

完成实验后,等齿距啮合实验结果曲线如图 10 所示,不等齿距啮合实验结果曲线如图 11 所示。

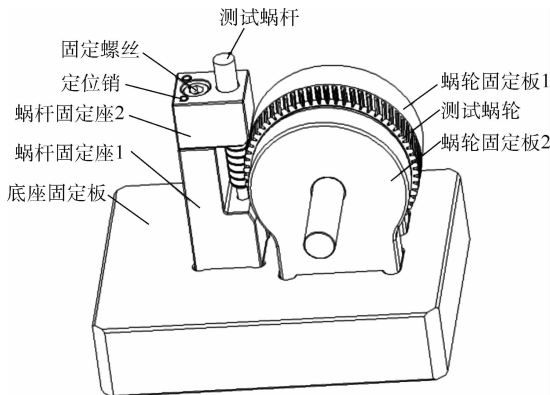


图 8 测试工作台

Fig. 8 Test fixture

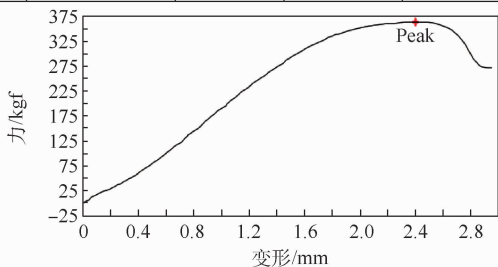


图 9 力学性能仪

Fig. 9 Mechanical properties meter

力学性能试验报告

客户名称		测试人员		试验日期	2020/8/11 15:35:19
材料名称	等齿距	试验标准		测试规格	0.2*1.5
试验速度	20.000 mm/min	标距	100.000 mm		



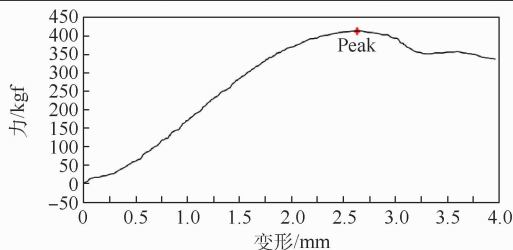
No.	设定值/%	最大力/kgf	最大压缩/mm	持压时间/min
1	6 000.000	364.010	3.086	0.000

图 10 等齿距啮合实验结果曲线

Fig. 10 Curve of results of equal pitch meshing experiment

力学性能试验报告

客户名称		测试人员		试验日期	2020/8/11 15:26:12
材料名称	不等齿距	试验标准		测试规格	0.2*1.5
试验速度	20.000 mm/min	标距	100.000 mm		



No.	设定值/%	最大力/kgf	最大压缩/mm	持压时间/min
1	6 000.000	413.850	4.220	0.000

图 11 不等齿距啮合实验结果曲线

Fig. 11 Curve of results of unequal pitch meshing experiment

从数据输出曲线分析,等齿距的金属蜗杆与斜齿轮啮合能承受 364.01 kgf 的力(1 kgf = 9.8 N),而不等齿距的金属蜗杆与斜齿轮啮合能承受 413.85 kgf 的力,承载能力提高 13.69%。

## 4 结 论

1) 本文在传统等齿距蜗杆斜齿轮啮合理论的基础上,推导了其不等齿距啮合方法,并建立了相应的实验模型。

2) 等齿距啮合时,各啮合齿的齿根弯曲变形率不同,反映了各齿承受的载荷不同,则齿轮的最大承载力受限于某一个齿所能承受的最大载荷,此时可以通过调整蜗杆的齿距,使各齿的齿根弯曲变形率相同,亦即使各齿的承载相同,从而得到蜗杆斜齿轮不等齿距啮合的思想,并在传统等齿距啮合的基础上推导了不等齿距啮合时蜗杆的调整量。

3) 通过实例可知,等齿距蜗杆齿距需减小 0.06 mm,并由静态强度实验可知,此时塑料齿轮的承载能力可提高 13.69%,验证了不等齿距啮合方法是合理的。

## 参考文献 (References)

[1] 云永璇,胡泓,塔静宁. 塑料斜齿轮与钢制蜗杆啮合力变化规律的研究[J]. 机械传动,2018,42(11):27-32.  
YUN Y H, HU H, TA J N. Research of the variation of meshing force of the plastic helical gear and steel worm[J]. Journal of Mechanical Transmission, 2018, 42(11): 27-32 (in Chinese).

[2] 欧阳志喜,石照耀. 塑料齿轮设计与制造[M]. 北京:化学工业出版社,2011:343-350.  
OUYANG Z X, SHI Z Y. Design and manufacturing of plastic gears[M]. Beijing: Chemical Industry Press, 2011: 343-350 (in Chinese).

[3] LETZELTER E, VAUJANY J P. Quasi static load sharing model in the case of nylon 6/6 cylindrical gears[J]. Materials & Design, 2009, 30(10): 4360-4368.

[4] 姜碧琼,任重义,段建中. 双圆弧齿轮啮合点载荷分配研究[J]. 机械传动,2017,41(4):58-61.  
JIANG B Q, REN Z Y, DUAN J Z. Research of the load distribution at the meshing point of double circular arc gear[J]. Journal of Mechanical Transmission, 2017, 41(4): 58-61 (in Chinese).

[5] 符双学,周长江,韩旭. 典型工况下啮合齿间有效载荷分布及齿轮强度分析[J]. 机械传动,2016,40(8):102-106.  
FU S X, ZHOU C J, HAN X. Payload distribution and gear strength analysis of gear meshing process under the typical contact condition[J]. Journal of Mechanical Transmission, 2016, 40(8): 102-106 (in Chinese).

[6] 王博. 塑料斜齿轮钢制蜗杆传动强度与传动性能研究[D]. 长春:吉林大学,2016:56-58.

WANG B. Transmission strength and performance study of plastic helical gear and steel worm drive[D]. Changchun: Jilin University, 2016: 56-58 (in Chinese).

[7] 孙训方. 材料力学[M]. 5版. 北京:科学出版社,2009:205-210.  
SUN X F. Mechanics of materials[M]. 5th ed. Beijing: Science Press, 2009: 205-210 (in Chinese).

[8] 刘鸿文. 材料力学[M]. 5版. 北京:高等教育出版社,2011:286-290.  
LIU H W. Mechanics of materials[M]. 5th ed. Beijing: Higher Education Press, 2011: 286-290 (in Chinese).

[9] ELKHOLY A H. Tooth load sharing in high-contact ratio spur gears[J]. ASME Journal of Mechanisms, Transmissions, and Automation in Design, 1985, 107(1): 11-16.

[10] FALAH A H, ELKHOLY A H. Load and stress analysis of cylindrical worm gearing using tooth slicing method[J]. Transactions of the Canadian Society for Mechanical Engineering, 2006, 30(1): 97-112.

[11] LITVIN F L, FUENTES A. Gear geometry and applied theory[M]. Cambridge: Cambridge University Press, 2004: 441-475.

[12] SHIGLEY J A, MISCHEKE C R, BUDYNAS R G. Mechanical engineering design[M]. 9th ed. New York: McGraw-Hill Inc, 2010: 721-780.

[13] 郝一舒,李磊. 基于 COSMOS/Works 的塑料斜齿轮与钢制蜗杆啮合特性研究[J]. 机械设计, 2007, 24(2): 56-59.  
HAO Y S, LI L. Research on the characteristics of engagement between plastic helical gear and steel worm based on COSMOS/Works[J]. Journal of Machine Design, 2007, 24(2): 56-59 (in Chinese).

[14] 刘舸. 圆柱蜗杆斜齿轮传动的理论分析及试验研究[D]. 重庆:重庆大学,2004:78.  
LIU G. Theoretical analysis and experimental investigation of an involute cylindrical worm-helical gear drive[D]. Chongqing: Chongqing University, 2004: 78 (in Chinese).

[15] 张生明. 斜齿轮与蜗杆配合方法和工艺探析[J]. 中国医疗设备, 2009, 24(12): 24-26.  
ZHANG S M. Discussion on coordinated methods and technique of bevel wheel and worm[J]. Information of Medical Equipment, 2009, 24(12): 24-26 (in Chinese).

[16] 李磊. 塑料蜗轮与钢制蜗杆的啮合性能研究[D]. 上海:同济大学,2007:83.  
LI L. Research of the gearing characteristics between plastic worm wheel and steel worm[D]. Shanghai: Tongji University, 2007: 83 (in Chinese).

[17] 魏敏,魏勇. 汽车微电机中蜗杆斜齿轮啮合传动设计[J]. 装备制造技术, 2012(9): 49-51.  
WEI M, WEI Y. The design of worm and spiral gear meshing transmission in the automobile micro motors[J]. Manufacturing Technology, 2012(9): 49-51 (in Chinese).

[18] 鲍洪,安琦. 渐开线直齿齿轮齿载荷及应力计算方法[J]. 华东理工大学学报(自然科学版), 2012, 38(1): 116-122.  
BAO H, AN Q. A method of calculating load and stress of involute spur gear[J]. Journal of East China University of Science and Technology (Natural Science Edition), 2012, 38(1): 116-122 (in Chinese).

- [19] HEINZ P B. Understanding right-angle gear backlash[J]. Hydrocarbon Processing,2017,1:54.
- [20] 罗光汉. 基于 Pro/E 的斜齿齿轮副变位系数与侧隙的优化设计[J]. 制造业自动化,2016,38(2):115-116.
- LUO G H. The optimization of modification coefficient and backlash of helical gear pairs based on Pro/E[J]. Manufacturing Automation,2016,38(2):115-116 (in Chinese).
- [21] 郝一舒,岳滨楠. 塑料斜齿轮与钢制蜗杆的啮合理论分析[J]. 机械传动,2009,33(5):9-11.
- HAO Y S, YUE B N. Analysis of the mesh theory between plastic helical gear and steel worm [J]. Journal of Mechanical Transmission,2009,33(5):9-11 (in Chinese).
- [22] 郝一舒,马士涛. 塑料斜齿轮与钢制蜗杆传动的非线性边界分析[J]. 机械设计,2008,25(7):45-48.
- HAO Y S, MA S T. Nonlinear boundary analysis of plastic helical gear and steel worm transmission [J]. Journal of Machine Design,2008,25(7):45-48 (in Chinese).

## Method of unequal pitch meshing of metal worm and plastic helical gear

REN Jihua<sup>1</sup>, SHI Zhaoyao<sup>1,\*</sup>, WANG Defeng<sup>2</sup>, YU Shipeng<sup>2</sup>

(1. College of Mechanical Engineering and Applied Electronics Technology, Beijing University of Technology, Beijing 100124, China;

2. Dongguan Xinghuo Gear Co., Ltd., Dongguan 523723, China)

**Abstract:** In order to further improve the bearing capacity of plastic gears in the transmission of metal worm and plastic helical gears, the force characteristics of the traditional equal-pitch worm helical gear meshing transmission were studied in depth, and the unequal pitch meshing of the worm and the helical gear was innovatively proposed. Based on beam bending theory and gear tooth deformation theory, the relationship between tooth surface load, deformation and contact stiffness is obtained. On the premise that the bending deformation rate of the gear tooth root is equal, the theoretical design method of unequal meshing is deduced, and the pitch adjustment of the worm in the case of unequal pitch meshing is obtained, which is verified by static strength experiments. The experimental results show that the unequal pitch design can increase the load-bearing capacity of the plastic helical gear by 13.69%.

**Keywords:** unequal tooth pitch; metal worm; plastic helical gear; meshing theory; tooth root bending stress

Received: 2020-12-10; Accepted: 2020-12-31; Published online: 2021-03-02 12:28

URL: [kns.cnki.net/kcms/detail/11.2625.V.20210302.0946.002.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210302.0946.002.html)

Foundation item: Key Program of National Natural Science Foundation of China (51635001)

\* Corresponding author. E-mail: shizhaoyao@bjut.edu.cn

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2020.0711

# 卷积神经网络卫星信号自动调制识别算法

崔天舒<sup>1,2</sup>, 崔凯<sup>1,2</sup>, 黄永辉<sup>1,2,\*</sup>, 赵文杰<sup>1</sup>, 安军社<sup>1,2</sup>

(1. 中国科学院国家空间科学中心 复杂航天系统电子信息技术重点实验室, 北京 100190;

2. 中国科学院大学 计算机科学与技术学院, 北京 100049)

**摘 要:** 自动调制识别是空间认知通信系统的关键技术,有助于实现自适应信号解调。深度神经网络虽然具有特征提取能力强的优势,但也存在参数众多、计算量大的问题,难以实现空间在轨应用。针对以上问题,提出了一种轻量化、高性能的卷积神经网络结构。网络先提取信号的同相正交相关特征,再提取时域特征,最后提取各通道特征均值进行分类。对11种调制方式分类的实验结果表明:当信噪比高于0 dB时,平均识别准确率能达到86.94%,较传统的高阶累积量的方法提高了31.54%;与目前高识别准确率的深度神经网络模型相比,仅使用不到10%的模型参数,在树莓派4B上计算速度平均提高了20倍。

**关键词:** 认知通信系统; 调制方式识别; 同相正交相关特征; 卷积神经网络; 深度学习  
**中图分类号:** V47; TN927+.21

**文献标志码:** A **文章编号:** 1001-5965(2022)06-0986-09

下一代卫星通信系统主要旨在增强全球网络吞吐量,同时确保服务质量<sup>[1]</sup>。NASA正在开展应用于未来通信体系架构的认知通信技术,这些技术有望降低网络的运营复杂度,增加科学数据返回能力。自动调制识别是空间认知通信系统的关键技术,有助于实现接收机自动配置、干扰抵消和频谱管理等功能<sup>[2]</sup>。

传统的自动调制识别算法主要分为2类:基于最大似然假设检验的方法<sup>[3-4]</sup>与基于特征识别的方法<sup>[5-7]</sup>。基于最大似然假设检验的方法是在对某类调制信号的统计特性进行分析的基础上得到某种判决准则,存在计算复杂度高和适用范围窄等缺点;基于特征识别的方法是设计分类器对专家人工设计的特征进行分类,经常采用的特征有小波域特征<sup>[5]</sup>、循环谱<sup>[6]</sup>和高阶统计量<sup>[7]</sup>。人工提取特征只能对特定的几种调制方式进行有效识别,存在通用性差的问题。

深度学习是一种数据驱动的人工智能算法,能从数据中自动提取数据特征并进行分类识别。近年来,随着深度学习在机器视觉<sup>[8]</sup>、语音识别<sup>[9]</sup>等研究领域取得了巨大的成功,基于深度学习的方法在自动调制识别领域的研究<sup>[9-19]</sup>也越来越多。目前,基于深度学习的自动调制识别算法主要有2类:①将AD采样后得到的原始同相正交(in-phase and quadrature, IQ)信号先转化成图像,如时频图<sup>[9,11]</sup>、星座图<sup>[12-13]</sup>、Choi-Williams时频分布<sup>[14]</sup>等,再用图像识别领域的深度学习算法对调制方式进行分类,虽然更像人类视觉的处理,可理解性更强,但是在IQ信号转换成图像的处理过程中带来额外的计算量;②直接采用神经网络对原始IQ信号进行处理<sup>[14-19]</sup>,由于原始IQ信号是数字通信系统后端处理的基本输入,相对于其他需要进行预处理的方法,其蕴含了更多的无损信息,而且无需增加预处理过程中额外的计算量,

收稿日期: 2020-12-24; 录用日期: 2021-03-12; 网络出版时间: 2021-04-27 10:47

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20210426.1620.001.html](http://kns.cnki.net/kcms/detail/11.2625.V.20210426.1620.001.html)

基金项目: 中国科学院复杂航天系统电子信息技术重点实验室自主部署基金(Y42613A32S)

\* 通信作者. E-mail: yonghui@nssc.ac.cn

**引用格式:** 崔天舒, 崔凯, 黄永辉, 等. 卷积神经网络卫星信号自动调制识别算法[J]. 北京航空航天大学学报, 2022, 48(6): 986-994. CUI T S, CUI K, HUANG Y H, et al. Convolutional neural network based algorithm for automatic modulation recognition of satellite signals[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(6): 986-994 (in Chinese).

更有利于嵌入式设备的应用。

由于体积、质量、功耗的约束,以及空间辐射、极端温度、维修困难等环境因素的影响,星载计算机的计算能力与存储空间相对于地面计算机具有非常大的差距。而深度神经网络虽然具有特征提取能力强、识别准确率高的优点,但也存在网络参数众多、计算量大的不足,因此,面向卫星在轨应用的神经网络结构,一定要具备轻量化和计算高效等特点。

本文面向星载在轨应用的自动调制识别系统,提出了一种基于 IQ 相关特征的卷积神经网络结构(convolutional neural network structure based on IQ correlation features, IQCNet),该网络以原始 IQ 信号为输入,先提取信号的 IQ 相关特征,再提取时域特征,最后采用全局平均池化获得各通道特征均值进行分类。相对于同样直接处理原始 IQ 信号的 CNN2<sup>[15]</sup>、CLDNN<sup>[16]</sup>及 CNN\_LSTM<sup>[20]</sup>等网络结构,本文算法具有信息利用率高、参数数量少和计算量小的特点,更有利于部署到星载计算机上进行应用。

## 1 卫星智能接收系统

### 1.1 卫星智能接收系统架构

图 1 为基于零中频架构的卫星智能接收系统。已训练完成的神经网络参数部署在自动调制识别系统上,可在没有任何先验信息的情况下识别接收信号的调制类型,并进行同步和解调来获得信号的数据内容。

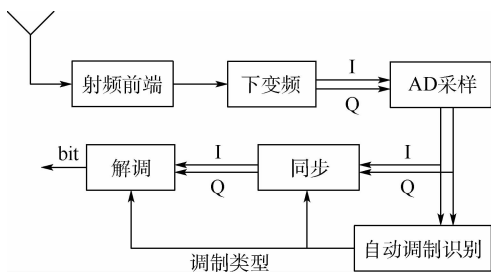


图 1 卫星智能接收系统

Fig. 1 Intelligent receiver system for satellite

### 1.2 调制信号模型

假定接收机输出的信号模型为

$$r[n] = \alpha e^{j(2\pi f_c t + \theta_0)} s[n] + w[n] \quad (1)$$

式中: $\alpha$ 为信号幅值; $r[n]$ 为第  $n$  个接收信号样本; $s[n]$ 第  $n$  个发射信号样本; $w[n]$ 表示加性高斯白噪声; $f_c$ 为载频; $\theta_0$ 为相位偏移。

识别调制信号共有 11 种,其中模拟调制方式包括 AM、PM 与 FM 等 3 种类型,数字调制方式包括 BPSK、QPSK、8PSK、CPFSK、GFSK、PAM4、

16QAM、64QAM 等 8 种类型。

## 2 IQCNet 网络结构

卷积神经网络(convolution neural network, CNN)<sup>[21]</sup>作为一种普遍使用的深度神经网络结构,由于其良好的特征提取能力,近年来逐步应用于电磁信号处理领域。卷积神经网络通过结合局部感知、权值共享、池化降采样和非线性映射等对数据特征逐层提取并高度抽象,提取信号特征后再进行分类。卷积神经网络由多个卷积层构成,每个卷积层包含多个卷积核,采用权值共享来减少网络参数规模。

针对原始 IQ 信号,本文提出了基于 IQ 相关特征的卷积网络结构 IQCNet,其充分利用了时序复信号的时域特征与复数特征。首先,采用卷积核构造 IQ 相关特征滤波器,提取 IQ 相关特征,达到充分利用复信号 I、Q 之间幅度及相位信息的目的,且将随后的二维数据降为了一维数据处理,计算量降低了 50%;然后,采用多层小卷积核替代大卷积核提取信号时域特征,既保证了时间尺度上的感受野,又降低了卷积核参数数量;最后,放弃采用多个全连接分类层,而是采用全局平均池化获得各通道特征值直接进行分类,在达到一定泛化能力的同时继续降低网络参数数量。

图 2 为 IQCNet 网络结构,输入数据是时间长度为  $N$  的 IQ 序列,输出为  $L$  个类别,编号为  $0 \sim L - 1$ 。IQCNet 在每个滤波通道内部,第 1 个卷积层采用(1,2)的滤波器提取 IQ 相关特征,使得 IQ 方向数据维度由  $N \times 2$  变为了  $N \times 1$ ,将后续处理的计算量降为了原来的一半;然后借鉴 VGG 网络<sup>[22]</sup>采用多层(3,1)小滤波器提取时域特征,在保证感受野的条件下减少卷积核参数数量;在每个滤波器后再采用最大池化降低数据维度,进一步降低时域方向计算量。

在分类阶段,借鉴 ResNet 网络<sup>[8]</sup>采用的自适应平均池化(adaptive average pooling)替代 Flatten,将每个特征通道的特征均值作为新的特征值,提高网络的泛化性能;再在分类时仅采用 1 个全连接层,以降低参数数量与计算量。

除此之外,在卷积层的输出与激活函数之间增加批归一化(batch normalization, BN)<sup>[23]</sup>操作,以增加模型的鲁棒性和训练速度,并防止过拟合;网络中采用 ReLU 作为激活函数,在分类层采用 Sigmoid 作为分类函数。

IQCNet 网络结构参数是可灵活配置的。IQCNet 网络由  $H$  个卷积层与 1 个全连接层构成,

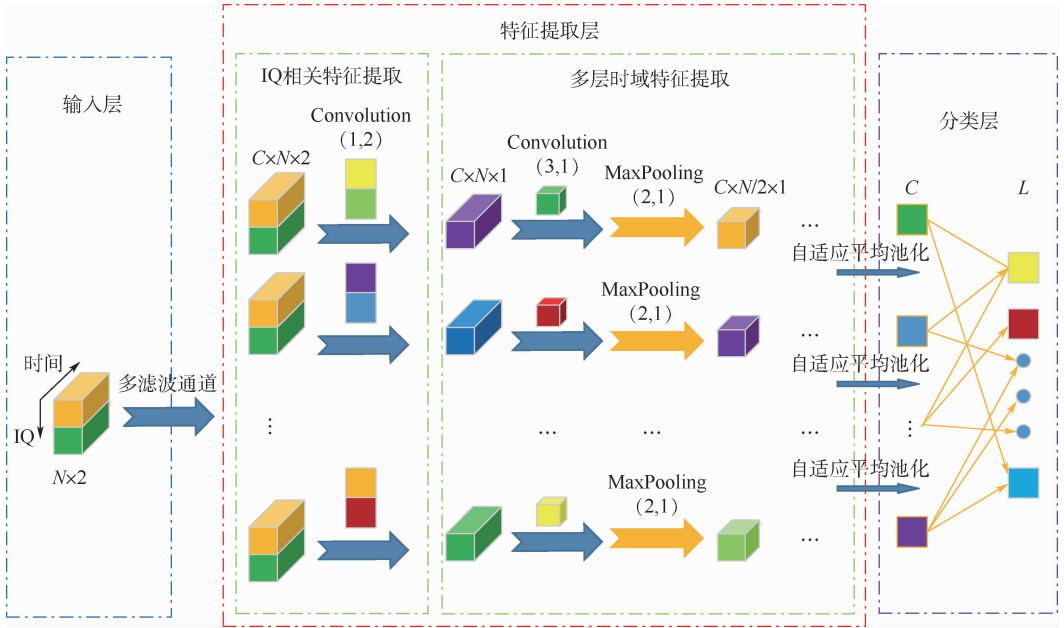


图2 IQCNet网络结构

Fig. 2 IQCNet network structure

且每个卷积层包含  $C$  个特征提取通道。为表示不同结构参数的 IQCNet 网络,用  $\text{IQCNet}(H, C)$  表示具体结构参数,其中  $H$  表示卷积层的层数,  $C$  表示每个卷积层的卷积核数量(通道数)。例如,  $\text{IQCNet}(2, 32)$  表示具有 2 个卷积层,且每层包含 32 个卷积核的 IQCNet 网络。表 1 为  $\text{IQCNet}(2, 32)$  的网络结构,其中数据输入格式为  $128 \times 2$ 。

表 1 IQCNet(2,32)网络结构

Table 1 IQCNet(2,32) network structure

层名称	输入尺寸	尺寸/步进	卷积核数量
Conv2d-1	$128 \times 2$	$1 \times 2/1 \times 1$	32
BatchNorm2d-1	$128 \times 1$		32
MaxPool2d-1	$128 \times 1$	$2 \times 1/2 \times 1$	32
Conv2d-2	$64 \times 1$	$1 \times 3/1 \times 1$	32
BatchNorm2d-2	$64 \times 1$		32
MaxPool2d-2	$64 \times 1$	$2 \times 1/2 \times 1$	32
AdaptiveAvgPool2d	$64 \times 1$		32
Linear-1	16		

### 3 对比算法

#### 3.1 高阶累积量

累积量(cumulants)是调制方式识别常用的高阶统计量特征,而 K 最近邻(K-nearest neighbor, KNN)分类算法是机器学习中常用的分类算法,选取累积量与 KNN 的组合(Cumulants + KNN)作为传统人工特征的方法的代表,以此作为比较基准。在 KNN 分类器中,选择  $k$  最近邻个数为 5。

对于一个具有零均值的信号  $x(t)$ ,  $p$  阶混合矩可表示为

$$\mathbf{M}_{pq} = E[X(t)^{(p-q)} X^*(t)^q] \quad (2)$$

$C_{pq}$  为随机变量  $X(t)$  的  $p$  阶联合累积量,可用混合矩  $\mathbf{M}_{pq}$  的多项式来表示, Cumulants + KNN 算法选取 8 个高阶累积量作为特征向量 ( $C_{20}, C_{21}, C_{40}, C_{41}, C_{42}, C_{60}, C_{63}, C_{80}$ ), 其累积量与联合矩的关系为

$$C_{20} = M_{20} \quad (3)$$

$$C_{21} = M_{21} \quad (4)$$

$$C_{40} = M_{40} - 3M_{20}^2 \quad (5)$$

$$C_{41} = M_{41} - 3M_{21}M_{20} \quad (6)$$

$$C_{42} = M_{42} - |M_{20}|^2 - 2M_{21}^2 \quad (7)$$

$$C_{60} = M_{60} - 15M_{40}M_{20} - 30M_{20}^3 \quad (8)$$

$$C_{63} = M_{63} - 6M_{41}M_{20} - 9M_{42}M_{21} - 18M_{20}^2M_{21} - 12M_{21}^3 \quad (9)$$

$$C_{80} = M_{80} - 28M_{20}M_{60} - 35M_{40}^2 - 420M_{20}^2M_{40} - 630M_{20}^4 \quad (10)$$

#### 3.2 对比网络结构

为验证 IQCNet 中提取 IQ 相关特征能否有效提高识别准确率,设计了一个除未提取 IQ 相关特征之外,网络其他部分与 IQCNet 完全相同的网络,并命名为  $\text{IQCNet-N}(H, C)$ 。将表 2 中的  $\text{IQCNet-N}(2, 32)$  与表 1 的  $\text{IQCNet}(2, 32)$  对比发现,两者的差别仅在于 IQCNet 第 1 个卷积层采用了 (1,2) 的滤波器提取 IQ 相关特征,并将数据维度值由 2 降到了 1,而  $\text{IQCNet-N}$  未做此处理。

此外,实验还选择了  $\text{CNN2}^{[15]}$ 、 $\text{CLDNN}^{[16]}$  及  $\text{CNN\_LSTM}^{[20]}$  等同样直接处理原始 IQ 信号的深度学习算法与 IQCNet 算法进行对比,网络参数如

表 3 所示。其中, CNN2 与 IQCNet 都是 CNN 结构,而 CLDNN 与 CNN\_LSTM 是结合 CNN 与深度长短期记忆网络(long short-term memory, LSTM)的网络模型。CNN2 为最早采用卷积神经网络识别调制方式的经典结构;CLDNN<sup>[24]</sup>是语音识别领域的经典结构,其将 CNN、LSTM 与 DNN 有效结合,充分利用了 CNN 频谱特征提取、LSTM 时间建模和 DNN 分类等各自的优势;而 CNN\_LSTM 是根据 CLDNN 改进的网络结构,在 RML2016. 10a 数据集使用更少参数取得了更高的识别准确率。

表 3 对比网络参数

Table 3 Parameters of networks for comparison

网络名称	卷积层数	卷积核尺寸	通道数	LSTM 层数	LSTM 单元数
CNN2	2	(1,3),(2,3)	256,80	0	0
CLDNN	3	(1,8)	50,50,50	1	50
CNN_LSTM	2	(1,3),(2,3)	128,32	1	128

## 4 实验验证

### 4.1 数据集

实验采用文献[16]中 GNU Radio 软件平台仿真生成的 IQ 信号数据集 RML2016. 10a 进行验证。

RML2016. 10a 数据集包含 11 种信号调制方式,其中包含 8 种数字调制方式(BPSK、QPSK、8PSK、16QAM、64QAM、GFSK、CPFSK、PAM4)和 3 种模拟调制方式(WB-FM、AM-SSB、AM-DSB)。数据集在 -20 ~ 18 dB 的信噪比环境下生成,每符号 8 个 IQ 采样点。每种调制方式在每个信噪比条件下的样本数为 1 000 个,故共有 22 000 个样本。每个样本的数据格式为 128 × 2,其中,2 表示 I、Q 两路采样信号,128 为采样点长度。

### 4.2 实验条件

在 Ubuntu 18.04.4 LTS 64 位操作系统上,使用 Python 编程语言及 Tensorflow 深度学习框架完成网络的搭建、训练和测试。其中,网络训练数据的批处理大小(batch size)为 1 024;损失函数采用交叉熵(cross entropy loss),优化器采用 Adam 优化器,学习率(learning rate)为 0.001,其余参数为 Tensorflow 中的默认参数。

实验中硬件设备共有 2 类:工作站和嵌入式设备。其中,工作站用于网络的训练,嵌入式设备用于计算算法的推理时间。实验设备信息如表 4 所示。

表 4 实验设备

Table 4 Experimental equipment

测试设备	CPU	GPU	内存/GB
PC	i9-7920X	RTX 2080Ti	64
Jetson Nano	Cortex-A57	128 个 CUDA 核	4
树莓派 4B	Cortex-A72	无	4

表 2 IQCNet-N(2,32)网络结构

Table 2 IQCNet-N(2,32) network structure

层名称	输入尺寸	尺寸/步进	卷积核数量
Conv2d-1	128 × 2	1 × 2/1 × 1	32
BatchNorm2d-1	128 × 2		32
MaxPool2d-1	128 × 2	2 × 1/2 × 1	32
Conv2d-2	64 × 2	1 × 3/1 × 1	32
BatchNorm2d-2	64 × 2		32
MaxPool2d-2	64 × 2	2 × 1/2 × 1	32
AdaptiveAvgPool2d	64 × 2		32
Linear-1	16		

### 4.3 评估方法

实验采用识别准确率 Accuracy 对网络模型的泛化能力进行性能度量,对于二分类问题,根据真实值和预测值的结果,可以将整个样本集分为真正例 TP、伪正例 FP、真反例 TN 和伪反例 FN。其中,  $Accuracy = (TP + TN) / (TP + TN + FP + FN)$ ,即所有分类正确的样本数与全体样本数的比例。

性能评估采用多次随机实验求平均值的方法获得测试结果,如图 3 所示。从数据集中随机抽取组成训练样本、验证样本与测试样本,使用随机参数初始化网络模型后,进行模型训练、验证与测试,将 N 次测试结果的平均值作为最终的测试结果,所有实验都进行 10 次随机训练与测试。

单次训练、验证与测试过程如图 4 所示。首先,采用无放回随机采样的方式,将数据集按 8:2 的比例生成训练样本与验证样本;然后,对训练样本进行训练,每个训练轮次(epoch)都进行验证,训练 80 个轮次后,选取在验证样本中获得最高识

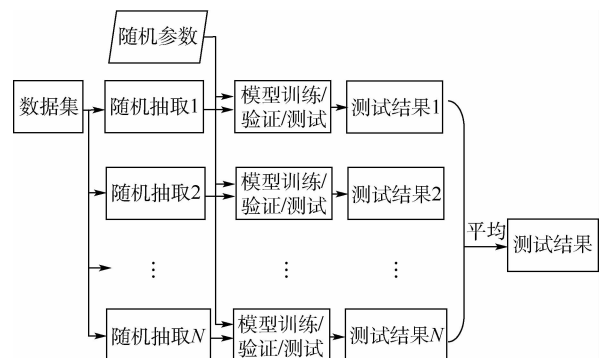


图 3 实验总体流程

Fig. 3 Flowchart of overall experiment process

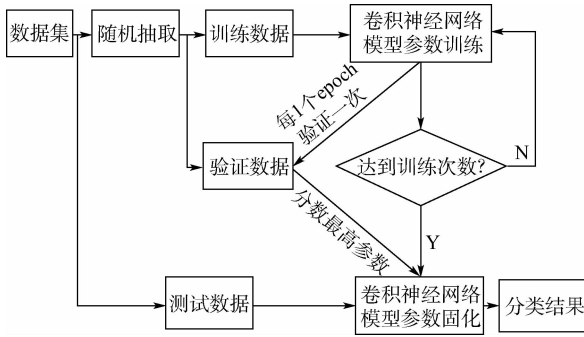


图4 实验详细流程

Fig. 4 Flowchart of experimental procedures

别准确率的参数作为最终模型参数;最后,采用全部样本中各信噪比下的样本作为测试样本,获得各信噪比下的识别准确率。

### 4.4 识别准确率

#### 4.4.1 IQ 相关特征有效性

实验分别选取了不同深度与卷积核数量的 IQCNet 与 IQCNet-N 网络进行对比,验证 IQCNet 提取的 IQ 相关特征是否有效,是否有利于提高调制方式的识别准确率。

图5为采用3、4和5个卷积层且每层包含16个卷积核的 IQCNet 与 IQCNet-N 网络结构的识别准确率。图6为采用3、4和5个卷积层且每层包含24个卷积核的 IQCNet 与 IQCNet-N 网络结构的识别准确率。图7为采用3、4和5个卷积层且每层包含32个卷积核的 IQCNet 与 IQCNet-N 网络结构的识别准确率。

由图5~图7可知,当网络深度和卷积核数量完全相同时,在各个信噪比下,IQCNet 网络比 IQCNet-N 的识别准确率都有明显的提高。

表5为信噪比分别在-20~18 dB与0~18 dB范围内时,IQCNet 与 IQCNet-N 在采用16、24和

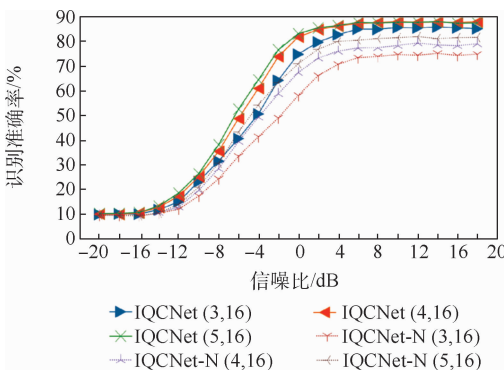


图5 16个卷积核时不同深度的网络结构识别准确率比较

Fig. 5 Comparison of recognition accuracy between different depths of network structure under 16 convolution kernels

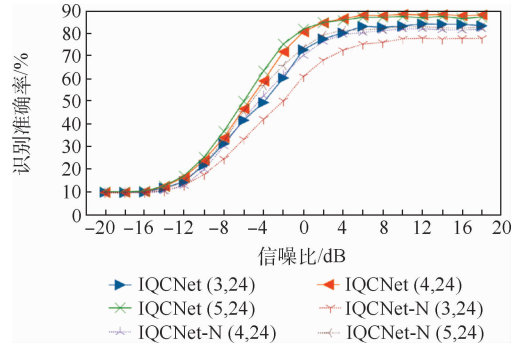


图6 24个卷积核时不同深度的网络结构识别准确率比较

Fig. 6 Comparison of recognition accuracy between different depths of network structure under 24 convolution kernels

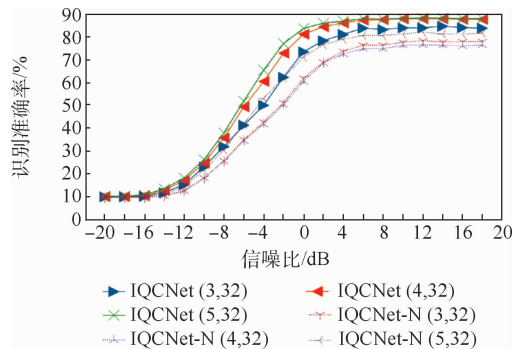


图7 32个卷积核时不同深度的网络结构识别准确率比较

Fig. 7 Comparison of recognition accuracy between different depths of network structure under 32 convolution kernels

32个卷积核时,在不同深度下的平均识别准确率,采用3、4和5个卷积层时不同卷积核数的平均识别准确率,以及不同深度不同卷积核数量下的全部平均准确率。

总体来看,IQCNet 与 IQCNet-N 的识别准确率都会随网络深度的增加而提高,但卷积核数量的增加并没有一定带来识别准确率的增加。但不管何等深度和卷积核数量下,IQCNet 都要比 IQCNet-N 的识别准确率高。

当信噪比在-20~18 dB范围内时,采用16个卷积核,IQCNet 比 IQCNet-N 的平均识别准确率高7.61%;而当采用24和32个卷积核时,IQCNet 比 IQCNet-N 的平均识别准确率分别提高了5.11%与7.75%。而若仅考察信噪比大于0 dB时的数据,当采用16个卷积核时,IQCNet 比 IQCNet-N 的平均识别率提高9.76%,当采用24和32个卷积核时,IQCNet 比 IQCNet-N 的平均识别率分别提高了6.57%与9.48%。

将不同网络深度与卷积核数量下的数据取均值,IQCNet 比 IQCNet-N 的平均识别准确率,在信

表 5 平均识别准确率

Table 5 Average recognition accuracy

网络名称	平均识别准确率/%	
	信噪比 -20 ~ 18 dB	信噪比 0 ~ 18 dB
IQCNet-N(3, -)	47.70	73.47
IQCNet-N(4, -)	50.33	76.49
IQCNet-N(5, -)	53.49	80.14
IQCNet(3, -)	54.36	82.36
IQCNet(4, -)	58.37	86.72
IQCNet(5, -)	59.27	86.82
IQCNet-N(-, 16)	50.07	75.90
IQCNet-N(-, 24)	51.75	78.38
IQCNet-N(-, 32)	49.70	75.82
IQCNet(-, 16)	57.68	85.66
IQCNet(-, 24)	56.87	84.94
IQCNet(-, 32)	57.45	85.30
IQCNet-N	50.51	76.70
IQCNet	57.33	85.30

噪比 -20 ~ 18 dB 段提高了 6.82%, 在信噪比 0 ~ 18 dB 段提高了 8.60%, 说明 IQCNet 提取的 IQ 相关特征显著提高了调制方式的识别准确率, 证明 IQCNet 提取的 IQ 相关特征是有效的。

#### 4.4.2 与其他方法对比

实验选取 Cumulants + KNN、IQCNet(4, 24) 与 CNN2、CLDNN 及 CNN\_LSTM 进行对比。

由图 8 可知, 基于传统的 Cumulants + KNN 方法的识别准确率明显低于深度学习方法。而在深度学习方法中, 整体来看, 将 CNN 与 LSTM 结合的方法的识别准确率高于一采用 CNN 的方法, 这主要是因为 LSTM 能有效提取时间序列特征, 提高了识别准确率。表 6 为 5 种方法信噪比在 -20 ~ 18 dB 与 0 ~ 18 dB 范围内的平均识别准确率。

在信噪比 -20 ~ 18 dB 条件下, IQCNet(4, 24) 平均识别准确率达到 58.15%, 在信噪比 0 ~ 18 dB 条件下, 平均识别准确率达到 86.94%; 相对于传统的 Cumulants + KNN 方法, 在 -20 ~ 18 dB 时, IQCNet(4, 24) 的平均识别准确率提高了 23.61%; 在信噪比 0 ~ 18 dB 时, IQCNet(4, 24) 的平均识别准确率提高了 31.54%。说明 IQCNet 相

对于传统 Cumulants + KNN 方法的识别准确率有明显的提升。

与其他深度网络对比, CNN\_LSTM 与 IQCNet 网络的结果优于其他 3 种方法。虽然在 -20 ~ 18 dB 时, IQCNet(4, 24) 的识别准确率低于 CNN\_LSTM, 但在信噪比 0 ~ 18 dB 段, IQCNet(4, 24) 略优于 CNN\_LSTM, 说明 IQCNet 已经达到了当前较高的识别准确率水平。

为了更好地说明 IQCNet 算法对不同调制信号的识别性能, 在图 9 与图 10 中分别展示了 IQCNet(4, 24) 在信噪比 0 dB 与 6 dB 时的混淆矩阵。可知, 除 16QAM 与 64QAM、AM-DSB 与 WB-FM 等调制类型容易混淆外, 其他调制类型都能较为准确地识别。

表 6 不同方法平均识别准确率

Table 6 Average recognition accuracy of different methods

方法	平均识别准确率/%	
	信噪比 -20 ~ 18 dB	信噪比 0 ~ 18 dB
Cumulants + KNN	34.54	55.40
CNN2	56.82	78.86
CLDNN	56.84	82.92
CNN_LSTM	60.19	86.89
IQCNet(4, 24)	58.15	86.94

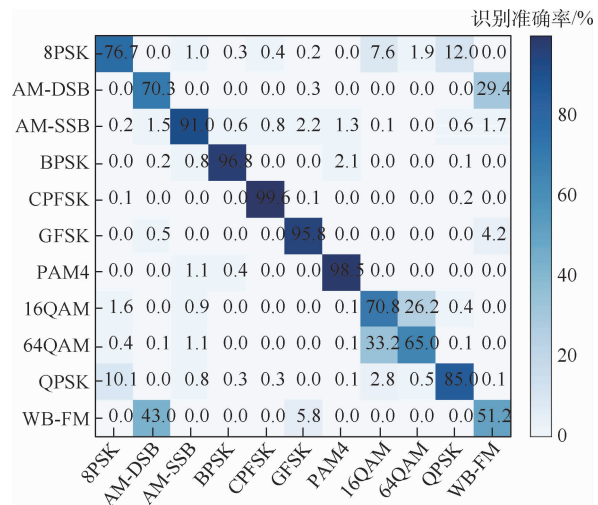


图 9 IQCNet 混淆矩阵 (信噪比为 0 dB, 识别准确率为 81.84%)

Fig. 9 IQCNet confusion matrix (SNR = 0 dB, Accuracy = 81.84%)

#### 4.5 网络参数与运行时间

实验选取了 IQCNet(4, 24) 与 CNN2、CLDNN 和 CNN\_LSTM 进行网络参数数量及网络训练时间、推理时间的对比。

由表 7 可知, IQCNet(4, 24) 的网络参数明显少于其他 3 种网络, CNN2、CLDNN 和 CNN\_LSTM 的网络参数分别为 IQCNet(4, 24) 的 901、12 和 18 倍, IQCNet 在达到相等识别准确率的条件下, 网络参数最少。

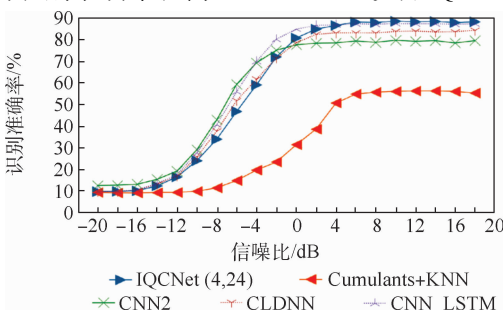


图 8 与其他方法的识别准确率比较

Fig. 8 Comparison of recognition accuracy with other methods

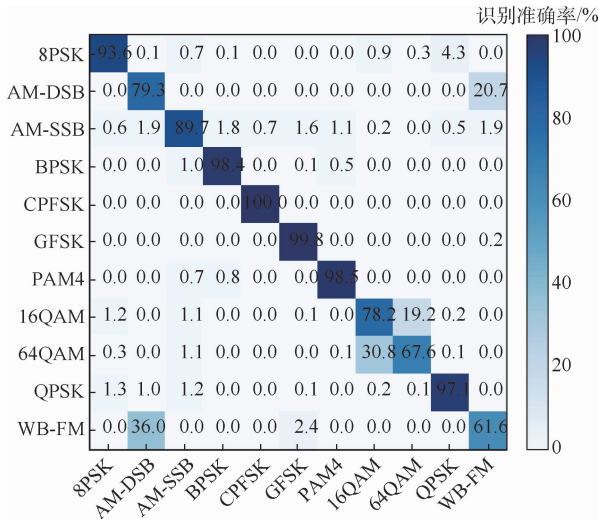


图10 IQCNet混淆矩阵(信噪比为6 dB, 识别准确率为87.62%)

Fig. 10 IQCNet confusion matrix (SNR = 6 dB, Accuracy = 87.62%)

表7 网络参数与计算时间

Table 7 Network parameters and compute time

网络名称	网络参数/个	RTX 2080Ti 训练时间/s	树莓派4B 推理时间/s	Jetson Nano 推理时间/s
CNN2	5 369 947	823	3 035	5 243
CLDNN	71 311	4 523	3 794	1 585
CNN_LSTM	108 971	2 673	3 564	990
IQCNet	5 963	171	181	473

## 5 结论

面向在轨应用的自动调制识别系统,提出了一种直接处理原始IQ信号的卷积神经网络结构,采用RML2016.10a数据集的11种信号调制方式进行验证,10次随机实验后得到以下结果:

1) 与IQCNet-N相比,IQCNet在信噪比-20~18 dB条件下,将平均识别准确率提高了6.82%,在信噪比0~18 dB条件下,将平均识别准确率提高了8.60%,说明IQCNet提取的IQ相关特征是有效的,能显著提高调制方式识别准确率。

2) 在信噪比-20~18 dB与0~18 dB条件下,IQCNet平均识别准确率分别达到了58.15%与86.94%,较传统基于Cumulants+KNN方法,同等信噪比条件下,识别准确率分别提高了23.61%与31.54%。

3) 与目前识别准确率较高的CNN\_LSTM相比,IQCNet与之达到了相同水平的识别准确率,但仅使用了前者5.5%的网络参数,训练时间仅为前者的6.4%,树莓派4B与Jetson Nano平台下推理时间仅为前者的5%与47.8%,说明IQCNet的计算速度更快,计算量更小。

实验采用RTX2080Ti作为网络训练平台对各种网络进行训练。表7为1次全数据集的训练时间。CNN2、CLDNN和CNN\_LSTM的训练时间分别为IQCNet(4,24)的4.8、26.5和15.6倍,IQCNet的训练时间最短。实验选取树莓派4B与Jetson Nano 2种不同架构的嵌入式平台测试网络的推理时间,其中Jetson Nano采用了GPU的CUDA加速,而树莓派4B仅采用CPU进行运算。

表7为在以上2种平台下1次全数据集的推理时间。在树莓派4B平台上,CNN2、CLDNN和CNN\_LSTM的推理时间分别为IQCNet(4,24)的17、21和20倍,而在Jetson Nano平台上,CNN2、CLDNN和CNN\_LSTM的推理时间分别为IQCNet(4,24)的11、3和2倍。说明IQCNet的推理时间更短,更有利于嵌入式场景下部署。

通过与其他算法在识别准确率、模型参数和运行时间等方面的比较,证明IQCNet是一种轻量化、高效率的卷积网络结构,适合于部署到星载在轨应用的自动调制识别系统中。

## 参考文献 (References)

- [1] VIDAL O, VERELST G, LACAN J, et al. Next generation high throughput satellite system[C]//2012 IEEE First AESS European Conference on Satellite Telecommunications. Piscataway: IEEE Press, 2012: 1-7.
- [2] SMITH A, EVANS M, DOWNEY J. Modulation classification of satellite communication signals using cumulants and neural networks[C]//2017 Cognitive Communications for Aerospace Applications Workshop. Piscataway: IEEE Press, 2017: 1-8.
- [3] SILLS J A. Maximum-likelihood modulation classification for PSK/QAM[C]//IEEE Military Communications. Piscataway: IEEE Press, 1999: 217-220.
- [4] WEI W, MENDEL J M. Maximum-likelihood classification for digital amplitude-phase modulations[J]. IEEE Transactions on Communications, 2000, 48(2): 189-193.
- [5] HASSAN K, DAYOUB I, HAMOUDA W, et al. Automatic modulation recognition using wavelet transform and neural network[C]//2009 9th International Conference on Intelligent Transport Systems Telecommunications. Piscataway: IEEE Press, 2009: 234-238.
- [6] DONG S L, LI Z P, ZHAO L F. A modulation recognition algo-

- rithm based on cyclic spectrum and SVM classification [C] // 2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference. Piscataway: IEEE Press, 2020:2123-2127.
- [7] FLOHBERGER M, GAPPMIR W, KOUELKA O. Modulation classifier for signals used in satellite communications [C] // 2010 5th Advanced Satellite Multimedia Systems Conference and the 11th Signal Processing for Space Communications Workshop. Piscataway: IEEE Press, 2010:198-202.
- [8] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 770-778.
- [9] ARISOY E, SETHY A, RAMABHADHRAN B, et al. Bidirectional recurrent neural network language models for automatic speech recognition [C] // 2015 IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway: IEEE Press, 2015:5421-5425.
- [10] 周鑫,何晓新,郑昌文. 基于图像深度学习的无线电信号识别[J]. 通信学报, 2019, 40(7):114-125.  
ZHOU X, HE X X, ZHENG C W. Radio signal recognition based on image deep learning [J]. Journal on Communications, 2019, 40(7):114-125 (in Chinese).
- [11] KARRA K, KUZDEBA S, PETERSEN J. Modulation recognition using hierarchical deep neural networks [C] // 2017 IEEE International Symposium on Dynamic Spectrum Access Networks. Piscataway: IEEE Press, 2017:1-3.
- [12] PENG S L, JIANG H Y, WANG H X, et al. Modulation recognition using hierarchical deep neural networks [J]. IEEE Transactions on Neural Networks and Learning Systems, 2019, 30(3):718-727.
- [13] 查雄,彭华,秦鑫,等. 基于多端卷积神经网络的调制识别方法[J]. 通信学报, 2019, 40(11):30-37.  
ZHA X, PENG H, QIN X, et al. Modulation recognition method based on multi-inputs convolution neural network [J]. Journal on Communications, 2019, 40(11):30-37 (in Chinese).
- [14] ZHANG M, DIAO M, GUO L M. Convolutional neural networks for automatic cognitive radio waveform recognition [J]. IEEE Access, 2017, 5:11074-11082.
- [15] O'SHEA T J, CORGAN J, CLANCY T C. Convolutional radio modulation recognition networks [C] // International Conference on Engineering Applications of Neural Networks, 2016:213-226.
- [16] WEST N E, O'SHEA T. Deep architectures for modulation recognition [C] // 2017 IEEE International Symposium on Dynamic Spectrum Access Networks. Piscataway: IEEE Press, 2017:1-6.
- [17] O'SHEA T J, ROY T, CLANCY T C. Over-the-air deep learning based radio signal classification [J]. IEEE Journal of Selected Topics in Signal Processing, 2018, 12(1):168-179.
- [18] 查雄,彭华,秦鑫,等. 基于循环神经网络的卫星幅相信号调制识别与解调算法[J]. 电子学报, 2019, 47(11):2443-2448.  
ZHA X, PENG H, QIN X, et al. Satellite amplitude-phase signals modulation identification and demodulation algorithm based on the cyclic neural network [J]. Acta Electronica Sinica, 2019, 47(11):2443-2448 (in Chinese).
- [19] YAO T Y, CHAI Y, WANG S, et al. Radio signal automatic modulation classification based on deep learning and expert features [C] // 2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference. Piscataway: IEEE Press, 2020:1225-1230.
- [20] WU Y L, LI X J, FANG J. A deep learning approach for modulation recognition via exploiting temporal correlations [C] // 2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications. Piscataway: IEEE Press, 2018:1-5.
- [21] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [J]. Communications of the ACM, 2017, 60(6):84-90.
- [22] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [EB/OL]. (2015-04-10) [2020-12-01]. <https://arxiv.org/abs/1409.1556>.
- [23] IOFFE S, SZEGEDY C. Batch normalization: Accelerating deep network training by reducing internal covariate shift [C] // International Conference on Machine Learning. New York: ACM, 2015:448-456.
- [24] SAINATH T N, VINYALS O, SENIOR A, et al. Convolutional, long short-term memory, fully connected deep neural networks [C] // 2015 IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway: IEEE Press, 2015: 4580-4584.

# Convolutional neural network based algorithm for automatic modulation recognition of satellite signals

CUI Tianshu<sup>1,2</sup>, CUI Kai<sup>1,2</sup>, HUANG Yonghui<sup>1,2,\*</sup>, ZHAO Wenjie<sup>1</sup>, AN Junshe<sup>1,2</sup>

(1. Key Laboratory of Electronics and Information Technology for Space Systems, National Space Science Center, Chinese Academy of Sciences, Beijing 100190, China;

2. School of Computer Science and Technology, University of Chinese Academy of Sciences, Beijing 100049, China)

**Abstract:** Automatic modulation recognition is a key technology for spatial cognitive communication system, which helps to realize adaptive signal demodulation. Although the deep neural network has the advantage of strong feature extraction, it suffers from the problems of numerous parameters and large amount of calculation, and thus is difficult to be implemented in in-orbit applications. To mitigate these problems, we propose a lightweight, high-performance convolutional neural network structure. The network first extracts the in-phase and quadrature features of the signal, then the time domain features, and finally the mean value of each channel feature for classification. The experimental results of the classification of 11 modulation methods show that when the signal-to-noise ratio is higher than 0 dB, the average recognition accuracy can reach 86.94%, which is 31.54% higher than that of traditional cumulant methods. Compared with the current deep neural network model with high recognition accuracy, the network proposed uses only less than 10% of model parameters, and increases the calculation speed by an average of 20 times on Raspberry Pi 4B.

**Keywords:** cognitive communication system; modulation recognition; in-phase and quadrature correlation feature; convolutional neural network; deep learning

**Received:** 2020-12-24; **Accepted:** 2021-03-12; **Published online:** 2021-04-27 10:47

**URL:** [kns.cnki.net/kcms/detail/11.2625.V.20210426.1620.001.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210426.1620.001.html)

**Foundation item:** Laboratory Fund of Key Laboratory of Electronics and Information Technology for Space Systems, CAS (Y42613A32S)

\* **Corresponding author.** E-mail: yonghui@nssc.ac.cn

http://bhxb.buaa.edu.cn      jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2020.0699

# 产品协同设计中异构模型数据融合的有效策略

薛俊杰<sup>1,2</sup>, 周军华<sup>1,2,3</sup>, 施国强<sup>1,2,3</sup>, 宋晓<sup>4</sup>, 蒋炎红<sup>5</sup>, 全红艳<sup>6,\*</sup>

(1. 北京仿真中心 北京市复杂产品先进制造系统工程技术研究中心, 北京 100854;

2. 北京电子工程总体研究所 复杂产品智能制造系统技术国家重点实验室, 北京 100854;

3. 北京仿真中心 航天系统仿真重点实验室, 北京 100854; 4. 北京航空航天大学 网络空间安全学院, 北京 100083;

5. 华东师范大学 软件工程学院, 上海 200062; 6. 华东师范大学 计算机科学与技术学院, 上海 200062)

**摘 要:** 针对复杂产品设计中, 不同设计工具产生的模型数据之间的融合问题, 提出了一种工具间端到端的异构模型数据融合策略。利用数据库管理动态特性, 通过模型信息共享, 实现异构模型数据之间的融合。在 OpenMBEE 系统集成环境中, 通过建模工具 CREO 二次开发, 利用所提策略获取全生命周期设计中的动态模型属性信息, 通过 3D 模型编辑及重用功能测试, 验证了所提策略的有效性。利用自动获取可视化模型属性信息的智能算法, 设计一种基于 Transformer 模型与双向长短期记忆 (Bi-LSTM) 模型相结合的模型属性智能提取算法, 利用神经网络的多层感知特性, 通过对模型中属性文本信息进行深度学习、特征分析, 实现了对异构数据属性信息的自动提取功能。利用 CAMEO 建模工具设计的需求分析模型构建模型数据集, 验证了智能模型信息自动提取功能的有效性。

**关键词:** 异构数据; OpenMBEE; 数据交换; 多层感知; 深度学习

**中图分类号:** V221<sup>+</sup>.3; TB553

**文献标志码:** A

**文章编号:** 1001-5965(2022)06-0995-09

在复杂产品的设计中, 近年来出现了各种先进数字化设计技术和手段, 在这些新技术发展中, 数字孪生概念和技术越发引起人们的关注, 也成为智能设计制造中的一种手段和方法。数字孪生概念自 Grieves<sup>[1]</sup> 于 2005 年提出以后, 在复杂产品设计中充分发挥着作用, 其可以很大程度减少产品设计周期, 降低产品设计的复杂度。

在数字孪生技术的发展及推动下, 复杂产品设计呈现以下问题: 设计结构集成度高、各个领域的工具建模难于耦合、设计信息出现庞杂等。另外, 其也导致了设计中需求难以得到快速响应、系统层设计验证不足、设计过程难以获得一致性语义模型等现象。为了有效解决这些问题, 科研人员提出了基于模型的系统工程 (model-based sys-

tems engineering, MBSE)<sup>[2]</sup>。MBSE 以模型为设计中心, 在复杂产品设计中, 通过模型设计的不断演化、迭代, 实现来自不同工具模型的集成与协同<sup>[3]</sup>。在模型工具集成中, 如何处理来自于不同工具的异构模型数据之间的协同, 以及如何充分利用异构数据源的信息和不同阶段的模型, 使得各个工具在系统设计、分析、验证等过程中充分发挥其作用, 这些都是在实际应用中亟须解决的问题。

在复杂产品的集成设计中, NASA 基于 MBSE 理念, 研发了面向复杂产品的开源设计平台 OpenMBEE<sup>[4]</sup>, 在复杂产品设计中需要采用集成的设计架构, 包括需求分析模型的设计工具、物理仿真模型的设计工具、功能逻辑模型和数学分析

收稿日期: 2020-12-18; 录用日期: 2021-04-09; 网络出版时间: 2021-04-19 16:19

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20210419.1351.001.html

\* 通信作者。E-mail: hyquan@cs.ecnu.edu.cn

**引用格式:** 薛俊杰, 周军华, 施国强, 等. 产品协同设计中异构模型数据融合的有效策略[J]. 北京航空航天大学学报, 2022, 48(6): 995-1003. XUE J J, ZHOU J H, SHI G Q, et al. Effective strategy of heterogeneous model data fusion in product collaborative design[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(6): 995-1003 (in Chinese).

模型的设计工具,还包括产品测试过程的可视化仿真模型的设计工具,提供团队之间多模型协同设计和验证服务的环境。然而,在实际应用中却产生了各个工具之间的协同和数据共享的问题。如何有效处理多工具的异构模型数据之间的信息协同就是实际应用中亟待解决的问题。

由于复杂产品设计需要多工具协同,如需求分析的设计工具、物理仿真模型的设计工具、功能逻辑模型和数学分析模型的设计工具及分析和测试的可视化仿真的设计工具,探索了一种多工具集成的异构数据协同策略,采用 OpenMBEE 的集成环境框架,设计不同建模工具进行集成的方案。利用 CAMEO 工具作为产品需求分析的设计工具,采用面向对象的建模语言 SysML 描述模型结构和行为<sup>[5]</sup>,通过集成 CAMEO 建模工具、仿真模型工具 MATLAB 及 3D 可视化仿真工具 CREO,实现了系统工具的集成。在此基础上,设计不同工具之间的异构数据协同策略,在多层协同策略指导下,获取全生命周期设计中的动态模型属性信息,通过不同工具之间模型数据的通信,验证了所设计的数据协同机制的有效性。

研究发现,在复杂产品设计中,根据异构模型数据之间的协同需要,有必要从产品的需求模型中自动获取模型的基本信息(如模型的规格属性信息等),其是实现异构数据共享和复杂产品协同设计的基础。然而,由于不同工具的模型之间存在着数据格式的差异,在研究中借助于 XML 格式文件作为模型数据交换媒介,将需求设计模型存到 XML 后,再实现智能模型信息的自动提取功能。也就是说,不同工具的模型数据共享问题,可以通过对 XML 的文本分析来进一步解决。

实际上,在近年来的科学探索中,对于 XML 文本信息提取的研究,已经开展了一些工作。例如,Web 数据信息的提取<sup>[6]</sup>、从 XML 文档提取有序的子树信息<sup>[7]</sup>、XML 文档特征及信息的抽取<sup>[8]</sup>、基于传统机器学习策略的 XML 信息挖掘<sup>[9]</sup>、从 XML 进行兴趣信息的提取<sup>[10]</sup>。随着人工智能技术的迅猛发展,近年来出现了基于深度学习的 XML 文档信息挖掘与提取的研究<sup>[11-12]</sup>,然而利用现有方法尚不能满足 XML 模型信息的自动提取。本文研究了一种基于深度学习的模型属性信息自动提取的关键技术算法,采用 Transformer 神经网络进行设计,并结合双向长短期记忆(bi-directional LSTM, Bi-LSTM)模型策略进行设计,得到了较满意的结果。

## 1 异构数据的多层协同策略

本文研究了一种异构数据的多层协同策略,在 OpenMBEE 集成环境中,通过建模工具 CREO 进行二次开发,设计多层协同策略,获取全生命周期设计中的动态模型属性信息,通过 3D 模型编辑及重用的功能测试,验证策略的有效性。

### 1.1 协同设计组件及工具的集成方案

在异构数据的多层协同策略中,以 OpenMBEE 协同设计平台作为框架,集成多种组件及模型设计工具,如图 1 所示。

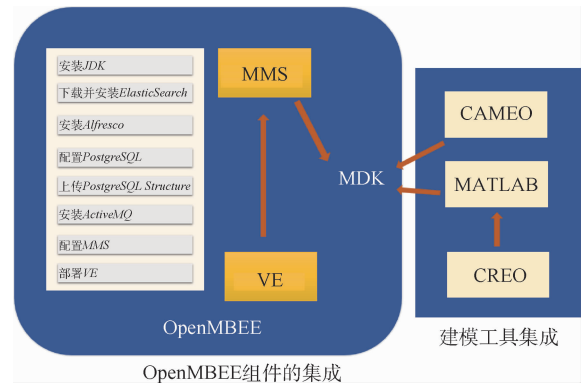


图 1 集成的组件及工具

Fig. 1 Integrated components and tools

集成设计环境中,组件及工具的集成步骤描述如下:

**步骤 1** OpenMBEE 平台环境的搭建。包括模型管理系统(model management system, MMS)及视图编辑器(view editor, VE)的部署。OpenMBEE 平台提供 3 个核心模块,也被称为 3 部组件:MMS、VE 和模型开发工具包(model development kit, MDK)。

在平台环境的搭建中,包括安装 JDK、下载并安装 ElasticSearch 搜索引擎等过程,如图 1 所示。MMS 是整个 OpenMBEE 平台的唯一数据来源,保证了数据的一致性和追溯性,用于存储来自 CAMEO 的 SysML 模型数据。为此,可以获得 SysML 描述的项目模型元素,如类、实例、属性、关系等<sup>[13]</sup>。在 MMS 的集成中,需要安装 ElasticSearch、Alfresco、Web 服务器 Apache Tomcat、对象关系数据库 PostgreSQL 和 Java 消息服务 ActiveMQ 代理。VE 模块的软件包部署在 Tomcat 服务器中,其利用浏览器作为可视化的界面工具,只有在 MMS 启动 Tomcat 服务和 PostgreSQL 以后,VE 才能得到正常访问。

**步骤 2** 建模工具的集成。为了满足复杂产品设计的需要,不仅要集成基本组件模块,也要集

成多种工具,包括需求分析设计工具 CAMEO、数学分析工具 MATLAB 及 3D 可视化仿真工具 CREO。在 CAMEO 工具集成中,利用 OpenMBEE 提供的 CAMEO MDK,通过 MDK 集成可以实现与 MMS 同步,并且验证构建系统组件的性能;在 MATLAB 集成时,关键是 OpenMBEE 提供的 MATLAB MDK 插件安装与设置,在策略探索中,借助 MATLAB 作为中间工具,实现 CREO 工具的集成,以获取模型动态数据。

## 1.2 多层协同策略

在集成的设计架构中,各个工具以协同合作模式,为用户提供多模型的设计和验证服务,目的是支持团队在模型设计中的协同。为了实现异构数据的通信,借助各个模块及工具,通过信息的交互实现异构数据的获取和更新。在设计时,需要设计各个工具及组件之间的认证接口,同时也要以协同数据的形式进行通信。充分利用 OpenMBEE 提供的 MDK 进行设计,实现不同工具之间的协调,以便集成系统可以实现对需求分析的模型设计、物理仿真模型的设计、功能逻辑模型和数学分析模型的设计及产品测试模型的设计。

为了实现各个模型工具的集成,先集成 OpenMBEE 的核心模块组件,即在 MMS 和 VE 模块实现集成后,能够正常工作的基础上,再对建模工具 CAMEO、MATLAB、CREO 进行集成,如图 2 所示。本文中设计了各个工具及组件之间的多层协同机制,建立异构数据协同建模的环境。

### 1) 建模工具与数据库间协同。

在进行 CAMEO 与 VE 之间的协同设计时,需要对模型数据做参数化处理过程。采用中间组件 MMS 解决用户模型数据存储问题,将 CAMEO 中建立模型的数据信息保存到 MMS 中,并利用 VE 组件进行可视化显示。图 3 展示了建模工具与数据库之间的协同。用户在 CAMEO 中进行需求和功能逻辑建模时,利用 MDK,将异构的模型数据分解成树状结构<sup>[14]</sup>,并存储在基于 PostgreSQL 的 MMS 组件数据库中,该过程称为模型数据的参数化。在 CAMEO 建模中建立参数化数据,为 MATLAB 对模型分析验证准备必要的模型数据,并进一步提供给 CREO 进行 3D 建模处理。

### 2) 分析工具与数据库间协同。

为了在异构数据建模过程中能够可视化显示构建的模型,需要建立 MATLAB 与 VE 之间的协同通信。设计时,借助于 OpenMBEE 的 MDK 插件,在 OpenMBEE 环境中部署 MATLAB,再通过 MATLAB MDK 中定义 MMS 类对象方式,实现

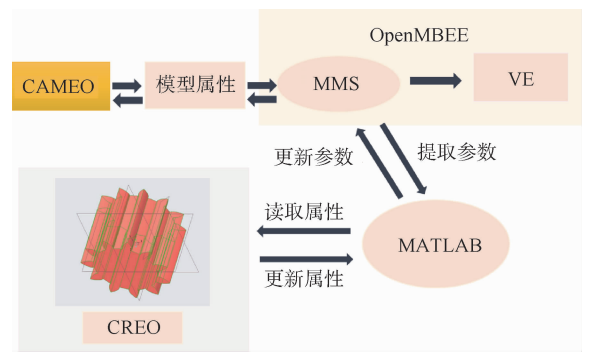


图 2 多工具集成及协同关系

Fig. 2 Multi tool integration and collaborative relationship

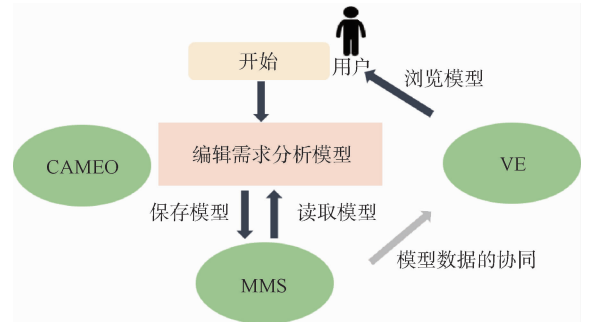


图 3 建模工具与数据库间协同

Fig. 3 Collaboration between modeling tools and databases

MATLAB 对 MMS 的登录、访问数据、更新数据的目的。这样便可以实现 MATLAB 对 MMS 数据中模型数据的提取。

### 3) 分析工具与可视化工具间协同。

在异构的数据源模型设计中,采用 CREO 作为产品设计中 3D 可视化的建模工具。在构建动态模型过程中,采用 CREO 的二次开发技术,在 CREO 中进行接口设计,可以读取 MATLAB 分析模型的数据。模型分析人员读取 MMS 中模型的基本信息,利用 VE 对模型参数进行编辑和修改,进一步保存到 MMS 数据库中。这时由于 MMS 和 VE 处于协同状态,3D 可视化设计人员可以通过与 MATLAB 之间的通信<sup>[15]</sup>获得模型信息,从而实现 3D 可视化功能,如图 4 所示。

### 4) 动态双向数据交换。

为了实现异构数据模型的构建,在集成各种建模工具的基础上,在多层协同机制控制下<sup>[16]</sup>,实现产品设计中动态双向数据交换。以 MMS 组件为数据存储中心,其他工具在工作时需要相互协同,即不仅需要相互接口,还需要对传输的数据进行处理转换和融合。也即需要将不同模块处理的数据转化为统一的标准形式,并存储在 MMS 中。

产品设计中,模型的建模通过以下方式实现了双向的数据交换功能。

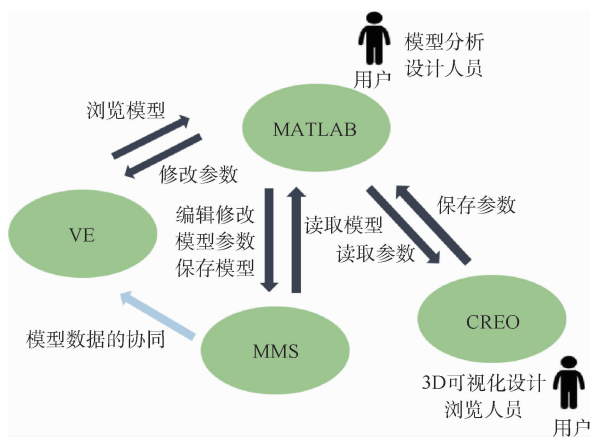


图4 分析工具与可视化工具间协同

Fig. 4 Collaboration between analysis tools and visualization tools

## ① 读取模型信息的数据流向。

a. 建模用户利用 CAMEO 建立基于 SysML 的需求分析模型,通过异构模型分解算法,模型将会被分解为树形结构,存储在 MMS 中,每个节点都自动生成专属标识 ID。VE 可通过接口获取 MMS 数据,显示到浏览器中。

b. 设计人员根据 MMS 中模型信息,利用 VE 可视化界面或者 MATLAB 工具,对模型进行实例化<sup>[16]</sup>,通过节点的专属 ID 给定 3D 模型属性的数值,并保存到 MMS 数据库中,数据同步显示到 VE 界面。

c. 3D 建模用户,利用 CREO 通过二次开发的接口,连接到 MATLAB。将要获取参数的节点 ID 传给 MATLAB, MATLAB 从 MMS 获取到模型的参数信息,再反馈给 CREO, CREO 根据 MATLAB 反馈得到的模型数据,重新渲染更新 3D 模型。

## ② 建模数据保存功能的数据流向。

首先,在 CREO 中对 3D 模型进行构建,编辑模型参数。其次,保存 3D 模型,向 MATLAB 发出存在的请求。然后,在 MATLAB 与 MMS 连接的情况下, MATLAB 向 MMS 发出保存模型信息的请求。最后,保存到 MMS 数据库中,并同步显示到 VE 界面中。

在双向的数据交换过程中,不同工具的模型需要相互配合。在异构数据源的设计中,研究算法时要考虑到模型数据的动态性协同特性。采用 OpenMBEE 的组件及工具进行实现,由于利用 MMS 对异构数据进行融合, MATLAB 可以直接获得 CAMEO 中创建的模型属性信息,并且当在 MATLAB 中创建仿真模型时,需要对模型进行参数化处理。同时,由于各个组件及工具之间存在动态链接机制<sup>[17]</sup>,可以形成多终端异构数据的相

关联动,当 MMS 中的模型参数发生变化时,可以对 MATLAB 运行仿真参数进行更新,进一步利用 CREO 工具实现模型数据可视化信息的动态更新,保持异构数据的一致性。

## 2 异构模型数据的智能提取算法

在产品设计中,为了实现模型异构数据信息的协同,利用 XML 数据交换的信息,设计了一种智能模型数据信息提取的算法。近年来,随着人工智能技术的发展,智能文本信息提取的关键技术研究取得了新的进展,特别是随着神经网络技术的飞跃,高效文本提取算法不断出现。Google 公司于 2017 年提出了一种基于神经网络的 Transformer 框架<sup>[18]</sup>,其可以充分学习得到文本的特征,并且能够更好地捕获文本之间上下文的依赖特征,在文本分析中起着重要作用<sup>[19-22]</sup>。Alaparthi 和 Mishra<sup>[22]</sup>提出一种基于 Transformer 的有效文本分析方法,采用 Transformer 作为编码器的主体模型结构,取得了较满意的文本分析结果,但是利用该方法对 XML 结构化模型数据文档进行分析,并提取模型属性的文本信息时,并不能得到满意的抽取结果。因此,基于该工作的启发,本文设计了一种基于 Transformer 模型与 Bi-LSTM 模型相结合的模型属性智能提取算法,实现模型异构数据的自动提取功能。

首先,由于 XML 是数据交换的通用格式文件,可以将需求模型工具对产品建模结果先存到 XML 文件中;然后,通过数据预处理清洗得到文本句子,利用多层感知的深度学习手段对 XML 文件的模型数据进行解析,实现智能模型信息自动提取功能。

## 2.1 数据的预处理

1) 清洗非标识符的 ASCII 字符。根据正则表达式规则,对 XML 文件中非标识符的 ASCII 字符进行过滤,如单引号、双引号、冒号、等号、尖括号、斜杠等,获得纯文本数据。

2) 数据平衡的预处理。研究中发现,由于 XML 文档中模型属性词汇所占的比例过小,使得在深度学习过程中,模型属性特征不明显,从而不能得到准确的模型属性预测结果。为了解决这一问题,在数据处理的研究过程中,将训练集中含有模型属性的单词数量在原来文本位置中复制  $N$  次,增强模型属性的特征,从而达到准确预测的目的。本文经过实验证明了  $N$  为 30 时,可以取得较满意的结果。

在多层感知的深度学习网络结构设计中,将

整体网络设计为基于 Transformer 的特征提取子网络 (Transformer based feature extraction sub-network, TF-Net) 和特征精细化子网络 (feature refinement subnetwork, FR-Net) 两部分。算法设计的步骤如下:

**步骤 1** 数据预处理。构建数据集,利用上述预处理步骤进行数据预处理。

**步骤 2** 神经网络构建。构建网络的拓扑结构。

**步骤 3** 依次进行预训练和特征优化两步训练。

**步骤 4** 模型的测试及分析。

### 2.2 神经网络结构设计

在 TF-Net 的设计中,采用现有研究的网络拓扑结构<sup>[22]</sup>,网络的基本结构如图 5 所示。利用定位嵌入层、分割嵌入层和编码嵌入层 3 层结构对文本特征进行分析和提取,得到初始化的学习特征。对于 TF-Net 输入是预处理后得到的 XML 的文本句子,输出结果是初始化特征结果,每个词的特征形式为 1 维向量,其具有 768 个分量。

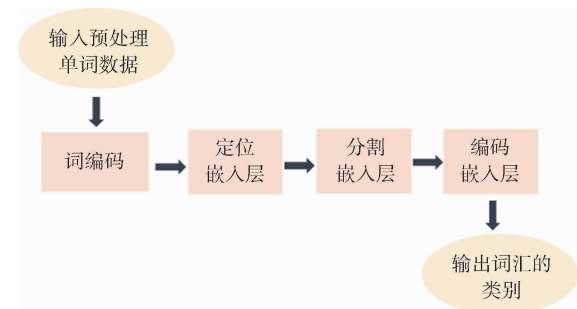


图 5 TF-Net 拓扑结构  
Fig. 5 Topology of TF-Net

在词编码中<sup>[23]</sup>,采用词编码的嵌入变换对输入词汇进行编码,主要包括嵌入变换功能和层特征的标准化处理;编码器的作用主要是特征学习和提取,由 6 个相同结构的隐藏单元相继连接构成。每个隐藏单元的功能为:通过神经单元的线性变换、神经元节点的激活、多头自注意力特征增强处理和跨层连接处理。

算法的网络拓扑结构如图 6 所示。对于预处理数据,经过 TF-Net 学习得到初始化特征  $T_1, T_2, \dots, T_N$  后 ( $C$  为句子开头的标记),再进一步利用整体网络结构 (包括 TF-Net 和 FR-Net) 进行特征精细优化。

FR-Net 设计主要包括 Bi-LSTM 单元与条件随机场 (conditional random field, CRF) 单元。在 Bi-LSTM 单元设计中<sup>[23]</sup> (见图 7,图中  $w$  表示权重),利用 Bi-LSTM 模型的文本序列前向传播学

习和后向传播学习机制,得到词汇序列的上下文特征,有利于 XML 的模型数据信息的准确提取。

在 CRF 单元的设计中,考虑到经 Bi-LSTM 单元处理后输出的句子特征表示词的类别特征,为了充分利用词和词之间的相互影响,采用 CRF 机制,考虑邻近词之间的依赖关系因素设计了 CRF 模块<sup>[24]</sup>。

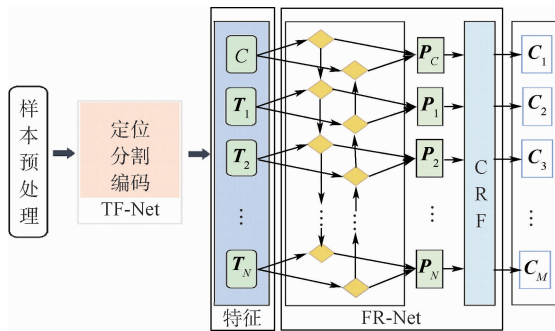


图 6 网络拓扑结构  
Fig. 6 Topology of network

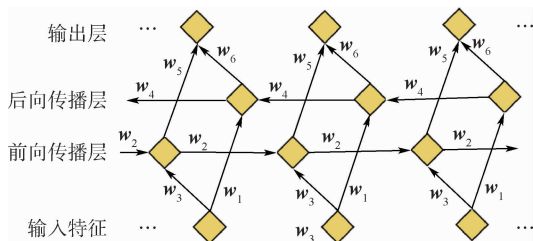


图 7 Bi-LSTM 单元结构  
Fig. 7 Structure of Bi-LSTM unit

将模型数据提取的过程看作 XML 文档中词的分类问题,并且目的是得到每个词的类别标签。每个词的标签设计为:需要提取的模型属性 (单词标签) 和不需要提取的其他单词标签,分别记为 1 和 0 的标记,对于输入的句子,输出结果记为一个标记序列 (即 Token 序列)。为了增强神经网络学习特征的性能,利用 CRF 考虑邻近标签的相互关系。在相关研究中,李丽双和郭元凯<sup>[25]</sup>提出一种基于 CRF 的文本分析策略,本文利用文献<sup>[25]</sup>的 CRF 模型进行设计,考虑随机产生的 Token 序列,选择整体概率最大的 Token 序列作为最优的句子标注结果。

### 2.3 神经网络结构的训练和模型属性的预测

为了提取模型参数信息,依次进行预训练和特征优化。在预训练中,对于 TF-Net 特征,根据 XML 模型文件的格式特点,通过建立数据集及标注进行约束,利用文本处理模型中的定位、分割和编码功能,获得文本分析编码特征。进一步对网络特征优化学习。利用学习得到的网络特征,实现 XML 数据模型属性的预测。在网络训练中,均

采用均方差损失对网络性能进行优化。算法的学习和预测过程如图 8 所示。

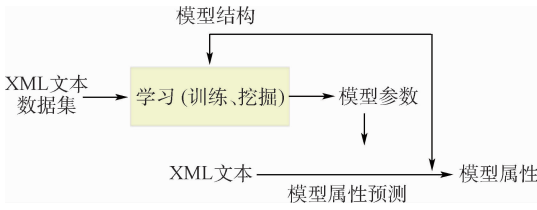


图 8 网络的学习和预测过程框架

Fig.8 Framework of network learning and prediction

### 3 实验与分析

#### 3.1 异构数据的多层协同策略的验证

为了验证异构数据模型构建策略的合理性,进行实验测试和分析。为了验证整体框架流程的合理性,构建协同环境。建立了基于 centos7 系统的实验环境,使用 CAMEO19.0 SP, MATLAB 2018b, MMS3.4.2, VE4.2.2, Visual studio 2019 进

行研究,做 2 组实验:一组是多层工具协同的验证,另一组是产品 3D 建模中动态数据交换功能的验证。

#### 1) 多层工具协同的验证。

实验中通过更改参数模型来测试多层工具协同的可行性。

首先,在 CAMEO 软件中建立模型,如图 9 所示。SysML 参数模型中包含模型属性名称,通过 CAMEO MDK 提交数据到数据存储中心。由于复杂产品设计中需求模型是多样的,并且参数模型的数据是庞大的,需要通过 SysML 描述语言对参数模型进行约束和展示。当用户登陆 VE 可视化编辑器后,可以浏览模型属性名,并且编辑修改模型的参数。同时, MATLAB 通过 MDK 登陆成功后,即成功连接到 MMS,并从 VE 可视化页面编辑模型参数,写入 MMS 中。在 MATLAB 中可以修改模型参数的数值,VE 可视化编辑页面也会同步更新。

Sine Wave	Amplitude	2
	Bias	3
	Frequency	4

图 9 多层工具协同的验证结果

Fig.9 Validation results of multi-layer tool collaboration

#### 2) 动态数据交换功能的验证。

实验建立在多层工具协同实现的基础上,通过 CREO 与其他多层软件工具的交互,验证动态数据交换功能。如图 2 所示,查看 CREO 中现有的模型参数,在建模软件中建立相应的参数模型。从多层工具协同实验可知,参数可以被用户在 VE 可视化编辑页面查看并编辑。此时,通过 CREO 的二次开发功能,进行参数的编辑与修改。VB. Net中直接调用 MATLAB MDK 中的函数,将

MATLAB 中的数据提取到 VB. Net 中,通过 CREO 提供的 VB 接口,实现从 MMS 中获取建模数据。在参数管理器中修改 3D 模型参数,修改后参数会更新到 MMS 数据存储中心。CREO 重新生成 3D 可视化模型,模型将会自动发生变化,查看当前模型参数,可见模型参数实现了动态更新的结果。通过以上实验结果,可以得出异构数据模型构建策略的合理性。将模型中异构数据重新参与 3D 可视化操作的构建,方便了在数字孪生的全生

命周期中,多端用户对于模型的操作。

本文的异构数据多层协同方案,已经在复杂产品设计的实际应用项目中得到了验证,证明了该策略的有效性。

### 3.2 异构模型数据的智能提取算法验证

在异构模型数据的智能提取算法验证中,采用 GPU GeForce GTX 1080 Ti 进行加速,在 CPU I7-9700K 环境进行实验,内存为 32 GB,硬件显存为 8 GB,采用 Tensorflow 1.4 的深度学习框架,用 Python3.7 编程语言编写实验。

#### 1) 数据集构建及训练方法。

为了构建数据集,使用 CAMEO 建立 50 个需求分析模型,存为 XML 文件形式,利用数据增强手段增强 XML 模型文件,共得到 500 个 XML 文档,将每个文档利用 Python 语言预处理提取 500 个文档中的句子,按照句子数量比例 3:1:1 划分为训练数据集、评估数据集与测试数据集,分别用于网络训练、计算网络性能的评价指标、对网络性能测试。实验过程中,网络的实现细节与使用的重要超参数如下:Dropout 率取 0.5,每个句子的最大长度为 512 Byte,如果不足,则通过补 0 方式补齐。输入模型的句子长度保持一致,每批次取样本的数量为 8。共进行 3 个周期的训练,每个周期进行 50 000 次训练。

#### 2) 算法功能测试及性能分析。

在算法验证中,一方面测试了本文所设计智能模型信息提取的功能,另一方面也对算法设计方案的性能进行分析。采用查准率、查全率和  $F_1$  分数作为性能指标对算法性能进行测试和分析。测试算法性能时,同时进行消融实验,分别对网络模型存在 Bi-LSTM 单元情况下的网络(记作 A 网络)和去掉 Bi-LSTM 单元情况下的网络(记为 B 网络)进行查准率、查全率和  $F_1$  分数的测试,得到的测试结果如图 10 所示。可以明显看出,利用 Bi-LSTM 机制,可以有效提高模型属性数据提取的准确性,利用文本之间的上下文特征,有利于提高文本分析的质量,改进文本定位的准确性。

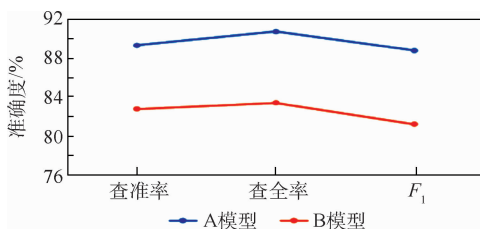


图 10 消融实验的性能对比

Fig. 10 Performance comparison of ablation experiments

测试中对 A 模型和 B 模型的时间性能进行了比较。统计 XML 文件中模型属性进行预测的平均时间,A 模型的预测时间为 0.099 91 s,而 B 模型的预测时间为 0.083 3 s,可以看出,由于 A 模型比 B 模型在拓扑结构上增加了 Bi-LSTM 单元,预测时间略有增加,但是却可以得到较高的信息提取准确率。

## 4 结 论

本文根据复杂产品设计中多工具集成的实际需要,探索了复杂产品设计中,不同设计工具产生的模型数据之间的融合问题。

1) 提出一种异构数据的多层协同策略,可以在协同设计环境中实现模型编辑及重用。

2) 研究一种异构模型数据的智能提取算法,具有较好的提取性能,在查全率达到 90% 以上的情况下,预测时间仅需要 0.099 91 s。

本文研究为多工具协作设计中异构数据模型之间协同问题提供了有效的解决方案。在进一步的研究中,可以考虑将智能模型信息提取算法结合到仿真集成平台,在复杂产品设计中实现端到端的协同控制。

### 参考文献 (References)

- [1] GRIEVES M W. Product life cycle management: The new paradigm for enterprises[J]. International Journal of Product Development, 2005, 2(1-2): 71-84.
- [2] BOSCHERT S, ROSEN R. Digital Twin—The simulation aspect [M]. Berlin: Springer, 2016: 59-74.
- [3] MOUSAVI B A, AZZOUC R, HEAVEY C, et al. A survey of model-based system engineering methods to analyse complex supply chains: A case study in semiconductor supply chain[J]. IFAC-PaperOnLine, 2019, 52(13): 1254-1259.
- [4] KRUSE B, BLACKBURN M. Collaborating with OpenMBEE as an authoritative source of truth environment[J]. Procedia Computer Science, 2019, 153: 277-284.
- [5] WEILKIENS T. SysML—The systems modeling language [M] // WEILKIENS T. Systems engineering with SysML/UML. Berlin: Springer, 2007: 223-270.
- [6] 谢慧敏. 基于 XML 的数据转换和发布的实现 [D]. 南京: 南京理工大学, 2007.
- [7] XIE H M. Implementation of data transformation and publishing based on XML [D]. Nanjing: Nanjing University of Technology, 2007 (in Chinese).
- [8] TAN H, HADZIC F, DILLON T S, et al. Tree model guided candidate generation for mining frequent subtrees from XML documents[J]. ACM Transactions on Knowledge Discovery from Data, 2008, 2(2): 1-43.
- [9] 郭丽红, 王箭. 基于 PCA 的 XML 文档特征提取方法 [J]. 计算机工程与设计, 2011, 32(11): 3894-3896.

- GUO L H, WANG J. Feature extraction method of XML document based on PCA [J]. Computer Engineering and Design, 2011, 32(11): 3894-3896 (in Chinese).
- [9] 潘有能. XML挖掘: 聚类, 分类与信息提取[M]. 杭州: 浙江大学出版社, 2012.
- PAN Y N. XML mining: Clustering, classification and information extraction[M]. Hangzhou: Zhejiang University Press, 2012 (in Chinese).
- [10] 邱实, 袁晓艳, 裴非, 等. 基于配置文件对XML数据进行字段提取及结构化方法: CN109885569A [P]. 2019-06-14.
- QIU S, YUAN X Y, PEI F, et al. Field extraction and structure method for XML data based on configuration file: CN109885569A [P]. 2019-06-14 (in Chinese).
- [11] SONG E, HAW S C. XML-REG: Transforming XML into relational using hybrid-based mapping approach [J]. IEEE Access, 2020, 8: 177623-177639.
- [12] BRAHMIA Z, HAMROUNI H, BOUAZIZ R. XML data manipulation in conventional and temporal XML databases: A survey [J]. Computer Science Review, 2020, 36: 100231.
- [13] MADNI A M, SIEVERS M. Model-based systems engineering: Motivation, current status, and needed advances [M]. Berlin: Springer, 2018: 311-325.
- [14] BONE M, BLACKBURN M, KRUSE B, et al. Toward an interoperability and integration framework to enable digital thread [J]. Systems, 2018, 6(4): 46.
- [15] BAYER T J, BENNETT M, DELP C L, et al. Update-concept of operations for integrated model-centric engineering at JPL [C] // 2011 Aerospace Conference. Piscataway: IEEE Press, 2011: 1-15.
- [16] CELLURA M, GUARINO F, LONGO S, et al. Modeling the energy and environmental life cycle of buildings: A co-simulation approach [J]. Renewable and Sustainable Energy Reviews, 2017, 80: 733-742.
- [17] GRAIGNIC P, VOSGIEN T, JANKOVIC M, et al. Complex system simulation: Proposition of a MBSE framework for design-analysis integration [J]. Procedia Computer Science, 2013, 16: 59-68.
- [18] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C] // Advances in Neural Information Processing Systems, 2017: 5998-6008.
- [19] SAHU S K, ANAND A. Drug-drug interaction extraction from biomedical texts using long short-term memory network [J]. Journal of Biomedical Informatics, 2018, 86: 15-24.
- [20] LI F, JIN Y, LIU W, et al. Fine-tuning bidirectional encoder representations from transformers (BERT)-based models on large-scale electronic health record notes: An empirical study [J]. JMIR Medical Informatics, 2019, 7(3): 14830.
- [21] GUAN W, SMETANNIKOV I, TIANXING M. Survey on automatic text summarization and transformer models applicability [C] // 2020 International Conference on Control, Robotics and Intelligent System, 2020: 176-184.
- [22] ALAPARTHI S, MISHRA M. Bidirectional encoder representations from transformers (BERT): A sentiment analysis odyssey [EB/OL]. (2020-06-02) [2020-12-16]. <https://arxiv.org/abs/2007.01127>.
- [23] DEVLIN J, CHANG M W, LEE K, et al. BERT: Pre-training of deep bidirectional transformers for language understanding [C] // NAACL-HLT 2019, 2019, 1: 4171-4186.
- [24] HUANG Z, XU W, YU K. Bidirectional LSTM-CRF models for sequence tagging [EB/OL]. (2015-08-09) [2020-12-16]. <https://arxiv.org/abs/1508.01991>.
- [25] 李丽双, 郭元凯. 基于CNN-BLSTM-CRF模型的生物医学命名实体识别 [J]. 中文信息学报, 2018, 32(1): 116-122.
- LI L S, GUO Y K. Biomedical named entity recognition based on CNN-BLSTM-CRF model [J]. Chinese Journal of Information, 2018, 32(1): 116-122 (in Chinese).

## Effective strategy of heterogeneous model data fusion in product collaborative design

XUE Junjie<sup>1,2</sup>, ZHOU Junhua<sup>1,2,3</sup>, SHI Guoqiang<sup>1,2,3</sup>, SONG Xiao<sup>4</sup>, JIANG Yanhong<sup>5</sup>, QUAN Hongyan<sup>6,\*</sup>

(1. Beijing Complex Product Advanced Manufacturing Engineering Research Center, Beijing Simulation Center, Beijing 100854, China;

2. State Key Laboratory of Intelligent Manufacturing System Technology, Beijing Institute of Electronic System Engineering, Beijing 100854, China;

3. Science and Technology on Space System Simulation Laboratory, Beijing Simulation Center, Beijing 100854, China;

4. School of Cyber Science and Technology, Beihang University, Beijing 100083, China;

5. Software Engineering Institute, East China Normal University, Shanghai 200062, China;

6. School of Computer Science and Technology, East China Normal University, Shanghai 200062, China)

**Abstract:** Aiming at the problem of model data fusion among different design tools in complex product design, the research explores the fusion strategy of end-to-end heterogeneous model data between tools. A multi-layer collaborative strategy of heterogeneous data is proposed, which uses the dynamic characteristics of database management and model attribute sharing to realize the integration of heterogeneous model data. In the system integration environment of OpenMBEE, through the secondary development of the modeling tool CREO, the strategy is employed to obtain the dynamic model attribute information in the whole life cycle design. The effectiveness of the strategy is verified by 3D model editing and reuse function testing. In order to realize the fusion of heterogeneous data, an intelligent algorithm to automatically obtain the attribute information of visual model is explored, based on Transformer model and bi-directional LSTM (Bi-LSTM) model. Utilizing the multi-layer perceptual characteristics of neural network, the algorithm realizes automatic extraction of heterogeneous data attribute information through deep learning and feature analysis of the attribute information in the model. The effectiveness of the intelligent model information extraction is verified by the model data set that is established with the requirement analysis models designed by modeling tool CAMEO.

**Keywords:** heterogeneous data; OpenMBEE; data exchange; multilayer perception; deep learning

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2020.0704

# 一种基于自动分区的海量科学数据计算框架

田杨, 晏海华\*

(北京航空航天大学 计算机学院, 北京 100083)

**摘 要:** 在科学研究领域, 存储容量、处理效率和分析精度并不能适应科学数据的指数级增长速度。通过对科学数据结构与标准的研究, 提出了一个海量科学数据计算框架 BSDF。提出了一种基于模型驱动的统一数据接口, 实现对异构科学数据的无差别访问; 提出了一种基于元数据的自动分区算法, 通过参数预取与超平面维度计算确定任务颗粒度。实验结果表明: 与 H5Spark 科学数据计算框架的基于 9 项基准测试的性能相比, BSDF 计算框架提升了 39% ~ 68%; 在特定领域 PKTM 的算法优化上, BSDF 达到了 41.62 倍的加速比。

**关键词:** 科学数据; 模型驱动; 分区算法; 叠前时间偏移; 软件工程; Spark

**中图分类号:** V221<sup>+</sup>.3; TB553

**文献标志码:** A

**文章编号:** 1001-5965(2022)06-1004-09

科学研究正一步步转变为数据驱动<sup>[1]</sup>。随着科学研究设备的迅猛发展与数据采集规模的指数级增长, 存储容量、处理效率与分析精度变得极为复杂和具有挑战性。根据国际数据公司 (IDC) 的调查显示, 2025 年全球数据量将达 163 ZB, 而原始科学勘探数据将达到千万亿规模。中国科学院在 2019 年发布地球大数据共享服务平台, 平台存储包括生物生态数据、大气海洋数据、基础地理数据及地面观测数据在内总计 5 PB 数据, 同时每年将以 3 PB 的数据量进行更新。腾讯的 Peacock 孔雀模型分析系统可对 10 亿 × 1 亿级别的大规模矩阵进行分解, 可从 PB 级海量样本数据中学习 10 万量级到 100 万量级的隐含语义, 以天为单位对模型进行训练与迭代。美国 Los Alamos 国家实验室研发的混合/硬件加速宇宙论代码 HACC 可以通过仿真将上万亿粒子模拟为 20 PB 左右的数据。截止到 2017 年 NASA 地球观测与信息系统管理着约 3.9 PB 的地球科学数据, 而且预计到 2022 年将会达到 47.7 PB。传统的科学研究方法在处理海量科学数据尤其是数据密集型计算时,

复杂度极高, 性能低下。因此, 如何高效地处理大规模科学数据, 提升存储能力和处理性能, 并结合数据挖掘与机器学习等技术进行深层次分析, 逐渐成为近年来科研机构的研究热点。

在跨学科研究领域, 高维度、非线性特征及时空相关性为科学数据分析带来了巨大的复杂难度。由于传统的数据统计分析工具 (如 MATLAB、SAS 等) 难以处理大规模数据, 领域科学家通常将海量科学数据分解成更小的数据块来解决复杂的科学问题。例如, 生物学家需要高额成本从未标记的图像样本数据中发现可用规律; 而生物路径建模、显微镜成像和地震图像迁移经常需要数天或数月才能完成原始数据重建或基于机器学习的模型训练。因此, 大规模科学数据研究与先进大数据技术之间存在着明显的技术鸿沟。

在高性能计算领域, 由于在大规模数据处理和分析方面的出色性能表现, 包括 Apache Hadoop 和 Apache Spark 在内的并行计算框架引起了学术界的极大关注。Apache Hadoop 和 Apache Spark

收稿日期: 2020-12-21; 录用日期: 2021-01-29; 网络出版时间: 2021-02-18 09:51

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20210216.1605.002.html](http://kns.cnki.net/kcms/detail/11.2625.V.20210216.1605.002.html)

\* 通信作者. E-mail: yhh@buaa.edu.cn

**引用格式:** 田杨, 晏海华. 一种基于自动分区的海量科学数据计算框架[J]. 北京航空航天大学学报, 2022, 48(6): 1004-1012.  
TIAN Y, YAN H H. A computing framework for massive scientific data based on auto-partitioning algorithm[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(6): 1004-1012 (in Chinese).

等并行计算框架通过 MapReduce 计算范式与机器学习扩展程序的无缝集成,以及并行计算和内存中计算等方式快速处理海量数据。与此同时,跨学科研究领域也提出了多种科学数据格式,如分层数据格式(HDF5)<sup>[2]</sup>和网络通用数据格式(NetCDF)<sup>[3]</sup>,并已被证明能够解决大规模数据存储和平台无关互操作性等问题。但是,在将大数据并行计算框架应用于科学数据处理时,会出现一定的技术壁垒与缺陷:①科学数据存在异构性与规范多样性,缺乏一种通用的科学数据读取方法以实现数据无差别与平台无关处理;②科学数据通常是以非结构化数据形式存储,缺乏一种统一有效的方法来定义数据读取的语义索引分割符。

为了解决以上问题,本文提出了一个海量科学数据计算框架 BSDF,能够支持大规模科学数据并行化处理与分析。BSDF 提供跨学科扩展,包括 Apache Spark<sup>[4]</sup> 扩展程序、数据检索与查询、分布式环境上的机器学习模型训练和特定领域的算法优化等。在 BSDF 计算框架的构建上,首先采用模型驱动的思想从元数据中提取数据抽象模型,然后通过模型转换与裁剪过滤及数据并行处理无关的参数,通过定义模型转换规则自动生成异构科学数据的读取接口。在此基础上,提出了一种自动分区算法,能够通过元数据的参数预取进而根据超平面计算任务分区颗粒度,在内存中建立 B-Tree 结构存储分区结果,实现对多维科学数据的高效分区与索引,以提升并行 I/O 性能与多任务并行效率。BSDF 提供了一个统一的数据接口(UAI)读取 HDF5、NetCDF 和 FITS 等通用科学数据格式,并使用自动分区算法优化大规模数据处理和分析性能。通过实验评估表明,BSDF 在基于 9 项基准测试程序的大规模数据计算中,相对于目前主流的 H5Spark 科学数据计算框架,总体性能得到显著提升。

## 1 科学数据标准与特征

科学数据是指为了实现特定的研究,通过特定的方法采集的数据<sup>[5]</sup>,通常包括科学观测与采集的原始数据、实验与仿真数据、计算与模拟数据及描述数据结构特征的元数据等。2015年,国务院印发的《促进大数据发展行动纲要》中指出,科学大数据是国家大数据战略的有机组成部分。科学数据正在成为科学实验与研究的新研究范式与计算资源、科学发现的论证基础,蕴藏着巨大的科研与社会价值。

科学大数据除了拥有大数据的 4V 特征以外,还拥有多层次逐级演化的特征。通过科学探测仪器、实验装置和计算机仿真产生的海量科学原始数据,通过复杂的去噪、验证及融合过程逐步生成领域科学家和算法工程师能够直接用于分析理解的实例数据,并与相关的模型或应用融合,形成最终易于理解并可直接使用的对象数据。例如,NASA 地球观测系统(earth observing system, EOS)将卫星观测所得的实验数据,按照不断加工与演化过程,分为 0 级、1A 级、1B 级、2 级、3 级和 4 级等 6 个逐级演化的级别。0 级为原始数据,存储着直接通过科学勘测、科研实验和计算机仿真所获取的数据;演化程度越高,数据的解释性越强,随之带来的重构与加工所需的成本与时耗也越高。科学数据主要具有以下特征:①原始数据以二进制流形式存储;②数据格式多样性;③层次化结构和高维度;④海量级数据规模。

科学数据在大数据领域是与互联网数据、流媒体数据、传感数据与商业大数据同等重要的战略资源,随着计算机硬件技术的迅猛发展,处理芯片性能、内存容量不断提高,硬件加速模块(GPU、MIC 或者 FPGA/ASIC 等)及大规模集群性能飞速进步;如果能够采用先进的并行计算与分布式处理技术对大规模科学数据提供通用的快速访问与分析,准确地提取具有科学研究价值的信息,大量缩减对高复杂度和异构的科学元数据的处理与加工成本,提高特定领域生产效率与决策能力,也将会对各领域的深入研究发展产生巨大的推动作用。

## 2 大规模科学数据计算框架 BSDF

### 2.1 BSDF 整体架构

本文设计并实现了一个海量科学数据计算框架 BSDF,可以直观地处理来自第一手勘探和实验的科学数据,解决海量数据处理和分析中存在的性能瓶颈。BSDF 通过提供一个统一数据接口以支持包括 HDF4/5、NetCDF、ADIOS、SEGY、FITS、CDF 和 XML 等多种格式的科学数据格式,并且通过模型驱动(MDA)工程的引入,根据异构科学数据元数据格式与标准的建模抽象出 UML 模型,采用统一模型转换语言(QVT)并结合模版解析语言(VTL)根据建模结果自动生成数据读取代码,直接嵌入框架。BSDF 框架可以根据科学数据元数据识别数据格式并映射相应的读取代码,读取的数据通过 RDD 连接器映射为 Apache Spark 数据结构 RDD。此外,BSDF 实现了一种自动分区

算法,优化了 Apache Spark 中的原生范围分区算法,根据科学元数据的结构特点,更为高效地计算

任务执行颗粒度,并支持海量高维数据处理与分析。BSDF框架的整体架构如图1所示。

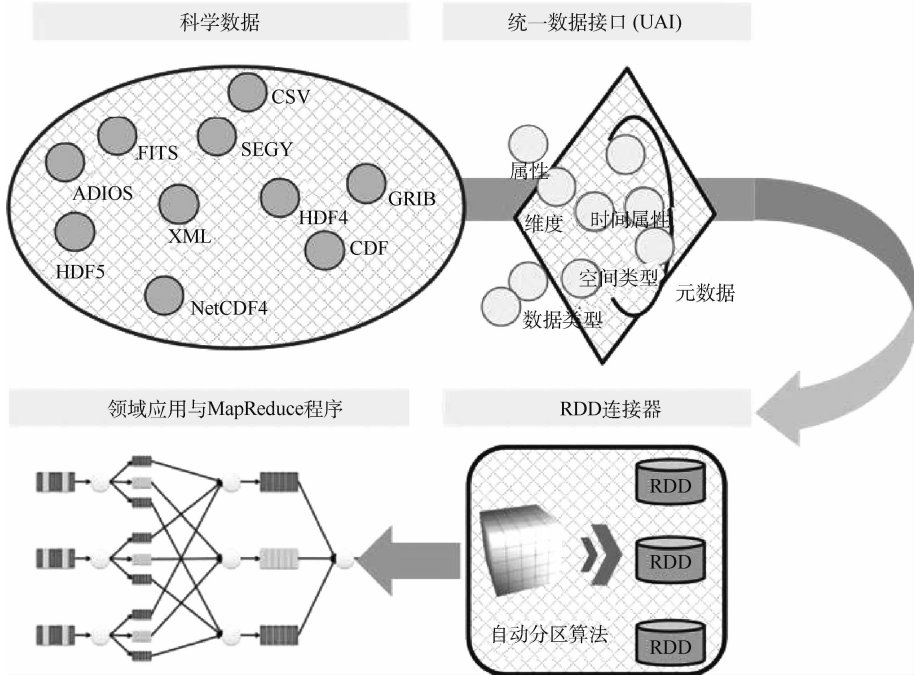


图1 BSDF整体架构示意图

Fig. 1 Schematic diagram of BSDF overall architecture

BSDF 主要由 3 个主要部分组成:

1) 统一数据接口。采用模型驱动的方法从元数据提取数据抽象模型,通过一系列模型转换与规则解析,自动化生成科学数据识别与读取接口。

2) 自动分区算法。通过元数据读取的信息和 Apache Spark 运行时的配置信息计算,根据超平面计算得到最优分区方案,并以 B-Tree 形式存储在主节点内存中,以引导 RDD 的生成与任务分发。

3) RDD 连接器。采用 Apache Spark 建立并行化集合方式从自动分区算法的分区结果生成 RDD,并通过分区索引映射相应的数据分区,最终将任务分发到各个节点。

BSDF 通过与 Apache Spark 并行计算框架的无缝连接来挖掘科学数据处理的性能与自动容错优势。为了提高易用性并简化复杂的科学数据处理流程,提供了 2 个 API 接口用于自动识别异构的科学数据,并且映射到应用程序完成隐式并行处理,通过自动分区算法的集成,实现自动化的任务粒度划分以达到性能优化。BSDF 提供的 API 包括:

① read()。科学数据读取接口。首先,通过统一数据接口来识别输入数据类型。然后,从读入的元数据中记录(如数据维度等)相关信息。此外,Apache Spark 的运行环境配置参数(如计算节点数量、节点计算核心个数等)也同时被读

入,用以计算分区结果,并以 Apache Spark 入口点 SparkContext 的形式作为参数传入。最后,该接口提供了一个可选参数“partition = auto”来供用户决定手动或者自动设置任务分区,如果该参数设置为“auto”,则会通过触发自动分区算法计算任务分区。read()接口的调用代码片段如下:

```
data = BSDF.read(SparkContext, FILE, auto)
```

② map()。RDD 映射接口。该接口的目的是将 read()接口中读入的科学数据自动转化为 BSDF 框架中的数据结构 Spark RDD。首先,读取 read()接口中计算的分区结果集合;然后,引用 Apache Spark 数组并行序列化方式对分区集合建立 RDD,并且与相应的分区数据块进行映射;最后,以“键/值”对的形式生成子 RDD,并分布在分布式环境中。Map()接口的代码片段如下:

```
rdds = BSDF.map(SparkContext, data)
```

BSDF 提供的读取和映射接口通过便捷的方式将统一数据接口和自动分区算法融合,消除了科学数据和 Spark RDD 两类数据结构之间的技术隔阂。最终生成的 RDD 可以与基于 MapReduce 的计算范式、基于 Apache Spark 的并行化处理无缝连接。

## 2.2 统一数据接口

在科学研究领域,由于探索设备与数据模型的多样性,将原始数据直接存储为关系型数据并

以 SQL 或结构化文件形式呈现难度极高。因此,本文提出一个统一数据接口,以汲取异构科学数据结构共性,并根据各特定领域的通用标准与库函数,实现对多种常用异构科学数据的统一访问与通用读取。本文研究了 17 种通用科学数据标准与格式<sup>[6-11]</sup>,并总结了 9 类通用特征,包含 8 种常用科学数据格式,覆盖 9 类包括地球物理、气象、生物和医学等在内的科研领域。统一数据接口整体逻辑架构如图 2 所示。

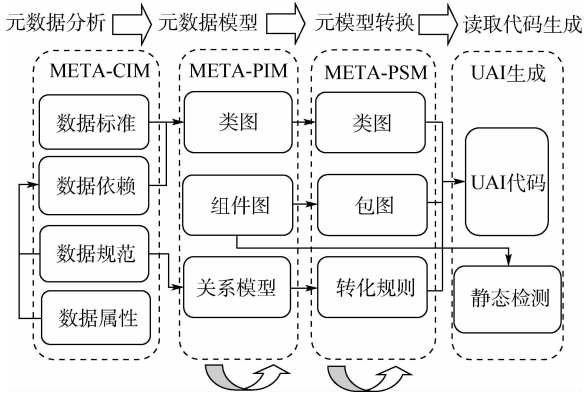


图 2 统一数据接口整体逻辑架构示意图

Fig. 2 Schematic diagram of overall logic architecture of UAI

UAI 引入软件工程的 MDA 思想,从科学数据元数据中抽象 UML 数据模型。在 MDA 的实现逻辑中,通过建立计算无关模型 (CIM) 描述数据标准与规格;通过分析建模,抽象元数据共性以生成平台无关模型 (PIM);通过 QVT 再次抽象元数据中的泛化与关联关系及与数据并行化处理相关的所有属性,经过裁剪和重构生成平台相关模型 (PSM);PSM 模型中不仅包括元数据中存储的数据格式、时空属性、位置尺度等信息外,还包括用于支持引导自动分区算法的维度信息等;与数据处理和分析无关的(如采集与设备等)信息将会在模型转换中有选择地过滤。通过 VTL 定义规则转化模版,通过 PSM 模型生成数据读取函数,并嵌入在 BSDF 计算框架中。目前,BSDF 支持通用的科学数据格式包括 HDF4/5、NetCDF3/4、FITS<sup>[12]</sup>、ADIOS<sup>[13]</sup>、CDF、SEG Y<sup>[14]</sup>、GRIB 及 XML 等,并且支持数据类型扩展。

### 2.3 自动分区算法

由于 Apache Spark 并行计算框架与基于 MapReduce 的计算范式缺乏对多维异构科学数据的通用语义索引与数据分区方法,难以对海量科学数据采取并行的策略来进行分布式处理与计算。正如 Gray 等<sup>[1]</sup>指出,科学家需要采用智能的索引与数据组织方式来引导科学数据查询。另外,

Apache Spark 针对结构化数据采用的基于行列哈希值的哈希分区算法和基于水塘随机抽样算法的范围分区算法,都可能会引发由于数据分区不均而造成性能损失等问题。因此,本文提出了一种自动分区算法,通过科学数据元数据的轻量级信息建模来计算任务粒度布局,并利用 Apache Spark 的内存中计算来加速基于 MapReduce<sup>[14]</sup> 的应用程序。自动分区算法如下。

#### 算法 1 自动分区算法。

输入:节点数量  $n_e$ 、节点核数  $n_c$ 、数据维度  $n_d$ 。

1. if ALL  $i \in n_d \text{ mod}(n_e \times n_c) \neq 0$  then
2.  $\min_{\text{dim}} = \arg \min_i (\text{abs}(i - n_e \times n_c))$
3.  $\text{id} = \min_{\text{dim}}$  在  $n_d$  中的对应的 index
4. if  $\min_{\text{dim}} \geq n_e \times n_c$  then
5.  $\text{partition}[\text{id}] = \min_{\text{dim}} / (n_e \times n_c)$
6. else
7.  $\text{partition}[\text{id}] = 1$
8. end if
9. else
10.  $\min_{\text{dim}} = \arg \min_i (\text{imod}(n_e \times n_c) \approx 0)$
11.  $\text{id} = \min_{\text{dim}}$  在  $n_d$  中的对应的 index
12.  $\text{partition}[\text{id}] = \min_{\text{dim}} / (n_e \times n_c)$
13. end if

#### 输出

由于分布式计算环境通常存在异构性,在此基础上,本文提出了针对高维科学数据的自动分区算法,包括节点级分区算法、更细粒度的节点核心级分区算法和非阻塞分区算法。自动分区算法思路为:在节点级分区算法中,高维数据分区所依据的分区维度是通过维度数组  $n_d = \{i_x, i_y, i_z, \dots, i_n\}$  决定的。如果某一维度能够被分配的节点数量  $n_w$  近似整除,那么输入数据在该维度逻辑可分,分区的结果为该维度上按照计算任务粒度数量平均切分的若干个超平面  $x_{\text{edge}}$ ,并均匀分配各个节点上。最终维度数组中可被分配节点数量整除的最小值将会被视为分区维度进行如下计算:

$$f(x) = x \text{ mod } n_w \tag{1}$$

$$x_{\text{edge}} = \min \{x \in n_d \mid f(x) \approx 0\} \tag{2}$$

$$n_{\text{partition}} = x_{\text{edge}} / n_w \tag{3}$$

令  $p = (\text{partition}_1, \text{partition}_2, \dots, \text{partition}_n)$ , 过滤为空的分区为

$$p_{\text{filter}} = \{x \in p \mid x \notin \emptyset\} \tag{4}$$

进而在节点计算核心级分区算法中,采用相同的维度数据计算和超平面切分法来获取节点计算核心级别的分区结果。通过计算所得分区结果通过分区过滤操作按照任务粒度平均分配到各个

计算节点的计算核心  $n_c$  上进行下一步操作。节点计算核心级分区算法计算流程为

$$g(x) = x \bmod n_c \tag{5}$$

$$x'_{edge} = \min \{ x' \in p_{filter} \mid g(x) \approx 0 \} \tag{6}$$

$$n'_{partition} = x'_{edge} / n_c \tag{7}$$

$$\text{令 } p' = (\text{partition}_1, \text{partition}_2, \dots, \text{partition}_n)$$

$$p'_{filter} = \{ x \in p' \mid x \notin \emptyset \} \tag{8}$$

尽管通过节点级和节点核心级分区算法计算的分区,以及通过分区过滤操作过滤为空的超平面空间,能够达到细粒度的任务分区,但是通过超平面划分的分区结果依然无法完全保证任务维度的分区平衡,这种情况在稀疏数据和倾斜数据中表现更为明显<sup>[15]</sup>。不平衡的分区结果在 MapReduce 计算范式会由于任务分配不均而引发性能缺失和掉队任务等问题。例如,在一次迭代过程中,分布式环境中的 Map 任务群需要等待最晚的一个任务执行完成才能执行 Reduce 操作。

为了解决 MapReduce 计算范式中的任务同步问题,借鉴 NOMAD 算法<sup>[16]</sup>中的非阻塞计算机制,并将其应用于第3层分区算法中。非阻塞分区算法遵守2条规则:①独立的分区将会按照节点或节点核心级别分组;②分组后分区将会平均分配到节点或节点核上。图3为自动分区应用于3D数据划分示意图。

由于自动分区算法直接通过元数据计算,相比范围分区和哈希分区能够更为快速地得到分区结果。与原始数据相比,元数据占据的磁盘空间要少得多,因此在数据分区步骤中可以减少数据遍历次数与 shuffle 阶段的磁盘 I/O。计算生成的分区结果将以 B-Tree 的形式缓存在 Master 节点内存中以便多次读取。自动分区算法同时提高了海量科学数据计算移植到 MapReduce 计算范式的可行性。与 Apache Spark 采用的哈希与范围分区算法相比,自动分区算法提供了一种更为严谨的方式来计算 Apache Spark 任务分区颗粒度,并且通过直接对科学元数据中的参数预取与超平面维度计算,降低任务分配不均与掉队风险,从而有

效提升了分区效率、任务吞吐量与内存利用率。

### 3 实验

通过9种基准测试程序来验证 BSDF 的性能与对科学数据的支持程序,包括 Mean、Aggregation、Sort、Logistic Regression、KMeans、Spark Collect、Matrix Multiplication、地震波程序偏移算法 PKTM。通过与目前成熟的支持科学数据计算的 H5Spark<sup>[17]</sup>平台相比较,分析整体性能和可扩展性。此外,为了验证自动分区算法所带来的性能提升,通过与 Apache Spark 原生范围分区方法比较,跟踪任务执行流水线并对比任务分区颗粒度。为了验证 BSDF 在特定科学研究领域的适用性与性能,将地震波偏移算法移植在 BSDF 平台、Hadoop 框架和 Apache Spark 框架,并综合比较性能。

#### 3.1 实验数据与基准测试程序

由于在实验中所对比的 H5Spark 平台仅支持 HDF5 和 NetCDF 格式的科学数据,引入3组科学数据(HDF5、NetCDF 和 SEGY),其中 SEGY(.sgy)格式数据代表地震数据。实验数据类型、数据量及应用的基准测试程序如表1所示。

搭建集群构建实验环境,集群包含32个计算节点(每个节点拥有10个主频为2.2 GHz 的 Intel Xeon Silver 4210 CPU 核心),并在 Master 节点和 Worker 节点均部署了 Apache Spark 3.0 并行计算框架。实验环境应用的基准测试程序如表2所示。

表1 实验数据与对应基准测试程序

Table 1 Experimental data and corresponding benchmark programs

数据类型	数据量/GB	基准测试程序
HDF5	19.23	Mean, Aggregation, Sort, KMeans, Logistic Regression, Alternating Least Squares, Spark Collect, Matrix Multiplication
NetCDF	20.13	Mean, Aggression, Sort, KMeans, Logistic Regression, Alternating Least Squares, Spark Collect, Matrix Multiplication
SEGY	4.95	PKTM

#### 3.2 BSDF 与 H5Spark 的性能与可扩展性比较

BSDF 与 H5Spark 的性能比较共引入了9项基准测试程序。图4为分布式环境中 BSDF 与 H5Spark 的总体性能对比。整体加速比大于1表示 BSDF 在该基准测试中性能优于 H5Spark,反之亦然。在实验比对结果中,Mean、Sort、Matrix Multiplication、

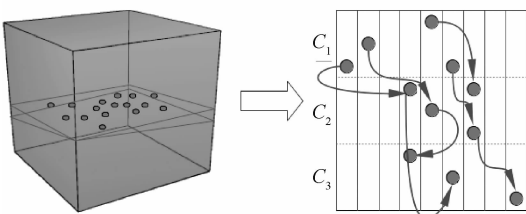


图3 自动分区算法在3D数据的分区示意图

Fig. 3 Schematic diagram of partitioning results of auto partitioning algorithms on 3D data

表 2 实验基准测试集列表

Table 2 Experimental benchmark test sets

序号	基准测试程序	基准测试程序描述
1	Mean	用户自定义函数,用于返回整个数据集的平均值
2	Aggregation	用户自定义函数,用于返回整个数据集的累加结果
3	Sort	Spark 内置函数,对输入数据进行排序
4	KMeans	聚类算法,用于将输入数据放入预定义的多个类中
5	Logistic Regression	分类算法,采用梯度斜率方法最小化逻辑回归模型训练
6	Alternating Least Squares	Spark MLlib 内置函数,与矩阵分解协同计算使得差平方和最小
7	Spark Collect	Spark 内置函数,将数据从各个节点拉取到驱动器节点
8	Matrix Multiplication	用户自定义函数,分布式矩阵相乘
9	PKTM	叠前时间偏移成像算法,专门用于地震数据处理

Alternating Least Squares、Spark Collect 和 Aggregation 的基准测试结果加速比达到 H5Spark 的 1.39 ~ 1.68 倍,而 PKTM 则获得了较高的 2.62 倍的加速比。

实验结果还显示,KMeans 基准测试在节点从 4 升至 16 时有明显的上下波动,并且在节点数为 8 时加速比小于 1。根据实验结果分析表明,并行化 KMeans 算法需要在 Map 阶段对每块分区数据求对应最近的聚类中心点,并在 Reduce 阶段对同一聚类中心的中间数据求和更新一个新的聚类中心点,此过程将重复迭代直到预设的  $K$  个聚类距离收敛。自动分区算法相较 H5Spark 会提供更为细粒度的分区,尤其在计算节点较多的情况下,直接影响到 reduce 的数据传输量和计算量。根据执行时间分析得到,自动分区算法中数据通信的时间占据了总体执行时间很大的比重。因此,自动分区算法在并行化 KMeans 时会随着节点数的提升而自动增加数据分区的颗粒度,而更为细致的数据分区颗粒度在 KMeans 聚类模型中并不代表着更优的性能。

除此以外,通过在 2 个 Worker 节点和 16 个 Worker 节点的性能结果比较来验证 BSDF 和 H5Spark 的可扩展性。图 5 为多节点环境根据 9 项基准测试总结的可扩展性比较结果。理论上,当可扩展性加速比达到 8 倍时,意味着性能随着计算节点的扩展达到了最优的线性加速比。BSDF 在 KMeans 和 PKTM 上均获得了更好的可扩展性增益,与 H5Spark 相比,整体加速比在 1.2 ~ 1.9 倍之间,体现了更优的可扩展性。

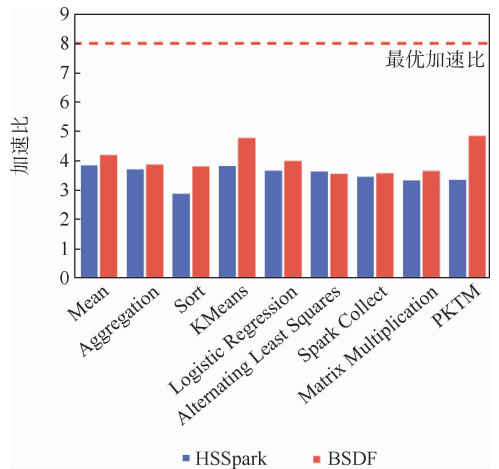


图 5 BSDF 与 H5Spark 在多节点环境可扩展性的比较  
Fig. 5 Overall scalability comparison between BSDF and H5spark in multi-node environments

### 3.3 自动分区算法性能分析

为了进一步验证自动分区算法在多节点环境下的任务分区性能优势,在具有 16 个计算节点的分布式环境中对比了自动分区算法和 Apache Spark 原生范围分区 (range partitioning) 算法之间的性能差异。通过 Spark Collect 方法从各个节点拉取总计 32.3 GB 的 FITS 格式数据。图 6 为在不同分区算法下的整体性能与任务分区流水线。

由图 6 可知,自动分区算法比 Apache Spark 范围分区算法在调用 Spark Collect 方法时获得了更优的整体性能。此外,自动分区算法在多节点

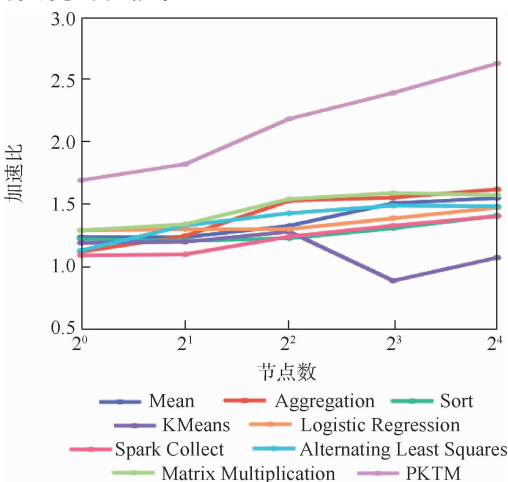


图 4 BSDF 与 H5Spark 整体性能比较  
Fig. 4 Overall performance comparison between BSDF and H5Spark

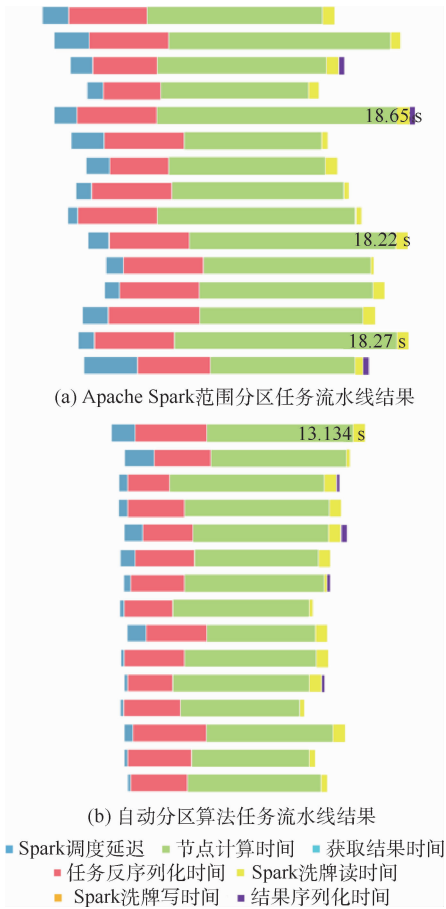


图6 Apache Spark 范围分区与自动分区算法对于 Spark Collect 任务的整体性能与任务流水线比较  
Fig. 6 Overall performance comparison between Apache Spark range partitioning and auto partitioning algorithm in both Spark Collect tasks and task pipelines

执行环境中能生成更为平衡的任务流水线,并降低了调度程序延迟。根据性能对比分析发现,性能差异主要由于 Apache Spark 范围分区采用了水塘抽样 (reservoir sampling)<sup>[18]</sup> 算法,该算法先从输入数据中随机抽取样本,再根据采样结果计算分区边界,因此,Apache Spark 范围分区算法只能保证一个近似平衡的任务划分结果,而不平衡分区将导致在多节点环境下由于任务分配不均而产生的掉队任务,从而影响整体性能。与自动分区算法相比,Apache Spark 范围分区算法需要在 Shuffle 阶段进行输入数据行索引的重排列,因此会产生额外的计算成本;而自动分区算法通过元数据中的数据维度直接指定分区边界来解决此类问题,并提供了一种基于多节点环境的更为细致的分区方式,从而有效减少了掉队任务的发生频次。

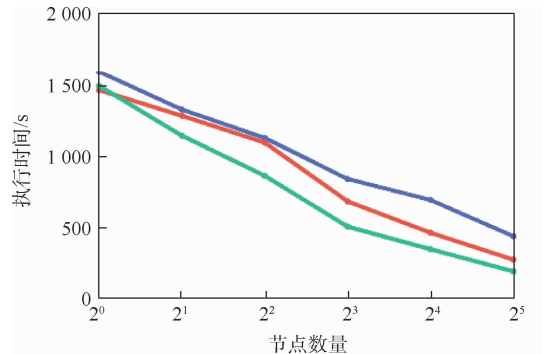
### 3.4 地震波偏移算法基准测试结果分析

为了验证在特定科学研究领域对于科学数据计算的可行性和性能结果,引入 PKTM 算法作为基准测试程序。PKTM 是一种地震勘探领域广泛

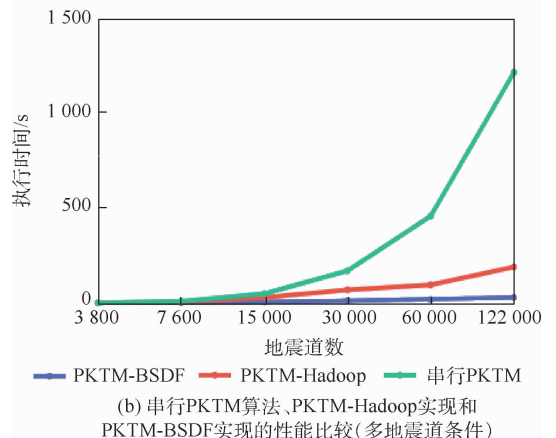
使用的成像算法,需要大量的迭代次数以完成时间偏移,在地质勘探与地震监测领域属于计算量庞大的数据密集型算法。

分别在 Apache Spark 3.0 和 Hadoop 3.2.0 框架中实现 PKTM 算法,并将其性能与 BSDF 算法进行比较,通过调整分布式环境中节点的数量验证在多节点环境下的性能差异。此外,在单节点多核环境中,通过 BSDF 并行化 PKTM 程序,与传统串行 PKTM 程序进行性能比较。在 PKTM 基准测试程序中,采用 CREWES MATLAB (afd vmodel) 生成 PKTM 中的速度模型  $velocity(x, z)$ 。实验所用的地震数据集包含 122 000 条地震道 (trace) 数据,每条地震道数据包含 9 000 个样本,总计 4.9 GB。

图 7 为 PKTM 算法在分布式环境中对应的 Hadoop、Spark 和 BSDF 算法实现的总体性能结果。在图 7(a) 多节点环境性能比较中,PKTM-BSDF 与 PKTM-Hadoop 算法实现相比,达到了 2.29 倍的加速比,与 Spark-PKTM 算法实现相比,达到了 1.43 倍的加速比。在图 7(b) 多核环境性能比对中,与 PKTM-Hadoop 实现相比,PKTM-BSDF



(a) Hadoop、Spark 和 BSDF 在 PKTM 实现上性能比较 (多节点环境下)



(b) 串行 PKTM 算法、PKTM-Hadoop 实现和 PKTM-BSDF 实现的性能比较 (多地震道条件)

图 7 PKTM 算法在多节点与多地震道下基于 Hadoop、Spark 和 BSDF 实现的性能比较  
Fig. 7 Performance comparison of Hadoop, Spark and BSDF based implementations of PKTM algorithm with multiple nodes and multiple seismic traces

在地震道数据为 122 000 条时达到了 6.52 倍的加速比,与 PKTM 的串行实现相比,达到了 41.62 倍的加速比。

实验结果表明,BSDF 在分布式和多核环境均具有良好的性能,并且随输入数据量的线性增加,性能的提升呈线性增长趋势。图 8 显示了 PKTM-BSDF 偏移成像结果。图中:ft 为英尺,1 ft = 0.304 8 m。

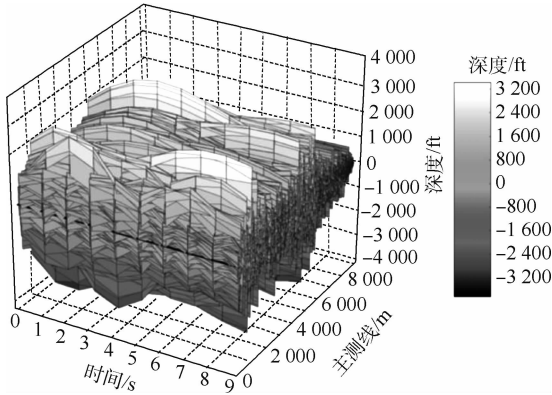


图 8 PKTM 地震数据时间偏移结果

Fig. 8 Results of PKTM time migration of seismic data

## 4 结 论

科学数据的格式多样性和高复杂度往往局限了领域科学家从海量科学数据中获取复杂规则与潜在联系。因此,在科学研究领域迫切需要一种通用便捷的访问方式,应用分布式计算来提升并行 I/O 和分析处理的能力<sup>[19]</sup>。本文结论如下:

1) 提出并实现了一个海量科学数据的高性能计算框架 BSDF。能够支持多种通用异构科学数据的自动化读取与并行处理。目前,BSDF 已支持包括 HDF5、NetCDF、ADIOS、SEG Y 和 FITS 在内的 8 种常用科学数据,涵盖了包括生物学、天文学、地震、物理、生态、基因组和气候研究领域在内的 9 个科学研究领域。

2) 提出一个科学数据统一访问接口。采用 MDA 思想对异构科学数据元数据进行建模,通过 VTL 定义规则转化模版,通过 PSM 模型生成数据读取函数,以接口形式封装在 BSDF 中,最终实现对异构和层次化科学数据的无差别访问。

3) 提出了一种基于元数据与集群计算资源的自动分区算法,通过元数据参数预取与超平面分割算法的统筹计算,建立数据分区策略来决定 MapReduce 任务粒度,并借鉴 NOMAD 的非阻塞算法,优化分区结果。

4) 在与 H5Spark 科学数据计算平台和 Apache Spark 范围分区算法的实验比较中,通过

9 项基准测试程序的性能和可扩展性评估,BSDF 达到了 1.39 ~ 1.68 倍的整体性能提升。通过特定领域 PKTM 算法实现与评估,验证了 BSDF 的性能优势与易用性。

在未来的工作中,计划采用 BSDF 对矩阵分解和深度卷积网络进行性能优化,并将科学数据处理过程迁移到 GPU 和 CPU 等复合环境中,以利用 GPU 的高吞吐量特性提升性能。

## 参考文献 (References)

- [ 1 ] GRAY J, LIU D T, NIETO-SANTISTEBAN M, et al. Scientific data management in the coming decade [ J ]. ACM SIGMOD Record, 2005, 34 ( 4 ) : 34-41.
- [ 2 ] The HDF Group. Hierarchical data format. Version 5 [ EB/OL ]. [ 2020-12-01 ]. <http://www.hdfgroup.org/HDF5>.
- [ 3 ] UCAR Community Programs. Network common data form (NetCDF) [ EB/OL ]. [ 2020-12-01 ]. <https://www.unidata.ucar.edu/software/netcdf>.
- [ 4 ] ZAHARIA M, CHOWDHURY M, FRANKLIN M J, et al. Spark: Cluster computing with working sets [ C ] // Proceedings of the 2nd USENIX Conference on Hot Topics in Cloud Computing. New York: ACM, 2010: 1-10.
- [ 5 ] SUTTON J, AUSTIN Z. Qualitative research: Data collection, analysis, and management [ J ]. The Canadian Journal of Hospital Pharmacy, 2015, 68 ( 3 ) : 226-231.
- [ 6 ] MIKAEL N, AMBJÖRN N, ERIK D, et al. Harmonization methodology for metadata models [ EB/OL ]. [ 2020-12-01 ]. <https://hal.archives-ouvertes.fr/hal-00591548>.
- [ 7 ] DPLA. Metadata application profile. Version 4. 0 [ EB/OL ]. [ 2020-12-01 ]. <http://dp.la/info/wpcontent/uploads/2015/03/MApV4.pdf>.
- [ 8 ] DIAMANTOPOULOS N, SGOUROPOULOU C, KASTRANTAS K, et al. Developing a metadata application profile for sharing agricultural scientific and scholarly research resources [ C ] // Research Conference on Metadata and Semantic Research. Berlin: Springer, 2011: 453-466.
- [ 9 ] RILEY J. Understanding metadata: What is metadata, and what is it for? [ M ] // WOOLCOTT L. Baltimore: National information standards organization. Oxford: Taylor, 2017: 669-670.
- [ 10 ] BARGMEYER B E, GILLMAN D W. Metadata standards and metadata registries: An overview [ EB/OL ]. [ 2020-12-01 ]. <https://www.bls.gov/osmr/research-papers/2000/pdf/st000010.pdf>.
- [ 11 ] JONES M B, BERKLEY C, BOJILOVA J, et al. Managing scientific metadata [ J ]. IEEE Internet Computing, 2001, 5 ( 5 ) : 59-68.
- [ 12 ] HANISCH R J, FARRIS A, GREISEN E W, et al. Definition of the flexible image transport system ( FITS ) [ J ]. Astronomy & Astrophysics, 2001, 376 ( 1 ) : 359-380.
- [ 13 ] PARK J K. Improving the performance of HDFS by reducing I/O using adaptable I/O system [ C ] // 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT). Piscataway: IEEE Press, 2016: 3139-3144.

- [14] SEG Technical Standards Committee. SEG-Y\_r2.0; SEG-Y revision 2.0 data exchange format[S]. [S. I.]: Society of Exploration Geophysicists, 2017.
- [15] DEAN J, GHEMAWAT S. MapReduce: Simplified data processing on large clusters[J]. Communications of the ACM, 2008, 51(1): 107-113.
- [16] YUN H, YU H F, HSIEH C J, et al. NOMAD: Non-locking, stochastic multi-machine algorithm for asynchronous and decentralized matrix completion[J]. Proceedings of the VLDB Endowment, 2013, 7(11): 975-986.
- [17] LIU J, RACAH E, KOZIOL Q, et al. H5Spark: Bridging the I/O gap between Spark and scientific data formats on HPC systems [C]// Proceedings of the Cray Users Group, 2016.
- [18] AGARWAL A, CHAPPELLE O, DUDÍK M, et al. A reliable effective tera scale linear learning system[J]. Journal of Machine Learning Research, 2014, 15(1): 1111-1133.
- [19] TIAN Y, LIU C, YAN H H. Accelerate large-scale seismic data Kirchhoff time migration in spark[C]// 2018 4th International Conference on Information Management (ICIM). Piscataway: IEEE Press, 2018: 41-45.

## A computing framework for massive scientific data based on auto-partitioning algorithm

TIAN Yang, YAN Haihua \*

(School of Computer Science and Engineering, Beihang University, Beijing 100083, China)

**Abstract:** In the scientific research field, storage capacity, processing efficiency and analysis accuracy cannot keep pace with the exponential growth rate of scientific data. Thus, a massive scientific data calculation framework named BSDF is proposed based on scientific data structure and standards. A unified data interface based on model-driving is integrated to implement indiscriminate access to heterogeneous scientific data. Then an auto-partitioning algorithm based on scientific metadata is proposed, which determines task granularities through parameter prefetching and hyperplane dimension calculation. Experimental results show that compared with the performance of the H5Spark framework, that of the BSDF is increased by 39% – 68% in nine benchmark tests. In the optimization of the domain-specific PKTM algorithm, a speedup ratio is increased by 41.62 times.

**Keywords:** scientific data; model-driving; partitioning algorithm; pre-stack Kirchoff time migration; software engineering; Spark

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2020.0727

# 移动机器人自主动态避障方法

张贺<sup>1</sup>, 缪存孝<sup>1,\*</sup>, 唐友军<sup>1</sup>, 闫晓强<sup>1</sup>, 时彦洋<sup>1</sup>, 余远金<sup>2</sup>

(1. 北京科技大学 机械工程学院, 北京 100083; 2. 北京理工大学 自动化学院, 北京 100081)

**摘 要:** 针对基于水流场的全向移动机器人动态避障过程中可能出现的避障失效和避障路径过长问题, 提出了一种基于速度斥力场改进人工势场方法的全向移动机器人自主动态避障方法。详细分析基于水流场改进的人工势场法在动态避障过程中存在的移动机器人从障碍物前方绕过而引起的避障路径过长或避障失效问题; 在基于水流场的人工势场避障方法中, 根据移动机器人和动态障碍物的相对速度, 引入速度斥力场, 使得移动机器人从动态障碍物的后方通过, 实现全向移动机器人的安全自主动态避障。通过仿真和室内避障实验, 验证了所提自主动态避障方法的有效性和实用性。

**关键词:** 移动机器人; 人工势场; 水流场; 速度斥力场; 动态避障

**中图分类号:** TP242.6

**文献标志码:** A **文章编号:** 1001-5965(2022)06-1013-09

自主避障一直是移动机器人研究领域的热点和难点<sup>[1-4]</sup>, 现有多种避障方法, 包括: ①基于机器学习的避障方法, 如神经网络算法<sup>[5]</sup>、强化学习算法<sup>[6]</sup>, 这类算法一般比较复杂, 对处理器硬件要求较高; ②基于几何模型搜索的避障方法, 如A\*算法<sup>[7-8]</sup>、D\*算法<sup>[9-11]</sup>等, 这类算法大多应用于静态路径规划; ③基于群体智能的避障方法, 如粒子群优化算法<sup>[12-13]</sup>、遗传算法<sup>[14]</sup>等, 这类算法往往需要多次迭代收敛, 实时性不高; ④基于虚拟势场的避障方法, 这类算法结构比较简单, 方便底层进行实时控制, 但是地形较复杂时, 容易进入局部最优解。

人工势场(artificial potential field, APF)法具有结构简单、计算量小、实时性好等优点, 广泛应用于无人系统避障问题中, 但是人工势场法存在易陷入局部极小值、路径振荡、目标不可达等问题。针对这些问题, 国内外学者提出众多解决方

案<sup>[15-18]</sup>。郭一聪等<sup>[15]</sup>将改进势场法应用于无人机三维路径规划, 利用相对距离改进势场函数, 解决目标不可达问题, 并采用记忆性合力的方法解决局部路径振荡问题, 但只针对静态避障进行了仿真研究; 陈麒杰等<sup>[16]</sup>将人工势场算法应用于无人机集群避障, 在传统人工势场的基础上引入无人机组间的相互作用力, 同时定义虚拟的前置形心, 解决了极小值问题和局部最小值问题, 并进行了仿真研究; 崔宝侠和宋佳瑞<sup>[17]</sup>在传统人工势场法的引力公式中引入动态目标的速度和加速度, 并重新定义斥力方向, 使其与引力的夹角不大于90°, 避免机器人在避障中进入局部极小值, 同时提高机器人的动态避障能力, 但是该方法需要在几种不同的情况下分别设计斥力的方向, 导致方法比较复杂, 同时仿真实验只验证了静态避障能力和动态目标追踪能力; 杜婉茹等<sup>[18]</sup>针对传统人工势场法存在的动态避障精度不高、

收稿日期: 2021-01-02; 录用日期: 2021-04-20; 网络出版时间: 2021-05-28 10:55

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20210527.1709.001.html](http://kns.cnki.net/kcms/detail/11.2625.V.20210527.1709.001.html)

基金项目: 国家重点研发计划(2018YFC0810500); 国家自然科学基金(60277005); 北京理工大学青年教师学术启动计划(2020CX04104)

\*通信作者. E-mail: miaocx@ustb.edu.cn

**引用格式:** 张贺, 缪存孝, 唐友军, 等. 移动机器人自主动态避障方法[J]. 北京航空航天大学学报, 2022, 48(6): 1013-1021.  
ZHANG H, MIAO C X, TANG Y J, et al. Dynamic obstacle avoidance method for mobile robots[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(6): 1013-1021 (in Chinese).

易出现抖动现象及目标不可达问题,提出一种利用局部障碍及目标点运动信息改进人工势场的路径规划算法,并通过与传统人工势场法进行对比仿真,验证了该算法在复杂环境下的优越性;张大志等<sup>[19]</sup>提出一种基于水流场改进的人工势场方法,解决了传统人工势场存在的局部极小值、路径振荡及目标不可达等问题,但是该方法在进行动态避障时,存在移动机器人从动态障碍物前方绕过,增加避障路程和避障失效概率的问题。

针对基于水流场改进的人工势场法进行动态避障时存在的问题,本文引入动态障碍物的速度斥力场,对动态障碍物进行提前躲避,解决了由于动态障碍物引起牵引运动可能导致动态避障失效和避障路径过长的问题。通过仿真和室内全向移动机器人动态避障实验,验证了本文方法的有效性和实用性。

## 1 基于水流场的人工势场法及问题

### 1.1 基于水流场的人工势场法原理

人工势场法依据传感器获取外部环境信息以构建虚拟势力场,包括势场函数和力函数<sup>[18]</sup>:

$$U_{att} = \frac{1}{2}K_{att}(x - x_g)^2 \quad (1)$$

$$U_{rep} = \begin{cases} \frac{1}{2}K_{rep}\left(\frac{1}{d} - \frac{1}{d_0}\right)^2(x - x_g)^n & d \leq d_0 \\ 0 & d > d_0 \end{cases} \quad (2)$$

$$F_{att} = -\nabla U_{att} = -K_{att}|x - x_g| \quad (3)$$

$$U_{rep} = \begin{cases} -(-y + xi) \\ -(y - xi) \\ -(\cos \gamma \sqrt{x^2 + y^2} + \sin \gamma \sqrt{x^2 + y^2}) \end{cases}$$

式中: $U_{rep}$ 为水流场坐标系下斥力场; $x$ 为障碍物在水流场坐标系下的横坐标; $y$ 为障碍物在水流场坐标系下的纵坐标; $\gamma$ 为在水流场坐标系中从 $X$ 轴的正方向向障碍物至目标点的矢量逆时针旋转的角度。

水流场方向的斥力值 $F_{rep1}$ 为

$$F_{rep1} = K_{rep}\left(\frac{1}{d} - \frac{1}{d_0}\right)\frac{(x - x_g)^n}{d^2} \quad (6)$$

当障碍物处于水流场坐标系的左半侧时,给移动机器人奔向目标点的斥力,当障碍物处于水流场的右半侧时,将斥力方向改为障碍物朝向目

$$F_{rep} = -\nabla U_{rep} = \begin{cases} \frac{n}{2}K_{rep}\left(\frac{1}{d} - \frac{1}{d_0}\right)^2(x - x_g)^{n-1} & d \leq d_0 \\ 0 & d > d_0 \end{cases} \quad (4)$$

式中: $U_{att}$ 为引力势场函数; $U_{rep}$ 为斥力势场函数; $F_{att}$ 为引力; $F_{rep}$ 为斥力; $K_{att}$ 为引力系数; $K_{rep}$ 为斥力系数; $x$ 为移动机器人当前位置; $x_g$ 为移动机器人目标位置; $d$ 为移动机器人与障碍物之间的距离; $d_0$ 为障碍物的作用半径; $n$ 为待优化的参数。

为解决斥力场过大且与引力场角度为钝角时导致移动机器人后退运动进而引起的振荡运动问题,采用水流场思想,改变斥力场方向,以移动机器人为坐标原点, $X$ 轴指向目标点建立水流场坐标系,障碍物位置与水流斥力场方向关系如图1所示。图中: $U_{repi}(i=1,2,3,\dots)$ 为水流斥力场矢量,水流场坐标系下斥力场作用复势场可表示为<sup>[19]</sup>

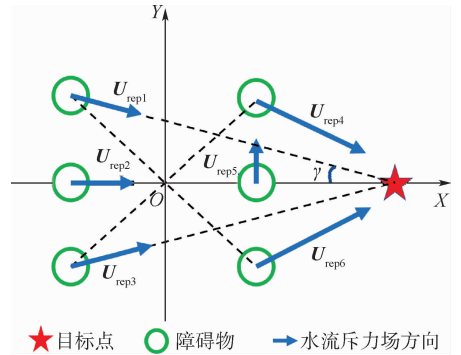


图1 障碍物位置与水流斥力场方向的关系  
Fig.1 Relation between position of obstacles and direction of repulsive force in flow field

$$\begin{cases} x \geq 0, y > 0 \\ x \geq 0, y \leq 0 \\ x < 0 \end{cases} \quad (5)$$

标点的方向。

基于水流场改进的人工势场法,可解决传统人工势场法路径振荡、陷入局部极小值等问题。但是,该方法在进行动态避障时会出现移动机器人从障碍物运动方向的前方通过完成避障,此行为将增大动态避障失效的概率,以下对此进行详细讨论。

### 1.2 基于水流场的人工势场法存在的问题

采用基于水流场的人工势场法进行动态避障,当动态障碍物 $(x_{ob}, y_{ob})$ 于水流场坐标系的右半侧,运动过程中,受到的水流场中的斥力为 $F_{rep1}$ ,其在

水流场坐标系下的  $XY$  轴分量可表示为

$$\begin{cases} F_{\text{rep1}x} = \|F_{\text{rep1}}\| \cos \beta \\ F_{\text{rep1}y} = \|F_{\text{rep1}}\| \sin \beta \end{cases} \quad (7)$$

式中:  $\beta = \arctan\left(\frac{-y_{\text{ob}}}{x_{\text{ob}}}\right)$ 。

当障碍物从  $Y$  轴负方向向  $Y$  轴正方向运动时,如图 2 所示,  $0 < \beta < \pi/2$ ,斥力场在  $Y$  轴的分量  $F_{\text{rep1}y} > 0$ ,该斥力会导致移动机器人沿着动态障碍物的运动方向运动。避障过程中,移动机器人受到的势场引力、势场斥力和运动路径如图 2 所示。在合力的作用下,移动机器人将从动态障碍物的前方绕过以完成避障过程,运动轨迹表现出一种被动态障碍物牵引运动的现象。显然,由于牵引运动的存在,不仅会增加避障路径的长度,而且会增加移动机器人与动态障碍物相撞的概率,从而引起动态避障失败。

当障碍物从  $Y$  轴正方向向  $Y$  轴负方向运动时,  $-\pi/2 < \beta < 0$ ,斥力场在  $Y$  轴的分量  $F_{\text{rep1}y} < 0$ ,

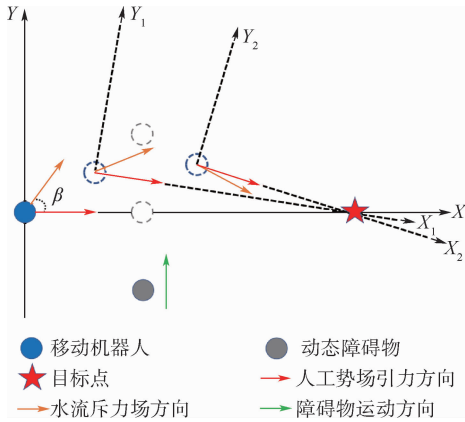


图 2 力和运动示意图

Fig.2 Schematic diagram of force and motion

该斥力同样会导致移动机器人沿着动态障碍物的运动方向运动。

## 2 基于速度斥力场的动态避障方法

为解决基于水流场的人工势场法在动态避障中存在的问题,考虑移动机器人与动态障碍物的相对速度,提出一种基于速度斥力场的动态避障方法。通过获取动态障碍物的位置、速度信息,构建速度斥力场,速度斥力场函数为

$$U_{\text{repr}}(X, X_{\text{ob}}, \mathbf{V}, \mathbf{V}_{\text{ob}}) = \begin{cases} K_v \mathbf{V}_{\text{or}} \sin \theta & \rho(X, X_{\text{ob}}) \leq \rho_{\text{obv}} \\ 0 & \rho(X, X_{\text{ob}}) > \rho_{\text{obv}} \end{cases} \quad (8)$$

式中:  $X$  和  $X_{\text{ob}}$  分别为移动机器人和动态障碍物的位置;  $K_v$  为速度斥力场系数,正实数;  $\mathbf{V}_{\text{or}}$  为移动机器人速度  $\mathbf{V}$  和动态障碍物速度  $\mathbf{V}_{\text{ob}}$  的矢量差(见图 3);  $\theta$  为  $\mathbf{V}_{\text{or}}$  与移动机器人和当前动态障碍物的位置连线的夹角;  $\rho(X, X_{\text{ob}})$  为移动机器人与动态障碍物的距离;  $\rho_{\text{obv}}$  为动态障碍物速度斥力场的影响半径。

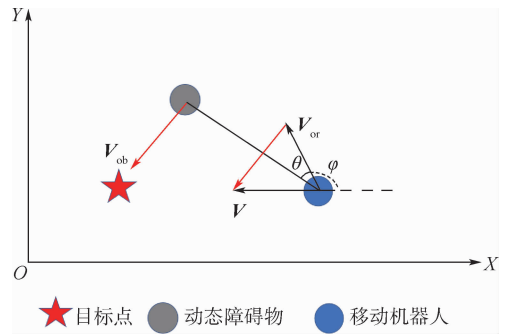


图 3 速度斥力场示意图

Fig.3 Schematic diagram of velocity repulsion field

速度斥力函数为

$$\rho(X, X_{\text{obs}}) \leq \rho_0 \cap \theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \quad (9)$$

其他

度斥力为零。

移动机器人在避障过程中受到的合力势场为

$$U_{\text{total}} = \begin{cases} U_{\text{att}} + U_{\text{rep1}} + U_{\text{rep2}} & \theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \\ U_{\text{att}} + U_{\text{rep1}} & \text{其他} \end{cases} \quad (10)$$

式中:  $U_{\text{total}}$  为合力势场;  $U_{\text{att}}$  为人工引力场;  $U_{\text{rep1}}$  为水流斥力场;  $U_{\text{rep2}}$  为速度斥力场。

移动机器人在合力势场中的受力如图 4 所示。

$$F_{\text{repr}} = -\text{grad } U_{\text{repr}}(X, X_{\text{ob}}, \mathbf{V}, \mathbf{V}_{\text{ob}}) = \begin{cases} \frac{K_v \mathbf{V}_{\text{or}} (-\sin \theta \cos \varphi \mathbf{X}_e + \cos \theta \sin \varphi \mathbf{Y}_e)}{\sqrt{X_{\text{or}}^2 + Y_{\text{or}}^2}} \\ 0 \end{cases}$$

式中:  $\varphi$  为世界坐标系下  $X$  轴与  $\mathbf{V}_{\text{or}}$  形成的夹角;  $X_{\text{or}}$  和  $Y_{\text{or}}$  分别为世界坐标系下移动机器人和动态障碍物的横向和纵向距离;  $\mathbf{X}_e$  和  $\mathbf{Y}_e$  为单位矢量值。

当  $\theta \in (-\pi/2, \pi/2)$  时,动态障碍物沿运动方向的延长线和移动机器人运动方向的延长线会有交集,彼此受到影响,便产生了速度斥力场,使移动机器人能及时躲避动态障碍物。当超出这个角度范围时,两者则朝着相互远离的方向运动,速

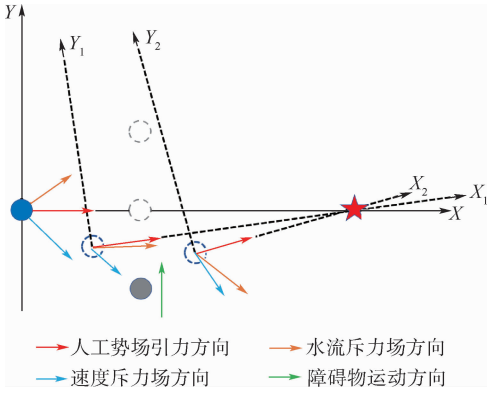


图4 移动机器人在合力势场中的受力示意图

Fig.4 Schematic diagram of forces on mobile robot in resultant force potential field

引入速度斥力场后,当存在动态障碍物时,移动机器人受到的速度斥力为  $F_{rep2}$ ,其在水流场坐标系下的  $XY$  轴分量可表示为

$$\begin{cases} F_{rep2x} = \|F_{rep2}\| \cos \alpha \\ F_{rep2y} = \|F_{rep2}\| \sin \alpha \end{cases} \quad (11)$$

式中:  $\alpha$  为速度斥力方向与  $X$  轴的夹角。

当动态障碍物从  $Y$  轴负方向向  $Y$  轴正方向运动时,如图4所示,  $-\pi < \alpha < 0$ ,斥力场在  $Y$  轴的分量  $F_{rep2y} < 0$ ,此斥力与动态障碍物运动方向相反,使得移动机器人能够从动态障碍物的后方通过,保证移动机器人能安全地完成动态避障过程。

对于动态障碍物从  $Y$  轴正方向向  $Y$  轴负方向运动的情况,  $0 < \alpha < \pi$ ,斥力场在  $Y$  轴的分量  $F_{rep2y} > 0$ ,此斥力会同样引导移动机器人从动态障碍物后方完成避障。

### 3 仿真实验

建立  $16 \times 20$  的仿真地图,分别对基于水流场的人工势场法(称为方案A)和引入速度场的方法(称为方案B)进行动态避障仿真实验。为测试速度斥力场的作用,在地图中设置2个动态障碍物,一个沿  $Y$  轴正方向运动,另一个沿  $X$  轴负方向运动。设定移动机器人的起点为  $(0,8)$ ,目标点为  $(20,9)$ 。仿真的算法参数设置如表1所示。

表1 算法参数

Table 1 Algorithm parameters

参数	数值
引力系数 $K_{att}$	10
斥力系数 $K_{rep}$	30
速度斥力场系数 $K_v$	40
待优化参数 $n$	2
速度斥力场影响半径 $\rho_{obv}$	1.5

2种方案仿真关键时刻的曲线分别如图5和图6所示。仿真结果的数据对比如表2所示。

通过对比图5和图6实验结果,当遇到沿  $Y$  轴正向运动的障碍物时,方案A中移动机器人从动态障碍物运动的方向避障,与理论分析结果相符,这种动态避障策略增加了二者相撞的风险,且为了完成避障动作,移动机器人需向动态障碍物运动方向绕行,增加了路径长度与运行时间;方案B中移动机器人从动态障碍物的运动反方向进行避障,降低了避障风险,且移动机器人路径长度和运行时间较短。2种方案仿真结果对比如表2所示。

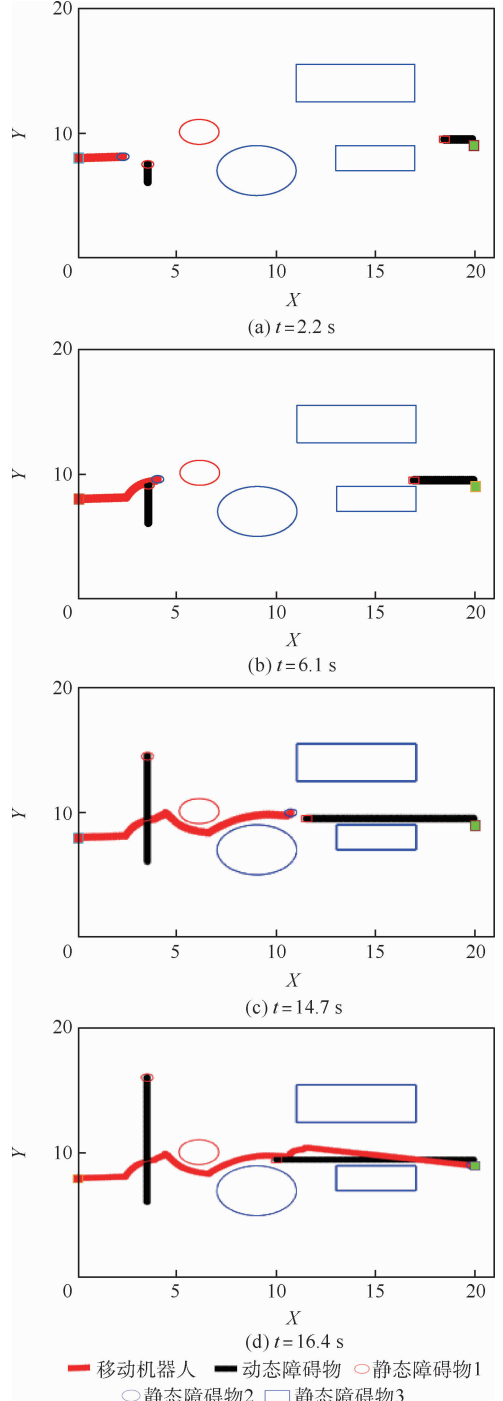


图5 方案A关键时刻仿真曲线

Fig.5 Key moments simulation curves of Plan A

对于静态障碍物,方案 B 与方案 A 相同,避障效果相似。仿真结果表明,引入速度斥力场可有效改善移动机器人的动态避障能力,避障效率显著提高。

在仿真中,2 种方案在不同方向上的受力如图 7 和图 8 所示。可知,由于速度斥力场的存在,X 方向和 Y 方向上,方案 B 受力的幅值相比于方案 A 较小,因此在遇到动态障碍物时轨迹更加平稳,且在进行避障行为时,Y 方向上总受

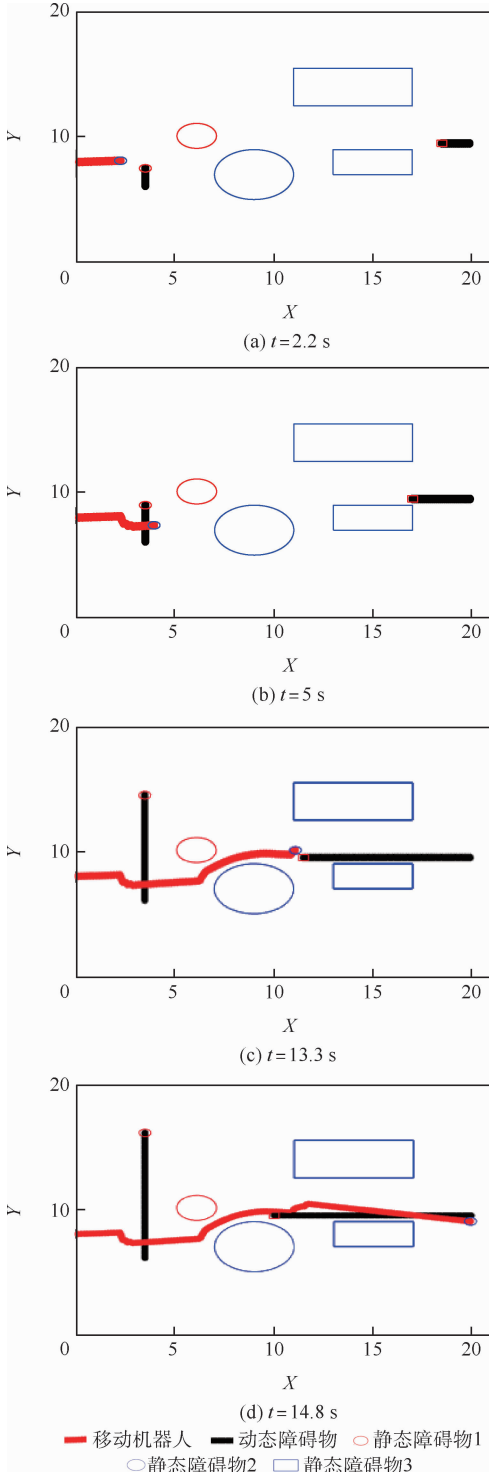


图 6 方案 B 关键时刻仿真曲线

Fig. 6 Key moments simulation curves of Plan B

表 2 仿真对比

Table 2 Simulation comparison

参数	方案 A	方案 B
路径长度	32.8	29.6
仿真运行时间/s	16.4	14.8

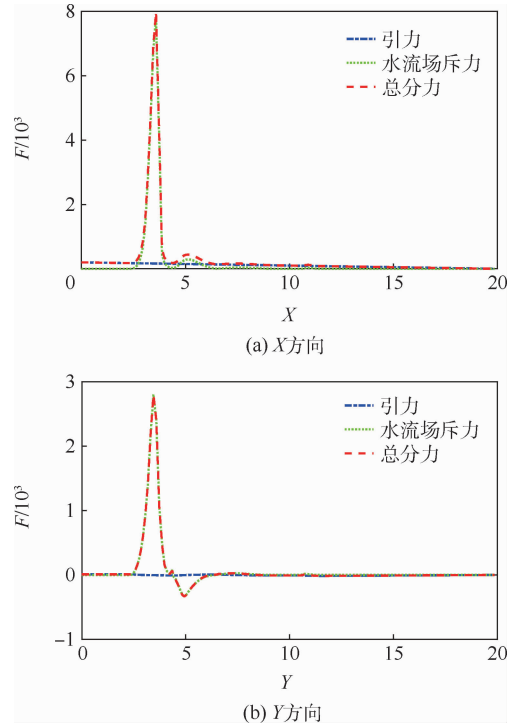


图 7 移动机器人在方案 A 中的仿真受力

Fig. 7 Force diagram during simulation of Plan A of mobile robot

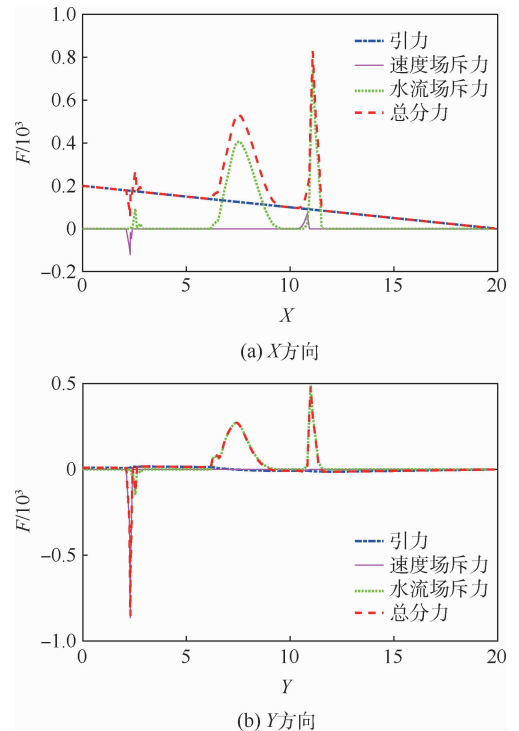


图 8 移动机器人在方案 B 中的仿真受力

Fig. 8 Force diagram during simulation of Plan B of mobile robot

力的方向方向反向,可使得移动机器人从动态障碍物的后方绕过进行避障,降低了相撞的概率。

### 4 实验验证

在室内搭建基于 Optitrack 定位系统的动态避障实验平台,包括 Optitrack 室内定位系统、Xbee 通信系统和 2 辆全向移动机器人,如图 9 所示。实验场地大小设置为 3 m × 3 m,实验中,以一辆全向移动机器人模拟动态障碍物,沿 X 轴方向运动,另一辆移动机器人进行动态避障,按方案 A 和方案 B 所提出的避障策略进行对比实验。

关键时刻的实验场景与同时刻的可视化界面如图 10 和图 11 所示。

方案 A 实验结果如图 10 所示,可见当移动机器人进行动态避障时,会从动态障碍物的前方绕过;方案 B 实验结果如图 11 所示,移动机器人进入到动态障碍物的影响范围开始避障行为(见图 11(b)),受速度斥力场的影响,移动机

器人从动态障碍物的后方绕过以规避障碍物(见图 11(c)),当离开动态障碍物的影响范围后,速度斥力场为 0,移动机器人朝着目标点移动并最终达到目标点(见图 11(d))。实验中,2 种方案在不同方向上的受力如图 12 和图 13 所示。

由实验可知,本文提出的动态避障策略能使移动机器人成功避开动态障碍物,并安全到达目标点,且路径比较平滑,没有出现剧烈抖动,证明了提出的动态避障方法的有效性与实用性。

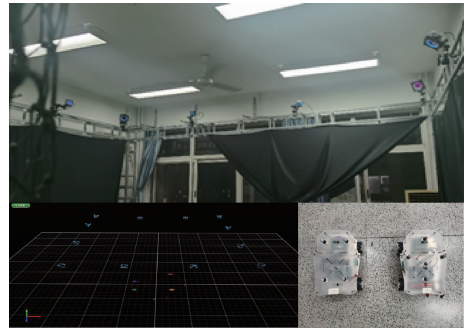


图 9 实验平台  
Fig. 9 Experiment platform

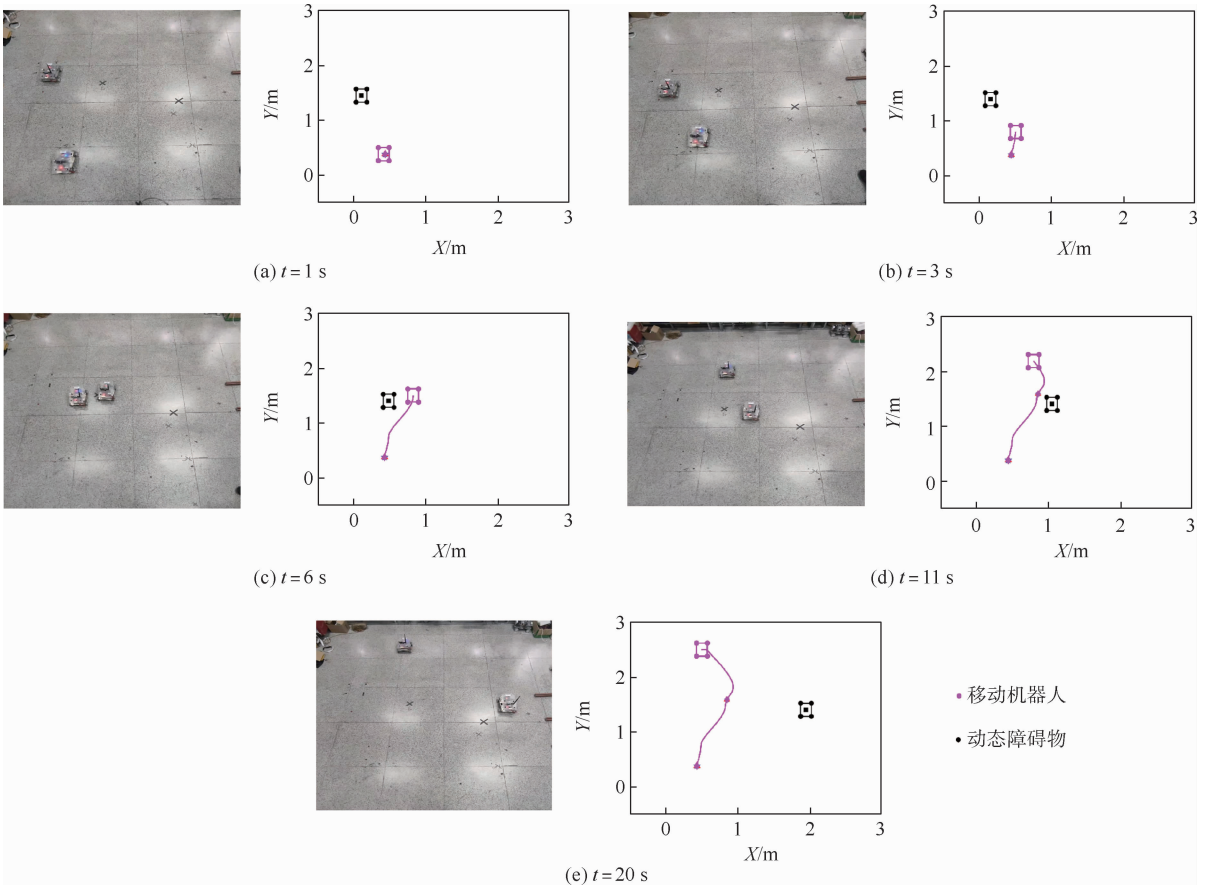


图 10 方案 A 关键时刻实验示意图

Fig. 10 Schematic diagram of experiment results at key moments of Plan A

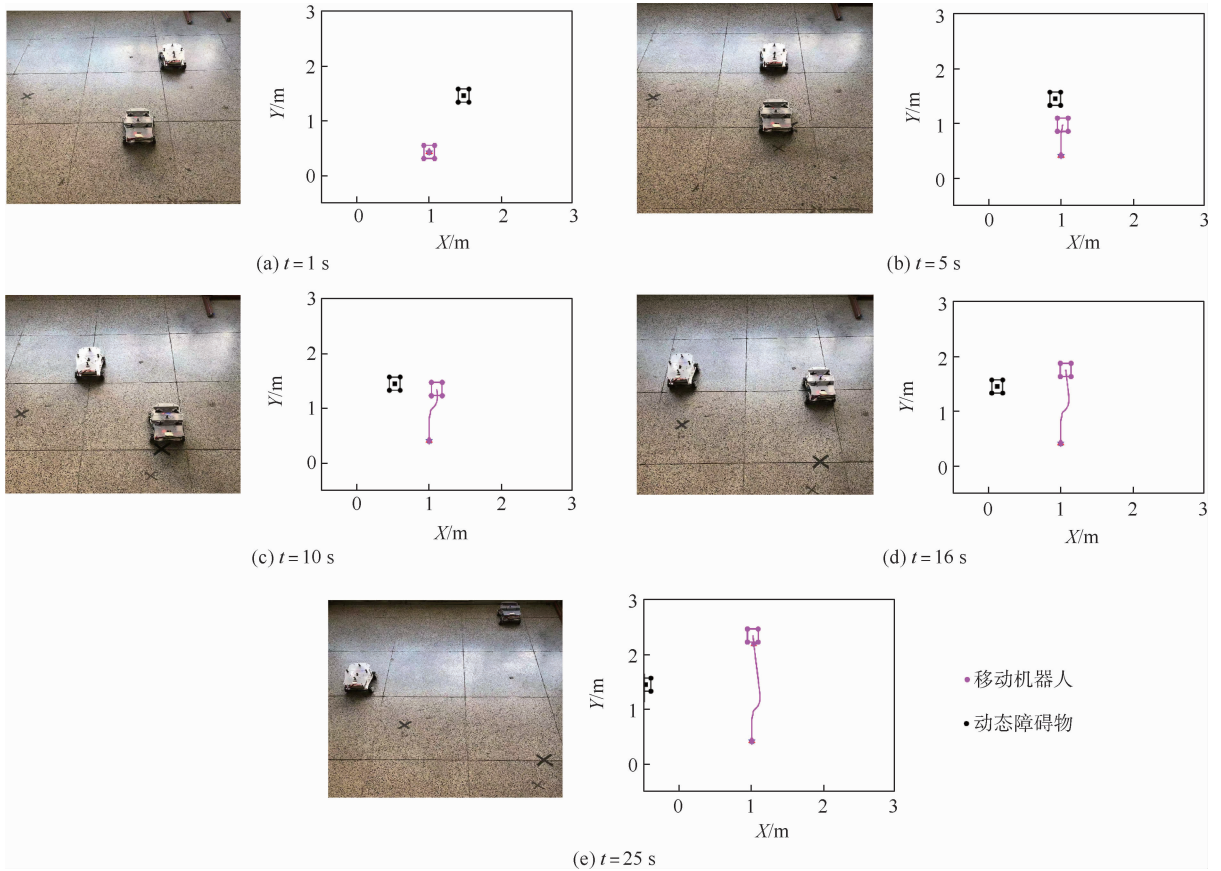


图 11 方案 B 关键时刻实验示意图

Fig. 11 Schematic diagram of experiment results at key moments of Plan B

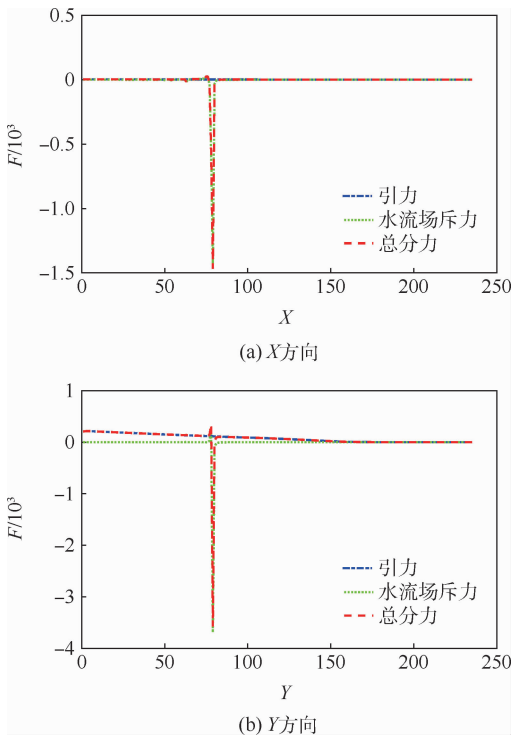


图 12 移动机器人在方案 A 中的实验受力  
Fig. 12 Force diagram during experiment of Plan A of mobile robot

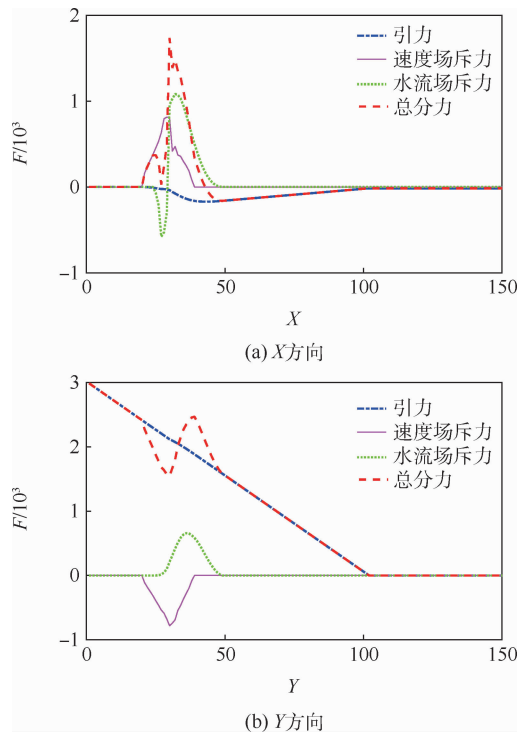


图 13 移动机器人在方案 B 中的实验受力  
Fig. 13 Force diagram during experiment of Plan B of mobile robot

## 5 结 论

1) 针对基于水流场的人工势场法在移动机器人动态避障中引起的牵引问题,通过动态障碍物的位置、速度信息,构建速度斥力场,使得移动机器人能够从动态障碍物后方通过以完成避障,不仅缩短了路径长度和运行时间,而且降低了碰撞概率,提高了避障成功率。

2) 通过仿真和室内移动机器人动态避障实验,实现了有效避障,验证了所提方法的有效性和实用性。

### 参考文献 (References)

- [1] 庞师坤,梁晓锋,李英辉,等.基于零空间行为法的自主水下机器人避障策略[J].上海交通大学学报,2020,54(3):295-304.  
PANG S K, LIANG X F, LI Y H, et al. Collision avoidance strategy for autonomous underwater vehicle based on null-space-based behavioral approach[J]. Journal of Shanghai Jiao Tong University, 2020, 54(3):295-304 (in Chinese).
- [2] 付丽霞,任玉洁,张勇,等.基于改进平滑A\*算法的移动机器人路径规划[J].计算机仿真,2020,37(8):271-276.  
FU L X, REN Y J, ZHANG Y, et al. Path planning of mobile robot based on improved smoothing A\* algorithms[J]. Computer Simulation, 2020, 37(8):271-276 (in Chinese).
- [3] MONTEIL O, OROZCO-ROSAS U, SEPÚLVEDA R. Path planning for mobile robots using bacterial potential field for avoiding static and dynamic obstacles[J]. Expert Systems with Applications, 2015, 42(12):5177-5191.
- [4] 杨明辉,吴垚,张勇,等.室内动态环境下的移动机器人自主避障策略[J].中南大学学报(自然科学版),2019,50(8):1833-1839.  
YANG M H, WU Y, ZHANG Y, et al. Autonomous obstacle avoidance strategy for mobile robots in indoor dynamic environment[J]. Journal of Central South University (Science and Technology), 2019, 50(8):1833-1839 (in Chinese).
- [5] 李卫硕,孙剑,陈伟.基于BP神经网络机器人实时避障算法[J].仪器仪表学报,2019,40(11):204-211.  
LI W S, SUN J, CHEN W. Real-time obstacle avoidance algorithm for robots based on BP neural network[J]. Chinese Journal of Scientific Instrument, 2019, 40(11):204-211 (in Chinese).
- [6] RAMEZANI D A, LEE D J. An end-to-end deep reinforcement learning-based intelligent agent capable of autonomous exploration in unknown environments[J]. Sensors, 2018, 18(10):3575.
- [7] 庞磊,曹志强,喻俊志.基于A\*和TEB融合的行人感知无碰撞跟随方法[J].航空学报,2021,42(4):524909.  
PANG L, CAO Z Q, YU J Z. A pedestrian-aware collision-free following approach for mobile robots based on A\* and TEB[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(4):524909 (in Chinese).
- [8] 槐创锋,郭龙,贾雪艳,等.改进A\*算法与动态窗口法的机器人动态路径规划[J].计算机工程与应用,2021,57(8):244-248.  
HUAI C F, GUO L, JIA X Y, et al. Improved A\* algorithm and dynamic window method for robot dynamic path planning[J]. Computer Engineering and Applications, 2021, 57(8):244-248 (in Chinese).
- [9] 程志,张志安,乐伟扬,等.基于D\* Lite算法的三维路径规划研究[J].传感器与微系统,2020,39(12):71-73.  
CHENG Z, ZHANG Z A, YUE W Y, et al. Research on path planning in 3D terrain based on D\* Lite algorithm[J]. Transducer and Microsystem Technologies, 2020, 39(12):71-73 (in Chinese).
- [10] 徐开放.基于D\* Lite算法的移动机器人路径规划研究[D].哈尔滨:哈尔滨工业大学,2017.  
XU K F. Research on mobile robot path-planning based on D\* Lite algorithm[D]. Harbin: Harbin Institute of Technology, 2017 (in Chinese).
- [11] 张贺,胡越黎,王权,等.基于改进D\*算法的移动机器人路径规划[J].工业控制计算机,2016,29(11):73-74.  
ZHANG H, HU Y L, WANG Q, et al. Path planning of mobile robot based on improved D\* algorithm[J]. Industrial Control Computer, 2016, 29(11):73-74 (in Chinese).
- [12] 朱战霞,靖飒,仲剑飞,等.基于碰撞检测的空间冗余机械臂避障路径规划[J].西北工业大学学报,2020,38(1):183-190.  
ZHU Z X, JING S, ZHONG J F, et al. Obstacle avoidance path planning of space redundant manipulator based on a collision detection algorithm[J]. Journal of Northwestern Polytechnical University, 2020, 38(1):183-190 (in Chinese).
- [13] LIN Y H, HUANG L C, CHEN S Y, et al. The optimal route planning for inspection task of autonomous underwater vehicle composed of MOPSO-based dynamic routing algorithm in currents[J]. Applied Ocean Research, 2018, 75:178-192.
- [14] 魏彤,龙琛.基于改进遗传算法的移动机器人路径规划[J].北京航空航天大学学报,2020,46(4):703-711.  
WEI T, LONG C. Path planning for mobile robot based on improved genetic algorithm[J]. Journal of Beijing University of Aeronautics and Astronautics, 2020, 46(4):703-711 (in Chinese).
- [15] 郭一聪,刘小雄,章卫国,等.基于改进势场法的无人机三维路径规划方法[J].西北工业大学学报,2020,38(5):977-986.  
GUO Y C, LIU X X, ZHANG W G, et al. 3D path planning method for UAV based on improved artificial potential field[J]. Journal of Northwestern Polytechnical University, 2020, 38(5):977-986 (in Chinese).
- [16] 陈麒杰,晋玉强,王陶昱.基于改进人工势场算法的无人机群避障算法研究[J].导航定位与授时,2020,7(6):109-113.  
CHEN Q J, JIN Y Q, WANG T Y. Research on obstacle avoidance algorithm of UAV group based on improved artificial potential field algorithm[J]. Navigation Positioning and Timing, 2020, 7(6):109-113 (in Chinese).
- [17] 崔宝侠,宋佳瑞.未知环境下机器人避障及动态目标追踪[J].沈阳工业大学学报,2018,40(3):292-298.

- CUI B X, SONG J R. Obstacle avoidance and dynamic target tracking of robot in unknown environment[J]. Journal of Shenyang University of Technology, 2018, 40(3):292-298 (in Chinese).
- [18] 杜婉茹,王潇茵,田涛,等. 面向未知环境及动态障碍的人工势场路径规划算法[J]. 计算机科学, 2021, 48(2):250-256.
- DU W R, WANG X Y, TIAN T, et al. Artificial potential field path planning algorithm for unknown environment and dynamic obstacles[J]. Computer Science, 2021, 48(2):250-256 (in Chinese).
- [19] 张大志,刘万辉,缪存孝,等. 全向移动机器人动态避障方法[J]. 北京航空航天大学学报, 2021, 47(6):1115-1123.
- ZHANG D Z, LIU W H, MIAO C X, et al. Dynamic obstacle avoidance method for omnidirectional mobile robots[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(6):1115-1123 (in Chinese).

## Dynamic obstacle avoidance method for mobile robots

ZHANG He<sup>1</sup>, MIAO Cunxiao<sup>1,\*</sup>, TANG Youjun<sup>1</sup>, YAN Xiaoqiang<sup>1</sup>, SHI Yanyang<sup>1</sup>, YU Yuanjin<sup>2</sup>

(1. School of Mechanical Engineering, University of Science and Technology Beijing, Beijing 100083, China;

2. School of Automation, Beijing Institute of Technology, Beijing 100081, China)

**Abstract:** This paper proposes an autonomous dynamic obstacle avoidance method for omnidirectional mobile robot by introducing velocity repulsion field to solve the existing problems, an improvement from water flow field based artificial potential field obstacle avoidance method. This paper presents a detailed analysis of problems of artificial potential field improved by water flow field, such as a too long avoidance path or avoidance failure caused by the mobile robot moving in front of the obstacles. To solve the above problems, the velocity repulsion field, in line with the relative velocity of the mobile robot and the dynamic obstacle, is introduced in artificial potential field obstacle avoidance method based on water flow field. With the omni-directional mobile robot moving in rear of the obstacles, a safe and autonomous dynamic obstacle avoidance is fully realized. The effectiveness and practicability of the autonomous dynamic obstacle avoidance algorithm are verified through simulation and indoor obstacle avoidance experiment.

**Keywords:** mobile robot; artificial potential field; water flow field; velocity repulsion field; dynamic obstacle avoidance

**Received:** 2021-01-02; **Accepted:** 2021-04-20; **Published online:** 2021-05-28 10:55

**URL:** [kns.cnki.net/kcms/detail/11.2625.V.20210527.1709.001.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210527.1709.001.html)

**Foundation items:** National Key R & D Program of China (2018YFC0810500); National Natural Science Foundation of China (60277005); Beijing Institute of Technology Research Fund Program for Young Scholars (2020CX04104)

\* **Corresponding author.** E-mail: miaocx@ustb.edu.cn

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2020.0696

# 基于 VMD-MPE 的开关磁阻电机功率 变换器故障诊断

张竞文<sup>1</sup>, 熊立新<sup>1,2,\*</sup>, 马宏昌<sup>2</sup>, 边敦新<sup>1</sup>

(1. 山东理工大学 电气与电子工程学院, 淄博 255000; 2. 山东科汇电力自动化股份有限公司, 淄博 255087)

**摘 要:** 故障诊断是提升开关磁阻电机(SRM)调速系统可靠性的重要技术。针对功率变换器故障信号非线性不稳定、有效信息易被噪声掩盖的特点,提出了一种新的故障特征提取方法。对直流母线电流进行变分模态分解,得到若干本征模态分量,取多尺度有效模态分量排列熵平均值作为特征向量,输入支持向量机分类器进行故障识别。为验证所提方法的可行性,建立仿真模型,与传统的小波分析等故障诊断方法进行对比;搭建了开关磁阻电机实验台架,测试了开路、短路故障状态。仿真和实验结果表明:所提方法可减小噪声影响,提高故障识别准确率。

**关键词:** 开关磁阻电机(SRM); 功率变换器; 故障诊断; 变分模态分解; 多尺度排列熵  
**中图分类号:** V242.44; TM352

**文献标志码:** A **文章编号:** 1001-5965(2022)06-1022-08

开关磁阻电机(switched reluctance motor, SRM)是近30年发展起来的新型调速系统,已广泛应用于多个领域<sup>[1]</sup>。功率变换器作为调速系统的核心,主要由电力电子器件构成,经常处于高频工作状态或恶劣的工作环境中,若出现电气故障时不及时处理,会导致系统故障进一步恶化,产生不可修复的损伤,因此,SRM功率变换器故障的准确诊断具有重要意义<sup>[2]</sup>。

现有文献提出的故障诊断方法主要有3类,分别基于逻辑判断<sup>[3-4]</sup>、硬件<sup>[5-6]</sup>或算法<sup>[7-9]</sup>实现。前2种方法通过检测一定位置的状态量进行诊断,需要安装多个额外的传感器,不仅增加了成本,也使得系统更加复杂;基于算法的故障诊断是对功率变换器的状态信号进行采样再通过一系列数学变换来实现,具有覆盖范围广、成本低的优点,但算法选择困难,易受噪声干扰。文献[10]利用小波变换对电机三相电流进行分析,提取低

频小波信号能量值作为故障特征,但小波变换方法需选择合适的基函数与原始信号进行运算,无法自适应变化;文献[11]将经验模态分解(empirical mode decomposition, EMD)引入SRM功率变换器故障诊断,计算奇异值能量和矩阵能量熵参数作为特征向量,而自适应EMD高度依赖极值包络插值,运算数据量较大,易产生模态混叠现象。变分模态分解(variational mode decomposition, VMD)方法理论基础坚实,对故障信号具有更好的鲁棒性,已在机械故障诊断、预测及降噪中得到应用<sup>[12-13]</sup>,但尚未应用到SRM功率变换器故障诊断中。

本文研究了将VMD用于SRM功率变换器故障诊断的方法,并引入多尺度排列熵(multi-scale permutation entropy, MPE)规则以更好地量化非线性不稳定故障信号在不同尺度下的故障信息,提取故障特征向量后应用支持向量机(support vector machine, SVM)实现故障诊断。建立了MAT-

收稿日期: 2020-12-16; 录用日期: 2021-03-13; 网络出版时间: 2021-04-01 17:21

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20210401.1459.002.html

基金项目: 国家重点研发计划(2017YFB0902800)

\* 通信作者。E-mail: xionglx@sdut.edu.cn

**引用格式:** 张竞文, 熊立新, 马宏昌, 等. 基于VMD-MPE的开关磁阻电机功率变换器故障诊断[J]. 北京航空航天大学学报, 2022, 48(6): 1022-1029. ZHANG J W, XIONG L X, MA H C, et al. Fault diagnosis of switched reluctance motor power converter based on VMD-MPE[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(6): 1022-1029 (in Chinese).

LAB/Simulink 模型,并与其他方法进行仿真对比,使用一台 45 kW SRM 搭建了开路、短路故障模拟台架,并进行了实际实验。仿真与实验结果表明,本文方法可以有效降低噪声对识别结果的干扰,准确识别功率变换器的故障类型。

### 1 功率变换器及其故障类型

本文以常见的三相不对称半桥功率变换器为研究对象,如图 1 所示。

以 A 相为例,S1、S2 分别为 A 相斩波管和位置导通管。A 相导通期间,S2 导通,S1 斩波,电源通过斩波控制给绕组供电;关断阶段,S1、S2 关断,电流经 D1、D2 实现反压续流,能量回馈电源。正压励磁、零压续流和反压回馈通路如图 2 所示。

功率变换器故障主要是功率开关管单管故障,包括斩波管、位置导通管的开路和短路故障 4 种故障类型<sup>[14]</sup>,对应的电流通路如图 3 所示。

图 3(a)为 A 相位置导通管 S2 短路,此时基本不影响电机相绕组导通励磁,但在关断期间无法对相绕组施加反压以强制续流电流快速回馈电源,易产生制动转矩。图 3(b)为斩波管 S1 短路,此时导通期间绕组电压无法斩波控制,关断期间无法反压强制续流,且电源直接加在绕组导致相电流过大。图 3(c)、(d)为 A 相开路故障,此时 A 相无法导通,系统进入缺相运行状态,产生转矩脉动。

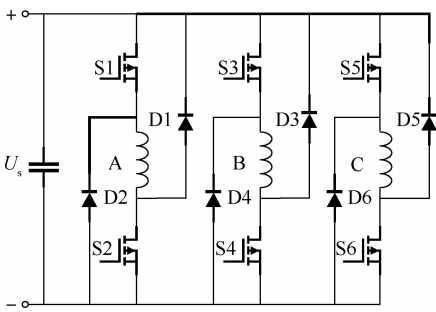


图 1 不对称半桥功率变换器拓扑结构

Fig. 1 Asymmetric half-bridge power converter topology

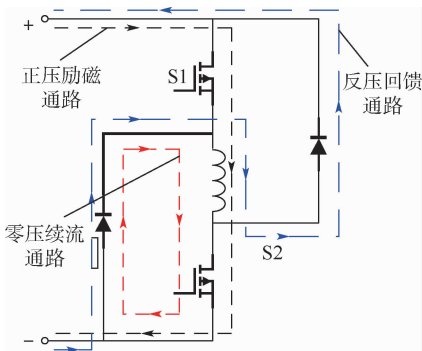


图 2 正常导通 A 相电流通路

Fig. 2 Current path of phase A under normal conduction

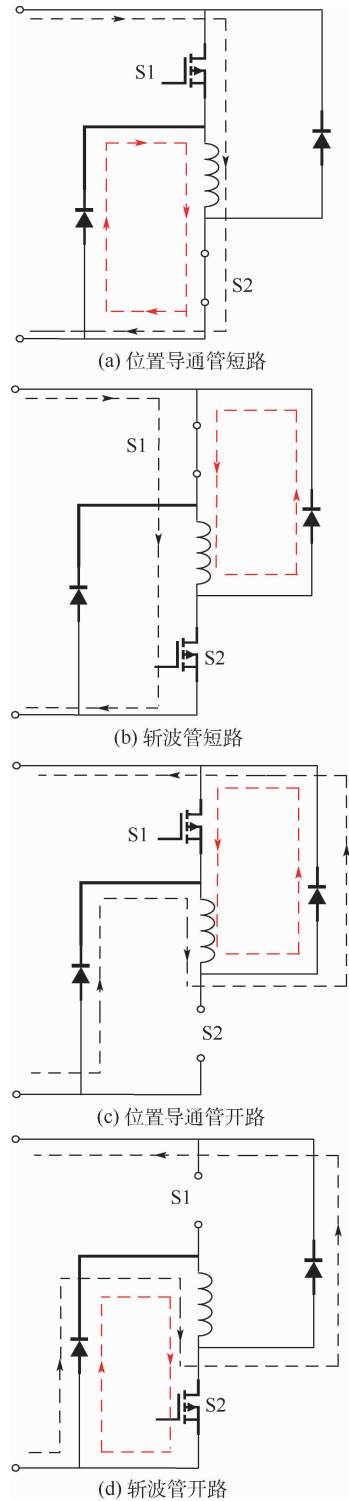


图 3 功率变换器单相故障

Fig. 3 Single phase fault of power converter

### 2 故障诊断方案

#### 2.1 变分模态分解

VMD 的基本思想是将故障产生的非线性不稳定信号分解为一组包含故障信息的模态分量,具体实施策略是:根据设置的模态分量个数  $K$  对原始信号  $f(x)$  进行分解,得到估计带宽之和最小的  $K$  个本征模态分量 (intrinsic mode function,

IMF),  $K$  为预设模态个数。通过希尔伯特变换得到各模态的解析信号单边频谱并移到基带, 对解析信号梯度的平方范数  $L^2$  估计带宽, 由此得到受约束的变分问题<sup>[15]</sup>:

$$\begin{aligned} \min_{\{\mathbf{u}_k\}, \{\omega_k\}} & \left\{ \sum_k \left\| \partial_t \left[ \left( \delta(t) + \frac{j}{\pi t} \right) * \mathbf{u}_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \\ \text{s. t.} & \sum_{k=1}^K \mathbf{u}_k = f(t) \end{aligned} \quad (1)$$

式中:  $\{\mathbf{u}_k\} = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_K\}$  为分解得到的  $K$  个模态分量;  $\{\omega_k\} = \{\omega_1, \omega_2, \dots, \omega_K\}$  为各模态的中心频率;  $k = 1, 2, \dots, K$ ;  $\delta(t)$  为狄拉克函数;  $\partial_t$  表示对  $t$  求偏导; “\*” 表示卷积运算;  $f(t)$  为原始信号。

为求最优解问题, 引入二次惩罚因子  $\alpha$  和拉格朗日乘数算子  $\lambda$ , 得到增广拉格朗日函数:

$$\begin{aligned} L(\{\mathbf{u}_k\}, \{\omega_k\}, \lambda) = & \alpha \sum_k \left\| \partial_t \left[ \left( \delta(t) + \frac{j}{\pi t} \right) * \mathbf{u}_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \\ & \left\| f(t) - \sum_k \mathbf{u}_k(t) \right\|_2^2 + \langle \lambda(t), f(t) - \sum_k \mathbf{u}_k(t) \rangle \end{aligned} \quad (2)$$

式中:  $\langle \rangle$  表示内积计算。

VMD 算法流程如下:

**步骤 1** 参数初始化, 使  $\{\mathbf{u}_k^1\}$ 、 $\{\omega_k^1\}$ 、 $\lambda^1$ 、 $n$  初始值为 0。

**步骤 2**  $n = n + 1$ , 更新  $\mathbf{u}_k$  和  $\omega_k$ , 从 1 循环到  $K$ 。

$$\hat{\mathbf{u}}_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i \neq k} \hat{\mathbf{u}}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k)^2} \quad (3)$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{\mathbf{u}}_k(\omega)|^2 d\omega}{\int_0^\infty |\hat{\mathbf{u}}_k(\omega)|^2 d\omega} \quad (4)$$

**步骤 3** 更新  $\lambda$ 。

$$\hat{\lambda}^{n+1}(\omega) = \hat{\lambda}^n(\omega) + \sigma \left( \hat{f}(\omega) - \sum_k \hat{\mathbf{u}}_k^{n+1}(\omega) \right) \quad (5)$$

**步骤 4** 给定  $\varepsilon > 0$ , 若满足式 (6), 则迭代终止。

$$\sum_k \|\hat{\mathbf{u}}_k^{n+1} - \hat{\mathbf{u}}_k^n\|_2^2 / \|\hat{\mathbf{u}}_k^n\|_2^2 < \varepsilon \quad (6)$$

式中:  $\sigma$  为保真度系数;  $\varepsilon$  为判别精度。

## 2.2 多尺度排列熵

MPE 可对时间序列进行多尺度粗粒化, 通过计算不同尺度的排列熵 (permutation entropy, PE), 从多个尺度反映故障信号的复杂度, 易于提取故障特征向量<sup>[16]</sup>。对给定长度为  $N$  的离散时间序列进行粗粒化处理, 得到如下序列:

$$\mathbf{y}_j^{(s)} = \frac{1}{s} \sum_{i=(j-1)s+1}^j \mathbf{x}_i \quad j = 1, 2, \dots, [N/s] \quad (7)$$

式中:  $s$  为尺度因子;  $[N/s]$  表示对  $\mathbf{y}_j^{(s)}$  取整。

对粗粒化序列进行重构:

$$\mathbf{Y}_l^{(s)} = \{\mathbf{y}_l^{(s)}, \mathbf{y}_{l+\tau}^{(s)}, \dots, \mathbf{y}_{l+(m-1)\tau}^{(s)}\} \quad (8)$$

式中:  $l = 1, 2, \dots, N - (m-1)\tau$ ,  $\tau$  为延迟时间,  $m$  为嵌入维数。

将  $\mathbf{Y}_l^{(s)}$  按升序排列, 对任意  $l$  都有唯一符号序列  $\mathbf{S}(r) = \{l_1, l_2, \dots, l_m\}$ ,  $r = 1, 2, \dots, R$ , 且  $R \leq m!$ , 计算每种符号序列出现的概率  $P_r$ 。

定义 MPE 为

$$H_p(m) = - \sum_{r=1}^R P_r \ln P_r \quad (9)$$

由式 (9) 可知,  $H_p(m)$  最大值为  $\ln(m!)$ , 将  $H_p(m)$  进行归一化得

$$H_p = H_p(m) / \ln(m!) \quad (10)$$

## 2.3 故障诊断流程

针对 SRM 功率变换器故障信号非线性不稳定且易受噪声干扰的特性, 本文采用 VMD-MPE 法提取故障特征向量, 应用 SVM 分类器对特征向量进行识别分类实现故障诊断。具体步骤如下:

**步骤 1** 对采集的各状态特征信号进行变分模态分解, 得到  $K$  个本征模态分量 IMF。

**步骤 2** 计算 IMF 分量与原始信号的相关系数, 剔除弱相关分量, 提取各尺度的有效 IMF 分量排列熵平均值作为特征向量。

**步骤 3** 将得到的特征向量分为 2 组: 一组作为已知训练样本, 另一组作为测试样本。将训练样本输入 SVM 进行训练, 得到 SVM 分类器。

**步骤 4** 将测试样本输入训练好的 SVM 分类器, 通过分类器的诊断结果来识别功率变换器故障类型。

## 3 建模仿真

本文应用 MATLAB/Simulink 建立了 SRM 调速闭环系统仿真模型, 采用阶跃信号和逻辑关系构造了开路与短路故障发生模块, 分别采集功率变换器正常工作、斩波管短路、位置导通管短路和开路 4 种状态下的直流母线电流信号, 验证 VMD-MPE 法的有效性。

### 3.1 参数确定

变分模态分解前需确定模态分量的个数  $K$ ,  $K$  值过小会滤除原始信号的重要特征信息,  $K$  值过大会使相邻模态分量的中心频率相近, 造成模态混叠, 影响诊断结果<sup>[17]</sup>。选择正常状态下直流母线电流进行变分模态分解, 不同  $K$  值对应的中心频率如表 1 所示。

可见,  $K$  为 9 时有 2 个相近的中心频率 1.66 kHz 和 1.54 kHz, 对应的模态分量可能会出现模态混叠, 故将  $K$  设为 8 并对正常工作的母线电流进行变分模态分解, 得到的模态分量从高频到低频排列如图 4 所示。

进一步计算 8 个模态分量与原始信号的相关系数  $\gamma$ , 得到表 2。剔除相关系数小于 0.1 的 IMF1, 选择 IMF2 ~ IMF8 为有效模态分量进行特征提取。

计算 MPE 需设定嵌入维数  $m$ 、延迟时间  $\tau$  和尺度因子  $s$ 。Christoph 和 Bernd<sup>[18]</sup> 建议嵌入维数  $m$  取值范围为 3 ~ 7。若  $m$  取值过小, 则 MPE 不能准确反映信号突变性, 若取值过大, 则会导致时间序列均匀化, 无法反映时间序列的细节信息。本文对  $m$ 、 $\tau$ 、 $s$  进行了仿真, 如图 5 和图 6 所示, 综

表 1 不同  $K$  值对应的中心频率

Table 1 Center frequency of different  $K$  values

$K$ 值	中心频率/kHz
5	1.33, 0.54, 0.26, 0.08, 0.002
6	1.61, 0.90, 0.47, 0.25, 0.08, 0.002
7	1.99, 1.34, 0.81, 0.45, 0.25, 0.08, 0.002
8	2.13, 1.54, 1.04, 0.67, 0.43, 0.24, 0.08, 0.002
9	2.36, 1.66, 1.54, 0.96, 0.63, 0.42, 0.24, 0.08, 0.002

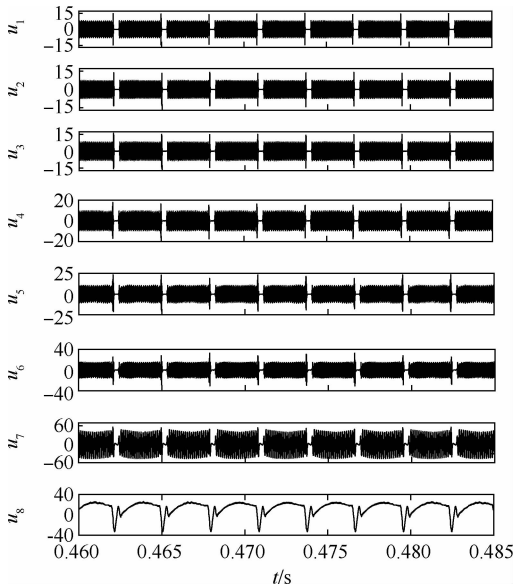


图 4 变分模态分解结果

Fig. 4 Variational mode decomposition results

表 2 模态分量对应的相关系数

Table 2 Correlation coefficients of modal components

IMF	$\gamma_i$
IMF1	0.088 6
IMF2	0.102 7
IMF3	0.126 0
IMF4	0.159 3
IMF5	0.208 5
IMF6	0.310 5
IMF7	0.813 3
IMF8	0.460 6

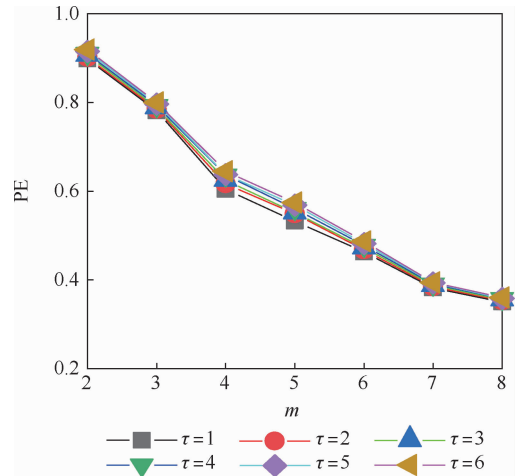


图 5 采样信号在不同延迟时间下的排列熵

Fig. 5 PE of sampled signals under different time delays

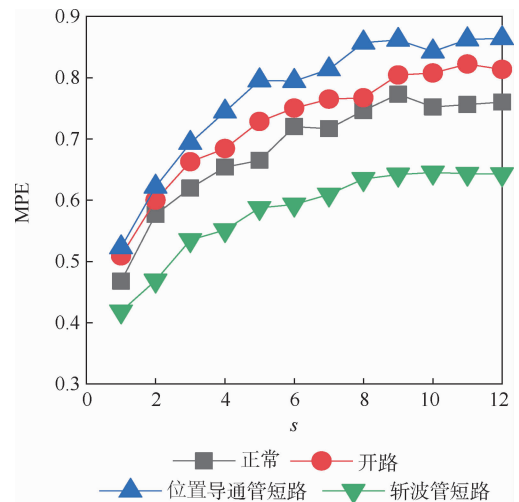


图 6 尺度因子的影响

Fig. 6 Influence of scale factor

合考虑后将嵌入维数  $m$  设为 6, 时间延迟  $\tau$  设为 1, 尺度因子  $s$  设为 10。

### 3.2 故障识别

分别采集 SRM 功率变换器正常工作、斩波管短路、位置导通管短路和开路 4 种状态下的直流母线电流样本各 30 组, 对每组样本进行变分模态分解提取有效模态分量, 取各尺度下的有效模态分量排列熵平均值作为特征向量。每种状态的特征向量随机取 25 组数据作为训练样本训练 SVM 分类器, 其余作为测试样本。正常工作、开路、位置导通管短路和斩波管短路 4 种状态的类别标签为“1”、“2”、“3”和“4”。SVM 训练过程中通过交叉验证法选择最优惩罚系数  $c$  和核函数中的  $g$ , 本文选择  $c$  为 0.353 6,  $g$  为 11.313 7, SVM 对测试样本分类结果如图 7 所示。可以看出, SVM 分类结果准确率为 100%。

### 3.3 方法对比

在同样的数据样本下, 分别使用 EMD 和平

稳小波变换 (stationary wavelet transform, SWT) 结合 MPE 的特征提取方法进行仿真, 结果如表 3 所示。可以看出, VMD-MPE 法识别准确率最高。

EMD 和 SWT 的故障分类结果如图 8 所示。可以看出, EMD 把部分开路故障识别为斩波管短路, 而 SWT 把部分斩波管短路识别为正常工作状态, 虽识别准确率较 EMD 高, 但斩波管短路危害较大, 不及时处理会产生严重后果。

为进一步验证诊断方法在噪声环境下的鲁棒性, 本文对 SRM 功率变换器的采样信号添加不同的高斯白噪声, 再使用 3 种方法对样本进行分类识别, 其结果如表 4 所示。可以看出, VMD-MPE 法抗干扰能力强, 对噪声有良好的鲁棒性。

为比较各方法在故障诊断实时性方面的差异, 根据方法数据需求分别设定 EMD-MPE 与 VMD-MPE 的采样时长为 20 ms, SWT-MPE 的采样时长为 10 ms, 各方法故障诊断时间曲线如图 9 所示。

从图 9 可以看出, EMD-MPE 法故障诊断用时最长, 平均为 0.050 6 s; SWT-MPE 法因算法简单用时最短, 平均为 0.012 2 s; VMD-MPE 法居中, 平均为 0.022 5 s, 但故障识别准确率最高。

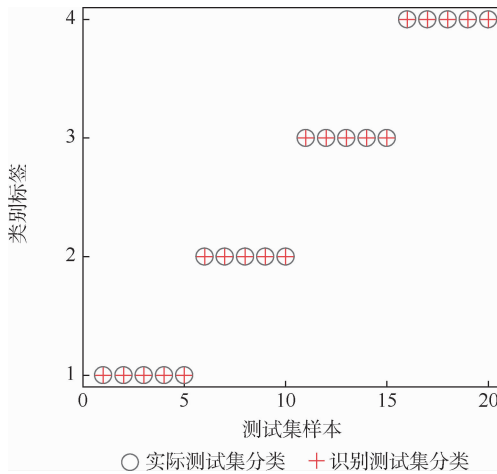


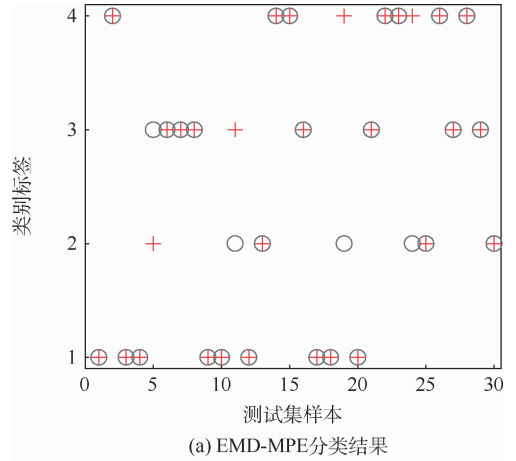
图 7 SVM 分类结果

Fig. 7 SVM classification results

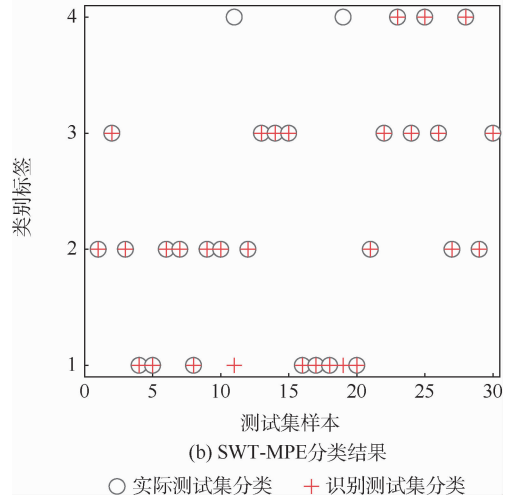
表 3 不同诊断方法的识别准确率

Table 3 Recognition accuracy rate of different diagnostic methods

诊断方法	识别准确率/%
EMD-MPE	86.67
SWT-MPE	93.33
VMD-MPE	100



(a) EMD-MPE分类结果



(b) SWT-MPE分类结果

图 8 EMD-MPE 和 SWT-MPE 分类结果

Fig. 8 EMD-MPE and SWT-MPE classification results

表 4 噪声环境下分类结果

Table 4 Classification result under noisy environment

诊断方法	识别准确率/%
EMD-MPE	76.67
SWT-MPE	90
VMD-MPE	96.67

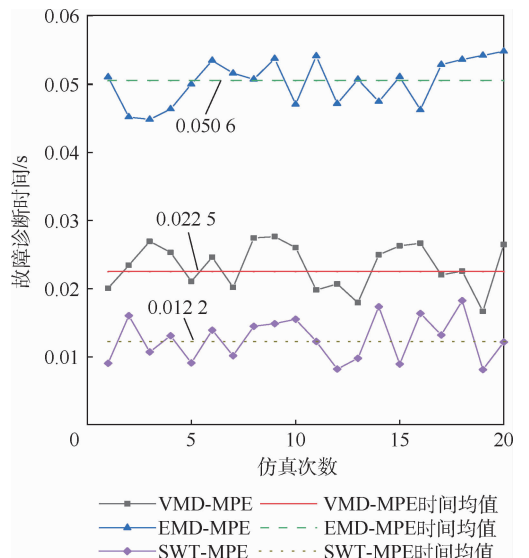


图 9 故障诊断时间对比

Fig. 9 Comparison of fault diagnosis time

### 4 台架实验

使用一台 45 kW、1 500 r/min SRM 搭建了实验平台,进行 A 相 IGBT 故障模拟实验,图 10 为实验系统框图。采用不对称半桥功率变换器,200 V 直流电源驱动,DSP 芯片 TMS320F2812 配合外围逻辑电路控制 IGBT 高频通断,实现电流斩波闭环调速,如图 11 所示。

测试时,将开通角和关断角设置为  $0^\circ$  和  $28^\circ$ ,采样频率为 500 kHz,采集额定转速 1 500 r/min 正常状态、开路状态、上管短路状态时电流波形如图 12 所示。图中: $I_d$  为直流母线电流, $I_a$ 、 $I_b$ 、 $I_c$  为对应 SRM 绕组电流。

提取尺度因子为 5 的各尺度 VMD 有效分量平均排列熵值如图 13 所示。可知,相同运行状态的熵值随转速变化不大,且故障后的熵值明显增大。

将实验得到的特征向量输入 SVM 进行训练和测试,其结果如表 5 所示。可知,采用 EMD-MPE 法时,对开路故障的识别准确率较低,综合识别准确率为 80%。采用 SWT-MPE 法时,识别正常和开路运行状态的准确率较高,但上管短路故障识别准确率较低,综合识别准确率为 91.67%。VMD-MPE 法识别准确率明显高于另外 2 种方法,可有效提高 SRM 功率变换器故障诊断的准确率。

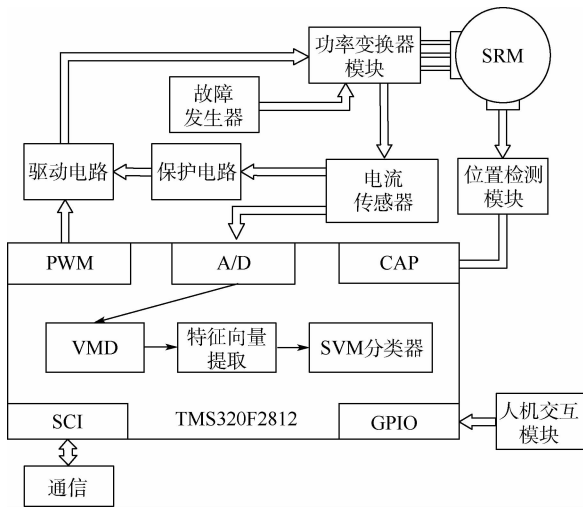


图 10 实验系统框图

Fig. 10 Experimental system block diagram

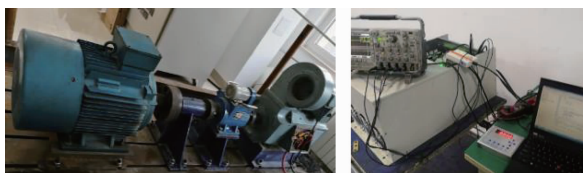


图 11 实验系统

Fig. 11 Experiment system

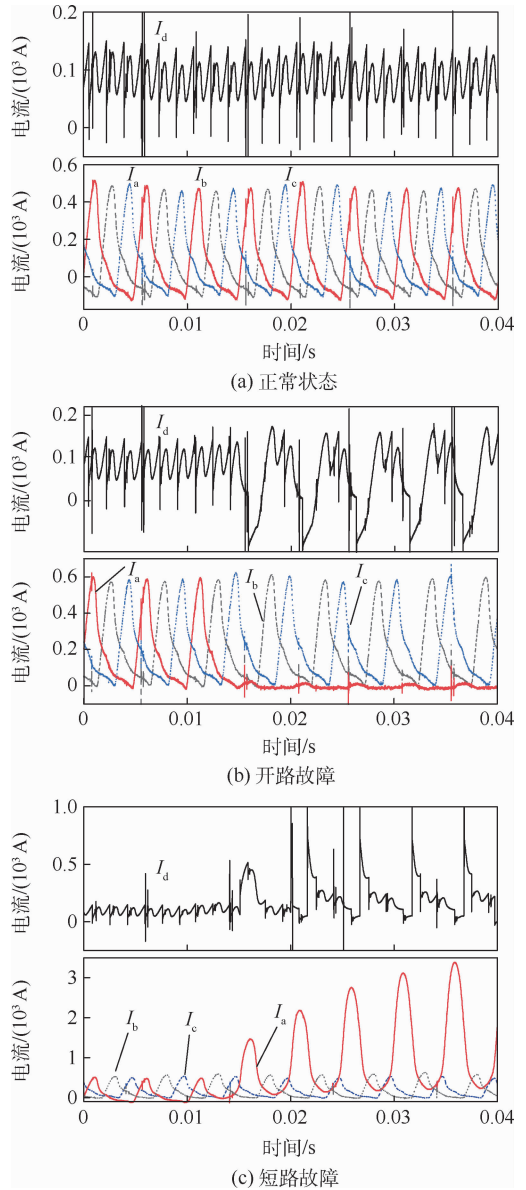


图 12 实测电流

Fig. 12 Measured current

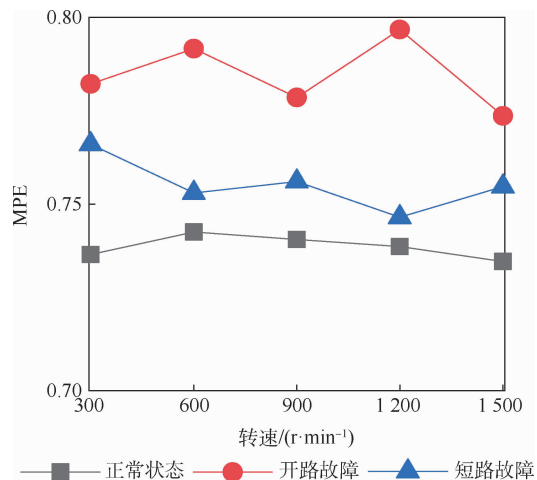


图 13 直流母线电流有效分量平均排列熵

Fig. 13 Average permutation entropy of effective component of DC bus current

表5 不同诊断方法分类结果

Table 5 Classification results of different diagnostic methods

诊断方法	故障类型	测试样本数	识别结果			综合识别准确率/%
			正常	开路	短路	
EMD-MPE	正常	20	18	1	1	80
	开路	20	2	14	4	
	短路	20	1	3	16	
SWT-MPE	正常	20	19	1		91.67
	开路	20	1	19		
	短路	20	3		17	
VMD-MPE	正常	20	20			100
	开路	20		20		
	短路	20			20	

本文样机中 TMS320F2812 工作频率为 120 MHz,可在 2~3 ms 内处理电流数据并产生故障诊断结果,计及采样电流时长 20 ms,可知该方法在线辨识最久需要 23 ms,实时性满足工程需要。

## 5 结论

本文提出了一种基于 VMD-MPE 的 SRM 功率变换器故障诊断方法。

1) 采集功率变换器直流母线电流进行变分模态分解,取多尺度有效模态分量排列熵平均值作为特征向量,输入 SVM 分类器可有效识别故障。

2) 仿真和实际台架实验结果均表明本文方法与传统的 EMD、SWT 方法相比,不需要选择小波基函数,抗噪声能力强,故障诊断准确率高,适合于工程应用。

## 参考文献 (References)

- [1] 管旻珺,朱妹妹,张磊,等.六相开关磁阻电机  $N+2$  功率变换器断路故障状态下的中点电流特性分析[J].电工技术学报,2019,39(25):407-416.  
GUAN M J, ZHU S S, ZHANG L, et al. Analysis of midpoint current characteristics of six-phase switched reluctance motor  $N+2$  power converter under open circuit fault[J]. Transactions of China Electrotechnical Society, 2019, 39(25):407-416 (in Chinese).
- [2] FERNÃO P V, CORDEIRO A, FOITO D, et al. A multilevel fault-tolerant power converter for a switched reluctance machine drive[J]. IEEE Access, 2020, 8:21917-21931.
- [3] CHEN H, HAN G Q, SHI X Q, et al. Phase current digital analysis of power converter for freewheeling diode fault diagnosis on switched reluctance motor drive[J]. IEEE Transactions on Industrial Electronics, 2019, 66(8):6613-6624.
- [4] 肖丽,范书瑞,王博文,等.基于电流分析的开关磁阻电机功率变换器故障检测方法[J].电机与控制学报,2018,22(4):67-74.  
XIAO L, FAN S R, WANG B W, et al. A fault detection method for switched reluctance motor power converter based on current analysis[J]. Electric Machines and Control, 2018, 22(4):67-74 (in Chinese).
- [5] RO H S, KIM D H, JEONG H G, et al. Tolerant control for power transistor faults in switched reluctance motor drives[J]. IEEE Transactions on Industry Applications, 2015, 51(4):3187-3197.
- [6] 靳志欣,宋建成,曲兵妮,等.开关磁阻电机驱动系统功率变换器的故障检测[J].煤炭学报,2014,39(1):186-191.  
JIN Z X, SONG J C, QU B N, et al. Fault detection of power converter in switched reluctance motor drive system[J]. Journal of China Coal Society, 2014, 39(1):186-191 (in Chinese).
- [7] 赵刚,周波.基于小波包分析的双凸极无刷直流发电机系统故障诊断[J].中国电机工程学报,2008,28(8):121-126.  
ZHAO G, ZHOU B. Fault diagnosis of doubly salient brushless DC generator system based on wavelet packet analysis[J]. Chinese Society for Electrical Engineering, 2008, 28(8):121-126 (in Chinese).
- [8] GAN C, WU J H, YANG S Y, et al. Fault diagnosis scheme for open-circuit faults in switched reluctance motor drives using fast Fourier transform algorithm with bus current detection[J]. IET Power Electronics, 2016, 9(1):20-30.
- [9] 甘醇,吴建华,杨仕友,等.基于小波包能量分析的开关磁阻电机功率变换器故障诊断[J].中国电机工程学报,2014,34(9):1415-1422.  
GAN C, WU J H, YANG S Y, et al. Fault diagnosis of switched reluctance motor power converter based on wavelet packet energy analysis[J]. Chinese Society for Electrical Engineering, 2014, 34(9):1415-1422 (in Chinese).
- [10] 杨欣荣,蒋林,王婧林,等.基于小波变换的无刷直流电机逆变器故障诊断[J].电测与仪表,2017,54(5):113-118.  
YANG X R, JIANG L, WANG J L, et al. Fault diagnosis of brushless DC motor inverter based on wavelet transform[J]. Electrical Measurement and Instrumentation, 2017, 54(5):113-118 (in Chinese).
- [11] YANG R K, MA R Q, BAI P. Application of HHT in SRM fault feature extraction[C]//2017 Electronic and Automation Control Conference. Piscataway:IEEE Press, 2017:57-63.
- [12] 田书,康智慧.基于改进变分模态分解和 SVM 的断路器机械故障振动分析[J].振动与冲击,2019,38(23):90-95.

- TIAN S, KANG Z H. Vibration analysis of circuit breaker mechanical fault based on improved variational mode decomposition and SVM[J]. *Vibration and Shock*, 2019, 38(23):90-95 (in Chinese).
- [13] 张杏莉, 卢新明, 贾瑞生, 等. 基于变分模态分解及能量熵的微震信号降噪方法[J]. *煤炭学报*, 2018, 43(2):356-363. ZHANG X L, LU X M, JIA R S, et al. Microseismic signal denoising method based on variational modal decomposition and energy entropy[J]. *Journal of China Coal Society*, 2018, 43(2):356-363 (in Chinese).
- [14] HU Y H, GAN C, CAO W P, et al. Flexible fault-tolerant topology for switched reluctance motor drives[J]. *IEEE Transactions on Power Electronics*, 2016, 31(6):4654-4668.
- [15] DRAGOMIRETSKIY K, ZOSSO D. Variational mode decomposition[J]. *IEEE Transactions on Signal Processing*, 2014, 62(3):531-544.
- [16] AZIZ W, ARIF M. Multiscale permutation entropy of physiological time series[C]//2005 Pakistan Section Multitopic Conference. Piscataway: IEEE Press, 2005:1-66.
- [17] 潘震, 黄国勇, 吴漫. 基于参数优化VMD和增强多尺度排列熵的单向阀故障诊断[J]. *振动与冲击*, 2020, 39(15):118-125. PAN Z, HUANG G Y, WU M. Fault diagnosis of one-way valve based on parameter optimization VMD and enhanced multiscale permutation entropy[J]. *Vibration and Shock*, 2020, 39(15):118-125 (in Chinese).
- [18] CHRISTOPH B, BERND P. Permutation entropy: A natural complexity measure for time series[J]. *Physical Review Letters*, 2002, 88(17):174102.

## Fault diagnosis of switched reluctance motor power converter based on VMD-MPE

ZHANG Jingwen<sup>1</sup>, XIONG Lixin<sup>1,2,\*</sup>, MA Hongchang<sup>2</sup>, BIAN Dunxin<sup>1</sup>

(1. School of Electrical and Electronic Engineering, Shandong University of Technology, Zibo 255000, China;

2. Shandong Kehui Electric Power Automation Co., Ltd., Zibo 255087, China)

**Abstract:** Fault diagnosis is an important technology to improve the reliability of the switched reluctance motor (SRM) speed control system. To address the non-linear and unstable fault signal of the switched reluctance motor power converter, and the problem that effective information is easily covered by noise, a new fault feature extraction method is proposed. The DC bus current is subjected to variational mode decomposition to obtain several intrinsic mode functions. The average value of the permutation entropy of the multi-scale effective modal components is taken as the feature vector, and is input into the support vector machine classifier for fault identification. In order to verify the feasibility of the proposed algorithm, a simulation model was established and compared with traditional fault diagnosis algorithms such as wavelet analysis; meanwhile, a switched reluctance motor experiment bench was built to test the open circuit and short circuit fault states. The simulation and experimental results show that the method proposed in this paper can reduce the influence of noise and improve the accuracy of fault identification rate.

**Keywords:** switched reluctance motor (SRM); power converter; fault diagnosis; variational mode decomposition; multiscale permutation entropy

Received: 2020-12-16; Accepted: 2021-03-13; Published online: 2021-04-01 17:21

URL: [kns.cnki.net/kcms/detail/11.2625.V.20210401.1459.002.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210401.1459.002.html)

Foundation item: National Key R & D Program of China (2017YFB0902800)

\* Corresponding author. E-mail: [xionglx@sdut.edu.cn](mailto:xionglx@sdut.edu.cn)

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2020.0700

# 基于多任务辅助推理的近眼视线估计方法

王小东<sup>1,2</sup>, 谢良<sup>2,3,\*</sup>, 闫慧炯<sup>2</sup>, 闫野<sup>2,3</sup>, 印二威<sup>2,3</sup>, 李卫国<sup>1</sup>

(1. 北京航空航天大学 软件学院, 北京 100083; 2. 天津(滨海)人工智能创新中心, 天津 300450;  
3. 军事科学院 国防科技创新研究院, 北京 100071)

**摘 要:** 眼动交互是头戴式虚拟现实(VR)/增强现实(AR)设备的关键操控方式, 如何进行高精度、高鲁棒性的非标定视线估计是当前VR/AR眼动交互的核心问题之一, 高效、鲁棒的非标定视线估计需要大量的眼图训练数据和高效的算法结构做支撑。在现有基于深度学习的近眼视线估计方法的基础上, 通过添加多任务辅助推理模块, 增加网络结构的多阶段输出, 进行多任务联合训练, 在不增加视线估计测试耗时的前提下, 有效提升视线估计精度。在模型训练时, 从视线估计网络结构的多个中间阶段引出多个眼部特征的辅助推理并行网络头, 包括眼动图像的语义分割、虹膜边界框及眼部轮廓信息, 为原始视线估计网络提供多阶段中继监控, 在不增加训练数据的基础上, 有效提升视线估计网络的测试精度。在国际公开数据集Acom-14与OpenEDS2020上的验证实验表明, 与无辅助推理的网络相比, 所提方法精度分别得到了21.74%与18.91%的效果提升, 平均角度误差分别减少到1.38°与2.01°。

**关键词:** 视线估计; 增强现实(AR); 人机交互; 多任务学习; 辅助推理

**中图分类号:** TP391

**文献标志码:** A

**文章编号:** 1001-5965(2022)06-1030-08

头戴式虚拟现实(VR)/增强现实(AR)设备能为用户提供丰富的生活工作交流体验, 也被称为新一代移动互联设备。眼动交互作为人类观测世界最直接的方式, 是头戴式VR/AR设备的关键交互方式之一。眼动交互需要精确鲁棒的视线估计做支撑, 通过检测用户实时视觉注视方向, 可有效理解用户的行为意图。当前, 基于头戴式VR/AR设备的眼动交互已经在多种应用中部署, 如基于视线的目标选择<sup>[1]</sup>、注意力监控<sup>[2]</sup>和视网膜中心凹式渲染等<sup>[3-4]</sup>。

近眼视线估计眼动交互技术主要分为标定式视线估计方法和非标定式视线估计方法。标定式视线估计方法主要基于人眼注视几何模型, 使用前需进行繁琐的眼动校准, 佩戴位置发生变化时

需重新标定, 难以适用于长时佩戴、剧烈运动等场景, 且无法应对头部晃动、环境变化等情况。非标定视线估计方法无需标定, 不受佩戴位置、注视目标和照明条件的影响, 但其依赖大量的眼动图像数据和高效的网络学习结构。随着眼动交互商用硬件设备的不断发展, 已累计出现了一定量的真实、仿真眼动图像数据集, 使得基于学习的非标定视线估计方法逐渐成为研究热点。基于学习的非标定视线估计在桌面式眼动仪研究领域已表现出良好的性能, 但面向头戴式VR/AR设备仍有许多不足, 且目前头戴式VR/AR设备的计算性能有限, 设计简洁高效的眼动模型至关重要。

非标定式视线估计方法通常使用深度学习以

收稿日期: 2020-12-18; 录用日期: 2021-04-09; 网络出版时间: 2021-04-15 12:57

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20210415.1129.001.html](http://kns.cnki.net/kcms/detail/11.2625.V.20210415.1129.001.html)

基金项目: 国家自然科学基金(61901505)

\* 通信作者. E-mail: xielnudt@gmail.com

**引用格式:** 王小东, 谢良, 闫慧炯, 等. 基于多任务辅助推理的近眼视线估计方法[J]. 北京航空航天大学学报, 2022, 48(6): 1030-1037. WANG X D, XIE L, YAN H J, et al. Near-eye gaze estimation based on multitasking auxiliary reasoning[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(6): 1030-1037 (in Chinese).

端到端的方式预测注视方向,但其精度有限。在深度学习中引入眼睛特征已被证明可有效提高视线估计精度,但带来精度提升的同时,网络复杂度也相应增加。基于上述问题,本文致力于寻找一种在不提升网络复杂度的前提下可有效引入眼睛特征的视线估计方法。

本文借用多任务学习思想,提出了面向头戴式VR/AR设备的近眼非标定视线估计网络结构ARGazeNet。ARGazeNet以现有的轻量级卷积神经网络(CNN)作为基础网络模型学习注视方向,在网络结构的中间阶段引出多个辅助推理模块,训练阶段以眼动图像的语义分割、虹膜边界框及眼部轮廓信息作为监督信息进行多任务辅助推理。在推理阶段,抛弃辅助推理模块而仅保留训练好的基础网络模型,从而实现在不增加测试耗时的前提下,视线估计精度的提升。本文在近眼数据集Acomo-14及OpenEDS2020上进行了精度测试实验,结果显示,相比未使用辅助推理的深度神经网络,本文方法取得了21.74%与18.91%的精度提升,证明了方法的有效性。

## 1 视线估计相关工作

非标定眼动视线估计方法通常先将从眼部图像中提取的高维向量作为图像特征,再将图像特征映射到视线点或视线方向。与传统的基于几何模型的标定眼动视线估计方法相比,基于学习的非标定眼动视线估计方法无需用户进行繁琐的校准,使用更为方便,且通过大样本眼动数据训练,能够适用于低分辨率的眼部图像。Zhang等<sup>[5]</sup>将基于学习的非标定眼动视线估计方法与基于几何模型的标定眼动视线估计方法在商用眼球跟踪仪器上进行了比较,结果表明,基于学习的方法比基于几何模型的方法具有更好的性能。在基于学习的视线估计方法中,使用神经网络是最有效的方法。Zhang等<sup>[6]</sup>提出了基于深度学习的视线估计方法,并提出了桌面式视线估计领域中应用最广泛的数据集MPIIGaze,随后Zhang等<sup>[7]</sup>又对其进行了网络改进。Cheng等<sup>[8]</sup>基于双眼对视线估计的不同贡献,提出了一种非对称回归方法。Krafka等<sup>[9]</sup>提出了一个itrack模型,针对桌面式眼动仪,将人脸图像、左右眼图像和头部位置同时通过4个网络分支进行处理,融合后输出屏幕视线点。

利用眼睛特征(如眼睛关键点定位)是传统几何方法的核心思想,也启发了一些利用附加语义信息来提高视线估计精度的研究。Park等<sup>[10]</sup>提出了一种基于人眼图像表示的视线估计方法,

将人眼抽象为一个简单的眼球图形,以提高视线估计的性能。Park等<sup>[11]</sup>使用沙漏网络来预测关键点的热图,从而进行视线估计。Yu等<sup>[12]</sup>提出了一种基于约束模型的视线估计方法,在视线估计时检测眼睛关键点的位置,同时学习上述2个任务,并利用互补的信息来提高性能。以上基于眼睛外观关键点特征定位学习的视线估计方法是面向桌面式眼动仪设计的,使用的数据集主要是由普通单目相机捕获的RGB图像生成的数据集,该数据集将整个面部图像或从整个面部图像中提取的眼睛图像作为输入,导致眼睛图像分辨率低且眼睛的大量关键特征缺失,从而导致预测精度往往偏低。

随着头戴式VR/AR设备能更加方便地捕捉到高分辨率的近眼图像,基于近眼图像的视线估计研究成为热点,近年来出现了许多基于近眼图像的视线估计研究。Facebook在2019—2020年连续举办OpenEDS竞赛,并公布了2个大型近眼图像数据集OpenEDS<sup>[13]</sup>和OpenEDS2020<sup>[14]</sup>。在比赛中,产生了用于眼睛图像分割<sup>[15]</sup>、使用时间信息的视线估计<sup>[16]</sup>和未来图像视线预测的多个模型。NVIDIA<sup>[17]</sup>提出了一个包含 $2 \times 10^6$ 合成图像和 $2.5 \times 10^6$ 近眼图像的近眼数据集,并训练了一个低延迟神经网络,实现了受试者之间 $2.06^\circ$ ( $\pm 0.44^\circ$ )的视线估计精度。Magic Leap<sup>[18]</sup>提出了一个新的混合现实数据集MagicEyes,包含587个被试,共80多万幅图像,并提出使用多任务模型,包括眼睛分割、眨眼检测、情感分类、红外LED闪烁测试、瞳孔和视力估计等,将EyeNet模型与传统的基于几何的视线估计方法相结合。虽然以上数据集为VR/AR眼动非标定视线估计提供了帮助,但如何将多任务之间的关系充分挖掘,从而有效提升视线估计精度的问题,仍未得到很好的解决。现有工作大多采用串联的形式,在视线估计任务前增加眼球区域估计任务,导致不能充分发挥多任务之间的内部关联特性,并且计算速度较慢。将多任务辅助推理并行连接到视线估计网络模型中,在有效提升视线估计精度的同时,不增加算法测试耗时,可以很好地应对头戴式VR/AR设备计算资源有限的问题。

## 2 模型

本文提出了基于多任务辅助推理的VR/AR视线估计网络模型ARGazeNet,使用CNN模型进行近眼图像的视线估计,获得眼部坐标系下的视线向量的俯仰角与偏航角,并在多个中间层引出

辅助推理并行网络头,用来预测眼部语义分割图、虹膜边界框信息与眼部关键点轮廓等辅助眼部特征,为视线估计网络提供多种中继监控,在不增加视线估计测试耗时的前提下,有效提升视线估计精度。

ARGazeNet 主要分为两部分:一个基础 CNN 模型和多个中继辅助推理模块。

ARGazeNet 网络结构如图 1 所示。ARGaze-

Net 使用预处理过的近眼图像作为输入,将其输入到基于 Efficientnet-B0 的主干网络中,以输出俯仰角和偏航角 2 个角度。同时,经过前 4 层卷积后,将数据引入辅助推理模块,分别使用语义分割图、眼部轮廓图和虹膜边界框坐标进行中继监督学习。在视线估计测试时,去除辅助推理模块,仅使用基础网络部分进行视线估计,不增加测试耗时。

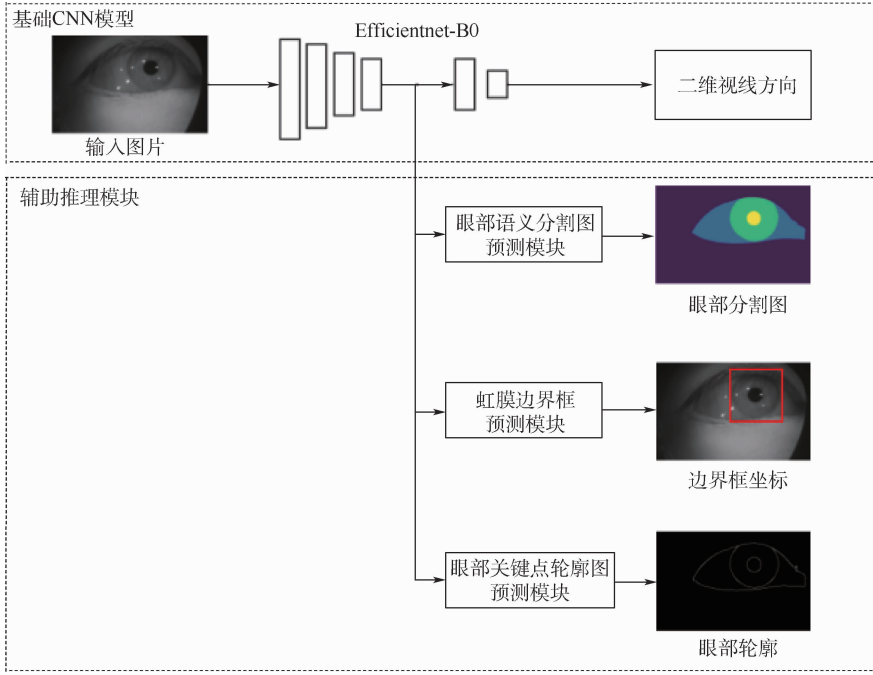


图 1 ARGazeNet 网络结构

Fig. 1 Network architecture of ARGazeNet

## 2.1 基础 CNN 模型

CNN 已经被证明具有非凡的特征提取能力,本文利用深度学习的优势,设计了一个基于 CNN 的图像特征提取器。考虑到眼动的实时处理速度要求( $\geq 60$  Hz),以及头戴式 VR/AR 设备性能较低等因素,采用了目前公认图像处理能力较强且资源消耗较少的 EfficientNet<sup>[19]</sup> 网络作为基础网络架构。

针对头戴式 VR/AR 设备获得的眼动灰度图像, Efficientnet-B0 网络的输入通道数被修改为 1。相比桌面式眼动仪需要在神经网络中添加额外的全连接层结合头部姿态信息以进行最终注视方向输出,基于头戴式眼动仪在 Efficientnet-B0 网络的模型后直接输出眼动注视方向的俯仰角与偏航角,无需额外的网络层。

使用三维单位向量之间的空间角作为评价指标,通过以下公式转换,将预测的俯仰角和偏航角映射到欧氏空间中的三维单位向量空间:

$$\begin{cases} x = \sin \theta \cos \delta \\ y = \sin \theta \sin \delta \\ z = \cos \theta \end{cases} \quad (1)$$

式中: $\theta$  和  $\delta$  分别为俯仰角和偏航角; $(x, y, z)$  为欧氏空间内的三维单位向量。

角度误差计算方法为

$$\phi = \arccos(x\hat{x} + y\hat{y} + z\hat{z}) \quad (2)$$

式中: $x, y, z$  为由俯仰角和偏航角真值转化而来的三维单位矢量数值; $\hat{x}, \hat{y}, \hat{z}$  为由模型输出转化而来的三维单位矢量数值。

使用均方误差损失作为基础模型损失函数进行训练,具体公式为

$$l_{\text{base}} = \frac{1}{n} \sum_{i=1}^n [(\theta_i - \hat{\theta}_i)^2 + (\delta_i - \hat{\delta}_i)^2] \quad (3)$$

式中: $n$  为单个 batch 中图片的数量; $\theta_i$  和  $\delta_i$  分别为第  $i$  张图片俯仰角和偏航角的真值; $\hat{\theta}_i$  和  $\hat{\delta}_i$  分别为第  $i$  张图片俯仰角和偏航角的模型预测值。

## 2.2 辅助推理模块

基于基础模型提出了辅助推理模块,在不增

加测试耗时的前提下,提升眼动注视方向估计精度。具体网络细节如图 2 所示。图中:k 指卷积核,后面的数字指卷积核大小。BN 指批量归一化,Swish 指 Swish 激活函数,MBCConv 指增加了 squeeze-and-excitation 优化的 mobile inverted bottleneck,c 指卷积核的通道数,interpolate 指使用插值函数进行上采样差分运算,AdaptiveAvg Pool2d 指自适应 2D 平均池化。在 ARGazeNet 网络中,使用了 3 个辅助推理模块,分别以预测虹膜边界框信息、眼部轮廓信息及眼部语义分割图作为监督信息。基础 CNN 模型共含有 7 个卷积模块组,并使用前 4 个卷积模块组进行权重共享。

1) 虹膜边界框预测模块。考虑到预测目标为虹膜边界坐标,预测任务与目标检测相似,但同时仅有单个检测目标(即人眼虹膜这一单个检测

目标),因此借鉴并改造经典目标检测网络结构 FasterRcnn,从主干网络 Efficientnet-B0 模型的前 4 层引出中继图像特征,作为虹膜边界预测网络头的输入,边界检测网络头由 3 个卷积模块组及单个全连接层组成。对于边界检测中继损失函数,使用以下 smoothL1loss 函数进行监督训练:

$$\begin{cases} l_{iris} = \frac{1}{n} \sum_{i=1}^n z_i \\ z_i = \begin{cases} 0.5(x_i - \hat{x}_i)^2 & |x_i - \hat{x}_i| < 1 \\ |x_i - \hat{x}_i| - 0.5 & \text{其他} \end{cases} \end{cases} \quad (4)$$

式中:n 为一个 batch 中虹膜边界顶点的数量;x<sub>i</sub> 和  $\hat{x}_i$  分别为第 i 个虹膜边界顶点坐标的真值和预测值;z<sub>i</sub> 为第 i 个虹膜边界顶点坐标的损失值。

虹膜边界顶点坐标的真值通过 OpenCV 目标跟踪和人工修正获得。

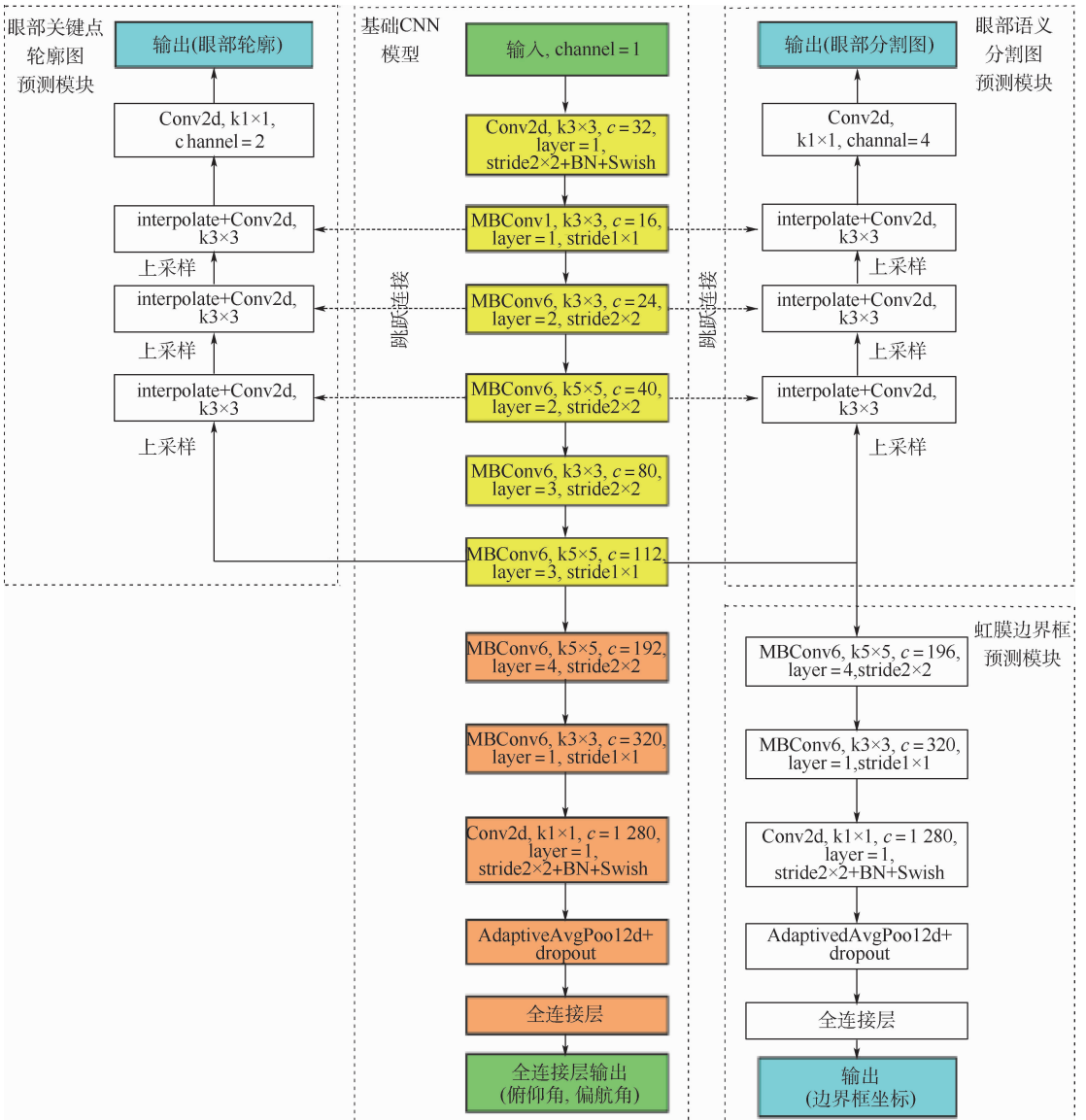


图 2 ARGazeNet 的架构细节  
Fig. 2 Architecture details of ARGazeNet

2) 眼部语义分割图预测模块。OpenEDS 比赛中证明了 RITnet 网络结构<sup>[15]</sup>在近眼语义分割任务中极其优异的性能,因此借鉴 RITnet 网络结构<sup>[15]</sup>,对图片进行下采样后,再进行上采样与跳跃连接操作(skip connection)得到语义分割图。直接使用主干网络的4个共享卷积模块作为下采样层,使用如图2所示的4个卷积组作为上采样网络结构,并将上采样网络组与下采样网络进行跳跃连接。使用如下标准交叉熵损失作为眼球语义分割损失函数:

$$l_{\text{seg}} = - \sum_{k,(r,c)} I_k(T_k(r,c)) \lg \left( \frac{e^{P_k(r,c)}}{\sum_{j=1}^n e^{P_j(r,c)}} \right) \quad (5)$$

式中: $r,c$ 表示图片的像素坐标; $T$ 为像素点的语义类型真值; $P_k$ 为像素点的神经网络预测值; $I_k$ 为第 $k$ 类语义类型的指示函数。

语义分割图的真值使用经过训练的原始 RITnet 网络再结合人工修正获得。RITnet 网络<sup>[15]</sup>是在 OpenEDS 数据集上<sup>[13]</sup>训练生成,并在 2019 年 OpenEDS 语义分割挑战赛上取得了 95.3% 的准确率。

3) 眼部关键轮廓图预测模块。相比于眼部语义分割,对于关键轮廓图预测任务,需要更加关注显著性目标的边缘信息,BASNet 网络<sup>[20]</sup>正是专注于此类任务的网络结构。因此,借鉴 BASNet 的网络结构,网络结构层使用了与眼部语义分割图预测模块相同的4个卷积组与跳跃连接作为网络头。使用如下交叉熵损失、SSIM 损失与 IOU 损失融合的损失函数进行监督训练,引导网络在像素级、图像块级和全图级3个层面共同学习预测图片与眼部关键轮廓图之间的映射关系:

$$l_{\text{cou}} = l_{\text{bce}} + l_{\text{ssim}} + l_{\text{iou}} \quad (6)$$

式中: $l_{\text{bce}}$ 为二值交叉熵损失,其为像素级别的损失,表示为

$$l_{\text{bce}} = - \sum_{(r,c)} |G(r,c) \lg(S(r,c)) + (1 - G(r,c)) \lg(1 - S(r,c))| \quad (7)$$

式中: $G$ 为像素点的轮廓标签,当像素点位于轮廓上标签为1,否则标签为0; $S$ 为当前像素点位于轮廓上的预测概率。

$l_{\text{ssim}}$ 表示结构相似性损失,其为像素块级别上的损失,为 Qin 等<sup>[20]</sup>所创用以关注每个像素的局部领域,表示为

$$l_{\text{ssim}} = 1 - \frac{(2\mu_a\mu_b + C_1)(2\sigma_{ab} + C_2)}{(\mu_a^2 + \mu_b^2 + C_1)(\sigma_a^2 + \sigma_b^2 + C_2)} \quad (8)$$

式中: $a$ 和 $b$ 分别为从预测图像和真值图像中提取出来的 $N \times N$ 个区域; $\mu_a, \mu_b$ 和 $\sigma_a, \sigma_b$ 分别为

$a, b$ 的均值和方差; $\sigma_{ab}$ 为 $a, b$ 之间的协方差; $C_1 = 0.01^2, C_2 = 0.03^2$ ,以避免分母为0。

$l_{\text{iou}}$ 表示交叠率损失,其为全图级别的损失,表示为

$$l_{\text{iou}} = 1 - \frac{\sum_{r=1}^H \sum_{c=1}^W S(r,c)G(r,c)}{\sum_{r=1}^H \sum_{c=1}^W |S(r,c) + G(r,c) - S(r,c)G(r,c)|} \quad (9)$$

式中: $H$ 和 $W$ 分别为图片的高度和宽度。

眼部关键轮廓图的真值通过对眼部语义分割图的真值进行 canny 边缘检测算法获取。

## 3 实验

### 3.1 眼动精度测试

#### 3.1.1 数据集

针对提出的视线估计 ARGazeNet 模型,利用公开数据集 Acom-14 和 OpenEDS2020 进行了性能测试。

1) Acom-14。数据集 Acom-14 共包含了8个被试,总计14个 trial,每个 trail 中有49个视线点。每个视线点有1000张压缩过眼部图片,并有视线方向标签 gazeX(偏航角)和 gazeY(俯仰角)。数据集是用安装在 Oculus DK2 上的 300 Hz 红外相机收集的,视场角 FOV 为 $40^\circ \times 40^\circ$ 。为便于进行训练,数据集图像已被裁剪并调整分辨率为 $127 \times 127$ 的大小。

2) OpenEDS2020。使用 OpenEDS2020: Gaze Prediction Dataset 的一个子集,共使用40个独立的被试,其中32个目标用于训练,8个被试用于测试。训练子集由每只眼睛、模式类型和匿名被试的10个100帧序列组成,每个被试共有4000张图像,共128000张图像。

#### 3.1.2 对比方法

主要对比的非标定眼动视线估计方法是 Kim 模型<sup>[17]</sup>,该模型由6个卷积层组成,输入分辨率为 $127 \times 127$ ,第一层包含8个通道,最后用一个全连接层来输出水平和垂直视线角。所有卷积层都使用 $3 \times 3$ 卷积核,步长为 $2 \times 2$ 像素,即每层中图像分辨率减半,不使用池化层或填充,每层中输出通道数增加1.5倍。每一个卷积层后面都有一个 dropout 层, $p = 0.1$ ,以防止过拟合,网络结构如图3所示。

#### 3.1.3 测试细节

在模型训练时,先对图像进行了预处理,借鉴 Park 等<sup>[10]</sup>提出的图像归一化的思想,将图像分辨

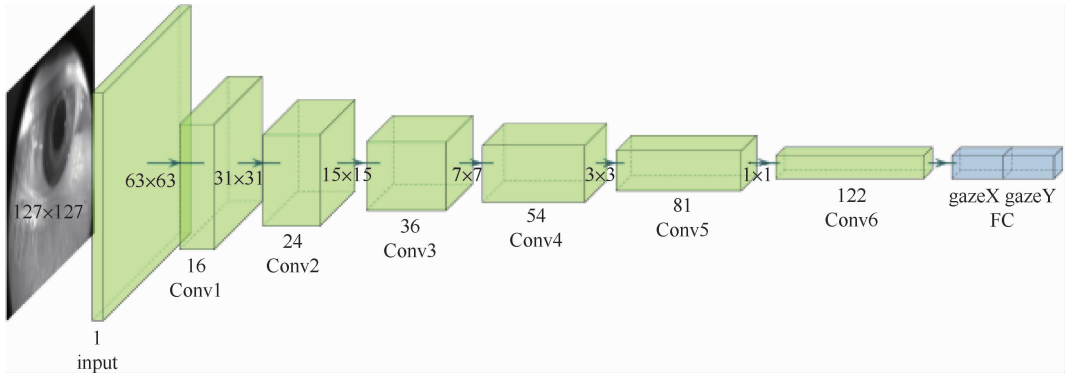


图 3 用于对比实验的 Kim 模型

Fig. 3 Kim model for comparison experiments

率压缩为  $127 \times 127$ , 与 Kim 模型的分辨率保持一致, 再在  $-60^\circ \sim 60^\circ$  之间随机翻转, 在 40 个像素范围内随机平移, 最终对图像信息进行归一化处理, 进行 3 倍的数据扩展, 以提高模型在不同人佩戴头戴式设备时的鲁棒性。在 GTX1660Ti 上进行训练, batch size 为 64, 模型输入大小为  $64 \times 1 \times 127 \times 127$ , 训练 epoch 数为 100, 使用 Adam 优化器进行训练, 学习率固定为 0.001。

对以原始图像为输入及没有任何辅助推理模块的 ARGazeNet 进行训练; 基于多任务辅助推理模块对 ARGazeNet 网络进行优化训练。同时, 为了展示辅助推理任务与视线估计任务的相关性, 与图像语义分割图作为输入的 ARGazeNet 进行视线估计对比。

### 3.1.4 测试结果

在基于 CUDA 的深度学习优化 GPU 核心架构 cuDNN<sup>[21]</sup> 上实现了本文的网络结构, 并对 Kim 模型进行复现, 在单台 Intel Core i7-9700K CPU、Nvidia GeForce GTX 1660 Ti 显卡及 INTEL SSD-SCKKW480H6 固态硬盘的硬件条件下进行实验, 实验结果如表 1 所示。ARGazeNet 模型在数据集 Acom-14 和 OpenEDS2020 中的平均角度误差最低, 分别为  $1.38^\circ$  和  $2.01^\circ$ 。相比于不含辅助推理模块的 ARGazeNet 模型, 分别得到了 21.74% 和 18.91% 的效果提升。

实验表明, ARGazeNet 网络结构的视线估计误

表 1 数据集 Acom-14 和 openEDS2020 上跨主体评估的平均角度误差

Table 1 Mean angular error in cross-subjects on Acom-14 and openEDS2020 datasets

模型	平均角度误差/ $(^\circ)$	
	Acom-14	openEDS2020
Kim 模型 <sup>[17]</sup>	2.51	3.51
不含辅助推理模块的 ARGazeNet	1.68	2.39
ARGazeNet	1.38	2.01
使用分割图作为输入的 ARGazeNet	5.94	4.95

差均小于 Kim 模型, 使用辅助推理模块的 ARGazeNet 的精度高于没有辅助推理模块的 ARGazeNet 网络。

同时可以看出, 当仅使用语义分割图作为视线估计输入时, 视线估计误差可以达到  $5^\circ \sim 6^\circ$ , 这表明关键眼球轮廓信息与视线估计矢量之间的强关联性, 也可以证明使用其来辅助视线估计推理的有效性。但从中也可以看出, 仅使用语义分割图作为输入, 方法存在较大的误差, 原因在于: 虽然语义分割图是高度浓缩的眼球关键特征轮廓, 但并不包括单个眼球细节。因此, 与 Park 等<sup>[11]</sup> 提出的使用神经网络学习眼部关键信息, 再使用眼部关键信息进行视线估计推理的思路不同, 将多种眼部关键信息作为辅助监督信息, 在为眼动注视方向估计网络提供辅助监督信息的同时, 不限制视线估计网络的性能, 从而有效提升视线估计精度。

### 3.2 算法耗时对比

3.1 节实验证明了 ARGazeNet 的可行性, 视线估计模型不仅要求精度高, 而且要求模型部署后的运行效率。在 3.1.4 节所述的实验平台条件下, 测试了单张图片的模型推理时间与包含图片读取、预处理所有操作的总耗时, 如表 2 所示。可以看到, 与 Kim 模型、没有辅助推理模块的 ARGazeNet 相比, ARGazeNet 的测试时间没有明显增加, 足以在 100 Hz 的硬件设备中实现全实时运行。

表 2 测试硬件上的推理性能

Table 2 Inference performance on tested hardware

模型	总耗时/ms
Kim 模型	2.34
不含辅助推理模块的 ARGazeNet	2.47
ARGazeNet	2.47

### 3.3 消融实验

为了研究辅助推理模块对视线估计误差的影响, 进行了 6 组消融实验: 仅使用单个辅助推理任务、使用 2 个辅助推理任务组合, 实验结果如表 3

表3 辅助推理模块的影响

Table 3 Effectiveness of auxiliary reasoning module

模型	平均角度误差/(°)	
	Acomo-14	openEDS2020
仅使用基础 CNN 模型	1.68	2.39
添加眼部语义分割图预测模块	1.60	2.20
添加虹膜边界框预测模块	1.46	2.30
添加眼部关键轮廓图预测模块	1.63	2.32
添加眼部语义分割图预测模块 与虹膜边界框预测模块	1.41	2.15
添加眼部语义分割图预测模块 与眼部关键轮廓图预测模块	1.54	2.18
添加虹膜边界框预测模块与眼 部关键轮廓图预测模块	1.43	2.26
ARGazeNet	1.38	2.01

所示。结果表明,使用了辅助推理模块的网络模型相较于基础 CNN 模型,效果都较好;同时,使用多个辅助推理模块比全部使用单一辅助推理模块的效果要好,而且用于预测的任务越复杂,模型性能越好。

## 4 结 论

本文提出了一种新的基于辅助推理的非标定的近眼视线估计方法。通过实验证明:

1) 可有效地进行视线估计,在数据集 Acomo-14 与 OpenEDS2020 上平均角度误差分别为  $1.38^\circ$  与  $2.01^\circ$ 。

2) 使用辅助推理模块比普通 CNN 有显著的优势,能使网络更加专注于眼睛关键特征的提取,提高了模型的泛化能力。

3) 有效辅助推理信息越多,语义信息越充分,模型精度越高。

## 参考文献 (References)

- [1] KYTÖ M, ENS B, PIUMSOMBOON T, et al. Pinpointing: Precise head- and eye-based target selection for augmented reality [C] // Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, 2018:1-14.
- [2] SMITH P, SHAH M, DA VITORIA LOBO N. Determining driver visual attention with one camera [J]. IEEE Transactions on Intelligent Transportation Systems, 2003, 4(4):205-218.
- [3] ALBERT R, PATNEY A, LUEBKE D, et al. Latency requirements for foveated rendering in virtual reality [J]. ACM Transactions on Applied Perception, 2017, 14(4):1-13.
- [4] PATNEY A, SALVI M, KIM J, et al. Towards foveated rendering for gaze-tracked virtual reality [J]. ACM Transactions on Graphics, 2016, 35(6):1-12.
- [5] ZHANG X, SUGANO Y, BULLING A. Evaluation of appearance-based methods and implications for gaze-based applications [C] // Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 2019:1-13.
- [6] ZHANG X, SUGANO Y, FRITZ M, et al. Appearance-based gaze estimation in the wild [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2015:4511-4520.
- [7] ZHANG X, SUGANO Y, FRITZ M, et al. MPIIGaze: Real-world dataset and deep appearance-based gaze estimation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 41(1):162-175.
- [8] CHENG Y, ZHANG X, LU F, et al. Gaze estimation by exploring two-eye symmetry [J]. IEEE Transactions on Image Processing, 2020, 29:5259-5272.
- [9] KRAFKA K, KHOSLA A, KELLNHOFER P, et al. Eye tracking for everyone [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016:2176-2184.
- [10] PARK S, SPURR A, HILLIGES O. Deep pictorial gaze estimation [C] // Proceedings of the European Conference on Computer Vision (ECCV). Berlin: Springer, 2018:721-738.
- [11] PARK S, ZHANG X, BULLING A, et al. Learning to find eye region landmarks for remote gaze estimation in unconstrained settings [C] // Proceedings of the 2018 ACM Symposium on Eye Tracking Research and Applications. New York: ACM, 2018:1-10.
- [12] YU Y, LIU G, ODOBEZ J M. Deep multitask gaze estimation with a constrained landmark-gaze model [C] // Proceedings of the European Conference on Computer Vision (ECCV). Berlin: Springer, 2018:456-474.
- [13] GARBIN S J, SHEN Y, SCHUETZ I, et al. OpenEDS: Open eye dataset [EB/OL]. (2019-05-17) [2020-12-01]. <https://arxiv.org/abs/1905.03702>.
- [14] PALMERO C, SHARMA A, BEHRENDT K, et al. OpenEDS2020: Open eyes dataset [EB/OL]. (2020-05-08) [2020-12-01]. <https://arxiv.org/abs/2005.03876>.
- [15] CHAUDHARY A K, KOTHARI R, ACHARYA M, et al. RITnet: Real-time semantic segmentation of the eye for gaze tracking [C] // 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). Piscataway: IEEE Press, 2019:3698-3702.
- [16] PALMERO C, KOMOGORTSEV O V, TALATHI S S. Benefits of temporal information for appearance-based gaze estimation [C] // Proceedings of the 2020 ACM Symposium on Eye Tracking Research and Applications. New York: ACM, 2020:1-5.
- [17] KIM J, STENGEL M, MAJERICIK A, et al. NVGaze: An anatomically-informed dataset for low-latency, near-eye gaze estimation [C] // Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 2019:1-12.
- [18] WU Z, RAJENDRAN S, VAN AS T, et al. EyeNet: A multi-task deep network for off-axis eye gaze estimation [C] // 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). Piscataway: IEEE Press, 2019:3683-3687.
- [19] TAN M, LE Q. EfficientNet: Rethinking model scaling for convolutional neural networks [C] // International Conference on Machine Learning, 2019:6105-6114.

- [20] QIN X, ZHANG Z, HUANG C, et al. BASNet: Boundary-aware salient object detection[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2019: 7479-7489.
- [21] CHETLUR S, WOOLLEY C, VANDERMERSCH P, et al. cuDNN: Efficient primitives for deep learning[EB/OL]. (2014-12-18) [2020-12-01]. <https://arxiv.org/abs/1410.0759>.

## Near-eye gaze estimation based on multitasking auxiliary reasoning

WANG Xiaodong<sup>1,2</sup>, XIE Liang<sup>2,3,\*</sup>, YAN Huijiong<sup>2</sup>, YAN Ye<sup>2,3</sup>, YIN Erwei<sup>2,3</sup>, LI Weiguo<sup>1</sup>

(1. College of Software, Beihang University, Beijing 100083, China;

2. Tianjin (Binhai) Artificial Intelligence Innovation Center, Tianjin 300450, China;

3. Defense Innovation Institute, Academy of Military Sciences, Beijing 100071, China)

**Abstract:** Eye-tracking interaction is the key control method for head-mounted virtual reality (VR)/augmented reality (AR) devices and non-calibrated gaze estimation is one of the core problem in current VR/AR eye-tracking interactions. Efficient and robust non-calibrated gaze estimation requires a large amount of training data and an efficient network structure. Based on the existing deep-learning-based near-eye gaze estimation method, by adding multitasking auxiliary reasoning and increasing the multi-stage output of the network structure for joint multi-task training, we achieve an effective improvement of gaze estimation accuracy without increasing the refer time compared to the original gaze estimation network. During model training, multiple intermediate stages of the gaze estimation network structure are used to derive multiple parallel network headers for auxiliary reasoning about eye features, including semantic segmentation of eye images, iris border frames, and eye contour information, to provide multi-stage relay monitoring for the original gaze estimation network, which effectively improves the generalization capability of the gaze estimation network without increasing the training data. Experiments on the open datasets Acom0-14 and OpenEDS2020 show that the accuracy of the algorithm is improved by 21.74% and 18.91%, respectively, and the average gaze estimation error is reduced to 1.38 degrees and 2.01 degrees, compared with the network without auxiliary reasoning.

**Keywords:** gaze estimation; augmented reality (AR); human-computer interaction; multitask learning; auxiliary reasoning

**Received:** 2020-12-18; **Accepted:** 2021-04-09; **Published online:** 2021-04-15 12:57

**URL:** [kns.cnki.net/kcms/detail/11.2625.V.20210415.1129.001.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210415.1129.001.html)

**Foundation item:** National Natural Science Foundation of China (61901505)

\* **Corresponding author.** E-mail: [xielnudt@gmail.com](mailto:xielnudt@gmail.com)

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2020.0701

# 输入受限下多无人机三维协同路径跟踪控制

刘荣华<sup>1</sup>, 刘树光<sup>1,\*</sup>, 张博洋<sup>1</sup>, 王欢<sup>1</sup>, 李伟<sup>2</sup>

(1. 空军工程大学 装备管理与无人机工程学院, 西安 710038; 2. 空军装备部驻洛阳地区第二军事代表室, 洛阳 471009)

**摘 要:** 针对多无人机在三维空间的协同路径跟踪问题, 设计了基于反步法的协同控制器。考虑无人机在飞行过程中的输入饱和问题, 加入辅助控制系统, 以确保系统在输入受限下仍能保持良好的控制性能。将无人机六自由度非线性模型反馈线性化处理, 同时考虑无人机飞行时受到的外界不确定扰动及自身模型存在的未建模动态, 利用径向基函数在线估计补偿, 提高了系统的鲁棒性和抗干扰能力。为了解决反步法需要对虚拟控制量求导导致的控制器复杂问题, 引入一阶滤波器, 避免对虚拟控制量的求导。利用图论解决无人机机间通信问题, 基于一致性理论实现了多无人机的协同控制。基于 Lyapunov 稳定性理论, 证明了系统的稳定性, 仿真结果表明, 所设计的路径跟踪协同控制器能够达到良好的协同跟踪控制效果。

**关键词:** 多无人机; 输入受限; 协同控制; 路径跟踪; 反步法

**中图分类号:** V221<sup>+</sup>.3

**文献标志码:** A **文章编号:** 1001-5965(2022)06-1038-12

多无人机的协同路径跟踪由于在民用和军事领域的需求愈加凸显<sup>[1]</sup>, 近年来得到了诸多研究者的广泛关注。与轨迹跟踪控制(trajjectory tracking control, TTC)不同, 路径跟踪控制(path following control, PFC)是无人机在控制器的作用下以一定控制精度跟踪一条与时间无关的参数化路径, 其是一种不考虑时间、强调几何位置关系的静态跟踪。路径跟踪控制一般可应用于无人机的巡航、侦察、避障、地形跟踪等任务。

对于多无人机协同路径跟踪控制问题, 近年来涌现了大量研究成果, 常用于实现无人机路径跟踪控制的方法有反步法<sup>[2-4]</sup>、模型预测控制<sup>[5-7]</sup>、反馈线性化<sup>[8-10]</sup>、自适应滑模控制<sup>[11-12]</sup>等。文献[13]指出, 现有跟踪控制大多使用反步法。基于反步法的控制器的控制性能良好, 足以为其其他类型的控制器提供性能参考。为了处理恒定的风扰动影响, 文献[3]通过估计外界干扰改进了该控制器, 以减轻外界干扰影响, 实验结果验证了

所设计控制器的鲁棒性。文献[4]探讨了多无人机协同运输的路径跟踪协同控制问题, 通过数值仿真验证了该控制系统具有较高的路径跟踪精度和较低的协同误差。文献[14]对固定翼无人机的协同路径跟踪控制问题进行了研究, 所提控制算法有效实现了固定队型的路径跟踪。文献[15]证明了多无人机协同路径跟踪控制在约束条件下的闭环系统渐近稳定性。文献[16]采用混合高阶控制器, 解决了固定翼无人机在有速度限制下的多无人机协同路径跟踪控制问题。文献[17]采用视线制导算法, 通过非线性固定时间观测器估计并补偿不确定项, 实现了二维平面内无人机的路径跟踪容错控制。文献[18]提出了具有有界未知扰动的一种连续非线性鲁棒控制方法, 实现了基于四旋翼无人机的二维平面分布式时变编队控制。

以上研究多基于二维平面的三自由度模型, 少有对无人机三维空间的路径跟踪控制问题的研

收稿日期: 2020-12-18; 录用日期: 2021-03-05; 网络出版时间: 2021-03-16 14:56

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20210316.1444.003.html](http://kns.cnki.net/kcms/detail/11.2625.V.20210316.1444.003.html)

基金项目: 国家自然科学基金(61973253); 国防科技重点实验室基金(6142219200301); 航空科学基金(20155896025)

\* 通信作者。E-mail: dawn418@126.com

**引用格式:** 刘荣华, 刘树光, 张博洋, 等. 输入受限下多无人机三维协同路径跟踪控制[J]. 北京航空航天大学学报, 2022, 48(6): 1038-1049. LIU R H, LIU S G, ZHANG B Y, et al. 3D cooperative path following control of multi-UAVs with input saturation [J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(6): 1038-1049 (in Chinese).

究,且很少考虑无人机模型自身的未建模动态及外界非线性时变扰动等不确定因素对系统稳定性的影响。同时,无人机执行器的输入限制(如升降舵面、发动机推力的饱和)是无人机控制器设计时必须关注的问题,若不考虑输入约束,在无人机实际飞行中将难以达到预期控制效果甚至影响系统的稳定性,前人鲜有对多无人机协同路径跟踪控制问题中输入限制问题进行研究。另外,运用反步法设计跟踪控制器需要对虚拟控制量进行复杂的求导运算,将导致控制器结构复杂,不利于控制器的设计。

针对上述问题,本文主要研究工作如下:①与前人研究较多的二维模型不同的是,本文通过对三维六自由度非线性模型的反馈线性化处理,将简化的无人机三维模型应用于路径跟踪控制问题,设计三维空间的路径跟踪协同控制器。②无人机往往在不确定复杂环境下工作,本文利用神经网络逼近环境的不确定扰动及自身模型存在的未建模动态,设计基于自适应动态面控制技术的三维多无人机协同路径跟踪控制策略。③考虑输入限制,设计无人机跟踪控制算法,以确保在输入饱和下系统的稳定性。④基于图论设计了分布式的路径跟踪协同控制策略,可以有效降低多无人机间的通信负载。⑤将反步法控制器应用于多无人机协同路径跟踪控制问题,为了有效解决反步法在对虚拟控制输入求导时的微分爆炸问题,引入一阶滤波器,以避免无人机控制器结构的复杂性问题。基于 Lyapunov 稳定性理论,证明了多无人机跟踪误差最终一致有界,并通过仿真验证了控制算法的有效性。

## 1 预备知识

### 1.1 无人机模型

本文采用无人机质点模型<sup>[19]</sup>描述无人机的飞行运动,在惯性坐标系中定义相关变量,如图 1 所示。

为了简化模型,做以下假设:①无人机的推力方向与飞行速度方向一致;②大地水平,即不考虑地球曲率的影响;③无人机的质点位置时不变。于是,无人机的运动模型可以描述为<sup>[20]</sup>

$$\begin{cases} \dot{x}_i = V_{gi} \cos \gamma_i \cos \chi_i \\ \dot{y}_i = V_{gi} \cos \gamma_i \sin \chi_i \\ \dot{h}_i = V_{gi} \sin \gamma_i \end{cases} \quad (1)$$

式中: $i = 1, 2, \dots, n$  为无人机的编号, $n$  为编队中

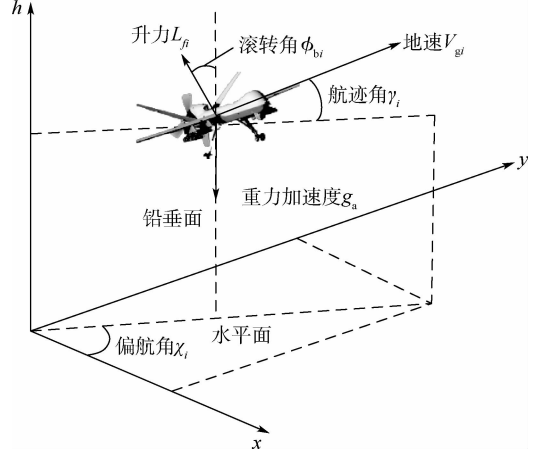


图 1 无人机模型

Fig. 1 UAV model

无人机的数量; $(x_i, y_i, h_i)$  为无人机  $i$  在惯性坐标系中的位置坐标; $V_{gi}$  为地速; $\gamma_i$  为航迹角; $\chi_i$  为偏航角。

无人机的动力学模型可以描述为

$$\begin{cases} \dot{V}_{gi} = \frac{T_i - D_{gi}}{m_i} - g_a \sin \gamma_i \\ \dot{\gamma}_i = \frac{g_a}{V_{gi}} (n_{gi} \cos \phi_{bi} - \cos \gamma_i) \\ \dot{\chi}_i = \frac{L_i \sin \phi_{bi}}{m_i V_{gi} \cos \gamma_i} \end{cases} \quad (2)$$

式中: $T_i$  为无人机发动机产生的推力控制量; $D_{gi}$  为气动阻力; $m_i$  为无人机的质量; $g_a$  为重力加速度; $L_i$  为升力; $\phi_{bi}$  为无人机方向舵和副翼产生滚转角控制量; $n_{gi} = L_i / (g_a m_i)$  为无人机升降舵产生的过载控制量。

上述无人机的非线性模型可以通过反馈线性化转化为线性的时变二阶积分模型<sup>[21]</sup>,则无人机的动力学模型可以由式(1)和式(2)转化为

$$\begin{cases} \ddot{x}_i = a_{xi} \\ \ddot{y}_i = a_{yi} \\ \ddot{h}_i = a_{hi} \end{cases} \quad (3)$$

式中: $a_{xi}$ 、 $a_{yi}$ 、 $a_{hi}$  分别为无人机二阶积分模型中的控制量,与无人机实际控制量的关系为

$$\begin{cases} \phi_{bi} = \text{artan} \left( \frac{(a_{yi} \cos \chi_i - a_{xi} \sin \chi_i) / (\cos \gamma_i (a_{hi} + g_a) - \sin \gamma_i (a_{xi} \cos \chi_i + a_{yi} \sin \chi_i))}{\cos \gamma_i (a_{hi} + g_a)} - \frac{\sin \gamma_i (a_{xi} \cos \chi_i + a_{yi} \sin \chi_i)}{g_a \cos \phi_{bi}} \right) \\ n_{gi} = \frac{\cos \gamma_i (a_{hi} + g_a)}{g_a \cos \phi_{bi}} - \frac{\sin \gamma_i (a_{xi} \cos \chi_i + a_{yi} \sin \chi_i)}{g_a \cos \phi_{bi}} \\ T_i = \sin \gamma_i (a_{hi} + g_a) m_i + D_{gi} + \cos \gamma_i (a_{xi} \cos \chi_i + a_{yi} \sin \chi_i) m_i \end{cases} \quad (4)$$

偏航角  $\chi_i$  和航迹角  $\gamma_i$  可以表示为

$$\begin{cases} \tan \chi_i = \frac{\dot{y}_i}{\dot{x}_i} \\ \sin \gamma_i = \frac{\dot{h}_i}{V_{gi}} \end{cases} \quad (5)$$

基于无人机二阶积分模型(3),并考虑无人机自身未建模动态 $\Delta_i$ ,时变有界大气扰动 $\kappa_i(t) = [\kappa_{xi}, \kappa_{yi}, \kappa_{hi}]^T$ ,其中,  $|\kappa_i(t)| \leq \kappa_M, \kappa_M = [\kappa_{xM}, \kappa_{yM}, \kappa_{hM}]^T$ 为正常值向量,可以将无人机模型描述为如下向量形式:

$$\begin{cases} \dot{p}_i = v_i \\ M_i \dot{v}_i = \tau_i - D_i(v_i)v_i - \Delta_i(v_i) + \kappa_i(t) \end{cases} \quad (6)$$

式中: $p_i = [x_i, y_i, h_i]^T \in \mathbf{R}^3$ 为无人机的位置向量; $v_i = [v_{xi}, v_{yi}, v_{hi}]^T \in \mathbf{R}^3$ 为无人机的速度向量; $\tau_i = [\tau_{xi}, \tau_{yi}, \tau_{hi}]^T \in \mathbf{R}^3$ 为控制输入向量; $M_i = M_i^T \in \mathbf{R}^{3 \times 3}$ 为无人机的惯性矩阵; $D_i \in \mathbf{R}^{3 \times 3}$ 为非线性阻力系数矩阵。

在无人机模型(6)中,设计控制律 $\tau_i$ ,通过控制加速度 $\dot{v}_i = [\dot{v}_{xi}, \dot{v}_{yi}, \dot{v}_{hi}]^T$ ,进而控制无人机的飞行状态。

## 1.2 输入受限模型

对于无人机的飞行控制而言,特别是当无人机需要进行较大机动来实现对路径的准确跟踪时,往往需要较大的控制输入来实现高性能的跟踪控制。但受限于无人机机动性能,实际控制输入需要限制在一定范围内;同时,过大的控制输入也可能导致执行器受损或者寿命下降。考虑输入受限是设计控制器必须考虑的问题,否则可能导致实际飞行控制系统不稳定。

考虑控制输入受限:

$$-\tau_{imax} \leq \tau_i \leq \tau_{imax} \quad (7)$$

式中: $\tau_{imax} = [\tau_{ximax}, \tau_{yimax}, \tau_{himax}]$ 为控制输入的最大值。

于是,实际的控制输入可描述为

$$\tau_i = \begin{cases} \text{sgn}(\tau_i(t)) \tau_{imax} & |\tau_i(t)| \geq \tau_{imax} \\ \tau_i(t) & |\tau_i(t)| < \tau_{imax} \end{cases} \quad (8)$$

式中: $\tau_i$ 表示受限后的输入, $\tau_i(t)$ 表示实际的输入指令。

## 1.3 图论概述

无向图由含有 $n$ 个节点的有限节点集 $V = \{1, 2, \dots, n\}$ 和含有 $m$ 条边的有限边集 $\{i, j\} \in E$ 组成。如果边 $\{i, j\}$ 包含于集合 $E$ ,则称节点 $i$ 和节点 $j$ 邻接。若存在路径连接节点 $i$ 和节点 $j$ ,则所经过的节点均连通;若任意节点 $i, j$ 连通,则图连通。定义图 $G$ 的邻接矩阵 $A = [a_{ij}] \in \mathbf{R}^{n \times n}$ , $A$ 的行列分别由节点表示,若 $\{i, j\} \in E$ 即节点 $i, j$

邻接时, $a_{ij} = 1$ ,否则 $a_{ij} = 0$ 。定义图 $G$ 的度矩阵 $\Delta = \text{diag}(A\mathbf{1}), \mathbf{1} = [1]_{n \times 1}$ ,对角线元素 $\{i, i\}$ 等于与节点 $i$ 相邻接的节点数量。定义图 $G$ 的Laplacian矩阵为 $L = \Delta - A$ , $L$ 为对称矩阵,满足 $L\mathbf{1} = 0$ ,于是0为矩阵 $L$ 的一个特征值, $\mathbf{1}$ 为对应的特征向量。对于连通图而言,其他特征值都为正值,则 $\text{rank } L = n - 1$ ,存在矩阵 $G \in \mathbf{R}^{n \times (n-1)}$ 使得 $L = GG^T$ 。对于矩阵 $G$ , $\text{rank } G = n - 1, G^T\mathbf{1} = 0$ 。

## 1.4 RBF神经网络

无人机在复杂环境飞行时常受到不确定扰动的影响,加上自身模型存在非线性未建模动态,使得无人机的飞行控制存在不稳定因素。研究表明,径向基函数(radial basis function, RBF)神经网络能够在任意精度下逼近任何非线性函数<sup>[22]</sup>。本文利用RBF神经网络在线估计,并补偿不确定扰动及未建模动态,解决了上述问题。

RBF神经网络通常由输入层、隐含层、输出层3层组成。网络的输出由加权函数 $W^T h(\xi)$ 实现, $W = [w_1, w_2, \dots, w_m]^T \in \mathbf{R}^m$ 为输出层的权值矩阵, $m$ 为输出节点个数, $h(\xi) = [h_j]^T \in \mathbf{R}^m$ 为每个隐含层的输出函数向量, $\xi = [\xi_i]^T \in \mathbf{R}^n$ 为神经网络的输入向量, $\xi_i$ 为神经网络的第 $i$ 个输入, $h_j(\xi)$ 为第 $j$ 个神经元的输出:

$$h_j = \exp\left(-\frac{\|\xi - c_j\|^2}{2b_j^2}\right) \quad (9)$$

式中: $c = [c_{ij}] = \begin{bmatrix} c_{11} & \dots & c_{1m} \\ \vdots & & \vdots \\ c_{n1} & \dots & c_{nm} \end{bmatrix}$ 为隐含层第 $j$ 个

神经元高斯基函数中心点的坐标向量, $i = 1, 2, \dots, m; b = [b_1, b_2, \dots, b_m]^T, b_j$ 为隐含层第 $j$ 个神经元高斯基函数中心点的宽度。

## 2 控制器设计及稳定性分析

多无人机协同路径跟踪控制可分为控制层和协同层2层。其中,控制层主要执行路径跟踪控制算法,实现单无人机的路径跟踪;协同层执行协调控制算法,实现路径参数、参考速度的协调一致。控制结构如图2所示。本节将依次实现控制层和协同层的控制器设计,并进行稳定性分析。

### 2.1 单无人机路径跟踪控制器设计及稳定性证明

**定义1** 路径跟踪问题<sup>[23]</sup>。定义第 $i$ 架无人机的期望路径为 $p_{di}(\theta_i(t)) = [x_{di}(\theta_i), y_{di}(\theta_i), h_{di}(\theta_i)]^T \in \mathbf{R}^3$ ,以虚拟弧长 $\theta_i \in [0; \theta_{fi}]$ 参数化描述,其中, $\theta_{fi}$ 为虚拟总弧长。路径跟踪控制的目标是:设计跟踪控制器,使得无人机的位置 $p_i(t) =$

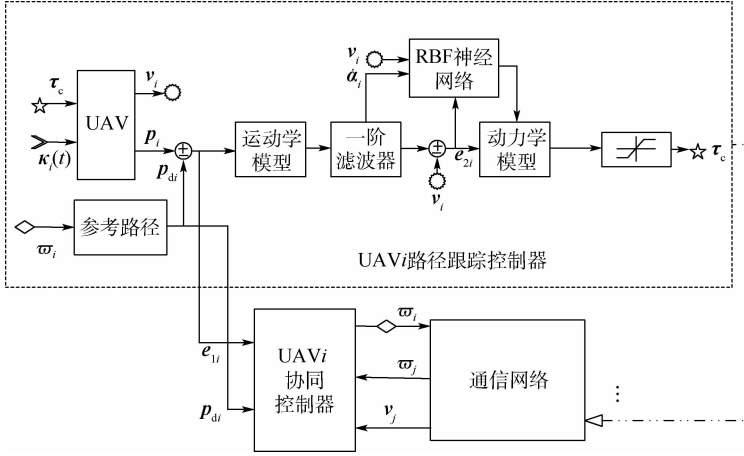


图2 控制结构

Fig.2 Control architecture

$[x_i(t), y_i(t), h_i(t)]^T \in \mathbf{R}^3$  收敛至期望路径  $p_{di}(\theta_i)$ 。值得注意的是,通过参数  $\theta_i$  对路径的描述,使得期望路径与时间无关,进而将跟踪问题转化为强调几何位置的静态路径跟踪问题。

定义跟踪误差为

$$e_{1i} = p_i - p_{di} \quad (10)$$

$$\zeta_i = \dot{\theta}_i - v_{di}(\theta_i) \quad (11)$$

式中:  $e_{1i}$  为位置跟踪误差;  $\zeta_i$  为路径参数速度跟踪误差;  $v_{di}$  为与路径参数  $\theta_i$  相关的期望跟踪速度。

对  $e_{1i}$  求导有

$$\dot{e}_{1i} = v_i - p_{di}^{\theta_i}(v_{di} + \zeta_i) \quad (12)$$

$$p_{di}^{\theta_i} = \partial p_{di} / \partial \theta_i \quad (13)$$

考虑到位置跟踪误差,定义第 1 个 Lyapunov 备选函数为

$$V_{1i} = \frac{1}{2} e_{1i}^T e_{1i} \quad (14)$$

对式(14)求导,由式(12)可得

$$\dot{V}_{1i} = e_{1i}^T v_i - e_{1i}^T p_{di}^{\theta_i}(v_{di} + \zeta_i) \quad (15)$$

引入虚拟控制量  $\alpha_i$ :

$$\alpha_i = -K_{1i} e_{1i} + p_{di}^{\theta_i} v_{di} \quad (16)$$

式中:  $K_{1i} \in \mathbf{R}^{3 \times 3}$  为正常值的对角矩阵。

定义虚拟速度误差为

$$e_{2i} = v_i - \alpha_i \quad (17)$$

则

$$\dot{V}_{1i} = -e_{1i}^T K_{1i} e_{1i} + e_{1i}^T e_{2i} - e_{1i}^T p_{di}^{\theta_i} \zeta_i \quad (18)$$

在反步设计中,求  $\dot{\alpha}_i$  时会出现微分爆炸现象,通过采用低通滤波器求  $\dot{\alpha}_i$  可克服这一缺点。虚拟控制量  $\alpha_i$  为  $\bar{x}_{2i}$  通过低通滤波器  $1/(s+1)$  的输出,定义  $\bar{x}_{2i} = -K_{1i} e_{1i} + p_{di}^{\theta_i} v_{di}$ , 有

$$s \dot{\alpha}_i + \alpha_i = \bar{x}_{2i} \quad (19)$$

所产生的滤波误差  $q_i = \alpha_i - \bar{x}_{2i}$ 。考虑到滤波误差,定义第 2 个 Lyapunov 备选函数为

$$V_{2i} = V_{1i} + \frac{1}{2} q_i^T q_i \quad (20)$$

对  $V_{2i}$  求导,由式(18)有

$$\dot{V}_{2i} = -e_{1i}^T K_{1i} e_{1i} + e_{1i}^T e_{2i} - e_{1i}^T p_{di}^{\theta_i} \zeta_i + q_i^T \dot{q}_i \quad (21)$$

针对无人机在实际飞行控制过程中存在的输入饱和现象,由式(8)设计如下辅助系统:

$$\dot{w}_i = \begin{cases} -K_{wi} w_i - \frac{f_{1i}}{\|w_i\|^2} + K_{3i} \Delta \tau_i & \|w_i\| \geq w_0 \\ 0 & \|w_i\| < w_0 \end{cases} \quad (22)$$

式中:  $w_i \in \mathbf{R}^3$  为该辅助系统的状态变量;  $K_{wi} \in \mathbf{R}^{3 \times 3}$ 、 $K_{3i} \in \mathbf{R}^{3 \times 3}$  为正常值的对角矩阵;  $w_0$  为设计参数;  $f_{1i} = (|S_i^T \Delta \tau_i| + 0.5 \Delta \tau_i^T \Delta \tau_i) w_i$ ,  $S_i = e_{2i}$  为滑模函数,  $\Delta \tau_i = \tau_i - \tau_{ci}$ ,  $\tau_i$  为实际控制输入,  $\tau_{ci}$  为期望控制输入。该辅助系统会在系统出现输入饱和时实现内部调节。当输入饱和消除时,  $\Delta \tau_i = 0$ ,  $\dot{w}_i = -K_{wi} w_i$ , 状态变量  $w_i$  将指数衰减至  $\|w_i\| < w_0$ , 使得  $\dot{w}_i = 0$ ; 当饱和再次出现时,  $\|w_i\| \geq w_0$ , 辅助系统将再次调节。

由式(6)和式(17)可得

$$M_i \dot{e}_{2i} = \tau_i - D_i(v_i) v_i - \Delta_i(v_i) + \kappa_i(t) - M_i \dot{\alpha}_i \quad (23)$$

考虑到虚拟速度误差和辅助系统影响,定义第 3 个 Lyapunov 备选函数为

$$V_{3i} = V_{2i} + \frac{1}{2} e_{2i}^T M_i e_{2i} + \frac{1}{2} w_i^T w_i \quad (24)$$

对式(24)求导,由式(20)~式(23)可得

$$\begin{aligned} \dot{V}_{3i} = & -e_{1i}^T K_{1i} e_{1i} - e_{1i}^T p_{di}^{\theta_i} \zeta_i + q_i^T \dot{q}_i - \\ & |e_{2i}^T \Delta \tau_i| + w_i^T (K_{wi} - 0.5 K_{3i}) w_i + \\ & e_{2i}^T (e_{1i} - D_i(v_i) v_i - \Delta_i(v_i) - M_i \dot{\alpha}_i) - \\ & |e_{2i}^T \kappa_i(t)| + e_{2i}^T \tau_i \end{aligned} \quad (25)$$

由双曲正切函数  $\tanh(\cdot)$  的性质有:  $|\mathbf{x}_1| - \mathbf{x}_2 \tanh(\mathbf{x}_1/\mathbf{x}_2) \leq 0.2478\mathbf{x}_2, \mathbf{x}_1, \mathbf{x}_2 \in \mathbf{R}^3$ , 代入式(25)有

$$\begin{aligned} \dot{V}_{3i} \leq & -\mathbf{e}_{1i}^T \mathbf{K}_{1i} \mathbf{e}_{1i} - \mathbf{e}_{1i}^T \mathbf{p}_{di}^{\theta_i} \zeta_i + \mathbf{q}_i^T \dot{\mathbf{q}}_i + \\ & \mathbf{e}_{2i}^T (\mathbf{e}_{1i} + \boldsymbol{\tau}_i - \mathbf{f}_i(\cdot)) - |\mathbf{e}_{2i}^T \Delta \boldsymbol{\tau}_i| - \\ & \mathbf{w}_i^T (\mathbf{K}_{wi} - 0.5\mathbf{K}_{3i}) \mathbf{w}_i + 0.2478\boldsymbol{\rho}^T \boldsymbol{\kappa}_M \end{aligned} \quad (26)$$

式中:  $\boldsymbol{\rho} = [\rho_1, \rho_2, \rho_3]^T \in \mathbf{R}^3$  为正常数。

考虑到  $\Delta \boldsymbol{\tau}_i = \boldsymbol{\tau}_i - \boldsymbol{\tau}_{ci}$ , 期望控制律为

$$\boldsymbol{\tau}_{ci} = -\mathbf{e}_{1i} - \mathbf{K}_{4i}(\mathbf{e}_{2i} - \mathbf{w}_i) + \mathbf{f}_i(\cdot) \quad (27)$$

$$\begin{aligned} \mathbf{f}_i(\cdot) = & \mathbf{D}_i(\mathbf{v}_i)\mathbf{v}_i + \Delta_i(\mathbf{v}_i) + \mathbf{M}_i \dot{\boldsymbol{\alpha}}_i - \\ & \tanh(\mathbf{e}_{2i}) \boldsymbol{\kappa}_i(t) \end{aligned} \quad (28)$$

式中:  $\mathbf{K}_{4i} \in \mathbf{R}^{3 \times 3}$  为对角线元素为正常值的对角阵。

**注1** 文献[18,24]均未考虑无人机执行器的输入限制, 本文考虑真实无人机存在的执行器输入受限非线性因素, 在反步法设计的基础上引入辅助控制系统, 使得控制器的设计更加符合实际无人机。

由于  $\mathbf{f}_i(\cdot)$  为未知非线性向量, 本文采用RBF神经网络去在线估计  $\mathbf{f}_i(\cdot)$ :

$$\mathbf{f}_i(\cdot) = \mathbf{W}_i^T \mathbf{h}(\boldsymbol{\xi}_i) + \boldsymbol{\varepsilon}_i \quad (29)$$

式中:  $\boldsymbol{\xi}_i = [1, \mathbf{v}_i^T, \boldsymbol{\alpha}_i^T]^T \in \mathbf{R}^7$  为第  $i$  架无人机控制器中神经网络的输入;  $\mathbf{W}_i$  为神经网络的权值;  $\boldsymbol{\varepsilon}_i$  为误差估计值, 满足  $\boldsymbol{\varepsilon}_i \leq \boldsymbol{\varepsilon}_M, \boldsymbol{\varepsilon}_M$  为正常数向量。

选择控制律为

$$\boldsymbol{\tau}_{ci} = -\mathbf{e}_{1i} - \mathbf{K}_{4i}(\mathbf{e}_{2i} - \mathbf{w}_i) + \hat{\mathbf{W}}_i^T \mathbf{h}(\boldsymbol{\xi}_i) \quad (30)$$

式中:  $\hat{\mathbf{W}}_i$  为  $\mathbf{W}_i$  的估计值, 其自适应律为

$$\dot{\hat{\mathbf{W}}}_i = \Gamma_W [-\mathbf{h}(\boldsymbol{\xi}_i) \mathbf{e}_{2i}^T - k_w \hat{\mathbf{W}}_i] \quad (31)$$

其中:  $\Gamma_W \in \mathbf{R}, k_w \in \mathbf{R}$  为正常值。

神经网络逼近误差为  $\tilde{\mathbf{W}}_i^T = \hat{\mathbf{W}}_i^T - \mathbf{W}_i^T$ , 代入式(26)有

$$\begin{aligned} \dot{V}_{3i} \leq & -\mathbf{e}_{1i}^T \mathbf{K}_{1i} \mathbf{e}_{1i} - \mathbf{e}_{2i}^T \mathbf{K}_{4i} \mathbf{e}_{2i} - \mathbf{e}_{1i}^T \mathbf{p}_{di}^{\theta_i} \zeta_i + \\ & \mathbf{e}_{2i}^T \mathbf{K}_{4i} \mathbf{w}_i + \mathbf{e}_{2i}^T \tilde{\mathbf{W}}_i^T \mathbf{h}(\boldsymbol{\xi}_i) + \mathbf{q}_i^T \dot{\mathbf{q}}_i - \\ & \mathbf{w}_i^T (\mathbf{K}_{wi} - 0.5\mathbf{K}_{3i}) \mathbf{w}_i - \mathbf{e}_{2i}^T \boldsymbol{\varepsilon}_i + 0.2478\boldsymbol{\rho}^T \boldsymbol{\kappa}_M \end{aligned} \quad (32)$$

考虑到神经网络逼近误差, 定义第4个 Lyapunov 备选函数为

$$V_{4i} = V_{3i} + \frac{1}{2} \text{tr}(\tilde{\mathbf{W}}_i^T \Gamma_W^{-1} \tilde{\mathbf{W}}_i) \quad (33)$$

对式(33)求导有

$$\begin{aligned} \dot{V}_{4i} \leq & -\mathbf{e}_{1i}^T \mathbf{K}_{1i} \mathbf{e}_{1i} - \mathbf{e}_{2i}^T \mathbf{K}_{4i} \mathbf{e}_{2i} - \mathbf{e}_{1i}^T \mathbf{p}_{di}^{\theta_i} \zeta_i + \\ & \mathbf{e}_{2i}^T \mathbf{K}_{4i} \mathbf{w}_i - k_w \text{tr}(\tilde{\mathbf{W}}_i^T \tilde{\mathbf{W}}_i) + \mathbf{q}_i^T \dot{\mathbf{q}}_i - \end{aligned}$$

$$\mathbf{w}_i^T (\mathbf{K}_{wi} - 0.5\mathbf{K}_{3i}) \mathbf{w}_i - \mathbf{e}_{2i}^T \boldsymbol{\varepsilon}_i + 0.2478\boldsymbol{\rho}^T \boldsymbol{\kappa}_M \quad (34)$$

**注2** 文献[24]未考虑无人机在实际飞行中可能遇到的各类不确定扰动, 文献[3]仅考虑了恒定风对无人机飞行的影响。本文则充分考虑无人机在实际飞行中可能遇到的时变未知扰动及无人机模型自身存在的非线性动态, 文献[3,24]所设计控制算法无法有效解决上述问题。

设计路径参数  $\theta_i$ 。考虑到路径参数速度跟踪误差  $\zeta_i = \dot{\theta}_i - v_{di}(\theta_i)$ , 设计第5个 Lyapunov 备选函数为

$$V_{5i} = V_{4i} + \frac{1}{2\lambda\beta} \zeta_i^2 \quad (35)$$

式中:  $\lambda, \beta$  为正常数。

对式(35)求导有

$$\begin{aligned} \dot{V}_{5i} \leq & -\mathbf{e}_{1i}^T \mathbf{K}_{1i} \mathbf{e}_{1i} - \mathbf{e}_{2i}^T \mathbf{K}_{4i} \mathbf{e}_{2i} + \mathbf{q}_i^T \dot{\mathbf{q}}_i - \\ & \left( \mathbf{e}_{1i}^T \mathbf{p}_{di}^{\theta_i} - \frac{1}{\lambda\beta} \zeta_i \right) \zeta_i + \mathbf{e}_{2i}^T \mathbf{K}_{4i} \mathbf{w}_i - \\ & k_w \text{tr}(\tilde{\mathbf{W}}_i^T \tilde{\mathbf{W}}_i) + 0.2478\boldsymbol{\rho}^T \boldsymbol{\kappa}_M - \\ & \mathbf{w}_i^T (\mathbf{K}_{wi} - 0.5\mathbf{K}_{3i}) \mathbf{w}_i - \mathbf{e}_{2i}^T \boldsymbol{\varepsilon}_i \end{aligned} \quad (36)$$

取  $\dot{\zeta}_i = \lambda\beta \mathbf{e}_{1i}^T \mathbf{p}_{di}^{\theta_i} - \lambda\zeta_i$ , 则

$$\begin{aligned} \dot{V}_{5i} \leq & -\mathbf{e}_{1i}^T \mathbf{K}_{1i} \mathbf{e}_{1i} - \mathbf{e}_{2i}^T \mathbf{K}_{4i} \mathbf{e}_{2i} - \frac{1}{\beta} \zeta_i^2 + \\ & \mathbf{e}_{2i}^T \mathbf{K}_{4i} \mathbf{w}_i - k_w \text{tr}(\tilde{\mathbf{W}}_i^T \tilde{\mathbf{W}}_i) + \mathbf{q}_i^T \dot{\mathbf{q}}_i - \\ & \mathbf{w}_i^T (\mathbf{K}_{wi} - 0.5\mathbf{K}_{3i}) \mathbf{w}_i - \mathbf{e}_{2i}^T \boldsymbol{\varepsilon}_i + 0.2478\boldsymbol{\rho}^T \boldsymbol{\kappa}_M \end{aligned} \quad (37)$$

由于  $\dot{\zeta}_i = \ddot{\theta}_i - v_{di}^{\theta_i}(t)$ , 可得  $\theta_i$  的反馈控制率为

$$\ddot{\theta}_i = \lambda\beta \mathbf{e}_{1i}^T \mathbf{p}_{di}^{\theta_i} - \lambda\zeta_i + v_{di}^{\theta_i}(t) \quad (38)$$

为了证明路径跟踪控制系统的稳定性, 给出如下定理。

**定理1** 考虑由虚拟控制器(16)、辅助系统(22)、实际控制律(30)、RBF参数自适应律(31)、反馈控制律(38)及滤波器(19)组成的闭环控制系统, 对于给定的任意  $\iota_1 > 0$ , 若闭环系统满足初始条件:  $\Omega = \{[\mathbf{e}_{1i}, \mathbf{e}_{2i}, \tilde{\mathbf{W}}_i, \mathbf{q}_i]^T : S \leq \iota_1\}$ , 则存在可调整的控制参数  $\mathbf{K}_{1i}, \mathbf{K}_{wi}, \mathbf{K}_{3i}, \mathbf{K}_{4i}, \varsigma, \Gamma_W, k_w$ , 使得该闭环系统的所有信号半全局一致有界, 同时位置跟踪误差  $\mathbf{e}_{1i} = \mathbf{p}_i - \mathbf{p}_{di}$ 、路径参数速度跟踪误差  $\zeta_i = \dot{\theta}_i - v_{di}(\theta_i)$  均收敛于任意正常数。

**证明** 由式(19)可得

$$\dot{\mathbf{q}}_i = -\frac{1}{\varsigma} \mathbf{q}_i + \mathbf{H}_i(\cdot) \quad (39a)$$

式中:  $\mathbf{H}_i(\cdot) \leq \mathbf{H}_i(\cdot)_M$  为  $\mathbf{e}_{1i}, \mathbf{e}_{2i}, \mathbf{q}_i$  的有界函数。

利用 Young's 不等式有

$$\begin{cases} \mathbf{e}_{2i}^T \mathbf{K}_{4i} \mathbf{w}_i \leq \frac{1}{2} \lambda_{\min}^2(\mathbf{K}_{4i}) \|\mathbf{e}_{2i}\|^2 + \frac{1}{2} \|\mathbf{w}_i\|^2 \\ -k_w \text{tr}(\tilde{\mathbf{W}}_i^T \tilde{\mathbf{W}}_i) \leq -\frac{k_w}{2} \|\tilde{\mathbf{W}}_i\|_F^2 + \|\mathbf{W}_i\|_F^2 \\ \mathbf{q}_i^T \dot{\mathbf{q}}_i \leq -\frac{1}{\zeta} \|\mathbf{q}_i\|^2 + \frac{1}{2} \|\mathbf{q}_i\|^2 + \frac{1}{2} \|\mathbf{H}_i\|^2 \\ -\mathbf{e}_{2i}^T \boldsymbol{\varepsilon}_i \leq \frac{1}{2} \|\mathbf{e}_{2i}\|^2 + \frac{1}{2} \|\boldsymbol{\varepsilon}_M\|^2 \end{cases} \quad (39b)$$

式中:  $\|\cdot\|_F$  为  $F$  范数。

于是,取  $S$  为式(37),可写为

$$\begin{aligned} \dot{V}_{5i} \leq & -\mathbf{e}_{1i}^T \mathbf{K}_{1i} \mathbf{e}_{1i} - \mathbf{e}_{2i}^T \mathbf{K}_{4i} \mathbf{e}_{2i} - \frac{1}{\beta} \zeta_i^2 + \\ & \frac{1}{2} \lambda_{\min}^2(\mathbf{K}_{4i}) \|\mathbf{e}_{2i}\|^2 + \frac{1}{2} \|\mathbf{w}_i\|^2 - \\ & \frac{k_w}{2} \|\tilde{\mathbf{W}}_i\|_F^2 + \frac{k_w}{2} \|\mathbf{W}_i\|_F^2 - \frac{1}{\zeta} \|\mathbf{q}_i\|^2 + \frac{1}{2} \|\mathbf{q}_i\|^2 + \\ & \frac{1}{2} \|\mathbf{H}_i\|^2 + \frac{1}{2} \|\mathbf{e}_{2i}\|^2 + \frac{1}{2} \|\boldsymbol{\varepsilon}_M\|^2 - \\ & \mathbf{w}_i^T (\mathbf{K}_{wi} - 0.5 \mathbf{K}_{3i}) \mathbf{w}_i + 0.247 \, 8 \boldsymbol{\rho}^T \boldsymbol{\kappa}_M \leq \\ & -\lambda_{\min}(\mathbf{K}_{1i}) \|\mathbf{e}_{1i}\|^2 - \left[ \frac{1}{\zeta} - \frac{1}{2} \right] \|\mathbf{q}_i\|^2 - \\ & \left[ \lambda_{\min}(\mathbf{K}_{wi} - 0.5 \mathbf{K}_{3i}) - \frac{1}{2} \right] \|\mathbf{w}_i\|^2 - \frac{k_w}{2} \|\tilde{\mathbf{W}}_i\|_F^2 + \\ & \Theta_i(\cdot) - \left[ \lambda_{\min}(\mathbf{K}_{4i}) - \frac{1}{2} \lambda_{\min}^2(\mathbf{K}_{4i}) - \frac{1}{2} \right] \|\mathbf{e}_{2i}\|^2 \end{aligned} \quad (39c)$$

式中:  $\Theta_i(\cdot) = 0.247 \, 8 \boldsymbol{\rho}^T \boldsymbol{\kappa}_M + \frac{1}{2} \|\boldsymbol{\varepsilon}_M\|^2 + \frac{k_w}{2} \|\mathbf{W}_i\|_F^2$ 。

选取合适的  $\mathbf{K}_{1i}, \mathbf{K}_{wi}, \mathbf{K}_{3i}, \mathbf{K}_{4i}, \zeta, k_w$  使得  $\lambda_{\min}(\mathbf{K}_{1i}) \geq 0, \lambda_{\min}(\mathbf{K}_{4i}) - \frac{1}{2} \lambda_{\min}^2(\mathbf{K}_{4i}) - \frac{1}{2} \geq 0, \frac{1}{\zeta} - \frac{1}{2} \geq 0,$

$\lambda_{\min}(\mathbf{K}_{wi} - 0.5 \mathbf{K}_{3i}) - \frac{1}{2} \geq 0$ 。注意到,  $\|\mathbf{e}_{1i}\| \geq$

$$\sqrt{\frac{\Theta_i(\cdot)}{\lambda_{\min}(\mathbf{K}_{1i})}}, \|\mathbf{e}_{2i}\| \geq \sqrt{\frac{\Theta_i(\cdot)}{\lambda_{\min}(\mathbf{K}_{4i}) - \frac{1}{2} \lambda_{\min}^2(\mathbf{K}_{4i}) - \frac{1}{2}}},$$

$$\|\mathbf{q}_i\| \geq \sqrt{\frac{2\zeta \Theta_i(\cdot)}{2 - \zeta}}, \|\mathbf{w}_i\| \geq \sqrt{\frac{\Theta_i(\cdot)}{\lambda_{\min}(\mathbf{K}_{wi} - 0.5 \mathbf{K}_{3i}) - \frac{1}{2}}}$$

或  $\|\tilde{\mathbf{W}}_i\|_F \geq \sqrt{\frac{2\Theta_i(\cdot)}{k_w}}$  使得  $\dot{V}_{5i} \leq 0$ , 则该闭环系统的所有信号半全局一致有界。 证毕

### 2.2 协同控制器设计及稳定性证明

本节依据 2.1 节路径跟踪控制器,设计编队协同控制器,实现多无人机的协同路径跟踪控制。

协同路径跟踪控制问题可以通过协调各智能体的速度,以保持路径参数的协同一致,进而实现

多智能体的固定队型<sup>[25]</sup>。对于多无人机系统而言,一致性问题是最为基础的问题,指的是各无人机的状态最终收敛于相同的静态或者动态的值。本文中一致性指的是多无人机系统中各无人机的速度及路径参数最终协同一致。由文献[26]启发,为了解决上述问题,可以通过定义参数  $\boldsymbol{\omega}_i$  再参数化  $\mathbf{p}_{di}(\theta_i)$  实现,则当  $\boldsymbol{\omega}_1 = \boldsymbol{\omega}_2 = \dots = \boldsymbol{\omega}_n$  时,实现了协同控制。定义  $\boldsymbol{\omega}_{i,j} = \boldsymbol{\omega}_i - \boldsymbol{\omega}_j$  为第  $i$  架无人机和第  $j$  架无人机路径参数化距离,则实现协同控制的条件转化为  $\boldsymbol{\omega}_{i,j} = 0, i, j \in \{1, 2, \dots, n\}$ 。对于路径  $\mathbf{p}_{di}(\theta_i)$  的再参数化,由  $\theta_i = \theta_i(\boldsymbol{\omega}_i)$  表示,定义  $R_i(\boldsymbol{\omega}_i) = \partial \theta_i / \partial \boldsymbol{\omega}_i$ 。  $\boldsymbol{\omega}_i$  的动态响应为

$$\dot{\boldsymbol{\omega}}_i = \dot{\theta}_i / R_i(\boldsymbol{\omega}_i) \quad (40)$$

选择虚拟长机  $l$ , 其对应期望路径由路径参数  $\theta_l = \boldsymbol{\omega}_l$  参数化为  $\mathbf{p}_{dl}$ 。对于  $l, R_l(\boldsymbol{\omega}_l) = 1$ , 期望路径参考速度为  $v_l = \dot{\boldsymbol{\omega}}_l$ 。由式(40),各无人机的期望速度为  $v_{di} = R_i(\boldsymbol{\omega}_i) v_l$ 。定义速度跟踪误差为

$$\zeta_i = \dot{\theta}_i - R_i(\boldsymbol{\omega}_i) v_l \quad (41)$$

其动态方程为

$$\dot{\zeta}_i = u_i = \ddot{\theta}_i - \frac{d}{dt}(R_i(\boldsymbol{\omega}_i) v_l) \quad (42)$$

由式(41)和式(40)可写为

$$\dot{\boldsymbol{\omega}}_i = \zeta_i / R_i(\boldsymbol{\omega}_i) + v_l \quad (43)$$

将式(42)和式(43)写为向量形式:

$$\begin{cases} \dot{\boldsymbol{\zeta}} = \mathbf{u} \\ \dot{\boldsymbol{\omega}} = \mathbf{R}(\boldsymbol{\omega})^{-1} \boldsymbol{\zeta} + v_l \mathbf{1} \end{cases} \quad (44)$$

为了实现协同控制,需要设计控制律  $\mathbf{u}$ , 使得  $\boldsymbol{\omega}_1 = \boldsymbol{\omega}_2 = \dots = \boldsymbol{\omega}_n$ 。从图论知识可以看出,每架无人机可以由节点表示,无人机间的通信联系可以由相应节点的边来表示。节点  $i$  的邻接节点集合包含了与无人机  $i$  通信的所有无人机。为了满足网络节点的通信限制,控制器的设计仅基于自身状态或者与其邻接无人机的交互信息。无人机间的通信关系网络如图 3 所示。

定义协同误差  $\boldsymbol{\eta} = \mathbf{G}^T \boldsymbol{\omega} \in \mathbf{R}^{n-1}$ , 其中矩阵  $\mathbf{G}$  为 1.3 节根据 Laplacian 矩阵得到的特征矩阵,以此实现多无人机间的网络通信。 $\boldsymbol{\eta} = \mathbf{0}$  等效于  $\boldsymbol{\omega}_{i,j} = 0, i, j \in \{1, 2, \dots, n\}$ 。由 1.3 节知识,式(44)可改写为

$$\begin{cases} \dot{\boldsymbol{\zeta}} = \mathbf{u} \\ \dot{\boldsymbol{\eta}} = \mathbf{G}^T \mathbf{R}(\boldsymbol{\omega})^{-1} \boldsymbol{\zeta} \end{cases} \quad (45)$$

注 3 式(44)和式(45)给出了本文协同控制的基本含义,  $\boldsymbol{\zeta}$  表示各无人机在考虑通信网络后的速度跟踪误差,  $\boldsymbol{\eta}$  表示考虑通信网络后的路径参数协同误差,当所设计的控制律使得  $\boldsymbol{\zeta}, \boldsymbol{\eta}$  渐

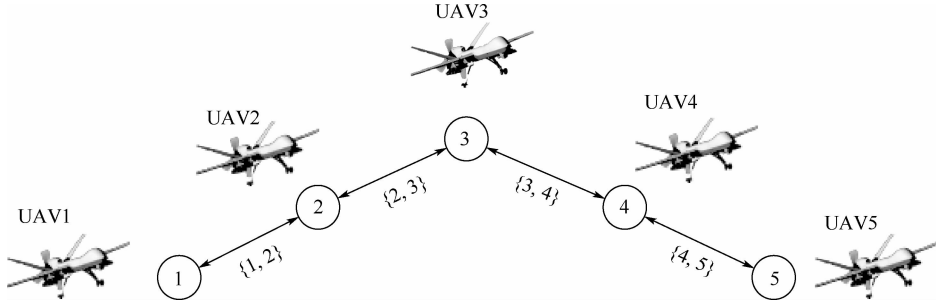


图3 无人机编队通信网络

Fig. 3 Graph induced by communication network of UAVs

近趋于0时,即可实现无人机编队速度意义、路径参数在时间上的协同一致。

**定义2** 协同控制问题<sup>[25]</sup>。编队参考速度取为虚拟长机速度  $v_l$ , 假设所有期望路径  $\mathbf{p}_{di}(\theta_i)$  光滑且其对  $\theta_i$  的导数有界。定义协同控制问题为设计  $\mathbf{u}$  的反馈控制律, 其中  $u_i$  为自身状态和参数  $\boldsymbol{\omega}_j, j \in N_i$  的函数, 使得  $\boldsymbol{\zeta}, \boldsymbol{\eta}$  渐近趋于0。

基于2.1节路径跟踪控制器设计思路, 考虑速度跟踪误差和协同误差, 定义 Lyapunov 备选函数  $V_c$  为

$$V_c = \sum_{i=1}^n \frac{d_i}{e_i} V_{4i} + \frac{1}{2} \boldsymbol{\eta}^T \boldsymbol{\eta} + \frac{1}{2} \boldsymbol{\zeta}^T \mathbf{E}^{-1} \boldsymbol{\zeta} \quad (46)$$

式中:  $\mathbf{E} = \text{diag}[e_i]_{n \times 1}, e_i > 0; \mathbf{D} = \text{diag}[d_i]_{n \times 1}, d_i > 0$ 。

令  $\boldsymbol{\mu}_i = -\mathbf{e}_{1i}^T \mathbf{p}_{di}^{\theta_i} \boldsymbol{\zeta}_i, \boldsymbol{\mu} = [\boldsymbol{\mu}_i]_{n \times 1}$ , 对式(46)求导有

$$\begin{aligned} \dot{V}_c \leq & \boldsymbol{\zeta}^T (\mathbf{E}^{-1} \mathbf{u} + \mathbf{R}^{-1} \mathbf{G}^T \boldsymbol{\eta} + \mathbf{E}^{-1} \mathbf{D} \boldsymbol{\mu}) - \\ & \sum_{i=1}^n \frac{d_i}{e_i} (\mathbf{e}_{1i}^T \mathbf{K}_{1i} \mathbf{e}_{1i} + \mathbf{e}_{2i}^T \mathbf{K}_{4i} \mathbf{e}_{2i} - \mathbf{e}_{2i}^T \mathbf{K}_{4i} \mathbf{w}_i + \\ & k_w \text{tr}(\tilde{\mathbf{W}}_i^T \tilde{\mathbf{W}}_i) - \mathbf{q}_i^T \dot{\mathbf{q}}_i + \mathbf{w}_i^T (\mathbf{K}_{wi} - 0.5 \mathbf{K}_{3i}) \mathbf{w}_i + \\ & \mathbf{e}_{2i}^T \boldsymbol{\varepsilon}_i - 0.247 \, 8 \boldsymbol{\rho}^T \boldsymbol{\kappa}_M) \end{aligned} \quad (47)$$

设计  $\mathbf{u}$  的反馈控制律为

$$\mathbf{u} = -\mathbf{E} \mathbf{R}^{-1} \mathbf{G}^T \boldsymbol{\eta} - \mathbf{D} \boldsymbol{\mu} - \mathbf{B} \boldsymbol{\zeta} \quad (48)$$

式中:  $\mathbf{B} = \text{diag}[b_i]_{n \times 1}, b_i > 0$ 。

代入式(47)有

$$\begin{aligned} \dot{V}_c \leq & - \sum_{i=1}^n \frac{d_i}{e_i} (\mathbf{e}_{1i}^T \mathbf{K}_{1i} \mathbf{e}_{1i} + \mathbf{e}_{2i}^T \mathbf{K}_{4i} \mathbf{e}_{2i} - \mathbf{e}_{2i}^T \mathbf{K}_{4i} \mathbf{w}_i + \\ & k_w \text{tr}(\tilde{\mathbf{W}}_i^T \tilde{\mathbf{W}}_i) - \mathbf{q}_i^T \dot{\mathbf{q}}_i + \mathbf{w}_i^T (\mathbf{K}_{wi} - 0.5 \mathbf{K}_{3i}) \mathbf{w}_i + \\ & \mathbf{e}_{2i}^T \boldsymbol{\varepsilon}_i - 0.247 \, 8 \boldsymbol{\rho}^T \boldsymbol{\kappa}_M) - \boldsymbol{\zeta}^T \mathbf{E}^{-1} \mathbf{B} \boldsymbol{\zeta} \end{aligned} \quad (49)$$

于是, 由式(47)可得考虑通信网络限制的分布式控制律  $u_i$  为

$$u_i = -d_i \boldsymbol{\mu}_i - b_i \boldsymbol{\zeta}_i - \frac{e_i}{R_i} \sum_{j \in N_i} (\boldsymbol{\omega}_i - \boldsymbol{\omega}_j) \quad (50)$$

为了证明协同路径跟踪控制器的稳定性, 给出如下定理。

**定理2** 考虑由虚拟控制器(16)、辅助系

统(22)、RBF 参数自适应律(31)、反馈控制律(48)、分布式控制律(50)及滤波器(19)组成的闭环控制系统, 对于给定的任意  $\iota_2 > 0$ , 若闭环系统满足初始条件:  $\Omega = \{[\mathbf{e}_{1i}, \mathbf{e}_{2i}, \tilde{\mathbf{W}}_i, \mathbf{q}_i, \boldsymbol{\zeta}, \boldsymbol{\eta}]^T : S \leq \iota_2\}$ , 则存在可调整的控制参数  $\mathbf{K}_{1i}, \mathbf{K}_{wi}, \mathbf{K}_{3i}, \mathbf{K}_{4i}, \varsigma, \Gamma_w, k_w, \mathbf{E}, \mathbf{B}$ , 使得该闭环系统的所有信号半全局一致有界, 同时位置跟踪误差  $\mathbf{e}_{1i} = \mathbf{p}_i - \mathbf{p}_{di}$ , 路径参数速度跟踪误差  $\boldsymbol{\zeta}_i = \dot{\theta}_i - v_{di}(\theta_i)$ , 协同误差  $\boldsymbol{\eta} = \mathbf{G}^T \boldsymbol{\omega}$  均收敛于任意正常数。

**证明** 由定理1证明推导过程及式(39c)、式(49)可为

$$\begin{aligned} \dot{V}_c \leq & - \sum_{i=1}^n \frac{d_i}{e_i} \left( -\lambda_{\min}(\mathbf{K}_{1i}) \|\mathbf{e}_{1i}\|^2 - \right. \\ & \left[ \lambda_{\min}(\mathbf{K}_{4i}) - \frac{1}{2} \lambda_{\min}^2(\mathbf{K}_{4i}) - \frac{1}{2} \right] \|\mathbf{e}_{2i}\|^2 - \\ & \left[ \frac{1}{\varsigma} - \frac{1}{2} \right] \|\mathbf{q}_i\|^2 - \frac{k_w}{2} \|\tilde{\mathbf{W}}_i\|_F^2 + \Theta_i(\cdot) - \\ & \left. \left[ \lambda_{\min}(\mathbf{K}_{wi} - 0.5 \mathbf{K}_{3i}) - \frac{1}{2} \right] \|\mathbf{w}_i\|^2 \right) - \boldsymbol{\zeta}^T \mathbf{E}^{-1} \mathbf{B} \boldsymbol{\zeta} \end{aligned} \quad (51)$$

同样的, 通过选取合适的参数  $\mathbf{K}_{1i}, \mathbf{K}_{wi}, \mathbf{K}_{3i}, \mathbf{K}_{4i}, \varsigma, \Gamma_w, k_w, \mathbf{E}, \mathbf{B}$ , 满足  $\mathbf{E}^{-1} \mathbf{B} \geq 0, \lambda_{\min}(\mathbf{K}_{1i}) \geq 0, \lambda_{\min}(\mathbf{K}_{4i}) - \frac{1}{2} \lambda_{\min}^2(\mathbf{K}_{4i}) - \frac{1}{2} \geq 0, \frac{1}{\varsigma} - \frac{1}{2} \geq 0, \lambda_{\min}(\mathbf{K}_{wi} - 0.5 \mathbf{K}_{3i}) - \frac{1}{2} \geq 0$ , 可以使得  $\dot{V}_c \leq 0$ , 则该闭环系统的所有信号半全局一致有界。证毕

### 3 仿真实验

为了验证本文设计控制器的有效性及稳定性分析的正确性, 利用 MATLAB/Simulink 分别对第2节设计的路径跟踪控制器及协同控制器进行仿真实验。通过对文献[24]所设计控制算法的复现, 对比验证了在考虑外界非线性时变扰动、无人机模型自身非线性动态及无人机实际存在的输入受限非线性因素后, 本文算法的有效性。

设定虚拟长机的期望路径为螺旋上升曲线:

$$p_{dl}(\theta_l) = \begin{bmatrix} a \cos \theta_l \\ a \sin \theta_l \\ \theta_l + 3 \end{bmatrix} \quad (52)$$

式中: $a$ 为期望路径水平面投影半径; $\theta_l$ 为期望路径的路径参数。

无人机的惯性矩阵  $M_i = \text{diag}[20]_3$ ,期望路径的半径  $a = 100$  m。仿真从  $t = 0$  开始,无人机的初始位置为  $p_{0i} = [80, -10, 0]^T$  m,初始速度  $\dot{p}_{0i} = [0, 0, 0]^T$  m/s。无人机自身的未建模动态为  $\Delta_i = [v_x^2 + 0.03v_y, v_y^2 + 0.03v_x, v_h^3]^T$ ,外界环境时变有界大气扰动为  $\kappa_i(t) = [0.6 \sin t, 0.6 \cos t, 0.5 \sin(t + \pi/3)]^T$ ,则  $\kappa_M = [0.6, 0.6, 0.5]^T$ ,控制输入的最大值为  $\tau_{imax} = 1\ 000$  N。

将路径跟踪控制器的基本参数设置为: $K_{1i} = \text{diag}[1, 1, 1]$ ,  $K_{wi} = \text{diag}[150, 150, 90]$ ,  $K_{3i} = \text{diag}[3, 3, 3]$ ,  $K_{4i} = \text{diag}[200, 200, 200]$ 。RBF神经网络中,  $\Gamma_w = 0.5, k_w = 0.05$ ,隐含层高斯基函数宽度  $b_j = 1.7$ 。低通滤波器的时间常数  $\tau = 0.01$ 。

假设5架无人机进行协同路径跟踪飞行,由图3所示的通信网络,根据1.3节定义,节点邻接时,  $a_{ij} = 1$ ,否则  $a_{ij} = 0$ 。则邻接矩阵为

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$E = 0.1I_3, D = 0.01I_3, B = I_3$ ,虚拟长机的路径参考速度  $v_l = 1$  m/s。

根据以上参数设置,分别进行单无人机路径跟踪仿真和多机协同路径跟踪仿真。

### 3.1 单无人机路径跟踪仿真结果分析

图4为单无人机对螺旋上升期望路径的跟踪控制效果。从跟踪效果可以看出,本文控制算法相较于文献[24]的控制算法,在考虑外界非线性扰动、无人机模型自身非线性动态及无人机实际存在的输入受限非线性因素后,能够以较为平滑的控制过程实现对期望路径的跟踪。

图5为2种控制算法在3个坐标轴方向的跟踪误差对比。可以看出,本文跟踪控制算法在较短时间内实现了路径跟踪误差的快速收敛,最终趋近于0,并且整个收敛过程平滑过渡。图6为RBF神经网络对无人机模型未建模动态及外界时变扰动的逼近效果。从仿真结果来看,RBF对所给的复杂非线性不确定部分具有良好的逼近效果,实现了对模型自身未建模动态和外界时变扰动的有效补偿。图7为考虑输入受限前后无人机

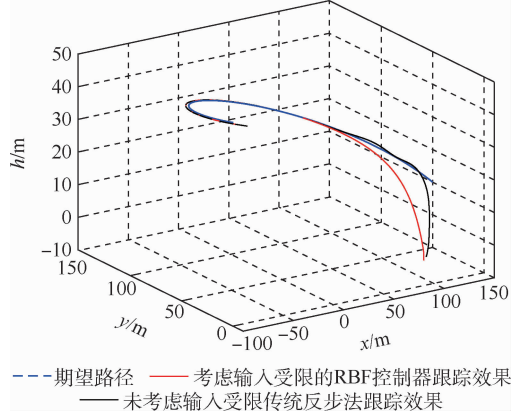


图4 考虑外界扰动后2类控制器控制效果对比  
Fig.4 Comparison of control effects of two kinds of controllers after considering external disturbance

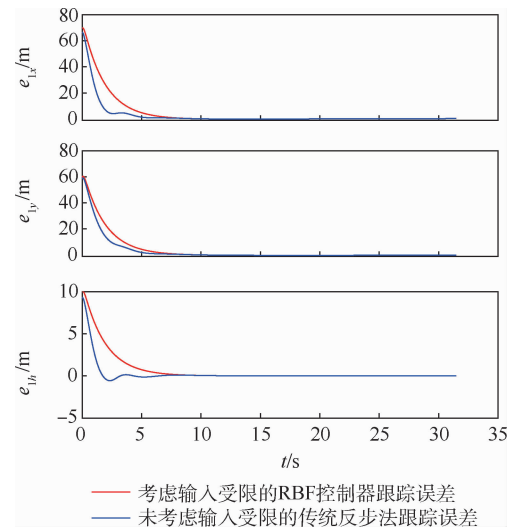


图5 考虑外界扰动后2类控制器路径跟踪误差  
Fig.5 Two kinds of controller path tracking errors after external disturbance are considered

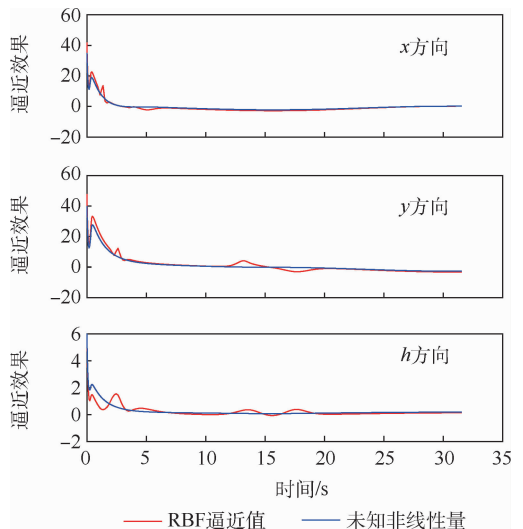


图6 神经网络对不确定项逼近效果  
Fig.6 Neural network approximation performance to uncertain terms

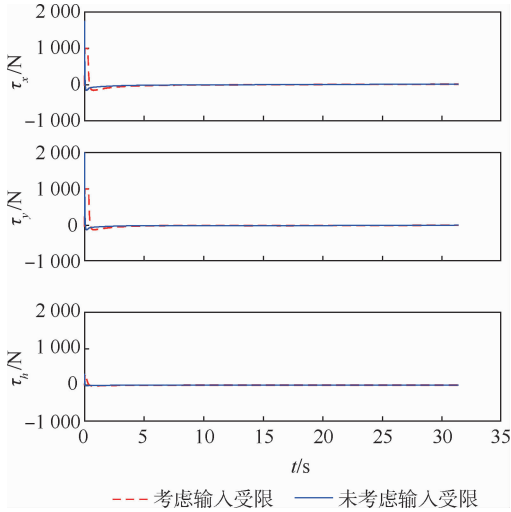


图7 考虑输入受限前后控制输入对比

Fig.7 Comparison of control input before and after considering input saturation

控制输入的对比。可以看出,考虑输入受限前控制输入最大可达到 2 000 N,考虑输入受限后对控制输入有了明显约束,保持在  $\tau_{i,max} = 1\ 000\ N$  以内,控制输入更加平稳,这也使得无人机的跟踪控制过程更为平稳。

与本文控制算法形成鲜明对比的是,文献[24]所提控制算法在考虑外界非线性时变扰动、未建模动态、控制输入受限等非线性因素后,对路径的跟踪在短时间内猛烈收敛,这是由于未考虑无人机实际存在控制输入受限约束。在跟踪过程中存在“抖动”,这是由于所涉及的非线性动态与速度高度关联,控制初期由于位置误差较大,飞行速度较大,使得非线性量  $f_i(\cdot)$  值较大。随着跟踪误差减小,飞行速度保持稳定,  $f_i(\cdot)$  开始收敛于有界值,  $f_i(\cdot)$  的剧烈变化导致该算法在控制过程中存在明显的“抖动”现象。经过上述分析,即使该算法最终能够实现对期望路径的跟踪,使得跟踪误差趋于 0,也是在本文仿真设计的非线性动态与飞行速度高度关联的条件下。若实际存在的非线性量与速度关联度小,则该算法可能无法实现对期望路径的跟踪。

本文所设计的考虑输入受限的 RBF 路径跟踪控制算法,较好地解决了上述存在的问题,使得无人机能够在考虑外界时变有界扰动、模型自身非线性未建模动态及无人机实际存在的控制输入受限非线性因素条件下,通过设计辅助控制系统,利用 RBF 神经网络估计并补偿非线性量,实现了对期望路径的平滑且稳定的跟踪控制。

### 3.2 多机协同路径跟踪仿真结果分析

对数量为 5 的多无人机编队进行协同路径跟

踪控制仿真验证。无人机编队从同一机场起飞,假设无人机的起飞间隔时间忽略不计,在协同控制器的作用下实现对路径的跟踪控制,并生成“一”字型编队。在编队生成后,通过改变编队无人机的期望路径参数,在 15 s 时,对不同无人机的路径参数做  $K_{li}(t(\text{end}) - 15)$  连续化等差递增处理,使得无人机的期望路径参数具有可控差异,进而实现多无人机的编队可控变换与重构。其中,  $K_{li}$  为实现编队重构各无人机的路径参数系数,  $t(\text{end})$  为无人机飞行实时时刻。仿真结果如图 8 ~ 图 10 所示。

从图 8 多机协同路径跟踪效果可以看出,  $t = 0\ s$  时刻各无人机从初始位置起飞后,迅速对各自路径进行跟踪,并实现协同控制,协同误差的变化如图 9 所示。由于飞行初期各无人机对各自期望路径的距离有所差异,路径参数的一致成为协同控制的主要目标。为了实现速度及路径参数的协同一致,各无人机的分布式控制器先以不同的速度实现路径参数的一致,从图 10 速度的变化可以看出,UAV1、UAV5 由于离各自期望路径最远,在短时间内速度迅速增大,而 UAV3 离期望路径最近,初始速度相对较小。从图 9 协同误差变化可

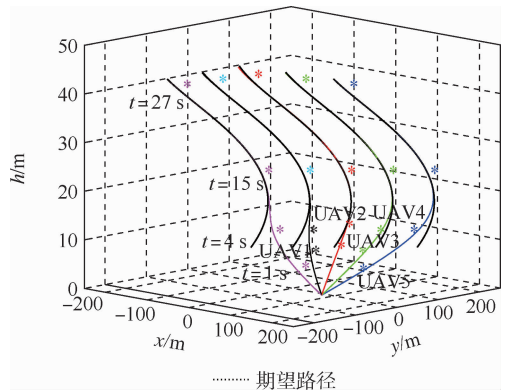


图8 协同跟踪控制效果

Fig.8 Cooperative path following result

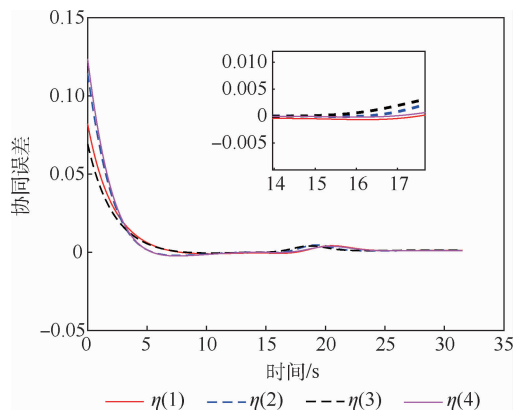


图9 协同误差

Fig.9 Path variables coordination errors

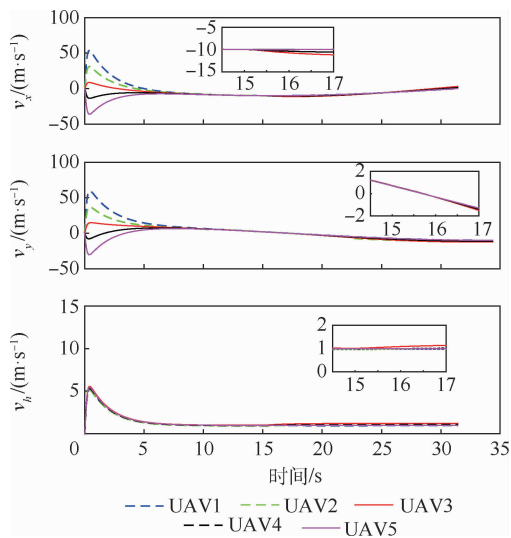


图10 编队协同飞行速度

Fig. 10 Cooperative speed of each UAV on each axis

以看出,在  $t = 4$  s 时刻,协同误差趋于 0;与此同时,在各无人机实现了对各自期望路径的跟踪,各无人机的速度一致成为此时协同控制的主要目标。从图 10 协同速度可以看出,在  $t = 9$  s 时刻,速度趋于一致,此时协同误差也收敛至 0。基于此,可以推断出多无人机编队在  $t = 9$  s 时刻实现了路径参数和飞行速度一致的协同控制,生成并保持了“一”字型编队。

在仿真验证中, $t = 15$  s 时刻对各无人机的路径参数做了等差连续化处理,以此实现多无人机的编队队形变换,生成并保持了“三角形”编队,实现了多无人机协同控制的队形重构。在此之前,各无人机已实现了对期望路径的跟踪控制,在进行队形重构时,只需要通过分布式控制器实现对速度的适当调节,以适应路径参数的等差处理,即可实现路径参数与速度一致的协同控制,并生成新队形。从图 10 编队飞行速度的变化可以看出,UAV3 作为“三角形”编队顶点,由于其路径参数增量最大,在  $t = 15$  s 时刻其飞行速度变化也相对较大;UAV2、UAV4 处于“三角形”等腰边中点处,路径参数增量较小,故而在  $t = 15$  s 时刻其速度变化较小;UAV1、UAV5 处于“三角形”底边两端,路径参数增量为 0,故其速度基本保持不变。从图 9 协同误差的变化可以看出,由于对路径参数做了等差增量处理,协同误差有所增加。通过分布式控制器对各无人机速度的调节,各无人机逐渐适应对路径参数的等差增量处理,协同误差于  $t = 24$  s 时刻趋于 0,实现了路径参数的协同一致。与此同时,各无人机速度也于  $t = 24$  s 时刻再次实现了协同一致。可以推断,自  $t = 15$  s 时刻开始队形重构后,多无人机于  $t = 24$  s 时刻实现

了路径参数与速度的协同一致,并由“一”字型编队重构为“三角形”编队。

## 4 结论

1) 为实现在三维空间中的路径跟踪控制,本文对固定翼无人机三维六自由度非线性模型反馈线性化处理,并运用反步法设计了一种无人机在三维空间中路径与时间无关的跟踪控制器。

2) 考虑无人机控制输入饱和,引入辅助控制系统,确保了在控制输入受限的非线性因素下系统仍能保持良好的控制性能。

3) 考虑无人机受外界不确定时变有界扰动和自身模型的未建模动态,利用 RBF 神经网络在线估计补偿,使控制器具有良好的鲁棒性和抗干扰能力。

4) 利用图论和一致性理论,基于路径跟踪控制器设计了多无人机的协同控制器,实现了多无人机的协同跟踪控制,最终实现了路径参数与速度一致的协同控制,并实现了编队生成、保持、重构。

## 参考文献 (References)

- [1] CHUNG S J, PARANJPE A A, DAMES P, et al. A survey on aerial swarm robotics [J]. IEEE Transactions on Robotics, 2018, 34(4): 837-855.
- [2] CABECINHAS D, CUNHA R, SILVESTRE C. Rotorcraft path following control for extended flight envelope coverage [C] // Proceedings of the IEEE Conference on Decision and Control. Piscataway: IEEE Press, 2009: 3460-3465.
- [3] CABECINHAS D, CUNHA R, SILVESTRE C. A globally stabilizing path following controller for rotorcraft with wind disturbance rejection [J]. IEEE Transactions on Control Systems Technology, 2015, 23(2): 708-714.
- [4] KLAUSEN K, FOSSEN T, JOHANSEN T, et al. Cooperative path-following for multirotor UAVs with a suspended payload [C] // 2015 IEEE Conference on Control Applications (CCA). Piscataway: IEEE Press, 2015: 1354-1360.
- [5] AKKINAPALLI V S, NIERMEYER P, LOHMANN B, et al. Adaptive nonlinear design plant uncertainty cancellation for a multirotor [C] // 2016 International Conference on Unmanned Aircraft Systems (ICUAS). Piscataway: IEEE Press, 2016: 1102-1110.
- [6] RUCCO A, AGUIAR A P, HAUSER J. Trajectory optimization for constrained UAVs: A virtual target vehicle approach [C] // 2015 International Conference on Unmanned Aircraft Systems (ICUAS). Piscataway: IEEE Press, 2015: 236-245.
- [7] RUCCO A, AGUIAR A P, PEREIRA F L, et al. A predictive path-following approach for fixed-wing unmanned aerial vehicles in presence of wind disturbances [C] // Robot 2015: Second Iberian Robotics Conference. Berlin: Springer, 2016: 623-634.

- [ 8 ] SINGH S, PADHI R. Automatic path planning and control design for autonomous landing of UAVs using dynamic inversion [ C ] // Proceedings of the American Control Conference. Piscataway: IEEE Press, 2009: 2409-2414.
- [ 9 ] AKHTAR A, WASLANDER S, NIELSEN C. Fault tolerant path following for a quadrotor [ C ] // Proceedings of the IEEE Conference on Decision and Control. Piscataway: IEEE Press, 2013: 847-852.
- [ 10 ] BYRNES C I, ISIDORI A. Asymptotic stabilization of minimum phase nonlinear systems [ J ]. IEEE Transactions on Automatic Control, 1991, 36(10): 1122-1137.
- [ 11 ] EDWARDS C, SPURGEON S K. Sliding mode control: Theory and applications [ M ]. London: Taylor & Francis, 1988: 43-50.
- [ 12 ] SHAH M Z, SAMAR R, BHATI A. Lateral track control of UAVs using the sliding mode approach: From design to flight testing [ J ]. Transactions of the Institute of Measurement and Control, 2015, 37(4): 457-474.
- [ 13 ] ROZA A, MAGGIORE M. Path following controller for a quadrotor helicopter [ C ] // Proceedings of the American Control Conference. Piscataway: IEEE Press, 2012: 4655-4660.
- [ 14 ] XARGAY E, KAMINER I, PASCOAL A, et al. Time-critical cooperative path following of multiple unmanned aerial vehicles over time-varying networks [ J ]. Journal of Guidance, Control, and Dynamics, 2013, 36(2): 499-516.
- [ 15 ] WANG Y, WANG D, ZHU S. Cooperative moving path following for multiple fixed-wing unmanned aerial vehicles with speed constraints [ J ]. Automatica, 2019, 100: 82-89.
- [ 16 ] CHEN H, CONG Y, WANG X, et al. Coordinated path-following control of fixed-wing unmanned aerial vehicles [ J ]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2022, 52(4): 2540-2554.
- [ 17 ] 崔正阳, 王勇. 无人机固定时间路径跟踪容错制导控制研究 [ J ]. 北京航空航天大学学报, 2021, 47(8): 1619-1627.
- CUI Z Y, WANG Y. Fault tolerant fixed-time path following control of UAV [ J ]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(8): 1619-1627 (in Chinese).
- [ 18 ] 鲜斌, 李宏图. 分布式多无人机的时变编队非线性控制设计 [ J ]. 控制与决策, 2021, 36(10): 2490-2496.
- XIAN B, LI H T. Time-varying formation nonlinear control of distributed multiple UAVs [ J ]. Control and Decision, 2021, 36(10): 2490-2496 (in Chinese).
- [ 19 ] MENON P K, SWERIDUK G D, SRIDHAR B. Optimal strategies for free-flight air traffic conflict resolution [ J ]. Journal of Guidance, Control, and Dynamics, 1999, 22(2): 202-211.
- [ 20 ] WANG J, XIN M. Integrated optimal formation control of multiple unmanned aerial vehicles [ J ]. IEEE Transactions on Control Systems Technology, 2013, 21(5): 1731-1744.
- [ 21 ] MENON P K A. Short-range nonlinear feedback strategies for aircraft pursuit-evasion [ J ]. Journal of Guidance, Control, and Dynamics, 1989, 12(1): 27-32.
- [ 22 ] HARTMAN E J, KEELER J D, KOWALSKI J M. Layered neural networks with Gaussian hidden units as universal approximations [ J ]. Neural Computation, 1990, 2: 210-215.
- [ 23 ] AGUIAR A P, HESPANHA J P. Trajectory-tracking and path-following of underactuated autonomous vehicles with parametric modeling uncertainty [ J ]. IEEE Transactions on Automatic Control, 2007, 52(8): 1362-1379.
- [ 24 ] RUBI B, PEREZ R, MORCEGO B. A survey of path following control strategies for UAVs focused on quadrotors [ J ]. Journal of Intelligent & Robotic Systems, 2020, 98: 241-265.
- [ 25 ] OLFATI-SABER R, MURRAY R M. Consensus problems in networks of agents with switching topology and time-delays [ J ]. IEEE Transactions on Automatic Control, 2004, 49(9): 1520-1533.
- [ 26 ] GHABCHELOO R, PASCOAL A, SILVESTRE C, et al. Nonlinear coordinated path following control of multiple wheeled robots with bidirectional communication constraints [ J ]. International Journal of Adaptive Control and Signal Processing, 2007, 21: 133-157.

### 3D cooperative path following control of multi-UAVs with input saturation

LIU Ronghua<sup>1</sup>, LIU Shuguang<sup>1,\*</sup>, ZHANG Boyang<sup>1</sup>, WANG Huan<sup>1</sup>, LI Wei<sup>2</sup>

(1. Equipment Management and Unmanned Aerial Vehicle Engineering College, Air Force Engineering University, Xi'an 710038, China;

2. Second Military Representative Office of Air Force Equipment Department in Luoyang, Luoyang 471009, China)

**Abstract:** Aiming at the problem of multi-UAV cooperative path following in three-dimensional space, a cooperative controller was designed based on backstepping. To avoid control input saturation, an auxiliary controller was introduced to ensure control performance of the system. The six degrees of freedom nonlinear model of the UAV was feedback linearized, and the uncertain disturbances of the UAV and the unmodeled dynamics of the model itself were estimated and compensated online by radial basis function, thus improving the robustness of the system and its capacity to resist disturbances. Meanwhile, a first-order filter was introduced to avoid the derivation of the virtual control variable in the backstepping method. The graph theory was adopted to solve the communication problem between UAVs, and the consensus theory to realize the cooperative control of multiple UAVs. Finally, the Lyapunov stability theory was used to prove the stability of the system. The simulation results show that the designed cooperative path following controller can achieve good cooperative tracking control effects.

**Keywords:** multi-UAVs; input saturation; cooperative control; path following; backstepping

**Received:** 2020-12-18; **Accepted:** 2021-03-05; **Published online:** 2021-03-16 14:56

**URL:** [kns.cnki.net/kcms/detail/11.2625.V.20210316.1444.003.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210316.1444.003.html)

**Foundation items:** National Natural Science Foundation of China (61973253); National Defense Science and Technology Key Laboratory Foundation of China (6142219200301); Aeronautical Science Foundation of China (20155896025)

\* **Corresponding author.** E-mail: dawny418@126.com

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2020.0710

# 一种改进 PSO-ARMA 半球谐振陀螺 温度误差建模方法

吴宗收, 汪立新\*, 李新三, 李灿

(火箭军工程大学 导弹工程学院, 西安 710025)

**摘 要:** 为减小半球谐振陀螺(HRG)在温度效应下产生的漂移,建立了温度漂移补偿模型,对与温度有关的不确定性漂移进行了补偿。提出了一种改进 PSO-ARMA 建模方法,对不确定性漂移进行了补偿。改进的 PSO-ARMA 建模方法将惯性权值递减策略引入到反向学习粒子群优化(PSO)算法中,提高算法跳出局部、快速收敛的能力,在建模时利用改进的 PSO 算法对 ARMA 参数寻优,以提高模型的精度。利用半球谐振陀螺升温实验数据进行了检验,经该模型补偿后,陀螺输出精度可达  $0.07^\circ/\text{h}$ ,且较传统 ARMA 建模方法精度提高了一倍。

**关键词:** 半球谐振陀螺(HRG); ARMA; 粒子群优化(PSO)算法; 反向学习; 温度漂移补偿模型

**中图分类号:** V221+.3; TB553

**文献标志码:** A **文章编号:** 1001-5965(2022)06-1050-07

半球谐振陀螺(hemispherical resonator gyro, HRG)作为一种新型的惯性敏感器,具有可靠性高、稳定性强、使用寿命长等特点,是目前精度最高的固体波动陀螺<sup>[1-2]</sup>。HRG 核心部件为由高纯熔融石英材料制成的半球结构的谐振子,且封装在真空金属罩中。由于 HRG 谐振子自身振动产生的热量及工作环境温度的变化产生的温度漂移,极大地影响了 HRG 的使用精度<sup>[3]</sup>。

HRG 温度漂移的解决办法主要有温度稳定法和温度补偿法<sup>[4]</sup>。温度稳定法是将 HRG 置于一个温度受控的环境里,保持其工作时周围环境温度恒定,从而减少温度变化引起的漂移,但是由于半球谐振子封装在真空金属罩中,其热交换只能依赖热辐射及谐振子与外罩之间支撑杆的热传递,陀螺仪需要很长时间才能达到温控环境的预设温度,不能满足快速启动的需要,温控系统的存

在也势必增加系统的体积和成本。温度补偿法是根据 HRG 的输出特性,建立其随温度变化的数学模型<sup>[5]</sup>,补偿温度变化导致的误差,可以在不增加体积、质量、成本的前提下补偿 HRG 温度漂移带来的误差,提高其使用精度。

为了提高 HRG 使用精度,科研人员对其温度漂移补偿做了许多工作。周小刚等<sup>[6]</sup>对 HRG 温度漂移的影响因素进行了分析,并建立了温度误差补偿模型;李广胜等<sup>[7]</sup>利用系统辨识理论建立了 HRG 温度漂移的 AR 多变量模型;王旭等<sup>[8]</sup>通过逐步线性回归的方法建立了基于 HRG 谐振子固有频率的温度补偿模型,提高了陀螺在大温度范围内的使用性能;周强等<sup>[9]</sup>设计了 HRG 温度补偿控制电路,一定程度上提高了陀螺对温度误差的抑制能力。

本文针对以上问题,进行了半球谐振陀螺温

收稿日期: 2020-12-23; 录用日期: 2021-05-04; 网络出版时间: 2021-05-17 16:09

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20210517.1521.002.html](http://kns.cnki.net/kcms/detail/11.2625.V.20210517.1521.002.html)

基金项目: 陕西省自然科学基金基础研究计划(2020JQ491); 陕西省高校科协青年人才托举计划(20200109)

\*通信作者. E-mail: 3115000134@mail2.gdut.edu.cn

**引用格式:** 吴宗收, 汪立新, 李新三, 等. 一种改进 PSO-ARMA 半球谐振陀螺温度误差建模方法[J]. 北京航空航天大学学报, 2022, 48(6): 1050-1056. WU Z S, WANG L X, LI X S, et al. An improved PSO-ARMA method for temperature error modeling of hemispherical resonator gyroscope [J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(6): 1050-1056 (in Chinese).

度试验,提出了陀螺温度漂移确定性补偿模型;在对确定性温度漂移进行补偿后,利用改进的时间序列方法对不确定性漂移进行了补偿,以实现陀螺输出漂移补偿的最佳效果。

### 1 半球谐振陀螺温度误差分析

HRG 温度误差主要来自于 2 个方面:①HRG 控制系统的温度误差。由于 HRG 控制电路板封装在真空环境的壳体内部,在控制电路工作过程中所产生的热量很难进行有效的散热,从而电路及 HRG 的温度发生变化,引起电路测量误差及 HRG 输出误差。②陀螺组件的温度误差。在 HRG 正常运行时,外界环境温度的改变及 HRG 谐振子振动所产生的热量都会导致陀螺温度的变化。由于 HRG 谐振子由高纯熔融石英材料制成,当其温度发生变化时,HRG 谐振子的弹性模量、泊松比、材料密度等一系列物理属性均会发生改变,同时温度的变化会导致 HRG 谐振频率的变化及其  $Q$  值的改变<sup>[10]</sup>,从而影响陀螺的输出精度。

研究表明,HRG 的谐振频率与温度密切相关。在半球壳体不发生转动,谐振子处于  $n=2$  的四波腹振动状态时,有

$$f = \omega_n = 1.5127 \frac{h}{R^2} \left[ \frac{E}{(1 + \mu)\rho} \right]^{\frac{1}{2}} \quad (1)$$

式中:  $f, \omega_n$  为谐振子的谐振频率,  $n$  为谐振子环向振型的波节点数;  $R$  为谐振子中半径;  $h$  为半球谐振子厚度;  $E$  为弹性模量;  $\mu$  为材料的泊松比;  $\rho$  为材料的密度。

谐振子结构如图 1 所示。图中:  $d$  为谐振子支撑杆直径。

当有温度因素的影响时,式(1)可表示为

$$f(T) = \omega_2 = 1.5127 \frac{h(T)}{R^2(T)} \left[ \frac{\|E(T)\|}{(1 + \mu(T))\rho(T)} \right]^{\frac{1}{2}} \quad (2)$$

由式(1)可知,半球谐振子的谐振频率与多项

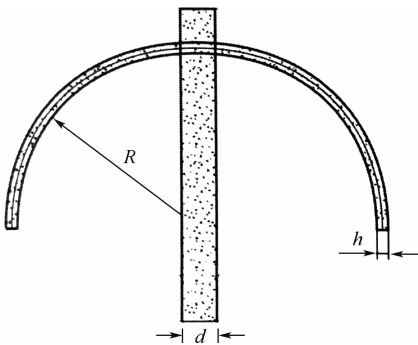


图 1 HRG 谐振子结构示意图

Fig. 1 Simple structure diagram of HRG harmonic oscillator

因素有关,考虑到温度的变化对上述各物理量都有所影响,因此对诸多因素逐项分析将会大大增加分析的难度<sup>[11]</sup>。为了简化分析,在温度  $T_0$  处将  $f(T)$  展开成三阶泰勒级数,可表示为

$$f(T) = f(T_0) + \left. \frac{\partial f}{\partial T} \right|_{T=T_0} (T - T_0) + \left. \frac{\partial^2 f}{2! \partial T^2} \right|_{T=T_0} (T - T_0)^2 + \left. \frac{\partial^3 f}{3! \partial T^3} \right|_{T=T_0} (T - T_0)^3 \quad (3)$$

由上可知,很难用数学解析的方法得到式(3)中的各项系数,在实际应用中通常采用标定实验确定相关系数,并通过温度补偿的方式减少由温度引起的陀螺漂移。

### 2 基于改进 PSO-ARMA 半球谐振陀螺温度误差建模

为了提高 ARMA 模型预测精度,克服矩估计法、最小二乘法等常用 ARMA 建模方法存在的计算过程繁琐、限制条件多的缺点,本文将改进的粒子群优化 (particle swarm optimization, PSO) 算法对 ARMA 模型参数进行优化估计,通过全局搜索得到模型的最佳参数,建立 HRG 温度误差模型。

#### 2.1 PSO 算法

PSO 算法是一种基于群体智能的并行高效优化算法<sup>[11]</sup>,其记忆特性使粒子可以动态跟踪当前搜索情况并调整搜索方向,不需要交叉和变异等操作<sup>[12]</sup>,在参数寻优问题中具有精度高、过程简单、收敛快等优点,适合于模型优化。粒子在解空间中的个体位置由跟踪个体最佳位置 Pbest 和群体最佳位置 Gbest 来更新,从而实现对候选解的进化<sup>[13]</sup>。第  $i$  个粒子的速度为  $V_i = [V_{i1}, V_{i2}, \dots, V_{id}]^T$ , 其个体最佳位置 Pbest 为  $P_i = [P_{i1}, P_{i2}, \dots, P_{id}]^T$ , 群体最佳位置 Gbest 为  $P_g = [P_{g1}, P_{g2}, \dots, P_{gd}]^T$ 。

在下次迭代中,粒子通过 Pbest 和 Gbest 更新自身的速度与位置,更新公式为

$$V_{id}^{k+1} = wV_{id}^k + c_1r_1(P_{id}^k - X_{id}^k) + c_2r_2(P_{gd}^k - X_{id}^k) \quad (4)$$

$$X_{id}^{k+1} = X_{id}^k + V_{id}^{k+1} \quad (5)$$

式中:  $d = 1, 2, \dots, D; i = 1, 2, \dots, n; k$  为当前迭代次数;  $V_{id}$  为在第  $d$  维空间当前粒子的速度;  $c_1, c_2$  为学习因子,为非负常数;  $r_1, r_2$  为区间  $[0, 1]$  内的均匀随机数;  $w$  为惯性权值。

#### 2.2 反向学习策略

在 PSO 算法初始化时,随机生成的粒子距离最优解越近越有利于算法的收敛。由于初始粒子

种群的随机性,无法预知初始粒子与最优解之间的距离,往往导致算法无效搜索。反向学习 (opposition-based learning, OBL) 是 2005 年由 Tizhoosh<sup>[14]</sup> 提出的一种智能算法领域的新技术,其基本思想为:对于粒子的一个可行解,可以生成其反向解,由于反向解相较可行解更靠近全局最优的可能性在概率上约占 50%,如果将可行解与反向解各自构造的解空间进行合并,在构造的合解空间内进行搜索,粒子获得全局最优解的能力将大大提升。

**定义 1** 设每个粒子  $x_i(t) = [x_{i1}, x_{i2}, \dots, x_{id}]$  的反向解为  $\tilde{x}_j = [\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_d]$ ,  $x_j \in [a_j, b_j]$ , 其中

$$\tilde{x}_j = a_j + b_j - x_j \quad (6)$$

**定义 2**  $D$  维空间内的普通粒子  $x_i(t)$  对应的个体极值即为精英粒子<sup>[15]</sup>, 设精英粒子的反向解为

$$\tilde{x}_{ij} = k(a_j + b_j) - x_{ij} \quad (7)$$

式中:  $x_{ij} \in [a_j, b_j]$ ;  $k \in [0, 1]$  为一般化系数,使粒子获得更优的反向解<sup>[16]</sup>。

### 2.3 改进的 PSO-ARMA 方法

为提高 PSO 算法的收敛速度,避免陷入局部最优,将惯性权重递减策略引入到反向学习 PSO 算法中,并将改进的 PSO 算法应用于 ARMA 模型的阶数  $p, q$  定阶寻优。

惯性权值  $w$  表示粒子当前速度受历史速度的影响程度,一种线性惯性权值递减策略的  $w$  计算公式为

$$w = w_s - (w_s - w_e) \left( \frac{i}{\text{maxiter}} \right) \quad (8)$$

式中:  $w_s$  为惯性权值最大值;  $w_e$  为惯性权值最小值;  $i$  为当前迭代次数;  $\text{maxiter}$  为最大迭代次数。在惯性权值  $w$  变化的影响下,粒子的速度与搜索能力也产生与之相应的变化。当惯性权值  $w$  较大时,粒子的飞行速度也较大,粒子的全局搜索能力较好;当惯性权值  $w$  随迭代次数的增加而减小时,粒子的飞行速度也相应变小,便于粒子群快速聚集,此时粒子的局部搜索能力较好。

改进的 PSO-ARMA 方法参数寻优流程如图 2 所示。

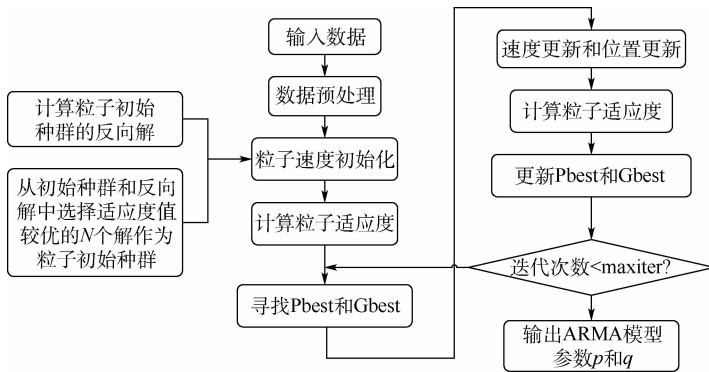


图 2 改进的 PSO-ARMA 流程

Fig. 2 Flowchart of improved PSO-ARMA

改进的 PSO-ARMA 方法主要步骤如下:

**步骤 1** 设置 PSO 算法参数。当前迭代次数  $k$ 、搜索空间维数  $D$ 、学习因子  $c_1$  和  $c_2$ 、粒子飞行速度上下界  $V_{\max}$ 、 $V_{\min}$  等参数,生成粒子初始位置及速度,随机初始化  $N$  个粒子种群。

**步骤 2** 生成反向粒子种群,合并正向与反向 2 个种群,并计算初始适应度值,选择  $N$  个适应度值较优的粒子作为初始种群,将每个粒子当前适应度值作为个体最佳位置 Pbest,所有 Pbest 中最优的设置作为群体最佳位置 Gbest。

**步骤 3** 根据 Pbest、Gbest 及惯性权值递减规则,由式(4)、式(5)更新粒子的速度与位置。

**步骤 4** 更新个体最佳位置 Pbest 及群体最佳位置 Gbest,并且根据更新后的 Pbest 和 Gbest 计算各粒子的适应度值,通过与历史群体最佳位

置的比较,得到新的 Gbest。

**步骤 5** 迭代运算,不断循环步骤 3 与步骤 4,在达到预设迭代次数后停止搜索,得到 ARMA 模型的最优  $p, q$  阶数。

在以上步骤中,需要合理设置算法的各项参数:

1) 学习因子  $c_1, c_2$  分别代表粒子向 Pbest 和 Gbest 靠拢的能力,  $c_1, c_2$  过小或过大将导致粒子容易陷入局部最优或算法无法收敛。本文中设置  $c_1 = c_2 = 2$ 。

2) 种群规模及迭代次数  $\text{maxiter}$  的选择。 $\text{maxiter}$  选择过小将影响寻优结果的精度,过大将导致方法运行时间过长。种群规模由以往经验确定。本文中设置迭代次数  $k = 50$ ,种群规模  $N = 10$ 。

3) 粒子的速度与位置范围。ARMA 模型在

实际应用中阶数  $p, q$  一般较低,因此将粒子的位置下限与上限分别设置为  $[1, 1], [3, 3]$ 。

ARMA 模型中,使用 AIC 准则对 ARMA 进行定阶,选择均方根误差 RMSE 作为改进 PSO-ARMA 模型的适应度函数,具体如下:

$$RMSE = \sqrt{\frac{1}{n} \sum_{k=1}^n (x_k - x_{k-pred})^2} \quad (9)$$

式中:  $x_k$  为第  $k$  个样本的实际值;  $x_{k-pred}$  为第  $k$  个样本的预测值;  $n$  为样本数量。

### 3 实验验证与分析

本节进行 HRG 温度漂移实验。实验时将 HRG 置于温箱中,初始温度常温 ( $25^\circ\text{C}$ ),温箱环境最高温度设置为  $70^\circ\text{C}$ ,测试陀螺在不同温度点的输出。固定 HRG 位置,陀螺通电,程序运行 5 min 后开始采集数据,采样间隔为 5 s。保持温箱温度  $25^\circ\text{C}$  恒定约 1 h,后匀速升温至  $70^\circ\text{C}$ 。在达到  $70^\circ\text{C}$  后保温约 2 h,将陀螺断电并关闭温箱。

在相同实验条件下,对同一 HRG 进行多次实验,数据说明该型号 HRG 具有较好的一致性。图 3 为其中一次温度变化实验的陀螺输出曲线。可知,工作温度发生变化时,固定位置 HRG 输出产生漂移且漂移较大。因此,可对温度变化下 HRG 输出数据建立补偿模型,以减小温度变化对 HRG 输出精度的影响。

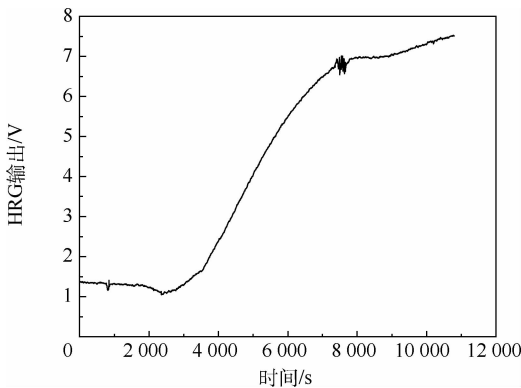


图 3 HRG 输出数据变化曲线  
Fig. 3 Curve of HRG output data

#### 3.1 温度模型及 ARMA 模型预测

HRG 温度漂移补偿包含确定性部分补偿与不确定性部分补偿 2 部分。不确定性温度漂移需建立在确定性漂移补偿基础之上,因此,应先进行确定性漂移补偿,再进行非确定性漂移补偿。由文献 [1] 可知,在 HRG 实际工作时,由于热传导的不均匀及熔融石英材料物理属性的变化,其工作温度及温度变化率也会随之变化,因此,可建立 HRG 温度漂移补偿公式为

$$\omega = \omega_0 + K_1 T + K_2 T' + K_3 (T')^2 + K_4 T T' \quad (10)$$

式中:  $\omega_0$  为陀螺常温稳定工作时刻的输出;  $T$  为陀螺谐振子温度;  $T'$  为温度变化率;  $K_1, K_2, K_3, K_4$  为与温度及温度变化率相关的系数。

多次采集分析该型号 HRG 输出数据,可得到  $\omega_0 = -0.001732 \text{ V}$ ,经最小二乘拟合辨识各系数:  $K_1 = 0.000658 \text{ V}/^\circ\text{C}$ ;  $K_2 = -0.063452 \text{ V} \cdot (\text{s}/^\circ\text{C})$ ;  $K_3 = -1.4647 \text{ V} \cdot (\text{s}/^\circ\text{C})^2$ ;  $K_4 = 0.004368 \text{ V} \cdot \text{s}/(^\circ\text{C})^2$ 。

对 HRG 工作稳定后  $25 \sim 70^\circ\text{C}$  的升温过程,利用式 (10) 对陀螺升温数据补偿,结果如图 4 所示。常温环境中输出精度为  $0.19^\circ/\text{h}$  的该型号 HRG,在环境温度变化达到约  $45^\circ\text{C}$  时,其输出漂移约为  $40^\circ/\text{h}$ ,经补偿后,由温度影响引起的陀螺漂移为  $0.4^\circ/\text{h}$ ,补偿效果较为理想。

对经确定性补偿后得到的漂移数据作平稳性检验,再对得到的平稳随机序列建立 ARMA 模型。由于 ARMA 模型在实际应用中一般阶次较低,确定  $p, q$  的最大阶数为 3。使用枚举方法根据 AIC 准则对 ARMA 模型定阶, AIC ( $p, q$ ) 值如表 1 所示。

对比表 1 中数据,在  $p = 2, q = 2$  时有 AIC 最小值,选择 ARMA (2, 2) 模型,自回归参数为  $\varphi_1 = -1.4125, \varphi_2 = 0.4127$ ,滑动平均参数为  $\theta_1 = -1.0206, \theta_2 = 0.1303$ 。使用该时间序列模型补

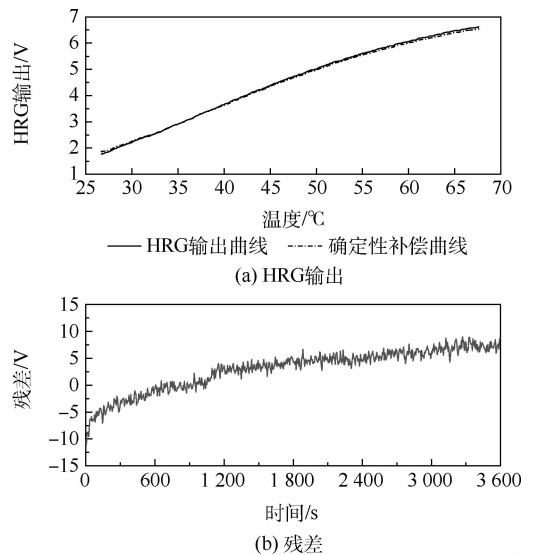


图 4 确定性部分补偿结果  
Fig. 4 Results of deterministic drift compensation

表 1 AIC ( $p, q$ ) 值

Table 1 Values of AIC ( $p, q$ )

$q$	$p$		
	1	2	3
1	-21.527	-21.834	-21.848
2	-21.744	-21.856	-21.833
3	-21.791	-21.853	-21.851

偿后,得到完整的 HRG 温度漂移模型,补偿结果如图 5 所示,补偿后的 HRG 输出精度为  $0.12^\circ/\text{h}$ 。

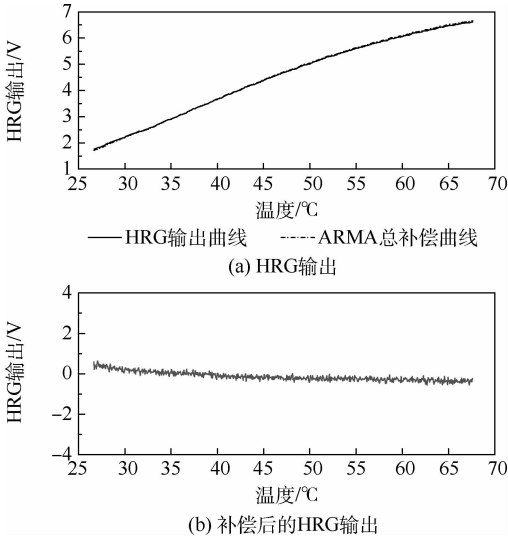


图 5 ARMA 总补偿结果

Fig. 5 Overall results of ARMA compensation

### 3.2 改进的 PSO-ARMA 模型预测

按照 2.3 节中的参数预设与方法步骤,使用引入惯性权重递减策略的反向学习 PSO 算法,对 3.1 节中得到的 HRG 不确定部分温度漂移数据建立 ARMA 模型并进行参数寻优,在迭代完成后,根据适应度最佳规则选出群体最优位置,得到最佳预测模型为  $\text{ARMA}(1,3)$ ,自回归参数为  $\varphi = -0.9991$ ,滑动平均参数为  $\theta_1 = -0.5843$ 、 $\theta_2 = -0.0772$ 、 $\theta_3 = -0.1362$ 。经该方法对温度漂移的不确定性部分补偿后,HRG 输出精度达到  $0.07^\circ/\text{h}$ ,较 3.1 节中补偿结果精度约提高了一倍,补偿结果如图 6 所示。

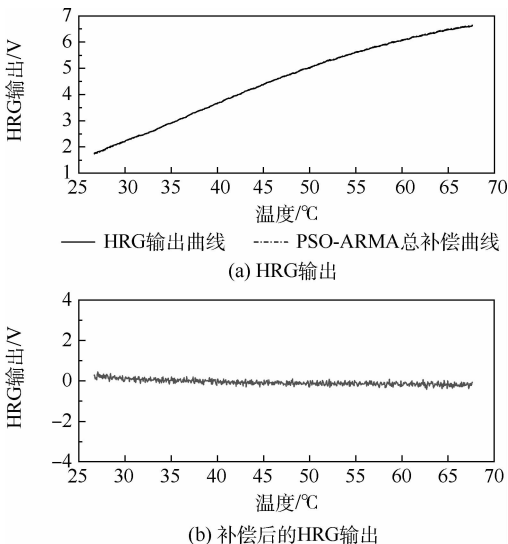


图 6 PSO-ARMA 总补偿结果

Fig. 6 Overall results of PSO-ARMA compensation

### 3.3 方法验证

为检验改进的 PSO-ARMA 方法所建模型的适用性,在相同实验条件下进行该型号 HRG 的升温实验,应用式(10)及 3.2 节所建立  $\text{ARMA}(1,3)$  模型对温度漂移进行补偿,总补偿结果如图 7 所示。经补偿,陀螺输出精度可达到  $0.06^\circ/\text{h}$ ,说明改进的 PSO-ARMA 方法建立的 ARMA 模型能够较好地补偿该陀螺温度漂移中的不确定性部分。

为检验改进 PSO 算法的快速性与逃出局部最优的能力,对 HRG 升温过程,同时使用改进 PSO-ARMA 方法与基于 LDWPSO<sup>[17]</sup> 的参数寻优方法进行比对,如图 8 所示。通过对比可以看出,LDWPSO 算法在 50 次的迭代步数内,其适应度值的变化始终处于跳变的状态,最终未达到收敛的稳定;改进的 PAO-ARMA 方法在第 20 次迭代后其寻优过程即趋于稳定,未出现过早收敛的问题。

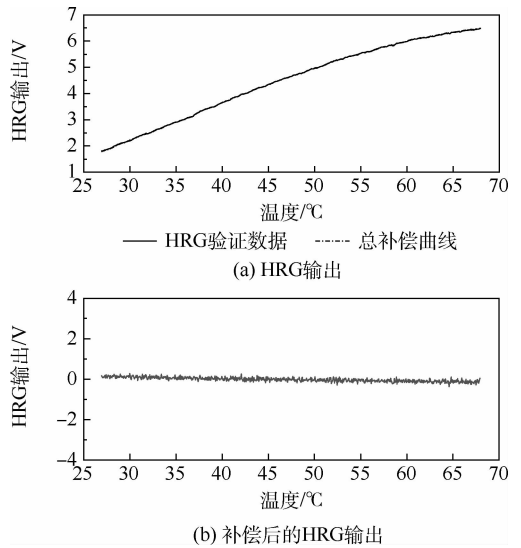


图 7 实验验证曲线

Fig. 7 Experimental verification curves

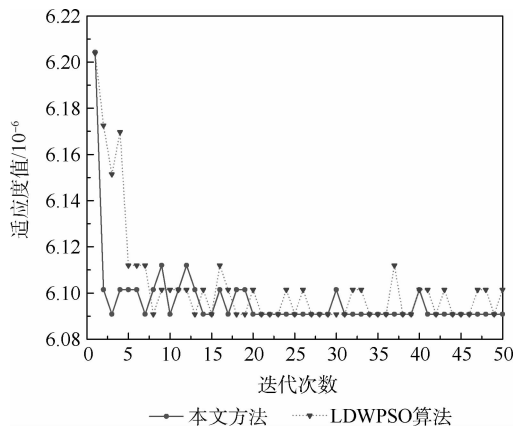


图 8 两种算法寻优过程对比

Fig. 8 Comparison of optimization process between two algorithms

## 4 结 论

1) 对 HRG 温度漂移的确定性部分进行了补偿,经补偿后,陀螺精度可达  $0.4^\circ/\text{h}$ 。

2) 应用改进的 PSO-ARMA 方法建立了 AR-MA 模型,对温度漂移的不确定性部分进行补偿。对温度漂移总补偿后,陀螺输出精度可达  $0.07^\circ/\text{h}$ ,补偿效果较为理想。

3) 进行了改进的 PSO-ARMA 方法与基于 LDWPSO 算法寻优过程的对比,前者在较少迭代次数内寻优过程便趋于稳定,说明该方法能够有效地跳出局部最优,同时有较好的收敛能力。

### 参考文献 (References)

- [1] 刘付成,赵万良,宋丽君.半球谐振陀螺惯性传感器及其空间应用[M].北京:中国宇航出版社,2019.  
LIU F C,ZHAO W L,SONG L J. Inertial sensor of hemispherical resonator gyroscope and its space application[M]. Beijing: China Aerospace Press,2019(in Chinese).
- [2] 樊尚春,刘广玉.半球谐振陀螺耦合振动的近似解析分析[J].北京航空航天大学学报,1997,23(6):708-713.  
FAN S C,LIU G Y. Approximately analysing of coupling vibration of resonator for HRG[J]. Journal of Beijing University of Aeronautics and Astronautics,1997,23(6):708-713(in Chinese).
- [3] 刘吉利,李建朋,武志忠,等.半球谐振陀螺零偏温度漂移的自补偿[J].空间控制技术与应用,2018,44(1):69-73.  
LIU J L,LI J P,WU Z Z,et al. Self-compensation of bias temperature drift in hemispherical resonator gyro[J]. Aerospace Control and Application,2018,44(1):69-73(in Chinese).
- [4] YATSENKO Y A,CHIKOVANI V V,YANCHUK A N. Study of thermal drift model for metallic cylindrical coriolis vibratory gyro[C]//10th Saint Petersburg International Conference on Integrated Navigation Systems,2003:337-339.
- [5] 刘宇,路永乐,黎蕾蕾.固态振动陀螺与导航技术[M].北京:中国宇航出版社,2010.  
LIU Y,LU Y L,LI L L. Solid-state vibration gyro and navigation technology[M]. Beijing:China Aerospace Press,2010(in Chinese).
- [6] 周小刚,汪立新,余婧,等.半球谐振陀螺温度补偿与实验研究[J].宇航学报,2010,31(4):1083-1087.  
ZHOU X G,WANG L X,SHE Q,et al. HRG temperature compensation and experiment research[J]. Journal of Astronautics,2010,31(4):1083-1087(in Chinese).
- [7] 李广胜,蒋英杰,孙志强,等.基于 AR 多变量模型的半球谐振陀螺温度漂移建模[J].传感技术学报,2009,22(10):1442-1445.  
LI G S,JIANG Y J,SUN Z Q,et al. Modelling temperature drift of HRG based on AR multi-variable model[J]. Chinese Journal of Sensors and Actuators,2009,22(10):1442-1445(in Chinese).
- [8] WANG X,WU W Q,FANG Z,et al. Temperature drift compensation for hemispherical resonator gyro based on natural frequency[J]. Sensors,2012,12(5):6434-6446.
- [9] 周强,覃施甦,方海斌,等.半球谐振陀螺温度特性及补偿分析[J].压电与声光,2015,37(5):818-820.  
ZHOU Q,QIN S S,FANG H B,et al. The temperature characteristic and compensating analysis of HRG[J]. Piezoelectrics & Acoustooptics,2015,37(5):818-820(in Chinese).
- [10] 任顺清,赵洪波.谐振子密度偏差引起的频率裂解的分析[J].哈尔滨工业大学学报,2012,44(3):13-16.  
REN S Q,ZHAO H B. Analysis of frequency cracking of resonator under the density error[J]. Journal of Harbin Institute of Technology,2012,44(3):13-16(in Chinese).
- [11] ANAND B,AAKASH I,AKSHAY,et al. Improvisation of particle swarm optimization algorithm[C]//2014 International Conference on Signal Processing and Integrated Networks (SPIN). Piscataway:IEEE Press,2014:20-24.
- [12] 肖素琼,罗可.具备反向学习和局部学习能力的磷虾群算法[J].计算机工程与应用,2018,54(18):34-39.  
XIAO S Q,LUO K. Krill herd optimization algorithm with reverse-learning and local-learning behavior[J]. Computer Engineering and Applications,2018,54(18):34-39(in Chinese).
- [13] 宋刚,张云峰,包芳勋,等.基于粒子群优化 LSTM 的股票预测模型[J].北京航空航天大学学报,2019,45(12):2533-2542.  
SONG G,ZHANG Y F,BAO F X,et al. Stock prediction model based on particle swarm optimization LSTM[J]. Journal of Beijing University of Aeronautics and Astronautics,2019,45(12):2533-2542(in Chinese).
- [14] TIZHOOSH H R. Opposition-based learning: A new scheme for machine intelligence[C]//International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-IAWTIC'06). Piscataway:IEEE Press,2005:695-701.
- [15] 韩江,闵杰.基于精英反向学习的烟花爆炸式免疫遗传算法[J].合肥工业大学学报(自然科学版),2020,43(4):433-437.  
HAN J,MIN J. A fireworks explosive immune genetic algorithm based on elite opposition-based learning[J]. Journal of Hefei University of Technology (Natural Science),2020,43(4):433-437(in Chinese).
- [16] 张超.一种精英反向学习的花授粉算法[J].西安工程大学学报,2017,31(6):847-856.  
ZHANG C. A flower pollination algorithm based on elite opposition-based learning[J]. Journal of Xi'an Polytechnic University,2017,31(6):847-856(in Chinese).
- [17] SHI Y,EBERHART R. A modified particle swarm optimizer[C]//1998 IEEE International Conference on Evolutionary Computation Proceedings. IEEE World Congress on Computational Intelligence. Piscataway:IEEE Press,1998:69-73.

## An improved PSO-ARMA method for temperature error modeling of hemispherical resonator gyroscope

WU Zongshou, WANG Lixin<sup>\*</sup>, LI Xinsan, LI Can

(School of Missile Engineering, Rocket Force Engineering University, Xi'an 710025, China)

**Abstract:** To reduce the drift of hemispherical resonator gyroscope (HRG) caused by temperature effect, a temperature drift compensation model is established, which compensates the deterministic drift related to temperature. An improved PSO-ARMA modeling method is proposed to compensate the uncertain drift. This method introduces the decreasing inertia weight strategy into the opposition-based learning particle swarm optimization (PSO) algorithm, improving the algorithm's ability to jump out of the local and converge fast. In the modeling process, the improved PSO algorithm is used to optimize the ARMA parameters, thus improving the model accuracy. The experimental data of HRG temperature rise are used for verification. After the model compensation, the output accuracy of HRG can reach  $0.07^\circ/\text{h}$ , which is twice as high as that of traditional ARMA modeling method.

**Keywords:** hemispherical resonator gyroscope (HRG); ARMA; particle swarm optimization (PSO) algorithm; opposition-based learning; temperature drift compensation model

**Received:** 2020-12-23; **Accepted:** 2021-05-04; **Published online:** 2021-05-17 16:09

**URL:** [kns.cnki.net/kcms/detail/11.2625.V.20210517.1521.002.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210517.1521.002.html)

**Foundation items:** Natural Science Basic Research Program of Shaanxi (2020JQ-491); Shaanxi Provincial Association for Science and Technology of Colleges and Universities Young Talents Lifting Program (20200109)

**\* Corresponding author.** E-mail: 3115000134@mail2.gdut.edu.cn

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2020.0713

# 基于 AK-IS 法的航空齿轮泵滑动 轴承可靠性分析

高宁, 李华聪\*, 洪林雄, 曹瑞, 符江锋

(西北工业大学 动力与能源学院, 西安 710072)

**摘 要:** 为研究高转速、低介质黏度运行的航空齿轮泵滑动轴承动压润滑可靠性, 通过将 Reynolds 润滑方程与影响矩阵耦合, 建立了考虑轴瓦弹性变形的航空齿轮泵滑动轴承弹流润滑(EHL)模型。考虑滑动轴承尺寸公差和运行工况等引起的不确定性, 将与滑动轴承润滑特性密切相关的动压润滑的压力峰值作为可靠性判据, 采用自适应 Kriging 和重要抽样法相结合的 AK-IS 法对滑动轴承动压润滑特性进行了可靠性及灵敏度计算。研究结果表明: 考虑轴瓦弹性变形的压力峰值比刚性轴瓦压力峰值降低 15.04%, 表明滑动轴承弹流润滑对轴承动压润滑的影响不能忽略; 基于 AK-IS 法的航空齿轮泵滑动轴承动压润滑可靠性分析具有准确性和高效性; 各不确定因素对动压润滑可靠性的影响程度不同, 其中轴承的半径间隙对可靠性最敏感, 转速对可靠性最不敏感。

**关键词:** 航空齿轮泵; 滑动轴承; 弹流润滑(EHL); AK-IS 法; 可靠性分析

**中图分类号:** TH113.2; TB115

**文献标志码:** A **文章编号:** 1001-5965(2022)06-1057-08

滑动轴承作为航空齿轮泵的主要支撑部件, 其良好的润滑特性关系到航空齿轮泵性能的进一步强化。随着现代航空燃油齿轮泵设计目标的提高, 难免存在会使航空齿轮泵滑动轴承无法安全工作进而影响到整个航空齿轮泵正常工作的加工误差、工况差异等随机因素, 因此, 航空齿轮泵滑动轴承的可靠性分析显得尤为重要。

国内外学者基于流体动压润滑理论采用数值仿真技术针对滑动轴承展开了大量仿真研究<sup>[1]</sup>。近年来, 为了更加符合轴承实际工作环境, 国内外学者考虑了更多影响因素开展了滑动轴承的多场耦合仿真研究<sup>[2]</sup>, 建立了热固耦合润滑模型<sup>[3]</sup>、气水湍流模型<sup>[4]</sup>、表面微凸体模型<sup>[5-6]</sup>等, 并进行相应因素作用下的润滑分析。然而这些研究只针对轴承性能与结构设计等方面进行定性分析, 即

在滑动轴承结构性能计算过程中将参数作为确定量, 没有考虑参数不确定性的影响, 且这些模型多适用于已有产品的测试与检验, 难以应用在研发人员设计或者再设计时对产品进行分析。

目前, 针对轴承可靠性分析的研究较少。金燕和刘少军<sup>[7]</sup>考虑油膜热效应, 以热接触疲劳极限衡量可靠性, 建立了基于人工神经网络的航空轴承可靠性分析模型。冯吉路等<sup>[8]</sup>针对数控机床主轴系统以轴端轴心的振动幅值为基础, 建立了基于 AK-MCS (adaptive Kriging-Monte Carlo sampling) 法主轴系统振动可靠性分析模型。胡明用等<sup>[9]</sup>提出了基于 MCS 法的轴承转子系统随机振动可靠性分析模型。以上轴承可靠性分析模型中并未将与润滑特性密切相关的压力作为可靠性判据, 且计入轴瓦表面弹性变形的滑动轴承弹

收稿日期: 2020-12-26; 录用日期: 2021-05-07; 网络出版时间: 2021-05-12 16:38

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20210512.1620.003.html](http://kns.cnki.net/kcms/detail/11.2625.V.20210512.1620.003.html)

基金项目: 国家科技重大专项 (2017-V-0013-0065)

\* 通信作者. E-mail: [lihuacong@nwpu.edu.cn](mailto:lihuacong@nwpu.edu.cn)

**引用格式:** 高宁, 李华聪, 洪林雄, 等. 基于 AK-IS 法的航空齿轮泵滑动轴承可靠性分析[J]. 北京航空航天大学学报, 2022, 48(6): 1057-1064. GAO N, LI H C, HONG L X, et al. Reliability analysis of journal bearings inside aero-gear pump based on AK-IS method[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(6): 1057-1064 (in Chinese).

流动压润滑可靠性分析也很少涉及。

在计入轴瓦弹性变形的航空齿轮泵滑动轴承动压润滑可靠性计算中,由于采用有限元法计算弹性变形,计算量较大,而且弹流润滑 (elasto-hydrodynamic lubrication, EHL) 数学模型为强非线性系统,以此构建的极限状态方程难以用显示表达,采用 MCS 等方法计算效率较低,耗时长<sup>[10]</sup>,而基于方差缩减技术的重要抽样 (importance sampling, IS) 法等仍需要调用大量有限元计算,难以满足实际工程需求<sup>[11]</sup>。为了有效解决此问题,支持向量机<sup>[12]</sup>、响应面法<sup>[13]</sup>、人工神经网络<sup>[7]</sup>、Kriging 模型<sup>[14]</sup>等代理模型被广泛应用。其中,Kriging 模型由于对非线性程度较高的问题拟合效果较好而被广泛应用。为了用较少的样本点构建精确的代理模型,提高可靠性计算精度,研究人员提出了如 EFF (expected feasibility function)、H,U,IEGO (improved efficient global optimization) 等主动学习函数,并用以选择符合要求的样本点逐步更新 Kriging 模型<sup>[15]</sup>。这种情况下,AK-MCS 法由于适用范围广,计算效率高被广泛采用<sup>[8]</sup>,但对于小失效概率问题,更新 Kriging 需要较大的样本池,导致自适应学习函数选择样本点的计算时间增加,更新代理模型较为费时<sup>[16]</sup>。

鉴于上述问题,本文考虑轴瓦在动压力下的弹性变形建立了低黏度、高转速工作条件下的航空齿轮泵滑动轴承弹流润滑模型。利用有限差分法和有限元法相结合对航空齿轮泵弹流润滑数学模型进行耦合求解。采用重要抽样法与自适应 Kriging 相结合的 AK-IS<sup>[17]</sup>法进行了航空齿轮泵滑动轴承的可靠性和可靠性灵敏度分析。本文方法可作为航空齿轮泵滑动轴承设计人员在设计阶段的轴承弹流压润滑可靠性计算方法,同时为后续基于可靠性优化设计方法做铺垫。

## 1 航空齿轮泵滑动轴承弹流润滑模型

### 1.1 Reynolds 润滑方程

图 1 为航空齿轮泵滑动轴承三维模型。在齿轮泵工作期间内,齿轮轴-轴承之间形成了可对齿轮轴产生动压支撑作用的楔形间隙油膜。描述该滑动轴承润滑行为的 Reynolds 润滑方程为

$$\frac{\partial}{R^2 \partial \theta} \left( \frac{h^3}{\eta} \cdot \frac{\partial p}{\partial \theta} \right) + h^3 \frac{\partial}{\partial y} \left( \frac{h^3}{\eta} \cdot \frac{\partial p}{\partial y} \right) = 6U \frac{dh}{Rd\theta} \quad (1)$$

式中: $h$  为油膜厚度; $\eta$  为介质黏度; $\theta$  为柱坐标系下的角度; $y$  为轴向坐标; $p$  为油膜压力; $R$  为齿轮轴半径; $U$  为齿轮轴的旋转速度。

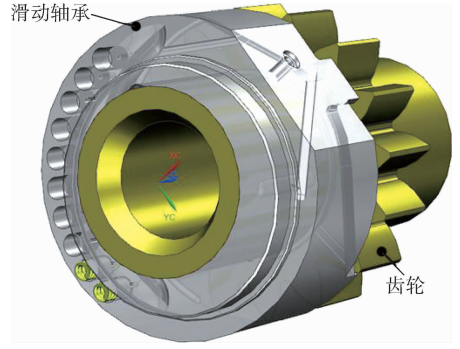


图 1 航空齿轮泵滑动轴承三维模型

Fig. 1 3D model of journal bearing of aero-gear pump

采用 Reynolds 边界条件,其形式为

$$p = 0 \text{ 且 } \frac{\partial p}{\partial \theta} = 0 \quad (2)$$

$$p|_{y=\pm B/2} = p_0 \quad (3)$$

式中: $B$  为轴承宽度; $p_0$  为环境压力。

### 1.2 油膜厚度方程

在齿轮泵工作过程中,齿轮轴-轴承间液膜压力会产生影响轴承油膜厚度的弹性变形,此时油膜厚度方程为

$$h = h_0 + \delta \quad (4)$$

式中: $h_0$  为不计入轴承弹性变形的油膜厚度, $h_0 = C(1 + \varepsilon \cos \theta)$ , $C$  为半径间隙, $\varepsilon$  为偏心率; $\delta$  为油膜压力作用下各节点径向位移对油膜厚度产生的影响。

#### 1.2.1 轴承弹性变形计算

通常采用有限元法求解轴承弹性变形,但该方法在轴承计算中每一步都需通过有限元法求解变形,计算耗时较长,因此本文采用影响矩阵法实现流固双向耦合<sup>[18]</sup>。单位力载荷作用下研究对象的弹性变形用矩阵的形式表示称为影响矩阵。由于研究对象表面弹性变形很小,是一个线性系统,可以应用影响矩阵线性叠加计算,其弹性变形计算方程为

$$[\delta_{rs}] = [K_{ij}^{rs}] \times [p_{ij}] \quad (5)$$

式中: $[\delta_{rs}]$  为实际油膜压力作用下轴承表面上各节点的弹性变形量矩阵; $[K_{ij}^{rs}]$  为影响矩阵,其通过轴承的有限元分析计算并整理后获得,表示作用在节点 $(i,j)$  的单位力使节点 $(r,s)$  产生的径向位移; $[p_{ij}]$  为作用在轴承表面上各节点的实际油膜压力矩阵。

#### 1.2.2 影响矩阵的建立

本文采用 ANSYS 软件的二次开发语言 APDL 进行参数化建模、网格划分及计算求解。将轴承表面节点按一定顺序排列,依次计算出对每一个节点施加单位力时内表面所有节点的径向变形,

并将其作为影响矩阵的列,以全部节点施加单位力时某一节点的径向变形作为影响矩阵的行,以此构成影响矩阵。图 2 为轴承内表面某一节点受单位力(1 N)时的变形。

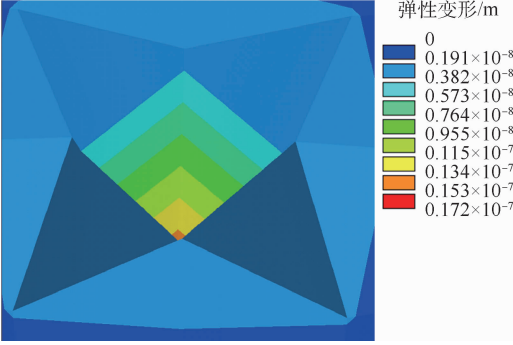


图 2 单位力作用下节点的弹性变形

Fig. 2 Elastic displacements of bearing loaded by 1 N

## 2 AK-IS 可靠性分析法

### 2.1 自适应 Kriging 模型

#### 2.1.1 Kriging 模型

Kriging 模型是一种估计方差最小的无偏估计模型,其有效性与随机误差无关,可以很好地拟合非线性成度较高的问题。Kriging 模型可以表示为多项式和随机过程的组合,即

$$g_k(\mathbf{x}) = \sum_{i=1}^q f_i(\mathbf{x})\beta_i + z(\mathbf{x}) \quad (6)$$

式中: $g_k(\mathbf{x})$ 为预测 Kriging 模型; $f(\mathbf{x})$ 为 Kriging 基函数; $\beta = [\beta_1, \beta_2, \dots, \beta_q]$ 为回归函数待定系数, $q$ 为基函数个数; $z(\mathbf{x})$ 为服从期望为 0、方差为  $\sigma^2$  的高斯随机过程,其协方差为

$$\text{Cov}[z(\mathbf{x}^{(i)}), z(\mathbf{x}^{(j)})] = \sigma^2[\mathbf{R}(\mathbf{x}^{(i)}, \mathbf{x}^{(j)})] \quad (7)$$

其中: $\mathbf{R}(\mathbf{x}^{(i)}, \mathbf{x}^{(j)})$ 为样本点  $\mathbf{x}^{(i)}$ 、 $\mathbf{x}^{(j)}$  之间的相关函数,其用高斯型相关函数可表示为

$$\mathbf{R}(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}) = \exp\left(-\sum_{k=1}^M \theta_k |\mathbf{x}_k^{(i)} - \mathbf{x}_k^{(j)}|^2\right) \quad (8)$$

式中: $M$ 为输入变量维度; $\theta_k (k=1, 2, \dots, M)$ 为待求相关参数,可通过求极大似然估计的最大值得到

$$\max F(\boldsymbol{\theta}) = -\frac{M \ln(\hat{\sigma}^2) + \ln \|\mathbf{R}\|}{2} \quad (9)$$

式中: $\hat{\sigma}$ 为方差的估计量。

已知  $m$  个训练样本数据响应值构成列向量  $\mathbf{g}$ , 则  $\hat{\beta}$  和  $\hat{\sigma}^2$  的估计值为

$$\hat{\beta} = (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F})^{-1} \mathbf{F}^T \mathbf{R}^{-1} \mathbf{g} \quad (10)$$

$$\hat{\sigma}^2 = (\mathbf{g} - \mathbf{F}\hat{\beta})^T \mathbf{R}^{-1} (\mathbf{g} - \mathbf{F}\hat{\beta}) / m \quad (11)$$

式中: $\mathbf{F}$ 为训练样本处的回归模型矩阵; $m$ 为序列

样本个数。

上述求解得到的 Kriging 模型,可用于预测未知点  $\mathbf{x}$  处的响应值,计算公式为

$$\mu_{g_k}(\mathbf{x}) = \mathbf{f}^T(\mathbf{x})\hat{\beta} + \mathbf{r}^T(\mathbf{x})\mathbf{R}^{-1}(\mathbf{g} - \mathbf{F}\hat{\beta}) \quad (12)$$

预测方差为

$$\hat{\sigma}_{g_k}^2(\mathbf{x}) = \sigma^2 \{1 - \mathbf{r}^T(\mathbf{x})\mathbf{R}^{-1}\mathbf{r}(\mathbf{x}) + [\mathbf{F}^T \mathbf{R}^{-1}\mathbf{r}(\mathbf{x}) - \mathbf{f}(\mathbf{x})]^T (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F})^{-1} [\mathbf{F}^T \mathbf{R}^{-1}\mathbf{r}(\mathbf{x}) - \mathbf{f}(\mathbf{x})]\} \quad (13)$$

式中: $\mathbf{r}^T(\mathbf{x}) = \{\mathbf{R}(\mathbf{x}, \mathbf{x}^{(1)}), \mathbf{R}(\mathbf{x}, \mathbf{x}^{(2)}), \dots, \mathbf{R}(\mathbf{x}, \mathbf{x}^{(m)})\}$ 。

#### 2.1.2 主动学习函数

本文采用 IEGO 方法<sup>[14]</sup>,使学习函数的最优选取策略优先考虑极限状态面附近的点和方差较大的点,加快了构建 Kriging 的速度,提高了可靠性的计算精度。IEGO 学习函数定义式为

$$E(I^*(\mathbf{x})) = E(\max(\hat{G}_{\min}^* - G^*(\mathbf{x}), 0)) = \int_0^{+\infty} I(\mathbf{x}) \left\{ \frac{1}{\sqrt{2\pi}\hat{\sigma}_c(\mathbf{x})} \cdot \exp\left[-\frac{(\hat{G}_{\min}^* - I(\mathbf{x}) - \hat{G}^*(\mathbf{x}))^2}{2\hat{\sigma}_c^2(\mathbf{x})}\right] \right\} dI(\mathbf{x}) = (\hat{G}_{\min}^* - \hat{G}^*(\mathbf{x}))\Phi\left(\frac{\hat{G}_{\min}^* - \hat{G}^*(\mathbf{x})}{\hat{\sigma}_c(\mathbf{x})}\right) + \hat{\sigma}_c(\mathbf{x})\varphi\left(\frac{\hat{G}_{\min}^* - \hat{G}^*(\mathbf{x})}{\hat{\sigma}_c(\mathbf{x})}\right) \quad (14)$$

式中: $\hat{G}_{\min}^*$ 为响应值的最小值,其值为  $\hat{G}_{\min}^* = \min(|\hat{G}_1^*|, |\hat{G}_2^*|, \dots, |\hat{G}_m^*|)$ ;  $I^*(\mathbf{x})$ 为对函数的改善程度,  $I^*(\mathbf{x}) = \max((\hat{G}_{\min}^* - G^*(\mathbf{x}), 0)$ ;  $\hat{G}^*(\mathbf{x})$ 和  $\hat{\sigma}_c(\mathbf{x})$ 分别为 Kriging 代理模型预测的均值和标准差; $\Phi$ 为标准正态分布的累积分布函数; $\varphi$ 为标准正态分布的概率密度函数。

### 2.2 重要抽样可靠性及可靠性灵敏度分析

针对高维小失效概率问题,采用 MCS 法抽样效率较低,计算量巨大。重要抽样法通过引入合理的重要抽样密度函数,增加了样本点落入失效域的概率,以此来获得较快的收敛速度和较高的抽样效率。重要抽样法的失效概率表达式为

$$P_f = \int_{R^n} I_F(\mathbf{x})f_X(\mathbf{x})d\mathbf{x} = \int_{R^n} I_F(\mathbf{x})\frac{f_X(\mathbf{x})}{h_X(\mathbf{x})}h_X(\mathbf{x})d\mathbf{x} = E\left(I_F(\mathbf{x})\frac{f_X(\mathbf{x})}{h_X(\mathbf{x})}\right) \quad (15)$$

式中: $I_F(\mathbf{x})$ 为失效域指示函数; $f_X(\mathbf{x})$ 为输入变量  $\mathbf{X} = [X_1, X_2, \dots, X_n]^T$  的联合概率密度函数; $h_X(\mathbf{x})$ 为重要抽样密度函数。

利用重要抽样密度函数  $h_X(\mathbf{x})$ 抽取  $N$  个样本,

记为 $[\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]^T$ , 则失效概率的估计值 $\hat{P}_f$ 为

$$\hat{P}_f = \frac{1}{N} \sum_{j=1}^N \left[ I_F(\mathbf{x}_j) \frac{f_X(\mathbf{x}_j)}{h_X(\mathbf{x}_j)} \right] \quad (16)$$

失效概率 $\hat{P}_f$ 的期望 $E(\hat{P}_f)$ 和方差 $\text{Var}(\hat{P}_f)$ 为

$$E(\hat{P}_f) = E\left(\frac{1}{N} \sum_{j=1}^N \left( I_F(\mathbf{x}_j) \frac{f_X(\mathbf{x}_j)}{h_X(\mathbf{x}_j)} \right)\right) = \hat{P}_f \quad (17)$$

$$\text{Var}(\hat{P}_f) = \frac{1}{N-1} \left[ \frac{1}{N} \sum_{j=1}^N \left( I_F(\mathbf{x}_j) \frac{f_X^2(\mathbf{x}_j)}{h_X^2(\mathbf{x}_j)} \right) - \hat{P}_f^2 \right] \quad (18)$$

失效概率估计值的变异系数 $\text{Cov}[\hat{P}_f]$ 为

$$\text{Cov}[\hat{P}_f] = \sqrt{\text{Var}(\hat{P}_f) / E(\hat{P}_f)} \quad (19)$$

重要抽样可靠性局部灵敏度为

$$\frac{\partial P_f}{\partial \theta_{X_i}^{(k)}} = \int_{R^n} I_F(\mathbf{x}) \frac{\partial f_X(\mathbf{x})}{\partial \theta_{X_i}^{(k)}} \cdot \frac{1}{h_X(\mathbf{x})} h_X(\mathbf{x}) d\mathbf{x} = E\left( \frac{I_F(\mathbf{x})}{h_X(\mathbf{x})} \cdot \frac{\partial f_X(\mathbf{x})}{\partial \theta_{X_i}^{(k)}} \right) \quad (20)$$

式中: $\theta_{X_i}^{(k)}$ 为输入变量 $X_i$ 的第 $k$ 个参数。

利用重要抽样密度函数 $h_X(\mathbf{x})$ 抽取 $N$ 个样本, 记为 $[\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]^T$ , 则可靠性局部灵敏度估计值 $\partial \hat{P}_f / \partial \theta_{X_i}^{(k)}$ 为

$$\frac{\partial \hat{P}_f}{\partial \theta_{X_i}^{(k)}} = \frac{1}{N} \sum_{j=1}^N \left( \frac{I_F(\mathbf{x}_j)}{h_X(\mathbf{x}_j)} \cdot \frac{\partial f_X(\mathbf{x}_j)}{\partial \theta_{X_i}^{(k)}} \right) \quad (21)$$

可靠性局部灵敏度估计值 $\partial \hat{P}_f / \partial \theta_{X_i}^{(k)}$ 的期望

$E(\partial \hat{P}_f / \partial \theta_{X_i}^{(k)})$ 和方差 $\text{Var}(\partial \hat{P}_f / \partial \theta_{X_i}^{(k)})$ 为

$$E\left(\frac{\partial \hat{P}_f}{\partial \theta_{X_i}^{(k)}}\right) = E\left(\frac{1}{N} \sum_{j=1}^N \left( \frac{I_F(\mathbf{x}_j)}{h_X(\mathbf{x}_j)} \cdot \frac{\partial f_X(\mathbf{x}_j)}{\partial \theta_{X_i}^{(k)}} \right)\right) = \frac{\partial \hat{P}_f}{\partial \theta_{X_i}^{(k)}} \quad (22)$$

$$\text{Var}\left(\frac{\partial \hat{P}_f}{\partial \theta_{X_i}^{(k)}}\right) = \text{Var}\left(\frac{1}{N} \sum_{j=1}^N \left( \frac{I_F(\mathbf{x}_j)}{h_X(\mathbf{x}_j)} \cdot \frac{\partial f_X(\mathbf{x}_j)}{\partial \theta_{X_i}^{(k)}} \right)\right) = \frac{1}{N} \text{Var}\left(\frac{I_F(\mathbf{x})}{h_X(\mathbf{x})} \cdot \frac{\partial f_X(\mathbf{x})}{\partial \theta_{X_i}^{(k)}}\right) \approx \frac{1}{N-1} \cdot \left( \frac{1}{N} \sum_{j=1}^N \left( \frac{I_F(\mathbf{x}_j)}{h_X(\mathbf{x}_j)} \cdot \frac{\partial f_X(\mathbf{x}_j)}{\partial \theta_{X_i}^{(k)}} \right)^2 - \left( \frac{\partial \hat{P}_f}{\partial \theta_{X_i}^{(k)}} \right)^2 \right) \quad (23)$$

可靠性局部灵敏度的变异系数为

$$\text{Cov}(\partial \hat{P}_f / \partial \theta_{X_i}^{(k)}) = \sqrt{\text{Var}(\partial \hat{P}_f / \partial \theta_{X_i}^{(k)})} / E(\partial \hat{P}_f / \partial \theta_{X_i}^{(k)}) \quad (24)$$

### 2.3 AK-IS 可靠性分析模型

AK-IS 法是自适应 Kriging 模型与重要抽样结合的方法, 其基本思路为: 利用改进一次二阶矩法求得设计点并构造重要抽样密度函数, 在由重

要抽样密度函数抽取的样本池中利用学习函数自适应地建立 Kriging 代理模型, 基于收敛的 Kriging 代理模型代替真实功能函数进行可靠性分析。具体的计算流程如图 3 所示。

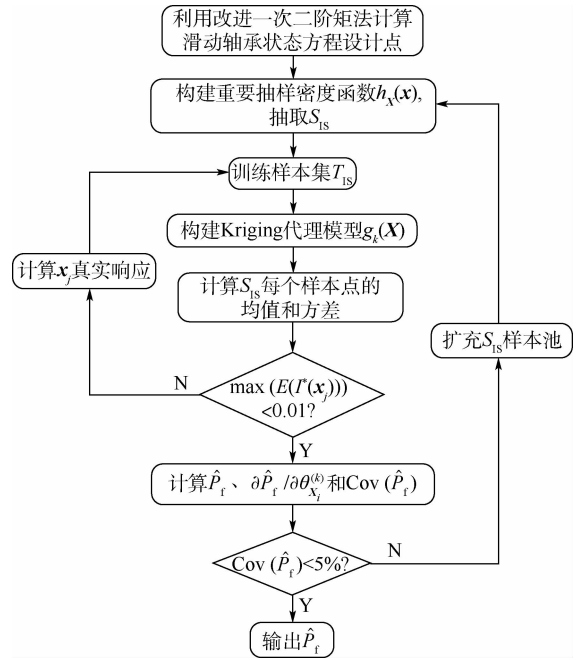


图 3 AK-IS 法计算可靠性流程

Fig. 3 Reliability calculation process of AK-IS method

AK-IS 法具体执行步骤如下:

**步骤 1** 采用改进一次二阶矩法求得极限状态方程的设计点。

**步骤 2** 构建重要抽样密度函数并生成样本池 $S_{IS}$ 。将输入变量的原始密度函数 $f_X(\mathbf{x})$ 的抽样中心 $P$ 平移到 $P^*$ 重新构建重要抽样密度函数 $h_X(\mathbf{x})$ , 并利用 $h_X(\mathbf{x})$ 抽取 $N_{IS}$ 个样本生成样本池。过程示意如图 4 所示。

**步骤 3** 生成初始训练样本集 $T_{IS}$ 。在输入变量空间中采用拉丁超立方产生的样本点 $n_L$ 和求解设计点时迭代过程中产生的样本点 $n_p$ 。及其

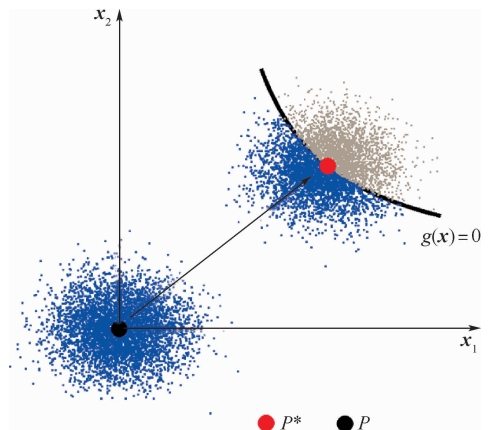


图 4 构建 $h_X(\mathbf{x})$ 的示意图

Fig. 4 Schematic diagram of construction  $h_X(\mathbf{x})$

对应真实响应共同构成训练样本集  $T_{IS}$ 。

**步骤 4** 利用  $T_{IS}$  构建初始 Kriging 模型  $g_k(\mathbf{X})$ 。基于 DACE 工具箱建立 Kriging 模型。

**步骤 5** 在  $S_{IS}$  中选择最佳学习样本点。基于已构建的  $g_k(\mathbf{X})$ , 在样本池  $S_{IS}$  选择最佳样本点, 定义为

$$\mathbf{x}^{IEGO} = \min_{\mathbf{x} \in S_{IS}} E(I^*(\mathbf{x})) \quad (25)$$

**步骤 6** 判断学习过程是否收敛。当  $\max(E(I^*(\mathbf{x}_j))) < 0.01$  时, 自适应学习过程停止, 到步骤 7 计算失效概率。否则将最佳样本点加入训练样本集  $T_{IS}$  构建新的  $g_k(\mathbf{X})$ , 返回步骤 3。

**步骤 7** 利用当前 Kriging 模型  $g_k(\mathbf{X})$  计算失效概率。利用  $g_k(\mathbf{X})$  计算样本池  $S_{IS}$  中每一个样本点对应的失效域指示函数  $I_F(\mathbf{x}^{(j)})$ , 并利用式 (16) 求得失效概率  $\hat{P}_f$ 。

$$I_F(\mathbf{x}^{(j)}) = \begin{cases} 0 & g_k(\mathbf{x}^{(j)}) > 0 \\ 1 & g_k(\mathbf{x}^{(j)}) \leq 0 \end{cases} \quad (\mathbf{x}^{(j)} \in S_{IS}) \quad (26)$$

**步骤 8** 判断计算结果的收敛性。利用失效概率估计值的变异系数判断 AK-IS 法的收敛性, 其计算公式为

$$\text{Cov}(\hat{P}_f) = \sqrt{\frac{1}{N_{IS}-1} \left[ \frac{1}{N_{IS}} \sum_{j=1}^{N_{IS}} \left( I_F(\mathbf{x}_j) \frac{f_{I_F}^2(\mathbf{x}_j)}{h_{I_F}^2(\mathbf{x}_j)} \right) - \hat{P}_f^2 \right]} / \hat{P}_f \quad (27)$$

当  $\text{Cov}(\hat{P}_f) < 5\%$ , 即计算结果收敛, 否则继续按每次增加 1 000 个样本点, 增大样本池  $S_{IS}$  返回步骤 2 直至收敛。

### 3 航空齿轮泵滑动轴承可靠性及可靠性灵敏度分析

#### 3.1 航空齿轮泵滑动轴承弹流润滑压力分析

某型航空齿轮泵滑动轴承的几何尺寸及运行参数的均值和标准差如表 1 所示<sup>[19]</sup>。当滑动轴承在稳态下工作时, 油膜压力会使轴瓦产生弹性变形, 轴瓦弹性变形影响了油膜厚度, 进而影响油膜产生的动压力。通过有限差分法和有限元法耦合得到滑动轴承在均值下的轴瓦弹性变形, 如图 5 所示, 对应的压力分布及峰值压力如图 6 所示。可以看出, 轴瓦弹性变形与动压润滑油膜压力形状相同, 由于轴瓦弹性变形改变了轴承间隙膜厚分布, 进而影响了油膜压力。

与不考虑轴瓦弹性的圆周方向压力对比如图 7 所示。可以看出, 弹性轴瓦的动压润滑油膜压力沿圆周方向呈现“山峰”形状, 其峰值压力与

表 1 滑动轴承随机变量分布参数

Table 1 Random variable distribution parameters of journal bearing

参数	分布形式	平均值	标准差
轴承宽度 $B/\text{mm}$	Normal	30	0.011
齿轮轴半径 $R/\text{mm}$	Normal	9.97	0.01
半径间隙 $C/\text{mm}$	Normal	0.03	0.005
转速 $n/(\text{r} \cdot \text{min}^{-1})$	Normal	7 000	500
介质黏度 $\eta/(\text{Pa} \cdot \text{s})$	Normal	$9.66 \times 10^{-4}$	0.1
偏心率 $\varepsilon$	Normal	0.8	0.01

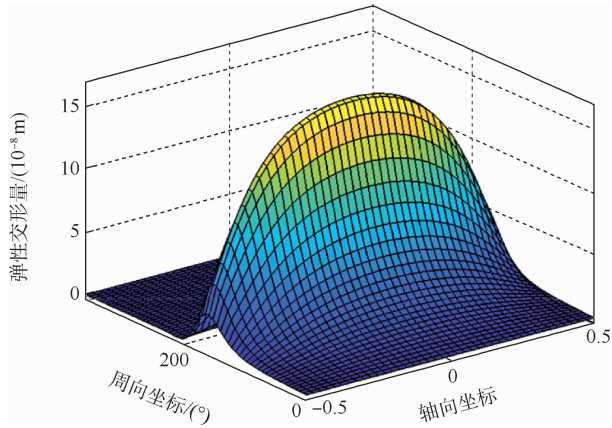


图 5 轴承变形

Fig. 5 Bearing deformation

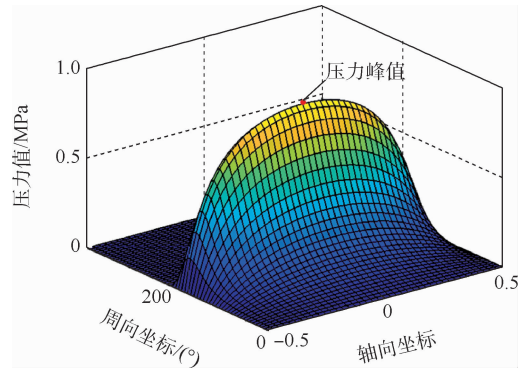


图 6 弹流动压润滑压力分布及压力峰值

Fig. 6 Pressure distribution and pressure peak value of EHL

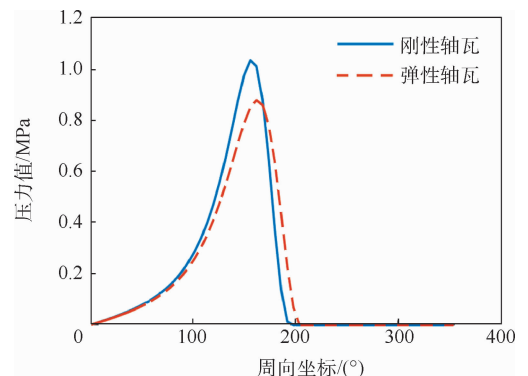


图 7 周向油膜压力分布对比

Fig. 7 Comparison of circumferential oil film pressure distribution

刚性轴瓦相比降低了 15.04%。因为刚性轴瓦油膜压力峰值较大,如果利用其进行滑动轴承可靠性分析,可靠性会降低,所以在进行航空齿轮泵滑动轴承时必须考虑轴瓦弹性变形。

### 3.2 计算结果及分析

航空齿轮泵滑动轴承动压润滑性能可靠性是衡量滑动轴承性能的主要指标,会受到滑动轴承几何参数和工况参数的影响。在这些参数发生变化时,滑动轴承由轴颈与轴瓦相对运动产生的动压力峰值必然发生变化。当齿轮泵径向载荷超出滑动轴承动压油膜承载力时,即压力峰值超过设计指标时,固体壁面直接接触导致轴承的灼伤、胶合等故障。通过滑动轴承弹流润滑分析,找到滑动轴承动压润滑压力最大处,即为最容易失效的部位,针对该位置构建极限状态功能函数:

$$G(X) = P_m - P(X) \quad (28)$$

式中: $P_m$ 为许用压力; $P(\cdot)$ 为滑动轴承弹流耦合求解的峰值压力; $X$ 为输入变量。

在对航空齿轮泵滑动轴承进行可靠性分析时,采用 MCS 法调用真实弹流润滑模型  $10^7$  次对滑动轴承进行次数值模拟,并将其结果作为精确解与 IS、AK-MCS 法进行对比,可靠性分析结果如表 2 所示。表中: $N_{\text{call}}$ 为滑动轴承弹流润滑模型调用次数; $N$ 为备选样本池规模; $\xi$ 为相对误差。可知,MCS、IS、AK-MCS、AK-IS 四种方法失效概率估计值较为接近;与 MCS 法相比,IS 法计算精度最高,AK-IS 法和 AK-MCS 法相近,相对误差仅 0.432%。与 AK-MCS 法相比,AK-IS 法备选样本池较小,仅为 AK-MCS 法的 0.85%,大大缩短了自适应学习建立 Kriging 代理模型的时间,提高了构建 Kriging 代理模型的效率;同时 AK-IS 法所调用有限元弹流润滑模型的次数明显少于另外 3 种方法,仅调用真实函数 49 次,比 IS 法少 9 951 次,比 AK-MCS 法少 151 次,计算效率明显提升,对航空齿轮泵滑动轴承可靠性分析有较好的实用性。

根据可靠性灵敏度计算公式,通过失效概率  $\hat{P}_f$  对上述 6 个变量的均值和方差求偏导,得到灵敏度计算结果,如表 3 所示。

表 2 滑动轴承可靠度计算结果

Table 2 Calculation results of reliability of sliding bearing

方法	$N_{\text{call}}$	$N$	$\hat{P}_f/10^{-4}$	$\xi/\%$
MCS	$10^7$		3.22	0
IS	$10^3$		3.230 2	0.317
AK-MCS	200	1 290 000	3.233 2	0.41
AK-IS	49	11 000	3.233 9	0.432

表 3 滑动轴承灵敏度计算结果

Table 3 Calculation of results of sensitivity analysis of sliding bearing

参数	数值	参数	数值
$\partial P_f/\partial \mu_B$	$1.55 \times 10^{-2}$	$\partial P_f/\partial \sigma_B$	$7.68 \times 10^{-6}$
$\partial P_f/\partial \mu_R$	0.261	$\partial P_f/\partial \sigma_R$	$1.97 \times 10^{-3}$
$\partial P_f/\partial \mu_C$	$-2.32 \times 10^2$	$\partial P_f/\partial \sigma_C$	$7.78 \times 10^2$
$\partial P_f/\partial \mu_n$	$2.15 \times 10^{-7}$	$\partial P_f/\partial \sigma_n$	$6.66 \times 10^{-8}$
$\partial P_f/\partial \mu_\eta$	1.422 476	$\partial P_f/\partial \sigma_\eta$	0.625
$\partial P_f/\partial \mu_\varepsilon$	0.012 163	$\partial P_f/\partial \sigma_\varepsilon$	$3.57 \times 10^{-3}$

从  $\partial P_f/\partial \mu_x$  中可以看出,变量  $B$ 、 $R$ 、 $n$ 、 $\eta$ 、 $\varepsilon$  的均值灵敏度为正值,即随着变量  $B$ 、 $R$ 、 $n$ 、 $\eta$ 、 $\varepsilon$  的均值的增大,失效概率随之增大,航空齿轮泵滑动轴承趋于危险;变量  $C$  的灵敏度为负值,即随着变量  $C$  均值的增大,失效概率随之减小,滑动轴承趋于安全。从  $\partial P_f/\partial \sigma_x$  中可以看出,变量  $B$ 、 $R$ 、 $C$ 、 $n$ 、 $\eta$ 、 $\varepsilon$  的标准差灵敏度为正值,其失效概率随变量的标准差增大而增大,可靠性减小。在这 6 个变量中,可靠性对变量  $C$  最为敏感, $B$ 、 $R$ 、 $\eta$ 、 $\varepsilon$  次之,对  $n$  最不敏感。在滑动轴承设计过程中,应注意半径间隙  $C$  不确定性的控制,在滑动轴承运行过程中可放宽对转速  $n$  的限制,保证轴承有较高的可靠性。

## 4 结论

本文采用 AK-IS 法对航空齿轮泵滑动轴承进行可靠性分析,得到如下结论:

1) 考虑航空齿轮泵滑动轴承轴瓦弹性变形时,油膜的压力峰值明显低于轴瓦刚性时的油膜压力峰值,表明轴承的弹性变形在轴承可靠性分析中的影响不能忽视,也说明本文采用弹流润滑模型进行可靠性分析结果的准确性。

2) 采用 AK-IS 法对滑动轴承进行动压润滑可靠性分析,其备选样本池规模仅为 AK-MCS 法的 0.85% (减少电脑计算存储问题),调用真实函数次数仅为 IS 法的 4.9%,与 MCS 法相比相对误差仅为 0.432%,极大改善了航空齿轮泵滑动轴承动压润滑可靠性的计算效率。

3) 通过灵敏度计算结果,定量描述了航空齿轮泵滑动轴承的 6 个变量对可靠性的影响程度。其中, $B$ 、 $R$ 、 $n$ 、 $\eta$ 、 $\varepsilon$  均值和方差与滑动轴承动压润滑可靠性都为负相关,为确保轴承足够的可靠性,应该在滑动轴承设计和运行时控制其不确定性。

4) 提出的基于弹流动压润滑的航空齿轮泵滑动轴承可靠性分析方法能有效提高轴承可靠性

评估的效率和精度,可为后续开展基于可靠性的航空齿轮泵滑动轴承优化设计提供理论基础。

### 参考文献 (References)

- [1] HUANG P. Numerical calculation of lubrication [M]. Singapore: John Wiley & Sons Singapore Pte. Ltd., 2013.
- [2] TAKABI J, KHONSARI M M. On the dynamic performance of roller bearings operating under low rotational speeds with consideration of surface roughness [J]. Tribology International, 2015, 86: 62-71.
- [3] 符江锋, 李华聪, 樊丁, 等. 一种自冷却结构燃油泵滑动轴承润滑特性分析 [J]. 北京航空航天大学学报, 2018, 44(10): 2052-2062.
- FU J F, LI H C, FAN D, et al. Lubrication characteristics analysis of a type of self-cooling structure fuel pump sliding bearings [J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(10): 2052-2062 (in Chinese).
- [4] CHENG F, JI W X. A new model of water-gas turbulent lubrication for analysis of the static and dynamic characteristics in a journal bearing [J]. Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology, 2016, 230(12): 1439-1451.
- [5] LIN Q Y, BAO Q K, LI K J, et al. An investigation into the transient behavior of journal bearing with surface texture based on fluid-structure interaction approach [J]. Tribology International, 2018, 118: 246-255.
- [6] 王琳, 张瑜, 陈国定. 高速高比压织构滑动轴承热弹流润滑分析 [J]. 西北工业大学学报, 2019, 37(4): 751-756.
- WANG L, ZHANG Y, CHEN G D. Thermo-elastohydrodynamic study on textured journal bearing with high-speed and high-specific-pressure [J]. Journal of Northwestern Polytechnical University, 2019, 37(4): 751-756 (in Chinese).
- [7] 金燕, 刘少军. 基于人工神经网络的航空轴承疲劳可靠性分析 [J]. 东北大学学报 (自然科学版), 2018, 39(6): 850-855.
- JIN Y, LIU S J. Fatigue reliability analysis of aviation bearings based on ANN [J]. Journal of Northeastern University (Natural Science), 2018, 39(6): 850-855 (in Chinese).
- [8] 冯吉路, 孙志礼, 赵坚, 等. 基于 AK-MCS 法的主轴系统振动可靠性分析 [J]. 振动与冲击, 2019, 38(18): 135-140.
- FENG J L, SUN Z L, ZHAO J, et al. Vibration reliability analysis on a spindle system based on AK-MCS method [J]. Journal of Vibration and Shock, 2019, 38(18): 135-140 (in Chinese).
- [9] 胡明用, 李昌, 韩兴, 等. 基于 Monte-Carlo 法的滚动轴承-转子系统非线性随机振动可靠性分析 [J]. 轴承, 2018(8): 27-31.
- HU M Y, LI C, HAN X, et al. Reliability analysis on nonlinear random vibration of rolling bearing-rotor system based on Monte-Carlo method [J]. Bearing, 2018(8): 27-31 (in Chinese).
- [10] 李洪双, 马远卓. 结构可靠性分析与随机优化设计的统一方法 [M]. 北京: 国防工业出版社, 2015: 51-55.
- LI H S, MA Y Z. Unified methods for structural reliability analysis and stochastic optimization design [M]. Beijing: National Defense Industry Press, 2015: 51-55 (in Chinese).
- [11] ZHANG F, LU Z Z, CUI L J, et al. Reliability sensitivity algorithm based on stratified importance sampling method for multiple failure modes systems [J]. Chinese Journal of Aeronautics, 2010, 23(6): 660-669.
- [12] TSIRIKOGLOU P, ABRAHAM S, CONTINO F, et al. A hyper-parameters selection technique for support vector regression models [J]. Applied Soft Computing, 2017, 61: 139-148.
- [13] 洪林雄, 李华聪, 彭凯, 等. 基于高效搜索方法的可靠性分析改进响应面法 [J]. 北京航空航天大学学报, 2020, 46(1): 95-102.
- HONG L X, LI H C, PENG K, et al. Improved response surface method of reliability analysis based on efficient search method [J]. Journal of Beijing University of Aeronautics and Astronautics, 2020, 46(1): 95-102 (in Chinese).
- [14] KAYMAZ I. Application of Kriging method to structural reliability problems [J]. Structural Safety, 2005, 27(2): 133-151.
- [15] 洪林雄, 李华聪, 彭凯, 等. 基于改进学习策略的 Kriging 模型结构可靠度算法 [J]. 西北工业大学学报, 2020, 38(2): 412-419.
- HONG L X, LI H C, PENG K, et al. Structural reliability algorithms of Kriging model based on improved learning strategy [J]. Journal of Northwestern Polytechnical University, 2020, 38(2): 412-419 (in Chinese).
- [16] 吕震宙, 宋述芳, 李璐祎. 结构/机构可靠性设计基础 [M]. 西安: 西北工业大学出版社, 2019: 216-236.
- LYU Z Z, SONG S F, LI L Y. Reliability design basis of structural mechanism [M]. Xi'an: Northwestern Polytechnical University Press, 2019: 216-236 (in Chinese).
- [17] ECHARD B, GAYTON N, LEMAIRE M, et al. A combined importance sampling and Kriging reliability method for small failure probabilities with time-demanding numerical models [J]. Reliability Engineering & System Safety, 2013, 111: 232-240.
- [18] ZHOU H, DU R L, XIE A H, et al. Investigations of the micro surface shape for the gear-shaft/journal-bearing interface in water hydraulic internal gear pumps [J]. Advances in Mechanical Engineering, 2017, 9(11): 168781401774257.
- [19] 孙志礼, 陈良玉. 实用机械可靠性设计理论与方法 [M]. 北京: 科学出版社, 2003: 55-60.
- SUN Z L, CHEN L Y. Practical mechanical reliability theory and method [M]. Beijing: Science Press, 2003: 55-60 (in Chinese).

## Reliability analysis of journal bearings inside aero-gear pump based on AK-IS method

GAO Ning, LI Huacong<sup>\*</sup>, HONG Linxiong, CAO Rui, FU Jiangfeng

(College of Power and Energy, Northwestern Polytechnical University, Xi'an 710072, China)

**Abstract:** To study the hydrodynamic lubrication reliability of journal bearings inside aero-gear pumps running at high speed and with low medium viscosity, an elastohydrodynamic lubrication (EHL) model considering elastic deformation of bearing bush is established by coupling Reynolds lubrication equation with influence matrix. Considering the uncertainty caused by the size tolerance and operating conditions of journal bearings, the peak pressure of hydrodynamic lubrication, which is closely related to the lubrication characteristics of journal bearings, is taken as the reliability criterion. The AK-IS method combining adaptive Kriging and importance sampling is used to calculate the reliability and sensitivity of the hydrodynamic lubrication characteristics of journal bearings. The results show that considering bearing deformation, the pressure peak value of journal bearings is 15.04% lower than that of rigid bearings, and hence the influence of EHL on hydrodynamic lubrication of bearings cannot be neglected. The results also show the accuracy and efficiency of the AK-IS based reliability analysis of hydrodynamic lubrication of journal bearings. Moreover, the degree of the effect of various uncertainties on the reliability of hydrodynamic lubrication is different, with the radius clearance of bearings being the most sensitive to reliability, and the rotational speed the least.

**Keywords:** aero-gear pump; journal bearing; elastohydrodynamic lubrication (EHL); AK-IS method; reliability analysis

**Received:** 2020-12-26; **Accepted:** 2021-05-07; **Published online:** 2021-05-12 16:38

**URL:** [kns.cnki.net/kcms/detail/11.2625.V.20210512.1620.003.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210512.1620.003.html)

**Foundation item:** National Science and Technology Major Project (2017-V-0013-0065)

**\* Corresponding author.** E-mail: [lihuacong@nwpu.edu.cn](mailto:lihuacong@nwpu.edu.cn)

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2020.0715

# 吸气流动控制对翼身融合飞机气动特性的影响

贾媛, 曹翔, 吴江浩\*

(北京航空航天大学 交通科学与工程学院, 北京 100083)

**摘要:** 以采用分布式动力的翼身融合飞机为研究对象, 探究了吸气流动控制方式(吸气位置和吸气动量)对飞机起飞和巡航状态下气动特性的影响规律, 解释了吸气流动控制影响翼身融合飞机气动特性的机理。研究表明: 起飞大攻角状态下, 采用外翼段吸气方案(吸气位置为  $0.05c$ , 吸气动量为  $0.02$ ), 飞机最大升力系数与无吸气状态相比提升  $7.16\%$ ; 巡航状态下, 采用中心体段吸气方案(吸气位置为  $0.6c$ , 吸气动量为  $0.0125$ ), 可改善动力系统的压力分布, 飞机升阻比与无吸气状态相比最大提升  $2.14\%$ 。

**关键词:** 翼身融合; 流动控制; 吸气位置; 吸气动量; 气动特性

**中图分类号:** V221.3

**文献标志码:** A

**文章编号:** 1001-5965(2022)06-1065-07

翼身融合 (blended-wing-body, BWB) 飞机是一种具有高度可行性和巨大潜力的客运飞机布局<sup>[1-3]</sup>。翼身融合布局的高升阻比带来更少的燃油消耗, 不仅为航空公司创造更大的经济效益, 其低排放低污染还切合时代绿色航空业发展的基本要求<sup>[4]</sup>。但翼身融合飞机也存在很多挑战。翼身融合飞机设计飞行马赫数大, 存在起飞所需速度较大、进场速度和姿态角过大等问题, 影响客机运行安全<sup>[5-6]</sup>。由于中心体部分发动机抽吸机身边界层带来进气面畸变和总压恢复系数分布不均, 影响发动机效率。对于起降状态, 现已有三段翼布局改善低速大攻角的气动特性<sup>[7]</sup>, 但需要引入诸多机械装置和部件才能实现, 增加飞机起飞重量, 如果可以通过微小的吹吸气装置实现上翼面的流动控制, 对于减小起飞重量十分有益。同时, 因难以在中心体部分外加机械装置来实现流动控制, 微小型的吹吸气流动控制装置对于巡航气动性能改善十分具有研究价值<sup>[8]</sup>。

采用简单的气流控制方式实现流动控制是未来的研究方向<sup>[9]</sup>。吹吸气流动控制技术是现阶段

研究较为深入且效果较为明显的流动控制手段。这种主动流动控制手段, 具有在高效率飞行条件下工作同时非设计工况下关闭从而保障原有气动效能的优越性<sup>[10]</sup>。张志勇等<sup>[11]</sup>针对 NACA0012 翼型开展了低雷诺数下吹吸气射流作用的研究, 贾宝佳<sup>[12]</sup>也对 NACA0012 翼型作了吹吸气相关研究, 结果表明, 在吹吸气共同作用下, 后缘涡分离消失, 抑制分离效果显著。Huang 等<sup>[13]</sup>的研究表明, 吸气机制比吹气机制更具有优势。Genc 和 Kaynak<sup>[14]</sup>采用数值模拟技术研究了 NACA2415 翼型在  $20 \times 10^4$  雷诺数模型下的吹吸气流动控制, 发现较小的吹气动量对于分离的抑制效果没有大吹气动量的效果明显。Agarwal 等<sup>[15]</sup>在对 NACA0015 翼型进行吹气流动控制研究时, 发现吹气控制的效果十分依赖于吹气动量系数。Wahidi 和 Bridges<sup>[16]</sup>的研究表明, 当动量系数增大到一定值时, 机翼表面的阻力呈现出了非常明显的减小趋势。

本文以 350 座分布式动力的翼身融合飞机 (BWB-350) 为研究构型, 对机翼表面的主动吸气

收稿日期: 2020-12-26; 录用日期: 2021-01-30; 网络出版时间: 2021-02-22 16:25

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20210222.1603.002.html](http://kns.cnki.net/kcms/detail/11.2625.V.20210222.1603.002.html)

\* 通信作者。E-mail: buaawjh@buaa.edu.cn

**引用格式:** 贾媛, 曹翔, 吴江浩. 吸气流动控制对翼身融合飞机气动特性的影响[J]. 北京航空航天大学学报, 2022, 48(6): 1065-1071. JIA Y, CAO X, WU J H. Influence of suction flow control on aerodynamic characteristics of blended-wing-body aircraft[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(6): 1065-1071 (in Chinese).

流动控制进行研究。通过搭建带有吸气流动控制系统的综合模型,探究此类流动控制技术对飞机气动特性的影响。探究不同吸气系统排布方式和吸气参数对全机的气动特性的影响,并提炼适用于此类飞机的吸气系统设计原则。

## 1 模型与方法

### 1.1 模型定义

#### 1.1.1 飞机模型

本文所采用的分布式动力 BWB-350 飞机模型取自闫万方等<sup>[2]</sup>的研究。BWB-350 飞机主要是由中心体、融合段、外翼段组成,如图 1 所示,全机翼展为 68 m,等效梯形面积为 560 m<sup>2</sup>。

外翼段翼型为超临界翼型 NASA SC(2)-0706。表 1 给出了飞机的总体参数,表 2 给出了飞机主要几何参数。

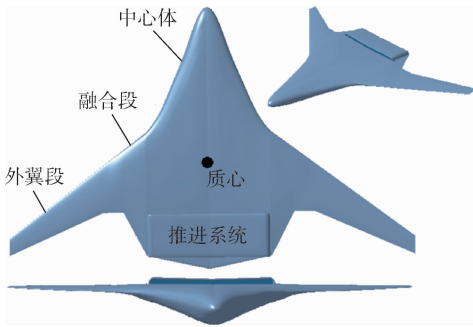


图 1 带有边界层吸入的翼身融合飞机

Fig. 1 Blended-wing-body aircraft with boundary layer suction

表 1 BWB-350 总体参数

Table 1 Overall parameters of BWB-350

参数	数值
最大航程/km	14 800
巡航高度/m	11 000
巡航马赫数	0.85
起飞离地速度/(m·s <sup>-1</sup> )	84
进场速度/(m·s <sup>-1</sup> )	72
最大起飞重量/kg	232 000

表 2 BWB-350 几何参数

Table 2 Geometric parameters of BWB-350

参数	数值
翼展/m	68
参考面积/m <sup>2</sup>	560
外翼前缘后掠角/(°)	36
展弦比	8.14
平均气动弦长/m	10.6
重心距前缘距离/m	27.3

#### 1.1.2 吸气流动控制方案

本文研究吸气技术对 BWB-350 飞机起飞状态和巡航状态下气动性能的影响。选取吸气位置

和吸气动量系数  $C_{m\_suc}$  两个变量,分别探究其对升力系数和升阻比的影响。吸气流动控制方案的布置示意图如图 2 所示。8 条点划线代表吸气槽,位于翼段的控制面上相同弦长位置处。将吸气槽的位置进行分组,第 1 组为 L1,第 2 组为 L2。L1 沿展向布置于外翼段全展长,L2 沿展向布置于中心体全展长。吸气槽用“L+组别+位置”表示,如“L1-0.5”表示该吸气槽位于第 1 组,位置在  $L=0.5c$  处, $L$  为不同控制面翼型上距离前缘的弦向距离, $L=0.5c$  代表吸气槽距离前缘的弦向距离为  $0.5c$ , $c$  为翼型弦长。为了不影响机翼及机身的结构强度设计,在实际工程中该方法由不连续的吸气小孔代替。

本文在 BWB-350 飞机上搭建吸气流动控制模型。起飞状态下,依次分别打开 L1 中吸气槽。巡航状态下,打开 L2 中的吸气槽,探究 L2 中 4 种吸气位置对升阻比的影响。

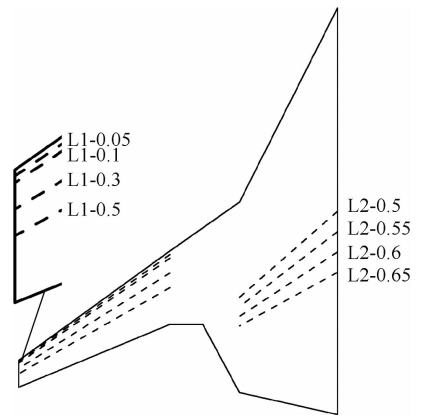


图 2 中心体及外翼段吸气布置示意图

Fig. 2 Schematic diagram of suction arrangement of center body and outer wing section

### 1.2 计算方法

以三维定常雷诺平均 Navier-Stokes (RANS) 作为控制方程,选用基于压力的隐式耦合求解方法的二阶精度求解,湍流模型为 S-A 模型,物面采用无滑移边界条件,发动机入口边界条件为压力出口,尾喷管出口边界条件为质量流量入口。计算网格为 C-H 结构网格。

$$C_{m\_suc} = \frac{hV_c\rho_c}{cV_\infty\rho_\infty} \quad (1)$$

式中: $C_{m\_suc}$  为动量系数; $h$  为吸气宽度; $V_c$  为吸气槽吸气速度; $\rho_c$  为吸气槽吸气密度; $V_\infty$  为无穷远处(远场)来流速度; $\rho_\infty$  为无穷远处(远场)来流密度。

$$AC = \frac{X_{AC}}{c} \quad (2)$$

式中:AC 为全机气动中心。

### 1.3 方法验证

如表3~表6所示,分别在马赫数  $Ma$  为 0.21,迎角  $\alpha$  为  $10^\circ$  和马赫数  $Ma$  为 0.85,迎角  $\alpha$  为  $2.6^\circ$  时,对本文所采用的计算网格进行验证。选取边界层第1层网格高度  $y_{\max}^+$  不同的网格进行比较,在  $y_{\max}^+$  为 2 时,计算结果趋于相对稳定。在第1层网格高度为  $2 \times 10^{-5}$  的基础上,网格密度为  $150 \times 10^4$  时,飞机的升力系数  $C_L$  和阻力系数  $C_d$  趋于稳定,故后续计算选用密度为  $150 \times 10^4$ 、第1层网格高度为  $2 \times 10^{-5}$  的网格。

表3 不同  $y_{\max}^+$  网格的计算结果对比 ( $Ma=0.21, \alpha=10^\circ$ )

Table 3 Comparison of calculation results of different  $y_{\max}^+$  grids ( $Ma=0.21, \alpha=10^\circ$ )

网格密度/ $10^4$	第1层网格高度	$y_{\max}^+$	$C_L$	$C_d$
150	$1 \times 10^{-3}$	140	0.984 2	0.116 43
150	$1 \times 10^{-4}$	16	0.989 3	0.117 39
150	$2 \times 10^{-5}$	2	0.989 9	0.117 89
150	$1 \times 10^{-5}$	0.8	0.989 7	0.117 92

表4 不同网格密度计算结果对比 ( $Ma=0.21, \alpha=10^\circ$ )

Table 4 Comparison of calculation results of different overall grid densities ( $Ma=0.21, \alpha=10^\circ$ )

网格密度/ $10^4$	第1层网格高度	$y_{\max}^+$	$C_L$	$C_d$
70	$2 \times 10^{-5}$	2	0.985 3	0.116 41
100	$2 \times 10^{-5}$	2	0.986 2	0.116 93
150	$2 \times 10^{-5}$	2	0.989 9	0.117 89
230	$2 \times 10^{-5}$	2	0.989 8	0.117 84

表5 不同  $y_{\max}^+$  网格的计算结果对比 ( $Ma=0.85, \alpha=2.6^\circ$ )

Table 5 Comparison of calculation results of different  $y_{\max}^+$  grids ( $Ma=0.85, \alpha=2.6^\circ$ )

网格密度/ $10^4$	第1层网格高度	$y_{\max}^+$	$C_L$	$C_d$
150	$1 \times 10^{-3}$	140	0.353 7	0.016 50
150	$1 \times 10^{-4}$	16	0.354 0	0.015 25
150	$2 \times 10^{-5}$	2	0.354 1	0.015 50
150	$1 \times 10^{-5}$	0.8	0.354 1	0.015 50

表6 不同网格密度计算结果对比 ( $Ma=0.85, \alpha=2.6^\circ$ )

Table 6 Comparison of calculation results of different overall grid densities ( $Ma=0.85, \alpha=2.6^\circ$ )

网格密度/ $10^4$	第1层网格高度	$y_{\max}^+$	$C_L$	$C_d$
70	$2 \times 10^{-5}$	2	0.354 5	0.016 50
100	$2 \times 10^{-5}$	2	0.353 9	0.015 75
150	$2 \times 10^{-5}$	2	0.354 1	0.015 50
230	$2 \times 10^{-5}$	2	0.354 1	0.015 49

### 1.4 飞行条件

本文研究所涉及的起飞和巡航状态的飞行条件参数如表7所示。

表7 飞行条件

Table 7 Flight conditions

状态	高度/m	马赫数	MFR	静压/Pa	静温/K	密度/ $(\text{kg} \cdot \text{m}^{-3})$
起飞	0	0.21	1.60	101 325	288.2	1.224 9
巡航	11 000	0.85	0.68	22 700	216.7	0.363 9

注:MFR为质量流率。

## 2 起飞状态

本文通过在外翼段添加一段吸气槽,主动吸气,抑制大攻角下外翼段流动的分流,延缓失速,提升最大升力系数。本节探究在外翼段布置吸气槽的最佳位置和最佳的吸气动量,确定起飞状态下外翼段的最佳控制方案。

### 2.1 吸气位置

在外翼段分别设置 L1-0.05、L1-0.1、L1-0.3、L1-0.5 四种位置方案,位置如图2所示。吸气角度  $b=90^\circ$ ,垂直于壁面。吸气宽度  $h=0.005c$ ,沿展向布置于外翼段全段,选取吸气动量系数  $=0.02$ ,飞行马赫数  $Ma=0.21$ 。发动机  $MFR=1.6$ 。

为从 0.05c、0.1c、0.3c、0.5c 四个位置分析出最佳吸气位置,图3(a)为起飞状态不同吸气位置下升力系数随攻角变化的曲线。其中,无吸气代表 BWB-350 飞机气动特性,选为基准模型状态,0.05c 表示布置 L1-0.05 后 BWB-350 飞机的气动特性。基准模型在  $\alpha=12^\circ$  以前,升力线斜率基本保持不变,线性段最大升力系数为 1.299。相较于其他位置吸气布置,吸气位置布置为 0.05c 时,线性段升力系数最大,为 1.392,比基准模型增加 7.16%。基准模型在  $13^\circ$  时,升力系数线性段结束,开始陡降,发生失速。0.05c 吸气流动控制方案,升力系数从  $14^\circ$  时才出现下降,失速攻角从  $12^\circ$  延缓至  $13^\circ$ 。在攻角大于  $9^\circ$  以后,0.1c、0.3c、0.5c 的吸气布置位置升力线斜率开始有不同程度的下降。随着吸气位置的后移,流动控制对于最大升力系数的提升效果逐渐变差,但由于后移吸气位置,升力系数没有出现明显的陡降,说明吸气位置的后移会使得失速发生的较为平稳,没有出现明显全翼段突然失速的过程。因此,对于起飞阶段来说,吸气位置布置于 0.05c 的方案线性段最长,最大升力系数也较无吸气基准状态有 8% 左右的提升,失速攻角推迟  $1^\circ$ ,效果最佳。

如图3(b)所示,攻角大于  $6^\circ$  时,吸气方案减

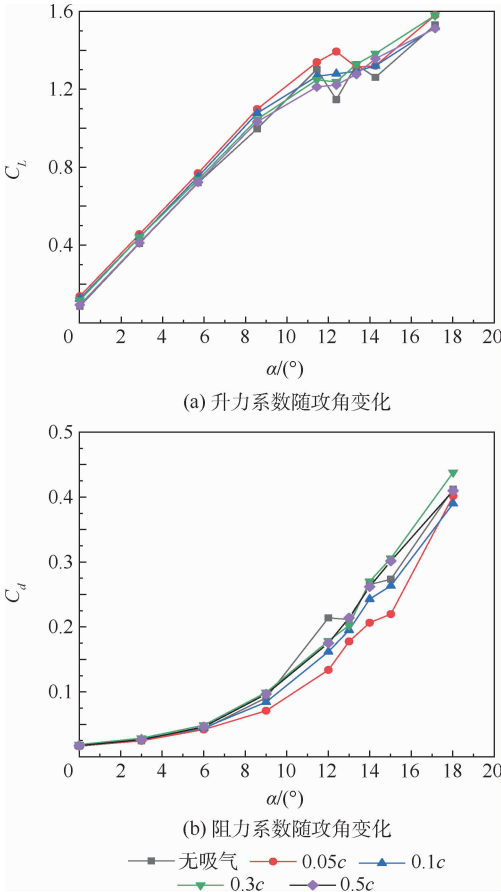


图3 不同弦向位置下的气动特性系数对比  
Fig.3 Comparison of aerodynamic characteristic coefficients at different chordal positions

阻效果明显, 12°时减阻效果最为明显。这是因为在无吸气的情况下, 此时翼面已经出现了很大的分离泡, 而吸气方案下的翼面仍然保持着稳定的表面流动, 没有分离迹象, 而吸气能够降低上翼面的流速, 减小逆压梯度, 使得分离减弱流动稳定, 吸气位置越靠前逆压梯度越小。0.05c 位置吸气方案的阻力减小最为突出, 效果最差是 0.5c 的吸气位置, 对比无吸气方案的阻力系数, 各吸气方案均有不同程度的减阻效果。

选取 0.05c 位置处的吸气方案, 计算俯仰力矩系数, 俯仰力矩系数范围为 -0.10 ~ 0.05, 力矩系数较小, 易于配平。

2.2 吸气动量

设置吸气动量系数 0.002 5、0.005、0.01、0.02、0.03 五种吸气方案, 起飞吸气位置基于 2.1 节研究选取 L1-0.05。吸气宽度  $h = 0.005c$ , 飞行马赫数  $Ma = 0.21$ 。

图 4(a) 为起飞状态下不同吸气动量系数下升力系数随攻角变化的曲线。可知, 起飞阶段下, 基准模型在 13°时, 升力系数线性段结束, 升力系数开始陡降, 发生失速。

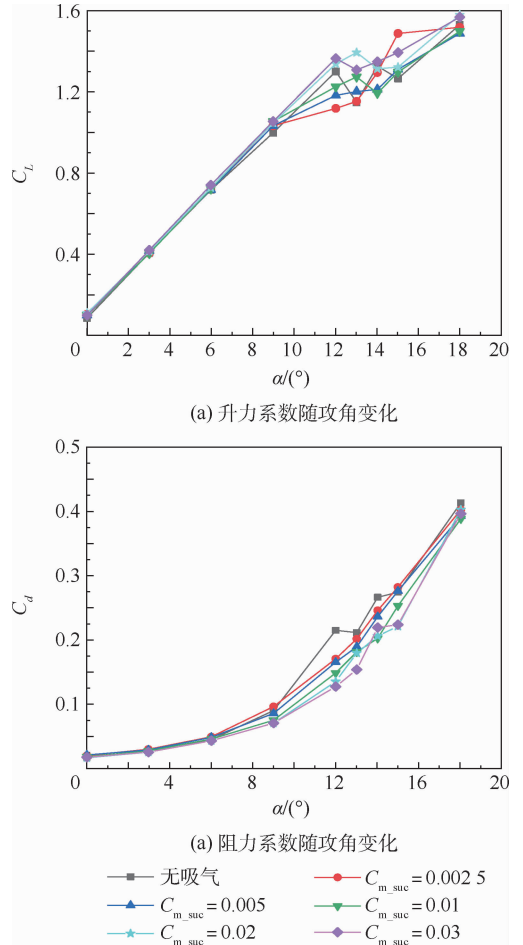


图4 不同动量系数下的气动特性系数对比  
Fig.4 Comparison of aerodynamic characteristic coefficients under different momentum coefficients

当吸气动量系数在 0.01 及以下时, 升力系数在攻角大于 9°以后, 升力线斜率出现一定程度的下降。当吸气动量系数在 0.01 以上时, 升力系数从 14°时才出现下降, 失速攻角从 12°延缓至 13°。吸气动量为 0.02 时, 线性段升力系数最大, 为 1.392, 相较于基准模型最大升力系数从 1.299 增至 1.392, 增量为 7.16%。对于起飞阶段来说, 吸气动量系数设置为 0.02 的方案线性段最长, 最大升力系数也较无吸气基准状态有 8% 左右的提升, 失速攻角推迟 1°, 综合来看效果最好。

如图 4(b) 所示, 与无吸气方案相比, 吸气方案均有不同程度的减阻效果, 0.03 动量系数吸气方案对于阻力的减小效果最好, 13°时减阻效果最为明显, 其次为 0.02 动量系数的方案。在失速攻角前, 随着动量系数的增大, 阻力系数相较于基准无吸气状态不断减小。考虑增升减阻效果, 采用 0.02 动量系数方案。

选取 0.05c 吸气位置、0.02 吸气动量方案, 计算俯仰力矩系数, 俯仰力矩系数范围为 -0.10 ~ 0.05, 力矩系数较小, 易于配平。

综上,起飞状态下, L1-0.05 位置、0.02 动量系数的吸气方案效果较好。

### 3 巡航状态

由于半埋式发动机边界层抽吸带来了动力系统进气畸变和总压分布不均问题,且中心体部分无法加装机械装置来实现气动性能的改善,吸气流动控制成为了十分有效的控制手段。本节从吸气流动控制入手,研究不同吸气位置和吸气动量对翼身融合飞机巡航动力系统气动特性的影响。

#### 3.1 吸气位置

在中心体设置 L2-0.5、L2-0.55、L2-0.6、L2-0.65 四种位置方案,位置如图 2 所示。吸气动量系数设定为 0.007 5,吸气角度为 90°,吸气宽度  $h = 0.005c$ ,巡航攻角设定为 2.6°,巡航马赫数设定为 0.85,动力系统进气质量流率 MFR 设定为 0.68。

添加吸气流动控制后,动力系统前方压力分布及整流罩部分没有明显变化,如图 5 所示,表明吸气流动控制对动力系统部分没有不良气动影响。对比无吸气的情况,随着吸气流动控制的引入,压差阻力减小,升阻比提高。随着吸气位置后移至 0.6c,吸气位置前的低压区增大,升力上升,吸气位置后的高压区范围增加,阻力减小。

当加入吸气槽分别位于 0.5c、0.55c、0.6c、0.65c 时,升力系数  $C_L$  分别为 0.351、0.352、0.354、0.358。阻力系数  $C_d$  如图 6(a) 所示。计算飞机的升阻比  $K$ ,如图 6(b) 所示,当吸气槽位于 0.6c 时,飞机获得的最大升阻比为 22.33。俯仰力矩系数  $C_m$  在 -0.02 ~ 0.02 之间波动,不对控制机构提出过高的配平要求。

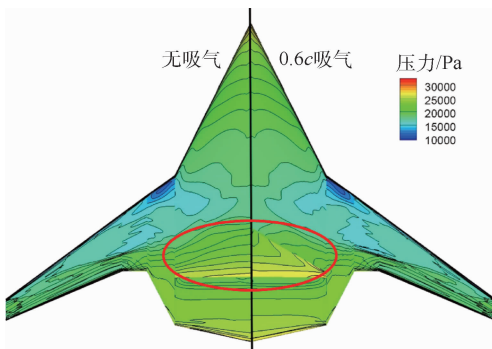


图 5 相同飞行条件下无吸气和 0.6c 吸气压力云图对比

Fig. 5 Comparison of pressure contour of no-suction and 0.6c suction under the same flight conditions

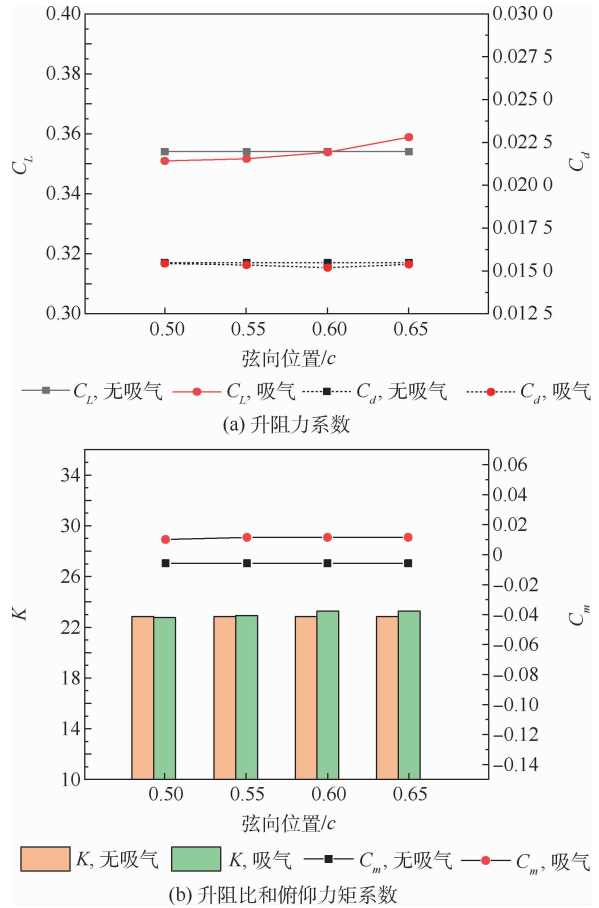


图 6 巡航状态时不同弦向位置下的气动特性系数对比  
Fig. 6 Comparison of aerodynamic characteristic coefficients at different chordal positions in cruising state

#### 3.2 吸气动量

选取升阻比最高的 L2-0.6 吸气位置作为本节的研究基础。设置  $C_{m\_suc} = 0.002 5、0.005、0.007 5、0.01、0.012 5$  五种吸气动量系数方案,吸气角度为 90°,吸气宽度  $h = 0.005c$ ,巡航攻角设定为 2.6°,巡航马赫数设定为 0.85,动力系统进气质量流率 MFR 设定为 0.68。

如图 7 所示,在吸气位置前方,随着吸气动量系数增加,低压区的范围增加,中心体部分的升力增加。在吸气位置后方,随着吸气动量系数的增加,高压区的范围增加,此部分高压区位于中心体部分背风区,高压区的存在有利于减小压差阻力,因此随着动量系数的增加,全机阻力减小。

当加入吸气动量分别为 0.002 5、0.005、0.007 5、0.01、0.012 5 时,升力系数  $C_L$  分别为 0.349、0.351、0.352、0.353、0.354。当吸气动量为 0.012 5 时,可以获得最大的升力系数。阻力系数  $C_d$  如图 8(a) 所示。计算飞机的升阻比,如图 8(b) 所示,吸气流动控制的升阻比与无吸气时的升阻比大致相等,当吸气动量系数为 0.012 5 时,飞机获得最大的升阻比 23.34。吸气流动控

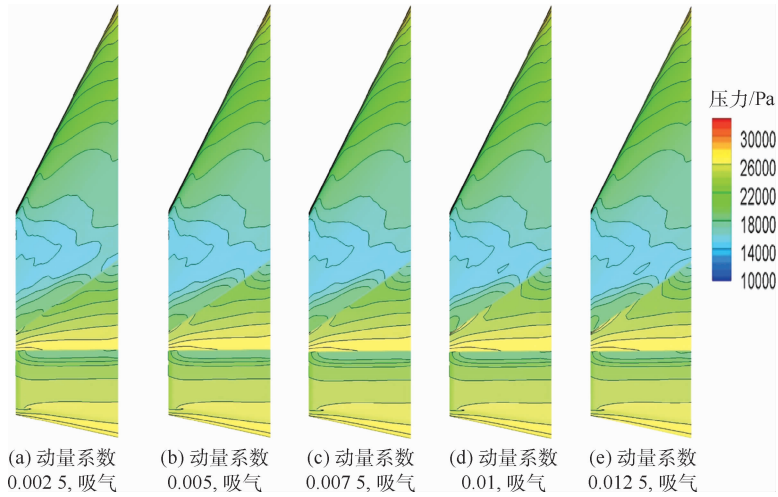


图7 不同吸气动量系数下中心体部分压力对比

Fig. 7 Comparison of pressures in central body under different suction momentum coefficients

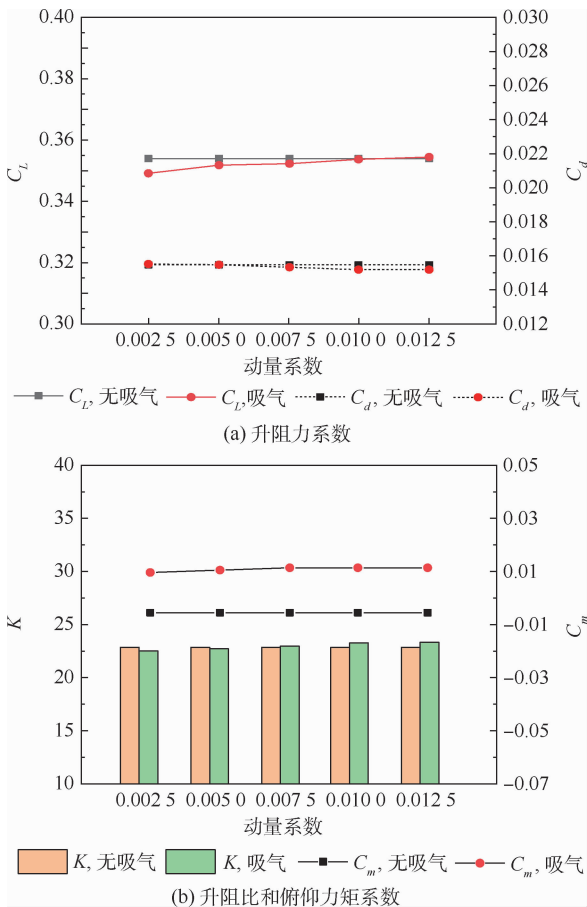


图8 巡航状态时不同动量系数下的气动特性系数对比

Fig. 8 Comparison of aerodynamic characteristic coefficients at different momentum coefficients in cruising state

制时,飞机的俯仰力矩系数稳定在 0.01,不对控制机构提出过高的配平要求。

综上,对于巡航状态中心体动力系统的流动控制,L2-0.6位置、0.0125吸气动量系数方案较优。

## 4 结论

本文以分布式动力的翼身融合布局飞机为研

究对象,研究吸气流动控制对于起飞状态气动特性、巡航状态动力系统气动特性的影响。研究了起飞状态(不同攻角、 $MFR = 1.6$ )下吸气流动控制对外翼段流动分离的抑制效果,探究不同吸气位置和吸气动量对流动分离和气动特性的影响。在巡航状态( $\alpha = 2.6^\circ$ ,  $MFR = 0.68$ ),研究了中心体上不同吹吸气位置和吸气动量对动力系统进气性能和气动特性的改善效果。本文研究主要得到以下结论:

1) 提出了一种针对BWB-350飞机的吸气流动控制方案,此方案能够有效提升该飞机的气动特性。

2) 起飞状态下,外翼段采用L1-0.05吸气位置、0.02吸气动量,飞机最大升力系数与无吸气状态相比提升7.16%,降低起飞距离,提升飞机起降安全。

3) 巡航状态下,中心体采用L2-0.6吸气位置、0.0125吸气动量,升阻比与无吸气状态相比最大提升2.14%,提高经济性。

## 参考文献 (References)

- [1] 索欣诗.翼身融合布局大型客机总体方案综合分析评价与优化[D].南京:南京航空航天大学,2017:1-20.  
SUO X S. Integrated analysis, evaluation and optimization in conceptual design of blended wing body commercial aircraft [D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2017:1-20 (in Chinese).
- [2] 闫万方,吴江浩,张艳来.分布式推进关键参数对BWB飞机气动特性影响[J].北京航空航天大学学报,2015,41(6):1055-1065.  
YAN W F, WU J H, ZHANG Y L. Effects of distributed propulsion crucial variables on aerodynamic performance of blended wing body aircraft [J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(6):1055-1065 (in Chinese).
- [3] LIEBECK R H. Design of the blended wing body subsonic transport [J]. Journal of Aircraft, 2004, 41(1):10-25.
- [4] 朱自强,王晓璐,吴宗成,等.民机的一种新型布局形式:翼

- 身融合体飞机[J]. 航空学报, 2008, 29(1): 49-59.  
ZHU Z Q, WANG X L, WU Z C, et al. A new type of transport-blended wing body aircraft[J]. Acta Aeronautica et Astronautica Sinica, 2008, 29(1): 49-59 (in Chinese).
- [5] 邓海强. 翼身融合布局无人机总体多学科设计优化研究[D]. 南京: 南京航空航天大学, 2017: 6-17.  
DENG H Q. Multidisciplinary design optimization for preliminary design of unmanned aerial vehicle with blended wing body[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2017: 6-17 (in Chinese).
- [6] 王刚, 张彬乾, 张明辉, 等. 翼身融合民机总体气动技术研究进展与展望[J]. 航空学报, 2019, 40(9): 623046.  
WANG G, ZHANG B Q, ZHANG M H, et al. Research progress and prospect for conceptual and aerodynamic technology of blended-wing-body civil aircraft[J]. Acta Aeronautica et Astronautica Sinica, 2019, 40(9): 623046 (in Chinese).
- [7] 刘铁中. 三翼面布局飞机低速气动性能的试验研究[D]. 北京: 北京航空航天大学, 1999: 6-12.  
LIU T Z. Experimental research on low-speed aerodynamic performance of aircraft with three wings[D]. Beijing: Beihang University, 1999: 6-12 (in Chinese).
- [8] 于哲慧, 刘沛清. CJ818 高升力构型吹吸气流动控制研究[J]. 民用飞机设计与研究, 2009(S1): 12-19.  
YU Z H, LIU P Q. Research on blowing and suction flow control of CJ818 high-lift configuration[J]. Civil Aircraft Design and Research, 2009(S1): 12-19 (in Chinese).
- [9] 朱佳晨, 史志伟, 孙琪杰, 等. 基于主动流动控制技术的高超声速翼型前段气动特性的数值模拟研究[C]//第十届全国流体力学学术会议, 2018.  
ZHU J C, SHI Z W, SUN Q J, et al. Numerical simulation research on the aerodynamic characteristics of hypersonic airfoil front section based on active flow control technology[C]//The 10th National Conference on Fluid Mechanics, 2018 (in Chinese).
- [10] OGINO K, MAMORI H, FUKUSHIMA N, et al. Direct numerical simulation of Taylor-Couette turbulent flow controlled by a traveling wave-like blowing and suction[J]. International Journal of Heat and Fluid Flow, 2019, 80: 108463.
- [11] 张志勇, 王团团, 陈志华, 等. 低雷诺数下吹吸气射流对 NACA0012 翼型气动性能的影响[J]. 空气动力学学报, 2020, 38(1): 58-65.  
ZHANG Z Y, WANG T T, CHEN Z H, et al. The effect of blowing/suction jet on the aerodynamic performance of airfoil NACA0012 at low Reynolds number[J]. Acta Aerodynamica Sinica, 2020, 38(1): 58-65 (in Chinese).
- [12] 贲宝佳. 一种新型吹吸气相结合的方法控制流动分离[J]. 科技创新与应用, 2017(2): 56.  
BEN B J. A new method of combined blowing and suction to control flow separation[J]. Technology Innovation and Application, 2017(2): 56 (in Chinese).
- [13] HUANG L, HUANG P G, LEBEAU R P, et al. Numerical study of blowing and suction control mechanism on NACA0012 airfoil[J]. Journal of Aircraft, 2004, 41(5): 1005-1013.
- [14] GENÇ M, KAYNAK Ü. Control of flow separation and transition point over an aerofoil at low  $Re$  number using simultaneous blowing and suction: AIAA 2009-3672[R]. Reston: AIAA, 2009.
- [15] AGARWAL G, REDINIOTIS O, TRAUB L. An experimental investigation on the effects of pulsed air blowing separation control on NACA0015: AIAA 2008-737[R]. Reston: AIAA, 2008.
- [16] WAHIDI R, BRIDGES D. Effects of distributed suction on an airfoil at low Reynolds number[C]//40th Fluid Dynamics Conference and Exhibit. Reston: AIAA, 2010.

## Influence of suction flow control on aerodynamic characteristics of blended-wing-body aircraft

JIA Yuan, CAO Xiang, WU Jianghao\*

(School of Transportation Science and Engineering, Beihang University, Beijing 100083, China)

**Abstract:** This paper studies a blended-wing-body aircraft with distributed propulsion, explores the influence of suction control (with suction position and suction momentum variables) on the aerodynamic characteristics of aircraft in take-off and cruise state and explains the mechanism of the influence of suction flow control on the aerodynamic characteristics of the blended-wing-body aircraft. The results show that under the condition of high take-off attack angle, compared with the non-aspirated state, the maximum lift coefficient of the aircraft is increased by 7.16% when the aspirator is located in the outer wing segment (chord position being  $0.05c$ , and inspiratory momentum being  $0.02$ ). In the cruise state, when the aspirator is located in the centrosome (chord position being  $0.6c$ , and inspiratory momentum being  $0.0125$ ), the pressure distribution of the powertrain is improved, and the lift-drag ratio of the aircraft is increased by 2.14% compared with that of the non-getter state.

**Keywords:** blended-wing-body; flow control; suction position; suction volume; aerodynamic performance

Received: 2020-12-26; Accepted: 2021-01-30; Published online: 2021-02-22 16:25

URL: [kns.cnki.net/kcms/detail/11.2625.V.20210222.1603.002.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210222.1603.002.html)

\* Corresponding author. E-mail: buaawjh@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2020.0721

# 基于 KF-LESO-PID 洛伦兹惯性稳定平台控制

熊颖<sup>1</sup>, 刘强<sup>1,\*</sup>, 任元<sup>2</sup>, 樊亚洪<sup>3</sup>, 孙津济<sup>4</sup>

(1. 北京石油化工学院 精密电磁装备与先进测量技术研究所, 北京 102617; 2. 航天工程大学 宇航科学与技术系, 北京 101416;  
3. 北京控制工程研究所 空间智能控制技术重点实验室, 北京 100190; 4. 北京航空航天大学 惯性技术重点实验室, 北京 100083)

**摘 要:** 为克服现有惯性稳定平台使用机械轴承干扰量大, 使用气/液浮轴承难度高, 使用磁阻力磁轴承线性度差的缺点, 提出一种基于洛伦兹力偏转磁轴承的新型洛伦兹惯性稳定平台(LISP)。为克服耦合效应和承载摩擦谐振干扰对平台偏转通道高频姿态补偿控制的影响, 提出一种基于 LESO-PID 结合卡尔曼滤波(KF)反馈的数字控制方案。根据洛伦兹力磁轴承(LFMB)支承偏转系统结构特点, 建立了 LISP 转子偏转动力学模型; 利用模型分析径向两自由度偏转特性, 提出在 PID 控制器的基础上, 引入线性扩张状态观测器(LESO)和卡尔曼滤波反馈以抑制摩擦谐振干扰及耦合效应; 搭建了以 DSP 和 FPGA 为核心的数字控制系统, 并以离散形式将控制方法进行数字化实现。采用对数频率特性判据和 Nichols 曲线对所提控制方法的稳定性进行分析, 通过仿真比较引入 LESO-KF 前后转子偏转通道的稳定性。实验结果表明: PID 控制条件下在高频时失真, 引入 LESO-KF 后明显降低噪声及干扰, 同时还可对系统内部状态参数进行实时观测。实验结果验证了所提控制方法对摩擦谐振干扰及耦合效应的抑制作用。

**关键词:** 洛伦兹惯性稳定平台(LISP); 洛伦兹力磁轴承(LFMB); LESO-PID 控制; 卡尔曼滤波(KF)反馈; 稳定控制

**中图分类号:** V448.2; TP273

**文献标志码:** A **文章编号:** 1001-5965(2022)06-1072-10

随着中国“三航”国防尖端科技的高速发展, 对轻量化、高精度及快速响应的惯性设备的需求也越来越迫切<sup>[1-2]</sup>。惯性平台系统为导航系统和姿态稳定系统提供测量基准, 而其中武器载荷稳定平台稳定控制性能事关战场主动权, 因此, 研制具备轻量化、高频姿态补偿和大角度快速响应的平台系统是发展现代化装备的必然需求。

目前, 在国内外已应用的高精度、多功能的动载体稳定平台中, 几乎都是使用多轴陀螺稳定平台和磁悬浮框架平台作为输出力矩执行机构, 影响平台系统精度和响应速度的主要因素是偏转结构。与传统机械多轴稳定平台相比, 采用洛伦兹

力偏转磁轴承的稳定平台具有无接触、无摩擦、线性度高等优点<sup>[3-4]</sup>, 是轻量载荷姿态稳定系统的理想执行机构。

洛伦兹惯性稳定平台(Lorentz inertial stability platform, LISP)转子系统结构采用双自由度主动控制, 与现有的磁悬浮惯性稳定平台(MSISP)相比具有特殊性: 一方面, 其承重及滑动结构采用球面滚珠轴承进行支承与滑动, 消除了悬浮轴承的耦合与额外功耗, 增加了高频的可能性; 另一方面, 相比于 MSISP 通过框架间接输出力矩<sup>[5]</sup>, LISP 通过洛伦兹力磁轴承(Lorentz force magnetic bearing, LFMB)对转子系统施加直驱力矩, 其动力学特性发生了变

收稿日期: 2020-12-30; 录用日期: 2021-04-02; 网络出版时间: 2021-04-23 16:06

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20210423.1118.001.html](http://kns.cnki.net/kcms/detail/11.2625.V.20210423.1118.001.html)

基金项目: 北京市自然科学基金(3212004); 北京市青年拔尖人才培养计划(2017000026833ZK22); “十三五”时期北京市属高校高水平教师队伍建设支持计划(CIT&TCD201804034)

\* 通信作者. E-mail: liuqiangbuaa@163.com

**引用格式:** 熊颖, 刘强, 任元, 等. 基于 KF-LESO-PID 洛伦兹惯性稳定平台控制[J]. 北京航空航天大学学报, 2022, 48(6): 1072-1081. XIONG Y, LIU Q, REN Y, et al. Lorentz inertial stability platform control based on KF-LESO-PID[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(6): 1072-1081 (in Chinese).

化。此外,为获得较大的偏转范围,采用球面转子结构,而球面在力矩输出上平衡了几何力学耦合<sup>[6]</sup>,但存在较多的其他耦合因素。

为抑制稳定平台的耦合谐振效应及摩擦等非线性的干扰因素,实现平台的稳定控制,国内外学者近年来提出了多种控制方法。文献[7]提出了一种由电流环、速度环、稳定环构成的三闭环控制方法,其中双速度环提高了系统对低频扰动的抑制能力,但由于内外环反馈信号不一致且采用 PI 调节,易降低系统信噪比,严重时影响稳定性。为解决普通闭环 PID 误差大和参数整定的问题,文献[8]针对传统 PID 控制抗干扰性能不高的问题提出了基于 BP 算法的改进模糊神经网络 PID 复合控制方案。文献[9]提出了基于模糊策略的 PID 控制器,通过适时修改参数从而提高稳定精度和水准。尽管以上复合 PID 控制器能提高系统响应与抗干扰性能,但文献[8-9]的复杂控制方法均以 PID 参数调节为中心,控制对象均为机械支承的框架机构,相比于 LISP 而言其复杂度过高,同时上述方法缺乏基本预知调节与降频减噪的能力。针对经典 PID 控制存在的参数调整难、自适应调节能力差的问题,国内外学者针对平台提出多种先进控制方法。文献[10]介绍了一种基于神经网络的惯性平台控制方法。文献[11]提出了基于迭代学习算法的伺服电机控制器。针对三轴惯性平台,文献[12]提出了一种变结构的控制方法。文献[13]针对陀螺稳定系统中非线性和控制对象不确定性的扰动因素,设计了分数阶 PI (FOPI)控制方法。文献[14]提出了一种滑模变结构控制惯性平台的方法。文献[15]设计了基于模型的改进型线性二次高斯最优控制器。以上

先进控制方法在提高系统响应的同时增加了运算量与故障率,偏离了 LISP 的设计期望,且控制对象同样为框架或电机结构。而本文研究对象为采用 LFMB 支承的 LISP 转子系统,关于 LFMB 的文献仅涉及利用 LFMB 进行内力矩输出与偏转角度控制<sup>[16-17]</sup>。特殊的结构给 LISP 这种新概念稳定平台转子偏转带来了控制挑战。因此,对 LFMB 偏转的 LISP 进行稳定偏转控制具有重要的工程应用价值。

为有效提高 LISP 可靠性、实用性及快速机动能力,本文提出了针对 LISP 的基于卡尔曼滤波 (KF)反馈和线性扩张状态观测器 (LESO)的 PID 控制。在建立 LISP 线性化动力学方程的基础上,设计 LESO 并推导出其稳定域,利用 LESO 进行动力学模型状态观测与误差补偿,引入卡尔曼滤波进行降噪滤波的同时增强自适应调节能力,搭建了以 DSP 和 FPGA 为核心的数字控制系统,采用对数频率特性判据和 Nichols 曲线分析全系统稳定性,并通过扰动实验对该控制方法进行了验证。

### 1 LISP 结构与工作原理

基于机械式平台承载力大的优点与磁悬浮式平台精度高的优点,提出一种混合式稳定平台 LISP,平台采用高研磨度润滑球面轴承进行支承,以及线性度高的 LFMB 进行偏转,实现横纵减摇的同时有效进行精确指向。LFMB 的偏转技术解决了稳定平台主动控制偏转产生的摩擦和磨损问题,可大幅提高稳定平台偏转速度和回转频率。LISP 结构如图 1 所示。平台系统主要由 LFMB (LFMB 径向二自由度偏转)、万向球面滚珠轴承

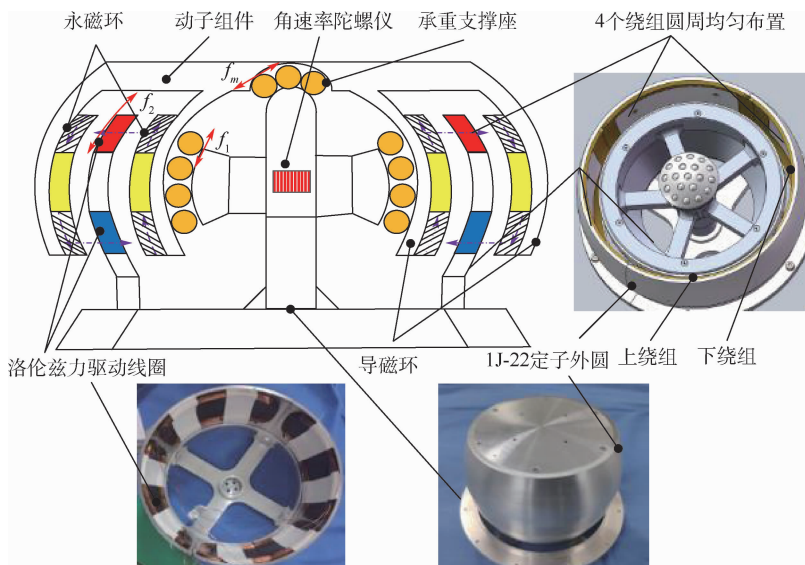


图 1 LISP 结构示意图

Fig. 1 Diagram of LISP structure

(承重及辅助偏转)、球形动子(执行机构)和内房结构(角速率陀螺仪及线路机械接口)组成。

当平台受外界干扰振动时,角速率陀螺仪测量动子偏转角度和速率信号,并将其反馈至控制器。通过控制器调整线圈电流,LFMB产生与干扰力矩相反的电磁力矩,进而将动子回复至预设平衡位置,实现载荷和平台稳定。

## 2 LISP 数学建模

### 2.1 LFMB 磁路分析

不考虑磁路漏磁及线圈产生磁场的影响,由图1中LISP结构可得LFMB等效磁路,如图2所示。图中: $F_{p1} \sim F_{p4}$ 分别为内外上下4个径向充磁磁钢磁动势; $R_{p1} \sim R_{p4}$ 分别为对应的磁钢磁阻; $R_{iu}$ 和 $R_{id}$ 分别为上下经过绕组气隙磁阻; $R_{ui}$ 和 $R_{uo}$ 分别为内外导磁环磁阻; $\Phi_u$ 和 $\Phi_d$ 分别为上下气隙磁通量。

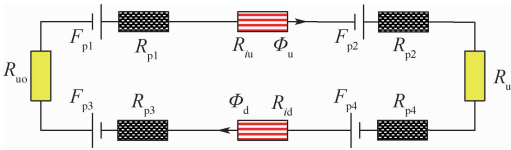


图2 LFMB 等效磁路

Fig.2 LFMB equivalent magnetic circuit

气隙磁通量  $\Phi$  与磁感应强度  $B$  的计算如下:

$$\begin{cases} \Phi = \frac{\sum_{i=1}^4 F_{pi}}{\sum_{i=1}^4 R_{pi} + \sum_{i=1}^4 R_{iu} + \sum_{i=1}^4 R_{id}} \\ \mathbf{B} = [B_u, B_d]^T = \text{diag}(A_u, A_d)^{-1} \cdot \Phi \end{cases} \quad (1)$$

式中: $F_{pi} = H_c l_{pi}$ ,  $H_c$  为磁钢矫顽力,  $l_{pi}$  为磁钢的磁化长度; $\Phi$  为穿过上下气隙磁通量; $B$  为磁感应强度,  $B_u$  和  $B_d$  分别为上下气隙磁感应强度;  $A_u$  和  $A_d$  分别为对应的磁钢等效截面积。

### 2.2 电磁力矩偏转模型

LFMB产生的电磁力遵循安培力定律,对其采用微元法。定义 $\alpha, \beta$ 分别为绕 $x, y$ 轴偏转角,每组 $N$ 匝定子绕组对应的周向圆心角为 $\varphi$ ,在线圈周向上取微元长度 $dl$ ,其对应的微元角为 $d\varphi$ ,球形半径为 $L_r$ ,LFMB微元长度线圈所受洛伦兹力为

$$df_i = 2NI_r L_r B d\varphi \quad (2)$$

式中: $B$ 为 $B$ 的向量范数,  $B = \|\mathbf{B}\|$ 为气隙磁感应强度大小;  $I_i (i = 1, 2, 3, 4)$ 为电流大小,第 $i$ 段线圈产生洛伦兹力 $df_i$ 。

单组线圈上下部分产生的有效电磁合力矩为

$$T_i = \int_0^\phi 2NI_r L_r^2 B d\varphi \quad i = 1, 2, 3, 4 \quad (3)$$

式中: $\phi$ 为单组微元偏转线圈圆心角积分上限。

沿圆心对称组成 $x, y$ 力偶矩,绕 $x$ 轴和 $y$ 轴偏转的电磁力偶矩为

$$\begin{bmatrix} M_\alpha \\ M_\beta \end{bmatrix} = \begin{bmatrix} 2NB(I_1 \pm I_3)L_r^2\phi \\ 2NB(I_2 \pm I_4)L_r^2\phi \end{bmatrix} \quad (4)$$

假设永磁环长度足够长,径向充磁磁密均匀,考虑承重轴承的差动滑动摩擦及阻尼系数,有

$$\begin{bmatrix} J_x \\ J_y \end{bmatrix} \cdot [\alpha'' \quad \beta''] = \begin{bmatrix} M_\alpha \\ M_\beta \end{bmatrix} - M_f \quad (5)$$

式中: $J$ 为转动惯量。

考虑到轴向承重摩擦,根据滚动轴承摩擦力矩一般计算公式:

$$M_f = \mu F d / 2 \quad (6)$$

式中: $M_f$ 为摩擦力矩,  $N \cdot \text{mm}$ ;  $\mu$ 为轴承的摩擦系数;  $d$ 为轴承的内径,  $\text{mm}$ ;  $F$ 为轴承所承受的载荷,  $N$ 。

联立式(1)、式(4)、式(5)和式(6),有

$$\begin{bmatrix} J_x \\ J_y \end{bmatrix} \cdot [\alpha'' \quad \beta''] = \begin{bmatrix} 2N \text{diag}(A_u, A_d)^{-1} \Phi (I_1 \pm I_3) L_r^2 \phi \\ 2N \text{diag}(A_u, A_d)^{-1} \Phi (I_2 \pm I_4) L_r^2 \phi \end{bmatrix} - \mu F d / 2 \quad (7)$$

式(7)为LISP电磁力矩偏转模型。

## 3 LISP 偏转控制系统

### 3.1 LESO 设计

为方便设计LESO及分析稳定性,简化式(7)对象,设定统一形式为

$$J\ddot{\theta} = u(t) + d(t) + f(t) \quad (8)$$

式中: $J$ 为转动惯量;  $u(t)$ 为控制输入;  $\theta$ 为偏转角度;  $d(t)$ 为外加干扰;  $f(t)$ 为系统内未知动态。

式(8)可写为

$$\ddot{\theta} = bu(t) + w(t) + f \quad (9)$$

式中: $b = 1/J$ 为控制增益;  $f$ 的导数存在且有界。

令 $x_3(t) = f + w(t) + bu(t)$ ,进一步有

$$\begin{cases} \dot{x}_1(t) = x_2(t) \\ \dot{x}_2(t) = x_3(t) \\ \dot{x}_3(t) = m(t) \\ y(t) = x_1(t) \end{cases} \quad (10)$$

式中: $\mathbf{x} = [x_1, x_2] = [\theta, \theta']$ ;  $m(t) = f' + w'(t) + bu'(t)$ 。

将式(10)改写为

$$\begin{cases} \dot{x} = Ux + O(bu + w(t) + f) \\ y = Wx \end{cases} \quad (11)$$

式中:  $x = [x_1, x_2] = [\theta, \theta']$ ;  $U = [0, 1; 0, 0]$ ;  $O = [0, 1]$ ;  $W = [1, 0]$ 。

由系统可观矩阵  $V = [W; WU] = [1, 0; 0, 1]$ ,  $\text{rank } V = E$  知系统可观, 可对系统配置观测器。参考文献[18], 针对 LISP 将 LESO 设计为

$$\begin{cases} \dot{\hat{x}}_1(t) = \hat{x}_2(t) + \frac{\alpha_1}{\varepsilon}(y - \hat{x}_1(t)) \\ \dot{\hat{x}}_2(t) = bu(t) + \hat{x}_3(t) + \frac{\alpha_2}{\varepsilon^2}(y - \hat{x}_1) \\ \dot{\hat{x}}_3(t) = \frac{\alpha_3}{\varepsilon^3}(y - \hat{x}_1) \end{cases} \quad (12)$$

式中:  $\varepsilon$  为 LESO 增益; 参数  $\alpha_1, \alpha_2, \alpha_3$  为增益系数。

式(12)与式(11)作差有

$$\begin{cases} \dot{e}_1(t) = e_2(t) - \frac{\alpha_1}{\varepsilon}e_1(t) \\ \dot{e}_2(t) = e_3(t) - \frac{\alpha_2}{\varepsilon^2}e_1(t) \\ \dot{e}_3(t) = m(t) - \frac{\alpha_3}{\varepsilon^3}e_1(t) \end{cases} \quad (13)$$

式中:  $e_i = x_i - \hat{x}_i, i = 1, 2, 3$ 。

对式(13)作拉普拉斯变换有

$$\begin{cases} G_1(s) = \frac{1}{s^3 + s^2\alpha_1/\varepsilon + s\alpha_2/\varepsilon^2 + \alpha_3/\varepsilon^3} \\ G_2(s) = \frac{s + \alpha_1/\varepsilon}{s^3 + s^2\alpha_1/\varepsilon + s\alpha_2/\varepsilon^2 + \alpha_3/\varepsilon^3} \\ G_3(s) = \frac{s^2 + s\alpha_1/\varepsilon + \alpha_2/\varepsilon^2}{s^3 + s^2\alpha_1/\varepsilon + s\alpha_2/\varepsilon^2 + \alpha_3/\varepsilon^3} \end{cases} \quad (14)$$

式中:  $G_i(s) = E_i(s)/H(s), i = 1, 2, 3$ 。

### 3.2 LESO 稳定性判定

为了保证状态观测误差在有界范围内, 须使得式(14)中系统特征方程的特征根(闭环传递函数的极点)在复平面左半平面, 即  $D(s) = s^3 + s^2\alpha_1/\varepsilon + s\alpha_2/\varepsilon^2 + \alpha_3/\varepsilon^3 = 0$ , 列写 Routh 表:

$$\begin{array}{ccc|c} s^3 & 1 & \alpha_2/\varepsilon^2 & \\ s^2 & \alpha_1/\varepsilon & \alpha_3/\varepsilon^3 & \\ s & (\alpha_1\alpha_2 - \alpha_3/\varepsilon^3)/(\alpha_1/\varepsilon) & 0 & \\ s^0 & \alpha_3/\varepsilon^3 & 0 & \end{array} \quad (15)$$

由 Routh 判据充分必要条件得 LESO 稳定性条件为

$$\begin{cases} \alpha_1 > 0 \\ \alpha_3 > 0 \\ \alpha_1\alpha_2 - \alpha_3 > 0 \end{cases} \quad (16)$$

通过频域计算解出各状态观测误差对于干扰上

界, 有

$$\begin{cases} |G_1(j\omega)| = \frac{\varepsilon^3}{\sqrt{(\alpha_3 - \alpha_1\varepsilon^2\omega^2)^2 + (\alpha_2\varepsilon\omega - \varepsilon^3\omega^3)^2}} \\ |G_2(j\omega)| = \frac{\varepsilon^2\sqrt{\alpha_1^2 + \varepsilon^2\omega^2}}{\sqrt{(\alpha_3 - \alpha_1\varepsilon^2\omega^2)^2 + (\alpha_2\varepsilon\omega - \varepsilon^3\omega^3)^2}} \\ |G_3(j\omega)| = \frac{\varepsilon\sqrt{(\alpha_2 - \varepsilon^2\omega^2)^2 + \alpha_1^2\varepsilon^2\omega^2}}{\sqrt{(\alpha_3 - \alpha_1\varepsilon^2\omega^2)^2 + (\alpha_2\varepsilon\omega - \varepsilon^3\omega^3)^2}} \end{cases} \quad (17)$$

对式(17)取极限  $\varepsilon \rightarrow 0$ , 有

$$\begin{cases} \lim_{\varepsilon \rightarrow 0} |G_1(j\omega)/\varepsilon^3| = 1/\alpha_3 \rightarrow \lim_{\varepsilon \rightarrow 0} |G_1(j\omega)| = o(\varepsilon^3) \\ \lim_{\varepsilon \rightarrow 0} |G_2(j\omega)/\varepsilon^2| = \alpha_1/\alpha_3 \rightarrow \lim_{\varepsilon \rightarrow 0} |G_2(j\omega)| = o(\varepsilon^2) \\ \lim_{\varepsilon \rightarrow 0} |G_3(j\omega)/\varepsilon| = \alpha_2/\alpha_3 \rightarrow \lim_{\varepsilon \rightarrow 0} |G_3(j\omega)| = o(\varepsilon) \end{cases} \quad (18)$$

对于平台系统, 忽略时变  $h(t)$  频率  $\omega$  的影响, 当  $\varepsilon$  足够小,  $\varepsilon\omega \rightarrow 0$ , 式(18)可写为

$$\begin{cases} \lim_{\varepsilon\omega \rightarrow 0} |G_1(j\omega)| = \varepsilon^3/\alpha_3 \\ \lim_{\varepsilon\omega \rightarrow 0} |G_2(j\omega)| = \alpha_1\varepsilon^2/\alpha_3 \\ \lim_{\varepsilon\omega \rightarrow 0} |G_3(j\omega)| = \alpha_2\varepsilon/\alpha_3 \end{cases} \quad (19)$$

同时为防止峰值产生, 达到观测器最佳收敛效果, 设计  $\varepsilon$  为<sup>[19-20]</sup>

$$\frac{1}{\varepsilon} = R = \begin{cases} \kappa \left( 1 - \frac{1 - e^{-\lambda_1 t}}{1 + e^{-\lambda_2 t}} \right) & 0 \leq t \leq t_{\max} \\ \kappa & t > t_{\max} \end{cases} \quad (20)$$

式中:  $\kappa, \lambda_1$  和  $\lambda_2$  为正实数。

### 3.3 LISP-KF 复合 LESO-PID 设计

卡尔曼滤波是改善控制器相位滞后的有效方式。为此, 采用基于 LESO-PID 卡尔曼滤波(KF-LESO-PID)反馈控制律。其中, PID 控制器为基本的控制器, 通过 LESO 将被控对象内部状态提取, 并以一定可调数据量参与 PID 控制, 卡尔曼滤波通过调节增益可过滤噪声和扰动因素。如图 3 所示, 输入给定角度  $\theta$ , 经过 LESO-PID 及测量反馈输出响应,  $N$  为随机干扰量。

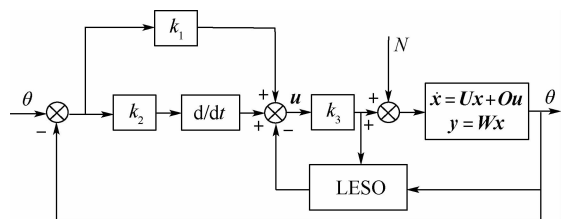


图 3 LESO-PID 控制框图

Fig. 3 Block diagram of LESO-PID control

对于离散域线性系统:

$$\begin{cases} \mathbf{x}(k) = \mathbf{U}\mathbf{x}(k-1) + \mathbf{O}(\mathbf{u}(k) + \mathbf{w}(k)) \\ \mathbf{y}_v(k) = \mathbf{W}\mathbf{x}(k) + \mathbf{v}(k) \end{cases} \quad (21)$$

式中:  $\mathbf{w}(k)$  为过程噪声信号;  $\mathbf{v}(k)$  为测量噪声信号。

为模拟数字滤波处理, 采用离散卡尔曼推算:

$$\mathbf{M}_n(k) = \frac{\mathbf{P}(k)^- \mathbf{C}^T}{\mathbf{C}\mathbf{P}(k-1)^- \mathbf{C}^T + \mathbf{R}_k} \quad (22)$$

式中:

$$\begin{cases} \mathbf{P}(k)^- = \mathbf{U}\mathbf{P}(k-1)\mathbf{U}^T + \mathbf{O}\mathbf{Q}_k\mathbf{O}^T \\ \mathbf{P}(k) = (\mathbf{E}_n - \mathbf{M}_n(k)\mathbf{C})\mathbf{P}(k)^- \end{cases} \quad (23)$$

离散卡尔曼状态方程为

$$\begin{cases} \mathbf{x}(k) = \mathbf{U}\mathbf{x}(k-1) + \mathbf{M}_n(k)(\mathbf{y}_v(k) - \mathbf{W}\mathbf{U}\mathbf{x}(k-1)) \\ \mathbf{y}_e(k) = \mathbf{W}\mathbf{x}(k) \end{cases} \quad (24)$$

误差协方差为:  $\mathbf{E}\text{cov}(k) = \mathbf{W}\mathbf{P}(k)\mathbf{W}^T$ 。

卡尔曼离散滤波结构如图4所示。

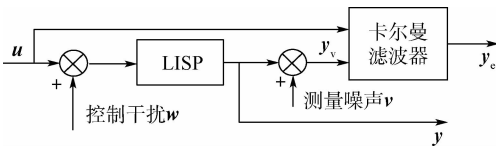


图4 LISP-KF 结构  
Fig. 4 LISP-KF structure

### 3.4 LISP 偏转控制系统模型及稳定性

LFMB 对平台动子径向偏转控制过程如图5所示, 控制通道主要包括传感器、卡尔曼低通滤波器、LESO-PID 控制器和功率放大器。图5的闭合回路中, 卡尔曼低通滤波器将传感器采集到的转子偏转角信息  $\alpha, \beta$  进行降噪处理后传递至控制器, LESO-PID 控制器根据偏转指令信息  $\alpha_r, \beta_r$  解算微信号送功率放大器, 功率放大器将微信号转化为控制电流, 驱动 LFMB 对转子产生偏转控制力矩。

根据图5中的偏转控制系统框图可知, 偏转通道的控制电流表达式为

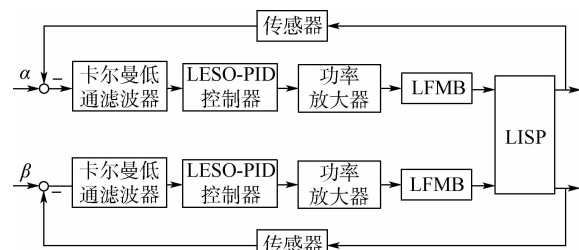


图5 LISP 偏转控制系统框图

Fig. 5 Block diagram of LISP deflection control system

$$\begin{cases} i_\alpha(t) = -k_s G_a G_c G_k \alpha(t) \\ i_\beta(t) = -k_s G_a G_c G_k \beta(t) \end{cases} \quad (25)$$

式中:  $k_s$  为传感器比例增益;  $G_a, G_c, G_k$  分别为卡尔曼低通滤波器、LESO-PID 控制器、功率放大器的输入-输出传递函数拉普拉斯变换算子。

对于 LISP 偏转转子, 结合式(7)和式(25), 得到 LISP 转子偏转动力学方程组为

$$\begin{cases} J_x \ddot{\alpha} = 4NB(-k_s G_a G_c G_k \alpha(t)) L_r^2 \phi - \mu Fd/2 \\ J_y \ddot{\beta} = 4NB(-k_s G_a G_c G_k \beta(t)) L_r^2 \phi - \mu Fd/2 \end{cases} \quad (26)$$

对式(26)进行拉普拉斯变换可得

$$\begin{cases} J_x (s^2 \alpha(s) - s\alpha(0) - \dot{\alpha}(0)) = 4NB(-k_s G_a G_c G_k \alpha(s)) L_r^2 \phi - \mu Fd/2 \\ J_y (s^2 \beta(s) - s\beta(0) - \dot{\beta}(0)) = 4NB(-k_s G_a G_c G_k \alpha(s)) L_r^2 \phi - \mu Fd/2 \end{cases} \quad (27)$$

令  $\theta = \alpha + j\beta$ , 将式(27)第2个方程乘以  $j$  与第1个方程相加可得

$$\begin{aligned} J_r (s^2 \theta(s) - s\theta(0) - \dot{\theta}(0)) = \\ 4NB(-k_s G_a G_c G_k) L_r^2 \phi \theta(s) - (1 + j)\mu Fd/2 \end{aligned} \quad (28)$$

进一步得到

$$\theta(s) = \frac{-(1 + j)\mu Fd/2}{J_r s^2 + 4NBk_s G_a G_c G_k L_r^2 \phi} \quad (29)$$

转子闭环系统特征方程为

$$J_r s^2 + 4NBk_s G_a G_c G_k L_r^2 \phi = 0 \quad (30)$$

等效被控对象和等效控制通道传递函数分别为

$$\begin{cases} G_{oe}(s) = 1/(J_r s^2) \\ G_{ce}(s) = 4NBk_s G_a G_c G_k L_r^2 \phi \end{cases} \quad (31)$$

系统等效开环传递函数表达式为

$$G_k = G_{oe}(s)G_{ce}(s) = 4NBk_s G_a G_c G_k L_r^2 \phi / (J_r s^2) \quad (32)$$

采用 Bode 图对系统稳定性进行分析, 根据经典控制理论, 传递函数系数为实数, 正负频率特性关于零频率对称, 只需绘制正频率特性。使用的动子系统主要参数如表1所示, 控制通道主要参数根据文献[21]中的方法进行设置。

绘制的开环传递函数 Bode 图如图6所示, 开环增益 Nichols 曲线如图7所示。根据 Nyquist 稳定判据<sup>[22]</sup>及与 Bode 图的关联性, 由图6知 Bode 相频穿越  $(2k + 1 + v/2)$  的次数为零, 由式(32)知开环极点无正实根, 因此  $N = 0, Q = 0$ , 推得  $Z = Q - N = 0$ , 故闭环系统稳定。

表 1 系统实际参数与调试参数

Table 1 System actual parameters and debugging parameters

系统调试参数	数值	物理尺寸参数	数值
功放增益 $k_g/(V \cdot A^{-1})$	0.31	动子质量 $m/kg$	2.1
传感器比例增益 $k_s/(V \cdot m^{-1})$	9 800	线圈有效长度 $L/mm$	71.2
优化比例系数 $k_p$	17.13	转动惯量 $J_x/(kg \cdot m^2)$	0.005 76
优化积分系数 $k_i$	0.55	线圈匝数 $N$	150
优化微分系数 $k_d$	0.47	线圈电阻 $R/\Omega$	8.3
功放截止频率 $w_g/Hz$	320	磁感应强度大小 $B/T$	0.3
滤波截止频率 $w_f/Hz$	350	$X/Y$ 额定工作范围/ $(^\circ)$	$\pm 20$

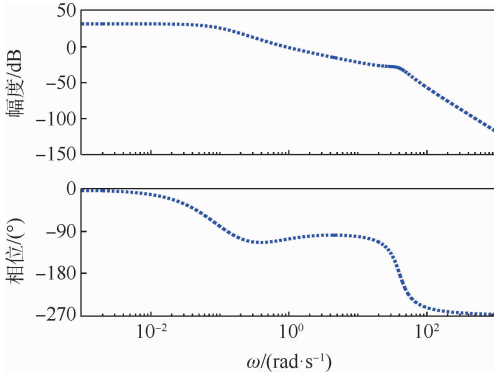


图 6 等效开环传递函数 Bode 图

Fig. 6 Bode diagram of equivalent open-loop transfer function

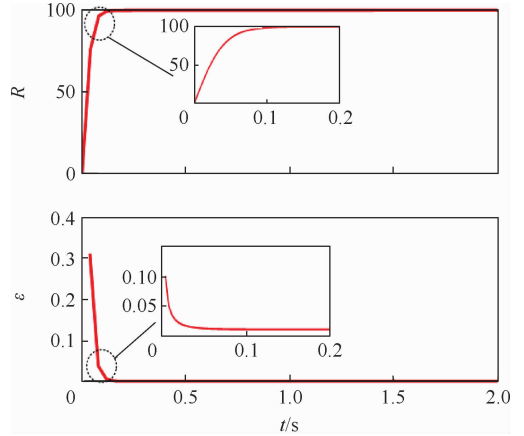


图 8 模拟参数  $R$  与扰动参数  $\varepsilon$  取值曲线

Fig. 8 Simulation parameter  $R$  and perturbation parameter  $\varepsilon$  value curves

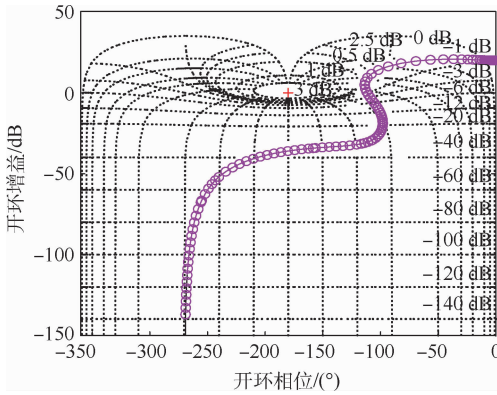


图 7 LISP 系统 Nichols 曲线

Fig. 7 Nichols curves of LISP system

### 4 LISP 仿真与实验校验

按式 (20) 设计  $\varepsilon$ , 令  $\lambda_1 = \lambda_2 = 50, \mu = 100, R$  及  $\varepsilon$  的变化如图 8 所示。

为体现 LESO 在理想被控对象上的观测效果,对式 (8) 系统模型进行扩张观测仿真。在 Simulink 下建立仿真程序,根据式 (12) 和式 (16) 规则编写 S 函数,并在固定初始化输入条件下进行仿真。

LISP 角度和角速度观测曲线分别如图 9 和图 10 所示,相应的系统未知动态曲线如图 11 所示。由图 9 可知,该 LESO 观测器能快速跟踪上理想偏转角度,0.5 s 后误差小于  $0.01^\circ$ ;图 10 中,角速度呈现正弦周期变化,所设计的观测器在 0.5 s 后无误差跟随角速度信号。

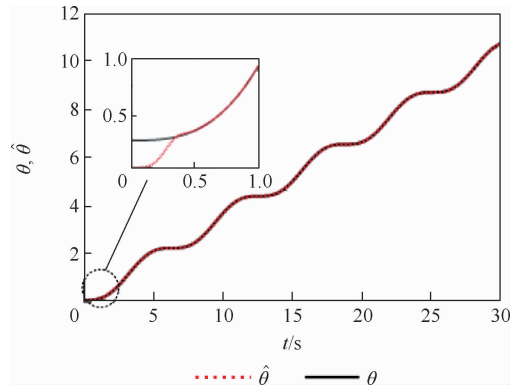


图 9 LFMB 角度  $\theta$  扩张观测

Fig. 9 LFMB angle  $\theta$  expansion observation

由图 11 可知,系统内部一般情况下动态误差  $\xi$  与不可计量因素很小,观测结果与结论相似。上述 3 类观测的状态观测效果均达到理想要求,且无明显失真和迟滞。因此,通过选择合适的增益与带宽,所设计的 LESO 观测器在针对平台系统理想模型上满足观测指标和需求。

图 12 为 LISP 实际控制时 LESO 观测的干扰信号与对应误差。可知,干扰信号的观测几乎能完全跟踪模拟实际干扰,说明通过选择合适的 LESO 增益系数能很好地对系统的干扰进行观测与补偿,进而实现带载的精确偏转。

为体现 LESO 观测器在实际控制中具有优良

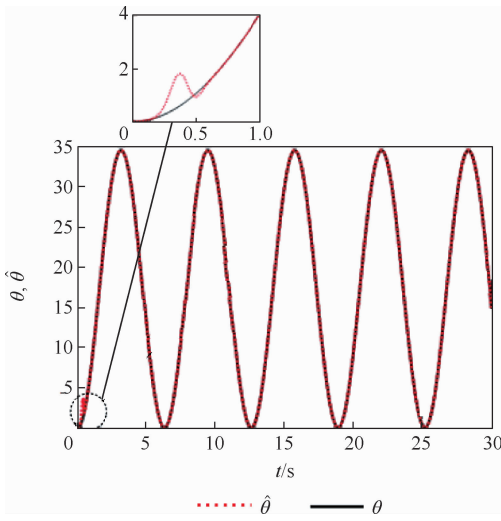


图 10 角速度  $\hat{\theta}$  扩张观测

Fig. 10 Angular velocity  $\hat{\theta}$  expansion observation

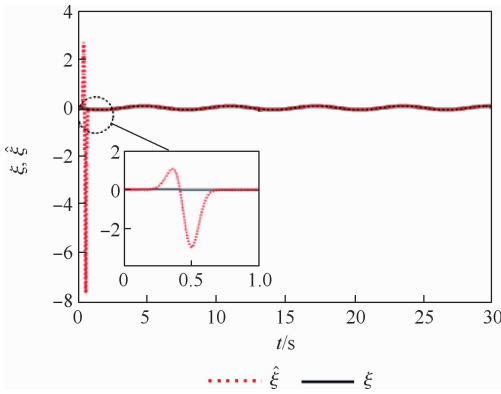


图 11 系统未知动态  $\xi$  扩张观测曲线

Fig. 11 Unknown dynamic  $\xi$  expansion observation curve in system

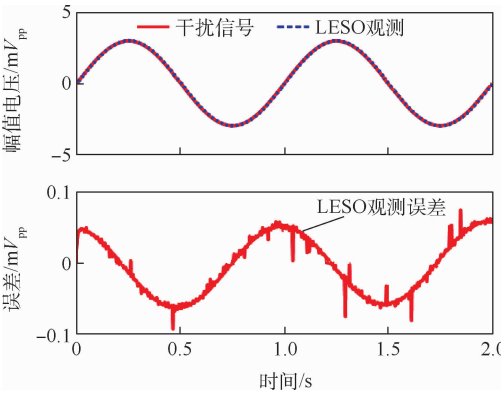


图 12 LESO 干扰信号观测与误差

Fig. 12 LESO interference signal observation and error

的高频动态补偿控制性能,加入 KF-PID 控制器并以连续系统仿真。仿真参数如表 1 所示。对比项数为是否加入 LESO 高频动态补偿控制,如图 13和图 14 所示。

对比图 13 与图 14 可以看出, KF-LESO-PID 控制对输入信号的跟踪优于 KF-PID 控制,其跟踪误差在  $10^{-3} V_{pp}$  范围内。对于同一个输入信

号,2 种控制方法的控制输出误差相差一个量级标准, KF-PID 控制器控制输出的变化率趋于零,基本误差维持在  $0.1 V_{pp}$  以内, KF-LESO-PID 控制则是将基本误差维持在  $0.01 V_{pp}$  以内。

通过以上针对式(7)的仿真及其对比可发现,本文提出的控制策略综合了 LESO 与卡尔曼滤波的优点,在轻量载荷的作用下,设计的控制器不仅能实现状态量和干扰量的观测,对载荷干扰进行前馈补偿,还能有效抑制平台谐振。此外,基于 LESO 的 PID 控制器结构简单,调节参数物理意义明确,更易于实现平台稳定控制。

为验证模型和控制策略的正确性,根据设计指标和仿真结果进行偏转测试实验,实验系统如图 15 所示。

图 16 为系统连续运行下工作噪声测量图。可知,连续系统噪声来源随机不确定,噪声电压峰值显示达到  $0.6 m V_{pp}$ ,均值稳定在  $0.25 m V_{pp}$  上下。由于平台采用的 MEMS 陀螺仪标度因素稳定性为  $0.05\%$ ,除带来一定的干扰外,瞬时噪声易造成系统暂时性失稳。

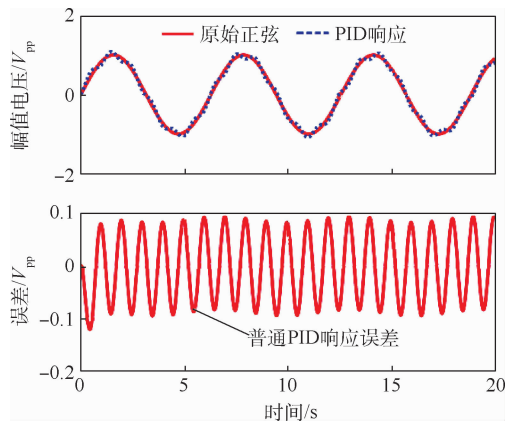


图 13 普通 KF-PID 响应曲线

Fig. 13 Response curves of common KF-PID

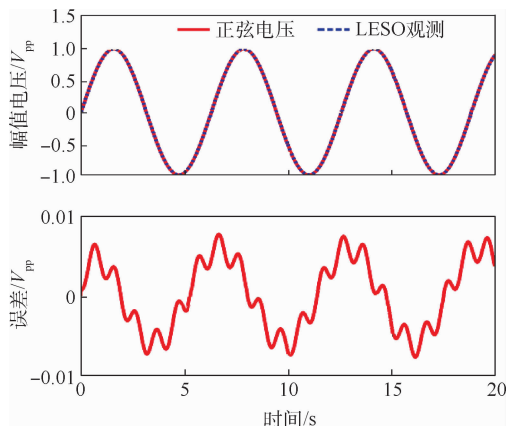


图 14 含 LESO KF-PID 响应曲线

Fig. 14 Response curves of KF-PID with LESO

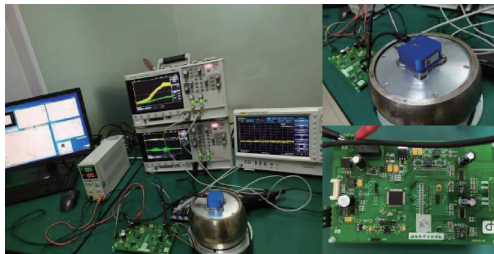


图 15 LISP 控制系统实验平台

Fig. 15 LISP control system experiment platform

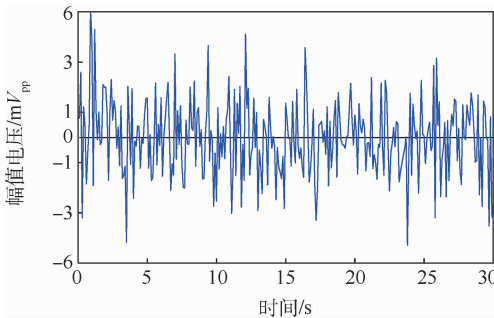


图 16 LISP 连续系统噪声测量

Fig. 16 LISP continuous system noise measurement

图 17 为 KF-PID 与单纯 PID 控制连续系统输出状态。为表现噪声在连续系统的影响,将标称值设定在 1.2 V。可知,低压状态下噪声较为明显,造成的输出漂移容易影响后级定位与指向。普通 PID 的输出峰值达到 2 V 以上,稳定状态的振荡频率在 1.8 Hz 左右,而加入离散卡尔曼滤波处理后峰值在 1.65 V 左右,稳定状态振荡频率在 0.4 Hz 左右。通过对比可知,卡尔曼滤波环节对前后级功放具有显著降低信噪的功能。

为进一步验证 LESO 的功能性及 KF-LESO-PID 的优势,对在 KF-PID 基础上加入 LESO 补偿控制后进行系统状态观测。图 18 为偏转矢量测试曲线,不考虑仪器测量等误差,在 2.3 s 以前双矢量跟踪有一定偏差,但 LESO 补偿功能与卡尔曼自适应趋势使得整体趋势逼近设定矢量运动状态。

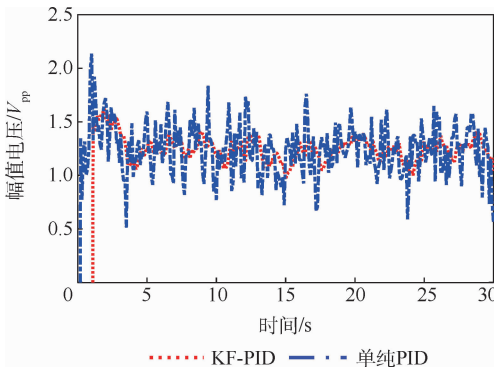


图 17 KF-PID 低压实际测量曲线

Fig. 17 KF-PID low pressure actual measurement curves

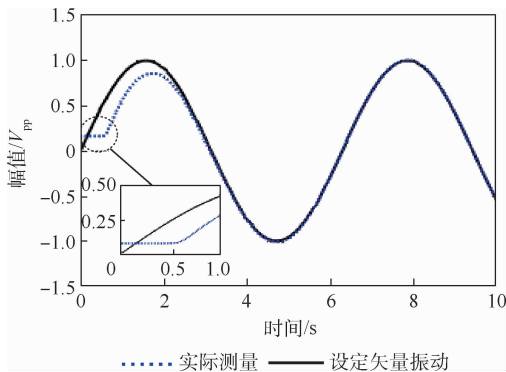


图 18 LISP 矢量示踪曲线

Fig. 18 LISP vector tracer curves

为进一步展现矢量运动趋势与 LESO 对矢量的跟踪效果,对可观测系统的多状态量进行扩张观测。图 19 为 LISP 系统位置与速度矢量观测与实际工况下对比曲线。可知,在低频正弦运动下,0.5 s 时刻实现双曲线拟合,而速度虽有一定抖动,但拟合时间不超过 1 s,拟合过程在单周期中占比不超过 20%,可以实现快速稳定运动。

通过仿真对比可以发现,根据特定 LISP 提出的控制策略结合了 LESO-PID 控制策略和卡尔曼自适应滤波的特点,在轻量载荷及环境噪声作用下,设计的控制器不仅能够实现内部状态量和干扰的观测,对载荷干扰及噪声进行补偿控制,还能有效抑制系统内部不确定因素对系统的不利影响,具有一定的自适应学习能力。此外,基于 LESO 的 PID 控制器,结构简单且调节参数物理意义明确,更易于实现 LISP 主动控制。

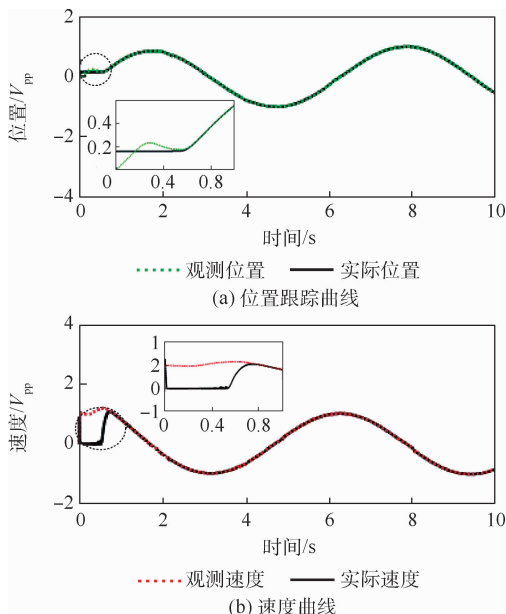


图 19 位置与速度矢量拟合曲线

Fig. 19 Position and velocity vector fitting curves

## 5 结 论

1) 针对耦合效应和承载摩擦谐振干扰对新型洛伦兹惯性稳定平台偏转通道高频姿态补偿控制影响,提出基于 LESO-PID 的主动控制策略,在此基础上结合卡尔曼滤波对全域进行自适应去噪和稳频。

2) 该复合控制策略使得稳频振荡频率降低了 3.5 倍,追踪拟合周期不超过 20%,实现了 LISP 主动高频控制,有效提高了响应速度和精度。

3) 综合分析,提出的复合 KF-LESO-PID 控制策略具有重要的工程应用价值。

### 参考文献 (References)

- [1] HILKERT J M. Inertially stabilized platform technology concepts and principles [J]. IEEE Control Systems Magazine, 2008, 28(1): 26-46.
- [2] MASTEN M K. Inertially stabilized platforms for optical imaging systems[J]. IEEE Control Systems Magazine, 2008, 28(1): 47-64.
- [3] ZHAO Y, LIU Q, MA L M, et al. Novel Lorentz force-type magnetic bearing with flux congregating rings for magnetically suspended gyrowheel[J]. IEEE Transactions on Magnetics, 2019, 55(12): 1-8.
- [4] XU G F, CAI Y W, REN Y, et al. Application of a new Lorentz force-type tilting control magnetic bearing in a magnetically suspended control sensitive gyroscope with cross-sliding mode control[J]. Transactions of the Japan Society for Aeronautical and Space Sciences, 2018, 61(1): 40-47.
- [5] 夏长峰, 蔡远文, 任元, 等. 磁悬浮控制敏感陀螺转子偏转通道稳定控制方法[J]. 控制理论与应用, 2020, 37(7): 1535-1543.
- XIA C F, CAI Y W, REN Y, et al. Stable control method for rotor tilt channel in magnetically suspended control and sensing gyro[J]. Control Theory & Applications, 2020, 37(7): 1535-1543 (in Chinese).
- [6] LIU Q, LI H, WANG W, et al. Analysis and experiment of 5-DOF decoupled spherical vernier-gimballing magnetically suspended flywheel (VGMSFW) [J]. IEEE Access, 2020, 8: 111707-111717.
- [7] 李志俊, 包启亮, 毛耀, 等. 惯性平台稳定回路多闭环串级控制[J]. 光电工程, 2010, 37(5): 19-24.
- LI Z J, BAO Q L, MAO Y, et al. Multi-closed loops cascade control for stabilization of inertia platform[J]. Opto-Electronic Engineering, 2010, 37(5): 19-24 (in Chinese).
- [8] ZHOU X Y, LI Y T, JIA Y, et al. An improved fuzzy neural network compound control scheme for inertially stabilized platform for aerial remote sensing applications[J]. International Journal of Aerospace Engineering, 2018, 2018: 7021038.
- [9] ZHANG Y S, YANG T, LI C Y, et al. Fuzzy-PID control for the position loop of aerial inertially stabilized platform [J]. Aerospace Science and Technology, 2014, 36: 21-26.
- [10] TSAI M S, LIN M T, YAU H T. Development of command-based iterative learning control algorithm with consideration of friction, disturbance, and noise effects [J]. IEEE Transactions on Control Systems Technology, 2006, 14(3): 511-518.
- [11] GE S S, LEE T H, ZHAO Q. Real-time neural network control of a free gyro stabilized mirror system [C] // Proceedings of the 1997 American Control Conference. Piscataway: IEEE Press, 1997: 1076-1080.
- [12] SHTESSEL Y B. Sliding mode stabilization of three axis inertial platform [C] // Proceedings of 26th Southeastern Symposium on System Theory. Piscataway: IEEE Press, 1994: 54-58.
- [13] WANG L, LING M X, WANG D Z, et al. Line-of-sight stabilization system based on fractional-order control [C] // 2008 2nd International Symposium on Systems and Control in Aerospace and Astronautics. Piscataway: IEEE Press, 2008: 1-4.
- [14] 贾琳, 孟卫锋. 滑模变结构控制在惯性平台稳定回路中的应用 [J]. 科学技术与工程, 2009, 9(2): 433-436.
- JIA L, MENG W F. Sliding mode variable structure control in the stabilization loop of inertial platform [J]. Science Technology and Engineering, 2009, 9(2): 433-436 (in Chinese).
- [15] 李红光, 鱼云岐, 宋亚民. 最优控制在车载惯性平台稳定回路中的应用 [J]. 应用光学, 2007, 28(3): 251-256.
- LI H G, YU Y Q, SONG Y M. Application of optimal control for stabilization loop of vehicle inertial platform [J]. Journal of Applied Optics, 2007, 28(3): 251-256 (in Chinese).
- [16] WANG C E, TANG J Q. Design and mathematical analysis of a novel reluctance force-type hybrid magnetic bearing for flywheel with gimballing capability [J]. Mathematical Problems in Engineering, 2013, 2013: 836058.
- [17] TANG J Q, XIANG B, WANG C E. Rotor's suspension for vernier-gimballing magnetically suspended flywheel with conical magnetic bearing [J]. ISA Transactions, 2015, 58: 509-519.
- [18] 王新华, 陈增强, 袁著祉. 基于扩张观测器的非线性不确定系统输出跟踪 [J]. 控制与决策, 2004, 19(10): 1113-1116.
- WANG X H, CHEN Z Q, YUAN Z Z. Output tracking based on extended observer for nonlinear and uncertain systems [J]. Control and Decision, 2004, 19(10): 1113-1116 (in Chinese).
- [19] 王新华, 刘金琨. 微分器设计及应用: 信号滤波与求导 [M]. 北京: 电子工业出版社, 2010: 152-153.
- WANG X H, LIU J K. Differentiator design and application: Signal filtering and differentiation [M]. Beijing: Publishing House of Electronics Industry, 2010: 152-153 (in Chinese).
- [20] AHRENS J H, KHALIL H K. High-gain observers in the presence of measurement noise: A switched-gain approach [J]. Automatica, 2009, 45(4): 936-943.
- [21] 夏长峰, 蔡远文, 任元, 等. MSCSG 转子系统的扩展双频 Bode 图稳定性分析方法 [J]. 宇航学报, 2018, 39(2): 168-176.
- XIA C F, CAI Y W, REN Y, et al. Stability analysis method with extended double-frequency Bode diagram for rotor of MSCSG [J]. Journal of Astronautics, 2018, 39(2): 168-176 (in Chinese).
- [22] GARCIA-SANZ M. The Nyquist stability criterion in the Nichols chart [J]. International Journal of Robust and Nonlinear Control, 2016, 26(12): 2643-2651.

# Lorentz inertial stability platform control based on KF-LESO-PID

XIONG Ying<sup>1</sup>, LIU Qiang<sup>1,\*</sup>, REN Yuan<sup>2</sup>, FAN Yahong<sup>3</sup>, SUN Jinji<sup>4</sup>

(1. Institute of Precision Electromagnetic Equipment and Advanced Measurement Technology,  
Beijing Institute of Petrochemical Technology, Beijing 102617, China;

2. Department of Aerospace Science and Technology, Space Engineering University, Beijing 101416, China;

3. Science and Technology on Space Intelligent Control Laboratory, Beijing Institute of Control Engineering, Beijing 100190, China;

4. Key Laboratory of Inertial Technology, Beihang University, Beijing 100083, China)

**Abstract:** To overcome the disadvantages of the existing inertial stabilization platform such as large interference for using mechanical bearings, high difficulty for using air/liquid bearings and poor linearity for using magnetic resistance magnetic bearings, a new Lorentz inertial stability platform (LISP) based on Lorentz force deflection magnetic bearing is proposed. To suppress the influence of coupling effect and load-bearing friction resonance interference on the high-frequency attitude compensation control of the platform deflection channel, a digital control scheme based on LESO-PID combined with Kalman filter (KF) feedback is proposed. According to the structural characteristics of rotor tilt supported by Lorentz force magnetic bearing (LFMB), a dynamics model for LISP deflection is established; the tilting relationship of two radial channels is analyzed with the established model, and the linear extended state observer (LESO) and Kalman filter feedback control is introduced into PID controller to suppress friction resonance interference and coupling effects; a digital control system based on DSP and FPGA is construed, and the control method is digitalized in a discrete form. The stability of the proposed control method is analyzed by logarithmic frequency characteristic criterion and Nichols curve, and the stabilities of the rotor deflection channel before and after importing LESO-Kalman are compared through simulation. Experimental results show that with traditional PID, the rotor system causes serious distortion at high frequency, while the system reduces noise and interference greatly after importing LESO-Kalman control. Meanwhile, the internal state parameters of the system can be monitored in real time. Experimental results verify the effectiveness of proposed control method to suppress the frictional resonance interference and coupling effects.

**Keywords:** Lorentz inertial stability platform (LISP); Lorentz force magnetic bearing (LFMB); LESO-PID control; Kalman filter (KF) feedback; stable control

**Received:** 2020-12-30; **Accepted:** 2021-04-02; **Published online:** 2021-04-23 16:06

**URL:** [kns.cnki.net/kcms/detail/11.2625.V.20210423.1118.001.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210423.1118.001.html)

**Foundation items:** Beijing Municipal Nature Science Foundation (3212004); Youth Top Talent Training Funded Project of Beijing (2017000026833ZK22); Support Project of High-Level Teachers in Beijing Municipal Universities in the Period of 13th Five-Year Plan (CIT&TCD201804034)

\* **Corresponding author.** E-mail: liuqiangbuaa@163.com

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2020.0724

# 基于 ESO 的电液位置伺服系统反步滑模控制

张震阳<sup>1</sup>, 汪成文<sup>1,2,3,\*</sup>, 郭新平<sup>4</sup>, 陈帅<sup>1</sup>

(1. 太原理工大学 机械与运载工程学院, 太原 030024;

2. 太原理工大学 新型传感器与智能控制教育部/山西省重点实验室, 太原 030024;

3. 浙江大学 流体动力与机电系统国家重点实验室, 杭州 310058; 4. 中南大学 机电工程学院, 长沙 410083)

**摘 要:** 针对阀控电液位置伺服系统未建模摩擦力、参数不确定性和外部随机干扰造成的复合扰动问题,提出一种基于扩张状态观测器(ESO)的反步滑模控制方法。ESO的设计可以对作动器速度、加速度和复合扰动进行在线估计,解决工程应用中难以测定的问题;基于 ESO 估计值和位移反馈信号进行反步滑模控制器设计,通过构造包含反步设计误差、滑模函数和观测器误差的 Lyapunov 函数,对所提控制方法进行稳定性证明;为验证所提方法的有效性,进行了 AMESim 和 MATLAB/Simulink 联合仿真,与 PID 控制器、传统的反步滑模控制器和基于 ESO 的滑模控制器的控制效果进行对比,并对仿真数据进行了分析。研究结果表明:所提方法可以有效抑制系统复合扰动,位移跟踪精度高,鲁棒性强。

**关键词:** 阀控缸;复合扰动;扩张状态观测器(ESO);反步滑模控制;位移跟踪控制  
**中图分类号:** V221+.3; TP271

**文献标志码:** A **文章编号:** 1001-5965(2022)06-1082-09

电液伺服系统具有功重比高、响应速度快、控制精度高等优点,被广泛应用于挖掘机<sup>[1-2]</sup>、电液负载模拟器<sup>[3-4]</sup>、工业机器人<sup>[5-6]</sup>等工业领域。然而,电液伺服系统具有参数不确定和非线性的特性,并且在作动器运动过程中可能受到外部随机干扰,不利于位移跟踪控制。为使电液伺服系统具有良好的控制性能,学者们对系统非线性和补偿外部随机干扰的控制方法进行了研究。Yao 等<sup>[7]</sup>针对非对称液压缸提出了一种自适应鲁棒控制方法,不仅考虑了惯性负载参数变化和液压系统参数不确定性,而且对外部不确定干扰和较难建模摩擦力进行了补偿。贾鹤鸣和宋文龙<sup>[8]</sup>提出了一种基于干扰观测器的滤波反步控制方法,可以对未知复合干扰进行在线估计,通过干扰抑制方法解决观测器误差带来的问题。Wang

等<sup>[9]</sup>基于液压系统动力学降阶模型进行干扰观测器和反步控制器的设计,解决了系统非匹配不确定性问题,并对系统内部压力动力学稳定性进行了证明。蒲明等<sup>[10]</sup>针对高阶非线性系统,通过改进的高阶滑模微分器对干扰进行估计,基于估计值进行递阶 Terminal 滑模控制器的设计,理论上可达到跟踪误差任意小。Guan 和 Pan<sup>[11]</sup>提出了一种自适应滑模控制方法解决电液伺服系统参数不确定性造成的非线性问题。Na 等<sup>[12-13]</sup>基于高阶滑模观测器设计了一种新的自适应控制器。文献[14-15]提出一种基于干扰观测器的反步位置跟踪控制器,对系统未知复合干扰进行估计,降低了反步递推过程的复杂程度。文献[16-19]提出自适应反步滑模控制方法,通过参数自适应得到电液伺服系统不确定参数,通过反步滑模的设

收稿日期: 2020-12-31; 录用日期: 2021-04-23; 网络出版时间: 2021-05-18 12:28

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20210517.1802.007.html

基金项目: 国家重点研发计划(2019YFB2004502); 山西省重点研发计划(201903D121069); 山西省回国留学人员科研资助项目(HGKY2019016); 流体动力与机电系统国家重点实验室开放基金(GZKF-201720)

\* 通信作者. E-mail: cwwang@yeah.net

**引用格式:** 张震阳, 汪成文, 郭新平, 等. 基于 ESO 的电液位置伺服系统反步滑模控制[J]. 北京航空航天大学学报, 2022, 48(6): 1082-1090. ZHANG Z Y, WANG C W, GUO X P, et al. Backstepping sliding mode control of electro-hydraulic position servo system based on ESO[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(6): 1082-1090 (in Chinese).

计抑制系统非匹配随机干扰,但控制器需要速度等信号的输入,增加了工程实践难度。韩京清等<sup>[20-21]</sup>针对一类存在不确定性的系统,提出了一种扩张状态观测器(extended state observer, ESO)对系统不确定性进行估计。文献[22-24]针对系统进行了滑模控制器的设计,保证滑模函数有限时间内收敛到滑模面,系统的鲁棒性较强。

基于以上研究,ESO 可以对系统状态和不确定扰动进行在线估计,反步法和滑模控制器的结合可以抑制系统扰动,增强系统鲁棒性;针对电液伺服系统在工程实践中速度、加速度和扰动信号的难以测定问题和阀控电液位置伺服系统非线性和参数不确定的特性,本文提出一种基于 ESO 的反步滑模控制方法,只需一个位移信号的反馈,既能够对速度、加速度和复合扰动信号进行观测,又能够对系统未建模摩擦力、随机干扰和参数不确定性造成的复合扰动进行抑制,使阀控电液位置伺服系统控制精度高,鲁棒性强。

首先,以 Brunovsky 标准型对阀控电液位置伺服系统进行数学建模,将系统非线性和不确定性部分整定到一个复合扰动函数中。然后,进行 ESO 的设计,对作动器速度、加速度和复合扰动进行在线估计。最后,基于观测器估计值设计了反步滑模控制方法,并对所提控制策略进行了稳定性证明,利用 AMESim 和 MATLAB/Simulink 进行联合仿真验证,并对仿真数据结果进行了分析。

### 1 系统描述

电液位置伺服系统结构简图如图 1 所示。其中,单向阀、定量泵、电机和油箱为对整个系统供油的元件,溢流阀起到定压溢流作用;双出杆对称液压缸为作动器,位移信号通过位移传感器进行反馈;控制器基于位移误差对伺服阀进行控制,最终实现作动器的位移跟踪控制。

由牛顿第二定律可得惯性负载的动力学方程为

$$P_L A = m\ddot{x}_p + b\dot{x}_p + F_d(t) \quad (1)$$

式中: $x_p$  为负载位移; $A$  为活塞作用面积;液压缸两腔压差  $P_L = P_1 - P_2$  ( $P_1$ 、 $P_2$  分别为液压缸两腔的压力); $m$  为负载质量; $b$  为黏性摩擦系数; $F_d(t)$  为系统未建模摩擦力和未知干扰力。

液压缸的压力-流量方程为

$$\dot{P}_L = \frac{4\beta_e}{V_t}(Q_L - A\dot{x}_p - C_t P_L) \quad (2)$$

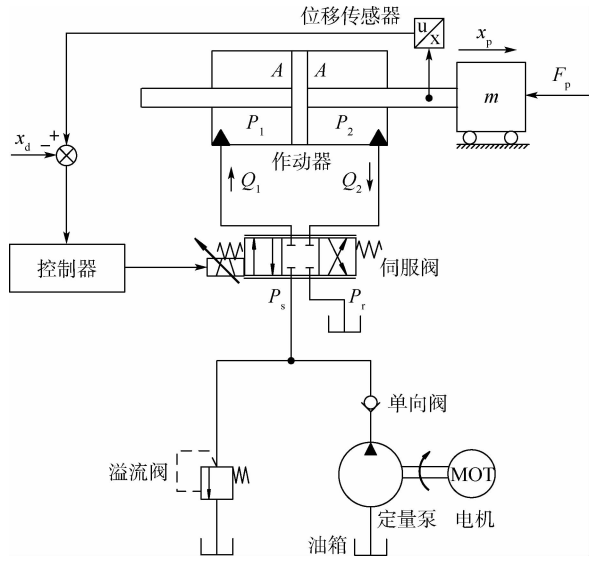


图 1 电液位置伺服系统结构

Fig. 1 Structure of electro-hydraulic position servo system

式中: $\beta_e$  为油液体积弹性模量; $V_t$  为液压缸油腔总的体积; $Q_L = (Q_1 + Q_2)/2$  为负载流量( $Q_1$  和  $Q_2$  分别为液压缸进油流量和回油流量); $C_t$  为液压缸总的内泄漏系数。

定义符号函数:

$$\text{sgn}(x) = \begin{cases} 1 & x \geq 0 \\ -1 & x < 0 \end{cases} \quad (3)$$

负载流量-伺服阀阀芯位移方程为

$$Q_L = C_d w x_v \sqrt{\frac{P_s - \text{sgn}(x_v) P_L}{\rho}} \quad (4)$$

式中: $C_d$  为伺服阀节流口的流量系数; $w$  为伺服阀节流口面积梯度; $x_v$  为阀芯位移; $P_s$  为供油压力; $\rho$  为液压油密度。

由于伺服阀的响应频率远大于液压缸动作的频宽,阀芯位移与控制输出可以近似等同于线性关系,即

$$x_v = k_{xv} u \quad (5)$$

式中: $k_{xv}$  为伺服阀增益; $u$  为系统的控制输出。

定义状态变量  $\mathbf{x} = [x_1, x_2, x_3]^T = [x_p, \dot{x}_p, \ddot{x}_p]^T$ , 系统状态方程为

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = x_3 \\ \dot{x}_3 = R_1 x_2 + R_3 u + \zeta \end{cases} \quad (6)$$

式中:  $R_1 = -\frac{4A^2\beta_e}{mV_t} + \frac{b^2}{m^2}$ ;  $R_3 = \frac{4A\beta_e C_d w k_{xv}}{mV_t} \sqrt{\frac{P_s - \text{sgn}(u) P_L}{\rho}}$ , 定义系统复合扰动  $\zeta = \Delta R_1 x_2 + R_2 P_L + \Delta R_3 u - \frac{b}{m} F_d(t) + \dot{F}_d(t)$ ,  $R_2 = -\frac{4A\beta_e C_t}{mV_t} - \frac{bA}{m^2}$ , 由于黏性摩擦系数  $b$ 、油液体积弹性模量  $\beta_e$ 、

等参数无法得到其准确值,本文  $R_1$  为根据近似参数得到的近似值,  $\Delta R_1$  为其误差。

伺服阀说明书中都会给出额定压降  $\Delta p_{drop} = 7 \text{ MPa}$  下的流量公式:

$$Q_r = C_d w k_{xv} u_{max} \sqrt{\frac{\Delta p_{drop}}{\rho}} \quad (7)$$

式中:  $Q_r$  为伺服阀额定流量;  $u_{max}$  为伺服阀最大控制输入。

由于参数  $C_d$ 、 $w$ 、 $\rho$  的准确值较难得到,将式(7)代入  $R_3$  得

$$R_3 = \frac{4A\beta_e Q_r}{mV_1 u_{max}} \sqrt{\frac{P_s - \text{sgn}(u)P_L}{\Delta p_{drop}}} \quad (8)$$

从式(8)可以看出,  $R_3$  的值随符号函数和负载压力的变化在一定范围内变化,由于  $R_3$  含有系统不确定参数和关于  $u$  的符号函数,而本文所提的控制方法可以对误差进行补偿,为了简化方法,  $R_3$  取其中间近似值,  $\Delta R_3$  为其误差。

## 2 控制器设计及稳定性分析

采用 ESO 的电液位置伺服系统反步滑模控制策略框图如图 2 所示。

图 2 中,电液位置伺服系统存在参数不确定性、模型误差和未知干扰等造成的复合扰动;电液位置伺服系统只需检测作动器位移信号,基于位移反馈信号,ESO 便可以对作动器速度、加速度及复合扰动进行在线估计;基于观测器估计值和位移反馈信号进行反步滑模控制器的设计,最终实现电液位置伺服系统的位移跟踪控制。在对基于 ESO 的电液位置伺服系统反步滑模控制器进行设计的过程中做出如下假设:①位移指令信号三阶可导;②  $\zeta$  有界且可导。

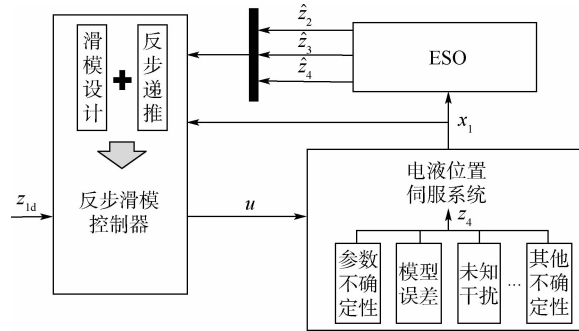


图 2 基于 ESO 的反步滑模控制策略框图

Fig.2 Block diagram of backstepping sliding mode control strategy based on ESO

### 2.1 扩张状态观测器的设计

定义  $z = [z_1, z_2, z_3, z_4]^T = [x_p, \dot{x}_p, \ddot{x}_p, \zeta]^T$ , 则式(6)经过状态扩张后的数学模型为

$$\begin{cases} \dot{z}_1 = z_2 \\ \dot{z}_2 = z_3 \\ \dot{z}_3 = R_1 z_2 + R_3 u + z_4 \\ \dot{z}_4 = h(t) \end{cases} \quad (9)$$

式中:  $h(t)$  为有界的不确定函数。

定义变量  $\hat{z}_i$  ( $i = 1, 2, 3, 4$ ) 分别为 ESO 对状态  $z_i$  ( $i = 1, 2, 3, 4$ ) 的估计,对式(9)构造的 ESO 为

$$\begin{cases} \dot{\hat{z}}_1 = \hat{z}_2 + q_1(z_1 - \hat{z}_1) \\ \dot{\hat{z}}_2 = \hat{z}_3 + q_2(z_1 - \hat{z}_1) \\ \dot{\hat{z}}_3 = R_1 \hat{z}_2 + R_3 u + \hat{z}_4 + q_3(z_1 - \hat{z}_1) \\ \dot{\hat{z}}_4 = q_4(z_1 - \hat{z}_1) \end{cases} \quad (10)$$

式中:  $q_i$  ( $i = 1, 2, 3, 4$ ) 为正数。

定义  $\tilde{z}_i = z_i - \hat{z}_i$  为 ESO 估计误差,则由式(9)减去式(10)可得估计误差的动态方程为

$$\begin{cases} \dot{\tilde{z}}_1 = -q_1 \tilde{z}_1 + \tilde{z}_2 \\ \dot{\tilde{z}}_2 = -q_2 \tilde{z}_1 + \tilde{z}_3 \\ \dot{\tilde{z}}_3 = -q_3 \tilde{z}_1 + R_1 \tilde{z}_2 + \tilde{z}_4 \\ \dot{\tilde{z}}_4 = -q_4 \tilde{z}_1 + h(t) \end{cases} \quad (11)$$

由式(11)可得

$$\dot{\tilde{z}} = \mathbf{A}\tilde{z} + \mathbf{H} \quad (12)$$

$$\text{式中: } \mathbf{A} = \begin{bmatrix} -q_1 & 1 & 0 & 0 \\ -q_2 & 0 & 1 & 0 \\ -q_3 & R_1 & 0 & 1 \\ -q_4 & 0 & 0 & 0 \end{bmatrix}; \mathbf{H} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ h(t) \end{bmatrix}, h(t)$$

有界,即存在  $\phi > 0$  使得  $\|\mathbf{H}\|_2 \leq \phi$ 。

当  $q_i > 0$  ( $i = 1, 2, 3, 4$ ) 时,可以使  $\det|\lambda \mathbf{I} - \mathbf{A}|$  满足为 Hurwitz 多项式,即  $\mathbf{A}$  为 Hurwitz 矩阵,因此,存在一个实对称正定矩阵  $\mathbf{P}$  和正定矩阵  $\mathbf{Q}$  使得:  $\mathbf{A}^T \mathbf{P} + \mathbf{P} \mathbf{A} = -\mathbf{Q}$ , 即存在  $\varepsilon > 0$  使得  $\|\tilde{z}\|_1 \leq \varepsilon$  成立。

### 2.2 反步滑模控制器设计及稳定性分析

基于 ESO 对速度、加速度及复合扰动的估计值进行控制器的设计,并证明了位移跟踪误差、滑模函数和观测器误差一致最终有界。在下面的推导过程中,  $k_i > 0$  ( $i = 1, 2, 3$ ),  $c_1 > 0$ ,  $c_2 > 0$  为控制器参数。

1) 定义  $e_1 = z_1 - z_{1d}$  为位移跟踪误差,  $z_{1d}$  为位移指令信号,对  $e_1$  求导得

$$\dot{e}_1 = z_2 - \dot{z}_{1d} \quad (13)$$

定义 Lyapunov 函数:

$$V_1 = \frac{1}{2} e_1^2 \quad (14)$$

对式(14)求导得

$$\dot{V}_1 = e_1 \dot{e}_1 = e_1(z_2 - \dot{z}_{1d}) \quad (15)$$

定义  $e_2 = z_2 - z_{2d}$ ,  $z_{2d}$  为虚拟控制变量,为保证

$\dot{V}_1$  负定,令  $z_{2d} - \dot{z}_{1d} = -k_1 e_1$ , 得

$$z_{2d} = -k_1 e_1 + \dot{z}_{1d} \quad (16)$$

由式(15)和式(16)得

$$\dot{V}_1 = -k_1 e_1^2 + e_1 e_2 \quad (17)$$

2) 定义 Lyapunov 函数为

$$V_2 = V_1 + \frac{1}{2} e_2^2 \quad (18)$$

对式(18)求导得

$$\dot{V}_2 = \dot{V}_1 + e_2 \dot{e}_2 = -k_1 e_1^2 + e_2(z_3 - \dot{z}_{2d} + e_1) \quad (19)$$

式中:  $\dot{z}_{2d} = -k_1(z_2 - \dot{z}_{1d}) + \ddot{z}_{1d}$ 。控制器设计时,  $z_2$

为观测器估计值  $\hat{z}_2$ , 定义  $\hat{z}_{2d} = -k_1(\hat{z}_2 - \dot{z}_{1d}) + \ddot{z}_{1d}$ ,

$$\hat{e}_2 = \hat{z}_2 - z_{2d}$$

定义  $e_3 = z_3 - z_{3d}$ ,  $z_{3d}$  为虚拟控制变量,为保证  $\dot{V}_2$  负定,令  $z_{3d} - \hat{z}_{2d} + e_1 = -k_2 \hat{e}_2$ , 得

$$z_{3d} = -k_2 \hat{e}_2 + \hat{z}_{2d} - e_1 \quad (20)$$

由式(20)和  $\dot{z}_{2d} - \hat{z}_{2d} = -k_1 \bar{z}_2$ ,  $e_2 - \hat{e}_2 = \bar{z}_2$  得

$$z_3 = e_3 + z_{3d} = e_3 - k_2 e_2 - e_1 + \dot{z}_{2d} + (k_1 + k_2) \bar{z}_2 \quad (21)$$

将式(21)代入(19)得

$$\dot{V}_2 = -k_1 e_1^2 - k_2 e_2^2 + e_2 e_3 + (k_1 + k_2) e_2 \bar{z}_2 \quad (22)$$

3) 定义滑模函数为

$$s = c_1 e_1 + c_2 e_2 + e_3 \quad (23)$$

对式(23)求导得

$$\dot{s} = c_1 \dot{e}_1 + c_2 \dot{e}_2 + \dot{e}_3 = c_1 \dot{e}_1 + c_2 \dot{e}_2 + R_1 z_2 + R_3 u + z_4 - \dot{z}_{3d} \quad (24)$$

式中:  $\dot{z}_{3d} = -k_2[\dot{\hat{z}}_2 + k_1(z_2 - \dot{z}_{1d}) - \dot{z}_{1d}] - k_1(\dot{\hat{z}}_2 - \ddot{z}_{1d}) + \ddot{z}_{1d} - z_2 - \dot{z}_{1d}$ 。

为解决系统状态  $z_i (i=2,3,4)$  在工程应用中难以测定的问题,在所提出控制器设计过程中使用的是 ESO 对  $z_i (i=2,3,4)$  的估计  $\hat{z}_i (i=2,3,4)$ , 定义  $\hat{z}_3 - z_{3d} = \hat{e}_3$ , 得到实际控制器滑模函数为

$$\hat{s} = c_1 e_1 + c_2 \hat{e}_2 + \hat{e}_3 \quad (25)$$

定义  $\hat{z}_{3d} = -k_2[\dot{\hat{z}}_2 + k_1(\hat{z}_2 - \dot{z}_{1d}) - \dot{z}_{1d}] - k_1(\dot{\hat{z}}_2 - \ddot{z}_{1d}) + \ddot{z}_{1d} - \hat{z}_2 - \dot{z}_{1d}$ , 由式(24)和式(25)可得设计控制律为

$$u = -\frac{1}{R_3} \{c_1(\hat{z}_2 - \dot{z}_{1d}) + c_2[\hat{z}_3 + k_1(\hat{z}_2 - \dot{z}_{1d}) - \dot{z}_{1d}] + R_1 \hat{z}_2 + \hat{z}_4 - \hat{z}_{3d} + k_3 \hat{s}\} \quad (26)$$

由  $e_i - \hat{e}_i = z_i - z_{id} - (\hat{z}_i - z_{id}) = z_i - \hat{z}_i = \bar{z}_i$

( $i=2,3$ ) 得

$$s - \hat{s} = c_2 \bar{z}_2 + \bar{z}_3 \quad (27)$$

由式(22)、式(26)和式(27)得

$$\dot{s} = (c_1 + k_1 c_2 + k_3 c_2 + R_1 - k_1 k_2 - 1) \bar{z}_2 + (c_2 + k_3) \bar{z}_3 + \bar{z}_4 - k_3 s \quad (28)$$

定义 Lyapunov 函数为

$$V = V_2 + \frac{1}{2} s^2 + \bar{z}^T P \bar{z} \quad (29)$$

对式(29)求导,由式(23)和式(28)得

$$\begin{aligned} \dot{V} = \dot{V}_2 + s \dot{s} - \bar{z}^T Q \bar{z} + H^T P \bar{z} + \bar{z}^T P H = & -k_1 e_1^2 - (k_2 + c_2) e_2^2 - k_3 s^2 - \bar{z}^T Q \bar{z} - c_1 e_1 e_2 + \\ & (k_1 + k_2) e_2 \bar{z}_2 + s e_2 + s[(c_1 + k_1 c_2 + k_3 c_2 + R_1 - \\ & k_1 k_2 - 1) \bar{z}_2 + (c_2 + k_3) \bar{z}_3 + \bar{z}_4] + H^T P \bar{z} + \bar{z}^T P H \end{aligned} \quad (30)$$

由  $|\bar{z}_i| \leq \varepsilon$ ,  $\|H\|_2 \leq \phi$  和 Young 不等式得

$$\begin{aligned} \dot{V} \leq & -k_1 e_1^2 - (k_2 + c_2) e_2^2 - k_3 s^2 - \|Q\|_F \|\bar{z}\|_2^2 + \\ & c_1 |e_1| |e_2| + (k_1 + k_2) \varepsilon e_2 + s e_2 + \alpha s + \\ & \phi \varepsilon \|P\|_F + \phi \varepsilon \|P\|_F \leq -\left(k_1 - \frac{1}{2} c_1\right) e_1^2 - \\ & \left[k_2 + c_2 - \frac{1}{2} c_1 \eta - \frac{1}{2} (k_1 + k_2) \eta - \frac{1}{2} \eta\right] e_2^2 - \left(k_3 - \frac{1}{2} \eta - \frac{1}{2} \eta\right) s^2 - \|Q\|_F \|\bar{z}\|_2^2 + \\ & \frac{1}{2} \eta (k_1 + k_2) \varepsilon^2 + \frac{1}{2} \eta \alpha^2 + \phi \varepsilon \|P\|_F + \phi \varepsilon \|P\|_F \end{aligned} \quad (31)$$

式中:  $\alpha = \max\{c_1 + k_1 c_2 + k_3 c_2 + R_1 - k_1 k_2 - 1, c_2 + k_3, 1\}$ ;  $\varepsilon; \eta > 0$  为不等式参数。

由式(31)可得

$$\dot{V} \leq -\gamma \left(\frac{1}{2} e_1^2 + \frac{1}{2} e_2^2 + \frac{1}{2} s^2 + \bar{z}^T P \bar{z}\right) + \sigma = -\gamma V + \sigma \quad (32)$$

式中:  $\gamma = \min\left\{k_1 - \frac{1}{2} c_1, k_2 + c_2 - \frac{1}{2} c_1 \eta - \frac{1}{2} (k_1 + k_2) \eta - \frac{1}{2} \eta, k_3 - \frac{1}{2} \eta - \frac{1}{2} \eta, \frac{\|Q\|_F}{\lambda_{\max}(P)}, 2\right\}$ ,  $\lambda_{\max}(P)$

为  $P$  的最大特征值,  $\frac{c_1}{2k_1} < \eta < \frac{2(k_2 + c_2)}{c_1 + k_1 + k_2 + 1}$ ,

$$k_3 > \frac{1}{2} \left(\frac{1}{\eta} + \eta\right); \sigma = \frac{1}{2} \eta (k_1 + k_2) \varepsilon^2 + \frac{1}{2} \eta \alpha^2 + \phi \varepsilon \|P\|_F + \phi \varepsilon \|P\|_F$$

由式(32)得  $V(t) \leq V(0) e^{-\gamma t} + \frac{\sigma}{\gamma}$ , 则  $s, e_i, \bar{z}_i$

( $i=1,2,3$ ) 一致最终有界,收敛域为  $\Omega = \{s, e_i,$

$$\bar{z}_i \mid |s| \leq \sqrt{\frac{2\sigma}{\gamma}}, |e_i| \leq \sqrt{\frac{2\sigma}{\gamma}}, |\bar{z}_i| \leq \sqrt{\frac{\sigma}{\gamma \|P\|_F}}\}$$
, 通

过调节观测器和控制器参数可使位移跟踪误差任意小。

### 3 仿真与分析

为对本文所提控制方法进行验证,采用 MATLAB/Simulink 和 AMESim 进行联合仿真,仿

真模型搭建如图 3 所示。在 AMESim 模型中,考虑了电液位置伺服系统的内泄漏、静摩擦、库伦摩擦、黏性摩擦等对系统性能产生影响的因素,参数设置如表 1 所示。同时,在 AMESim 模型中可施加外部干扰力,在 MATLAB/Simulink 模型中进行位移指令信号输入、控制器算法编写及仿真数据采集。

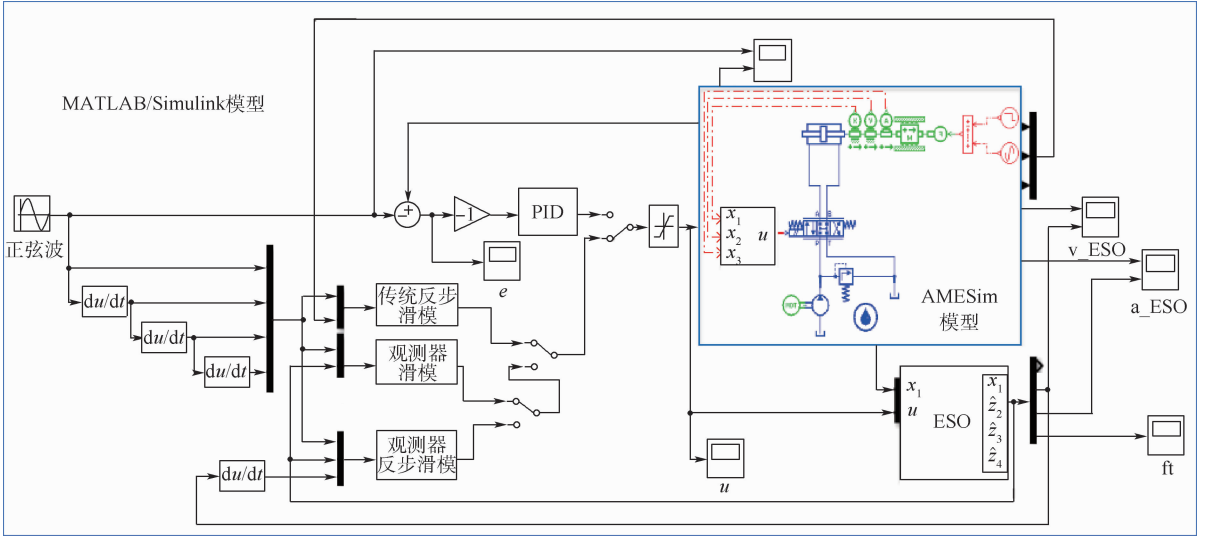


图 3 MATLAB/Simulink 和 AMESim 联合仿真模型

Fig.3 Co-simulation model of MATLAB/Simulink and AMESim

表 1 AMESim 模型中电液位置伺服系统参数

Table 1 Parameters of electro-hydraulic position servo system in AMESim model

参数	数值
液压缸行程/m	$\pm 0.1$
负载质量/kg	50
库伦摩擦力/N	20
静摩擦力/N	22
黏性摩擦系数/( $N \cdot (m \cdot s^{-1})^{-1}$ )	20
活塞杆直径/mm	25
伺服阀额定压降/Pa	$7 \times 10^6$
活塞直径/mm	40
泵排量/( $cm^3 \cdot r^{-1}$ )	6.4
内泄漏系数/( $m^3 \cdot s^{-1} \cdot Pa^{-1}$ )	$1.17 \times 10^{-13}$
油液体积弹性模量/Pa	$7 \times 10^8$
油腔总体积/ $m^3$	$1.53 \times 10^{-4}$
伺服阀额定电流/mA	40
伺服阀额定流量/( $L \cdot min^{-1}$ )	38

为证明本文所提控制方法的有效性,采用 3 种控制器作为对比:①在工程领域应用广泛的 PID 控制器;②传统反步滑模控制器;③基于 ESO 的滑模控制器。为保证仿真更接近现实工况,在 MATLAB/Simulink 中使用 Band-Limited White Noise 模块,在位移传感器反馈信号中加入噪声。

系统在外部强干扰力作用下,对 4 种控制器进行仿真对比, $z_{1d} = 0.02\sin(2\pi t)$  为位移指令信号,从 0 ~ 6 s 施加  $F_{sin} = 2500\sin(4\pi t)$  的正弦干扰力,在第 3 s 开始施加  $F_{step} = 4000$  N 的阶跃干扰力。

1) 本文控制方法中 ESO 参数的值为: $q_1 = 400, q_2 = 2 \times 10^4, q_3 = 1 \times 10^6, q_4 = 3 \times 10^{10}$ ,基于 ESO 估计值的反步滑模控制器参数的值为: $k_1 = 60, k_2 = 80, k_3 = 500, c_1 = 20, c_2 = 1800$ 。

2) PID 参数的确定基于位移指令信号进行试调,从较小的值进行调整,通过权衡系统的瞬态响应和稳态性能,最终确定: $K_p = 4300, K_i = 1600, K_d = 0$ 。

3) 传统反步滑模控制器设计为

$$e'_2 = z_2 - z'_{2d} \quad (33)$$

式中: $z'_{2d} = -k_1 e_1 + \dot{z}_{1d}$ 。

$$e'_3 = z_3 - z'_{3d} \quad (34)$$

式中: $z'_{3d} = -k_2 e'_2 - e_1 - k_1(z_2 - \dot{z}_{1d}) + \ddot{z}_{1d}$ 。

滑模函数为

$$s' = c'_1 e_1 + c'_2 e'_2 + e'_3 \quad (35)$$

该控制器控制律设计为

$$u' = -\frac{1}{R_3} \{ c'_1(z_2 - \dot{z}_{1d}) + c'_2 [z_3 + k'_1(z_2 - \dot{z}_{1d}) - \ddot{z}_{1d}] + R_1 z_2 - \dot{z}'_{3d} + k'_3 s' + k'_4 \text{sgn}(s') \} \quad (36)$$

式中:  $\dot{z}'_{3d} = -k'_2[z_3 + k'_1(z_2 - \dot{z}_{1d}) - \ddot{z}_{1d}] - k'_1(z_3 - \dot{z}_{1d}) - \ddot{z}_{1d} - z_2 - \dot{z}_{1d}$ 。

综合考虑系统瞬态响应和稳态性能,控制器参数值设置为:  $c'_1 = 300, c'_2 = 180, k'_1 = 300, k'_2 = 100, k'_3 = 50, k'_4 = 40\ 000$ 。

4) 基于 ESO 的滑模控制器的设计。

ESO 参数值与本文控制方法中 ESO 参数值相同,其滑模函数为

$$s'' = c''_1 e_1 + c''_2 e_2'' + e_3'' \quad (37)$$

式中:  $e_2'' = \hat{z}_2 - \dot{z}_{1d}; e_3'' = \hat{z}_3 - \ddot{z}_{1d}$ 。

控制律设计为

$$u'' = -\frac{1}{R_3} [c''_1(\hat{z}_2 - \dot{z}_{1d}) + c''_2(\hat{z}_3 - \ddot{z}_{1d}) + R_1 \hat{z}_2 - \ddot{z}_{1d} + \hat{z}_4 + k''_3 s''] \quad (38)$$

综合考虑系统瞬态响应和稳态性能,控制器参数值设置为:  $c''_1 = 2\ 000, c''_2 = 800, k''_3 = 1\ 200$ 。

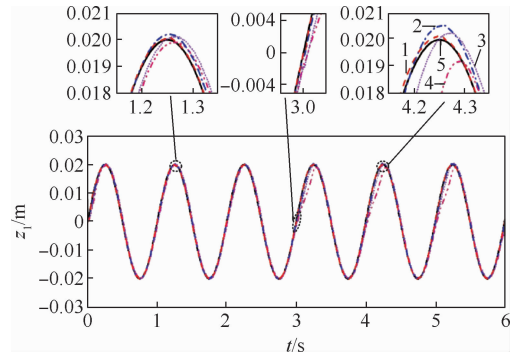
图 4 ~ 图 8 为外部干扰力作用下各控制器仿真结果对比。图 4 和图 5 分别为位移跟踪结果对比和位移跟踪误差结果对比,仿真结果表明,在前 3 s 只有正弦干扰力作用下,4 个控制器都能保证较好的位移跟踪。其中,基于 ESO 的反步滑模控制器(即本文控制方法)位移跟踪误差最小,传统反步滑模控制器位移跟踪效果最差;在第 3 s 突然受到外部阶跃干扰力时,基于 ESO 的反步滑模控制器的位移跟踪波动最小且调节时间最短;在第 3 ~ 6 s 系统受到外部正弦和阶跃干扰力的综合作用下,基于 ESO 的反步滑模控制器位移跟踪效果依然最好,基于 ESO 的滑模控制器控制效果次之,PID 控制器、传统反步滑模控制器跟踪误差较大且在作动器换向时误差出现较大波动。图 6 为控制器输出对比。结果表明,本文控制策略在外部正弦和阶跃干扰力综合作用下依然可以保证稳定的控制输出。图 7 和图 8 分别为 ESO 对作动器速度和加速度的估计。结果表明,ESO 跟踪效果理想,在受到强干扰力时对速度和加速度的估计波动小。图 9 为 ESO 对系统复合扰动的估计。结果表明,ESO 可随外部干扰力的变化对复合扰动进行在线估计,表明其对系统复合扰动的估计是有效的。

为定量分析 3 个控制器的控制效果及确保对比的公平性,通过对位移跟踪误差数据分析,完成控制器的性能评价。现引入以下 3 个性能指标:

1) 绝对误差的平均值。

$$E_{MAE} = \frac{1}{N} \sum_{i=1}^N |e_1(i)|$$

2) 绝对误差的标准差。



1—本文控制方法;2—基于 ESO 的滑模控制;3—位移指令信号;4—PID 控制;5—传统反步滑模控制。

图 4 位移跟踪结果对比

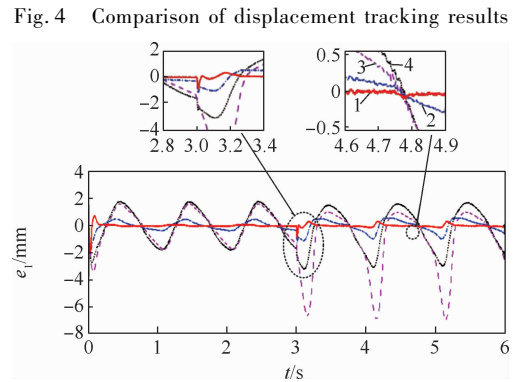


Fig. 4 Comparison of displacement tracking results

1—传统反步滑模控制;2—本文控制方法;3—基于 ESO 的滑模控制;4—PID 控制。

图 5 位移跟踪误差对比

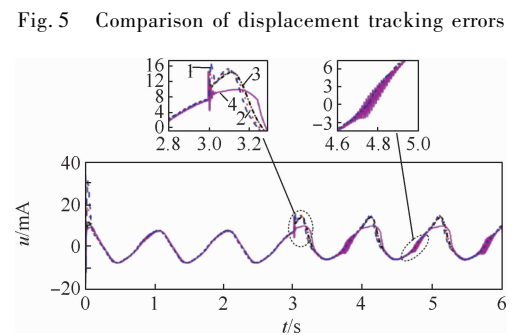
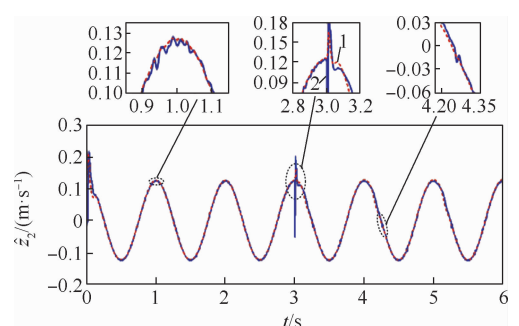


Fig. 5 Comparison of displacement tracking errors

1—本文控制方法;2—基于 ESO 的滑模控制;3—PID 控制;4—传统反步滑模控制。

图 6 控制输出对比

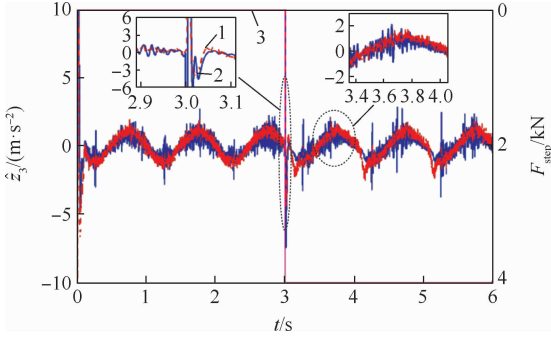
Fig. 6 Comparison of control outputs



1—速度估计;2—作动器速度。

图 7 速度估计结果

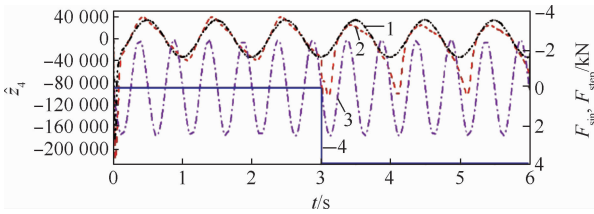
Fig. 7 Velocity estimation results



1—加速度估计;2—作动器加速度;3—阶跃干扰力。

图8 加速度估计结果

Fig.8 Acceleration estimation results



1—未加外力复合扰动估计;2—施加外力复合扰动估计;  
3—正弦干扰力;4—阶跃干扰力。

图9 复合扰动估计结果

Fig.9 Compound disturbance estimation results

表2 性能评价指标对比

Table 2 Comparison of performance evaluation indexes

控制方法	$E_{MAE}$	$E_{RMSE}$	$E_{ITAE}$
PID 控制	$1.2 \times 10^{-3}$	$7.1068 \times 10^{-4}$	22.897 8
传统反步滑模控制	$1.4 \times 10^{-3}$	$1.5 \times 10^{-3}$	28.429 4
基于 ESO 的滑模控制	$3.2992 \times 10^{-4}$	$2.5476 \times 10^{-4}$	6.233 7
基于 ESO 的反步滑模控制	$6.5989 \times 10^{-5}$	$1.3493 \times 10^{-4}$	1.020 2

### 4 结论

针对阀控电液位置伺服系统未建模摩擦力、参数不确定性和外部随机干扰造成的复合扰动问题,本文提出了一种基于 ESO 的反步滑模控制方法,对其稳定性进行了理论分析,得到滑模函数和观测器误差一致最终有界。根据联合仿真结果及误差数据分析可以得到如下结论:

1) 通过 ESO 对速度、加速度和复合扰动估计结果可得,在受到强扰动时,基于 ESO 的反步滑模控制方法可以对速度、加速度和复合扰动进行有效估计。

2) 基于 ESO 的反步滑模控制方法只需测定作动器位移信号,对解决工程中速度、加速度和扰动信号的测定问题具有一定的意义。

3) 作动器在受到外部强干扰力时,与 PID 控制器、传统反步滑模控制器和基于 ESO 的滑模控制器相比,基于 ESO 的反步滑模控制方法位移跟踪精度高,并能有效抑制复合扰动,鲁棒性强。

$$E_{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (|e_1(i)| - E_{MAE})^2}$$

3) 时间乘绝对误差的积分。

$$E_{ITAE} = \sum_{i=1}^N iT_s |e_1(i)|$$

式中: $T_s$  为仿真步长, $T_s = 0.001 s$ 。

通过数据分析得到的结果如表 2 所示。绝对误差的平均值反映了控制误差的平均水平,其中,基于 ESO 的反步滑模控制器的控制精度相对于 PID 控制器、传统反步滑模控制器和基于 ESO 的滑模控制器分别提高了约 94.5%、95.3% 和 80.0%;绝对误差的标准差反映了控制误差的离散程度,其中,基于 ESO 的反步滑模控制器控制误差离散程度明显最小,基于 ESO 的滑模控制器次之,传统反步滑模控制器控制误差离散程度最大;时间乘绝对误差的积分指标是把控制误差以时间加权,对初始误差不敏感,对后期误差很敏感,从分析结果可得,基于 ESO 的反步滑模控制器  $E_{ITAE}$  最小,即在作动器运行时受到外部干扰力时位移跟踪误差小,鲁棒性强。

### 参考文献 (References)

[ 1 ] LI Y H, HE L Y. Counterbalancing speed control for hydrostatic drive heavy vehicle under long down-slope [ J ]. IEEE/ASME Transactions on Mechatronics, 2015, 20(4): 1533-1542.

[ 2 ] 董致新, 黄伟男, 葛磊, 等. 泵阀复合进出口独立控制液压挖掘机特性研究 [ J ]. 机械工程学报, 2016, 52(12): 173-180.

DONG Z X, HUANG W N, GE L, et al. Research on the performance of hydraulic excavator with pump and valve combined separate meter in and meter out circuits [ J ]. Journal of Mechanical Engineering, 2016, 52(12): 173-180 (in Chinese).

[ 3 ] WANG C W, JIAO Z X, QUAN L. Nonlinear robust dual-loop control for electro-hydraulic load simulator [ J ]. ISA Transactions, 2015, 59: 280-289.

[ 4 ] WANG C W, JIAO Z X, WU S, et al. Nonlinear adaptive torque control of electro-hydraulic load system with external active motion disturbance [ J ]. Mechatronics, 2014, 24(1): 32-40.

[ 5 ] PI Y J, WANG X Y. Trajectory tracking control of a 6-DOF hydraulic parallel robot manipulator with uncertain load disturbances [ J ]. Control Engineering Practice, 2011, 19(2): 185-193.

[ 6 ] 付永领, 韩旭, 杨荣荣, 等. 电动静液作动器设计方法综述

- [J]. 北京航空航天大学学报, 2017, 43(10): 1939-1952.
- FU Y L, HAN X, YANG R R, et al. Review on design method of electro-hydrostatic actuator[J]. Journal of Beijing University of Aeronautics and Astronautics, 2017, 43(10): 1939-1952 (in Chinese).
- [7] YAO B, BU F P, JOHN R, et al. Adaptive robust motion control of single-rod hydraulic actuators: Theory and experiments[J]. IEEE/ASME Transactions on Mechatronics, 2000, 5(1): 79-91.
- [8] 贾鹤鸣, 宋文龙. 基于非线性干扰观测器的  $L_2$  滤波反步控制[J]. 中南大学学报(自然科学版), 2014, 45(8): 2640-2647.
- JIA H M, SONG W L.  $L_2$  filter backstepping control based on nonlinear disturbance observer[J]. Journal of Central South University (Science and Technology), 2014, 45(8): 2640-2647 (in Chinese).
- [9] WANG C W, ZHANG Z Y, WANG H, et al. Disturbance observer based output feedback control of hydraulic servo system considering mismatched uncertainties and internal pressure dynamics stability[J]. IET Control Theory and Applications, 2020, 14(8): 1046-1056.
- [10] 蒲明, 吴庆宪, 姜长生, 等. 非匹配不确定高阶非线性系统递阶终端滑模控制[J]. 自动化学报, 2012, 38(11): 1777-1793.
- PU M, WU Q X, JIANG C S, et al. Recursive Terminal sliding mode control for higher-order nonlinear system with mismatched uncertainties[J]. Acta Automatica Sinica, 2012, 38(11): 1777-1793 (in Chinese).
- [11] GUAN C, PAN S X. Adaptive sliding mode control of electro-hydraulic system with nonlinear unknown parameters[J]. Control Engineering Practice, 2008, 16(11): 1275-1284.
- [12] NA J, REN X M, ZHENG D D. Adaptive control for nonlinear pure-feedback systems with high-order sliding mode observer[J]. IEEE Transactions on Neural Networks and Learning Systems, 2013, 24(3): 370-382.
- [13] NA J, LI Y P, HUANG Y B, et al. Output feedback control of uncertain hydraulic servo systems[J]. IEEE Transactions on Industrial Electronics, 2020, 67(1): 490-500.
- [14] 石胜利, 李建雄, 方一鸣. 具有输入饱和的电液伺服系统反步位置跟踪控制[J]. 中南大学学报(自然科学版), 2016, 47(10): 3369-3374.
- SHI S L, LI J X, FANG Y M. Backstepping position tracking control for electro-hydraulic servo system with input saturation[J]. Journal of Central South University (Science and Technology), 2016, 47(10): 3369-3374 (in Chinese).
- [15] LIU J Y. Observer-based backstepping dynamic surface control for stochastic nonlinear strict-feedback systems[J]. Neural Computing & Applications, 2014, 24(5): 1067-1077.
- [16] JI X H, WANG C W, ZHANG Z Y, et al. Nonlinear adaptive position control of hydraulic servo system based on sliding mode back-stepping design method[J]. Proceedings of the Institution of Mechanical Engineers. Part I-Journal of Systems and Control Engineering, 2021, 235(4): 474-485.
- [17] YIN X X, LIN Y G, LI W, et al. Adaptive sliding mode backstepping pitch angle control of a variable-displacement pump controlled pitch system for wind turbines[J]. ISA Transactions, 2015, 58: 629-634.
- [18] 方一鸣, 焦宗夏, 王文宾, 等. 轧机液压伺服位置系统的自适应反步滑模控制[J]. 电机与控制学报, 2011, 15(10): 95-100.
- FANG Y M, JIAO Z X, WANG W B, et al. Adaptive backstepping sliding mode control for hydraulic servo position system of rolling mill[J]. Electric Machines and Control, 2011, 15(10): 95-100 (in Chinese).
- [19] 吉鑫浩, 汪成文, 陈帅, 等. 阀控电液位置伺服系统滑模反步控制方法[J]. 中南大学学报(自然科学版), 2020, 51(6): 1518-1525.
- JI X H, WANG C W, CHEN S, et al. Sliding mode backstepping control method for valve-controlled electro-hydraulic position servo system[J]. Journal of Central South University (Science and Technology), 2020, 51(6): 1518-1525 (in Chinese).
- [20] 韩京清. 一类不确定对象的扩张状态观测器[J]. 控制与决策, 1995, 10(1): 85-88.
- HAN J Q. Extended state observer for a class of uncertain plants[J]. Control and Decision, 1995, 10(1): 85-88 (in Chinese).
- [21] HAN J Q, HUANG Y. Analysis and design for the second order nonlinear continuous extended states observer[J]. Chinese Science Bulletin, 2000, 21(3): 1938-1944.
- [22] SHI P, XIA Y Q, LIU G P, et al. On designing of sliding-mode control for stochastic jump systems[J]. IEEE Transactions on Automatic Control, 2006, 51(1): 97-103.
- [23] 郭新平, 汪成文, 刘华, 等. 基于扩张状态观测器的泵控电液伺服系统滑模控制[J]. 北京航空航天大学学报, 2020, 46(6): 1159-1168.
- GUO X P, WANG C W, LIU H, et al. Sliding mode control of pump controlled electro-hydraulic servo system based on extended state observer[J]. Journal of Beijing University of Aeronautics and Astronautics, 2020, 46(6): 1159-1168 (in Chinese).
- [24] CHEN H P, JYH C R, JUHNG P S. Sliding mode control with varying boundary layers for an electro-hydraulic position servo system[J]. International Journal of Advanced Manufacturing Technology, 2005, 26(1-2): 117-123.

## Backstepping sliding mode control of electro-hydraulic position servo system based on ESO

ZHANG Zhenyang<sup>1</sup>, WANG Chengwen<sup>1,2,3,\*</sup>, GUO Xinping<sup>4</sup>, CHEN Shuai<sup>1</sup>

(1. College of Mechanical and Vehicle Engineering, Taiyuan University of Technology, Taiyuan 030024, China;

2. Key Laboratory of Advanced Transducers and Intelligent Control System, Ministry of Education and Shanxi Province, Taiyuan University of Technology, Taiyuan 030024, China;

3. State Key Laboratory of Fluid Power and Mechatronic Systems, Zhejiang University, Hangzhou 310058, China;

4. College of Mechanical and Electrical Engineering, Central South University, Changsha 410083, China)

**Abstract:** The backstepping sliding mode control method based on extended state observer (ESO) is proposed, which can solve the compound disturbance problem caused by unmodeled friction force, parameter uncertainty and external random disturbance. The ESO is designed to estimate the velocity and acceleration of actuator. The backstepping sliding mode controller is designed based on the displacement feedback signal and the estimated values of ESO. By constructing Lyapunov function including backstepping design error, sliding mode function and observer error, the stability of the proposed control method is proved. In order to verify the effectiveness of the proposed method, AMESim and MATLAB/Simulink co-simulation is carried out to compare with PID controller, traditional backstepping sliding mode controller and sliding mode controller based on ESO, and the simulation data is analyzed. The results show that the proposed method can effectively suppress the compound disturbance of the system, with higher precision and stronger robustness.

**Keywords:** valve controlled cylinder; compound disturbance; extended state observer (ESO); backstepping sliding mode control; displacement tracking control

**Received:** 2020-12-31; **Accepted:** 2021-04-23; **Published online:** 2021-05-18 12:28

**URL:** [kns.cnki.net/kcms/detail/11.2625.V.20210517.1802.007.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210517.1802.007.html)

**Foundation items:** National Key R & D Program of China (2019YFB2004502); Key Research and Development Program of Shanxi Province (201903D121069); Research Project Supported by Shanxi Scholarship Council of China (HGKY2019016); Open Foundation of the State Key Laboratory of Fluid Power and Mechatronic Systems (GZKF-201720)

\* **Corresponding author.** E-mail: cwwang@yeah.net

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2020.0725

# 多四旋翼无人机系统分布式分层编队合围控制

郑伟铭<sup>1</sup>, 徐扬<sup>2,3</sup>, 罗德林<sup>1,\*</sup>

(1. 厦门大学 航空航天学院, 厦门 361102; 2. 西北工业大学 民航学院, 西安 710072;

3. 西北工业大学太仓长三角研究院, 苏州 215400)

**摘 要:** 针对带有多个领航者与跟随者的欠驱动四旋翼无人机群系统, 提出了一种分布式分层编队合围控制方法。设计分层分布式有限时间滑模估计器, 实现在仅有部分领航者获取到期望轨迹的条件下, 每架无人机都能生成其满足控制需求的估计位置信息。针对六自由度欠驱动四旋翼无人机模型的特点, 提出一种无人机位置层和姿态层的分层控制方法, 实现了无人机对所生成的估计位置的跟踪控制, 该方法采用高阶导数逼近算法, 防止在求解期望角速度的过程中出现微分爆炸。所提方法能在满足姿态稳定收敛的条件下实现有效的编队合围控制。通过数值仿真验证了所提方法的有效性。

**关键词:** 编队合围控制; 编队控制; 四旋翼无人机; 欠驱动系统; 分布式控制

**中图分类号:** V221+.3; TB553

**文献标志码:** A

**文章编号:** 1001-5965(2022)06-1091-15

近年来, 无人机理论和技术得到了快速发展, 无人机能够实现的功能愈发强大, 在一些领域正渐渐取代载人飞行器, 尤其是在一些人类难以到达的环境和具有挑战的项目中有着很高的效率<sup>[1-4]</sup>。然而, 由于任务需求和工作环境变得不断繁杂, 单架无人机在任务执行能力上显露出了局限性。多无人机群系统不但能发挥无人机的优势, 还能避免单无人机的局限性, 在部分个体发生故障的情况下还可以完成多样的任务, 是未来无人机发展的重要方向<sup>[5-8]</sup>。

当前的多无人机协同控制主要还停留在一致性、跟踪问题和编队问题<sup>[9-11]</sup>, 这些基本的协同控制问题与工程实际问题难以结合。多无人机群系统的编队合围控制是由多个领航者无人机的编队控制和多个跟随者无人机的合围控制组成, 能够满足多种任务需求, 有效避免了基本协同控制问题的单一性。此外, 通过改变

系统内领航者和跟随者数量、各无人机之间的通信拓扑结构, 即可将编队合围问题转换为基本的协同控制问题<sup>[12]</sup>。

文献[13]给出了编队合围控制的运动模型, 即每个领航者都能跟踪到期望轨迹, 而每个跟随者都能分布到由领航者围成的凸包内, 由于每个领航者都需要实时获取到期望轨迹, 该方案属于分散式工作模式。文献[14-15]给出了分布式编队合围控制的运动模型, 即领航者无需获取期望轨迹, 只需要获取到邻居间分布式的期望相对位置即可完成编队, 跟随者同样能够分布到由领航者围成的凸包内, 该方案无需获取外界提供的期望轨迹, 更符合实际工程的需求, 但是其局限性在于只能对静止的期望状态进行处理。

对于多无人机群系统, 无论是基本协同控制问题还是编队合围控制问题, 大多研究都是将无

收稿日期: 2020-12-31; 录用日期: 2021-02-06; 网络出版时间: 2021-02-22 16:47

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20210222.1636.003.html](http://kns.cnki.net/kcms/detail/11.2625.V.20210222.1636.003.html)

基金项目: 航空电子系统综合技术重点实验室和航空科学基金联合资助项目(20185568005)

\* 通信作者。E-mail: luodelin1204@xmu.edu.cn

**引用格式:** 郑伟铭, 徐扬, 罗德林. 多四旋翼无人机系统分布式分层编队合围控制[J]. 北京航空航天大学学报, 2022, 48(6): 1091-1105. ZHENG W M, XU Y, LUO D L. Distributed hierarchical formation-containment control of multiple quadrotor UAV systems [J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(6): 1091-1105 (in Chinese).

人机简化为低阶的线性模型,不能真实地表示无人机的机动特性<sup>[16-19]</sup>。一些学者还将高阶积分器质点模型<sup>[20-21]</sup>、欧拉-拉格朗日模型<sup>[22-23]</sup>考虑到编队合围控制问题中,但是这些现有的研究模型都是全驱动的,而一些工程应用中的模型往往是欠驱动的。

四旋翼无人机作为无人机的一个重要类别,能够在狭隘的空间内完成垂直起降、全方位导航等任务,且具有较好的机动性。四旋翼无人机飞行所需的升力和姿态变化所需的力矩由4个电机控制转速的螺旋桨提供,只要求解出所需的升力和力矩,就可以反解出每个电机的转速。将升力和具有3个维度的力矩作为输入,四旋翼无人机在空间中位置和姿态变化的6个状态作为输出,因此,四旋翼无人机是一种典型的六自由度欠驱动系统,其模型可分为三输入、三输出的全驱动旋转部分和单输入、三输出的欠驱动平动部分,这2个部分之间还具有耦合关系项。四旋翼无人机在三维空间飞行时,非线性和欠驱动特性严重影响其控制性能<sup>[24]</sup>。因此,将六自由度欠驱动四旋翼模型与编队合围控制方法相结合,可以有效提高编队系统的机动性。

基于上述情况,本文在图论、李雅普诺夫稳定性理论、群系统编队合围控制、四旋翼无人机控制等研究成果的基础上,针对带有多个领航者与跟随者的欠驱动四旋翼无人机群系统,在信息从领航者单向流向跟随者的拓扑结构下,开展对分布式分层编队合围控制问题的研究。①将六自由度欠驱动四旋翼无人机模型与编队合围控制相结合,将群系统分为领航者层和跟随者层,分别设计分布式有限时间滑模估计器,使得每架无人机都能估计出满足编队合围控制要求的估计位置;②对于每架无人机,其内部又分为位置控制层与姿态控制层,设计控制方法使得每架无人机的实际位置跟踪到其生成的估计位置,姿态收敛到稳态值,从而实现实际意义上的编队合围;③无人机控制方法中所需的高阶速度导数信息较为难求,本文设计了高阶导数逼近算法,使得定义的新变量能够近似逼近所需的高阶导数,防止微分爆炸。

## 1 问题描述

### 1.1 符号含义

$R^{m \times n}$ 和 $R^n$ 分别为 $m \times n$ 维和 $n \times 1$ 的欧几里得空间; $I_n$ 和 $\mathbf{1}_n$ 分别为 $n$ 阶单位方阵和 $n$ 维元素为1

的列向量; $\mathbf{0}_{m \times n}$ 为 $m \times n$ 维元素为0的矩阵; $\otimes$ 表示Kronecker乘积; $\odot$ 表示四元数乘积; $\lambda_{\max}(\cdot)$ 和 $\lambda_{\min}(\cdot)$ 分别为矩阵的最大特征值和最小特征值; $\text{sgn}(\cdot)$ 代表符号函数; $\text{diag}\{x_1, x_2, \dots, x_n\}$ 表示由标量元素 $x_1, x_2, \dots, x_n$ 构成的对角矩阵;

$\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|$ 和 $\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$ 分别为向量 $\mathbf{x}$ 的1-范数和2-范数, $\mathbf{x} = [x_1, x_2, \dots, x_n]^T \in R^n$ 。

**定义1**  $S(\bar{\mathbf{x}}) = [0, -\bar{x}_3, \bar{x}_2; \bar{x}_3, 0, -\bar{x}_1; -\bar{x}_2, \bar{x}_1, 0]$ 为与向量 $\bar{\mathbf{x}} = [\bar{x}_1, \bar{x}_2, \bar{x}_3]^T \in R^3$ 相关的斜对称矩阵。

**定义2** 如果任意 $\mathbf{x} \in C$ 和 $\mathbf{y} \in C$ ,对于任意 $Z \in [0, 1]$ 均有 $(1-z)\mathbf{x} + z\mathbf{y} \in C$ 成立,则称集合 $C \subset R^n$ 是一个凸集。有限点集 $P = \{p_1, p_2, \dots, p_n\}$ 构成的凸包 $\text{Co}(P)$ 是包含点集 $P$ 内所有点的最小凸集, $\text{Co}(P) = \left\{ \sum_{i=1}^n \kappa_i p_i \mid p_i \in P, \kappa_i > 0, \sum_{i=1}^n \kappa_i = 1 \right\}$ 。

### 1.2 图论基础

令 $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ 表示有向图,该图由点集 $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$ 与边集 $\mathcal{E} \subseteq \{(i, j) : i, j \in \mathcal{V}, i \neq j\}$ 组成。图 $\mathcal{G}$ 的邻接矩阵为 $\mathbf{M} = [m_{ij}] \in R^{n \times n}$ , $m_{ij}$ 为节点 $v_i$ 与 $v_j$ 之间的连接权重。如果节点 $v_j$ 能够获取到节点 $v_i$ 的信息,则 $m_{ij} = 1$ ,否则 $m_{ij} = 0$ 。记节点 $v_i$ 的邻居集合为 $\mathcal{N}_i = \{j \in \mathcal{V} : m_{ij} = 1\}$ 。图 $\mathcal{G}$ 的拉普拉斯矩阵表示为 $\mathbf{L} = [l_{ij}] \in R^{n \times n}$ , $l_{ii} = \sum_{j=1, j \neq i}^n m_{ij}$ , $l_{ij} = -m_{ij}$ 。如果图 $\mathcal{G}$ 中存在节点 $v_i$ ,使得从 $v_i$ 出发抵达任意其他节点的有向路径存在,则该图包含一个有向生成树,且 $v_i$ 被称为根节点。

对于一个四旋翼无人机编队系统,把第1~ $N$ 架无人机记为领航者,第 $N+1$ ~ $N+M$ 架无人机记为跟随者。分别令领航者和跟随者的下标集合为 $\mathcal{L} = \{1, 2, \dots, N\}$ 和 $\mathcal{F} = \{N+1, N+2, \dots, N+M\}$ 。用拉普拉斯矩阵 $\bar{\mathbf{L}} \in R^{(N+M+1) \times (N+M+1)}$ 来表示该编队的拓扑连接方式。 $\bar{\mathbf{L}}$ 的定义为

$$\bar{\mathbf{L}} = \begin{bmatrix} 0 & \mathbf{0}_{1 \times N} & \mathbf{0}_{1 \times M} \\ \mathbf{L}_{\text{dl}} & \mathbf{L}_{\text{ll}} & \mathbf{0}_{N \times M} \\ \mathbf{0}_{M \times 1} & \mathbf{L}_{\text{lf}} & \mathbf{L}_{\text{ff}} \end{bmatrix} \quad (1)$$

式中: $\mathbf{L}_{\text{ll}}$ 表示领航者和领航者之间的连接; $\mathbf{L}_{\text{lf}}$ 表示领航者到跟随者的连接; $\mathbf{L}_{\text{ff}}$ 表示跟随者和跟随者之间的连接; $\mathbf{L}_{\text{dl}} = [-m_{10}, -m_{20}, \dots, -m_{N0}]^T$ ,若第 $i$ 个领航者能够获取到期望位置与期望速度,则 $m_{i0} = 0$ ,否则 $m_{i0} > 0, i \in \mathcal{L}$ 。

定义一个虚拟智能体 0, 则领航者对期望位置与速度的获取能力可以等效表示为领航者与虚拟智能体 0 的连接情况。

**假设 1** 领航者之间、跟随者之间的信息流是双向的, 领航者到跟随者之间的信息流是单向的。

**假设 2** 对于每个领航者, 虚拟智能体 0 到该领航者有一条有向路径; 对于每个跟随者, 至少存在一个领航者到该跟随者有一条有向路径。

**引理 1**<sup>[25]</sup> 若多智能体系统的拓扑连接满足假设 1 和假设 2, 则  $L_{\text{in}}$  和  $L_{\text{ff}}$  是正定对称的。  $-L_{\text{in}}^{-1}L_{\text{in}}$  和  $-L_{\text{ff}}^{-1}L_{\text{ff}}$  的所有元素非负且每一行的和为 1。

### 1.3 系统模型

考虑具有  $N + M$  架欠驱动四旋翼无人机的编队系统, 其中每一架无人机可以被建模为六自由度刚体运动模型。该模型在惯性坐标系  $\mathcal{F}_I = \{O_I x_I y_I z_I\}$  和机体坐标系  $\mathcal{F}_B = \{O_B x_B y_B z_B\}$  下描述。惯性坐标系的原点位于地球上某点, 三轴指向北-东-地; 机体坐标系的原点位于无人机的质心, 三轴指向前-右-下, 如图 1 所示。使用单位四元数  $\mathbf{Q}_i = [\mathbf{q}_i^T, \eta_i]^T \in \mathbb{Q} = \{\mathbf{Q}_i \in \mathbb{R}^4 \mid \mathbf{q}_i^T \mathbf{q}_i + \eta_i^2 = 1\}$  和旋转矩阵  $\mathbf{R}(\mathbf{Q}_i)$  来表示第  $i$  架无人机的姿态, 其关系为<sup>[26]</sup>

$$\mathbf{R}(\mathbf{Q}_i) = (\eta_i^2 - \mathbf{q}_i^T \mathbf{q}_i) \mathbf{I}_3 + 2\mathbf{q}_i^T \mathbf{q}_i - 2\eta_i \mathbf{S}(\mathbf{q}_i) \quad (2)$$

式中:  $\mathbf{q}_i$  为三维列向量, 表示该四元数的矢量部分;  $\eta_i$  为该四元数的标量部分。

则第  $i$  架四旋翼无人机的运动学方程可以表示为

$$\begin{cases} \dot{\mathbf{p}}_i = \mathbf{v}_i \\ \dot{\mathbf{v}}_i = g\hat{\mathbf{e}}_3 - \frac{T_i}{m_i} \mathbf{R}(\mathbf{Q}_i)^T \hat{\mathbf{e}}_3 \end{cases} \quad (3)$$

式中:  $\mathbf{p}_i = [p_{i,x}, p_{i,y}, p_{i,z}]^T$  和  $\mathbf{v}_i = [v_{i,x}, v_{i,y}, v_{i,z}]^T$  分别为无人机在惯性坐标系  $\mathcal{F}_I$  中的位置和速度;



图 1 四旋翼无人机示意图

Fig. 1 Schematic diagram of quadrotor UAV

$g$  为重力加速度;  $\hat{\mathbf{e}}_3 = [0, 0, 1]^T$ ;  $m_i$  为无人机的质量;  $T_i$  为无人机沿着机体坐标系  $\mathcal{F}_B$  的  $z_B$  轴方向上的推力输入;  $i \in \mathcal{L} \cup \mathcal{F}$ 。

第  $i$  架四旋翼无人机的动力学方程可以表示为

$$\begin{cases} \dot{\mathbf{Q}}_i = \frac{1}{2} \mathbf{G}(\mathbf{Q}_i) \boldsymbol{\omega}_i \\ \mathbf{J}_i \dot{\boldsymbol{\omega}}_i = \boldsymbol{\tau}_i - \mathbf{S}(\boldsymbol{\omega}_i) \mathbf{J}_i \boldsymbol{\omega}_i \end{cases} \quad (4)$$

式中:  $\boldsymbol{\omega}_i = [\omega_{i,x}, \omega_{i,y}, \omega_{i,z}]^T$  为无人机在机体坐标系  $\mathcal{F}_B$  中的角速度;  $\mathbf{G}(\mathbf{Q}_i) = [\eta_i \mathbf{I}_3 + \mathbf{S}(\mathbf{Q}_i), -\mathbf{q}_i]^T$ ;  $\mathbf{J}_i = \text{diag}\{J_{i,x}, J_{i,y}, J_{i,z}\}$  和  $\boldsymbol{\tau}_i = [\tau_{i,x}, \tau_{i,y}, \tau_{i,z}]^T$  分别为惯性张量矩阵和无人机在机体坐标系  $\mathcal{F}_B$  中的力矩输入;  $i \in \mathcal{L} \cup \mathcal{F}$ 。

令  $[f_{i,1}, f_{i,2}, f_{i,3}, f_{i,4}]^T$  表示第  $i$  架四旋翼无人机中 4 个旋翼产生的推力, 则推力输入  $T_i$  和力矩输入  $\boldsymbol{\tau}_i = [\tau_{i,x}, \tau_{i,y}, \tau_{i,z}]$  可以表示为

$$\begin{bmatrix} T_i \\ \tau_{i,x} \\ \tau_{i,y} \\ \tau_{i,z} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & -d_i & 0 & d_i \\ d_i & 0 & -d_i & 0 \\ C_i & -C_i & C_i & -C_i \end{bmatrix} \begin{bmatrix} f_{i,1} \\ f_{i,2} \\ f_{i,3} \\ f_{i,4} \end{bmatrix} \quad (5)$$

式中:  $d_i$  为每个旋翼与无人机重心的距离;  $C_i$  为反扭矩系数;  $i \in \mathcal{L} \cup \mathcal{F}$ 。

### 1.4 控制目标

本文设计了一个分层控制框架: ①在第 1 层中, 只有部分领航者无人机可以连接到虚拟智能体 0, 从而获取到期望轨迹。利用分布式有限时间滑模估计器, 使得每一架领航者无人机都能够估计出期望位置并生成队形, 同时, 估计出的速度可以收敛到期望轨迹的导数值。②在第 2 层中, 跟随者无人机接收由领航者传递而来的估计信息, 通过分布式有限时间滑模估计器, 使得每一架跟随者无人机生成的估计位置可以收敛到由领航者无人机估计位置构成的凸包内, 生成的估计速度可以收敛到领航者的估计速度。

对于每一架无人机的控制, 又将其设计为位置控制层和姿态控制层。在位置控制层中, 设计控制力  $T_i$ , 使其实际的位置与速度跟踪到估计值。在姿态控制层中, 设计控制力矩  $\boldsymbol{\tau}_i$  使其姿态收敛。值得注意的是, 在此设计过程中, 采用了一种高阶导数的逼近算法, 避免了求取期望角速度过程中出现微分爆炸。由所求得每一架无人机的控制力和控制力矩, 即可对其位置和姿态进行控制, 进而实现多四旋翼无人机系统的编队包围控制, 该系统的控制架构如图 2 所示。

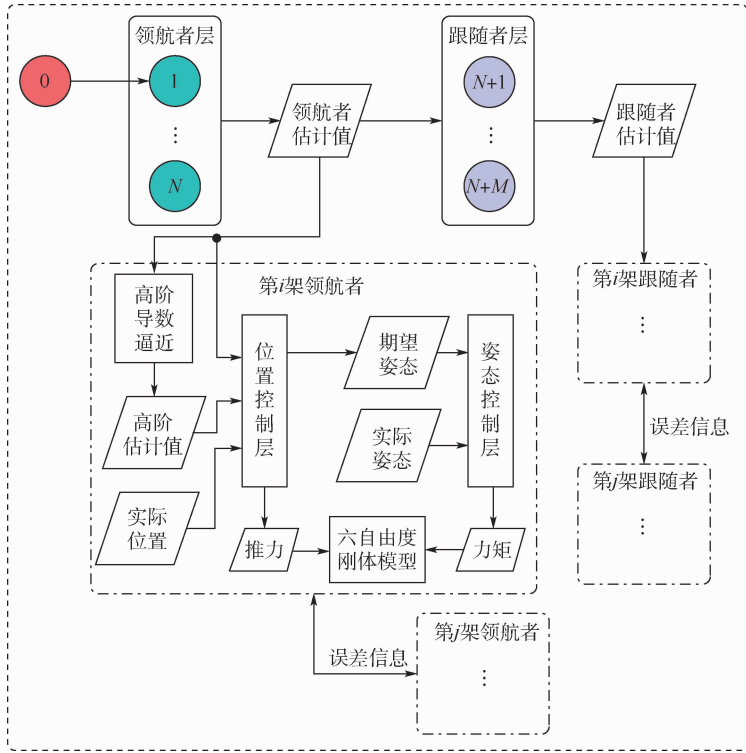


图2 编队合围控制架构

Fig.2 Formation-containment control architecture

## 2 分层分布式编队合围估计

### 2.1 领航者层

为使所有领航者无人机都能估计出期望位置和期望速度,对于第*i*架领航者无人机,设计如下有限时间滑模估计器:

$$\begin{cases} \dot{\hat{p}}_i = \hat{v}_i - \alpha_1 \text{sgn} \left( \sum_{j=0}^N m_{ij} (\hat{p}_i - \hat{p}_j - \delta_{ij}) \right) \\ \dot{\hat{v}}_i = -\alpha_2 \text{sgn} \left( \sum_{j=0}^N m_{ij} (\hat{v}_i - \hat{v}_j) \right) \end{cases} \quad (6)$$

式中: $\alpha_1$ 和 $\alpha_2$ 为正常数增益;常向量 $\delta_{ij} = \delta_{i0} - \delta_{j0}$ , $\delta_{i0}$ 和 $\delta_{j0}$ 分别为第*i*架和第*j*架领航者无人机相对于虚拟智能体0的位置偏差, $i, j \in \mathcal{L}$ ;  $\hat{p}_i$ 和 $\hat{v}_i$ 分别为第*i*架领航者无人机的位置估计和速度估计。记虚拟智能体0的估计值 $\hat{p}_0 = p_d, \hat{v}_0 = v_d$ 。

**假设3** 本文中,期望位置 $p_d$ 、期望速度 $v_d$ 及其各阶导数 $\dot{v}_d, \ddot{v}_d, \dddot{v}_d$ 都是有界的,并且存在一个非负常数 $\sigma_d$ 使得 $\|\dot{v}_d\|_2 \leq \sigma_d < g$ 。

**定理1** 令 $\alpha_1 > 0$ 且 $\alpha_2 > \sigma_d$ ,若假设1~假设3成立,则领航者无人机可以在有限时间内精确估计出期望位置和期望速度,并生成队形,即 $\hat{p}_i \rightarrow p_d + \delta_{i0}, \hat{v}_i \rightarrow v_d, i \in \mathcal{L}$ 。

**证明** 对于每一架领航者无人机,定义其位置估计误差为 $\tilde{p}_i = \hat{p}_i - p_d - \delta_{i0}$ ,速度估计误差为 $\tilde{v}_i = \hat{v}_i - v_d, i \in \mathcal{L}$ 。

记 $\tilde{p}_\mathcal{L} = [\tilde{p}_1^T, \tilde{p}_2^T, \dots, \tilde{p}_N^T]^T, \tilde{v}_\mathcal{L} = [\tilde{v}_1^T, \tilde{v}_2^T, \dots, \tilde{v}_N^T]^T$ ,选取以下李雅普诺夫函数:

$$V_1 = \frac{1}{2} \tilde{v}_\mathcal{L}^T (\mathbf{L}_\Pi \otimes \mathbf{I}_3)^T \tilde{v}_\mathcal{L} \quad (7)$$

$$V_2 = \frac{1}{2} \tilde{p}_\mathcal{L}^T (\mathbf{L}_\Pi \otimes \mathbf{I}_3)^T \tilde{p}_\mathcal{L} \quad (8)$$

由式(26)和引理1可得

$$\begin{aligned} \dot{V}_1 &= \tilde{v}_\mathcal{L}^T (\mathbf{L}_\Pi \otimes \mathbf{I}_3)^T \dot{\tilde{v}}_\mathcal{L} = \\ & [(\mathbf{L}_\Pi \otimes \mathbf{I}_3)^T \tilde{v}_\mathcal{L}]^T [-\alpha_2 \text{sgn}((\mathbf{L}_\Pi \otimes \mathbf{I}_3) \tilde{v}_\mathcal{L}) - (\mathbf{1}_N \otimes \mathbf{I}_3) \dot{v}_d] = \\ & [(\mathbf{L}_\Pi \otimes \mathbf{I}_3)^T \tilde{v}_\mathcal{L}]^T [-\alpha_2 \text{sgn}((\mathbf{L}_\Pi \otimes \mathbf{I}_3) \tilde{v}_\mathcal{L}) - (-\mathbf{L}_\Pi^{-1} \mathbf{L}_{d1} \otimes \mathbf{I}_3) \dot{v}_d] \end{aligned} \quad (9)$$

根据文献[27],对于任意向量 $x \in R^3$ 满足:

$$\|x\|_1 \geq \|x\|_2 \quad (10)$$

则有

$$\begin{aligned} \dot{V}_1 &\leq -(\alpha_2 - \|\mathbf{L}_\Pi^{-1} \mathbf{L}_{d1}\|_2 \sigma_d) \|(\mathbf{L}_\Pi \otimes \mathbf{I}_3)^T \tilde{v}_\mathcal{L}\|_1 \leq \\ & -(\alpha_2 - \sigma_d) \|(\mathbf{L}_\Pi \otimes \mathbf{I}_3)^T \tilde{v}_\mathcal{L}\|_2 \leq \\ & -(\alpha_2 - \sigma_d) \sqrt{\tilde{v}_\mathcal{L}^T (\mathbf{L}_\Pi \otimes \mathbf{I}_3)^2 \tilde{v}_\mathcal{L}} \leq \\ & -(\alpha_2 - \sigma_d) \lambda_{\min}(\mathbf{L}_\Pi) \|\tilde{v}_\mathcal{L}\|_2 \leq \end{aligned}$$

$$-\sqrt{2}(\alpha_2 - \sigma_d) \frac{\lambda_{\min}(\mathbf{L}_{\parallel})}{\sqrt{\lambda_{\max}(\mathbf{L}_{\parallel})}} V_1^{1/2} \quad (11)$$

**引理 2**<sup>[28]</sup> 若存在正实数  $c > 0$  和  $\beta \in (0, 1)$  使得连续可微的正定李雅普诺夫函数  $V(x)$  满足:

$$\dot{V}(x) + cV(x)^\beta \leq 0 \quad (12)$$

则系统在有限时间内稳定,收敛时间为

$$T(x_0) \leq \frac{V(x_0)^{1-\beta}}{c(1-\beta)} \quad (13)$$

故当时间  $t \geq t_1$  时,有  $\hat{\mathbf{v}}_i \rightarrow \mathbf{v}_d$ , 对于  $i \in \mathcal{L}$  成立。 $t_1$  的表达式为

$$t_1 = \frac{\sqrt{2}\sqrt{\lambda_{\max}(\mathbf{L}_{\parallel})}}{(\alpha_2 - \sigma_d)\lambda_{\min}(\mathbf{L}_{\parallel})} V_1^{1/2}(0) \quad (14)$$

当  $t \geq t_1$  时,对式(8)求导可得

$$\begin{aligned} \dot{V}_2 &= \tilde{\mathbf{p}}_{\mathcal{L}}^T (\mathbf{L}_{\parallel} \otimes \mathbf{I}_3)^T [ -\alpha_1 \text{sgn}((\mathbf{L}_{\parallel} \otimes \mathbf{I}_3) \tilde{\mathbf{p}}_{\mathcal{L}}) ] \leq \\ &- \alpha_1 \|(\mathbf{L}_{\parallel} \otimes \mathbf{I}_3) \tilde{\mathbf{p}}_{\mathcal{L}}\|_1 \leq \\ &- \alpha_1 \|(\mathbf{L}_{\parallel} \otimes \mathbf{I}_3) \tilde{\mathbf{p}}_{\mathcal{L}}\|_2 \leq \\ &- \alpha_1 \sqrt{\tilde{\mathbf{p}}_{\mathcal{L}}^T (\mathbf{L}_{\parallel} \otimes \mathbf{I}_3)^2 \tilde{\mathbf{p}}_{\mathcal{L}}} \leq \\ &- \alpha_1 \lambda_{\min}(\mathbf{L}_{\parallel}) \| \tilde{\mathbf{p}}_{\mathcal{L}} \|_2 \leq \\ &-\sqrt{2} \alpha_1 \frac{\lambda_{\min}(\mathbf{L}_{\parallel})}{\sqrt{\lambda_{\max}(\mathbf{L}_{\parallel})}} V_2^{1/2} \end{aligned} \quad (15)$$

故当时间  $t \geq t_2$  时,有  $\hat{\mathbf{p}}_i \rightarrow \mathbf{p}_d + \delta_{i0}$ , 对于  $i \in \mathcal{L}$  成立。 $t_2$  的表达式为

$$t_2 = t_1 + \frac{\sqrt{2}\sqrt{\lambda_{\max}(\mathbf{L}_{\parallel})}}{\alpha_1 \lambda_{\min}(\mathbf{L}_{\parallel})} V_2^{1/2}(t_1) \quad (16)$$

因此,当时间  $t \geq t_2$  时,对于每个领航者而言,其期望位置和期望速度可以用生成的估计位置  $\hat{\mathbf{p}}_i$  和估计速度  $\hat{\mathbf{v}}_i$  来替代。

证毕

### 2.2 跟随者层

为使所有跟随者无人机的估计位置位于领航者无人机估计位置构成的凸包里,且估计速度与领航者收敛一致,对于第  $i$  架跟随者无人机,设计如下有限时间滑模估计器:

$$\begin{cases} \dot{\hat{\mathbf{p}}}_i = \hat{\mathbf{v}}_i - \alpha_3 \text{sgn} \sum_{j=1}^{N+M} m_{ij} (\hat{\mathbf{p}}_i - \hat{\mathbf{p}}_j) \\ \dot{\hat{\mathbf{v}}}_i = -\alpha_4 \text{sgn} \sum_{j=1}^{N+M} m_{ij} (\hat{\mathbf{v}}_i - \hat{\mathbf{v}}_j) \end{cases} \quad (17)$$

式中:  $\alpha_3$  和  $\alpha_4$  为正常数;  $\hat{\mathbf{p}}_i$  和  $\hat{\mathbf{v}}_i$  分别为第  $i$  架跟随者无人机的位置估计和速度估计,  $i \in \mathcal{F}$ 。

令  $\hat{\mathbf{p}}_{\mathcal{L}} = [\hat{\mathbf{p}}_1^T, \hat{\mathbf{p}}_2^T, \dots, \hat{\mathbf{p}}_N^T]^T$  表示领航者估计位置,  $\hat{\mathbf{p}}_{\mathcal{F}} = [\hat{\mathbf{p}}_{N+1}^T, \hat{\mathbf{p}}_{N+2}^T, \dots, \hat{\mathbf{p}}_{N+M}^T]^T$  表示跟随者估计位置。定义一个位置向量为

$$\mathbf{h}_{\mathcal{F}} = [\mathbf{h}_{N+1}, \mathbf{h}_{N+2}, \dots, \mathbf{h}_{N+M}]^T = -(\mathbf{L}_{\text{ff}}^{-1} \mathbf{L}_{\text{ff}} \otimes \mathbf{I}_3) \hat{\mathbf{p}}_{\mathcal{L}} \quad (18)$$

由引理 1 和文献[25]可知,位置向量  $\mathbf{h}_{\mathcal{F}}$  的含义是凸包  $\text{Co}(\hat{\mathbf{p}}_{\mathcal{L}})$  的内部位置集合。

**定理 2** 令  $\alpha_3 > 0$  且  $\alpha_4 > \sigma_d$ , 若假设 1 ~ 假设 3 成立,则跟随者无人机的估计位置可以在有限时间内分布于领航者无人机估计位置构成的凸包里,即  $\hat{\mathbf{p}}_i \rightarrow \mathbf{h}_i$ 。同理可求得跟随者无人机的估计速度,即  $\hat{\mathbf{v}}_i \rightarrow \dot{\mathbf{h}}_i, i \in \mathcal{F}$ 。

**证明** 对于每一架跟随者无人机,定义其位置估计误差为  $\tilde{\hat{\mathbf{p}}}_i = \hat{\mathbf{p}}_i - \mathbf{h}_i$ , 速度估计误差为  $\tilde{\hat{\mathbf{v}}}_i = \hat{\mathbf{v}}_i - \dot{\mathbf{h}}_i, i \in \mathcal{L}$ 。

记  $\tilde{\hat{\mathbf{p}}}_{\mathcal{F}} = [\tilde{\hat{\mathbf{p}}}_{N+1}^T, \tilde{\hat{\mathbf{p}}}_{N+2}^T, \dots, \tilde{\hat{\mathbf{p}}}_{N+M}^T]^T, \tilde{\hat{\mathbf{v}}}_{\mathcal{F}} = [\tilde{\hat{\mathbf{v}}}_{N+1}^T, \tilde{\hat{\mathbf{v}}}_{N+2}^T, \dots, \tilde{\hat{\mathbf{v}}}_{N+M}^T]^T$ , 选取以下李雅普诺夫函数:

$$V_3 = \frac{1}{2} \tilde{\hat{\mathbf{v}}}_{\mathcal{F}}^T (\mathbf{L}_{\text{ff}} \otimes \mathbf{I}_3)^T \tilde{\hat{\mathbf{v}}}_{\mathcal{F}} \quad (19)$$

$$V_4 = \frac{1}{2} \tilde{\hat{\mathbf{p}}}_{\mathcal{F}}^T (\mathbf{L}_{\text{ff}} \otimes \mathbf{I}_3)^T \tilde{\hat{\mathbf{p}}}_{\mathcal{F}} \quad (20)$$

与式(9)~式(16)的证明类似,可得,当时间  $t \geq t_3$  时,有  $\hat{\mathbf{v}}_i \rightarrow \dot{\mathbf{h}}_i$  对于  $i \in \mathcal{F}$  成立。 $t_3$  的表达式为

$$t_3 = \frac{\sqrt{2}\sqrt{\lambda_{\max}(\mathbf{L}_{\text{ff}})}}{(\alpha_4 - \sigma_d)\lambda_{\min}(\mathbf{L}_{\text{ff}})} V_3^{1/2}(0) \quad (21)$$

当  $t \geq t_4$  时,有  $\hat{\mathbf{p}}_i \rightarrow \mathbf{h}_i$  对于  $i \in \mathcal{F}$  成立。 $t_4$  的表达式为

$$t_4 = t_3 + \frac{\sqrt{2}\sqrt{\lambda_{\max}(\mathbf{L}_{\text{ff}})}}{\alpha_1 \lambda_{\min}(\mathbf{L}_{\text{ff}})} V_4^{1/2}(t_3) \quad (22)$$

注意到,当时间  $t \geq \max\{t_2, t_4\}$  时,对于  $i \in \mathcal{L}$ , 有  $\hat{\mathbf{p}}_i \rightarrow \mathbf{p}_d + \delta_{i0}, \hat{\mathbf{v}}_i \rightarrow \mathbf{v}_d$  成立; 对于  $i \in \mathcal{F}$ , 有  $\hat{\mathbf{p}}_i \rightarrow \mathbf{h}_i, \hat{\mathbf{v}}_i \rightarrow \dot{\mathbf{h}}_i \rightarrow \mathbf{v}_d$  成立。换句话说,每架无人机的期望位置和期望速度可以用生成的估计位置  $\hat{\mathbf{p}}_i$  和估计速度  $\hat{\mathbf{v}}_i$  替代。只要每架无人机的实际位置  $\mathbf{p}_i$  和实际速度  $\mathbf{v}_i$  分别对  $\hat{\mathbf{p}}_i$  和  $\hat{\mathbf{v}}_i$  完成了跟踪,则该多无人机系统即可形成编队合围控制。证毕

## 3 估计值分布式跟踪控制

### 3.1 位置控制层

本节中,每架无人机都对其估计位置  $\hat{\mathbf{p}}_i$  和估计速度  $\hat{\mathbf{v}}_i$  进行跟踪。由第 2 节可知,当时间  $t \geq \max\{t_2, t_4\}$  时,每架无人机就相当于对期望位置和期望速度进行跟踪。

将式(3)中第  $i$  架四旋翼无人机的加速度重写为

$$\dot{\mathbf{v}}_i = \mathbf{A}_i - \frac{T_i}{m_i} (\mathbf{R}(\mathbf{Q}_i)^T - \mathbf{R}(\mathbf{Q}_{i,d})^T) \hat{\mathbf{e}}_3 \quad (23)$$

式中:  $\mathbf{A}_i$  为中间控制输入, 定义为

$$\mathbf{A}_i = g\hat{\mathbf{e}}_3 - \frac{T_i}{m_i}\mathbf{R}(\mathbf{Q}_{i,d})^T\hat{\mathbf{e}}_3 \quad (24)$$

其中:  $\mathbf{Q}_{i,d} = [\mathbf{q}_{i,d}^T, \boldsymbol{\eta}_{i,d}^T]^T$  为无人机期望姿态的单位四元数。

**引理 3**<sup>[29]</sup> 考虑中间控制输入  $\mathbf{A}_i = [\Lambda_{i,x}, \Lambda_{i,y}, \Lambda_{i,z}]^T$  满足:

$$\mathbf{A}_i \neq [0, 0, \epsilon_1]^T \quad \epsilon_1 \geq g \quad (25)$$

则从式(24)中提取出第  $i$  架四旋翼无人机的推力输入大小:

$$T_i = m_i \sqrt{\Lambda_{i,x}^2 + \Lambda_{i,y}^2 + (g - \Lambda_{i,z})^2} \quad (26)$$

期望姿态:

$$\begin{cases} \mathbf{q}_{i,d} = \frac{m_i}{2T_i\eta_{i,d}}[\Lambda_{i,y}, -\Lambda_{i,x}, 0]^T \\ \eta_{i,d} = \sqrt{0.5[1 + m_i(g - \Lambda_{i,z})/T_i]} \end{cases} \quad (27)$$

若  $\mathbf{A}_i$  可微, 则可得到第  $i$  架四旋翼无人机的期望角速度为

$$\boldsymbol{\omega}_{i,d} = \mathbf{H}(\mathbf{A}_i)\dot{\mathbf{A}}_i \quad (28)$$

式中:

$$\mathbf{H}(\mathbf{A}_i) = \frac{1}{\lambda_{i,1}^2\lambda_{i,2}} \cdot \begin{bmatrix} -\Lambda_{i,x}\Lambda_{i,y} & -\Lambda_{i,y}^2 + \lambda_{i,1}\lambda_{i,2} & \lambda_{i,2}\Lambda_{i,y} \\ \Lambda_{i,x}^2 - \lambda_{i,1}\lambda_{i,2} & \Lambda_{i,x}\Lambda_{i,y} & -\Lambda_{i,x}\Lambda_{i,y} \\ \lambda_{i,1}\Lambda_{i,y} & -\lambda_{i,1}\Lambda_{i,x} & 0 \end{bmatrix} \quad (29)$$

$$\lambda_{i,1} = T_i/m_i \quad (30)$$

$$\lambda_{i,2} = \lambda_{i,1} + g - \Lambda_{i,z} \quad (31)$$

对于  $i \in \mathcal{L} \cup \mathcal{F}$ , 定义第  $i$  架无人机的位置跟踪误差  $\tilde{\mathbf{p}}_i = \mathbf{p}_i - \hat{\mathbf{p}}_i$ , 速度跟踪误差  $\tilde{\mathbf{v}}_i = \mathbf{v}_i - \hat{\mathbf{v}}_i$ , 再定义新的误差变量为

$$\boldsymbol{\zeta}_i = \tilde{\mathbf{p}}_i - \boldsymbol{\Xi}_i \quad (32)$$

$$\mathbf{z}_i = \tilde{\mathbf{v}}_i - \dot{\boldsymbol{\Xi}}_i \quad (33)$$

式中:  $\boldsymbol{\Xi}_i$  为一个中间变量,  $\boldsymbol{\Xi}_i$  及其导数  $\dot{\boldsymbol{\Xi}}_i$  可以任意初始化。

将式(23)中的中间控制输入设计为

$$\mathbf{A}_i = \dot{\tilde{\mathbf{v}}}_i - k_{\Xi}\tanh(\boldsymbol{\Xi}_i) - k_d\tanh(\dot{\boldsymbol{\Xi}}_i) \quad (34)$$

$$\ddot{\boldsymbol{\Xi}}_i = \mathbf{A}_i - \mathbf{u}_i - \dot{\tilde{\mathbf{v}}}_i \quad (35)$$

式中:  $k_{\Xi}$  和  $k_d$  为正常数增益;  $\mathbf{u}_i$  为待求辅助输入。

注意到, 需确保式(34)的中间控制输入  $\mathbf{A}_i$  有界, 即满足以下不等式:

$$\|\mathbf{A}_i\|_2 \leq \sigma_d + \sqrt{3}(k_{\Xi} + k_d) \quad (36)$$

则需要设计合适的  $k_{\Xi}$  和  $k_d$  使得下列约束条件

成立:

$$\sqrt{3}(k_{\Xi} + k_d) < g - \epsilon_2 \quad (37)$$

式中: 常数  $\epsilon_2$  满足

$$\sigma_d \leq \epsilon_2 < g \quad (38)$$

如此便可使得  $\|\mathbf{A}_i\|_2 < g$ , 即满足了引理 3 的条件, 且推力输入  $T_i$  也有界, 即

$$T_i \leq m_i[g + \sigma_d + \sqrt{3}(k_{\Xi} + k_d)] = \bar{T}_i \quad (39)$$

根据文献[30], 如果  $\mathbf{u}_i$  全局有界且渐近收敛于  $\mathbf{0}$ , 那么变量  $\boldsymbol{\Xi}_i$  及其导数  $\dot{\boldsymbol{\Xi}}_i$  有界且渐近收敛于  $\mathbf{0}$ , 因此, 只要设计待求辅助输入  $\mathbf{u}_i$  使  $\boldsymbol{\zeta}_i, \mathbf{z}_i \rightarrow \mathbf{0}$ , 则可实现  $\tilde{\mathbf{p}}_i, \tilde{\mathbf{v}}_i \rightarrow \mathbf{0}$ , 即实现对估计值的跟踪。

设计辅助输入为

$$\mathbf{u}_i = \begin{cases} -k_z\mathbf{z}_i - k_{\zeta}\boldsymbol{\zeta}_i - k_s \sum_{j=1}^N m_{ij}(\boldsymbol{\zeta}_i - \boldsymbol{\zeta}_j) & i \in \mathcal{L} \\ -k_z\mathbf{z}_i - k_{\zeta}\boldsymbol{\zeta}_i - k_s \sum_{j=N+1}^{N+M} m_{ij}(\boldsymbol{\zeta}_i - \boldsymbol{\zeta}_j) & i \in \mathcal{F} \end{cases} \quad (40)$$

式中:  $k_z, k_{\zeta}$  和  $k_s$  为正常数增益。

### 3.2 姿态控制层

对于  $i \in \mathcal{L} \cup \mathcal{F}$ , 考虑式(27)给出的期望姿态的提取值  $\mathbf{Q}_{i,d}$ , 定义第  $i$  架无人机的姿态跟踪误差  $\tilde{\mathbf{Q}}_i = [\tilde{\mathbf{q}}_i^T, \tilde{\eta}_i]^T = \mathbf{Q}_{i,d}^{-1} \odot \mathbf{Q}_i$ , 由文献[24]可知, 姿态误差  $\tilde{\mathbf{Q}}_i = [0, 0, 0, 1]^T$  表示精确的姿态跟踪。由式(4)可得

$$\dot{\tilde{\mathbf{Q}}}_i = \frac{1}{2}\mathbf{G}(\tilde{\mathbf{Q}}_i)\tilde{\boldsymbol{\omega}}_i \quad (41)$$

式中:

$$\tilde{\boldsymbol{\omega}}_i = \boldsymbol{\omega}_i - \mathbf{R}(\tilde{\mathbf{Q}}_i)\boldsymbol{\omega}_{i,d} \quad (42)$$

$$\mathbf{R}(\tilde{\mathbf{Q}}_i) = \mathbf{R}(\mathbf{Q}_i)\mathbf{R}(\mathbf{Q}_{i,d})^T \quad (43)$$

其中:  $\tilde{\boldsymbol{\omega}}_i$  为第  $i$  架无人机的角速度跟踪误差。

对式(28)求导得

$$\dot{\boldsymbol{\omega}}_{i,d} = \mathbf{H}(\mathbf{A}_i, \dot{\mathbf{A}}_i)\dot{\mathbf{A}}_i + \mathbf{H}(\mathbf{A}_i)\ddot{\mathbf{A}}_i \quad (44)$$

式中: 中间控制输入  $\mathbf{A}_i$  的导数为

$$\begin{aligned} \dot{\mathbf{A}}_i &= \dot{\tilde{\mathbf{v}}}_i - k_{\Xi}(1 - \tanh^2(\boldsymbol{\Xi}_i))\dot{\boldsymbol{\Xi}}_i - \\ &\quad k_d(1 - \tanh^2(\dot{\boldsymbol{\Xi}}_i))\ddot{\boldsymbol{\Xi}}_i \\ \ddot{\mathbf{A}}_i &= \dot{\tilde{\mathbf{v}}}_i - 2k_{\Xi}\tanh(\boldsymbol{\Xi}_i)(\tanh^2(\boldsymbol{\Xi}_i) - 1)\dot{\boldsymbol{\Xi}}_i - \\ &\quad (k_{\Xi}(1 - \tanh^2(\boldsymbol{\Xi}_i)) + 2k_d\tanh(\dot{\boldsymbol{\Xi}}_i)) \cdot \\ &\quad (\tanh^2(\dot{\boldsymbol{\Xi}}_i) - 1)\ddot{\boldsymbol{\Xi}}_i - k_d(1 - \tanh^2(\dot{\boldsymbol{\Xi}}_i)) \cdot \\ &\quad (\dot{\mathbf{A}}_i - \dot{\mathbf{u}}_i - \dot{\tilde{\mathbf{v}}}_i) \end{aligned} \quad (46)$$

当时间  $t \geq \max\{t_2, t_4\}$  时, 对于  $i \in \mathcal{L} \cup \mathcal{F}$ , 有  $\hat{\mathbf{v}}_i \rightarrow \mathbf{v}_{d,i}$  成立, 则式(34)、式(45)和式(46)中的  $\dot{\tilde{\mathbf{v}}}_i \rightarrow \dot{\mathbf{v}}_d, \ddot{\tilde{\mathbf{v}}}_i \rightarrow \ddot{\mathbf{v}}_d, \ddot{\boldsymbol{\Xi}}_i \rightarrow \ddot{\mathbf{v}}_d$ 。为了防止微分爆炸, 具体求

解过程在 3.3 节中描述。

定义矩阵  $\Phi_i$  满足:

$$\Phi_i \cdot \tilde{q}_i = (\mathbf{R}(\mathbf{Q}_i)^T - \mathbf{R}(\mathbf{Q}_{i,d})^T) \hat{e}_3 \quad (47)$$

则有

$$\Phi_i = 2\mathbf{R}(\mathbf{Q}_i)^T \begin{bmatrix} 0 & \tilde{\eta}_i & -\tilde{q}_{i,1} \\ -\tilde{\eta}_i & 0 & -\tilde{q}_{i,2} \\ \tilde{q}_{i,1} & \tilde{q}_{i,2} & 0 \end{bmatrix} \quad (48)$$

式中:  $[\tilde{q}_{i,1}, \tilde{q}_{i,2}, \tilde{q}_{i,3}]^T = \tilde{q}_i$  为姿态跟踪误差四元数  $\tilde{\mathbf{Q}}_i$  的向量部分。

由式(3)、式(33)、式(35)和式(47)可得

$$\dot{\mathbf{z}}_i = \mathbf{u}_i - \frac{T_i}{m_i} \Phi_i \cdot \tilde{q}_i \quad (49)$$

根据式(32)、式(33)、式(40)、式(49)则可求得式(46)中的  $\dot{\mathbf{u}}_i$ 。

定义函数:

$$\Theta_i(\cdot) = \mathbf{S}(\omega_i) \mathbf{J}_i \omega_i - \mathbf{J}_i \mathbf{S}(\tilde{\omega}_i) \cdot$$

$$\mathbf{R}(\tilde{\mathbf{Q}}_i) \omega_{i,d} + \mathbf{J}_i \mathbf{R}(\tilde{\mathbf{Q}}_i) \dot{\omega}_{i,d} \quad (50)$$

则第  $i$  架四旋翼无人机的力矩输入表示为

$$\boldsymbol{\tau}_i = \Theta_i(\cdot) + \mathbf{J}_i \dot{\boldsymbol{\beta}}_i - k_q \tilde{q}_i - k_\psi (\tilde{\omega}_i - \boldsymbol{\beta}_i) \quad (51)$$

式中:

$$\boldsymbol{\beta}_i = -k_\beta \tilde{q}_i + \frac{T_i}{k_q m_i} \Phi_i^T \mathbf{z}_i \quad (52)$$

$$\dot{\boldsymbol{\beta}}_i = -\frac{k_\beta}{2} (\tilde{\eta}_i \mathbf{I}_3 + \mathbf{S}(\tilde{q}_i)) \tilde{\omega}_i + \frac{T_i}{k_q m_i} \cdot \frac{d(\Phi_i^T \mathbf{z}_i)}{dt} + \frac{m_i}{k_q T_i} (g \hat{e}_3 - \Lambda_i)^T \dot{\Lambda}_i \Phi_i^T \mathbf{z}_i \quad (53)$$

式中:  $k_q, k_\psi$  和  $k_\beta$  为正常数增益。

### 3.3 高阶导数逼近

设计以下高阶导数逼近算法:

$$\begin{cases} \dot{\boldsymbol{\chi}}_{0,i} = -\frac{\boldsymbol{\chi}_{0,i} - \hat{\mathbf{v}}_i}{\zeta_0} - \frac{\xi_0(\boldsymbol{\chi}_{0,i} - \hat{\mathbf{v}}_i)}{\sqrt{\|\boldsymbol{\chi}_{0,i} - \hat{\mathbf{v}}_i\|_2 + \Delta_0}} \\ \dot{\boldsymbol{\chi}}_{1,i} = -\frac{\boldsymbol{\chi}_{1,i} - \dot{\boldsymbol{\chi}}_{0,i}}{\zeta_1} - \frac{\xi_1(\boldsymbol{\chi}_{1,i} - \dot{\boldsymbol{\chi}}_{0,i})}{\sqrt{\|\boldsymbol{\chi}_{1,i} - \dot{\boldsymbol{\chi}}_{0,i}\|_2 + \Delta_1}} \\ \dot{\boldsymbol{\chi}}_{2,i} = -\frac{\boldsymbol{\chi}_{2,i} - \dot{\boldsymbol{\chi}}_{1,i}}{\zeta_2} - \frac{\xi_2(\boldsymbol{\chi}_{2,i} - \dot{\boldsymbol{\chi}}_{1,i})}{\sqrt{\|\boldsymbol{\chi}_{2,i} - \dot{\boldsymbol{\chi}}_{1,i}\|_2 + \Delta_2}} \\ \dot{\boldsymbol{\chi}}_{3,i} = -\frac{\boldsymbol{\chi}_{3,i} - \dot{\boldsymbol{\chi}}_{2,i}}{\zeta_3} - \frac{\xi_3(\boldsymbol{\chi}_{3,i} - \dot{\boldsymbol{\chi}}_{2,i})}{\sqrt{\|\boldsymbol{\chi}_{3,i} - \dot{\boldsymbol{\chi}}_{2,i}\|_2 + \Delta_3}} \end{cases} \quad (54)$$

式中:  $\zeta_0, \zeta_1, \zeta_2, \zeta_3$  为正常数;  $\xi_0, \xi_1, \xi_2, \xi_3$  及  $\Delta_0, \Delta_1, \Delta_2, \Delta_3$  为待设计常数。对于  $i \in \mathcal{L} \cup \mathcal{F}, \boldsymbol{\chi}_{0,i}, \boldsymbol{\chi}_{1,i}, \boldsymbol{\chi}_{2,i}, \boldsymbol{\chi}_{3,i}$  分别为第  $i$  架无人机对其估计速度  $\hat{\mathbf{v}}_i$  的 0~3 阶导数的逼近值。

**定理 3** 令  $\Delta_r \ll 0$  且  $\zeta_r > 0$ , 选取合适的正参数  $\xi_r$ , 在假设 3 成立的情况下, 每架无人机生成的

估计速度  $\hat{\mathbf{v}}_i$  的 0~3 阶导数可以在时间  $t \geq \max\{t_2, t_4\}$  之后的有限时间内分别近似逼近于  $\boldsymbol{\chi}_{r,i}, r=0, 1, 2, 3$ 。

**证明** 定义逼近误差为

$$\begin{cases} \mathbf{e}_{0,i} = \boldsymbol{\chi}_{0,i} - \hat{\mathbf{v}}_i \\ \mathbf{e}_{1,i} = \boldsymbol{\chi}_{1,i} - \dot{\boldsymbol{\chi}}_{0,i} \\ \mathbf{e}_{2,i} = \boldsymbol{\chi}_{2,i} - \dot{\boldsymbol{\chi}}_{1,i} \\ \mathbf{e}_{3,i} = \boldsymbol{\chi}_{3,i} - \dot{\boldsymbol{\chi}}_{2,i} \end{cases} \quad (55)$$

选取以下李雅普诺夫函数:

$$V_5 = \sum_{i=1}^{N+M} \sum_{r=0}^3 W_{r,i} = \sum_{i=1}^{N+M} \sum_{r=0}^3 \frac{1}{2} \mathbf{e}_{r,i}^T \mathbf{e}_{r,i} \quad (56)$$

则有

$$\begin{aligned} \dot{W}_{0,i} &= \mathbf{e}_{0,i}^T \dot{\mathbf{e}}_{0,i} = \\ &= -\frac{\mathbf{e}_{0,i}^T \mathbf{e}_{0,i}}{\zeta_0} - \frac{\xi_0 \mathbf{e}_{0,i}^T \mathbf{e}_{0,i}}{\|\mathbf{e}_{0,i}\|_2^{1/2} + \Delta_0} - \mathbf{e}_{0,i}^T \hat{\mathbf{v}}_i = \\ &= -\frac{\mathbf{e}_{0,i}^T \mathbf{e}_{0,i}}{\zeta_0} - \frac{\xi_0}{\|\mathbf{e}_{0,i}\|_2^{1/2} + \Delta_0} [(\|\mathbf{e}_{0,i}\|_2 + \Delta_0^2) \cdot \\ &(\|\mathbf{e}_{0,i}\|_2^{1/2} + \Delta_0)(\|\mathbf{e}_{0,i}\|_2 - \Delta_0) + \Delta_0^4] - \mathbf{e}_{0,i}^T \hat{\mathbf{v}}_i \leq \\ &= -\xi_0 \|\mathbf{e}_{0,i}\|_2^{3/2} + \xi_0 \Delta_0 \|\mathbf{e}_{0,i}\|_2 + \xi_0 \Delta_0^3 + \sigma_d \|\mathbf{e}_{0,i}\|_2 \end{aligned} \quad (57)$$

选取  $\Delta_0 \ll 0$  使  $\xi_0 \Delta_0^3 \ll 0$  成立, 令  $\gamma_0 = \xi_0 / \sigma_d$ ,

则有

$$\begin{aligned} \dot{W}_{0,i} &\leq -\sigma_d (\gamma_0 \|\mathbf{e}_{0,i}\|_2^{3/2} - \gamma_0 \Delta_0 \|\mathbf{e}_{0,i}\|_2 - \|\mathbf{e}_{0,i}\|_2) \leq \\ &= -\sigma_d \left( \gamma_0 - \frac{\gamma_0 \Delta_0 + 1}{W_{0,i}^{1/4}} \right) W_{0,i}^{3/4} \end{aligned} \quad (58)$$

选取正常数  $\gamma_0$ , 使

$$\gamma_0 - \frac{\gamma_0 \Delta_0 + 1}{W_{0,i}^{1/4}} > 0 \quad (59)$$

成立, 则能保证  $W_{0,i}$  的有限时间性能。

同理可得

$$\dot{W}_{r,i} \leq -\sigma_{r,d} \left( \gamma_r - \frac{\gamma_r \Delta_r + 1}{W_{r,i}^{1/4}} \right) W_{r,i}^{3/4} \quad r = 1, 2, 3 \quad (60)$$

成立。其中,  $\sigma_{r,d} \geq \|\dot{\boldsymbol{\chi}}_{r,i}\|_2$  为已知正常数, 设计参数  $\gamma_r$  为  $\gamma_r = \xi_r / \sigma_{r,d}, r = 1, 2, 3$ 。

**引理 4** 对于任意向量  $\mathbf{l}(x) = [l_1, l_2, \dots, l_n] \in R^n$  及常数  $0 < \Omega \leq 1$ , 有

$$\begin{aligned} (|l_1| + |l_2| + \dots + |l_n|)^\Omega &\leq |l_1|^\Omega + \\ &|l_2|^\Omega + \dots + |l_n|^\Omega \end{aligned} \quad (61)$$

则有

$$\dot{V}_5 = \sum_{i=1}^{N+M} \sum_{r=0}^3 \dot{W}_{r,i} \leq -\gamma \left( \sum_{i=1}^{N+M} \sum_{r=0}^3 W_{r,i} \right)^{\frac{3}{4}} = -\gamma V_5^{3/4} \quad (62)$$

式中: 参数  $\gamma$  表达式为

$$\gamma = \min \left\{ \sigma_d \left( \gamma_0 - \frac{\gamma_0 \Delta_0 + 1}{W_{0,i}^{1/4}} \right), \sigma_{1,d} \left( \gamma_1 - \frac{\gamma_1 \Delta_1 + 1}{W_{1,i}^{1/4}} \right), \right. \\ \left. \sigma_{2,d} \left( \gamma_2 - \frac{\gamma_2 \Delta_2 + 1}{W_{2,i}^{1/4}} \right), \sigma_{3,d} \left( \gamma_3 - \frac{\gamma_3 \Delta_3 + 1}{W_{3,i}^{1/4}} \right) \right\} \quad (63)$$

则当时间  $t \geq t_5$  时, 有  $\chi_{r,i} \rightarrow \hat{v}_i^{(r)} \rightarrow v_d^{(r)}$  ( $r = 0, 1, 2, 3$ ) 对于  $i \in \mathcal{L} \cup \mathcal{F}$  成立, 即  $\chi_{1,i}, \chi_{2,i}, \chi_{3,i}$  可以代替式(34)、式(45)和式(46)中的  $\hat{v}_i, \ddot{v}_i$  和  $\ddot{v}_i$ 。  $t_5$  的表达式为

$$t_5 = \max \{ t_2, t_4 \} + \frac{4}{\gamma} V_5^{1/4} (\max \{ t_2, t_4 \}) \quad (64)$$

证毕

本节提出的高阶导数逼近算法求得的无人机估计速度的 0 ~ 3 阶导数的逼近值是用于替代式(34)、式(45)、式(46)中的高阶速度信息, 以进行式(28)、式(44)的计算。在无人机的姿态解算过程中, 常常需要求解变量的微分甚至高阶微分, 直接求解微分除了难以计算之外, 还存在难以避免的噪声问题, 若直接用差分法求解微分容易导致系统崩溃。在实际系统中, 如车辆的运动控制、航空航天控制、机器人控制也常需要求解信号的微分<sup>[31]</sup>, 本文提出的算法可以视为一种求解信号微分的方法。

### 3.4 系统稳定性分析

**定理 4** 考虑由  $N$  个领航者和  $M$  个跟随者构成的四旋翼无人机群系统, 假设控制参数  $k_{\Xi}$  和  $k_d$  满足式(37)。第  $i$  架无人机的推力输入由式(26)给出, 力矩输入由式(51)给出。各无人机的通信拓扑满足假设 1 ~ 假设 3, 则从任意初始条件开始, 总有: 当  $t \geq t_5$  时, 对于  $i \in \mathcal{L}$ , 有  $p_i \rightarrow \hat{p}_i \rightarrow p_d + \delta_{i0}, v_i \rightarrow \hat{v}_i \rightarrow v_d$  成立; 对于  $i \in \mathcal{F}$ , 有  $p_i \rightarrow \hat{p}_i \rightarrow h_i, v_i \rightarrow \hat{v}_i \rightarrow \dot{h}_i \rightarrow v_d$  成立。

换句话说, 每架无人机都对其生成的估计值  $\hat{p}_i$  和  $\hat{v}_i$  完成了跟踪, 领航者无人机形成编队, 跟随者无人机位于领航者编队内部。另外, 每架无人机的姿态收敛, 即姿态四元数误差  $\tilde{Q}_i \rightarrow [0, 0, 0, 1]^T$ , 角速度误差  $\tilde{\omega}_i \rightarrow \mathbf{0}$ , 该多四旋翼无人机系统完成编队合围控制。

**证明** 定义  $\psi_i = \tilde{\omega}_i - \beta_i$ , 由式(2)、式(42)和式(43)可得

$$\dot{\psi}_i = \dot{\tilde{\omega}}_i - R(\tilde{Q}_i) \dot{\omega}_{i,d} + S(\tilde{\omega}_i) R(\tilde{Q}_i) \omega_{i,d} - \dot{\beta}_i \quad (65)$$

根据式(4)可得

$$J_i \dot{\psi}_i = \tau_i - \Theta_i(\cdot) - J_i \dot{\beta}_i \quad (66)$$

选取以下李雅普诺夫函数:

$$V_6 = \sum_{i=1}^N V_{6,i} \quad (67)$$

式中:

$$V_{6,i} = \frac{1}{4} k_s \sum_{j=1}^N m_{ij} (\varsigma_i - \varsigma_j)^T (\varsigma_i - \varsigma_j) + \frac{1}{2} z_i^T z_i + \frac{1}{2} k_{\zeta} \varsigma_i^T \varsigma_i + \frac{1}{2} \psi_i^T J_i \psi_i + k_q [\tilde{q}_i^T \tilde{q}_i + (1 - \tilde{\eta}_i)^2] \quad (68)$$

根据假设 1, 领航者之间的信息流是双向的, 即  $m_{ij} = -m_{ji}, i, j \in \mathcal{L}$ , 则有

$$\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N m_{ij} (z_i - z_j)^T (\varsigma_i - \varsigma_j) = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N m_{ij} z_i^T (\varsigma_i - \varsigma_j) - \frac{1}{2} \sum_{j=1}^N \sum_{i=1}^N m_{ji} z_i^T (\varsigma_j - \varsigma_i) = \sum_{i=1}^N \sum_{j=1}^N m_{ij} z_i^T (\varsigma_i - \varsigma_j) \quad (69)$$

则对式(67)求导可得

$$\dot{V}_6 = \frac{1}{2} k_s \sum_{i=1}^N \sum_{j=1}^N m_{ij} (z_i - z_j)^T (\varsigma_i - \varsigma_j) + \sum_{i=1}^N \left[ z_i^T \cdot \left( u_i - \frac{T_i}{m_i} \Phi_i \tilde{q}_i \right) \right] + k_{\zeta} \sum_{i=1}^N z_i^T \varsigma_i + \sum_{i=1}^N \psi_i^T (\tau_i - \Theta_i(\cdot) - J_i \dot{\beta}_i) + k_q \sum_{i=1}^N \tilde{q}_i^T (\psi_i + \beta_i) = k_s \sum_{i=1}^N \sum_{j=1}^N m_{ij} z_i^T (\varsigma_i - \varsigma_j) + \sum_{i=1}^N \left[ -k_z z_i^T z_i - k_{\zeta} z_i^T \varsigma_i - k_s \sum_{j=1}^N m_{ij} z_i^T (\varsigma_i - \varsigma_j) - \frac{T_i}{m_i} z_i^T \Phi_i \cdot \tilde{q}_i \right] + k_{\zeta} \sum_{i=1}^N z_i^T \varsigma_i + \sum_{i=1}^N \psi_i^T (-k_q \tilde{q}_i - k_{\psi} \psi_i) + k_q \sum_{i=1}^N \tilde{q}_i^T (\psi_i - k_{\beta} \tilde{q}_i + \frac{T_i}{k_q m_i} \Phi_i^T z_i) = \sum_{i=1}^N (-k_z z_i^T z_i - k_{\psi} \psi_i^T \psi_i - k_q k_{\beta} \tilde{q}_i^T \tilde{q}_i) \quad (70)$$

对于  $i \in \mathcal{L}$ , 显然  $\varsigma_i, z_i, \psi_i$  与  $\tilde{q}_i$  有界。根据式(35)和式(49)可知,  $\ddot{\Xi}_i$  与  $\dot{z}_i$  有界。根据式(39)和式(52)可知,  $\beta_i$  有界, 则  $\tilde{\omega}_i$  有界。根据式(4)和式(41)可知,  $\dot{\tilde{q}}_i$  有界。根据式(51)和式(66)可知,  $\dot{\psi}_i$  有界。因此,  $\dot{V}_6$  有界。根据 Barbalat 引理可知, 当  $t \rightarrow \infty$  时,  $\varsigma_i \rightarrow 0, z_i \rightarrow 0, \psi_i \rightarrow 0$  且  $\tilde{q}_i \rightarrow 0$ , 因此,  $\dot{V}_6$  半负定, 每架领航者无人机的姿态收敛, 根据式(32)和式(33), 领航者可以对其估计值实现跟踪。

对于  $i \in \mathcal{F}$ , 根据假设 1, 跟随者之间的信息流也是双向的, 与上述证明相似, 跟随者也可以对其

估计值实现跟踪,且姿态收敛。因此,该多四旋翼无人机系统完成了编队包围控制。 证毕

**注 1** 本文未考虑避障和避碰问题,在今后的研究中拟采用团队已有工作<sup>[32]</sup>中的人工势场算法来解决在复杂障碍环境中的四旋翼无人机编队问题。

### 4 仿真验证

不失一般性,考虑一个具有 6 架领航者无人机和 6 架跟随者无人机的平面型编队和一个具有 8 架领航者无人机和 8 架跟随者无人机的立方体编队,其通信拓扑如图 3 所示。图 4 给出了这 2 种编队的航迹。对于图 4(b)的立方体编队,每架无人机的控制参数和初始状态如表 1 所示。

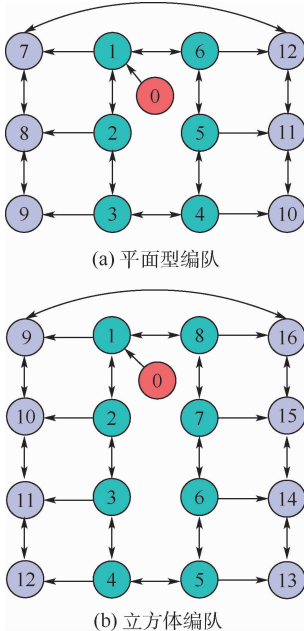


图 3 通信拓扑

Fig. 3 Communication topology

由图 4 可以看出,无论是平面型编队还是立方体编队,都可以完成控制目标。对于图 4(b)所示的立方体编队,仿真结果如图 5 ~ 图 11 所示。

图 5 与图 7 分别表明,领航者无人机约在 10 s 左右可以精确估计出期望的位置和速度,跟随者无人机则需要大约 13 s。这意味着在 13 s 以后,每架无人机对估计位置信息的跟踪等效于对期望位置信息的跟踪。图 6 与图 8 分别表明,领航者约在 14 s 左右可以精确跟踪到估计位置与估计速度,跟随者则需要大约 17 s。这意味着在 17 s

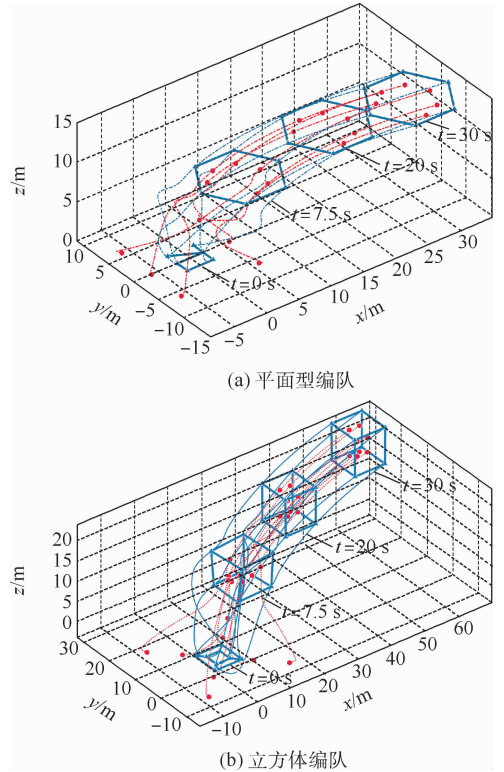


图 4 两种编队在空间中的航迹

Fig. 4 Trajectories of two type of formations in space

表 1 仿真条件

Table 1 Simulation conditions

控制参数	初始状态
$\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 8$	$p_1(0) = [-2 \ 2 \ 0]^T, p_2(0) = [2 \ 2 \ 0]^T$
$k_{\Xi} = k_d = 1.5$	$p_3(0) = [2 \ -2 \ 0]^T, p_4(0) = [-2 \ -2 \ 0]^T$
$k_z = k_C = 4$	$p_5(0) = [-4 \ 4 \ 0]^T, p_6(0) = [4 \ 4 \ 0]^T$
$k_s = 2$	$p_7(0) = [4 \ -4 \ 0]^T, p_8(0) = [-4 \ -4 \ 0]^T$
$k_q = k_{\beta} = 20$	$p_9(0) = [-6 \ 6 \ 0]^T, p_{10}(0) = [6 \ 6 \ 0]^T$
$k_{\psi} = 0.1$	$p_{11}(0) = [6 \ -6 \ 0]^T, p_{12}(0) = [-6 \ -6 \ 0]^T$
$\zeta_0 = \zeta_1 = \zeta_2 = \zeta_3 = 0.4$	$p_{13}(0) = [-12 \ 12 \ 0]^T, p_{14}(0) = [12 \ 12 \ 0]^T$
$\xi_0 = \xi_1 = \xi_2 = \xi_3 = 0.2$	$p_{15}(0) = [12 \ -12 \ 0]^T, p_{16}(0) = [-12 \ -12 \ 0]^T$
$\Delta_0 = \Delta_1 = \Delta_2 = \Delta_3 = -1$	$v_i(0) = [0 \ 0 \ 0]^T, i \in 1, 2, \dots, 16$
	$\omega_i(0) = [0 \ 0 \ 0]^T, i \in 1, 2, \dots, 16$
	$Q_i(0) = [0 \ 0 \ 0 \ 1]^T, i \in 1, 2, \dots, 16$

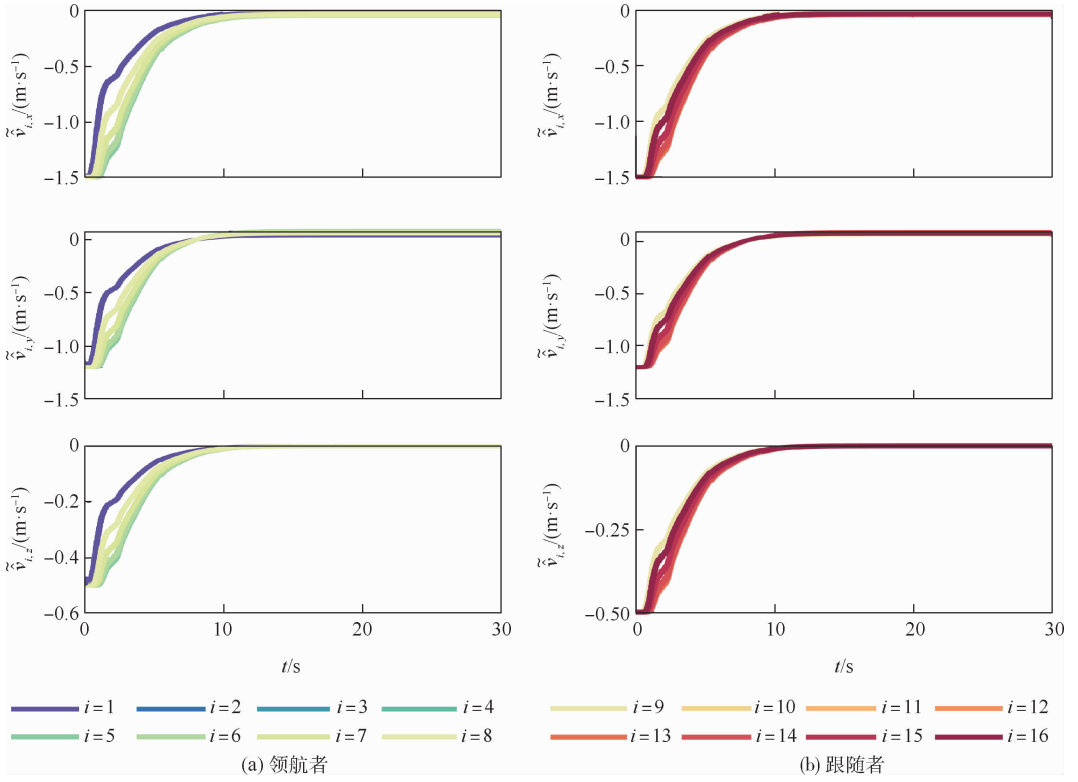


图 5 领航者与跟随者的期望速度估计误差

Fig. 5 Estimated errors of leader and follower desired velocities

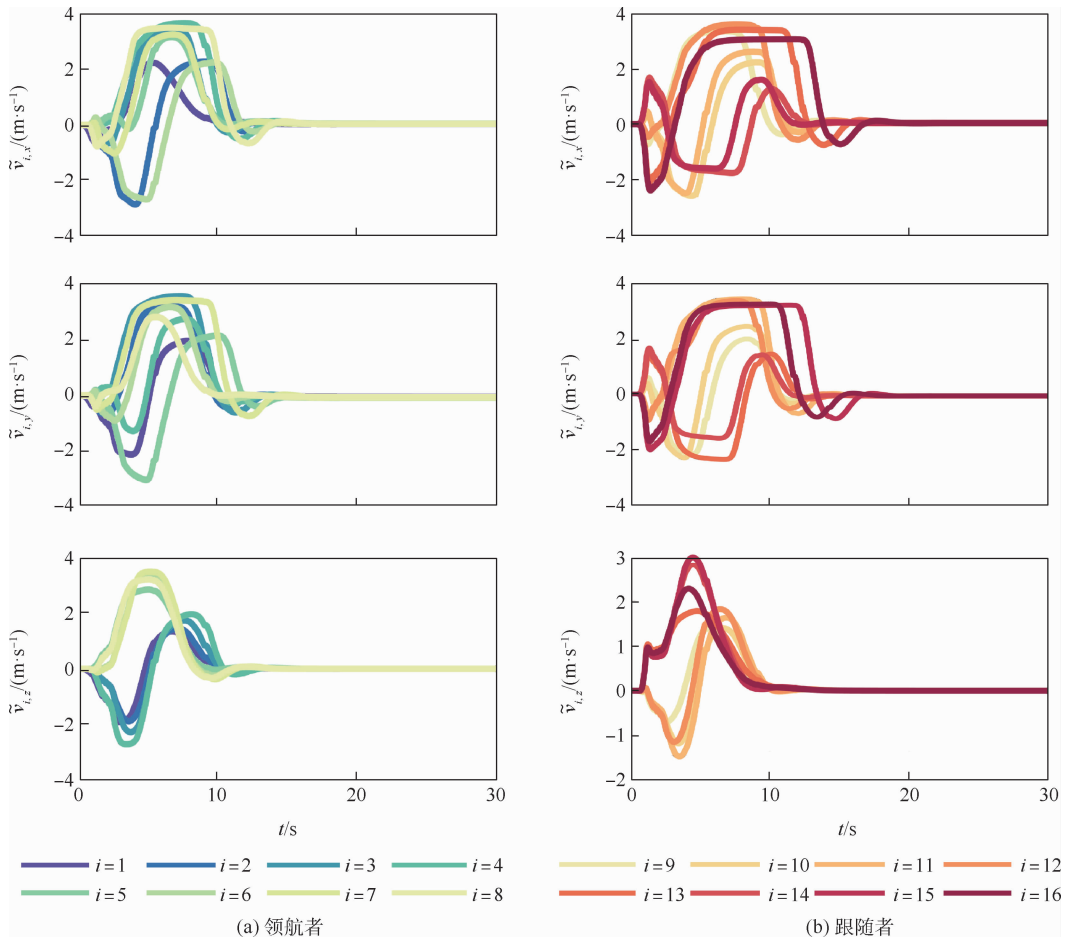


图 6 领航者与跟随者的估计速度跟踪误差

Fig. 6 Tracking errors of leader and follower estimated velocities

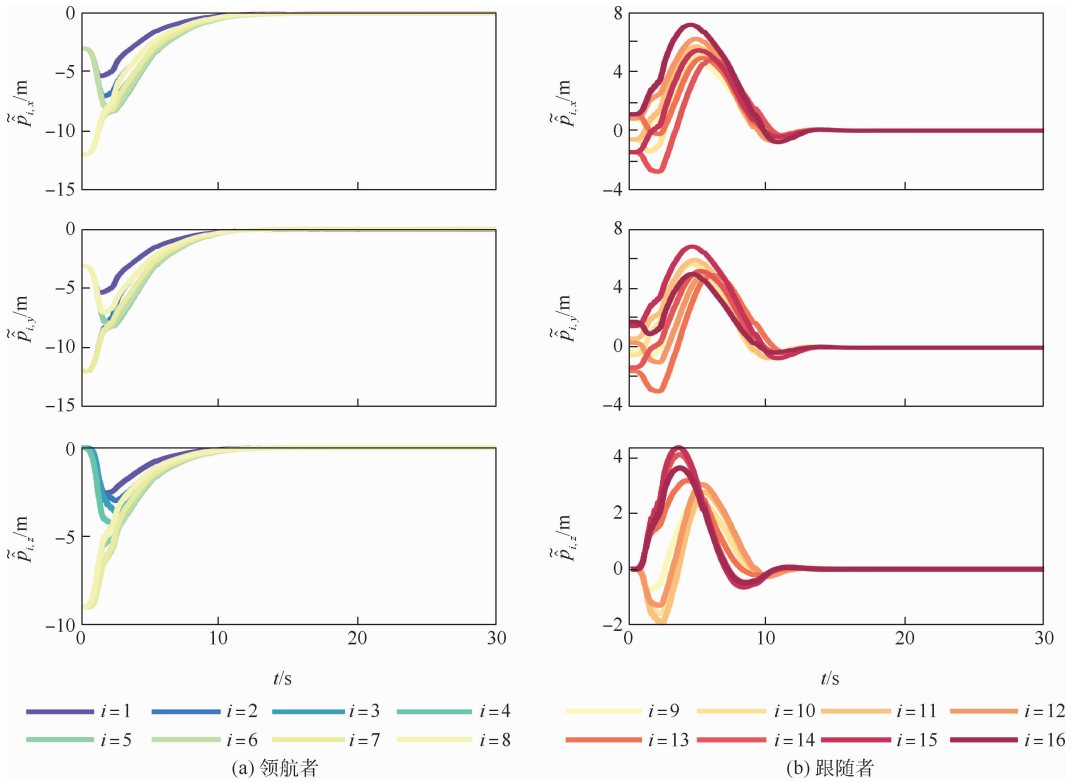


图 7 领航者与跟随者的期望位置估计误差

Fig. 7 Estimated errors of leader and follower desired positions

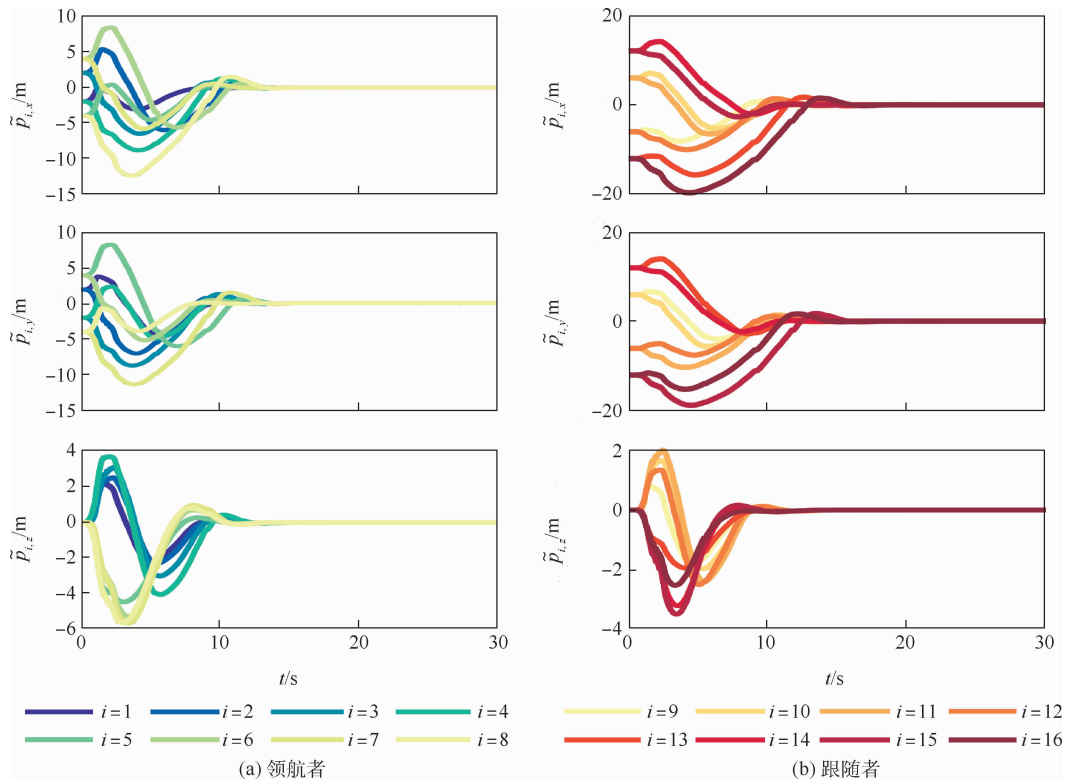


图 8 领航者与跟随者的估计位置跟踪误差

Fig. 8 Tracking errors of leader and follower estimated positions

以后,集群系统完成了编队合围。图 9 和图 10 表明了每架无人机的姿态能够收敛。图 11 给出了一架领航者无人机与一架跟随者无人机对其估计

速度的 1~3 阶导数逼近误差,可知大约在 12 s 以后,每架无人机可以对估计速度的高阶导数进行精确逼近,有效防止了微分爆炸。图 12(a)为

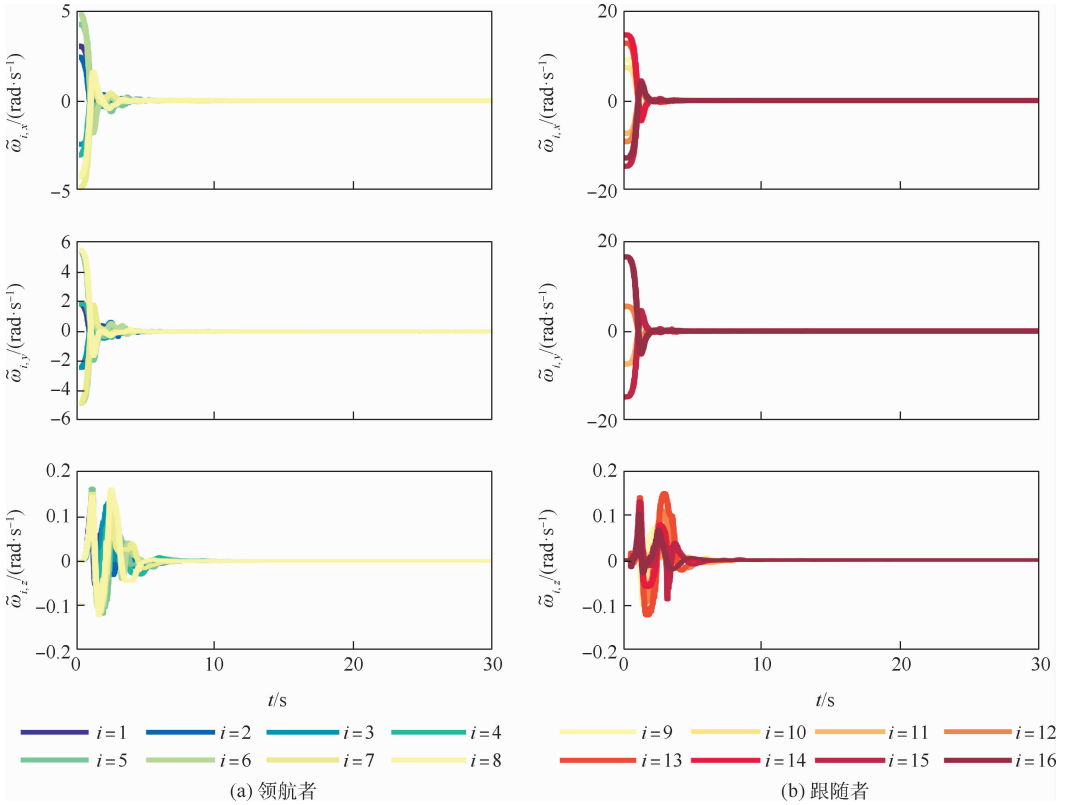


图 9 领航者与跟随者的角速度跟踪误差

Fig. 9 Tracking errors of leader and follower angular velocities

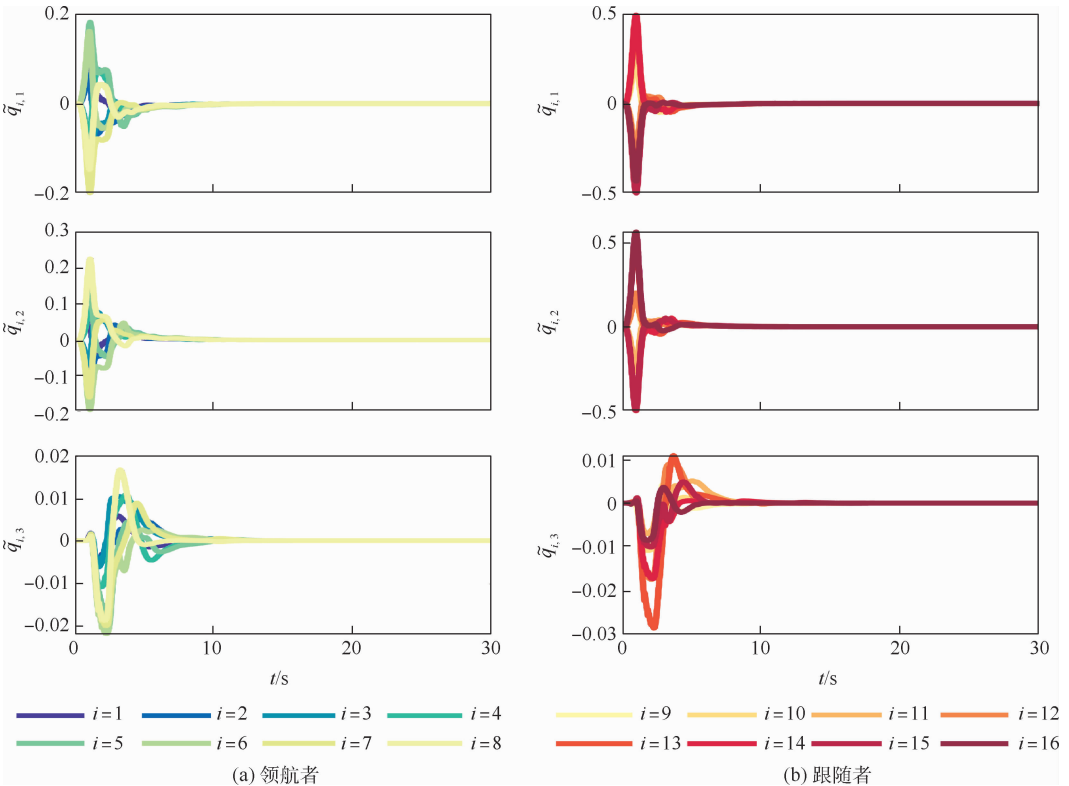


图 10 领航者与跟随者的姿态跟踪误差

Fig. 10 Tracking errors of leader and follower attitude

图 3(a) 编队在二维平面上的旋转航迹,图 12(b) 为图 3(b) 编队在空间中的螺旋曲线航迹。仿真

结果表明,在航迹较为复杂的情况下,所设计的有限时间滑模估计器和控制律仍能满足控制要求。

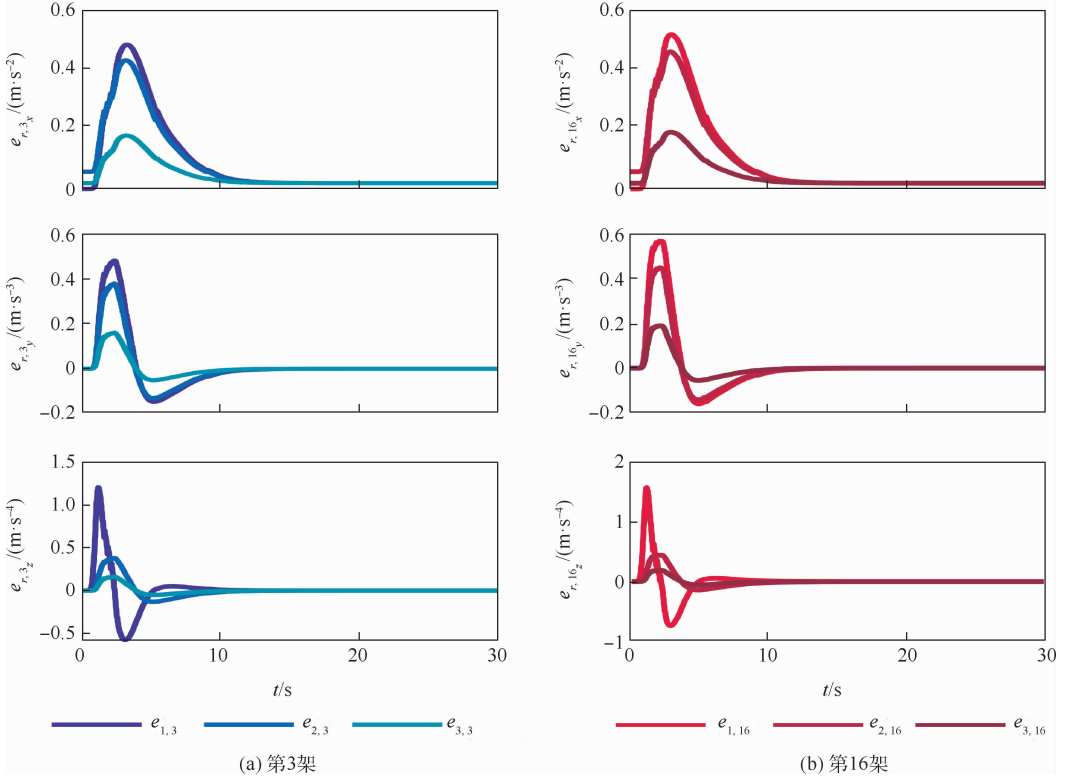


图 11 第 3 架与第 16 架四旋翼无人机对估计速度 1~3 阶导数的逼近误差

Fig. 11 Approximation errors of the first to third derivatives of the 3rd and 16th quadrotor UAVs' estimated velocities

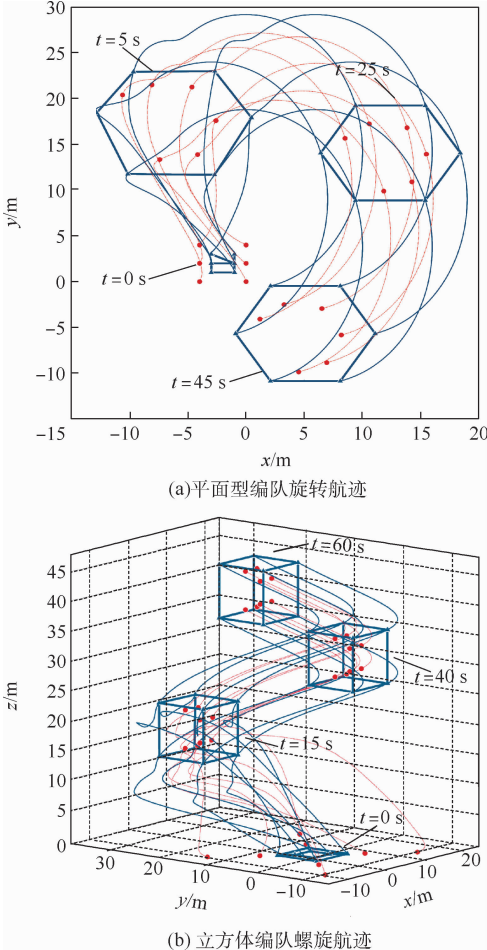


图 12 两种编队包围航迹

Fig. 12 Formation-containment trajectories of two type of formations

### 5 结 论

对于一个只有部分领航者能获取到给定期望信息的多领航者与多跟随者构成的四旋翼无人机群系统,能够实现以下 3 个指标:

1) 每架无人机能够在有限时间内实现对给定期望信息的估计,且领航者无人机能够估计出编队队形位置,跟随者无人机能够估计出编队内部的位置。

2) 每架无人机能够在有限时间内对其生成的估计信息的高阶导数进行近似逼近,该逼近结果可用于每架无人机的位置与姿态跟踪控制。

3) 每架无人机的实际位置能够精确跟踪到其生成的估计位置,且每架无人机能够实现姿态收敛,即整个多四旋翼无人机系统完成了编队包围控制。

后续工作将研究多四旋翼无人机系统的时变编队包围控制,以及在四旋翼模型的基础上加入不确定性并且考虑含有外部扰动情况下的编队包围控制。

### 参考文献 (References)

[ 1 ] 韩亮,任章,董希旺,等. 多无人机协同控制方法及应用研究[J]. 导航定位与授时, 2018, 5(4): 1-7.  
 HAN L, REN Z, DONG X W, et al. Research on cooperative

- control method and application for multiple unmanned aerial vehicles[J]. *Navigation Positioning and Timing*, 2018, 5(4): 1-7 (in Chinese).
- [2] 马思迁,董朝阳,马鸣宇,等. 基于自适应通信拓扑四旋翼无人机编队重构控制[J]. *北京航空航天大学学报*, 2018, 44(4): 841-850.  
MA S Q, DONG C Y, MA M Y, et al. Formation reconfiguration control of quadrotor UAVs based on adaptive communication topology[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2018, 44(4): 841-850 (in Chinese).
- [3] 徐扬,罗德林,周路鹏,等. 基于增益矩阵的二阶集群系统鲁棒分布式立体编队控制[J]. *中国科学:技术科学*, 2020, 50(4): 461-474.  
XU Y, LUO D L, ZHOU L P, et al. A gain matrix approach for robust distributed 3D formation control with second order swarm systems[J]. *Scientia Sinica: Technologica*, 2020, 50(4): 461-474 (in Chinese).
- [4] 任章,郭栋,董晔旺,等. 飞行器集群协同制导控制方法及应用研究[J]. *导航定位与授时*, 2019, 6(5): 1-9.  
REN Z, GUO D, DONG X W, et al. Research on the cooperative guidance and control method and application for aerial vehicle swarm systems[J]. *Navigation Positioning and Timing*, 2019, 6(5): 1-9 (in Chinese).
- [5] 陈杰,辛斌. 有人/无人系统自主协同的关键科学问题[J]. *中国科学:信息科学*, 2018, 48(9): 1270-1274.  
CHEN J, XIN B. Key scientific problems in the autonomous cooperation of manned-unmanned systems[J]. *Scientia Sinica: Informationis*, 2018, 48(9): 1270-1274 (in Chinese).
- [6] 段海滨,张岱峰,范彦铭,等. 从狼群智能到无人机集群协同决策[J]. *中国科学:信息科学*, 2019, 49(1): 112-118.  
DUAN H B, ZHANG D F, FAN Y M, et al. From wolf pack intelligence to UAV swarm cooperative decisionmaking[J]. *Scientia Sinica: Informationis*, 2019, 49(1): 112-118 (in Chinese).
- [7] LUO D L, SHAO J, XU Y, et al. Coevolution pigeon-inspired optimization with cooperation-competition mechanism for multi-UAV cooperative region search[J]. *Applied Sciences*, 2019, 9(5): 827.
- [8] 周绍磊,祁亚辉,张雷,等. 切换拓扑下无人机集群系统时变编队控制[J]. *航空学报*, 2017, 38(4): 320452.  
ZHOU S L, QI Y H, ZHANG L, et al. Time-varying formation control of UAV swarm systems with switching topologies[J]. *Acta Aeronautica et Astronautica Sinica*, 2017, 38(4): 320452 (in Chinese).
- [9] 胡春鹤,陈宗基. 基于Helly定理的多智能体最短时间一致性[J]. *北京航空航天大学学报*, 2015, 41(9): 1701-1707.  
HU C H, CHEN Z J. Helly-theorem-based time-optimal consensus for multi-agent systems[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2015, 41(9): 1701-1707 (in Chinese).
- [10] REN W. Consensus strategies for cooperative control of vehicle formations[J]. *IET Control Theory & Applications*, 2007, 1(2): 505-512.
- [11] 姚辉,席建祥,王成,等. 二阶多智能体系统自抗扰编队跟踪与避撞控制[J]. *北京航空航天大学学报*, 2020, 46(5): 960-977.  
YAO H, XI J X, WANG C, et al. Active disturbance rejection based formation tracking and collision avoidance control for second-order multi-agent system[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2020, 46(5): 960-977 (in Chinese).
- [12] 陈亮名. 考虑约束的多智能体 Euler-Lagrange 系统编队-包含控制[D]. 哈尔滨:哈尔滨工业大学, 2019.  
CHEN L M. Formation-containment control of multi-agent Euler-Lagrange systems with constraints[D]. Harbin: Harbin Institute of Technology, 2019 (in Chinese).
- [13] DONG X W. Formation and containment control for high-order linear swarm systems[M]. Berlin: Springer, 2016.
- [14] LI C J, CHEN L M, GUO Y N, et al. Formation-containment control for networked Euler-Lagrange systems with input saturation[J]. *Nonlinear Dynamics*, 2018, 91(2): 1307-1320.
- [15] CHEN L M, LI C J, MEI J, et al. Adaptive cooperative formation-containment control for networked Euler-Lagrange systems without using relative velocity information[J]. *IET Control Theory & Applications*, 2017, 11(9): 1450-1458.
- [16] DONG X W, YU B C, SHI Z Y, et al. Time-varying formation control for unmanned aerial vehicles: Theories and applications [J]. *IEEE Transactions on Control Systems Technology*, 2015, 23(1): 340-348.
- [17] DONG X W, ZHOU Y, REN Z, et al. Time-varying formation tracking for second-order multi-agent systems subjected to switching topologies with application to quadrotor formation flying[J]. *IEEE Transactions on Industrial Electronics*, 2017, 64(6): 5014-5024.
- [18] ZHOU Y, DONG X W, LU G, et al. Time-varying formation control for unmanned aerial vehicles with switching interaction topologies[C] // 2014 International Conference on Unmanned Aircraft Systems (ICUAS). Piscataway: IEEE Press, 2014: 1203-1209.
- [19] ZHANG W Q, DONG C Y, RAN M P, et al. Fully distributed time-varying formation tracking control for multiple quadrotor vehicles via finite-time convergent extended state observer[J]. *Chinese Journal of Aeronautics*, 2020, 33(11): 2907-2920.
- [20] XU Y, ZHAO S Y, LUO D L, et al. Affine formation maneuver control of high-order multi-agent systems over directed networks [J]. *Automatica*, 2020, 118: 109004.
- [21] ZHAO S Y. Affine formation maneuver control of multiagent systems[J]. *IEEE Transactions on Automatic Control*, 2018, 63(12): 4140-4155.
- [22] SHAO X L, TIAN B, YANG W. Fixed-time trajectory following for quadrotors via output feedback[J]. *ISA Transactions*, 2021, 110: 213-224.
- [23] LI D Y, ZHANG W, HE W, et al. Two-layer distributed formation-containment control of multiple Euler-Lagrange systems by output feedback[J]. *IEEE Transactions on Cybernetics*, 2019, 49(2): 675-687.
- [24] YU Y S, DING X L. A global tracking controller for underactuated aerial vehicles: Design, analysis, and experimental tests on quadrotor [J]. *IEEE/ASME Transactions on Mechatronics*, 2016, 21(5): 2499-2511.

- [25] MENG Z Y, REN W, YOU Z. Distributed finite-time attitude containment control for multiple rigid bodies[J]. *Automatica*, 2010, 46(12):2092-2099.
- [26] SHUSTER M D. A survey of attitude representations[J]. *Journal of Astronautical Sciences*, 1993, 41(4):439-517.
- [27] HORN R A, JOHNSON C R. Matrix analysis[M]. Cambridge: Cambridge University Press, 2012.
- [28] BHAT S P, BERNSTEIN D S. Finite-time stability of continuous autonomous systems[J]. *SIAM Journal on Control and Optimization*, 2000, 38(3):751-766.
- [29] ROBERTS A, TAYEBI A. Adaptive position tracking of VTOL UAVs[J]. *IEEE Transactions on Robotics*, 2011, 27(1):129-142.
- [30] ZHU B, HUO W. Nonlinear control for a model-scaled helicopter with constraints on rotor thrust and fuselage attitude[J]. *Acta Automatica Sinica*, 2014, 40(11):2654-2664.
- [31] LI C Y, JING W X, GAO C S. Adaptive backstepping-based flight control system using integral filters[J]. *Aerospace Science and Technology*, 2009, 13(2-3):105-113.
- [32] XU Y, LUO D L, LI D Y, et al. Target-enclosing affine formation control of two-layer networked spacecraft with collision avoidance[J]. *Chinese Journal of Aeronautics*, 2019, 32(12):2679-2693.

## Distributed hierarchical formation-containment control of multiple quadrotor UAV systems

ZHENG Weiming<sup>1</sup>, XU Yang<sup>2,3</sup>, LUO Delin<sup>1,\*</sup>

(1. School of Aerospace Engineering, Xiamen University, Xiamen 361102, China;

2. School of Civil Aviation, Northwestern Polytechnical University, Xi'an 710072, China;

3. Yangtze River Delta Research Institute of NPU, Taicang, Suzhou 215400, China)

**Abstract:** For the under-actuated quadrotor UAV swarm systems with multiple leaders and followers, a distributed hierarchical formation-containment control method is proposed. First, a hierarchical distributed finite-time sliding mode estimator is designed to achieve that each UAV can generate estimated position information that meets the control needs under the condition that only some leaders can obtain the desired trajectory. Then, considering the research object is an under-actuated six-degree-of-freedom quadrotor UAV model, a hierarchical control method of the UAV position layer and the attitude layer is proposed, which realizes the tracking control of the generated estimated position. This method adopts a high-order derivative approximation algorithm to prevent differential explosions in the process of solving the desired angular velocity. The given method can realize the effective formation-containment under the condition of satisfying the stable convergence of attitude. Finally, the accuracy and effectiveness of the proposed method are verified through numerical simulation.

**Keywords:** formation-containment control; formation control; quadrotor UAV; under-actuated system; distributed control

**Received:** 2020-12-31; **Accepted:** 2021-02-06; **Published online:** 2021-02-22 16:47

**URL:** [kns.cnki.net/kcms/detail/11.2625.V.20210222.1636.003.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210222.1636.003.html)

**Foundation item:** Jointly Supported by the Science and Technology on Avionics Integration Laboratory and the Aeronautical Science Foundation of China (20185568005)

\* **Corresponding author.** E-mail: luodelin1204@xmu.edu.cn

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2020.0728

# 采用改进导重法的拓扑结构灰度单元过滤技术

高翔<sup>1,2</sup>, 王林军<sup>1,2,\*</sup>, 杜义贤<sup>1,2</sup>, 付君健<sup>1,2</sup>

(1. 三峡大学 水电机设备设计与维护湖北省重点实验室, 宜昌 443002; 2. 三峡大学 机械与动力学院, 宜昌 443002)

**摘 要:** 针对二分法计算拉格朗日乘子时收敛速度较慢的问题, 提出了拉格朗日乘子计算方法, 应用于优化准则(OC)法和导重(GW)法 2 种密度更新方法, 并与二分法进行了对比。建立体积约束下柔度最小的拓扑优化模型; 通过固体各向同性材料惩罚(SIMP)法或材料属性有理近似(RAMP)法计算单元的弹性模量; 通过所提方法计算拉格朗日乘子, 并通过导重法更新单元密度; 通过 Heaviside 投影函数减少灰度单元的数量。计算结果表明: 虽然所提方法对有限元分析次数并没有显著改进, 但计算拉格朗日乘子所用 CPU 时间少于二分法, 且密度更新次数降低至 50% 以下; 在 2 个数值算例中, 采用 SIMP 模型时, 导重法所得结构柔度比 OC 法更小, 能够得到刚度更高的结构。

**关键词:** 二分法; 优化准则(OC)法; 导重(GW)法; Heaviside 投影函数; 拓扑优化  
**中图分类号:** TH122

**文献标志码:** A **文章编号:** 1001-5965(2022)06-1106-09

拓扑优化旨在对结构进行轻量化设计。在保证结构的力学性能、固有频率、制造约束、可靠性等约束条件的情况下, 使得结构的质量最轻。

自 1904 年 Michell 提出桁架结构拓扑优化以来, 很多学者相继提出了不同的拓扑优化方法<sup>[1]</sup>。Bendsøe 和 Kikuchi<sup>[2]</sup>提出了均匀化方法和固体各向同性材料惩罚(solid isotropic material with penalization, SIMP)法<sup>[3-4]</sup>。Rietz<sup>[5]</sup>提出了材料属性有理近似(rational approximation of material properties, RAMP)法。Stolpe 和 Svanberg<sup>[6]</sup>证明了 RAMP 法一定能构造凸函数, 而 SIMP 法不一定能构造凸函数。优化准则(optimal criteria, OC)法<sup>[7-8]</sup>和导重(guide-weight, GW)法<sup>[9-11]</sup>均为准则法, 收敛速度较快, 但 OC 法的密度更新公式需要随着拓扑优化模型的变化而改变, 因此导重法的通用性更好<sup>[12]</sup>。

焦洪宇等<sup>[13]</sup>通过导重法对轮毂等类周期性结

构进行了拓扑优化设计。陈树勋等<sup>[14]</sup>通过导重法求解了多约束条件下的拓扑优化问题。李枝东和刘辛军<sup>[15]</sup>通过导重法求解了柔度最小化问题及质量最小化问题, 并与 ANSYS 所得结果进行了对比。任毅如等<sup>[16]</sup>通过导重法求解了自重载荷条件下的拓扑优化问题。陈祥和刘辛军<sup>[17]</sup>通过 ANSYS 求解了基于 RAMP 法和导重法的拓扑优化问题。

Andreassen 等<sup>[18]</sup>编写了 SIMP 法和 OC 法的 88 行代码, 并介绍了灰度单元的抑制方法。秦浩星等<sup>[19]</sup>采用了龙凯和赵红伟<sup>[20]</sup>提出的 Heaviside 投影函数, 并对导重法的灰度抑制方法进行了研究。杜义贤<sup>[21]</sup>、昌俊康<sup>[22]</sup>和高翔<sup>[23]</sup>等均提出了不同的 Heaviside 投影函数, 并将其应用于灰度单元的抑制。陈垂福和杨晓翔<sup>[24]</sup>对导重法密度更新公式中的移动步长进行了研究, 并提出了一种根据结构最优性指标的步长计算公式。

文献[2-11]奠定了拓扑优化的基础理论。文

收稿日期: 2021-01-02; 录用日期: 2021-03-29; 网络出版时间: 2021-04-12 17:04

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20210412.1556.002.html

基金项目: 国家自然科学基金(51775308); 水电机设备设计与维护湖北省重点实验室开放基金(2019KJX12)

\* 通信作者。E-mail: ljwang2006@126.com

**引用格式:** 高翔, 王林军, 杜义贤, 等. 采用改进导重法的拓扑结构灰度单元过滤技术[J]. 北京航空航天大学学报, 2022, 48(6): 1106-1114. GAO X, WANG L J, DU Y X, et al. Gray element filtering technology of topology structure based on improved guide-weight method[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(6): 1106-1114 (in Chinese).

献[12-17]对导重法进行了研究,但是并未考虑结构中的灰度单元。文献[18-23]对变密度法拓扑优化模型中灰度单元的抑制问题进行了研究,但文献[19-23]的研究仅限于 OC 法,并未对导重法进行研究。虽然文献[24]对导重法密度更新公式中的步长因子进行了研究,但并未显著提高迭代速度。因此,在文献[18-23]的基础上,本文对导重法的灰度单元问题进行了研究。同时,本文对导重法的拉格朗日乘子的更新公式进行了改进,提高了拉格朗日乘子的计算效率,将密度的更新次数降低至原始次数的 50% 以下。

## 1 拓扑优化理论

### 1.1 插值模型

离散变量给拓扑优化问题带来了一些不便,因此提出了不同的弹性模量插值公式,将离散变量转化为连续变量,降低了问题的求解难度。在 SIMP 法的基础上,RAMP 法<sup>[5-6]</sup>、SINH 法<sup>[25]</sup>相继被提出并做出了改进。二者的优点分别为:①当弹性模量位于 Hashin-Shtrikman 边界的下界时,RAMP 法仍然能够构造凸函数,而 SIMP 不法可以;②当惩罚系数  $\gamma$  逐渐增加时,SINH 法所得结构的体积与实际体积的误差更小,而 SIMP 法的体积大于 SINH 法和实际体积。

SIMP、RAMP、SINH 三种插值模型对应的弹性模量分别为

$$E_e(x_e) = E_{\min} + x_e^\gamma(E_0 - E_{\min}) \quad (1)$$

$$E_e(x_e) = E_{\min} + \frac{x_e}{1 + \gamma(1 - x_e)}(E_0 - E_{\min}) \quad (2)$$

$$E_e(x_e) = E_{\min} + \frac{\sinh(\gamma x_e)}{\sinh(\gamma)}(E_0 - E_{\min}) \quad (3)$$

SIMP、RAMP、SINH 三种插值模型对应的插值曲线如图 1 所示。

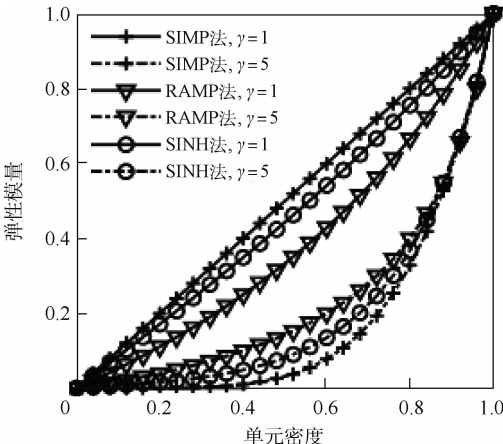


图 1 SIMP 法和 RAMP 法的弹性模量对比

Fig. 1 Elastic module comparison of SIMP and RAMP methods

### 1.2 最小柔度拓扑优化模型

在体积约束下的柔度最小化问题中,变密度法的拓扑优化模型为

$$\begin{cases} \min c = \mathbf{U}^T \mathbf{K} \mathbf{U} = \sum_{e=1}^N E_e(x_e) \mathbf{u}_e^T \mathbf{k}_0 \mathbf{u}_e \\ \text{s. t. } g = \frac{V}{V_0} - f = 0 \\ \mathbf{K} \mathbf{U} = \mathbf{F} \\ 0 \leq x_e \leq 1 \end{cases} \quad (4)$$

式中: $c$  为结构柔度; $\mathbf{K}$  为全局刚度矩阵; $\mathbf{k}_0$  为单元刚度矩阵; $\mathbf{U}$  为全局位移列向量; $\mathbf{u}_e$  为第  $e$  个单元的位移列向量; $\mathbf{F}$  为载荷列向量; $V$  为当前结构的体积; $V_0$  为初始结构的体积; $f$  为期望的体积分数; $g$  为体积约束。

SIMP、RAMP、SINH 三种插值模型对应的柔度的灵敏度分别为

$$\frac{\partial c}{\partial x_e} = - \sum_{e=1}^N \gamma x_e^{\gamma-1} (E_0 - E_{\min}) \mathbf{u}_e^T \mathbf{k}_0 \mathbf{u}_e \quad (5)$$

$$\frac{\partial c}{\partial x_e} = - \sum_{e=1}^N \frac{1 + \gamma}{[1 + \gamma(1 - x_e)]^2} (E_0 - E_{\min}) \mathbf{u}_e^T \mathbf{k}_0 \mathbf{u}_e \quad (6)$$

$$\frac{\partial c}{\partial x_e} = - \sum_{e=1}^N \gamma \frac{\cosh(\gamma x_e)}{\sinh(\gamma)} (E_0 - E_{\min}) \mathbf{u}_e^T \mathbf{k}_0 \mathbf{u}_e \quad (7)$$

SIMP 法、RAMP 法、SINH 法 3 种插值模型中,RAMP 法和 SINH 法对应的拓扑结构近似,所以本文仅采用 SIMP 法和 RAMP 法 2 种方法进行对比。

无论采用 SIMP 法还是 RAMP 法,体积的灵敏度均为

$$\partial V / \partial x_e = 1 \quad (8)$$

### 1.3 密度更新模型

由于 OC 法的密度更新公式需要随优化模型做出修改,导重法的通用性优于 OC 法。OC 法和导重法均属于准则法,且根据库恩塔克条件可求得对应的密度更新公式。OC 法和导重法对应的密度更新公式分别为

$$x_e = \begin{cases} \max(0, x_e - m) & x_e B_e^\eta \leq \max(0, x_e - m) \\ \min(1, x_e + m) & x_e B_e^\eta \geq \min(1, x_e + m) \\ x_e B_e^\eta & \text{其他} \end{cases} \quad (9)$$

$$x_e = \begin{cases} 0 & x_e \leq 0 \\ 1 & x_e \geq 1 \\ m[G_e/(\lambda H_e)] + (1 - mx_e) & \text{其他} \end{cases} \quad (10)$$

式中: $m$  为移动步长; $\eta$  为阻尼系数; $\lambda$  为拉格朗

日乘子;  $B_e = \frac{-\partial c/\partial x_e}{\lambda \partial V/\partial x_e}$  为启发式更新因子;  $G_e = -x_e \frac{\partial c}{\partial x_e}$  为单元导重;  $H_e = \frac{\partial V}{\partial x_e}$  为单元容重。

#### 1.4 拉格朗日乘子更新

二分法是一种最简单的一维搜索算法,但是计算效率较低。进退法是一种收敛速度更快的一维搜索算法,通常用于寻找函数的极小值<sup>[26]</sup>。本文对进退法进行了修改,并将普通二分法替换为本文提出的方法。

1) 普通二分法。

步骤1 令  $\lambda_{\min}^1 = 0, \lambda_{\max}^1 = 1 \times 10^9$ 。

步骤2 令  $\lambda^{k+1} = (\lambda_{\min}^k + \lambda_{\max}^k)/2$ , 并更新单元密度。

步骤3 若  $g < 0$ , 则  $\lambda_{\min}^{k+1} = \lambda^{k+1}$ ; 否则,  $\lambda_{\max}^{k+1} = \lambda^{k+1}$ 。

步骤4  $k = k + 1$ , 转步骤2。当  $(\lambda_{\max}^k - \lambda_{\min}^k)/(\lambda_{\min}^k + \lambda_{\max}^k) \leq 1 \times 10^{-3}$  时, 迭代终止。

2) 本文方法。

步骤1 令  $\lambda_{\min}^1 = 0, \lambda_{\max}^1 = 1 \times 10^9, h^1 = 1, h$  为步长。

步骤2 令  $\lambda^{k+1} = \lambda_{\min}^k + h^k$ , 并更新单元密度。

步骤3 若  $g < 0$ , 则  $\lambda_{\min}^{k+1} = \lambda^{k+1}, h^{k+1} = 2h^k$ ; 否则,  $\lambda_{\max}^{k+1} = \lambda^{k+1}, h^{k+1} = h^k/4$ 。

步骤4  $k = k + 1$ , 转步骤2。当  $(\lambda_{\max}^k - \lambda_{\min}^k)/(\lambda_{\min}^k + \lambda_{\max}^k) \leq 1 \times 10^{-3}$  时, 迭代终止。

#### 1.5 灰度过滤技术

变密度法所得拓扑优化结构中大量的灰度单元, 导致结构的边界比较模糊。因此, 需要采取一些措施, 提高拓扑优化结构边界的清晰度。首先, 使用密度过滤方法抑制拓扑优化结构的棋盘格现象。其次, 使用 Heaviside 投影函数对灰度单元进行抑制, 减少结构中灰度单元的数量, 以获得具有清晰边界的拓扑优化结构。

1) 密度过滤以后的灵敏度和单元密度分别为

$$\frac{\partial \phi}{\partial x_j} = \sum_{e \in N_j} \frac{\partial \phi}{\partial \hat{x}_j} \cdot \frac{\partial \hat{x}_j}{\partial x_j} = \sum_{e \in N_j} \frac{1}{\sum_{i \in N_e} H_{ei}} H_{je} \frac{\partial \phi}{\partial \hat{x}_e} \quad (11)$$

$$\hat{x}_e = \frac{1}{\sum_{i \in N_e} H_{ei}} \sum_{i \in N_e} H_{ei} x_i \quad (12)$$

式中:  $N_e$  代表符合条件  $\Delta(e, i) < r_{\min}$  的区域;  $H_{ei} = \max(0, r_{\min} - \Delta(e, i))$  为卷积算子;  $\hat{x}_e$  为密度过滤后的单元密度。为了避免类似的公式重复出现, 体积和柔度的灵敏度均用  $\phi$  表示。

2) 通过 Heaviside 投影函数过滤灰度单元时, 体积和柔度的灵敏度为

$$\frac{\partial \phi}{\partial x_j} = \sum_{e \in N_j} \frac{\partial \phi}{\partial \hat{x}_e} \cdot \frac{\partial \hat{x}_e}{\partial x_e} = \sum_{e \in N_j} \frac{1}{\sum_{i \in N_e} H_{ei}} H_{je} \frac{\partial \phi}{\partial x_e^{\text{new}}} \cdot \frac{\partial x_e^{\text{new}}}{\partial \hat{x}_e} \quad (13)$$

式中:  $x_e^{\text{new}}$  为通过 Heaviside 投影函数过滤灰度单元后的单元密度。

文献[27]指出, 传统 Heaviside 投影函数为

$$\begin{cases} x_e^{\text{new}} = 1 - \exp(-\beta \hat{x}_e) + \hat{x}_e \exp(-\beta) \\ \frac{\partial x_e^{\text{new}}}{\partial \hat{x}_e} = \beta \exp(-\beta \hat{x}_e) + \exp(-\beta) \end{cases} \quad (14)$$

式中:  $\beta$  为投影函数的系数。

文献[23]提出的 Heaviside 投影函数为

$$\begin{cases} x_e^{\text{new}} = 1 - \frac{\exp[\beta(1 - \hat{x}_e)] - 1}{\exp(\beta) - 1} \\ \frac{\partial x_e^{\text{new}}}{\partial \hat{x}_e} = \frac{\beta \exp[\beta(1 - \hat{x}_e)]}{\exp(\beta) - 1} \end{cases} \quad (15)$$

## 2 算例验证

### 2.1 单载荷简支梁的拓扑优化

简支梁宽 120 mm, 高 60 mm, 下侧一端固定一端简支, 载荷作用于下端中心。该简支梁的结构示意图如图 2 所示。

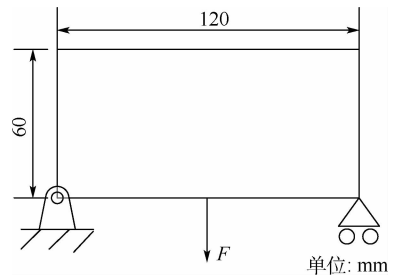


图2 二维简支梁的结构示意图

Fig. 2 Schematic structure of two-dimensional simply supported beam

设  $F = 1 \text{ N}$ ,  $E_{\min} = 1 \times 10^{-9} \text{ MPa}$ ,  $E_0 = 1 \text{ MPa}$ ,  $f = 0.5$ , 过滤半径  $r_{\min} = 2.5$ ,  $\gamma = 3$ 。参数  $\beta$  由 1 逐渐递增且每 50 步乘以 2,  $\beta$  增加到 512 后停止增加。SIMP 法的计算结果如表 1 和表 2 所示, RAMP 法的计算结果如表 3 和表 4 所示。SIMP 法的迭代过程如图 3 和图 4 所示, RAMP 法的迭代过程如图 5 和图 6 所示。

由表 1、表 2、图 3 和图 4 可知, 当采用 SIMP 法时:

1) 将二分法替换为本文方法后, 虽然外层迭代(有限元分析)次数没有显著减少, 但是 CPU 时间明显减少, 且迭代总数(密度更新)减少至 50% 以下。

表 1 OC 法求解 SIMP 模型所得结果(简支梁)

Table 1 Results from OC method of solving SIMP model (simply supported beam)

Heaviside 投影函数	OC 法(二分法)				拓扑优化结构	OC 法(本文方法)				
	迭代次数		结构 柔度	CPU 时间/s		迭代次数		结构 柔度	CPU 时间/s	拓扑优化结构
	外层	总数				外层	总数(改进)			
式(14)	445	20 761	14.447 6	63.187 0		469	9 392 (54.76%)	14.444 8	62.510 5	
式(15)	479	22 332	14.438 5	65.354 4		481	9 813 (56.06%)	14.438 6	64.020 7	

表 2 导重法求解 SIMP 模型所得结果(简支梁)

Table 2 Results from GW method of solving SIMP model (simply supported beam)

Heaviside 投影函数	导重法(二分法)				拓扑优化结构	导重法(本文方法)				
	迭代次数		结构 柔度	CPU 时间/s		迭代次数		结构 柔度	CPU 时间/s	拓扑优化结构
	外层	总数				外层	总数(改进)			
式(14)	517	23 821	14.362 4	70.171 0		517	10 470 (56.05%)	14.362 4	68.878 6	
式(15)	559	25 780	14.369 8	75.572 8		559	11 389 (55.82%)	14.369 8	74.951 0	

表 3 OC 法求解 RAMP 模型所得结果(简支架)

Table 3 Results from OC method of solving RAMP model (simply supported beam)

Heaviside 投影函数	OC 法(二分法)				拓扑优化结构	OC 法(本文方法)				
	迭代次数		结构 柔度	CPU 时间/s		迭代次数		结构 柔度	CPU 时间/s	拓扑优化结构
	外层	总数				外层	总数(改进)			
式(14)	489	23 001	14.463 7	67.421 9		491	10 227 (55.54%)	14.463 5	65.436 4	
式(15)	488	22 925	14.618 5	67.756 6		485	9 469 (58.70%)	14.620 0	64.710 9	

表 4 导重法求解 RAMP 模型所得结果(简支梁)

Table 4 Results from GW method of solving RAMP model (simply supported beam)

Heaviside 投影函数	导重法(二分法)				拓扑优化结构	导重法(本文方法)				
	迭代次数		结构 柔度	CPU 时间/s		迭代次数		结构 柔度	CPU 时间/s	拓扑优化结构
	外层	总数				外层	总数(改进)			
式(14)	554	25 591	14.552 1	74.083 1		553	11 551 (54.86%)	14.552 1	71.908 0	
式(15)	490	22 507	14.509 1	67.468 5		490	10 158 (54.87%)	14.509 1	64.695 8	

2) 采用 OC 法时,式(15)可得到柔度更小的拓扑优化结构;采用导重法时,式(14)可得到柔度更小的拓扑优化结构。

3) 采用式(14)时,二分法和本文方法的迭代曲线不完全重合;采用式(15)时,二分法和本

文方法的迭代曲线完全重合。

由表 3、表 4、图 5 和图 6 可知,当采用 RAMP 法时:

1) 将二分法改为本文方法后,虽然总时间下降不显著,但是迭代总数减少至 50% 以下。

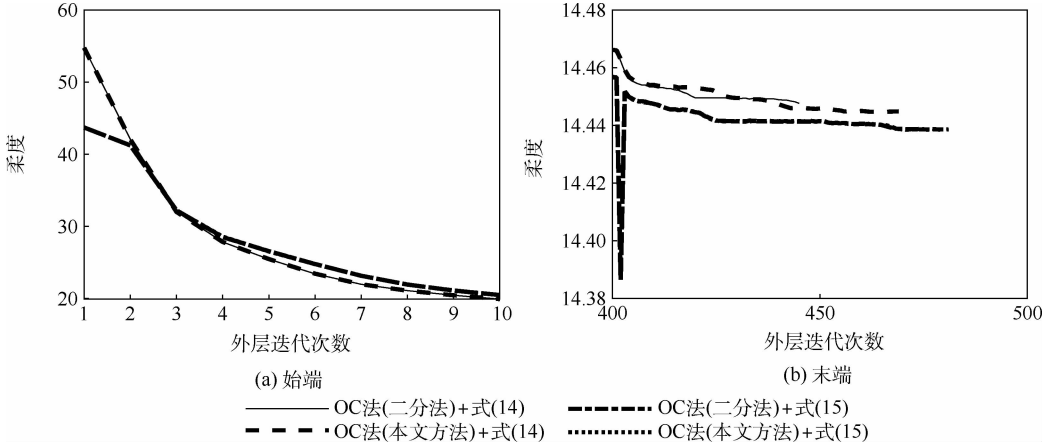


图 3 OC 法求解 SIMP 模型 (简支梁)

Fig. 3 OC method for solving SIMP model ( simply supported beam)

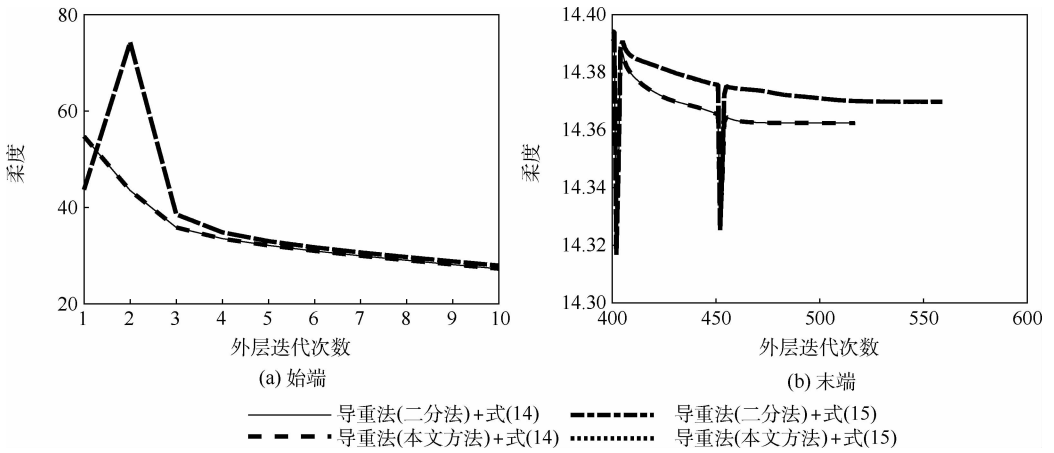


图 4 导重法求解 SIMP 模型 (简支梁)

Fig. 4 GW method for solving SIMP model ( simply supported beam)

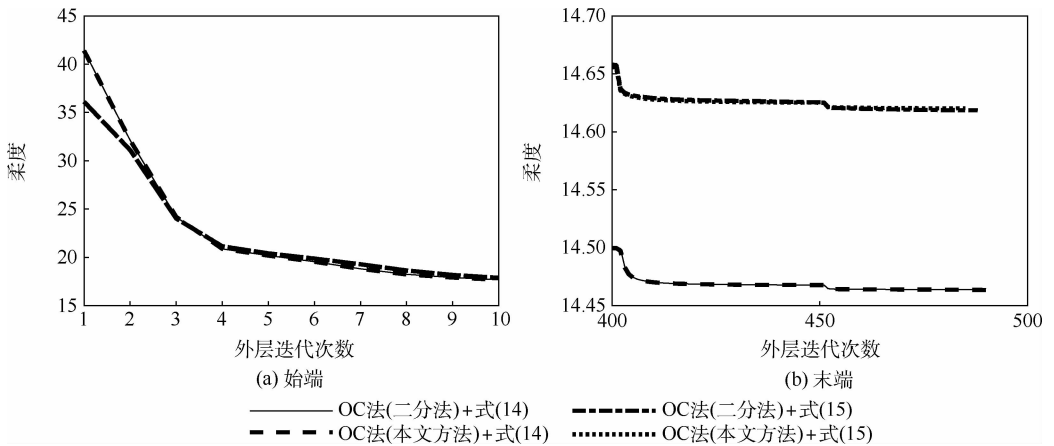


图 5 OC 法求解 RAMP 模型 (简支梁)

Fig. 5 OC method for solving RAMP model ( simply supported beam)

2) 采用 OC 法时,式(14)所得结构的柔度更小;采用导重法时,式(15)所得结构的柔度更小。

3) 无论采用式(14)还是式(15),二分法和本文方法的迭代曲线均完全重合。

### 2.2 双载荷悬臂梁的拓扑优化

悬臂梁宽 120 mm,高 60 mm,左端固定,下端

面右下角及 2/3 处承受集中力  $F$ 。该悬臂梁的结构示意图如图 7 所示。

设  $F = 1 \text{ N}$ ,  $E_{\min} = 1 \times 10^{-9} \text{ MPa}$ ,  $E_0 = 1 \text{ MPa}$ ,  $f = 0.5$ ,  $r_{\min} = 2.5$ ,  $\gamma = 3$ 。SIMP 法的计算结果如表 5 和表 6 所示,RAMP 法的计算结果如表 7 和表 8 所示。SIMP 法的迭代过程如图 8 和图 9 所示,RAMP 法的迭代过程如图 10 和图 11 所示。

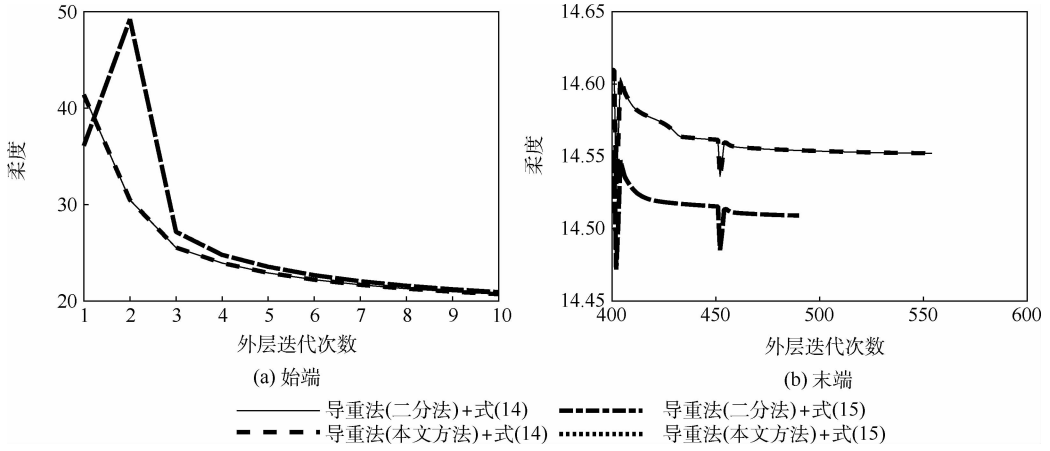


图 6 导重法求解 RAMP 模型(简支梁)

Fig. 6 GW method for solving RAMP model (simply supported beam)

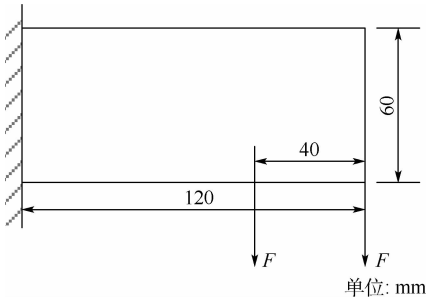


图 7 二维悬臂梁的结构示意图

Fig. 7 Schematic structure of two-dimensional cantilever beam

由表 5、表 6、图 8 和图 9 可知,当采用 SIMP 法时:

1) 本文改进后,总迭代次数降低至原来的

50% 以下,且 CPU 时间更少。

2) 采用导重法时,式(15)所得拓扑优化结构的柔度低于式(14)。

3) 无论采用式(14)还是式(15),二分法和本文方法的迭代曲线均完全重合。

由表 7、表 8、图 10 和图 11 可知,当采用 RAMP 法时:

1) 本文改进后,总迭代次数降低至原来的 50% 以下,且 CPU 时间更少。

2) 当采用导重法求解 RAMP 拓扑优化模型时,式(15)所得拓扑优化结构的柔度小于式(14)。

3) 无论采用式(14)还是式(15),二分法和本文方法的迭代曲线均完全重合。

表 5 OC 法求解 SIMP 模型所得结果(悬臂梁)

Table 5 Results from OC method of solving SIMP mode (cantilever beam)

Heaviside 投影函数	OC 法(二分法)				拓扑优化结构	OC 法(本文方法)				
	迭代次数		结构 柔度	CPU 时间/s		迭代次数		结构 柔度	CPU 时间/s	拓扑优化结构
	外层	总数				外层	总数(改进)			
式(14)	474	20 470	156.289 4	67.346 9		379	7 059 (65.52%)	156.724 5	48.525 3	
式(15)	395	17 089	156.632 5	56.269 1		396	7 262 (57.50%)	156.632 5	52.822 6	

表 6 导重法求解 SIMP 模型所得结果(悬臂梁)

Table 6 Results from GW method of solving SIMP model (cantilever beam)

Heaviside 投影函数	导重法(二分法)				拓扑优化结构	导重法(本文方法)				
	迭代次数		结构 柔度	CPU 时间/s		迭代次数		结构 柔度	CPU 时间/s	拓扑优化结构
	外层	总数				外层	总数(改进)			
式(14)	519	22 002	156.118 3	70.463 3		519	9 509 (56.78%)	156.117 9	65.420 6	
式(15)	528	22 369	156.096 8	68.503 2		528	9 646 (56.88%)	156.096 5	66.566 9	

表7 OC法求解 RAMP 模型所得结果(悬臂梁)

Table 7 Results from OC method of solving RAMP model (cantilever beam)

Heaviside 投影函数	OC法(二分法)					OC法(本文方法)				
	迭代次数		结构 柔度	CPU 时间/s	拓扑优化结构	迭代次数		结构 柔度	CPU 时间/s	拓扑优化结构
	外层	总数				外层	总数(改进)			
式(14)	528	22 913	155.547 8	74.144 5		531	9 680 (57.75%)	155.547 7	67.407 0	
式(15)	445	19 299	155.547 6	58.749 6		445	8 447 (56.23%)	155.547 8	56.336 7	

表8 导重法求解 RAMP 模型所得结果(悬臂梁)

Table 8 Results from GW method of solving RAMP mode (cantileve beam)

Heaviside 投影函数	导重法(二分法)					导重法(本文方法)				
	迭代次数		结构 柔度	CPU 时间/s	拓扑优化结构	迭代次数		结构 柔度	CPU 时间/s	拓扑优化结构
	外层	总数				外层	总数(改进)			
式(14)	579	24 683	155.949 9	74.149 1		579	11 262 (54.37%)	155.949 9	71.664 8	
式(15)	546	23 200	155.926 1	69.964 2		546	10 469 (54.88%)	155.926 0	67.687 6	

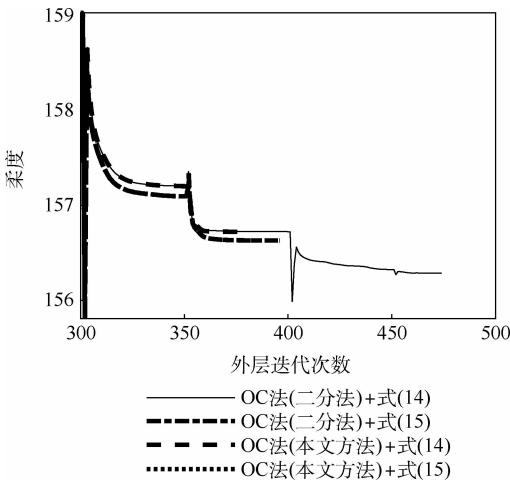


图8 OC法求解 SIMP 模型(悬臂梁)

Fig.8 OC method for solving SIMP model (cantilever beam)

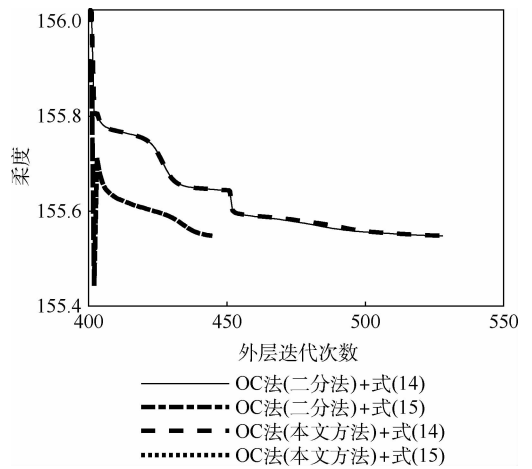


图10 OC法求解 RAMP 模型(悬臂梁)

Fig.10 OC method for solving RAMP model (cantilever beam)

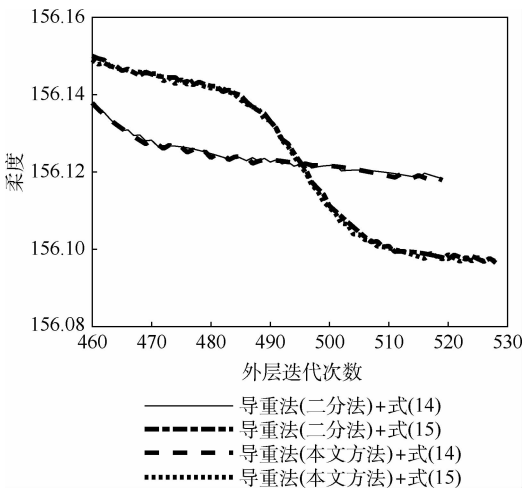


图9 导重法求解 SIMP 模型(悬臂梁)

Fig.9 GW method for solving SIMP model (cantilever beam)

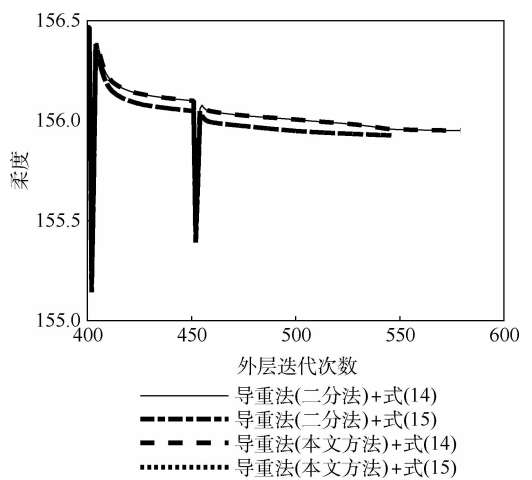


图11 导重法求解 RAMP 模型(悬臂梁)

Fig.11 GW method for solving RAMP model (cantilever beam)

### 3 结 论

本文提出了一种新的拉格朗日乘子迭代方法,并将其应用于 OC 法和导重法 2 种方法。采用 SIMP 法和 RAMP 法 2 种弹性模量插值方法,对比了本文方法与二分法。2 个算例的计算结果表明:

1) 本文改进后,OC 法和导重法的总迭代次数下降至 50% 以下,且 CPU 时间也有明显下降。

2) 算例 1 中,同时采用 SIMP 法和导重法时,所得拓扑优化结构的柔度最小;算例 2 中,采用 SIMP 法时,与 OC 法相比,导重法所得拓扑优化结构的柔度更小。因此,与 OC 法相比,导重法能显著降低 SIMP 模型所得结构的柔度,提高刚度。

本文仅对普通二维算例进行了研究,以后可将本文方法应用于三维算例,以及柔性机构算例、散热结构算例、声学结构算例。

#### 参考文献 (References)

- [ 1 ] MICHELL A G M. The limits of economy of material in frame structures[J]. Philosophical Magazine, 1904, 8(47): 589-597.
- [ 2 ] BENDSØE M P, KIKUCHI N. Generating optimal topologies in structural design using a homogenization method[J]. Computer Methods in Applied Mechanics and Engineering, 1988, 71(2): 197-224.
- [ 3 ] BENDSØE M P. Optimal shape design as a material distribution problem[J]. Structural Optimization, 1989, 1: 193-202.
- [ 4 ] BENDSØE M P, SIGMUND O. Material interpolation schemes in topology optimization[J]. Archive of Applied Mechanics, 1999, 69: 635-654.
- [ 5 ] RIETZ A. Sufficiency of a finite exponent in the SIMP (power law) method[J]. Structural and Multidisciplinary Optimization, 2001, 21: 159-163.
- [ 6 ] STOLPE M, SVANBERG K. An alternative interpolation scheme for minimum compliance topology optimization[J]. Structural and Multidisciplinary Optimization, 2001, 22: 116-124.
- [ 7 ] ROZVANY G I N, ZHOU M. The COC algorithm, Part I: Cross-section optimization or sizing[J]. Computer Methods in Applied Mechanics and Engineering, 1991, 89(1-3): 281-308.
- [ 8 ] ZHOU M, ROZVANY G I N. The COC algorithm, Part II: Topological, geometry and generalized shape optimization[J]. Computer Methods in Applied Mechanics and Engineering, 1991, 89(1-3): 309-336.
- [ 9 ] 叶尚辉, 陈树勋. 天线结构优化设计的最佳准则法[J]. 西北电讯工程学院学报, 1982, 9(1): 11-28.  
YE S H, CHEN S X. An optimality criterion method for antenna structure design[J]. Journal of Northwest Telecommunication Engineering Institute, 1982, 9(1): 11-28 (in Chinese).
- [ 10 ] 陈树勋. 精密复杂结构的几种现代设计方法[M]. 北京: 北京航空航天大学出版社, 1992.  
CHEN S X. Several modern design methods for precise and complex structures[M]. Beijing: Beihang University Press, 1992 (in Chinese).
- [ 11 ] 陈树勋. 工程结构系统的分析、综合与优化设计[M]. 香港: 中国科学文化出版社, 2008.  
CHEN S X. Analysis, synthesis and optimization of engineering structural system[M]. Hongkong: China Science Culture Publishing House, 2008 (in Chinese).
- [ 12 ] 安宗文, 孙道明, 秦浩星. 基于导重法结合小波分析的连续体结构拓扑优化[J]. 兰州理工大学学报, 2015, 41(6): 37-42.  
AN Z W, SUN D M, QIN H X. Topology optimization of continuum structure based on guide-weight method combined with wavelet analysis[J]. Journal of Lanzhou University of Technology, 2015, 41(6): 37-42 (in Chinese).
- [ 13 ] 焦洪宇, 李英, 胡顺安, 等. 基于导重法的结构类周期性布局优化方法研究[J]. 机械工程学报, 2020, 56(13): 218-230.  
JIAO H Y, LI Y, HU S A, et al. Study of structural periodic-like layout optimization based on guide-weight method[J]. Journal of Mechanical Engineering, 2020, 56(13): 218-230 (in Chinese).
- [ 14 ] 陈树勋, 韦齐峰, 黄锦成, 等. 利用导重法进行结构轻量化设计[J]. 工程力学, 2016, 33(2): 179-187.  
CHEN S X, WEI Q F, HUANG J C, et al. Lightening structural design using guide weight method[J]. Engineering Mechanics, 2016, 33(2): 179-187 (in Chinese).
- [ 15 ] 李枝东, 刘辛军. 导重法求解单工况的拓扑优化问题[J]. 机械工程学报, 2011, 47(15): 107-114.  
LI Z D, LIU X J. Guide-weight method on solving topology optimization problems under single load case[J]. Journal of Mechanical Engineering, 2011, 47(15): 107-114 (in Chinese).
- [ 16 ] 任毅如, 向剑辉, 何杰, 等. 基于导重法的自重载荷下悬臂梁结构拓扑优化[J]. 北京航空航天大学学报, 2021, 47(7): 1338-1344.  
REN Y R, XIANG J H, HE J, et al. Topology optimization of cantilever beam with self-weight based on guide-weight method[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(7): 1338-1344 (in Chinese).
- [ 17 ] 陈祥, 刘辛军. 基于 RAMP 插值模型结合导重法求解拓扑优化问题[J]. 机械工程学报, 2012, 48(1): 135-140.  
CHEN X, LIU X J. Solving topology optimization problems based on ramp method combined with guide-weight method[J]. Journal of Mechanical Engineering, 2012, 48(1): 135-140 (in Chinese).
- [ 18 ] ANDREASSEN E, CLAUSEN A, SCHEVENELS M, et al. Efficient topology optimization in MATLAB using 88 lines of code[J]. Structural and Multidisciplinary Optimization, 2011, 43: 1-16.
- [ 19 ] 秦浩星, 安宗文, 孙道明. 改进的导重法求解拓扑优化问题及灰度过滤技术[J]. 计算机辅助设计与图形学学报, 2015, 27(10): 2001-2007.  
QIN H X, AN Z W, SUN D M. Improved guide-weight method on solving topology optimization problems and gray-scale filtering method[J]. Journal of Computer-Aided Design & Computer Graphics, 2015, 27(10): 2001-2007 (in Chinese).
- [ 20 ] 龙凯, 赵红伟. 抑制灰度单元的改进优化准则法[J]. 计算机辅助设计与图形学学报, 2010, 22(12): 2197-2201.  
LONG K, ZHAO H W. A modified optimality criterion method for gray elements suppression[J]. Journal of Computer-Aided Design & Computer Graphics, 2010, 22(12): 2197-2201 (in Chinese).
- [ 21 ] 杜义贤, 张严, 李涵钊, 等. 具有极化特性的拓扑优化节点密度插值方法[J]. 机械设计, 2018, 35(3): 80-85.

- DU Y X, ZHANG Y, LI H Z, et al. Topology optimization of node density interpolation method with polarization properties [J]. *Journal of Machine Design*, 2018, 35 (3): 80-85 (in Chinese).
- [22] 昌俊康. 连续体结构拓扑优化设计及其在天线结构设计中的应用[D]. 西安: 西安电子科技大学, 2007.
- CHANG J K. Research on topology optimization of continuum structure and its application in design of antenna structure[D]. Xi'an: Xidian University, 2007 (in Chinese).
- [23] 高翔, 王林军, 杜义贤. 改进的抑制灰度单元的拓扑优化方法[J]. *计算机辅助设计与图形学学报*, 2020, 32 (12): 2003-2012.
- GAO X, WANG L J, DU Y X. An improved topology optimization method for suppressing gray elements[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2020, 32 (12): 2003-2012 (in Chinese).
- [24] 陈垂福, 杨晓翔. 基于步长因子改进的导重法求解拓扑优化问题[J]. *计算力学学报*, 2018, 35 (1): 14-20.
- CHEN C F, YANG X X. Based step factor improved guide-weight method for solving topology optimization problems [J]. *Chinese Journal of Computational Mechanics*, 2018, 35 (1): 14-20 (in Chinese).
- [25] 薛红军, 陶才勇, 党思娜. 基于双曲正弦函数的拓扑优化在热传导结构设计中的应用[J]. *航空工程进展*, 2020, 11 (3): 370-379.
- XUE H J, TAO C Y, DANG S N. Application of topology optimization based on sinh function in the design of heat conduction structure[J]. *Advances in Aeronautical Science and Engineering*, 2020, 11 (3): 370-379 (in Chinese).
- [26] 刘惟信. 机械最优化设计[M]. 北京: 清华大学出版社, 1994.
- LIU W X. Mechanical optimization design [M]. Beijing: Tsinghua University Press, 1994 (in Chinese).
- [27] GUEST J K, PRÉVOST J H, BELYTSCHKO T. Achieving minimum length scale in topology optimization using nodal design variables and projection functions [J]. *International Journal for Numerical Methods in Engineering*, 2004, 61 (2): 238-254.

## Gray element filtering technology of topology structure based on improved guide-weight method

GAO Xiang<sup>1,2</sup>, WANG Linjun<sup>1,2,\*</sup>, DU Yixian<sup>1,2</sup>, FU Junjian<sup>1,2</sup>

(1. Hubei Key Laboratory of Hydroelectric Machinery Design and Maintenance, China Three Gorges University, Yichang 443002, China;

2. College of Mechanical and Power, China Three Gorges University, Yichang 443002, China)

**Abstract:** Aiming at the problem of slow convergence speed when calculating Lagrange multiplier by bisection method, the proposed Lagrange multiplier calculation method is applied to the optimal criteria (OC) method and the guide-weight (GW) method to update the density, and the results of the proposed method are compared with the bisection method. The topology optimization model with the smallest compliance under the volume constraint is established. The elastic modulus of element density is calculated by the solid isotropic material with penalization (SIMP) or rational approximation of material properties (RAMP) method. The multiplier is calculated by the proposed method in this paper, and the element density is updated by the GW method. The number of gray element is reduced by the Heaviside projection function. The computational results show that; the proposed method does not significantly reduce the number of finite element analysis, but the CPU time used by the proposed method to calculate the Lagrange multiplier is less than that of the bisection method, and the number of density updates is reduced to less than 50% than before. In addition, in the two numerical examples, when SIMP model is adopted, the structure obtained by the GW method has smaller compliance than the OC method.

**Keywords:** bisection method; optimal criteria (OC) method; guide-weight (GW) method; Heaviside projection function; topology optimization

**Received:** 2021-01-02; **Accepted:** 2021-03-29; **Published online:** 2021-04-12 17:04

**URL:** [kns.cnki.net/kcms/detail/11.2625.V.20210412.1556.002.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210412.1556.002.html)

**Foundation items:** National Natural Science Foundation of China (51775308); Open Foundation of Hubei Key Laboratory of Hydroelectric Machinery Design and Maintenance (2019KJX12)

\* **Corresponding author.** E-mail: ljwang2006@126.com

# 《北京航空航天大学学报》征稿简则

《北京航空航天大学学报》是北京航空航天大学主办的以航空航天科学技术为特色的综合性自然科学学术期刊(月刊)。本刊以反映航空航天领域研究成果与动态、促进学术交流、培养科技人才和推动科技成果向社会生产力转化为办刊宗旨。本刊为中国自然科学核心期刊,并被 Ei Compendex 等国内外权威文献检索系统收录。本刊向国内外公开发行人,为进一步提高办刊质量和所刊出文章的学术水平,特制定本简则。

## 1 论文作者及内容

1.1 本刊面向海内外所有学者。

1.2 主要刊载与航空航天科学技术有关的材料科学及工程、飞行器设计与制造、宇航科学与工程、信息与电子技术、控制技术和自动化工程、流体力学和动力工程、计算机科学及应用技术、可靠性工程与失效分析等领域的研究文章。航空航天科学技术民用方面以及具有航空航天工程背景的应用数学、应用物理、应用力学和工程管理等方面的文章也在本刊优先考虑之列。

## 2 来稿要求

2.1 论文应具有创新性、科学性、学术性和可读性。

2.2 论文为原创作品,尚未公开发表过,并且不涉及泄密问题。若发生侵权或泄密问题,一切责任由作者承担。

2.3 主题明确,数据可靠,图表清晰,逻辑严谨,文字精练,标点符号正确。

2.4 文稿撰写顺序:中文题名(一般不超过 20 个汉字),作者中文姓名、单位、所在城市、邮政编码,中文摘要(包括目的、方法、结果及结论),中文关键词(5~8 个),中图分类号,英文题名,作者英文姓名、单位、所在城市、邮政编码、国别,英文摘要,英文关键词,引言,正文,参考文献。首页下角注明基金项目名称及编号,作者信息。

2.5 作者请登录本刊网页进行在线投稿。

## 3 稿件的审核、录用与版权

3.1 来稿须经专家两审和主编、编委讨论后决定刊用与否。

3.2 若来稿经审查后认定不宜在本刊发表,将及时告知作者。如果在投稿满 3 个月后仍未收到本刊任何通知,作者有权改投它刊。在此之前,请勿一稿多投,否则一切后果自负。

3.3 来稿一经刊登,即赠送单行本。

3.4 来稿一经作者签字并在本刊刊出,即表明所有作者都已经认可其版权转至本刊编辑部。本刊在与国内外文献数据库或检索系统进行交流及合作时,不再征询作者意见。

邮寄地址:100191 北京市海淀区学院路 37 号 北京航空航天大学学报编辑部

办公地点:北京航空航天大学办公楼 405,407,409 房间

电 话:(010)82315594,82338922,82314839,82315426

E-mail: jbuaa@buaa.edu.cn

http://bhxb.buaa.edu.cn

http://www.buaa.edu.cn

《北京航空航天大学学报》  
第五届编辑委员会

主任(主编): 赵沁平

(以下按姓氏笔画为序)

副主任(副主编): 丁希仑 王少萍 孙志梅 李秋实 李焕喜 杨嘉陵  
苗俊刚 相艳 徐立军 钱德沛 曹晋滨

编委: 马殿富 王琪 王聪 邓小燕 王青云 王荣明 刘宇  
刘红 江洁 刘强 闫鹏 朱天乐 刘铁钢 齐铂金  
陈万春 邹正平 苏东林 杨世春 沈成平 邱志平 宋知人  
杨树斌 张晓林 杨晓奕 杨继萍 李惠峰 吴新开 张瑞丰  
杨照华 宋凝芳 周锐 林宇震 林贵平 战强 姚仰平  
胡庆雷 赵秋红 段海滨 赵巍胜 席平 郭宏 徐洁  
徐世杰 郭洪波 康锐 翟锦 熊华钢

北京航空航天大学学报

Beijing Hangkong Hangtian Daxue Xuebao

(原《北京航空学院学报》)

(月刊 1956年创刊)

第48卷第6期 2022年6月

JOURNAL OF BEIJING UNIVERSITY OF  
AERONAUTICS AND ASTRONAUTICS  
(JBUEAA)

(Monthly, Started in 1956)

Vol.48 No.6 June 2022

主管单位 中华人民共和国工业和信息化部  
主办单位 北京航空航天大学  
主编 赵沁平  
编辑出版 《北京航空航天大学学报》编辑部  
邮编 100083  
地址 北京市海淀区学院路37号  
印刷 北京科信印刷有限公司  
发行 《北京航空航天大学学报》编辑部  
发行范围 国内外发行  
联系电话 (010) 82315594 82338922  
82314839  
电子信箱 jbuua@buaa.edu.cn

Administrated by Ministry of Industry and Information  
Technology of the People's Republic of China  
Sponsored by Beijing University of Aeronautics  
and Astronautics (BUAA)  
(Beijing 100083, P. R. China)  
Chief Editor ZHAO Qinping  
Edited and Published by Editorial Board of JBUEAA  
Printed by Beijing Kexin Printing Co., Ltd.  
Distributed by Editorial Board of JBUEAA  
Telephone (010) 82315594 82338922  
82314839  
E-mail jbuua@buaa.edu.cn  
http://bhxb.buaa.edu.cn

中国标准连续出版物号: ISSN 1001-5965  
CN 11-2625/V

国内定价: 50.00元/期

