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

ISSN 1001-5965
CN 11-2625/V

北京航空航天大學学报

JOURNAL OF BEIJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS



带软时间窗的多调度中心
UAV 安全运输线路和航迹协同规划

魏 明 孙雅茹 孙 博 王盛杰

10 2025
Vol.51

北京航空航天大学学报

第 51 卷　第 10 期　（总第 392 期）　2025 年 10 月

目　　次

带软时间窗的多调度中心 UAV 安全运输线路和航迹协同规划 …… 魏明，孙雅茹，孙博，王盛杰（3233）
基于肢体动作预测的动态可变人机协作装配 …… 吴海彬，宋晨阳，周世璇（3243）
面向点融合运行的多航空器低碳轨迹协同规划 …… 王超，李雯清，孙石磊，徐晨洋（3253）
基于 FY-3E 的星载 GNSS-R 海面高度反演模型 ……
…… 张云，鹿麒，张月维，秦甘尧，胡秀清，杨光林（3262）
基于人工势场的虚拟编组自适应模型预测控制 …… 林俊亭，倪铭君（3273）
基于子模型的复合材料主承力混杂连接结构失效分析 …… 李星，黎增山，邓凡臣，聂磊，张田（3286）
点阵结构散热特性及其轻量化评价 …… 胡建军，窦若尘，张欣，姚静，孔祥东（3299）
多角度偏振成像仪数据校正方法对比与评价 …… 王羿，孟炳寰，叶晨，骆冬根，姚萍萍，洪津（3307）
复合材料蒙皮搭接结构电气性能分析 …… 杨占刚，魏宇昊，石旭东（3313）
FAR 25 附录 C 结冰气象探测数据处理方法 ……
…… 张文强，陈奕屹，雷国强，王光雨，支亚非，毛雪瑞（3324）
改进超螺旋滑模自抗扰的 PMSM 无传感器控制 …… 陈德海，刘祥，王海峰（3337）
考虑双时间窗特性的机场多车型摆渡车调度优化 ……
…… 张文义，唐雨拉尔，王旭兰，周静，边可，刘志硕（3345）
双端异步 DoS 攻击下基于数模联动的 ICPS 多模态综合安全控制 … 李亚洁，李钢，李炜，路晨静（3354）
富氧火炬式电点火器出口火焰稳定性研究 …… 杨进慧，王希杰，徐世洋，王晓丽（3368）
CAD/CAM 模型驱动的型腔特征自适应数控工艺设计方法 ……
…… 徐昌鸿，黄瑞，张树生，梁嘉宸，卞荣（3374）
基于 PIV 的冲压发动机喷管推力的重构测量方法 …… 刘顺，孔祥瑞，徐惊雷，高波，刘海龙（3385）
同质竞争下补贴策略对多机场航线网络演化博弈 …… 吴维，林芷伊，王兴隆（3392）
基于高精度算法的二级轻气炮性能分析 …… 项高翔，伍佳豪，吴梦佳（3405）
火箭推进剂连接器在脉动风载下非合作自动对接策略 ……
…… 李泳峄，王南，吴新跃，吴鹏辉，姚博丹（3415）
基于 LSTM 和海图约束的舰船航迹预测方法 …… 赵汝健，杨威，巫震宇，曾虹程，陈杰，马雷（3424）
基于 OMEGA 理论的燃油计量装置污染卡滞特性分析 ……
…… 李文强，王彬，杨军杰，李柯柯，叶志锋（3433）
空心涡轮叶片残芯检测的双模成像方法 …… 于洋懿，杨祎罡（3443）
基于分层深度强化学习的无人机混合路径规划 …… 吕超，李慕宸，欧家骏（3451）
无人机数据采集任务中的航迹规划与资源分配优化 ……
…… 雷耀麟，丁文锐，罗祎喆，王玉峰，刘思琪，张芷兰（3460）
复杂光照环境下民用飞机飞行员关键点检测方法 …… 刘豪，孙有朝，吴红兰，张海燕（3471）
低压双流体细水雾削减三元锂离子电池热失控风险性能 …… 曾祥淑，梁家鑫，孙强，贺元骅（3482）
含滚道剥落缺陷的滚动轴承动力学建模及分析 …… 王潇，马军，熊新，陈靖宇（3492）

基于 LSTM-DDPG 算法的四翼变掠角飞行器主动变形决策 …… 彭余萧，何真，仇靖雯（3504）
不确定危险区规避神经网络多段式弹道规划方法 …… 谢蕾蕨，王旭刚，顾镇镇（3515）
运载火箭助推段大攻角剖面高精度弹道解析解 …… 张珂洋，王鹏，于琦，余文斌，陈万春（3524）
加权几何平均组合预测方法对数均方误差的分解 …… 朱双磊，陈华友，张琦，王一帆（3535）
耗氧型催化惰化系统整体式反应器性能研究 ……
…… 郭晨阳，刘祎，刘豪正，王俊杰，高经诚，冯诗愚（3547）
基于改进一致性算法的无人机编队队形保持控制器设计 ……
…… 万茹，马梓元，龚华军，王新华，张帅（3555）
微型直接甲醇燃料电池多级退化参量融合预测 …… 董杰，苏雨临，张大骋（3567）
一种考虑滑流效应的螺旋桨设计方法及应用 …… 陈胜久，杨佑绪，张兴翠，吴逸飞（3578）
基于注意力机制改进的 PSO-BiLSTM 刀具磨损预测 …… 杨沛东，黄华，尉卫卫，郭宝岛（3589）

期刊基本参数：CN 11-2625/V * 1956 * m * A4 * 366 * zh * P * ￥50.00 * 400 * 36 * 2025-10

（**编　辑**　张　嵘　孙　芳　万玉敏　卞欢欢　王　茜　李亚泰）

JOURNAL OF BEIJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS

Vol. 51 No. 10 (Sum 392) October 2025

CONTENTS

Cooperative planning for safe transportation routes and flight paths of UAV with multiple dispatching centers and soft time windows ····· WEI Ming, SUN Yaru, SUN Bo, WANG Shengjie (3233)
Dynamically changeable human-robot collaborative assembly based on limb motion prediction ····· WU Haibin, SONG Chenyang, ZHOU Shixuan (3243)
Multi-aircraft low-carbon trajectories cooperative planning for point merge operation ····· WANG Chao, LI Wenqing, SUN Shilei, XU Chenyang (3253)
Spaceborne GNSS-R sea surface height inversion model using FY-3E ····· ZHANG Yun, LU Qi, ZHANG Yuewei, QIN Ganyao, HU Xiuqing, YANG Guanglin (3262)
Adaptive model predictive control of virtual coupled based on artificial potential field ····· LIN Junting, NI Mingjun (3273)
Sub-model based failure analysis of composite primary load-bearing hybrid joint structure ····· LI Xing, LI Zengshan, DENG Fanchen, NIE Lei, ZHANG Tian (3286)
Heat dissipation characteristics and lightweight evaluation of lattice structure ····· HU Jianjun, DOU Ruochen, ZHANG Xin, YAO Jing, KONG Xiangdong (3299)
Comparison and evaluation of data correction methods for directional polarimetric cameras ····· WANG Yi, MENG Binghuan, YE Chen, LUO Donggen, YAO Pingping, HONG Jin (3307)
Electrical properties analysis of composite materials skin bonding structures ····· YANG Zhangang, WEI Yuhao, SHI Xudong (3313)
Data processing methodology of icing meteorological detection in FAR 25 Appendix C ····· ZHANG Wenqiang, CHEN Yiyi, LEI Guoqiang, WANG Guangyu, ZHI Yafei, MAO Xuerui (3324)
Modified sensorless control of PMSM based on super-twisting sliding mode auto disturbance rejection ····· CHEN Dehai, LIU Xiang, WANG Haifeng (3337)
Optimization of multi-type airport ferry vehicles schedules with double service time windows characteristic ····· ZHANG Wenyi, TANG Yulaer, WANG Xulan, ZHOU Jing, BIAN Ke, LIU Zhishuo (3345)
ICPS multi-modal integrated security control based on data-model linkage under dual-end asynchronous DoS attacks ····· LI Yajie, LI Gang, LI Wei, LU Chenjing (3354)
Study on flame stability of oxygen-rich torch igniter ····· YANG Jinhui, WANG Xijie, XU Shiyang, WANG Xiaoli (3368)
Adaptive NC process planning approach for pocket features driven by CAD/CAM models ····· XU Changhong, HUANG Rui, ZHANG Shusheng, LIANG Jiachen, BIAN Rong (3374)
Reconstruction measurement method for ramjet nozzle thrust based on PIV ····· LIU Shun, KONG Xiangrui, XU Jinglei, GAO Bo, LIU Hailong (3385)
Evolutionary game of subsidy strategy on multi-airport route network under homogeneous competition ····· WU Wei, LIN Zhiyi, WANG Xinglong (3392)
Performance analysis of two-stage light gas gun based on higher-order code ····· XIANG Gaoxiang, WU Jiahao, WU Mengjia (3405)
Non-cooperative automatic docking strategy of propellant umbilical connector for launch vehicle under fluctuating wind load ····· LI Yongyi, WANG Nan, WU Xinyue, WU Penghui, YAO Bodan (3415)
Ship track prediction method based on LSTM and nautical chart constraints ····· ZHAO Rujian, YANG Wei, WU Zhenyu, ZENG Hongcheng, CHEN Jie, MA Lei (3424)
Contamination lock characteristics analysis of fuel metering unit based on OMEGA theory ····· LI Wenqiang, WANG Bin, YANG Junjie, LI Keke, YE Zhifeng (3433)
Bi-modal imaging method for detection of residual core in hollow turbine blades ····· YU Yangyi, YANG Yigang (3443)
UAV hybrid path planning based on hierarchical deep reinforcement learning ····· LYU Chao, LI Muchen, OU Jiajun (3451)
Trajectory planning and resource allocation optimization in UAV data collection missions ····· LEI Yaolin, DING Wenrui, LUO Yizhe, WANG Yufeng, LIU Siqi, ZHANG Zhilan (3460)
Key points detection method for civil aircraft pilot in complex lighting environments ····· LIU Hao, SUN Youchao, WU Honglan, ZHANG Haiyan (3471)
Performance of low-pressure dual-fluid water mist to reduce thermal runaway risk of ternary lithium-ion battery ····· ZENG Xiangshu, LIANG Jiaxin, SUN Qiang, HE Yuanhua (3482)
Dynamic modeling and analysis of rolling bearing with raceway spalling defect ····· WANG Xiao, MA Jun, XIONG Xin, CHEN Jingyu (3492)

Active deformation decision-making for four-wing variable sweep aircraft based on LSTM-DDPG algorithm
…… PENG Yuxiao，HE Zhen，QIU Jingwen （3504）
Multi-stage trajectory planning method for hazard zone avoidance under uncertainty based on neural networks
…… XIE Fanwei，WANG Xugang，GU Zhenzhen （3515）
Analytical solutions for ascent trajectory under high angle-of-attack conditions
…… ZHANG Keyang，WANG Peng，YU Qi，YU Wenbin，CHEN Wanchun （3524）
Decomposition of logarithm mean square error of weighted geometric mean combined forecasting method
…… ZHU Shuanglei，CHEN Huayou，ZHANG Qi，WANG Yifan （3535）
Research on oxygen consumption based inerting monolithic catalyst reactor performance
…… GUO Chenyang，LIU Yi，LIU Haozheng，WANG Junjie，GAO jingcheng，FENG Shiyu （3547）
Design of UAV formation-keeping controller based on improved consistency algorithm
…… WAN Ru，MA Ziyuan，GONG Huajun，WANG Xinhua，ZHANG Shuai （3555）
Ensemble-based prediction using multi-level degradation parameters for micro direct methanol fuel cells
…… DONG Jie，SU Yulin，ZHANG Dacheng （3567）
A propeller design method considering slipstream effect and its application
…… CHEN Shengjiu，YANG Youxu，ZHANG Xingcui，WU Yifei （3578）
Tool wear prediction based on attention mechanism and PSO-BiLSTM
…… YANG Peidong，HUANG Hua，YU Weiwei，GUO Baodao （3589）

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0509

带软时间窗的多调度中心UAV安全运输线路和航迹协同规划

魏明[1,2,*]，孙雅茹[1]，孙博[1]，王盛杰[2]

(1. 中国民航大学 空中交通管理学院，天津 300300；2. 中国民用航空飞行学院 民航飞行技术与飞行安全重点实验室，广汉 618300)

摘　　要：针对面向物流配送的无人机(UAV)运输线路和航迹规划问题，建立了一种协同规划双层模型。上层模型考虑客户时间窗、无人机载重、能耗和路径风险等约束因素，计算无人机的出发时间、调度中心及其访问客户的顺序和调度时间表，实现无人机调度成本最低；下层模型考虑障碍物、无线电干扰与无人机坠落代价等多重安全因素，计算无人机在任意调度中心和客户之间可行的最短飞行航迹。根据问题特征，设计求解该问题的嵌入A*算法和贪婪策略的两阶段深度强化学习(DRL)算法。通过案例计算最佳的无人机运输线路及其航迹规划方案，分析关键参数的变化对调度结果的影响，并与遗传算法(GA)、差分进化(DE)算法及粒子群优化(PSO)算法进行对比，验证所提算法的有效性和正确性。

关　键　词：航空运输；无人机运输线路；无人机航迹；双层模型；深度强化学习；A*算法

中图分类号：U8

文献标志码：A　　**文章编号**：1001-5965(2025)10-3233-10

与传统车辆配送方式相比，无人机(unmanned aerial vehicle, UAV)能够突破交通拥堵和地形等地面障碍物的时空制约，高效地将货物从起点运送到目的地，已成为城市物流配送领域的发展趋势之一。面向物流配送的无人机运输核心任务主要有：①根据客户和调度中心的时空分布，考虑无人机的性能约束，制定无人机运输线路和运行计划，实现运输成本最低；②根据障碍物空间分布，寻找无人机在任意客户点和调度中心之间的可行安全飞行通道，实现航迹最短。显然，无人机运输线路和航迹规划之间相互影响，局部最优，整体未必最佳。鉴于该问题属于车辆路径问题的拓展，为一类NP-hard问题，引起了国内外学者的广泛关注。

目前，现有文献分别探索无人机运输路线和航迹规划问题。前者主要借鉴传统的车辆路径问题，考虑无人机的性能差异(如能耗、安全和载重等)，以物品运输交付时间成本最小[1-2]、服务覆盖客户数量最大[3]、整体运输成本最小[4]等为目标，探讨多行程无人机配送线路设计与交付时间表优化问题，并利用多种算法求解，包括改进退火算法[1]、改进遗传算法(genetic algorithm, GA)[2]、启发式算法[3]、精确算法[4]等。后者一般分为二维[5-7]和三维[8-16]航迹规划，考虑障碍物及威胁分布[6-8,10,13-15]、能耗[8,11]、坠地风险[9,11]和无冲突飞行[16]等因素，以任务完成效率最高[5]、任务完成时间最短[12]、运行总成本最小[7-9,11,16]、飞行距离最短[13-14]、能量消耗最小[15]等为目标，探讨单[5-7,10-11,13-14]、多[7-9,12-13,15-16]无人机协同路径规划问题，求解算法一般分为深度学习[5,14-15]、启发式算法[6,11,16]和群智能算法[7-10,12-13]，其中，启发式算法包括改进快速搜索随机树算法[6]、A*-人工势场算法[11]、稀疏A*算法[16]，群智能算法包括改进蝙蝠算法[7]、灰狼算法[8]、改进粒子群优化(particle swarm

收稿日期：2023-08-04；录用日期：2023-09-03；网络出版时间：2024-01-08 08:26
网络出版地址：link.cnki.net/urlid/11.2625.V.20240105.1557.004
基金项目：民航飞行技术与飞行安全重点实验室开放基金(FZ2021KF06)；教育部人文社科项目(20YJCZH176)
*通信作者. E-mail：mwei@cauc.edu.cn

引用格式：魏明，孙雅茹，孙博，等. 带软时间窗的多调度中心UAV安全运输线路和航迹协同规划[J]. 北京航空航天大学学报，2025，51(10)：3233-3242. WEI M，SUN Y R，SUN B，et al. Cooperative planning for safe transportation routes and flight paths of UAV with multiple dispatching centers and soft time windows[J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51(10)：3233-3242 (in Chinese).

optimization, PSO)算法[9]、生物启发路径规划算法[10]、蚁群算法[12]、混合灰狼优化和微分进化算法[13]。

可知,现有研究不足之处主要有:①多数将无人机运输线路设计、调度中心选取、出发时间和航迹规划分开研究,亟待寻求它们之间的耦合关系;②尚未挖掘无人机性能、线路与航迹规划、多重安全因素、运输成本之间的内在联系。综上,本文针对无人机物流配送问题,研究基于深度强化学习(deep reinforcement learning, DRL)算法的无人机运输线路和航迹协同规划的双层模型。上层模型考虑无人机的性能,利用深度强化学习算法求解带时间窗的多调度中心无人机运输线路问题;下层模型综合考虑坠落风险和无线电干扰等多重安全约束因素,利用A*算法求解任意客户点与调度中心之间的可行航迹,据此,基于贪婪算法确定无人机的出发地和时间。通过数值仿真,验证了本文模型和算法的有效性。

1 无人机运输线路和航迹协同规划双层模型

面向物流配送的无人机运输线路和航迹规划都属于NP-hard的路径规划问题,两者之间相互影响,即无人机运输线路一旦确定,据此可以生成无人机航迹;同时,无人机在任意调度中心与客户之间的不同航迹也影响无人机运输线路走向。显然,考虑客户的时空分布、无人机的性能和安全等约束,两者局部最优,未必整体最优。因此,本文借助双层规划技术探讨面向物流配送的无人机运输线路与航迹规划之间的耦合关系。

1.1 上层模型:带时间窗的多调度中心无人机运输线路规划

在上层模型中,根据客户的位置、时间窗和需求量,安排多架无人机从不同调度中心出发和返回,确定各无人机访问客户的顺序,并计算出发和到达时间。当无人机提前或滞后访问客户时,设置一定惩罚费用,且无人机的能耗取决于其当前的载重和飞行里程。考虑无人机的性能、安全风险等因素,在需求不可拆分、静态需求、无人机速度和飞行高度固定的假设前提下,构建无人机运输线路规划涉及路径安排、调度中心选取、出发时间计算的混合整数规划模型,实现调度成本最低。

根据问题描述,建立上层模型的混合整数线性规划模型,如下:

$$\min \sum_{s\in S}\sum_{k\in K_s}\left\{\left[\sum_{i,j\in I\cup S}\left(c_5F\left(t_i^k,d_{ij}^k\right)+c_3\sum_{i,j\in I\cup S}d_{ij}^k\right)x_{ij}^k\right]+c_4\sum_{i\in I}x_{si}^k+\sum_{i\in I}y_i^k\left(c_1\cdot\max\left(t_i^s-t_i^k,0\right)+c_2\cdot\max\left(t_i^k-t_i^e,0\right)\right)\right\} \tag{1}$$

约束条件如式(2)~式(4)。

$$\sum_{s\in S}\sum_{k\in K_s}y_i^k=1 \qquad i\in I \tag{2}$$

$$\sum_{i\in I}x_{si}^k=\sum_{i\in I}x_{is}^k=1 \qquad s\in S,k\in K_s \tag{3}$$

$$\sum_{i,j\in I\cup S}x_{ij}^k=\sum_{i,j\in I\cup S}x_{ji}^k=y_i^k \qquad k\in K_s,s\in S \tag{4}$$

$$t_i^k+\frac{d_{ij}^k}{v}-\left(1-x_{ij}^k\right)M\geqslant t_j^k \qquad i,j\in I\cup S,k\in K_s,s\in S \tag{5}$$

$$t_i^k+\frac{d_{ij}^k}{v}+\left(1-x_{ij}^k\right)M\leqslant t_j^k \qquad i,j\in I\cup S,k\in K_s,s\in S \tag{6}$$

$$t_s^k=T_s^k \qquad s\in S,k\in K_s \tag{7}$$

$$l_i^k-q_i-\left(1-x_{ij}^k\right)M\geqslant l_j^k \qquad i,j\in I\cup S,k\in K_s,s\in S \tag{8}$$

$$l_i^k-q_i+\left(1-x_{ij}^k\right)M\leqslant l_j^k \qquad i,j\in I\cup S,k\in K_s,s\in S \tag{9}$$

$$\sum_{i\in I}y_i^kq_i\leqslant l_{\max} \qquad s\in S,k\in K_s \tag{10}$$

$$p_i^k-F\left(l_i^k,d_{ij}^k\right)-\left(1-x_{ij}^k\right)M\geqslant p_j^k \qquad i,j\in I\cup S,k\in K_s,s\in S \tag{11}$$

$$p_i^k-F\left(l_i^k,d_{ij}^k\right)+\left(1-x_{ij}^k\right)M\leqslant p_j^k \qquad i,j\in I\cup S,k\in K_s,s\in S \tag{12}$$

$$p_i^k\geqslant 0 \qquad i\in I\cup S,k\in K_s,s\in S \tag{13}$$

$$\sum_{i,j\in I\cup S}F\left(l_i^k,d_{ij}^k\right)x_{ij}^k\leqslant p_{\max} \qquad k\in K_s,s\in S \tag{14}$$

式中:I为客户集合;S为调度中心集合;K_s为任意调度中心s的无人机集合;x_{ij}^k为决策变量,表示无人机k是否依次访问客户或调度中心i和j;y_i^k为决策变量,表示无人机k是否访问客户i;T_s^k为决策变量,表示无人机k在调度中心s的出发时间;q_i为客户i需求量;$[t_i^s,t_i^e]$为客户i时间窗;t_i^k、l_i^k和p_i^k分别为无人机k到达客户i的时间、载重和电量;$l_{\max}$和$p_{\max}$分别为无人机的最大载重和最大电量;d_{ij}^k为无人机k从客户或调度中心i到达j的飞行距离;v为无人机飞行速度;$F\left(l_i^k,d_{ij}^k\right)$为无人机$k$在节点$i$以载重$l_i^k$飞行航程$d_{ij}^k$到达节点$j$的能耗,计算公式参考文献[1];$c_1$和$c_2$分别为提前和滞后服务客户的单位时间惩罚费用;$c_3$和$c_4$分别为无人机的单位里程费用和固定费用;$c_5$为电价;$M$为一个较大的常数。

在上层模型中,目标函数(1)追求无人机的运输费用最少,包括机队规模成本、能耗成本、里程成本和时间窗惩罚成本。式(2)~式(14)为约束条件,其中,式(2)将每个客户指派给唯一无人机;式(3)表示每架无人机从某调度中心出发并返回;

式(4)表示某客户仅能被一架无人机访问, 一架无人机同时只能访问一个客户; 式(5)和式(6)表示无人机途径相邻节点(客户或调度中心)的达到时间关系; 式(7)表示无人机从调度中心的出发时间; 式(8)～式(10)表示无人机途径相邻节点的载重变化, 并限制其不能超过最大值; 式(11)～式(13)表示无人机到达相邻节点的电能变化关系, 并确保有电; 式(14)表示无人机的总能耗不超过其上限。

1.2　下层模型: 考虑多重安全因素的无人机航迹规划

针对上层模型的无人机途径任意调度中心i与客户j的顺序确定时($x_{ij}^k=1$), 将其周边空域根据地理信息进行网格化(每个网格相当于一个航路点), 假设无人机在固定高度层飞行并沿着网格中心线飞行, 可以提前获取每个网格的障碍物、坠机代价、无线电干扰信号强度。根据无人机k的起点i和终点j, 考虑多重安全因素, 构建无人机航迹规划的 0-1 整数规划模型, 计算一条最短航程的航迹d_{ij}^k。

根据上述问题, 建立下层模型的 0-1 整数线性规划模型如下:

$$d_{ij}^k=\min\sum_{m,n\in A}z_{mn}^k d_{mn} \tag{15}$$

约束条件如式(16)～式(17)所示。

$$\sum_{n\in A}z_{mn}^k-\sum_{n\in A}z_{nm}^k=\begin{cases}1 & m=i\\ -1 & m=j\\ 0 & m\neq i,j\end{cases} \tag{16}$$

$$z_{mn}^k\leqslant a_{mn}\qquad m,n\in A \tag{17}$$

式中: A为网格集合; z_{mn}^k为决策变量, 表示无人机k是否依次访问相邻网格 m 和 n; d_{mn}为网格 m 和 n 之间距离; a_{mn}为考虑多重安全因素条件下的任意 2 个相邻网格m和n之间的连通性, 其计算过程为

$$a_{mn}=\begin{cases}1 & b_m=b_n=0\wedge f_m\leqslant R\wedge f_n\leqslant R\wedge w_m\leqslant W\wedge w_n\leqslant W\\ 0 & b_m=1\vee b_n=1\vee f_m\geqslant R\vee f_n\geqslant R\vee w_m\geqslant W\vee w_n\geqslant W\end{cases}\qquad m,n\in A \tag{18}$$

其中: b_m表示网格m是否为障碍物; f_m为网格m的无人机坠毁代价, 其值取决于无人机坠毁率、地面保护系数、网格面积及人口密度[17]; w_m为网格 m 受无线电信号干扰强度, 取决于网格附近的多个信号源功率r_l(信号源数量为L), 即$w_m=\sum_{l=1}^{L}\frac{r_l}{4\pi d_{ml}^2}$; R和W分别为无人机坠毁代价和无线电信号干扰阈值。

在下层模型中, 式(15)是目标函数, 表示极小化某无人机在起讫点之间最短可行航迹。式(16)和式(17)是约束条件, 其中, 式(16)表示无人机航迹依次途径 2 个网络的流量守恒约束; 式(17)表示无人机航迹只能途径可连通的相邻网格。

2　基于深度强化学习的两阶段算法

根据上下层模型之间关系, 当上层模型的无人机配送线路确定后, 据此可以确定下层模型的无人机航迹。鉴于精确算法无法在短时间内求解大规模的上下层 NP-hard 问题, 本文设计一种嵌入 A*算法和贪婪算法的两阶段深度强化学习求解框架, 如图 1 所述。第 1 阶段, 利用改进指针网络构造上层模型中的无人机运输线路; 第 2 阶段, 利用 A*算法搜索无人机访问任意节点之间的飞行通道, 并基于贪婪算法确定无人机的出发地点和时间。基于上述过程, 引入 REINFORCE 算法进行模型训练, 通过 actor-critic 网络计算模型评价函数, 不断寻优从而找到更优解。

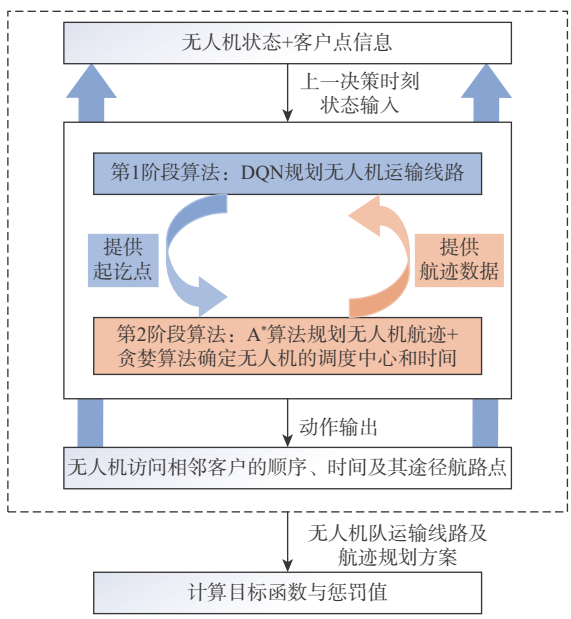


图 1　基于深度强化学习的两阶段求解框架

Fig. 1　Two-stage solution framework based on deep reinforcement learning

2.1　第 1 阶段算法: 深度强化学习算法

根据第 1 阶段无人机运输线路规划的问题特征, 可以将其描述为马尔可夫决策过程, 在此基础上, 利用深度 Q 网络(deep Q-network, DQN)的指针网络构造问题的解, 并结合强化学习进行模型训练。求解组合优化问题的指针网络一般包含循环神经网络(recurrent neural network, RNN)编码器和解码器, 多用于处理静态系统的问题。而本文提到

的无人机运输线路规划问题中，客户点的需求是动态变化的，即一旦客户点需求被满足，其值将会变为0，为克服这一问题，Nazari等[18]改进指针网络结构，如图2所示。为简化计算，忽略RNN编码过程，将客户点的位置、需求及时间窗信息输入嵌入层，映射得到初始高维特征向量，并在RNN解码器中储存解码序列信息，基于上下文的注意力机制，从输入中提取可变长度向量$\boldsymbol{a}^{\omega}$和$\boldsymbol{c}^{\omega}$，计算每个客户节点的选择概率，在每个解码时刻ω选择概率最大的节点加入解中，从而构造多架无人机访问全部客户的顺序。

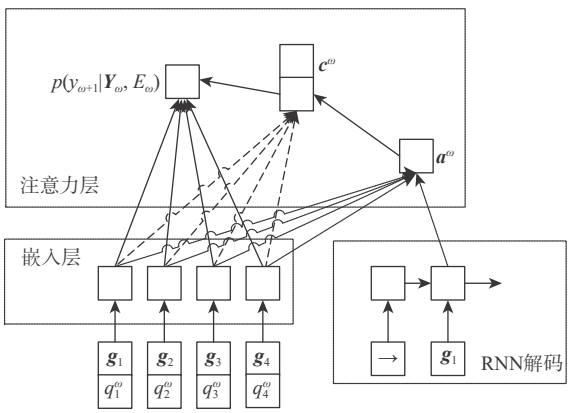


图2 指针网络结构

Fig. 2 Pointer network structure

基于深度强化学习算法求解无人机运输线路规划问题的思想在于：将其抽象为动态规划问题，把每个阶段根据状态和序列的变化与深度强化学习算法的策略、状态、动作和奖励等映射，通过不确定性环境对选择动作的评价来指导学习，不断试错和选择，进而逐步改进动作选择策略，最终生成一个适于当前任务环境的无人机配送线路方案。具体如下：

1）输入状态。将输入隐藏层的状态空间表示为$E_{\omega}=\{\boldsymbol{G}_i^{\omega}, i=1,2,\cdots,I\}$，该集合包括客户在每个决策时刻$\omega(\omega\in\Omega)$的位置$\boldsymbol{g}_i$及其对应的需求量$q_i^{\omega}$，即$\{\boldsymbol{G}_i^{\omega}=(\boldsymbol{g}_i,q_i^{\omega}), i\in I\}$，$\boldsymbol{g}_i$为客户位置信息；同时，由RNN解码器对状态集合中的客户位置进行解码并储存序列，随着决策的步进，状态转移函数如下：

$$\begin{cases} q_i^{\omega+1}=\max\left(0,q_i^{\omega}-l_i^{\omega}\right) \\ q_j^{\omega+1}=q_j^{\omega} \qquad j\neq i \\ l_i^{\omega+1}=\max\left(0,l_i^{\omega}-q_i^{\omega}\right) \end{cases} \qquad i,j\in I,\omega\in\Omega \tag{19}$$

式中：第1个和第2个情形表示无人机访问被选中客户点和未选中客户点的需求信息更新；第3个情形表示无人机的载重量在每一次决策后进行信息更新。

2）动作输出。当无人机在决策时刻选择节点时，基于隐藏层映射到高维空间产生的表征向量和RNN解码序列，结合注意力机制，产生一个无人机访问任意节点的概率分布，在每一次动作都选择概率最大的节点，以概率链规则来阐述在初始状态生成的解序列$\boldsymbol{Y}$的概率$P(\boldsymbol{Y}|E_0)$，计算方法如下：

$$P(\boldsymbol{Y}|E_0)=\prod_{\omega\in\Omega}p(y_{\omega+1}|\boldsymbol{Y}_{\omega},E_{\omega}) \tag{20}$$

$$E_{\omega+1}=f(y_{\omega+1},E_{\omega}) \tag{21}$$

式(20)来自于注意力机制，即

$$p(y_{\omega+1}|\boldsymbol{Y}_{\omega},E_{\omega})=\mathrm{softmax}(\xi(h_{\omega},E_{\omega})) \tag{22}$$

式中：y_{ω}为每一步的动作选择结果；E_{ω}为在ω时刻下所有输入的状态集合；$\xi(h_{\omega},E_{\omega})$表示仿射函数；$h_{\omega}$表示在进行动作决策之前的RNN解码器状态。

3）奖励函数。根据目标函数，采取梯度上升的方式寻找近最优解，将智能体的收益Λ定义为无人机队运行成本的相反数，即

$$\begin{aligned}\Lambda=-\sum_{s\in S}\sum_{k\in K_s}\Bigg\{\Bigg[\sum_{i,j\in I\cup S}\left(c_5F\left(l_i^k,d_{ij}^k\right)+c_3\sum_{i,j\in I\cup S}d_{ij}^k\right)x_{ij}^k\Bigg]+\\ c_4\sum_{i\in I}x_{si}^k+\sum_{i\in I}y_i^k\Bigg(c_1\cdot\max\left(t_i^s-t_i^k,0\right)+\\ c_2\cdot\max\left(t_i^k-t_i^e,0\right)\Bigg)\Bigg\}-\varpi M\end{aligned}$$

式中：$M\gg c_1,c_2,c_3,c_4,c_5$；ϖ为无人机超出其续航能力运行记号。

4）终止条件。在构造无人机运输线路问题解的阶段，动作决策的终止条件是没有未满足需求的客户节点。

2.2 第2阶段算法：A*算法与贪婪算法

基于航路点的GIS信息及无线电干扰情况，结合式(18)，计算避开障碍物、高无线电信号干扰和坠地风险区域的可行航路网，获取无人机在相邻网格之间的连通情况。在此基础上，当第1阶段算法找到无人机访问任意客户之间顺序后，基于A*算法的最短路思想搜寻无人机在客户点之间的可行飞行航迹，该算法以客户点坐标为起点和终点，将航路网中的网格节点作为搜索节点，通过优先扩展在可行航路网中距离最近的节点，快速搜寻出满足安全飞行要求的最短可行航迹。利用贪婪策略确定无人机的出发调度中心调度中心和时间，即从调度中心集合中，依次计算各中心至首个客户点的航迹长度及所需飞行时间，选择总初始航时最短的中心作为出发地；同时，根据首个客户点的服务时间窗口及无人机电池续航能力确定具体出发时刻。

2.3 基于两阶段深度强化学习算法的训练过程

引入REINFORCE算法，该算法以反向传播作为策略参数更新方法，智能体在探索解空间时，用θ参数化随机策略，使用对策略参数期望回报的梯度估计来迭代地改进策略，主要训练过程如图3所示。根据随机概率分布生成问题实例，抽取N个样

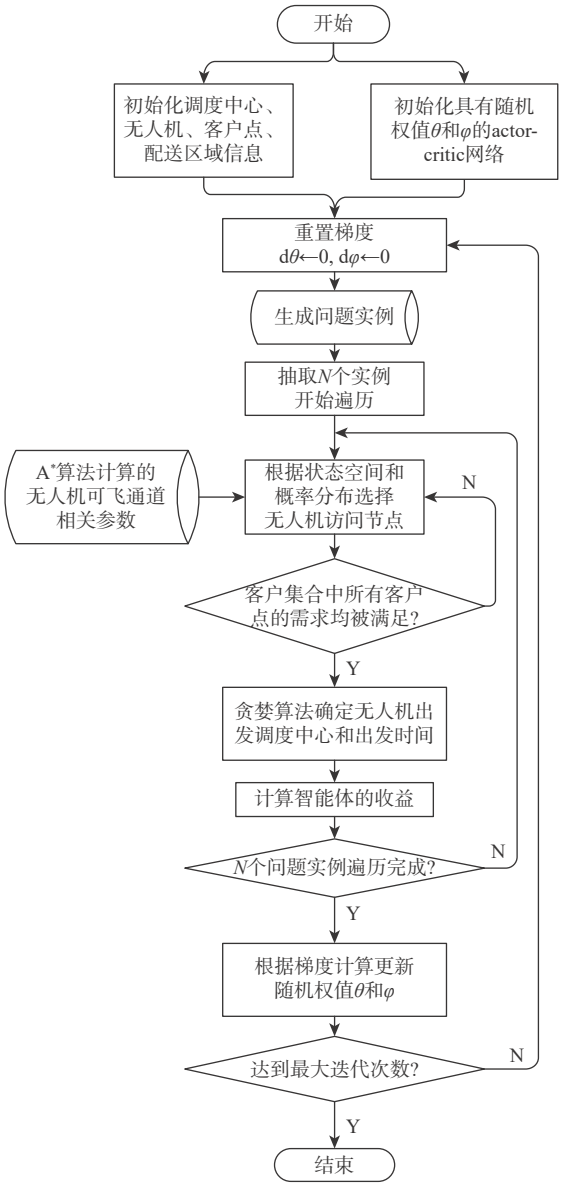


图 3　模型训练的算法流程

Fig. 3　Algorithm flow for model training

本问题, 并使用蒙特卡罗模拟生成相对于当前策略的可行解, 在所有 N 个问题的解码终止后, 利用 actor-critic 网络计算相应的收益 Λ 和策略梯更新, 从而更新 actor-critic 网络。

3　数值实验

3.1　算例基本情况

如图 4 所示, 将 6 km×6 km 的区域按照 200 m 划分为若干个栅格, 其中分布 27 个客户(D1～D27)、3 个调度中心(S1～S3)和 2 个无线电干扰源。客户信息如表 1 所示, 涉及位置、需求量和时间窗, 调度中心位置分别为 S1(8,8), S2(15,25), S3(26,4), 如 S2(15,25) 含义为调度中心 S2 位于坐标为 (15,25) 的栅格中。2 个无线电干扰源位于 (7,22) 和 (19,7), 其功率分别为 50 kW 和 70 kW, 黑色栅格表示受无线电信号干扰区域。每个栅格的无人机坠落代价根

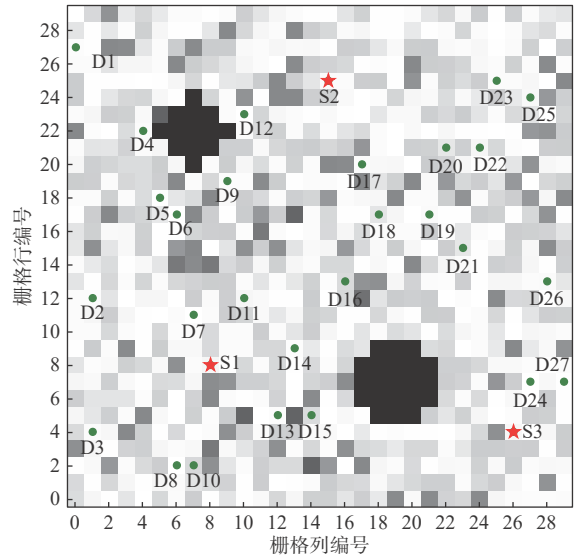


图 4　考虑障碍物分布、坠落代价和无线电干扰的配送区域 GIS 栅格化

Fig. 4　GIS rasterization of distribution area considering obstacle distribution, crash cost, and radio interference

表 1　客户基本信息

Table 1　Customer basic information

需求点(坐标)	最早时间	最晚时间	需求量/kg
D1(0,27)	06:24	06:29	1
D2(1,12)	06:20	06:25	2
D3(1,4)	06:00	06:05	1
D4(4,22)	06:26	06:31	3
D5(5,18)	06:27	06:32	1
D6(6,17)	06:27	06:32	2
D7(7,11)	06:20	06:25	1
D8(6,2)	06:02	06:07	1
D9(9,19)	06:12	06:17	1
D10(7,2)	06:02	06:07	2
D11(10,12)	06:15	06:21	1
D12(10,23)	06:10	06:15	1
D13(12,5)	06:03	06:08	1
D14(13,9)	06:03	06:08	2
D15(14,5)	06:20	06:25	3
D16(16,13)	06:16	06:21	1
D17(17,20)	06:18	06:23	1
D18(18,17)	06:17	06:22	2
D19(21,17)	06:30	06:35	2
D20(22,21)	06:35	06:40	1
D21(23,15)	06:31	06:36	2
D22(24,21)	06:33	06:38	1
D23(25,25)	06:34	06:39	1
D24(27,7)	06:26	06:31	2
D25(27,24)	06:34	06:39	1
D26(28,13)	06:25	06:30	2
D27(29,7)	06:27	06:32	2

据其人口与地面遮盖物的计算获取，灰色程度表示坠落代价。采用设计速度 16 m/s、最大续航 15 km、自重 36 kg 的六旋翼物流无人机进行配送，其他模型参数为：c_1为 2 元/min，c_2为 100 元/次，c_3为 1.5 元/km，c_4为 2 000 元/架，c_5为 2 元/(kW·h)。

3.2 最佳方案

采用 Pytorch10.2 和 Python3.9 编写两阶段深度强化学习算法的程序，运行环境为单张 GPU（RTX3080，显存 12 GB），在模型训练阶段设置 epoch 为 30 轮，每轮 epoch 的训练规模为 10 000 次，每次训练抓取 256 个样本，使用学习率为 1×10^{-4} 的 Adam 优化器[19]进行模型训练。当 R=1.4 时，求解上下层模型的解寻优迭代曲线，如图 5 所示，最佳无人机运输线路如表 2 和图 6 所示，与之对应的航迹如图 7 所示。可知：①共需 5 架无人机为 27 个客户运输货物，滞后和等待时间分别为 0 min 和 44.66 min，总配送成本为 10 163.249 元，包括固定成本 10 000 元，飞行损耗成本 66.765 元，能耗成本 7.108 元，无人机等待成本 89.376 元。②以无人机 UAV2 为例，其运输线路为：S1—D3—D8—D10—D2—D7—S1，出发时间为 06:03:00，相应的航迹为 S1[(8,8), (7,7), (6,7), (5,7), (4,6),(3,5),(2,4),(1,4)]—D3[(1,4),(2,3),(3,2),(4,2),(5,1), (6,2)]—D8[(6,2),(7,2)]—D10[(7,2),(6,3),(5,4),(5,5), (4,6),(3,7),(2,8),(2,9),(1,10),(1,11),(1,12)]—D2[(1,12), (2,12),(3,12),(4,12),(5,12),(6,12),(7,11)]—D7[(7,11),(7, 10),(7,9),(8,8)]—S1(8,8)。

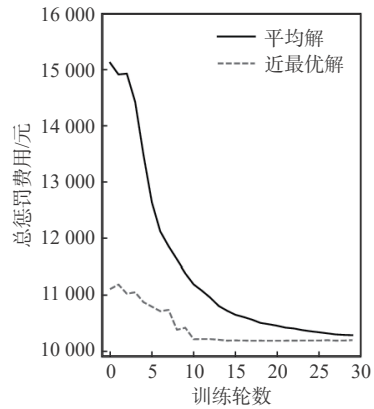


图 5 基于深度强化学习算法的寻优过程曲线

Fig. 5 Optimization process curve based on deep reinforcement learning algorithm

表 2 最佳调度方案

Table 2 The optimal scheduling scheme

无人机	线路及时间规划	航程/km	滞后时间/min	等待时间/min	能耗/(kW·h)
UAV1	S1(06:05:00)—D13(06:06:06)—D14(06:07:07)—D9(06:09:37)—D16(06:14:13)—D11(06:17:20)—D15(06:19:08)—S1(06:21:30)	11.022	0	5.00	0.900
UAV2	S1(06:03:00)—D3(06:04:48)—D8(06:06:01)—D10(06:06:13)—D2(06:08:50)—D7(06:21:20)—DS1(06:30:42)	7.560	0	19.83	0.596
UAV3	S2(06:20:00)—D17(06:21:13)—D18(06:21:56)—D20(06:23:14)—D22(06:35:25)—D25(06:36:18)—D23(06:36:48)—S2(06:39:04)	6.994	0	11.78	0.534
UAV4	S2(06:13:00)—D12(06:14:13)—D1(06:16:39)—D5(06:26:18)—D6(06:27:18)—D4(06:28:31)—DS2(06:31:21)	9.891	0	8.05	0.798
UAV5	S3(06:26:00)—D27(06:27:00)—D26(06:28:21)—D24(06:29:41)—D19(06:32:17)—D21(06:32:52)—DS3(06:35:25)	9.043	0	0	0.726

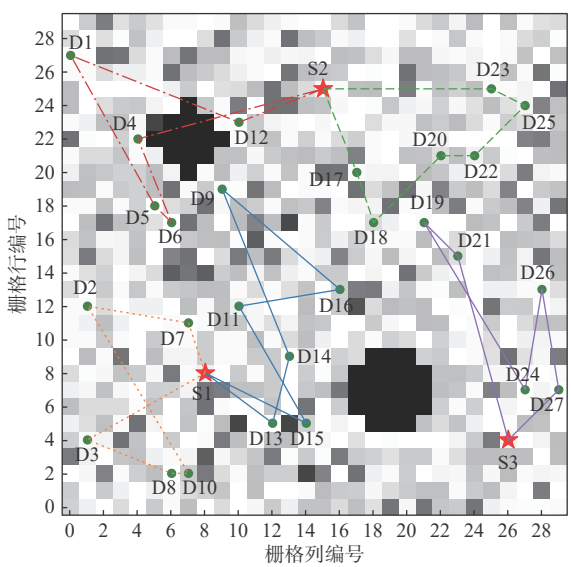


图 6 上层模型的无人机队运输线路可视化

Fig. 6 Visualization of UAV transport routes in upper-level model

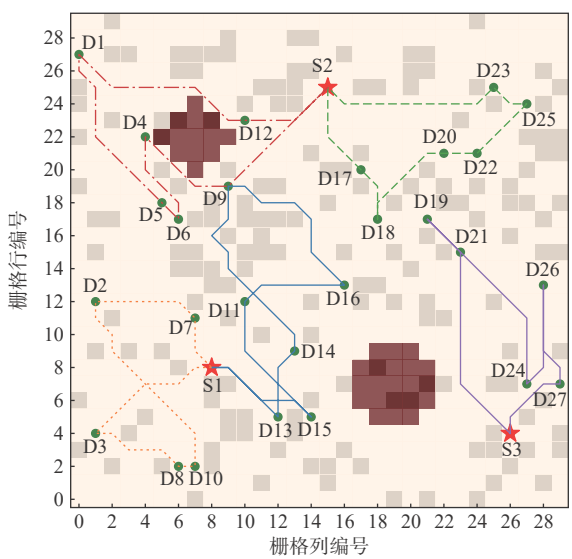


图 7 下层模型的无人机队航迹可视化

Fig. 7 Visualization of UAV flight paths in lower-level model

3.3　灵敏度分析

在上述基础上, 进一步分析不同坠落代价阈值、固定/灵活出发时间、单/多调度中心、无线电干扰源布局对配送方案的影响, 如图 8～图 11 所示。

可得如下结论:

1) R 值的减少将引起禁飞区域扩大, 导致无人机可能需要绕飞, 从而引起总航程和能耗成本增加, 并且总滞后和等待时间之和也随之增加。从图 8 可知, 与 R=1.4 相比, R=1.1 情形下的无人机配送方案在总航程上增加了 20.4%, 滞后时间不再是 0, 等待时间增加了 13.2%, 能耗增加了 18.9%。

2) 灵活的出发时间可以减少无人机到达客户点的等待与滞后时间。从图 9 可知, 与固定出发时间的配送方案相比, 虽然灵活出发时间条件下的配送方案总航程不变, 但其总滞后和等待时间之和减少 61.1%, 故总调度成本也相应减少 1.3%。

3) 无人机配送在多调度中心情形下允许机队在临近的调度中心出发执行任务, 从而降低无人机航程和能耗成本。从图 10 可知, 与单调度中心相比, 本文的总航程减少了 21.7%, 引起能耗减少 22.7%, 从而总调度成本也相应减少。

4) 无线电干扰源信号强度的增加将引起禁飞

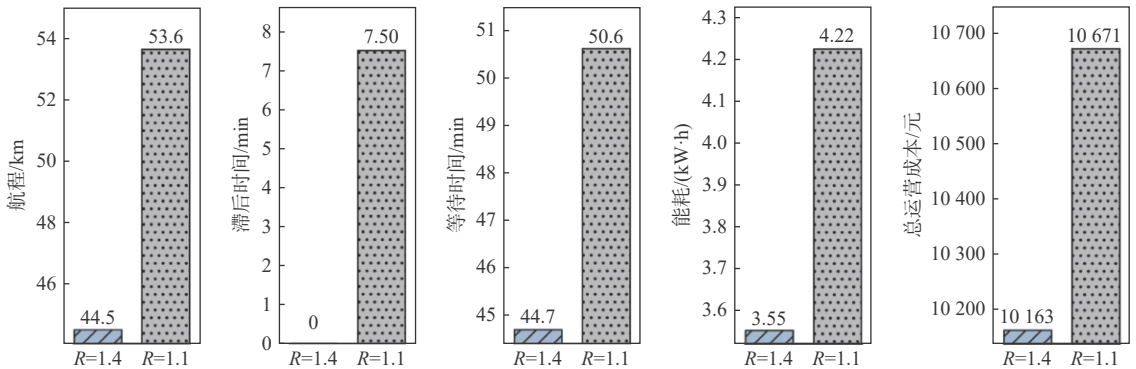


图 8　不同坠落代价阈值对双层模型结果的影响

Fig. 8　Effect of different crash cost thresholds on results of two-layer model

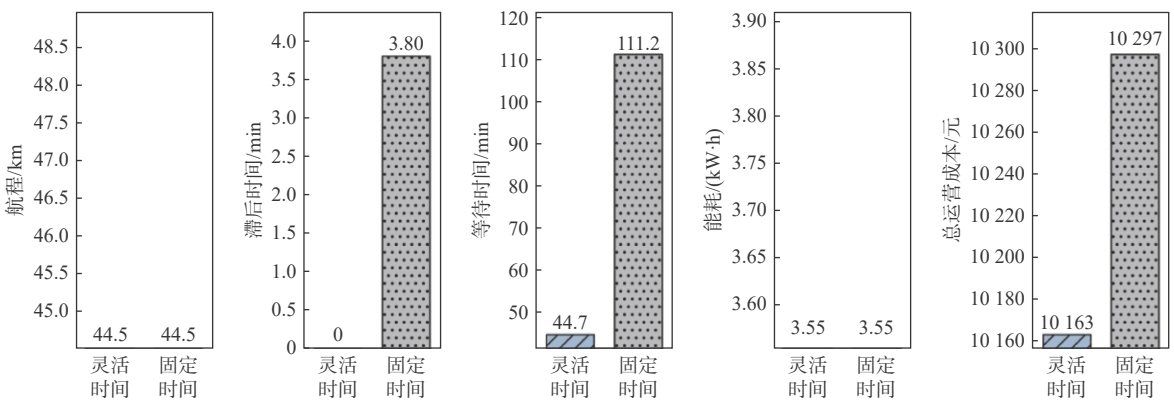


图 9　灵活和固定出发时间对双层模型结果的影响

Fig. 9　Effect of flexible and fixed departure time on results of two-layer model

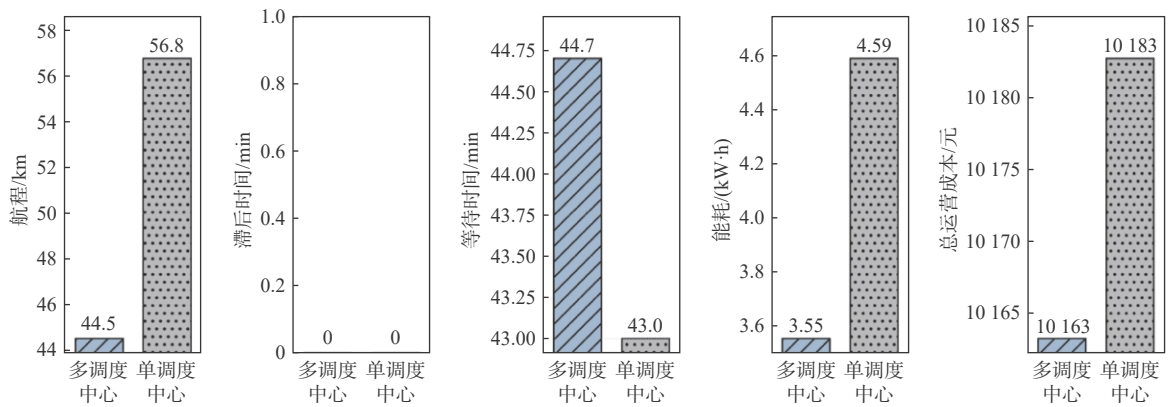


图 10　单/多调度中心对双层模型结果的影响

Fig. 10　Effect of single/multiple dispatch centers on results of two-layer model

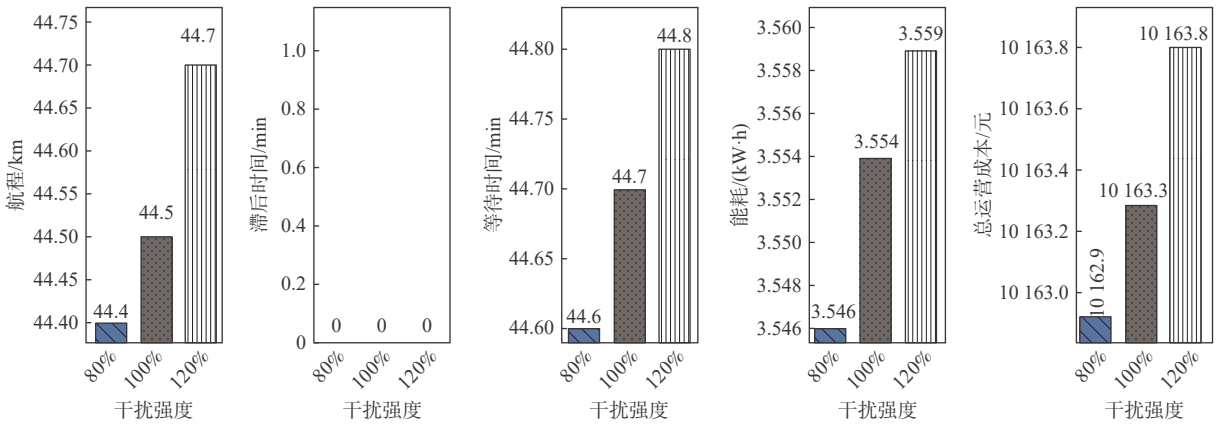


图 11 不同无线电干扰强度对双层模型结果的影响

Fig. 11 Effect of different radio interference intensities on results of two-layer mode

区域变多，从而无人机需要绕行避开风险区，引起航程、能耗和总成本也逐步增加。从图 11 可知，当无人机信号强度按照 20% 逐步增加时，总航程、无人机等待时间、能耗增量分别为 0.23% 和 0.45%、0.22% 和 0.22%、0.23% 和 0.14%。

3.4 算法性能分析

为验证本文算法的有效性，利用深度强化学习算法和典型的智能算法求解不同规模算例，求解结果如表 3 所示。可知：①在求解质量方面，由于本文的深度强化学习算法遵循可行解空间构造可行解，在训练中出现无效解的概率较低，故其到目前产生的最好解，即近最优解、平均解和最差解均优于 GA、差分进化 (differential evolution, DE) 算法和 PSO 算法；②在求解时间方面，虽然本文的深度强化学习算法需要较长的模型训练时间，当其一旦被训练好，求解时间优于 GA、DE 和 PSO 算法。

表 3 不同订单规模的多种算法性能对比

Table 3 Performance comparison of multiple algorithms for different order sizes

客户数量	算法	运营成本/元				求解时间/s
		近最优解	平均解	最差解	标准差	
27	DRL	10 163.25	10 283.63	10 696.34	53.25	10.26
	GA	10 589.28	15 206.35	45 646.75	4 796.39	44.34
	DE	13 087.94	27 396.45	51 437.46	5 068.57	45.87
	PSO	11 035.69	13 387.29	41 638.84	3 678.77	900.44
50	DRL	26 576.69	27 492.59	28 675.39	636.35	30.64
	GA	26 824.57	59 375.78	97 305.03	11 372.49	86.75
	DE	36 127.49	80 543.46	100 837.54	15 453.55	88.45
	PSO	42 198.26	114 385.58	189 627.29	10 783.89	1 703.45
100	DRL	40 783.58	47 295.49	53 652.03	982.32	42.58
	GA	54 782.36	88 547.57	238 693.72	32 981.49	148.29
	DE	62 642.23	100 458.28	219 051.03	20 682.68	209.61
	PSO	80 681.56	189 672.67	286 319.57	26 890.33	3 662.82

4 结 论

本文研究一种面向物流配送的带时间窗的多调度中心无人机运输线路和航迹协同规划的双层模型，及其基于深度强化学习的两阶段算法，考虑障碍物、无线电干扰与无人机坠落代价等多重安全因素的影响，根据客户和调度中心的时空分布，计算最佳的无人机运输线路、运行时间及其航迹。研究表明：

1）与单调度中心、固定出发时间的模型结果相比，本文的多调度中心、灵活出发时间可以减少总航程、能耗及延误费用。

2）多重安全因素的不同时空分布会影响无人机运输线路及其航迹布局，当其约束增加时，其航程及其能耗成本增加。

3）与传统的典型智能算法相比，本文提出的两

阶段深度强化学习算法框架在求解质量和时间方面具有优势。

参考文献（References）

[1] DORLING K, HEINRICHS J, MESSIER G G, et al. Vehicle routing problems for drone delivery[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2017, 47(1): 70-85.

[2] WANG X D, LIU Z Y, LI X P. Optimal delivery route planning for a fleet of heterogeneous drones: a rescheduling-based genetic algorithm approach[J]. Computers & Industrial Engineering, 2023, 179: 109179.

[3] SONGBD, PARK K, KIM J. Persistent UAV delivery logistics: MILP formulation and efficient heuristic[J]. Computers & Industrial Engineering, 2018, 120: 418-428.

[4] CHENG C, ADULYASAK Y, ROUSSEAU L M. Drone routing with energy function: formulation and exact algorithm[J]. Transportation Research Part B: Methodological, 2020, 139: 364-387.

[5] 张建行, 康凯, 钱骅, 等. 面向物联网的深度 Q 网络无人机路径规划[J]. 电子与信息学报, 2022, 44(11): 3850-3857.
ZHANG J H, KANG K, QIAN H, et al. UAV trajectory planning based on deep Q-network for Internet of things[J]. Journal of Electronics & Information Technology, 2022, 44(11): 3850-3857(in Chinese).

[6] 袁一帆, 吴德伟, 戴传金, 等. 基于 RRT 算法改进的无人机航迹规划研究[J]. 战术导弹技术, 2022(5): 126-133.
YUAN Y F, WU D W, DAI C J, et al. Research on UAV's path planning based on improved RRT algorithm[J]. Tactical Missile Technology, 2022(5): 126-133(in Chinese).

[7] LIN N, TANG J C, LI X W, et al. A novel improved bat algorithm in UAV path planning[J]. Computers, Materials & Continua, 2019, 61(1): 323-344.

[8] XU C, XU M, YIN C J. Optimized multi-UAV cooperative path planning under the complex confrontation environment[J]. Computer Communications, 2020, 162: 196-203.

[9] SHAO S K, PENG Y, HE C L, et al. Efficient path planning for UAV formation via comprehensively improved particle swarm optimization[J]. ISA Transactions, 2020, 97: 415-430.

[10] ZHOU Y M, SU Y, XIE A H, et al. A newly bio-inspired path planning algorithm for autonomous obstacle avoidance of UAV[J]. Chinese Journal of Aeronautics, 2021, 34(9): 199-209.

[11] 刘光才, 马寅松, 齐福强, 等. 基于改进 A*-人工势场法的城市物流无人机路径规划[J]. 飞行力学, 2022, 40(6): 16-23.
LIU G C, MA Y S, QI F Q, et al. Flight path planning for urban logistics UAV based on improved A*-APF algorithm[J]. Flight Dynamics, 2022, 40(6): 16-23(in Chinese).

[12] CHEN J C, LING F Y, ZHANG Y, et al. Coverage path planning of heterogeneous unmanned aerial vehicles based on ant colony system[J]. Swarm and Evolutionary Computation, 2022, 69: 101005.

[13] YU X B, JIANG N J, WANG X M, et al. A hybrid algorithm based on grey wolf optimizer and differential evolution for UAV path planning[J]. Expert Systems with Applications, 2023, 215: 119327.

[14] YU X B, LUO W G. Reinforcement learning-based multi-strategy cuckoo search algorithm for 3D UAV path planning[J]. Expert Systems with Applications, 2023, 223: 119910.

[15] 司鹏搏, 吴兵, 杨睿哲, 等. 基于多智能体深度强化学习的无人机路径规划[J]. 北京工业大学学报, 2023, 49(4): 449-458.
SI P B, WU B, YANG R Z, et al. UAV path planning based on multi-agent deep reinforcement learning[J]. Journal of Beijing University of Technology, 2023, 49(4): 449-458(in Chinese).

[16] LIU X X, SU Y Z, WU Y, et al. Multi-conflict-based optimal algorithm for multi-UAV cooperative path planning[J]. Drones, 2023, 7(3): 217.

[17] 韩鹏, 赵嶷飞. 基于飞行环境建模的 UAV 地面撞击风险研究[J]. 中国安全科学学报, 2020, 30(1): 142-147.
HAN P, ZHAO Y F. Study on ground impact risk of UAV based on flight environment[J]. China Safety Science Journal, 2020, 30(1): 142-147(in Chinese).

[18] NAZARI M, OROOJLOOY A, SNYDER L, et al. Deep reinforcement learning for solving the vehicle routing problem[C]//Proceedings of the 32nd International Conference on Neural Information Processing Systems. New York: ACM, 2018: 9861-9871.

[19] KINGMA D P, BA J. Adam: a method for stochastic optimization [EB/OL].(2017-01-30)[2023-08-01].https://arxiv.org/abs/1412.6980.

Cooperative planning for safe transportation routes and flight paths of UAV with multiple dispatching centers and soft time windows

WEI Ming[1, 2, *], SUN Yaru[1], SUN Bo[1], WANG Shengjie[2]

(1. School of Air Traffic Management, Civil Aviation University of China, Tianjin 300300, China;
2. Key Laboratory of Flight Techniques and Flight Safety, Civil Aviation Flight University of China, Guanghan 618300, China)

Abstract: A two-layer collaborative planning model is established for the problem of unmanned aerial vehicle (UAV) transportation routes and flight paths planning for logistics distribution. In the upper-level model, considering constraint factors such as customer time windows, UAV load, energy consumption, and path risk, the UAV's departure dispatching center, access sequence, and customer arrival times were calculated to minimize UAV dispatching costs. In the lower-level model, considering multiple safety factors such as obstacles, radio interference, and UAV crash costs, the shortest feasible flight path between any dispatching center and customers was calculated. A two-stage deep reinforcement learning (DRL) algorithm, incorporating the A^* algorithm and a greedy strategy, was designed to solve the problem based on its characteristics. Finally, a case studywas presented where the optimal UAV transportation route and flight path planning scheme were calculated, and the impact of changes in key parameters on the scheduling results was analyzed. The validity and accuracy of this paper were verified by comparison with genetic algorithm (GA), differential evolution (DE) and particle swarm optimization (PSO) algorithms.

Keywords: air transport; unmanned aerial vehicle transportation route; unmanned aerial vehicle flight path; two-layer model; deep reinforcement learning; A^* algorithm

Received: 2023-08-04; **Accepted**: 2023-09-03; **Published Online**: 2024-01-08 08:26

URL: link.cnki.net/urlid/11.2625.V.20240105.1557.004

Foundation items: Open Fund of Key Laboratory of Flight Techniques and Flight Safety (FZ2021KF06); The Ministry of Education of Humanities and Social Science Project (20YJCZH176)

* **Corresponding author**. E-mail: mwei@cauc.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0543

基于肢体动作预测的动态可变人机协作装配

吴海彬[1,2,*]，宋晨阳[1]，周世璇[1]

(1. 福州大学 先进制造学院，泉州 362200；2. 福建省特种智能装备安全与测控重点实验室，福州 350003)

摘 要：在复杂的装配领域中，单纯依靠机器人很难完成，进行人机协作装配才能确保装配过程顺利进行，但人机协作过程中，机器人缺乏根据人的不同操作作出响应的能力。针对产品组装过程中存在多种不同的组装方案和顺序的情况，提出一种通过识别和预测操作者肢体动作，使机器人根据操作者不同的选择做出对应配合的装配动作，实现动态可变的人机协作装配方案。利用惯性测量单元(IMU)采集操作者的肢体动作信息，提取惯性测量单元信号特征，为提高识别准确率，提出一种同时利用时域和时频域特征的粒子群优化(PSO)支持向量机(SVM)算法，用于肢体动作识别；同时，提出一种参数可变的隐马尔可夫模型(HMM)实时预测肢体动作序列，在监督训练和零训练状态下，推断出操作者的未来意图，实现装配序列的动态可变性。实验结果表明：在减速器人机协作装配实验中，所提方案肢体动作平均识别率达到96.7%，并且能够有效预测操作者肢体动作，实现装配顺序可变的动态人机协作装配，显著提升了装配系统的适应性，降低了人机协作装配的复杂度。

关 键 词：运动控制；可穿戴传感器；人机协作装配；顺序动态可变；隐马尔可夫模型

中图分类号：TB181

文献标志码：A **文章编号**：1001-5965(2025)10-3243-10

人机协作(human-robot collaboration, HRC)是当前机器人领域的一个重要发展方向。机器人虽然广泛地应用于车间生产，但缺乏认知能力，在一些复杂装配环境或柔性生产车间中，仍需要人与机器人配合完成关键步骤的操作。人的动作复杂多样且具有随机性，并且人可以主观选择可行的不同操作方法和路线，机器人在与人协作过程中应实时识别和预测人的动作意图，并提供相应的帮助，这对提高生产效率和实现柔性化生产都有着重大意义[1-2]。

目前，在人机协作领域中，一些研究主要试图利用视觉传感器[3-5]、脑电传感器[6]、肌电传感器[7]、惯性测量单元(inertial measurement unit, IMU)等方式从人体动作信息中识别人的意图。肌电传感器硬件成本低廉且佩戴便捷，但存在受肌肉状态影响大、识别效果不稳定等问题。脑电传感器实时性好，但脑电信号随机性强，提取信息相对困难。IMU相对于其他传感器，具有设备轻巧、不受光线限制、便于佩戴等优点，因而展现出巨大的应用潜力。Syed等[8]在物流产线中使用IMU采集人体动作数据，通过XGBoost分类器进行识别，活动识别平均准确率为78.61%。Wang等[9]通过带有IMU和肌电传感器的可穿戴式设备采集人体运动数据，采用隐马尔可夫模型(hidden Markov model, HMM)识别出人的物料交接意图。

运动预测主要分为2种形式：运动轨迹预测和操作序列预测。在运动轨迹预测方面，陈友东等[10]提出一种基于长短期记忆(long short-term memory, LSTM)网络的动作预测方法，根据人体动作的初

收稿日期：2023-08-24；**录用日期**：2023-12-29；**网络出版时间**：2024-01-31 17:46
网络出版地址：link.cnki.net/urlid/11.2625.V.20240131.1038.001
基金项目：国家重点研发计划(2018YFB1308603)；福建省科技重大专项专题项目(2024HZ026020)
*通信作者. E-mail：wuhb@fzu.edu.cn

引用格式：吴海彬，宋晨阳，周世璇. 基于肢体动作预测的动态可变人机协作装配[J]. 北京航空航天大学学报，2025，51(10)：3243-3252.
WU H B，SONG C Y，ZHOU S X. Dynamically changeable human-robot collaborative assembly based on limb motion prediction[J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51(10)：3243-3252 (in Chinese).

始部分对动作的终点做出预测。Liu等[11]探索了一种基于卷积神经网络(convolutional neural network, CNN)和LSTM网络的视觉信号深度学习系统,利用LSTM网络提取人体运动的时间模式,在运动发生前自动输出预测结果。以上运动轨迹预测方法对计算机的计算速率要求较高,且获得的轨迹数据可能会受环境干扰出现缺失或位置偏移的现象。在操作序列预测方面,一些研究旨在通过对人和机器人执行任务的离散动作进行分析,利用动作之间的转移关系来预测后续动作。为实现人机无缝交互,Kwon等[12]提出了基于决策理论的主动规划框架,以选择最佳行动和最佳时机改善人机交互的手段,并开发了复合节点时态贝叶斯网络,能够根据当前工作流程预测工人行为和将要完成的任务。Hawkins等[13]通过贝叶斯网络推理预测操作之后的行为。Liu等[14]基于人体运动识别技术,引入HMM生成运动转移概率矩阵,对人体运动进行识别与建模,实现了对工人任务序列的预测。Cramer等[15]采用“物体启示”的方法,运用kinect相机识别操作者的手进入的区域,根据这些物体的启示来确定当前和预测未来操作员意图。Zanchettin等[16]提出了一种识别和预测协同装配过程中人的活动模式序列的算法,对人执行某项操作之前的剩余时间进行估计,以便及早推断出人何时会要求特定的协作操作,同时允许机器人执行替代的自主任务。以上研究尽管可以预测后续动作,但并不能有效实现操作序列动态可变。

为解决人机协作过程中无法有效实现装配序列动态可变性的问题,本文提出了一种同时利用时域和时频域特征的粒子群优化(particle swarm optimization, PSO)的支持向量机(support vector machine, SVM)肢体动作识别算法,将肢体动作识别的动态动作序列作为人体运动预测的输入;同时,提出一种基于参数可变HMM的肢体动作序列预测算法,在装配过程中进行增量训练,实时改变HMM的参数,进而实现装配顺序可变的动态人机协作装配,避免了传统方案需要大量训练来适应不同装配方案及限制操作者灵活性和自主性的缺点,提高了人机协作装配过程的效率,更好地发挥了系统中操作者的主观能动性。

1 人机协作装配系统整体方案

本文搭建的基于肢体动作识别预测的人机协作装配系统总体方案如图1所示。系统由3个模块组成:模块1为便携式IMU传感器,用于收集工人装配动作时的加速度和欧拉角信号;模块2为计算机,对IMU数据进行预处理及特征提取,并使用基于PSO的SVM算法识别人机协作装配动作意图,进一步结合肢体动作识别序列和参数可变的HMM算法调用相应机器人运动控制程序;模块3为工业机器人,执行计算机发送的控制程序,从而协助完成装配任务。

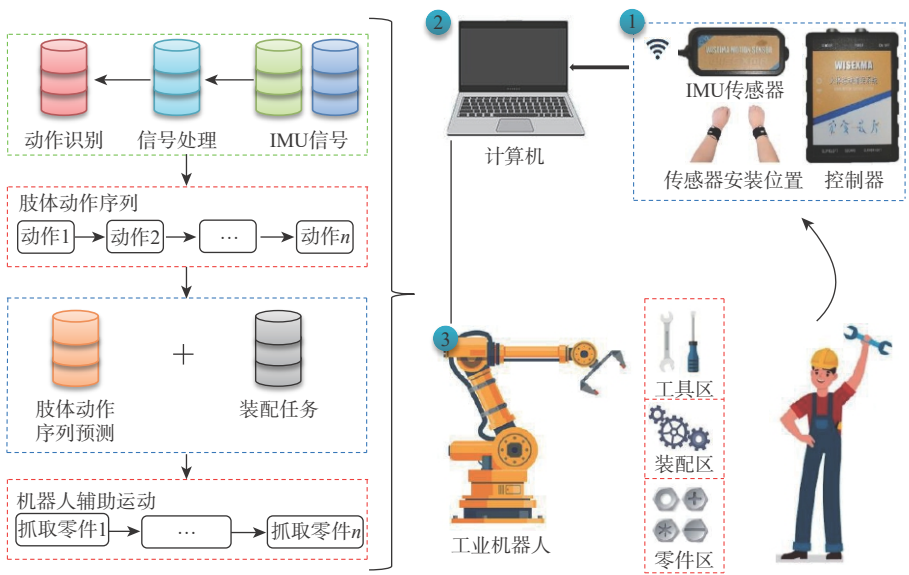


图1 人机协作装配系统总体方案

Fig. 1 Overall scheme of human-robot collaborative assembly system

以某减速器的人机协作装配为例,图2为某减速器的部分装配序列。传感器对人的肢体动作进行识别预测,并按照可行的装配序列为操作工人提供相应配合操作,而在此过程中,工人可以根据实际情况动态调整装配过程,更好地发挥人的主观能动性,高效完成装配任务。

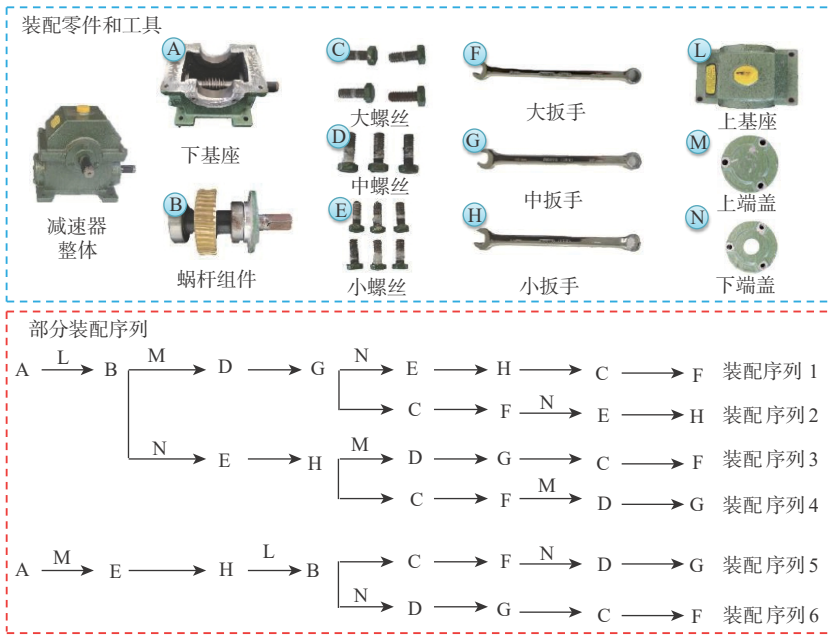


图 2　某减速器的部分装配序列

Fig. 2　Partial assembly sequence of a reducer

2　肢体动作识别

2.1　数据预处理及特征提取

为剔除信号中的重力加速度和零点漂移现象的影响, 需对 IMU 中 3 轴加速度计进行重力加速度补偿和零点漂移补偿, 使无运动状态下的数据尽量处于零点位置。

选取恰当的特征值对于运动模式识别系统的精度至关重要。时域特征能够很好地表示各种不同的人体肢体动作, 计算复杂度低, 过程简单; 频域特征表示信号的频域特点, 建立在快速傅里叶变换的基础上, 计算复杂, 耗时长; 而时频域特征可以弥补频域特征中缺乏的时域信息, 能更全面地刻画信号的特征。因此, 采用同时提取时域信号和时频域信号特征的方法, 提取 9 轴 IMU 传感器信号中加速度信号和欧拉角信号的均值、标准差、峰值、峰值因子、均方差作为时域特征, 将 db1 小波包进行 3 层系数分解, 提取 3 层系数分解的低频能量值作为信号时频域特征, 如表 1 所示。其中, 均值反映信号平均强度, 标准差反映信号分散程度, 峰值反映信号的幅度, 波形因子反映信号有效值 (均方根) 与整流平均值的比值, 峰值因子反映信号峰值在波形中的极端程度, 小波包能量特征提取信号中主要信息的低频信息。表 1 中: x_i为信号的时域序列, $i=1,2,\cdots,N$, N为样本点数, $d_l^{j,k}$为信号应用小波包进行信号分解后在节点(j,k)的分解信号, l为节点信号$d_l^{j,k}$的离散点编号, m 为节点信号的离散点总数。

表 1　特征提取

Table 1　Feature extraction

特征	公式
均值	$\bar{x}=\frac{1}{N}\sum_{i=1}^{N}x_i$
峰值	$x_{\mathrm{p}}=\max\{\lvert x_1\rvert,\lvert x_2\rvert,\cdots,\lvert x_N\rvert\}$
标准差	$\sigma_x=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})^2}$
均方根	$x_{\mathrm{rms}}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}x_i^2}$
峰值因子	$C=\frac{x_{\mathrm{p}}}{x_{\mathrm{rms}}}$
波形因子	$W=\frac{x_{\mathrm{rms}}}{\frac{1}{N}\sum_{i=1}^{N}\lvert x_i\rvert}$
小波包能量	$E_{j,k}=\sum_{l=1}^{m}\lvert d_l^{j,k}\rvert^2$

归一化的效果是原始数据被规整到 [0,1], f_i为信号的特征序列, F_i为信号归一化后的特征序列, 采用的归一化映射如下:

$$F_i=\frac{f_i-f_{\min}}{f_{\max}-f_{\min}} \tag{1}$$

式中: $f_i,F_i\in\mathbf{R}$; $f_{\min}=\min(f_i)$; $f_{\max}=\max(f_i)$; $F_i\in[0,1]$; $i=1,2,\cdots,n$。

2.2　肢体动作识别算法

SVM 是一种具有较好学习能力和泛化性能的机器学习算法, 其能够有效解决传统方法中遇到的非线性、小样本和高维度问题[17]。本文使用 MATLAB 软件, 基于 LIBSVM 软件包实现动作识别[18]。为使测试集的分类准确率维持在较高的水平, 避免过学习和欠学习状况发生, 需对分类器中的关键参数惩

罚因子 c 和核函数参数 g 进行寻优。采用 PSO 算法对搭建的 SVM 分类器[19]进行参数优化。采用基于交叉验证 (cross validation, CV) 思想的网格搜索法寻找最佳参数 c 和 g，即全局最优解，但在更大范围寻找最佳参数 c 和 g 会很费时，采用启发式算法就可以不必遍历网格中的所有参数点，也能找到全局最优解。PSO 算法是一种基于群体智能的演化计算技术，通过粒子在解空间追随最优的例子进行搜索。利用 PSO 对 SVM 参数进行优化的算法流程如图 3 所示。

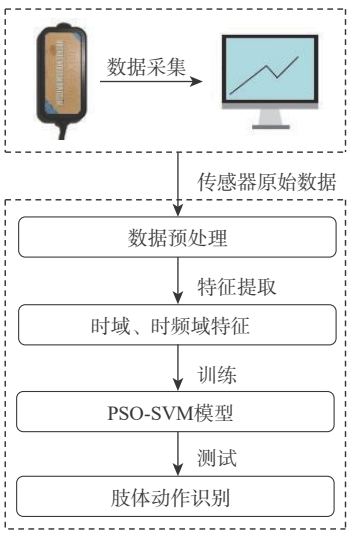


图 3 PSO 优化 SVM 参数的算法流程

Fig. 3 Algorithm flow of SVM parameters optimized by PSO

3 肢体动作预测

HMM 是一种与时间序列有关的概率统计模型，描述由一个隐藏的马尔可夫链随机生成不可观测的状态随机序列，再由各状态生成一个观测从而产生观测随机序列的过程[20]。HMM 算法分为 2 个阶段：建模阶段和预测阶段。建模阶段是以数据集为基础对模型进行训练，求解出模型参数 λ；预测阶段将模型参数 λ 输入模型后进行预测。本文对算法进行了改进，提出了一种参数可变的 HMM 的肢体动作序列预测算法，对 HMM 中的前向算法进行扩展，使其可以预测工人的未来意图；Viterbi 算法用于推断装配任务序列，在零训练肢体动作序列状态下，随机生成模型参数，通过 Baum-Welch 算法在工人装配任务过程中增量训练肢体动作序列，对模型参数进行迭代，从而适应工人的偏好，预测操作人员当前意图及未来操作意图，实现装配顺序可变的动态人机协作装配。本文改进的 HMM 流程如图 4 所示。

3.1 建模阶段

HMM 的核心是隐马尔可夫链，定义随机顺序

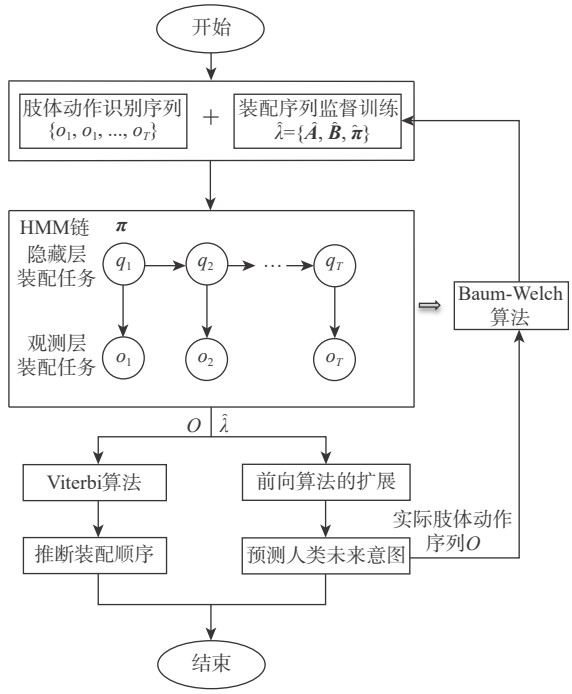


图 4 改进的 HMM 算法流程

Fig. 4 Flow of improved HMM algorithm

的 T 种行为为观测层，隐藏层代表装配任务，如图 5 所示。状态集合 Q 表示不同装配任务状态 q_i 的集合 (隐状态集合)，观测集合 V 表示肢体动作识别状态 o_i 的集合 (显状态集合)，状态序列 I 表示装配任务的序列，观测序列 O 表示肢体动作识别状态的序列。

$$Q=\{q_1,q_2,\cdots,q_N\} \tag{2}$$

$$V=\{o_1,o_2,\cdots,o_M\} \tag{3}$$

$$I=\{i_1,i_2,\cdots,i_T\} \tag{4}$$

$$O=\{o_1,o_2,\cdots,o_T\} \tag{5}$$

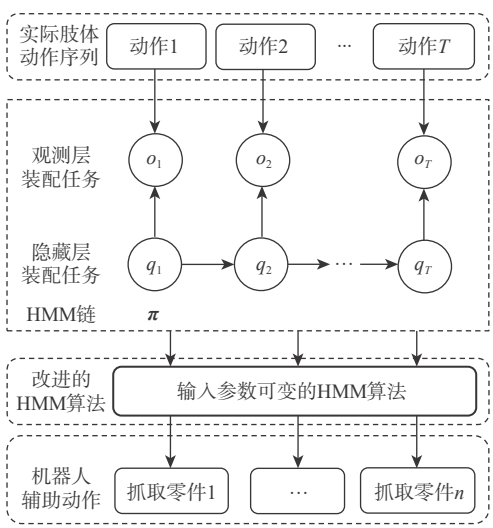


图 5 肢体动作序列预测 HMM 结构

Fig. 5 HMM structure prediction based on a limb movement sequence

式中: N为可能的隐状态数; M为可能的显状态数。

3.2　预测阶段

HMM 预测阶段的输入为实时肢体动作识别序列O和参数λ, 为防止零概率的情况, 对模型参数λ包括的初始状态矩阵$\boldsymbol{\pi}$、状态转移矩阵$\boldsymbol{A}$、观测矩阵 (发射矩阵)$\boldsymbol{B}$每次重估完后, 对参数λ进行改进, 即

$$\lambda=\{\boldsymbol{A},\boldsymbol{B},\boldsymbol{\pi}\}+\{\boldsymbol{A}',\boldsymbol{B}',\boldsymbol{\pi}'\} \tag{6}$$

对 HMM 中的前向算法进行扩展, 使其可以预测人的下一步意图。初始化模型首个动作前向概率$\alpha_1(i)$, 再迭代计算$\alpha_{t+1}(j)$, 最后计算观测序列O的概率$P(O|\lambda)$, 算法步骤如下:

$$\alpha_1(i)=\pi_i b_i(o_1)\qquad 1\leqslant i\leqslant N \tag{7}$$

$$\alpha_{t+1}(j)=\left(\sum_{1\leqslant i\leqslant N}\alpha_t(i)a_{ij}\right)b_j(o_{t+1}) \tag{8}$$

$$P(O|\lambda)=\sum_{i=1}^{N}\alpha_T(i) \tag{9}$$

式中: $\alpha_t(j)$表示在时刻t隐状态为q_j时, 观测状态的序列为$o_1,o_2,\cdots,o_t$的概率; a_{ij}为矩阵$\boldsymbol{A}$中的元素; $b_j(k)$为矩阵$\boldsymbol{B}$中的元素; π_i为矩阵$\boldsymbol{\pi}$中的元素; $1\leqslant t\leqslant T-1$, $1\leqslant k\leqslant M$, $1\leqslant j\leqslant N$。

对 Baum-Welch 算法进行改进, 使用 Baum-Welch 算法进行迭代计算, 得到重估模型参数$\hat{\lambda}=\{\hat{\boldsymbol{A}},\hat{\boldsymbol{B}},\hat{\boldsymbol{\pi}}\}$, 将参数$\hat{\lambda}$作为下一次人机协作装配的 HMM 参数, 实时对参数λ进行迭代, 从而实现装配顺序可变的动态人机协作装配。算法步骤如下:

1) 初始化模型。对于$n=0$, 选取$\boldsymbol{A}^{(0)}$、$\boldsymbol{B}^{(0)}$、$\boldsymbol{\pi}^{(0)}$得到如下模型:

$$\lambda^{(0)}=\{\boldsymbol{A}^{(0)},\boldsymbol{B}^{(0)},\boldsymbol{\pi}^{(0)}\} \tag{10}$$

2) 递推。设置迭代次数$n=1,2,\cdots$, 得

$$a_{ij}^{(n+1)}=\sum_{t=1}^{T-1}\xi_t(i,j)\Big/\sum_{t=1}^{T-1}\gamma_t(i) \tag{11}$$

$$b_j^{(n+1)}(k)=b_j^{(n)}(k) \tag{12}$$

$$\pi_i^{(n+1)}=\gamma_1(i) \tag{13}$$

式中: $\xi_t(i,j)$为在时刻t处于状态q_i, 且时刻$t+1$处于状态q_j的概率; $\gamma_t(i)$为在时刻t时处于状态q_i的概率。

从而得到迭代后的参数为

$$\lambda^{(n+1)}=\{\boldsymbol{A}^{(n+1)},\boldsymbol{B}^{(n+1)},\boldsymbol{\pi}^{(n+1)}\} \tag{14}$$

$$\hat{\lambda}=\{\hat{\boldsymbol{A}},\hat{\boldsymbol{B}},\hat{\boldsymbol{\pi}}\} \tag{15}$$

3) 计算时刻$t+1$每个隐藏状态联合概率S_{t+2}, 其中, 概率最高的隐藏状态代表操作者的意图, 从而实现对操作者未来意图的预测, 如下:

$$S_2=\arg\max_{1\leqslant i\leqslant N}\left(\sum\alpha_1(i)a_{ij}\right)\qquad 1\leqslant j\leqslant N \tag{16}$$

$$S_{t+2}=\arg\max_{1\leqslant i\leqslant N}\left(\sum\alpha_{t+1}(j)a_{ij}\right)\qquad 1\leqslant t\leqslant T-1,1\leqslant j\leqslant N \tag{17}$$

4　实　验

4.1　肢体动作识别实验

定义符合装配过程的 8 种装配状态, 分别为装基座 (AS1)、装配蜗杆 (AS2)、装配大螺丝 (AS3)、装配上端盖 (AS4)、装配下端盖 (AS5)、拧大螺丝 (AS6)、拧中螺丝 (AS7)、拧小螺丝 (AS8), 因此, 装配步骤不同任务状态的集合如下: $Q=\{\text{AS1},\text{AS2},\text{AS3},\text{AS4},\text{AS5},\text{AS6},\text{AS7},\text{AS8}\}$。

定义装配过程的 8 种手臂指令集, 肢体动作说明如图 6 所示, 分别为左手拿下基座 (LTB)、右手拿蜗杆 (RTW)、左手拿大螺丝 (LBS)、左手拿中螺丝 (LMS)、左手拿小螺丝 (LSS)、右手拿大扳手 (RBT)、右手拿中扳手 (RMT)、右手拿小扳手 (RST), 因此, 肢体动作识别的集合态为: $V=\{\text{LTB},\text{RTW},\text{LBS},\text{LMS},\text{LSS},\text{RBT},\text{RMT},\text{RST}\}$。

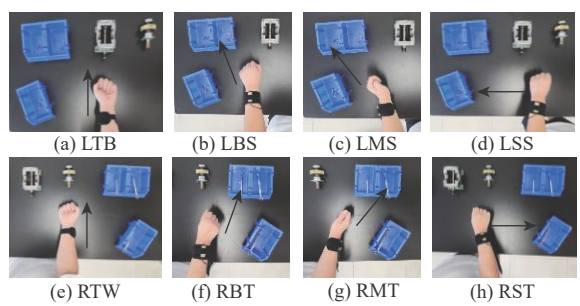


图 6　肢体动作

Fig. 6　Limb movement

定义机器人 3 个肢体动作, 分别为机器人拿上基座 (RTP)、机器人拿上端盖 (RUC)、机器人拿下端盖 (RDC)。同时, 规定预测到 RTW 动作时, 机器人执行 RTP 动作, 预测到 LMS 动作时, 机器人执行 RUC 动作, 预测到 LSS 动作时, 机器人执行 RDC 动作。

邀请 5 名被试人员参加数据采集实验, 不考虑被试人员的年龄和经验。在实验过程中, 数据采集频率为 50 Hz, 被试人员依次做出 8 个肢体动作, 每个肢体动作持续 2 s, 保留采集过程中间 1.6 s 的数据, 共采集 560 个样本作为训练集。继续邀请 3 名被试人员参加数据采集实验, 共采集 240 个样本作为测试集。为保持 IMU 数据的一致性, 要求所有被试人员以相同的位置佩戴。

将 IMU 平稳向上放置一段时间, 加速度信号在z轴方向受重力加速度的影响最大, 如图 7 所示, 加速度值约为 9.8 m/s^2, 进行重力补偿后, x、y、z 三

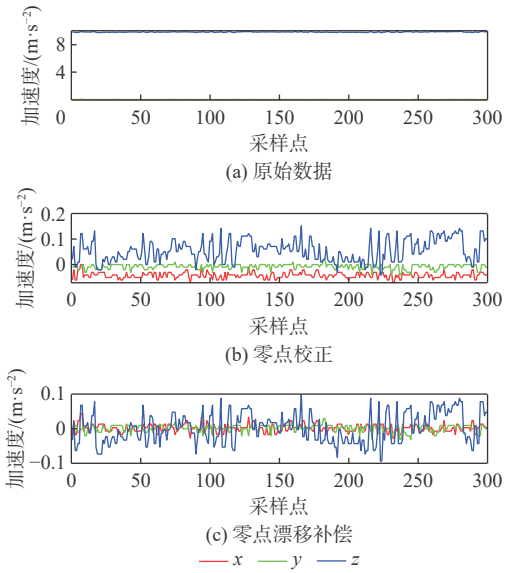


图 7 数据预处理

Fig. 7 Data preprocessing

轴方向的加速度补偿到零点位置左右。对于零点漂移的补偿，采用均值作为零点漂移的补偿值，加速度信号的精度区间提高到 [−0.1, 0.1] m/s^2 左右。

采用基于 PSO 的 SVM 算法进行离线状态识别。在 PSO 算法参数寻优过程中，设置 PSO 算法的种群数量为 20，终止迭代次数为 200，适应度曲线如图 8 所示，得到训练集最优参数组合为 c=79.986 5，g=1.031 5，训练集的平均识别率为 97.857 1%。

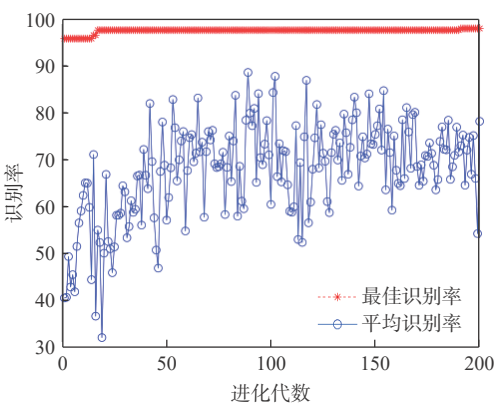


图 8 PSO 寻找最佳参数的识别率曲线

Fig. 8 Recognition rate curves of PSO to find the best parameters

如图 9 所示，PSO-SVM 算法对 8 个肢体动作的测试集平均识别率为 96.7%。因此，采用基于时域和时频特征的 PSO-SVM 的肢体动作识别算法进行肢体动作识别，并将肢体动作识别序列作为 HMM 的输入是可行的。

4.2 肢体动作序列预测实验

4.2.1 监督训练肢体动作序列预测实验

某齿轮减速器的装配任务由 11 个部分组成，

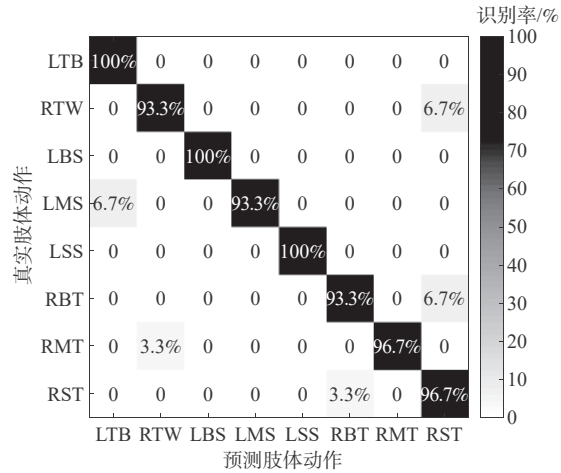


图 9 PSO-SVM 肢体动作分类混淆矩阵

Fig. 9 PSO-SVM limb action classification confusion matrix

其中，操作人员负责 8 个装配任务，机器人负责 3 个装配任务，共产生 66 个潜在的装配任务序列，实验场景如图 10 所示。

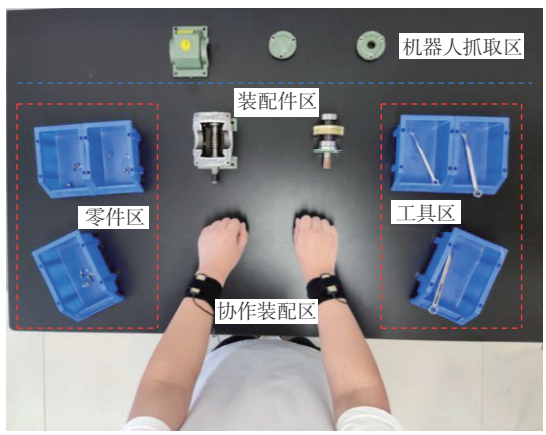


图 10 减速器装配实验场景

Fig. 10 Scene diagram of reducer assembly experiment

邀请 4 名被试人员每人组装某齿轮减速器 20 次，共采集 80 个样本。在标记的 80 组肢体动作序列数据集上监督训练学习 HMM 参数，得到 $\boldsymbol{A}$、$\boldsymbol{B}$、$\boldsymbol{\pi}^{\mathrm{T}}$，如表 2～表 4 所示。

表 2 初始状态矩阵参数

Table 2 Initial state matrix parameter

状态	数值
LTB	1
RTW	0
LBS	0
LMS	0
LSS	0
RBT	0
RMT	0
RST	0

如图 11 所示，HMM 模型在经过多个装配序列

表 3　状态转移矩阵参数

Table 3　State transition matrix parameter

状态	数值							
	AS1	AS2	AS3	AS4	AS5	AS6	AS7	AS8
AS1	0	$\frac{31}{40}$	0	0	$\frac{9}{40}$	0	0	0
AS2	0	0	$\frac{21}{80}$	$\frac{41}{80}$	$\frac{9}{40}$	0	0	0
AS3	0	0	0	$\frac{11}{40}$	$\frac{7}{20}$	$\frac{3}{16}$	$\frac{1}{10}$	$\frac{7}{80}$
AS4	0	0	$\frac{9}{20}$	0	$\frac{3}{20}$	$\frac{1}{10}$	$\frac{9}{40}$	$\frac{9}{40}$
AS5	0	$\frac{3}{20}$	$\frac{7}{40}$	$\frac{3}{20}$	0	$\frac{19}{80}$	$\frac{1}{20}$	$\frac{19}{80}$
AS6	0	0	0	$\frac{2}{55}$	$\frac{1}{55}$	0	$\frac{34}{55}$	$\frac{18}{55}$
AS7	0	0	$\frac{5}{58}$	0	$\frac{3}{58}$	$\frac{10}{29}$	0	$\frac{15}{29}$
AS8	0	$\frac{6}{47}$	$\frac{4}{47}$	$\frac{3}{47}$	0	$\frac{18}{47}$	$\frac{16}{47}$	0

表 4　发射矩阵参数

Table 4　Transmit matrix parameter

状态	数值							
	LTB	RTW	LBS	LMS	LSS	RBT	RMT	RST
AS1	1	0	0	0	0	0	0	0
AS2	0	1	0	0	0	0	0	0
AS3	0	0	1	0	0	0	0	0
AS4	0	0	0	1	0	0	0	0
AS5	0	0	0	0	1	0	0	0
AS6	0	0	0	0	0	1	0	0
AS7	0	0	0	0	0	0	1	0
AS8	0	0	0	0	0	0	0	1

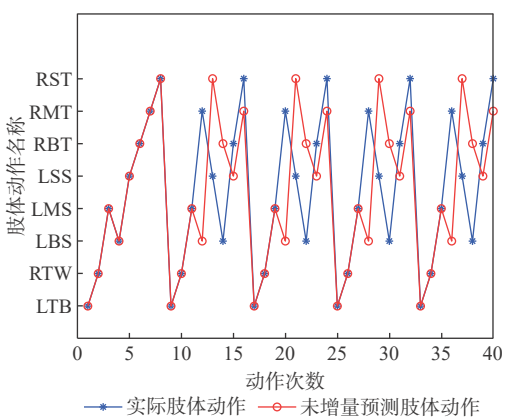


图 11　肢体动作序列未增量训练预测

Fig. 11　Prediction of limb action sequence without incremental training

数据集监督训练参数后, 输入经过训练后的肢体动作序列, 可以正确预测操作者的意图肢体动作序列。但当输入未经训练的肢体动作序列后, 推断的意图不符合预期的装配协作方式。肢体动作序列是否可以有效预测, 与前期监督算法训练的参数 λ有关, 对于未经过训练的肢体动作序列, 无法进行正确的预测。

为实现对未知肢体动作序列的识别, 在操作过程中对模型进行增量训练。如图 12 所示, 在第 1 个装配序列过程中, 可以有效推断已经经过训练的肢体动作序列。当操作者更换装配序列顺序时, 通过改进的 Baum-Welch 算法, 对已经完成的一组肢体动作序列过程进行非监督的增量训练, 其增量训练的数据是上一次整个装配过程实际肢体动作识别序列, 设置迭代次数n为 6 次。在第 2 个动态装配序列过程中, 预测的意图肢体动作序列逐渐接近实际肢体动作序列。在装配过程中训练 3 次数据后, 可以看出, 从一个装配序列转向另一个装配序列, 达到了预期的装配协作, 实现了装配顺序的可变性。

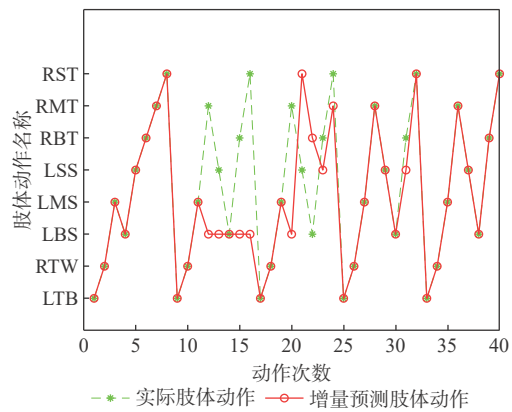


图 12　肢体动作序列增量训练预测

Fig. 12　Incremental training prediction of limb action sequence

图 13 为模型监督训练肢体动作序列后, 通过在装配过程中增量训练数据获得的部分状态转移概率变化。可知, 当操作人员完成一次装配任务, 由于在过程中进行了增量训练, AS4 转向 AS3、AS7 转向 AS8、AS5 转向 AS6、AS3 转向 AS5 的概率增加至 1 附近, 而 AS4 转向 AS7、AS7 转向 AS5、AS5 转向 AS3、AS3 转向 AS6 的概率减小至 0 附近。当操作员改变装配序列时, 经过 4 次增量训练后, AS4 转向 AS7、AS7 转向 AS5、AS5 转向 AS3、AS3 转

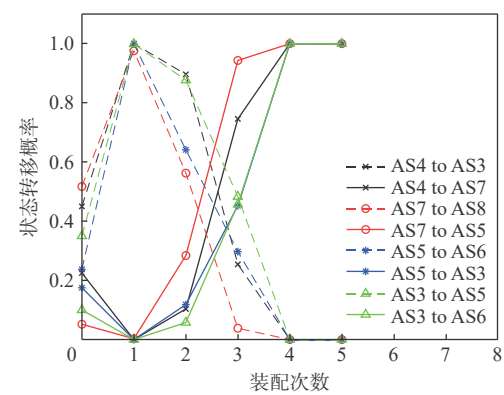


图 13　监督训练肢体动作序列增量训练部分状态转移概率

Fig. 13　State transition probability of incremental training part of supervised training limb action sequence

向AS6的概率逐渐增加至1附近，但AS4转向AS3、AS7转向AS8、AS5转向AS6、AS3转向AS5的概率逐渐减小至0附近。因此，装配过程的增量训练可以实时改变状态转移矩阵的参数，也可实时改变初始状态矩阵的参数，进而实现具有装配顺序可变的动态人机协作装配。此外，这种协作将随着时间的推移得到改善，并更好地适应个别操作者。

4.2.2 零训练肢体动作序列预测实验

模型初始状态不需要对潜在装配序列统计及监督训练，即零训练装配序列状态下，随机生成HMM的模型参数λ。为验证参数可变的HMM算法的泛化性，随机生成模型参数λ作为HMM参数的输入，进行肢体动作序列预测验证。随机采用了一种肢体动作序列进行预测，如图14所示，实际肢体动作序列和预测肢体动作序列存在较大误差，装配过程中，模型未进行增量训练，无法正确有效地预测正确的装配动作序列。

如图15所示，首个装配过程中，由于未进行增量训练，实际肢体动作序列和预测肢体动作序列存在误差。完成第1组装配后，通过改进的Baum-Welch算法在过程中非监督增量训练1组肢体动作序列，可以看出，实际肢体动作序列和预测肢体动作序列相同，达到了预期的装配协作顺序。当更改装配顺序时，在过程中增量训练2组肢体动作序列后，通过参数可变的HMM算法预测的意图肢体动作序列和实际肢体动作序列相同，便达到了预期的装配协作。因此，只需要确定整个装配过程中装配动作数，不需要对潜在的装配序列预先统计及监督训练，说明本文改进的HMM算法具有一定的泛化性。

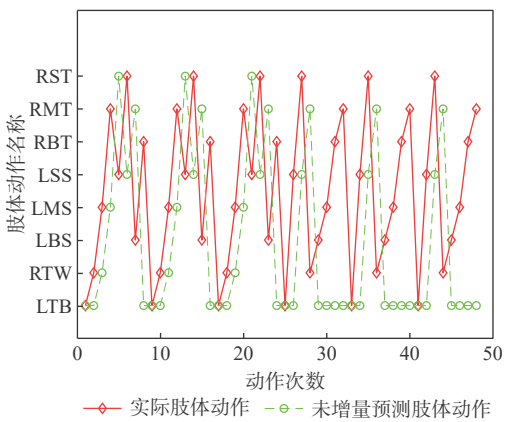


图14 零训练肢体动作序列未增量训练预测

Fig. 14 Prediction of zero-training limb action sequence without incremental training

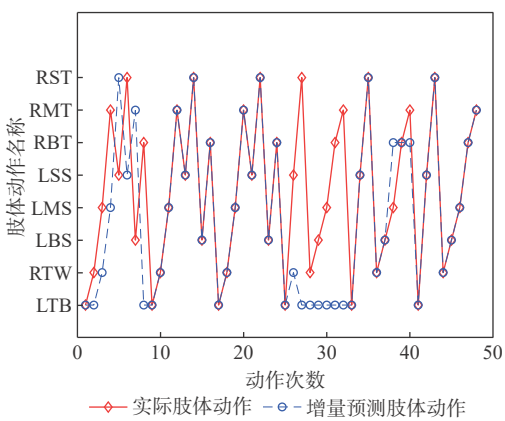


图15 零训练肢体动作序列增量训练预测

Fig. 15 Incremental training prediction of zero training limb action sequence

图16为零训练肢体动作序列后，模型随机生成模型参数，在装配过程中进行增量训练所得的部分状态转移概率变化。可知，当操作员做完3次装配任务，在过程中增加增量训练，AS1转向AS2、AS2转向AS4、AS4转向AS7、AS8转向AS3的概率逐渐增加至1附近，而AS1转向AS5、AS2转向AS3、AS4转向AS6、AS8转向AS2的概率徘徊在0附近。当操作员改变装配序列时，经过3次增量训练，而AS1转向AS2、AS2转向AS4、AS4转向AS7、AS8转向AS3的概率逐渐降低至0附近，但AS1转向AS5、AS2转向AS3、AS4转向AS6、AS8转向AS2的概率逐渐增加至1附近。由于在多组装配过程中，肢体动作LSS和RST的先后顺序未改变，AS5转向AS8的概率逐渐增加至1附近。因此，即使模型在零训练肢体动作序列状态下，只要在装配过程进行增量训练，就可以实时改变状态转移矩阵的参数及初始状态矩阵的参数，进而实现装配顺序可变的动态人机协作装配，减轻前期监督训练负担，提高人机协作装配模式的泛化性。

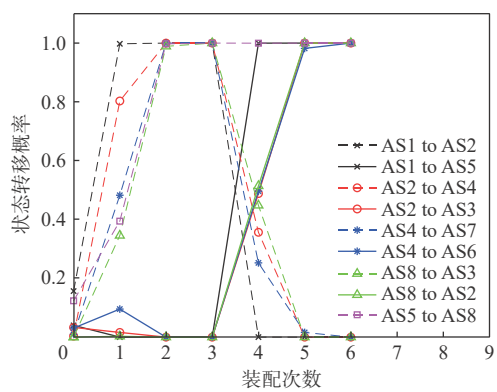


图16 零训练肢体动作序列增量训练部分状态转移概率

Fig. 16 State transition probability of incremental training part of zero training limb action sequence

5 结 论

1）采集IMU传感器加速度和欧拉角信号，对

操作员的 LTB、RTW、LBS、LMS、LSS、RBT、RMT、RST 肢体动作进行分析, 提取传感器时域和时频域特征, PSO-SVM 算法离线识别肢体动作的正确率达到 96.7%。结果表明, 基于时域和时频域特征的 PSO-SVM 算法进行肢体动作识别, 并将肢体动作识别序列作为 HMM 的输入是可行的。

2) 参数可变的 HMM 算法监督训练参数 λ 后, 设置 6 次迭代次数并在过程中增量训练序列数据, 可以实时改变状态转移矩阵和初始状态矩阵的参数, 进而将一个装配序列成功转变为另一个装配序列。结果表明, 参数可变的 HMM 算法可以实现装配顺序的可变性。

3) 零训练肢体动作序列实验中, 利用参数可变的 HMM 算法实现从一个装配序列到另一个装配序列的成功过渡, 减少装配序列监督训练负担, 提高人机协作装配模式的泛化性。

本文仅考虑肢体动作序列的预测, 提前预测到某些肢体动作, 机器人执行协作任务, 在未来的工作中, 可以考虑肢体动作识别技术融合更精确时间预测。

参考文献（References）

[1] ZHOU J, LI P G, ZHOU Y H, et al. Toward new-generation intelligent manufacturing[J]. Engineering, 2018, 4(1): 11-20.

[2] 黄海丰, 刘培森, 李擎, 等. 协作机器人智能控制与人机交互研究综述[J]. 工程科学学报, 2022, 44(4): 780-791.
HUANG H F, LIU P S, LI Q, et al. Review: intelligent control and human-robot interaction for collaborative robots[J]. Chinese Journal of Engineering, 2022, 44(4): 780-791(in Chinese).

[3] WANG P, LIU H Y, WANG L H, et al. Deep learning-based human motion recognition for predictive context-aware human-robot collaboration[J]. CIRP Annals, 2018, 67(1): 17-20.

[4] 常玉青. 人机协作中基于多 Kinect 的人体行为识别研究[D]. 长春: 吉林大学, 2018: 5-7.
CHANG Y Q. Research on human behavior recognition based on multi-Kinect in human-computer cooperation[D]. Changchun: Jilin University, 2018: 5-7(in Chinese).

[5] LU G L, ZHOU Y Q, LI X Y, et al. Efficient action recognition via local position offset of 3D skeletal body joints[J]. Multimedia Tools and Applications, 2016, 75(6): 3479-3494.

[6] SHI M, YANG C Y, ZHANG D L. A novel human-machine collaboration model of an ankle joint rehabilitation robot driven by EEG signals[J]. Mathematical Problems in Engineering, 2021, 2021(1): 5564235.

[7] ÇOBAN M, GELEN G. Realization of human-robot collaboration in hybrid assembly systems by using wearable technology[C]//Proceedings of the 6th International Conference on Control Engineering & Information Technology. Piscataway: IEEE Press, 2018: 1-6.

[8] SYED A S, SHERHAN Z, SHEHRAM M, et al. Using wearable sensors for human activity recognition in logistics: a comparison of different feature sets and machine learning algorithms[J]. International Journal of Advanced Computer Science and Applications, 2020, 11(9): 644-649.

[9] WANG W T, LI R, DIEKEL Z M, et al. Controlling object hand-over in human-robot collaboration via natural wearable sensing[J]. IEEE Transactions on Human Machine Systems, 2019, 49(1): 59-71.

[10] 陈友东, 刘嘉蕾, 胡澜晓. 人机协作中人的动作终点预测[J]. 北京航空航天大学学报, 2019, 45(1): 35-43.
CHEN Y D, LIU J L, HU L X. Human motion end point prediction in human-robot collaboration[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(1): 35-43(in Chinese).

[11] LIU Z T, LIU Q, XU W J, et al. Deep learning-based human motion prediction considering context awareness for human-robot collaboration in manufacturing[J]. Procedia CIRP, 2019, 83: 272-278.

[12] KWON W Y, SUH I H. Planning of proactive behaviors for human-robot cooperative tasks under uncertainty[J]. Knowledge-Based Systems, 2014, 72: 81-95.

[13] HAWKINS K P, VO N, BANSAL S, et al. Probabilistic human action prediction and wait-sensitive planning for responsive human-robot collaboration[C]//Proceedings of the 13th IEEE-RAS International Conference on Humanoid Robots. Piscataway: IEEE Press, 2013: 499-506.

[14] LIU H Y, WANG L H. Human motion prediction for human-robot collaboration[J]. Journal of Manufacturing Systems, 2017, 44: 287-294.

[15] CRAMER M, CRAMER J, KELLENS K, et al. Towards robust intention estimation based on object affordance enabling natural human-robot collaboration in assembly tasks[J]. Procedia CIRP, 2018, 78: 255-260.

[16] ZANCHETTIN A M, CASALINO A, PIRODDI L, et al. Prediction of human activity patterns for human-robot collaborative assembly tasks[J]. IEEE Transactions on Industrial Informatics, 2018, 15(7): 3934-3942.

[17] 祁亨年. 支持向量机及其应用研究综述[J]. 计算机工程, 2004, 30(10): 6-9.
QI H N. Support vector machines and application research overview[J]. Computer Engineering, 2004, 30(10): 6-9(in Chinese).

[18] CHANG C C, LIN C J. LIBSVM[J]. ACM Transactions on Intelligent Systems and Technology, 2011, 2(3): 1-27.

[19] YE M Y, YAN X A, JIA M P. Rolling bearing fault diagnosis based on VMD-MPE and PSO-SVM[J]. Entropy, 2021, 23(6): 762.

[20] TONG L N, SONG Q J, GE Y J, et al. HMM-based human fall detection and prediction method using tri-axial accelerometer[J]. IEEE Sensors Journal, 2013, 13(5): 1849-1856.

Dynamically changeable human-robot collaborative assembly based on limb motion prediction

WU Haibin[1, 2, *], SONG Chenyang[1], ZHOU Shixuan[1]

(1. School of Advanced Manufacturing, Fuzhou University, Quanzhou 362200, China;
2. Fujian Key Laboratory of Special Intelligent Equipment Safety Measurement and Control, Fuzhou 350003, China)

Abstract: Human-robot collaborative assembly is necessary to ensure that the assembly process goes successfully in some complex assembly areas where it is challenging to rely entirely on robots to complete the assembly. However, in the process of human-robot collaboration, the robot cannot respond to different human operations. We propose a dynamic variable human-robot collaborative assembly scheme to address the challenges in complex assembly domains. The scheme achieves dynamic variability in the assembly process by recognizing and predicting the operator's limb movements, and the robot responds accordingly. We use an inertial measurement unit (IMU) to collect motion information and propose a particle swarm optimization (PSO) support vector machine (SVM) algorithm for accurate limb motion detection. In addition, we introduce a parameter-variable hidden Markov model (HMM) to predict action sequences in real-time. This allows the robot to infer the operator's future intentions. The decelerator human-robot collaborative assembly experiment showed that the suggested approach obtained an average limb motion recognition rate of 96.7%. It was also able to effectively predict operator actions, enabling dynamic human-robot collaborative assembly with adjustable assembly sequences. This significantly enhanced the adaptability of the assembly system and reduced the complexity of human-robot collaborative assembly.

Keywords: motion control; wearable sensors; human-machine collaborative assembly; dynamically variable sequence; hidden Markov model

Received: 2023-08-24; **Accepted**: 2023-12-29; **Published Online**: 2024-01-31 17:46

URL: link.cnki.net/urlid/11.2625.V.20240131.1038.001

Foundation items: National Key Research and Development Program of China (2018YFB1308603); Fujian Provincial Science and Technology Major Project of China (2024HZ026020)

* **Corresponding author**. E-mail: wuhb@fzu.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0575

面向点融合运行的多航空器低碳轨迹协同规划

王超*，李雯清，孙石磊，徐晨洋
(中国民航大学 空中交通管理学院，天津 300300)

摘 要：为解决在高交通密度终端区实施连续下降运行(CDO)的技术难题，提出一种面向点融合系统(PMS)的多航空器低碳轨迹协同规划方法。分析了基于管制员人工决策的点融合运行的额外碳排放产生机理；以个体航空器低碳4D轨迹为研究对象，建立了一种考虑PMS平飞弧段的CDO多阶段最优控制模型；以总体运行时间最小为目标，考虑进场交通流冲突约束，建立了多航空器CDO 4D轨迹协同规划模型。以实际PMS及航迹数据为例开展验证实验，结果表明：所提方法共减少12.2%的空中飞行时间，降低15.7%的燃油消耗和CO_2排放，不仅能给出支持高效运行的航空器排序结果，还能提供达成该排序的可行低碳4D轨迹。

关 键 词：航空运输；4D轨迹；最优控制理论；连续下降运行；点融合系统；排序调度
中图分类号：V355
文献标志码：A **文章编号**：1001-5965(2025)10-3253-09

随着人们对空中交通运行优化减排途径的不断探索，非繁忙时段连续下降运行(continuous descent operation, CDO)的节油减排效果已经得到实证[1-4]，但如何在高交通密度终端区实施CDO仍是一个需要解决的难题。

为实现航空器优化排序与降低环境影响的双重目标，研究者尝试将进场航空器优化排序问题与CDO结合起来，主要形成2种方法。一种方法是在获得航空器所需到达时间的基础上生成优化的CDO剖面。Solak等[5]采用随机规划方法优化了进场交通流汇聚点的位置和数量，并在此基础上设计了CDO程序。Pawelek等[6]通过混合整数规划方法求解进场航空器调度问题，在获得所需到达时间的基础上再计算各航空器的CDO剖面。吕波等[7]在获取所需到达时间的基础上提出了基于卡尔曼滤波的进场航空器4D轨迹规划模型。Toratani[8]提出了一种基于汇聚优化的新型进场管理算法，采用滚动时域控制法对动态进场航班进行优化排序，建立最优控制模型生成航空器轨迹，由于缺少结构化航线的约束，水平轨迹过于机动，难以满足可飞性的要求。另一种方法是直接进行多航空轨迹优化。Dahlberg等[9]认为进场航线的预战术规划是智能空中交通的发展方向，设计了基于网格的进场航线树状结构，并建立了混合整数规划模型来求解。Gui等[10]提出了一种结合水平路径延伸策略和CDO垂直剖面优化的航空器调度模型，促进了繁忙终端区内CDO程序的实施。Sáez等[11]以标准进场航线为研究对象，构建了多个备选水平进场航线及其CDO剖面，采用混合整数规划模型实现了汇聚轨迹优化；并在文献[11]的基础上将终端区划分为预排序区与动态轨迹区，提高了飞行轨迹可预测性和CDO运行效率，但预排序区的存在可能增加航空器飞行时间[12]。

点融合系统(point merge system, PMS)可在一定程度上减少雷达引导产生的轨迹不确定性，降低空中交通复杂性，因而得到广泛应用。随着点融合

收稿日期：2023-09-07；**录用日期**：2023-11-03；**网络出版时间**：2023-11-22 11:48
网络出版地址：link.cnki.net/urlid/11.2625.V.20231121.1154.003
基金项目：天津市应用基础研究多元投入基金重点项目(21JCZDJC00780)
* **通信作者**. E-mail：wangch@cauc.edu.cn

引用格式：王超，李雯清，孙石磊，等. 面向点融合运行的多航空器低碳轨迹协同规划[J]. 北京航空航天大学学报，2025，51(10)：3253-3261.
WANG C，LI W Q，SUN S L，et al. Multi-aircraft low-carbon trajectories cooperative planning for point merge operation[J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51(10)：3253-3261 (in Chinese).

运行模式的逐渐推广，将点融合运行与 CDO 结合起来的方法开始涌现。Liang 等[13]设计了多层点融合进场航线结构，尝试将 CDO 与 PMS 结合起来，采用模拟退火算法实现了水平航线和垂直剖面协同优化，解决了排序和轨迹同时优化问题，但缺少非线性最优控制模型对下降剖面优化的支撑。张军峰等[14]构建了基于 PMS 的帝国竞争优化调度算法，并研究了 PMS 的 4D 轨迹规划问题，是对未来基于轨迹运行的有效探索。杨磊等[15]设计了一种基于 PMS 理念的新型倒皇冠形进场航线结构，提出了面向 CDO 的航空器进场轨迹多目标优化模型，并采用非支配排序遗传算法求解，扩展了 PMS 的设计理念，但未构建 PMS CDO 的最优控制模型，对轨迹燃油消耗和温室气体排放优化不充分。

以上研究多以进场航空器调度排序为出发点，未能明确调度排序的本质是一个多航空器轨迹规划问题。本文提出一种面向点融合运行的进场航空器排序和轨迹同时优化方法。首先，提出 PMS 中的 CDO 问题，建立水平备选航线集合，构建水平备选航线的最优控制模型；然后，将多航空器轨迹规划问题转化为以运行时间最小为目标的 4D 轨迹组合优化问题，进而同时获得多航空器优化排序。低碳 4D 轨迹规划的研究有助于充分发挥 PMS 的减排作用，促进实现低碳空中交通控制系统的完全自动化，并最终实现基于轨迹的运行。

1 点融合系统中的连续下降运行问题

1.1 点融合系统

PMS 是欧洲航行安全组织提出的一种基于卫星导航的新型进场航线结构，为进场航空器排序和连续下降提供了一种系统化方法，近年来逐渐得到推广应用。

典型的 PMS 由内外排序弧及融合点 (merge point, MP) 组成，如图 1 所示。排序弧是一段与融合点距离相等的飞行弧段，用于调节航空器飞行航线长度。出于建立足够的垂直间隔的考虑，内排序弧高度一般高于外排序弧若干个高度层，并要求航空器在弧上飞行时高度与速度不变。融合点是终端区内设置的一个定位点。从不同方向到达的航空器沿其进场航线飞入排序弧，在收到管制员转弯指令后，直飞融合点。航空器在直飞融合点时形成汇聚，因此，融合点可以作为航空器排序的一个参考点。

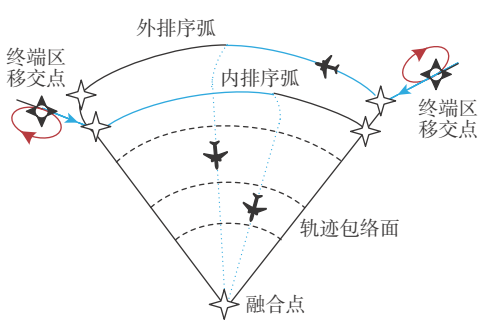


图 1 点融合系统

Fig. 1 Point merge system

进场航空器在终端区内飞行轨迹可以概括为 2 个部分：终端区移交点处等待航线（见图 1 中红色线段）和进场飞行航线（见图 1 中蓝色线段）。其中，进场飞行航线又包括进场航段、排序弧和直飞融合点航段 3 个子段。

1.2 点融合运行的下降剖面

在已知未来飞行距离的情况下，机载飞行管理系统可以给出理想的垂直剖面。考虑到 PMS 的平飞弧段，理想的 CDO 过程可划分为 5 个独立阶段，如图 2 所示。图中，左轴表示飞行高度 h，下轴表示航空器持续下降运行时间 t，上轴表示航空器与融合点的水平距离 d，$t^{(p)}$ 为第 p 阶段结束时间，$d^{(p)}$ 为航空器在第 p 阶段末端与融合点的水平距离，h_c 为航空器进入终端区时的平飞高度，h_a 为航空器在排序弧上的飞行高度，h_{mp} 为航空器到达融合点的限制高度。

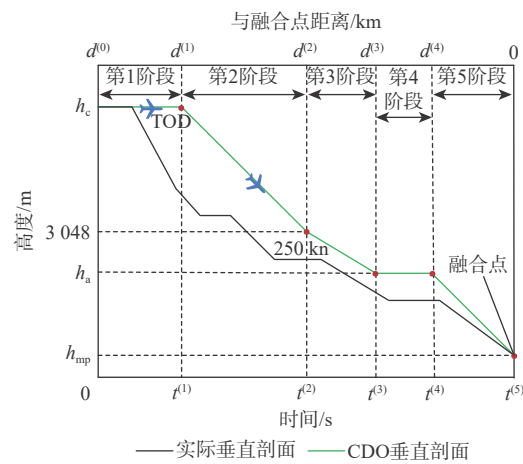


图 2 点融合系统的 CDO 程序

Fig. 2 CDO pattern of PMS

由于航空器在终端区移交点处等待航线上保持飞行高度不变，CDO 过程在进场飞行阶段才得以实际执行。CDO 前 3 个阶段对应图 1 中的进场航段。在第 1 阶段，航空器从移交点进入终端区，保持等表速平飞至下降顶点 (top of descend, TOD)；大型商用航空器飞行操作规章一般要求航空器低于 3 048 m（相当于 10 000 ft）时校正空速不能超过 250 kn，因此，按照该标准划分第 2 阶段；在第 3 阶段，航空器减速下降至排序弧高度；与传统 CDO 不同的是，点

融合运行下降过程存在排序弧平飞段, 因此, 增加了第 4 阶段, 在该阶段航空器速度、高度均保持不变; 第 5 阶段对应直飞融合点航段, 在该阶段航空器减速飞行到融合点, 并需达到融合点处速度和高度的限制要求。

然而, 在实际运行中, 多架航空器向融合点汇聚时可能导致飞行冲突。在此情形下, 管制员采取改变高度和航向等指令对航空器实施控制, 造成未来飞行距离不确定, 导致先进机载飞行管理系统无法发挥对下降剖面的优化作用, 使得实际垂直轨迹偏离其理想剖面, 产生额外的燃油消耗和 CO_2 排放。因此, 如何消除多架进场航空器排序造成的飞行路线不确定性, 减少低空平飞段, 在实现个体轨迹垂直剖面优化的基础上实现多航空器轨迹协同规划的问题被提出。

2　进场航空器 4D 轨迹优化

航空器能对下降剖面优化的前提是事先确定剩余飞行航线的长度。在 PMS 中, 航空器剩余飞行距离的不确定性来源于其在排序弧上的飞行长度。由于人类认知能力的限制, 管制员无法提前预测完整的航空器进场航线。因此, 为对进场航线进行预战术规划, 可通过对排序弧等距分割的方式创建若干水平备选航线, 并求解水平备选航线的优化下降剖面, 最终形成进场航空器的低碳 4D 轨迹集合。

2.1　水平备选航线生成

通过对排序弧等距分割的方式构建航空器水平备选航线集合 J, 其过程如图 3 所示。

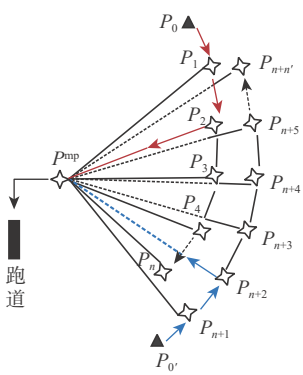


图 3　水平备选进场航线生成过程

Fig. 3　Horizontal alternative arrival route generation process

首先, 假设内排序弧弧长为 S, 外排序弧弧长为 S', 水平管制间隔标准为 S_0。由式 (1) 来确定内外排序弧分割点个数:

$$\begin{cases} n = \left\lfloor \dfrac{S}{S_0} \right\rfloor + 1 \\ n' = \left\lfloor \dfrac{S'}{S_0} \right\rfloor + 1 \end{cases} \tag{1}$$

式中: n 为内排序弧分割点个数(对应图 3 中点 $P_1,\cdots,P_n$); n'为外排序弧分割点个数(对应图 3 中点 $P_{n+1},\cdots,P_{n+n'}$)。

其次, 在确定每条排序弧上分割点的数量后, 可根据式 (2) 分别确定内排序弧和外排序弧中相邻分割点之间的弧长 ΔS 和 $\Delta S'$。

$$\begin{cases} \Delta S = \dfrac{S}{n-1} \\ \Delta S' = \dfrac{S'}{n'-1} \end{cases} \tag{2}$$

最后, 根据分割点和每段弧长构建水平备选进场航线集合 J。$j\in J, j=\{1,2,\cdots,n,n+1,n+2,\cdots,n+n'\}$, 其中, 内排序弧水平备选航线共 n 条, 外排序弧水平备选航线共 n'条; 当 $j\in\{1,2,\cdots,n\}$时, j 表示内排序弧相关的水平备选进场航线, 航线 j 由序列点 $P_0P_1\cdots P_jP^{mp}$ 构成, 如图 3 中红色线航线表示 j=2 的情况; 当 $j\in\{n+1,n+2,\cdots,n+n'\}$时, 表示外排序弧相关的水平备选进场航线, 航线 j 由序列点 $P_{0'}P_{n+1}\cdots P_jP^{mp}$ 组成, 如图 3 中蓝色航线表示 j=n+2 的情况。

通过对排序弧进行等距分割, 可为每架到达航空器提供多条可供选择的水平备选航线。针对每条水平备选航线 j, 可对其 CDO 轨迹优化问题进行建模和求解, 以确定航空器使用该航线时的最佳垂直剖面。

2.2　基于水平备选航线的垂直剖面最优控制模型

在给定机型、进场高度、进场速度等初始条件下, 点融合运行的 CDO 垂直剖面优化问题可转化为多阶段非线性最优控制问题。在给定水平备选航线 j 水平距离的基础上, 考虑排序弧平飞段, 并根据航空器运行规则确定相应约束条件, 建立符合 PMS 特点的 CDO 剖面多阶段非线性最优控制模型。

1) 目标函数。航空器燃油消耗与 CO_2 排放量存在简单比例关系[16], 航油的 CO_2 排放指数为 3.15 kg/kg, 即每燃烧 1 kg 燃油生成的 CO_2 量(kg)。以 CO_2 排放最少为目标对航空器进场航线垂直剖面进行优化, 目标函数为

$$\min E = 3.15\left(f^{(1)}\frac{d^{(0)}-d^{(1)}}{v_{\mathrm{T}}^{(1)}} + f^{(4)}\frac{d^{(3)}-d^{(4)}}{v_{\mathrm{T}}^{(4)}} + \sum_{p=2,3,5}\int_{t^{(p-1)}}^{t^{(p)}} f^{(p)}\mathrm{d}t \right) \tag{3}$$

式中: E 为 CO_2 排放量; $f^{(p)}$ 为第 p 阶段燃油消耗率; $v_{\mathrm{T}}^{(p)}$为航空器在第 p 阶段平飞真空速。

2) 控制变量。在优化控制模型中, 航空器状态 $\boldsymbol{X}$ 可由真空速 v_{T}、航空器垂直速度分量 $\dot{h}$、水平速度分量 $\dot{d}$定义, $\boldsymbol{X}=[v_{\mathrm{T}},\dot{h},\dot{d}]^{\mathrm{T}}$, 如式 (4) 所示。由于下滑角 γ 决定航空器状态, 其可作为模型中唯一的控制变量。

$$\begin{cases} v_{\mathrm{T}} = \dfrac{1}{m_0}(T-D) - g\sin\gamma \\ \dot{h} = v_{\mathrm{T}}\sin\gamma \\ \dot{d} = v_{\mathrm{T}}\cos\gamma \end{cases} \tag{4}$$

式中：T为发动机推力；D为阻力；m_0为航空器质量；g为重力加速度。

推力T和阻力D分别为

$$\begin{cases} T = C_{\mathrm{des}}C_{\mathrm{T1}}\left(1 - \dfrac{h}{C_{\mathrm{T2}}} + C_{\mathrm{T3}}h^2\right) \\ D = \left[C_{\mathrm{D0}} + C_{\mathrm{D2}}\left(\dfrac{2m_0 g}{\rho v_{\mathrm{T}}^2 S_{\mathrm{w}}}\right)^2\right]\rho v_{\mathrm{T}}^2 S_{\mathrm{w}}/2 \end{cases} \tag{5}$$

式中：C_{des}为推力修正系数；C_{T1}为第一推力系数；C_{T2}为第二推力系数；C_{T3}为第三推力系数；C_{D0}为零升阻力系数；C_{D2}为诱导阻力系数；ρ为大气密度；S_{w}为机翼参考面积。各系数和机型有关，均由BADA文件中的数据得到[17]。

3）约束条件。考虑到高度、温度、静压等因素对实际空速的影响，校正空速更能准确地反映航空器操纵特性，因此，在垂直剖面最优控制问题中，使用校正空速来表示航空器速度。在国际标准大气条件下，校正空速v_{C}与真空速v_{T}的转换关系为

$$v_{\mathrm{T}} = v_{\mathrm{C}} \Bigg/ \sqrt{\left(1 - \frac{0.006\,5h}{288.15}\right)^{4.257\,9}} \tag{6}$$

① 阶段边界约束。飞行距离、高度等参数在各阶段的取值须限制在相应范围内，如下：

$$\begin{cases} d^{(5)} = 0 \\ h_0^{(1)} = h_{\mathrm{c}} \\ h_{\mathrm{f}}^{(2)} = 3\,048\ \mathrm{m} \\ h_{\mathrm{f}}^{(5)} = h_{\mathrm{mp}} \end{cases} \tag{7}$$

式中：$h_0^{(p)}$为第p阶段初始高度；$h_{\mathrm{f}}^{(p)}$为第p阶段末端高度。

② 下降率约束。考虑到航空器飞行安全、操作稳定性及旅客舒适性，商用航空器的平均下降率β必须限制在一定范围内，如下：

$$\beta_{\min}^{(p)} \leqslant \beta^{(p)} \leqslant \beta_{\max}^{(p)} \qquad p \in [2,3,5] \tag{8}$$

式中：$\beta^{(p)}$为第p阶段平均下降率；$\beta_{\max}^{(p)}$、$\beta_{\min}^{(p)}$分别为第p阶段平均下降率的上限、下限。

③ 速度约束。航空器在终端区内下降时，在不同阶段须满足相应的速度限制，如下：

$$\begin{cases} 250\ \mathrm{kn} \leqslant v_{\mathrm{C}}^{(p)} \leqslant v_{\mathrm{C}}^{(0)} & p \in [1,2] \\ v_{\mathrm{f}}^{(5)} \leqslant v_{\mathrm{C}}^{(p)} \leqslant 250\ \mathrm{kn} & p \in [3,5] \\ v_{\mathrm{C}}^{(4)} = v_{\mathrm{C}}^{\mathrm{arc}} \\ v_{\mathrm{f}}^{(5)} = v_{\mathrm{C}}^{\mathrm{mp}} \end{cases} \tag{9}$$

式中：$v_{\mathrm{C}}^{(p)}$为航空器在第p阶段时校正空速；$v_{\mathrm{C}}^{(0)}$为进入终端区时校正空速；$v_{\mathrm{f}}^{(p)}$为第p阶段末端校正空速；$v_{\mathrm{C}}^{\mathrm{arc}}$为排序弧上校正空速；$v_{\mathrm{C}}^{\mathrm{mp}}$为融合点处校正空速。

④ 下滑角约束。为避免陡峭下滑角给旅客带来的不适，下滑角须限制在一定范围内，如下：

$$\gamma_{\min} \leqslant \gamma \leqslant \gamma_{\max} \tag{10}$$

式中：$\gamma_{\max}$、$\gamma_{\min}$分别为下滑角的上限、下限。

⑤ 连接约束。最优控制模型将航空器完整的飞行过程划分为不同阶段，而实际飞行却是连续不可分割的，因此，航空器在相邻航段的临界点前后须保证状态一致，如下：

$$X_{\mathrm{f}}^{(p-1)} = X_0^{(p)} \qquad p \in [2,3,4,5] \tag{11}$$

式中：$p \in \{2,3,4,5\}$，下标0表示当前阶段p下的初始时刻，下标f表示当前阶段p下的最终时刻。

在不考虑其他航空器影响的条件下，以上模型给出了给定水平备选航线的最优下降剖面。给定水平备选航线和最优下降剖面共同构成完整的最优4D轨迹，其中，水平备选航线决定了航空器的飞行距离，最优下降剖面决定了航空器的飞行高度和速度。利用该模型，给定执飞机型和水平备选航线便可获得理想条件下的最优4D轨迹集合。

3 多航空器4D轨迹协同规划

以往的进场航空器优化排序问题研究一般仅给出排序结果，而未给出实现该排序的各航空器飞行轨迹，从这个角度来看，其是一个尚未得到完全解决的问题，本质是多航空器轨迹规划问题。

第2节输出的个体最优值并未考虑其他航空器的影响，在实际运行中，多航空器间存在冲突的可能性。因此，在第2节构建的水平备选进场4D轨迹集合的基础上，建立多航空器4D轨迹组合优化模型，为每架航空器提供一条排放最小的4D轨迹，最终获得总体运行时间最小、无冲突、排序优化的多航空器低碳4D轨迹可行解集。

3.1 决策变量与目标函数

1）决策变量。假设I为进入终端区的航空器集合。$i \in I$，$i=\{1,2,\cdots,m,m+1,m+2,\cdots,m+m'\}$。$i \in \{1,2,\cdots,m\}$时，表示从相关移交点沿其进场航线进入内排序弧的航空器，共m架；$i \in \{m+1, m+2,\cdots, m+m'\}$时，表示从另一移交点沿其进场航线进入外排序弧的航空器，共m'架。进入相同排序弧的多架进场航空器构成了同一交通流。

为标识航空器i是否选择航线j，引入0-1变量x_{ij}，如下：

$$x_{ij} = \begin{cases} 1 & \text{航空器}i\text{选择航线}j \\ 0 & \text{航空器}i\text{未选择航线}j \end{cases} \tag{12}$$

2）目标函数。为实现进场效率最大化，以全部航空器总运行时间最少为目标构造目标函数，如下：

$$\min Z=\sum_{i\in I}\sum_{j\in J}x_{ij}(w_i+t_{ij}^{(5)}) \tag{13}$$

式中：w_i 为航空器 i 在终端区移交点的等待时间，对应图 1 的第 1 部分航线时间；$t_{ij}^{(5)}$ 为航空器 i 在水平备选航线 j 上的飞行时间，对应图 1 第 2 部分航线时间。

3.2　约束条件

1）航线唯一性约束。从移交点进入内排序弧的航空器能且只能选择一条内排序弧水平备选航线，从移交点进入外排序弧的航空器能且只能选择一条外排序弧水平备选航线，如下：

$$\begin{cases}\sum\limits_{j=1}^{n}x_{ij}=1 & \forall i=1,2,\cdots,m\\ \sum\limits_{j=n+1}^{n+n'}x_{ij}=1 & \forall i=m+1,m+2,\cdots,m+m'\end{cases} \tag{14}$$

2）移交点等待时间约束。假设航空器 i 在终端区移交点处执行标准等待程序，标准等待时长为 1 圈 300 s，等待时间 w_i 取决于航空器 i 的标准等待圈数 c_i，如下：

$$w_i=300c_i \qquad c_i\in \mathrm{N} \tag{15}$$

3）同一交通流航空器尾随间隔约束。为判断航空器 i、i'是否属于同一交通流，引入 0-1 变量 $e_{ii'}$，如下：

$$e_{ii'}=\begin{cases}1 & \text{航空器}i,i'\text{属于同一交通流}\\ 0 & \text{航空器}i,i'\text{属于不同交通流}\end{cases} \tag{16}$$

同一交通流内任意 2 架航空器在对应排序弧入口处的时间间隔不能小于规定的安全间隔标准，如下：

$$\begin{cases}t_i^{\mathrm{arc}}=x_{ij}(t_i^{\mathrm{se}}+w_i+t_{ij}^{(3)})\\ \left|t_i^{\mathrm{arc}}-t_{i'}^{\mathrm{arc}}\right|\geqslant s_{ii'}e_{ii'}\end{cases} \tag{17}$$

式中：t_i^{se} 为航空器 i 到达终端区移交点时间；$t_{ij}^{(3)}$ 为航空器 i 从移交点到排序弧入口的飞行时间，即在水平备选航线 j 上完成 CDO 前 3 个阶段的时间；t_i^{arc}、$t_{i'}^{\mathrm{arc}}$ 分别为航空器 i、i'到达排序弧入口的时刻；$s_{ii'}$ 为航空器 i 和 i'之间须遵循的安全间隔标准。

4）不同交通流的汇聚冲突约束。在融合点处，来自不同排序弧的航空器之间存在冲突的可能性，因此，相继飞行的航空器之间的时间间隔不能小于规定的安全间隔标准，如下：

$$\begin{cases}t_i^{\mathrm{mp}}=x_{ij}(t_i^{\mathrm{se}}+w_i+t_{ij}^{(5)})\\ \left|t_i^{\mathrm{mp}}-t_{i'}^{\mathrm{mp}}\right|\geqslant s_{ii'}\end{cases} \tag{18}$$

式中：t_i^{mp}、$t_{i'}^{\mathrm{mp}}$ 分别为航空器 i、i'到达融合点处时间。

4　实例验证

4.1　点融合系统及水平备选航线

以上海终端区 1 号 PMS 为研究对象。该 PMS 主要汇聚 MATNU 方向和 AND&BK 方向进场的交通流；其中，MATNU 方向进场的交通流沿进场航线 MAT81A 进入内排序弧，在内排序弧上飞行高度为 2 400 m，AND&BK 方向进场的交通流沿进场航线 AND82A 进入外排序弧，在外排序弧上飞行高度为 2 100 m。航空器在排序弧上飞行限制速度为 230 kn。

在上海终端区内运行的机型主要以重型机和中型机为主，因此，以重型机和中型机之间最大的尾流间隔标准作为排序弧分割指标，结合 ICAO 发布的无风情况下各类航空器之间的最小尾流间隔标准规范[18] 可知，该指标 S_0=9.26 km。

对排序弧进行等距分割，结果如图 4 所示。内排序弧被分割为 4 段，生成 5 个内排序弧分割点，从 MATNU 进场的航班有 5 条内排序弧水平备选航线，航线编号为 j=1,2,⋯,5；外排序弧被分割为 5 段，生成 6 个外排序弧分割点，从 AND&BK 进场的航班有 6 条外排序弧水平备选航线，航线编号为 j=6,7,⋯,11。

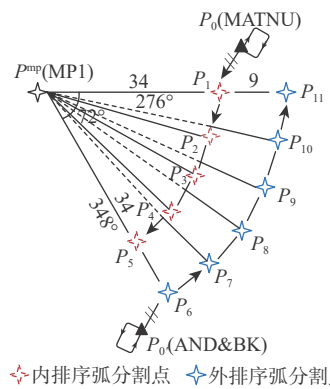


图 4　1 号点融合系统的水平备选航线

Fig. 4　Horizontal alternative routes of No.1 PMS

4.2　个体航空器连续下降运行轨迹优化

根据 2.2 节提出的最优控制模型，以执飞机型为 A320 的航班进场过程为例，对其在 6 条外排序弧水平备选航线下的垂直剖面进行优化。该航班在终端区移交点的高度为 6 900 m，速度为 300 kn。优化过程中，使用 BADA[17] 飞机性能数据，A320 机型的质量为 64 000 kg，下滑角 $\gamma\in[2°,6°]$，下降率 $\beta\in[3,9]$ m/s。优化后，轨迹的燃油消耗和 CO_2 排放由最优控制模型求解计算，航班实际轨迹燃油消耗和 CO_2 排放由美国联邦航空局的航空环境设计软件计算。

以 CO_2 消耗最少为目标，采用 GPOPS 求解该

最优控制模型。高度剖面优化结果如图5所示。可以看出，与实际垂直剖面相比，在不同水平备选进场航线下，航空器使用CDO程序可以在较高的高度保持平飞，下降过程中的平飞段显著减少。

图6给出航空器在航线j=7下执行CDO程序的校正空速与下滑角。可以看出：①航空器下降后尽可能多地保持低速飞行状态，以减少燃油消耗和CO_2排放；②航空器从TOD点高度下降到3 048 m（10 000 ft）期间，下滑角变化较大，高度低于3 048 m，航空器校正空速不高于250 kn；③航空器连续下降到排序弧期间，减速至排序弧上限制速度并保持等速平飞，此时下滑角变化较小。

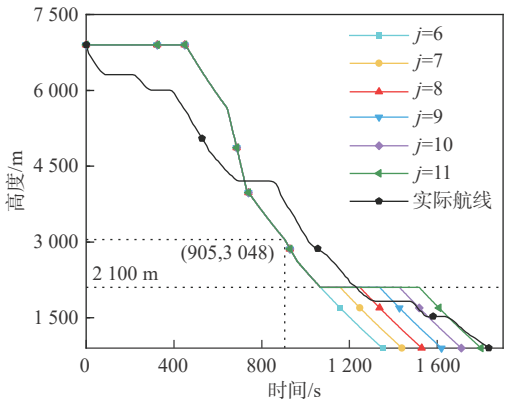


图5 A320机型的实际和优化后垂直剖面

Fig. 5 Actual and optimized vertical profile of A320

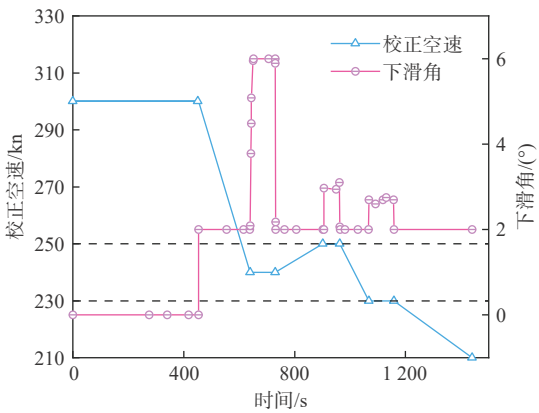


图6 校正空速与下滑角

Fig. 6 Calibrate air speed and flight path angle

表1给出了实际与在不同水平备选航线下优化后的CO_2排放和飞行时间对比。可以看出，所有备选进场航线的CDO时间及排放均优于实际情况，燃油消耗降低6.0%～34.7%，飞行时间减少1.9%～26.4%，CO_2排放减少6.0%～34.7%。

表1 实际与优化后CO_2排放和飞行时间

Table 1 Actual and optimized CO_2 emissions and flight time

航线	飞行时间/s	燃油消耗/kg	CO_2排放/kg
实际航线	1 835	249.14	784.79
j=6	1 350	162.72	512.50
j=7	1 440	177.02	557.61
j=8	1 530	191.32	602.66
j=9	1 620	205.62	647.70
j=10	1 710	219.92	692.75
j=11	1 800	234.22	737.79

同理，可对任一航空器执飞任一进场水平备选航线进行连续下降剖面优化。

4.3 多航空器低碳轨迹规划模型求解

选取上海浦东国际机场2023年5月7日11:00—13:00时间段内28架进场航空器的历史轨迹，其中，从AND&BK进场的航空器有15架，从MATNU进场的航空器有13架，中型机22架，重型机6架。通过轨迹数据处理得到不同航班机型、进场时刻、进场高度等初始条件，具体数据如表2所示。

表2 28个航班初始信息

Table 2 Initial conditions of 28 flights

航空器编号i	机型	进场方向	进场时刻	进场高度/m	进场速度/kn	航空器编号i	机型	进场方向	进场时刻	进场高度/m	进场速度/kn
1	A321	MATNU	11:01	4 800	250	15	A321	AND&BK	11:58	7 200	310
2	B737	MATNU	11:07	4 800	270	16	A319	AND&BK	12:00	7 500	300
3	A320	AND&BK	11:07	7 500	280	17	A321	AND&BK	12:01	7 500	300
4	B787	MATNU	11:08	5 100	260	18	A319	MATNU	12:02	5 400	300
5	A330	MATNU	11:17	5 400	300	19	A320	AND&BK	12:10	6 900	300
6	A330	AND&BK	11:19	7 200	320	20	A321	AND&BK	12:15	6 900	260
7	B777	AND&BK	11:20	6 000	270	21	B737	MATNU	12:22	5 400	260
8	A320	AND&BK	11:36	6 600	260	22	B737	AND&BK	12:27	7 200	300
9	A321	AND&BK	11:38	6 900	280	23	A330	MATNU	12:31	5 100	270
10	A320	AND&BK	11:38	7 500	320	24	A320	MATNU	12:43	5 400	280
11	A321	AND&BK	11:48	7 200	310	25	A320	MATNU	12:46	5 100	300
12	B787	MATNU	11:48	4 800	260	26	A321	MATNU	12:55	5 100	240
13	B737	AND&BK	11:50	6 600	270	27	A320	MATNU	12:59	5 400	290
14	A320	MATNU	11:55	5 100	240	28	A320	AND&BK	13:00	7 200	320

经统计, 28 架航空器在终端区内的实际总运行时间为 620.8 min, 总燃油消耗为 6 996.18 kg, CO_2 总排放为 22 037.97 kg。

使用上述方法求得 28 架航空器在相应水平备选航线下的最优下降剖面, 并得到燃油消耗量、CO_2 排放量及 CDO 运行时间。

在 MATLAB 2018b 的环境下进行实验, 将 28 架航空器在不同水平备选航线下的 CDO 时间作为轨迹规划模型的输入, 并使用遗传算法进行求解。其中, 遗传算法种群规模设为 100, 交叉概率设为 0.8, 变异概率设为 0.2, 每次迭代保留的最优种群比例设为 10%, 最大迭代次数设为 500。

最终实验结果为: 优化后的 28 架航空器总运行时间为 545.1 min, 燃油消耗为 5 897.21 kg, CO_2 排放为 18 576.21 kg。与实际相比, 总运行时间下降了 12.2%, 燃油消耗和 CO_2 排放均减少了 15.7%。

根据实验结果中各航空器的等待时间及所选择的水平备选航线绘制多航空器轨迹规划时空图, 如图 7 所示。圆圈上方数字表示该航空器选择的航线, 如数字 1 表示该航空器选择的航线 j=1; 黄色实线表示该航空器在终端区移交点处无等待; 黄色虚线表示该航空器在终端区移交点处等待。

由图 7 可知: ①在优化后的多航空器进场航线中, 飞行距离较短的航线被选择次数多(如航线 1 和航线 6), 提高了空域利用效率, 降低燃油消耗, 减少 CO_2 排放; ②为保持航空器之间的安全间隔, 算法对距离较长的航线也有所选择(如航线 4 和航线 9); ③航空器在交通流内部及融合点处无冲突, 保证了航空器调度的安全性。

图 8 对比了 28 架航空器实际水平轨迹与优化后水平轨迹。在实际运行中, 管制员未能充分发挥 PMS 的排序潜力, 实际轨迹仍存在秩序性差、不确定性强等不足。优化后的航线可有效协同进场航班, 轨迹一致性大幅提高。

图 9 对比了 28 架航空器实际垂直剖面与优化后垂直剖面。实际下降剖面开始下降时间较早, 且在低高度层仍存在大量平飞段, 而优化后航空器的下降顶点比较靠后, 避免在较低高度层平飞造成的

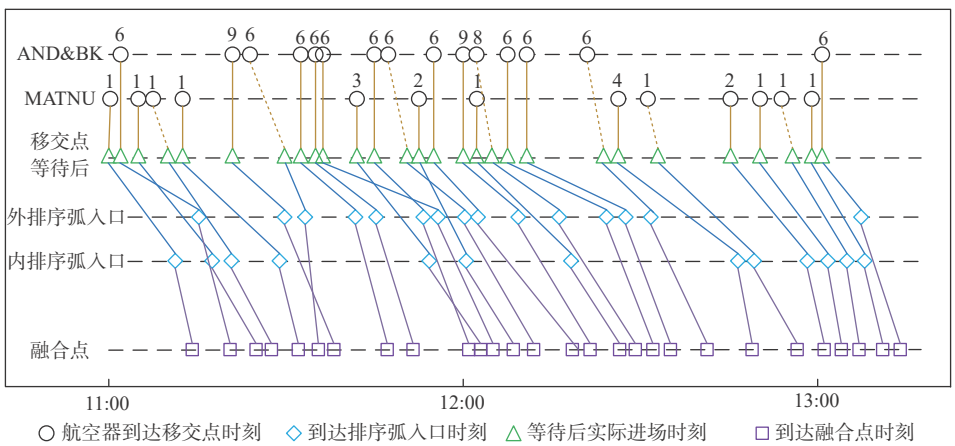


图 7　规划后 28 架进场航空器轨迹时空图

Fig. 7　Time space diagram of planned trajectory of 28 arrival aircraft

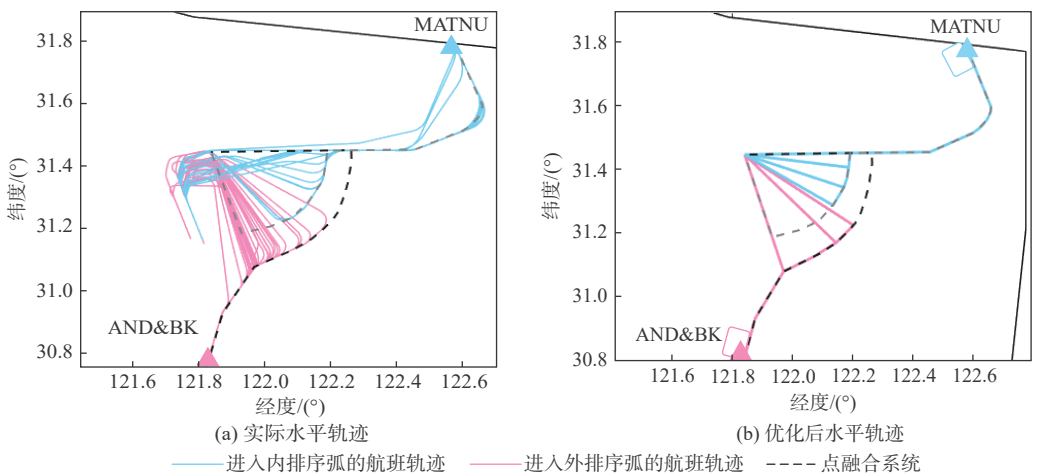


图 8　实际与优化后水平轨迹对比

Fig. 8　Comparison between actual and optimized horizontal trajectory

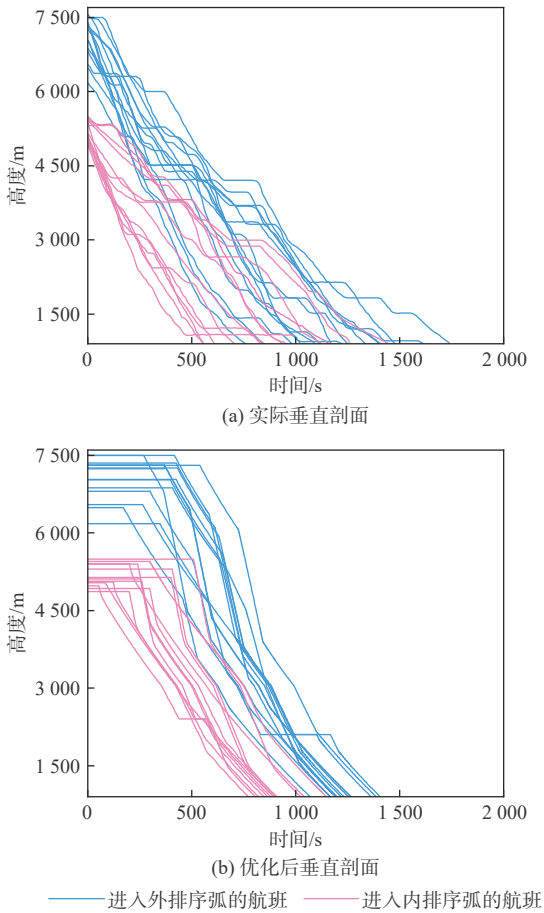


图 9 实际与优化后垂直剖面对比

Fig. 9 Comparison between actual and optimized vertical profile

能源浪费，可有效降低燃油消耗和 CO_2 排放。

以上实验结果分析可知，多航空器低碳 4D 轨迹规划可带来以下优点：①进场航空器排序结果具有可达性，优化结果给出了具体进场轨迹，增强飞行员情景态势感知能力；②提高了飞行轨迹一致性，有效降低了进场交通复杂性；③实现了航空器排序和轨迹同时优化，保证了减排与进场效率最大化指标的达成。

5 结 论

本文针对终端区实际交通运行条件下如何提高进场航空器排序效率和降低环境影响的难点问题开展研究，给出了实现空中交通低碳运行的可行解决方案，得到以下结论：

1）提出了面向点融合运行的多航空器 CDO 4D 轨迹规划方法，能同时实现进场航空器排序和轨迹规划，该方法不仅可以给出运行效率最大的航空器排序结果，还能提供实现该排序结果的具体低碳飞行轨迹。

2）多航空器低碳轨迹协同规划模型估计了进场航空器运行碳排放基准，揭示了目前基于管制员人工决策的点融合运行模式仍具有较大减排提升潜力。

参考文献（References）

[1] ICAO. Continuous descent operations (CDO) manual: Doc 9931[S]. Montréal: ICAO, 2010: 21-36.

[2] PARK S G, CLARKE J P. Optimal control based vertical trajectory determination for continuous descent arrival procedures[J]. Journal of Aircraft, 2015, 52(5): 1469-1480.

[3] DE OLIVEIRA R F, BÜSKENS D C. Benefits of optimal flight planning on noise and emissions abatement at the Frankfurt Airport: AIAA 2012-4482[R]. Reston: AIAA, 2012.

[4] ELLERBROEK J, INAAD M, HOEKSTRA J. Fuel and emission benefits for continuous descent approaches at schiphol[C]//Proceedings of the International Conference on Research in Air Transportation. Nice: ICRAT, 2018: 2-9.

[5] SOLAK S, CHEN H. Optimal metering point configurations for optimized profile descent based arrival operations at airports[J]. Transportation Science, 2018, 52(1): 150-170.

[6] PAWEŁEK A, LICHOTA P, DALMAU R, et al. Fuel-efficient trajectories traffic synchronization[J]. Journal of Aircraft, 2019, 56(2): 481-492.

[7] 吕波, 王超. 改进的扩展卡尔曼滤波在航空器 4D 航迹预测算法中的应用[J]. 计算机应用, 2021, 41(增刊 1): 277-282.
LU B, WANG C. Application of improved extended Kalman filtering in aircraft 4D trajectory prediction algorithm[J]. Journal of Computer Applications, 2021, 41(Sup 1): 277-282(in Chinese).

[8] TORATANI D. Application of merging optimization to an arrival manager algorithm considering trajectory-based operations[J]. Transportation Research Part C: Emerging Technologies, 2019, 109: 40-59.

[9] DAHLBERG J, GRANBERG T A, POLISHCHUK T, et al. Capacity-driven automatic design of dynamic aircraft arrival routes [C]//Proceedings of the IEEE/AIAA 37th Digital Avionics Systems Conference. Piscataway: IEEE Press, 2018: 1-9.

[10] GUI D D, LE M L, HUANG Z C, et al. Optimal aircraft arrival scheduling with continuous descent operations in busy terminal maneuvering areas[J]. Journal of Air Transport Management, 2023, 107: 102344.

[11] SÁEZ R, PRATS X, POLISHCHUK T, et al. Traffic synchronization in terminal airspace to enable continuous descent operations in trombone sequencing and merging procedures: an implementation study for Frankfurt Airport[J]. Transportation Research Part C: Emerging Technologies, 2020, 121: 102875.

[12] SÁEZ R, POLISHCHUK T, SCHMIDT C, et al. Automated sequencing and merging with dynamic aircraft arrival routes and speed management for continuous descent operations[J]. Transportation Research Part C: Emerging Technologies, 2021, 132: 103402.

[13] LIANG M, DELAHAYE D, MARÉCHAL P. Integrated sequencing and merging aircraft to parallel runways with automated conflict resolution and advanced avionics capabilities[J]. Transportation Research Part C: Emerging Technologies, 2017, 85: 268-291.

[14] 张军峰, 游录宝, 周铭, 等. 基于点融合系统的多目标进场排序与调度[J]. 北京航空航天大学学报, 2023, 49(1): 66-73.
ZHANG J F, YOU L B, ZHOU M, et al. Multi-objective arrival sequencing and scheduling based on point merge system[J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49(1):

66-73(in Chinese).

[15] 杨磊, 李文博, 刘芳子, 等. 柔性空域结构下连续下降航迹多目标优化[J]. 航空学报, 2021, 42(2): 324157.
YANG L, LI W B, LIU F Z, et al. Multi-objective optimization of continuous descending trajectories in flexible airspace[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(2): 324157(in Chinese).

[16] ICAO. Carbon emissions calculator methodology: Version 11[S]. Montréal: ICAO, 2014: 6-7.

[17] EUROCONTROL. User manual for the base of aircraft data (BADA): Revision 3.6[S]. Belgium: EUROCONTROL Experimental Centre, 2004: 6-33.

[18] ICAO. Manual on airspace planning methodology for the determination of separation minima: Doc 9689-AN/953[S]. Montréal: ICAO, 1998: 55-60.

Multi-aircraft low-carbon trajectories cooperative planning for point merge operation

WANG Chao*，LI Wenqing，SUN Shilei，XU Chenyang

(College of Air Traffic Management，Civil Aviation University of China，Tianjin 300300，China)

Abstract：In order to solve the technical problem of continuous descent operation (CDO) in the high traffic density terminal control area, a multi-aircraft low-carbon trajectory collaborative planning method for a point merge system (PMS) is proposed. Firstly, the mechanism of additional carbon emission based on the artificial decision-making of the air traffic controller was analyzed. Secondly, a multi-stage optimal control model of CDO was developed, taking into account the level flight arc segment of the PMS, using the low-carbon 4D trajectory of a single aircraft as the research object. Then, with the objective of minimum overall operation time, the 4D trajectory collaborative optimization model of multiple aircraft CDO was established with considering the arrival traffic flows' flight conflict constraint. Finally, the validation experiment was carried out with the real PMS and its air traffic track data. In addition to providing the aircraft landing sequence that supports efficient operation, the experimental findings demonstrate that this method reduces the overall flight duration by 12.2% and CO_2 emissions by 15.7%. It also provides plausible low-carbon 4D paths to accomplish this landing sequencing.

Keywords：air transportation; 4D trajectory; optimal control theory; continuous descent operation; point merge system; sequencing and scheduling

Received：2023-09-07；**Accepted**：2023-11-03；**Published Online**：2023-11-22 11：48

URL：link.cnki.net/urlid/11.2625.V.20231121.1154.003

Foundation item：Key Project of Applied Basic Research Multi-Investment Fund of Tianjin Municipal (21JCZDJC00780)

* **Corresponding author**. E-mail：wangch@cauc.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0540

基于 FY-3E 的星载 GNSS-R 海面高度反演模型

张云[1,2]，鹿麒[1,2]，张月维[3,*]，秦甘尧[1,2]，胡秀清[4]，杨光林[4]

(1. 上海海洋大学 信息学院，上海 201306；2. 上海市海洋智能信息与导航遥感工程技术研究中心，上海 201306；
3. 广州气象卫星地面站，广州 510640；4. 国家卫星气象中心，北京 100081)

摘 要：海面高度 (SSH) 在海洋学和气象学等领域中具有重要意义。针对目前缺乏利用中国自主研发风云三号 E 星 (FY-3E) 的全球导航卫星掩星探测仪-Ⅱ型 (GNOS-Ⅱ) 特有的不均匀分布时延-多普勒 (DDM) 数据实现海面测高研究的问题，采用 DDM 海面高度反演技术，使用丹麦 DTU18 模型和海潮模型验证反演精度，综合传统物理模型和机器学习模型，分别对星载北斗导航卫星系统 (BDS) 和全球定位系统 (GPS) 反射信号的海面高度反演性能进行评估。结果显示，由于 BDS B1Ⅰ信号的码片分辨率高于 GPS，使用 BDS 的物理模型进行全球海面高度反演的最大平均绝对误差 (MAE) 约为 3.0 m，明显优于 GPS 反演结果 (最大 MAE 约为 5.0 m)。通过随机森林 (RF) 和卷积神经网络 (CNN) 模型，GPS 和 BDS 均能实现较好的反演结果，最佳 MAE 均约为 0.4 m。与英国 TDS-1 数据的反演结果相比，GPS 反演精度在物理模型上提高约 15%，验证了 FY-3E GNOS-Ⅱ的全球导航卫星系统反射信号 (GNSS-R) 遥感数据的有效性。研究成果对于推广国产 FY-3E 的 GNSS-R 海面测高应用具有重要意义。

关 键 词：全球导航卫星系统反射信号；风云三号 E 星；海面高度反演；机器学习；特征选择
中图分类号：P715.6
文献标志码：A **文章编号**：1001-5965(2025)10-3262-11

海面高度 (sea surface height, SSH) 对于海洋学和气象学等领域至关重要[1]。全球导航卫星系统反射信号 (global navigation satellite system-reflectometry, GNSS-R) 技术是一种快速发展的遥感技术，可以利用从海洋表面反射的全球导航卫星系统 (global navigation satellite system, GNSS) 信号反演海面高度[2]。相对于机载和岸基接收平台[3-5]，星载 GNSS-R 在全球覆盖和时空多样性方面具有显著优势[6]。英国国家航天中心 (British National Space Center, BNSC) 灾害监测联盟 (Disaster Monitoring Coalition, DMC) 的 UK-DMC 卫星、萨里卫星技术有限公司 (Surrey Satellite Technology Ltd., SSTL) 的 TDS-1 卫星和美国国家航空航天局 (National Aeronautics and Space Administration, NASA) 的旋风全球导航卫星系统 (cyclone global navigation satellite system, CYGNSS) 提供的数据为研究星载 GNSS-R 海面高度反演提供了宝贵的机会。

2015 年，刘风玲等使用 UK-DMC 数据评估了星载 GNSS-R 在海洋探测中的可行性[7]。2016 年，Clarizia 等使用 TDS-1 数据进行了星载 GNSS-R 海面高度反演的初步研究，发现结果与丹麦技术大学 DTU10 全球平均海平面高度模型 (下文简称 DTU10 模型) 基本一致，均方根误差 (root mean square error, RMSE) 约为 8 m[8]。2018 年，Mashburn 等利用 TDS-1 数据，使用 DTU10 模型与电离层模型、对流层模型等建立 HiFi 模型作为验证，全球海面高度反演结果

收稿日期：2023-08-23；**录用日期**：2023-10-29；**网络出版时间**：2023-11-15 13:57
网络出版地址：link.cnki.net/urlid/11.2625.V.20231114.1744.002
基金项目：国家自然科学基金 (42176175,42271335)；国家重点研发计划 (2019YFD0900805)
* **通信作者**. E-mail：weixing1132@126.com

引用格式：张云，鹿麒，张月维，等. 基于 FY-3E 的星载 GNSS-R 海面高度反演模型 [J]. 北京航空航天大学学报，2025，51（10）：3262-3272. ZHANG Y，LU Q，ZHANG Y W，et al. Spaceborne GNSS-R sea surface height inversion model using FY-3E [J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51（10）：3262-3272（in Chinese）.

的残差为 6.4 m[9]。2021 年, 张云等使用 TDS-1 数据, 采用时延-多普勒 (delay-Doppler map, DDM) 海面高度反演技术, 同时优化了全球定位系统反射信号 (global position system reflectometry, GPS-R) 技术的误差修正模型, 相对于 DTU15 模型, 优化后反演模型的平均绝对误差 (mean absolute error, MAE) 为 6.05 m, 精度提高了约 29%[10]。

近年来, 机器学习算法因其强大的非线性拟合能力和易于构建的特性, 逐渐应用于 GNSS-R 海面测高领域。2021 年, Wang 等开发了一种新型的机器学习融合模型用于海面高度反演, 并将其应用于分析波罗的海的机载数据, 与 DTU15 模型相比, RMSE 约为 0.23 m, 决定系数 R^2 约为 0.75[11]。2022 年, 张云等分析了 TDS-1 数据, 提出了主成分分析-支持向量回归机 (principal component analysis combined with support vector regression, PCA-SVR) 和卷积神经网络 (convolutional neural network, CNN) 2 种星载海面高度反演模型, 与 DTU15 模型相比, 模型反演结果的平均 MAE 在 1 m 以下[12]。2022 年, 张云等根据星载 DDM 海面测高原理, 使用 CYGNSS 数据中的 DDM 等相关特征参数, 采用反向传播 (back propagation, BP) 神经网络和 CNN 模型对不同卫星号的数据进行海面高度反演, 与 DTU18 模型相比, 平均 MAE 约为 1 m[13]。

2021 年 7 月 5 日, 风云三号 E 星 (FengYun-3E, FY-3E) 在中国酒泉卫星发射中心成功发射。FY-3E 是中国自主研发的第二代极轨气象卫星[14], 搭载有全球导航卫星掩星探测仪-Ⅱ型 (GNSS occultation sounderⅡ, GNOS-Ⅱ), 可同时接收北斗导航卫星系统 (Beidou navigation satellite system, BDS) 和全球定位系统 (global positioning system, GPS) 的海面反射信号, 其主要任务是进行全球导航卫星掩星信号探测和海洋反射信号探测[15]。2022 年, 杨光林等介绍了 FY-3E 的 DDM 数据结果、海面风速反演模型和相关结果, 结果表明, 当海面风速小于 20 m/s 时, FY-3E 反演结果的精度小于 2 m/s[16]。

本文使用 FY-3E 获取的 L1 级 DDM 数据进行海面高度反演。基于星载 DDM 海面测高原理, 将传统物理模型和机器学习模型应用于 FY-3E 数据进行海面高度反演, 并与 TDS-1 数据的反演结果相对比, 验证 FY-3E 遥感数据的可靠性。同时, 本文评估了使用 GPS 和 BDS 不同卫星的物理模型和机器学习模型的海面测高性能, 并且评估了星载 BDS 反射信号海面测高性能。

1　数据来源

1.1　FY-3E GNOS-Ⅱ数据

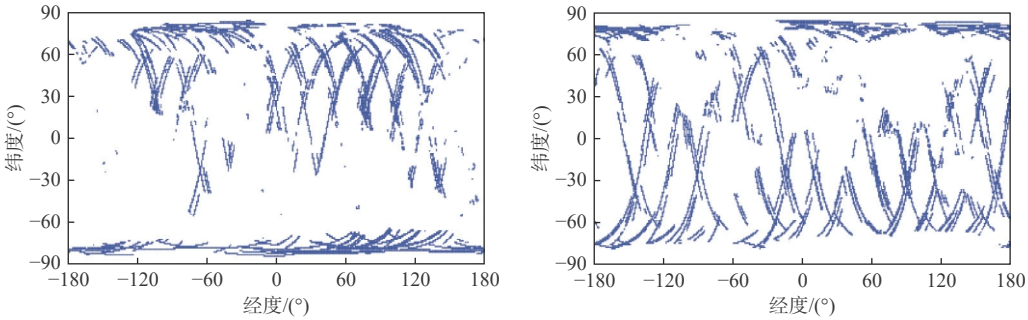
FY-3E 卫星是全球首颗民用晨昏轨道气象卫星, 其轨道高度为 836 km, 倾角为 98.5°。FY-3E 搭载的 GNOS-Ⅱ是风云三号 C 星和 D 星上 GNOS 的升级, 具备掩星探测和反射探测 2 种功能, 可以接收来自 BDS、GPS、GLONASS 和 Galileo 的信号。本文使用 GNOS-Ⅱ L1 GNSS 反射产品, 包括 DDM 波形数据、镜面反射点数据及接收机 (FY-3E) 和发射机 (GNSS) 的相关信息, 这些数据以 HDF5 格式存储。本文主要研究 BDS 和 GPS 的海面反射信号特性, 表 1 展示了 BDS B1Ⅰ信号与 GPS L1 信号的比较, BDS B1Ⅰ信号的伪码速率是 GPS L1 信号的 2 倍, 而 BDS B1Ⅰ信号的码片宽度是 GPS L1 信号的 1/2。

表 1　BDS B1Ⅰ信号与 GPS L1 信号的比较

Table 1　Comparison between BDS B1Ⅰ signals and GPS L1 signals

信号	频点/MHz	带宽/MHz	波长/m	调制方式	伪码速率/MHz	码片个数	码片宽度/μs
BDS B1Ⅰ	1 561.098	4.092	0.192	BPSK	2.046	2 046	0.489
GPS L1	1 575.42	2.046	0.191	BPSK	1.023	1 023	0.978

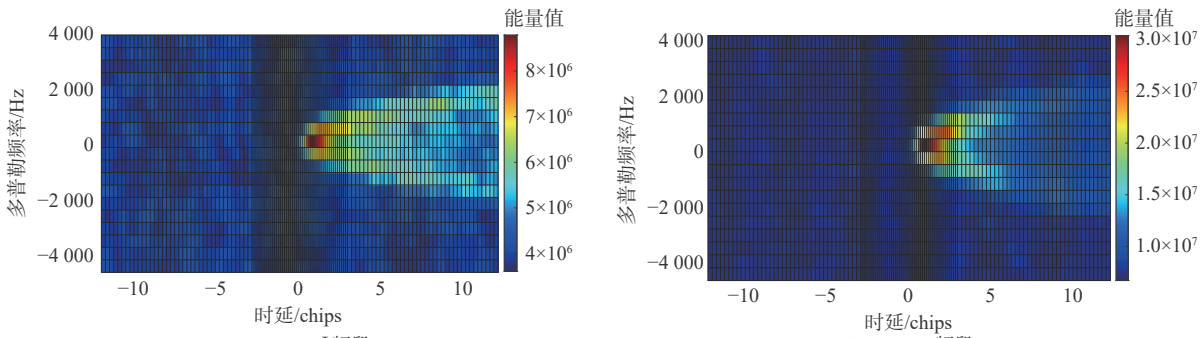
FY-3E GNOS-Ⅱ可以接收 8 通道反射信号, DDM 采样频率为 1 Hz, 每天可以接收约 9 万个来自陆地和约 12 万个来自海洋的反射信号, 图 1 展示了 2022 年 8 月 25 日的 FY-3E 镜面反射点分布。图 2 展示了 FY-3E L1 级反射产品中典型的海面 DDM 数据。其中, 时延维的范围为−12.25～12.125 chip (码片), 分为 122 个时延维 (lag); 多普勒频率维的范围为−5 000～4 500 Hz, 分为 20 个多普勒频率维, 分辨率为 500 Hz。与 CYGNSS 和 TDS-1 等的 DDM 不同, FY-3E 的 DDM 时延维分辨率呈不均匀分布。在镜面反射点附近的−2.875～2.875 chip 范围内, DDM 的时延分辨率为 1/8 chip (1 lag=1/8 chip); 而在−12.25～−2.875 chip 和 2.875～12.125 chip 范围内, DDM 的时延分辨率为 1/4 chip (1 lag=1/4 chip)。FY-3E 的 DDM 在镜面反射点附近提供了更高采样率的数据, 从而可以提高可观测值反演结果的准确性[16]。根据表 1, 由于 BDS B1Ⅰ信号的码片宽度是 GPS L1 信号的 1/2, BDS B1Ⅰ在镜面反射点附近的时延分辨率为 18.3 m, 而 GPS L1 的时延分辨率为 36.6 m。理论上, BDS B1Ⅰ的测高精度高于 GPS L1。



(a) 镜面反射点位于陆地的分布 (89 133个)　(b) 镜面反射点位于海洋的分布 (120 901个)

图 1　FY-3E 数据在 2022 年 8 月 25 日镜面反射点分布

Fig. 1　Distribution of specular points for FY-3E data on August 25th, 2022



(a) BDS B1I频段　(b) GPS L1频段

图 2　FY-3E 反射产品中典型的海面 DDM 图像

Fig. 2　Typical sea surface DDM images in FY-3E reflection products

1.2　DTU18 模型验证数据

由于缺乏全球范围的真实海面高度数据，当前的 GNSS-R 海面测高研究主要使用 DTU 全球平均海平面高度模型和 DTU 全球海洋潮汐模型作为海面高度的参考数据。DTU 全球平均海平面高度模型是由 Jason-1、ERS-1、ICESat 等多颗具有海面测高任务卫星的数据计算得来，空间分辨率为 1′×1′或 2′×2′。通过与英国和挪威的 GPS 水准潮位计进行广泛比较证实，在世界大多数地区，其垂直精度通常优于 10 cm[17]。DTU 全球海洋潮汐模型根据输入的儒略日 (Julian Day) 时间和经纬度生成具体的海潮高度信息，通常在±2.5 m 以内，空间分辨率为 0.125°×0.125°[18](这 2 个模型的数据可以从 http://www.space.dtu.dk 下载)。本文选取 DTU18 模型和 DTU 全球海洋潮汐模型代替真实海面高度作为海面高度验证模型[11]，其海面高度结果 H_{DTU} 为

$$H_{\mathrm{DTU}} = h_{\mathrm{MSS18}} + h_{\mathrm{Tide}} \tag{1}$$

式中：h_{MSS18}为 DTU18 模型结果；h_{Tide}为 DTU 全球海洋潮汐模型结果。

2　海面测高原理与模型

2.1　星载 GNSS-R 海面测高的几何原理

在地球表面为水平的理想情况下，可建立图 3 所示的星载 GNSS-R 测高几何路径延迟模型。根据几何光学原理，接收机卫星到镜面反射点的距离 h 求解如下：

$$\rho = \rho_{\mathrm{r}} + \rho_{\mathrm{i}} = \rho'_{\mathrm{r}} + \rho_{\mathrm{i}} \tag{2}$$

$$\rho'_{\mathrm{r}} + \rho_{\mathrm{i}} = 2h\cos\left(\frac{\pi}{2} - A\right) \tag{3}$$

$$h = \frac{\rho}{\sin A} \tag{4}$$

式中：ρ为反射信号相对于直射信号总的延迟路径；ρ_{r}为总延迟路径中的入射信号部分；ρ_{i}为总延迟路

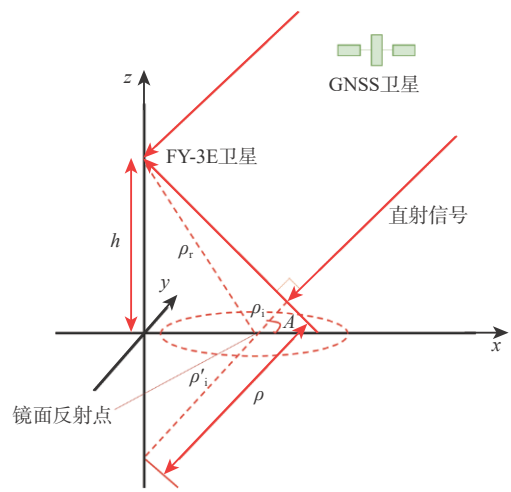


图 3　星载 GNSS-R 测高几何路径延迟模型

Fig. 3　Geometric path delay model for spaceborne GNSS-R altimetry

径中的直射信号部分; ρ_r'为ρ_r的镜像; A为反射信号的仰角。

基于 GNSS 的 L 波段电磁波的直线传播特性, 星载 GNSS-R 测高的几何原理如图 4 所示。选取 WGS84 参考椭球面, GNSS 卫星为信号发射机, FY-3E 为信号接收机, 可以计算海平面反射信号延迟和椭球面反射信号延迟的差值 D 为

$$D = c[(R_2 - D_{\mathrm{dir}}) - (R_1 - D_{\mathrm{dir}})] = c(R_2 - R_1) \quad (5)$$

式中: c 为光速; D_{dir} 为直射信号的延迟; R_1 为经过海平面反射信号的延迟; R_2 为经过椭球面反射信号的延迟。

在星载情况下, GNSS 卫星与 FY-3E 卫星距离地球足够远, 故海平面和椭球面的仰角差异较小, 均可视为A。综上, 可以通过确定延迟差 D 和反射信号的仰角A来计算海面高度 H[19]:

$$H = H_2 - H_1 = \frac{D}{2\sin A} \quad (6)$$

式中: H_1为 FY-3E 卫星到海平面的距离; H_2为 FY-3E 卫星到椭球面的距离。

将式 (6) 反演得到的 H 和式 (1) 的 DTU 海面高度验证模型结果 H_{DTU} 对比, 以进行反演精度的评估。

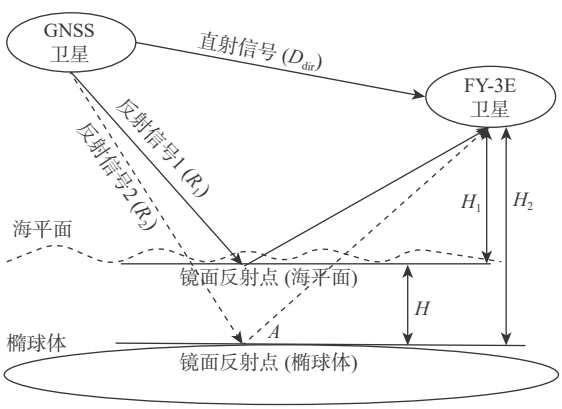


图 4　星载 GNSS-R 测高的几何原理

Fig. 4　Geometric schematic diagram of spaceborne GNSS-R altimetry

2.2　DDM 海面测高的物理模型

FY-3E 的 L1 级反射产品能够筛选出由海面上的镜面反射点生成的 DDM 信息, 典型的海上 DDM 散射功率分布如图 2 所示。DDM 的时延维中心点, 即开环跟踪点, 是理想情况下椭球面镜面反射点的跟踪点, 用于确定椭球面反射信号相对于直射信号在波形窗口内的延迟; 海面上的镜面反射点的重跟踪点通过 DER 方法计算, 即选择波形前沿的一阶导数的峰值位置作为重跟踪点, 用于确定海面反射信号相对于直射信号在波形窗口内的延迟[10]。通过计算开环跟踪点对应的延迟 D_{OLTP}与重跟踪点对应的延迟 D_{retrack}之间的差值, 可以得到延迟距离差 D:

$$D = c(D_{\mathrm{OLTP}} - D_{\mathrm{retrack}}) \quad (7)$$

鉴于 1/8 chip 的 DDM 时延分辨率仍然无法满足海面测高的精度要求, 进行了线性插值处理以获得更精确的时延 (插值步长为 20 lags)。图 5 为经过插值后的结果 (总 lags=2 421)。图 5(a) 展示了图 2(a) 中 BDS-DDM 数据的多普勒频率为 0 的时延一维功率谱切片; 图 5(b) 展示了图 2(b) 中 GPS-DDM 数据的多普勒频率为 0 的时延一维功率谱切片。通过式 (7) 可以计算出延迟距离差 D, 通过式 (6) 计算出海面高度 H。图 6 为基于上述物理模型的 DDM 海面测高流程。物理模型的结果可以评估 FY-3E 遥感数据的可靠性。此外, 通过物理模型可以提取与海面高度有关的关键特征, 用于构建机器学习模型的输入特征向量, 这有助于提高模型的性能和泛化能力。

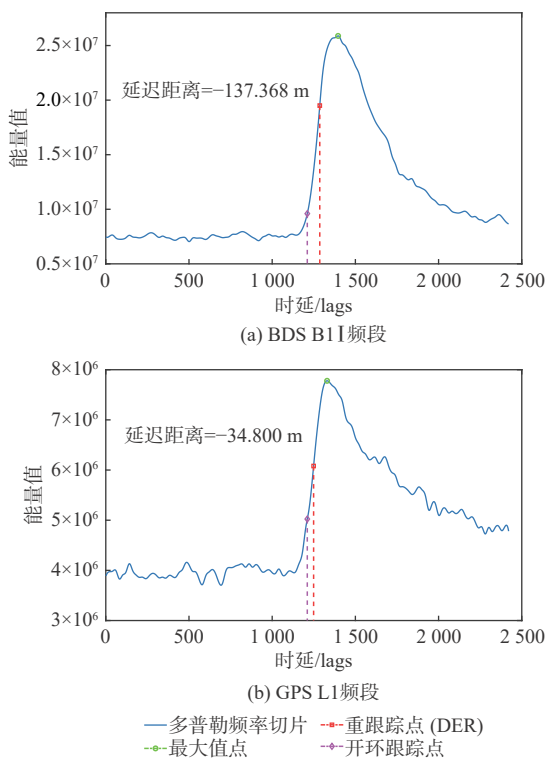


图 5　插值后的多普勒频率为 0 的时延一维功率谱切片

Fig. 5　One-dimensional power spectrum slice with zero Doppler shift after interpolation

2.3　机器学习模型

选取随机森林 (random forest, RF) 和 CNN 模型, 在发射机为 BDS 和 GPS 这 2 种情况下分别进行海面高度反演。RF 是一种经典的机器学习模型, 其利用多棵决策树进行分类并预测, 既可以用于分

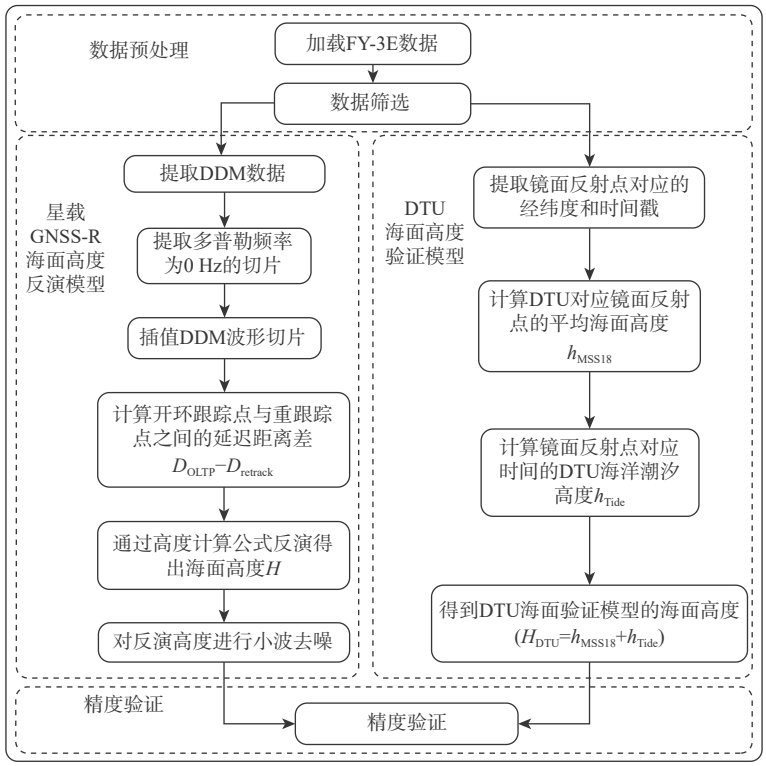


图6 物理模型反演流程

Fig. 6 Inversion proess of physical model

类问题，也可以用于回归问题。RF决策树的生成具有随机性，不仅在生成样本时采取有放回的随机抽取方式，而且在样本中的特征选择也具有随机性，这2种随机性使RF模型具有不易过拟合的优点，因此，在抵抗噪声方面表现相对较好。在回归问题中，森林中的每棵决策树都会对训练好的模型输入样本进行回归预测，从而各自产生一个预测结果。RF将所有决策树的预测结果取平均值作为最终的预测值。图7展示了机器学习数据处理流程及RF训练模型结构，其中，训练集和验证集的比例为8:2。DTU数据在机器学习模型训练中作为标签，此外，还将其作为真实数据进行反演精度评估。

在图7的特征参数选择过程中，物理模型是重要依据。FY-3E数据直接提供了镜面反射点和能量峰值的DDM坐标，作为计算反射信号路径延迟的关键参数，它们和镜面反射点处的入射角一起被选用为机器学习模型的输入参数。DDM的能量值会受到延迟和多普勒频率的影响，这是由接收机和发射机之间的高速相对运动、方向动态变化及卫星高度和间距较大等引起的。因此，接收机和发射机的速度也被选择为输入参数，值得注意的是，它们的位置参数可能导致模型学习到地理位置与海面高度之间的关系，从而相当于输出海面高度的插值信息，因此，位置信息没有被选为输入参数。此外，星载GNSS-R观测的时空跨度较大，GNSS信号更容易受到干扰，复杂的海况和海面粗糙度等因素都会影响反射信号的DDM波形，因此，也将信噪比和接收机天线增益作为输入参数。

2.4 数据处理

本文使用DDM海面测高物理模型反演了2022年8月7日、12日和17日的海面高度；使用2022年8月1日至6日作为训练数据，建立机器学习模型反演了8月7日、12日和17日的海面高度。为得到更好的结果，本文对数据集进行以下质量控制：

1）使用FY-3E数据质量控制标志筛选数据。

2）剔除位于陆地部分的镜面反射点观测值。

在物理模型中，DDM波形对海面高度反演结果具有重要影响，因此，物理模型中引入了以下筛选步骤：

1）剔除DDM峰值处信噪比过低的数据。

2）剔除DDM中有明显双峰的数据。

3）选取DDM切片峰值索引在60～71 lags之间的数据。

表2显示了经过数据质量控制后物理模型和机器学习模型中的实验数据量。由于物理模型引入了额外的筛选步骤，物理模型反演的数据量约为机器学习模型的1/2。

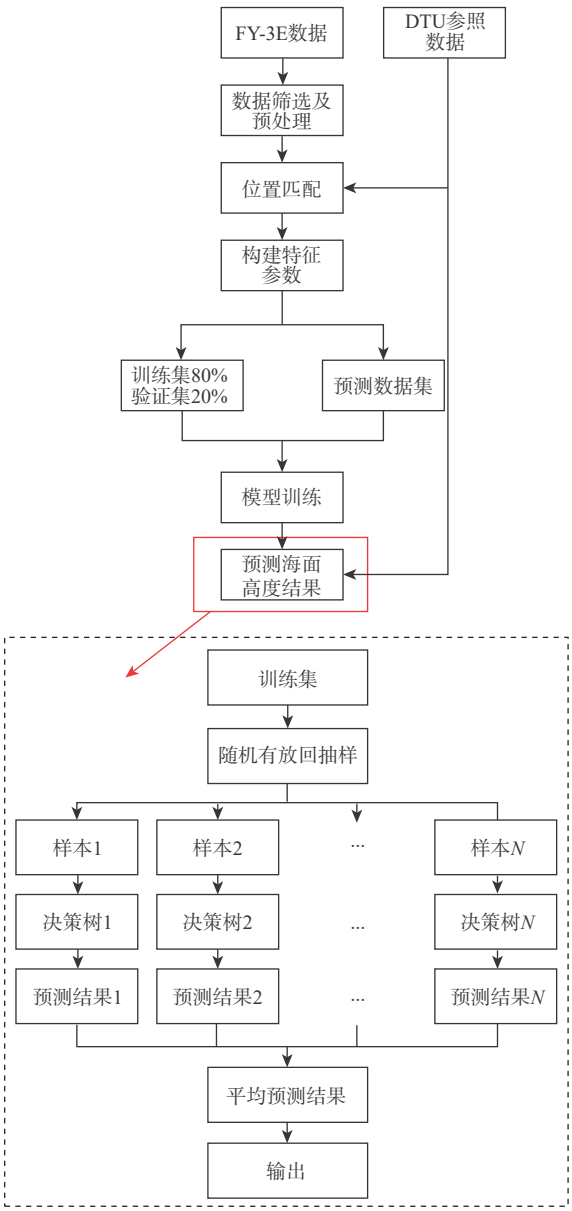


图 7　机器学习数据处理流程及 RF 训练模型结构

Fig. 7　Machine learning data processing process and random forest training model structure

表 2　数据质量控制后的数据数量

Table 2　Amount of data after data quality control

模型	BDS/GPS	训练集数据量	反演数据量		
		2022-08-01—2022-08-06	2022-08-07	2022-08-12	2022-08-17
机器学习模型	BDS	266 954	60 912	50 339	57 926
	GPS	260 100	54 432	52 695	45 151
物理模型	BDS		31 784	24 447	28 800
	GPS		34 135	31 823	30 368

3　反演结果

3.1　物理模型反演结果

根据 2.2 节中的物理模型，利用 FY-3E 的 L1 反射产品反演海面高度，采用 MAE 和 RMSE 评估精度，计算公式如下：

$$M=\frac{1}{m}\sum_{i=1}^{m}|(y_i-y_i')| \tag{8}$$

$$R=\sqrt{\frac{1}{m}\sum_{i=1}^{m}(y_i-y_i')^2} \tag{9}$$

式中：M 为 MAE；R 为 RMSE；m 为样本数量；y_i 为第 i 个样本的海面高度反演结果；y_i' 为相应的 DTU 海面高度验证模型的海面高度。

图 8 展示了单个文件“FY3E_GNOSR_ORBT_L1_20220812_2338_RFLC5_V0”中 BDS 反射信号的海面高度反演结果，MAE 为 2.816 m，RMSE 为 3.841 m。图 9 展示了单个文件“FY3E_GNOSR_ORBT_L1_20220812_2338_RFLC5_V0”中 GPS 反射信号的海面高度反演结果，MAE 为 5.215 m，RMSE 为 5.894 m。

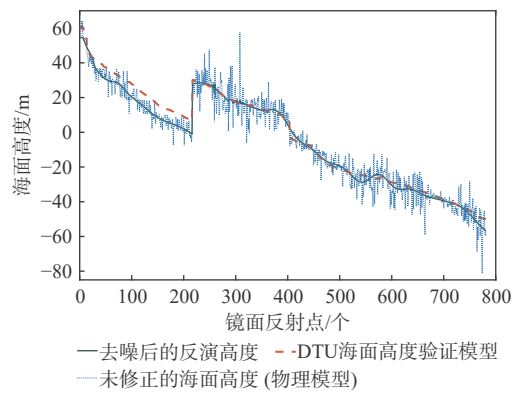


图 8　单个文件海面高度反演结果(BDS)

Fig. 8　Single file sea surface height inversion results (BDS)

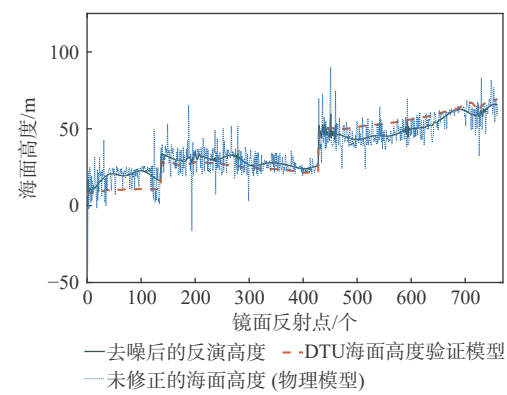


图 9　单个文件海面高度反演结果(GPS)

Fig. 9　Single file sea surface height inversion results (GPS)

从图 8 和图 9 可以看到，物理模型反演结果的趋势大致与 DTU 模型结果一致，但物理模型反演结果表现出噪声特性，对反演精度产生了影响。通过对反演得到的海面高度数据进行小波变换和去噪，明显提高了反演精度。小波去噪是一种基于小波分析的信号去噪方法，通过分解信号的小波系

数，对高频噪声进行滤波处理，再通过反向小波变换重构信号，其数学表达式为

$$x_{\text{denoised}}(t) = W^{-1}\{\text{Threshold}(W\{x(t)\})\} \quad (10)$$

式中：$W()$表示小波变换；$W^{-1}()$ 表示反向小波变换；Threshold() 表示阈值函数，用于对小波系数进行滤波处理；$x(t)$为原始信号；x_{denoised}为去噪后的信号。

本文使用 Db10 小波作为小波基，并根据通用阈值准则进行软阈值处理。Daubechies 小波族是小波分析中广泛采用的一组小波基函数，常用于对信号进行多尺度分解和重构以起到滤波作用。Db10 小波的平滑性介于高度平滑的小波 (如 Db4) 和更细节的小波 (如 Db20) 之间，在本文中的效果较好。由于通过物理模型反演得到的海面高度信号中存在较多小幅噪声，而通用阈值函数 sqtwolog 能有效抑制较小的噪声并保留信号的细节信息，本文采用该函数确定每个小波系数的阈值。为了在保留主要特征的同时保留足够的细节信息，将信号经过小波变换分解成 5 个层次，并使用基于一级系数的电平噪声的单一估计来重新缩放信号。从图 8 和图 9 可以看到，经过小波去噪后，海面高度曲线明显比去噪前的结果更加平滑。

图 10 显示了 2022 年 8 月 7 日、12 日和 17 日在发射机为 BDS 时的物理模型反演误差 (残差) 的地理分布。8 月 7 日海面高度的 MAE 为 2.635 m, RMSE 为 3.702 m; 8 月 12 日海面高度的 MAE 为 2.812 m, RMSE 为 3.939 m; 8 月 17 日海面高度的 MAE 为 3.501 m, RMSE 为 4.631 m。平均 MAE 为 2.983 m, 平均 RMSE 为 4.091 m。

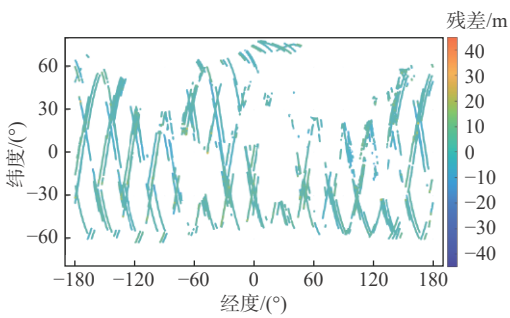


图 10 发射机为 BDS 时的误差 (残差) 地理分布(物理模型，2022 年 8 月 7 日、12 日和 17 日)

Fig. 10 Geographical distribution of errors (residual) when transmitter is BDS (physical model, 2022-08-07, 2022-08-12, 2022-08-17)

图 11 显示了和图 10 相同时间内发射机为 GPS 时的物理模型反演误差 (残差) 的地理分布。8 月 7 日海面高度的 MAE 为 5.539 m, RMSE 为 6.801 m; 8 月 12 日海面高度的 MAE 为 5.260 m, RMSE 为 6.468 m; 8 月 17 日海面高度的 MAE 为 4.702 m, RMSE 为 5.867 m。平均 MAE 为 5.167 m, 平均 RMSE 为 6.379 m。

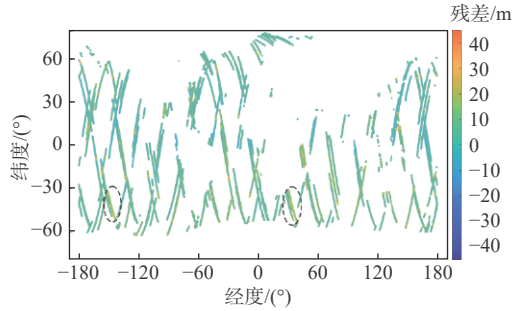


图 11 发射机为 GPS 时的误差 (残差) 地理分布(物理模型，2022 年 8 月 7 日、12 日和 17 日)

Fig. 11 Geographical distribution of errors (residual) when transmitter is GPS (physical model, 2022-08-07, 2022-08-12, 2022-08-17)

表 3 汇总了 2 种发射机情况下反演结果的精度，反演结果证明了基于 FY-3E 的 BDS 或 GPS 的 DDM 信号的海面高度反演的有效性。可以看出，BDS 的反演精度优于 GPS，符合 BDS B1Ⅰ信号码片分辨率更优的特点。在相近地理位置，GPS 反演结果经常存在一个轨道的大部分样本点精度低于 BDS 反演结果的情形 (如图 11 椭圆处所示)。此外，如表 2 所示，在物理模型的质量控制中，剔除了更多 BDS 信号源中的低质量数据，而 GPS 则相对较少。这些差异可能与两者信号的调制参数有关。BDS 的反演 MAE 大于 2 m, GPS 大于 4 m, 与 TDS-1 数据反演结果相比[10]，GPS 总反演精度在物理模型上提高了约 15%。虽然 FY-3E 物理模型的反演精度相对较好，但进一步的误差估计有望继续提高反演精度。

表 3 物理模型反演结果精度

Table 3 Accuracy of physical model inversion results

反演日期	MAE/m		RMSE/m	
	BDS	GPS	BDS	GPS
2022-08-07	2.635	5.539	3.702	6.801
2022-08-12	2.812	5.260	3.939	6.468
2022-08-17	3.501	4.702	4.631	5.867

3.2 RF 模型反演结果

根据 2.3 节，使用 FY-3E 的 L1 级反射产品 (2022 年 8 月 1 日至 6 日) 和 DTU 海面高度验证模型进行 RF 模型训练，实现海面高度外推反演。所有结果都通过 2 次模型训练结果取平均得到。图 12 显示了 2022 年 8 月 7 日、12 日和 17 日在发射机为 BDS 时的 RF 模型误差 (残差) 的地理分布。8 月 7 日海面高度的 MAE 为 0.439 m, RMSE 为 0.663 m; 8 月 12 日海面高度的 MAE 为 0.531 m, RMSE 为

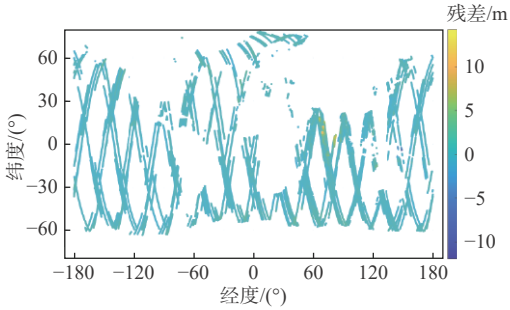


图 12　发射机为 BDS 时的误差 (残差) 地理分布
（RF 模型, 2022 年 8 月 7 日、12 日和 17 日）

Fig. 12　Geographical distribution of errors (residual) when transmitter is BDS (RF model, 2022-08-07, 2022-08-12, 2022-08-17)

0.743 m; 8 月 17 日海面高度的 MAE 为 0.481 m, RMSE 为 0.813 m。总 MAE 为 0.484 m, 总 RMSE 为 0.740 m。

图 13 显示了和图 12 相同时间内发射机为 GPS 时的 RF 模型误差 (残差) 的地理分布。8 月 7 日海面高度的 MAE 为 0.471 m, RMSE 为 0.816 m; 8 月 12 日海面高度的 MAE 为 0.597 m, RMSE 为 0.989 m; 8 月 17 日海面高度的 MAE 为 0.467 m, RMSE 为 0.908 m。总 MAE 为 0.512 m, 总 RMSE 为 0.904 m。

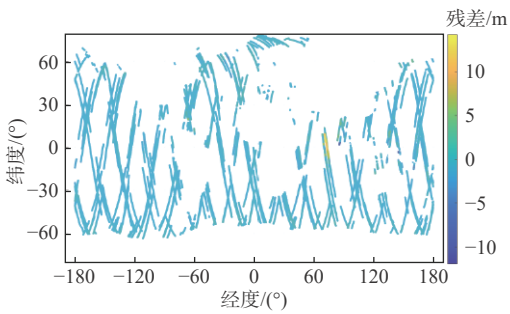


图 13　发射机为 GPS 时的误差 (残差) 地理分布
（RF 模型, 2022 年 8 月 7 日、12 日和 17 日）

Fig. 13　Geographical distribution of errors (residual) when transmitter is GPS (RF model, 2022-08-07, 2022-08-12, 2022-08-17)

表 4 汇总了 2 种发射机情况下反演结果的精度。R^2 均大于 0.99, 说明 BDS 和 GPS 的 RF 预测结果与 DTU 数据相吻合。虽然总体 MAE 达到了约 0.5 m, 但每次训练的精度差值可能达到 10 cm, 说明 RF 模型在外推反演中存在一定的不确定性。

3.3　CNN 模型反演结果

为进一步验证 FY-3E 数据在不同机器学习模型中的反演表现, 进一步使用 CNN 模型对海面高度进行反演。CNN 是一种经典的前馈神经网络, 其最大的优点在于可以使用卷积核对大规模数据进行压缩提炼, 从而在海量数据中挖掘出不同特征参

表 4　RF 模型反演结果的精度

Table 4　Accuracy of RF model inversion results

预测日期	MAE/m		RMSE/m	
	BDS	GPS	BDS	GPS
2022-08-07	0.439	0.471	0.663	0.816
2022-08-12	0.531	0.597	0.743	0.989
2022-08-17	0.481	0.467	0.813	0.908

数之间的非线性关系。在海面高度反演领域, CNN 在 TDS-1[12] 和 CYGNSS[13] 数据上都有相对较好的表现。图 14 为本文使用的 CNN 模型结构。

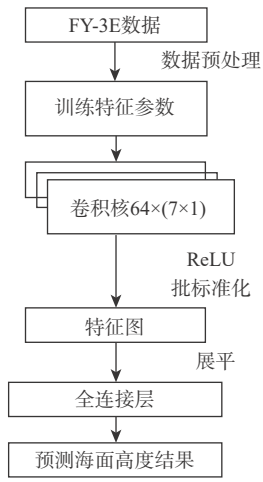


图 14　CNN 模型结构

Fig. 14　Structure of CNN model

与 RF 模型相同, 使用 2022 年 8 月 1 日至 6 日的数据训练 CNN 模型, 反演 2022 年 8 月 7 日、12 日和 17 日的海面高度。图 15 显示了发射机为 BDS 时的 CNN 模型误差 (残差) 的地理分布。8 月 7 日海面高度的 MAE 为 0.429 m, RMSE 为 0.573 m; 8 月 12 日海面高度的 MAE 为 0.578 m, RMSE 为 0.745 m; 8 月 17 日海面高度的 MAE 为 0.496 m, RMSE 为 0.675 m。总 MAE 为 0.501 m, 总 RMSE 为 0.664 m。

图 16 显示了发射机为 GPS 时的 CNN 模型误

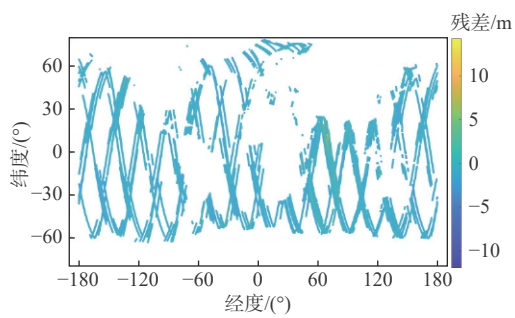


图 15　发射机为 BDS 时的误差 (残差) 地理分布
（CNN 模型, 2022 年 8 月 7 日、12 日和 17 日）

Fig. 15　Geographical distribution of errors (residual) when transmitter is BDS (CNN model, 2022-08-07, 2022-08-12, 2022-08-17)

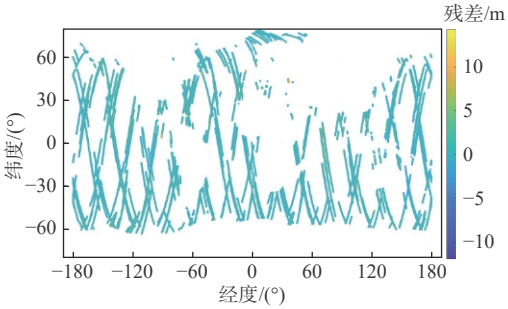


图16 发射机为GPS时的误差(残差)地理分布(CNN模型, 2022年8月7日、12日和17日)

Fig. 16 Geographical distribution of errors (residual) when transmitter is GPS (CNN model, 2022-08-07, 2022-08-12, 2022-08-17)

差(残差)的地理分布。8月7日海面高度的MAE为0.495 m, RMSE为0.652 m; 8月12日海面高度的MAE为0.593 m, RMSE为0.766 m; 8月17日海面高度的MAE为0.435 m, RMSE为0.584 m。总MAE为0.508 m, 总RMSE为0.667 m。

表5总结了2种发射机情况下反演结果的精度, 且R^2均大于0.99。结果显示, CNN模型的表现与RF模型相似, 均优于物理模型。与物理模型相比, 机器学习模型能够从大量的训练集数据中学习各特征参数之间的关系, 可以捕捉到物理模型无法精确建模的复杂非线性关系。此外, 机器学习模型可以通过训练来适应数据中普遍存在的噪声, 表现出更强的鲁棒性。但机器学习模型的性能通常受到数据预处理和特征工程的影响, 而基于物理模型的特征参数选择可以使机器学习模型更好地适应任务和数据集。通常情况下, CNN等神经网络模型需要大量的数据支持, 但本文使用的FY-3E数据与CYGNSS[13]相比, 6天窗口的样本规模存在百万级差距, 对CNN模型的反演精度可能有一定的影响。在今后的研究中, 通过增加更多的训练天数, 有望获得精度更高、鲁棒性更强的泛化结果。此外, 对于印度洋北部的一部分地区, 物理模型和机器学习模型的误差都相对较大。该现象可能与DTU模型所示的北印度洋平均海面高度偏低及气旋的季节扰动有关, 难以避免由此产生的反演异常值, 需要进一步研究。

表5 CNN模型反演结果精度

Table 5 Accuracy of CNN model prediction results

预测日期	MAE/m		RMSE/m	
	BDS	GPS	BDS	GPS
2022-08-07	0.429	0.495	0.573	0.652
2022-08-12	0.578	0.593	0.745	0.766
2022-08-17	0.496	0.435	0.675	0.584

4 结 论

本文将中国自主研发的FY-3E GNOS-Ⅱ载荷提供的GNSS-R L1级非均匀DDM数据用于海面高度反演。使用DDM海面测高传统物理模型和机器学习模型, 分别以BDS和GPS为发射机信号源, 成功地使用FY-3E数据进行了海面高度反演, 并将反演结果与DTU18模型和DTU全球海洋潮汐模型进行对比得到反演精度。

1) 利用星载GNSS-R测高的几何原理, 开发了基于FY-3E L1级DDM数据的DER重跟踪模型, 建立了DDM海面高度反演的物理模型, 并利用2022年8月7日、12日和17日的FY-3E数据实现了BDS和GPS的海面高度反演。BDS (MAE为2～3 m)的反演结果明显优于GPS反演结果(MAE为4～5 m), 这与BDS B1Ⅰ信号的码片分辨率更高有关。与TDS-1的物理模型反演精度(约6.05 m)相比较, FY-3E物理模型的反演精度略好, 也证明了FY-3E的DDM数据在海面高度反演上的有效性。

2) 根据物理模型, 选择14个与海面高度相关的特征参数输入, 设计了RF海面高度反演模型。利用2022年8月1日至6日这6天的数据训练模型, 反演了8月7日、12日和17日的海面高度。结果表明, GPS或BDS的反演精度均约为0.5 m, 明显优于物理模型, 可满足海洋渔业等领域的业务需求。

3) 使用与RF模型相同的训练数据构建了CNN海面高度反演模型。与RF模型相比, 总MAE基本相同(仅相差几厘米), 而总RMSE有了一定的改善(BDS和GPS分别改善了约10%和26%), CNN模型有一定的优势。

4) 机器学习模型在3天(2022年8月7日、12日和17日)的反演精度无明显差异, 表明本文的2个机器学习模型(使用6天训练数据)在短时间范围内的反演中具有良好的鲁棒性和泛化能力。

在星载GNSS-R海面测高物理模型中, 不可避免地存在系统硬件、外部环境及几何模型计算导致的误差, 为了实现更高精度定位测高, 在未来的研究中需要对各类误差(电离层误差、轨道误差等)进行分析并建立相应的误差模型, 实现误差纠正; 在机器学习模型中, 随着FY-3E星载GNSS-R数据量的持续增加, 能够利用更多、覆盖更长时间范围的数据来研究海面高度反演模型, 从而进一步提高FY-3E海面测高的性能。

致 谢 感谢丹麦技术大学提供了DTU18全球平均海平面高度模型和DTU全球海洋潮汐模型。

感谢上海航天电子技术研究所的周勃高级工程师、秦瑾博士和孟婉婷对反射信号方面提出的建议。

参考文献（References）

[1] 陈锐志, 王磊, 李德仁, 等. 导航与遥感技术融合综述[J]. 测绘学报, 2019, 48(12): 1507-1522.
CHEN R Z, WANG L, LI D R, et al. A survey on the fusion of the navigation and the remote sensing techniques[J]. Acta Geodaetica et Cartographica Sinica, 2019, 48(12): 1507-1522(in Chinese).

[2] MARTIN-NEIRA M. A passive reflectometry and interferometry system (PARIS): application to ocean altimetry[J]. ESA Journal, 1993, 17(4): 331-335.

[3] 张云, 张杨阳, 孟婉婷, 等. 机载 GNSS 反射信号海面测高模型的研究[J]. 海洋学报, 2020, 42(3): 149-156.
ZHANG Y, ZHANG Y Y, MENG W T, et al. Research on sea surface altimetry model of airborne GNSS reflected signal[J]. Haiyang Xuebao, 2020, 42(3): 149-156(in Chinese).

[4] ZHANG Y, TIAN L M, MENG W T, et al. Feasibility of code-level altimetry using coastal BeiDou reflection (BeiDou-R) setups[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2015, 8(8): 4130-4140.

[5] ZHANG Y, LI B B, TIAN L M, et al. Phase altimetry using reflected signals from BeiDou GEO satellites[J]. IEEE Geoscience and Remote Sensing Letters, 2016, 13(10): 1410-1414.

[6] 李黄, 夏青, 尹聪, 等. 我国 GNSS-R 遥感技术的研究现状与未来发展趋势[J]. 雷达学报, 2013, 2(4): 389-399.
LI H, XIA Q, YIN C, et al. The current status of research on GNSS-R remote sensing technology in China and future development[J]. Journal of Radars, 2013, 2(4): 389-399(in Chinese).

[7] 刘风玲, 张云, 孟婉婷, 等. UK-DMC 卫星接收机 GNSS 反射信号的应用分析[J]. 遥感信息, 2015, 30(1): 90-95.
LIU F L, ZHANG Y, MENG W T, et al. Application analysis of airborne GNSS reflected data from UK-DMC satellite receiver[J]. Remote Sensing Information, 2015, 30(1): 90-95(in Chinese).

[8] CLARIZIA M P, RUF C, CIPOLLINI P, et al. First spaceborne observation of sea surface height using GPS-reflectometry[J]. Geophysical Research Letters, 2016, 43(2): 767-774.

[9] MASHBURN J, AXELRAD P, LOWE S T, et al. Global ocean altimetry with GNSS reflections from TechDemoSat-1[J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(7): 4088-4097.

[10] 张云, 马德皓, 孟婉婷, 等. 基于 TechDemoSat-1 卫星的 GPS 反射信号海面高度反演[J]. 北京航空航天大学学报, 2021, 47(10): 1941-1948.
ZHANG Y, MA D H, MENG W T, et al. Sea surface height inversion of GPS reflected signal based on TechDemoSat-1 satellite[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(10): 1941-1948(in Chinese).

[11] WANG Q, ZHENG W, WU F, et al. A new GNSS-R altimetry algorithm based on machine learning fusion model and feature optimization to improve the precision of sea surface height retrieval[J]. Frontiers in Earth Science, 2021, 9: 730565.

[12] ZHANG Y, HUANG S, HAN Y L, et al. Machine learning methods for spaceborne GNSS-R sea surface height measurement from TDS-1[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2022, 15: 1079-1088.

[13] ZHANG Y, LU Q, JIN Q, et al. Global sea surface height measurement from CYGNSS based on machine learning[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2022, 16: 841-852.

[14] ZHANG P, HU X Q, LU Q F, et al. FY-3E: the first operational meteorological satellite mission in an early morning orbit[J]. Advances in Atmospheric Sciences, 2022, 39(1): 1-8.

[15] SUN Y Q, WANG X Y, DU Q F, et al. The status and progress of Fengyun-3E GNOS Ⅱ mission for GNSS remote sensing[C]//Proceedings of the IEEE International Geoscience and Remote Sensing Symposium. Piscataway: IEEE Press, 2019: 5181-5184.

[16] YANG G L, BAI W H, WANG J S, et al. FY3E GNOS Ⅱ GNSS reflectometry: mission review and first results[J]. Remote Sensing, 2022, 14(4): 988.

[17] ANDERSEN O B, KNUDSEN P. The DTU2010MSS mean sea surface in the arctic-for and with Cyrosat-2 data[J]. Cryosat Validation Workshop, 2011, 693: 28.

[18] PICCIONI G, ANDERSEN O B, STENSENG L. SAR altimetry for mean sea surface determination in the arctic DTU15MSS[C]//Proceedings of the Sentinel-3 for Science Workshop. Venice : European Space Agency, 2015: 75.

[19] GARRISON J L, KATZBERG S J. The application of reflected GPS signals to ocean remote sensing[J]. Remote Sensing of Environment, 2000, 73(2): 175-187.

Spaceborne GNSS-R sea surface height inversion model using FY-3E

ZHANG Yun[1, 2], LU Qi[1, 2], ZHANG Yuewei[3, *], QIN Ganyao[1, 2], HU Xiuqing[4], YANG Guanglin[4]

(1. College of Information Technology, Shanghai Ocean University, Shanghai 201306, China;
2. Shanghai Engineering Research Center of Marine Intelligent Information and Navigation Remote Sensing, Shanghai 201306, China;
3. Guangzhou Meteorological Satellite Ground Station, Guangzhou 510640, China;
4. National Satellite Meteorological Centre, Beijing 100081, China)

Abstract: Sea surface height (SSH) is of great significance in fields such as oceanography and meteorology. Given the current absence of research utilizing the unevenly distributed delay-Doppler map (DDM) data from China's independently developed FengYun-3E (FY-3E) global navigation satellite system occultation sounder Ⅱ (GNOS-Ⅱ), this paper employs DDM-based SSH retrieval technology and validates the retrieval accuracy using the Danish DTU18 global mean SSH model and tidal model. The SSH retrieval performance of reflected signals from the global positioning system (GPS) and the spaceborne Beidou navigation satellite system (BDS) is evaluated individually in this research using both conventional physical models and machine learning techniques. The results demonstrate that due to the higher chip resolution of BDS B1Ⅰ signals compared to GPS, the maximum mean absolute error (MAE) of the global SSH inversion using the physical model of BDS is about 3.0 m, which is significantly better than that of the GPS (maximum MAE is about 5.0 m). Both GPS and BDS may produce good inversion results using the random forest (RF) and convolutional neural network (CNN) models; their best MAE is approximately 0.4 m. Compared with the inversion results of the UK TDS-1 data, the GPS inversion accuracy is improved by about 15% in the physical model, which verifies the validity of the global navigation satellite system-reflectometry (GNSS-R) remote sensing data of the FY-3E GNOS-Ⅱ. The research findings in this paper are of great significance for promoting the application of domestic FY-3E GNSS-R SSH measurements.

Keywords: global navigation satellite system-reflectometry; FengYun-3E; sea surface height inversion; machine learning; feature selection

Received: 2023-08-23; **Accepted**: 2023-10-29; **Published Online**: 2023-11-15 13:57

URL: link.cnki.net/urlid/11.2625.V.20231114.1744.002

Foundation items: National Natural Science Foundation of China (42176175,42271335); National Key Research and Development Program of China (2019YFD0900805)

* **Corresponding author**. E-mail: weixing1132@126.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0544

基于人工势场的虚拟编组自适应模型预测控制

林俊亭[1,*]，倪铭君[1,2]

(1. 兰州交通大学 自动化与电气工程学院，兰州 730070；2. 新疆交通职业技术大学，乌鲁木齐 831400)

摘 要：现今，列车高速度、高密度追踪控制对编队列车运行的安全性提出更高的要求。为满足人们对列车运行过程中自适应性和准确性的需求，提出一种基于人工势场的虚拟编组(VC)自适应模型预测控制(MPC)方法。将VC列车作为研究对象，采用MPC方法建立基于列车平衡态的动力学模型，以控制精度和平稳性、安全性为优化目标，并将基于人工势场设置的防撞函数加入目标函数，从而实现编队的防撞控制；分析不同时域参数对系统控制精度和计算效率的影响作用，设计对应的适应度函数，基于遗传算法(GA)求得不同工况下的最优时域参数组合，并制定时域参数更新策略，在确保列车编组准确控制的同时提高系统的实时性；在MATLAB平台上搭建4列车追踪运行场景，仿真验证所提方法的有效性。结果表明：相较于传统的模型预测控制器，基于人工势场的模型预测控制器在间隔控制上准确度提高了94.8%，可有效避免列车间发生碰撞，保证了列车运行的安全性；另外，采用自适应控制律的控制器可根据列车运行状态对系统进行实时调整，在确保高控制精度的前提下，计算效率提高10%。研究结果验证了所提方法的可行性，提高了控制器的综合控制性能，并为进一步优化编队控制和保障列车安全运行提供参考。

关 键 词：虚拟编组；模型预测控制；人工势场；遗传算法；列车追踪运行优化

中图分类号：U284.4

文献标志码：A **文章编号**：1001-5965(2025)10-3273-13

虚拟编组(virtual coupling, VC)作为一种新型的不需要物理连接的编组方式，编队内列车借助无线通信交互信息实现协同控制，在增加编组灵活性的同时，可采用更小的理想追踪间隔运行，从而提高铁路运输效率[1]。与此同时，VC对列车间隔保持提出了高要求，在列车实际追踪过程中，需避免与前车间隔为负值发生碰撞的情况。因此，为提高协同控制的安全性，实现编队列车的防撞控制成为学者研究的热点方向[2]。

在高速铁路高密度、小间隔运营条件下，引进列车群编队运行控制思想，使编队内部列车追踪运行更紧密，提高了线路的空间利用率。刘俊囡[3]研究编队列车群协同优化方案，分别采用多目标文化差分算法和人群搜索算法进行求解，实现对编队列车发车间隔时间和速度控制策略曲线协同优化。Bai等[4]基于状态信息传输建立拓扑结构，构建多列车的分布式协同控制律，确保编队内列车能够精准追踪期望曲线。

然而，在列车群协同控制中，因系统抗扰性较弱，易受多变的线路运行环境影响，导致追踪控制的准确性降低。而模型预测控制(model predictive control, MPC)方法因建模简单、控制灵活等优点得到了广泛应用，并且具有在线优化的特点，适用于高速运行的列车，有助于降低控制的偏差度。汤旻安等[5]针对列车追踪精度不高的问题，提出了基于混合系统神经网络反馈补偿控制的MPC方法，结

收稿日期：2023-08-25；**录用日期**：2023-10-17；**网络出版时间**：2023-11-20 11:03
网络出版地址：link.cnki.net/urlid/11.2625.V.20231118.1404.001
基金项目：国家自然科学基金(52162050)
*通信作者. E-mail：linjt@lzjtu.edu.cn

引用格式：林俊亭，倪铭君. 基于人工势场的虚拟编组自适应模型预测控制[J]. 北京航空航天大学学报，2025，51(10)：3273-3285.
LIN J T，NI M J. Adaptive model predictive control of virtual coupled based on artificial potential field[J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51(10)：3273-3285 (in Chinese).

果表明，MPC方法具有良好的跟踪能力，可较大程度地减小误差。杨辉等[6]为了弥补建模中高度非线性和不确定性带来的误差，提出基于模型补偿的高速列车状态反馈预测控制方法，提高了系统的控制精度。文献[5-6]均基于移动闭塞的列车群追踪控制。由于VC属于目前前沿的研究课题，将MPC方法应用到VC的研究相对较少。本文在文献[7]的基础上进行改进完善，文献[7]采用MPC方法对虚拟编组列车组(virtually coupled train set, VCTS)进行准确追踪控制，但与文献[5-6]相同，均未考虑追踪间隔为负值的情况，即存在相邻列车发生碰撞的不安全因素，因此，在编队列车追踪的安全性上具有一定的提升空间。

人工势场(artificial potential field, APF)计算简单且实时性好，便于实时避障和轨迹控制，并且在汽车、船舶等行业具有一定的应用基础，可将其引入列车编队控制中实现防撞控制。周红梅[8]在障碍物表面构造β-型虚拟智能体，结合人工势场法，利用弹性距离的思想，实现微型无人机避障控制。朱梁鹏[9]选取中间介质点作为跟踪虚拟参考点，在人工势场的作用下，构造避碰和避障斥力势场函数，从而实现编队控制。吴宝库[10]对人工势场法进行改进，针对目标不可达和局部极小值的问题，分别对斥力场进行优化并引入模拟退火算法，使多自动引导车编队可协同控制运行。钟嘉灏[11]针对无人机动态航迹规划中突发的局部威胁问题，设计了基于动态窗避障策略下的改进人工势场法，实现无人机局部动态航迹规划。熊俊涛等[12]针对人工势场法范围斥力影响最短路径规划的问题，提出了一种方向惩罚避障函数设置方法，进一步优化了规划路径长度，提高采摘效率。而在VCTS中，可根据列车间的通信关系及误差对编队的影响作用，设计基于人工势场的防撞函数，由于本文避障针对动态的邻间列车，在实现实时追踪控制的同时，在动态避碰方面也得到了改进。

传统的MPC方法多采用固定时域的控制方法，而时域参数对系统的控制精度和计算效率均有影响。因此，从时域参数优化的角度出发，可采用变时域参数来改善控制效果。丁伟兵[13]分析了2种工况下时域参数对控制器性能的影响，采用遗传算法(genetic algorithm, GA)求得最优时域参数，结果表明，变时域参数的MPC方法在确保控制精度的同时提高了计算的实时性。张志伟[14]采用GA综合考虑各因素设计适应度函数，针对不同的速度和路径，确定参数的适配方案，相较于固定时域参数控制器，运算时间明显缩短。白国星等[15]通过分析预测时域、速度与控制性能之间的关系，实现动态变换预测时域，采用三次多项式拟合获得了最佳预测时域的控制律，进一步提高了非线性MPC系统的实用性。杨阳阳[16]采用GA对时域参数进行优化设计，确定各行驶速度及道路工况下的最优时域参数组合，提高了控制器对系统的可控性与可观性。

本文以高速列车为研究对象，提出了一种基于人工势场的虚拟编组自适应模型预测控制方法。首先，根据VC列车的运行方式搭建列车动力学模型，为避免列车间发生碰撞，采用人工势场法建立基于速度和间隔的防撞函数，以控制精度、平稳性和安全性为目标设计目标函数，实现列车安全稳定运行。其次，采用GA对时域参数进行优化设计，使系统在不同工况下选取最优时域参数组合，通过变时域的控制策略进一步优化MPC方法的控制效果。仿真结果表明，本文方法能够有效避撞，并通过时域参数的自适应调节提高了系统的综合性能。

1 基于人工势场的MPC防撞控制

1.1 列车动力学模型

在建立纵向动力学模型时，本文将列车视为单质点[17]进行研究，忽略坡道、弯曲阻力等阻力的作用，假设列车在平直轨道上行驶，在运行过程中进行受力分析，其动力学公式如下：

$$\begin{cases} \dot{s}_i(t) = v_i(t) \\ m\dot{v}_i(t) = u_i(t) - (c_0 + c_1 v_i(t) + c_2 v_i^2(t)) \end{cases} \quad i = 1, \cdots, n \tag{1}$$

式中：$s_i(t)$为列车i在t时刻的位置，m；$v_i(t)$为列车i在t时刻的速度，m/s；$u_i(t)$为列车i在t时刻牵引/制动系统下产生的实际作用力，N；c_0、c_1、c_2为戴维斯系数；n为编队中列车的数量；m为列车的质量，kg。

在VCTS中，采用基于“领导者-跟随者”的控制方案[18]，可将“前后车速度一致，并以VC安全追踪间隔运行”的状态作为平衡态。领导车追踪参考曲线，其中，VCTS的理想追踪间隔为d_r，在设置d_r时将预留一定的安全裕量，目标速度为v_r，且$\dot{d}_r = v_r$，u_r为期望牵引/制动力，假设领导车位移误差为$\hat{s}_1 = s_1 - d_r$，速度误差为$\hat{v}_1 = v_1 - v_r$，控制变量误差为$\hat{u}_1 = u_1 - u_r$。而跟随车则以前车为追踪目标，即与前车状态量的偏差为误差值，则定义跟随车的位移误差为$\hat{s}_i = s_{i-1} - s_i - L - d_r$，$i = 2,3,\cdots,n$，$L$为列车长，速度误差为$\hat{v}_i = v_{i-1} - v_i$，控制误差变量为$\hat{u}_i = u_{i-1} - u_i$。图1为VC下的列车追踪运行图。

对式(1)进行泰勒展开并将其线性化，可得围

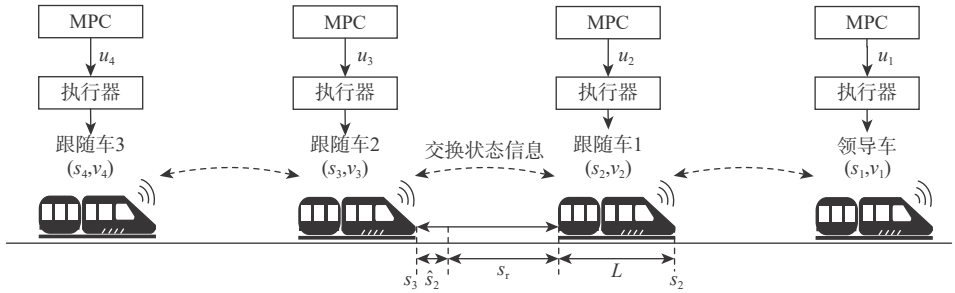


图 1　虚拟编组下的列车追踪运行图

Fig. 1　Train tracking operation chart under virtual coupled

绕平衡态的线性化连续动力学方程为

$$\begin{cases}\dot{\hat{s}}_i(t)=\hat{v}_i(t)\\ m\dot{\hat{v}}_i(t)=\hat{u}_i(t)-c_1\hat{v}_i(t)-2c_2\hat{v}_i(t)v_r(t)\end{cases}\quad i=1,2,\cdots,n \tag{2}$$

选取 $\boldsymbol{x}_i=[\hat{s}_i,\hat{v}_i]^{\mathrm{T}}$ 为状态变量, $\hat{u}_i$ 为控制变量, 对式 (2) 进行一阶差商离散化, 可得状态方程为

$$\dot{\boldsymbol{x}}_i(k+1)=\boldsymbol{A}\boldsymbol{x}_i(k)+\boldsymbol{B}u_i(k) \tag{3}$$

式中: $\boldsymbol{A}=\begin{bmatrix}0 & 1\\ 0 & -\dfrac{c_1+2c_2v_r}{m}\end{bmatrix}$; $\boldsymbol{B}=\begin{bmatrix}0\\ \dfrac{1}{m}\end{bmatrix}$。

在预测时域 N_p 内预测的系统状态 $\boldsymbol{X}_i(k)=[\boldsymbol{x}_i(k+1|k),\cdots,\boldsymbol{x}_i(k+j|k),\cdots,\boldsymbol{x}_i(k+N_p|k)]^{\mathrm{T}}$, 输入控制序列为 $\boldsymbol{U}_i(k)=[u_i(k|k),u_i(k+1|k),\cdots,u_i(k+N_p-1|k)]^{\mathrm{T}}$。在 VCTS 中, 为使编队列车快速达到理想控制状态, 以速度作为系统的输出, 对应输出矩阵为 $\boldsymbol{W}$, 具体形式如下:

$$\begin{cases}\begin{cases}\boldsymbol{X}_i(k+1)=\boldsymbol{\psi}\boldsymbol{X}_i(k)+\boldsymbol{\theta}\boldsymbol{U}_i(k)\\ \boldsymbol{Y}_i(k)=\boldsymbol{W}\boldsymbol{X}_i(k)\end{cases} & i=1,\cdots,n\\ \boldsymbol{\psi}=\left[\boldsymbol{A},\boldsymbol{A}^2,\cdots,\boldsymbol{A}^{N_p}\right]^{\mathrm{T}}\\ \boldsymbol{\theta}=\begin{bmatrix}\boldsymbol{B} & 0 & 0 & \cdots & 0\\ \boldsymbol{AB} & \boldsymbol{B} & 0 & \cdots & 0\\ \vdots & \vdots & \vdots & & \vdots\\ \boldsymbol{A}^{N_p-1}\boldsymbol{B} & \boldsymbol{A}^{N_p-2}\boldsymbol{B} & \boldsymbol{A}^{N_p-3}\boldsymbol{B} & \cdots & \boldsymbol{B}\end{bmatrix}\\ \boldsymbol{W}=\begin{bmatrix}0 & 1 & 0 & 0 & \cdots & 0 & 0\\ 0 & 0 & 0 & 1 & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots\\ 0 & 0 & 0 & 0 & \cdots & 0 & 1\end{bmatrix}_{N_p\times 2N_p}\end{cases} \tag{4}$$

1.2　防撞函数设计

高速列车领域中, 人工势场法[19] 是将本列车的运行环境抽象为势力场, 将障碍物作为高势能点, 目标点作为最低势能点。在 VCTS 运行过程中, 前车不仅目标点需要靠近, 同时障碍物也不能发生碰撞。因此, 列车在合力的作用下追踪运行, 最终朝着最低势能点运行, 即实现 VC 控制目标, 以相对速度差为零和最小间隔追踪运行。

在设计势力场前, 需先明确各列车间的信息交互关系, 即建立通信拓扑图[20], 只有建立通信关系的列车间才存在人工势场。在“领导者-跟随者”模型中, 前车采集本车状态信息并通过无线通信发送给后车, 后车接收前车信息做出动作决策, 并反馈给前车, 同时也将自身信息传递给跟随者。由此可知, 本车与前后相邻的列车均存在通信关系, 即用连接权重 C_{ij} 来表示, 若列车间存在交互关系, 则 $C_{ij}=1$, 否则取 0。

为实现 VCTS 追踪运行, 需实时获取列车的状态信息, 从而产生相应的作用力。当列车间隔过大时, 列车间的引力效果更明显, 随着列车不断靠近目标前车, 引力相应减小; 当列车间隔达到理想最小追踪间隔时, 列车间的势场不产生作用力; 当列车间隔小于理想追踪间隔时, 为避免事故发生, 列车间会产生斥力作用, 列车间隔越小, 斥力越大。因此, 选取 tanh( ) 为势场力函数, 可根据列车间状态信息产生斥力或引力。VCTS 的人工势场原理如图 2 所示。

列车间的距离势场函数为

$$J_{is}=\begin{cases}-\displaystyle\sum_{j=1}^{n}k_sC_{ij}\cdot\tanh(s_i-d_r) & i=1\\ -\displaystyle\sum_{j=1}^{n}k_sC_{ij}\cdot\tanh(s_{ij}-L-d_r) & i=2,\cdots,n\end{cases} \tag{5}$$

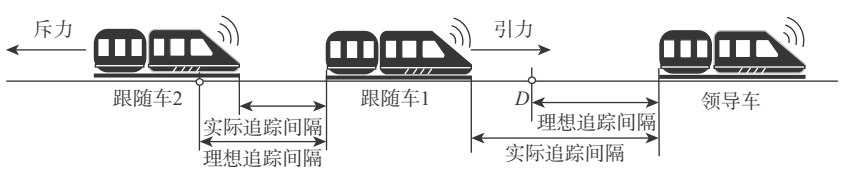


图 2　虚拟编组列车组的人工势场原理

Fig. 2　Schematic diagram of artificial potential field of virtual coupled train set

式中：k_s为距离加权系数；s_i-d_r和$s_{ij}-L-d_r$表示相邻2个列车与理想追踪间隔的误差，s_{ij}为列车i和列车j之间的实际运行间隔。当误差值为0时，列车间达到平衡态，J_{is}为0。当误差值大于0时，列车在引力作用下间隔不断减小。当误差值小于0时，列车在斥力作用下间隔不断增大。

为实现最终追踪控制目标，不可忽视速度误差，因此，引入列车间的速度势场函数，公式如下：

$$J_{iv}=\begin{cases}-\sum_{j=1}^{n}k_vC_{ij}\cdot\tanh(v_i-v_r) & i=1\\ -\sum_{j=1}^{n}k_vC_{ij}\cdot\tanh(v_j-v_i) & i=2,\cdots,n\end{cases} \tag{6}$$

式中：k_v为速度加权系数；v_i-v_r和v_j-v_i表示邻间列车的速度误差，作用原理同距离势场函数。

为有效避免列车间发生碰撞，进一步提高控制器的安全性，结合距离、速度的合势场函数为

$$J_{i\mathrm{APF}}=J_{is}+J_{iv} \tag{7}$$

1.3 优化问题的构建

为实现对VCTS的稳定安全追踪控制，即在减少与参考曲线偏差的同时实现防撞控制，目标函数应包括误差、控制量和合势场3项，且运行时还需满足列车性能的约束，则目标函数和约束条件定义如下：

$$\begin{cases}\min\limits_{u_i(k)}J(\boldsymbol{x}_i(k),u_i(k))=\sum\limits_{j=0}^{N_p-1}|\boldsymbol{Y}_i(k+j)-\boldsymbol{Y}_{i\mathrm{ref}}(k+j)|_{\boldsymbol{Q}}^2+\\ \sum\limits_{j=0}^{N_c}|u_i(k+j)|_{\boldsymbol{R}}^2+\sum\limits_{j=0}^{N_p-1}|J_{i\mathrm{APF}}(k+j)|_{\boldsymbol{P}}^2\\ \text{s.t.}\begin{cases}0\leqslant v_i(k+j|k)\leqslant v_{\lim}\\ u_{\min}\leqslant u_i(k+j|k)\leqslant u_{\max}\end{cases}\\ \forall j=0,1,\cdots,N_p-1\end{cases} \tag{8}$$

式中：N_p和N_c分别为系统的预测时域和控制时域；$\boldsymbol{Q}$、$\boldsymbol{R}$、$\boldsymbol{P}$为目标函数中各项的权重。目标函数的第1项为列车实际轨迹与参考曲线的误差值；第2项为系统的控制量，用于确保列车的稳定控制；第3项为基于人工势场的合势场函数，通过综合考虑列车间间隔差和速度差实现防撞控制。约束条件中，$v_{\lim}$为线路限制速度，$u_{\min}$和$u_{\max}$分别为列车的最小制动力和最大牵引力。

2 自适应时域参数的MPC设计

时域参数是MPC的关键设计参数，其预测时域N_p和控制时域N_c分别为系统的预测输出和控制输入的步长，该参数的选取直接影响控制器的性能[21]。为完善控制器的性能，在上述所设计的基于人工势场的MPC基础上，分析不同N_p和N_c对系统控制精度和实时性的影响。本文选取CRH380AL型高速列车作为研究对象，仿真场景为4列VC列车在平直轨道上追踪运行，总运行时间为620 s，参考曲线设置为：0～100 s时以0.5 m/s^{2}加速度加速运行，100～200 s时以50 m/s巡航运行，200 s时以0.7 m/s^{2}加速运行40 s直至速度为78 m/s，再以该速度从240 s巡航运行至500 s，500 s时以0.6 m/s^{2}减速度减速至42 m/s，560 s时保持42 m/s运行60 s。

2.1 时域参数对控制器性能影响分析

2.1.1 预测时域对控制器的影响

预测时域表示控制器对未来状态的预测程度，随着N_p的增大，获取系统未来状态的数量逐渐增加，因此，在控制精度上得到提升。本节选取$N_c=5,N_p=5,10,15,20$进行仿真验证。

图3为不同N_p下各列车与理想追踪间隔的误差。图3(a)中，在列车1启动初期，$N_p=5,N_c=5$时产生了0.285 1 m的超调间隔误差，这是由于N_p较小对未来状态考虑较少，系统控制精度较低，在前0～6 s内产生0.068 29 m的误差浮动，但短时间内可消除误差波动并趋于稳定。列车在100 s时加速运行，图3(c)的放大图中，因列车运行状态发生改变而产生较明显的误差波动，当N_p由5增加至20时，误差波动范围减小，达到稳定状态的时间缩短，稳定性增强。如图3(d)的局部放大图所示，327～387 s时列车4处于巡航状态，$N_p=5,10,15,20$时对应的误差分别为−11.84 m、−11.81 m、−11.74 m和−11.72 m，N_p增大使误差减小。因此，随着N_p的增大，预测未来状态的数量逐渐增加，误差和误差波动范围均逐渐减小，并且稳定性也有所提升。

图4为不同N_p下各列车与前车速度的误差。图4(a)、图4(b)、图4(d)的放大图中，列车分别处于启动、加速和减速运行阶段，而图4(c)的放大图中，列车处于巡航阶段。在图4(a)中，$N_p=5$时产生−0.459 7 m/s的超调量，因N_p过小，获取列车状态信息过少而产生较大的误差；$N_p=20$时也产生0.507 5 m/s的激增，因N_p过大，会增加未来时刻误差的权重，降低当前时刻误差的权重，此时控制量不再是最优值，从而产生激增。在图4(a)、图4(b)、图4(d)中，因列车运行状态改变而产生的误差，$N_p=5$时误差波动范围最大，$N_p=10$次之，$N_p=20$最小，且收敛至稳定的速度也随着N_p增大而加快。图4(c)的放大图中，$N_p=5$时，误差在$(-4\times10^{-7},4\times10^{-7})$m/s范围内波动，其余$N_p$下误差近似为0。

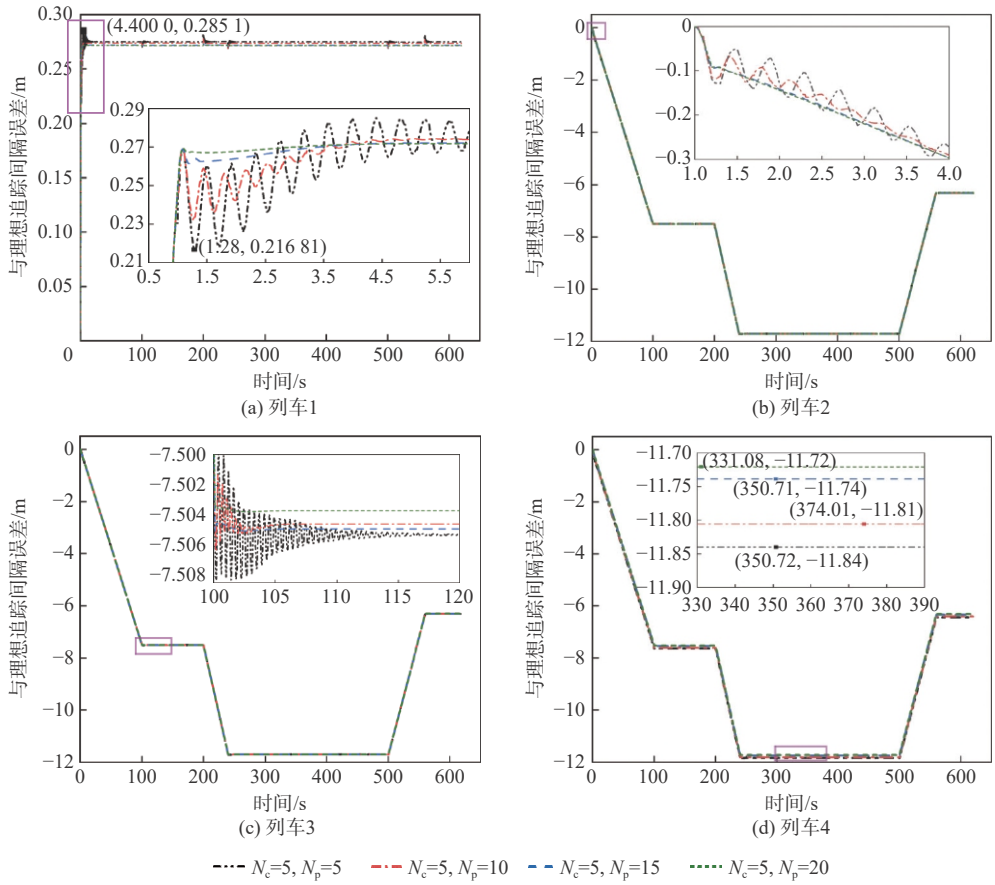


图 3　不同预测时域下各列车与理想追踪间隔的误差

Fig. 3　Error from ideal tracking interval of each train in different prediction time domains

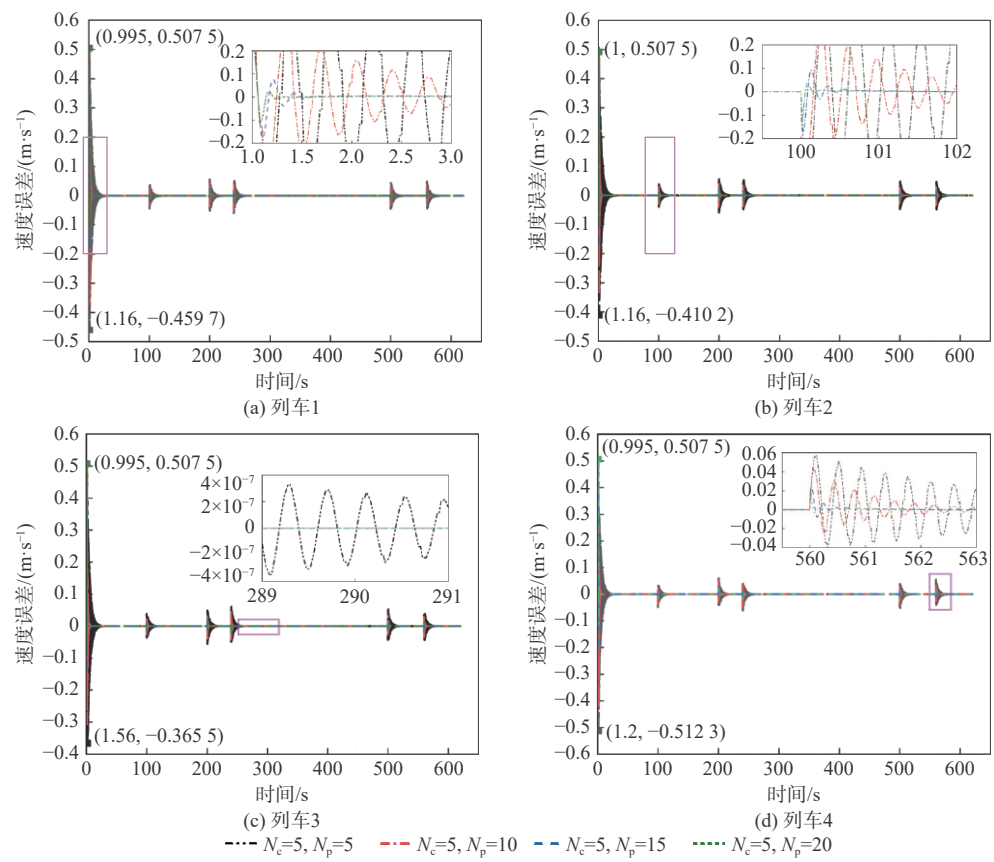


图 4　不同预测时域下各列车与前车速度的误差

Fig. 4　Error of speed with front train of each train in different prediction time domains

图5为不同N_p下各列车的计算时间。可知，随着N_p的增大，控制器的计算时间小幅度增加，但呈现无规律波动。主要原因是：N_p的增大使控制器在求解优化问题时需要预测更多时刻的状态信息，增加了目标函数的计算维度。以图5(a)列车1为例，N_p为5,10,15,20时，平均计算时间分别为0.013 521 s、0.014 366 s、0.015 526 s和0.016 03 s，总计算时间分别为251.942 8 s、267.685 s、289.302 s和298.685 1 s，N_p每增加5，增长率分别为6%、8%和3%。可得，计算时间随着N_p的增大而逐渐增加，但相对增幅较小，即N_p对计算效率的影响相对较小，且其余列车同样遵循该规律。

综上所述，N_p过小时，由于对列车状态预测较少而产生误差超调量。N_p过大时，又因增加了远时刻状态量的权重系数而影响当前时刻的控制精度。N_p增加，控制的准确性逐渐提高，但N_p对计算时间的影响相对较小，因此，需设置合适的N_p来完善MPC的控制性能和实时性。

2.1.2 控制时域对控制器的影响

控制器需在N_c内求解优化问题，得出最优控制序列。随着N_c增加，系统需要计算的控制向量维度增加，使追踪的准确性也逐渐提高。本节选取$N_p=20, N_c=5,10,15,20$进行仿真验证。

图6为不同N_c下各列车与理想追踪间隔的误差。由于在MPC方法中，仅将控制向量的第1个元素作用于系统，N_c对控制精度的影响不大。图6(b)的局部放大图为列车2起步初始时刻，1.4～1.6 s选取$N_c=5$时产生的最大误差波动范围为0.057 38 m，在短时间内误差波动逐渐减小并趋于稳定，且$N_c=5$的误差波动最大，$N_c=20$最小。图6(a)、图6(d)的放大图中，列车处于加速和减速运行阶段，图6(a)中，列车1在100 s加速，N_c取5时产生0.003 1 m幅度的误差波动，其他较大的N_c误差幅度均比$N_c=5$时小。图6(d)中，列车4在560 s时减速产生小幅度波动，N_c取5和20时，两者间误差值为0.035 05 m，此时，不同N_c的控制差值以0.01 m为单位。图6(c)的放大图中，列车处于巡航状态，$N_c=5,10,15,20$的间隔误差分别为−11.705 49 m、−11.705 44 m、−11.705 27 m和−11.705 18 m，随着N_c增加，误差减小，并且不同N_c间误差在10^{-4}等级内。由此可知，随着N_c增加，系统跟踪控制性能提升，稳定性增强，但不同N_c作用效果差异较小。

图7为不同N_c下各列车与前车速度的误差。

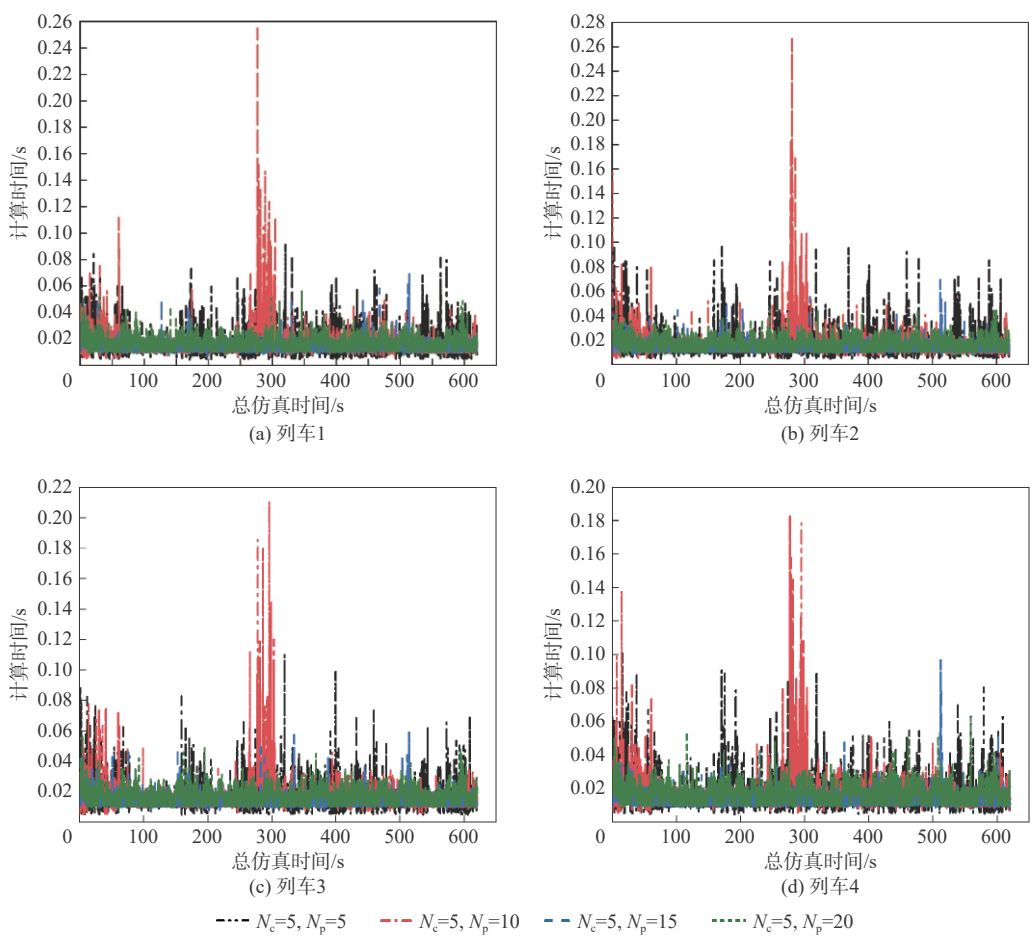


图5 不同预测时域下各列车的计算时间

Fig. 5 Computation time of each train in different prediction time domains

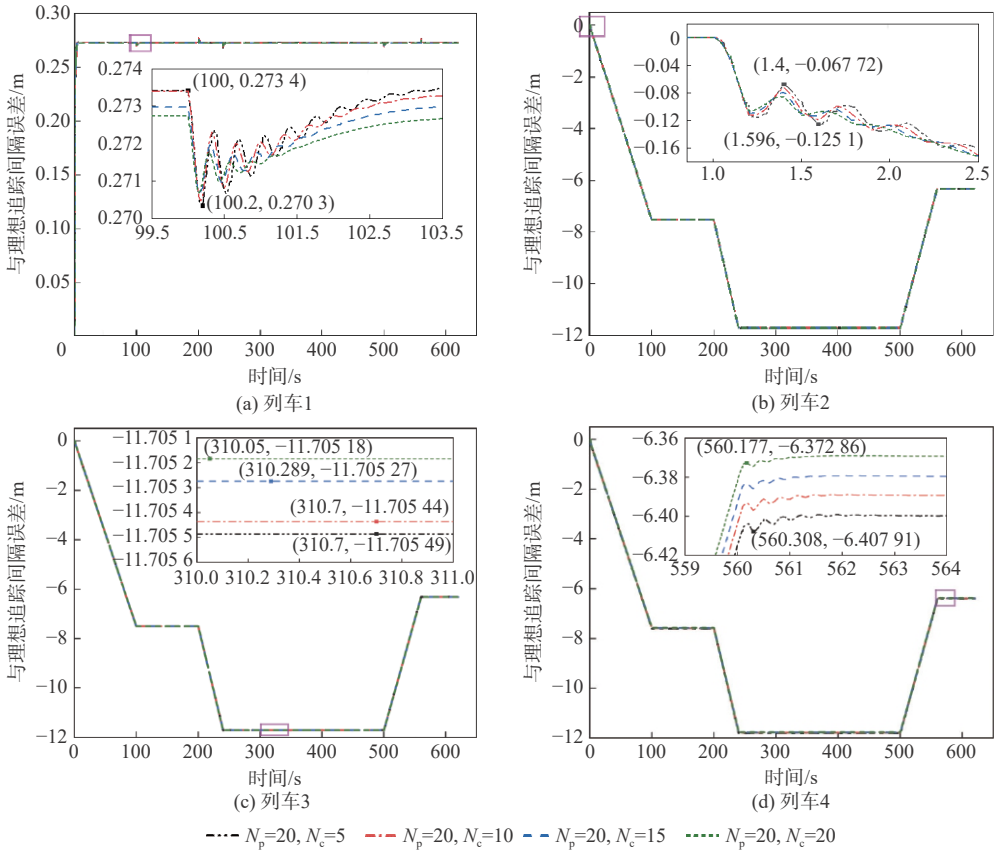


图 6　不同控制时域下各列车与理想追踪间隔的误差

Fig. 6　Error from ideal tracking interval of each train in different control time domains

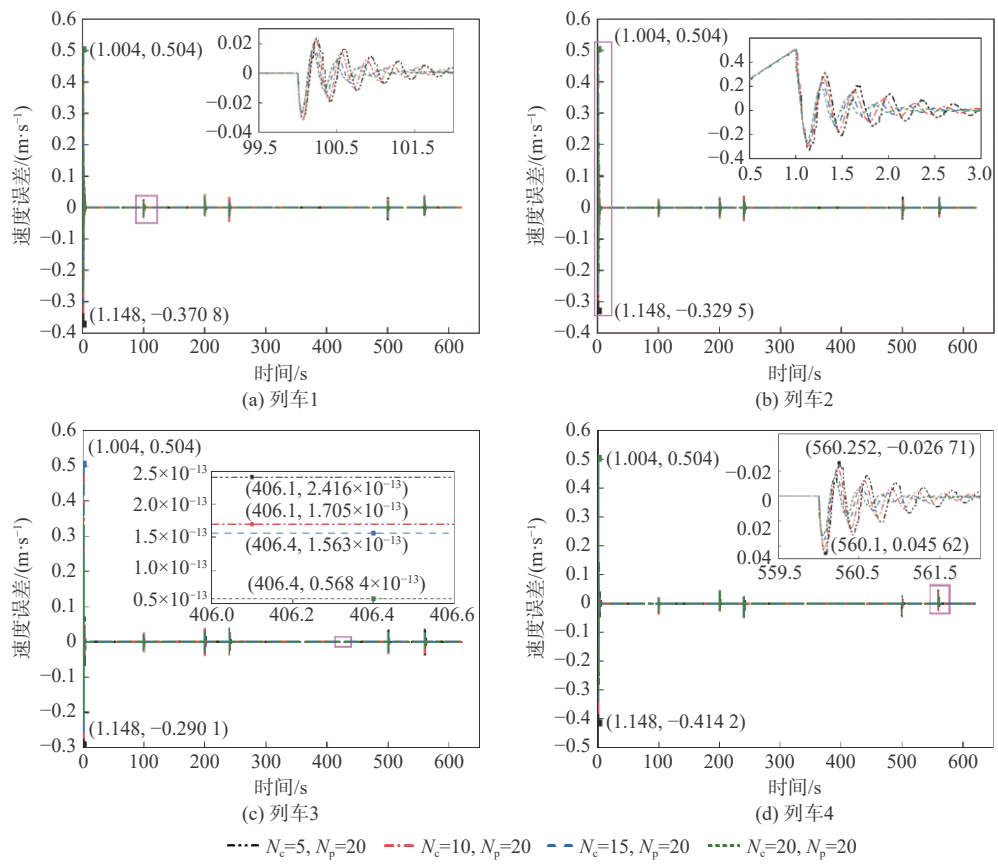


图 7　不同控制时域下各列车与前车速度的误差

Fig. 7　Error of speed with front train of each train in different control time domains

在图 7(a) 的放大图中，因列车状态改变而产生误差，并且随着N_c的增大，误差波动范围逐渐减小，趋于稳定的速度增快。图 7(d) 的放大图为列车 4 在 560 s 减速运行时的间隔误差，$N_c=5$时最大间隔误差幅度为 0.072 33 m/s，取较大N_c时的误差幅度均小于该值，因此，增加N_c可提高系统控制的准确度。图 7(c) 的放大图中，列车 3 处于巡航阶段，不同N_c下速度误差以10^{-13}为误差单位，在$N_c=20$时最小，误差为$0.568\,4\times10^{-13}$ m/s，$N_c=5$时误差为2.416×10^{-13} m/s，这两者间差值最大但也控制在$1.847\,6\times10^{-13}$ m/s 内。综上所述，不同N_c下虽在控制精度上有一定差异，但各值间的误差均控制在小范围内，即N_c对系统的控制性能影响较小。

图 8 为不同N_c下各列车的计算时间。相较于N_p，N_c的改变对系统实时性的影响较大，这是由于N_c的增加使控制序列的维度增加，从而增加了滚动优化运算的复杂度。在图 8(a) 中，$N_c=5$时的平均运算时间为 0.016 03 s，N_c每增加 5，控制时域$N_c=10,15,20$对应的平均计算时间分别为 0.017 228 s、0.019 303 s 和 0.022 762 s，N_c以 5 为涨幅，由 5 增加至 20 的增长率分别为 7%、12% 和 18%。在图 8(b)、图 8(c)、图 8(d) 中，即列车 2～4 亦符合该规律，随着N_c的增大，系统的计算时间逐渐增加，波动幅度也有小幅度上升，因此，N_c对系统计算时间影响较大，N_c过大会降低计算效率，在设计控制器时需根据需求选择合适的N_c。

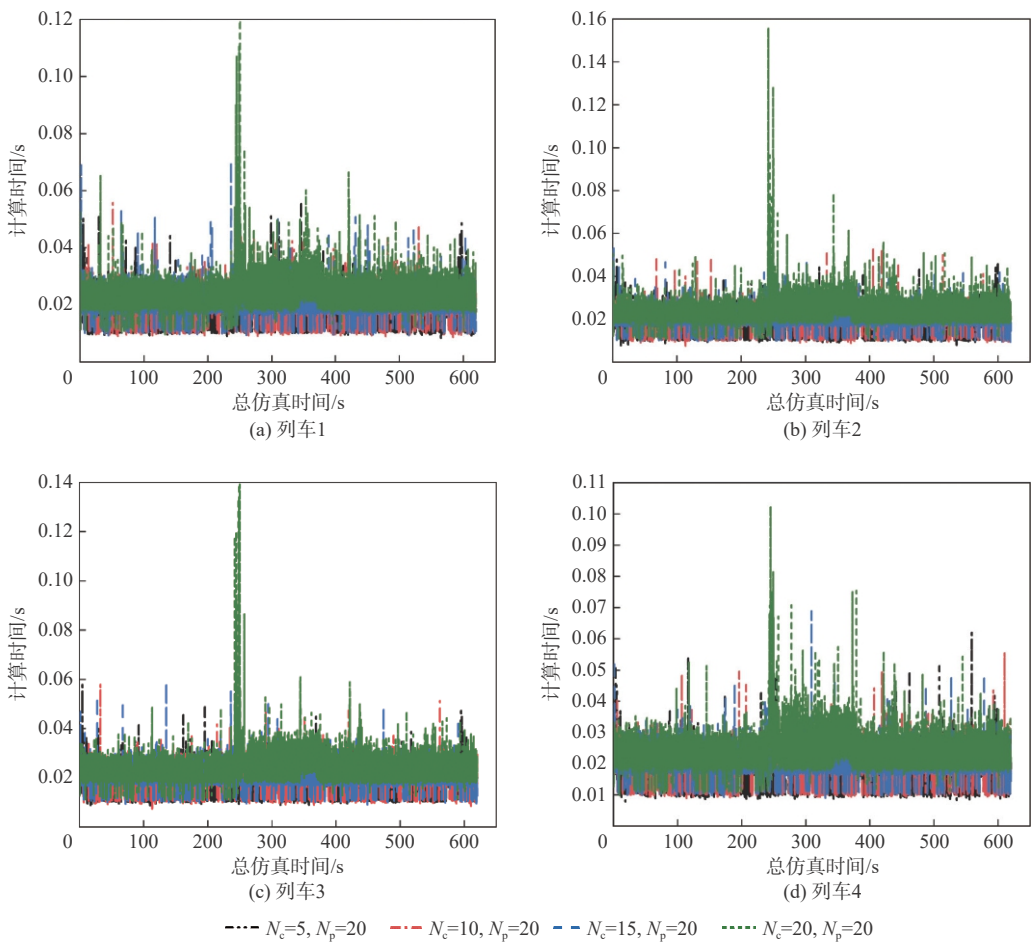


图 8　不同控制时域下各列车的计算时间

Fig. 8　Calculation time of each train in different control time domains

2.2　基于遗传算法的自适应 MPC 设计

2.1 节已分析N_p和N_c对 VCTS 追踪性能的影响，因此，选择合适的控制器时域参数至关重要。本文采用 GA 进行参数优化，GA 是利用“物竞天择、适者生存”的达尔文进化论观点优化控制，是一种全局优化算法[22]，适用于本文需在全运行时长内根据列车状态对控制器参数优化求解的情况，使系统可在 VCTS 不同运行工况下求得对应的最优参数组合，同时也为后续设计自适应更新策略提供依据。重要的是，高速列车对追踪运行的时效性有较高要求，而 GA 搜索能力强，能够快速找到最优组合，可以满足参数优化中实时性高的需求。基于 GA 对 MPC 中的N_p和N_c这 2 个参数优化时，一般需要进行以下步骤：

1) 初始化种群。在进行 GA 操作时, 采用随机数方法产生若干个初始解作为初始种群, 对染色体的每一维变量赋值, 并基于该种群进行遗传、交叉和变异操作[23]。

2) 适应度函数设计。适应度函数是度量某个物种对生存环境的适应程度, 即相当于 GA 中的目标函数。本文中, 参数优化不仅需要考虑 VCTS 在追踪过程中实际运行轨迹与参考轨迹的偏差, 即距离误差和速度误差, 还要考虑控制器的计算时间。由于这 3 个量的单位和数量级不一致, 采用极值化的方法进行无量纲化处理。

$$\gamma_i' = \frac{\gamma_i}{B-S} \tag{9}$$

式中: $\gamma_i(i=1,2,3)$分别为间隔误差的绝对值、速度误差的绝对值和计算时间; B 为多组试验的最大值; S 为多组试验的最小值。

选取 $k_i(i=1,2,3)$为加权系数, 则设计适应度函数为

$$J_{GA} = k_1\gamma_1' + k_2\gamma_2' + k_3\gamma_3' \tag{10}$$

3) 约束条件设置。由于在第 2 节设计控制器时已对列车性能和控制量进行约束, 在 GA 中只需对优化参数约束限制, 在减小迭代范围的同时也可缩短优化时间。本文选取的参数范围为: $N_p \in [5,20]$, $N_c \in [5,20]$, 且 $N_c \leqslant N_p$。

在全运行时长内, 基于 GA 求得的控制参数是全局最优参数, 该时域组合为固定值, 无法根据列车状态信息实时调整控制, 为增强控制器与被控对象的关联互通性, 考虑时域参数与控制器之间的自适应性。由于编队在运行状态转换时会产生误差波动, 选择在 42,50,78 m/s 这 3 个巡航速度下分别进行仿真, 采用 GA 可得各巡航速度下的最优时域参数组合。随后, 依据不同速度设计自适应控制律, 使控制器能够在线调整, 增强了灵活性, 并提高了综合性能。则制定时域参数更新策略如下:

$$[N_p, N_c] = \begin{cases} [9,5] & v < 42\ \text{m/s} \\ [9,6] & 42\ \text{m/s} \leqslant v < 50\ \text{m/s} \\ [15,9] & 50\ \text{m/s} \leqslant v < 78\ \text{m/s} \end{cases} \tag{11}$$

基于 GA 的自适应 MPC 控制框图如图 9 所示。

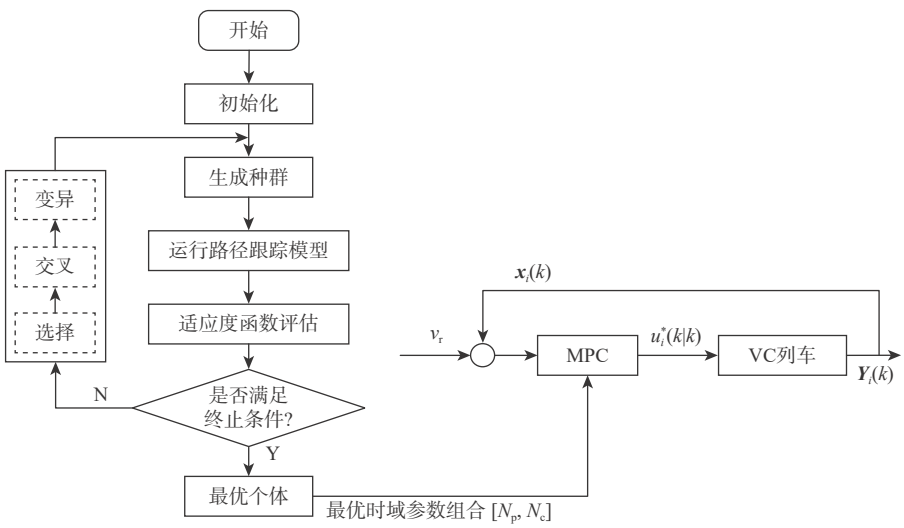


图 9　基于遗传算法的自适应 MPC 控制框图

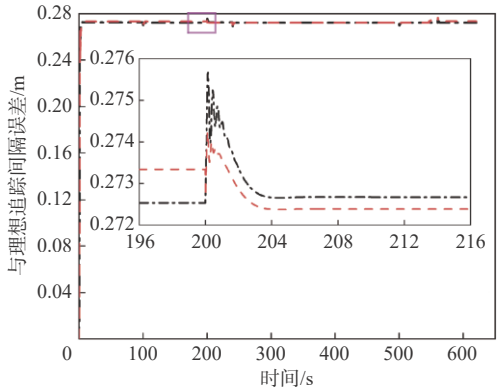
Fig. 9　Control block diagram of adaptive MPC based on genetic algorithm

3　仿真验证

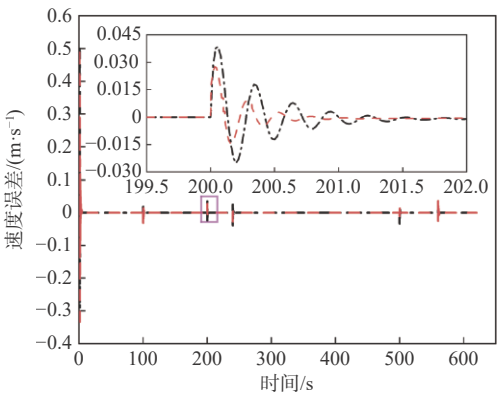
3.1　自适应时域参数控制器性能仿真验证

为验证时变参数对系统控制的有效性, 设计 2 种控制器, 即控制器 1 为基于 GA 的固定时域参数 $[N_p, N_c] = [13,11]$的 MPC, 该参数是 GA 全局寻优求解出的最优值; 控制器 2 为采用 2.2 节控制律的自适应时域参数 MPC, 该控制器可根据列车的运行状态实时选取合适的控制参数。本节以 VCTS 中列车 1 为例, 从速度误差、间隔误差、计算时间 3 个角度进行对比验证。

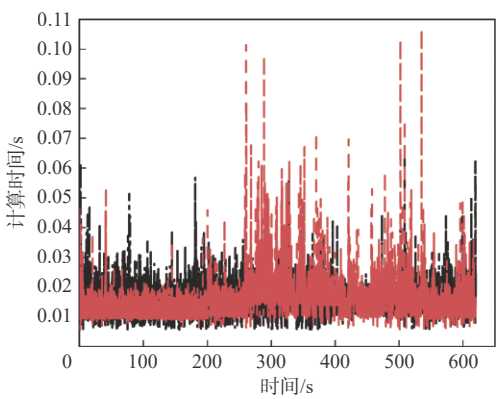
图 10 为 2 种控制器的综合控制性能对比。在控制精度方面, 图 10(a)、(b) 放大图中, 列车 1 在 200 s 时转为高速运行, 根据时域参数更新策略, 此时控制器 2 需更新时域参数组合。在 200 s 时, 控制器 2 的误差明显减小, 并且其误差值和波动范围均小于控制器 1, 在后续巡航阶段仍维持较小的误差值。在计算时间方面, 0～240 s 和 546～620 s 列车处于低速运行, 控制器 2 采用较小的时域参数, 在确保控制精度的前提下提高计算效率, 从图 10(c) 可看出, 控制器 2 的计算时间小于控制器 1, 而 240～546 s 列车高速运行, 此时需增大时域参数来



(a) 列车1与理想追踪间隔的误差



(b) 列车1与前车速度的误差



(c) 列车1的计算时间

控制器1 控制器2

图 10 2种控制器的综合控制性能对比

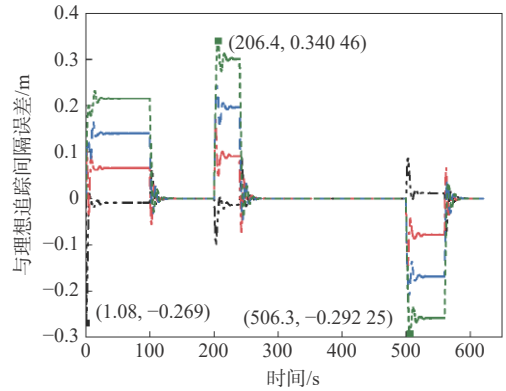
Fig. 10 Comparison of combined control performance of two controllers

提高控制的准确性，但总计算时间上，由控制器1的278 s降低至控制器2的250 s，计算效率提高10%。因此，相较于控制器1，控制器2可根据列车状态实时改变参数组合，使控制精度和计算效率达到最优平衡状态。

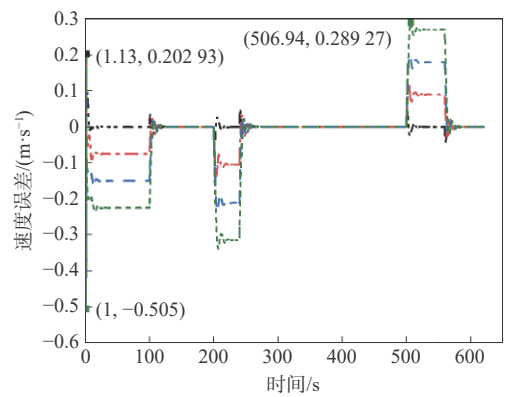
3.2 防撞的控制效果

为验证在目标函数中加入防撞项$J_{i\mathrm{APF}}$对编队列车控制精度的影响，采用基于人工势场的MPC与MPC进行对比仿真，从速度误差和间隔误差2个方面验证各列车防撞控制效果。

图11和图12分别为4列车基于人工势场的



(a) 与理想追踪间隔误差-时间曲线



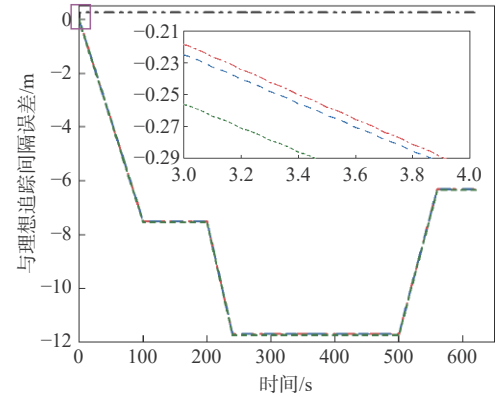
(b) 速度误差-时间曲线

列车1 列车2 列车3 列车4

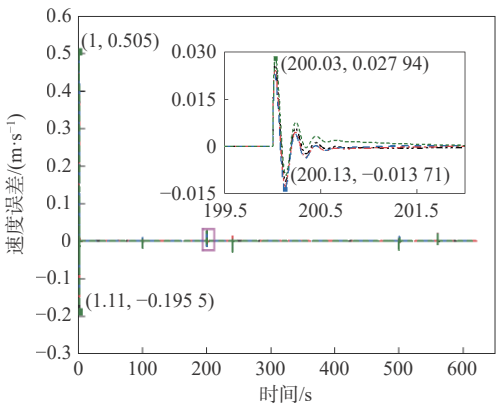
图 11 基于人工势场的MPC误差曲线

Fig. 11 Error curves for MPC based on artificial potential fields

MPC和传统MPC的误差曲线。图11(a)中，列车1在启动初期产生−0.269 m的激增，在列车运行状态改变的时刻，即t = 100,200,240,500,560 s时，均会产生误差波动，并且随着列车列数的增加，间隔误差增大，列车4在t = 200,500s时分别产生0.340 46 m和−0.292 25 m的误差，其余时刻的间隔误差均控制在[−0.292 25,0.340 46]m范围内。而图12(a)中，4列车的间隔误差则控制在[−12,0.5] m范围内，由局部放大图可知，在初始时刻，列车4误差最大，列车1的误差最小，但其误差范围都远大于图11(a)的误差范围，在间隔控制上准确度提高了94.8%。因此，在目标函数中加入防撞项$J_{i\mathrm{APF}}$，可有效缩小列车间的追踪间隔误差，并且由于安全裕量的存在，可在一定程度上减小列车发生碰撞的情况，提高了系统安全性。图11(b)与图12(b)中，速度误差范围变化不明显，可分别控制在[−0.505,0.289 27]m/s和[−0.195 5,0.505]m/s范围内，在状态转换时出现误差波动，对列车1的控制精度最高，随着列数的增加，造成误差积累，但最终均趋于稳定状态。因此，基于人工势场的MPC控制器，在间隔控制方面，可以大幅度减小间隔误差，并能防止列车间发生碰撞的情况；在速度控制方面，也能保证较高的准确度。



(a) 与理想追踪间隔误差-时间曲线



(b) 速度误差-时间曲线

图 12　基于 MPC 的误差曲线

Fig. 12　Error curves based on MPC

4　结　论

1) 基于列车平衡态搭建动力学模型, 为使编队列车能够保持安全间隔追踪运行且不发生碰撞, 在目标函数中引入基于人工势场构造的防撞函数, 仿真结果表明, 加入防撞项的 MPC 可有效减小列车间隔误差为负值的情况, 并且由于安全裕量的存在, 可避免列车发生碰撞。

2) 为增加控制器的自适应性, 引入 GA 求解列车不同速度下的最优时域参数组合。首先, 分析控制时域参数对 MPC 性能的影响, 为设置 GA 约束范围提供参考; 随后, 综合考虑控制精度和计算时间的作用设计适应度函数, 使系统保持良好的控制性能。仿真结果表明, 自适应参数的 MPC 可根据列车运行状态实时改变控制参数, 从而实现在确保控制精度的同时提高系统计算的效率。

在列车追踪过程中, 状态转换时仍存在较明显的误差波动, 并且列车运行环境复杂多变, 需加强列车与环境的交互反馈作用, 后续将针对该部分进行更深入的研究, 进一步提高系统的稳定性和实用性。

参考文献（References）

[1] 张友兵, 刘岭. 基于虚拟编组的列车控制系统研究[J]. 铁道工程学报, 2022, 39(3): 94-100.
ZHANG Y B, LIU L. Research on the train control system based on virtual coupling[J]. Journal of Railway Engineering Society, 2022, 39(3): 94-100(in Chinese).

[2] PAN D, ZHENG Y P. Dynamic control of high-speed train following operation[J]. PROMET-Traffic & Transportation, 2014, 26(4): 291-297.

[3] 刘俊囡. 高速列车编队多智能体运行控制优化方法[D]. 北京: 北京交通大学, 2020.
LIU J N. Optimization method of multi-agent operation control for high-speed train formation[D]. Beijing: Beijing Jiaotong University, 2020(in Chinese).

[4] BAI W Q, LIN Z L, DONG H R, et al. Distributed cooperative cruise control of multiple high-speed trains under a state-dependent information transmission topology[J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 20(7): 2750-2763.

[5] 汤旻安, 王攀琦. 基于混合系统建模预测控制的列车自动驾驶优化运行[J]. 铁道科学与工程学报, 2019, 16(6): 1527-1534.
TANG M A, WANG P Q. Automatic train optimization operation based on hybrid system modeling predictive control[J]. Journal of Railway Science and Engineering, 2019, 16(6): 1527-1534(in Chinese).

[6] 杨辉, 童英赫, 付雅婷, 等. 基于模型补偿的高速列车状态反馈预测控制[J]. 铁道科学与工程学报, 2020, 17(10): 2460-2468.
YANG H, TONG Y H, FU Y T, et al. State feedback predictive control of high-speed train based on model compensation[J]. Journal of Railway Science and Engineering, 2020, 17(10): 2460-2468(in Chinese).

[7] 林俊亭, 倪铭君. 基于扩张状态观测器的虚拟编组触发模型预测控制[J]. 交通运输系统工程与信息, 2023, 23(4): 134-146.
LIN J T, NI M J. Trigger model predictive control based on extended state observers for virtual coupling[J]. Journal of Transportation Systems Engineering and Information Technology, 2023, 23(4): 134-146(in Chinese).

[8] 周红梅. 微型无人机集群编队控制与协作算法研究[D]. 成都: 电子科技大学, 2020: 50-59.
ZHOU H M. Research on formation control and cooperation algorithm of micro UAV cluster[D]. Chengdu: University of Electronic Science and Technology of China, 2020: 50-59(in Chinese).

[9] 朱梁鹏. 基于路径跟踪的多智能车辆编队控制研究[D]. 常州: 江苏理工学院, 2021: 56-66.
ZHU L P. Research on multi-intelligent vehicle formation control based on path tracking[D]. Changzhou: Jiangsu University of Technology, 2021: 56-66(in Chinese).

[10] 吴宝库. 考虑避障的多 AGV 编队协同控制与路径规划设计[D]. 沈阳: 沈阳大学, 2021: 47-58.
WU B K. Cooperative control and path planning design of multi-AGV formation considering obstacle avoidance[D]. Shenyang: Shenyang University, 2021: 47-58(in Chinese).

[11] 钟嘉灏. 面向多目标多无人机任务分配与航迹规划算法研究[D]. 成都: 电子科技大学, 2022: 59-74.

ZHONG J H. Research on task assignment and path planning algorithm for multi-target and multi-UAV[D]. Chengdu: University of Electronic Science and Technology of China, 2022: 59-74(in Chinese).

[12] 熊俊涛, 李中行, 陈淑绵, 等. 基于深度强化学习的虚拟机器人采摘路径避障规划[J]. 农业机械学报, 2020, 51(增刊 2): 1-10.
XIONG J T, LI Z X, CHEN S M, et al. Obstacle avoidance planning of virtual robot picking path based on deep reinforcement learning[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(Sup 2): 1-10(in Chinese).

[13] 丁伟兵. 园艺电动拖拉机局部路径规划与跟踪控制研究[D]. 镇江: 江苏大学, 2022: 28-45.
DING W B. Research on local path planning and tracking control of horticultural electric tractor[D]. Zhenjiang: Jiangsu University, 2022: 28-45(in Chinese).

[14] 张志伟. 基于模型预测控制的四轮转向 AGV 路径跟踪横向控制研究[D]. 上海: 上海交通大学, 2020: 68-83.
ZHANG Z W. Study on lateral control of path tracking for four-wheel-steeling AGV based on model predictive control[D]. Shanghai: Shanghai Jiao Tong University, 2020: 68-83(in Chinese).

[15] 白国星, 孟宇, 刘立, 等. 基于可变预测时域及速度的车辆路径跟踪控制[J]. 中国机械工程, 2020, 31(11): 1277-1284.
BAI G X, MENG Y, LIU L, et al. Path tracking control of vehicles based on variable prediction horizon and velocity[J]. China Mechanical Engineering, 2020, 31(11): 1277-1284(in Chinese).

[16] 杨阳阳. 基于模型预测控制的路径跟踪控制方法研究[D]. 镇江: 江苏大学, 2018: 31-49.
YANG Y Y. Research on path tracking control method based on model predictive control[D]. Zhenjiang: Jiangsu University, 2018: 31-49(in Chinese).

[17] YANG Y M, YAN F. Research on train dynamic coupling strategy based on distributed model predictive control[J]. Journal of Physics: Conference Series, 2022, 2183(1): 012029.

[18] LIU Y F, ZHOU Y, SU S, et al. An analytical optimal control approach for virtually coupled high-speed trains with local and string stability[J]. Transportation Research Part C: Emerging Technologies, 2021, 125: 102886.

[19] ZHANG T R, XU J N, WU B K. Hybrid path planning model for multiple robots considering obstacle avoidance[J]. IEEE Access, 2022, 10: 71914-71935.

[20] HUANG Z W, WANG P P, ZHOU F, et al. Cooperative tracking control of the multiple-high-speed trains system using a tunable artificial potential function[J]. Journal of Advanced Transportation, 2022, 2022: 3639586.

[21] 吴施鹏. 基于遗传算法的无人驾驶车辆路径跟随控制及研究[D]. 常州: 江苏理工学院, 2021: 40-63.
WU S P. Genetic algorithm-based path-following control and research of driverless vehicles[D]. Changzhou: Jiangsu University of Technology, 2021: 40-63(in Chinese).

[22] YU D J, DENG F, WANG H Y, et al. Real-time weight optimization of a nonlinear model predictive controller using a genetic algorithm for ship trajectory tracking[J]. Journal of Marine Science and Engineering, 2022, 10(8): 1110.

[23] WANG M, CHEN J X, YANG H W, et al. Path tracking method based on model predictive control and genetic algorithm for autonomous vehicle[J]. Mathematical Problems in Engineering, 2022, 2022: 4661401.

Adaptive model predictive control of virtual coupled based on artificial potential field

LIN Junting[1, *], NI Mingjun[1, 2]

(1. School of Automation and Electrical Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China;
2. Xinjiang Jiaotong Vocational and Technical University, Urumqi 831400, China)

Abstract: The safety of operating formation trains is now subject to stricter standards due to the high speed and high density tracking control of trains. To meet the need for precision and adaptability during train operation, an adaptive model predictive control (MPC) technique based on an artificial potential field is created for virtual coupled (VC) systems. First, a VC train is used as the research object, and the MPC method is used to create a dynamic model based on the equilibrium state of the train, with control accuracy and smoothness as the optimization objectives. Next, a collision avoidance function based on an artificial potential field setting is added to the objective function, allowing for the realization of the formation's collision avoidance control. In order to improve the real-time performance of the system and ensure accurate control of the train formation, it is also necessary to analyze the impact of various time domain parameters on the control accuracy and computational efficiency of the system, design the corresponding adaptation function, and base this on a genetic algorithm (GA) that can find the best combinations of time domain parameters under various working conditions. On the MATLAB platform, a 4-train tracking operation scenario is developed as a last step to test the effectiveness of the suggested approach. The study's findings demonstrate that the artificial potential field-based model predictive controller has interval control accuracy that is 94.8% greater than that of the conventional model predictive controller, successfully preventing train collisions and ensuring the safety of train operation. Furthermore, the adaptive control law controller may adjust the system in real-time based on the train's operational status, resulting in a 10% increase in computing efficiency and excellent control accuracy. The study's findings support the viability of the suggested control approach, enhance the controller's overall control performance, and serve as a guide for future improvements in formation control and ensuring safe train operation.

Keywords: virtually coupled; model predictive control; artificial potential field; genetic algorithm; train tracking operation optimization

Received: 2023-08-25; **Accepted**: 2023-10-17; **Published Online**: 2023-11-20 11:03
URL: link.cnki.net/urlid/11.2625.V.20231118.1404.001
Foundation item: National Natural Science Foundation of China (52162050)
* **Corresponding author**. E-mail: linjt@lzjtu.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0555

基于子模型的复合材料主承力混杂连接结构失效分析

李星[1]，黎增山[2,*]，邓凡臣[3]，聂磊[1]，张田[1]

(1. 中国商飞北京民用飞机技术研究中心 民用飞机结构与复合材料北京市重点实验室，北京 102211；2. 北京航空航天大学航空科学与工程学院，北京 100191；3. 中国飞机强度研究所 复合材料结构强度研究室，西安 710065)

摘 要：以民机复合材料机翼翼根下壁板长桁对接结构为研究对象，通过建立单钉连接三维子模型并考虑材料损伤，分析连接失效机理及非线性载荷-变形关系；通过考虑非线性连接刚度的整体有限元模型，计算试验件非线性钉载分配，并预测破坏载荷。解决了直接建立全尺寸三维模型、考虑损伤等引起的计算效率低、收敛性差等问题。试验结果表明：预测破坏载荷为试验破坏载荷的93.6%，预测破坏模式为单侧多排紧固件同时剪切断裂，与试验破坏模式一致，验证了所提方法的有效性。相比之下，采用线性连接刚度模型预测的破坏载荷为试验破坏载荷的81.9%，预测失效模式为第1排紧固件断裂，预测破坏载荷、预测破坏模式均与试验结果存在较大差异。

关 键 词：复合材料；混杂连接；钉载分配；子模型；失效分析

中图分类号：V257; TB332

文献标志码：A **文章编号**：1001-5965(2025)10-3286-13

随着复合材料在民机主承力结构中的广泛应用[1]，关键连接结构逐渐引起人们的重视。以某型民机外翼-中央翼连接结构为例，包含复合材料、钛合金、铝合金等零部件组成的复杂连接，质量可占外翼盒段总重的4%～5%，其设计水平对民机安全性、经济性均有重要影响。在工程设计中，通常按线性计算复合材料多排紧固件连接结构的钉载分配[2]；与之相比，金属材料多排紧固件连接结构设计中，由于金属的塑性特点，在考虑一定系数的条件下，可认为多排紧固件载荷平均分配(静强度设计)[2]。按当前设计理念和方法，复合材料连接结构相比金属连接结构，其结构效率并不占明显优势。

为了在保证安全性的前提下进一步发掘复合材料结构的减重潜力，国内外研究学者针对复合材料连接结构的载荷分配和失效机理开展了充分的研究和探讨。Hart-Smith[3]基于试验结果和分析，提出了考虑初始间隙、预紧力等因素的复合材料机械连接非线性载荷-变形曲线，并应用于复合材料胶接、螺接混合连接结构的失效分析中。Gray等[4]建立了二维复合材料层板、三维紧固件的混合有限元模型，分析复合材料多钉连接结构的非线性特征，同时采用ABAQUS建立了用于模拟复合材料紧固件连接的UEL自定义单元，并应用于多钉连接结构的钉载分配模拟。Zhang等[5]应用二维非线性有限元方法，分析了复合材料多钉连接结构在剪切载荷作用下的力学响应及失效机理。Sasikumar等[6]采用三维分析模型，研究了高低温、摩擦系数、螺栓夹紧力、钉孔间隙等因素对复合材料-铝合金混杂连接结构的影响。李玺等[7]结合Puck准则和刚度幂指数衰减法研究了复合材料多钉连接结构的渐进损伤过程。黄河源等[8]建立了一种复合材料损伤本构模型，并应用于孔边失效分析，定义了连接结构的非线性损伤函数，对连接结构非线性段力学响应进行了准确模拟。国内相关学者也充分讨

收稿日期：2023-08-29；录用日期：2023-11-10；网络出版时间：2023-12-13 17:20
网络出版地址：link.cnki.net/urlid/11.2625.V.20231213.1442.001
基金项目：国家重点研发计划 (2022YFB4301002)
* 通信作者. E-mail：lizengshan@buaa.edu.cn

引用格式：李星，黎增山，邓凡臣，等. 基于子模型的复合材料主承力混杂连接结构失效分析 [J]. 北京航空航天大学学报，2025，51（10）：3286-3298. LI X，LI Z S，DENG F C，et al. Sub-model based failure analysis of composite primary load-bearing hybrid joint structure [J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51（10）：3286-3298（in Chinese）.

论了环境温度、间隙/干涉、预紧力等因素对复合材料-金属混杂机械连接结构钉载分配的影响[9-12]。

国内外已有的复合材料连接结构载荷分配及失效机理研究中, 通常以简单连接结构为研究对象, 开展详细的建模分析。实际工程结构连接件众多, 连接关系复杂, 直接采用全尺寸三维渐进损伤分析存在计算效率低、收敛性差等缺点。为解决以上问题, 本文提出一种基于子模型的复合材料混杂连接结构失效分析方法, 以某型民机复合材料外翼-中央翼主承力连接结构为研究对象, 探讨了连接结构钉载分配的非线性演化过程及失效机理, 并通过试验进行验证。研究内容来源于大型民机复合材料机翼研制攻关项目, 可为民机复合材料主承力混杂连接结构的设计、分析提供参考和借鉴。

1　研究对象及分析方法

1.1　研究对象

研究对象取自某型民机复合材料机翼翼根下壁板对接区域, 外翼、中央翼均设计为 T 型长桁加筋壁板, 长桁腹板在端部斜削, 翼根下壁板对接(1#肋处)采用琴键式水平三叉接头、外部对接带板的连接形式, 示意图如图 1 所示。对于翼根下壁板对接结构, 主要承受外翼壁板的拉伸载荷, 以紧固件连接的形式将外翼壁板载荷通过水平三叉接头和对接带板传递至中央翼壁板, 从而达到受力平衡。

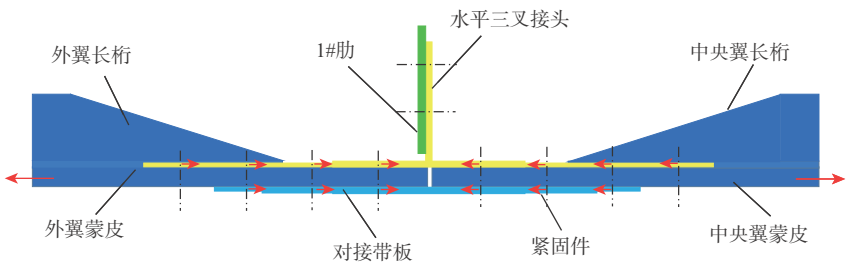


图 1　复合材料机翼翼根下壁板对接剖面

Fig. 1　Cross section of composite wing root lower panel butt joint

真实翼根下壁板连接结构存在后掠角和上反角, 分别会对翼根连接的传载大小、传载比例产生影响[13]。考虑到试验难度及研究目的主要为混杂连接结构的非线性钉载分配及失效机理, 试验件忽略后掠角、上反角影响, 两侧长桁采用对称设计, 以外翼侧长桁对接结构为试验考核对象。

外翼、中央翼壁板均为 T800 级碳纤维环氧复合材料, 蒙皮铺层比例为 45/50/10, 长桁铺层比例为 60/30/10; 水平三叉接头、外部对接带板材料为 Ti-6Al-4V; 1# 肋材料为 7050-T7451。外翼、中央翼侧分别采用 4 排 14# 钛合金高锁螺栓连接, 且水平三叉接头、对接带板分别与 4 排、3 排紧固件连接, 以避免局部刚度突变。试验件沿展向包含 4 列紧固件, 整体尺寸约为 1 900 mm(长)×361 mm(宽)×172 mm(高), 如图 2 所示。

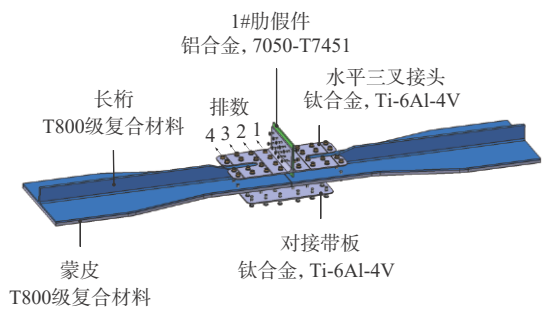


图 2　翼根长桁对接试验件示意图

Fig. 2　Schematic diagram of wing root stringer joint specimen

1.2　分析方法

翼根下壁板长桁对接试验件在蒙皮面内包含 32 颗高锁螺栓, 若采用考虑接触、损伤等因素的全三维有限元模型建模, 模型规模大且计算效率较低。因此, 提出了基于单紧固件连接子模型的建模方法, 通过子模型分析研究金属塑性、复合材料损伤对钉孔连接承载能力及刚度的影响, 并最终集成至全尺寸连接模型中。

子模型建模思路如图 3 所示。对于长桁对接

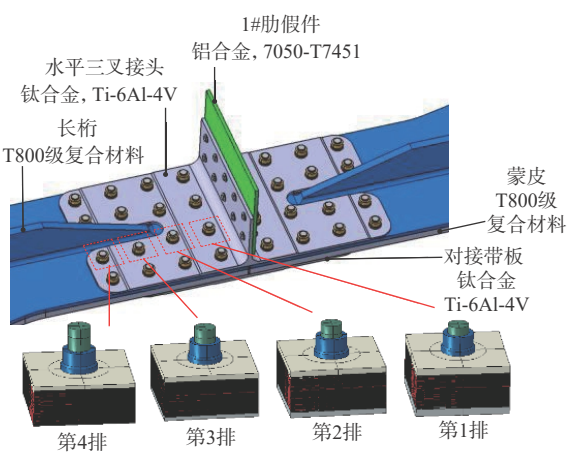


图 3　翼根长桁对接试验件子模型建模思路

Fig. 3　Sub-model modeling method of wing root stringer joint specimen

结构的每排紧固件，根据实际钉间距、钉边距和被连接件厚度，分别选取典型区域建立三维有限元模型，子模型在蒙皮面内尺寸约为 50 mm×50 mm。单个子模型包含螺栓、螺母、水平三叉接头、复合材料壁板、对接带板共 5 个零件，体现了单紧固件的连接关系和实际尺寸。

2 子模型计算分析

2.1 子模型建模

子模型采用 ABAQUS 商用软件建模，以第 1 排典型连接区域的子模型为例，建模细节如图 4 所示。各部件均采用 C3D8R 单元建模，材料属性如表 1 所示[14-15]。钉孔挤压区域网格细化，按约 0.2 mm 的网格密度建模，其他区域按 1.0 mm 的网格密度建模，共包含 46.9 万节点，44.3 万单元。

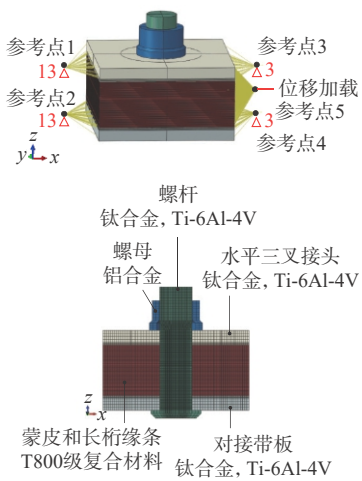


图 4 子模型示意图

Fig. 4 Schematic diagram of sub-model

根据图 1 中的受力分析可知，对于单个紧固件的连接区域，复合材料蒙皮和长桁缘条受力方向与水平三叉接头、对接带板相反。为尽可能真实地模拟单个紧固件连接区域的受力形式和边界条件，约束子模型中沿 y 方向对称面上所有节点的 2 方向自由度；在子模型左侧建立参考点 1、参考点 2，分别与水平三叉接头、对接带板的左侧截面以 coupling 绑定，并约束参考点的 13 自由度；在子模型右侧建立参考点 3、参考点 4，分别与水平三叉接头、对接带板的右侧截面以 coupling 绑定，并约束参考点的 3 自由度；建立参考点 5 与复合材料蒙皮和长桁缘条的右侧截面以 coupling 绑定，并在 1 方向按位移加载，监测参考点 5 的位移和对应的支反力。各零件之间建立接触，接触方式为硬接触，螺母与螺栓之间采用 tie 绑定，以便于各子模型的参数化建模。

为避免计算收敛性问题，子模型采用考虑几何非线性的 Dynamic, Explicit 分析步，并施加质量放大。按 0.5% 的时长间隔输出计算结果以监测子模型渐进损伤过程，并用于计算子模型非线性连接刚度。

表 1 有限元模型材料性能参数[14-15]

Table 1 Material property parameters of finite element model[14-15]

材料	计算参数	数值
复合材料	沿纤维方向弹性模量E_{11}/GPa	163.5
	垂直纤维方向弹性模量E_{22}/GPa	9.00
	面内剪切模量G_{12}/GPa	4.14
	面内泊松比v_{12}	0.319
	厚度t/mm	0.191
	单层0°拉伸强度/MPa	3 071.0
	单层0°压缩强度/MPa	1 747.0
	单层90°拉伸强度/MPa	88.0
	单层90°压缩强度/MPa	271.0
	单层面内剪切强度/MPa	131.0
Ti-6Al-4V	拉压平均弹性模量E/MPa	116 522
	剪切模量G/MPa	44 816
	泊松比v	0.31
	板材拉伸极限强度/MPa	931
	板材剪切极限强度/MPa	545
	紧固件拉伸极限强度/MPa	1 103
	紧固件剪切极限强度/MPa	655
7050-T7451	拉压平均弹性模量E/MPa	71 016
	剪切模量G/MPa	26 890
	泊松比v	0.33
	拉伸极限强度/MPa	510
	剪切极限强度/MPa	290

紧固件刚度的影响因素主要包括温度引起的热应力、螺栓预紧力、摩擦系数、钉孔间隙等。根据参考文献 [6]，连接结构在受载过程中包括“无滑动-滑动-接触”3 个阶段，螺栓预紧力和摩擦系数主要对“无滑动”阶段起主要影响；钉孔间隙对“滑动”阶段、钉孔接触面积有显著影响，进而影响螺栓连接刚度；温度热应力导致的变形对螺栓预紧力、钉孔间隙均产生影响。根据参考文献 [10]，小的钉孔干涉量、合适的预紧力可有效提高钉孔挤压强度。综合以上研究成果，可见随着螺栓载荷增大，连接结构进入螺栓与被连接件接触受力阶段后，以上因素对螺栓刚度、连接结构失效模式的影响较小，因此，暂不针对以上影响因素开展讨论，按常温、无预紧力及摩擦系数、钉孔无间隙/干涉开展

子模型建模。

为在子模型中分析连接结构的失效机理及非线性刚度特性, 考虑了金属材料的塑性变形、断裂及复合材料失效。钛合金、铝合金塑性性能来自MMPDS[15], 单向拉伸应力-应变曲线如图 5 所示。在 ABAQUS 中, 以数据列表的形式定义铝合金、钛合金材料的 Plastic 属性, 当塑性应变达到图 5 中的断裂应变时, 通过 Ductile Damage 判定材料的最终断裂。定义塑性数据时采用真实应力σ和真实应变ε。材料试验数据以名义应力σ_{nom}和名义应变$\varepsilon_{\mathrm{nom}}$的形式给出, 通过式(1)、式(2)转换为真实应力和真实应变。

$$\sigma=\sigma_{\mathrm{nom}}(1+\varepsilon_{\mathrm{nom}}) \tag{1}$$

$$\varepsilon=\ln(1+\varepsilon_{\mathrm{nom}}) \tag{2}$$

复合材料层内失效采用 Hashin-Rotem 失效准则, 该准则复合材料层板层内失效分为纤维失效和基体失效, 并采用单层应力来判断其是否失效。失效准则表达式如下。

纤维失效:

$$\begin{cases}\left(\dfrac{\sigma_{11}}{X_{\mathrm{t}}}\right)\geqslant 1.0 & \sigma_{11}>0\\ \left(\dfrac{\sigma_{11}}{X_{\mathrm{c}}}\right)\geqslant 1.0 & \sigma_{11}<0\end{cases} \tag{3}$$

基体失效:

$$\begin{cases}\left(\dfrac{\sigma_{22}}{Y_{\mathrm{t}}}\right)^2+\left(\dfrac{\sigma_{12}}{S_{12}}\right)^2\geqslant 1.0 & \sigma_{22}>0\\ \left(\dfrac{\sigma_{22}}{Y_{\mathrm{c}}}\right)^2+\left(\dfrac{\sigma_{12}}{S_{12}}\right)^2\geqslant 1.0 & \sigma_{22}<0\end{cases} \tag{4}$$

式中: X_{t}为与单层纤维体积相关的纤维拉伸强度; X_{c}为与单层纤维体积相关的纵向压缩强度; Y_{t}为单层横向就位拉伸强度; Y_{c}为单层横向就位压缩强度; S_{12}为单层面内就位剪切强度; σ_{ij}为单层平面内应力。

纤维或基体失效后, 分别将相关的材料参数衰减为原值的 1%, 以避免数值收敛性问题。通过编写用户自定义场变量子程序 VUSDFLD 实现复合材料损伤本构的定义, 具体损伤退化方案如表 2 所示。

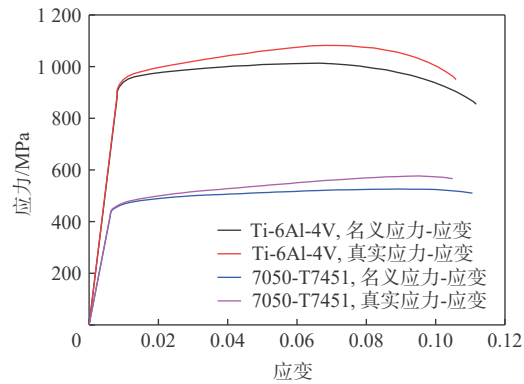


图 5　金属材料单向拉伸塑性曲线

Fig. 5　Unidirectional tensile plastic curves of metal materials

表 2　复合材料损伤后性能退化方案

Table 2　Performance degradation scheme of composite material after damage

失效模式	退化方案								
	E_{11}	E_{22}	E_{33}	G_{12}	G_{13}	G_{23}	v_{12}	v_{13}	v_{23}
完好	1	1	1	1	1	1	1	1	1
纤维失效	0.01	1	1	0.01	0.01	1	0.01	0.01	1
基体失效	1	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
纤维失效+基体失效	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

2.2　子模型计算结果

对 2.1 节建立的子模型求解计算直至子模型最终破坏, 提取图 4 中参考点 5 的加载位移及对应的支反力, 即可获得子模型的载荷-位移曲线。

第 1～4 排紧固件连接子模型加载点的载荷-位移曲线如图 6 所示。试验件复合材料壁板厚度不变, 第 1～3 排紧固件处为双剪连接, 随着水平三叉接头、对接带板的厚度逐渐变小, 子模型载荷-位移曲线斜率出现小幅下降; 第 4 排紧固件处为复合材料壁板与水平三叉接头的单剪连接, 其载荷-位移曲线斜率显著低于第 1～3 排。对于各连接子模

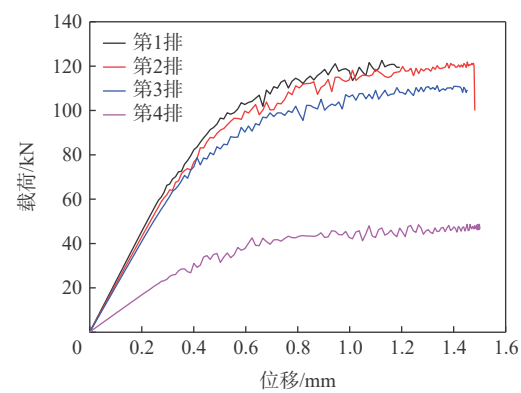


图 6　子模型加载点载荷-位移曲线

Fig. 6　Load-displacement curves of sub-model loading point

型，在初始阶段载荷随位移线性增大；而后随着复合材料损伤、金属塑性变形的出现，载荷-位移曲线出现了较长的非线性段。第1～4排连接子模型的破坏载荷分别为122.6 kN、121.8 kN、111.2 N、48.7 kN。

由于第1～3排连接关系较为相似，以第1排连接子模型为例，其剖视图的渐进损伤过程如图7所示，分别给出了65 kN、95 kN和最终破坏(122.6 kN)情况下的子模型应力、塑性应变、复合材料纤维失效、复合材料基体失效云图。由于第1排连接处水平三叉接头、对接带板厚度相同(5 mm)，为对称的双剪连接，应力集中主要出现在复合材料壁板与水平三叉接头/对接带板的上、下连接界面区域，金属塑性变形、复合材料损伤也首先在应力集中区出现。随着载荷的增大，金属塑性、复合材料损伤区域逐渐增大，螺栓在应力集中区及中部受弯区域也出现塑性变形；最终螺栓在复合材料壁板-对接带板连接界面处断裂，此时复合材料孔边沿厚度方向均出现明显损伤。

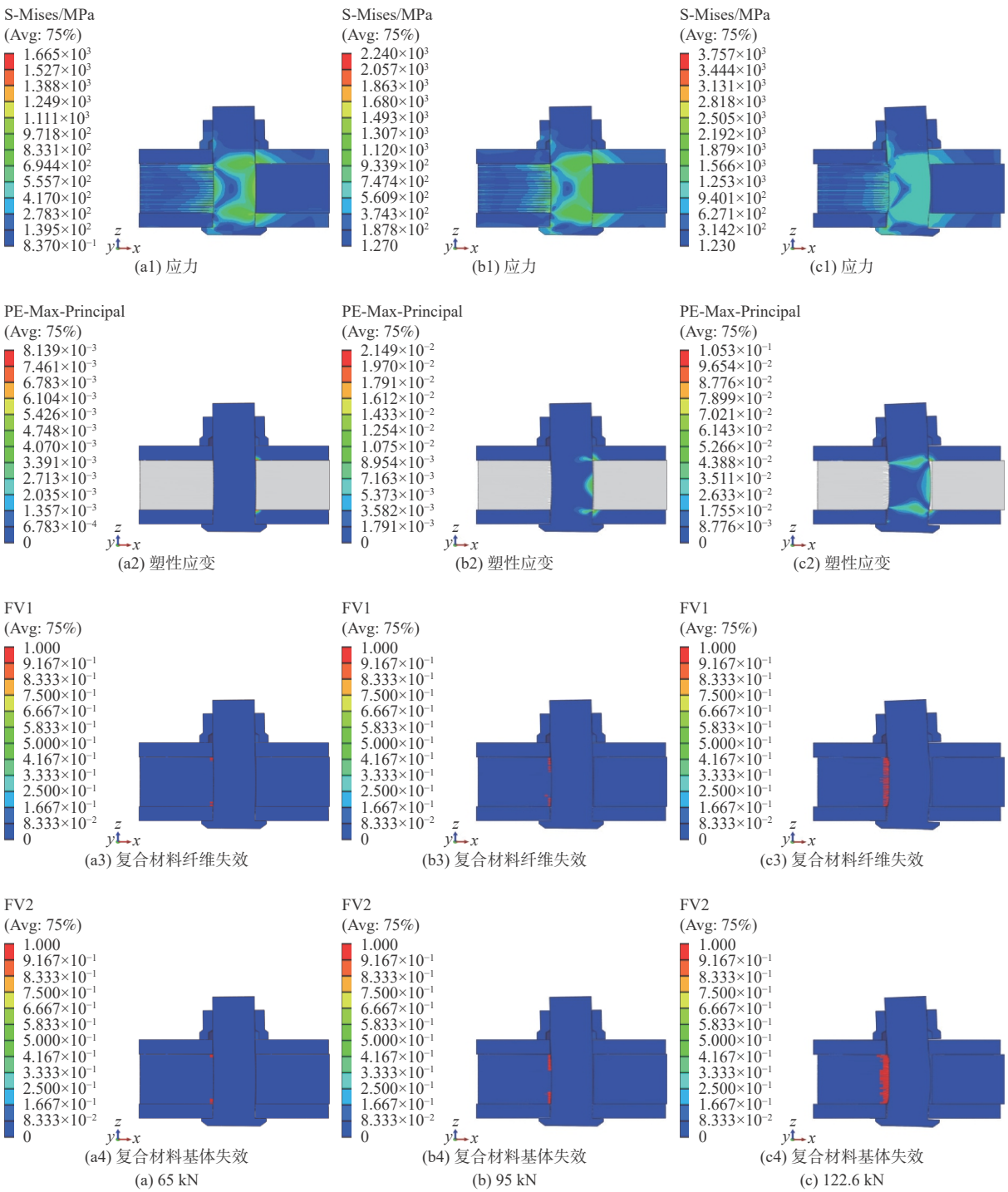


图7 第1排连接子模型渐进损伤过程

Fig. 7 Progressive damage of the 1st row joint sub-model

第 4 排连接子模型剖视图的渐进损伤过程如图 8 所示。此处为复合材料壁板与水平三叉接头的单剪连接, 应力集中主要出现在复合材料壁板与水平三叉接头连接区域。复合材料纤维失效、基体失效起始于应力集中区, 且水平三叉接头与螺栓接触位置出现较为明显的塑性变形。随着加载位移的增大, 复合材料损伤区域、水平三叉接头塑性区不断扩大, 直至子模型载荷不随位移继续增加, 达到最大承载能力。

2.3　孔边金属塑性变形与复合材料损伤的影响分析

以第 1 排连接为例, 分别建立仅考虑被连接件金属塑性变形及最终断裂、仅考虑被连接件复合材料损伤的子模型, 用于分析金属塑性、复合材料损伤对子模型连接刚度及最终承载能力的影响。各子模型分析得到的载荷-位移曲线如图 9 所示。相比原模型, 仅考虑金属塑性变形模型、仅考虑复合材料损伤模型的载荷-位移曲线线性段刚度相同,

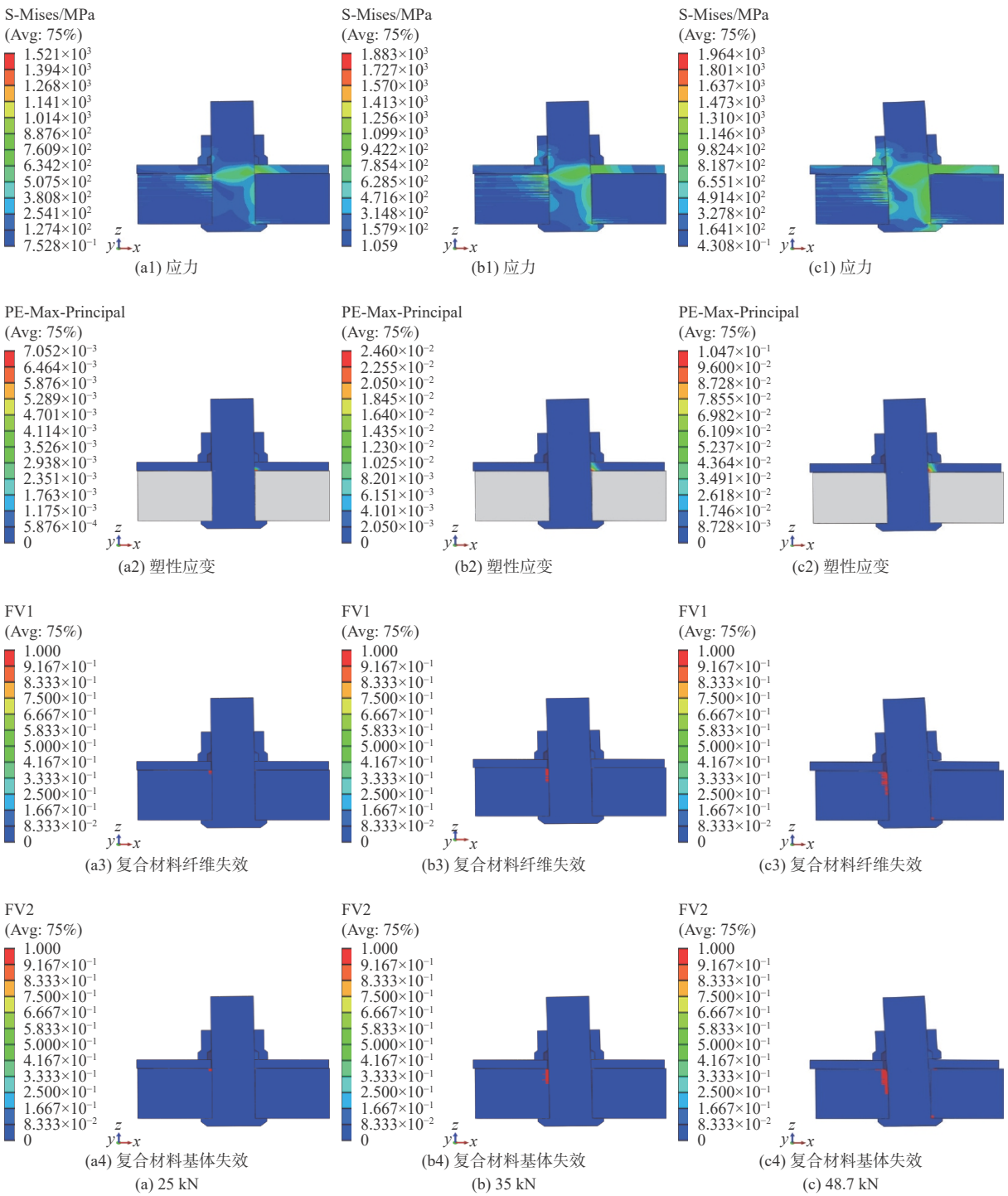


图 8　第 4 排连接子模型渐进损伤过程

Fig. 8　Progressive damage of the 4th row joint sub-model

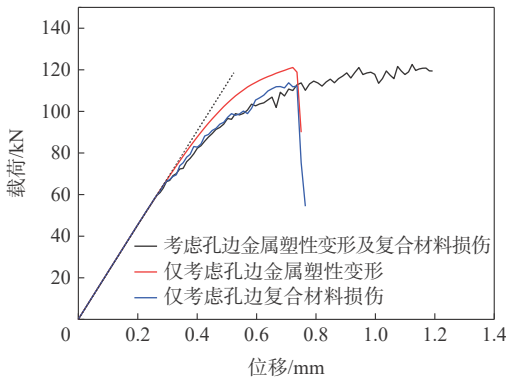
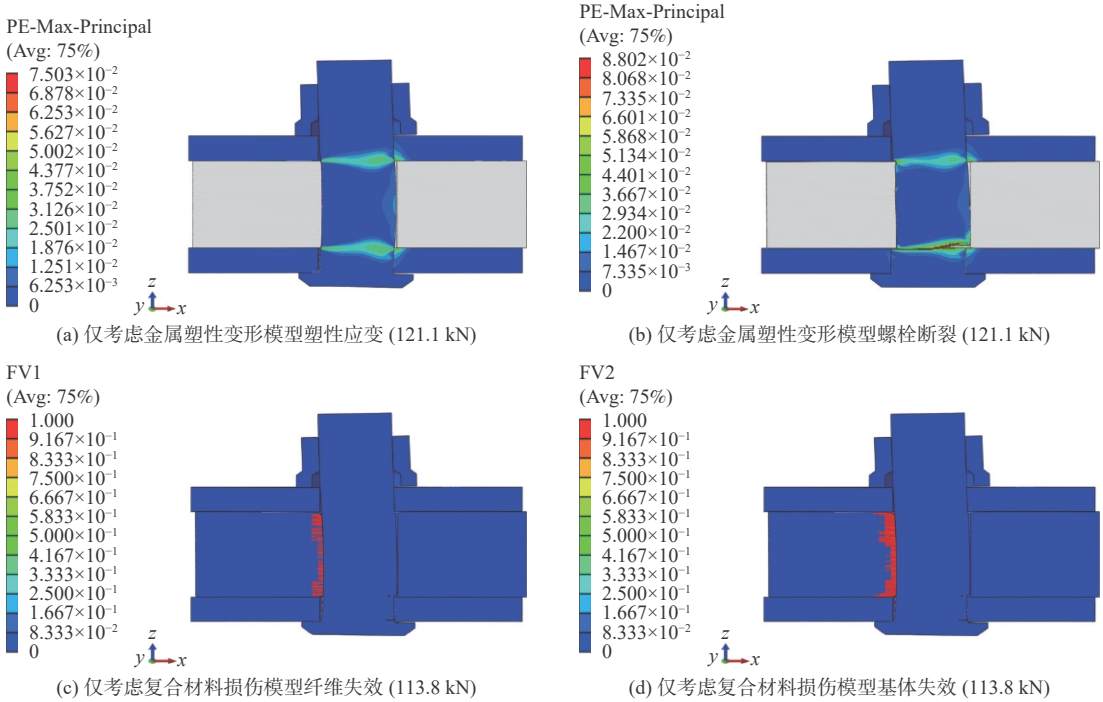


图9 第1排子模型考虑不同损伤类型的载荷-位移曲线

Fig. 9 Load-displacement curves of the 1st row sub-model considering different damage types

且基本在相同载荷(约65 kN)下刚度出现非线性;进入非线性段后,仅考虑金属塑性变形、仅考虑复合材料损伤模型的刚度高于原模型,最大承载能力分别为原子模型的98.8%、92.8%。

考虑不同损伤类型子模型的破坏形式如图10所示。对于仅考虑金属塑性变形模型,复合材料孔边没有损伤引起的“软化”,螺栓在复合材料壁板与水平三叉接头、对接带板的连接界面处存在明显应力集中并发生显著的塑性变形,最终出现螺栓剪切断裂。对于仅考虑复合材料损伤模型,复合材料孔边出现大范围的纤维失效和基体失效,最终螺栓在复合材料壁板与对接带板的连接界面处断裂失效。由此可见,对于当前连接构型,金属塑性变形、复合材料损伤均会引起孔边明显“软化”,共同决定了载荷-位移曲线的非线性特性;连接结构最终破坏由螺栓剪切断裂驱动。



(a) 仅考虑金属塑性变形模型塑性应变 (121.1 kN)

(b) 仅考虑金属塑性变形模型螺栓断裂 (121.1 kN)

(c) 仅考虑复合材料损伤模型纤维失效 (113.8 kN)

(d) 仅考虑复合材料损伤模型基体失效 (113.8 kN)

图10 第1排连接考虑不同损伤类型子模型的破坏形式

Fig. 10 Failure modes of the 1st row sub-model considering different damage types

2.4 基于子模型的连接刚度定义

根据子模型分析结果,计算得出各连接件之间的非线性连接刚度,为整体模型中连接单元刚度定义提供输入。以第1排连接为例,子模型非线性连接刚度计算如图11所示。参考复合材料连接结构挤压响应试验方法[16],各被连接件之间的相对变形主要由孔边挤压变形导致,被连接件本身的变形为小量,因此,参考点的位置选取应捕捉每2个被连接件之间的钉孔挤压变形。图11中,在紧固件左侧的蒙皮和长桁缘条沿厚度方向中间位置定义点a,在紧固件右侧的水平三叉接头处定义点b,点a、b之间在迭代步i沿x方向的相对变形量记为δ_{1i};在紧固件右侧的对接带板处定义点c,点a、c之间在迭代步i沿x方向的相对变形量记为δ_{2i}。在蒙皮和长桁缘条-水平三叉接头、蒙皮和长桁缘条-对接带板的连接界面处分别提取螺栓截面的Free-body剪切载荷,分别记为$F_{\text{bolt1-}i}$、$F_{\text{bolt-2}i}$。迭代步i的连接刚度按式(5)计算:

$$\begin{cases} D_{1i} = F_{\text{bolt-1}i}/\delta_{1i} \\ D_{2i} = F_{\text{bolt-2}i}/\delta_{2i} \end{cases} \quad i = 1,2,\cdots \tag{5}$$

式中:D_{1i}为迭代步i下蒙皮和长桁缘条-水平三叉接

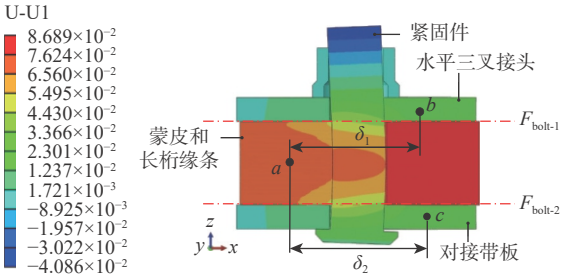


图 11　子模型连接刚度计算示意图

Fig. 11　Schematic diagram of sub-model joint stiffness calculation

头之间的连接刚度, N/mm; D_{2i}为迭代步i下蒙皮和长桁缘条-对接带板之间的连接刚度, N/mm。

3　整体模型计算分析

3.1　整体有限元模型

翼根下壁板长桁对接整体有限元模型如图 12 所示。壁板、水平三叉接头、对接带板采用 S4R 单元模拟, 并赋予对应的材料属性。紧固件连接采用 Cartesian+Cardan 连接模拟, 共包含 6 个自由度, 分别为 K_1～K_6。轴向刚度 K_1为线性刚度, 定义为 $K_1 = E_{bolt}A_{bolt}/L$, 其中, E_{bolt}和A_{bolt}分别为紧固件的弹性模量和截面积, L为被连接件中面之间的距离; K_2为沿试验件加载方向的面内剪切刚度, 按照 2.4 节子模型计算结果, 在 ABAQUS 的 Connector Section 中以数据列表的形式定义其非线性连接刚度; K_3为垂直于试验件加载方向的面内剪切刚度, 对试验件的整体传力影响较小, 按 Huth 公式[17]定义线性刚度值:

$$\frac{1}{K_3}=\left(\frac{t_1+t_2}{2d_{bolt}}\right)^{m_1}\frac{m_2}{n}\left(\frac{1}{t_1E_1}+\frac{1}{nt_2E_2}+\frac{1}{2t_1E_{bolt}}+\frac{1}{2nt_2E_{bolt}}\right)\quad(6)$$

式中: t_1、t_2为被连接件厚度; d_{bolt}为紧固件直径; m_1、m_2为钉的连接类型系数, 对于复合材料结构, m_1=2/3, m_2=4.2; n与钉的剪切形式相关, 单剪时, n=1, 双剪时, n=2; E_1、E_2为层板垂直于试验加载方向的面内等效弹性模量。

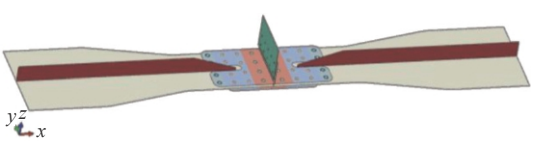


图 12　整体有限元模型示意图

Fig. 12　Schematic diagram of global finite element model

K_4为绕紧固件轴线方向的转动刚度, K_5、K_6分别为沿试验件加载方向、垂直试验件加载方向的转动刚度。为避免连接单元面内剪切与转动的耦合作用并参考已有经验[5], 限制连接单元的转动变形, 将K_4～K_6均定义为 10^9 N/mm。

各零件间定义 surface to surface 的接触形式, 并指定无摩擦的硬接触。模型左侧通过参考点绑定端部并施加固支边界条件; 右侧通过参考点绑定端部, 约束除 x 方向以外的所有自由度, 并在 x 方向施加载荷; 1#肋假件端部约束 z 方向变形。定义几何非线性静态分析步开展计算。

3.2　破坏强度预测

对于典型结构件级机械连接结构, 其失效模式包括紧固件剪切、被连接件挤压、被连接件端部剪切、复合材料结构挤压-旁路耦合、紧固件拉脱等[2]。实际民机结构设计中, 通过钉边距、钉间距的设计原则, 基本避免了被连接件端部剪切、复合材料挤压-旁路耦合等失效模式的出现; 对于其他失效模式, 根据子模型分析结果, 试验件最终破坏形式为紧固件剪切断裂, 因此, 主要针对此失效模式开展分析讨论。

基于子模型计算得到的各排连接结构非线性连接刚度, 采用整体有限元模型计算得到的 1～4 排紧固件钉载如图 13 所示。在加载初始阶段, 各排紧固件钉载保持线性增加, 且第 1 排连接钉载高于 2～4 排; 在加载到约 82% 预测破坏载荷时, 第 1 排紧固件钉载出现明显的非线性降低, 同时第 2、3 排钉载增大。第 4 排紧固件由于刚度较弱, 在加载过程中保持了较好的线性。最终第 1～3 排紧固件在较为接近的钉载下达到紧固件极限剪切强度(即紧固件断裂失效), 导致试验件破坏。

为与非线性连接刚度计算结果进行对比, 采用子模型线性段连接刚度开展整体有限元模型计算, 钉载计算结果如图 14 所示。对于线性连接刚度模型, 各排紧固件钉载随载荷的增大线性增加, 第 1 排紧固件(复合材料壁板-水平三叉接头、复合材

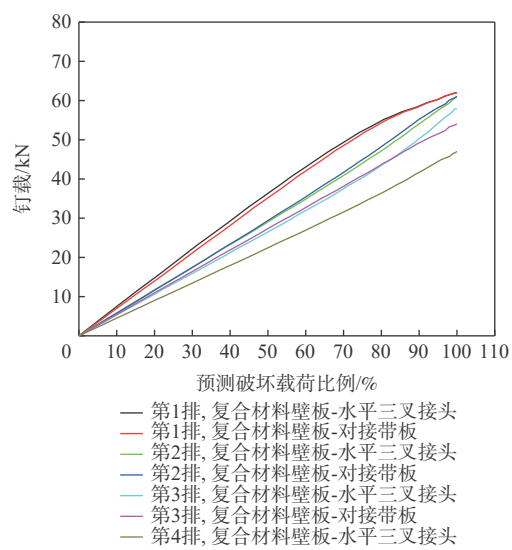


图 13　非线性连接刚度钉载计算结果

Fig. 13　Bolt load calculation results using non-linear joint stiffness

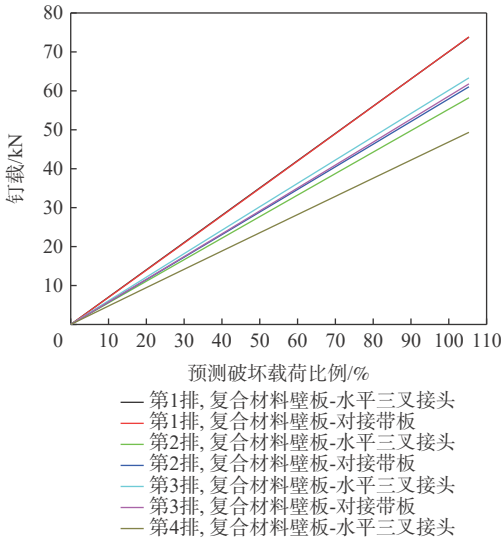


图 14 线性连接刚度钉载计算结果

Fig. 14 Bolt load calculation results using linear joint stiffness

料壁板-对接带板)钉载显著高于第 2～4 排,且在约 87.5% 预测破坏载荷时,第 1 排紧固件钉载达到紧固件极限剪切强度,导致试验件破坏。

4 算例验证

4.1 试验加载及结果

试验在 600 t MTS 试验机上进行,图 15 为试验加载示意图。试验件两端的蒙皮、长桁与夹具通过紧固件连接,保证加载中心与试验件端部刚心重合;试验件中部肋假件与侧向支持夹具连接,侧向支持夹具可上下自由滑动,以模拟实际飞机结构中 1#肋对翼根下壁板连接结构的边界支撑条件。

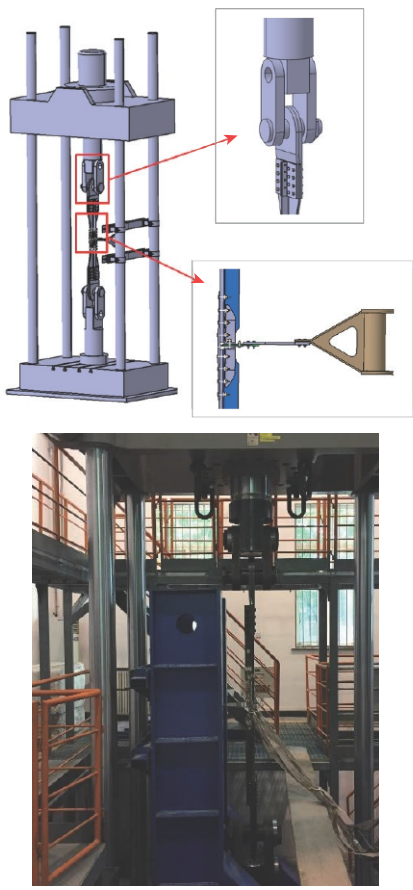


图 15 试验加载示意图

Fig. 15 Schematic diagram of test loading

最终加载到 106.8% 预测破坏载荷后试验件失效,对应的试验件远端平均应变约为 4 150 με。现场观测到的试验件破坏形貌如图 16 所示。试验件上侧(加载侧)的 4 排共 16 颗高锁螺栓均发生断裂,复合材料壁板孔边出现了显著的挤压变形,水平三叉接头、对接带板在第 3、4 排孔边出现较为明显的塑性变形,第 1、2 排孔边塑性变形不明显。试验结果表明,试验件在最终破坏前紧固件连接发生了明显的非线性变形,并最终导致多排紧固件同时破坏。

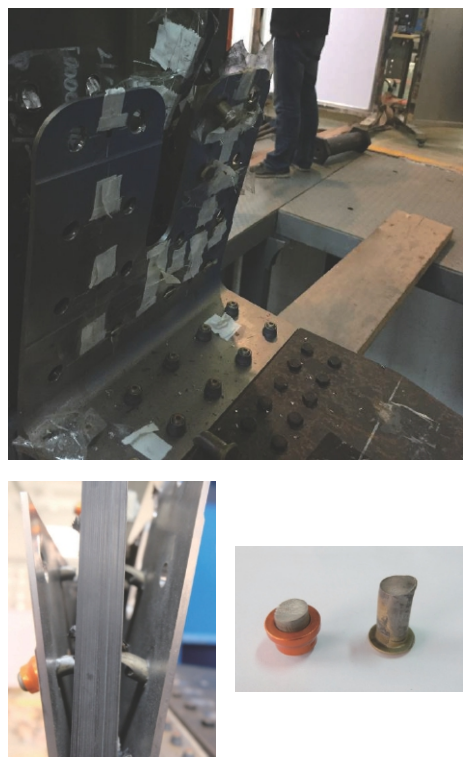


图 16 试验件破坏模式

Fig. 16 Failure mode of test specimen

4.2 应变结果对比

试验件应变布置方案如图 17 所示。采用非线性连接刚度模型计算得到的应变与试验应变对比如图 18～图 20 所示。对于试验件远端应变及加载方向对称面处应变,随着载荷加载基本保持了线性,计算应变与试验应变具有较好的一致性。可见,钉载分配的非线性变化并未影响试验件远端的总载荷,也未显著改变水平三叉接头、对接带板之间的传载比例。

对于第 1、2 排紧固件之间的应变,试验应变在

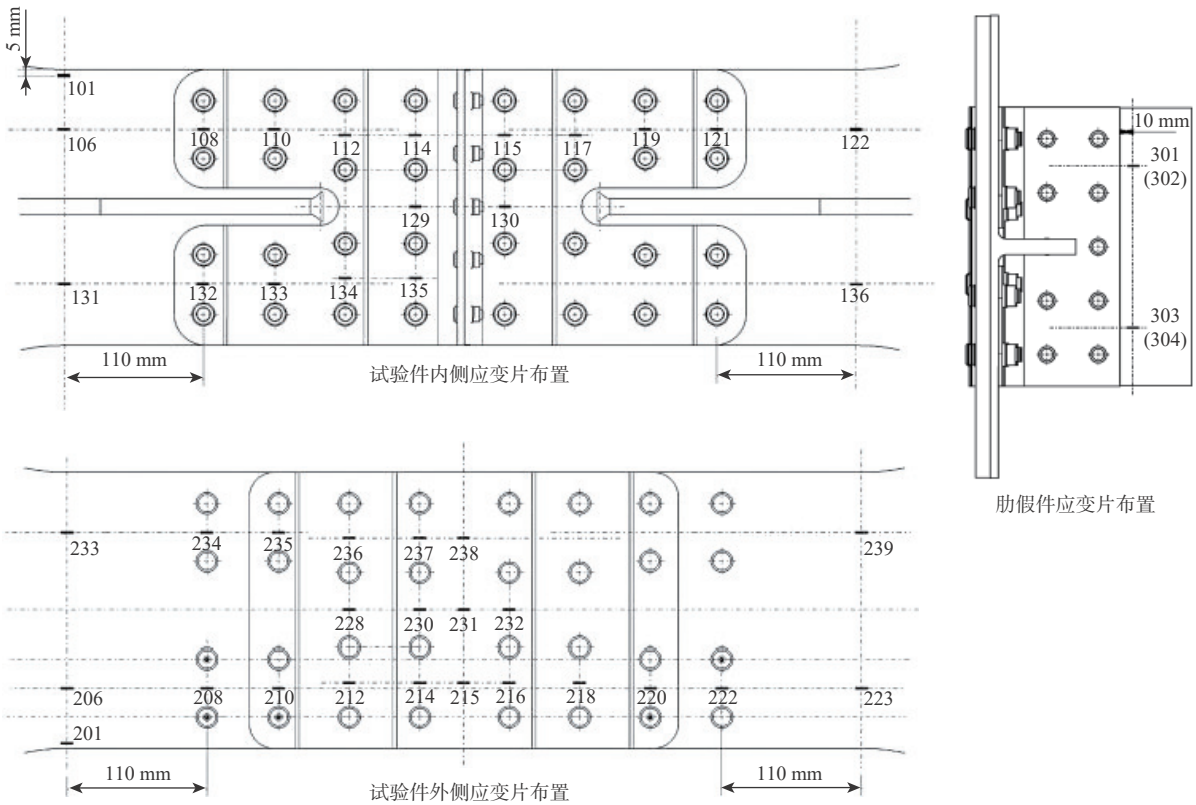


图 17　试验件应变片布置

Fig. 17　Specimen strain gages arrangement

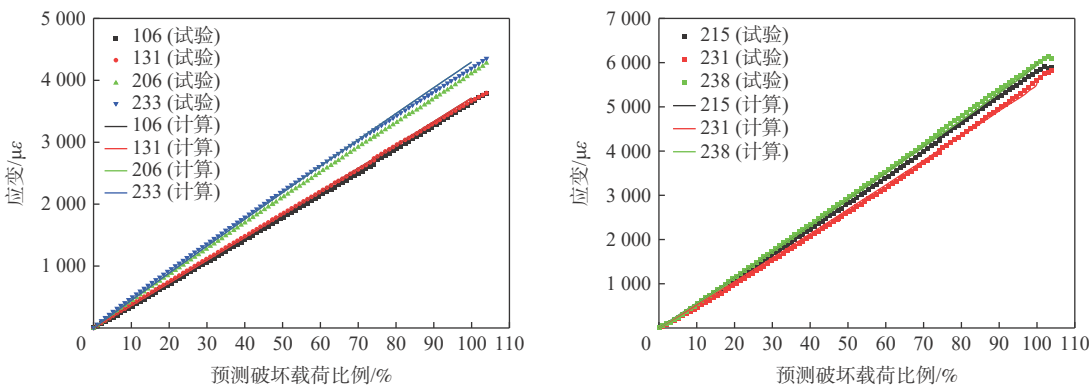


图 18　试验件远端和对称面处应变计算结果与试验结果对比

Fig. 18　Comparison of calculated strain results with experimental results at location of far field and symmetry plane of specimen

约 71.3% 预测破坏载荷时出现了较明显的非线性增大, 计算应变也出现了非线性变化, 但与试验应变存在一定的差异。认为第 1、2 排连接处应变的非线性增大由 2 个方面因素导致: ①由于紧固件钉载出现非线性, 第 1、2 排连接钉载相比线性模型减小, 导致其旁路载荷增大, 因此, 应变值增大; ②水平三叉接头、对接带板出现了局部塑性变形。

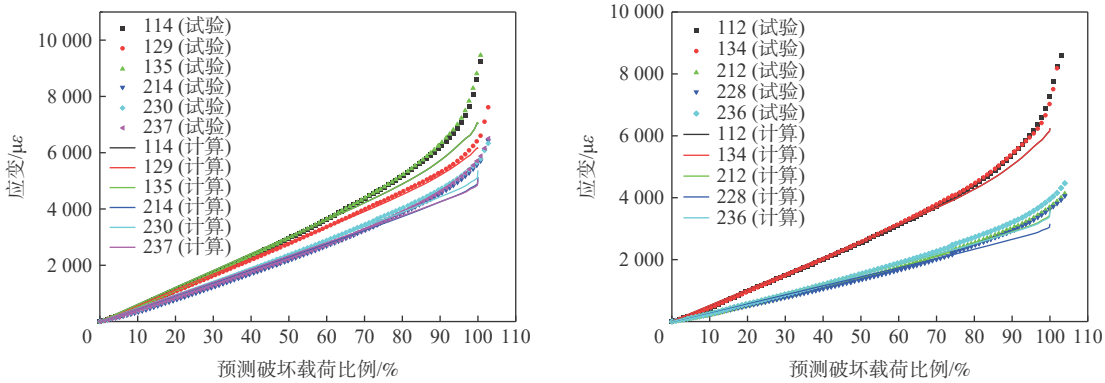
第 3、4 排连接试验应变的非线性相比第 1、2 排有所减弱, 原因是此处连接旁路载荷变小, 且金属件应力水平降低, 因此, 未发生显著的塑性变形。

整体来说, 计算应变与试验应变在线性段一致性较高, 表明模型已较准确地模拟了试验件的传载特性; 由于子模型中已经考虑金属塑性变形对紧固件刚度影响, 整体有限元模型中未重复考虑金属塑性, 导致计算应变与试验应变在非线性段存在一定的偏差。

4.3　结果分析

结合子模型分析结果及破坏强度预测, 整体有限元模型考虑非线性连接刚度后, 预测破坏载荷为试验破坏载荷的 93.6%, 预测破坏模式为多排紧固件同时断裂, 与试验破坏模式一致, 分析应变与试验应变一致性较好, 验证了基于子模型的连接结构分析方法的有效性。

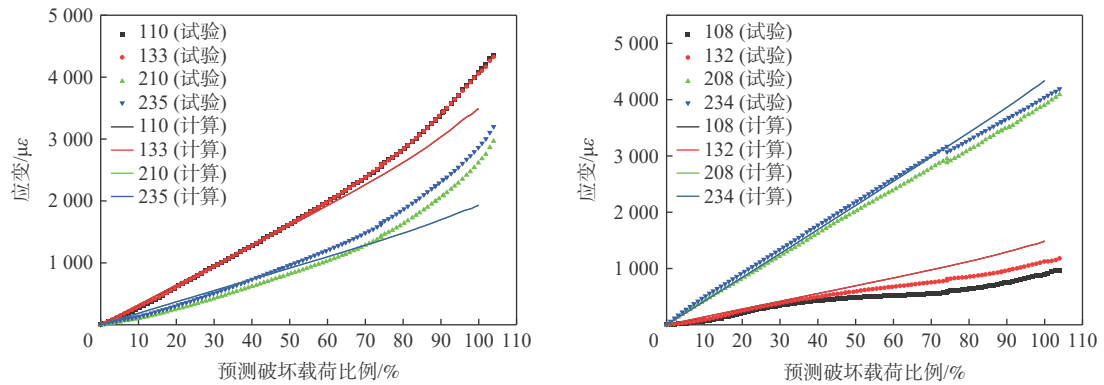
与之相比, 整体有限元模型采用线性连接刚



(a) 第1排连接 (b) 第2排连接

图 19 试验件第 1 排和第 2 排连接处应变计算结果与试验结果对比

Fig. 19 Comparison of calculated strain results with experimental results at location of the 1st row joint and the 2nd row joint of specimen



(a) 第3排连接 (b) 第4排连接

图 20 试验件第 3 排和第 4 排连接处应变计算结果与试验结果对比

Fig. 20 Comparison of calculated strain results with experimental results at location of the 3rd row joint and the 4th row joint of specimen

度，预测破坏载荷为试验破坏载荷的 81.9%，预测破坏模式为第 1 排紧固件断裂，与试验结果存在较大的偏差。试验监测应变在加载过程中出现了明显的非线性特征，不符合线性分析假设。

当前民机复合材料连接结构设计中，通常采用线性连接刚度模型计算钉载分配，并采用孔边挤压强度（考虑分散性、环境影响等）判定连接结构失效，往往得到较为保守的结果，导致连接设计强度远低于结构试验破坏强度。结合以上计算结果可知，在翼根长桁对接结构分析模型中考虑非线性钉载分配，相比线性模型预测破坏强度提高了 14.3%。随着民机复合材料连接结构应用实践和信心的积累及复合材料性能的进一步稳定，该分析方法应用于民机复合材料连接结构设计，可有效实现结构减重。

5 结 论

1）以民机复合材料机翼翼根下壁板长桁对接结构为研究对象，提出了一种基于子模型的复合材料混杂连接结构非线性钉载分配及失效机理分析方法，用于预测结构的破坏载荷及失效模式，并采用试验进行了验证。预测破坏载荷为试验破坏载荷的 93.6%；预测破坏模式为单侧多排紧固件同时剪切断裂，与试验破坏模式一致，验证了方法的有效性。

2）典型连接结构的子模型分析表明，金属塑性变形、复合材料损伤均会引起孔边明显“软化”，共同决定了子模型载荷-位移曲线的非线性特性。

3）整体有限元模型考虑非线性连接刚度后，第 1～3 排紧固件的钉载在约 82% 预测破坏载荷后趋近一致，最终在较为接近的钉载下达到紧固件极限剪切强度导致试验件破坏。采用线性连接刚度模型预测的破坏载荷为试验破坏载荷的 81.9%，预测失效模式为第 1 排紧固件断裂，预测载荷、破坏模式均与试验结果存在较大差异。

参考文献（References）

[1] 杜善义, 关志东. 我国大型客机先进复合材料技术应对策略思考[J]. 复合材料学报, 2008, 25(1): 1-10.

DU S Y, GUAN Z D. Strategic considerations for development of advanced composite technology for large commercial aircraft in China[J]. Acta Materiae Compositae Sinica, 2008, 25(1): 1-10(in Chinese).

[2] 谢鸣九. 复合材料连接技术[M]. 上海: 上海交通大学出版社, 2016: 129-235.
XIE M J. Joints for composites materials[M]. Shanghai: Shanghai Jiao Tong University Press, 2016: 129-235(in Chinese).

[3] HART-SMITH L J. Bonded-bolted composite joints[J]. Journal of Aircraft, 1985, 22(11): 993-1000.

[4] GRAY P J, MCCARTHY C T. A highly efficient user-defined finite element for load distribution analysis of large-scale bolted composite structures[J]. Composites Science and Technology, 2011, 71(12): 1517-1527.

[5] ZHANG F, HU Z D, GAO L M, et al. Investigation on in-plane shear behavior of large-size composite plates with multi-bolt joints [J]. Composite Structures, 2020, 232: 111553.

[6] SASIKUMAR A, GUERRERO J M, QUINTANAS-COROMINAS A, et al. Numerical study to understand thermo-mechanical effects on a composite-aluminium hybrid bolted joint[J]. Composite Structures, 2021, 275: 114396.

[7] 李玺, 李亚智, 磨承杰, 等. 基于连续损伤力学的复合材料多钉连接失效分析和试验研究[J]. 机械强度, 2022, 44(4): 911-920.
LI X, LI Y Z, MO C J, et al. Failure analysis and experimental study of composite multi-pin connections based on continuum damage mechanics[J]. Journal of Mechanical Strength, 2022, 44(4): 911-920 (in Chinese).

[8] 黄河源, 赵美英, 万小朋, 等. 一种复合材料螺栓连接结构非线性刚度模型及应用[J]. 西北工业大学学报, 2018, 36(1): 66-73.
HUANG H Y, ZHAO M Y, WAN X P, et al. A composite bolted joints non-linear stiffness model and its application[J]. Journal of Northwestern Polytechnical University, 2018, 36(1): 66-73(in Chinese).

[9] 蔡启阳, 赵琪. 环境温度和间隙对复合材料-金属混合结构机械连接钉载分配的影响[J]. 复合材料学报, 2021, 38(12): 4228-4238.
CAI Q Y, ZHAO Q. Effects of temperature and clearance fit on the load distribution of composite-metal hybrid structures[J]. Acta Materiae Compositae Sinica, 2021, 38(12): 4228-4238(in Chinese).

[10] 张岐良, 曹增强. 复合材料螺接性能的影响因素研究[J]. 航空学报, 2012, 33(4): 755-762.
ZHANG Q L, CAO Z Q. Study on factors influencing the performance of composite bolted connections[J]. Acta Aeronautica et Astronautica Sinica, 2012, 33(4): 755-762(in Chinese).

[11] 雷凯, 王彬文, 成竹, 等. 飞机多钉壁板混合结构热应力分析与验证[J]. 应用力学学报, 2023, 40(1): 40-47.
LEI K, WANG B W, CHENG Z, et al. Thermal stress analysis and verification of aircraft multi-bolt hybrid panel structure[J]. Chinese Journal of Applied Mechanics, 2023, 40(1): 40-47(in Chinese).

[12] 贾宝惠, 方嘉晨, 武涛, 等. 预紧力影响下的复合材料螺接连接件失效安全性分析[J]. 机械科学与技术, 2025, 44(7): 1290-1300.
JIA B H, FANG J C, WU T, et al. Failure safety analysis of composite screw connection under the influence of pre-tightening force [J]. Mechanical Science and Technology for Aerospace Engineering, 2025, 44(7): 1290-1300(in Chinese).

[13] 汤平, 李星. 插入式机翼下壁板对接附加弯矩研究[J]. 航空学报, 2019, 40(2): 522436.
TANG P, LI X. Additional bending moment in lower panel butt-joint of inserted wing[J]. Acta Aeronautica et Astronautica Sinica, 2019, 40(2): 522436(in Chinese).

[14] 高阳. 某型飞机中央翼关键连接区混杂结构试验仿真与分析[D]. 南京: 南京航空航天大学, 2020: 19-20.
GAO Y. Experimental simulation and analysis of hybrid structure in key connection area of central wing of an aircraft[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2020: 19-20(in Chinese).

[15] RICE R C, JACKSON J, BAKUCKAS J, et al. Metallic materials properties development and standardization (MMPDS)[M]. Washington, D. C. : U. S. Department of Transportation Federal Aviation Administration, 2003.

[16] ASTM. Standard test method for bearing response of polymer matrix composite laminates: ASTM D5961/D5961M-23[S]. Philadelphia: ASTM International, 2017: 1-31.

[17] HUTH H. Influence of fastener flexibility on the prediction of load transfer and fatigue life for multiple-row joints[C]//Proceedings of the Fatigue in Mechanically Fastened Composite and Metallic Joints. Philadelphia: ASTM, 1986: 221-250.

Sub-model based failure analysis of composite primary load-bearing hybrid joint structure

LI Xing[1], LI Zengshan[2,*], DENG Fanchen[3], NIE Lei[1], ZHANG Tian[1]

(1. Beijing Key Laboratory of Commercial Aircraft Structures and Composite Materials, COMAC Bejing Aircraft Technology Research Institute, Beijing 102211, China;
2. School of Aeronautics Science and Engineering, Beihang University, Beijing 100191, China;
3. Composite Structure and Strength Laboratory, Aircraft Strength Research Institute of China, Xi'an 710065, China)

Abstract: Taking the lower panel stringer joint structure in the composite wing root of civil aircraft as the research object, a three-dimensional sub-model of a single-bolt joint considering material damage was established, and the joint failure mechanism and the nonlinear load-deformation relationship were analyzed using the sub-model. After calculating the nonlinear fastener load distribution, the global finite element model was used to estimate the failure load while taking the nonlinear joint stiffness into account. By using the above method, the problems such as low computational efficiency and poor convergence caused by direct establishment of a full-size 3D model and considering damage are solved. The test confirms the efficacy of the suggested approach since the results indicate that the predicted failure mode is the simultaneous shear fracture of single-side multi-row fasteners, which is consistent with the test failure mode, and the predicted failure load is 93.6% of the test failure load. In contrast, the failure load predicted by the model with linear joint stiffness is 81.9% of the test failure load, and the predicted failure mode is the failure of the first row fasteners, which is obviously different from the test results.

Keywords: composite material; hybrid joint; bolt load distribution; sub-model; failure analysis

Received: 2023-08-29; **Accepted**: 2023-11-10; **Published Online**: 2023-12-13 17:20

URL: link.cnki.net/urlid/11.2625.V.20231213.1442.001

Foundation item: National Key Research and Development Program of China (2022YFB4301002)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0572

点阵结构散热特性及其轻量化评价

胡建军[1,2,*]，窦若尘[2]，张欣[1]，姚静[2]，孔祥东[2]
(1. 燕山大学 建筑工程与力学学院，秦皇岛 066044；2. 燕山大学 机械工程学院，秦皇岛 066044)

摘 要：航空电静液驱动执行器(EHA)作为航空舵面控制的核心部件，对高功重比的追求要求对其部件进行高效散热与轻量化综合设计。基于此，通过实验方法对4种典型轻量化点阵结构的散热特性进行研究，得到了不同点阵散热模块的温度分布特性，为点阵结构散热与轻量化性能评价提供可靠数据。针对点阵结构散热特性，提出点阵结构散热轻量化系数评价指标，用以定量评估不同散热点阵结构的轻量化特性。根据构建的评价指标对4种点阵结构进行评价，结果表明：OT点阵结构散热性能与轻量化综合特性最优。与翅片结构相比，OT点阵结构散热系数为翅片的1.2倍，但其散热轻量化系数仅约为翅片的1/3。研究结果为综合考虑散热和轻量化要求的点阵结构筛取提供了参考。

关 键 词：点阵结构；散热；轻量化；实验研究；评价
中图分类号：TH122; TK124
文献标志码：A **文章编号**：1001-5965(2025)10-3299-08

航空电静液驱动执行器(electro-hydrostatic actuator, EHA)将机械、电气、液压、测试和控制等高度集成，作为未来大型客机和先进战机飞行舵面控制的核心部件，是实现其智能驱动的关键，高功重比是衡量航空EHA的关键技术指标之一[1-6]。因此，亟需探讨航空EHA轻量化设计及制造新原理、新方法和新技术。同时，航空EHA工作于高压、高速、高频、重载工况，易引起系统温度升高，造成航空EHA动态性能恶化、可靠性降低、寿命缩短。因此，研究高效散热结构也成为亟需解决的关键问题。

三维点阵结构因具有轻质和散热优[7-10]等功能特性一直受到广泛的关注。针对点阵结构散热性能，Khaderi等[11]对Gyroid点阵进行了等截面模型简化，并计算了其等效导热系数；Shen等[12]通过数值模拟表明，用点阵结构代替传统的翅片，具有更好的散热能力和相似的压降；Son等[13]对增材制造(additive manufacturing, AM)制作的四面体晶格通道进行了实验和数值研究，通过对摩擦系数和努塞尔数的实验测量，验证了该方法的有效性，并研究了孔隙度对性能的影响；Liang等[14]通过实验对体心立方(body-centered cubic, BCC)点阵结构和Kagome点阵结构的流体与端壁之间的传热进行了分析，实验结果证明了这2种点阵结构的端壁换热性能优于针形翅片结构；Takezawa[15]和Vaissier[16]等分别通过理论推导出点阵具有高效各向同性导热系数的最佳点阵结构胞元和一种最大化点阵结构比表面积设计方法，以此提高点阵结构散热性能。

此外，针对点阵结构的轻量化特性，Cheng[17]和张龙[18]等均采用拓扑优化方法对点阵结构进行轻量化设计；Valdevit等[19]基于力学和导热性能对多种点阵结构进行优化设计，得到六边形结构的综合性能最优；Moon等[20]以承载能力和轻量化为目标，对比了Kagome、金字塔和钻石型点阵，指出Kagome型点阵性能最优，并将其应用在无人机机翼；Xu等[21]以轻量化和提高隔音性能为目标，采用遗传算法对金字塔点阵进行了优化，并研究了材

收稿日期：2023-09-06；**录用日期**：2024-02-23；**网络出版时间**：2024-04-04 07:23
网络出版地址：link.cnki.net/urlid/11.2625.V.20240403.0916.001
基金项目：国家自然科学基金重大项目(51890881)；河北省自然科学基金(2020203028)；中央引导地方科技发展资金项目(226Z1902G)
*通信作者. E-mail：kewei729@163.com

引用格式：胡建军，窦若尘，张欣，等. 点阵结构散热特性及其轻量化评价[J]. 北京航空航天大学学报，2025，51(10)：3299-3306.
HU J J，DOU R C，ZHANG X，et al. Heat dissipation characteristics and lightweight evaluation of lattice structure[J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51(10)：3299-3306 (in Chinese).

料、上下面板厚度对性能的影响。

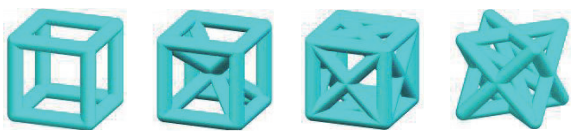
综上所述，目前基于简单点阵结构的散热理论分析与点阵结构设计较多，但将点阵结构散热性能与轻量化特性相结合，并对轻量化特性量化分析的研究很少。因此，本文通过实验对几种点阵结构的散热性能进行探究，并提出散热轻量化系数为点阵结构的筛选与应用提供理论依据。

1 点阵设计及加工

1.1 点阵结构设计

三维点阵结构的几何特征具有复杂性，其几何特征可由杆径、胞元长度、胞元宽度、胞元高度及杆件倾角表示。本文重点研究图 1 所示的 4 种典型正六面体空间内由若干杆与节点组成的点阵结构，故其几何特征主要为杆径与胞元长度。

点阵散热结构建模方法分为基本单元法和拓扑优化法，采用基本单元法对以上 4 种点阵胞元进行建模。如图 2 所示，以立方体点阵散热模块模型为例，模型由若干点阵胞元和基板 2 部分组成，为排除边界散热的影响，基板尺寸选为 100 mm×100 mm×2 mm。



(a) 立方体点阵 (b) 体心点阵 (c) 面心点阵 (d) OT点阵

图 1 4 种正六面体点阵胞元

Fig. 1 Four kinds of regular hexahedral lattice cells

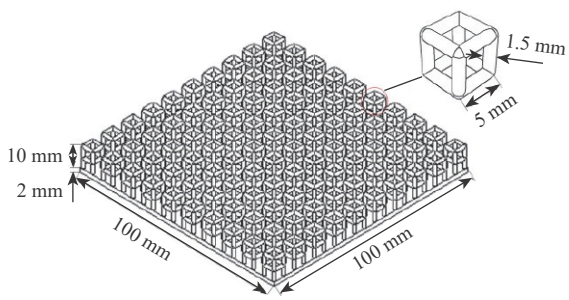


图 2 立方体点阵散热模块模型

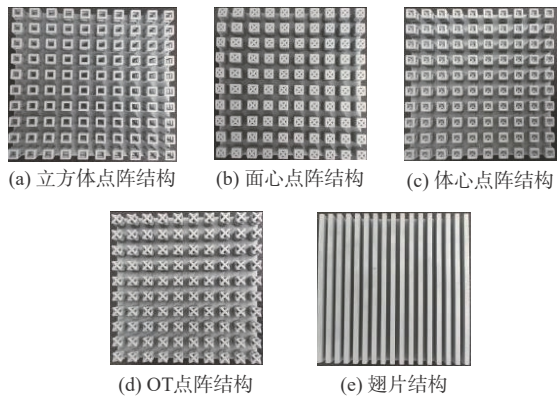
Fig. 2 Cube lattice heat dissipation module model

为比较点阵结构与传统翅片结构的散热性能，增加了翅片结构散热作为实验对照组。翅片模型基板尺寸与点阵散热模块基板尺寸相同，翅片的尺寸参数取自现有 EHA 电机壳体翅片参数，高度为 10 mm，宽度为 3 mm，间距为 3 mm。

受限于 3D 打印工艺精度的约束，本文中点阵结构的最小杆径取 0.5 mm，其他参数参考翅片结构确定，同时，为避免点阵结构的孔隙率过小，调整点阵结构的胞元长度为 5 mm，点阵层数为 2 层，点阵结构总高度为 10 mm。

1.2 3D 打印模型

实验所选材料为铝合金 AlSi10Mg，采用选区激光熔化工艺制造，其成形机理为：在金属粉床上，激光热源按照给定的扫描路径熔化特定区域的粉末材料，成形一层后工作台下降指定高度，同时铺置新的粉末，如此往复直到成形结束[22]。4 种点阵散热模块及翅片散热模块的 3D 打印件如图 3 所示。



(a) 立方体点阵结构 (b) 面心点阵结构 (c) 体心点阵结构

(d) OT点阵结构 (e) 翅片结构

图 3 散热模块 3D 打印实物图

Fig. 3 3D printed physical drawing of heat dissipation structure

2 实验系统

实验系统由点阵结构散热性能实验装置及数据采集系统共同组成。

2.1 实验方案

为探究点阵结构的散热及轻量化性能，针对点阵胞元类型、点阵胞元杆径及出口风速对点阵散热模块散热性能影响进行实验研究。采用控制变量的方法，分别改变点阵类型、杆径及出口风速完成第 1～3 组实验，具体实验工况如表 1 所示。通过测量基板温度、加热板电压与电流等参数，得到散热系数、热阻等参数，进而得到点阵结构散热轻量

表 1 点阵结构散热性能实验工况

Table 1 Working condition of lattice structure heat dissipation characteristic experiment

分组	出口风速/(m·s^{-1})	加热功率/W	杆径/mm	点阵类型
第1组	5	35	1.5	立方体点阵
	5	35	1.5	体心点阵
	5	35	1.5	面心点阵
	5	35	1.5	OT点阵
	5	35	3(翅厚)	翅片
第2组	5	35	0.5	OT点阵
	5	35	1	OT点阵
	5	35	1.5	OT点阵
第3组	1	35	1.5	OT点阵
	3	35	1.5	OT点阵
	5	35	1.5	OT点阵

化性能指标。

实验前对实验装置进行了密闭性测试和加热均匀性测试, 测试结果表明装置满足实验要求。实验中, 当温度波动小于 0.5 ℃ 并保持 200 s 后认为达到稳态, 可以进行实验数据读取和分析。

2.2　实验装置

实验系统原理如图 4 所示, 点阵结构散热性能实验系统如图 5 所示, 其中, 风机提供空气循环动力, 通过调频装置对入口流量进行调节。加热板的加热功率通过调压器调节电压来实现, 并通过万用表测量的电流和电压来计算加热功率。

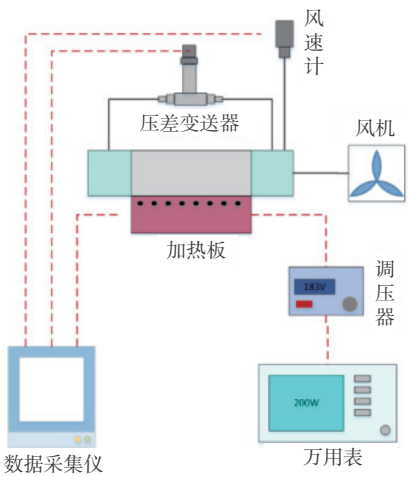
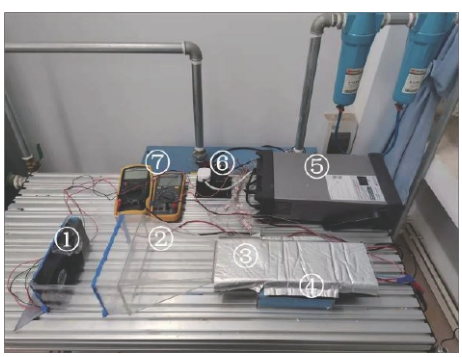


图 4　点阵结构散热性能实验系统原理

Fig. 4　Schematic diagram of lattice structure heat dissipation characteristic experiment system

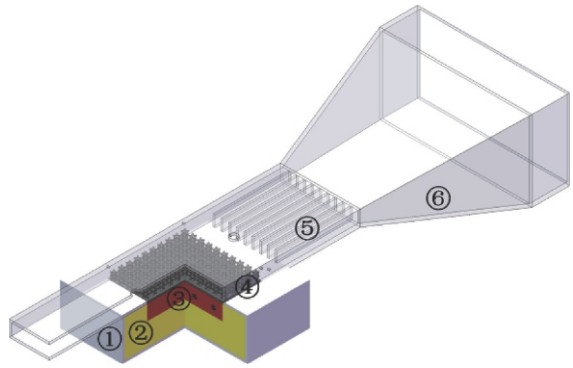


①. 风机; ②. 通风管道; ③. 保温棉; ④. 加热板壳体; ⑤. 安捷伦数据采集仪; ⑥. 调压器; ⑦. 万用表。

图 5　点阵结构散热性能实验台

Fig. 5　Lattice structure heat dissipation characteristic experiment bench

图 6 为主要实验段内部结构示意图。轴流风机出口的流动比较复杂, 为保证流量分配均匀, 风机采用抽吸的方式提供动力。点阵基板与加热板上表面紧密相连, 中间均匀涂抹一层 1 mm 厚度的导热硅脂(导热系数为 2.0 W/(m·K))。为保证加热组件产生的热量全部通过点阵结构散热性能结构,



①. 加热板壳体; ②. 泡沫保温板; ③. 加热板; ④. 点阵散热板; ⑤. 整流板; ⑥. 通风管道。

图 6　主要实验段结构示意图

Fig. 6　Structural diagram of main experimental sections

设计加热板体积为 100 mm×100 mm×20 mm, 其周围除与散热结构接触的上表面均覆盖一层厚度为 30 mm 的高密度 XPS 挤塑板作为保温层, 在加热板周围布置一层厚度为 3 mm 的泡沫保温板, 同时, 在主要实验段的通风管道外包裹一层保温棉作为保温层, 减少热量损失。空气在进入通风管道后, 先进入由若干平行板组成的整流段, 使通过点阵散热模块的空气均匀稳定, 减少湍流度。

根据 EHA 散热需求换算, 选取加热板功率为 25 W、35 W 和 45 W。

假设通风管道出口风速 v 为 0～6 m/s, 已知出口矩形面积 S 为 100 mm×10 mm, 则体积流量为

$$Q = Sv \tag{1}$$

由式(1)可得点阵结构散热性能实验所需流量为 0～0.36 m^3/min。风机的额定功率应略有富余, 因此, 选择额定电压为 12 V、额定电流为 1 A 的轴流风机, 在预实验中, 风机可保证提供实验所需的风速。

2.3　数据采集

图 7 为实验装置与数据采集系统之间的连接图, 其中, 测量仪器的型号及参数如表 2 所示。实验中, 采用 omage K 型热电偶测温。实验设置 12 个热电偶测点, 通过热电偶胶水将 omage K 型热电偶的测温端固定在点阵基板表面, 按 2 个正交方向布置, 反映点阵基板整体温度分布特征。其中, 测点 I 和 O 分别用于测量实验通风管道的进出口空气温度, 安装在进出口管中心位置。测点 A、B、C、D、E 测量点阵基板沿风速方向温度, 将热电偶均匀安装点阵基板表面中线上。测量点 a、b、c、d、e 测量点阵基板垂直空气流动方向温度变化, 将热电偶均匀安装在点阵基板表面垂直于空气流动方向的中线上。

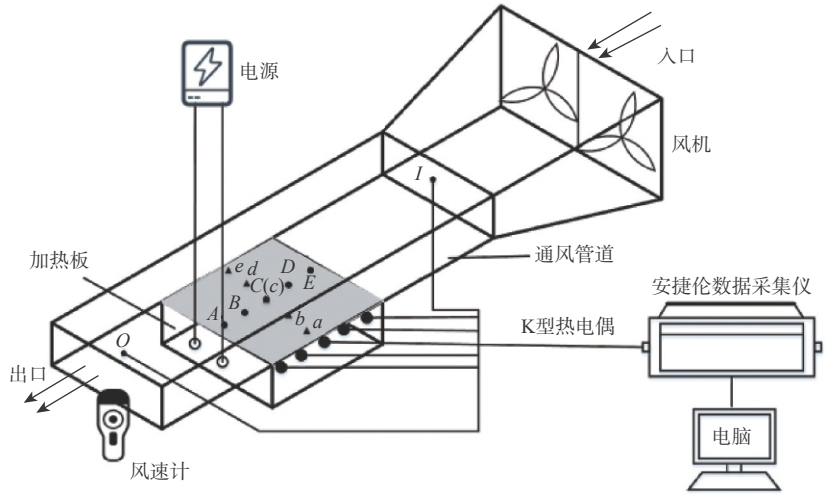


图 7 点阵结构散热性能实验数据采集系统

Fig. 7 Data acquisition system for lattice structure heat dissipation characteristic experiment

表 2 测量仪器型号及参数

Table 2 Measuring instrument model and parameters

设备	型号	参数
安捷伦数据采集仪	34972A	精度: 0.004%
欧米伽热线风速计	HHF-SD1	工作温度0～50 ℃; 分辨率: 0.1 ℃; 精度: 0.8 ℃
数字万用表	Fluke 101	精度: ±(1%+1)
热电偶	K-Type	测温范围: −200～260 ℃; 精度: 0.75%

加热板采用电加热方式, 加热功率通过分别测量加热板电流 I_r 与电压 U 检测, 实际加热板加热功率 P 的计算公式为

$$P = UI_r \tag{2}$$

温度信号由安捷伦数据采集仪实时采集, 采集时间间隔设定为 5 s。点阵散热模块的热阻 R 为点阵散热模块与周围环境的温差 ΔT 与加热板加热功率的比值, 其计算公式为

$$R = \Delta T / P \tag{3}$$

管道出口风速由风速计每 10 min 记录一次, 并在进口截面处多点测量取平均值。环境温度、加热板输出电压和电流用数字万用表每 10 min 记录一次。

2.4 误差分析

实验过程中存在不可避免的系统误差, 温度测量采用的 K 型热电偶精度为 0.75%, 风速误差小于 1%, 数字万用表误差小于 1%。安捷伦数据采集仪误差较小, 可忽略不计。因此, 可根据式(4)计算总误差 W_Y 为 1.6%。

$$W_Y = \left[\left(\frac{\partial Y_1}{\partial x_1}w_1\right)^2 + \left(\frac{\partial Y_2}{\partial x_2}w_2\right)^2 + \cdots + \left(\frac{\partial Y_n}{\partial x_n}w_n\right)^2\right]^{1/2} \tag{4}$$

式中: Y 为实验中变量的变化范围; x_i 为变量的实际测量值; w_i 为测量仪器 i 的准确度等级。

3 结果与分析

3.1 点阵散热模块温度分布特性

以杆径为 1.5 mm 的 OT 点阵散热模块为例, 分析点阵散热模块总体温度分布特性。

为观察点阵散热模块底板温度随时间的瞬时变化情况, 绘制了加热功率为 35 W、风速为 5 m/s 时, 点阵散热模块基板各测温点温度随时间的变化情况, 如图 8 所示。随着加热板不断提供热量, 各点温度均不断升高, 最终达到稳定状态。可以看出, 测温点 1 温度最低, 测温点 5 温度最高且沿着空气流动方向, 相邻 2 个测温点间温差逐渐减小, 这是因为沿着空气流动方向点阵散热模块与空气的温差逐渐减小, 使得后方测温点位置的换热效果较差。

为反映点阵散热模块基板稳态温度的分布特点, 绘制了上述 OT 点阵结构散热性能实验各测温点稳态温度沿不同轴向的分布情况, 如图 9 所示。基板温度沿不同方向存在明显的分布差异性, 在空气流动方向上, 温度一直呈现上升并逐渐减缓的趋势, 这是由于流体与底板之间的换热温差沿此方向逐渐减小导致。在垂直于空气流动方向上, 温度沿

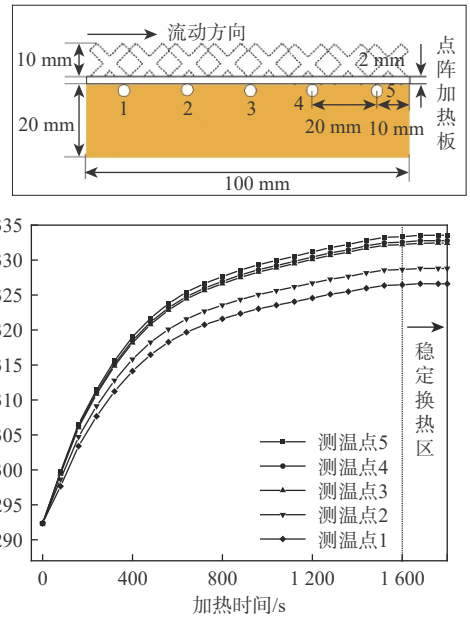


图 8　点阵底板温度随时间变化情况

Fig. 8　Temperature change of lattice bottom plate with time

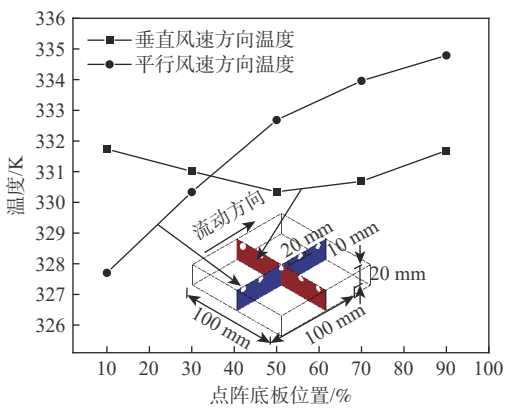


图 9　点阵底板不同方向稳态温度分布情况

Fig. 9　Steady state temperature distribution of lattice bottom plate in different directions

中心位置 (50% 处) 呈对称分布, 主要原因是流动在此方向也是对称的, 同时受边界因素影响, 点阵散热模块两边所处流道流速较低, 因此, 两边温度略高。对比 2 个方向的温度变化, 沿流动方向的温度变化差值约为 7.4 ℃, 沿垂直流动方向的温度变化差值约为 1.5 ℃。由此可知, 点阵散热模块在垂直于流动方向上具有良好的温度分布均匀性, 从而证明, 点阵散热模块大小选取合适, 可以部分排除边界散热影响。

3.2　点阵结构散热轻量化系数定义

本文中, 点阵结构与壳体为增材制造一体成形, 材质相同(导热系数相同), 内部边界条件可视为恒定热流密度, 因此, 散热过程仅受点阵结构参数、外表面对流换热系数及环境温度影响。同时, 在温度不高的散热场合, 其热辐射可以忽略。因此, 可推导其对流换热系数 k 为

$$k=\frac{A_z}{S_j\Delta T} \tag{5}$$

其中: A_z 为总散热功率; S_j 为点阵散热模块(点阵或翅片及基板)与流体接触的对流换热面积; ΔT 为散热达到稳定后, 点阵散热模块基板平均温度与空气平均温度的差值。

本文引入点阵结构散热, 希望增大散热面积强化散热, 同时还具备轻量化的优点。因此, 参照汽车轻量化指标设计思路[23], 将汽车质量替换为点阵体积(也可折算成质量), 将车身的扭转刚度替换为点阵表面对流换热系数, 将车身脚印面积替换为点阵结构与基板的接触面积, 构造出散热轻量化系数 L 如下:

$$L=\frac{V}{kS_d} \tag{6}$$

式中: V 为散热结构(点阵或翅片)总体积; S_d 为散热结构与基板的接触面积。

点阵结构散热轻量化系数可以综合评价点阵结构的散热性能与轻量化特性, 为点阵结构的选取提供依据。

3.3　点阵结构散热的轻量化评价

点阵散热模块的散热方式为强迫对流换热, 其中, 风速对换热效果有明显影响。为分析风速对点阵散热模块散热性能的影响, 图 10 给出 OT 点阵结构及翅片在加热功率为 35 W、出口风速分别为 1, 3,5 m/s 条件下的热阻。图中: r 为点阵结构的杆件半径。可以看出, 点阵散热模块与翅片散热模块的热阻随风速增大而减小。在风速较小为 1 m/s 时, 翅片与点阵散热模块的热阻相近, 翅片略小于点阵结构, 热阻差值约为 0.1 K/W。而随着风速的增大, 点阵散热模块热阻逐渐小于翅片的热阻, 差值增大至约 0.35 K/W。这是因为随着风速的增大, 对流换热能力也随之增强, 点阵结构的特殊空隙形状增大

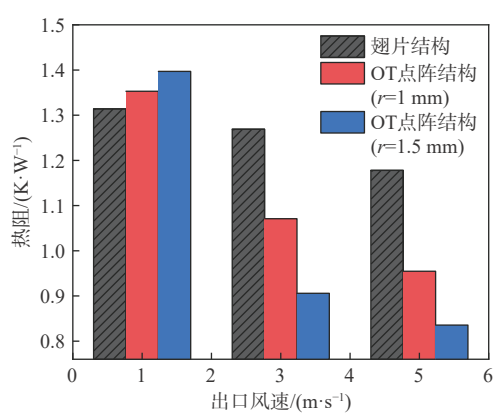


图 10　热阻随风速变化情况

Fig. 10　Thermal resistance changes with wind speed

了当量散热面积，同时，点阵杆柱对空气流动存在扰动，从而进一步强化散热，因此，随着风速的增加，点阵散热模块热阻减小更明显。对于不同杆径的 OT 点阵散热模块，杆径为 1.5 mm 的点阵散热模块散热面积远大于杆径为 1 mm 的点阵散热模块，故其散热热阻也更小。

杆径作为点阵结构的重要参数之一，其大小决定了点阵结构的孔隙率。孔隙率可定义为点阵结构中孔隙的体积与点阵结构所占包络体积的比值。本文中，散热结构（点阵和翅片）所占包络体积取为相同，则孔隙率与散热结构实际体积呈负相关关系，即在散热轻量化系数 L 指标中隐含了孔隙率因素，体积越小，孔隙率越大。在加热功率为 35 W、出口风速为 5 m/s 的条件下，对点阵胞元大小为 5 mm、杆径分别为 0.5,1,1.5 mm 的 OT 点阵散热模块进行实验研究，结果如图 11 所示。可以看出，点阵散热模块的散热系数和散热轻量化系数都随杆径的增大而增大，但增幅变化却不相同。散热系数增幅由 61.5% 减小为 3.3%，散热轻量化系数增幅由 18.6% 增大为 40%，这是因为随着杆径的增大，改变的点阵散热模块的散热性能影响参数包括散热面积和孔隙率。其中，散热面积随杆径增大而增大，故对点阵散热模块散热效果起正向作用。点阵结构孔隙率随着杆径增大逐渐减小，孔隙率的减小会导致空气的渗透率变低，从而降低点阵散热模块的对流换热效果，故对点阵散热模块散热效果起负向作用。综上所述，当杆径过小时，点阵结构较轻，但其散热性能较差。反之，当杆径过大时，散热性能虽然增强，但其整体质量又过大，点阵结构杆径的选取受散热面积和孔隙率的双重影响。以上分析表明，利用散热轻量化系数可以为杆径的选取提供理论依据。

图 12 给出了不同类型点阵结构的散热系数与散热轻量化系数。可以看出，在 5 种散热结构中，按散热系数比较为：OT 点阵结构>体心点阵结构>翅片结构>立方体点阵结构>面心点阵结构。究其原因：OT 点阵结构相较于翅片结构增大了散热面积，同时由于其杆柱搭建的空隙的存在，增强了空气扰动，OT 点阵结构相较于翅片结构散热能力更强，OT 点阵结构散热模块散热系数为翅片的 1.2 倍。面心点阵结构散热系数小于翅片结构的原因在于：面心点阵结构杆柱的分布导致空气渗透率相较于其他点阵结构更低，散热效果相对较差。而对于立方体点阵结构，其杆柱与加热板的接触面积较小，且相较于其他点阵结构散热面积较小，导致其散热效果较差。按散热轻量化系数比较，OT 点阵结构<立方体点阵结构<体心点阵结构<面心点阵结构<翅片结构。可以看出，4 种点阵结构平均散热轻量化系数相较于翅片结构可减小 61%，且在几种点阵结构中，OT 点阵结构散热轻量化系数最小，仅约为翅片结构散热轻量化系数的 1/3。综上所述，OT 点阵结构具有最大散热系数 110 W/(m^2K)及最小的散热轻量化系数 4.2×10^{-6} m^3K/W，其是在本文研究的几种点阵结构中，综合考虑散热能力及轻量化条件下的最优结构。

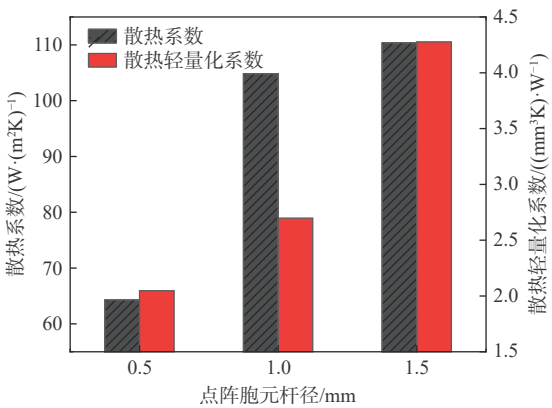


图 11 点阵结构散热性能及轻量化特性随杆径变化

Fig. 11 Heat dissipation performance and lightweight characteristics of lattice structure change with rod diameter

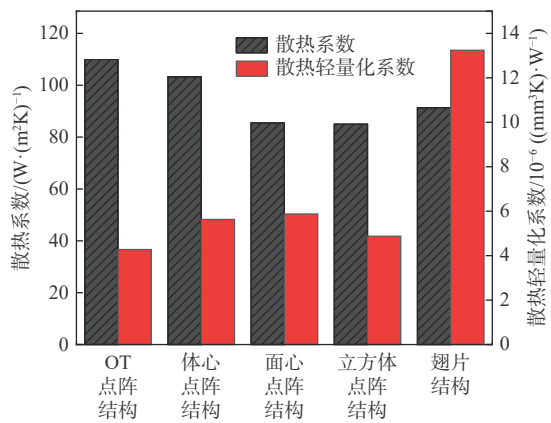


图 12 不同类型点阵结构散热性能及轻量化特性

Fig. 12 Heat dissipation performance and lightweight characteristics of different types of lattice structure

4 结 论

本文采用实验研究了轻量化点阵结构的散热性能，比较点阵结构散热性能随参数变化情况及其轻量化特性。

1）得到点阵散热模块底板的温度分布特性，其温度在各个方向分布情况与经验相符，为后续研究的可靠性提供依据。

2）研究表明，点阵结构相较于翅片散热性能受风速影响更为显著，增大风速时，点阵结构散热性能有明显提升，在具有高风速的对流换热工况下，

更适合应用点阵结构。

3) 提出散热轻量化系数, 并对各类型点阵结构散热轻量化性能进行量化评价, 结果表明, OT 点阵结构作为最优散热结构, 其散热系数为翅片结构的 1.2 倍, 散热轻量化系数为翅片结构的 1/3。

4) 研究表明, 轻量化点阵结构的散热性能受散热面积和孔隙率的双重影响, 依据散热轻量化系数对点阵杆径的选取要综合考虑 2 个因素, 根据具体需求选出合适的杆径大小。

参考文献（References）

[1] 许巍, 赵俊伟, 袁本立, 等. 飞行器热防护与利用一体化系统实验与模拟[J]. 航空动力学报, 2022, 37(3): 555-563.
XU W, ZHAO J W, YUAN B L, et al. Experiment and simulation on aircraft thermal protection and utilization integrated system[J]. Journal of Aerospace Power, 2022, 37(3): 555-563(in Chinese).

[2] 朱喜华, 李颖晖, 刘聪, 等. 基于自适应核主元分析的 EHA 系统传感器故障检测[J]. 推进技术, 2014, 35(6): 838-845.
ZHU X H, LI Y H, LIU C, et al. Sensor fault detection for EHA system based on adaptive kernel principal component analysis[J]. Journal of Propulsion Technology, 2014, 35(6): 838-845(in Chinese).

[3] 李永欣, 郭长春, 李凯伦, 等. SLM 制备的小长径比四棱锥点阵填充轻量化研究[J]. 机械工程学报, 2021, 57(24): 132-138.
LI Y X, GUO C C, LI K L, et al. Filling lightweight research on pyramid lattice with small slenderness ratio fabricated by SLM[J]. Journal of Mechanical Engineering, 2021, 57(24): 132-138(in Chinese).

[4] 王伟, 袁雷, 王晓巍. 飞机增材制造制件的宏观结构轻量化分析[J]. 飞机设计, 2015, 35(3): 24-28.
WANG W, YUAN L, WANG X W. Macro-structural lightweight analysis for aircraft parts made by additive manufacturing technology[J]. Aircraft Design, 2015, 35(3): 24-28(in Chinese).

[5] LIAN W L, NIU W J, LIN L J. A passive cooling design for an aircraft electromechanical actuator by using heat pipes[J]. Applied Thermal Engineering, 2021, 184: 116248.

[6] OENLER C D. Fluid flow and heat transfer simulation of an electro hydraulic actuator[D]. Delft: Delft University of Technology, 2016.

[7] 白晓辉, 刘存良, 孟宪龙, 等. 八面体桁架结构在内冷通道中的流动传热特性研究[J]. 推进技术, 2022, 43(7): 286-295.
BAI X H, LIU C L, MENG X L, et al. Flow and heat transfer characteristics of octet truss structure in internal cooling channel[J]. Journal of Propulsion Technology, 2022, 43(7): 286-295(in Chinese).

[8] 邓昊宇, 王春洁. 三维点阵结构等效热分析与优化方法[J]. 北京航空航天大学学报, 2019, 45(6): 1122-1128.
DENG H Y, WANG C J. Equivalent thermal analysis and optimization method for three-dimensional lattice structure[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(6): 1122-1128(in Chinese).

[9] EKADE P, KRISHNAN S. Fluid flow and heat transfer characteristics of octet truss lattice geometry[J]. International Journal of Thermal Sciences, 2019, 137: 253-261.

[10] 易长炎, 柏龙, 陈晓红, 等. 金属三维点阵结构拓扑构型研究及应用现状综述[J]. 功能材料, 2017, 48(10): 10055-10065.
YI C Y, BAI L, CHEN X H, et al. Review on the metal three-dimensional lattice topology configurations research and application status[J]. Journal of Functional Materials, 2017, 48(10): 10055-10065(in Chinese).

[11] KHADERI S N, DESHPANDE V S, FLECK N A. The stiffness and strength of the gyroid lattice[J]. International Journal of Solids and Structures, 2014, 51(23-24): 3866-3877.

[12] SHEN B B, LI Y, YAN H B, et al. Heat transfer enhancement of wedge-shaped channels by replacing pin fins with Kagome lattice structures[J]. International Journal of Heat and Mass Transfer, 2019, 141: 88-101.

[13] SON K N, WEIBEL J A, KUMARESAN V, et al. Design of multifunctional lattice-frame materials for compact heat exchangers[J]. International Journal of Heat and Mass Transfer, 2017, 115: 619-629.

[14] LIANG D, CHEN W, JU Y C, et al. Comparing endwall heat transfer among staggered pin fin, Kagome and body centered cubic arrays[J]. Applied Thermal Engineering, 2021, 185: 116306.

[15] TAKEZAWA A, KOBASHI M, KOIZUMI Y, et al. Porous metal produced by selective laser melting with effective isotropic thermal conductivity close to the Hashin-Shtrikman bound[J]. International Journal of Heat and Mass Transfer, 2017, 105: 564-572.

[16] VAISSIER B, PERNOT J P, CHOUGRANI L, et al. Parametric design of graded truss lattice structures for enhanced thermal dissipation[J]. Computer-Aided Design, 2019, 115: 1-12.

[17] CHENG L, LIU J K, LIANG X, et al. Coupling lattice structure topology optimization with design-dependent feature evolution for additive manufactured heat conduction design[J]. Computer Methods in Applied Mechanics and Engineering, 2018, 332: 408-439.

[18] 张龙, 李昂, 赵云鹏, 等. 一种全封闭蒙皮点阵支撑结构的优化设计与试验验证[J]. 机械工程学报, 2021, 57(22): 35-42.
ZHANG L, LI A, ZHAO Y P, et al. Optimal design and experimental verification of an enclosed skin lattice support structure[J]. Journal of Mechanical Engineering, 2021, 57(22): 35-42(in Chinese).

[19] VALDEVIT L, HUTCHINSON J W, EVANS A G. Structurally optimized sandwich panels with prismatic cores[J]. International Journal of Solids and Structures, 2004, 41(18-19): 5105-5124.

[20] MOON S K, TAN Y E, HWANG J, et al. Application of 3D printing technology for designing light-weight unmanned aerial vehicle wing structures[J]. International Journal of Precision Engineering and Manufacturing-Green Technology, 2014, 1(3): 223-228.

[21] XU X , JIANG Y, LEE H P. Engineering optimization multi-objective optimal design of sandwich panels using a genetic algorithm multi-objective optimal design of sandwich panels using a genetic algorithm[J]. Engineering Optimization, 2017, 49(10): 1665-1684.

[22] 侯伟, 陈静, 储松林, 等. 选区激光熔化成形 $AlSi_{10}Mg$ 组织与拉伸性能的各向异性研究[J]. 中国激光, 2018, 45(7): 67-77.
HOU W, CHEN J, CHU S L, et al. Anisotropy of microstructure and tensile properties of $AlSi_{10}Mg$ formed by selective laser melting[J]. Chinese Journal of Lasers, 2018, 45(7): 67-77(in Chinese).

[23] 姚再起, 马芳武, 刘强, 等. 汽车轻量化评价方法研究[J]. 中国工程科学, 2014, 16(1): 36-39.
YAO Z Q, MA F W, LIU Q, et al. Evaluation methods of vehicle lightweight[J]. Engineering Sciences, 2014, 16(1): 36-39(in Chinese).

Heat dissipation characteristics and lightweight evaluation of lattice structure

HU Jianjun[1, 2, *], DOU Ruochen[2], ZHANG Xin[1], YAO Jing[2], KONG Xiangdong[2]

(1. School of Civil Engineering and Mechanics, Yanshan University, Qinhuangdao 066044, China;
2. School of Mechanical Engineering, Yanshan University, Qinhuangdao 066044, China)

Abstract: As the core component of aviation rudder control, the pursuit of a high power-to-weight ratio for an electro-hydrostatic actuator (EHA) requires efficient heat dissipation and a lightweight integrated design for its components. Reliable data for assessing lattice structure heat dissipation and lightweight performance is provided by this paper's experimental study of the heat dissipation properties of four common lightweight lattice structures and the temperature distribution properties of various lattice heat dissipation modules. Based on the heat dissipation characteristics of lattice structures, an evaluation index of the heat dissipation lightweight coefficient of lattice structures is proposed to quantitatively evaluate the lightweight characteristics of different heat dissipation lattice structures. The four types of lattice structures are assessed using the evaluation index developed in this research, and the findings indicate that the OT lattice structure has the best combined lightweight and heat dissipation performance. Compared with the fin structure, the heat dissipation coefficient of the OT lattice structure is 1.2 times that of the fin, but its heat dissipation lightweight coefficient is only about 1/3 of that of the fin. This study provides a theoretical basis for the screening of the lattice structure considering both heat dissipation and lightweight.

Keywords: lattice structure; heat dissipation; lightweight; experimental research; evluation

Received: 2023-09-06; **Accepted**: 2024-02-23; **Published Online**: 2024-04-04 07:23

URL: link.cnki.net/urlid/11.2625.V.20240403.0916.001

Foundation items: The General Program (Major Research Plan) of National Natural Science Foundation of China (51890881); Natural Science Foundation of Hebei Province (2020203028); Central Guiding Local Science and Technology Development Fund Projects (226Z1902G)

* **Corresponding author**. E-mail: kewei729@163.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0527

多角度偏振成像仪数据校正方法对比与评价

王羿[1,2]，孟炳寰[1,2,*]，叶晨[1,2]，骆冬根[1,2]，姚萍萍[1,2]，洪津[1,2]

(1. 中国科学院合肥物质科学研究院，合肥 230031；2. 中国科学院通用光学定标与表征技术重点实验室，合肥 230031)

摘 要：多角度偏振成像仪(DPC)在数据校正过程主要包括本底校正和帧转移校正2部分。光电探测器的图像采集是一个随机抽样的过程，数据校正过程需对抽样结果进行运算，从而引入噪声。结合DPC数据校正过程，主要研究本底校正和帧转移效应校正对信噪比(SNR)的影响，分析2种帧转移效应校正方法的情况，分别对SNR进行建模和仿真，并用大气环境监测卫星DPC实测数据进行验证。结果表明：本底校正和帧转移效应校正均会使噪声增加，降低SNR；本底校正时应选择大于5帧本底探测数据；2种帧转移效应校正方法相比，列总量比例法优于暗行法，最大可以使SNR增加15.8。

关 键 词：多角度偏振成像仪；数据校正；信噪比；帧转移效应；电荷耦合器件

中图分类号：V19

文献标志码：A **文章编号**：1001-5965(2025)10-3307-06

信噪比通常指信号与噪声的比值，是评估测量结果有效性的重要指标，反映了数据质量，直接影响了数据的应用效果。为研究信噪比，国内外学者针对不同类型的遥感器和不同类型的数据开展了大量的研究工作。其研究方法主要归纳为两大类：数学模型评估法和数据统计评估法。

数学模型评估法指根据仪器系统设计和特点，对信号值和噪声值进行建模，得到信噪比与电子数或电流的关系，再通过实测值计算信噪比。美国Terra卫星搭载的中分辨率成像光谱仪(moderate resolution imaging spectroradiometer, MODIS)[1]，NPP卫星搭载的可见光红外成像辐射仪(visible infrared imaging radiometer, VIIRS)[2]及环境减灾二号卫星搭载的偏振扫描大气校正仪(polarized scanning atmospheric corrector, PSAC)[3]等仪器都使用数学模型评估信噪比。用该方法评估可以得到比较准确的信噪比。

数据统计评估方法通常在无法建立模型的情况下使用，利用已有的若干组数据进行统计得到信号值和噪声值：信号值为信号的期望，通常采用多次测量的平均值代替；噪声值用于衡量信号的随机波动，用标准偏差进行计算。期望值除以标准偏差得到信噪比。为提高信噪比评估的准确性，相关学者提出了局部方差法[4]、求地学统计量法[5]和人工神经网络等[6]方法提高评估的准确性。对于分时偏振探测仪器信噪比计算，如Fougnie等[7]基于云场景和海洋场景计算POLDER仪器的在轨噪声，从而评估仪器的信噪比。

以上方法均是根据仪器测试数据进行信噪比评估，然而在仪器数据校正和处理过程中，当涉及几组测量数据之间计算的情形，则会引入噪声。数据校正通常包括本底校正和帧转移效应校正，本底校正需要用感光图像和本底图像相减，帧转移效应校正也需要对感光图像进行处理，这2种处理过程均会使噪声增加，使信噪比变小，因此，研究校正过程的信噪比是十分有意义的。本文介绍了多角度

收稿日期：2023-08-17；**录用日期**：2023-10-13；**网络出版时间**：2023-11-17 14:48
网络出版地址：link.cnki.net/urlid/11.2625.V.20231116.1500.002
基金项目：王宽诚教育基金“先进偏振遥感技术与应用国际团队”(GJTD-2018-15)
*通信作者. E-mail：bhmeng@aifom.ac.cn

引用格式：王羿，孟炳寰，叶晨，等. 多角度偏振成像仪数据校正方法对比与评价[J]. 北京航空航天大学学报，2025，51(10)：3307-3312. WANG Y，MENG B H，YE C，et al. Comparison and evaluation of data correction methods for directional polarimetric cameras[J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51(10)：3307-3312 (in Chinese).

偏振成像仪探测工作原理，并介绍了针对仪器所采用的探测器的数据校准过程，对几种校正方法进行分析并得到结论。

1 多角度偏振成像仪及数据校正过程简介

1.1 多角度偏振成像仪器原理简介

多角度偏振成像仪是中国科学院安徽光学精密机械研究所研制的一种用于云、气溶胶和水汽探测的分时偏振探测仪器，该仪器已搭载于多型号卫星并在轨运行，包括高分五号卫星、高光谱观测卫星、大气环境监测卫星和陆地生态系统碳监测卫星等。

多角度偏振成像仪由广角镜头、镜片转轮和面阵探测器3部分组成，其仪器组成如图1所示。



图1 多角度偏振成像仪组成

Fig. 1 Composition of directional polarimetric camera

仪器工作时转轮沿一个方向旋转，当镜头、转轮上的通道位置匹配时，探测器进行曝光成像。转轮上各通道安装不同波段的滤光片，相邻3个通道安装偏振方向不同的偏振片实现不同波段和偏振的探测。在转轮上设置一个不透光的通道（称为本底通道），每次通过该通道仍控制探测器曝光，得到的图像用于本底校正。

1.2 数据校正过程简介

遥感仪器数据校正的目的，是将仪器输出的数字量计数值或灰度值（digital number,DN）转换为物理量（如辐亮度、偏振度等）的过程，多角度偏振成像仪数据校正过程由仪器设计特点决定，主要包括本底校正和帧转移效应校正2部分[8]，仪器采用了面阵帧转移型电荷耦合器件（charge coupled device, CCD），且没有快门装置。具备这种设计方案的科学型相机都需要开展本底校正和帧转移效应校正工作。经过校正后得到的DN除以仪器响应度（即绝对辐射定标系数）可得到仪器入瞳处的光谱辐亮度。

遥感仪器原始图像（指模数转换后的DN）由3部分组成，曝光时光电子转换的DN、暗电流产生的DN和电路偏置电压经模数转换后的DN。本底校正指原始图像扣除暗电流和电路偏置的过程，其结果反映了探测器实际收到的光子数。多角度偏振成像仪本底校正是用曝光通道DN扣除本底通道DN，本底通道DN包含暗电流和电路偏置2部分。

帧转移效应校正目的是校正探测器的拖尾效应。拖尾效应是帧转移型面阵CCD探测器的固有特性。在没有快门装置的情况下，帧转移型探测器的电荷包在转移过程中会感光并产生光电子。这部分光电子不属于设计的曝光时间内产生的，在探测器行数较多、曝光时间短的情况下，这部分光电子产生的DN甚至可以占到总输出DN的20%，因此，拖尾效应产生的影响不可忽略。扣除拖尾效应造成的DN增加，才能得到有效信号。在进行信噪比计算和研究时，信号应指有效信号，即帧转移效应校正后得到的信号，而用含拖尾效应的信号计算信噪比是不适当的。

帧转移效应校正的方法一般有2种[9-10]，分别命名为列总量比例法和暗行法，列总量比例法计算一列所有像素DN之和，乘以帧转移系数得到帧转移过程增加的DN。其中，帧转移系数定义为拖尾效应产生的DN（或电子数）占总DN（或电子数）的百分比。暗行法利用探测器不感光像元在帧转移过程中感光的特点，计算帧转移过程增加的DN。

2 数据校正过程信噪比模型建立和仿真

2.1 信噪比建模

仪器信噪比中的信号和噪声都用电子数进行计算[11-15]，在信噪比建模过程中，定义 $n(i,j,k)$ 为像元对应的电子数，其中，i 表示图像行序号，j 表示图像列序号，k 为成像图像或本底图像序号；并用 σ 表示噪声、σ_{d} 表示暗电流噪声、R_{SN} 表示信噪比。

2.1.1 信号建模

多角度偏振成像仪原始信号值（即仪器输出的DN）由3部分组成：实际信号、本底信号（即暗电流和电路偏置）及帧转移效应产生的信号。用于信噪比建模的数值为实际信号。

本底校正过程，信号值计算式为[8]

$$n'(i,j)=n(i,j,k)-\frac{1}{X}\sum_{k=0}^{X}n_{\mathrm{d}}(i,j,k) \tag{1}$$

式（1）表示原始信号值扣除 X 个本底平均后的值，其中，X 为本底图像数量。计算后的 n' 由于只有一幅图像（即 k=1），可以对 k 进行省略。

帧转移效应校正分为列总量比例法和暗行法2种, 其中暗行法信号值计算式为[9]

$$n''(i,j)=n'(i,j)-\frac{1}{Y}\sum_{j=0}^{Y}n'(i,j) \tag{2}$$

式中: Y 为暗行的数量。

另一种校正方法—列总量比例法信号值计算式为

$$n'''(i,j)=n'(i,j)-\delta\sum_{i=0}^{Z}n'(i,j) \tag{3}$$

式中: Z 为探测器的总行数; δ 为帧转移时间系数, 由行转移时间和积分时间共同决定。在完成本底信号和帧转移效应产生的信号扣除后, 得到的信号值即为实际信号, 用于后续信噪比计算。

2.1.2　噪声建模

多角度偏振成像仪噪声分为原始信号散粒噪声、电路产生的正态分布随机噪声和量化噪声[11]。其中, 电路产生的噪声和量化噪声小于散粒噪声1～2个数量级, 建模时忽略。

预处理过程中本底校正和帧转移校正, 相当于用随机试验的结果进行计算, 会产生噪声, 这2类噪声均为散粒噪声。结合以上分析, 噪声建模时, 需考虑3部分: 原始信号散粒噪声、本底校正过程叠加的散粒噪声和帧转移校正过程叠加的散粒噪声。

原始信号散粒噪声符合泊松分布, 方差与期望相等, 因此, 噪声(即标准差)计算式为

$$\sigma(i,j)=\sqrt{n(i,j)} \tag{4}$$

本底校正是原始信号扣除本底的过程, 本底信号噪声主要表现为暗电流散粒噪声, 同样符合泊松分布。式(1)用 X 个本底进行平均, 因此, 方差为期望的 $1/X$。本底校正产生的噪声计算式为

$$\sigma_{\mathrm{d}}(i,j)=\sqrt{\frac{1}{X}\left(\frac{1}{X}\sum_{k=0}^{X}n_{\mathrm{d}}(i,j,k)\right)} \tag{5}$$

帧转移校正过程产生噪声的原理与本底校正相同, 2种校正方法不同产生的噪声也不同。暗行法采用 Y 个暗行数据进行校正, 产生的噪声计算式为

$$\sigma''(i,j)=\sqrt{\frac{1}{Y}\left(\frac{1}{Y}\sum_{i=0}^{Y}n(i,j)\right)} \tag{6}$$

列总量比例法采用总行数 Z 个数据进行校正, 产生的噪声计算式为

$$\sigma'''(i,j)=\sqrt{\frac{1}{Z}\left(\delta\sum_{i=0}^{Z}n'(i,j)\right)} \tag{7}$$

数据校正过程噪声模型为3部分的合成, 即

$$\sigma(i,j)=\sqrt{n(i,j)+\sigma_{\mathrm{d}}^2+\sigma_{\mathrm{ftc}}^2} \tag{8}$$

式中: σ_{ftc} 为帧转移校正产生的噪声, 根据不同的帧转移效应校正方法由式(6)或式(7)得到。因此, 数据校正过程产生的噪声是受校正方法影响的。通常探测器总行数 Z 较大, 列总量比例法产生的噪声值较小, 可以认为帧转移校正过程不会产生噪声。

结合信号建模和噪声建模结果, 信噪比计算式为

$$R_{\mathrm{SN}}=\frac{n_{\mathrm{ac}}(i,j)}{\sigma(i,j)} \tag{9}$$

式中: n_{ac} 为校正后的电子数, 根据不同的帧转移效应校正方法由式(2)或式(3)得到, σ 为式(8)的计算结果。

2.2　信噪比仿真

根据信噪比建模可知, 本底数量不同, 帧转移效应校正方法不同, 引入的噪声不同。设计一组典型场景, 分析仪器在不同入瞳辐亮度情况下的信噪比变化情况。

观测场景模拟如下: 陆地有云场景, 其中, 云覆盖率为60%, 晴空占比为40%。选择865 nm波段, 本底数量为5, 晴空场景归一化辐亮度取值1.085, 云场景取值0.245。该场景模拟了典型陆地含云的情况, 波段及本底数量选择靠近实际应用。其中, 晴空场景和云场景归一化辐亮度的数值是根据POLDER/PARASOL全球覆盖数据统计得到。参与统计的数据为2012年6月22日UTC时间0时至24时的载荷1级数据。全球一天的数据可以覆盖冬夏不同季节、陆地海洋不同地表、云和晴空不同大气条件及各种太阳观测几何。可以认为辐亮度数据统计具有代表性。

图像和探测器参数按照多角度偏振成像仪设计值选取如下: 单列1 024个像素, 按照观测场景, 该列中有614个像素(即60%)的归一化辐亮度取值1.085, 剩下410个像素(40%)取值0.245。积分时间设计为20 ms, 行转移时间为6 μs, 帧转移系数为0.235。

利用单色辐照度和绝对辐射定标系数可以把归一化辐亮度转换为图像DN, 再根据系统增益将DN转换为电子数。仿真场景下不同辐亮度计算结果如表1所示。

计算帧转移效应产生的灰度值和电子数, 如表2所示。

根据以上计算结果, 结合信噪比过程建模, 在暗电流噪声不考虑的情况下, 计算不同辐亮度仪器信噪比, 得到结果如表3所示。

由表3可知, 列总量比例法计算结果与理论最

表1 辐亮度计算结果

Table 1 Results of radiance calculation

场景	归一化辐亮度	辐亮度/($W\cdot m^{-2}\cdot sr^{-1}\cdot \mu m^{-1}$)	信号灰度值	信号电子数
陆地晴空平均值	0.245	75.000	3 722.4	35 896
陆地有云最大值	1.085	332.771	16 516.1	159 268
归一化辐亮度为0.1	0.100	30.612	1 519.3	14 651
归一化辐亮度为0.2	0.200	61.340	3 044.4	29 358
归一化辐亮度为0.3	0.300	92.066	4 569.4	44 064
归一化辐亮度为0.4	0.400	122.449	6 077.4	58 606
归一化辐亮度为0.5	0.500	153.351	7 611.1	73 395
归一化辐亮度为0.6	0.600	184.132	9 138.8	88 127
归一化辐亮度为0.7	0.700	214.286	10 635.4	102 559

大信噪比一致，且大于暗行法，将两者信噪比之差

表2 帧转移效应计算结果

Table 2 Results of frame transfer effect calculation

项目	灰度值	电子数
晴空场景	3 722.4	35 896
有云场景	16 516.1	159 268
场景平均值	8 839.9	85 245
帧转移效应	2 077.4	20 033

进行归一化并作为纵坐标，以归一化辐亮度作为横坐标进行绘图，结果如图2所示。

由图2可知，归一化辐亮度越小，2种方法计算的信噪比之差最大，最大值达到13.6%，辐亮度越大则越接近，利用暗行法计算的信噪比始终小于列总量比例法，最小值为2.3%。

表3 信噪比仿真计算结果

Table 3 Signal-to-noise ratio simulation results

归一化辐亮度值	信号电子数	帧转移电子数	理论最大信噪比	噪声等效电子数		信噪比	
				暗行法	列总量比例法	暗行法	列总量比例法
0.1	14 651	20 033	121.0	140.2	104.5	121.0	121.0
0.2	29 358	20 033	171.3	185.4	158.4	171.3	171.3
0.3	44 064	20 033	209.9	221.5	198.9	209.9	209.9
0.4	58 605	20 033	242.1	252.2	232.4	242.1	242.1
0.5	73 395	20 033	270.9	280.0	262.1	270.9	270.9
0.6	88 128	20 033	296.9	305.2	288.8	296.9	296.9
0.7	102 560	20 033	320.2	328.0	312.7	320.2	320.2

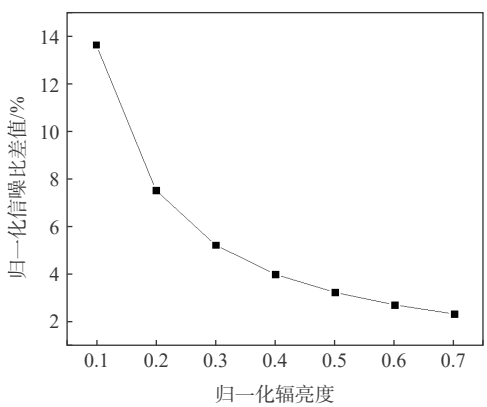


图2 信噪比曲线

Fig. 2 Signal-to-noise ratio curve

3 测试数据验证及讨论

使用大气环境监测卫星多角度偏振成像仪在轨实测数据进行验证。选择2022年4月27日卫星飞过红海和撒哈拉沙漠的场景进行分析。场景如图3所示。

图3场景大部分被沙漠覆盖，归一化辐亮度大于陆地晴空平均值，且小于云场景，扣除图中云场景并进行统计可得，该场景平均灰度值为8 872.1，与表2中平均灰度值8 839.9相差小于1%，可以用于检验模型。同时该场景包括低反射率场景（河流），可以检验归一化辐亮度较小时的信噪比。从卫星数据中分别提取无云覆盖的沙漠和河流进行计算，得到的结果如表4所示。

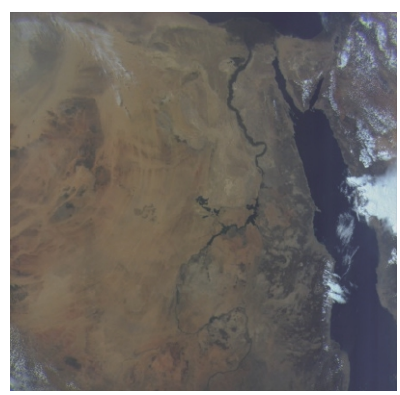


图3 计算信噪比的场景

Fig. 3 Scene for Signal-to-noise ratio calculation

由表4可知，2种不同场景，用暗行法计算的信噪比均偏小，两者之差分别为9.5和15.8，相对偏差为3.4%和13.2%。与仿真计算结果（见图2）相比分别相差0.02和0.04，仿真与实测结果一致。另

表 4　卫星数据计算结果

Table 4　Calculation results using satellite data

场景	信号电子数	帧转移电子数	最大信噪比理论值	本底校正后信噪比	暗行法信噪比	列总量比例法信噪比	相对差值/%
沙漠	77 064	21 034	277.60	276.20	266.4	275.9	3.4
沙漠中河流	14 504	19 832	120.43	119.15	103.1	118.9	13.2

外, 表 4 中理论最大信噪比和本底校正后信噪比相差不大, 说明采用 5 帧平均的方法进行本底校正对信噪比影响很小, 也验证了仿真时忽略暗电流噪声的合理性。

4　结　论

1) 本文对多角度偏振成像仪数据校正过程进行了建模、仿真, 实测数据的验证与仿真情况一致性好, 证明了信噪比计算模型的正确性。

2) 在仪器设计时, 应尽量降低帧转移效应的影响, 可以采取至少 3 种方法: 增加积分时间、缩短行转移时间及采用机械快门阻挡电荷包转移过程中的光照。

后续可以围绕帧转移系数对信噪比的影响这一研究方向继续开展工作。

参考文献（References）

[1] ANGAL A, XIONG X X, SUN J Q, et al. On-orbit noise characterization of MODIS reflective solar bands[J]. Journal of Applied Remote Sensing, 2015, 9: 094092.

[2] TWEDT K A, LEI N, XIONG X J. On-orbit noise characterization of the SNPP VIIRS reflective solar bands[C]//Proceedings of the Earth Observing Systems XXII. Bellingham: SPIE, 2017.

[3] 董浩, 刘振海, 赵鑫鑫, 等. 环境减灾二号卫星偏振扫描大气校正仪在轨信噪比评估[J]. 光学学报, 2022, 42(24): 2401005.
DONG H, LIU Z H, ZHAO X X, et al. On-orbit SNR evaluation of polarized scanning atmospheric corrector onboard HJ-2 satellites[J]. Acta Optica Sinica, 2022, 42(24): 2401005(in Chinese).

[4] 高连如, 张兵, 张霞, 等. 基于局部标准差的遥感图像噪声评估方法研究[J]. 遥感学报, 2007, 11(2): 201-208.
GAO L R, ZHANG B, ZHANG X, et al. Study on the method for estimating the noise in remote sensing images based on local standard deviations[J]. Journal of Remote Sensing, 2007, 11(2): 201-208 (in Chinese).

[5] CHAPPELL A, SEAQUIST J W, EKLUNDH L. Improving the estimation of noise from noaa avhrr ndvi for Africa using geostatistics[J]. International Journal of Remote Sensing, 2001, 22(6): 1067-1080.

[6] DELVIT J M, LEGER D, ROQUES S, et al. Signal-to-noise ratio assessment from nonspecific views[C]//Proceedings of the Image and Signal Processing for Remote Sensing VII. Bellingham: SPIE, 2002.

[7] FOUGNIE B, BRACCO G, LAFRANCE B, et al. PARASOL in-flight calibration and performance[J]. Applied Optics, 2007, 46(22): 5435-5451.

[8] 涂碧海. 大气气溶胶多角度偏振成像仪数据校正及检验方法研究[D]. 合肥: 中国科学技术大学, 2020: 29-85.
TU B H. Research on data correction and inspection method of atmospheric aerosol multi-angle polarization imager[D]. Hefei: University of Science and Technology of China, 2020: 29-85(in Chinese).

[9] 李扬, 骆冬根, 孙亮, 等. 帧转移型面阵 CCD 相机拖尾评价标准和校正方法[J]. 光子学报, 2022, 51(4): 0404001.
LI Y, LUO D G, SUN L, et al. Evaluation criteria and correction method of frame transfer CCD camera smearing[J]. Acta Photonica Sinica, 2022, 51(4): 0404001(in Chinese).

[10] 翁建文, 袁银麟, 郑小兵, 等. 星载偏振相机帧转移模糊效应校正方法研究[J]. 光学学报, 2019, 39(12): 1228005.
WENG J W, YUAN Y L, ZHENG X B, et al. Correction method for frame-transfer blurring effect of spaceborne polarization camera[J]. Acta Optica Sinica, 2019, 39(12): 1228005(in Chinese).

[11] European Machine Vision Association. Implementation of EMVA 1288 standard release 4.0 for characterization of image sensors [EB/OL].(2021-06-16)[2023-08-04].https://www.emva.org/wp-content/uploads/EMVA1288Linear_4.0Release.pdf.

[12] 曹金彦, 贺小军, 陈茂胜, 等. 航天 TDICCD 相机谱间串扰分析与处理[J]. 半导体光电, 2024, 45(6): 953-959.
CAO J Y, HE X J, CHEN M S, et al. Spectral crosstalk analysis and processing for aerospace TDICCD camera[J] Semiconductor Optoelectronics, 2024, 45(6): 953-959(in Chinese).

[13] 关怡然, 李新, 张权, 等. 高光谱紫外辐照度仪 CCD 电路系统的设计及测试[J]. 量子电子学报, 2024, 41(2): 246-256.
GUAN Y R, LI X, ZHANG Q, et al. Design and test of CCD circuit system of hyperspectral UV irradiance monitor[J]. Chinese Journal of Quantum Electronics, 2024, 41(2): 246-256(in Chinese).

[14] 黄辉, 熊健, 刘祥, 等. 基于图像信噪比的 PSP 图像平均幅数确定方法[J]. 计算机测量与控制, 2023, 31(12): 237-243.
HUANG H, XIONG J, LIU X, et al. PSP image average number determination method based on image signal-to-noise ratio[J]. Computer Measurement & Control, 2023, 31(12): 237-243(in Chinese).

[15] 朱树旺, 赵开春. 分焦平面偏振图像传感器信噪比标定[J]. 光学精密工程, 2023, 31(7): 1012-1021.
ZHU S W, ZHAO K C. Calibration of signal-to-noise ratio of sub focal plane polarization image sensors[J]. Optics and Precision Engineering, 2023, 31(7): 1012-1021(in Chinese).

Comparison and evaluation of data correction methods for directional polarimetric cameras

WANG Yi[1, 2], MENG Binghuan[1, 2, *], YE Chen[1, 2], LUO Donggen[1, 2], YAO Pingping[1, 2], HONG Jin[1, 2]

(1. Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei 230031, China;
2. Key Laboratory of Optical Calibration and Characterization, Chinese Academy of Science, Hefei 230031, China)

Abstract: The data correction process of a directional polarimetric camera (DPC) primarily includes two components: background correction and frame transfer effect correction. The image acquisition process of the photodetector is a random sampling procedure, and mathematical processing methods are applied to the sampling results during data correction, which introduces noise. This paper mainly discussed the influence of dark background correction and frame transfer effect correction on the signal-to-noise ratio (SNR) during DPC data processing. Modeling and simulation of the SNR for two frame transfer effect correction methods were conducted. Validation was performed using actual data from the DPC onboard an atmospheric monitoring satellite. The results demonstrate that both background correction and frame transfer effect correction increase noise, thus reducing the SNR. Background correction should involve more than five background data sets. Compared to the two frame transfer effect correction methods, the sum of column ratio method outperforms the dark row method, achieving a maximum increase in SNR of 15.8.

Keywords: multi-angular polarimeter directional polarimetric camera; data correction; signal-to-noise ratio; frame transfer effect; charge coupled device

Received: 2023-08-17; **Accepted**: 2023-10-13; **Published Online**: 2023-11-17 14:48

URL: link.cnki.net/urlid/11.2625.V.20231116.1500.002

Foundation item: K.C.Wong Education Foundation "International Team of Advanced Polarization Remote Sensing Technology and Applications" (GJTD-2018-15)

* **Corresponding author**. E-mail: bhmeng@aifom.ac.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0507

复合材料蒙皮搭接结构电气性能分析

杨占刚[1,2]，魏宇昊[1]，石旭东[1,*]

(1. 中国民航大学 电子信息与自动化学院，天津 300300；2. 天津市城市空中交通系统技术与装备重点实验室，天津 300300)

摘 要：现代商业飞机中复合材料的使用率逐步上升，然而较低的电导率限制了其在蒙皮结构中的应用，一般还需搭接额外的导电结构以增强蒙皮内电流的导通性能。为分析复合材料飞机蒙皮电气特性及其影响因素，对蒙皮结构之间的搭接进行建模，并通过建立等效电路分析搭接结构电气特性影响机理。基于所构建的复合材料飞机蒙皮典型搭接结构模型，计算结构件类型、搭接方式对蒙皮电气特性的影响，并分析不同搭接方式受雷电间接效应的影响情况。计算及仿真结果表明:结构件的材质、形状及尺寸对蒙皮电气性能有一定影响，采取恰当的搭接方式可有效减小蒙皮结构阻抗、提升飞机雷电间接效应屏蔽性能。

关 键 词：复合材料；飞机蒙皮；搭接方式；电气特性；雷电间接效应

中图分类号：V261.97

文献标志码：A **文章编号**：1001-5965(2025)10-3313-11

随着航班运载能力需求的不断提升，飞机制造过程中的减重设计愈发重要[1]。目前，飞机蒙皮框架由最初的金属材料占主导逐渐变更为以碳纤维增强复合材料 (carbon fiber reinforced plastics, CFRP) 为主的构造方式[2]。与传统的金属材料相比，CFRP具有质量轻、使用时间长、环境适应性好等优势，能有效减轻飞机质量，满足飞机运载能力增加的需求[3]。

传统飞机的金属机身结构一般为电的良导体，在机身蒙皮的电流传导过程中起着低阻抗通路的作用，可视为理想的接地等电位体，因此,在飞机闪电防护及静电屏蔽等方面均有良好的表现[4]；而CFRP电阻各向异性，其最小方向的阻值仍比铝材料阻值高60倍[5]，直接用于电气搭接会引入远大于标准值的电阻，造成蒙皮电流回路阻抗偏大，影响电流的正常导通。因此,需通过搭接额外的导电结构，提升蒙皮内电流导通的性能，降低CFRP较低的导电率带来的影响[6]。

根据蒙皮装配流程，除CFRP外，复合材料飞机蒙皮的组成还包括金属连接件、延展金属箔、树脂、紧固件及凹形垫圈等，这些组成部分统称为结构件，一般通过搭接技术连为一个整体[7]。其中，搭接指的是使用机械或化学方法，将蒙皮内各板材结构件，通过紧固件及凹形垫圈实现两两或多个结构之间的连接，形成一条完整稳定的低阻抗电气通路，以保障电流在蒙皮内的导通及泄放[8]。若蒙皮各结构之间搭接不良，会引入较大的结构阻抗，给电流的传导与泄放、抗雷击和电磁屏蔽等方面带来不利影响[9]。

针对复合材料飞机蒙皮搭接特性的研究，主要集中于蒙皮整体回流网络的电气性能分析和力学性能表现，而较少关注蒙皮搭接结构对电气性能的影响[10-12]。Revel等通过对一定频率范围内飞机的电流回流进行了数值模拟，对复合材料蒙皮搭接及机体内部复合材料接地网和回流网进行了结构优化[13]。Bandinelli等研究了结构类型的影响，分析了

收稿日期：2023-08-08；**录用日期**：2023-09-28；**网络出版时间**：2023-10-17 19:02
网络出版地址：link.cnki.net/urlid/11.2625.V.20231016.1519.001
基金项目：国家自然科学基金 (51407185)；天津市航空装备安全性与适航技术创新中心开放基金 (JCZX-2023-KF-06)
* **通信作者**. E-mail：xdshi@cauc.edu.cn

引用格式：杨占刚，魏宇昊，石旭东. 复合材料蒙皮搭接结构电气性能分析 [J]. 北京航空航天大学学报，2025，51（10）：3313-3323.
YANG Z G，WEI Y H，SHI X D. Electrical properties analysis of composite materials skin bonding structures[J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51（10）：3313-3323（in Chinese）.

结构件的选择对蒙皮所在回路电气回流性能的影响，但局限于整体接地结构的分析，而未考虑蒙皮结构搭接件选择的影响[14]。Goleanu 等对包含复合材料蒙皮搭接结构的接地回流网络建模，通过连接电缆的电流值和压降变化分析最佳电流返回路径，但同样未考虑蒙皮搭接结构自身电气性能的影响[15-16]。隋政等使用部分元等效电路算法，分析了结构件的阻抗对接地回流网内压降的影响，但未对蒙皮结构之间搭接阻抗的产生原因进行深入分析[17-18]。

由上述分析可见，数值仿真计算可以较全面地考虑各单独结构件和各结构连为一个整体后的电学特性，也可以有效计算搭接结构受雷电侵袭后的间接效应防护性能[19-20]。大量学者利用该方法对复合材料飞机接地回流网络电气性能做了分析，但在复合材料飞机蒙皮自身电气性能及相关影响因素方面仍需进一步研究。因此，本文针对复合材料飞机蒙皮搭接结构进行电气特性研究，分析结构件类型、搭接方式对蒙皮电气性能和飞机雷电间接效应屏蔽性能的影响，为蒙皮电学性能的提升提供一定参考。

1 蒙皮搭接结构电气特性影响因素

1.1 蒙皮搭接结构电流路径及影响机理

复合材料飞机蒙皮搭接结构实物如图 1 所示，实际的金属连接件有铝板、钛板等，CFRP 板材有 CMS-CP-308、CMS-CP-309 等型号。

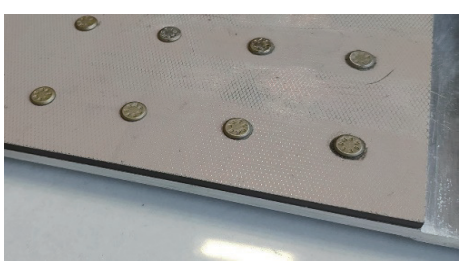


图 1 复合材料飞机蒙皮搭接实物

Fig. 1 Physical odject of composite aircraft skin lap

对于常见单盖板对接蒙皮搭接结构，电流通过延展金属箔的侧面进入当前蒙皮组件，此后大多数电流通过紧固件进入下方的金属连接件中，在金属连接件中再流向与下一块蒙皮相搭接的区域，流入该处的绝大部分电流再通过第 2 组紧固件流进下一块蒙皮的延展箔中，最后，从第 2 块金属箔的侧面流出，其路径示意如图 2 所示。

在电流的传导过程中，部分电流在经过左侧紧固件后，会进入 CFRP 板材中，即与螺杆直接接触的部分 CFRP 板材与螺杆间存在并联关系，共同传导电流进入下端金属连接件，此时复合材料飞机蒙皮搭接结构的电路简化示意如图 3 所示。其中，Z_1 和 Z_1' 表示左右两侧延展金属铜箔的阻抗值，Z_2 和 Z_2' 为左右两侧与 CFRP 板材直接接触的紧固件螺杆的阻抗值，Z_3 与 Z_3' 为左右两侧螺杆进入金属连接件部分的阻抗值，Z_4 与 Z_4' 为左右两侧 CFRP 板材阻抗值，Z_5 为金属连接件的阻抗值。

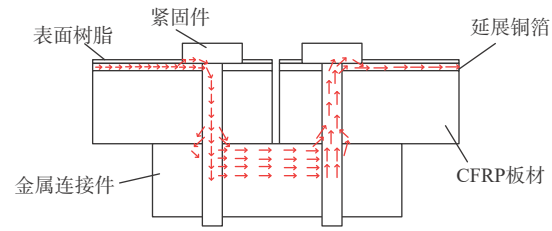


图 2 蒙皮内部电流主要路径示意图

Fig. 2 Schematic diagram of main current path inside skin

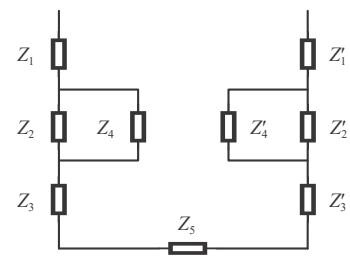


图 3 CFRP 板材与金属连接件等效电路

Fig. 3 Equivalent circuit of CFRP sheets and metal connectors

图 3 中，Z_2 与 CFRP 板材阻抗 Z_4 并联，设螺杆支路上的电流为 I_1，CFRP 板材内流通的电流为 I_2，则有 $I_1Z_2=I_2Z_4$。设输入并联结构的总电流为 I，且满足 $I=I_1+I_2$，则并联电路的熵增为

$$P = P_1 + P_2 = I_1^2 Z_2 + I_2^2 Z_4 \tag{1}$$

根据并联电路分流特点，式 (1) 可展开为

$$P = (I - I_2)^2 Z_2 + I_2^2 Z_4 = I^2 Z_2 + I_2^2 Z_2 + I_2^2 Z_4 - 2 I I_2 Z_4 \tag{2}$$

考虑到注入蒙皮的并联电路总电流 I 一般为恒定值，Z_4 数值较大且保持不变，故流经 CFRP 板材的电流 I_2 一般较小，因此，紧固件螺杆的阻抗 Z_2 对并联区域熵增的影响最大，同时对流入 CFRP 板材内的电流大小也会产生影响。

此外，受空间布局的影响，复合材料飞机蒙皮在不同舱段处会采用不同的搭接方式，如图 4 所示。当搭接方式不同时，引入结构件的数量、接触表面的数目不同，导致蒙皮搭接结构内的电流流向、整体结构阻抗等电气性能表现也不同。

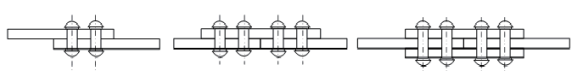


图 4 复合材料飞机蒙皮常见搭接方式

Fig. 4 Common lap splicing methods for composite aircraft skins

采用单搭接方式时, 蒙皮搭接结构内部阻抗分布可简化如图 5 所示, 整体结构阻抗为

$$Z = Z_1'' + \frac{Z_2''Z_4''}{Z_2'' + Z_4''} + Z_3'' + Z_5'' \tag{3}$$

式中: Z_1''、Z_2''、Z_3''、Z_4''及Z_5''分别为单搭接时延展金属铜箔、与 CFRP 板材直接接触的紧固件螺杆、进入金属连接件中紧固件螺杆、CFRP 板材及金属连接件的阻抗。由式 (3) 可知与单盖板对接方式相比, 单搭接时蒙皮整体结构的阻抗值会发生明显变化。

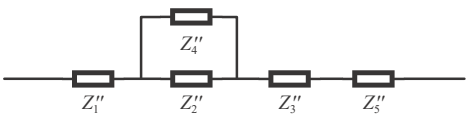


图 5　单搭接时等效电路

Fig. 5　Equivalent circuit for single lap connection

1.2　复合材料对飞机电磁屏蔽作用机理

复合材料飞机电磁屏蔽特性与其蒙皮搭接结构及外部电磁脉冲相关, 其中, 外界施加的电磁脉冲为电磁干扰源, 机载设备或内部线缆为被干扰对象, 电磁干扰传播途径一般分为传导耦合和辐射耦合[21]。此外, 电磁脉冲耦合一般分为前门耦合和后门耦合, 即前者通过天线耦合, 后者通过孔缝耦合。通过电磁屏蔽手段减弱外部电磁脉冲对飞机内部的耦合是对间接效应进行有效防护的方法之一, 其原理可从电磁能量损耗的角度来分析, 电磁波穿透介质的过程如图 6 所示。

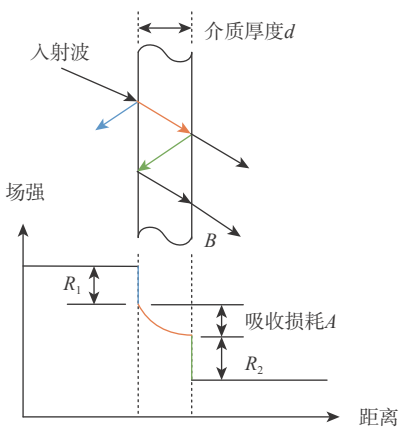


图 6　电磁波穿过介质时的电磁损耗

Fig. 6　Electromagnetic loss in electromagnetic waves passing through medium

电磁波在穿过介质时的电磁损耗为 S_E, 则根据图 6 可知电磁损耗 S_E 为

$$S_E = R_1 + A + R_2 + B = R + A + B \tag{4}$$

$$\delta = \sqrt{\frac{2\rho}{\omega\mu}} = \sqrt{\frac{2}{2\pi f\sigma\mu}} = \sqrt{\frac{1}{\pi f\mu\sigma}} \tag{5}$$

$$A = 8.69d/\delta \tag{6}$$

$$R = 1 - 2\sqrt{2\omega\varepsilon_0/\sigma} \tag{7}$$

式中: R_1 为电磁波穿过第 1 层界面时的损耗; R_2 为电磁波穿过第 2 层界面时的损耗; R 为反射总损耗; A 为吸收损耗; B 为多次反射修正因子; ρ 为电阻率; ω 为电磁波的角频率; f 为频率; μ 为磁导率; σ 为电导率; d 为电磁波穿透的介质厚度; ε_0 为相对介电常数。

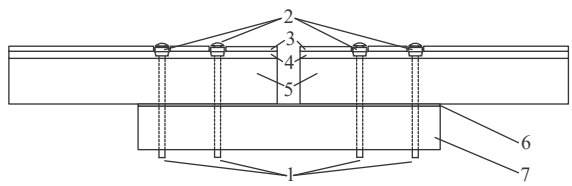
当吸收损耗 A 大于 10 dB 时, 可认为多次反射修正因子 B 可忽略[22], 则由式 (5) ～式 (7) 可知, 电导率 σ 越低, 趋肤深度 δ 越大, 吸收损耗 A 与反射损耗 R 越小。

由于复合材料的电导率低、趋肤深度大, 电流产生的电磁场穿透复合材料蒙皮时的电磁损耗较小, 导致飞机电磁屏蔽性能较弱。此外, 由于飞机蒙皮在搭接时会存在一定宽度的缝隙, 外部电磁脉冲将会通过这些缝隙对飞机进行后门耦合, 进一步降低了飞机间接效应屏蔽效能。

2　结构件类型对蒙皮电气特性影响

2.1　蒙皮典型搭接结构建模

基于蒙皮搭接结构典型构型, 利用 COMSOL 软件构建复合材料飞机蒙皮典型搭接结构模型, 采用单盖板对接, 紧固件使用带凹形垫圈的沉头螺栓, 各结构件之间的组合关系如图 7 所示。



1. 螺栓; 2. 凹形垫圈; 3. 树脂; 4. 延展金属箔; 5. CFRP; 6. 金属氧化膜; 7. 金属连接件。

图 7　粗糙凸起点侧视图

Fig. 7　Side view of rough raised points

图 7 中各结构件的尺寸为: 树脂与延展金属箔的高度均为 0.5 mm, CFRP 板材、金属连接件及其表面氧化膜的高度分别为 10.5 mm、5.8 mm 及 0.2 mm, 各六面体结构件的长、宽均为 500 mm×300 mm, 紧固件的尺寸如表 1 所示。

复合材料飞机蒙皮典型搭接结构模型中, 各结构件材质如表 2 所示, 其中, CFRP 的电导率为 1.5×10^4 S/m, 钢的电导率为 4.032×10^6 S/m。

将蒙皮搭接结构接地面设置为金属连接件下底面, 在左右 2 块延展金属铜箔外侧各施加 1 A 的直流源激励, 观测对下端金属连接件及蒙皮搭接结

表 1 螺栓与凹形垫圈尺寸

Table 1 Bolt and concave washer size

参数	数值	参数	数值
沉头螺栓总长度/mm	18	凹形垫圈高度/mm	2.2
沉头螺杆直径/mm	6.4	凹形垫圈内径/mm	6.45
沉头螺栓头直径/mm	9	凹形垫圈外径/mm	19
沉头螺栓头厚度/mm	0.5	凹形垫圈厚度/mm	0.4

表 2 结构件材质及其电导率

Table 2 Materials of structural parts and their electrical conductivity

结构件名称	材质
延展金属箔	铜
CFRP板材	CFRP
金属氧化膜	氧化铝
金属连接件	铝
沉头螺栓	钢
凹形垫圈	钛

构电气性能的影响情况，蒙皮搭接结构内电流流向示意图如图 8 所示。

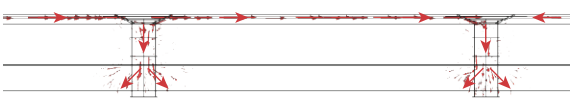


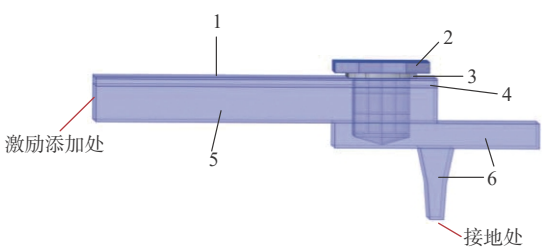
图 8 螺栓附近区域电流流向

Fig. 8 Current flow in area near bolt

由图 8 可知，施加的电流激励通过延展金属铜箔进入蒙皮后，红色箭头所示的内部电流较为集中的通过螺栓进入金属连接件中，部分电流通过具有一定电导率的 CFRP 板材进行传导；同时在紧固件螺杆附近区域，金属连接件与 CFRP 板材间的氧化膜上下两端的电流变化较大，最终电流在金属连接件的下表面接地端流出。

2.2 搭接模型有效性验证

通过与 Simulink 中建立的等效电路法所得电阻值比较来验证本文建立的 COMSOl 模型的有效性。构建典型蒙皮搭接中的单搭接模型如图 9 所示，紧固件使用带平垫圈的平头螺栓，并设置 1 A 电流的激励源及接地。



1. 树脂；2. 螺栓；3. 平垫圈；4. 金属铜层；5. CFRP；6. 氧化铝层。

图 9 单搭接模型侧视图

Fig. 9 Side view of single lap model

图 9 中的各结构件尺寸如下：树脂层、金属铜层、CFRP 层长宽均为 400 mm×100 mm，高度分别为 0.5 mm、2 mm、10 mm。上层氧化铝层长、宽、高分别为 400 mm×62 mm×7 mm，下端为 400 mm×62 mm×27 mm。紧固件尺寸如表 3 所示。

表 3 单搭接模型紧固件尺寸

Table 3 Single lap model fastener size

参数	数值	参数	数值
螺栓总长度/mm	23.14	垫圈高度/mm	1.5
螺杆直径/mm	16	垫圈内径/mm	16
螺栓头边长/mm	13.86	垫圈外径/mm	20
螺栓头厚度/mm	3	垫圈厚度/mm	1.5

其中，设置树脂、铜、铝电导率分别为 1×10^{-14}，5.998×10^{7}，3.03×10^{7} S/m。在 Simulink 中构建如图 10 所示等效电路，各等效电阻如表 4 所示。

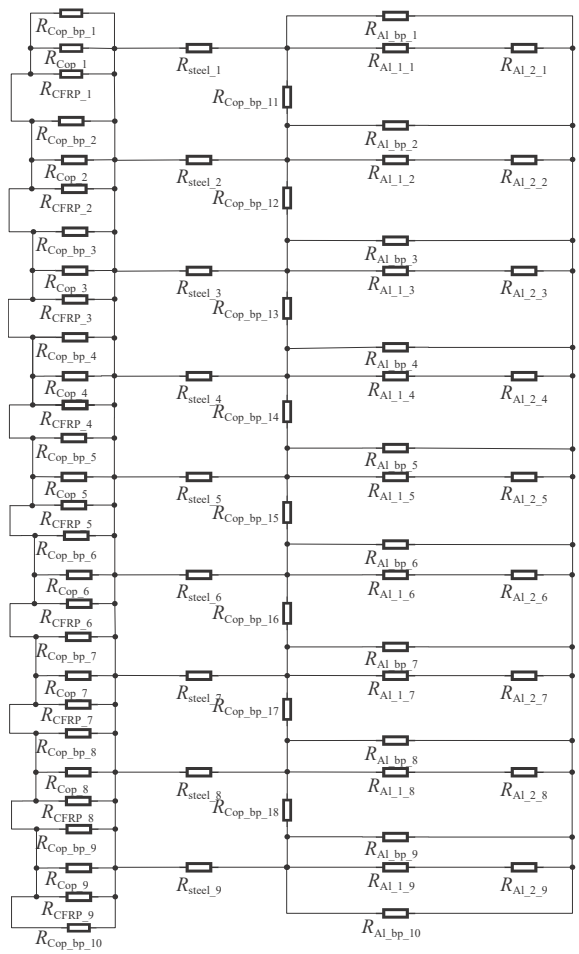


图 10 单搭接模型等效电路

Fig. 10 Equivalent circuit of single lap model

由于树脂层阻值过大并且与铜层并联后对整体阻值影响较小，因此，可将其忽略。R_{cop}、R_{CFRP} 分别表示铜层、CFRP 层从激励源处到每个紧固件对应的通道电阻，$R_{cop_by_1}$～$R_{cop_by_10}$ 表示剩余铜层从激

表 4　单搭接模型等效阻值情况

Table 4　Equivalent resistance of single lap model

参数	阻值/Ω	参数	阻值/Ω
R_{cop}	3.96×10^{-5}	R_{Al_1}	1.10×10^{-5}
$R_{cop_by_1}$～$R_{cop_by_10}$	1.58×10^{-5}	$R_{Al_by_1}$～$R_{Al_by_10}$	1.89×10^{-5}
R_{CFRP}	3.17×10^{-2}	R_{Al_2}	5.96×10^{-6}
$R_{cop_by_2}$～$R_{cop_by_9}$	3.96×10^{-5}	$R_{Al_by_2}$～$R_{Al_by_9}$	4.72×10^{-6}
R_{steel}	2.04×10^{-5}	$R_{cop_by_11}$～$R_{cop_by_18}$	1.25×10^{-5}

励处到紧固件两侧通道电阻, $R_{cop_by_11}$～$R_{cop_by_18}$ 表示铜层在紧固件两端通道电阻, R_{steel} 表示螺栓与垫圈的电阻, R_{Al_1} 与 R_{Al_2} 分别表示紧固件经过铝层到接地处的通道电阻, R_{Al_by} 表示剩余铝层到接地处的通道电阻。

最终, Simulink 计算等效电路阻值为 4.405×10^{-6} Ω, COMSOL 中的仿真结果为 4.220×10^{-6} Ω, 基本保持一致, 验证了 COMSOL 中本文构建模型的准确性。

2.3　金属连接件材质与厚度影响分析

金属连接件的材质与结构尺寸会对蒙皮搭接结构的电气特性产生一定影响, 常见的连接件材质有铝、铜、钛等, 其电导率如表 5 所示。

表 5　金属连接件材质及对应电导率

Table 5　Metal connector material and corresponding conductivity

材质	电导率/(S·m^{-1})
铜	5.998×10^{7}
铝	3.03×10^{7}
氧化铝	0.01
钛	7.41×10^{5}

其中, 对铝表面生成的氧化铝薄膜进行额外考虑, 同时增加 CFRP 板材作为连接件进行对比分析, 获得使用不同材质连接件时复合材料飞机蒙皮整体搭接结构的阻抗模值如图 11 所示。

由图 11 可知, 随着金属连接件电导率的减小, 蒙皮整体搭接结构的阻抗模值逐渐增加; 当不考虑氧化铝的存在时, 由于不同材质的连接件电导率差异较小, 导致整体结构的阻抗变化并不明显; 当引入电导率为 0.01 S/m 的氧化铝时, 即在金属连接件与 CFRP 板材之间引入了较大的膜电阻, 引起整体搭接结构的阻抗模值明显上升。

出于减重考虑, 金属连接件多选择铝制结构, 但需对其表面氧化膜进行处理以减小膜电阻。对铝制金属连接件的厚度影响进行分析, 不同厚度蒙皮搭接结构的阻抗模值及铝板-氧化铝接触面电流密度模的极大值变化, 如图 12 所示。

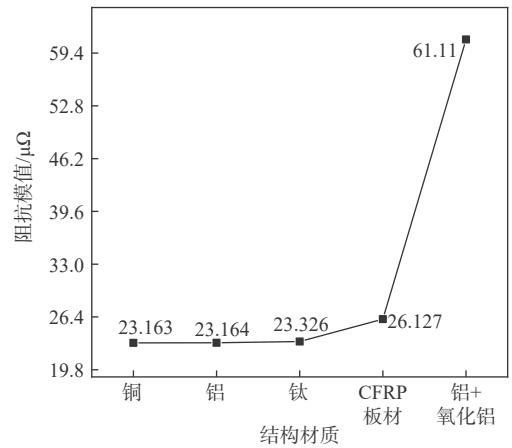
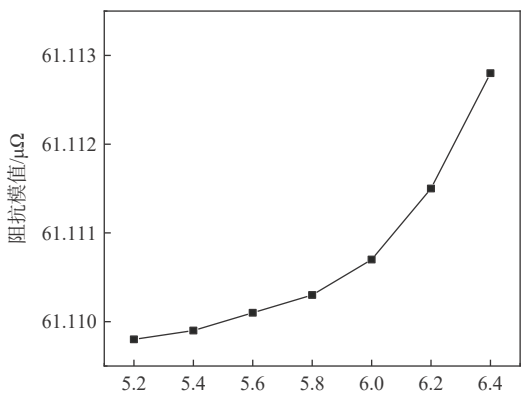


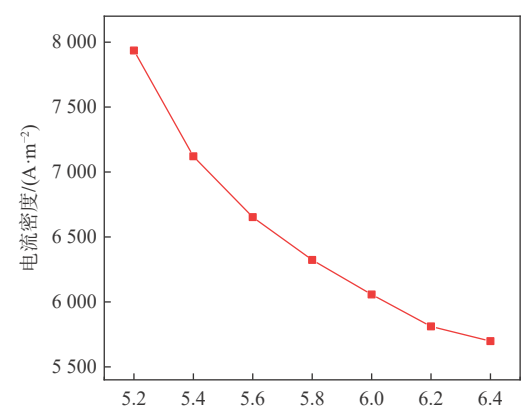
图 11　金属连接件材质对蒙皮整体结构阻抗影响

Fig. 11　Effect of metal connector material on overall structural impedance of skin

由图 12 可知, 随着金属连接件厚度增加, 搭接结构的阻抗模值逐渐上升且上升速率逐渐加快, 同时接触面上电流密度模的极大值将降低。



(a) 蒙皮搭接结构阻抗变化



(b) 铝-氧化铝接触面上电流密度变化

图 12　金属连接件材质对蒙皮整体结构阻抗影响

Fig. 12　Effect of metal connector material on overall structural impedance of skin

2.4　金属箔材质与厚度影响分析

延展金属箔是与进入蒙皮内的电流首先接触

的结构件，也是电流主要流经结构之一，其材质、尺寸对蒙皮导电性能影响较大。常用的金属箔材质有铝、铜、镍、钛等，电导率如表6所示。

表6 金属箔材质及对应电导率

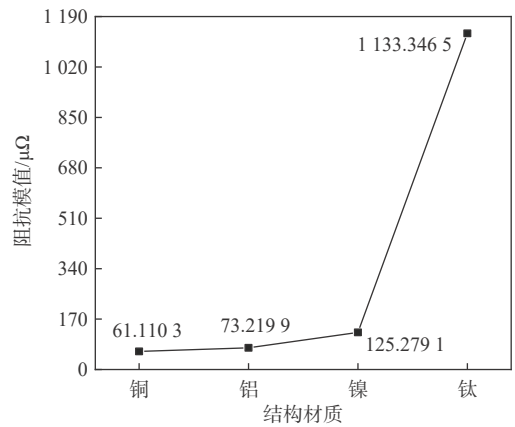
Table 6 Metal foil material and corresponding conductivity

材质	电导率/(S·m^{-1})
铜	5.998×10^{7}
铝	3.03×10^{7}
镍	1.43×10^{7}
钛	7.41×10^{5}

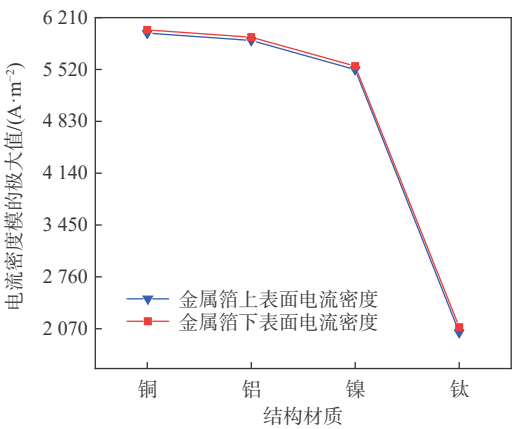
不同材质金属箔蒙皮搭接结构的阻抗变化及金属箔上、下表面电流密度模值变化情况如图13所示。可见，当使用高电导率的铜制、铝制金属箔时，金属箔表面流过的电流较大，蒙皮整体搭接结构阻抗模值较小，电气性能越好。

金属箔的厚度也会对蒙皮搭接结构的电气性能产生影响，计算不同厚度延展金属铜箔蒙皮整体搭接结构的阻抗模值，结果如图14所示。

由图14可知，金属铜箔的厚度与蒙皮整体搭接结构的阻抗模值成反比。考虑到厚度增加会提升蒙皮质量，因此，需要在保证阻抗满足需求的前提下，尽量减小延展金属铜箔的厚度。



(a) 蒙皮搭接结构阻抗变化



(b) 金属箔上下表面电流密度变化

图13 金属连接件材质对蒙皮整体结构阻抗影响

Fig. 13 Effect of metal connector material on overall structural impedance of skin

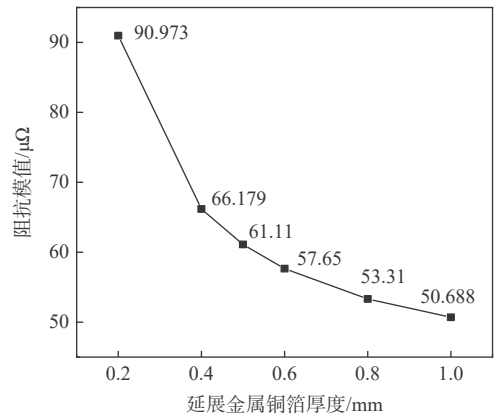
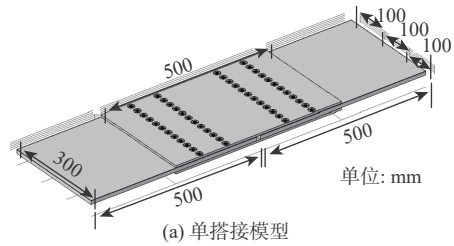


图14 延展金属铜箔厚度对蒙皮阻抗影响

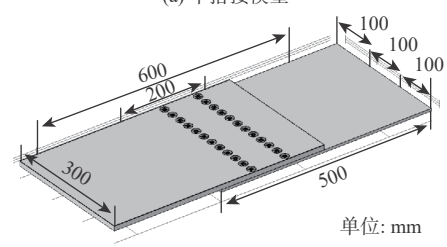
Fig. 14 Effect of extended metal copper foil thickness on skin impedance

3 搭接方式对蒙皮电气特性影响

除典型搭接结构模型中采用的单盖板对接外，飞机蒙皮还可采取单搭接和双盖板对接的搭接方式，对应模型如图15所示。



(a) 单搭接模型



(b) 双盖板对接模型

图15 蒙皮搭接结构模型

Fig. 15 Model of skin lap structure

与典型搭接结构模型相比，图15(a)中的单搭接模型去除了另外一侧的CFRP板材、延展金属铜箔及表面树脂，且紧固件减少为2列；图15(b)中的双盖板对接模型增加了最上层的金属连接件及其表面氧化膜。

对单盖板对接时的蒙皮整体搭接结构进行电气特性分析，结果如图16所示。

由图16可知，当采取单盖板对接的搭接方式时，表面电流密度在螺杆附近区域数值较大，达到

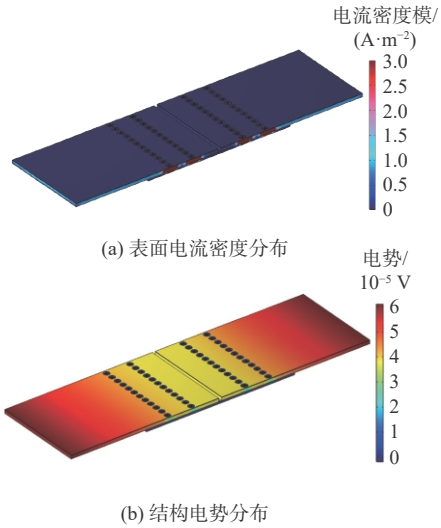


图 16　单盖板对接时蒙皮结构电气性能表现

Fig. 16　Electrical performance of skin under single cover docking

6.045×10^4 A/m^2。整体搭接结构上的电势在表面树脂及 CFRP 板材上分布较大。

3.1　单搭接方式

单搭接时的蒙皮电气性能如图 17 所示。采取该种搭接方式时, 除 CFRP 板材与金属连接件的接触区域, 结构电势分布规律基本与单盖板对接时一致。但采用单搭接时 CFRP 板材会承担更多的电流传输功能, 其表面电流密度更大, 模的极大值也从单盖板对接时的 1.265×10^4 A/m^2 增加到 3.075×10^4 A/m^2; 此外, 单搭接还会使 CFRP 板材与金属连接件的接触面上产生额外电场, 对邻近区域产生电磁干扰影响。

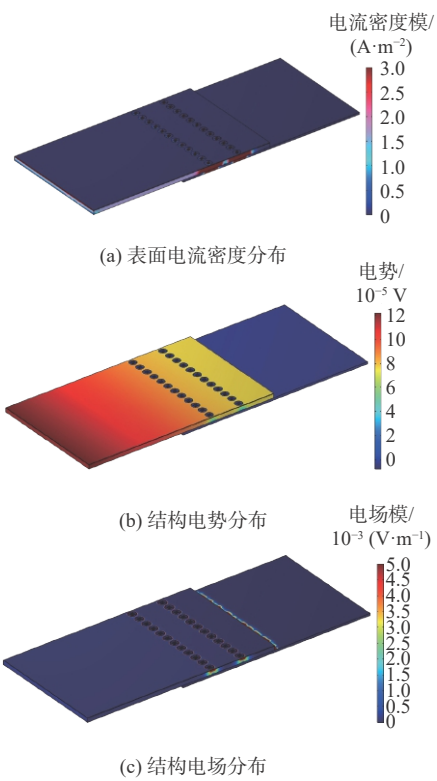


图 17　蒙皮单搭接时的电气性能表现

Fig. 17　Electrical performance of skin under single lap joint

3.2　双盖板对接方式

双盖板对接方式下, 考虑到上层金属连接件表面也会产生氧化膜, 当其与紧固件等接触时会对表面电流的导通产生影响, 因此, 在金属连接件表面增加一层 0.2 mm 厚的氧化铝薄膜。

采用双盖板对接方式时, 表面电流密度分布、结构电势分布及电场模分布与另外 2 种搭接方式相似; 差异在于采用双盖板对接时上层树脂与金属连接件接触区域电场模值较大, 达到 8.456×10^3 A/m^2, 需额外对该区域增加保护措施。同时上层金属连接件、尤其是在与紧固件的搭接区域中电流密度模值也较大。

对该搭接方式下的蒙皮内部电流进行进一步分析, 取左侧两排紧固件区域的电流流向如图 18 所示。

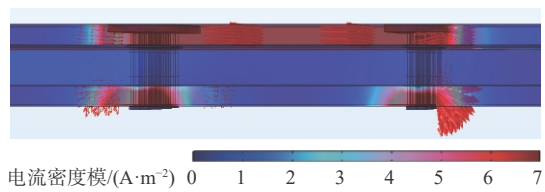


图 18　双盖板对接时紧固件附近电流流向及密度分布

Fig. 18　Current flow direction and density distribution near the fastener under double cover docking

由图 18 可知, 当搭接结构上表面增加一层金属连接件后, 从延展金属箔中进入的电流在流经第 1 排紧固件后会直接进入上层连接件, 少量电流通过第 1 排紧固件流入下层金属连接件; 上层连接件内的电流在通过第 2 排紧固件后, 才开始密集的通过螺杆流入下层金属连接件。因此, 若采用此搭接方式, 应加强第 2 排紧固件结构强度的设计; 同时确保上层金属连接件自身结构的阻抗值小于下层金属连接件的阻抗值。

3.3　不同搭接方式下蒙皮自身电气特性对比

通过 3 种搭接方式可知, 不同搭接方式下的电势分布规律基本一致; 无论采用哪种搭接方式, 在结构件的连接处电场模值均较大, 且电导率较低的结构件上电流密度模值较大。因此, 在进行搭接方式设计时, 需要对各结构件连接区域做重点防护。

对 3 种搭接方式下的蒙皮整体结构阻抗模值、电流密度模的极大值、结构电场模的极大值进行对比分析, 结果如表 7 所示。

由表 7 可知, 由于单搭接方式仅有下侧电流导通通道, 蒙皮搭接结构连接处电场模的极大值、表面电流密度模的极大值均为最大。采用双盖板对

表 7 搭接方式对蒙皮电气性能的影响

Table 7 Effect of lap method on electrical performance of skin

搭接方式	蒙皮整体搭接结构阻抗/μΩ	表面电流密度模的极大值/(A·m^{-2})	电场模的极大值/(V·m^{-1})
单搭接	122.216	30 750	0.705
单盖板对接	61.114	12 650	0.396
双盖板对接	20.935	13 870	0.034

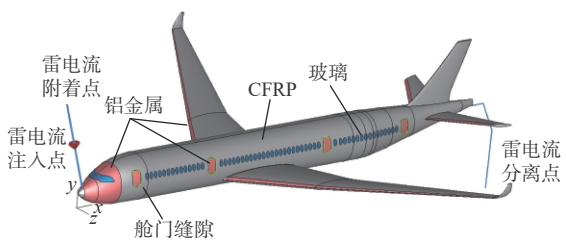
接时，其上层金属连接件可有效增加电流导通通道，其整体结构阻抗模值、电场模的极大值均为最小，使蒙皮的电气性能最为优异。

综上，由于单搭接方式仅对一侧 CFRP 板材进行连接，蒙皮的电气性能表现较差；双盖板对接电气性能优异，但因引入额外的金属连接件导致蒙皮质量增加，并且无保护的外侧金属连接件易受高温、高湿度的影响发生氧化而影响蒙皮电气性能，因此，应优先选择单盖板对接的搭接方式。

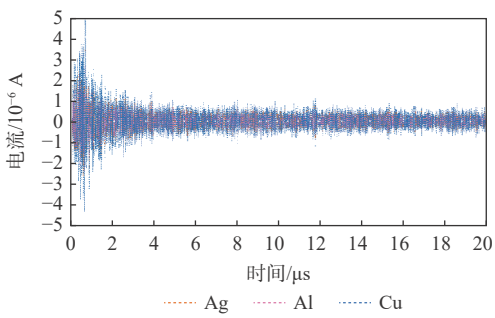
3.4 搭接方式对蒙皮雷电间接效应影响分析

复合材料机身由多段蒙皮搭接而成，不同搭接方式产生的缝隙大小一般不同。考虑到单搭接方式在蒙皮搭接中受结构区域限制较大，因此，舱段搭接区域仅考虑单盖板对接和双盖板对接。

为验证雷电间接效应的准确性，利用 CST 软件构建 A350 模型，仿真雷电流对不同材质屏蔽层下芯线电流的影响，结果如图 19 与表 8 所示，与文献 [23] 保持一致，验证了仿真的有效性。



(a) A350模型



(b) 不同材质下感应的芯线电流

图 19 仿真模型及结果

Fig. 19 Simulation model and results

采用如图 20 所示的双指数函数雷电电磁脉冲[24]波形，其中，雷电上升时间 T_r=1.4×10^{-3} μs，下降时间 T_f=2.5×10^{-2} μs，幅值为 65 000，激励时间为 0.25 μs。

表 8 不同材质屏蔽层下芯线电流情况

Table 8 Core current under different shield materials

材质	电流/10^{-6} A
铜	4.777
铝	2.647
银	3.436

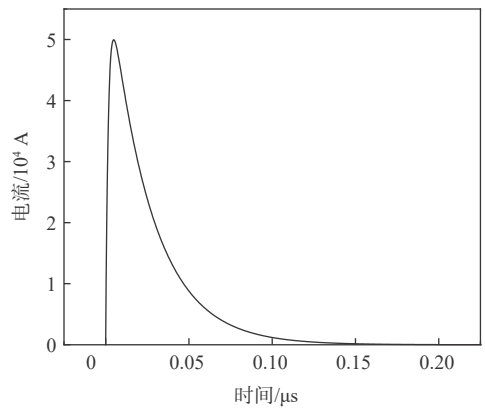


图 20 雷电电磁脉冲信号波形

Fig. 20 Waveform of lightning electromagnetic pulse signal

基于典型的机头击入机尾击出的雷电流路径，通过 CST 软件建立 CFRP 蒙皮飞机整机模型如图 21 所示。

雷电经过机体时空间电磁场情况如图 22 所示。由图可见，电磁能量会通过飞机舷窗玻璃及舱门缝隙进入机身内部。图 22(a) 中红色圆圈部分表示在遭遇雷击时，飞机机头、机翼末端及机尾处电

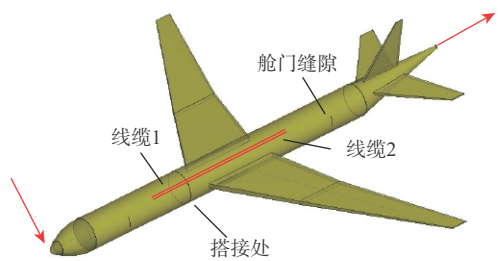
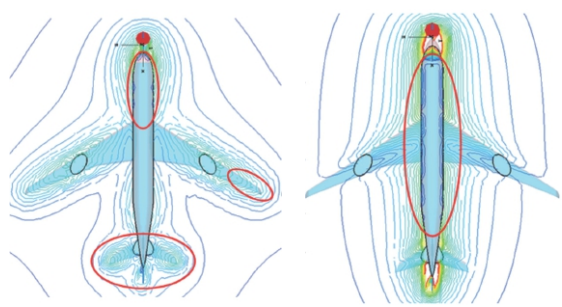


图 21 复合材料飞机整机模型

Fig. 21 Complete composite aircraft model



(a) 机身空间电场情况 (b) 机身空间磁场情况

图 22 机体周围空间电磁场情况

Fig. 22 Electromagnetic field conditions around airframe

场强度较大。图 22(b) 中红色圆圈部分则表示遭遇雷击时, 机身段空间磁场强度较大。

在机身中心处分别敷设单芯线、屏蔽层接地的 RG58 同轴电缆, 对比分析 2 种线缆上因雷电间接效应产生的感应电流, 验证蒙皮搭接模型的雷电间接效应屏蔽性能, 结果如图 23 所示。

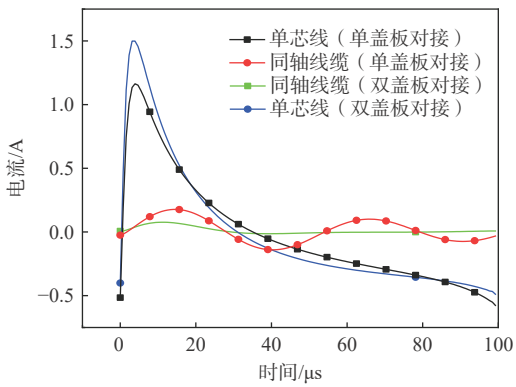


图 23　整机蒙皮内部线缆感应电流

Fig. 23　Cable-induced current inside entire aircraft skin

由图 23 可知, 2 种搭接方式下同轴线缆的感应电流均远小于单芯线。

对当前复合材料飞机中常用的同轴线缆, 双盖板对接时的感应电流峰值最小, 单盖板对接时的感应电流峰值最大。

由表 9 可知, 蒙皮采用双盖板对接时的线缆感应电流更小, 即采取双盖板对接的蒙皮搭接方式对雷电间接效应的屏蔽性能表现最优; 且双盖板对接情况下, 紧固件处的电流密度也较小, 与前述搭接结构自身电气特性分析结果相一致。

表 9　搭接方式对线缆感应电流和紧固件电流的影响

Table 9　Effect of lap splice method on cable-induced current and fastener current

搭接方式	同轴线缆感应电流峰值/mA	紧固件电流密度峰值/(dBA·m^{-1})
双盖板对接	76.519	80.938
单盖板对接	177.062	88.414

4　结　论

本文通过建立复合材料飞机蒙皮搭接结构模型, 对蒙皮的电气特性进行研究, 以优化结构内电流传导性能为目标, 分析影响蒙皮搭接结构阻抗等电气参数的因素, 得到以下结论:

1) 蒙皮搭接结构自身电气特性主要受搭接结构件类型、搭接方式影响, 不同的搭接方式对蒙皮雷电间接效应影响也不同。

2) 搭接结构件材质产生的影响主要取决于结构件自身的电导率, 综合考虑结构自身质量所带来的影响, 选取铝制金属连接件时最适宜, 但需进一步对铝制结构表面氧化膜进行处理; 同时金属连接件和延展金属箔结构尺寸也会对蒙皮搭接结构阻抗值产生影响。

3) 不同搭接方式对蒙皮电气特性影响较大, 采用双盖板对接时搭接结构自身电气性能表现最优; 整机蒙皮搭接结构采取双盖板对接时, 飞机雷电间接效应的屏蔽性能也更好, 但综合考虑电学特性及结构功能等多种因素, 应优先选择单盖板对接的搭接方式。

致　谢　感谢各位老师及同学们在本文理论探讨、实验仿真及写作过程中提供的宝贵建议和帮助。在此, 向参与和关心本文工作的朋友们表示诚挚的谢意!

参考文献（References）

[1] JOSÉ A, T G F, BENTO S D M, et al. An innovative approach for integrated airline network and aircraft family optimization[J]. Chinese Journal of Aeronautics, 2020, 33(2): 634-663.

[2] JAWAID M, THARIQ M. Sustainable composites for aerospace applications[M]. Cambridge: Woodhead Publishing, 2018: 1-18.

[3] PARVEEZ B, KITTUR M I, BADRUDDIN I A, et al. Scientific advancements in composite materials for aircraft applications: a review[J]. Polymers, 2022, 14(22): 5007.

[4] JONES C E, SZTYKIEL M, PENA-ALZOLA R, et al. Correction: grounding topologies for resilient, integrated composite electrical power systems for future aircraft applications[C]//Proceedings of the AIAA Propulsion and Energy Forum. Reston: AIAA, 2019.

[5] 赵中杰. 结构-导电复合材料的制备及其导电性能研究[D]. 哈尔滨: 哈尔滨工业大学, 2019.
ZHAO Z J. Preparation and conductivity of structure-conductive composites[D]. Harbin: Harbin Institute of Technology, 2019(in Chinese).

[6] CALADO E A, LEITE M, SILVA A. Selecting composite materials considering cost and environmental impact in the early phases of aircraft structure design[J]. Journal of Cleaner Production, 2018, 186: 113-122.

[7] AMAN F, CHERAGHI S H, KRISHNAN K K, et al. Study of the impact of riveting sequence, rivet pitch, and gap between sheets on the quality of riveted lap joints using finite element method[J]. The International Journal of Advanced Manufacturing Technology, 2013, 67(1): 545-562.

[8] 康永刚, 李春生, 陈希多, 等. 航空大壁板装配连接局部变形数值建模与仿真分析[J]. 航空制造技术, 2020, 63(3): 45-52.
KANG Y G, LI C S, CHEN X D, et al. Numerical modeling and simulation analysis of local deformation for aircraft large-walled panel assembly connection[J]. Aeronautical Manufacturing Technology, 2020, 63(3): 45-52(in Chinese).

[9] JONES C E, NORMAN P J, GALLOWAY S J, et al. Electrical model of carbon fibre reinforced polymers for the development of electrical protection systems for more-electric aircraft[C]//Proceedings of the 18th European Conference on Power Electronics and

Applications. Piscataway: IEEE Press, 2016: 1-10.

[10] 刘国春, 郭荣辉, 秦文峰. 民用飞机复合材料结构制造与维修[M]. 北京: 清华大学出版社, 2020: 19.
LIU G C, GUO R H, QIN W F. Manufacture and maintenance of composite structure of civil aircraft[M]. Beijing: Tsinghua University Press, 2020: 19(in Chinese).

[11] MALINOWSKI P H, WANDOWSKI T, OSTACHOWICZ W M. Characterisation of CFRP adhesive bonds by electromechanical impedance[C]//Proceedings of the Health Monitoring of Structural and Biological Systems. Belingham: SPIE, 2014: 906415.

[12] MALINOWSKI P H, OSTACHOWICZ W M, BRUNE K, et al. Study of electromechanical impedance changes caused by modifications of CFRP adhesive bonds[J]. Fatigue & Fracture of Engineering Materials & Structures, 2017, 40(10): 1592-1600.

[13] REVEL I, PICHE A, PERES G, et al. Modeling strategy for functional current return in large CFRP structures for aircraft applications[C]//Proceedings of the International Symposium on Electromagnetic Compatibility-EMC Europe. Piscataway: IEEE Press, 2008: 1-5.

[14] BANDINELLI M, MORI A, GALGANI G, et al. A surface PEEC formulation for high-fidelity analysis of the current return networks in composite aircrafts[J]. IEEE Transactions on Electromagnetic Compatibility, 2015, 57(5): 1027-1036.

[15] GOLEANU A L, DUNAND M, GUICHON J M, et al. Towards the conception and optimisation of the current return path in a composite aircraft[C]//Proceedings of the IEEE International Systems Conference. Piscataway: IEEE Press, 2010: 466-471.

[16] GOLEANU A L, GUICHON J M, DUNAND M, et al. Design and optimization techniques for the current return path in a composite aircraft[C]//Proceedings of the 14th European Conference on Power Electronics and Applications. Piscataway: IEEE Press, 2011: 1-10.

[17] 刘建英, 隋政, 张起浩, 等. 复合材料飞机接地回流网络建模与阻抗分析[J]. 北京航空航天大学学报, 2021, 47(5): 885-893.
LIU J Y, SUI Z, ZHANG Q H, et al. Modeling and impedance analysis of composite material aircraft grounded return network[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(5): 885-893(in Chinese).

[18] 杨占刚, 隋政, 张起浩, 等. 复合材料飞机接地回流网络网内压降分析[J]. 航空学报, 2022, 43(1): 324859.
YANG Z G, SUI Z, ZHANG Q H, et al. Voltage drop in composite aircraft grounding and current return network[J]. Acta Aeronautica et Astronautica Sinica, 2022, 43(1): 324859(in Chinese).

[19] HEPPE A J. Computation of potential at surface above an energized grid or other electrode, allowing for non-uniform current distribution[J]. IEEE Transactions on Power Apparatus and Systems, 1979, 98(6): 1978-1989.

[20] KOUTEYNIKOFF P. Numerical computation of the grounding resistance of substations and towers[J]. IEEE Transactions on Power Apparatus and Systems, 1980, 99(3): 957-965.

[21] MOUPFOUMA F, TSE W, JALALI M. More about lightning induced effects on systems in a composite aircraft[C]//Proceedings of the SAE Technical Paper Series. Warrendale: SAE International, 2013: 151-162.

[22] 龙奕, 胡皓全, 赵家升, 等. 电磁屏蔽效能分析与实例计算[C]//第17届全国电磁兼容学术会议论文集. 广州: 中国通信学会, 2007: 68-73.
LONG Y, HU H Q, ZHAO J S, et al. Electromagnetic shielding effectiveness analysis and example calculation[C]//Proceedings of the 17th National Conference on Electromagnetic Compatibility. Guangzhou: China Institute of Communications, 2007: 68-73(in Chinese).

[23] 周秦汉. 复合材料飞机雷电间接效应仿真研究[D]. 南京: 南京航空航天大学, 2019.
ZHOU Q H. Simulation study on lightning indirect effect of composite aircraft[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2019(in Chinese).

[24] United States of America Department of Defense. Electromagnetics environmental effects requirements for systems: MIL-STD-464A[S]. Washington, D. C.: United States of America Department of Defense, 2002.

Electrical properties analysis of composite materials skin bonding structures

YANG Zhangang[1, 2], WEI Yuhao[1], SHI Xudong[1, *]

(1. College of Electronic Information and Automation, Civil Aviation University of China, Tianjin 300300, China;
2. Key Laboratory of Technology and Equipment of Tianjin Urban Air Transportation System, Tianjin 300300, China)

Abstract: The use of composite materials in modern commercial aircraft is gradually increasing. However, their low electrical conductivity limits their application in skin structures. To enhance the conductivity of current within the skin, additional conductive structures are often bonded. This study modeled the bonding structures of composite aircraft skins and analyzed the electrical properties and influencing factors by establishing equivalent circuits. Using a typical bonding structure model for composite aircraft skin, the effects of structural types and bonding modes on the electrical properties were calculated, and the impact of lightning indirect effects on different bonding modes was analyzed. The results of calculations and simulations show that the material, shape, and size of structural components affect the electrical properties of the skin. Adopting appropriate bonding modes can effectively reduce structural impedance and enhance the shielding performance against lightning indirect effects.

Keywords: composite material; aircraft skin; bonding mode; electrical properties; lightning indirect effect

Received: 2023-08-08; **Accepted**: 2023-09-28; **Published Online**: 2023-10-17 19:02

URL: link.cnki.net/urlid/11.2625.V.20231016.1519.001

Foundation items: National Natural Science Foundation of China (51407185); Supported by the Open Fund of Tianjin Aviation Equipment Safety and Airworthiness Technology Innovation Center (JCZX-2023-KF-06)

* **Corresponding author**. E-mail: xdshi@cauc.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0569

FAR 25 附录 C 结冰气象探测数据处理方法

张文强[1]，陈奕屹[1]，雷国强[1]，王光雨[1]，支亚非[2]，毛雪瑞[1,*]

(1. 北京理工大学 机电学院，北京 100081；2. 成都飞机工业（集团）有限责任公司，成都 610091)

摘 要：结冰会破坏飞机的气动外形，导致升力下降阻力增加，飞机的操纵性和稳定性下降，严重影响飞行安全。考虑到 FAR 25 附录 C 的使用有一定的地区和季节的局限性，同时更多更复杂的结冰环境的发现需要对附录 C 进行不断地修订，因此，总结和探索气象探测的数据处理方法能够对编制和改进中国的结冰适航审定规章提供指导。介绍附录 C 的编制背景，总结传统结冰气象探测仪器的测量原理及数据处理方法，重点解释了如何将测量到的气象数据转化为可供查询的包线图；从结冰数据库建立、提升仪器测量精度、无人机防冰 3 方面给出当前面临的挑战与未来发展趋势，从而为气象探测技术的发展和未来发展方向提供科学的指导。

关 键 词：FAR25 附录 C；结冰；气象探测；数据处理；防除冰

中图分类号：V219

文献标志码：A **文章编号**：1001-5965(2025)10-3324-13

结冰是影响飞行器安全的重要因素之一，根据对来自美国国家运输安全委员会 (National Transportation Safety Board, NTSB) 和航空安全报告系统 (Aviation Safety Reporting System, ASRS) 的数据统计分析，2009—2018 年间美国共有 1 740 起由恶劣天气引发的航空事故，这其中由于结冰导致的事故占到了 9.6%[1]。飞机必须配备防除冰系统，从而保证飞机在结冰气象环境中能够安全飞行。防除冰系统的设计取决于飞机所遇到的结冰环境，如飞机不同部位的结冰与不同的云层和天气环境有关系[2]，飞机防除冰系统需要的加热功率则和空气中的液态水含量（liquid water content, LWC）等直接相关。因此，需要编制一份大气飞行结冰气象条件指南，用以查询在何种气象条件下可能会发生结冰，美国联邦航空局（Federal Aviation Administration, FAA）推出的 FAR 25 附录 C[3] 就是在这个背景下应运而生。

FAR 25 附录 C 编制首先根据飞机结冰的位置、发生结冰的云层和持续时间等对结冰气象进行分类；然后，确定具有指导意义的结冰条件参数，影响结冰的参数有很多，其定义也比较丰富，历史上还曾经用过结冰强度等参数表征结冰条件，后来逐渐废弃不用。结冰参数的选择最终趋于采用直接与结冰相关、可以用仪器直接测量、不需要复杂处理即可获得的参数，如 LWC、气压高度等信息。在结冰条件参数确定以后，需要给出这些参数之间的内在关联和拟合关系，因此，需要大量收集气象数据，在 FAR 25 附录 C 编制的时代，主要通过飞行试验获得；最后，利用数据处理方法制定出结冰气象条件特征包线。

中国的运输类飞机适航审定规章是在 FAR 25 附录 C 的基础上修订而来。然而，一方面，FAR 25 附录 C 的使用有一定的地区和季节的局限性，其编制更多是参考了美国北部春冬季节的结冰情况，因此，其推广和应用还需要考虑使用者当地的气候环境；另一方面，FAR 25 附录 C 也在不断地增加更多更复杂的结冰环境，因此，需要针对新发现的结冰气象条件对 FAR 25 附录 C 进行修订。

目前已经存在相当数量的文献用于指导如何使用 FAR 25 附录 C[4-7]，但是 FAR 25 附录 C 的形成过程，编制的过程中使用了哪些具体的气象探测数据处理方法，还缺乏比较详细的介绍，处于知其然

收稿日期：2023-09-04；录用日期：2023-12-05；网络出版时间：2023-12-28 15:10
网络出版地址：link.cnki.net/urlid/11.2625.V.20231227.1647.005
*通信作者. E-mail：xmao@bit.edu.cn

引用格式：张文强，陈奕屹，雷国强，等. FAR 25 附录 C 结冰气象探测数据处理方法 [J]. 北京航空航天大学学报，2025，51（10）：3324-3336. ZHANG W Q，CHEN Y Y，LEI G Q，et al. Data processing methodology of icing meteorological detection in FAR 25 Appendix C [J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51（10）：3324-3336（in Chinese）.

而不知其所以然的局面。因此, 本文围绕 FAR 25 附录 C 的制定、各类结冰气象探测仪器的测量原理及相应的数据处理方法展开, 核心在于阐明将观测气象数据向 FAR 25 附录 C 参考包线图的转化。本文对中国自主制定和优化结冰条件特征包线、完善适航审定规章具有一定的指导意义。

1　FAR 25 附录 C 形成背景及其演化过程

过冷微液滴广泛存在于低于 3.6 km 的大气云层中[2]。当环境温度处于−40～0 ℃(233～273 K)时, 云层中含有大量没有自然冻结而维持过冷状态的液滴。飞机在起降过程中穿越云层时, 过冷液滴会与飞机表面撞击而结冰, 从而破坏飞机的气动外型并影响发动机的工作, 对飞行安全构成严重威胁。为保证飞机在结冰的气象环境中仍然能够安全稳定地飞行, 飞机的适航性和运营要求必须考虑结冰气象条件对飞机的影响。从第二次世界大战开始, 发达航空国家开始对大气飞行结冰气象条件和防冰系统进行研究, 形成标准用以指导飞机防冰系统的设计, 提升飞行安全性。

根据飞机不同部位的结冰严重程度, 可以将产生结冰的气象条件进行分类, 然后, 给出每种气象条件中影响结冰参数的相互关系和分布范围。这些参数包括云层水平范围、LWC、过冷液滴平均有效直径(mean effective diameter, MED)和空气温度等。1945—1948 年, 美国国家航空咨询委员会(National Advisory Committee for Aeronautics, NACA)通过采集的结冰云层数据研究确定了基本的结冰条件特征包线, 这些数据包线后来被 FAA 采纳, 成为联邦航空认证法规的基础, 并于 1965 年发布《运输类飞机适航标准》FAR 25 附录 C[3], 给出了飞机能够安全飞行的自然结冰气象条件和各阶段积冰的定义[6]。为提高新型号飞机在结冰条件下飞行的安全等级, FAA 随后颁发了一系列的修正案, 引入运输类飞机在结冰条件下评估性能和操稳品质的适航标准, 给出了起飞、航路、待机、进场和着陆各个阶段积冰的定义[8]。图 1 描述了结冰条件 FAR 25 附录 C 演化过程的时间线。

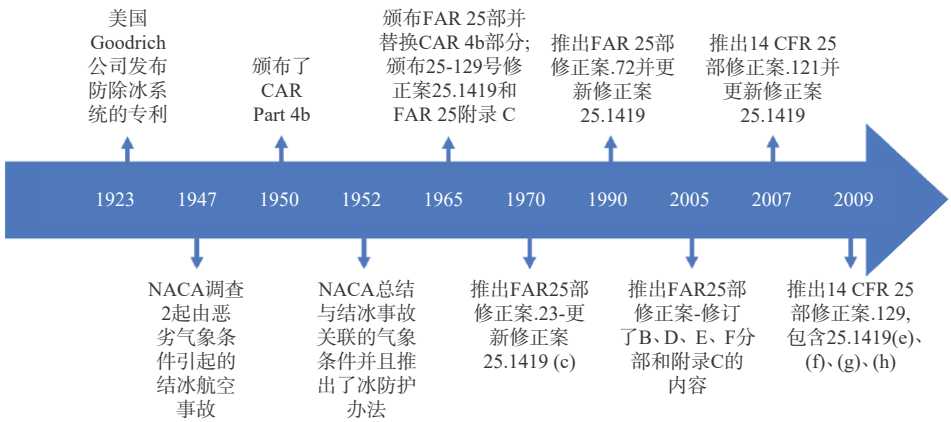


图 1　结冰条件演化过程及防冰系统发展的时间线

Fig. 1　Evolution of icing conditions and development timeline of anti-icing system

FAR 25 附录 C 大气结冰环境自然结冰气象条件包括连续最大结冰、间断最大结冰和最大起飞结冰。连续最大结冰气象条件主要发生在层状云中, 其云层的 LWC、过冷液滴 MED 和空气温度 3 个变量的互相关系如图 2 所示[3]。其中, 气压高度为海拔 6.7 km, 最大垂直范围为 1.98 km, 水平范围为 32.3 km (17.4 n mile)。图 3[3] 为连续最大结冰气象条件的高度-温度包线范围。对于水平范围为 32.3 km 以外的连续最大结冰气象条件, 云层的 LWC 等于图 2 中的水含量乘图 4[3] 中的系数。间断最大结冰气象主要存在于对流云和积云中, 其 LWC、MED 和空气温度 3 个变量的相互关系如图 5 所示[3], 图 5 中的气压高度范围为 1.2～6.7 km, 水平范围为 4.8 km。图 6[3] 为间断最大结冰气象条件的高度-温度包线

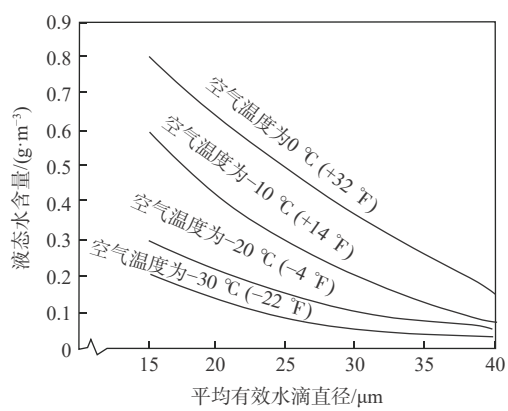


图 2　连续最大(层云)大气结冰状态 LWC 与 MED 的关系[3]

Fig. 2　Continuous maximum (stratiform clouds) atmospheric Icing conditions (liquid water content vs mean effective drop diameter)[3]

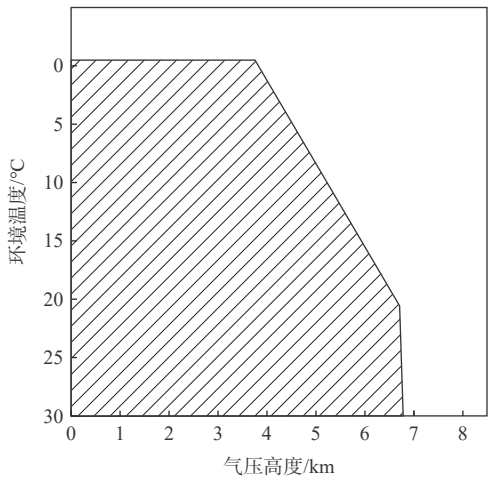


图3 连续最大(层云)大气结冰状态环境温度与气压高度的关系[3]

Fig. 3 Continuous maximum (stratiform clouds) atmospheric icing conditions (ambient temperature vs pressure altitude)[3]

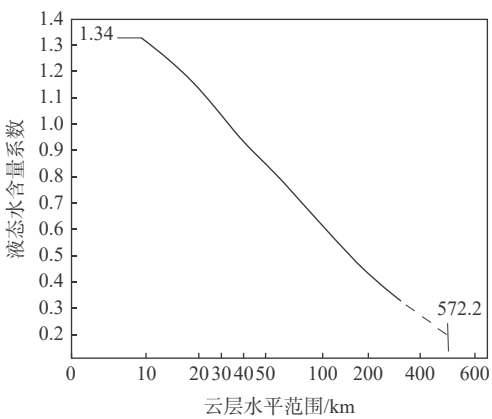


图4 连续最大(层云)大气结冰状态LWC系数与云层水平范围的关系[3]

Fig. 4 Continuous maximum (stratiform clouds) atmospheric icing conditions (liquid water content factor vs cloud horizontal extent)[3]

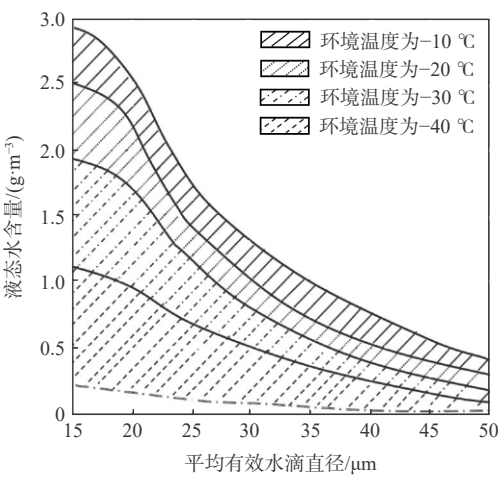


图5 间断最大(积云)大气结冰状态LWC与MED的关系[3]

Fig. 5 Intermittent maximum (cumuliform clouds) atmospheric icing conditions (liquid water content vs mean effective drop diameter)[3]

范围,对于水平范围超出4.8 km(2.6 n mile)的连续最大结冰气象条件,云层的LWC等于图5中的水含量乘图7[3]中纵坐标对应的系数。

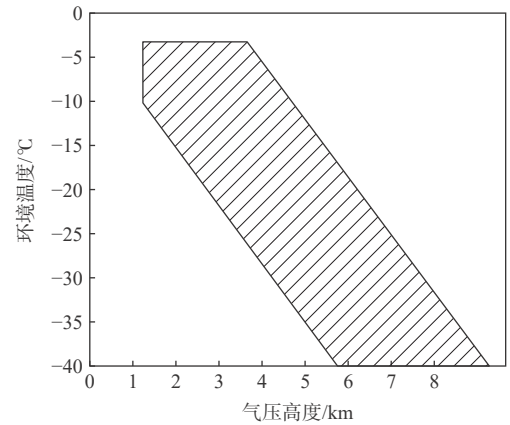


图6 间断最大(积云)大气结冰状态环境温度与气压高度的关系[3]

Fig. 6 Intermittent maximum (cumuliform clouds) atmospheric icing conditions (ambient temperature vs pressure altitude)[3]

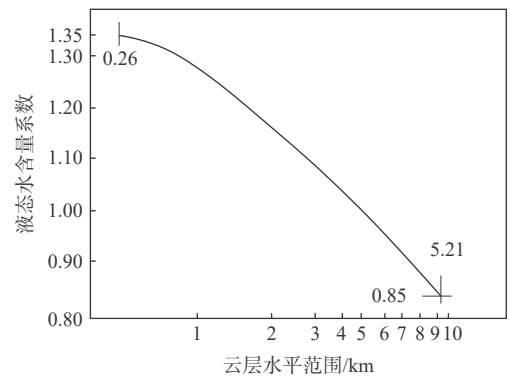


图7 间断最大(积云)大气结冰状态LWC系数与云层水平范围之间的关系[3]

Fig. 7 Intermittent maximum (cumuliform clouds) atmospheric icing conditions (variation of liquid water content factor with cloud horizontal extent)[3]

2 FAR 25附录C图表编制方法

在FAR 25附录C形成时期,已经有了大量的飞行观测数据,这些数据来源不一、测量方式不同、数据包含的内容不同、测量地点也各异,因此,需要可靠的方法将这些数据编制成具有统一格式、方便查询的图表。FAR 25附录C的编制方法是:把观测获得的气象数据根据结冰环境进行分类,在此基础上,筛选出对结冰有最直接影响的结冰参数,探寻这些结冰参数之间的影响规律,从而获得结冰气象条件的包线,建立起具有较高可靠性的经验性图表。当原始气象数据库中个别条目中的记录不能覆盖到图表需要的参数范围或部分参数缺失时,就根据经验估计或理论模型推测出这些参数,最终形成标准的“LWC-MED-气压高度-环境温

度-云层水平范围”的结冰条件单元。

飞行结冰的气象条件主要被分为 4 种, 如表 1 所示[2], 其中, V为平均空速。对于结冰气象条件的描述, 除以上信息外, 还有一个重要的参数是该气象状况出现的可能性。因此, FAR 25 附录 C 中根据美国国内的飞行记录, 将以上 4 种气象进一步分成 2 类, 分别用 M 和 N 指代, M 指所有可能的天气状况中, 气象参数出现的最大值, N 则指正常或平均的天气状况。图 2(Class III-M)针对的是第 3 种类型中气象参数最大值, 这种情况往往发生在飞机穿越层状云时[9-10]。FAR 25 附录 C 形成时期的气象条件参数主要靠机载测量仪器获得, 主要包括: 自由空气温度、LWC、MED、液滴粒径分布和冰颗粒浓度。此外, 空速由空速指示器测量并由记录器进行记录; 飞行的海拔由高度计测量和高度记录仪进行记录[11]。图 2和图 5 的数据来自于 NACA[1],而其原始数据则大都来自于 1945—1946年冬天美国 Ames 实验室利用 C-46 飞机携带机载设备开展的自然结冰条件下的飞行试验[12]。此次试验一共采集了 232 次可靠的结冰探测数据, 这些数据包括测量的地点、云层的类型、云层的组分、日期和时间、LWC、MED、液滴直径的最大值和最小值、气温和气压高度等; 这些数据根据云层的类型(层云和积云)和其中是否含有冰晶, 绘制成结果如图 8 所示[9], 其中, V_i为结冰速度, 图中数据采用在高度为 3.048 km, 温度为−9.4 ℃, 空速为 89 m/s 情况下, 直径为 0.076 2 m 的圆柱积冰速度为不同结冰强度的划分依据。

表 1　4 种结冰气象类型和其特性[2]

Table 1　Classes of flight icing conditions and their features[2]

气象类型	归类	水平飞行距离/km	持续时间/min (V=80 m/s)	含水量	适用飞机部位
类型I	瞬时	0.8	0.17	非常高	发动机进气道
类型II	连续	4.8	1	高	发动机进气道、风挡及其他需要保持清晰视野的部件
类型III	连续	连续飞行	较长	适中	飞机机翼和尾翼
类型IV	冰雨	160.9	30	雨滴较大, 温度接近冰点, 但是空气总体含水量低	机身上用于测量飞行速度的皮托管等

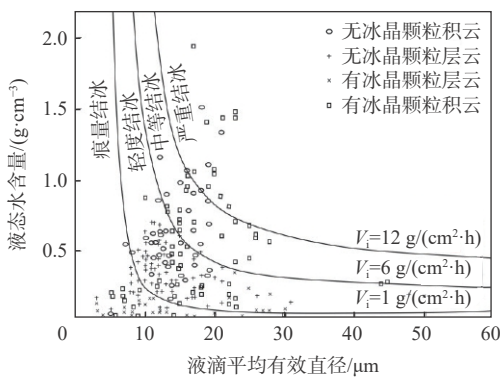


图 8　液滴 MED 和 LWC 飞行试验测量数据分布[9]

Fig. 8　Distribution of flight measurement data for MED and LWC [9]

本文以图 2 中液滴平均直径为 15 μm, 含水量为 0.8 g/m^3 的点为例介绍其计算过程: 从图 8 中可以看出, 层状云中最大的 LWC 为 0.7 g/m^3, 通过对飞行时间和 LWC 的关系推测, 0.8 g/m^3 为比较合理的 LWC, 再根据飞行数据中 LWC 和 MED(见图 8)及其与温度的关系, 从而得出其直径为 15 μm, 温度为 −6.7 °C。

以图 5 中气温为 0 °C, LWC 为 2.5 g/m^3, 平均直径为 20 μm 的点为例, 获取该点数据的步骤如下。

1) 气温的确定

气温取决于云层出现的地点和季节及云层所处的海拔高度。以美国北部冬季为典型, 考虑 1.83～2.44 km 高度的云层, 可以得知云层中的平均温度为 0 °C。

2) LWC 的确定

LWC 的大小w可以通过大量的飞行试验测试获得, 在测量数据量比较丰富的情况下, 不需要非常复杂的数据处理或条件假设即可获得比较准确的结冰条件参数之间的关系。但是在数据不完整或缺失的情况下, 还需要理论计算来补充。其计算式为[9]

$$w = (X_t - X_s)\rho \tag{1}$$

式中: X_t为当地地面的饱和混合比; X_s为云层中的饱和混合比; ρ为干空气的密度。该表达式表明地面空气上升至云层高度时, 空气中的水蒸气冷凝成为液态水的质量。根据该表达式可以计算出空气中的含水量为 2.5 g/cm^3。

3) MED 的确定

MED 的计算依据如图 8 所示, 在图中可以看到, 对应积云的数据空气含水量超过 1.2 g/cm^3 的点一共有 7 个, 这 7 个点的直径范围为 17～23 μm, 取其平均值 20 μm 作为该状态下的液滴平均直径。图 9[13] 为图 3 和图 6 绘制所需要的原始数据来源, 也是来自于飞行试验中发生结冰时对应的气压高度和温度数据, 根据其发生频率制定的包线, 图中数字表示结冰次数。图 10 为图 4 和图 7 绘制所需要的原始数据来源, 该图由 1 038 次飞行测量数据

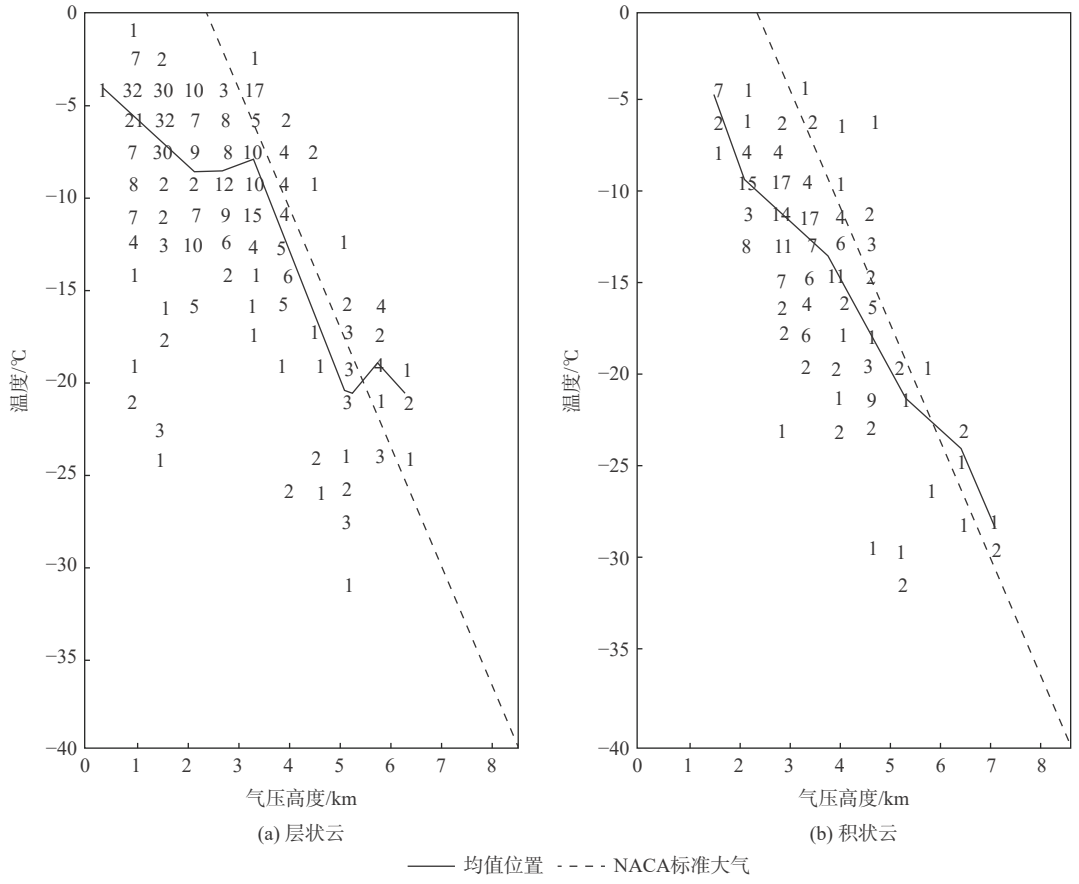


图9 飞行实测的气压高度和气温的关系[13]

Fig. 9 Rolationship between pressure altitude and temperature from flight measurements[13]

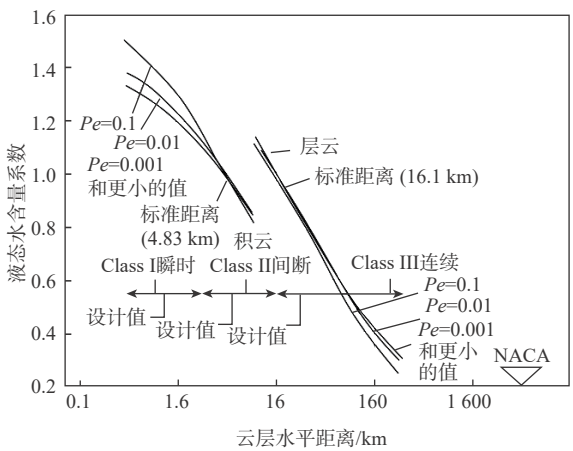


图10 水平距离和LWC系数的关系[14]

Fig. 10 Relationship between horizontal distance and LWC coefficient[14]

总结而成[14]，其中，*Pe*为Peclet数，表示液态水的空间分布与输送特性，其影响LWC因子的变化趋势，越小扩散主导更强，液态水分布更平缓。

3 FAR 25附录C形成时期气象条件参数测量

3.1 待测参数

1）自由空气温度。自由空气温度的测量主要采用铜-康铜热电偶，用毫伏表测量热电偶的电压差。热电偶的偶丝连接处使用挡板防护以防止结冰，同时采用铠装以防止辐射误差。飞机上的热电偶在高速气流中会受到气动加热的作用，因此，测量到的温度还需要通过以下表达式进行校正[9]：

$$T = T_{\mathrm{M}} - 269.52\alpha U^2 \tag{2}$$

式中：单位采用华氏温度；T为自由气流的真实温度；T_{M}为热电偶直接测量所得的温度；U为当地速度，单位为cm/s。α常数与热电偶的安装方式有关。在理想状态下，对于绝热压缩的气流滞止点，$\alpha = 1.0$。通常情况下α的值可以取0.93 ± 0.05。这个数值需要在固定的空域和飞行高度及不同的飞行速度条件下多次测量获得。

2）LWC。LWC的测量采用旋转多圆柱测量仪，实物如图11所示[15]。该仪器可以对液滴尺寸、液滴尺寸分布及LWC进行测量。其原理是不同尺寸的圆柱体，其单位正面（迎风）面积上收集液滴的数量不同。液滴在高速气流中绕圆柱运动的轨迹计算可以参考文献[10]。圆柱体上单位正面面积所收集的冰量用收集效率表示，即收集到的冰的质量与圆柱体扫过的空间中包含的液态水的质量之比。理论表明，收集效率是液滴大小、圆柱尺寸、

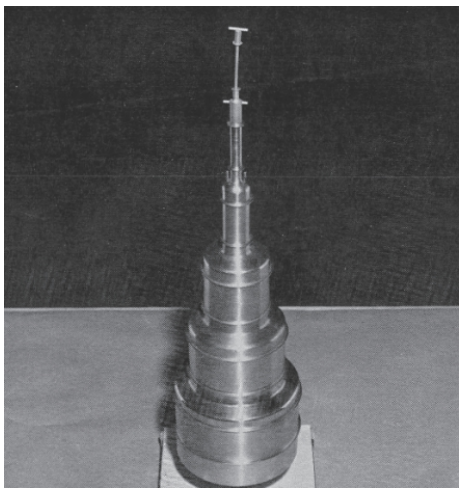


图 11　旋转多圆柱测量仪[15]

Fig. 11　Rotating multicylinders gauge[15]

空速、空气密度和空气黏度的函数[16]。通过将每个圆柱上收集的冰的质量与收集效率的理论计算值进行比较, 即可得出 LWC 和 MED。实际上, 该测量方法并不能准确的得到液滴的尺寸分布规律, 但可以大致估测一下粒径分布的不均匀程度。具体方式是, 先假设一个合理的液滴尺寸分布规律, 然后, 代入到理论计算中得到收集效率, 再与测量得到的飞行数据进行比较, 以此进行迭代和改进, 得到最终近似的尺寸分布规律[16]。

在测量时多圆柱测量仪从机身伸出, 如图 12 所示[16], 测量仪暴露在来流中一定的时间, 然后, 收回、拆卸和称重。测量中默认所有撞击在圆柱上的过冷液滴都会完全冻结在圆柱上。圆柱在测量期间会保持旋转, 以保证液滴围绕着圆柱表面均匀结冰, 从而保持仪器的横截面为圆形[17]。

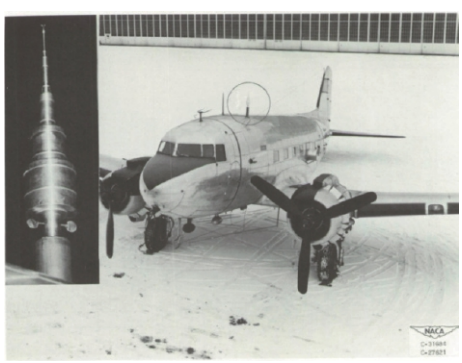


图 12　旋转多圆柱测量仪在飞机上的安装(在机身顶部伸出)[16]

Fig. 12　Installation of a rotating multicylinder measuring instrument on aircraft (protruding from top of fuselage)[16]

3.2　测量步骤

多圆柱测量仪的使用步骤如下[17]:

1) 在已知的气压、飞行高度、风速和气温条件下, 将一组不同尺寸的圆柱体暴露在含有过冷液滴的云层测量一段时间(通常为 1～5 min);

2) 将多圆柱测量仪拆卸, 把每个圆柱(和其表面携带的冰层)存放在单独的防水容器中;

3) 称量单个圆柱在结冰前后 2 种条件下的质量, 获得收集的冰的质量;

4) 将每个圆柱上面结冰的质量, 与在该飞行条件下获得的理论计算值进行对比, 获得云层 LWC、液滴尺寸和液滴尺寸分布。具体注意事项见文献 [17]。

3.3　数据处理

在进行数据处理之前, 还需要测量飞行参数(包括当地空速 U、压力高度和气流温度 T_{M})。数据处理需要的主要参数有: 理论计算获得的收集效率、多圆柱测量仪获得的结冰数据和 φ/LU 参数。其中, φ 为一个无量纲参数, 该参数的数值通过查询图表获得, L 为圆柱平均直径。具体的数据处理方法遵循以下步骤, 如图 13 所示。

1) 将机载仪器直接测量得到的空速经过校正, 得到真实平均空速 V, 单位 cm/s。该过程主要修正机身外形对于测量的影响, 实验[18] 表明, 未经校正的测量数据比真实值高出约 12%。

2) 将机载仪器直接测量到的气流温度 T_{M} 进行校正, 得到真实的气流温度 T。和式 (2) 不一样的是, 文献 [17] 采用了另一个表达式

$$T = T_{\mathrm{M}} - rV^2\frac{a_{\mathrm{s}}}{a_{\mathrm{d}}} \times 301.4 \times 10^{-4} \tag{3}$$

式中: 温度恢复因子 r 为一个无量纲的常数, 该常数和温度探针的形状、热传导率及安装位置有关, 通常可以取 0.85; 气温直减率 $\frac{a_{\mathrm{s}}}{a_{\mathrm{d}}}$ 为压力高度和温度 T_{M} 的函数, 可以通过查询图表获得(注: 图表来自于文献 [17], 但文献中未说明图表数据来源, 笔者在文献 [17] 中寻找到其中一个图表的原始数据, 据此推测此数据处理方法中使用的所有图表应该来自于对实验数据的拟合)。

3) $\frac{\varphi}{LU}$ 参数计算。$\frac{\varphi}{LU}$ 主要表征海拔和温度对于液滴尺寸和尺寸分布计算的影响。

4) 计算平均圆柱直径 L。本文平均圆柱直径是指表面结冰以后的圆柱直径, 其计算式为[17]

$$L = \frac{1}{2}\left(L_0 + \sqrt{\frac{m}{\rho_{\mathrm{i}}\pi l} + L_0^2}\right) \tag{4}$$

式中: ρ_{i} 为冰的密度, 一般可以取 0.85 g/cm^3; l 为圆柱的长度; m 为圆柱表面结冰的质量, g; L_0 为圆柱未结冰时的原始直径。

5) 计算单位迎风面积的结冰质量[17]:

$$I = 1.118 \times 10^4 \times \frac{m}{Ll} \tag{5}$$

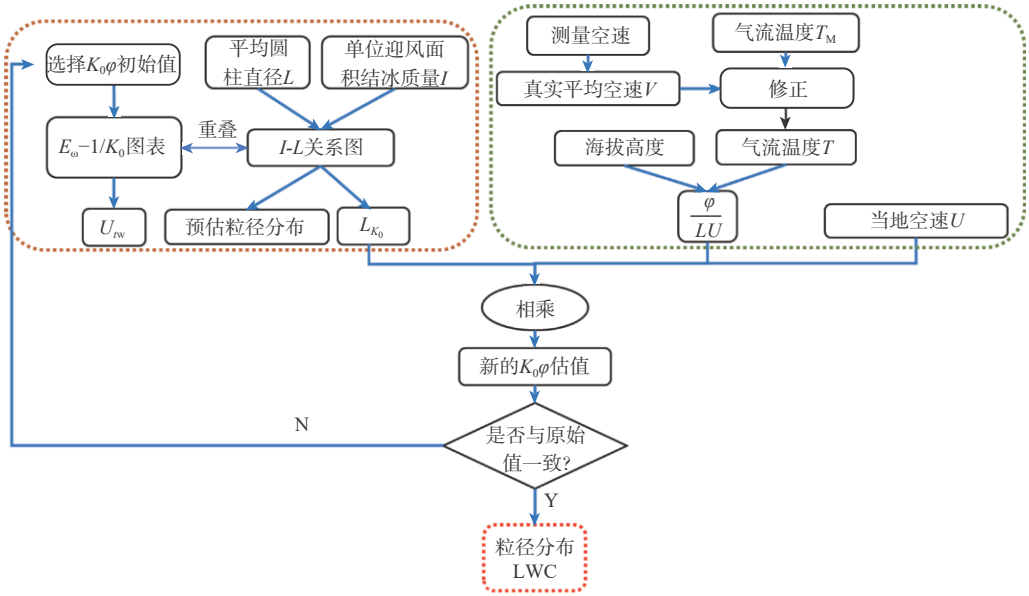


图 13 结冰参数数据处理步骤

Fig. 13 Procedure for post-processing of icing parameter

6) 基于以上计算结果，可以绘制I-L的关系图，采用对数坐标轴，如图 14 所示[17]。

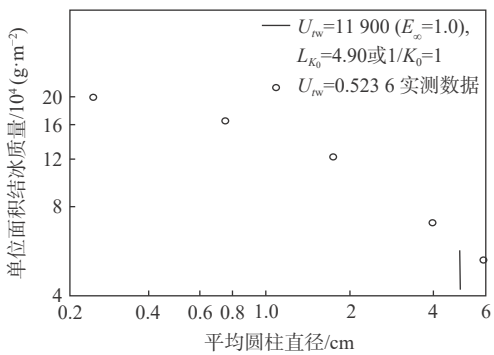


图 14 单位正面面积上的结冰质量随着平均圆柱半径的变化关系[17]

Fig. 14 Relationship between icing mass per unit frontal area and mean cylindrical radius[17]

7) 匹配计算。在这一步需要将第 5 步中实验测量所得到的I与在该飞行海拔高度和大气条件下用理论计算出来的结果相对比，从而得到 LWC、液滴尺寸和粒径分布。理论计算获得的是 cE_ω−1/K_0 图表，如图 15 所示[17]，即将E_ω表示为1/K_0的函数，图 15 中 c 为常数，不同常数下cE_ω−1/K_0关系不同。E_ω为无量纲的收集效率，K_0为表征惯性的参数。A、B、C、D、E为不同常数下cE_ω−1/K_0之间的关系。无量纲参数φ的定义为[17]

$$\varphi=\frac{18\rho_a LV'}{\mu\rho_w} \tag{6}$$

式中：ρ_a为空气的密度，g/cm³；V'为空速，cm/s；μ为空气黏度，g/(cm·s)；ρ_w为液滴的密度，g/cm³；在$K_0\varphi$取不同值的时候，可以通过cE_ω−1/K_0图表得到

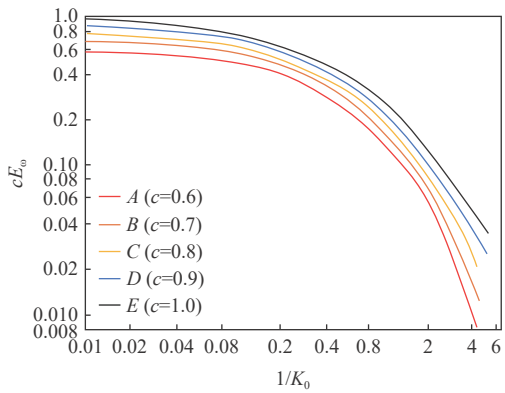


图 15 cE_ω-1/K_0图表[17]

Fig. 15 cE_ω-1/K_0 plot[17]

不同粒径尺寸所对应的E_ω值。匹配的具体步骤如下：

① 选择合适的$K_0\varphi$的初始值；除非有经验表明别的数值能够取得更好的效果，否则一般选择 3 000；

② 根据选取的$K_0\varphi$的值，将对应的$cE_\omega-1/K_0$图表与第 5 步中实验数据生成的L-I图叠加在一起（在 FAR 25 附录 C 制定的年代，需要放在透光的桌面上，让光线穿过 2 张图纸以方便透视），保持 2 张图的横纵坐标平行，然后移动实验的图，直到实验图中的数据点和$cE_\omega-1/K_0$图中的曲线达到最大重叠（即尽可能保证实验的点都落在$cE_\omega-1/K_0$图中某一条曲线上），该曲线所对应的粒径分布即为暂定的结果。在重叠的I-L上的数据点中，选取1/K_0 =1 所对应的L值，记为LK_0；选取E_ω=1 所对应的I值，记为U_{tw}（t为圆柱暴露在空气中的时间）。

③ 将步骤 3 中获得的$\frac{\varphi}{LU}$与L_{K_0}、U这 2 个参数

相乘, 得到新的$K_0\varphi$估测值, 再利用表 2 中对应关系, 得到最终的标准$K_0\varphi$值。

表 2　$K_0\varphi$估测值和标准值对应关系[17]

Table 2　Correspondence between estimated and standard values of $K_0\varphi$[17]

估测或计算的$K_0\varphi$范围	对应的标准$K_0\varphi$值
200～440	200
441～1 750	1 000
1 751～5 600	3 000
5 601～30 000	10 000

如果得到的标准的$K_0\varphi$值和第 1 步中选择的$K_0\varphi$初始值一致, 那么L_{K_0}和U_{rw}就在第 8 步使用。如果不相同, 那么以上匹配过程就要进行迭代, 直到$K_0\varphi$初始选取的值和标准值相同为止。

8) 计算液滴直径和 LWC。在第 7 步中, 已经获得了U_{rw}、L_{K_0}、粒径尺寸分布和$K_0\varphi$。所有的这些参数的获取都是基于$K_0\varphi$的标准值和初始值。通过L_{K_0}、U和T的值, 查询相应的图表即可获得未经校正的液滴直径。MED 则是利用$K_0\varphi$的估测值、$K_0\varphi$的标准值在对应的图表中获得校正因子, 与未经校正的液滴直径相乘得到。LWC 则利用U_{rw}除以U和t获得。

4　现代电子/光学探测器和数据处理

现代测量设备为结冰飞行试验中的参数测量提供了更加精确的手段, 基于光学、声学和电磁波的各种新型测量设备得到了有效的应用, 针对这些新型的仪器, 人们发展出了各种新的数据处理方法。现代的测量手段也为防冰带来了新的思路, 即从过去的被动防冰模式转化为主动防冰模式: 过去飞机上的结冰探测都是靠传感器感受到结冰后提供告警信号, 是一种被动式的结冰探测。而现代的探测系统(如波音 787)利用处理器和先进的算法, 能够综合空气湿度、总温、静压和飞机迎角等参数进行计算, 对比内置的大气数据模型, 从而具备提前探测空气中潜在结冰条件的能力, 提供主动的预测式结冰探测。

4.1　机载光学测量设备

过去结冰试验研究中使用了多种测量云层中液滴直径的设备, 根据测量方式主要可以分为接触式测量和非接触式测量。其中, 光学测量是最常用的非接触式测量设备。过去美国国家航空航天局(National Aeronautics and Space Administration, NASA)开展的实验测量中有 2 种仪器最为常用: 前向散射光谱仪(forward scattering spectrometer probe, FSSP)和光学阵列仪(optical array probe, OAP), 如图 16 所示[15], 上方为 FSSP, 下方为 OAP。FSSP 用于测量较小范围尺寸的液滴, 测量范围为 2～47 μm; OAP 可以用来测量 15～450 μm范围内的液滴[11]。

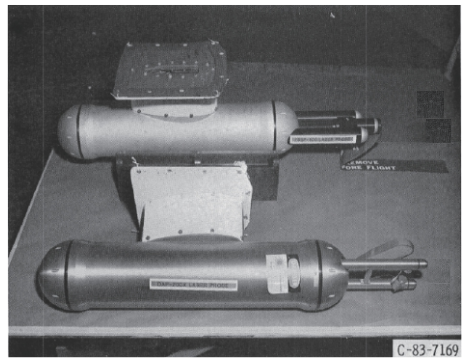


图 16　液滴尺寸测量仪器(FSSP 和 OAP)[15]

Fig. 16　Droplet size measuring instruments (FSSP and OAP)[15]

FSSP 使用 Mie 散射模型计算通过粒子的大小和直径。待测量的液滴通过 FSSP 的聚焦激光束时, 探测器上的入射功率和液滴的尺寸大小呈线性关系。测量时假定一次只有一个液滴处于探头光学元件的景深内。仪器将测量的液滴的尺寸分类计入 15 个匣(bin), 如 FSSP 总的液滴直径测量范围是 2～47 μm的话, 相邻的 2 个匣尺寸就相差 3 μm。将一段时间内的记录汇总在一起便可以生成液滴尺寸分布。通常连续采集 50 s, 就可以获得粒子数量对应粒子大小的柱状图, 仪器的数据分析程序根据这个原始记录就可以计算出被测量粒子的分布特性。FSSP 容易受到云层环境的影响, 测量时会发生仪器冻结、数据丢包及将较小的冰晶计为液态液滴等问题。FSSP 在使用前需要进行标定, 最常用的方法就是用已知直径的小玻璃球作为待测粒子进行标定。

OAP 通过测量粒子形成的投影的直径大小来确定液滴本身的直径。当液滴通过激光束时, 其阴影通过光学透镜扩大并投射在光电二极管阵列上。液滴直径由被遮挡住的二极管的数量来确定, 同时阴影的大小也和系统的放大倍数有关。文献 [15] 中使用的 OAP 有 30 个匣, 能够测量的液滴直径范围为 15～450 μm, 每个匣的梯度为 15 μm。与 FSSP 一样, 需要假设每次只有一个液滴出现在探头的景深内。为消除测量过程中信号波动造成的影响, 保证统计结果准确, 特别是对液滴较大的情况, 在云层中的测量时间需要持续 100 s。

这 2 种仪器除了可以测量液滴直径的大小和分布规律之外, 也可以用来计算 LWC, 如下[11]:

$$w=\frac{\rho_w\pi}{6}\sum\frac{N_i d_i^3}{S_{Ai}Vt} \tag{7}$$

式中：ρ_w为水的密度；N_i为第i个匣子中统计出的液滴数量；d_i为第i个匣子中的中径大小；S_{Ai}为第i个匣子的横截面积；t为测量的采样时间。对于FSSP来说，每个匣子的S_{Ai}的大小相同，而对于OAP来说，不同匣子的S_{Ai}大小不同。

此外，常用的还有相位多普勒粒子分析仪（phase Doppler particle analyzer, PDPA），如图17所示[19]。相位多普勒粒子分析仪可以同时测量雾化液滴的尺寸、速度及LWC。该系统基于光散射干涉原理，通过小型非接触式光学探头进行测量。PDPA测量速度的原理与激光多普勒测速仪（laser Doppler velocimetry, LDV）一致，使用激光产生相干（固定频率）光束以照亮流动中的运动粒子。这种粒子运动产生的多普勒频移与其运动速度成正比。而其测量粒子尺寸的原理是：用两束激光交叉确定系统测量空间，当粒子通过测量空间时，其将光束中的光散射到位于偏离轴收集角的一个多检测器接收探头中。来自不同检测器的多普勒脉冲信号之间的相移与球形粒子的粒径成正比[20]。PDPA仪器计算LWC的方式与FSSP[21]等一致。测量方式为单点测量，通过多次采集获得整体流场的数据。该仪器的主要优点是易于操作，不需要复杂的校正过程；对测量环境背景光线引起的误差不敏感。

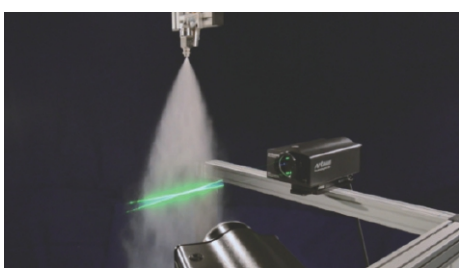


图17 相位多普勒粒子分析仪[19]

Fig. 17 Phase Doppler particle analyzer[19]

此外近些年来还有云粒子探头（cloud droplet probe, CDP）及快速云粒子探头（fast cloud droplet probe, FCDP）等，如图18所示[22-23]，该探头原理类似于FSSP，也是一种前向散射探头，测量范围为[24] 2～50 μm。

后向散射云探测器（back scattered cloud detectors, BCP）等测量仪器，其基本原理是穿过被采样空间的粒子，其散射光的强度和粒子自身的尺寸、折射率、粒子形状及相对于入射光束的收集角度呈正比[21]。因此，通过测量散射光的强度即可获得粒子的大小。该仪器相比FSSP和OAP具有尺寸小质量轻的优势，如图19所示[21]，已经应用在一些小型客机和A330/340等宽体客机上进行测量，如图20所示[21]。



(a) 云粒子探头[22]



(b) 快速云粒子探头[23]

图18 云粒子探头和快速云粒子探头

Fig. 18 Cloud particle probe and fast cloud particle probe

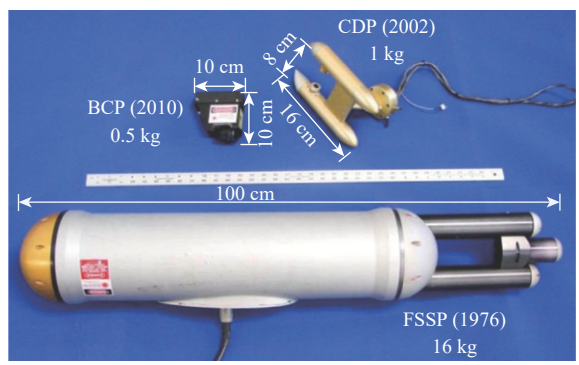


图19 BCP和CDP及FSSP的对比[21]

Fig. 19 Comparison of BCP, CDP and FSSP[21]

4.2 雷达测量

FSSP、OAP、PDPA、CDP等测量方法，不管是接触式还是非接触式，都需要飞机穿过含水量较高的云层进行测量，因此，会增大飞机结冰的风险，造成安全隐患。后来人们发展了利用地面和机载雷达进行云层LWC及液滴尺寸分布的测量方法。

4.2.1 LWC测量

早期雷达探测使用单波段雷达[25]，这种探测方式只能获得空气中的LWC，难以获得别的参数，而且误差较大。后来发展了基于双波段雷达的测量方式[26]。该测量方式需要同时准确测量云层中的气温，适用于云层中的液滴尺寸不超过雷达波长的情况。双波段雷达同时使用Ka（波长为0.86 cm）和X（波长为3 cm）2个波段。其中，Ka波段的强度在穿过云层中的液态水以后会被减弱，而X则不会受


(a) 探头在入口舱口


(b) 探头在右下方无线电舱内


(c) 探头在空中客车上


(d) 空中客车窗口特写

图 20　安装在飞机上的 BCP 探头[21]

Fig. 20　BCP mounted on airplane[21]

到影响。Ka波段减弱的原因主要来自于信号的吸收和散射。对于比较小的液滴来说, 散射造成的信号损失可以忽略不计, 因此, 信号的减弱主要由吸收引起。吸收损失和空气中的 LWC 成正比, 在 −10 ℃ 的情况下, 可以表示为[27]

$$w = 0.74A \tag{8}$$

式中: w为 LWC, g/cm^3; A 为 Ka 波段雷达信号穿透云层所经历的单位距离的衰减, dB/km。其中, A云层中液滴对单程Ka波段雷达信号造成的减弱随着气温的降低, 相应的减弱也会降低, 因此, 式(8)还需要根据气温进行调整。

4.2.2　雷达估计尺寸

由于瑞利散射, 雷达信号的反射率因子Z(或者简称为反射率)和液滴尺寸分布的六阶矩成正比, 信号的减弱和其三阶矩成正比。这规律在$D < \lambda/16$时成立, D为粒子直径, λ为雷达波长。对于以上 2 种波长, 瑞利散射定律对于小于 1 mm 直径的液滴成立[27]。反射率因子 Z 和 A 表达式分别为

$$Z = \int D^6 N(D)\mathrm{d}D = D^6 \tag{9}$$

$$A = 4.343\times10^{-3}\frac{\pi}{6}k\varepsilon_{\mathrm{r}}''\left|\frac{3}{\varepsilon_{\mathrm{r}}+2}\right|^2 \cdot \int D^3 N(D)\mathrm{d}D = c_1 D^3 \tag{10}$$

式中: 常数$c_1 = 4.343\times10^{-3}$; $\pi/6k\varepsilon_{\mathrm{r}}''|3/(\varepsilon_{\mathrm{r}}+2)|^2$的值为 7.12×10^{-4} dB/km; ε_{r}和$\varepsilon_{\mathrm{r}}''$分别为水相对电容率的实数部分和虚数部分。$Z/A$可以定义为一个新的变量, 叫雷达估计尺寸(radar-estimated size, RES), 其表达式为 [27]

$$D_{\mathrm{RES}} = \left(\frac{D^6}{D^3}\right)^{\frac{1}{3}} \tag{11}$$

结合方程 (9)～ 方程 (11), RES可以用雷达测量到的参数定义为[27]

$$D_{\mathrm{RES}} = \left(7.12\times10^{-4}Z/A\right)^{\frac{1}{3}} \tag{12}$$

式中: D_{RES} 的单位为mm; Z的单位为mm^6/mm^3。一般来说, RES 的值比平均体积直径(mean volume diameter, MVD)要偏大, 但是实际上, RES 在实际应用中更有优势: 一方面来说, RES 的计算比较容易, 并不需要事先获取粒径尺寸的分布规律; 而且真正容易造成飞机结冰的正是空气中比较大的液滴, 因此, RES 作为结冰严重程度的衡量标准更有价值; 另一方面, 当液滴粒径的分布比较窄的时候, 实验测量结果表明 RES 和 MVD是一致的。

4.2.3　液滴尺寸分布

对于液滴尺寸分布$N(D)$的测量, 其指导思想是: 首先需要人为定义一种分布形式, 如指数分布或 gamma 分布, gamma 分布的公式可以写为[27]

$$N(D) = aD^{\sigma}\exp(-bD^{\gamma}) \tag{13}$$

式中: a、σ、b和γ这 4 个参数用来确定液滴尺寸分布规律。这种方程形式可以通过调整 4 个参数值的大小, 灵活获取不同的分布曲线。当然也可以采用指数形式的分布, 如下[27]:

$$N(D) = aD^{\sigma}\exp(-bD^{\gamma}) \tag{14}$$

基于以上分布, 可以推导出相应的 MVD 和 RES 的值[27]:

$$D_{\mathrm{MVD}} = \left(\frac{3.67+\sigma}{\gamma b}\right)^{1/\gamma} \tag{15}$$

$$D_{\mathrm{RES}} = b^{-1/\gamma}\left\{\frac{\Gamma\left[(\alpha+7)/\gamma\right]}{\Gamma\left[(\alpha+4)/\gamma\right]}\right\}^{1/3} \tag{16}$$

式中: gamma 积分$\Gamma\left[(\alpha+7)/\gamma\right] = \int D^7 D^{\alpha}\exp(-bD^{\gamma})\mathrm{d}D$。采用不同的$a$、$\sigma$、$b$和$\gamma$参数组合, 会获得不同的 MVD、RES 和 LWC 的值, 用这些数值去和真实飞行试验中所测量到的值相对比, 从而寻找到能够和实验数据一致的a、σ、b和γ等参数。

值得一提的是, 这种测量方式的精度会受到云层中冰晶含量的影响, 这是因为Ka波段的衰减需要X波段的反射率信息进行辅助。冰晶并不会影响Ka波段雷达信号的衰减, 但是会对雷达信号产生散射, 从而对其反射率产生影响。而且近些年来[25], 高空中的冰晶造成的航空发动机内部结冰问题逐渐引起重视, 而冰晶的尺寸非常微小, 探测难度大, 因此, 现在出现了三波段的雷达探测技术, 在双波

段雷达的基础上，增加了波长更短的 W 波段，测量原理和双波段雷达基本相同[26]。

现代自然结冰环境的测量往往需要多种测量方式结合，图 21[28] 为 2018 年 8 月开展的 NASA/FAA HIWC RADAR II 飞行测试试验。一架 NASA AFRC DC-8 飞机同时搭载了多种测量仪器，包括用于测量冰晶水含量的第 2 代等速探头（iso-kinetic probe II, IKP2），用于测量云层液滴粒径分布的 3 个粒子探针：云滴探针 II（cloud drop probe, CDP-2）、二维立体（two-dimensional stereoscopic, 2D-S）成像仪和降水粒子成像仪（precipitation imaging probe, PIP）。IKP2 用于测量体积总含水量（total water content,TWC）。CDP-2、2D-PIP 都是现代的大气科学界常用的机载云探测器。其他的机载配套探头还有空气总温（total air temperature, TAT）探头、3 个总含水量探头，用以研究不同位置所测量得到的总含水量的差异，背景湿度系统（background humidity systems, BHS）和皮托管探头和测量气流方向的探头，其安装位置如图 21 所示，安装在 NASA AFRC DC-8 飞行实验室的 HIWC RADAR II 仪器上[28]，具体的介绍可以参考文献 [28]，其中，MMS 表示大气基本参数测量系统。

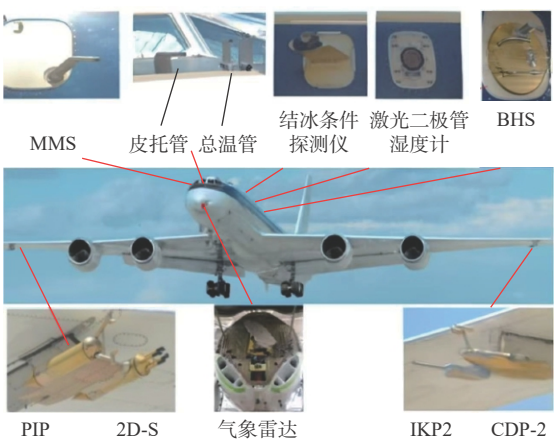


图 21 安装在 NASA AFRC DC-8 飞行实验室的 HIWC RADAR II 仪器[28]

Fig. 21 HIWC RADAR II instruments at NASA AFRC DC-8 flight laboratory[28]

5 挑战及发展趋势

目前 FAR 25 附录 C 的使用还有一定的地区和季节局限性，其定制更多是参考了早期美国北部春冬季节的飞行结冰情况；近年来，随着全球气候变化，以及更复杂的结冰环境（如过冷大水滴结冰、冰晶结冰及冰晶/过冷水滴混合相结冰等）被发现，都会导致 FAR 25 附录 C 的使用受到制约。因此，有必要在国内开展结冰气象的探测工作。其工作重点一方面在于结冰数据库的建立，这需要大量的飞行探测和先进的机载探测仪器作为支撑；另一方面在于如何利用当前飞速发展的数据科学，结合多种测量方式，提高测量精度。不管是传统的还是现代的光学测量仪器，其测量结果之间往往存在差距，需要进行相互印证。雷达结冰探测的难度在于其结果的同步验证，即利用雷达获取到待测云层中的结冰数据的同时，需要有机载仪器在该云层中进行同步测量，从而为雷达数据提供验证和标定[26]。欧洲 SENS4ICE EU 计划也已经开展了基于第二代气象卫星图像的结冰气象识别工作[29]，空间和时间的分辨率分别为 3 km 和 15 min。同样的，该方法也面临着识别结果的验证难题。

此外，结冰天气持续时间短，可能发生结冰的云层难以捕捉，给测量带来了难度；由于时间分辨率较低，也给基于卫星图像的结冰气象识别方法带来了挑战。过冷大水滴的结冰气象对于探测飞机本身也有严重威胁，作业危险性较高。利用无人机探测结冰气象有一定的应用前景（见 https://www.sohu.com/a/670262240_121123814），然而，当前无人机受制于机载能源和尺寸的限制，尚无可靠的防除冰系统，因此，其相比有人机对于结冰的防护更差。

6 结 论

1）本文介绍了用于指导飞机防除冰系统设计的 FAR 25 附录 C 的形成背景及演化过程，对影响飞行安全的结冰气象条件的分类和区别做出了详细的解释。

2）本文重点回顾了 FAR 25 附录 C 早期用于测量结冰气象条件的探测仪器和测量原理，同时以具体的例子阐释了数据处理步骤和绘制过程，从而对编制中国结冰气象环境的标准化研究工作提供指导。

3）在现代电子/光学云粒子探测仪器广泛应用的背景下，本文总结了这些测量仪器的原理与优缺点。雷达数据探测方法与其他方法相比，飞行器可以不必飞入结冰与层中，从而提升了探测的安全性。

4）本文详细介绍了雷达探测的原理和相应的数据的处理方法，基于目前的研究进展，讨论了结冰气象探测目前的发展方向和工作难点。

参考文献（References）

[1] LONG T. Analysis of weather-related accident and incident data associated with section 14 CFR part 91 operations[J]. Collegiate Aviation Review International, 2022, 40(1): 25-39.

[2] ALUN R J, LEWIS W. Recommended values of meteorological factors to be considered in the design of aircraft ice-prevention equipment: NACA-TN-1855[R]. Washington, D. C.: National Aero-

nautics and Space Administration, 1949.

[3] REGULATIONS F A. Part 25-Airworthiness standards: transport category airplanes Appendix C: 4080-29-FR-18291[R]. Washington, D. C.: Federal Aviation Administration, 1965.

[4] JECK R K. Icing design envelopes (14 CFR Parts 25 and 29, Appendix C) converted to a distance-based format: DOT/FAA/AR-00/30 [R]. Washington, D. C.: Federal Aviation Administration, 2002.

[5] JECK R. Converting appendix C to other variables: DOT/FAA/AR-00/30 [R]. Warrendal: SAE International, 2003.

[6] JECK R K. A history and interpretation of aircraft icing intensity definitions and FAA rules for operating in icing conditions: DOT/FAA/AR-01/91 [R]. Washington, D. C.: Federal Aviation Administration, 2001.

[7] Federal Aviation Administration. AC 25.1419-1A - certification of transport category airplanes for flight in icing conditions: 25.1419-1A [R]. Washington, D. C.: Federal Aviation Administration, 2002.

[8] 季发明, 黄子彰, 浅析飞机适航规章 FAR25 第 121 修正案[J], 黑龙江省政法管理干部学院学报, 2015, 26: 66.
JI F M, HUANG Z Z. A brief analysis of the aircraft airworthiness regulation FAR25 amendment to 121[J]. Journal of Heilongjiang Provincial Political and Legal Management Cadre College, 2015, 26: 66(in Chinese).

[9] LEWIS W. A flight investigation of the meteorological conditions conducive to the formation of ice on airplanes: NACA-TN-1393[R]. Washington, D. C.: National Advisory Committee for Aeronautics, 1947.

[10] LANGMUIR I, BLODGETT K. Mathematical investigation of water droplet trajectories: 5418[R]. Champaign: University of Illinlis at Urbana-Champaign, 1946.

[11] IDE R. Comparison of liquid water content measurement technquies in an icing wind tunnel: DTIC ADA371598[R]. Washington, D. C.: National Aeronautics and Space Administration, 1999.

[12] HILL E. Overview of federal aviation administration aviation safety research for aircraft icing[C]//Proceedings of the 44th AIAA Aerospace Sciences Meeting and Exhibit. Reston: AIAA, 2006.

[13] HACKER, P T, ROBERT G D. A summary of meteorological conditions associated with aircraft icing and a proposed method of selecting design criterions for ice-protection equipment: NACA-TN-2569[R]. Washington, D. C.: National Advisory Committee for Aeronautics, 1951.

[14] LEWIS W, NORMAN R B. A probability analysis of the meteorological factors conducive to aircraft icing in the United States: NACA-TN-2738[R]. Washington, D. C.: National Advisory Committee for Aeronautics, 1952.

[15] IDE R F, PAUL RICHTER G. Comparison of icing cloud instruments for 1982-1983 icing season flight pro gram: A82-16098[R]. Washington, D. C.: National Aeronautics and Space Administration, 1984.

[16] BRUN R J, LEWIS W, PERKINS P J, et al. Impingement of cloud droplets on a cylinder and procedure for measuring liquid-water content and droplet sizes in supercooled clouds by rotating multicylinder method: NACA-TR-1215 [R]. Washington, D. C.: National Advisory Committee for Aeronautics, 1955.

[17] LEWIS W, PORTER J P, RINALDO J B. Procedure for measuring liquid-water content and droplet sizes in supercooled clouds by rotating multicylinder method: NACA-RM-E53D23[R]. Washington, D. C.: National Advisory Committee for Aeronautics, 1953.

[18] LEWIS W, DWIGHT B K, CHARLES P S. A further investigation of the meteorological conditions conducive to aircraft icing: NACA-TN-1424[R]. Washington, D. C.: National Advisory Committee for Aeronautics, 1947.

[19] ALBRECHT H E, BORYS M, DAMASCHKE N, et al. LaserDoppler and phase Doppler measurement techniques[M]. Berlin: Springer, 2003.

[20] ESPOSITO B, MARRAZZO M. Application of PDPA system with different optical configuration to the IWT calibration[C]//Proceedings of the 45th AIAA Aerospace Sciences Meeting and Exhibit. Reston: AIAA, 2007.

[21] RUDOFF R E, BACHALO W, OLDENBURG J, et al. Liquid water content measurements using the phase doppler particle analyzer in the NASA lewis icing research tunnel: AIAA PAPER 93-0298 [R]. Reston: AIAA, 1993.

[22] 陈舒越, 郭向东, 王梓旭, 等. 结冰风洞过冷大水滴粒径测量初步研究[J]. 实验流体力学, 2021, 35(3): 22-29.
CHEN S Y, GUO X D, WANG Z X, et al. Preliminary research on size measurement of super-cooled large droplet in icing wind tunnel[J]. Journal of Experiments in Fluid Mechanics, 2021, 35(3): 22-29(in Chinese).

[23] GLIENKE S, MEI F. Fast cloud droplet probe (FCDP) instrument handbook: DOE/SC-ARM-TR-238[R]. Washington, D. C.: United States Department Energy, 2020.

[24] FABER S, FRENCH J R, JACKSON R. Laboratory and in-flight evaluation of measurement uncertainties from a commercial cloud Droplet probe (CDP)[J]. Atmospheric Measurement Techniques, 2018, 11(6): 3645-3659.

[25] MASON J, STRAPP W, CHOW P. The ice particle threat to engines in flight[C]//Proceedings of the 44th AIAA Aerospace Sciences Meeting and Exhibit. Reston: AIAA, 2006.

[26] REEHORST A, BRINKER D, POLITOVICH M, et al. Progress towards the remote sensing of aircraft icing hazards[C]//Proceedings of the Remote Sensing Applications for Aviation Weather Hazard Detection and Decision Support. Bellingham: SPIE, 2008.

[27] VIVEKANANDAN J, ZHANG G, POLITOVICH M K. An assessment of droplet size and liquid water content derived from dual-wavelength radar measurements to the application of aircraft icing detection[J]. Journal of Atmospheric and Oceanic Technology, 2001, 18(11): 1787-1798.

[28] RATVASKY T, HARRAH S, STRAPP J W, et al. Summary of the high ice water content (HIWC) RADAR flight campaigns: DOT/FAA/TC–19/28[R]. Warrendal: SAE International, 2019.

[29] ZOLLO A L, BUCCHIGNANI E. A Tool for remote detection and nowcasting of in-flight icing using satellite data[C]//Proceedings of the International Conference on Icing of Aircraft, Engines, and Srrucetukes. Warreadale: SAE International, 2023.

Data processing methodology of icing meteorological detection in FAR 25 Appendix C

ZHANG Wenqiang[1]，CHEN Yiyi[1]，LEI Guoqiang[1]，WANG Guangyu[1]，ZHI Yafei[2]，MAO Xuerui[1,*]

(1. School of Mechatronical Engineering，Beijing Institute of Technology，Beijing 100081，China；
2. Corporation Chengdu Aircraft Industry Co., Ltd.，Chengdu 610091，China)

Abstract：Icing can change the shape of the airplane, decrease the lift, and increase the drag. The maneuverability and stability of the airplane are crippled, which severely threatens the flight safety. The application of Appendix C is restricted by regional and seasonal limitations. In addition, a more complex icing condition has been discovered recently and thus Appendix C needs to be consistently revised. This report can provide instructions for the compilation and improvements of China's icing airworthiness regulations. This study introduces the methods for data postprocessing, the mechanism of the conventional icing meteorological detection sensors, and the creation of the FAR 25 Appendix C. Appendix C provides a detailed explanation of how the plots were derived from the meteorological detection data. Finally, the current challenges and future development trends are presented from three aspects: the establishment of icing databases, improving the measurement accuracy of instruments, and anti-icing of unmanned aerial vehicles, thus providing scientific guidance for the development and future direction of meteorological detection technology.

Keywords：FAR25 Appendix C; icing; meteorological detection; data processing; anti-deicing

Received：2023-09-04；**Accepted**：2023-12-05；**Published Online**：2023-12-28 15：10
URL：link.cnki.net/urlid/11.2625.V.20231227.1647.005

* **Corresponding author**. E-mail：xmao@bit.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0578

改进超螺旋滑模自抗扰的PMSM无传感器控制

陈德海[1,2]，刘祥[1]，王海峰[2,*]

(1. 江西理工大学 电气工程与自动化学院，赣州 341000；2. 中国科学院 赣江创新研究院，赣州 341000)

摘 要：针对永磁同步电机矢量控制系统中无传感器控制策略的抗扰动动态性能不足的问题。对超螺旋滑模算法(STA)进行研究，提出一种改进超螺旋滑模自抗扰控制(ISTA-ADRC)算法，该算法在原有超螺旋滑模算法的基础上引入改进Sigmoid函数和线性项。相较于传统超螺旋滑模算法，所提算法具有更好的快速性和鲁棒性。将所提算法嵌入自抗扰控制技术中，提出以改进超螺旋滑模自抗扰为核心的无传感器控制策略，提升无传感器控制法的抗扰性能。仿真实验结果表明:所提算法相比于传统的PI、线性自抗扰(LADRC)和滑模自抗扰算法，具有响应速度快、抖振小、抗扰能力强、稳态性能好的优点。

关 键 词：永磁同步电机；自抗扰控制；超螺旋滑模算法；无位置传感器；跟踪微分器

中图分类号：TM341

文献标志码：A **文章编号**：1001-5965(2025)10-3337-08

高质量发展阶段的社会，对电机运行性能提出了越来越高的要求，特别是对于需要在复杂环境下使用的，如海上大型风力发电机这类特殊工况的电机来说，急需高性能的控制方法来提供支撑。在近年来的高性能的控制方法中,针对永磁同步电机(permanent magnet synchronous motor, PMSM)的矢量控制系统中传统的PI控制和各种新型智能控制方法还存在不足之处：PI控制算法结构简单且易于实现，但由于永磁同步电机的非线性，PI的实际效果会受到参数影响从而无法提供理想的抗扰性能[1-3]；目前较为先进的滑模自抗扰控制，较好的改善了系统的抗扰性能，提高了系统的跟踪精度，然而其在原点邻域内存在拐点不够平滑，容易产生高频抖振，限制抗扰能力且对负载突变的自适应能力不足[4-6]。

为解决复杂环境下的电机运行问题，国内外学者将目光放到了无传感器控制技术上。目前，无传感器控制技术主要分为反电动势法、磁链估计法和电流采样法等[7-9]。但其都不能对外界干扰如噪音等扰动因素做出很好的应对，因此，学者们想到了自抗扰控制技术。在对自抗扰控制的研究中，文献[10]针对系统的扰动，研究了一种线性自抗扰控制(linear auto disturbance rejection control, LADRC)器，较好的改善了系统的抗扰性能。文献[11]提出了将自抗扰技术与滑模控制相结合的方法，减少了需要调节的参数个数，提高了控制效果。文献[12]提出了一种基于新型非线性函数的改进自抗扰控制，提高了系统的跟踪精度。

本文以此为基础提出一种改进超螺旋滑模自抗扰控制(improve super twisting sliding mode auto disturbance rejection control, ISTA-ADRC)算法，并在速度环处针对负载的突变设计滑模观测器，将观测的负载力矩引入到控制系统，构造模型补偿自抗扰控制器(model compensated active disturbance rejection controller, MCADRC)，其次，以线性扩张状态观测器(linear extended state observer, LESO)为核

收稿日期：2023-09-12；**录用日期**：2023-11-21；**网络出版时间**：2023-12-06 08:41
网络出版地址：link.cnki.net/urlid/11.2625.V.20231205.1341.002
基金项目：国家A类重大攻关项目(E210E001010)
*通信作者. E-mail：wanghf@mail.iee.ac.cn

引用格式：陈德海，刘祥，王海峰. 改进超螺旋滑模自抗扰的PMSM无传感器控制[J]. 北京航空航天大学学报，2025，51(10)：3337-3344.
CHEN D H，LIU X，WANG H F. Modified sensorless control of PMSM based on super-twisting sliding mode auto disturbance rejection[J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51(10)：3337-3344 (in Chinese).

心实现无位置传感器的控制。仿真结果表明，本文算法与传统的LADRC、滑模自抗扰及PI控制器相比，本文算法在抗干扰性及精度方面都有良好优势。

1 PMSM数学模型

本文以表贴式永磁同步电机为例，在理想状态下建立α-β静止坐标系下的电流方程为

$$\begin{cases}\dfrac{\mathrm{d}i_\alpha}{\mathrm{d}t}=-\dfrac{R_\mathrm{s}i_\alpha}{L_\mathrm{s}}+\dfrac{u_\alpha}{L_\mathrm{s}}-\dfrac{E_\alpha}{L_\mathrm{s}}\\ \dfrac{\mathrm{d}i_\beta}{\mathrm{d}t}=-\dfrac{R_\mathrm{s}i_\beta}{L_\mathrm{s}}+\dfrac{u_\beta}{L_\mathrm{s}}-\dfrac{E_\beta}{L_\mathrm{s}}\end{cases}\tag{1}$$

式中：u_α、u_β和i_α、i_β分别为α-β静止坐标系上的电压和电流分量；$E_\alpha=-\omega_\mathrm{e}\varphi_\mathrm{f}\sin\theta_\mathrm{e}$和$E_\beta=\omega_\mathrm{e}\varphi_\mathrm{f}\cos\theta_\mathrm{e}$分别为$\alpha$-$\beta$轴上的反电动势，$\theta_\mathrm{e}$为转子位置角，$\omega_\mathrm{e}$为电角速度，$\varphi_\mathrm{f}$为永磁体磁链；$L_\mathrm{s}$和$R_\mathrm{s}$分别为电机上的定子电感和电阻。

d-q坐标系下的电流方程为

$$\begin{cases}\dfrac{\mathrm{d}i_d}{\mathrm{d}t}=-\dfrac{R_\mathrm{s}}{L_d}i_d+\omega_\mathrm{e}\dfrac{L_q}{L_d}i_q+\dfrac{u_d}{L_d}\\ \dfrac{\mathrm{d}i_q}{\mathrm{d}t}=-\dfrac{R_\mathrm{s}}{L_q}i_q-\omega_\mathrm{e}\dfrac{L_d}{L_q}i_d-\dfrac{\omega_\mathrm{e}\varphi_\mathrm{f}}{L_q}+\dfrac{u_q}{L_q}\end{cases}\tag{2}$$

式中：u_d、u_q和i_d、i_q分别为d-q轴的电压和电流分量；L_d、L_q为d-q轴的电感分量。

PMSM的运动方程为

$$\begin{cases}\dfrac{\mathrm{d}\omega_\mathrm{m}}{\mathrm{d}t}=\dfrac{T_\mathrm{e}}{J}-\dfrac{T_\mathrm{L}}{J}-\dfrac{B\omega_\mathrm{m}}{J}\\ T_\mathrm{e}=1.5p_\mathrm{n}\varphi_\mathrm{f}i_q\end{cases}\tag{3}$$

式中：p_n为极对数；T_e和T_L分别为电机的电磁转矩和负载转矩；J和B分别为电机的转动惯量和黏性摩擦系数；ω_m为电机输出的机械角速度。

2 改进自抗扰控制器的设计

2.1 LADRC的数学模型

传统LADRC的数学模型由跟踪微分器(tracking differentiator, TD)、线性扩张状态观测器、线性状态误差反馈控制律(linear state error feedback, LSEF)3部分表示。

TD数学模型为

$$\begin{cases}e_0=v_1-\omega_\mathrm{m}^*\\ \dfrac{\mathrm{d}v}{\mathrm{d}t}=-re_0\end{cases}\tag{4}$$

式中：ω_m^*为输入参考转速；v_1为过渡后的参考速度信号；e_0为误差信号；r为速度因子，当r值增大时，跟踪速度也会增快，但若r取值过大，会导致系统不稳定。

根据式(3)可构造线性扩张状态观测器：

$$\begin{cases}e_1=z_1-\omega_\mathrm{m}\\ \dfrac{\mathrm{d}z_1}{\mathrm{d}t}=z_2-\beta_1e_1+b_0u\\ \dfrac{\mathrm{d}z_2}{\mathrm{d}t}=-\beta_2e_1\end{cases}\tag{5}$$

式中：u为控制量；e_1为转速的误差信号；z_1为观测器对速度的估计；z_2为观测器对扰动的估计；β_1、β_2为观测器的误差校正增益；$b_0=1.5p_\mathrm{n}\varphi_\mathrm{f}$。

文献[13]给出了一阶LADRC的参数整定方法：$\beta_1=2\omega_0$；$\beta_2=\omega_0{}^2$，其中，ω_0为观测器的带宽。

LSEF数学模型为

$$\begin{cases}e_2=v_1-z_1\\ u_0=k_\mathrm{p}e_2\\ u=(u_0-z_2)/b_0\end{cases}\tag{6}$$

式中：u_0为等效控制量；e_2为估计转速与过渡参考转速之间的误差信号；k_p为比例增益系数。

传统的一阶LADRC速度环如图1所示。

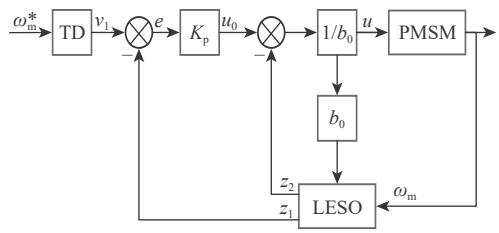


图1 传统一阶LADRC速度环控制结构

Fig. 1 Structure of a traditional first-order LADRC speed loop control

2.2 优化超螺旋算法设计的自抗扰速度控制器

线性自抗扰控制器相对于非线性自抗扰控制器确实具有调参量较少的优势，但同时也可能降低系统的响应速度。这是因为线性自抗扰控制器基于线性化的系统模型设计，可能无法完全适应非线性的系统动态特性。而超螺旋滑模算法在运动轨迹收敛时能够避免传统滑模控制中的抖振现象，从而降低系统的振荡幅度，提高控制的平稳性[14-15]。

2.2.1 传统超螺旋滑模控制算法及优化

设滑模面$s=y-y^*$，则传统超螺旋滑模控制器可表示为

$$\begin{cases}u=-m|s|^c\mathrm{sign}(s)+u_s\\ \dot{u}_s=-n\mathrm{sign}(s)\end{cases}\tag{7}$$

式中：m、n为控制器的增益系数；c为待设计数；u_s为切换控温。

然而，传统的超螺旋滑模控制算法虽然具有较高的鲁棒性和较快的响应速度，但其使用的符号函数具有非光滑性，会使得输入在某些情况下产生高频振荡和不连续的问题，这会对系统的稳定性造成不利影响。针对此，本文将引入线性项并采取改进的sigmoid函数来替代符号函数，其表达式为

$$\text{sigmoid}(x)=\frac{2}{1+\mathrm{e}^{-ax}}-1 \tag{8}$$

优化后的超螺旋滑模控制算法为

$$\begin{cases} u=-m|s|^{c}\text{sigmoid}(s)-k_1 s+u_s \\ \dot{u}_s=-n\text{sigmoid}(s)-k_2 s \end{cases} \tag{9}$$

式中: k_1s 和 k_2s 分别为比例项和积分项; c 一般取 0.5。

2.2.2　改进速度环 ADRC 的设计

根据 2.2.1 节得到的优化后的超螺旋滑模控制算法, 重新设计一阶 ADRC 控制的速度环。由式 (9) 可得改进的 TD 表达式为

$$\begin{cases} e_0=v_1-\omega_{\mathrm{m}}^{*} \\ \dfrac{\mathrm{d}v_1}{\mathrm{d}t}=v_2-m_1\text{sigmoid}(e_0)|e_0|^{c}-k_1 s \\ \dfrac{\mathrm{d}v_2}{\mathrm{d}t}=-n_1\text{sigmoid}(e_0)-k_2 s \end{cases} \tag{10}$$

由式 (3) 可构造扩张系统

$$\begin{cases} \dot{\omega}_{\mathrm{m}}=b_0u+f \\ \dot{f}=h(t) \end{cases} \tag{11}$$

式中: f为总扰动。

由式 (5) 与式 (8) 构建改进后的线性扩张状态观测器:

$$\begin{cases} e_1=z_1-\omega_{\mathrm{m}} \\ \dfrac{\mathrm{d}z_1}{\mathrm{d}t}=z_2-\beta_1e_1+b_0u \\ \dfrac{\mathrm{d}z_2}{\mathrm{d}t}=-\beta_2\text{sigmoid}(e_1) \end{cases} \tag{12}$$

式中: $\beta_1>0$, $\beta_2>0$。当 β_1、β_2 取合适值时, $z_1\to\omega_{\mathrm{m}}$, $z_2\to f$。

对式 (12) 的稳定性进行分析, 假设 $h(t)$ 为定值且 $h(t)\neq0$, 令 $e_1=z_1-\omega_{\mathrm{m}}$, $e_2=z_2-f$。由式 (11) 和式 (12) 可得误差系统:

$$\begin{cases} \dot{e}_1=\dot{z}_1-\dot{\omega}_{\mathrm{m}}=e_2-\beta_1e_1 \\ \dot{e}_2=\dot{z}_2-\dot{f}=-\beta_2\text{sigmoid}(e_1)-h(t) \end{cases} \tag{13}$$

再令 $e_{21}=e_1$, $e_{22}=\mathrm{d}e_1/\mathrm{d}t$, 代入式 (13) 可得

$$\begin{cases} \dot{e}_{21}=e_{22} \\ \dot{e}_{22}=-\beta_2\text{sigmoid}(e_{21})-\beta_1e_{22}-h(t) \end{cases} \tag{14}$$

取式 (14) 的李雅普诺夫 (Lyapunov) 函数:

$$V_1=\int_0^{e_{21}}2\beta_2\text{sigmoid}(x)\mathrm{d}x+e_{22}^2 \tag{15}$$

由积分中值定理可得, 式 (15) 中一定存在一点 $d\in[0,\ e_{21}]$ 使得积分函数为 $2\beta_2\text{sigmoid}(d)e_{21}$, 又 sigmoid 函数在 $(0,+\infty)$ 上恒大于等于 0, 且 e_{22}^2 恒大于等于 0, 故可以得到

$$V_1=2\beta_2\text{sigmoid}(d)e_{21}+e_{22}^2\geqslant0 \tag{16}$$

求导可得

$$\dot{V}_1=-2e_{22}(\beta_1e_{22}+h(t)) \tag{17}$$

由式 (17) 分析可得, 当 $h(t)<0$ 时, 若 $e_{22}>-h(t)/\beta_1$ 或 $e_{22}<h(t)/\beta_1$ 时, 有 $\dot{V}_1<0$。当且仅当 $e_{22}=0$ 时, 有 $\dot{V}_1=0$, 故系统是渐进稳定的。

结合式 (6) 和式 (9), 可以得到超螺旋滑模算法的非线性状态误差反馈控制律:

$$\begin{cases} e_2=z_1-v_1 \\ u_0=-m_2|e_2|^{c}\text{sigmoid}(e_2)-k_3e_2+u_s \\ \dot{u}_s=-n_2\text{sigmoid}(e_2)-k_4e_2 \\ u=(u_0-z_2)/b_0 \end{cases} \tag{18}$$

式中: k_3、k_4、m_2、n_2为增益系数。

同样对该系统进行稳定性分析, 取类二次型 Lyapunov 函数 $V(u_0,u_s)=\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{\xi}$, 其中 $\boldsymbol{\xi}^{\mathrm{T}}=[|e|^{(1/2)},e,u_s]$, $\boldsymbol{P}$ 为实对称矩阵:

$$\boldsymbol{P}=\begin{bmatrix} 4n_2+m_2^2 & m_2k_3 & -m_2 \\ m_2k_3 & 2k_4+k_3^2 & -k_3 \\ -m_2 & -k_3 & 2 \end{bmatrix}$$

根据二次型函数定义 $V(z_1,z_2)$ 是连续正定且径向无界的, 除 $z_1=z_2=0$ 外, 处处可微。所以有

$$\lambda_{\min}\{\boldsymbol{P}\}\|\boldsymbol{\xi}\|_2^2\leqslant V(x)\leqslant\lambda_{\max}\{\boldsymbol{P}\}\|\boldsymbol{\xi}\|_2^2 \tag{19}$$

式中: $\|\boldsymbol{\xi}\|_2^2=|z_1|+z_1^2+z_2^2$为欧几里得范数。

对 Lyapunov 函数求导可得

$$\dot{V}(z_1,z_2)=\dot{\boldsymbol{\xi}}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{\xi}+\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{P}\dot{\boldsymbol{\xi}}=\frac{-1}{|z_1|^{1/2}}\left(\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\xi}+\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{\xi}\right) \tag{20}$$

式中:

$$\boldsymbol{A}=\begin{bmatrix} 2n_2+m_2^2 & 0 & -m_2 \\ 0 & 2k_4+5k_3^2 & -3k_3 \\ -m_2 & -3k_3 & 1 \end{bmatrix}$$

$$\boldsymbol{B}=\begin{bmatrix} n_2+2m_2^2 & 0 & 0 \\ 0 & k_4+k_3^2 & -3k_3 \\ 0 & -3k_3 & 1 \end{bmatrix}$$

可见若 $\boldsymbol{A}>\boldsymbol{0}$, $\boldsymbol{B}>\boldsymbol{0}$, 那么 $\dot{V}<0$, 可见当 m_2、n_2、k_3、k_4 均大于 0 且满足式(21), 即可证明系统稳定。

$$4n_2k_4>\left(8n_2+9m_2^2\right)k_3 \tag{21}$$

然而, 当速度环受到的扰动较大时, 会增加观测器对扰动的估计负担, 进一步影响观测的精确度。故本文设计了一个新型负载转矩观测器。将观测出来的扰动反馈到速度环之中进行补偿, 从而增加系统的抗扰性能。

由式 (3) 可以构建滑模负载转矩观测器, 其表达式为

$$\begin{cases} e_3=\omega_{\mathrm{m}}-\hat{\omega}_{\mathrm{m}} \\ \dot{\hat{\omega}}_{\mathrm{m}}=\dfrac{T_{\mathrm{e}}}{J}-\dfrac{B\hat{\omega}_{\mathrm{m}}}{J}-\dfrac{\hat{T}_{\mathrm{L}}}{J}-k_5\text{sigmoid}(e_3) \\ \dot{\hat{T}}_{\mathrm{L}}=k_6\text{sigmoid}(e_3) \end{cases} \tag{22}$$

式中: $\hat{\omega}_{\mathrm{m}}$为角速度估计值; $\hat{T}_{\mathrm{L}}$为负载转矩估计值。

最终的速度环控制器结构如图2所示。

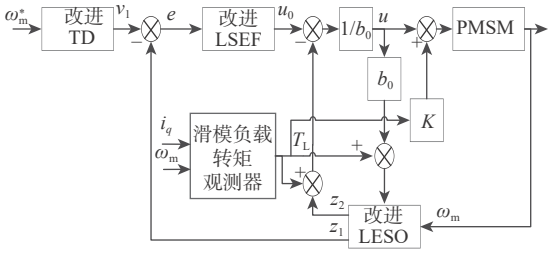


图2 改进一阶ADRC控制的速度环结构

Fig. 2 Structure of an improved first-order ADRC speed loop control

2.3 自抗扰电流环的设计

由于PMSM的双闭环控制系统中电流环主要受到参数和输入误差的影响，所以无需再设计负载观测器对其进行补偿，电流环结构如图3所示。

可见一阶自抗扰系统的结构是一致的，故可以根据第2节的内容使用式(2)进行设计。

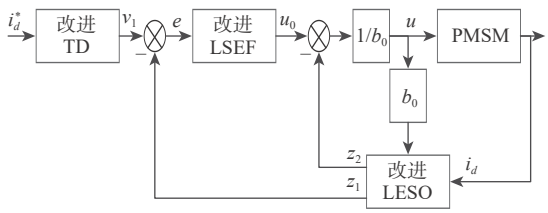


图3 改进ADRC电流环结构

Fig. 3 Improved structure of ADRC current loop

3 无位置传感器的设计

传统的机械接触式传感器容易因外界的各种因素而损坏，从而导致永磁同步电机故障，不利于系统的稳定性。文献[16]提出了滑模观测器来解决这一问题。然而，传统的滑模观测器使用的符号函数会使系统出现高频的抖振。为减少抖振提升抗扰性能，本文使用LESO来对反电动势进行观测。

根据式(1)可以构建LESO反电动势观测器，其表达式为

$$\begin{cases} E_\alpha = z_{11} - i_\alpha \\ \dot{z}_{11} = z_{21} - \beta_1 E_\alpha + b_0 u + f_1 \\ \dot{z}_{21} = -\beta_2 E_\alpha \end{cases} \tag{23}$$

$$\begin{cases} E_\beta = z_{12} - i_\beta \\ \dot{z}_{12} = z_{22} - \beta_1 E_\beta + b_0 u + f_2 \\ \dot{z}_{22} = -\beta_2 E_\beta \end{cases} \tag{24}$$

式中：z_{11}、z_{12}为α-β轴上的电流观测值；f_1+z_{21}、f_2+z_{22}为α-β轴上总扰动估计值，其中，$f_1=-R_s i_\alpha/L_s$，$f_2=-R_s i_\beta/L_s$，因此，可以从z_{21}和z_{22}提取出α-β轴上的反电动势估计值$\hat{E}_\alpha$与$\hat{E}_\beta$：

$$\begin{cases} \hat{E}_\alpha = -L_s z_{21} \\ \hat{E}_\beta = -L_s z_{22} \end{cases} \tag{25}$$

得到反电动势估计值后，为提高转速和转子位置的精确度，本文采用经过归一化的锁相环结构来提取PMSM中的转子信息。由于LESO对反电动势的估计会对相位造成一定的延迟，所以需要对转子位置进行补偿。位置补偿表达式为

$$\theta_{\text{lag}} = 2\arctan\left(\hat{\omega}_e/\omega_0\right) \tag{26}$$

式中：$\hat{\omega}_e$为估计电角度；θ_{lag}为滞后补偿角度。

图4为归一化锁相环的结构，其中，E为反电动势。

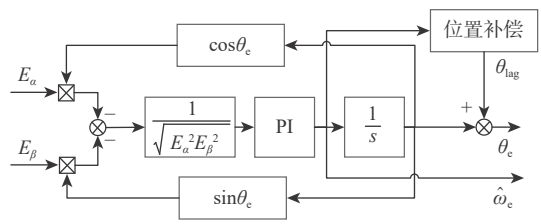


图4 归一化锁相环结构

Fig. 4 Structure of quadrature phase-locked loop

4 仿真实验验证

为验证本文设计的ISTA-ADRC在PMSM双闭环控制系统中对转速突变和负载突变时的抗扰性能的可行性和有效性，基于MATLAB/Simulink搭建仿真模型。图5为基于优化超螺旋算法改进的ADRC的PMSM无传感器矢量控制框图。

控制器方面，本文使用了传统PI控制器、LADRC、传统的STA-ADRC和改进的ISTA-ADRC在无传感器控制的情况下做对比实验验证。其中，转速环中，PI的控制参数：k_p=0.04，k_i=2；LADRC控制参数：k_p=100，ω_0=300；STA-ADRC的控制参数：m_2=500，n_2=10，ω_0=200。ISTA-ADRC的控制参数：m_2=600，n_2=50，ω_0=1 000。电流环中，PI的控制参数：k_p=9.35，k_i=3 162.5；LADRC的控制参数：k_p=800，ω_0=4 000；STA-ADRC的控制参数：m_2=800，n_2=5 000，ω_0=4 000。ISTA-ADRC的控制参数：m_2=2 000，n_2=50，ω_0=8 000。

表1为仿真实验所用的PMSM具体参数。

首先，为验证电机在无传感器控制下的各控制器在启动阶段和稳定阶段的转速响应效果，给定电机参考转速为1 000 rad/min，在空载启动的情况下，电机转速n与时间t的关系曲线如图6所示。

由图6可知，在启动阶段，传统PI控制的转速有较大的超调且稳定较慢，LADRC几乎无超调，但是其系统响应的快速性较差，STA-ADRC虽有较快

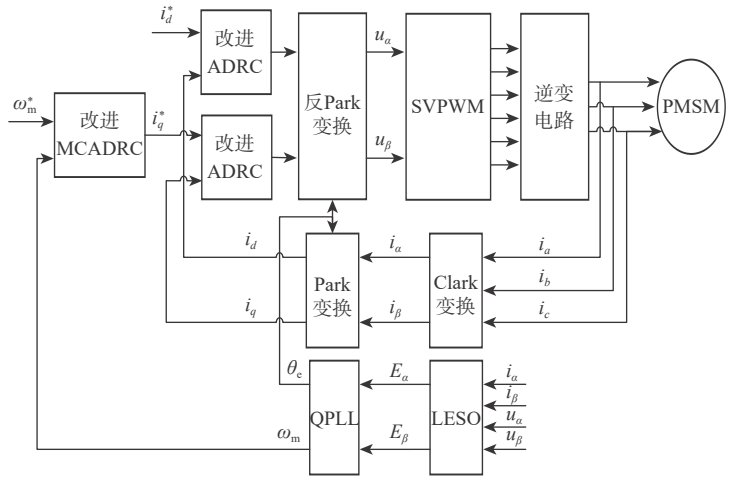


图 5　基于改进 STA-ADRC 的 PMSM 结构

Fig. 5　PMSM structure of improved STA-ADRC

表 1　PMSM 参数

Table 1　PMSM parameters

参数	数值
转动惯量J/(kg·m^2)	0.003
永磁体磁链φ_f/Wb	0.175
d-q轴电感L_{dq}/mh	8.500
定子电阻R_s/Ω	2.875
极对数p_n	4.000
阻尼系数B/ (N·m·s)	0.008

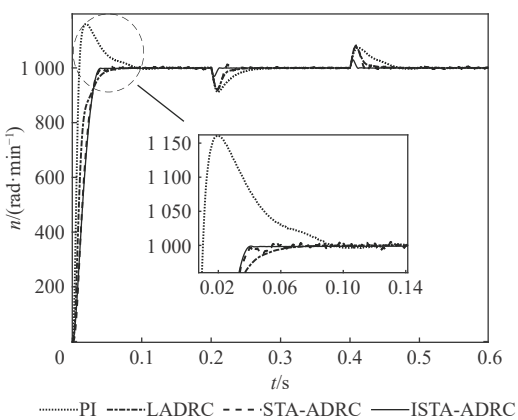


图 6　各控制器下的电机转速响应曲线

Fig. 6　Motor speed response curves under different controllers

的响应速度, 但也有较大的抖振。而本文的 ISTA-ADRC 在响应速度和稳态性能方面都具有显著的改善效果, 较好的提高了系统的抗扰性能。

其次, 为验证改进的 STA-ADRC 在负载突变时的抗扰性能, 先将系统仿真时间设为 0.6 s, 电机参考转速 n 设置在 1 000 rad/min, 然后, 在 0.2 s 时给系统突然增加 5, 10, 25 N 的负载, 0.4 s 时减去相应负载。最终可以得到各控制器在面对不同大小的负载扰动时的控制效果, 如图 7 所示。

由图 7(a) 可知, PI 控制面对 5 N 的负载扰动时的转速波动为 89 rad/min, 稳定时间为 0.1 s, LADRC 的转速波动为 79 rad/min, 稳定时间为 0.08 s, 传统 STA-ADRC 的转速波动为 82 rad/min, 稳定时间为 0.06 s, 而改进后算法的转速超调量为 3.1%, 稳定时间为 0.02 s。同样, 由图 7(c) 可知, 面对 10 N 的负载突变, 改进后的算法在超调与稳定时间上仍优于其他控制器。而从图 7(e) 可见, 在面对 25 N 的负载突变时, STA-ADRC 在稳定时转速有较小的抖振。因此, 本文 ISTA-ADRC 的抗负载扰动的动态性能在负载为 0～25 N 的范围内有较大的优势。

图 8 为在空载启动时的 q 轴电流 I 的波形, 由图 8 可知, PMSM 在空载启动的情况下, q 轴电流会迅速达到最大值, 而 PI 控制的电流是传统 LADRC 的 2 倍, 是 STA-ADRC 与 ISTA-ADRC 的 3 倍。这不可避免的导致了硬件设计成本的增加。而相较于 STA-ADRC, 本文 ISTA-ADRC 在稳定时没有较大的抖振。故本文控制算法在空载启动时也有良好的性能。

为验证本文的 ISTA-ADRC 在转速突变时的抗扰性能。在转速突变的仿真实验中, 首先, 设置仿真时间为 0.8 s, 电机参考转速为 1 000 rad/min; 其次, 在 0.2 s 加入恒定为 10 N 负载的同时使转速突降到 600 rad/min; 然后, 于 0.4 s 时突增到 1 600 rad/min, 0.6 s 时下降到 800 rad/min; 最终, 得到电机在转速扰动下的转速响应曲线如图 9 所示。

由图 9 可知, 同 10 N 负载的情况下, 传统的 PI 控制在面对突增 800 rad/min 工况时的转速波动为 402 rad/min, 稳定时间为 0.1 s, LADRC、传统 STA-ADRC 和 ISTA-ADRC 几乎无超调, 稳定时间分别为 0.07 s、0.055 s 和 0.038 s。

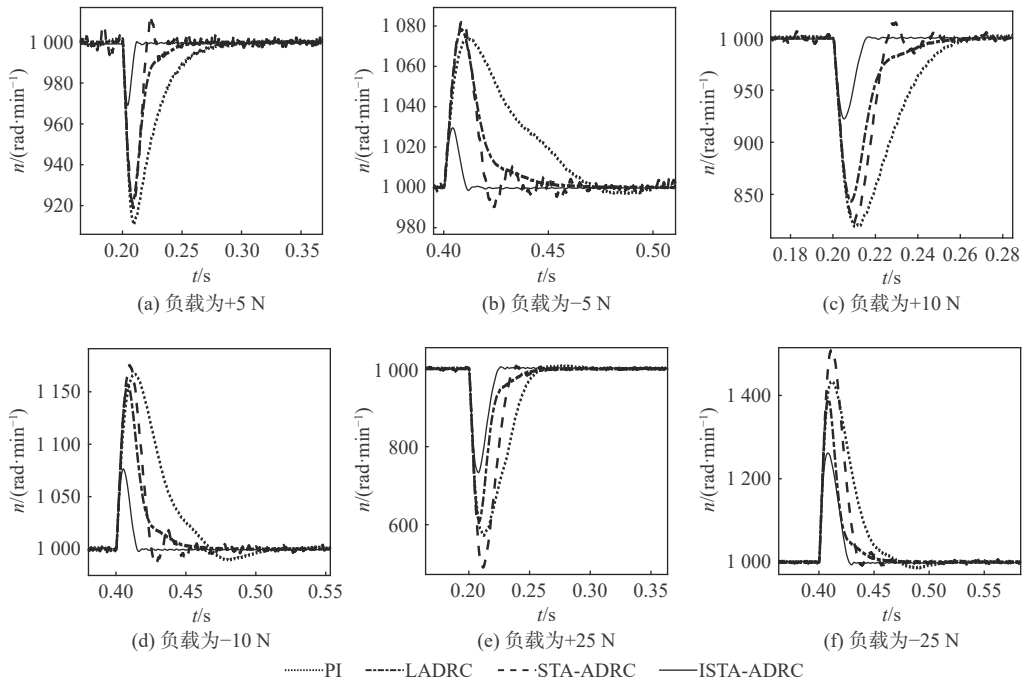


图 7 突增和突降负载的转速抗扰性能曲线

Fig. 7 Speed resilience curves under sudden load increase and decrease

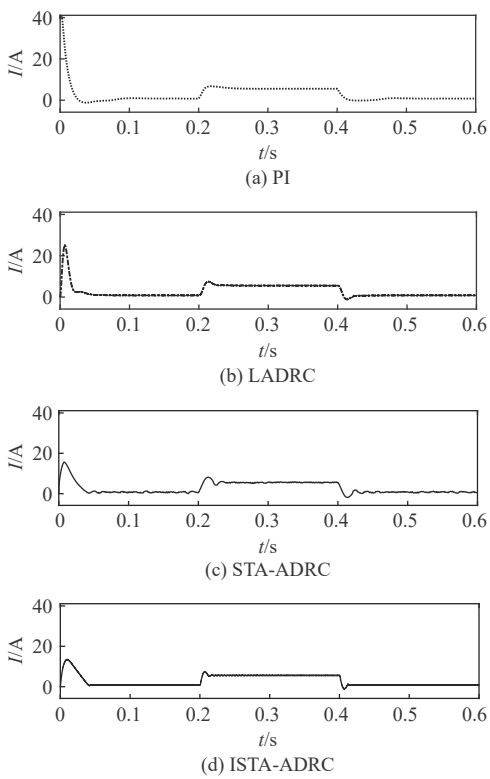


图 8 q 轴电流波形对比

Fig. 8 Comparison of q-axis current waveforms

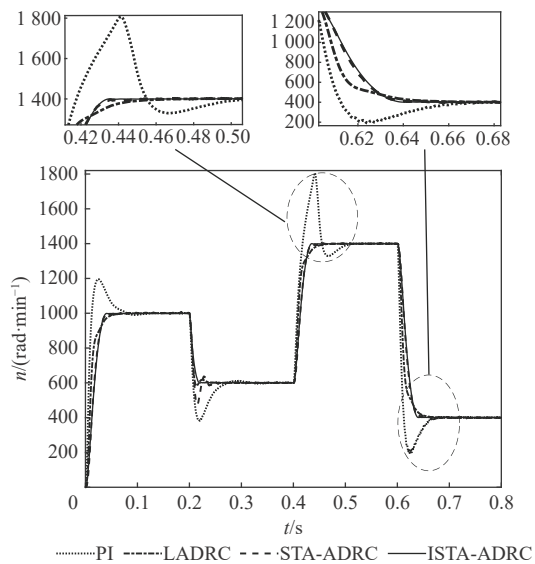


图 9 转速突变时的响应曲线

Fig. 9 Response curves of speed transients

在突降 1 000 rad/min 的工况时，PI 超调量为 20.9%，稳定时间为 0.68 s，LADRC、传统 STA-ADRC 和 ISTA-ADRC 的稳定时间分别为 0.089 s、0.078 s 和 0.068 s，且 STA-ADRC 在稳定时有约 1 rad/min 的抖振。故对比得出本文 ISTA-ADRC 的抗速度扰动的动态性能在转速波动为−1 000 rad/min 至 +800 rad/min 的区间内有良好的效果。

最后，使用改进算法验证改进的 LESO 在无传感器控制下对电机转速和转子位置的估计精度。设定观测器的带宽 ω_0 为 6 000。图 10 为观测的转子位置精度。图 11 为 LESO 估计转速与实际转速的误差。

由图 10 可知，通过改进 LESO 观测的反电动势在经过归一化处理和转子位置补偿的正交锁相环所提取出来的转子位置角度误差较小。而由图 11 可知，基于改进 LESO 估计的稳态转速波动在 ±0.06 rad/min 之间，故该算法抖振小，精度高。

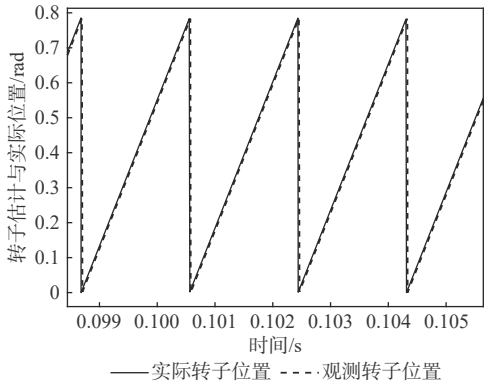


图 10　LESO 估计转子位置与实际位置波形对比

Fig. 10　Comparison of estimated rotor position and actual position waveforms with LESO

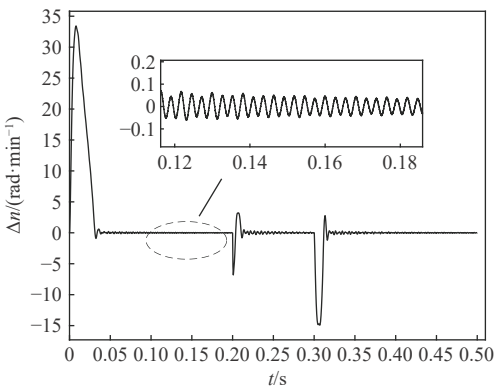


图 11　LESO 估计转速误差

Fig. 11　Error in estimated speed with LESO

5　结　论

1) 本文算法运用到自抗扰控制中使得系统在对 0～25 N 的负载突变及−1 000～+800 rad/min 区间内的转速突变有了更强的抗干扰能力, 提高了整个系统的控制性能。例如: 系统在面对 5 N 的负载扰动时, 转速超调量为 3.1%, 稳定时间为 0.02 s。

2) LESO 无位置传感器与正交锁相环的设计将转速的稳态误差限制在了±0.06 rad/min, 使系统在恶劣环境下依然能具有较好的稳态和动态性能。

3) 多种工况下的对比实验分析表明, 本文算法响应速度快、稳态性能好、抖振小、抗扰能力强, 能进一步降低硬件成本, 减少系统维护成本。

参考文献（References）

[1] 张硕, 李雪荣, 崔星, 等. 永磁同步电机参数扰动抑制方法[J]. 北京理工大学学报, 2022, 42(2): 184-191.
ZHANG S, LI X R, CUI X, et al. Parameter disturbance suppression method of PMSM[J]. Transactions of Beijing Institute of Technology, 2022, 42(2): 184-191(in Chinese).

[2] WANG H Z, XU S, HU H S. PID controller for PMSM speed control based on improved quantum genetic algorithm optimization[J]. IEEE Access, 2023, 11: 61091-61102.

[3] GAO P, ZHANG G M, OUYANG H M, et al. An adaptive super twisting nonlinear fractional order PID sliding mode control of permanent magnet synchronous motor speed regulation system based on extended state observer[J]. IEEE Access, 2020, 8: 53498-53510.

[4] PINDORIYA R M, TEJAN K V, RAJPUROHIT B S. Speed control of sensorless PMSM drive using adaptive current control prediction technique[J]. Electrical Engineering, 2023, 105(2): 1209-1221.

[5] LIU G, CHEN B D, WANG K, et al. Selective current harmonic suppression for high-speed PMSM based on high-precision harmonic detection method[J]. IEEE Transactions on Industrial Informatics, 2019, 15(6): 3457-3468.

[6] XIE F, HONG W J, QIU C M. Speed fluctuation suppression of PMSM using active disturbance rejection and feedback compensation control[J]. IET Electric Power Applications, 2021, 15(8): 1056-1067.

[7] 王志福, 田昊, 黄丽敏, 等. 基于卡尔曼观测器的电机模拟器接口电流自抗扰无差拍预测控制算法[J]. 北京理工大学学报, 2023, 43(9): 912-925.
WANG Z F, TIAN H, HUANG L M, et al. Kalman observer-based active disturbance rejection dead-beat predictive control algorithm for electric machine emulator interface current[J]. Transactions of Beijing Institute of Technology, 2023, 43(9): 912-925(in Chinese).

[8] FU X H, GU S D, HE H, et al Compensation and rejection of cogging torque of PMSM based on novel internal model observer in series with extended state observer[J]. Proceedings of the Chinese Society of Electrical Engineering, 2022, 42(23): 8729-8739.

[9] SHEN H Y, WU X Z, WANG, H F, et al. Suppression of third harmonic caused by SPWM for a nine-phase open-end winding permanent magnet synchronous motor[J]. Proceedings of the Chinese Society of Electrical Engineering, 2023, 43(5): 1992-2000.

[10] 康尔良, 吴炳道, 禹聪. 基于 LADRC 控制的永磁同步电机无位置传感器研究[J]. 电机与控制学报, 2023, 27(2): 69-78.
KANG E L, WU B D, YU C. Research on sensorless of permanent magnet synchronous motor based on LADRC control[J]. Electric Machines and Control, 2023, 27(2): 69-78(in Chinese).

[11] 杜海明, 高函, 胡智宏, 等. 永磁同步电机超螺旋滑模自抗扰调速系统设计[J]. 重庆理工大学学报 (自然科学), 2022, 36(10): 216-222.
DU H M, GAO H, HU Z H, et al. Design of the active disturbance rejection speed regulation system of super-twisting sliding mode for permanent magnet synchronous motors[J]. Journal of Chongqing University of Technology (Natural Science), 2022, 36(10): 216-222 (in Chinese).

[12] 邱凌烽, 杨凯, 杨帆, 等. 基于改进重复自抗扰控制的永磁同步电机谐波抑制[J]. 组合机床与自动化加工技术, 2023(1): 1-5.
QIU L F, YANG K, YANG F, et al. Harmonic suppression of permanent magnet synchronous motor based on revised repetitive-auto disturbance rejection control[J]. Modular Machine Tool & Automatic Manufacturing Technique, 2023(1): 1-5(in Chinese).

[13] GAO Z Q. Scaling and bandwidth-parameterization based controller tuning[C]//Proceedings of the American Control Conference. Piscataway: IEEE Press, 2003: 4989-4996.

[14] 宋昱霖, 李洪文, 邓永停. 永磁同步电机快速超螺旋滑模控制[J]. 控制工程, 2023, 30(1): 62-69.
SONG Y L, LI H W, DENG Y T. Fast super twisting sliding mode

control of permanent magnet synchronous motor[J]. Control Engineering of China, 2023, 30(1): 62-69(in Chinese).

[15] LIU L, JIN D S, SI J W, et al. A novel nonsingular fast terminal sliding mode observer combining I-F method for wide-speed sensorless control of PMSM drives[J]. IET Power Electronics, 2023, 16(5): 843-855.

[16] 程勇, 李思卿, 李森豪. 永磁同步电机干扰观测复合滑模控制技术[J]. 电机与控制应用, 2023, 50(3): 39-46.
CHENG Y, LI S Q, LI S H. Compound sliding mode control technology of permanent magnet synchronous motor based on disturbance observation[J]. Electric Machines & Control Application, 2023, 50(3): 39-46(in Chinese).

Modified sensorless control of PMSM based on super-twisting sliding mode auto disturbance rejection

CHEN Dehai[1, 2], LIU Xiang[1], WANG Haifeng[2, *]

(1. School of Electrical Engineering and Automation, Jiangxi University of Science and Technology, Ganzhou 341000, China;
2. Ganjiang Innovation Academy, Chinese Academy of Sciences, Ganzhou 341000, China)

Abstract: The dynamic performance of the sensorless control strategy in the vector control system of a permanent magnet synchronous motor is insufficient. Firstly, the super twisting sliding mode algorithm (STA) is studied, and super twisting sliding mode auto disturbance rejection control (STA-ADRC) algorithm is proposed, which introduces an improved Sigmoid function and linear term on the basis of the original super spiral sliding mode algorithm. Compared with the traditional super spiral sliding mode algorithm, the proposed algorithm has better rapidity and robustness. Secondly, a sensorless control strategy based on enhanced super twisting sliding mode auto disturbance rejection control is suggested to enhance the disturbance rejection performance of the sensorless control. According to the simulation results, the proposed algorithm has the following benefits over the conventional PI, liner auto disturbance rejection control (LADRC), and sliding mode auto disturbance rejection control: quick response time, minimal buffeting, robust disturbance immunity, and good steady-state performance.

Keywords: permanent magnet synchronous motor; auto disturbance rejection control; super twisting sliding mode algorithm; sensorless; tracking differentiator

Received: 2023-09-12; **Accepted**: 2023-11-21; **Published Online**: 2023-12-06 08:41
URL: link.cnki.net/urlid/11.2625.V.20231205.1341.002
Foundation item: National Class A Major Research and Development Project (E210E001010)
* **Corresponding author**. E-mail: wanghf@mail.iee.ac.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0579

考虑双时间窗特性的机场多车型摆渡车调度优化

张文义，唐雨拉尔，王旭兰，周静，边可，刘志硕*
(北京交通大学 综合交通运输大数据应用技术交通运输行业重点实验室，北京 100044)

摘 要：为提升摆渡车运行的灵活性、增强车辆调度方案的能效和实践可操作性，综合考虑多车型和机场摆渡车运输服务的双时间窗特性，从任务点视角构建机场摆渡车调度问题的混合整数线性规划模型。该模型可视为一个考虑多车型和双时间窗且需求可拆分的车辆路径问题，是一个复杂的NP-hard问题，为此，提出一种改进的自适应大邻域搜索(ALNS)算法。以北京首都国际机场东飞行区为背景的案例仿真结果表明：所提算法具有良好的稳定性和优化效果，200次迭代下10轮次计算实验结果的标准差与平均值之比仅为5.6%，优化前后对应的最优目标函数值下降达54%；多车型调度方案显著优于人工和单一车型调度方案，在车辆容量利用能效上较人工、大车型和小车型方案分别提升了98.3%、31.3%和22.2%，在运行总成本上较人工、大车型和小车型方案则分别下降了48.3%、23.4%和23.5%。

关 键 词：航空运输；机场；摆渡车调度；多车型；双时间窗；自适应大邻域搜索
中图分类号：U268.6
文献标志码：A **文章编号**：1001-5965(2025)10-3345-09

国内民航运输需求的迅速攀升给有关系统的保障能力和效率提出了新的更高要求，然而，中国民航运输机场的摆渡车等稀缺资源调度目前仍主要依靠人工，难以有效应对大规模复杂运行场景，且无法满足管理科学化、智能化和精细化的发展要求，严重制约了机场管理效能的进一步提升。鉴于此，研发科学高效、适应能力更强的机场摆渡车智能化调度方法和系统是中国民航领域目前亟待解决的关键问题，具有重要的理论和实践价值。

国内外有关机场摆渡车调度问题的研究近年来快速增长。Kuhn等[1]采用融入分支定界程序的遗传算法研究了单车型摆渡车调度问题；杨文东等[2]应用仿真优化方法探讨了多目标摆渡车调度问题；冯霞等[3]探讨了加油车和摆渡车的协同调度问题；何丹妮[4]设计一种两阶段启发式算法来优化摆渡车调度，以减少用车数量和行驶距离；Han等[5]基于虚拟航班和摆渡车场为节点构建摆渡车运力网络，构建了混合整数规划模型，以最小化摆渡车数量；Zhao等[6]引入虚拟航班，经过简化将问题转化为可有效求解的网络最大流问题；Han等[7]则利用遗传算法来求解双目标摆渡车调度问题；Liu等[8]提出一种双目标混合整数规划模型，并设计了改进的非支配排序遗传算法对模型进行求解；Zhu等[9]探讨了航班到达时间不确定条件下机场场面摆渡车和加油车的协同调度问题；冯明端等[10]面向机场地面特种设备车辆，构建了带有容量限制和时间窗的联合调度模型；Han等[11]提出一种两阶段优化方法来提升航班不确定下的摆渡车调度准点率；Bao等[12]研究了以最小化用车规模为目标的机场燃油和电动特种车辆混合调度问题，并提出一种改进的自适应大邻域搜索(adaptive large neighborhood search, ALNS)算法来求解。

收稿日期：2023-09-12；**录用日期**：2023-10-13；**网络出版时间**：2023-11-06 11:13
网络出版地址：link.cnki.net/urlid/11.2625.V.20231103.1634.001
基金项目：国家自然科学基金(72271018)；北京首都国际机场股份有限公司资助项目(T20L00690)
* **通信作者**. E-mail：zhsliu@bjtu.edu.cn

引用格式：张文义，唐雨拉尔，王旭兰，等. 考虑双时间窗特性的机场多车型摆渡车调度优化[J]. 北京航空航天大学学报，2025，51(10)：3345-3353. ZHANG W Y，TANG Y L E，WANG X L，et al. Optimization of multi-type airport ferry vehicles schedules with double service time windows characteristic[J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51(10)：3345-3353 (in Chinese).

现有有关机场摆渡车调度问题的研究大多只关注单车型问题；此外，为简化问题建模和求解，相关约束也未考虑机场摆渡车运输服务的双时间窗特性，从而使得研究问题与实际不尽相符。本文在现有研究基础上，进一步考虑多类型摆渡车和摆渡车运输的双服务时间窗要求，构建了更符合中国实际且适应场景更广的机场摆渡车调度优化模型，并提出改进的自适应大邻域搜索算法，基于北京首都国际机场的案例研究结果表明，本文算法稳定有效，并且多车型调度方案显著优于人工和单一车型调度方案。本文可为机场摆渡车调度提供辅助决策支持，助力机场降本增效，同时提升其管理的科学化和智能化水平。

1 问题描述

机场摆渡车调度问题旨在给未来一定时间内要进/离港的远机位航班调派摆渡车，使得各航班及时被服务且用车数量（或包含空驶距离）最小。现有研究一般只讨论同一类型（按载客容量划分）摆渡车的调度问题，并且假定每个航班只有一个服务时间窗。在此场景下，单一航班的摆渡运输任务可直接被拆分为固定数量的任务单元（即车辆往返航班停机位和航站楼间的一次“上客-行驶-卸客”作业过程）。例如，假定某一进港航班载有乘客340人，而摆渡车载客容量为100人，那么服务该航班将需要4辆车，对应于4个独立任务单元。如此一来，原问题即可转化为带单一时间窗约束的标准车辆路径问题[13]。

本文的摆渡车调度问题与现有研究有2个显著区别：①本文考虑不同种类车型；②本文还考虑了摆渡车运输服务的双时间窗特性。前者可减少单一车型调度因“大车小用”导致的资源浪费，进而节约成本；后者使得问题更符合实际，所得方案的可操作性更强。然而，考虑二者也将使得研究问题的复杂性明显提升。考虑多车型使得单一航班的用车方案不再固定，问题将转变为需求可拆分这一更复杂情形，并且各航班本身的运输成本（即往返航班与航站楼间的摆渡车运行成本）也将因此转为可变。摆渡车服务的双时间窗特性源于中国民用航空局颁布的《航班安全运行保障标准》中的有关规定[14]：“对于进港摆渡车，首车应在航班计划/预计到港时间前5 min到达机位或指定等待区域，其他摆渡车应在前车驶离后2 min内到位；对于出港摆渡车，首车不应晚于相应时间点（随机型不同而变化）前到达登机口或指定等待区域，其他摆渡车应在前车驶离后2 min内到位”。规定对于首辆摆渡车与后续车辆的到达时间要求明显不同，并且无法用单一服务时间窗来刻画。因此，本文为各航班（注：过站航班被定义为一个进港航班和一个离港航班）定义了2个服务时间窗，即前置时间窗$[0,T_{\text{begin}}-\mu]$和后置时间窗$[T_{\text{begin}},T_{\text{end}}]$，其中：$T_{\text{begin}}$对于进港航班表示其计划/预计到港时间，对于离港航班则表示规定中要求的时间点；μ为时间提前量，对于进港和离港航班分别取值为5 min和0；$T_{\text{end}}-T_{\text{begin}}$表示航班后置服务时间窗长度，本文假定其值随航班乘客数量梯次递增，如每增加100人，其值增加2 min。规定中要求后续车辆的到发间隔不超过2 min，无折扣处置该规定将给问题建模和求解带来较大挑战，后置时间窗可视为在不显著影响服务水平下对该规定的简化近似处理。由此，对于任一航班，必须有一辆摆渡车在前置时间窗内到达，其他车辆只需在前置和后置时间窗内选择其一到达即可，因此，理论上允许后置时间窗内无车辆到达，这也符合管理规定要求。考虑双服务时间窗使得车辆的到访时间决策变得更加灵活，将增加研究难度。

如上所述，若将航班摆渡运输任务（一个航班任务可包括多个任务单元，取决于航班载客量和用车方案）视为客户点，本文的摆渡车调度问题将可视为一个考虑多车型和双时间窗且需求可拆分的车辆路径问题，其目的是寻找可行且总成本最低的车辆调度方案（含车辆数量、类型、服务航班次序和时间）。

图1为本文摆渡车调度问题的一个示例，图中包含2种摆渡车、10个待服务航班（如图中各飞机图标所示，每个图标对应一个航班摆渡运输任务）。图1采用不同线型的线段来标示不同车辆服务（或访问）相邻航班的顺序，车辆到访某航班代表执行

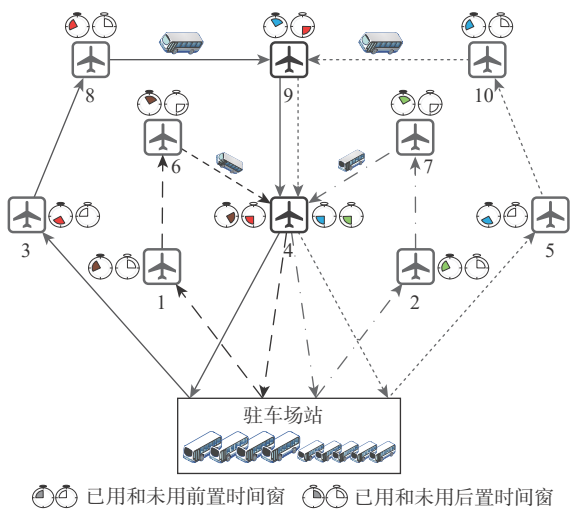


图1 本文摆渡车调度问题示例

Fig. 1 An illustrative example for the ferry vehicle scheduling problem

该航班的一个摆渡任务单元（即一次“上客-行驶-卸客”作业过程），同线型线段依次相连形成一辆摆渡车的完整服务路径，路径数量、用车类型、航班组成及访问次序的不同将对应不同的摆渡车调度方案。如图 1 所示的方案共启用了 4 辆摆渡车，其中，大型车和小型车各 2 辆；受航班载客量、位置和系统总成本等的影响，航班 4 和航班 9 被分派多辆车进行服务（前者用到 2 种车型，后者只用到大车），其他航班只被分派一辆大型或小型车进行服务；受有关运行保障规定约束[14]，所有航班的前置时间窗内都有车辆到访服务。

2　数学建模

本文问题建模时考虑的主要假设有：①各航班的机位位置和航站楼接驳口、载客量、计划/预计到/离港时刻等信息固定且已知；②机场场面上任意两点（含机位、航站楼接驳口等）间的摆渡车行驶路线固定且已知，相应时长和距离将作为参数纳入本文模型中；③摆渡车按容量划分为多种类型，不同类型车辆的行驶速度相同且已知；④同类车辆的上客（和下客）时长相同，不受旅客数量的影响；⑤摆渡车场站只有一个，并且车辆中途无需返回场站补能。

基于上述假设，本文考虑双服务时间窗特性的机场多车型摆渡车调度问题可表示为如下的混合整数线性规划模型：

$$\min g(x)=\sum_{v\in V}\sum_{f\in F}\sum_{k\in K}c_{\mathrm{fix}}^{k}\delta_{k}^{v}x_{o,f}^{v}+\sum_{v\in V}\sum_{i\in F_o}\sum_{j\in F_d}\sum_{k\in K}e_k\delta_k^v(l_{\mathrm{deadHD}}^{i,j}+\lambda_k l_{\mathrm{move}}^{i})x_{i,j}^{v} \tag{1}$$

$$\sum_{f\in F}x_{o,f}^{v}\leqslant 1 \qquad \forall v\in V \tag{2}$$

$$\sum_{i\in F_o}x_{i,f}^{v}\leqslant 1 \qquad \forall v\in V,f\in F \tag{3}$$

$$\sum_{f\in F}x_{o,f}^{v}=\sum_{f\in F}x_{f,d}^{v} \qquad \forall v\in V \tag{4}$$

$$\sum_{i\in F_o}x_{i,f}^{v}=\sum_{j\in F_d}x_{f,j}^{v} \qquad \forall v\in V,f\in F \tag{5}$$

$$\sum_{v\in V}\sum_{i\in F_o}\sum_{k\in K}R_k\delta_k^v x_{i,f}^{v}\geqslant q_f \qquad \forall f\in F \tag{6}$$

$$\sum_{v\in V}y_{f,1}^{v}\geqslant 1 \qquad \forall f\in F \tag{7}$$

$$y_{f,1}^{v}+y_{f,2}^{v}=\sum_{i\in F_o}x_{i,f}^{v} \qquad \forall v\in V,f\in F \tag{8}$$

$$T_f^v\leqslant T_{\mathrm{begin}}^{f}-u_f+M(1-y_{f,1}^{v}) \qquad \forall v\in V,f\in F \tag{9}$$

$$T_f^v\geqslant T_{\mathrm{begin}}^{f}-M(1-y_{f,2}^{v}) \qquad \forall v\in V,f\in F \tag{10}$$

$$T_f^v\leqslant T_{\mathrm{end}}^{f}+M(1-y_{f,2}^{v}) \qquad \forall v\in V,f\in F \tag{11}$$

$$T_f^v\geqslant T_o^v+T_{\mathrm{move}}^{o,f}-M(1-x_{o,f}^{v}) \qquad \forall v\in V,f\in F \tag{12}$$

$$T_f^v\leqslant T_o^v+T_{\mathrm{move}}^{o,f}+M(1-x_{o,f}^{v}) \qquad \forall v\in V,f\in F \tag{13}$$

$$T_{f'}^v\geqslant T_{\mathrm{begin}}^{f}+s_f+T_{\mathrm{move}}^{f,f'}-M(2-x_{f,f'}^{v}-y_{f,1}^{v}) \qquad \forall v\in V,f,f'\in F \tag{14}$$

$$T_{f'}^v\geqslant T_f^v+s_f+T_{\mathrm{move}}^{f,f'}-M(2-x_{f,f'}^{v}-y_{f,2}^{v}) \qquad \forall v\in V,f,f'\in F \tag{15}$$

$$T_d^v\geqslant T_{\mathrm{begin}}^{f}+s_f+T_{\mathrm{move}}^{f,d}-M(2-x_{f,d}^{v}-y_{f,1}^{v}) \qquad \forall v\in V,f\in F \tag{16}$$

$$T_d^v\geqslant T_f^v+s_f+T_{\mathrm{move}}^{f,d}-M(2-x_{f,d}^{v}-y_{f,2}^{v}) \qquad \forall v\in V,f\in F \tag{17}$$

$$\sum_{f\in F}x_{o,f}^{v}=0\Rightarrow T_o^v=W \qquad \forall v\in V \tag{18}$$

$$\sum_{f\in F_o}x_{i,f}^{v}=0\Rightarrow T_f^v=W \qquad \forall v\in V,f\in F \tag{19}$$

$$x_{i,j}^{v}=\{0,\ 1\} \qquad \forall v\in V,i\in F_o,j\in F_d,i\neq j \tag{20}$$

$$y_{f,1}^{v},y_{f,2}^{v}=\{0,\ 1\} \qquad \forall v\in V,f\in F \tag{21}$$

$$\begin{cases}T_f^v\geqslant 0 & \forall v\in V,f\in F\\ T_d^v\geqslant T_o^v & \forall v\in V\end{cases} \tag{22}$$

式中：F为待服务航班集合；f、f'为航班索引；o和d分别为摆渡车的逻辑起始和终到场站，两者对应于同一个物理场站；F_o为F与o点的并集，索引为i或j；F_d为F与d点的并集，索引为i或j；V为可用的摆渡车集合，索引为v或v'；K为车型的集合，索引为k；q_f为航班f的乘客数；T_{begin}^{f}、T_{end}^{f}为航班f的后置到达服务时间窗的起、止时间，其对于进港航班表示航班计划/预计到港时间，对于离港航班则表示规定[14]中要求的时间点；$[0,T_{\mathrm{begin}}^{f}-\mu_f]$表示航班$f$的前置到达服务时间窗，$\mu$为时间提前量，进港航班取值为 5 min、离港航班取值为 0；s_f为服务航班f的作业时长，包括上客、行驶至下车点和下客耗时 3 部分；δ_k^v为车辆是否为k型摆渡车，是则取 1，否则取 0；c_{fix}^{k}为k型车的固定启用成本；e_k为k型车行驶单位里程所需开支（如能源耗费等）；R_k为k型车的载客容量；l_{move}^{i}为摆渡车往返$i\in F_o$与航站楼接驳口的标准行驶里程，并且令$l_{\mathrm{move}}^{o}=0$；$l_{\mathrm{deadHD}}^{i,j}$为车辆从节点$i$结束位置到节点$j$开始位置的行驶距离；$T_{\mathrm{move}}^{i,j}$为摆渡车从节点$i$结束位置行驶至节点$j$开始位置的行驶时长；$\lambda_k$为$k$型车载客行驶时的单位里程费用增

益因子；M、W为大的正实数，前者用于模型约束线性化，后者用于给变量赋初值；$y^v_{f,1}$、$y^v_{f,2}$为0～1变量，分别表示车辆v是否在前置时间窗和后置时间窗内服务航班f，是则取1、否则取0；$x^v_{i,j}$为0～1变量，表示车辆v是否连续访问节点i和j，是则取值为1，否则取0，$\boldsymbol{x}$为其向量形式，另令$x^v_{o,d}=0$；T^v_f为车辆v到达航班f作业起点的时刻，若车辆v不服务航班f，其值为W；T^v_o、T^v_d为非负实数，分别表示摆渡车v离开和返回驻车场的时刻，若车辆v不服务任何航班，其取值为W。

目标函数(1)表示运营总成本：等号右侧第1项表示车辆启用成本，其值取决于不同类型车辆的启用数量；等号右侧第2项表示车辆总行驶成本，具体包括空驶成本和车辆往返航班与航站楼间的摆渡运输成本2部分。在既有研究中，由于只考虑一种车型，各航班的用车数量固定不变，各航班的摆渡运输成本因此成为一个固定值，故而不必列于目标函数中。考虑多车型后，航班的用车数量和类型将不再固定不变，而不同类型摆渡车的单位行驶里程耗费又不相同，这就使得各航班的摆渡运输成本存在进一步优化的空间，因此，必须列于目标函数中。考虑到载客前后车身质量的明显变化会影响车辆单位里程能耗，因此，目标函数(1)中的航班摆渡运输成本是基于载客情形来计算的，以更准确地反映车辆行驶成本。

式(2)～式(23)为模型的约束条件：式(2)表示可启用的车辆数目有限，并且没有车辆必须要被启用；式(3)表示任一一辆摆渡车至多服务同一航班一次；式(4)和式(5)分别表示车辆在场站和航班节点上的流量守恒约束；式(6)表示调派的摆渡车总装载容量不能低于被服务航班的乘客总数；式(7)～式(11)反映了机场摆渡车运输服务的双时间窗特性和要求，式(7)表示必须有车辆在前置时间窗内到达航班服务地点，式(8)表示服务某航班的车辆只能在前置和后置时间窗中选其一对该航班进行服务，式(9)表示前置服务车辆须在目标航班的前置时间窗内到达，式(10)和式(11)共同表示后置服务车辆须在目标航班的前置时间窗内到达；式(12)～式(17)表示车辆的连续行为活动(车辆的行为获得包括离开场站、服务航班和返回场站3种)在时间上必须满足可接续，式(12)和式(13)表示车辆离开场站继而服务首个航班，式(14)和式(15)表示车辆相继服务2个航班，式(16)和式(17)表示车辆服务完最后一个航班继而返回场站；式(18)和式(19)给特殊条件下的部分决策变量赋值，以降低变量数目；式(20)～式(22)给出了离散和连续决策变量的基本取值范围。

由以上数学模型可见，考虑多车型，使得所有决策变量需要增加角标v，这意味着决策变量的数量将大大增加；考虑双服务时间窗，使得问题需要另行增设两类决策变量(即$y^v_{f,1}$和$y^v_{f,2}$)，用以刻画车辆拟选择到访的时间窗。除决策变量增加和扩维以外，考虑多车型和双时间窗还将使得问题数学模型的约束数量增加，进而急剧加大问题的研究难度。

3 求解算法

众所周知，标准的带时间窗约束的车辆路径问题是一个NP-hard问题，目前尚缺乏可求解此类较大规模问题的精确算法[13]。本文的摆渡车调度问题相当于在标准带时间窗车辆路径问题基础上进一步考虑多车型、双时间窗与需求可拆分等复杂因素，这使得问题决策维度、变量数和约束数量均急剧增加，由此问题的计算复杂度将剧烈上升，可精确求解的问题规模将急速下降。因此，对于实际中或具有一定规模的问题而言，可行的解法只能是启发式算法，而在众多此类算法中，本文选择了相对较新的ALNS算法，ALNS算法在大邻域搜索基础上构造了一套自适应机制来动态选用更适宜的邻域算子来生成新解，从而提升算法效果[15]。相比于禁忌搜索、模拟退火算法等小邻域搜索算法，ALNS算法所有破坏和修复算子可显著扩大解空间的搜索范围，从而增强算法的全局寻优能力；较之遗传算法和一般邻域搜索算法盲目无差别选择邻域算子，ALNS算法根据历史表现和使用频次等信息来选择邻域算子，这种相互竞争的自适应选择机制有利于生成更好的邻域结构，从而让算法有更大概率找到更好的解；此外，这种动态算子选择机制可以限制无效或低效邻域算子的执行，从而允许ALNS算法引入更多的邻域算子(而不必担心无效搜索浪费计算资源)，此举可进一步增强搜索的多样性和算法跳出局部最优的能力。因为这些突出优点，ALNS算法已被成功应用于车辆调度、资源分派等一系列复杂组合优化问题的求解中[16]。本文结合问题的自身特点对标准ALNS算法进行适当改进，得到了求解本文摆渡车调度问题的ALNS算法，本文算法流程如图2所示。

初始解的质量对于算法效果通常具有重要影响，本文采用一种贪婪启发式规则来构造ALNS算法初始解，该规则遵循“先大后小”原则，首先，将待服务航班按最早可服务时间按升序排列，然后，依此顺序为各航班安排符合时间约束的摆渡车，且优先选用大车。为提高车辆利用效率，安排大车时

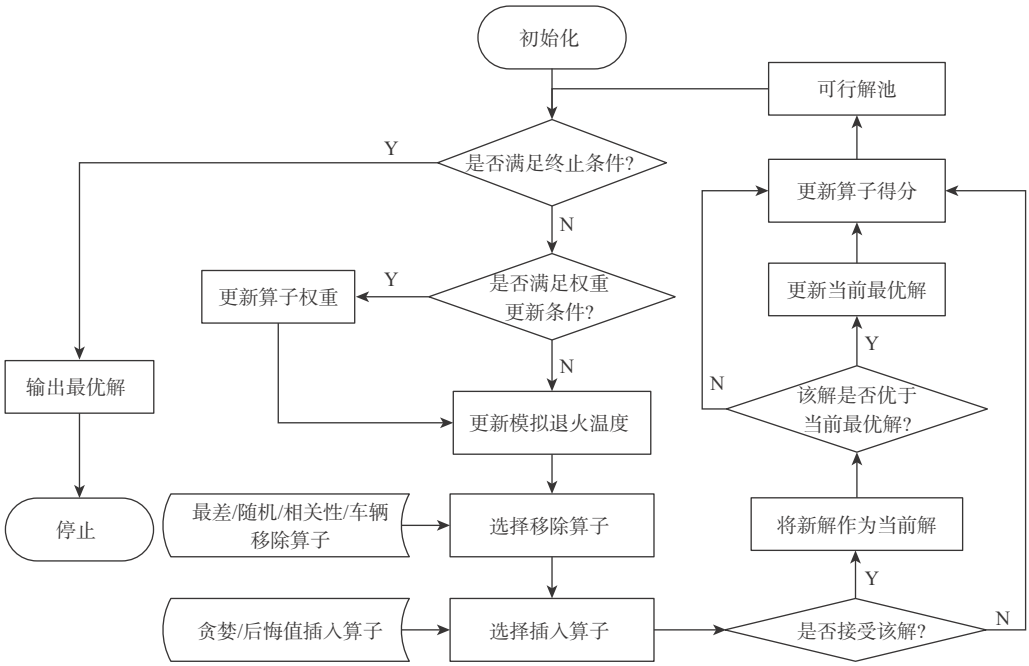


图 2　本文算法流程

Fig. 2　Flow of the proposed algorithm

还需考虑其满载率，即当大、小车均可用时，若某航班剩余乘客数量小车即可容纳，则弃大车、选小车；若场面上无现车可用，则从场站派用新车。破坏算子方面，本文在传统随机移除、最坏移除和相关性移除基础上增设了车辆移除算子，用以降低高成本车辆解的生存机率；修复算子方面，本文采用贪婪插入和后悔插入 2 种。需要指出：本文 ALNS 算法邻域操作生成的解只有满足模型所有约束条件时才可能得以保留，因此，本文算法最终得到的解必定是正确可行的。

4　实例分析

4.1　案　例

本文实例选取北京首都国际机场 T3 航站楼所在东飞行区 2018 年 9 月 1 日 6：00—12：00 之间的远机位进/离港航班数据对本文模型和算法进行测算分析，该时段共包括 7 个进港航班和 28 个出港航班，数据给出了各远机位航班的机型、停机位编号、到港时间、离港时间等信息。由于数据未提供各航班的载客数量信息，本文案例研究中将按照平均值取 80% 在区间 [70%, 90%] 内随机生成各航班的客座率。图 3 为首都国际机场东飞行区的远机位分布，其中，No. 1～No. 10 为远机位置机口。

本文案例设置大、小 2 种类型的摆渡车，其中，大型车和小型车的单车固定启用成本分别为 1 500 元和 800 元、单车载客容量分别为 120 人和 50 人、单位行驶里程成本分别为 3 元/km 和 1.6 元/km。摆渡车在场面上的平均行驶速度设为 20 km/h，不同类型车辆的载客行驶单位成本增益系数 λ 均取值为 1.2。ALNS 算法参数设置如下：初始温度设为 0.8，停止温度为 0.1，温度下降系数设为 0.995，给定温度下的算法迭代次数设为 5 次。

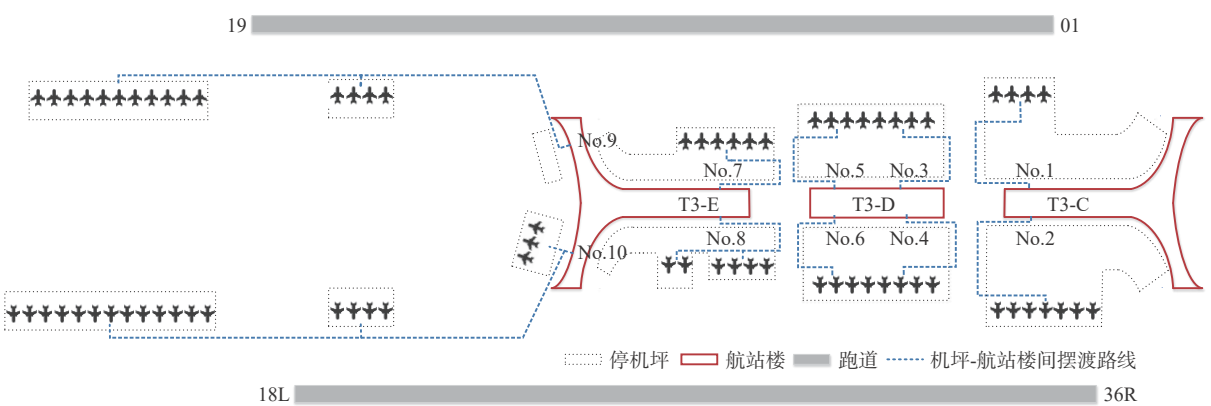
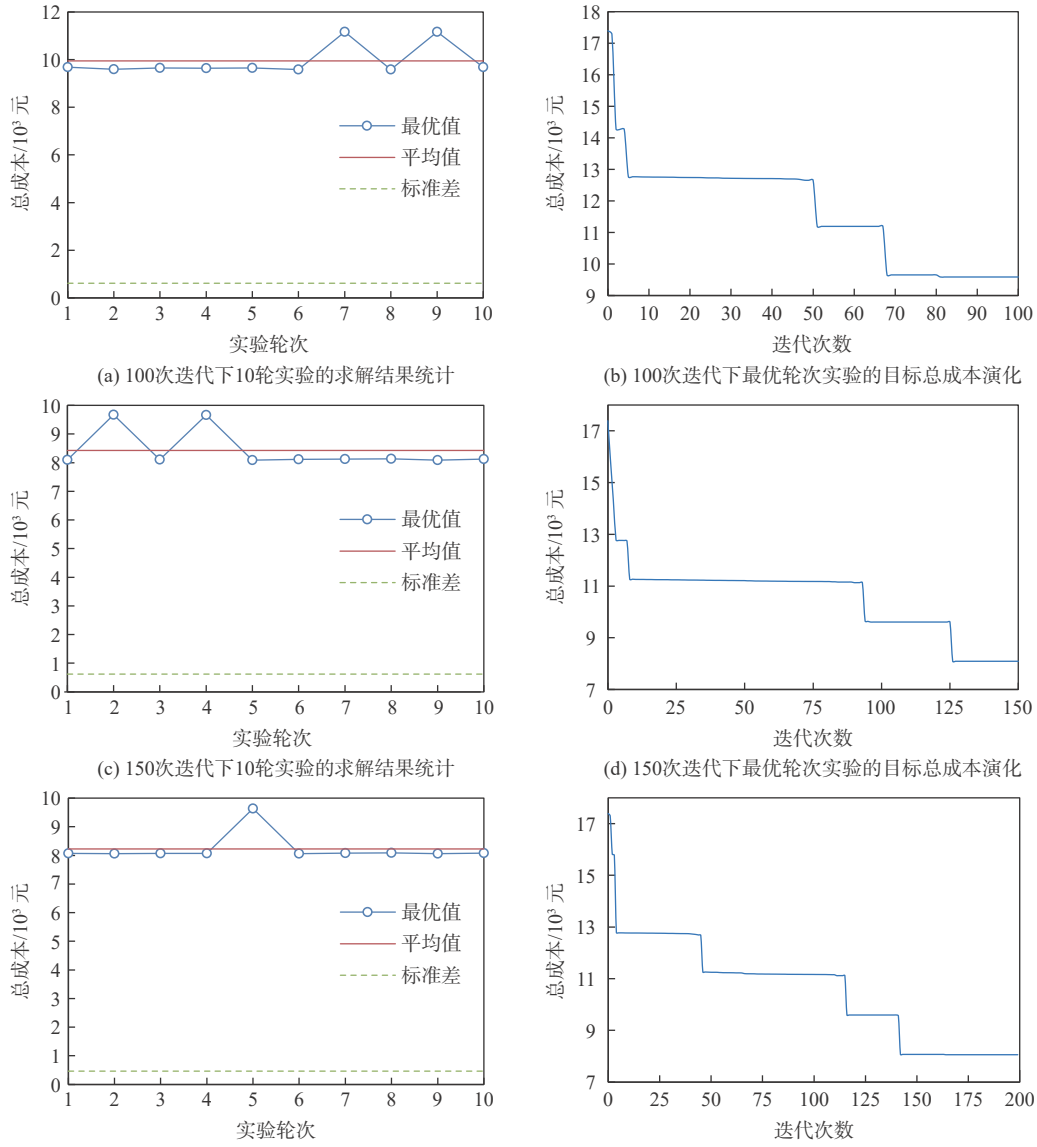


图 3　首都机场东区远机位分布与摆渡路线

Fig. 3　Layout and running routes of far stands in eastern airfield of BCIA

4.2 本文算法求解效果分析

本节针对不同终止迭代次数各自开展10轮计算实验，以检验本文ALNS算法的求解效果。图4(a)、4(c)和4(e)分别给出了终止迭代次数为100、150和200时10轮实验得到的最优解及其有关统计指标；图4(b)、4(d)和4(f)分别给出了终止迭代次数为100、150和200时最优实验对应目标函数值的下降过程。



(a) 100次迭代下10轮实验的求解结果统计　(b) 100次迭代下最优轮次实验的目标总成本演化

(c) 150次迭代下10轮实验的求解结果统计　(d) 150次迭代下最优轮次实验的目标总成本演化

(e) 200次迭代下10轮实验的求解结果统计　(f) 200次迭代下最优轮次实验的目标总成本演化

图4　不同终止迭代次数下的实验结果统计(左列)与最优目标函数值下降过程(右列)

Fig. 4　Experimental result statistics (left column) and evolution of optimal objective function values (right column) under different iterative tolerances

由图4可得，当终止迭代次数取100、150和200时，10轮实验所得最优解的标准差与平均值比值分别仅为6.2%、7.4%和5.6%，波幅较小，表明本文算法稳定性较好；最优实验所得最优解的目标函数值较初始解降幅随迭代次数增加而上升，分别达45%、53%和54%，表明本文算法对本文案例问题的优化效果较好。

4.3 调度方案对比

表1与表2分别给出了不同车辆调度方案的计算结果及多车型调度方案跟人工调度与单一车型调度方案的目标函数对比。此处人工调度用于模拟现实情况，其调度方法与本文算法的初始解生成过程类似，不同点在于这里只使用大车(因为现实情况就是如此)；单一车型调度则可视为本文算法的一种特例，相当于第2节所示模型中只有一种车型。

由表1和表2可以看出，本文多车型调度方案的各项子目标(即用车成本与行驶费用)均明显小于人工调度和单一车型调度结果，对应的总成本较之人工调度、大车型和小车型调度方案改进(即下

表 1　不同调度方案的计算结果

Table 1　Results between different scheduling schemes

调度方案	用车数量/辆		用车成本/元	行驶费用/元	总成本/元
	大车	小车			
人工调度	9		13 500	656	14 156
单一大车型调度	6		9 000	546	9 546
单一小车型调度		11	8 800	750	9 550
多车型联合调度	3	3	6 900	410	7 310

表 2　多车型调度方案与其他 3 种方案的结果对比

Table 2　Results comparison between multi-type vehicle scheduling schemes and other three schemes

对比调度策略	改进幅度/%		
	用车成本	行驶费用	总成本
人工调度	48.9	37.5	48.3
单一大车型调度	23.3	24.9	23.4
单一小车型调度	21.6	45.3	23.5

降）分别达到 48.3%、23.4% 和 23.5%，这表明人工调度方案尚存在较大的成本节约空间，采用混合车型比单一车型更更有助节约运营成本。

从车辆利用效能角度来分析导致各方案性能差异的更深层次原因。图 5 为不同调度方案的车辆满载率热力图。图中：列表示航班标号，共有



图 5　不同调度方案的满载率热力图

Fig. 5　Thermodynamic chart of vehicle full-load ratios for different scheduling strategies

35个航班；行表示分派的车型（xB表示编号为x的大型车；xS表示编号为x的小型车）；Avg表示平均满载率，具体包括车辆级平均满载率和航班级平均满载率，有关计算方式定义如下：

1）“车辆单次满载率”等于“车辆在某航班的载客人数”与“车辆容量”之比；

2）“车辆平均满载率”等于车辆服务航班的单次满载率的算术平均值，设某车共服务3个航班，对应满载率分别为80%、70%和60%，那么该车的平均满载率为70%；

3）“航班级平均满载率”等于“航班载客人数”与“服务航班的车辆总容量”之比，如某航班载客180人，服务该航班的3辆车的容量分别为100，50，50人，那么该航班的平均车辆满载率为180/(100+50+50)=90%。

由图5可见，相比于单一小车型和多车型调度方案，只使用大车的人工调度和单一大车型调度方案的淡色格子（意味着车辆单次满载率较低）数量更多，对应车辆满载率的平均值也更低；此外，较之3种经优化得到的方案，人工调度方案的空格率明显更高，这意味着其车辆复用效果可能不太理想。

表3给出了各方案的车辆总体平均满载率和车辆平均容量服务率，前者等于“航班总人数”与“所有航班的车辆总容量”之比（假设需要服务5个航班，总载客数为1 000人，某方案给各航班分派的车辆总容量分别为200，250，250，250，150人，该方案的车辆总体平均满载率则为1 000/(200+250×3+150)=95.24%），后者等于“航班总人数”与“启用车辆容量之和”之比（比如，假定需服务5个航班，总载客人数为1 000，某方案共启用了3辆车且其容量分别为100、100和50，那么该方案的车辆平均容量服务率为1 000/(100+100+50)=4），这2个指标分别反映了机场摆渡车的利用率和复用率，两者互为补充、综合反映了各调度方案的车辆利用能效。由表3可知，利用小车型的后2种方案的车辆利用率更高；对于车辆容量服务率，优化过的方案的效果比人工调度好，单一小车型方案比单一大车型的效果更好，而多车型方案中小车的合理利用使得大车的优势得到更充分的发挥，使其总体效果最好。由表3可计算出，多车型联合调度方案的车辆容量服务率（集中反映车辆容量利用能效）相较于人工、单一大车和单一小车调度方案分别提升98.3%、31.3%和22.2%，改进效果显著。综合本节结果及分析，采用混合车型可以有效避免人工或单一车型调度方案存在的“大车小用”或“小车多用”问题，从而助力企业进一步降本增效。

表3 不同调度方案的车辆利用能效对比

Table 3 Vehicle utilization efficiency comparison between different scheduling schemes

调度方式	车辆总体平均满载率/%	车辆平均容量服务率/（人·单位容量$^{-1}$）		
		大车	小车	总体
人工调度	82.05	7.17		7.17
单一大车型调度	82.05	10.83		10.83
单一小车型调度	90.78		11.64	11.64
多车型联合调度	90.78	18.33	10.33	14.22

5 结 论

本文综合考虑多车型和摆渡车运输服务的双时间窗特性，从任务点视角构建了机场摆渡车调度问题的混合整数线性规划模型，并以北京首都国际机场东区为背景开展了数值实验研究，所得主要研究结论如下：

1）研究考虑多类型摆渡车调度问题，实例结果表明多车型调度方案在车辆利用效能方面显著优于人工调度和单一车型方案，可以有效降低或避免单一车型调度方案中存在的“大车小用”或“小车多用”问题，可为相关企业降本增效提供更大潜力。

2）模型约束建模纳入更多民航行业管理规定、目标计算考虑空驶/载客车重差异，使得目标函数的费用计算更准确，本文方案的实效和可操作性将更强。

3）贪婪式人工调度仍存在较大的降本提效空间，不管采取何种车型，本文优化所得方案总能带来明显的成本节约和能效提升，而其中尤以多车型方案的效果最好。

参考文献（References）

[1] KUHN K, LOTH S. Airport service vehicle scheduling[J]. Air Traffic Control Quarterly, 2010, 18(1): 63-83.

[2] 杨文东, 陶婧婧, 贾玉平. 机坪摆渡车实时调度系统仿真[J]. 南京航空航天大学学报, 2013, 45(6): 854-858.
YANG W D, TAO J J, JIA Y P. Simulation of real-time scheduling of airport ferry bus[J]. Journal of Nanjing University of Aeronautics & Astronautics, 2013, 45(6): 854-858(in Chinese).

[3] 冯霞, 任子云. 基于遗传算法的加油车和摆渡车协同调度研究[J]. 交通运输系统工程与信息, 2016, 16(2): 155-163.
FENG X, REN Z Y. Collaborative scheduling of fuelling vehicle and ferry vehicle based on genetic algorithm[J]. Journal of Transportation Systems Engineering and Information Technology, 2016, 16(2): 155-163(in Chinese).

[4] 何丹妮. 大型机场航班过站地面服务车辆调度问题研究: 以摆渡车为例[D]. 北京: 北京交通大学, 2018.
HE D N. Research on the scheduling problem of ground service vehicles for large airport flights: taking car ferry as an example[D]. Beijing: Beijing Jiaotong University, 2018(in Chinese).

[5] HAN X, ZHAO P X, MENG Q C, et al. Optimal scheduling of air-

port ferry vehicles based on capacity network[J]. Annals of Operations Research, 2020, 295(1): 163-182.

[6] ZHAO P X, HAN X, WAN D. Evaluation of the airport ferry vehicle scheduling based on network maximum flow model[J]. Omega, 2021, 99: 102178.

[7] HAN X, ZHAO P X, KONG D X. A bi-objective optimization of airport ferry vehicle scheduling based on heuristic algorithm: a real data case study[J]. Advances in Production Engineering & Management, 2022, 17(2): 183-192.

[8] LIU Y H, WU J J, TANG J, et al. Scheduling optimisation of multi-type special vehicles in an airport[J]. Transportmetrica B: Transport Dynamics, 2022, 10(1): 954-970.

[9] ZHU S R, SUN H J, GUO X. Cooperative scheduling optimization for ground-handling vehicles by considering flights' uncertainty[J]. Computers & Industrial Engineering, 2022, 169: 108092.

[10] 冯明端, 肖雪, 周航. 机场地面保障多车型车辆联合调度模型研究[J]. 武汉理工大学学报 (交通科学与工程版), 2023, 47(1): 67-72.
FENG M D, XIAO X, ZHOU H. Research on the multi-type joint scheduling model of airport ground support vehicles[J]. Journal of Wuhan University of Technology (Transportation Science & Engineering), 2023, 47(1): 67-72(in Chinese).

[11] HAN X, ZHAO P X, KONG D X. Two-stage optimization of airport ferry service delay considering flight uncertainty[J]. European Journal of Operational Research, 2023, 307(3): 1103-1116.

[12] BAO D W, ZHOU J Y, ZHANG Z Q, et al. Mixed fleet scheduling method for airport ground service vehicles under the trend of electrification[J]. Journal of Air Transport Management, 2023, 108: 102379.

[13] BRÄYSY O, GENDREAU M. Vehicle routing problem with time windows, part I: route construction and local search algorithms[J]. Transportation Science, 2005, 39(1): 104-118.

[14] 中国民用航空局. 航班安全运行保障标准: 民航发〔2020〕4 号[S]. 北京: 中国民用航空局, 2020.
Civil Aviation Administration of China. Flight safety operation guarantee standards: CAAC〔2020〕No. 4[S]. Beijing: Civil Aviation Administration of China, 2020(in Chinese).

[15] ROPKE S, PISINGER D. An adaptive large neighborhood search heuristic for the pickup and delivery problem with time windows[J]. Transportation Science, 2006, 40(4): 455-472.

[16] WINDRAS MARA S T, NORCAHYO R, JODIAWAN P, et al. A survey of adaptive large neighborhood search algorithms and applications[J]. Computers & Operations Research, 2022, 146: 105903.

Optimization of multi-type airport ferry vehicles schedules with double service time windows characteristic

ZHANG Wenyi，TANG Yulaer，WANG Xulan，ZHOU Jing，BIAN Ke，LIU Zhishuo*

(Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport，
Beijing Jiaotong University，Beijing 100044，China)

Abstract: A mixed-integer linear programming model is developed for the airport ferry vehicle scheduling problem from the perspective of task units, taking into account both the double time windows characteristics of the airdrome ferry vehicle transport service and multi-type vehicles in order to increase the flexibility of airport ferry vehicle operation and the efficiency and practicality of the vehicle scheduling schemes. The resultant model can be viewed as a split-delivery vehicle routing problem with multi-type vehicles and double service time windows, which is a complex NP-hard problem. To solve such a challenging problem, a tailored adaptive large neighborhood search (ALNS) algorithm is proposed. Computational experiments are conducted on the eastern airfield of the Beijing Capital International Airport. The experimental results demonstrate that the proposed algorithm shows excellent stability and effectiveness. The ratio of the standard deviation to the average value of the calculation results for 10 iterations under 200 iterations is only 5.6%, and the optimal objective function value decreases by 54% before and after optimization. In addition, the multi-type vehicle schedule performs much better than both the man-made and the single-type schedules, which produces 98.3%, 31.3% and 22.2% of vehicle capacity efficiency increase and creates 48.3%, 23.4% and 23.5% of total operation cost saving in comparison to the man-made, single-big-type, and single-small-type schedules, respectively.

Keywords: air transportation; airport; ferry vehicle scheduling; multi-type vehicles; double time windows; adaptive large neighborhood search

Received：2023-09-12；**Accepted**：2023-10-13；**Published Online**：2023-11-06 11：13

URL：link.cnki.net/urlid/11.2625.V.20231103.1634.001

Foundation items：National Natural Science Foundation of China (72271018)；Research Foundation of Beijing Captial International Airport Co., Ltd. (T20L00690)

* **Corresponding author**. E-mail：zhsliu@bjtu.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0818

双端异步 DoS 攻击下基于数模联动的 ICPS 多模态综合安全控制

李亚洁，李钢，李炜*，路晨静
(兰州理工大学 电气工程与信息工程学院，兰州 730050)

摘　　要：针对一类双端异步拒绝服务 (DoS) 攻击与执行器故障共存的工业信息物理系统 (ICPS)，通过将数据驱动技术与模型机理解析方法相融合，对多模态综合安全控制与通讯间的协同设计问题进行研究。设计一种触发阈值可随系统行为动态变化的自适应离散事件触发通讯机制 (ADETCS)，并构建可同时抵御双端异步 DoS 攻击与执行器故障的 ICPS 多模态综合安全控制架构；针对不同能量等级的 DoS 攻击，采用“分而治之”的思想，借助长短期记忆 (LSTM) 网络与弹性控制方法，提出基于数模联动的主-被动协同混合容侵策略。基于 Lyapunov 稳定性理论进行观测器与控制器的推证，进而采用 K-Means++聚类算法及模糊融合方法，在线对不同模态下的控制器进行加权融合，实现不同控制模态间的软切换；通过四容水箱实例，验证了双端异步 DoS 攻击下基于数模联动的多模态综合安全控制方法的正确性。实验结果表明：数模联动的方法增强了 ICPS 抵御双端异步 DoS 攻击的能力，多模态综合安全控制器的设计实现了控制模态与 ADETCS 间的双向自适应协同控制。

关　键　词：工业信息物理系统；双端异步 DoS 攻击；数模联动；自适应离散事件触发通讯机制；多模态综合安全控制
中图分类号：TP277
文献标志码：A　　**文章编号**：1001-5965(2025)10-3354-14

工业信息物理系统 (industrial cyber-physical systems, ICPS) 是通过自动化、计算机、通信等技术进行集成，进而管控工业过程的新型智能系统，且已在多个领域得到了广泛应用[1-2]。随着 ICPS 中信息系统与物理系统间的融合程度日益加深，ICPS 将面临信息层恶意网络攻击与物理层元部件失效故障的双重威胁[3]。仅就网络安全而言，近年来已报道多起 ICPS 遭受恶意网络攻击的事件，如伊朗核设备因 Stuxnet 病毒攻击而长期停运[4]、中东遭遇火焰病毒攻击导致行业隐私信息泄露[5] 等。此外，随着 ICPS 的信息化与智能化水平的进一步提升，以其为中心的系统规模也在不断扩大，所需感知和传输的数据也随之增加，使得有限网络带宽资源与海量数据信息传输间的矛盾日益凸显。

ICPS 是物理层与信息层的深度融合产物，而开放的网络外部环境使得 ICPS 在通信过程中极易受到恶意攻击，当两者的实时交互受阻时，无疑会导致 ICPS 性能受限，甚至整体失稳。ICPS 遭受的网络攻击主要有拒绝服务 (denial of service, DoS) 攻击[6-7]、虚假数据注入攻击[8-9] 和数据重放攻击[10] 等。其

收稿日期：2024-11-15；**录用日期**：2025-04-11；**网络出版时间**：2025-05-14 13:03
网络出版地址：link.cnki.net/urlid/11.2625.V.20250514.1025.001
基金项目：国家自然科学基金 (62163022)；甘肃省科技重大专项计划 (23ZDGA007)；甘肃省产业支撑计划项目 (2024CYZC-18)；甘肃省联合科研基金项目 (24JRRA827)；兰州市科技计划项目 (2024-2-5)
* **通信作者**. E-mail：liwei@lut.edu.cn

引用格式：*李亚洁，李钢，李炜，等. 双端异步 DoS 攻击下基于数模联动的 ICPS 多模态综合安全控制 [J]. 北京航空航天大学学报，2025，51（10）：3354-3367. LI Y J，LI G，LI W，et al. ICPS multi-modal integrated security control based on data-model linkage under dual-end asynchronous DoS attacks [J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51（10）：3354-3367 （in Chinese）.*

中, DoS 攻击尽管属于非隐蔽性攻击, 但由于其策略简单且易于实施, 在实际工业系统中也较为常见。然而, 目前针对 DoS 攻击的防御研究, 仅就控制层面而言, 大多仅考虑单端通讯网络遭受攻击的情形, 未考虑工业系统中更为常见的双端异步 DoS 攻击, 且对攻击的防御主要采用弹性控制, 即基于机理模型的鲁棒性来应对, 而通过采用数据驱动技术与机理模型技术相结合的策略来抵御网络攻击的情形则鲜有考虑。

随着 ICPS 的规模性与复杂度日趋增大, 系统的物理元件(如执行器、传感器)发生故障的概率也随之增加。加之受到有限网络带宽资源与海量数据信息传输间的矛盾驱动, 学者们在离散事件触发通讯机制 (discrete event-triggered communication scheme, DETCS) 下展开了对 ICPS 容错控制的研究, 能够兼顾网络服务质量和系统运行性能, 并取得了一定的研究成果[11-12]。尽管 DETCS 通过预先设定的触发阈值实现了对传输数据的筛选, 节约了通讯资源, 但固定的触发阈值难以适应于动态变化的系统运行状态。若恒定的事件触发参数过大, 系统稳态运行时, 会出现“事件触发条件”长时间不被满足, 经网络传输数据的间隔变大的情况。进而在时延理论框架下设计控制器时, 将致使求解空间变小, 导致系统性能变差; 反之, 若恒定的事件触发参数过小, 则于节约通讯资源无补。因此, DETCS 很难实现系统性能与资源占用间的动态优化平衡。为克服上述不足, 学者们提出了自适应离散事件触发通信机制 (adaptive discrete event-triggered communication scheme, ADETCS), 可根据系统运行工况自适应调整触发阈值, 以节约更多的网络通讯资源[13-14]。尽管通讯机制具备了自适应触发的功能, 遗憾的是在设计求解系统控制器时, 依旧使用固定触发参数来实现控制器与触发权矩阵的协同设计, 并未在 ADETCS 下使控制器也具有真正的自适应能力, 更未考虑与 ADETCS 相对应的多模态综合安全控制问题。

受上述问题驱动, 本文针对同时存在双端异步 DoS 攻击和执行器故障的 ICPS, 通过数据驱动与模型机理的融合联动, 研究了 ICPS 多模态综合安全控制与通讯的协同设计问题, 主要创新点包括:

1) 为进一步实现对通讯资源的节约, 本文设计了一种基于反正切函数可双向自适应变化的 ADETCS, 用触发参数的阈值函数来取代 DETCS 中固定不变的触发阈值, 可以根据系统的运行状况自适应变化, 进一步实现系统控制性能与通讯资源节约之间的优化折中平衡。

2) 考虑双端异步 DoS 攻击, 采用主被动协同的数模联动混合容侵策略, 对传感端小能量 DoS 攻击以被动容侵机理模型的自身鲁棒性能加以应对, 对传感端大能量 DoS 攻击及执行端 DoS 攻击, 采用长短期记忆(long short-term memory, LSTM)神经网络预测模型, 实时补偿因 DoS 攻击而丢失的数据包, 将基于数据驱动的主动容侵与机理解析的被动容侵相结合, 提升系统应对双端异步 DoS 攻击的能力。

3) 为提升控制器可随通讯机制中触发参数自适应变化的能力, 采用 K-Means++聚类算法, 基于历史数据, 离线对系统在不同工况下的控制模态进行合理划分, 而后采用模糊融合方法, 在线得到不同模态下控制器输出并进行加权融合, 从而以信息融合的方式实现不同模态之间的软切换, 旨在改善系统控制策略对通讯机制的适应性。

1　系统描述

1.1　综合安全控制架构的构建

为与工业现场更加贴近, 本文考虑双侧网络均遭受 DoS 攻击并存在执行器故障情形的 ICPS。为缓解有限带宽与海量数据传输间的矛盾, 加之提升控制器可随系统行为自适应变化能力的期冀, 在 ADETCS 下构建了可同时抵御 DoS 攻击与执行器故障的 ICPS 多模态综合安全控制架构, 如图 1 所示。

在图 1 中, 先做如下约定: 若 DoS1 攻击造成的系统时延接近或超过系统的最大允许时延则定义为大能量 DoS 攻击, 反之, 则定义为小能量 DoS 攻击。$\lambda=1$和$\eta=1$分别表示系统传感端和执行端检测到大能量 DoS 攻击, $\lambda=0$和$\eta=0$分别表示系统传感端和执行端未检测到大能量 DoS 攻击; 为更加真实的还原与反映现场应用中发生 DoS 攻击的实际情况, 图 1 中的 DoS 攻击具有异步特性, 其中, 传感端 DoS1 和执行端 DoS2 均为随机发生, 且两者间并无时间关联。ICPS 多模态综合安全控制架构主要包括:

1) 智能传感单元。首先, 引入 ADETCS, 通过动态变化的触发参数阈值函数实现对传输数据的筛选, 可降低数据在通信网络中的传输率, 从而节约网络资源。

2) 智能控制单元。满足 ADETCS 的触发条件, 经传感侧网络传输的数据将保存在第 1 个数据缓存器中。攻击检测器 1 检测传感端是否遭受大能量 DoS 攻击, 若检测到大能量 DoS 攻击, 则激活数据重构补偿器 1, 利用 LSTM 网络模型进行实时数据重构与补偿, 并将其送至观测器; 否则, 直接将接

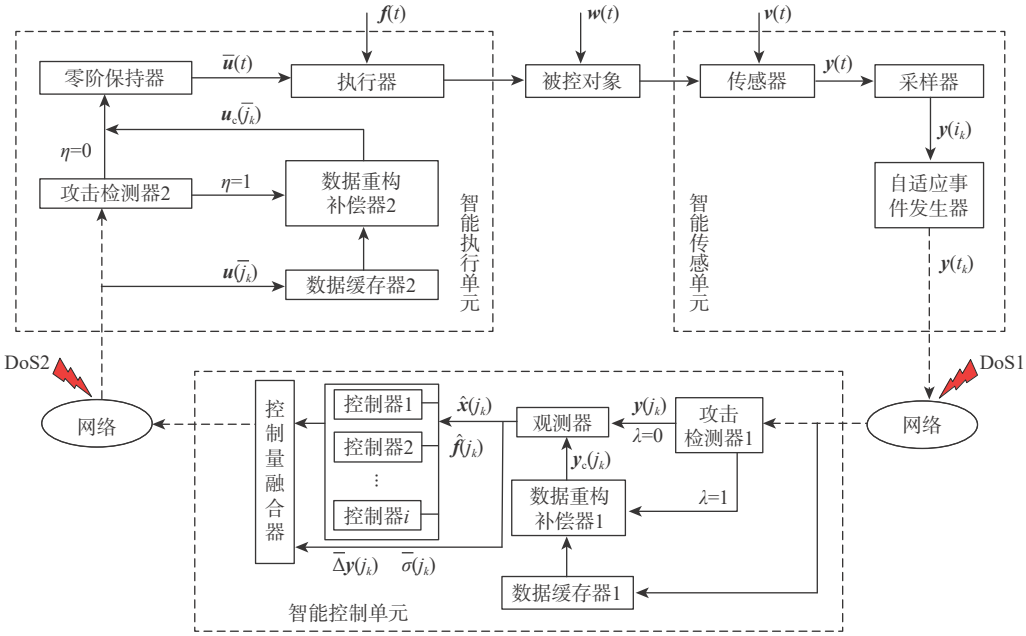


图 1 基于 ADETCS 的双端异步 DoS 攻击下 ICPS 多模态综合安全控制架构

Fig. 1 ICPS multi-modal integrated security control architecture under double-ended asynchronous DoS attacks based on ADETCS

收的数据送入观测器以实现对系统状态、执行器故障的准确估计。接着，将估计结果送入控制器，用以控制量的计算，同时将触发参数、触发误差送入控制量融合器，进而利用提前设计好的模糊融合规则，对控制量进行加权融合以实现不同控制模态间的软切换，最后，将融合后的控制量传输至执行单元。

3）智能执行单元。智能执行单元的数据缓存器 2 用于存储成功传输至执行单元的历史数据，攻击检测器 2 用以检测执行端网络中是否遭受 DoS 攻击，若检测到 DoS 攻击，则触发数据重构补偿器 2，基于 LSTM 网络模型对因 DoS 攻击而丢失的数据包进行实时补偿，再将其传送至执行器。若没有攻击则直接将接收的数据传输至执行器。

注 1 在此架构中采用了“分而治之”的策略，在智能控制单元和智能执行单元中都独立集成了攻击检测器和数据重构补偿器，不仅能够独立识别传感端与执行端的 DoS 攻击，并且能对 DoS 攻击造成的数据丢失进行实时补偿，以期实现对双端异步 DoS 攻击的精准容侵。

注 2 由于对传感端小能量 DoS 攻击的被动容侵是基于机理模型 (观测器、控制器) 自身的鲁棒性而实现的。因此，为进一步提升系统安全性，只对传感端 DoS 攻击区分大、小能量等级，并采用主-被动协同数模联动的混合容侵策略加以应对，而对执行端 DoS 攻击则不区分能量等级，完全采用基于数据驱动的主动容侵策略来应对。

1.2 被控对象的描述

针对同时存在有限能量 DoS 攻击和执行器时变故障的 ICPS 描述如下：

$$\begin{cases}\dot{\boldsymbol{x}}(t)=\boldsymbol{A}\boldsymbol{x}(t)+\boldsymbol{B}\boldsymbol{u}(t)+\boldsymbol{E}_f\boldsymbol{f}(t)+\boldsymbol{B}_w\boldsymbol{w}(t)\\ \boldsymbol{y}(i_kh)=\boldsymbol{C}\boldsymbol{x}(i_kh)+\boldsymbol{E}_v\boldsymbol{v}(i_kh)\end{cases}\tag{1}$$

式中：$\boldsymbol{x}(t)\in\mathbf{R}^n$为系统状态向量；$\boldsymbol{u}(t)\in\mathbf{R}^{n_u}$为控制输入向量；$\boldsymbol{w}(t)\in\mathbf{R}^{n_w}$为系统扰动；$\boldsymbol{y}(i_k)\in\mathbf{R}^m$为系统采样输出向量；$\boldsymbol{v}(i_k)\in\mathbf{R}^{n_v}$为测量噪声；$\boldsymbol{f}(t)\in\mathbf{R}^{n_f}$为时变且连续的故障向量，且假定故障的导数范数有界，因此，存在常数f_1且$\|\dot{\boldsymbol{f}}(t)\|\leqslant f_1$；$\boldsymbol{A}$、$\boldsymbol{B}$、$\boldsymbol{C}$、$\boldsymbol{E}_f$、$\boldsymbol{B}_w$、$\boldsymbol{E}_v$为已知的矩阵。

1.3 ADETCS 的设计

图 1 引入 ADETCS，设计如下自适应事件触发条件，用以对系统数据进行筛选：

$$\boldsymbol{e}^{\mathrm{T}}(i_kh)\boldsymbol{\Phi}\boldsymbol{e}(i_kh)\geqslant\sigma(t_kh)\boldsymbol{y}^{\mathrm{T}}(i_kh)\boldsymbol{\Phi}\boldsymbol{y}(i_kh)\tag{2}$$

式中：$\boldsymbol{\Phi}$为事件触发权矩阵；$\sigma(t_kh)$为待设计的触发参数阈值函数；$\boldsymbol{y}(t_kh)$为上一时刻满足事件触发条件的系统输出值，$i_kh=t_kh+lh,l\in\mathbf{N}$，$h$为采样周期。输出误差为$\boldsymbol{e}(i_kh)=\boldsymbol{y}(i_kh)-\boldsymbol{y}(t_kh)$，$\boldsymbol{y}(i_kh)$为当前时刻系统输出的采样值。

本文设计的自适应触发参数阈值函数依照以下规则实现动态自适应变化：

$$\begin{cases}\sigma(t_kh)=\min\{\max[\sigma_m,\vartheta\sigma_m],\sigma_{\max}\}\\ \vartheta=\mu-\dfrac{2\alpha}{\pi}\mathrm{atan}(|\Delta|)\\ \Delta=\beta(\|\boldsymbol{y}(t_kh)\|-\|\boldsymbol{y}(t_{k-1}h)\|)\end{cases}\tag{3}$$

式中：$\sigma_m>0$，σ_m和$\sigma_{\max}$为$\sigma(t_kh)$的上界和下界；$\mu>0$，$\alpha>0$，$\beta>0$为给定的常数。

注 3 本文设计了 ADETCS 以取代 DETCS，且所设计的阈值函数是基于反正切函数，结合μ、α、β以调节阈值函数的取值范围。当系统受到扰动、故

障及攻击时, $\|\boldsymbol{y}(t_k h)\|$与$\|\boldsymbol{y}(t_{k-1}h)\|$之间的误差就会变大, $\vartheta\sigma_m$将变小, 此时$\sigma(t_k h)<\sigma(t_{k-1}h)$, 需采用较小的触发参数$\sigma(t_k h)$以增加数据传输频率, 在及时的控制作用下使系统更快的回到稳定状态; 反之, 则采用更大的触发参数$\sigma(t_k h)$, 以减少数据的传输频率, 从而节约网络资源。触发参数$\sigma(t_k h)$的双向自适应变化, 使得系统数据传输频率能够根据系统行为动态调整, 实现了通讯资源节约与系统控制性能的双向兼顾。

1.4　双端异步 DoS 攻击下数据传输时序时延分析

在 ADETCS 下, 考虑双端异步 DoS 攻击对系统数据传输的影响, 系统中数据传输的时序分析如图 2 所示。

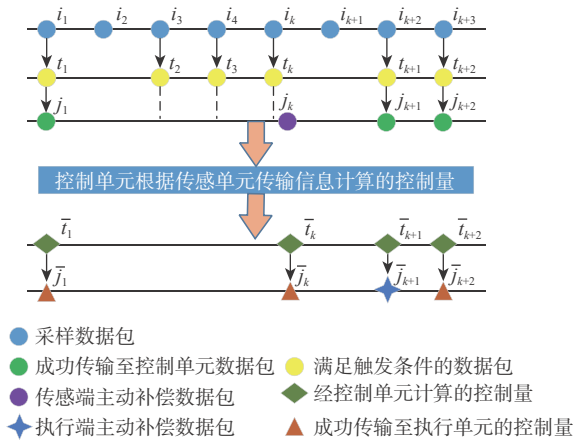


图 2　双端异步 DoS 攻击下数据传输时序图

Fig. 2　Timing diagram of data transmission under double-ended asynchronous DoS attack

图 2 假设采样器的采样周期为h, 得到采样数据序列$\{i_1,i_2,\cdots,i_k,\cdots\}$, 再经过自适应事件发生器筛选后满足触发条件的数据序列为$\{t_1,t_2,\cdots,t_k,\cdots\}$, 经传感端无线网络成功传输至控制单元或补偿后的数据序列为$\{j_1,j_2,\cdots,j_k,\cdots\}$。自适应事件发生器的输入和输出分别为$\boldsymbol{y}(i_k)\in\mathbf{R}^m$,$\boldsymbol{y}(t_k)\in\mathbf{R}^m$。图 1 中, 在有限能量 DoS 攻击下经过补偿后的输出为$\boldsymbol{y}_c(j_k)\in\mathbf{R}^m$。若传感端未检测到大能量 DoS 攻击, 此时传输至观测器的数据序列为$\boldsymbol{y}(j_k)\in\mathbf{R}^m$。$\{\bar{t}_1,\bar{t}_2,\cdots,\bar{t}_k,\cdots\}$为控制单元根据传感单元传输信息计算的控制量序列, $\{\bar{j}_1,\bar{j}_2,\cdots\bar{j}_k,\cdots\}$为经控制端无线网络成功传输至执行单元或补偿后的数据序列。在有限能量 DoS 攻击下经过补偿后的控制量为$\boldsymbol{u}_c(\bar{j}_k)\in\mathbf{R}^m$。若执行端未检测到 DoS 攻击, 此时传输至执行器的数据序列为$\boldsymbol{u}(\bar{j}_k)\in\mathbf{R}^m$。

图 1 中被控对象是连续的, 然而, 经过等周期采样和 ADETCS 的筛选, 传至控制单元进行估计的量变为离散数据, 因此, 该系统属于采样数据系统, 可将其转化为对应的时滞系统。通过借鉴文献 [15] 中的时滞系统分析方法, 将非均匀传输属性转化为系统时滞。因此, 在 ADETCS 下, 考虑到传感端时延会受到有限能量 DoS 攻击的影响, 定义其时延函数为

$$\tau(t)=t-j_k\quad t\in[j_k,j_{k+1})\tag{4}$$

式中: 时延函数满足$h\leqslant\tau(t)\leqslant h_m$, h_m为系统最大允许时延, 其取值决定于 ADETCS 的触发间隔和小能量 DoS 攻击发生的频率及持续时间。

注 4　图 1 中, 由于执行端对 DoS 攻击的应对是完全基于数据驱动技术的主动容侵, 因此, 由 DoS 攻击而丢失的数据包是被实时补偿的, 在对双端异步 DoS 攻击下的系统时延进行分析时, 不考虑执行端 DoS 攻击对系统时延的影响。

2　双端异步 DoS 攻击的主-被动协同数模联动混合容侵策略

2.1　基于 DoS 攻击能量分级的检测机制设计

根据数据传输时序图, 在j_{k-1}和j_k时刻传输至控制单元的系统量测输出量分别是$\boldsymbol{y}(j_{k-1})$和$\boldsymbol{y}(j_k)$, T为连续 2 次满足 ADETCS 触发条件的传输间隔, 即, $T=j_k-j_{k-1}$。若h_1和h_2分别是观测器和控制器在既定性能条件下的最大允许时延, 则系统的最大允许时延$h_{\mathrm{M}}=\min(h_1,h_2)$。为进一步提升系统安全性, 设计了安全系数$\theta$, $\theta\in(0,1)$, 并将θh_{M}作为区分大、小能量 DoS 攻击的阈值。若$T<\theta h_{\mathrm{M}}$, 判定系统没有遭受 DoS 攻击或小能量 DoS 攻击; 若$T>\theta h_{\mathrm{M}}$, 则判定系统存在大能量 DoS 攻击。

2.2　数据驱动技术的引入

随着 ICPS 信息化、智能化程度的不断提升, 使得以 ICPS 为核心的体系规模越来越庞大, 而 ICPS 运行中产生的海量数据则为数据驱动技术的引入提供了数据基础。数据驱动技术依赖大量数据进行训练和优化, 且可根据系统历史数据对因 DoS 攻击而丢失的数据进行有效的预测和补偿, 从而为 ICPS 抵御大能量 DoS 攻击提供了可行之径。其中, LSTM 网络模型的门控机制允许模型决定何时更新记忆、何时读取记忆、何时忘记非关键信息, 这使得 LSTM 网络可以处理具有复杂时序依赖性的数据, 且 LSTM 网络模型能够处理实际工程中时序数据可能包含间隔和延迟等非连续性数据, 从而在预测含有间隔或延迟的时序数据时仍能保持较高的准确性。这也使得 LSTM 网络预测模型可以根据遭受小能量 DoS 攻击后的系统历史数据, 准确的重构补偿因大能量 DoS 攻击而丢失的具有时序特征的系统输出数据。

2.3　长短期记忆神经网络

LSTM 网络是 Hochreiter 等在 1997 年设计的[16],

在原有循环神经网络(recurrent neural networks, RNN)隐藏层神经节点的基础上改进了网络的结构,增加存储历史信息的记忆单元结构。其中,遗忘门决定网络结构中传入信息的保留或丢弃;输入门决定传入信息在更新当前LSTM网络单元状态中的权重;输出门决定当前LSTM网络的信息传递比重。因此,这种结构决定了LSTM网络能准确预测具有时序依赖性的时间序列。

2.4 基于数模联动下主-被动协同的传感端DoS攻击混合容侵策略设计

若传感端遭受的小能量DoS攻击所引发的时延处于$(h,\theta h_{\mathrm{M}})$之内,则该攻击对系统性能不会造成大的影响,可采用基于机理模型自身的鲁棒性应对;当传感端的传输时延间隔$T>\theta h_{\mathrm{M}}$时,表示系统已遭受大能量DoS攻击,此时,DoS攻击造成的时延超过了系统允许的最大时延,被动容侵策略已经失效,则需要采用基于数据驱动的主动容侵方法来实现对因大能量DoS攻击而丢失的数据进行实时重构补偿。当传感端攻击检测器1检测到大能量DoS攻击时,数据重构补偿器1中的LSTM预测模型将根据数据缓存器1中成功传输的历史数据,对因大能量DoS攻击而丢失的数据进行实时补偿。需要说明的是用于训练LSTM的系统数据是在小能量DoS攻击和执行器故障工况下采集的,尽可能多的考虑了不同的工况。具体步骤如下:

步骤1 将小能量DoS攻击和执行器故障下的系统输出数据用作LSTM的训练集和测试集,用于建立和训练精准的预测模型,并集成到数据重构补偿器1中;

步骤2 攻击检测器1若检测到大能量DoS攻击,则触发重构补偿器1,对因大能量DoS攻击而丢失的数据进行实时在线重构,反之,直接输出当前系统输出值。传入观测器的测量输出数据为

$$\bar{\boldsymbol{y}}(j_k)=\begin{cases}\boldsymbol{y}(j_k) & \lambda=0\\ \boldsymbol{y}_{\mathrm{c}}(j_k) & \lambda=1\end{cases} \tag{5}$$

式中:$\bar{\boldsymbol{y}}(j_k)$同时用于传入控制量融合器中触发参数$\bar{\sigma}(j_k)$及事件触发误差$\Delta\bar{\boldsymbol{y}}(j_k)$的计算,保证传入控制量融合器的$\bar{\sigma}(j_k)$和$\Delta\bar{\boldsymbol{y}}(j_k)$具有相同的时序特征。

2.5 基于数据驱动的执行端DoS攻击主动容侵策略设计

当攻击检测器2检测到DoS攻击时,数据重构补偿器2利用数据缓存器2中成功传输的历史数据,以实时补偿的方式抵御执行端网络中的DoS攻击,具体步骤如下:

步骤1 使用未受DoS攻击的控制量作为LSTM的训练集与测试集,用于建立和训练LSTM网络预测模型,并集成到数据重构补偿器2中;

步骤2 攻击检测器2若检测到DoS攻击,则触发数据重构补偿器2对因DoS攻击而丢失的数据进行实时补偿,反之,直接输出当前控制量:

$$\bar{\boldsymbol{u}}(\bar{j}_k)=\begin{cases}\boldsymbol{u}(\bar{j}_k) & \eta=0\\ \boldsymbol{u}_{\mathrm{c}}(\bar{j}_k) & \eta=1\end{cases} \tag{6}$$

3 双端异步DoS攻击下鲁棒观测器的设计方法

由式(4)的时延函数及式(5)可得

$$\bar{\boldsymbol{y}}(t)=\begin{cases}\boldsymbol{y}(t-\tau(t)) & \lambda=0\\ \boldsymbol{y}_{\mathrm{LSTM}}(t-\tau(t)) & \lambda=1\end{cases} \tag{7}$$

设计如下鲁棒观测器,以实现对系统状态和故障的准确估计:

$$\begin{cases}\dot{\hat{\boldsymbol{x}}}(t)=\boldsymbol{A}\hat{\boldsymbol{x}}(t)+\boldsymbol{B}\boldsymbol{u}(t)+\boldsymbol{E}_f\hat{\boldsymbol{f}}(t)-\boldsymbol{L}(\hat{\boldsymbol{y}}(t)-\bar{\boldsymbol{y}}(t))\\ \hat{\boldsymbol{y}}(t)=\boldsymbol{C}\hat{\boldsymbol{x}}(t-\tau(t))\\ \dot{\hat{\boldsymbol{f}}}(t)=-\boldsymbol{F}(\hat{\boldsymbol{y}}(t)-\bar{\boldsymbol{y}}(t))\end{cases} \tag{8}$$

式中:$\hat{\boldsymbol{x}}(t)$、$\hat{\boldsymbol{y}}(t)$和$\hat{\boldsymbol{f}}(t)$分别为状态估计值、观测器输出和故障估计值;$\boldsymbol{L}$和$\boldsymbol{F}$分别为待设计的状态和故障增益矩阵。

定义:$\boldsymbol{e}_x(t)=\hat{\boldsymbol{x}}(t)-\boldsymbol{x}(t)$, $\boldsymbol{e}_y(t)=\hat{\boldsymbol{y}}(t)-\bar{\boldsymbol{y}}(t)$, $\boldsymbol{e}_f(t)=\hat{\boldsymbol{f}}(t)-\boldsymbol{f}(t)$得到如下误差系统:

$$\begin{cases}\dot{\boldsymbol{e}}_x(t)=\boldsymbol{A}\boldsymbol{e}_x(t)+\boldsymbol{E}_f\boldsymbol{e}_f(t)-\boldsymbol{L}\boldsymbol{C}\boldsymbol{e}_x(t-\tau(t))+\\ \qquad\boldsymbol{L}\boldsymbol{E}_v\boldsymbol{v}(t-\tau(t))-\boldsymbol{B}_w\boldsymbol{w}(t)\\ \dot{\boldsymbol{e}}_f(t)=-\boldsymbol{F}\boldsymbol{C}\boldsymbol{e}_x(t-\tau(t))+\boldsymbol{F}\boldsymbol{E}_v\boldsymbol{v}(t-\tau(t))-\dot{\boldsymbol{f}}(t)\end{cases} \tag{9}$$

为便于分析,将$\boldsymbol{e}_x(t)$和$\boldsymbol{e}_f(t)$增广为一个整体。定义$\bar{\boldsymbol{e}}(t)=\left[\boldsymbol{e}_x^{\mathrm{T}}(t)\quad\boldsymbol{e}_f^{\mathrm{T}}(t)\right]^{\mathrm{T}}$,结合式(9)可得增广误差系统:

$$\begin{aligned}\dot{\bar{\boldsymbol{e}}}(t)=&\bar{\boldsymbol{A}}\bar{\boldsymbol{e}}(t)-\bar{\boldsymbol{L}}\bar{\boldsymbol{C}}\bar{\boldsymbol{e}}(t-\tau(t))-\bar{\boldsymbol{B}}_w\bar{\boldsymbol{w}}(t)+\\ &\bar{\boldsymbol{L}}\boldsymbol{E}_v\boldsymbol{v}(t-\tau(t))\end{aligned} \tag{10}$$

式中:

$$\begin{cases}\bar{\boldsymbol{A}}=\begin{bmatrix}\boldsymbol{A} & \boldsymbol{E}_f\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix},\bar{\boldsymbol{e}}(t)=\begin{bmatrix}\boldsymbol{e}_x(t)\\ \boldsymbol{e}_f(t)\end{bmatrix},\bar{\boldsymbol{C}}=\begin{bmatrix}\boldsymbol{C} & \boldsymbol{0}\end{bmatrix}\\ \bar{\boldsymbol{L}}=\begin{bmatrix}\boldsymbol{L}\\ \boldsymbol{F}\end{bmatrix},\bar{\boldsymbol{B}}_w=\begin{bmatrix}\boldsymbol{B}_w & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}\end{bmatrix},\bar{\boldsymbol{w}}(t)=\begin{bmatrix}\boldsymbol{w}(t)\\ \dot{\boldsymbol{f}}(t)\end{bmatrix}\end{cases}$$

定理1 给定正数$\bar{h}$、γ_1、$n_i(i=1,2,3)$,若存在正定对称矩阵$\boldsymbol{P}$及适维矩阵$\boldsymbol{M}$、$\boldsymbol{Y}$满足线性矩阵不等式:

$$\begin{bmatrix}\boldsymbol{S}_{11} & \boldsymbol{S}_{12} & \boldsymbol{S}_{13} & \boldsymbol{S}_{14} & \boldsymbol{S}_{15} & \boldsymbol{S}_{16} & \boldsymbol{0}\\ & \boldsymbol{S}_{22} & \boldsymbol{S}_{23} & \boldsymbol{S}_{24} & \boldsymbol{S}_{25} & \boldsymbol{S}_{26} & \boldsymbol{S}_{27}\\ & & \boldsymbol{S}_{33} & \boldsymbol{S}_{34} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\\ & & & \boldsymbol{S}_{44} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\\ & & & & \boldsymbol{S}_{55} & \boldsymbol{S}_{56} & \boldsymbol{S}_{57}\\ & & & & & \boldsymbol{S}_{66} & \boldsymbol{0}\\ & & & & & & \boldsymbol{S}_{77}\end{bmatrix}<\boldsymbol{0} \tag{11}$$

$$\begin{bmatrix} \boldsymbol{S}'_{11} & \boldsymbol{S}'_{12} & \boldsymbol{S}'_{13} & \boldsymbol{S}'_{14} & \boldsymbol{S}'_{15} & \boldsymbol{S}'_{16} & \boldsymbol{0} & \boldsymbol{M}^{\mathrm{T}} \\ & \boldsymbol{S}'_{22} & \boldsymbol{S}'_{23} & \boldsymbol{S}'_{24} & \boldsymbol{0} & \boldsymbol{S}'_{26} & \boldsymbol{S}'_{27} & \boldsymbol{M}^{\mathrm{T}} \\ & & \boldsymbol{S}'_{33} & \boldsymbol{S}'_{34} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{M}^{\mathrm{T}} \\ & & & \boldsymbol{S}'_{44} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{M}^{\mathrm{T}} \\ & & & & \boldsymbol{S}'_{55} & \boldsymbol{S}'_{56} & \boldsymbol{S}'_{57} & \boldsymbol{0} \\ & & & & & \boldsymbol{S}'_{66} & \boldsymbol{0} & \boldsymbol{0} \\ & & & & & & \boldsymbol{S}'_{77} & \boldsymbol{0} \\ & & & & & & & \boldsymbol{S}'_{88} \end{bmatrix} < \boldsymbol{0} \tag{12}$$

式中: $\boldsymbol{S}_{11}=(1+\bar{h}n_2)(\boldsymbol{P}\overline{\boldsymbol{A}}+\overline{\boldsymbol{A}}^{\mathrm{T}}\boldsymbol{P})+\bar{h}^2n_3\overline{\boldsymbol{A}}^{\mathrm{T}}\boldsymbol{P}\overline{\boldsymbol{A}}-(\pi^2/4)n_3\boldsymbol{P}+\boldsymbol{I}+3\boldsymbol{M}+3\boldsymbol{M}^{\mathrm{T}}$; $\boldsymbol{S}_{12}=-(1+\bar{h}n_2)\boldsymbol{Y}\overline{\boldsymbol{C}}-\boldsymbol{M}^{\mathrm{T}}+[n_2+(\pi^2/4)n_3]\cdot\boldsymbol{P}-\bar{h}^2n_3\overline{\boldsymbol{A}}^{\mathrm{T}}\boldsymbol{Y}\overline{\boldsymbol{C}}-\bar{h}n_2\overline{\boldsymbol{A}}^{\mathrm{T}}\boldsymbol{P}+3\boldsymbol{M}$; $\boldsymbol{S}_{13}=3\boldsymbol{M}-7\boldsymbol{M}^{\mathrm{T}}$; $\boldsymbol{S}_{14}=3\boldsymbol{M}+14\boldsymbol{M}^{\mathrm{T}}$; $\boldsymbol{S}_{15}=\boldsymbol{Y}\boldsymbol{E}_v+\bar{h}n_2\boldsymbol{Y}\boldsymbol{E}_v+\bar{h}^2n_3\overline{\boldsymbol{A}}^{\mathrm{T}}\boldsymbol{Y}\boldsymbol{E}_v$; $\boldsymbol{S}_{16}=-\boldsymbol{P}\overline{\boldsymbol{B}}_w-\bar{h}^2n_3\overline{\boldsymbol{A}}^{\mathrm{T}}\boldsymbol{P}\overline{\boldsymbol{B}}_w-\bar{h}n_2\boldsymbol{P}\overline{\boldsymbol{B}}_{\mathrm{w}}$; $\boldsymbol{S}_{22}=-[2n_2+(\pi^2/4)n_3]\cdot\boldsymbol{P}-\boldsymbol{M}-\boldsymbol{M}^{\mathrm{T}}+\bar{h}n_2(\boldsymbol{Y}\overline{\boldsymbol{C}}+\overline{\boldsymbol{C}}^{\mathrm{T}}\boldsymbol{Y}^{\mathrm{T}})$; $\boldsymbol{S}_{23}=-\boldsymbol{M}-7\boldsymbol{M}^{\mathrm{T}}$; $\boldsymbol{S}_{24}=-\boldsymbol{M}+14\boldsymbol{M}^{\mathrm{T}}$; $\boldsymbol{S}_{25}=-\bar{h}n_2\boldsymbol{Y}\boldsymbol{E}_{\mathrm{v}}$; $\boldsymbol{S}_{26}=\bar{h}n_2\boldsymbol{P}\overline{\boldsymbol{B}}_{\mathrm{W}}+\bar{h}^2n_3\bar{\boldsymbol{C}}^{\mathrm{T}}\boldsymbol{Y}^{\mathrm{T}}\bar{\boldsymbol{B}}_w$; $\boldsymbol{S}_{27}=\bar{h}^2n_3\bar{\boldsymbol{C}}^{\mathrm{T}}\boldsymbol{Y}^{\mathrm{T}}$; $\boldsymbol{S}_{33}=-7\boldsymbol{M}-7\boldsymbol{M}^{\mathrm{T}}$; $\boldsymbol{S}_{34}=-7\boldsymbol{M}+14\boldsymbol{M}^{\mathrm{T}}$; $\boldsymbol{S}_{44}=14\boldsymbol{M}+14\boldsymbol{M}^{\mathrm{T}}$; $\boldsymbol{S}_{55}=-\gamma_1^2\boldsymbol{I}$; $\boldsymbol{S}_{56}=\bar{h}n_2\boldsymbol{E}_v^{\mathrm{T}}\boldsymbol{Y}^{\mathrm{T}}\bar{\boldsymbol{B}}_w$; $\boldsymbol{S}_{57}=-\bar{h}^2n_3\boldsymbol{E}_v{}^{\mathrm{T}}\boldsymbol{Y}^{\mathrm{T}}$; $\boldsymbol{S}_{66}=\bar{h}^2n_3\bar{\boldsymbol{B}}_w^{\mathrm{T}}\boldsymbol{P}\bar{\boldsymbol{B}}_w-\gamma_1^2\boldsymbol{I}$; $\boldsymbol{S}_{77}=-\bar{h}^2n_3\boldsymbol{P}$; $\boldsymbol{S}'_{11}=\boldsymbol{P}\bar{\boldsymbol{A}}+\bar{\boldsymbol{A}}^{\mathrm{T}}\boldsymbol{P}-(\pi^2/4)n_3\boldsymbol{P}+\boldsymbol{I}+3\boldsymbol{M}^{\mathrm{T}}+3\boldsymbol{M}+(\bar{h}^2n_3+\bar{h}n_1)\bar{\boldsymbol{A}}^{\mathrm{T}}\boldsymbol{P}\bar{\boldsymbol{A}}$; $\boldsymbol{S}'_{12}=[(\pi^2/4)n_3+n_2]\boldsymbol{P}-\boldsymbol{Y}\bar{\boldsymbol{C}}+3\boldsymbol{M}-(\bar{h}^2n_3+\bar{h}n_1)\bar{\boldsymbol{A}}^{\mathrm{T}}\boldsymbol{Y}\bar{\boldsymbol{C}}-\boldsymbol{M}^{\mathrm{T}}$; $\boldsymbol{S}'_{13}=3\boldsymbol{M}-7\boldsymbol{M}^{\mathrm{T}}$; $\boldsymbol{S}'_{14}=3\boldsymbol{M}+14\boldsymbol{M}^{\mathrm{T}}$; $\boldsymbol{S}'_{15}=\boldsymbol{Y}\boldsymbol{E}_v+(\bar{h}^2n_3+\bar{h}n_1)\bar{\boldsymbol{A}}^{\mathrm{T}}\boldsymbol{Y}\boldsymbol{E}_v$; $\boldsymbol{S}'_{16}=-(\bar{h}^2n_3+\bar{h}n_1)\bar{\boldsymbol{A}}^{\mathrm{T}}\boldsymbol{P}\bar{\boldsymbol{B}}_w-\boldsymbol{P}\bar{\boldsymbol{B}}_{\mathrm{W}}$; $\boldsymbol{S}'_{22}=-[2n_2+(\pi^2/4)n_3]\boldsymbol{P}-\boldsymbol{M}-\boldsymbol{M}^{\mathrm{T}}$; $\boldsymbol{S}'_{23}=-\boldsymbol{M}-7\boldsymbol{M}^{\mathrm{T}}$; $\boldsymbol{S}'_{24}=-\boldsymbol{M}+14\boldsymbol{M}^{\mathrm{T}}$; $\boldsymbol{S}'_{26}=\bar{h}n_1\bar{\boldsymbol{C}}^{\mathrm{T}}\boldsymbol{Y}^{\mathrm{T}}\bar{\boldsymbol{B}}_w$; $\boldsymbol{S}'_{27}=\bar{h}^2n_3\bar{\boldsymbol{C}}^{\mathrm{T}}\boldsymbol{Y}^{\mathrm{T}}$; $\boldsymbol{S}'_{33}=-7\boldsymbol{M}-7\boldsymbol{M}^{\mathrm{T}}$; $\boldsymbol{S}'_{34}=-7\boldsymbol{M}+14\boldsymbol{M}^{\mathrm{T}}$; $\boldsymbol{S}'_{44}=14\boldsymbol{M}+14\boldsymbol{M}^{\mathrm{T}}$; $\boldsymbol{S}'_{55}=-\gamma_1^2\boldsymbol{I}$; $\boldsymbol{S}'_{56}=-(\bar{h}^2n_3+\bar{h}n_1)\boldsymbol{E}_v^{\mathrm{T}}\boldsymbol{Y}^{\mathrm{T}}\bar{\boldsymbol{B}}_w$; $\boldsymbol{S}'_{57}=-\bar{h}^2n_3\boldsymbol{E}_v^{\mathrm{T}}\boldsymbol{Y}^{\mathrm{T}}-\bar{h}n_1\boldsymbol{E}_v^{\mathrm{T}}\boldsymbol{Y}^{\mathrm{T}}$; $\boldsymbol{S}'_{66}=(\bar{h}^2n_3+\bar{h}n_1)\bar{\boldsymbol{B}}_w^{\mathrm{T}}\boldsymbol{P}\bar{\boldsymbol{B}}_w-\gamma_1^2\boldsymbol{I}$; $\boldsymbol{S}'_{77}=-(\bar{h}^2n_3+\bar{h}n_1)\boldsymbol{P}$; $\boldsymbol{S}'_{88}=-\boldsymbol{P}^{-1}/9\bar{h}n_1$。

则存在状态增益矩阵$\boldsymbol{L}$和故障增益矩阵$\boldsymbol{F}$, 使得增广误差系统式 (10) 在无扰动时渐近稳定, 在有扰动时满足性能指标: $\|\bar{\boldsymbol{e}}(t)\|_2^2\leqslant\gamma_1^2(\|\bar{\boldsymbol{w}}(t)\|_2^2+\bar{h}\|\boldsymbol{v}(j_k)\|_2^2)$, $\boldsymbol{L}$、$\boldsymbol{F}$可通过$\bar{\boldsymbol{L}}=\boldsymbol{P}^{-1}\boldsymbol{Y}=\begin{bmatrix}\boldsymbol{L}^{\mathrm{T}} & \boldsymbol{F}^{\mathrm{T}}\end{bmatrix}^{\mathrm{T}}$求取。

证明 构造如下 Lyapunov-Krasovskii 泛函:

$$\boldsymbol{V}(t)=\boldsymbol{V}_1(t)+\boldsymbol{V}_2(t)+\boldsymbol{V}_3(t)+\boldsymbol{V}_4(t)+\boldsymbol{V}_5(t) \tag{13}$$

式中:

$$\boldsymbol{V}_1(t)=\bar{\boldsymbol{e}}^{\mathrm{T}}(t)\boldsymbol{P}\bar{\boldsymbol{e}}(t)$$

$$\boldsymbol{V}_2(t)=(\bar{h}-\tau(t))\boldsymbol{\varphi}_1^{\mathrm{T}}\boldsymbol{S}\boldsymbol{\varphi}_1(t)+\int_{t-\tau(t)}^{t}\bar{\boldsymbol{e}}^{\mathrm{T}}(s)\boldsymbol{S}\bar{\boldsymbol{e}}(s)\mathrm{d}s$$

$$\boldsymbol{V}_3(t)=\bar{h}^2\int_{t-\tau(t)}^{t}\dot{\bar{\boldsymbol{e}}}^{\mathrm{T}}(s)\boldsymbol{Q}\dot{\bar{\boldsymbol{e}}}(s)\mathrm{d}s$$

$$\boldsymbol{V}_4(t)=-(\pi^2/4)\int_{t-\tau(t)}^{t}(\bar{\boldsymbol{e}}(s)-\bar{\boldsymbol{e}}(t_{kh}))^{\mathrm{T}}\boldsymbol{Q}(\bar{\boldsymbol{e}}(s)-\bar{\boldsymbol{e}}(t_{kh}))\mathrm{d}s$$

$$\boldsymbol{V}_5(t)=\int_{-\tau(t)}^{0}\int_{t+\theta}^{t}\dot{\bar{\boldsymbol{e}}}^{\mathrm{T}}(s)\boldsymbol{R}\dot{\bar{\boldsymbol{e}}}(s)\mathrm{d}s\mathrm{d}\theta$$

其中: $\boldsymbol{P}$、$\boldsymbol{S}$、$\boldsymbol{Q}$、$\boldsymbol{R}$为正定对称矩阵; $\boldsymbol{\varphi}_1(t)=\bar{\boldsymbol{e}}(t)-\bar{\boldsymbol{e}}(t_{kh})$。

限于篇幅, 定理 1 后续证明过程在此不再详述, 证明中采用了新型 Bessel-Legendre 不等式[17] 和线性凸组合引理[18]。

当给定观测器性能指标γ_1时, 由式 (14) 可以求得观测器最大允许时延h_1:

$$\begin{cases}\max\limits_{\boldsymbol{P},\boldsymbol{M},\boldsymbol{Y},\gamma_1} h_1 \\ \text{s.t.}\,\boldsymbol{P}>\boldsymbol{0}\end{cases} \tag{14}$$

证毕

4　双端异步 DoS 攻击下控制器模态划分与控制量融合方法

结合 ICPS 在工业现场的应用需求, 科学合理的控制模态个数确定与不同控制模态划分, 不仅能提升 ICPS 的控制性能, 还能有效减少控制器结构的冗余度与复杂度。虽然通过专家先验知识可以大致确定系统模态数目, 但仍有知识不足、缺乏理论依据和定量分析的局限。此外, 多模态硬切换下的消抖技术虽已层出不穷, 但其柔性仍有待提升。鉴于此, 结合数据聚类与信息融合技术, 本文设计了一种基于系统离线模态划分与多控制量在线融合的柔性切换机制。

4.1　基于 K-Means++算法的控制器模态划分

若既要改善控制器对通讯机制的适应性又要提高系统控制性能, 本文将引入基于数据驱动技术的 K-Means++算法。选取可以反映系统运行状态的触发误差和触发参数 2 个指标进行聚类, 从而确定系统在不同运行工况下所对应的控制模态的个数。

Arthur 等[19] 提出 K-means++聚类算法旨在解决传统 K-means 算法的初始中心点选择问题。该算法通过引入概率分布, 智能选择初始聚类中心, 来提高聚类结果的质量和稳定性。基于 K-Means++算法聚类的具体步骤如下:

步骤 1　从数据集X中随机选择一中心点C_i;

步骤 2　对于数据集X中的每个x_i, 计算与其相距最近的中心C_j的欧氏距离$D_i(x)$, 并添加到S_j:

$$\begin{cases}D_i(x)=\sqrt{(a_i-a_j)^2+(b_i-b_j)^2} \\ S_j=\sum\limits_{x_i\in X}D_i(x)^2\end{cases} \tag{15}$$

式中: (a_i,b_i)和(a_j,b_j)分别为x_i和C_j的坐标。

步骤 3　依据概率P_{o}选取一个新的中心C_{j+1}, 其中: $P_{\mathrm{o}}=D_i(x)^2/S_j$。

步骤 4　重复上述步骤, 直到达到迭代次数或最优中心点不再变化, 输出符合条件的聚类结果。

4.2　多模态综合安全控制器的设计

在基于 K-Means++算法的控制模态划分完成

后，当有$\bar{n}$个控制器模态时，其中，第l个模态的综合安全控制策略为

$$\boldsymbol{u}_i(t)=\boldsymbol{K}_l\hat{\boldsymbol{x}}(t-\tau(t))-\boldsymbol{B}^{+}\boldsymbol{E}_f\hat{\boldsymbol{f}}(t-\tau(t)) \tag{16}$$

式中：$t\in[j_k,j_{k+1}]$，控制器增益矩阵为$\boldsymbol{K}_l\in\boldsymbol{R}^{n_u\times n}$，故障调节矩阵$\boldsymbol{B}^{+}\in\boldsymbol{R}^{n_u\times n}$满足$(\boldsymbol{I}-\boldsymbol{B}\boldsymbol{B}^{+})\boldsymbol{E}_f=\boldsymbol{0}$，将式(16)代入式(1)可得闭环ICPS模型：

$$\begin{aligned}\dot{\boldsymbol{x}}(t)=&\boldsymbol{A}\boldsymbol{x}(t)+\boldsymbol{B}\boldsymbol{K}_i\boldsymbol{x}(t-\tau(t))+\boldsymbol{B}\boldsymbol{K}_l\boldsymbol{e}_x(t-\tau(t))-\\&\boldsymbol{E}_f\boldsymbol{e}_f(t-\tau(t))+\boldsymbol{B}_w\boldsymbol{w}(t)+\tau(t)\boldsymbol{E}_f\dot{\boldsymbol{f}}(t)\end{aligned} \tag{17}$$

定理2 给定正数m_i、$\bar{n}_i(i=1,2,3)$、$\bar{h}$、γ_2、σ_l，若存在对称正定矩阵$\tilde{\boldsymbol{P}}$及适维矩阵$\tilde{\boldsymbol{M}}$、$\boldsymbol{K}_l$、$\boldsymbol{Q}_1$、$\boldsymbol{Q}_3$、$\boldsymbol{Q}_5$、$\boldsymbol{Q}_7$满足矩阵不等式式(18)～式(20)：

$$\begin{bmatrix}\boldsymbol{Z}_{11}&\boldsymbol{Z}_{12}\\&\boldsymbol{Z}_{22}\end{bmatrix}<\boldsymbol{0} \tag{18}$$

$$\begin{bmatrix}\boldsymbol{Z}'_{11}&\boldsymbol{Z}'_{12}\\&\boldsymbol{Z}'_{22}\end{bmatrix}<\boldsymbol{0} \tag{19}$$

式中：

$$\boldsymbol{Z}_{11}=\begin{bmatrix}\boldsymbol{\Pi}_{11}&\boldsymbol{\Pi}_{12}&\boldsymbol{\Pi}_{13}&\boldsymbol{\Pi}_{14}&\boldsymbol{\Pi}_{15}&\boldsymbol{\Pi}_{16}\\&\boldsymbol{\Pi}_{22}&\boldsymbol{\Pi}_{23}&\boldsymbol{\Pi}_{24}&\boldsymbol{\Pi}_{25}&\boldsymbol{\Pi}_{26}\\&&\boldsymbol{\Pi}_{33}&\boldsymbol{\Pi}_{34}&\boldsymbol{0}&\boldsymbol{0}\\&&&\boldsymbol{\Pi}_{44}&\boldsymbol{0}&\boldsymbol{0}\\&&&&\boldsymbol{\Pi}_{55}&\boldsymbol{\Pi}_{56}\\&&&&&\boldsymbol{\Pi}_{66}\end{bmatrix}$$

$$\boldsymbol{Z}_{12}=\begin{bmatrix}\boldsymbol{\Pi}_{17}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}\\\boldsymbol{\Pi}_{27}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{\Pi}_{210}\\\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}\\\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}\\\boldsymbol{\Pi}_{57}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{\Pi}_{510}\\\boldsymbol{\Pi}_{67}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}\end{bmatrix}$$

$$\boldsymbol{Z}_{22}=\begin{bmatrix}\boldsymbol{\Pi}_{77}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}\\\boldsymbol{0}&\boldsymbol{\Pi}_{88}&\boldsymbol{0}&\boldsymbol{0}\\\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{\Pi}_{99}&\boldsymbol{0}\\\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{\Pi}_{1010}\end{bmatrix}$$

$$\boldsymbol{Z}'_{11}=\begin{bmatrix}\boldsymbol{\Pi}'_{11}&\boldsymbol{\Pi}'_{12}&\boldsymbol{\Pi}'_{13}&\boldsymbol{\Pi}'_{14}&\boldsymbol{\Pi}'_{15}&\boldsymbol{\Pi}'_{16}\\&\boldsymbol{\Pi}'_{22}&\boldsymbol{\Pi}'_{23}&\boldsymbol{\Pi}'_{24}&\boldsymbol{0}&\boldsymbol{\Pi}'_{26}\\&&\boldsymbol{\Pi}'_{33}&\boldsymbol{\Pi}'_{34}&\boldsymbol{0}&\boldsymbol{0}\\&&&\boldsymbol{\Pi}'_{44}&\boldsymbol{0}&\boldsymbol{0}\\&&&&\boldsymbol{\Pi}'_{55}&\boldsymbol{\Pi}'_{56}\\&&&&&\boldsymbol{\Pi}'_{66}\end{bmatrix}$$

$$\boldsymbol{Z}'_{12}=\begin{bmatrix}\boldsymbol{\Pi}'_{17}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{M}^{\mathrm{T}}\\\boldsymbol{\Pi}'_{27}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{\Pi}'_{210}&\boldsymbol{M}^{\mathrm{T}}\\\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{M}^{\mathrm{T}}\\\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{M}^{\mathrm{T}}\\\boldsymbol{\Pi}'_{57}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{\Pi}'_{510}&\boldsymbol{0}\\\boldsymbol{\Pi}'_{67}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}\end{bmatrix}$$

$$\boldsymbol{Z}'_{22}=\begin{bmatrix}\boldsymbol{\Pi}'_{77}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}\\&\boldsymbol{\Pi}'_{88}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}\\&&\boldsymbol{\Pi}'_{99}&\boldsymbol{0}&\boldsymbol{0}\\&&&\boldsymbol{\Pi}'_{1010}&\boldsymbol{0}\\&&&&\boldsymbol{\Pi}'_{1111}\end{bmatrix}$$

其中：$\boldsymbol{\Pi}_{11}=(1+\bar{h}n_1)(\tilde{\boldsymbol{P}}\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}}\tilde{\boldsymbol{P}})+3\tilde{\boldsymbol{M}}+3\tilde{\boldsymbol{M}}^{\mathrm{T}}+\boldsymbol{I}+[\bar{h}^2n_3+(\bar{h}^2/4)m_3+(\bar{h}^2/4)m_4]\boldsymbol{A}^{\mathrm{T}}\tilde{\boldsymbol{P}}\boldsymbol{A}+(\bar{h}^2/4)m_2\tilde{\boldsymbol{P}}-(\pi^2/4)n_3\tilde{\boldsymbol{P}}$；$\boldsymbol{\Pi}_{12}=(1+\bar{h}n_1)\boldsymbol{K}_{1l}+(n_1+[\pi^2/4)n_3\tilde{\boldsymbol{P}}-(\bar{h}^2/4)m_2]\tilde{\boldsymbol{P}}-\bar{h}n_1\boldsymbol{A}^{\mathrm{T}}\tilde{\boldsymbol{P}}+3\tilde{\boldsymbol{M}}-\tilde{\boldsymbol{M}}^{\mathrm{T}}+\bar{h}^2n_3\boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}_{1l}+[(\bar{h}^2/4)m_3+(\bar{h}^2/4)m_4]\cdot\boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}_{1l}$；$\boldsymbol{\Pi}_{13}=3\tilde{\boldsymbol{M}}-7\tilde{\boldsymbol{M}}^{\mathrm{T}}$；$\boldsymbol{\Pi}_{14}=3\tilde{\boldsymbol{M}}+14\tilde{\boldsymbol{M}}^{\mathrm{T}}$；$\boldsymbol{\Pi}_{15}=(1+\bar{h}n_1)\boldsymbol{K}_{1l}+\bar{h}^2n_3m_3\boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}_{1l}+[(\bar{h}^2/4)m_3+(\bar{h}^2/4)m_4]m_3\boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}_{1l}$；$\boldsymbol{\Pi}_{16}=-\tilde{\boldsymbol{P}}\boldsymbol{E}_f-\bar{h}n_1\tilde{\boldsymbol{P}}\boldsymbol{E}_f-[\bar{h}^2n_3+(\bar{h}^2/4)m_3+(\bar{h}^2/4)m_4]\boldsymbol{A}^{\mathrm{T}}\cdot\tilde{\boldsymbol{P}}\boldsymbol{E}_f$；$\boldsymbol{\Pi}_{17}=(1+\bar{h}n_1)\tilde{\boldsymbol{P}}\boldsymbol{B}_w+\left(\bar{h}^2n_3+(\bar{h}^2/4)m_3\right)\boldsymbol{A}^{\mathrm{T}}\tilde{\boldsymbol{P}}\boldsymbol{B}_w+(\bar{h}^2/4)m_4\boldsymbol{A}^{\mathrm{T}}\tilde{\boldsymbol{P}}\boldsymbol{B}_w$；$\boldsymbol{\Pi}_{22}=-\bar{h}n_1(\boldsymbol{K}_{1l}+\boldsymbol{K}_{1l}^{\mathrm{T}})-\tilde{\boldsymbol{M}}+[-2n_1+(\bar{h}^2/4)m_2-(\Pi^2/4)n_3]\tilde{\boldsymbol{P}}-\tilde{\boldsymbol{M}}^{\mathrm{T}}$；$\boldsymbol{\Pi}_{23}=-\tilde{\boldsymbol{M}}-7\tilde{\boldsymbol{M}}^{\mathrm{T}}$；$\boldsymbol{\Pi}_{24}=-\tilde{\boldsymbol{M}}+14\tilde{\boldsymbol{M}}^{\mathrm{T}}$；$\boldsymbol{\Pi}_{25}=-\bar{h}n_1\boldsymbol{K}_{1l}$；$\boldsymbol{\Pi}_{26}=\bar{h}n_1\tilde{\boldsymbol{P}}\boldsymbol{E}_f-\left[\bar{h}^2n_3+(\bar{h}^2/4)m_3+(\bar{h}^2/4)m_4\right]\boldsymbol{K}_{1l}^{\mathrm{T}}\boldsymbol{E}_f$；$\boldsymbol{\Pi}_{27}=\left[\bar{h}^2n_3+(\bar{h}^2/4)m_3+(\bar{h}^2/4)m_4\right]\boldsymbol{K}_{1l}^{\mathrm{T}}\boldsymbol{B}_w+\bar{h}n_1\tilde{\boldsymbol{P}}\boldsymbol{B}_w$；$\boldsymbol{\Pi}_{210}=\left[\bar{h}^2n_3+(\bar{h}^2/4)m_3+(\bar{h}^2/4)m_4\right]\boldsymbol{K}_{1l}^{\mathrm{T}}$；$\boldsymbol{\Pi}_{33}=-7\tilde{\boldsymbol{M}}-7\tilde{\boldsymbol{M}}^{\mathrm{T}}$；$\boldsymbol{\Pi}_{34}=-7\tilde{\boldsymbol{M}}+14\tilde{\boldsymbol{M}}^{\mathrm{T}}$；$\boldsymbol{\Pi}_{44}=14\tilde{\boldsymbol{M}}+14\tilde{\boldsymbol{M}}^{\mathrm{T}}$；$\boldsymbol{\Pi}_{55}=-\gamma_2^2\boldsymbol{I}$；$\boldsymbol{\Pi}_{56}=-(\bar{h}^2/4)m_4\boldsymbol{K}_{1l}^{\mathrm{T}}\boldsymbol{E}_f-\left(\bar{h}^2n_3+(\bar{h}^2/4)m_3\right)\boldsymbol{K}_{1l}^{T}\boldsymbol{E}_f$；$\boldsymbol{\Pi}_{57}=(\bar{h}^2/4)m_3\boldsymbol{K}_{1l}^{\mathrm{T}}\boldsymbol{B}_w+(\bar{h}^2n_3+(\bar{h}^2/4)m_4)\boldsymbol{K}_{1l}^{\mathrm{T}}\boldsymbol{B}_w$；$\boldsymbol{\Pi}_{510}=\bar{h}^2n_3\boldsymbol{K}_{1l}^{\mathrm{T}}+\left[(\bar{h}^2/4)m_3+(\bar{h}^2/4)m_4\right]\cdot\boldsymbol{K}_{1l}^{\mathrm{T}}$；$\boldsymbol{\Pi}_{66}=\bar{h}^2n_3\boldsymbol{E}_f^{\mathrm{T}}\tilde{\boldsymbol{P}}\boldsymbol{E}_f+\left[(\bar{h}^2/4)m_3+(\bar{h}^2/4)m_4\right]\boldsymbol{E}_f^{\mathrm{T}}\tilde{\boldsymbol{P}}\boldsymbol{E}_f-\gamma_2^2\boldsymbol{I}$；$\boldsymbol{\Pi}_{67}=-\left[\bar{h}^2n_3+(\bar{h}^2/4)m_3+(\bar{h}^2/4)m_4\right]\boldsymbol{E}_f\tilde{\boldsymbol{P}}\boldsymbol{B}_w$；$\boldsymbol{\Pi}_{77}=\left[\bar{h}^2n_3+(\bar{h}^2/4)m_3+(\bar{h}^2/4)m_4\right]\boldsymbol{B}_w^{\mathrm{T}}\tilde{\boldsymbol{P}}\boldsymbol{B}_w-\gamma_2^2\boldsymbol{I}$；$\boldsymbol{\Pi}_{88}=\sigma\boldsymbol{\Phi}$；$\boldsymbol{\Pi}_{99}=-\boldsymbol{\Phi}$；$\boldsymbol{\Pi}_{1010}=-(\bar{h}^2/4)m_3\tilde{\boldsymbol{P}}-\left[\bar{h}^2n_3+(\bar{h}^2/4)m_4\right]\tilde{\boldsymbol{P}}$；$\boldsymbol{\Pi}'_{11}=\tilde{\boldsymbol{P}}\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}}\tilde{\boldsymbol{P}}+\boldsymbol{I}+\left[(\bar{h}^2/4)m_2-(\Pi^2/4)n_3+\bar{h}m_1\right]\tilde{\boldsymbol{P}}+3\tilde{\boldsymbol{M}}^{\mathrm{T}}+\left[\bar{h}^2n_3+(\bar{h}^2/4)m_3+(\bar{h}^2/4)m_4+\bar{h}n_2\right]\boldsymbol{A}^{\mathrm{T}}\tilde{\boldsymbol{P}}\boldsymbol{A}+3\tilde{\boldsymbol{M}}$；$\boldsymbol{\Pi}'_{12}=\boldsymbol{K}_{1l}+\left[n_1-(\bar{h}^2/4)m_2+(\Pi^2/4)n_3\right]\tilde{\boldsymbol{P}}+3\tilde{\boldsymbol{M}}+[\bar{h}^2n_3+(\bar{h}^2/4)m_3+(\bar{h}^2/4)m_4+\bar{h}n_2]A^{\mathrm{T}}\boldsymbol{K}_{1l}-\tilde{\boldsymbol{M}}^{\mathrm{T}}$；$\boldsymbol{\Pi}'_{13}=3\tilde{\boldsymbol{M}}-7\tilde{\boldsymbol{M}}^{\mathrm{T}}$；$\boldsymbol{\Pi}'_{14}=3\tilde{\boldsymbol{M}}+14\tilde{\boldsymbol{M}}^{\mathrm{T}}$；$\boldsymbol{\Pi}'_{15}=\boldsymbol{K}_{1l}+\left[\bar{h}^2n_3+(\bar{h}^2/4)m_3+(\bar{h}^2/4)m_4+\bar{h}n_2\right]\boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}_{1l}$；$\boldsymbol{\Pi}'_{16}=-\left[(\bar{h}^2/4)m_3+(\bar{h}^2/4)m_4+\bar{h}n_2\right]\cdot\boldsymbol{A}^{\mathrm{T}}\tilde{\boldsymbol{P}}\boldsymbol{E}_f-\bar{h}^2n_3\boldsymbol{A}^{\mathrm{T}}\tilde{\boldsymbol{P}}\boldsymbol{E}_f-\tilde{\boldsymbol{P}}\boldsymbol{E}_f$；$\boldsymbol{\Pi}'_{17}=-\bar{h}^2n_3\boldsymbol{A}^{\mathrm{T}}\tilde{\boldsymbol{P}}\boldsymbol{B}_w-[(\bar{h}^2/4)m_4+\bar{h}n_2+(\bar{h}^2/4)m_3]\boldsymbol{A}^{\mathrm{T}}\tilde{\boldsymbol{P}}\boldsymbol{B}_w+\tilde{\boldsymbol{P}}\boldsymbol{B}_w$；$\boldsymbol{\Pi}'_{22}=[-2n_1+(\bar{h}^2/4)m_2-(\Pi^2/4)n_3]\tilde{\boldsymbol{P}}-\tilde{\boldsymbol{M}}-\tilde{\boldsymbol{M}}^{\mathrm{T}}$；$\boldsymbol{\Pi}'_{23}=-\tilde{\boldsymbol{M}}-7\tilde{\boldsymbol{M}}^{\mathrm{T}}$，$\boldsymbol{\Pi}'_{24}=-\tilde{\boldsymbol{M}}+14\tilde{\boldsymbol{M}}^{\mathrm{T}}$；$\boldsymbol{\Pi}'_{26}=-\bar{h}^2n_3\boldsymbol{K}_{1l}^{\mathrm{T}}\boldsymbol{E}_f-[\bar{h}^2n_3+(\bar{h}^2/4)m_3+(\bar{h}^2/4)m_4+\bar{h}n_2]\boldsymbol{K}_{1l}^{\mathrm{T}}\boldsymbol{E}_f$；$\boldsymbol{\Pi}'_{27}=[\bar{h}^2n_3+(\bar{h}^2/4)m_3+(\bar{h}^2/4)m_4+\bar{h}n_2]\boldsymbol{K}_{1l}^{\mathrm{T}}\boldsymbol{B}_w$；$\boldsymbol{\Pi}'_{210}=[\bar{h}^2n_3+(\bar{h}^2/4)m_3+(\bar{h}^2/4)m_4+\bar{h}n_2]\boldsymbol{K}_{1l}^{\mathrm{T}}$；$\boldsymbol{\Pi}'_{210}=\left[\bar{h}^2n_3+(\bar{h}^2/4)m_3+(\bar{h}^2/4)m_4+\bar{h}n_2\right]\boldsymbol{K}_{1l}^{\mathrm{T}}$；$\boldsymbol{\Pi}'_{33}=-7\tilde{\boldsymbol{M}}-7\tilde{\boldsymbol{M}}^{\mathrm{T}}$；$\boldsymbol{\Pi}'_{34}=-7\tilde{\boldsymbol{M}}+14\tilde{\boldsymbol{M}}^{\mathrm{T}}$；$\boldsymbol{\Pi}'_{44}=14\tilde{\boldsymbol{M}}+14\tilde{\boldsymbol{M}}^{\mathrm{T}}$；$\boldsymbol{\Pi}'_{55}=-\gamma_2^2\boldsymbol{I}$；$\boldsymbol{\Pi}'_{56}=-(\bar{h}^2/4)m_4\boldsymbol{K}_{1l}^{\mathrm{T}}\boldsymbol{E}_f-\left[\bar{h}^2n_3+(\bar{h}^2/4)m_3+\bar{h}n_2\right]\cdot\boldsymbol{K}_{1l}^{\mathrm{T}}\boldsymbol{E}_f$；$\boldsymbol{\Pi}'_{57}=\bar{h}^2n_3\boldsymbol{K}_{1l}^{\mathrm{T}}\boldsymbol{B}_w+\left[(\bar{h}^2/4)m_3+(\bar{h}^2/4)m_4+\bar{h}n_2\right]\cdot\boldsymbol{K}_{1l}^{\mathrm{T}}\boldsymbol{B}_w$；$\boldsymbol{\Pi}'_{510}=\left[\bar{h}^2n_3+(\bar{h}^2/4)m_3+(\bar{h}^2/4)m_4+\bar{h}n_2\right]\boldsymbol{K}_{1l}^{\mathrm{T}}$；$\boldsymbol{\Pi}'_{66}=\left[\bar{h}^2n_3+(\bar{h}^2/4)m_3+(\bar{h}^2/4)m_4+\bar{h}n_2\right]\boldsymbol{E}_f^{\mathrm{T}}\tilde{\boldsymbol{P}}\boldsymbol{E}_f-\gamma_2^2\boldsymbol{I}$；$\boldsymbol{\Pi}'_{67}=-\left[\bar{h}^2n_3+(\bar{h}^2/4)m_3+\bar{h}n_2\right]\boldsymbol{E}_f^{\mathrm{T}}\tilde{\boldsymbol{P}}\boldsymbol{B}_w+(\bar{h}^2/4)m_4\boldsymbol{E}_f^{\mathrm{T}}\tilde{\boldsymbol{P}}\boldsymbol{B}_w$；$\boldsymbol{\Pi}'_{77}=(\bar{h}^2/4)m_4\boldsymbol{B}_w^{\mathrm{T}}\tilde{\boldsymbol{P}}\boldsymbol{B}_w-\gamma_2^2\boldsymbol{I}+\left(\bar{h}^2n_3+(\bar{h}^2/4)m_3++\bar{h}n_2\right)\cdot\boldsymbol{B}_w^{\mathrm{T}}\tilde{\boldsymbol{P}}\boldsymbol{B}_w$；$\boldsymbol{\Pi}'_{88}=\sigma\boldsymbol{\Phi}$；$\boldsymbol{\Pi}'_{99}=-\boldsymbol{\Phi}$；$\boldsymbol{\Pi}'_{1111}=-\tilde{\boldsymbol{P}}^{-1}/9\bar{h}n_2$；$\boldsymbol{\Pi}'_{1010}=-\bar{h}^2n_3\tilde{\boldsymbol{P}}-\left[(\bar{h}^2/4)m_3+(\bar{h}^2/4)m_4+\bar{h}n_2\right]\tilde{\boldsymbol{P}}$。

$$\begin{cases}\begin{bmatrix}\boldsymbol{Q}_1 & \boldsymbol{E}_f^{\mathrm{T}}\tilde{\boldsymbol{P}}\\ & m_1\tilde{\boldsymbol{P}}\end{bmatrix}>\boldsymbol{0},\begin{bmatrix}\boldsymbol{Q}_3 & n_1\boldsymbol{E}_f^{\mathrm{T}}\tilde{\boldsymbol{P}}\\ & m_2\tilde{\boldsymbol{P}}\end{bmatrix}>\boldsymbol{0}\\ \begin{bmatrix}\boldsymbol{Q}_5 & n_3\boldsymbol{E}_f^{\mathrm{T}}\tilde{\boldsymbol{P}}\\ & m_3\tilde{\boldsymbol{P}}\end{bmatrix}>\boldsymbol{0},\begin{bmatrix}\boldsymbol{Q}_7 & n_2\boldsymbol{E}_f^{\mathrm{T}}\tilde{\boldsymbol{P}}\\ & m_4\tilde{\boldsymbol{P}}\end{bmatrix}>\boldsymbol{0}\end{cases} \tag{20}$$

则存在控制器增益矩阵$\boldsymbol{K}_l$和事件触发矩阵$\boldsymbol{\Phi}_l$, 使得 ICPS 闭环系统式 (17) 无扰动时渐进稳定, 有扰动时满足$\boldsymbol{H}_\infty$性能指标: $\|\boldsymbol{x}(t)\|_2^2\leqslant\gamma_2^2[\|\bar{\boldsymbol{w}}(t)\|_2^2+\bar{h}(\|\boldsymbol{e}_x(j_k)\|_2^2+\|\boldsymbol{e}_f(j_k)\|_2^2)]$, $\boldsymbol{K}_l=(\tilde{\boldsymbol{P}}\boldsymbol{B})^{-1}\boldsymbol{K}_{1l}$与$\boldsymbol{\Phi}_l$可协同求取。

定理 2 的证明与定理 1 类似, 不再赘述。

当给定控制器性能指标γ_2, 由式 (21) 可以得出控制器最大允许时延h_2:

$$\begin{cases}\underset{\tilde{\boldsymbol{P}},\tilde{\boldsymbol{M}},\boldsymbol{K}_1,\boldsymbol{Q}_1,\boldsymbol{Q}_3,\boldsymbol{Q}_5,\gamma_2}{\mathrm{Max}} & h_2\\ \mathrm{s.t.}\boldsymbol{P}>\boldsymbol{0}\end{cases} \tag{21}$$

4.3　基于模糊融合方法建立控制器融合机制

基于 K-Means++算法及本文设计的不同模态下的控制器, 多模态控制器既要自适应于触发参数的变化, 又要消除不同模态间切换引起的抖振。选取可以反映系统运行状态的触发参数和触发误差作为模糊融合的输入, 进而根据模糊规则和反模糊化处理得到模糊融合的输出, 包括不同模态的权重系数。当前时刻对应的自适应综合安全控制策略为

$$\boldsymbol{u}(t)=\sum_{l=1}^{s}\kappa_l\boldsymbol{u}_l(t)\Big/\sum_{l=1}^{s}\kappa_l\quad l=1,2,\cdots,\tilde{n} \tag{22}$$

式中: κ_l为第l个模态的权重系数; $\tilde{n}$为模态数。

5　仿真实验与结果分析

5.1　实验描述

为验证本文方法的有效性, 采用工业常见的四容水箱系统[20]进行验证, 模型参数为

$$\begin{cases}\boldsymbol{A}=\begin{bmatrix}-0.016 & 0 & 0.042 & 0\\ 0 & -0.011 & 0 & 0.033\\ 0 & 0 & -0.042 & 0\\ 0 & 0 & 0 & -0.033\end{bmatrix}\\ \boldsymbol{B}=\begin{bmatrix}0.083 & 0\\ 0 & 0.063\\ 0 & 0.048\\ 0.031 & 0\end{bmatrix},\boldsymbol{B}_w=\begin{bmatrix}0.01\\ 0.01\\ 0\\ 0.01\end{bmatrix}\\ \boldsymbol{E}_v=\begin{bmatrix}0.01\\ 0\\ 0.01\\ 0.01\end{bmatrix},\boldsymbol{C}=0.5\boldsymbol{I}_4,\boldsymbol{E}_f=\begin{bmatrix}0.083\\ 0\\ 0\\ 0.031\end{bmatrix}\end{cases} \tag{23}$$

初始状态为$\boldsymbol{x}(0)=[\,4\quad 4\quad 2\quad 2\,]^{\mathrm{T}}$, 仿真过程中存在的扰动$\boldsymbol{w}(t)$及噪声$\boldsymbol{v}(i_k)$为服从$N(0.1,0.01)$的高斯白噪声。仿真时间为 800 s, 采样器的采样周期为 0.1 s。假定在 200 ～300 s 之间系统传感端将被注入小能量 DoS 攻击, 在 400 ～600 s 之间传感端将被注入大能量 DoS 攻击。假定执行端 DoS 攻击发生在 400 ～600 s 之间。

执行器时变故障为

$$f(t)=\begin{cases}0 & t\leqslant 250\ \mathrm{s}\\ 2+2\sin 0.01\pi(t-250) & 250\ \mathrm{s}<t\leqslant 800\ \mathrm{s}\end{cases} \tag{24}$$

5.2　LSTM 数据重构模型评价

模拟四容水箱系统在双端异步 DoS 攻击下的实际运行工况, 考虑执行器故障, 分别采集 8 000 组传感端遭受小能量 DoS 攻击的系统输出数据, 及执行端未遭受 DoS 攻击的控制量数据作为训练和测试样本, 进行 LSTM 网络预测模型的训练, 以得到传感端和执行端的重构补偿模型。为客观评价 LSTM 网络预测模型的优越性, 选取文献 [21] 中的 PD 数据重构补偿方法与本文提出的基于 LSTM 网络的模型进行重构补偿方法进行对比。同时, 为便于分析, 只选取传感端系统输出数据的预测结果进行对比, 图 3 和图 4 为基于不同模型的数据重构补偿误差。

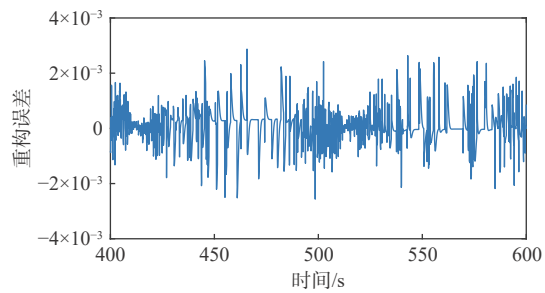


图 3　基于 LSTM 网络的重构误差

Fig. 3　Reconstruction error based on LSTM networks

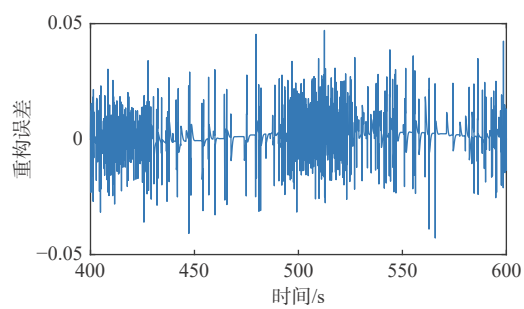


图 4　基于 PD 的重构误差

Fig. 4　Reconstruction error based on PD

由图 3 和图 4 可知, 本文基于 LSTM 网络模型的重构误差在±0.003以内, 而基于 PD 的重构补偿误差在±0.05左右。究其原因, LSTM 网络是一种循环神经网络模型, 可以有效地捕捉时间序列数据中的长期依赖关系。相比之下, PD 补偿方法难以捕

捉复杂的时间序列与时序特征，且模型参数需要反复试探选择且无明显规律可循，因此，难以对丢失的数据进行准确补偿。此外，LSTM网络模型的数据重构效果也可以通过后续状态估计误差及系统响应输出曲线的变化趋势来反映，即当数据重构误差较大时，系统响应输出和状态估计误差会出现较明显的波动。

5.3 基于K-Means++的模态划分结果分析

为确定最优控制器模态数量，采用K-Means++算法将系统在不同工况下的触发误差及事件触发参数进行聚类，确定最优控制器模态个数。图5和图6为确定最优聚类个数的仿真结果，给出了采用K-Means++算法将聚类个数从2～9的数值变化趋势，以确定最佳的控制模态个数。

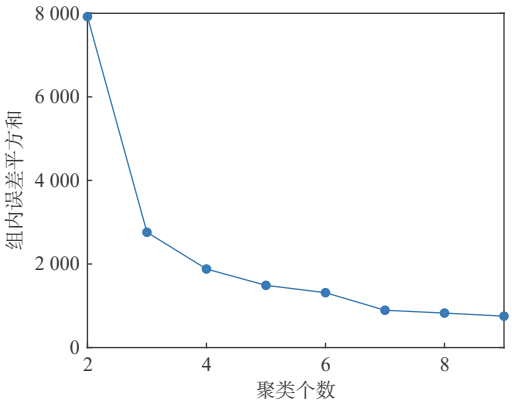


图5 手肘法

Fig. 5 Elbow method

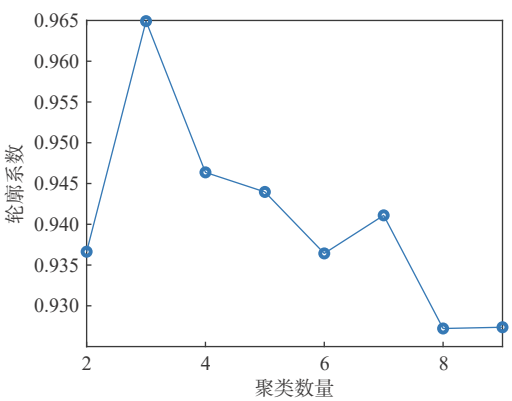


图6 轮廓系数法

Fig. 6 Contour coefficient method

由图5和图6可知，将聚类个数选择为2～9，当聚类数为3时，对应轮廓系数值最大且组内误差平方和下降幅度也最大，因此，最佳的控制模态个数为3。采用K-Means++算法将事件触发误差及事件触发参数按3类进行聚类，可得到每个模态的范围及聚类中心，如图7所示。

由图7可知，每个控制器模态所对应的聚类中心及每个触发误差e_y和触发参数$\bar{\sigma}(t_k h)$的范围，如表1所示。

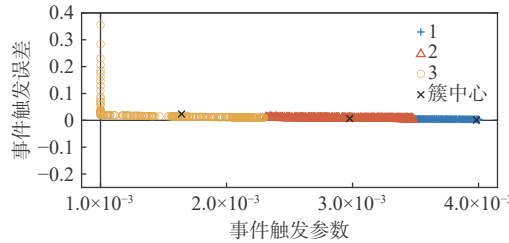


图7 基于K-Means++的聚类结果

Fig. 7 Plot of clustering results based on K-Means++

表1 不同控制模态下的参数

Table 1 Parameters under different control modes

模态	$\bar{\sigma}(t_k h)$聚类中心	e_y聚类中心	$\bar{\sigma}(t_k h)$范围	e_y范围
C_1	0.004	0.001 9	(0.003 5,0.004)	(0,0.005)
C_2	0.003	0.006 9	(0.002 3,0.003 5)	(0.005,0.009 7)
C_3	0.001 6	0.002 34	(0.001,0.002 3)	(0.009 7,0.36)

由表1可知，不同模态下触发参数和触发误差的聚类中心及每个模态中触发参数和触发误差对应的范围。其中，模态C_1的触发误差较大，触发参数较小；模态C_3的触发误差较小，触发参数较大；模态C_2则介于模态C_1和模态C_3之间。模态C_1、模态C_3、模态C_2分别对应系统运行处于非稳定状态、稳定状态和抖动状态。

5.4 状态、故障估计与多模态综合安全控制仿真

5.4.1 状态与故障估计方法的有效性验证

对于本文所设计的多目标约束下状态与故障观测器，当限定参数$\gamma_1=\gamma_2=2.8$，由式(14)和式(21)可得$h_1=1.62$，$h_2=1.93$，则可以求得系统最大允许时延$h_{\mathrm{M}}=\min(h_1,h_2)=1.62$。由定理1，给定参数$n_1=0.3$，$n_2=2.8$，$n_3=0.5$，$\theta=0.5$，$\bar{h}=0.5h_{\mathrm{M}}=0.81$，可求得观测器增益矩阵$\boldsymbol{L}$和故障估计增益矩阵$\boldsymbol{F}$如下：

$$\begin{cases}\boldsymbol{L}=\begin{bmatrix}-4.583\,7 & -14.515\,8 & -16.487\,3 & 22.500\,1\\ -0.052\,2 & 2.718\,9 & -0.114\,0 & 0.183\,1\\ 0.911\,3 & 1.502\,9 & 2.586\,2 & -2.520\,6\\ -2.224\,6 & -4.537\,4 & -6.342\,2 & 9.711\,1\end{bmatrix}\\ \boldsymbol{F}=[2.278\,1 \quad 4.379\,6 \quad 4.498\,5 \quad -6.915\,0]\end{cases} \tag{25}$$

针对传感端因大能量DoS攻击而未成功传输的系统数据，基于本文模型进行数据重构补偿。将系统输出数据送入观测器，实现对系统状态和执行器故障的准确估计。系统状态和执行器故障的估计误差如图8和图9所示。

由图8和图9可知，在采用基于数模联动的主-被动协同混合容侵策略后，基于定理1中的方法来

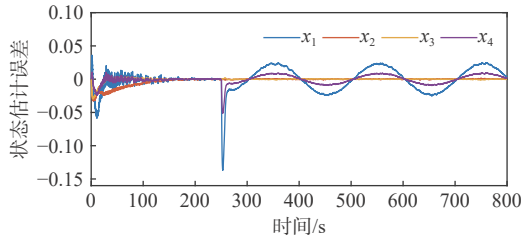


图 8　系统状态估计误差

Fig. 8　Estimation error of system state

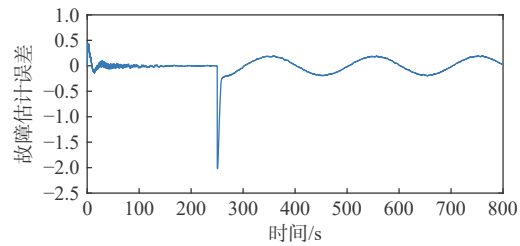


图 9　执行器故障估计误差

Fig. 9　Estimation error of actuator failure

设计的鲁棒观测器可以精确地估计出系统的运行状态及遭遇到的执行器故障, 且对小能量 DoS 攻击和外部扰动具有较强的鲁棒性, 即实现了对小能量 DoS 攻击的被动容侵。

5.4.2　多模态综合安全控制方法有效性验证

由$(\boldsymbol{I}-\boldsymbol{B}\boldsymbol{B}^{+})\boldsymbol{E}_f=\boldsymbol{0}$求得故障调节矩阵如下:

$$\boldsymbol{B}^{+}=\begin{bmatrix}11.955\,0 & -10.573\,2 & 0 & 0.249\,6\\ 0 & 0 & 0 & 0\end{bmatrix} \tag{26}$$

通过 K-Means++算法将事件触发参数按 3 类进行聚类得到每个模态的聚类中心后, 以事件触发参数的聚类中心 0.001 6、0.003、0.004 为依据进行控制器的设计。不依赖控制模态的前提下, 给定如下参数$n_1=0.8, n_2=1.8, n_3=1.4, m_1=1.9$, $m_2=1.5, m_3=1.2, m_4=1.2, \bar{h}=0.81$。下一步则根据触发参数$\sigma_1=0.001\,6$, $\sigma_2=0.003$, $\sigma_3=0.004$得到不同控制模态下的控制器。由定理 2 可得

$$\boldsymbol{\Phi}_1=10^{-8}\times\begin{bmatrix}4.917\,4 & 0 & 0 & 0\\ 0 & 4.917\,4 & 0 & 0\\ 0 & 0 & 4.917\,4 & 0\\ 0 & 0 & 0 & 4.917\,4\end{bmatrix} \tag{27}$$

$$\boldsymbol{K}_1=\begin{bmatrix}-1.036\,1 & -0.023\,2 & -1.067\,0 & -0.211\,4\\ -0.009\,4 & -1.107\,4 & -0.264\,8 & -0.803\,3\end{bmatrix} \tag{28}$$

$$\boldsymbol{\Phi}_2=10^{-8}\times\begin{bmatrix}4.833\,9 & 0 & 0 & 0\\ 0 & 4.833\,9 & 0 & 0\\ 0 & 0 & 4.833\,9 & 0\\ 0 & 0 & 0 & 4.833\,9\end{bmatrix} \tag{29}$$

$$\boldsymbol{K}_2=\begin{bmatrix}-1.092\,7 & 0.028\,7 & -0.974\,0 & -0.243\,6\\ 0.015\,5 & -1.153\,1 & -0.341\,0 & -0.736\,8\end{bmatrix} \tag{30}$$

$$\boldsymbol{\Phi}_3=10^{-8}\times\begin{bmatrix}2.567\,3 & 0 & 0 & 0\\ 0 & 2.567\,3 & 0 & 0\\ 0 & 0 & 2.567\,3 & 0\\ 0 & 0 & 0 & 2.567\,3\end{bmatrix} \tag{31}$$

$$\boldsymbol{K}_3=\begin{bmatrix}-1.092\,5 & 0.029\,2 & -0.973\,9 & -0.242\,9\\ 0.015\,9 & -1.152\,6 & -0.340\,8 & -0.736\,5\end{bmatrix} \tag{32}$$

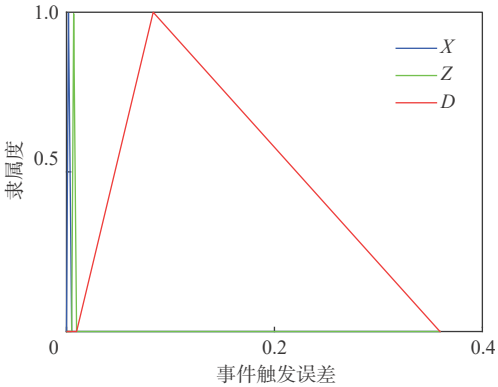
为确保传入控制量融合器的事件触发误差e_y和事件触发参数$\bar{\sigma}(t_k h)$具有相同的时序特征, 因此, 采用式 (5) 观测器实际传入数据$\bar{\boldsymbol{y}}(j_k)$用于传入控制量融合器中触发参数$\bar{\sigma}(t_k h)$及事件触发误差$e_y$的计算。通过表 1 中不同模态的范围和聚类中心, 确定模糊融合的输入及输出, 如图 10 所示, 图中, X、Z、D 分别表示中间范围、较小、较大的触发参数或触发误差。

图 10 为模糊融合的输入输出隶属函数曲线, 其中, 输入变量触发误差和触发参数的隶属度函数选用三角型隶属度函数, 选取三角形的顶点为各模态的聚类中心以确保在聚类中心的隶属度为 1, 且隶属度由聚类中心向两侧依次递减。输出变量则选用更为平滑的高斯隶属度函数, 目的是防止输出变量模糊化时出现模糊值突变的情况, 避免出现由模态切换引起的抖振, 进而达到系统具有良好控制性能的目标。

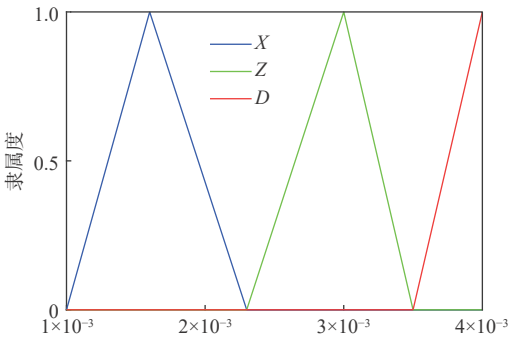
在确定模糊融合的输入输出后, 根据触发误差和触发参数的变化规律, 建立模糊规则, 如表 2 所示。

建立模糊规则后, 为进一步消除模态切换引起的抖振, 在反模糊化时采用重心法, 从而将模糊集合的信息以更平滑的方式表达出来, 使得反模糊化后的输出更加连续和稳定, 反模糊化后的输出值为不同模态下控制量的加权平均值, 实现了不同控制模态间的软切换, 使得模态间的切换更加平滑。κ_l将事件触发误差e_y和事件触发参数$\bar{\sigma}(t_k h)$作为输入变量, 根据预先设计好的模糊规则进行实时计算, 从而得到每个时刻控制模态的权重系数, 进而以信息融合软切换的方式实现不同模态间的切换。

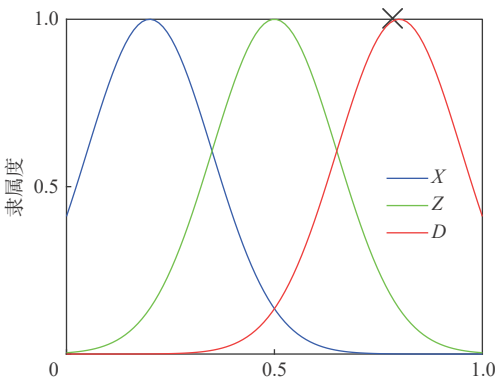
为凸显上述基于模糊融合的多模态综合安全控制器的优势, 分别给出了不同融合策略下的多模态控制系统输出响应及单模态控制系统输出响应, 如图 11 所示。其中, 采用文中模糊融合软切换下多模态控制系统的输出响应, 如图 11(a) 所示, 采用



(a) 事件触发误差的隶属度函数曲线



(b) 事件触发参数的隶属度函数曲线



(c) 控制器的隶属度函数曲线

图 10 输入输出的隶属度函数曲线

Fig. 10 Membership function curves of input and output

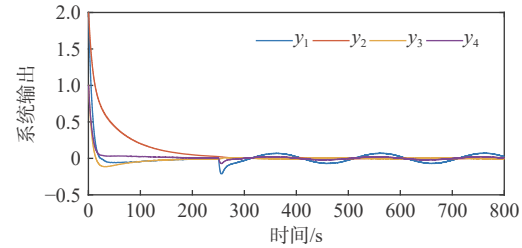
表 2 模糊规则

Table 2 Fuzzy rules

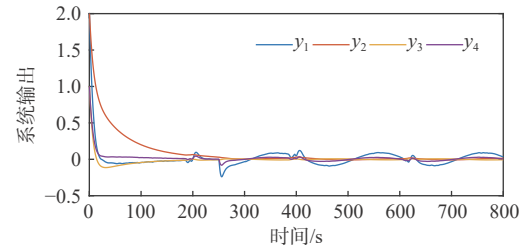
$\sigma(t_k h)$	$\Delta\bar{y}(j_k)$		
	D	*Z*	*X*
D	*D*/*D*	*D*/*Z*	*D*/*X*
Z	*Z*/*D*	*Z*/*Z*	*Z*/*X*
X	*X*/*D*	*X*/*Z*	*X*/*X*

文献 [22] 以区间阈值进行不同模态间的切换，采用表 1 所示的触发参数和触发误差的具体区间，以区间阈值进行不同模态间的切换，得到系统输出响应如图 11(b) 所示，同时还给出了单模态下的系统输出响应如图 11(c) 所示。

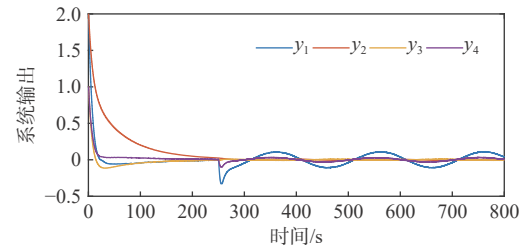
观察图 11(a) 中 4 种情形下系统输出响应可



(a) 软切换下多模态控制系统输出



(b) 以区间阈值硬切换下多模态控制系统输出[22]



(c) 单模态控制器下的系统输出

图 11 多/单模态控制器下的系统输出

Fig. 11 System output under multimodal/unimodal controllers

得：当$t \leqslant 200$ s时，系统输出快速趋于稳定状态；当200 s $< t \leqslant 250$ s时，在系统中存在小能量 DoS 攻击，系统输出响应依然趋于稳定状态，且无明显波动；当250 s $< t \leqslant 400$ s时，在 250 s 时加入故障的瞬间，系统输出产生了较小的波动，但在很短的时间内收敛至稳定状态；当 400 s $< t \leqslant 800$ s时，系统在遭受执行器故障的情形下，又遭受了双端异步 DoS 攻击，系统输出依然保持稳定且波动范围保持在±0.2以内。由图 11(b) 可知，基于区间阈值硬切换下多模态控制器的系统响应输出可以看出，在 100 s 内系统由初始状态快速收敛到平衡状态，但分别在 200 s、400 s、600 s 左右出现明显波动，究其原因是采用固定区间阈值进行不同控制器模态间的切换引起了系统抖振，进而造成系统输出响应的波动。相较于图 11(c)，单模态控制器下的系统输出响应可以看出，多模态控制系统的输出在 250 s 加入故障瞬间的波动幅值有所减小，且在 600 s 后更加平稳。

本节结果表明，本文设计的鲁棒观测器和多模态综合安全控制器对传感端小能量 DoS 攻击具有较强鲁棒性，实现对 DoS 攻击的被动容侵。同时，结合基于数据驱动的主动补偿策略可以对双端异步 DoS 攻击进行精准容侵，且多模态控制器可实现对执行器故障的有效调节，确保系统有效抵御执行

器故障与双端异步 DoS 攻击的影响。

5.5　ADETCS 对通讯资源的节约分析

本文设计的 ADETCS 中事件触发参数变化曲线, 如图 12 所示。

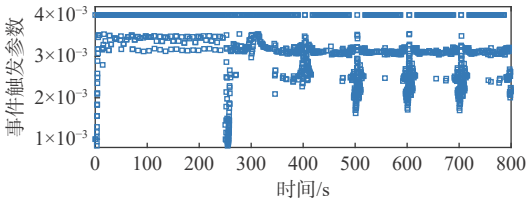


图 12　事件触发参数的变化曲线

Fig. 12　Curve of changes in event triggered parameters

由图 12 可知, 在系统从初始到趋于稳定状态, 再到 250 s 时系统内加入执行器故障, 事件触发参数可随系统运行状况而自适应变化。当系统状态出现波动时, 事件触发参数随之自适应减小, 以传输更多数据, 进而使系统快速恢复稳定; 而当系统平稳运行时, 事件触发参数则自适应增大, 以节约通讯资源。

不同触发机通讯制下数据传输量的比较, 如表 3 所示。

表 3　不同触发机通讯制下数据传输量的比较

Table 3　Comparison of data transmission under different trigger machine communication systems

触发机制	数据传输量	传输周期/s
DETCS[21]	4 201	0.191
ADETCS[23]	3 985	0.208
本文ADETCS	3 751	0.213

由表 3 可知, DETCS 驱动下的文献 [21] 在 800 s 内传输了 4 201 个数据, 平均传输周期为 0.191 s, ADETCS 驱动下触发参数阈值是根据系统运行状态灵活变化的。与 DETCS 相比, 本文 ADETCS 在 800 s 内仅需要传输 3 751 个数据, 平均传输周期为 0.213 s, 较文献 [21] 节约更多的网络通讯资源。与文献 [23] 的 ADETCS 相比, 本文 ADETCS 由于设计了触发阈值的上界 $\sigma_{\max}$, 所以在系统稳定时传输更少的数据量, 进一步实现了对通讯资源的节约。总之, ADETCS 较 DETCS 能节约更多的网络通讯资源, 究其原因是, 将可随系统运行状态而自适应变化的触发参数选择机制植入 ADETCS 中, 解决了由 DETCS 中触发参数阈值固定不变而威胁系统稳定性的问题, 且在网络通讯资源节约方面取得了更好的效果。

6　结　论

本文综合考虑了 ICPS 面临的物理安全、信息安全、通讯资源节约等问题, 以同时存在双端异步 DoS 攻击和执行器故障的 ICPS 为研究对象, 研究了其综合安全控制与网络通信的协同设计问题, 得出如下结论:

1) 用 ADETCS 替代了传统的 DETCS, 并通过机理解析和数据驱动方法的融合联动形成主-被动协同的混合容侵策略来抵御双端异步 DoS 攻击, 进而构建了新型 ADETCS 下多模态 ICPS 综合安全控制架构。

2) 对于由传感端大能量 DoS 攻击和执行端 DoS 攻击而导致的数据丢失, 研究了基于数据驱动 LSTM 网络主动容侵策略对其进行主动补偿的方法, 并推证出了具有小能量 DoS 攻击被动容侵能力的鲁棒观测器和多模态综合安全控制器, 使得 ICPS 对于双端异步 DoS 攻击的抵御能力得到了明显提升。

3) 针对现存控制器的设计多采用某一固定触发参数, 未能真正自适应于系统行为变化的问题, 结合多模态控制思想, 先采用 K-Means++算法对系统不同工况下的工作模态进行划分, 进而采用模糊理论对不同模态控制器输出进行融合, 最终形成了多模态柔性切换策略, 确保 ICPS 的性能得到进一步的提升。

参考文献 (References)

[1] WU Y Y, ZHANG L W, XIE W, et al. Robust mixed H_2/H_∞ model predictive control for cyber-physical systems with input saturation and energy-bounded disturbance[J]. IEEE Transactions on Industrial Informatics, 2024, 20(6): 8328-8337.

[2] 薛威峰, 孙子文. 数据注入攻击下 ICPS 的自适应事件触发弹性控制[J]. 控制理论与应用, 2024, 41(5): 866-874.
XUE W F, SUN Z W. Adaptive event triggered resilient control of ICPS under data injection attack[J]. Control Theory & Applications, 2024, 41(5): 866-874(in Chinese).

[3] JI Z D, CHEN C L, HE J P, et al. Edge sensing and control co-design for industrial cyber-physical systems: observability guaranteed method[J]. IEEE Transactions on Cybernetics, 2022, 52(12): 13350-13362.

[4] ZUO F L, LUO Z X, YU J Z, et al. Vulnerability detection of ICS protocols via cross-state fuzzing[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2022, 41(11): 4457-4468.

[5] 梁宏, 刘佳男, 李勇. “火焰”病毒分析与防范[J]. 信息网络安全, 2012, 12(8): 157-159.
LIANG H, LIU J N, LI Y. Analysis and protection of flame virus[J]. Netinfo Security, 2012, 12(8): 157-159(in Chinese).

[6] YANG H Q, LI T S, LONG Y, et al. Event-triggered distributed secondary control with model-free predictive compensation in AC/DC networked microgrids under DoS attacks[J]. IEEE Transactions on Cybernetics, 2024, 54(1): 298-307.

[7] YANG F S, LIANG X H, GUAN X H. Resilient distributed economic dispatch of a cyber-power system under DoS attack[J]. Frontiers of Information Technology & Electronic Engineering, 2021, 22(1): 40-50.

[8] 彭大天, 董建敏, 蔡忠闽, 等. 假数据注入攻击下信息物理融合系统的稳定性研究[J]. 自动化学报, 2019, 45(1): 196-205.
PENG D T, DONG J M, CAI Z M, et al. On the stability of cyber-physical systems under false data injection attacks[J]. Acta Automatica Sinica, 2019, 45(1): 196-205(in Chinese).

[9] GONG Z Y, YANG F S. Secure tracking control via fixed-time convergent reinforcement learning for a UAV CPS[J]. IEEE/CAA Journal of Automatica Sinica, 2024, 11(7): 1699-1701.

[10] 张正道, 杨佳佳, 谢林柏. 基于辅助信息补偿和控制信号编码的重放攻击检测方法[J]. 自动化学报, 2023, 49(7): 1508-1518.
ZHANG Z D, YANG J J, XIE L B. Replay attack detection method based on auxiliary information compensation and control signal coding[J]. Acta Automatica Sinica, 2023, 49(7): 1508-1518 (in Chinese).

[11] 杨中林, 魏自航, 林相泽. 线性系统的事件触发输出反馈有限时间有界控制[J]. 控制理论与应用, 2022, 39, (8): 1497-1505.
YANG Z L, WEI Z H, LIN X Z. Event-triggered output feedback finite-time bounded control of linear systems[J]. Control Theory and Applications, 2022, 39, (8): 1497-1505(in Chinese).

[12] 陈烁, 樊渊. 基于采样的线性不确定系统事件触发控制[J]. 控制工程, 2021, 28(5): 999-1004.
CHEN S, FAN Y. Sampling-based event-triggered control for linear uncertain systems[J]. Control Engineering of China, 2021, 28(5): 999-1004(in Chinese).

[13] HUANG S, ZONG G D, XU N, et al. Adaptive dynamic surface control of MIMO nonlinear systems: a hybrid event triggering mechanism[J]. International Journal of Adaptive Control and Signal Processing, 2024, 38(2): 437-454.

[14] ZHAO C H, SHEN Y J, PANG H X. Adaptive output-feedback control for switched nonlinear systems with measurement sensitivity based on double-side event-triggering mechanisms[J]. International Journal of Systems Science, 2024, 55(11): 2349-2372.

[15] FRIDMAN E. A refined input delay approach to sampled-data control[J]. Automatica, 2010, 46(2): 421-427.

[16] HOCHREITER S, SCHMIDHUBER J. Long short-term memory [J]. Neural Computation, 1997, 9(8): 1735-1780.

[17] ZHANG X M, HAN Q L, SEURET A, et al. Overview of recent advances in stability of linear systems with time-varying delays[J]. IET Control Theory & Applications, 2019, 13(1): 1-16.

[18] ZHANG X M, HAN Q L. Global asymptotic stability analysis for delayed neural networks using a matrix-based quadratic convex approach[J]. Neural Networks, 2014, 54: 57-69.

[19] ARTHUR D,VASSILVITSKII S.K-means++: the advantages of careful seeding[C]//Proceedings of the 18th Annual ACM-SIAM Symposium on Discrete Algorithms.New York: ACM, 2007:1027-1035.

[20] 邱爱兵, 吉虹钢, 顾菊平. 非均匀采样数据系统时变故障估计与调节最优集成设计[J]. 自动化学报, 2014, 40(7): 1493-1504.
QIU A B, JI H G, GU J P. Optimal integrated design of time-varying fault estimation and accommodation for nonuniformly sampled data systems[J]. Acta Automatica Sinica, 2014, 40(7): 1493-1504(in Chinese).

[21] 李炜, 韩小武, 李亚洁. 基于 DoS 攻击能量分级的 ICPS 综合安全控制与通信协同设计[J]. 信息与控制, 2022, 51(3): 257-270.
LI W, HAN X W, LI Y J. Co-design of integrated security control and communication of ICPS based on DoS attack energy classification[J]. Information and Control, 2022, 51(3): 257-270(in Chinese).

[22] 程雪. ADETCS 下 ICPS 多模态综合安全控制与通讯协同设计[D]. 兰州: 兰州理工大学, 2023: 42-50.
CHENG X. Collaborative design of multimodal integrated safety control and communication of ICPS under ADETCS[D]. Lanzhou: Lanzhou University of Technology, 2023: 42-50(in Chinese).

[23] PENG C, ZHANG J, YAN H. Adaptive event-triggering load frequency control for network-based power system[J]. IEEE Transactions on Industrial Electronics, 2016, 63(8): 1685-1694.

ICPS multi-modal integrated security control based on data-model linkage under dual-end asynchronous DoS attacks

LI Yajie，LI Gang，LI Wei*，LU Chenjing

(School of Electrical and Information Engineering，Lanzhou University of Technology，Lanzhou 730050，China)

Abstract：Data-driven technology and model mechanism analysis techniques are combined to study the problem of multi-modal integrated safety control and communication collaborative design, with an eye toward a type of industrial cyber-physical system (ICPS) that is susceptible to dual-end asynchronous denial of service (DoS) attacks and actuator failures. Firstly, an adaptive discrete event triggered communication scheme (ADETCS) with a trigger threshold that can dynamically change with the system behavior is designed, and an ICPS multi-modal integrated safety control architecture that can simultaneously resist asynchronous dual-end DoS attacks and actuator failures is constructed. Secondly, an active-passive collaborative hybrid intrusion tolerance strategy based on data-model linkage is proposed. Then, combined with the idea of “divide and conquer”, a tolerance method for dual-end asynchronous DoS attacks is proposed with the help of a long short-term memory (LSTM) networks and elastic control. Thirdly, the observer and controller are deduced based on Lyapunov stability theory, and then the K-Means++ clustering algorithm and fuzzy fusion method are used to perform weighted fusion of controllers under different modes online, thus realizing soft switching between different control modes. Finally, the effectiveness of the proposed method is verified through a four-tank example. According to the experimental results, the data-model linkage method improves ICPS's resistance to dual-end asynchronous DoS attacks, and the multi-modal integrated safety controller's design enables two-way adaptive cooperative control between the ADETCS and control mode.

Keywords：industrial cyber-physical systems; dual-end asynchronous DoS attack; data-model linkage; adaptive discrete event-triggered communication scheme; multi-modal integrated safety control

Received：2024-11-15；**Accepted**：2025-04-11；**Published Online**：2025-05-14 13：03

URL：link.cnki.net/urlid/11.2625.V.20250514.1025.001

Foundation items：National Natural Science Foundation of China (62163022); Major Science and Technology Special Project of Gansu Province in China (23ZDGA007); Industrial support plan project of Gansu Province in China (2024CYZC-18); Joint Research Fund Project of Gansu Province in China (24JRRA827); Science and Technology Plan Project of Lanzhou City in China (2024-2-5)

* **Corresponding author**. E-mail：liwei@lut.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0528

富氧火炬式电点火器出口火焰稳定性研究

杨进慧*，王希杰，徐世洋，王晓丽
(北京航天动力研究所，北京 100076)

摘　　要：富氧火炬式电点火器在低压、宽工况及推进剂入口温度大幅变化等极限工况下存在燃气温度偏低、不均匀及出口流速过快等火焰不稳定性现象。基于此，采用试验仿真方法研究了富氧火炬式电点火器缩进、扩口、扩口凹腔3种火焰稳定结构对出口火焰稳定性的影响。扩口凹腔结构将中心燃气向壁面引流，降低掺混区内流速延长燃气停留时间，并在凹腔结构内形成回流区促进补燃。极限工况下，相较于缩进结构，扩口凹腔结构火焰出口中心燃气温度下限由1 000 K提升至1 200 K，与设计工况温度差由高于100 K降至50 K以内，出口马赫数由1.4降至声速，有效提升了点火器的出口火焰稳定性。

关 键 词：点火器；极限工况；火焰稳定；缩进；扩口；扩口凹腔
中图分类号：V434.1
文献标志码：A　　**文章编号**：1001-5965(2025)10-3368-06

液氢液氧、液氧煤油、液氧甲烷等大推力液体火箭发动机需提供初始点火热源使非自燃推进剂点火燃烧，其中，火炬式电点火器因其点火能量高、可多次启动、结构简单、便于维护等特点被广泛采用，如美国的RL-10系列发动机[1]、航天飞机主发动机[2]、欧洲的Vinci发动机、日本的LE系列发动机[3]、俄罗斯的RD-0120发动机等[4]。高可靠点火技术是液体火箭发动机工作的第1步，其直接影响火箭的可靠性。

火炬式电点火器主要工作在发动机的起动段，推进剂入口压力低，温度接近气液两相点，物性参数变化剧烈，导致点火器流量及混合比变化范围很大[5]，尤其在上面级液体火箭发动机真空点火起动时，处于低压、低温和微重力环境，对点火器出口火焰的稳定性提出更高要求。点火器出口燃气射流的温度及均匀性是表征点火能量的重要指标，可用于评价不同环境下的点火器出口火焰能否满足发动机点火需求。

火炬式电点火器可分为富燃点火器和富氧点火器。富燃点火器一般采用膜冷却方式，点火流量大，结构简单，但混合比需小于当量混合比[6-7]；富氧点火器[8-10]的燃料一小部分参与燃烧，大部分作为冷却剂对点火室身部进行冷却，中心燃气为富氧状态，混合比适应范围较宽。富氧点火器中心燃气流量较小，点火能量低，需在点火器引火管出口处组织中心燃气与冷却燃料进一步补燃提高出口火焰温度，以避免液体火箭发动机出现点火失败或点火延迟。

富氧点火器中心燃气及冷却剂出口速度均可达声速，补燃区域距离短，试验中发现出口火焰随工况变化、分布不均匀、火焰温度不够等不稳定状态。航空发动机[11-13]及超燃冲压发动机[14-19]领域点火稳定性问题研究较多，凹腔结构作为火焰稳定器应用广泛，其通过回流区增强氧燃掺混达到稳焰目的，还可以扩宽点火混合比范围。本文将凹腔结构引入富氧火炬式电点火器，并结合扩口、缩进等

收稿日期：2023-08-17；**录用日期**：2023-11-03；**网络出版时间**：2023-11-14 09:35
网络出版地址：link.cnki.net/urlid/11.2625.V.20231113.0937.001
基金项目：中国航天科技集团自主研发项目
* **通信作者**. E-mail：yangjinhui@buaa.edu.cn

引用格式：杨进慧，王希杰，徐世洋，等. 富氧火炬式电点火器出口火焰稳定性研究[J]. 北京航空航天大学学报，2025，51(10)：3368-3373. YANG J H，WANG X J，XU S Y，et al. Study on flame stability of oxygen-rich torch igniter[J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51(10)：3368-3373 (in Chinese).

火焰稳定措施，采用试验与仿真相结合的方式研究了点火器出口火焰在低压、宽工况及变推进剂入口温度条件下的稳定性。

1　点火器模型及计算方法

富氧火炬式电点火器点火介质为液氧气氢，气氢进入点火器后一分为二：①经氢喷嘴与液氧掺混燃烧；②经排放冷却通道对点火器身部及下游燃气导管进行热防护[10]。点火器中心富氧燃气与冷却氢在燃气导管出口处进一步补燃燃烧，为增强燃气与冷却氢的掺混，如图 1 所示，共设置了缩进、燃气出口扩口、扩口结合稳焰凹腔 3 种补燃腔结构。

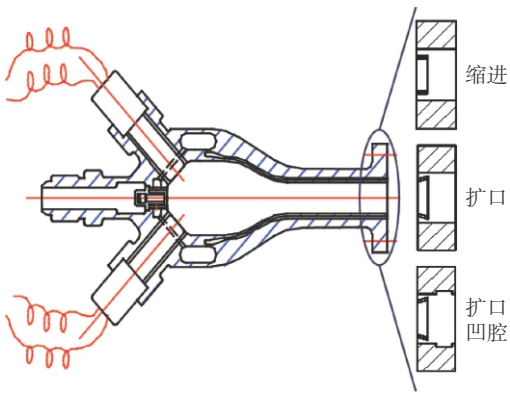


图 1　点火器火焰稳定结构示意图[10]

Fig. 1　Torch igniter flame stability reburning chamber[10]

采用 Fluent 仿真分析软件开展富氧点火器三维燃烧流动分析，采用结构化网格划分，并对出口补燃腔进行网格加密，如图 2 所示，经网格无关性验证后网格数量约为 1.2×10^6（与 1.5×10^6 网格压力相差最大 0.713%）。氢氧入口设置为质量流量入口，燃气出口为压力出口，壁面为无滑移绝热壁面。采用理想气体非预混燃烧（probability denstity function, PDF）模型评估燃烧流动情况，湍流采用标准 k-ε 湍流模型封闭方程组，近壁区域用标准壁面函数法进行处理。

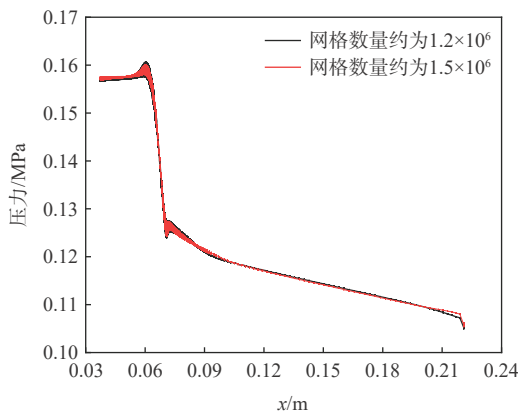


图 2　网格无关性验证

Fig. 2　Grid independence verification

2　模型验证

本文开展了富氧点火器缩进补燃腔结构的点火试验，试验包含点火器和引火管 2 部分，通过室压及引火管温度压力分析点火器内燃烧及热防护情况，如图 3 所示，引火管出口粘贴细铜丝定性分析缩进腔的补燃效果。

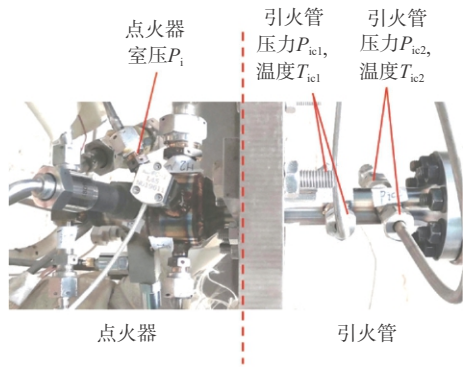


图 3　富氧点火器试验

Fig. 3　Oxygen-rich torch igniter experiment

富氧点火器试验压力温度如图 4 所示，试验中氧流量受供应系统影响存在波动，氧流量增大，中心燃气混合比增加，燃烧状态更加偏离当量混合比，冷却氢换热量减少而温度下降。由于结构限制，无法直接通过温度压力测量研究缩进腔内的补燃状态。

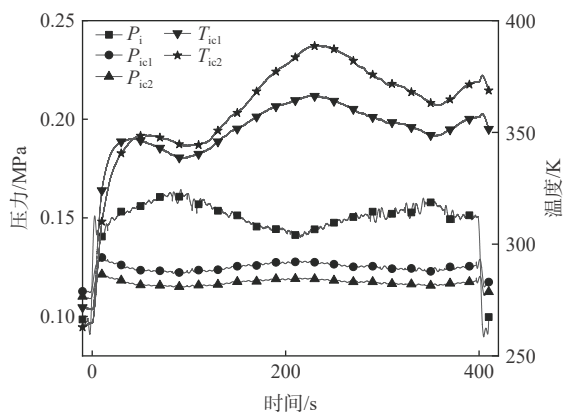


图 4　富氧点火器试验曲线

Fig. 4　Curves of oxygen-rich torch igniter experiments

点火器仿真结果与试验对比如图 5 所示，室压及冷却通道压力与试验误差最大为 4.2%，中心燃气计算混合比为 24，与试验中心燃气混合比 25.4 相差 5.5%。由于冷却氢径向温度梯度大，难以将仿真结果与试验温度进行对比。图 5 中提取了燃气侧壁面温度，最高约为 960 K，位于点火器喉部，引火管段温度降低，出口处壁面温度约为 725 K，试验后检查点火器仅喉部附近有轻微发黄，与仿真结果基本相符。

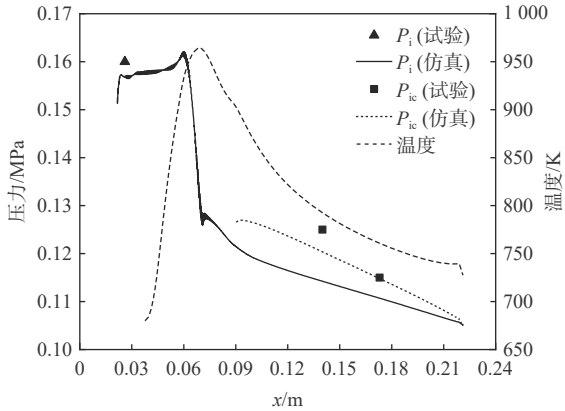


图5 仿真模型验证

Fig. 5 Validation of simulation model

富氧点火器试验中心燃气导管与引火管为缩进补燃腔结构，引火管出口粘贴的细铜丝在点火过程中从近圆心处烧断，如图6所示，径向向外发黑程度逐渐变弱为铜本色，说明中心燃气温度高于冷却氢与中心燃气的补燃温度，点火过程中缩进补燃腔内补燃效果不显著。

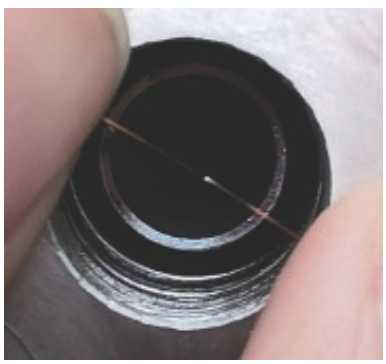


图6 缩进补燃腔出口细铜丝烧断情况

Fig. 6 Fusing of copper wire at recess reburning chamber outlet

3 点火器出口火焰稳定性分析

基于上述仿真模型及方法研究了缩进、扩口、扩口凹腔3种结构对富氧点火器出口火焰稳定性的影响，如图7所示，缩进结构中心燃气与冷却氢平行流出，仅在剪切层补燃，扩口结构能够将中心燃气向壁面引流，并降低掺混区内的冷却氢及中心燃气流速，凹腔结构通过流通面积局部增大将流速进一步降低，中心燃气静温随之增加，温度均匀性得之改善；另一方面凹腔内形成回流区，加强了中心燃气与冷却氢的换热与掺混，加之燃气停留时间延长，补燃区域随之扩大。

针对不同环境压力、混合比及推进剂入口温度，进一步研究3种稳焰结构下的出口燃气射流温度及均匀性。

3.1 环境压力影响

图8和图9对比分析了常压环境和低压环境

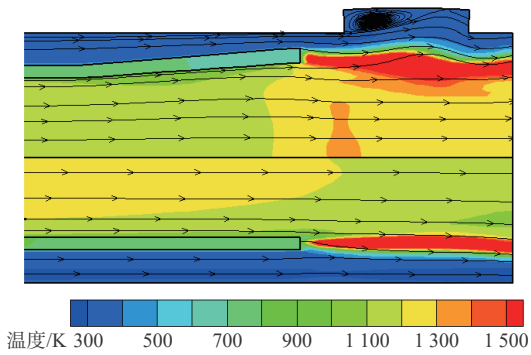


图7 扩口凹腔结构的局部流场分布

Fig. 7 Flow field distribution for flared cavity

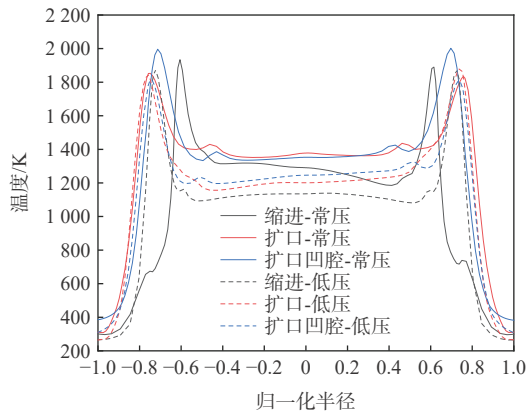


图8 环境压力对火焰出口温度的影响

Fig. 8 Effect of ambient pressure on flame temperature

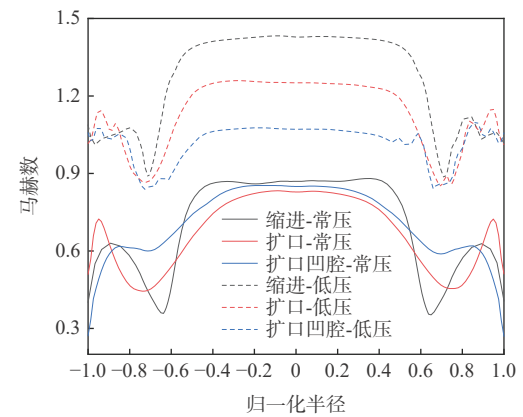


图9 环境压力对火焰出口速度的影响

Fig. 9 Effect of ambient pressure on flame velocity

(环境压力为0.01 MPa)下3种稳焰结构在设计工况(中心混合比为20，推进剂入口氢温度为250 K，氧温度为91 K)的出口燃气温度及速度，图中，归一化半径为燃气出口任选一条直径的径向坐标与引火管内径之商。

常压环境下，扩口及扩口凹腔结构出口燃气中心温度相较缩进结构最高可提高约230 K，高温区域增大21%，燃气均匀性得到改善；扩口结构不会破坏冷却气膜，故而出口壁面温度与缩进结构基本

一致，扩口凹腔结构虽然出口壁面温度升高 100 K，但远低于烧蚀温度，对冷却效果影响不大。常压环境下出口燃气速度均小于声速，相比于缩进结构，扩口及扩口凹腔结构中心燃气降速不明显；低压环境下，扩口凹腔结构相较于缩进结构对于点火器出口火焰稳定性改善更为明显，在扩口及凹腔结构共同作用下掺混区内流速由超声速降至亚声速，燃气停留时间延长，与冷却氢掺混补燃更为充分，点火器出口中心温度最高提升约 250 K，中心燃气马赫数从 1.4 降至 1.05。扩口凹腔稳焰结构在提高中心燃气温度同时降低了射流速度，有助于增强点火燃气射流与下游燃烧装置推进剂的掺混燃烧，可显著提高点火器在低压环境下的点火稳定性。

3.2　中心混合比影响

低压环境下，扩口凹腔结构对点火器出口火焰稳定性改善最为显著，图 10 和图 11 进一步讨论了扩口凹腔相较于缩进结构在不同中心混合比（10、20、40）工况下的火焰稳定性（环境压力为 0.01 MPa，推进剂入口氢温度为 250 K，氧温度为 91 K），计算中氧流量及冷却氢流量不变，改变中心氢流量进行分析。

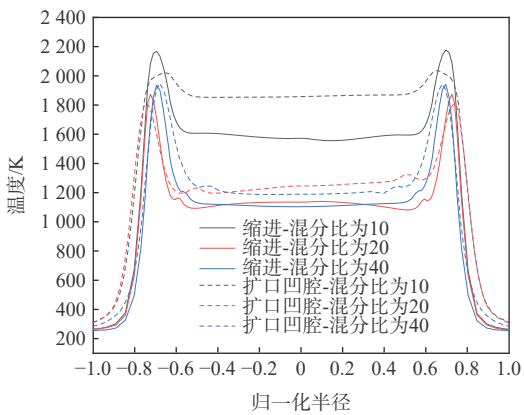


图 10　中心混合比对火焰出口温度的影响

Fig. 10　Effect of central mixture ratio on flame temperature

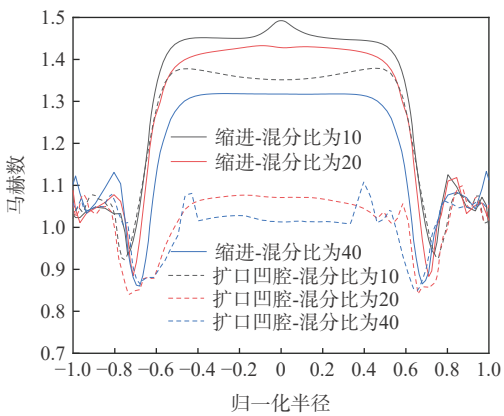


图 11　中心混合比对火焰出口速度的影响

Fig. 11　Effect of central mixture ratio on flame velocity

混合比为 10 工况理论燃气总温为 3 126 K，扩口凹腔结构相较于缩进结构出口马赫数降低了 0.15，中心燃气与冷却氢的补燃增强效果显著，燃气温度从 1 600 K 提升至 1 900 K，但壁面温度仅上升 50 K，凹腔结构未破坏壁面的冷却气膜；混合比为 20 和混合比为 40 工况理论燃气总温分别为 2 840 K 和 2 275 K，二者相差 565 K，但 2 种火焰稳定结构下燃气出口温度相差均小于 50 K，主要由于混合比为 40 工况下，虽然中心燃气温度低，但氧含量高提升了补燃区温度，掺混后缩小了温度差异。缩进结构的中心燃气马赫数均大于 1.3，扩口凹腔结构将出口流速降至声速附近，混合比为 40 工况下中心燃气温度提升至 1 200 K，充分保证了点火器在宽工况范围下的火焰稳定性。

3.3　推进剂入口温度影响

火炬式电点火器推进剂如果从推力室或喷管换热后供应，其推进剂入口温度变化范围可近 100 K，图 12 和图 13 研究了在低压（环境压力为 0.01 MPa）、大富氧（混合比为 40）、氢氧入口温度大幅变化极

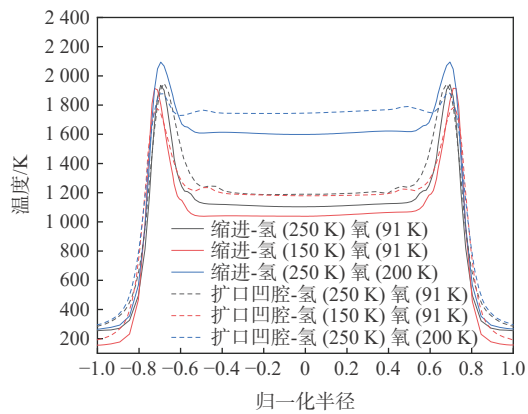


图 12　推进剂入口温度对火焰出口温度的影响

Fig. 12　Effect of propellant inlet temperature on flame temperature

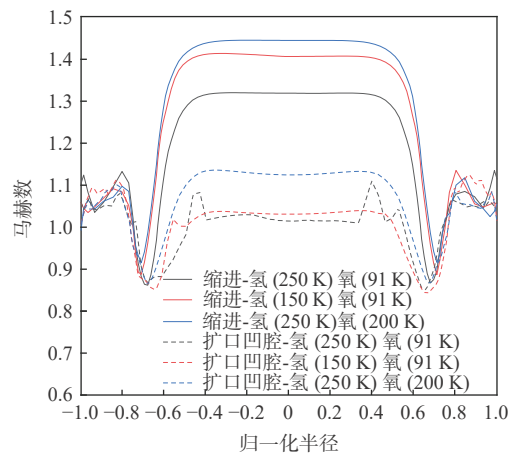


图 13　推进剂入口温度对火焰出口速度的影响

Fig. 13　Effect of propellant inlet temperature on flame velocity

限工况下的点火器出口火焰稳定性。

混合比不变情况下，由于中心燃气为富氧状态，氧入口温度升高点火器出口燃气温度大幅提高，在扩口凹腔结构增强补燃降速作用下中心火焰温度可达1 800 K，马赫数由声速加速至1.1；冷氢工况下，氢入口温度降低至150 K，与设计工况（氢250 K，氧91 K）相比，缩进结构中心燃气温度降至1 000 K，与设计工况相差超过100 K；扩口凹腔结构仅补燃区域略有减小，中心燃气温度及速度基本保持不变，保证在1 200 K以上，相较于缩进结构高温区域扩张了10.6%，出口马赫数由1.4降至声速，有效降低了氢入口温度降低可能引发的点火能量不足风险。

4 结 论

本文研究了缩进、扩口及扩口凹腔3种稳焰结构对富氧火炬式电点火器极限工况下出口火焰稳定性的影响，结果表明，扩口凹腔稳焰结构效果最为显著。

1）常压环境下，扩口及扩口凹腔结构均可改善燃气均匀性；低压环境下，扩口凹腔稳焰结构在提高中心燃气温度同时降低了射流速度，出口燃气中心温度最高提升约250 K。

2）低压大富氧工况（混合比为40）下，扩口凹腔结构通过增强中心燃气与冷却氢的掺混燃烧保证中心燃气温度与设计工况（混合比为20）出口温度相差小于50 K，提高点火器对宽工况的适应性。

3）低压大富氧冷氢极限工况下，扩口凹腔结构维持点火器火焰出口中心燃气温度在1 200 K以上，出口马赫数为声速，与设计工况基本一致，钝化了冷氢入口对点火器火焰出口稳定性影响。

本文将航空发动机与超燃冲压发动机的凹腔稳焰技术引入富氧火炬式电点火器，可显著改善点火器在低压宽工况环境下的火焰稳定性，提高点火能量，保证发动机的点火可靠性，对液体火箭发动机的真空点火技术具有一定的参考意义。

参考文献（References）

[1] VAN LERBERGHE W M, EMDEE J L, FOUST R R. Enhanced reliability features of the RL10E-1 engine[J]. Acta Astronautica, 1997, 41(4-10): 197-207.

[2] LIANG P Y. Modeling of SSME fuel preburner ASI[C]//Proceedings of the 10th Workshop for Computational Fluid Dynamic Applications in Rocket Propulsion. Washington, D. C. : NASA, 1992: 1013-1031.

[3] FUKUSHIMA Y, LMOTO T. Lessons learned in the development of the LE-5 and LE-7[C]//Proceedings of the 30th Joint Propulsion Conference and Exhibit. Reston: AIAA, 1994.

[4] 吴宏斌. 国外氢氧火箭发动机的火药点火器和电点火器[J]. 导弹与航天运载技术, 1994(5): 32-42.
WU H B. Pyrotechnic igniters and electrical spark igniters for LOX/LH_2 rocket engines[J]. Missiles and Space Vehicles, 1994(5): 32-42(in Chinese).

[5] 郑大勇, 陶瑞峰, 胡骏. 氢氧发动机真空点火及高空模拟试验[J]. 导弹与航天运载技术, 2014(5): 38-43.
ZHENG D Y, TAO R F, HU J. Vacuum ignition characteristics and altitude simulation testing of the cryogenic LOX/LH_2 rocket engine[J]. Missiles and Space Vehicles, 2014(5): 38-43(in Chinese).

[6] WANG Y, YANG J H, CAI G B, et al. Effects of injection distance and droplet size on an LH2/LOX torch igniter[J]. Aerospace Science and Technology, 2022, 121: 107334.

[7] 杨进慧, 王朝晖, 左安军, 等. 氢氧火炬式电点火器燃烧流动分析[J]. 导弹与航天运载技术, 2019(3): 45-48.
YANG J H, WANG Z H, ZUO A J, et al. Combustion flow analysis of the hydrogen-oxygen torch igniter[J]. Missiles and Space Vehicles, 2019(3): 45-48(in Chinese).

[8] 王铁岩, 郑孟伟, 蔡振宇. 气氧甲烷火炬式点火器稳态仿真分析[J]. 导弹与航天运载技术, 2016(4): 63-66.
WANG T Y, ZHENG M W, CAI Z Y. Steady-state simulation of GOX/methane spark torch igniter[J]. Missiles and Space Vehicles, 2016(4): 63-66(in Chinese).

[9] 张锋, 于涵, 尚帅, 等. 气氧/气甲烷火炬式电点火器燃烧仿真与热试[J]. 火箭推进, 2022, 48(4): 36-42.
ZHANG F, YU H, SHANG S, et al. Combustion simulation and hot-fire test of GOX/GCH_4 spark torch igniter[J]. Journal of Rocket Propulsion, 2022, 48(4): 36-42(in Chinese).

[10] 孙纪国, 王珏. 高混合比火炬式电点火器试验研究[J]. 推进技术, 2000, 21(1): 33-35.
SUN J G, WANG J. Experimental study on high mixture ratio torch ignitor[J]. Journal of Propulsion Technology, 2000, 21(1): 33-35(in Chinese).

[11] WU H, CHEN Q, SHAO W W, et al. Combustion of hydrogen in an experimental trapped vortex combustor[J]. Journal of Thermal Science, 2009, 18(3): 256-261.

[12] CHEN S, ZHAO D. Numerical study of non-reacting flowfields of a swirling trapped vortex ramjet combustor[J]. Aerospace Science and Technology, 2018, 74: 81-92.

[13] LI M Y, HE X M, ZHAO Y L, et al. Effect of strut length on combustion performance of a trapped vortex combustor[J]. Aerospace Science and Technology, 2018, 76: 204-216.

[14] CAI Z, WANG T Y, SUN M B. Review of cavity ignition in supersonic flows[J]. Acta Astronautica, 2019, 165: 268-286.

[15] 马文杰, 孙明波, 邵文清, 等. 基于Driscoll凹腔稳焰模型的超声速燃烧火焰稳定尺度效应研究[J]. 推进技术, 2021, 42(8): 1865-1875.
MA W J, SUN M B, SHAO W Q, et al. Scaling effect of supersonic combustion flame stabilization based on driscoll cavity blowout limits model[J]. Journal of Propulsion Technology, 2021, 42(8): 1865-1875(in Chinese).

[16] 陈兴良, 景婷婷, 朱韶华, 等. 支板/凹腔组合稳焰器耦合机制研究[J]. 推进技术, 2022, 43(11): 198-207.
CHEN X L, JING T T, ZHU S H, et al. Coupling mechanism of strut/cavity combined flame stabilizer[J]. Journal of Propulsion

Technology, 2022, 43(11): 198-207(in Chinese).

[17] SHI L, YANG X, YANG Y Y, et al. Experimental study on rocket jet-driven ignition and scramjet combustion in a kerosene-fueled RBCC combustor[J]. Aerospace Science and Technology, 2022, 126: 107643.

[18] 杨揖心. 后缘突扩型凹腔超声速流动模式与稳焰机理研究[D]. 长沙: 国防科技大学, 2018.
YANG Y X. Study on supersonic flow pattern and flame stabilization mechanism of cavity with sudden expansion of trailing edge[D]. Changsha: National University of Defense Technology, 2018(in Chinese).

[19] 丁猛. 基于凹腔的超声速燃烧火焰稳定技术研究[D]. 长沙: 国防科学技术大学, 2005.
DING M. Study on flame stabilization technology of supersonic combustion based on concave cavity[D]. Changsha: National University of Defense Technology, 2005(in Chinese).

Study on flame stability of oxygen-rich torch igniter

YANG Jinhui*，WANG Xijie，XU Shiyang，WANG Xiaoli

(Beijing Aerospace Propulsion Institute，Beijing 100076，China)

Abstract: Flame instability appeared under wide extreme sub-atmospheric conditions with varied inlet propellant temperatures, manifested as inadequate ignition energy, uneven gas temperatures, and excessive gas velocity. This paper investigated three flame stability methods: recess, flared outlet and flared cavity, through experiments and simulations. The flared cavity directs the central gas to mix and react with the cooling hydrogen, which also slows down the mixture and enhances reburning in the recirculation zone. Compared to the recess igniter under extreme conditions, the center gas temperature of the flared cavity increases from 1 000 K to 1 200 K. The temperature difference from the design point decreases from 100 K to 50 K, and the outlet Mach number decelerates from 1.4 to sonic. The flared cavity significantly improves the flame stability of the oxygen-rich torch igniter.

Keywords: igniter; extreme condition; flame stability; recess; flared outlet; flared cavity outlet

Received：2023-08-17；**Accepted**：2023-11-03；**Published Online**：2023-11-14 09：35

URL：link.cnki.net/urlid/11.2625.V.20231113.0937.001

Foundation item：Research Projects of China Aerospace Science and Technology Corporation

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0564

CAD/CAM 模型驱动的型腔特征自适应数控工艺设计方法

徐昌鸿[1,*]，黄瑞[2]，张树生[3]，梁嘉宸[3]，卞荣[1]

(1. 南京工程学院 工程训练中心、应用技术学院，南京 211167；2. 河海大学 人工智能与自动化学院，常州 213022；
3. 西北工业大学 机电学院，西安 710072)

摘 要：为实现对已有数控(NC)工艺设计成果的有效重用，提出一种计算机辅助设计/计算机辅助制造(CAD/CAM)模型驱动的型腔特征自适应数控工艺设计方法。基于中轴变换(MAT)提取型腔特征三维CAD模型中隐藏的高层次数控工艺信息，有效表征刀具切削过程；针对已有特征实例进行可重用性分析，构建一个融合多属性数控工艺要素的特征相似性评价模型；基于检索返回的相似特征实例，自适应重用其关联的CAM模型工艺参数生成新特征的数控工艺。实验结果表明：所提方法能够自动、高效地获得特征数控工艺，提升工艺设计效率85%以上，满足企业工艺设计智能化的工程需求。

关 键 词：数控工艺；重用；型腔特征；自适应；相似性
中图分类号：TP391
文献标志码：A **文章编号**：1001-5965(2025)10-3374-11

随着基于模型的定义(model based definition, MBD)技术在工业界的应用愈加广泛，零件三维计算机辅助设计(computer aided design, CAD)、CAD模型及其关联的计算机辅助制造(computer aided manufacturing, CAM)模型文件已经成为各制造企业进行产品设计制造的主要数据载体，不断生成并存储在企业的数据库中[1]。虽然企业产品在不断推陈出新，但是80%的新特征数控工艺都可以通过重用已有相似CAM模型的工艺数据得到[2-3]。因此，如何快速、准确地查找与新特征高度相似的特征实例(即数据库中已存储的制造特征及其关联的CAD/CAM模型)，并重用其CAM模型携带的工艺数据，进而高效地获得新特征数控工艺，对于提升企业创新能力、缩短研制周期具有极其重要的作用[4]。

近30年来，基于制造特征的三维CAD模型检索技术的出现，为这一难题提供了有效的解决途径[5-7]。其核心是构建包含典型制造数据(如尺寸公差、拓扑结构、精度要求等)的特征相似度评价模型，进行新特征和已有实例的相似性精细化度量[8]。通过交互式/半自动式修订检索到的相似实例工艺参数，生成新特征的数控工艺。Zhang等[9]提出一种基于集成产品信息模型的相似制造特征检索方法。首先，对比2个CAD模型制造特征的非几何因素相似性，包括特征类型、加工资源、制造方法等；然后，计算模型整体几何形状、空间方向、特征分布等几何因素相似性。Li等[10-11]提出一种三维CAD模型的层次化表征方法。该方法将STEP格式表示的三维CAD模型表征为层次化特征依赖图，较好地保留了模型的层次特性，通过图的分割、聚类、匹配算法，实现制造特征/零件整体粒度的相似度计

收稿日期：2023-09-01；**录用日期**：2024-01-10；**网络出版时间**：2024-01-15 21:13
网络出版地址：link.cnki.net/urlid/11.2625.V.20240115.1106.001
基金项目：国家自然科学基金(52075148)；江苏省高等学校基础科学(自然科学)研究项目(23KJB460017)；南京工程学院高层次引进人才科研启动基金资助项目(YKJ202128)
* **通信作者**. E-mail：xuchanghonglz@163.com

引用格式：徐昌鸿，黄瑞，张树生，等. CAD/CAM模型驱动的型腔特征自适应数控工艺设计方法[J]. 北京航空航天大学学报，2025，51(10)：3374-3384. XU C H，HUANG R，ZHANG S S，et al. Adaptive NC process planning approach for pocket features driven by CAD/CAM models[J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51(10)：3374-3384 (in Chinese).

算和相似实例重用。Hong 等[12]提出一种基于 B-rep 表示的三维 CAD 模型多粒度相似度计算方法。首先, 基于形状分布图粗略计算两模型的几何形状相似度, 然后, 考虑切削体、加工面数量、特征数量等因素, 进行各匹配制造特征相似度的精细度量, 最后, 综合计算实现三维 CAD 模型相似性评价。Hou 等[13]为提升注塑模具抽芯机构三维 CAD 模型设计效率, 缩短研制时间, 提出一种层次化相似特征检索方法。该方法分别计算抽芯机构的特征类别、作用力、工作距离和外形分布等信息, 构建三层次相似度计算模型, 并基于模糊理论分配不同层次的权重。然而, 现有方法主要考虑 CAD 模型中显式的、表层的特征信息, 忽略了与实际加工过程密切相关的制造信息(如刀具选取、刀具轨迹、加工策略、切削面积等)对于工艺重用的价值, 导致检索结果主要适用于对已有设计模型资源的重用, 而无法支撑数控工艺的有效重用。实际加工中, 几何、拓扑相似并不意味着工艺相似。形状完全相同的 2 个特征, 如果其加工精度不一样, 其数控工艺可能存在较大差异; 而形状拓扑相差较大的 2 个特征, 其数控工艺却可能存在很强的可重用性[14]。因此, 现有方法检索得到的相似特征, 其 CAM 模型工艺数据并不能很好地适用于新特征, 仍然需要工艺人员进一步的编辑修正, 效率低、精度差, 无法在制造业得到切实的落地应用。

为解决这一问题, 本文以 2.5D 型腔特征为研究对象, 提出一种 CAD/CAM 模型驱动的型腔特征自适应数控工艺设计方法。首先, 计算型腔特征轮廓曲线图的中轴并将其作为刀具轨迹, 结合刀具切削运动过程分析提取 CAD 模型中隐式存在的高层次数控工艺信息; 然后, 从工艺重用角度出发, 构建一个融合多属性数控工艺信息的特征相似性评价模型, 检索得到相似特征实例; 最后, 通过自适应重用相似特征实例 CAM 模型工艺数据, 自动生成新特征数控工艺。

1　基本概念与本文方法概述

1.1　基本概念

定义 1　CAM 模型。CAM 模型是指基于零件三维 CAD 模型采用 CAM 软件编制的工艺文件, 如 CATIA 软件中的.CATProcess 文件、UG 软件中的.prt 文件等, 由多个有序加工操作 O_{P} 组成, 每个 O_{P} 关联一个制造特征 F_i 的 CAD 模型 $F_{i\mathrm{CAD}}$, 包括一系列重要的数控工艺数据, 表示为

$$\begin{cases} <\mathrm{CAM}> ::= \bigcup_{i=1}^{p} O_{\mathrm{P}i} \\ <O_{\mathrm{P}i}> ::= F_{i\mathrm{CAD}} <O_{\mathrm{t}}> \cup <M_{\mathrm{s}}> \cup \\ \qquad\qquad <M> \cup <T> \cup <\mathrm{TAD}> \cdots \end{cases} \tag{1}$$

式中: O_{t} 表示加工操作类型; M_{s} 表示加工阶段; M 表示加工机床; T 表示加工刀具; TAD 表示进刀方向。

图 1 为在 CATIA 铣削加工模块中设计的 CAM 模型, 包含多个加工操作。对于加工操作 Roughing_Pocketing_F16, O_{t} 为型腔铣(Pocketing), M_{s} 为粗加工(Roughing), 加工机床 M 为 3 轴铣削机床, 加工刀具为 D32 端铣刀, 进刀方向为轴向进刀, 关联制造特征 $F_{16\mathrm{CAD}}$。

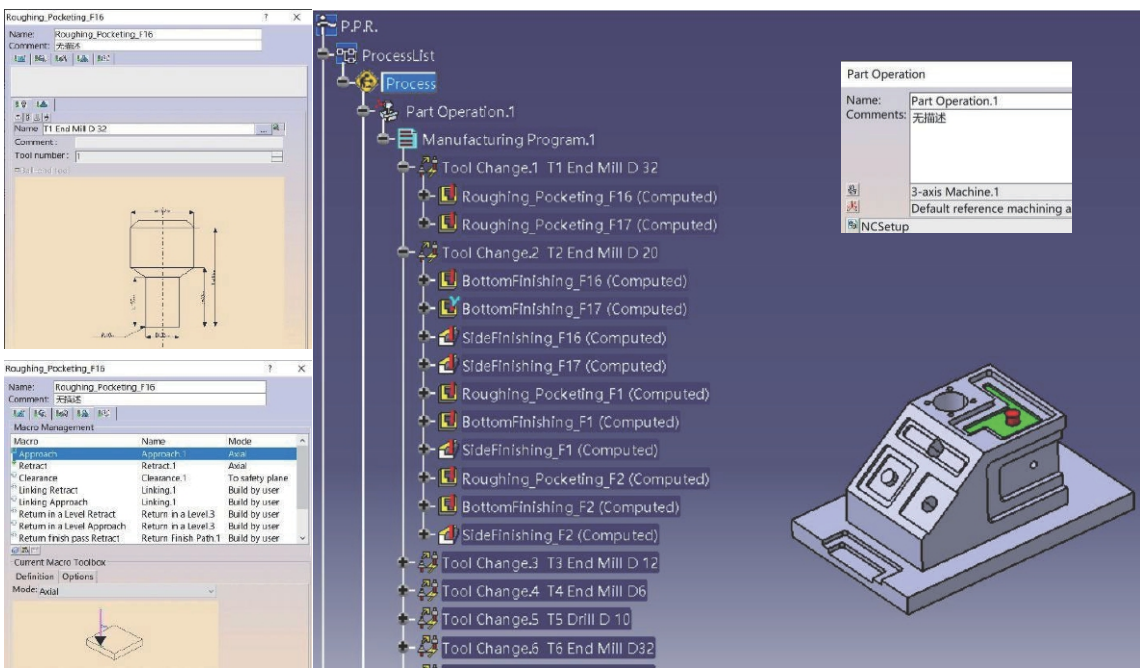


图 1　CATIA 软件中的 CAM 模型

Fig. 1　CAM model in CATIA software

定义 2 轮廓曲线图。型腔特征 F 的轮廓曲线图 Q 反映了特征的轮廓形状及分布，由一系列二维轮廓点组成，是中轴计算要求的数据格式[15]，主要包括型腔特征内轮廓曲线 Q_1 和凸台特征外轮廓曲线 Q_2，表示为

$$\begin{cases} <Q>::=<Q_1>\cup<Q_2> \\ <Q_2>::=\bigcup_{i=1}^{q} Q_2(B_i) \end{cases} \tag{2}$$

式中：B_i 为型腔特征内的第 i 个凸台；$Q_2(B_i)$ 为第 i 个凸台的外轮廓曲线；q 为凸台特征数量。

图 2(a)为型腔特征 F，B_1 和 B_2 为 2 个凸台特征。图 2(b)中蓝色线为轮廓曲线，包括 Q_1、$Q_2(B_1)$ 和 $Q_2(B_2)$，红色线为中轴线。

定义 3 候选刀具集(candidate tools set, CTS)。特征 F 的候选刀具集 F_{CTS} 由满足 F 粗加工要求的所有刀具按一定顺序(刀具切削直径递减)组成，表示为

$$<F_{CTS}>::=\bigcup_{i=1}^{M} D_i|D_i>D_{i+1} \tag{3}$$

式中：M 为刀具数量；D_i 为刀具切削直径。

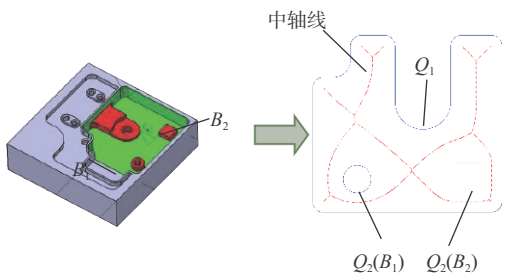


图 2 型腔特征的轮廓曲线

Fig. 2 Boundary curves of pocket feature

1.2 本文方法概述

图 3 为本文方法的总体流程，主要包括三维 CAD 模型与数控加工关联分析、面向加工过程的相似特征实例检索、CAM 模型驱动的数控工艺自适应重用 3 部分，具体地：

1) CAD 模型与数控加工关联分析。该部分参考文献 [16] 中特征识别算法获得三维 CAD 模型中的特征及其轮廓曲线图，此部分不做详细描述。计算特征的中轴变换表征刀具轨迹，依据数控加工规则提取 CAD 模型中隐藏的数控工艺信息，包括刀轨长度、退刀次数和切削面积等。

2) 面向加工过程的相似特征实例检索。该部分基于提取的数控工艺信息，针对数据库中已有特征实例进行可重用性分析，构建多属性工艺要素融合的特征相似度评价模型，实现面向数控加工过程的特征相似性精细度量与快速检索。

3) CAM 模型驱动的数控工艺自适应重用。该部分基于检索得到的相似特征实例，综合考虑现实生产环境的不确定性和动态性，提出数控工艺自适应重用机制，自适应修改相似实例 CAM 模型的工

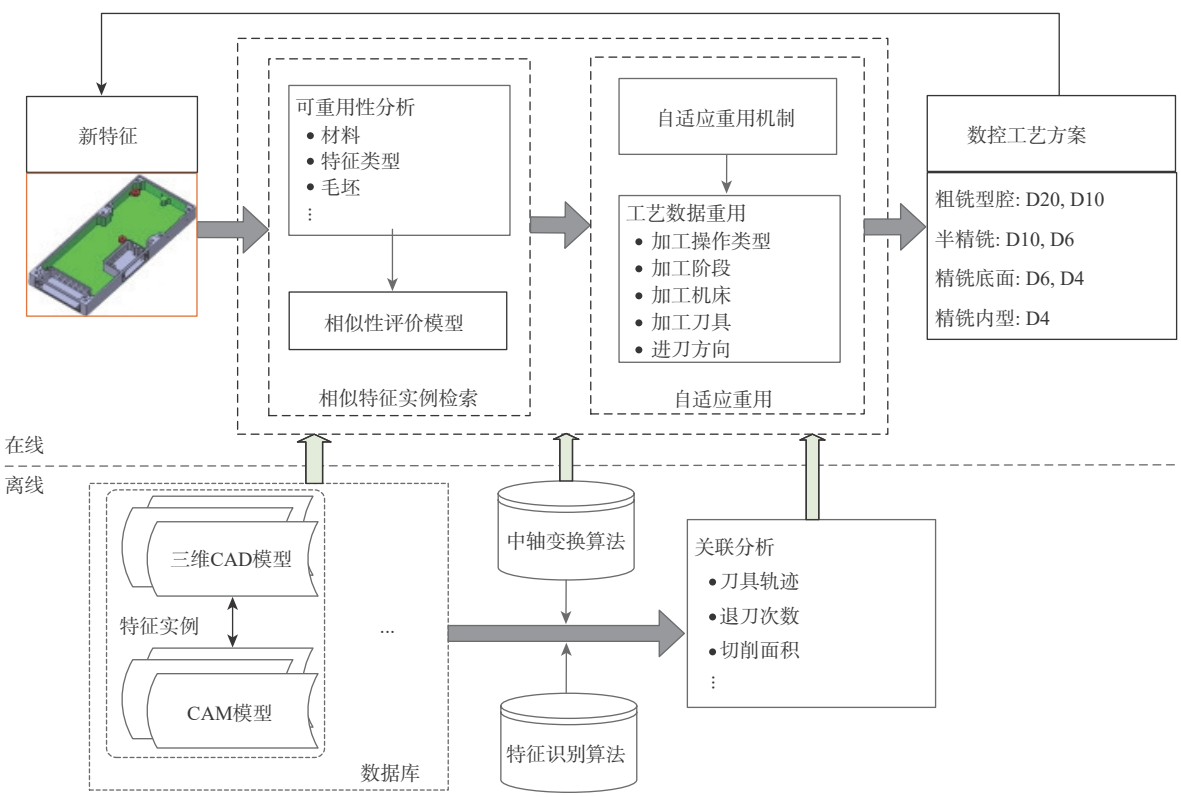


图 3 本文方法总体流程

Fig. 3 Overview of the proposed method

艺参数, 使之适用于新特征。

2　CAD 模型与数控加工关联分析

由于中轴变换具有拓扑对称性、均等分割性等特点, 已经被广泛采用作为刀具轨迹的有效参考[17]。因此, 本文采用文献 [18] 中轴变换算法, 计算型腔特征的中轴作为刀具轨迹的有效表征, 分析提取其中隐式存在的数控工艺信息, 实现三维 CAD 模型与加工过程的紧密关联。

2.1　候选刀具集计算

由文献 [18] 可知, 中轴线是由一系列连续的、与特征轮廓曲线相切的圆, 其圆心按顺序连接形成, 各圆心坐标 $O(x_i,y_i)$ 及直径 d_i 易于获得, 其中, 直径最大值和最小值分别为 $d_{\max}$ 和 $d_{\min}$。在加工过程中, 刀具沿中轴进行切削, 若刀具半径 $D_i>d_{\max}-2\delta$ (δ 为粗加工余量), 则刀具将与特征轮廓干涉过切, 影响零件质量。另一方面, 若 $D_i\leqslant d_{\min}-2\delta$, 虽然刀具能够沿中轴完整加工整个特征区域, 但是 D_i 过小会增加额外的切削时间, 影响加工效率。

综上, 特征 F 的候选刀具集 F_{CTS} 应位于区间 $[d_{\max}-2\delta, d_{\min}-2\delta]$ 范围内。由于刀具直径通常为整数(小数情况较少, 本文暂不考虑), 因此, F_{CTS} 定义为

$$\begin{cases} F_{\mathrm{CTS}}=\{D_i \mid D_i=D_{\max}-i \quad i=0,1,\cdots,(D_{\max}-D_{\min}) \\ D_{\max}=\lfloor d_{\max}-2\delta \rfloor \\ D_{\min}=\lfloor d_{min}-2\delta \rfloor \end{cases} \tag{4}$$

式中: $D_{\max}$ 和 $D_{\min}$ 分别为最大和最小候选刀具直径; $\lfloor d_{\max}-2\delta \rfloor$和$\lfloor d_{\min}-2\delta \rfloor$分别为不大于 $d_{\max}-2\delta$ 和 $d_{\min}-2\delta$ 的最大整数。

2.2　数控工艺信息提取

2.2.1　刀轨长度

对于一把可行粗加工刀具 $D_i\in F_{\mathrm{CTS}}$, 其切削特征 F 时的刀具轨迹为若干段连续中轴曲线, 满足条件 $d_i\geqslant D_i-2\delta$(d_i 为此段中轴上的相切圆直径), 则刀轨长度 $L(F,D_i)$ 计算式为

$$\begin{cases} L(F,D_i)=\sum_{n=1}^{N} L_n \\ L_n=\sum \sqrt{(x_{i+1}-x_i)^2+(y_{i+1}-y_i)^2} \end{cases} \tag{5}$$

式中: N 为刀具轨迹对应的中轴线数量; (x_i,y_i) 为中轴线上的圆心坐标。

2.2.2　退刀次数

由于刀具运动轨迹由多段中轴线组成, 刀具在切削运动过程中沿一条中轴线完成加工后, 需要进行退刀操作移动到下一段刀轨继续加工。因此, 刀具 D_i 加工 F 时的退刀次数 $R(F,D_i)$ 等于刀具轨迹中轴线的数量。

2.2.3　切削面积

加工过程中的材料去除量是衡量加工效率的一个重要因素。由于 2.5D 型腔特征内部高度相同, 因此, 本文计算切削面积作为材料去除量的有效表征。刀具 D_i 加工 F 时的切削面积 $S(F,D_i)$ 可以看做以中轴相切圆圆心为运动轨迹, 以切削直径 D_i 为圆, 运动过程中覆盖的区域面积, 如图 4 所示, 通过 OpenGL 函数接口计算可得[19], 计算式为

$$\begin{cases} S(F,D_i)=\sum_{j=1}^{R(F,D_i)} S_j(F,D_i) \\ S_j(F,D_i)=f(P_j(F,D_i),D_i,\delta) \end{cases} \tag{6}$$

式中: $P_j(F,D_i)$ 为刀具 D_i 加工 F 时的第 j 段刀具轨迹中轴线; $S_j(F,D_i)$ 为 D_i 沿 $P_j(F,D_i)$ 加工时的切削面积。

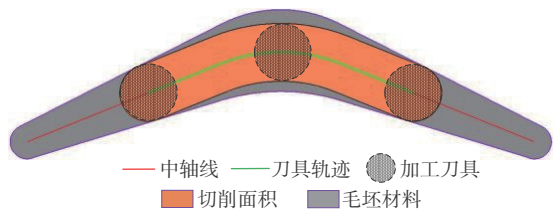


图 4　切削面积计算示意图

Fig. 4　Schematic diagram of computing cutting area calculation

3　面向加工过程的相似特征实例检索

3.1　特征实例可重用性分析

2.2 节提取的数控工艺信息反映了加工过程中的本质内涵, 对于工艺重用要求下的特征相似性计算具有重要的参考价值。为提高运算效率, 本文在相似性计算之前先对数据库中的特征实例进行可重用性分析, 如表 1 所示, 快速筛选排除不满足工艺重用要求的实例, 分为 3 步:

步骤 1　判断新特征 F_{A} 和特征实例 F_{B} 的材料、特征类型、毛坯类型等是否都相同, 如果相同则进行后续判断, 否则排除 F_{B};

步骤 2　判断 F_{B} 的加工精度是否不低于 F_{A}, 如

表 1　可重用性分析因素

Table 1　Elements of reusability analysis

因素	内容	可重用要求
材料M_{aterial}	GGr15、45钢…	$M_{\mathrm{aterial}}(F_{\mathrm{A}})=M_{\mathrm{aterial}}(F_{\mathrm{B}})$
特征类型F_{eature}	型腔、孔、凸台…	$F_{\mathrm{eature}}(F_{\mathrm{A}})=F_{\mathrm{eature}}(F_{\mathrm{B}})$
毛坯类型B_{lank}	铸件、锻件、机加件…	$B_{\mathrm{lank}}(F_{\mathrm{A}})=B_{\mathrm{lank}}(F_{\mathrm{B}})$
精度要求	尺寸公差等级I_{T}、表面粗糙度R_{a}	$I_{\mathrm{T}}(F_{\mathrm{B}})\leqslant I_{\mathrm{T}}(F_{\mathrm{A}})\wedge R_{\mathrm{a}}(F_{\mathrm{B}})\leqslant R_{\mathrm{a}}(F_{\mathrm{A}})$
刀具要求	粗加工刀具序列、候选刀具集	$T_{\mathrm{R}}(F_{\mathrm{B}})\subseteq F_{\mathrm{ACTS}}$

果满足则进行后续判断，否则排除 F_B；

步骤 3 设 F_B 的 CAM 模型中粗加工刀具序列为 $T_R(F_B)$，只有当 F_B 的粗加工刀具全部在 F_A 的候选刀具集内，F_B 的工艺才可能被 F_A 重用，否则，将发生刀具干涉过切等，不具备重用性。因此，需要判断是否满足 $T_R(F_B)\subseteq F_{ACTS}$，如果满足则进行后续相似性计算，否则，排除 F_B。

通过可重用性分析可以大幅度缩小相似性计算的数据量，有效提升本文方法的检索效率。

3.2 相似性评价模型

针对满足条件的可重用特征实例，从工艺相似性要求出发，构建一个融合多属性数控工艺要素的特征相似性评价模型，如表 2 所示，包括刀具轨迹、退刀次数、切削面积等高层次工艺信息和精度要求等低层次信息。

表 2 多属性数控工艺要素及其内涵

Table 2 Multi-attribute factors of NC process

信息层次	工艺要素	工艺相似性内涵
高层次	刀具轨迹	刀具切削运动
	退刀次数	加工连续性
	切削面积	材料去除量
低层次	精度要求	加工质量

3.2.1 刀具轨迹相似性

刀具轨迹是刀具切削过程的直观描述。若已有特征实例数控工艺可被新特征重用，则二者的刀具轨迹也具有高度的相似性。本文将刀轨长度作为刀具轨迹的有效表征，设特征实例 F_B 的粗加工刀具序列 $T_R(F_B)=\{D_i\}$（$1\leqslant i\leqslant K$，K 为粗刀具数量），则新特征 F_A 和特征实例 F_B 的刀具轨迹相似度 θ_L 计算式为

$$\theta_L=\frac{1}{K}\sum_{i=1}^{K}\left(1-\frac{|L(F_A,D_i)-L(F_B,D_i)|}{L(F_A,D_i)+L(F_B,D_i)}\right) \tag{7}$$

3.2.2 退刀次数相似性

退刀次数反映了加工过程的连续性，其相似性也是判断数控工艺可重用性的一个重要因素。新特征 F_A 和特征实例 F_B 的退刀次数相似度 θ_R 计算式为

$$\theta_R=\frac{1}{K}\sum_{i=1}^{K}\left(1-\frac{|R(F_A,D_i)-R(F_B,D_i)|}{R(F_A,D_i)+R(F_B,D_i)}\right) \tag{8}$$

3.2.3 切削面积相似性

切削面积相似性描述了 F_A 和 F_B 在材料去除量方面的相似程度。若 F_B 的数控工艺可以被 F_A 重用，则二者的切削面积也必然存在一定的相似性。同理，切削面积相似度 θ_A 计算式为

$$\theta_A=\frac{1}{K}\sum_{i=1}^{K}\left(1-\frac{|S(F_A,D_i)-S(F_B,D_i)|}{S(F_A,D_i)+S(F_B,D_i)}\right) \tag{9}$$

3.2.4 精度要求相似性

特征的加工精度要求反映加工质量，对于加工策略（顺/逆铣削、走刀方式等）和切削参数（进给量、进给速度等）的选择具有重要影响，如精度要求低的特征要求切削速度大而进给量和背吃刀量小，而精度要求高的特征则相反。通过 CAM 软件的接口函数，可以得到 F_A 和 F_B 的匹配尺寸公差等级集 $\{(I_{Ti}(F_A),\ I_{Ti}(F_B)),1\leqslant i\leqslant U\}$ 和匹配粗糙度集 $\{(R_{aj}(F_A),\ R_{aj}(F_B)),1\leqslant j\leqslant V\}$。精度要求相似度 θ_P 计算式为

$$\theta_P=\frac{1}{2}\left[\frac{1}{U}\sum_{i=1}^{U}\left(1-\frac{|I_{Ti}(F_A)-I_{Ti}(F_B)|}{I_{Ti}(F_A)+I_{Ti}(F_B)}\right)+\frac{1}{V}\sum_{j=1}^{V}\left(1-\frac{|R_{aj}(F_A)-R_{aj}(F_B)|}{R_{aj}(F_A)+R_{aj}(F_B)}\right)\right] \tag{10}$$

综上，新特征 F_A 和特征实例 F_B 的相似度 $\psi(F_A,F_B)$ 定义为上述相似度值的加权和，表示为

$$\begin{cases}\psi(F_A,F_B)=\omega_L\theta_L+\omega_R\theta_R+\omega_A\theta_A+\omega_P\theta_P\\ \omega_L+\omega_R+\omega_A+\omega_P=1\end{cases} \tag{11}$$

式中：权重可由工艺人员启发式设定，高层次工艺信息的权重 ω_L、ω_R、ω_A 不低于低层次工艺信息权重 ω_P。

4 CAM 模型驱动的数控工艺自适应重用

由于生产制造环境的动态变化，通过相似性计算返回的最相似特征实例，其 CAM 模型关联的工艺数据可能无法直接被新特征采纳重用。因此，本文提出一种数控工艺自适应重用机制，如图 5 所示，依据生产制造环境及新特征的结构特点和加工要求，采用自适应操作（如添加、删除、更改等）更改相似实例 CAM 模型各加工操作的工艺数据，自动生成新特征的数控工艺。

4.1 加工操作类型和加工阶段

加工操作类型 O_t 由加工特征类型决定，如型腔特征大多采用型腔铣操作，面特征多采用平面铣操作等。由于 F_A 和 F_B 特征类型相同，故 O_t 可直接被重用。同时，加工阶段 M_s 与工艺参数关系密切，如粗加工要求较高的进给率和切削深度，精加工正好相反。因此，只有相同加工阶段的工艺才存在重用可能性，即 F_B 的加工阶段 M_s 可直接被 F_A 重用。

4.2 加工机床

加工机床 M 参数主要包括机床精度、回转直径、转速级数、电机功率、最高/最低转速等。由于 F_A 和 F_B 特征类型相同，且 F_B 的加工精度不低于

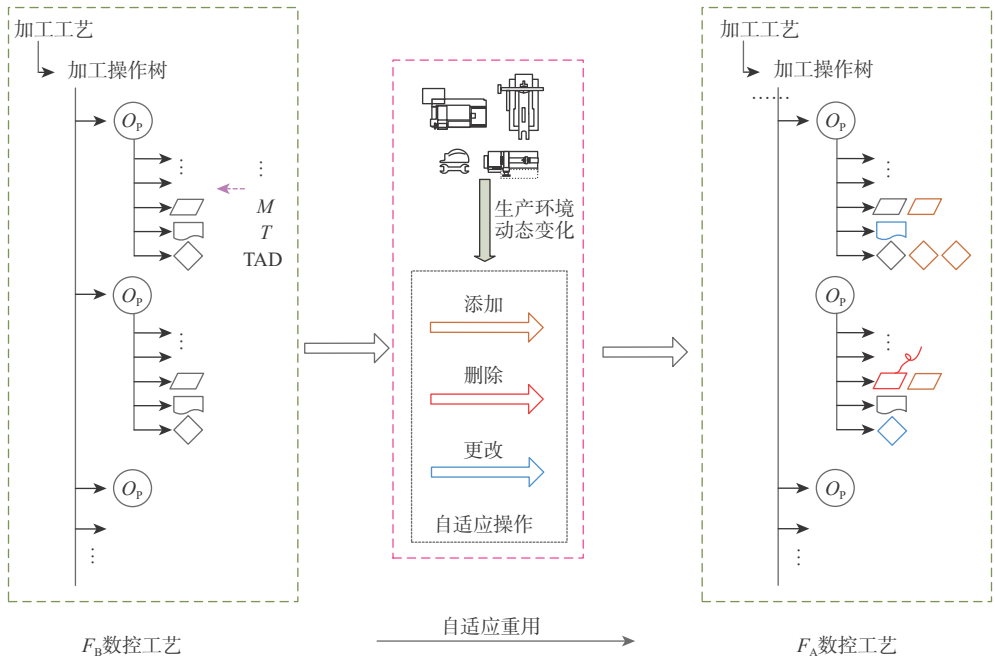


图 5　数控工艺自适应重用机制

Fig. 5　Adaptive reuse mechanism of numerical control process

F_A, 因此, F_B 的加工机床理论上可以被 F_A 直接重用。但是, 由于生产车间处于动态变化过程中, 可能出现某一时间段内 F_B 的加工机床处于不可用阶段(如保养、维修、检查等), 则应依据当前车间中的可用机床资源自适应选取生成 F_A 的加工机床。特别地, 若当前存在多个机床可满足 F_A 的加工要求, 则自适应添加作为 F_A 的机床资源候选集, 提升工艺多样性及生产加工柔性。

4.3　加工刀具

加工刀具 T 参数主要有刀具名称、刀具类型、刀具编号、切削直径、刃长等。由于 F_B 的粗加工刀具序列 $T_R(F_B)$ 决定了 F_A 和 F_B 的刀具轨迹相似度, 对于相似实例检索具有重要影响, 因此, 具有唯一性。若当前车间刀具库中存在 $T_R(F_B)$, 则直接重用作为 F_A 粗加工刀具序列。类似地, 若 $T_R(F_B)$ 处于不可用状态, 则自适应选择刀具库中相似度最高的粗加工刀具。

若 F_B 精加工刀具序列 $T_F(F_B)$ 在刀具库中处于可用状态, 则 F_A 直接重用其刀具参数生成 $T_F(F_A)$, 若不可用则为 F_A 自适应添加刀具库中所有可用精加工刀具。

4.4　进刀方向

进刀方向 TAD 规定了刀具进入工件进行切削加工的路径方向。由于不同特征实例的 CAD 建模和 CAM 编程的步骤、软件、方法等不尽相同, 因此, F_B 的 TAD 通常不能直接被 F_A 重用, 需要依据 F_A 的拓扑结构和加工要求自适应构建。若 F_A 为封闭型腔, 则自适应计算特征底面法式反方向作为 TAD; 若 F_A 为开型腔, 则自适应构建所有可行 TAD 作为 TAD 候选集。

5　算法验证与讨论

为有效验证本文方法的合理性, 以 Microsoft Visual Studio 2008 为集成开发工具, 在 CATIA CAA 环境中实现了本文方法。同时, 与某大型航空制造企业合作构建了一个包含 500 多个实例的数据库, 对本文方法进行验证测试。

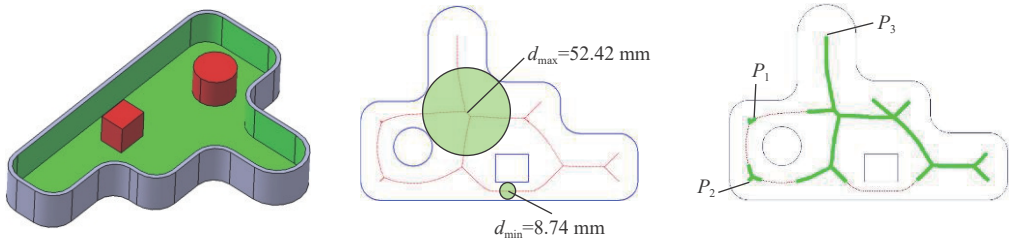
5.1　三维 CAD 模型数控工艺信息提取实例

图 6 为三维 CAD 模型数控工艺信息提取实例。图 6(a)为型腔特征 F, 计算其中轴变换参数可得 d_{max}=52.42 mm, d_{min}=8.74 mm (见图 6(b))。设 δ=1, 则 D_{max}=⌊52.42−2⌋=50 mm, D_{min}=⌊8.74−2⌋=6 mm。依据式 (4), 可得候选刀具集 F_{CTS}=[50,6]。由于 D20∈F_{CTS}, 其粗加工刀具轨迹为图 6(c)中绿色实线, 依据式 (5) 可得刀轨长度 $L(F,\mathrm{D20})$=168.89 mm。可以看出, 由于刀具轨迹由 3 段不连续中轴线 P_1、P_2 和 P_3 组成, 因此, 退刀次数 $R(F,\mathrm{D20})$=3。图 6(d)为 D32 的粗加工刀具轨迹, 共有 2 段中轴线组成, 计算可得其切削面积 $S(F,\mathrm{D32})$=4 809.57 mm^2(见图 6(e)黄色区域)。

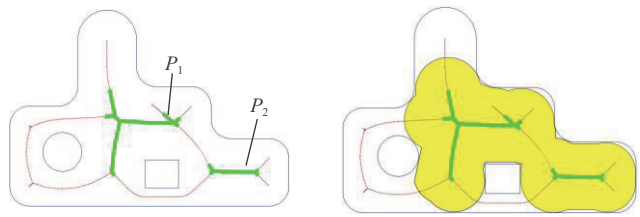
5.2　相似特征检索实例

5.2.1　可重用性分析

如图 7 所示, 新特征 F_A 和特征实例 F_B 的材料(45 钢)、特征类型(型腔特征)、毛坯类型(机加件)都相同, 且存在 4 对匹配尺寸公差 (I_T(12), I_T(11))、(I_T(12), I_T(11))、(I_T(13), I_T(12))、(I_T(14), I_T(12)) 和粗糙



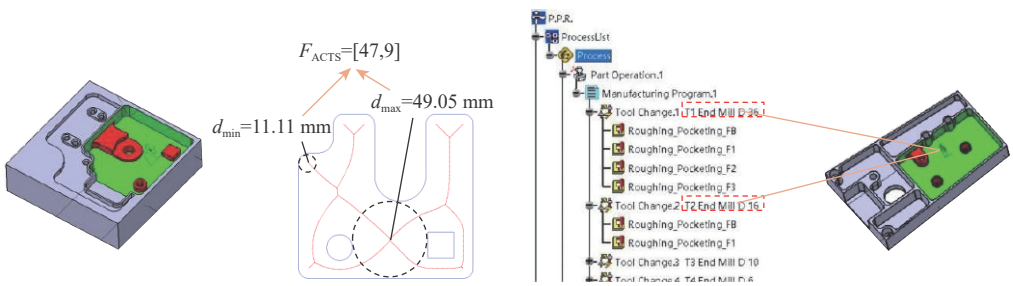
(a) 型腔特征F_1 (b) 中轴变换及其参数 (c) D20刀具轨迹



(d) D32刀具轨迹 (e) D32切削面积

图 6 三维 CAD 模型数控工艺信息提取实例

Fig. 6 Example of NC process extraction case for 3D CAD model



(a) 新特征F_A及其中轴 (b) 特征实例F_BCAM模型

图 7 可重用性分析计算

Fig. 7 Reusability analysis calculations

度 (R_a(6.3), R_a(6.3)), 满足 $I_T(F_B)\leqslant I_T(F_A)\wedge R_a(F_B)\leqslant R_a(F_A)$。通过计算 F_A 的中轴变换参数可得 d_{max}=49.05 mm, d_{min}=11.11 mm, 则依据式 (4) 计算 F_A 的候选刀具集 F_{ACTS}=[47,9]。F_B 的 CAM 模型关联粗加工刀具序列 $T_R(F_B)$={D36,D16}, 满足 $T_R(F_B)\subseteq F_{ACTS}$。因此, F_B 满足工艺可重用性要求, 可以进行后续相似性计算。

5.2.2 特征相似性计算

图 8 为选用 $T_R(F_B)$={D36,D16}分别加工 F_A 和 F_B 时的刀具轨迹, 基于中轴变换可得多属性数控工艺要素计算结果如表 3 所示。本文设定 $\omega_L=\omega_R=\omega_A$=0.3, ω_P=0.1, 则由式 (11) 可得 F_A 和 F_B 相似度 $\psi(F_A, F_B)$=0.86, 表明 F_B 的数控工艺具有较高的重用价值。如果采用传统的 D2 形状分布算法[20], 则二者相似度仅为 0.43, 其数控工艺重用价值无法得到体现。因此, 相比于传统方法, 本文方法能够更加有效、精确地检索满足数控工艺重用实际需求的相似特征实例。

5.3 特征检索实例

针对图 9(a)所示新特征 F_2, 其检索结果如图 9(b)所示, 共返回 5 个相似特征实例, 相似度由高到低排列。虽然实例 1 和实例 2 的几何拓扑、加工精度完全相同, 但是二者粗加工刀具不同, 导致与新特征在刀具轨迹、退刀次数、切削面积等方面存在明显差异, 因此, 整体相似度相差较大。实例 2 和实例 3 的几何形状、粗加工刀具都相同, 但是

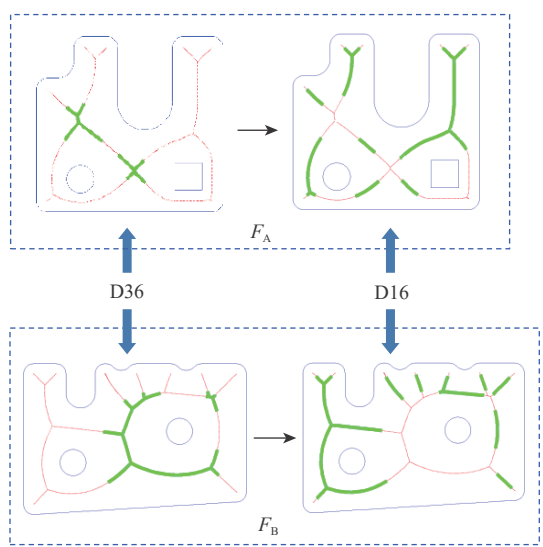


图 8 F_A 和 F_B 刀具轨迹

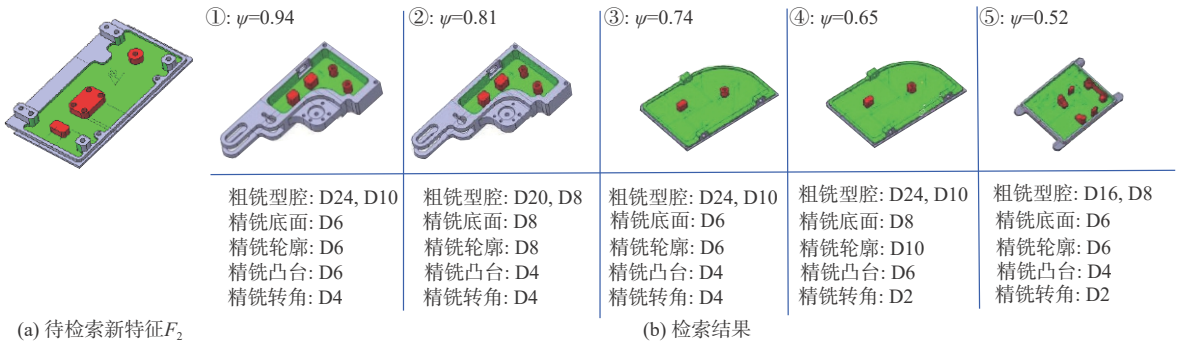
Fig. 8 Tool paths of F_A and F_B

表 3　相似度计算结果

Table 3　Result of similarity calculation

特征	L(D36)/mm	L(D16)/mm	R(D36)	R(D16)	S(D36)/mm^2	S(D16)/mm^2	I_T	R_a	ψ
F_A	73.45	221.98	2	7	3 938.52	2 742.31	(12,11) (12,11)	6.3	0.86
F_B	136.51	323.05	2	8	5 436.13	4 128.95	(13,12) (14,12)	6.3	

注: L(D36), L(D16)为θ_L=0.76, ω_L=0.3时数值; R(D36), R(D16)为θ_R=0.97, ω_R=0.3时数值; S(D36), S(D16)为θ_A=0.82, ω_A=0.3时数值; I_T和R_a为θ_P=0.98, ω_P=0.1时数值。



(a) 待检索新特征F_2　　(b) 检索结果

图 9　特征检索实例

Fig. 9　Retrieval case for features

精度要求与新特征相差较大, 相似度也不相同。实例 5 的刀具轨迹和加工精度都与新特征存在明显不同, 因此, 相似度较低, 其数控工艺可重用价值不高。可以看出, 如果只考虑几何形状, 则无法准确、有效地反映检索结果数控工艺的重用价值。本文方法综合考虑与加工过程密切相关的数控工艺要素, 检索结果对于数控工艺重用具有较好的指导作用。

5.4　数控工艺自适应重用实例

图 10 为相似特征数控工艺自适应重用实例。对于新特征 F_A, 检索返回相似度最高的实例 F_B(相似度 $\psi(F_A,F_B)$=0.92)。首先, 对于 F_A, 直接重用 F_B 的 O_t 和 M_s。由于当前可用加工资源库中存在多台加工机床(M02,M03), 则自适应生成 F_A 的加工机床候选集{M02,M03}。针对 F_B 的粗加工刀具 D32, 由于当前资源库中不存在相应刀具, 则自适应选择 D28 作为 F_A 的粗加工刀具。同理, 对于 F_B 的轮廓精铣刀具 D12, 依据库中刀具情况自适应生成 F_A 的轮廓精铣刀具集{D10,D8,D6}。由于坐标系设定不同, F_A 的进刀方向自适应更改为(1,0,0)。

表 4 为本文方法与交互式工艺设计方法的效率对比。可以看出, 交互式方法需要人工创建 5 个加工操作完成 F_A 的工艺设计, 超过 120 次交互操作, 耗时超过 20 min。然而, 本文方法基于型腔特征自身属性及加工要求, 通过自适应重用相似实例 CAM

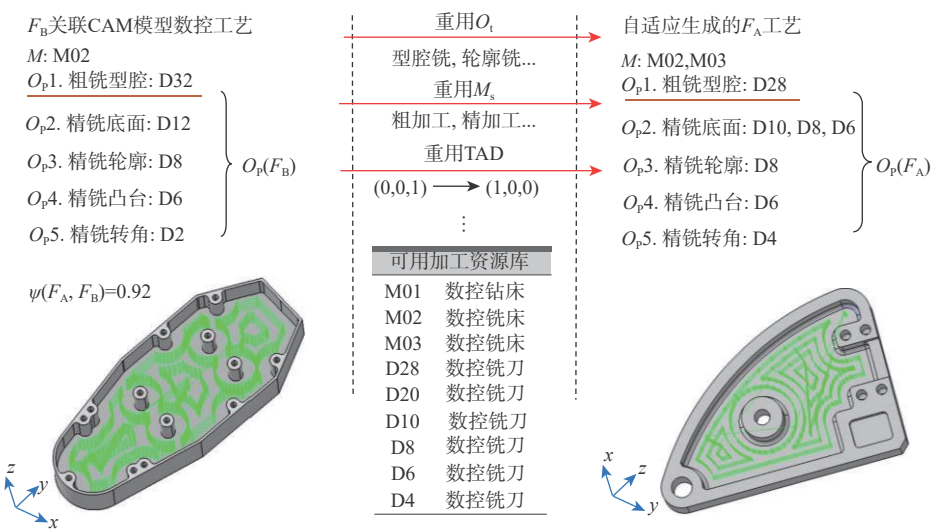


图 10　相似特征数控工艺自适应重用

Fig. 10　Adaptive reuse of NC process of similar feature

表 4 数控工艺设计效率分析

Table 4 Efficiency analysis of NC process planning

方法	加工操作数量	人工交互次数	编程时间/min
交互式方法	5	≥120	≥20
本文方法	0	≤15	≤3

模型关联的数控工艺参数，可自动生成 F_A 的加工操作，仅需少于 15 次人工交互，耗时不高于 3 min，整体数控工艺设计效率提升超过 85%。

5.5 讨 论

表 5 为本文方法与现有典型方法进行对比分析。本文和文献 [21] 都是制造特征粒度的相似实例检索，且本文方法基于中轴变换表征刀具轨迹，结合数控加工规则自动挖掘隐含在三维 CAD 模型中的高层次隐式数控工艺信息，整个过程对工艺人员的交互依赖性小，自动化程度高。同时，由于在相似性计算之前进行可重用性分析，快速筛除不符合重用要求的实例，缩小计算范围，显著提升了检索效率，能够有效支撑数控工艺设计资源的合理重用。文献 [21] 主要面向产品设计重用，相似性计算以几何拓扑关系为主，无法有效支撑数控工艺重用。文献 [22-23] 面向零件整体粒度的检索，但是相似度计算过程中忽视了制造领域相关要素，自动化程度低，难以适用于特征级的数控工艺重用。文献 [24] 主要面向装配工艺重用，对制造语义挖掘不充分。此外，文献 [22-24] 未考虑可重用性分析，需要将新特征与实例库中所有实例进行一一匹配计算，运算量大，检索效率低。

表 5 本文方法与已有方法的比较

Table 5 Comparison of the proposed method and existing approaches

方法	检索粒度	隐式信息挖掘	自动化程度	检索效率	重用领域
本文方法	制造特征	充分	良好	高	数控工艺设计
文献[21]	制造特征	不充分	一般	高	产品设计
文献[22-23]	零件整体		一般	低	
文献[1]	制造特征	不充分	一般	低	装配工艺设计

本文方法主要针对 2.5D 型腔，不涉及复杂曲面。实际加工中，由于复杂曲面型腔特征可以看做 2.5D 型腔的逐层叠加组合。因此，本文方法原理对于需要多轴加工的复杂曲面工艺设计是可行的，需要结合多轴加工刀具及零件运动过程，探索开发适合复杂曲面型腔特征的中轴变换算法及相似性评价模型。

6 结 论

1）本文通过计算 2.5D 型腔特征的中轴变换，有效表征其刀具轨迹，依据中轴关键参数，如相切圆圆心坐标、直径等，结合数控加工规则，实现隐式数控工艺信息的有效挖掘。

2）本文方法在检索计算之前，通过分析材料、特征类型、毛坯类型、加工精度、刀具序列等，判断已有实例数控工艺是否满足可重用要求，从而快速筛除不可用实例，有效缩小计算范围。

3）从实际加工过程角度出发，构建一个融合多属性数控工艺要素的特征相似性评价模型，实现数控工艺可重用实例的有效检索。针对生产制造环境的动态变化，自适应修改相似实例 CAM 模型的加工操作类型、加工阶段、加工机床、刀具、进刀方向等工艺参数，使之适用于新特征。

4）通过多个实例验证了本文方法的实用性，避免了传统方法对工艺要素考虑不足、自动化程度低的不足。

下一步工作考虑结合已有数据挖掘算法，更深入地挖掘探索 CAD 模型中蕴含的其他属性数控工艺要素（如加工顺序、装夹方式等），提升本文方法的实用性；研究适合多轴加工特征的中轴变换算法，实现其数控工艺自适应重用，拓展本文方法的应用范围。

参考文献（References）

[1] 邓乐武, 黄玉璐, 王浩然, 等. 基于 MBD 的飞机测试工艺数字化定义方法[J]. 北京航空航天大学学报, 2021, 47(11): 2181-2188.
DENG L W, HUANG Y L, WANG H R, et al. Digital definition method of aircraft test process based on MBD[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(11): 2181-2188(in Chinese).

[2] SONG I, HAN S. Implementation of the direct integration from CAM to CAE for the PCB simulation[J]. Computers in Industry, 2013, 64(8): 1014-1021.

[3] 黄波, 张树生, 黄瑞, 等. 面向宏观工艺重用的三维 CAD 模型检索方法[J]. 计算机集成制造系统, 2020, 26(12): 3386-3398.
HUANG B, ZHANG S S, HUANG R, et al. Macro process reuse oriented 3D CAD model retrieval method[J]. Computer Integrated Manufacturing Systems, 2020, 26(12): 3386-3398(in Chinese).

[4] ZHANG C, ZHOU G H, LU Q, et al. Graph-based knowledge reuse for supporting knowledge-driven decision-making in new product development[J]. International Journal of Production Research, 2017, 55(23): 7187-7203.

[5] GAO Y, TANG J H, LI H J, et al. View-based 3D model retrieval with probabilistic graph model[J]. Neurocomputing, 2010, 73(10-12): 1900-1905.

[6] HERZOG V, SUWELACK S. Bridging the gap between geometry and user intent: retrieval of CAD models *via* regions of interest[J]. Computer-Aided Design, 2023, 163: 103573.

[7] 杜雨佳, 李海生, 姚春莲, 等. 基于三元组网络的单图三维模型检索[J]. 北京航空航天大学学报, 2020, 46(9): 1691-1700.
DU Y J, LI H S, YAO C L, et al. Monocular image based 3D mod-

el retrieval using triplet network[J]. Journal of Beijing University of Aeronautics and Astronautics, 2020, 46(9): 1691-1700(in Chinese).

[8] 于勇, 顾黎, 印璞, 等. MBD 模型本体建模及检索技术研究与应用[J]. 北京航空航天大学学报, 2017, 43(2): 260-269.
YU Y, GU L, YIN P, et al. Research and implementation of ontology modeling and retrieval technology of MBD model[J]. Journal of Beijing University of Aeronautics and Astronautics, 2017, 43(2): 260-269(in Chinese).

[9] ZHANG R Z, ZHOU X H. Similarity assessment of mechanical parts based on integrated product information model[J]. Journal of Computing and Information Science in Engineering, 2011, 11(1): 1.

[10] LI Z, ZHOU X H, LIU W, et al. A similarity-based reuse system for injection mold design in automotive interior industry[J]. The International Journal of Advanced Manufacturing Technology, 2016, 87(5): 1783-1795.

[11] LI Z, ZHOU X H, LIU W. A geometric reasoning approach to hierarchical representation for B-rep model retrieval[J]. Computer-Aided Design, 2015, 62: 190-202.

[12] HONG T, LEE K, KIM S. Similarity comparison of mechanical parts to reuse existing designs[J]. Computer-Aided Design, 2006, 38(9): 973-984.

[13] HOU B K, PENG J H, HE S, et al. A similarity-based approach for the variant design of core-pulling mechanism in injection mold design[J]. The International Journal of Advanced Manufacturing Technology, 2021, 115(1): 329-344.

[14] 花豹, 周彬, 顾星海, 等. 基于图神经网络的工艺表格语义相似性度量[J]. 计算机集成制造系统, 2022, 28(12): 3805-3821.
HUA B, ZHOU B, GU X H, et al. Semantic similarity measurement of process table based on graph neural network[J]. Computer Integrated Manufacturing Systems, 2022, 28(12): 3805-3821(in Chinese).

[15] 杨梦媛, 李迎光, 许可. 基于中轴变换的型腔高速铣削刀具轨迹生成方法[J]. 航空学报, 2021, 42(10): 524979.
YANG M Y, LI Y G, XU K. An approach of tool path generation based on medial transformation for high-speed pocketing[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(10): 524979(in Chinese).

[16] HUANG R, JIANG J F, HUANG B, et al. Multilevel structured NC machining process model based on dynamic machining feature for process reuse[J]. The International Journal of Advanced Manufacturing Technology, 2019, 104(5): 2045-2060.

[17] ERBER M, ROSNITSCHEK T, HARTMANN C, et al. Geometry-based assurance of directional solidification for complex topology-optimized castings using the medial axis transform[J]. Computer-Aided Design, 2022, 152: 103394.

[18] CHEN Z C, FU Q. An optimal approach to multiple tool selection and their numerical control path generation for aggressive rough machining of pockets with free-form boundaries[J]. Computer-Aided Design, 2011, 43(6): 651-663.

[19] QIAN B, FAN H R, ZHANG J R. Fast calculation algorithm for region recognition and model interference ratio in the STL model based on voxel mapping decoupling[J]. The International Journal of Advanced Manufacturing Technology, 2022, 119(3): 1553-1578.

[20] RASOULZADEH S, SENK V, KÖNIGSBERGER M, et al. A novel integrative design framework combining 4D sketching, geometry reconstruction, micromechanics material modelling, and structural analysis[J]. Advanced Engineering Informatics, 2023, 57: 102074.

[21] BAI J, GAO S M, TANG W H, et al. Design reuse oriented partial retrieval of CAD models[J]. Computer-Aided Design, 2010, 42(12): 1069-1084.

[22] BICKEL S, SCHLEICH B, WARTZACK S. A novel shape retrieval method for 3D mechanical components based on object projection, pre-trained deep learning models and autoencoder[J]. Computer-Aided Design, 2023, 154: 103417.

[23] LUPINETTI K, GIANNINI F, MONTI M, et al. Content-based multi-criteria similarity assessment of CAD assembly models[J]. Computers in Industry, 2019, 112: 103111.

[24] LEE H, RYU K. Product and design feature-based similar process retrieval and modeling for mold manufacturing[J]. The International Journal of Advanced Manufacturing Technology, 2021, 115(3): 703-714.

Adaptive NC process planning approach for pocket features driven by CAD/CAM models

XU Changhong[1,*], HUANG Rui[2], ZHANG Shusheng[3], LIANG Jiachen[3], BIAN Rong[1]

(1. Engineering Training Center, School of Applied Technology, Nanjing Institute of Technology, Nanjing 211167, China;
2. College of Artificial Intelligence and Automation, Hohai University, Changzhou 213022, China;
3. School of Mechanical Engineering, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: To realize the effective reuse of existing numerical control (NC) process, an adaptive NC processes planning approach for pocket features driven by computer aided design/computer aided manufacturing (CAD/CAM) models was presented. Firstly, using the medial axis transform (MAT), the deeper NC process information concealed in 3D CAD models was recovered to explain the tool cutting process. Finally, based on the returned similar feature case through retrieval, the corresponding process parameters of the associated CAM model were reused adaptively to generate the NC process for the query feature. According to experimental results, the proposed method could automatically and effectively get the NC process for features and boost the NC process planning's efficiency by over 85%, satisfying the requirements of intelligent NC process planning for entertainment.

Keywords: numerical control process; reuse; pocket features; adaptive; similarity

Received: 2023-09-01; **Accepted**: 2024-01-10; **Published Online**: 2024-01-15 21:13

URL: link.cnki.net/urlid/11.2625.V.20240115.1106.001

Foundation items: National Natural Science Foundation of China (52075148); Jiangsu Provincial Natural Science Foundation of Higher Education, China (23KJB460017); The Scientific Research Foundation for the High-level Personnel of Nanjing Institute of Technology (YKJ202128)

* **Corresponding author**. E-mail: xuchanghonglz@163.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0535

基于PIV的冲压发动机喷管推力的重构测量方法

刘顺[1,2]，孔祥瑞[1,3]，徐惊雷[4]，高波[1]，刘海龙[1,*]

(1. 江苏大学 能源与动力工程学院，镇江 212013；2. 中国空气动力研究与发展中心 气动噪声控制重点实验室，绵阳 621000；
3. 奇瑞汽车股份有限公司，芜湖 241007；4. 南京航空航天大学 能源与动力学院，南京 210016)

摘 要：冲压发动机喷管推力系数的准确测量十分重要。传统的测力天平受环境噪声等外界因素的影响较大，因此，考虑通过测量喷管出口截面的速度、压力和密度分布，直接根据定义获得推力系数。以典型过膨胀状态下单边膨胀喷管的复杂超声速内流场为研究对象，提出一套基于粒子图像测速(PIV)的多物理场和推力系数的非接触式重构测量方法。通过PIV虚拟实验，综合评估了所提方法的理论精度。结果表明：基于PIV测得的多物理场，利用定义式直接测得的喷管推力系数约为0.779，相比于参考值0.771，相对误差仅为1.04%，从而验证了所提方法的可行性和准确性。

关 键 词：粒子图像测速；冲压发动机喷管；过膨胀；推力系数；重构测量
中图分类号：V211.71
文献标志码：A **文章编号**：1001-5965(2025)10-3385-07

冲压发动机是超声速飞行器重要的吸气式推进系统。喷管作为主要部件之一，其推力系数的准确测量十分重要。传统的测力天平不仅操作复杂，而且精度受环境噪声等外界因素的影响较大[1-2]。因此，若能直接测得喷管出口截面的速度、压力和密度分布，则能够直接根据定义获得推力系数，既直观又可靠。但是，过膨胀状态下的单边膨胀喷管(single expansion ramp nozzle, SERN)会产生大尺度分离，并诱导激波反射/相交、激波/附面层干扰等复杂流动现象。此时，常规的测压探针在严重非均匀的喷管出口流场中难以保证测量精度[3]，且空间分辨率很低，并将诱导产生干扰激波[4-6]。所以，高精度、高分辨率、低干扰的超声速流场定量测量技术非常关键。

非接触式光学定量测量技术，以粒子图像测速(particle image velocimetry, PIV)为代表，在波系结构复杂的超声速流场中具有很大优势[7]。PIV通过测量跟随流体运动的示踪粒子，在同一时刻记录大量空间点上的速度信息，从而提供详细的流场空间结构，并且几乎不会对流场本身产生干扰，因此，在冲压发动机内流场中得到广泛应用[8-9]。但是，PIV能够准确测得流场的速度分布，而压力场、密度场和温度场同样非常重要。

随着对多物理场同步测量需求的日益增加，近20年来，一种基于PIV的多物理场重构测量技术逐渐发展成熟[10]。其基本思路是：根据流体力学控制方程建立流场的速度与压力、密度和温度之间的定量关系，基于PIV测得的高精度速度场，通过特定的数值求解方法，获得其他多物理场。换言之，可以通过PIV直接测得的速度场，间接地对其他物理场进行高分辨率的全场非接触式同步测量，此即基于PIV的多物理场重构测量方法。由于全场的压力分布格外具有吸引力和代表性，因此，该技术多以基于PIV的压力场重构测量方法展现[10]。

最初的研究主要集中在不可压缩流场，其基本原理是：将动量方程中的压力项与其他项分离，代

收稿日期：2023-08-23；录用日期：2023-09-23；网络出版时间：2023-11-07 16:06
网络出版地址：link.cnki.net/urlid/11.2625.V.20231107.0903.002
基金项目：国家自然科学基金(12202164)；气动噪声控制重点实验室开放课题(ANCL20230307)；安徽省博士后研究人员科研活动经费资助项目(2022B600)
*通信作者. E-mail：leo@ujs.edu.cn

引用格式：刘顺，孔祥瑞，徐惊雷，等. 基于PIV的冲压发动机喷管推力的重构测量方法[J]. 北京航空航天大学学报，2025，51(10)：3385-3391.
LIU S，KONG X R，XU J L，et al. Reconstruction measurement method for ramjet nozzle thrust based on PIV[J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51(10)：3385-3391 (in Chinese).

入 PIV 测得的速度场，进而获得流场的压力梯度分布，然后，给定适当的边界条件，通过空间积分[11]（spatial integration, SI）法或 Poisson 方法[12]进行数值求解，最终得到流场的压力分布。超声速流场额外多出了密度未知量，使得压力场重构测量的难度大大增加：激波上游及远离附面层、剪切层和尾迹的大部分区域都属于等熵流动，此时速度场和压力场存在直接对应的代数关系[13]；在其他区域，由于激波属于流场间断面，因此，通过采用守恒型动量方程，并结合气体状态方程，可以得到适用于含有较弱激波的超声速流场的压力重构基本原理，最后使用类似的空间积分法得到重构压力场，称为守恒型空间积分（conservative spatial integration, CSI）方法[14]。然而，随着激波强度增加，流场间断也愈发剧烈，使得 CSI 方法在激波处的求解精度逐渐下降，且误差会向激波下游迅速累积，最终导致无法获得有效的重构压力场[15]。因此，自 2017 年以来，作者致力于基于 PIV 的超声速多物理场重构算法研究，并已取得重要成果[15-17]。其中，基于守恒型动量方程，借鉴计算流体力学（computational fluid dynamics, CFD）中的 MacCormack 求解技术，提出针对含有复杂波系的超声速多物理场重构测量方法，命名为 MacCormack 方法[15]，解决了较强激波后重构精度迅速恶化的问题，使得基于 PIV 的喷管推力非接触式重构测量成为可能。

本文以典型的冲压发动机过膨胀 SERN 的复杂超声速内流场为背景，提出一套基于 PIV 的压力场和推力的非接触式重构测量方法。由于流场包含复杂的激波/膨胀波系，其参考压力场很难通过理论或其他成熟的测量设备直接获取，所以无法直接使用 PIV 实验数据验证本文方法。因此，本文将利用自主搭建的 PIV 虚拟实验室，基于 CFD 数值速度场人工合成 PIV 粒子图像和噪声速度场，并以 CFD 获取的压力场和推力为参考，综合评估本文方法的可行性、准确性和理论精度。

1 喷管模型和数值流场

1.1 喷管模型

本文采用简化的二维 SERN 构型，如图 1 所示[18]。模型使用等长直壁设计，喉道高度 h 为 20 mm，上壁面扩张角为 25°，下壁面水平，设计落压比为 20，膨胀面积比为 2.896。

1.2 数值流场

以上述 SERN 喷管工作在落压比 3 时的过膨胀内流场为例，借鉴 Yu 等[18]在过膨胀喷管流动分离现象中的数值模拟经验，获得用于定量评估本文方法精度的参考速度场和压力场。

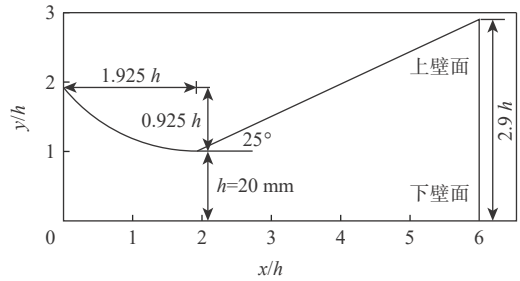
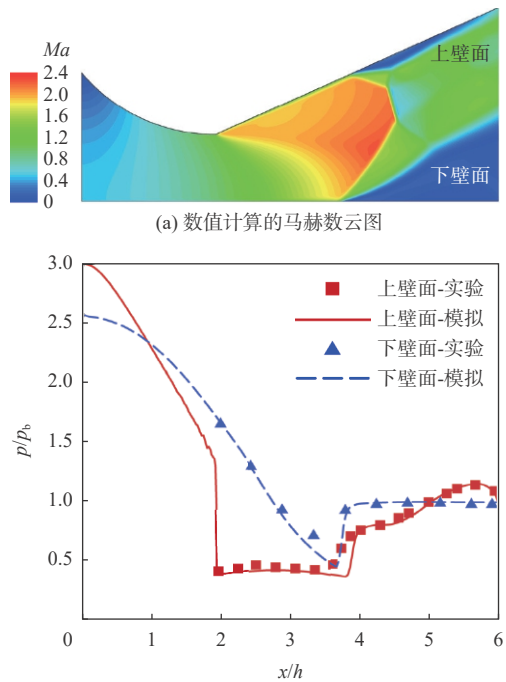


图 1 简化的 SERN 构型[18]

Fig. 1 Simplified SERN configuration[18]

具体地，基于密度基求解器，对二维定常可压缩雷诺平均 Navier-Stokes 方程进行隐式时间推进求解；采用 k-ε RNG 湍流模型，并结合增强型壁面函数；使用二阶迎风差分格式，对流通量采用 Roe 通量差分格式（Roe-FDS）。边界条件给定进口总压为 93.2 kPa，背压 p_b 为 31.1 kPa，总温为 300 K。网格划分及其无关性验证见文献 [18]，此处不再赘述。图 2 给出了数值计算的马赫数云图，以及壁面压力分布与实验测量结果的对比。



(a) 数值计算的马赫数云图

(b) 数值计算与实验测量对比

图 2 数值计算与实验测量的结果对比

Fig. 2 Comparison of CFD and experimental results

整体而言，两者吻合良好，仅分离点位置略有偏差，主要由分离泡的振荡引起。由于喷管处于严重的过膨胀状态，导致下壁面产生较大的分离，形成分离激波，并进一步诱导上壁面产生激波/附面层干扰，最终形成复杂的马赫反射。

2 重构测量方案和理论验证

本节首先简要介绍针对超声速激波流场的压力重构原理。然后，基于第 1 节 SERN 和数值参考

流场，建立对应的压力场重构测量方案，并从理论上评估其可行性和重构精度。

2.1　超声速压力场重构测量原理

引言中已经提到，目前针对超声速流场的压力重构方法主要有 2 种，一种是基于守恒型动量方程的 CSI 方法，另一种是基于时间推进求解方案的 MacCormack 方法。

激波上游及远离附面层、剪切层和尾迹的大部分区域属于等熵流动，此时速度场和压力场存在如下简单的代数关系[13]：

$$\frac{p}{p_\infty}=\left[1+\frac{\gamma-1}{2}Ma_\infty^2\left(1-\frac{V^2}{V_\infty^2}\right)\right]^{\frac{\gamma}{\gamma-1}} \tag{1}$$

式中：p、γ、Ma 和 V 分别为压力、比热比、马赫数和速度大小，下标∞表示自由来流。其他非等熵区域则必须采用 CSI 方法或 MacCormack 方法。

对于 CSI 方法，由于存在额外的密度未知量 ρ，因此，还需引入其他方程。假设流场处于绝热状态，且比热比为常数，则流场温度 T 可直接由以下代数方程获得[13]：

$$\frac{T}{T_\infty}=1+\frac{\gamma-1}{2}Ma_\infty^2\left(1-\frac{V^2}{V_\infty^2}\right) \tag{2}$$

考虑到黏性项对重构压力场的影响很小[14]，因此，定常守恒型动量方程可以简化为

$$\frac{\partial p}{\partial x_i}=-\frac{\partial\left(\rho u_i u_j\right)}{\partial x_j} \tag{3}$$

式中：u_i 和 u_j 为速度分量。

结合气体状态方程，使用压力和温度替换密度，得到 CSI 方法的基本方程[14]：

$$\left(\delta_{ij}+\frac{u_i u_j}{RT}\right)\frac{\partial\ln(p/p_\infty)}{\partial x_j}=-\frac{1}{RT}\left(\frac{\partial u_i u_j}{\partial x_j}-\frac{u_i u_j}{T}\cdot\frac{\partial T}{\partial x_j}\right) \tag{4}$$

式中：R 为气体常数；δ_{ij} 为克罗内克符号。

MacCormack 方法的原理基于守恒型 Navier-Stokes 方程，具体包括状态方程、动量方程和能量方程，连续方程已经隐含在守恒型动量方程中，因此，基本方程如下[15]：

$$\begin{cases}p=\rho RT\\ e=c_v T\end{cases} \tag{5}$$

$$\frac{\partial \boldsymbol{U}}{\partial t}=-\frac{\partial \boldsymbol{F}}{\partial x}-\frac{\partial \boldsymbol{G}}{\partial y} \tag{6}$$

式中：e 和 c_v 分别为内能和定容比热容；$\boldsymbol{U}$、$\boldsymbol{F}$ 和 $\boldsymbol{G}$ 给定为

$$\boldsymbol{U}=\begin{bmatrix}\rho(u+v)\\ E\end{bmatrix},\boldsymbol{F}=\begin{bmatrix}\rho u^2+\rho uv+p\\ u(E+p)+q_x\end{bmatrix},$$

$$\boldsymbol{G}=\begin{bmatrix}\rho uv+\rho v^2+p\\ v(E+p)+q_y\end{bmatrix}$$

其中：u 和 v 分别为 x 和 y 方向的速度分量，且：

$$E=\rho\left(e+\frac{u^2+v^2}{2}\right),q_x=-K\frac{\partial T}{\partial x},q_y=-K\frac{\partial T}{\partial y}$$

其中：K 为导热系数。

使用 MacCormack 时间推进求解技术，得到超声速重构压力场：

$$\boldsymbol{U}^{t+\Delta t}=\boldsymbol{U}^t+\left(\partial\boldsymbol{U}/\partial t\right)_{\mathrm{av}}\Delta t \tag{7}$$

式中：$(\partial\boldsymbol{U}/\partial t)_{\mathrm{av}}$ 为 $\partial\boldsymbol{U}/\partial t$ 在时间步长 Δt 内的平均值。具体的理论分析和执行细节详见文献 [15]。

2.2　本文重构测量方法

本节将综合利用 MacCormack 方法在激波流场中的高精度及 CSI 方法的灵活性，建立基于 PIV 的冲压发动机喷管内流压力场和推力系数的非接触式、高分辨率重构测量方案。

将 SERN 的过膨胀内流场分为 3 个区域，如图 3 所示。区域 1 虽然不规则，但是流场处于等熵膨胀状态，因此，可以直接使用等熵法（方程式(1)）获取压力分布。区域 2 包含马赫反射的主要波系结构，所以采用 MacCormack 方法更为合适。受常规 PIV 空间分辨率的限制，剪切层内部通常无法精确解析。因此，为确保 MacCormack 方法的边界条件具有较高精度，本文将区域 2 分为 2 个矩形子区（虚线框所示），尽量保证重构边界远离剪切层核心区。左侧边界条件基于区域 1 的重构结果插值给定（第 1 类边界条件），上、下两侧边界条件由方程式(4)给定（第 2 类边界条件），右侧边界直接给定无反射条件。区域 3 不规则，并以低速分离流场为主，因此，采用更加灵活的 CSI 方法：以重构压力场已知的区域 2 下边界为起始位置，沿垂直于剪切层的方向逐行向下壁面进行多路径空间推进求解[11]，这在很大程度上降低了剪切层内的速度测量误差和脉动速度对压力场重构精度的影响。此外，为防止起始边界上激波点处的较大误差沿积分路径扩散，每行的积分都由右向左依次执行。受壁面反光的影响，PIV 在近壁区的测量误差较大。为模拟这种实际情况，以上各区域边界距离壁面 2 mm。

2.3　理论验证

本节将基于纯净的数值速度场，如图 4 所示，评

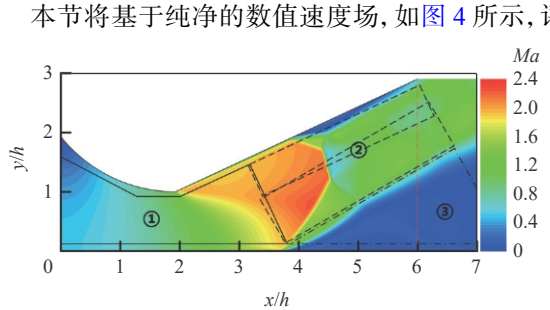
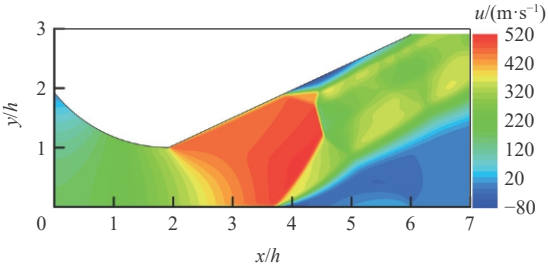
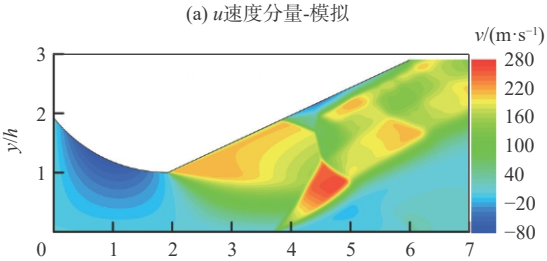


图 3　重构方案示意图

Fig. 3　Reconstruction scheme



(a) u速度分量-模拟



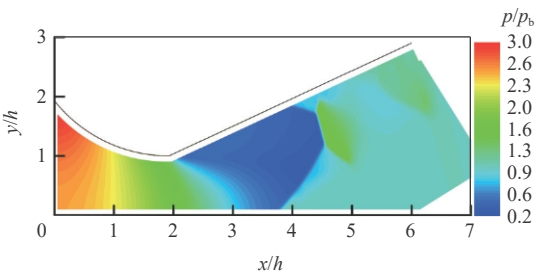
(b) v速度分量-模拟

图4 纯净数值速度场

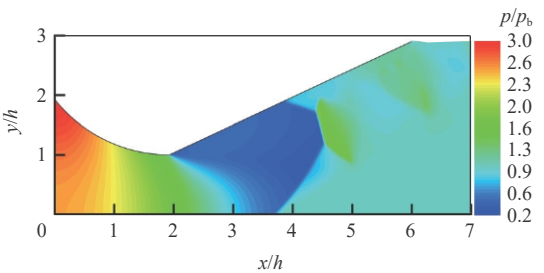
Fig. 4 Clean numerical velocity field

估本文方法的可行性和理论精度。为模拟实际PIV的数据结构，上述各个区域的速度场均由0.5 mm的网格插值而得。

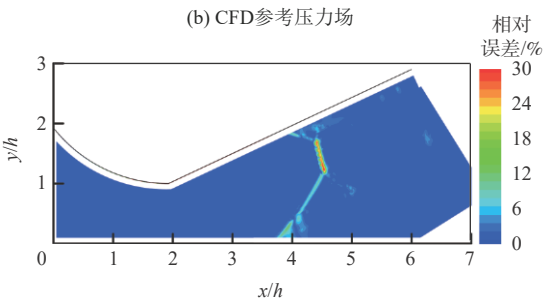
图5为重构压力场和CFD参考压力场的对比，以无量纲压比p/p_b显示。可以发现，基于3个区域组合而得的重构压力云图没有明显间断。由图5(c)可知，重构误差集中在激波“内部”，主要由CFD有限的网格分辨率所引起的速度误差导致；在激波“区域”之外，重构精度始终较高。区域3采用了CSI方法，且每行都由右向左执行积分操作，使得起始边界上激波点处重构误差较大的区域限制在左侧角落。总体而言，基于时均速度场得到的重构压力场与参考结果吻合良好，整场的均方根相对误差约为0.4%，且分离区、剪切层和附面层处的重构误差没有显著增大或向四周扩散，尤其是下壁面大分离区的均方根相对误差仅为0.5%，这从侧面证明了速度脉动项对本文方法的影响非常有限[14,17]，并初步验证了其可行性与高精度。



(a) 重构压力场



(b) CFD参考压力场



(c) 相对误差分布

图5 重构压力场与CFD参考压力场的对比

Fig. 5 Comparison of reconstructed pressure field and CFD reference pressure field

3 基于虚拟PIV实验的评估

3.1 虚拟PIV实验

PIV粒子图像基于自主编写的合成图生成算法获取[15]，分辨率为1 120像素 × 2 560像素，对应的目标视场为70 mm × 160 mm。电荷耦合器(charge coupled device, CCD)的像素密度为156.25像素/mm，饱和电平和填充比都设置为0.75，噪声服从均值为15、标准差为5(单位：灰度)的高斯分布。激光片光的强度同样服从高斯分布，厚度设置为1.5 mm。激光脉冲时间间隔Δt设置为2 μs。经过统计，合成图中粒子图像的平均直径为2.5像素，标准差为0.3像素。单位像素中的粒子个数N_{ppp}为0.03，粒子图像密度N_s为0.15，低于散斑极限0.3。粒子的平均物理直径设置为0.5 μm，密度设置为2.6×10^3 kg/m^3，使其平均松弛时间为2.02 μs。以上设置使得合成图的信噪比SNR约为6.5，如图6所示。可以发现，粒子浓度在激波下游增大，膨胀区的粒子浓度较低，分离区内的粒子很少，这都与实际的PIV粒子图像相符。尽管合成图的质量相比于实验数据仍然偏高，但目前依旧是评估重构测量算法精度的最佳方案。

使用开源PIVlab[19]对合成图进行后处理，通过多尺寸查访区变形迭代互相关算法(查访区变形采用样条曲线插值法，亚像素值使用高斯法计算)，



(a) 粒子图像局部放大图　(b) 整场粒子图像

图6 PIV粒子合成图示例

Fig. 6 Synthetic particle image of PIV

逐次进行 3 次迭代处理, 最后一次迭代的查访区尺寸为 32 像素 × 32 像素, 重叠率为 75%, 使得相邻速度矢量的间距为 0.5 mm。最终, 基于 20 对粒子图像得到收敛的虚拟 PIV 速度场及其对应的绝对误差, 如图 7 所示。

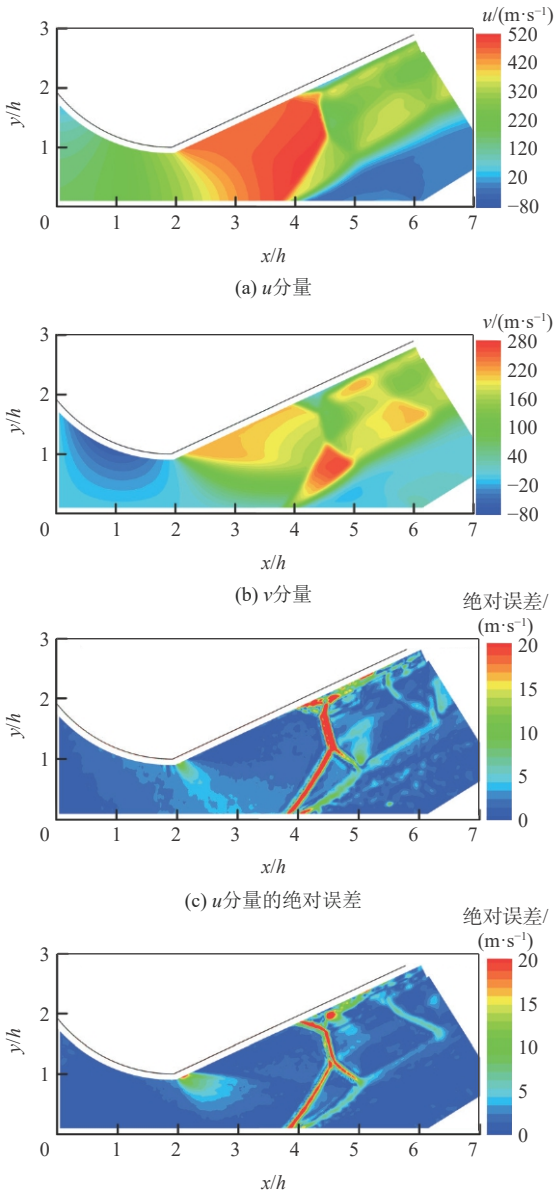


图 7　虚拟 PIV 速度场的 u/v 分量及其对应的绝对误差

Fig. 7　u/v components of synthetic PIV velocity field and corresponding absolute error distribution

由图 7 可知, 受粒子松弛效应和 PIV 有限时空分辨率的影响, 在速度梯度较大的区域, 包括激波、膨胀区和剪切层, 都出现了较为明显的误差。由于区域 2 的上下边界位于附面层和剪切层附近, 因此, 会对 MacCormack 方法的精度产生一定的潜在影响。总之, 上述虚拟 PIV 速度场不仅含有复杂的激波/膨胀波系, 而且加入了一系列 PIV 实验中存在的典型误差, 这将对本文方法带来很大挑战。

3.2　压力场和推力系数的重构测量

基于上述虚拟 PIV 速度场, 运用本文方法, 得到对应的压力场及其相对误差分布, 如图 8 所示。

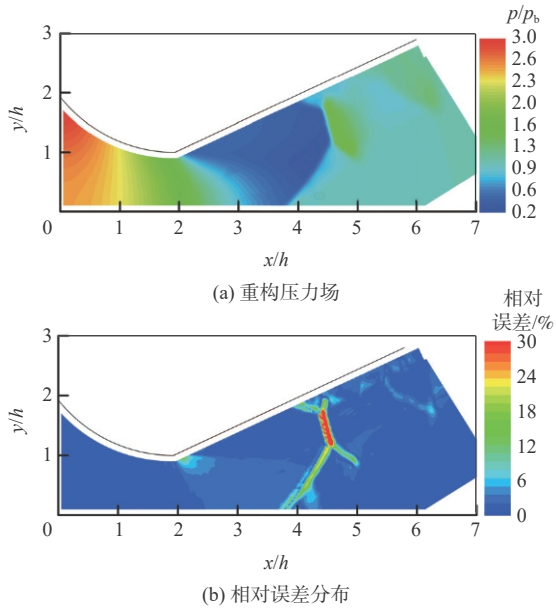
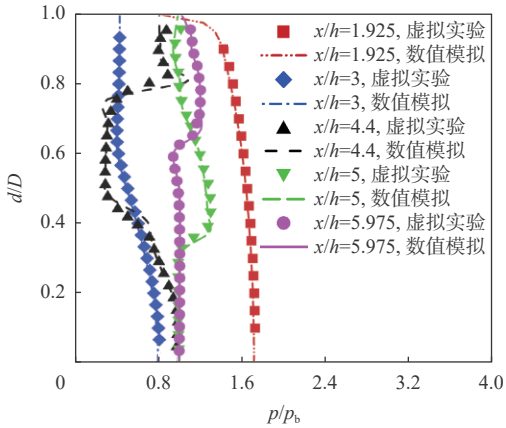


图 8　重构压力场及其相对误差分布

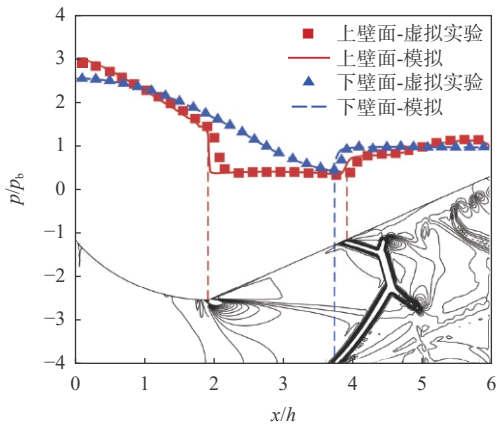
Fig. 8　Reconstructed pressure field and corresponding relative error distribution

可以发现, 喉道前等熵膨胀区的速度精度高, 使得等熵法的重构误差小于 0.5%; 喉道附近的剧烈膨胀导致粒子松弛效应加剧, 速度场误差迅速增大, 使得压力重构精度显著下降, 膨胀扇起点附近的相对误差甚至达到 10%, 但是在远处逐渐下降至 3.0% 以内, 从而保证了区域 2 中 MacCormack 方法左侧边界条件的精度。此外, 区域 2 的下边界距剪切层核心区较远(见 2.2 节), 且上边界处附面层的速度误差对重构压力精度的影响范围主要限制在当地局部流场, 最终使得马赫反射之后主流区的重构误差大都低于 1%, 同时保证了区域 3 中 CSI 方法边界条件的精度。尽管剪切层中的粒子松弛效应较为明显, 但是对法向速度分量(接近 0)的影响有限, 使得 CSI 方法沿垂直于剪切层方向的积分精度较高[14], 保证了区域 3 的压力场重构精度, 均方根相对误差约为 0.8%。以上结果充分验证了重构方案的准确性, 这也是基于 PIV 的喷管推力系数重构测量的基础。

图 9(a)为内流场 5 个典型位置处(x/h = 1.925, 3, 4.4, 5, 5.975, 沿高度方向(d/D)上的重构压力分布和对应的参考值。可以发现, 基于重构压力场, 沿壁面法向进行外插, 能够获得较高精度的壁面压力分布, 如图 9(b)所示。除激波反射点和喉道附近的剧烈膨胀区之外, 外插结果的相对误差大都在 3% 以内, 从而验证了基于 PIV 的壁面压力测量在



(a) 竖直截面



(b) 壁面

图 9 竖直截面和壁面处重构压力分布与参考值的对比

Fig. 9 Comparison of reconstructed pressure distributions and reference results at vertical sections and walls

SERN 中的可行性。

在获取了 SERN 出口截面(x/h=5.975)速度和压力分布的基础上，得到

$$C_{fx}=\frac{\dot{m}u+(p-p_{\mathrm{b}})A_x}{\dot{m}\sqrt{\dfrac{2\gamma}{\gamma-1}RT_t\left[1-\left(\dfrac{p_{\mathrm{b}}}{p_t}\right)^{\frac{\gamma-1}{\gamma}}\right]}} \tag{8}$$

式中：A_x 为出口截面积在 x 方向的投影；$\dot{m}$ 为质量流量；下标 t 表示滞止参数。

可以直接根据式(8)得到喷管的推力系数 C_{fx}，约为 0.779，相比于 CFD 参考值 0.771，相对误差仅为 1.04%。

4 结 论

本文以 MacCormack 方法、CSI 方法和等熵法为基础，提出一种基于 PIV 的冲压发动机喷管内流多物理场和推力系数的非接触式重构测量方法，并通过虚拟 PIV 实验进行了较为全面的理论分析。结论如下：

1）基于纯净速度场对本文方法进行初步评估，所得重构压力场的均方根相对误差仅为 0.4%，证明了本文方法较高的理论精度。

2）利用虚拟 PIV 实验，基于含有多种测量误差模型的合成速度场，对本文方法的精度和准确性进行全面分析：虽然粒子松弛效应导致激波、附面层和剪切层附近的速度场含有较大误差，但是压力场的重构误差依然主要限制在当地局部流场；基于已获得的重构压力场，沿壁面法向进行外插，能够获得精度较高的壁面压力分布，相对误差大都在 3% 以内；得益于出口截面上高精度的速度/压力分布，基于定义式直接获得喷管的推力系数约为 0.779，相对误差仅为 1.04%。

总之，本文方法在未来冲压发动机喷管的风洞试验中展示出巨大潜力：只要能够获取高精度的 PIV 速度场，就可以实现速度、压力、温度、密度和推力系数的同步定量测量；无需结合多种传统的测量设备，尤其是操作复杂且测量精度易受环境干扰的测力天平；相应地，该方法对 PIV 的测量精度提出了较高要求，须在实验中进行精细的标定、测试与数据后处理。

参考文献（References）

[1] LV Z, XU J L, YU K K, et al. Experimental and numerical investigations of a scramjet nozzle at various operations[J]. Aerospace Science and Technology, 2020, 96: 105536.

[2] 汪丰, 徐惊雷, 汪阳生. 并联式 TBCC 排气系统模态转换过程的流场特性研究[J]. 实验流体力学, 2019, 33(3): 68-75.
WANG F, XU J L, WANG Y S. Study of flow field characteristics of an over-under TBCC exhaust system during mode transition process[J]. Journal of Experiments in Fluid Mechanics, 2019, 33(3): 68-75(in Chinese).

[3] HOHN O M, GÜLHAN A. Experimental investigation of sidewall compression and internal contraction in a scramjet inlet[J]. Journal of Propulsion and Power, 2017, 33(2): 501-513.

[4] TANI K, KANDA T, KUDOU K J. Aerodynamic performance of scramjet inlet models with a single strut[J]. Journal of Propulsion and Power, 2006, 22(4): 905-912.

[5] SMART M K. Experimental testing of a hypersonic inlet with rectangular-to-elliptical shape transition[J]. Journal of Propulsion and Power, 2001, 17(2): 276-283.

[6] FLOCK A K, GÜLHAN A. Experimental investigation of the starting behavior of a three-dimensional scramjet intake[J]. AIAA Journal, 2015, 53(9): 2686-2693.

[7] RAFFEL M, WILLERT C E, SCARANO F, et al. Particle image velocimetry: a practical guide[M]. Berlin: Springer, 2018.

[8] XU J L, SHA J, SHI Z Y, et al. PIV experimental study and numerical simulation of the over-expanded SERN exit jet[C]//Proceedings of the 46th AIAA Aerospace Sciences Meeting and Exhibit. Reston: AIAA, 2008.

[9] XU J L, LIN C F, SHA J. Particle image velocimetry study of the impinging height effect on an overexpanded supersonic impinging

free jet at *Ma* = 1.754. Part 1: global coherent structure[J]. Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering, 2012, 226(1): 64-73.

[10] VAN OUDHEUSDEN B W. PIV-based pressure measurement[J]. Measurement Science and Technology, 2013, 24(3): 032001.

[11] BAUR T. PIV with high temporal resolution for the determination of local pressure reductions from coherent turbulence phenomena [C]//Proceedings of the 3rd International Workshop on PIV. Berlin: Springer, 1999: 101-106.

[12] GURKA R, LIBERZON A, HEFETZ D, et al. Computation of pressure distribution using PIV velocity data[C]//Proceedings of the 3rd International Symposium on Particle Image Velocimetry. Bristol: Measurement Science & Technology , 1999, 2: 1-6.

[13] ANDERSON J D. Fundamentals of aerodynamics[J]. AIAA Journal, 2010, 48(12): 2983-2983.

[14] VAN OUDHEUSDEN B W. Principles and application of velocimetry-based planar pressure imaging in compressible flows with shocks[J]. Experiments in Fluids, 2008, 45(4): 657-674.

[15] LIU S, XU J L, YU K K. MacCormack's technique-based pressure reconstruction approach for PIV data in compressible flows with shocks[J]. Experiments in Fluids, 2017, 58(6): 64.

[16] LIU S, XU J L. Assessment of pressure field and performance parameter reconstruction from velocity data for ramjet inlet[J]. Aerospace Science and Technology, 2021, 110: 106454.

[17] LIU S, XU J L, GONG H, et al. Assessment of streamline-based pressure reconstruction for PIV data in supersonic flows with shock waves[J]. Experiments in Fluids, 2020, 61(7): 161.

[18] YU Y, XU J L, MO J W, et al. Principal parameters in flow separation patterns of over-expanded single expansion RAMP nozzle[J]. Engineering Applications of Computational Fluid Mechanics, 2014, 8(2): 274-288.

[19] THIELICKE W, SONNTAG R. Particle image velocimetry for MATLAB: accuracy and enhanced algorithmsin PIVlab[J]. Journal of Open Research Software, 2021, 9(1): 12.

Reconstruction measurement method for ramjet nozzle thrust based on PIV

LIU Shun[1, 2], KONG Xiangrui[1, 3], XU Jinglei[4], GAO Bo[1], LIU Hailong[1, *]

(1. School of Energy and Power Engineering, Jiangsu University, Zhenjiang 212013, China;
2. Laboratory of Aerodynamic Noise Control, China Aerodynamics Research and Development Center, Mianyang 621000, China;
3. Chery Automobile Co., Ltd., Wuhu 241007, China;
4. College of Energy and Power Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

Abstract: Accurate measurement of the thrust coefficient of ramjet nozzles is crucial. Traditional force balances are highly susceptible to external factors such as ambient noise and signal interference. Therefore, the paper proposed directly obtaining the thrust coefficient based on the definition by measuring the velocity, pressure, and density distributions at the nozzle outlet. A non-intrusive reconstruction measurement scheme based on particle image velocimetry (PIV) was proposed for multiple physical fields and the thrust coefficient of a typical single expansion ramp nozzle under over-expansion conditions. The theoretical accuracy of the reconstruction scheme was comprehensively evaluated using synthetic PIV experiments. The results indicate that the thrust coefficient of the nozzle, measured using the reconstructed multi-physical fields from PIV, is about 0.779, with a relative error of only 1.04% compared to the reference value of 0.771. This confirms the feasibility and accuracy of the proposed method.

Keywords: particle image velocimetry; ramjet nozzle; over-expansion; thrust coefficient; reconstruction measurement

Received: 2023-08-23; **Accepted**: 2023-09-23; **Published Online**: 2023-11-07 16:06

URL: link.cnki.net/urlid/11.2625.V.20231107.0903.002

Foundation items: National Natural Science Foundation of China (12202164); Open Project of Laboratory of Aerodynamic Noise Control (ANCL20230307); Anhui Province Postdoctoral Researcher Research Activity Funding Project (2022B600)

* **Corresponding author**. E-mail: leo@ujs.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0477

同质竞争下补贴策略对多机场航线网络演化博弈

吴维[1,*]，林芷伊[2]，王兴隆[1]

(1. 中国民航大学 民航飞联网重点实验室，天津 300300；2. 北京航空航天大学 电子信息工程学院，北京 100191)

摘 要：为实现区域内多机场基于差异化定位的高质量协同发展，研究机场差异化补贴策略对多机场航线网络演化影响，进而确定最佳补贴策略。基于旅客、航司、机场间的竞争博弈关系，构建了双层演化博弈模型。在上层博弈模型中，考虑旅客自学能力对票价的影响，构建融合自学习机制的Logit旅客选择模型，利用Hotelling定价模型分析同一航线航司间票价竞争对旅客选择行为的影响，进而确定在竞争条件下航司最佳定价策略；在下层博弈模型中，基于复制动态方程分析各机场补贴与航司间竞争性选择航线优化过程，确定机场间协同补贴策略与航线网络协同效果。结果表明：对于转移航线的航司，吸引“渗流”旅客的优势票价折扣区间为0.6～0.75；同航线竞争的航司票价折扣集中在0.6～0.85之间，可避免出现低价竞争带来的收益共损；通过机场差异化补贴实现航线网络优化，不同机场均存在基于差异化功能定位的最佳补贴区间。

关 键 词：航空运输；差异化补贴策略；航线网络优化；双层演化博弈；复制动态方程
中图分类号：V355.2
文献标志码：A **文章编号**：1001-5965(2025)10-3392-13

城市规模扩大与区域经济发展，使得机场也出现了聚集，机场群孕育而生。由于机场间长期存在隐形壁垒，强大的“虹吸效用”使机场群出现了国际枢纽机场资源紧张、航线主骨干通道容量不足，而机场群内的国内枢纽机场、区域机场资源过剩、航线次骨干通道有待利用的现象。为此，研究多机场的协同发展成为很多学者研究重点。

1）研究旅客出行行为与航线网络演变关系。文献[1]考虑高铁和航空运输竞争下构建旅客出行选择偏好模型，探讨不同竞争策略下路网/航线间的竞争关系。文献[2]基于多家航司间的竞争，构建航司航线选择收益模型，分析在竞争视角下机场收费标准对航司建立航线网络的影响。文献[3]利用旅客出行数据，实证分析航司选择不同等级机场对其机场航线网络演化的影响。文献[4]分析票价定价策略与航线网络组合之间的关系，并设计了航司定价算法。文献[5]考虑在同质化竞争日益激烈情况下，分析航空联盟与航线网络之间的关系，并构建指标体系评价其紧密程度。文献[6]考虑不断加剧的高铁与民航竞争关系，采用博弈理论构建竞争与合作的模型。文献[7]研究了多机场内旅客在不同机场间的选择行为，阐述不同条件下旅客的“渗流”现象。文献[8]统计分析多机场内各机场航线网络特征，基于重叠度、频次等指标构建了评价模型，分析旅客选择的可替代性及机场间的竞争程度。文献[9]针对中距离旅客选择行为特征建立博弈模型，分析在中距范围内高铁与航空的竞争关系，并用逆向归纳法求解双方博弈策略。

2）针对机场、航司2个航线网络演变推动主体，研究两者相互作用对旅客出行和航线网络规划的影响。文献[10]围绕机场发展策略对航司制定票价策略的影响，构建了机场与航线的选择偏好模型，采用二维策略图进行求解，为优化双方策略提供基础。文献[11]系统分析了航司、机场与旅客之

收稿日期：2024-06-21；**录用日期**：2024-08-16；**网络出版时间**：2024-09-13 10:00
网络出版地址：link.cnki.net/urlid/11.2625.V.20240912.1744.004
基金项目：中央高校基本科研业务费专项资金中国民航大学专项(3122025098)
*通信作者. E-mail：wwu@cauc.edu.cn

引用格式：吴维，林芷伊，王兴隆. 同质竞争下补贴策略对多机场航线网络演化博弈[J]. 北京航空航天大学学报，2025，51(10)：3392-3404.
WU W，LIN Z Y，WANG X L. Evolutionary game of subsidy strategy on multi-airport route network under homogeneous competition[J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51(10)：3392-3404 (in Chinese).

间的共生关系, 从生态视角建立动力学方程, 探讨三者之间关系, 并制定协同发展最优策略。文献 [12] 针对新冠疫情影响后的航线恢复问题, 研究不同类型机场采取不同补贴策略对彼此航线网络的影响, 为制定机场不同补贴策略提供基础。文献 [13-14] 考虑多利益主体影响下航司构建航线网络模式, 利用博弈论和逆向归纳法进行求解, 确定不同利益相关策略及最佳航线网络模式。

综上, 现有研究主要针对航线网络结构特征、旅客出行行为选择、航司发展战略等角度分析航线网络演变特征及各因素间的相互影响, 主要采用数据统计分析、出行行为建模、动力学建模等方法, 为研究航线网络优化提供了思路和方法。但随着区域内多机场的发展及航司间的竞争加剧, 航线网络演化面临新的挑战, 需要对影响因素和研究方法进一步完善, 主要问题包括: ①多机场同质化竞争日益严重, 降低竞争增强协作的关键是差异化的航线网络调整(即航司航线网络构建要匹配机场的功能定位), 而针对基于多机场协同定位的主动引导式航线网络优化研究不足; ②研究多机场旅客出行行为对航线网络的影响时主要基于票价因素决定旅客选择行为, 而实质上旅客作为决策主体具有智能学习决策特征, 其会根据其他旅客收益、历史决策经验、航司特征等动态修正出行行为, 而现有研究弱化了旅客自身选择特征的影响; ③研究航司航线优化时, 主要从补贴、自身发展战略出发, 对多机场协同补贴策略及航司同质化竞争对航线网络构建的影响考虑不足。

为此, 本文针对多机场、多航司同质化竞争下, 分析补贴策略对多机场航线网络优化引导效能, 以最佳补贴策略引导航司构建匹配机场功能定位的航线网络, 实现多机场协同发展。综合考虑旅客、航司和机场间的同质竞争、异质博弈, 构建了双方、双层演化博弈模型。

1　问题描述与建模思路

随着城市群的发展, 机场也呈现出了聚集现象, 出现了相距比较近的多机场, 由于航司间竞争造成机场航线网络趋同, 叠加距离较近使得机场间差异化定位执行效果不佳, 造成多机场发展不均衡、资源浪费、运营效率低等问题。引导机场按照差异化功能定位实现协同发展, 就成为多机场发展的必由之路。例如, 多机场内可按照机场特点分为国际枢纽机场、国内枢纽机场和区域机场, 进而打造差异化航线网络。国际枢纽机场重点打造国际航线, 疏解与自身定位不符的航线; 国内枢纽机场重点打造国内干线网络; 区域机场主要打造区域内支线。但多数机场都在开航相同的航线, 造成各机场航线网络趋同, 同质化严重。航线网络构建是市场行为, 需要通过补贴策略引导航司以差异化机场定位为基础优化航线布局。针对已具备机场群特征的现有航线网络, 通过补贴引导航司调整其在不同机场的航线网络, 打造出协同化的航线网络, 该过程如图 1 所示。

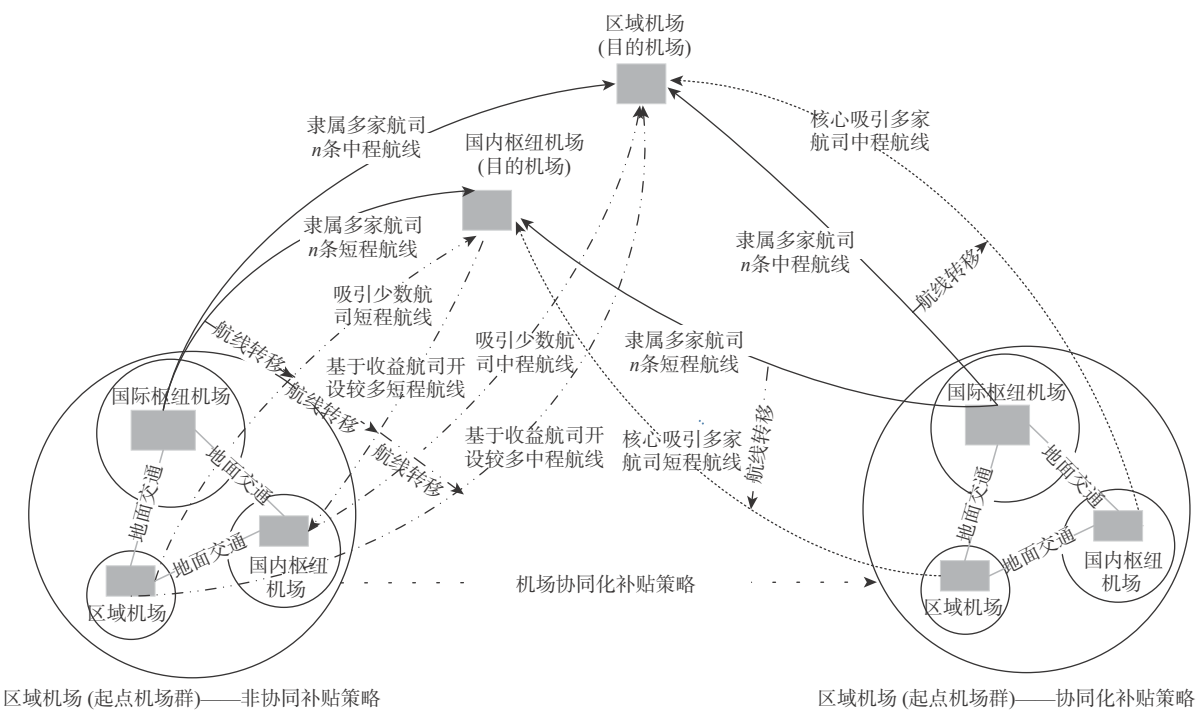


图 1　机场间航线网络协同演化示意图

Fig. 1　Collaborative evolution diagram of route network between airports

图1左侧表示现有机场群航线网络特点，机场群中有不同定位的机场，由于航司为争夺热门航线和航线补贴，造成机场间航线重叠度高，机场差异化定位没有体现（如国际枢纽机场开辟很多中短程航线，国内枢纽机场开辟很多远程航线），为此，通过以机场差异化功能实现为目标构建机场间协同化航线补贴策略，使得航司在追求收益的同时打造了机场差异化航线网络，降低同质化竞争，提升机场协同性，国际枢纽机场以远程航线为主，国内枢纽机场以中程航线为主，区域机场以短程航线为主的差异化共赢航线网络。

上述演变过程是基于多机场功能已经明确下，由机场主导完成航线网络布局，为此，机场通过制定补贴策略，引导航司开辟、调整航线实现其定位。因此，上述市场行为的核心是航司，而航司调整航线网络的关键是收益，收益主要由客票收入和补贴收入构成，其值大小受2方面博弈影响。一方面，客票收益本质是旅客与航司间的博弈，不同票价叠加航司间竞争造成吸引旅客数量不同形成不同收益曲线，航司会通过票价折扣不断扩大腹地旅客和“渗流”旅客人数；另一方面，机场为实现自身功能定位制定最小补贴策略吸引航司开辟航线，而航司希望通过调整航线获取最大补贴收益，两者之间产生了博弈，为此，本文构建了双层演化博弈模型。

2 上层博弈模型

上层博弈模型在考虑旅客选择不同公司航线折扣收益、旅客间相互学习及同一航线不同航司间竞争下，分析航司定价策略对旅客选择概率的影响，以确定航司在最佳收益时的定价策略。

2.1 融合自学习改进Logit旅客选择模型

该博弈过程的博弈主体是非腹地旅客与航司。机场在建设时有主要服务的行政范围旅客，不在其服务行政范围而选择该机场起降的为非腹地旅客。非腹地旅客对价格比较敏感，会受到航司折扣票价的影响，根据价格作出每次出行是否为渗流旅客的决策。

该演化博弈过程的基本假设如下：

H1：在固定航线下，旅客总人数为选择民航出行旅客且为可估算值，所有旅客均可掌握航司的票价信息作出是否渗流的决策，并根据收益进行策略调整。

H2：假设旅客在未转移前享受的机票价格为A^*，并且各航司会根据实际情况制定折扣票价吸引旅客实现由非腹地旅客向“渗流”旅客的身份转变。由于旅客在选择是否成为“渗流”旅客的出行策略时，更多关注新交通起止点（origin-destination，OD）航线市场的总体票价情况，选取新市场各航司的折扣票价均值作为旅客转移后所享受的折扣票价$\overline{A}$。

H3：旅客在实现由非腹地旅客向“渗流”旅客的身份转变时，会产生额外的交通成本G，包括2个机场间的货币成本G_{M}及时间成本G_{T}。货币成本G_{M}主要考虑交通出行所需的现金成本，包括城际间出行成本（如以高铁出行成本计算）及市内出行成本（如以网约车和公共交通占比计算平均成本或最低交通成本）；时间成本G_{T}是旅客到非腹地机场增加的额外出行成本，在对其量化时考虑旅客对时长敏感性，即设定时间阈值，超过该阈值设置较大时间成本，在阈值内可按当地平均小时工资作为参考。

H4：航司对航线调整后，腹地旅客会增加中转或候机成本Q。

根据H1～H4，构建旅客出行选择的成本矩阵，如表1所示，表中每种选择策略对应2行成本，第1行为旅客i成本，第2行为旅客j成本。

表1 基于出行策略的旅客效益矩阵

Table 1 Passenger benefit matrix based on travel strategy

策略种类	旅客i效益	旅客j效益
渗流与渗流	$\overline{A}(t)+G$	$\overline{A}(t)+G$
渗流与非渗流	$\overline{A}(t)+G$	A^*+Q
非渗流与渗流	A^*+Q	$\overline{A}(t)+G$
非渗流与非渗流	A^*+Q	A^*+Q

通过购票软件，旅客可以全面掌握票价价格，叠加其他旅客经验分享及自身经验提升，都会影响其出行行为决策，为更加客观表征旅客学习行为，引入经验加权吸引力模型（experience-weighted attraction，EWA）学习算法[15]。该算法具有高效率、低误差的特点，可计算旅客的出行选择概率，能将旅客经验有效传递给下一期决策，有效表征学习经验价值。

EWA学习算法通过吸引力指数B来更新旅客选择策略k的概率，策略对应的吸引力指数值越大，则该策略被选择的可能性越高。计算仿真周期第t期旅客i选择策略k时的吸引力指数$B_i^k(t)$为

$$B_i^k(t)=\frac{\varphi\cdot N(t-1)\cdot B_i^k(t-1)+\pi_i^k(t)}{N(t)} \tag{1}$$

式中：k为旅客选择策略，取值0或1，1表示旅客决策为渗流旅客，0表示决策为非渗流旅客；φ为上一期吸引力指数对下一期决策的影响程度；$N(t)$为第t期的经验权重，其更新方式为

$$N(t)=\rho\cdot N(t-1)+1 \tag{2}$$

其中: ρ为旅客经验对下期决策的影响程度。

$\pi_i^k(t)$为第t期旅客i选择策略k的收益函数, 表示为

$$\pi_i^k(t)=\begin{cases}A^*-\overline{A}-G & k=1\\ -Q & k=0\end{cases} \tag{3}$$

通过融合吸引力指数对 Logit 模型进行改进, 计算旅客在不同策略下的选择概率$P_i^k(t+1)$为

$$P_i^k(t+1)=\frac{1}{1+\sum_{k=0}^{1}\theta^k}\cdot\frac{\mathrm{e}^{\lambda\cdot B_i^k(t)}}{\sum_{k=0}^{1}\mathrm{e}^{\lambda\cdot B_i^k(t)}}+\frac{\theta^k}{1+\sum_{k=0}^{1}\theta^k} \tag{4}$$

式中: θ^k为旅客对不同策略的依赖程度, 取值增加表示旅客对于选择相应航司的信念越强; 在博弈过程中, 旅客还会将本期航司提供的折扣票价$\overline{A}(t)$与心理价格$\widetilde{A}(t)$比较后作决策, λ反映了旅客的理性程度, 若折扣票价高于预期价格, 随差值增大, 旅客对吸引力指数的敏感程度越弱, 旅客理性程度越弱, 若折扣票价低于预期价格, 随差值增大, 旅客对吸引力指数的敏感程度越强, 旅客理性程度越强, 其更新公式为

$$\lambda=\frac{A^*-\overline{A}(t)}{A^*-\widetilde{A}(t)}\lambda_0 \tag{5}$$

其中: λ_0为旅客理性程度初值。

旅客会根据本期折扣票价及心理价格预测下一期的预期价格, 同时, $\widetilde{A}(t)$会受到季节、旅客出行需求变化、燃油波动等影响, 其更新公式为

$$\widetilde{A}(t+1)=\alpha\cdot\overline{A}(t)+(1-\alpha)\cdot\widetilde{A}(t) \tag{6}$$

式中: α为平滑系数, $0<\alpha<1$, 取值越大, 表明旅客受折扣票价的影响越大, 航司更容易通过票价折扣促使旅客实现由腹地旅客向渗流旅客的身份转变, 反之, 旅客对心理价格的信念越强, 航司需要付出更大的票价折扣来实现旅客转移。

由于旅客由原航线转移至新航线时会增加货币成本、时间成本等交通成本, 转移后其愿意支付的最高价格将会有所降低, 即旅客的预期价格必然会小于未转移前的机票价格; 同时, 所有旅客理性独立且根据自身利益最大化原则决策, 航司会通过折扣票价满足旅客利益不变或增加的目的, 促使其实现转移, 故而排除$A^*\geqslant\widetilde{A}$及$A^*\geqslant\overline{A}$。

2.2　基于航司竞争的 Hotelling 定价模型

当一条航线面临多家航司竞争时, 航司定价策略会受到彼此竞争的影响, 为此, 研究在相互竞争下航司定价策略。

博弈过程的假设如下:

H5: 基于 Hotelling 定价模型, 假定在长度为 1 的线性航空运输市场内, 相同的起点与终点间存在由不同航司运营的相同航线, R_0为既有航司, R_1为转移航司, R_0、R_1分别位于点 0、1 处。

H6: 航空运输市场内存在不同出行选择偏好的“渗流”旅客, 分别均匀分布在[0,1]区间内, “渗流”旅客 i 所在位置为x_i, 代表对不同航司航线的偏好程度, “渗流”旅客与航司之间的距离差代表损失效用, 反映了 2 类航司由于准点率、市场响应速度、品牌效应等潜在因素差异造成对旅客吸引能力不同的情况, 从而反映 Hotelling 定价模型中的“距离”因素。

H7: 假设市场内任何位置的“渗流”旅客均能掌握各航司的票价信息, 从而选择不同航司开设的航线出行。

H8: 对于既有航司而言, 其在该 OD 对航线内存在一定数量的腹地旅客N_0。

H9: “渗流”旅客在决策时, 航司间无定价差异, 如航司常旅客计划等。

“渗流”旅客选择不同航司的效用$U_0(t)$、$U_1(t)$分别为

$$U_0(t)=A_0(t)+E_0-(E_1-E_0)x \tag{7}$$

$$U_1(t)=A_1(t)+E_1-(E_1-E_0)(1-x) \tag{8}$$

式中: $A_0(t)$和$A_1(t)$分别为仿真第t期既有航司和转移航司的折扣票价; E_0和E_1分别为既有航司和转移航司可为旅客带来的相对潜在效用; x 为旅客在既有航司与转移航司形成的 [0,1] 效用分布区间相对既有航司的效用距离。

E_0和E_1为综合考虑准点率、市场响应速度、品牌效应、旅客服务质量等为指标, 采用熵权法修正的 TOPSIS 模型计算, 每项指标y对应的信息熵e_y为

$$e_y=-\frac{1}{\ln m}\sum_{y=1}^{m}\frac{Z_{\tilde{x}y}}{\sum_{\tilde{x}=1}^{m}Z_{\tilde{x}y}}\ln\left(\frac{Z_{\tilde{x}y}}{\sum_{\tilde{x}=1}^{m}Z_{\tilde{x}y}}\right) \tag{9}$$

式中: $Z_{\tilde{x}y}$为由指标数据建立原始指标体系矩阵经标准化至非负区间所得的标准化矩阵; $\tilde{x}=1,2,\cdots,m$, $y=1,2,\cdots,n$, m为评价目标总数, n为指标总数。

由信息熵得指标y的权重W_y为

$$W_y=\frac{1-e_y}{n-\sum_{y=1}^{n}e_y} \tag{10}$$

加权规范化矩阵H_{mn}为

$$H_{mn}=\left[h_{\tilde{x}y}\right]_{m\times n}=W_y\times Z_{\tilde{x}y} \tag{11}$$

根据加权规范化矩阵中每个指标的最大值

$h_y^{\max}$及最小值$h_y^{\min}$计算各节点航司的相对潜在效用值E_x，$0 \leqslant E_x \leqslant 1$，值越大，相应的航司可为旅客提供更高的潜在效用。

$$E_x = \frac{\sqrt{\sum_{y=1}^{n}\left(h_{\bar{x}y} - h_y^{\min}\right)^2}}{\sqrt{\sum_{y=1}^{n}\left(h_{\bar{x}y} - h_y^{\min}\right)^2} + \sqrt{\sum_{y=1}^{n}\left(h_{\bar{x}y} - h_y^{\max}\right)^2}} \tag{12}$$

如图2所示，存在位置x_m表示旅客选择开辟新航线的航司和既有航司出行的无差异效用点，即$U_0 = U_1$：

$$x_m(t) = \frac{A_0(t) - A_1(t)}{2(E_1 - E_0)} \tag{13}$$

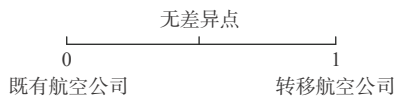


图2 Hotelling定价模型示意图

Fig. 2 Schematic diagram of the Hotelling model

当旅客位于$(0, x_m)$时，将会选择既有航司出行；当旅客位于$(x_m, 1)$时，将会选择转移航司出行。因此，既有航司和转移航司在该OD段航线的“渗流”旅客市场份额分别为

$$n_0 = \int_0^{x_m}\left(N^* \cdot P_i^{k=1}(t)\right)\mathrm{d}x = N^* \cdot P_i^{k=1}(t) \cdot x_m \tag{14}$$

$$n_1 = \int_{x_m}^{1}\left(N^* \cdot P_i^{k=1}(t)\right)\mathrm{d}x = N^* \cdot P_i^{k=1}(t) \cdot (1 - x_m) \tag{15}$$

式中：N^*为原航线旅客总人数。

因此，根据旅客的策略选择，既有航司收益Π_0和转移航司收益Π_1分别为

$$\Pi_0 = A_0(t) \cdot (n_0 + N_0) = A_0(t)\left[\left(N^* \cdot P_i^{k=1}(t) \cdot x_m\right) + N_0\right] \tag{16}$$

$$\Pi_1 = A_1(t) \cdot n_1 = A_1(t) \cdot N^* \cdot P_i^{k=1}(t) \cdot (1 - x_m) \tag{17}$$

式中：N_0为同一航线内既有航司的腹地旅客总人数。

在博弈过程中，航司会根据旅客转移情况及其与既有航司竞争的情况，综合评估下一期的降价策略。因此，票价调整策略为

$$A_0(t+1) = A_0(t) + \beta \cdot f^0(t)\Big[(A^* - A_0(t)) \cdot \left(1 + f^0(t)\right) + (A_0(t) - A_{\min})\left(1 - f^0(t)\right)\Big] \tag{18}$$

$$f^0(t) = \frac{P_0^{k=1}(t) - \widetilde{P}_0^{k=1}(t)}{\left|P_0^{k=1}(t) - \widetilde{P}_0^{k=1}(t)\right|} \quad P_0^{k=1}(t) \neq \widetilde{P}_0^{k=1}(t) \tag{19}$$

$$A_1(t+1) = A_1(t) + \beta \cdot f^1(t)\Big[(A^* - A_1(t)) \cdot \left(1 + f^1(t)\right) + (A_1(t) - A_{\min})\left(1 - f^1(t)\right)\Big] \tag{20}$$

$$f^1(t) = \frac{P_1^{k=1}(t) - \widetilde{P}_1^{k=1}(t)}{\left|P_1^{k=1}(t) - \widetilde{P}_1^{k=1}(t)\right|} \quad P_1^{k=1}(t) \neq \widetilde{P}_1^{k=1}(t) \tag{21}$$

式中：β为票价变化幅度；$A_{\min}$为航司可接受的票价底限；$f^0(t)$和$f^1(t)$分别为涨价和降价判别函数，取值为1或−1；$P_0^{k=1}(t)$和$P_1^{k=1}(t)$分别为第t期选择既有航司和转移航司的“渗流”旅客概率，两者相等时票价维持不变；$\widetilde{P}_0^{k=1}(t)$和$\widetilde{P}_1^{k=1}(t)$分别为第t期既有航司和转移航司预期吸引的“渗流”旅客概率，更新公式为

$$A_0(t) \cdot \left(\widetilde{P}_0^{k=1}(t) \cdot N^* + N_0\right) = C_0 A^* \gamma \tag{22}$$

$$A_1(t) \cdot \widetilde{P}_1^{k=1}(t) N^* = C_1 A^* \gamma \tag{23}$$

其中：γ为成本比例系数；C_0和C_1分别为既有航司和新入驻航司的载客能力。

3 下层博弈模型

该层是基于上层博弈模型航司最佳收益票价下，分析机场间协同补贴策略与航司航线调整情况的博弈过程。

3.1 模型假设

该演化博弈过程作出如下假设：

H10：由于是以研究航线网络优化（即航线转移）为目的，该层航线所属的航司均为上层的转移航司，且将上层最佳收益票价作为输入。Π_R^g表示航司R转移中程航线至机场g的最佳客票收益，$\tilde{\Pi}_R^g$表示航司R转移短程航线至机场g的最佳客票收益，其中，$g = 1,2$，$g = 1$表示中小型机场，$g = 2$表示大型机场。

H11：基于市场行为，航司会比较不同规模机场设置的机场补贴金额，自愿性选择转移中程航线和短程航线至大型机场或中小型机场，采取“转移中程航线至中小型机场”策略的概率为p，采取“转移中程航线至大型机场”策略的概率为$1-p(0 \leqslant p \leqslant 1)$，采取“转移短程航线至中小型机场”策略的概率为$q$，采取“转移短程航线至大型机场”策略的概率为$1-q(0 \leqslant q \leqslant 1)$。

H12：航司R将不同航线转移至机场g的主要成本包括转移成本D_R^g（转移中程航线）、$\tilde{D}_R^g$（转移短程航线），机场收费I^g，航线运行成本L_R（运营中程航线）、$\tilde{L}_R$（运营短程航线）。

H13：多家航司转移航线选择相同的机场时会共享机场补贴S，为避免航线过度聚集阻碍多机场协同化发展，机场补贴总额会随航线聚集减少至δS，δ为待转移航线系数，$0 < \delta < 1$。S_R^g为机场g对航司R中程航线补贴，$\tilde{S}_R^g$为机场g对航司R短程航

线补贴。

H14: 机场间协同补贴经费主要来自政府和机场自筹, 经调研, 机场自筹经费不会超过其收入的40%, 计算式为

$$S^g \leqslant 0.4I^g + \omega V^g \tag{24}$$

式中: I^g为航线业务收入; V^g为机场航线网络优化潜在收入与政府航线补贴; ω为差异化航线达成率。

3.2　构建效用函数与复制动态方程

根据 H10～H14, 航司不同转移策略对应的效益矩阵如表 2 所示。

表 2　航司转移策略效用矩阵

Table 2　Utility matrix of airline transfer strategies

策略种类	航司R优化中程航线	航司R优化短程航线
中小机场至中小机场	$\Pi_R^{g=1}+\delta S_R^{g=1}-D_R^{g=1}-L_R-I^{g=1}$	$\tilde{\Pi}_R^{g=1}+\delta\tilde{S}_R^{g=1}-\tilde{D}_R^{g=1}-\tilde{L}_R-I^{g=1}$
中小机场至大型机场	$\Pi_R^{g=1}+S_R^{g=1}-D_R^{g=1}-L_R-I^{g=1}$	$\tilde{\Pi}_R^{g=2}+\tilde{S}_R^{g=2}-\tilde{D}_R^{g=2}-\tilde{L}_R-I^{g=2}$
大型机场至中小机场	$\Pi_R^{g=2}+S_R^{g=2}-D_R^{g=2}-L_R-I^{g=2}$	$\tilde{\Pi}_R^{g=1}+\tilde{S}_R^{g=1}-\tilde{D}_R^{g=1}-\tilde{L}_R-I^{g=1}$
大型机场至大型机场	$\Pi_R^{g=2}+\delta S_R^{g=2}-D_R^{g=2}-L_R-I^{g=2}$	$\tilde{\Pi}_R^{g=2}+\delta\tilde{S}_R^{g=2}-\tilde{D}_R^{g=2}-\tilde{L}_R-I^{g=2}$

航司R选择转移中程航线至中小机场或大型机场的期望收益U_{R_1}、U_{R_2}及平均期望收益$\overline{U}_R$分别为

$$\begin{cases} U_{R_1}=\Pi_R^{g=1}+\left[q(\delta-1)+1\right]S_R^{g=1}-D_R^{g=1}-L_R-I^{g=1} \\ U_{R_2}=\Pi_R^{g=2}+\left[q(1-\delta)+\delta\right]S_R^{g=2}-D_R^{g=2}-L_R-I^{g=2} \\ \overline{U}_R=pU_{R_1}+(1-p)U_{R_2} \end{cases} \tag{25}$$

航司R转移中程航线的策略选择的复制动态方程为

$$F(p)=\mathrm{d}p/\mathrm{d}t=p\left(U_{R_1}-\overline{U}_R\right)=p(p-1)\left(U_{R_2}-U_{R_1}\right) \tag{26}$$

p的一阶导数为

$$F'(p)=(2p-1)\left(U_{R_2}-U_{R_1}\right) \tag{27}$$

根据微分方程稳定性定理, 航司 R 选择转移中程航线至中小机场的策略处于稳定状态必须满足: $F(p)=0$且$F'(p)<0$, 此时求得阈值q^*为

$$q^*=\frac{\Delta\Pi_R-\Delta D_R-\Delta I+S_R^{g=1}-\delta S_R^{g=2}}{(1-\delta)\left(S_R^{g=1}+S_R^{g=2}\right)} \tag{28}$$

$$\Delta\Pi_R=\Pi_R^{g=1}-\Pi_R^{g=2} \tag{29}$$

$$\Delta D_R=D_R^{g=1}-D_R^{g=2} \tag{30}$$

$$\Delta I=I^{g=1}-I^{g=2} \tag{31}$$

定理 1　当$q<q^*$时, 航司 R 转移中程航线的稳定策略是转移至中小型机场; 当$q>q^*$时, 航司 R 转移中程航线的稳定策略是转移至大型机场; 当$q=q^*$时, 不能确定航司 R 转移中程航线的稳定策略。

证明　令

$$K(q)=U_{R_2}-U_{R_1}=\Delta D_R+\Delta I-\Delta\Pi_R+\delta S_R^{g=2}-S_R^{g=1}+q(1-\delta)\left(S_R^{g=1}+S_R^{g=2}\right) \tag{32}$$

故$K(q)$为关于q的增函数, 当$q<q^*$时, $K(q)<0$, $F(p)|_{p=1}=0$且$F'(p)|_{p=1}<0$, 则$p=1$具有稳定性; 当$q>q^*$时, $K(q)>0$, $F(p)|_{p=0}=0$且$F'(p)|_{p=0}<0$, 则$p=0$具有稳定性; 当$q=q^*$时, $K(q)=0$, $F(p)=0$且$F'(p)=0$, 此时无法确定航司R转移中程航线的稳定策略。　**证毕**

航司 R 选择转移短程航线至中小型机场或大型机场的期望收益$U_{R_1'}$、$U_{R_2'}$及平均期望收益$\overline{U}_{R'}$分别为

$$\begin{cases} U_{R_1'}=\tilde{\Pi}_R^{g=1}+\left[p(\delta-1)+1\right]\tilde{S}_R^{g=1}-\tilde{D}_R^{g=1}-\tilde{L}_R-I^{g=1} \\ U_{R_2'}=\tilde{\Pi}_R^{g=2}+\left[p(1-\delta)+\delta\right]\tilde{S}_R^{g=2}-\tilde{D}_R^{g=2}-\tilde{L}_R-I^{g=2} \\ \overline{U}_{R'}=qU_{R_1'}+(1-q)U_{R_2'} \end{cases} \tag{33}$$

航司 R 转移短程航线的策略选择的复制动态方程为

$$F(q)=\mathrm{d}q/\mathrm{d}t=q\left(U_{R_1'}-\overline{U}_{R'}\right)=q(q-1)\left(U_{R_2'}-U_{R_1'}\right) \tag{34}$$

q的一阶导数为

$$F'(q)=(2q-1)\left(U_{R_2'}-U_{R_1'}\right) \tag{35}$$

根据微分方程稳定性定理, 航司 R 选择转移短程航线至中小机场的策略处于稳定状态必须满足: $F(q)=0$且$F'(q)<0$, 此时求得阈值p^*为

$$p^*=\frac{\Delta\tilde{\Pi}_R-\Delta\tilde{D}_R-\Delta I+\tilde{S}_R^{g=1}-\delta\tilde{S}_R^{g=2}}{(1-\delta)\left(\tilde{S}_R^{g=1}+\tilde{S}_R^{g=2}\right)} \tag{36}$$

$$\Delta\tilde{\Pi}_R=\tilde{\Pi}_R^{g=1}-\tilde{\Pi}_R^{g=2} \tag{37}$$

$$\Delta\tilde{D}_R=\tilde{D}_R^{g=1}-\tilde{D}_R^{g=2} \tag{38}$$

定理 2　当$p<p^*$时, 航司 R 转移短程航线的稳定策略为转移至中小机场; 当$p>p^*$时, 航司 R 转移短程航线的稳定策略为转移至大型机场; 当$p=p^*$时, 不能确定航司 R 转移短程航线的稳定策略。

证明　令

$$K(p)=U_{R_2'}-U_{R_1'}=\Delta\tilde{D}_R+\Delta I-\Delta\tilde{\Pi}_R+\delta\tilde{S}_R^{g-2}-\tilde{S}_R^{g=1}+p(1-\delta)\left(\tilde{S}_R^{g=1}+\tilde{S}_R^{g-2}\right)\tag{39}$$

故$K(p)$为关于p的增函数，当$p<p^*$时，$K(p)<0$，$F(q)|_{q=1}=0$且$F'(q)|_{q=1}<0$，则$q=1$具有稳定性；当$p>p^*$时，$K(p)>0$，$F(q)|_{q=0}=0$且$F'(q)|_{q=0}<0$，则$q=0$具有稳定性；当$p=p^*$时，$K(p)=0$，$F(q)=0$且$F'(q)=0$，此时无法确定航司R转移短程航线的稳定策略。**证毕**

通过解复制动态方程可得博弈的5个均衡点，分别为(0,0)、(0,1)、(1,0)、(1,1)、(p^*,q^*)，根据Lyapunov第一法则判别均衡点状态。根据复制动态方程构造的雅可比矩阵为

$$\boldsymbol{J}=\begin{bmatrix}\dfrac{\partial F(p)}{p} & \dfrac{\partial F(p)}{q}\\ \dfrac{\partial F(q)}{p} & \dfrac{\partial F(q)}{q}\end{bmatrix}=\begin{bmatrix}(2p-1)\left(U_{R_2}-U_{R_1}\right) & p(p-1)(1-\delta)\left(S_R^{g=1}+S_R^{g=2}\right)\\ q(q-1)(1-\delta)\left(S_{R'}^{g=1}+S_{R'}^{g=2}\right) & (2q-1)\left(U_{R_2'}-U_{R_1'}\right)\end{bmatrix}\tag{40}$$

各均衡点的稳定性分析如表3所示，根据稳定性分析确定博弈双方可采用的方案。

表3 均衡点的稳定性分析

Table 3 Stability analysis of equilibrium points

均衡解	特征值	稳定性
(0,0)	$\Delta\Pi_R-\Delta D_R-\Delta I+S_R^{g=1}-\delta S_R^{g=2}$ $\Delta\tilde{\Pi}_R-\Delta\tilde{D}_R-\Delta I+\tilde{S}_R^{g=1}-\delta\tilde{S}_R^{g=2}$	不确定
(0,1)	$\Delta\Pi_R-\Delta D_R-\Delta I-S_R^{g=2}+\delta S_R^{g=1}$ $-\Delta\tilde{\Pi}_R+\Delta\tilde{D}_R+\Delta I+\tilde{S}_R^{g=2}-\delta\tilde{S}_R^{g=1}$	不确定
(1,0)	$-\Delta\Pi_R+\Delta D_R+\Delta I-S_R^{g=1}+\delta S_R^{g=2}$ $\Delta\tilde{\Pi}_R-\Delta\tilde{D}_R-\Delta I+\tilde{S}_R^{g=1}-\delta\tilde{S}_R^{g=2}$	不确定
(1,1)	$-\Delta\Pi_R+\Delta D_R+\Delta I+S_R^{g=2}-\delta S_R^{g=1}$ $-\Delta\tilde{\Pi}_R+\Delta\tilde{D}_R+\Delta I+\tilde{S}_R^{g=2}-\delta\tilde{S}_R^{g=1}$	不确定
(p^*,q^*)	Ω_1 Ω_2	不确定

4 博弈模型仿真分析

上层博弈模型中，旅客会根据当期航司提供的折扣票价分析不同策略的收益情况，并根据预期价格及往期经验完成策略更新，同时航司会根据旅客的策略选择优化折扣票价，从而实现在收益最大化情况下的最佳定价策略；下层博弈模型中，在航司获得最佳客票收益时分析机场提供不同补贴金额的均衡点，从而确定机场间协同补贴策略与航线网络优化的最佳效果。

4.1 仿真案例

京津冀协同发展已上升为国家重大战略，以京津冀机场群为研究背景，根据城市发展定位及综合交通的整体发展情况，对京津冀机场群的航线网络进行优化调整。通过采取差异化补贴引导符合机场功能定位的航线转移，实现机场群航线的最优配置。仿真以津冀2地机场承接北京2场的溢出航线为目标，溢出航线分属2家不同规模的航司，中程航线与短程航线比例分为2∶1、2∶3，其中，中程航线24条，短程航线24条，共计航线48条。仿真以天津滨海国际机场为国内枢纽机场代表，石家庄正定国际机场为区域机场代表，将48条溢出航线转移至天津滨海国际机场与石家庄正定国际机场。天津滨海国际机场致力于吸引中程航线，石家庄正定国际机场致力于吸引短程航线。上下层仿真模型初始参数及敏感度如表4和表5所示。

根据中国铁路12306官方网站查询城际铁路出

表4 上层博弈模型初值设置

Table 4 Initial value setting of upper level game model

参数	参数基准值	变化后参数	参数变化率/%	旅客选择概率变化率/%		初值敏感指数/%	
				转移航司	既有航司	转移航司	既有航司
φ	0.5	0.75	50	−2.66	6.86	−5.32	13.72
ρ	0.5	0.25	−50	−1.16	3.14	2.32	−6.28
θ^k	25	50	100	−0.60	−1.01	−0.60	−1.01
α	0.5	0.75	50	0	0	0	0

行成本与高德地图 APP 计算城内网约车出行成本，按天津、河北 2 地的平均工资水平确定时间成本，具体参数如表 6 所示。

表 5　下层博弈模型初值设置

Table 5　Initial value settings for lower level game models

φ	ρ	θ^k	α
0.5	0.5	25	0.5

表 6　渗流旅客转移成本参数

Table 6　Passenger transfer cost parameters

渗流路线	城际出行成本/元	城际出行时间/min	城内出行成本/元	城内出行时间/min	时间成本/(元·h^{-1})
北京—天津	20～180	30～60	60	90	60
北京—石家庄	40～230	60～130	40	70	60

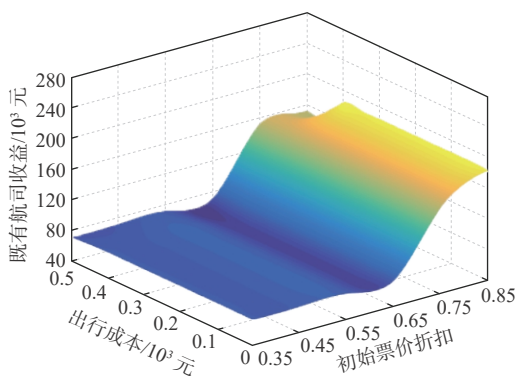
4.2　上层博弈模型求解

4.2.1　票价折扣及出行成本对旅客选择行为的影响

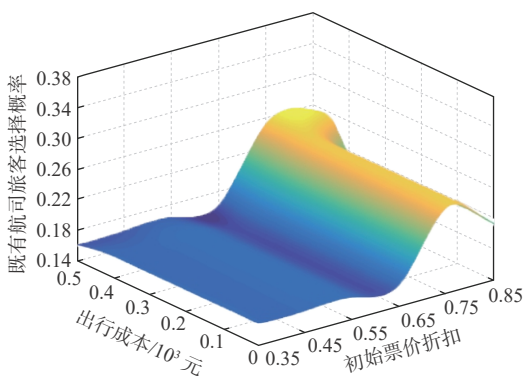
图 3(a)、图 3(c)为在全票价为 1 500 元(国内航线代表性票价)时，既有航司与转移航司在不同票价折扣和旅客不同出行成本下的收益曲线。2 类航司在初始票价折扣为 0.75～0.85 之间均存在收益峰值，其票价折扣为 0.75，小于该值(即高票价折扣、低折扣票价)时，由于高票价折扣导致 2 类航司的收益在不同旅客出行成本下都出现了下降。从旅客出行成本看，较低出行成本有利于提升航司收益，且票价折扣与旅客出行成本成反比，即较低的票价折扣对出行成本增加影响较小；当票价折扣为 0.75～0.85 之间时，旅客出行成本控制在 400 元以内有利于航司收益增加。

图 3(b)、图 3(d)为在全票价为 1 500 元时，初始票价折扣、旅客出行成本对旅客选择既有航司、转移航司的概率曲线。可知，从旅客选择概率波动程度看，转移航司的波动幅度为 0.24～0.35，既有航司的波动幅度为 0.15～0.26；0.75 以上的低票价折扣区间为既有航司吸引“渗流”旅客的优势区间，相较于 0.35～0.60 的高票价折扣区间旅客选择概率较高；0.60 以下的高票价折扣区间为转移航司吸引“渗流”旅客的优势区间，相较于 0.65～0.85 的折扣区间对旅客的吸引力更强。整体来看，2 类航司的旅客选择概率波动幅度相同，转移航司和既有航司分别在高票价折扣区间和低票价折扣区间具有吸引“渗流”旅客优势，两者形成“跷跷板”效应。

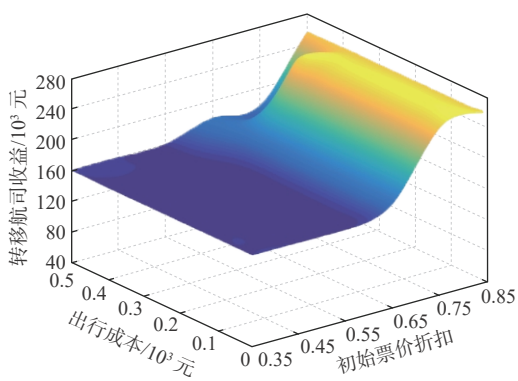
综合航司收益和旅客选择概率可知，对于转移航司而言，维持高收益的主要影响因素为高折扣票价，较高的售价可有效弥补旅客人数减少所带来的损失，并对收益产生显著作用，但同时转移航司在高票价折扣区间内凭借着旅客资源的竞争优势也



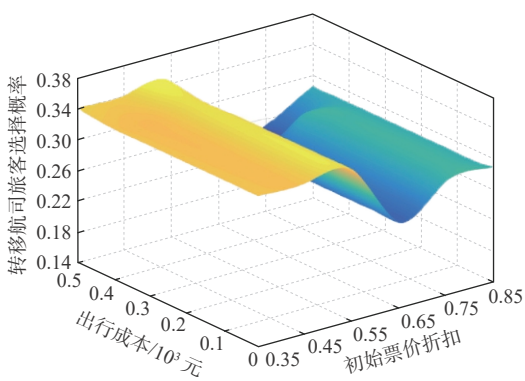
(a) 既有航司在不同初始票价折扣和旅客不同出行成本下的收益曲线



(b) 初始票价折扣、旅客出行成本对旅客选择既有航司影响的概率曲线



(c) 转移航司在不同初始票价折扣和旅客不同出行成本下的收益曲线



(d) 初始票价折扣、旅客出行成本对旅客选择既有航司影响的概率曲线

图 3　航司初始票价折扣与收益、旅客选择概率关系

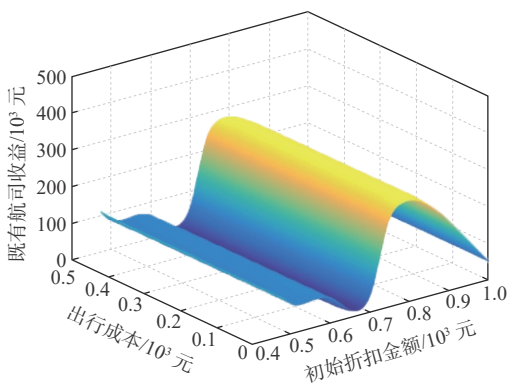
Fig. 3　The relationship between initial ticket price discount and revenue, passenger selection probability of airlines

获得了较为可观的收益；对于既有航司，维持高收益的主要影响因素为航司可为旅客提供的较高潜在效用，较高的潜在效用可在票价折扣较低的情况下吸引旅客选择，显著的价格优势及旅客资源的竞争优势为既有航司带来了较高的收益。对比2类航司的收益可知，在0.6～0.7折扣区间内，既有航司和转移航司出现了明显的收益增长，因此，为实现航司间收益最佳，转移航司与既有航司票价折扣要维持在0.6～0.85之间，以避免出现低价竞争带来的收益共损。

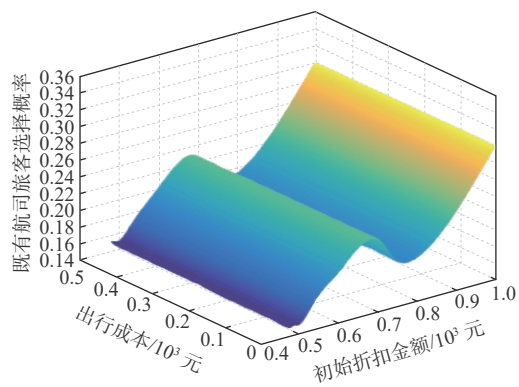
4.2.2 不同折扣金额与出行成本对旅客选择行为的影响

图4表示在不同的初始票价下转移航司和既有航司初始折扣分别为0.5和0.6时，票价与出行成本、旅客选择概率间关系。在固定折扣时，航程越长，机票价格越高，相同折扣的对应优惠金额就越大。由图4(a)、图4(c)可知，初始票价对应的折扣金额在400～700元(即初始票价较低)、航程较短时，转移航司的收益与航程成负相关，即航程越短，收益越高，而既有航司收益基本保持不变，对于非腹地旅客的吸引力微弱，仅靠腹地旅客来维持较低的收益；当初始票价对应折扣金额为800元时，2类航司的收益都出现了明显的峰值，且在700～800元的折扣金额区间内，转移航司的收益增长率要明显高于既有航司；当初始票价对应折扣金额在800元以上(即初始票价较高)、属于长航程航线时，2类航司收益均存在下降趋势，随着航程增长，收益减少。

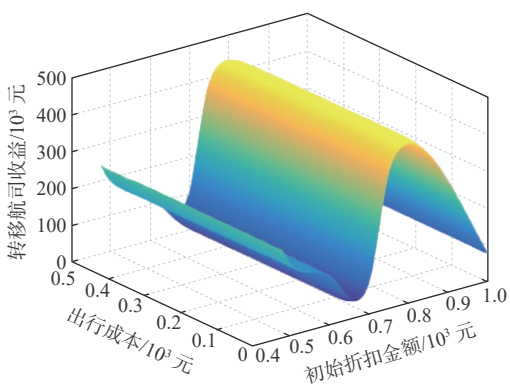
由图4(b)、图4(d)可知，既有航司初始折扣金额在900元以上(即长航程航线)时，对非腹地旅客具有较强吸引力，而转移航司初始折扣金额在400～800元区间(对应中短程航线)时，对非腹地旅客吸引能力最强，可知，在优惠折扣固定的情况下，随着航程的增长，非腹地旅客在选择航司出行时会更注重航司可为其提供的潜在效用为较长航程提供保障；对于既有航司，航司的收益和旅客选择概率变化趋势相反，说明机票售价对于维持既有航司的高收益至关重要，即使初始折扣金额增大吸引了更多的腹地旅客选择，但过大的折扣金额对于航司仍有不可忽视且难以抵消的削减作用；对于转移航司，航司的收益和旅客选择概率呈现了类似的变化趋势，说明非腹地旅客的转移情况是影响收益的关键因素。整体来看，在转移航司和既有航司初始折扣分别为0.5和0.6的情况下，二者对非腹地旅客的吸引优势形成了“跷跷板”效应，既有航司和转移航司分别在长航程和短航程时存在吸引非腹地旅



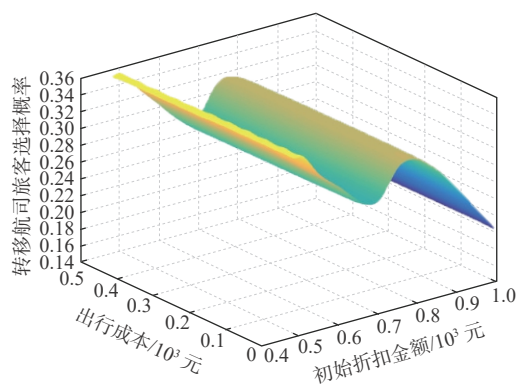
(a) 既有航司在不同初始折扣金额和旅客不同出行成本下的收益曲线



(b) 初始折扣金额、旅客出行成本对旅客选择既有航司影响的概率曲线



(c) 转移航司在不同初始折扣金额和旅客不同出行成本下的收益曲线



(d) 初始折扣金额、旅客出行成本对旅客选择既有航司影响的概率曲线

图4 航司初始折扣金额与收益、旅客选择概率关系

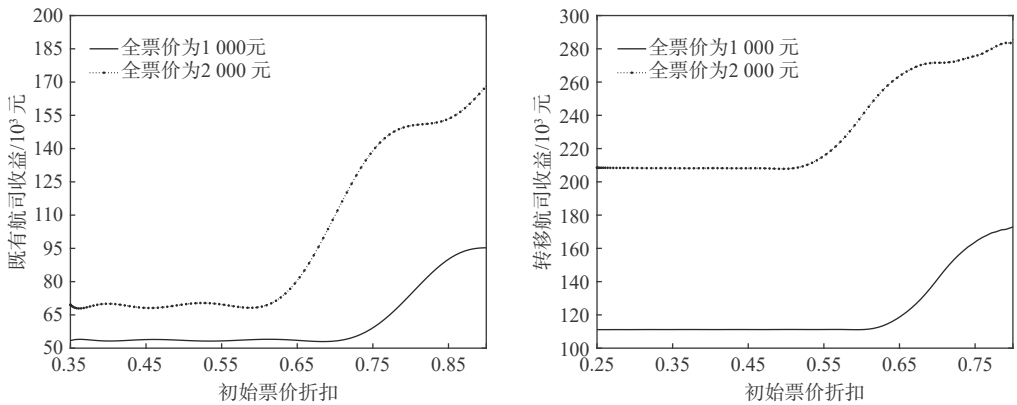
Fig. 4 The relationship between initial discount amount of airlines and revenue, passenger selection probability

客的竞争优势。

进一步表明航司对市场的熟悉程度是影响航司收益变化情况的关键因素, 也是航司根据不同航线特征选择折扣比例的依据, 既有航司在市场内已存在部分腹地旅客的支撑, 故其对票价金额的敏感性更强, 而转移航司在增加了航线转移成本的压力下, 为维持公司的可持续发展对旅客选择概率变化更敏感。同时, 曲线的波动体现旅客对 2 类航司的选择概率与票价之间的矛盾, 只有实现航司的航线开设、票价折扣与航司自身实际情况之间的最佳匹配才能达到航司收益的峰值。

根据京津冀协同发展战略部署对京津冀“三地四场”的定位, 北京大兴国际机场和北京首都国际机场进行航线网络优化, 按照国内航线里程分类标准: 短程航线为小于 1 000 km, 中程航线为 3 000 km 左右, 长程航线为 11 000 km 左右。其中, 长程航线一般用于国际航线且起止机场相对固定, 故不考虑其在机场群内的航线转移, 短程航线和中程航线相应初始票价主要为 1 000 元和 2 000 元, 作为下层博弈模型航司的中短程航线在机场采取不同补贴策略下转移的输入参数。

由图 5 可知, 初始票价为 1 000 元、票价折扣为 0.90, 以及初始票价为 2 000 元、票价折扣为 0.80 时, 既有航司收益较高; 初始票价为 1 000 元、票价折扣为 0.80, 以及初始票价为 2 000 元、票价折扣为 0.67 时, 转移航司收益较高。



(a) 既有航司在不同初始票价折扣下的收益曲线　　(b) 转移航司在不同初始票价折扣下的收益曲线

图 5　航司最佳定价策略

Fig. 5　Best pricing strategy for airlines

4.3　下层博弈模型求解

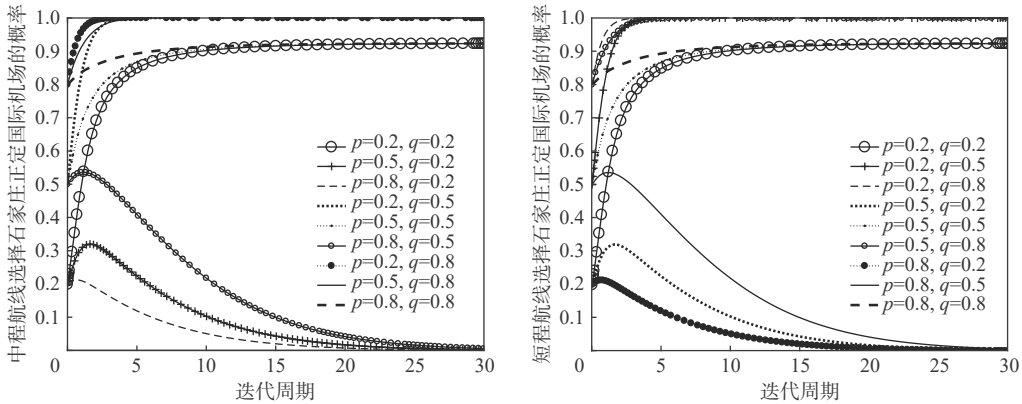
为验证各均衡点的稳定性, 现结合实际情况对模型进行赋值, 并利用 MATLAB 2016b 进行数值仿真。航班运营成本包括固定成本及变动成本, 以每年飞行 650 h 计算, A320 的运营成本为 7 416 美元/h (约 53 618 元/h), 取中小机场航线变动成本为运营成本的 19%, 大型机场航线变动成本为运营成本的 21%, 即$D_R^{g=1}=D_{R'}^{g=1}=10\,187.42$/h, $D_R^{g=2}=D_{R'}^{g=2}=11\,259.78$元/h; 根据民航发 [2017]18 号《关于印发民用机场收费标准调整方案的通知》, 以 A320 最大起飞重量 73.5 t、容纳乘客 150 人、业载 12 t 计算, 一类一级机场(北京首都国际机场、上海浦东国际机场、广州白云国际机场 3 个机场)的内地航司内地航班航空性业务收费项目的机场收费为 8 900 元/架次, 二级机场的内地航司内地航班航空性业务收费项目的机场收费为 10 655 元/架次。因此, 可以判断$\left(\Pi_R^{g=2}-\Pi_R^{g=1}\right)+\left(D_R^{g=1}-D_R^{g=2}\right)+\left(I^{g=1}-I^{g=2}\right)<0$, $\left(\Pi_{R'}^{g=2}-\Pi_{R'}^{g=1}\right)+\left(D_{R'}^{g=1}-D_{R'}^{g=2}\right)+\left(I^{g=1}-I^{g=2}\right)<0$。

4.3.1　策略选择的初始概率对博弈均衡的影响

为分析博弈双方策略选择的初始概率对演化博弈过程及结果的影响, 取$S^{g=1}=10\,000$ 元($S_R^{g=1}=S_{R'}^{g=1}=5\,000$ 元), $S^{g=2}=6\,000$ 元($S_R^{g=2}=S_{R'}^{g=2}=3\,000$ 元), $\delta=0.7$, 复制动态方程组随时间演化 30 次。由图 6 可知, 若博弈初始短程航线转移至中小型机场的倾向程度强于中程航线($q>p$), 则博弈的稳定点为$E_{SS}(0,1)$; 若博弈初始中程航线转移至中小型机场的倾向程度强于短程航线($p>q$), 则博弈的稳定点为$E_{SS}(1,0)$; 若博弈初始中程航线与短程航线的策略选择概率相同($p=q$), 则博弈的稳定点为$E_{SS}(0.92,0.92)$。因此, 在 2 类航线转移策略初始选择概率不同时, 若竞争对手对策略的倾向程度高于博弈者本身时, 博弈者会放弃该策略采取其他航线转移策略。因此, 为实现航线网络内的航线优化, 机场需根据自身功能定位判断不同航程航线的补贴需求, 理性设置机场补贴以吸引相应的航线转移。

4.3.2　待转移航线系数δ对演化博弈的影响

为分析待转移航线系数δ的变化对演化博弈过程及结果的影响, 对δ分别赋值为 0.5、0.7、0.9, 取$p=0.5$(即中程航线)随机选择航线转移策略时, 复制动态方程组随时间演化 30 次结果如图 7 所示。



(a) 中程航线选择大型机场策略选择的初始概率对博弈均衡的影响曲线

(b) 短程航线选择中小型机场策略选择的初始概率对博弈均衡的影响曲线

图6 策略选择的初始概率对演化博弈的影响

Fig. 6 The impact of initial probability of strategy selection on evolutionary games

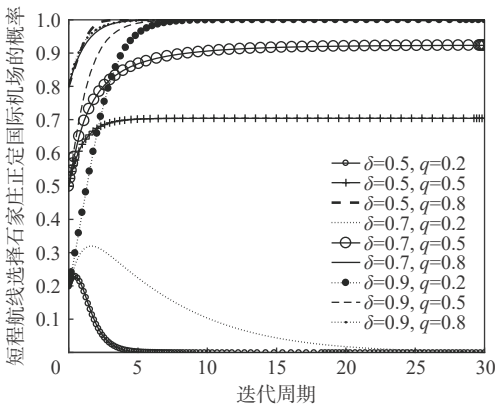


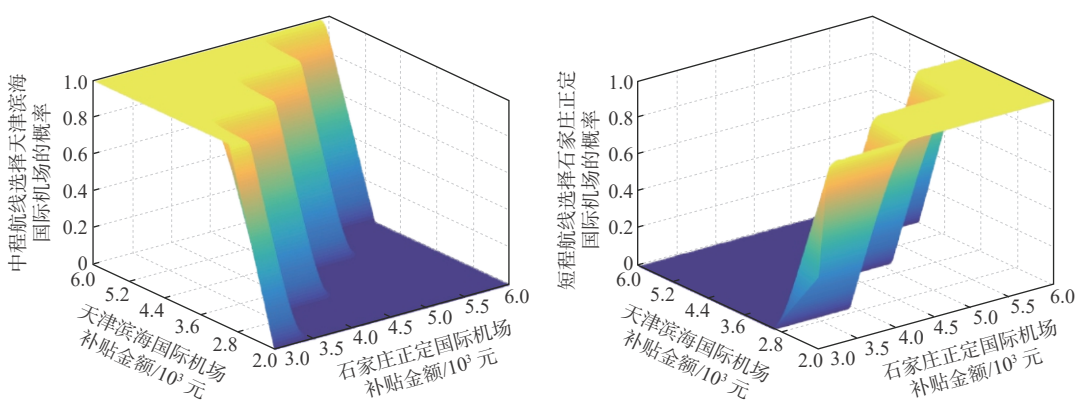
图7 待转移航线系数对演化博弈的影响

Fig. 7 The impact of pending transfer route coefficient on evolutionary games

可知，当短程航线转移至中小型机场倾向程度较高时（$q>0.8$），待转移航线系数对演化速度影响微弱且不会影响演化博弈的结果；当短程航线随机选择航线转移策略时（$q=0.5$），随待转移航线系数增加，短程航线转移至中小型机场的混合策略选择概率值明显提升，直至最终由混合策略变为纯策略选择；当短程航线博弈初始倾向转移至大型机场时（$q<0.5$），待转移航线系数的增加有利于减缓短程航线向大型机场转移的演化速度，甚至改变演化博弈结果。因此，为疏解大型机场的非核心功能，实现机场群航线网络的优化，可适当提升待转移航线系数、放宽机场补贴对短程航线的限制程度，以吸引短程航线转移至中小型机场。

4.3.3 机场补贴策略对演化博弈的影响

为分析2类机场不同补贴策略对演化博弈结果的影响，针对不同机场补贴策略，利用复制动态方程组随时间演化100次，结果如图8所示。图8(a)为大型机场以吸引中程航线为目标，在中小型机场不同机场补贴金额的情况下，大型机场不同机场补贴金额对中程航线的吸引情况。可知，当中小型机场及大型机场的补贴范围位于$S_R^{g=2}=1.02S_R^{g=1}-0.9$与$S_R^{g=2}=1.36S_R^{g=1}-0.81$这2条平面之间时，中程航线的策略选择为混合策略，且曲面处于陡增范围，2类机场的竞争激烈；当$S_R^{g=2}>1.4S_R^{g=1}-1.19$时，大型机场补贴处于绝对优势区间，可根据机场自身功能定位吸引中程航线转移；而$S_R^{g=2}=1.02S_R^{g=1}-0.9$时，大型机场对中程航线的吸引力微弱，几乎不存



(a) 大型机场对中程航线的吸引情况

(b) 中小型机场对短程航线的吸引情况

图8 机场补贴对演化博弈的影响

Fig. 8 The impact of airport subsidies on evolutionary games

在竞争力。为此, 考虑中小机场补贴策略的影响, 大型机场的最佳补贴策略应在 $S_R^{g=2} > 1.4S_R^{g=1} - 1.19$ 范围内, 且根据中小型机场具体补贴金额进行调整, 以较低的机场补贴实现符合自身功能定位航线的高效转移。

图 8(b)为中小机场以吸引短程航线为目标, 在大型机场不同机场补贴金额的情况下, 中小机场不同机场补贴金额对短程航线的吸引情况。可知, 在吸引短程航线时, 中小机场与大型机场的竞争激烈区间在 $S_{R'}^{g=2} > 0.82S_{R'}^{g=1} - 0.46$ 与 $S_{R'}^{g=2} < 1.06S_{R'}^{g=1} - 2.33$ 这 2 个平面之间的补贴范围内。当 $S_{R'}^{g=2} > 0.82S_{R'}^{g=1} - 0.46$ 时, 大型机场对于短程航线的吸引力更强, 具有绝对优势; 而 $S_{R'}^{g=2} < 1.06S_{R'}^{g=1} - 2.33$ 时, 中小机场处于吸引短程航线的绝对优势区间, 吸引的航线符合自身的功能定位。因此, 考虑到大型机场补贴策略的影响, 中小机场的最佳补贴策略应在 $S_{R'}^{g=2} > 0.82S_{R'}^{g=1} - 0.46$ 范围内, 且根据大型机场的具体补贴金额进行灵活调整, 通过较低的机场补贴实现符合自身功能定位航线的高效转移。

5　结　论

本文通过构建双层博弈模型, 以旅客、航司及机场三方分别作为上下层演化博弈的主体, 研究旅客与航司、航司与机场之间的演化稳定策略, 得到以下结论:

1) 航司在转移航线时, 通过折扣票价来吸引非腹地旅客, 在相同航线(认为全票价相同)与既有航司进行票价折扣竞争时, 应避免打“价格战”, 理性的高折扣票价可以使其通过较高的售价有效弥补旅客人数减少所带来的损失, 并对收益产生显著提升, 0.75 以下的低票价折扣区间是其吸引“渗流”旅客的优势区间。

2) 针对同一航线, 航司间在通过票价争夺旅客时, 为实现航司间收益最佳, 转移航司与既有航司票价折扣要维持在 0.6～0.85 之间, 以避免出现低价竞争带来的收益共损。

3) 不同机场间的机场补贴策略位差优势明显, 在以吸引中程航线为目标时, 大型机场的最佳补贴策略应在 $S_R^{g=2} > 1.4S_R^{g=1} - 1.19$ 范围内; 在以吸引短程航线为目标时, 中小机场的最佳补贴策略应在 $S_{R'}^{g=2} < 1.06S_{R'}^{g=1} - 2.33$ 范围内, 同时, 机场在进行机场补贴策略选择时, 应考虑到自身功能定位兼顾竞争对手的补贴金额情况灵活调整, 以较低的机场补贴促进多机场航线网络优化, 实现协同发展。

参考文献（References）

[1] SU M, LUAN W X, YUAN L Y, et al. Sustainability development of high-speed rail and airline: understanding passengers' preferences: a case study of the Beijing–Shanghai corridor[J]. Sustainability, 2019, 11(5): 1352.

[2] TERAJI Y, MORIMOTO Y. Price competition of airports and its effect on the airline network[J]. Economics of Transportation, 2014, 3(1): 45-57.

[3] DOBRUSZKES F, GIVONI M, VOWLES T. Hello major airports, goodbye regional airports? recent changes in European and US low-cost airline airport choice[J]. Journal of Air Transport Management, 2017, 59: 50-62.

[4] 周蔷, 王忠群, 龚本刚. 基于航线网络的定价、超售与舱位控制策略研究[J]. 兰州大学学报 (自然科学版), 2016, 52(1): 126-134.
ZHOU Q, WANG Z Q, GONG B G. Strategy research on the pricing, overbooking and cabin inventory control of airline network[J]. Journal of Lanzhou University (Natural Sciences), 2016, 52(1): 126-134(in Chinese).

[5] PENG I C, LU H A. Coopetition effects among global airline alliances for selected Asian airports[J]. Journal of Air Transport Management, 2022, 101: 102193.

[6] 骆嘉琪, 匡海波, 冯涛, 等. 基于两阶段博弈模型的高铁民航竞合关系研究[J]. 系统工程理论与实践, 2019, 39(1): 150-164.
LUO J Q, KUANG H B, FENG T, et al. Research on the co-opetition between high speed rail and civil aviation based on two stage game model[J]. Systems Engineering-Theory & Practice, 2019, 39(1): 150-164(in Chinese).

[7] KIM A M, RYERSON M S. A long drive: Interregional airport passenger “leakage” in the U. S[J]. Tourism Management, 2018, 65: 237-244.

[8] LIAO W, CAO X S, LI S C. Competition and sustainability development of a multi-airport region: a case study of the Guangdong-Hong Kong-Macao Greater Bay Area[J]. Sustainability, 2019, 11(10): 2958.

[9] 卞骞, 宋丽英, 毛保华, 等. 基于市场进入博弈的高速铁路与航空竞争分析[J]. 交通运输系统工程与信息, 2019, 19(5): 20-27.
BIAN Q, SONG L Y, MAO B H, et al. High speed rail and air transport competition: a market-entry game approach[J]. Journal of Transportation Systems Engineering and Information Technology, 2019, 19(5): 20-27(in Chinese).

[10] YANG C W, LU J L, HSU C Y. Modeling joint airport and route choice behavior for international and metropolitan airports[J]. Journal of Air Transport Management, 2014, 39: 89-95.

[11] MINATO N, MORIMOTO R. Designing the commercial sustainability of unprofitable regional airports using system dynamics analysis [J]. Research in Transportation Business & Management, 2011, 1(1): 80-90.

[12] HOU M, WANG K, YANG H J. Hub airport slot re-allocation and subsidy policy to speed up air traffic recovery amid COVID-19 pandemic-case on the Chinese airline market[J]. Journal of Air Transport Management, 2021, 93: 102047.

[13] 陈欣, 张珍, 邱瑞, 等. 基于动态博弈的补贴模式对机场群航线网络结构的影响研究[J]. 四川大学学报 (自然科学版), 2022, 59(5): 201-207.
CHEN X, ZHANG Z, QIU R, et al. Research on the impact of subsidization on the airline networks of multi-airport systems based on dynamic game theory[J]. Journal of Sichuan University (Natural

Science Edition), 2022, 59(5): 201-207(in Chinese).

[14] 张珍. 基于博弈论的机场群管理模式和航线网络选择策略研究[D]. 南京: 南京财经大学, 2023.
ZHANG Z. Research on airport group management mode and route network selection strategy based on game theory[D]. Nanjing: Nanjing University of Finance & Economics, 2023(in Chinese).

[15] 魏夕凯, 马本. 农村生活垃圾分类治理的奖惩激励机制: 基于复杂网络演化博弈模型[J]. 中国环境科学, 2022, 42(8): 3822-3831.
WEI X K, MA B. Reward and punishment incentive mechanism of domestic waste classification in rural China: based on complex network evolutionary game model[J]. China Environmental Science, 2022, 42(8): 3822-3831(in Chinese).

Evolutionary game of subsidy strategy on multi-airport route network under homogeneous competition

WU Wei[1,*], LIN Zhiyi[2], WANG Xinglong[1]

(1. Key Laboratory of Internet of Aircraft, Civil Aviation University of China, Tianjin 300300, China;
2. School of Electronic Information Engineering, Beihang University, Beijing 100191, China)

Abstract: To enhance the high-quality collaborative development of multiple airports in the region based on differentiated positioning, the impact of airport subsidy strategies on the evolution of the multi-airport route network was studied, and the optimal subsidy strategy was determined. A double-layer evolutionary game model was constructed based on the competitive game relationships among passengers, airlines, and airports. In the upper-level model, the influence of passengers' self-learning ability on fares was considered, and a passenger Logit choice model incorporating a self-learning mechanism was constructed. The Hotelling model was then used to analyze the impact of fare competition among airlines on the same route on passenger choice behavior, thereby determining the optimal pricing strategy for airlines under competitive conditions. In the lower-level model, the evolutionary game process of competitive choices between airport subsidies and airlines was analyzed based on the replicator dynamic equation. The synergistic effect of the inter-airport collaborative subsidy strategy and the route network was determined.The results show that for airlines transferring routes, the advantageous discount range for attracting leakage passengers is between 0.6 and 0.75. The fare discounts for airlines competing on the same route should be concentrated between 0.6 and 0.85 to avoid the loss of revenue caused by low price competition. Additionally, the optimization of the route network through airport subsidies reveals that each airport has an optimal subsidy range based on its functional positioning.

Keywords: air transportation; differentiated subsidy strategy; route network optimization; double-layer evolutionary game; replicated dynamic equation

Received: 2024-06-21; **Accepted**: 2024-08-16; **Published Online**: 2024-09-13 10:00

URL: link.cnki.net/urlid/11.2625.V.20240912.1744.004

Foundation item: The Fundamental Research Funds for the Central Universities of Civil Aviation University of China Special Project (3122025098)

* **Corresponding author**. E-mail: wwu@cauc.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0580

基于高精度算法的二级轻气炮性能分析

项高翔[1, 2, 3, *]，伍佳豪[1]，吴梦佳[1]

(1. 西北工业大学 力学与土木建筑学院，西安 710072；2. 西北工业大学深圳研究院，深圳 518063；
3. 西北工业大学重庆科创中心，重庆 401135)

摘 要：针对超高速发射技术中弹丸速度与能量转换效率的优化问题，采用经典内弹道模型对一维非定常可压缩流动下药室内火药燃烧状态、活塞运动和气室内气体流动状态、弹丸运动进行研究。药室部分通过四阶 Runge-Kutta 算法求解，气室部分求解 Euler 方程并采用 WENO 格式进行高精度激波捕捉。分别对不同初始注气条件、气室尺寸、发射管尺寸、火药质量下的二级轻气炮进行数值模拟，得到相关参数对轻气炮性能影响规律。研究发现：泵管内轻质气体能明显提高轻气炮的性能，气室尺寸、管径及长度对活塞底部压力和速度影响较大，但对弹丸速度影响较小；增加火药质量、发射管尺寸及管径可显著增大弹丸出膛速度；为新型二级轻气炮的设计制造提供重要理论和数据支撑。

关 键 词：二级轻气炮；内弹道；数学模型；WENO 格式；数值模拟
中图分类号：V221$^+$.3; TB553
文献标志码：A **文章编号**：1001-5965(2025)10-3405-10

早期的火炮产生的大摩尔质量的火药气体限制了其弹丸的速度(3 km/s 以下)，弹丸在被火药气体膨胀加速的过程中，很大一部分能量用来加速气体本身并导致最终作用在弹丸上的能量减少了很多[1]。为提高弹丸速度，美国于 1946 年研制出了第一门由氦气和氢气作为推进气体的轻气炮[2]。20 世纪 50 年代美国国家航空咨询委员会(National Advisory Committee for Aeronautics, NACA)利用轻气炮对火箭和再入飞行器进行研究，在此之间利用氦气可以将目标加速到 4.5 km/s[3]。此后，NACA 及其继任者，美国国家航空航天局(National Aeronautics and Space Administration, NASA)对轻气炮进行了 2 个方面的加强改进：①将氦气驱动气体用氢气代替[4-5]；②在泵管高压尾端增加一个收缩的锥形截面[6]。接着，美国、加拿大和英国的许多弹道研究所又对轻气炮进行了改进[4-7]，使这项高速发射技术得到了迅速发展，仅用了 13 年时间就将过去火药炮所能达到的最高发射速度 3 km/s 提高到 9.144 km/s。几十年来，国际上研制了各种各样的模型发射器，主要包括单级炮、二级轻气炮等，被广泛的应用在超高速空气动力学[8-10]、超高速气动物理学，超高速撞击现象，如空间碎片和微流星体对航天器部件的撞击模拟、超高压环境下材料的动态载荷的研究中。

二级轻气炮是采用火药燃烧，驱动气室中的活塞运动，经过压缩轻质气体在高压锥段处形成高压高声速的驱动介质，从而进一步驱动弹丸实现更高速度加载的一种高速火炮。目前对二级轻气炮的计算除实验外，还有理论近似计算、特征线法和数值模拟[11-13]。由于计算机技术的快速发展[14-17]，二级轻气炮研究领域开始大量采用数值模拟方法进行模拟计算，陈大年[18]采用准一维不定常流的 Lagrangian 算法对二级轻气炮运行气动过程进行分析

收稿日期：2023-09-12；**录用日期**：2023-10-14；**网络出版时间**：2023-11-17 12:50
网络出版地址：link.cnki.net/urlid/11.2625.V.20231116.1504.003
基金项目：国家自然科学基金 (12202365)；中国博士后科学基金 (2021M692633,2022T150534)；广东省基础与应用基础研究基金 (2022A1515011565,2023A1515010031)；重庆市自然科学基金 (2022NSCQ-MSX5709)
* **通信作者**. E-mail：xianggx@nwpu.edu.cn

引用格式：项高翔，伍佳豪，吴梦佳. 基于高精度算法的二级轻气炮性能分析 [J]. 北京航空航天大学学报，2025，51（10）：3405-3414.
XIANG G X，WU J H，WU M J. Performance analysis of two-stage light gas gun based on higher-order code[J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51（10）：3405-3414（in Chinese）.

并归纳了影响轻气炮的一些基本规律。文献[19-21]运用经典Runge-Kutta方法求解药室方程，二阶MacCormack格式和Harten二阶TVD格式求解轻气室方程，通过2部分计算交替进行，实现了二级轻气炮发射过程的数值模拟，并分析不同差分格式对计算结果的影响。刘海明[22]建立了二级轻气炮药室部分的一维两相流内弹道物理数学方程，固相采用MacCormack差分格式，气相采用TVD格式求解方程，模拟了整个二级轻气炮的内弹道过程。朱斌[23]建立7.62 mm口径步枪膛口流场二维模型，采用基于非结构网格的有限体积法，利用Fluent软件对其进行计算，得到了与实验照片比较吻合的计算结果。

本文采用四阶Runge-Kutta算法和高精度WENO激波捕捉格式，分别建立药室和气室气体压力动态模型，WENO格式可以有效捕捉激波间断，具有较高计算精度，且方法简单、鲁棒性好、计算效率高，是目前最流行的激波捕捉算法。通过对本文建立的数学模型求解，得到了不同初始注气条件、泵管尺寸、发射管尺寸、火药质量下的弹道性能。

1 数学物理模型及计算方法

1.1 数学物理模型

二级轻气炮的内弹道问题十分复杂，因为其涉及2个气室内的压力变化、一级活塞和弹丸的运动规律等诸多问题。图1为一个火药驱动活塞型二级轻气炮的示意图，电发火头点燃火药，药室火药燃气压力达到活塞启动压力C_{p0}时，药室和气室之间的膜片破裂，燃气推动活塞并压缩泵管气室内的轻气体，当轻气室压力达到弹丸膜片破裂压力O_{p}时，气室和发射管之间的膜片破裂，弹丸开始运动。活塞、弹丸同时向前运动一段时间后，活塞先挤入压缩管与发射管间的锥形过渡段而停止运动，弹丸继续沿发射管向前运动直至从炮口射出。

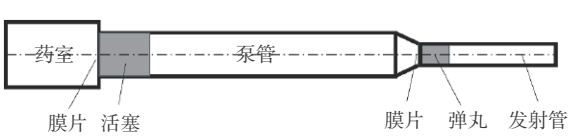


图1 二级轻气炮示意图

Fig. 1 Diagrammatic sketch of secondary light-gas gun

根据上述描述，提出如下假设：

1）气体为理想气体；

2）忽略轻气与管道的热损失和摩擦损失(由次要功系数计算)；

3）假定气室气体余容为常数。

1.2 计算方法

1）药室部分。根据经典火药燃烧理论，在1.1节假设条件下可得[24-25]：

$$\begin{cases}\dfrac{\partial\psi}{\partial t}=\begin{cases}x(1+2\lambda z+3\mu z^2)\dfrac{u_1}{e_1}p_{\mathrm{s}}^n & z<1\\ x_{\mathrm{s}}(1+2\lambda_{\mathrm{s}}z)\dfrac{u_1}{e_1}p_{\mathrm{s}}^n & 1\leqslant z\leqslant Z_k\\ 0 & z\geqslant Z_k\end{cases}\\ \dfrac{\mathrm{d}z}{\mathrm{d}t}=\begin{cases}\dfrac{u_1}{e_1}p_{\mathrm{s}}^n & z<Z_k\\ 0 & z\geqslant Z_k\end{cases}\\ \dfrac{\mathrm{d}v_{\mathrm{p}}}{\mathrm{d}t}=\dfrac{S(p_{\mathrm{s}}-p_{\mathrm{ps}})}{\varphi_1 m_{\mathrm{p}}}\\ \dfrac{\mathrm{d}l_{\mathrm{p}}}{\mathrm{d}t}=v_{\mathrm{p}}\\ S\,p_{\mathrm{s}}(l_{\mathrm{p}}+l\psi)=f\omega\psi-\dfrac{\theta}{2}\varphi_1 m_{\mathrm{p}}v_{\mathrm{p}}^2\\ l_\psi=l_0\left[1-\dfrac{\Delta}{\rho_{\mathrm{p}}}(1-\psi)-\alpha_1\Delta\psi\right]\end{cases}\tag{1}$$

式中：ψ为火药相对已燃体积的形状函数；v_{p}为活塞速度；l_{p}为燃烧位置；z为火药；f为火药力；ω为装药量；Δ为药室装填密度；l_0为药室容积缩径长；u_1为燃速系数；e_1为火药初始厚度的一半；n为燃速指数；m_{p}为活塞的质量；φ_1为活塞次要功系数；p_{s}为药室的平均压力；p_{ps}为作用在活塞面上的气室压力；S为活塞的横截面积；l_{p}为活塞路程；θ为火药相关常数；l_ψ为药室自由容积缩径长；ρ_{p}为火药密度；α_1为火药气体余容。

对式(1)采用四阶Runge-Kutta算法求解：

$$\begin{cases}y_{i,k+1}=y_{i,k}+h(K_{i1}+2K_{i2}+2K_{i3}+K_{i4})/6\\ K_{i1}=f_i(x_i,y_{1k},\cdots,y_{nk})\\ K_{i2}=f_i(x_i+h/2,y_{1k}+hK_{i1}/2,\cdots,y_{nk}+hK_{n1}/2)\\ K_{i3}=f_i(x_i+h/2,y_{1k}+hK_{i2}/2,\cdots,y_{nk}+hK_{n2}/2)\\ K_{i4}=f_i(x_i+h,y_{1k}+hK_{i3},\cdots,y_{nk}+hK_{n3})\end{cases}\tag{2}$$

式中：h为计算步长，取$h=x_{k+1}-x_k$。

2）气室部分。在理想气体，无黏的假设下，忽略轻气与管道的热损失和摩擦损失，假定气体余容为常数，可得到如下简化方程[22, 24, 26]：

$$\frac{\partial\boldsymbol{U}}{\partial t}+\frac{\partial\boldsymbol{F}}{\partial x}=\boldsymbol{H}\tag{3}$$

$$\begin{cases}\boldsymbol{U}=\begin{bmatrix}U_1\\U_2\\U_3\end{bmatrix}\\ \boldsymbol{F}=\begin{bmatrix}U_2\\U_2^2/U_1+P\\U_3U_2/U_1+PU_2/U_1\end{bmatrix}\\ \boldsymbol{H}=-\dfrac{1}{A}\cdot\dfrac{\mathrm{d}A}{\mathrm{d}x}\begin{bmatrix}U_2\\U_2^2/U_1\\U_3U_2/U_1+PU_2/U_1\end{bmatrix}\end{cases}\tag{4}$$

$$\begin{cases} U_1=\rho \\ U_2=\rho u \\ U_3=\rho\left(e+u^2/2\right) \\ e=C_{\mathrm{v}}T \\ p(1/\rho-\alpha)=RT \end{cases} \tag{5}$$

式中: C_{v} 为定容比热容; α 为轻气余容, 一般理想气体取 0; A 为发射管截面变化, 可由过渡管的角度计算; R 为气体常数, J/(kg·K)。

方程对流项的离散形式:

$$(U^{n+1}-U^n)/\mathrm{d}t=-(F_{i+1/2}-F_{i-1/2})/\mathrm{d}x \tag{6}$$

采用差分方法, 一般先将方程的通量 $\boldsymbol{F}$ 分裂为正、负通量:

$$\boldsymbol{F}=\boldsymbol{F}^{+}+\boldsymbol{F}^{-} \tag{7}$$

再采用 WENO 格式分别重构正、负通量, 得到 $F^{+}_{i+1/2}$、$F^{-}_{i+1/2}$, 进而:

$$F_{I+1/2}=F^{+}_{i+1/2}+F^{-}_{i+1/2} \tag{8}$$

最后, Runge-Kutta 时间推进, 得到下一步值 U^{n+1}。

WENO 格式具有迎风性质, 为避免负耗散导致数值振荡, 求解非线性问题, 需要将通量分裂为正、负方向, 并保证:

$$\begin{cases} \dfrac{\mathrm{d}\boldsymbol{F}^{+}}{\mathrm{d}\boldsymbol{U}}\geqslant 0 \\ \dfrac{\mathrm{d}\boldsymbol{F}^{-}}{\mathrm{d}\boldsymbol{U}}\leqslant 0 \end{cases} \tag{9}$$

LF 分裂方法如下:

$$F_i^{\pm}=\frac{1}{2}(F_i\pm\alpha_i U_i)\quad i=0,mx \tag{10}$$

式中: $\alpha_i=(|u_i+c_i|)$为 Local LF 当地 LF 分裂方法, $\alpha_i=\max(|u_i+c_i|)$为 Global LF 全局 LF 分裂方法, 一般采用全局 LF 方法, 分辨率最差, 但鲁棒性最好[27-28]。

采用 WENO 格式, 该格式可有效捕捉激波间断, 具有较高计算精度, 且方法简单, 鲁棒性好, 计算效率高, 是目前最流行的激波捕捉算法。

3)弹丸部分。在 1.1 节假设条件下可得运动方程[22]:

$$\frac{\mathrm{d}l_{\mathrm{s}}}{\mathrm{d}t}=v_{\mathrm{s}} \tag{11}$$

$$\frac{\mathrm{d}v_{\mathrm{s}}}{\mathrm{d}t}=\frac{A(p_{\mathrm{d}}-p_{cj})}{\phi_{\mathrm{s}}m_{\mathrm{s}}} \tag{12}$$

式中: l_{s} 为弹丸路程; v_{s} 为弹丸速度; p_{d} 为气室压力; ϕ_{s}为弹丸次要功系数; m_{s} 为弹丸质量; p_{cj} 为弹丸气体压力。

弹前气体压力[29]:

$$\frac{p_{cj}}{p_{\mathrm{q}}}=(1+\frac{\gamma-1}{2}\cdot\frac{v_{\mathrm{s}}}{C_{\mathrm{q}}})^{\frac{2\gamma}{\gamma-1}} \tag{13}$$

有激波情况, 根据激波前后关系的公式:

$$\frac{p_{cj}}{p_{\mathrm{q}}}=1+\left(\frac{v_{\mathrm{s}}}{c_{\mathrm{q}}}\right)^2\frac{(\gamma+1)\gamma}{4\gamma}+\gamma\frac{v_{\mathrm{s}}}{C_{\mathrm{q}}}\sqrt{1+\left(\frac{(\gamma+1)v_{\mathrm{s}}}{4C_{\mathrm{q}}}\right)^2} \tag{14}$$

式中: p_{q} 为初始外部环境气体条件压力; C_{q} 为声速; γ 为比热比。

2　计算结果分析

结合数学物理模型和计算方法, 编写二级轻气炮数值模拟程序, 初始条件选取为: 泵管初始压力 C_{P}=0.93 MPa, 气室直径 C_{D}=130 mm, 气室管长 C_{L}=27.7 m, 气室内气体为氢气; 发射管管径 O_{D}=37 mm, 发射管管长 O_{L}=10.0 mm; 泵管至发射管锥管过渡角 β=8°, 火药质量为 4.75 kg。计算结果与实验数据进行比较, 结果如图 2～图 4 所示。由图可知, 数值模拟的弹丸出膛速度和实验结果非常吻合, 验证了数值模拟程序的可靠性。图 3 为低精度 Upwind 格式和高精度 WENO 格式对比, 由图可知, 高精度格式与实验结果吻合较好, 低精度格式与实验差别较大。

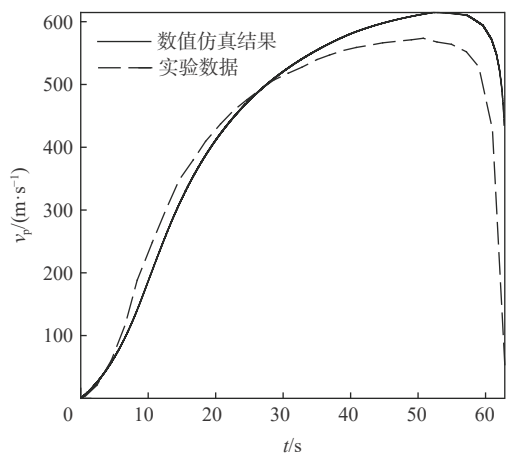


图 2　活塞速度与时间的关系

Fig. 2　Relationship between piston velocity and time

二级轻气炮的内弹道过程涉及火药燃烧、活塞运动、轻气气体流动、弹丸运动 4 部分的相互作用, 发射过程复杂, 影响参数众多, 为找到一些主要参数对二级轻气炮内弹道过程的影响, 就必须对轻气炮进行大量调试; 如果采用实验的方式, 成本过高, 因此, 有必要采用数值模拟的方式更改参数, 研究参数的改变对二级轻气炮内弹道过程的影响, 从而为二级轻气炮的研制及实验提供理论指导。本文采用参数设置为初始条件作为对照, 更改气室初始注气条件、气室尺寸、发射管尺寸、火药质量等

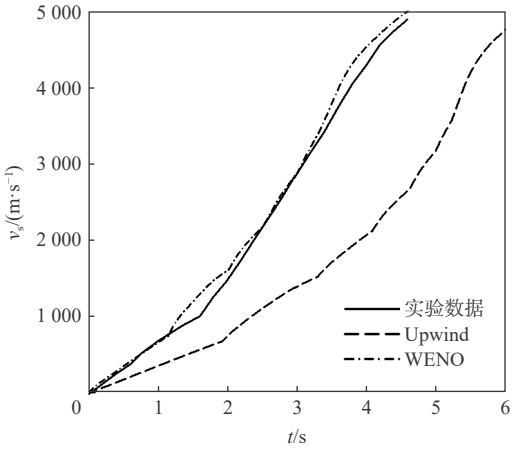


图 3 不同差分格式下弹丸速度与时间的关系

Fig. 3 Relationship between projectile velocity and time under different difference schemes

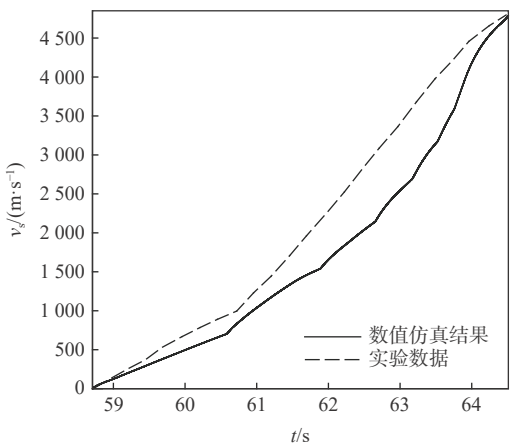


图 4 弹丸速度与时间的关系

Fig. 4 Relationship between projectile velocity and time

参数，分析各参数对二级轻气炮发射性能的影响。

2.1 气室参数对轻气炮性能的影响

轻气炮的出现就是为解决弹丸在被火药气体膨胀加速的过程中，很大一部分能量用来加速气体本身，导致最终作用在弹丸上的能量大量减少。因此，选用低摩尔质量的气体作为发射工质，从理论上可以大大提高火炮的发射性能。

在其他条件不变的情况下，更改气室内的气体，将氢气分别换成氮气和空气，数值模拟结果如图 5 所示，由图可知，氢气和氮气下活塞速度最大值远高于空气下的活塞速度，说明火药燃烧的能量中，加速空气消耗的能量占了很大一部分，和已知理论相符，进一步验证了数值模拟程序的可靠性。

二级轻气炮的气室主要由 3 个参数组成，分别是气室直径、气室长度、气室至发射管间高压锥管的过渡角角度。

其他条件不变，添加 2 个气室直径 C_D 分别为 125 mm 和 145 mm，计算结果如图 6～图 8 所示。由图 6 可知，管径越大，活塞的加速性能越好。由

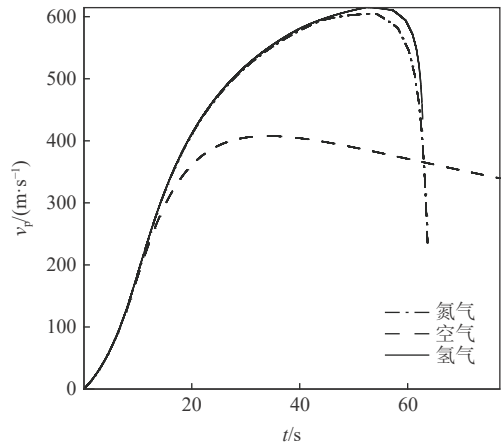


图 5 不同初始注气条件下活塞速度与时间的关系

Fig. 5 Relationship between piston velocity and time for different initial gas injection conditions

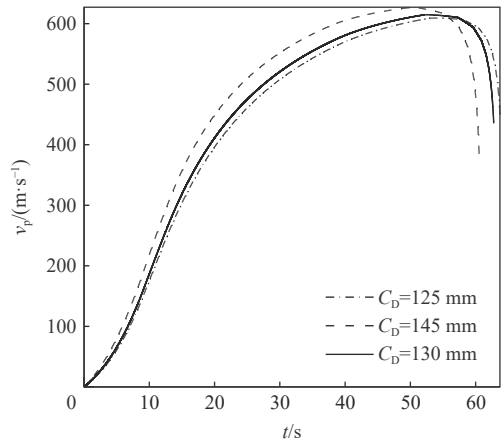


图 6 不同气室直径下活塞速度与时间的关系

Fig. 6 Relationship between piston velocity and time

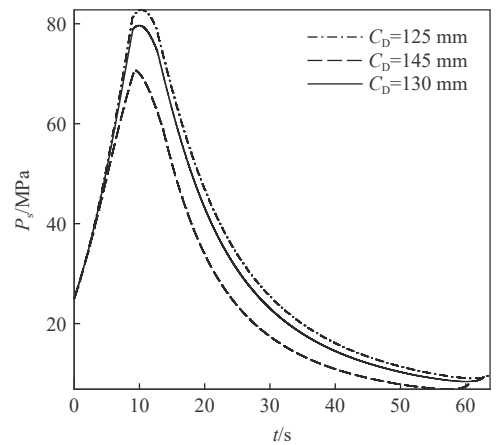


图 7 不同气室直径下活塞后压力与时间的关系

Fig. 7 Relationship between piston back pressure and time

图 7 可知，气室直径越大活塞后压力 P_s 越大，因此，导致活塞加速度和速度越大，气室与发射管间膜片破裂较早，弹丸出膛越快，但出膛速度相差不大。

理论指出，如果气室容积、气体质量和压缩速率保持不变，那么活塞速度应与气室直径的平方成反比。为提供相同的活塞能量，活塞质量按直径比

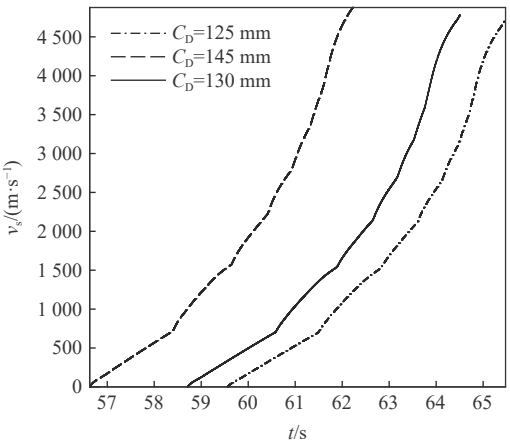


图 8　不同气室直径下弹丸速度与时间的关系

Fig. 8　Relationship between projectile velocity and time

的 4 次方变化, 即

$$\frac{m_{p2}}{m_{p1}}=\left(\frac{C_{D2}}{C_{D1}}\right)^4 \tag{15}$$

因此, 增加气室直径时, 提供相同活塞能量的条件下, 活塞质量会迅速增大。此外, 实验数据上也常常关注气室直径 C_D 与发射管管径 O_D 之比, 现有的二级轻气炮使用的 C_D/O_D 值大约在 3～5 内[30]。图 9～图 11 为测试了不同 C_D/O_D 值的弹丸出膛速度, 活塞压力和弹底压力 P_r, 初始 O_D 为 37 mm, 测试结果发现 C_D/O_D 越小活塞后压力越大, 弹底最大压力越大, 同时 C_D/O_D 在 3～4 间弹丸出膛速度相差不大, 而 C_D/O_D 增加到 5 时弹丸出膛速度下降的较多, 考虑到气室直径较小时, 活塞速度高烧蚀严重, 因此, 一般二级轻气炮的 C_D/O_D 在 3～5 之间为宜。

其他条件不变, 添加 2 个气室长度分别为 26.7 m 和 28.7 m, 计算结果如图 12 和图 13 所示。由图 12 可知, 活塞运动时间内前半段的加速过程基本一致, 但在后半段, 管长更长的情况下, 活塞所获得的加速时间越长, 从而活塞最大速度略高一些。值得注意的是, 图 13 显示气室长度的改变对弹丸出膛速度并无明显影响, 反而气室长度越长, 弹丸出膛

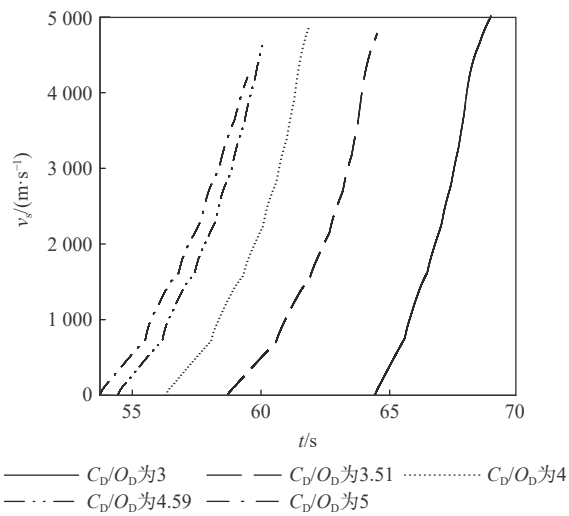


图 9　不同 C_D/O_D 下弹丸速度与时间的关系

Fig. 9　Relationship between projectile velocity and time for Different C_D/O_D

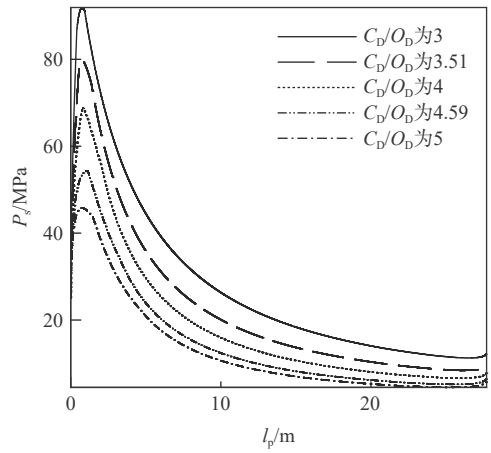


图 10　不同 C_D/O_D 下活塞后压力与活塞路程的关系

Fig. 10　Relationship between piston back pressure and piston stroke for different C_D/O_D

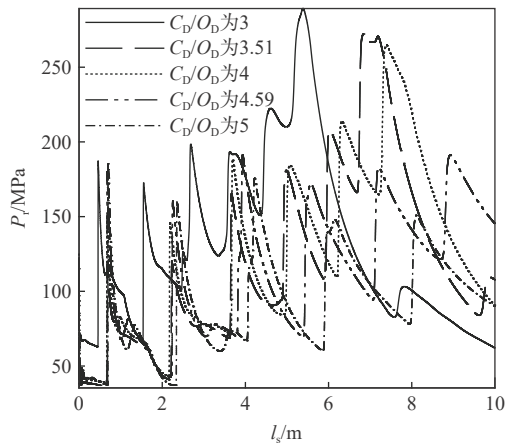


图 11　不同 C_D/O_D 下弹底压力与弹丸路程的关系

Fig. 11　Relationship between projectile back pressure and projectile stroke for different C_D/O_D

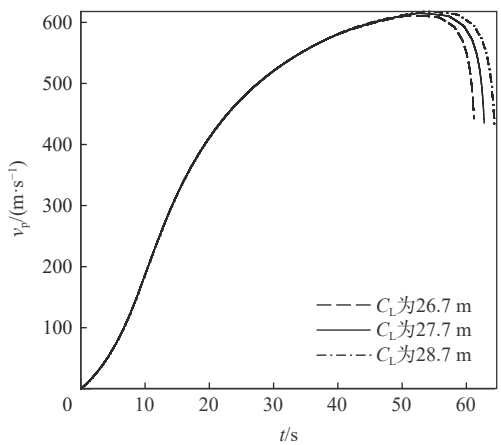


图 12　不同气室长度下活塞速度与时间的关系

Fig. 12　Relationship between piston back velocity and time for different gas chamber lengths

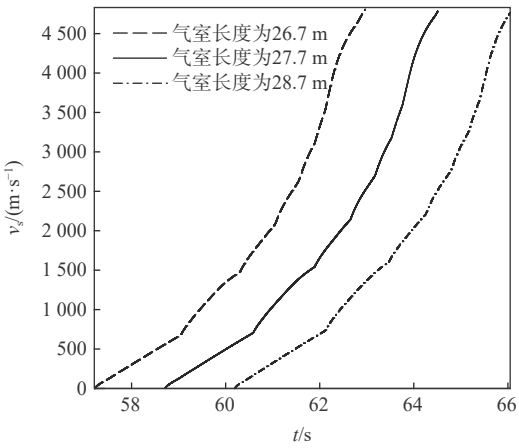


图 13 不同气室长度下弹丸速度与时间的关系

Fig. 13 Relationship between projectile velocity and time for different gas chamber lengths

时间越长，有可能导致火药燃烧完后，出现活塞后压力小于活塞前压力而出现活塞减速到0甚至往回运动的情况。因此，气室长度应随其他参数的设置而选取较为合适的值。此外，过长的气室长度也会导致烧蚀严重，因此，气室长度不宜过长。

实验研究表明，采用变形活塞和细长的锥形过渡段获得的最大弹速比垂直过渡的二级轻气炮得到的最大弹速要高得多，最大发射速度通常在过渡角角度小于15°时得到[30]。考虑这一情况，其他条件不变，过渡角角度更改为7°和10°，计算结果如图14和图15所示。由图14可知，在小范围内锥管过渡角的增加对提高弹丸出膛速度有一定的帮助。由于：

$$l=\frac{C_{\mathrm{D}}-O_{\mathrm{D}}}{2\tan\beta} \tag{16}$$

式中：l为高压锥管的长度；过渡角β在C_{D}和O_{D}不变的情况下相当于增加了高压锥管的长度，从而起到了类似于增大气室长度的作用；同时在增大过渡角

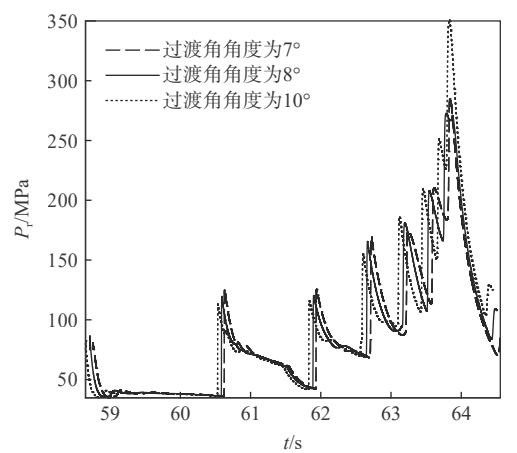


图 14 不同过渡角角度下弹底压力与时间的关系

Fig. 14 Relationship between projectile back pressure and time at different transition angles

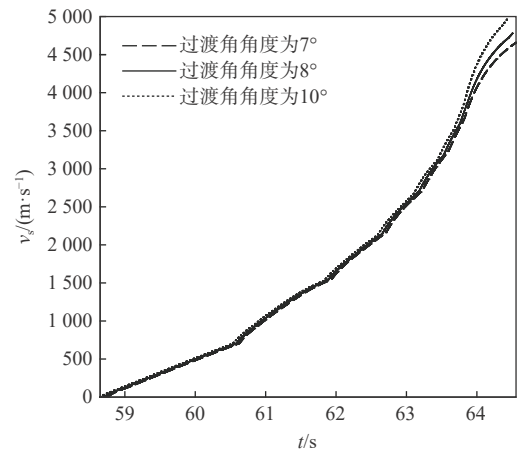


图 15 不同过渡角角度下弹丸速度与时间的关系

Fig. 15 Relationship between projectile velocity and time at different transition angles

角度时，气体的压缩更为明显，图14为过渡角角度改为10°时，弹丸底部最大压力激增，因此，8°左右可能是较为合适的过渡角角度。

2.2 发射管尺寸对轻气炮性能的影响

二级轻气炮的发射管部分主要由发射管管径和发射管长度组成，发射管管径即通常所说的火炮口径，在实际应用中存在规格的限制。目前比较流行的火炮口径有30, 76, 103, 203 mm等，故将发射管管径更改为76 mm和203 mm，其他条件不变，计算结果如图16和图17所示。图16和图17反映了随着发射管管径增大，弹丸加速度增大，但加速时间变短，因此，弹丸出膛速度不一定随着发射管管径增大而增大，甚至可能出现由于加速时间过短，弹底压力过大导致活塞返回（此情况出现在O_{D}=103 mm，O_{L}=10 m时，当设置O_{L}=20 m时，O_{D}=103 mm可获得弹丸正常射出结果）但可以考虑在增大发射管管径的同时加长发射管管长，从而获得更好的收益；如发射管管径为76 mm，更改发射管管长为20 m

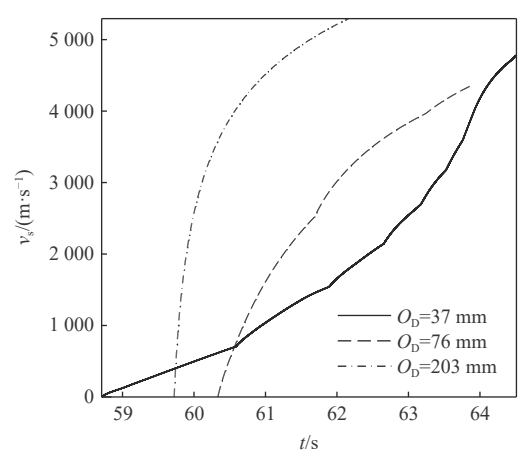


图 16 不同发射管管径下弹丸速度与时间的关系

Fig. 16 Relationship between projectile velocity and time for different launch tube diameters

后, 其弹丸出膛速度将高于发射管管径为 37 mm、发射管长为 10 m 的弹丸出膛速度, 如图 18 所示。

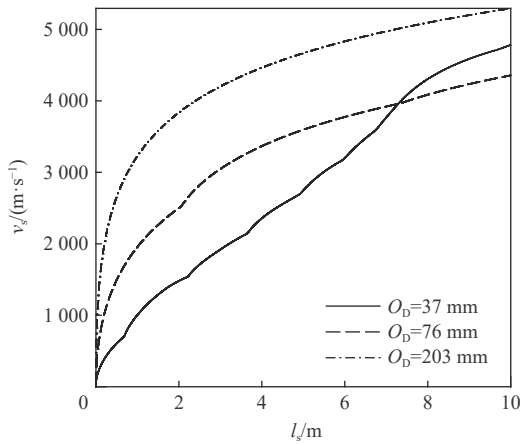


图 17　不同发射管管径下弹丸速度与弹丸行程的关系

Fig. 17　Relationship between projectile velocity and projectile stroke for different launch tube diameters

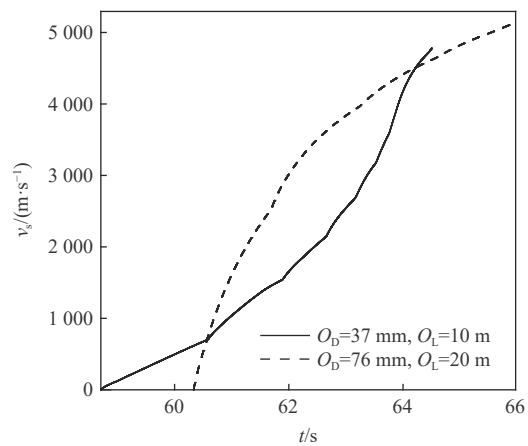


图 18　同时更改发射管管长弹丸速度与时间的关系

Fig. 18　Relationship between projectile velocity and time while simultaneously changing length of launch tube

其他条件不变, 发射管管长更改为 8 m 和 12 m, 计算结果如图 19 所示。可以看出, 管长增长部分是前部分的延续, 不同管长的前部分弹丸速度几乎一样。因此, 在其他参数合适或需要更长加速时间的情况下, 发射管管长增长可以增加弹丸的加速过程, 从而获得更高的出膛速度; 在经济允许, 其他参数设置合适的情况下, 适当增加发射管管长可以有效提高弹丸出膛速度。现有发射管长度一般是根据经验确定, 约为口径的 300～400 倍[20], 数值模拟结果也支持这一结论, 其他条件不变, 在此范围内轻气炮发射性能较好, 随着发射管长度增长弹丸出膛速度增加的斜率变低(见图 19), 因此, 过长的发射管长度也不适宜。

2.3　火药质量对轻气炮性能的影响

火药是二级轻气炮的能量来源, 火药质量的改变直接影响了活塞运行速度, 进而影响轻气流动从而最终影响弹丸的出膛速度, 因此, 火药质量对二级轻气炮的性能影响很大。

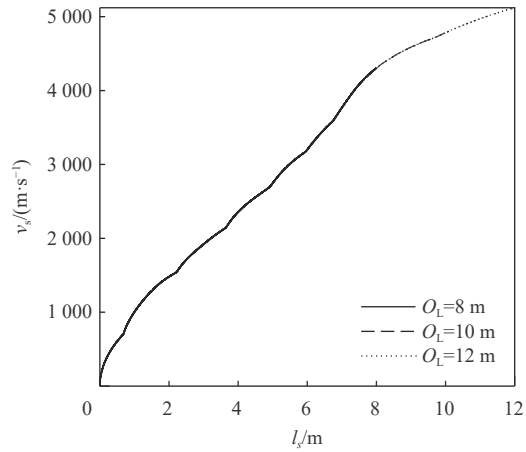


图 19　不同发射管管长下弹丸速度与弹丸行程的关系

Fig. 19　Relationship between projectile velocity and projectile stroke for different launch tube lengths

在其他条件不变下, 火药质量更改为 4 kg 和 5 kg, 计算结果如图 20～图 22 所示。可以看出, 随

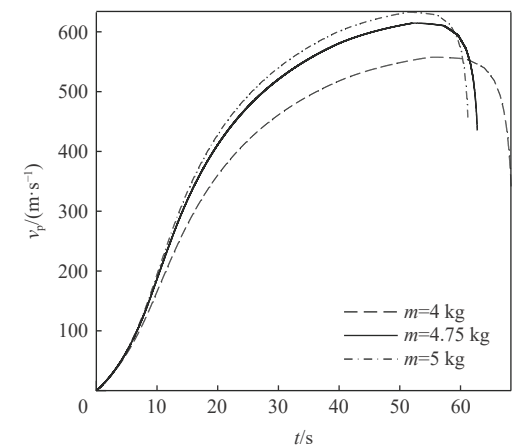


图 20　不同火药质量下活塞速度与时间的关系

Fig. 20　Relationship between piston back pressure and time for different powder masses

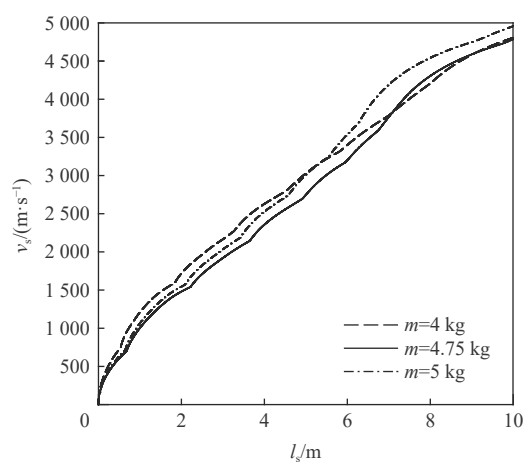


图 21　不同火药质量下弹丸速度与弹丸路程的关系

Fig. 21　Relationship between projectile velocity and projectile stroke for different powder qualities

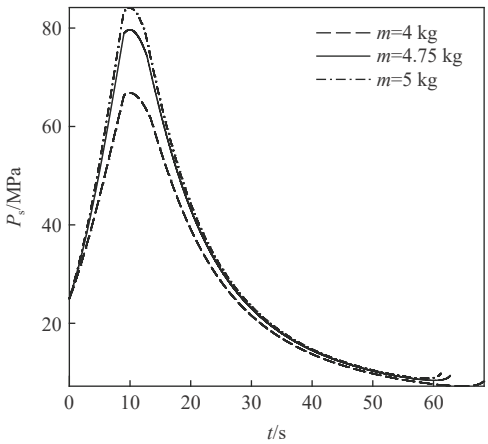


图 22 不同火药质量下活塞后压力与时间的关系

Fig. 22 Relationship between piston back pressure and time for different powder qualities

着火药质量的增加，活塞速度和弹丸速度都增大，但图 22 显示随着火药质量的增加，活塞后压力也迅速增大，那么就对药室的承载能力提出了要求。因此，火药质量虽然可以增大弹丸出膛速度，但也受到安全性的制约。

3 结 论

1）本文数值模拟程序可以对不同参数下的二级轻气炮内弹道过程进行模拟，与已有实验数据对比可以证明，本文程序是可靠的。

2）气室初始注气条件，注入氢气、氦气等轻质气体能明显提高二级轻气炮的性能。

3）发射管管径影响活塞后压力，发射管管径越大压力越大，活塞加速度和速度越大，弹丸出膛越快，但出膛速度相差不大，一般气室直径与发射管管径之比在 3～5 之间为宜，可考虑与气室长度、过渡角角度和发射管尺寸等参数协调设置；气室长度对活塞最大速度有影响，对出膛速度影响不大，考虑到气室烧蚀问题气室长度不宜过长，设置合理即可；过渡角在小范围内增大可以提高弹丸出膛速度，但是弹底压力激增，需要考虑到高压锥管的承载能力，一般以 8°左右为宜。

4）发射管管径增大可显著增大弹丸出膛速度，但需同时设置合适的发射管管长以保证弹丸有充足的加速距离；发射管管长在经济允许，其他参数设置合适的情况下，增加发射管管长可以保证弹丸在发射管中获得充足的加速时间，从而有效地提高弹丸出膛速度。

5）火药质量的增加可以有效提高活塞和弹丸的运行速度，但是会增大药室的气压，因此，火药质量的增加需要考虑到药室的安全性。

参考文献（References）

[1] BERGGREN R E, REYNOLDS R M. The light-gas-gun model launcher[J]. Ballistic Range Technology, 1970: 9-54.

[2] CROZIER W D, HUME W. High-velocity, light-gas gun[J]. Journal of Applied Physics, 1957, 28(8): 892-894.

[3] CHARTERS A C, DENARDO B P, ROSSOW V J. Development of a piston-compressor type light-gas gun for the launching of free-flight models at high velocities: NACA-TN-4143[R]. Washington, D. C. : NASA, 1957: 93R14291.

[4] CHARTERS A C. Development of the high-velocity gas-dynamics Gun[J]. International Journal of Impact Engineering, 1987, 5(1-4): 181-203.

[5] STEPHENSON W B, KNAPP R E. Performance of a two-stage launcher using hydrogen: AEDC TDR 62-32[R]. Chicago: Arnold Engineering Development Center, 1962: 386079.

[6] CURTIS J S. An accelerated reservoir light-gas gun: NASA TN D-1144[R]. Washington, D. C. : NASA, 1962.

[7] KAMIYA T, UCHIKAWA H, SHIRAKI K. Performance assessment of debris shield with hyper velocity impact testing for the japanese experiment module (JEM) and HII transfer vehicle (HTV)[C]//Proceedings of the 56th International Astronautical Congress of the International Astronautical Federation, the International Academy of Astronautics, and the International Institute of Space Law. Reston: AIAA, 2005.

[8] 滕宏辉, 杨鹏飞, 张义宁, 等. 斜爆震发动机的流动与燃烧机理[J]. 中国科学: 物理学 力学 天文学, 2020, 50(9): 129-151.
TENG H H, YANG P F, ZHANG Y N, et al. Flow and combustion mechanism of oblique detonation engines[J]. Scientia Sinica (Physica, Mechanica & Astronomica), 2020, 50(9): 129-151(in Chinese).

[9] 滕宏辉, 姜宗林. 斜爆轰的多波结构及其稳定性研究进展[J]. 力学进展, 2020, 50(1): 43-50.
TENG H H, JIANG Z L. Progress in multi-wave structure and stability of oblique detonations[J]. Advances in Mechanics, 2020, 50(1): 43-50(in Chinese).

[10] 罗凯, 汪球, 李逸翔, 等. 基于高温气体效应的磁流体流动控制研究进展[J]. 力学学报, 2021, 53(6): 1515-1531.
LUO K, WANG Q, LI Y X, et al. Research progress on magneto-hydrodynamic flow control under test conditions with high temperature real gas effect[J]. Chinese Journal of Theoretical and Applied Mechanics, 2021, 53(6): 1515-1531(in Chinese).

[11] 韩信, 张子健, 马凯夫, 等. 超燃冲压发动机喷管推力性能理论预测[J]. 气体物理, 2022, 7(1): 1-8.
HAN X, ZHANG Z J, MA K F, et al. Theoretical prediction on the nozzle thrust of scramjets[J]. Physics of Gases, 2022, 7(1): 1-8(in Chinese).

[12] 沈欢, 张子健, 刘云峰, 等. 超燃冲压发动机推进性能理论分析[J]. 气体物理, 2018, 3(1): 12-19.
SHEN H, ZHANG Z J, LIU Y F, et al. Analysis on the propulsion performance of scramjet engine[J]. Physics of Gases, 2018, 3(1): 12-19(in Chinese).

[13] 李逸翔, 汪球, 罗凯, 等. 高超声速 MHD 球头激波脱体距离理论求解[J]. 力学学报, 2021, 53(9): 2493-2500.
LI Y X, WANG Q, LUO K, et al. Theoretical analysis on hyperson-

ic mhd shock stand-off distance of blunt body[J]. Chinese Journal of Theoretical and Applied Mechanics, 2021, 53(9): 2493-2500(in Chinese).

[14] 姚轩宇, 王春, 喻江, 等. JF12 激波风洞高 Mach 数超燃冲压发动机实验研究[J]. 气体物理, 2019, 4(5): 25-31.
YAO X Y, WANG C, YU J, et al. High-Mach-number scramjet engine tests in JF12 shock tunnel[J]. Physics of Gases, 2019, 4(5): 25-31 (in Chinese).

[15] 邓放, 韩桂来, 孟宝清, 等. 超声速后台阶流动/射流相互作用的数值模拟[J]. 气体物理, 2018, 3(6): 41-50.
DENG F, HAN G L, MENG B Q, et al. Numerical simulation of supersonic flow/jet interaction at the double backward-facing step[J]. Physics of Gases, 2018, 3(6): 41-50(in Chinese).

[16] 郁雪辰, 杨鹏飞, 滕宏辉. 一维爆轰波在扰动来流中传播的数值研究[J]. 气体物理, 2022, 7(4): 1-9.
XI X C, YANG P F, TENG H H. Numerical study of one-dimensional detonation propagation in perturbed inflow[J]. Physics of Gases, 2022, 7(4): 1-9(in Chinese).

[17] 吴锦涛, 吕一品, 董刚, 等. 并行建表化学加速算法研究及其在气相斜爆轰模拟中的应用[J]. 气体物理, 2022, 7(1): 9-21.
WU J T, LYU Y P, DONG G, et al. Parallel chemistry acceleration algorithms and application in numerical simulations of gaseous oblique detonation wave[J]. Physics of Gases, 2022, 7(1): 9-21(in Chinese).

[18] 陈大年. 二级轻气炮内弹道的数值模拟与性能分析[J]. 爆炸与冲击, 1989, 9(1): 37-42.
CHEN D N. Numerical simulation and performance analysis of the two-stage light gas Gun[J]. Explosion and Shock Waves, 1989, 9(1): 37-42(in Chinese).

[19] 管小荣, 徐诚. 二级轻气炮发射过程数学模型和计算方法[J]. 南京理工大学学报 (自然科学版), 2007, 31(1): 22-26.
GUAN X R, XU C. Mathematical model and computing method for launch process of two-stage light-gas Gun[J]. Journal of Nanjing University of Science and Technology (Natural Science), 2007, 31(1): 22-26(in Chinese).

[20] 冯福全. 二级轻气炮内弹道仿真研究[D]. 南京: 南京理工大学, 2008.
FENG F Q. Simulation study on interior ballistics of two-stage hydrogen Gun[D]. Nanjing: Nanjing University of Science and Technology, 2008(in Chinese).

[21] 庄宇, 陆欣. 二级轻气炮内弹道过程数学建模及数值仿真[J]. 兵器装备工程学报, 2016, 37(12): 75-79.
ZHUANG Y, LU X. Mathematical modeling and numerical simulation for interior ballistics process of two-stage light-gas Gun[J]. Journal of Ordnance Equipment Engineering, 2016, 37(12): 75-79 (in Chinese).

[22] 刘海明. 二级轻气炮内弹道过程计算机仿真[D]. 南京: 南京理工大学, 2006.
LIU H M. Computer simulation of interior ballistic process of two-stage light gas Gun[D]. Nanjing: Nanjing University of Science and Technology, 2006(in Chinese).

[23] 朱斌. 与内弹道耦合的多管发射膛口非定常流场数值模拟[D]. 南京: 南京理工大学, 2006.
ZHU B. Numerical simulation of unsteady flow field in muzzle of multi-tube launcher coupled with interior ballistics[D]. Nanjing: Nanjing University of Science and Technology, 2006(in Chinese).

[24] 庄宇. 轻气发射装置内弹道性能数值仿真及优化设计[D]. 南京: 南京理工大学, 2017.
ZHUANG Y. Numerical simulation and optimization design of interior ballistic performance of light gas launcher[D]. Nanjing: Nanjing University of Science and Technology, 2017(in Chinese).

[25] 金志朋. 枪炮内弹道学[M]. 北京: 北京理工大学出版社, 2007.
JIN Z P. Interior ballistics of guns[M]. Beijing: Beijing Institute of Technology Press, 2007(in Chinese).

[26] 冯福全. 二级轻气炮内弹道仿真研究[D]. 南京: 南京理工大学, 2008.
FENG F Q. Simulation of the internal ballistics of a two-stage light gas gun[D]. Nanjing: Nanjing University of Science and Technology, 2008(in Chinese).

[27] SHU C W, OSHER S. Efficient implementation of essentially non-oscillatory shock-capturing schemes[J]. Journal of Computational Physics, 1988, 77(2): 439-471.

[28] JIANG G S, SHU C W. Efficient implementation of weighted ENO schemes[J]. Journal of Computational Physics, 1996, 126(1): 202-228.

[29] 吴应湘, 郑之初, 林同骥. 轻气炮性能计算[J]. 兵工学报, 1993, 14(4): 79-84.
WU Y X, ZHENG Z C, LIN T J. Performance analysis of chambered launchers[J]. Acta Armamentarii, 1993, 14(4): 79-84(in Chinese).

[30] 王金贵. 气体炮原理及技术[M]. 北京: 国防工业出版社, 2001: 96-106.
WANG J G. Principle and technology of gas cannon[M]. Beijing: National Defence Industry Press, 2001: 96-106(in Chinese).

Performance analysis of two-stage light gas gun based on higher-order code

XIANG Gaoxiang[1, 2, 3, *], WU Jiahao[1], WU Mengjia[1]

(1. School of Mechanics, Civil Engineering and Architecture, Northwestern Polytechnical University, Xi'an 710072, China;
2. Research & Development Institute of Northwestern Polytechnical University in Shenzhen, Shenzhen 518063, China;
3. Innovation Center NPU Chongqing, Northwestern Polytechnical University, Chongqing 401135, China)

Abstract: Focusing on the optimization of projectile velocity and energy conversion efficiency in hypervelocity launch technology, this paper uses the classical interior ballistic model to investigate the combustion in the chamber and the piston motion. The two components create a mathematical model for the interior ballistic launching process of a two-stage light gas cannon by describing the flow characteristics in the light gas chamber and the projectile velocity using a dimensional unstable compressible fluid model. The basic equations in the powder chamber are solved numerically with the fourth-order Runge-Kutta method, and the equations in the light gas chamber are solved with the Euler method and the WENO scheme is used to capture the shock wave. The effects of the initial gas injection conditions, light gas chamber size, launch tube size and gunpowder quality on the performance of two-stage light gas guns are studied. It is found that the performance of a light gas gun can be improved obviously by using light gas in the pump tube. The size, diameter and length of the gas chamber have significant influence on the pressure and velocity at the bottom of the piston, while their influence on the velocity of the projectile is small. Increasing the mass of propellant, the size and diameter of the launch tube can significantly increase the velocity of the projectile. This work offers crucial data and theoretical support for the light gas gun's engineering application and experimental investigation.

Keywords: two-stage light gas gun; interior ballistics; mathematical model; WENO scheme; numerical simulation

Received: 2023-09-12; **Accepted**: 2023-10-14; **Published Online**: 2023-11-17 12:50

URL: link.cnki.net/urlid/11.2625.V.20231116.1504.003

Foundation items: National Natural Science Foundation of China (12202365); China Postdoctoral Science Foundation (2021M692633,2022T150534); Guangdong Basic and Applied Basic Research Foundation (2022A1515011565,2023A1515010031); Natural Science Foundation of Chongqing (2022NSCQ-MSX5709)

* **Corresponding author**. E-mail: xianggx@nwpu.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0506

火箭推进剂连接器在脉动风载下非合作自动对接策略

李泳峄*，王南，吴新跃，吴鹏辉，姚博丹
(北京航天发射技术研究所，北京 100076)

摘 要：为解决运载火箭推进剂连接器在脉动风载激励干扰下与箭上接口的动态非合作自动对接问题，提高自动对接系统鲁棒性及响应速度。分析连接器自动对接任务剖面，制定了由检测、控制、位姿调整单元等组成的技术路线；分析脐带臂风载激励响应特性，获得了平均风速分别为10 m/s、15 m/s、20 m/s脉动风载激励下脐带臂的摆动情况；通过自动对接仿真模型开展常规PID、PID+干扰观测器、分段PID+干扰观测器控制策略仿真对比分析，搭建自动对接联合试验系统，进行控制策略验证，成功完成了自动对接全流程试验。结果表明：分段PID+干扰观测器控制策略能有效降低脉动风载激励下脐带臂摆动对自动对接的干扰影响，跟踪性能改善效果明显，对接可靠，各工况下对接跟踪误差均在±38 mm以内，完成对接时间不超过3 min。

关 键 词：自动对接；非合作；连接器；对接策略；运载火箭
中图分类号：V553.1; V554+.5
文献标志码：A **文章编号**：1001-5965(2025)10-3415-09

在航天运载火箭发射支持系统中，地面推进剂、供气等管路与运载火箭间大部分需通过脐带臂上的连接器与箭上不同高度的接口进行对接，实现推进剂加注、泄出、供气等功能。目前，连接器普遍通过人工方式与火箭接口完成对接，操作易受外界风载、光照等因素影响，且对接时二者相对姿态持续变化，对操作人员及地面保障条件要求苛刻，在紧急关机中止发射时，对接延误易引发箭上推进剂贮箱爆炸、发射设施损毁及人员伤亡等事故[1-2]。

随着世界航天技术的不断发展，各航天大国对火箭发射准备流程安全性、自动化程度及无人值守水平提出更高的要求，其中，对推进剂连接器的自动对接技术开展了持续的研究。俄罗斯方面，其旋风号、天顶号运载火箭应用了锥杆-伞式自动对接技术，但连接器对接部位位于火箭尾部，位置相对静止，对接难度低，但增加了箭上推进剂输送系统复杂度[1-2]。美国方面，其运载火箭偏向采用零秒脱落技术，部分连接器具备自动对接能力，2007年起，美国航空航天局、肯尼迪航天中心开展了“自动对接地面脐带连接器系统”、线性脐带对接锁紧机构等研究[2-6]。欧洲航天局的阿里安5运载火箭连接器操作自动化水平较高，可实现在总装厂房的对接及射前零秒脱落[7]。在国内，北京航天发射技术研究所从20世纪90年代开始从事连接器自动对接技术的研究，近年来，以未来重型运载火箭为牵引，全面开展了基于柔性直角坐标机器人的连接器自动对接技术研究，并完成了常温、低温工况下对接、加注、脱落全流程试验验证[8-10]；上海宇航系统工程研究所、南京理工大学、上海交通大学等单位对连接器自动对接技术中的机构运动学及动力学特性、SCARA机构、液压系统响应特性开展了研究[11-13]。此外，针对连接器自动对接控制策略方面的研究较少，在与之类似的动态跟踪、随动控制等问题上，由于传统的PID控制方法、滑模控制方法

收稿日期：2023-08-02；录用日期：2023-09-22；网络出版时间：2023-10-16 09:50
网络出版地址：link.cnki.net/urlid/11.2625.V.20231013.1613.002
*通信作者. E-mail：lgslyy@126.com

引用格式：李泳峄，王南，吴新跃，等. 火箭推进剂连接器在脉动风载下非合作自动对接策略[J]. 北京航空航天大学学报，2025，51(10)：3415-3423. LI Y Y，WANG N，WU X Y，et al. Non-cooperative automatic docking strategy of propellant umbilical connector for launch vehicle under fluctuating wind load[J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51(10)：3415-3423 (in Chinese).

等难以兼顾响应速度及抗干扰性能，众多学者开展了相关的控制策略、方法研究，其中，原浩等为解决双轴直驱平台伺服系统同步进给问题，提出交叉耦合迭代学习控制器与自适应加速度控制结合的控制方法[14]；针对随动系统在随机力矩干扰时输出信号畸变问题，姜尚等对基于干扰观测器与模糊前馈补偿的PID控制策略进行研究，提高了跟踪性能与鲁棒性[15]；针对磁悬浮控制敏感陀螺转子偏转耦合问题，尹增愿等采用自抗扰控制器抑制外部干扰，系统抗干扰性能及鲁棒性明显提高[16]。可以看出，目前对推进剂连接器自动对接技术的研究取得了一定成果，但主要是基于火箭尾部相对静止的箭上接口边界，且未考虑真实环境中外界风载作用下脐带臂、火箭持续摆动的干扰影响，相应的自动对接技术方案、对接策略等存在测试边界不覆盖的风险，制约了自动对接技术在航天发射场的进一步推广应用。

基于此，首先，本文分析了连接器自动对接技术的任务剖面、技术难点，提出解决的技术路线；其次，分析了脉动风载激励下的脐带臂响应特性；搭建出包含自动对接装置、脐带臂、火箭的联合仿真模型，并开展多种控制策略仿真分析；最后，研制了自动对接联合试验系统，进行了控制策略的试验验证，解决了在外界脉动风载干扰下火箭推进剂连接器与箭上接口的非合作自动对接控制策略问题，提供了连接器自动对接工程应用的关键技术支撑。

1 连接器自动对接技术

1.1 应用背景

在如图1所示的火箭地面发射支持系统中，推进剂输送管路铺设至火箭各贮箱加注口对应的脐带臂上，通过管路末端连接器与箭上接口进行对接，实现推进剂加注、泄出、排气等功能。

火箭推进剂连接器自动对接技术应用的任务剖面包括以下。

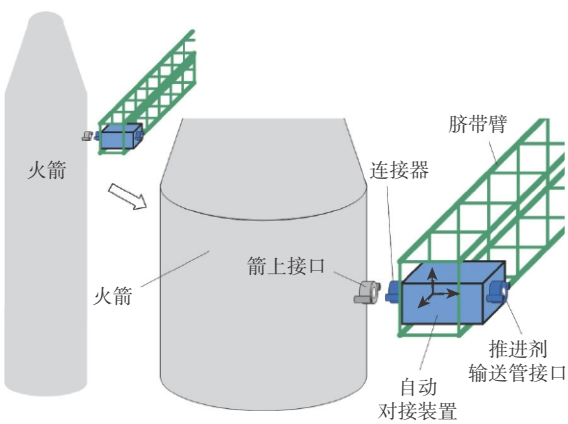


图1 推进剂连接器自动对接应用示意

Fig. 1 Automatic docking application of umbilical connectors

1）在技术区，进行连接器与箭上接口对接匹配测试。

2）在发射区，射前连接器与箭上接口对接，进行推进剂加注；加注完成后，连接器与箭上接口自动脱落，脐带臂摆开。

3）判断火箭发射状态，满足条件则点火发射，火箭升空；异常需推迟或中止发射时，脐带臂需重新摆回，连接器与箭上接口进行紧急二次对接，进行推进剂泄出、紧急排气等处理，保证火箭及地面发射设备安全。

综上，连接器自动对接技术主要用于解决常规发射准备流程中的对接任务，以及异常情况下避免人工抢险的紧急二次对接任务。

1.2 技术难点

连接器自动对接主要技术难点在于：

1）对接装置所处的脐带臂、火箭等均为大型细长结构，在外界脉动风载激励下会持续随机摆动，箭地接口相对位姿持续变化，与常规的静态对接或合作模式下的动态对接不同，属于非合作模式动态对接；

2）箭地接口间有较高的推进剂密封要求，对接精度要求高；

3）连接器对接过程中，需通过对接装置进行位姿动态调整，动作时会引入激励，进一步加剧了脐带臂的摆动。

上述诸多技术难点对连接器自动对接的控制策略提出更高的鲁棒性和响应速度要求。

1.3 技术路线

参照国内现役火箭人工完成连接器对接方式，制定了由检测、控制、位姿调整单元组成的自动对接装置技术路线，依次代替人工对接过程中眼睛、大脑、手臂的作用[8-10]。连接器自动对接工作流程如下：

1）检测单元实时获取连接器与箭上接口相对位姿偏差；

2）控制单元接收反馈的相对位姿偏差信息，解算出位姿调整单元执行机构的控制量；

3）位姿调整单元实时调整连接器位姿，与箭上接口位姿一致后，进行对接插入动作；

4）连接器与箭上接口锁紧后，与自动对接装置脱离；

5）自动对接装置复位，流程结束。

2 风载激励下脐带臂响应分析

2.1 脐带臂结构特性

脐带臂作为地面发射支持系统的一部分，通常为细长桁架式结构，主要用于承载火箭推进剂加

注、供气等管路及自动对接装置, 射前摆开让出火箭起飞通道, 应急抢险时可自动摆回。首先, 脐带臂的结构设计时刚度和强度按照最大风速下的静风压载荷进行校核; 其次, 考虑自然风随机脉动载荷作用下的抖振, 避免引发结构局部或整体破坏; 此外, 脐带臂的一阶频率应避开火箭一阶固有频率, 避免自动对接连接器位姿调整产生的激励所引发的共振风险。

以国内某型号火箭推进剂加注脐带臂模块为例, 离地高度超 50 m, 主体为采用圆形钢管焊接组成的桁架结构, 桁架截面高度、宽度尺寸分别为 2 m、1 m, 长度为 25 m, 如图 2 所示。

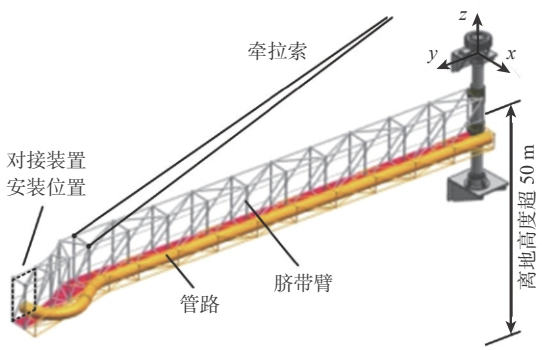


图 2　脐带臂模块(含管路、对接装置等)示意图

Fig. 2　Umbilical arm module (including pipeline and docking device) diagram

该脐带臂的结构自重及附加的管路、自动对接装置等总质量超 30 000 kg。在脐带臂端头设置有牵拉索, 用于减小脐带臂等结构重量引起的下沉变形量。

根据脐带臂模态分析结果, 前 3 阶频率分别为 1.26 Hz、2.58 Hz、6.90 Hz, 对应的振型分别为绕 z 向摆动、绕 x 轴摆动、绕 y 轴扭转。其中, 一阶频率远离了火箭的一阶固有频率, 符合对接装置对脐带臂结构特性的要求。

2.2　风载荷计算

发射支持系统的脐带臂等结构由于离地较高, 且迎风面积较大, 容易受到外界环境中自然风影响。而自然风本质上是一种动力随机载荷, 存在脉动特性, 在脉动风载的激励下, 脐带臂结构将会产生持续的动态摆动。

结合如图 2 分析, 由于脐带臂 x 向的迎风面积最大, 且与连接器对接插入方向一致, 故沿该方向的顺风向脉动风载激励对连接器自动对接过程起主要影响。因此, 本文为提高风载荷分析效率, 将三维风场降低为一维风场, 不考虑沿 y 向、z 向脉动风载对自动对接过程的影响, 但考虑了沿脐带臂 y 向上脉动风速的空间相关性。

参考脉动风速功率谱模拟公式, 利用谐波合成法可计算获得参考高度为 10 m、参考平均风速分别为 10 m/s、15 m/s、20 m/s 条件下该脐带臂对应高度的脉动风速时程[17-18]。

通过式(1)可计算获得该脐带臂受到的顺风向脉动风载荷, 分别包括平均风速及脉动风速下 2 部分载荷:

$$F = CqA \tag{1}$$

$$q = \bar{q} + q_{\mathrm{f}}(t) \tag{2}$$

$$\bar{q} = \rho \bar{U}(z)^2/2 \tag{3}$$

$$q_{\mathrm{f}}(t) = \rho \bar{U}(z) v_{\mathrm{f}}(t) + \rho v_{\mathrm{f}}(t)^2/2 \tag{4}$$

式中: F为风载荷; C为风力系数; q为总风压; A为迎风面积; $\bar{q}$为平均风压; ρ为空气密度(取 1.25 kg/m^3); $q_{\mathrm{f}}(t)$为脉动风压; $\bar{U}(z)$为高度z处的平均风速; $v_{\mathrm{f}}(t)$为脉动风速。

2.3　摆动响应分析

本文利用 ANSYS 有限元软件, 采用梁、杆单元建立了脐带臂结构、钢丝绳等模型, 在脐带臂根部施加固定约束, 将上述沿 x 向的顺向风脉动载荷施加至脐带臂模型各节点上, 通过计算可获得脐带臂各节点的响应摆动位移。

图 3 为自动对接安装位置节点处的摆动位移仿真结果, 可以看出, 平均风速为 10 m/s、15 m/s、20 m/s

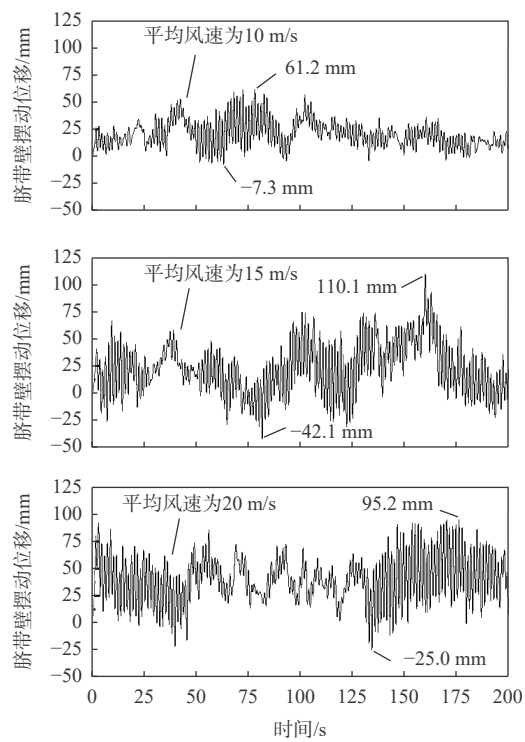


图 3　顺风向脉动风载激励下脐带臂摆动位移时程片段

Fig. 3　Displacement response of umbilical arm under fluctuating wind load in downwind direction

条件下，在截取的200 s计算时程内，脐带臂自动对接装置安装位置节点处的最大摆动位移分别达到了68.5 mm、152.2 mm、120.2 mm，而由于脉动风载的随机性及脐带臂响应影响，15 m/s、20 m/s条件下的风致摆动位移出现了差异。

3 自动对接控制策略仿真分析

3.1 自动对接装置方案

本文自动对接装置技术方案包括检测单元、控制单元和位姿调整单元，如图4所示。根据自动对接工作流程，综合考虑任务可靠性、流程优化、系统复杂程度等因素，采用非接触式测量+主动位姿调整+对接后被动随动总体策略，该技术方案、策略在前期研究中已得到了验证[10]。

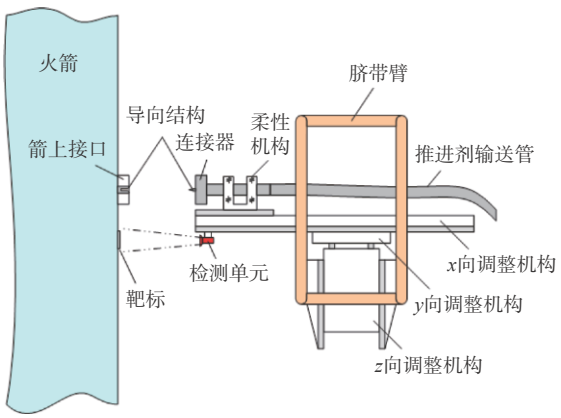


图4 自动对接装置技术方案图

Fig. 4 Technical scheme of umbilical connector automatic docking device

对接装置中，主动位姿调整环节，采用工业领域常用的三自由度直角坐标机器人方案，由x、y、z这3个方向调整机构对连接器的3个平动自由度进行主动调整；通过柔性机构夹持，连接器可实现3个转动自由度的自适应被动调整。连接器与箭上接口间设计了锥杆+锥伞导向结构，当二者各向位置偏差小于±30 mm时，可满足连接器对接精度要求。

3.2 联合仿真模型搭建

联合仿真采用基于多物理领域统一建模语言Modelica的系统设计与仿真验证平台M Works软件，搭建了跨机械、电气、电机、控制、测量、算法等专业的自动对接系统联合仿真模型，包含自动对接装置模型、脐带臂模型、火箭模型等，可快速开展方案迭代、参数优化、算法验证等工作，如图5所示，其中：

1）自动对接装置模型中，包括3个自由度的导轨滑块模型、滚珠丝杠传动模型、交流伺服电机模型、控制单元模型及实时检测连接器与箭上接口相对位姿的检测单元模型等。

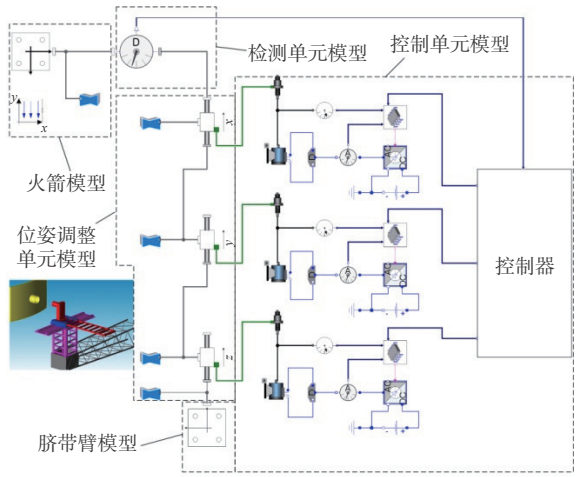


图5 自动对接系统联合仿真模型

Fig. 5 Co-simulation model of automatic docking system

2）脐带臂模型中，由于自动对接装置质量相对脐带臂较轻，忽略连接器调整时对脐带臂的附加激励，仅考虑顺风向脉动风载激励的摆动位移。

3）火箭模型中，等效简化为箭上接口结构的正弦运动，频率与火箭一阶固有频率相同，幅值根据箭上接口处的风致摆动情况进行设定。

3.3 控制策略选择

本文自动对接策略采取基于误差反馈的PID控制，各运动轴交流伺服电机速度环控制模型表达式为

$$V(t)=K_{\mathrm{p}}e(t)+K_{\mathrm{i}}\sum_{j=1}^{t}e(t)+K_{\mathrm{d}}(e(t)-e(t-1)) \tag{5}$$

式中：$V(t)$为第t时刻电机转速；$e(t)$、$e(t-1)$分别为第t、$t-1$时刻实时位置误差；K_{p}、K_{i}、K_{d}分别为PID控制器比例、积分、微分系数[19]。

考虑顺风向脉动风载激励下脐带臂摆动对自动对接引入的干扰影响，为保证自动对接控制的跟踪性能及鲁棒性，参考自适应控制、前馈补偿控制、自抗扰控制等方法，在PID控制核心理念基础上，选取PID+干扰观测器及分段PID+干扰观测器的控制策略进行分析、对比。

干扰观测器的控制类似于前馈控制，在自动对接过程中，对由于脐带臂风致摆动引起的未知干扰进行观测，并根据观测结果在控制输入端进行前馈补偿，提高了系统响应特性，并且能减小或消除未知干扰对系统的影响。干扰观测器一般包含被控对象的逆标称模型及滤波器。根据干扰观测器控制原理，设定实时误差$e(t)$经PID控制器解算后的控制输入为u，被控对象自动对接装置传递函数为$G(s)$、名义模型为$G_{\mathrm{n}}(s)$，干扰为d，低通滤波器传递

函数为$Q(s)$, 控制输出为y, 则控制输入到输出的传递函数$G_{\mathrm{i}}(s)$和干扰到输出的传递函数$G_{\mathrm{d}}(s)$的表达式分别为

$$G_{\mathrm{i}}(s)=\frac{G(s)G_{\mathrm{n}}(s)}{G_{\mathrm{n}}(s)+Q(s)(G(s)-G_{\mathrm{n}}(s))} \tag{6}$$

$$G_{\mathrm{d}}(s)=\frac{G(s)G_{\mathrm{n}}(s)(1-Q(s))}{G_{\mathrm{n}}(s)+Q(s)(G(s)-G_{\mathrm{n}}(s))} \tag{7}$$

由式(6)和式(7)可知, 滤波器对干扰抑制性能起关键作用, 滤波器$Q(s)$趋向于 1 时, 被控对象将完全呈现名义模型特性, 干扰将得到完全抑制。考虑脐带臂在脉动风载激励下摆动产生的干扰频率较低, 滤波器设置为一阶低通滤波器, 滤波器截止频率设为 1.5 Hz[15]。

分段 PID 控制策略中, 主要根据第t时刻连接器与箭上接口实时位置误差$e(t)$分段设置 PID 控制器比例、积分、微分系数, 以获得对接过程中对接准备、对接逼近、跟踪随动 3 个不同阶段所需的跟踪性能。实时位置误差$e(t)$分为 3 个区间: $|e(t)|\geqslant 200$; $30\leqslant|e(t)|<200$; $0<|e(t)|<30$。

3.4　仿真结果

通过自动对接联合仿真模型, 开展了不同脉动风载条件、火箭运动规律条件下自动对接控制策略的仿真分析, 主要结果如下:

1) 在一定的火箭运动规律下, 随着平均风速的增加, 脐带臂摆动干扰对自动对接跟踪精度影响加剧, 跟踪精度下降。以火箭正弦运动速度为 150 mm/s 边界为例, 在常规 PID 控制策略下, 0 m/s、10 m/s、15 m/s、20 m/s 平均风速条件下的动态跟踪误差分别为−10.7～21.2 mm、−17.1～23.3 mm、−43.1～48.1 mm、−42.2～45.6 mm。

2) 在一定的火箭运动规律及 0 m/s、10 m/s、15 m/s、20 m/s 平均风速条件下, 相对常规 PID 控制策略, PID+干扰观测器、分段 PID+干扰观测器 2 种控制策略下的自动对接跟踪误差明显改善, 跟踪精度平均提高约 35%、78%。

图 6 为平均风速为 15 m/s 下脐带臂摆动及火箭正弦运动速度为 150 mm/s 边界下不同控制策略对应的跟踪误差对比, 常规 PID、PID+干扰观测器、分段 PID+干扰观测器 3 种控制策略下进入跟踪允许误差区间时间基本相当(分别为 4.45 s、4.18 s、4.17 s), 过程最大跟踪误差分别为±48.8mm、±21.8 mm、±10.4 mm, 跟踪精度明显提高。此外, 在常规 PID 控制策略情况下, 跟踪误差超差明显, 在截取的 200 s 时程内上限超差 38 次, 下限超差 47 次, 引起连接器与箭上接口间相对姿态持续变化, 未能达到稳定对接插入状态, 影响对接可靠性。

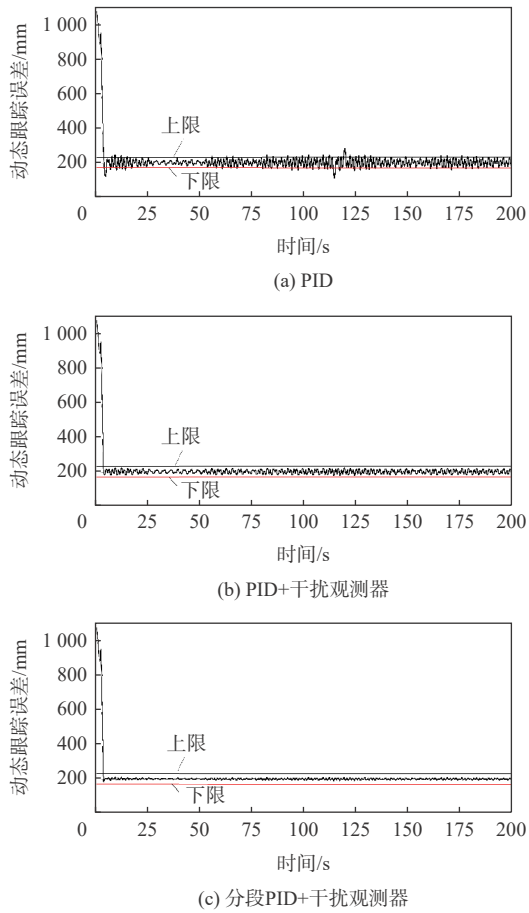


图 6　不同控制策略下跟踪误差对比

Fig. 6　Comparison of tracking errors under different control strategies

因此, 本文优选分段 PID+干扰观测器作为后续自动对接试验系统的控制策略。

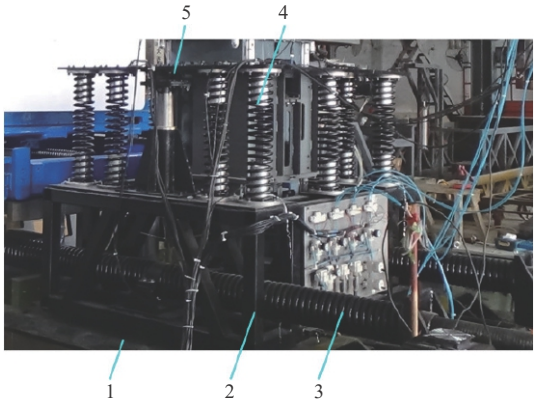
4　风载下自动对接联合试验分析

4.1　脐带臂模拟装置

由于真实脐带臂体积、质量规模较大, 制作困难, 本文研制了脐带臂模拟装置, 如图 7 所示, 在水平、垂直方向分别设置有弹簧组、阻尼器和导轨结构, 用于等效模拟脐带臂的模态特性。在模拟装置的水平运动方向上设置有伺服电动缸、压力传感器及采用工控机+运动控制卡方式的电控系统, 用于等效模拟顺风向脉动风载对脐带臂的激励。测试结果表明, 该装置一阶频率约为 1 Hz(理论值为 1.26 Hz), 符合设计允许偏差不大于±0.5 Hz 要求。

模拟装置采用交流伺服电机电流环 PID 闭环控制方式, 根据图 3 不同脉动风载激励下的脐带臂理论响应位移, 计算出交流伺服电机输出力矩及电动缸的输出力, 通过压力传感器实时反馈实现装置模拟摆动的高精度闭环控制。

通过试验测试, 可分别获得平均风速为 10 m/s、15 m/s、20 m/s 条件下脐带臂的风致摆动响应试验



1. 底座；2. 水平移动座；3. 水平弹簧组；4. 垂直弹簧组；5. 对接装置安装座。

图 7 脐带臂模拟装置

Fig. 7 Analog device of umbilical arm

值。图 8 为脐带臂摆动位移理论、试验对比情况，10 m/s、15 m/s、20 m/s 这 3 种平均风速条件下，在截取的 200 s 时程内脐带臂试验摆动位移与理论摆动位移基本一致，通过对试验值、理论值逐个时刻进行对比，二者最大误差均不超过 8%，该装置可以准确模拟脐带臂在脉动风载激励下的摆动情况。

4.2 自动对接联合试验系统

为验证脉动风载激励摆动干扰下连接器非合作自动对接控制策略，本文研制并搭建了自动对接联合试验系统，如图 9 所示，底座上分别相对安装

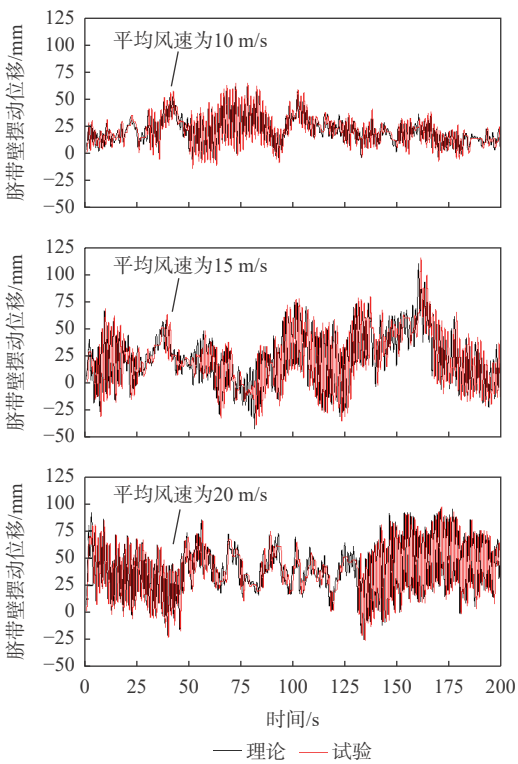
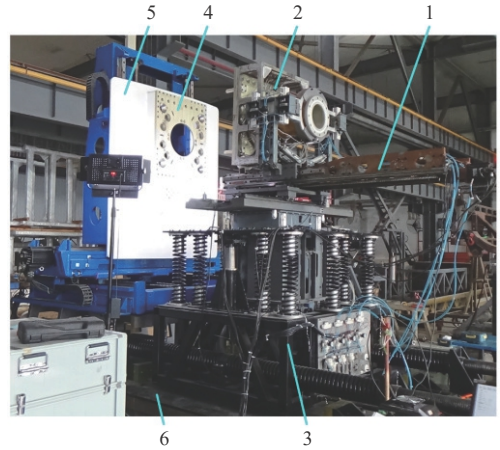


图 8 脐带臂模拟装置摆动位移理论、试验值对比

Fig. 8 Comparison between theoretical and experimental swinging displacements of umbilical arm



1. 自动对接装置；2. 连接器；3. 脐带臂模拟装置；4. 箭上接口；5. 箭体模拟运动台；6. 底座。

图 9 自动对接联合试验系统

Fig. 9 Automatic docking co-experimental system

脐带臂模拟装置、箭体模拟运动台，自动对接装置安装在脐带臂模拟装置上，通过柔性机构固定连接器，箭上接口安装在箭体模拟运动台上，与连接器对接插入方向一致。

自动对接装置中，检测单元主要由高速数字相机、图像采集与数据处理控制器、靶标等组成，采用基于模板匹配算法，获得箭上接口相对连接器的位置偏差信息，进而反馈至控制系统；控制单元由上位机、下位机、交流伺服电机、驱动器等组成，下位机采用了基于 xPC 内核的 SpeedGoat 快速原型机，通过定制的 PID 控制器、干扰观测器、滤波器等模块能高效开展控制策略设置及参数调试，且控制周期可高达 1 ms，可以满足自动对接实时精确控制的要求；位姿调整单元中 x、y、z 这 3 个方向均通过交流伺服电机驱动滚珠丝杠实现运动，通过多组气缸组成的柔性机构对连接器进行夹持定位，并实现对接过程中姿态的自适应调整。

连接器、箭上接口主要由 1 路低温连接器、8 路供气插头、锥杆-锥伞导向结构等组成。

箭体模拟运动台主要用于模拟火箭及箭上接口在风载等外界条件作用下 x、y、z 这 3 个方向的摆动情况，通过其电控系统程序可以设定不同的模拟运动规律（正/余弦、幅值、频率），单向最大运动范围为±500 mm，运动速度为±725 mm/s。

4.3 试验步骤

自动对接联合试验系统调试完成后，开展不同脉动风载激励下的自动对接控制策略验证试验，主要试验步骤如下：

步骤 1 启动脐带臂模拟装置，设定平均风速为 0 m/s；

步骤 2 启动箭体模拟运动台，设定摆动幅

值、频率参数;

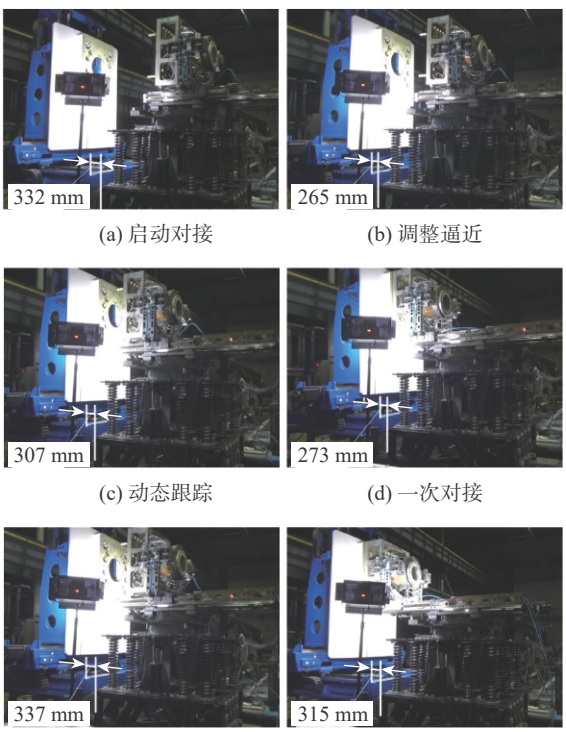
步骤 3　启动自动对接装置, 依次完成箭上接口靶标检测、三自由度位置调整逼近、动态跟踪随动、一次对接插入、脱落、二次对接、脱落等动作;

步骤 4　自动对接装置复位, 停止脐带臂模拟装置、箭体模拟运动台, 记录结果;

步骤 5　依次调整步骤 1 中平均风速为 10 m/s、15 m/s、20 m/s, 重复步骤 2～步骤 4。

4.4　试验结果分析

图 10 为脐带臂模拟装置设定平均风速为 15 m/s、箭体模拟运动台设定正弦运动速度为 150 mm/s 工况下的连接器对接过程, 成功完成了自动对接全流程试验。此外, 从脐带臂模拟装置与底座的相对位置测量值变化情况可以看出, 该装置在脉动风载激励下沿顺风向持续摆动, 最大摆动量约为 154 mm, 与理论计算结果(152.2 mm)基本一致, 对自动对接装置实时跟踪、对接控制干扰明显。



(a) 启动对接　(b) 调整逼近

(c) 动态跟踪　(d) 一次对接

(e) 对接后脱落　(f) 二次对接

图 10　脉动风载激励干扰下自动对接试验过程

Fig. 10　Automatic docking process under fluctuating wind load disturbance

通过脐带臂模拟装置模拟平均风速分别为 10 m/s、15 m/s、20 m/s 条件下的摆动干扰, 开展了系列自动对接联合试验, 并对自动对接过程的最大跟踪误差、箭地接口状态、对接耗时等进行对比分析, 结果表明: 在常规 PID 控制策略下, 脐带臂摆动干扰对自动对接控制影响明显, 跟踪误差基本稳定在±50 mm～±65 mm 之间, 对接过程中多次出现连接器与箭上接口间脱离现象, 与仿真结果中跟踪误差超差现象一致, 导致对接流程中止, 对接可靠性差, 无法完成对接。

在分段 PID+干扰观测器控制策略下, 脐带臂摆动干扰对自动对接控制影响明显改善, 跟踪误差基本稳定在±16 mm～±38 mm 之间, 连接器插入箭上接口过程顺畅, 多次对接过程中连接器与箭上接口间姿态稳定, 无脱离现象, 对接可靠, 完成对接时间均在 3 min 以内。该控制策略能满足脐带臂风致摆动干扰下的自动对接需求。

5　结　论

1) 火箭推进剂连接器与箭上接口间的对接受脉动风载激励干扰影响, 属于非合作模式动态对接, 对对接控制的鲁棒性及响应速度要求高。

2) 本文通过对某型号火箭脐带臂风载激励响应特性分析, 平均风速为 10 m/s、15 m/s、20 m/s 脉动风载激励下对接装置安装位置的摆动位移分别达到 68.5 mm、152.2 mm、120.2 mm。

3) 本文建立自动对接联合仿真模型, 开展了常规 PID、PID+干扰观测器、分段 PID+干扰观测器控制策略分析, 结果表明, 常规 PID 抗干扰性较差, 跟踪精度不能满足对接精度要求, 无法完成对接; PID+干扰观测器表现出一定的抗干扰能力, 跟踪精度有所提高; 分段 PID+干扰观测器抗干扰能力较好, 对接过程稳定, 跟踪精度改善效果明显。

4) 本文搭建了自动对接联合试验系统, 开展了不同平均风速下连接器与箭上接口的自动对接控制策略验证试验, 结果与仿真基本一致, 分段 PID+干扰观测器控制策略合理、有效, 抗干扰性强, 多次成功完成了连接器自动对接全流程试验, 平均风速为 10 m/s、15 m/s、20 m/s 脉动风载激励下最大跟踪误差在±38 mm 以内, 完成对接时间不超过 3 min, 对接可靠。

本文对脐带臂进行等效模拟, 干扰与真实工况还有差别; 后续在具体火箭型号工程研制中, 将结合发射场使用边界, 开展脉动风载激励干扰下自动对接联合仿真、对接控制参数优化及可靠性提升等工作, 加快实现推进剂连接器自动对接技术的工程应用。

参考文献 (References)

[1] 何家声. 航天发射地面支持技术[M]. 北京: 北京理工大学出版社, 2015.
HE J S. Space launch ground support technology[M]. Beijing: Beijing Insititute of Technology Press, 2015(in Chinese).

[2] 符锡理. 运载火箭脐带自动脱落连接器[J]. 国外导弹与航天运载

器, 1988(9): 60-67.
FU X L. Automatic drop connector for umbilical cord of launch vehicle[J]. Missiles and Space Vehicles, 1988(9): 60-67(in Chinese).

[3] DANDAGE S R, HERMAN N A, GODFREY S E, et al. Design and development of the space shuttle tail service masts: NASA-1979-21357[R]. Washington, D. C. : NASA, 1979.

[4] National Aeronautics and Space Administration. The perfect mate for safe fueling[EB/OL]. (2004-01-01)[2016-03-20]. https://spinoff.nasa.gov.

[5] GOSSELIN A M. Automated ground umbilical systems (AGUS) project: NASA-20130011368[R]. Washington, D. C. : NASA, 2007.

[6] MANLEY W C, TAMASY G J, MALONEY P. Ares I linear mate umbilical plate and collet: NASA-N20130008828[R]. Washington, D. C. : NASA, 2012.

[7] 白文龙, 李泳峄, 翟旺, 等. 运载火箭连接器自动对接技术应用分析[J]. 导弹与航天运载技术, 2017(6): 65-71.
BAI W L, LI Y Y, ZHAI W, et al. Application analysis of automated mating technique used on launch vehicle filling connectors [J]. Missiles and Space Vehicles, 2017(6): 65-71(in Chinese).

[8] 李泳峄, 吴新跃, 翟旺, 等. 基于柔性直角坐标机器人的火箭连接器自动对接过程分析及试验[J]. 导弹与航天运载技术, 2019(1): 116-121.
LI Y Y, WU X Y, ZHAI W, et al. Analysis and experiment on the launch vehicle's umbilical connector autonomous mating process based on flexible rectangular robot[J]. Missiles and Space Vehicles, 2019(1): 116-121(in Chinese).

[9] 李泳峄, 翟旺, 张国栋, 等. 火箭推进剂输送连接器自动对接装置及其动态特性研究[J]. 机床与液压, 2020, 48(4): 13-18.
LI Y Y, ZHAI W, ZHANG G D, et al. Research on dynamic characteristics of the autonomous mating equipment of launch vehicle's umbilical connector[J]. Machine Tool & Hydraulics, 2020, 48(4): 13-18(in Chinese).

[10] 李泳峄, 吴新跃, 李道平, 等. 运载火箭气液组合连接器动态自动对接技术[J]. 北京航空航天大学学报, 2021, 47(9): 1774-1779.
LI Y Y, WU X Y, LI D P, et al. Dynamic automatic docking technology of gas-liquid composite umbilical connector for launch vehicle[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(9): 1774-1779(in Chinese).

[11] 黄小妮, 顿向明, 张育林, 等. 运载火箭推进剂加注自动对接与脱离机器人本体设计[J]. 机器人, 2010, 32(2): 145-149.
HUANG X N, DUN X M, ZHANG Y L, et al. The auto-docking and auto-separating mechanism designed for the rocket fuel loading[J]. Robot, 2010, 32(2): 145-149(in Chinese).

[12] 游嘉伟, 顿向明, 山磊, 等. 运载火箭推进剂加注机器人机构设计[J]. 机电一体化, 2016, 22(4): 41-44.
YOU J W, DUN X M, SHAN L, et al. Mechanism design of carrier rocket propellant filling robot[J]. Mechatronics, 2016, 22(4): 41-44 (in Chinese).

[13] 张鑫, 李志刚, 徐华, 等. 运载火箭自动对接连接器液压系统研究[J]. 兵器装备工程学报, 2016, 37(3): 146-150.
ZHANG X, LI Z G, XU H, et al. Research of hydraulic system of automatic butt-joint connector for carrier booster[J]. Journal of Ordnance Equipment Engineering, 2016, 37(3): 146-150(in Chinese).

[14] 原浩, 赵希梅. 一种双轴直驱平台新型同步渐近跟踪控制设计[J]. 机械工程学报, 2023, 59(5): 271-279.
YUAN H, ZHAO X M. Design of a novel synchronous asymptotic tracking control for dual axis direct drive platform[J]. Journal of Mechanical Engineering, 2023, 59(5): 271-279(in Chinese).

[15] 姜尚, 田福庆, 梁伟阁, 等. 基于干扰观测器与模糊前馈补偿的随动系统 PID 控制策略[J]. 火炮发射与控制学报, 2017, 38(2): 29-34.
JIANG S, TIAN F Q, LIANG W G, et al. PID control strategy of servo system based on disturbance observer and fuzzy feed-forward compensation[J]. Journal of Gun Launch & Control, 2017, 38(2): 29-34(in Chinese).

[16] 尹增愿, 蔡远文, 任元, 等. 磁悬浮转子状态反馈解耦自抗扰控制方法[J]. 北京航空航天大学学报, 2022, 48(7): 1210-1221.
YIN Z Y, CAI Y W, REN Y, et al. Decoupled active disturbance rejection control method for magnetically suspended rotor based on state feedback[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(7): 1210-1221(in Chinese).

[17] 居龙, 黎定仕, 王南, 等. 脉动风载荷下发射平台摆杆的动态响应研究[J]. 工程机械, 2021, 52(2): 20-25.
JU L, LI D S, WANG N, et al. Dynamic response research on umbilical arm of launch platform under stochastic wind load[J]. Construction Machinery and Equipment, 2021, 52(2): 20-25(in Chinese).

[18] 邱惠清, 吴志君. 集装箱起重机的脉动风场仿真及响应分析[J]. 振动 测试与诊断, 2013, 33(1): 44-48.
QIU H Q, WU Z J. Simulation of fluctuating velocity field and response analysis of container crane[J]. Journal of Vibration, Measurement& Diagnosis, 2013, 33(1): 44-48(in Chinese).

[19] 董霞, 陈康宁, 李天石. 机械控制理论基础[M]. 西安: 西安交通大学出版社, 2005.
DONG X, CHEN K N, LI T S. Theoretical basis of mechanical control[M]. Xi'an: Xi'an Jiaotong University Press, 2005(in Chinese).

Non-cooperative automatic docking strategy of propellant umbilical connector for launch vehicle under fluctuating wind load

LI Yongyi*, WANG Nan, WU Xinyue, WU Penghui, YAO Bodan

(Beijing Institute of Space Launch Technology, Beijing 100076, China)

Abstract: To address the challenge of dynamic non-cooperative automatic docking between the umbilical connector and the onboard interface of a launch vehicle under fluctuating wind load disturbances, and to enhance the robustness and response speed of the docking system, this paper first analyzed the workflow of the automatic docking task and proposes a technical scheme composed of a detection module, a control module, and a pose adjustment module. Second, the wind-induced response characteristics of the umbilical arm were analyzed, and its swing behavior under average wind speeds of 10, 15, and 20 m/s was obtained. Finally, a comparative simulation analysis was conducted using conventional PID, PID with disturbance observer, and piecewise PID with disturbance observer control strategies with a co-simulation model. An integrated experimental system for automatic docking was developed to verify the effectiveness of the control strategies. Full-process docking experiments were completed. The results show that the piecewise PID combined with a disturbance observer effectively reduces the impact of umbilical arm swing under fluctuating wind loads, significantly improves tracking performance and docking reliability, maintains tracking errors within ±38 mm, and achieves docking within 3 min under various operation conditions.

Keywords: automatic docking; non-cooperative; umbilical connector; docking strategy; launch vehicle

Received: 2023-08-02; **Accepted**: 2023-09-22; **Published Online**: 2023-10-16 09:50

URL: link.cnki.net/urlid/11.2625.V.20231013.1613.002

* **Corresponding author**. E-mail: lgslyy@126.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0516

基于LSTM和海图约束的舰船航迹预测方法

赵汝健[1,2]，杨威[1,*]，巫震宇[2]，曾虹程[1]，陈杰[1]，马雷[3]

(1. 北京航空航天大学 电子信息工程学院，北京 100191；2. 北京遥感信息研究所，北京 100011；
3. 中国科学院自动化研究所，北京 100190)

摘　　要：针对现有方法对舰船航迹尤其是对具有点位稀疏、灵活机动特征的军用舰船进行预测时，航迹特征提取不完整、预测准确性与可靠性不理想的情况,综合考虑航迹的多维度特性、航迹间关联特征及舰船海上航行海图约束(NCC)等，基于船舶自动识别系统(AIS)数据提出一种改进的长短期记忆(LSTM)人工神经网络舰船位置预测方法。针对军用舰船航迹特点，预测时将历史航迹通过三次样条插值的方法生成等间隔点位数据；通过航行区域地图栅格化处理，定义可通航栅格，建立海图约束来提升预测效果。基于LSTM设计网络时，通过设置自定义损失函数、预测点位进行栅格匹配等方法将海图约束融入模型训练和预测过程。基于南海海域AIS数据的仿真实验结果表明：所构建的网络可以有效地预测舰船航迹，尤其是对于具有高度机动性的军用舰船。与传统预测方法相比，所提方法在预测准确性和预测可靠性2个方面均有所改善。

关 键 词：航迹预测；三次样条插值；海图约束；栅格化处理；长短期记忆网络
中图分类号：V221^{+}.3; TB553
文献标志码：A　　**文章编号**：1001-5965(2025)10-3424-09

航迹预测是一种利用船舶的历史和当前的位置、航速、航向等信息，预测其未来一段时间内的运动状态和轨迹的技术。现阶段大多数的航迹预测研究主要基于民用船舶展开，与民用船舶相比，军用舰船具有航行自由度高、海上活动范围广阔、机动能力强、航线易受时局影响变化等特点，导致其航线具有较强的随机性，增加了预测的难度。随着信息获取技术的飞速发展，可以通过各种手段收集大量的舰船航行数据，这些数据为深入挖掘舰船海上航行规律和行为模式从而进行航迹预测提供了有力支撑。

关于舰船航迹预测的研究已有众多方法被提出，主要包括基于统计学的方法，如卡尔曼滤波[1-3]、粒子滤波器[4-5]、模糊逻辑[6]和隐马尔可夫模型[7]等方法，上述方法需要依赖目标的运动模型和先验信息，而这些在实际情况中往往难以获得或不够准确。基于神经网络的方法，如反向传播(back propagation, BP)神经网络[8]具有强大的学习能力，可以学习和存储大量的输入-输出模式映射关系，而不需要事先了解这些映射关系的数学表达式，但在训练过程可能会陷入局部最优解，而非全局最优解。径向基函数(radial basis function, RBF)网络[9]能够利用数据自身的特征进行预测，无需建立目标运动模型，但其结构简单，不能充分利用航迹数据中的时空信息。长短期记忆(long short-term memory, LSTM)网络[10]作为一种能够处理时序数据的神经网络，近年来被广泛应用于舰船航迹预测，该网络能够有效地学习航迹数据中的长期依赖和动态变化，提高了预测的精度和稳定性，成为当前舰船航迹预测的主流方法之一。文献[11]提出使用卷积神经网络和

收稿日期：2023-08-08；**录用日期**：2023-12-29；**网络出版时间**：2024-01-23 20:39
网络出版地址：link.cnki.net/urlid/11.2625.V.20240123.1121.002
基金项目：国家自然科学基金(62271028,62101014)
*通信作者. E-mail：yangweigigi@sina.com

引用格式：赵汝健，杨威，巫震宇，等. 基于LSTM和海图约束的舰船航迹预测方法[J]. 北京航空航天大学学报，2025，51(10)：3424-3432. ZHAO R J，YANG W，WU Z Y，et al. Ship track prediction method based on LSTM and nautical chart constraints[J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51(10)：3424-3432 (in Chinese).

LSTM 网络构建时空特征融合网络, 该方法能够充分提取航迹数据中的时空特征, 消除了传统方法需要大量先验信息的弊端, 实现了基于数据驱动的舰船航迹预测。文献 [12] 则提出先用聚类分析解决数据不均衡, 再用双向长短期记忆网络提取时序和双向特征, 最后用注意力机制加强长期依赖的方法, 相比其他方法, 在舰船航迹预测的准确性和稳定性方面都有显著的提升。

上述基于 LSTM 网络的预测方法都是基于历史航迹的时间、经纬度、航速、航向 4 个维度特征, 在具有相对固定航线的民用领域有一定的应用效果, 但对航线相对灵活的军用舰船时预测效果不够理想, 经常预测出不可航行或不符合航迹路径规划规则的位置。为解决上述问题, 应考虑海图的航道约束这个变量, 本文提出一种基于 LSTM 和海图约束的军用舰船航迹预测方法, 采用“航迹插值, 生成海图约束, 基于 LSTM 进行预测”的思想, 将航迹等间隔插值、生成海图约束与航迹预测 3 个阶段相结合。

1　舰船航迹数据处理

有关舰船航迹数据处理的相关流程步骤, 如图 1 所示, 主要包括数据预处理(数据清洗与数据筛选)、三次样条插值[13]、归一化处理 3 个部分。

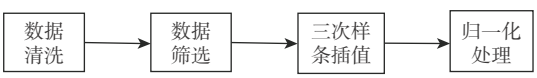


图 1　数据处理流程

Fig. 1　Data processing flow

1.1　自动识别系统数据预处理

原始的自动识别系统(automatic identification system, AIS)数据可能存在不符合实验要求的信息, 如重复、异常和错误等, 需要对其进行预处理, 以剔除错误和冗余的数据, 预处理主要包括:

1) 数据清洗。针对 AIS 数据中存在重复的数据, 通过比较相邻 2 个数据项的时间和地理信息, 当其相同时, 删除重复的数据; 为保证后续顺利插值, 如果相邻轨迹点记录的时间差小于 1 min, 则只保留第 1 个轨迹点; 当 mmsi 编码为 0 或航行速度大于 51.44 m/s 时, 可以认为异常数据, 需要删除该条数据。

2) 数据筛选。本文以南海区域内舰船预测为背景, 因此, 需筛选出符合下列要求的数据: 经度正常数据范围为东经 105°～120°、纬度正常数据范围为北纬 4°～21°、吃水正常数据范围为 0～50 m、航速正常数据范围为 0～25.72 m/s、航向正常数据范围为 0～360°。

1.2　三次样条插值

针对军用舰船 AIS 点位稀疏的问题, 本文采用三次样条插值方法, 将预处理后的历史数据插值成等间隔数据。

三次样条插值是一种常用的插值方法, 其通过构造一个三次多项式来拟合数据点, 使得拟合曲线在每个数据点处都具有连续的一阶和二阶导数。这样可以保证插值曲线的光滑性。具体来说, 假设有 n 个数据点 (x_1,y_1)、(x_2,y_2)、⋯、(x_n,y_n), 其中, $x_1<x_2<\cdots<x_n$。那么可以构造 n-1 个三次多项式:

$$S_i(x)=a_ix^3+b_ix^2+c_ix+d_i \quad i=1,2,\cdots,n-1 \tag{1}$$

使得对于任意的 i=1,2,⋯,n−1 都有 $S_i(x_i)=y_i$、$S_{i+1}(x_{i+1})=y_{i+1}$。

此外, 为保证曲线的光滑性, 还需要满足以下条件: 对于任意的 i=1,2,⋯,n−2, 有

$$\begin{cases} S_i'(x_{i+1})=S_{i+1}'(x_{i+1}) \\ S_i''(x_{i+1})=S_{i+1}''(x_{i+1}) \end{cases} \tag{2}$$

这样就可以保证拟合曲线在每个数据点处都具有连续的一阶和二阶导数。根据上述条件, 可以列出一个包含 4(n-1) 个未知数和 4(n-1) 个方程的方程组, 并求解出所有的未知数。然后, 利用求得的三次多项式计算出经度、纬度、航向、航速等信息, 对任意一个插值点进行插值。为保证后续归一化操作, 防止出现极大值或极小值, 并根据舰船航行实际, 插值过程中需限定经度插值范围为 105°～120°、纬度插值范围为北纬 4°～21°、航速插值范围为 0～50/h、航向插值范围为 0～360°。如图 2 所示, 插值后的舰船轨迹变为等间隔数据, 且采用该方法对原始数据的干扰较小。



图 2　三次样条插值法效果

Fig. 2　Effect of cubic spline interpolation

另外, 由于部分点位间隔时间过长, 为避免过多的插值数据影响了真实的航迹信息, 在处理间隔大于 12 h 的点位时, 本文并没有进行插值。通过这种方式, 本文能够在保证插值精度的同时, 避免对真实数据造成过多的干扰。

1.3　归一化处理

由于不同的特征向量可能在数量级上差别很大, 为使得表征不同属性的各数据之间具有可比性, 采用最小-最大归一化方法对航迹数据进行变

换处理。转换表达式如下：

$$\boldsymbol{X}_{\mathrm{n}}=(\boldsymbol{X}_{\mathrm{i}}-\boldsymbol{X}_{\min})/(\boldsymbol{X}_{\max}-\boldsymbol{X}_{\min}) \tag{3}$$

式中：$\boldsymbol{X}_{\mathrm{i}}$为原始值；$\boldsymbol{X}_{\min}$和$\boldsymbol{X}_{\max}$分别为最小值和最大值；$\boldsymbol{X}_{\mathrm{n}}$为归一化后的数据。处理后，原始数据值范围转换为[0, 1]。

2 构建海图约束

船舶在海上航行时，除受气象水文因素的影响外，还受到海图约束。海图是一种图形表示，显示了一个海域或水域及其相邻的海岸或岸线，主要包括水深、陆地高度（地形）、海底自然特征、海岸细节、航行障碍、航标的位置和特征、锚地和其他特征。海图的最主要作用是根据适航性来确定适航区域，适航区域可以有效防止航迹预测时预测出不可航行位置。但由于海图约束的信息复杂且可能不适合实际航行情况，无法直接用来确定适航区域，本文采用一种基于舰船历史航迹确定适航区域的方法，通过统计分析区域内所有舰船的航行数据，构建通航栅格得到一个能够反映真实适航情况的海图约束。因此，本文中所述不可航行栅格并非严格意义上的不能航行，只是根据历史数据没有相关舰船位此活动。

经综合比较，本文选择以南海军用舰船航迹预测为背景进行可通航栅格环境模型构建，南海尤其是南沙群岛岛礁众多，南沙群岛是南海四大群岛中位置最南、岛礁最多、分布最广的一组群岛，由230多个岛、洲、礁、沙和滩组成，露出水面的仅占总数的1/5。这些岛礁大多是由珊瑚礁形成的环礁或台礁，涨潮时大部分礁体没入海中，形成了复杂的海底地形和水文环境。因此，舰船在南沙群岛海域航行时，海图约束更加明显。

2.1 设置通航条件

构建海图约束时，由于限制条件仅为吃水深度，因此，可以同时使用民用船舶和军用舰船航行数据，以达到最大程度构建适航栅格的目的。针对民用船舶数据进行吃水深度条件筛选前，应采用1.1节所述方法对数据进行预处理并进行等间隔采样。

对于军用舰船数据，本文主要通过分析外军经常位南海活动且有大量公开AIS数据的海洋监测船和补给舰的航行数据来设置通航条件。

通过AIS历史数据可知，经常位南海活动的美军“能干”号、“有效”号、“忠诚”号3艘海洋监测船的吃水范围为5.8～7.6 m，刘易斯克拉克级干货弹药船和亨利凯泽级补给舰吃水范围为9.12～10.97 m。而根据公开资料查证可知，美现役尼米兹级航母的最大吃水深度为11.3 m左右，阿利伯克级驱逐舰为9.3 m左右，提康德罗加级巡洋舰为9.7 m左右。综上，以吃水大于5.5 m为标准，保留符合条件的AIS数据用来构建符合美军主要舰艇最低通航需求的通航栅格。

2.2 海图约束构建

本文基于采样后的民用船舶和插值后的军用舰船AIS数据，根据设定的栅格划分方法把航行区域分割成可通航栅格环境模型，形成适航区域。

1）栅格划分。栅格划分需要遵循以下原则：首先，可通航性是最基本的要求。舰船需要在各种海况下都能安全、稳定地航行，这就需要航道的宽度能满足舰船的通航需求。其次，军用舰船在航行过程中可能需要执行各种战术动作，如紧急转向、紧急停车等。这就需要航道的宽度和深度都能满足舰船执行这些动作的空间需求。

同时，舰船在航行过程中可能会偏离预定航线，因此，在确定航道宽度时需要考虑一定的安全冗余。

基于以上原则，以满足舰船完成一次转圈航行设计一个通航栅格，可知舰船吨位越大航速越快转弯半径越大。如图3所示，以尼米兹级航母为例，其在最常用航速10.288 m/s条件下的回转半径约为500 m，舰长为332.9 m、舰宽为40.8 m，再考虑0.5倍的安全冗余，故栅格长宽均以1 500 m为宜，然后，用墨卡托投影坐标将南海海域范围栅格化。

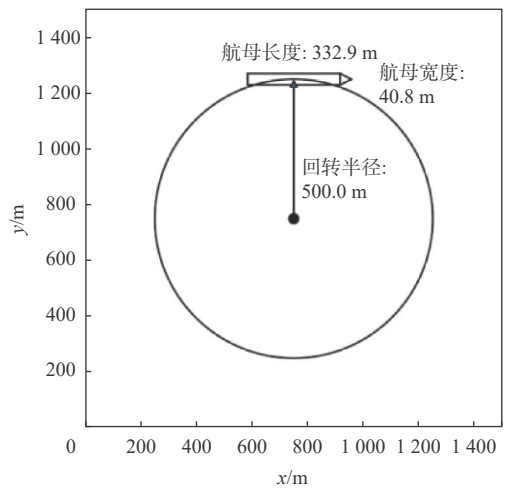


图3 栅格划分标准

Fig. 3 Grid division standard

2）数据聚类。为消除AIS数据中公共通航区内的离散点对航道约束准确性的影响，本文采用基于密度的噪声空间聚类（density-based spatial clustering of applications with noise, DBSCAN）算法[14]对AIS数据中的异常点进行处理。

DBSCAN 参数主要包括 eps 和 min_samples, eps 是距离阈值, 定义了 2 个点被认为是邻居的最大距离, 本文设置的栅格大小为 1 500 m×1 500 m, 为捕捉到栅格间的连续性, 故将 eps 参数的值设置为栅格半径即 750 m。min_samples 是包含的点数阈值, 定义了一个点被认为是核心点的邻居数量的最小值, 为能够检测到小的连续区域, 故将 min_samples 设置为一个较小的值 3。

最后, 将 AIS 数据密集的区域聚合为一个个簇, 根据聚类核心点是否密度相连, 筛选出噪声点。这种方法能有效地消除离散点的影响, 提高了航道约束的准确性, 聚类处理前后对比效果如图 4 所示。

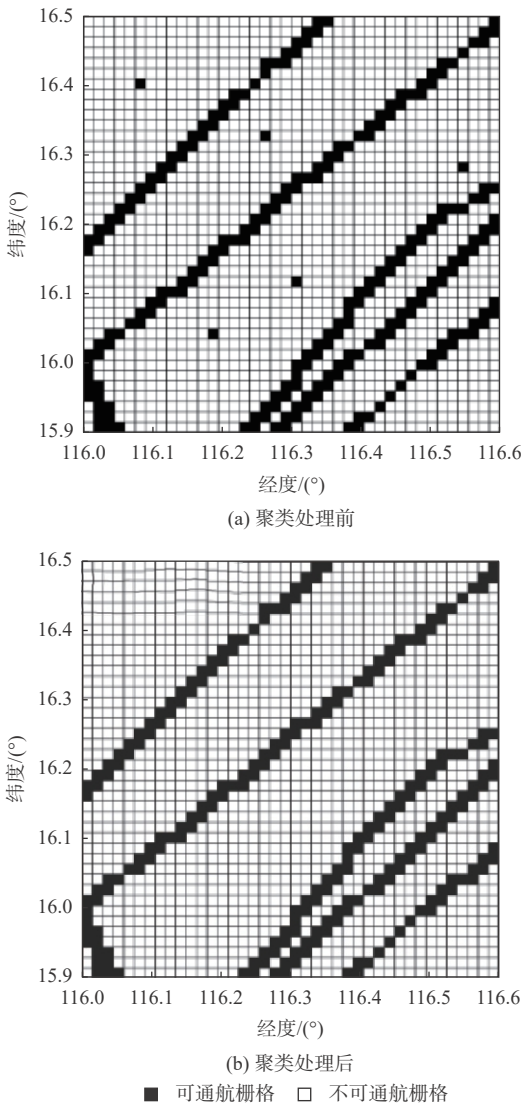


图 4　数据聚类处理效果对比

Fig. 4　Comparison of data clustering processing results

3) 约束构建。根据聚类后的 AIS 数据判断各栅格内是否有聚类核心点, 若有, 则将其标记为黑色即可通航栅格, 若无, 则将其标记为白色即不可通航栅格, 如图 4(b)所示。最终形成南海海域可通航环境分布即海图约束, 如图 5 所示。

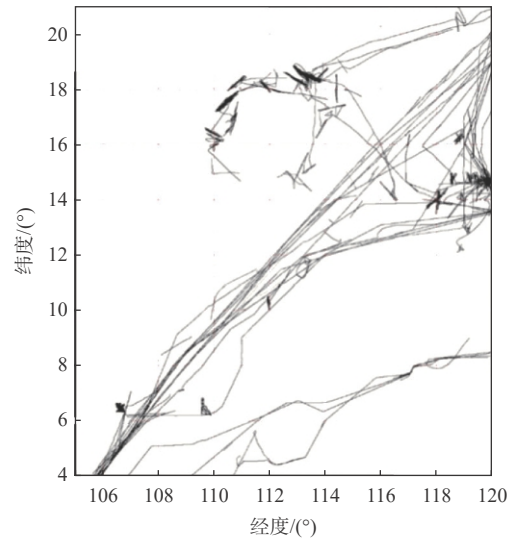


图 5　南海海域可通航环境分布

Fig. 5　Distribution of navigable environment in the South China Sea

3　舰船轨迹预测模型

3.1　LSTM 网络

LSTM 网络是一种能够处理时间序列数据的神经网络, 是为克服一般循环神经网络在处理长时间序列时出现的问题而专门设计出来的。

LSTM 网络有 3 个门来控制信息的流动, 分别是遗忘门、输入门和输出门。

遗忘门主要是对上一个节点传进来的信息进行选择性忘记, 在时间步 t 的遗忘门的激活值 f_t 表示为

$$f_t = \sigma(W_{\mathrm{f}}[h_{t-1}, x_t] + b_{\mathrm{f}}) \tag{4}$$

式中: σ 为 Sigmoid 非线性激活函数; W_{f} 和 b_{f} 分别为遗忘门的网络参数和偏置; h_{t-1} 为上一时间步的隐藏状态; x_t 为当前时间步的输入。

输入门主要是向单元状态添加新信息。当前网络输入值 x_t 进入到单元状态 c_t 的比重 i_t 表示为

$$i_t = \sigma(W_{\mathrm{i}}[h_{t-1}, x_t] + b_{\mathrm{i}}) \tag{5}$$

式中: W_{i} 和 b_{i} 分别为输入门的网络参数和偏置。

当前时刻的候选单元状态 $\tilde{c}_t$:

$$\tilde{c}_t = \tanh(W_{\mathrm{c}}[h_{t-1}, x_t] + b_{\mathrm{c}}) \tag{6}$$

式中: W_{c} 和 b_{c} 分别为记忆单元的网络参数和偏置。

当前时刻的单元状态 c_t:

$$c_t = f_t c_{t-1} + i_t \tilde{c}_t \tag{7}$$

输出门决定哪些状态被当成当前状态的输出。

单元状态 c_t 到当前输出 h_t 的比重 o_t 表示为

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \tag{8}$$

式中：W_o 和 b_o 分别为输出门的网络参数和偏置。

LSTM 网络的输出值表示为

$$h_t = o_t \tanh(c_t) \tag{9}$$

3.2 模型构建

1）确定数据维度。给定包含有限个航迹点的历史航迹 T，可表征为由一系列离散航迹点组成的序列：

$$T = (\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_n) \tag{10}$$

式中：$\boldsymbol{p}_n$ 为航迹点。

通常每个航迹点 $\boldsymbol{p}_n$ 可以用一个四维向量对其表征：

$$\boldsymbol{p}_n = [l_{ng}, l_{at}, s, c] \tag{11}$$

式中：l_{ng}、l_{at}、s、c 分别为经度、纬度、航速、航向。

由于 1.1 节军用舰船数据预处理的过程中未对时间间隔大于 12 h 的进行插值，为保证预测的准确性，本文引入一个时间间隔的维度，让 LSTM 网络模型能够学习到船舶航迹在不同时间间隔下的运动趋势，更适合船舶预测回归问题。通过上述处理每个航迹点 $\boldsymbol{p}_n$ 的表征更新如下：

$$\boldsymbol{p}_n^{'} = [l_{ng}, l_{at}, s, c, t] \tag{12}$$

式中：t 为时间间隔。

2）设置自定义损失函数。为将海图约束融入到模型训练过程中，本文设置了自定义损失函数，用于在模型训练过程中惩罚模型预测出不可航行的位置。首先，读取海图约束文件，并将其转换为常量张量；然后，在自定义损失函数中，计算预测位置对应的通航栅格的灰度值；最后，使用均方误差（mean squared error, MSE）作为损失函数计算损失值，其计算式如下：

$$M = \frac{1}{n}\sum_{i=l}^{n}\left(y_{ture,i} - y_{pred,i}\right)^2 \tag{13}$$

式中：n 为样本数量；$y_{ture,i}$ 为第 i 个样本的真实值；$y_{pred,i}$ 为第 i 个样本的预测值。

上述方法可以在模型训练过程中对通航栅格进行匹配，通过惩罚模型预测出不可航行的位置来提高模型的准确性。

3）进行预测。当前目标 o_i 处于航迹点 $\boldsymbol{p}_t$，舰船航迹预测问题就是根据当前 k 个时刻组成的序列，预测 t+1 时刻目标最有可能处于的航迹点 $\boldsymbol{p}_{t+1}$：

$$\boldsymbol{p}_{t+1} = \max \boldsymbol{P}(\boldsymbol{p}_t | o_i, (\boldsymbol{p}_{t-k}, \boldsymbol{p}_{t-k+1}, \cdots, \boldsymbol{p}_t)) \tag{14}$$

依此类推，直到完成对最后一个航迹点 $\boldsymbol{p}_{t+1}$ 的预测。

4）栅格匹配。为防止预测位置为不可通航区域，需要对预测位置进行栅格匹配。如果预测位置为不可通航区域，则使用启发式算法来寻找最优的可通航栅格，并将预测位置更新为该栅格的位置。

经综合比较，本文采用 A*算法[15]，其融合了最佳优先搜索和 Dijkstra 算法的优点，能够在保证最优解的同时提高搜索效率。具体流程是：首先，获取所有可通航栅格的坐标，并将其作为 A*算法的输入；然后，使用 A*算法寻找从当前预测位置到最近可通航栅格的最短路径，并返回路径上的所有节点；最后，取路径上的最后一个节点作为最近的可通航栅格，并更新预测结果。

该算法的核心思想是使用一个启发函数来估计从当前节点到终点的距离，从而更快地找到最短路径。主要采用了欧几里得距离公式：

$$h(n) = \mathrm{sqrt}\left((g[0] - n[0])^2 + (g[1] - n[1])^2\right) \tag{15}$$

式中：sqrt 表示平方根运算，g 和 n 分别为终点和当前节点的坐标；g [0] 和 n [0] 分别为终点和当前节点的横坐标；g [1] 和 n [1] 分别为终点和当前节点的纵坐标。

3.3 本文方法设计

本文提出的基于 LSTM 和海图约束的舰船航迹预测方法流程分为 5 步：

1）运用三次样条插值方法将稀疏舰船航迹进行修复；

2）运用可通航栅格环境模型，采用 DBSCAN 密度聚类算法将符合筛选条件的 AIS 航迹点聚类为可通航栅格；

3）构建基于 LSTM 的舰船航迹预测模型，主要在模型训练过程对数据进行改进，增加了时间间隔维度，设置了自定义损失函数；在预测阶段增加了栅格匹配过程。将实验数据按比例划分成训练集、测试集和验证集，用训练集和测试集对网络进行训练；

4）输入验证集到训练好的网络中，得到航迹预测结果；

5）设计对比实验，分析方法性能。

基于 LSTM 和海图约束的舰船航迹预测模型如图 6 所示。

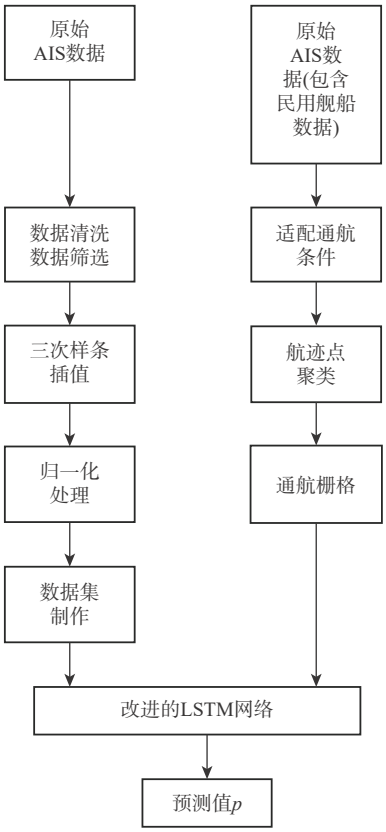


图 6　基于 LSTM 和海图约束的舰船航迹预测模型

Fig. 6　Ship trajectory prediction model based on LSTM and nautical chart constraints

4　实验验证

4.1　实验数据

本文使用位于南海海域活动的军用舰船、民用船舶真实 AIS 数据, 时间跨度为 2016—2023 年, 军用舰船种类主要包括补给舰、海洋监测船等, 数据格式为 CSV 格式, 所使用的字段为: MMSI、TIME、LONGITUDE、LATITUDE、SOG、COG、Draft。其中, 用于南海海图约束构建的数据为符合吃水标准的军用、民用船舶数据约 1 000 000 条; 用于模型训练、测试、验证的为军用舰船数据约 92 000 条, 三次样条插值后约 480 000 条, 按照 6:2:2 比例划分。

由于本文设计网络的数据输入维度为 5, 数据规模也较大, 故设置隐藏层节点数为 108、训练轮数为 200。具体网络参数设置如表 1 所示。

4.2　对比与分析

为有效评估本文方法的效果, 设计对比试验, 与使用原始数据集未融入海图约束的 BP 网络(记为 BP)和 LSTM 网络(记为 LSTM)进行对比。

为更好的反应模型的效果, 本文设计了单步预测和多步预测来进行实验验证。

单步预测是指基于之前的输入序列预测下一个时间步长的位置。由于军用舰船航行较为灵活, 历史数据中主要存在短期的依赖关系, 因此, 本文设置用于单步预测的数据步长为 10。规则如下, 从第 1 个位置开始, 前 10 个位置 (真实位置) 预测第 11 个位置, 然后, 第 2 个位置到第 11 个位置 (真实位置) 为一组, 预测第 12 个位置, 以此类推, 最后, 基于预测值和对应的真实值进行度量分析。因前期插值时以 min 为单位进行插值, 结合 LSTM 网络的特点, 预测值之间的时间步长为 1 min。

表 1　网络参数设置

Table 1　Network parameter settings

参数	数值
隐藏层节点数	108
训练轮数	200
批大小	64
Dropout	0.3
最小学习率	0.001

实验采用 MSE[16] 来衡量不同模型的精确度, 以最大误差、最小误差、误差均值来衡量不同模型的稳定性,用预测值和真实值之间的点位距离来直观的反应预测效果。不同模型的预测结果对比如表 2 所示。

由表 2 可知, 本文方法在 MSE、最大误差、最小误差、平均误差等指标上均优于对比方法, 其中, MSE 指标在经度、维度、航速、航向 4 个维度上较 LSTM 网络分别减少 0.009 4、0.000 5、0.023 5、27.131 6, 较 BP 网络分别减少 0.035 2、0.018 6、0.019、44.658 0, 尤其在航向预测上提升非常明显, 这是因为不同于民用船舶的航线相对固定, 军用舰船在航行过程中往往会根据形势随机应变, 航向变化较为频繁, 基于原有网络预测航向比较困难, 而改进网络中加入了海图约束, 使得航向预测时更加准确, 因航向预测误差导致的经纬度预测误差也随之明显降低。综上, 改进模型在精确度和稳定性 2 个方面都取得了较好效果。

如图 7 所示, 本文方法针对某艘船进行预测时, 虽然航迹灵活多变, 但预测点位更接近真实航迹点位, 如图 7(a) 所示; 距离误差均值比原有模型大幅度减少, 如图 7(b) 所示; 航向预测与真实航向拟合性更好, 如图 7(c) 所示。

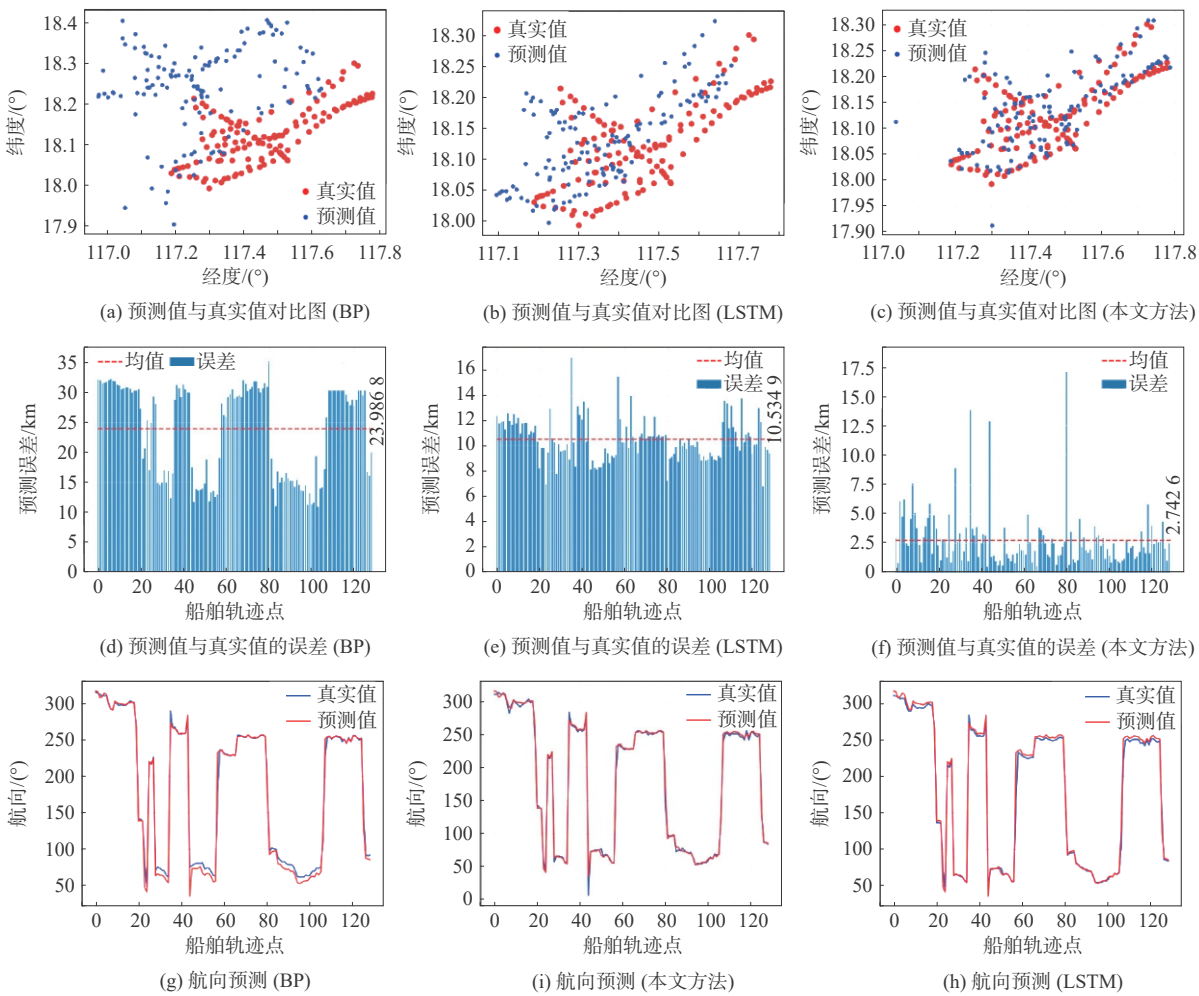
上述结果证明, 本文方法可以适用于军用舰船 AIS 点位稀疏的情况, 并在舰船灵活机动过程中预测结果也不会发生较大偏差, 预测结果具有一定的可靠性。

另外, 本文还设计了多步预测实验, 多步预测是基于之前的输入序列预测未来特定时间跨度的

表 2 不同方法结果对比

Table 2 Comparison of results of different methods

对比参数	方法	经度/(°)	维度/(°)	航速/(m·s^{-1})	航向/(°)	点位距离/km
MSE	BP	0.035 7	0.019 3	0.094 9	78.789 0	637.389 0
	LSTM	0.009 9	0.001 2	0.097 2	61.262 6	113.707 2
	本文方法	0.000 5	0.000 7	0.085 1	34.131 0	13.527 2
最大误差	BP	0.257 1	0.282 7	0.942 9	55.810 3	35.893 9
	LSTM	0.155 9	0.065 7	0.987 2	62.208 3	17.023 1
	本文方法	0.151 5	0.024 2	0.728 0	32.044 7	18.416 8
最小误差	BP	0.029 4	0.000 5	0.000 5	0.022 6	10.980 6
	LSTM	0.033 6	0.000 4	0.000 6	0.031 8	6.827 3
	本文方法	0.000 1	0.000 1	0.000 5	0.013 5	0.503 1
平均误差	BP	0.181 1	0.122 8	0.111 9	4.962 9	23.986 7
	LSTM	0.098 0	0.012 3	0.106 6	4.075 8	10.534 9
	本文方法	0.013 3	0.008 2	0.088 7	3.177 5	2.742 6



(a) 预测值与真实值对比图 (BP)　(b) 预测值与真实值对比图 (LSTM)　(c) 预测值与真实值对比图 (本文方法)

(d) 预测值与真实值的误差 (BP)　(e) 预测值与真实值的误差 (LSTM)　(f) 预测值与真实值的误差 (本文方法)

(g) 航向预测 (BP)　(i) 航向预测 (本文方法)　(h) 航向预测 (LSTM)

图 7 航迹预测(单步)的可视化效果

Fig. 7 Visualization of track prediction (single step)

位置。同样设置用于多步预测的数据步长为 10, 规则如下, 从第 1 个位置开始, 前 10 个位置 (真实位置) 预测第 11 个位置, 然后, 第 2 个位置到第 11 个位置 (预测位置) 为一组, 预测第 12 个位置, 以此循环预测更长时间的值, 其误差会随时间的延长而增加, 最后, 基于预测值和对应的真实值进行度量分析。设定预测时间跨度为 30 min, 根据 3.2 节叙述可知, 模型将预测出 30 个点位, 预测效果对比如图 8 所示。由图 8 可知, 多步预测情况下, 相较对比方法, 本文方法仍能取得较好的预测效果。

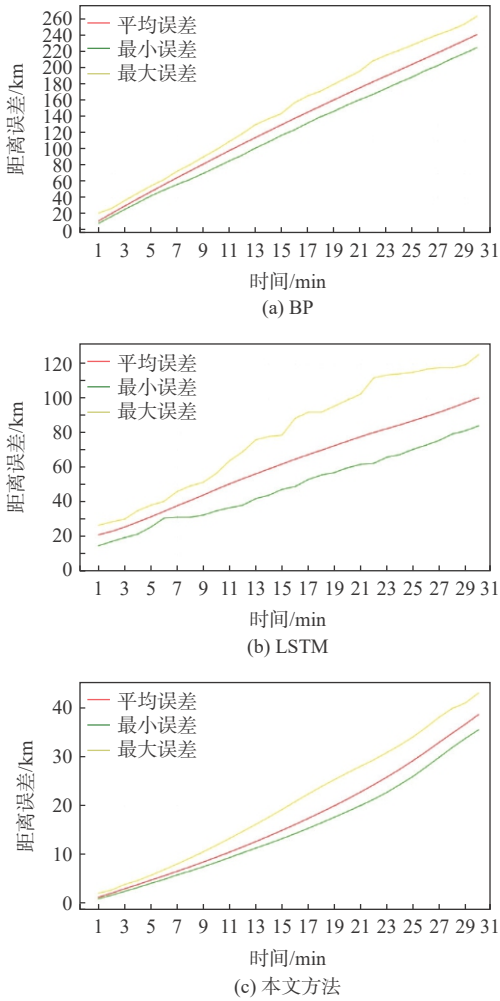


图 8　航迹预测(多步)距离误差对比

Fig. 8　Track prediction (multi-step) distance error comparison

5　结　论

本文以海上舰船位置预测为研究对象, 主要结论如下:

1) 本文方法采用“航迹插值, 生成海图约束, 基于 LSTM 进行预测”的思想, 将航迹等间隔插值、生成海图约束与航迹预测 3 个阶段相结合。

2) 以军用舰船航行随机性大为背景设计模型, 将 AIS 数据聚类成为适航性约束, 提高预测针对性和准确性。

3) 对数据进行改造, 采用三次样条插值方法将舰船稀疏数据插值成等间隔数据, 采用增加时间间隔维度的方法使模型能学习到不同时间间隔的船舶航迹运动趋势。

4) 对模型进行改进, 通过设置自定义损失函数和预测结果栅格匹配, 将海图约束融入到模型训练和预测全过程, 提高预测的准确性。

今后的工作中, 针对不同来源的船舶航行数据如描述事件的文本信息等, 通过多源异构数据融合技术开展进一步研究。

参考文献（References）

[1] 徐铁, 蔡奉君, 胡勤友, 等. 基于卡尔曼滤波算法船舶 AIS 轨迹估计研究[J]. 现代电子技术, 2014, 37(5): 1-6.
XU T, CAI F J, HU Q Y, et al. Ship trajectory prediction based on Kalman filter algorithm[J]. Modern Electronics Technique, 2014, 37(5): 1-6(in Chinese).

[2] 成光, 刘卫东, 魏尚俊, 等. 基于卡尔曼滤波的目标估计和预测方法研究[J]. 计算机仿真, 2006(1): 355-362.
CHENG G, LIU W D, WEI S J, et al. Research on target estimation and prediction method based on Kalman Filter[J]. Computer Simulation, 2006(1): 355-362(in Chinese).

[3] WAN E A, VAN DER MERWE R. The unscented Kalman filter for nonlinear estimation[C]//Proceedings of the IEEE Adaptive Systems for Signal Processing, Communications, and Control Symposium. Piscataway: IEEE Press, 2000: 153-158.

[4] 张蕴绮, 郭发勇, 朱梓清, 等. 改进的多特征粒子滤波目标跟踪算法研究[J]. 计算机测量与控制, 2023, 31(12): 106-113.
ZHANG Y Q, GUO F Y, ZHU Z Q, et al. Research on improved multi feature particle filter target tracking algorithm[J]. Computer Measurement and Control, 2023, 31(12): 106-113(in Chinese).

[5] ARULAMPALAM M S, MASKELL S, GORDON N, et al. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking[J]. IEEE Transactions on Signal Processing, 2002, 50(2): 174-188.

[6] 李鹏飞, 喻建平, 李良群, 等. 基于模糊自适应 α-β 滤波的机动目标跟踪测[J]. 系统工程与电子技术, 2008(11): 135-142.
LI P F, YU J P, LI L Q, et al. Maneuvering target tracking measurement based on fuzzy adaptive α-β filtering[J]. System Engineering and Electronics, 2008(11): 135-142(in Chinese).

[7] 孙红, 陈锁. 一种聚类隐马尔可夫模型的时空轨迹预测算法[J]. 小型微型计算机系统, 2019, 40(3): 35-41.
SUN H, CHEN S. A spatiotemporal trajectory prediction algorithm based on clustering hidden Markov model[J]. Mini-Micro Systems, 2019, 40(3): 35-41(in Chinese).

[8] 徐婷婷, 柳晓鸣, 杨鑫. 基于 BP 神经网络的船舶航迹实时预测[J]. 大连海事大学学报, 2012, 38(1): 9-11.
XU T T, LIU X M, YANG X. BP neural network-based ship track real-time prediction[J]. Journal of Dalian Maritime University, 2012, 38(1): 9-11(in Chinese).

[9] 甄荣, 金永兴, 胡勤友, 等. 基于 AIS 信息和 BP 神经网络的船舶航行行为预测[J]. 中国航海, 2017, 40(2): 6-10.
ZHEN R, JIN Y X, HU Q Y, et al. Vessel behavior prediction based on AIS data and BP neural network[J]. Navigation of China, 2017, 40(2): 6-10(in Chinese).

[10] HOCHREITER S, SCHMIDHUBER J. Long short-term memory [J]. Neural Computation, 1997, 9(8): 1735-1780.

[11] 郝延彪, 万浦波, 王文博. 基于时空特征融合的舰船航迹预测方法[J]. 海军航空工程学院学报, 2021, 36(2): 191-198.
HAO Y B, WAN P B, WANG W B. A ship track prediction method based on spatial-temporal feature fusion[J]. Journal of Naval Aeronautical and Astronautical University, 2021, 36(2): 191-198(in Chinese).

[12] 蒋通, 崔良中, 刘立国, 等. 基于聚类分析和 Att-Bi-LSTM 的舰船航迹预测方法[J]. 计算机仿真, 2022, 39(8): 1-5.
JIANG T, CUI L Z, LIU L G, et al. The method of ship track prediction based on cluster analysis and att-Bi-LSTM[J]. Computer Simulation, 2022, 39(8): 1-5(in Chinese).

[13] 张黎翔, 朱怡安, 陆伟, 等. 基于 AIS 数据的船舶轨迹修复方法研究[J]. 西北工业大学学报, 2021, 39(1): 119-125.
ZHANG L X, ZHU Y A , LU W, et al. A detection and restoration approach for vessel trajectory anomalies based on AIS[J]. Journal of Northwestern Polytechnical University, 2021, 39(1): 119-125(in Chinese).

[14] SANDER J, ESTER M, KRIEGEL H P, et al. Density-based clustering in spatial databases: the algorithm GDBSCAN and its applications[J]. Data Mining and Knowledge Discovery, 1998, 2(2): 169-194.

[15] 石浩, 乔继潘, 季盛. 一种基于 AIS 数据的船舶航线自动规划方法[J]. 上海船舶运输科学研究所学报, 2021, 44(1): 25-30.
SHI H, QIAO J P, JI S. An automatic navigation route planning for ships based on AIS data[J]. Journal of Shanghai Ship and Shipping Research Institute, 2021, 44(1): 25-30(in Chinese).

[16] 权波, 杨博辰, 胡可奇, 等. 基于 LSTM 的船舶航迹预测模型[J]. 计算机科学, 2018, 45(S2): 126-131.
QUAN B, YANG B C, HU K Q, et al. Ship track prediction model based on LSTM[J]. Computer Science, 2018, 45(S2): 126-131(in Chinese).

Ship track prediction method based on LSTM and nautical chart constraints

ZHAO Rujian[1, 2], YANG Wei[1, *], WU Zhenyu[2], ZENG Hongcheng[1], CHEN Jie[1], MA Lei[3]

(1. School of Electronic Information Engineering, Beihang University, Beijing 100191, China;
2. Beijing Institute of Remote Sensing Information, Beijing 100011, China;
3. Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China)

Abstract: To address the issues of insufficient trajectory feature extraction, low prediction accuracy, and stability in existing methods for ship trajectory prediction, especially for military ships with sparse points and flexible maneuvering characteristics, this paper proposed an improved long short-term memory (LSTM) artificial neural network ship position prediction method. This approach was based on automatic identification system (AIS) data, considering the multi-dimensional features of the trajectory, inter-trajectory correlation features, and nautical chart constraints (NCC) for ships sailing at sea. For military ship trajectories, historical trajectories were interpolated by cubic spline interpolation to generate equidistant point data for prediction. The navigation area map was rasterized, with navigable grids defined to establish map constraints and improve prediction accuracy. Finally, when designing the LSTM-based network, chart constraints were integrated into the model training and prediction process by using a custom loss function, grid matching for predicted points, and other methods. Simulation results based on AIS data from the South China Sea show that the proposed network can effectively predict ship trajectories, especially for military ships with high maneuverability. The proposed method outperforms traditional prediction methods in both prediction accuracy and stability.

Keywords: track prediction; cubic spline interpolation; nautical chart constraint; rasterization; long short-term memory networks

Received: 2023-08-08; **Accepted**: 2023-12-29; **Published Online**: 2024-01-23 20:39
URL: link.cnki.net/urlid/11.2625.V.20240123.1121.002
Foundation items: National Natural Science Foundation of China (62271028,62101014)
* **Corresponding author**. E-mail: yangweigigi@sina.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0510

基于OMEGA理论的燃油计量装置污染卡滞特性分析

李文强[1]，王彬[1,*]，杨军杰[2]，李柯柯[2]，叶志锋[1]

(1. 南京航空航天大学 能源与动力学院，南京 210016；2. 中国航发西安动力控制科技有限公司，西安 710077)

摘 要：为获得燃油中污染颗粒对燃油计量装置(FMU)工作特性影响的定量规律，并探索一种面向非标燃油部件污染卡滞特性的分析方法，基于OMEGA理论将污染颗粒对燃油部件的作用转化为卡滞力，建立了燃油污染工况下燃油部件和装置的动力学模型。开展了固体污染度等级为GJB 420B-7/8/9级时FMU的工作特性仿真，获得了各等级燃油对其关键部件性能的影响规律。结果表明：所建电液伺服阀污染卡滞模型获得的负载流量响应与标准电液伺服阀污染卡滞模型的结果基本吻合，GJB 420B-9级燃油中计量活门最大卡滞力约140 N，响应时间为无污染时的2倍以上，计量流量最大相对误差增加4倍以上。所提方法可为非标燃油部件和系统的污染卡滞建模提供参考。

关 键 词：OMEGA理论；燃油计量装置；卡滞力；工作特性；燃油等级

中图分类号：V233.2+4

文献标志码：A **文章编号**：1001-5965(2025)10-3433-10

燃油计量装置(fuel metering unit, FMU)是航空发动机燃油调节系统的主要功能单元，其工作特性直接决定发动机的状态调节水平和可靠性。工作环境是影响其工作特性诸多因素中的关键因素之一，当前，工程中频繁出现的因油液污染导致的燃油调节性能退化或故障严重影响了发动机的控制水平和可靠性。水污染物和固体颗粒是燃油中存在的2类主要污染物[1]，本文主要讨论燃油中固体颗粒对装置的影响。FMU中的液压元件长期暴露于燃油中，其中的固体污染物颗粒(以下简称污染颗粒)极易滞留于通流阀口或运动间隙，从而加速阀内运动副磨损或堵塞阀口和间隙，对FMU工作过程及性能造成影响[2]。从破坏机理层面，一方面，污染颗粒易划伤液压元件(如泵和阀)的运动副表面，导致元件密封间隙处泄漏量增加，削减输出流量或执行元件输出[3-4]；另一方面，部分污染颗粒经过阀类元件间隙时会被截留，使运动间隙减小或消失，最终导致阀芯卡滞，使精密液压控制元件失效，无法完成既定动作[5]。统计显示，因油液污染直接或间接引起的故障数占液压系统总故障数的70%以上[6]，油液污染后，系统或部件的特性及抗污染能力直接决定系统性能和可靠性。污染对FMU主要燃油部件特性的影响规律可为评估燃油调节系统性能并提出有效抑制或补偿措施提供参考依据，对估计和提高航空发动机燃油附件可靠性具有重要意义。

目前，燃油系统的污染研究主要集中在污染防治和污染对系统寿命及可靠性的影响分析2个方面。污染防治方面，学者们主要基于燃油系统的污染传递原理和污染物测定方法，通过分析污染颗粒变化规律[7]及测定燃油的固体污染度等级(以下简称燃油等级)[8-9]，指导更换油液或设计和优化配置过滤器[10-12]，从而使元件工作在燃油等级安全的油液中。影响分析方面，研究者们主要基于污染机理和理论对典型元件所受的污染影响开展定量化分析，据此评估其可靠性。张育洋[13]通过试验确定了柱塞泵污染磨损寿命模型，并基于该模型对柱塞泵进行了寿命预测和可靠性分析。刘勇等[14]从颗粒

收稿日期：2023-08-04；录用日期：2023-08-25；网络出版时间：2023-10-23 13:24
网络出版地址：link.cnki.net/urlid/11.2625.V.20231023.1008.001
基金项目：国家科技重大专项(J2019-Ⅴ-0016-0111)
*通信作者. E-mail：binwang@nuaa.edu.cn

引用格式：李文强，王彬，杨军杰，等. 基于OMEGA理论的燃油计量装置污染卡滞特性分析[J]. 北京航空航天大学学报，2025，51(10)：3433-3442. LI W Q，WANG B，YANG J J，et al. Contamination lock characteristics analysis of fuel metering unit based on OMEGA theory[J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51(10)：3433-3442 (in Chinese).

破碎的微观角度出发建立了齿轮泵流量劣化模型，并进行了试验验证。王洪波等[15]通过试验获得电液伺服阀的污染卡滞敏感度系数，并结合加速寿命试验方法对阀的寿命进行了评估。郑长松等[16]基于滤饼过滤理论和颗粒分布参数影响因子建立了滑阀的污染卡紧力模型，试验验证了该模型的有效性。

当前，以燃油中污染颗粒的浓度和尺寸划分燃油等级[17]，为保证航空发动机燃油部件不受燃油污染的影响，设计人员对各类燃油元件的适用燃油等级进行了规定。当因过滤器失效或工作时间过长使燃油等级超出规定等级时，污染颗粒对燃油元件的定量化影响难以清晰表征。另外，有关污染对航空发动机FMU工作特性影响的定量化仿真研究报道较少，而传统试验法成本高，无法满足现代高性能发动机对燃油附件高可靠设计方法的迫切需求。本文通过将燃油中污染颗粒对阀类元件造成的卡滞影响转化为卡滞力建立FMU的污染卡滞模型，开展装置中关键阀类元件的定量化仿真分析，获得燃油污染时装置的动态特性，为发动机非标燃油部件抗污染评估和设计提供理论方法参考。

1 FMU污染卡滞模型

1.1 FMU工作原理

典型FMU的构成原理如图1所示，主要由计量活门、等压差活门、回油活门、电液伺服阀和位移传感器组成[18]。FMU的主要功能是将经过主燃油泵增压后的燃油定量地供给燃烧室，其工作原理为：通过电液伺服阀、计量活门和位移传感器组成的位置闭环系统实现对计量活门开度的精确控制，等压差活门与回油活门联合调节实现计量活门前后压差恒定，由薄壁小孔流量公式可知此时计量流量仅与计量活门开度成比例，由此实现燃油流量的精确控制[19]。

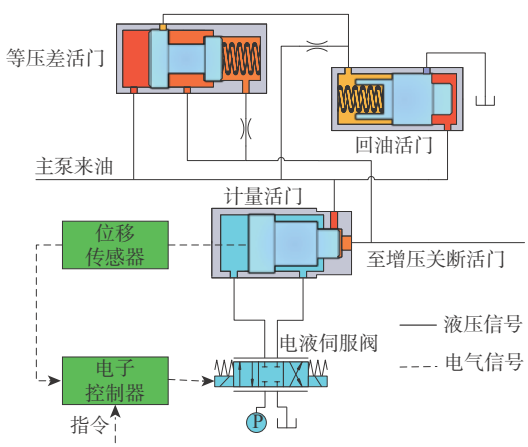


图1 FMU原理

Fig. 1 FMU schematic diagram

1.2 阀芯污染卡滞模型

为实现计量功能和保证控制性能，FMU中阀类元件的结构设计较为精密，使其易受燃油中污染颗粒的影响。例如，电液伺服阀驱动计量活门运动，是实现位置闭环控制的核心控制元件，属高精密液压件，因液压放大器运动范围小且灵敏度高，污染颗粒易滞留于功率级阀芯的运动副间隙，极大影响电液伺服阀性能和可靠性[20]。当污染颗粒淤积于阀芯的运动副间隙而使间隙完全消失时，会对阀芯的启动产生附加的卡滞效应，这种效应对于阀芯类似静摩擦力，即仅在阀芯受驱运动前存在，当阀芯开始运动后，间隙又被重新建立，此时卡滞效应消失。若阀芯的驱动力无法克服污染颗粒淤积产生的作用于阀芯上的卡滞力时，计量活门被卡死，元件功能失效。

OMEGA理论描述燃油中污染颗粒对阀芯的卡滞特性[15]，将每计量个特定尺度区间的污染颗粒对阀芯的卡滞影响用污染敏感度影响因子表达。据此，燃油中所有污染颗粒对电液伺服阀(或活门)阀芯的卡滞影响表现为污染敏感度影响因子和污染颗粒总数乘积的累积。

单一尺度区间的污染颗粒对阀芯产生的卡滞影响可描述为

$$\frac{\mathrm{d}Z_i(t)}{\mathrm{d}t_\mathrm{j}} = S_i q n_i \tag{1}$$

式中：Z_i为燃油中污染颗粒尺度位于第i个区间的污染颗粒对阀芯产生的卡滞力；S_i为阀芯对污染颗粒尺度位于第i个区间的污染颗粒卡滞敏感度；q为阀芯径向间隙泄漏流量；n_i为污染颗粒尺度位于第i个区间的污染颗粒浓度；t_j为静置时间。

式(1)表明，当泄漏流量和燃油等级不变时，卡滞力随静置时间增加线性增加，此时可将式(1)改写为增量式：

$$\Delta Z_i = S_i q n_i \Delta t_\mathrm{j} \tag{2}$$

式(2)用于试验计算污染敏感度，计算式为

$$S_i = \frac{1}{n_i}\left(\frac{Z_i}{q t_\mathrm{j}} - \sum_{g=1}^{i-1} S_g n_g\right) \tag{3}$$

式中：S_g为阀芯对尺度位于第g个区间的污染颗粒的卡滞敏感度；n_g为尺度位于第g个区间的污染颗粒浓度。

综上可得所有污染颗粒对阀芯的卡滞力为

$$Z = q\sum_{i=1}^{m} S_i n_i \Delta t_\mathrm{j} \tag{4}$$

式中：m为燃油中污染颗粒尺度分布区间数。

在阀芯静置时间t_j内系统处于静止状态，燃油

中污染颗粒以一定沉淀速率堆积于阀芯径向间隙中, 使卡滞力以一定的速率增加, 静置时间越长, 卡滞力增加越明显。通常情况下, 燃油中污染颗粒对阀芯的卡滞效应需足够静置时间方有显著表征, 因此, 燃油部件的污染仿真具有耗时长的特点。由式 (1) 可知, 可根据加速寿命试验的思想引入加速因子以缩短仿真时系统的静置用时, 此时式 (4) 可写为

$$Z = A_{\mathrm{F}} q \sum_{i=1}^{m} S_i n_i \Delta t_{\mathrm{j}} \tag{5}$$

式中: A_{F} 为加速因子, 表示将污染颗粒浓度增加为原来的 A_{F} 倍。

阀芯工作特性由式 (6) 所示的阀类元件动力学平衡方程和流量连续方程描述, 燃油中污染颗粒对阀芯造成的卡滞效应通过该式体现。

$$\begin{cases} \sum_{f=1}^{K} F_f(t) = 0 \\ \sum_{l=1}^{L} Q_l(t) + \sum Q_q(t) = 0 \end{cases} \tag{6}$$

式中: F_f 为无污染时阀芯所受的第 f 个外力 (如压差力、液动力、阻尼力等); K 为外力总数; Q_l 为阀类元件第 l 个通流孔通过的流量; L 为阀类元件通流孔数; Q_q 为因泄漏和压缩等产生的其他流量; t 为系统工作时间。

根据卡滞力表现特征, 其对阀芯的动力学模型的影响可表示为

$$Z(t) = \begin{cases} q \sum_{i}^{m} S_i n_i \Delta t_{\mathrm{j}} & v = 0 \\ 0 & v \neq 0 \end{cases} \tag{7}$$

式中: v 为阀芯运动速度。

综上可得阀芯的污染卡滞模型:

$$\begin{cases} \sum_{f=1}^{K} F_f(t) + Z(t) = 0 \\ \sum_{l=1}^{L} Q_l(t) + \sum Q_q(t) = 0 \end{cases} \tag{8}$$

当瞬时驱动力不足以克服污染颗粒对阀芯的卡滞效应时, 元件的工作特性会受到影响[21], 如响应滞后, 为便于分析这种影响, 定义电液伺服阀和活门死区分别如式 (9) 和式 (10) 所示:

$$D_{\mathrm{d}} = \frac{I_{\min}}{I_{\max}} \times 100\% \tag{9}$$

$$D_{\mathrm{h}} = \frac{t_z}{T} \times 100\% \tag{10}$$

式中: $I_{\min}$ 为使电液伺服阀产生负载流量的最小电流; $I_{\max}$ 为电液伺服阀饱和电流; t_z 为活门滞后时长, 即产生位移的最小时间; T 为活门工作时长。

1.3　仿真方法

FMU 由传感器和多个相互耦合的阀类元件组成, 以其中的电液伺服阀为例, 由式 (8) 可知, 对其进行污染卡滞仿真需建立阀芯的动力学平衡方程和流量连续方程。工程应用软件具有高效快捷的建模能力, 且不同软件具有各自优势, 本文采用 AMESim/Simulink 联合仿真方法对其进行建模仿真。FMU 污染卡滞仿真方法如图 2 所示。

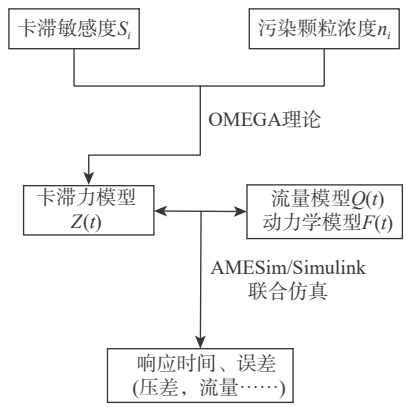


图 2　FMU 污染卡滞仿真方法

Fig. 2　Method of FMU contamination lock simulation

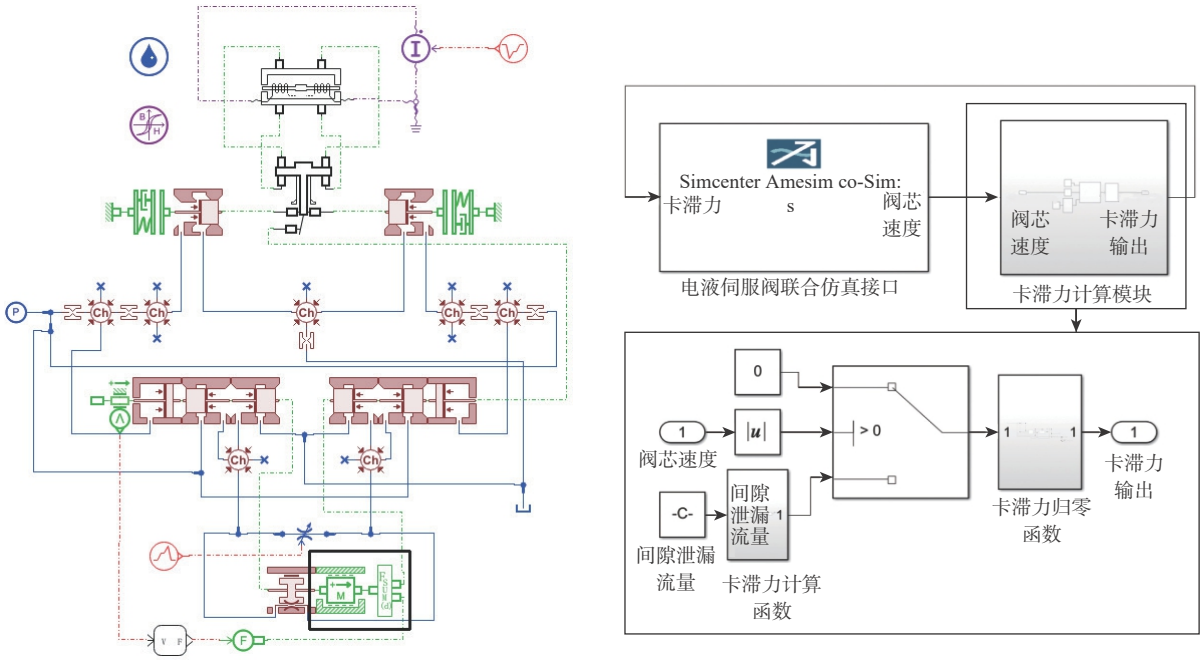
2　FMU 部件仿真验证

为验证采用联合仿真方法对 FMU 开展污染卡滞定量化分析的合理性和有效性, 基于第 1 章方法对装置中的电液伺服阀开展污染卡滞仿真研究, 将仿真结果与文献 [22] 中的典型电液伺服阀污染卡滞仿真结果进行对比。

2.1　电液伺服阀污染卡滞模型建立

基于第 1 章所述数学模型和理论分析建立的电液伺服阀污染卡滞特性联合仿真模型如图 3 所示。图 3(a) 为典型的 2 级喷嘴挡板电液伺服阀, 图 3(b) 为卡滞力计算模型。卡滞力对电液伺服阀的影响由阀芯动力学方程描述, 对应图 3(a) 中方框模型所示。

GJB 7078—2010[23] 规定了航空发动机燃油系统附件固体污染颗粒浓度等级验收水平应参照 GJB 420B—2015[17] 中所规定的等级进行确定, 且规定调节器类元件的污染度等级在 GJB 420B-6 级及以下。采用 GJB 420B-9 级燃油对该电液伺服阀进行污染卡滞特性仿真, 污染颗粒浓度如表 1 所示。该型电液伺服阀的污染卡滞敏感度如表 2 所示。由表 1 和表 2 可知, 污染卡滞敏感度的颗粒尺度区间与燃油中污染颗粒浓度的颗粒尺度区间不匹配, 这是因为用于污染试验的标准污染粉尘与测试燃油等级使用的仪器校准方法和所用标准物质不同。为简化计算, 在难以确定某一尺度区间内各尺度的污染颗粒分布的情况下, 假设尺度区间内各尺度的污染颗粒浓度服从均匀分布。



(a) 电液伺服阀系统模型　　(b) 卡滞力计算模块及联合仿真接口模型

图 3　电液伺服阀污染卡滞模型

Fig. 3　Contamination lock model of electrohydraulic servo-valve

表 1　GJB 420B-7/8/9 级燃油污染颗粒浓度

Table 1　Particle concentration of pollutants in class GJB 420B-7/8/9 fuel

污染颗粒尺度/μm	污染颗粒浓度/(个·(100 mL)$^{-1}$)		
	GJB 420B-7	GJB 420B-8	GJB 420B-9
>1	100 000	200 000	400 000
>5	38 900	77 900	156 000
>15	6 920	13 900	27 700
>25	1 200	2 450	4 900
>50	212	424	848
>100	32	64	128

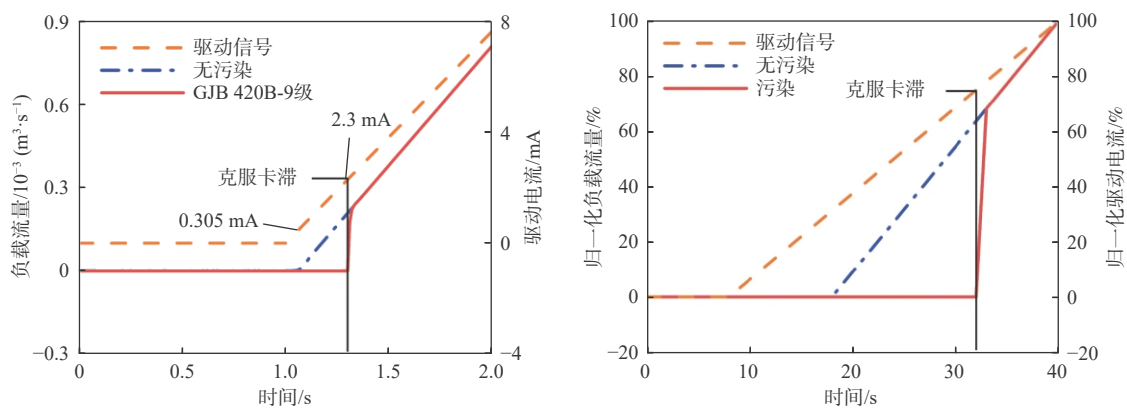
表 2　污染卡滞敏感度

Table 2　Contamination lock sensitivity

污染颗粒尺度/μm	卡滞敏感度/(N·粒$^{-1}$)
0～5	$4.564\,86\times10^{-8}$
5～10	$1.276\,56\times10^{-7}$
10～20	$3.302\,82\times10^{-7}$
20～30	$1.223\,69\times10^{-5}$
30～40	$1.150\,79\times10^{-5}$
40～50	$1.505\,53\times10^{-6}$
50～60	$3.336\,44\times10^{-6}$
60～70	0
70～80	0

2.2　仿真验证

通过分析文献 [22] 中电液伺服阀卡滞时的负载流量特性，选择对电液伺服阀施加斜坡驱动信号，根据负载流量 Q_f 曲线判断电液伺服阀污染卡滞模型的合理性和有效性。图 4 为电液伺服阀受污染卡滞影响时的负载流量曲线，图 4(a) 为本文方法取得的电液伺服阀污染卡滞负载流量曲线，该电液伺服阀的饱和电流为 15.3 mA，图 4(b) 为文献 [22] 中使用 Hypneu 软件获得的标准电液伺服阀的负载流量曲线。由



(a) AMESim/Simulink模型结果　　(b) Hypneu标准模型结果[22]

图 4　电液伺服阀负载流量特性

Fig. 4　Load flow characteristics of electrohydraulic servo-valve

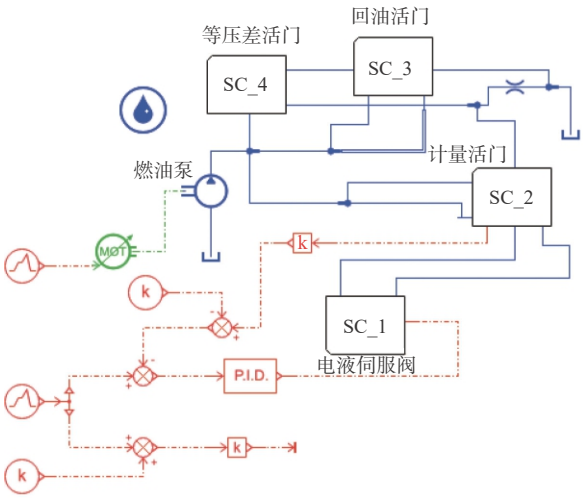
图 4(a) 可知, 燃油无污染时, 该电液伺服阀工作时的死区约为 2%, 而 GJB 420B-9 级燃油中因污染颗粒产生的卡滞力使死区增大至 15%。当驱动电流增大至可克服污染颗粒作用在阀芯上的卡滞力时, 负载流量特性恢复到无污染时的状态, 系统正常工作。

对比图 4(a) 和图 4(b) 发现, 燃油污染时, 本文方法获得的负载流量特性与文献 [22] 中的电液伺服阀污染卡滞负载流量特性表现特征相似, 即在起动阶段两者均表现出死区特性, 仅当驱动电流足以克服卡滞力时才可产生负载流量。由此得出, 根据 OMEGA 理论采用联合仿真方法对阀类元件开展污染卡滞特性仿真分析合理有效。

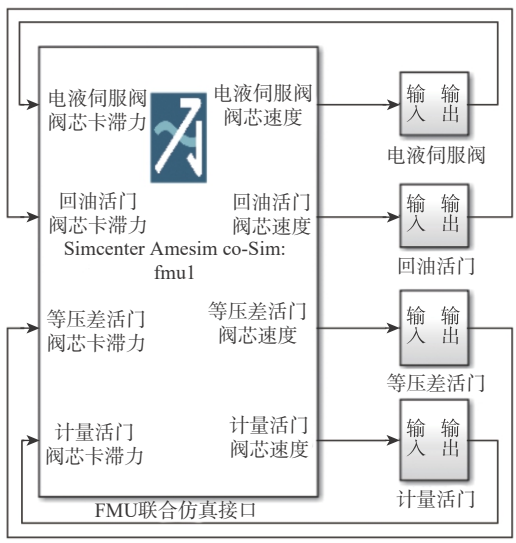
3　FMU 污染卡滞仿真分析

3.1　仿真模型建立

对 FMU 进行污染卡滞仿真, 逐个建立等压差活门、回油活门、计量活门和电液伺服阀的污染卡滞子模型, 构建整个装置模型。FMU 污染卡滞模型由系统模型和卡滞力计算模型 2 部分组成, 如图 5 所示。



(a) 系统模型



(b) 卡滞力计算模块及联合仿真接口模型

图 5　FMU 污染卡滞联合仿真模型

Fig. 5　FMU contamination lock co-simulation model

由式 (5) 可知, 污染卡滞力模型由燃油元件对燃油中污染颗粒的卡滞敏感度和燃油等级决定, 仿真采用的燃油等级分别为 GJB 420B-7/8/9。

3.2　结果分析

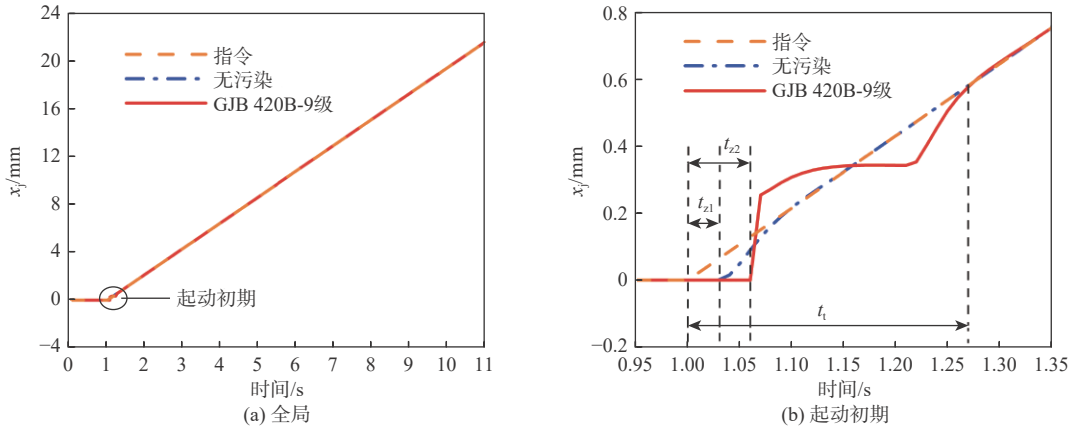
本文以某涡扇发动机主 FMU 为例, 仿真时关注的主要参数有计量活门位移 x_j、计量前后压差 P_j、等压差活门位移 x_d、回油活门位移 x_h、滞后时间 t_z、调节时间 t_t、计量流量 Q_j 和卡滞力 Z。仿真工况为: 计量活门指令在 10 s 内从零线性开至全开度, 在 GJB 420B-9 级燃油环境中进行污染卡滞仿真分析。

仿真获得的计量活门阀芯位移响应曲线如图 6 所示, 其中, t_{z1} 为固有滞后时间, t_{z2} 为污染滞后时间。图中显示, 计量活门存在 0.3% 的固有死区, 燃油无污染时计量活门可快速跟随指令信号, 满足 FMU 的调节功能要求。图 6(b) 显示, 相比无污染, GJB 420B-9 级燃油时计量活门的死区增加 1 倍, 达 0.6%, 且小开度时计量活门无法跟随指令信号, 存在明显的爬行现象, 需经过 0.28 s 以上才可完全跟随指令信号。

压差调节组件的响应曲线如图 7 所示, 其中, t_{t1} 为正常调节时间, t_{t2} 为污染调节时间。由图 7(a) 和图 7(b) 可知, 无污染时, 燃油计量前后压差可在 0.1 s 内调节至设定的平衡值, 而 GJB 420B-9 级燃油中压差调节时出现了明显震荡, 产生了大于 10 MPa 的超调, 调节过程显著变缓。由等压差活门阀芯平衡原理可知, 计量前后压差与其实时位移直接相关, 图 7(c) 和图 7(d) 的等压差活门位移响应曲线可解释这一特性。燃油无污染时, 等压差活门的固有死区为 0.5%, GJB 420B-9 级燃油时死区增大至 0.9%, 阀芯位移响应明显滞后, 且实时压差在克服卡滞力瞬时, 阀芯受力突变, 导致其位移出现明显震荡, 使计量活门前后压差产生上述变化。

另外, 计量活门前后压差的稳定受回油活门工作特性的影响。图 7(e) 和图 7(f) 为 GJB 420B-9 级燃油时回油活门的位移响应。图中显示, 无污染时, 回油活门固有死区约为 0.4%, 而 GJB 420B-9 级燃油时死区增加至 1.1%, 卡滞力使位移响应明显滞后。因回油活门受经过等压差活门节流的计量后燃油的油压驱动, 1.2～1.6 s 区间, 等压差活门位移有明显震荡, 使回油活门位移也发生了较大变化。

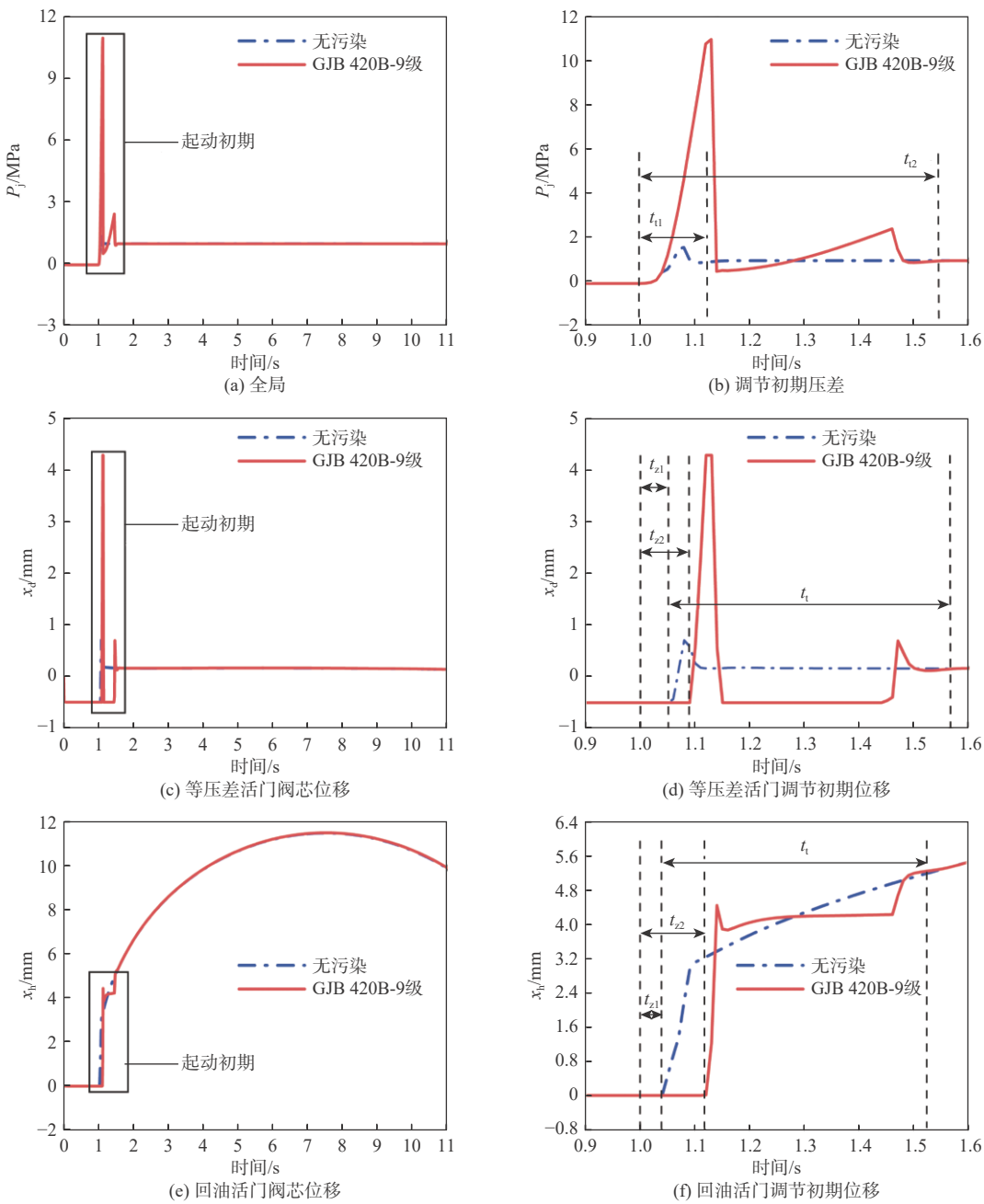
计量后流量曲线如图 8 所示。可知, GJB 420B-



(a) 全局　(b) 起动初期

图 6　GJB 420B-9 级燃油时计量活门位移响应

Fig. 6　Displacement response of metering valve with class GJB 420B-9 fuel



(a) 全局　(b) 调节初期压差

(c) 等压差活门阀芯位移　(d) 等压差活门调节初期位移

(e) 回油活门阀芯位移　(f) 回油活门调节初期位移

图 7　GJB 420B-9 级燃油时压差调节组件响应

Fig. 7　Response of differential pressure regulating assembly with class GJB 420B-9 fuel

9 级燃油中由于计量活门、等压差活门和回油活门存在上述的表现特征, FMU 起动的初始阶段计量流量出现明显震荡, 与无污染时的计量后流量相比, 最大相对误差达 468%。

污染颗粒产生的卡滞力是造成 FMU 起动阶段工作性能恶化的根本原因, 图 9 为 3 个活门所受卡滞力的变化特征。0～1 s 内系统处于静置状态, 此时卡滞力以一定速率线性增加, 1 s 时计量活门开始受控产生位移, 但此时各活门的驱动力难以克服卡滞力, 卡滞力继续增加, 其中, 计量活门卡滞力最大增加至 140 N, 当各活门的驱动力足以克服卡滞力时, 活门阀芯开始运动, 滞留于间隙内污染物被挤压破碎, 卡滞力消失。

由此得出, GJB 420B-9 级燃油中的污染颗粒对阀芯产生的卡滞力会使计量活门、等压差活门及回油活门在调节初始阶段产生明显震荡, 使响应过程延缓。相较于回油活门, 污染颗粒对计量活门和等压差活门的影响较大, 会极大限制 FMU 在起动阶段对小流量燃油的计量精度。

相同静置时间内不同等级燃油对 FMU 关键工作参数的动态影响如图 10 所示。图 10(a) 为计量

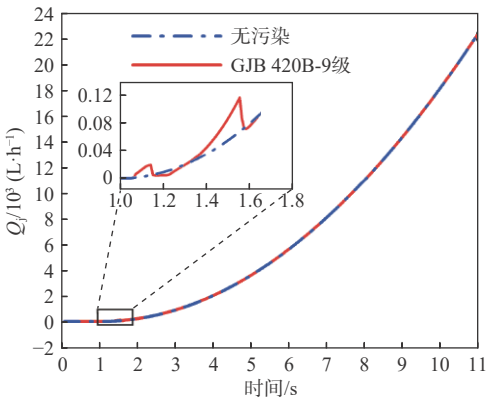


图 8　计量后燃油流量

Fig. 8　Metered fuel flow rate

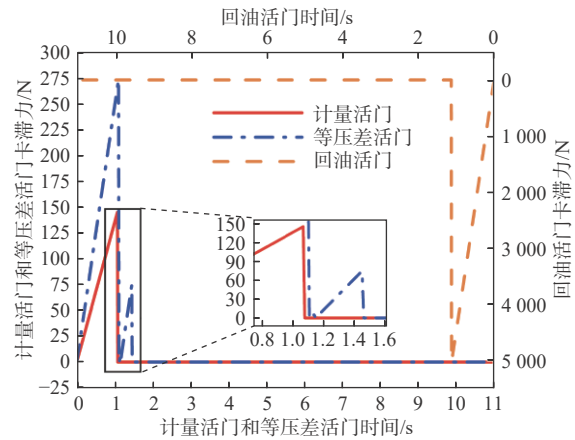
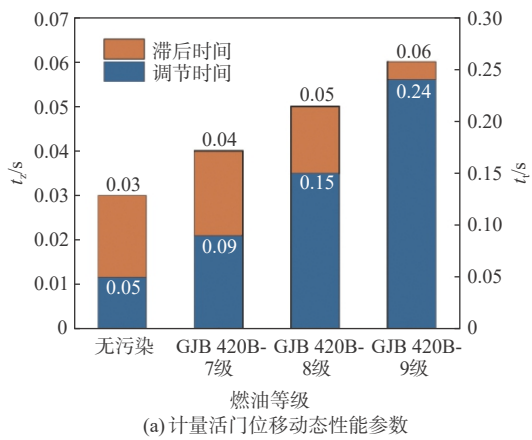
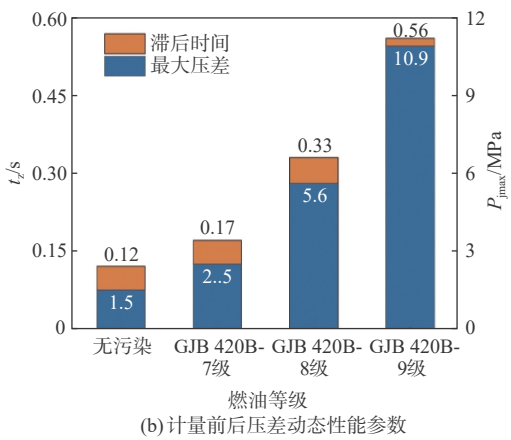


图 9　活门卡滞力

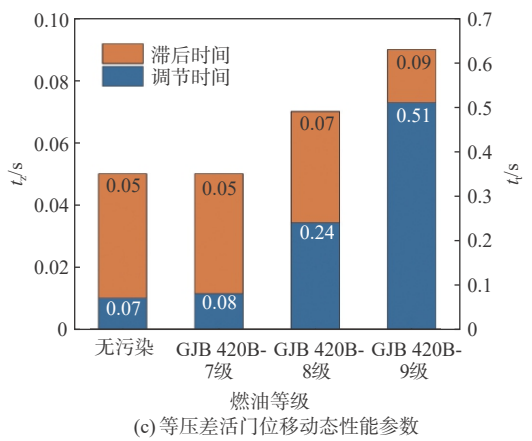
Fig. 9　Valve lock force



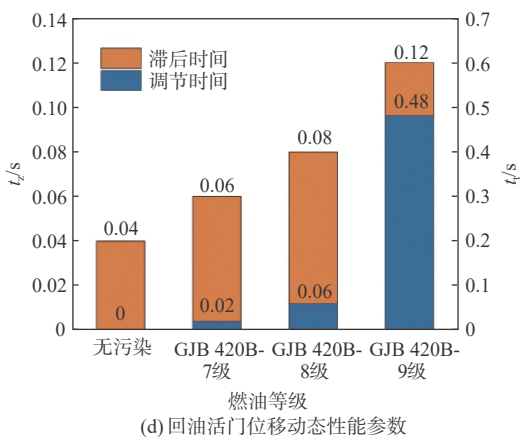
(a) 计量活门位移动态性能参数



(b) 计量前后压差动态性能参数



(c) 等压差活门位移动态性能参数



(d) 回油活门位移动态性能参数

图 10　不同等级燃油时 FMU 动态性能参数

Fig. 10　Dynamic performance parameters of FMU with different fuel classes

活门的阀芯位移动态性能参数，随着燃油等级的提高，计量活门的滞后和调节时间均明显增加，且调节时间随燃油等级提高增加更长。图 10(b) 为计量前后压差动态性能参数，滞后时间和超调量随燃油等级提高显著增加。图 10(c) 为等压差活门位移动态性能参数，GJB 420B-7 级燃油时滞后时间和调节时间未见明显增加，而 GJB 420B-8 级和 GJB 420B-9 级燃油时滞后时间和调节时间均明显增加。图 10(d) 为回油活门位移动态性能参数，随燃油等级增加，响应愈发滞后，相应地，调节时间大幅增加。仿真时，将燃油等级增加至 GJB 420B-10 级时，计量活门和等压差活门被完全卡死，系统失效。综上，GJB 420B-9 级燃油时，上述各元件的滞后时间和调节时间均大幅增加，其工作特性恶化程度明显，工程中需重点关注该燃油等级范围调节时 FMU 性能的监测或补偿。

4 结 论

基于 OMEGA 理论获得了燃油中污染颗粒对阀类元件产生的卡滞力模型，建立了某涡扇发动机主 FMU 的污染卡滞模型，在 GJB 420B-9 级燃油中开展了装置的污染卡滞仿真，分析了不同等级燃油对计量装置工作特性的影响规律，主要结论如下：

1）建立了阀类元件的污染卡滞模型，通过对比污染时本文所建电液伺服阀模型与文献 [22] 中标准电液伺服阀模型的负载流量特性，验证了所提卡滞力模型和仿真方法的合理性和有效性，该方法可为非标燃油部件和系统的污染卡滞特性分析提供参考。

2）工作在高污染等级燃油中的 FMU 关键元件能够克服卡滞力时，调节初始阶段装置受影响程度较大，具体表现为响应滞后，调节过程延缓。因计量活门前后压差调节过程中存在大幅超调，调节初始阶段难以实现小流量精确计量。GJB 420B-9 级燃油中，计量活门响应时间增加 1 倍以上，计量流量最大相对误差增大 4 倍以上。

3）燃油等级增加会使 FMU 的工作特性迅速恶化，甚至使活门阀芯完全卡死而致系统失效。FMU 的回油活门抗污染能力较强，设计时应重点针对电液伺服阀、计量活门和等压差活门开展抗污染设计或控制系统补偿。

参考文献（References）

[1] 杨文举, 邵垒, 刘卫华, 等. 飞机燃油箱水污染物数值建模[J]. 北京航空航天大学学报, 2024, 50(11): 3578-3586.
YANG W J, SHAO L, LIU W H, et al. Numerical modeling of water contaminants in aircraft fuel tank[J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50(11): 3578-3586 (in Chinese).

[2] 赵维义, 李岩, 王占勇, 等. 飞机液压系统油液的污染与控制[J]. 机床与液压, 2012, 40(10): 157-159.
ZHAO W Y, LI Y, WANG Z Y, et al. Pollution and control of oil in hydraulic system of aircraft[J]. Machine Tool & Hydraulics, 2012, 40(10): 157-159(in Chinese).

[3] 李志强. 高效精油过滤器的设计及在线检测和监控[D]. 南京: 南京理工大学, 2016: 6-9.
LI Z Q. Design and on-line detection and monitoring of high-efficiency essential oil filter[D]. Nanjing: Nanjing University of Science and Technology, 2016: 6-9(in Chinese).

[4] 郭善新, 陈达贵, 张禹, 等. 高压齿轮泵污染磨损的寿命模型研究与验证[J]. 流体机械, 2018, 46(5): 13-17.
GUO S X, CHEN D G, ZHANG Y, et al. Research and verification of life model for pollution wear of high pressure gear pump[J]. Fluid Machinery, 2018, 46(5): 13-17(in Chinese).

[5] 罗宇轩, 林振浩, 母娟, 等. 阀芯卡滞的研究进展[J]. 机电工程, 2022, 39(7): 865-876.
LUO Y X, LIN Z H, MU J, et al. Research progress of valve core clamping[J]. Journal of Mechanical & Electrical Engineering, 2022, 39(7): 865-876(in Chinese).

[6] 郑帅可. 液压元件污染敏感度分析及测试系统研制[D]. 杭州: 浙江大学, 2020.
ZHENG S K. Analysis of contamination sensitivity of hydraulic components and development of testing system[D]. Hangzhou: Zhejiang University, 2020(in Chinese).

[7] 卢继霞, 郭瑞宇, 王成, 等. 基于固体颗粒尺寸分布的微孔滤膜堵塞机理分析[J]. 机械工程学报, 2018, 54(20): 145-151.
LU J X, GUO R Y, WANG C, et al. Fouling mechanism analysis on micro-pore filter membrane based on the solid particle size distribution[J]. Journal of Mechanical Engineering, 2018, 54(20): 145-151 (in Chinese).

[8] 刘阁, 张梅菊, 王文瑾, 等. 光阻法油液污染度检测传感器的设计与实现[J]. 测控技术, 2022, 41(3): 18-25.
LIU G, ZHANG M J, WANG W J, et al. Design and implementation of oil contamination detection sensor based on photoresist method[J]. Measurement & Control Technology, 2022, 41(3): 18-25 (in Chinese).

[9] SRIDHAR V, CHANA K S, SINGH D. Computational and experimental study of a platinum thin-film based oil condition and contamination sensor[C]//Proceedings of the ASME Turbo Expo 2017 Turbomachinery Technical Conference and Exposition. New York: ASME, 2017: 1-7.

[10] 姚金勇, 李娇, 张坤, 等. 液压系统污染性能仿真与优化设计[J]. 液压与气动, 2014, 38(3): 48-53.
YAO J Y, LI J, ZHANG K, et al. Pollution performance simulation and optimization design for hydraulic system[J]. Chinese Hydraulics & Pneumatics, 2014, 38(3): 48-53(in Chinese).

[11] 张人才. 主动控制油液污染来提高航空液压系统的可靠性及寿命[D]. 北京: 中国地质大学 (北京), 2017: 47-53.
ZHANG R C. Actively control oil pollution to improve the reliability and life of aviation hydraulic system[D]. Beijing: China University of Geosciences (Beijing), 2017: 47-53(in Chinese).

[12] JI H, NIE S L, BAI X R. Simulation on mechanism of contamination mitigation through Higee and hydrocyclone techniques in fluid power system[J]. Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering, 2018, 232(1): 77-93.

[13] 张育洋. 基于污染磨损试验的液压柱塞泵寿命评估技术研究[D]. 西安: 西安电子科技大学, 2019.
ZHANG Y Y. Research on life assessment technology of hydraulic piston pump based on pollution wear test[D]. Xi'an: Xidian University, 2019(in Chinese).

[14] 刘勇, 马彪, 张赛飞, 等. 油液污染颗粒引起的齿轮泵劣化失效研究[J]. 农业机械学报, 2015, 46(2): 316-321.
LIU Y, MA B, ZHANG S F, et al. Failure analysis of degradation induced by contaminant particles in gear pump[J]. Transactions of the Chinese Society for Agricultural Machinery, 2015, 46(2): 316-321(in Chinese).

[15] 王洪波, 黄智鹏, 张向英, 等. 基于加速寿命试验方法的伺服阀污染卡滞寿命分析[J]. 液压与气动, 2022, 46(7): 131-135.
WANG H B, HUANG Z P, ZHANG X Y, et al. Analysis of fouling and sticking life of servo valve based on accelerated life test method[J]. Chinese Hydraulics & Pneumatics, 2022, 46(7): 131-135(in Chinese).

[16] 郑长松, 葛鹏飞, 李芸辉, 等. 液压滑阀污染卡紧及滤饼形成机制研究[J]. 润滑与密封, 2014, 39(8): 14-19.
ZHENG C S, GE P F, LI Y H, et al. Contaminant lock force and filter cake forming mechanism of hydraulic spool valves[J]. Lubrication Engineering, 2014, 39(8): 14-19(in Chinese).

[17] 中国航空工业第二集团. 航空工作液固体污染度分级: GJB 420B—2015[S]. 北京: 总装备部军标出版发行部, 2016: 1-3.
The Second Group of China Aviation Industry. Solid particle contamination classes for fluid of aviation: GJB 420B—2015[S]. Beijing: Military Standard Publishing and Distribution Department of General Equipment Department, 2016: 1-3(in Chinese).

[18] ZHANG Y, WANG S P, SHI J, et al. Evaluation of thermal effects on temperature-sensitive operating force of flow servo valve for fuel metering unit[J]. Chinese Journal of Aeronautics, 2020, 33(6): 1812-1823.

[19] 来晨阳, 郭迎清, 于华锋. 基于 RF-SVR 的燃油计量装置性能衰退检测和剩余寿命估计方法[J]. 航空动力学报, 2019, 34(7): 1624-1632.
LAI C Y, GUO Y Q, YU H F. Fuel metering unit performance degradation detection and remaining useful life estimation method based on RF-SVR[J]. Journal of Aerospace Power, 2019, 34(7): 1624-1632(in Chinese).

[20] 陆亮, 夏飞燕, 阎耀保, 等. 小球式旋转直驱压力伺服阀卡滞机理研究[J]. 浙江大学学报 (工学版), 2019, 53(7): 1265-1273.
LU L, XIA F Y, YIN Y B, et al. Spool stuck mechanism of ball-type rotary direct drive pressure servo valve[J]. Journal of Zhejiang University (Engineering Science), 2019, 53(7): 1265-1273(in Chinese).

[21] 朴英, 张绍基. 某涡扇发动机加力供油系统排故试验及仿真研究[J]. 航空动力学报, 1995, 10(4): 354-356.
PIAO Y, ZHANG S J. Troubleshooting tests and simulation for a reheat fuel system of a turbofan engine[J]. Journal of Aerospace Power, 1995, 10(4): 354-356(in Chinese).

[22] WANG X H, YAN Y J, LI Q X. Contaminant lock life prediction of double flapper-nozzle servo valve[J]. Advanced Materials Research, 2014, 933: 953-958.

[23] 中国人民解放军驻 113 厂军事代表室, 中国一航第 113 厂. 航空发动机燃油系统附件污染度要求: GJB 7078—2010[S]. 北京: 总装备部军标出版发行部, 2010: 1-2.
Military Representative Office of the Chinese People's Liberation Army Stationed at Factory 113, No.113 Factory of Aviation Industry Corporation of China. Fuel systems, components, aero-engine, requirements for contamination: GJB 7078—2010[S]. Beijing: Military Standard Publishing and Distribution Department of General Equipment Department, 2010: 1-2(in Chinese).

Contamination lock characteristics analysis of fuel metering unit based on OMEGA theory

LI Wenqiang[1], WANG Bin[1,*], YANG Junjie[2], LI Keke[2], YE Zhifeng[1]

(1. College of Energy and Power Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China;
2. Xi'an Aero-Engine Controls Technology Company Limited, Aero Engine Corporation of China, Xi'an 710077, China)

Abstract: To quantify the impact of contamination particles on the operational characteristics of a fuel metering unit (FMU) and to explore an analytical method for the contamination lock characteristics of non-standard fuel components, OMEGA theory was applied to translate the effect of these particles into lock forces on fuel components. A dynamic model for the components and units under fuel contamination conditions was developed. The operational characteristics of an FMU using class GJB 420B-7/8/9 fuel were simulated, and the effects of each fuel class on key components were analyzed. The results show that the load flow response from the servo-valve contamination lock model aligns closely with the standard servo-valve lock model. The maximum lock force of the metering valve reaches 140 N for class GJB 420B-9 fuel. The response time is more than twice as long as that under non-contaminated conditions, and the maximum relative error in the metered fuel increases by a factor of more than four. The method proposed in this paper provides a valuable reference for modeling contamination lock effects in non-standard fuel components and systems.

Keywords: OMEGA theory; fuel metering unit; lock force; operational characteristic; fuel class

Received: 2023-08-04; **Accepted**: 2023-08-25; **Published Online**: 2023-10-23 13:24
URL: link.cnki.net/urlid/11.2625.V.20231023.1008.001
Foundation item: National Science and Technology Major Project (J2019-Ⅴ-0016-0111)
* **Corresponding author**. E-mail: binwang@nuaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0526

空心涡轮叶片残芯检测的双模成像方法

于洋懿[1,2]，杨祎罡[1,2,*]

(1. 清华大学 工程物理系，北京 100084；2. 清华大学 粒子技术与辐射成像教育部重点实验室，北京 100084)

摘 要：空心涡轮叶片的残芯检测对于航空发动机的生产制造具有重要价值。目前，基于电子加速器的双模成像方法已经实现了中子与X射线的同光路融合成像，可以从材料属性角度对叶片的基体材料与残芯材料进行分辨。针对双模成像中的射线能谱硬化问题，采用自动分段阈值设定的方法，实现了无需了解叶片几何构造先验知识和人工参与判断，直接检测出叶片中残芯存在区域的自动化无损检测新方法。当单个像素的中子计数不低于10^4时，双模成像方法对残芯识别的误差不超过1%。

关 键 词：双模成像；中子成像；无损检测；残余型芯；空心涡轮叶片

中图分类号：V447.1；O571.53

文献标志码：A **文章编号**：1001-5965(2025)10-3443-08

对空心涡轮叶片中的残余型芯(简称残芯)进行有效地检测，对于保证航空发动机的正常工作具有重要意义[1]。为了满足该需求，数十年来，开展了包括内窥镜法、流量法[2]、X射线成像[3]、中子成像[4-5]等不同方法的应用研究。在这些技术中，基于射线透射成像的方法在检测效果和检测速度方面均具有突出的优点，具备大规模推广的可能性。

为了开展射线透射成像，所使用的射线应兼具2方面的能力：穿透性和灵敏度。穿透性是指射线不应被叶片基体材料本身强烈衰减，以保证有足够多的射线能够穿过叶片并在探测器内形成计数，满足统计精度要求。灵敏度是指应该有足够多的射线被残芯(通过吸收或散射)衰减掉，以使探测器记录到的计数有显著下降。很显然，穿透性和灵敏度之间存在矛盾关系，适用的射线应能兼顾这2种能力。

X射线和中子射线是目前常见的2种具备穿透力的射线类型，基于它们的成像技术均得到了广泛的研究。相对来说，X射线成像技术更加成熟，在工业生产中已经得到了大量应用[6-7]。

单一射线成像技术在物质识别方面存在不足。为了解决该问题，本文提出一种基于电子加速器的融合X射线和中子成像的方法，发展了物质分析算法，实现了像素级别的物质识别能力，对空心涡轮叶片中残芯的自动检出具有重要意义。

1 射线成像检测

1.1 X射线成像技术及其不足

X射线在穿透物体时，其衰减遵循指数下降规律(Lambert-Beer定律[8])：

$$I_{\mathrm{X}} = I_{\mathrm{X},0}\mathrm{e}^{-N\sigma_{\mathrm{X}}D} \tag{1}$$

式中：$I_{\mathrm{X},0}$和I_{X}分别为探测器收集到的入射和透射X射线的强度，1/(cm^2·s)；N为被穿透物体的原子数密度，1/cm^3；D为被穿透物体的厚度，cm；σ_{X}为X射线与物质原子反应的平均微观截面，cm^2。

基于式(1)，在实际中通过测量$I_{\mathrm{X},0}$和I_{X}可以实现对分析叶片的测量，有

$$N\sigma_{\mathrm{X}}D = \ln\left(I_{\mathrm{X},0}/I_{\mathrm{X}}\right) \tag{2}$$

式(2)可进一步改写为

收稿日期：2023-08-17；录用日期：2023-09-14；网络出版时间：2023-10-10 13:16
网络出版地址：link.cnki.net/urlid/11.2625.V.20231009.1047.001
基金项目：国家自然科学基金(11735008)
*通信作者. E-mail：yangyigang@mail.tsinghua.edu.cn

引用格式：于洋懿，杨祎罡. 空心涡轮叶片残芯检测的双模成像方法[J]. 北京航空航天大学学报，2025，51(10)：3443-3450.
YU Y Y，YANG Y G. Bi-modal imaging method for detection of residual core in hollow turbine blades[J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51(10)：3443-3450 (in Chinese).

$$\mu_{X,m}t_m = \ln\left(I_{X,0}/I_X\right) \tag{3}$$

式中：$\mu_{X,m}$为X射线与叶片中原子发生相互作用时的质量衰减系数，cm^2/g；t_m为叶片的质量厚度，g/cm^2。

作为高能光子，X射线在与原子发生光原子反应时，主要是与原子核的库仑场和核外电子发生相互作用，在X射线能量一定的情况下，式(1)～式(3)中σ_X和$\mu_{X,m}$均由物质的原子类型决定，且随着原子序数Z的变化而平滑变化，因此，X射线技术对原子序数差异不大的物质无法实现区分。此外，由于高Z原子具有更强的核库仑场和更多的核外电子，则高Z原子的σ_X和$\mu_{X,m}$也均更大，因此，X射线对于高Z物质更加敏感。

在空心涡轮叶片残芯检测应用场景中，基体材料(镍基高温合金[9])中的Ni元素的原子序数Z_{Ni}=28，而残芯的主要构成为相对原子序数较低的氧化铝(Z_O=8，Z_{Al}=13)或氧化硅(Z_{Si}=14)。在使用X射线对其进行检测时，相对于基体材料，残芯的σ_X和$\mu_{X,m}$均更小，通过测量$I_{X,0}$和I_X，利用式(3)得到的结果为

$$\ln\left(I_{X,0}/I_X\right) = \mu_{X,m}t_m = \mu_{X,m}^{core}t_m^{core} + \mu_{X,m}^{Ni}t_m^{Ni} \tag{4}$$

式中：上标core和Ni分别表示残芯和基体材料。由于$\mu_{X,m}^{core}t_m^{core} \ll \mu_{X,m}^{Ni}t_m^{Ni}$，在开展X射线成像时，对于残芯的分析灵敏度不高，统计噪声对$\ln\left(I_{X,0}/I_X\right)$的影响会严重干扰对残芯含量$t_m^{core}$的判断，这也使得X射线成像技术在叶片残芯检测的应用中受到了来自物理规律层面的限制。图1为利用平均能量为276.2 keV的X射线对镍基高温合金叶片的成像检测结果。

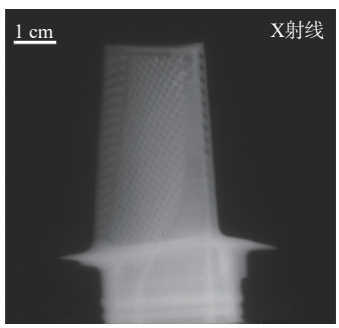


图1 含残芯空心涡轮叶片的X射线图像

Fig. 1 X-ray image of hollow turbine blade containing residual core

1.2 中子成像技术及其特点

式(4)是利用射线进行残芯检测的物理基础，由于t_m^{core}总是很小(这是由生产工艺决定的)，在$\ln\left(I_{X,0}/I_X\right)$的统计噪声总是存在的前提下，为了获得更好的残芯分析结果，需要在一定程度上扭转$\mu_{X,m}^{core}t_m^{core} \ll \mu_{X,m}^{Ni}t_m^{Ni}$的情况，从而使残芯对射线的衰减更加显著。与X射线不同，中子不是与原子核的库仑场或核外电子发生作用，而是通过核力与原子核发生作用。作为短程力，核力只有当中子非常靠近物质原子核时才会变得显著，这导致中子的波函数在核内与核外出现巨大的差异。在波函数于核边界处连续且平滑的约束条件下，不同的原子核对于入射中子波函数的影响也是不同的，该影响既可类似式(2)中的σ_X，将截面表示为σ_n，其中，下标n表示中子，也可以用散射长度b_n(单位：fm，10^{-15} m)表示散射波函数的振幅[10]。此时，原子核内的核子(中子、质子)数会显著影响入射中子在核内的波函数，进而影响截面σ_n或散射长度b_n，即使是相邻元素(甚至是相邻核素)，也会因此出现巨大的σ_n或b_n差异(本文只考虑热中子能区的中子，中子波函数的相干现象相比于其他相互作用的影响可忽略不计，暂时不对b_n开展讨论，可在单晶材料晶格取向分析的应用中再对b_n进行讨论)，这就为式(4)中残芯权重的增大提供了可能。具体地，可以在残芯中添加一定量截面σ_n巨大的中子敏感核素，在不影响制造工艺的情况下显著提升中子分析的灵敏度。

根据核素数据[11]，^{155,157}Gd、^{10}B、^{6}Li等是常见的大截面中子衰减核素，又尤以^{155,157}Gd的截面为最大，因此，Gd元素成为了优选的残芯掺杂元素。参考实际的中子成像检测案例，通过在残芯材料中掺入一定量(质量分数为3%[4])的Gd_2O_3，可显著提升残芯材料对中子的衰减，因此，有

$$\ln\left(I_{n,0}/I_n\right) = \mu_{n,m}t_m = \mu_{n,m}^{core(Gd)}t_m^{core(Gd)} + \mu_{n,m}^{Ni}t_m^{Ni} \tag{5}$$

式中：$I_{n,0}$和I_n分别为探测器收集到的入射和透射中子的强度；上标core(Gd)表示掺入Gd_2O_3后的残芯；$\mu_{n,m}$为中子与叶片中原子发生相互作用时的质量衰减系数，cm^2/g。

基体材料和残芯对中子和X射线的质量衰减系数及典型的质量厚度如表1所示。

1.3 单一射线法的不足

图2为引入Gd_2O_3材料后含有残芯模拟物的叶片的中子成像检测结果。可以看出，中子成像方法显著提升了对残芯的检出能力。然而，参考表1中的数据，由于残芯的含Gd量和残芯本身的尺寸都不可能很大，相对于基体材料，残芯对中子的衰减并不占优，基体材料仍是中子衰减的主要影响因素。因此，仅仅通过分辨图2各像素的灰度值(对应中子衰减量)来进行残芯的判断是不可靠的，因为“小质量厚度基体材料+残芯”和“大质量厚度基体材料”对中子的衰减可能是一样的，这是由物理规律的限制决定的。

由式(4)和式(5)可知，在每次测量的单一结果

表 1　利用 X 射线或中子进行残芯检测时的质量衰减系数、质量厚度和射线衰减数据

Table 1　Mass attenuation coefficient, mass thickness, and attenuation data for residual core inspection using X-rays and neutrons

射线种类	质量衰减系数/($cm^2 \cdot g^{-1}$)		典型质量厚度/($g \cdot cm^{-2}$)		射线衰减		残芯衰减占比/%
	基体材料	残芯	基体材料	残芯	基体材料	残芯	
X射线	0.262	0.156	4.39	0.054 0	1.156	0.008 4	0.73
中子	0.188	3.320	4.39	0.054 0	1.004	0.180 0	17.9

注: 表中X射线与中子的质量衰减系数根据清华大学9 MeV双模成像系统的能谱与ENDF/B-Ⅷ数据共同计算得到, 其中, 典型的基体材料厚度设定为5 mm, 残芯厚度设定为200 μm。

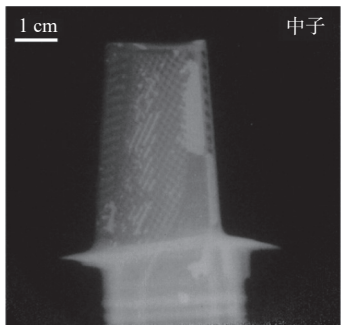


图 2　含残芯空心涡轮叶片的中子图像

Fig. 2　Neutron image of hollow turbine blade containing residual core

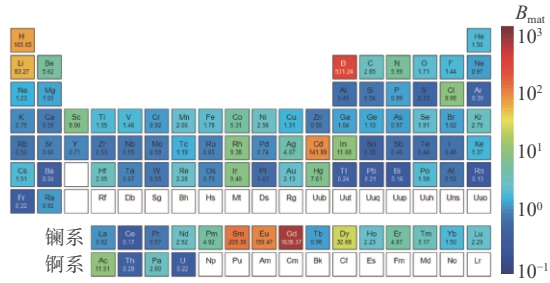


图 3　各元素的B_{mat}值

Fig. 3　B_{mat} values of each element

$\ln\left(I_{X,0}/I_X\right)$或$\ln\left(I_{n,0}/I_n\right)$中, 体现材料属性和材料空间分布的 2 个未知量(μ_m和t_m)耦合在一起, 即利用单一射线进行成像, 无法通过像素信息获得μ_m并实现物质识别。因此, 中子成像技术虽然相对于 X 射线成像技术显著提升了残芯在叶片中的对比度, 但仍然不能保证对残芯的直接检出。

为了在材料的类型上区分残芯与叶片基体, 提升对空心涡轮叶片中残芯的检测能力, 将 2 种射线的成像结果进行融合是一种可行的方法。

2　原理和方法

将式 (4) 和式 (5) 做比, 得到式 (6), 可以看出, 2 个公式共有的未知量t_m被消去, 只留下 2 个质量衰减系数的比值B_{mat}(无量纲数)。由于质量衰减系数由原子/原子核的固有属性决定, 因此, 根据式 (6), 每个元素(核素)都具有不同的B_{mat}值, 据此可以实现对物质属性的识别。图 3 为元素周期表中各元素对应的B_{mat}值。

$$\frac{\ln\left(I_{n,0}/I_n\right)}{\ln\left(I_{X,0}/I_X\right)}=\frac{\mu_{n,m}t_m}{\mu_{X,m}t_m}=\frac{\mu_{n,m}}{\mu_{X,m}}=B_{mat} \tag{6}$$

在残芯材料中添加的 Gd 元素相比基体材料中含有的 Ni、Cr 等元素, 其B_{mat}值高约 3 个数量级。这保证了在双模测量中对于空心涡轮叶片中残芯的高灵敏度。

2.1　双模成像的实现方法

国际上的双模成像技术主要依托反应堆中子源或散裂中子源来开展[12-15], 这些大型中子源提供的中子具有良好的准直比(通常>100:1), 因此, 中子束流可以近似地被看做平行束。然而, X 光机提供的 X 射线束流是锥束, 与中子束流的成像光路存在差异。在实际测量中, 这种差异将导致 2 种射线的平面成像结果存在视差而无法直接进行融合。

本文提出一种基于电子加速器的中子光子双模成像方法[16-17]。具体方案是: 利用电子加速器产生的高能电子脉冲(9 MeV, 脉宽 5 μs)沿图 4 中 x 轴负方向轰击金属靶, 通过韧致辐射产生高能 X 射线, 其中, 部分 X 射线在光中子转换体内通过光核反应制造出光中子。这些被制造出的中子和 X 射线在后续的靶材料内发生散射, 并在与入射电子方向垂直的 z 轴方向被引出, 成为成像 X 射线或成像中子。电子加速器工作在脉冲模式下, 中子与 X 射线可被视为同时产生。对于 X 射线, 其飞行速度为光速, 到达 5 m 外的探测器处仅需要 16.7 ns, 而对于热中子, 其平均飞行速度约 2 200 m/s, 到达探测器处需要 2.27 ms, 因此, 可以使用飞行时间法在探测器处对 X 射线与中子进行分辨, 并先后得到成像结果。

双模成像方案的示意如图 4 所示。成像 X 射线和中子的产生与测量使用了相同的射线源和探测器系统。这保证了 X 射线和中子的成像几何是严格同光路的。因此, 该双模成像系统测量得到的双射线图像具备在二维平面进行逐像素匹配融合的条件, 使平面的双模成像检测成为可能。

2.2　物质属性判别方法

将式 (6) 提出的材料属性判别方法应用于残芯检测的场景时, 纯基体材料部分和含有残芯的部分将会对应 2 种不同的B_{mat}值, 分别为

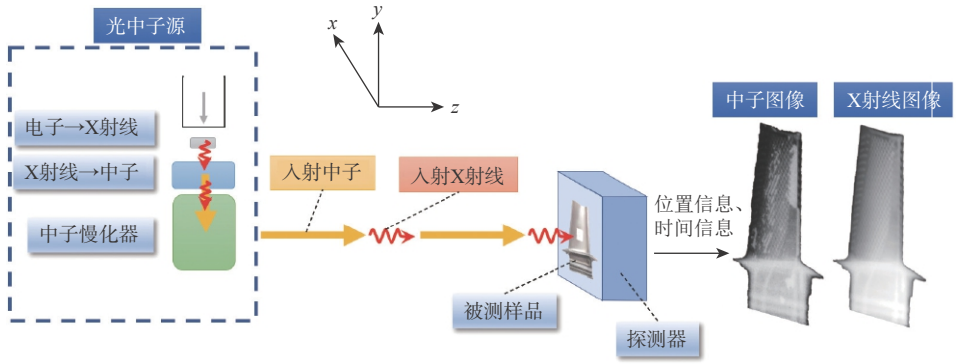


图4 基于电子加速器的双模成像测量示意图

Fig. 4 Bi-modal imaging measurement based on e-LINAC

$$B_{\mathrm{mat}}=\frac{\mu_{\mathrm{n,m}}^{\mathrm{Ni}}t_{\mathrm{m}}^{\mathrm{Ni}}}{\mu_{\mathrm{X,m}}^{\mathrm{Ni}}t_{\mathrm{m}}^{\mathrm{Ni}}}=\frac{\mu_{\mathrm{n,m}}^{\mathrm{Ni}}}{\mu_{\mathrm{X,m}}^{\mathrm{Ni}}} \tag{7}$$

$$B_{\mathrm{mat}}=\frac{\mu_{\mathrm{n,m}}^{\mathrm{Ni}}t_{\mathrm{m}}^{\mathrm{Ni}}+\mu_{\mathrm{n,m}}^{\mathrm{core}}t_{\mathrm{m}}^{\mathrm{core}}}{\mu_{\mathrm{X,m}}^{\mathrm{Ni}}t_{\mathrm{m}}^{\mathrm{Ni}}+\mu_{\mathrm{X,m}}^{\mathrm{core}}t_{\mathrm{m}}^{\mathrm{core}}} \tag{8}$$

根据表1数据，由于残芯的质量厚度远小于基体材料，且2种材料的X射线质量衰减系数接近，因此，$\mu_{\mathrm{X,m}}^{\mathrm{core}}t_{\mathrm{m}}^{\mathrm{core}}<<\mu_{\mathrm{X,m}}^{\mathrm{Ni}}t_{\mathrm{m}}^{\mathrm{Ni}}$，于是式(8)的$B_{\mathrm{mat}}$值可化简为

$$B_{\mathrm{mat}}=\frac{\mu_{\mathrm{n,m}}^{\mathrm{Ni}}t_{\mathrm{m}}^{\mathrm{Ni}}+\mu_{\mathrm{n,m}}^{\mathrm{core}}t_{\mathrm{m}}^{\mathrm{core}}}{\mu_{\mathrm{X,m}}^{\mathrm{Ni}}t_{\mathrm{m}}^{\mathrm{Ni}}+\mu_{\mathrm{X,m}}^{\mathrm{core}}t_{\mathrm{m}}^{\mathrm{core}}}\approx\frac{\mu_{\mathrm{n,m}}^{\mathrm{Ni}}}{\mu_{\mathrm{X,m}}^{\mathrm{Ni}}}+\frac{\mu_{\mathrm{n,m}}^{\mathrm{core}}t_{\mathrm{m}}^{\mathrm{core}}}{\mu_{\mathrm{X,m}}^{\mathrm{Ni}}t_{\mathrm{m}}^{\mathrm{Ni}}} \tag{9}$$

式(9)表明，含有残芯部分的B_{mat}值是在纯基体材料结果的基础上，额外增加了残芯部分的贡献。由于残芯的中子质量衰减系数比基体材料的X射线质量衰减系数高约1个数量级，在残芯部分的质量厚度远小于基体材料的情况下，该额外贡献项仍可以显著改变B_{mat}值的大小。分析表1数据可知，200 μm大小的残芯可使式(9)的B_{mat}值比式(7)的B_{mat}值大17.9%，这表明了双模成像方法可以在几乎不损失分析灵敏度(表1中的纯中子衰减为17.9%)的情况下同时提供X射线分析和中子分析能力。

在实际测量中，由于存在计数统计误差，基体材料与残芯部分的B_{mat}值很难是如式(7)和式(9)所展示的单一固定数值，而将会表现为以该数值为期望的概率密度分布，由中心极限定理，可表示为高斯函数。受限于成像束线的注量率和测量时间，每个像素测量到的粒子数也将是有限的，低的像素计数会导致相对更大的高斯分布展宽，这将对不同材料B_{mat}值之间的分辨构成不利的影响。

根据式(5)，变量$I_{\mathrm{n,0}}$和I_{n}均受统计误差的影响。$I_{\mathrm{n,0}}$是无样品时某像素所累计的总中子计数，其精度可以通过事先长时间的刻度测量来改善，由于事先测量的时间可以很长，总计数相对很高，其误差的影响在计算过程中通常可以忽略，此时，可以主要考虑I_{n}统计误差的贡献(由于双模束线中X射线的束流强度比中子高约3个数量级，因此，也忽略X射线部分的贡献)，其统计误差为：$\Delta I_{\mathrm{n}}=\sqrt{I_{\mathrm{n}}}$。式(6)中$B_{\mathrm{mat}}$值的误差可被计算为

$$\frac{\Delta B_{\mathrm{mat}}}{B_{\mathrm{mat}}}=\frac{1}{\sqrt{\left(\mu_{\mathrm{n,m}}t_{\mathrm{m}}\right)^{2}I_{\mathrm{n,0}}\exp\left(-\mu_{\mathrm{n,m}}t_{\mathrm{m}}\right)}} \tag{10}$$

以表1中的数值进行计算，引入统计误差(按照单个像素计数10^5进行估计)，纯基体材料和含有不同厚度残芯材料的概率密度分布情况如图5所示。图中，蓝色区域为基体材料分布，右侧的虚线分别是在基体材料的基础上，含有200～600 μm残芯时的分布。可以看出，残芯的厚度越大，其B_{mat}值与基体材料之间的区分度越大，因此，对小尺寸残芯进行分析计算，就确定了检测精度的下限。

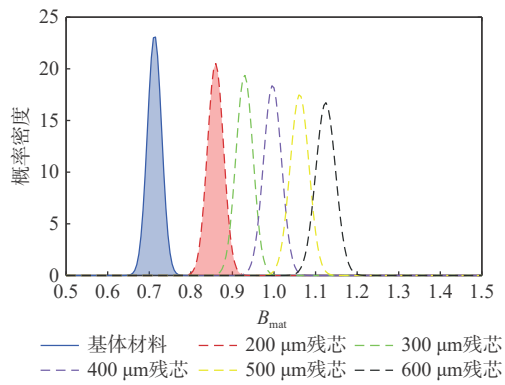


图5 纯基体材料与含有不同厚度残芯的概率密度分布

Fig. 5 Probability density distribution of substrate and residual core materials with varying core thicknesses

参考中子成像检测残芯的工作[4]，以200 μm的尺寸作为检测标准。采用受试者工作特征(receiver operating characteristic, ROC)曲线对纯基体材料与含有200 μm残芯2种情况进行分析，不同测量条件(单像素中子计数)下残芯和基体材料的ROC曲线如图6所示。根据生产检测对于检测精度的要求，可以选择合适的测量时间和检测阈值，确定具体的检测方案。

2.3 能谱硬化与校正

双模成像系统提供的成像X射线和中子射线

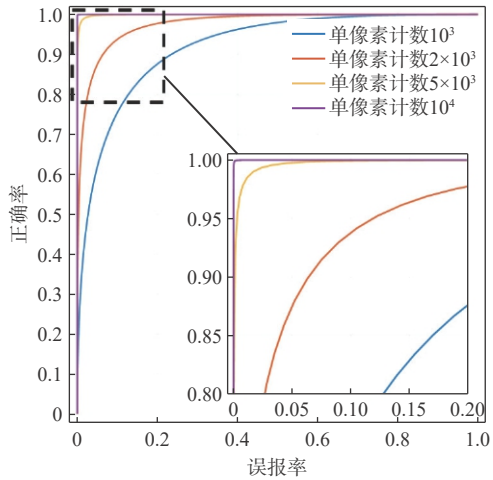
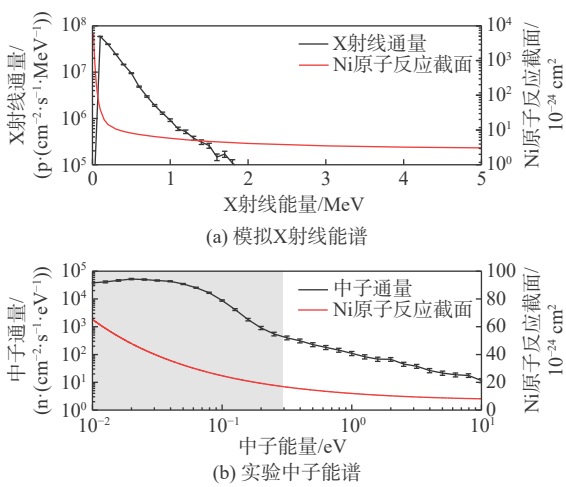


图 6　不同像素计数情况下(5 mm 厚纯基体材料与含有 200 μm 残芯)的 ROC 曲线

Fig. 6　ROC curves under different pixel counting conditions (5 mm thick pure substrate and 200 μm thick residual core)

具有一定的能量分布, 且同种材料对于不同能量射线的反应截面不同, 较低能量的射线成分被衰减的更多, 这导致射线的能谱向高能部分偏移[18]。因此, 材料的等效质量衰减系数会随材料厚度的增加而降低, 从而使 B_{mat} 值的分布随厚度的变化而变化, 影响残芯判别过程。

双模成像系统的 X 射线模拟能谱如图 7(a) 所示。这一能量区间 Ni 金属截面的主要贡献是光电效应, 其截面随能量的升高而降低, 当光子的能量从 50 keV 增加到 500 keV 时, 光子总截面下降约 2 个数量级。射线的等效质量衰减系数将随着样件厚度的增大而持续减小。中子能谱利用飞行时间法测量得到, 如图 7(b) 所示, 其中, 10～300 meV(图中标记为灰色的区域)为成像用热中峰的主要范围。当中子能量从 10 meV 增加到 300 meV 时,



(a) 模拟X射线能谱

(b) 实验中子能谱

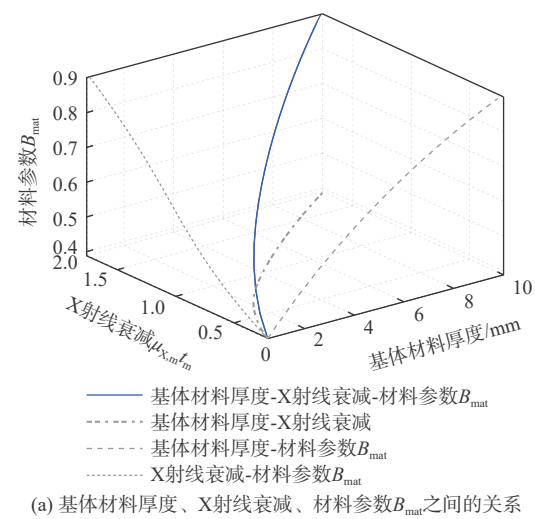
图 7　实验中子能谱和模拟 X 射线能谱

Fig. 7　Experimental neutron energy spectrum and simulated X-ray energy spectrum

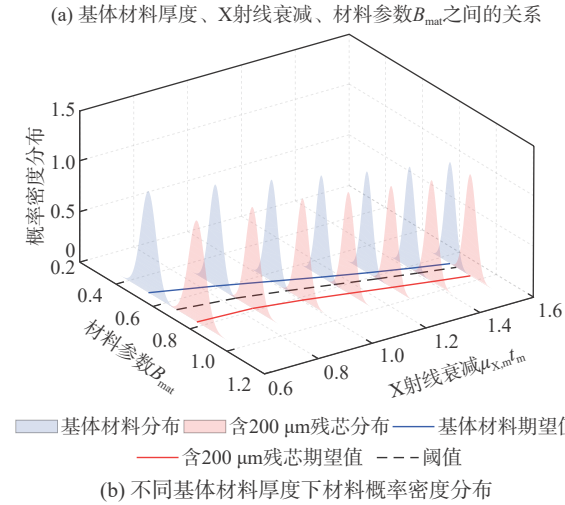
Ni 的截面降低为原来的 27%。与 X 射线相比, 中子的截面变化幅度相对较小, 其产生的硬化效应影响也相对更小。由于 2 种射线的硬化效应随厚度变化的规律不同, B_{mat} 值也会随材料厚度变化而发生偏移。

为了实现射线硬化的校正, 使用图 7 中的能谱和截面计算出 Ni 基高温合金材料在不同厚度下的 X 射线和中子的质量衰减系数。并在此基础上, 计算出 B_{mat} 值随厚度的变化关系。由于实际样品的厚度未知, 使用 X 射线衰减量来反映实际厚度的大小。图 8(a) 为合金基体材料的厚度、X 射线衰减量及 B_{mat} 值三者之间的关系。根据这一映射关系可以确定出不同 X 射线衰减下 B_{mat} 值的分布规律, 以进行能谱硬化的校正。

对于掺杂质量分数为 3% 的 Gd_2O_3、尺寸为 200 μm 的陶瓷残芯的情况, 将不同 X 射线衰减下 B_{mat} 值的概率密度分布函数绘制于图 8(b)。可以看出, 随着基体材料 X 射线衰减(基体材料厚度)的增加, 材料 B_{mat} 值的期望也逐渐增加。随着基体材料厚度的增加, 相同尺寸残芯与基体材料之间的区分



(a) 基体材料厚度、X射线衰减、材料参数B_{mat}之间的关系



(b) 不同基体材料厚度下材料概率密度分布

图 8　被测量材料的特征

Fig. 8　Characteristics of tested materials

度降低，这将导致更大的误差。图中黑色虚线为按照 1% 漏报率(漏报率指将存在残芯的区域错误判定为不含残芯的概率，相应的，误报率指将不含残芯区域误识别为含有残芯的概率)设定的判别阈值。这一阈值将随基体材料 X 射线衰减的变化而发生变化。因此，需要根据 X 射线衰减对 B_{mat} 值进行动态阈值的设定。

在检测保证足够低的漏报率前提下，可以绘制出不同厚度下的误报率。通过增加单个像素的计数率可以显著降低对 B_{mat} 值的测量误差。图 9 显示，当中子计数从每像素 1×10^3 提升到 1×10^4 时，判别阈值与基体材料间的距离从约 1 倍标准差增加到了超过 3 倍标准差，5 mm 处的误报率从 15.87% 降低到了 0.1% 以下，检出率接近含有残芯的真实水平。

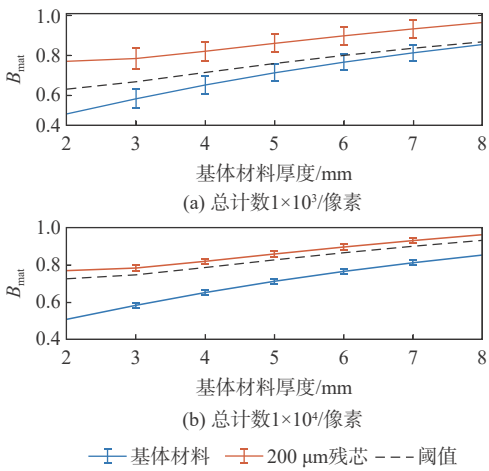


图 9 不同总计数下 B_{mat} 与基体材料厚度的关系

Fig. 9 Relationship between B_{mat} values and substrate thickness under different total counts

3 实验验证

3.1 实验装置与测量

清华大学课题组于 2020 年建设了基于 9 MeV 电子加速器的双模成像装置[15]。成像探测器选用中子敏感的微通道板 (micro channel plate, MCP) 探测器。系统利用飞行时间法分辨中子与 X 射线，整体装置的照片如图 10 所示，电子加速器和成像探测器的参数信息如表 2 所示。

在该装置上开展了空心涡轮叶片残芯检测的双模成像原型实验。被测物是由 Ni 基高温合金制作的空心涡轮叶片，在其内部空腔填充 Gd_2O_3 粉末作为残芯的模拟物。实验的样品和成像结果如图 11 所示。

3.2 厚度分区方法与识别测量结果

根据 2.2 节提出的方法，对残芯区域的识别需要用到叶片基体材料 B_{mat} 值的概率密度分布。对于实际叶片，残芯的厚度与分布随机且未知，因此，只

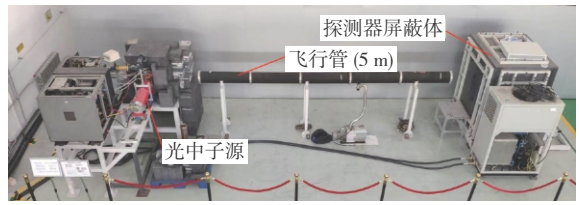


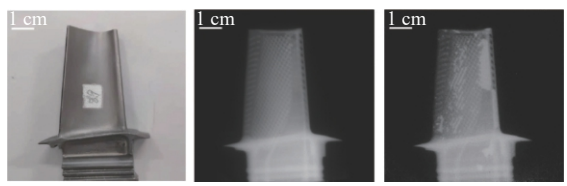
图 10 双模成像装置

Fig. 10 Bi-modal imaging system

表 2 双模成像装置参数

Table 2 Parameters of bi-modal imaging system

电子加速器参数	数值	成像探测器参数	数值
电子能量/MeV	9	探测器类型	nMCP
重复频率/Hz	100	灵敏区尺寸/(mm×mm)	95×95
中子总产额/(n·s^{-1})	2×10^{11}	分辨率/μm	＜200
中子通量(10 m飞行距离)/(cm^{-2}·s^{-1})	2 500	读出方式	sCMOS
光子通量(10 m飞行距离)/(cm^{-2}·s^{-1})	1.03×10^6	像素尺寸/μm	～50
准直比	100/1	像素数/(像素×像素)	2 160×2 560



(a) 空心涡轮叶片样品实物 (b) X射线成像结果 (c) 中子成像结果

图 11 空心涡轮叶片样品实物、X 射线成像结果和中子成像结果

Fig. 11 Photo of hollow turbine blade and its X-ray and neutron imaging results

能根据基体材料的分布情况计算单边阈值以划分材料属性。校正射线硬化需要划分厚度区间，近似地认为在每个厚度区间内，材料的 B_{mat} 值同分布。厚度区间的判断基于图像 X 射线的衰减强度。

在单一区间内，含有残芯部分的 B_{mat} 值显著高于基体材料部分。选取 B_{mat} 值低的区域可以避免基体材料的拟合数据中出现残芯区域而引入误差(图 12 中 B_{mat} 值分布极大值以下的灰色区域)。拟合函数采用高斯分布函数，得到拟合分布结果后，再按照误差标准计算出阈值所在位置(图 12 中虚线)。大于阈值的区域判定为残芯区，其他为基体材料区。

将单层的残芯判断方法推广到各个不同的质量厚度区间，数据分析流程如下：

1) 融合各像素的 X 射线和中子的衰减，得到该像素的 B_{mat} 值，并形成叶片的 B_{mat} 值分布图。

2) 根据 X 射线衰减对叶片的质量厚度区间进行划分，单个区间内像素数目>10^4，使统计误差影响<1%。

3) 针对步骤 2)中的各个质量厚度区间，分析由纯基体材料(无残芯)决定的 B_{mat} 值的概率密度分布，得到期望值 $E(B_{\mathrm{matNi}})$ 和标准偏差 $\sigma(B_{\mathrm{matNi}})$。

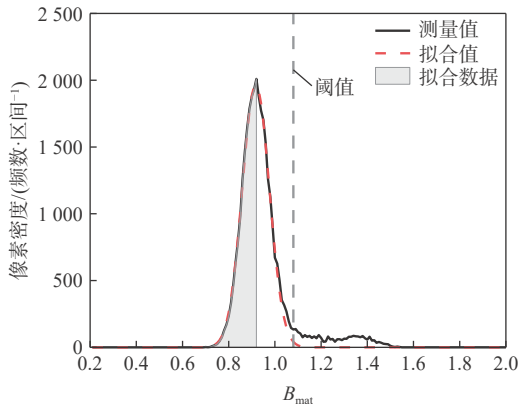


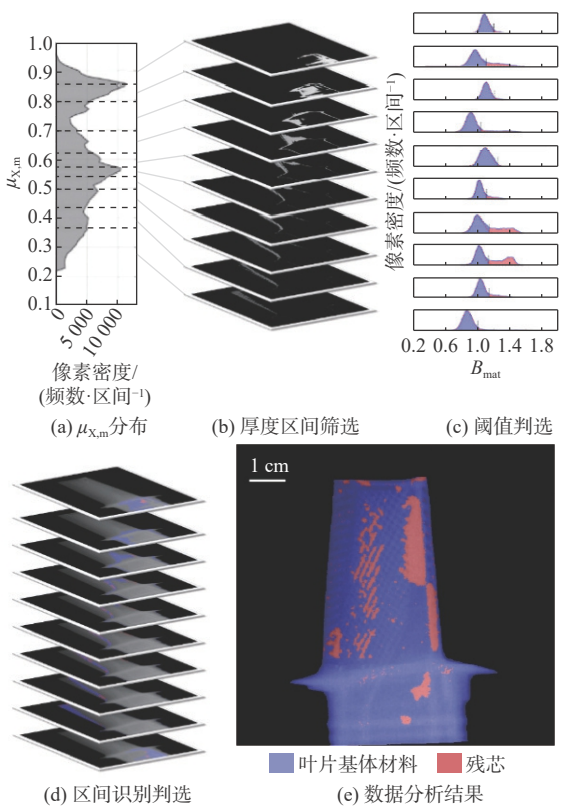
图 12　质量厚度区间 0.43～0.49 g/cm^2 内的测量数据与基体材料拟合分布

Fig. 12　Distribution of measurement data and substrate material fitting in the mass thickness range of 0.43-0.49 g/cm^2

4) 针对步骤 3)所选定的质量厚度区间, 按照当前数据精度, 考虑到当前测量条件下的最佳分辨效果, 选取 2 倍标准偏差的阈值, 设定 $E(B_{matNi})+2\sigma(B_{matNi})$为基体材料与残芯的分辨阈值, 据此, 将当前区间内的像素划分为纯基体材料部分和含有残芯的部分。此时, 误报率≤2.23%。

5) 对每个质量厚度层开展从步骤 2)～步骤 4)的工作, 得到最终的残芯分布图像。

整体的判别流程及结果如图 13 所示。通过对每个质量厚度区间的分析, 最终得到了叶片中残芯的位置分布。



(a) $\mu_{X,m}$分布　(b) 厚度区间筛选　(c) 阈值判选

(d) 区间识别判选　(e) 数据分析结果

图 13　不同质量厚度叶片分层数据的处理及分析结果

Fig. 13　Data processing and analysis results for layers of blades with various mass thickness

4　结　论

1) 对中子/光子双模成像方法在航空发动机空心涡轮叶片残芯检测中的应用开展了研究, 提出并实现了针对残芯的自动检出方法。为基于电子加速器的航空发动机空心涡轮叶片残芯检测提供了一个可行的技术方案。

2) 双模成像方法在融合中子与 X 射线衰减信息并得到反映材料物质属性的 B_{mat} 值后, 根据由 X 射线衰减分析得到的材料质量厚度信息自动确定了物质的分辨阈值, 在叶片的几何结构未知的情况下, 仍可以免人工干预地准确分辨基体材料和残芯。

3) 当单个像素的中子计数不小于 10^4 时, 双模成像方法对残芯识别的误差不超过 1%。

致谢　感谢清华大学工程物理系核技术及应用研究所与北京固鸿科技有限公司对本文工作的支持。

参考文献（References）

[1] BARTON J P, FARNY G, PERSON J L, et al. Neutron radiography: proceedings of the second world conference Paris, France, June16-20, 1986[M]. Berlin: Springer, 1986: 16-20.

[2] 李彪, 祝强, 金磊, 等. 空心涡轮叶片内腔残留型芯的检测技术[C]//2020 中国铸造活动周论文集. 北京: 中国机械工程学会, 2020: 158-166.
LI B, ZHU Q, JIN L, et al. Detection technologies for residual ceramic cores in hollow turbine blades[C]//Proceedings of the 2020 China Foundry Week. Beijing: Chinese Mechanical Engineering Society, 2020: 158-166(in Chinese).

[3] 王婵, 李泽, 罗辅全, 等. 涡轮叶片内腔滞留物射线检测[C]//陕西省第十届无损检测年会论文集. 西安: 陕西省机械工程学会, 2006: 111-121.
WANG C, LI Z, LUO F Q, et al. Turbine blade inner cavity stagnant material radiographic inspection[C]//Proceedings of the 10th Annual NDT Conference in Shaanxi Province. Xi'an: The Shaanxi Provincial Institute of Mechanical Engineering, 2006: 111-121(in Chinese).

[4] 高祥熙, 尹伟, 顾国红, 等. 基于钆掺杂法的空心涡轮叶片残芯冷中子照相检测[J]. 无损检测, 2021, 43(6): 1-5.
GAO X X, YIN W, GU G H, et al. Cold neutron radiography testing of residual core in hollow turbine blade based on gadolinium doping method[J]. Nondestructive Testing Technologying, 2021, 43(6): 1-5(in Chinese).

[5] 王倩妮, 郭广平, 顾国红, 等. 航空发动机叶片残余型芯中子照相检测[J]. 失效分析与预防, 2021, 16(1): 76-82.
WANG Q N, GUO G P, GU G H, et al. Neutron radiography detection of residual core in turbine blades[J]. Failure Analysis and Prevention, 2021, 16(1): 76-82(in Chinese).

[6] GARBE U, AHUJA Y, IBRAHIM R, et al. Industrial application experiments on the neutron imaging instrument DINGO[J]. Physics Procedia, 2017, 88: 13-18.

[7] MARTZ H E, LOGAN C M, SCHNEBERK D J, et al. X-ray imaging: fundamentals, industrial techniques and applications[M]. Raton: CRC Press, 2016.

[8] MAIER A, STEIDL S, CHRISTLEIN V, et al. Medical imaging systems: an introductory guide[M]. Beilin: Springer, 2018.

[9] MCLEAN M. Nickel-base superalloys: current status and potential [J]. Philosophical Transactions of the Royal Society of London Series A: Physical and Engineering Sciences, 1995, 351(1697): 419-433.

[10] YIP S. Nuclear radiation interactions[M]. Singapore: World Scientific Publishing Company, 2014.

[11] BROWN D A, CHADWICK M B, CAPOTE R, et al. ENDF/B-VIII. 0: the 8th major release of the nuclear reaction data library with CIELO-project cross sections, new standards and thermal scattering data[J]. Nuclear Data Sheets, 2018, 148: 1-142.

[12] LEHMANN E H, MANNES D, KAESTNER A P, et al. The XTRA option at the NEUTRA facility: more than 10 years of bi-modal neutron and X-ray imaging at PSI[J]. Applied Sciences, 2021, 11(9): 3825.

[13] SINHA V, SRIVASTAVA A, LEE H K. A novel method for NDT applications using NXCT system at the Missouri University of Science & Technology[J]. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 2014, 750: 43-55.

[14] TENGATTINI A, LENOIR N, ANDÒ E, et al. NeXT-Grenoble, the neutron and X-ray tomograph in Grenoble[J]. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 2020, 968: 163939.

[15] LAMANNA J M, HUSSEY D S, BALTIC E, et al. Neutron and X-ray tomography (NeXT) system for simultaneous, dual modality tomography[J]. Review of Scientific Instruments, 2017, 88(11): 113702.

[16] 杨祎罡, 王学武, 张智, 等. 双模探测方法、控制器和系统: CN110779939B[P]. 2018-07-11.
YANG Y G, WANG X W, ZHANG Z, et al. Bi-modal detection method, controller and system: CN110779939B[P]. 2018-07-11(in Chinese).

[17] YU Y Y, ZHANG R Q, LU L, et al. The bimodal neutron and X-ray imaging driven by a single electron linear accelerator[J]. Applied Sciences, 2021, 11(13): 6050.

[18] BROOKS R A, DI CHIRO G. Beam hardening in X-ray reconstructive tomography[J]. Physics in Medicine and Biology, 1976, 21(3): 390-398.

Bi-modal imaging method for detection of residual core in hollow turbine blades

YU Yangyi[1, 2], YANG Yigang[1, 2, *]

(1. Department of Engineering Physics, Tsinghua University, Beijing 100084, China;
2. Key Laboratory of Particle and Radiation Imaging of Ministry of Education, Tsinghua University, Beijing 100084, China)

Abstract: Residual core inspection is crucial in the production of hollow turbine blades. The e-LINAC-based bi-modal imaging method combines neutron and X-ray imaging along the same optical path, enabling fusion analysis to differentiate between base and core materials based on their properties. This study addresses the beam hardening issue in bi-modal imaging by developing an automatic segmental threshold setting method, which enables a novel automated non-destructive testing approach that does not require prior knowledge of blade geometry or manual intervention, and directly identifies the residual core regions within the blade. When the neutron count per pixel is no less than 10^4, the error in residual core identification using the bi-modal imaging method does not exceed 1%.

Keywords: bi-modal imaging; neutron imaging; non-destructive testing; residual core; hollow turbine blade

Received: 2023-08-17; **Accepted**: 2023-09-14; **Published Online**: 2023-10-10 13:16

URL: link.cnki.net/urlid/11.2625.V.20231009.1047.001

Foundation item: National Natural Science Foundation of China (11735008)

* **Corresponding author**. E-mail: yangyigang@mail.tsinghua.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0550

基于分层深度强化学习的无人机混合路径规划

吕超[1]，李慕宸[2]，欧家骏[3,*]

(1. 北京航空航天大学 航空科学与工程学院，北京 100191；2. 战略支援部队航天系统部装备部装备项目管理中心，北京 100193；
3. 北京航空航天大学 无人系统研究院，北京 100191)

摘 要：无人机应用中需要在非净空环境中实现自主安全飞行，其中障碍物存在部分已知、部分未知的特点，需在基于已知障碍物进行全局路径规划的同时，开展针对未知障碍物的自主避障。为实现障碍物半已知环境下的安全飞行，提出基于分层深度强化学习的混合路径规划方法。该方法将自主感知避障和趋向全局路径规划航迹点飞行2个无人机自主飞行中需要执行的子任务，利用分层模型有机结合，实现功能的有效复合。分层深度强化学习模型中，避障和导航2个子任务模型可分别单独训练，并通过2个训练后的模型对系统状态进行抽象，在此基础上训练顶层模型，实现对2个任务输出的有效调度。试验表明：所提分层深度强化学习方法可在趋向全局路径规划航迹点飞行和自主感知避障的子任务控制器基础上，实现功能扩展，完成更为复杂的任务，可在降低模型训练难度的同时，保持模型任务执行能力。

关 键 词：无人机；路径规划；自主避障；深度强化学习；分层模型
中图分类号：V219
文献标志码：A **文章编号**：1001-5965(2025)10-3451-09

无人机相对于常规有人驾驶飞行器，具备低成本、机动灵活、可损耗等优点，特别适宜于在危险、重复、枯燥的场景中应用，因此，在军民领域发挥出了愈发重要的作用。在无人机技术不断发展的同时，其应用场景不断扩展，无人机需要到城市、森林和山区等存在障碍物的非净空环境中飞行。无人机所面临的飞行环境越来越复杂，其中甚至存在大量未知的障碍物。无人机能否有效识别和利用障碍物信息，规划出合理的飞行路线，避免碰撞坠毁，不仅关系到无人机本身的安全和任务执行成败，更关系到周边人员和设施的安全。

无人机在非净空环境条件下实现安全飞行，需获得飞行路线上可能的障碍物信息，根据障碍物信息做出准确的路径规划。在实际应用环境中，受地图测绘和应用场景开发水平等因素限制，实际应用场景中的障碍物信息在飞行前一般难以全面掌握，障碍物信息部分已知，部分未知。因此，无人机的自主飞行过程包含2方面要素：基于已知障碍物的全局路径规划和针对未知障碍物的自主避障。

针对各类无人平台的路径规划问题，有关学者开展了大量研究工作。文献[1]基于改进的A*算法实现路径规划。文献[2-3]将改进的快速扩展随机树(rapidly-exploring random tree, RRT)方法应用于无人机和机器人领域，实现路径规划。文献[4]将同步定位与建图(simultaneous localization and mapping, SLAM)方法和人工势场法相结合，提出了一种室内机器人自主导航方法。文献[5]基于SLAM方法实现无人平台的自主感知与路径规划。文献[6-7]使用深度神经网络，学习路径规划的方法，实现对近似最优路径的获取。文献[8-9]基于深度强化学习方法，解决全局路径规划问题。

在自主避障方面，文献[10]根据传感器数据，

收稿日期：2023-08-28；录用日期：2023-10-20；网络出版时间：2023-11-01 14:10
网络出版地址：link.cnki.net/urlid/11.2625.V.20231101.1247.003
*通信作者. E-mail：oujiajun@buaa.edu.cn

引用格式：吕超，李慕宸，欧家骏. 基于分层深度强化学习的无人机混合路径规划[J]. 北京航空航天大学学报，2025，51(10)：3451-3459.
LYU C，LI M C，OU J J. UAV hybrid path planning based on hierarchical deep reinforcement learning[J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51(10)：3451-3459 (in Chinese).

从第一人称视角构建地图，并规划机器人的动作序列实现避障。文献[11]为实现避障能力，提出了一种深度对抗Q神经网络，在从RGB图像生成的深度信息存在大量噪声的前提下，该网络机制可以有效学习如何避障。文献[12]采用分层架构，利用强化学习训练模型获取避障策略，控制无人机动作实现自主飞行。文献[13]提出了一个基于演员-评论家模型的机器人避障导航模型。文献[14]利用全卷积残差网络对障碍物进行识别，得到深度图像，并采用深度强化方法规划避障路径，同时构建环境的二维地图。

在前期研究中，有关学者一般将无人机自主飞行分解为路径规划和自主感知避障2个单独问题进行研究。本文为实现无人机在半已知环境下的安全飞行，提出基于分层深度强化学习的无人机混合路径规划方法，兼顾全局路径规划和自主感知避障，实现功能复合。

1 无人机混合路径规划问题描述

本文研究的无人机混合路径规划问题中，设定无人机飞行任务目标点已经预先获得，将飞行任务分解为2个子任务，分别为基于趋向航迹规划点飞行子任务和基于传感器数据进行自主感知避障子任务，2个子任务同时运行，并添加顶层控制器对子任务执行进行控制，完成无人机安全飞行的总任务目标，如图1所示。

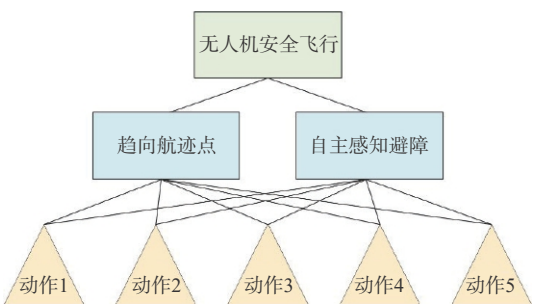


图1 无人机混合路径规划架构

Fig. 1 UAV hybrid path planning architecture

将无人机飞行任务定义为总任务M，并将其分解为子任务集合$\{M_0,M_1,M_2\}$。其中，M_0定义为根任务，其处在任务分层的顶层，即无人机在存在障碍物环境下的安全飞行任务。M_1和M_2为子任务，分别为无人机向全局路径规划航迹点飞行任务和无人机根据传感器数据自主感知避障任务，以元组$\langle T_i,A_i,R_i\rangle$定义子任务，具体定义如下：

1）T_i为子任务M_i的终止状态，对某一特定子任务而言，终止状态和活动状态S_i共同构成了系统所有的状态集合$\mathcal{S}$，当系统当前状态$s\in S_i$时，子任务M_i的策略π_i执行，当系统状态进入T_i时，子任务M_i也终止。

2）A_i为子任务M_i的动作集合，不同子任务间的动作集合可以完全不同、部分相同甚至完全相同，M_i的策略π_i在A_i中选取系统动作，用以更新系统状态。

3）R_i为子任务的回报函数，表示子任务M_i执行动作$a\in A_i$，子任务系统所获得的回报。由于子任务系统与整体系统存在差异，子任务回报与根任务回报不同。该回报用于子任务策略的学习过程，根任务的回报可与子任务回报关联，也可完全独立。

2 混合路径规划模型设计

2.1 设计框架

利用分层深度强化学习，针对无人机在半已知环境下的安全飞行任务建立分层模型，模型设计框架如图2所示。无人机的自主安全飞行作为顶层任务，被划分为趋向全局路径规划航迹点飞行和根据传感器信息自主感知避障2个子任务。2个子任务学习得到的策略分别可以控制无人机选取不同的飞行动作进行执行，进而实现对无人机飞行轨迹的调整。无人机安全飞行任务作为顶层任务，其策略控制2个子任务策略的生效与否，实现无人机在趋向航迹点飞行和开展自主感知避障2个子任务之间的实时调度。

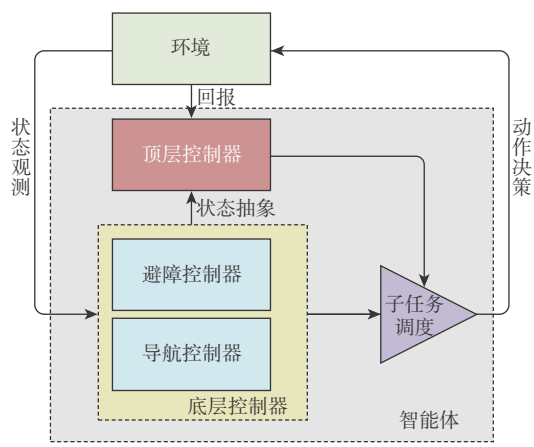


图2 分层深度强化学习模型设计框架

Fig. 2 The design framework of the hierarchical deep reinforcement learning model

无人机分层深度强化学习模型中包含趋向航迹点和自主感知避障2个子任务的底层控制器，2个底层控制器分别设计、建模、训练和并行执行。对任意一个子任务M_i而言，其决策过程可以看做马尔可夫决策过程，其模型可由一个四元组$\langle S_i,A_i,P_i^{\pi_i},R_i\rangle$表示。其中，$P_i^{\pi}$表示在状态$S_i$下执行动作$a\in A_i$后，环境转移到下一个状态$S_{i+1}$的概率。系统在子任务内运行，对动作空间内各个动作

$a \in A_i$进行评价，获得每个动作对应的Q值（在状态S_i下采取动作$a \in A_i$所能获得的长期预期累积奖励的估计）。根任务M_0决策过程为半马尔可夫决策过程，其运行根据状态输入，评估当前状态是否为某一子任务的终止状态T_i，进而决策各子任务平稳状态持续时间。

2.2　控制器设计

混合路径规划控制器模型基于分层深度强化学习模型设计，包含 2 个底层控制器，分别为避障控制器和导航控制器，对应于自主感知避障和趋向航迹点飞行 2 个子任务的执行，2 个子任务共享相同的无人机动作空间。另有一个顶层控制器，负责 2 个子任务的调度，实现无人机的安全飞行。分层模型中，避障控制器和导航控制器将系统状态信息映射到动作空间中，分别为$Q_1 = \{Q'_1, Q'_2, \cdots, Q'_5\}$和$Q_2 = \{Q''_1, Q''_2, \cdots, Q''_5\}$，以此实现子任务抽象，顶层控制器基于抽象后的信息，得到最终的$Q = \{Q_1, Q_2, \cdots, Q_5\}$，进而决策无人机最终飞行动作，如图 3 所示。

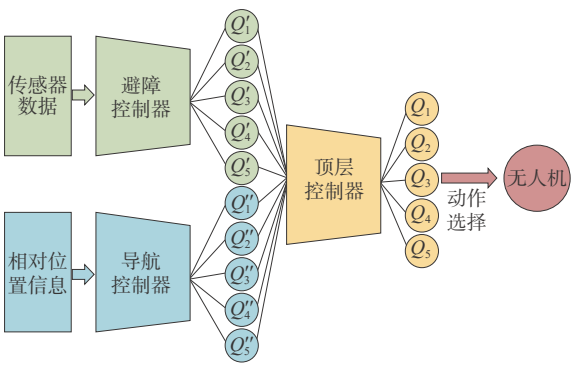


图 3　分层控制器示意图

Fig. 3　The diagram of hierarchical controller

2.3　神经网络结构

控制器神经网络结构基于深度竞争循环双 Q 神经网络（deep dueling recurrent double Q-network, D3RQN）设计，其融合了深度循环 Q 神经网络（deep recurrent Q-network, DRQN）[15]、深度竞争 Q 神经网络（dueling deep Q-network, DuelingDQN）[16] 和深度双 Q 神经网络（double deep Q-network, DoubleDQN）[17] 的优势。D3RQN 结构如图 4 所示。具体网络结构在 DuelingDQN 的基础上，将最后一层全连接（fully connected, FC）层替换为长短期记忆（long short-term memory, LSTM）层。

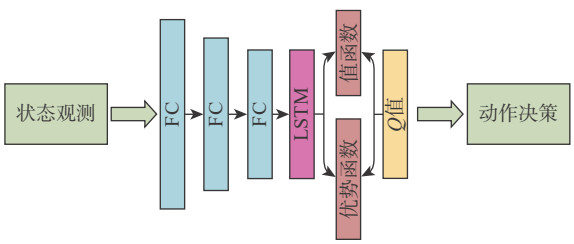


图 4　深度竞争循环双 Q 神经网络示意图

Fig. 4　The diagram of the D3RQN

在 D3RQN 中，分别计算优势函数和值函数，最后共同构成对 Q 值的估计，可表示为

$$Q^{\pi}(s,a) = A^{\pi}(s,a) + V^{\pi}(s) \tag{1}$$

式中：π为控制策略；$A(s,a)$ 为优势函数；$V(s)$ 为价值函数；s和a分别为状态和动作。

针对部分可观马尔可夫决策过程，系统状态s难以得到，仅能获得对系统状态s的观测$\boldsymbol{o}$，因此，以观测的集合来近似系统状态，可表示为

$$\Omega_t = \{\boldsymbol{o}_0, \boldsymbol{o}_1, \boldsymbol{o}_2, \cdots, \boldsymbol{o}_t\} \tag{2}$$

式中：Ω_t为t时刻前的观测集合；$\boldsymbol{o}_t$为t时刻时的观测。同时，结合 D3RQN 中 LSTM 层可以处理序列数据的特点，t时刻的 Q 值可表示为

$$Q^{\pi}(\Omega_t, a_t) = A^{\pi}(\Omega_t, a_t) + V^{\pi}(\Omega_t) \tag{3}$$

D3RQN 采取与 DoubleDQN 相似的学习策略，分别建立当前网络和目标网络。基于该设计，D3RQN 下损失函数公式设计如下：

$$L_t(\theta) = E_{\Omega_t, a_t, r_t, \Omega_{t+1}}\left[\left(y_t^{\mathrm{D3RQN}} - Q(\Omega_t, a_t; \theta)\right)^2\right] \tag{4}$$

$$y_t^{\mathrm{D3RQN}} = r_t + \gamma \max_{a_{t+1}} Q\left(\Omega_{t+1}, a_{t+1}; \theta^{-}\right) \tag{5}$$

式中：E为期望值；r_t为 t 时刻获得的即时奖励；γ 为折扣因子；θ和θ^{-}分别为当前网络和目标网络参数。神经网络模型可以通过最小化损失函数来学习值函数估计，并基于学习得到的策略进行动作决策，如下：

$$a_t = \pi(\Omega_t) = \arg\max_{a} Q(\Omega_t, a) \tag{6}$$

3　混合路径规划决策

3.1　状态空间与动作空间

基于分层深度强化学习的无人机混合路径规划，针对二维平面内开展路径规划，设计仿真环境，环境场景尺度为 640 m×480 m。无人机机载传感器数据、无人机位姿数据和目标航迹点坐标共同构成系统对环境的观测。无人机机载传感器选取前侧视激光雷达，其可直接获取特定方向上无人机与障碍物间的距离。激光雷达与无人机机体固连，采用 32 线束激光雷达，视场角为无人机机体坐标系前向±45°，激光雷达最大探测距离为 200 m，激光雷达示意图如图 5 所示。

无人机在飞行过程中保持恒定速度向前飞行，通过分层深度强化学习模型控制无人机飞行航向，使无人机在飞行过程中实现避障并飞抵目标位置。无人机动作空间定义如表 1 所示，其中，线速

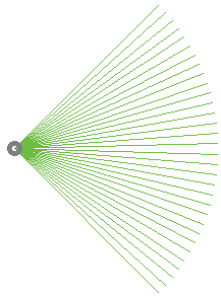


图5 激光雷达示意图

Fig. 5 The diagram of the lidar

表1 无人机动作空间定义

Table 1 The definition of the UAV action space

动作编号	线速度/(m·s^{-1})	航向控制量/(°)
1	(10, 0, 0)	0
2	(10, 0, 0)	+5
3	(10, 0, 0)	−5
4	(10, 0, 0)	+10
5	(10, 0, 0)	−10

度单位定义为单位时间内无人机在地图图像上移动的图像像素距离。

3.2 控制器参数

3.2.1 避障控制器参数

避障控制器以32线束激光雷达数据为输入，对应避障控制器网络模型第一层输入尺度为32。同时考虑激光雷达最远探测距离为200 m，对传感器数据进行归一化，使控制器输入在0～1之间。避障控制器网络模型主要参数如表2所示。

表2 避障控制器网络模型主要参数

Table 2 The main parameters of obstacle avoidance controller network model

层号	类别	输入尺度	输出尺度	激活函数
1	FC	32	128	ReLU
2	FC	128	128	ReLU
3	FC	128	128	ReLU
4	LSTM	128	128	
5($A(s,a)$)	FC	128	5	ReLU
5($V(s)$)	FC	128	1	ReLU

3.2.2 导航控制器参数

导航控制器以无人机与目标点的相对位置作为输入，相对位置包含东向(x轴向右)和北向(y轴向上)2个分量。考虑到仿真环境的场景最大尺度，将无人机与目标点的相对位置参数除以500做近似归一化。导航控制器网络模型主要参数如表3所示。

表3 导航控制器网络模型主要参数

Table 3 The main parameters of navigation controller network model

层号	类别	输入尺度	输出尺度	激活函数
1	FC	2	32	ReLU
2	FC	32	32	ReLU
3	FC	32	32	ReLU
4	LSTM	32	32	
5(A(s,a))	FC	32	5	ReLU
5(V(s))	FC	32	1	ReLU

3.2.3 顶层控制器参数

顶层控制器以2个底层控制器对各动作选项的Q值估计作为输入，将其重新映射到无人机的动作空间上，基于2个底层控制器子任务输出，对无人机最终动作进行宏观决策。顶层控制器网络模型主要参数如表4所示。

表4 顶层控制器网络模型主要参数

Table 4 The main parameters of top controller network model

层号	类别	输入尺度	输出尺度	激活函数
1	FC	10	128	ReLU
2	FC	128	128	ReLU
3	FC	128	128	ReLU
4	LSTM	128	128	
5($A(s,a)$)	FC	128	5	ReLU
5($V(s)$)	FC	128	1	ReLU

3.3 回报函数设计

3.3.1 避障回报函数

在无人机自主感知避障子任务中，系统所获得的回报定义如下：

$$R_{\mathrm{OA}}=\begin{cases} d_{\mathrm{nst}} & \text{安全} \\ -10 & \text{碰撞} \end{cases} \tag{7}$$

式中：d_{nst}为无人机到最近障碍物的距离，当无人机与最近障碍物的距离大于20 m时，认为无人机处在安全状态，当小于安全距离时，考虑无人机发生碰撞。

3.3.2 导航回报函数

在趋向航迹点飞行子任务中，无人机位置随着每个动作执行而更新，同时获得相应的回报，回报函数被定义为

$$R_{\mathrm{Nav}}=\begin{cases} 10 & \text{到达} \\ K(d_{\mathrm{curr}}-d_{\mathrm{next}}-d_{\mathrm{step}}) & \text{其他} \end{cases} \tag{8}$$

式中：d_{curr}为无人机当前位置与目标位置间的距离；d_{next}为无人机下一步位置与目标位置之间的距离；d_{step}为无人机动作步长；K为常数，取值为0.1。根据回报函数的设计，无人机在达到目标位置前，所

获得的回报总为负值, 以避免进行无意义的徘徊。

3.3.3　顶层控制回报函数

顶层控制器混合了自主感知避障和趋向目标点飞行 2 种策略, 无人机随着飞行动作的执行, 获得的回报由自主感知避障和趋向目标点飞行 2 种策略同时组成, 定义为

$$R_{\text{Hybrid}} = R_{\text{OA}} + R_{\text{Nav}} \tag{9}$$

3.4　混合路径规划方法

基于分层深度强化学习的无人机混合路径规划方法实现过程如下:

输入: 避障控制器模型、导航控制器模型和顶层控制器模型; 当前无人机位置 $\boldsymbol{L}_{\text{c}}$ 和目标位置 $\boldsymbol{L}_{\text{g}}$。

重设初始状体: 路径 p=$\{\boldsymbol{L}_{\text{c}}\}$。

While $\boldsymbol{L}_{\text{c}}$≠$\boldsymbol{L}_{\text{g}}$

do

获得当前观测 Ω_t

避障控制器根据当前观测 Ω_t 获得 Q_1 值

导航控制器根据无人机位置 $\boldsymbol{L}_{\text{c}}$ 和目标位置 $\boldsymbol{L}_{\text{g}}$ 获得 Q_2 值

顶层控制器根据 Q_1 和 Q_2 值选取最优动作 a

更新当前位置 $\boldsymbol{L}_{\text{c}}$, 并将其添加到路径中

End while

输出: 路径 p。

该方法以无人机的实时位置、观测数据及目标位置为输入。底层的避障控制器根据观测数据生成作为顶层控制器输入的 Q_1 值; 导航控制器根据无人机位置和目标位置同样生成作为顶层控制器输入的 Q_2 值; 顶层控制器根据 Q_1、Q_2 值选择最优动作, 无人机执行该动作后到达下一位置。如此往复, 直至到达目标点。

4　仿真验证

无人机混合路径规划仿真场景内可随机生成若干障碍物, 障碍物数量、位置、尺寸可定义, 无人机初始位姿和目标航迹点位置可随机生成。

4.1　训练设置

混合路径规划模型采用分步训练策略, 首先训练 2 个底层控制器模型, 分别实现自主感知避障和趋向目标点飞行功能, 然后在底层模型的基础上训练顶层模型, 实现混合路径规划目标。为实现基于分层深度强化学习的混合路径规划模型训练, 针对自主感知避障、趋向目标点飞行和顶层混合策略 3 项内容训练分别设置了 3 个训练场景。3 个场景分别对应训练模型避障、导航和顶层控制器能力, 控制器训练参数如表 5 所示。

1) 场景 1: 仿真训练场景要素包括无人机及机载传感器和障碍物, 无人机在场景中自由飞行, 要求在飞行过程中避开障碍物, 并尽可能延长无碰撞飞行时间, 旨在训练底层避障控制器性能。

2) 场景 2: 仿真场景中无障碍物, 仅有无人机和目标点, 无人机位姿和目标点初始位置在训练过程中不断随机生成, 且位姿信息可被控制器获取, 以此训练导航控制器, 实现引导无人机抵达目标点的能力。

3) 场景 3: 针对混合路径规划顶层控制器, 仿真场景中包含无障碍物、无人机和目标点, 训练过程中障碍物固定不变, 无人机位姿和目标点初始位置随机生成。利用在场景 1 和场景 2 中训练得到的子任务底层控制器, 将系统状态信息抽象为对动作的 2 套 Q 估计值, 将底层控制器进行抽象后的 Q 估计值作为顶层控制器的输入数据, 训练顶层控制器对底层子任务的协同能力。

4.2　训练及测试结果

4.2.1　自主感知避障

避障控制器训练曲线如图 6 所示。为了测试避障控制器的性能, 采用障碍物数量、布置方式和尺寸与训练场景不同的测试场景, 对避障控制器能力进行了测试, 测试结果如图 7 所示, 以红色细线显示无人机避障飞行轨迹。可见, 无人机飞行轨迹可自主绕开障碍物飞行。开展 1 000 次避障飞行测试, 设定无人机安全飞行 200 个时间步不发生碰撞, 即认为单次飞行测试成功。最终结果为: 无人机避障飞行成功率为 96.8%。

4.2.2　导航控制

导航控制器训练曲线如图 8 所示。为了测试导航控制器引导无人机趋向目标点飞行的能力, 测试过程中随机生成无人机位姿和目标点位置, 对导航控制器能力进行了测试, 测试结果如图 9 所示。可见, 无人机可快速飞行目标点。开展 1 000 次导航飞行测试, 设定当无人机在 200 步动作内飞抵目

表 5　控制器训练参数

Table 5　The parameters of controller training

场景	地图尺度/(m×m)	障碍物数量	训练批量	折扣因子	学习率	输入数据维度	动作时间间隔/s	网络更新频率/步	优化器	优化器参数
场景1	640×480	4	64	0.99	1×10^{-4}	32	0.5	400	Adam	[0.9,0.999,0]
场景2	640×480	0	64	0.99	1×10^{-4}	2	0.5	400	Adam	[0.9,0.999,0]
场景3	640×480	4	32	0.99	1×10^{-4}	34	0.5	400	Adam	[0.9,0.999,0]

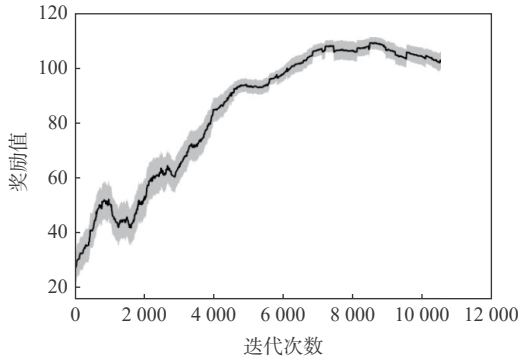


图6 避障控制器训练曲线

Fig. 6 Training curve of obstacle avoidance controller

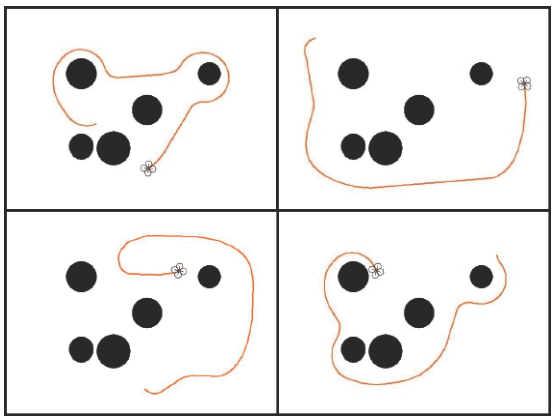


图7 自主感知避障测试结果

Fig. 7 Autonomous perception obstacle avoidance test results

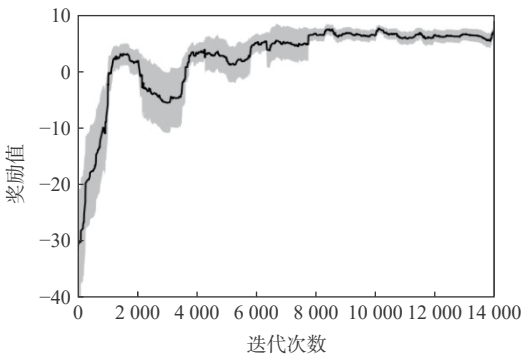


图8 导航控制器训练曲线

Fig. 8 Training curve of navigation controller

标点而不飞出地图边界时，即认为单次飞行测试成功。最终结果为：无人机导航飞行成功率为99.9%。

4.2.3 顶层控制

顶层控制器训练曲线如图10所示。为了测试顶层控制器的性能，采用障碍物数量、布置方式和尺寸与训练场景不同的测试场景，同时随机生成无人机位姿和目标点位置，同步测试无人机避障及飞行目标点的能力，测试结果如图11所示。可见，无人机在避开障碍物飞行的同时，完成路径规划达到目标点。开展1 000次飞行测试，设定当无人机在200步动作内飞抵目标点而不飞出地图边界或与障

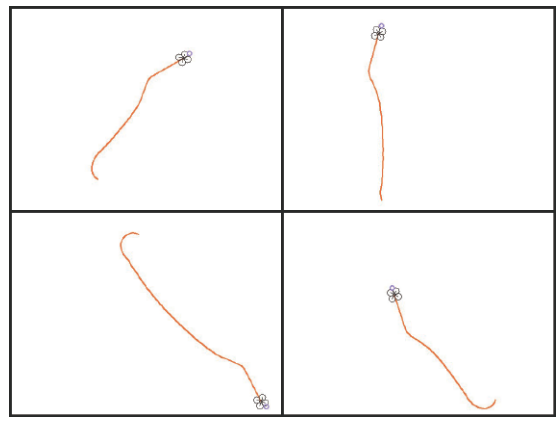


图9 导航控制测试结果

Fig. 9 Navigation test result

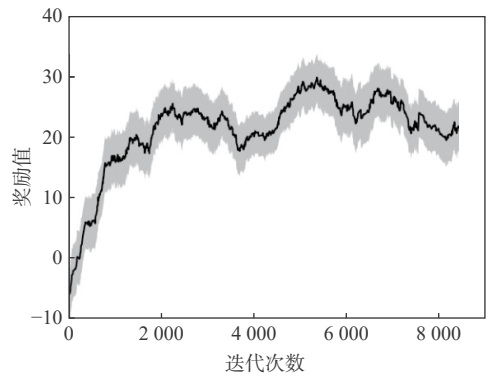


图10 顶层控制器训练曲线

Fig. 10 Training curve of top controller

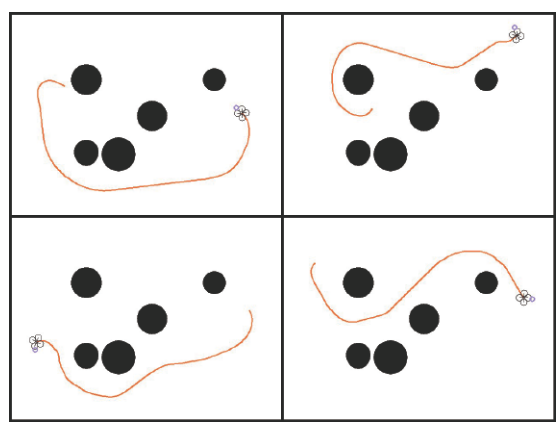


图11 顶层控制测试结果

Fig. 11 Top controller test results

碍物发生碰撞时，即认为单次飞行测试成功。最终结果为：无人机飞行成功率为94.7%。

4.3 结果分析

试验结果显示，底层控制器网络模型可在仿真环境中独立完成自主感知避障和趋向航迹点飞行任务。利用底层控制器网络模型对系统状态的抽象信息，顶层控制器可调度2个底层模型，最终实现整个分层模型任务目标。

在路径规划过程中分析各控制器决策，2个底

层控制器根据场景情况，分别执行底层控制决策，顶层控制器根据具体情况，在 2 个底层控制器决策结果冲突时，做出独立判断，并实现最终任务目标，如图 12 所示。在靠近目标点飞行和避开障碍物飞行 2 种机制下，2 个底层控制器的输出显然不同，甚至可能截然相反。蓝色曲线代表顶层控制器的动作输出，其综合 2 个底层控制器的输出，复合后作出决策。顶层控制器在某些时刻可能选择与其中 1 个底层控制器输出保持一致，也可能与 2 个底层控制器输出完全相反。

此外，为了进一步证明本文方法的有效性，将其与 RRT 方法进行对比。如图 13 所示，共设置 2 种场景，2 种方法均可实现从绿色起点到红色目标点的全局路径规划，在过程中躲避黑色障碍物。与基于 RRT 的方法相比，本文方法执行效率更高，不需要大量搜索与试错，适应不同环境能力较强，且规划的路径更平滑。

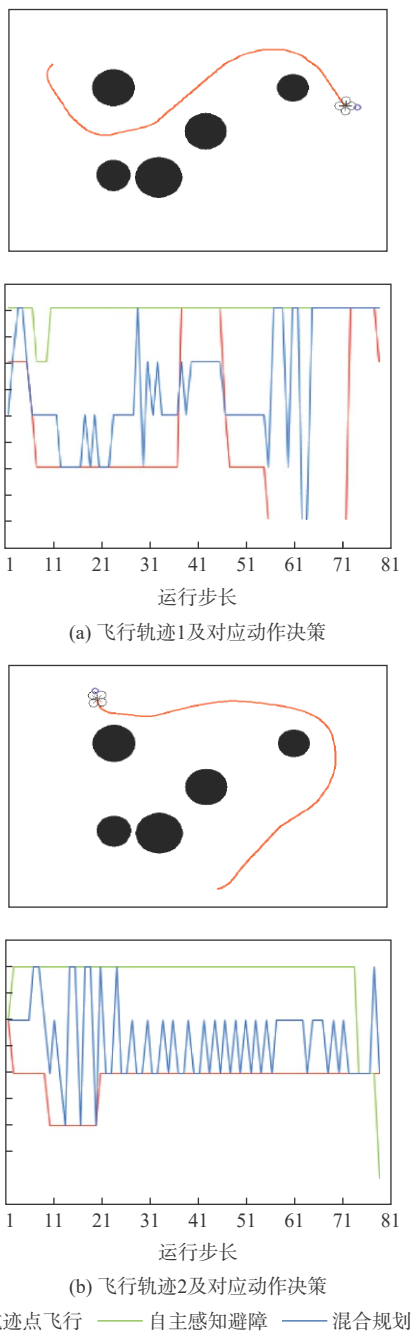


图 12　飞行轨迹及各控制器动作决策

Fig. 12　The flight path and action decision of each controller

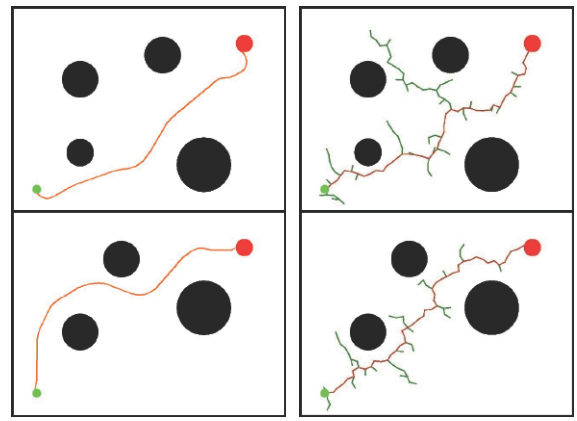


图 13　混合路径规划方法与 RRT 规划方法结果对比

Fig. 13　Comparison of results of hybrid path planning algorithm and RRT planning algorithm

分层模型体现了以下特点：①状态抽象。通过底层控制器对系统状态进行抽象，顶层控制器根据对系统抽象的结果进行决策，系统当前状态下与所解决问题无关的状态将被忽略。②迁移能力。分层深度强化学习模型本质上将问题进行分解，将任务分解为若干个小问题，并分而治之，先构建若干子任务功能模块，并复合这些模块实现最终目标。模块之间可根据任务目标进行灵活组合，或单独运行发挥作用。③策略灵活。普通深度强化学习模型，在每个决策点均需要根据策略进行决策，而分层深度强化学习模型可根据场景和任务，实现在不同策略间的自动切换，如自主感知避障和飞向目标点的策略切换，进而形成一个可根据场景灵活变化的策略。

5　结　论

本文围绕无人机在存在障碍物的非净空环境下的安全飞行问题，提出了一种基于分层深度强化学习的混合路径规划方法，使无人机在飞向目标位置的同时可自主避开障碍物，主要结论如下：

1）提出了基于分层深度强化学习的混合路径规划方法，其将自主感知避障和趋向全局路径规划航迹点飞行 2 个无人机自主飞行中需要执行的子任务，利用分层模型有机结合，实现功能的有效复合。

2）提出的分层深度强化学习模型中，避障和导航 2 个子任务控制器策略可分别单独训练，并通过 2 个训练后的控制器模型对系统状态进行抽象，在

此基础上训练顶层控制器，实现对2个任务输出的有效调度。

3）本文提出的分层深度强化学习模型可在子任务控制器基础上，实现功能扩展，完成更为复杂的任务，进而在降低模型训练难度的同时，保持模型任务执行能力。

在进一步的研究工作中，可以将更多的无人机功能扩展到分层模型中，实现更为复杂的无人机功能。

参考文献（References）

[1] 支琛博, 张爱军, 杜新阳, 等. 改进A*算法的移动机器人全局路径规划研究[J]. 计算机仿真, 2023, 40(2): 486-491.
ZHI C B, ZHANG A J, DU X Y, et al. Research on global path planning of mobile robot based on improved A* algorithm[J]. Computer Simulation, 2023, 40(2): 486-491(in Chinese).

[2] 栾添添, 王皓, 孙明晓, 等. 基于动态变采样区域RRT的无人车路径规划[J]. 控制与决策, 2023, 38(6): 1721-1729.
LUAN T T, WANG H, SUN M X, et al. Path planning of unmanned vehicle based on dynamic variable sampling area RRT[J]. Control and Decision, 2023, 38(6): 1721-1729(in Chinese).

[3] 李樾, 韩维, 陈清阳, 等. 基于快速扩展随机树算法的多无人机编队重构方法研究[J]. 西北工业大学学报, 2019, 37(3): 601-611.
LI Y, HAN W, CHEN Q Y, et al. Research on formation reconfiguration of UAVs based on RRT algorithm[J]. Journal of Northwestern Polytechnical University, 2019, 37(3): 601-611(in Chinese).

[4] HARIK E H, KORSAETH A. Combining hector SLAM and artificial potential field for autonomous navigation inside a greenhouse[J]. Robotics, 2018, 7(2): 22.

[5] RHODES C, LIU C J, CHEN W H. Autonomous source term estimation in unknown environments: from a dual control concept to UAV deployment[J]. IEEE Robotics and Automation Letters, 2022, 7(2): 2274-2281.

[6] YU X Q, WANG P, ZHANG Z X. Learning-based end-to-end path planning for lunar rovers with safety constraints[J]. Sensors, 2021, 21(3): 796.

[7] WU K Y, ESFAHANI M A, YUAN S H, et al. TDPP-Net: achieving three-dimensional path planning via a deep neural network architecture[J]. Neurocomputing, 2019, 357: 151-162.

[8] 黄昱洲, 王立松, 秦小麟. 一种基于深度强化学习的无人小车双层路径规划方法[J]. 计算机科学, 2023, 50(1): 194-204.
HUANG Y Z, WANG L S, QIN X L. Bi-level path planning method for unmanned vehicle based on deep reinforcement learning[J]. Computer Science, 2023, 50(1): 194-204(in Chinese).

[9] 封硕, 舒红, 谢步庆. 基于改进深度强化学习的三维环境路径规划[J]. 计算机应用与软件, 2021, 38(1): 250-255.
FENG S, SHU H, XIE B Q. 3D environment path planning based on improved deep reinforcement learning[J]. Computer Applications and Software, 2021, 38(1): 250-255(in Chinese).

[10] GUPTA S, TOLANI V, DAVIDSON J, et al. Cognitive mapping and planning for visual navigation[J]. International Journal of Computer Vision, 2020, 128(5): 1311-1330.

[11] XIE L H, WANG S, MARKHAM A, et al. Towards monocular vision based obstacle avoidance through deep reinforcement learning[EB/OL]. (2017-06-29) [2023-08-20]. https://arxiv.org/abs/1706.09829.

[12] SINGLA A, PADAKANDLA S, BHATNAGAR S. Memory-based deep reinforcement learning for obstacle avoidance in UAV with limited environment knowledge[J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 22(1): 107-118.

[13] ZHU Y K, MOTTAGHI R, KOLVE E, et al. Target-driven visual navigation in indoor scenes using deep reinforcement learning[C]// Proceedings of the IEEE International Conference on Robotics and Automation. Piscataway: IEEE Press, 2017: 3357-3364.

[14] WEN S H, ZHAO Y F, YUAN X, et al. Path planning for active SLAM based on deep reinforcement learning under unknown environments[J]. Intelligent Service Robotics, 2020, 13(2): 263-272.

[15] HAUSKNECHT M, STONE P. Deep recurrent Q-learning for partially observable MDPS[C]//Proceedings of the 2015 International Conference on Learning Representations. New York: ACM, 2015: 29-37.

[16] WANG Z Y, SCHAUL T, HESSEL M, et al. Dueling network architectures for deep reinforcement learning[C]//Proceedings of the 33rd International Conference on Machine Learning. New York: ACM, 2016, 48: 1995-2003.

[17] VAN H H, GUEZ A, SILVER D. Deep reinforcement learning with double Q-learning[C]//Proceedings of the AAAI Conference on Artificial Intelligence. New York: AAAI Press, 2016, 30(1): 2094-2100.

UAV hybrid path planning based on hierarchical deep reinforcement learning

LYU Chao[1]，LI Muchen[2]，OU Jiajun[3,*]

(1. School of Aeronautic Science and Engineering，Beihang University，Beijing 100191，China；
2. Equipment Project Management Center，Aerospace System Department of Strategic Support Force Equipment Department，Beijing 100193，China；
3. Institute of Unmanned System，Beihang University，Beijing 100191，China)

Abstract：In the application of UAV, it is necessary to realize autonomous and safe flight in a non-clearance environment. For both known and unknown obstacles, it is necessary to carry out autonomous obstacle avoidance for unknown obstacles while making global path planning based on known obstacles. To achieve safe flight in the semi-known obstacle environment, a hybrid path planning method based on hierarchical deep reinforcement learning is proposed, which naturally combines autonomous perception, obstacle avoidance, and global path planning, two sub-tasks that UAV must perform in autonomous flight, to achieve an efficient combination of functions. In the hierarchical deep reinforcement learning model, the obstacle avoidance and navigation sub-task models can be trained separately, and the system state can be abstracted through the two trained models. On this basis, the top-level model can be trained to achieve effective scheduling of the two sub-task outputs. The obstacle avoidance and navigation sub-task models can be trained independently in the hierarchical deep reinforcement learning model, and the system state can be abstracted through the two trained models. It can reduce the difficulty of model training while maintaining the ability to execute the model task.

Keywords：UAV；path planning；autonomous obstacle avoidance；deep reinforcement learning；hierarchical model

Received：2023-08-28；**Accepted**：2023-10-20；**Published Online**：2023-11-01 14：10
URL：link.cnki.net/urlid/11.2625.V.20231101.1247.003

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0531

无人机数据采集任务中的航迹规划与资源分配优化

雷耀麟[1,2]，丁文锐[3]，罗祎喆[4,*]，王玉峰[3]，刘思琪[5]，张芷兰[1]

(1. 北京航空航天大学 电子信息工程学院，北京 100191；2. 中国电子科技集团第五十四研究所，石家庄 050081；
3. 北京航空航天大学 无人系统研究院，北京 100191；4. 郑州大学 计算机与人工智能学院，郑州 450053；
5. 国网河南省电力公司新乡供电公司，新乡 453000)

摘 要：针对无人机（UAV）在应急场景中执行数据采集任务时，其电池容量有限、缓存空间有限，以及地面目标优先级动态变化所导致的无人机航迹规划与资源分配效果较差的问题，提出一种基于深度强化学习的无人机航迹规划与资源分配联合优化方法。考虑无人机任务中的通信、计算、飞行、数据缓存过程，构建相应的数学问题模型；针对无人机航迹规划与资源分配问题构建马尔可夫过程模型，设计相应的状态和行为描述及用于平衡无人机能耗和数据采集信息量的加权奖励函数；与贪婪算法和遗传算法等智能优化方法进行仿真对比。结果表明：所提方法能够使无人机在较短的任务时间内，以消耗相似或者较低的能量为代价，较大提升对地面用户的数据采集量。

关 键 词：无人机；资源分配；航迹规划；强化学习；移动边缘计算
中图分类号：TN929.5
文献标志码：A **文章编号**：1001-5965(2025)10-3460-11

无人机(unmanned aerial vehicle, UAV)作为不需要人类飞行员操控的轻量级飞机，具备低成本、鲁棒性强、可商购[1]等优点。得益于其高机动性与高空通信优势，无人机在数据采集任务中的应用越来越广泛。例如，在电力巡检过程中[2]，通过规划无人机资源，并对无人机任务进行合理规划，可有效减少电力物联网的业务延迟和能耗。此外，无人机数据采集可用于无线传感器网络中的广域数据采集[3]，在传统方式中，每个传感器节点通过多跳传输将其监测到的数据传送到融合中心，其间，不仅需要传输自己的数据，还需要中继其他节点的数据，因此，传感器的电池可能会迅速耗尽，导致多跳网络失联。而通过将无人机作为移动融合中心，每个传感器节点都可以直接将观测数据发送给无人机，极大提升了无线传感网络数据采集的效率，具有较高的研究价值[4]。

无人机航迹规划与资源分配对数据采集有着至关重要的影响，一般优化的目标为最低能耗、最短飞行时间、最低时延等。无人机航迹规划常常与资源分配组合，从而最大化资源利用率，构成的组合优化问题的决策空间很大，因此，优化复杂度很高。无人机航迹规划与资源分配优化方法可以划分为基于图论的方法、基于优化方法的方法、基于人工智能的方法。其中，基于图论的航迹规划方法包括沃罗诺伊图算法、概率路线图算法、希尔伯特曲线算法等，该类方法将无人机执行任务所在的地理空间转换为图并搜索源和目的地之间的路径，优化目标一般为最大采集数据量、最小数据采集时间、最小能耗等[5-7]。基于优化方法的方法包括动态规划、分支定界和连续凸逼近算法，此类方法能够找到问题的最优解或次优解，但对于一些所有多项式复杂程度的非确定性问题都能在多项式时间复杂度内归约到的非确定性多项式时间完全问题，时间复杂度是指数级的，求解难度很高。当决策空间很大时，可以应用启发式方法来寻找可行解。在

收稿日期：2023-08-17；**录用日期**：2023-09-14；**网络出版时间**：2023-10-19 18:38
网络出版地址：link.cnki.net/urlid/11.2625.V.20231019.1355.002
基金项目：国家自然科学基金企业创新发展联合基金 (U20B2042)
* **通信作者**. E-mail：luoyizhe@zzu.edu.cn

引用格式：雷耀麟，丁文锐，罗祎喆，等. 无人机数据采集任务中的航迹规划与资源分配优化 [J]. 北京航空航天大学学报，2025，51（10）：3460-3470. LEI Y L，DING W R，LUO Y Z，et al. Trajectory planning and resource allocation optimization in UAV data collection missions [J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51（10）：3460-3470（in Chinese）.

文献 [8] 中, 传感器被部署在一条直线上, 将无人机的轨迹划分为不重叠的区间进行数据采集。使用动态规划优化数据采集的间隔, 可以最大限度地减少无人机的飞行时间, 并优化无人机的飞行速度和传感器发射功率。为了从尽可能多的传感器采集数据, Samir 等[9] 提出了高复杂度的分支、归约和定界算法, 以找到相对小规模场景下的全局最优解, 开发了一种基于逐次凸逼近的有效求解算法, 以获得更大网络的结果, 并提出了一种扩展算法, 在无人机初始和最终位置先验已知的情况下, 进一步最小化无人机的飞行距离。Zhan 等[10] 研究了多个物联网设备的通信调度和发射功率分配的联合优化及在无人机物联网系统上进行数据采集的无人机航迹优化问题, 提出了一种联合优化方法, 以最大限度地减少受无人机能量约束的所有设备的最大能耗。

在应急场景中, 环境往往是动态变化的, 如受灾区域的地面目标优先级会在任务执行过程中随着卫星监测获取的新信息或相关人员的手动修改而变化。传统的优化方法往往难以解决这类问题, 而强化学习方法是解决这类动态规划问题的有效途径之一, 因此, 对强化学习方法在无人机数据采集任务规划中的应用进行研究具有重要的意义和价值。Nguyen 等[11] 设计了一种无人机辅助的物联网数据采集系统, 采用基于强化学习的训练方法, 在特定的覆盖范围内优化无人机轨迹和吞吐量。Fu 等[12] 研究了支持无线充电功能的无人机在数据采集任务中的应用, 提出了一种基于强化学习的方法对无人机进行规划, 使其从分散在物理环境中的传感器设备中采集传感器数据。Bayerlein 等[13] 提出一种基于双深度 Q 网络的无人机航迹规划算法, 用于在城市环境中从物联网设备采集数据。

综上所述, 现有无人机数据采集相关研究仍存在 2 点局限性: ①多数研究仍以最优化理论或启发式智能优化方法实现无人机任务期间的航迹规划和资源分配。然而, 随着传感器网络规模不断扩展及用户服务体验的不断升级, 数据采集呈现出“高密度-低时延”的特点, 传统优化方法逐渐不能满足其实时性的要求; ②多数基于深度强化学习的研究是在无人机能量有限或工作时间有限的条件下优化无人机航迹, 忽略了任务过程中的资源分配(计算资源和通信资源)问题, 且大多未考虑无人机缓存空间的局限性。对此, 本文提出一种基于深度强化学习的无人机航迹规划与资源分配联合优化方法, 相关创新工作可总结如下:

1) 在现有无人机资源分配模型的基础上, 进一步设计无人机数据采集模型和无人机边缘缓存模型, 并充分考虑无人机飞行、通信、计算、数据采集、缓存等物理过程, 构建无人机航迹规划与资源分配数学模型。

2) 在上述工作基础上, 进行马尔可夫决策过程构建, 设计相应的状态-行为空间, 并提出无人机能耗与数据采集信息量加权的训练奖励函数, 对两者进行平衡与调整。

3) 设计基于软决策/评价(soft actor critic, SAC)算法[14] 的决策算法训练框架, 通过仿真实验, 证明了本文方法可使无人机在消耗相近或者较小能量的情况下, 有效提升数据采集量。

1　任务场景分析与建模

无人机数据采集任务场景如图 1 所示, 任务场景包含 1 个无人机、1 个地面处理单元和 N 个地面目标, 无人机以拍摄图像的方式采集地面目标数据, 地面处理单元可以接收无人机发送的数据并处理。无人机在一片固定大小的区域内执行数据采集任务时, 由于无人机缓存空间有限, 需要在采集数据的同时机上处理数据或传输至地面处理单元进行处理。考虑到无人机电池能量有限, 而无人机空中飞行、盘旋、机上数据处理和数据传输都会消耗电池能量, 因此, 需要优化无人机航迹规划和资源分配。

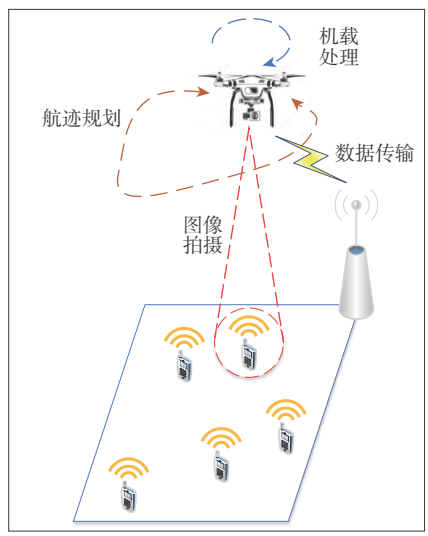


图 1　无人机数据采集任务工作环境示意图

Fig. 1　Work environment for UAV data collection tasks

能量有限、缓存空间有限的无人机在执行数据采集任务时, 需要合理的航迹规划和资源分配策略, 以高效分配无人机的计算资源和通信资源, 将采集到的数据及时机上处理或传输至地面处理单元处理。对上述问题进行工作环境建模时, 需要建立信道模型、能耗模型、缓存模型等, 如图 2 所示。

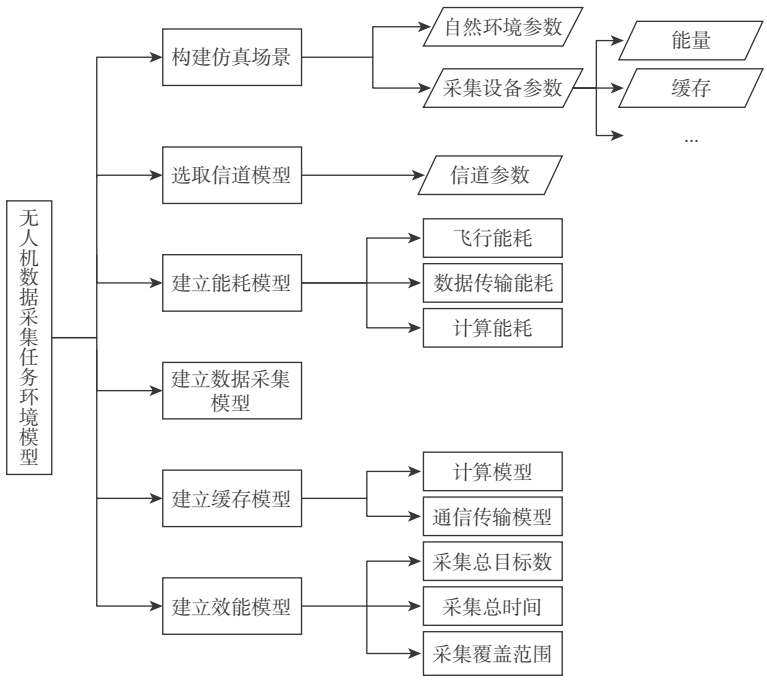


图2 无人机数据采集任务环境模型结构

Fig. 2 Environmental model structure for UAV data collection tasks

1）任务场景描述。构建无人机数据采集任务环境模型时，假设工作环境为一个正方形区域，无人机在其中执行数据采集任务，该区域内存在一定数量的目标，无人机通过高分辨率相机采集目标的高清晰度图像或视频数据。

2）信道模型构建。对无人机数据采集任务的信道进行建模时，由于无人机在高空完成数据采集，信道模型只需要考虑无人机与地面处理单元间的信道，该信道可以建模为直视信道，则数据传输速率[15]为

$$S_t = B\log_2\left(1+\frac{g_0 P_{\mathrm{Tr},t}}{\left(H^2+d_{\mathrm{u},B,t}^2\right)\sigma^2}\right) \tag{1}$$

式中：B为通信带宽；g_0为以1 W传输功率传输数据时在1 m位置处的接收功率；$P_{\mathrm{Tr},t}$为无人机传输功率；H为无人机所处高度；$d_{\mathrm{u},B,t}$为无人机与地面处理单元间的水平距离；σ^2为噪声功率。

3）数据采集模型构建。在数据采集过程建模中，对无人机任务时间进行离散处理，将其划分为T个时间段，保证每个时间段的时长足够大以完成数据采集任务。用$\mathcal{T}$描述时间段的集合，即$\mathcal{T}=\{1,2,\cdots,T\}$，对每个时间段$t\in\mathcal{T}$存在约束$t\geqslant\delta$，其中，$\delta$为单个目标的数据采集时间，每个时间段包括飞行时间和数据采集时间，分别用$t_{\mathrm{f},t}$和$t_{\mathrm{s},t}$来描述，为便于处理，将数据采集时间$t_{\mathrm{s},t}$设为δ。

假设无人机的最大水平覆盖半径为$R_{\max}$，在该范围内无人机能够拍摄到所包含目标节点的高清晰度图像。为了方便描述任意时间段t内无人机与目标间的数据采集关系，引入一个二进制指针如下：

$$I_{m,t}=\{0,1\} \qquad \forall m\in\mathcal{M},\forall t\in\mathcal{T} \tag{2}$$

式中：$I_{m,t}=1$表示无人机在时间段t内对目标m进行数据采集；$\mathcal{M}$为地面目标的集合，$\mathcal{M}=\{1,2,\cdots,M\}$。对同一目标有多次采集时，只有第一次采集结果有效，且无人机仅在剩余缓存足够完成一次数据采集时才能进行采集，因此，对该采集变量存在约束：

$$I_{m,t}d_{\mathrm{u},m,t}\leqslant R_{\max} \qquad \forall m\in\mathcal{M},\forall t\in\mathcal{T} \tag{3}$$

$$C_t>\chi \qquad \forall t\in\mathcal{T} \tag{4}$$

$$\sum_{t=1}^{T} I_{m,t}\leqslant 1 \qquad \forall m\in\mathcal{M},\forall t\in\mathcal{T} \tag{5}$$

式中：$d_{\mathrm{u},m,t}$为在时间段t内无人机与目标点m的水平距离；C_t为无人机的剩余缓存；χ为一次数据采集的数据量。式(3)表示在无人机覆盖范围内的目标才可被采集数据，式(4)表示当无人机剩余缓存足够存储一次数据采集的数据量时才可进行采集，式(5)则表示对一个目标的数据采集只有第一次是有效的。

假设无人机每次采集固定大小的数据，用χ来描述，并引入一个二进制指针用于描述无人机在时间段t之前是否已完成对目标m的数据采集任务，如下：

$$I_{\mathrm{c},t}=\begin{cases}1 & \sum_{m=1}^{M} I_{m,t}>0\\ 0 & \text{其他}\end{cases} \tag{6}$$

4) 能耗模型构建。无人机能耗模型主要包括飞行能耗、数据传输能耗和计算能耗[16]。

飞行能耗指无人机终端执行任务时的机动能耗, 参考目前主流研究, 无人机的飞行能耗主要与各时刻速度有关。假设无人机终端在 t 时刻的速度为 v_t, 为便于处理, 本节假设无人机在每个时间段内的速度保持不变, 且忽略无人机的加速与减速过程, 即无人机的飞行功率 P_{f} 保持恒定。同理, 无人机采集目标数据时在目标上空盘旋的功率 P_{h} 也恒定, 则可通过式 (7) 计算各时刻飞行能耗:

$$E_{\mathrm{f},t}=P_{\mathrm{f}}t_{\mathrm{f},t}+P_{\mathrm{h}}t_{\mathrm{h},t}\qquad\forall t\in\mathcal{T}\tag{7}$$

式中: $t_{\mathrm{h},t}$ 为时刻 t 内无人机的盘旋时间。

数据传输能耗指系统中无人机与中心处理单元间进行数据传输时的能耗。数据传输能耗与传输功率相关, 由于无人机仅在空中盘旋时进行数据传输, 数据传输能耗可以表示为

$$E_{\mathrm{Tr},t}=P_{\mathrm{Tr},t}t_{\mathrm{s},t}\qquad\forall t\in\mathcal{T}\tag{8}$$

式中: $P_{\mathrm{Tr},t}\in\left[0,P_{\mathrm{Tr,max}}\right]$ 为无人机在时间段 t 内的传输功率。

计算能耗指无人机对缓存中数据进行机上处理时的能耗。假设在时间段 t 内无人机 CPU 主频为 $F_{\mathrm{c},t}$ (单位为 cycle/s), 当其处理采集到的数据时, 其计算所消耗的能量[17] 可以表示为

$$E_{\mathrm{c},t}=\eta_{\mathrm{c}}F_{\mathrm{c},t}^{3}\left(t_{\mathrm{f},t}+t_{\mathrm{s},t}\right)\qquad\forall t\in\mathcal{T}\tag{9}$$

式中: η_{c} 为无人机边缘计算节点有效转换率; $F_{\mathrm{c},t}\in\left[0,F_{\mathrm{c,max}}\right]$。

5) 缓存模型构建。建立缓存模型时, 无人机先在任务执行过程中采集数据, 然后可以机上处理数据, 也可以将采集到的数据传回给地面处理单元进行处理。依据现有文献的建模过程, 机上处理数据量可以描述为

$$D_{\mathrm{c},t}=F_{\mathrm{c},t}\left(t_{\mathrm{f},t}+t_{\mathrm{s},t}\right)/\zeta\qquad\forall t\in\mathcal{T}\tag{10}$$

式中: ζ 为常量, 表示每输入 1 bit 数据后 CPU 的转数。

采集数据假设每次采集过程时间恒定为 δ, 数据量恒定为 χ。

6) 地面目标优先级动态转移模型构建。建立地面目标优先级动态转移模型时, 选择将其建模为马尔可夫链, 即下一时刻的目标优先级只与当前时刻的目标优先级有关, 目标动态转移矩阵 $\boldsymbol{B}$ 不随时间改变。则 t 时刻的目标优先级为

$$\boldsymbol{B}_t=\boldsymbol{B}_{t-1}\boldsymbol{B}\tag{11}$$

式中: $\boldsymbol{B}_{t-1}$ 为 t–1 时刻地面目标的优先级矩阵。

7) 算法性能评价模型构建。对于无人机数据采集任务, 需要建立算法性能评价模型来评估无人机在数据采集任务中的完成效果。选取回合总奖励、采集总信息量、飞行总能耗作为评价指标, 构成算法性能评价模型。

考虑到地面目标优先级动态变化, 且不同优先级目标的信息量不同, 因此, 为每一优先级分配相应的信息量。应急场景下, 通常会对受灾区域进行多次监测, 本文的优化目标是在单次监测时间内最大化信息量的同时减少能耗, 因此, 设计目标函数如下:

$$O_{\mathrm{bj}}=\sum_{t\in\mathcal{T}}w_t-\alpha E_t\tag{12}$$

式中: α 为能耗惩罚系数; E_t 为时间段 t 内的总能耗; w_t 为采集目标的信息量, 即各目标优先级权重。

综上所述, 无人机数据采集任务中的航迹规划与资源分配优化问题数学模型可以表示为如下形式:

$$\begin{cases}\max\limits_{l_t,\theta_t,P_{\mathrm{Tr},t},F_{\mathrm{c},t}}\dfrac{1}{T}\displaystyle\sum_{t\in\mathcal{T}}w_t-\alpha E_t\\ \text{s.t.}\begin{cases}\mathrm{C1}:\ 0\leqslant\left\|c_{\mathrm{u},t+1}-c_{\mathrm{u},t}\right\|\leqslant l_{\max}\qquad\forall t\in\mathcal{T}\\ \mathrm{C2}:\ I_{m,t}\in\{0,1\}\qquad\forall m\in\mathcal{M},\forall t\in\mathcal{T}\\ \mathrm{C3}:\ I_{\mathrm{c},t}\in\{0,1\}\qquad\forall t\in\mathcal{T}\\ \mathrm{C4}:\ I_{m,t}d_{\mathrm{u},m,t}\leqslant R_{\max}\qquad\forall m\in\mathcal{M},\forall t\in\mathcal{T}\\ \mathrm{C5}:\ \displaystyle\sum_{t=1}^{T}I_{m,t}\leqslant 1\qquad\forall m\in\mathcal{M},\forall t\in\mathcal{T}\\ \mathrm{C6}:\ \displaystyle\sum_{t=1}^{T}E_t\leqslant E_{\max}\qquad\forall t\in\mathcal{T}\\ \mathrm{C7}:\ 0\leqslant S_t t+D_{\mathrm{c},t}\leqslant C_{\max}-C_t\qquad\forall t\in\mathcal{T}\\ \mathrm{C8}:\ I_{\mathrm{c},t}\chi\leqslant C_t\qquad\forall t\in\mathcal{T}\end{cases}\end{cases}\tag{13}$$

式中: l_t 为 t 时刻的飞行距离; θ_t 为 t 时刻的飞行方位角; C1 中 $c_{\mathrm{u},t+1}$ 为 t+1 时刻无人机位置, $c_{\mathrm{u},t}$ 为 t 时刻无人机位置, 该式限制每一时间段内无人机最大飞行距离为 $l_{\max}$; C6 表示无人机任务过程中电池能耗不能超过最大能量储备 $E_{\max}$; C7 表示时间段 t 内无人机机上计算数据量与通信传输数据量不能超过机上缓存数据量, 其中, $C_{\max}$ 为无人机机上缓存空间大小; C8 表示剩余缓存不足时, 不可进行数据采集。

2　基于智能优化的航迹规划与资源分配优化

2.1　遗传算法

遗传算法[18] 是基于生物进化理论的一种优化算法, 其通过模拟自然选择、交叉和变异等生物进化过程, 以产生更好的解决方案。与其他优化方法相比, 遗传算法具有 3 大优势: ①全局优化能力强: 可以搜索全局最优解, 而不仅仅是局部最优解;

②适应性强：遗传算法可以应用于各种类型的问题，且可以根据问题特性调整算法参数以获得更好的性能；③并行性强：遗传算法可以使用多个处理器或计算机同时计算，以提高求解效率。

具体到本文问题，每一个体是对无人机航路点和资源分配编码的结果，编码方式采用二进制编码，优化变量维度为$4T$，在每一时间段内优化无人机航迹（横坐标x与纵坐标y）和资源分配（计算频率与传输功率）。编码位数依据$\log_2(V_{\mathrm{ub}}-V_{\mathrm{lb}}/A+1)$计算得到，$V$为变量，$V_{\mathrm{ub}}$和$V_{\mathrm{lb}}$分别为变量上界和下界，$A$为变量优化精度。通过种群内的选择、交叉、编译得到子代，从而实现无人机航迹规划与资源分配的联合优化，并将目标函数最大化的子代作为最优解，具体算法流程如下：

初始化交叉概率p_{c}、变异概率p_{m}、适应度函数、种群规模M、迭代次数。

对无人机航迹规划与资源分配编码，随机产生第一代种群。

Do

Do

根据每个个体的适应度从种群中选取2个个体。

if random(0, 1) < p_{c}

对选取的2个个体依交叉概率进行交叉产生新个体。

end if

if random(0, 1) < p_{m}

对选取的2个个体依变异概率进行变异产生新个体。

end if

将新产生的个体加入到新种群中。

until 新种群中个体数量达到M。

新种群代替旧种群。

从新种群所有个体提取出编码的无人机航迹与资源分配结果，计算各个体的适应度。

until 达到指定迭代次数或某一个体适应度达到指定值。

2.2 贪婪算法

贪婪算法的目标是最大化每一时刻奖励，采用的优化方法为遗传算法。假设每一时刻内地面目标优先级保持不变，从而每一时刻内的决策变成一个静态规划问题，算法流程如下：

初始化环境。

for each step

初始化交叉概率p_{c}、变异概率p_{m}、适应度函数、种群规模M、迭代次数。

对无人机航迹与资源分配编码，随机产生第一代种群。

Do

使用遗传算法优化无人机下一时刻位置及资源分配结果。

until 达到指定迭代次数或某一个体适应度达到指定值。

执行遗传算法优化所得结果，进入下一状态。

end for

每一个体是对无人机航路点和资源分配编码的结果，编码方式采用二进制编码，优化变量维度为4，分别为无人机航迹（横坐标x和纵坐标y）和资源分配（计算频率和传输功率），编码位数依据$\log_2(V_{\mathrm{ub}}-V_{\mathrm{lb}}/A+1)$计算得到，优化目标为$w_t-\alpha E_t+\beta D_t$，其中，$\beta D_t$为对时间段$t$内处理数据量的奖励，以避免剩余缓存不足时贪婪算法失效，β为数据处理奖励系数，为一常数。通过种群内的选择、交叉、编译得到子代，从而实现无人机航迹规划与资源分配的联合优化，并将目标函数最大化的子代作为最优解。

3 基于深度强化学习的航迹规划与资源分配优化

智能优化方法虽然在多数资源优化问题中有较好的表现，但仍存在一定局限性，例如，对于高维度、非线性和非凸优化问题，遗传算法的求解效率可能会受到影响。此外，智能优化方法的时间复杂度往往较高，在时效性方面往往较差。为此，本文进一步引入深度强化学习机制，提出一种基于深度强化学习的无人机航迹规划与资源分配联合优化方法，并将无人机与环境的交互过程建模为马尔可夫决策过程，通过最大化每回合累积奖励的方法来提高无人机对地面目标的数据采集能力。

3.1 马尔可夫决策过程建模

为将深度强化学习应用于所提无人机航迹与资源优化问题，需要将优化问题建模为马尔可夫决策过程。在马尔可夫决策过程中存在一个智能体与环境交互，该过程可以表示为一个四元组$[s(t), a(t), r(t), s(t+1)]$。其中，$s(t)$为智能体在时隙$t$中的状态，状态可以是离散的或连续的，有限的或无限的；$a(t)$为智能体可选择的行为；$r(t)$为在某个状态下，执行某个动作后，环境反馈给智能体的奖赏或惩罚，即$s(t)a(t)\rightarrow r(t)$，$s(t+1)$为神经网络施加行为$a(t)$后在下一个时隙中的状态。深度强化学习的目标是学习到一个最大化累积奖励的策略$\pi: s\rightarrow a$，即一个行为函数，输入当前状态可输出最优动作，

其是一个状态到动作的映射。在本文所提问题中，上述马尔可夫过程的状态、行为和奖励定义如下。

3.1.1　状态定义

智能体需感知环境状态，从而做出决策，环境状态包括：

1）无人机当前位置，其水平面内的二维坐标可表示为：$c_{\mathrm{u},t}=(x_t,y_t)$。

2）无人机剩余电池能量，当前时刻无人机电池能量通过上一时刻电池能量减去上一时刻内所消耗能量得到。

3）无人机剩余缓存，每时刻无人机会将采集到的数据存储到缓存空间，也会通过机上处理和传输至地面处理单元处理。

4）地面目标采集列表，地面目标图像信息若已采集，则标为 1，若未采集，则标为 −1，由所有地面目标的采集标志位拼接得到地面目标采集列表。

5）地面目标优先级，地面目标具有不同的优先级，不同优先级的信息量不同，由所有地面目标的优先级拼接得到地面目标优先级列表。

综上所述，无人机航迹规划与资源分配优化问题在 t 时刻的状态可以表示为

$$s(t)=\left\{c_{\mathrm{u},t},E_t,C_t,L_t,\boldsymbol{B}_t\right\} \tag{14}$$

其中：L_t 为目标服务列表，指示地面目标是否已完成数据采集。

3.1.2　行为定义

智能体做出的具体决策行为包含：

1）无人机坐标，在时刻 t 决定下一时刻无人机位置，用二维坐标 (x_{t+1},y_{t+1}) 表示，决策变量也可以采用速度和航向，本节直接采用坐标，更便于处理。

2）无人机计算频率和传输功率，计算频率对应无人机计算资源，传输功率对应无人机通信资源，直接影响各时刻无人机机上处理数据量和传输数据量。

综上所述，智能体在 t 时刻的决策动作可以表示为

$$a(t)=\left\{x_t,y_t,P_{\mathrm{Tr},t},F_{\mathrm{c},t}\right\} \tag{15}$$

3.1.3　奖励定义

目标函数对应深度强化学习方法中的奖励函数，在 t 时刻的奖励 $r(t)$ 可以表示为

$$r(t)=w_t-\alpha E_t \tag{16}$$

能耗惩罚系数 α 意味着无人机的目标是在保证对所有目标完成数据采集的同时最小化能耗。

3.2　基于 SAC 的无人机航迹规划与资源分配决策

考虑到上述决策行为变量为连续变量，采用深度强化学习方法中的 SAC 算法，该算法是解决连续动作空间的常用深度强化学习方法之一。

SAC 算法是一种无模型的深度强化学习方法，结合了决策/评价方法和最大熵强化学习的思想。SAC 算法通过使用深度神经网络来表示策略和值函数，可以解决高维连续动作空间的复杂任务。

SAC 算法主要包括策略网络和价值网络，如图 3 所示。策略网络用于智能体选择动作，输入为当前状态，输出为动作分布（通常为高斯分布），价值网络为智能体估计状态-动作对 Q 值（价值）的网络。SAC 算法使用 2 个独立的价值网络，以降低过拟合风险并提高训练稳定性，通过引入熵的概念，在学习到一个高效策略的同时，也鼓励探索策略的随机性，提高了算法的鲁棒性和采样效率。本节所提基于 SAC 的无人机航迹规划与资源分配决策方法如下：

初始化环境：策略网络参数 ϕ、价值网络参数 φ 与 $\hat{\varphi}$，清空经验缓存池 $\mathcal{D}$。

将价值网络参数赋值给目标价值网络 $\hat{\theta}\leftarrow\theta$

Do

观测当前所处状态，依据策略网络输出动作，即无人机航迹坐标与其资源分配的决策结果 $a\leftarrow\pi_\phi(\cdot|s)$。

在环境中执行该动作，飞往指定位置，分配指定计算资源和通信资源处理数据。

进入下一状态，并观测状态、奖励及回合是否结束。

将前一时刻状态、所执行动作、当前时刻状态、所得奖励、结束信号存入经验缓存池中。

if 回合结束

then 重置环境

for $i=1$: K（梯度更新次数）

随机从经验缓存池中选取一批数据，更新策略网络和价值网络参数。

end for

end if

until 收敛

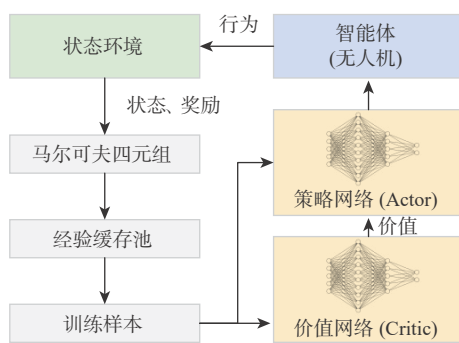


图 3　SAC 算法与环境交互示意图

Fig. 3　Interaction between SAC algorithm and environment

经验缓存池$\mathcal{D}$用于存放智能体采样结果，智能体根据当前策略π选择一个动作并执行，观察环境返回的新状态和奖励。

参数更新包含策略网络和价值网络的参数更新。其中，使用新的状态-动作对和奖励来更新价值网络的参数时，通常使用均方误差作为损失函数，如下：

$$J_Q(\theta)=e_{(s(t),a(t))\sim\mathcal{D}}\left[\frac{1}{2}(Q_\theta(s(t),a(t))-(r(s(t),a(t))+\gamma e_{s(t)}[V_{\hat{\theta}}(s(t+1))]))^2\right] \tag{17}$$

式中：e为期望值；$Q_\theta(s(t),a(t))$为价值网络对状态动作对$(s(t),a(t))$的Q值预测值；$r(s(t),a(t))$为环境反馈奖励；γ为折扣因子；$V_{\hat{\theta}}(s(t+1))$为通过目标价值网络计算得到的状态价值，可由式(18)计算得到：

$$V_{\hat{\theta}}(s(t+1))=e_{a_t\sim\pi}\left[Q_{\hat{\theta}}(s(t),a(t))-\alpha\lg\pi(s(t),a(t))\right] \tag{18}$$

式中：$Q_{\hat{\theta}}(s(t),a(t))$为目标价值网络的预测$Q$值；$\lg\pi(s(t),a(t))$为策略熵，有时也会使用一个单独的状态价值网络直接预测状态价值。

使用价值网络评估动作的价值，可以最大化期望的Q值和策略熵，从而更新策略网络的参数，其损失函数为

$$J_\pi(\phi)=e_{s_t\sim\mathcal{D}}\left\{e_{a_t\sim\pi}[\alpha\lg\pi(s(t),a(t))-Q_\theta(s(t),a(t))]\right\} \tag{19}$$

最后，使用软更新策略来更新价值目标网络的参数，以微小幅度不断更新参数，更新方式如下：

$$\hat{\theta}=\tau\theta+(1-\tau)\hat{\theta} \tag{20}$$

式中：τ为软更新系数，一般取0.005。

4 仿真验证

4.1 仿真参数设置

执行无人机数据采集任务的场景为200 m×200 m的矩形区域，区域内有20个地面目标，无人机以匀速飞行，飞行速度为20 m/s，采集地面目标图像数据所需时间为10 s，最大水平覆盖半径为10 m；无人机机上最大计算频率为10 GHz，单次数据采集的数据量为5 GB；最大数据传输功率为100 W，通信带宽为20 MHz，飞行功率为220 W，空中盘旋功率为160 W；无人机电池能量为200 kJ，飞行高度为100 m，平均为每个地面目标的数据采集分配20 s时间，总时间为400 s，超过总时间则采集任务终止。

无人机与地面处理单元初始位置皆为[0,0]，地面目标位置为随机生成。地面目标优先级分为4级，最高为Ⅳ级，最低为Ⅰ级，Ⅰ级目标共7个，Ⅱ级目标和Ⅲ级目标各5个，Ⅳ级目标共3个，目标优先级会动态变化，优先级动态转移矩阵如表1所示，各优先级目标的信息量如表2所示。

表1 地面目标优先级动态转移矩阵

Table 1 Ground target priority dynamic transfer matrix

优先级	优先级概率			
	Ⅰ	Ⅱ	Ⅲ	Ⅳ
Ⅰ	0.8	0.1	0.08	0.02
Ⅱ	0	0.8	0.16	0.04
Ⅲ	0	0	0.9	0.1
Ⅳ	0	0	0	1

表2 各优先级目标的信息量

Table 2 Information volume of targets with different priorities

优先级	信息量
Ⅰ	1
Ⅱ	2
Ⅲ	3
Ⅳ	4

本节所采用的遗传算法、SAC算法、贪婪算法的相关参数如表3所示。其中，遗传算法和SAC算法对数据采集过程中所有时刻的无人机航迹与资源分配进行优化，贪婪算法使用的基线算法为遗传算法，对数据采集过程中某一时刻的无人机航迹与资源分配进行优化。考虑到地面目标优先级的动态变化特性，3种算法具有不同的特点：①在遗传算法中，每一次种群评价过程需要测试多个回合，本节设为100；②由于优化单一时刻内的无人机航迹与资源分配不需要考虑环境的动态特性，优化问题更易于求解，因此，贪婪算法中种群大小和迭代次数均小于遗传算法，求解速度更快，但全局的优化耗时比遗传算法长；③SAC算法作为一种常用的处

表3 无人机航迹规划与资源分配联合优化方法参数

Table 3 Parameters for joint optimization algorithm in UAV trajectory planning and resource allocation

算法	种群测试回合数	种群大小	迭代次数	学习率	网络结构	缓存池大小	批大小	折扣因子	软更新系数	训练步数
遗传算法	100	200	1 000							
SAC算法				3×10^{-4}	128×128×128	500 000	256	0.99	0.005	5 000 000
贪婪算法		100	200							

理连续动作空间的强化学习方法, 实验结果对超参数的依赖相比其他强化学习方法相对较小。

深度强化学习的训练环境基于 Pytorch 1.10.1 设计开发, 所使用的编程语言为 python3.7, 并利用开源深度强化学习工具包 StableBaselines3 对方法框架进行搭建。训练硬件环境为第 12 代 Intel Core i5-12600KF 3.69 GHz、32 GB RAM、NVIDIA GeForce RTX 3060。使用的遗传算法基于智能优化方法库 Scikit-opt 进行开发。

4.2　仿真结果及分析

采用遗传算法和贪婪算法作为基线方法, 与 SAC 算法进行性能对比, 验证所提深度强化学习方法的有效性, 评价指标使用每回合的平均奖励、采集信息量和能耗。

对无人机不同初始缓存空间下航迹规划与资源分配优化方法的性能进行对比, 由于地面目标总数为 20, 对每一目标的数据采集数据量为 5 GB, 则最多需要 100 GB 存储空间即可存储所有目标数据。因此, 本节在无人机存储空间分别为 8 GB、16 GB、32 GB、64 GB、128 GB, 能耗惩罚系数 $\alpha=1\times10^{-4}$, 贪婪算法数据处理奖励系数 $\beta=1\times10^{-9}$ 的条件下测试算法优化性能, 测试结果如表 4 所示。可知, 所提 SAC 算法在奖励、采集信息量方面均优于遗传算法和贪婪算法, 但所消耗能量相比于遗传算法略有上升, 综合评价优于另外 2 种对比基线方法。

由于环境的动态特性, 即地面目标优先级动态变化特性的存在, 不同回合内算法性能有一定差距, 因此, 需要通过测试多个回合评估实际性能。为更清晰地分析各算法间的性能差异, 各算法性能分析如图 4～图 6 所示, 对比指标分别为回合总奖励均值、回合总能耗均值、回合采集总信息量均值。

由图 4～图 6 可知: ①SAC 算法的每回合总奖励和采集总信息量显著优于遗传算法和贪婪算法, 而总能耗相比遗传算法大, 可能原因是目标函数中更多地考虑总信息量, 对能耗的惩罚系数 α 相对较

表 4　不同算法在不同初始缓存下的性能对比

Table 4　Performance comparison of different algorithms with varying initial cache sizes

算法	缓存/GB	奖励	能耗/kJ	采集信息量
遗传算法	8	52.79±3.26	83.10±0	61.10±3.27
	16	55.04±3.02	64.40±0	61.48±3.02
	32	53.18±3.31	48.95±0	58.08±3.31
	64	55.37±3.42	37.82±0	59.15±3.42
	128	54.77±3.42	36.32±0	58.40±3.42
贪婪算法	8	36.39±3.29	143.32±23.10	50.72±3.87
	16	46.56±3.71	172.39±18.79	63.80±3.73
	32	47.10±3.93	179.48±20.72	65.05±3.52
	64	49.11±3.41	119.23±15.29	61.03±3.04
	128	50.90±3.03	96.51±14.85	61.55±3.11
SAC算法	8	62.10±2.42	88.86±6.58	70.97±2.33
	16	64.83±2.93	71.29±13.10	71.96±2.28
	32	66.25±1.91	59.17±5.51	72.17±1.86
	64	63.12±2.79	45.44±8.72	67.66±2.19
	128	63.24±2.14	40.08±5.89	67.24±2.16

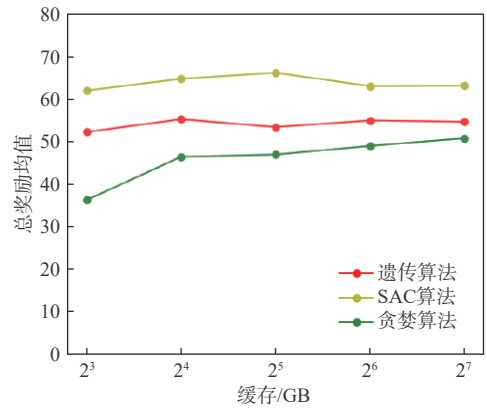


图 4　不同算法总奖励均值对比

Fig. 4　Comparison of average total rewards for different algorithms

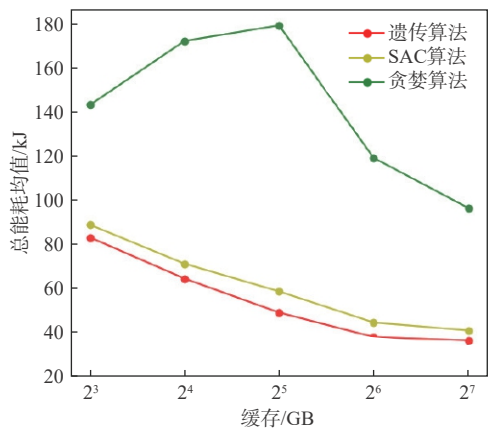


图 5　不同算法的总能耗均值对比

Fig. 5　Comparison of average total energy consumption for different algorithms

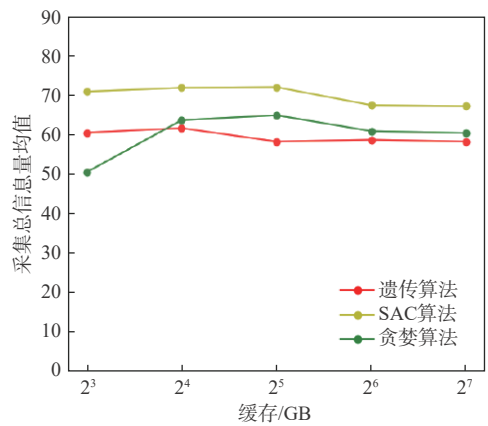


图 6　不同算法的采集总信息量均值对比

Fig. 6　Comparison of average total information collected for different algorithms

小，此处$\alpha=1\times10^{-4}$。不同于遗传算法的全局优化，SAC算法在训练过程中会更关注当下获得的奖励，对未来可能获得的奖励会通过折扣因子γ进行衰减，因此，过小的α会导致算法忽略对能耗的惩罚；②贪婪算法在总奖励和总能耗2项指标上的表现均为最差，且总能耗远高于其他2种算法，因为目标优先级存在动态变化及能耗惩罚较小，每一时间段内贪婪算法总会选择高优先级的目标，忽略了可能带来的能耗，从而导致有限时间内采集信息量更低，能耗更高。但是，贪婪算法在采集总信息量指标上的表现随着缓存增加而超越遗传算法，而仍低于SAC算法，这是因为贪婪算法的任务时间更长，采集到了更多高优先级目标的信息，却未能采集完成全部目标的数据。

为测试α值对算法性能的影响，在无人机初始缓存为32 GB的条件下分别测试$\alpha=1\times10^{-5}$，$\alpha=1\times10^{-4}$，$\alpha=1\times10^{-3}$时的算法性能，测试结果如图7～图9所示。

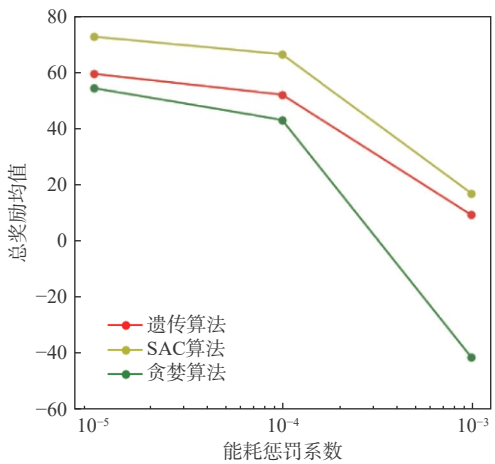


图7 不同能耗惩罚系数下不同算法总奖励均值对比

Fig. 7 Comparison of total rewards for different algorithms under varying energy consumption penalty coefficients

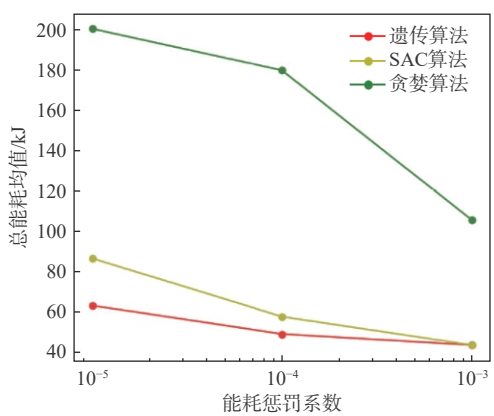


图8 不同能耗惩罚系数下不同算法总能耗均值对比

Fig. 8 Comparison of total energy consumption for different algorithms under varying energy consumption penalty coefficients

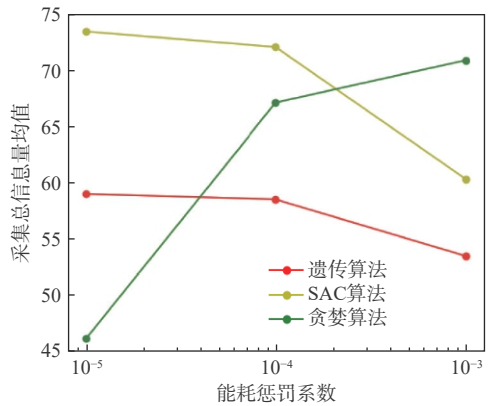


图9 不同能耗惩罚系数下不同算法采集总信息量均值对比

Fig. 9 Comparison of total information collected for different algorithms under varying energy consumption penalty coefficients

由图7～图9可知：①SAC算法的每回合总奖励依旧高于遗传算法和贪婪算法，而飞行总能耗则逐步降低，在$\alpha=1\times10^{-3}$时，能耗控制与遗传算法相近，而采集总信息量仍然高于遗传算法，可见，增大α使SAC算法更加关注能耗的影响，为了降低能耗而牺牲了部分信息量；②贪婪算法在总奖励和总能耗2项指标上的表现仍然为最差，$\alpha=1\times10^{-5}$时，对能耗惩罚过小，导致无人机会在采集完成全部目标数据前耗尽电池能量；$\alpha=1\times10^{-3}$时，总奖励最低，而总信息量最高，是以超出SAC算法和遗传算法1.5倍能耗和更长的任务时间为代价达到的。

5 结 论

1）本文围绕无人机数据采集任务开展相应的航迹规划与资源分配优化，充分考虑了无人机飞行、通信、计算、数据采集等物理过程，并进行了相关数学模型构建。

2）在现有SAC框架的基础上，设计了无人机能耗与数据采集信息量加权的训练奖励函数，可通过合理分配权重进而对两者进行平衡与调整。

3）对比仿真实验结果表明，相较于遗传算法和贪婪算法，所提的基于深度强化学习的无人机航迹规划与资源分配联合优化方法可使无人机在消耗相近或者较小能量的情况下，采集到较大的数据量。

参考文献（References）

[1] 胡欣颖. 面向信息采集场景的无人机轨迹规划与通信资源分配研究[D]. 北京: 北京邮电大学, 2021: 1-3.
HU X Y. Research on UAV trajectory planning and communication resource allocation for information acquisition scenes[D]. Beijing: Beijing University of Posts and Telecommunications, 2021: 1-3(in Chinese).

[2] LI Z R, ZHANG Y W, WU H, et al. Design and application of a

UAV autonomous inspection system for high-voltage power transmission lines[J]. Remote Sensing, 2023, 15(3): 865.

[3] DONG M X, OTA K, LIN M, et al. UAV-assisted data gathering in wireless sensor networks[J]. The Journal of Supercomputing, 2014, 70(3): 1142-1155.

[4] GONG J, CHANG T H, SHEN C, et al. Flight time minimization of UAV for data collection over wireless sensor networks[J]. IEEE Journal on Selected Areas in Communications, 2018, 36(9): 1942-1954.

[5] CHEN X, CHEN X M. The UAV dynamic path planning algorithm research based on Voronoi diagram[C]//Proceedings of the 26th Chinese Control and Decision Conference. Piscataway: IEEE Press, 2014: 1069-1071.

[6] HSU D, LATOMBE J C, KURNIAWATI H. On the probabilistic foundations of probabilistic roadmap planning[J]. International Journal of Robotics Research, 2006, 25(7): 627-643.

[7] LIU X. Four alternative patterns of the Hilbert curve[J]. Applied Mathematics and Computation, 2004, 147(3): 741-752.

[8] MU X D, LIU Y W, GUO L, et al. Energy-constrained UAV data collection systems: NOMA and OMA[J]. IEEE Transactions on Vehicular Technology, 2021, 70(7): 6898-6912.

[9] SAMIR M, SHARAFEDDINE S, ASSI C M, et al. UAV trajectory planning for data collection from time-constrained IoT devices[J]. IEEE Transactions on Wireless Communications, 2020, 19(1): 34-46.

[10] ZHAN C, LAI H. Energy minimization in internet-of-things system based on rotary-wing UAV[J]. IEEE Wireless Communications Letters, 2019, 8(5): 1341-1344.

[11] NGUYEN K K, DUONG T Q, DO-DUY T, et al. 3D UAV trajectory and data collection optimisation *via* deep reinforcement learning[J]. IEEE Transactions on Communications, 2022, 70(4): 2358-2371.

[12] FU S, TANG Y J, WU Y, et al. Energy-efficient UAV-enabled data collection via wireless charging: a reinforcement learning approach [J]. IEEE Internet of Things Journal, 2021, 8(12): 10209-10219.

[13] BAYERLEIN H, THEILE M, CACCAMO M, et al. UAV path planning for wireless data harvesting: a deep reinforcement learning approach[C]//Proceedings of the IEEE Global Communications Conference. Piscataway: IEEE Press, 2020: 1-6.

[14] HAARNOJA T, ZHOU A, ABBEEL P, et al. Soft actor-critic: off-policy maximum entropy deep reinforcement learning with a stochastic actor[C]//Proceedings of the International Conference on Machine Learning. Piscataway: IEEE Press, 2018: 1861-1870.

[15] ISLAM S K, HAIDER M R. Sensors and low power signal processing[M]. Berlin: Springer Science & Business Media, 2009.

[16] LIU Q, SHI L, SUN L L, et al. Path planning for UAV-mounted mobile edge computing with deep reinforcement learning[J]. IEEE Transactions on Vehicular Technology, 2020, 69(5): 5723-5728.

[17] WU Q Q, ZENG Y, ZHANG R. Joint trajectory and communication design for multi-UAV enabled wireless networks[J]. IEEE Transactions on Wireless Communications, 2018, 17(3): 2109-2121.

[18] LAMBORA A, GUPTA K, CHOPRA K. Genetic algorithm-a literature review[C]//Proceedings of the International Conference on Machine Learning, Big Data, Cloud and Parallel Computing. Piscataway: IEEE Press, 2019: 380-384.

Trajectory planning and resource allocation optimization in UAV data collection missions

LEI Yaolin[1, 2], DING Wenrui[3], LUO Yizhe[4, *], WANG Yufeng[3], LIU Siqi[5], ZHANG Zhilan[1]

(1. School of Electronics and Information Engineering, Beihang University, Beijing 100191, China;
2. China Electronics Technology Group Corporation 54th Research Institute, Shijiazhuang 050081, China;
3. Institute of Unmanned System, Beihang University, Beijing 100191, China;
4. School of Computer and Artificial Intelligence, Zhengzhou University, Zhengzhou 450053, China;
5. State Grid Xinxiang Electric Power Supply Company, Xinxiang 453000, China)

Abstract: A joint optimization method for unmanned aerial vehicle (UAV) trajectory planning and resource allocation based on deep reinforcement learning was proposed to address the challenges of limited battery capacity, limited cache space, and dynamic changes in ground target priorities during data collection tasks in emergency scenarios. First, a mathematical model was developed by considering the communication, computation, flight, and data caching processes in UAV missions. Then, a Markov process model was established for UAV trajectory planning and resource allocation, with corresponding state and action descriptions. A weighted reward function was designed to balance UAV energy consumption and data collection volume. Finally, simulations were conducted to compare the proposed method with greedy algorithms and genetic algorithms. The results show that the proposed method can significantly improve the amount of data collected from ground users within a shorter task time, at a similar or lower energy cost for UAVs.

Keywords: unmanned aerial vehicle; resource allocation; trajectory planning; reinforcement learning; mobile edge computing

Received: 2023-08-17; **Accepted**: 2023-09-14; **Published Online**: 2023-10-19 18:38

URL: link.cnki.net/urlid/11.2625.V.20231019.1355.002

Foundation item: National Natural Science Foundation of China-Enterprise Innovation and Development Joint Fund (U20B2042)

* **Corresponding author**. E-mail: luoyizhe@zzu.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0566

复杂光照环境下民用飞机飞行员关键点检测方法

刘豪，孙有朝*，吴红兰，张海燕
(南京航空航天大学 民航学院，南京 211106)

摘 要：基于计算机视觉技术对民用飞机飞行员操纵行为进行识别和监控，对于确保民用航空运行安全具有重要的现实意义。提出一种复杂光照环境下的飞行员关键点检测方法。鉴于民用飞机驾驶舱复杂光照环境，提出图像亮度调节模块，该模块分级判断图像亮度均值，实现不同亮度图像特征融合，提升运行速度，并最大化保留图像细节特征；鉴于较多的关键点定位是精确行为识别的基础，提出轻量化的飞行员肢体关键点和手部关键点检测网络，该网络在高分辨率分支运用次序交换注意力模块，以缓解原始视觉注意力计算成本随输入分辨率增加呈二次增长的问题，联合部署飞行员肢体和手部关键点检测网络，选择典型的飞行动作进行实验验证；进行大量消融实验，定量和定性地探讨不同组件(图像亮度调节模块、次序交换注意力模块)对模型性能的影响，建立所提方法与预测结果之间的可解释关系。提出的模型在飞行员关键点检测数据集和MS COCO val2017数据集上的AP分别达到81.9%和72.8%，兼顾精度和实时性。

关 键 词：民用飞机；智能驾驶舱；复杂光照；飞行员；关键点检测；视觉注意力
中图分类号：V249.1；TB553
文献标志码：A **文章编号**：1001-5965(2025)10-3471-11

人工智能和其他新兴技术与航空工业的融合引领着民航业的技术进步[1-3]。在民用飞机飞行过程中，重新构建驾驶舱中人机交互模式，成为重大的变革方向之一。实时识别飞行员行为可以为决策者提供关键信息，减少飞行操纵过程中的人为差错[4-5]。目前的方法包括可穿戴传感技术[6]和基于计算机视觉的非接触技术[7]。虽然可穿戴传感技术较为成熟，但由于飞行员需要佩戴传感设备，其使用可能会干扰人机交互过程。而计算机视觉技术具有自然、无接触的特点，逐渐成为民用飞机驾驶舱飞行员行为识别的主流技术。民用飞机飞行员关键点检测是进行操纵行为监测的基本任务之一，然而由于驾驶舱内部复杂光照环境和个体差异等原因，使得飞行员关键点检测任务非常具有挑战性。

复杂光照环境对驾驶舱内部图像的表现分为以下3个方面：①在不同的飞行任务场景中，驾驶舱暴露在不同等级亮度的光环境中；②驾驶舱内部亮度环境在一次飞行任务中变化过快；③受整体曝光和暗度影响，部分图像具有较多的噪点。此外，由于民用飞机驾驶舱本身结构部件繁多，在复杂光照环境下，部分仪表区域细节难以被精确描述。

为改善复杂光照环境引起的图像亮度变化问题，已经有一些研究成果，这些方法包括基于分布映射方法[8]、基于模型优化方法[9-10]和基于深度学习方法[11-12]。基于分布映射方法作为一种直观的低照度增强方案被提出，通过映射数值分布放大较小像素值，该方法一般被分为直方图均衡法[13-14]和基于S曲线方法[15]。基于模型优化方法依赖于对现实情况的假设，这类方法的代表是基于Retinex理论方法[10]。近年来，基于深度学习的低照度图像增强方法成为主流。从实现目的来看，该方法分为2类：亮度增强方法[16]和增强亮度与去除噪声相结

收稿日期：2023-09-04；**录用日期**：2023-10-13；**网络出版时间**：2023-11-22 09:44
网络出版地址：link.cnki.net/urlid/11.2625.V.20231121.1058.002
基金项目：国家自然科学基金-民航联合研究基金(U2033202,U1333119)；国家自然科学基金(52172387)
*通信作者. E-mail：sunyc@nuaa.edu.cn

引用格式：刘豪，孙有朝，吴红兰，等. 复杂光照环境下民用飞机飞行员关键点检测方法[J]. 北京航空航天大学学报，2025，51(10)：3471-3481. LIU H，SUN Y C，WU H L，et al. Key points detection method for civil aircraft pilot in complex lighting environments[J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51(10)：3471-3481 (in Chinese).

合的方法[17]。然而,尽管现有的图像亮度调节技术表现良好,但其在复杂光照环境下的实际应用问题值得进一步考虑。特别是在民用飞机驾驶舱内复杂环境下的关键点定位问题。因此,本文的目标是提出一种在复杂光照下调节驾驶舱图像亮度的方法,以提高图像亮度,同时保留不同亮度下的图像细节,为飞行员关键点定位提供图像基础。

人体关键点定位的目的是预测图像中关键点的位置[18]。近年来,随着人工智能技术的发展,人体关键点定位研究取得重大突破。目前的方法从不同角度提高人体关键点定位精度,如网络架构设计[19]、多尺度特征提取[20-21]和注意力机制[22-23]。全局相关性对于人体姿态估计至关重要,而卷积神经网络(convolutional neural network, CNN)的局部性会限制其捕捉远距离相互作用的能力。为解决该问题,一些研究将Transformer架构纳入人体关键点定位领域。Yang等[24]将注意力层置于CNN主干网后端,以有效建模全局特征依赖关系。Dosovitskiy等[25]使用改进的视觉Transformer结构,以回归关键点热图。此外,一些研究将Transformer应用于三维人体关键点定位领域。Lin等[26]通过引入Transformer将多视角的特征结合起来。

足够数量的关键点位置信息是识别复杂行为的基础,因此,捕捉民用飞机飞行员的操纵动作需要较多的关键点信息。对于“打开自动导航按钮”、“打开起落架”等操作动作,需要联合手部关键点检测网络和肢体关键点检测网络。然而,在一张图像上进行多个关键点识别,会带来较大的计算负担,降低复杂行为识别的实时性。

为降低联合模型参数量和复杂度,提高模型推理的实时性,刘豪等[27]提出结合全局上下文信息的关键点定位模型,该模型继承高分辨率网络(high-resolution network, HRNet)的多尺度网络主干,使用深度可分离卷积联合Transformer。然而,视觉自注意力的计算瓶颈是Softmax注意模块,该模块的计算成本随输入分辨率的增加呈二次增长[28]。Yuan等[22]提出高分辨率Transformer(high-resolution Transformer, HRFormer)模型,该方法修改HRNet主干的内部结构,使用基于窗口的Transformer模块代替卷积模块,通过在前向反馈网络(forward feedback network, FFN)层中引入3×3深度可分离卷积,实现在断开的图像窗口之间交换信息。然而,将自注意力限制在固定大小的局部窗口内,损害视觉Transformer的非局部注意能力,降低全局信息交互。此外,还有一些方法通过采用结构化稀疏注意力,将Softmax注意分解成2个函数近似逼近Softmax注意[29]。

基于以上文献,在复杂光照环境下,精确和实时地进行飞行员关键点检测存在以下挑战:

1)由于视觉技术关键点定位对光照环境十分敏感,需要设计一种快速且保留图像局部细节的图像亮度调节方法。

2)复杂动作识别需要肢体和手部关键点,更多的节点造成模型推理速度下降,设计轻量化的关键点检测模型成为现实需求。

3)设计一种低复杂度的全局交互自注意力模型,同时近似达到原始视觉Transformer的特征提取能力。

基于以上考虑,本文提出一种民用飞机复杂光照环境下的飞行员关键点检测模型。首先,针对民用飞机驾驶舱复杂的光照环境,提出图像亮度调节模块,该模块融合多幅不同亮度的图像,有效地保留图像局部区域的细节信息。同时,引入轻量级的飞行员肢体关键点检测网络,该网络在不同分支运用次序交换注意力和原始注意力,以降低模型参数量和复杂度。其次,在FFN层中额外插入深度卷积,以此扩大模型的感受野,从而增强局部感知能力。基于肢体关键点检测模型,设计手部关键点检测网络。飞行员肢体和手部关键点检测网络被联合部署。最后为验证所提模型的有效性及不同组件对关键点定位的影响,在模型与预测结果之间建立可解释的关系。通过消融实验定量和定性地探讨设计组件(图像亮度调节模块、次序交换注意力模块)对模型性能的影响,使预测结果具有可追溯性。

本文旨在提高复杂光照环境下飞行员关键点定位的准确性和实时性,为民用飞机驾驶舱的人机交互任务提供坚实的基础。

1 本文方法

本文方法由3个部分组成:用于调整图像亮度的图像亮度调节模块、飞行员肢体关键点检测网络和飞行员手部关键点检测网络。

1.1 图像亮度调节模块

本文提出的分级图像亮度调节模块如图1所示。将RGB图像转换到色相、饱和度、亮度(hue, saturation, value, HSV)空间,对图像亮度情况进行计算与分类。根据图像中亮度分量,计算分量均值得到$\overline{M}_{\mathrm{V}}$。设置亮度类别参数$\sigma_1$、$\sigma_2$,将$\overline{M}_{\mathrm{V}}$与$\sigma_1$、$\sigma_2$进行比较。若$\overline{M}_{\mathrm{V}}<\sigma_1$,为弱光图像,则需要进行亮度增强处理;若$\sigma_1<\overline{M}_{\mathrm{V}}<\sigma_2$,则为亮度适中图像,不需要进行亮度调节处理;若$\overline{M}_{\mathrm{V}}>\sigma_2$,则为强光图像,需要进行亮度减弱处理。亮度类别参数通

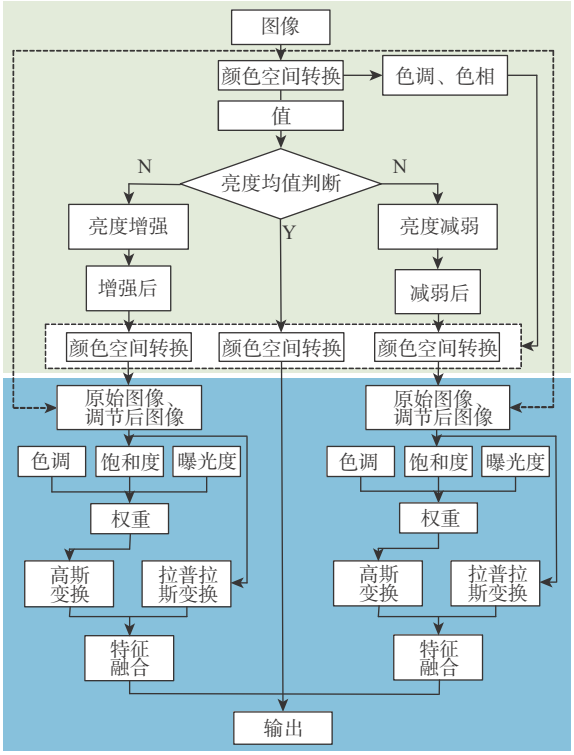


图 1　图像亮度调节模块整体结构

Fig. 1　Architecture for image brightness adjustment module

常设置为σ_1=1/3, $\sigma_2=2/3$。

1.1.1　增强和减弱图像亮度

伽马变换具有改善图像视觉效果、运算简单的优点。然而在以往的研究中，伽马变换参数需要手动设置，在进行调节时导致图像局部亮度不均匀。

针对上述问题，构造一种自适应的伽马变换方法。该方法利用亮度分量的分布特性自适应地调整伽马函数的参数。对于原始输入图像，伽马定义如下：

$$F(i,j)=255(f(i,j)/f_{\max})^{\gamma} \tag{1}$$

式中：$f(i,j)$ 为待处理图像亮度分量；$f_{\max}$ 为最大像素值；γ 为用于控制图像亮度的可变数值。改进的伽马变换将利用自适应的 γ 值对图像进行不同程度的调节。计算图像部分 7×7 领域内的局部均值 μ 及平均偏差 c，并将其归一化。

对于低对比度输入图像，γ 计算方法如下：

$$\gamma=-\log_2((\bar{\mu}+\bar{c})/2) \tag{2}$$

式中：$\bar{\mu}$ 为均值归一化后的结果；$\bar{c}$为偏差归一化后的结果。

对于高对比度输入图像，γ 计算方法如下：

$$\gamma=\exp(((\bar{\mu}+\bar{c})-1)/2) \tag{3}$$

如图 1 所示，在图像亮度调整后，通过将值分量与色调和饱和度分量融合，进一步处理得到最终的 RGB 图像。图 2 为驾驶舱图像亮度增强和减弱前后的结果，展示了本文方法的有效性。



图 2　图像亮度调节的可视化效果

Fig. 2　Visualization of image brightness adjustment

1.1.2　图像融合

为进行图像融合，需要检测出每张图像细节丰富、质量高的区域，通过构建图像质量评估指标进行检测。根据图像灰度信息，采用拉普拉斯滤波[30]，选取滤波响应绝对值为色调指标。

在过暗和过亮图像中，饱和度低的区域所含细节内容较少。针对该问题，对于每个像素点，根据图像的饱和度信息，分别计算每个像素点R、G、B通道内的标准差，以此得到饱和度指标 A，如下：

$$m_{\mathrm{u}}=(B_{\mathrm{R}}+B_{\mathrm{G}}+B_{\mathrm{B}})/3 \tag{4}$$

$$A=\left[\frac{(B_{\mathrm{R}}-m_{\mathrm{u}})^2+(B_{\mathrm{G}}-m_{\mathrm{u}})^2+(B_{\mathrm{B}}-m_{\mathrm{u}})^2}{3}\right]^{1/2} \tag{5}$$

式中：B_{R}、B_{G}、B_{B}为各通道数值经过归一化后的结果。

根据与 0.5 的接近程度，对各个颜色通道数值点进行加权，以此保持数值不接近 0 (曝光不足) 或 1 (曝光过度)，加权权重与高斯曲线一致。以 R 通道归一化数值为例，具体描述如下：

$$B_{\mathrm{R}}^{\mathrm{e}}=\exp\left(-\frac{(B_{\mathrm{R}}-0.5)^2}{2\rho^2}\right) \tag{6}$$

式中：ρ为高斯常量，取 0.2；$B_{\mathrm{R}}^{\mathrm{e}}$为加权数值结果。

为了考虑多个颜色通道，将高斯曲线分别应用于每个通道，并将结果相乘，得到曝光度指标λ，如下：

$$\lambda=B_{\mathrm{R}}^{\mathrm{e}}B_{\mathrm{G}}^{\mathrm{e}}B_{\mathrm{B}}^{\mathrm{e}} \tag{7}$$

分别计算得到原始图像和调节后图像的权重矩阵$\boldsymbol{C}_{i,j,l}$、$\boldsymbol{S}_{i,j,l}$、$\boldsymbol{E}_{i,j,l}$。将每张图像所得的权重矩阵相乘，得到每张图像的融合权重矩阵$\boldsymbol{J}_{i,j,l}$，如下：

$$\boldsymbol{J}_{i,j,l}=\boldsymbol{C}_{i,j,l}\boldsymbol{S}_{i,j,l}\boldsymbol{E}_{i,j,l} \tag{8}$$

式中：i、j 表示像素(i,j)，l 表示第 l 幅图像(本文中 $l=2$，即原始图像和亮度调节后图像)。

如图 1 所示，为最大化保留图像信息，将高斯滤波作用于每张图像的融合权重，将拉普拉斯滤波

作用于每张图像。然后，将高斯滤波与拉普拉斯滤波的结果逐点相乘。原始图像和调节后图像两者相加得到输出图像$\boldsymbol{N}\{\boldsymbol{r}\}_{ij}$，如下：

$$\boldsymbol{N}\{\boldsymbol{r}\}_{ij}=255\sum_{l}\boldsymbol{G}\{\boldsymbol{J}\}_{ij,l}\boldsymbol{\psi}\{\boldsymbol{I}\}_{ij,l} \tag{9}$$

式中：$\boldsymbol{G}\{\boldsymbol{J}\}_{ij,l}$为高斯滤波；$\boldsymbol{\psi}\{\boldsymbol{I}\}_{ij,l}$为拉普拉斯滤波。

图3为驾驶舱图像亮度融合效果。图4上部分为图像融合过程。



(a) 不同亮度序列

(b) 融合效果

图3 不同亮度图像融合的可视化效果

Fig. 3 Visualization of different brightness image fusion

1.2 飞行员肢体关键点检测模型

1.2.1 HRNet-Former

鉴于HRNet在人体关键点检测任务中展示的出色效果，本文提出HRNet-Former模块，遵循HRNet[20]的设计，共分为4个阶段。每个阶段从一个高分辨率主干开始，逐步从高分辨率到低分辨率增加分支。每个阶段从上到下各层分支所含特征信息的分辨率逐步降低，特征信息通道数成2倍上升。通过多尺度融合模块重复交换不同分辨率的信息。如图4所示，利用图像亮度调节模块对输入图像进行亮度调节，并将模块的输出$\boldsymbol{T}\in\mathbf{R}^{3\times H\times W}$（$H$为图像高度，$W$为图像宽度）作为HRNet-Former模块的输入。

1.2.2 原始注意力模块和次序交换注意力模块

各阶段的结构具体描述如表1所示，在高分辨率分支中使用次序交换注意力模块，在低分辨率分支中使用原始注意力模块，具体细节如下。

与局部卷积的CNN结构相比，自我注意机制具有全局特征交互的特点。为准备Transformer的输入，将第1阶段输出$\boldsymbol{Z}\in\mathbf{R}^{U\times H\times W}$转换为二维序列$\boldsymbol{X}\in\mathbf{R}^{L\times D}$（$L=H\times W$；$U$为通道数；$D$为特征维度，设置与通道数相同），如下：

$$\boldsymbol{X}=\mathrm{Img2Sqe}(\boldsymbol{\Phi}) \tag{10}$$

式中：Img2Sqe(·)表示图像特征转换序列特征操作；$\boldsymbol{\Phi}$为经过第1阶段卷积主干网络处理后的输出

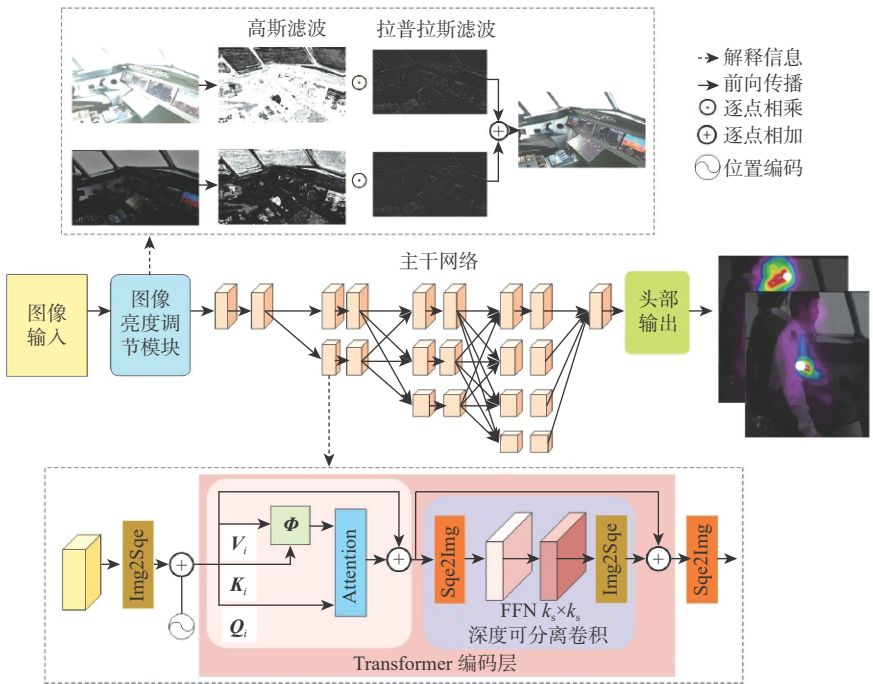


图4 面向复杂光照环境飞行员肢体关键点检测整体架构

Fig. 4 Overall architecture of pilot limb key points detection for complex lighting environment

表 1　HRNet-Former 主干结构

Table 1　HRNet-Former backbone structure

尺寸	阶段1	阶段2	阶段3	阶段4
64 × 48	$\begin{bmatrix}1\times1,64\\3\times3,64\\1\times1,\ 256\end{bmatrix}\times2$	$\begin{bmatrix}\text{次序交换注意力模块},32\\\text{头=1},\text{深度线性模块}=7\times7\end{bmatrix}\times2$	$\begin{bmatrix}\text{次序交换注意力模块},32\\\text{头=1},\text{深度线性模块}=7\times7\end{bmatrix}\times2$	$\begin{bmatrix}\text{次序交换注意力模块},32\\\text{头=1},\text{深度线性模块}=7\times7\end{bmatrix}\times2$
32 × 24		$\begin{bmatrix}\text{次序交换注意力模块},64\\\text{头=1},\text{深度线性模块}=7\times7\end{bmatrix}\times2$	$\begin{bmatrix}\text{次序交换注意力模块},64\\\text{头=1},\text{深度线性模块}=7\times7\end{bmatrix}\times2$	$\begin{bmatrix}\text{次序交换注意力模块},64\\\text{头=1},\text{深度线性模块}=7\times7\end{bmatrix}\times2$
16 × 12			$\begin{bmatrix}\text{次序交换注意力模块},128\\\text{头=2},\text{深度线性模块}=5\times5\end{bmatrix}\times2$	$\begin{bmatrix}\text{次序交换注意力模块},128\\\text{头=2},\text{深度线性模块}=5\times5\end{bmatrix}\times2$
8 × 6				$\begin{bmatrix}\text{原始注意力模块},256\\\text{头=4}\end{bmatrix}\times2$

特征图。

为保留每个特征的空间位置信息，将位置信息 $\boldsymbol{P}\in\mathbf{R}^{L\times D}$ 嵌入到二维序列 $\boldsymbol{X}^{\mathrm{I}}$ 中。该序列经过线性变换得到查询向量 $\boldsymbol{Q}\in\mathbf{R}^{L\times D}$、键向量 $\boldsymbol{K}\in\mathbf{R}^{L\times D}$ 及值向量 $\boldsymbol{V}\in\mathbf{R}^{L\times D}$。原始注意力计算 $\boldsymbol{Q}$ 和 $\boldsymbol{V}$ 之间成对相似度，导致复杂度为 $o(L^2D)$。原始注意力公式如下：

$$\mathrm{Attention}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V})=\mathrm{Softmax}\left(\boldsymbol{Q}\boldsymbol{K}^{\mathrm{T}}/\sqrt{d_{\mathrm{k}}}\right)\boldsymbol{V}\tag{11}$$

式中：Softmax(·)表示激活注意窗口操作；$\sqrt{d_{\mathrm{k}}}$用于缩放点乘运算结果。

考虑到在高分辨率分支中，特征维度 D 远小于 L，在高分辨率分支中运用次序交换注意力。值得注意的是，在低分辨率分支运用原始注意力，各个分支具体描述如表 1 所示。

如图 5 所示，次序交换注意力设法通过适当的交换操作，通过先计算 $\boldsymbol{K}^{\mathrm{T}}\boldsymbol{V}$ 来改变模型复杂度，从而导致复杂度为 $o(L\,D^2)$。然后，进行单头次序交换注意力运算以获取序列中全局交互关系，如下：

$$\mathrm{Attention}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V})=\boldsymbol{Q}\cdot\mathrm{Softmax}\left(\boldsymbol{K}^{\mathrm{T}}\boldsymbol{V}/\sqrt{d_{\mathrm{k}}}\right)\tag{12}$$

利用多个单头注意力与权重矩阵 $\boldsymbol{\Gamma}^{\mathrm{O}}$ 进行多头注意力运算，如下：

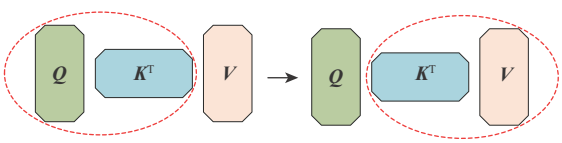
$$\mathrm{MultiAtt}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V})=\mathrm{Concat}(q_1,q_2,\cdots,q_t)\boldsymbol{\Gamma}^{\mathrm{O}}\tag{13}$$

式中：$q_1,q_2,\cdots,q_t$ 为第 1, 2, ⋯, t 个注意力头的输出；Concat(·)表示拼接操作，将全部 t 个注意力头的输出在特征维度 D 上进行拼接，形成特征矩阵。

经过多头注意力计算的特征与序列 $\boldsymbol{X}^{\mathrm{I}}\in\mathbf{R}^{L\times D}$ 进行融合，得到 $\boldsymbol{\varpi}\in\mathbf{R}^{L\times D}$，如下：

$$\boldsymbol{\varpi}=\boldsymbol{X}^{\mathrm{I}}+\mathrm{MSA}(\boldsymbol{X}^{\mathrm{I}})\tag{14}$$

式中：MSA(·)表示多头注意力计算操作。



(a) 原始注意力　　(b) 次序交换注意力

图 5　次序交换注意力示意图

Fig. 5　Schematic diagram of order-switched attention

然而，次序交换注意力将 Softmax 注意窗口由 $L\times L$ 固定为 $D\times D$ 尺寸，减少了局部特征细节。因此，为增强次序交换注意力模块的局部感知能力，本文提出 FFN-DW（深度线性）模块，通过在 FFN 层中额外插入深度卷积来增强局部注意。图 6 为 FFN-DW 模块的结构，该模块先使用 1×1 卷积，再使用卷积核尺寸（$k_{\mathrm{s}}\times k_{\mathrm{s}}$）深度卷积，最后，使用 1×1 卷积。

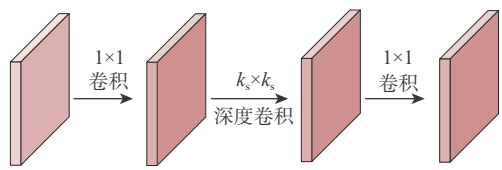


图 6　FFN-DW 模块结构

Fig. 6　Structure of FFN-DW module

融合信息 $\boldsymbol{\varpi}$ 进入 FFN-DW 模块完成特征信息交互得到 $\boldsymbol{Y}\in\mathbf{R}^{L\times D}$，如下：

$$\boldsymbol{Y}=\mathrm{Img2Sqe}(\mathrm{FFN\text{-}DW}(\mathrm{Seq2Img}(\boldsymbol{\varpi})))\tag{15}$$

$$\boldsymbol{O}_{\mathrm{f}}=\mathrm{Sqe2Img}(\boldsymbol{\varpi}+\boldsymbol{Y})\tag{16}$$

式中：Sqe2Img(·) 表示序列转换图形操作。

对于下一阶段，序列 $\boldsymbol{Y}$ 经过 Sqe2Img(·) 操作得到下一个图像特征 $\boldsymbol{O}_{\mathrm{f}}\in\mathbf{R}^{D\times H\times W}$，如式(16)所示。最后，Transformer 编码层特征经过 1 × 1 卷积变换，该输出维度与关键点个数一致，生成关键点热图。

1.3　飞行员手部关键点检测模型

本文提出一种手部关键点检测试验网络，主要由手掌检测模块和手部关键点检测模块组成。手掌检测器由 TPH-YOLOv5[31] 训练得到，手部关键点检测模型为 1.2 节飞行员肢体关键点检测模型。

2　飞行员关键点检测实验验证

2.1　飞行员肢体关键点检测数据集

在驾驶舱模拟平台上采集关键点检测数据集。根据操作标准，在文献 [32] 的基础上，共采集

6 800张飞行员姿态图像，包括5 200张训练集、800张验证集和800张测试集。图像数据经过预处理，高度为640像素，宽度为480像素。飞行员肢体标注17个关键点，分别为鼻子、右眼、左眼、右耳、左耳、右肩、左肩、右肘、左肘、右腕、左腕、右臀、左臀、右臀膝中部、左臀膝中部、右膝盖和左膝盖。

2.1.1 评估标准

使用物体关键点相似度(object keypoint similarity, OKS)[33]评估方式为飞行员肢体关键点评估标准。包括：OKS为0.50,0.55,…,0.90,0.95时10个阈值点的平均准确率(average precision, AP)，OKS为0.50, 0.55,…,0.90,0.95时10个阈值点的平均召回率(average recall, AR)。该评估标准方法如下：

$$a_{\mathrm{OKS}}=\sum_{p}\frac{\exp(-\kappa_p^2/(2s^2\kappa_p^2))\delta(v_p>0)}{\sum_{p}\delta(v_p>0)} \tag{17}$$

式中：a_{OKS}为物体关键点相似度；p为飞行员肢体关键节点的类型；Ω_p为人工标注坐标与预测坐标之间的欧氏距离；v_p为飞行员肢体关键节点的可见信息；s为关键节点尺度；κ_p为每个关键节点相关控制衰减常数；δ为布尔指示函数，当括号内的条件成立时，函数值为1，否则为0。

2.1.2 实验环境配置

实验训练环境为：Ubuntu 20.04 LST 64位系统，2张英伟达3090显卡，Pytorch 1.8.1学习框架。在飞行员肢体关键点检测数据集上训练时，将图像裁剪为256像素×192像素大小。采用AdamW优化器。实验初始学习率为1×10^{-3}，第140周期时学习率为1×10^{-4}，第170周期时学习率衰减为1×10^{-5}，模型共训练210周期。每个GPU批量大小为32。使用图像旋转和水平翻转等方式进行数据样本增强。

2.1.3 结果分析

为评估本文模型的有效性，在飞行员肢体关键点检测数据集上，与几个基准模型进行比较。表2为比较结果(未使用图像亮度调节模块)，包括关键点检测网络参数量和浮点运算速度GFLOPs。

表2 不同模型在飞行员肢体关键点检测数据集上的性能比较

Table 2 Performance comparison of different models on the pilot limb key point detection dataset

模型	主干网络	输入尺寸/(像素×像素)	参数量	浮点运算速度/(10^9 s^{-1})	AP/%	AR/%
SimpleBaseline[34]	ResNet-50	256×192	34.0×10^6	8.9	77.9	78.5
HRNet[20]	HRNet-W32	256×192	28.5×10^6	7.1	80.8	81.6
TransPose[24]	HRNet-W32	256×192	8.0×10^6	10.2	81.3	83.2
PoseUR[35]	HRFormer-B	256×192	28.8×10^6	12.6	82.1	83.8
HRFormer[22]	HRFormer-S	256×192	2.5×10^6	1.3	80.6	82.2
本文	HRNet-Former	256×192	4.5×10^6	3.8	81.9	83.3

本文模型在图像输入尺寸为256像素×192像素的情况下进行训练，AP为81.9%，与SimpleBaseline相比，高出4%。此外，本文方法的AP比HRNet高1.1%，参数量减少24.0×10^6。与TransPose相比，本文方法参数量减少3.5×10^6，AP高0.6%。然而，与PoseUR相比，本文方法AP低0.2%，但参数量却显著减少24.3×10^6。总体而言，本文方法在参数数量与准确性之间取得平衡，为飞行员行为识别提供实时基础。

此外，如表3所示，在单张英伟达3090显卡上统计推理速度(未使用图像亮度调节模块)。由表3可知，相较于HRFormer，虽然本文模型的参数量和复杂度高于HRFormer，但本文模型推理速度较快。这可能与HRFormer本身的移位机制有关，该机制带来计算成本的损耗。同时也表明本文模型的现实工程意义。

表3 不同模型推理速度比较

Table 3 Comparison of the inference speed of reasoning for different models

模型	参数量	浮点运算速度/(10^9 s^{-1})	推理速度/(帧·min^{-1})
MobileNetV2[36]	9.6×10^6	1.97	53.4
HRNet[20]	7.6×10^6	1.70	28.6
HRFormer[22]	2.5×10^6	1.30	33.1
本文	4.5×10^6	3.80	44.3

2.2 MS COCO关键点检测数据集

MS COCO关键点检测数据集有超过20万张图像、25万个人体实例和17个关键点。

2.2.1 评估标准

实验在MS COCO数据集上采用OKS评估标准，如式(17)所示。

2.2.2 实验环境设置

实验使用的训练环境如下：Ubuntu 20.04 LST 64位系统，4张英伟达3090显卡，Pytorch 1.8.1学习框架。在MS COCO关键点检测数据集上训练时，将图像裁剪为256像素×192像素大小。采用AdamW优化器。实验初始学习率为1×10^{-3}，第160周期时学

习率为 1×10^{-4}, 第 210 周期时学习率衰减为 1×10^{-5}, 模型共训练 230 周期。每个 GPU 批量大小为 32。

2.2.3　结果分析

模型在 MS COCO val2017 数据集上进行验证。表 4 为本文模型和其他模型的比较结果。当输入图像尺寸为 256 像素×192 像素时, 本文模型的 AP 为 72.8%(为公平比较, 以上结果未使用图像亮度调节模块), 与 CPN 模型相比, 提高 4.2%。相较于 SimpleBaseline 模型, 本文模型参数量为其 13.2%, 浮点运算速度为其 42.7%, AP 提高 2.4%; 与 HRNet 模型相比, 本文模型参数量下降 84.2%, 浮点运算速度下降 46.5%, 然而, AP 仅降低 0.6%。

表 4　不同模型在 MS COCO val2017 数据集的性能比较

Table 4　Performance comparison of the different models on MS COCO val2017 dataset

模型类型	模型	输入尺寸/(像素×像素)	参数量	浮点运算速度/(10^9 s^{-1})	AP/%	AR/%
复杂网络模型	CPN[19]	256×192	27.0×10^6	6.2	68.6	
	SimpleBaseline[34]	256×192	34.0×10^6	8.9	70.4	76.3
	HRNet[20]	256×192	28.5×10^6	7.1	73.4	78.9
	DAEK[21]	128×96	63.6×10^6	3.6	71.9	77.9
	TransPose[24]	256×192	8.0×10^6	10.2	74.2	78.0
	PoseUR[35]	256×192	28.8×10^6	4.48	74.7	
轻量网络模型	MobileNetV2[36]	256×192	9.6×10^6	1.48	64.6	70.7
	ShuffleNetV2[37]	256×192	7.6×10^6	1.28	59.9	66.4
	Lite-HRNet-18[38]	256×192	1.1×10^6	0.2	64.8	71.2
	HRFormer[22]	256×192	2.5×10^6	1.3	70.9	76.6
	本文	256×192	4.5×10^6	3.8	72.8	78.4

在轻量网络模型上, 相较于 MobileNetV2 和 ShuffleNetV2, 本文模型的参数量分别减少 5.1×10^6 和 3.1×10^6, AP 分别提升 8.2% 和 12.9%。相较于 HRFormer 模型, 参数量增加 2.0×10^6, AP 提升 2.7%。

2.3　FreiHAND 手部关键点数据集

足够数量的关键点定位是识别复杂动作的基础, 本文使用关键点检测模型识别飞行员手部信息。在手部关键点数据集方面, 本文采用 FreiHAND[39] 数据集, 其中包含 130 240 个训练样本和 3 960 个验证样本。每个训练样本提供 RGB 图像 (224 像素 × 224 像素) 和 21 个手部关键点的位置信息。

2.3.1　评估标准

使用标准度量正确关键点的比例 (percentage of correct keypoints, PCK)评估模型手部关键点定位性能, 如下:

$$p_{\mathrm{CK}}=\sum_b \delta(h_{bn}\leqslant e)\Big/\sum_b 1 \tag{18}$$

式中: b 为第 b 个手掌; n 为手掌上第 n 个关键点; h_{bn} 为第 b 个手掌的第 n 个关键点预测值与标注值的欧式距离。若式 (18) 中括号内条件成立, 则 δ 为 1, 反之为 0。

2.3.2　结果分析

在 FreiHAND 数据集上进行实验验证, 表 5 为本文方法和其他方法的比较结果(未使用图像亮度调节模块)。

表 5　不同模型在 FreiHAND 数据集上的性能比较

Table 5　Comparison of the different models on FreiHAN dataset

模型	参数量	PCK/%	推理速度/(帧·min^{-1})
MobileNetV2[36]	9.6×10^6	81.48	49.5
MobileNetV3[36]	8.7×10^6	83.61	41.1
ShuffleNetV2[37]	7.6×10^6	82.86	45.3
SRHandNet[40]	16.3×10^6	95.43	28.4
本文	4.8×10^6	98.84	38.6

在实时性上, 本文模型低于 MobileNetV2、MobileNetV3 和 ShuffleNetV2, 然而, 其 PCK 相对更高。与网络模型 SRHandNet 相比, 本文模型的 PCK 提升 3.41%。

2.4　联合部署肢体和手部关键点检测网络

本文在 1 张图像上完成飞行员肢体和手部关键点检测模型的联合部署。该联合方法共提取飞行员 59 个身体关键节点。其中, 手部关键节点数为 42 个(每只手 21 个关键节点), 肢体关键节点数为 17 个。图 7 为肢体和手部关键点联合检测的可视化结果。如图 7 所示, 选取典型的飞行动作进行实验验证, 包括点击触摸屏、推拉油门杆、打开起落架和打开自动导航按钮。联合部署的模型显示良好的实时检测能力, 为飞行员行为识别提供基础。表 6 为本文联合部署模型和其他方法的比较结果, 其中, 联合模型来自表 2 和表 5。相较于表

中其他模型，本文方法在参数量较小的条件下，实现了先进的实时性能。图8为本文方法和部分联合部署模型在一块英伟达3090显卡上的实时推理效果。



(a) 点击触摸屏　(b) 推拉油门杆

(c) 打开起落架　(d) 打开自动导航按钮

图7　飞行员肢体和手部关键点联合检测的可视化效果

Fig. 7　Visualisation of the joint detection of key points of the pilot's limb and hands

表6　联合模型的实时性比较

Table 6　Real-time comparison for joint models

联合模型(肢体+手部)	参数量	推理速度/(帧·min^{-1})
HRNet+MobileNetV2	38.1×10^6	7.6
HRFormer+MobileNetV2	12.1×10^6	11.2
HRNet+ShuffleNetV2	36.1×10^6	7.4
HRFormer+ShuffleNetV2	10.1×10^6	11.1
HRNetV1+SRHandNet	44.8×10^6	5.4
HRFormer+SRHandNet	18.8×10^6	8.8
本文	9.7×10^6	14.3

2.5　消融实验

2.5.1　图像亮度调节模块消融实验

为检验图像亮度调节模块对飞行员关键点检测性能的影响，使用2组不同亮度水平的图像进行实验。对所提出模型的输出进行概率激活，形成关键点热图。在进行图像亮度减弱和增强时，图9(a)和图9(b)分别为模型预测的可视化效果。色条从上到下的颜色变化对应不同的置信度，红色区域表示最高置信度，白色圆圈为人工标注位置。图中水平方向为飞行员肢体关键点检测相对应的关键点类型。根据关键点估计热图，可以定性地看出图像亮度调节模块对模型识别能力的增益。

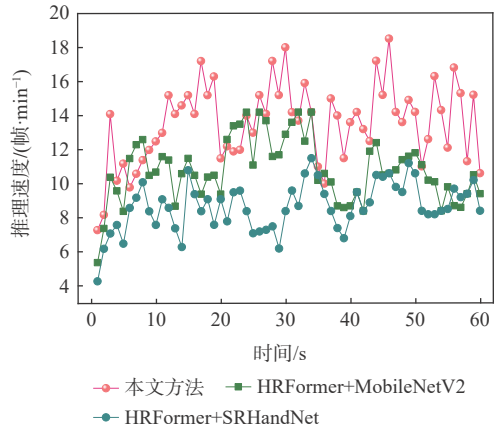


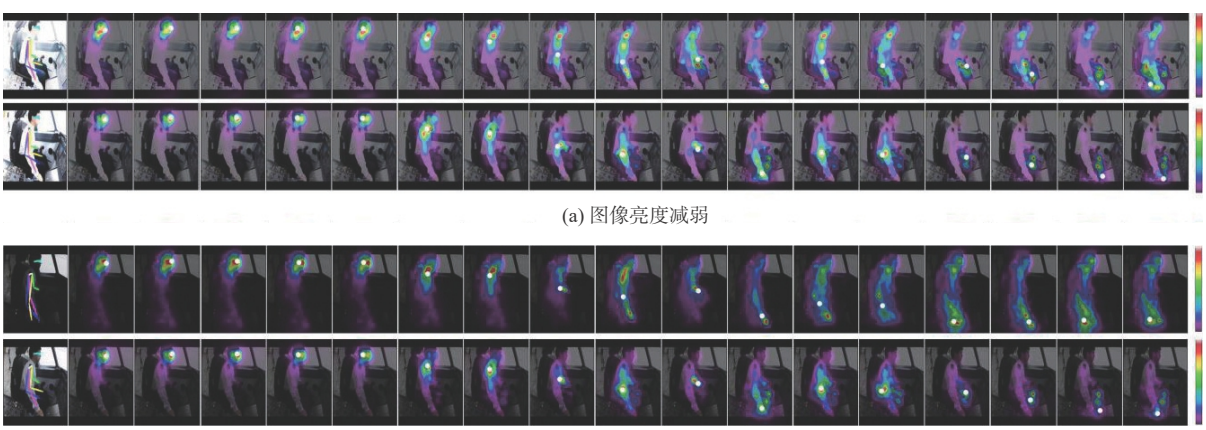
图8　联合模型的实时性效果

Fig. 8　The real-time effect of the joint models

此外，为进一步量化图像亮度调节模块对模型识别性能的影响，建立800幅图像，包括曝光和低亮度，其中每组各400张。使用训练完成的模型进行关键点检测。从表7可以看出，图像亮度调节模块对模型的识别性能有显著影响。对曝光和低亮度图像进行调整后，模型预测AP分别提高10.4%和7.8%。由此可以得出，图像亮度调节模块对于提高复杂光照条件下的关键点定位精度具有积极意义。

2.5.2　次序交换注意力模块消融实验

为研究Transformer模块中次序交换注意力模块对模型性能的影响，使用3种模型在飞行员肢体



(a) 图像亮度减弱

(b) 图像亮度增强

图9　图像亮度调节模块下的飞行员肢体关键点输出比较

Fig. 9　Comparison of pilot limb key point outputs under the image brightness adjustment module

表 7　图像亮度调节前后模型性能比较

Table 7　Comparison of model performance before and after image brightness adjustment

图像亮度调节	输入尺寸/(像素×像素)	AP/%	推理速度/(帧·min^{-1})
原始图像(曝光)	256×192	74.2	25.1
经过图像调节模块(曝光)	256×192	84.6	25.1
原始图像(弱光)	256×192	75.3	26.3
经过图像调节模块(弱光)	256×192	83.1	26.3

表 8　不同模型在飞行员肢体关键点数据集上的性能比较

Table 8　Performance comparison of different models on the pilot's limb key point dataset

模型	参数量	浮点运算速度/(10^9 s^{-1})	AP/%	推理速度/(帧·min^{-1})
原始注意力模块	5.0×10^6	8.56	82.2	21.3
次序交换(无FFN-DW层)	4.1×10^6	3.34	74.4	39.7
本文	4.5×10^6	3.80	82.3	35.4

关键点数据集进行验证。3 种模型分别为原始注意力模块、次序交换(无 FFN-DW 层)、本文模型。

当输入图像尺寸为 256 像素×192 像素时，表 8 为不同模型在飞行员肢体关键点数据集上的性能比较。相较于原始注意力模块，本文模型可以近似达到其识别能力，同时具有较低的复杂度和较高的推理速度。相较于次序交换(无 FFN-DW 层)，在 FFN 层中插入深度卷积可以感知更多的局部信息。图 10 为不同模型在飞行员关键点定位的可视化结果，该顺序依次为鼻子、右眼、左眼、右耳、左耳、右肩、左肩。

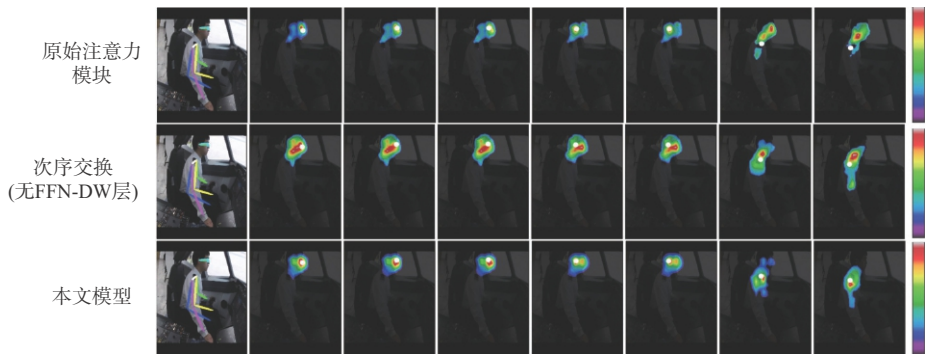


图 10　不同模型在飞行员关键点定位的可视化结果

Fig. 10　Visualisation of pilot key point detection results with different models

3　结　论

针对复杂光照下民机飞行员关键点检测问题，提出一种融合图像亮度自适应调整与轻量化关键点检测网络的联合方法，通过系统实验验证，得出以下结论：

1) 图像亮度调节模块有效提升图像质量和模型鲁棒性。所提图像亮度调节模块通过自适应伽马变换和多曝光融合策略，有效缓解过曝、欠曝及噪声问题，在低亮度和曝光图像上的 AP 分别提升 7.8% 和 10.4%。

2) 轻量化 HRNet-Former 网络兼顾检测精度和推理效率。构建的 HRNet-Former 网络引入次序交换注意力机制和深度可分离卷积，在参数量仅为 4.5×10^6 时，在自建数据集和 MS COCO val2017 数据集上的 AP 分别达到 81.9% 和 72.8%，兼顾精度和实时性(44.3 帧/min)。

3) 联合检测网络支持复杂飞行行为识别。肢体与手部关键点检测网络的联合部署实现了 59 个关键点的同步检测，参数量为 9.7×10^6、推理速度为 14.3 帧/min，为复杂行为识别提供有效支持。

本文为智能驾驶舱中飞行员行为监控提供了有效的技术方案，未来将探索多模态信息融合，以进一步提升系统实用性和可靠性。

参考文献（References）

[1] 杨志刚, 张炯, 李博, 等. 民用飞机智能飞行技术综述[J]. 航空学报, 2021, 42(4): 525198.
YANG Z G, ZHANG J, LI B, et al. Reviews on intelligent flight technology of civil aircraft[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(4): 525198(in Chinese).

[2] LIANG B H, CHEN Y, WU H. A conception of flight test mode for future intelligent cockpit[C]//Proceedings of the Chinese Automation Congress. Piscataway: IEEE Press, 2020: 3260-3264.

[3] JIN Z B, LI D C, XIANG J W. Robot pilot: a new autonomous system toward flying manned aerial vehicles[J]. Engineering, 2023, 27: 242-253.

[4] KELLY D, EFTHYMIOU M. An analysis of human factors in fifty controlled flight into terrain aviation accidents from 2007 to 2017 [J]. Journal of Safety Research, 2019, 69: 155-165.

[5] DE SANT'ANNA D A L M, DE HILAL A V G. The impact of human factors on pilots' safety behavior in offshore aviation companies: a Brazilian case[J]. Safety Science, 2021, 140: 105272.

[6] WANG A G, CHEN H H, ZHENG C D, et al. Evaluation of random forest for complex human activity recognition using wearable

sensors[C]//Proceedings of the International Conference on Networking and Network Applications. Piscataway: IEEE Press, 2020: 310-315.

[7] YAN S J, XIONG Y J, LIN D H. Spatial temporal graph convolutional networks for skeleton-based action recognition[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2018, 32(1): 7444-7452.

[8] KIM Y T. Contrast enhancement using brightness preserving bi-histogram equalization[J]. IEEE Transactions on Consumer Electronics, 1997, 43(1): 1-8.

[9] LAND E H, MCCANN J J. Lightness and retinex theory[J]. Journal of the Optical Society of America, 1971, 61(1): 1.

[10] SHEN L, YUE Z H, FENG F, et al. MSR-net: low-light image enhancement using deep convolutional network[EB/OL]. (2017-11-07)[2023-09-01]. https://arxiv.org/abs/1711.02488.

[11] WEI C, WANG W, YANG W, et al. Deep retinex decomposition for low-light enhancement[EB/OL]. (2018-08-14)[2023-09-01]. https://arxiv.org/abs/1808.04560.

[12] YASARLA R, SINDAGI V A, PATEL V M. Semi-supervised image deraining using Gaussian processes[J]. IEEE Transactions on Image Processing, 2021, 30: 6570-6582.

[13] LV F F, LIU B, LU F. Fast enhancement for non-uniform illumination images using light-weight CNNs[C]//Proceedings of the 28th ACM International Conference on Multimedia. New York: ACM, 2020: 1450-1458.

[14] KANSAL S, PURWAR S, TRIPATHI R K. Image contrast enhancement using unsharp masking and histogram equalization[J]. Multimedia Tools and Applications, 2018, 77(20): 26919-26938.

[15] RAO B S. Dynamic histogram equalization for contrast enhancement for digital images[J]. Applied Soft Computing, 2020, 89: 106114.

[16] YUAN L, SUN J. Automatic exposure correction of consumer photographs[C]//Proceedings of the Computer Vision-ECCV. Berlin: Springer, 2012: 771-785.

[17] ZHANG J A, LIU R S, MA L, et al. Principle-inspired multi-scale aggregation network for extremely low-light image enhancement[C]//Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway: IEEE Press, 2020: 2638-2642.

[18] 孔英会, 秦胤峰, 张珂. 深度学习二维人体姿态估计方法综述[J]. 中国图象图形学报, 2023, 28(7): 1965-1989.
KONG Y H, QIN Y F, ZHANG K. Deep learning based two-dimension human pose estimation: a critical analysis[J]. Journal of Image and Graphics, 2023, 28(7): 1965-1989(in Chinese).

[19] CHEN Y L, WANG Z C, PENG Y X, et al. Cascaded pyramid network for multi-person pose estimation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 7103-7112.

[20] SUN K, XIAO B, LIU D, et al. Deep high-resolution representation learning for human pose estimation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2019: 5686-5696.

[21] BULAT A, KOSSAIFI J, TZIMIROPOULOS G, et al. Toward fast and accurate human pose estimation via soft-gated skip connections[C]//Proceedings of the 15th IEEE International Conference on Automatic Face and Gesture Recognition. Piscataway: IEEE Press, 2020: 8-15.

[22] YUAN Y H, FU R, HUANG L, et al. HRFormer: high-resolution Transformer for dense prediction[EB/OL]. (2021-11-07)[2023-09-01]. https://arxiv.org/abs/2110. 09408.

[23] XIONG Z N, WANG C X, LI Y, et al. Swin-Pose: swin Transformer based human pose estimation[C]//Proceedings of the IEEE 5th International Conference on Multimedia Information Processing and Retrieval. Piscataway: IEEE Press, 2022: 228-233.

[24] YANG S, QUAN Z B, NIE M, et al. TransPose: keypoint localization via Transformer[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2021: 11782-11792.

[25] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16×16 words: Transformers for image recognition at scale[EB/OL]. (2021-06-03)[2023-09-01]. https://arxiv.org/abs/2010. 11929.

[26] LIN K, WANG L J, LIU Z C. End-to-end human pose and mesh reconstruction with Transformers[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2021: 1954-1963.

[27] 刘豪, 吴红兰, 房宇轩. 结合全局上下文信息的高效人体姿态估计[J]. 计算机工程, 2023, 49(7): 102-109.
LIU H, WU H L, FANG Y X. Efficient human pose estimation combining global contextual information[J]. Computer Engineering, 2023, 49(7): 102-109(in Chinese).

[28] CAI H, LI J Y, HU M Y, et al. EfficientViT: multi-scale linear attention for high-resolution dense prediction[EB/OL]. (2024-02-06)[2023-09-01]. https://arxiv.org/abs/2205.14756.

[29] XU W J, XU Y F, CHANG T, et al. Co-scale conv-attentional image Transformers[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2021: 9961-9970.

[30] OGDEN J M, ADELSON E H, BERGEN J R, et al. Pyramid-based computer graphics[J]. RCA Engineer, 1985, 30(5): 4-15.

[31] ZHU X K, LYU S C, WANG X, et al. TPH-YOLOv5: improved YOLOv5 based on Transformer prediction head for object detection on drone-captured scenarios[C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2021: 2778–2788.

[32] 吴红兰, 刘豪, 孙有朝. 基于视觉 Transformer 飞行员姿态估计[J]. 北京航空航天大学学报, 2024, 50(10): 3100-3110.
WU H L, LIU H, SUN Y C. Vision Transformer-based pilot pose estimation[J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50(10): 3100-3110(in Chinese).

[33] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft COCO: common objects in context[C]//Proceedings of the Computer Vision-ECCV. Berlin: Springer, 2014: 740-755.

[34] XIAO B, WU H P, WEI Y C. Simple baselines for human pose estimation and tracking[C]//Proceedings of the Computer Vision-ECCV. Berlin: Springer, 2018: 472-487.

[35] MAO W A, GE Y T, SHEN C H, et al. PoseUR: direct human pose regression with Transformers[C]//Proceedings of the Computer Vision-ECCV. Berlin: Springer, 2022: 72-88.

[36] SANDLER M, HOWARD A, ZHU M L, et al. MobileNetV2: inverted residuals and linear bottlenecks[C]//Proceedings of the IEEE/

CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 4510-4520.

[37] MA N N, ZHANG X Y, ZHENG H T, et al. ShuffleNet V2: practical guidelines for efficient CNN architecture design[C]//Proceedings of the Computer Vision-ECCV. Berlin: Springer, 2018: 122-138.

[38] YU C Q, XIAO B, GAO C X, et al. Lite-HRNet: a lightweight high-resolution network[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2021: 10435-10445.

[39] ZIMMERMANN C, CEYLAN D, YANG J M, et al. FreiHAND: a dataset for markerless capture of hand pose and shape from single RGB images[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2019: 813-822.

[40] WANG Y G, ZHANG B W, PENG C. SRHandNet: real-time 2D hand pose estimation with simultaneous region localization[J]. IEEE Transactions on Image Processing, 2020, 29: 2977-2986.

Key points detection method for civil aircraft pilot in complex lighting environments

LIU Hao，SUN Youchao*，WU Honglan，ZHANG Haiyan

(College of Civil Aviation，Nanjing University of Aeronautics and Astronautics，Nanjing 211106，China)

Abstract：The recognition and monitoring of pilot maneuvering behaviors in civil aircraft based on computer vision is of great practical significance to ensure the safety of civil aviation operations. In this paper, a pilot key point detection model in complex lighting environments is proposed. Firstly, considering the complex lighting environment in the cockpit of civil aircraft, an image brightness adjustment module is proposed. This module increases the retention of image detail features while simultaneously increasing operation speed by hierarchically determining the average value of image brightness and achieving the fusion of image features of varying brightness. Second, given that a higher number of key point localizations are fundamental to accurate behavior recognition, a lightweight pilot limb key point, and hand key point detection network is proposed. The network employs a sequential exchange of attention modules in the high-resolution branch to alleviate the problem of quadratic growth of the computational cost of raw vision attention with increasing input resolution. In addition, the pilot limb and hand key points detection networks are jointly deployed and typical flight maneuvers are selected for experimental validation. In order to establish an interpretable relationship between the methodology and the prediction results, comprehensive ablation experiments are finally carried out to investigate both quantitatively and qualitatively the effects of various components (image brightness adjustment module, order exchange attention module) on the model performance. The proposed model achieves an AP of 81.9% on the pilot limb key point dataset and 72.8% on the MS COCO val 2017 dataset, balancing accuracy and real-time performance.

Keywords：civil aircraft; intelligent cockpit; complex lighting; pilot; key points detection; vision attention

Received：2023-09-04；**Accepted**：2023-10-13；**Published Online**：2023-11-22 09：44

URL：link.cnki.net/urlid/11.2625.V.20231121.1058.002

Foundation items：Joint Fund of National Natural Science Foundation of China and Civil Aviation Administration of China (U2033202,U1333119); National Natural Science Foundation of China (52172387)

* **Corresponding author**. E-mail：sunyc@nuaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0520

低压双流体细水雾削减三元锂离子电池热失控风险性能

曾祥淑，梁家鑫，孙强，贺元骅*
(中国民用航空飞行学院 民机火灾科学与安全工程四川省重点实验室，广汉 618307)

摘 要：为削减气体、高压细水雾等灭火剂抑灭锂离子电池火灾导致二次灾害的短板问题，研究低压双流体细水雾抑灭锂离子电池燃爆技术。以中小型储能系统常用的三元锂离子电池为对象，研究了低压双流体细水雾对锂离子电池在热失控前不同温度下(90 ℃、145 ℃和200 ℃)的外观形貌变化、产热过程抑制和电性能防护的影响。结果表明：在施加低压双流体细水雾情况下，锂离子电池极耳附近鼓包厚度分别减小了0 cm、0.01 cm和0.15 cm，电池直接受热大面的受损状况有所减小；电池非直接受热大面温度降低幅度很大，分别降低了38 ℃、43 ℃和31 ℃；电池充电容量分别提高了487.54 mAh、4116.69 mAh和6 230.06 mAh，放电容量分别提高了565.07 mAh、4 325.11 mAh和6297.45 mAh。因此，低压双流体细水雾不仅可有效防护遭遇热失控锂离子电池的外观形貌、热量抑制和电性能，而且可有效防止锂离子电池热失控灾害传统抑灭技术导致的二次灾害。

关 键 词：低压双流体细水雾；三元锂离子电池；产热抑制；电性能防护；二次灾害
中图分类号：X932; TM911.11
文献标志码：A **文章编号**：1001-5965(2025)10-3482-10

安全问题始终是储能行业面临的一大挑战，随着储能市场快速发展，安全性问题也逐渐凸显。三元锂离子电池在容量与安全性方面比较均衡，综合性能非常优异，能量密度高，使用寿命长[1]，因此，三元锂离子电池被广泛地应用于航空航天、新能源、中小型储能系统等领域。但是，被放置于受限空间(密闭箱体)的锂离子电池在机械滥用、电滥用和热滥用等条件下，可能会发生热失控[2-5]，锂离子电池内部发生一系列的化学反应，放出大量的热量和有毒有害可燃气体，形成严重的火灾爆炸危险[6-8]。为破解高发频发的锂离子电池储能系统热灾害风险，人们开发了多种储能锂离子电池系统抑灭火技术，如《消防设施通用规范》[9]就规定有干粉、气体、水喷淋和细水雾等4种灭火技术。

Meng[10]和Wang[11]等研究了粉体和气体灭火剂抑灭锂离子电池热失控的效果，发现粉体和气体灭火剂冷却锂离子电池热失控能力差强人意，既不能抑制电池组二次灾害发生[12-14]，又可能诱发新的风险。如七氟丙烷灭火后产生的分解产物氟化氢(HF)有一定的毒性，在潮湿条件下具有腐蚀性，达到一定浓度对人体及设备会造成损害[15]。全氟己酮抑灭锂离子电池火灾必须在其空间内连续喷洒才有效果，但不能削减二次灾害且使用成本偏高[16]。胡振恺等[17]研究了水喷淋灭火系统抑灭锂离子电池火灾的效果，发现其降温效果明显，但灭火后储能系统设备损坏严重，且无法有效抑制热失控电池内部温度，存在二次复燃风险。

Larsson等[18]研究了高压细水雾抑制不同荷电

收稿日期：2023-08-08；录用日期：2023-10-13；网络出版时间：2023-11-14 08:30
网络出版地址：link.cnki.net/urlid/11.2625.V.20231110.1706.003
基金项目：国家自然科学基金(U2033206)；四川省科技计划项目(2021YFSY0001)
*通信作者. E-mail：he_yuan_hua@163.com

引用格式：曾祥淑，梁家鑫，孙强，等. 低压双流体细水雾削减三元锂离子电池热失控风险性能[J]. 北京航空航天大学学报，2025，51(10)：3482-3491. ZENG X S，LIANG J X，SUN Q，et al. Performance of low-pressure dual-fluid water mist to reduce thermal runaway risk of ternary lithium-ion battery[J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51(10)：3482-3491 (in Chinese).

状态下的锂离子电池热失控效果, 所使用的高压细水雾管道流量大, 用水量有限, 对锂离子电池热失控有一定的抑制作用, 但因喷射时间短, 降温冷却效果不显著。Zhang 等[19] 研究了高压细水雾抑灭单电池和多电池实验, 发现细水雾可以有效扑灭单电池火灾, 降低电池表面的最高温度, 但由于喷水时间较短且水量有限, 细水雾的冷却效率总体不明显。韩路豪等[20] 开展了 10 MPa 的高压细水雾抑制锂离子电池热失控实验, 研究表明, 施加细水雾时间越长, 冷却效果越好, 但发现有导致水渍污染的风险, 实验中高压细水雾仅释放 40 s, 其用水量即达到了 10 L, 如果按照法规规定的 30 min 来施加细水雾, 用水量将达到 450 L。丰汉军等[21] 自建全尺寸锂离子电池灭火实验平台, 研究了高压细水雾的灭火过程及效果, 表明高压细水雾可快速扑灭明火, 抑制热失控蔓延, 高压细水雾连续喷射约 10 min, 电池组未出现复燃现象, 但将电池组拆开后发现其遭到了严重破坏, 未产生热失控的电池也因水量较大遭到损坏。Liu 等[22] 研究了低压细水雾对控制电池热失控的有效性, 发现当细水雾在达到热失控临界温度之前开启时, 可以有效控制热失控发生。赵军超等[23] 使用低压细水雾灭火系统开展锂离子电池单体火灾实验, 结果表明, 低压细水雾灭火系统有效抑制了锂离子电池热失控早期风险, 在抑制过程中, 都未发生复燃和热失控传播的迹象, 证明其冷却效果较好。

综上所述, 惰性气体类和粉体类这些主要依赖化学抑制作为灭火机理的灭火剂, 对电池热失控的抑制效果较弱[24], 虽然在一定程度上有利于抑制电池储能系统热灾害, 但不能有效避免锂离子电池热灾害导致的二次灾害, 如 2018 年 4 月美国亚利桑那储能电站的二次燃爆; 水喷淋系统虽然能有效扑灭锂离子电池火灾, 但会导致储能电站水渍灾害、未受热电池毁损等二次灾害, 高压细水雾灭火系统建设费用偏高、运维成本偏高。可见, 开发一种既可以削减锂离子电池热灾害风险, 又可以避免热失控抑灭二次灾害, 且经济适用的锂离子电池热灾害抑灭技术显得十分重要。因此, 本文探究利用市政自来水管网低压双流体细水雾抑灭锂离子电池热灾害的技术, 以促进锂离子电池环保型、经济型抑灭火新技术发展。

1　实验设置

1.1　实验样品

样品电池及电池物性参数如图 1 和表 1 所示。三元锂离子电池宽度为65 mm, 厚度为 7.0 mm, 长

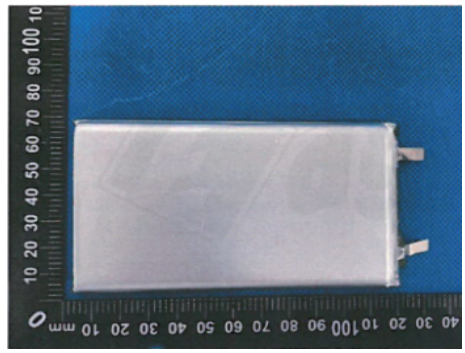


图 1　样品电池

Fig. 1　Sample battery

表 1　锂离子电池物性参数

Table 1　Physical property parameters of lithium-ion batteries

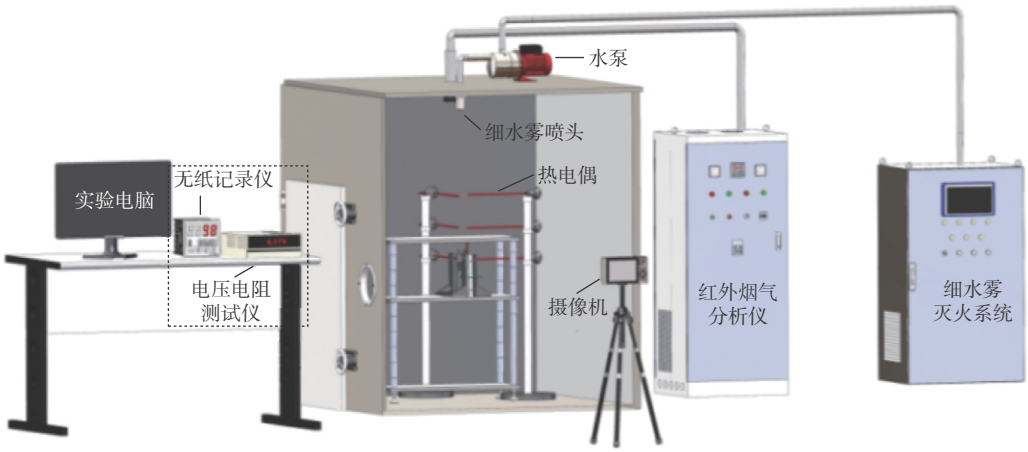
参数	状态/数值
正极材料	NCM523
负极材料	石墨
电池质量/g	151.4±1
额定容量/mAh	10 000
标定电压/V	4.2
荷电状态/%	100
电池尺寸/(mm×mm×mm)	7.0×65×115

度为 115 mm。电池电压的工作范围为 2.8～4.2 V, 额定容量为 10 000 mAh, 标定电压为 4.2 V。单体电池的质量为 (151.4±1) g, 根据 GB 38031—2020[25] 电池相关要求, 在实验测试前, 电池单体需要进行预处理循环, 以确保实验对象的性能处于激活和稳定的状态。

1.2　实验平台

实验采用 1∶1 真实密闭储能舱, 锂离子电池位于储能舱内的储能柜中, 低压双流体细水雾喷头位于电池正上方, 实验舱的一侧面由钢化玻璃制成, 主要用于实验观测, 实验主要由电池加热系统、温度实时监测系统、电压电阻测试仪、细水雾灭火系统和烟气净化系统组成。细水雾灭火系统为自主研发, 气体入口压力和液体入口压力分别约为 0.4 MPa 和 0.04 MPa。红外烟气分析仪型号为 SMC-9021, 可分析气体为 SO_2、NO、NO_2、O_2、CO、CO_2, 精度为±0.2%。电压电阻测试仪型号为 HK3560, 电阻为 1×10^{-10}～3 kΩ(精度为±0.3%), 电压为 1×10^{-4}～60 V (精度为±0.01%)。无纸记录仪型号为 RX6048C, 温度为－200～1 200 ℃ (精度为±0.01 ℃)。电池充放电设备型号为 CI-ZWJ-4S-I-10。平台示意图如图 2 所示。

如图 2(a)所示, 舱体的尺寸为 2 000 mm×2 000 mm×1 670 mm, 放置锂离子电池的架子尺寸为



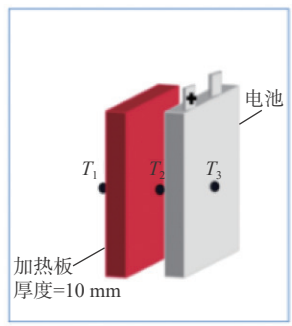
(a) 实验整体示意图



(b) 电池充放电设备



(c) 电压和温度测试设备



(d) 热电偶布置图

图2 实验平台示意图

Fig. 2 Experimental platform

495 mm×175 mm×630 mm, 架体总高 630 mm, 分为3层, 每层高度为 210 mm, 尺寸参考中航锂电公司实际生产产品大小。实验使用图 2(a)所示低压双流体细水雾抑灭火系统, 其主要特点是利用双流体细水雾喷头进行水的雾化, 此喷头具有水和气 2 个流动通道, 气体一般采用空气, 气体入口压力约为 0.4 MPa, 液体入口压力约 0.04 MPa, 喷头距离电池约 50 cm, 喷射角度为 180°, 根据《细水雾灭火系统技术规范》[26]相关要求, 喷射细水雾时长至少为 30 min, 在锂离子电池四周形成一个环形的细水雾区域。图 2(b)为电池充放电设备。实验使用图 2(c)所示的无纸记录仪和电压电阻测试仪对锂离子电池的温度和电压电阻实时监测和记录。实验平台搭建时安装了气体过滤系统, 实验中产生的气体可以经过过滤系统转化为安全气体无害排放至空气中。如图 2(d)所示, T_1、T_2 和 T_3 分别代表热源的温度、电池直接受热大面温度和电池非直接受热大面温度。电池加热主要利用电加热板, 加热板功率为 300 W, 厚度为 10 mm, 均匀给电池加热, 同时使用高温胶带将探头直径为 1 mm 的热电偶固定在如图 2(d)所示的中心位置。实验过程中, 使用不锈钢丝对加热板和电池进行固定, 为防止电池倾斜和受热电池形变空间受限, 不锈钢丝固定电池和加热板时会预留电池热失控时最大应变空间。

1.3 实验流程

实验流程如图 3 所示。利用电池充放电平台将新拆封的锂离子电池以 0.5 C 的充放电倍率放电至截止电压 2.8 V, 静置 10 min 后, 再充电至最大截至电压 4.2 V, 按此步骤, 充放电循环 3 次至 100%荷电状态, 静置 12 h 后用于实验。所有实验均在常温 25℃ 的情况下开展。当锂离子电池受热面温度 T_2 分别达到 90 ℃、145 ℃ 和 200 ℃ 时关闭加热板,

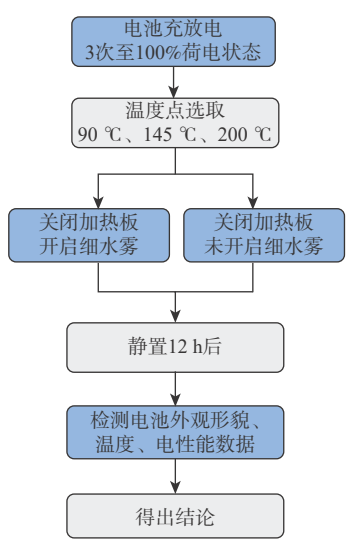


图3 实验流程

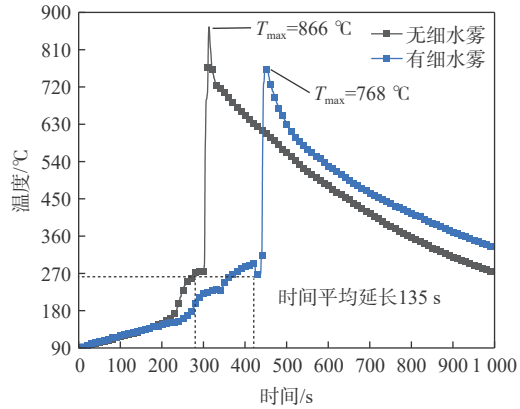
Fig. 3 Experimental flowchart

分别展开开启和不开启低压双流体细水雾实验。实时监测锂离子电池的产热情况和电压电阻的变化情况, 在低压双流体细水雾施加 30 min 后电池静置 12 h, 一直到锂离子电池表面水分完全蒸发, 再将受热后的锂离子电池与实验前新电池以同样的方式进行充放电性能测试, 得到实验后受热锂离子电池的电性能相关数据。为保证实验准确性, 每组实验重复 3 次, 加上空白对照组实验, 共计 18 次实验。

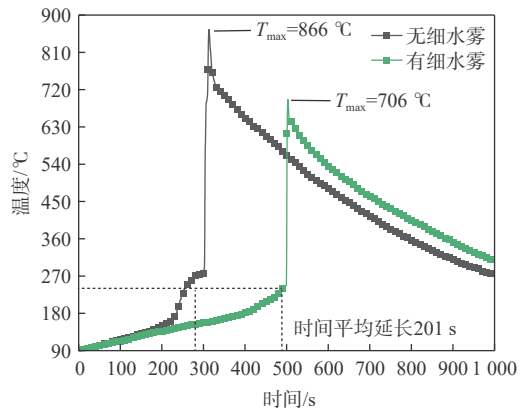
2　结果与讨论

2.1　细水雾施加气液比流量分析

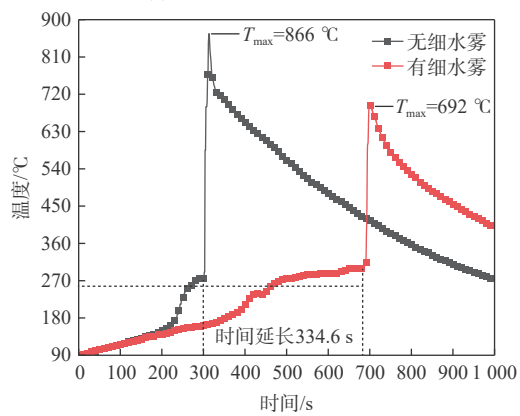
锂离子电池火灾抑灭与灭火剂的流量选取息息相关, 不同流量决定了灭火剂的抑灭火效率。本文所选的低压双流体细水雾包含水和空气 2 种灭火介质, 抑灭受限空间锂离子电池火灾应该探明内在流量配比关系。经多次实验验证, 锂离子电池热失控的平均触发条件为 220℃ 左右, 本文所选温度点基于该触发条件。为选用合适的细水雾气液比, 选取 325 L/min/1.25 L/min、300 L/min/1.50 L/min 和 275 L/min/1.75 L/min 3 个细水雾气液比进行实验探究, 细水雾气液比为 325 L/min/1.25 L/min 时, 锂离子电池热失控时间平均延后 135 s, 峰值温度降低了 98 ℃; 气液比为 300 L/min/1.50 L/min 时, 锂离子电池热失控时间平均延后 201 s, 峰值温度降低 160 ℃。根据国家标准 GB 38031—2020[25], 单个电池或电池包发生热失控前要至少预留出 5 min 的时间, 多次实验下证明气液比 275 L/min/1.75 L/min(双流体细水雾空气和细水雾出口流量值分别为 275 L/min 和 1.75 L/min)能够使锂离子电池热失控时间平均延后 334.6 s, 峰值温度降低 174 ℃, 而 30 min 用水量仅 52.5 L, 相比水喷淋和高压细水雾用水量节约 70%～80% 左右, 故本次实验选用的低压双流体细水雾气液比为 275 L/min/1.75 L/min, 如图 4 所示。



(a) 气液比为325 L/min/1.25 L/min



(b) 气液比为300 L/min/1.50 L/min



(c) 气液比为275 L/min/1.75 L/min

图 4　不同气液比下电池热失控延长时间对比

Fig. 4　Comparison of battery thermal runaway extension time under different gas-liquid ratios

2.2　细水雾保护电池外观形貌结果分析

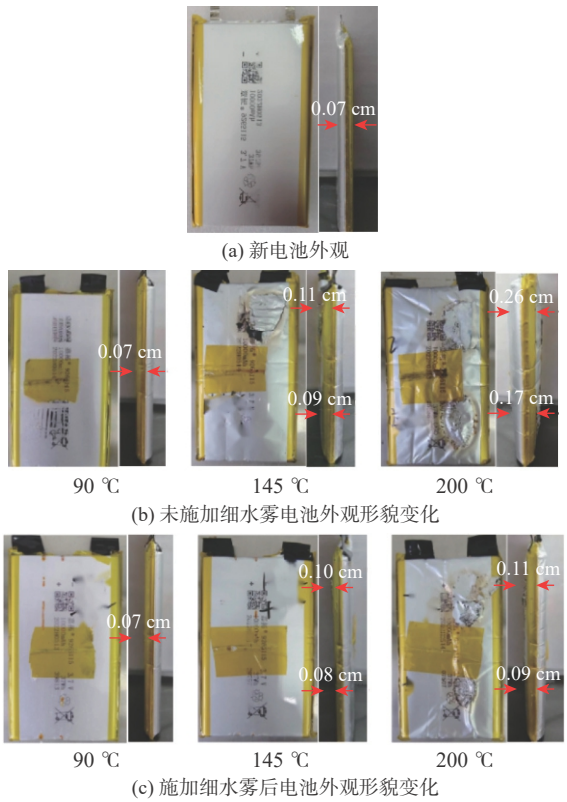
铝塑膜是软包装锂离子电池电芯封装的关键材料和最为关键的一个安全环节, 其是由多种塑料、铝箔和黏合剂组成的高强度、高阻隔的多层复合材料。单片电池组装后用铝塑膜密封, 形成一个电池, 铝塑膜起保护电芯的作用。当外界温度在 170～190 ℃ 时, 铝塑膜会发生变形蜷缩, 当温度在 230 ℃ 左右时, 铝塑膜表面会有明显损坏, 在有外力拉扯的情况下极易发生断裂[27]。

按照实验方案, 先在电池受热面 T_2 达到 90 ℃、145 ℃ 和 200 ℃ 时, 关闭加热板, 自然冷却后, 观察电池外观形貌的变化情况; 再在电池受热面 T_2 达到 90 ℃、145 ℃ 和 200 ℃ 时, 关闭加热板, 同时开启细水雾, 细水雾施加30 min 后, 对比观察电池外观形貌的变化情况。

三元锂离子电池有无施加细水雾的外观形貌观测如图 5 所示, 图 5(a)～图 5(c)均由电池正面和侧面图像 2 部分组成。

图 5(a)为新电池外观。在未受热时, 电池表面整齐光滑, 无任何破损的现象, 电池的正常厚度为 0.07 cm。

图 5(b)为 90 ℃、145 ℃ 和 200 ℃ 未施加细水雾时电池的外观破损情况。加热至 90 ℃ 时的电池



(a) 新电池外观

(b) 未施加细水雾电池外观形貌变化

(c) 施加细水雾后电池外观形貌变化

图 5 有无细水雾电池外观形貌对比

Fig. 5 Battery morphology comparison with or without water mist

外观上看不出明显的破损；加热至 145 ℃ 时，明显可以看出极耳附近产生破损撕裂的情况，锂离子电池也有轻微鼓包的现象，距离极耳较近的地方，电池鼓包厚度为 0.11 cm，较新电池厚度增加了 0.04 cm，远离极耳的部分电池鼓包厚度为 0.09 cm，相比新电池厚度增加了 0.02 cm。产生破损撕裂的原因是铝塑膜高温变形破裂，锂离子电池出现轻微鼓包的原因是锂离子电池在内部温度 120～130 ℃ 左右会发生副反应，锂离子与电解液中的溶剂碳酸乙烯酯（ethylene carbonate, EC）、碳酸丙烯酯（propylene carbonate, PC）、碳酸二甲酯（dimethyl carbonate, DMC）发生化学反应，产生少量的 C_2H_4、C_3H_6 和 C_2H_6 可燃气体[28]；加热至 200 ℃ 时，锂离子电池表面也有较为严重的破损，表面最外层的铝膜已经损坏，锂离子电池严重鼓包膨胀，锂离子电池极耳附近鼓包厚度达 0.26 cm，远离极耳的部分鼓包厚度达 0.17 cm。随着温度的升高，锂离子电池内部温度达到 150～200 ℃ 时，正极物质氧化锂钴开始分解，产生较多的 O_2 和其他微量气体，进而导致锂离子电池膨胀严重[29]。

图 5(c)分别为 90 ℃、145 ℃ 和 200 ℃ 施加细水雾时电池的外观破损情况。加热至 90 ℃ 时的电池外观无明显破损；加热至 145 ℃ 时，电池表面无破损，且电池整体鼓包现象不明显，极耳附近鼓包厚度达 0.10 cm，远离极耳的部分鼓包厚度达 0.08 cm，较未施加细水雾时电池 2 个位置的鼓包厚度分别减少了 0.01 cm；加热至 200 ℃ 时，锂离子电池有轻微鼓包膨胀，极耳附近鼓包厚度达 0.11 cm，远离极耳的部分鼓包厚度达 0.09 cm，较未施加细水雾时电池 2 个位置的鼓包厚度分别减少了 0.15 cm 和 0.08 cm。电池表面也存在破损，较未施加细水雾时破损面积偏小，出现这种情况的原因是细水雾的及时施加抑制了热量的产生，进而延缓了锂离子电池内部产生化学反应，且延缓效果较好。另外，极耳附近形变较为明显，原因可能是锂离子电池极耳处产生的热量较多。整体来看，施加双流体细水雾对锂离子电池外观形貌有一定的保护作用。

2.3 细水雾抑制电池产热结果分析

一般加热板内部装有加热丝或加热盘，在通电时，加热丝或加热盘会因为电阻而发热，从而提供稳定的加热能力。在加热板给锂离子电池加热过程中，锂离子电池紧贴加热板的受热面会接收加热板产生的热量，这些热量经过一定厚度的锂离子电池，将热量传递给锂离子电池非受热面，热量传递过程需要一定的时间，存在热量损耗[30]，因此，在实验数据可以看出 T_1、T_2 和 T_3 呈递减的趋势。按照实验方案，分别在 90 ℃、145 ℃ 和 200 ℃ 这 3 个温度点进行实验，对双流体细水雾施加 30 min 过程中不同测温点（T_1、T_2 和 T_3）进行实时监测，监测结果分别如图 6～图 8 所示。

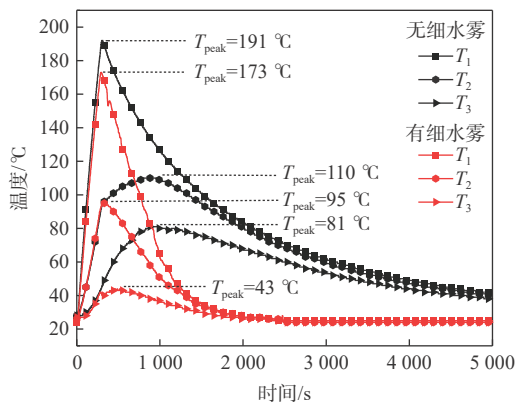


图 6 90 ℃ 有无细水雾不同测温点对比

Fig. 6 Comparison of temperature measurement points with or without water mist at 90 ℃

图 6 为 90 ℃ 有无细水雾温度的数据对比（温度误差为±2.5 ℃）。没有施加细水雾时热源的峰值温度为 191 ℃，电池直接受热大面的峰值温度为 110 ℃，电池未直接受热大面的峰值温度为 81 ℃，当细水雾施加 30 min 后，T_1、T_2 和 T_3 的温度都快速下降到常温状态，细水雾施加期间，热源的峰值温度为 173 ℃，电池直接受热大面的峰值温度为 95 ℃，

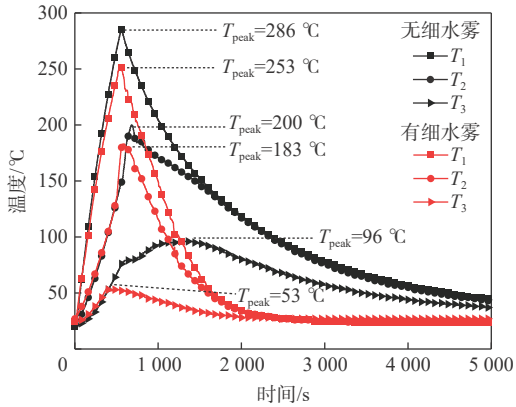


图 7　145 ℃ 有无细水雾不同测温点对比

Fig. 7　Comparison of temperature measurement points with or without water mist at 145 ℃

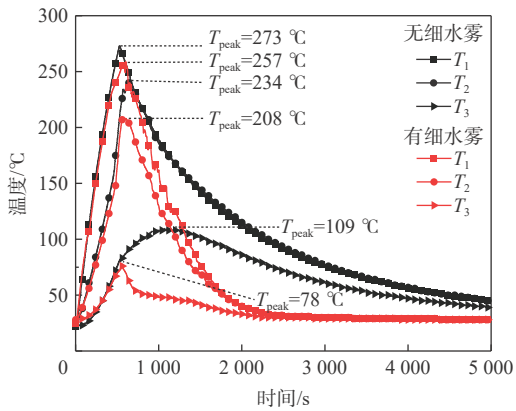


图 8　200 ℃ 有无细水雾不同测温点对比

Fig. 8　Comparison of temperature measurement points with or without water mist at 200 ℃

电池非直接受热大面的峰值温度为 43 ℃, 由于细水雾较好的冷却降温作用, 热源降低了 18 ℃, 电池直接受热大面温度降低了 15 ℃, 电池非直接受热大面温度降低了 38 ℃。

图 7 为 145 ℃ 有无细水雾温度的数据对比。未施加细水雾时热源的峰值温度为 286 ℃, 电池直接受热大面的峰值温度为 200 ℃, 电池非直接受热大面的峰值温度为 96 ℃, 施加细水雾后, 热源温度降低了 33 ℃, 峰值温度为 253 ℃, 电池直接受热大面温度降低了 17 ℃, 峰值温度为 183 ℃, 电池非直接受热大面温度降低了 43 ℃, 峰值温度为 53 ℃。

图 8 为 200 ℃ 有无细水雾温度的数据对比。未施加细水雾时热源的峰值温度为 273 ℃, 电池直接受热大面的峰值温度为 234 ℃, 电池非直接受热大面的峰值温度为 109 ℃, 施加细水雾后, 热源温度降低了 16 ℃, 其峰值温度为 257 ℃, 电池直接受热大面温度降低了 26 ℃, 其峰值温度为 208 ℃, 电池非直接受热大面温度降低了 31 ℃, 其峰值温度为 78 ℃。

整体来看, 不同温度点在关闭加热板前, 温升曲线几乎一致, 说明热源加热速率一致性较好。双流体细水雾主要是利用气相冷却的原理来降低锂离子电池的温度, 由于细水雾粒径很小, 受热后可迅速汽化, 从受热锂离子电池区域吸收大量的热, 同时空气的进入会有扰流作用, 气体在空间内形成一定的紊流循环。电池非受热面在细水雾的影响下降温幅度最大, 原因是电池非直接受热大面完全裸露在空间中, 能够最大面积接触到细水雾。表面积相对较大, 更易于汽化, 并带走更多的热量[31]。

2.4　细水雾保护电池电性能结果分析

锂离子电池在受热后, 其电压电阻、锂离子电池充放电容量相应也会受到一定的影响。Lu 等[28]在研究三元锂离子电池热失控时提出随着持续加热, 锂离子电池会出现高温容量衰减和一定程度的自放电, 同时锂离子电池的电压也会出现微小的降低情况, 电阻呈逐渐增大的趋势。按照实验方案, 对施加细水雾和未施加细水雾的不同温度电池的电压电阻进行实时监测, 测试结果如图 9～图 11 所示。

图 9 为 90 ℃ 有无细水雾电池的电压电阻对

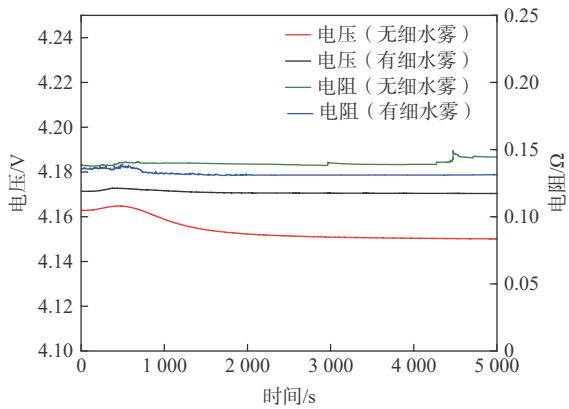


图 9　90 ℃ 有无细水雾电压电阻对比

Fig. 9　Comparison of voltage resistance with or without water mist at 90 ℃

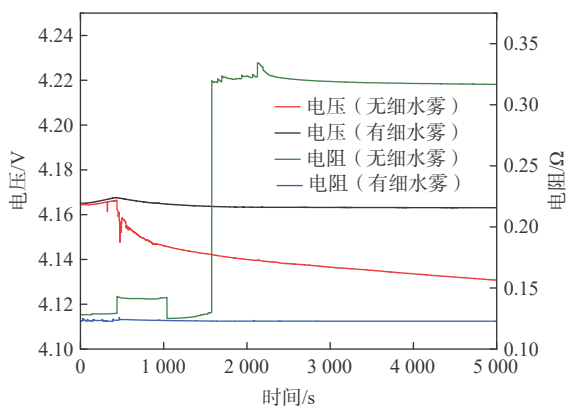


图 10　145 ℃ 有无细水雾电压电阻对比

Fig. 10　Comparison of voltage resistance with or without water mist at 145 ℃

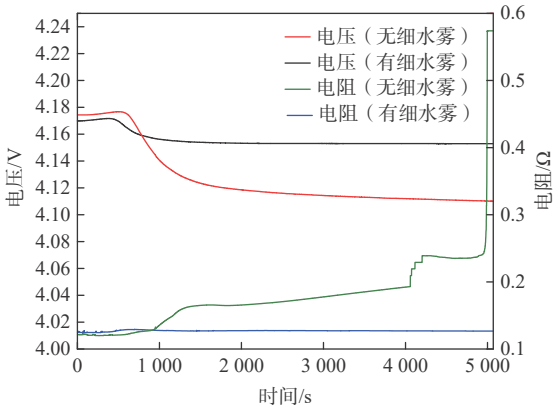


图 11 200 ℃ 有无细水雾电压电阻对比

Fig. 11 Comparison of voltage resistance with or without water mist at 200 ℃

比。前期电池的电压电阻有微小波动属于正常现象。本文重点观测整体的趋势，在 90 ℃ 时，未施加细水雾，电池的电压由最初监测的 4.163 V 下降至 4.148 V，随着温度的升高，下降了 0.015 V，电池的电阻由最初监测的 0.138 Ω 升高至 0.144 Ω，增加了 0.006 Ω；在施加细水雾的过程中，电压从最初的 4.173 V 下降至 4.170 V，几乎可以看作没有下降，是稳定状态，电阻一直在 0.131 Ω 上下波动，施加细水雾后的电压和电阻整体是平稳的，无太大的起伏波动。

图 10 为 145 ℃ 有无细水雾电池的电压电阻对比。在 145 ℃ 时，未施加细水雾，电池的电压由最开始的 4.164 V 下降至 4.130 V，下降了 0.034 V，电池的电阻由最开始的 0.128 Ω 升高至 0.317 Ω，施加细水雾后，电压和电阻都趋于一条直线，电压在 4.165 V 左右微小波动，电阻在 0.122 Ω 左右微小波动。

图 11 为 200 ℃ 有无细水雾电池的电压电阻对比。在 200 ℃ 时，未施加细水雾，电池的电压由最开始的 4.174 V 下降至 4.110 V，下降了 0.064 V，电池的电阻由最开始的 0.121 Ω 升高至 0.547 Ω，施加细水雾后，电压有轻微下降趋势，从 4.169 V 下降至 4.152 V，后期和电阻一样，还是趋于一条直线，电阻在 0.127 Ω 左右微小波动。

锂离子电池随着受热温度升高，电池产生副反应破坏了电池内的化学平衡，导致电池电压降低，当分子热运动频率随温度上升，锂离子电池内部电子流动时将受到更多次的热运动分子的碰撞和影响而降低运动速度，即电阻增大。综上，电池在 90 ℃、145 ℃ 和 200 ℃ 施加细水雾后，电压比未喷细水雾要高，电阻比未喷细水雾要低。电压方面，在 3 个不同的温度点施加细水雾后，电压比未施加有所提升，且一致性较好；电阻方面，随着温度升高，未施加细水雾电阻变大，施加细水雾后各温度点电阻相对变小。同时，在实验过程中，由于低压双流体细水雾选择的气液比合适，未造成电池短路的现象。说明低压双流体细水雾对受热电池的电压、电阻有很好的保护作用。

锂离子电池在完成受热实验后，静置 12 h，待到电池表面水分充分蒸发和热量完全散去，对受热程度不同的电池进行充放电性能测试，和前期充放电的倍率保持一致，得到的结果如图 12 所示。

图 12(a)为电池在不同温度下有无细水雾充电容量对比。新电池的容量为 10 000 mAh。未施加细水雾时，电池在不同程度受热后(90 ℃、145 ℃ 和 200 ℃)充电容量分别为 9 492.52 mAh、5 860.34 mAh 和 1 818.20 mAh，随着温度的升高，电池受到不同程度的损坏，电池的性能也随之降低。施加细水雾后，电池在受热后充电容量分别为 9 980.06 mAh、9 977.03 mAh 和 8 048.26 mAh，电池充电容量分别提高了 487.54 mAh、4 116.69 mAh 和 6 230.06 mAh，充电容量随温度升高会有轻微的下降，但相比未施加细水雾的电池，整体有很大的提升。图 12(b)为电池在不同温度下有无细水雾放电容量对比。未施加细水雾时，电池在温度为 90 ℃、145 ℃ 和 200 ℃ 时放电容量分别为 9 347.47 mAh、5 538.01 mAh 和

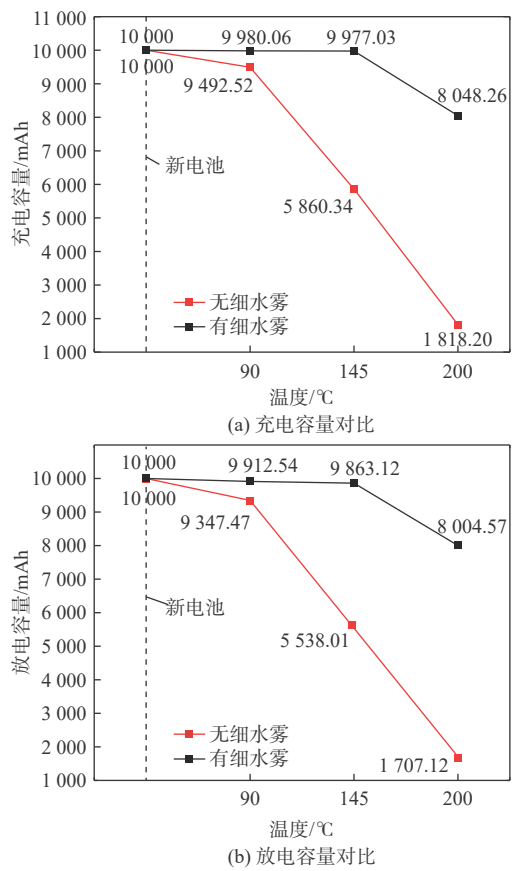


图 12 不同温度有无细水雾电池充放电容量对比

Fig. 12 Comparison of charging and discharging capacity of batteries with or without water mist at different temperatures

1 707.12 mAh, 施加细水雾后, 电池放电容量分别为 9 912.54 mAh、9 863.12 mAh 和 8 004.57 mAh, 分别提高了 565.07 mAh、4 325.11 mAh 和 6 297.45 mAh。表 2 为不同受热程度有无细水雾电池充放电性能对比。进一步说明了低压双流体细水雾对锂离子电池的充放电性能有很好的提升效果。

表 2　不同受热程度有无细水雾电池充放电性能对比

Table 2　Comparison of charge and discharge performance of batteries with or without water mist at different heating degrees

温度/℃	实验阶段	充电容量/mAh	放电容量/mAh	充放电效率/%
90	无细水雾	9 492.52	9 347.47	98.47
	有细水雾	9 980.06	9 912.54	99.32
145	无细水雾	5 860.34	5 538.01	94.50
	有细水雾	9 977.03	9 863.12	98.86
200	无细水雾	1 818.20	1 707.12	93.89
	有细水雾	8 048.26	8 004.57	99.46

3　结　论

针对低压双流体细水雾对锂离子电池热失控灾害抑制性能影响问题, 本文详细对锂离子电池的外观形貌、产热抑制和电性能防护 3 个方面开展了实验研究, 得到以下结论:

1) 锂离子电池在受热过程中, 内部和外部会产生不同程度的损伤, 对比 90 ℃、145 ℃ 和 200 ℃ 空白实验, 施加低压双流体细水雾后, 极耳附近电池鼓包厚度分别减少 0 cm、0.01 cm 和 0.15 cm; 远离极耳的部分, 电池鼓包厚度分别减少 0 cm、0.01 cm 和 0.08 cm, 电池表面受损面积也有所减小, 说明细水雾对受热锂离子电池的外观形貌具有一定的保护作用。

2) 施加细水雾后, 电池非受热面温度降低幅度最大, 在 90 ℃、145 ℃ 和 200 ℃ 分别降低了 38 ℃、43 ℃ 和 31 ℃。双流体细水雾能够有效地抑制锂离子电池的热量, 从而减少电池内部化学反应和外部的损伤, 抑制锂离子电池产热效果良好。

3) 受热后的锂离子电池电性能随着温度的升高, 电压降低至 4.110 V, 电阻增大至 0.547 Ω, 电池的充放电性能降低, 在施加细水雾后, 电池的各方面性能有了一定程度的提升, 说明细水雾对电池的电性能具有一定的防护作用。

三元锂离子电池普遍应用于中小型储能系统, 在热滥用、机械滥用、电滥用等条件下极易发生热失控, 甚至爆炸等二次灾害, 本文对热失控前的产热和电性能进行研究, 可以为中小型的储能系统消防灭火设计提供一定的基础依据。

参考文献（References）

[1] 王鹏博, 郑俊超. 锂离子电池的发展现状及展望[J]. 自然杂志, 2017, 39(4): 283-289.
WANG P B, ZHENG J C. The present situation and expectation of lithium-ion battery[J]. Chinese Journal of Nature, 2017, 39(4): 283-289(in Chinese).

[2] 徐晓楠. 灭火剂与应用[M]. 北京: 化学工业出版社, 2006: 81-82.
XU X N. Fire extinguishing agent and application[M]. Beijing: Chemical Industry Press, 2006: 81-82(in Chinese).

[3] WANG Q S, PING P, ZHAO X J, et al. Thermal runaway caused fire and explosion of lithium ion battery[J]. Journal of Power Sources, 2012, 208: 210-224.

[4] FENG X N, OUYANG M G, LIU X, et al. Thermal runaway mechanism of lithium ion battery for electric vehicles: a review[J]. Energy Storage Materials, 2018, 10: 246-267.

[5] WHITTINGHAM M S. History, evolution, and future status of energy storage[J]. Proceedings of the IEEE, 2012, 100(SI): 1518-1534.

[6] HENRIKSEN M, VAAGSAETHER K, LUNDBERG J, et al. Explosion characteristics for Li-ion battery electrolytes at elevated temperatures[J]. Journal of Hazardous Materials, 2019, 371: 1-7.

[7] CHEN S C, WANG Z R, WANG J H, et al. Lower explosion limit of the vented gases from Li-ion batteries thermal runaway in high temperature condition[J]. Journal of Loss Prevention in the Process Industries, 2020, 63: 103992.

[8] LIU X, WU Z B, STOLIAROV S I, et al. Heat release during thermally-induced failure of a lithium ion battery: impact of cathode composition[J]. Fire Safety Journal, 2016, 85: 10-22.

[9] 中华人民共和国住房和城乡建设部, 国家市场监督管理总局. 消防设施通用规范: GB 55036—2022[S]. 北京: 中国计划出版社, 2022: 12-19.
Ministry of Housing and Urban-Rural Development of the People's Republic of China, State Administration for Market Regulation. General code for fire protection facilities: GB 55036—2022[S]. Beijing: China Planning Press, 2022: 12-19(in Chinese).

[10] MENG X D, YANG K, ZHANG M J, et al. Experimental study on combustion behavior and fire extinguishing of lithium iron phosphate battery[J]. Journal of Energy Storage, 2020, 30: 101532.

[11] WANG Q, CHEN J Z, LI M, et al. Eltoprazine prevents levodopa-induced dyskinesias by reducing causal interactions for theta oscillations in the dorsolateral striatum and substantia nigra pars reticulate[J]. Neuropharmacology, 2019, 148: 1-10.

[12] LI Y L, LIN Z L. A generalized piecewise quadratic Lyapunov function approach to estimating the domain of attraction of a saturated system[J]. IFAC-PapersOnLine, 2015, 48(11): 120-125.

[13] WANG Q S, SHAO G Z, DUAN Q L, et al. The efficiency of heptafluoropropane fire extinguishing agent on suppressing the lithium titanate battery fire[J]. Fire Technology, 2016, 52(2): 387-396.

[14] SI R J, LIU D Q, XUE S Q. Experimental study on fire and explosion suppression of self-ignition of lithium ion battery[J]. Procedia Engineering, 2018, 211: 629-634.

[15] HYNES R. Inhibition of premixed hydrogen-air flames by 2-H heptafluoropropane[J]. Combustion and Flame, 1998, 113(4): 554-565.

[16] 张乃平, 马永飞, 杨孟霖, 等. 锂电池火灾灭火技术研究综述[J]. 中国安全生产科学技术, 2022, 18(7): 47-53.
ZHANG N P, MA Y F, YANG M L, et al. Review on fire extinguishing technology research of lithium battery[J]. Journal of Safety Science and Technology, 2022, 18(7): 47-53(in Chinese).

[17] 胡振恺, 李勇琦, 彭鹏. 电池储能系统火灾预警与灭火系统设计[J]. 消防科学与技术, 2020, 39(10): 1434-1438.
HU Z K, LI Y Q, PENG P. Design of fire warning system and fire extinguishing system of battery energy storage system[J]. Fire Science and Technology, 2020, 39(10): 1434-1438(in Chinese).

[18] LARSSON F, ANDERSSON P, BLOMQVIST P, et al. Characteristics of lithium-ion batteries during fire tests[J]. Journal of Power Sources, 2014, 271: 414-420.

[19] ZHANG L, DUAN Q L, LIU Y J, et al. Experimental investigation of water spray on suppressing lithium-ion battery fires[J]. Fire Safety Journal, 2021, 120: 103117.

[20] 韩路豪, 王子阳, 何骁龙, 等. 细水雾释放策略对大容量三元锂离子电池热失控火灾抑制效果的实验研究[J]. 储能科学与技术, 2023, 12(5): 1664-1674.
HAN L H, WANG Z Y, HE X L, et al. The effect of water mist strategies on thermal runaway fire suppression of large-capacity NCM lithium-ion battery[J]. Energy Storage Science and Technology, 2023, 12(5): 1664-1674(in Chinese).

[21] 丰汉军, 邹丽, 张宇航. 高压细水雾抑制磷酸铁锂电池模块火灾试验研究及设计讨论[J]. 给水排水, 2022, 58(11): 114-119.
FENG H J, ZOU L, ZHANG Y H. Experimental research on fire suppression of lithium iron phosphate battery module by high pressure water mist and discussion on design[J]. Water & Wastewater Engineering, 2022, 58(11): 114-119(in Chinese).

[22] LIU T, TAO C F, WANG X S. Cooling control effect of water mist on thermal runaway propagation in lithium ion battery modules[J]. Applied Energy, 2020, 267: 115087.

[23] 赵军超, 付阳阳, 薛峰, 等. 低压细水雾系统扑救锂电池箱火灾的试验研究[J]. 安全与环境学报, 2022, 22(4): 1897-1903.
ZHAO J C, FU Y Y, XUE F, et al. Experimental study on suppressing lithium battery package fires by low-pressure water mist system[J]. Journal of Safety and Environment, 2022, 22(4): 1897-1903 (in Chinese).

[24] KRITZER P, DORING H, EMERMACHER B. Improved safety for automotive lithium batteries: an innovative approach to include an emergency cooling element[J]. Advances in Chemical Engineering and Science, 2014, 4(2): 197-207.

[25] 国家市场监督管理总局, 国家标准化管理委员会. 电动汽车用动力蓄电池安全要求: GB 38031—2020[S]. 北京: 中国标准出版社, 2020: 5-6.
State Administration for Market Regulation, National Standardization Administration. Electric vehicles traction battery safety requirements: GB 38031—2020[S]. Beijing: China Standard Press, 2020: 5-6(in Chinese).

[26] 中华人民共和国住房和城乡建设部, 中华人民共和国国家质量监督检验检疫总局. 细水雾灭火系统技术规范: GB 50898—2013[S]. 北京: 中国计划出版社, 2013: 10-11.
Ministry of Housing and Urban-Rural Development of the People's Republic of China, General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China. Technical code for water mist fire extinguishing system: GB 50898—2013[S]. Beijing: China Planning Press, 2013: 10-11(in Chinese).

[27] 王雪峰, 纪小刚. 软包锂电池铝塑膜热封强度工艺参数研究[J]. 实验力学, 2023, 38(2): 222-230.
WANG X F, JI X G. Study on the heat-sealing process parameters of aluminum plastic film for pouch lithium battery[J]. Journal of Experimental Mechanics, 2023, 38(2): 222-230(in Chinese).

[28] LU L G, HAN X B, LI J Q, et al. A review on the key issues for lithium-ion battery management in electric vehicles[J]. Journal of Power Sources, 2013, 226: 272-288.

[29] JIANG J, DAHN J R. ARC studies of the thermal stability of three different cathode materials: $LiCoO_2$; $Li[Ni_{0.1}Co_{0.8}Mn_{0.1}]O_2$; and $LiFePO_4$, in $LiPF_6$ and LiBoB EC/DEC electrolytes[J]. Electrochemistry Communications, 2004, 6: 39-43.

[30] 王淮斌, 李阳, 王钦正, 等. 三元锂离子动力电池热失控及蔓延特性实验研究[J]. 工程科学学报, 2021, 43(5): 663-675.
WANG H B, LI Y, WANG Q Z, et al. Experimental study on the thermal runaway and its propagation of a lithium-ion traction battery with NCM cathode under thermal abuse[J]. Chinese Journal of Engineering, 2021, 43(5): 663-675(in Chinese).

[31] GRANT G, BRENTON J, DRYSDALE D. Fire suppression by water sprays[J]. Progress in Energy and Combustion Science, 2000, 26(2): 79-130.

Performance of low-pressure dual-fluid water mist to reduce thermal runaway risk of ternary lithium-ion battery

ZENG Xiangshu，LIANG Jiaxin，SUN Qiang，HE Yuanhua*

(Civil Aircraft Fire Science and Safety Engineering Key Laboratory of Sichuan Province，
Civil Aviation Flight University of China，Guanghan 618307，China)

Abstract：To address the shortcomings of gas, high-pressure water mist, and other fire extinguishing agents in suppressing lithium battery fires, this paper investigated the low-pressure dual-fluid water mist for ternary lithium-ion battery fire suppression. The effects of low-pressure dual-fluid water mist on the appearance, heat generation inhibition, and electrical performance protection of lithium batteries at different temperatures (90 °C, 145 °C, and 200 °C) before thermal runaway were examined. The results show that under the application of low-pressure dual-fluid water mist, the thickness of the bulge near the pole lugs decreased by 0 cm, 0.01 cm, and 0.15 cm, and the extent of damage to the battery surface was reduced. The temperature of the non-heated surface of the battery decreases the most, by 38 °C, 43 °C, and 31 °C. The battery's charging capacity increases by 487.54 mAh、4116.69 mAh and 6230.06 mAh, and the discharge capacity increases by 565.07 mAh、4325.11 mAh and 6297.45 mAh. These findings demonstrate that low-pressure dual-fluid water mist not only effectively protects the appearance, morphology, heat suppression, and electrical properties of lithium batteries, but also proves its reliability in inhibiting the thermal runaway process during the early stages.

Keywords：low-pressure dual-fluid water mist; ternary lithium-ion battery; heat generation inhibition; electrical performance protection; secondary disaster

Received：2023-08-08；**Accepted**：2023-10-13；**Published Online**：2023-11-14 08：30

URL：link.cnki.net/urlid/11.2625.V.20231110.1706.003

Foundation items：National Natural Science Foundation of China (U2033206); Sichuan Provincial Science and Technology Program (2021YFSY0001)

* **Corresponding author**. E-mail：he_yuan_hua@163.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0565

含滚道剥落缺陷的滚动轴承动力学建模及分析

王潇[1,2]，马军[1,2,*]，熊新[1,2]，陈靖宇[1,2]

(1. 昆明理工大学 信息工程与自动化学院，昆明 650500；
2. 昆明理工大学 云南省先进装备智能控制及应用国际联合实验室，昆明 650500)

摘　　要：为了深入分析轴承滚道剥落缺陷引发冲击性振动响应的机理，动力学模型的构建成为必要且有效的手段。针对现有轴承动力学模型中的缺陷激励函数难以准确刻画实际剥落缺陷的问题，提出一种考虑缺陷三维尺寸的激励函数来构建含滚道剥落缺陷的轴承动力学模型。以椭圆形状模拟缺陷三维形貌，在此基础上构建缺陷激励函数，并将其融入动力学模型。分析滚动体通过不同尺寸缺陷过程中产生的额外接触位移量的变化情况。通过仿真和实验验证模型的正确性，并深入探究缺陷尺寸与振动响应间的关系。与矩形激励函数模型对比，证明所提模型的有效性和适用性。实验结果表明：所构建的模型与理论、实测信号误差均在±5% 以内；缺陷尺寸的变化将引起额外接触位移的改变，进而使轴承产生不同的振动响应；相比矩形激励函数模型，所提模型可有效模拟滚动轴承在不同尺寸剥落缺陷下的振动响应。相关研究结果可为轴承故障诊断及性能评估提供理论支持。

关　键　词：滚动轴承；滚道剥落；缺陷尺寸；动力学建模；振动响应分析
中图分类号：TH133.33
文献标志码：A　　**文章编号：**1001-5965(2025)10-3492-12

滚动轴承广泛应用于航空航天、交通运输等领域的重要机械设备中，作为各类机械的“关节”，其起到了支撑旋转结构的作用[1]。滚动轴承在运行过程中常处于高转速、强载荷的工作环境下，极易产生滚道的点蚀、磨损与剥落等缺陷，影响设备运行安全。其中，由疲劳引发的剥落是轴承最常见的失效形式之一[2]。因此，构建轴承动力学模型，深入分析滚道剥落缺陷引发冲击的机理及其振动响应对于轴承的故障诊断和性能评估等具有非常重要的意义。

为了探究缺陷引发冲击性振动响应的机理，国内外学者针对轴承动力学模型的构建进行了深入研究，并取得了显著成果。Epps[3] 提出了滚动轴承的缺陷激励函数及动力学模型，并分析了带有缺陷的滚动轴承的振动响应特性。Sawalhi 等[4] 将单元谐振器融入到动力学模型中以模拟缺陷引发的高频共振现象，建立了含剥落缺陷的滚动轴承 5 自由度动力学模型。胡爱军等[5] 基于 5 自由度动力学模型分析了轴承滚道多点故障引起的振动特性变化规律。Cui 等[6] 通过动力学模型仿真了滚道剥落缺陷扩展过程不同阶段的振动响应结果，以辅助滚动轴承性能评估和寿命预测。张耀强等[7] 利用脉冲序列模拟缺陷激励，并分析了外圈局部缺陷引起的轴承振动响应特性。Wu 等[8] 通过建立动力学模型，实现了滚动轴承滚道剥落缺陷引发的双冲击现象的仿真，并基于模型完成了缺陷尺寸定量评估。Zhang 等[9] 构建了 2 自由度动力学模型，仿真分析了缺陷长度、转速、载荷等因素对轴承动力学响应的影响。胡爱军等[10] 修正了时变接触刚度和接触力，建立模型仿真出了滚动体通过缺陷区域的多次

收稿日期：2023-09-04；录用日期：2023-09-22；网络出版时间：2023-10-23 17:42
网络出版地址：link.cnki.net/urlid/11.2625.V.20231023.1423.004
基金项目：国家自然科学基金 (62163020,62173168)；云南省基础研究计划 (202101BE070001-055)；云南省重大科技专项 (202202AD080005)
*通信作者. E-mail：491941203@qq.com

引用格式：王潇，马军，熊新，等. 含滚道剥落缺陷的滚动轴承动力学建模及分析 [J]. 北京航空航天大学学报，2025，51（10）：3492-3503. WANG X，MA J，XIONG X，et al. Dynamic modeling and analysis of rolling bearing with raceway spalling defect [J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51（10）：3492-3503（in Chinese）.

冲击现象。常斌全等[11]分析了滚动体通过缺陷区域的进入、撞击、离开、载荷补偿事件,建立了多事件激励的滚动轴承动力学模型。上述研究为了简化计算,将缺陷等效为矩形或用脉冲序列直接模拟,使得滚动体通过缺陷区域过程的运动轨迹与实际存在一定的差距。

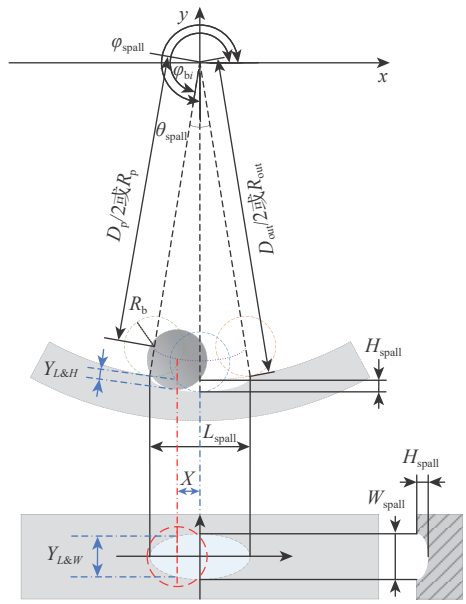
为了进一步提升模型的保真度,学者们结合了实际缺陷的形貌及滚动体经过缺陷区域的运动特性来建立模型。李志农等[12]利用圆弧代替了矩形缺陷,证明了考虑变形量渐变释放过程的必要性。李昊泽等[13]分析了滚动体经过缺陷区域时的间隙变化过程来构建激励函数。Sawalhi[14]和王震[15]等考虑了缺陷表面粗糙度因素,建立了非规则缺陷的轴承动力学模型,并分析了非规则缺陷对振动特性的影响。以上研究利用了非矩形形状来模拟剥落缺陷形貌,但仅在缺陷长度、深度二维尺寸上进行仿真分析。雷春丽等[16]考虑缺陷长度、宽度、深度三维尺寸构建缺陷激励函数,并分析了矩形、三角形、梯形及复合形状等不同形貌缺陷演化过程的位移激励变化规律,但并未进一步分析缺陷尺寸与轴承振动加速度信号间的关系。除此之外,振动传递路径[17]、冲击力[18]、摩擦[19]等因素也被融入到了轴承动力学模型中,提高了仿真模型的保真度。

综上所述,目前轴承动力学建模研究中将滚道剥落缺陷简化为矩形的方法难以描述复杂形状剥落缺陷形貌,且现有缺陷激励函数普遍未考虑缺陷三维尺寸的影响;同时,缺陷三维尺寸对轴承振动特性影响的研究中也未深入分析缺陷尺寸与加速度信号间的关系。针对以上问题,本文提出一种考虑缺陷三维尺寸的激励函数,并构建含滚道剥落缺陷的轴承动力学模型;开展含 4 类不同尺寸缺陷的轴承振动响应实验与分析,并利用模型辅助分析不同缺陷尺寸下振动响应发生变化的原因。此外,进一步分析缺陷尺寸与轴承振动加速度响应间的关系。最后,与矩形激励函数模型进行对比,证明本文模型模拟不同尺寸剥落缺陷振动响应的有效性和适用性。

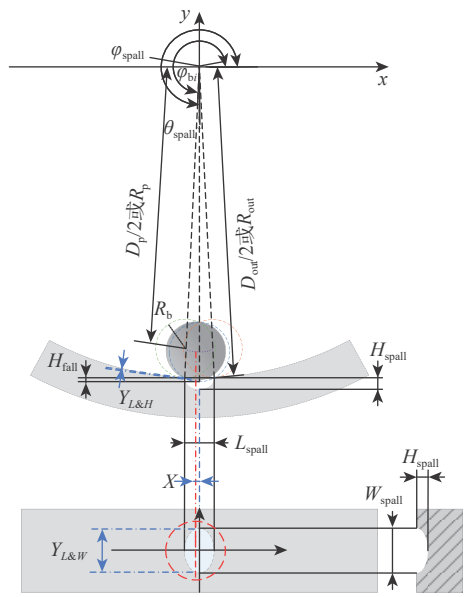
1　滚道剥落缺陷激励函数构建

当滚动体通过缺陷区域时,滚动体向剥落区偏移,产生额外接触位移,使得总接触位移量和接触力发生变化,进而引起冲击性振动。缺陷激励函数则是刻画该过程中额外接触位移量的必要工具。同时,文献 [20] 研究了剥落缺陷进出口的微观图像,结果也表明滚道剥落缺陷形貌具有连续变化过程。因此,本文以椭圆形模拟剥落缺陷形貌,考虑其三维尺寸构建缺陷激励函数以表征滚动体通过缺陷区域引发的额外接触位移量变化过程。

图 1 为滚动体通过缺陷区域过程示意图,分别用周向长度 L_{spall}、轴向宽度 W_{spall}、径向深度 H_{spall} 来表示剥落缺陷的尺寸大小。



(a) 大尺寸缺陷



(b) 小尺寸缺陷

图 1　滚动体通过缺陷区域过程

Fig. 1　Process of balls passing through the defect area

基于缺陷长度 L_{spall} 和深度 H_{spall} 构建缺陷底面函数 B。缺陷在 L_{spall}-H_{spall} 平面的投影为半椭圆形,记变量 X 为滚动体经过缺陷区域过程中其几何中心与缺陷中心的水平距离,变量 $Y_{L\&H}$ 为缺陷底部距离滚道表面的高度。将 L_{spall} 和 H_{spall} 代入椭圆标准方程,推导得出 X 和 $Y_{L\&H}$ 之间的关系如下:

$$\frac{4X^2}{L_{\mathrm{spall}}^2}+\frac{Y_{L\&H}^2}{H_{\mathrm{spall}}^2}=1 \tag{1}$$

得到 $Y_{L\&H}$ 的表达式(2),并用 B 代替表示为缺陷底面到滚道平面距离变化的函数:

$$B = Y_{L\&H} = \left| \sqrt{1 - \frac{4X^2}{L_{\text{spall}}^2}} \right| H_{\text{spall}} \tag{2}$$

当滚动体通过小尺寸缺陷时,滚动体主要与缺陷边缘发生接触,则需要计算该情况下滚动体最大下落量 H_{fall},用于判断滚动体通过缺陷区域过程中是否会与缺陷底面接触。同时为简化计算,以新的椭圆表征此下落过程,并将该过程中滚动体的下落量记作 $\Delta H_{L\&H}$,具体如下:

$$H_{\text{fall}} = \sqrt{R_{\text{b}}^2 - \frac{L_{\text{spall}}^2}{4}} - R_{\text{b}} \tag{3}$$

$$\Delta H_{L\&H} = \left| \sqrt{1 - \frac{4X^2}{L_{\text{spall}}^2}} \right| H_{\text{fall}} \tag{4}$$

式中:R_{b} 为滚动体半径。

记变量 $Y_{L\&W}$ 为滚动体通过缺陷过程中瞬时位置对应的轴向缺陷宽度。与下落量分析不同的是,缺陷在 L_{spall}-W_{spall} 平面的投影为椭圆。将 L_{spall} 和 W_{spall} 代入椭圆方程,经变换化简可得到 $Y_{L\&W}$ 的表达式(5)。进而可计算得到该位置滚动体因与缺陷轴向边缘接触产生的下落量 $\Delta H_{L\&W}$,具体如式(6)所示。

$$Y_{L\&W} = W_{\text{spall}} \sqrt{\frac{1}{4} - \frac{X^2}{L_{\text{spall}}^2}} \tag{5}$$

$$\Delta H_{L\&W} = \sqrt{R_{\text{b}}^2 - \frac{Y_{L\&W}^2}{4}} - R_{\text{b}} \tag{6}$$

式(2)、式(4)和式(5)中的 X 可由滚动体角度和轴承几何参数推导得到,具体过程如下。

由图 1 可知,缺陷范围角度 θ_{spall} 可由缺陷长度 L_{spall} 及轴承滚道尺寸表示为

$$\theta_{\text{spall}} = \frac{L_{\text{spall}}}{R_{\text{out}}} \tag{7}$$

式中:R_{out} 为外圈滚道半径。

变量 X 包含式(8)所示的滚动体“进入”缺陷区域和“离开”缺陷区域 2 个阶段:

$$X = \begin{cases} X_{\text{entry}} & \varphi_{\text{spall}} - \dfrac{\theta_{\text{spall}}}{2} \leqslant \varphi_{\text{b}i} \leqslant \varphi_{\text{spall}} \\ X_{\text{exit}} & \varphi_{\text{spall}} < \varphi_{\text{b}i} \leqslant \varphi_{\text{spall}} + \dfrac{\theta_{\text{spall}}}{2} \end{cases} \tag{8}$$

式中:X_{entry} 为滚动体进入缺陷并到达中心位置过程中其几何中心与缺陷中心的水平距离;X_{exit} 为滚动体通过缺陷中心位置到离开缺陷过程中几何中心与缺陷中心的水平距离;φ_{spall} 为剥落缺陷的位置角度;$\varphi_{\text{b}i}$ 为第 i 个滚动体的瞬时位置角度。

对于“进入”过程,滚动体瞬时位置距离缺陷入口点的角度差 $\Delta\varphi_{\text{entry}}$ 表示为

$$\Delta\varphi_{\text{entry}} = \varphi_{\text{b}i} - \left(\varphi_{\text{spall}} - \frac{\theta_{\text{spall}}}{2} \right)$$
$$\varphi_{\text{spall}} - \frac{\theta_{\text{spall}}}{2} \leqslant \varphi_{\text{b}i} \leqslant \varphi_{\text{spall}} \tag{9}$$

滚动体与缺陷中心的距离 X_{entry} 为

$$X_{\text{entry}} = -\left(\frac{L_{\text{spall}}}{2} - \Delta\varphi_{\text{entry}} R_{\text{out}} \right)$$
$$\varphi_{\text{spall}} - \frac{\theta_{\text{spall}}}{2} \leqslant \varphi_{\text{b}i} \leqslant \varphi_{\text{spall}} \tag{10}$$

将 X_{entry} 代入式(2)、式(4)和式(5),进而计算得到“进入”过程滚动体瞬时位置对应的 B、$\Delta H_{L\&H}$ 和 $\Delta H_{L\&W}$ 的数值。

同理,“离开”过程中,滚动体瞬时位置距离缺陷出口点的角度差 $\Delta\varphi_{\text{exit}}$ 为

$$\Delta\varphi_{\text{exit}} = \left(\varphi_{\text{spall}} + \frac{\theta_{\text{spall}}}{2} \right) - \varphi_{\text{b}i}$$
$$\varphi_{\text{spall}} < \varphi_{\text{b}i} \leqslant \varphi_{\text{spall}} + \frac{\theta_{\text{spall}}}{2} \tag{11}$$

$$X_{\text{exit}} = \frac{L_{\text{spall}}}{2} - \Delta\varphi_{\text{exit}} R_{\text{out}}$$
$$\varphi_{\text{spall}} < \varphi_{\text{b}i} \leqslant \varphi_{\text{spall}} + \frac{\theta_{\text{spall}}}{2} \tag{12}$$

将 X_{exit} 代入式(2)、式(4)和式(5)后,计算得到“离开”过程滚动体瞬时位置对应的 B、$\Delta H_{L\&H}$ 和 $\Delta H_{L\&W}$ 的数值。

最终,综合判断滚动体通过缺陷全过程中 B、$\Delta H_{L\&H}$ 和 $\Delta H_{L\&W}$ 的数值大小,取三者中最小值为滚动体瞬时位置额外接触位移量 δ_{d},如式(13)所示。需要注意的是,δ_{d} 在非缺陷区域时均为 0。

$$\delta_{\text{d}} = \min(B, \Delta H_{L\&H}, \Delta H_{L\&W}) \tag{13}$$

2 滚动轴承动力学模型构建

在建立动力学模型过程中,仅考虑轴承的径向运动和绕轴旋转,忽略轴承各部件的轴向运动和滚动体滑动、摩擦力及缺陷粗糙形貌的影响,同时,轴承滚动体和滚道满足 Hertz 接触理论。基于此,所建立的滚动轴承动力学模型示意图如图 2 所示,具体包括外圈、内圈、保持架、滚动体、单元谐振器等部分,其中,外圈固定,内圈随转轴旋转,保持架与滚动体随内圈同向转动。

保持架与滚动体公转速度 ω_{c} 与内圈转速 ω_{s} 具有如下关系:

$$\omega_{\text{c}} = \left(1 - \frac{D_{\text{b}} \cos\alpha}{D_{\text{p}}} \right) \frac{\omega_{\text{s}}}{2} \tag{14}$$

式中:D_{b} 为滚动体直径;D_{p} 为节圆直径;α 为接触角,深沟球轴承中为 0[21]。

每个滚动体瞬时位置角度 $\varphi_{\text{b}i}$ 可表示为

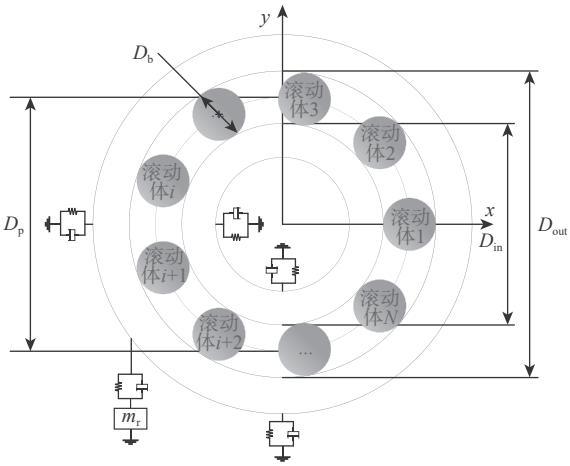


图 2　滚动轴承动力学模型示意图

Fig. 2　Diagram of the dynamic model of rolling bearing

$$\varphi_{bi}=\omega_c t+\frac{2\pi\,(i-1)}{N_{ball}}+\varphi_0 \tag{15}$$

式中：N_{ball} 为滚动体个数；i 为滚动体编号；φ_0 为第 1 个滚动体的初始角度。

如图 3 所示，轴承转动时，滚道与滚动体发生挤压接触，产生总接触位移量 δ_i 为

$$\delta_i=\left(x_s-x_p\right)\cos\varphi_{bi}+\left(y_s-y_p\right)\sin\varphi_{bi}-\delta_c-\delta_d \tag{16}$$

式中：x、y 分别为轴承在水平、竖直方向振动产生的位移量，下标 s、p 分别表示轴承内圈、外圈；δ_c 为径向游隙产生的位移；δ_d 为由缺陷产生的额外接触位移量，即式(13)。

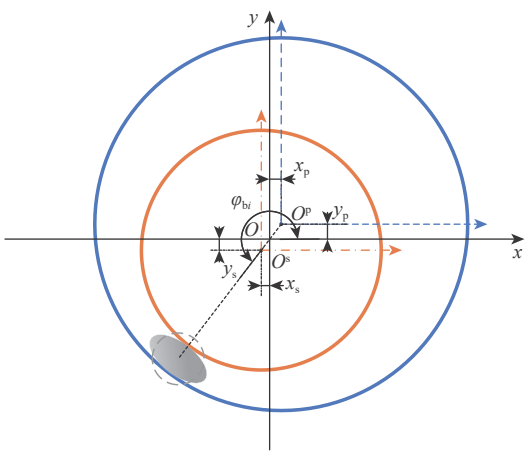


图 3　滚道与滚动体接触变形

Fig. 3　Contact deformation between raceway and balls

根据赫兹接触理论得到滚动体与滚道间的赫兹接触力为

$$\begin{cases} f_x=K_b\sum\limits_{i=1}^{N_{ball}} h\delta_i^n\cos\varphi_{bi}\quad h=0,\delta_i<0 \\ f_y=K_b\sum\limits_{i=1}^{N_{ball}} h\delta_i^n\sin\varphi_{bi}\quad h=1,\delta_i\geqslant 0 \end{cases} \tag{17}$$

式中：K_b 为滚动体与滚道的等效接触刚度；h 为接触判断系数；n 为接触变形系数，深沟球轴承一般取 3/2[21]。

根据牛顿第二定律，列出滚动轴承动力学模型微分方程组为[6]

$$\begin{cases} m_s\ddot{x}_s+c_s\dot{x}_s+k_s x_s+f_x=F_{sx} \\ m_s\ddot{y}_s+c_s\dot{y}_s+k_s y_s+f_y=F_{sy}+m_s g \\ m_p\ddot{x}_p+c_p\dot{x}_p+k_p x_p-f_x=0 \\ m_p\ddot{y}_p+\left(c_p+c_r\right)\dot{y}_p+\left(k_p+k_r\right)y_p- \\ \qquad k_r y_r-c_r\dot{y}_r-f_y=m_p g \\ m_r\ddot{y}_r+c_r\left(\dot{y}_r-\dot{y}_p\right)+k_r\left(y_r-y_p\right)=0 \end{cases} \tag{18}$$

式中：m、c、k 分别为滚动轴承的质量、阻尼、刚度，下标 r 表示单元谐振器；g 为重力加速度；F_s 为径向负载，f 为赫兹接触力投影分量，下标 x 和 y 分别表示水平方向和竖直方向。

此外，由于轴承产生的振动从支撑座传递到传感器的过程中会产生一定的衰减，因此，考虑传递路径衰减作用的影响，借鉴文献 [17,22] 中对传递路径的建模方法，利用汉宁窗函数模拟传递路径的衰减。

第 i 个滚动体的传递路径为[17]

$$S_i(t)=A\left[a-b\cos(\omega_c t+\varphi_{si})\right] \tag{19}$$

式中：a 和 b 为汉宁窗系数，具体数值设定可见文献 [17]；φ_{si} 为第 i 个滚动体与传感器位置的初始角度差；A 为轴承支座传递路径衰减系数，经过多次实验，本文将其设定为 0.02。

综合得到传感器接收到的振动信号为[22]

$$\begin{cases} z_x(t)=\ddot{x}(t)\sum\limits_{i=1}^{N_{ball}} S_i(t)/N_{ball} \\ z_y(t)=\ddot{y}(t)\sum\limits_{i=1}^{N_{ball}} S_i(t)/N_{ball} \end{cases} \tag{20}$$

3　仿真分析与验证

3.1　额外接触位移分析

以表 1 所示轴承几何参数进行仿真，不同尺寸的缺陷导致滚动体通过缺陷区域时产生不同的额外接触位移量，根据仿真结果分析，可分为图 4 所示的 4 个模式。以图 5 所示的 4 种缺陷尺寸对应的额外接触位移量变化曲线为例，分析不同模式下滚

表 1　轴承几何参数

Table 1　Geometric parameter of bearing

参数	数值
内滚道直径D_{in}/mm	47.0
外滚道直径D_{out}/mm	31.0
节圆直径D_p/mm	39.0
滚动体直径D_b/mm	8
滚动体个数N_{ball}	9
接触角α/(°)	0

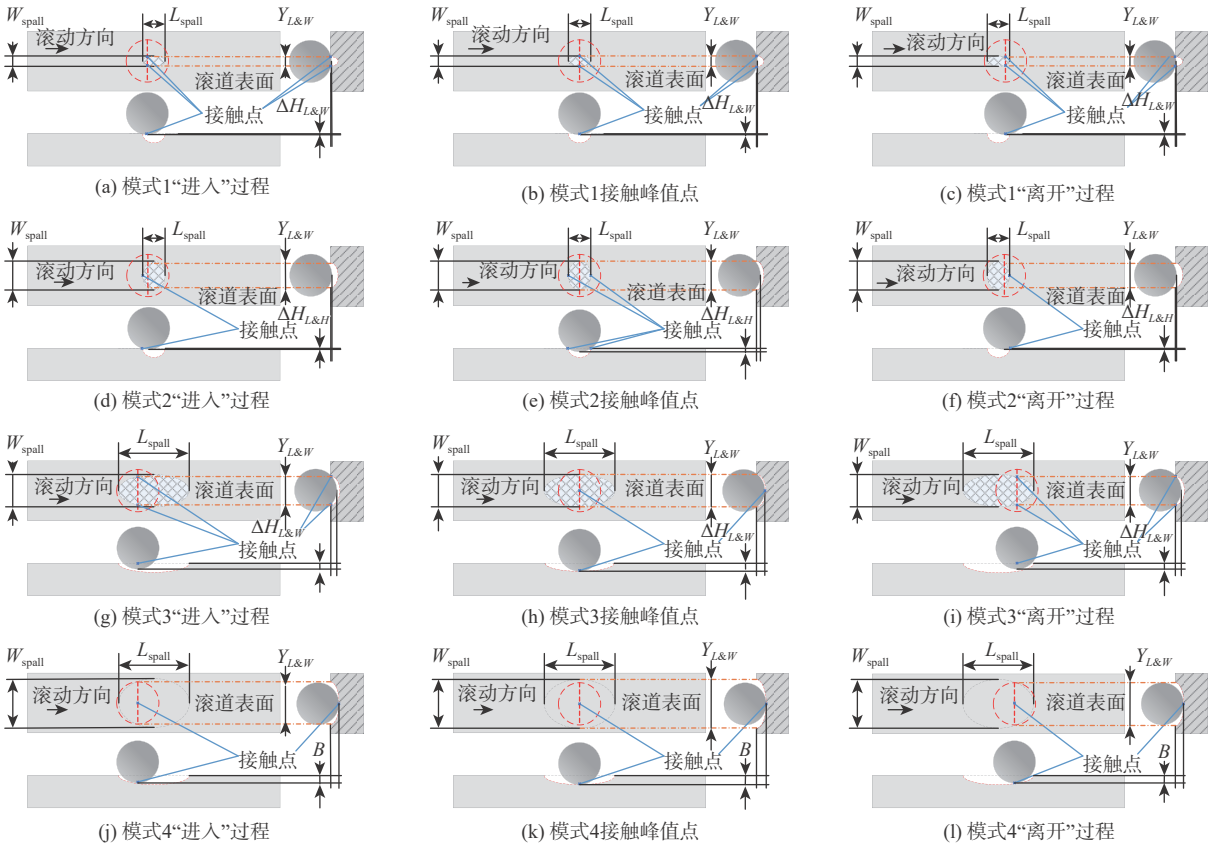


图 4 接触模式分析

Fig. 4 Analysis of contact modes

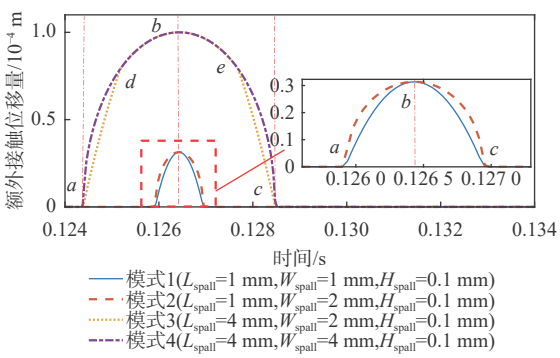


图 5 4种模式额外接触位移量曲线

Fig. 5 Additional contact displacement curves under four modes

动体通过缺陷区域的过程。

1) 模式 1: 缺陷长度、宽度均远小于滚动体直径, 滚动体通过缺陷区域的全过程与缺陷轴向边缘接触, 额外接触位移量小于滚动体瞬时位置处的缺陷底面到滚道面的距离($\Delta H_{L\&W}$<min(B, $\Delta H_{L\&H}$)), 如图 4(a)～图 4(c)所示。额外接触位移变化曲线对应图 5 中的模式 1, “进入”过程为曲线 *ab* 段, 接触峰值点对应 *b* 点, “离开”过程为 *bc* 段, 最大接触位移量小于缺陷深度。

2) 模式 2: 滚动体通过缺陷区域全过程仅与缺陷前后边缘接触, 产生的额外接触位移量小于缺陷的实际深度($\Delta H_{L\&H}$<min(B, $\Delta H_{L\&W}$)), 如图 4(d)～图 4(f)所示。对应图 5 中的模式 2, *ab* 段曲线为“进入”过程, *b* 点为接触峰值点, *bc* 段为“离开”过程, 全过程产生的额外接触位移为半椭圆形, 最大接触位移量也小于缺陷深度。

3) 模式 3: 滚动体经过缺陷时, 首先与缺陷边缘接触, 但未触碰缺陷底面($\Delta H_{L\&W}$<min($\Delta H_{L\&H}$, B)), 当滚动体经过缺陷中部区域时, 与缺陷底面接触(B<$\Delta H_{L\&H}$, $\Delta H_{L\&W}$), 具体过程如图 4(g)～图 4(i)所示。额外接触位移变化曲线对应图 5 中的模式 3, 与前面 2 种模式不同, 该过程为一个分段函数, *ad* 段曲线滚动体与缺陷边缘接触, *d* 点和 *e* 点为滚动体与缺陷边缘、底面接触的临界点, *de* 段曲线滚动体与缺陷底面接触, *b* 点时到达接触峰值点, 产生的最大接触位移量等于缺陷深度 H_{spall}, *e* 点之后恢复与缺陷边缘的接触。

4) 模式 4: 滚动体通过缺陷区域全过程均与缺陷底面发生接触, 此时的额外接触位移量完全由缺陷底面到滚道平面距离 B 决定(B<min($\Delta H_{L\&H}$, $\Delta H_{L\&W}$)), 如图 4(j)～图 4(l)所示。其额外接触位移变化曲线对应图 5 中的模式 4, 接触变化过程与模式 2 相似, 但不同的是, 该模式接触峰值 *b* 点对应的最大接触位移量与缺陷深度 H_{spall} 相等。

3.2 模型求解与验证

轴承动力学模型参数如表 2 所示。本文以外

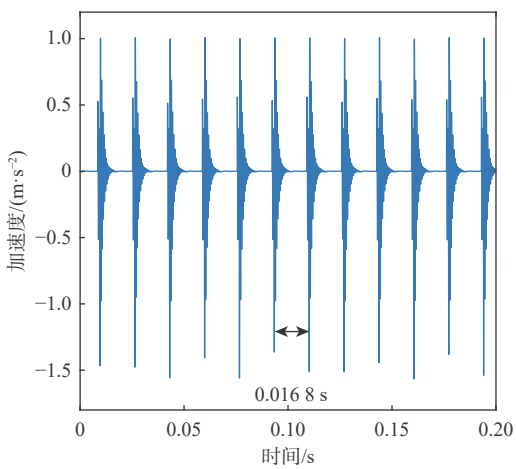
表 2　动力学模型参数

Table 2　Parameters of dynamic model

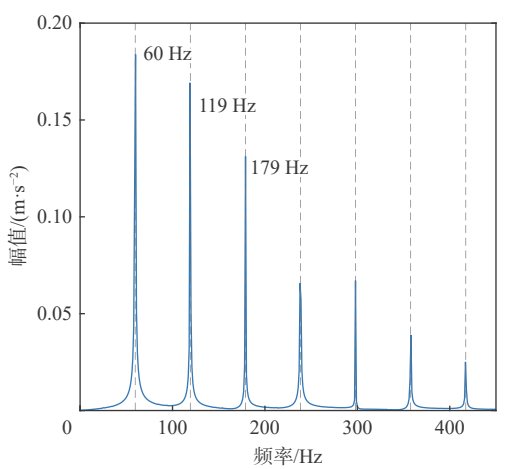
参数	数值
内圈及转轴质量m_s/kg	1.2
单元谐振器质量m_r/kg	1
外圈及底座刚度k_p/(N·m^{-1})	15.105 6×10^6
内圈及转轴阻尼c_s/(N·s·m^{-1})	2 376.8
单元谐振器阻尼c_r/(N·s·m^{-1})	9 424.8
外圈及底座质量m_p/kg	7.2
内圈及转轴刚度k_s/(N·m^{-1})	4.241×10^4
单元谐振器刚度k_r/(N·m^{-1})	8.882 6×10^9
外圈及底座阻尼c_p/(N·s·m^{-1})	2 210.7
等效接触刚度K_b/(N·m$^{-1.5}$)	1.897 8×10^{10}

圈滚道剥落缺陷为例, 进行仿真实验分析。设置仿真步长为 1/25 600, 利用 4 阶龙格库塔法求解得到振动响应结果。

借鉴文献 [13,23] 中测量时域信号冲击间隔及故障特征频率的方法来验证模型的正确性。图 6(a)和图 6(b)为转速 1 000 r/min 情况下得到的仿真加速度信号及其包络谱。测得 2 次冲击时间间隔为 0.016 8 s, 与相邻 2 个滚动体通过缺陷区域所需理论时间(0.016 8 s)一致。同时, 可观察到包络谱中明显的故障特征频率(60 Hz)及其倍频, 与故障特征频率理论值[24](59.62 Hz)误差仅有 0.64%, 初步验证了模型的正确性。



(a) 时域加速度

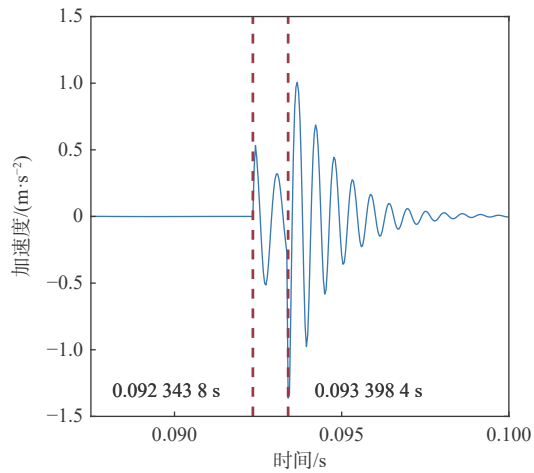


(b) 包络谱

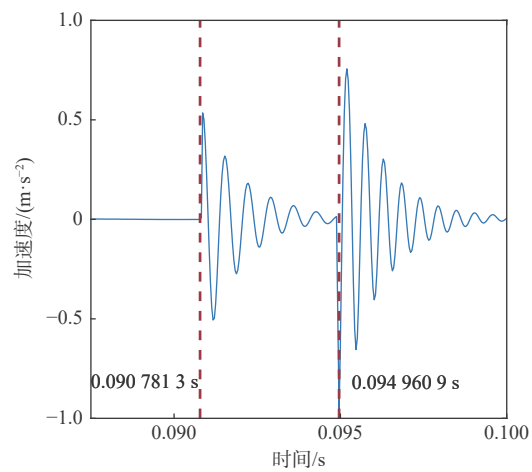
图 6　仿真振动加速度信号及频谱

Fig. 6　Simulation vibration acceleration signal and its frequency spectrum

如图 7 所示, 测量缺陷长度分别为 1 mm 和 4 mm 情况下的时域加速度信号双冲击时间间隔, 并以此推算缺陷尺寸[8,25]。1 mm 下的缺陷长度估计值为 1.03 mm, 与设定值误差为 3%, 4 mm 情况下缺陷长度估计值为 4.08 mm, 与设定值误差为 2%, 从双冲击现象的维度验证了模型的正确性。



(a) 缺陷长度为1 mm



(b) 缺陷长度为4 mm

图 7　仿真振动加速度信号局部放大图

Fig. 7　Local amplification diagram of simulation vibration acceleration signal

3.3　实验分析与验证

为了进一步验证模型的正确性, 使用自建滚动轴承实验平台采集实验数据与仿真结果进行对比。如图 8 所示, 实验平台包括驱动电机、转轴、支撑轴承、测试轴承、加速度传感器等部件。测试轴承几何尺寸与仿真参数一致。设置采样频率为

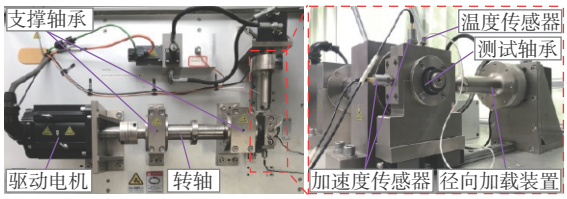


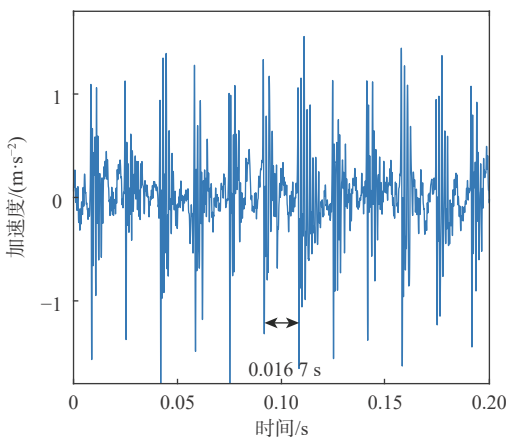
图 8 自建滚动轴承实验平台

Fig. 8 Self-built rolling bearing experimental platform

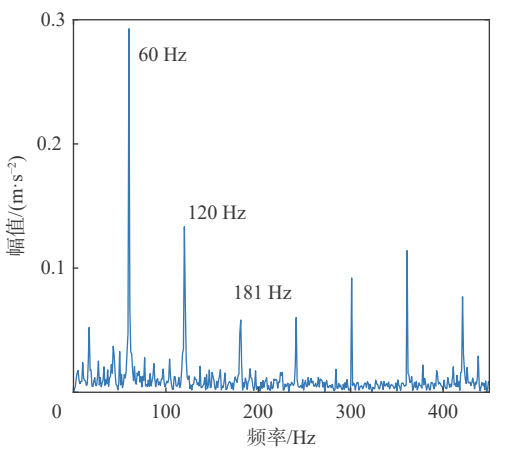
25.6 kHz, 驱动电机转速为 1 000 r/min。

图 9 为实测加速度信号及其包络谱。对比图 6(a)与图 9(a)可知, 相邻 2 个滚动体经过缺陷引发的 2 次冲击时间间隔为 0.016 7 s, 仿真信号与其误差为 0.60%。对比图 6(b)与图 9(b)包络谱, 均可明显观察到 60 Hz 故障特征频率, 仿真与实测包络谱的 2 倍频、3 倍频分别存在−0.83%、−1.1% 的误差, 分析原因为: 实验中受到轴承实际加工尺寸、电机转速误差等因素的影响, 导致仿真与实际存在误差。总体而言, 仿真与实测信号在时域冲击间隔及故障特征频率上高度符合, 验证了模型的正确性。

将实测信号经过低通滤波器去除部分噪声, 基



(a) 时域加速度

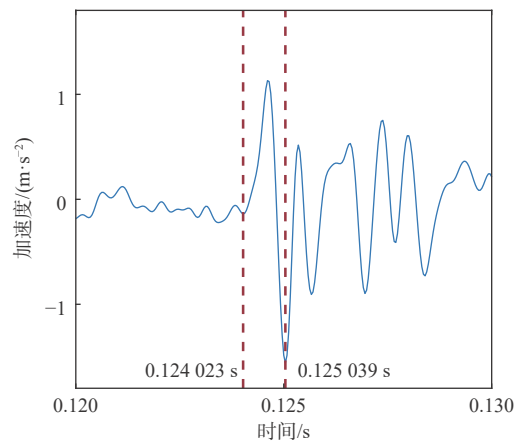


(b) 包络谱

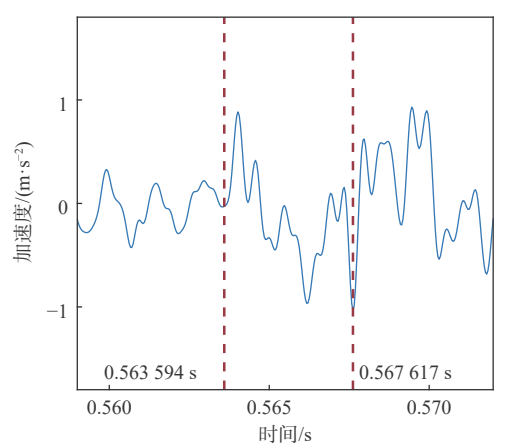
图 9 实测振动加速度信号及频谱

Fig. 9 Measured vibration acceleration signal and its frequency spectrum

于双冲击现象进行缺陷长度的估计。图 10(a)为缺陷长度为 1 mm 时的实测加速度信号。计算得到的缺陷长度评估值为 0.99 mm, 与实际缺陷长度误差为−1%, 模型仿真结果与实测评估值误差为 4.04%; 图 10(b)为缺陷长度为 4 mm 时的加速度信号。经计算得到评估值为 3.93 mm, 与实际缺陷长度误差为−1.75%, 模型与实测评估值误差为 3.82%。随机选取信号中 10 次不同位置的冲击进行定量评估, 得到结果如表 3 所示。2 类不同尺寸情况下模型与实际尺寸、实测评估值间的平均误差均在±5% 以内, 证明了模型的正确性。



(a) 缺陷长度为1 mm



(b) 缺陷长度为4 mm

图 10 实测振动加速度信号局部放大图

Fig. 10 Local amplification diagram of measured vibration acceleration signal

3.4 缺陷尺寸对振动响应影响分析

利用电火花方法在轴承滚道上分别加工出图 11 所示的 4 种不同尺寸剥落缺陷, 4 个试件缺陷深度均为 0.1 mm, 尺寸分别对应了 3.1 节分析的 4 种不同接触模式。

将 4 个试件安装到自建实验台进行实验测试, 在与前文一致的实验条件下采集得到如图 12 所示的加速度振动信号。

表 3　双冲击定量评估结果

Table 3　Quantitative evaluation results based on dual-impulse

编号	1 mm缺陷评估值/mm	4 mm缺陷评估值/mm	评估值与实际误差(1 mm)/%	评估值与实际误差(4 mm) /%	模型与评估值误差(1 mm) /%	模型与评估值误差(4 mm) /%
1	0.88	4.01	−12.00	0.25	17.05	1.75
2	1.03	3.85	3.00	−3.75	0	5.97
3	1.03	3.89	3.00	−2.75	0	4.88
4	1.07	3.97	7.00	−0.75	−3.74	2.77
5	0.95	3.82	−5.00	−4.50	8.42	6.81
6	0.95	4.20	−5.00	5.00	8.42	−2.86
7	0.99	3.82	−1.00	−4.50	4.04	6.81
8	0.92	3.86	−8.00	−3.50	11.96	5.70
9	0.96	4.01	−4.00	0.25	7.29	1.75
10	1.07	4.01	7.00	0.25	−3.74	1.75

注: 1 mm缺陷评估值、4 mm缺陷评估值、评估值与实际误差(1 mm)、评估值与实际误差(4 mm) 、模型与评估值误差(1 mm)和模型与评估值误差(4 mm)的平均值分别为0.985 mm、3.944 mm、−1.50%、−1.40%、4.97%和3.53%。



(a) 试件1 (L_{spall}=1 mm W_{spall}=1 mm)　(b) 试件2 (L_{spall}=1 mm W_{spall}=2 mm)　(c) 试件3 (L_{spall}=4 mm W_{spall}=2 mm)　(d) 试件4 (L_{spall}=4 mm W_{spall}=4 mm)

图 11　含有不同缺陷的测试轴承

Fig. 11　Test bearings with different defects

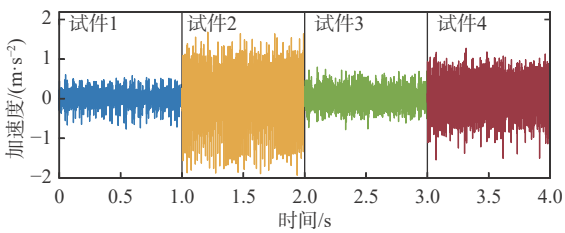


图 12　不同尺寸缺陷的实测振动加速度信号

Fig. 12　Measured vibration acceleration signals under different size defects

由图 12 可知, 试件 1 和试件 2 缺陷长度均为 1 mm, 随着宽度的增加, 振动加速度响应幅值也随之增加。同样在缺陷长度为 4 mm 的试件 3 和试件 4 中, 振动加速度响应幅值也会随缺陷宽度的扩展而增加。而对比缺陷宽度同为 2 mm 的试件 2 和试件 3 振动响应, 可观察到, 即使试件 3 的缺陷长度相比试件 2 更大, 但其加速度信号幅值相比试件 2 更小。

为了分析 4 个试件的振动响应存在差异的原因, 分别对各个试件进行动力学仿真。图 13 为 4 个不同尺寸缺陷下对应的仿真加速度信号。对比试件 1 与试件 2 的加速度仿真信号幅值, 可观察到, 后者在幅值上更大。同样, 试件 4 对应的仿真

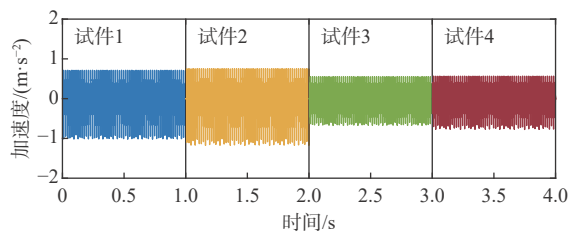


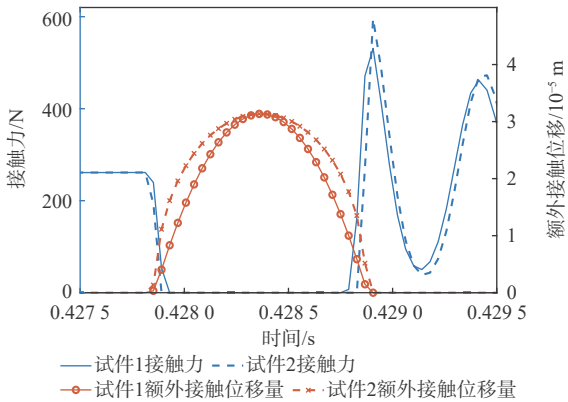
图 13　不同尺寸缺陷的仿真振动加速度信号

Fig. 13　Simulated vibration acceleration signal under different size defects

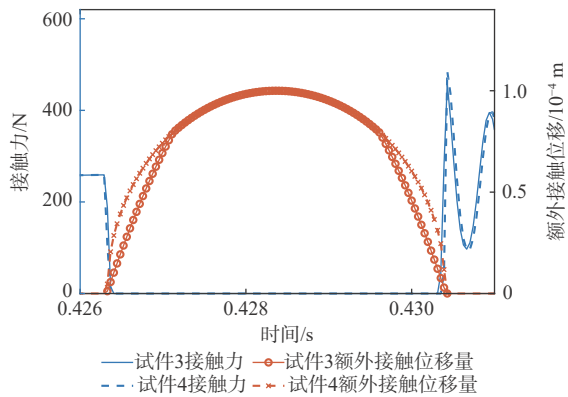
信号幅值相比试件 3 更大。试件 2 与试件 3 相比, 后者幅值出现了降低。以上所得的仿真规律与实测规律一致。

图 14 为 4 个试件的额外接触位移量与接触力变化对比曲线。由图 14(a)可知, 试件 1 和试件 2 分别对应 3.1 节的模式 1 和模式 2, 当宽度增加时, 额外接触位移曲线在滚动体“进入”和“离开”缺陷区域部分越陡峭, 斜率越大, 去应力和恢复应力阶段经历的时间则越短, 从而导致接触力变化急剧, 因此, 恢复应力过程中的接触力更大, 导致振动加速度幅值也变大。图 14(b)中试件 3 和试件 4 分别对应模式 3 和模式 4, 同属于长度固定、宽度扩展的情况, 具有与试件 1 和试件 2 相同的规律。由图 14(c)可知, 对比试件 2 与试件 3, 当缺陷宽度固定、长度增加时, 额外接触位移跨度、幅度也随之增大, 但该过程中接触力却有所减小, 对应的振动响应幅值降低, 并不随缺陷长度的增加而增大。

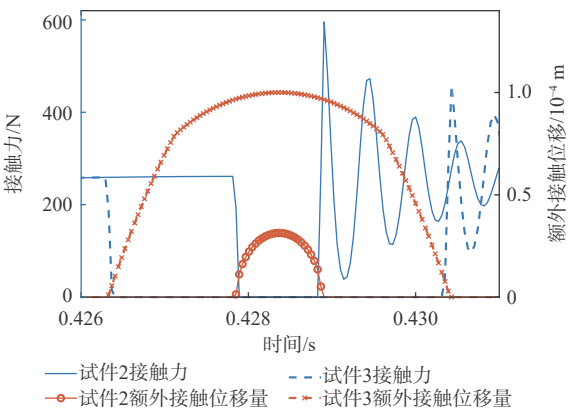
为了深入探究缺陷尺寸与振动响应间的关系, 分别对缺陷宽度变化、长度变化 2 种情况下的振动加速度信号进行统计分析。鉴于峰峰值能够有效表征加速度信号振动幅度的大小, 选用该指标作为



(a) 试件1和试件2



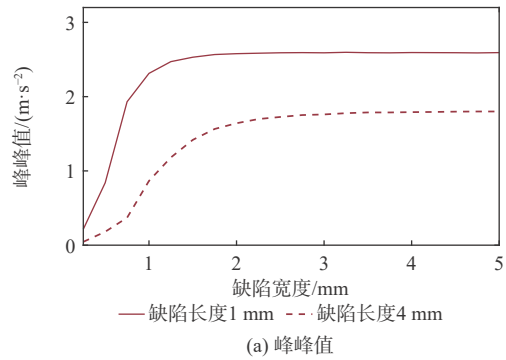
(b) 试件3和试件4



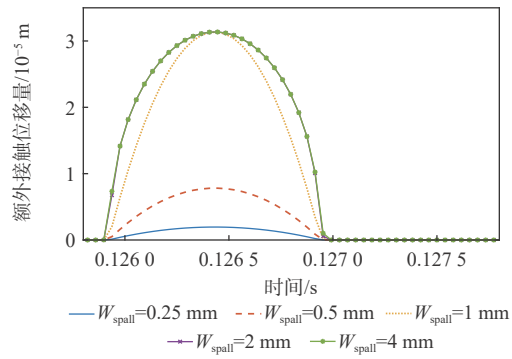
(c) 试件2和试件3

图 14 额外接触位移量与接触力变化曲线对比

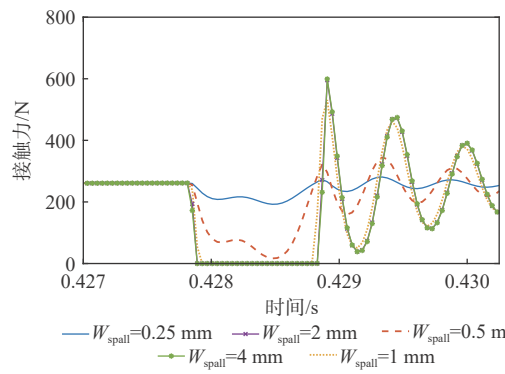
Fig. 14 Comparison of the curves of additional contact displacement and contact force



(a) 峰峰值



(b) 额外接触位移量



(c) 接触力

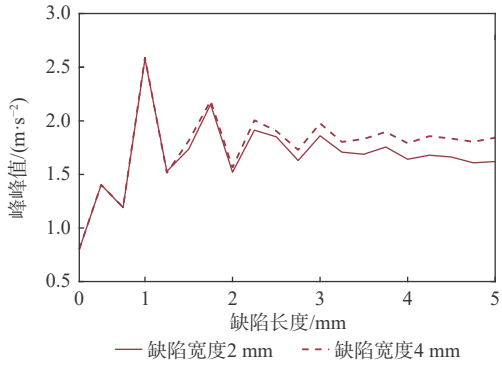
图 15 缺陷宽度变化过程振动响应分析

Fig. 15 Vibration response analysis of defect width change process

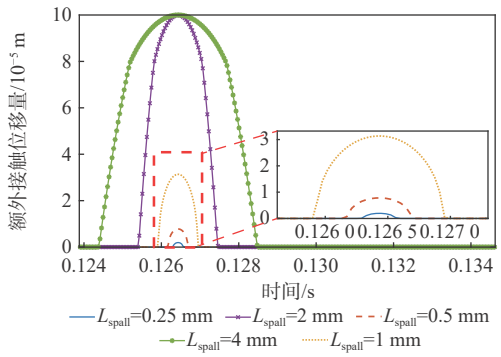
参考依据。

图 15 为缺陷长度分别为 1 mm、4 mm 时，宽度变化过程的峰峰值曲线及宽度为 1 mm 时对应的额外接触位移量和接触力。由于缺陷长度固定，滚动体“进入”与“离开”缺陷区域事件发生点位置相同，则额外接触位移跨度不会随宽度变化而变化。当缺陷宽度增长时，额外接触位移幅度增大，曲线逐渐陡峭，导致接触力最大值增大，振动加速幅值变大，从而使峰峰值也随之增大。当宽度继续增加，接触模式由模式 1 变为模式 2 之后，额外接触位移量不再发生改变，产生的振动加速度幅值不再改变，峰峰值趋于稳定。

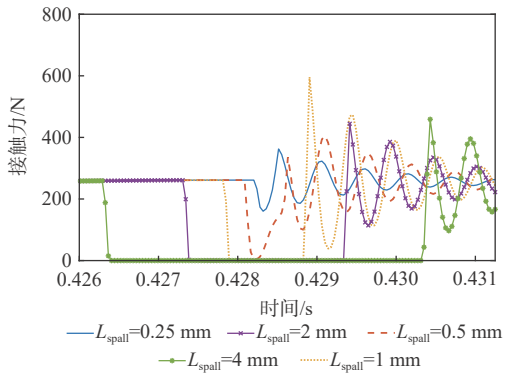
图 16 为缺陷宽度为 2 mm、4 mm 时，长度变化过程的峰峰值曲线及长度为 2 mm 时对应的额外接触位移量和接触力。在缺陷宽度固定的前提下，随着长度增加，峰峰值与缺陷宽度间的关系并不是线性和单调的，整体呈现先递增后递减并趋于平稳的趋势。由于缺陷长度的扩展将影响到额外接触位移量的跨度和幅度，同时在跨度扩展的过程中，滚动体“进入”与“离开”缺陷区域事件发生点处于载荷区的不同位置，所受的接触力大小也不一致，因此，多种因素导致峰峰值与缺陷长度间呈现非线性与非单调的关系。当缺陷长度继续扩展，宽度、深度保持不变，额外接触位移进入模式 3，滚动体通过缺陷区域的过程将更平缓，引起振动响应幅值也趋于



(a) 峰峰值



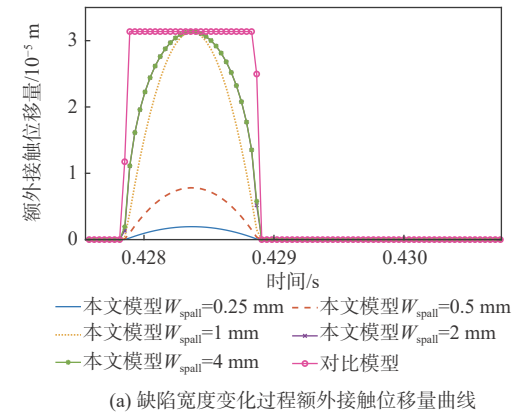
(b) 额外接触位移



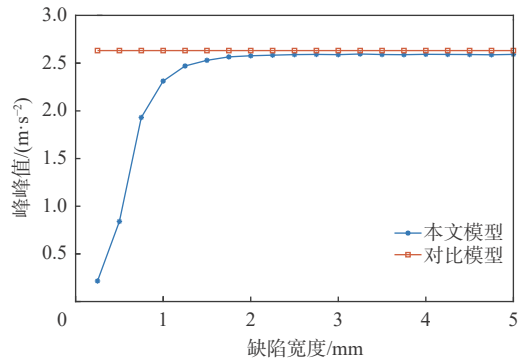
(c) 接触力

图 16　缺陷长度变化过程振动响应分析

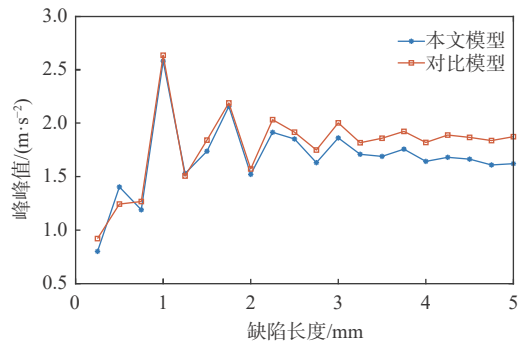
Fig. 16　Vibration response analysis of defect length change process



(a) 缺陷宽度变化过程额外接触位移量曲线



(b) 缺陷宽度变化过程峰峰值曲线



(c) 缺陷长度变化过程峰峰值曲线

图 17　模型对比与分析

Fig. 17　Model comparison and analysis

稳定。

3.5　模型对比分析

为了说明所建模型的优势，选取目前常用的矩形激励函数模型[5,9]与本文模型进行对比。从分析缺陷宽度、长度扩展过程中 2 个模型的峰峰值及额外接触位移量的区别入手，对比 2 种模型刻画缺陷尺寸变化过程的能力。图 17 为缺陷宽度变化过程中额外接触位移量曲线、振动加速度峰峰值曲线及缺陷长度变化过程的振动加速度峰峰值曲线。由图 17(a)可知，当宽度扩展时，传统矩形激励函数无法随宽度的变化而变化，且在“进入”与“离开”缺陷区域时均为阶跃响应，不能够表征滚动体落入缺陷中的连续变化过程，且只包含单一的接触模式。同时，由图 17(b)可知，矩形激励函数模型峰峰值的变化也无法体现缺陷宽度对振动响应的影响。

在缺陷长度变化的过程中，矩形激励函数均以方波激励的形式产生额外接触位移量，仅具有一种接触模式。如图 17(c)所示，随着缺陷长度增加，本文模型的振动响应峰峰值小于对比模型，且差距逐渐变大，即说明随着缺陷长度的扩展，本文模型相比矩形激励函数更能将不同长度缺陷下振动响应的区别体现出来。综上，相比矩形激励函数模型，本文模型更适合用于模拟不同尺寸的剥落缺陷，从而辅助轴承振动响应变化规律的分析。

4　结　论

1）以椭圆形为基础提出了考虑缺陷三维尺寸

的缺陷激励函数，构建了含滚道剥落缺陷的轴承动力学模型。通过分析不同尺寸缺陷引起的额外接触位移量曲线，将滚动体通过缺陷区域的接触模式分为4类，为本文分析缺陷尺寸与振动响应关系建立了基础。

2）以外圈剥落缺陷为例进行模型验证，模型与理论值、实测信号误差均在±5%以内，证明了模型的正确性。

3）利用4个含不同尺寸缺陷的轴承试件开展了实体实验，结合动力学仿真结果分析了实测振动响应发生变化的机理。同时，进一步分析了缺陷宽度、长度扩展过程中振动响应与缺陷尺寸的关系。当缺陷宽度增加时，轴承振动加速度幅值大小会随之递增；而缺陷长度增加的过程中，振动加速度幅值则呈现非线性、非单调的趋势。

4）通过与矩形缺陷激励函数模型对比，本文模型能够根据缺陷三维尺寸得到不同模式的额外接触位移量曲线，并刻画出滚动体经过缺陷区域时额外接触位移量的连续变化过程，更适合用于模拟不同尺寸剥落缺陷轴承的振动响应。

参考文献（References）

[1] 王凡, 马军, 王晓东, 等. 一种改进的TVD电机轴承故障特征提取方法[J]. 振动与冲击, 2023, 42(10): 203-214.
WANG F, MA J, WANG X D, et al. An improved TVD fault feature extraction method for motor bearing[J]. Journal of Vibration and Shock, 2023, 42(10): 203-214(in Chinese).

[2] ZHANG H C, BORGHESANI P, RANDALL R B, et al. A benchmark of measurement approaches to track the natural evolution of spall severity in rolling element bearings[J]. Mechanical Systems and Signal Processing, 2022, 166: 108466.

[3] EPPS I K. An investigation into vibrations excited by discrete faults in rolling element bearings[D]. Christchurch: University of Canterbury, 1991.

[4] SAWALHI N, RANDALL R B. Simulating gear and bearing interactions in the presence of faults Part I. The combined gear bearing dynamic model and the simulation of localised bearing faults[J]. Mechanical Systems and Signal Processing, 2008, 22(8): 1924-1951.

[5] 胡爱军, 许莎, 向玲, 等. 滚动轴承外圈多点故障特征分析[J]. 机械工程学报, 2020, 56(21): 110-120.
HU A J, XU S, XIANG L, et al. Characteristic analysis of multi-point faults on the outer race of rolling element bearing[J]. Journal of Mechanical Engineering, 2020, 56(21): 110-120(in Chinese).

[6] CUI L L, WANG X, WANG H Q, et al. Remaining useful life prediction of rolling element bearings based on simulated performance degradation dictionary[J]. Mechanism and Machine Theory, 2020, 153: 103967.

[7] 张耀强, 陈建军, 唐六丁, 等. 考虑外圈局部缺陷的滚动轴承非线性动力特性[J]. 航空学报, 2009, 30(4): 751-756.
ZHANG Y Q, CHEN J J, TANG L D, et al. Nonlinear dynamic characteristics of rolling element bearing with localized defect on outer ring[J]. Acta Aeronautica et Astronautica Sinica, 2009, 30(4): 751-756(in Chinese).

[8] WU R Q, WANG X F, NI Z X, et al. Dual-impulse behavior analysis and quantitative diagnosis of the raceway fault of rolling bearing[J]. Mechanical Systems and Signal Processing, 2022, 169: 108734.

[9] ZHANG F L, ZHANG Y W, GUAN J Y, et al. Fault dynamic modeling and characteristic parameter simulation of rolling bearing with inner ring local defects[J]. Shock and Vibration, 2021: 5077366.

[10] 胡爱军, 吉新星, 向玲, 等. 滚动轴承非线性时变参数动力学模型与故障机理研究[J]. 机械工程学报, 2022, 58(19): 139-147.
HU A J, JI X X, XIANG L, et al. Nonlinear time-varying parameter dynamic model of rolling bearing and failure mechanism research[J]. Journal of Mechanical Engineering, 2022, 58(19): 139-147(in Chinese).

[11] 常斌全, 刻昌锋, 苑浩, 等. 多事件激励的滚动轴承动力学建模[J]. 振动与冲击, 2018, 37(17): 16-24.
CHANG B Q, YAN C F, YUAN H, et al. Dynamic modeling for rolling bearings under multi-event excitation[J]. Journal of Vibration and Shock, 2018, 37(17): 16-24(in Chinese).

[12] 李志农, 李云龙, 任帅, 等. 局部点蚀故障滚动体的滚动轴承动力学模型研究[J]. 振动工程学报, 2020, 33(3): 597-603.
LI Z N, LI Y L, REN S, et al. Research on dynamic model of rolling bearing with local pitting fault in rolling bearing element[J]. Journal of Vibration Engineering, 2020, 33(3): 597-603(in Chinese).

[13] 李昊泽, 贺雅, 冯坤, 等. 考虑时变激励的滚动轴承局部故障动力学建模[J]. 航空学报, 2022, 43(8): 625176.
LI H Z, HE Y, FENG K, et al. Dynamic modeling of rolling bearing local fault considering time-varying excitation[J]. Acta Aeronautica et Astronautica Sinica, 2022, 43(8): 625176(in Chinese).

[14] SAWALHI N, RANDALL R B. Simulating gear and bearing interactions in the presence of faults. Part Ⅱ: Simulation of the vibrations produced by extended bearing faults[J]. Mechanical Systems and Signal Processing, 2008, 22(8): 1952-1966.

[15] 王震, 杨正伟, 何浩浩, 等. 非规则轴承故障的动力学建模与仿真[J]. 北京航空航天大学学报, 2021, 47(8): 1580-1593.
WANG Z, YANG Z W, HE H H, et al. Dynamic modeling and simulation of irregular bearing failure[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(8): 1580-1593(in Chinese).

[16] 雷春丽, 宋瑞哲, 樊高峰, 等. 含局部缺陷的角接触球轴承时变位移激励及动力学建模[J]. 北京航空航天大学学报, 2025, 51(3): 753-762.
LEI C L, SONG R Z, FAN G F, et al. Study on time-varying displacement excitation and dynamic modeling of local defects in angular contact ball bearings[J]. Journal of Beijing University of Aeronautics and Astronautics, 2025, 51(3): 753-762(in Chinese).

[17] 吴兴国. 数模联合驱动的全寿命滚动轴承数字孪生模型研究[D]. 重庆: 重庆大学, 2022: 12-22.
WU X G. Research on digital twin model of full-life rolling bearing driven by digital-analog combination[D]. Chongqing: Chongqing University, 2022: 12-22(in Chinese).

[18] 罗茂林, 郭瑜, 伍星. 考虑冲击力的球轴承外圈剥落缺陷双冲击现象动力学建模[J]. 振动与冲击, 2019, 38(14): 48-54.
LUO M L, GUO Y, WU X. Dynamic modeling of the dual-impulse behavior produced by a spall on the outer race of a ball bearing considering impact forces[J]. Journal of Vibration and Shock, 2019,

38(14): 48-54(in Chinese).

[19] 丁东升, 刘永强, 杨绍普. 外圈滚道剥离故障对滚动轴承动力学响应的影响分析[J]. 振动与冲击, 2022, 41(8): 141-148.
DING D S, LIU Y Q, YANG S P. Analysis of the influence of outer ring raceway peeling fault on the dynamic response of a rolling bearing[J]. Journal of Vibration and Shock, 2022, 41(8): 141-148(in Chinese).

[20] SAWALHI N, RANDALL R B. Vibration response of spalled rolling element bearings: observations, simulations and signal processing techniques to track the spall size[J]. Mechanical Systems and Signal Processing, 2011, 25(3): 846-870.

[21] HARRIS T A, KOTZALAS M N. Rolling bearing analysis: advanced concepts of bearing technology[M]. 5th ed. New York: Taylor and Francis, 2007: 1-7.

[22] LIU L B, LIANG X H, ZUO M J. Vibration signal modeling of a planetary gear set with transmission path effect analysis[J]. Measurement, 2016, 85: 20-31.

[23] 雷春丽, 宋瑞哲, 樊高峰, 等. 考虑冲击激励的局部缺陷角接触球轴承振动特性[J]. 航空动力学报, 2025, 40(2): 208-218.
LEI C L, SONG R Z, FAN G F, et al. Vibration characteristics of angular contact ball bearing with local defect considering impact excitation[J]. Journal of Aerospace Power, 2025, 40(2): 208-218(in Chinese).

[24] LI X, MA J, WANG X D, et al. An improved local mean decomposition method based on improved composite interpolation envelope and its application in bearing fault feature extraction[J]. ISA Transactions, 2020, 97: 365-383.

[25] CUI L L, ZHANG Y, ZHANG F B, et al. Vibration response mechanism of faulty outer race rolling element bearings for quantitative analysis[J]. Journal of Sound and Vibration, 2016, 364: 67-76.

Dynamic modeling and analysis of rolling bearing with raceway spalling defect

WANG Xiao[1, 2], MA Jun[1, 2, *], XIONG Xin[1, 2], CHEN Jingyu[1, 2]

(1. Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, China;
2. Yunnan International Joint Laboratory of Intelligent Control and Application of Advanced Equipment, Kunming University of Science and Technology, Kunming 650500, China)

Abstract: In order to deeply analyze the mechanism of impact vibration response triggered by bearing raceway spalling defects, the construction of dynamic models becomes a necessary and effective means. The dynamic model of a bearing with a raceway spalling defect is constructed by taking into account the three-dimensional size of the defect. This is done in order to address the issue that the defect excitation function in the current bearing dynamic model finds it challenging to accurately describe the actual spalling defect. Firstly, the three-dimensional morphology of the defect is simulated in an elliptical shape. On this basis, the defect excitation function is constructed and integrated into the dynamic model. Secondly, the variation of the additional contact displacement generated during the passage of the rolling element through the defects of different sizes is analyzed. Then, the correctness of the model is verified by simulation and actual experiments, and the relationship between defect size and vibration response is investigated in depth. Finally, the validity and applicability of the proposed model are demonstrated by comparing it with the rectangular excitation function model. The experimental results show that the error between the model, the theory and the measured signal is within ±5%. The change in defect size will cause the change of additional contact displacement, which will lead to a different vibration response of the bearing. The suggested model is capable of accurately simulating the vibration response of rolling bearings with varying sizes of spalling flaws as compared to the rectangular excitation function model. The relevant research results can provide theoretical support for bearing fault diagnosis and performance evaluation.

Keywords: rolling bearings; raceway spalling; defect size; dynamic modeling; vibration response analysis

Received: 2023-09-04; **Accepted**: 2023-09-22; **Published Online**: 2023-10-23 17:42

URL: link.cnki.net/urlid/11.2625.V.20231023.1423.004

Foundation items: National Natural Science Foundation of China (62163020,62173168); Yunnan Fundamental Research Projects (202101BE070001-055); Major Project of Science and Technology of Yunnan Province (202202AD080005)

* **Corresponding author**. E-mail: 491941203@qq.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0513

基于 LSTM-DDPG 算法的四翼变掠角飞行器主动变形决策

彭余萧，何真*，仇靖雯
(南京航空航天大学 自动化学院，南京 211106)

摘 要：针对变体飞行器主动变形控制问题，提出一种基于长短期记忆 (LSTM) 网络深度确定性策略梯度 (DDPG) 算法的智能变形控制方法；以一种串置翼构型的四翼变掠角飞行器为研究对象，利用 OPENVSP 软件计算其几何模型和气动参数，并建立了飞行器动力学模型；针对四翼变掠角飞行器的加速爬升过程，设计了基于 LSTM-DDPG 算法学习框架，并在对称变形条件下，针对纵向轨迹跟踪进行主动变形决策训练。仿真结果表明：应用于主动变形控制过程中的 LSTM-DDPG 算法可以快速收敛并达到更高的平均奖励，且训练获得的主动变形控制器在四翼变掠角飞行器的轨迹跟踪任务中具有良好的控制效果。

关 键 词：变体飞行器；飞行控制；深度强化学习；深度确定性策略梯度；长短期记忆递归神经网络

中图分类号：V249.1

文献标志码：A **文章编号**：1001-5965(2025)10-3504-11

变体飞行器可以通过改变气动布局，使得飞行器在应对各种飞行阶段、飞行环境时均能保持最佳的飞行效能[1]；也可利用变形机构来辅助甚至取代传统舵面，用以简化飞行器结构并提高操纵性[2]。国内外对变体飞行器的研究多集中于前者，而对于将变形部件取代传统舵面作为飞行器操纵机构的研究相对不足。

变体飞行器的主动变形控制问题是本文研究的重点。主动变形控制是指根据不同的飞行状态，通过机翼形变产生的力、力矩及气动变化，代替传统舵面达到控制飞行器完成飞行任务的目的。但由于变体飞行器的动力学存在强非线性、强时变和不确定性等问题，传统非线性控制算法的设计和计算都较为复杂而困难。越来越多的研究集中于智能控制方法。文献 [3] 针对参数不确定的变后掠角飞行器的纵向动力学问题提出一种新的自适应神经网络反步控制器。文献 [4] 研究了一种基于强化学习的变后掠翼飞行器的运动切换控制策略。文献 [5] 设计了一种改进的深度确定性策略梯度算法，通过引入优先级扫描算法来优化变形策略。文献 [6] 采用解码器方法减小强化学习网络规模。文献 [7] 基于先验知识生成训练数据，训练深度神经网络作为自主变形知识库。其中，强化学习算法是变体飞行器主动变形控制的较优方案。

传统的强化学习方法在处理复杂的飞行器控制问题时，往往面临着维数爆炸、无法处理连续动作空间等挑战。为此，深度强化学习的概念于 2013 年被 Deepmind 团队提出[8]，深度 Q 学习网络 (deep Q-learning network, DQN) 及其派生算法双深度 Q 学习网络[9](double Q-learning network, double DQN) 算法、对偶深度 Q 学习网络[10](dueling deep Q-learning network, dueling DQN) 算法，以及可信域策略优化[11]

收稿日期：2023-08-08；**录用日期**：2023-10-23；**网络出版时间**：2023-11-21 15:04
网络出版地址：link.cnki.net/urlid/11.2625.V.20231120.1506.001
基金项目：国家自然科学基金 (61873126)
* **通信作者**. E-mail：hezhen@nuaa.edu.cn

引用格式：彭余萧，何真，仇靖雯. 基于 LSTM-DDPG 算法的四翼变掠角飞行器主动变形决策 [J]. 北京航空航天大学学报，2025，51（10）：3504-3514. PENG Y X，HE Z，QIU J W. Active deformation decision-making for four-wing variable sweep aircraft based on LSTM-DDPG algorithm [J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51（10）：3504-3514（in Chinese）.

(trust region policy optimization, TRPO) 算法, 近端策略优化[12](proximal policy optimization, PPO) 等算法在各类控制环境中显示了优越的性能。其中, 深度确定性策略梯度(Deep deterministic policy gradient, DDPG)[13] 算法是一种基于异步优势演员-评论家[14] (asynchronous advantage actor-critic, A3C) 框架的深度强化学习算法, 可以处理连续的动作空间问题。然而, DDPG 算法不能很好地处理具有复杂动态特性的问题, 难以理解飞行器产生的时间序列数据中的模式和依赖关系。

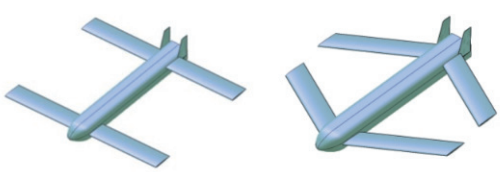
为解决这个问题, 本文在 DDPG 算法基础上引入了长短期记忆 (long short-term memory, LSTM 递归神经网络[15]), 提出一种 LSTM-DDPG 算法。LSTM 是一种特殊的递归神经网络 (recurrent neural network, RNN), 其可以处理时间序列依赖性的问题[16]。通过将 LSTM 引入到 DDPG 算法中, 可以更好地处理变体飞行器控制问题中的时间序列依赖性问题。

本文以一类新型的串置翼结构的四翼变掠角飞行器为研究对象, 提出并实现一种基于 LSTM-DDPG 算法的变体飞行器主动变形控制策略, 能够使飞行器在无副翼、无平尾的情况下有效地跟踪预定轨迹, 同时对环境变化具有良好的适应性。这项研究为飞行器的主动变形控制提供了一种新的、有效的解决方案。

1　变体飞行器动力学分析

1.1　四翼变掠角飞行器的变体构型

本文的研究对象为一种具备串置翼结构的四翼变掠角飞行器, 其由中央机体、2 对翼及垂直尾翼构成, 如图 1 所示。4 个主翼主动进行变掠角动作, 改变飞机气动分布, 从而达到操纵飞机的作用[17]。在各类变形策略中, 掠角对称变形动作简单、稳定性强, 通过同步调整左右翼的掠角, 从而对飞行器纵向运动产生影响, (见图 1)。具体而言, 该变体动作通过改变升力分布来产生俯仰力矩, 从而控制飞行器的俯仰姿态。本文的研究重点是四翼变掠角飞行器在缺少副翼和水平尾翼的情况下, 如何通过建模和控制策略实现其纵向飞行的稳定性。



(a) 掠角为0°　　(b) 掠角为30°

图 1　掠角对称变形下飞行器形态

Fig. 1　Aircraft configuration under sweepback symmetrical deformation

在起飞阶段, 该飞行器采用无掠角的状态, 如图 1(a) 所示, 主要适用于起飞和低速巡航任务。而在高速飞行中, 如图 1(b) 所示, 前后翼均对称调整至一定掠角, 以此减小翼展并相应降低阻力。

四翼变掠角飞行器的控制输入定义为其 4 个主翼的掠角变化, 即各机翼坐标系相对于机体坐标系围绕 z 轴的转动角度[18], 如图 2 所示。取轴 x_i 平行于机翼坐标系平面 Ox_by_b, 轴 y_i 通过机翼转动中心, 指向飞行器右方, 每个主翼的掠角分别用 a_1、a_2、a_3、a_4 表示。

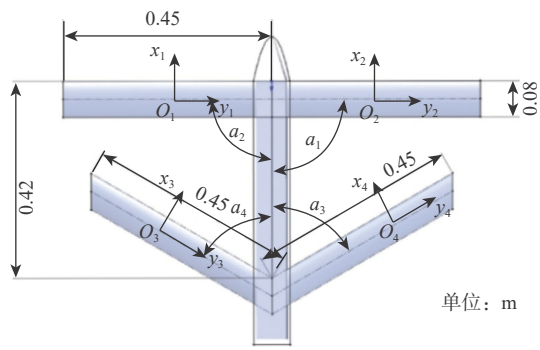


图 2　飞行器掠角控制输入定义

Fig. 2　Definition of sweep angle control inputs

取顺时针为正值, 所以在对称变形时 $a_1 = -a_2 < 0, a_3 = -a_4 < 0$, 且 $|a_i| \in [0°, 30°]$。将控制量简化表示为变形率 $\lambda_i \in [0,1]$[19]:

$$\lambda_i = \begin{cases} 1 & a_i > 30° \\ a_i/30, & 0 \leqslant a_i \leqslant 30° \\ 0 & a_i < 0° \end{cases} \tag{1}$$

式中: $i = 1,2$, 分别表示前翼和后翼。

表 1 列出了四翼变掠角飞行器的各项关键参数。

表 1　四翼变掠角变体飞行器关键参数

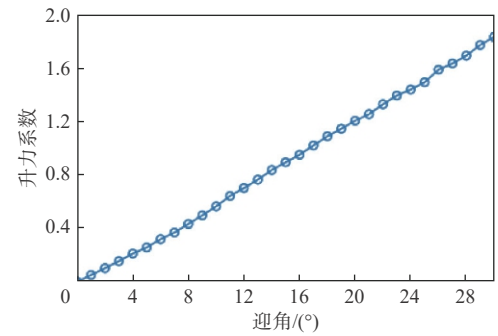
Table 1　Key parameters of quad-wing variable sweep angle variant aircraft

参数	数值
飞行器总质量m/kg	1.668
质心与机翼旋转中心间距l/m	0.140
面Oxz与前翼旋转中心间距a_c/m	0.165
面Oxz与后翼旋转中心间距a_w/m	0.235
面Oxz与机翼旋转中心间距b_l/m	0.040
机翼质量m_a/kg	0.080
翼展b/m	0.893
平均气动弦长c_a/m	0.077
前后翼间距S_l/m	0.420
翼面积S_m/m^2	0.135

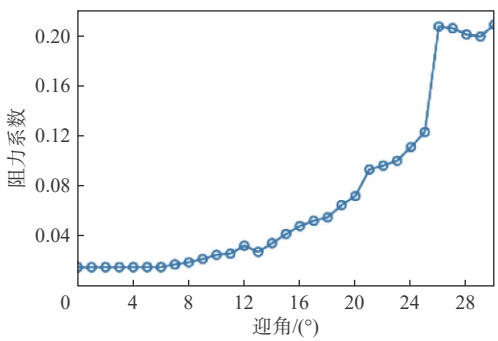
1.2　气动特性分析

本文中四翼变掠角飞行器的气动分析将采用 OPENVSP 软件。基于表 1 所述参数, 在 OPENVSP

软件中构建了该飞行器的几何模型，同时选用NACA 0012型号的翼型。仿真飞行条件包括迎角、马赫数、雷诺数、重心、参考面积，选取掠角作为变量，取若干不同数值计算气动参数。图3给出了掠角为0°，马赫数为0.06，雷诺数为10 000条件下升力系数和阻力系数随迎角的变化。



(a) 升力系数与迎角的关系



(b) 阻力系数与迎角的关系

图3 气动系数与迎角的关系

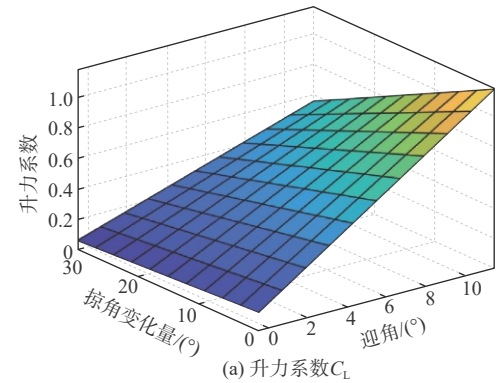
Fig. 3 Relationship between aerodynamic coefficients and angle of attack

由于掠角改变会导致气动力的非线性变化，在不同掠角下进行多次仿真后，获得了掠角变化对纵向气动系数影响的非线性数学模型：

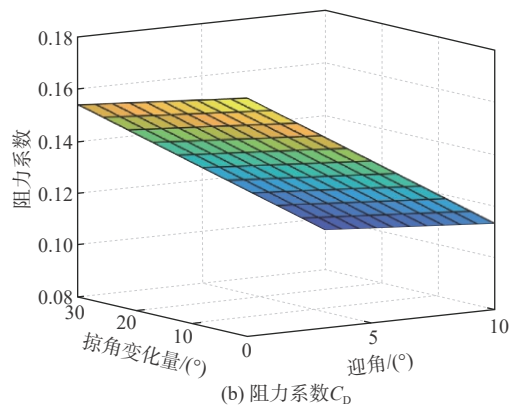
$$\begin{cases} C_L = 0.01(47.95 - 4.077\lambda_1^2 - 4.579\lambda_2^2 + 16.89\lambda_1^2\lambda_2 + \\ \quad 17.44\lambda_1\lambda_2^2 - 20.41\lambda_1\lambda_2 - 16.23\lambda_1^2\lambda_2^2)(9.448\alpha + 0.339\,7) \\ C_D = 0.001(83.58 - 5.229\lambda_2 - 0.129\,6\lambda_2^2 - 4.34\lambda_1 - \\ \quad 9.409\lambda_2 + 3.595\lambda_1\lambda_2)(28.85\alpha^2 + 0.236\,3\alpha + 0.842\,9) \\ C_m = 0.01(-8.103 - 11.67\lambda_1^2 - 4.525\lambda_2^2 - \\ \quad 27.26\lambda_1 + 37.47\lambda_2)(-8.207\alpha^2 + 10.03\alpha + 0.34) + \\ \quad (7.248\lambda_1 + 36.89\lambda_2 - 69.24)q \end{cases} \tag{2}$$

式中：C_L为飞行器升力系数；C_D为阻力系数；C_m为俯仰力矩系数；α为迎角；q表示俯仰角速率。

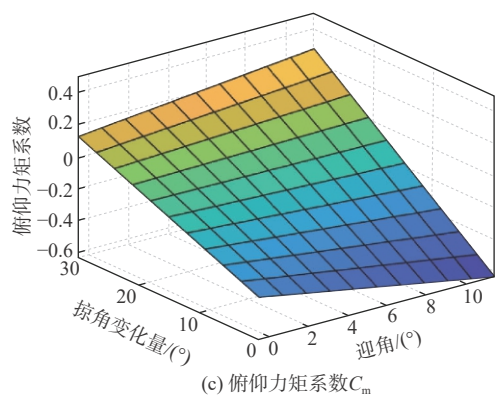
图4给出了当飞行器的侧向力、滚转力矩和偏航力矩均为0时，掠角角度和迎角变化幅度对飞行器俯仰力矩系数、阻力系数、升力系数的影响。该图表明，升力系数对迎角较为敏感，且与迎角呈正相关，小迎角下与掠角变化相关性较小；阻力系数随迎角变化较小，与迎角、掠角变化均呈负相关。俯仰力矩受掠角影响较大，随着掠角增大而增大。当迎角$\alpha = 6°$时，该对称变形产生的最大抬头力矩为0.316 N·m，最大低头力矩为−0.618 N·m。



(a) 升力系数C_L



(b) 阻力系数C_D



(c) 俯仰力矩系数C_m

图4 变形量和迎角对气动系数的影响

Fig. 4 Effects of deformation and angle of attack on aerodynamic coefficients

可见，掠角变化对升力和阻力的影响相对较小，但对力矩影响较大，表明将掠角变化作为控制量，作机翼掠角对称变形以控制飞行器俯仰运动的方案是有意义的。

1.3 变体飞行器动力学建模

飞行器纵向动力学方程可表示为

$$\dot{\boldsymbol{X}} = f(\boldsymbol{X}, \boldsymbol{u}) \tag{3}$$

式中：f为非线性函数模型；$\boldsymbol{X}$和$\boldsymbol{u}$分别为状态量和控制量。

掠角对称变形策略下飞行器主要通过同步调

整 2 对翼掠角来控制飞行姿态, 同时辅以发动机推力变化共同控制飞行。因此, 式 (3) 中的控制输入 $\boldsymbol{u}=[\lambda_1,\lambda_2,T]$包含 2 对掠角 λ_1、λ_2和推力 T; 输出状态量 $\boldsymbol{X}=[h,q,V,\mu,\alpha]$包括飞行高度 h、迎角 α等信息。定义升力 L、阻力 D 和俯仰力矩 M 为

$$\begin{cases} L=\dfrac{1}{2}\rho V^2 S C_L \\ D=\dfrac{1}{2}\rho V^2 S C_D \\ M=\dfrac{1}{2}\rho V^2 S c_a C_M \end{cases} \tag{4}$$

式中: ρ为空气密度; V为飞行速度; S 为机翼面积。本文聚焦于飞行器的起飞阶段和低速平飞的纵向运动过程, 不考虑飞行器横侧向运动。在机翼掠角对称变形过程中, 飞行器的质心和焦点随机翼变形而发生位移, 进而产生俯仰力矩。此外, 机翼在改变掠角的同时, 也会产生额外的惯性力 $F_{x\mathrm{dyn}}$、$F_{z\mathrm{dyn}}$及附加力矩 M_{dyn}。假设横侧向的姿态均为 0, 即侧滑角、滚转角和偏航角均满足等于零的条件下, 得到动力学方程式(3)的具体形式为

$$\begin{cases} \dot{h}=V\sin\mu \\ \dot{q}=(M_{\mathrm{dyn}}+M)/J_y \\ \dot{V}=(-D-mg\sin\mu+F_{x\mathrm{dyn}}\cos\alpha+F_{z\mathrm{dyn}}\sin\alpha+ \\ \qquad T\cos\alpha+2m_{\mathrm{a}}(a_{\mathrm{c}}-a_{\mathrm{w}})\sin\alpha\dot{q})/m \\ \dot{\mu}=(L-mg\cos(-\mu)-F_{z\mathrm{dyn}}\cos\alpha+F_{x\mathrm{dyn}}\sin\alpha+ \\ \qquad T\sin\alpha-2m_{\mathrm{a}}(a_{\mathrm{c}}-a_{\mathrm{w}})\cos\alpha\dot{q})/mV \\ \dot{\alpha}=q-\dot{\mu} \end{cases} \tag{5}$$

式中: J_y为转动惯量; M_{dyn}为由对称变形引起的附加惯性俯仰力矩; $F_{x\mathrm{dyn}}$和$F_{z\mathrm{dyn}}$分别为飞行器在 x和 z轴方向的附加变形力, 其表达式为

$$\begin{cases} F_{z\mathrm{dyn}}=mg\cos\theta-D\sin\alpha-L\cos\alpha \\ F_{x\mathrm{dyn}}=2m_{\mathrm{a}}l(\dot{a}_1^2\cos a_1-\ddot{a}_1\cos a_1+\dot{a}_3^2\cos a_3-\ddot{a}_3\cos a_3) \\ J_y=0.024\,2+4\times2.6\times10^{-5}+4m_{\mathrm{a}}c^2+ \\ \qquad m_{\mathrm{a}}(l_1^2+l_2^2+l_3^2+l_4^2) \\ M_{\mathrm{dyn}}=2m_{\mathrm{a}}ql(\dot{a}_3l_2\cos a_3-\dot{a}_1l_1\cos a_1)+ \\ \qquad 2m_{\mathrm{a}}l(q\sin\theta+F_z/m)(\sin a_1+\sin a_3)- \\ \qquad 2m_{\mathrm{a}}cl(\dot{a}_1\cos a_1-\ddot{a}_3\cos a_3-\dot{a}_1^2\sin a_1+\dot{a}_3^2\sin a_3) \end{cases} \tag{6}$$

考虑模型的不确定性, 具体地表现为升力系数的参数不确定性, 写成如下形式:

$$\begin{aligned} C_L=&0.01(47.95-4.077\lambda_1^2-4.579\lambda_2^2+ \\ &16.89\lambda_1^2\lambda_2+17.44\lambda_1\lambda_2^2-20.41\lambda_1\lambda_2- \\ &16.23\lambda_1^2\lambda_2^2)(9.448\alpha+0.339\,7)+\varepsilon \end{aligned} \tag{7}$$

式中: ε为不确定参数且界为常值。

采用鸽群优化 (pigeon-inspired optimization, PIO) 算法对飞行器爬升过程进行参考轨迹优化设计, 规定飞行器的初始状态为 $\boldsymbol{X}_0=[198;-0.004;16;0.126;0.129]$, 末端状态为 $\boldsymbol{X}_{\mathrm{end}}=[250;0;37;0;0]$, 选取种群数量 N 为 100, 优化维度 n 为 20。通过设定评估标准为最快达到成功位置, 建立了一个效能评价函数 $f_{\mathrm{itness}}(t)$。该函数利用 2 个主要的更新机制——方向感知机制和地形分析机制, 以在不同阶段调整粒子的状态。最终得到优化结果$\{\boldsymbol{X}_{\mathrm{r}},\boldsymbol{u}_{\mathrm{r}}\}$, 即为爬升平飞段的参考轨迹。

对于给定的参考轨迹 $\boldsymbol{X}_{\mathrm{r}}=[h_{\mathrm{r}},q_{\mathrm{r}},V_{\mathrm{r}},\mu_{\mathrm{r}},\alpha_{\mathrm{r}},l_{\mathrm{r}}]$和 $\boldsymbol{u}_{\mathrm{r}}=[\lambda_{1\mathrm{r}},\lambda_{2\mathrm{r}},T_{\mathrm{r}}]$, 可以进一步得到变体飞行器的增量模型:

$$\Delta\dot{\boldsymbol{X}}=f(\Delta\boldsymbol{X}+\boldsymbol{X}_{\mathrm{r}},\Delta\boldsymbol{u}+\boldsymbol{u}_{\mathrm{r}})-f(\boldsymbol{X}_{\mathrm{r}},\boldsymbol{u}_{\mathrm{r}}) \tag{8}$$

式中: $\Delta\boldsymbol{X}=\boldsymbol{X}-\boldsymbol{X}_{\mathrm{r}}$; $\Delta\boldsymbol{u}=\boldsymbol{u}-\boldsymbol{u}_{\mathrm{r}}$。

2　基于 LSTM-DDPG 算法的主动变形控制

变体飞行器在完成飞行任务的过程中, 飞行状态与变形动作均为连续的过程。因此, 本文采用 DDPG 算法处理连续动作空间问题, 结合 LSTM 网络[15] 的独特记忆单元, 增强变体飞行器飞行过程中对环境变化作出实时应对的能力。同时, 为解决 DDPG 算法训练时动作空间大、收敛速度慢的问题, 在强化学习的仿真环境中加入参考轨迹, 将动作空间定义为当前动作和参考轨迹的误差(基于参考轨迹的增量动力学模型(见式 (8))), 从而减小动作搜索空间。

2.1　爬升轨迹跟踪控制的 MDP 过程建模

2.1.1　动作空间与状态空间选取

状态空间由当前的飞行状态变量组成, 选取俯仰角速率q、飞行速度V、迎角α、航迹角μ、飞行高度h及水平位移s为状态变量, 状态变量在训练的过程中逐步更新直至策略收敛。动作空间为机翼掠角$[\lambda_1,\lambda_2]$与参考轨迹$[\lambda_{1\mathrm{r}},\lambda_{2\mathrm{r}}]$之差。

2.1.2　奖励函数设计

为加快收敛速度、引导飞行任务完成, 采用稀疏奖励和非稀疏奖励[20] 相结合的方式设计一种复合奖励函数。奖励函数中的各项设计如下:

1) 稀疏奖励项r_{c}表示飞行器爬升终点的约束量, 即飞行器的终点状态在允许的误差范围内。r_{c}设计为

$$\begin{aligned} r_{\mathrm{c}}=&(t<t_{\mathrm{f}})\&\&(|h-h_{\mathrm{target}}|\leqslant\delta_h)\&\&(|q|\leqslant\delta_q)\&\& \\ &(|V-V_{\mathrm{target}}|\leqslant\delta_v)\&\&(|\mu-\mu_{\mathrm{target}}|\leqslant\delta_\mu) \end{aligned} \tag{9}$$

式中: 符号 && 为逻辑运算符与; t_{f}为完成任务的最大时间; h_{target}、V_{target}和q_{target}分别为飞行高度、飞行速度和俯仰角速率的目标值; δ_h、δ_v、δ_μ和δ_q分别为高度、速度、航迹倾斜角和俯仰角速率的最大容许

误差范围。

2）稀疏奖励项r_o为飞行器爬升过程的约束量，惩罚飞行器在爬升过程中超过状态阈值。

$$\begin{cases} r_o = t > t_f \| h > h_{max} \| h < h_{min} \| \\ V > V_{max} \| V < V_{min} \| q > q_{max} \| \\ \mu > \mu_{max} \| \mu < \mu_{min} \| \alpha > \alpha_{max} \end{cases} \tag{10}$$

式中：符号‖为逻辑运算符或；h_{max}和h_{min}分别为飞行高度约束的最大值和最小值；V_{max}和V_{min}分别为基于飞行器本身特性的飞行速度约束的最大值和最小值；q_{max}为容许的最大俯仰角速率；α_{max}为容许的最大迎角值；μ_{min}和μ_{max}分别为航迹角约束的最小值和最大值。

3）将上述稀疏奖励项综合，设计约束奖励项R_f为

$$R_f = 600r_c - 100r_o \tag{11}$$

4）非稀疏动作奖励函数R_a表示对动作平滑程度的奖励项。机翼的频繁转动会对驱动器和机械结构产生负担，故设计R_a鼓励控制量平滑变化，以减小机翼掠角变化产生的抖动，R_a设计为

$$R_a = w_3 |\Delta a|^2 \tag{12}$$

式中：Δa为动作的变化量；w_3为负值，引导动作λ_1和λ_2缓慢变化。

5）非稀疏状态奖励函数R_m表示对飞行器状态靠近目标轨迹的奖励。该奖励项基于视线法 (line of sight, LoS)[21] 设计，假设在解耦的垂直平面跟踪路径，以飞行器为圆心，半径为R的圆与目标路径相交，取接近速度方向的交点(x_d, h_d)如图5所示，可得目标的航向角角度$\Delta\psi = \arctan(\sqrt{R^2 - x_d^2}/x_d)$。将当前时刻飞行器位置向量定义为$\boldsymbol{X}_{UAV} = [s_1, h_1]$，目标位置向量定义为$\boldsymbol{X}_{target} = [s_2, h_2]$。则相对距离$D$和航向角$\psi$可以定义为

$$\begin{cases} D = \sqrt{(s_1 - s_2)^2 + (h_1 - h_2)^2} \\ \psi = \arctan \dfrac{h_1 - h_2}{s_1 - s_2} \end{cases} \tag{13}$$

设计状态奖励函数R_m为

$$R_m = 0.8e^{-5\times10^{-4}D} + 0.2e^{-2\times10^{-3}\psi} \tag{14}$$

综合上述各项奖励得到总体奖励函数R_{total}，以满足对目标约束、过程约束、跟踪性能及变形稳定性的评价。总体奖励函数R_{total}设计为

$$R_{total} = R_f + R_a + R_m \tag{15}$$

2.1.3 马尔可夫过程模型

本文算法的MDP流程涵盖状态S、动作A、奖励R和折扣率γ等元素。飞行器在t时刻得到状态$\boldsymbol{s}_t \in S$，基于该状态选择动作$a_t \in A$并执行，获得奖励

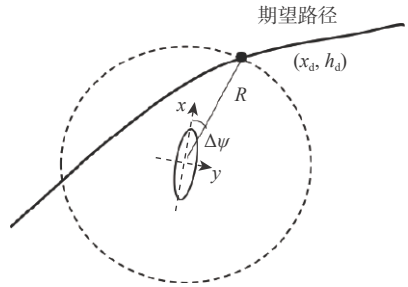


图5 LoS路径跟踪方法原理图

Fig. 5 LoS path tracking method

$r_{t+1} \in R$；状态转移由动作和动力学模型共同完成，在$t+1$时刻得到状态$\boldsymbol{s}_{t+1}$。终止条件设为飞行器成功到达目标状态或超出某状态阈值。

在策略优化类型的方法中，考虑确定性情况，则飞行器爬升任务过程模型的优化目标如下，代表最大化策略收益：

$$J_G(\pi) = E_{\tau\sim p(\tau|\pi)}\left[\sum_{k=0}^{T_{step}} \gamma^k R_{total\,t+k+1}\right] \tag{16}$$

式中：$\tau = \{(s_0, a_0, r_1), (s_1, a_1, r_2), \cdots, T_{step}\}$为从状态$\boldsymbol{s}_0$出发，基于策略$\pi$的一系列状态动作对。

2.2 LSTM-DDPG算法框架

LSTM网络是一种改进的RNN，其算法框架如图6所示，其中，x_t为输入，σ和tanh分别为sigmoid函数和双曲正切函数，f_t、i_t和o_t分别为遗忘门、输入门和输出门的运算结果。c_t'为当前储存单元的候选者，c_t为更新后的单元状态，h_t为隐藏层的状态。$\boldsymbol{w}_f$、$\boldsymbol{w}_i$、$\boldsymbol{w}_c$、$\boldsymbol{w}_o$分别为各部分的权重矩阵，$\boldsymbol{b}_f, \boldsymbol{b}_i, \boldsymbol{b}_c, \boldsymbol{b}_o$分别为遗忘门、输入门、候选单元、输出门的偏置向量。因此，c_t和h_t分别为

$$\begin{cases} c_t = f_t * c_{t-1} + i_t * c_t' \\ h_t = o_t * \tanh(c_t) \end{cases} \tag{17}$$

其中，

$$\begin{cases} f_t = \sigma(\boldsymbol{w}_f \cdot [h_{t-1}, x_t] + \boldsymbol{b}_f) \\ i_t = \sigma(\boldsymbol{w}_i \cdot [h_{t-1}, x_t] + \boldsymbol{b}_i) \\ c_t' = \tanh(\boldsymbol{w}_c \cdot [h_{t-1}, x_t] + \boldsymbol{b}_c) \\ o_t = \sigma(\boldsymbol{w}_o \cdot [h_{t-1}, x_t] + \boldsymbol{b}_o) \end{cases} \tag{18}$$

在LSTM-DDPG算法中，首先，建立带有LSTM层的Actor和Critic网络，以处理飞行器飞行过程中具有时间序列关联性的问题，新网络结构如图7所示。actor网络的输入层由飞行器的状态变量和高度误差组成，由七维向量$\boldsymbol{s}_t$表征。该网络的隐藏层由2层LSTM网络和1层全连接网络组成，LSTM网络包含128个神经元，用以学习序列特征，增强实时性[22]。全连接层包含256个神经元。选用ReLU函数作为隐藏层之间的激活函数，规避梯度

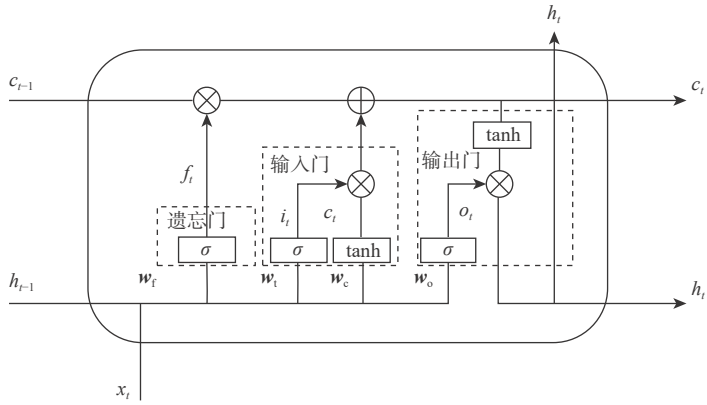


图 6　LSTM 算法框架

Fig. 6　LSTM algorithm framework

消失问题[23]。输出层为二维的动作空间$[\delta_{\lambda_1},\delta_{\lambda_2}]$, 选用 tanh 激活函数以将输出限制在指定范围内。

critic 网络的输入层由状态量$s_t=[q,V,\mu,\alpha,h,s,\Delta h]$以及 actor 网络的输出动作$[\delta_{\lambda_1},\delta_{\lambda_2}]$组成, 合并为一个 9 维数组。隐藏层由 2 层 128 神经元的 LSTM 网络和 1 层 200 神经元的全连接网络组成。输出层仅有 1 个神经元, 表征动作价值函数$Q(s,a)$。选用 ReLU 函数作为各层之间的激活函数。

翼变掠角飞行器 LSTM-DDPG 算法框架如图 8 所示。

传统 DDPG 算法在大的状态与动作空间中探索效率较低, 易陷入局部最优解, 且训练耗时长。为此, 本文算法引入基于行为克隆 (behavior cloning, BC) 的预训练框架。BC 算法所需的专家数据由标称模型得出的参考轨迹组成, 这些轨迹包含不同的初始条件, 由 PIO 算法迭代得出, 代表该环境下的优秀数据。以专家数据为样本, 采用经典 AC 算法作离线训练, 得到预训练策略网络和价值 Q 网络。这样, 一方面能够提高算法的搜索效率, 另一方面

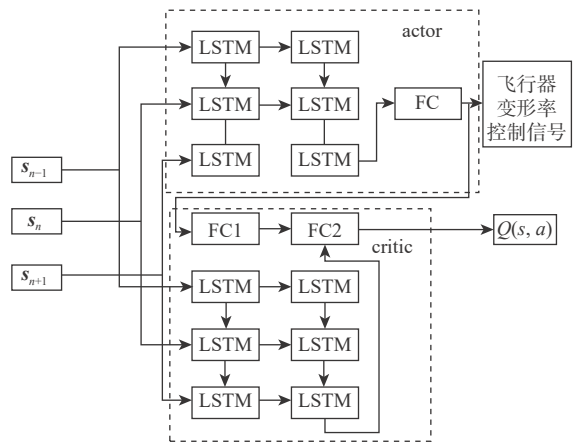


图 7　LSTM-DDPG 算法网络结构

Fig. 7　LSTM-DDPG algorithm network architecture

BC 算法与 DDPG 算法相结合能够面对更多的状态信息及状态信息的变化。

在算法的第 2 阶段, 将预训练得到的初始策略网络参数输入到 actor 网络中, 进行新一阶段的在线训练。由于 LSTM-DDPG 是一种深度确定性策略方法, 在 actor 网络的输出动作中加入奥恩斯坦-

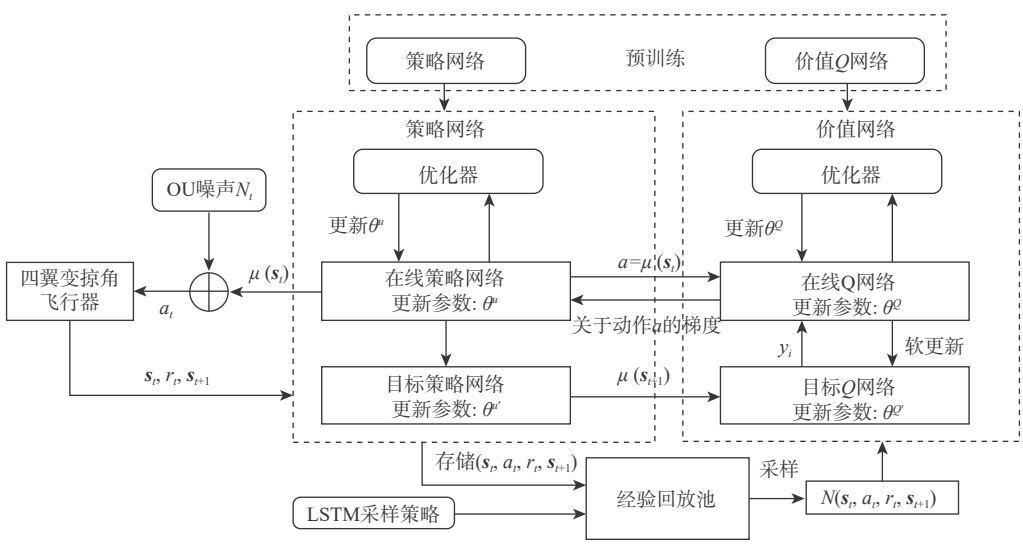


图 8　四翼变掠角飞行器 LSTM-DDPG 算法学习框架

Fig. 8　LSTM-DDPG algorithm learning framework for four-wing variable sweep angle aircraft

乌伦贝克(Ornstein-Uhlenbeck, OU)噪声$\mathcal{N}_t$以增强策略的探索性。对每个时间步,根据当前的状态$\boldsymbol{s}_t$和噪声选择动作a_t,执行动作,观察奖励和新的状态,并将这些数据存入经验回放池(replay buffer, RB)。当经验回放池内的数据达到训练阈值时,取N个小批量数据来更新actor网络和critic网络的参数。当前critic网络的参数更新采用时间差分算法,最小化损失函数$\mathcal{L}$的均方误差:

$$\mathcal{L}=\frac{1}{N}\sum_i(y_i-Q(s_i,a_i|\theta^Q))^2 \tag{19}$$

当前actor网络的参数通过计算采样的策略梯度更新:

$$\nabla_{\theta^\mu}\mathcal{J}\approx\frac{1}{N}\sum_i\nabla_a Q(s_i,a|\theta^Q)|_{a=\mu(s_i)}\nabla_{\theta^\mu}\mu(s_i|\theta^\mu) \tag{20}$$

为保证训练过程稳定,LSTM-DDPG算法使用了目标网络的框架,目标网络$Q(s,a|\theta^{Q'})$和$\mu(s|\theta^{\mu'})$的参数更新采用软更新,由当前网络$Q(s,a|\theta^Q)$和$\mu(s|\theta^\mu)$的参数加权平均运算得出:

$$\begin{cases}\theta^{Q'}\leftarrow\tau\theta^Q+(1-\tau)\theta^{Q'}\\ \theta^{\mu'}\leftarrow\tau\theta^\mu+(1-\tau)\theta^{\mu'}\end{cases} \tag{21}$$

其中,目标网络的参数更新速率受更新率τ限制,增加学习过程的稳定性。

综上,LSTM-DDPG总体算法流程如下所示。

输入:专家数据$\{s_t,s_{t+1},r_t,a_t,a_{t+1}\}$。

输出:目标网络。

预训练阶段:

1: 收集专家数据$\{s_t,s_{t+1},r_t,a_t,a_{t+1}\}$;

2: 随机初始化预训练Q网络$Q(s,a|\theta^Q_{\text{pre}})$

3: 随机初始化策略网络$\mu(s|\theta^\mu_{\text{pre}})$;

4: 离线训练专家数据;

正式训练阶段:

5: 随机初始化Q网络$Q(s,a|\theta^Q)$;

6: 初始化策略网络$\mu(s|\theta^\mu)$:$\theta^\mu\leftarrow\theta^\mu_{\text{pre}}$

7: 初始化目标网络:$\theta^{Q'}\leftarrow\theta^Q,\theta^{\mu'}\leftarrow\theta^\mu$

8: for episode = 1:MaxEpisode do:

9: 初始化随机噪声$\mathcal{N}$,获取初始状态s_0

10: for t = 1:T do:

11: 选择动作$a_t=\mu(s_t|\theta^\mu)+\mathcal{N}_t$;

12: 执行动作a_t;

13: 得到奖励r_t和新状态s_{t+1};

14: 将$(\boldsymbol{s}_t,a_t,r_t,\boldsymbol{s}_{t+1})$存进经验回放池RB;

15: 从RB中随机批量序列采样:

$$\left\{\begin{matrix}(s_{i-1},a_{i-1},r_{i-1},s_i)\\(s_i,a_i,r_i,s_{i+1})\\(s_{i+1},a_{i+1},r_{i+1},s_{i+2})\end{matrix}\right\}\quad i=1,2,\cdots,N$$

16: 用目标网络计算:

$y_i=r_i+\gamma Q'(s_{i+1},\mu'(s_{i+1}|\theta^{\mu'})|\theta^{Q'})$

17: 根据最小损失更新Q网络参数:

$\mathcal{L}=\frac{1}{N}\sum_i(y_i-Q(s_i,a_i|\theta^Q))^2$

$\theta^Q=\theta^Q+\tau\nabla_{\theta^Q}\mathcal{L}$

18: 根据策略梯度更新策略网络参数:

$\nabla_{\theta^\mu}\mathcal{J}\approx\frac{1}{N}\sum_i\nabla_a Q(s_i,a|\theta^Q)|_{a=\mu(s_i)}\nabla_{\theta^\mu}\mu(s_i|\theta^\mu)$

19: 软更新目标网络参数:

$\theta^{Q'}\leftarrow\tau\theta^Q+(1-\tau)\theta^{Q'}$

$\theta^{\mu'}\leftarrow\tau\theta^\mu+(1-\tau)\theta^{\mu'}$

20: end for t

21: end for episode

值得注意的是,引入LSTM之后,网络的输入和输出不再是单个状态或动作,而是一定长度的连续数据,因此,训练数据不能从经验回放池RB中随机采样。本文采用一种随机批量序列采样方法,将经验回放池RB中3个连续的状态转移数据作为一组训练单元;当RB中数据量满足训练阈值后,算法会从RB中随机采样N组训练单元,用于网络的一批次训练输入。这种设计可以有效地捕获飞行器控制过程中的时间相关性,提高模型对环境的理解和反应能力。

在实际应用中,控制器即训练得到的目标actor网络,负责在线计算每步状态到动作的映射,作为系统控制量输入。同时系统会继续收集实际飞行数据,以便进一步训练控制器。由于神经网络前向传播的运算量远小于反向传播,基于前向传播的控制量运算耗时极短,可以满足本环境中每0.02 s计算一次的工程需要。

3 仿真与分析

3.1 算法参数设置

状态空间取值为飞行状态实际值。动作空间取值为$a=[\delta_{\lambda_1},\delta_{\lambda_2}]\in[-0.1,0.1]$,代表动作的选取约束在距离参考轨迹正负0.1范围内。根据变体飞行器的飞行任务与参考轨迹数据,推力T取定值7.3 N,训练时间步长取0.02 s,每步的机翼掠角变化量最大值为1.5°。

奖励函数的参数如表2所示。

在LSTM-DDPG算法训练中使用的超参数如表3所示。

3.2 算法对比

为验证LSTM网络及预训练结构对训练结果的优越性,基于MATLAB的强化学习工具箱进行仿真验证。仿真结果如图9所示,图9为DDPG算法、预训练DDPG算法,及LSTM-DDPG算法在四

表 2　奖励函数参数

Table 2　Reward function parameters

奖励函数参数	数值
h_{target}/m	232
q_{target}/(rad·s^{-1})	0
V_{target}/(m·s^{-1})	32
δ_h/m	1
δ_v/(m·s^{-1})	0.5
δ_μ/rad	0.1
δ_q/(rad·s^{-1})	0.1
(h_{max},h_{min})/m	(235,198)
(V_{max},V_{min})/(m·s^{-1})	(33,20)
q_{max}/(rad·s^{-1})	1
α_{max}/rad	π/4

表 3　LSTM-DDPG 算法超参数设置

Table 3　LSTM-DDPG algorithm hyperparameter settings

超参数	数值
目标网络更新率	0.001
Actor网络学习率	0.000 1
Critic网络学习率	0.001
经验回放池容量	10^6
小批量样本数	256
折扣因子	0.991
最大幕数	80 000
最大步数	800
采样时间/s	0.02

翼变掠角飞行器起飞爬升控制任务中的表现。可以观察到, 传统 DDPG 算法的训练收敛速度最慢, 在前期进行了大量无用的探索, 在训练接近 1.7 万幕奖励才开始上升; 相较之下, 加入预训练后在 1.3 万幕附近探索得到最优动作, 平均回报收敛在 500 左右, 在收敛速度上提升 23% 左右, 验证了预训练结构对算法训练速度的提升作用。

在加入 LSTM 网络后, 算法训练效果得到进一步提升, 在 0.8 万幕附近探索得到最优动作, 在预训练 DDPG 算法基础上提升近 38%; 同时平均回报加速收敛至 570 处, 较预训练 DDPG 算法提升 14% 左右。此外, 算法收敛后的稳定性表现如图 10 所示, 传统 DDPG 算法收敛后平均回报仅为 315, 却有标准差为 200 的波动量, 其中, 回报值大于 500 的幕数占比仅为 13.5%; 加入预训练后平均回报提高到 500, 回报标准差下降至 82.5; 加入 LSTM 网络结构后, 回报标准差进一步降低至 24.5, 同时回报值大于 500 的幕数占比高达 98%。可以看出, LSTM-DDPG 表现最为稳定, 波动最小, 高回报幕数占比极大。这主要归因于 LSTM 网络结构的引入, 其能有效地处理和利用飞行器起飞爬升过程中的序列关系, 从而使策略网络能更准确、更快速地分辨出优秀数据, 并学习得到最优策略, 避免进行不必要的探索和低效率的网络参数更新, 提高了数据的利用效率。

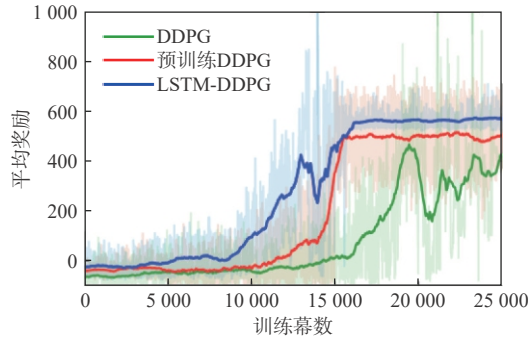


图 9　训练平均奖励对比

Fig. 9　Comparison of average training rewards

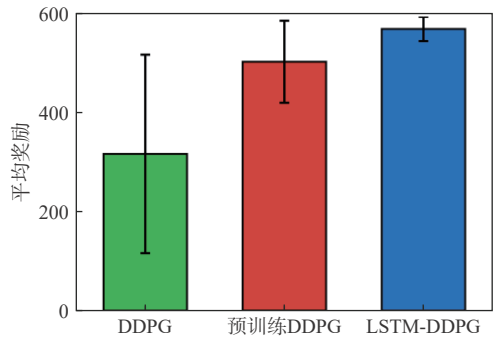
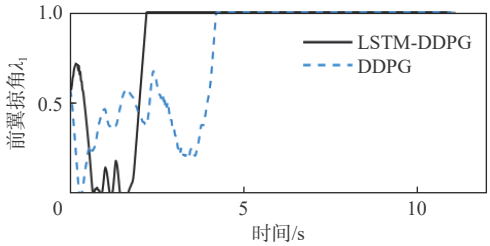


图 10　各算法收敛后表现对比

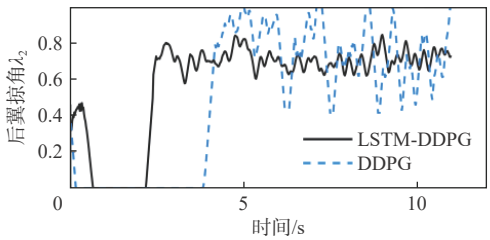
Fig. 10　Comparison of each algorithm performance after convergence

综上, 仿真结果验证了 LSTM-DDPG 算法在处理四翼变掠角飞行器起飞爬升控制任务时, 对比传统 DDPG 和预训练 DDPG 算法, 具有显著的优越性。这强有力地证明了 LSTM 结构的有效性和必要性, 为此类任务的强化学习算法设计提供了有价值的参考。

图 11 为完成该飞行任务的控制量变化曲线。对于 DDPG 算法, 虽然能够成功控制飞行器完成起飞爬升任务, 但是其输出的控制量波动较大。这种较大的波动可能会对飞行器的稳定性带来负面影响, 也造成了图 10 所示算法收敛后波动较大, 高回报幕数占比少于 LSTM-DDPG 算法的情况。在实际操作中, 频繁且大幅度的控制量变化可能会导致飞行器性能下降, 甚至可能会对机翼结构产生破坏。相比之下, 采用 LSTM-DDPG 算法的飞行器在起飞爬升过程中, 其控制量的变化明显更为稳定, 波动较小。这可以归因于 LSTM 网络的加入, 其能有效地理解和利用飞行器起飞爬升过程中的时间序列信息, 从而在预测未来的控制量时, 可以更为稳定且准确。可见爬升阶段主要依靠后翼变化, 到达一定高度后, 加速阶段主要靠前翼调节。图 12 为飞行过程中飞行器的外形, 在爬升过程中, 后翼掠角逐渐增大, 产生抬头力矩, 增加升力; 爬升到最



(a) 前翼掠角变形量控制曲线



(b) 后翼掠角变形量控制曲线

图 11 动作量变化曲线

Fig. 11 Action variation curves

高点后，后翼掠角达到最大值，此时依靠前翼掠角的变化控制飞行器的低头力矩使飞行器保持平飞状态。

3.3 轨迹跟踪仿真结果

为验证 LSTM-DDPG 算法的鲁棒性，向飞行器动力学方程中引入升力不确定性的系数 $\varepsilon = 0.1$，以模拟复杂环境下变体飞行器运行的高不确定性。选取加速爬升阶段轨迹跟踪控制为目标，进一步训

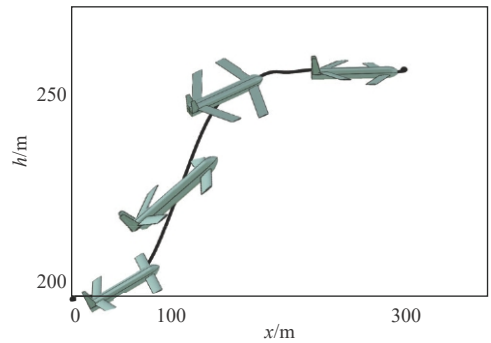
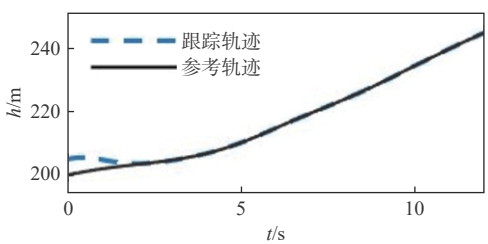


图 12 飞行过程中外形变化

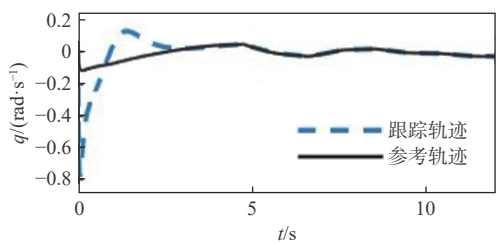
Fig. 12 Shape variation during flight

练 LSTM-DDPG 网络，选取训练集为 1 000 幕，测试其跟踪参考轨迹的效果。加入高度的初始误差以测试其跟踪性能，设置初始状态 $\boldsymbol{s}_0 = [205, 0, 16, 0.128\,9, 0.128\,9, 0]$，可得飞行器各项飞行状态的跟踪表现如图 13 所示。

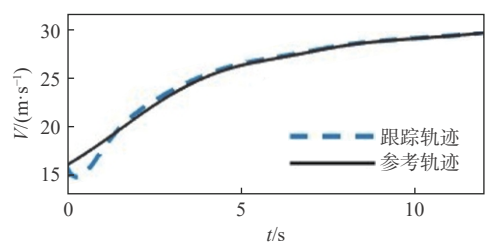
由图 13 可知，在 LSTM-DDPG 算法的训练后，即使在升力系数存在不确定性的情况下，变体飞行器依然可以快速而准确地跟踪目标轨迹，完成加速爬升的任务。各项状态跟踪的调节时间均控制在 2.4 s 以内，末端误差均小于 0.7%；其中，飞行高度跟踪曲线超调量为 0.2%，末端误差为 0.11%。验证了 LSTM-DDPG 算法对变形过程模型不确定性的强鲁棒性。这种表现得益于 LSTM 网络的记忆单



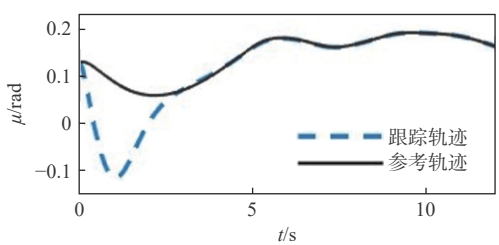
(a) 飞行高度跟踪曲线



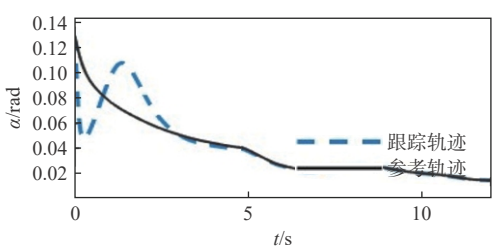
(b) 俯仰角速率跟踪曲线



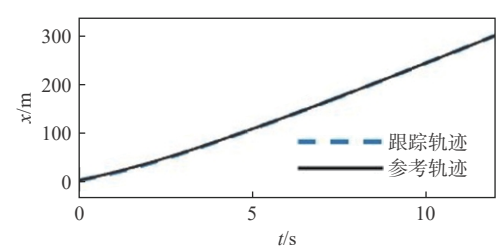
(c) 飞行速度跟踪曲线



(d) 航迹角跟踪曲线



(e) 迎角跟踪曲线



(f) 水平位移跟踪曲线

图 13 初始偏差下的跟踪曲线

Fig. 13 Tracking curves under initial deviation

元结构, 能够有效地存储并利用过去的信息, 从而在进行动作选择时, 能够更全面地考虑到当前动作对后续状态转移的影响。这进一步证实了 LSTM 网络结构在处理复杂、不确定性高的任务时的优越性, 展示了 LSTM-DDPG 算法的卓越性能和应用潜力。

综上所述, LSTM-DDPG 算法在处理无人机加速爬升控制任务表现出强大的性能, 能够有效处理复杂的飞行环境, 在训练效率和控制稳定性方面较传统 DDPG 算法有更好的表现。这些结果进一步证明了将 LSTM 网络应用于深度强化学习框架的有效性, 为未来更为复杂的无人机控制任务提供了有力的理论支持和实践参考。

4　结　论

本文针对变体飞行器在复杂环境中飞行控制的问题, 提出一种新型的深度强化学习算法——LSTM-DDPG。本文不仅进行了理论框架的构建, 而且在实践中通过详细的仿真实验验证了其有效性。

1) 针对一种由前后 2 对变掠角机翼代替传统副翼和升降舵的变体飞行器, 构建了精确的飞行器动态模型, 为无人机的加速爬升任务提供了准确的飞行状态预测。

2) 将 LSTM 网络结合到 DDPG 算法中, 构建了 LSTM-DDPG 算法, 以此来处理强化学习中的序列依赖问题。通过 LSTM 网络, 飞行器可以有效地存储和利用过去的环境信息, 从而作出更好的决策。这种结构为处理大量历史信息的任务提供了一种新的有效方法。

3) 设计了基于飞行状态和控制量的综合奖励函数, 并进行了算法训练及加速爬升阶段变形控制的性能测试。

4) 仿真结果表明, 无论是在复杂环境下飞行稳定性方面, 还是在轨迹跟踪能力方面, LSTM-DDPG 算法都表现出了显著的优势, 收敛速度较传统 DDPG 算法提升约 53%, 平均回报提升约 14%。在引入不确定性的复杂飞行环境中, LSTM-DDPG 算法能有效地处理初始偏差, 显示出较强的鲁棒性。且实际操作中, 控制量频繁大幅度变化所导致的飞行器性能下降问题, 可以通过使用 LSTM-DDPG 算法来有效解决。

综合来看, LSTM-DDPG 算法在变体飞行器控制问题中具有高效且稳定的性能。LSTM 的引入极大地增强了算法对环境信息的利用能力, 提高了飞行器在复杂环境下的控制性能。这一结果进一步证实了 LSTM 网络在深度强化学习框架中的应用价值, 并为未来无人机在更为复杂的任务和环境中的控制提供了有力的理论支持和实践参考。在实际工程运用中, 下一步还需要详细考虑该算法对机体结构强度、转动机构和驱动器负载的影响; 建立更精细的非定常气动模型并优化参数; 基于算法运算量评估各类机载平台。在未来的研究中, 将进一步优化 LSTM-DDPG 算法, 拓展其在更广泛场景中的应用。

参考文献（References）

[1] BARBARINO S, BILGEN O, AJAJ R M, et al. A review of morphing aircraft[J]. Journal of Intelligent Material Systems and Structures, 2011, 22(9): 823-877.

[2] GUO T H, HOU Z X, ZHU B J. Dynamic modeling and active morphing trajectory-attitude separation control approach for gull-wing aircraft[J]. IEEE Access, 2017, 5: 17006-17019.

[3] WU Z H, LU J C, SHI J P, et al. Robust adaptive neural control of morphing aircraft with prescribed performance[J].Mathematical Problems in Engineering, 2017, 2017(1): 1401427.

[4] GONG L G, WANG Q, HU C H, et al. Switching control of morphing aircraft based on Q-learning[J]. Chinese Journal of Aeronautics, 2020, 33(2): 672-687.

[5] XU D, HUI Z, LIU Y, et al. Morphing control of a new bionic morphing UAV with deep reinforcement learning[J]. Aerospace science and technology, 2019, 92: 232-243.

[6] YANG Z C, TAN J B, WANG X Q, et al. Reinforcement learning-based robust tracking control application to morphing aircraft[C]// Proceeding of the American Control Conference. Piscataway: IEEE Press, 2023: 2757-2762.

[7] SONG Z G, LIN Y C, LV R L, et al. Research on UAV autonomous deformation strategy based on deep learning[C]//Proceeding of the 5th International Symposium on Autonomous Systems (ISAS). Piscataway: IEEE Press, 2022: 1-7.

[8] Mnih V, Kavukcuoglu K, Silver D, et al. Playing atari with deep reinforcement learning[J/OL]. arXiv preprint, [2013-12-19]. https://doi.org/10.48550/arXiv.1312.5602

[9] VAN HASSELT H, GUEZ A, SILVER D. Deep reinforcement learning with double Q-learning[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2016, 30(1): 2094-2100.

[10] WANG Z, SCHAUL T, HESSEL M, et al. Dueling network architectures for deep reinforcement learning[C]//Proceedings of the International Conference on Machine Learning. New York: ICML, 2016: 1995-2003.

[11] SCHULMAN J, LEVINE S, ABBEEL P, et al. Trust region policy optimization[C]//Proceedings of the International Conference on Machine Learning. New York: ICML, 2015: 1889-1897.

[12] SCHULMAN J, WOLSKI F, DHARIWAL P, et al. Proximal policy optimization algorithms[EB/OL]. (2017-08-28)[2023-10-23]. https://arxiv.org/abs/1707.06347.

[13] LILLICRAP T P, HUNT J J, PRITZEL A, et al. Continuous control with deep reinforcement learning[C]//Proceedings of 4th International Conference on Learning Representations.Washington, D.c.:

ICIR, 2016.

[14] KONDA V, TSITSIKLIS J. Actor-critic algorithms[J]. Advances in Neural Information Processing Systems, 1999, 12: 1008-1014.

[15] 马少捷, 惠俊鹏, 王宇航, 等. 变形飞行器深度强化学习姿态控制方法研究[J]. 航天控制, 2022, 40(6): 3-10.
MA S J, HUI J P, WANG Y H, et al. Research on attitude control method of morphing aircraft based on deep reinforcement learning[J]. Aerospace Control, 2022, 40(6): 3-10 (in Chinese).

[16] LI B, YANG Z P, CHEN D Q, et al. Maneuvering target tracking of UAV based on MN-DDPG and transfer learning[J]. Defence Technology, 2021, 17(2): 457-466.

[17] LI R Z, WANG Q, LIU Y-A, et al. Morphing strategy design for UAV based on prioritized sweeping reinforcement learning[C]//Proceedings of the 46th Annual Conference of the IEEE Industrial Electronics Society. Piscataway: IEEE Press, 2020: 2786-2791.

[18] 仇靖雯, 何真, 黄赞. 四翼变掠角飞行器模糊建模及受约束飞行控制[J]. 飞行力学, 2023, 41(4): 10-18.
QIU J W, HE Z, HUANG Z. Fuzzy modeling and constrained flight control of four-wing variable sweep aircraft[J]. Flight Dynamics, 2023, 41(4): 10-18, 28 (in Chinese).

[19] 高良. 弹射式变掠角串置翼飞行机器人设计及控制方法研究[D]. 哈尔滨: 哈尔滨工业大学, 2020.
GAO L. Research on design and control of catapult launched tandem-wing flying robot with variable sweep[D]. Harbin: Harbin Institute of Technology, 2020(in Chinese).

[20] 杨惟轶, 白辰甲, 蔡超, 等. 深度强化学习中稀疏奖励问题研究综述[J]. 计算机科学, 2020, 47(3): 182-191.
YANG W Y, BAI C J, CAI C, et al. Survey on sparse reward in deep reinforcement learning[J]. Computer Science, 2020, 47(3): 182-191(in Chinese).

[21] KIM J J, CHA S H, RYU M, et al. Pre-training framework for improving learning speed of reinforcement learning based autonomous vehicles[C]//Proceedings of the International Conference on Electronics, Information, and Communication. Piscataway: IEEE Press, 2019: 1-2.

[22] YU Y, SI X S, HU C H, et al. A review of recurrent neural networks: LSTM cells and network architectures[J]. Neural Computation, 2019, 31(7): 1235-1270.

[23] CHAO W, HAN D W, JIE X W. Multi-rotor UAV autonomous tracking and obstacle avoidance based on improved DDPG[C]// Proceedings of the 2nd International Conference on Artificial Intelligence and Computer Engineering. Piscataway: IEEE Press, 2021: 261-267.

Active deformation decision-making for four-wing variable sweep aircraft based on LSTM-DDPG algorithm

PENG Yuxiao，HE Zhen*，QIU Jingwen

(College of Automation Engineering，Nanjing University of Aeronautics and Astronautics，Nanjing 211106，China)

Abstract：This paper presented an intelligent deformation control method based on the long short-term memory (LSTM) deep deterministic policy gradient (DDPG) algorithm, addressing the active deformation control challenges in variable configuration aircraft. A four-wing variable sweep aircraft with a tandem-wing configuration was studied, and its geometric model and aerodynamic parameters were calculated through OPENVSP, which was then used to establish the aircraft's dynamics model. The LSTM-DDPG algorithm learning framework was designed for the accelerated climb process of the four-wing variable sweep aircraft. Under symmetrical deformation conditions, active deformation decision training was performed for longitudinal trajectory tracking. Simulation results show that the LSTM-DDPG algorithm applied to the active deformation control process converges quickly and achieves higher average rewards. Moreover, the trained active deformation controller exhibits good control performance in the trajectory tracking tasks of the four-wing variable sweep aircraft.

Keywords：morphing aircraft; flight control; deep reinforcement learning; deep deterministic policy gradient; long short-term memory recurrent neural network

Received：2023-08-08；**Accepted**：2023-10-23；**Published Online**：2023-11-21 15：04

URL：link.cnki.net/urlid/11.2625.V.20231120.1506.001

Foundation item：National Natural Science Foundation of China (61873126)

* **Corresponding author**. E-mail：hezhen@nuaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0521

不确定危险区规避神经网络多段式弹道规划方法

谢蕾蕨，王旭刚*，顾镇镇
(南京理工大学 能源与动力工程学院，南京 210094)

摘 要：基于最优控制的弹道规划方法可以最大程度地发挥出超远程滑翔制导炮弹的飞行能力，然而在面对不确定的战场环境时，这种数值方法耗时、不易收敛的缺点导致其难以在线应用。针对这一问题，提出一种深度神经网络弹道规划方法，利用深度神经网络的非线性映射能力近似伪谱法计算模型以减少弹载计算机的运算负荷。根据弹体和环境的多种随机状态，在3维空间内以多阶段高斯伪谱法(MGPM)基于连续性条件将准接触点连接起来，形成满足路径约束的射程最优弹道样本数据库；以最优弹道数据样本库为基础，深度神经网络离线学习弹体在不同状态下的最优动作，以此映射出最优弹道规划计算模型。仿真结果表明：所提方法可以在随机状态下快速生成近似最优轨迹，具有良好的实时性和鲁棒性，可适用于解决在线弹道规划问题。

关 键 词：深度神经网络；超远程滑翔制导炮弹；在线弹道规划；多阶段高斯伪谱法；危险区规避

中图分类号：TJ413.+6

文献标志码：A **文章编号**：1001-5965(2025)10-3515-09

超远程滑翔式制导炮弹可以仅依靠气动力在滑翔段作宽时域、宽空域飞行，滑翔段的弹道规划一直以来都是学者们研究的热点。随着现代防空技术的发展，战场环境日益复杂化，传统离线式弹道规划已经难以适应现代战争的需求。当面对不确定作战场景时[1]，具备实时规划系统可以极大地提高滑翔制导炮弹的战场生存率，但这同时也对弹载计算机的性能提出了更高的要求，在线弹道规划的快慢和质量直接决定了能否完成打击任务。

最优轨迹规划可以最大程度上发挥出无动力飞行器的滑翔能力，在这一背景下，结合最优控制理论的弹道规划有了广泛的应用前景。最优弹道规划的本质是一个复杂、高度非线性、受到严格过程约束、终端约束限制的最优控制问题[2-5]，可以用间接法和直接法2种方法来求解，国内外在这一领域已有十分成熟的数值求解方法。间接法基于变分法或庞特里亚金极值原理，其优点是解的高保真度，缺点是收敛半径小，必须提供精确的初值猜测值和提供协态变量的初始猜测值，其复杂的求解方式导致很少应用。直接法的基本原理是采用参数化方法将连续最优控制问题转化为离散形式，再通过数值方法求解非线性规划问题获得最优轨迹。随着现代计算机性能的发展，直接法的应用变得更加广泛。文献[6]针对存在危险区的复杂环境，提出一种基于准接触点的规划策略，这种策略在危险区边缘生成准接触点，将连线穿过危险区的配点分隔成2个阶段，然后，采用内点约束的方式将每个阶段连接起来，可以有效避免规划轨迹误入危险区。文献[7]针对战术火箭的多阶段高斯伪谱法(Multi-stage Gaussian Pseudospectral method, MGPM)弹道规划问题，提出一种初值生成器的迭代策略，通过提供初值猜测，进一步提升解的生成效率。文

收稿日期：2023-08-08；**录用日期**：2023-09-14；**网络出版时间**：2023-10-24 09:21
网络出版地址：link.cnki.net/urlid/11.2625.V.20231023.1500.005
基金项目：中央高校基本科研业务费专项资金(30919011401)
*通信作者. E-mail：wxgnets@163.com

引用格式：*谢蕾蕨，王旭刚，顾镇镇. 不确定危险区规避神经网络多段式弹道规划方法[J]. 北京航空航天大学学报，2025，51(10)：3515-3523. XIE F W，WANG X G，GU Z Z. Multi-stage trajectory planning method for hazard zone avoidance under uncertainty based on neural networks[J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51(10)：3515-3523 (in Chinese).*

献 [8] 针对高超音速再入轨迹优化问题，在传统参数化求解的基础上提出一种改进的麻雀搜索算法对初值生成器进行优化，进一步地提高了计算精度和收敛速度。直接法可以解决复杂环境下的规划问题，但要设计出一种能够可靠且实时应用的最优或接近最优的规划策略仍然十分困难[9]，设计主要面临的问题是飞行器模型固有的非线性、初始条件的不确定性及在线计算的高要求[10]。

近年来，人工智能技术的发展为在线弹道规划提供了新的可能性[11-13]。深度学习依靠强大的非线性数据拟合能力，可以从弹道数据样本库中挖掘出隐含的最优控制规律实现计算模型的复刻。国内外学者一直致力于利用深度学习技术探索一种高自主性、最优动作实时映射的方法，建立起深度神经网络 (deep neural network, DNN) 与最优控制理论之间的联系[14]。文献 [15-16]利用训练有素的 DNN 为间接射击法提供精确的初始状态值，从而实现“热启动”，大大提高了算法的实时性，不过其本质是数值求解法，实时性上仍有所限制。文献 [17-18] 利用 DNN 映射初始状态和代价变量，保证了间接法的高效性和成功率，大大提高了用于航天器任务的机载轨迹优化算法的实时性和稳定性；文献 [19] 设计一种分类和回归神经网络，将着陆飞行状态分为 bang-bang 或 non-bang-bang 推力模式，再分别设计不同的神经网络制导指令控制器得到近似最优轨迹。文献 [20-21] 主要体现在轨迹样本库上，分别利用多目标优化算法和凸优化生成符合各自要求的轨迹数据库，以便神经网络完成训练。

上述研究已经证明了神经网络在轨迹规划上的可行性，然而却少有人进行基于神经网络技术在不确定危险区下的研究，因此，本文的研究重点主要面向在不确定危险区下如何利用神经网络快速生成在线弹道规划指令，主要内容分为 3 个环节，在初始扰动和随机环境的作用下，重复利用 MGPM 策略生成大规模规划轨迹数据集；在获得的轨迹数据集上训练网络，调节网络参数，以便训练出合适的神经网络控制方案；经过多组仿真验证了本文提出的深度神经网络规划方法的可行性和可靠性。本文方法可以在多种复杂环境下进行实时在线弹道规划，且结果高度近似于最优数值计算弹道。

1 问题描述

1.1 应用背景介绍

由于地缘政治、军事斗争和防御等实际问题的存在，超远程滑翔制导炮弹在进行长时间远距离飞行时需要避开一些禁飞区，离线式弹道规划固然可以完成这一点，然而战场环境的本质属性具有不确定性，这种属性一直是困扰滑翔制导炮弹弹道规划的主要难题。这种不确定性在这里主要包括弹体自身的不确定性和作战环境的不确定性，当使用最优理论进行弹道规划设计时，大气密度的不确定性、气动参数的不确定性及危险禁飞区位置的不确定性都会导致离线式最优弹道规划难以进行实际应用。因此，本文基于实际战场环境进行作战假设，超远程滑翔制导炮弹在进入敌我防区边界线后，指令系统发现危险区位置发生改变，这时需要根据新的危险区位置进行在线弹道规划，在完成规避的同时，最大限度地保留滑翔能力，作战示意图如图 1 所示。

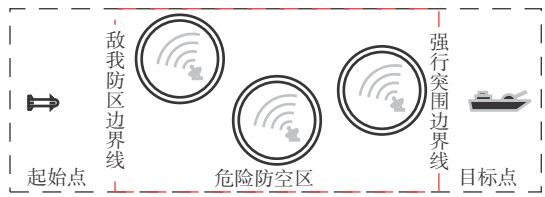


图 1 远程滑翔制导炮弹作战示意图

Fig. 1 Long-range glide-guided projectile operation

1.2 弹道学建模

本文的主要研究目标是 DNN 弹道规划方法在滑翔段的适用性，采用典型滑翔质点运动模型即可。3 维空间的 3 自由度外弹道模型为

$$\begin{cases} \dot{v} = \left(-S\rho v^2 C_x/2 - mg\sin\theta\right)/m \\ \dot{\theta} = \left(S\rho v^2 C_y/2 - mg\cos\theta\right)/(mv) \\ \dot{\psi}_v = -\dfrac{S\rho v^2 C_z}{2mv\cos\theta} \\ \dot{x} = v\cos\theta\cos\psi_v \\ \dot{y} = v\sin\theta \\ \dot{z} = -v\cos\theta\sin\psi_v \end{cases} \tag{1}$$

式中：超远程滑翔制导炮弹的特征面积 S 为 0.070 686 m^2，重量 m 为 300 kg；x、y、z 分别为炮弹在地面坐标系的位置坐标；v、θ、ψ_v、ρ 分别为炮弹在飞行过程中的速度、弹道倾角、弹道偏角及当前空气密度；C_x、C_y、C_z 分别为其阻力系数、升力系数、侧向力系数，其具体计算过程为

$$\begin{cases} C_x = C_{x0}(1 + k_1(\alpha^2 + \beta^2)) + C_{x\delta_0}(1 + k_2(\delta_z^2 + \delta_y^2)) \\ C_y = C_y^{\alpha}\alpha + C_y^{\delta_z}\delta_z \\ C_z = C_z^{\beta}\beta + C_z^{\delta_y}\delta_y \end{cases} \tag{2}$$

式中：C_{x0}、$C_{x\delta}$ 分别为弹体和尾翼组合体的零升阻力系数、舵的零升阻力系数；k_1、k_2 分别为弹体和尾翼组合体的诱导阻力系数、舵的诱导阻力系数；α、β、δ_z、δ_y 分别为攻角、侧滑角、升降舵偏角和方向舵偏角；C_y^{α}、C_z^{β} 分别为攻角的升力系数、侧滑角的侧向力系数，因为弹体为轴对称外型，所以两者数值相等；$C_y^{\delta_z}$、$C_z^{\delta_y}$ 分别为升降舵偏角的升力系数、

方向舵偏角的侧向力系数, 同理也是数值相等。

假设超远程滑翔制导炮弹在任意飞行过程中都处于瞬时平衡状态, 则攻角、侧滑角和舵偏角的关系式为

$$\begin{cases} \dfrac{\alpha}{\delta_z} = -\dfrac{m_z^{\delta_z}}{m_z^{\alpha}} \\ \dfrac{\beta}{\delta_y} = -\dfrac{m_y^{\delta_y}}{m_y^{\beta}} \end{cases} \tag{3}$$

式中: $m_z^{\delta_z}$、$m_z^{\delta_y}$分别为升降舵、方向舵操纵效率; m_z^{α}、m_y^{β}分别为俯仰、偏航静稳定导数, 两者的比值结果为炮弹配平比, 可以将舵机控制量转化为攻角、侧滑角控制量, 整个质点运动的控制向量只有$[\alpha,\beta]$。

以上便是整个超远程滑翔制导炮弹的运动学模型。

1.3　约束参数

超远程滑翔制导炮弹在飞行过程中具有宽域飞行的特点, 但其控制量本身是有限制的。因此, 控制约束参数如表 1 所示。

表 1　控制约束参数

Table 1　Control constraint parameters

参数	数值
$[\alpha_{\min},\alpha_{\max}]/(°)$	[0,12]
$[\beta_{\min},\beta_{\max}]/(°)$	[−12,12]

敌危险禁飞区一般可看作无限高的圆柱体, 设置为

$$\sqrt{(x(t)-x_{\mathrm{d}}^i)^2+(z(t)-z_{\mathrm{d}}^i)^2} \geqslant r_{\mathrm{d}}^i \qquad i=1,2,\cdots,N \tag{4}$$

式中: $(x_{\mathrm{d}}^i,z_{\mathrm{d}}^i)$为第$i$个危险禁飞区圆柱的中心位置坐标; r_{d}^i为其半径。

2　弹道样本库的建立

本节的目的为超远程滑翔制导炮弹在复杂环境和约束限制条件下生成滑翔能力最优的离线弹道样本数据库, 同时解决离散规划的情况下不满足实际路径约束将如何进行处理。

2.1　最优控制问题框架

将滑翔段弹道规划问题转化为最优控制问题, 寻找射程最大的弹道规划曲线, 最小化性能指标为

$$\begin{cases} J=\phi(s(t_{\mathrm{f}}))+\displaystyle\int_{t_0}^{t_{\mathrm{f}}} L(s(k),u(k),k)\mathrm{d}t \\ \dot{s}(t)=f(s(t),u(t),t) \end{cases} \tag{5}$$

式中: $s=(x,y,z,v,\theta,\psi_v)\in\mathbf{R}^6$为状态变量; $\phi(s(t_{\mathrm{f}}))$为终端指标函数; L为动态指标函数; 时间$t\in[t_0,t_{\mathrm{f}}]$, 其中, t_{f}为自由终端时间; J为性能函数; $\dot{s}(t)$为运动学微分方程, 详细如 (1) 式所示。

边界约束和过程约束为

$$\begin{cases} B_{\min}\leqslant B(s(t_0),s(t_{\mathrm{f}}))\leqslant B_{\max} \\ C_{\min}\leqslant C(s(t),s(t))\leqslant C_{\max} \end{cases} \tag{6}$$

式中: $B_{\min}$和$B_{\max}$分别为其边界约束的最小值和最大值; $C_{\min}$和$C_{\max}$分别为过程约束的最小值和最大值。本文的过程约束可以分为 2 类, 一类是路径约束, 即危险禁飞区的实际位置, 另一类是在工程中的控制量约束, 具体约束参数如表 1 所示。

2.2　基于高斯伪谱法的离散化

最优控制问题的伪谱法求解方法是以全局插值多项式为基函数的方式来近似整个优化时域内的状态变量和控制变量。由于伪谱法的离散化节点分布于区间$[-1,1]$, 所以需要引入τ作为新的独立变量, 转换式如下:

$$t=\frac{t_{\mathrm{f}}-t_0}{2}\tau+\frac{t_{\mathrm{f}}+t_0}{2} \tag{7}$$

式中: t_0为起始时间; t_{f}为终端时间。

转换后, τ成为最优控制问题的时间变量, 再进行后续配点和离散化, 高斯伪谱法(Gaussian Pseudospectral method, GPM)的N阶配点取为 Legendre 多项式$P_N(t)$在$[-1,1]$内关于零对称的$N+1$个实根, 简称为 LG 点, 其中, −1为函数逼近引入量。从而重新定义 Lagrange 插值多项式$L_i^{(k)}(\tau)(i=0,1,\cdots,N^{(k)})$, 将每个子区间变量离散到配点处, 就得到$N+1$个离散状态变量$(S_0^{(k)},S_1^{(k)},\cdots,S_N^{(k)})$和离散控制变量$(U_0^{(k)},U_1^{(k)},\cdots,U_N^{(k)})$。

再利用 Lagrange 插值多项式逼近$s^{(k)}(\tau)$和$u^{(k)}(\tau)$, 即:

$$\begin{cases} s^{(k)}(\tau)\approx X^{(k)}(\tau)=\displaystyle\sum_{i=0}^{N^{(k)}} L_i^{(k)}(\tau)S_i^{(k)} \\ u^{(k)}(\tau)\approx U^{(k)}(\tau)=\displaystyle\sum_{i=0}^{N^{(k)}} L_i^{(k)}(\tau)U_i^{(k)} \end{cases} \tag{8}$$

式中: $L_i^{(k)}(\tau)$为 Lagrange 插值基函数。将状态变量通过插值多项式参数化后, 对状态的微分可近似为

$$\dot{s}^{(k)}(\tau)\approx\dot{S}^{(k)}=\sum_{i=0}^{N^{(k)}}\dot{L}_i^{(k)}(\tau)S_i^{(k)}=\sum_{i=0}^{N^{(k)}}D_{ij}^{(k)} \tag{9}$$

式中: $k=0,1,\cdots,N^{(k)}$; $D_{ij}^{(k)}$为微分矩阵。

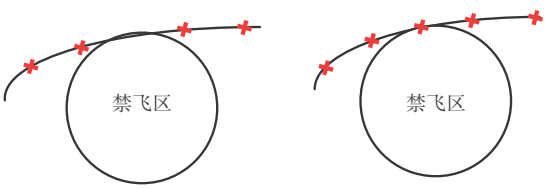
至此, 最优控制问题已经转换为非线性规划问题, 再用序列二次规划算法求解相应问题即可。

2.3　多阶段高斯伪谱法

针对复杂不确定的战场环境, 采用直接法的思想建立弹道样本库, GPM 在求解过程中是可以进行路径约束这一环节, 也就是可以在伪谱法中直接完成危险区绕飞的工作, 但 GPM 是将连续最优问题离散化后再进行求解的数值方法, LG 点是 Legendre 多项式的求解根, 求解规律是越靠近边界

处的LG点会更加密集，而区间的LG点相对稀疏。当进行全过程求解时，极易出现配点在危险区的边界，2个配点的连线却与危险区相交的情形，如图2(a)所示。

基于以上问题，本文使用以准接触点为分段节点的GPM规划策略[6]，提高规划弹道样本库的精准度。准接触点的规划算法思想如下：使用较少LG点进行粗略规划，在此时的LG点十分稀疏，观察相邻2个LG点的连线是否与圆弧相交，若与圆弧相交，则战场约束区的边界会被最优规划轨迹分成一个优弧和一个劣弧，显然，准接触点应该选取于战场约束区边界圆的劣弧上。为了有利于产生更平滑的轨迹并且减轻对机动性的需求，准接触点选取为劣弧中点。以准接触点为分段节点将制导攻击阶段进一步分段后，战场约束区附近的轨迹与LG点分布情况如图2(b)所示。



(a) 轨迹与危险区相交　(b) 轨迹穿过准接触点

图2　规划曲线在禁飞区附近的配点状态

Fig. 2　Distribution state of planning curve near no-fly zone

在滑翔多阶段规划过程中存在着2种情况：一是从当前状态规划到准接触点位置，二是从当前状态规划到落点位置，整个弹道规划过程便由两阶段规划而形成。GPM和MGPM的具体弹道规划如图3所示。

2.4　生成数据库

超远程滑翔制导炮弹的上升段在火箭助推作用下是一个复杂的过程，滑翔段起点在各种因素的影响下具有极不确定性，其具体参数实际上是难以与地面设置的顶点参数相一致。因此，考虑到样本库的更多情形，设置随机滑翔顶点初始状态参数和落速约束如表2所示。

考虑到滑翔弹的机动能力有限，若完全不考虑危险区的参数限制，弹道样本库在生成期间会产生极大的困难，为保障滑翔弹具有足够的规避能力，本文的样本库建立于如下作战场景，如表3所示。

基于以上战场模型，随机生成1 035条弹道$T_{\mathrm{raj}}=\{T_1,T_2,\cdots,T_{1035}\}$，1 035条弹道样本数据如图4所示。

表2　初始状态参数

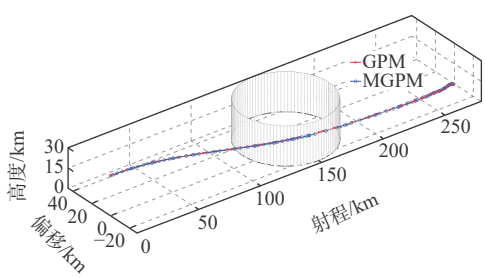
Table 2　Initial state parameters

参数	数值
$v/(\mathrm{m\cdot s^{-1}})$	[1 450,1 550]
$\theta/(°)$	[−3,3]
$\psi_{\mathrm{v}}/(°)$	[−3,3]
y/km	[28,32]
z/km	[−1,1]
$v_{\mathrm{f}}/(\mathrm{m\cdot s^{-1}})$	[300,350]

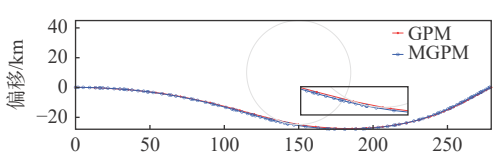
表3　危险区状态参数

Table 3　Danger zone status parameters

参数	数值
$x_{\mathrm{d}}^1/\mathrm{km}$	[100,130]
$x_{\mathrm{d}}^2/\mathrm{km}$	[130,160]
$z_{\mathrm{d}}^1/\mathrm{km}$	[−10,10]
$z_{\mathrm{d}}^2/\mathrm{km}$	[−10,10]
$r_{\mathrm{d}}^1/\mathrm{km}$	[30,40]
$r_{\mathrm{d}}^2/\mathrm{km}$	[20,30]



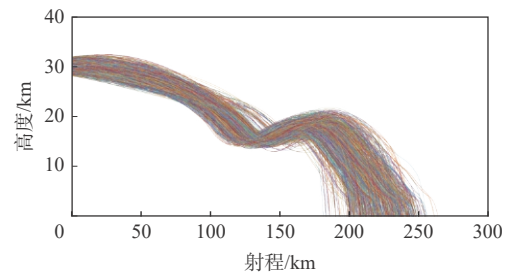
(a) GPM和MGPM规避曲线



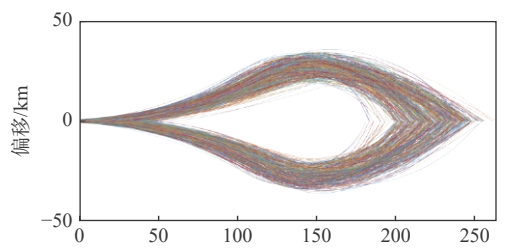
(b) GPM和MGPM规避平面局部放大图

图3　示例弹道仿真

Fig. 3　Example ballistic simulation



(a) 弹道样本库正视图



(b) 弹道样本库俯视图

图4　弹道数据样本库

Fig. 4　Ballistic data sample library

3　训练神经网络学习

本文利用 DNN 以实现超远程滑翔制导炮弹的实时弹道规划。离线训练阶段，通过采用复杂的 MGPM 轨迹优化算法来求解规避危险区的最优射程问题，生成由丰富的状态-动作对组成的训练数据集。在数据集中，随机的初始状态使得轨迹覆盖了一个很大的包络，可靠地包含了弹道规划中的多种状态空间。

本文将危险区的状态和飞行器的相关状态作为输入量，控制作为输出量，分别记为 O^{threat}、O^{agent} 和 O^{control}：

$$\begin{cases} O^{\text{threat}} = \left(x_1^{\text{threat}}, z_1^{\text{threat}}, r_1^{\text{threat}}, x_2^{\text{threat}}, z_2^{\text{threat}}, r_2^{\text{threat}}\right) \\ O^{\text{agent}} = (x, y, z, v, \theta, \psi_{\text{v}}) \\ O^{\text{control}} = (\alpha, \beta) \end{cases} \tag{10}$$

式中：$O_i^{\text{threat}} \in \mathbf{R}^6$ 为危险区的状态信息；$O^{\text{agent}} \in \mathbf{R}^6$ 为弹体的状态信息；$O^{\text{control}} \in \mathbf{R}^2$ 为输出控制量。

为使得不同量级的数据具有可比性，同时保持着数据之间彼此的相对关系，数据样本在输入前先进行归一化处理：

$$D_{\text{norm}} = \frac{D - D_{\min}}{D_{\max} - D_{\min}} \tag{11}$$

式中：D 为真实样本值；$D_{\min}$ 为这类数据样本的最小值；$D_{\max}$ 为这类数据样本的最大值；D_{norm} 为归一化之后的变量。

DNN 是一个全连接的前馈神经网络，主体通常包括输入层、若干隐藏层、输出层 3 种结构。相邻 2 层通过权值形式连接起来，每个神经元的输入参数通过激活函数进行处理，ReLU 函数具有计算速度快的特点，收敛速度远快于其他激活函数，十分符合本文机载计算机的计算要求，其计算式如下：

$$\text{ReLU}(D) = \max(0, D) \tag{12}$$

式中：$\max(0, D)$ 表示输入大于 0 的值全部不变，小于 0 的值压缩到 0 值。

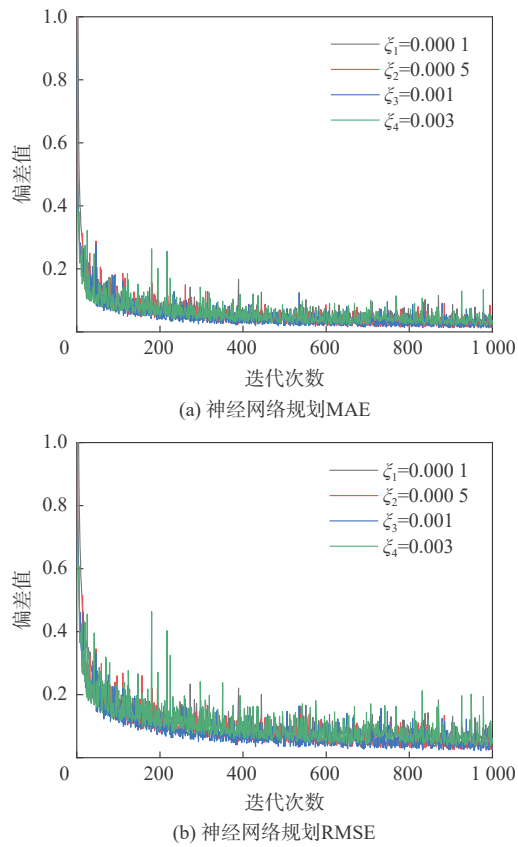
训练环境基于 Python 语言和 Pytorch 框架，训练设备基于 Intel(R)2.60 GHz 和 NVIDIA2070 显卡。综合考虑运算时间及准确性的要求，本文设置 3 层隐藏层，分别为 256-512-512 个神经单元，batch-size 设置为 2 000，意为每次训练从训练集中抽取的数量。在训练过程中采用 Adam 优化器进行梯度下降，其参数设置为：$\beta_1 = 0.9$，$\beta_2 = 0.999$，$\varepsilon = 10^{-8}$，其中，β_1 为 1 次估计指数衰减率，β_2 为 2 次估计指数衰减率，参数 ε 的作用是防止算法计算中出现 0。

DNN 通常采用平均绝对误差（mean absolute error, MAE）和均方根误差 (root mean square error, RMSE) 作为神经网络性能参考指标：

$$\begin{cases} f_{\text{MAE}}(D, h) = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} |h(D^{(i)}) - y^{(i)}| \\ f_{\text{RMSE}}(D, h) = \sqrt{\dfrac{1}{n}\displaystyle\sum_{i=1}^{n} (h(D^{(i)}) - y^{(i)})^2} \end{cases} \tag{13}$$

式中：$h(D^{(i)})$ 表示通过神经网络计算得到的输出值；$y^{(i)}$ 为真实数据样本。

学习率 ξ 作为监督学习中重要的超参，其决定着目标函数能否收敛到全局最小值以及何时收敛到最小值。暂时没有一种通用的理论技术指导如何合理设置有效学习率。本文设置 $\xi_1 = 0.000\,1$，$\xi_2 = 0.000\,5$，$\xi_3 = 0.001$，$\xi_4 = 0.003$ 这 4 种学习率，通过训练对比实验得到神经网络的相对较优学习率，如图 5 所示。



(a) 神经网络规划MAE

(b) 神经网络规划RMSE

图 5　在不同学习率下的 DNN 训练损失变化

Fig. 5　Variation of DNN training loss under different learning rates

在神经网络训练过程中，并不能保证在最大训练回合数得到的网络就是整个训练过程的性能最优网络，在过程中的某个时刻的网络性能可能比最后时刻的网络性能更好。基于以上原因，本文提出最优网络保存的策略，以 RMSE 作为参考性能指标，实时保存训练过程中的最优性能网络，如表 4 所示。

如表 4 所示，神经网络弹道规划 DNN 在 $\xi = 0.001$ 取到最小值，以此作为最优网络进行后面的仿真训练。

表 4 RMSE 最优网络性能

Table 4 Optimal network performance based on RMSE

学习率ξ	RMSE
0.000 1	0.053 6
0.000 5	0.032 0
0.001	0.026 5
0.003	0.058 5

4 仿真实验

对于实际的弹道规划而言，MAE 和 RMSE 只能作为网络性能参考，并不是弹道规划的实际评估指标，网络中的细微误差对弹道性能的影响有多大还需进一步的仿真验证。因此，本文设置多组仿真实验验证网络规划的鲁棒性及表现性能。

4.1 弹道规划网络性能

本节的目的是验证神经网络规划 DNN 的鲁棒性。在线弹道规划的本意就是可以在任意时刻根据当前状态信息实时调整弹道，形成多段式弹道规划。

当机载计算机判断当前状态需要重新弹道规划时，规划系统迅速规划一段全新弹道，由于这个规划环境是不确定的，也就导致 DNN 的每段规划路径长度是不定的，其具体长度可以根据当前机载计算机的负载状态而定。为更好地评价神经网络规划弹道的规划性能，本文的规划仿真统一采用 0.5 s 作为神经网络的一个时间步长，神经网络前向传播 50 次，总共规划 25 s 临时飞行路径作为一段弹道。本节进行了 100 组在不同状态和环境下的蒙特卡罗仿真实验，得到神经网络规划弹道与 GPM 的数值计算弹道的相比结果，具体效果如图 6 所示。

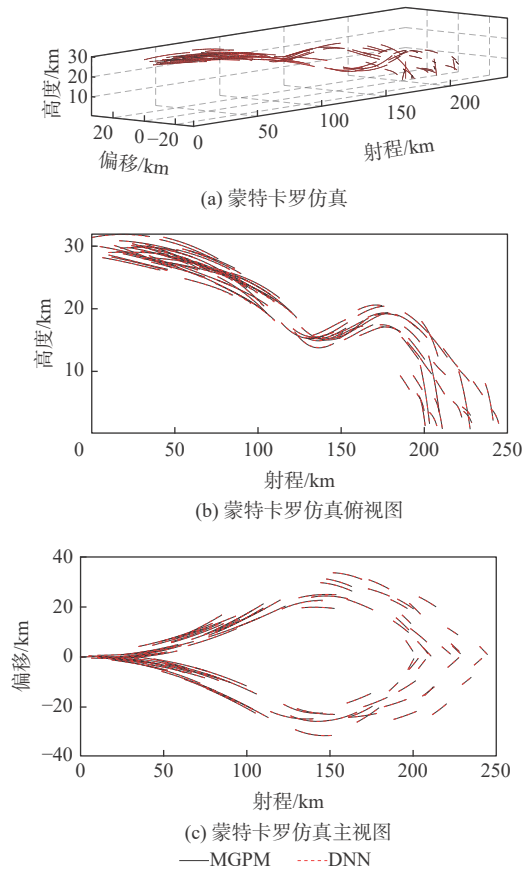


图 6 100 组蒙特卡罗仿真实验

Fig. 6 100 sets of Monte Carlo simulation experiments

图 6 的 2 种规划曲线分布于随机初始条件下的不同弹道阶段，从最短 6 km 规划路径到最长 42 km 规划路径，其平均规划长度值约为 23 km，图 7 为 MGPM 和神经网络规划方法在同一状态条件下完成规划后的末点的 3 轴偏差散点图。图 8 为 100 组实验的 3 轴平均偏差随规划时间的变化曲线，25 s 后，x 轴的平偏差值为 38.2 m，y 轴的平均偏差值为 55.3 m，z 轴的平均偏差值为 41.7 m，2 种方法的最后平均规划末点差值为 79.1 m，这个数值相较于平均总体规划长度而言，仅为它的 0.35%。从图 6～图 8 的结论可以验证神经网络规划的良好鲁棒性。

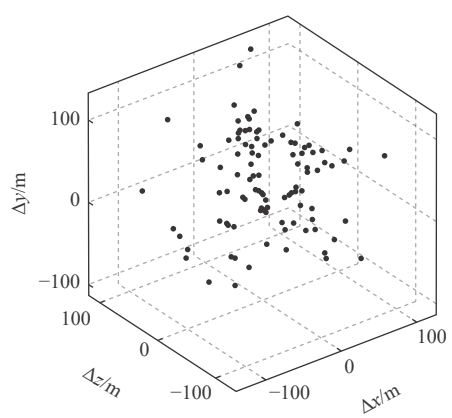


图 7 25 s 后两种弹道规划方法的偏差散点图

Fig. 7 Trajectory planning error scatter plot after 25 s

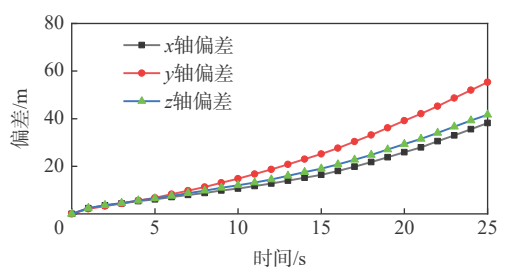


图 8 随规划时间变化的平均偏差

Fig. 8 Average deviation over planning time

4.2 全过程多段式弹道规划仿真

本节的仿真工作集中于弹道规划全过程，观察 DNN 在整个弹道规划中的适用性。设置战场环境参数如表 5 所示。

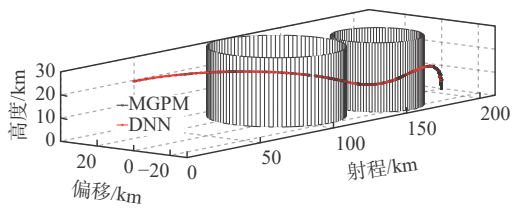
神经网络规划过程由于其不确定性，本身是一个需要不断调整的过程，实际弹道规划中，由于控制机构的操纵性能及其他随机过程扰动，导致无动力滑翔弹难以长时间方案弹道飞行。因此，本文采用多段式弹道规划策略，在每段规划结束后，给状态参数添加随机扰动以模拟弹体的不确定性，每次规划弹道根据当前状态进行实时规划。具体多段式弹道规划曲线如图 9 所示。

图 9 和表 6 展示了全过程弹道总共 13 个阶段的规划过程及每个阶段结束后 2 种规划方法的偏

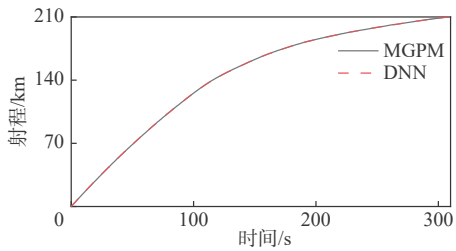
表 5　起始状态参数和战场环境参数设置

Table 5　Settings of starting state parameters and battlefield environment parameters

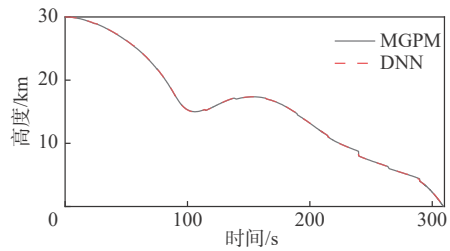
参数设置	数值
高度位置y/m	30 000
偏移位置z/m	0
起始速度v/(m·s^{-1})	1 500
起始弹道倾角θ/(°)	0
起始弹道偏角ψ_v/(°)	0
落速约束v_{f}/(m·s^{-1})	330
危险区1位置$(x_{\mathrm{d}}^1, z_{\mathrm{d}}^1, r_{\mathrm{d}}^1)$/km	(110,10,30)
危险区2位置$(x_{\mathrm{d}}^2, z_{\mathrm{d}}^2, r_{\mathrm{d}}^2)$/km	(160,−5,20)



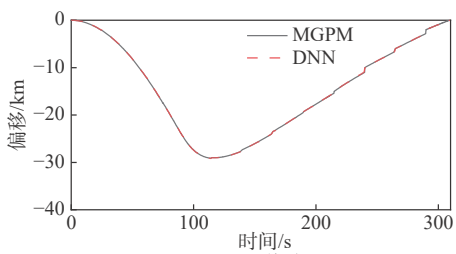
(a) 多段式全过程弹道规划3维弹道图



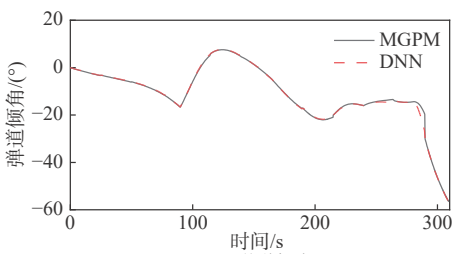
(b1) 射程



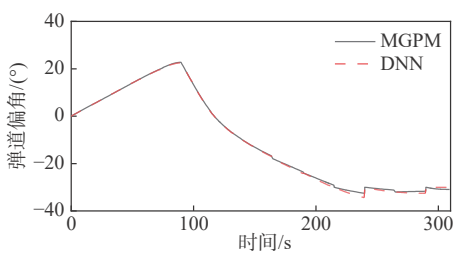
(b2) 高度



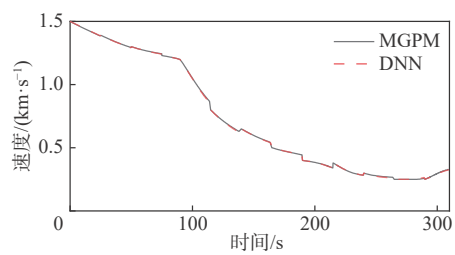
(b3) 偏移



(b4) 弹道倾角

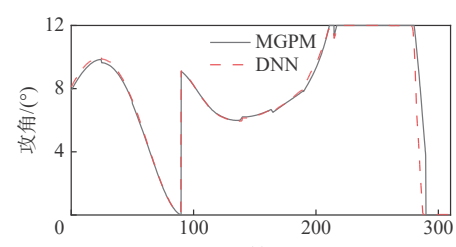


(b5) 弹道偏角

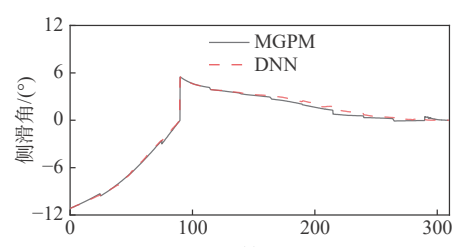


(b6) 速度

(b) 多段式全过程弹道规划状态参数变化曲线



(c1) 攻角



(c2) 侧滑角

(c) 多段式全过程弹道规划控制参数变化曲线

图 9　多段式全过程弹道规划

Fig. 9　Multi-stage ballistic planning

差。神经网络输出和精确的最优轨迹计算结果之间不可避免的会出现偏差，虽然规划末点存在差异，但相较其规划空间的大小而言是一个小量值。在规划效率方面，每段神经网络规划耗时仅为0.05 s，13段合计耗时为0.65 s，由于这个时间分布于整个飞行过程中，进一步降低了弹载计算机的负荷。而MGPM平均弹道规划时间需要10 s左右，并且收敛性无法得到有效保障。本文弹道规划为显得阶段更加直观，对于超过200 km只进行了13段规划，在实际作战场景可根据当前弹载计算机的负荷采用更多次实时规划进一步减少误差。

表6 DNN和MGPM在每阶段结束后的数值差

Table 6 Difference between DNN and MGPM at the end of each stage

阶段	Δx/m	Δy/m	Δz/m	$\Delta\theta$/(°)	$\Delta\psi_v$/(°)	Δv/(m·s^{-1})
Phase 1	83.28	43.89	56.79	0.10	0.00	1.77
Phase 2	−5.53	69.58	24.09	0.15	0.08	−3.52
Phase 3	18.17	91.55	63.50	0.19	−0.14	−2.03
Phase 4	0.00	53.03	63.43	0.08	−0.23	−0.42
Phase 5	−60.09	4.54	−43.20	0.51	−0.26	−10.69
Phase 6	−57.86	24.07	−10.35	0.03	−0.21	−7.14
Phase 7	−17.65	15.84	12.28	−0.14	−0.27	−1.98
Phase 8	9.15	−1.89	65.98	−0.16	−0.73	1.83
Phase 9	−33.29	28.03	71.96	−0.03	1.20	0.59
Phase 10	−39.32	−28.61	125.19	−0.02	−1.86	−1.36
Phase 11	−113.22	−36.39	−3.32	1.36	−0.87	−5.64
Phase 12	−37.82	−133.62	42.23	−7.35	−0.88	9.06
Phase 13	35.80	0.00	−26.37	−0.37	0.90	−3.85

5 结 论

1）本文通过使用MGPM解决了常规伪谱法在复杂环境下，规划的路径曲线存在着越过危险区边界的问题，提高了弹道样本库的规避质量。

2）本文大量仿真实验已经验证了经过良好训练的DNN在复杂环境下完成弹道的自主规划，规划效果与数值规划误差极小，成功将时间成本转移到了地面训练过程。并且这种多段式神经网络规划的策略可以实时应对飞行过程中的各种误差扰动，规划速度极快，可满足在线弹道规划的时间要求。

参考文献（References）

[1] 沈林成, 陈璟, 王楠. 飞行器任务规划技术综述[J]. 航空学报, 2014, 35(3): 593-606.
SHEN L C, CHEN J, WANG N. Overview of air vehicle mission planning techniques[J]. Acta Aeronautica et Astronautica Sinica, 2014, 35(3): 593-606 (in Chinese).

[2] BETTS J T. Survey of numerical methods for trajectory optimization[J]. Journal of Guidance, Control, and Dynamics, 1998, 21(2): 193-207.

[3] SHIRAZI A, CEBERIO J, LOZANO J A. Spacecraft trajectory optimization: a review of models, objectives, approaches and solutions[J]. Progress in Aerospace Sciences, 2018, 102: 76-98.

[4] 张远龙, 谢愈. 滑翔飞行器弹道规划与制导方法综述[J]. 航空学报, 2020, 41(1): 023377.
ZHANG Y L, XIE Y. Review of trajectory planning and guidance methods for gliding vehicles[J]. Acta Aeronautica et Astronautica Sinica, 2020, 41(1): 023377 (in Chinese).

[5] 黄长强, 国海峰, 丁达理. 高超声速滑翔飞行器轨迹优化与制导综述[J]. 宇航学报, 2014, 35(4): 369-379.
HUANG C Q, GUO H F, DING D L. A survey of trajectory optimization and guidance for hypersonic gliding vehicle[J]. Journal of Astronautics, 2014, 35(4): 369-379 (in Chinese).

[6] ZHAO J, ZHOU R. Reentry trajectory optimization for hypersonic vehicle satisfying complex constraints[J]. Chinese Journal of Aeronautics, 2013, 26(6): 1544-1553.

[7] 刘超越, 张成. 基于高斯伪谱法的二级助推战术火箭多阶段轨迹优化[J]. 兵工学报, 2019, 40(2): 292-302.
LIU C Y, ZHANG C. Multi-stage trajectory optimization of tactical two-stage booster rocket based on Gauss pseudospectral method[J]. Acta Armamentarii, 2019, 40(2): 292-302 (in Chinese).

[8] SU Y, DAI Y, LIU Y. A hybrid hyper-heuristic whale optimization algorithm for reusable launch vehicle reentry trajectory optimization[J]. Aerospace Science and Technology, 2021, 119: 107200.

[9] CHAI R Q, SAVVARIS A, TSOURDOS A, et al. Optimal tracking guidance for aeroassisted spacecraft reconnaissance mission based on receding horizon control[J]. IEEE Transactions on Aerospace and Electronic Systems, 2018, 54(4): 1575-1588.

[10] CONWAY B A. A survey of methods available for the numerical optimization of continuous dynamic systems[J]. Journal of Optimization Theory and Applications, 2012, 152(2): 271-306.

[11] 杨希祥, 李晓斌, 肖飞, 等. 智能优化算法及其在飞行器优化设计领域的应用综述[J]. 宇航学报, 2009, 30(6): 2051-2061.

YANG X X, LI X B, XIAO F, et al. Overview of intelligent optimization algorithm and its application in flight vehicles optimization design[J]. Journal of Astronautics, 2009, 30(6): 2051-2061 (in Chinese).

[12] 黄旭星, 李爽, 杨彬, 等. 人工智能在航天器制导与控制中的应用综述[J]. 航空学报, 2021, 42(4): 524201.
HUANG X X, LI S, YANG B, et al. Spacecraft guidance and control based on artificial intelligence: review[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(4): 524201 (in Chinese).

[13] 魏毅寅, 郝明瑞, 范宇. 人工智能技术在宽域飞行器控制中的应用[J]. 宇航学报, 2023, 44(4): 530-537.
WEI Y Y, HAO M R, FAN Y. The application of artificial intelligence technology in wide-field vehicle control[J]. Journal of Astronautics, 2023, 44(4): 530-537 (in Chinese).

[14] EFFATI S, PAKDAMAN M. Optimal control problem via neural networks[J]. Neural Computing and Applications, 2013, 23(7): 2093-2100.

[15] SHI J L, WANG J B, SU L F, et al. A neural network warm-started indirect trajectory optimization method[J]. Aerospace, 2022, 9(8): 435.

[16] LI X, WANG L, WANG H, et al. A warm-started trajectory planner for fixed-wing unmanned aerial vehicle formation[J]. Applied Mathematical Modelling, 2023, 122: 200-219.

[17] SÁNCHEZ-SÁNCHEZ C, IZZO D. Real-time optimal control via deep neural networks: study on landing problems[J]. Journal of Guidance, Control, and Dynamics, 2018, 41(5): 1122-1135.

[18] CHENG L, WANG Z B, JIANG F H, et al. Fast generation of optimal asteroid landing trajectories using deep neural networks[J]. IEEE Transactions on Aerospace and Electronic Systems, 2019, 56(4): 2642-2655.

[19] WANG J B, MA H J, LI H X, et al. Real-time guidance for powered landing of reusable rockets via deep learning[J]. Neural Computing and Applications, 2023, 35(9): 6383-6404.

[20] CHAI R Q, TSOURDOS A, SAVVARIS A, et al. Real-time reentry trajectory planning of hypersonic vehicles: a two-step strategy incorporating fuzzy multiobjective transcription and deep neural network[J]. IEEE Transactions on Industrial Electronics, 2019, 67(8): 6904-6915.

[21] DAI P, FENG D Z, FENG W H, et al. Entry trajectory optimization for hypersonic vehicles based on convex programming and neural network[J]. Aerospace Science and Technology, 2023, 137: 108259.

Multi-stage trajectory planning method for hazard zone avoidance under uncertainty based on neural networks

XIE Fanwei，WANG Xugang*，GU Zhenzhen

(School of Energy and Power Engineering，Nanjing University of Science and Technology，Nanjing 210094，China)

Abstract：The trajectory planning method based on optimal control can maximize the flight capability of ultra-long-range glide-guided projectiles. However, when faced with uncertain battlefield conditions, this numerical method, which is time-consuming and prone to poor convergence, makes online applications difficult. To address this issue, a deep neural network-based trajectory planning method was proposed. This method utilized the non-linear mapping ability of deep neural networks to approximate the pseudospectral method computation model, reducing the computational load on the onboard computer of the projectile. The implementation of the method was primarily divided into two steps. The first step connected pre-contact points in three-dimensional space based on continuity conditions, considering various random states of the projectile and environment, using the multi-stage Gaussian pseudospectral method (MGPM), forming a range-optimal trajectory sample database that satisfies path constraints. The second step involved mapping out the optimal trajectory planning model by having the deep neural network learn the optimal actions of the projectile under different states offline, based on the optimal trajectory data sample library. Simulation results show that the proposed method can quickly generate near-optimal trajectories under random, demonstrating good real-time performance and robustness, making it suitable for solving online trajectory planning problems.

Keywords：deep neural network；ultra-long-range glide-guided projectile；online trajectory planning；multi-stage Gaussian pseudospectral method；hazard zone avoidance

Received：2023-08-08；**Accepted**：2023-09-14；**Published Online**：2023-10-24 09：21

URL：link.cnki.net/urlid/11.2625.V.20231023.1500.005

Foundation item：The Fundamental Research Funds for the Central Universities (30919011401)

* **Corresponding author**. E-mail：wxgnets@163.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0529

运载火箭助推段大攻角剖面高精度弹道解析解

张珂洋[1]，王鹏[2]，于琦[1]，余文斌[1,*]，陈万春[1]

(1. 北京航空航天大学 宇航学院，北京 102206；2. 北京宇航系统工程研究所，北京 100076)

摘 要：针对运载火箭助推段预测校正制导方法中在线快速轨迹预测困难的问题，提出一种适用于大攻角剖面的运载火箭助推段高精度弹道解析解。推导得到以质量为自变量的纵平面简化动力学模型，并创新性地将攻角的正弦值设计为关于质量的多项式。因为大攻角机动的影响，简化动力学模型依旧是高度非线性的，无法直接求得解析解。通过受力分析，构造近似多项式替代原方程中非线性强但模值较小的项，并将真实值与近似值之差作为摄动小量。根据摄动理论可以对动力学模型进行分阶以获得可以解析求解的子系统。对子系统积分可以得到速度、弹道倾角、航程和高度的解析解。仿真实验表明：在大攻角条件下，所提出解析解的精度比现有解析解至少提升了85%。

关 键 词：解析解；助推段弹道；大攻角剖面；运载火箭；摄动法

中图分类号：V448.1

文献标志码：A **文章编号**：1001-5965(2025)10-3524-11

运载火箭主动段关机点参数，特别是速度、弹道倾角和高度，将直接决定入轨精度，因此主动段弹道规划与制导是运载火箭设计的核心[1]。火箭在主动段受力复杂，约束条件多，且速度、高度和弹道倾角相互耦合，给弹道规划和制导带来了巨大的挑战[2]。近年来，随着数学手段的发展，越来越多的新方法被应用在运载火箭上升段制导问题上。在文献 [3] 中，Cheng 等针对复杂上升轨迹规划问题，基于 Newton-Kantorovich/伪谱方法，将其转化为凸规划问题，利用现有的凸优化方法迭代求解。Sun 等[4]研究了离散约束和动力学方程的凸方法，提出了一种基于缩小信赖域约束的迭代算法，提高了优化效率，但由于轨迹优化方法对计算能力要求较高，还无法在线应用。预测校正方法[5-6]因其鲁棒性强、与轨迹优化法相比计算量相对较小等特点逐渐成为运载火箭主动段制导律的研究热点。该方法常用数值积分算法预测终端状态，根据状态偏差校正当前控制剖面[7]，然而弹道预测的计算效率将直接决定制导律能否满足在线应用的要求。

为了减轻弹载计算机的负担以实现在线制导，弹道解析解凭借较高的计算效率被越来越多的学者所重视[8-9]。但由于推力的存在，助推段动力学方程相互耦合且非线性强，很难获得高精度的解析解。早在 1903 年，齐奥尔科夫斯基在忽略重力和气动力的假设下，推导得到了火箭速度与火箭质量、推进剂比冲之间的关系[10]，这是最早的运载火箭主动段解析解。胡锦川在其博士论文中，忽略地球自转，在零攻角假设下推导考虑气动力影响的运载火箭主动段速度、高度和弹道倾角解析解[11]。Zhao 等在文献 [11] 的基础上，在零攻角处对各状态量使用一阶泰勒级数展开，通过正则摄动法求得各状态增量的解析解，进而推导得到非零攻角下的运载火箭解析解[12]，但随着攻角的增加，使用一阶泰勒级数近似的精度将显著降低，因此 Zhao 等推导得到的解析解只适用于小攻角场景。

固体火箭发动机一般无法进行推力调节或提

收稿日期：2023-08-17；录用日期：2023-10-24；网络出版时间：2023-11-01 16:05
网络出版地址：link.cnki.net/urlid/11.2625.V.20231101.1418.004
*通信作者. E-mail：yuwenbin_buaa@163.com

引用格式：张珂洋，王鹏，于琦，等. 运载火箭助推段大攻角剖面高精度弹道解析解 [J]. 北京航空航天大学学报，2025，51（10）：3524-3534. ZHANG K Y，WANG P，YU Q，et al. Analytical solutions for ascent trajectory under high angle-of-attack conditions [J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51（10）：3524-3534（in Chinese）.

前关机, 因此需要在高空执行大攻角机动以实现能量管理[13], 其攻角剖面的最大值有时甚至会超过60°, 此时无法再像文献 [13] 和文献 [14] 那样使用零攻角假设来简化动力学模型, 而大攻角条件下的助推段动力学方程具有高度非线性的特点, 各状态量耦合严重, 解析求解具有较大难度。因此本文在考虑地球自转的前提下, 研究了一种运载火箭在大攻角剖面下的弹道解析解。首先, 基于量级分析对助推段动力学方程进行合理简化, 建立以质量为自变量的运载火箭纵平面运动简化方程; 其次, 根据攻角的实际规律创新性地设计正弦攻角剖面, 使用切比雪夫插值多项式近似余弦攻角剖面; 巧妙设计简单多项式对动力学模型中的复杂非线性项进行近似, 利用辅助飞行变量求解近似多项式的系数; 然后, 使用正则摄动法对原系统分阶并解析求解零阶和一阶子系统; 最后, 通过数值仿真实验验证了在大攻角的飞行条件下, 与现有解析解相比, 助推段弹道解析解的精度至少提升了 85%。

1　助推段动力学模型的简化

假设地球为一旋转的均质圆球, 运载火箭质心的运动方程[14] 如下:

$$\frac{\mathrm{d}\lambda}{\mathrm{d}t}=\frac{V\cos\gamma\sin\psi}{R\cos\phi} \tag{1}$$

$$\frac{\mathrm{d}\phi}{\mathrm{d}t}=\frac{V\cos\gamma\cos\psi}{R} \tag{2}$$

$$\frac{\mathrm{d}R}{\mathrm{d}t}=V\sin\gamma \tag{3}$$

$$\frac{\mathrm{d}V}{\mathrm{d}t}=-\frac{D}{m}+\frac{P}{m}\cos\alpha\cos\beta+\omega_{\mathrm{e}}^2R\cos^2\phi\sin\gamma-g\sin\gamma-\omega_{\mathrm{e}}^2R\sin\phi\cos\phi\cos\gamma\cos\psi \tag{4}$$

$$\frac{\mathrm{d}\gamma}{\mathrm{d}t}=\frac{L}{mV}+\frac{P\sin\alpha}{mV}-\frac{g\cos\gamma}{V}+\frac{V\cos\gamma}{R}+\frac{\omega_{\mathrm{e}}^2R}{V}\cdot\left(\cos^2\phi\cos\gamma+\sin\phi\cos\phi\sin\gamma\cos\psi\right)+2\omega_{\mathrm{e}}\cos\phi\sin\psi \tag{5}$$

$$\frac{\mathrm{d}\psi}{\mathrm{d}t}=\frac{Z}{mV\cos\gamma}-\frac{P\cos\alpha\sin\beta}{mV\cos\gamma}+\frac{V\cos\gamma\sin\psi\tan\phi}{R}+\frac{\omega_{\mathrm{e}}^2R\sin\phi\cos\phi\sin\psi}{V\cos\gamma}+2\omega_{\mathrm{e}}(\sin\phi-\cos\phi\tan\gamma\cos\psi) \tag{6}$$

$$\frac{\mathrm{d}m}{\mathrm{d}t}=-q_m \tag{7}$$

式中: R为火箭质心到地心的距离; λ和ϕ分别为经度和纬度; V为相对地球的飞行速度; γ为当地弹道倾角; ψ为当地航向角(北偏东为正); m为火箭质量; D、L和Z分别为气动阻力、气动升力和气动横向力; α和β分别为攻角和侧滑角; g为当地重力加速度; ω_{e}为地球自转角速度; q_m为火箭质量流量; P为当地推力, 其与真空推力P_0、当地大气压强P_{a}和火箭发动机尾喷管截面积S_N的关系为

$$P=P_0-S_NP_{\mathrm{a}} \tag{8}$$

由式 (1) ～ 式 (8) 可知, 助推段动力学模型具有高度非线性的特点, 难以获得精确的解析解。因此, 有必要根据助推段弹道的特点, 对上述方程进行适当的简化。

在主动段飞行过程中, 火箭主要在纵向平面内进行机动转弯, 由地球自转引起的横向运动范围较小且容易修正, 因此可以忽略横向运动, 仅考虑纵向平面内的运动, 即仅以航程x、地心距R、飞行速度V、弹道倾角γ和质量m为状态量。

首先, 对速度方程(即式(4))进行简化, 地球自转角速度[15]$\omega_{\mathrm{e}}\approx7.292\,116\times10^{-5}$ rad/s, 主动段飞行高度H一般不超过500 km, 因此$R=R_{\mathrm{e}}+H<6\,871$ km且$\omega_{\mathrm{e}}^2R<0.04$ m/s^2, 其中, $R_{\mathrm{e}}=6\,371$ km为地球平均半径。而火箭的推重比一定大于 1, 即$\omega_{\mathrm{e}}^2R\ll g<P/m$, 所以地球自转对飞行速度的影响可以忽略。其次, 对弹道倾角方程(式 (5))进行简化。$\omega_{\mathrm{e}}R/V$项可以视为火箭所处位置的地球自转线速度与火箭飞行速度之比, 该比值一定远小于 1, 因此在方程中, 含ω_{e}^2R/V的项对弹道倾角的影响远小于科氏力项, 可以忽略。此外, 在文献 [16] 中已经证明了当飞行器在纵向平面内飞行时, $\cos\phi\sin\psi$为常数, 即$\cos\phi\sin\psi=\cos\phi_0\sin\psi_0$, 其中, ϕ_0和ψ_0分别表示ϕ和ψ的初值。

以米诺陶 IV 运载火箭为例, 其质量流量为常值[17], 结合式 (7) 可知, 常值质量流量下质量和时间可做线性变换, 因此可以使用质量代替时间作为自变量, 如此一来, 可以减少质心运动方程组的个数, 简化动力学模型, 将式 (1) ～ 式 (6) 除以式 (7) 可得

$$\frac{\mathrm{d}x}{\mathrm{d}m}=-\frac{V}{q_m}\cos\gamma \tag{9}$$

$$\frac{\mathrm{d}R}{\mathrm{d}m}=-\frac{V}{q_m}\sin\gamma \tag{10}$$

$$\frac{\mathrm{d}V}{\mathrm{d}m}=-\frac{1}{q_m}\left(-\frac{D}{m}+\frac{P_0-S_NP_{\mathrm{a}}}{m}\cos\alpha-g\sin\gamma\right) \tag{11}$$

$$\frac{\mathrm{d}\gamma}{\mathrm{d}m}=-\frac{1}{q_m}\left(-\frac{g\cos\gamma}{V}+2\omega_{\mathrm{e}}\cos\phi_0\sin\psi_0\right)-\frac{1}{q_m}\left[\frac{L}{mV}+\frac{(P_0-S_NP_{\mathrm{a}})\sin\alpha}{mV}+\frac{V\cos\gamma}{R}\right] \tag{12}$$

2　攻角剖面设计

在纵向平面内, 火箭通过调节攻角来改变推力方向, 进而调整轨迹。为了推导弹道解析解, 需要

设计一个攻角剖面。传统的运载火箭助推段一般直接对攻角剖面进行规划[18]，但由于在运载火箭的飞行动力学模型中，攻角α以三角函数的形式出现（如式(11)和式(12)所示），所以本文创新性地将$\sin\alpha$看作一个整体进行规划，如此一来，可以大幅降低解析解推导的复杂程度。此外，运载火箭在级间分离和跨声速飞行时通常都要求攻角收敛到零，且火箭每一级的攻角一般遵循“快速增长，缓慢降低”的变化规律。因此，在满足攻角变化规律的前提下，为了简化解析解推导，创新性地将攻角剖面设计为

$$\sin\alpha = Q_{\sin\alpha} = c_0 + c_1 m + c_2 m^2 + c_3 m^3 \tag{13}$$

式中：c_0、c_1、c_2和c_3为攻角剖面的系数。

此外，式(11)中还需要用到$\cos\alpha$。此处构造切比雪夫插值多项式$Q_{\cos\alpha}$来近似$\cos\alpha$，如下：

$$\cos\alpha \approx Q_{\cos\alpha} = \sum_{i=0}^{N} k_i^{\cos\alpha} m^i \tag{14}$$

图1对比了$\cos\alpha$的真实剖面与近似多项式的结果。从图中可以看出，近似剖面具有较高的拟合精度，可以满足解析解的推导需要。

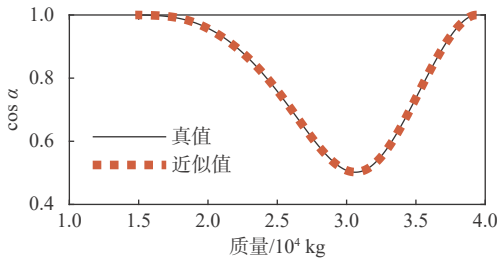


图1 攻角余弦值近似解的精度

Fig. 1 Accuracy of approximate solution of cosα

3 辅助飞行变量

为了利用摄动法推导式(9)～式(12)的近似解析解，需要先使用关于质量的多项式函数对部分高度非线性的状态量进行近似。由于无法获知终端状态的准确值，在求解近似函数的系数时需要事先引入一些辅助飞行变量。为了与真实飞行状态进行区分，使用上标“^”来表示辅助飞行变量。通过对速度、弹道倾角和地心距的微分方程进行进一步简化，辅助飞行变量只与初始状态和质量相关，因而便于求解。

首先对速度微分方程（式(11)）进行简化。该式中涉及3种作用力：阻力、推力和重力，其中阻力远小于推力和重力。因此，可以粗略构造一个关于质量的一次函数$Q_{\hat{D}}$来近似辅助阻力$\hat{D}$，形式如下：

$$\hat{D} \approx Q_{\hat{D}} = k_0^{\hat{D}} + k_1^{\hat{D}} m \tag{15}$$

式中：$k_0^{\hat{D}}$和$k_1^{\hat{D}}$为拟合系数，可由当前阻力值和预测的终端阻力值求解。由于运载火箭主动段的关机高度较高，大气密度较小，预测的终端阻力可以假设为0。

将重力加速度g和弹道倾角γ视为常数，分别以初始值代替，并忽略大气压强对发动机推力的影响，则式(11)可以简化为

$$\frac{\mathrm{d}\hat{V}}{\mathrm{d}m} \approx -\frac{1}{q_m}\sum_{i=-1}^{N-1} f_i^{\hat{V}} m^i \tag{16}$$

式中：$f_i^{\hat{V}}$为多项式系数，其表达式形式复杂但可以通过多项式运算简单地计算得到。

对式(16)积分可得辅助速度解析解为

$$\hat{V} = \sum_{i=0}^{N} F_i^{\hat{V}} m^i + F_{\ln}^{\hat{V}} \ln m \tag{17}$$

式中：$F_i^{\hat{V}}$和$F_{\ln}^{\hat{V}}$为多项式系数。

然后，对弹道倾角微分方程（式(12)）进行简化。该式中涉及5种作用力：升力、推力、重力、由地球曲率引起的离心力和由地球自转引起的惯性力。因为气动力和由大气压强引起的推力变化对运载火箭飞行的影响很小，可以粗略使用关于质量的一次多项式来近似$0.5\rho V S_{\mathrm{ref}}/m$和$P_{\mathrm{a}}/mV$项：

$$\frac{0.5\hat{\rho}\hat{V}S_{\mathrm{ref}}}{m} \approx Q_{\hat{\rho}\hat{V}} = k_0^{\hat{\rho}\hat{V}} + k_1^{\hat{\rho}\hat{V}} m \tag{18}$$

$$\frac{\hat{P}_{\mathrm{a}}}{m\hat{V}} \approx Q_{\hat{P}_{\mathrm{a}}(0)} = k_0^{\hat{P}_{\mathrm{a}}} + k_1^{\hat{P}_{\mathrm{a}}} m \tag{19}$$

式中：S_{ref}为气动参考面积。

此外，由于已经获得$\hat{V}$的解析表达式，可以构造一个关于质量的二次多项式来近似$1/\hat{V}$，如下所示：

$$\frac{1}{\hat{V}} \approx Q_{\hat{V}} = k_0^{\hat{V}} + k_1^{\hat{V}} m + k_2^{\hat{V}} m^2 \tag{20}$$

与式(15)同理，在求解式(18)和式(19)的拟合系数时，可以假设终端密度和终端大气压强的值均为0。

众所周知，气动力系数随攻角和马赫数变化。由于运载火箭具有轴对称的气动外形，当给定马赫数时，其升力系数一般可以表示为攻角的奇数次幂级数形式。而在区间$[-\pi/2,\pi/2]$内，$\sin\alpha$和α为单调关系，因此可以利用线性函数对升力系数进行近似，如下：

$$C_L \approx Q_{C_L} = k_1^{C_L}\sin\alpha = k_1^{C_L} Q_{\sin\alpha} \tag{21}$$

运载火箭主动段的高度变化相对地球半径而言属于小量，可以将$\hat{V}\cos\hat{\gamma}/\hat{R}$项和$\hat{g}\cos\hat{\gamma}/\hat{V}$项中的$\hat{R}$和$\hat{g}$视为常数，分别以初始值代替。由于辅助飞行变量是真实飞行状态的粗略近似，本文可以将$\cos\hat{\gamma}$视为常数。综上，式(12)可被简化为

$$\frac{\mathrm{d}\hat{\gamma}}{\mathrm{d}m}\approx\sum_{i=-1}^{N}f_i^{\hat{\gamma}}m^i+f_{\ln}^{\hat{\gamma}}\ln m \tag{22}$$

式中: $f_i^{\hat{\gamma}}$和$f_{\ln}^{\hat{\gamma}}$为多项式系数。

对式 (22) 积分可得辅助弹道倾角解析解为

$$\hat{\gamma}=\gamma_0+f_{-1}^{\hat{\gamma}}\ln\frac{m}{m_0}+\sum_{i=0}^{N}\frac{f_i^{\hat{\gamma}}}{i+1}\left(m^{i+1}-m_0^{i+1}\right)+f_{\ln}^{\hat{\gamma}}(m\ln m-m_0\ln m_0)-f_{\ln}^{\hat{\gamma}}(m-m_0) \tag{23}$$

最后推导辅助地心距。因为地心距微分方程 (式 (10)) 中含有$\sin\hat{\gamma}$项, 需要先对该项进行简化。实际上, 由于$\hat{\gamma}$的解析式精度并不高, $\sin\hat{\gamma}$的精度也不高。进而, 在求解辅助地心距的过程中, 无法利用高次多项式对$\sin\hat{\gamma}$进行高精度近似。为了简化运算, 采用关于m的线性多项式粗略近似$\sin\hat{\gamma}$项:

$$\sin\hat{\gamma}\approx k_0^{\sin\hat{\gamma}}+k_1^{\sin\hat{\gamma}}m \tag{24}$$

将式 (17) 和式 (24) 代入式 (10) 可得辅助地心距微分方程:

$$\frac{\mathrm{d}\hat{R}}{\mathrm{d}m}=\sum_{i=0}^{N+1}f_i^{\hat{R}}m^i-\frac{F_{\ln}^{\hat{V}}k_0^{\sin\hat{\gamma}}}{q_m}\ln m-\frac{F_{\ln}^{\hat{V}}k_1^{\sin\hat{\gamma}}}{q_m}m\ln m \tag{25}$$

式中: $f_i^{\hat{R}}$为多项式系数。

对式 (25) 进一步积分可得辅助地心距的解析解为

$$\hat{R}=R_0-\frac{F_{\ln}^{\hat{V}}k_1^{\sin\hat{\gamma}}}{2q_m}\left(m^2\ln m-m_0^2\ln m_0\right)+\sum_{i=0}^{N+1}\frac{f_i^{\hat{R}}}{i+1}\left(m^{i+1}-m_0^{i+1}\right)+\frac{F_{\ln}^{\hat{V}}k_1^{\sin\hat{\gamma}}}{4q_m}\left(m^2-m_0^2\right)-\frac{1}{q_m}\cdot\left[F_{\ln}^{\hat{V}}k_0^{\sin\hat{\gamma}}(m\ln m-m_0\ln m_0)-F_{\ln}^{\hat{V}}k_0^{\sin\hat{\gamma}}(m-m_0)\right] \tag{26}$$

4　动力学模型的分解

在本节中, 利用辅助飞行变量来构造近似多项式, 以替换式 (9) ～ 式 (12) 中具有高度非线性, 但模值较小的项, 然后根据摄动理论将真实值与近似值之差作为摄动小量进行修正。

首先, 将式 (9) ～ 式 (12) 改写成如下形式:

$$\frac{\mathrm{d}x}{\mathrm{d}m}=-V\left[Q_{\cos\gamma(1)}-\varepsilon\left(Q_{\cos\gamma(1)}-\cos\gamma\right)\right]/q_m \tag{27}$$

$$\frac{\mathrm{d}R}{\mathrm{d}m}=-V\left[Q_{\sin\gamma}-\varepsilon\left(Q_{\sin\gamma}-\sin\gamma\right)\right]/q_m \tag{28}$$

$$\frac{\mathrm{d}V}{\mathrm{d}m}=\frac{Q_{\hat{D}}-\varepsilon(Q_{\hat{D}}-D)}{q_m m}+\frac{Q_g-\varepsilon\left(Q_g-g\sin\gamma\right)}{q_m}-\frac{P_0Q_{\cos\alpha}}{q_m m}+\frac{S_NQ_{\cos\alpha}}{q_m}\left[Q_{P_{\mathrm{a}}(1)}-\varepsilon\left(Q_{P_{\mathrm{a}}(1)}-\frac{P_{\mathrm{a}}}{m}\right)\right] \tag{29}$$

$$\frac{\mathrm{d}\gamma}{\mathrm{d}m}=\frac{S_NQ_{\sin\alpha}}{q_m}\left(Q_{P_{\mathrm{a}}(2)}-\varepsilon\left(Q_{P_{\mathrm{a}}(2)}-\frac{P_{\mathrm{a}}}{mV}\right)\right)-\frac{Q_{C_L}-\varepsilon\left(Q_{C_L}-C_L\right)}{q_m}\left(Q_{\hat{\rho}\hat{V}}-\varepsilon\left(Q_{\hat{\rho}\hat{V}}-\frac{\rho VS_{\mathrm{ref}}}{2m}\right)\right)-\frac{P_0Q_{\sin\alpha}}{q_m m}\left(Q_{\hat{V}}-\varepsilon\left(Q_{\hat{V}}-\frac{1}{V}\right)\right)-\frac{2\omega_{\mathrm{e}}\cos\phi_0\sin\psi_0}{q_m}-\left\{\frac{V}{q_m\bar{R}}-\frac{\bar{g}}{q_m}\left[Q_{\hat{V}}-\varepsilon\left(Q_{\hat{V}}-\frac{1}{V}\right)\right]\right\}\cdot\left[Q_{\cos\gamma(2)}-\varepsilon\left(Q_{\cos\gamma(2)}-\cos\gamma\right)\right] \tag{30}$$

式中: $\varepsilon=1$, 用于实施摄动法; $\bar{R}$为$\hat{R}$初始和终端值的均值; $\bar{g}$由$\bar{R}$计算得到。

在航程和地心距微分方程 (式 (27) 和式 (28)) 中, 使用如下所示的方法近似$\cos\gamma$和$\sin\gamma$:

$$\begin{cases}\cos\hat{\gamma}\approx Q_{\cos\gamma(1)}=\cos\hat{\gamma}^*-\sin\hat{\gamma}^*(\gamma-\hat{\gamma}^*)\\ \sin\hat{\gamma}\approx Q_{\sin\gamma}=\sin\hat{\gamma}^*+\cos\hat{\gamma}^*(\gamma-\hat{\gamma}^*)\end{cases} \tag{31}$$

式中: $\hat{\gamma}^*$为辅助弹道倾角初始和终端值的均值。

在式 (29) 和式 (30) 中, 分别使用关于质量的一次函数Q_g和$Q_{\cos\gamma(2)}$来近似$g\sin\gamma$和$\cos\gamma$项, 如下:

$$\begin{cases}\hat{g}\sin\hat{\gamma}\approx Q_g=k_0^g+k_1^gm\\ \cos\hat{\gamma}\approx Q_{\cos\gamma(2)}=k_0^{\cos\gamma}+k_1^{\cos\gamma}m\end{cases} \tag{32}$$

分别使用关于质量的二次函数$Q_{P_{\mathrm{a}}(1)}$和$Q_{P_{\mathrm{a}}(2)}$来近似P_{a}/m和P_{a}/mV项, 形式如下:

$$\begin{cases}\hat{P}_{\mathrm{a}}/m\approx Q_{P_{\mathrm{a}}(1)}=k_0^{P_{\mathrm{a}}}+k_1^{P_{\mathrm{a}}}m+k_2^{P_{\mathrm{a}}}m^2\\ \hat{P}_{\mathrm{a}}/m\hat{V}\approx Q_{P_{\mathrm{a}}(2)}=k_0^{P_{\mathrm{a}2}}+k_1^{P_{\mathrm{a}2}}m+k_2^{P_{\mathrm{a}2}}m^2\end{cases} \tag{33}$$

利用第 3 节中介绍的辅助飞行变量, 可以确定上述多项式的系数。根据摄动法的要求, 将状态量展开为一阶级数, 即

$$\begin{cases}x=x^{(0)}+\varepsilon x^{(1)}\\ R=R^{(0)}+\varepsilon R^{(1)}\\ V=V^{(0)}+\varepsilon V^{(1)}\\ \gamma=\gamma^{(0)}+\varepsilon\gamma^{(1)}\end{cases} \tag{34}$$

式中: 上角标 "(0)" "(1)" 分别表示零阶和一阶状态量。将式 (34) 代入式 (27) ～ 式 (30), 并根据正则摄动法将其展开成ε的幂级数, 合并同阶的ε项可得状态量的零阶微分方程:

$$\frac{\mathrm{d}x^{(0)}}{\mathrm{d}m}=-\frac{V^{(0)}}{q_m}\left[(\cos\hat{\gamma}^*+\hat{\gamma}^*\sin\hat{\gamma}^*)-\gamma^{(0)}\sin\hat{\gamma}^*\right] \tag{35}$$

$$\frac{\mathrm{d}R^{(0)}}{\mathrm{d}m}=-\frac{V^{(0)}}{q_m}\left[(\sin\hat{\gamma}^*-\hat{\gamma}^*\cos\hat{\gamma}^*)+\gamma^{(0)}\cos\hat{\gamma}^*\right] \tag{36}$$

$$\frac{\mathrm{d}V^{(0)}}{\mathrm{d}m}=\frac{Q_{\hat{D}}}{mq_m}-\frac{P_0Q_{\cos\alpha}}{mq_m}+\frac{S_NQ_{\cos\alpha}Q_{P_{\mathrm{a}}(1)}+Q_g}{q_m} \tag{37}$$

$$\frac{\mathrm{d}\gamma^{(0)}}{\mathrm{d}m}=-\frac{1}{q_m}\left(\frac{P_0Q_{\sin\alpha}Q_{\hat{V}}}{m}-S_NQ_{\sin\alpha}Q_{P_{\mathrm{a}}(2)}\right)-\frac{1}{q_m}\cdot\left(Q_{C_L}Q_{\hat{\rho}\hat{V}}+V^{(0)}Q_{\cos\gamma(2)}/\bar{R}-\bar{g}Q_{\hat{V}}Q_{\cos\gamma(2)}\right)-2\omega_{\mathrm{e}}\cos\phi_0\sin\psi_0/q_m \tag{38}$$

状态量的一阶微分方程为

$$\frac{\mathrm{d}x^{(1)}}{\mathrm{d}m}=\frac{V^{(0)}}{q_m}\left(\gamma^{(1)}\sin\hat{\gamma}^*-\cos\gamma^{(0)}\right)+\frac{V^{(0)}-V^{(1)}}{q_m}\cdot\left[(\cos\hat{\gamma}^*+\hat{\gamma}^*\sin\hat{\gamma}^*)-\gamma^{(0)}\sin\hat{\gamma}^*\right] \tag{39}$$

$$\frac{\mathrm{d}R^{(1)}}{\mathrm{d}m}=\left[(\sin\hat{\gamma}^*-\hat{\gamma}^*\cos\hat{\gamma}^*)+\gamma^{(0)}\cos\hat{\gamma}^*\right]\cdot\frac{V^{(0)}-V^{(1)}}{q_m}-\frac{V^{(0)}}{q_m}\left(\gamma^{(1)}\cos\hat{\gamma}^*+\sin\gamma^{(0)}\right) \tag{40}$$

$$\frac{\mathrm{d}V^{(1)}}{\mathrm{d}m}=-\frac{1}{q_m}\left[\frac{Q_{\hat{D}}-D^{(0)}}{m}+\left(Q_g-g^{(0)}\sin\gamma^{(0)}\right)\right]-\frac{1}{q_m}\left[S_NQ_{\cos\alpha}\left(Q_{P_{\mathrm{a}}(1)}-\frac{P_{\mathrm{a}}^{(0)}}{m}\right)\right] \tag{41}$$

$$\frac{\mathrm{d}\gamma^{(1)}}{\mathrm{d}m}=\frac{Q_{C_L}}{q_m}\left(Q_{\hat{\rho}\hat{V}}-\frac{0.5\rho^{(0)}V^{(0)}S_{\mathrm{ref}}}{m}\right)+\frac{\left(Q_{C_L}-C_L\right)Q_{\hat{\rho}\hat{V}}}{q_m}+\frac{P_0Q_{\sin\alpha}}{mq_m}\left(Q_{\hat{V}}-\frac{1}{V^{(0)}}\right)-\frac{S_NQ_{\sin\alpha}}{q_m}\left(Q_{P_{\mathrm{a}}(2)}-\frac{P_{\mathrm{a}}^{(0)}}{mV^{(0)}}\right)+\frac{V^{(0)}}{\bar{R}q_m}\left(Q_{\cos\gamma(2)}-\cos\gamma^{(0)}\right)-\frac{V^{(1)}}{\bar{R}q_m}Q_{\cos\gamma(2)}-\frac{\bar{g}Q_{\hat{V}}}{q_m}\left(Q_{\cos\gamma(2)}-\cos\gamma^{(0)}\right)-\frac{\bar{g}Q_{\cos\gamma(2)}}{q_m}\left(Q_{\hat{V}}-\frac{1}{V^{(0)}}\right) \tag{42}$$

5 解析解推导

5.1 零阶解析解

首先，将式 (32) 和式 (33) 代入式 (37)，则速度零阶微分方程可被简化为

$$\frac{\mathrm{d}V^{(0)}}{\mathrm{d}m}=\sum_{i=-1}^{N+2}f_i^{V^{(0)}}m^i \tag{43}$$

式中：$f_i^{V^{(0)}}$为多项式系数。对上式积分并化简可得速度零阶解：

$$V^{(0)}=\sum_{i=0}^{N+3}F_i^{V^{(0)}}m^i+F_{\ln}^{V^{(0)}}\ln m \tag{44}$$

式中：$F_i^{V^{(0)}}$和$F_{\ln}^{V^{(0)}}$为化简后的多项式系数。

将式 (32) 和式 (33) 代入式 (38)，弹道倾角零阶微分方程可被简化为

$$\frac{\mathrm{d}\gamma^{(0)}}{\mathrm{d}m}=\sum_{i=-1}^{N+4}f_i^{\gamma^{(0)}}m^i+f_{\ln,0}^{\gamma^{(0)}}\ln m+f_{\ln,1}^{\gamma^{(0)}}m\ln m \tag{45}$$

式中：$f_i^{\gamma^{(0)}}$、$f_{\ln,0}^{\gamma^{(0)}}$和$f_{\ln,1}^{\gamma^{(0)}}$为多项式系数。

对上式积分并化简可得弹道倾角零阶解：

$$\gamma^{(0)}=\sum_{i=0}^{N+5}F_i^{\gamma^{(0)}}m^i+F_{\ln,0}^{\gamma^{(0)}}\ln m+F_{\ln,1}^{\gamma^{(0)}}m\ln m+F_{\ln,2}^{\gamma^{(0)}}m^2\ln m \tag{46}$$

式中：$F_i^{\gamma^{(0)}}$、$F_{\ln,0}^{\gamma^{(0)}}$、$F_{\ln,1}^{\gamma^{(0)}}$和$F_{\ln,2}^{\gamma^{(0)}}$为化简后的系数。

观察可知，地心距和航程的零阶微分方程在形式上十分相似，只需要将式 (35) 中的$K_{\cos\gamma(0)}=\cos\hat{\gamma}^*+\hat{\gamma}^*\sin\hat{\gamma}^*$项替换为式 (36) 中的$K_{\sin\gamma(0)}=\sin\hat{\gamma}^*-\hat{\gamma}^*\cos\hat{\gamma}^*$项，再将$K_{\cos\gamma(1)}=\sin\hat{\gamma}^*$项替换为$K_{\sin\gamma(1)}=-\cos\hat{\gamma}^*$项，即可获得式 (36)。由于替换项在推导零阶解时均视为常数，而常数替换后不会改变零阶解的推导过程和解的形式。因此，先推导航程零阶解，然后对其进行常数替换即可得到地心距零阶解。

将式 (44) 和式 (46) 代入式 (35) 可得

$$\frac{\mathrm{d}x^{(0)}}{\mathrm{d}m}=\sum_{i=0}^{2N+8}f_{2,i}^{x^{(0)}}m^i-\frac{K_{\cos\gamma(0)}}{q_m}F_{\ln}^{V^{(0)}}\ln m+\frac{K_{\cos\gamma(1)}}{q_m}\sum_{i=0}^{N+5}f_{3,i}^{x^{(0)}}m^i\ln m+K_{\cos\gamma(1)}\cdot\frac{F_{\ln}^{V^{(0)}}\ln^2 m}{q_m}\left(F_{\ln,0}^{\gamma^{(0)}}+F_{\ln,1}^{\gamma^{(0)}}m+F_{\ln,2}^{\gamma^{(0)}}m^2\right) \tag{47}$$

式中：$f_{2,i}^{x^{(0)}}$、$f_{3,i}^{x^{(0)}}$为多项式系数。对上式积分并进行代数化简可得航程零阶解：

$$x^{(0)}=\sum_{i=0}^{2N+9}F_{1,i}^{x^{(0)}}m^i+\sum_{i=1}^{N+6}F_{2,i}^{x^{(0)}}m^i\ln m+F_{\ln,1}^{x^{(0)}}m\cdot\ln^2 m+F_{\ln,2}^{x^{(0)}}m^2\ln^2 m+F_{\ln,3}^{x^{(0)}}m^3\ln^2 m \tag{48}$$

式中：$F_{1,i}^{x^{(0)}}$、$F_{2,i}^{x^{(0)}}$、$F_{\ln,1}^{x^{(0)}}$、$F_{\ln,2}^{x^{(0)}}$和$F_{\ln,3}^{x^{(0)}}$为化简后的多项式系数。

式 (36) 与式 (35) 形式相似，所以$R^{(0)}$的形式也与$x^{(0)}$相似。这里直接给出地心距零阶解，见式 (49)。此处只要对航程零阶解推导过程中的常量进行替换，即可得到式 (49) 中的系数表达式。

$$R^{(0)}=\sum_{i=0}^{2N+9}F_{1,i}^{R^{(0)}}m^i+\sum_{i=1}^{N+6}F_{2,i}^{R^{(0)}}m^i\ln m+F_{\ln,1}^{R^{(0)}}m\cdot\ln^2 m+F_{\ln,2}^{R^{(0)}}m^2\ln^2 m+F_{\ln,3}^{R^{(0)}}m^3\ln^2 m \tag{49}$$

5.2 一阶解析解

5.2.1 归一化质量

根据摄动法理论，一阶解是对零阶解的修正，为保证解析解的精度，需尽可能降低一阶解的误差。在前文推导过程中，近似多项式均以真实质量m为自变量。但m的范围一般在$10^3\sim10^4$ kg，其高阶幂次运算会产生较大的舍入误差，而在一阶解推导过程中，解析解的阶数会更高。因此，为了提高一阶解的精度，在推导时引入归一化质量$\tilde{m}$，其与真实质量m的关系为

$$m=\frac{m_0+m_f}{2}+\frac{m_0-m_f}{2}\tilde{m} \tag{50}$$

式中：m_0和m_f分别为初始质量和终端质量。为了后文推导方便，引入K_{m0}和K_{m1}：

$$\begin{cases}K_{m0}=\dfrac{m_0+m_f}{2}\\[2ex]K_{m1}=\dfrac{m_0-m_f}{2}\end{cases} \tag{51}$$

则基于m和$\tilde{m}$的近似多项式有如下所示的转换

关系:

$$\sum_{i=0}^{N} k_i m^i = \sum_{i=0}^{N}\left[K_{m1}^i \sum_{j=i}^{N}\left(k_j C_j^{j-i} K_{m0}^{j-i}\right)\tilde{m}^i\right] \tag{52}$$

$$\sum_{i=0}^{N} \tilde{k}_i \tilde{m}^i = \sum_{i=0}^{N}\left\{\left(\frac{1}{K_{m1}}\right)^i \sum_{j=i}^{N}\left[k_j C_j^{j-i}\left(-\frac{K_{m0}}{K_{m1}}\right)^{j-i}\right] m^i\right\} \tag{53}$$

式中: C_j^{j-i}代表组合数运算。

5.2.2　速度一阶解

式 (41) 中含有需要通过零阶解求得的阻力 $D^{(0)}$、大气压强 $P_{\mathrm{a}}^{(0)}$、重力加速度 $g^{(0)}$和 $\sin\gamma^{(0)}$, 结合式 (44)、式 (46) 和式 (48) 可知, 上述各项形式极其复杂, 必须对其进行预处理, 否则无法推导解析解。

与式 (22) 和式 (38) 中对升力 L的处理类似, 将零阶阻力 $D^{(0)}$拆分为动压、参考面积与阻力系数乘积的形式, 即

$$D^{(0)} = 0.5\rho^{(0)}\left(V^{(0)}\right)^2 S_{\mathrm{ref}} C_D^{(0)} \tag{54}$$

与升力系数的分析同理, 阻力系数也可以表示为 $\sin\alpha$的偶数次幂级数形式:

$$C_D = k_0^{C_D}(Ma) + k_2^{C_D}(Ma)\sin^2\alpha + k_4^{C_D}(Ma)\cdot\sin^4\alpha \tag{55}$$

为了保证精度且更符合实际, 对零阶阻力系数 $C_D^{(0)}$而言, 假设上式中 $k_0^{C_D}$、$k_2^{C_D}$、$k_4^{C_D}$均为零阶马赫数 $Ma^{(0)}$的函数。$Ma^{(0)}$是零阶速度除以由零阶高度计算得到的零阶声速, 所以其形式也较为复杂。为了便于推导, 构造切比雪夫插值多项式分别近似以下项:

$$Ma^{(0)} \approx \tilde{Q}_{Ma}^{(0)} = \sum_{i=0}^{N} \tilde{k}_i^{Ma^{(0)}} \tilde{m}^i \tag{56}$$

$$\frac{0.5\rho^{(0)}\left(V^{(0)}\right)^2}{m} \approx \tilde{Q}_{\rho V}^{(0)} = \sum_{i=0}^{N} \tilde{k}_i^{\rho V^{(0)}} \tilde{m}^i \tag{57}$$

$$\frac{P_{\mathrm{a}}^{(0)}}{m} \approx \tilde{Q}_{P_{\mathrm{a}}}^{(0)} = \sum_{i=0}^{N} \tilde{k}_i^{P_{\mathrm{a}}^{(0)}} \tilde{m}^i \tag{58}$$

$$g^{(0)}\sin\gamma^{(0)} \approx \tilde{Q}_g^{(0)} = \sum_{i=0}^{N} \tilde{k}_i^{g^{(0)}} \tilde{m}^i \tag{59}$$

使用式 (56) 和式 (57) 所示近似多项式的另一个优势在于, 推导弹道倾角一阶解时可以直接使用上述拟合公式, 减少需要近似变量的数目, 既降低了计算量, 又提升了精度。

利用式 (52) 将 $Q_{\sin\alpha}$、$Q_{\cos\alpha}$和 Q_g转换成归一化质量 $\tilde{m}$的函数 $\tilde{Q}_{\sin\alpha}$、$\tilde{Q}_{\cos\alpha}$和 $\tilde{Q}_g$。再将其与式 (15)、式 (55) ～ 式 (59) 代入式 (41) 可得

$$\frac{\mathrm{d}V^{(1)}}{\mathrm{d}m} = -\frac{k_0^{\hat{D}}}{q_m}\frac{1}{K_{m0}+K_{m1}\tilde{m}} + \sum_{i=0}^{2N+12} \tilde{f}_i^{V^{(1)}} \tilde{m}^i \tag{60}$$

式中: $\tilde{f}_i^{V^{(1)}}$为多项式系数。对上式进行积分可得

$$V^{(1)} = \sum_{i=0}^{2N+13} \tilde{F}_i^{V^{(1)}} \tilde{m}^i + F_{\ln}^{V^{(1)}} \ln(K_{m0}+K_{m1}\tilde{m}) \tag{61}$$

式中: $\tilde{F}_i^{V^{(1)}}$和 $F_{\ln}^{V^{(1)}}$为多项式系数。

5.2.3　弹道倾角一阶解

为保证精度, 在式 (21) 的基础上对式 (42) 中的升力系数 C_L使用关于 $\sin\alpha$的三次函数近似。采用与阻力系数 C_D相似的方法, 将拟合系数假设为 $Ma^{(0)}$的函数, 即

$$C_L = k_1^{C_L}(Ma)\sin\alpha + k_3^{C_L}(Ma)\sin^3\alpha \tag{62}$$

然后, 对式 (42) 中的 $\frac{P_0 Q_{\sin\alpha}}{m q_m}\left(Q_{\hat{V}} - \frac{1}{V^{(0)}}\right)$项进行处理。该项分母中含有真实质量 m, 且 $V^{(0)}$的形式较为复杂, 若直接转换为归一化质量 $\tilde{m}$, 则无法积分得到解析解。所以先使用切比雪夫插值多项式近似 $1/V^{(0)}$, 形式如下:

$$\frac{1}{V^{(0)}} \approx Q_V^{(0)} = \sum_{i=0}^{N} k_i^{V^{(0)}} m^i \tag{63}$$

将式(13)、式(20) 和式(63) 代入 $\frac{P_0 Q_{\sin\alpha}}{m q_m}\left(Q_{\hat{V}} - \frac{1}{V^{(0)}}\right)$, 并利用式 (52) 将 $Q_{\rho\hat{V}}$、$Q_{\hat{V}}$、$Q_{\cos\gamma(2)}$、$Q_{P_{\mathrm{a}}(2)}$、$V^{(0)}$和该项转换成归一化质量 $\tilde{m}$的函数。再使用切比雪夫插值多项式近似 $\cos\gamma^{(0)}$, 形式如下:

$$\cos\gamma^{(0)} \approx \tilde{Q}_{\cos\gamma}^{(0)} = \sum_{i=0}^{N} \tilde{k}_i^{\cos\gamma^{(0)}} \tilde{m}^i \tag{64}$$

由于速度一阶解 $V^{(1)}$远小于速度零阶解 $V^{(0)}$, 式 (42) 中的 $\frac{V^{(1)}}{R q_m} Q_{\cos\gamma(2)}$项可以被忽略。将式 (21)、式 (56) 和式 (57)、式 (62)、式 (64) 代入式 (42), 并将 $\gamma^{(1)}$分解为 $\gamma_1^{(1)}$和 $\gamma_2^{(1)}$两部分, 整理可得

$$\frac{\mathrm{d}\gamma_1^{(1)}}{\mathrm{d}m} = \frac{1}{q_m}\frac{P_0 c_0\left(k_0^{\hat{V}} - k_0^{V^{(0)}}\right)}{K_{m0}+K_{m1}\tilde{m}} + \sum_{i=0}^{N_1} \tilde{f}_i^{\gamma^{(1)}} \tilde{m}^i \tag{65}$$

$$\frac{\mathrm{d}\gamma_2^{(1)}}{\mathrm{d}m} = \ln(K_{m0}+K_{m1}\tilde{m}) \sum_{i=0}^{N} \tilde{f}_{\ln,i}^{\gamma^{(1)}} \tilde{m}^i \tag{66}$$

式中: $N_1 = \max\{2N+3, N+11\}$。观察式 (66) 可以发现 $\ln(K_{m0}+K_{m1}\tilde{m})\sum_{i=0}^{N} \tilde{f}_{\ln,i}^{\gamma^{(1)}} \tilde{m}^i$项较难求得解析解, 且其阶数相对较小, 所以使用式 (53) 的转换关系, 式 (66) 可被改写为

$$\frac{\mathrm{d}\gamma_2^{(1)}}{\mathrm{d}m} = \ln m \sum_{i=0}^{N} f_{\ln,i}^{\gamma^{(1)}} m^i \tag{67}$$

分别对式 (65) 和式 (67) 进行积分, 将二者相加可得弹道倾角一阶解:

$$\gamma^{(1)}=\gamma_1^{(1)}+\gamma_2^{(1)}=F_{\ln}^{\gamma^{(1)}}\ln(K_{m0}+K_{m1}\tilde{m})+\sum_{i=0}^{N_1+1}\tilde{F}_i^{\gamma^{(1)}}\tilde{m}^i+\sum_{i=0}^{N+1}F_i^{\gamma^{(1)}}m^i(i\ln m-1)\tag{68}$$

式中：$F_{\ln}^{\gamma^{(1)}}$、$\tilde{F}_i^{\gamma^{(1)}}$和$F_i^{\gamma^{(1)}}$为多项式系数。

5.2.4 航程一阶解

式(39)中含有$\cos\gamma^{(0)}$项，结合式(46)可知其积分后无法得到解析解，因此利用式(64)对$\cos\gamma^{(0)}$进行近似。将式(44)、式(46)、式(61)和式(68)代入式(39)，经过代数运算可知其中包含以下项：

$$\begin{cases}\displaystyle\sum_{i=0}^{N+3}\tilde{F}_i^{V^{(0)}}\tilde{m}^i\sum_{i=0}^{N+1}F_i^{\gamma^{(1)}}m^i(i\ln m-1)\\ \ln(K_{m0}+K_{m1}\tilde{m})\displaystyle\sum_{i=0}^{N+1}F_i^{\gamma^{(1)}}m^i(i\ln m-1)\\ \ln(K_{m0}+K_{m1}\tilde{m})\displaystyle\sum_{i=0}^{2N+15}k_i^{x^{(1)}}\tilde{m}^i\end{cases}\tag{69}$$

前两项积分后无法得到解析解，第三项先使用式(53)进行转换后再积分，虽然可以得到解析解，但阶数较高会引入较大舍入误差，导致解析解失效。因此，需要再次使用切比雪夫插值多项式近似$V^{(0)}\gamma^{(1)}$和$\left(V^{(0)}-V^{(1)}\right)\gamma^{(0)}$项，形式如下：

$$\begin{cases}V^{(0)}\gamma^{(1)}\approx\tilde{Q}_{V\gamma1}^{(1)}=\displaystyle\sum_{i=0}^{N}\tilde{k}_i^{V\gamma1}\tilde{m}^i\\ \left(V^{(0)}-V^{(1)}\right)\gamma^{(0)}\approx\tilde{Q}_{V\gamma2}^{(1)}=\displaystyle\sum_{i=0}^{N}\tilde{k}_i^{V\gamma2}\tilde{m}^i\end{cases}\tag{70}$$

将式(44)、式(46)、式(61)、式(64)、式(68)和式(70)代入式(39)，可以获得与弹道倾角一阶微分方程形式类似的表达式，采用相同的推导方法，将$x^{(1)}$分解为$x_1^{(1)}$和$x_2^{(1)}$两部分，其表达式分别为

$$\frac{\mathrm{d}x_1^{(1)}}{\mathrm{d}m}=\sum_{i=0}^{2N+13}\tilde{f}_i^{x^{(1)}}\tilde{m}^i+\frac{\cos\hat{\gamma}^*+\hat{\gamma}^*\sin\hat{\gamma}^*}{q_m}\cdot\left(F_{\ln}^{V^{(0)}}-F_{\ln}^{V^{(1)}}\right)\ln(K_{m0}+K_{m1}\tilde{m})\tag{71}$$

$$\frac{\mathrm{d}x_2^{(1)}}{\mathrm{d}m}=\ln(K_{m0}+K_{m1}\tilde{m})\sum_{i=0}^{N}\tilde{f}_{\ln,i}^{x^{(1)}}\tilde{m}^i\tag{72}$$

与式(66)处理类似，对式(72)先使用式(53)的转换关系可得

$$\frac{\mathrm{d}x_2^{(1)}}{\mathrm{d}m}=\ln m\sum_{i=0}^{N}f_{\ln,i}^{x^{(1)}}m^i\tag{73}$$

分别对式(71)和式(73)积分，二者求和可得航程一阶解：

$$x^{(1)}=x_1^{(1)}+x_2^{(1)}=F_{\ln}^{x^{(1)}}\left(\tilde{m}+\frac{K_{m0}}{K_{m1}}\right)\ln(K_{m0}+K_{m1}\tilde{m})+\sum_{i=0}^{2N+14}\tilde{F}_i^{x^{(1)}}\tilde{m}^i+\sum_{i=0}^{N+1}F_i^{x^{(1)}}m^i(i\ln m-1)\tag{74}$$

式中：$F_{\ln}^{x^{(1)}}$、$\tilde{F}_i^{x^{(1)}}$和$F_i^{x^{(1)}}$为多项式系数。

5.2.5 地心距一阶解

式(40)中含有$\sin\gamma^{(0)}$项，结合式(46)可知其积分后无法得到解析解，因此需要先对其进行近似：

$$\sin\gamma^{(0)}\approx\tilde{Q}_{\sin\gamma}^{(0)}=\sum_{i=0}^{N}\tilde{k}_i^{\sin\gamma^{(0)}}\tilde{m}^i\tag{75}$$

将式(44)、式(46)、式(61)、式(68)、式(70)和式(75)代入式(40)，可以获得与航程一阶微分方程形式类似的表达式，采用相同的推导方法，将$R^{(1)}$分解为$R_1^{(1)}$和$R_2^{(1)}$两部分，其表达式分别为

$$\frac{\mathrm{d}R_1^{(1)}}{\mathrm{d}m}=\sum_{i=0}^{2N+13}\tilde{f}_i^{R^{(1)}}\tilde{m}^i+\frac{\sin\hat{\gamma}^*-\hat{\gamma}^*\cos\hat{\gamma}^*}{q_m}\cdot\left(F_{\ln}^{V^{(0)}}-F_{\ln}^{V^{(1)}}\right)\ln(K_{m0}+K_{m1}\tilde{m})\tag{76}$$

$$\frac{\mathrm{d}R_2^{(1)}}{\mathrm{d}m}=\ln(K_{m0}+K_{m1}\tilde{m})\sum_{i=0}^{N}\tilde{f}_{\ln,i}^{R^{(1)}}\tilde{m}^i\tag{77}$$

与式(66)处理类似，对式(77)先使用式(53)的转换关系可得

$$\frac{\mathrm{d}R_2^{(1)}}{\mathrm{d}m}=\ln m\sum_{i=0}^{N}f_{\ln,i}^{R^{(1)}}m^i\tag{78}$$

分别对式(76)和式(78)积分，二者求和可得地心距一阶解：

$$R^{(1)}=R_1^{(1)}+R_2^{(1)}=F_{\ln}^{R^{(1)}}\left(\tilde{m}+\frac{K_{m0}}{K_{m1}}\right)\ln(K_{m0}+K_{m1}\tilde{m})+\sum_{i=0}^{2N+14}\tilde{F}_i^{R^{(1)}}\tilde{m}^i+\sum_{i=0}^{N+1}F_i^{R^{(1)}}m^i(i\ln m-1)\tag{79}$$

式中：$F_{\ln}^{R^{(1)}}$、$\tilde{F}_i^{R^{(1)}}$和$F_i^{R^{(1)}}$为多项式系数。

6 结果与讨论

6.1 阶次分析

在本节中，为了说明一阶解的必要性，将对比零阶解和一阶解的结果，并以数值弹道仿真作为基准。以米诺陶IV运载火箭[19]的第一级为飞行器模型，仿真的初始条件为：初始航程$x_0=840$ m；初始高度$H_0=6\,520$ m；初始速度$v_0=472$ m/s；初始弹道倾角$\gamma_0=79°$；初始质量$m_0=77\,937$ kg；质量流量$q_m=797.2$ kg/s；真空推力$P_0=2\,204.5$ kN；发动机工作时间为43 s；仿真攻角剖面参数为$c_0=-6.968\,3$；$c_1=4.086\,4\times10^{-4}$；$c_2=-7.752\,0\times10^{-9}$；$c_3=4.691\,0\times10^{-14}$，如图2所示。仿真结果曲线如图3所示。表1展示了终端误差及计算效率。其中，数值仿真结果是对式(1)～式(7)采用四阶龙格库塔法进行数值积分得到的。在弹道仿真中，使用USSA—1976标准大气模型[20]。

随着发动机工作，火箭质量逐渐减小，因此在

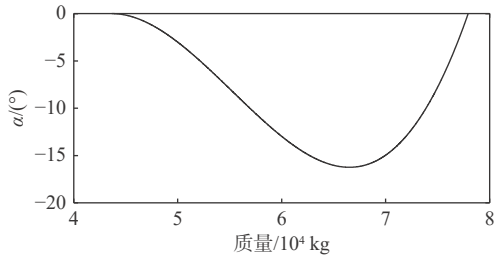


图 2　算例一攻角剖面

Fig. 2　AOA profile for case 1

图 2 和图 3 中, 曲线右端点是初始状态, 左端点是终端状态。由图 2 可知, 攻角剖面满足初始和终端时刻收敛到零的要求, 符合工程实际规律。从表 1 的结果可以看出, 一阶近似解的终端误差均小于零阶近似解。针对弹道倾角和高度而言, 一阶近似解的修正更加明显。针对速度而言, 由于零阶近似解的精度已经较高, 一阶近似解的精度提升并不明显。但即使较小的助推段关机点速度误差也会导致较大的落点误差, 因此使用一阶解修正是有意义的。从计算时间上看, 解析解的计算时间比数值仿真低 4 个数量级, 极大地提高了助推段弹道预测的效率。一阶解的计算时间约为零阶解的二倍。即使如此, 与一阶解析解精度提升的意义相比, 其计算效率降低的不足是完全可以承受的。

6.2　不同助推段解析解对比

为了展示本文解析解的优势, 将其与现有的主动段解析解进行对比, 分别是文献 [13] 和文献 [14] 中的解析解。为了更全面地验证运载火箭的实际工作场景, 本文算例中使用与前文不同的初始条件。以米诺陶 IV 运载火箭的第三级为飞行器模型, 仿真初始条件为: 初始航程 $x_0 = 11\,640$ m; 初始高度 $H_0 = 45\,500$ m; 初始速度 $v_0 = 1\,600$ m/s; 初始弹道倾角 $\gamma_0 = 72°$; 初始质量 $m_0 = 39\,380.2$ kg; 质量流量 $q_m = 403.7$ kg/s; 真空推力 $P_0 = 1\,223.3$ kN; 发动机工作时间为 61 s; 攻角剖面参数: $c_0 = -2.071\,3$, $c_1 = 3.788\,3\times10^{-4}$, $c_2 = -2.087\,6\times10^{-8}$, $c_3 = 3.197\,5\times10^{-13}$, 如图 4 所示。仿真对比结果如图 5 所示。

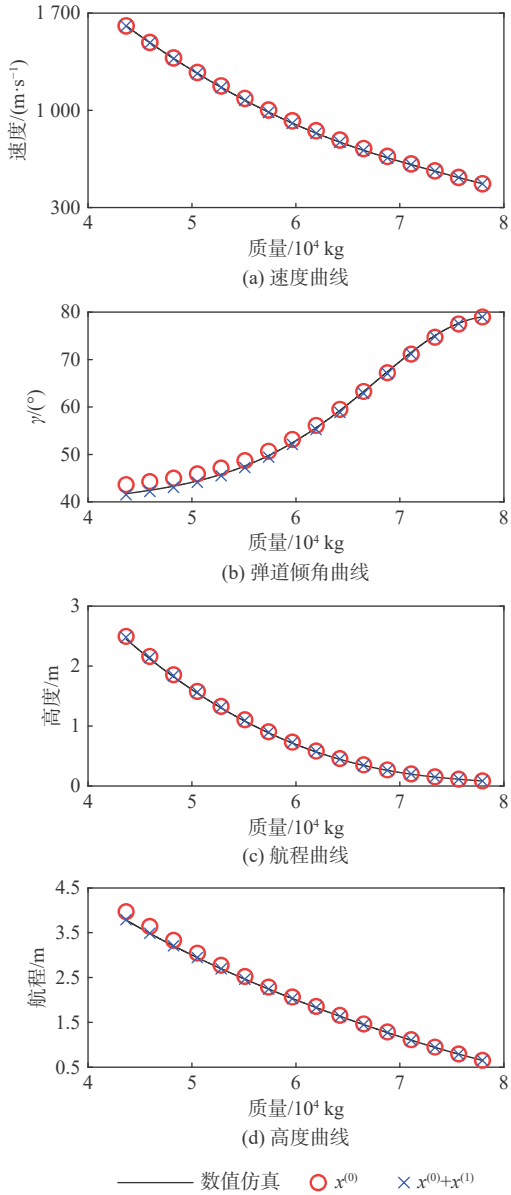


图 3　算例一解析解精度对比

Fig. 3　Analytical solution results and trajectory simulation for case 1

文献 [13] 中的解析解是基于零攻角假设推导的, 因此在弹道倾角和高度上均有最大误差。在大攻角场景中, 文献 [13] 中解析解的弹道倾角误差达到40°, 高度误差达到54 km。这对于助推段弹道预

表 1　算例一解析解终端误差及计算时间

Table 1　Terminal results and computing time for case 1

速度/(m·s⁻¹)			速度相对误差1/%	速度相对误差2/%	弹道倾角/(°)			弹道倾角相对误差1/%	弹道倾角相对误差2/%
数值仿真	$x^{(0)}$	$x^{(0)}+x^{(1)}$			数值仿真	$x^{(0)}$	$x^{(0)}+x^{(1)}$		
1 609.5	1 607.6	1 607.9	−0.120 2	−0.098 9	41.728 4	43.475 6	41.435 4	4.187 0	−0.702 3

航程/km			航程相对误差1/%	航程相对误差2/%	高度/km			高度倾角相对误差1/%	高度倾角相对误差2/%	计算时间/s		
数值仿真	$x^{(0)}$	$x^{(0)}+x^{(1)}$			数值仿真	$x^{(0)}$	$x^{(0)}+x^{(1)}$			数值仿真	$x^{(0)}$	$x^{(0)}+x^{(1)}$
24.507 8	24.995 0	24.819 7	1.988 0	1.272 8	37.833 5	39.671 7	37.869 5	4.858 6	0.095 0	0.819 4	$1.563\,4\times10^{-5}$	$3.472\,0\times10^{-5}$

注: 相对误差1是$x^{(0)}$相对数值仿真的误差; 相对误差2是$x^{(0)}+x^{(1)}$相对数值仿真的误差。

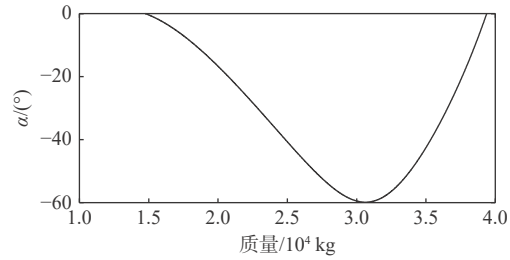
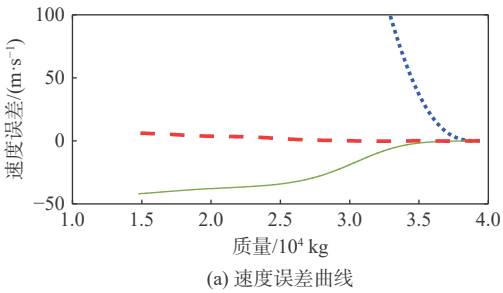
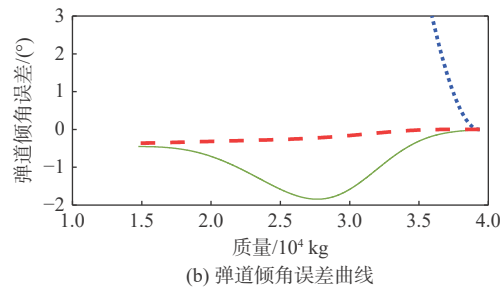


图4 算例二攻角剖面

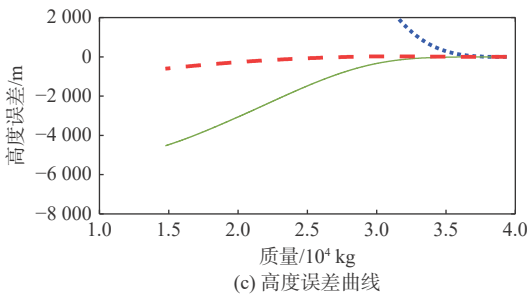
Fig. 4 AOA profile for case 2



(a) 速度误差曲线



(b) 弹道倾角误差曲线



(c) 高度误差曲线



图5 算例二不同解析解误差对比

Fig. 5 Errors in different analytical solutions for case 2

测问题而言，是不可接受的结果。针对速度而言，文献 [13] 中解析解的速度误差也会超过400 m/s。因此，针对助推段弹道预测的问题，文献 [13] 中的解析解并不适合绝大多数场景。

虽然文献 [14] 中解析解弹道倾角的终端误差与本文解析解基本一致，但中间时刻的误差会显著大于本文解析解。本文解析解在速度和高度上的误差远小于文献 [14] 中的解析解，终端精度至少可以提升85%。

为了进一步验证本文提出的解析解对不同初始条件和不同量级攻角剖面的适应性，采用新的算例对本文解析解、文献 [13] 和文献 [14] 的解析解进行对比分析。仿真对象为米诺陶 IV 运载火箭的第三级，初始条件为：初始航程 $x_0 = 12\,000$ m；初始高度 $H_0 = 34\,000$ m；初始速度 $v_0 = 1\,826$ m/s；初始弹道倾角 $\gamma_0 = 60°$。图6展示了攻角剖面，由图可知，此算例中运载火箭以小攻角飞行。仿真结果曲线见图7，即使在小攻角情况下，本文解析解的预测精度也明显高于现有的两种解析解，终端精度至少可以提升89%。

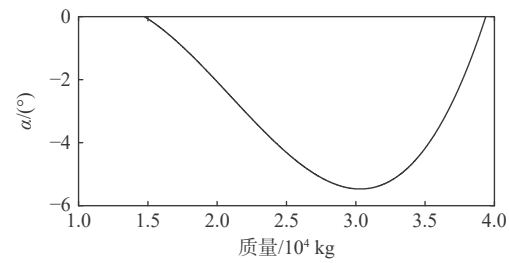
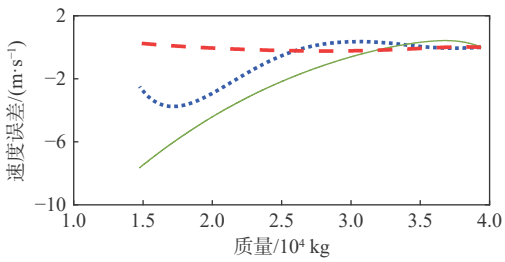
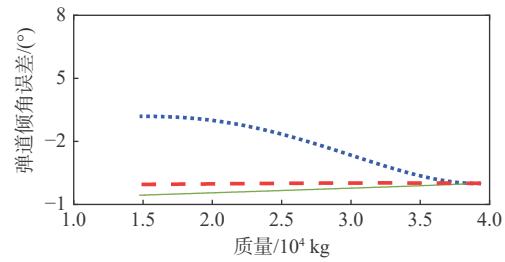


图6 算例三攻角剖面

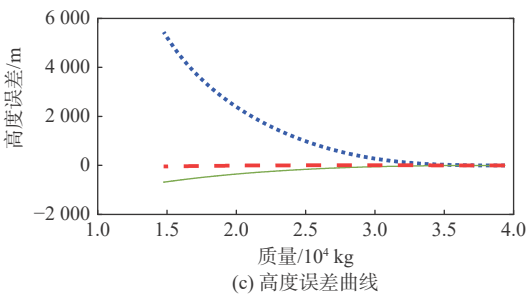
Fig. 6 AOA profile of case 3



(a) 速度误差曲线



(b) 弹道倾角误差曲线



(c) 高度误差曲线

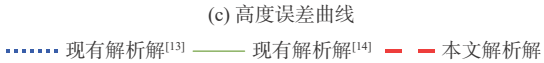


图7 算例三不同解析解误差对比

Fig. 7 Errors in different analytical solutions for case 3

7 结 论

本文针对大攻角下运载火箭助推段在线快速

预测困难的问题, 基于摄动法近似求解大攻角剖面下高度非线性的助推段动力学模型, 得到了高精度的弹道解析解, 仿真结果表明:

1) 一阶解析解的最大相对误差不超过 1.28%, 计算时间与数值仿真相比降低了 4 个数量级, 大幅提升了运载火箭助推段弹道预测的速度和精度。

2) 在大攻角机动飞行的条件下, 所提出的解析解在速度、高度和弹道倾角上的精度与现有解析解相比至少提升了 85%。

在今后的研究中, 可以基于该高精度弹道解析解设计能够在线应用的运载火箭助推段制导方法, 当存在状态偏差时能在保证终端约束的前提下满足多种复杂约束条件。

参考文献（References）

[1] 马树微, 李静琳, 陈曦, 等. 多级固体运载火箭分级多学科设计优化[J]. 北京航空航天大学学报, 2016, 42(3): 542-550.
MA S W, LI J L, CHEN X, et al. Stage-wise multidisciplinary design optimization for multi-stage solid launch vehicle[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(3): 542-550 (in Chinese).

[2] 程晓明, 尚腾, 徐帆, 等. 基于序列凸规划的运载火箭轨迹在线规划方法[J]. 宇航学报, 2021, 42(2): 202-210.
HENG X M, SHANG T, XU F, et al. Online trajectory planning for launch vehicles with successive convex programming[J]. Journal of Astronautics, 2021, 42(2): 202-210 (in Chinese).

[3] CHENG X M, LI H F, ZHANG R. Efficient ascent trajectory optimization using convex models based on the Newton-Kantorovich/pseudospectral approach[J]. Aerospace Science and Technology, 2017, 66: 140-151.

[4] SUN Z X, CHAO T, WANG S Y, et al. Convex method for ascent trajectory optimization using iterative narrowing trust region[C]// Proceedings of the 2018 IEEE CSAA Guidance, Navigation and Control Conference. Piscataway: IEEE Press, 2018: 1-6.

[5] PRABHAKAR N, KUMAR I, TATA SK, et al. A predictive explicit guidance scheme for ballistic missiles[J]. Defence Science Journal, 2013, 63(5): 456-461.

[6] CAO R, LIU Y B, LU Y P. Robust multiple model predictive control for ascent trajectory tracking of aerospace vehicles[J]. IEEE Transactions on Aerospace and Electronic Systems, 2022, 58(2): 1333-1351.

[7] 李惠峰, 谢陵. 基于预测校正方法的 RLV 再入制导律设计[J]. 北京航空航天大学学报, 2009, 35(11): 1344-1348.
LI H F, XIE L. Reentry guidance law design for RLV based on predictor-corrector method[J]. Journal of Beijing University of Aeronautics and Astronautics, 2009, 35(11): 1344-1348 (in Chinese).

[8] YU W B, CHEN W C, JIANG Z G, et al. Omnidirectional autonomous entry guidance based on 3-D analytical glide formulas[J]. ISA Transactions, 2016, 65: 487-503.

[9] YU W B, CHEN W C, JIANG Z G, et al. Analytical entry guidance for no-fly-zone avoidance[J]. Aerospace Science and Technology, 2018, 72: 426-442.

[10] SPRENGER A R, JAHANSHAHI S, IVLEV A V, et al. Time-dependent inertia of self-propelled particles: the Langevin rocket[J]. Physical Review E, 2021, 103(4): 042601.

[11] 胡锦川. 高超声速飞行器平稳滑翔弹道特性及全程弹道规划与制导方法研究[D]. 北京: 北京航空航天大学, 2015.
HU J C. Research on steady-glide trajectory characteristics and algorithm design of trajectory planning and explicit guidance for hypersonic vehicle[D]. Beijing: Beihang University, 2015(in Chinese).

[12] ZHAO S L, CHEN W C, YANG L. Endoatmospheric ascent optimal guidance with analytical nonlinear trajectory prediction[J]. International Journal of Aerospace Engineering, 2022, 2022(1): 5729335.

[13] ZHANG Q, XU Z. Autonomous ascent guidance with multiple terminal constraints for all-solid launch vehicles[J]. Aerospace Science and Technology, 2020, 97: 105633.

[14] YU Q, CHEN W C, YU W B. Approximate analytical solutions for launch-vehicle ascent trajectory[J]. IEEE Transactions on Aerospace and Electronic Systems, 2023, 59(4): 4033-4048.

[15] VEPA R. Flight dynamics, simulation, and control: for rigid and flexible aircraft[M]. 2nd ed. Boca Raton: CRC Press, 2023.

[16] YU W B, CHEN W C, JIANG Z G, et al. Analytical entry guidance based on pseudo-aerodynamic profiles[J]. Aerospace Science and Technology, 2017, 66: 315-331.

[17] 孙勇强, 佟明羲, 王鹏飞, 等. 国外固体运载火箭主动力系统发展研究[J]. 宇航总体技术, 2022, 6(4): 56-64.
SUN Y Q, TONG M X, WANG P F, et al. Research on the development of the main propulsion system of foreign solid launch vehicles[J]. Astronautical Systems Engineering Technology, 2022, 6(4): 56-64(in Chinese).

[18] 陈克俊, 刘鲁华, 孟云鹤. 远程火箭飞行动力学与制导[M]. 北京: 国防工业出版社, 2014.
CHEN K J, LIU L H, MENG Y H. Launch vehicle flight dynamics and guidance[M]. Beijing: National Defense Industry Press, 2014(in Chinese).

[19] NORTHROP GRUMMAN INNOVATION SYSTEMS. Minotaur IV•V•VI User's Guide[EB/OL]. (2023-07-30)[2013-06-09]. https://www.manualsdir.com/manuals/569830/orbital-minotaur-rimibotaur-v-minotaur-iv.html.

[20] 陈赟. 运载火箭助推段轨迹优化与制导方法研究[D]. 武汉: 华中科技大学, 2019.
CHEN B. Research on trajectory optimization and guidance method for the booster of launching vehicle[D]. Wuhan: HuaZhong University of Science and Technology, 2019(in Chinese).

Analytical solutions for ascent trajectory under high angle-of-attack conditions

ZHANG Keyang[1], WANG Peng[2], YU Qi[1], YU Wenbin[1,*], CHEN Wanchun[1]

(1. School of Astronautics, Beihang University, Beijing 102206, China;
2. Beijing Institute of Aerospace Systems Engineering, Beijing 100076, China)

Abstract: To rapidly predict the ascent trajectory of launch vehicles, high-accuracy analytical solutions for ascent trajectory were proposed under high angle-of-attack (AOA) conditions. First, a simplified longitudinal-plane dynamics model with mass as the independent variable was developed. The sine of AOA was innovatively expressed as a polynomial of mass. Due to the high AOA, the simplified dynamics model remains highly nonlinear, which prevents it from being analytically solved directly. Approximate polynomials are introduced to replace the strongly nonlinear but relatively small terms in the original equations through force analysis. Furthermore, the difference between the true values and the approximations was treated as a minor perturbation. The dynamics model was divided into analytically solvable subsystems according to the perturbation theory. Analytical solutions for velocity, flight-path angle, downrange, and altitude were derived by solving the subsystems. Simulation results confirm that the proposed solutions are at least 85% more accurate than existing solutions under high AOA conditions.

Keywords: analytical solution; ascent trajectory; high angle of attack; launch vehicle; perturbation method

Received: 2023-08-17; **Accepted**: 2023-10-24; **Published Online**: 2023-11-01 16:05
URL: link.cnki.net/urlid/11.2625.V.20231101.1418.004

* **Corresponding author**. E-mail: yuwenbin_buaa@163.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0553

加权几何平均组合预测方法对数均方误差的分解

朱双磊[1]，陈华友[1,2,*]，张琦[2]，王一帆[3]

(1. 安徽大学 大数据与统计学院，合肥 230601；2. 安徽大学 数学科学学院，合肥 230601；
3. 安徽大学 纽约石溪学院，合肥 230039)

摘 要：针对非线性的加权几何平均组合预测模型，引入对数均方误差，并提出了对数均值偏差、模型解释力系数、对数离差误差方差及预测方法差异性测度等概念。在此基础上，将组合预测对数均方误差分解为对数均值偏差、因模型解释力不足而产生的对数方差以及基于模型解释力系数的对数离差方差3个组成部分，从理论上探讨了对数均方误差的来源。同时，将组合预测对数均方误差分解为单项预测方法对数均方误差的加权平均和预测方法差异性测度的加权平均两部分，获得了有益的结论：提高单项预测方法的精度和预测方法差异性测度水平有利于减少加权几何平均组合预测方法的对数均方误差，为组合预测单项方法的遴选提供了理论支持。通过实际案例，分析了各组成部分以及单项预测方法之间的差异性测度对组合预测对数均方误差的影响。

关 键 词：组合预测；加权几何平均；对数均方误差；分解；差异性测度
中图分类号：O211.67
文献标志码：A **文章编号**：1001-5965(2025)10-3535-12

随着社会经济系统的快速发展，预测的复杂性和不确定性也在显著增加；在实际的应用中单个方法有多个影响因素，包括模型的表达差异、数据来源的差异，很难得到一个稳定的预测结果。因此结合不同单项预测模型，利用算术加权或几何加权平均，可以融合多元信息，避免预测产生的风险，这就是组合预测提出的动机。Bates等[1]在1969年提出了组合预测的概念，发现组合预测方法相对于单项预测方法在预测的稳定性上有明显的提高[2-3]，并且在最优组合预测模型中，若至少有2种单项预测方法的最优权重大于0，则在某种误差准则下组合预测方法表现较优。由于组合预测能够有效地减少预测误差，因此受到了国内外学者的广泛研究，并且取得了大量的显著成果[4-9]，在多种领域得到了普遍的应用[10-14]。文献[15]基于相关系数建立了最优组合预测模型；文献[16]探讨了基于Theil不等系数的组合预测模型及其若干性质；文献[17]构建了基于向量夹角余弦的组合预测模型；文献[18]基于预测有效度建立了改进的组合预测方法；以上根据不同准则建立的组合预测模型，探究了组合预测的有效性，并提出了优性组合预测等若干概念。然而传统的组合预测模型对单项方法是固定权的加权平均，但是由于单项方法随着时间的变化表现出“时好时坏”的特点，即传统的组合预测不能反映单项预测方法的精度随着时间变化的特征。因此，文献[19]建立了基于诱导有序加权平均(induced ordered weighted averaging, IOWA)算子的组合预测模型；文献[20]在非线性的诱导有序加权调和平均(Induced ordered weighted harmonic averaging, IOWHA)算子的基础上建立了组合预测模型；文献[21]探讨了基于诱导有序几何加权平均(induced ordered weighted geometric averaging, IOWGA)算子

收稿日期：2022-08-29；**录用日期**：2023-10-29；**网络出版时间**：2023-11-24 15:24
网络出版地址：link.cnki.net/urlid/11.2625.V.20231123.1407.003
基金项目：国家自然科学基金（72371001）；安徽大学创新训练项目
*通信作者. E-mail：huayouc@126.com

引用格式：朱双磊，陈华友，张琦，等. 加权几何平均组合预测方法对数均方误差的分解[J]. 北京航空航天大学学报，2025，51(10)：3535-3546. ZHU S L，CHEN H Y，ZHANG Q，et al. Decomposition of logarithm mean square error of weighted geometric mean combined forecasting method[J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51(10)：3535-3546 (in Chinese).

组合预测方法的若干性质。文献 [22]在不确定环境下,以最大化灰色关联度的精度准则,建立区间组合预测模型,按照 Shapley 值法进行单项预测模型的权重求解,并通过实例说明了组合预测在灰色关联度信息准则下的有效性。文献 [23]为了解决传统组合预测带来的过拟合问题,定义了全新的刻画组合预测模型的极端偏差指标,并以此为目标函数建立了组合预测,有效地提高了模型的泛化能力。随着计算机技术的发展,人们发现将智能算法与组合预测相结合能够取得更好的预测效果,文献 [24] 采用变分模态分解-稀疏自编码器 (variable mode decomposition-sparse autoencoder, VMD-SAE) 提取原始风速特征对数据进行平滑处理和信息提取,再将高阶模糊认知映射 (high-order fuzzy cognitive mapping, HFCM) 神经网络和批量梯度下降优化算法结合建立组合预测模型,利用不同方法的优势实现准确稳定的风速预测结果。文献 [25] 利用 3 种分解方法对数据进行处理,并采用基于最大综合灰色关联度聚类 (comprehensive grey correlation degree clustering, CGCD)对数据分量重构,从而建立了将分解-重构-集成方法与改进的自然全局优化算法相结合的农产品价格预测最优组合框架,有效地降低了预测计算成本,提高了预测性能。文献 [26] 针对短时交通流提出利用自适应噪声完全集合经验模态分解 (complete ensemble empirical mode decomposition with adaptive noise, CEEMDAN), 将时间序列分解为具有不同频率的模态函数的组合预测方法。文献 [27] 建立了基于误差分解和线性非线性框架的组合预测模型,通过自回归滑动平均模型 (autoregressive integrated moving average model, ARMIA)线性建模获得残差,利用非线性的多层感知器和支持向量机对残差分解后的子序列组合预测。Thomson 等[28] 通过误差平方和的分解得到了均方误差的组成部分,对减少组合预测均方误差有着重要的意义。文中结论显示,单项方法所体现的信息越不相关,则组合预测误差会减少越多。同时,给出了一致性等相关指标,以此来衡量组合中单项方法之间所产生的信息不相关水平和差异程度。然而文献 [28] 是针对线性加权平均建立的组合预测模型,探讨了均方误差的 3 个组成部分,但单项方法在组合预测中的贡献度是不同的。因此,文献 [29] 分解了加权算术平均组合预测模型的均方误差,发现简单算术平均是加权算术平均均方误差分解的一个特例。同样, Kang 等[30] 将加权算术平均组合预测均方误差分解为单项方法均方误差的加权平均和单项方法之间的多样性指标,得到组合预测的均方误差一定小于或等于单个预测方法均方误差的加权平均,并且具有更大多样性的组合预测方法将具有更低的总体均方误差,即预测方法池中单项方法的种类越多,整体预测精度越高。

上述文献是基于简单或加权算术平均所构建的组合预测模型进行均方误差的组分来源分析。实际上除了线性组合之外,更多的预测系统满足非线性特征。由加权几何平均形式建立的组合预测模型是非线性的一种有效表达。文献 [31] 通过实例给出了该种形式组合模型有效性的一个良好的例证。且文献 [32] 表明了加权算术平均组合预测模型中的误差是绝对误差;加权几何平均组合预测模型中的误差是对数误差,反映的是预测的相对误差;在赋权的过程中加权算术平均组合预测只考虑了绝对误差而没有考虑相对误差,加权几何平均组合预测只考虑了相对误差而没有考虑绝对误差;所以 2 种组合方法在提高预测精度上是从 2 个不同的角度出发的。并且通过分析 2 种误差的关系,可知 2 种组合方法产生的误差的绝对值大小对其在组合预测中最优权重上有着决定性的影响。因此,在选择单项预测方法上,这 2 种不同形式的组合预测模型是完全不同的。已有文献 [28-30] 仅对加权算术平均组合预测均方误差进行了分解,即从绝对误差的角度进行的分解;而目前还没有从相对误差的角度对误差的来源进行探究。所以,有必要对加权几何平均组合预测所产生的对数均方误差的组分来源进行探究。根据文献 [29-30], 为了探讨对数均方误差的来源,定义了对数均值偏差等 3 个指标,进而推导出组合预测模型对数均方误差的组分的恒等式。并且,将组合预测的对数均方误差分解为单项预测方法对数均方误差的加权平均和各单项方法之间的差异性指标。最后通过案例,分析讨论了组合预测对数均方误差各组成部分和单项预测方法差异性对组合预测对数均方误差的影响。

1 最优组合预测模型的构建

设$X=\{x_t,t=1,2,\cdots,N\}$为实际原始时间序列,不同单项方法的预测值序列为$X_i=\{x_{it},t=1,2,\cdots,N\}$, $i=1,2,\cdots,m$, 同时设m种单项预测参与组合预测,相应的权系数为$w_1,w_2,\cdots,w_m$, 且满足:

$$\sum_{i=1}^{m} w_i = 1 \qquad w_i \geqslant 0, i = 1,2,\cdots,m \tag{1}$$

根据加权几何平均数的公式,令

$$\hat{x}_t = \prod_{i=1}^{m} x_{it}^{w_i} \qquad t = 1,2,\cdots,N \tag{2}$$

则称 $\hat{x}_t$ 为加权几何平均的组合预测值, 令 $\hat{X} = \{\hat{x}_t, t = 1,2,\cdots,N\}$ 为 t 时刻的组合预测序列。将式 (2) 变为

$$\ln \hat{x}_t = \sum_{i=1}^{m} w_i \ln x_{it} \qquad t = 1,2,\cdots,N \tag{3}$$

设 $\ln X = \{\ln x_t, t = 1,2,\cdots,N\}$ 为实际值对数序列, 第 i 种单项方法在第 t 时刻的预测值对数序列为 $\ln X_i = \{\ln x_{it}, t = 1,2,\cdots,N\}, i = 1,2,\cdots,m$, $\ln \hat{X} = \{\ln \hat{x}_t, t = 1,2,\cdots,N\}$ 为 t 时刻的组合预测值对数序列。

定义 1　令

$$e_{it} = \ln x_t - \ln x_{it} \tag{4}$$

$$e_t = \ln x_t - \ln x_{it} = \sum_{i=1}^{m} w_i (\ln x_t - \ln x_{it}) = \sum_{i=1}^{m} w_i e_{it} \tag{5}$$

则称 e_{it} 为第 i 种单项预测方法在第 t 时刻的对数误差, e_t 为第 t 时刻组合预测方法的对数误差。

定义 2　令

$$E_i = \sum_{t=1}^{N} e_{it}^2 \qquad i = 1,2,\cdots,m \tag{6}$$

$$E = \sum_{t=1}^{N} e_t^2 = \sum_{t=1}^{N} \left(\sum_{i=1}^{m} w_i e_{it} \right)^2 \tag{7}$$

则称 E_i 为第 i 种单项预测方法的对数误差平方和; E 为组合预测方法的对数误差平方和。

显然, 从组合预测误差的角度衡量预测的有效性, E 越小表示组合预测方法的精度越高。因此以误差平方和来确定组合预测中单项方法的权重系数, 就要极小化误差平方和; 于是基于误差平方和的加权几何平均组合预测模型可表示为

$$\begin{aligned} &\min E(w) = \sum_{t=1}^{N} \left(\sum_{i=1}^{m} w_i e_{it} \right)^2 \\ &\text{s.t.} \begin{cases} \sum_{i=1}^{m} w_i = 1 \\ w_i \geqslant 0 \qquad i = 1,2,\cdots,m \end{cases} \end{aligned} \tag{8}$$

式 (8) 最后求得的解就是该模型的最优权重。

2　相关概念和性质

在对预测性能的优劣情况的评判, 均方误差是一个常规有意义的指标; 对于加权几何平均组合预测模型, 引入相对均方误差 (L) 作为新的评价标准。

定义 3　令

$$L_i = \frac{1}{N} \sum_{t=1}^{N} (\ln x_{it} - \ln x_t)^2 \qquad i = 1,2,\cdots,m \tag{9}$$

$$L = \frac{1}{N} \sum_{t=1}^{N} (\ln \hat{x}_t - \ln x_t)^2 \tag{10}$$

式中: L_i 为第 i 种单项方法产生的对数均方误差; L 为加权几何组合预测方法产生的对数均方误差。

定义 4　令

$$\overline{\ln X_i} = \frac{1}{N} \sum_{t=1}^{N} \ln x_{it} \qquad i = 1,2,\cdots,m \tag{11}$$

$$\overline{\ln \hat{X}} = \frac{1}{N} \sum_{t=1}^{N} \ln \hat{x}_t \tag{12}$$

$$\overline{\ln X} = \frac{1}{N} \sum_{t=1}^{N} \ln x_t \tag{13}$$

则称 $\overline{\ln X_i}$ 为第 i 种单项预测方法的对数均值序列, $\overline{\ln \hat{X}}$ 为加权几何组合预测方法的对数均值序列, $\overline{\ln X}$ 为实际值的对数均值序列。

由式 (11)～式 (13) 可知:

$$\sum_{t=1}^{N} (\ln x_{it} - \overline{\ln X_i}) = \sum_{t=1}^{N} \ln x_{it} - N\overline{\ln X_i} = 0 \tag{14}$$

$$\sum_{t=1}^{N} (\ln \hat{x}_t - \overline{\ln \hat{X}}) = \sum_{t=1}^{N} \ln \hat{x}_t - N\overline{\ln \hat{X}} = 0 \tag{15}$$

$$\sum_{t=1}^{N} (\ln x_t - \overline{\ln X}) = \sum_{t=1}^{N} \ln x_t - N\overline{\ln X} = 0 \tag{16}$$

性质 1　组合预测与各个单项方法对数均值序列成立以下等式:

$$\overline{\ln \hat{X}} = \sum_{i=1}^{m} w_i \overline{\ln X_i} \tag{17}$$

证明　由式 (3)、式 (11)、式 (12) 可知:

$$\overline{\ln \hat{X}} = \frac{1}{N} \sum_{t=1}^{N} \ln \hat{x}_t = \frac{1}{N} \sum_{t=1}^{N} \left(\sum_{i=1}^{m} w_i \ln x_{it} \right) = \sum_{i=1}^{m} w_i \left(\frac{1}{N} \sum_{t=1}^{N} \ln x_{it} \right) = \sum_{i=1}^{m} w_i \overline{\ln X_i}$$

从而式 (17) 结论成立。

定义 5　令

$$B_i = (\overline{\ln X_i} - \overline{\ln X})^2 \qquad i = 1,2,\cdots,m \tag{18}$$

$$B = (\overline{\ln \hat{X}} - \overline{\ln X})^2 \tag{19}$$

式中: B_i、B 分别为单项方法对应的对数均值偏差和组合预测对应的对数均值偏差。

显然, B_i 表示单项方法对数序列均值与实际值对数序列均值之差, 该值越大表示差异越大, B 表示相应组合预测的情形。

定义6 令

$$\mathrm{Cov}(\ln X_i,\ln X)=\frac{1}{N}\sum_{t=1}^{N}(\ln x_{it}-\overline{\ln X_i})(\ln x_t-\overline{\ln X}) \tag{20}$$

$$\mathrm{Cov}(\ln \hat{X},\ln X)=\frac{1}{N}\sum_{t=1}^{N}(\ln \hat{x}_t-\overline{\ln \hat{X}})(\ln x_t-\overline{\ln X}) \tag{21}$$

$$V(\ln X)=\frac{1}{N}\sum_{t=1}^{N}(\ln x_t-\overline{\ln X})^2 \tag{22}$$

式中：$\mathrm{Cov}(\ln X_i,\ln X)$、$\mathrm{Cov}(\ln \hat{X},\ln X)$分别为单项方法与组合预测对应的对数协方差；$V(\ln X)$为实际对数序列的方差。

性质2 $\mathrm{Cov}(\ln X_i,\ln X)$与$\mathrm{Cov}(\ln \hat{X},\ln X)$之间成立如下等式：

$$\mathrm{Cov}(\ln \hat{X},\ln X)=\sum_{i=1}^{m}w_i\mathrm{Cov}(\ln X_i,\ln X) \tag{23}$$

证明 由式(20)、式(21)和性质1可知：

$$\mathrm{Cov}(\ln \hat{X},\ln X)=\frac{1}{N}\sum_{t=1}^{N}(\ln \hat{x}_t-\overline{\ln \hat{X}})(\ln x_t-\overline{\ln X})=$$
$$\frac{1}{N}\sum_{t=1}^{N}\left(\sum_{i=1}^{m}w_i\ln x_{it}-\sum_{i=1}^{m}w_i\overline{\ln X_i}\right)(\ln x_t-\overline{\ln X})=$$
$$\frac{1}{N}\sum_{t=1}^{N}\left(\sum_{i=1}^{m}w_i(\ln x_{it}-\overline{\ln X_i}\right)(\ln x_t-\overline{\ln X}))=$$
$$\sum_{i=1}^{m}w_i\left(\frac{1}{N}\sum_{t=1}^{N}(\ln x_{it}-\overline{\ln X_i}\right)(\ln x_t-\overline{\ln X}))=$$
$$\sum_{i=1}^{m}w_i\mathrm{Cov}(\ln X_i,\ln X)$$

从而性质2成立。

定义7 令

$$P_i=\frac{\mathrm{Cov}(\ln X_i,\ln X)}{V(\ln X)} \tag{24}$$

$$P=\frac{\mathrm{Cov}(\ln \hat{X},\ln X)}{V(\ln X)} \tag{25}$$

式中：P_i为单项方法i对应的模型解释力系数；P为组合预测对应的模型解释力系数。

从方差的角度分析，在定义7中P_i表示单项方法对实际值对数序列的解释程度；同理P表示组合预测方法对实际值对数序列的解释程度。

性质3 两类模型解释力系数满足以下等式：

$$P=\sum_{i=1}^{m}w_iP_i \tag{26}$$

证明 由式(25)和性质2可知：

$$P=\frac{\mathrm{Cov}(\ln \hat{X},\ln X)}{V(\ln X)}=\frac{\sum_{i=1}^{m}w_i\mathrm{Cov}(\ln X_i,\ln X)}{V(\ln X)}=$$
$$\sum_{i=1}^{m}w_i\frac{\mathrm{Cov}(\ln X_i,\ln X)}{V(\ln X)}=\sum_{i=1}^{m}w_iP_i$$

从而性质3成立。

定义8 令

$$R_i=(1-P_i)^2V(\ln X) \tag{27}$$

$$R=(1-P)^2V(\ln X) \tag{28}$$

式中：R_i和R分别为单项方法和组合预测在模型解释力缺乏的条件下所产生的对数方差。

定义8给出了R_i和R误差来源的原因分析。

定义9 令

$$u_{it}=(\ln x_{it}-\overline{\ln X_i})-P_i(\ln x_t-\overline{\ln X})$$
$$i=1,2,\cdots,m;t=1,2,\cdots,N \tag{29}$$

$$\hat{u}_t=(\ln \hat{x}_t-\overline{\ln \hat{X}})-P(\ln x_t-\overline{\ln X})$$
$$t=1,2,\cdots,N \tag{30}$$

$U_i=\{u_{it},t=1,2,\cdots,N\}$和$\hat{U}=\{\hat{u}_t,t=1,2,\cdots,N\}$分别定义为单项方法与组合预测对应的对数离差误差序列。

性质4 上述两类对数离差误差序列成立如下等式，即

$$\hat{u}_t=\sum_{i=1}^{m}w_iu_{it}\qquad t=1,2,\cdots,N \tag{31}$$

证明 由式(17)、式(26)、式(29)和式(30)可知：

$$\hat{u}_t=(\ln \hat{x}_t-\overline{\ln \hat{X}})-P(\ln x_t-\overline{\ln X})=$$
$$\left(\sum_{i=1}^{m}w_i\ln x_{it}-\sum_{i=1}^{m}w_i\overline{\ln X_i}\right)-\sum_{i=1}^{m}P_i(\ln x_t-\overline{\ln X})=$$
$$\sum_{i=1}^{m}w_i\left[(\ln x_{it}-\overline{\ln X_i})-P_i(\ln x_t-\overline{\ln X})\right]=\sum_{i=1}^{m}w_iu_{it}$$

即性质4得证。

性质5 单项方法与组合预测对应的对数离差误差序列的均值为0，即

$$\begin{cases}\overline{U}_i=0\\ \overline{\hat{U}}=0\end{cases} \tag{32}$$

证明 由均值，以及式(14)～式(16)和式(29)知：

$$\overline{U}_i=\frac{1}{N}\sum_{t=1}^{N}u_{it}=$$
$$\frac{1}{N}\sum_{t=1}^{N}(\ln x_{it}-\overline{\ln X_i})-\frac{P_i}{N}\sum_{t=1}^{N}(\ln x_t-\overline{\ln X})=0$$

同理，由式(14)～式(16)和式(30)可知：

$$\overline{\hat{U}}=\frac{1}{N}\sum_{t=1}^{N}\hat{u}_t=$$

$$\frac{1}{N}\sum_{t=1}^{N}(\ln\hat{x}_t-\overline{\ln\hat{X}})-\frac{P}{N}\sum_{t=1}^{N}(\ln x_t-\overline{\ln X})=0$$

即性质 5 得证。

定义 10　令

$$D_i=V(U_i) \tag{33}$$

$$D=V(\hat{U}) \tag{34}$$

式中: D_i和D分别为单项方法与组合预测i所产生的对数离差误差的方差。

定义 10 给出了D_i和D对数离差误差的波动性的度量。

3　加权几何平均组合预测模型对数均方误差的分解

定理 1　加权几何平均组合预测模型的对数均方误差来源于 3 个组分之和, 即

$$L=B+R+D \tag{35}$$

式中: B为对应的对数均值偏差; R为对应的对数方差; D为对应的对数离差误差的方差。

证明　根据对数均方误差表达式, 可得

$$L=\frac{1}{N}\sum_{t=1}^{N}(\ln\hat{x}-\ln x_t)^2=$$

$$\frac{1}{N}\sum_{t=1}^{N}\left[(\ln\hat{x}_t-\overline{\ln\hat{X}})+(\overline{\ln\hat{X}}-\overline{\ln X})-(\ln x_t-\overline{\ln X})\right]^2=$$

$$\frac{1}{N}\sum_{t=1}^{N}(\ln\hat{x}_t-\overline{\ln\hat{X}})^2+\frac{1}{N}\sum_{t=1}^{N}(\overline{\ln\hat{X}}-\overline{\ln X})^2-$$

$$\frac{1}{N}\sum_{t=1}^{N}(\ln x_t-\overline{\ln X})^2+\frac{2(\overline{\ln\hat{X}}-\overline{\ln X})}{N}\sum_{t=1}^{N}(\ln\hat{x}_t-\overline{\ln\hat{X}})-$$

$$\frac{2(\overline{\ln\hat{X}}-\overline{\ln X})}{N}\sum_{t=1}^{N}(\ln x_t-\overline{\ln X})-$$

$$\frac{2}{N}\sum_{t=1}^{N}(\ln\hat{x}_t-\overline{\ln\hat{X}})(\ln x_t-\overline{\ln X})=$$

$$V(\ln\hat{X})+B+V(\ln X)+0-0-2\mathrm{Cov}(\ln\hat{X},\ln X)=$$

$$V(\ln\hat{X})+B+V(\ln X)-2\mathrm{Cov}(\ln\hat{X},\ln X)$$

由式 (25) 可知:

$$\mathrm{Cov}(\ln\hat{X},\ln X)=PV(\ln X)$$

因此,

$$L=V(\ln\hat{X})+B+V(\ln X)-PV(\ln X)=$$

$$V(\ln X)+B+(1-2P)V(\ln X) \tag{36}$$

由式 (28) 可知:

$$R=(1-P)^2V(X)=(1-2P+P^2)V(X)=$$

$$(1-2P)V(X)+P^2V(X) \tag{37}$$

结合式 (36) 和式 (37) 可得

$$L=V(\ln\hat{X})+B+V(\ln X)-PV(\ln X)=$$

$$V(\ln\hat{X})+B+V-P^2V(\ln X) \tag{38}$$

由式 (30)、式 (32)、式 (34) 及式 (25) 可知:

$$D=V(\hat{U})=\frac{1}{N}\sum_{t=1}^{N}(\hat{u}_t-\overline{\hat{U}})^2=$$

$$\frac{1}{N}\sum_{t=1}^{N}[(\hat{u}_t)^2-2\hat{u}_t\overline{\hat{U}}+(\overline{\hat{U}})^2]=\frac{1}{N}\sum_{t=1}^{N}(\hat{u}_t)^2=$$

$$\frac{1}{N}\sum_{t=1}^{N}(\ln\hat{x}_t-\overline{\ln\hat{X}})^2-$$

$$\frac{2P}{N}\sum_{t=1}^{N}(\ln\hat{x}_t-\overline{\ln\hat{X}})(\ln x_t-\overline{\ln X})+$$

$$\frac{2P}{N}\sum_{t=1}^{N}(\ln x_t-\overline{\ln X})^2=$$

$$V(\ln\hat{X})-2P\mathrm{Cov}(\ln\hat{X},\ln X)+P^2V(\ln X)=$$

$$V(\ln\hat{X})-2P^2V(X)+P^2V(\ln X)=V(\ln\hat{X})-P^2V(X)$$

将上式代入式 (38) 可得

$$L=B+R+D \tag{39}$$

因此定理 1 得证。

从定理 1 可以看出, 加权几何平均组合预测方法的对数均方误差主要来自 3 部分。①表示组合预测方法对应的对数均值偏差; ②表示的是组合预测方法对应的对数方差; ③表示的是对数离差误差。

4　基于差异性测度的对数均方误差的来源解析

定义 11　令

$$C_{ij}=\frac{1}{N}\sum_{t=1}^{N}(\ln x_{it}-\ln x_{jt})^2 \tag{40}$$

式中: C_{ij}为组合预测方法池中的第 i 与第 j 种单项预测方法之间差异性指标。显然, C_{ij}表示 2 种单项预测方法之间差异性的程度, 反映两者之间信息独立性的程度。

定义 12　令

$$B_{ij}=(\overline{\ln X_i}-\overline{\ln X_j})^2 \tag{41}$$

式中: B_{ij}为第 i 与第 j 种单项预测方法之间的对数均值偏差。

从均值的角度分析, 定义 12 中的B_{ij}反映了 2 种单项预测方法对数序列的均值之间的差异性, 即反映了 2 种单项预测方法预测值对数序列中心

的差异性。

定义 13 令

$$R_{ij}=(P_i-P_j)^2V(\ln X) \tag{42}$$

式中：R_{ij}为第 i 与第 j 种单项预测方法预测值对数序列因模型解释力不足而产生的对数方差。

从定义 13 可以看出，R_{ij}反映了 2 个单项预测方法之间的解释力性能之间的差异性程度，当其等于 0 时表示 2 种方法的解释力相同，即对实际值的解释程度相同。

定义 14 令

$$D_{ij}=V(U_i-U_j) \tag{43}$$

式中：D_{ij}为第 i 与第 j 种单项预测方法之间的对数离差误差方差。从方差的角度，D_{ij}表示 2 种单项方法的误差的变动，即 2 种预测方法差异的波动性。

定理 2 加权几何平均组合预测和单项方法的对数值序列加权误差平方和满足下述关系式：

$$\sum_{i=1}^{m}w_i(\ln x_{it}-\ln\hat{x}_i)^2=\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}w_iw_j(\ln x_{it}-\ln x_{jt})^2 \tag{44}$$

证明 已知权重满足$\sum_{i=1}^{m}w_i=1$，则有

$$\sum_{i=1}^{m}w_i(\ln x_{it}-\ln\hat{x}_t)^2=$$

$$\sum_{i=1}^{m}w_i\left[(\ln x_{it})^2-2\ln x_{it}\ln\hat{x}_t+(\ln\hat{x}_t)^2\right]=$$

$$\sum_{i=1}^{m}w_i(\ln x_{it})^2-2\ln\hat{x}_t\sum_{i=1}^{m}w_i\ln x_{it}+$$

$$\sum_{i=1}^{m}w_i(\ln\hat{x}_t)^2=\sum_{i=1}^{m}w_i(\ln x_{it})^2-2(\ln\hat{x}_t)^2+(\ln\hat{x}_t)^2=$$

$$\sum_{i=1}^{m}w_i(\ln x_{it})^2-(\ln\hat{x}_t)^2=\sum_{i=1}^{m}w_i(\ln x_{it})^2-$$

$$\left[\sum_{i=1}^{m}w_i^2(\ln x_{it})^2+2\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}w_iw_j\ln x_{it}\ln x_{jt}\right]=$$

$$\sum_{i=1}^{m}w_i(l-w_i)(\ln x_{it})^2-2\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}w_iw_j\ln x_{it}\ln x_{jt}=$$

$$w_1(w_2+w_3+\cdots+w_m)(\ln x_{1t})^2+\cdots+w_m(w_1+$$

$$w_3+\cdots+w_{m-1})(\ln x_{mt})^2-2w_1w_2\ln x_{1t}\ln x_{2t}-$$

$$2w_1w_3\ln x_{1t}\ln x_{3t}-\cdots-2w_{m-1}w_m\ln x_{m-1t}\ln x_{mt}=$$

$$[w_1w_2(\ln x_{1t})^2+w_1w_2(\ln x_{2t})^2-$$

$$2w_1w_2\ln x_{1t}\ln x_{2t}]+\cdots+[w_{m-1}w_m(\ln x_{m-1t})^2+$$

$$w_{m-1}w_m(\ln x_{mt})^2-2w_{m-1}w_m\ln x_{m-1t}\ln x_{mt}]=$$

$$w_1w_2(\ln x_{1t}-\ln x_{2t})^2+w_1w_3(\ln x_{1t}-\ln x_{3t})^2+\cdots+$$

$$w_{m-1}w_m(\ln x_{m-1t}-\ln x_{mt})^2=$$

$$\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}w_iw_j(\ln x_{it}-\ln x_{jt})^2$$

因此定理 2 得证。

类似于定理 2 的证明，可得定理 3～定理 5 的结论也是成立的，证明从略。

定理 3 单项方法与加权几何平均组合预测的对数均值序列，满足如下等式：

$$\sum_{i=1}^{m}w_i(\overline{\ln X_i}-\overline{\ln\hat{X}})^2=\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}w_iw_j(\overline{\ln X_i}-\overline{\ln X_j})^2 \tag{45}$$

定理 4 单项预测方法预测值与实际值序列的模型解释力系数与加权几何平均组合预测值与实际值序列的模型解释力系数，满足如下等式：

$$\sum_{i=1}^{m}w_i(P_i-P)^2=\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}w_iw_j(P_i-P_j)^2 \tag{46}$$

定理 5 单项方法和加权几何平均组合方法对应的对数离差误差序列满足下列等式：

$$\sum_{i=1}^{m}w_i(U_i-\hat{U})^2=\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}w_iw_j(U_i-U_j)^2 \tag{47}$$

定理 6 加权几何平均组合预测方法的对数均方误差可分解为如下形式：

$$L=\sum_{i=1}^{m}w_iL_i-\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}w_iw_jC_{ij} \tag{48}$$

式中：L_i为第 i 种单项预测方法的对数均方误差；C_{ij}为单项方法之间的差异性指标。

证明 由式 (10) L的定义可知：

$$L=\frac{1}{N}\sum_{t=1}^{N}(\ln\hat{x}_t-\ln x_t)^2=$$

$$\frac{1}{N}\sum_{t=1}^{N}\left[(\ln\hat{x}_t)^2-2\ln\hat{x}_t\ln x_t+(\ln x_t)^2\right]=$$

$$\frac{1}{N}\sum_{t=1}^{N}\left[\sum_{i=1}^{m}w_i(\ln\hat{x}_t)^2-2\sum_{i=1}^{m}w_i\ln x_{it}\ln x_t+\right.$$

$$\left.\sum_{i=1}^{m}w_i(\ln x_t)^2\right]=\frac{1}{N}\sum_{t=1}^{N}\left[\sum_{i=1}^{m}w_i[(\ln\hat{x}_t)^2-\right.$$

$$\left.2\ln x_{it}\ln x_t+(\ln x_t)^2]\right]=\frac{1}{N}\sum_{t=1}^{N}\sum_{i=1}^{m}w_i(\ln x_{it}-\ln x_t)^2-$$

$$\frac{1}{N}\sum_{t=1}^{N}\sum_{i=1}^{m}w_i[(\ln x_{it})^2-(\ln\hat{x}_t)^2]=$$

$$V(\ln\hat{X})+B+V(\ln X)+0-0-2\mathrm{Cov}(\ln\hat{X},\ln X)=$$

$$V(\ln\hat{X})+B+V(\ln X)-2\mathrm{Cov}(\ln\hat{X},\ln X) \tag{49}$$

因为 $\ln\hat{x}_t=\sum_{i=1}^{m}w_i\ln x_{it}$，所以有 $(\ln\hat{x}_t)^2=\sum_{i=1}^{m}w_i\ln x_{it}\ln\hat{x}_t$，从而可得

$$\frac{1}{N}\sum_{t=1}^{N}\sum_{i=1}^{m}w_i[(\ln x_{it})^2-(\ln \hat{x}_t)^2]=$$
$$\frac{1}{N}\sum_{t=1}^{N}\left[\sum_{i=1}^{m}w_i(\ln x_{it})^2-2\sum_{i=1}^{m}w_i\ln x_{it}\ln \hat{x}_t+\sum_{i=1}^{m}w_i(\ln \hat{x}_t)^2\right]=\frac{1}{N}\sum_{t=1}^{N}\sum_{i=1}^{m}w_i(\ln x_{it}-\ln \hat{x}_t)^2 \quad (50)$$

注意到式 (40) 和式 (44), 从而有

$$\frac{1}{N}\sum_{t=1}^{N}\sum_{i=1}^{m}w_i(\ln x_{it}-\ln \hat{x}_t)^2=$$
$$\frac{1}{N}\sum_{t=1}^{N}\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}w_iw_j(\ln x_{it}-\ln x_{jt})^2=$$
$$\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}w_iw_j\left[\frac{1}{N}\sum_{t=1}^{N}(\ln x_{it}-\ln x_{jt})^2\right]=$$
$$\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}w_iw_jC_{ij} \quad (51)$$

结合式(50)和式(51), 则有

$$\frac{1}{N}\sum_{t=1}^{N}\sum_{i=1}^{m}w_i\left[(\ln x_{it})^2-(\ln \hat{x}_t)^2\right]=\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}w_iw_jC_{ij} \quad (52)$$

将式 (52) 代入式 (49) 可得

$$L=\sum_{i=1}^{m}w_iL_i-\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}w_iw_jC_{ij}$$

从而定理 6 得证。

从定理 6 的式 (48) 可以看出, 加权几何平均组合预测的对数均方误差可以分解成 2 部分之差。第一部分是单项预测方法对数均方误差的加权平均, 第二部分为单项预测方法之间的差异性指标对应的加权平均。显然第二部分是非负的, 要减少加权几何平均组合预测的对数均方误差, 可以通过 2 个途径, 一方面, 尽量选取较为准确的单项预测方法参与到组合预测中, 即降低第 i 种单项预测方法 L_i; 另一方面, 要增加第二部分量的大小, 即增加单项方法之间的差异性指标 C_{ij}, 该定理表明: 当单项方法的对数均方误差越小, 单项方法的差异性越大, 组合预测将具有更低的对数均方误差, 即有更高的预测精度。

定理 7　组合预测方法的对数均值偏差 B 可分解为如下形式:

$$B=\sum_{i=1}^{m}w_iB_i-\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}w_iw_jB_{ij} \quad (53)$$

证明　由式 (17)、式 (19) 可知:

$$B=(\overline{\ln \hat{X}}-\overline{\ln X})^2=(\overline{\ln \hat{X}})^2-2(\overline{\ln X})(\overline{\ln \hat{X}})+(\overline{\ln X})^2=$$
$$\sum_{i=1}^{m}w_i(\overline{\ln \hat{X}})^2-2\sum_{i=1}^{m}w_i\overline{\ln X_i\ln X}+\sum_{i=1}^{m}w_i(\overline{\ln X})^2=$$
$$\sum_{i=1}^{m}w_i\left[(\overline{\ln \hat{X}})^2-2\overline{\ln X_i\ln X}+(\overline{\ln X})^2\right]=$$
$$\sum_{i=1}^{m}w_i\left[(\overline{\ln X_i})^2-2\overline{\ln X_i\ln X}+(\overline{\ln X})^2-(\overline{\ln X_i})^2+(\overline{\ln \hat{X}})^2\right]=\sum_{i=1}^{m}w_i(\overline{\ln X_i}-\overline{\ln X})^2-$$
$$\sum_{i=1}^{m}w_i\left[(\overline{\ln X_i})^2-(\overline{\ln \hat{X}})^2\right]=$$
$$\sum_{i=1}^{m}w_iB_i-\sum_{i=1}^{m}w_i\left[(\overline{\ln X_i})^2-(\overline{\ln \hat{X}})^2\right] \quad (54)$$

由式 (19) 可知: $\overline{\ln \hat{X}}=\sum_{i=1}^{m}w_i\overline{\ln X_i}$, 所以有 $(\overline{\ln \hat{X}})^2=\sum_{i=1}^{m}w_i\overline{\ln X_i}\overline{\ln \hat{X}}$, 从而可得

$$\sum_{i=1}^{m}w_i\left[(\overline{\ln X_i})^2-(\overline{\ln \hat{X}})^2\right]=$$
$$\sum_{i=1}^{m}w_i(\overline{\ln X_i})^2-2\sum_{i=1}^{m}w_i\overline{\ln X_i}\overline{\ln \hat{X}}+\sum_{i=1}^{m}w_i(\overline{\ln \hat{X}})^2=$$
$$\sum_{i=1}^{m}w_i(\overline{\ln X_i}-\overline{\ln \hat{X}})^2 \quad (55)$$

将式 (41) 和式 (45) 代入式 (55) 可得

$$\sum_{i=1}^{m}w_i(\overline{\ln X_i}-\overline{\ln \hat{X}})^2=$$
$$\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}w_iw_j(\overline{\ln X_i}-\overline{\ln X_j})^2=\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}w_iw_jB_{ij} \quad (56)$$

则有

$$\sum_{i=1}^{m}w_i\left[(\overline{\ln X_i})^2-(\overline{\ln \hat{X}})^2\right]=\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}w_iw_jB_{ij} \quad (57)$$

将式 (57) 代入式 (54) 可得

$$B=\sum_{i=1}^{m}w_iB_i-\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}w_iw_jB_{ij}$$

定理 7 得证。

从定理 7 中可以看出, 加权几何平均组合预测的对数均值偏差一定小于或等于单项预测方法对数的均值偏差的加权平均。当单项方法与实际值对数序列的中心差异越小, 且各单项预测方法对数序列中心的差异性越大, 则组合预测对数序列与实际值对数序列的中心差异性就越小。

定理 8　组合预测对应的对数方差 R 可分解为如下形式:

$$R=\sum_{i=1}^{m}w_iR_i-\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}w_iw_jR_{ij} \tag{58}$$

证明 由式(26)～式(28)可知:

$$R=(1-P)^2V(\ln X)=(1-2P+P^2)V(\ln X)=$$
$$\left(1-2\sum_{i=1}^{m}w_iP_i+\sum_{i=1}^{m}w_iP_i^2-\sum_{i=1}^{m}w_iP_i^2+P^2\right)\cdot$$
$$V(\ln X)=\sum_{i=1}^{m}w_i(1-2P_i+P_i^2)V(\ln X)-$$
$$\left(\sum_{i=1}^{m}w_iP_i^2-P^2\right)V(\ln X)=$$
$$\sum_{i=1}^{m}w_i(1-P_i)V(\ln X)-\sum_{i=1}^{m}w_i(P_i^2-P^2)V(\ln X)=$$
$$\sum_{i=1}^{m}w_iR_i-\sum_{i=1}^{m}w_i(P_i^2-P^2)V(\ln X) \tag{59}$$

注意到, $P=\sum_{i=1}^{m}w_iP_i$, 所以有 $P^2=\sum_{i=1}^{m}w_iP_iP$, 因此可得

$$\sum_{i=1}^{m}w_i(P_i^2-P^2)V(\ln X)=\left(\sum_{i=1}^{m}w_iP_i^2-2\sum_{i=1}^{m}w_iP_iP+\sum_{i=1}^{m}w_iP^2\right)V(\ln X)=\sum_{i=1}^{m}w_i(P_i-P)^2V(\ln X) \tag{60}$$

将式(42)和式(46)代入式(60)可得

$$\sum_{i=1}^{m}w_i(P_i-P)^2V(\ln X)=\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}w_iw_j(P_i-P_j)^2V(\ln X)=\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}w_iw_jR_{ij} \tag{61}$$

则有

$$\sum_{i=1}^{m}w_i(P_i^2-P^2)V(\ln X)=\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}w_iw_jR_{ij} \tag{62}$$

将式(62)代入式(59)可得

$$R=\sum_{i=1}^{m}w_iR_i-\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}w_iw_jR_{ij}$$

定理8得证。

从定理8中可以看出,由于第二项的非负性,加权几何平均组合预测因模型解释力不足而产生的对数方差一定小于或等于单项预测方法的对数方差的加权平均。当单项预测方法之间的解释能力越大时,组合预测因模型解释力不足而产生的对数方差就越小,组合预测所产生的不可解释的误差就越小。

定理9 组合预测方法预测值对数序列基于解释力系数的离差误差方差D可分解为如下形式:

$$D=\sum_{i=1}^{M}w_iD_i-\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}w_iw_jD_{ij} \tag{63}$$

证明 由式(31)、式(33)、式(34)可知:

$$D=V(\hat{U})=\frac{1}{N}\sum_{t=1}^{N}(\hat{U}-\overline{\hat{U}})^2=\frac{1}{N}\sum_{t=1}^{N}(\hat{U})^2=$$
$$\frac{1}{N}\sum_{i=1}^{m}\hat{u}_t^2=\frac{1}{N}\sum_{i=1}^{m}(2\hat{u}_t^2-\hat{u}_t^2)=$$
$$\frac{1}{N}\sum_{t=1}^{N}\sum_{i=1}^{m}w_i(2u_{it}\hat{u}_t-\hat{u}_t^2)=$$
$$\frac{1}{N}\sum_{t=1}^{N}\sum_{i=1}^{m}w_i[u_{it}^2-(u_{it}^2-2u_{it}\hat{u}_t+\hat{u}_t^2)]=$$
$$\frac{1}{N}\sum_{t=1}^{N}\sum_{i=1}^{m}w_iu_{it}^2-\frac{1}{N}\sum_{t=1}^{N}\sum_{i=1}^{m}w_i(u_{it}-\hat{u}_t)^2=$$
$$\sum_{i=1}^{m}w_i\left(\frac{1}{N}\sum_{t=1}^{N}u_{it}^2\right)-\frac{1}{N}\sum_{t=1}^{N}\sum_{i=1}^{m}w_i(u_{it}-\hat{u}_t)^2=$$
$$\sum_{i=1}^{m}w_iV(U_i)-\frac{1}{N}\sum_{t=1}^{N}\sum_{i=1}^{m}w_i(u_{it}-\hat{u}_t)^2 \tag{64}$$

由式(32)可知: $\overline{U}_i=0,\overline{\hat{U}}=0$, 因此有

$$V(U_i-U_j)=\frac{1}{N}\sum_{i=1}^{N}\left[\left(u_{it}-u_{jt}\right)-\left(\overline{u_{it}}-\overline{u_{jt}}\right)\right]^2=$$
$$\frac{1}{N}\sum_{i=1}^{N}\left(u_{it}-u_{jt}\right)^2 \tag{65}$$

因此,由式(43)、式(47)和式(65)可知:

$$\frac{1}{N}\sum_{t=1}^{N}\sum_{i=1}^{m}w_i(u_{it}-\hat{u}_t)^2=$$
$$\frac{1}{N}\sum_{t=1}^{N}\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}w_iw_j(u_{it}-u_{jt})^2=$$
$$\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}w_iw_jV(U_i-U_j)=\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}w_iw_jD_{ij} \tag{66}$$

将式(66)代入式(64)可得

$$D=\sum_{i=1}^{M}w_iD_i-\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}w_iw_jD_{ij}$$

定理9得证。

定理9中表明加权几何平均组合预测基于解释力系数的对数离差误差一定小于或等于单项预测方法基于解释力系数的对数离差误差的加权平均。从第二项可以发现,当单项预测方法之间基于解释力系数的对数离差误差方差的波动性越大,组合预测的对数离差误差方差就越小。

5　案例分析

为了说明均方误差各组件以及差异性对其的影响, 本文使用文献 [33] 预测的 2000—2015 年的人口数据进行实证分析。分别将文献 [27] 所用的 4 种方法表示为: F1, F2, F3, F4。将这些单项方法进行加权几何平均组合, 通过式 (8) 求得各单项方法的权重系数; F12、F13、F14、F23、F24、F34 表示单项方法的两两组合, F123、F124、F134、F234 为 3 个单项方法的组合, 使用 4 个单项方法的组合为 F1234。实际值、单项方法的预测值以及对数误差, 分别如表 1 和表 2 所示。其中 e_1、e_2、e_3、e_4 分别表示单项预测方法与实际值之间的对数误差。

表 1　实际值及单项方法预测值

Table 1　Actual value and predicted value of single method

实际值	预测值			
	F1	F2	F3	F4
6 278	6 335	6 280	6 325	6 256
6 325	6 379	6 321	6 372	6 311
6 369	6 370	6 454	6 419	6 362
6 410	6 415	6 522	6 466	6 415
6 461	6 660	6 459	6 512	6 467
6 516	6 663	6 514	6 559	6 519
6 593	6 736	6 590	6 605	6 571
6 676	6 683	6 792	6 651	6 624
6 741	6 858	6 725	6 696	6 676
6 795	6 822	6 904	6 742	6 728
6 827	6 963	6 835	6 787	6 780
6 876	6 996	6 844	6 832	6 833
6 902	6 914	7 051	6 877	6 885
6 929	7 067	6 926	6 922	6 937
6 936	6 931	7 122	6 966	6 989
6 949	7 117	6 948	7 010	7 041

表 3 给出了 4 种单项预测 F1、F2、F3、F4 的对数均方误差和相应的分解指标; 以及不同单项方法进行加权几何平均组合时所产生的不同对数均方误差和相应的组分来源, 进一步验证了理论分析的结果。

表 4 给出了单项预测方法的两两之间的差异性测度的指标值, 可见 F1 与 F2 之间的差异性测度最大, 而 F3 与 F4 之间的差异性测度较小。

从图 1 中可以很直观地可以看出, 组合预测方法的对数均方误差小于组合中的单项方法的对数均方误差 (例如: $L_{12}<L_1$, $L_{12}<L_2$), 说明组合预测降低了预测的对数均方误差。在组合预测中, 对于 F123 可划分为 F12-3、F13-2、F23-1 进行比较 (即在 2 种单项预测方法组合的基础上增加一个方法), 有

表 2　对数误差

Table 2　Logarithmic error

e_1	e_2	e_3	e_4
−0.009 0	−0.000 3	−0.007 5	0.003 5
−0.008 5	0.000 6	−0.007 4	0.002 2
−0.000 2	−0.013 3	−0.007 8	0.001 1
−0.000 8	0.017 3	−0.008 7	−0.000 8
−0.030 3	0.000 3	−0.007 9	−0.000 9
−0.022 3	0.000 3	−0.006 6	−0.000 5
−0.021 5	0.000 5	−0.001 8	0.003 3
−0.001 0	−0.017 2	0.003 8	0.007 8
−0.017 2	0.002 4	0.006 7	0.009 7
−0.004	−0.015 9	0.007 8	0.009 9
−0.019 7	−0.001 2	0.005 9	0.006 9
−0.017 3	0.004 7	0.006 4	0.006 3
−0.001 7	−0.021 4	0.003 6	0.002 5
−0.019 7	0.000 4	0.001 0	−0.001 2
0.000 7	−0.026 5	−0.004 3	−0.007 6
−0.023 9	0.000 1	−0.008 7	−0.013 2

表 3　对数均方误差及其分解指标

Table 3　Log-mean-square error and its decomposition index

方法	L	B	R	D
F1	$24.983\,8\times10^{-5}$	$15.075\,5\times10^{-5}$	$0.115\,7\times10^{-5}$	$9.792\,6\times10^{-5}$
F2	$13.826\,1\times10^{-5}$	$4.201\,4\times10^{-5}$	$0.228\,4\times10^{-5}$	$9.396\,4\times10^{-5}$
F3	$4.130\,3\times10^{-5}$	$0.253\,6\times10^{-5}$	$1.428\,4\times10^{-5}$	$2.448\,3\times10^{-5}$
F4	$3.812\,1\times10^{-5}$	$0.331\,9\times10^{-5}$	$0.063\,0\times10^{-5}$	$3.417\,2\times10^{-5}$
F12	$8.831\,1\times10^{-5}$	$7.317\,6\times10^{-5}$	$0.183\,7\times10^{-5}$	$1.329\,7\times10^{-5}$
F13	$3.991\,1\times10^{-5}$	$0.588\,6\times10^{-5}$	$1.156\,4\times10^{-5}$	$2.246\,0\times10^{-5}$
F14	$2.967\,3\times10^{-5}$	$0.025\,0\times10^{-5}$	$0.070\,6\times10^{-5}$	$2.871\,7\times10^{-5}$
F23	$3.461\,1\times10^{-5}$	$0.670\,3\times10^{-5}$	$0.729\,5\times10^{-5}$	$2.061\,2\times10^{-5}$
F24	$2.766\,4\times10^{-5}$	$0.001\,9\times10^{-5}$	$0.092\,8\times10^{-5}$	$2.671\,7\times10^{-5}$
F34	$3.128\,4\times10^{-5}$	$0.007\,6\times10^{-5}$	$0.163\,3\times10^{-5}$	$2.957\,5\times10^{-5}$
F123	$3.346\,9\times10^{-5}$	$1.097\,8\times10^{-5}$	$0.568\,7\times10^{-5}$	$1.680\,4\times10^{-5}$
F124	$2.274\,0\times10^{-5}$	$0.268\,7\times10^{-5}$	$0.094\,5\times10^{-5}$	$1.910\,8\times10^{-5}$
F134	$2.788\,0\times10^{-5}$	$0.063\,3\times10^{-5}$	$0.016\,4\times10^{-5}$	$2.708\,4\times10^{-5}$
F234	$2.536\,9\times10^{-5}$	$0.054\,3\times10^{-5}$	$0.013\,8\times10^{-5}$	$2.468\,8\times10^{-5}$
F1234	$2.249\,2\times10^{-5}$	$0.286\,2\times10^{-5}$	$0.015\,1\times10^{-5}$	$1.941\,6\times10^{-5}$

表 4　差异性测度的指标值

Table 4　The index value for measures of variability

差异测度	C
F1,F2	$3.911\,30\times10^{-4}$
F1,F3	$2.455\,35\times10^{-4}$
F1,F4	$3.148\,69\times10^{-4}$
F2,F3	$1.630\,16\times10^{-4}$
F2,F4	$1.890\,72\times10^{-4}$
F3,F4	$0.334\,08\times10^{-4}$

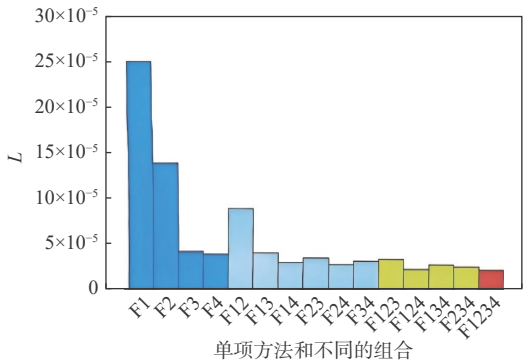


图1 对数均方误差 L

Fig. 1 Log-mean-square error L

$L_{123}<L_{12}$、$L_{123}<L_{13}$、$L_{123}<L_{23}$，同理3种单项方法的组合也是类似的结果；并且所有组合预测方法中 $L_{1\,234}$ 是最小的；说明随着方法多样性的增加，预测的均方误差是在减小的。

从图2中可以发现，单项方法中F1和F2有较高的 B，在组合预测方法中F12有较高的 B，并且这3种方法的 B 都远高于其他方法。并且发现含有F1和F2的组合预测方法的对数均值偏差相对于其他组合预测方法是偏大的，再结合图1可以发现相对应的对数均方误差也是高于其他方法的。因此，具有较高的 B 可能会带来较大的对数均方误差，即预测值对数序列与实际值对数序列的中心差异性较大时会带来较大的对数均方误差。

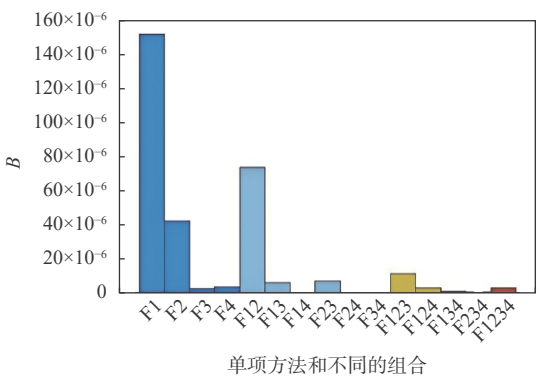


图2 对数均值偏差 B

Fig. 2 Log-mean bias B

从图3可以明显地看出，组合预测的 R 小于所组合的单项方法的 R。结合图1，对于单项方法，F3有着最高 R，却有着比F1和F2更低的 L；对于2种单项方法的组合中F13有着最高的 R，却有着比F12更小的 L，同样在3种单项方法中也有着类似的情况。因此 R 在减小对数均方误差的作用方面并没有起到显著的作用。

从图4中可以很直观地看出，组合预测方法的 D 小于组合中的单项方法的 D，说明组合预测降低了预测的对数离差误差。并且该指标在 L 中占比

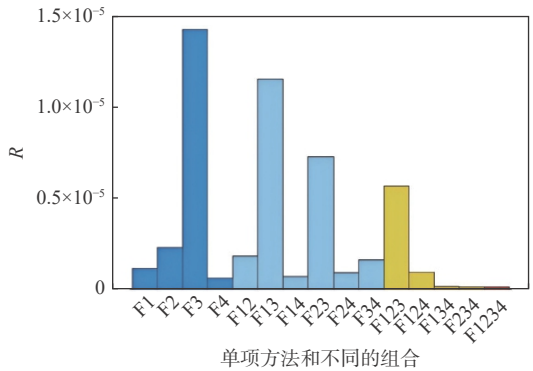


图3 因模型解释力不足而产生的对数方差 R

Fig. 3 Logarithmic variance due to insufficient explanatory power of the model R

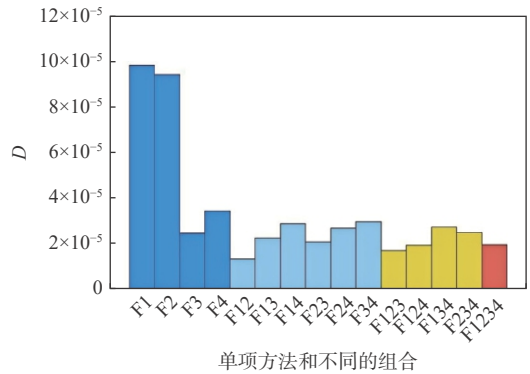


图4 基于解释力系数的对数离差误差 D

Fig. 4 Logarithmic deviation error based on explanatory power coefficient D

是最大的，因此对组合预测的对数均方误差起着主导作用，可以通过降低 D 更有效地降低组合预测的对数误差，即降低预测对实际值不可解释的误差。

6 结 论

本文针对非线性的加权几何平均组合预测模型，从两个方面探讨了对数均方误差的来源。

1）通过引入相关概念，对加权几何平均组合预测模型产生的对数均方误差进行分解，得到对数均方误差来自3个方面。

2）从组合预测单项方法的遴选角度出发，将组合预测对数均方误差分解为单项方法均方误差的加权和差异性指标。

3）通过案例并结合对数均方误差的分解表达式可知，通过增加组合中单项方法的差异性，提高各单项方法的精度这2个途径，可以有效地降低组合预测对数误差。

本文基于加权几何平均进行的研究，而组合预测还有其他的形式，以及基于各种不同指标建立的组合模型；未来可进一步探讨其他非线性的加权调和平均组合预测模型的误差来源解析。

参考文献（References）

[1] BATES J M, GRANGER C W J. The combination of forecasts[J]. Journal of the Operational Research Society, 1969, 20(4): 451-468.

[2] ATIYA A F. Why does forecast combination work so well?[J]. International Journal of Forecasting, 2020, 36(1): 197-200.

[3] SEGURA-HERAS J V, BERMÚDEZ J D, CORBERÁN-VALLET A, et al. Analysis of weighting strategies for improving the accuracy of combined forecasts[J]. Mathematics, 2022, 10(5): 725.

[4] 唐小我, 马永开, 曾勇, 等. 现代组合预测和组合投资决策方法及应用[M]. 北京: 科学出版社, 2003.
TANG X W, MA Y K, ZENG Y, et al. Modern portfolio forecasting and portfolio investment decision-making method and its application[M]. Beijing: Science Press, 2003(in Chinese).

[5] 陈华友. 组合预测方法有效性理论及其应用[M]. 北京: 科学出版社, 2008.
CHEN H Y. Validity theory of combined forecasting method and its application [M]. Beijing: Science Press, 2008(in Chinese).

[6] 汪同三, 张涛. 组合预测: 理论、方法及应用[M]. 北京: 社会科学文献出版社, 2008.
WANG T S, ZHANG T. Combinatorial forecasting: Theory, method and application[M]. Beijing: Social Sciences Academic Press, 2008(in Chinese).

[7] ELLIOTT G, GRANGER C W J, TIMMERMANN A. Handbook of economic forecasting (volume1) [M]. North-Holland: Elsevier, 2006.

[8] ARMSTRONG J. Principles of forecasting[M]. Boston: Kluwer Academic, 2001.

[9] BUNN D W. Forecasting with more than one model[J]. Journal of Forecasting, 1989, 8(3): 161-166.

[10] LV W D, QI J P. Stock market return predictability: a combination forecast perspective[J]. International Review of Financial Analysis, 2022, 84: 102376.

[11] FAMELITI S P, SKINTZI V D. Statistical and economic performance of combination methods for forecasting crude oil price volatility[J]. Applied Economics, 2022, 54(26): 3031-3054.

[12] 林虹, 张永, 杨兴雨, 等. 考虑组合预测股价的泛证券投资组合选择策略[J]. 管理工程学报, 2023, 37(5): 130-141.
LIN H, ZHANG Y, YANG X Y, et al. Universal portfolio selection strategy based on combination forecasting price[J]. Journal of Industrial Engineering and Engineering Management, 2023, 37(5): 130-141(in Chinese).

[13] 徐小峰, 余乐安, 林姿汝, 等. 基于特征融合的生鲜商品短期销量组合预测[J]. 管理科学学报, 2022, 25(12): 102-123.
XU X F, YU A L, LIN Z R, et al. Combination forecasting of short-term sales for fresh products based on feature fusion[J]. Journal of Management Sciences in China, 2022, 25(12): 102-123(in Chinese).

[14] 刘金培, 张了丹, 朱家明, 等. 非结构性数据驱动的混合分解集成碳交易价格组合预测[J]. 运筹与管理, 2023, 32(3): 149-154.
LIU J P, ZHANG L D, ZHU J M, et al. Unstructured data driven carbon price combined forecast based on hybrid decomposition-integration[J]. Operations Research and Management Science, 2023, 32(3): 149-154(in Chinese).

[15] 陈华友. 基于相关系数的优性组合预测模型研究[J]. 系统工程学报, 2006, 21(4): 353-360.
CHEN H. Research on properties of superior combined forecasting models based on correlation coefficients[J]. Journal of Systems Engineering, 2006, 21(4): 353-360(in Chinese).

[16] 陈华友. 基于 Theil 不等系数的组合预测模型的性质[J]. 电子科技大学学报, 2004, 33(1): 105-108.
CHEN H Y. Properties of combination forecasting models based on theil coefficient[J]. Journal of University of Electronic Science and Technology of China, 2004, 33(1): 105-108(in Chinese).

[17] 陈华友, 盛昭瀚, 刘春林. 基于向量夹角余弦的组合预测模型的性质研究[J]. 管理科学学报, 2006, 9(2): 1-8.
CHEN H Y, SHENG Z H, LIU C L. Research on properties of combination forecasting model based on vectorial angle cosine[J]. Journal of Management Sciences in China, 2006, 9(2): 1-8(in Chinese).

[18] 陈华友. 基于预测有效度的组合预测模型研究[J]. 预测, 2001, 20(3): 72-73,68.
CHEN H Y. Research on combination forecasting model based on effective measure of forecasting methods[J]. Forecasting, 2001, 20(3): 72-73,68 (in Chinese).

[19] 陈华友, 刘春林. 基于 IOWA 算子的组合预测方法[J]. 预测, 2003, 22(6): 61-65.
CHEN H Y, LIU C L. A kind of combination forecasting method based on induced ordered weighted averaging (IOWA) operators[J]. Forecasting, 2003, 22(6): 61-65(in Chinese).

[20] 陈华友, 刘春林, 盛昭瀚. IOWHA 算子及其在组合预测中的应用[J]. 中国管理科学, 2004, 12(5): 35-40.
CHEN H Y, LIU C L, SHENG Z H. Induced ordered weighted harmonic averaging (IOWHA) operator and its application to combination forecasting method[J]. Chinese Journal of Management Science, 2004, 12(5): 35-40(in Chinese).

[21] 陈华友, 盛昭瀚. 一类基于 IOWGA 算子的组合预测新方法[J]. 管理工程学报, 2005, 19(4): 36-39.
CHEN H Y, SHENG Z H. A kind of new combination forecasting method based on induced ordered weighted geometric averaging (IOWGA) operator[J]. Journal of Industrial Engineering and Engineering Management, 2005, 19(4): 36-39(in Chinese).

[22] 陈勤勤, 陈华友, 韩冰. 基于灰色关联度和 Shapley 值的区间组合预测模型及其应用[J]. 安徽大学学报 (自然科学版), 2023, 47(4): 16-24.
CHEN Q Q, CHEN H Y, HAN B. Interval combination forecasting model based on grey correlation degree and Shapley value and its application[J]. Journal of Anhui University (Natural Science Edition), 2023, 47(4): 16-24(in Chinese).

[23] 张逸飞, 成晟, 韩一杰, 等. 整体与极端偏差之权衡: 一种改进的组合预测优化方法[J]. 系统工程理论与实践, 2023, 43(6): 1837-1851.
ZHANG Y F, CHENG S, HAN Y J, et al. Trade-off between the global and extreme bias: an improved optimization approach in forecast combination[J]. Systems Engineering-Theory & Practice, 2023, 43(6): 1837-1851(in Chinese).

[24] HU Y H, GUO Y S, FU R. A novel wind speed forecasting combined model using variational mode decomposition, sparse auto-encoder and optimized fuzzy cognitive mapping network[J]. Energy, 2023, 278: 127926.

[25] ZENG L L, LING L W, ZHANG D B, et al. Optimal forecast combination based on PSO-CS approach for daily agricultural future prices forecasting[J]. Applied Soft Computing, 2023, 132: 109833.

[26] 聂铃, 张剑, 胡茂政. 基于 CEEMDAN 分解的短时交通流组合预测[J]. 计算机工程与应用, 2022, 58(11): 279-286.
NIE L, ZHANG J, HU M Z. Short-term traffic flow combination prediction based on CEEMDAN decomposition[J]. Computer Engineering and Applications, 2022, 58(11): 279-286(in Chinese).

[27] CHEN W, XU H L, CHEN Z S, et al. A novel method for time series prediction based on error decomposition and nonlinear combination of forecasters[J]. Neurocomputing, 2021, 426(22): 85-103.

[28] THOMSON M E, POLLOCK A C, öNKAL D, et al. Combining forecasts: performance and coherence[J]. International Journal of Forecasting, 2019, 35(2): 474-484.

[29] 张倩昀, 陈华友, 江立辉. 加权算术平均组合预测方法均方误差的分解[J]. 合肥学院学报 (综合版), 2020, 37(2): 7-13.
ZHANG Q Y, CHEN H Y, JIANG L H. Decomposition of mean square error of weighted arithmetic mean combined forecasting method[J]. Journal of Hefei University(Comprehensive Ed), 2020, 37(2): 7-13(in Chinese).

[30] KANG Y F, CAO W, PETROPOULOS F, et al. Forecast with forecasts: Diversity matters[J]. European Journal of Operational Research, 2022, 301(1): 180-190.

[31] 周传世, 罗国民. 加权几何平均组合预测模型及其应用[J]. 数理统计与管理, 1995, 14(3): 17-19.
ZHOU C S, LUO G M. Weight geometric average combination forecast model and its applications[J]. Journal of Applied Statistics and Management, 1995, 14(3): 17-19(in Chinese).

[32] 杨桂元, 唐小我, 马永开. 最优加权几何平均组合预测方法研究[J]. 统计研究, 1996, 13(2): 55-58.
YNAG G Y, TANG X W, MA Y K. Research on optimal weighted geometric average combined forecasting method[J]. Statistical Research, 1996, 13(2): 55-58(in Chinese).

[33] 王自成, 陈华友, 韩冰, 等. 基于 TOPSIS 的最大贴近度的最优组合预测模型及其性质[J]. 中国科学技术大学学报, 2019, 49(9): 751-761.
WANG Z C, CHEN H Y, HAN B, et al. An optimal combination forecasting model based on maximum closeness of TOPSIS and its properties[J]. Journal of University of Science and Technology of China, 2019, 49(9): 751-761(in Chinese).

Decomposition of logarithm mean square error of weighted geometric mean combined forecasting method

ZHU Shuanglei[1], CHEN Huayou[1, 2, *], ZHANG Qi[2], WANG Yifan[3]

(1. School of Big Data and Statistics, Anhui University, Hefei 230601, China;
2. School of Mathematical Sciences, Anhui University, Hefei 230601, China;
3. Stony Brook Institute at Anhui University, Hefei 230039, China)

Abstract: In addition to proposing the ideas of logarithm mean deviation, logarithm explanatory power coefficient, logarithm deviation error variance, and diversity measure of prediction methods, the nonlinear weighted geometric mean combination prediction model was introduced to the logarithm mean square error. On this basis, the logarithm mean square error of the combined prediction is decomposed into three components: the logarithm mean error, the logarithm variance due to the insufficient explanatory power of the model and the logarithm deviation variance based on the logarithm explanatory power coefficient. The source of the logarithm mean square error is discussed theoretically. Simultaneously, the logarithm mean square error of combination forecasting is broken down into the weighted average of the logarithm mean square error of single forecasting methods and the weighted average of diversity measure of forecasting methods. This yields the helpful conclusion that increasing the accuracy of single forecasting methods and the diversity measure level of forecasting methods helps to lower the logarithm mean square error of weighted geometric mean combined forecasting methods. A case study analyzed the influence of the diversity measure of each component and single forecasting method on the logarithm mean square error of combined forecasting.

Keywords: combination forecasting; weighted geometric mean; logarithm mean squared error; decomposition; diversity measure

Received: 2022-08-29; **Accepted**: 2023-10-29; **Published Online**: 2023-11-24 15:24

URL: link.cnki.net/urlid/11.2625.V.20231123.1407.003

Foundation items: National Natural Science Foundation of China (72371001); Anhui University Innovation Training Program

* **Corresponding author**. E-mail: huayouc@126.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0562

耗氧型催化惰化系统整体式反应器性能研究

郭晨阳[1]，刘祎[2]，刘豪正[1]，王俊杰[1]，高经诚[1]，冯诗愚[1,*]

(1. 南京航空航天大学 航空学院 飞行器环境与生命保障工信部重点实验室，南京 210016；
2. 中航工业集团有限公司金城南京机电液压工程研究中心 航空机电系统综合航空科技重点实验室，南京 211106)

摘　　要：对 $Pd/\gamma\text{-}Al_2O_3$ 催化剂进行了实验测量，拟合得到了其反应动力学方程，建立了整体式反应器二维瞬态模型，并采用 Modelica 语言编程求解，开展动态性能仿真，分析了耗氧型惰化系统中飞温和催化剂高温失活现象对整体式反应器性能的影响,研究了飞行过程中氧气体积分数变化对反应器的影响。研究结果表明：壁面冷却可有效降低飞温现象，但同时会降低反应器催化效率。飞温现象导致的催化剂涂层不可逆失活发生在反应最剧烈的反应器出口侧壁面上，小面积的失活会导致催化效率降低 80%。飞机爬升阶段油箱内的氧浓度增加可有效提升反应器内的燃油蒸汽转化率，但也需要避免高氧浓度对反应器涂层加速失活的影响。

关 键 词：催化燃烧；反应器；燃油系统；惰化系统；催化剂失活
中图分类号：V228
文献标志码：A　　**文章编号**：1001-5965(2025)10-3547-08

耗氧型燃油惰化系统的功能是将油箱上部气相空间的燃油蒸汽和空气抽吸进入催化反应器，使其发生无焰燃烧，然后将氧浓度降低后的混合气体重新注入到油箱内部，从而保证油箱氧浓度在安全范围内。催化反应器是耗氧型燃油惰化系统的核心部件，其性能直接影响到惰化效果[1-2]。

常见的反应器有固定床层式和整体式等多种。这些反应器在化工领域已经有了较广泛的研究和应用，如汽车尾气处理、加氢脱硫等强吸/放热反应、脂交换反应等[3]，但机载条件对整体式反应器又提出了不同的要求[4-5]。化工领域中操作条件稳定，机载条件下，起飞、爬升、巡航和降落等飞行过程中，油箱上部空间内燃油蒸汽浓度、温度和压力差异大，要求机载催化反应器操作工况范围宽，且在兼顾较小重量和尺寸的前提下，具备较强的动态调节能力，动态调节能力中最关键的一点是对燃油蒸汽浓度或氧气浓度突然升高导致升温的耐受能力[6]。

固定床层式反应器在内部填充有催化剂颗粒，气体流动复杂，热量难以导出，易出现局部温度升高的飞温现象，导致催化剂烧结，比表面积降低，活性下降，降低反应器催化效率，减少使用寿命[7]。整体式反应器内部为金属基或陶瓷基微小通道，表面附着催化剂，相比床层式反应器，整体式反应器流道简单，催化剂面积大，催化效率高[8-10]，更加适合机载耗氧型惰化系统。

目前在工业领域对整体式反应器的研究相对充分，例如在反应器仿真方面，Wanker 等[11] 研究了单通道内气相空间和催化剂涂层之间传热与传质的关系，Depcik 等[12] 对涂层物理性质和涂层内气体扩散对反应器的影响进行了研究，Guo 等[13] 研究了反应器通道几何形状和壁厚对催化效率的影响。但是，飞机在飞行条件下造成的温度压力变化导致耗氧型惰化反应器工况复杂，且催化的是燃油蒸

收稿日期：2023-09-01；**录用日期**：2024-03-29；**网络出版时间**：2024-08-15 09:34
网络出版地址：link.cnki.net/urlid/11.2625.V.20240814.1628.002
基金项目：国家自然科学基金中国民用航空民航联合研究基金 (U1933121)；国家自然科学基金 (52372331)；中央高校基本科研业务费专项基金；江苏省高校优势学科建设工程；南京航空航天大学研究生科研与实践创新计划项目 (xcxjh20220102)
* **通信作者**. E-mail：shiyuf@nuaa.edu.cn

引用格式：郭晨阳，刘祎，刘豪正，等. 耗氧型催化惰化系统整体式反应器性能研究 [J]. 北京航空航天大学学报，2025，51（10）：3547-3554. GUO C Y，LIU Y，LIU H Z，et al. Research on oxygen consumption based inerting monolithic catalyst reactor performance [J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51（10）：3547-3554 （in Chinese）.

汽，其成分复杂[14-15]，现有的研究不能直接使用。

有鉴于此，首先，通过实验测量了耗氧型惰化反应器中常用的 Pd/γ-Al_2O_3 催化剂的催化特性和高温失活特性，拟合出其反应动力学方程，然后，建立机载整体式反应器二维仿真模型，通过 Modelica 语言进行编程求解，开展反应器动态性能仿真。得到不同工况下整体式反应器的转化率、温度分布、高温条件下的失活特性，以及真实飞行状态下反应器的催化效率和失活特性。

1 反应动力学方程测定

1.1 实验装置

反应动力学实验测试装置如图 1 所示。氮气和氧气按照不同比例混合后进入油箱上部空间，与其中的燃油蒸汽充分混合，流经预热器，加热至工况所设定的温度，然后进入马弗炉内，来流中的燃油蒸汽在 Pd/γ-Al_2O_3 催化剂作用下发生无焰燃烧，反应产物进入尾气处理装置中进行处理。在马弗炉前后各有一个取样口，用针管对待测气体进行取样，注入到气相色谱仪中进行成分分析，得到反应前后的气体组分，进行对比即可得出该催化剂在指定工况下的催化反应速率。

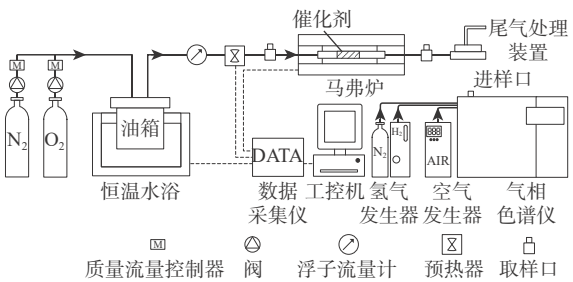


图 1 反应动力学实验台示意

Fig. 1 Reaction dynamics experiment diagram

燃油浓度通过调节恒温水浴的温度控制，反应温度通过调节预热器和马弗炉温度控制，马弗炉带有隔热装置确保催化剂保持在要求温度内。试验用设备清单及参数见表 1。

表 1 实验台使用的设备型号及性能参数

Table 1 Equipment model and performance parameters in the experiment used in the laboratory bench

设备	型号	工作范围	精度范围
浮子流量计		6～60 mL/min	±4%
注射泵	SP-1000	0.1～999.9 mL	±3%
马弗炉	SK2-2-13单管	功率2 kW	
气相色谱仪	GC9790	总烃	±0.1%
反应器	非标定制		
温控表	HKDN	0～5 V	±0.5%F.S.
数据采集仪	LR-8400-21	DC±100 V	500 mV

1.2 实验方法

采用指数型的阿伦尼乌斯公式作为反应动力学方程，如下所示：

$$r = k_0 e^{-\frac{E}{RT}} C_{CH}^m C_{O_2}^n \tag{1}$$

式中：r 为反应速率，mol/(m^3·s)；k_0 为指前因子，与温度无关，m^3/(mol·s)；E 为反应活化能，J/mol；R 为摩尔气体常数，J/(mol·K)；C_{CH} 和 CO_2 分别为燃油蒸汽浓度和氧气浓度，mol/m^3；m、n 为燃油蒸汽、氧气反应级数。

本文实验测得的催化剂颗粒反应速率 r，mol/(m^3·s)，通过式(2)转化为催化剂涂层反应速率 r_1，mol/(m^3·s)。

$$r_1 = \frac{S}{S_1} r(1-\varepsilon)\rho \tag{2}$$

式中：S 为实验用颗粒催化剂的比表面积，m^2/g；S_1 为催化剂涂层的比表面积，m^2/g；r 为测得的颗粒催化剂反应速率，mol/(m^3·s)；ε 为实验用颗粒催化剂的孔隙率；ρ 为催化剂颗粒密度，g/m^3。

通过实验测得 Pd/γ-Al_2O_3 催化剂在不同反应温度、燃油蒸汽浓度和氧气浓度下的反应速率，进行数据拟合，即可得出反应动力学方程中的 4 个未知参数。

实验采用催化剂体积为 5 mL，气体流量为 200 mL/min，反应在常压下进行。反应温度设置 5 组，150 ℃ 起每隔 25 ℃ 为一组。油箱恒温水浴的温度设置 5 组，20 ℃ 起每隔 10 ℃ 设置一组。来流氧气体积分数设置 6 组，0.02、0.05、0.1、0.2、0.5、0.8。实验共进行 150 组。

高温失活实验在 300～800 ℃ 每隔 100 ℃ 设置一个实验点，共进行 6 组，到达指定温度后测量一次碳氢物转化率，温度稳定 10 min 后测量一次，后降温至 200 ℃ 再测量一次反应速率，以判断是否为永久性失活。

1.3 实验结果

得到的部分实验结果如图 2～图 4 所示，图 2 是反应温度与催化剂反应速率的关系。

图 2 反应温度上升时，反应速率增速先大后小，高温时逐渐平缓。将燃油箱温度从 20 ℃ 提升至 60 ℃ 时，反应器进口处燃油蒸汽浓度从 2.07×10^{-5} mol/m^3 增长至 7.05×10^{-5} mol/m^3，增长了 250%，同时反应速率也增长了 300%。

高温失活实验结果如图 3 所示，反应温度为 600 ℃ 时转化率最高，为 59.24%，温度升至 700 ℃ 时转化率开始下降，至 800 ℃ 时大幅降低，且温度降低后未见转化率回升，因此可认为本文实验使用的 Pd/γ-Al_2O_3 催化剂的最佳反应温度为 600 ℃，大于 600 ℃ 时开始发生烧结现象，活性表面积降低，催化性能开始下降，此时发生的是小面积烧结，温

度降低后转化率回升, 800 ℃ 时催化剂大面积失活, 严重影响工作效率, 与郑婷婷等[16]的研究结论相符。

图 4 为氧气体积分数对反应速率的影响曲线。氧气体积分数小于 0.1 时, 反应速率增加迅速, 大于 0.1 时, 进入富氧燃烧阶段, 反应速率对氧气体积分数敏感度降低, 故采用低氧气体积分数 (<0.1) 和高氧气体积分数 (≥0.1) 两段对反应动力学方程进行拟合。

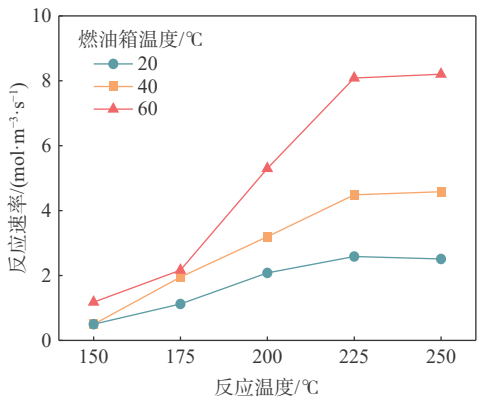


图 2　反应温度与反应速率关系

Fig. 2　Curves of reaction rate and reaction temperature under different tank temperature

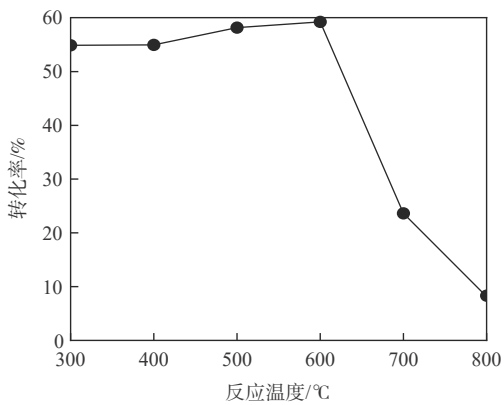


图 3　高温条件下反应温度与转化率关系

Fig. 3　Curves of reaction temperature and convention rate under high temperature

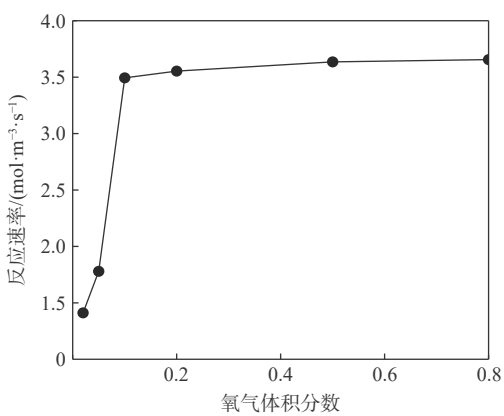


图 4　反应速率随氧气体积分数变化

Fig. 4　Curve of reaction rate and oxygen volume fraction

将实验得到的 150 组数据代入式(1)和式(2)中, 拟合结果如表 2 所示。

表 2　参数拟合结果

Table 2　Parameter fitting result

氧气体积分数	指前因子k_0/$(m^3(mol·s)^{-1})$	活化能E/$(J·mol)^{-1}$	m	n
<0.1(低)	13 370.72	23 422.56	0.319	0.659
≥0.1(高)	25 027.53	28 285.78	0.798	0

图 5 为拟合结果与实验对比, 可以看出该拟合误差在±20% 以内, 在低反应温度区间内 (≤175 ℃) 预测结果最优, 误差可缩小至 5%～10%。

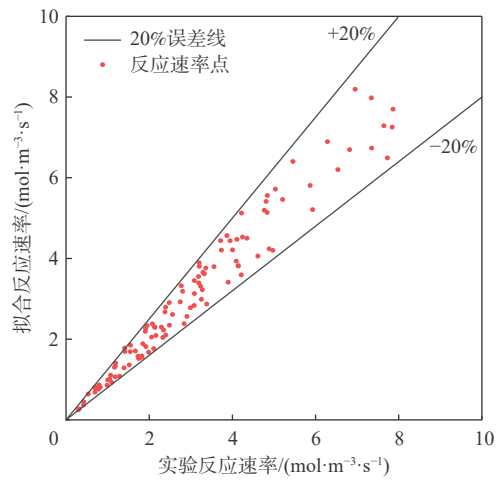


图 5　拟合反应速率与实验反应速率对比

Fig. 5　Comparison between fitting reaction rate and experiment reaction rate

2　反应器数值模拟

2.1　物理模型与控制方程

该仿真的物理模型如图 6 所示, 整体式反应器主结构材质为耐热铁铬合金, 内部管壁涂有 Pd/γ-Al_2O_3 催化剂涂层。为了简化模型, 忽略催化剂涂层厚度, 认为反应在反应器通道壁面上发生。

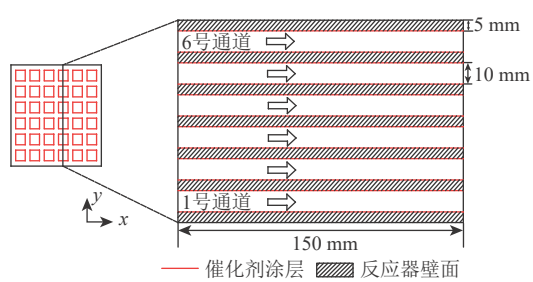


图 6　反应器物理模型

Fig. 6　Reactor physics model

该反应器的物理参数如表 3 所示。

该模型采用的控制方程如下。

连续性方程:

表 3 反应器物理参数

Table 3 Reactor physics parameter

物理参数	数值
管壁密度/(kg·m^{-3})	2 100
管壁比热容/(J·kg^{-1}·K^{-1})	520
管壁导热系数/(W·m^{-1}·K^{-1})	17.6

$$\frac{\partial \rho_g}{\partial t} = -\nabla(\rho_g \boldsymbol{v}) + S \tag{3}$$

式中：$\boldsymbol{v}$ 为气体速度，m/s；ρ_g 为气体密度，g/m^3；S 为化学反应源项。

动量方程：

$$\rho \frac{\partial \boldsymbol{v}}{\partial t} + \rho(\boldsymbol{v}\nabla)\boldsymbol{v} = -\nabla p + \eta \nabla^2 \boldsymbol{v} \tag{4}$$

式中：p 为气体压力，Pa；η 为气体动力黏度，Pa·s。

气相能量方程：

$$\frac{\partial}{\partial t}\left(c_p \rho_g T_g\right) + \nabla\left[\boldsymbol{v}\left(\rho_g E + p\right)\right] = \nabla\left[k_f \nabla T_g - h_i \boldsymbol{J}_i + \tau \boldsymbol{v}\right] + h_v (T_g - T_s) \tag{5}$$

式中：c_p 为气体定压比热容，J/(g·K)；ρ_g 为气体密度，g/m^3；T_g 为气体温度，K；E 为气体能量，J；p 为气体压力，Pa；k_f 为气体导热系数，W/(m·K)；h_i 为气体组分 i 的焓变，J；$\boldsymbol{J}_i$ 为气体组分 i 的质量扩散通量，g/(m^2·s)；h_v 为气体与壁面的对流换热系数，W/(m^2·K)；τ 为剪切应力，Pa；T_s 为壁面温度，K。

固相能量方程：

$$\frac{\partial}{\partial t}(\rho_s c_s T_s) = k_s \nabla^2 T_s + h_v (T_g - T_s) + H_r \tag{6}$$

式中：ρ_s 为管壁密度，g/m^3；c_s 为壁面材料比热容，J/(g·K)；k_s 为管壁导热系数，W/(m·K)；H_r 为反应热，J。

气相状态方程：

$$p = \rho_g R T_g \sum_i \frac{a_i}{M_i} \tag{7}$$

式中：a_i 为第 i 组分气体质量分数；M_i 为第 i 组分气体摩尔质量。

求解过程中，对以上控制方程进行离散。时间项采用隐式格式，固相能量方程中的扩散项使用中心差分格式，气相连续性方程和动量方程的对流项使用一阶迎风格式，扩散项使用中心差分格式，收敛标准为 10^{-5}。对模型进行空间离散时使用均匀结构化正交网格。

使用控制方程中的源项对化学反应进行模拟，使用固相能量方程中的能量源项模拟化学反应热，认为反应热全部传递给固相壁面后由对流换热传递给气相，使用连续性方程中的源项模拟化学反应消耗或产生的物质。

该仿真共进行 4 组，第 1 组为绝热边界条件动态仿真，第 2 组为强制对流边界条件动态仿真，第 3 组为催化剂失活动态仿真，第 4 组为飞机爬升时进口燃油蒸汽浓度增大条件下的失活仿真。

第 1、3、4 组采用绝热边界条件，第 2 组采用流速 0.1 m/s、温度 20 ℃ 的空气强制对流边界条件。第 3、4 组中，根据实验得出的结论以及郑婷婷等[16]的研究，将催化剂失活温度设置为 800 ℃，认为当催化剂微元温度达到 800 ℃ 时不再具有催化能力。

第 1、2、3 组中计算的进口条件为温度 140 ℃，速度 0.1 m/s，燃油蒸汽体积分数 0.004 5，氧气体积分数 0.23，氮气体积分数 0.765 5。

飞机油箱内不进行惰化处理时，随着海拔的增加，油箱内压力降低，溶解在燃油内的氧气溢出，气相空间氧气体积分数逐渐增大[17]，为了模拟这一过程以得出第 4 组中进口氧气体积分数变化，需要计算飞机油箱内气相空间氧气体积分数变化情况，如下式[2]所示。

$$\Delta c = \frac{\Delta(\beta_0 V_F)}{V_A} \tag{8}$$

$$\beta_O = \frac{101\,325 - p_v}{101\,325} \cdot \frac{2.31(980-\rho)}{1\,000} e^{\frac{0.639(700-T_F)}{T_F}\ln(3.333\beta_{O0})} \tag{9}$$

$$V_F = m_F/\rho \tag{10}$$

中：Δc 为气相空间氧浓度变化量，kg/(m^3·s)；β_O 为氧气阿斯特瓦尔德系数，表示单位体积燃油内溶解的氧气体积；V_F 为燃油体积，m^3；V_A 为气相空间体积，m^3；p_v 为燃油饱和蒸汽压，Pa；ρ 为燃油密度，g/m^3；T_F 为燃油温度，K；m_F 为燃油质量，g。

将上述公式与标准大气温度、压力公式联立并求解后得到飞机在 200 s 内从海平面 (1 013 hPa, 293 K) 匀速上升至海拔 4 000 m(616 hPa, 263 K) 时，油箱上部气相空间内氧气体积分数变化，如图 7 所示，第 4 组仿真中采用该曲线作为进口氧气体积分

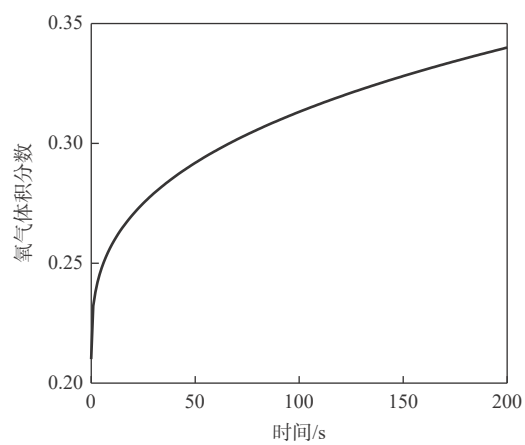


图 7 氧气浓度变化曲线

Fig. 7 Oxygen concentration change curve

数进行仿真。燃油蒸汽体积分数为 0.045。

2.2　网格无关性验证

采用 7 种网格布置分别进行仿真, 计算所有管道的平均燃油蒸汽转化率 σ, 结果如图 8 所示。

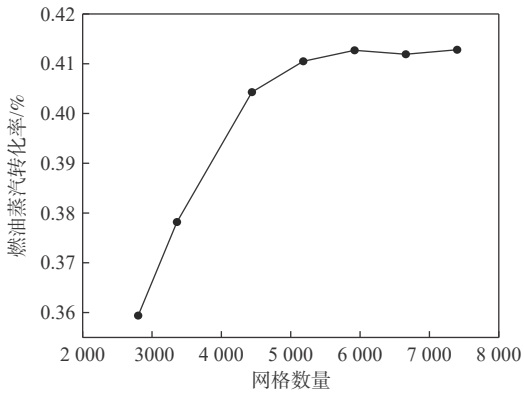


图 8　网格无关性验证

Fig. 8　Grid independence verification

网格数量超过 6 000 后, 燃油蒸汽转化率较为稳定, 满足仿真计算要求, 因此 6 600 为满足网格无关性的最小网格数。

2.3　模型可靠性验证

根据黄敬敬等[18]的实验数据对本模型进行可靠性验证。设定反应器温度分别为 400 ℃、450 ℃、500 ℃、550 ℃ 进行稳态仿真, 结果如图 9 所示。数据对比显示, 本仿真可靠性较好, 模拟结果可信。

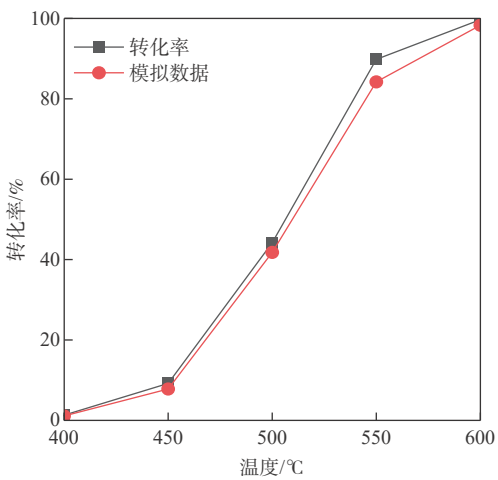


图 9　模拟数据与实验数据对比

Fig. 9　Comparison between simulation data and experimental data

3　仿真结果与分析

绝热边界条件仿真 150 s 后, 温度分布如图 10 所示。

从图 10 中可以观察到, 在反应器末端存在较为明显的温度不均匀分布, 在平行于气流方向上, 气流温度随着反应在壁面的进行逐渐升高, 高温区分布在每根管道的后端; 垂直于气流方向上, 高温区位于管壁附近, 并且在靠近反应器中心的第 3、4、5 号管壁处均有较大的温升产生, 可以观察到较为明显的局部高温现象。

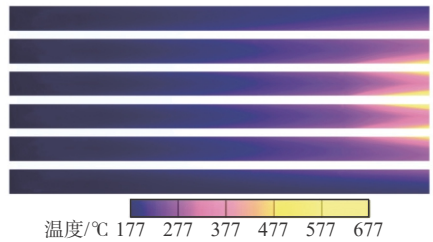


图 10　反应器温度分布

Fig. 10　Reactor temperature distribution

图 11 为距气体进口 5 mm、10 mm、15 mm 处的气体径向温度分布情况, 每段曲线代表在对应长度处该管道径向温度分布情况。在外壁面绝热的条件下, 反应器内部热量无法有效排出, 造成气流沿流动方向温度升高, 催化反应速率变大, 进一步提高反应器温度。

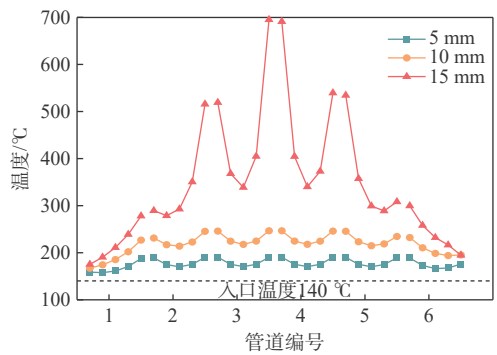


图 11　绝热边界条件下反应器径向气体温度分布

Fig. 11　Reactor temperature normal distribution in different section under adiabatic boundary condition

图 12 为不同位置的气体温度随时间变化情况, 在 130 s 后, 出口处气体温度增速突然增大, 并在 150 s 仿真结束时温度达到 700 ℃。

反应器出现温度激增的现象, 是空间上边界绝热、散热困难和时间上指数型阿伦尼乌斯方程出现

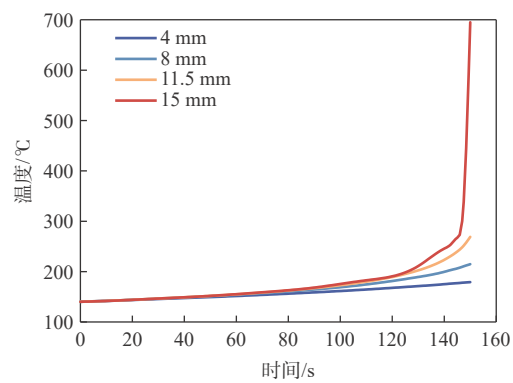


图 12　管道内气体温度随时间变化

Fig. 12　Gas temperature in pipeline change over time

温度的指数爆炸共同作用的结果。

图 13 为第 2 组强制对流边界条件反应器的温度分布情况。

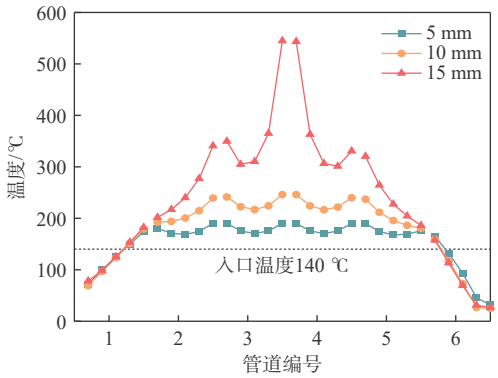


图 13 对流边界条件下反应器径向气体温度分布

Fig. 13 Reactor temperature normal distribution in the radial direction of the reactor under convection boundary condition

在外壁添加流速 0.1 m/s、温度 20 ℃ 空气后，反应器降温明显，越靠外侧的管道降温幅度越大。添加外壁冷却可以有效降低反应器的最高温度，避免因局部温度过高导致催化涂层变性失活问题的产生。但是由于反应器高温区位于中心管道，热量的传导需要经过多层气固相界面，等效热阻过大，导致降低核心温度需要大量冷量，使图 13 中原本位于 500～700 ℃ 最佳反应温度区间的涂层温度降低，最外层涂层温度最低只有 20 ℃，无法有效进行催化反应，在减少涂层失活现象发生的同时，也将平均转化率从 54% 降至 13%，严重影响了反应器的催化性能。

在第 3 组对催化剂涂层失活的仿真中，图 14 展示了涂层失活现象和时间的关系。由于催化反应释放的热量在壁面的累积，失活现象首先出现于反应器出口处，并逆气流方向向反应器内部延伸。

图 15 展示了反应器累积失活长度和燃油蒸汽转化率的关系，140 s 时开始出现失活现象，小面积

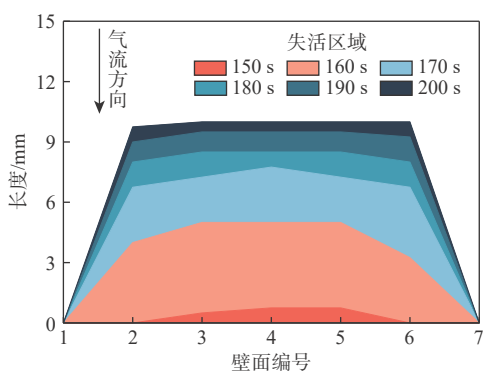


图 14 永久活区域长度随时间增大

Fig. 14 The length of the Catalyst deactivation area increase over time

的失活增大了燃油蒸汽的转化率，这表明伴随着小规模失活发生时温度处于最佳反应区间内的催化剂涂层长度较大。随着反应继续进行，温度持续升高，在小规模失活时处于最佳反应温度区间的催化剂涂层开始陆续发生失活现象，失活长度增速变大。170 s 时失活区域面积扩大至 35%，随后失活区域长度增加速度逐渐减小。失活区域长度增加速度减小的原因是未失活区域过短，无法提供足够的催化强度使来流升温。如图 15 所示，200 s 时各壁面内失活区域长度最长达到 10 mm，失活长度占总催化剂涂层长度的 50%，燃油蒸汽转化率从 51% 降低至 10%。由于后半段 50% 的催化剂涂层失活导致转化率降低了 4/5。

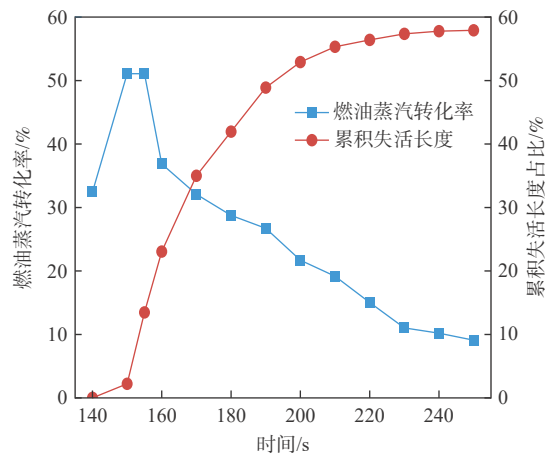


图 15 燃油蒸汽转化率和失活面积关系

Fig. 15 Curves of fuel steam conversion rate and catalyst deactivation area relationship

第 4 组仿真中，模拟了飞机爬升过程中进口氧气体积分数逐渐增大时反应器的状态，图 16 展示了爬升状态与稳定状态（即进口氧气体积分数基本保持在 0.21）碳氢物转化率的差异。随时间增大的氧气体积分数对碳氢物转化率有明显的提升作用，前 50 s 内 2 条曲线走势基本相同，但爬升状态比稳定状态高 30% 左右。对碳氢物转化率起促进作用的高氧浓度同时也加速了催化剂涂层的失活，如图 17 所示，爬升状态下催化剂涂层完全失活的时间比稳定状态下，提前了 15 s 左右，这一点也反映图 16 中 50～100 s 的区间内，爬升状态的碳氢物转化率衰退速度大于稳定状态，并在 100 s 后低于稳定状态。

值得注意的是，在第一节反应动力学实验得出的反应动力学方程内，大于 0.1 氧气体积分数条件下反应速率与氧气体积分数无关，但在第 4 组仿真中仍然出现了碳氢物转化率随着氧气体积分数升高而增大的情况的出现。这说明，在壁面反应消耗氧气后，一定范围内较高的氧气体积分数仍然可以通过增加管道中心与壁面氧气体积分数梯度的方

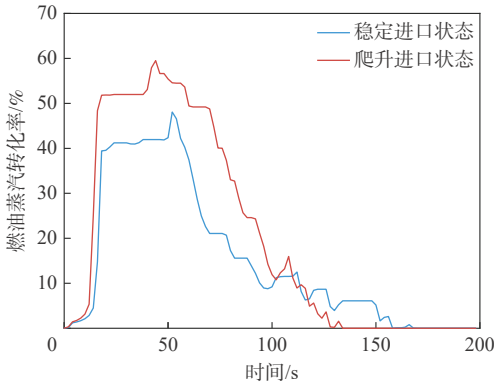


图 16　不同状态进口氧气体积分数对碳氢物转化率的影响

Fig. 16　Influence of hydrocarbon conversion efficiency rate under different inlet oxygen volume fraction

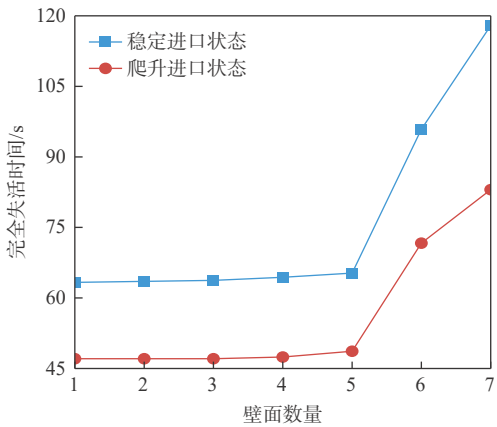


图 17　不同状态进口氧气体积分数对壁面完全失活时间的影响

Fig. 17　Influence of completely deactivation time under different inlet oxygen volume fraction

式提升反应器工作效率。

4　结　论

通过反应动力学实验测定 Pd/γ-Al_2O_3 催化剂对 RP-3 航空煤油催化燃烧的反应动力学方程, 基于测得的反应动力学方程用 Modelica 语言对整体式反应器进行了建模仿真, 得到了以下结论。

1) 飞温现象在整体式反应器内主要发生在靠近反应器中心的出口处, 一定温度范围内的飞温现象有利于反应的加速进行。但整体式反应器内飞温现象难以控制。

2) 对整体式反应器外壁施加冷却可以有效降低核心的最高温度, 但由于多层相界面的存在, 需要将外层管道冷却至极低的温度, 才能有效降低核心温度, 通过外壁面对反应器中心的冷却使反应器催化效率降低 80%。

3) 低燃油蒸汽浓度下反应器后端催化涂层高温失活现象的发生依赖于前段反应器内发生的催化反应对气体的加热, 随着失活现象从后向前发展, 未失活区域变短, 无法提供足够的催化强度使来流升温至失活温度, 失活速度逐渐降低。

4) 飞机爬升阶段氧气体积分数的增大可以有效提高碳氢物转化率, 同时也会加快涂层的失活。

参考文献 (References)

[1] 王明波, 邵垒, 潘俊, 等. 耗氧型燃油箱惰化技术历史和现状[J]. 航空科学技术, 2016, 27(7): 1-7.
WANG M B, SHAO L, PAN J, et al. History and current status of oxygen consumption based fuel tank inerting technology[J]. Aeronautical Science & Technology, 2016, 27(7): 1-7 (in Chinese).

[2] 刘卫华, 冯诗愚. 飞机燃油箱惰化技术[M]. 北京: 科学出版社, 2018: 2-3.
LIU W H, FENG S Y. Aircraft fuel tank inerting technology[M]. Beijing: Science Press, 2018: 2-3. (in Chinese).

[3] 王衡, 王昌红. 近几年整体式催化剂的开发与应用[J]. 工业催化, 2019, 27(10): 15-21.
WANG H, WANG C H. Development and application of monolithic catalyst in recent years[J]. Industrial Catalysis, 2019, 27(10): 15-21.

[4] JOHNSON R W, ZAKI R, YATES S F. Advanced carbon dioxide fuel tank inerting system: US7905259 [P]. 2011-03-15.

[5] 冯诗愚, 刘冠男, 江荣杰, 等. 飞机燃油箱机载惰化技术研究现状与发展趋势[J]. 航空动力学报, 2021, 36(3): 616-625.
FENG S Y, LIU G N, JIANG R J, et al. Research status and development trend of aircraft fuel tank on-board inerting technology[J]. Journal of Aerospace Power, 2021, 36(3): 616-625 (in Chinese).

[6] 彭孝天, 冯诗愚, 周利彪, 等. 温度对耗氧型惰化系统产水性能影响[J]. 航空动力学报, 2020, 35(8): 1622-1627.
PENG X T, FENG S Y, ZHOU L B, et al. Effect of temperature on water production performance of oxygen-consuming inerting system[J]. Journal of Aerospace Power, 2020, 35(8): 1622-1627 (in Chinese).

[7] 李乃珍. 焦炉煤气加氢脱硫 FeMo/Al_2O_3 催化剂积碳失活与性能优化研究[D]. 太原: 太原理工大学, 2022.
LI N Z. Study on carbon deposit deactivation and performance optimization of FeMo/Al_2O_3 hydrodesulfurization catalyst in coke oven gas[D]. Taiyuan: Taiyuan University of Technology, 2022 (in Chinese).

[8] WALKER S, JUNG W, ROBERTSON S. Demonstration of a novel catalyst based green on board inert gas generation system (GOBIGGS) for fuel tank inerting[C]//Proceedings of the American Helicopter Society 69th Annual Forum. Fairfax: American Helicopter Society, 2013, 69(1): 550-559.

[9] 孙浩程, 赵朝成, 陈亚男, 等. 催化燃烧法处理挥发性有机物的研究进展[J]. 现代化工, 2015, 35(6): 57-61.
SUN H C, ZHAO C C, CHEN Y N, et al. Research progress in catalytic combustion of volatile organic compounds[J]. Modern Chemical Industry, 2015, 35(6): 57-61.

[10] 卢泽湘, 范立委. 金属基整体式催化剂的研究进展[J]. 广州化工, 2010, 38(6): 9-12.
LU Z X, FAN L W. Research progress of metallic monolithic catalysts[J]. Guangzhou Chemical Industry, 2010, 38(6): 9-12 (in

Chinese).

[11] WANKER R, BERG M, RAUPENSTRAUCH H, et al. Numerical simulation of monolithic catalysts with a heterogeneous model and comparison with experimental results from a wood-fired domestic boiler[J]. Chemical Engineering & Technology, 2000, 23(6): 535-542.

[12] DEPCIK C, SRINIVASAN A. One + one-dimensional modeling of monolithic catalytic converters[J]. Chemical Engineering & Technology, 2011, 34(12): 1949-1965.

[13] GUO W Y, WU W Z, LUO M, et al. Modeling of diffusion and reaction in monolithic catalysts for the methanol-to-propylene process[J]. Fuel Processing Technology, 2013, 108: 133-138.

[14] PENG X T, WANG H M, HUANG L, et al. Performance of an oxygen-consuming inerting system for an aircraft fuel tank with RP-3 aviation fuel in flight[J]. Areospace Science and Technology, 2022, 123: 107446.

[15] Peng X , Feng S , Chaoyue L I , et al. Effect of fuel type on the performance of an aircraft fuel tank oxygen-consuming inerting system[J]. Chinese Journal of Aeronautics, 2021, 34(3): 82-93.

[16] 郑婷婷, 何俊俊, 吴乐刚, 等. 汽车尾气三效催化剂的失活 [C]//第八届全国工业催化技术及应用年会论文集. 西安: 工业催化, 2011: 11-14.

ZHENG T T. HE J J. WU L G, et al. Deactivation of automobile exhaust three-way catalyst[C]// Proceedings of the 8th National Annual Conference on Industrial Catalysis Technology and Application. Xi'an: Industrial Catalysis, 2011: 11-14 (in Chinese).

[17] 闫红敏. 军用飞机燃油箱的氮气惰化特性研究[D]. 沈阳: 沈阳航空工业学院, 2006.

YAN H M. Study on nitrogen inerting characteristic of military aircraft fuel tank[D]. Shenyang: Shenyang Institute of Aeronautics Engineering, 2006 (in Chinese).

[18] 黄敬敬, 贾志刚, 刘翻艳. Ce 改性 Pd 基整体式催化剂的结构特征及其甲烷催化燃烧性能[J]. 工业催化, 2013, 21(5): 23-29.

HUANG J J, JIA Z G, LIU F Y. Catalytic combustion of lean methane over Pd/γ-Al_2O_3/Cord monolith catalysts modified by Ce[J]. Industrial Catalysis, 2013, 21(5): 23-29 (in Chinese).

Research on oxygen consumption based inerting monolithic catalyst reactor performance

GUO Chenyang[1], LIU Yi[2], LIU Haozheng[1], WANG Junjie[1], GAO jingcheng[1], FENG Shiyu[1,*]

(1. Key Laboratory of Aircraft Environment and Life Support of Ministry of Industry and Information Technology, College of Aerospace Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China;
2. Aviation Key Laboratory of Science and Technology on Aero Electromechanical System Integration, Jincheng Nanjing Engineering Institute of Aircraft System, Aviation Industry Corporation of China, Nanjing 211106, China)

Abstract: An experiment of Pd/γ-Al_2O_3 performance is carried out to get its reaction dynamics function. A two-dimensional monolithic catalyst transient model is established and solved with Modelica. Catalyst transition performance has been studied by simulation. The influence of high temperature and catalyst deactivation on a monolithic catalyst is researched. It was investigated how the reactor was affected by the oxygen content while in flight. The result shows that wall cooling can reduce high temperatures and reactor efficiency at the same time. The catalyst deactivation due to high temperature occurs on the surface near the outlet of the reactor, where the sharpest reaction takes place. Deactivation in Less than half of the catalyst surface area cuts down the efficiency by 80%. The conversion rate of fuel steam can be effectively increased by the increase in oxygen concentration brought on by the loss of pressure during aircraft ascent; however, precautions should be taken to prevent the acceleration of catalyst deactivation caused by high oxygen concentration.

Keywords: catalytic combustion; monolithic catalyst; fuel system; inerting system; catalyst inactivation

Received: 2023-09-01; **Accepted**: 2024-03-29; **Published Online**: 2024-08-15 09:34

URL: link.cnki.net/urlid/11.2625.V.20240814.1628.002

Foundation items: National Natural Science Foundation of China - China Civil Aviation Joint Research Foundation (U1933121); National Natural Science Foundation of China (52372331); Special Fund for basic scientific research funds of central universities; Jiangsu Province superior discipline construction project; Graduate Research and Practice Innovation Program of Nanjing University of Aeronautics and Astronautics (xcxjh20220102)

* **Corresponding author**. E-mail: shiyuf@nuaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0505

基于改进一致性算法的无人机编队队形保持控制器设计

万茹，马梓元，龚华军，王新华*，张帅
(南京航空航天大学 自动化学院，南京 211100)

摘 要：针对传统一致性算法中存在的大增益收敛振荡以及当编队队形受到平动转动同时施加的问题进行了深入研究。在传统一致性编队算法基础上引入虚拟长机，定义新的范数来扩展原有的一致性算法，以解决线性一致性算法中大增益收敛振荡的问题。针对传统一致性编队过程中平动转动的施加导致队形无法保持的问题，引入最大距离作为参考控制器的反馈，修正参考输入轨迹，以确保编队队形的同时，实现平动转动过程中的良好跟踪队形。结合开源飞控 Ardupilot 和开源仿真平台 Airsim 进行了联合半物理仿真，进一步验证了改进算法的可行性。研究结果表明：所提算法在解决传统一致性协议中存在的问题方面具有显著的改进，并在实际应用中具有较好的可行性。

关 键 词：一致性算法；无人机编队；分布式编队；队形保持；Airsim 仿真
中图分类号：V279; V249
文献标志码：A **文章编号**：1001-5965(2025)10-3555-12

近年来，多无人机协同控制的理论和技术迅速发展，功能愈发强大。当前的多无人机协同控制主要还停留在一致性问题、跟踪问题和编队问题[1-3]。目前有关多机编队国内外研究成果丰富，且大部分依据旋翼无人机进行展开。较早的一项研究成果可追溯到 2013 年 Kumar 教授组织 GRASP 研究团队实现了 20 架四旋翼的协同轨迹规划与跟踪飞行任务[4]；从目前无线传感器通信能力来看，分布式通信策略因所需带宽窄，自主性强且计算成本低成为近年被广泛采用的编队通信方式[5]。分布式策略[6]（distributed control）各机与邻居节点通信，编队中不存在中心节点，这降低了对控制中心的依赖，增加了编队的冗余性。Kuriki 等[7]将分布式模型预测控制与一致性算法相结合，在考虑机间碰撞耦合约束的情况下，各机独立规划自己的轨迹实现具有避撞能力的协同编队。针对多无人机的队形保持控制，也已经提出了较多的策略，其中包括：领航-跟随者策略[8]、虚拟结构策略[9]、基于行为策略[10]等，每一种控制方法均有其优缺点。领航-跟随法[11]作为目前最成熟的传统编队控制算法，其主要思想为引入一个“领航者”按照预先设定的轨迹飞行，其他无人机与领航机交换状态信息从而更新自己状态，最终达到一致。虚拟结构法本质上是一种集中式控制方法，可靠性不高。行为法[12]本质上是分布式控制结构，便于添加或减少编队数量，在避碰处理上直观简单；缺点是很难对系统进行严格的数学分析和理论推导证明，无人机的各种行为难以用数学方式进行表述分析。而一致性编队[13]只需要与邻居节点交换信息，保证各机在共同协作的某些状态变量上达成一致意见即可实现大规模的分布式协同编队。该理论由 Ren[14]提出，并且论证了领航-跟随法、行为控制法和虚拟结构法可以被统一到一致性理论的框架中。针对无人机运动学模型特点和远距离成形问题，苟进展等[15]提出了一种基于改进一

收稿日期：2023-08-02；录用日期：2024-01-21；网络出版时间：2024-02-05 09:56
网络出版地址：link.cnki.net/urlid/11.2625.V.20240202.1847.005
*通信作者. E-mail：xhwang@nuaa.edu.cn

引用格式：万茹，马梓元，龚华军，等. 基于改进一致性算法的无人机编队队形保持控制器设计[J]. 北京航空航天大学学报，2025，51（10）：3555-3566. WAN R，MA Z Y，GONG H J，et al. Design of UAV formation-keeping controller based on improved consistency algorithm[J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51（10）：3555-3566（in Chinese）.

致性算法的无人机集结-成形策略，仿真结果表明：所提出的编队控制算法能使无人机在满足机动性约束的情况下，形成稳定队形。但其仅考虑了编队成形过程，而在队形保持过程中仍然存在线性一致性算法大增益收敛振荡问题。刘志江等[16]提出了一致性算法和领航跟随法相结合的无人机编队控制方法，其在无人机的闭环控制上存在不足。已有一致性编队的研究较为深入，但在编队保持跟踪过程中，仍然存在线性一致性算法大增益收敛振荡，且很少有能考虑在平动转动同时叠加到参考输入的时候队形能否保持跟踪的问题。

本文的创新性总结如下：

1）在传统一致性算法的基础上引入虚拟长机，并定义了新范数。

2）基于一致性算法模块引入最大距离作为参考控制器的反馈，并修正参考输入轨迹。

3）搭建了用于验证改进后的一致性算法的开源飞控 Ardupilot 与仿真器 Airsim 的联合半物理仿真平台。

1 无人机编队模型建立

1.1 无人机动力学模型

四旋翼无人机在空间中可以被看作是一个六自由度的刚体，其平动和转动可以用牛顿-欧拉方程来描述：

$$\boldsymbol{f}=m\boldsymbol{a} \tag{1}$$

$$\boldsymbol{\tau}=I\dot{\boldsymbol{\upsilon}}+\boldsymbol{\upsilon}\times(I\boldsymbol{\upsilon}) \tag{2}$$

式中：$\boldsymbol{f}$为四旋翼受到的合外力；m为四旋翼的质量；$\boldsymbol{a}$为加速度；$\boldsymbol{\tau}$为合外力矩；$\boldsymbol{\upsilon}$为四旋翼的角速度，表示为$\boldsymbol{\upsilon}=\begin{bmatrix}p & q & r\end{bmatrix}^{\mathrm{T}}$，$p$、$q$、$r$分别为绕$x$轴旋转的滚转角速率，绕$y$轴旋转的俯仰角速率，绕$z$轴旋转的航向角速率；$\boldsymbol{I}$为惯性对角矩阵，表示为

$$\boldsymbol{I}=\begin{bmatrix}I_{xx} & 0 & 0\\ 0 & I_{yy} & 0\\ 0 & 0 & I_{zz}\end{bmatrix}$$

对于式(1)而言，合外力主要有4个电机带动螺旋桨产生的推力f_t，四旋翼自身重力mg，气动阻力f_{aer}，其中，g为重力加速度；f_t需要从机体坐标系转到大地坐标系。

四旋翼的六自由度非线性模型见下式，其中ϕ为绕x轴旋转的滚转角，θ为绕y轴旋转的俯仰角，ψ为绕z轴旋转的偏航角，$\boldsymbol{I}_M$为螺旋桨转动惯性积，τ_{roll}、τ_{pitch}、τ_{yaw}分别是滚转、俯仰、航向转矩，$k_{\mathrm{aer},x}$、$k_{\mathrm{aer},y}$、$k_{\mathrm{aer},z}$分别为x轴、y轴、z轴气动阻力常数。

$$\begin{cases}\ddot{x}=\dfrac{f_t}{m}(\sin\phi\sin\psi+\cos\phi\cos\psi\sin\theta)-\dfrac{k_{\mathrm{aer},x}|\dot{x}|x}{m}\\ \ddot{y}=\dfrac{f_t}{m}(\cos\phi\sin\theta\sin\psi-\cos\psi\sin\phi)-\dfrac{k_{\mathrm{aer},y}|\dot{y}|y}{m}\\ \ddot{z}=\dfrac{f_t}{m}\cos\theta\cos\phi-g-\dfrac{k_{\mathrm{aer},z}|\dot{z}|z}{m}\\ \ddot{\phi}=\dfrac{I_{yy}-I_{zz}}{I_{xx}}qr-\dfrac{I_M}{I_{xx}}q\Omega+\dfrac{\tau_{\mathrm{roll}}}{I_{xx}}\\ \ddot{\theta}=\dfrac{I_{zz}-I_{xx}}{I_{yy}}pr+\dfrac{I_M}{I_{yy}}p\Omega+\dfrac{\tau_{\mathrm{pitch}}}{I_{yy}}\\ \ddot{\psi}=\dfrac{I_{xx}-I_{yy}}{I_{zz}}qr+\dfrac{\tau_{\mathrm{yaw}}}{I_{zz}}\end{cases} \tag{3}$$

为避免四旋翼运动时出现超出自身物理性能限制的平动和转动情况，对四旋翼的加速度进行限制。大地坐标系下z轴加速度取决于4个螺旋桨总推力f_t的大小，所以z轴加速度限制为

$$\ddot{z}_{\max}=\frac{f_{t,\max}-k_{\mathrm{aer},z}|\dot{z}|z}{m}-g \tag{4}$$

$$\ddot{z}_{\min}=\frac{f_{t,\min}-k_{\mathrm{aer},z}|\dot{z}|z}{m}-g=\frac{-k_{\mathrm{aer},z}|\dot{z}|z}{m}-g \tag{5}$$

x轴和y轴方向加速度的限制取决于俯仰角θ和滚转角ϕ的范围，若$-\theta_{\max}\leqslant\theta\leqslant\theta_{\max}$、$-\phi_{\max}\leqslant\phi\leqslant\phi_{\max}$，则有

$$\ddot{x}_{\max}=-\ddot{x}_{\min}=g\tan\theta_{\max}-\frac{k_{\mathrm{aer},x}|\dot{x}|x}{m} \tag{6}$$

$$\ddot{y}_{\min}=-\ddot{y}_{\max}=-g\frac{\tan\phi_{\max}}{\cos\theta_{\max}}-\frac{k_{\mathrm{aer},y}|\dot{y}|y}{m} \tag{7}$$

记最大加速度限制为$\boldsymbol{a}_{\mathrm{quadpos}}$，最小加速度限制为$\boldsymbol{a}_{\mathrm{quadneg}}$，写成矢量形式则有

$$\begin{cases}\boldsymbol{a}_{\mathrm{quadpos}}=\begin{bmatrix}\ddot{x}_{\max}\\ \ddot{y}_{\max}\\ \ddot{z}_{\max}\end{bmatrix}\\ \boldsymbol{a}_{\mathrm{quadneg}}=\begin{bmatrix}\ddot{x}_{\min}\\ \ddot{y}_{\min}\\ \ddot{z}_{\min}\end{bmatrix}\end{cases} \tag{8}$$

1.2 分布式编队通信网络拓扑模型建立

本文用图论相关知识对分布式编队通信网络建立模型，并用矩阵形式表述。假设编队网络中有n架无人机，每架无人机都是一个通信节点。

引入虚拟长机，虚拟长机与编队中其他节点的联结关系是预设的，通过扩展其邻接矩阵以扩展其编队通信网络拓扑模型。虚拟长机并不是真正的无人机，但包含编队参考状态信息，通过分布式的方式流转到其他节点。为了便于统一管理编队保持中的各无人机的期望位置和实际位置等状态信息，以虚拟长机作为编队队形的中心和新参考坐标系的原点。

以4架无人机为例建立如图1所示通信拓扑网络。不考虑虚拟长机和考虑虚拟长机的邻接矩阵

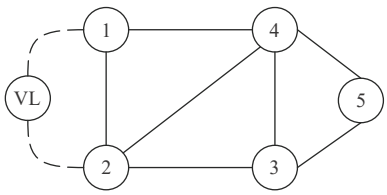


图 1　含虚拟长机编队的无向通信网络

Fig. 1　Undirected communication network with virtual leader formation

分别为

$$\boldsymbol{A}=\boldsymbol{A}^{\mathrm{T}}=\begin{bmatrix}0&1&0&1&0\\1&0&1&1&0\\0&1&0&1&1\\1&1&1&0&1\\0&0&1&1&0\end{bmatrix}$$

$$\boldsymbol{A}_{VL}=\begin{bmatrix}0&1&0&1&0&0\\1&0&1&1&0&0\\0&1&0&1&1&1\\1&1&1&0&1&1\\0&0&1&1&0&0\\0&0&1&1&0&0\end{bmatrix}\in\mathbf{R}^{(n+1)\times(n+1)}$$

每架无人机均有一个自己的一致性算法模块, 通过通信网络中的可连接的边, 将自身信息传递给邻居节点。整体网络通信拓扑控制框图如图 2 所示。

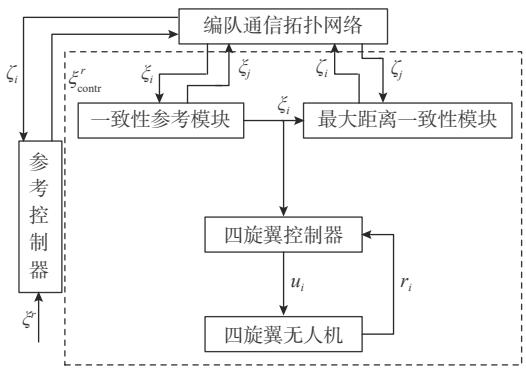


图 2　改进一致性编队算法控制框图

Fig. 2　Control block of improved consistency formation algorithm

其中, $\boldsymbol{\xi}^{\mathrm{r}}$为参考坐标系下的原始参考轨迹状态信息, $\boldsymbol{\xi}^{\mathrm{r}}_{\mathrm{contr}}$为经过参考控制器修正变换以后的期望控制状态信息, $\boldsymbol{\xi}_i$为第i架无人机在参考坐标系下的状态信息, $\boldsymbol{\xi}_j$为第j架无人机在参考坐标系下的状态信息, ζ_i为第i架无人机的最大距离信息, ζ_j为第j架无人机的最大距离信息, u_i为每架无人机自身控制器的输出。

参考控制器会把给定的原始的参考轨迹$\boldsymbol{\xi}^{\mathrm{r}}$以及接收到的编队网络中的最大距离反馈信息$\boldsymbol{\xi}_i$转变为期望控制状态信息$\boldsymbol{\xi}^{\mathrm{r}}_{\mathrm{contr}}$。$\boldsymbol{\xi}^{\mathrm{r}}_{\mathrm{contr}}$是一个修正的输出量, 最终通过一致性参考模块传递到编队网络中, 最终编队网络的各无人机会收敛到相同的状态, 保证无人机状态的一致性。修正的输出量$\boldsymbol{\xi}^{\mathrm{r}}_{\mathrm{contr}}$由于考虑了无人机物理学运动特性的限制的原因, 所以最终在进行队形控制与保持时会减少位置误差, 这是因为最大加速度一般发生在距离队形中心最远的那架无人机上。而参考控制器的队形反馈信息就是来源于每一架无人机的最大距离状态信息$\boldsymbol{\xi}_i$。最大距离状态信息是由最大距离一致性模块提供的, 该模块像一致性参考模块一样, 也会是每一架无人机都含有这个模块, 有自己的最大距离状态信息, 而参考控制器只有一个。

2　改进一致性算法设计

本文设计的改进一致性分布式协同编队保持算法主要包括 3 个部分: 一致性算法模块、基于最大距离的反馈模块和一致性参考控制器模块。

2.1　加入虚拟长机的改进一致性算法模块的设计

一致性参考模块的作用就是保证各无人机最终收敛到设定的参考轨迹状态$\boldsymbol{\xi}^{\mathrm{r}}$。轨迹状态是预先给出的。在一致性算法的作用下, 最终编队中各无人机状态趋于一致, 可得到: $\boldsymbol{\xi}_1=\cdots\boldsymbol{\xi}_n=\boldsymbol{\xi}^{\mathrm{r}}$, 加入虚拟长机后传统线性一致性算法被扩展成如下形式。

$$\dot{\boldsymbol{\xi}}_i=\sum_{j=1}^{n}a_{ij}\left\|\boldsymbol{\xi}_i-\boldsymbol{\xi}_j\right\|_{\mathrm{new}}\left[\dot{\boldsymbol{\xi}}_j-k_1\left(\boldsymbol{\xi}_i-\boldsymbol{\xi}_j\right)\right]+\frac{1}{\eta_i}a_{i(n+1)}\left[\dot{\boldsymbol{\xi}}^{\mathrm{r}}-k_2\left(\boldsymbol{\xi}_i-\boldsymbol{\xi}^{\mathrm{r}}\right)\right]\tag{9}$$

其中, 式 (9) 中第一项表示邻居节点对控制输入的影响, 无人机之间解算出距离, 通过开根号的方式与调整式 (11) 中的增益项, 抑制较大的位置误差, 保证解算出的位置项大小不会突变, 保证响应的稳定性; 第二项表示虚拟长机与编队中其他节点存在联结的影响, 并指定轨迹状态$\boldsymbol{\xi}^{\mathrm{r}}$作为参考的位置信息, 邻居节点的状态信息是控制输入的反馈, k_1和k_2为大于 0 的系数, 用于控制影响比率。η_i为扩展邻接矩阵$\boldsymbol{A}_{VL}$的第i行元素之和, 确保了每一架无人机有较为公平的权重系数而不是受到连接边数的影响。当队形中满足$\boldsymbol{\xi}_1=\cdots=\boldsymbol{\xi}_n=\boldsymbol{\xi}^{\mathrm{r}}$时, 编队内$n$架无人机达到状态一致, 此时$\dot{\boldsymbol{\xi}}_i=0$, 表示所有多旋翼不发生位置改变, 即编队队形已保持稳定。

在一致性协议中的$\boldsymbol{I}_{ij}$为一个单位矢量, 沿着$\boldsymbol{\xi}_i$到$\boldsymbol{\xi}_j$连线,

$$\boldsymbol{I}_{ij}=\frac{\boldsymbol{\xi}_i-\boldsymbol{\xi}_j}{\left\|\boldsymbol{\xi}_i-\boldsymbol{\xi}_j\right\|}\tag{10}$$

$$\left\|\boldsymbol{\xi}_i-\boldsymbol{\xi}_j\right\|_{\mathrm{new}}=\frac{1}{k_0}\left[\sqrt{k_0\left\|\boldsymbol{\xi}_i-\boldsymbol{\xi}_j\right\|+b^2}-b\right]\tag{11}$$

式 (11) 为定义的新范数, 其中: k_0和b这 2 个参数的作用是调节位置控制量的变化快慢和响应速率, $k_0\in(0,1)$。新范数作了限幅处理, 无人机之间距

离的远近不会使得该参数产生突变。新范数由于选择的参数不同其影响方式也不同，从而可以使得编队控制的运动方式更加平滑高效。

2.2 基于最大距离的反馈模块设计

2.1节在传统线性一致性算法的基础上加入虚拟长机，并对协议做了扩展。本节考虑编队运动保持过程中的物理运动特性因素，为编队中的各无人机设计最大距离参考模块与参考控制器模块。

最大距离参考模块的作用是通过考虑编队中每一架无人机在参考坐标系下的最大距离，反映出编队运动过程中队形不同位置的加速度大小，在考虑加速度限制的情况下，便于后续输入到参考控制器做反馈修正实际参考控制状态$\boldsymbol{\xi}^{\mathrm{r}}_{\mathrm{contr}}$，最大距离参考向量可被表示成如下形式：

$$\boldsymbol{\xi}^{\mathrm{r}}=\begin{bmatrix}\max\left(\boldsymbol{r}^{d}_{\mathrm{rel},i}\right)\\ \min\left(\boldsymbol{r}^{d}_{\mathrm{rel},i}\right)\end{bmatrix}\in\mathbf{R}^{6}\quad i=\{1,\cdots,n\}\tag{12}$$

不考虑邻居节点无人机的影响情况下，每一架无人机可以求得自己在参考坐标系下的最大距离向量，形式如下：

$$\boldsymbol{\xi}_i=\begin{bmatrix}\boldsymbol{d}_{\max,i}\\ \boldsymbol{d}_{\min,i}\end{bmatrix}\in\mathbf{R}^{6}\tag{13}$$

式中：$\boldsymbol{d}_{\max,i}$为第i架无人机的最大正距离；$\boldsymbol{d}_{\min,i}$为第i架无人机的最大负距离。图3是二维平面坐标系中第i架无人机的各轴最大距离求解示意。

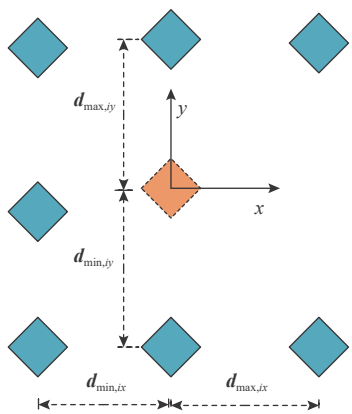


图3 最大距离求解示意

Fig. 3 Solution of the maximum inter-agent distance

其中，橙色菱形表示当前第i架无人机，蓝色菱形是相邻无人机。$\boldsymbol{d}_{\max,ix}$与$\boldsymbol{d}_{\min,ix}$分别是当前无人机x轴方向距离编队队形边缘的最大与最小距离，$\boldsymbol{d}_{\max,iy}$与$\boldsymbol{d}_{\min,iy}$分别是当前无人机y轴方向距离编队队形边缘的最大与最小距离，推广到三维空间可得

$$\boldsymbol{d}_{\max,i}=\left[\boldsymbol{d}_{\max,ix},\boldsymbol{d}_{\max,iy},\boldsymbol{d}_{\max,iz}\right]^{\mathrm{T}}\tag{14}$$

$$\boldsymbol{d}_{\min,i}=\left[\boldsymbol{d}_{\min,ix},\boldsymbol{d}_{\min,iy},\boldsymbol{d}_{\min,iz}\right]^{\mathrm{T}}\tag{15}$$

根据分布式协同设计原则，最大距离向量应考虑邻居节点信息，故写成如下形式：

$$\boldsymbol{\xi}_i=\begin{bmatrix}\max\left(\boldsymbol{d}_{\max,i},\boldsymbol{d}_{\max,j}\right)\\ \max\left(\boldsymbol{d}_{\min,j},\boldsymbol{d}_{\min,i}\right)\end{bmatrix}\qquad j\in N_i\tag{16}$$

式中：N_i为第i架无人机对应的邻居节点集合；$\boldsymbol{\xi}_i$为最大距离状态向量，初始化为$\boldsymbol{\xi}_i=\begin{bmatrix}\boldsymbol{r}^{d}_{\mathrm{rel},i} & \boldsymbol{r}^{d}_{\mathrm{rel},i}\end{bmatrix}^{\mathrm{T}}$，如果通信拓扑网络为一棵有向生成树[17]，编队中各无人机的$\boldsymbol{\xi}_i$可以在有限时间内收敛到$\boldsymbol{\xi}^{\mathrm{r}}$。参考控制器模块会接收来自编队网络的最大距离参考向量$\boldsymbol{\xi}^{\mathrm{r}}$作为反馈，将输入的$\boldsymbol{\xi}^{\mathrm{r}}$修正成期望控制状态$\boldsymbol{\xi}^{\mathrm{r}}_{\mathrm{contr}}$。$\boldsymbol{\xi}^{\mathrm{r}}_{\mathrm{contr}}$在大地坐标系下可被表示成如下形式：

$$\boldsymbol{\xi}^{\mathrm{r}}_{\mathrm{contr}}=\begin{bmatrix}x_c & y_c & z_c & \alpha & \beta & \gamma\end{bmatrix}^{\mathrm{T}}\tag{17}$$

大地坐标系转参考坐标系如图4所示。

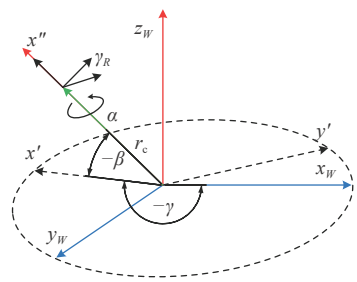


图4 大地坐标系转参考坐标系

Fig. 4 Conversion from geodetic to reference coordinate system

由于有$\boldsymbol{\xi}_i$输入至参考控制器，记$\boldsymbol{d}_{\max}=\begin{bmatrix}d_1 & d_2 & d_3\end{bmatrix}^{\mathrm{T}}$，$\boldsymbol{d}_{\min}=\begin{bmatrix}d_4 & d_5 & d_6\end{bmatrix}^{\mathrm{T}}$，则有：$\boldsymbol{\xi}_i=\begin{bmatrix}d_1 & d_2 & d_3 & d_4 & d_5 & d_6\end{bmatrix}^{\mathrm{T}}$。考虑编队最大距离信息情况，可得最大加速度：

$$\boldsymbol{a}_{\max 1}=\ddot{\boldsymbol{r}}_c+\dot{\boldsymbol{\omega}}\times\boldsymbol{d}_{\max}+\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\boldsymbol{d}_{\max})\tag{18}$$

$$\boldsymbol{a}_{\max 2}=\ddot{\boldsymbol{r}}_c+\dot{\boldsymbol{\omega}}\times\boldsymbol{d}_{\min}+\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\boldsymbol{d}_{\min})\tag{19}$$

实际角速度为

$$\begin{aligned}\boldsymbol{\omega}=&\begin{bmatrix}\cos\gamma & -\sin\gamma & 0\\ \sin\gamma & \cos\gamma & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}\cos\beta & 0 & \sin\beta\\ 0 & 1 & 0\\ -\sin\beta & 0 & \cos\beta\end{bmatrix}\begin{bmatrix}\dot\alpha\\ 0\\ 0\end{bmatrix}+\\ &\begin{bmatrix}\cos\gamma & -\sin\gamma & 0\\ \sin\gamma & \cos\gamma & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}0\\ \dot\beta\\ 0\end{bmatrix}+\begin{bmatrix}0\\ 0\\ \dot\gamma\end{bmatrix}=\\ &\begin{bmatrix}\cos\gamma\cos\beta & -\sin\gamma & 0\\ \sin\gamma\cos\beta & \cos\gamma & 0\\ -\sin\beta & 0 & 1\end{bmatrix}\begin{bmatrix}\dot\alpha\\ \dot\beta\\ \dot\gamma\end{bmatrix}=\boldsymbol{R}^{W}_{E}\dot{\delta}\end{aligned}\tag{20}$$

对上式求偏导，角加速度矢量为

$$\begin{aligned}\dot{\boldsymbol{\omega}}=&\boldsymbol{R}^{W}_{E}\ddot{\delta}+\dot{\boldsymbol{R}}^{W}_{E}\dot{\delta}=\\ &\Big[\ddot\alpha\cos\beta\cos\gamma-\dot\alpha\dot\beta\sin\beta\cos\gamma-\\ &\dot\alpha\cos\beta\dot\gamma\sin\gamma-\ddot\beta\sin\gamma-\dot\beta\dot\gamma\cos\gamma\ddot\alpha\cos\beta\sin\gamma-\\ &\dot\alpha\dot\beta\sin\beta\sin\gamma+\dot\alpha\cos\beta\dot\gamma\cos\gamma+\ddot\beta\cos\gamma-\\ &\dot\beta\dot\gamma\sin\gamma-\cos\beta\dot\beta\dot\alpha-\sin\beta\ddot\alpha+\ddot\gamma\Big]\end{aligned}\tag{21}$$

结合加速度限制式 (8) 可得

$$\boldsymbol{a}_{\text{quadneg}} \leqslant \boldsymbol{a}_{\max 1} \leqslant \boldsymbol{a}_{\text{quadpos}} \tag{22}$$

$$\boldsymbol{a}_{\text{quadneg}} \leqslant \boldsymbol{a}_{\max 2} \leqslant \boldsymbol{a}_{\text{quadpos}} \tag{23}$$

如果编队运动过程中产生的加速度$\boldsymbol{a}_{\max 1}$、$\boldsymbol{a}_{\max 2}$有任意一个超过式 (20) 或式 (23) 的物理运动特性的限制, 此时会产生位置误差。

2.3　一致性参考控制器模块设计

参考控制器模块的设计初衷就是修正期望参考轨迹$\boldsymbol{\xi}^{\text{r}}$, 保证输入到编队网络中的控制状态信息$\boldsymbol{\xi}^{\text{r}}_{\text{contr}}$, 不会让编队中各无人机发生因自身加速度限制从而产生编队保持跟踪误差的情况, 因此采用 PD 控制器加误差修正向量对期望控制状态信息$\boldsymbol{\xi}^{\text{r}}_{\text{contr}}$进行修正。

图 5 为参考控制器模块设计框图。

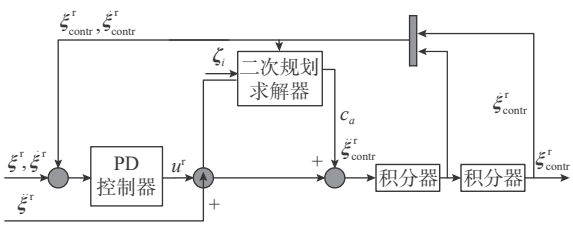


图 5　参考控制器模块设计框图

Fig. 5　Design block of reference controller module

参考控制器表达式如下:

$$\ddot{\boldsymbol{\xi}}^{\text{r}}_{\text{contr}} = \ddot{\boldsymbol{\xi}}^{\text{r}} + \boldsymbol{k}^{\text{r}}_d\left(\dot{\boldsymbol{\xi}}^{\text{r}} - \dot{\boldsymbol{\xi}}^{\text{r}}_{\text{contr}}\right) + \boldsymbol{k}^{\text{r}}_p\left(\boldsymbol{\xi}^{\text{r}} - \boldsymbol{\xi}^{\text{r}}_{\text{contr}}\right) + \boldsymbol{c}_a = \boldsymbol{u}^{\text{r}} + \boldsymbol{c}_a \tag{24}$$

式中: $\boldsymbol{c}_a \in \mathbf{R}^6$是误差修正向量, 如果编队运动中满足式 (8) 的加速度限制, 修正向量为零向量; 如果$\boldsymbol{a}_{\max 1}$或$\boldsymbol{a}_{\max 2}$超出了加速度限制, 那么误差修正向量会通过求解二次规划问题得到。$\boldsymbol{c}_a$与 PD 控制器的输出$\boldsymbol{u}^{\text{r}}$的加和就是最终参考坐标系下的期望控制加速度信息$\ddot{\boldsymbol{\xi}}^{\text{r}}_{\text{contr}}$。下面推导如何把加速度限制条件转化到$\boldsymbol{c}_a$的不等式约束条件。由加速度限制公式可得

$$\begin{aligned}\boldsymbol{a}_{\max 1} = \ddot{\boldsymbol{r}}_c + \dot{\boldsymbol{\omega}} \times \boldsymbol{d}_{\max} + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{d}_{\max}) \leqslant \boldsymbol{a}_{\text{quadpos}} = \\ \ddot{\boldsymbol{r}}_c + \dot{\boldsymbol{\omega}} \times \boldsymbol{d}_{\max} \leqslant \boldsymbol{a}_{\text{quadpos}} - \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{d}_{\max})\end{aligned} \tag{25}$$

根据式 (21) 可将$\dot{\boldsymbol{\omega}} \times \boldsymbol{d}_{\max}$化为

$$\begin{aligned}\dot{\boldsymbol{\omega}} \times \boldsymbol{d}_{\max} = \left(\boldsymbol{R}^W_E \ddot{\delta} + \dot{\boldsymbol{R}}^W_E \dot{\delta}\right) \times \boldsymbol{d}_{\max} = \\ \boldsymbol{R}^W_E \ddot{\delta} \times \boldsymbol{d}_{\max} + \dot{\boldsymbol{R}}^W_E \dot{\delta} \times \boldsymbol{d}_{\max}\end{aligned} \tag{26}$$

记 $\boldsymbol{G}_1\ddot{\delta} = \boldsymbol{R}^W_E \ddot{\delta} \times \boldsymbol{d}_{\max}$, $\boldsymbol{h}_1 = \dot{\boldsymbol{R}}^W_E \dot{\delta} \times \boldsymbol{d}_{\max}$, 所以有$\dot{\boldsymbol{\omega}} \times \boldsymbol{d}_{\max} = \boldsymbol{G}_1\ddot{\delta} + \boldsymbol{h}_1$, 其中:

$$\boldsymbol{G}_1 = \begin{bmatrix} d_3\cos\beta\sin\gamma + d_2\sin\beta & \cos\gamma d_3 & -d_2 \\ d_1\sin\beta - d_3\cos\beta\cos\gamma & d_3\sin\gamma & d_1 \\ d_2\cos\beta\cos\gamma - d_1\cos\beta\sin\gamma & -d_2\sin\gamma - d_1\cos\gamma & 0 \end{bmatrix}$$

联合上式, 可得

$$\ddot{\boldsymbol{r}}_c + \boldsymbol{G}_1\ddot{\delta} + \boldsymbol{h}_1 \leqslant \boldsymbol{a}_{\text{quadpos}} - \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{d}_{\max}) \tag{27}$$

上式左边写成矩阵形式则有:

$$\begin{bmatrix} \boldsymbol{I}_3 & \boldsymbol{G}_1 \end{bmatrix} \ddot{\boldsymbol{\xi}}^{\text{r}}_{\text{contr}} \leqslant \boldsymbol{a}_{\text{quadpos}} - \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{d}_{\max}) - \boldsymbol{h}_1 \tag{28}$$

又因为$\ddot{\boldsymbol{\xi}}^{\text{r}}_{\text{contr}} = \boldsymbol{u}^{\text{r}} + \boldsymbol{c}_a$, 式 (28) 可写为

$$\tilde{\boldsymbol{G}}_1 \boldsymbol{c}_a \leqslant \boldsymbol{a}_{\text{quadpos}} + \tilde{\boldsymbol{h}}_1 \tag{29}$$

式 (29) 中, $\tilde{\boldsymbol{G}}_1 = \begin{bmatrix} \boldsymbol{I}_3 & \boldsymbol{G}_1 \end{bmatrix}$, $\tilde{\boldsymbol{h}}_1 = -\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{d}_{\max}) - \boldsymbol{h}_1 - \begin{bmatrix} \boldsymbol{I}_3 & \boldsymbol{G}_1 \end{bmatrix} \boldsymbol{u}^{\text{r}}$, 同理可以推出另外一侧加速度的限制产生的不等式关系:

$$-\tilde{\boldsymbol{G}}_1 \boldsymbol{c}_a \leqslant -\boldsymbol{a}_{\text{quadneg}} - \tilde{\boldsymbol{h}}_1 \tag{30}$$

根据$\boldsymbol{a}_{\max 2}$的不等式关系, 同样可以得到:

$$\tilde{\boldsymbol{G}}_2 \boldsymbol{c}_a \leqslant \boldsymbol{a}_{\text{quadpos}} + \tilde{\boldsymbol{h}}_2 \tag{31}$$

$$-\tilde{\boldsymbol{G}}_2 \boldsymbol{c}_a \leqslant -\boldsymbol{a}_{\text{quadneg}} - \tilde{\boldsymbol{h}}_2 \tag{32}$$

式中: $\tilde{\boldsymbol{G}}_2 = \begin{bmatrix} \boldsymbol{I}_3 & \boldsymbol{G}_2 \end{bmatrix}$; $\tilde{\boldsymbol{h}}_2 = -\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{d}_{\min}) - \boldsymbol{h}_2 - \begin{bmatrix} \boldsymbol{I}_3 & \boldsymbol{G}_2 \end{bmatrix} \boldsymbol{u}^{\text{r}}$; $\boldsymbol{h}_2 = \dot{\boldsymbol{R}}^W_E \dot{\delta} \times \boldsymbol{d}_{\min}$;

$$\boldsymbol{G}_2 = \begin{bmatrix} d_6\cos\beta\sin\gamma + d_5\sin\beta & \cos\gamma d_6 & -d_5 \\ d_4\sin\beta - d_6\cos\beta\cos\gamma & d_6\sin\gamma & d_4 \\ d_5\cos\beta\cos\gamma - d_4\cos\beta\sin\gamma & -d_5\sin\gamma - d_4\cos\gamma & 0 \end{bmatrix}$$

联合上述 4 个不等式 (29)～(32) 可得

$$\boldsymbol{G}_a \boldsymbol{c}_a \leqslant \boldsymbol{h}_a \tag{33}$$

展开为

$$\begin{bmatrix} \tilde{\boldsymbol{G}}_1 \\ -\tilde{\boldsymbol{G}}_1 \\ \tilde{\boldsymbol{G}}_2 \\ -\tilde{\boldsymbol{G}}_2 \end{bmatrix} \boldsymbol{c}_a \leqslant \begin{bmatrix} \boldsymbol{a}_{\text{quadpos}} \\ -\boldsymbol{a}_{\text{quadneg}} \\ \boldsymbol{a}_{\text{quadpos}} \\ -\boldsymbol{a}_{\text{quadneg}} \end{bmatrix} + \begin{bmatrix} \tilde{\boldsymbol{h}}_1 \\ -\tilde{\boldsymbol{h}}_1 \\ \tilde{\boldsymbol{h}}_2 \\ -\tilde{\boldsymbol{h}}_2 \end{bmatrix} \tag{34}$$

由此问题就化为

$$\begin{cases} \min\limits_{\boldsymbol{c}_a} \dfrac{1}{2}\boldsymbol{c}^{\text{T}}_a \boldsymbol{H}_a \boldsymbol{c}_a, \quad \boldsymbol{H}_a = \begin{bmatrix} h_1 & & \\ & \ddots & \\ & & h_6 \end{bmatrix} \in \mathbf{R}^{6\times 6} \\ \text{s.t. } \boldsymbol{G}_a \boldsymbol{c}_a \leqslant \boldsymbol{h}_a \end{cases} \tag{35}$$

最终解算出的修正量$\boldsymbol{c}_a$与 PD 控制器的输出量进行加和后二次积分可得最终的输入修正参考状态$\boldsymbol{\xi}^{\text{r}}_{\text{contr}}$, 因此在 2.1 节设计的一致性协议可被修改为

$$\begin{aligned}\ddot{\boldsymbol{\xi}}_i = \sum_{j=1}^{n} \boldsymbol{a}_{ij} \left\|\boldsymbol{\xi}_i - \boldsymbol{\xi}_j\right\|_{\text{new}} \left[\dot{\boldsymbol{\xi}}_j - \boldsymbol{k}_1\left(\boldsymbol{\xi}_i - \boldsymbol{\xi}_j\right)\right] + \\ \frac{1}{\eta_i} \boldsymbol{a}_{i(n+1)} \left[\dot{\boldsymbol{\xi}}^{\text{r}}_{\text{contr}} - \boldsymbol{k}_2\left(\boldsymbol{\xi}_i - \boldsymbol{\xi}^{\text{r}}_{\text{contr}}\right)\right]\end{aligned} \tag{36}$$

3　仿真验证

3.1　改进一致性算法的仿真验证

本节实验对比验证传统线性一致性算法和改

进邻接矩阵权重的一致性算法性能。

设定编队期望高度为4 m，偏航角为π/4，速度为2.5 m/s，编队队形为如图6所示的菱形，编队整体在xOy平面的运动轨迹与x、y轴各有45°的夹角。

图7～图9给出了改进前后的控制响应曲线。图10给出了编队XY平面飞行轨迹。

传统一致性算法和改进一致性算法都可使编队收敛至指定状态。在高度响应曲线对比上，改进一致性算法较原算法整体没有较大的超调，且收敛速度更快。在图8(a)与图8(b)的速度响应曲线的对比上，传统一致性算法控制的无人机在速度状态调节中速度最大超调量达到0.6 m/s，6 s以前出现较大的振荡调节，6 s以后速度状态收敛并趋于一致；而采用改进一致性编队控制算法，4架无人机有较小超调，但并未出现振荡调节的情况，6 s以后速度状态达成一致。

综合对比，考虑邻接矩阵权重和加入虚拟长机的改进一致性算法在编队各个无人机状态收敛调节中有较小的超调量和较快的调节响应；加入状态微分信息的改进一致性协议可以有效保持设定编队队形进行运动。

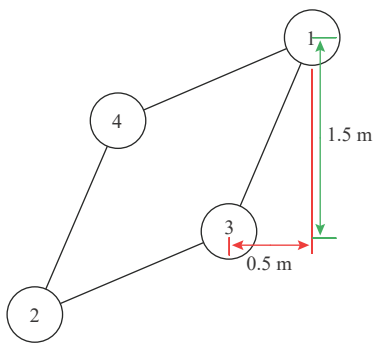


图6 4机设定菱形编队

Fig. 6 Diamond formation of four UAVs

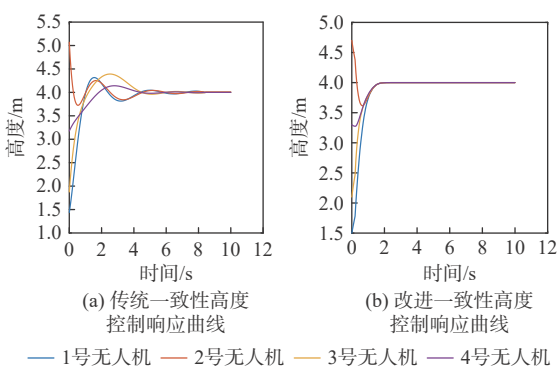


图7 传统一致性与改进一致性算法高度控制响应曲线对比

Fig. 7 Height control response curves of traditional and improved consistency algorithms

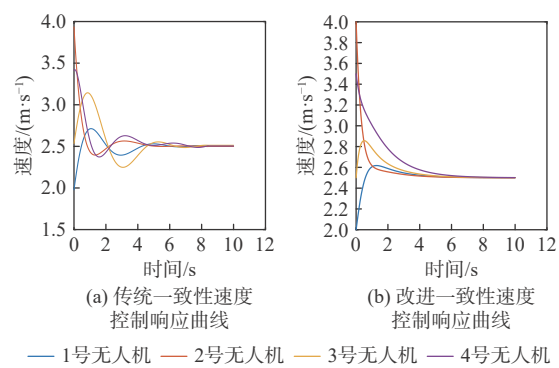


图8 传统一致性与改进一致性算法速度控制响应曲线对比

Fig. 8 Speed control response curves of traditional and improved consistency algorithms

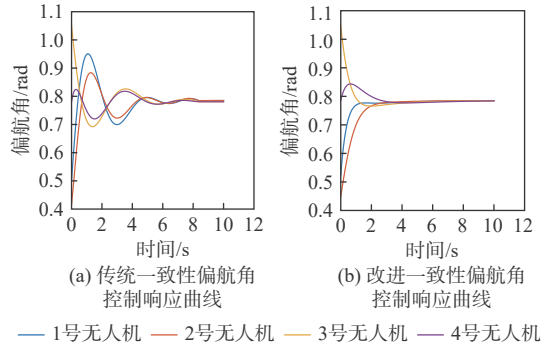


图9 传统一致性与改进一致性算法偏航角控制响应曲线对比

Fig. 9 Yaw angle control response curves of traditional and improved consistency algorithms

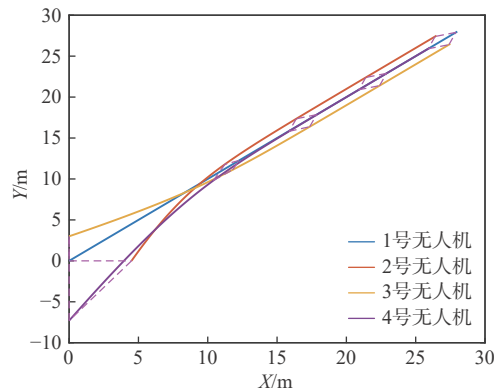


图10 编队xy平面飞行轨迹

Fig. 10 Formation flight trajectories on xy plane

3.2 加入最大距离反馈模块的效果验证

本节验证加入最大距离反馈模块以后对编队保持跟踪性能的影响。编队通信拓扑网络由文献[18]提出的算法随机创建生成。

设置8架无人机参与编队，虚拟长机与1号无人机连接，随机生成的通信无向图和最小生成树如图11和图12所示。

设置期望编队队形为“一字形”，8架无人机沿

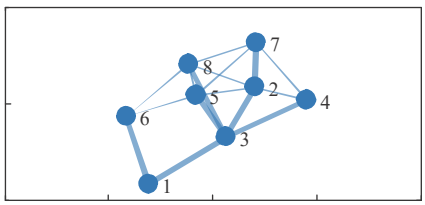


图 11　通信拓扑网络

Fig. 11　Communication topologynetwork

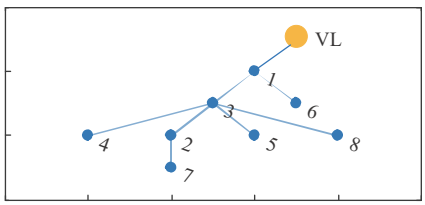


图 12　最小生成树

Fig. 12　Minimum spanning tree

着y轴对称横一字排开, 间距均为 3 m, 三维空间队形如图 13 所示。

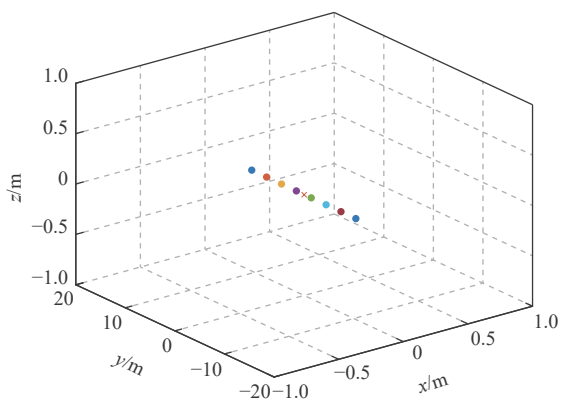


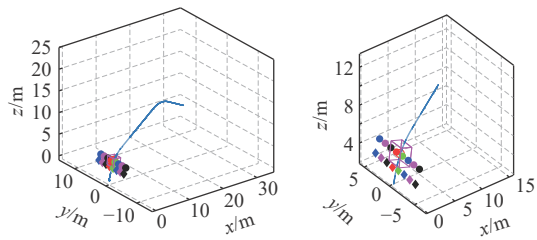
图 13　编队初始“横一字”队形

Fig. 13　Initial horizontal-line formation

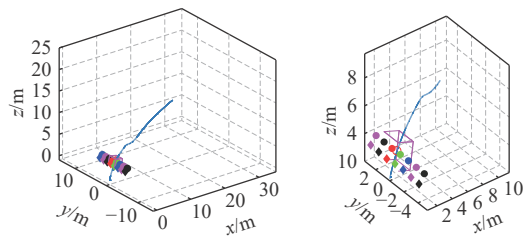
使用时变参考状态量$\boldsymbol{\xi}^{\mathrm{r}}(t)$的二阶导数作为参考轨迹输入驱动量$\ddot{\boldsymbol{\xi}}^{\mathrm{r}}(0)=\left[\begin{array}{cccccc}\ddot{x}_c^{\mathrm{r}} & \ddot{y}_c^{\mathrm{r}} & \ddot{z}_c^{\mathrm{r}} & \ddot{\alpha}_c^{\mathrm{r}} & \ddot{\beta}_c^{\mathrm{r}} & \ddot{\gamma}_c^{\mathrm{r}}\end{array}\right]^{\mathrm{T}}$, 初始状态设置为$\boldsymbol{\xi}^{\mathrm{r}}(0)=\left[\begin{array}{cccccc}0 & 0 & 0 & 0 & 0 & 0\end{array}\right]^{\mathrm{T}}$, $\dot{\boldsymbol{\xi}}^{\mathrm{r}}(0)=\left[\begin{array}{cccccc}0 & 0 & 0 & 0 & 0 & 0\end{array}\right]^{\mathrm{T}}$。

算法改进前后编队跟踪控制结果对比如图 14～图 17 所示。

图 14～图 17 中菱形表示当前编队队形, 圆形表示对应的 50 个采样点以后的期望编队队形, 粉色的正方体表示以编队队形中心为原点所建立的参考坐标系框架, 蓝色的曲线表示传统一致性编队和改进一致性编队的参考运动轨迹, 因为改进一致性算法加入$\boldsymbol{c}_a$向量修正输入的参考轨迹, 所以原始参考状态$\boldsymbol{\xi}^{\mathrm{r}}$与修正的输出量$\boldsymbol{\xi}_{\mathrm{contr}}^{\mathrm{r}}$此时不相等, 而传统一致性算法中并未考虑编队整体转动中最大加速度限制, 因此没有加入轨迹的误差修正向量$\boldsymbol{c}_a$,



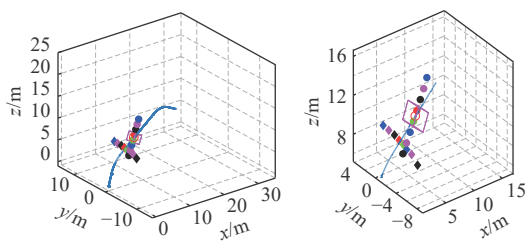
(a) t=2 s时传统一致性编队队形



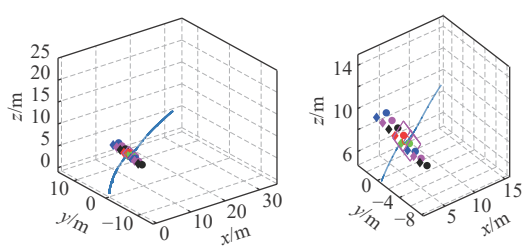
(b) t=2 s时传统一致性编队队形

图 14　t = 2 s 时传统与改进一致性算法队形轨迹跟踪过程

Fig. 14　Formation tracking traditional and improved consistency algorithms at t=2 s



(a) t=5 s时传统一致性编队队形



(b) t=5 s时改进一致性编队队形

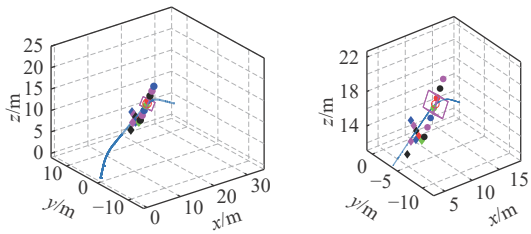
图 15　t = 5 s 时传统与改进一致性算法队形轨迹跟踪过程

Fig. 15　Formation tracking traditional and improved consistency algorithms at t=5 s

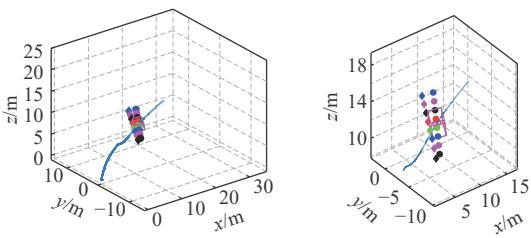
此时$\boldsymbol{\xi}^{\mathrm{r}}=\boldsymbol{\xi}_{\mathrm{contr}}^{\mathrm{r}}$。

从仿真结果来看, 前 2 s 编队轨迹运动跟踪中, 可以看到两种编队算法都能很好地对期望编队队形进行保持, 因为此时在设定的参考状态中只有平动状态; 在 5 s 时看到传统一致性算法对期望运动队形的跟踪有明显保持滞后, 产生较大的跟踪误差和队形保持误差, 因为此时参考运动状态在平动的基础上添加了额外的转动, 结果表明加入最大距离修正的参考控制器可有效保持队形。

修正前后的期望加速度与上下限加速度以z方



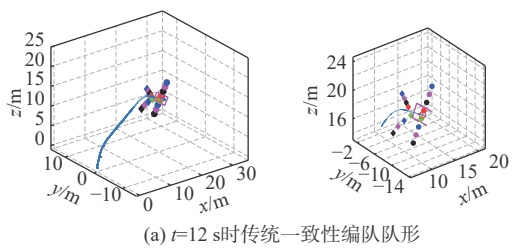
(a) t=9 s时传统一致性编队队形



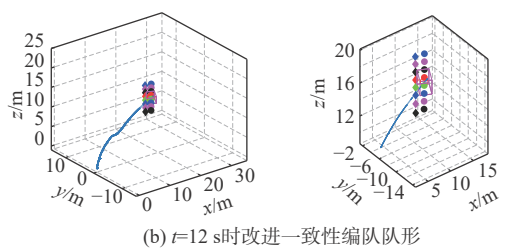
(b) t=9 s时改进一致性编队队形

图 16 t = 9 s 时传统与改进一致性算法队形轨迹跟踪过程

Fig. 16 Formation tracking traditional and improved consistency algorithms at t=9 s



(a) t=12 s时传统一致性编队队形

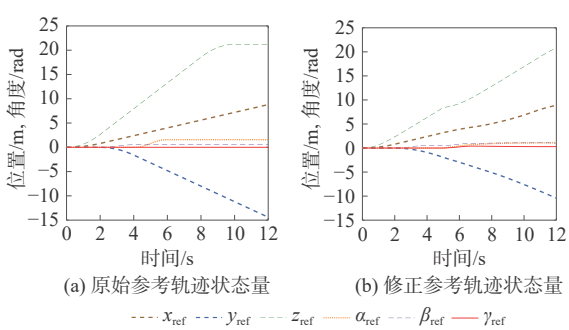


(b) t=12 s时改进一致性编队队形

图 17 t = 12 s 时传统与改进一致性算法队形轨迹跟踪过程

Fig. 17 Formation tracking at traditional and improved consistency algorithms t=12 s

向为例，x、y方向同理。图 18 为修正前后参考轨迹状态量，而从图 19～图 21 可知各机的x、y、z三轴修正后参考轨迹基本收敛在同一参考状态量，表明本节扩展的改进一致性协议是收敛的；根据图 22(a)与图 22(b) 的对比可知，各机在没有加入参考控制



(a) 原始参考轨迹状态量 (b) 修正参考轨迹状态量

图 18 修正前后参考轨迹状态量对比图

Fig. 18 Comparison of reference trajectory states before and after correction

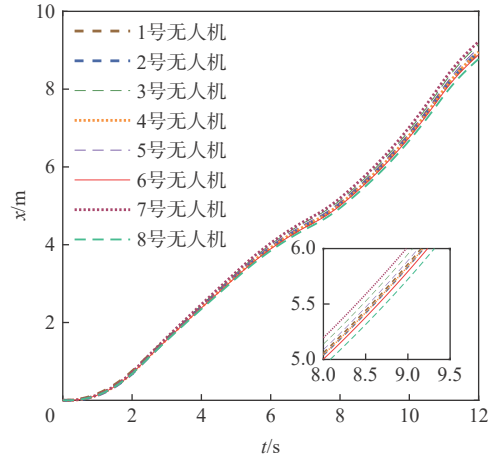


图 19 修正后编队各机x方向参考运动轨迹

Fig. 19 Corrected x-direction reference trajectories for each UAV

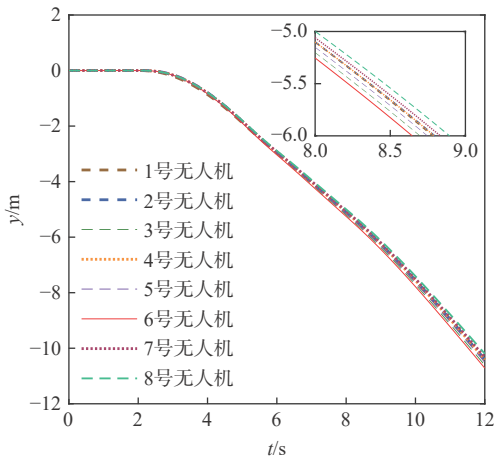


图 20 修正后编队各机y方向参考运动轨迹

Fig. 20 Corrected y-direction reference trajectories for each UAV

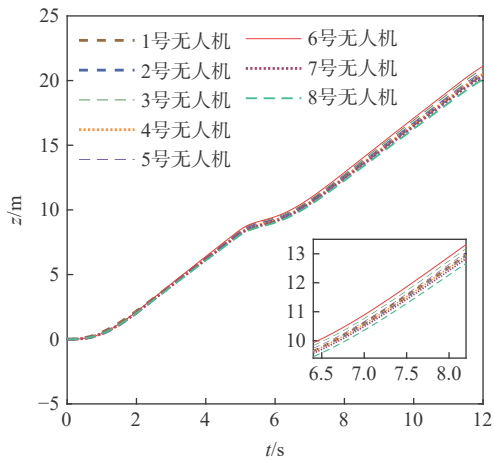


图 21 修正后编队各机z方向参考运动轨迹

Fig. 21 Corrected z-direction reference trajectories for each UAV

器修正参考输入轨迹的情况下, 给定的期望加速度有超过编队无人机自身加速度上下限的情况, 这也是编队各无人机在轨迹跟踪过程中无法保持队形的原因, 由z轴的期望加速度在第 5 s 左右甚至出现了接近−12 m/s 的加速度; 而考虑最大距离限制因素以后, 通过二次规划求解修正参考输入状态, 保证了期望加速度不会超过无人机的加速度上下限, 从而确保了编队在平动转动同时施加过程中的队形保持能力。

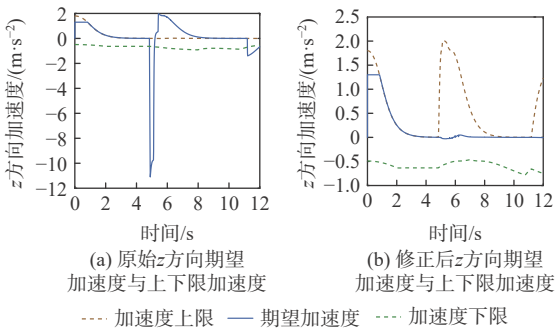


图 22　修正前后z方向期望加速度与上下限加速度

Fig. 22　Expected and upper/lower bound acceleration in z-direction before and after correction

3.3　加入最大距离反馈模块的效果验证

本节选用开源飞控 Ardupilot 作为半物理仿真的飞控软件部分实现, 选取了微软公司基于游戏引擎开发的开源跨平台仿真器 Airsim, 总体框架图如图 23 所示。

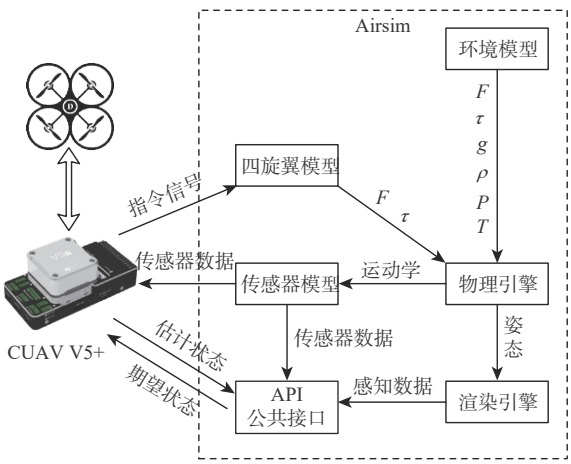
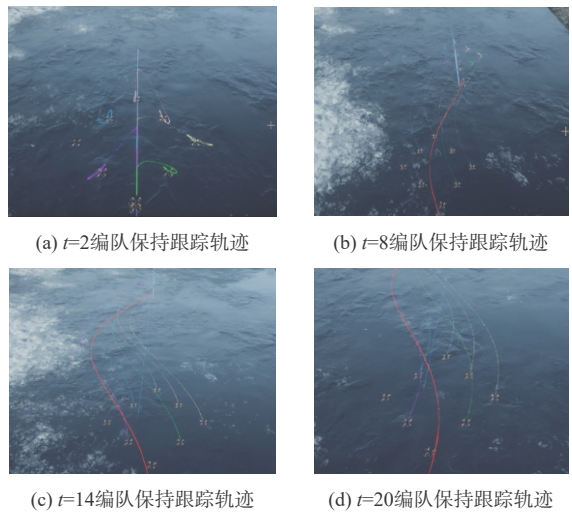


图 23　Airsim 总体框架

Fig. 23　Overall framework of AirSim simulation system

设置 8 机四旋翼进行编队保持飞行半物理仿真实验, 如图 24 所示。

由图 24(a)～由图 24(d) 的视景仿真以及图 25～图 27 参数曲线可知, 8 架无人机在一致性协议驱动下, 初始“一字形”转成了“菱形”后, 各机高度状态



(a) t=2编队保持跟踪轨迹　(b) t=8编队保持跟踪轨迹

(c) t=14编队保持跟踪轨迹　(d) t=20编队保持跟踪轨迹

图 24　不同时刻 8 机编队保持跟踪轨迹

Fig. 24　Eight-UAV formation tracking at different time

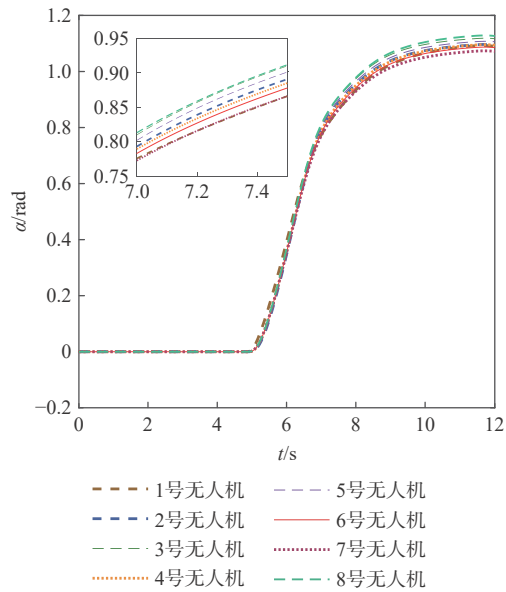


图 25　编队各机滚转角值

Fig. 25　Roll angle of each UAV in formation

统一到了设定的同一高度, 并按照预先设定的摆动轨迹保持飞行, 队形保持了良好的控制精度, 并未出现较大的队形跟踪误差。这说明, 本文提出的改进一致性算法有着良好的编队保持跟踪能力, 在工程实际落地中具有可行性与实践性。

设置队形变换目标为将 13 架无人机由初始“M”型队形变换到“A”型队形, 1～13 号机初始坐标为 (0,0,10), (6,0,10), (12,0,10), (18,0,10), (12,3,10), (6,6,10), (0,9,10), (6,12,10), (12,15,10), (18,18,10), (12,18,10), (6,18,10), (0,18,10); 最终队形的坐标为 (60,60,50), (64,63,50), (68,66,50), (72,69,50), (76,72,50), (80,75,50), (76,78,50), (72,81,50), (68,84,50), (64,87,50), (60,90,50), (68,69,50), (68,72,50) , 初始队形

和最终队形如图 28 和图 29所示。

队形变换视景仿真结果如图 30 所示。

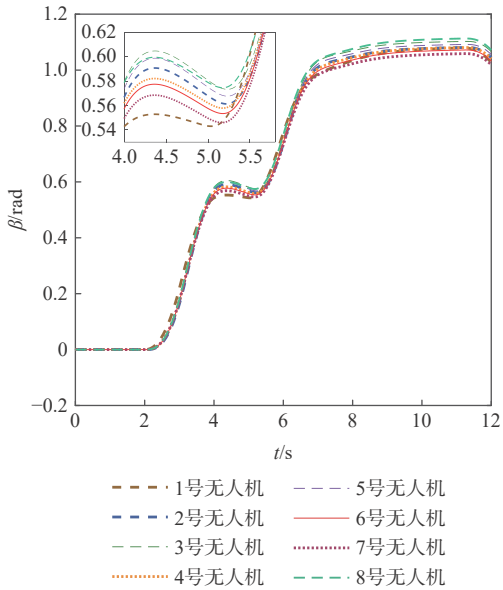


图 26 编队各机俯仰角值

Fig. 26 Pitch angle of each UAV in formation

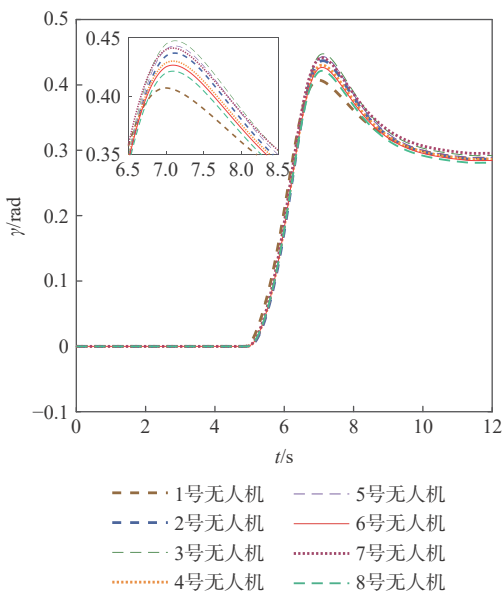


图 27 编队各机偏航角值

Fig. 27 Yaw angle of each UAV in formation

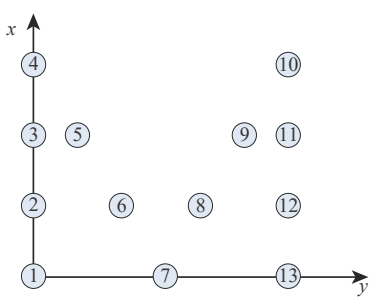


图 28 编队初始队形

Fig. 28 Initial formation

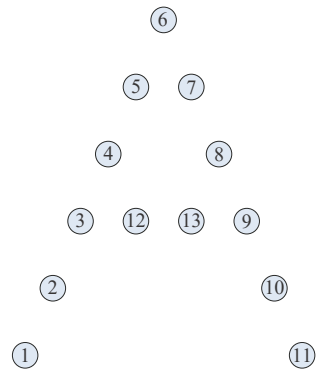
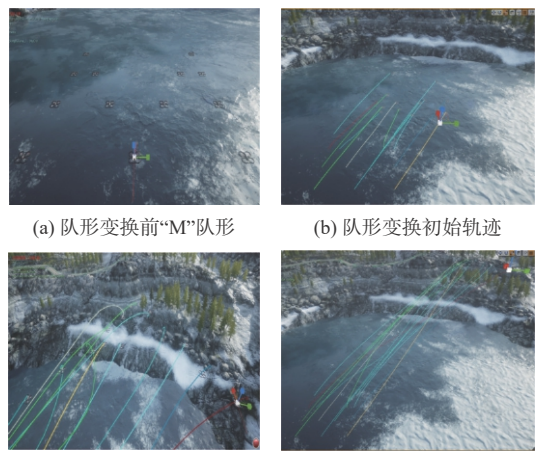


图 29 编队最终队形

Fig. 29 Final formation



(a) 队形变换前“M”队形 (b) 队形变换初始轨迹

(c) 队形变换后“A”队形 (d) 队形变换整体轨迹图

图 30 队形变换视景仿真

Fig. 30 Visual simulation of formation transformation

4 结 论

1）本文通过引入虚拟长机、定义新范数改进的一致性算法模块，解决了线性一致性算法大增益收敛振荡问题。

2）在改进算法的基础上设计的参考控制器，有效改善了传统一致性编队过程中平动转动同时施加导致队形无法保持的问题，确保编队队形的同时，使得平动转动过程中也可以保持良好的跟踪队形。

3）对比 MATLAB 数字仿真，由于实际半物理工程验证中考虑了环境模型、传感器噪声、飞控硬件平台算力等因素的影响，实际仿真曲线存在一定噪声，编队保持跟踪过程中略有低频振荡。但半物理仿真结果与 MATLAB 数字仿真结果一致，实验结果表明本文研究的分布式编队队形保持算法可用于实际工程飞行。

参考文献（References）

[1] 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).

[2] REN W. Consensus strategies for cooperative control of vehicle formations[J]. IET Control Theory& Applications, 2007, 1(2): 505-512.

[3] 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).

[4] KUSHLEYEV A, MELLINGER D, POWERS C, et al. Towards a swarm of agile micro quadrotors[J]. Autonomous Robots, 2013, 35(4): 287-300.

[5] 周文惠, 齐瑞云, 姜斌. 面向突发故障的分布式多无人机任务重规划方法[J]. 控制与决策, 2023, 38(05): 1373-1385.
ZHOU W H, QI R Y, JIANG B. A mission replanning method of distributed multiple unmanned aerial vehicles for pop-up faults[J]. Control and Decision, 2023, 38(05): 1373-1385(in Chinese).

[6] REN W, SORENSEN N. Distributed coordination architecture for multi-robot formation control[J]. Robotics and Autonomous Systems, 2008, 56(4): 324-333.

[7] KURIKI Y, NAMERIKAWA T. Formation control with collision avoidance for a multi-UAV system using decentralized MPC and consensus-based control[J]. SICE Journal of Control, Measurement, and System Integration, 2015, 8(4): 285-294.

[8] YUAN W, CHEN Q Y, HOU Z X, et al. Multi-UAVs formation flight control based on leader-follower pattern[C]//Proceedings of the 2017 36th Chinese Control Conference. Piscataway: IEEE Press, 2017: 1276-1281.

[9] 李正平, 鲜斌. 基于虚拟结构法的分布式多无人机鲁棒编队控制[J]. 控制理论与应用, 2020, 37(11): 2423-2431.
LI Z P, XIAN B. Robust distributed formation control of multiple unmanned aerial vehicles based on virtual structure[J]. Control Theory & Applications, 2020, 37(11): 2423-2431(in Chinese).

[10] 邱华鑫, 段海滨, 范彦铭. 基于鸽群行为机制的多无人机自主编队[J]. 控制理论与应用, 2015, 32(10): 1298-1304.
QIU H X, DUAN H B, FAN Y M. Multiple unmanned aerial vehicle autonomous formation based on the behavior mechanism in pigeon flocks[J]. Control Theory & Applications, 2015, 32(10): 1298-1304(in Chinese).

[11] CHALLA V R, RATNOO A. Analysis of UAV kinematic constraints for rigid formation flying[C]//Proceedings of the AIAA Guidance, Navigation, and Control Conference. Reston: AIAA, 2016: 2105.

[12] CROWTHER W J. Rule-based guidance for flight vehicle flocking[J]. Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering, 2004, 218(2): 111-124.

[13] 任伟, 兰德尔. W. 比尔德. 多航行体协同控制中的分布式一致性: 理论与应用[M]. 吴晓锋, 译. 北京: 电子工业出版社, 2014: 5-7.
REN W, BEARD R W. Distributed consensus in multi-vehicle cooperative control [M]. WU X F, translated. Beijing: Publishing House of Electronics Industry, 2014: 5-7(in Chinese).

[14] REN W. Consensus based formation control strategies for multi-vehicle systems[J]. Proceedings of the American Control Conference, 2006: 4237-4242.

[15] 苟进展, 梁天骄, 陶呈纲, 等. 基于一致性理论的无人机编队控制与集结方法[J]. 北京航空航天大学学报, 2024, 50(5): 1646-1654.
GOU J Z, LIANG T J, TAO C G, et al. A Method of UAV Formation control and assembly based on consistency theory[J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50(5): 1646-1654(in Chinese).

[16] 刘志江, 宋友, 李国宾, 等. 一致性算法和领航跟随法结合的多机编队控制[J]. 航天控制, 2022, 40(6): 46-52.
LIU Z J, SONG Y, LI G B, et al. A Multi-aircraft formation control based on consistency algorithm and pilot following method[J]. Aerospace Control, 2022, 40(6): 46-52(in Chinese).

[17] BOYD S, GHOSH A, PRABHAKAR B, et al. Analysis and optimization of randomized gossip algorithms[C]//Proceedings of the 2004 43rd IEEE Conference on Decision and Control. Piscataway: IEEE Press, 2005: 5310-5315.

[18] DOUGLAS B W. Introduction to graph theory[M]. 2ed. Hoboken: Prentice Hall, 2000.

Design of UAV formation-keeping controller based on improved consistency algorithm

WAN Ru，MA Ziyuan，GONG Huajun，WANG Xinhua*，ZHANG Shuai

(College of Automation Engineering，Nanjing University of Aeronautics and Astronautics，Nanjing 211100，China)

Abstract：This paper conducted in-depth research on the problems of large-gain convergence oscillation and the simultaneous application of translational and rotational motions in traditional consistency algorithms. First, based on the traditional consistent formation algorithm, a virtual leader was introduced, and a new norm was defined to extend the original algorithm, aiming to solve the issue of convergence oscillation caused by large gains in linear consistency algorithms. Second, to address the issue of formation loss during translational and rotational movements, the maximum inter-agent distance was introduced as feedback to the reference controller. This allowed for modification of the reference input trajectory and ensured stable formation tracking during motion. Finally, a joint semi-physical simulation was conducted using the open-source flight control system Ardupilot and the simulation platform AirSim to verify the feasibility of the improved algorithm. The results indicate that the proposed algorithm significantly improves the limitations of traditional consistency protocols and demonstrates good feasibility in practical applications.

Keywords：consistency algorithm；UAV formation；distributed formation；formation keeping；AirSim simulation

Received：2023-08-02；**Accepted**：2024-01-21；**Published Online**：2024-02-05 09：56

URL：link.cnki.net/urlid/11.2625.V.20240202.1847.005

* **Corresponding author**. E-mail：xhwang@nuaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0517

微型直接甲醇燃料电池多级退化参量融合预测

董杰[1]，苏雨临[2]，张大骋[1, 3, *]

(1. 昆明理工大学 信息工程与自动化学院，昆明 650500；2. 北京航空航天大学 软件学院，北京 100191；
3. 云南省绿色能源与数字电力量测及控保重点实验室，昆明 650500)

摘 要：微型直接甲醇燃料电池(μDMFC)膜电极性能会随运行时间和工况衰退，影响其效率和使用寿命。准确的健康状态(SOH)估计与剩余使用寿命(RUL)预测可以确保μDMFC安全运行。输出电压的退化趋势受运行工况影响而波动，传统基于退化趋势回归的预测方法无法准确捕获这种随机变化。在数据驱动和机理模型相结合的思路下，提出基于输出电压和等效电路模型(ECM)的RUL融合预测方法。针对动态工况，引入负载电流这一退化协变量，利用随机过程重构了未来时刻的负载变化，结合ECM参量的退化趋势，实现对输出电压的准确估计和RUL预测，并在中国轻型汽车测试工况(CLTC)下进行验证。实验结果表明：基于多级退化参量的融合预测方法能更好地适应变工况，RUL预测精度、准确度分别为88.18%和85.71%，高于单一退化模型最优值3.27%和14.28%。

关 键 词：微型直接甲醇燃料电池；动态工况；剩余使用寿命；多级退化参量；融合预测
中图分类号：TM911.4
文献标志码：A **文章编号**：1001-5965(2025)10-3567-11

微型直接甲醇燃料电池(micro direct methanol fuel cell, μDMFC)是一种将燃料及氧化剂中的化学能直接转化为电能的装置，具有高效率、低排放、低噪声、易于控制等优点，在航空航天、运输、便携式设备和分布式发电等领域广受青睐[1-2]。然而，电化学反应中产生的一些副产物和副反应会使其性能衰退，降低燃料电池的效率和寿命。因此，为了减少副产物和副反应的影响，研究人员不断改进电极材料和催化剂，优化燃料电池的结构设计和操作条件，以提高燃料电池的稳定性和使用寿命[3]。除了在材料和结构上的改进以外，预测和健康管理(prognostics and health management, PHM)也是提高μDMFC耐久性的一种可行方法。作为PHM的一个关键步骤，剩余使用寿命(remaining useful life, RUL)预测可以估计系统达到故障前的时间，帮助决策者尽早主动避免设备停机[4]。该技术在优化维护策略和提高设备可靠性方面得到了广泛关注[5]。

目前，燃料电池的RUL预测方法主要有基于模型驱动、数据驱动和信息物理融合的预测方法。基于模型驱动的方法是对燃料电池的退化过程进行建模并用于预测，模型的种类可分为退化机理模型、经验模型或半经验模型[6]。机理模型可以帮助研究人员更好地理解燃料电池内部的反应过程，优化设计运行参数，提高燃料电池的效率和可靠性。2016年，Zhang等[7]建立了一个典型的半经验模型，其基于实验数据和经验设计了电池功率自恢复模型，并且基于电化学阻抗谱(electrochemical impedance spectroscopy, EIS)的机理分析结果设置了模型的参数值。2019年，Futter等[8]建立了质子交换膜燃料电池(proton exchange membrane fuel cell, PEMFC)的

收稿日期：2023-08-08；**录用日期**：2023-09-28；**网络出版时间**：2023-10-20 16:34
网络出版地址：link.cnki.net/urlid/11.2625.V.20231020.1308.004
基金项目：国家自然科学基金(62103174)；云南省基础研究计划(202202AD080006)
* **通信作者**. E-mail：dacheng.zhang@kust.edu.cn

引用格式：董杰，苏雨临，张大骋. 微型直接甲醇燃料电池多级退化参量融合预测[J]. 北京航空航天大学学报，2025，51(10)：3567-3577. DONG J，SU Y L，ZHANG D C. Ensemble-based prediction using multi-level degradation parameters for micro direct methanol fuel cells[J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51(10)：3567-3577 (in Chinese).

二维物理模型用于模拟质子交换膜的化学退化过程，并通过加速应力试验数据对模型进行了验证，结果表明该模型可以有效地预测质子交换膜在不同工作压力、湿度和工作电压下的退化速率。数据驱动的预测利用统计学或机器学习等方法从大量燃料电池老化实验数据中学习电池的退化趋势，并利用学习结果预测电池寿命。2014 年，Jouin 等[9]提出了一种基于粒子滤波算法，通过考虑退化状态的连续概率分布实现 RUL 预测。2020 年，Vichard 等[10]提出了一种基于回声状态网络的方法，实现了不同温度条件下对 PEMFC 的退化预测。信息物理融合的预测方法是前两种类型的融合。这种方法以物理方程为基础，通过学习大量数据总结规律来估计参数随时间的变化，既保留模型的解释能力，又充分挖掘数据信息，有针对性地融合多种模型，最大化利用模型互补优势，从而提高预测精度。2014 年，Dong 等[11]建立了支持向量回归-粒子滤波器（support vector regression - particle filter, SVR-PF）的电池健康状态（state of health, SOH）监测模型，引入容量退化参数，以实时确定电池的健康状态。2019 年，Zhang 等[12]指出，当燃料电池处于动态工况条件下运行时，单一退化指标无法单独提供准确和可靠的剩余使用寿命预测，提出了一种基于输出电压和 SOH 的集成预测模型，将 2 种模型的预测结果动态分配权重。结果表明，这种融合预测方法在动态测试环境下优于单一模型。

在实际应用中，输出电压易于实时测量，通常被用作退化指标，但这一外部退化指标易受工况变化影响，导致较大的预测误差；而通过测量 EIS 可以得到燃料电池内部阻抗等内部退化参量，建立退化机理模型，但考虑测量复杂性和成本，只能周期性测量。因此，仅使用单一类别的退化数据，难以实现精确的 RUL 预测。本文在信息物理融合的基础上采用多级退化指标对中国轻型汽车测试循环（China light vehicle test cycle, CLTC）工况下运行的 μDMFC 进行健康状态估计和 RUL 预测。基于等效电路模型（equivalent circuit model, ECM）结合外部监测数据（输出电压）的退化趋势模型（degradation trend model, DTM）通过评估不同模型的预测性能，动态加权融合得到最终预测结果。

1 实验设置

1.1 μDMFC 制备与实验平台

μDMFC 主要由阳极端板、阳极集流板、膜电极组件、阴极端板、阴极集流板和胶垫片等部件构成。两侧端板采用亚克力板加工制成，膜电极由气体扩散层、催化层和质子交换膜组成。阳极端板有一容积为 1 cm^3 的储液腔为电池储存甲醇溶液，阴极端板打通 1 cm^2 方孔，使阴极集流板直接与空气接触。阳极集流板采用厚度为 1.33 mm，栅形沟槽开孔率为 38.5% 的 304 不锈钢，与储液腔连通为甲醇提供流道，收集传导电流。阴极极集流板采用厚度为 1.33 mm，孔形沟槽开孔率为 38.5% 的泡沫不锈钢与空气接触，为阴极提供氧气通道，并且起着收集传导电流、合理分配阴极反应水的作用[13]。在封装之前，用去离子水对电极进行超声波清洗，然后将各组件按图 1 进行组装制备出 μDMFC 单体。

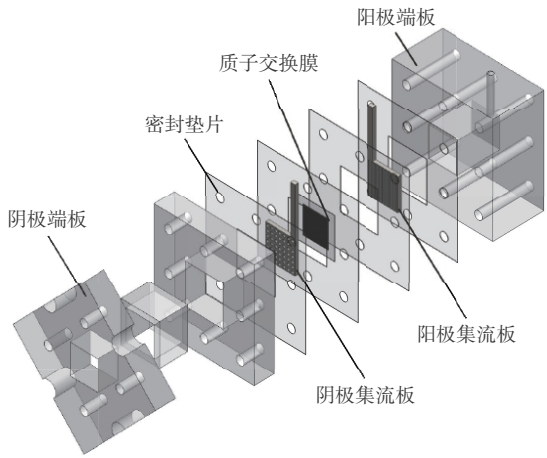


图 1 直接甲醇燃料电池单芯结构

Fig. 1 Structure of μDMFC single cell

将制备好的 μDMFC 加入去离子水，无渗透后放入恒温箱，在 60 ℃ 下处理 3 h，以形成电联。接着在 60 ℃、1 mol/L 甲醇溶液下以 50 mA 的电流放电 1 h，然后以 10 mA 的电流梯度递增放电至电压下降为 0 完成活化，以获得最佳的催化剂活性[14]。完成活化后即可进行测试，测试平台由直流电子负载（IT8511A+）、电化学工作站（CHI660E）和恒温干燥箱（101-0A）组成。其中，直流电子负载用来记录电压、电流数据，电化学工作站用来测试电化学阻抗谱。加速老化测试平台如图 2 所示。

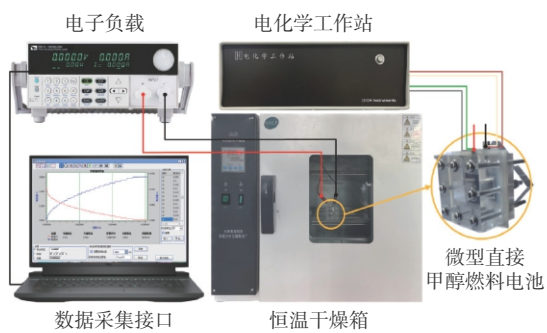


图 2 直接甲醇燃料电池加速老化测试平台

Fig. 2 μDMFC accelerated aging test platform

μDMFC 活化完成后, 对其性能进行检测, 以选出性能最优的单体电池进行后续实验。将活化好的 μDMFC 放置在 60 ℃ 恒温箱内, 测试其在不同甲醇溶液浓度(0.5～4.0 mol/L)下的性能, 如图 3 所示。

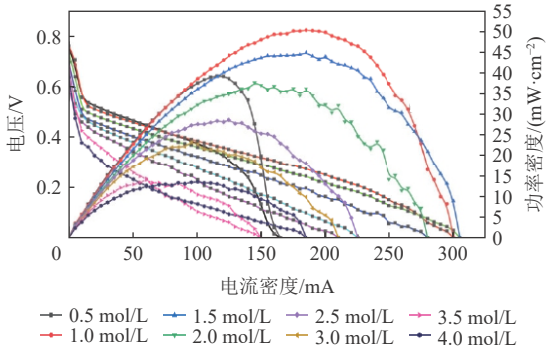


图 3　不同甲醇溶液浓度的输出性能比较

Fig. 3　Output performance comparison with different methanol concentrations

从图 3 中可以看出, 当甲醇浓度为 1 mol/L 时, 电流密度、最大功率和开路电压均取得最大值, 该浓度下燃料电池性能最好, 最大功率密度为 50.35 mW/cm^2。随着甲醇溶液浓度的升高, 由于高浓度甲醇的渗透作用, 电池性能变差, 因此选用 1 mol/L 浓度的甲醇溶液进行后续加速老化实验。

1.2　动态运行工况

实验通过控制负载电流模拟燃料电池汽车在中国轻型汽车测试循环 CLTC 工况下的长期使用情况。如表 1 中所示, 共包括低速(1 部)、中速(2 部)和高速(3 部)3 个速度区间, 1 部为市区工况, 2 部为市郊工况, 3 部为高速工况; 最常见的 4 种运行状态为减速、匀速、加速和怠速, 其中怠速状态并入停机时间[15]。由于单堆燃料电池的功率极限, 在μDMFC 的加速老化测试中, 依据 CLTC 运行工况对负载电流进行同比例缩小, 低速、中速和高速 3 个速度区间共包含 9 种行驶状态的负载电流分别为 0.018 A、0.038 A、0.053 A、0.068 A、0.089 A、0.096 A、0.114 A、0133 A 和 0.138 A, 占比与 CLTC 运行工况一致。在实际的汽车运行过程中, 9 种状态随机发生并时刻变化, 于是采用马尔可夫过程对 9 种负载电流进行随机分配, 使其尽可能服从车辆实际行驶情况。

本文实验采用被动式μDMFC, 甲醇溶液不能被持续供应, 储液腔中的甲醇溶液每次耗尽需要进行换液。图 4 为一次放电周期, 每次持续约为 2 h。

图 4 中, 红色虚线为基于马尔可夫随机过程的 9 种负载电流工况, 蓝色实线为 9 种电流所对应的输出电压。当燃料供应不足时应为电池更换新的甲醇溶液, 以此开始新一轮运行。

表 1　基于 CLTC 工况的负载电流设置

Table 1　Load current settings based on CLTC operating conditions

速度区间	运行状态	占比/%	平均车速/(km·h^{-1})	实验负载电流/A
低速	减速	23.67	11.85	0.018
	匀速	21.36	24.91	0.038
	加速	22.18	35.02	0.053
	怠速	32.79	0	0
中速	减速	26.67	44.52	0.068
	匀速	28.29	55.53	0.089
	加速	27.32	62.49	0.096
	怠速	17.72	0	0
高速	减速	19.56	75.13	0.114
	匀速	57.11	87.71	0.133
	加速	19.78	90.92	0.138
	怠速	3.56	0	0

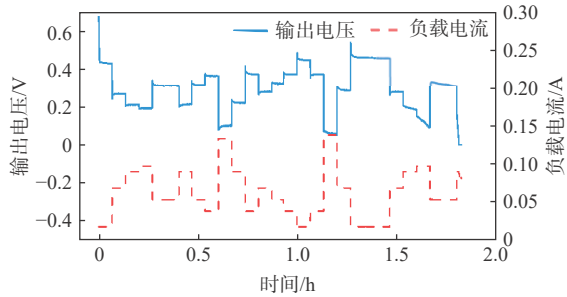


图 4　μDMFC 一次放电循环

Fig. 4　A discharge cycle of μDMFC

1.3　数据描述

μDMFC 平均每天运行 10 h(其余为停机时间), 持续 85 天直至燃料电池失效。共获取了 2 种不同类型的数据。

1.3.1　外部退化指标

输出电压以 1 Hz 采样频率实时测量, 将每次的放电数据整合到一起可以得到μDMFC 全寿命周期的放电数据。累计测得 850 h 的输出电压数据, 每隔 1 h 取 1 个测量值即足够表征出电压的退化趋势, 如图 5 所示。由于该加速老化实验在变工况下进行, 因此电压值在小时间窗内波动较大, 但电压值整体呈下降趋势, 且大电流下的电压下降情况更为明显, 这是由μDMFC 内部膜电极组件老化等不可逆因素导致的。

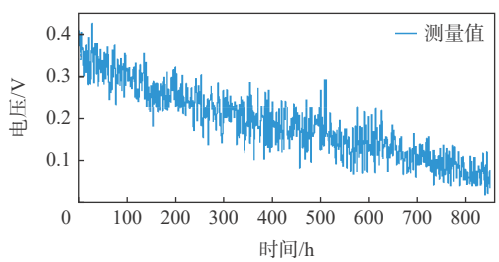


图 5　μDMFC 输出电压

Fig. 5　μDMFC output voltage

1.3.2 内部退化参量

燃料电池在长时间运行过程中内部参数(如电化学反应面积、内部阻抗和催化效率)会受到影响,从而影响输出性能。燃料电池内部阻抗是由电化学反应、传质和电化学反应速率限制等因素引起的,内部阻抗的增加可以直接反映出燃料电池的老化和退化程度。燃料电池中的退化参量可以通过建立恰当的ECM来识别和量化。通过定期停机测量表征燃料电池的内部电化学行为的EIS结合ECM可获取内部阻抗的退化规律。在加速老化实验中,EIS每天停机测量一次,图6展示了其中5个时间点测量的EIS。可以看出,随着运行时间的增加,电池内部的总阻抗随之增大。

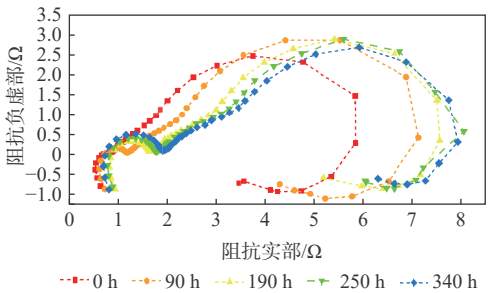


图6 不同时刻EIS测量示例

Fig. 6 EIS measurement examples at different time

2 模型描述与构建

2.1 退化趋势模型

本文选用长短期记忆网络(long short-term memory, LSTM)对电池退化参量的时序变化进行估计,LSTM是递归神经网络(recurrent neural network, RNN)的改进,在对时间序列的预测上,普通的RNN结构当训练深度过大时,由于反向传播链很长,梯度会以指数方式缩小,导致梯度信息消失。为了克服梯度消失问题,使用LSTM预测模型能够有效保留长时间记忆信息,改善循环神经网络的长程依赖问题[16]。

LSTM模块的结构如图7所示,记忆单元在信息的传输与输入的转换中使用的是加法和乘法,加

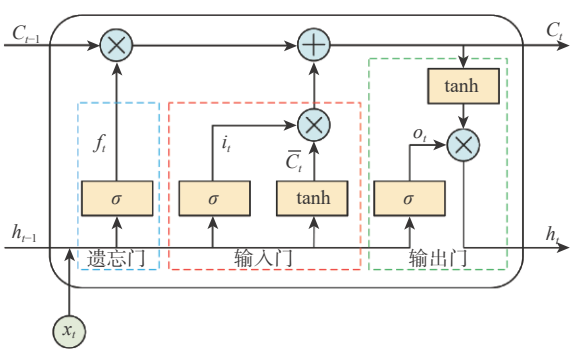


图7 LSTM神经元结构

Fig. 7 LSTM neuron structure

法操作在进行深度反向传播过程中,会保持恒定的误差。在LSTM中,信息流主要通过3种门控信号进行更新:遗忘门$\boldsymbol{f}_t$、输入门$\boldsymbol{i}_t$及输出门$\boldsymbol{o}_t$。

1)遗忘门使用σ函数来决定舍弃或保留前序状态$\boldsymbol{C}_{t-1}$的信息。

$$\boldsymbol{f}_t=\sigma(\boldsymbol{W}_f\left[h_{t-1}\quad x_t\right]+\boldsymbol{b}_f) \tag{1}$$

式中:σ为激活函数;$\boldsymbol{W}_f$和$\boldsymbol{b}_f$分别为对应的权重矩阵和偏置向量;x_t为t时刻退化参量的测量值,其作为神经网络的输入,通过与估计值比较以更新模型参数;h_t为燃料电池各参量退化模型停止更新后的预测值。遗忘门读取隐藏层h_{t-1}(即前一时刻燃料电池各参量的模型估计值)并输入当前时刻的x_t,然后输出一个值在0和1之间(0表示传递的数据完全丢失,1表示数据完全保留)。

2)输入门用来更新单元状态,σ层决定了网络将更新哪些值,tanh激活层创建一个新的估计值$\overline{\boldsymbol{C}}_t$,$\overline{\boldsymbol{C}}_t$表示为

$$\boldsymbol{i}_t=\sigma(\boldsymbol{W}_i\left[h_{t-1}\quad x_t\right]+\boldsymbol{b}_i) \tag{2}$$

$$\overline{\boldsymbol{C}}_t=\tanh(\boldsymbol{W}_C\left[h_{t-1}\quad x_t\right]+\boldsymbol{b}_C) \tag{3}$$

通过激活函数σ的信息与通过激活函数tanh的信息一并输入期望的单元状态$\overline{\boldsymbol{C}}_t$。

3)更新单元状态$\boldsymbol{C}_t$,即退化参量估计值。

$$\boldsymbol{C}_t=\boldsymbol{f}_t\boldsymbol{C}_{t-1}+\boldsymbol{i}_t\overline{\boldsymbol{C}}_t \tag{4}$$

将旧状态$\overline{\boldsymbol{C}}_t$与遗忘门$\boldsymbol{f}_t$相乘,剔除旧状态中需要丢弃的信息,将输入门$\boldsymbol{i}_t$和$\overline{\boldsymbol{C}}_t$的乘积相加,单元状态被更新。

4)输出门用来更新隐含状态h_t即退化参量的预测值。

$$\boldsymbol{o}_t=\sigma(\boldsymbol{W}_o\left[h_{t-1}\quad x_t\right]+\boldsymbol{b}_o) \tag{5}$$

$$h_t=\boldsymbol{o}_t\tanh(\boldsymbol{C}_t) \tag{6}$$

首先,将先前的隐含状态和当前输入传递到σ函数中,然后,新修改的单元状态通过tanh函数与σ函数输出相乘,确定新的隐含状态应保留的信息。

本文实验硬件环境为Intel Core i7-12700,深度学习框架为TensorFlow,网络优化器采用Adam,epoch=32, batch size=64。本文构建的LSTM预测模型算法如下:

算法1 LSTM预测算法。

1: 初始化燃料电池退化参量观测数据集D,迭代时间T,预测步数S

2: 搭建神经网络LSTM: model.add(LSTM)

3: for t< 1, ⋯, T

4: 数据归一化

5: 用数据集D训练LSTM模型model.fit(D)

6: 通过 LSTM 模型得到训练数据 train:

train = model.predict(D)

7: for s< 1, ⋯, S

8: 预测下一时间的参量值:

9: 递归运行步骤 8 获得长度为 S 的预测数据 predict

10: end for

11: end for

12: 计算迭代 T 次后的预测平均值

13: return 训练与预测的平均值

首先, 将输出电压归一化后输入 LSTM 神经网络, 初始化 LSTM 模型以及其权重与偏置, 按照式 (1)～式 (6) 前向计算方法进行前向传播; 在每次更新迭代中, LSTM 递归神经网络将进行反向传播, 在经历了前向传播与反向传播后计算出每个网络的误差项, 根据相应的误差项计算权重梯度, 并通过 Adam 优化算法更新权重, 建立数据的非线性映射关系得到训练好的 LSTM 模型。递归调用该模型进行预测。鉴于退化过程的随机性, 在同一预测时刻 t_p 进行了 10 次反复预测, 本文最终取 10 次预测结果的中值。

2.2　等效电路模型

EIS 可以表征电池内部阻抗的变化情况, μDMFC 的内部损耗主要包括欧姆损耗、活化损耗和传质损耗。这 3 种损耗分别对应着 EIS 中的高频、中频和低频这 3 个频段。典型的燃料电池电化学阻抗谱如图 8 所示。

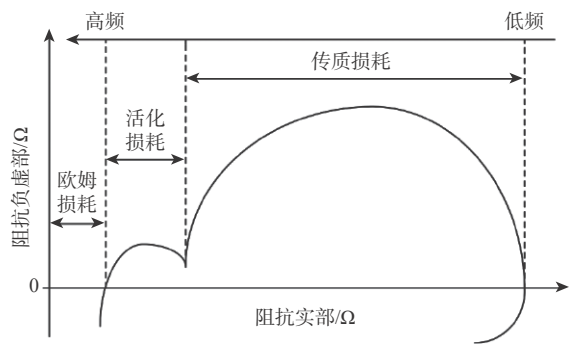


图 8　μDMFC 的典型电化学阻抗谱

Fig. 8　Typical EIS of μDMFC

其中, 欧姆损耗是由膜阻抗、扩散层以及部分催化层的接触阻抗引起的[17-18]; 活化损耗主要是由催化剂的活性引起的; 传质损耗是由于阴极处产生的大量液态水阻碍了空气向载体/催化剂层的传递, 即气体的有限扩散受到限制, 特别是在高电流放电条件下[19]。除了电池的内部阻抗外, EIS 中还反映了电池的外部感抗。除此之外电池内部还有容抗的产生, 因此将阻抗、容抗、感抗全部考虑后建立了如图 9 所示的等效电路模型。

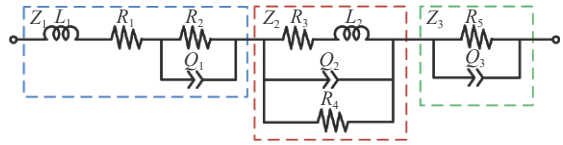


图 9　等效电路模型

Fig. 9　Equivalent circuit model

图 9 中, L_1 为高频电感, 由外部电缆与导线产生, R_1 为极板与膜之间的接触阻抗(欧姆阻抗)[17]; R_2 为膜和催化层之间的接触阻抗, Q_1 为膜和催化层之间的电容行为[18]; R_3 为电荷转移阻抗, R_5 为传质阻抗, Q_2 和 Q_3 表示阳极和阴极双层电容的充放电过程[20]; R_4 用于修正相位延迟, L_2 为低频电感, 由杂质和副产物如 CO 等产生[18,21]。常相位元件 Q 表征因电极表面多孔性分布不均匀所导致的有限区域呈现时间常数分布(阻抗谱变形)的现象[22]。根据 Kim 等[23] 的研究, 电容和电感随着老化几乎不变, 而电阻对老化比较敏感。因此, R_1、R_2、R_3、R_4 和 R_5 被认为是与老化相关的主要参数。根据本文所构建的 ECM, 电池的总内部损耗为

$$Z = Z_1 + Z_2 + Z_3 \tag{7}$$

式中: Z_1 由 L_1、R_1 和第一个 RC 并联电路串联组成表征欧姆损失; Z_2 由 R_3、Q_2 和与 R_4 和 L_2 并联组成表征活化损失; Z_3 由 R_5 和 Q_3 并联组成表征传质损失。由于省略了电感和电容, 电池的内部总阻抗为

$$R = R_1 + R_2 + (R_3R_4)/(R_3 + R_4) \tag{8}$$

式中: R_1 为欧姆阻抗; R_2 为活化阻抗; $(R_3R_4)/(R_3+R_4)$ 为传质阻抗。获得总阻抗的预测值后, 可以计算出电池电压的退化趋势为

$$E(t) = E_r(t) - i(t)R(t) \tag{9}$$

2.3　运行工况重构

根据式 (9), 若负载电流已知, 则可重现输出电压。本文通过马尔可夫状态转移过程构建全寿命周期加速老化测试的负载电流。由于 μDMFC 加速老化测试的负载电流依据 CLTC 运行工况利用马尔可夫过程随机生成的, 那么在未来时刻的负载电流可被视为每个对应的 CLTC 工况电流强度水平具有相同的概率。设负载电流为一组离散随机变量的集合$\{\boldsymbol{X_n}|n=1,2,\cdots\}$, 即马尔可夫链:

$$\boldsymbol{X_n} = X(1), X(2), \cdots, X(n), n = 1, 2, \cdots \tag{10}$$

将位于 n 时刻电流 $X(n)$ 的状态空间标记为 $\boldsymbol{S}$, 即 9 种 CLTC 工况下的模拟负载电流, 因此马尔可夫的状态空间 $\boldsymbol{S}$ 为

$$\boldsymbol{S} = \{i_1, i_2, \cdots, i_9\} \tag{11}$$

其中，负载电流 i_1 到 i_9 分别为：0.018 A、0.068 A、0.114 A、0.038 A、0.089 A、0.133 A、0.053 A、0.096 A 及 0.138 A。

9种负载电流对应 CLTC 工况占比作为概率，其概率转移矩阵 $\boldsymbol{P}$ 为

$$\boldsymbol{P}=(p_1,p_2,\cdots,p_9) \tag{12}$$

根据表1，转移概率 p_1～p_9 分别为 0.117、0.106、0.110、0.108、0.115、0.111、0.068、0.197 及 0.068。其中，每4种状态之间相互转移及其概率的马尔可夫链如图10所示。

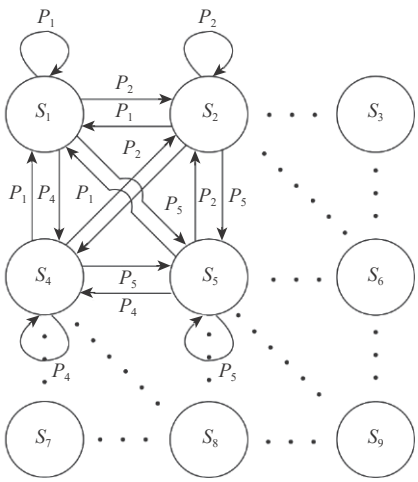


图10 负载电流重构的马尔可夫链

Fig. 10 Markov chain-based load current reconstruction

本文运用马尔可夫过程重构燃料电池的工况电流，由于放电后期，电池性能已部分退化，能达到的最大电流也逐渐降低，图10中展示的实际电流最大只能达到 0.114 A 左右。负载电流在 600 h 时刻的运行工况的重构结果如图11所示。

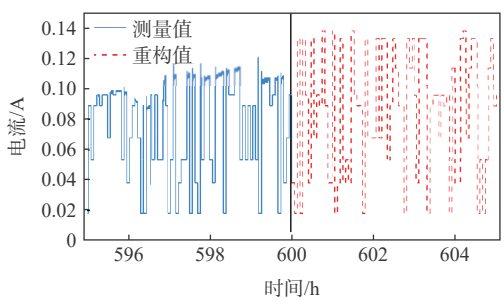


图11 t_p= 600 h 时刻的重构电流

Fig. 11 Current reconstruction at t_p=600 h

3 剩余使用寿命预测

μDMFC 的性能随着运行时间的下降，其相关退化指标参数也会随之变化，对 RUL 预测的问题实际上是估计当前时刻到燃料电池系统失效时刻（end of life, EOL）的运行时间 E_{OL}。如图12所示，其剩余使用寿命 RUL 定义为从预测时刻 t_p 到 EOL 之间的时间间隔 R_{UL}：

$$R_{\mathrm{UL}}=E_{\mathrm{OL}}-t_p \tag{13}$$

由于电压值容易测量且可以直接反映电池的性能，当输出电压值降低到失效阈值（failure threshold, FT）时视为燃料电池失效。本文以 μDMFC 的输出电压表征电池的退化状态，失效阈值为电压初始值的 10%[24]，即 μDMFC 输出电压随运行时间下降至 0.040 8 V 时，认为其失效。

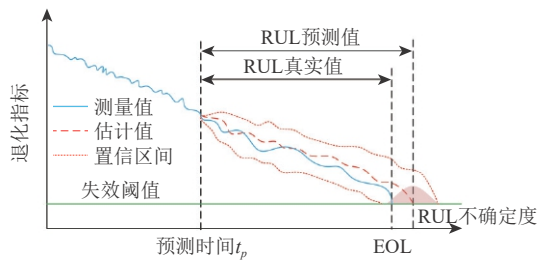


图12 RUL 预测原理

Fig. 12 RUL prediction illustration

3.1 基于退化趋势的 RUL 预测

通过前面描述的神经网络技术对 μDMFC 输出电压退化趋势进行了估计。图13展示了 600 h 时刻的输出电压下降趋势预测结果，RUL 预测值为 180 h（EOL 为 780 h），真实 RUL 为 239 h（EOL 为 839 h），相对误差为 24.7%。

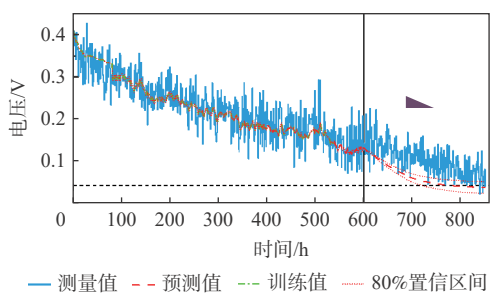


图13 t_p=600 h 时刻基于外部电压的 RUL 预测

Fig. 13 DTM-based RUL prediction at t_p=600 h

从图13中可看出，预测阶段单纯依靠退化趋势的回归拟合无法适应工况变化带来的电压波动。

3.2 基于等效电路模型的 RUL 预测

按照前面构建的 ECM 通过最小二乘法非线性拟合方法从 EIS 中识别模型参数得到 ECM 中各个电路元器件的值。图14展示了 270 h 的拟合示例，其中蓝色的点是实测得的 EIS，红色线条是拟合曲线。

不同时间点的 ECM 参数识别结果在表2中列出。其中，R_1、R_2、R_3、R_4、R_5 可以反映 μDMFC 内部电阻的变化。

从表2中可以看出，R_1、R_2、R_3、R_4 和 R_5 的阻值

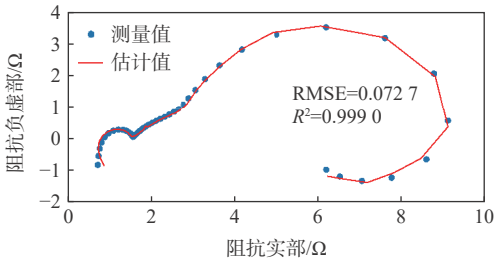


图 14　电化学阻抗谱拟合示例

Fig. 14　EIS fitting example

表 2　不同时间的 ECM 模型参数识别结果示例

Table 2　ECM parameters identification results at different time

时间/h	E_r/V	R_1/Ω	R_2/Ω	R_3/Ω	R_4/Ω	R_5/Ω
0	0.724	0.531	0.280	4.530	1.940	1.290
90	0.717	0.590	1.220	5.250	2.720	0.551
190	0.705	0.615	1.010	5.560	2.660	1.460
250	0.678	0.635	0.719	5.920	3.000	1.680
340	0.639	0.620	1.550	6.490	3.340	1.770

随着老化而呈现上升的趋势, 最大开路电压 E_r 随时间逐渐下降。将基于神经网络的预测方法应用于参数随时间变化的趋势估计, 退化参量 E_r、R_1、R_2、R_3、R_4 和 R_5 的估计和预测示例如图 15 所示。

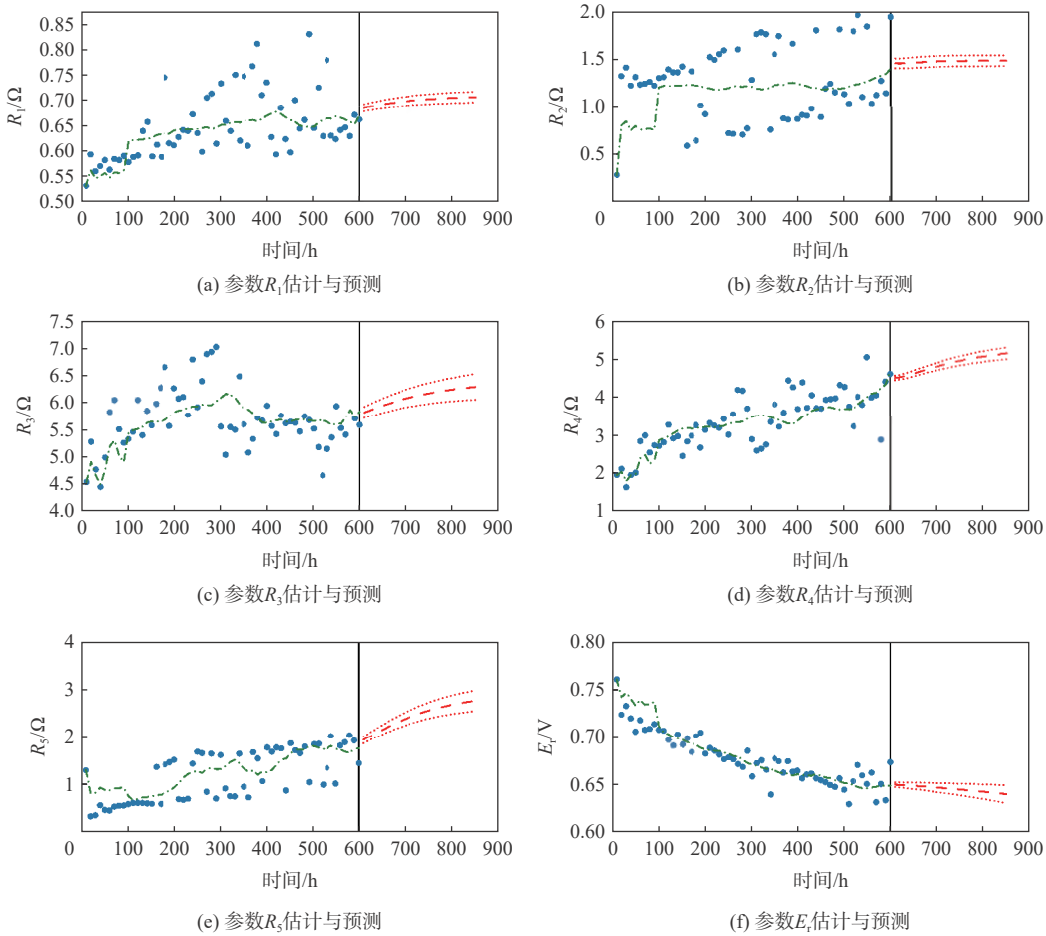


图 15　t_p=600 h 时刻基于电化学阻抗谱的参量估计与预测

Fig. 15　EIS-based covariate estimation at t_p=600 h

将 ECM 退化参量 R_1、R_2、R_3、R_4、R_5 的预测结果代入式 (9) 可以重构电压的退化趋势。图 16 为 600 h 的预测示例, RUL 的预测值为 193 h, 真实值为 239 h, 与真实值相差 46 h, 相对误差为 19.2%。

与基于外部退化指标(输出电压)的预测结果相比, 等效电路模型的预测方法可以更准确地描述 μDMFC 的退化趋势, 能更好地适应动态工况带来的电压波动情况, 预测结果更加接近实际电压的波

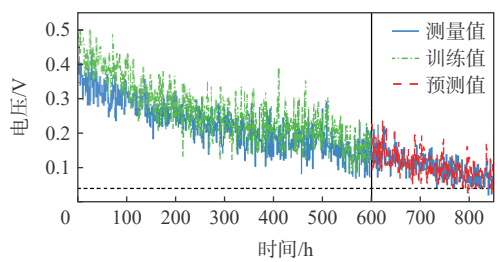


图 16　t_p=600 h 时刻基于等效电路模型的 RUL 预测

Fig. 16　ECM-based RUL prediction at t_p=600 h

动趋势。

3.3 模型融合预测

不同模型的预测具有不同的优势，5 基于外部退化指标(输出电压)的 RUL 预测快捷便利，只需要通过观察输出电压的退化趋势，不需要构建复杂的模型便可粗略地估计μDMFC 的 RUL；而基于 ECM 内部退化参量的预测方法可较准确地描述μDMFC 的退化特征，应对工况变化。因此，本文探究将 2 种模型相结合的预测方法，充分利用两者优势进一步提高预测性能。模型融合预测技术路线如图 17 所示。

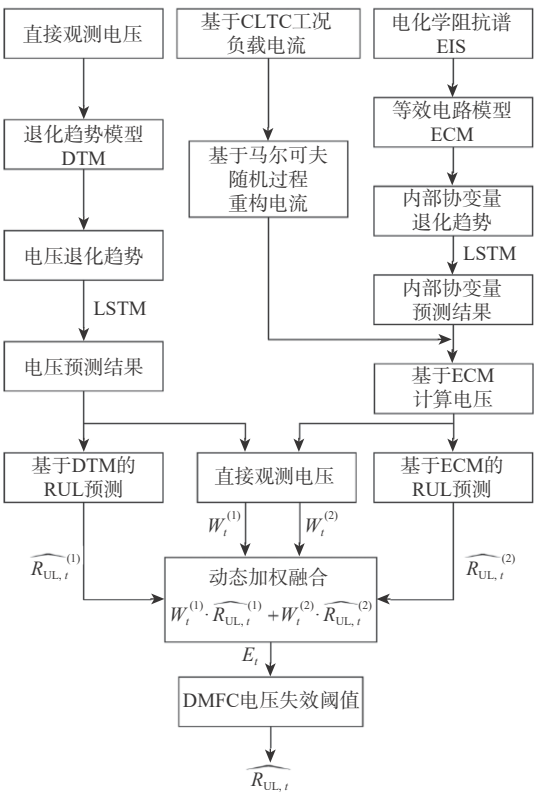


图 17 基于模型融合预测的技术路线

Fig. 17 Model ensemble-based prediction framework

通过外部退化指标(输出电压)可以预测电压的下降趋势；通过 ECM 预测内部阻抗退化趋势，可进行输出电压重构。所有的时间序列预测均采用 LSTM 技术实现，最后将不同模型的预测结果通过动态加权进行融合。

在模型的学习过程中，根据训练误差评估模型性能：

$$\varepsilon_{t_p}=\sum_{t=1}^{t_p}\frac{|x_t^*-x_t|}{x_t^*}\Big/t_p \tag{14}$$

式中：x_t^*为测量得到的每一个时刻的真实值；x_t为模型在对应时刻的训练值；t_p为预测时刻的时间步长；ε_{t_p}为预测模型在 1 到预测时刻 t_p 累积训练误差的平均值。

通过计算出的每个时间点的误差，按照训练模型的准确度对模型在融合时所占的权重进行分配，误差较低则在融合时权重大。融合权重如下：

$$w_{t_p}^{(m)}=1/\varepsilon_{t_p}^{(m)}\Big/\sum_{m=1}^{M}1/\varepsilon_{t_p}^{(m)} \tag{15}$$

式中：M为模型的数量(M=2)；m为不同的模型；$w_{t_p}^{(m)}$为局部权重，值是非负的，且总和为 1。

采用上述原则进行权重分配的融合方法，将基于外部退化指标(模型 1)的电压预测与等效电路模型(模型 2)的电压预测融合。图 18 展示了 600 h 时刻基于外部退化指标(输出电压)与基于等效电路模型 2 种预测方法的融合结果。

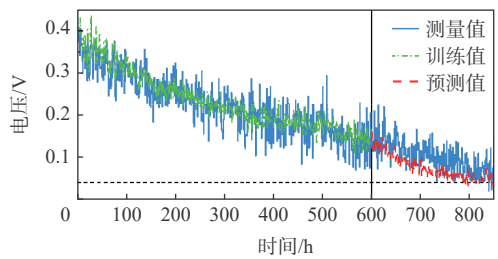


图 18 t_p=600 h 时刻模型 1 与模型 2 融合的 RUL 预测

Fig. 18 Ensemble-based RUL prediction (Model 1 and Model 2) at t_p=600 h

在图 18 中展示了 2 种模型融合的预测结果真实的剩余使用寿命为 239 h，预测值为 193 h。

为了进一步评估预测性能，从 t_p=400 h 开始，每 20 h 做一次 RUL 预测，直到燃料电池失效时刻。图 19 展示了 400～800 h(间隔 20 h)基于各模型的 RUL 预测结果。

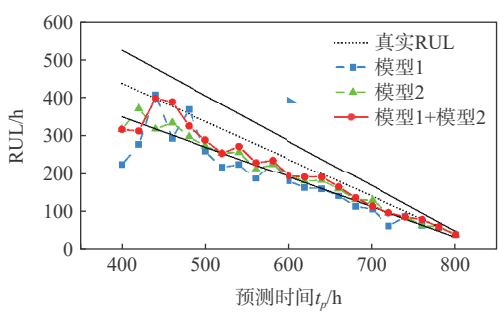


图 19 不同模型的 RUL 预测结果

Fig. 19 RUL prediction results by different models

黑色虚线是通过式 (13) 计算的真实 RUL。2 条实线形成了一个逐渐变窄的 80% 置信区间。可以看出，基于单一模型预测的 RUL 有一部分时间点的预测结果落在了置信区间外而 2 个模型融合的预测结果大多落在置信区间内且更接近真实 RUL，这表明在预测的准确度上，基于多级退化参量的融合预测方法具有更好的预测性能。

3.4　预测性能评价

为了评估不同模型的 RUL 的预测性能，使用领域内常用的评价指标[12]对各模型的 RUL 整体预测性能进行量化。

1）精度指标A_{cc}是建立在预测值R_{UL}和真实燃料电池剩余使用寿命R^*_{UL}的相对误差基础上的精度指标：

$$A_{cc,t}=1-\frac{\left|R^*_{UL,t}-\widehat{R_{UL,t}}\right|}{R^*_{UL,t}} \tag{16}$$

式中：t为预测时刻；$R^*_{UL,t}$为时刻t的真实值；$\widehat{R_{UL,t}}$为时刻t的预测值。通过式 (16) 的计算可以得到预测方法预测结果的相对精度。该精确度值越大表示预测值和真实值之间的误差越小。

2）准确度α_{A_c}表示预测值落在置信区间中的个数在所有预测中的占比，反映了预测准确度的总体情况：

$$\alpha_{A_c}=\frac{N_{in}}{N_{all}} \tag{17}$$

式中：N_{in}为预测值落在置信区间中的个数；N_{all}为总的预测数，该值越大表明预测处在合理范围内的值越多，其不确定度控制越好。

由于 RUL 预测的不确定性，本文从 400 h 开始，每隔 20 h 进行一次预测。3 种方法在不同时间点下预测所需时长，如图 20 所示。各模型 RUL 预测评价结果在表 3 中列出。

通过图 20 和表 3 可以发现，针对动态工况，融合预测方法的平均预测时长略高于模型 2，但其预测精度和准确度均比单一模型高，兼顾了 2 种模型的优点。

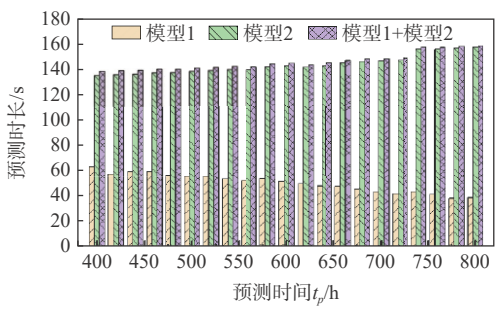


图 20　不同模型预测时长比较

Fig. 20　Time complexity comparison of different models

表 3　各模型 RUL 预测性能评价结果

Table 3　Prognostic performance evaluation for different models

模型	A_{cc}/%	α_{A_c}/%	平均预测时长/s
模型1	77.87	33.33	50.20
模型2	84.91	71.43	143.91
模型1与模型2融合	88.18	85.71	146.11

4　结　论

本文针对动态工况下μDMFC 的剩余使用寿命预测问题，提出了基于多退化指标融合预测的方法。

1）使用回归拟合直接对μDMFC 输出电压的预测容易实现，但因为输出电压的退化趋势受工况影响导致预测精度不高。

2）使用等效电路模型将μDMFC 退化过程分解为多个退化参量共同作用的结果后再反应到输出电压退化上，能够应对μDMFC 运行中的工况（负载电流）变化，从而获得更高的预测精度。使用加权融合的方法充分利用两者的优势，可进一步提高预测的性能。

3）在 CLTC 工况下的实验结果表明，基于多级退化参量的融合预测方法能更好地适应变工况环境，RUL 预测精度、准确度分别达到 88.18%、85.71%，高于单一退化模型最优值 3.27%、14.28%。

参考文献（References）

[1] JIAO K, XUAN J, DU Q, et al. Designing the next generation of proton-exchange membrane fuel cells[J]. Nature, 2021, 595(7867): 361-369.

[2] MARAPPAN M, PALANISWAMY K, VELUMANI T, et al. Performance studies of proton exchange membrane fuel cells with different flow field designs–review[J]. The Chemical Record, 2021, 21(4): 663-714.

[3] ESMAEILI N, GRAY E M, WEBB C J. Non - fluorinated polymer composite proton exchange membranes for fuel cell applications: a review[J]. ChemPhysChem, 2019, 20(16): 2016-2053.

[4] LEE J, WU F J, ZHAO W Y, et al. Prognostics and health management design for rotary machinery systems: reviews, methodology and applications[J]. Mechanical Systems and Signal Processing, 2014, 42(1-2): 314-334.

[5] LI X, ZHANG W, DING Q. Deep learning-based remaining useful life estimation of bearings using multi-scale feature extraction[J]. Reliability Engineering & System Safety, 2019, 182: 208-218.

[6] 刘浩. 质子交换膜燃料电池的寿命预测研究[D]. 杭州: 浙江大学, 2019: 1-14.
LIU H. Study on life prediction of proton exchange membrane fuel cell[D]. Hangzhou: Zhejiang University, 2019: 1-14 (in Chinese).

[7] ZHANG D, CADET C, BÉRENGUER C, et al. Some improvements of particle filtering based prognosis for PEM fuel cells[J]. IFAC-PapersOnLine, 2016, 49(28): 162-167.

[8] FUTTER G A, LATZ A, JAHNKE T. Physical modeling of chemical membrane degradation in polymer electrolyte membrane fuel cells: influence of pressure, relative humidity and cell voltage[J]. Journal of Power Sources, 2019, 410: 78-90.

[9] JOUIN M, GOURIVEAU R, HISSEL D, et al. Prognostics of PEM fuel cell in a particle filtering framework[J]. International Journal of Hydrogen Energy, 2014, 39(1): 481-494.

[10] VICHARD L, HAREL F, RAVEY A, et al. Degradation prediction of PEM fuel cell based on artificial intelligence[J]. International Journal of Hydrogen Energy, 2020, 45(29): 14953-14963.

[11] DONG H C, JIN X N, LOU Y B, et al. Lithium-ion battery state of health monitoring and remaining useful life prediction based on support vector regression-particle filter[J]. Journal of Power Sources, 2014, 271: 114-123.

[12] ZHANG D, BARALDI P, CADET C, et al. An ensemble of models for integrating dependent sources of information for the prognosis of the remaining useful life of proton exchange membrane fuel cells[J]. Mechanical Systems and Signal Processing, 2019, 124: 479-501.

[13] ZHAO Z G, ZHANG F, ZHANG Y H, et al. Performance optimization of μDMFC with foamed stainless steel cathode current collector[J]. Energies, 2021, 14(20): 6608.

[14] 薛瑞. 微型直接甲醇燃料电池高浓度传质阻挡层技术研究[D]. 哈尔滨: 哈尔滨工业大学, 2019: 46-62.
XUE R. Research on the mass transfer barrier layer of concentrated direct methanol fuel cell[D]. Harbin: Harbin Institute of Technology, 2019: 46-62(in Chinese).

[15] LIU Y, WU Z X, ZHOU H, et al. Development of China light-duty vehicle test cycle[J]. International Journal of Automotive Technology, 2020, 21(5): 1233-1246.

[16] 王鑫, 吴际, 刘超, 等. 基于 LSTM 循环神经网络的故障时间序列预测[J]. 北京航空航天大学学报, 2018, 44(4): 772-784.
WANG X, WU J, LIU C, et al. Exploring LSTM based recurrent neural network for failure time series prediction[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(4): 772-784(in Chinese).

[17] PASTOR-FERNÁNDEZ C, DHAMMIKA WIDANAGE W, MARCO J, et al. Identification and quantification of ageing mechanisms in Lithium-ion batteries using the EIS technique[C]//Proceedings of the 2016 IEEE Transportation Electrification Conference and Expo. Piscataway: IEEE Press, 2016: 1-6.

[18] YUAN T, ZOU Z Q, CHEN M, et al. New anodic diffusive layer for passive micro-direct methanol fuel cell[J]. Journal of Power Sources, 2009, 192(2): 423-428.

[19] SØRENSEN T S, KJELSTRUP S. A simple Maxwell-Wagner-Butler-Volmer approach to the impedance of the hydrogen electrode in a Nafion fuel cell[J]. Journal of Colloid and Interface Science, 2002, 248(2): 355-375.

[20] ZHU Y L, GAO L, LI J Y. A novel button-type micro direct methanol fuel cell with graphene diffusion layer[J]. Micromachines, 2019, 10(10): 658.

[21] YUAN T, YANG J, WANG Y L, et al. Anodic diffusion layer with graphene-carbon nanotubes composite material for passive direct methanol fuel cell[J]. Electrochimica Acta, 2014, 147: 265-270.

[22] JIANWEI G, ZONGQIANG M, JINGMING X. Studies on the electrochemical behavior of polymer electrolyte membrane fuel cell (PEMFC) by AC impedance method[J]. Chemical Journal of Chinese Universities, 2003, 24(8): 1477-1481.

[23] KIM T, KIM H, HA J, et al. A degenerated equivalent circuit model and hybrid prediction for state-of-health (SOH) of PEM fuel cell[C]//Proceedings of the 2014 International Conference on Prognostics and Health Management. Piscataway: IEEE Press, 2014: 1-7.

[24] 苏雨临, 连冠, 张大骋. 等效电路模型法预测动态工况下微型直接甲醇燃料电池剩余使用寿命[J]. 上海交通大学学报, 2024, 58(10): 1525-1532.
SU Y L, LIAN G, ZHANG D C. Equivalent circuit model-based prognostics for micro direct methanol fuel cell under dynamic operating conditions[J]. Journal of Shanghai Jiao Tong University, 2024, 58(10): 1525-1532(in Chinese).

Ensemble-based prediction using multi-level degradation parameters for micro direct methanol fuel cells

DONG Jie[1]，SU Yulin[2]，ZHANG Dacheng[1, 3, *]

(1. Faculty of Information Engineering and Automation，Kunming University of Science and Technology，Kunming 650500，China；
2. School of Software，Beihang University，Beijing 100191，China；
3. Yunnan Key Laboratory of Green Energy，Electric Power Measurement Digitalization，Control and Protection，Kunming 650500，China)

Abstract: The deterioration of the membrane electrode assembly in micro direct methanol fuel cells (μDMFC) limits cell efficiency and lifespan. Accurate prediction of the state of health (SOH) and remaining useful life (RUL) is essential for ensuring the safe and reliable operation of μDMFCs in industrial applications. The degradation trend of the output voltage fluctuates depending on the operating conditions. Traditional trend regression methods, however, are inadequate for capturing such stochastic fluctuations. Therefore, an RUL ensemble-based prediction method based on the output voltage and equivalent circuit model (ECM) was proposed under a combination of data-driven and mechanistic models. The degradation covariate of load current was introduced to account for dynamic operating conditions. The load changes in the future were reconstructed using a random process, and combined with the degradation trend of ECM parameters, the accurate estimation of the output voltage and RUL prediction were achieved. The proposed approach was validated using accelerated aging tests under the China light vehicle test cycle (CLTC). Experimental results show that the ensemble-based method can adapt to varying operating conditions. RUL prediction accuracy and precision are 88.18% and 85.71%, respectively, outperforming the best individual model by 3.27% and 14.28%.

Keywords: micro direct methanol fuel cell; dynamic operating conditions; remaining useful life; multi-level degradation parameter; ensemble-based prediction

Received：2023-08-08；**Accepted**：2023-09-28；**Published Online**：2023-10-20 16：34

URL：link.cnki.net/urlid/11.2625.V.20231020.1308.004

Foundation items：National Natural Science Foundation of China (62103174); Yunnan Fundamental Research Projects (202202AD080006)

* **Corresponding author**. E-mail：dacheng.zhang@kust.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0533

一种考虑滑流效应的螺旋桨设计方法及应用

陈胜久，杨佑绪*，张兴翠，吴逸飞
(南昌航空大学 飞行器工程学院，南昌 330063)

摘　　要：分布式电推进飞机大量采用螺旋桨，存在显著的桨-翼气动耦合效应。通过改变轴向诱导速度分布控制螺旋桨滑流，从而得到螺旋桨弦长和扭转角分布，提出一种考虑滑流效应的高效螺旋桨设计方法，并验证方法的可行性。计算所设计螺旋桨在孤立和分布式构型的巡航和悬停气动特性，并和最小诱导损失方法进行对比。结果表明：巡航升力系数相同时，所设计螺旋桨效率比最小诱导损失法高3.4%～6.6%；不同转速时，单独螺旋桨悬停效率比最小诱导损失法高10.4%～13.5%；分布式构型，设计螺旋桨悬停效率为比最小诱导损失法高13%。设计螺旋桨在巡航和悬停状态都能满足其设计要求，且都保持较高的效率运行。

关 键 词：滑流效应；分布式电推进；倾转螺旋桨；电动垂直起降；推进力/气动力耦合
中图分类号：V211.44
文献标志码：A　　**文章编号**：1001-5965(2025)10-3578-11

分布式电推进(distributed electric propulsion, DEP)[1-2]是由电机驱动分布在机翼或机身上的多个螺旋桨为飞机提供动力，充分利用推进力-气动力耦合效应减小机翼面积，增加升阻比，提高飞机经济性的一项新技术。其往往采用倾转螺旋桨，通过螺旋桨在一定角度范围内倾转实现飞机地面垂直起降、空中悬停和高速巡航前飞等多状态飞行，目前被新兴的电动垂直起降(electric vertical takeoff and landing, eVTOL)飞行器所广泛采用，如Archer的Midnight、Joby的S4，吉利沃飞的AE200、时的科技的E20、君腾航空的QJ50等。随着飞机能源种类从油动转变为电动，其动力系统的效率问题尤其显著。

相比于传统飞机，分布式电推进飞机的螺旋桨和其他部件之间存在复杂的气动干扰作用，严重影响动力系统的效率，国内外学者在这方面展开了许多相关的研究。Patterson等[3]、Patterson[4]发现螺旋桨诱导的近均匀轴向速度剖面可以对下游机翼产生最有利的影响。Droandi等[5]研究悬停状态2个并排无人机螺旋桨尾迹的相互作用，发现转子的推力系数与螺旋桨分离距离有关，随着分离距离减小，推力波动显著增大。Shukla等[6]研究一组并排旋翼的空气动力学特性，随着转子间距和雷诺数减少，转子间尾迹干扰效应增强，转子效率下降。Stokkermans等[7]研究并排桨和前后桨2种布局方式在悬停、过渡和巡航状态的螺旋桨相互作用。Young等[8]研究了旋翼-机翼距离、机翼后掠角、翼展和襟副翼迎角对倾转旋翼入流和性能的影响。对于螺旋桨与其他部件间的气动耦合效应，螺旋桨间的耦合效应会降低螺旋桨的气动性能。而桨翼间的耦合效应产生有利的影响，螺旋桨的滑流会显著增加下游机翼表面流动速度，直接影响飞机有效迎角[9]，改变机翼表面升力分布。通过合理的螺旋桨气动设计，充分利用桨翼间的有利耦合，从而提高螺旋桨在分布式电推进飞机的效率。

国内外学者对螺旋桨的气动设计方面进行了

收稿日期：2023-08-17；**录用日期**：2023-11-16；**网络出版时间**：2023-12-08 10:17
网络出版地址：link.cnki.net/urlid/11.2625.V.20231206.1445.006
基金项目：江西省“双千计划”人才项目(CK202006470)
*通信作者. E-mail：zgdy_1@163.com

引用格式：陈胜久，杨佑绪，张兴翠，等. 一种考虑滑流效应的螺旋桨设计方法及应用[J]. 北京航空航天大学学报，2025，51(10)：3578-3588. CHEN S J，YANG Y X，ZHANG X C，et al. A propeller design method considering slipstream effect and its application[J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51(10)：3578-3588 (in Chinese).

大量研究。Patterson 等[10]、Xue 等[11] 通过设计螺旋桨滑流提高机翼在低速状态的气动效率, 结果表明均匀螺旋桨滑流能提高机翼气动效率。范中允等[12] 基于片条理论逆向推导最佳环量分布, 建立以环量分布输入的螺旋桨快速设计方法。郭佳豪等[13] 通过改变螺旋桨径向拉力分布控制螺旋桨滑流, 提出一种给定拉力分布的螺旋桨设计方法, 发现拉力分布较为和缓的螺旋桨有利于增升减阻。D'Angelo[14] 提出了给定设计工况下设计最大效率螺旋桨的弦长和桨距角的数值方法, 项松等[15] 对 D'Angelo 方法进行改进, 提出一种迭代优化桨叶弦长分布的高效螺旋桨设计方法, 所设计螺旋桨的巡航效率超过 83%。唐伟等[16] 研究兼顾垂直起降和巡航前飞构型的螺旋桨设计问题, 与单设计点的螺旋桨进行对比, 发现 2 个设计点的动力系统整体效率下降。这些研究提出了螺旋桨设计方法, 但都只是研究了单螺旋桨的气动性能, 并且未结合机翼进行分布式构型的验证分析。

本文提出一种考虑滑流效应的螺旋桨设计方法, 通过不同精度计算方法验证螺旋桨性能, 并与最小诱导损失设计方法对比, 研究孤立和分布式构型螺旋桨在巡航和悬停工况的气动性能。

1　数值方法

1.1　螺旋桨快速设计方法

通过给定螺旋桨巡航条件, 包括入流速度 V、转速 N、螺旋桨半径 R、桨叶叶片数 B、空气密度 ρ、推力 T 等 6 个设计变量, 基于叶素动量理论[17] 和涡流理论[18], 确定螺旋桨轴向诱导因子 a 沿径向分布, 从而控制螺旋桨的滑流轴向速度大小, 并快速迭代得到螺旋桨的几何外形参数弦长 c 和扭转角 β 分布。

同时在计算扭转角分布时采用的方法是在设计前就给定桨叶截面的气动翼型, 因此需要先提供当地截面翼型的有效攻角 α。为简化设计, 将各截面的有效攻角设为该翼型最大升阻比所对应的攻角, 这可以保证弦长光滑, 通过这种方法设计的螺旋桨只是弦长和扭转角分布不一样, 并不会对诱导速度分布有影响[19]。

1.1.1　控制滑流速度

根据叶素动量理论, 得到推力与轴向诱导因子之间的关系, 如下式。螺旋桨桨叶任意截面径向站位 r 处的气流速度分量与攻角的关系如图 1 所示, α 为截面叶素的有效攻角, φ 为截面叶素几何入流角, Ω 为截面叶素的角速度, W 为截面叶素的入流速度, V_a 和 V_t 分别为轴向和切向诱导速度, a 和 b 分别为轴向和切向诱导速度因子。

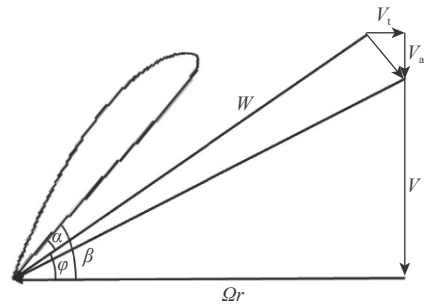


图 1　叶素速度三角形

Fig. 1　Blade element velocity triangle

$$T = 2\pi r^2 \rho V^2 (1+a)a \tag{1}$$

$$a = \frac{V_a}{V} \tag{2}$$

$$b = \frac{V_t}{\Omega r} \tag{3}$$

根据叶素动量理论有以下等式:

$$V^2(1+a)a = \Omega^2 r^2 (1-b)b \tag{4}$$

化简得

$$b = \frac{1-\sqrt{1-\dfrac{4V^2(1+a)a}{\Omega^2 r^2}}}{2} \tag{5}$$

$$a = \frac{-1+\sqrt{1+\dfrac{4\Omega^2 r^2 (1-b)b}{V^2}}}{2} \tag{6}$$

所以当轴向诱导因子 a 确定之后, 切向诱导因子 b 也会随之确定。但根据式 (5) 可知切向诱导因子 b 最大值为 0.5, 故当式 (5) 失效时需要采取式 (6) 来确定诱导因子。

1.1.2　几何参数分布

在确定轴向和切向诱导因子分布后, 可以通过式 (7) 计算得到径向站位 r 处的局部入流角 φ, 翼型攻角 α 在设计前已经给定, 再根据式 (8) 确定扭转角 β。

$$\varphi = \tan^{-1}\left(\frac{V(1+a)}{\Omega r(1-b)}\right) \tag{7}$$

$$\beta = \varphi + \alpha \tag{8}$$

由于叶尖损失, 螺旋桨叶尖不会承受太多的载荷, 同时这还会降低叶尖的速度。增大叶尖载荷可以通过增加叶尖附近的轴向诱导因子, 将修正的 Prandtl 叶尖损失系数 F[17] 应用于轴向诱导因子可以达到所需效果, 如式 (9) 所示, 其中 f 可由式 (10) 计算得到。

$$F = \frac{2}{\pi}\cos^{-1}(\mathrm{e}^{-f}) \tag{9}$$

$$f=\frac{B}{2}\frac{R'-r}{r\sin\varphi} \tag{10}$$

$$a'=\frac{a}{F} \tag{11}$$

式中：B为螺旋桨桨叶叶片数；R'通常使用的是螺旋桨半径R；φ为径向站位r处入流角，由式(10)得到的入流角是一个更新迭代的值，对于第一次迭代，假设所有的径向站位的F值为1。修正轴向诱导因子a'通过式(11)计算得到。

由于叶尖损失系数F在叶尖附近非常小，如果将传统形式的F直接应用于式(11)，将会使叶尖附近的轴向诱导因子增大，这可能导致叶片尖端附近得到不切实际的弦长和扭转角。因此，在计算式(10)时R'不能直接使用螺旋桨的半径R，R'值应该略大于R，计算得到的弦长和扭转角满足设计要求并且是合理的。

根据叶素理论和动量理论，由式(12)和式(13)这2个单独的推力T的表达式，得到局部弦长c的表达式(14)。在式(12)中的叶尖损失系数F时，输入式(10)中R'的值应为螺旋桨半径R。

$$\mathrm{d}T=4\pi r\rho V^2(1+a)aF\mathrm{d}r \tag{12}$$

$$\mathrm{d}T=\frac{B}{2}\rho W^2[C_\mathrm{l}\cos\varphi-C_\mathrm{d}\sin\varphi]c\mathrm{d}r \tag{13}$$

$$c=\frac{8\pi rV^2(1+a)F}{B[(\Omega r(1-b))^2+(V(1+a))^2][C_\mathrm{l}\cos\varphi-C_\mathrm{d}\sin\varphi]} \tag{14}$$

式中：C_l和C_d分别为截面叶素的升力系数和阻力系数。迭代计算得到螺旋桨各个径向点处的扭转角β和弦长c，其中迭代收敛条件是入流角φ两次计算差值小于0.5°。

1.1.3 螺旋桨空气动力学特性

螺旋桨悬停状态下的拉力系数C_T、功率系数C_P和悬停效率F_M分别为

$$C_T=\frac{T}{\rho\pi n^2R^4}C_P=\frac{P}{\rho\pi n^3R^5} \tag{15}$$

$$F_\mathrm{M}=C_T\sqrt{C_T/2.0}\Big/C_P \tag{16}$$

巡航状态下的拉力系数C_T、扭矩系数C_Q、功率系数C_P、螺旋桨效率η分别为

$$C_T=\frac{T}{\rho n^2D^4}C_Q=\frac{Q}{\rho n^2D^5}C_P=\frac{P}{\rho n^3D^5} \tag{17}$$

$$\eta=\frac{TV_\infty}{2\pi nQ} \tag{18}$$

式中：n为螺旋桨转速，其中$n=N/60$；D为螺旋桨直径；Q为扭矩；P为功率。

1.2 螺旋桨滑流快速计算方法

为了在螺旋桨早期设计阶段快速计算螺旋桨滑流，建立螺旋桨尾迹模型获得诱导速度分布情况。螺旋桨滑流的诱导速度场是不均匀的，且计算复杂。本文采用Veldhuis[20]提出的滑流管模型，该模型将螺旋尾迹简化成2个连续的涡度分布：轴向涡和切向涡，如图2所示。轴向涡平行于旋转轴，分布在同心放置的圆柱管上；切向涡分布在相同的圆柱表面上，但垂直于轴向涡。

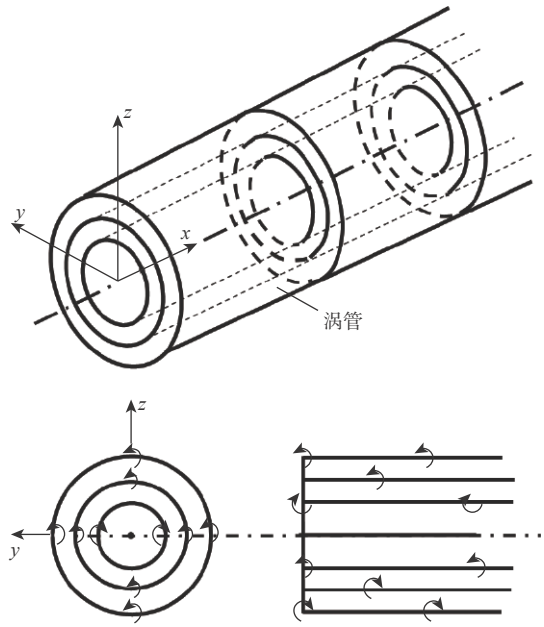


图2 滑流管模型

Fig. 2 Slipstream tube model

在滑流管模型中，涡流微元段$\mathrm{d}\boldsymbol{l}$对在任意点P处的诱导速度$\mathrm{d}\boldsymbol{V}_P$可以表示如下形式：

$$\mathrm{d}\boldsymbol{V}_P=\frac{\Gamma}{4\pi}\cdot\frac{\boldsymbol{M}\times\mathrm{d}\boldsymbol{l}}{\boldsymbol{M}^3} \tag{19}$$

式中：$\boldsymbol{M}$为从点P到微元段的向量；Γ表示微元段环量大小。

1)轴向涡的诱导速度：$u_x=0$

$$\begin{cases}v_x=\dfrac{\gamma_x r}{8\pi}\displaystyle\int_{\varphi=0}^{2\pi}\frac{k}{e^2+k^2}\left[1-\frac{x_0-x_p}{\sqrt{(x_0-x_p)^2+e^2+k^2}}\right]\mathrm{d}\psi\\ w_x=\dfrac{\gamma_x r}{4\pi}\displaystyle\int_{\varphi=0}^{2\pi}\frac{-e}{e^2+k^2}\left[1-\frac{x_0-x_p}{\sqrt{(x_0-x_p)^2+e^2+k^2}}\right]\mathrm{d}\psi\end{cases} \tag{20}$$

2)切向涡的诱导速度：

$$\begin{cases}u_t=\dfrac{\gamma_t r}{4\pi}\displaystyle\int_{\varphi=0}^{2\pi}\frac{e\sin\psi-k\cos\psi}{e^2+k^2}\left[1-\frac{x_0-x_p}{\sqrt{(x_0-x_p)^2+e^2+k^2}}\right]\mathrm{d}\psi\\ v_t=\dfrac{\gamma_t r}{4\pi}\displaystyle\int_{\varphi=0}^{2\pi}\sin\psi\Big/\sqrt{(x_0-x_p)^2+e^2+k^2}\mathrm{d}\psi\\ w_t=\dfrac{\gamma_t r}{4\pi}\displaystyle\int_{\varphi=0}^{2\pi}\cos\psi\Big/\sqrt{(x_0-x_p)^2+e^2+k^2}\mathrm{d}\psi\end{cases} \tag{21}$$

在式(20)和式(21)中：

$$\gamma_x = \frac{B}{2\pi r}\cdot\frac{\mathrm{d}\Gamma}{\mathrm{d}r}\gamma_t = \frac{\Omega B}{V_\infty}\cdot\frac{\mathrm{d}\Gamma}{\mathrm{d}r}$$

$$e = r\sin\psi - y_p$$

$$k = -r\cos\psi - z_p$$

其他细节可参考文献 [20]。

基于叶素动量和涡流理论的螺旋桨设计方法本身存在简化及假设, 设计结果一般与真实情况存在一定的偏差。在得到螺旋桨弦长分布和扭转角分布后, 通过上述螺旋桨滑流快速计算方法检验螺旋桨滑流的分布情况, 并对其迭代修正, 直到满足轴向诱导速度分布均匀的要求, 如图 3 所示。

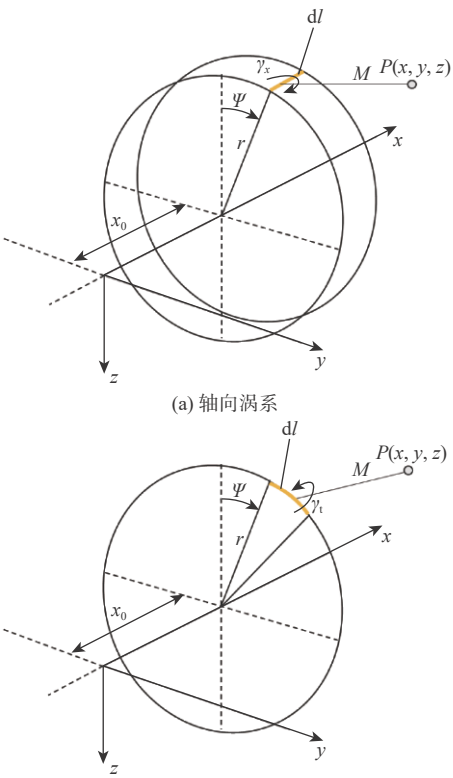


图 3　滑流管模型涡系在任意点 P 的诱导速度

Fig. 3　Induced velocity of slipstream tube model vortex system at any point P

1.3　螺旋桨快速设计流程

基于以上方法及分析, 提出考虑滑流效应螺旋桨快速设计方法, 步骤如下:

步骤 1　快速设计得到初始弦长和扭转角分布。

步骤 2　通过 1.2 节的螺旋桨滑流快速计算方法计算螺旋桨滑流轴向速度沿径向分布。

步骤 3　根据螺旋桨滑流快速计算方法计算的结果修正迭代设计螺旋桨滑流分布。

步骤 4　重复步骤 2～步骤 3, 直到得到满足轴向诱导速度分布均匀的结果。

步骤 5　将修正迭代后的螺旋桨的几何参数导入快速建模程序, 得到螺旋桨几何模型。

设计流程如图 4 所示。

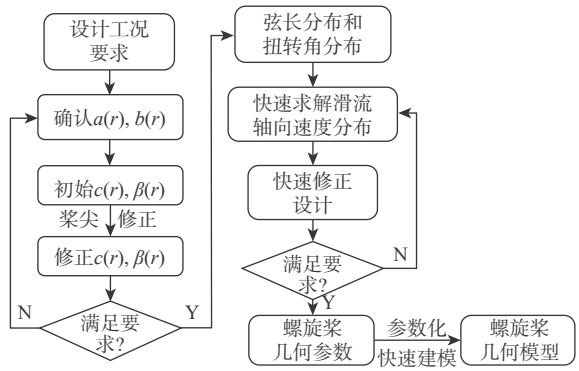


图 4　螺旋桨快速设计流程

Fig. 4　Fast design process of tilting propeller

2　数值方法验证

2.1　螺旋桨设计方法验证

巡航状态是飞机工作时间最长的飞行状态, 提高巡航状态的效率对提升整体飞行器经济性更加明显。针对巡航阶段要求设计螺旋桨, 取飞行高度 H=500 m, 此高度空气密度 ρ=1.156 6 kg/m^3, 飞行速度 V_0=60 m/s。螺旋桨半径 R=0.48 m, 桨叶数 B=5, 转速 N=3 000 r/min, 设计拉力 T=850 N。为简化设计过程, 桨叶整体采用 MH114 翼型。

为了验证本文考虑滑流 (considering slipstream, CST) 效应的螺旋桨设计方法的可靠性, 与传统最小诱导损失 (minimum induced loss, MIL) 方法[18] 对比, 分析 2 种方法设计螺旋桨的性能。图 5 是 2 种螺旋桨的几何模型对比, 为简化计算, 忽略桨毂的气动干扰的影响, 深黄色为 MIL 螺旋桨, 浅黄色为 CST 螺旋桨, 图 6 是根据设计工况设计的螺旋桨的弦长 c/R 和扭转角 β 沿径向分布。

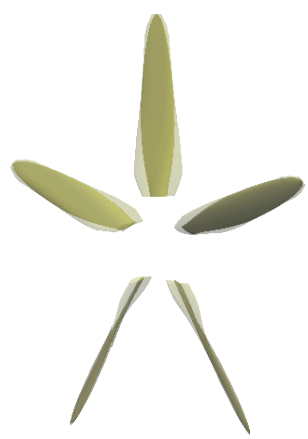


图 5　CST 和 MIL 螺旋桨几何模型对比

Fig. 5　Comparison of CST and MIL propeller geometric models

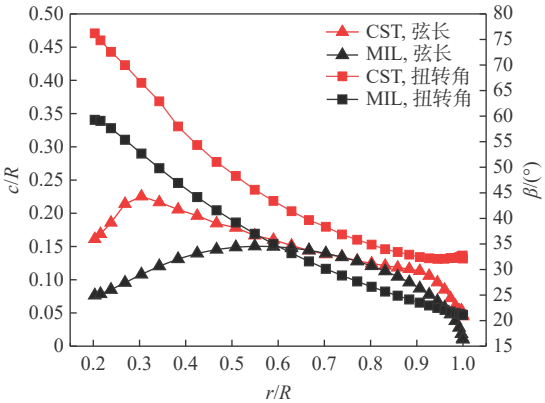


图 6 不同螺旋桨弦长和扭转角分布

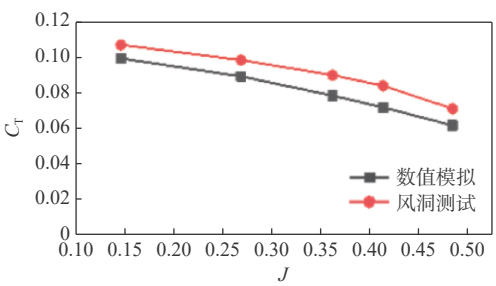
Fig. 6 Distribution of different propeller chord lengths and twist angles

2.2 CFD 数值模拟方法及验证

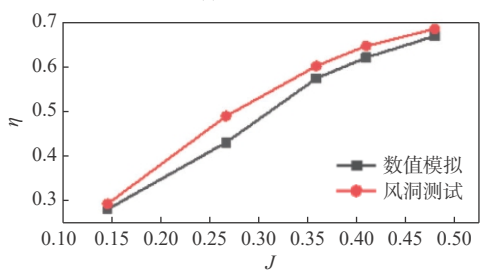
针对常规巡航雷诺数下的螺旋桨和螺旋桨/机翼干扰问题，本文采用耦合 k-ω 剪切应力运输 (shear stress transport, SST) 湍流模型准定常求解雷诺平均纳维-斯托克斯 (Reynolds-averaged Naiver-Stokes, RANS) 方程的刚体运动方法对螺旋桨性能进行数值计算。空间离散方法采用二阶迎风 Roe 格式，黏性通量采用二阶中心差分格式。在湍流模型中，使用了二次本构松弛方案。

为验证数值模拟方法的可行性，对文献 [21] 中的商业螺旋桨 APC 10×7E 型号进行数值模拟，并将计算结果与文献风洞测试结果进行对比。

数值模拟结果与风洞测试结果对比如图 7 所示。计算得到的结果在整体趋势上与风洞测试结果符合较好，具有良好的精度，验证了该螺旋桨数值模拟方法的可行性。



(a) 拉力系数对比



(b) 螺旋桨效率对比

图 7 螺旋桨性能系数对比

Fig. 7 Comparison of propeller performance coefficients

2.3 螺旋桨设计点计算结果分析

采用上述螺旋桨数值模拟方法计算设计的 2 个螺旋桨巡航状态的气动特性，计算结果如表 1 所示。同时，计算 CST 螺旋桨的滑流特性，得到螺旋桨下游 0.06 m 处诱导速度分布，将计算的诱导速度分布与诱导速度修正值、设计值进行对比，如图 8 所示。

表 1 不同螺旋桨巡航状态下的性能参数对比

Table 1 Comparison of performance parameters of different propellers in cruise state

螺旋桨类型	拉力T/N	力矩Q/(N·m)	效率η/%
CST螺旋桨	796	207	73.13
MIL螺旋桨	841	224	71.67

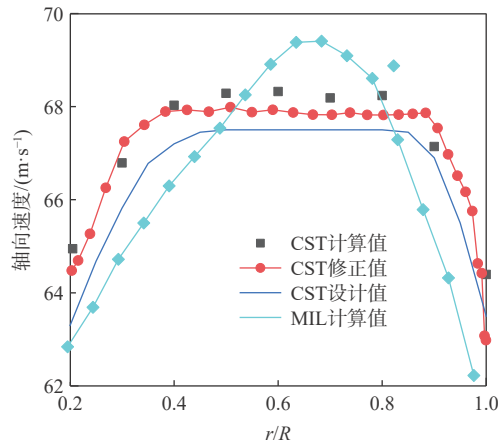


图 8 螺旋桨滑流轴向速度不同值对比

Fig. 8 Comparison of axial induced velocity in propeller slipstream

从表 1 螺旋桨性能参数计算结果可以看出，CST 螺旋桨比 MIL 螺旋桨效率高 1.46%。两者拉力结果与设计拉力存在一定偏差，这是由于设计时叶素升阻特性采用 Xfoil 求解，设计采用的叶素气动力与实际状态存在一定偏差。同时，螺旋桨拉力与设计值相差 6.35%，这是在设计的过程中靠近桨毂的区域输入的轴向诱导因子较小，以避免靠近桨毂部分的弦长过大。从图 8 的诱导轴向速度分布对比可以看出，数值模拟计算的螺旋桨诱导速度分布和诱导速度修正值变化趋势符合较好，说明设计的螺旋桨达到设计要求，设计方法是准确可靠的。

3 分布式构型中的应用

3.1 计算模型

为检验设计的螺旋桨在分布式多螺旋桨/机翼构型巡航状态下的应用情况，参考现有分布式螺旋桨电推进飞机的布局，简化后的计算模型如图 9 所示。分布式构型整体采用六螺旋桨/机翼构型，机翼采用平直机翼，翼展 l=10 m，弦长 c_1=1 m，安装角为 0°，无扭转。为简单统一设计，机翼采用和螺旋桨

一样的 MH114 翼型。6 个螺旋桨等距安装在机翼两侧, 两桨桨尖间距 0.3 m, 机翼两端翼尖各安装一个螺旋桨, 机翼左右两侧螺旋桨均以逆翼尖涡方向运行, 螺旋桨距机翼前缘水平距离 $d=0.354$ m, 具体螺旋桨位置及坐标系定义如表 2 和图 9 所示。数值模拟方法同上文单独螺旋桨数值模拟方法。

3.2　分布式螺旋桨/机翼构型气动特性对比

数值模拟结果如表 3 所示, 以单独机翼构型的气动特性作为参考, 在计算的巡航 0°攻角工况条件下, 构型 2 和构型 3 的增升效果显著, 其中构型 2 中机翼的升力系数增加 20.29%, 阻力系数增加 14.26%, 升阻比增加 6.18%; 构型 3 中机翼的升力系数增加 20.66%, 阻力系数增加 12.82%, 升阻比增加 6.95%, 构型 2 和构型 3 的气动特性得到改善。

表 2　分布式构型计算模型参数和计算条件

Table 2　Distributed configuration calculation model parameters and conditions

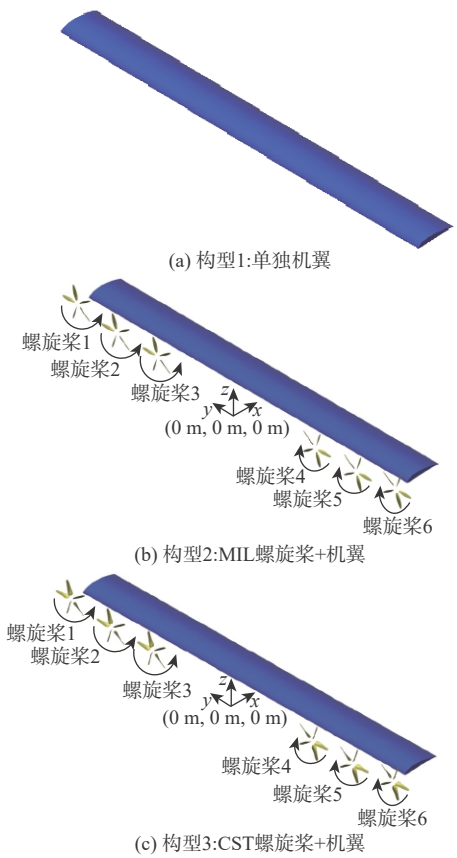
参数	数值
机翼展长/m	10
机翼弦长/m	1
螺旋桨半径R/m	0.48
螺旋桨轮毂半径R_{hub}/m	0.096
螺旋桨位置坐标	(0, ±5, 0.333)(0, ±3.74, 0.178)(0, ±2.48, 0.178)
巡航速度V_0/(m·s^{-1})	60
巡航高度/m	500
机翼攻角α_w/(°)	0、2、4、6
螺旋桨转速/(r·min^{-1})	3 000
马赫数Ma	0.176
雷诺数Re/10^6	7

表 3　3 种构型升阻特性对比

Table 3　Comparison of lift and drag characteristics of three configurations

攻角 α_w/(°)	C_l			C_d			C_l/C_d		
	构1	构2	构3	构1	构2	构3	构1	构2	构3
0	0.69	0.028	24.80	0.83	0.032	26.33	0.83	0.031	26.53
2	0.85	0.044	20.91	1.01	0.046	21.84	1.00	0.045	22.38
4	1.01	0.057	17.71	1.18	0.064	18.52	1.17	0.062	18.79
6	1.17	0.075	15.63	1.34	0.086	15.57	1.33	0.084	15.79

图 10 展示了 3 个构型在计算工况下的升阻特性, 可以看出 2 种螺旋桨/机翼构型气动特性和纯机翼均有改善且增加的效果相同, 这就说明 2 种螺旋桨产生的滑流对机翼的增升效果是一样的。在小攻角工况下, 构型 3 的机翼升阻比构型 2 大, 对比构型 2、3 的升阻特性, 构型 3 是通过增加小阻力和小升力的形式获得比构型 2 更高的升阻比。



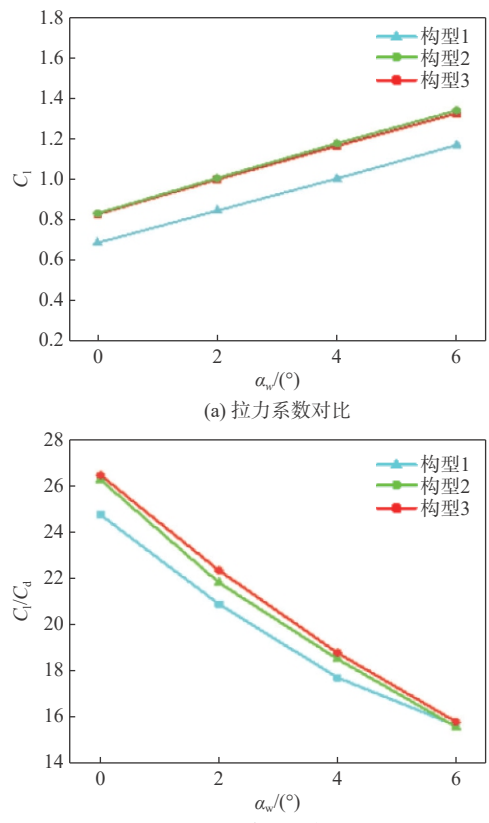
(a) 构型1:单独机翼

(b) 构型2:MIL螺旋桨+机翼

(c) 构型3:CST螺旋桨+机翼

图 9　机翼、螺旋桨/机翼构型计算模型和坐标系定义

Fig. 9　Wing, propeller/wing configuration calculation model and coordinate system definition



(a) 拉力系数对比

(b) 升阻比对比

图 10　3 种构型气动系数-攻角曲线

Fig. 10　Aerodynamic coefficient-angle of attack curves for three configurations

图11为CST螺旋桨和MIL螺旋桨在不同攻角状态下计算的推力、效率对比。推力方面，MIL螺旋桨比CST螺旋桨的推力大，但随着攻角的增加MIL螺旋桨推力均有所不同的下降，而CST螺旋桨的推力随着攻角的增加而增大。效率方面，CST螺旋桨效率比MIL螺旋桨效率高3.4%～6.6%，随着攻角的增加MIL螺旋桨效率均有所不同的下降，且CST螺旋桨的效率随着攻角的增大而增大，最大增幅2.3%。可以看出，当2种构型中螺旋桨滑流对机翼的增升效果一样时，构型3的CST螺旋桨用了较少的扭矩来产生低于MIL螺旋桨的推力以达到更高的效率。

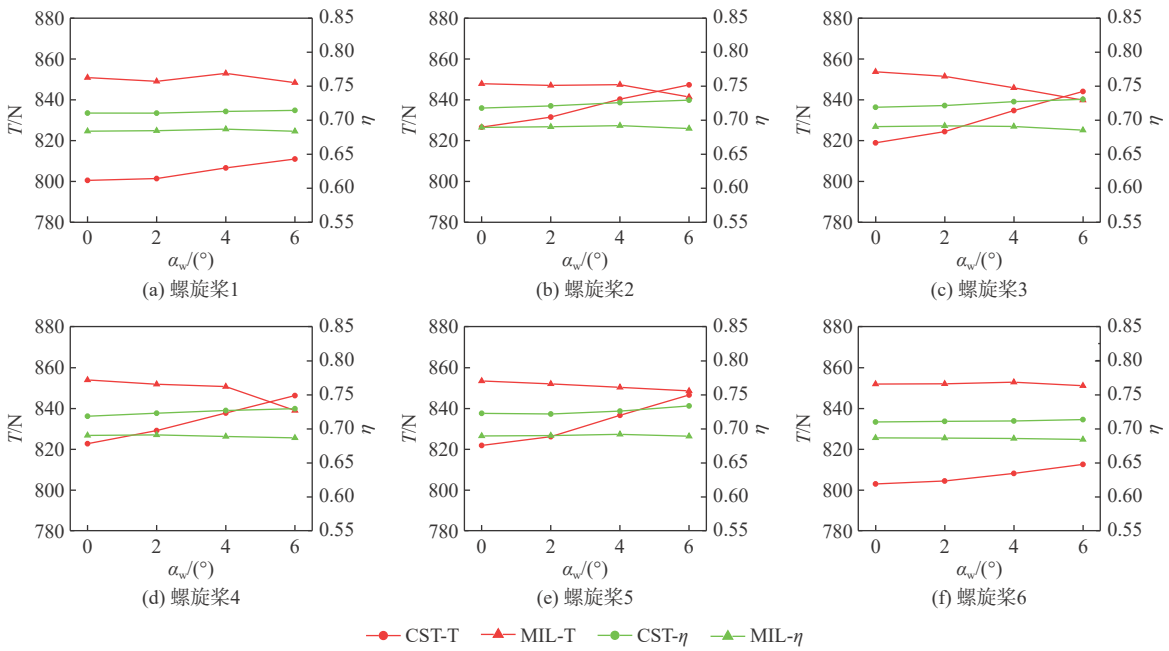


图11 不同设计方法设计的螺旋桨在不同攻角下的气动性能对比

Fig. 11 Comparison of aerodynamic performance of propellers designed by different methods at different angles of attack

3.3 分布式螺旋桨/机翼构型流场分析

为研究螺旋桨在分布式螺旋桨/机翼构型中对机翼气动特性的影响机理，图12展示了3个构型中机翼上表面压力系数分布对比及壁面流线分布。图中红色虚线为螺旋桨旋转轴的位置，构型1的机翼为无动力基准，构型2、3中的机翼左侧螺旋桨逆时针旋转、右侧螺旋桨顺时针旋转。相比于构型1来说，构型2和构型3的近壁面流线较为紊乱，并伴有明显的展向流动，这与螺旋桨滑流机翼的相互作用有关。对比构型2和构型3压力分布，可以看出构型3机翼上表面的低压区滑流影响的展向范围更大，数值上也更低。螺旋桨所处位置的两侧机翼前缘的压力分布也因为螺旋桨的旋转效应而不同，同时由于受到螺旋桨滑流加速影响，下

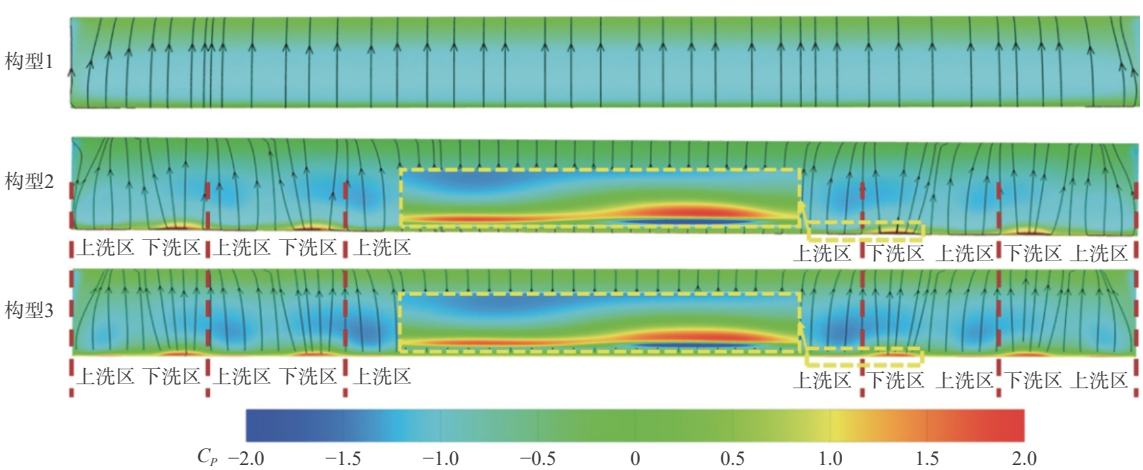
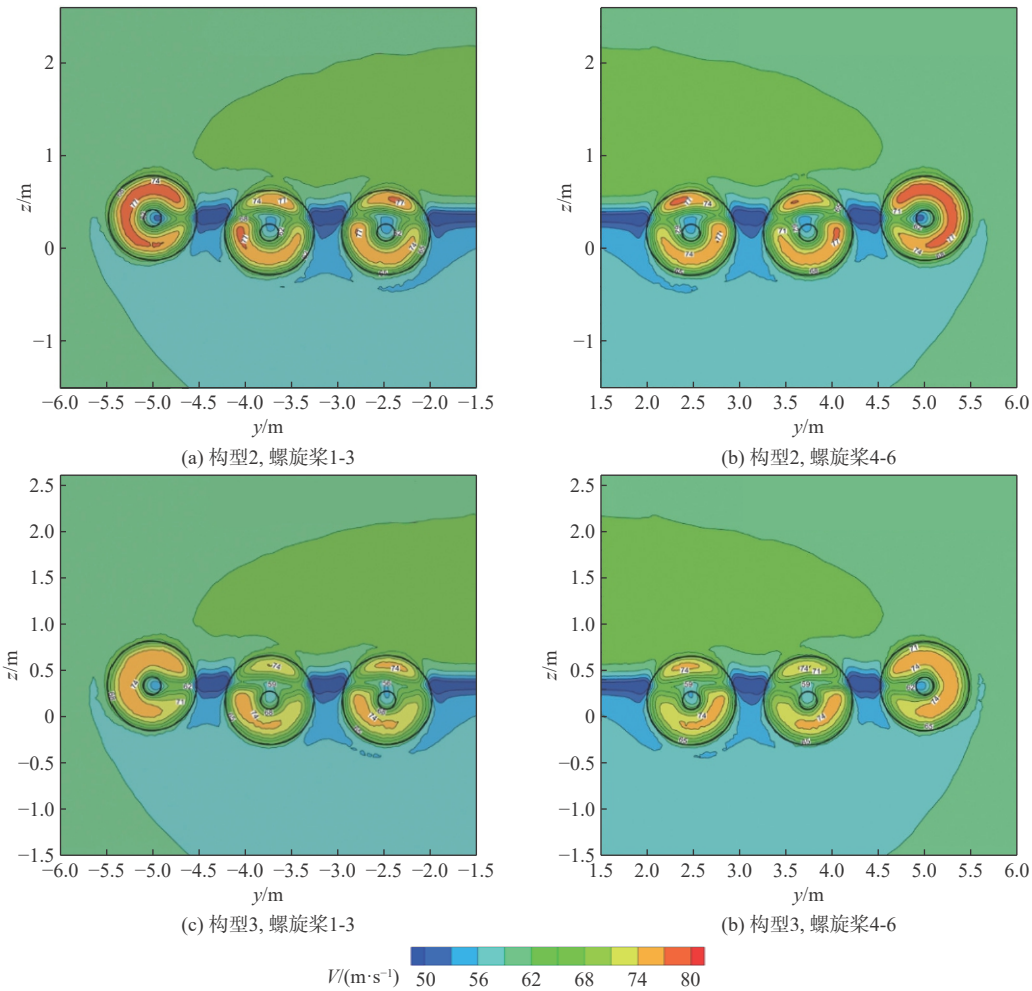


图12 2°攻角机翼上表面压力系数分布对比及近壁面流线分布

Fig. 12 Comparison of pressure coefficient distribution on the upper surface of wing at 2° angle of attack and streamline distribution near wall

洗区的机翼前缘上表面周围压力更高, 上洗区的机翼前缘下表面周围压力更高。

图 13 为构型 2 和构型 3 螺旋桨下游 x=0.28 m 处的截面速度分布, 其中图 13(a) 和图 13(b) 为构型 2MIL 螺旋桨结果, 图 13(c) 和图 13(d) 为构型 3CST 螺旋桨结果。由图可以看出: MIL 螺旋桨滑流区域内最大速度区域分布于沿径向居中的位置, 由桨毂至桨尖速度趋势由小变大再变小, CST 螺旋桨滑流区域内最大速度区域分布情况和 MIL 螺旋桨一样, 但 MIL 螺旋桨的最大速度值比 CST 螺旋桨大, 速度变化梯度大, CST 螺旋桨最大速度区域的范围更广同时速度更加均匀。



(a) 构型2, 螺旋桨1-3　(b) 构型2, 螺旋桨4-6

(c) 构型3, 螺旋桨1-3　(b) 构型3, 螺旋桨4-6

图 13　x=0.28 m 处的速度分布对比

Fig. 13　Comparison of velocity distribution at x=0.28 m

4　螺旋桨悬停工况气动特性计算

4.1　孤立螺旋桨悬停工况计算

螺旋桨在悬停工况下处于小扭转角状态, 而本文螺旋桨是以巡航工况下设计的, 扭转角过大, 悬停效率过低, 故对其做变距处理, 以检验螺旋桨在悬停工况的气动性能。为简化计算, 本文将直接采用变距 25°处理。该工况仍采用本文上述的数值模拟方法计算该模型悬停状态的气动特性, 计算工况为来流速度 0 m/s、空气密度 1.156 6 kg/m^3、螺旋桨转速范围 2 000～4 000 r/min 等多组转速的计算。

计算结果如表 4 所示, 表中列出了悬停工况不同转速下螺旋桨的气动性能参数。从表中可以看出, 随着螺旋桨转速的增大, 推力和扭矩也随之增大, 但 CST 螺旋桨的悬停效率保持稳定, 而 MIL 螺旋桨的悬停效率变化趋势先增大后减小。对比 2 种类型螺旋桨的悬停状态的气动性能, CST 螺旋

表 4　螺旋桨悬停状态的气动性能参数

Table 4　Aerodynamic performance parameters of propeller in hover state

转速/(r·min^{-1})	拉力/N		悬停效率/%	
	CST	MIL	CST	MIL
2 000	242	114	68.48	56.85
2 500	383	180	68.62	58.23
3 000	556	262	68.71	57.18
3 500	765	356	68.68	56.02
4 000	1 010	469	68.64	55.10

桨的悬停效率比 MIL 螺旋桨悬停效率高 10.40%～13.54%，在转速达到 4 000 r/min 时，2 个类型的螺旋桨悬停效率差值最大达到 13.54%。

螺旋桨在悬停时的悬停效率一般在 0.6～0.7[22]，可以看出 CST 螺旋桨的悬停效率稳定在 68%，转速空间比较大。当 2 种螺旋桨改变相同的桨距时，CST 螺旋桨用了较大的扭矩来产生高于 MIL 螺旋桨的推力以达到更高的效率。

4.2 分布式螺旋桨/机翼构型悬停计算

为进一步了解 CST 螺旋桨在分布式构型悬停状态的气动性能，将构型 2 和构型 3 的巡航状态转换成悬停状态，具体螺旋桨位置如表 5 所示，坐标系定义如图 14 所示。计算该构型悬停状态下的气动特性。

表 5 分布式构型悬停状态计算模型参数和计算条件

Table 5 Calculation model parameters and conditions for distributed configuration in hover state

参数	数值
螺旋桨坐标	(−0.045, ±5, 0.696)(−0.2, ±3.74, 0.696)(−0.2, ±2.48, 0.696)
速度V_1(m·s^{-1})	0
转速/(r·min^{-1})	4 000
雷诺数Re/10^6	2

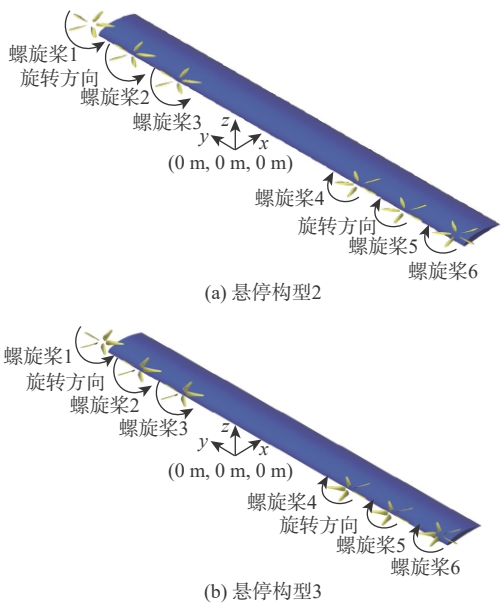


图 14 悬停状态的计算模型及坐标系定义

Fig. 14 Calculation model of hover state and definition of coordinate system

悬停工况的数值模拟计算结果如表 6 所示，与 MIL 螺旋桨相比，CST 螺旋桨的悬停效率提升 13%。但在分布式构型中的螺旋桨拉力比孤立构型拉力计算值小，同时螺旋桨的悬停效率比孤立螺旋桨低，与同转速下的孤立螺旋桨对比，分布式构型中螺旋桨的拉力计算值小 3.1%～4.1%、悬停效率减小 7.51%～8.28%。这说明，分布式螺旋桨/机翼构型对螺旋桨的气动性能有不利的影响。

为了解分布式构型对螺旋桨气动性能的影响，图 15 给出的是悬停构型的螺旋桨尾迹 Q 等值面，螺旋桨尾迹涡缠绕在机翼上，与机翼产生严重干扰。图 16 给出的是 x=0.2 m 截面速度分布和表面流线分布。在螺旋桨的两侧上方，悬停状态的气流被吸入进入桨盘，形成螺旋桨的滑流，在螺旋桨的外侧下方，循环气流形成 2 个很大的涡，机翼两侧

表 6 螺旋桨悬停状态的气动性能参数

Table 6 Aerodynamic performance parameters of configuration 4 propeller in hover state

类型	拉力/N		悬停效率/%	
	CST	MIL	CST	MIL
Prop1	978	479	61.13	48.35
Prop2	970	472	60.40	47.30
Prop3	978	480	61.10	48.51
Prop4	979	480	61.04	48.51
Prop5	969	466	60.53	46.40
Prop6	976	479	61.10	48.35

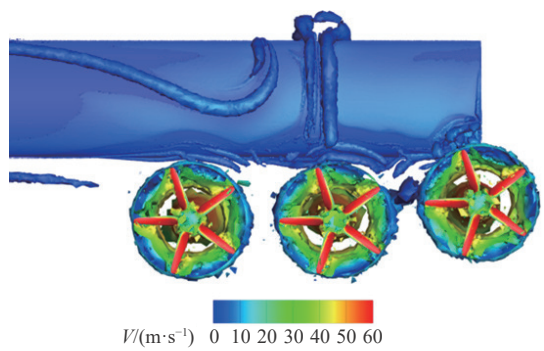


图 15 螺旋桨尾迹 Q 等值面

Fig. 15 Q-criterion contour of propeller wake under hovering

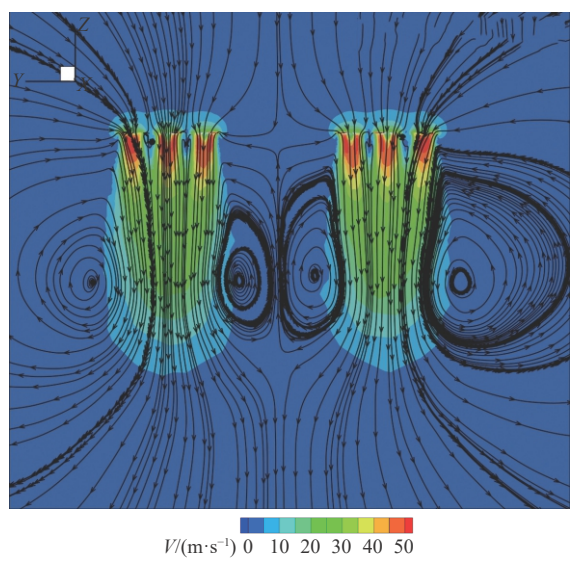


图 16 悬停构型 x=0.2 m 截面速度分布和表面流线分布

Fig. 16 Velocity and surface streamline distribution for hovering configuration at x=0.2 m section

螺旋桨的尾流汇集在机翼前对称面形成 2 个对称的涡, 严重影响螺旋桨的气动性能。

5　结　论

本文建立了考虑滑流效应的螺旋桨快速设计方法, 并通过低阶数值模拟计算对螺旋桨滑流的轴向速度进行快速修正。将设计螺旋桨应用于分布式构型, 研究其在巡航和悬停状态的气动性能, 得到的结论如下:

1) 基于叶素动量和涡流理论的考虑滑流效应螺旋桨快速设计方法, 设计的螺旋桨滑流轴向速度分布与高阶数值计算结果存在一定差距, 通过低阶数值模拟计算方法快速迭代修正, 得到与高阶数值计算结果更接近的螺旋桨。

2) 与最小诱导损失方法相比, 本文设计的 CST 螺旋桨在孤立构型中巡航效率提升了 1.46%; 分布式构型, 螺旋桨巡航效率提升了 3%～7%; 在计算的攻角范围内, 构型 2 中 MIL 螺旋桨效率随着攻角增大而减小, 构型 3 中 CST 螺旋桨效率随着攻角增大而增大。

3) 悬停工况时, 本文设计的 CST 螺旋桨在孤立构型的悬停效率比 MIL 螺旋桨的悬停效率高 10% 以上; 在不同转速下, 本文设计的 CST 螺旋桨的悬停效率始终保持 68% 附近; 本文设计的 CST 螺旋桨在分布式构型的悬停效率比 MIL 螺旋桨高 13%。

4) 相对于孤立螺旋桨, 分布式构型中的螺旋桨巡航和悬停效率更低, 巡航工况效率下降不明显, 悬停工况下降 8%, 说明桨—桨和桨—翼气动干扰会降低分布式构型的螺旋桨效率, 在设计中需合理布置螺旋桨之间和螺旋桨与机翼的相对位置。

致　谢　感谢气弹课题组的帮助与支持

参考文献（References）

[1] KIM H D, PERRY A T, ANSELL P J. A review of distributed electric propulsion concepts for air vehicle technology[C]//Proceedings of the 2018 AIAA/IEEE Electric Aircraft Technologies Symposium. Reston: AIAA, 2018: 4998.

[2] 黄俊. 分布式电推进飞机设计技术综述[J]. 航空学报, 2021, 42(3): 624037.
HUANG J. Survey on design technology of distributed electric propulsion aircraft[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(3): 624037 (in Chinese).

[3] PATTERSON M D, DERLAGA J M, BORER N K. High-lift propeller system configuration selection for NASA's SCEPTOR distributed electric propulsion flight demonstrator[C]//Proceedings of the 16th AIAA Aviation Technology, Integration, and Operations Conference. Reston: AIAA, 2016: 3922.

[4] Patterson M D. Conceptual design of high-lift propeller systems for small electric aircraft[D]. Atlanta: Georgia Institute of Technology, 2016.

[5] DROANDI G, ZANOTTI A, GIBERTINI G, et al. Experimental investigation of the rotor-wing aerodynamic interaction in a tiltwing aircraft in hover[J]. The Aeronautical Journal, 2015, 119(1215): 591-612.

[6] SHUKLA D, KOMERATH N. Multirotor drone aerodynamic interaction investigation[J]. Drones, 2018, 2(4): 43.

[7] STOKKERMANS T C A, USAI D, SINNIGE T, et al. Aerodynamic interaction effects between propellers in typical eVTOL vehicle configurations[J]. Journal of Aircraft, 2021, 58(4): 815-833.

[8] Young L A, Derby M R. Rotor/wing interactions in hover, NASA/TM-2002-211392[R]. Washington D.C.: NASA, 2002.

[9] Veldhuis L L. Review of propeller-wing aerodynamic interference [C]//Proceedings of the 24th International Congress of the Aeronautical Sciences. Beijing: ICAS, 2004: 2004-2006.

[10] PATTERSON M D, BORER N K, GERMAN B. A simple method for high-lift propeller conceptual design[C]//Proceedings of the 54th AIAA Aerospace Sciences Meeting. Reston: AIAA, 2016: 0770.

[11] XUE C, ZHOU Z. Propeller-wing coupled aerodynamic design based on desired propeller slipstream[J]. Aerospace Science and Technology, 2020, 97: 105556.

[12] 范中允, 周洲, 祝小平. 一种可任意给定环量分布的螺旋桨设计方法[J]. 航空动力学报, 2019, 32(2): 434-441.
FAN Z Y, ZHOU Z, ZHU X P. A design method for propeller with arbitrary circulation distribution[J]. Journal of Aerospace Power, 2019, 32(2): 434-441 (in Chinese).

[13] 郭佳豪, 周洲, 范中允. 一种给定拉力分布的螺旋桨设计方法及应用[J]. 航空动力学报, 2020, 35(6): 1238-1246.
GUO J H, ZHOU Z, FAN Z Y. A method of propeller design with given thrust distribution and it's application[J]. Journal of Aerospace Power, 2020, 35(6): 1238-1246 (in Chinese).

[14] D'ANGELO S, BERARDI F, MINISCI E. Aerodynamic performances of propellers with parametric considerations on the optimal design[J]. The Aeronautical Journal, 2002, 106(1060): 313-320.

[15] 项松, 王吉, 张利国, 等. 一种高效率螺旋桨设计方法[J]. 航空动力学报, 2015, 30(1): 136-141.
XIANG S, WANG J, ZHANG L G, et al. A design method for high efficiency propeller[J]. Journal of Aerospace Power, 2015, 30(1): 136-141 (in Chinese).

[16] 唐伟, 宋笔锋, 张玉刚, 等. 两个设计点的螺旋桨气动性能[J]. 航空动力学报, 2017, 32(2): 354-363.
TANG W, SONG B F, ZHANG Y G, et al. Aerodynamic performance of propeller with two design points[J]. Journal of Aerospace Power, 2017, 32(2): 354-363 (in Chinese).

[17] ADKINS C N, LIEBECK R H. Design of optimum propellers[J]. Journal of Propulsion and Power, 1994, 10(5): 676-682.

[18] LARRABEE E E. Practical design of minimum induced loss propellers[J]. SAE Transactions, 1979: 2053-2062.

[19] 薛臣, 周洲, 范中允, 等. 螺旋桨/机翼耦合下的目标螺旋桨滑流设计[J]. 航空动力学报, 2021, 36(1): 104-118.
XUE C, ZHOU Z, FAN Z Y, et al. Design of target propeller slipstream under propeller-wing interaction[J]. Journal of Aerospace Power, 2021, 36(1): 104-118 (in Chinese).

[20] VELDHUIS L L M. Propeller wing aerodynamic interference[D].

Delft: Technische Universiteit Delft, 2005.

[21] DANTSKER O D, CACCAMO M, DETERS R W, et al. Performance testing of APC electric fixed-blade UAV propellers[C]//Proceedings of the AIAA AVIATION 2022 Forum. Reston: AIAA, 2022: 4020.

[22] 王适存. 直升机空气动力学[M]. 北京: 国防工业出版社, 1959.
Wang S C, Helicopter aerodynamics[M]. Beijing: National Defence Industry Press, 1959.

A propeller design method considering slipstream effect and its application

CHEN Shengjiu，YANG Youxu*，ZHANG Xingcui，WU Yifei

(School of Flight Vehicle and Engineering，Nanchang Hangkong University，Nanchang 330063，China)

Abstract：Propellers are widely used in distributed electric propulsion aircraft, where a significant propeller-wing aerodynamic coupling effect exists. By changing the axial induced velocity distribution to control the propeller slipstream, the propeller chord length and torsion angle distributions were obtained. An efficient propeller design method considering the slipstream effect was proposed, and the feasibility of the method was verified. The cruise and hover aerodynamic characteristics of the designed propellers in isolated and distributed configurations were calculated and compared with the minimum induced loss method. The results show that, for the same cruising lift coefficient, the efficiency of the designed propeller is 3.4%–6.6% higher than that of the minimum induced loss method. The hovering efficiency of a single propeller is 10.4%–13.5% higher than that of the minimum induced loss method at different rotational speeds. In the distributed configuration, the designed propeller's hover efficiency is 13% higher than that of the minimum induced loss method. The designed propeller meets its design requirements in both cruise and hover states, maintaining high-efficiency operation.

Keywords：slipstream effect；distributed electric propulsion；tilting propeller；electric vertical takeoff and landing；aerodynamic-propulsion coupling

Received：2023-08-17；**Accepted**：2023-11-16；**Published Online**：2023-12-08 10：17

URL：link.cnki.net/urlid/11.2625.V.20231206.1445.006

Foundation items：Jiangxi Province "Double Thousand Plan" Talent Project (CK202006470)

* **Corresponding author**. E-mail：zgdy_1@163.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2023.0545

基于注意力机制改进的PSO-BiLSTM刀具磨损预测

杨沛东，黄华*，尉卫卫，郭宝岛
(兰州理工大学 机电工程学院，兰州 730050)

摘 要：针对刀具磨损故障诊断中存在的监测数据单一和特征信号处理效果差的问题，提出了一种基于注意力机制(AM)改进的粒子群算法(PSO)优化双向长短时记忆(BiLSTM)神经网络来实现端到端的刀具磨损预测方法。根据传感器信号进行多域特征提取，构建优质的信号输入样本；利用卡尔曼滤波对输入样本进行多传感器数据融合，得到鲁棒性更高的融合数据样本，在此基础上，通过PSO对BiLSTM网络进行超参数寻优，根据优化的超参数建立神经网络模型；基于注意力机制赋予输入影响权重，改进PSO-BiLSTM以获得更好的刀具磨损预测效果。对比实验结果验证了所提模型在刀具磨损预测中的可行性，其精度相比传统深度学习方法有较大的提升。

关 键 词：刀具磨损；卡尔曼滤波；粒子群算法；注意力机制；双向长短时记忆神经网络
中图分类号：TH164; V262.3+3
文献标志码：A **文章编号**：1001-5965(2025)10-3589-10

准确、可靠的刀具磨损预测是切削加工过程中的关键问题之一[1]，根据采用的监测方法不同，刀具磨损预测技术可分为直接法和间接法[2]。直接法是通过直接监测刀具的材料形状[3]或光学成像[4]来判断刀具磨损状况，但直接法监测方法昂贵并且易受操作环境的限制。间接法通过在现有的机床上安装传感器，监测加工过程中的力、振动和声发射等信号，并采用机器学习或深度学习智能预测刀具磨损情况。间接法实时性强，可通过预测模型自动预警，实现刀具磨损的智能监测。在刀具磨损监测与预测中，时域和频域是刀具磨损的主要特征[5]，能够有效表征刀具磨损状态。戴稳等[6]通过小波分解从多个频段的信号中获得信息，并以此对刀具的磨损状态进行识别；郑圆等[7]利用参数优化后的变分模式分解(variational mode decomposition, VMD)技术，准确地监测到轴承的故障，而且相比传统的VMD、经验模态分解(emprical mode decomposition, EEMD)以及快速谱峭度方法，该方法可以显著地增强故障特征提取的灵活性；李聪波等[8]提出了一种基于实时调整切削力功率阈值的刀具磨损在线监测系统，通过研究加工参数对切削功率的影响，能够快速发现刀具的磨损情况，减少刀具磨损产生的不良影响。尽管如此，由于上述人工智能方法对高维数据和复杂非线性关系的学习能力受到浅层网络结构的限制，必须依靠已有的经验进行多种尝试，人为地降低多域特征维度，容易造成潜在有效特征的损失，阻碍了刀具磨损预测精度的进一步提高[9]。

近年来，深度学习因其出色的数据特征提取能力，在图像识别、语言处理等领域得到了广泛的应用。深度学习从数据中自适应学习特征，不仅可以减少对先验知识和人类经验的依赖，而且使得识别精度和泛化性得到了极大的提升，从而改善了刀具磨损预测的工作效率。曹大理等[10]提出了一种增加卷积神经网络深度的方法，通过改进的网络能够更好地提取深层的特征，进而提高模型整体的预测精度。王丽华等[11]通过提取声发射信号特征，提出一种基于堆叠降噪自编码网络的数据驱动刀具磨损识别方法。结果表明，该模型具有较高的预测精

收稿日期：2023-08-25；**录用日期**：2023-11-13；**网络出版时间**：2023-12-13 12:12
网络出版地址：link.cnki.net/urlid/11.2625.V.20231212.1424.001
基金项目：国家自然科学基金(52365057,51965037)
*通信作者. E-mail：hh318872@126.com

引用格式：杨沛东，黄华，尉卫卫，等. 基于注意力机制改进的PSO-BiLSTM刀具磨损预测[J]. 北京航空航天大学学报，2025，51(10)：3589-3598. YANG P D，HUANG H，YU W W，et al. Tool wear prediction based on attention mechanism and PSO-BiLSTM[J]. Journal of Beijing University of Aeronautics and Astronautics，2025，51(10)：3589-3598 (in Chinese).

度。虽然这些深度学习方法比传统的机器学习方法更有效，但仍然存在以下不足：首先现有的文献也较多地使用长短时记忆（long short-term memory, LSTM）以及改进算法，但由于LSTM受到其前馈网络的限制，只能处理当前输入的数据，当处理具有长期依赖关系的时序数据时，传统的LSTM难以有效地捕捉长距离的依赖关系，不能同时处理前向和后向数据，容易出现梯度消失的问题。另外，虽然大部分研究中采用了叠加的神经网络，但是只使用了单一的振动信号进行研究，忽略了多传感器信号融合的问题，难以充分揭示信号的代表性特征，导致与刀具磨损相关的重要信息丢失，降低了预测精度[12]。

针对以上问题，本文提出了一种基于注意力机制 (attention mechanism, AM) 改进的粒子群优化双向长短期记忆网络 (particle swarm optimization-Bi-directional long short-term memory, PSO-BiLSTM) 的刀具磨损预测方法。首先在时域、频域和时频域进行多域特征提取，利用卡尔曼（Kalman）滤波的自适应特征融合方法，通过状态转移函数计算卡尔曼增益实现对特征信号的有效融合。然后通过粒子群优化 (particle swarm optimization, PSO) 算法对融合特征进行参数寻优，根据超参数取值范围随机初始化各粒子位置信息。其次根据粒子位置对应的超参数取值建立具有注意力机制的BiLSTM网络模型，通过注意力机制计算并合理分配网络权值，选择性学习关键信息。最后利用神经网络模型对优化的参数进行训练，实现刀具磨损的预测。与传统的BiLSTM结构不同的是，注意力机制可以在不同的时间步长关注不同权重的特征。该模型结合了强大的多传感器信号特征融合能力、优化的BiLSTM从时间序列数据中处理数据的能力和注意力机制选择性关注不同特征的能力，进一步提升BiLSTM模型在多传感器数据输入条件下的准确性和泛化性。

1 刀具磨损预测模型及理论

本文目的是在多传感器数据的前提下实现一种端到端的刀具磨损预测。为充分利用数据并且使模型具有较好的泛化能力，本文使用切削力传感器和振动传感器分别测量了刀具X、Y、Z轴向的切削力和振动信号的时间序列数据，将这些数据进行多域特征提取，然后根据Kalman滤波方法对多域特征提取的信号进行特征融合，将融合之后的特征组成模型的输入数据集，训练具有注意力机制的BiLSTM网络。将PSO算法确定的最优超参数输入到网络中，得到模型的最优配置之后进行刀具磨损的预测。实现过程如图1所示。

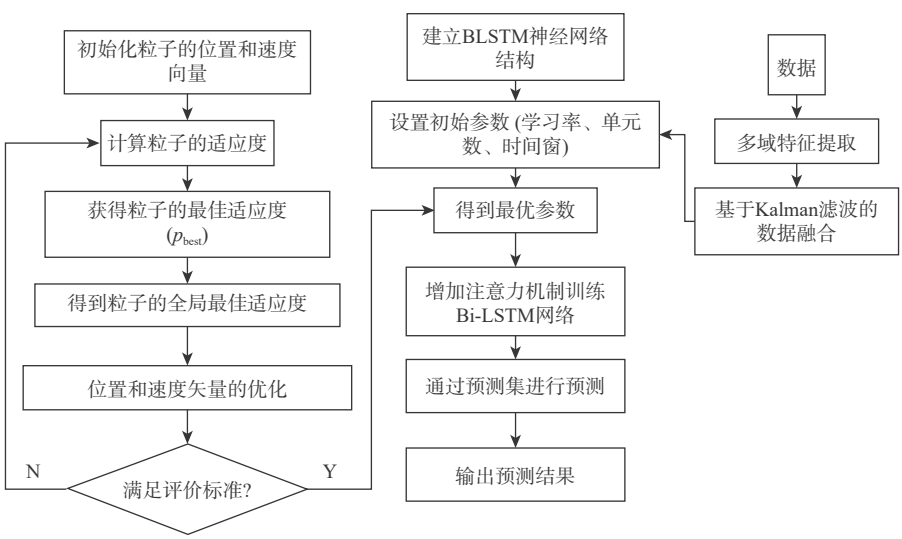


图1 模型框架

Fig. 1 Model framework

1.1 多传感器特征融合

Kalman滤波具有简单易用、低成本、高效率等优点，在特征融合方面得到了广泛的研究[13]。Kalman滤波利用系统噪声和观测噪声的统计特性，融合前以系统的观测量作为滤波器的输入，融合后以所要求的估计值作为滤波器的输出，滤波器的输入和输出由时间更新和观测更新联系在一起。

设多传感器有如下形式：

$$\boldsymbol{x}(k+1)=\boldsymbol{A}(k)\boldsymbol{x}(k)+\boldsymbol{w}(k) \tag{1}$$

$$\boldsymbol{z}(i,k)=\boldsymbol{C}(i,k)\boldsymbol{x}(k)+\boldsymbol{v}(i,k) \tag{2}$$

式中：$\boldsymbol{x}(k)\in\mathbf{R}^{N\times1}$为第$k$−1时刻的状态预测值；$\boldsymbol{A}(k)\in\mathbf{R}^{n\times n}$为状态转移矩阵；$\boldsymbol{z}(i,k)\in\mathbf{R}^{q_i\times n}(q_i\leqslant n)$为第$i$个

传感器的观察向量; $\boldsymbol{C}(i,k)\in\mathbf{R}^{q_i\times n}$ 为观察矩阵, 这些值可以用来描述系统的运行情况; $\boldsymbol{w}(k)\in\mathbf{R}^{n\times 1}$ 和 $\boldsymbol{v}(i,k)\in\mathbf{R}^{q_i\times 1}$ 分别为系统噪声和观测噪声序列, 系统噪声和观测噪声被建模为零均值的白噪声。式(1)为预测方程, 主要描述了如何从先前的状态和输入预测系统中得到下一个状态, 式(2)为观测方程, 主要描述系统的观测如何与状态之间建立联系。

状态向量初始值 $\boldsymbol{x}(0)$ 为一随机向量, 具有以下特征:

$$E\{\boldsymbol{x}(0)\}=\boldsymbol{x}_0 \tag{3}$$

$$E\{[\boldsymbol{x}(0)-\boldsymbol{x}_0][\boldsymbol{x}(0)-\boldsymbol{x}_0]^{\mathrm{T}}\}=P_0 \tag{4}$$

假设 $\boldsymbol{x}(0)$、$\boldsymbol{w}(k)$ 和 $\boldsymbol{v}(i,k)$ 之间存在显著的相关性, 其中 i=1, 2, …, N 分别表示不同的传感器。

1.2　双向长短时记忆神经网络

为了解决递归神经网络长期依赖的问题, 提出了一种新的递归网络架构, 如 LSTM、门控循环单元(gate recurrent unit, GRU) 以及相应的变体 BiLSTM。LSTM 网络是一种改变循环神经网络(recurrent neural network, RNN)模型循环结构的递归神经网络, 能有效地处理长序列输入, 具有更强的长期记忆和建模能力[14], 其结构如图 2所示。

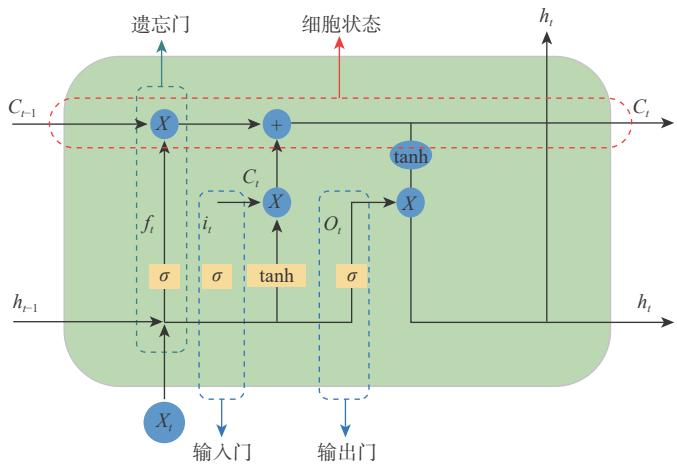


图 2　LSTM 单元结构

Fig. 2　LSTM cell structure

与传统的神经网络形式有所区别, LSTM 采用了“门”结构控制信息的遗忘与流动, 主要包括遗忘门、输入门和输出门。无论是增强信息还是减弱信息, 都需要将相应的数据传输到“门”结构, 从而完成对数据的管理。

1)遗忘门。LSTM 算法的第一步是由遗忘门来丢弃掉一些数据, 会从外界的输入 X_t 和 h_{t-1} 中提取一些有用的信息, 然后将这些信息输入到内部状态。

$$\boldsymbol{f}_t=\sigma(\boldsymbol{W}_{fC}\times\boldsymbol{C}_t+\boldsymbol{W}_{fh}\times\boldsymbol{h}_{t-1}+\boldsymbol{b}_f) \tag{5}$$

式中: $\boldsymbol{W}$ 为权重; $\boldsymbol{b}$ 为偏差。$\boldsymbol{W}_{fC}$ 为遗忘门对应 $\boldsymbol{X}_t$ 的权重, $\boldsymbol{b}_f$ 为遗忘门偏差, σ 为 Sigmoid 激活函数。

2)输入门。首先, 输入门 $\boldsymbol{i}_t$ 将会更新相关的指标, 其次, 输入节点 $\boldsymbol{g}_t$ 将会创建一个可供参考的状态 $\boldsymbol{C}_t$。如下:

$$\boldsymbol{i}_t=\sigma(\boldsymbol{W}_{iX}\times\boldsymbol{X}_t+\boldsymbol{W}_{iH}\times\boldsymbol{h}_{t-1}+\boldsymbol{b}_i) \tag{6}$$

$$\boldsymbol{g}_t=\tanh(\boldsymbol{W}_{gX}\times\boldsymbol{X}_t+\boldsymbol{W}_{gh}\times\boldsymbol{h}+\boldsymbol{b}_g) \tag{7}$$

式中: $\boldsymbol{b}_i$ 和 $\boldsymbol{b}_g$ 分别为输入门和输入节点偏差。

则由式(1)～式(4)可以将旧的内部状态 $\boldsymbol{C}_{t-1}$ 更新为 $\boldsymbol{C}_t$:

$$\boldsymbol{C}_t=\boldsymbol{g}_t\boldsymbol{i}_t+\boldsymbol{C}_{t-1}f_t \tag{8}$$

3)输出门。由输出门决定输出的信息, 如下:

$$\boldsymbol{o}_t=\sigma(\boldsymbol{W}_{oC}\times\boldsymbol{X}_t+\boldsymbol{W}_{oh}\times\boldsymbol{h}_{t-1}+\boldsymbol{b}_o) \tag{9}$$

式中: $\boldsymbol{b}_o$ 为输出门偏差。

最后, 结合输出门决定输出的信息, 如下:

$$\boldsymbol{h}_t=\tanh\boldsymbol{S}_t\times\boldsymbol{o}_t \tag{10}$$

式中: $\boldsymbol{S}_t$ 为单元格状态; $\boldsymbol{o}_t$ 为同步门。

虽然 LSTM 可以更好地捕捉远程依赖问题, 但是在预测模型中, LSTM 不能将信息从后到前进行编码, 只能从一个方向访问串行数据。与标准的单向 LSTM 相比, BiLSTM 可以同时获得历史和当前信息的相关性, 提高预测能力。完整的 BiLSTM 隐藏元素提供了正向和反向过程的输出联向量, 如下:

$$\boldsymbol{h}_t=\overrightarrow{h^t}\oplus\overleftarrow{h^t} \tag{11}$$

式中: $\overrightarrow{h^t}$ 和 $\overleftarrow{h^t}$ 分别为前向 LSTM 层和后向 LSTM 层; “⊕” 表示按元素求和, 用于对正向和反向输出的元素求和。其结构如图 3 所示。

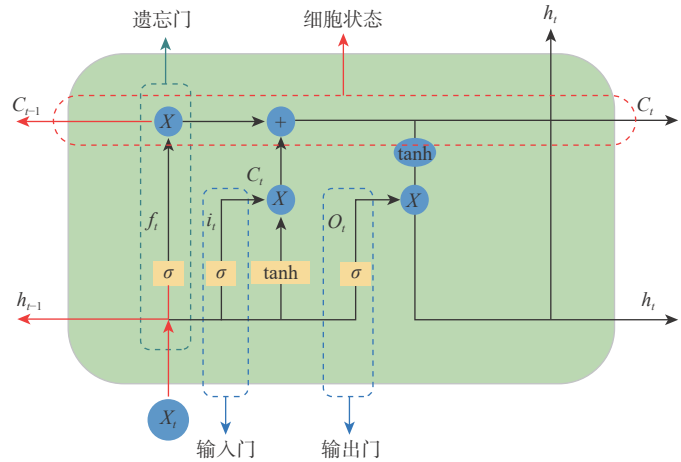


图 3 BiLSTM 单元结构

Fig. 3 BiLSTM cell structure

1.3 粒子群算法

PSO 算法是一种基于自然语言处理的智能算法，通过对当前解的迭代演变找到更加准确的结果[15]。和常见的蚁群算法、遗传算法和差分进化算法相比，PSO 算法具有更快的收敛速度，每个粒子通过迭代更新自己的位置和速度，利用群体中的信息快速地进行寻优。PSO 算法在 D-维空间中寻优的流程，如图 4 所示。

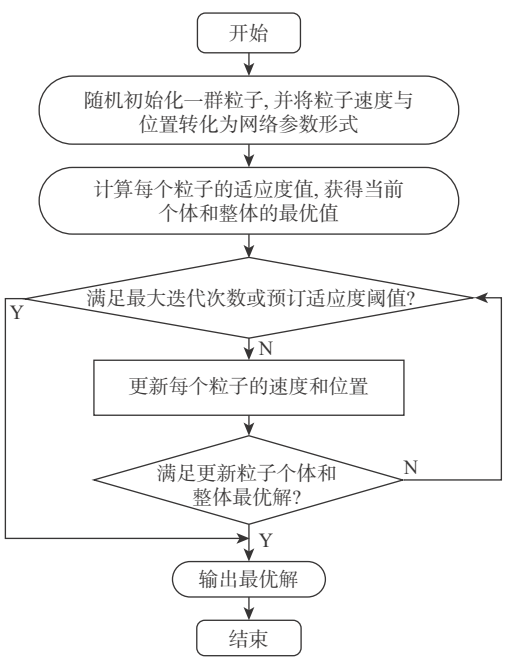


图 4 优化流程

Fig. 4 Optimize the process

具体步骤如下：

步骤 1 初始化 PSO 算法，包括种群大小 N，结束时间和其他参数等。并根据这些因素来确定每个粒子 i 的速度 $\boldsymbol{v}_i$ 与位置 $\boldsymbol{x}_i$。

$$\boldsymbol{v}_i=(v_{i1},v_{i2},\cdots,v_{iD}) \qquad i=1,2,\cdots,N \tag{12}$$

$$\boldsymbol{x}_i=(x_{i1},x_{i2},\cdots,x_{iD}) \qquad i=1,2,\cdots,N \tag{13}$$

步骤 2 定义适应度函数 $F(i)$，通过适应度函数可以评估每一个粒子 i 的解。

步骤 3 根据式 (14)更新个体最优解，根据式 (15)更新全局最优解。

$$\boldsymbol{p}_{\text{best}}=(p_{i1},p_{i2},\cdots,p_{iD}) \qquad i=1,2,\cdots,N \tag{14}$$

$$\boldsymbol{g}_{\text{best}}=(p_{g1},p_{g2},\cdots,p_{gD}) \qquad g=1,2,\cdots,N \tag{15}$$

式中：$\boldsymbol{p}_{\text{best}}$ 为第 i 个粒子群在进行迭代之前找出的最优解；$\boldsymbol{g}_{\text{best}}$ 为多次迭代后整个粒子群得到的最优解。

步骤 4 在更新粒子的速度与位置时，通过以下公式进行计算：

$$\boldsymbol{v}_{id}=w\boldsymbol{v}_{(i-1)d}+c_1r_1(\boldsymbol{p}_{(i-1)d})+c_2r_2(\boldsymbol{p}_{gd}-\boldsymbol{x}_{(i-1)d}) \tag{16}$$

$$\boldsymbol{x}_{id}=\boldsymbol{x}_{(i-1)d}+a\boldsymbol{v}_{(i-1)d} \tag{17}$$

式中：w 为惯性因子；c_1 和 c_2 分别为个体经验和群体经验的加速常速；r_1 和 r_2 为用于提高搜索的 [0,1] 随机性数字；α 为用于限制速度的约束因子。

步骤 5 循环步骤 2～步骤 4，直至满足优化终止条件，结束优化进程。

1.4 注意力机制

近年来，注意力机制已成为深度学习领域的一个研究热点，其源于对人类视觉注意的模拟[16]。注意力机制的核心思想在于权重的分配，即对重要信息赋予较高的权重，忽略无关信息，从而合理地改变外界对信息的关注。注意力机制的本质则是通过加权求和得到 Source 中的元素，Query 数据库和系数值用来计算相应的权重系数，如下：

$$\boldsymbol{A}(Q,S)=\sum_{L_x} i=1S_m(Q,k_i)\times V_i \tag{18}$$

式中：L_x 为数据源的长度；i 为权重系数。注意力机制的核心思想如图 5 所示。

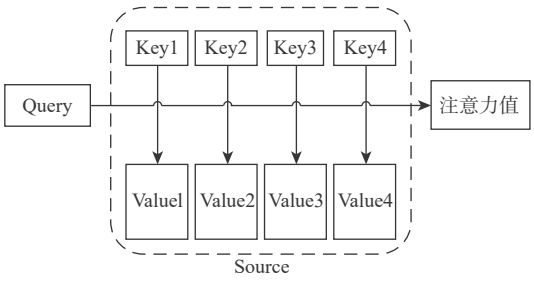


图 5　注意力机制的核心思想

Fig. 5　The core idea of attention mechanism

2　基于注意力机制改进的 PSO-BiLSTM 刀具磨损预测流程

基于注意力机制和 PSO-BiLSTM 的刀具磨损预测原理为: 首先通过使用传感器采集加工过程中 3 把刀具的切削信号和振动信号, 在此基础上将多域信号数据进行特征提取获得特征集。进一步利用 Kalman 滤波算法对特征提取的数据集进行特征融合, 从而在参数设置方面获得输入特征数, 采用标准的 PSO 算法, 同时增加了注意力机制, 减少后续模型的处理时间。根据式(19)所示的平均绝对百分比误差(mean absolute percentage error,MAPE)作为适应度函数, 对 BiLSTM 神经网络的隐藏层数和学习率参数进行优化。

$$M_{\mathrm{MAPE}}=\frac{1}{n}\sum_{i=1}^{n}\left|\frac{X_a(i)-X_p(i)}{X_a(i)}\right| \tag{19}$$

式中: n 为测试样本的长度; $X_a(i)$ 为第 i 时刻刀具磨损的真实情况; $X_{\mathrm{p}}(i)$ 为第 i 时刻刀具磨损的预测情况。表 1 为增加了注意力机制的 PSO-BiLSTM 模型的初始化参数设置。

表 1　PSO-BiLSTM 模型参数

Table 1　Parameters of the PSO-BiLSTM model

网络参数	初始值
初始化权重	[−0.5,0.5]
遗忘门偏置量	1
输入门偏置量	[0,0.8]
输出门偏置量	[0,0.8]
迭代次数	50
种群个数	5
学习因子c_1	2
学习因子c_2	2
隐藏层单元数m	[10,300]
学习率r	[0.01,0.15]

基于注意力机制改进的 PSO-BiLSTM 刀具磨损预测识别流程如下:

1)划分训练集和测试集。将传感器数据进行特征提取, 本文中将训练集分为 BiLSTM 训练集以及 PSO 寻优集, 目的是区别于传统 LSTM 方法中将数据集分为训练集及测试集, 从而形成一个完整的模型。

2) 多传感器数据融合。利用 Kalman 滤波对训练集进行数据融合, 通过算法进一步提高特征的鲁棒性。

3) 初始化 PSO 算法。在数据集分配结束后, 随机产生一组二维粒子种群, 包括初始学习率和 Dropout 概率。

4) 根据 BiLSTM 训练集和 PSO 寻优集, 可以在每一轮迭代中模拟预测过程。

5) 在 BiLSTM 网络进行反向传播算法来计算梯度时, 每个神经元的输出值是通过前向计算和反向计算的隐藏状态进行拼接而得到的。为了更好地进行标签的定位和分类, 所以设置注意力机制识别不同监测数据之间的区别, 从而在更深层次上提取数据中潜在的信息。

6) 学习率是最关键的超参数, 其次是隐藏层单元数。采用的 PSO 算法主要通过均方根误差函数搜索最佳适应度值下的最优学习率和隐藏单元层数。同时, 在优化算法中将时间窗引入数据集中, 提高数据处理的精度。采用 Adam 算法作为优化器, 更有效地优化模型权重, 以帮助模型更快地收敛并减少偏差, 参数 θ 遵循下式所示的更新规律。

$$\begin{cases} m_{mt}=\dfrac{m_t}{1-\beta_1^t} \\ v_{mt}=\dfrac{v_t}{1-\beta_2^t} \\ q_{i+1}=q_i-\dfrac{\eta}{\sqrt{v_{mt}}+\gamma}m_{mt} \end{cases} \tag{20}$$

式中: β_1为上一阶段的矩估计值的衰变速率; β_2为下一个阶段的矩估计指数的衰变速率; m_{mt} 为前一时刻的梯度平均值; v_{mt} 为后一时刻的梯度非中心方差值。本文将 β_1设置为 0.9, β_2设置为 0.999, γ 设置为 10^{-8}。

3　实验验证

3.1　刀具和工件材料

为了确保测量结果的准确性, 实验验证采用 PHM2010 数据挑战赛中的数据集[17]。铣削实验在松浦加工中心 MC-510V 设备上进行, 实验中机床型号为 Roders-TechRF, 刀具材料选用 HRC52 不锈钢, 刀具提供三齿碳化钨球头立铣刀。

切削力信号由安装在工作台和工件之间的测力仪采集, 振动信号通过安装在工件上的振动传感器进行采集, 信号采集频率设置为 50 kHz, 每次走刀采集 x、y、z 三向切削力和 x、y、z 三向振动共两

种原始信号，传感器安装位置及刀具磨损监控过程示意如图6所示。每次走刀完成后分别测得刀具三齿的磨损值，共计采集315次走刀样本。切削参数如表2所示。

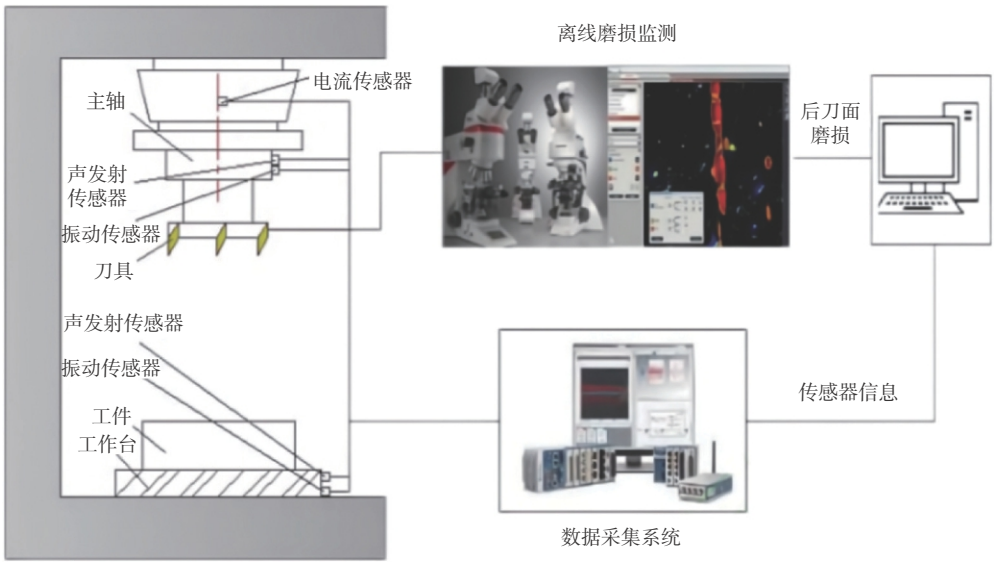


图6 刀具磨损系统示意

Fig. 6 Schematic diagram of the tool wear system

表2 实验加工参数

Table 2 Experimental processing parameters

主轴速度/($r\cdot min^{-1}$)	进给速度/($mm\cdot min^{-1}$)	径向切深/mm	轴向切深/mm	铣削方式
10 400	1 555	0.125	0.2	顺铣

为了尽可能提取到刀具磨损过程中隐含的信号特征，对每次走刀 x、y、z 这3个方向的前60 000个采样点数据进行特征提取。每次走刀振动信号可提取时域特征14个，频域特征6个，总共提取到时域特征和频域特征20个，3个方向共可获得60个特征，将其作为特征数据集，具体指标如表3所示。

表3 特征指标

Table 3 Characteristic index

序号	特征名称	序号	特征名称
1	峰峰值	11	峰值指标
2	方差	12	脉冲指标
3	均值	13	裕度指标
4	歪度	14	峭度指标
5	峭度	15	均值频率
6	均方值	16	频谱二阶矩
7	方根幅值	17	标准偏差频
8	均方根值	18	峭度频率
9	绝对均值	19	均方根频率
10	波形指标	20	中心频率

对采集到的传感器信号进行特征提取之后，通过计算各特征与刀具磨损的相关系数，剔除与刀具磨损相关性较低的特征。相关系数的计算式为

$$\begin{cases}\rho_{xy}=\dfrac{\mathrm{Cov}(x,y)}{s_x s_y}\\ \mathrm{Cov}(x,y)=E_{xy}-E_x\times E_y\end{cases}\tag{21}$$

式中：ρ 为相关系数，其值为[−1,1]。通常$|\rho|>0.8$表示为2个变量之间存在较高的相关性。根据各信号特征的相关系数，可以在降低特征向量维数的同时，综合利用多传感器信号在时域、频域、时频域的融合优势。最后选择降维指标为$|\rho|\geqslant 0.94$。

3.2 刀具磨损值预测结果评估

本文提出的刀具磨损预测方法是根据提出的神经网络对已知的数据进行拟合，然后通过刀具预测曲线预测刀具的剩余寿命。实验数据中，提取C1、C4和C6这3把刀具3个方向每次走刀60 000个信号，利用多域特征提取方法，直接从原信号中提取切削力信号和振动信号的特征。利用基于Kalman滤波的多传感器数据融合将提取的特征进行特征级融合。而切削力信号和振动信号的特征与刀具磨损趋势的关系并不明显，需要利用提出的PSO优化的BiLSTM模型自适应地挖掘隐藏在特征中的潜在特征，并利用注意力机制选择性地研究重要的时频域特征来预测刀具磨损。再分别使用未优化的LSTM模型和改进后的PSO-BiLSTM模型对不同参数下的刀具磨损值进行预测。未优化的LSTM模型在C1、C4和C6数据集上的预测结果如图7所示。

由图7可以明显看出，未经过PSO优化的LSTM未能将3把刀具的真实值准确地预测出来。优化的PSO-AM-BiLSTM模型在C1、C4和C6数据集上的预测结果如图8～图10所示。

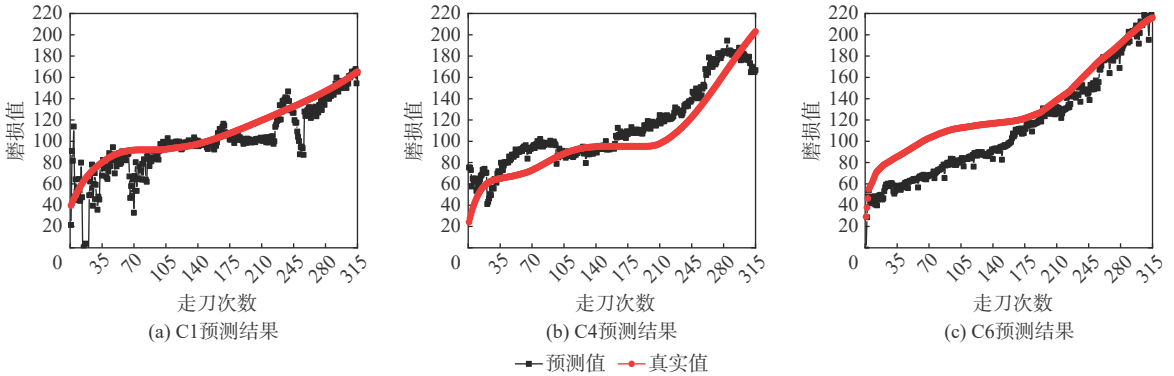


图 7　未优化的 LSTM 在 C1、C4 和 C6 上的预测效果

Fig. 7　Predictive effects of unoptimized LSTM on C1, C4 and C6

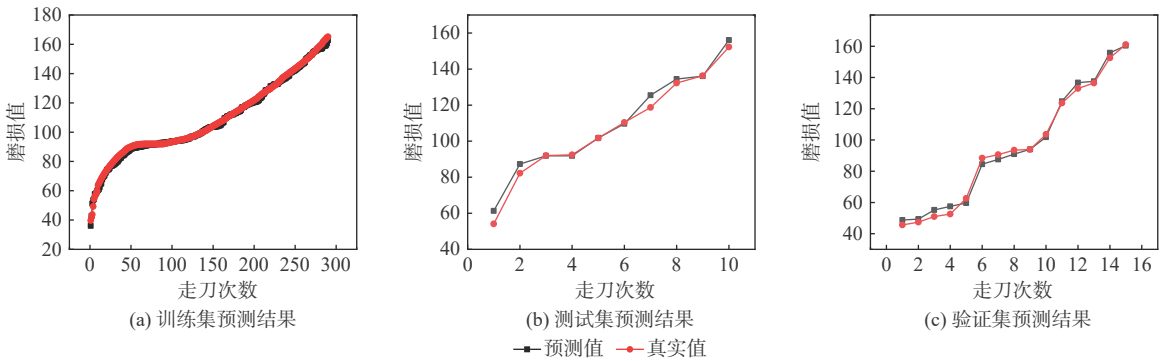


图 8　PSO-AM-BiLSTM 在 C1 刀具上的预测结果

Fig. 8　Prediction results of PSO-AM-BiLSTM on C1

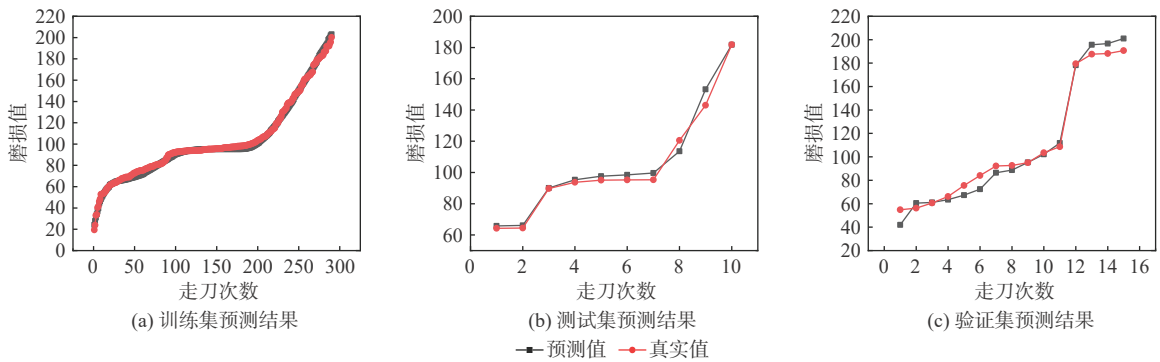


图 9　PSO-AM-BiLSTM 在 C4 刀具上的预测结果

Fig. 9　Prediction results of PSO-AM-BiLSTM on C4

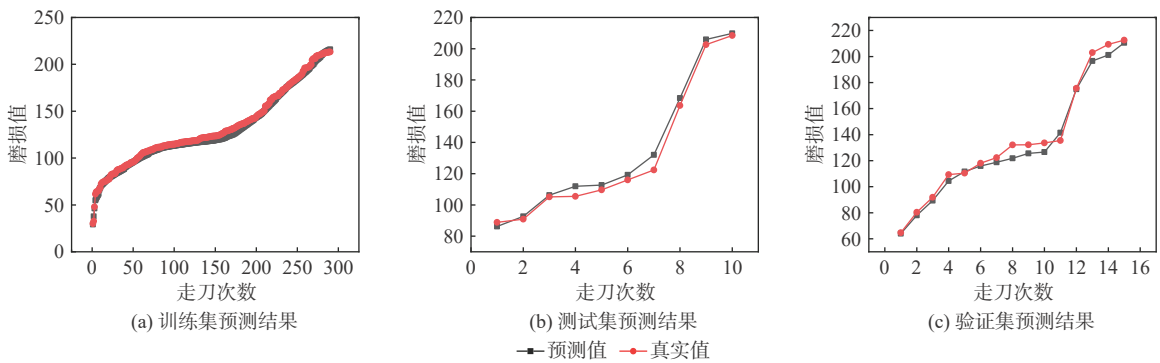


图 10　PSO-AM-BiLSTM 在 C6 刀具上的预测结果

Fig. 10　Prediction results of PSO-AM-BiLSTM on C6

可以看出, 经过改进的模型能够准确地预测刀具磨损曲线, 且其预测精度达到 99.34%。

以 C4 刀具的训练集预测结果为例, 在 0～100 次的区间, 刀具的磨损值开始增大, 这是因为刀具处于

磨损初期而导致的正常现象。在100～200次的区间，刀具磨损值变化率保持在5%以内，可以认为这一区间刀具磨损值保持不变，即处于刀具磨损的中期阶段。在200～300次的区间，刀具磨损进入稳定的退化期，变化率也趋于稳定，表明刀具正以相对稳定的速率进入磨损后期，直至损坏。即使在受到噪声的影响，特征融合之后模型的真实值和预测值的误差平均保持在5%以内，由此可以得出本文所提出的增加注意力机制的PSO-BiLSTM模型具有较好的预测效果。与未优化的模型相比，优化后的模型稳定性更好且精度更高。

为了验证方法的有效性，将增加了注意力机制的PSO-BiLSTM与PSO优化的CNN、反向传播神经网络(backpropagation neural network, BPNN)和LSTM方法进行对比。本文采用平均绝对百分比误差(mean absolute percentage error, MAPE)、均方根误差(root mean square error, RMSE)和决定系数R^2指标来量化估计不同深度学习的性能，如表4所示，增加了注意力机制的PSO-BiLSTM模型在C1、C4和C6测试数据集上的拟合能力明显优于PSO-LSTM模型、PSO-BPNN模型和PSO-CNN模型，MAPE和RMSE值最低。例如，与PSO-LSTM相比，改进的PSO-BiLSTM模型的RMSE分别减少了82.21%、55.14%和69.42%。与PSO-CNN模型相比，改进的PSO-BiLSTM模型的MAPE分别减少了48.94%、30.48%和36.74%，表明优化模型精度得到较大的提升。与PSO-BPNN模型相对比，改进的PSO-BiLSTM模型的R^2分别增长了12.47%、18.26%和16.33%。

为了避免实验的偶然性，进行消融实验来验证引入注意力机制的优势。选择未增加注意力机制的PSO-BiLSTM模型作为对比模型，2个模型的迭代次数都设置为30次，种群规模为50，模型收敛过程如图11所示。可以看出，增加了注意力机制的PSO-BiLSTM模型在全局都有很好的迭代效果，未增加注意力机制的PSO-BiLSTM虽然在前期具有较好的迭代效果，但是中后期还是无法避免局部寻优的局限性。结果证明增加了注意机制的PSO-BiLSTM模型不仅在前期有优秀的迭代效果，在后期也能更好地优化寻优结果。

表4 预测结果评价指标

Table 4 Evaluations of the prediction results

模型	RMSE			MAPE			R^2		
	C1	C4	C6	C1	C4	C6	C1	C4	C6
PSO-AM-BiLSTM	2.906	6.85	5.22	2.563	5.54	4.32	0.994	0.991	0.992
PSO-CNN	6.72	8.66	8.31	5.02	7.97	6.83	0.91	0.907	0.903
PSO-BPNN	11.384	13.027	17.51	10.34	12.76	11.85	0.87	0.81	0.83
PSO-LSTM	16.34	15.27	17.08	12.17	13.03	14.05	0.75	0.85	0.80

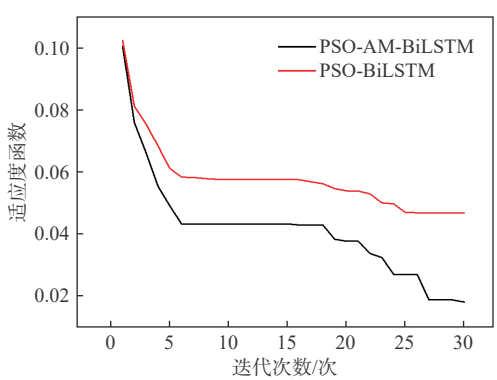


图11 算法迭代过程

Fig. 11 Process of optimization algorithms

通过以上结果发现，增加了注意力机制的PSO-BiLSTM模型可以有效地捕捉时间序列中的时频特性，相比未经优化的神经网络模型具有更好的预测效果。这表明增加了注意力机制的PSO-BiLSTM模型的刀具磨损预测结果更加符合实用性的标准。

4 结 论

本文提出了一种基于注意力机制改进的PSO-BiLSTM模型对刀具磨损寿命进行预测的方法，该方法以多传感器数据作为输入，进一步通过PSO算法对BiLSTM神经网络模型进行了优化。基于注意力机制对融合特征的重要性进行过滤，保留了长时信息的特征，实现了一种端到端的刀具磨损预测。并与未优化的深度学习进行了对比，可以得出以下结论：

1) 对多个传感器信号进行特征级数据融合，提高了预测模型的精度，发挥了多传感器信号的优势，这样既提高了工作效率，又增强了模型的泛化性。

2) 通过对比RMSE和MAPE等指标发现，采用基于注意力机制的BiLSTM网络能够在小规模样本下训练刀具磨损的预测，诊断效果好于传统的神经网络。

3) 将注意力机制引入 BiLSTM 网络中, 可以选择性地学习模型训练过程中的特征, 挖掘数据中的隐藏信息, 更好地预测刀具磨损寿命。

4) 通过对比相关系数后可以发现, 增加了注意力机制的 PSO-BiLSTM 神经网络相较于未优化的深度学习在预测上表现更出色, 极大地改善了 BiLSTM 模型在学习效率方面的泛化性与鲁棒性。

研究结果对于铣削加工的刀具磨损预测具有较强的参考意义。在实际加工过程中, 该方法还可以方便地应用于车削、钻孔等其他加工过程中的刀具磨损预测。本文方法目前只在小样本条件下完成验证, 今后可以扩大实验样本, 应用于不同刀具磨损预测技术中; 还可以尝试利用原始信号直接预测刀具磨损, 进一步提高预测的准确度和泛化能力。

参考文献 (References)

[1] CHEN Y X, JIN Y, JIRI G. Predicting tool wear with multi-sensor data using deep belief networks[J]. The International Journal of Advanced Manufacturing Technology, 2018, 99(5): 1917-1926.

[2] 谢振龙, 岳彩旭, 刘献礼, 等. 基于 EMD-SVM 的钛合金铣削过程刀具磨损监测[J]. 振动、测试与诊断, 2022, 42(5): 988-996.
XIE Z L, YUE C X, LIU X L, et al. Tool wear monitoring based on EMD-SVM in milling process of Ti-alloy[J]. Journal of Vibration, Measurement & Diagnosis, 2022, 42(5): 988-996(in Chinese).

[3] ZHOU J J, YU J B. Chisel edge wear measurement of high-speed steel twist drills based on machine vision[J]. Computers in Industry, 2021, 128: 103436.

[4] LI L H, AN Q B. An in-depth study of tool wear monitoring technique based on image segmentation and texture analysis[J]. Measurement, 2016, 79(2): 44-52.

[5] MORGAN J, O'DONNELL G E. Cyber physical process monitoring systems[J]. Journal of Intelligent Manufacturing, 2018, 29(6): 1317-1328.

[6] 戴稳, 张超勇, 孟磊磊, 等. 基于深度学习与特征后处理的支持向量机铣刀磨损预测模型[J]. 计算机集成制造系统, 2020, 26(9): 2331-2343.
DAI W, ZHANG C Y, MENG L L, et al. Support vector machine milling wear prediction model based on deep learning and feature re-processing[J]. Computer Integrated Manufacturing Systems, 2020, 26(9): 2331-2343(in Chinese).

[7] 郑圆, 胡建中, 贾民平, 等. 一种基于参数优化变分模态分解的滚动轴承故障特征提取方法[J]. 振动与冲击, 2020, 39(21): 195-202.
ZHENG Y, HU J Z, JIA M P, et al. A method for rolling bearing fault feature extraction based on parametric optimization VMD[J]. Journal of Vibration and Shock, 2020, 39(21): 195-202(in Chinese).

[8] 李聪波, 万腾, 陈行政, 等. 基于切削功率的数控车削批量加工刀具磨损在线监测[J]. 计算机集成制造系统, 2018, 24(8): 1910-1919.
LI C B, WAN T, CHEN X Z, et al. On-line monitoring method of tool wear for NC turning in batch processing based on cutting power[J]. Computer Integrated Manufacturing Systems, 2018, 24(8): 1910-1919(in Chinese).

[9] 何彦, 凌俊杰, 王禹林, 等. 基于长短时记忆卷积神经网络的刀具磨损在线监测模型[J]. 中国机械工程, 2020, 31(16): 1959-1967.
HE Y, LING J J, WANG Y L, et al. In-process tool wear monitoring model based on LSTM-CNN[J]. China Mechanical Engineering, 2020, 31(16): 1959-1967(in Chinese).

[10] 曹大理, 孙惠斌, 张纪铎, 等. 基于卷积神经网络的刀具磨损在线监测[J]. 计算机集成制造系统, 2020, 26(1): 74-80.
CAO D L, SUN H B, ZHANG J D, et al. In-process tool condition monitoring based on convolution neural network[J]. Computer Integrated Manufacturing Systems, 2020, 26(1): 74-80(in Chinese).

[11] 王丽华, 杨家巍, 张永宏, 等. 基于堆叠降噪自编码的刀具磨损状态识别[J]. 中国机械工程, 2018, 29(17): 2038-2045.
WANG L H, YANG J W, ZHANG Y H, et al. Tool wear condition recognition based on SDAE[J]. China Mechanical Engineering, 2018, 29(17): 2038-2045(in Chinese).

[12] LUO J H, ZHANG X. Convolutional neural network based on attention mechanism and Bi-LSTM for bearing remaining life prediction[J]. Applied Intelligence, 2022, 52(1): 1076-1091.

[13] ALJAMAL M A, ABDELGHAFFAR H M, RAKHA H A. Developing a neural-Kalman filtering approach for estimating traffic stream density using probe vehicle data[J]. Sensors, 2019, 19(19): 4325.

[14] 沈涛, 李舜酩. 针对滚动轴承故障的批标准化 CNN-LSTM 诊断方法[J]. 计算机集成制造系统, 2022, 28(12): 3946-3955.
SHEN T, LI S M. CNN-LSTM method with batch normalization for rolling bearing fault diagnosis[J]. Computer Integrated Manufacturing Systems, 2022, 28(12): 3946-3955(in Chinese).

[15] 吴飞, 农皓业, 马晨浩. 基于 PSO-LSTM 模型的刀具磨损预测方法[J]. 吉林大学学报 (工学版), 2023, 53(4): 984-997.
WU F, NONG H Y, MA C H. Tool wear predictionmethod based on PSO-LSTM model [J]. Journal of Jilin University(Engineering and Technology Edition), 2023, 53(4): 984-997(in Chinese).

[16] ZHANG J S, JIANG Y C, WU S M, et al. Prediction of remaining useful life based on bidirectional gated recurrent unit with temporal self-attention mechanism[J]. Reliability Engineering & System Safety, 2022, 221: 108297.

[17] PHM Society. PHM data challenge 2010[DB/OL]. (2023-07-25) [2010-05-18]. https://www.phmsociety. org/competition/phm/10.

Tool wear prediction based on attention mechanism and PSO-BiLSTM

YANG Peidong，HUANG Hua*，YU Weiwei，GUO Baodao

(School of Mechanical and Electrical Engineering，Lanzhou University of Technology，Lanzhou 730050，China)

Abstract：A bi-directional long short-term memory (BiLSTM) neural network optimization end-to-end tool wear prediction method based on attention mechanism (AM) and particle swarm optimization (PSO) algorithm is proposed to address the issues of single monitoring data and poor feature signal processing in tool wear fault diagnosis. Firstly, based on the sensor data, to construct high-quality signal input samples, extract the multi-domain feature. With multi-sensor data to obtain the fused data samples with higher robustness, the Kalman filter is used to fuse the input samples. On this basis, the hyperparameters of the BiLSTM are optimized by PSO, and the neural network model is built according to the optimized hyperparameters. Finally, based on the attention mechanism to give weights to the input influences, the PSO-BiLSTM is improved to obtain a better tool wear prediction. The suggested model's validity and feasibility in tool wear prediction are confirmed by comparative experimental findings, and its accuracy is significantly higher than that of classical deep learning. This approach offers a fresh viewpoint on tool wear prediction.

Keywords：tool wear；Kalman filter；particle swarm optimization；attention mechanism；bi-diretional long short-term memory neural network

Received：2023-08-25；**Accepted**：2023-11-13；**Published Online**：2023-12-13 12：12

URL：link.cnki.net/urlid/11.2625.V.20231212.1424.001

Foundation items：National Natural Science Foundation of China (52365057,51965037)

* **Corresponding author**. E-mail：hh318872@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技术核心期刊,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,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,公式符号标准规范,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(作者顺序始终不得更改),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,引言,正文,参考文献,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。首、末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,作者有权改投其他刊。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西配楼412~416房间
电　　话:(010)82316698,82316981,82315499,82317448
E-mail: jbuaa@ buaa. edu. cn
http://bhxb. buaa. edu. cn
http://www. buaa. edu. cn

集群成员期刊

《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

编辑部主任：张　嵘

编　　辑：孙　芳　万玉敏　王　茜　卞欢欢　李亚泰

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51卷 第10期　2025年10月

JOURNAL OF BEIJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS (JBUAA)

(Monthly, Started in 1956)

Vol.51 No.10　October 2025

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　100191
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) 82316698 82316981
82315499 82317448
电子信箱　jbuaa@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 100191, P. R. China)
Chief Editor　ZHAO Qinping
Edited and Published by　Editorial Board of JBUAA
Printed by　Beijing Kexin Printing Co., Ltd.
Distributed by　Editorial Board of JBUAA
Telephone　(010) 82316698 82316981
82315499 82317448
E-mail　jbuaa@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

ISSN 1001-5965

微信公众号　超星电子书

国内定价：50.00元／期