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

2024-01
Vol.50 No.1

北 京 航 空 航 天 大 学 学 报

第 50 卷　第 1 期　（总第 371 期）　2024 年 1 月

目　　次

低轨互联网星座发展研究 …… 吴树范，王伟，温济帆，吴岳东（1）

锂离子电池热失控气体毒性风险分析方法 …… 张青松，曲奕润，刘添添（12）

ML 组合的 CYGNSS 海面风速反演质量控制模型 ……
……张云，赵星宇，杨树瑚，孙聪，韩彦岭，尹继伟（20）

改进迁移学习的双分支卷积神经网络图像去雾 ……
…… 李云红，于惠康，马登飞，苏雪平，段娇娇，史含驰（30）

基于截断 p-shrinkage 范数的航空发动机数据重构 …… 张红梅，武江南，赵永梅，曾航，李全根（39）

多约束条件下四旋翼机动轨迹跟踪一体化控制方法 …… 王英勋，李欣，蔡志浩，赵江（48）

基于虚拟协商与滚动时域优化的分布式协同制导策略 …… 杨秀霞，姜子劼，张毅，王聪，于浩（61）

基于 QAR 数据的民机高高原进近着陆风险评估方法 …… 陈农田，满永政，李俊辉（77）

基于协同滤波轨迹预测的机动目标 RTPN 拦截制导律 ……
…… 李继广，陈欣，董彦非，屈高敏，赵成功，张阿龙（86）

无人机集群分组编队自适应跟踪-包含控制 …… 张毅，于浩，杨秀霞，姜子劼，王晨蕾（97）

恶劣天气下机场离场航班运行韧性评估及恢复 …… 王兴隆，赵俊妮，王进（110）

基于图对比学习的稳健交通流量预测 …… 刘伟，贾素玲（122）

星载 SAR 天线展开机构装配精度确信可靠性建模 …… 李云，张湘霄，孙富强，李晓阳，韩建超（134）

基于动态视觉传感器的无人机目标检测与避障 …… 蔡志浩，陈文军，赵江，王英勋（144）

板翅式换热器传递函数动态模型及参数确定方法 …… 梁兴壮，黄志远，艾凤明，袁振伟，汪箭（154）

电动静液作动器的非线性变阻尼积分滑模控制 …… 杨荣荣，张玲，赵家黎，付永领，张朋（163）

过冷大水滴动力学行为对翼型结冰的影响分析 …… 刘飞宇，邓甜（173）

升空平台相对测量误差对定位精度的影响及定位算法 ……
…… 欧阳晓凤，曾芳玲，吕大千，董天宝，韩宜静（187）

纤维排布对 CFRP 感应加热固化成型过程影响表征 ……
……付天宇，杨宁，古云飞，许家忠，蒋悦，李斌（198）

三维自由弯曲过程中临界偏距对管材成形精度影响 ……
……范鑫，程宗辉，李少鑫，舒送，程诚，郭训忠（208）

广义多边形凸包弹性线递支模拟算法 …… 崔钰萍，李子涵，郑国磊（216）

飞机地面除冰保障过程动态预测 …… 李彪，邢志伟，王立文（224）

自抗扰控制在陀螺电机稳速控制中的应用 …… 关伊哲，冯明（234）

空间可展开天线铰链热变形分析与试验验证 …… 杨智杰，张从发，赵瑞杰，王春洁，赵军鹏（243）

拓扑优化相变热沉在不同重力场的性能 …… 王仕淞，殷健宝，邢玉明，侯煦，王子贤（250）
基于预设性能的飞机全电刹车系统滑模控制 …… 谢明军，段京奇，马文冉，贾玉红（260）
基于平滑矩阵集优化的双基地 MIMO 雷达相干目标 DOD 和 DOA 联合估计 ……
…… 游致远，胡国平，周豪，郑桂妹（268）
基于飞机状态的备件动态规划技术 …… 康子旭，周栋，李会欣，郭子玥，陈承璋，宋子骋（276）
融合边界处理机制的学习型麻雀搜索算法 …… 王子恺，黄学雨，朱东林，闫少强，李权，郭伟（286）
LS-SVM 和抗差估计的 GNSS/INS 紧组合欺骗检测算法 ……
…… 柯晔，吕志伟，周玟龙，邓旭，商向永，武文博（299）
月基装备行走运动学分析与连续步态规划 …… 何天宇，董洋，檀傈锰，王岩，赵军鹏，王春洁（308）
基于领域概念图的航天新闻自动摘要模型 …… 黄浩宁，陈志敏，徐聪，张晓燕（317）
基于隐含非线性维纳退化过程的剩余寿命预测 ……
……杨家鑫，唐圣金，李良，孙晓艳，祁帅，司小胜（328）
爆炸冲击波作用后变形机翼模态数值模拟 …… 肖良丰，周兰伟，李向东（341）
2195-T8 铝锂合金疲劳多裂纹融合试验与仿真研究 …… 庞闯，刘德俊，田干，郭伟，刘德博（350）

期刊基本参数：CN 11-2625/V * 1956 * m * A4 * 358 * zh * P * ￥50.00 * 400 * 35 * 2024-01

（**编　辑**　张　嵘　孙　芳　苏　磊　卞欢欢　王　茜　李亚泰）

JOURNAL OF BEIJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS

Vol. 50 No. 1 （Sum 371） January 2024

CONTENTS

Review on development of LEO Internet constellation
……WU Shufan，WANG Wei，WEN Jifan，WU Yuedong （1）

Risk analysis method for thermal runaway gas toxicity of lithium-ion batteries
…… ZHANG Qingsong，QU Yirun，LIU Tiantian （12）

Quality control model of CYGNSS sea surface wind speed retrieval based on ML combination
…… ZHANG Yun，ZHAO Xingyu，YANG Shuhu，SUN Cong，HAN Yanling，YIN Jiwei （20）

Improved transfer learning based dual-branch convolutional neural network image dehazing
…… LI Yunhong，YU Huikang，MA Dengfei，SU Xueping，DUAN Jiaojiao，SHI Hanchi （30）

Aero-engine data reconstruction based on truncated p-shrinkage norm
…… ZHANG Hongmei，WU Jiangnan，ZHAO Yongmei，ZENG Hang，LI Quangen （39）

Integrated control method for quadrotors' aggressive trajectory tracking under multiple constraints
…… WANG Yingxun，LI Xin，CAI Zhihao，ZHAO Jiang （48）

Distributed cooperative guidance strategy based on virtual negotiation and rolling horizon optimization
…… YANG Xiuxia，JIANG Zijie，ZHANG Yi，WANG Cong，YU Hao （61）

Risk assessment method for civil aircraft approach and landing at high plateau based on QAR data
…… CHEN Nongtian，MAN Yongzheng，LI Junhui （77）

RTPN interception guidance law for maneuvering target based on collaborative filtering trajectory prediction
…… LI Jiguang，CHEN Xin，DONG Yanfei，QU Gaomin，ZHAO Chenggong，ZHANG Along （86）

Adaptive group formation tracking-containment control for UAV swarm
…… ZHANG Yi，YU Hao，YANG Xiuxia，JIANG Zijie，WANG Chenlei （97）

Resilience assessment and recovery of airport departure flights under severe weather
…… WANG Xinglong，ZHAO Junni，WANG Jin （110）

Robust traffic flow prediction based on graph contrastive learning
…… LIU Wei，JIA Suling （122）

Belief reliability modeling for assembly accuracy of spaceborne SAR antenna deployable mechanism
…… LI Yun，ZHANG Xiangxiao，SUN Fuqiang，LI Xiaoyang，HAN Jianchao （134）

Object detection and obstacle avoidance based on dynamic vision sensor for UAV
…… CAI Zhihao，CHEN Wenjun，ZHAO Jiang，WANG Yingxun （144）

Dynamic model by transfer function and parameter determination method of plate fin heat exchanger
…… LIANG Xingzhuang，HUANG Zhiyuan，AI Fengming，YUAN Zhenwei，WANG Jian （154）

Nonlinear variable damping integral sliding mode control for electro-hydrostatic actuator
…… YANG Rongrong，ZHANG Ling，ZHAO Jiali，FU Yongling，ZHANG Peng （163）

Influence of dynamic behavior of supercooled large droplets on airfoil icing
…… LIU Feiyu，DENG Tian （173）

Positioning accuracy and localization algorithm with relative measurement errors in blast-off platforms
…… OUYANG Xiaofeng，ZENG Fangling，LYU Daqian，DONG Tianbao，HAN Yijing （187）

Characterization of influence of fiber arrangement on CFRP induction heating curing process
…… FU Tianyu，YANG Ning，GU Yunfei，XU Jiazhong，JIANG Yue，LI Bin （198）

Effect of critical eccentricity on forming accuracy of tubes in 3D free bending process
…… FAN Xin，CHENG Zonghui，LI Shaoxin，SHU Song，CHENG Cheng，GUO Xunzhong （208）

Computing convex hull of a generic polygon with simulation of progressive support for an elastic line
…… CUI Yuping，LI Zihan，ZHENG Guolei （216）

Dynamic prediction for aircraft ground deicing operation process
…… LI Biao，XING Zhiwei，WANG Liwen （224）

Application of active disturbance rejection control in gyro motor steady speed control
…… GUAN Yizhe，FENG Ming （234）

Thermal deformation analysis and experimental verification of spatial deployable antenna hinge
…… YANG Zhijie，ZHANG Congfa，ZHAO Ruijie，WANG Chunjie，ZHAO Junpeng （243）

Topological optimzation of phase change heat sink performance in different gravity fields
…… WANG Shisong，YIN Jianbao，XING Yuming，HOU Xu，WANG Zixian （250）

Sliding mode control for electric braking systems of aircraft based on prescribed performance
…… XIE Mingjun，DUAN Jingqi，MA Wenran，JIA Yuhong （260）

Joint DOA and DOD estimation of bistatic MIMO radar coherent targets based on smoothing matrix sets optimization
…… YOU Zhiyuan，HU Guoping，ZHOU Hao，ZHENG Guimei （268）

Dynamic planning technology of civil spare parts based on aircraft state
…… KANG Zixu，ZHOU Dong，LI Huixin，GUO Ziyue，CHEN Chengzhang，SONG Zicheng （276）

Learning sparrow search algorithm of hybrids boundary processing mechanisms
…… WANG Zikai，HUANG Xueyu，ZHU Donglin，YAN Shaoqiang，LI Quan，GUO Wei （286）

Tightly-coupled GNSS/INS spoofing detection algorithm for LS-SVM and robust estimation
…… KE Ye，LYU Zhiwei，ZHOU Wenlong，DENG Xu，SHANG Xiangyong，WU Wenbo （299）

Kinematic analysis and continuous gait planning of lunar-based equipment in walking state
…… HE Tianyu，DONG Yang，TAN Limeng，WANG Yan，ZHAO Junpeng，WANG Chunjie （308）

Automatic summarization model of aerospace news based on domain concept graph
…… HUANG Haoning，CHEN Zhimin，XU Cong，ZHANG Xiaoyan （317）

Remaining useful life prediction based on implicit nonlinear Wiener degradation process
…… YANG Jiaxin，TANG Shengjin，LI Liang，SUN Xiaoyan，QI Shuai，SI Xiaosheng （328）

Numerical simulation of deformed airfoil modal after blast shock wave
…… XIAO Liangfeng，ZHOU Lanwei，LI Xiangdong （341）

Experimental and simulation study on fatigue multi crack fusion of 2195-T8 Al-Li alloy
…… PANG Chuang，LIU Dejun，TIAN Gan，GUO Wei，LIU Debo （350）

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0242

低轨互联网星座发展研究

吴树范，王伟*，温济帆，吴岳东
(上海交通大学 航空航天学院，上海 200240)

摘　　要：近年来，随着互联网用户激增，SpaceX、OneWeb等创新型企业纷纷计划打造低轨互联网星座，引发全球低轨卫星互联网星座的发展热潮。因此，研究国内外低轨互联网星座的发展情况及当前面临的技术难题有着很重要的意义。基于此，介绍了国外具有代表性的3个低轨互联网星座计划(Starlink、OneWeb、Lightspeed)及国内星座的最新发展情况；根据低轨互联网星座的特点，着重分析了低轨互联网星座面临的五大技术挑战：星座相对运动演化预报、星座自主导航与定轨、星座构型重构的路径规划、星座自组织协同构型控制、星座通信与网络服务；根据当前低轨互联网星座的发展趋势，对中国在低轨互联网星座的发展给出了建议。

关 键 词：低轨互联网星座；星座相对运动演化预报；星座自主导航与定轨；星座构型重构的路径规划；星座自组织协同构型控制

中图分类号：V221^{+}.3; TB553

文献标志码：A　　**文章编号**：1001-5965(2024)01-0001-11

国际电信联盟(International Telecommunication Union, ITU)的最新数据显示，全球互联网使用量强劲增长，截至2021年11月，全球互联网用户已达49亿，占世界人口的比重达63%。造成互联网用户激增的原因，除了发展中国家的快速增长，更主要是受新冠疫情(COVID-19)的影响。

虽然现在世界上几乎2/3的人口已经上网，但在全球范围内依旧有29亿人处于离线状态。居住地距离任何光纤电缆基础设施超过10 km的地方，全世界有超过50亿人；而距离超过25 km的地方，则有36亿人。受制于经济发展水平及铺设技术，目前，全世界互联网土地覆盖率仅有1/5。相比之下，卫星通信的广覆盖、不受空间限制、受自然灾害影响小等特征，使其可以作为地面通信的有效补充，为陆地、海洋和天空等各类用户提供互联网服务[1]。

与此同时，随着近年来各项技术的迅速发展，商业航天开始逐步走入人们的视野。其中，卫星制造备受关注。而近地球轨道(low Earth orbit, LEO)凭借其轨道高度低、传输延时短等特点备受互联网星座运营商的青睐，掀起了低轨互联网星座的浪潮。

本文紧随低轨互联网星座浪潮，对当前具有代表性的低轨互联网星座进行了介绍，分析了低轨互联网星座面临的五大技术挑战，并对中国在低轨互联网星座的发展给出了建议。

1　低轨互联网星座发展现状

低轨互联网星座是大规模星群的一类典型形式，由按照预定构型运行在不同轨道平面、高度和相位的数百、数千甚至上万颗卫星构成，是对地遥感、互联网通信等空间应用系统的重要发展方向，对社会经济、国家安全具有重大战略意义，成为当前各国争相研究和部署的热点。

1.1　国外低轨互联网星座

为全球宽带服务设计LEO星座一直是卫星运

收稿日期：2022-04-11；**录用日期**：2022-08-11；**网络出版时间**：2022-08-19 09:27
网络出版地址：link.cnki.net/urlid/11.2625.V.20220818.1438.001
基金项目：国家自然科学基金(11902175,U20B2056)
*通信作者. E-mail：wangwei215661@sjtu.edu.cn

引用格式：吴树范，王伟，温济帆，等. 低轨互联网星座发展研究[J]. 北京航空航天大学学报，2024，50(1)：1-11.
WU S F，WANG W，WEN J F，et al. Review on development of LEO Internet constellation[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(1)：1-11 (in Chinese).

营商的目标。早在20世纪90年代，摩托罗拉公司提出的“铱星星座”（Iridium satellite constellation）便开启了利用非地球同步卫星轨道 (non-geostationary satellite orbit, NGSO) 在太空提供通信服务的先河。然而，由于技术的不成熟及没有可行的盈利模式，该项目最终破产。经过20年的发展，全球通信、卫星制造、火箭发射等相关技术得到了长足发展，低轨卫星互联网走入大众视野，而这些也重新燃起了运营商从太空覆盖全球互联网的信心。当前，世界各国纷纷提出了自己的卫星互联网建设计划，本文将介绍其中最具代表性的3个星座。

1.1.1 Starlink 低轨互联网星座

Starlink 低轨互联网星座计划由 SpaceX 公司于2015年提出，将为全世界大部分地区提供高速的互联网，包括偏远地区。Starlink 低轨互联网星座计划预计将4.2万颗低轨卫星送上太空，其星座构型如图1所示。目前，已经开始实施的是第1阶段，第1阶段的1.2万颗卫星由分布在1 150 km高度的4 425颗低轨卫星和分布在340 km左右的7 518颗甚低轨卫星构成[1]。

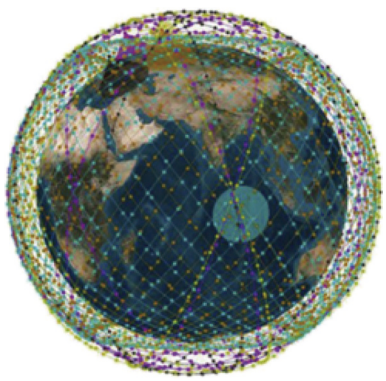


图1 Starlink 星座[1]

Fig. 1 Starlink constellation[1]

2021年，Starlink 低轨互联网星座计划共进行19次发射，目前在轨卫星总数接近1 900颗。性能上，对地终端宽带通信的实际网速达到约150 Mbit/s，不逊于地面光纤网速，并且具有星间激光通信的卫星开始常态化批量入轨[2]。应用上，Starlink 的测试版服务产品已在全球21个国家或地区推出。2022年，Starlink 低轨互联网星座计划进入两周一次的高密度发射期，所使用的猎鹰9号火箭一次可以将60颗卫星送入轨道，使得 Starlink 低轨互联网星座计划在更多国家开通服务成为可能，而且新一代卫星将配备星间链路，能够有效覆盖极地地区[3]。

1.1.2 OneWeb 低轨互联网星座

OneWeb 低轨互联网星座的市场主要面向企业和政府（包括电话网络运营商）。OneWeb 几经转手，目前由英国政府和印度 Bharti 公司联合收购。根据最新公布，该星座由648颗工作于 Ku/Ka 频段的低轨卫星构成，分布在轨道高度为1 200 km、倾角为87.9°的18个轨道面上，星座容量达到7 Tbit/s[4]，如图2所示。

2021年，OneWeb 计划共完成8次发射，截止2022年6月，在轨卫星数量达到428颗，占整个计划的60%以上。目前，OneWeb 的生产工厂达到了每天2颗卫星的生产速度[5]。



图2 OneWeb 星座[4]

Fig. 2 OneWeb constellation[4]

1.1.3 Lightspeed 星座计划

Lightspeed 星座计划由 Telesat 公司于2016年提出，并于2018年发射了该计划的首颗原型卫星。近年来，受各种原因影响，计划进展缓慢。2021年2月，为推动 Lightspeed 星座计划顺利进行，Telesat 公司宣布与欧洲 Thales Alenia Space 公司合作。该星座由298颗卫星构成，其中，78颗将部署到高度为1 015 km的极轨道，其余220颗将发射到高度为1 325 km的倾斜轨道，最终形成如图3所示的星座[6]。

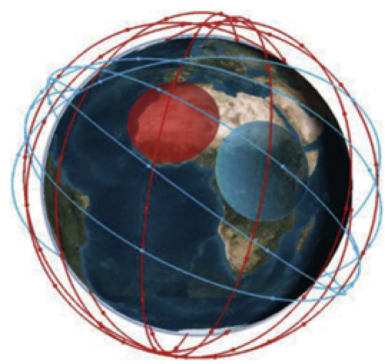


图3 Lightspeed 星座[6]

Fig. 3 Lightspeed constellation[6]

1.2 国内低轨互联网星座

作为“新基建”的七大重点领域，中国的低轨卫星互联网产业进入了国家统一统筹、规划和运营的新阶段。此外，随着中国卫星导航年会等学术会

议、展览活动的召开, 卫星互联网产业在国内饱受瞩目。2021 年 4 月 28 日挂牌成立的中国卫星网络集团有限公司更是标志着建设低轨卫星星座已经上升到国家意志, 开启了全面快速发展的时代。其中, 具有代表性的星座计划如表 1 所示。

表 1　国内代表性低轨互联网星座

Table 1　Representative LEO Internet constellations in China

单位	卫星星座	卫星数量/颗	已发射卫星数量/颗
中国航天科技集团公司	鸿雁星座	300	1
中国航天科工集团有限公司	虹云工程	156	1
中国航天科工集团有限公司	行云工程	80	2
中国电子科技集团有限公司	天象星座	120	2
银河航天(北京)科技有限公司	银河5G	650	1
北京未来导航科技有限公司	微厘空间	120	1

1.2.1　鸿雁星座

鸿雁星座全称为鸿雁全球卫星星座通信系统, 由中国航天科技集团公司于 2016 年提出。该系统计划由 300 颗低轨道小卫星及全球数据业务处理中心组成, 拥有多个 1 000 km 左右的轨道面。鸿雁星座具有星间链路, 能够使全球数据均可回传国内落地, 从而全面保证海内外的用户数据安全。首颗试验星于 2018 年成功发射。预计将建设成骨干星座系统, 为全球用户提供服务, 为国家战略和百姓民生提供一个完全自主可控的低轨信息网络保障。

1.2.2　虹云工程和行云工程

虹云工程共包含 156 颗卫星。该星座的首颗卫星于 2018 年 12 月完成验证, 待全部卫星升空后, 该星座将能够为广大用户提供高速的宽带移动通信服务。

行云工程计划由 80 颗行云小卫星组成系统, 将分为 α、β、γ 3 个建设阶段。其中, α 阶段计划组成行云二号 1 星和 2 星的系统, 同步开展试运营工作和示范工程建设; β 阶段计划完成小规模组网工程; γ 阶段将完成系统的完全构建, 为“一带一路”倡议提供服务和辅助。

1.2.3　银河 5G

银河 5G 由国内民营商业航天企业银河航天(北京)科技有限公司提出, 计划发射 650 颗小卫星, 通过研发高性能、低成本的 5G 卫星, 打造一个低轨宽带通信卫星星座。

2020 年 9 月, 中国以“GW”为代号正式向 ITU 提交了 2 个低轨互联网星座的轨道和频率申请资料, 2 个星座总计 12 992 颗卫星, 星座构型分布如表 2 所示[6]。

表 2　“GW”星座构型分布[6]

Table 2　“GW” constellation configuration distribution[6]

星座计划	星座子计划	轨道高度/km	轨道倾角/(°)	轨道面卫星数	轨道面	卫星总数/颗
GW-A59星座	GW-A59/1	590	85	16	30	6 080
	GW-A59/2	600	50	40	50	
	GW-A59/3	508	55	60	60	
GW-2星座	GW-2/1	1 145	30	48	36	6 912
	GW-2/2	1 145	40	48	36	
	GW-2/3	1 145	50	48	36	
	GW-2/4	1 145	60	48	36	

2　低轨互联网星座技术挑战

低轨互联网星座作为近年来新出现的一个庞大的空间系统, 具有距地近、信号优、低延时等优点。国际上, 对其自主协同感知导航、路径规划和构型控制的研究处于起步阶段, 直接相关的可供借鉴的研究较为有限。针对低轨互联网星座, 有以下 5 个方面的问题值得关注: 星座相对运动演化预报、星座自主导航与定轨、星座构型重构的路径规划、星座自组织协同构型控制、星座通信与网络服务。

2.1　星座相对运动演化预报

相对运动演化预报是星座动态拓扑切换控制、构型维持控制的前提。相关研究重点在于 3 方面: 星座相对运动解析建模、兼顾精度与计算量的在线预报方法、预报模型与方法的鲁棒性评估。

相对运动模型直接决定了星座运动与拓扑耦合推演的精度, 从而影响大规模星座稳定性分析和运行维持代价。相对运动建模从原理上可以分为动力法和几何法 2 类。动力法是以轨道动力学方程为出发点, 通过动力学分析建立相对运动模型, 最典型的是针对圆参考轨道、近距离相对运动提出的 HCW 方程。后续学者针对任意偏心率参考轨道和考虑 J_2 摄动、大气阻力摄动开展了拓展性研究[7]。然而, 由于建模原理的限制, 基于动力法的相对运动模型主要针对近距离编队, 其解析解难以分析复杂摄动影响。几何法以轨道根数为出发点, 通过构型的空间几何关系得到相对运动方程。Alfriend 等率先以几何法建立的相对运动模型为基础, 针对近圆、椭圆、开普勒、非开普勒轨道开展研究[8]。后续学者进一步推广到考虑大气阻力[9]、三体引力摄动[10-11]等情况。

高精度的轨道预报多采用高保真度动力学模型, 往往考虑地球形状、多体引力、大气、光压、固体潮、潮汐等因素, 轨道外推计算量巨大, 无法用于卫星在线预报。为了降低计算代价, 在线轨道预报大多基于正交多项式拟合快速外推星历。这些方

法需要大量高精度数据和高保真模型支持，不能适应低轨互联网星座的庞大卫星数量。为兼顾参数不确定下的在线模型外推预报准确性和快速性，人工神经网络（artificial neural network, ANN）被用于构建轨道外推快速预测响应面[12]。Jones 和 Weisman[13]综合考虑非确定因素影响，基于保真度协同，减少轨道外推状态参数概率密度函数精度损失，提出了多保真度不确定性推演自适应方法；Carlo 等[14]提出了基于保真协同调度的轨道转移优化工具。Toal[15]基于低保真模型函数构建初始响应面，利用 co-Kriging 预期改进的最大化来实现代理响应面优化，通过引入高保真模型数据，基于高斯过程修正比例因子进行修正低保真响应模型，实现轨道高精度外推预报。

在轨卫星轨道预报的鲁棒性评估多面向碰撞风险概率分析和空间碎片环境推演[16]，近年来低轨互联网星座成为热点对象[17]。基于参数统计特征和轨道外推的简化解析模型，会导致概率稀释问题，使设计人员和决策者错误估计风险水平[18]。为有效估计预报模型的鲁棒性，面向不确定性环境摄动参数的非连续性和突发特点，有学者基于 Dempster-Shafer 证据理论（Dempster-Shafer's theory, DST）对动力学预报中的认知不确定性进行分析[19]。侯黎强等[20]基于证据空间对火星探测器和激光空间碎片清除弱统计认知类型非确定性因素进行了量化分析，利用保真协同调度解决昂贵多学科模型精确模型计算代价和证据冲突计算维数膨胀问题，给出了置信和似然双水平度量鲁棒的高效多学科设计最优化设计方法。

上述卫星轨道外推预报的保真度协同和鲁棒性评估方法，均面向单星轨道的长期预报和碰撞概率分析问题，缺乏适用于低轨互联网星座的整体性与协同性。

2.2 星座自主导航与定轨

紧密卫星集群、传统中轨星座、传统卫星编队等不同空间星座，对应不同的自主导航问题和方法。与低轨互联网星座最为相关和能够借鉴的是传统中轨星座的自主导航与定轨方法，主要面向 GPS、北斗等中轨及含有中轨部分的导航星座。

自 20 世纪 80 年代，国外专家提出中轨导航星座自主定轨技术以来，一直致力于该技术的发展。至今，GPS 星座在轨卫星中，仍有多款卫星具备自主导航功能，并且对在轨卫星开展了一系列的星间链路测试[21]。

当前，国内外研究的目标主要围绕提高长期导航精度和数据处理效率 2 个方面。

1）在提高长期导航精度方面，主要有引入先验信息和增加额外观测基准 2 种思路。通过引入升交点经度、轨道倾角等先验信息，可有效抑制星座漂移误差的快速增长，但不能彻底消除，长期导航误差仍然发散[22]。为此，增加地面锚固站[23]、X 射线脉冲信号等空间基准[24]、惯性空间星间指向[25]等观测基准成为更有效的解决手段。李木子等[26]在星间测距的基础上，增加雷达测高和地面图像标志点方位等测量量，有效解决了混合星座长期漂移估计问题。可见，通过有限、局部增加星间定向等观测基准，可为低轨低轨互联网星座的构型漂移估计与补偿提供可行的研究思路。

2）提高数据处理效率的方法可分为优化星座整网测量通信结构和提升数据处理算法性能 2 个方面。快速高效地完成整网自主导航测量与数据分发是星座导航解算的基础。美国第三代 GPS 卫星 Block IIR 具有交联测距和在轨处理导航数据的能力。Block IIR 采用一种 TDMA 式的测量通信方式，设定 36 s 为一帧，每颗卫星分配 1.5 s，测距帧与数据帧交替进行[27]。由于传统的星座卫星数量较少，这种方式对整体的导航效率影响不大。但是，对于低轨互联网星座，由于该星座的卫星数量远远超过之前星座的卫星数量，星座内卫星的单帧测量通信时间会大幅度提升，星座内部的实时性将无法保障。因此，高效快速的低轨互联网星座整网测量与通信结构设计和优化，成为亟待解决的新问题，国内外尚未专门针对该问题开展研究。

目前，星座自主导航数据处理方法从算法结构上分为 2 种：集中式、分布式。一般来说，集中式处理模式下，由主星集中处理所有星间观测数据生成导航星历，并得到给定噪声条件下的最优解，但该结构对主星计算和通信能力要求极高，抗风险性差，因此，不适用于大规模星座的自主导航。分布式处理模式下，每颗卫星仅利用与之相关的观测量单独进行处理，将整网导航问题解耦为多个相互独立的子问题，有效提升了导航解算效率、可拓展性、容错性。文献 [28] 基于 GPS、BDS 星座和星间链路的特点，循序渐进研究了迭代级联扩展卡尔曼滤波 (iterative cascading extended Kalman filter, ICEKF)、提升扩展卡尔曼滤波 (improved cascading extended Kalman filter, IMCEKF)、Schmidt-Kalman 滤波 (Schmidt-Kalman filter, SKF)、状态平衡扩展卡尔曼滤波（balanced extended Kalman filter, BEKF）等分布式自主定轨算法和抗差自适应滤波方法，均可以有效提高导航精度、稳健性和星载处理效率。

星载全球导航卫星系统（global navigation satellite

system, GNSS)定轨方法是低轨互联网星座定轨的主要方法, 常用的 4 种方法的优缺点参见文献 [29]。

面对低轨互联网星座这样的巨型星座, 如何提高 GNSS 定轨的解算效率, 充分提高多频多系统观测数据的利用率, 是当前首先需要解决的问题。

2.3　星座构型重构的路径规划

星座构型重构路径规划的本质是: 在分布式求解框架下, 星座中的个体综合考虑安全防撞、任务维持等多约束条件, 燃料消耗、重构时间等多目标, 以及复杂摄动、测量与控制不确定、未知有界干扰等因素的鲁棒优化问题。

对于轨道机动路径的鲁棒优化, 大量学者在研究中考虑了轨道摄动、有界干扰等因素影响, 提出的方法主要分为直接法、间接法和混合法 3 种。Lee 和 Park[30] 在考虑非线性和 J_2 摄动的基础上, 通过变分法得到了近似燃料最优解。潘迅等[31] 结合多次同伦算法求解考虑 J_2 项摄动的燃料最优转移轨道。间接法的初值猜测等问题一直不能很好解决, 当任务时间较长时, 收敛较慢, 易于陷入局部最优解。直接法则通过离散化的方法, 将连续无穷维的最优化问题转换成有限维的非线性规划问题, 包括伪谱法[32-34]、虚拟域逆动力学法[35] 等。王有亮[36] 针对卫星编队飞行开展了一系列研究, 并尝试了不同的轨迹优化方法, 均取得了不错的效果。Morgan 等针对大规模星座的路径规划问题, 采用分布式序列凸优化方法寻找 J_2-invariant 轨道[37], 并将序列凸优化与模型预测控制结合[38], 提升了对不确定性和干扰的适应性。Foust 等[39] 提出了非线性动力学修正方法, 补偿了序列凸优化方法存在的线性化和离散化模型偏差, 保证较低计算量的同时, 提升了序列凸优化对非线性系统的性能。但是, 现有分布式鲁棒优化方法均未综合考虑低轨互联网星座面临的复杂动态拓扑下的分布式求解策略、信息交互时延、连续微推力、不确定与干扰等因素影响。

在多目标多约束优化方面, 传统多目标优化方法一般采用加权法、约束法、目标规划法等方法, 将多个目标函数合成一个单独有参数的目标函数, 在不同参数设置下得到考虑不同目标侧重的近似 Pareto 最优解集。近年来, 随着人工智能的发展, 针对多目标多约束优化的进化算法[40-42]、遗传算法[43-45]、模拟退火算法[46] 等在空间星座重构中得到了广泛应用。对于复杂重构机动过程, 可以分为多个任务阶段, 各阶段对应的优化目标和约束条件众多且复杂时变, 甚至不同阶段间目标与约束存在冲突。简单地将多目标加权合成为一个目标, 无法适应不同任务阶段的不同目标与约束, 得到的优化路径结果整体性能较为保守, 且在不同目标或约束冲突时很可能导致无解。不同阶段的约束和目标存在明显冲突。对此, 有学者研究了分段多目标规划的方法, 如 Soleymani[47] 和 Karaman[48] 等分别使用混合粒子群优化算法和遗传算法, 降低了星座构型重构的能源消耗问题。但是, 简单分阶段规划会导致问题复杂, 不适合在线规划求解。基于时序逻辑的路径规划方法能够很好地解决此类含有复杂时序任务的路径规划问题, 进而得到全局最优路径, 已经在多机器人系统[49]、无人机编队[50] 领域路径规划问题中得到应用, 然而针对星座重构等复杂多任务多约束路径规划的研究还处于探索阶段。

可见, 目前的星座变构路径规划方法难以满足低轨互联网星座的特殊性需要。

2.4　星座自组织协同构型控制

维持与重构是星座构型控制的 2 个重要方面, 研究的重点分别在于星座相对运动边界维持控制和基于局部信息的分布式控制。

低轨互联网星座相对运动边界维持控制的目标是维持相对运动有界不发散, 可以通过少量控制的方式实现长期在轨飞行[51]。采用能量匹配、长期项消除等方法获得长期自然维持的相对运动初始状态, 是一类减少燃料消耗的有效途径。其中, 能量匹配指成员星初始轨道对应动能和势能构成的总能量与期望轨道总能量保持一致, 此时成员星轨道能够自然跟随期望轨道, 有效降低相对运动边界发散速度。Chung 团队[51-52] 针对星座拓扑动态切换变连接下的构型控制, 提出了自适应动态网络图的相位同步控制方法。王兆魁等[53-54] 推导得到了 J_2 摄动下的有界相对运动, 通过解析或半解析的方法进行相对运动边界求解, 从而实现了少量控制下的安全、松散有界相对运动。

基于局部信息的分布式控制是大规模星座变构控制的必然途径。这是因为卫星数量多, 且星座成员的计算资源、通信容量、相对导航距离等受限, 每个成员只能获取周围邻居的相对信息, 且只能求解自身的控制指令。随着协同规则和通信能力的发展, 无中心自组织控制成为分布式控制的重要形式, 其实现手段是建立个体的行为规则, 通过个体行为的叠加实现整体意义上的涌现行为, 主要包括一致性控制[55]、分布式模型预测控制[56]、分布式自适应协同控制[57]、人工势函数控制[58] 等方法。刘付成等[59-60] 考虑星间拓扑连接关系和未知干扰影响, 采用多智能体系统一致性理论, 设计了自适应增益分布式控制算法, 实现了基于局部信息的多星构型分布协同控制。人工势函数方法具有

复杂环境下形式简单、计算效率高等优点[38,61]。McInnes[62]首次将人工势函数方法用于卫星在轨控制与自主交会任务；Mushet等[63]研究了地球同步轨道上自组织星座的自主任务分配问题，由人工势函数实现卫星机动控制；Spencer等[64-65]提出基于相对轨道根数的人工势函数设计方法，控制单颗卫星收敛到目标飞行区域。星座变构控制过程中还必须考虑内部成员间避碰和外部威胁避碰2个方面的问题。周亮[66]以航天器集群面对空间碎片等障碍物威胁的快速分散/重聚任务为例，使用改进的模拟退火算法（modified simulate anneal algorithm, MSAA）搜索星座重构问题的全局最优解，得到了一种航天器集群飞行的重构机动策略。人工势函数方法也可以有效实现大规模星座避碰控制，Palacios等[67]构造了高斯型排斥势函数以确保星间避碰，Nag和Summerer[68]基于人工势函数提出多种卫星集群避碰策略。王兆魁等[69-70]基于相对轨道根数实现了近距离星座的自组织重构、有界飞行与避碰控制。同样的，这些星座分布式控制方法都是针对近距离星座。

当前，星座自组织协同构型控制方法均属于近距离星座的多脉冲离散控制，没有针对低轨互联网星座特性的相关应用。

2.5 星座通信与网络服务

卫星网的跨域覆盖特性为移动通信网在海域、空域、山区等较困难地区的接入提供了有效的途径，地面移动通信和卫星通信融合已经成为一种发展趋势[71]。3GPP公开明确表示其将开展“基于5G的卫星网络部署方案”的研究；而欧洲航天局则成立了Sat5G组织，探索5G在卫星上应用的可行性方案。6G网络技术将能够提供无缝全覆盖的、泛在无线连接的和情景感知的智能服务与应用，并且实现包含地面、卫星、机载网络和海洋通信网络等全方位、多领域的无缝覆盖[4,72-73]。将低轨通信卫星之间有无星间链路作为分类依据，可以将低轨互联网星座划分为2种网络架构：“天星天网”和“天星地网”。“天星天网”的代表有Iridium、Starlink等星座，特点是：卫星作为网络传输节点，拥有星上数据处理功能，无需建立多个关口站，用户可以直接连接到卫星互联网，但是这种架构对星间链路及路由算法的要求很高。“天星地网”的代表有Globalstar、OneWeb等星座，特点是：星上是一个透明转发器，卫星之间没有组网，技术复杂度很低，在维护和管理方面有很大的优势，但是这种架构必须要克服地缘问题，在全球范围内构建出足够多的地面站，才可以达到一个完整的全球服务能力[74]。

低轨卫星通信采用2个分开且有一定频率间隔要求的频分双工（frequency division duplex, FDD）模式，分别完成实现地面站至卫星的上行链路和卫星至地面站的下行链路工作。要实现基于5G技术的低轨卫星通信，则要求系统必须对5G部分关键技术进行一定适应性的改进，并有针对性的优化和设计，如需要考虑卫星通信信道特有的大路损、大时延、大多普勒频偏、快速多普勒频偏变化率、城市高大建筑遮挡直视径信道（line of sight, LOS）、频繁的波束切换和星间切换、卫星姿态变化、毫米波信号的空间特性、高频相位噪声等特点[74]。但要本着最大程度复用/兼容地面5G技术标准的原则，低轨卫星通信系统在指标体系定义及空口技术规格定义上做相应的适配，如采用双向时延无关的FDD双工模式，具有较低峰均比（peak to average power ratio, PAPR）的DFT-s-OFDM波形，可容忍大多普勒频移的最小120 kHz子载波间隔，针对LOS信道简化传输模式的单流传输、多途径星历参数及卫星姿态信息更新等技术[75-77]。

随着星上信号再生处理、路由交换等技术的提高，以及载荷容量与算力的提升，卫星节点能够通过星间链路来实现信息的共享和网络状态的感知。对于不同处理能力的卫星，能够将其分为不同层级，采用协同方式来进行端到端的网络和业务管理。而地面站则以网关等形式，与技术成熟的地面蜂窝网络相结合，从而达到卫星网络与地面网络之间业务的互联互通。低轨卫星节点承担主要的数据传输、路由寻址、上传卫星动态等任务。作为天地网络融合提供服务的先决条件，卫星通信与地面移动通信体制逐渐兼容、频率资源实现复用、网络调度与服务质量统一管理等方面已经成为目前的研究热点[78]。

3 展望与建议

随着空间站建设，深空探测领域不断取得新的突破，中国正在从航天大国向航天强国迈进。同时，低轨互联网星座的发展预示着航天“轨道革命”的来临，为国家带来了新的机遇。机遇的背后，则是挑战，因此，在积极承接新机遇的同时，也要对随之而来的挑战和问题做好准备。结合上述五大技术挑战，以及发展现状，对中国的低轨互联网星座的发展建议如下：

1）针对星座相对运动演化预报挑战。对于低轨巨型星座成员星群相对运动演化预报，首先，需要分析大空间尺度相对运动的周期性松散有界特性。然后，构建成员星局部拓扑关联集合。最后，

提出基于保真度协同的成员间距离与运动边界快速预报方法,并基于证据理论评估预报模型与方法在复杂摄动与不确定性下的鲁棒性边界。

2) 针对星座自主导航与定轨挑战。导航方面,由于当前分布式自主导航数据处理方法对于数百上千颗卫星构成的低轨互联网星座,需要大幅提升测量更新与数据处理效率,并在单星计算能力约束下提高导航精度和稳健性。因此,需要引入低时延区域并行处理结构的概念,研究动态拓扑网络耦合下适应星上处理能力的降阶次优分布式并行数据处理方法。

3) 针对星座构型重构的路径规划挑战。由于低轨互联网星座的特殊性,首先,可以考虑分布式求解框架下的信息交互时延,以及连续微推力、不确定摄动与干扰等特性,提高分布式优化算法的鲁棒性。其次,可以考虑不同任务阶段的多目标多约束时序逻辑,避免约束无解和优化目标冲突,提高复杂重构多阶段路径优化的整体性能。

4) 针对星座自组织协同构型控制挑战。由于成员星座受到全球覆盖构型约束,其周期松散有界相对运动尺度远大于近距离星座,不满足近距离、圆轨道等简化假设;同时,成员卫星普遍采用电推进,轨道控制作用力为长时间连续微推力,不满足离散控制的短时脉冲推力假设。

因此,首先,可以从适应大空间尺度松散有界运动的能量匹配条件和连续微作用力下的运动边界控制方法开展研究。其次,也可以在人工势函数中引入大尺度相对轨道根数,扩展满足连续微推力约束的自组织控制状态量,构造新型势函数。

5) 针对星座通信与网络服务挑战。低轨互联网星座通信面向 B5G/6G,需要将技术结合未来网络的发展和应用需求,深入研究网络人工智能技术,从而实现包括服务按需提供、网络资源智能编排和满足网络架构弹性可重构需求等一系列目标。

安全性上,C 频段和 Ku 频段的卫星轨位资源正在日益减少,频率带宽变得越来越紧张,可以在 Ka 频段进行发展。Ka 频段的特点类似于 Ku 频段,虽然有雨衰较大的绝点,但是其增益较高,因此,可以通过增加适量天线口径的方式来削弱影响。

而在时间同步过程中,通信速率、信噪比、抗干扰能力等均为需要满足的参考指标,目前的低轨互联网星座卫星间的时间同步需要进一步研究。

6) 加强低轨互联网星座政策顶层设计。要清晰地认识到低轨卫星轨道和频谱资源作为一种不可再生的战略资源,是极其珍贵的。因此,不但需要加快低轨互联网星座的建设,更重要的是,需要积极地参与到国际相关法规的制定中,为中国的低轨互联网星座的发展奠定法律保障,为后续星座计划的轨道资源和频谱资源预留发展先机。

7) 构建中国特色低轨互联网星座。面对当前世界规模较大的低轨互联网星座计划,要仔细研判,既不能照搬国外的发展模式,也不能闭门造车。要充分汲取那些已经失败的案例的教训,基于国家的经济发展水平、技术实力、实际需求等方面,结合国家的技术优势(5G 技术、高铁技术等),在交叉领域构建具有中国特色的低轨互联网星座。

8) 做好技术储备积极应对各项挑战。随着各国低轨互联网星座进入高密度发射器,在轨卫星越来越多,空间碰撞的几率也大幅度提升。2021 年 12 月,据中国外交部证实,“Starlink 星座”在 2021 年 7 月、10 月曾 2 次靠近中国空间站,使得中国空间站不得不采取“紧急避险”的措施,而当时,中国航天员正在空间站执行任务。因此,发展在轨卫星自动避障技术、在轨卫星主动规避技术、低成本空间碎片批量移除技术等技术就显得尤为重要。

此外,由于低轨互联网星座规模宏大,发射成本低,并且失效卫星能够在短时间内完成补网发射,低轨互联网星座本身作为一种空间基础设施,就具有成为军用平台的可能性,而其所具有的强抗毁性更为其向军用领域发展奠定了基础。下一代太空架构是美军 2019 年成立的太空发展局推出的以近地轨道星座为主的大型星座网络,2020 年设计,2021 年进入招标研制,在该项目下,小卫星的优势发挥到极致。因此,中国需要在相关领域密切关注,提前做好技术储备来应对可能的挑战。

4　结束语

当前,随着商业航天的蓬勃发展,以 SpaceX、OneWeb 为代表的低轨互联网星座在世界掀起了一阵发展浪潮。本文介绍了具有代表性的国内外星座计划及当前进展,着重分析了低轨互联网星座面临的五大技术挑战:星座相对运动演化预报、星座自主导航与定轨、星座构型重构的路径规划、星座自组织协同构型控制、星座通信与网络服务。结合这五大技术挑战及中国的实际情况,给出了发展的建议。

参考文献(References)

[1] 徐冰玉, 李侠宇. Starlink 低轨卫星通信星座深度分析[J]. 信息通信技术与政策, 2021, 47(9): 16-20.
XU B Y, LI X Y. In-depth analysis of the satellite constellation Starlink[J]. Information and Communications Technology, 2021, 47(9): 16-20(in Chinese).

[2] 杨文翰, 花国良, 冯岩, 等. 星链计划卫星网络资料申报情况分析[J]. 天地一体化信息网络, 2021, 2(1): 60-68.
YANG W H, HUA G L, FENG Y, et al. Analysis of the Starlink satellite network filing application[J]. Space-Integrated-Ground Information Networks, 2021, 2(1): 60-68(in Chinese).

[3] CLARK S. SpaceX launches 25th mission for Starlink Internet network[EB/OL]. (2021-03-24)[2022-04-01]. https://spaceflightnow.com/2021/03/24/spacex-launches-25th-mission-to-build-out-starlink-internet-network/.

[4] PORTILLO I D, CAMERON B G, CRAWLEY E F. A technical comparison of three low Earth orbit satellite constellation systems to provide global broadband[J]. Acta Astronautica, 2019, 159: 123-135.

[5] OneWeb. OneWeb secures investment from SoftBank and Hughes network systems[EB/OL]. (2021-01-15) [2022-04-01]. https://oneweb.net/media-center/oneweb-secures-investment-from-softbank-and-hughes-network-systems.

[6] 阮永井, 胡敏, 云朝明. 低轨互联网星座构型设计与控制研究进展与展望[J]. 中国空间科学技术, 2022, 42(1): 1-15
RUAN Y J, HU M, YUN C M. Advances and prospects of the configuration design and control research of the LEO mega-constellations[J]. Chinese Space Science and Technology, 2022, 42(1): 1-15 (in Chinese).

[7] SCHWEIGHART S A, SEDWICK R J. High-fidelity linearized J_2 model for satellite formation flight[J]. Journal of Guidance, Control, and Dynamics, 2002, 25(6): 1073-1080.

[8] ALFRIEND K T, VADALI S R, GURFIL P, et al. Spacecraft formation flying: Dynamics, control and navigation[M]. Amsterdam: Elsevier, 2009.

[9] SCHAUB H. Relative orbit geometry through classical orbit element differences[J]. Journal of Guidance, Control, and Dynamics, 2004, 27(5): 839-848.

[10] SCHAUB H, ALFRIEND K T. J_2 invariant relative orbits for spacecraft formations[J]. Celestial Mechanics and Dynamical Astronomy, 2001, 79(2): 77-95.

[11] KOENIG A W, GUFFANTI T, DAMICO S. New state transition matrices for relative motion of spacecraft formations in perturbed orbits[C]//Proceedings of the Astrodynamics Specialist Conference. Reston: AIAA, 2016: 13-16.

[12] XU G Y, WANG D W. Nonlinear dynamic equations of satellite relative motion around an oblate earth[J]. Journal of Guidance, Control, and Dynamics, 2008, 31(5): 1521-1524.

[13] JONES B A, WEISMAN R. Multi-fidelity orbit uncertainty propagation[J]. Acta Astronautica, 2019, 155: 406-417.

[14] CARLO M D, MARTIN J M R, VASILE M. CAMELOT: Computational-analytical multi-fidelity low-thrust optimisation tool-box[J]. CEAS Space Journal, 2018, 10: 25-36.

[15] TOAL D J J. Some considerations regarding the use of multi fidelity Kriging in the construction of surrogate models[J]. Structural and Multidisciplinary Optimization, 2015, 51(6): 1223-1245.

[16] VIRGILI B B, DOLADO J C, LEWIS H G, et al. Risk to space sustainability from large constellations of satellites[J]. Acta Astronautica, 2016, 126: 154-162.

[17] ALFANO S, OLTROGGE D. Probability of collision: Valuation, variability, visualization, and validity[J]. Acta Astronautica, 2018, 148: 301-316.

[18] NEWMAN L, FRIGM R C, DUNCAN M G, et al. Evolution and implementation of the NASA robotic conjunction assessment risk analysis concept of operations[C]//Proceedings of the Advanced Maui Optical and Space Surveillance Technologies Conference. Red Hook: Curran Associates, Inc. , 2014.

[19] HOU L Q, CAI Y L, LIU J, et al. Variable fidelity robust optimization of pulsed laser orbital debris removal under epistemic uncertainty[J]. Advances in Space Research, 2016, 57(8): 1698-1714.

[20] HOU L Q, CAI Y L, ZHANG R Z, et al. Robust design of Mars entry micro-probe with evidence theory and multi-fidelity strategies[J]. Engineering Computations, 2014, 31(6): 1052-1073.

[21] 肖寅. 导航卫星自主导航关键技术研究[D]. 上海: 中国科学院研究生院 (上海技术物理研究所), 2015.
XIAO Y. Research on the key technology of automous navigation for navigation satellite[D]. Shanghai: University of Chinese Academy of Sciences (Shanghai Institute of Technical Physics), 2015 (in Chinese).

[22] 陈金平, 焦文海, 马骏, 等. 基于星间测距/轨道定向参数约束的导航卫星自主定轨研究[J]. 武汉大学学报 (信息科学版), 2005, 30(5): 439-443.
CHEN J P, JIAO W H, MA J, et al. Autonav of navigation satellite constellation based on crosslink range and orientation parameters constraining[J]. Geomatics and Information Science of Wuhan University, 2005, 30(5): 439-443(in Chinese).

[23] 苏天祥, 文援兰, 蓝柏强, 等. 基于锚固站的导航星座分布式自主定轨研究[J]. 大地测量与地球动力学, 2014, 34(3): 111-115.
SU T X, WEN Y L, LAN B Q, et al. Study on autonomous orbit determination of navigation constellation with distribution algorithm based on ground-based anchors[J]. Journal of Geodesy and Geodynamics, 2014, 34(3): 111-115(in Chinese).

[24] 李琳琳, 孙辉先. 基于星敏感器的星光折射卫星自主导航方法研究[J]. 系统工程与电子技术, 2004, 26(3): 353-357.
LI L L, SUN H X. Satellite autonomous navigation by stellar refraction based on a star sensor[J]. Systems Engineering and Electronics, 2004, 26(3): 353-357(in Chinese).

[25] PSIAKI M L. Absolute orbit and gravity determination using relative position measurements between two satellites[J]. Journal of Guidance, Control, and Dynamics, 2011, 34(5): 1285-1297.

[26] LI M Z, XU B, SUN J. Autonomous orbit determination for a hybrid constellation[J]. International Journal of Aerospace Engineering, 2018, 2018: 4843061.

[27] 张艳. 基于星间观测的星座自主导航方法研究[D]. 长沙: 国防科技大学, 2005.
ZHANG Y. Study on autonomous navigation of constellation using inter-satellite measurement[D]. Changsha: National University of Defense Technology, 2005(in Chinese).

[28] FERGUSON P A. Distributed estimation and control technologies for formation flying spacecraft by Philip Andrew Ferguson[D]. Cambridge: Massachusetts Institute of Technology, 2003.

[29] 刘伟平, 郝金明, 王智明. 几种 LEO 星载 GNSS 精密定轨方法的对比分析[J]. 测绘科学技术学报, 2014, 31(2): 140-144.
LIU W P, HAO J M, WANG Z M. Comparison and analysis of some methods of LEO space-borne GNSS precise orbit determination[J]. Journal of Geomatics Science and Technology, 2014, 31(2): 140-144(in Chinese).

[30] LEE S, PARK S Y. Approximate analytical solutions to optimal reconfiguration problems in perturbed satellite relative motion[J]. Journal of Guidance, Control, and Dynamics, 2011, 34(4): 1097-1111.

[31] 潘迅, 洋斌峰, 唐硕. 考虑 J_2 项摄动的小推力燃料最优转移轨道设计[J]. 哈尔滨工业大学学报, 2017, 49(10): 15-21.
PAN X, PAN B F, TANG S. Fuel-optimal low thrust trajectory design with J_2 perturbation[J]. Journal of Harbin Institute of Technology, 2017, 49(10): 15-21(in Chinese).

[32] WU B L, WANG D W, POH E K, et al. Nonlinear optimization of low-thrust trajectory for satellite formation: Legendre pseudo spectral approach[J]. Journal of Guidance, Control, and Dynamics, 2009, 32(4): 1371-1381.

[33] 韩威华, 甘庆波, 王晓光, 等. 摄动情况下有限推力轨道转移与交会联合优化[J]. 系统工程与电子技术, 2013, 35(7): 1486-1491.
HAN W H, GAN Q B, WANG X G, et al. Comprehensive optimization for orbit transfer and rendezvous of finite thrust with perturbation[J]. Systems Engineering and Electronics, 2013, 35(7): 1486-1491(in Chinese).

[34] HAN F, WANG Z, DUAN G R, et al. Trajectory plan for an ultra-short distance on-orbit service based on the Gaussian pseudo-spectral method[J/OL]. IEEE/CAA Journal of Automatica Sinica, 2018 (2018-02-13)[2022-04-01]. https://ieeexplore.ieee.org/document/8291087.

[35] 武海雷, 韩飞, 贺亮, 等. 翻滚目标逼近的虚拟域逆动力学轨迹规划[J]. 控制理论与应用, 2018, 35(10): 1453-1461.
WU H L, HAN F, HE L, et al. Trajectory planning based on inverse dynamics in virtual domain for approaching tumbling target[J]. Control Theory & Applications, 2018, 35(10): 1453-1461(in Chinese).

[36] 王有亮. 卫星编队飞行相对轨迹优化与控制[D]. 北京: 中国科学院大学, 2018.
WANG Y L. Relative trajectory optimization and control for satellite formation flying[D]. Beijing: University of Chinese Academy of Sciences, 2018(in Chinese).

[37] MORGAN D, CHUNG S J, HADAEGH F Y. Spacecraft swarm guidance using a sequence of decentralized convex optimization: AIAA 2012-4583[R]. Reston: AIAA, 2012.

[38] MORGAN D, CHUNG S, HADAEGH F Y, et al. Model predictive control of swarms of spacecraft using sequential convex programming[J]. Journal of Guidance, Control, and Dynamics, 2014, 37(6): 1725-1740.

[39] FOUST R, CHUNG S J, FRED Y, et al. Solving optimal control with nonlinear dynamics using sequential convex programming: AIAA 2019-0652[R]. Reston: AIAA, 2019.

[40] 黄海滨, 马广富, 庄宇飞, 等. 基于协同进化粒子群和 Pareto 最优解的卫星编队队形重构方法[J]. 航空学报, 2011, 32(11): 118-127.
HUANG H B, MA G F, ZHUANG Y F, et al. Satellite formation reconfiguration using co-evolutionary particle swarm optimization and Pareto optimal solution[J]. Acta Aeronautica et Astronautica Sinica, 2011, 32(11): 118-127(in Chinese).

[41] 张磊, 田百义, 周文艳, 等. 木星系多目标探测轨道设计研究[J]. 航天器工程, 2018, 27(1): 31-36.
ZHANG L, TIAN B Y, ZHOU W Y, et al. Research on Jovian system tour trajectory design[J]. Spacecraft Engineering, 2018, 27(1): 31-36(in Chinese).

[42] 汉京滨, 张雅声, 汤亚锋, 等. 基于 NSGA-Ⅱ的通信星座重构方法研究[J]. 兵器装备工程学报, 2019, 40(8): 74-79.
HAN J B, ZHANG Y S, TANG Y F, et al. Research on reconstruction method of communication constellation based on NSGA-Ⅱ[J]. Journal of Ordnance Equipment Engineering, 2019, 40(8): 74-79(in Chinese).

[43] SAVITRI T, KIM Y, JO S, et al. Satellite constellation orbit design optimization with combined genetic algorithm and semianalytical approach[J]. International Journal of Aerospace Engineering, 2017, 2017: 1235692.

[44] 张敬, 吴美平, 付晓锋. 遗传算法在星座多星接近轨道优化中的应用[J]. 空间科学学报, 2012, 32(1): 99-105.
ZHANG J, WU M P, FU X F. Genetic algorithm for orbital optimization to approach multiple constellation satellites[J]. Chinese Journal of Space Science, 2012, 32(1): 99-105(in Chinese).

[45] PAEK S, KIM S, DE WECK O. Optimization of reconfigurable satellite constellations using simulated annealing and genetic algorithm[J]. Sensors, 2019, 19(4): 765.

[46] FAKOOR M, BAKHTIARI M, SOLEYMANI M. Optimal design of the satellite constellation arrangement reconfiguration process[J]. Advances in Space Research, 2016, 58(3): 372-386.

[47] SOLEYMANI M, FAKOOR M, BAKHTIARI M. Optimal mission planning of the reconfiguration process of satellite constellations through orbital maneuvers: A novel technical framework[J]. Advances in Space Research, 2019, 63(10): 3369-3384.

[48] KARAMAN S, FRAZZOLI E. Linear temporal logic vehicle routing with applications to multi-UAV mission planning[J]. International Journal of Robust and Nonlinear Control, 2011, 21(12): 1372-1395.

[49] RAMIREZ-ATENCIA C, BELLO-ORGAZ C, R-MORENO M D, et al. Solving complex multi-UAV mission planning problems using multi-objective genetic algorithms[J]. Soft Computing, 2017, 21(17): 4883-4900.

[50] PENG H, LI C. Bound evaluation for spacecraft swarm on libration orbits with an uncertain boundary[J]. Journal of Guidance, Control, and Dynamics, 2017, 40(10): 2690-2698.

[51] MORGAN D, CHUNG S J. Swarm-keeping strategies for spacecraft under J_2 and atmospheric drag perturbations[J]. Journal of Guidance, Control, and Dynamics, 2012, 35(5): 1492-1506.

[52] CHUNG S J, BANDYOPADHYAY S, CHANG I, et al. Phase synchronization control of complex networks of Lagrangian systems on adaptive digraphs[J]. Automatica, 2013, 49(5): 1148-1161.

[53] DANG Z, WANG Z K, ZHANG Y. Improved initialization conditions and single impulsive maneuvers for J_2-invariant relative orbits[J]. Celestial Mechanics and Dynamical Astronomy, 2015, 121(3): 301-327.

[54] DANG Z, WANG Z K, ZHANG Y. Modeling and analysis of the bounds of periodical satellite relative motion[J]. Journal of Guidance, Control, and Dynamics, 2014, 37(6): 1984-1998.

[55] DANG Z, LI T, WANG Z, et al. Bounds on maximal and minimal distances for coplanar satellite relative motion under given initial conditions[J]. Aerospace Science and Technology, 2015, 46: 204-209.

[56] DANG Z, FAN L, WANG Z, et al. On the maximal and minimal distances of flying-around satellite formation[J]. Aircraft Engineering and Aerospace Technology, 2017, 89(6): 845-852.

[57] HE W, CHEN G, HAN Q, et al. Network-based leader-following

consensus of nonlinear multi-agent systems via distributed impulsive control[J]. Information Sciences, 2017, 380: 145-158.

[58] DU H, CHEN M Z Q, WEN G. Leader-following attitude consensus for spacecraft formation with rigid and flexible spacecraft[J]. Journal of Guidance, Control, and Dynamics, 2016, 39(4): 941-948.

[59] 刘付成, 梅杰, 马广富. 有向图中模块化航天器系统相对轨道的自适应分布式一致性[J]. 控制理论与应用, 2014, 31(2): 223-229.
LIU F C, MEI J, MA G F. Adaptive distributed consensus for relative orbit of modular spacecrafts under a directed graph[J]. Control Theory & Applications, 2014, 31(2): 223-229(in Chinese).

[60] 刘付成, 梅杰, 马广富. 带未知干扰的模块化航天器系统相对轨道的队形控制[J]. 控制与决策, 2014, 29(6): 985-990.
LIU F C, MEI J, MA G F. Formation control for relative translation of modular spacecraft with unknown disturbances[J]. Control and Decision, 2014, 29(6): 985-990(in Chinese).

[61] CHEN T, CHEN G. Distributed adaptive tracking control of multiple flexible spacecraft under various actuator and measurement limitations[J]. Nonlinear Dynamics, 2018, 91(3): 1571-1586.

[62] MCINNES C R. Autonomous proximity manoeuvring using artificial potential functions[J]. ESA Journal, 1993, 17(2): 159-169.

[63] MUSHET G, MINGOTTI G, COLOMBO C, et al. Self-organizing satellite constellation in geostationary earth orbit[J]. IEEE Transactions on Aerospace and Electronic Systems, 2015, 51(2): 910-923.

[64] SPENCER D A, LOVELL T A. Maneuver design using relative orbital elements[J]. Journal of the Astronautical Sciences, 2015, 62(4): 315-350.

[65] SPENCER D A. Automated trajectory control using artificial potential functions to target relative orbits[J]. Journal of Guidance, Control, and Dynamics, 2016, 39(9): 2142-2148.

[66] 周亮. 航天器集群飞行的轨道保持与重构机动[D]. 西安: 西北工业大学, 2016.
ZHOU L. Orbit maintenance and reconfiguration for spacecraft cluster flight[D]. Xi'an: Northwestern Polytechnical University, 2016 (in Chinese).

[67] PALACIOS L, CERIOTTI M, RADICE G. Close proximity formation flying via linear quadratic tracking controller and artificial potential function[J]. Advances in Space Research, 2015, 56(10): 2167-2176.

[68] NAG S, SUMMERER L. Behaviour based, autonomous and distributed scatter manoeuvres for satellite swarms[J]. Acta Astronautica, 2013, 82(1): 95-109.

[69] WANG Z K, XU Y, JIANG C, et al. Self-organizing control for satellite clusters using artificial potential function in terms of relative orbital elements[J]. Aerospace Science and Technology, 2019, 84: 799-811.

[70] XU Y, WANG Z K, ZHANG Y L. Bounded flight and collision avoidance control for satellite clusters using intersatellite flight bounds[J]. Aerospace Science and Technology, 2019, 94: 105425.

[71] SALARI A, SHIRVANIMOGHADDAM M, SHAHAB M B, et al. Clustering-based joint channel estimation and signal detection for grant-free NOMA[C]//Proceedings of the IEEE Globecom Workshops. Piscataway: IEEE Press, 2020: 1-6.

[72] 孟亚洁, 武林. 信息通信技术发展趋势展望[J]. 中国电信业, 2021(2): 12-15.
MENG Y J, WU L. Outlook on information and communication technology development trends[J]. China Telecommunications Trade, 2021(2): 12-15(in Chinese).

[73] LI P X, HE Y Z, CUI G F, et al. Asynchronous cooperative alpha for multi-receiver satellite communication networks[J]. IEEE Communications Letters, 2017, 21(6): 1321-1324.

[74] 陈文江, 陈麒安, 陈宏铭, 等. 低轨卫星通信的机遇、挑战与量测方案[J]. 中国集成电路, 2022, 31(6): 22-30.
CHEN W J, CHEN Q A, CHEN H M, et al. Opportunities, challenges and measurement schemes of LEO satellite communication[J]. China Integrated Circuit, 2022, 31(6): 22-30(in Chinese).

[75] 赵鹏. 我国低轨卫星通信产业发展现状及趋势分析[J]. 卫星应用, 2021(8): 20-23.
ZHAO P. Development status and trend analysis of low-orbit satellite communication industry in China[J]. Satellite Application, 2021(8): 20-23(in Chinese).

[76] 陈山枝. 关于低轨卫星通信的分析及我国的发展建议[J]. 电信科学, 2020, 36(6): 1-13.
CHEN S Z. Analysis of LEO satellite communication and suggestions for its development strategy in China[J]. Telecommunications Science, 2020, 36(6): 1-13(in Chinese).

[77] LIU J J, SHI Y P, ZHAO L, et al. Joint placement of controllers and gateways in SDN-enabled 5G-satellite integrated network[J]. IEEE Journal on Selected Areas in Communications, 2018, 36(2): 221-232.

[78] 张聪, 高峰. 卫星互联网未来应用场景及安全性分析[J]. 信息技术, 2022(3): 120-126.
ZHANG C, GAO F. The future application scenarios and security analysis of satellite Internet[J]. Information Technology, 2022(3): 120-126(in Chinese).

Review on development of LEO Internet constellation

WU Shufan，WANG Wei*，WEN Jifan，WU Yuedong

(School of Aeronautics and Astronautics，Shanghai Jiao Tong University，Shanghai 200240，China)

Abstract: In recent years, several innovative enterprises like SpaceX and OneWeb have initiated the LEO Internet constellations. Therefore, recognizing its importance to the internet usersis meaningful to review the state-of-the-art technologies and the relevant challenges. In this review, three representative LEO Internet constellation plans (Starlink, OneWeb, Lightspeed) and the latest development in China are firstly introduced. Then, according to the characteristics of LEO Internet constellations, five major challenges of the LEO Internet constellations are analyzed, i.e., prediction of relative motion evolution of constellation, autonomous navigation and orbiting of constellation, path planning for constellation configuration reconstruction, constellation self-organizing collaborative configuration control, constellation communication and network services. Finally, some suggestions are provided for the development of LEO Internet constellation in China.

Keywords: LEO Internet constellation; prediction of relative motion evolution of constellation; autonomous navigation and orbiting of constellation; path planning for constellation configuration reconstruction; constellation self-organizing collaborative configuration control

Received：2022-04-11；**Accepted**：2022-08-11；**Published Online**：2022-08-19 09：27

URL：link.cnki.net/urlid/11.2625.V.20220818.1438.001

Foundation items：National Natural Science Foundation of China (11902175,U20B2056)

* **Corresponding author**. E-mail：wangwei215661@sjtu.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0217

锂离子电池热失控气体毒性风险分析方法

张青松*，曲奕润，刘添添
(中国民航大学 民航热灾害防控与应急重点实验室，天津 300300)

摘　　要：为研究锂离子电池热失控气体中主要有毒物质的危害程度，提出一种基于风险评估理论的锂离子电池热失控气体毒性风险分析方法。以点燃参数表征锂离子电池热失控发生概率，利用2种有效剂量分数(FED)方程与气体传感器阵列检测结果，建立锂离子电池热失控气体毒性动力学模型，进而表征气体毒性造成的后果，分析不同荷电状态(SOC)下三元锂离子电池热失控气体的毒性风险。结果表明：高SOC锂离子电池更易进入热失控状态，热失控释放的CO、HF及气体总量随SOC的增加而增加；锂离子电池SOC越高，热失控释放气体毒性风险越大，100% SOC锂离子电池毒性风险约为25% SOC锂离子电池的8倍，需要11倍的新鲜空气稀释才能达到安全浓度。研究结果可为锂离子电池热失控早期预警及气体毒性评价提供数据参考。

关 键 词：三元锂离子电池；热失控；气体传感器阵列；风险评估；毒性分析
中图分类号：X949
文献标志码：A　　**文章编号**：1001-5965(2024)01-0012-08

锂离子电池因其能量密度高、使用寿命长、充电效率高等优点被广泛应用于电动汽车和便携式电子设备中。同时，伴随全球经济一体化进程的加快和各种电子产品的普及，锂离子电池及以锂离子电池为能源供给的电子设备在航空运输中的比重也日益增加[1-2]。随着锂离子电池作为一种很有前途的能源和储能进入交通运输市场，人们开始担心其安全性。锂离子电池在电滥用、机械滥用、热滥用条件下可能会发生热失控，即电池内部放热反应产生的热量超过了电池散热的速度，导致电池内部温度和压力迅速上升，进一步加快了反应速率，这种自加速过程最终导致电池破裂排出大量易燃、有毒气体并释放热量[3-5]。在大型锂离子电池组中，其中一个失效电池产生的热量会加热邻近电池，最终导致整个电池组出现热失控蔓延，后果将更为严重[6-8]。

锂离子电池火灾的一个严重风险因素是有毒和易燃气体的排放，在相对封闭的环境中，如在飞机、潜水艇、矿井、装有电池储能系统的航天器或家中，锂离子电池热失控产生的有毒气体会是更大的威胁[9]。迄今为止，有不少学者对锂离子电池热失控释放气体进行了大量研究。锂离子电池热失控排出的气体可以包含蒸发的溶剂和分解产物，常见的有CO(一种窒息气体)、CO_2(引起缺氧)、H_2、CH_4[10-12]，除此之外，还会释放出一些有毒化合物[13-15]，其中，HF因具有非常强的毒性而引起极大关注[16-18]。电池中的氟不仅来自锂盐，如$LiPF_6$，还来自电极胶黏剂，如聚偏氟乙烯(PVDF)，电极材料和涂层中的氟磷酸盐及涂有AlF_3的阴极，以及含氟添加剂，如阻燃剂[19-20]。Ribière等[21]利用火焰传播装置(fire propagation apparatus, FPA)研究了袋装电池热失控释放气体，发现了大量的有毒气体，包括HF、CO、NO、SO_2和HCl，认为HF是含氟电池排放的最关键的气体。

目前，对锂离子电池释放气体危险性评估的研

收稿日期：2022-04-02；**录用日期**：2022-04-19；**网络出版时间**：2022-04-25 10:12
网络出版地址：link.cnki.net/urlid/11.2625.V.20220424.1842.003
基金项目：国家自然科学基金民航联合研究基金重点支持项目(U2033204)
*通信作者. E-mail：nkzqsong@126.com

引用格式：张青松，曲奕润，刘添添. 锂离子电池热失控气体毒性风险分析方法[J]. 北京航空航天大学学报，2024，50(1)：12-19.
ZHANG Q S，QU Y R，LIU T T. Risk analysis method for thermal runaway gas toxicity of lithium-ion batteries[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(1)：12-19 (in Chinese).

究较少, 对气体毒性了解十分有限。Diaz 等[22] 利用傅里叶变换红外光谱(Fourier transform infrared spectrometer, FTIR)和离子色谱(ion chromatography, IC)测量废旧锂离子电池机械和热处理过程中产生的气体, 计算了总废气量, 并评估了其毒性, 结果发现, 使用 $LiFePO_4$ 作为阴极电池产生的有毒气体最多, 一节锂离子电池在 379 m^3 的空间发生热失控才能确保安全环境。Sun 等[13] 对商用锂离子电池的燃烧产物进行分析, 鉴定出 100 多种气体产物, 其中, 大多数气体对人体有害且会造成环境污染, 此外, 研究发现, 有毒气体的产生与电池荷电状态(state of charge, SOC)有关, 且 100% SOC 电池释放气体毒性最大。以上研究主要集中在电池排放有毒气体的类型, 电池热失控排放气体毒性在线动态分析的研究较少。

尽管有毒气体对火灾救援人员和受害者的生命构成严重威胁, 但中国锂离子电池安全法规中尚未包含气体毒性标准及风险分析方法。本文使用成本低、能量密度高、被广泛应用于便携式电子设备和电动汽车领域的 18650 型商用三元锂离子电池进行热失控实验, 分析电池热失控特征参数, 利用气体传感技术, 研究不同 SOC 三元锂离子电池热失控全过程 CO、CO_2、O_2 和 HF 气体释放规律。依据立即威胁生命与健康浓度 (immediately dangerous to life or health, IDLH) 初步分析热解气体毒性, 基于气体毒性动力学模型重点对窒息性气体 CO、CO_2 和主要刺激性气体 HF 的毒性风险进行分析, 并提出应对措施, 为应急救援疏散提供理论支撑。

1　分析方法

1.1　气体成分分析

利用自主搭建的气体传感器阵列系统在线监测锂离子电池热失控全过程释放气体体积分数变化, 该系统基于一系列多传感技术的气体传感器。CO、CO_2 传感器为 GascardNG 红外气体传感器, 包括 1 个红外源、1 个气体取样单元、2 个红外滤光片和 1 个匹配的孪生元件探测器, 可提供真正的双波长红外感应, 具有低漂移、高精度、低功耗、快速响应的优点。传感器量程均为 0～0.30, 测量精度为 1×10^{-3}%, 要求流量在 0.2～1.2 L/min 的范围内, 输出幅值为 4～20 mA 电流信号。O_2、HF 传感器为基于电化学原理的物联型传感器, HF 传感器量程为 0～1×10^{-5}, 测量精度为 1×10^{-8}, O_2 传感器量程为 0～0.25, 测量精度为 1×10^{-3}。

1.2　锂离子电池热失控气体毒性风险分析

风险指某种特定的危险事件发生的可能性与其产生后果的组合。因此, 在评价锂离子电池热失控气体毒性风险时, 既要考虑电池发生热失控的可能性, 又要考虑电池热失控气体毒性危害。张宇等[23] 在考虑电池热失控可能性时引入点燃参数的概念, 而毒性危害则引入 ISO 13571[24] 和 ISO 13344[25] 中描述的有毒气体模型来综合评估毒性效应。

点燃参数 I 计算公式为

$$I = \lg(T_{TR}\Delta t) \tag{1}$$

式中: Δt为安全阀开启到热失控开始的时间间隔; T_{TR} 为发生热失控温度。

ISO 13571[24] 中用有效剂量分数(fractional effective dose, FED)估计在毒气环境内人可承受的时间, FED 值为 1 时, 意味着 50% 的人将无法安全逃生。FED 模型中提到 CO 和 HCN 是 2 种主要窒息性气体, 窒息性气体会导致组织细胞缺氧, 无法维持正常生命活动, 并且吸入 CO_2 会刺激呼吸, 导致呼吸加深加快, 增加其他有毒气体的吸收速率。具体公式为

$$\begin{cases} X_{FED} = \sum_{t_1}^{t_2} \dfrac{\phi_{CO_2} V_{CO}}{35\,000}\Delta t_i + \sum_{t_1}^{t_2} \dfrac{(\phi_{CO_2} V_{HCN})^{2.36}}{1.2\times10^6}\Delta t_i \\ \phi_{CO_2} = \exp\left(V_{CO_2}/5\right) \end{cases} \tag{2}$$

式中: V为每种气体的体积分数; Δt_i为时间增量; X_{FED} 为待测气体窒息性毒性。

ISO 13344[25] 中利用 FED 模型估算待测有毒气体 30 min 致死性。根据电池热失控释放气体实际情况可简化为

$$L_{FED} = \left(\frac{V_{CO}}{LC_{50.CO}} + \frac{V_{HF}}{LC_{50.HF}}\right) V_{CO_2} + Z_A + \frac{21 - V_{O_2}}{21 - 5.4} \tag{3}$$

式中: L_{FED} 为待测有毒气体 30 min 致死性; $LC_{50.CO}$= 5 700 μL/L 为 CO 的 LC50 值; $LC_{50.HF}$=2 900 μL/L 为 HF 的 LC50 值; Z_A 为酸中毒因子, 表示为 $V_{CO_2}\times0.05$。

锂离子电池热失控气体毒性风险分析方法表示如下:

$$K = SH = \frac{1}{I}\cdot\frac{L_{FED}}{t_{X_{FED}=1}}$$

式中: K 为风险率; S 为电池发生热失控的可能性; H 为毒性危害; $t_{X_{FED}=1}$为从热失控开始到 X_{FED} 值为 1 所经历的时间, s。

2　验证实验设计

2.1　电池样品

为获得不同的 SOC, 使用蓝电电池测试系统 LANHE CT2001B, 以恒流-恒压的方式对新出厂未经充放电操作的电池分别充放至 25% SOC、50% SOC、75% SOC 和 100% SOC。为减弱电池之间容量、内阻和 SOC 的影响, 每次充放电测试重复 3 次。将充

电后的电池置于温度为24℃的恒温箱中24 h,以保证电池的稳定性。电池阳极为石墨,阴极为NCM($LiNi_{0.8}Co_{0.1}Mn_{0.1}O_2$),电解液成分为$LiPF_6$、碳酸乙烯酯(EC)、碳酸甲乙烯酯(EMC)、碳酸二甲酯(DMC),基本信息如表1所示。

表1 实验电池参数

Table 1 Experimental battery parameters

参数	数值
高度/mm	65.1
直径/mm	18.5
额定容量/(mA·h)	3 500
额定电压/V	3.635
两端电压/V	3.59±0.05(25% SOC)
	3.72±0.05(50% SOC)
	3.92±0.05(75% SOC)
	4.16±0.05(100% SOC)

2.2 实验装置

图1为锂离子电池热失控气体实时原位检测装置。本文系统主要由2部分组成,包括滥用装置密封实验舱和热解气体分析系统。

密封实验舱的体积为20 L,实验舱的顶端配备2个出气孔:一个出气孔连接压力传感器,用来监测实验舱内气压;另外一个出气孔和底部进气孔配合,采用真空泵作为动力牵引,使实验舱内部气体经过滤和除湿后不断在外气路循环。在对电池进行加热前,将密封装置四周螺丝拧紧密封好,以防止内外气体交换。



图1 实验平台示意图

Fig. 1 Schematic diagram of experimental platform

2.3 实验方法

如图1所示,将同类型的CO、CO_2传感器串联后与串联的气杯式HF、O_2传感器并联构成传感器阵列,在正式实验前,电池去除塑料包装后在侧表面焊接一个K型热电偶,在正负极焊接电线,记录加热过程中电池表面温度和电压变化,将锂离子电池装进加热套筒进行电连接,并固定在实验舱底部;在实验舱内部电池上方3 cm位置固定安装一个K型热电偶,以测量电池释放气体的温度。实验舱通过螺丝密封,实验前将传感器通电预热10 min,开启真空泵,将气体从实验舱中抽出,利用转子流量计控制通过传感器的2条气路流量均为0.4 L/min,当气路里气流稳定,所有热电偶达到环境温度后,打开加热套筒,利用直流电源为加热套筒提供40 W稳定功率,对锂离子电池进行持续加热,当电池进入热失控时,立即关掉加热套筒电源。

3 验证结果与讨论

3.1 不同SOC锂离子电池热失控特性

SOC由锂离子在电池中的分布决定,高SOC的电池Li^+主要嵌入在负极材料中,低SOC的电池Li^+主要嵌入在正极材料中,0% SOC的电芯,其石墨负极中,嵌锂碳含量很少,而嵌锂碳与黏结剂、电解液反应释放的能量是热失控的重要动力[26]。为表征不同SOC对电池热失控危险性的影响并计算点燃参数,根据实验结果,选取并分析了安全阀开启到热失控开始的时间间隔Δt、压力第1次下降温度T_{drop}、安全阀开启温度T_{vent}、热失控起始温度T_{TR}、气体最高温度T_{gmax}、最大压力P_{max}、热失控后装置内稳定压力P_1、最大压力上升速率dP/dt_{max}等8种关键参数,不同SOC下以上参数的值及计算得出的点燃参数I如表2所示。这些关键参数随电

表 2　不同 SOC 电池关键参数

Table 2　Key parameters of batteries at different SOC

SOC	Δt/s	T_{drop}/℃	T_{vent}/℃	T_{TR}/℃	T_{gmax}/℃	P_{max}/MPa	P_1/MPa	dP/dt_{max}	I
25% SOC	182±44	149.39±5.8	167.44±2.62	239.26±20.23	367.36±30.50	0.315 8±0.127	0.115 8±0.004	0.082 48±0.062	2.86
50% SOC	140±10	144.47±2.88	162.29±3.23	224.26±6.25	396.86±23.46	0.262 4±0.012	0.126 7±0.003	0.084 63±0.037	2.72
75% SOC	127±19	142.91±4.05	158.71±2.61	213.49±5.89	669.45±196.3	0.265 3±0.059	0.134 3±0.003	0.092 64±0.033	2.66
100% SOC	105±19	133.78±1.30	148.14±3.50	193.86±4.33	868.36±334.4	0.311 2±0.067	0.137 8±0.003	0.102 16±0.037	2.53

池 SOC 的增加呈现出较为明显的规律性。随着 SOC 的增加, 压力第 1 次下降温度、安全阀开启温度和热失控起始温度都有所下降, 热失控时间提前, 安全阀开启到热失控开始的时间间隔明显缩短, 说明 SOC 越高, 电池稳定性越差, 越容易达到热失控状态, 热失控起始温度越低。高 SOC 电池热失控气体最高温度高, 说明内部材料的能量水平高, 热失控时释放更多的热量。安全阀打开时压力受 SOC 影响不大, 但随 SOC 升高, 电池热失控达到的最高压力、热失控后装置内稳定压力及最大压力上升速率明显提高, 说明高 SOC 电池热失控反应更剧烈, 释放更多气体。点燃参数可以反应热失控的难易程度, 其随 SOC 升高而降低, 说明 SOC 越高越容易发生热失控。

3.2　锂离子电池热失控释放的气体量

为研究 SOC 对三元锂离子电池热失控释放气体量的影响, 可利用理想气体状态方程计算密封实验舱内的气体量:

$$n = \frac{PV}{RT}$$

式中: n 为密封实验舱内所有气体的摩尔数; P 为实验舱内的压力; V 为实验舱内气体体积; R 为气体常数, 为 8.314 J/(mol·K); T 为气体温度。

通过计算得到热滥用前实验舱内气体的摩尔数 n_0 和热失控后气体的摩尔数 n_{TR}, 释放气体的摩尔数 Δn 为两者的差值, 在标准情况下, 释放气体体积 ΔV 为 Δn 与气体摩尔体积(22.4 L/mol)的乘积, 不同 SOC 下以上参数的值如表 3 所示。通过对比发现, 锂离子电池热失控释放气体量随 SOC 的增加而增加, 100% SOC 电池热失控释放 0.274 mol 气体, 25% SOC 热失控释放 0.017 mol 气体。这是因为锂离子电池热失控过程中很多反应的程度取决于嵌锂度, 而高 SOC 负极中嵌入的锂离子数量多, 与电解液发生化学反应的程度高, 材料热稳定性低, 热失控反应更剧烈, 因此, 释放的能量和气体更多。

表 3　不同 SOC 释放气体量

Table 3　Amountof vented gases at different SOC

SOC	n_0/mol	n_{TR}/mol	Δn/mol	ΔV/L
25% SOC	0.838	0.855	0.017	0.381
50% SOC	0.834	0.928	0.094	2.106
75% SOC	0.836	1.081	0.245	5.488
100% SOC	0.838	1.112	0.274	6.138

3.3　基于传感器阵列的热失控气体分析

锂离子电池热失控释放气体爆炸是导致灾难性后果的重要因素, 因此, 了解电池热失控气体释放规律且及时捕捉到气体体积分数上升信号显得尤为重要。之前对锂离子电池热失控气体释放机理进行了大量研究[22,27-29], 图 2 记录了三元锂离子电池经历热滥用导致热失控全过程 O_2、CO、CO_2 和 HF 体积分数随时间的动态变化过程。每次实验开始前, 密封实验舱内充满空气, O_2 体积分数为 0.21。可以看出, 当安全阀破裂后, 释放出的气体就会被传感器捕获, 所有电池均检测到 CO 和 CO_2 气体的上升信号, 没有检测到 HF 气体, 这是因为安全阀打开前电池内部温度低, 化学反应慢, 程度低, 释放气体量少。之后热失控阶段气体体积分数直线上升, 这是因为此时电池内部发生非常剧烈不可控的链式放热反应, 电池内部温度急剧升高, 从而释放出大量气体。在电池热失控后才检测到 HF 信号, 推测是因为热滥用前期电池内部温度低, $LiPF_6$ 没有明显的热行为表现, 随着内部温度不断升高, 电池在之后热失控早期先生成中间产物 PF_5, 之后才会继续反应生成 HF。图 2 表明, 使用气体传感器监测锂离子电池热失控事件可以产生灵敏、准确的早期诊断结果, 可用于早期预警和应急管理。

3.4　热失控释放气体毒性分析

3.4.1　基于 IDLH 的热失控气体毒性初步分析

当锂离子电池进入热失控后, 会释放大量有毒、可燃气体。本文主要关注 CO、CO_2 和 HF 的生成规律, 并定量分析相应的气体毒性。

三元锂离子电池热失控后实验舱内 CO、CO_2 及两者总和 CO_x 的物质的量如图 3 所示。可以看出, 电池热失控释放的 CO 的量随 SOC 的增加而增加。高 SOC 释放的 CO_2 普遍多于低 SOC, 100% SOC 释放 CO_2 低于 75% SOC, 但释放的 CO 明显高于 75% SOC。热失控释放 CO_x 总物质的量随 SOC 的增加而增加, 25% SOC 释放 0.170 mol, 50% SOC

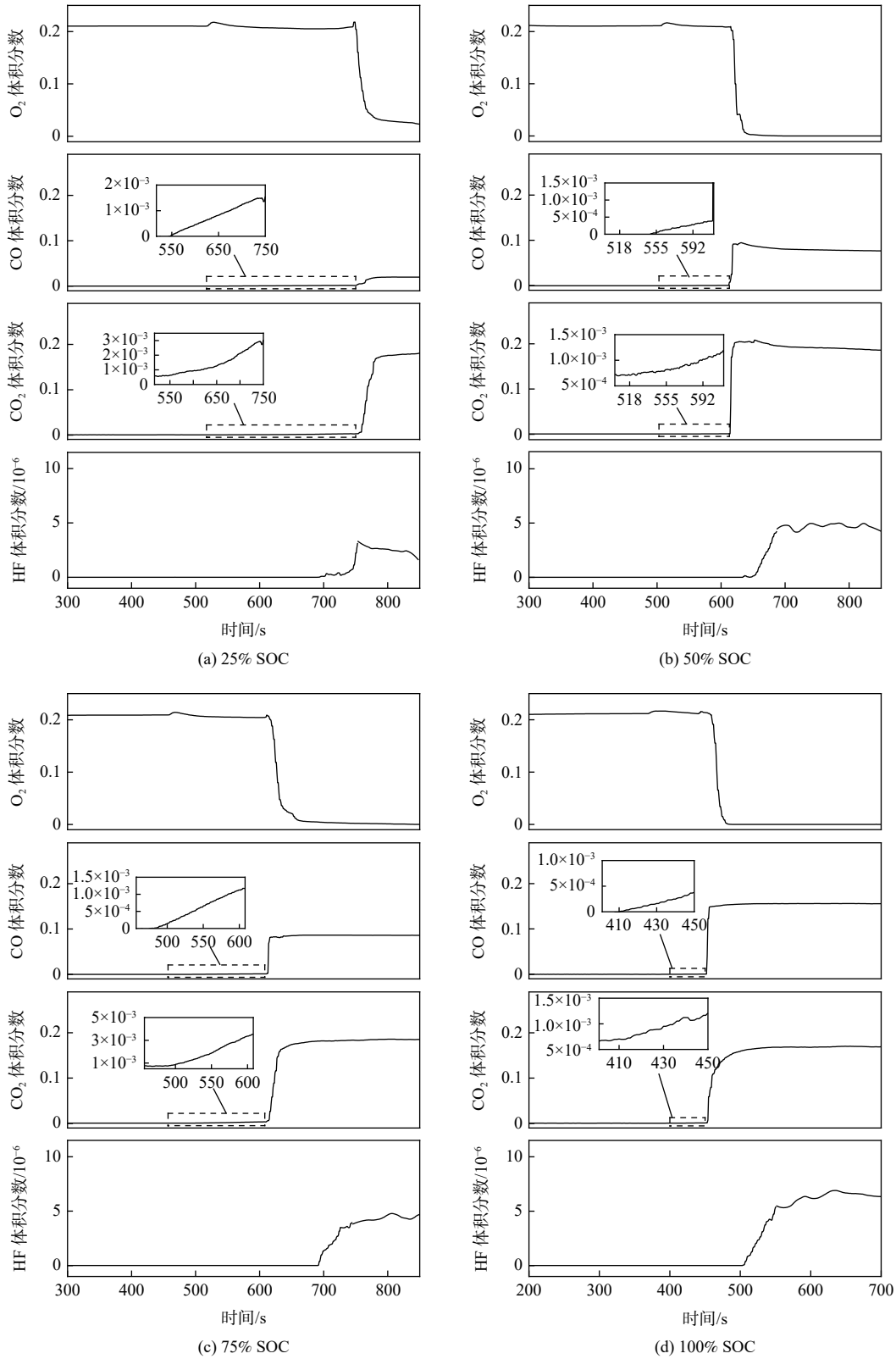


图 2 不同 SOC 电池热失控全过程气体分析

Fig. 2 Gas analysis of whole process of thermal runaway in different SOC batteries

释放 0.252 mol, 75% SOC 释放 0.292 mol, 100% SOC 释放 0.360 mol。CO/CO_2 比值可以从另一个角度反映热失控危险性, 25% SOC、50% SOC、75% SOC 和 100% SOC 电池 CO/CO_2 比值分别为 0.111、0.412、0.472 和 0.927, 可以看出, CO/CO_2 比值随 SOC 增高而增高, 表明高 SOC 电池热失控燃烧效率低。电池 SOC 越高, 热失控喷射越猛烈, 导致不完全燃烧, 通常会产生更多的有毒气体。

根据 IDLH 对 CO、CO_2 和 HF 的毒性进行初步评估, 这是对人体造成死亡或永久损害及丧失逃生

能力时的体积分数限值。具体 IDLH 值与不同 SOC 电池热失控后 3 种气体体积分数如表 4 所示。CO 和 HF 体积分数随 SOC 的增加而增加, 热失控后, 热解气体中 CO 和 CO_2 的体积分数远超 IDLH 限值, 而 100% SOC 电池释放 HF 最高体积分数为 6.29×10^{-6}, 未达到 IDLH 限值。可以看出, 对于一节 18650 型三元锂离子电池热解气体毒性危害主要由窒息性气体造成。

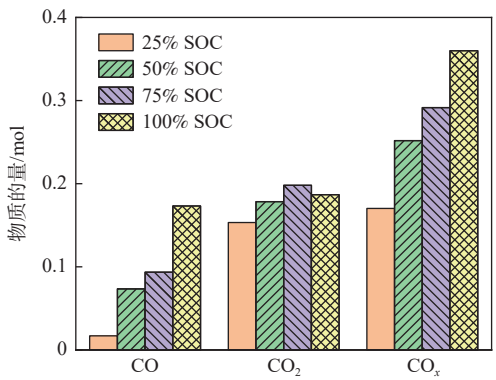


图 3　电池热失控后实验舱内 CO_x 物质的量

Fig. 3　Amount of CO_x products in experimental cabin after battery thermal runaway

表 4　不同 SOC 电池释放毒性气体体积分数

Table 4　Concentration of toxic gas released by batteries at different SOC

SOC	$V_{HF}/10^{-6}$	V_{CO}	V_{CO_2}
25% SOC	2.85	0.019 8	0.179 3
50% SOC	4.74	0.079 1	0.192 2
75% SOC	4.33	0.086 5	0.183 2
100% SOC	6.29	0.155 7	0.168 0

注: CO的IDLH限值为0.012, CO_2的IDLH限值为0.04, HF的IDLH限值为0.3。

3.4.2　基于气体毒性动力学模型的毒性风险分析

根据式(2)、式(3)中提到的 FED 模型计算方法, 结合图 2 中 CO、CO_2 和 O_2 实时体积分数数据, 绘制热失控气体毒性动力学模型, 如图 4 所示。可以看出, 在电池热失控后, 2 种 FED 模型都直线上升, 短时间内超过限值 1, 说明锂离子电池热失控释放气体毒性会迅速达到灾难性程度, 且 SOC 越高毒性越强。25% SOC、50% SOC、75% SOC、100% SOC 的电池分别在电池热失控 29 s、20 s、19 s、14 s 后, X_{FED} 值超过临界值 1, 热失控后 L_{FED} 值分别为 1.85、3.85、4.15、5.97。

利用表 2 中计算的点燃参数及图 4 给出的气体毒性动力学模型计算不同 SOC 电池热失控释放气体的毒性风险, 如表 5 所示。可以看出, SOC 越高, 热解气体毒性风险越大。假设 CO 的临界体积分数为 6.67×10^{-4}[30], 因为 6.67×10^{-4} 的 CO 就可能会导致高达 50% 的人体血红蛋白转化为碳氧血红蛋白。50% 的碳氧血红蛋白水平可能导致癫痫发作、昏迷和死亡。25% SOC、50% SOC、75% SOC 和 100% SOC 电池在 20 L 的密闭空间中发生热失控分别需要 549 L、2 446 L、3 122 L 和 5796 L 新鲜空气才能使 CO 体积分数降至 6.67×10^{-4} 安全值。

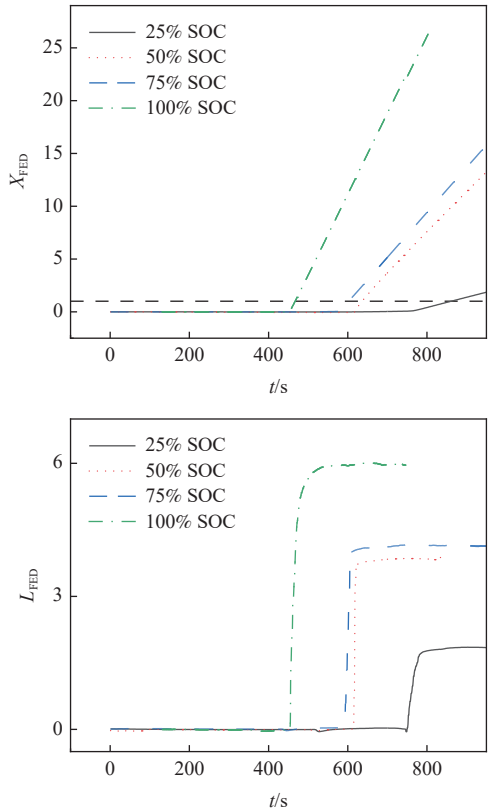


图 4　不同 SOC 电池热失控气体 FED 模型

Fig. 4　Thermal runaway gas FED modes by batteries at different SOC

表 5　不同 SOC 电池热失控气体毒性风险率

Table 5　Risk rate of thermal runaway gas toxicity of different SOC batteries

SOC	I	L_{FED}	$t_{X_{FED}=1}$/s	V
25% SOC	2.86	1.85	29	0.022
50% SOC	2.72	3.85	20	0.071
75% SOC	2.66	4.15	19	0.082
100% SOC	2.53	5.97	14	0.169

4　结　论

本文对不同 SOC 下三元锂离子电池热失控气体毒性进行全面分析, 主要结论如下:

1) 高 SOC 相对于低 SOC 电池具有热稳定性低、热失效快、热失控剧烈、释放气体更多等更严重的热危害特征。针对研究结果可合理制定排气

和热失控冲击处理的应对措施。

2）CO和CO_2气体信号均在电池安全阀打开后被捕捉到，且信号特征明显，电池热失控时才检测到HF信号，因此，CO和CO_2可作为电池热失控早期有效故障特征，为应急处置提供相对较长的时间。

3）HF、CO释放量与CO/CO_2比值随SOC的增加而增加，说明高SOC电池热失控燃烧效率低，会释放更多的有毒气体。100% SOC单体锂离子电池热失控释放HF体积分数为6.29×10^{-6}，未达到IDLH限值。全部测试电池的2种FED模型均在电池热失控后短时间超过临界值1，说明窒息性气体的影响比刺激性气体更显著，且高SOC电池热解气体毒性更强，会导致灾难性后果。

4）通过毒性风险分析，可知高SOC电池热失控释放气体毒性风险更大，一旦发生热失控需要更多的新鲜空气稀释。实际使用过程中，电池通常以模组形式存在，单节电池热失控易导致整个电池组发生多米诺热级联，具有更大的毒性风险。本文结果有助于进行有毒气体的风险评估，并为消防救援及安全排放提供定量参考。

参考文献（References）

[1] 齐创, 邝男男, 张亚军, 等. 高比能锂离子电池模组热扩散行为仿真研究[J]. 高电压技术, 2021, 47(7): 2633-2643.
QI C, KUANG N N, ZHANG Y J, et al. Simulation study on the thermal propagation behavior of high energy density lithium-ion battery module[J]. High Voltage Engineering, 2021, 47(7): 2633-2643 (in Chinese).

[2] 张青松, 曲奕润, 郝朝龙, 等. 三元锂离子电池热失控气体原位分析[J]. 高电压技术, 2022, 48(7): 2817-2825.
ZHANG Q S, QU Y R, HAO C L, et al. In-situ analysis of thermal runaway gas in ternary lithium-ion battery[J]. High Voltage Engineering, 2022, 48(7): 2817-2825(in Chinese).

[3] 张青松, 刘添添, 郝朝龙, 等. 锂离子电池热失控气体快速检测及危险性方法[J]. 北京航空航天大学学报, 2023, 49(9): 2227-2233.
ZHANG Q S, LIU T T, HAO C L, et al. Study on rapid analysis method of thermal runaway gas composition and risk of lithium ion battery[J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49(9): 2227-2233(in Chinese).

[4] 张青松, 罗星娜, 程相静, 等. 基于锂离子电池温降指数的细水雾添加剂筛选方法[J]. 北京航空航天大学学报, 2020, 46(6): 1073-1079.
ZHANG Q S, LUO X N, CHENG X J, et al. Method for screening fine water mist additive based on temperature drop index of lithium-ion battery[J]. Journal of Beijing University of Aeronautics and Astronautics, 2020, 46(6): 1073-1079(in Chinese).

[5] 张青松, 刘添添, 赵子恒. 锂离子电池热失控气体燃烧对热失控传播影响的量化方法研究[J]. 北京航空航天大学学报, 2023, 49(1): 17-22.
ZHANG Q S, LIU T T, ZHAO Z H. Study on quantitative method of the influence of thermal runaway gas combustion on thermal runaway propagation of lithium ion battery[J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49(1): 17-22(in Chinese).

[6] SAID A O, LEE C, STOLIAROVS I. Experimental investigation of cascading failure in 18650 lithium ion cell arrays: Impact of cathode chemistry[J]. Journal of Power Sources, 2020, 446: 227347.

[7] GAO S, LU L, OUYANG M, et al. Experimental study on module-to-module thermal runaway-propagation in a battery pack[J]. Journal of the Electrochemical Society, 2019, 166(10): A2065-A2073.

[8] FENG X, SUN J, OUYANG M, et al. Characterization of penetration induced thermal runaway propagation process within a large format lithium ion battery module[J]. Journal of Power Sources, 2015, 275: 261-273.

[9] LARSSON F, ANDERSSON P, BLOMQVIST P, et al. Author correction: Toxic fluoride gas emissions from lithium-ion battery fires[J]. Scientific Reports, 2018, 8(1): 5265.

[10] LAMB J, ORENDORFF C J, ROTH E P, et al. Studies on the thermal breakdown of common Li-ion battery electrolyte components [J]. Journal of the Electrochemical Society, 2015, 162(10): A2131-A2135.

[11] GOLUBKOV A W, FUCHS D, WAGNER J, et al. Thermal-runaway experiments on consumer Li-ion batteries with metal-oxide and olivin-type cathodes[J]. RSC Advances, 2014, 4(7): 3633.

[12] ABRAHAM D P, ROTH E P, KOSTECKI R, et al. Diagnostic examination of thermally abused high-power lithium-ion cells[J]. Journal of Power Sources, 2006, 161(1): 648-657.

[13] SUN J, LI J, ZHOU T, et al. Toxicity, a serious concern of thermal runaway from commercial Li-ion battery[J]. Nano Energy, 2016, 27: 313-319.

[14] LEBEDEVA N P, BOON-BRETT L. Considerations on the chemical toxicity of contemporary Li-ion battery electrolytes and their components[J]. Journal of the Electrochemical Society, 2016, 163(6): A821-A830.

[15] NEDJALKOV A, MEYER J, KOHRING M, et al. Toxic gas emissions from damaged lithium ion batteries—Analysis and safety enhancement solution[J]. Batteries, 2016, 2(1): 5.

[16] STURK D, ROSELL L, BLOMQVIST P, et al. Analysis of Li-ion battery gases vented in an inert atmosphere thermal test chamber[J]. Batteries, 2019, 5(3): 61.

[17] National Research Council. Acute exposure guideline levels for selected airborne chemicals (Volume 13)[M]. Washington, D.C.: National Academies Press, 2013.

[18] LIU P J, LI Y Q, MAO B B, et al. Experimental study on thermal runaway and fire behaviors of large format lithium iron phosphate battery[J]. Applied Thermal Engineering, 2021(1): 1-13.

[19] PARK Y U, SEO D H, KIM B, et al. Tailoring a fluorophosphate as a novel 4 V cathode for lithium-ion batteries[J]. Scientific Reports, 2012, 2: 704.

[20] LARSSON F, BERTILSSON S, FURLANI M, et al. Gas explosions and thermal runaways during external heating abuse of commercial lithium-ion graphite-$LiCoO_2$ cells at different levels of ageing[J]. Journal of Power Sources, 2018, 373: 220-231.

[21] RIBIÈRE P, GRUGEON S, MORCRETTE M, et al. Investigation on the fire-induced hazards of Li-ion battery cells by fire calorimetry[J]. Energy & Environmental Science, 2012, 5(1): 5271-5280.

[22] DIAZ F, WANG Y, WEYHE R, et al. Gas generation measurement

and evaluation during mechanical processing and thermal treatment of spent Li-ion batteries[J]. Waste Management, 2019, 84: 102-111.

[23] 张宇, 白伟, 史砚磊, 等. 基于热失控风险指数的锂电池安全评价方法[J]. 北京航空航天大学学报, 2021, 47(5): 912-918.
ZHANG Y, BAI W, SHI Y L, et al. Evaluation method of lithium battery safety based on thermal runaway risk index[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(5): 912-918(in Chinese).

[24] International Organization for Standardization. Life-threatening components of fire-guidelines for the estimation of time to compromised tenability in fires: ISO 13571[S]. Geneve: International Organization for Standardization, 2012.

[25] International Organization for Standardization. Estimation of the lethal toxic potency of fire effluents: ISO 13344[S]. Geneve: International Organization for Standardization, 2015.

[26] 王青松, 平平, 孙金华. 锂离子电池热危险性及安全对[M]. 北京: 科学出版社, 2017: 123-124.
WANG Q S, PING P, SUN J H. Thermal hazards and safety measures of lithium-ion batteries[M]. Beijing: Science Press, 2017: 123-124(in Chinese).

[27] YANG H, ZHUANG G V, ROSS JR P N. Thermal stability of $LiPF_6$ salt and Li-ion battery electrolytes containing $LiPF_6$[J]. Journal of Power Sources, 2006, 161(1): 573-579.

[28] KAWAMURA T, OKADA S, YAMAKI J I. Decomposition reaction of $LiPF_6$-based electrolytes for lithium ion-cells[J]. Power Sources, 2006, 156(2): 547-554.

[29] WILKEN S, TRESKOW M, SCHEERS J, et al. Initial stages of thermal decomposition of $LiPF_6$-based lithium ion battery electrolytes by detailed Raman and NMR spectroscopy[J]. RSC Advances, 2013, 3(37): 16359-16364.

[30] TIKUISIS P, KANE D, MCLELLAN T, et al. Rate of formation of carboxyhemoglobin in exercising humans exposed to carbon monoxide[J]. Journal of Applied Physiology, 1992, 72(4): 1311-1319.

Risk analysis method for thermal runaway gas toxicity of lithium-ion batteries

ZHANG Qingsong*, QU Yirun, LIU Tiantian

(Key Laboratory of Civil Aviation Thermal Hazards Prevention and Emergency Response, Civil Aviation University of China, Tianjin 300300, China)

Abstract: A method to analyse the toxicity of thermal runaway gases of batteries, referring to the risk assessment theory was proposed. It aims to investigate the hazard ratings of the main harmful substances in the thermal runaway gas of lithium-ion batteries.By using the fractional effective dose (FED) equation and a gas sensor array, the results of the method used to characterize the likelihood of thermal runaway of lithium batteries occurred were discovered.The thermal runaway gas toxicity kinetic model of the battery was demonstrated in order to identify the consequences of gas toxicity, while the toxicity risk of thermal runaway gas in ternary lithium-ion batteries under different states of charge (SOC) was analyzed accordingly. The results retrieved show that high SOC batteries are more likely to enter the thermal runaway state, and the total amount of CO, HF and gas released by thermal runaway increases according to the SOC of batteries. As a result, the risk of thermal runaway increases with battery SOC.A fully (100%) charged battery has about 8 times the toxicity risk of a 25% charged battery and requires 11 times the fresh air dilution to reach a safe concentration.The results provided data reference for the early warning of lithium battery thermal runaway and the toxicity evaluation of pyrolysis gas.

Keywords: ternary lithium-ion battery; thermal runaway; gas sensor array; risk assessment; toxicity analysis

Received: 2022-04-02; **Accepted**: 2022-04-19; **Published Online**: 2022-04-25 10:12

URL: link.cnki.net/urlid/11.2625.V.20220424.1842.003

Foundation item: Key Program of the Joint Fund for Civil Aviation Research Fund with National Natural Science Foundation of China (U2033204)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0220

ML组合的CYGNSS海面风速反演质量控制模型

张云[1,2]，赵星宇[1,2]，杨树瑚[1,2,*]，孙聪[3]，韩彦岭[1,2]，尹继伟[1,2]

(1. 上海海洋大学 信息学院，上海 201306；2. 上海海洋大学 上海市海洋智能信息与导航遥感工程中心，上海 201306；
3. 上海航天空间技术有限公司，上海 201109)

摘　　要：卷积神经网络(CNN)可用于气旋全球导航卫星系统(CYGNSS)的海面风速反演。虽然在模型训练前设置了质量控制指标来检测和削弱CYGNSS的异常观测数据，但CYGNSS观测数据中仍存在异常值导致模型反演精度降低，甚至出现错误反演结果。因此，提出一种基于机器学习(ML)组合的海面风速反演模型。在基于CNN回归模型的CYGNSS反演海面风速基础上，ML分类模型生成CNN回归结果的质量标志位，该标志位可以检测并删除CNN回归结果的异常值，进一步提高风速反演结果的数据质量，ML分类模型能够更好地考虑各种数据误差之间的相互作用，而不是单独使用每个条件的阈值，以达到更优的海面风速反演精度的效果。实验对比了Logistic回归(LR)、决策树(DT)、朴素贝叶斯模型、K最邻近(KNN)算法、神经网络(NN)模型、支持向量机(SVM)算法等6个分类模型，其中，基于KNN算法的分类模型对风速反演质量控制的效果最优。所提风速反演组合模型显著提高了反演结果的精度，在0～20 m/s区间内，异常样本过滤率为81.27%，在所有被过滤的数据中，过滤正确率为86.03%；风速反演误差的均方根误差从无ML分类模型的1.7 m/s降低到有ML分类模型的1.44 m/s，其中，训练样本为0～10 m/s的反演结果精度提升效果较为明显，证明了所提风速反演组合模型对风速质量控制的有效性。

关　键　词：气旋全球导航卫星系统；风速反演；质量控制；机器学习组合模型；卷积神经网络；K最邻近算法

中图分类号：V221^{+}.3; TB553

文献标志码：A　　**文章编号**：1001-5965(2024)01-0020-10

全球导航卫星系统反射信号(global navigation satellite system reflectometry, GNSS-R)利用导航卫星的信号反射原理进行遥感研究，能够在复杂天气下工作，广泛应用于海面测高[1]、海面测风[2]、海冰检测[3]、土壤湿度探测[4]等领域。海面风场与海洋表面大气环流有着密切的关系，是大气与海洋相互作用的一项重要参数。同时，近地面风是海洋与大气界面动量和能量的主要贡献者与指标。准确探测海风不仅能够预测海面大气环流趋势，同时也为人类海上作业提供可行性预测，保证人们的海上作业平稳安全进行。星载GNSS-R受天气环境影响较小，可以对全球范围内进行高时空分辨率的探测，气旋全球导航卫星系统(cyclone global navigation satellite system, CYGNSS)是近年来在GNSS-R领域被广泛应用的卫星系统，其主要对热带气旋和台风生命周期内的眼壁内和眼壁附近的海洋表面风进行频繁测量，将测量到的数据用于海面风场反演，对研究海洋风场遥感有重要意义。

目前，CYGNSS卫星数据海面风速反演的数据处理中，在影响数据质量的不同方面已经设置了许

收稿日期：2022-04-06；**录用日期**：2022-06-11；**网络出版时间**：2022-07-27 17:10
网络出版地址：link.cnki.net/urlid/11.2625.V.20220727.1416.001
基金项目：国家自然科学基金(41871325,42176175)；国家重点研发计划(2019YFD0900805)
*通信作者. E-mail：shyang@shou.edu.cn

引用格式：张云，赵星宇，杨树瑚，等. ML组合的CYGNSS海面风速反演质量控制模型[J]. 北京航空航天大学学报，2024，50(1)：20-29.
ZHANG Y，ZHAO X Y，YANG S H，et al. Quality control model of CYGNSS sea surface wind speed retrieval based on ML combination[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(1)：20-29 (in Chinese).

多质量控制 (quality control, QC) 标志[5], 以便在实验过程中选择达到质量要求的数据。虽然有多种严格的筛选条件来控制样本质量, 但还是会存在无法去除的异常值[5]。在工程层面, 主要的误差来源可归结为全球定位系统 (global positioning system, GPS) 发射功率估计、GPS 天线方向图、CYGNSS 指向星跟踪器/科学天线轴视偏差[5] 等方面。在仪器层面, 可能的误差包括科学天线增益误差、仪器噪声功率估计误差、校准负载温度误差和数模量化误差[5]。在测量层面, 由于高光点地理定位误差、观测数据对其他地球物理参数 (如波龄、膨胀等) 的依赖、从风速到观测数据的非唯一映射及与参考风场空间和时间插值的误差, 这部分异常值最终会影响风速反演精度。

到目前为止, 绝大部分 CYGNSS 海面风速反演研究中, 都利用了 CYGNSS 的 L1 数据中提供的质量控制标志[5]、观测值的距离校正增益 (range-corrected gain, RCG) 及卫星天线入射角等经验值作为输入数据的质量控制指标[6-7], 实现数据预处理去除观测数据异常值, 但是对于质量控制指标或方法没有做进一步的分析, 研究重点都集中在机器学习方法和特征参量的改进[8-10], 在 CYGNSS 官方文档[11] 中也主要介绍了 CYGNSS 的公开 L1 级数据中的信号、接收机和系统等相关观测量的质量控制标志。综上所述, 在目前的研究方法中, 都是利用观测数据的质量控制标志实现质量控制, 如去除仪器噪声功率所产生的误差、卫星高度超过额定高度所产生的误差等。但是, 依靠单一条件的阈值无法完全去除观测量中的异常值。因此, 可以进一步通过有效识别和消除反演结果的异常值来提高 CYGNSS 的反演精度, 但是关于建立反演结果的质量控制标志的研究还较为缺乏[12]。

针对这一问题, 本文在卷积神经网络 (convolutional neural networks, CNN) 反演 CYGNSS 海面风速[13] 的研究成果基础上, 提出使用质量控制标志筛选样本, 在 CNN 回归模型反演海面风速后, 增加机器学习 (machine learning, ML) 分类模型来有效识别并过滤异常值, 构成 ML 组合的海面风速反演质量控制网络模型。本文设计的基于 ML 分类模型的质量过滤器的目的仅是去除异常值, 而不是纠正反演后的偏差[7]。使用 CYGNSS 7 d 的数据进行 ML 组合模型训练, 反演结果表明, ML 组合模型的反演精度相较于单一 CNN 模型有了较明显的提高, 其中, CNN-KNN 分类模型的精度最高。由此证明了本文模型可以整体改善测试集风速反演质量, 从而提升风速反演精度。

1　数 据 集

1.1　CYGNSS 数据集

CYGNSS 卫星于 2016 年底发射, 共由 8 颗卫星组成, 每颗卫星重约 30 kg, 卫星轨道倾角为 35°, 轨道高度约为 510 km[14]。每颗卫星配置 1 根右旋圆极化天线接收发射卫星的直射信号和 2 根左旋圆极化天线接收来自海面或陆地的反射信号。CYGNSS 将 GPS 双基地雷达的全天候性能与时空采样特性结合, 解决了传统观测受天气与时空限制的局限性。CYGNSS 的飞行任务是了解海洋表面性质及海洋表面大气热力学, 改进飓风观测工作, 得到更准确的飓风信息。早期的科学活动侧重于风速反演算法的开发和数据产品的验证。美国国家航空航天局 (National Aeronautics and Space Administration, NASA) 于 2017 年 11 月大西洋台风季结束时开始非临时公开发布这些产品。CYGNSS 数据被划分为 L0～L3 等 4 个层级, 本文使用 2019 年 1 月 2 日、1 月 22 日、2 月 2 日、2 月 22 日、3 月 2 日、3 月 22 日、4 月 2 日、5 月 2 日、5 月 22 日、6 月 2 日的 CYGNSS L1 级数据进行分析。

1.2　ECMWF 数据集

欧洲中期天气预报中心 (European Centre for Medium-range Weather Forecasts, ECMWF) 再分析数据集将模型数据与各地观测数据结合成为一个完整的数据集, 用来生成对大气状态的最佳估计[15], 是目前应用最为广泛的海面风场数据。ECMWF 风场数据是时空分辨率为 1 h、0.25°, 海表面 10 m 高度处的经度方向和纬度方向风速, 使用的是 UTC 时间。因为 ECMWF 海面风场数据通过时空插值可以得到与 CYGNSS 同一位置相近时间 (小于 0.5 h) 的海面风速, 能够较准确描述海面的矢量风速, 所以选择 ECMWF 再分析风场数据作为真值参与模型的训练与结果验证。

2　方　法

2.1　ML 组合模型

本文结合 CNN 的优势, 在 CNN 训练海面风速反演模型基础上, 增加筛选样本质量标签的 ML 分类模型, 实现 CNN 反演结果的分类过滤, 得到最终过滤后的海面风速反演结果。与单独使用 CNN 模型反演风速相比, ML 组合模型能够有效筛掉异常数据, 进一步提高海面风速反演精度。

2.1.1　CNN 回归模型

CNN 是一类包含卷积计算且具有深度结构的前馈神经网络, 是深度学习的代表算法之一。与传

统的人工神经网络结构不同，其包含非常特殊的卷积层和降采样层。卷积层和前一层采用局部连接和权值共享的方式进行连接，大大降低了参数的数量。降采样层可以降低输入数据维度和网络复杂性，使CNN较其他神经网络方法更具有鲁棒性[16]。CNN模型的卷积操作和输出如下：

$$\boldsymbol{V}=\text{conv2}(\boldsymbol{W},\boldsymbol{X},\text{“valid”})+\boldsymbol{b} \tag{1}$$

$$\boldsymbol{Y}=\varphi(\boldsymbol{V}) \tag{2}$$

式中：conv2()为卷积运算函数；$\boldsymbol{W}$为卷积核矩阵；$\boldsymbol{X}$为输入矩阵；“valid”字段表示只进行有效卷积，对边界不进行处理；$\boldsymbol{b}$为偏置项矩阵；$\boldsymbol{V}$为卷积操作得到的矩阵；$\varphi(\boldsymbol{V})$为激活函数；$\boldsymbol{Y}$为CNN模型的输出矩阵。

2.1.2 ML分类模型

本文考虑了6种不同的ML分类模型：Logistic回归(Logistic regression, LR)[17]、决策树(decision tree, DT)、朴素贝叶斯模型(naive Bayes model, Bayes)[18]、K最邻近(K-nearest neighbor, KNN)算法、神经网络(neural network, NN)模型和支持向量机(support vector machine, SVM)算法[19]。

1）LR针对线性可分问题，根据现有数据对分类边界建立回归公式，以此进行分类。在本文实验中，利用梯度下降算法来确定最优回归系数。

2）DT学习的关键在于选择最优的划分属性，即尽可能地使同类样本划分到同一个类别中，以保持样本的纯度最高，这时需要用到信息熵的概念[20]。信息熵E为当前样本集的纯度，表示如下：

$$E(D)=-\sum_{c=1}^{C}p_c\log_2 p_c \tag{3}$$

式中：D为当前样本集；C为类别的总数；p_c为第c类样本在D这个样本集中的比例；$E(D)$为信息熵，值越小，则D的纯度越高[20]。

3）Bayes以贝叶斯定理为基础，即对于给出的待分类项，求解在此项出现的条件下各类别出现的概率，最大概率所对应的类别即为此待分类项的类别。

4）在KNN算法中，K值的选择是重要问题[21]。因此，本文实验中采用5折交叉验证法来选择最合适的K值。

5）NN模型中[22]采用具有2层隐藏层的前馈神经网络模型，隐藏层的节点数均为30个，网络的输入为26维样本特征，输出层有一个神经元，指明该样本为异常值或良好样本。

6）SVM是一种有监督的学习方式，算法核心是找到一个超平面，使得分类后样本中所有的点距离这个超平面有最大距离[23]。

2.1.3 ML组合模型的结构

ML组合模型反演海面风速的技术路线可简要概括为4个部分：①确定所使用的卫星数据和风速数据，对确定好的数据进行预处理；②将处理好的数据使用CNN回归模型进行训练，得到反演海面风速结果；③制作基于CNN反演海面风速结果的质量标志训练标签，实现ML分类模型训练，对反演结果进行质量分类；④过滤质量差的数据，保留剩余反演结果。

图1为本文使用的ML组合模型结构。根据大量的数据验证后，最终确定CNN卷积层数为3，未设置池化层，卷积核大小为3×1，隐藏层节点抛弃率dropout为0.3，卷积核个数均为32个，单次训练的参数个数batch-size为1 000，训练次数epochs为2 000。其中，ML分类模型部分为2.1.2节介绍的6种分类模型。

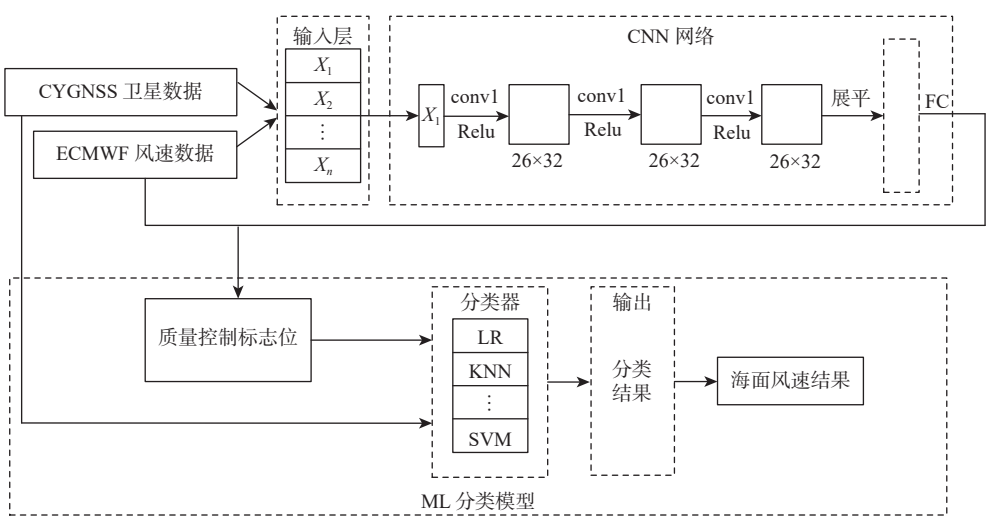


图1 ML组合模型结构

Fig. 1 Structure of ML combined model

ML 分类模型流程如图 2 所示。图中: $U_{\text{retrieval}}$ 为 CNN 模型反演的风速结果, U_{real} 为由 ECMWF 风速数据代表的真实风速。

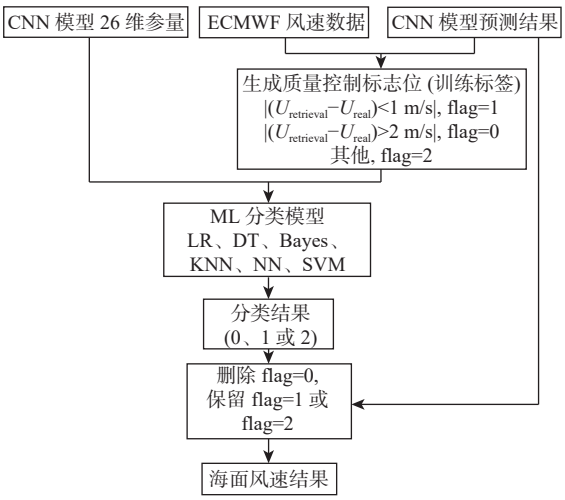


图 2　ML 分类模型流程

Fig. 2　Flowchart of ML classification model

图 2 中, 根据 CNN 训练模型的反演风速结果与 ECMWF 风速数据 (真值风速) 比较后生成风速质量控制标志, 作为 ML 分类器分类标签, 与 CYGNSS 特征参量一起进行 ML 分类模型训练。在 ML 分类模型的训练中, 根据经验, 若反演风速与真实风速之差的绝对值小于 1 m/s, 则风速质量控制标志设置为“1”; 若反演风速与真实风速之差的绝对值大于 2 m/s, 则风速质量控制标志设置为“0”; 其他情况下的风速质量控制标志设置为“2”。ML 分类模型的输出为每个预测风速值对应的标志位, 将 CNN 模型反演结果中对应标志位为 0 的数据筛除, 保留其余分类结果 (标志位为 1 和 2), 实现风速结果过滤功能, 提高风速结果质量。

ML 组合模型训练完成后进行海面风速反演的过程中, CNN 回归模型和 ML 分类模型可以分别输入相同 CYGNSS 的特征参量进行结果反演。但是, 2 个模型的主要目的不同, CNN 回归模型用来根据 CYGNSS 的特征参量反演风速, ML 分类模型则用来根据 CYGNSS 的特征参量得到风速结果的标志位 (0,1,2), 通过该标志位可以检测并过滤 CNN 反演结果中的异常值 (flag=0), ML 分类结果的意义在于能够表征基于 CNN 回归模型的风速结果的质量控制标志, 但是不能显示具体的风速结果, 因此, 只有与 CNN 回归结果结合使用才有意义。

2.2　特征量选择

CYGNSS 的 L1 级数据产品包括多种卫星可观测值, 如取决于风速及海面粗糙度的延迟多普勒平均图 (delay-Doppler map average, DDMA) 和前沿斜率 (leading edge slope, LES) 等[11,13]。在现有的质量控制标志上, 由于 GNSS 反射信号的二维延迟多普勒功率波形随风速变化较大, 导致风速反演误差变大[13], 可以选择与反射信号能量、反射信号延迟及时延多普勒图 (delay-Doppler map, DDM) 等相关的多维特征量进行训练, 以提高特征参量与风速之间的相关性。本文提取了 CYGNSS L1 级数据中 26 维与风速相关的特征参量用于优化测风模型精度[13]。其中, DDM 归一化双基地雷达散射截面、DDM 峰值多普勒列、LES、GPS 伪随机噪声码、直射信号码相位、飞行软件镜面反射点延迟、DDM 峰值延迟行、DDM 镜面反射点延迟行、镜面反射点多普勒列、距离校正增益、镜面反射点延迟校正及 DDM 信噪比共 12 维参量与反射信号能量值相关; CYGNSS/GPS 位置和速度、UTC 时间及入射角共 14 维参量与 CYGNSS 系统观测量相关。

2.3　模型训练与验证

本文实验在配备 Intel Core i7-5600H 处理器、16 GB RAM 和 NVIDIA GTX 1060 显卡的 PC 上实施。由于 PC 性能的限制, 无法使用千万级别的数据量进行训练, 为保证训练集特征量的时间跨度和多样性, 本文使用的训练集横跨 3 个月, 共 7 d (2019 年 1 月 2 日、1 月 22 日、2 月 2 日、2 月 22 日、3 月 2 日、3 月 22 日、4 月 2 日)的 CYGNSS L1 级数据与 ECMWF 数据做时间空间插值。为使训练数据更具有泛化性, 对插值后的数据集进行质量控制筛选, 并使用洗牌算法将 1 d 内的全部数据打乱, 每天随机选取 15 万条数据, 7 d 共 105 万条数据作为训练数据。为证明模型的泛化性能, 测试集横跨 1 个月, 共 3 d(2019 年 5 月 2 日、5 月 22 日、6 月 2 日), 每天与训练集同样方法随机选取 10 万条数据, 共 30 万条数据作为测试集。训练集与测试集风速分布直方图如图 3 所示。

实验前, 为增强各参量之间的相关性, 在使用任何对地观测的遥感卫星时均需进行严格的数据质量控制。按以下标准对 CYGNSS 数据集进行质量控制:

1) CYGNSS 的 L1 级数据集提供的数据质量控制标志, 该标志位通过比特位的组合计算定义了各类质量异常 (具体计算方法可查看 CYGNSS 数据文档)。

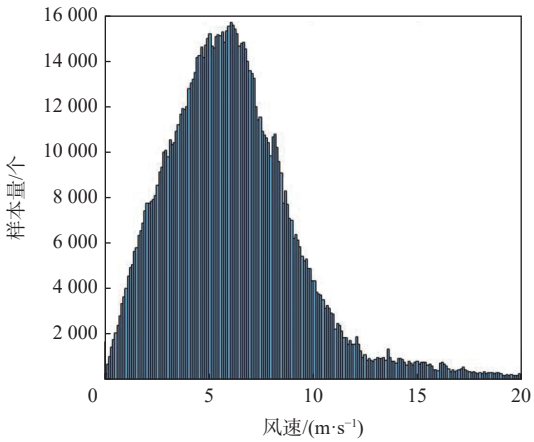
2) CYGNSS 观测值和风速匹配数据均为正值。

3) 观测值的 RCG 大于 10, RCG 的定义和描述见文献 [24]。

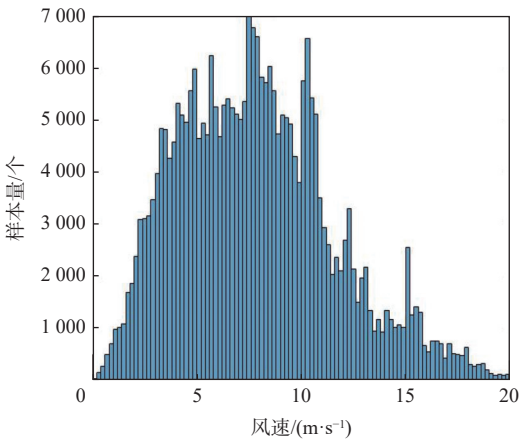
4) 卫星天线入射角小于 60°。

5) 镜面反射点在海上。

在精度验证部分, 本文采用 4 种标准对模型性



(a) 训练集数据分布



(b) 测试集数据分布

图3 训练集与测试集数据分布

Fig. 3 Training data distribution and test data distribution

能进行评估：①偏差 B；②平均绝对误差 M；③均方根误差 R；④相关指数 R^2。相应的计算公式如下：

$$B=\frac{1}{n}\left[\sum_{a=1}^{n}(\tilde{x}_a-x_a)\right] \tag{4}$$

$$M=\frac{1}{n}\sum_{a=1}^{n}(|\tilde{x}_a-x_a|) \tag{5}$$

$$R=\sqrt{\frac{1}{n}\sum_{a=1}^{n}(\tilde{x}_a-x_a)^2} \tag{6}$$

$$R^2=1-\frac{\sum_{a=1}^{n}(x_a-\tilde{x}_a)^2}{\sum_{a=1}^{n}(x_a-\bar{x}_a)^2} \tag{7}$$

式中：$\tilde{x}_a$为反演值；x_a为真实值；$\bar{x}_a$为真实值的平均值。

3 结果与讨论

3.1 ML 分类模型的过滤性能比较

通过已训练好的 CNN 模型对测试数据集进行海面风速反演，分别分析使用不同 ML 分类模型下的过滤性能来选择最优的过滤组合模型。图 4 为 CNN 和 ML 组合模型训练后真实风速与反演风速对比，具体样本量参照图 3。直线 $y=x$ 表示反演风速与真实风速相等，越靠近直线则反演效果越好。

由图 4 对比可知，经过 KNN 过滤后的反演数据较其他 5 个分类模型过滤后的反演数据更加集中于 $y=x$ 函数直线上，离群值与异常值过滤效果相较其他 3 个模型更加明显，且过滤前的相关系数为 0.69，过滤后的相关系数为 0.76，提升了 10.1%，说明 KNN 分类模型相对其他分类模型对 CYGNSS 质量控制有着显著的作用。采用反演结果的均方根误差作为实验的精度验证标准，如表 1 所示。单独使用 CNN 反演结果的均方根误差为1.70 m/s，使用 CNN-LR 模型后的均方根误差为 1.68 m/s，性能提升了 1.18%；使用 CNN-DT 模型后的均方根误差为 1.52 m/s，性能提升了 10.6%；使用 CNN-Bayes 模型后的均方根误差为 1.59 m/s，性能提升了 6.5%；使用 CNN-KNN 模型后的均方根误差为 1.44 m/s，性能提升了 15.3%；使用 CNN-NN 模型后的均方根误差为 1.49 m/s，性能提升了 12.4%；使用 CNN-SVM 模型后的均方根误差为 1.62 m/s，性能提升了 4.7%。通过对比可得，CNN-KNN 模型无论是从均方根误差还是反演结果的相关性上都优于其他 ML 组合模型。由此可基本确定 CNN-KNN 模型对本文实验的过滤性能是最优的。

通过大量实验并参考文献 [22] 后，对不同分类模型在相同训练集下的过滤性能进行比较，如表 2 所示。

表 2 中，过滤正确率表示被分类模型过滤掉的数据中异常样本 (反演风速与真实风速之差的绝对值大于 2 m/s) 的数量与被分类模型过滤掉的所有样本数量的比值，异常样本过滤率表示被过滤掉的异常样本的数量与测试集中所有异常样本数量的比值，表示为

$$A=\frac{f_{\text{bad}}}{f_{\text{all}}} \tag{8}$$

$$R_{\text{bad}}=\frac{f_{\text{bad}}}{b_{\text{all}}} \tag{9}$$

式中：A 为过滤正确率；f_{bad} 为过滤异常样本个数；f_{all} 为过滤总个数；R_{bad} 为异常样本过滤率；b_{all} 为异常样本总个数。

分析表 2 可知，LR 回归耗时最少，即时间复杂度最低，但过滤正确率和异常样本过滤率也是最低。其中，时间复杂度低是 LR 回归的优点之一，模型的训练耗时只取决于输入特征数量的多少；而且 LR 回归常用于解决线性问题，对于实验中涉及到

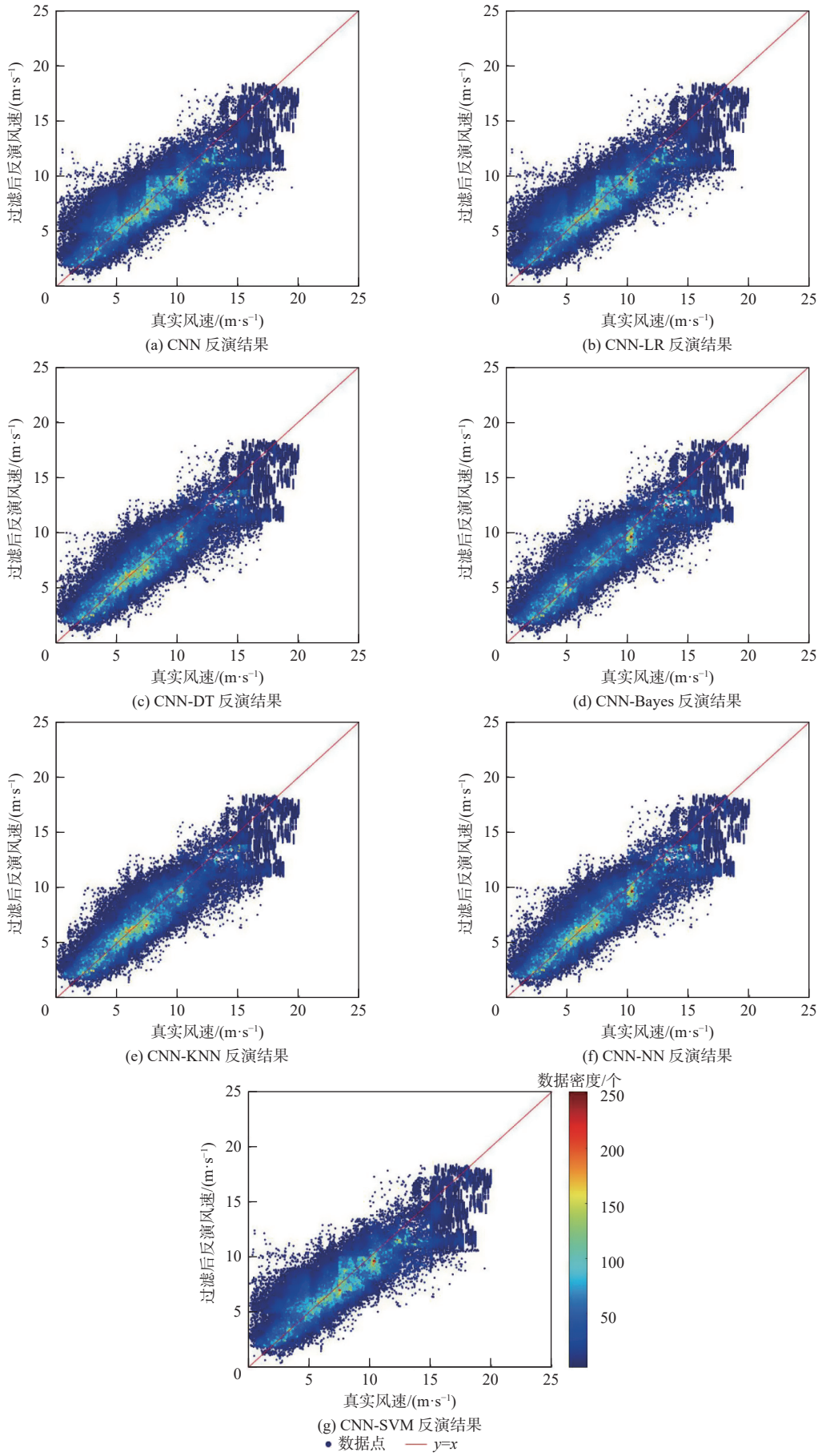


图 4　反演风速与真实风速对比散点图

Fig. 4　Scatter plot of inverted wind speed vs. true wind speed

表1 ML组合模型性能分析

Table 1 Performance analysis of ML combination model

模型	M/(m·s^{-1})	R/(m·s^{-1})	B/(m·s^{-1})	R^2
CNN	1.47	1.70	0.21	0.69
CNN-LR	1.46	1.68	1.32	0.71
CNN-DT	1.34	1.52	0.22	0.73
CNN-Bayes	1.41	1.59	0.19	0.73
CNN-KNN	1.19	1.44	0.17	0.76
CNN-NN	1.28	1.49	0.13	0.69
CNN-SVM	1.43	1.62	1.92	0.61

表2 分类模型过滤性能比较

Table 2 Comparison of filtering performance of classification models

分类模型	过滤正确率/%	异常样本过滤率/%	耗时/s
LR	16.43	14.25	21 598
DT	70.91	69.98	71 153
Bayes	80.95	32.29	40 426
KNN	86.03	81.27	66 311
NN	81.03	80.91	29 108
SVM	32.30	40.98	41 803

的非线性问题，使用LR回归处理起来存在一定的问题，导致过滤正确率及异常样本过滤率都很低。DT模型的思想是自顶向下采用递归的方式构造一种树结构，能够递归地选择最优特征，并且对训练集有很好的分割性，保证每个训练子集都有最好的分类过程。DT需要从根结点不断向叶子结点迭代，导致耗时最久；在过滤异常样本时，易忽略输入特征参量与风速标签之间的相关性，导致异常样本过滤率并不理想。Bayes对小样本的数据分类有较好的性能，但对训练实验中的多维大批量的数据较为吃力，使得最终的分类结果并未达到最佳效果。由于过滤输出的是风速质量标签，与训练特征集的相关性较差，导致NN模型没有达到预期的训练要求。SVM算法通过二次规划求解支持向量，求解二次规划将涉及到矩阵的相关计算，当输入特征维度较大时，将会耗费大量的空间和时间；同时，SVM主要解决二分类问题，虽然组合了多个SVM分类模型来实现多分类任务，但当任务中涉及多分类时，算法本身很难将多维参量映射到低维度的空间，导致多分类任务的实现存在一定困难。KNN算法是最简单也是最常用的分类算法之一，作为一种有监督的机器学习，其分类原理为：当预测一个新值x时，根据其距离最近的K个点中同一类别数最多的那一类来划分x，并且实验中采用交叉验证的方式选取K值，KNN具有简易用、精度高、对异常值不敏感等优点，适合本文大数据量的分类，虽然KNN计算量较大，每个待分类的样本都要计算其到全部点的距离，最终导致时间复杂度升高，但从结果的准确性上进行考虑，作为一种具有较高精度的ML分类模型，其所消耗的时间成本是值得的。总的来说，CNN-KNN模型为本文实验最优模型组合。

3.2 CNN-KNN模型性能分析

在本文实验中，将0～20 m/s的真实值风速划分为4个区间：0～5 m/s、5～10 m/s、10～15 m/s和15～20 m/s。分别对每个区间的反演性能做单独分析，如表3所示。其中，0～5 m/s风速区间内的CNN-KNN反演均方根误差为1.16 m/s，较未过滤前精度提升了37.3%；5～10 m/s风速区间内的CNN-KNN反演均方根误差为1.20 m/s，较未过滤前精度提升了3.2%；10～15 m/s风速区间内的CNN-KNN反演均方根误差为1.77 m/s，较未过滤前精度提升了4.8%；15～20 m/s风速区间内的CNN-KNN反演均方根误差为3.31 m/s，较未过滤前精度提升了2.9%。由于目前PC性能的限制，无法使用更大的数据量进行训练，导致15～20 m/s风速区间内数据量相对较少，不足以训练出适合高风速的ML组合模型，导致15～20 m/s的风速区间内反演误差较大。但从结果看，相较于单CNN模型，15～20 m/s区间内反演精度已有了进一步改善，随着PC性能的完善，这一风速区间甚至更高风速区间的反演精度将进一步提高。从反演整体效果来看，CNN-KNN模型的均方根误差满足CYGNSS风速反演的精度要求。

表3 CNN-KNN模型性能分析

Table 3 CNN-KNN model performance analysis

风速区间/(m·s^{-1})	M/(m·s^{-1})		R/(m·s^{-1})		B/(m·s^{-1})		R^2	
	CNN	CNN-KNN	CNN	CNN-KNN	CNN	CNN-KNN	CNN	CNN-KNN
0～5	1.51	1.01	1.85	1.16	0.97	0.78	0.58	0.78
5～10	1.12	0.97	1.24	1.2	0.24	0.13	0.72	0.75
10～15	1.67	1.69	1.86	1.77	1.3	1.18	0.57	0.59
15～20	2.94	3.02	3.41	3.31	2.15	1.98	0.18	0.19

为证明过滤后精度提升的同时有效数据样本量是否会明显减少的问题, 对过滤前后的风速结果做了进一步的有效数据占比分析。

总数为测试集样本量的 30 万条数据, 有效数据定义为反演风速结果与真实风速之差的绝对值小于 2 m/s 的数据。过滤前有效数据的个数为 23.86 万, 有效占比为 79.53%; 过滤后有效数据的样本个数为 22.97 万, 有效占比为 76.57%。过滤前后有效数据的绝对数量仅减少约 8 900 个, 有效数据留存率仅下降了约 2.96%, 相较于风速精度的提升, 过滤后有效数据占比只发生了较小的变化。

将异常样本数据按照真实值风速区间划分, 对每个风速区间进行分析。图 5 为异常样本在每个风速区间内的分布情况。

由图 5 可知, 在 0～10 m/s 的风速区间内, 分类模型对异常样本有较好的过滤性能, 0～10 m/s 风速区间内过滤的异常样本数量占全部过滤异常样本数量的比重达到 71.81%; 0～10 m/s 风速区间内未过滤的异常样本数量占全部未过滤异常样本数量的比重仅为 37.55%。而对于 10～20 m/s 的风速区间, 由于 ECMWF 数据集所测得的中高风速相对来说较少, 导致训练集中这部分数据占比较低, 不足以训练出一个具有鲁棒性的 KNN 分类模型, 在测试集测试分类模型过滤性能时, 对于 10～20 m/s 风速区间的异常样本过滤性能不如低风速区间效果好。

从整体过滤效果看, 使用 KNN 过滤后的模型在 *B*、*M* 及 *R* 上均优于过滤前及其他 ML 分类模型, 以此证明 CNN-KNN 模型能够较好地完成 CYGNSS 全球风速反演任务。

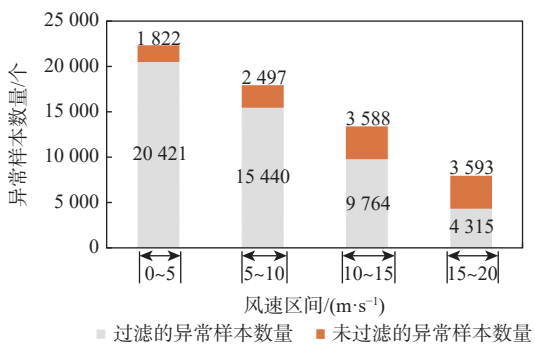


图 5　异常样本分布

Fig. 5　Distribution of abnormal samples

4　结　论

由于 CYGNSS 产品的质量控制基于各种工程、仪器和测量条件的固定阈值, 使得 CYGNSS 在 0～20 m/s 的风速区间反演风速具有一定的误差。基于此类误差, 本文提出构建一个 ML 组合模型, 利用 ML 分类模型输出 CNN 回归结果的质量控制标志, 来识别和删除 CNN 回归结果中的异常值, 削弱 CYGNSS 自身误差对风速反演结果带来的影响。

1）实验选取 2019 年 1 月 2 日、1 月 22 日、2 月 2 日、2 月 22 日、3 月 2 日、3 月 22 日、4 月 2 日的 CYGNSS L1 级数据作为训练集, 测试集选取 2019 年 5 月 2 日、5 月 22 日、6 月 2 日的 CYGNSS L1 级数据。对比过滤前后的风速反演精度可得, 本文提出的 6 种 ML 组合模型对于反演风速及误差控制方面相比于单 CNN 反演模型都具有一定的优势, 过滤后有效数据留存率仅下降了 2.96%, 而风速精度得到了 15.3% 的提升, KNN 过滤性能是其中最好的。实验结果证明了 ML 组合模型对于提高海面风速反演精度的有效性。

2）本文建立的 CNN-KNN 模型在 0～20 m/s 的风速范围内有较好的反演性能, 对于更高的风速区间, 由于数据量较少, 还不足以训练 KNN 分类模型得到较好的过滤效果。未来将会利用更好配置的 PC, 用更多的数据进行训练与验证。同时, 在保留尽可能多的样本的情况下, 将考虑提高该组合模型在更高风速下的性能的策略。随着时间的推移, 更多高风速数据的加入也有助于将该组合模型的可用性扩展到更高的风速范围。

致谢　感谢 NASA 网站提供的 CYGNSS 数据, 感谢欧洲中期天气预报中心提供的 ECMWF 数据。感谢北京航空航天大学的杨东凯教授和王峰博士在 GNSS-R 星载数据分析中提供的建议。感谢上海航天电子技术研究所的孟婉婷工程师对反射信号方面提供的建议。

参考文献（References）

[1] MASHBURN J, AXELRAD P, ZUFFADA C, et al. Improved GNSS-R ocean surface altimetry with CYGNSS in the seas of Indonesia[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 58(9): 6071-6087.

[2] HUANG F X, GARRISON J L, RODRIGUEZ-ALVAREZ N, et al. Sequential processing of GNSS-R delay-Doppler maps to estimate the ocean surface wind field[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(12): 10202-10217.

[3] ZHU Y C, TAO T Y, YU K G, et al. Sensing sea ice based on Doppler spread analysis of spaceborne GNSS-R data[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020, 13: 217-226.

[4] YIN C, LOPEZ-BAEZA E, MARTIN-NEIRA M, et al. Intercomparison of soil moisture retrieved from GNSS-R and from passive L-band radiometry at the Valencia anchor station[J]. Sensors (Basel, Switzerland), 2019, 19(8): 1900.

[5] CLARIZIA M P, RUF C S, JALES P, et al. Spaceborne GNSS-R

minimum variance wind speed estimator[J]. IEEE Transactions on Geoscience and Remote Sensing, 2014, 52(11): 6829-6843.

[6] VALENCIA E, ZAVOROTNY V U, AKOS D M, et al. Using DDM asymmetry metrics for wind direction retrieval from GPS ocean-scattered signals in airborne experiments[J]. IEEE Transactions on Geoscience and Remote Sensing, 2014, 52(7): 3924-3936.

[7] GLEASON S, RUF C, CLARIZIA M P, et al. Calibration and unwrapping of the normalized scattering cross section for the cyclone global navigation satellite system[J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(5): 2495-2509.

[8] 骆黎明, 白伟华, 孙越强. 基于树模型机器学习方法的 GNSS-R 海面风速反演[J]. 空间科学学报, 2020, 40(4): 595-601.
LUO L M, BAI W H, SUN Y Q. GNSS-R sea surface wind speed retrieval based on tree model machine learning method[J]. Journal of Space Science, 2020, 40(4): 595-601(in Chinese).

[9] CARDELLACH E, NAN Y, LI W. Variational retrievals of high winds using uncalibrated CyGNSS observables[J]. Remote Sensing, 2020, 12(23): 3930.

[10] SAÏD F, JELENAK Z, PARK J, et al. The NOAA track-wise wind retrieval algorithm and product assessment for CyGNSS[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 60: 1-24.

[11] GLEASON S. Level 1B DDM calibration : 148-0137[R]. Washington, D. C. : NASA, 2020.

[12] LI X, MECIKALSKI J R, LANG T J. A study on assimilation of CYGNSS wind speed data for tropical convection during 2018 January MJO[J]. Remote Sensing, 2020, 12(8): 1243.

[13] ZHANG Y, YIN J, YANG S, et al. High wind speed retrieval model of CYGNSS sea surface data based on machine learning[J]. Remote Sensing, 2021, 13(16): 3324.

[14] RUF C, CHANG P, CLARIZIA M P, et al. CYGNSS handbook[M]. Ann Arbor: Michigan Publishing, 2016.

[15] HERSBACH H, BELL B, BERRISFORD P, et al. ERA5 hourly data on single levels from 1940 to present[EB/OL]. (2018-10-23) [2022-04-01]. https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=form.

[16] ROSKA T, CHUA L O. The CNN universal machine: An analogic array computer[J]. IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing, 2015, 40(3): 163-173.

[17] 徐嘉兴, 李钢, 陈国良. 基于 logistic 回归模型的矿区土地利用演变驱动力分析[J]. 农业工程学报, 2012, 28(20): 247-255.
XU J X, LI G, CHEN G L. Analysis of driving force of land use evolution in mining area based on logistic regression model[J]. Chinese Journal of Agricultural Engineering, 2012, 28(20): 247-255 (in Chinese).

[18] 张增伟, 吴萍. 基于朴素贝叶斯算法的改进遗传算法分类研究[J]. 计算机工程与设计, 2012, 33(2): 750-753.
ZHANG Z W, WU P. Research on classification of improved genetic algorithm based on naive Bayes algorithm[J]. Computer Engineering and Design, 2012, 33(2): 750-753(in Chinese).

[19] ALIMJAN G, SUN T L, LIANG Y, et al. A new technique for remote sensing image classification based on combinatorial algorithm of SVM and KNN[J]. International Journal of Pattern Recognition and Artificial Intelligence, 2018, 32(7): 1859012.

[20] BOLLWEIN F, WESTPHAL S. A branch & bound algorithm to determine optimal bivariate splits for oblique decision tree induction[J]. Applied Intelligence, 2021, 51: 7552-7572.

[21] ALIMJAN G, SUN T L, JUMAHUN H, et al. A hybrid classification approach based on support vector machine and K-nearest neighbor for remote sensing data[J]. International Journal of Pattern Recognition and Artificial Intelligence, 2017, 31(10): 1750034.

[22] BALASUBRAMANIAM R, RUF C. Neural network based quality control of CYGNSS wind retrieval[J]. Remote Sensing, 2020, 12(17): 2859.

[23] SUN X K, LIU L, LI C F, et al. Classification for remote sensing data with improved CNN-SVM method[J]. IEEE Access, 2019, 7: 164507-164516.

[24] RUF C, GLEASON S, MCKAGUE D. Assessment of CYGNSS wind speed retrieval uncertainty[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2019, 12(1): 87-97.

Quality control model of CYGNSS sea surface wind speed retrieval based on ML combination

ZHANG Yun[1, 2], ZHAO Xingyu[1, 2], YANG Shuhu[1, 2, *], SUN Cong[3], HAN Yanling[1, 2], YIN Jiwei[1, 2]

(1. College of Information Technology, Shanghai Ocean University, Shanghai 201306, China;
2. Shanghai Marine Intelligent Information and Navigation Remote Sensing Engineering Technology Research Center, Shanghai Ocean University, Shanghai 201306, China;
3. Shanghai Aerospace Technology Co., Ltd., Shanghai 201109, China)

Abstract: Convolutional neural networks (CNN) can be used for sea surface wind speed retrieval of cyclone global navigation satellite system (CYGNSS). There are still anomalous values in the observation data of CYGNSS, despite the fact that numerous quality control indicators have been set up to detect and weaken the abnormal observation data of CYGNSS before model training, which results in a drop in model retrieval accuracy and even incorrect retrieval results. Therefore, this paper proposes a wind speed retrieval model based on machine learning (ML) combination. Based on the CYGNSS retrieval of sea surface wind speed based on the CNN regression model, the ML classification model generates the quality flag of the CNN regression result, which can detect and remove the outliers of the CNN regression results to further improve the data quality of the wind speed retrieval results, and the ML classification model can better consider the interaction between various data errors, instead of using the threshold for each condition individually, to achieve better results. The effect of retrieval accuracy of sea surface wind speed. Six classification models were compared in the experiments, including Logistic regression (LR), decision tree (DT), naive Bayes model, K-nearest neighbor (KNN) algorithm, neural network (NN) model, and support vector machine (SVM). It was ultimately determined that the classification model based on KNN algorithm had the best impact on the quality control of wind speed retrieval. The wind speed retrieval combined model significantly improves the accuracy of the retrieval results. In the range of 0−20 m/s, the filtering rate of abnormal samples is 81.27%, and in all filtered data, the filtering correct rate is 86.03%; the root mean square error of the error is reduced from 1.7 m/s for the classification model without ML to 1.44 m/s for the classification model with ML. Among them, the training sample is 0−10 m/s retrieval results, and the accuracy improvement effect is more obvious, which proves the effectiveness of the ML combination model proposed in this paper for wind speed quality control.

Keywords: cyclone global navigation satellite system; wind speed retrieval; quality control; machine learning combination model; convolutional neural networks; K-nearest neighbor algorithm

Received: 2022-04-06; **Accepted**: 2022-06-11; **Published Online**: 2022-07-27 17:10

URL: link.cnki.net/urlid/11.2625.V.20220727.1416.001

Foundation items: National Natural Science Foundation of China (41871325,42176175); National Key R & D Program of China (2019YFD0900805)

* **Corresponding author**. E-mail: shyang@shou.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0253

改进迁移学习的双分支卷积神经网络图像去雾

李云红*，于惠康，马登飞，苏雪平，段姣姣，史舍驰
(西安工程大学 电子信息学院，西安 710048)

摘 要：针对现有图像去雾算法存在去雾不彻底和图像颜色失真的问题，提出一种迁移学习子网络和残差注意力子网络相结合的图像去雾模型。采用迁移学习子网络的预训练模型增强样本的特征属性；构建双分支网络结构，并利用残差注意力子网络辅助迁移学习子网络训练网络模型的参数；利用尾部集成学习的方法融合双网络的特征，得到去雾图像的模型参数，完成图像恢复任务。实验结果表明：所提算法在 RESIDE 数据集和 O-HAZE 数据集上 PSNR 指标比 GCANet 分别提高了 1.87 dB 和 4.22 dB，在 O-HAZE 数据集上 SSIM 指标比 GCANet 提高了 6.7%。

关 键 词：图像去雾；迁移学习；卷积神经网络；注意力机制；集成学习
中图分类号：TP391.41
文献标志码：A **文章编号**：1001-5965(2024)01-0030-09

计算机视觉系统在人工智能领域中已经得到广泛应用，在智慧城市、军事探测、气象检测等领域中，图像去雾技术得到了快速发展。在雾霾、烟尘、水汽等恶劣天气情况下，由于大气中悬浮颗粒物的散射作用，导致地面水平能见度降低，使系统采集的图像出现模糊、颜色失真、锐度降低等问题，极大影响了人们的户外生活。

雾图清晰化是不适定问题[1]。由于自然场景下雾的分布不均匀、雾浓度级别不同，图像去雾研究仍然是很重要的挑战，也引起了研究者的普遍重视。目前，图像去雾算法主要分为两大类，分别为基于大气散射模型的方法和基于深度学习的方法。He 等[2]统计了多幅清晰图像中局部像素区域的颜色通道趋近于一个很小的值，并利用暗通道先验 (dark channel prior, DCP) 知识估计透射率，但对于天空、白色块等高亮区域会产生颜色失真，这种基于暗通道的方法去雾效果往往不理想。Zhu 等[3]利用雾浓度和像素饱和度的关系建立了颜色先验模型 (color attenuation prior, CAP)，并采用有监督的学习方法确定相关系数，利用景深估计透射图恢复雾图。Tan[4]利用马尔可夫模型建立了代价方程，采用图像分割技术进行优化，利用该先验条件估计大气光，增强图像对比度，生成清晰的图像。杨燕等[5]重构最小通道区间估计并结合自适应约束模型优化透射率，利用物理模型完成图像去雾任务。随着人工智能的兴起，研究人员训练卷积神经网络 (convolutional neural network, CNN) 参数，并利用网络测试来完成图像处理任务。基于深度学习的方法主要分为 2 种：①通过训练网络间接估计大气光和透射率，利用大气散射模型恢复图像；②采用端到端的卷积神经网络学习雾图和清晰图像的空间映射关系。Cai 等[6]首次将深度卷积去雾网络 DehazeNet 用来学习雾图的透射率，并使用双边线性函数提高模型收敛速度，依据大气散射模型恢复雾图。Ren 等[7]引出局部多补丁网络的结构，融合雾图不同尺度的特征，并采取跳跃连接的方式聚合上下文

收稿日期：2022-04-18；**录用日期**：2022-07-08；**网络出版时间**：2022-08-26 12:57
网络出版地址：link.cnki.net/urlid/11.2625.V.20220826.1036.002
基金项目：国家自然科学基金 (61902301)；陕西省科技厅自然科学基础研究计划重点项目 (2022JZ-35)；陕西省教育厅自然科学基础研究计划 (19JK0364)；国家级大学生创新创业训练计划 (202210709012)；陕西高校青年创新团队
* **通信作者**. E-mail：hitliyunhong@163.com

引用格式：李云红，于惠康，马登飞，等. 改进迁移学习的双分支卷积神经网络图像去雾 [J]. 北京航空航天大学学报，2024，50（1）：30-38.
LI Y H，YU H K，MA D F，et al. Improved transfer learning based dual-branch convolutional neural network image dehazing[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50（1）：30-38 （in Chinese）.

特征, 在合成雾图和真实雾图都取得了良好的效果。Li 等[8]将物理模型中大气光和透射率 2 个参数进行统一, 利用一体化去雾网络 (all-in-one dehazing network, AODNet) 学习新变量 $K(x)$ 和清晰图像之间的映射。间接的学习方法尽管在恢复任务中取得了不错的效果, 但会过度依赖大气散射模型, 对于浓雾图像和不均匀分布雾图往往效果不佳。Liu 等[9]将语义分割网络应用于图像去雾, 设计了一种网格去雾算法用于图像恢复任务, 通过一种新的多尺度特征提取和通道注意力机制有效缓解传统多尺度方法遇到瓶颈的问题。Chen 等[10]提出了一种能够聚合图像语义信息的门控聚合网络 (gated context aggregation network, GCANet), 利用平滑扩张卷积模块提高图像的感受野, 解决了网络伪影的问题。王殿伟等[11]利用图像频率约束循环生成对抗网络, 有效降低了雾对生成质量的影响。虽然直接去雾的方法比大气散射模型具有更好的鲁棒性, 然而在真实有雾的训练过程中会丢失高频细节信息, 忽略雾图中有用的信息。

目前, 大多数先进的算法(如 DCP、DehazeNet、GCANet 等)在合成场景的雾图上有着良好的表现, 但对于户外真实分布不均匀的雾图存在雾残留、颜色偏移等问题。本文针对上述问题, 提出了一种双分支迁移学习的卷积神经网络, 能够直接学习雾图和清晰图像的映射关系, 并且弥补了单分支卷积神经网络的图像去雾精度不高的缺陷, 有效解决了非均匀雾图场景下存在的雾残留问题。

1　双分支卷积神经网络模型

图 1 为基于迁移学习和残差注意力机制的双分支网络结构, 主要包含迁移学习子网络和残差注意力子网络 2 个子网络。其中, 迁移学习子网络在先前任务学习的基础上帮助当前任务, 从输入的雾图提取全局特征并加以表示, 还能够有效解决过拟合和数据不足的问题。残差注意力子网络利用线性叠加的思想, 避免梯度消失和梯度爆炸, 在图像训练任务中表现良好, 最终利用集成学习的方式融合不同图像特征子网络的特征, 并采用一个镜像填充层、一个卷积层和一个激活函数 Tanh 激活函数层输出清晰图像。

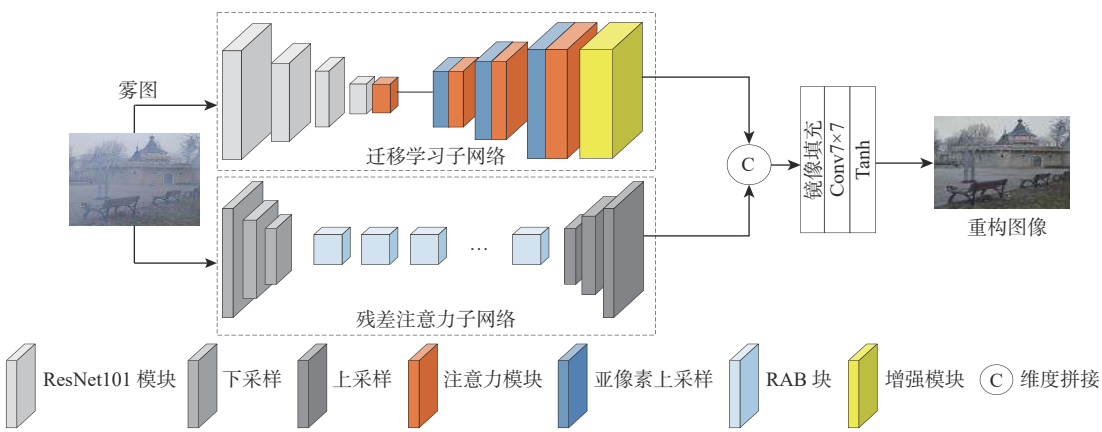


图 1　双分支网络结构

Fig. 1　Dual-branch network structure

1.1　迁移学习子网络

迁移学习子网络主要包括编码器和解码器 2 个部分。编码器采用 He 等提出的 ResNet[12], 其在 ImageNet 数据集分类任务中表现出色。本文选取 ResNet101 的 32 倍降采样部分作为编码器的结构, 能够有效提取图像特征, 在训练时加载 ImageNet 预训练模型, 相对参数随机初始化的模型训练, 预训练编码器一定程度上提高了模型精度。

在解码器中加入注意力模块, 由于编码器使用了下采样操作, 对获取到的特征造成一定程度的损失。因此, 在编码器与解码器之间采用跳跃连接的方式后进行上采样,有助于解码器获取更多的细节特征。迁移学习子网络尾部加入增强模块, 加强了模型对图像特征的学习。

1.1.1　注意力模块

注意力机制采用人类视觉注意力方式[13], 传统的注意力机制从大量信息中筛选出当前任务的重要特征, 但会丢失一些重要细节特征, 对于非均匀分布的雾图, 会降低去雾性能。因此, 本文提出了一种局部双分支结构的注意力模块, 如图 2 所示。

本文提出的注意力模块利用全局平均池化操作和全局最大池化操作, 压缩通道信息并关注不同通道的加权信息, 对特征图的大小压缩并去除冗余信息, 有效简化模型复杂度。首先, 利用全局平均池化和全局最大池化将特征图获取不同的通道描述符。全局平均池化操作和全局最大池化操作表

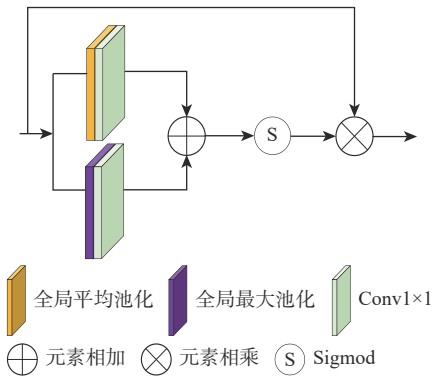


图2 注意力模块

Fig. 2 Attention block

示如下：

$$g_{ac}=H_{a}(\boldsymbol{F}_c)=\frac{1}{H\times W}\sum_{i=1}^{H}\sum_{j=1}^{W}X_c(i,j) \tag{1}$$

$$g_{mc}=H_{m}(\boldsymbol{F}_c)=\max(X_c(i,j)) \tag{2}$$

式中：$\boldsymbol{F}_c$为输入的特征图；$X_c(i,j)$为第c个通道X_c在(i,j)位置的值；H_a为平均函数；H_m为最大值函数。特征图的形状从$H\times W$变为1×1，分别经过全局平均池化和全局最大池化后进行1×1的卷积操作，再进行元素相加和Sigmod激活函数得到通道权值。输出特征图表示为

$$\boldsymbol{A}_c=\sigma(\mathrm{Conv}(\boldsymbol{g}_{ac})\oplus\mathrm{Conv}(\boldsymbol{g}_{mc})) \tag{3}$$

式中：σ表示Sigmod函数；Conv表示卷积操作；“$\oplus$”表示元素相加。

最后，将输入的特征图$\boldsymbol{F}_c$按照元素顺序与权值$\boldsymbol{A}_c$进行元素相乘得到输出结果$\boldsymbol{F}_c^*$，表示为

$$\boldsymbol{F}_c^*=\boldsymbol{F}_c\otimes\boldsymbol{A}_c \tag{4}$$

1.1.2 增强模块

迁移学习子网络的尾部加入了特征增强模块，如图3所示。该模块是一种金字塔池化卷积模块，特征图通过2个卷积操作后进入平均池化层下采样，输出1/4、1/8、1/16、1/32等4个不同尺度的特征图，再分别进行1×1卷积操作后进行上采样。不同尺度的特征会提供不同的感受野，有利于帮助图像恢复和重建，并且避免网络伪影，不同尺度的平均池化操作保证了网络的宽度，同时利用1×1卷积核有效减少了模型参数量。最终将不同尺度特征图上采样得到的特征图与开端卷积层的输出特征拼接起来，再通过3×3的卷积操作得到输出特征图。

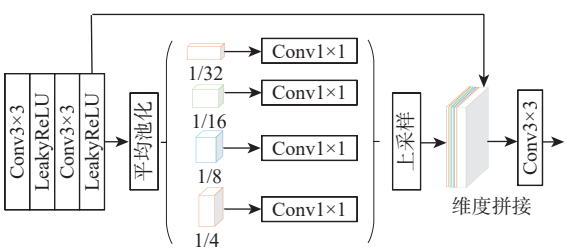


图3 增强模块

Fig. 3 Enhance block

1.2 残差注意力子网络

在迁移学习子网络中，编码器到解码器的全分辨率跳跃连接方式会导致网络丢失一些对恢复图像细节很重要的信息。因此，本文设计残差注意力子网络来提取精细的特征，以补充迁移学习网络的不足。如图4所示，残差注意力子网络主要由编码器、若干个残差注意力块(residual attention block, RAB)和解码器组成。为减少模型参数量，编码器对输入图像进行4倍下采样的卷积操作，主干网络由20个RAB块组成，解码器对特征图进行4倍上采样的反卷积操作来重建图像。

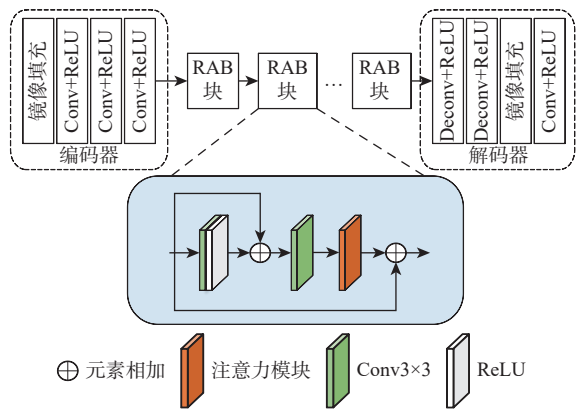


图4 残差注意力子网络

Fig. 4 Residual attention sub-net

采用RAB块可有效避免梯度消失和梯度爆炸的问题，提高当前模型提取更具体的特征，并且增强模型对当前数据的拟合能力。具有强大的拟合能力的模型易导致过拟合问题，因此，迁移学习分支的作用不可忽视。在一个RAB块中，主要包括2个卷积层、1个修正单元层(ReLU)和1个注意力模块。输入的特征图与特征图经过卷积和激活函数后进行第1次元素相加，再经过卷积层和注意力模块后与输入的特征图进行第2次元素相加，利用局部残差连接可以达到特征的重复叠加的效果，实现特征的重复利用。

1.3 参数评估

本文采用3种不同的损失函数用于规范网络训练中参数的学习方向，分别为平滑L_1损失、感知损失和多尺度结构相似损失。平滑L_1损失作为主要的规范函数，用于修正图像之间像素的差异，感知损失和多尺度结构相似损失结合用于规范人类视觉感知效果。

1）平滑L_1损失。平滑L_1损失克服了L_1损失

不稳定的缺点，使模型变得更加鲁棒，相对于 L_2 损失，对异常点、离群点不敏感，在许多图像恢复任务中表现出色。平滑 L_1 损失函数表示为

$$L_{\mathrm{s}}=\frac{1}{N}\sum_{i=1}^{N}S_{\mathrm{L1}}(\boldsymbol{y}_i-F(\boldsymbol{x}_i)) \tag{5}$$

$$S_{\mathrm{L1}}(x)=\begin{cases}0.5x^2 & |x|<1\\ |x|-0.5 & \text{其他}\end{cases} \tag{6}$$

式中：$\boldsymbol{x}_i$ 为输入雾图；$\boldsymbol{y}_i$ 为真值；$F()$ 表示网络模型；N 为像素的总数。

2）感知损失。在图像恢复任务中，感知损失[14]被广泛用于 2 幅图像之间的知觉差异，并且能够获得更加丰富的细节信息。本文利用感知损失获取去雾图像与清晰图像之间的高级感知信息来优化网络参数。感知损失函数表示为

$$L_{\mathrm{p}}=\sum_{i}\frac{1}{C_i\times H_i\times W_i}\|\boldsymbol{\varphi}_i(\boldsymbol{J})-\boldsymbol{\varphi}_i(F(\boldsymbol{x}))\|_2^2 \tag{7}$$

式中：$\boldsymbol{\varphi}_i$ 为 VGG-16 网络第 i 层的特征图，尺寸大小 $C_i\times H_i\times W_i$；$\boldsymbol{J}$ 为清晰图像；$F(\boldsymbol{x})$ 为去雾图像；设置 i=3, 8, 15。

3）多尺度结构相似损失。与感知损失不同，为使图像获得更高的视觉效果，并且保留高频信息，本文采用多尺度结构相似损失结合感知损失应用于模型训练。多尺度结构相似损失表示为

$$L_{\mathrm{m}}=1-\prod_{m=1}^{M}\left(\frac{2u_{\mathrm{p}}u_{\mathrm{g}}+C_1}{u_{\mathrm{p}}^2+u_{\mathrm{g}}^2+C_1}\right)^{\alpha_m}\left(\frac{2\sigma_{\mathrm{pg}}+C_2}{\sigma_{\mathrm{p}}^2+\sigma_{\mathrm{g}}^2+C_2}\right)^{\gamma_m} \tag{8}$$

式中：下标 p 和 g 分别表示去雾图像和清晰图像；u_{p} 和 σ_{p} 分别为去雾图像的均值和方差；u_{g} 和 σ_{g} 分别为清晰图像的均值和方差；σ_{pg} 为协方差；C_1 和 C_2 为变量；α_m 和 γ_m 为默认参数。

总损失函数定义为

$$L=\lambda_{\mathrm{s}}L_{\mathrm{s}}+\lambda_{\mathrm{p}}L_{\mathrm{p}}+\lambda_{\mathrm{m}}L_{\mathrm{m}} \tag{9}$$

式中：λ 为用于平衡网络损失函数的权重，利用总损失函数更新并优化网络参数，根据训练经验将 λ_{s}、λ_{p}、λ_{m} 分别设置为 1、0.01、0.5。

2　实验结果与分析

本文提出的迁移学习和残差注意力机制的双分支卷积神经网络在 RESIDE 合成数据集和 O-HAZE 真实数据集上分别进行了客观指标的评价和主观视觉效果的分析。选取图像峰值信噪比（peak signal to noise ratio, PSNR）和结构相似性（structural similarity index, SSIM）作为客观评价指标。

2.1　数据集

实验选取合成数据集和真实数据集来评估，合成数据集选取 RESIDE 数据集[15]中的室内训练集 ITS 和综合目标测试集 SOTS，ITS 训练集包含 13 990 张合成雾图和 1 399 张清晰图像，ITS 训练集主要用于训练，SOTS 数据集中包含 500 张室内雾图，实验选用 SOTS 室内测试集进行测试。由于雾天真实数据集获取难度大、成本高，其具有数量少和雾分布不均匀等特点。真实数据集采用 NTIRE2018 中的 O-HAZE 数据集[16]，包括 45 对雾图和清晰图像。在 RESIDE 数据集和 O-HAZE 数据集上，分别与 DCP、CAP、DehazeNet、AODNet 和 GCANet 进行对比实验。

2.2　实验条件

实验环境为：Ubuntu18.04 LTS 64-bit 操作系统，内存 16 GB，显卡为 NVIDIA RTX 3060。本文算法采用 PyTorch 深度学习框架实现，利用随机裁剪、上下翻转和水平翻转的方法扩充数据集，输入的图像被随机裁剪为 256×256 的区域。本文使用 ADAM 优化器，将 β_1 和 β_2 分别设置为 0.9 和 0.999。初始学习率为 0.000 1，批次大小设置为 8。在 RESIDE 数据集上训练时，epoch 设置为 100，采用动态衰减策略，在第 30 epoch、60 epoch 和 80 epoch 衰减 0.5。在 O-HAZE 数据集上训练时，epoch 设置为 1 000，在第 100 epoch、200 epoch、300 epoch 和 700 epoch 衰减 0.5。

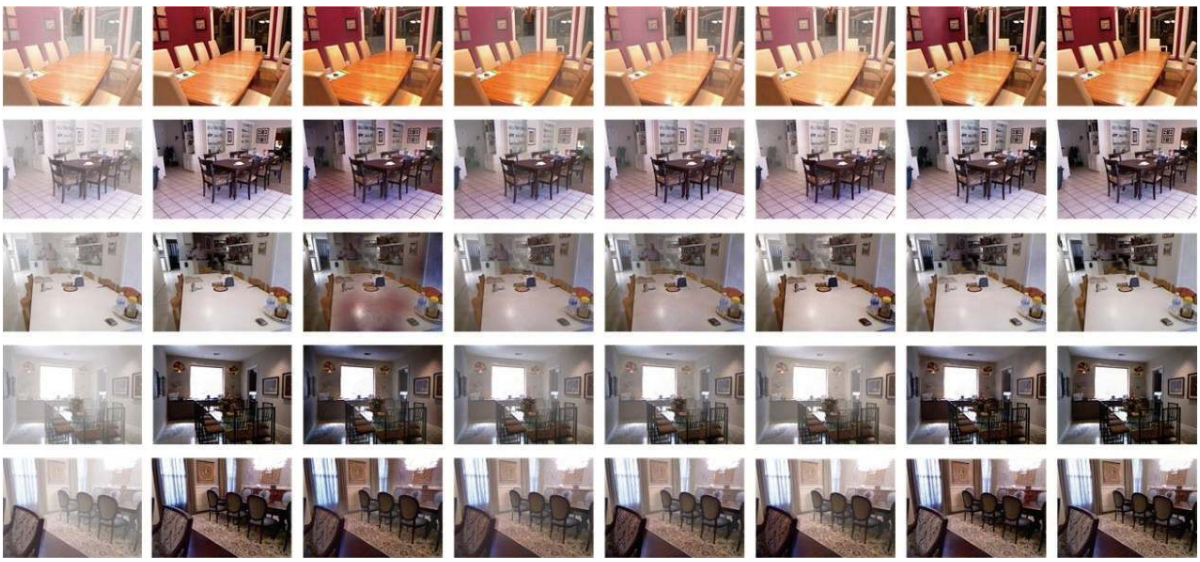
2.3　RESIDE 数据集的实验结果

图 5 为本文算法与现有算法（DCP、CAP、DehazeNet、AODNet、GCANet）在 SOTS 数据集上去雾效果对比。其中，DCP 算法去雾图像整体颜色偏暗，对于白色高亮区域会产生颜色失真（见图 5(c)中第 2、3 幅图）。CAP 算法在远景处有明显的雾残留（见图 5(d)中第 1、4、5 幅图），DehazeNet 在白色高亮区域存在轻微颜色失真，并且图像去雾不彻底（见图 5(e)中第 1、3 幅图）。AOD 算法存在的问题与 DehazeNet 算法类似，去雾后的图像会丢失一些细节信息并存在雾残留的问题。虽然 GCANet 去雾效果很好，但存在微量雾残留（见图 5(g)第 1 幅图中墙面泛白）和图像失真（见图 5(g)第 3 幅图中人物）的问题。综上，本文算法去雾效果视觉对比度高，去雾后的图像颜色更加自然，纹理细节更加显著。

为进一步验证本文算法的去雾效果，从客观角度来评价算法的性能，采取 PSNR 和 SSIM 指标对比其他 5 种算法。6 种算法在 SOTS 数据集上的 SSIM、PSNR 值如表 1 所示。可以看出，传统算法 DCP 和 CAP 指标较低，而深度学习 DehazeNet、AODNet 和 GCANet 网络的评价指标有所改善，本文算法在 PSNR 指标上比 GCANet 高出 1.87 dB。

2.4　O-HAZE 数据集的实验结果

相对于合成数据集，户外真实场景下采集有雾



(a) 雾图 (b) 清晰图像 (c) DCP (d) CAP (e) DehazeNet (f) AODNet (g) GCANet (h) 本文算法

图5 不同去雾算法在SOTS数据集上的去雾效果

Fig. 5 Dehazing effect of different dehazing algorithms on SOTS dataset

表1 不同去雾算法在SOTS数据集上的SSIM、PSNR值

Table 1 SSIM and PSNR values of SOTS dataset under different dehazing algorithms

算法	PSNR/dB	SSIM
DCP	16.62	0.817
CAP	19.05	0.840
DehazeNet	20.64	0.800
AODNet	19.06	0.850
GCANet	30.23	0.975
本文算法	32.10	0.94

图和无雾图难度大，采集的数据量少，因此，真实数据集在训练和测试时往往效果不理想。对于O-HAZE数据集，选取35对图像用于训练，5对图像用于验证，5对图像用于测试。在主观和客观指标上，与现有算法进行对比。图6为O-HAZE数据集去雾效果对比。从主观视觉上来看，由于DCP算法是基于暗通道先验的算法，去雾效果受到环境因素的影响，DCP在非均匀有雾的场景下产生了严重的颜色偏移现象，去雾效果不太理想。CAP和DehazeNet都存在大量雾残留的问题。AODNet虽然一定程度上有去雾的效果，但去雾后的图像发生了色彩偏移。GCANet网络对图像去雾后依然存在少量雾，图6(g)第1、2、5幅图远处存在少量雾。本文算法融入了迁移学习的分支，并导入预训练的模型参数，加快了模型的收敛速度，并使用集成学习融合双分支特征。由图6可以看出，本文算法去雾效果明显，去雾后的图像在颜色和物体细节方面更加接近真实清晰图像，没有出现大量雾残留的区域。

为进一步验证本文算法的去雾性能，6种算法



(a) 雾图 (b) 清晰图像 (c) DCP (d) CAP (e) DehazeNet (f) AODNet (g) GCANet (h) 本文算法

图6 不同去雾算法在O-HAZE数据集上的去雾效果

Fig. 6 Dehazing effect of different dehazing algorithms on O-HAZE dataset

在 O-HAZE 数据集上的 SSIM、PSNR 指标对比如表 2 所示。可以看出，DCP 和 CAP 传统算法的 PSNR 和 SSIM 的值相对较低，图像恢复质量不高，本文算法客观指标有所提升，表明图像恢复效果更好。本文算法的峰值信噪较 GCANet 高出 4.22 dB，结构相似性高出 0.067。

表 2　不同去雾算法在 O-HAZE 数据集上的 SSIM、PSNR 值

Table 2　SSIM and PSNR values of O-HAZE dataset under different dehazing algorithms

算法	PSNR/dB	SSIM
DCP	12.92	0.505
CAP	14.55	0.567
DehazeNet	16.27	0.604
AODNet	17.69	0.616
GCANet	18.41	0.680
本文算法	22.63	0.747

3　算法对比实验

3.1　模型结构对比分析

为验证网络分支和迁移学习的有效性，选取 O-HAZE 数据集训练 500 次进行对比实验，从模型结构角度进行对比分析，本文算法与使用预训练迁移学习子网络（TLN）、参数随机初始化的迁移学习子网络、残差注意力子网络（RAN）、参数随机初始化的迁移学习和残差注意力机制的双分支卷积神经网络（TLN+RAN）进行实验对比。图 7 为不同的去雾算法的对比。其中，NP 表示不使用 ImageNet 预训练。从图 7（c）、（d）可以看出，未使用 ImageNet 预训练参数恢复的图像颜色较深，且色彩饱和度不高。从图 7（g）可以看出，本文算法通过引入双分支网络并采用 ImageNet 预训练参数，获得更加丰富的去雾效果，图像恢复得更加自然。

为进一步说明迁移学习和多分支网络的有效性，表 3 为 PSNR 和 SSIM 指标的对比。TLN 网络使用 ImageNet 预训练参数明显比参数随机初始化在 PSNR 和 SSIM 指标上有不同程度的提高。对比单分支网络的结构也可以发现，双分支模型去雾性能表现的更好，表明了双分支网络结构的有效性。本文算法在 PSNR 和 SSIM 指标上分别达到了 21.94 dB 和 0.736。



(a) 雾图　(b) 清晰图像　(c) TLN(NP)　(d) TLN　(e) RAN　(f) TLN(NP)+RAN　(g) 本文算法

图 7　不同去雾算法的去雾效果

Fig. 7　Dehazing effect of different dehazing algorithms

表 3　不同去雾算法对性能的影响

Table 3　Effects of different dehazing algorithms on performance

算法	PSNR/dB	SSIM
TLN(NP)	19.87	0.661
TLN	20.56	0.706
RAN	17.34	0.589
TLN(NP)+RAN	20.92	0.716
本文	21.94	0.736

3.2　不同模型设置对网络性能的影响分析

为说明本文所使用模块的有效性，实验在 O-HAZE 数据集上分为 4 个模型（Model 1、Model 2、Model 3 和 Model 4）进行分析。其中，Model 1 指双子网不使用注意力模块和增强模块，Model 2 指双子网只使用注意力模块，Model 3 指双子网只使用增强模块，Model 4 为本文算法。表 4 为模型的客观指标。可以看出，使用注意力模块或增强模块能有效提高 SSIM，结合两者性能最优。

区域去雾效果如图 8 所示。Model 1 对于远景

表 4 不同模型对性能的影响

Table 4 Effects of different models on performance

模型	PSNR/dB	SSIM
Model 1	21.36	0.612
Model 2	21.50	0.700
Model 3	21.52	0.650
Model 4	22.63	0.747

区域存在模糊的问题，Model 2 和 Model 3 尽管有一定改善，但恢复后的图像质量不高，结合注意力模块和增强模块会更加关注图像远处有雾区域，去雾后的图像质量更高。

3.3 损失函数参数设置

设置不同的参数（λ_s、λ_p、λ_m）进行实验结果对比，由于平滑 L_1 损失作为最主要的范数，实验中将 λ_s 设置为 1，λ_p 设置为 0.05 和 0.01，λ_m 设置为 0.5 和 0.8，总损失函数采用不同的加权系数在 O-HAZE 数据集上的客观指标如表 5 所示。

由表 5 可得，引入 3 种不同的损失函数参数，相较于 1 种或 2 种参数在 PSNR 和 SSIM 性能上都有不同程度的提升，表明 3 种损失函数的结合对提升网络性能起到了明显的作用。图 9 为 Method 1、Method 3、Method 4、Method 8 和 Method 9 训练过程曲线对比结果。由此可以看出，Method 1 与 Method 3 相比，感知损失训练时能够产生更高的 PSNR 和 SSIM，Method 1 与 Method 4 相比，引入多尺度结构相似损失尽管会降低 PSNR，但是有助于改善 SSIM

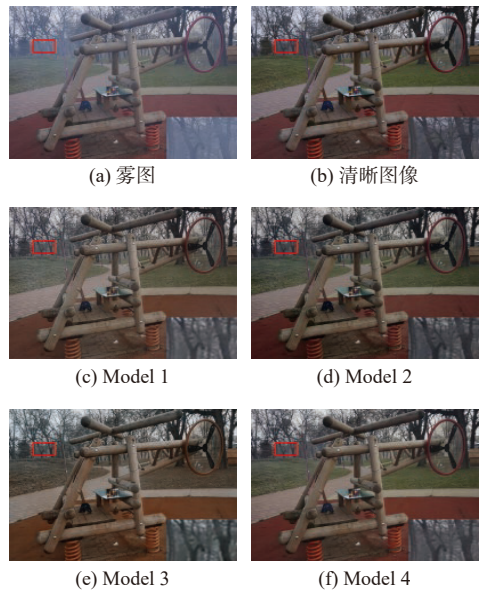


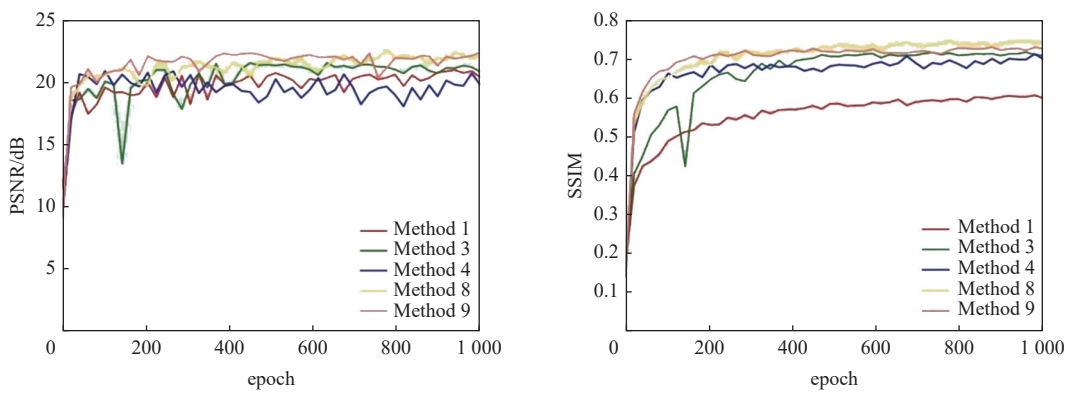
图 8 不同模型的区域去雾效果

Fig. 8 Regional dehazing effect of different models

表 5 不同加权系数对性能的影响

Table 5 Effect of different weighting coefficients on performance

方法	λ_s	λ_p	λ_m	损失函数	PRNR/dB	SSIM
Method 1	1	0	0	$L=L_s$	20.90	0.607
Method 2	1	0.05	0	$L=L_s+0.05L_p$	21.06	0.703
Method 3	1	0.01	0	$L=L_s+0.01L_p$	21.33	0.717
Method 4	1	0	0.5	$L=L_s+0.5L_m$	20.81	0.713
Method 5	1	0	0.8	$L=L_s+0.8L_m$	20.75	0.727
Method 6	1	0.05	0.5	$L=L_s+0.05L_p+0.5L_m$	21.35	0.733
Method 7	1	0.05	0.8	$L=L_s+0.05L_p+0.8L_m$	21.20	0.730
Method 8	1	0.01	0.5	$L=L_s+0.01L_p+0.5L_m$	22.63	0.747
Method 9	1	0.01	0.8	$L=L_s+0.01L_p+0.8L_m$	22.38	0.730



(a) PSNR 训练曲线 (b) SSIM 训练曲线

图 9 O-HAZE 数据集的训练曲线对比

Fig. 9 Comparison of training curves for O-HAZE dataset

指标, 当$\lambda_s = 1$, $\lambda_p = 0.01$, $\lambda_m = 0.5$时, 模型学习能力更强, PSNR 和 SSIM 指标达到了最优值。

4 结　论

1) 本文算法可实现较为优秀的去雾性能, 在 RESIDE 数据集上测试 500 张图像, 平均 PSNR 和 SSIM 指标分别能够达到 32.10 dB 和 0.94, 在 O-HAZE 数据集上平均 PSNR 和 SSIM 指标分别能够达到 22.63 dB 和 0.747。

2) 本文算法提出的 ResNet101 和残差网络的双分支结构搭配注意力机制和增强模块能够达到细节特征提取。

3) 对于环境较为复杂的场景, 本文算法可能受到光照、颜色、物体数量的影响, 需要进一步数据扩充或完善模型结构来处理这些场景。

为使本文算法能处理多个场景的雾图, 并应用到视频领域当中, 需要对模型性能进一步压缩并优化模型结构。

参考文献 (References)

[1] 禹晶, 徐东彬, 廖庆敏. 图像去雾技术研究进展[J]. 中国图象图形学报, 2011, 16(9): 1561-1576.
YU J, XU D B, LIAO Q M. Imagedefogging: A survey[J]. Journal of Image and Graphics, 2011, 16(9): 1561-1576(in Chinese).

[2] HE K, SUN J, TANG X. Single image haze removal using dark channel prior[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 33(12): 2341-2353.

[3] ZHU Q, MAI J, SHAO L. A fast single image haze removal algorithm using color attenuation prior[J]. IEEE Transactions on Image Processing, 2015, 24(11): 3522-3533.

[4] TAN R T. Visibility in bad weather from a single image[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2008: 1-8.

[5] 杨燕, 张金龙, 张浩文. 基于区间估计与透射率自适应约束的去雾算法[J]. 北京航空航天大学学报, 2022, 48(1): 15-26.
YANG Y, ZHANG J L, ZHANG H W. Dehazing algorithm based on interval estimation and adaptive constraints of transmittance[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(1): 15-26(in Chinese).

[6] CAI B, XU X, JIA K, et al. DehazeNet: An end-to-end system for single image haze removal[J]. IEEE Transactions on Image Processing, 2016, 25(11): 5187-5198.

[7] REN W, LIU S, ZHANG H, et al. Single image dehazing via multi-scale convolutional neural networks[C]//Proceedings of the European conference on Computer Vision. Berlin: Springer, 2016: 154-169.

[8] LI B, PENG X, WANG Z, et al. AOD-Net: All-in-one dehazing network[C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2017: 4770-4778.

[9] LIU X, MA Y, SHI Z, et al. GridDehazeNet: Attention-based multi-scale network for image dehazing[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2019: 7314-7323.

[10] CHEN D, HE M, FAN Q, et al. Gated context aggregation network for image dehazing and deraining[C]//Proceedings of the IEEE Winter Conference on Applications of Computer Vision. Piscataway: IEEE Press, 2019: 1375-1383.

[11] 王殿伟, 李顺利, 韩鹏飞. 基于特征约束 CycleGAN 的单幅图像去雾算法研究[J]. 激光与光电子学进展, 2021, 58(14): 272-278.
WANG D W, LI S L, HANG P F. Feature constraint CycleGAN for single image dehazing algorithm[J]. Laser & Optoelectronics Progress, 2021, 58(14): 272-278(in Chinese).

[12] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 770-778.

[13] NIU Z, ZHONG G, YU H. A review on the attention mechanism of deep learning[J]. Neurocomputing, 2021, 452: 48-62.

[14] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2017: 2223-2232.

[15] LI B, REN W, FU D, et al. Benchmarking single-image dehazing and beyond[J]. IEEE Transactions on Image Processing, 2019, 28(1): 492-505.

[16] ANCUTI C O, ANCUTI C, TIMOFTE R, et al. O-HAZE: A dehazing benchmark with real hazy and haze-free outdoor images[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition . Piscataway: IEEE Press, 2018: 754-762.

Improved transfer learning based dual-branch convolutional neural network image dehazing

LI Yunhong*, YU Huikang, MA Dengfei, SU Xueping, DUAN Jiaojiao, SHI Hanchi

(School of Electronics and Information, Xi'an Polytechnic University, Xi'an 710048, China)

Abstract: To address the problems of incomplete dehazing and image color distortion in the existing image dehazing algorithms, a dehazing network combining transfer learning sub-net and residual attention sub-net is proposed. First, the pre-trained model of the transfer learning subnet is adopted to enhance the feature attributes of the samples. Second, the structure of the dual-branch network is constructed, and the residual attention sub-network is used to assist the transfer learning sub-network to train the parameters of the network model. Finally, the method of tail ensemble learning is used to fuse the features of the dual network to obtain the model parameters of the dehazed image, so as to complete the image restoration task.The experimental results show that the algorithm proposed in the paper improves the PSNR index by 1.87 dB and 4.22 dB on the RESIDE dataset and the O-HAZE dataset respectively compared to GCANet, and the SSIM index on the O-HAZE dataset by 6.7% compared to GCANet.

Keywords: image dehazing; transfer learning; convolutional neural network; attention mechanism; ensemble learning

Received: 2022-04-18; **Accepted**: 2022-07-08; **Published Online**: 2022-08-26 12:57

URL: link.cnki.net/urlid/11.2625.V.20220826.1036.002

Foundation items: National Natural Science Foundation of China (61902301); Key Projects of Natural Science Basic Research Program of Shaanxi Provincial Department of Science and Technology (2022JZ-35); Natural Science Basic Research Program of Shaanxi Provincial Department of Education (19JK0364); China National University Student Innovation & Entrepreneurship Development Program (202210709012); The Youth Innovation Team of Shaanxi Universities

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0263

基于截断 p-shrinkage 范数的航空发动机数据重构

张红梅*，武江南，赵永梅，曾航，李全根
(空军工程大学 装备管理与无人机工程学院，西安 710051)

摘 要：针对航空发动机传感器的数据缺失问题，提出基于张量奇异值阈值 (TSVT) 的张量重构模型 LRTC-PTNN，对航空发动机的传感器数据进行重构。LRTC-PTNN 模型运用截断 p-shrinkage 范数的方式代替原始张量迹范数作为张量秩的凸包络，并根据 TSVT 的特性，计算了传感器之间的相关性，选取传感器截面作为重构精度最佳的数据输入方向，使用交替乘子法实现 LRTC-PTNN 算法。选取 NASA 提供的 PHM2008 数据集进行实验，对数据集进行标准化，并在重构后进行恢复，将多个时间序列个数相近的发动机传感器数据构建为高维张量的形式，设置 2 种传感器的数据缺失场景进行实验，结果表明：重构后数据的均方根误差和平均绝对百分比误差范围分别为 2.10～13.13 和 0.32～1.49，LRTC-PTNN 模型优于现有的基线模型，且在极端情况下有较强的鲁棒性。

关 键 词：航空发动机；数据缺失；张量；截断 p-shrinkage 范数；交替乘子法
中图分类号：V221+.3; TB553
文献标志码：A **文章编号**：1001-5965(2024)01-0039-09

航空发动机是复杂的空气动力学和热力学旋转机械系统，其在具有极高温度、压力、旋转和强烈振动的环境中工作，经常承受高负荷和热冲击，工作环境异常严酷。因此，在正常的操作条件下，航空发动机的关键部件会不断地累积退化，部件退化到一定程度后，航空发动机的性能就会偏离原来的水平，可能会导致各种故障，而航空发动机作为飞机最重要的部件之一，其可靠性对飞行的安全至关重要。航空发动机的健康管理[1]非常复杂且成本高昂，合理的维护可以在一定程度上解决该问题。通常，使用传感器对航空发动机进行状态监测，根据所得到的状态数据进行故障诊断[2]、故障隔离和剩余寿命预测[3-4]等。

但是，在实际的数据采集过程中，由于不可避免的事件发生(如传感器损坏、执行机构损坏等[5])，会导致数据采集中断，造成部分传感器数据缺失等问题，然而，完整有效的数据对航空发动机的状态检测与健康管理来说意义重大。在实际应用中，完整的退化信息是后续任务的必要输入。如果数据库中收集到的是不完整的缺失数据，不仅会造成实际获取与预先估计数据之间的差异，还会使计算的准确性降低，这就形成了数据处理误差，制约后续的预测和维护。

近年来，如何对缺失数据进行恢复是国内外学者研究的热点之一。对于现有的数据重构研究，可以分为 2 类：基于插值的方法和基于机器学习的方法。基于插值的方法有经典的众数、平均值、期望最大化(expectation-maximum, EM)算法，以及 Ni 等[6]提出的历史邻近补全、K 最近邻(K-nearest neighbors, KNN)插补等，但上述基于插值的方法除去 KNN 插补外，要求单个时间序列的内部特征呈线性或周期性，当数据缺失率较高而打破了这种内部特征时，基于插值的方法会出现较大误差。于是，基于机器学习的重构方法被广泛应用，而基于机器学习的重

收稿日期：2022-04-20；**录用日期**：2022-07-02；**网络出版时间**：2022-08-11 13:54
网络出版地址：link.cnki.net/urlid/11.2625.V.20220810.0855.001
基金项目：国家自然科学基金 (62002381)；空军工程大学创新实践基金 (CXJ2021097)
* **通信作者**. E-mail：zhm_plum@163.com

引用格式：张红梅，武江南，赵永梅，等. 基于截断 p-shrinkage 范数的航空发动机数据重构 [J]. 北京航空航天大学学报，2024，50（1）：39-47. ZHANG H M，WU J N，ZHAO Y M，et al. Aero-engine data reconstruction based on truncated p-shrinkage norm [J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50（1）：39-47（in Chinese）.

构方法又分为基于统计学习的方法和其他。基于统计学习的方法有 Cai 等[7]提出的奇异值阈值(singular value thresholding, SVT)模型，在此基础上，Liu 等[8]提出了低秩矩阵补全(low rank matrix completion, LRMC)模型与低秩张量补全(low rank tensor completion, LRTC)模型，之后，Chen 等[9-10]提出了时间正则化矩阵分解(time regularized matrix decomposition, TRMF)和贝叶斯时间矩阵分解(Bayesian time matrix decomposition, BTMF)模型，周媛等[11]提出了基于子空间的方法 PIMC 等。其他的传感器数据重构方法有 Elnour 等[12]提出的自动关联神经网络(auto-associative neural network, AANN)，Zhao 等[13]提出的基于加权在线序贯极限学习机(weighted online sequential extreme learning machine, WOS-ELM)等。然而，对于航空发动机的传感器数据重构仍面临如下问题：①衡量发动机中传感器数据之间的线性相关性特征；②充分利用这种相关性提升重构精度。

针对上述 2 个问题，本文运用张量模式对航空发动机的传感器数据进行重构，提出了基于截断 p-shrinkage 范数的低秩张量补全(low rank tensor completion based on truncated p-shrinkage norm, LRTC-PTNN)模型。将数据进行预处理，标准化到阈值区间；提出航空发动机退化趋势相似性假设，将数据构建为张量模式，计算传感器之间的线性相关性；为充分使用传感器内部的相关性信息，选取传感器与时间序列的切片方向作为数据的输入方向，运用 LRTC-PTNN 模型对缺失位置的数据进行恢复，进行数据的反向标准化处理。实验结果表明，基于张量的方法具有更高的重构精度与鲁棒性效果。

1 基于截断 p-shrinkage 范数的 LRTC-PTNN 模型

1.1 截断 p-shrinkage 范数

截断 p-shrinkage 范数作为秩最小化问题的替代方法，综合了截断核范数与 p-shrinkage 范数的优点，在对秩最小化问题的逼近过程中比单一范数更为紧致。截断 p-shrinkage 范数用截断的方式保留了大奇异值的完整特征，只针对截断后 $\min(n_1,n_2)-r$ 部分的奇异值进行 p-shrinkage 形式的收缩，且在收缩过程中，奇异值越小，获得的阈值惩罚越大，最大程度上保留了足够多的内部强相关性信息。

首先，描述矩阵的 p-shrinkage 范数定义[14]：给定一个矩阵 $\boldsymbol{X}\in\mathbf{R}^{m\times n}$，对于任何的 $p\in\mathbf{R}$，p-shrinkage 范数定义为

$$\|\boldsymbol{X}\|_p=\sum_{i=1}^{\min(m,n)}|\sigma_i|^{p-1} \tag{1}$$

式中：$\|\cdot\|_p$ 表示取矩阵的 p-shrinkage 范数；假定矩阵 $\boldsymbol{X}$ 的奇异值分解为 $\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}}$，则有 $\boldsymbol{\Sigma}=\mathrm{diag}(\sigma_i(\boldsymbol{X}))$，$\sigma_i$ 为矩阵的奇异值，即矩阵的核范数，并且奇异值由大到小排序为：$\sigma_1\geqslant\sigma_2\geqslant\cdots\geqslant\sigma_{\min}\{m,n\}\geqslant 0$；$p$ 为收缩参数，且当 $p=2$ 时，p-shrinkage 范数等价于矩阵核范数。

将 p-shrinkage 范数扩展到张量模式，对于给定的 $\boldsymbol{x}\in\mathbf{R}^{n_1\times n_2\times n_3}$，则张量模式下的 p-shrinkage 范数[15]可以定义如下：

$$\|\boldsymbol{x}\|_p=\frac{1}{n_3}\sum_{k=1}^{n_3}\|\overline{\boldsymbol{X}}^{(i)}\|_p=\frac{1}{n_3}\sum_{k=1}^{n_3}\sum_{i=1}^{\min\{n_1,n_2\}}|\overline{\sigma}_i|^{p-1} \tag{2}$$

根据张量的奇异值分解原理，将 $\boldsymbol{x}$ 沿第三维度 n_3 进行傅里叶变换，$\overline{\boldsymbol{X}}^{(i)}$ 表示变换后的正向切片，为矩阵形式，$|\overline{\sigma}_i|$ 表示正向切片 $\overline{\boldsymbol{X}}^{(i)}$ 进行奇异值分解后的奇异值，即张量核范数。

张量的截断核范数定义[16]为：给定一个三阶张量 $\boldsymbol{x}\in\mathbf{R}^{n_1\times n_2\times n_3}$ 和一个正整数 $t=\min(n_1,n_2)$，截断核范数为张量最小奇异值之和的形式，即

$$\begin{aligned}\|\boldsymbol{x}\|_{r,*}&=\frac{1}{n_3}\|\overline{\boldsymbol{X}}\|_{r,*}=\frac{1}{n_3}\sum_{j=1}^{n_3}\sum_{i=r+1}^{t}\left|\sigma_i\left(\overline{\boldsymbol{X}}^{(j)}\right)\right|=\\&\frac{1}{n_3}\sum_{j=1}^{n_3}\sum_{i=1}^{t}\left|\sigma_i\left(\overline{\boldsymbol{X}}^{(j)}\right)\right|-\frac{1}{n_3}\sum_{j=1}^{n_3}\sum_{i=1}^{r}\left|\boldsymbol{\sigma}_i\left(\overline{\boldsymbol{X}}^{(j)}\right)\right|=\\&\sum_{i=1}^{t}|\sigma_i(\boldsymbol{x})|-\sum_{i=1}^{r}|\boldsymbol{\sigma}_i(\boldsymbol{x})|\end{aligned} \tag{3}$$

式中：$\|\cdot\|_{r,*}$ 表示张量的截断核范数；$\boldsymbol{\sigma}_i(\boldsymbol{x})$ 为 $\boldsymbol{x}$ 的第 i 个奇异值，且截断参数 $r<\min(n_1,n_2)$。

结合 2 种张量的范数定义，给出张量的截断 p-shrinkage 范数定义：

$$\|\boldsymbol{x}\|_{p,r}=\frac{1}{n_3}\sum_{j=1}^{n_3}\sum_{i=r+1}^{t}|\overline{\sigma}_i|^{p-1} \tag{4}$$

式中：$\|\cdot\|_{p,r}$ 表示截断 p-shrinkage 范数，且在张量模式下，对于任何给定的 $\boldsymbol{x}$、$\boldsymbol{y}\in\mathbf{R}^{n_1\times n_2\times n_3}$，$\mu>0$，$p\in\mathbf{R}$，$r<\min(n_1,n_2)$，定义截断 p-shrinkage 范数的收缩阈值算子为 $S^{\mu}_{p,r}(\cdot)$，有

$$S^{\mu}_{p}(\boldsymbol{y})=\arg\min_{\boldsymbol{x}\in\mathbf{R}^{n_1\times n_2\times n_3}}\|\boldsymbol{x}\|_{p,r}+\frac{1}{2\mu}\|\boldsymbol{x}-\boldsymbol{y}\|_F^2 \tag{5}$$

由张量奇异值阈值(tensor singular value thresholding, TSVT)定理，$S^{\mu}_{p}(\boldsymbol{y})$ 是关于 $\boldsymbol{x}$ 优化问题的求解，μ 为奇异值阈值分解的阈值参数，p 为收缩参数，当 $-\infty<p<1$ 时，截断 p-shrinkage 范数收缩阈值满足输入越大，受到的惩罚越小的性质。$\|\cdot\|_F$ 表示 Frobenius 范数，对于 $S^{\mu}_{p}(\boldsymbol{y})$ 的求解有 $S^{\mu}_{p,r}(\boldsymbol{y})=\boldsymbol{U}S^{\mu}_{p,r}\cdot(\boldsymbol{\Sigma})\boldsymbol{V}^{\mathrm{T}}$，且

$$(S_{p,r}^{\mu}(\boldsymbol{\Sigma}))_{ijk}=$$
$$\operatorname{diag}\left(\left(\boldsymbol{\sigma}_1,\cdots,\boldsymbol{\sigma}_r,\left[\boldsymbol{\sigma}_{r+1}-\mu|\boldsymbol{\sigma}_{r+1}|^{p-1}\right]_+,\cdots,\right.\right.$$
$$\left.\left.\left[\boldsymbol{\sigma}_{\min\{m,n\}}-\mu\left|\boldsymbol{\sigma}_{\min\{m,n\}}\right|^{p-1}\right]_+\right)^{\mathrm{T}}\right) \tag{6}$$

对于任何$a\in\mathbf{R}$, 加函数$(\cdot)_+$的定义为$a_+=\max\{a,0\}$。

1.2　LRTC-PTNN 模型

基于 TSVT 的 LRTC 最早由 Zhang 和 Aeron 等[17]为解决图像的修复问题所提出, 之后常被扩展用于具有内部相关性数据的重构问题, 与基于多重截断核范数(tensor nuclear norm, TNN)的 LRTC 类似, 目的都在于用已知信息位置对未知位置进行填充, 使得重构值尽量逼近原始的真实值, 分别适用于不同的场景下, 都可看作是低秩矩阵补全模型的高阶扩展。对于航空发动机的数据重构问题, 运用基于张量平均秩概念的框架进行建模:

$$\begin{aligned}&\min_{\boldsymbol{M}}:\operatorname{rank}_a(\boldsymbol{y})\\&\text{s.t. }P_{\Omega}(\boldsymbol{y})=P_{\Omega}(\boldsymbol{x})\end{aligned} \tag{7}$$

式中: $\operatorname{rank}_a(\boldsymbol{y})$表示取张量的平均秩最小化问题, $\boldsymbol{y}$为重构张量; $\boldsymbol{x}$为原始的观测值张量。

对于Ω定义为

$$(P_{\Omega}(\boldsymbol{x}))_{ijk}=\begin{cases}x_{ij} & (i,j,k)\in\Omega\\0 & (i,j,k)\notin\Omega\end{cases} \tag{8}$$

式中: $P_{\Omega}(\boldsymbol{x})$为已知元素的下标集合, 即$x_{ijk}$为观测值, $P_{\bar{\Omega}}(\boldsymbol{x})$为$\boldsymbol{x}$中缺失元素的下标集合。

已知$\operatorname{rank}_a(\boldsymbol{y})$是非凸且不连续的, 式 (7) 是非凸问题, 难以求解, 根据上述 PTNN, 有: $\|\boldsymbol{x}\|_*<\|\boldsymbol{x}\|_p<\|\boldsymbol{x}\|_{p,r}<\operatorname{rank}_a(\boldsymbol{x})$, 于是用张量的 PTNN 来代替张量的核范数作为张量平均秩最小化最紧密的凸包络问题, 后续引入辅助张量变量$\boldsymbol{M}\in\mathbf{R}^{n_1\times n_2\times n_3}$, 给定约束条件限制, 原始的张量补全模型可以重写为

$$\begin{aligned}&\min_{\boldsymbol{M}}:\ \|\boldsymbol{y}\|_{p,r}\\&\text{s.t. }\begin{cases}\boldsymbol{M}=\boldsymbol{y}\\P_{\Omega}(\boldsymbol{M})=P_{\Omega}(\boldsymbol{x})\end{cases}\end{aligned} \tag{9}$$

对于该模型, 通过经典的交替乘子法(alternating direction method of multipliers, ADMM)进行求解, 根据式(9)来定义增强的拉格朗日函数, 如下:

$$L_{\rho}(\boldsymbol{y},\boldsymbol{M},\boldsymbol{Z},\rho)=\|\boldsymbol{y}\|_{p,r}+\frac{\rho}{2}\|\boldsymbol{M}-\boldsymbol{y}\|_F^2+\langle\boldsymbol{Z},\boldsymbol{M}-\boldsymbol{y}\rangle \tag{10}$$

式中: $\boldsymbol{Z}\in\mathbf{R}^{n_1\times n_2\times n_3}$为拉格朗日乘子; ρ为加入约束条件后的惩罚参数。根据 ADMM 框架的求解方式, 将$\boldsymbol{y}^l$、$\boldsymbol{M}^l$、$\boldsymbol{Z}^l$、ρ分别拆解为局部的子问题, 通过交替更新的方式来逼近问题的最优解, 在此过程中固定其他位置元素, 迭代更新的顺序如下:

$$\boldsymbol{y}^{l+1}:=\arg\min_{\boldsymbol{y}}L\left(\boldsymbol{y}^l,\boldsymbol{M}^l,\boldsymbol{Z}^l,\rho\right)$$
$$\boldsymbol{M}^{l+1}:=\arg\min_{\boldsymbol{M}}\min L\left(\boldsymbol{M}^l,\boldsymbol{y}^{l+1},\boldsymbol{Z}^l,\rho\right)$$
$$\boldsymbol{Z}^{l+1}:=\boldsymbol{Z}^l+\rho\left(\boldsymbol{y}^{l+1}-\boldsymbol{M}^{l+1}\right)$$

1) 更新$\boldsymbol{y}^{l+1}$, 固定$\boldsymbol{M}^l$和$\boldsymbol{Z}^l$的参数。

$$\begin{aligned}\boldsymbol{y}^{l+1}&=\arg\min_{\boldsymbol{y}}L\left(\boldsymbol{y}^l,\boldsymbol{M}^l,\boldsymbol{Z}^l,\rho\right)=\\&\|\boldsymbol{y}\|_{p,r}+\frac{\rho}{2}\|\boldsymbol{M}^l-\boldsymbol{y}^l\|_F^2+\left\langle\boldsymbol{Z}^l,\boldsymbol{M}^l-\boldsymbol{y}^l\right\rangle=\\&\|\boldsymbol{y}\|_{p,r}+\frac{\rho}{2}\left\|\boldsymbol{y}^l-\left\langle\boldsymbol{M}^l+\frac{1}{\rho}\boldsymbol{Z}^l\right\rangle\right\|_F^2=\\&S_{p,r}^{1/\rho}\left(\boldsymbol{M}^l+\frac{1}{\rho}\boldsymbol{Z}^l\right)\end{aligned} \tag{11}$$

$S_{p,r}^{1/\rho}\left(\boldsymbol{M}^l+\frac{1}{\rho}\boldsymbol{Z}^l\right)$为 TSVT[14] 分解, 其中, $\mu=1/\rho$。

2) 更新$\boldsymbol{M}^{l+1}$, 固定$\boldsymbol{y}^{l+1}$和$\boldsymbol{Z}^l$的参数。

$$\begin{aligned}\boldsymbol{M}^{l+1}&=\arg\min_{\boldsymbol{M}}L\left(\boldsymbol{M}^l,\boldsymbol{y}^{l+1},\boldsymbol{Z}^l,\rho\right)=\\&\frac{\rho}{2}\|\boldsymbol{M}^l-\boldsymbol{y}^{l+1}\|_F^2+\left\langle\boldsymbol{Z}^l,\boldsymbol{M}^l\right\rangle=\\&\frac{\rho}{2}\|\boldsymbol{M}^l\|_F^2-\rho\left\langle\boldsymbol{y}^{l+1},\boldsymbol{M}^l\right\rangle+\left\langle\boldsymbol{Z}^l,\boldsymbol{M}^l\right\rangle=\\&\boldsymbol{y}^{l+1}-\frac{1}{\rho}\boldsymbol{Z}^l\end{aligned} \tag{12}$$

再加入约束条件$P_{\Omega}(\boldsymbol{M})=P_{\Omega}(\boldsymbol{x})$, 则有

$$\boldsymbol{M}^{l+1}=\left(\boldsymbol{y}^{l+1}-\frac{1}{\rho}\boldsymbol{Z}^l\right)_{\bar{\Omega}}+P_{\Omega}(\boldsymbol{x}) \tag{13}$$

3) 求解$\boldsymbol{Z}^{l+1}$。

$$\boldsymbol{Z}^{l+1}=\boldsymbol{Z}^l+\rho\left(\boldsymbol{y}^{l+1}-\boldsymbol{M}^{l+1}\right) \tag{14}$$

设置迭代次数, 对于每次的迭代按照上述步骤进行更新, 寻求模型最优解, 则 LRTC-PTNN 算法的步骤如下:

输入: $\boldsymbol{M}$, $P_{\Omega}(\boldsymbol{M})=P_{\Omega}(\boldsymbol{x})$, $\boldsymbol{Z}=\boldsymbol{y}=\boldsymbol{0}$, ρ, l。

输出: $\boldsymbol{M}$。

1: for $i=1$ to l

2: 更新。

$$\boldsymbol{y}^{l+1}=\boldsymbol{S}_{p,r}^{1/\rho}\left(\boldsymbol{M}^l+\frac{1}{\rho}\boldsymbol{Z}^l\right)$$

3: 更新。

$$\boldsymbol{M}^{l+1}=\left(\boldsymbol{y}^{l+1}-\frac{1}{\rho}\boldsymbol{Z}^l\right)_{\bar{\Omega}}+P_{\Omega}(\boldsymbol{x})$$

4: 更新。

$$\boldsymbol{Z}^{l+1}=\boldsymbol{Z}^l+\rho\left(\boldsymbol{y}^{l+1}-\boldsymbol{M}^{l+1}\right)$$

5: if

$$\frac{\left\|\boldsymbol{M}^{l+1}-\boldsymbol{M}^l\right\|_F^2}{\|P_{\Omega}(\boldsymbol{M})\|_F^2}<\varepsilon$$

break

在整个数学模型的构建与求解过程中, 并未考

虑噪声的影响，LRTC-PTNN模型仅考虑给定数据的重构精度问题。

2 实 验

2.1 数据集

选取NASA给出的PHM2008实验数据集为例，该数据集与其他发动机退化数据集类似(参见Turbofan发动机退化模拟数据集)，只是未显示真实的RUL值，单个航空发动机的传感器数据用列向量表示为 $\boldsymbol{P}=[P_1,P_2,\cdots,P_N]$。PHM2008数据集中共有若干个航空发动机，且给定的发动机都属于同一型号，每个发动机中共有21个传感器，将这若干个发动机的传感器数据共同构建为矩阵的形式，大小为 $Y\in\mathbf{R}^{m\times 21}$，其中，每个发动机传感器数据的时间序列长度并不相同，但都表示发动机完整的寿命周期。

PHM2008实验数据集中，航空发动机内部的传感器分别为风扇入口处的总温度 T_2、LPC出口的总温度 T_{24}、HPC出口的总温度 T_{30}、LPT出口的总温度 T_{50}、风机入口压力 P_2、旁路管道中的总压力 P_{15}、HPC出口总压力 P_{30}、物理风扇速度 N_{f}、物理核心速度 N_{c}、发动机压力比 e_{pr}、HPC出口处的静压 P_{s30}、燃油流量与 P_{s30} 之比 φ、修正后的风扇转速 N_{Rf}、修正后的核心速度 N_{Rc}、旁路比 B_{BPR}、燃烧器空燃比 F_{farB}、渗漏焓 H_{htBleed}、需求的风扇速度 $N_{\mathrm{f_dmd}}$、要求校正的风扇速度 $P_{\mathrm{PCNfR_dmd}}$、HPT冷却剂的排放 W_{31} 和LPT冷却剂的排放 W_{32}。HPC与LPC分别为高压压气机和低压压气机，HPT和LPT分别为高压涡轮和低压涡轮。

2.2 数据预处理

本文提出的张量模型需要将数据构建为张量模式。已知每个发动机的传感器时间序列是不等长的，但发动机都属于同一型号，在此，对每个航空发动机的寿命退化数据做出相似性假设，构建出一个时间序列不等长的多发动机张量模式。在实际问题中，对于航空发动机某个时间段的传感器数据缺失问题，以该时间段为中心，围绕进行截取，截取为在缺失时间段周围等长的数据，并且时间序列尽可能长，足够长的时间序列可以很好地表征退化趋势，蕴含更多的信息。在此，实验过程中为了凸显模型的重构优势，对于PHM2008数据集，取传感器时间序列超过100个的发动机数据，并将其进行截取，最终构建为268×100×21的张量数据。然而，因为一些传感器之间数值差异过大，不方便进行数学计算，所以将所有传感器数据统一标准化到一个区间范围，再进行张量数据的构建与重构计算，后续以相反的方式恢复到各自的区间范围。整个航空发动机数据的模式构建过程如图1所示。

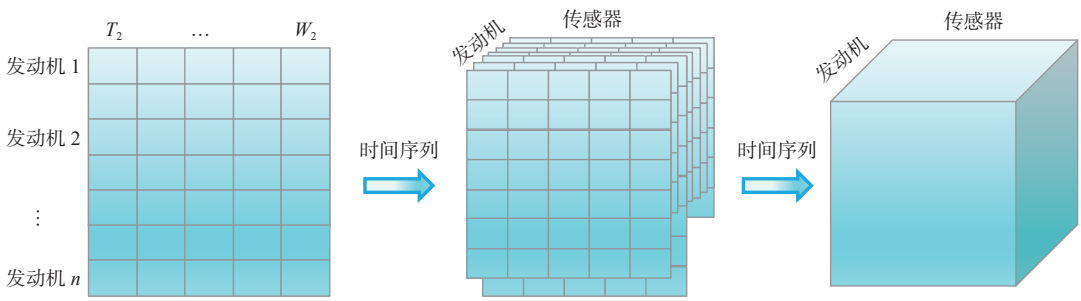


图1 张量模式构建

Fig. 1 Tensor pattern construction

2.3 基线模型选取

对于基线模型，选取经典的矩阵形式的重构模型进行对比，现行的3种矩阵形式的高精度重构模型有KNN插补、SVT和LRMC等。KNN插补模型是一种基于相似度插补的模型，SVT是基于矩阵奇异值阈值分解的重构模型，LRMC是经典的低秩矩阵重构模型。通过与基线模型的对比，凸显出用LRTC-PTNN模型将数据构建为高维进行重构的精度优势。

2.4 相关性分析

在数据预处理阶段，本文提出了航空发动机退化趋势相似性的假设，但NASA给定的数据集并无对每个发动机数据集采集时间与间隔的详细解释，相似性无法直观证明，而基于TSVT的LRTC模型对数据的重构效果与数据的输入方向有关，为选取合适的数据输入方向，使用皮尔森相关性系数来计算传感器之间的相关性，计算公式如下：

$$\rho_{x,y}=\frac{\mathrm{cov}(x,y)}{\sigma_x\sigma_y} \tag{15}$$

式中：$\mathrm{cov}(x,y)$为不同列向量之间的协方差；σ_x、σ_y为标准差；$\rho_{x,y}$为2个列向量之间的皮尔森相关系数，且$\rho_{x,y}\in(-1,1)$，分别表示传感器之间的负相关与正相关，通过热力图来直观呈现，如图2所示。

可以看到，除去传感器旁路比的相关性较弱

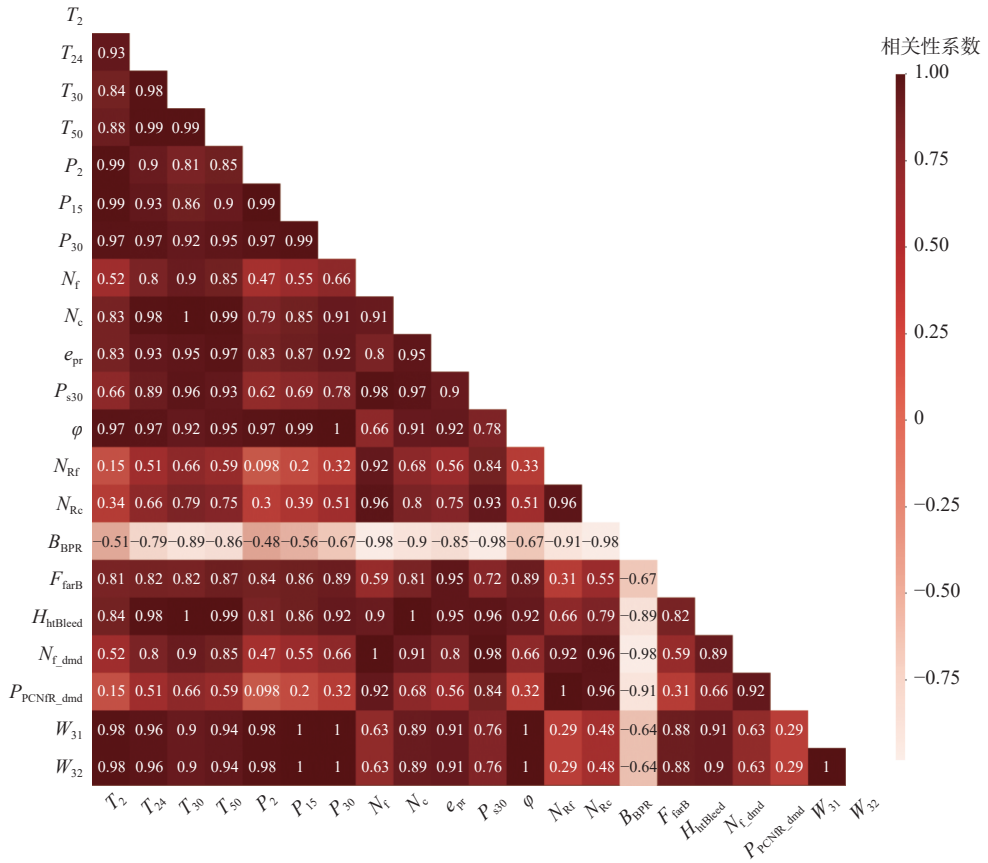


图 2　传感器之间的相关性热力图

Fig. 2　Thermal map of correlation between sensors

外，其余传感器之间的相关性较强，特别是邻近的传感器之间，在后续的 LRTC-PTNN 模型中选取传感器与时间序列的切面进行计算，保证有足够多的相关性信息可以使用。

2.5　对比方法

在实际的工程应用中，航空发动机数据缺失问题主要分为随机缺失与非随机缺失，在此，设置 2 种实验场景。PHM2008 实验数据集是一个完整的数据集，用于补全实验时，需要根据场景先对数据进行丢失处理，再通过重构的方式将数据集恢复，通过计算重构值与真实值的差值来对模型的优劣进行度量，使用均方根误差 e_{RMSE} 和平均绝对百分比误差 e_{MAPE} 的信息来判断补全的精度。定义如下：

$$\begin{cases} e_{\mathrm{RMSE}} = \sqrt{\dfrac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2} \\ e_{\mathrm{MAPE}} = \dfrac{1}{n}\sum_{i=1}^{n}\left|\dfrac{y_i-\hat{y}_i}{y_i}\right|\times 100\% \end{cases} \tag{16}$$

式中：y_i 为预测值；$\hat{y}_i$ 为真实值。

模型参数设置：参照 LRTC[8] 框架的参数设置方式，但由于 p-shrinkage 的缘故，初始ρ需要设置较大，设置为 1×10^{-9}，每次通过$\rho=\min\{1.05\rho,\rho_{\max}\}$来对$\rho$进行更新，对于截断参数$r$，随机缺失情况下设置为 0.1，非随机缺失情况下设置为 0.05，模型的迭代次数设置为 200，对于收敛条件，用 $\left\|\boldsymbol{M}^{l+1}-\boldsymbol{M}^{l}\right\|_F^2 / \|P_{\Omega}(\boldsymbol{M})\|_F^2 < \varepsilon$来设置模型的收敛精度，$\varepsilon$的值设置为 1×10^{-9}。对于 KNN 模型，设置l的值为 10，对于 LRMC、SVT 模型，初始值ρ设置为 1×10^{-5}，其余参数设置与 LRTC-PTNN 相同。

1）对比方案。设置在随机缺失（RM）与非随机缺失（NM）场景下，PHM2008 数据集的缺失率为 20%～70%，分别使用均方根误差和平均绝对百分比误差计算原始值与重构值的差值，在此过程中，均方根误差和平均绝对百分比误差的值越小，重构精度越高，通过与基线模型进行对比，突出模型优势，如表 1 所示。

2）实验分析。在随机缺失场景下，设置低缺失率 20%～30% 的情况下，KNN 的插补精度很高，而 LRMC 和 SVT 与 KNN 相比来说重构精度较差，但随着缺失率的增加，缺失率 40% 的情况下，KNN 已经发生了漂移情况，插补性能下降，远低于 LRTC 和 SVT。在极端情况下，3 种基于矩阵的重构方法精度都不高，效果最好的是 LRMC，但仍发生了严重的失真情况，相较来说，基于矩阵的重构方法稳定性最好的也是 LRMC 模型，证明了低秩特性对重构的影响，而 LRTC-PTNN 作为 LRMC 的高维扩展

与增强，表现出了极强的重构效果和稳定性效果，在整个随机缺失情况下，LRTC-PTNN的重构精度都优于基于矩阵的重构方法，极端场景下，LRTC-PTNN也发生了失真，但相较于基线的重构模型，LRTC-PTNN的失真并不算严重，在可控范围内。以极端缺失场景下（缺失率为70%）为例，比较KNN与LRTC-PTNN的重构效果，展示原始值与重构值的差值情况，并且为了防止对比冗余，选取4个传感器作为其中的典型案例代表，如图3所示。

从图3可以直观看到，KNN模型在随机缺失的极端情况下，在每个传感器中都有部分点的重构值与原始值差异较大，并不能够完全恢复所有的原始数据信息，而LRTC-PTNN相较于KNN来说，在大多数传感器中的重构值与原始值基本重合，只有少数重构点值与原始点值仍存在差异，但这种残差已经相对较小。因此，在随机缺失场景中，LRTC-PTNN模型的重构效果最好。

在非随机缺失场景下，如表2所示，设置低缺失率20%～40%情况下，KNN的插补精度在基于矩阵的方法中最高，但随着缺失率（50%～70%）的增加，KNN在缺失率为50%时开始出现严重失真情况，此时LRMC的重构精度最高，LRMC在缺失率为70%时才发生失真情况，LRMC的重构效果在3种方法中最为稳定，而SVT在缺失率的增加中重构同样平稳，但重构效果最差。与基于张量的重构算法对比，LRTC-PTNN在非随机缺失场景下与

表1 RM情况下的数据重构精度对比

Table 1 Comparison of data reconstruction accuracy at RM

缺失率/%	均方根误差				平均绝对百分比误差			
	KNN	LRMC	SVT	LRTC-PTNN	KNN	LRMC	SVT	LRTC-PTNN
20	2.51	3.61	7.96	2.37	0.152	0.80	1.82	0.37
30	4.30	3.98	8.36	2.47	0.26	0.91	1.87	0.38
40	13.71	5.22	9.38	2.92	1.39	1.04	1.91	0.44
50	28.68	7.87	12.69	3.83	3.63	1.32	2.17	0.56
60	38.02	15.49	21.89	6.17	4.93	1.77	3.03	0.82
70	41.67	30.62	65.08	13.13	5.22	2.68	7.27	1.49

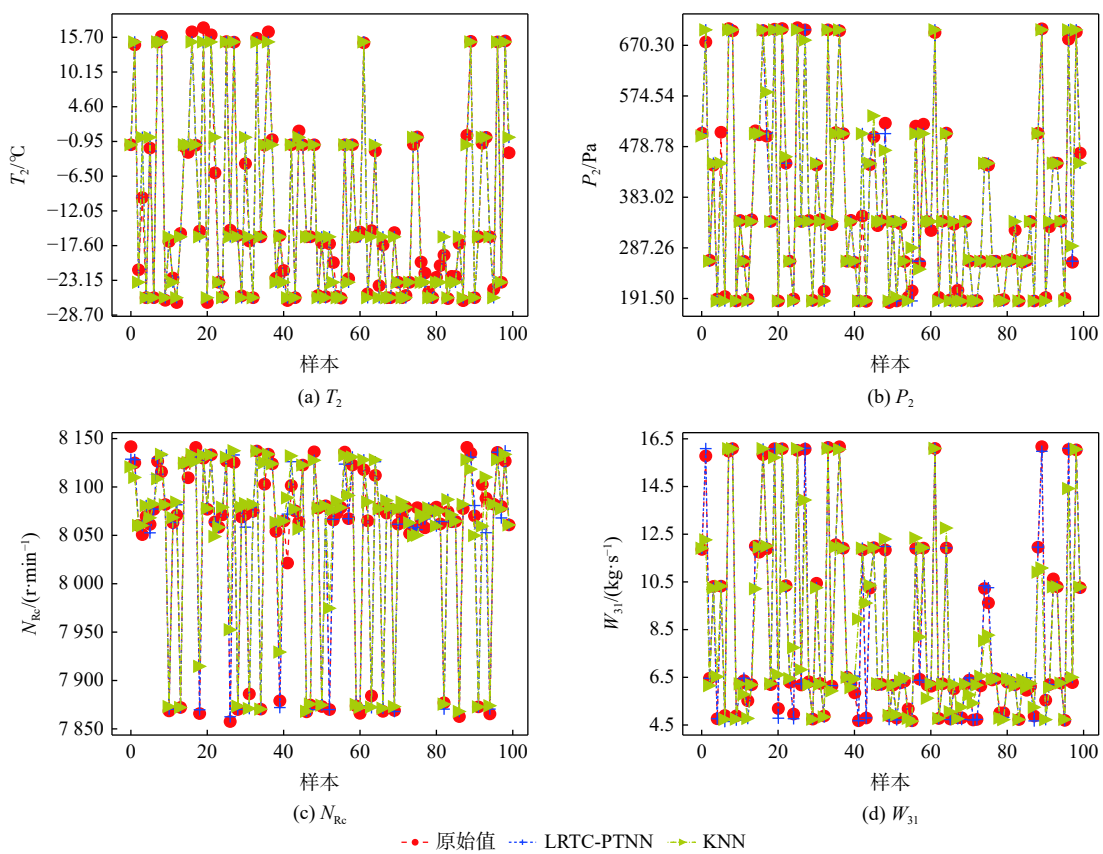


图3 RM情况下重构值与原始值对比

Fig. 3 Reconstructed value compared with original value in case of RM

在随机缺失场景下一样获得了很好的效果, 在整个的缺失率设置过程中, LRTC-PTNN 模型的重构精度都要优于基于矩阵的模型。同样以极端缺失场景下(70%)为例, 比较 KNN 与 LRTC-PTNN 的重构效果, 展示原始值与重构值的差值情况, 如图 4 所示。

从图 4 可以看到, 与随机缺失场景下相同, 并且保持与随机缺失场景下相同的实验设置。KNN 模型在非随机缺失的极端情况下, 在 T_2 与 W_{31} 中修复效果仍然较好, 而在 P_2 与 N_{Re} 中发生了严重的失真情况, 重构值与真实值差值较大, 而 LRTC-PTNN 模型与在随机缺失情况下类似, 在上述传感器中的重构值与原始值基本重合。因此, 在非随机缺失场景下, LRTC-PTNN 模型的重构效果最好。

在 2 种缺失情况下, 基于矩阵的方法 KNN 模型在低缺失率的情况下重构精度高于 LRMC 和 SVT, 但随着缺失率的增加, KNN 先发生失真, 其次是 SVT, 并且 SVT 的整体重构效果并不好, 相较来说, LRMC 在 3 种方法中最为稳定, 重构精度随着缺失率的增加下降更为平滑, 到了极端情况下才开始失真。当缺失率较低时, 数据集内部保留了大部分的发动机传感器数据, KNN 的方法通过计算相似度进行插补, 能够获得更多的信息, LRMC 与 SVT 都是基于低秩矩阵奇异值分解进行阈值迭代的模型, 阈值迭代的过程中因为是对所有的奇异值进行惩罚处理, 并无法保留主要的数据特征。高缺

表 2　NM 情况下的数据重构精度对比

Table 2　Comparison of data reconstruction accuracy at NM

缺失率/%	均方根误差				平均绝对百分比误差			
	KNN	LRMC	SVT	LRTC-PTNN	KNN	LRMC	SVT	LRTC-PTNN
20	2.57	3.59	7.80	2.10	0.14	0.78	1.77	0.32
30	2.57	3.52	7.59	2.43	0.15	0.91	1.88	0.38
40	3.00	5.23	10.27	2.65	0.20	1.10	1.96	0.43
50	15.37	7.51	11.51	4.28	0.90	1.32	2.10	0.63
60	28.81	12.03	29.01	5.12	2.90	1.76	3.85	0.76
70	32.94	29.75	72.40	11.66	3.22	2.72	8.23	1.34

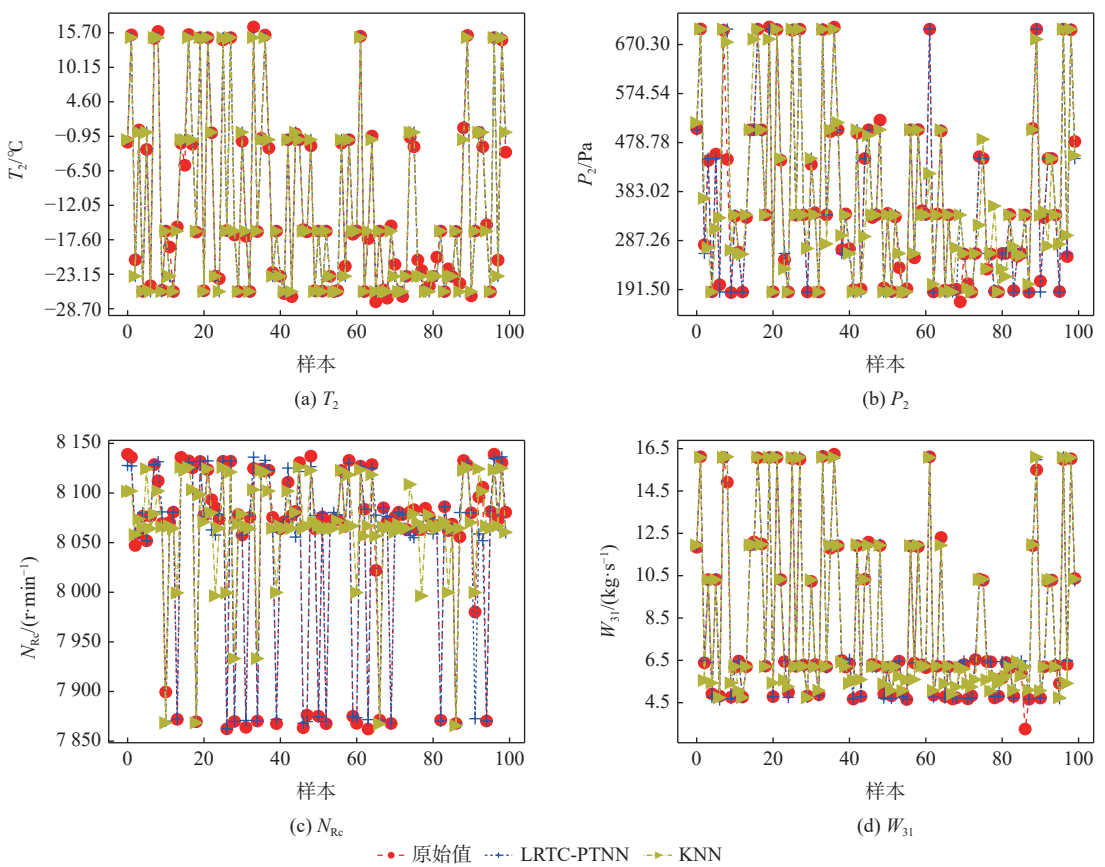


图 4　NM 情况下重构值与原始值对比

Fig. 4　Reconstructed value compared with original value in case of NM

失率情况下，KNN无法获得足够多的相似性信息，而LRMC与SVT又因为阈值迭代的方式相较来说更为稳定。在此，LRTC-PTNN表现出了很好的重构效果，在2种缺失的情况下重构精度都很高，虽然在极端情况下有失真情况，但相比于矩阵的重构模型效果较好，LRTC-PTNN一方面将数据构建为高维的紧凑结构能够保留数据内部足够多的相关性信息，其次选取传感器与时间序列作为切面，保证了在变换域求解时的低秩特性，充分利用了传感器的相关信息进行计算。

3 结 论

本文提出了截断p-shrinkage范数，并将其应用到LRTC框架中，提出了一了个航空发动机的传感器数据重构模型LRTC-PTNN，经过实验有如下结论：

1）航空发动机内部传感器之间的线性相关性极高，大多数传感器之间的相关性值高于0.8。

2）提出的LRTC-PTNN模型可以解决航空发动机传感器数据在不同缺失场景下的恢复问题，并且重构后数据的均方根误差和平均绝对百分比误差值分别介于2.10～13.13和0.32～1.49，重构值与原始值拟合度较高。

3）LRTC-PTNN模型的重构精度都要高于基于矩阵的模型，证明高维结构能够更充分地使用数据内部的相关性信息来提高重构精度。

4）LRTC-PTNN模型的缺点是要求数据基本等长，只能够选取时间序列相近长度的发动机数据进行预处理，后续将考虑引入多元时间序列的方式进行重构。

参考文献（References）

[1] HUANG Y, TANG Y F, VANZWIETEN J, et al. Reliable machine prognostic health management in the presence of missing data[J]. Concurrency and Computation Practice and Experience, 2020, 34(12): e5762.

[2] 孙瑞谦, 猴林峰, 韩小宝, 等. 考虑性能退化的航空发动机故障诊断量化评估[J]. 推进技术, 2022, 43(8): 337-348.
SUN R Q, GOU L F, HAN X B, et al. Quantitative evaluation of aeroengine fault diagnosis considering performance degradation[J]. Propulsion Technology, 2022, 43(8): 337-348(in Chinese).

[3] LI X Q, JIANG H K, LIU Y, et al. An integrated deep multiscale feature fusion network for aeroengine remaining useful life prediction with multisensor data[J]. Knowledge-based Systems, 2022, 235: 107652.

[4] 王昆, 郭迎清, 赵万里, 等. 基于SSAE和相似性匹配的航空发动机剩余寿命预测[J]. 北京航空航天大学学报, 2023, 49(10): 2817-2825.
WANG K, GUO Y Q, ZHAO W L, et al. Residual life prediction of aero-engine based on SSAE and similarity matching[J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49(10): 2817-2825(in Chinese).

[5] 孙浩, 郭迎清, 赵万里. 航空发动机传感器与执行机构信息重构算法[J]. 北京航空航天大学学报, 2020, 46(2): 331-339.
SUN H, GUO Y Q, ZHAO W L. Aircraft engine sensor and actuator information reconstruction algorithm[J]. Journal of Beijing University of Aeronautics and Astronautics, 2020, 46(2): 331-339(in Chinese) .

[6] NI D H, LEONARD J D, GUIN A, et al. Multiple imputation scheme for overcoming the missing values and variability issues in ITS data[J]. Journal of Transportation Engineering, 2005, 131(12): 931-938.

[7] CAI J F, CANDES E J, SHEN Z W, et al. A singular value thresholding algorithm for matrix completion[J]. SIAM Journal on Optimization, 2010, 20(4): 1956-1982

[8] LIU J, MUSIALSKI P, WONKA P, et al. Tensor completion for estimating missing values in visual data[C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2009: 2114-2121.

[9] CHEN X Y, HE Z, SUN L. A Bayesian tensor decomposition approach for spatiotemporal traffic data imputation[J]. Transportation Research Part C:Emerging Technologies, 2018, 98: 73-84.

[10] CHEN X Y, SUN L. Bayesian temporal factorization for multidimensional time series prediction[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(9): 4659-4673.

[11] 周媛, 左洪福, 何军. 信息缺失的航空发动机传感器数据重构[J]. 北京航空航天大学学报, 2016, 42(5): 891-898.
ZHOU Y, ZUO H F, HE J. Information missing aircraft engine sensor data refactoring[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(5): 891-898(in Chinese).

[12] ELNOUR M, MESKIN N, AL-NAEMI M. Sensor data validation and fault diagnosis using auto-associative neural network for HVAC systems[J]. Journal of Building Engineering, 2020, 27: 100935.

[13] ZHAO Z, LIU Z, SUN Y, et al. WOS-ELM-based double redundancy fault diagnosis and reconstruction for aeroengine sensor[J]. Journal of Control Science and Engineering, 2017, 2017: 1982879.

[14] LI Y F, SHANG K, HUANG Z H. A singular value p-shrinkage thresholding algorithm for low rank matrix recovery[J]. Computational Optimization & Applications, 2019, 73(2): 453-476.

[15] LIU C, SHAN H, CHEN C. Tensor p-shrinkage nuclear norm for low-rank tensor completion[J]. Neurocomputing, 2020, 387: 255-267.

[16] ZHANG D, HU Y, YE J, et al. Matrix completion by truncated nuclear norm regularization[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2012: 2192-2199.

[17] ZHANG Z M, AERON S. Exact tensor completion using t-SVD[J]. IEEE Transactions on Signal Processing, 2017, 65(6): 1511-1526.

Aero-engine data reconstruction based on truncated p-shrinkage norm

ZHANG Hongmei*，WU Jiangnan，ZHAO Yongmei，ZENG Hang，LI Quangen

(Equipment Management and Unmanned Aerial Vehicle Engineering College，Air Force Engineering University，Xi'an 710051，China)

Abstract：To address the data loss problem of aero engine sensors, a tensor reconstruction model LRTC-PTNN based on tensor singular value threshold (TSVT) is proposed to reconstruct the sensor data of aircraft engines. LRTC-PTNN uses truncation p-shrinkage norm to replace the original tensor trace norm as the convex envelope of tensor rank. According to the characteristics of TSVT, the correlation between sensors is calculated, and the data input direction with the best reconstruction accuracy is selected. The LRTC-PTNN algorithm was finally implemented using the alternating direction method of multipliers. Using the PHM2008 dataset provided by NASA for experiments, the dataset was standardized and restored after reconstruction, and the multiple time series similar number of engine sensor data were constructed into the form of high-dimensional tensor, and the data deletion scenarios of the two sensors were set for experiments. The results showed that the RMSE and MAPE values of the reconstructed data were between 2.10%−13.13% and 0.32%−1.49%, respectively; the LRTC-PTNN model was better than the existing baseline model; in extreme cases, the model also has strong robustness.

Keywords：aero engines；data loss；tensors；truncated p-shrinkage norm；alternating direction method of multipliers

Received：2022-04-20；**Accepted**：2022-07-02；**Published Online**：2022-08-11 13：54

URL：link.cnki.net/urlid/11.2625.V.20220810.0855.001

Foundation items：National Natural Science Foundation of China (62002381); Innovation Practice Fund of Air Force Engineering University (CXJ2021097)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0208

多约束条件下四旋翼机动轨迹跟踪一体化控制方法

王英勋，李欣，蔡志浩，赵江*
(北京航空航天大学 自动化科学与电气工程学院，北京 100191)

摘 要：四旋翼无人机高动态飞行需求的增加使其成为一个越来越热门的研究主题。针对四旋翼执行废墟裂缝和树林缝隙穿越任务时的大机动轨迹状态跟踪控制问题，提出基于模型预测控制的一体化控制方法，包含多约束条件下的高机动轨迹规划和多参考状态量的一体化跟踪控制，并通过飞行试验验证了所提方法相比前馈PID控制方法在跟踪规划的高机动轨迹时的优越性能。在飞行试验中，四旋翼成功穿越了60°滚转角的窄框，其实际滚转角达到60°大角度的同时，z轴误差仅为0.065 m。

关 键 词：多约束；一体化控制；模型预测控制；轨迹跟踪；轨迹规划
中图分类号：V249.1; TP249
文献标志码：A **文章编号**：1001-5965(2024)01-0048-13

无人机相比有人机具有体积小、造价低、使用方便、生存能力强等优势，最初主要应用于军事侦察、情报搜索、监视和目标追踪等方面。随着无人机技术的不断发展，无人机的民用场景也逐渐增加，包括灾害应急、教育、测绘、作物评估、交通监测等民用场景[1-4]。在多种无人机之中，四旋翼无人机以其独有的优势，在军用和民用方面的应用更加广泛。

四旋翼无人机具有较好的敏捷性和机动性，对于一些复杂的高动态任务场景，如在一些具有狭窄缝隙的环境中快速穿越飞行、跟踪具有较强机动性的目标、躲避突发动态威胁等，通过合理的硬件搭建和软件算法设计，能够最大限度地开发四旋翼的飞行潜力，使实现高动态飞行场景成为可能。随着应用场景复杂度的不断增加，四旋翼的高动态飞行需求也不断增加，成为一个越来越热门的研究主题。

为实现四旋翼的高动态飞行，许多学者在轨迹规划和飞行控制2方面进行了深入研究。美国国防部高级研究计划局的快速、轻量级、自主(fast lightweight autonomy, FLA)项目[5]以废墟建筑物内的救援或侦察任务为背景，实现了四旋翼在不需要外部通信或GPS情况下的室内和室外机动飞行。瑞士苏黎世大学机器人与感知实验室团队针对四旋翼在狭窄间隙中自主飞行问题，考虑几何、动态和感知约束规划了包括接近段和穿越段的穿框轨迹，实现了四旋翼自主且安全通过一个狭窄的倾斜框[6]。文献[7-10]重点研究了复杂环境的运动规划问题，提出多种有效的规划算法实现了四旋翼在复杂环境中可行轨迹的生成，并通过稳定的控制算法进行飞行验证。文献[7]通过B样条优化提高了轨迹的光滑性和间隙性。

在四旋翼的飞行控制算法研究方面，苏黎世大学团队对高机动轨迹的跟踪控制进行了深入研究。文献[11-13]主要研究了轨迹的精准跟踪控制问题，文献[11]提出一种角速率控制器和迭代推力混合的新方案，减少了偏航控制误差。文献[12]提出一种基于时变系统模型的线性二次调节(linear quadratic regulator, LQR)控制方法，实现了存在扰动时的轨迹跟踪控制。文献[13]考虑旋翼

收稿日期：2022-04-02；录用日期：2022-06-12；网络出版时间：2022-06-23 16:13
网络出版地址：link.cnki.net/urlid/11.2625.V.20220622.1758.004
*通信作者. E-mail：jzhao@buaa.edu.cn

引用格式：王英勋，李欣，蔡志浩，等. 多约束条件下四旋翼机动轨迹跟踪一体化控制方法[J]. 北京航空航天大学学报，2024，50(1)：48-60.
WANG Y X，LI X，CAI Z H，et al. Integrated control method for quadrotors' aggressive trajectory tracking under multiple constraints[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(1)：48-60 (in Chinese).

阻力模型, 大大降低了高速轨迹跟踪的误差。文献 [14-17] 进行了模型预测控制(model predictive control, MPC)在高机动轨迹跟踪方面的研究。文献 [14] 提出首个基于感知的模型预测控制方法, 改善感知和动作目标之间的冲突。文献 [15-16] 实现了非线性模型预测控制(nonlinear model predictive control, NMPC)对高机动轨迹的跟踪控制, 文献 [16] 提出的混合自适应模型预测控制方法能够精准地跟踪速度高达 70 km/h 的机动轨迹。文献 [17] 提出一种在线更新状态和权值的模型预测控制方法, 仿真验证了其相比固定权重时提高的 70% 轨迹跟踪精度。

目前, 四旋翼高机动轨迹跟踪的相关研究集中在位置量的精准跟踪。对于包含期望姿态的轨迹跟踪问题, 大部分学者选择优化高机动轨迹, 采用常规的比例-积分-微分(proportion-integral-differential, PID)控制器进行跟踪控制。本文以包含姿态的轨迹跟踪控制过程为研究重点, 针对四旋翼无人机在丛林、建筑物中穿越狭窄缝隙执行任务时需要解决的大机动轨迹精准状态跟踪控制问题, 设计了基于模型预测控制的一体化控制方法。其中, 反馈控制项保证了状态跟踪的稳定性, 基于 NMPC 的前馈控制项不仅提升了跟踪精度和实时性, 也保证了前馈控制的稳定性。二者融合后可同时对期望位置量、速度量、姿态量进行稳定、高实时性、高精度跟踪控制。本文提出一种包含一阶近似姿态控制器的四旋翼无人机非线性模型, 设计了一种多约束条件下的高机动轨迹规划算法, 提出了一种多参考状态量的一体化跟踪控制方法, 进行了一体化控制方法对多约束条件下规划的高机动穿框轨迹的跟踪飞行验证, 对比并表明本文方法相比常规前馈 PID 控制方法具有更加优越的性能。

1　多约束条件下四旋翼一体化控制方法

在一些特殊的任务场景下, 需要对四旋翼的飞行轨迹进行特定位置、速度、姿态的规划, 同时也要设计相应的控制律对上述多个参考状态量进行跟踪。本文方法包含了多约束条件下的高机动轨迹规划和多参考状态量的一体化跟踪控制。图 1 为本文方法的结构。

多约束条件下的高机动轨迹规划, 是在设定轨迹上某些点的位置、姿态、速度为定值的情况下, 考虑四旋翼的控制输入约束与轨迹的连续性、平滑性, 对多段轨迹基元进行合理时间分配的综合优化过程。规划的轨迹包含了期望位置量、速度量、加速度量及其多阶导数量, 可转换为参考状态量为后续的控制器提供参考输入。

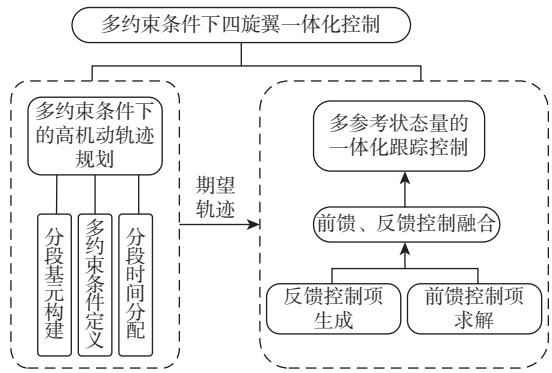


图 1　本文方法的结构

Fig. 1　Structure of proposed method

对于规划的多约束条件下的高机动轨迹, 利用本文方法进行跟踪, 本文方法可同时对期望位置量、速度量、姿态量进行跟踪控制, 通过权值系数调整实现不同参考状态量跟踪的平衡, 实现对位置量和姿态量同时地、一体化地跟踪控制过程。

模型预测控制的优势在于其预测行为, 主要步骤如下: ①建立合适的四旋翼动态模型; ②在某个时刻, 通过当前状态量和预测控制输入计算预测时域内的状态量; ③通过选择最优控制输入得到期望的状态量; ④在控制时域内执行最优控制输入; ⑤移动到下一个时刻, 从步骤②开始执行。

假设有如下四旋翼非线性系统:

$$\dot{\boldsymbol{x}}(t)=\boldsymbol{f}(\boldsymbol{x}(t),\boldsymbol{u}(t)) \tag{1}$$

式中: $\boldsymbol{x}(t)$和$\boldsymbol{u}(t)$分别为系统的状态量和控制输入; $\boldsymbol{f}(\boldsymbol{x}(t),\ \boldsymbol{u}(t))$为四旋翼动态模型, 并且对状态量和控制量可微。

假设所有的状态量是可观测的, 本文的目的是为了求解一个线性时变的、包含反馈项和前馈项的控制器。该控制器可表示为

$$\boldsymbol{u}(t)=\boldsymbol{u}_{\mathrm{ff}}(t)+\boldsymbol{K}(t)\boldsymbol{x}(t) \tag{2}$$

式中: $\boldsymbol{K}(t)$为反馈控制矩阵; $\boldsymbol{u}_{\mathrm{ff}}(t)$为前馈控制项。

该时变控制器的反馈项可在非线性系统线性化后, 通过线性二次调节控制方法进行求解; 前馈项可使用非线性模型预测控制方法, 通过迭代优化求解有限时域t_{f}最优控制问题, 使得以下代价函数最小:

$$J=h(\boldsymbol{x}(t_{\mathrm{f}}))+\int_0^{t_{\mathrm{f}}} l(\boldsymbol{x}(t),\boldsymbol{u}(t))\mathrm{d}t \tag{3}$$

式中: $l(\boldsymbol{x}(t),\boldsymbol{u}(t))$和$h(\boldsymbol{x}(t_{\mathrm{f}}))$分别为过程代价函数和终

止代价函数。

2 多约束条件下的四旋翼高机动轨迹规划

本文方法的首要阶段是多约束条件下的高机动轨迹规划。考虑四旋翼在执行穿越缝隙等具有大姿态角、大速度特性任务场景时的运动要求，四旋翼需要预规划一条满足多约束条件的穿越轨迹，作为后续一体化跟踪控制的参考状态量输入。四旋翼预规划的穿越轨迹需要具有合适的分段时间分配和尽可能小的控制输入，同时能够满足到达窄缝附近时的位置、姿态等约束。多约束条件下的轨迹规划问题可抽象为多段轨迹的最优控制问题进行求解。

2.1 关键点间的分段基元构建

首先进行多段轨迹基元的构建。由于本文规划的轨迹在某些位置点处需要满足特殊的状态需求，因此，可将轨迹上的这些特殊点定义为关键点，而关键点之间即为本节需要生成的分段轨迹基元。在轨迹基元生成过程中，考虑到四旋翼具有的微分平坦特性，即四旋翼输出量位置$\boldsymbol{p}$和偏航角ψ及其多阶导数能够表示所有状态和输入，因此，单段轨迹可用3个独立的一维时间参数化多项式函数表示[18]：

$$\boldsymbol{p}(t) := \left[p_x(t), p_y(t), p_z(t)\right]^{\mathrm{T}} \tag{4}$$

$$p_\mu(t) = \sum_{i=0}^{n} a_{\mu i} t^n \tag{5}$$

式中：$\mu \in \{x,y,z\}$；$a_{\mu i}$为多项式系数。期望轨迹的相关参数可表示为

$$\boldsymbol{\Phi}_\mu(t) = \left[p_\mu, \dot{p}_\mu, \ddot{p}_\mu, \dddot{p}_\mu, \ddddot{p}_\mu\right]^{\mathrm{T}} = \left[p_\mu, v_\mu, a_\mu, j_\mu, s_\mu\right]^{\mathrm{T}} \tag{6}$$

式中：p_μ为期望轨迹的位置；v_μ为速度；a_μ为加速度；j_μ为加速度的一阶导数加加速度；s_μ为为加速度的二阶导数加加加速度，为控制量。

期望轨迹的状态量选取为

$$\boldsymbol{s}(t) := \left[p_\mu, \dot{p}_\mu, \ddot{p}_\mu, \dddot{p}_\mu\right]^{\mathrm{T}} = \left[p_\mu, v_\mu, a_\mu, j_\mu\right]^{\mathrm{T}} \tag{7}$$

期望轨迹的控制量选取为

$$u_\mu(t) = s_\mu \tag{8}$$

选取以上状态量和控制量来建立最优控制问题，比较不同指标类型，由于轨迹的终止状态具有相关约束条件，可将积分型性能指标作为代价函数，以保证曲线的平滑，同时尽可能减小动力系统的压力。本文选取如下积分型性能指标作为代价函数：

$$J_\mu = \int_0^{t_{\mathrm{f}}} s_\mu(t)^2 \mathrm{d}t \tag{9}$$

此处先不考虑控制量的约束，仅给出如下的约束条件：

$$\dot{\boldsymbol{s}} = \boldsymbol{f}(\boldsymbol{s},u,t) = \left[v_\mu, a_\mu, j_\mu, s_\mu\right]^{\mathrm{T}} \tag{10}$$

$$\boldsymbol{s}(0) = \left[p_0, v_0, a_0, j_0\right]^{\mathrm{T}} \tag{11}$$

$$\boldsymbol{s}(t_{\mathrm{f}}) = \left[p_{\mathrm{f}}, v_{\mathrm{f}}, a_{\mathrm{f}}, j_{\mathrm{f}}\right]^{\mathrm{T}} \tag{12}$$

因此，将一个包含多约束条件的规划问题转化为一个典型的最优控制问题。针对上述积分型指标、t_{f}自由、末端固定的最优控制问题，可建立如下的哈密顿函数：

$$H(\boldsymbol{s},u,\boldsymbol{\lambda}) = s_\mu^2 + \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{f}(\boldsymbol{s},u) = s_\mu^2 + \lambda_1 v_\mu + \lambda_2 a_\mu + \lambda_3 j_\mu + \lambda_4 s_\mu \tag{13}$$

其中，$\boldsymbol{\lambda} = [\lambda_1,\lambda_2,\lambda_3,\lambda_4]^{\mathrm{T}}$。由正则方程$\dot{\boldsymbol{\lambda}} = -\dfrac{\partial H}{\partial \boldsymbol{s}}$可得：

$$\dot{\boldsymbol{\lambda}} = [0, -\lambda_1, -\lambda_2, -\lambda_3]^{\mathrm{T}} \tag{14}$$

联立可解得：

$$\boldsymbol{\lambda}(t) = \begin{bmatrix} -2\alpha \\ 2\alpha t + 2\beta \\ -\alpha t^2 - 2\beta t - 2\gamma \\ \dfrac{1}{3}\alpha t^3 + \beta t^2 + 2\gamma t + \rho \end{bmatrix} \tag{15}$$

其中，α、β、γ、ρ为积分常数。根据极小值条件：

$$H(\boldsymbol{s}^*,u^*,t) = \min_{u\in\Omega} H(\boldsymbol{s},u,t) \tag{16}$$

其中，Ω为控制量集合。将$\boldsymbol{\lambda}(t)$代入$H(\boldsymbol{s},u,t)$，并对u求导，即可得到期望轨迹的最优控制量u^*和最优状态量$\boldsymbol{s}^*$。期望轨迹状态可表示为

$$\boldsymbol{s}^*(t) = \begin{bmatrix} -\dfrac{\alpha}{5\,040}t^7 - \dfrac{\beta}{720}t^6 - \dfrac{\gamma}{120}t^5 - \dfrac{\rho}{48}t^4 + \dfrac{1}{6}j_0 t^3 + \dfrac{1}{2}a_0 t^2 + v_0 t + p_0 \\ -\dfrac{\alpha}{720}t^6 - \dfrac{\beta}{120}t^5 - \dfrac{\gamma}{24}t^4 - \dfrac{\rho}{12}t^3 + \dfrac{1}{2}j_0 t^2 + a_0 t + v_0 \\ -\dfrac{\alpha}{120}t^5 - \dfrac{\beta}{24}t^4 - \dfrac{\gamma}{6}t^3 - \dfrac{\rho}{4}t^2 + j_0 t + a_0 \\ -\dfrac{\alpha}{24}t^4 - \dfrac{\beta}{6}t^3 - \dfrac{\gamma}{2}t^2 - \dfrac{\rho}{2}t + j_0 \end{bmatrix} \tag{17}$$

由于末态时间t_{f}为定值, 可反解出α,β,γ,ρ的值:

$$\begin{bmatrix}\alpha\\ \beta\\ \gamma\\ \rho\end{bmatrix}=\frac{1}{t_{\mathrm{f}}^{7}}\begin{bmatrix}-\dfrac{1}{24t_{\mathrm{f}}^{3}} & -\dfrac{1}{6t_{\mathrm{f}}^{4}} & -\dfrac{1}{2t_{\mathrm{f}}^{5}} & -\dfrac{1}{2t_{\mathrm{f}}^{6}}\\ -\dfrac{1}{120t_{\mathrm{f}}^{2}} & -\dfrac{1}{24t_{\mathrm{f}}^{3}} & -\dfrac{1}{6t_{\mathrm{f}}^{4}} & -\dfrac{1}{4t_{\mathrm{f}}^{5}}\\ -\dfrac{1}{720t_{\mathrm{f}}} & -\dfrac{1}{120t_{\mathrm{f}}^{2}} & -\dfrac{1}{24t_{\mathrm{f}}^{3}} & -\dfrac{1}{12t_{\mathrm{f}}^{4}}\\ -\dfrac{1}{5\,040} & -\dfrac{1}{720t_{\mathrm{f}}} & -\dfrac{1}{120t_{\mathrm{f}}^{2}} & -\dfrac{1}{48t_{\mathrm{f}}^{3}}\end{bmatrix}^{-1}\cdot\begin{bmatrix}\Delta j\\ \Delta a\\ \Delta v\\ \Delta p\end{bmatrix}\tag{18}$$

其中:

$$\begin{bmatrix}\Delta j\\ \Delta a\\ \Delta v\\ \Delta p\end{bmatrix}=\begin{bmatrix}j_{\mathrm{f}}-j_0\\ a_{\mathrm{f}}-a_0-j_0t_{\mathrm{f}}\\ v_{\mathrm{f}}-v_0-\dfrac{1}{2}j_0{t_{\mathrm{f}}}^{2}-a_0t_{\mathrm{f}}\\ p_{\mathrm{f}}-p_0-\dfrac{1}{6}j_0{t_{\mathrm{f}}}^{3}-\dfrac{1}{2}a_0{t_{\mathrm{f}}}^{2}-v_0t_{\mathrm{f}}\end{bmatrix}\tag{19}$$

在对单段轨迹基元的生成过程进行说明后, 可以总结性地将多段轨迹表示为m段n阶的分段多项式函数:

$$p_{\mu}(t)=\begin{cases}\displaystyle\sum_{i=0}^{n}a_{\mu i1}t^{i} & t_0\leqslant t\leqslant t_1\\ \displaystyle\sum_{i=0}^{n}a_{\mu i2}t^{i} & t_1<t\leqslant t_2\\ \quad\vdots & \\ \displaystyle\sum_{i=0}^{n}a_{\mu im}t^{i} & t_{m-1}<t\leqslant t_m\end{cases}\tag{20}$$

对于多段轨迹的规划, 需要最小化的代价函数可以由上述基函数进行表示, 同样也可写出多段轨迹优化过程中的代价函数, 表示为规划时间区间$[t_0,t_m]$内轨迹控制量平方的积分:

$$\min\int_{t_0}^{t_m}\left|\frac{\mathrm{d}^{kp}p_{\mu}}{\mathrm{d}t^{kp}}\right|^{2}\mathrm{d}t\tag{21}$$

其中:

$$\begin{cases}p_{\mu}(t_i)=p_i & i=0,1,\cdots,m\\ \dfrac{\mathrm{d}^{k}p_x}{\mathrm{d}t^{k}}|_{t=t_j}=p_{xj}^{k}\text{或无约束} & \\ \quad j=0,1,\cdots,m;k=1,2,\cdots,kp & \\ \dfrac{\mathrm{d}^{k}p_y}{\mathrm{d}t^{k}}|_{t=t_j}=p_{yj}^{k}\text{或无约束} & \\ \quad j=0,1,\cdots,m;k=1,2,\cdots,kp & \\ \dfrac{\mathrm{d}^{k}p_z}{\mathrm{d}t^{k}}|_{t=t_j}=p_{zj}^{k}\text{或无约束} & \\ \quad j=0,1,\cdots,m;k=1,2,\cdots,kp & \end{cases}\tag{22}$$

多段轨迹规划要求时间区间内的自定义控制量$\dfrac{\mathrm{d}^{kp}p_{\mu}}{\mathrm{d}t^{kp}}$平方尽可能小。同时需要满足式(22)约束关系, 即在各个关键点处的位置量及位置量的各阶导数需要满足设定值$p_{\mu j}^{k}(\mu=x,y,z)$, 在没有设定值的情况下, 可不必进行约束。

上述多段轨迹规划问题可表示为二次规划问题:

$$\begin{cases}\min\boldsymbol{p}^{\mathrm{T}}\boldsymbol{H}\boldsymbol{p}+\boldsymbol{f}^{\mathrm{T}}\boldsymbol{p}\\ \text{s.t.}\qquad \boldsymbol{A}\boldsymbol{p}\leqslant\boldsymbol{b}\end{cases}\tag{23}$$

式中: $\boldsymbol{p}$为$3\times(n+1)\times m\times1$的向量, 对应了待求的$m$段轨迹的$n$阶多项式的$(n+1)$个多项式系数。二次规划问题的目标函数可由式(21)变换得到, 同时, 二次规划问题中的约束式可以用来满足位置量及其导数的相关约束, 从而满足对状态量和控制输入的相关物理约束。对于轨迹上关键点处需要满足设定值的要求, 可以写作约束式中的等式约束; 对于不需精确设定的条件, 可用约束式中的不等式约束表示。

2.2　基于轨迹需求的多约束定义

多约束轨迹规划过程中, 需要根据轨迹状态需求进行等式约束和不等式约束的定义。多约束条件包含了关键点等式约束和轨迹廊道不等式约束, 其中关键点等式约束包含了关键点处的设定值约束及关键点处的多阶连续性约束。

单段轨迹的代价函数为时间区间$[0,t_{\mathrm{f}}]$内控制量平方的积分, 也可写作:

$$J(t_{\mathrm{f}})=\int_0^{t_{\mathrm{f}}}p_{\mu}^{kp}(t)^{2}\mathrm{d}t=\boldsymbol{p}_i^{\mathrm{T}}\boldsymbol{Q}(t_{\mathrm{f}})\boldsymbol{p}_i\tag{24}$$

式中: $\boldsymbol{p}_i$为多项式$p_{\mu}(t)$的$(n+1)$个系数的向量; $\boldsymbol{Q}(t_{\mathrm{f}})$为控制量对$\boldsymbol{p}_i$中各个系数求导后平方且积分所得; t_{f}为固定值。

多段轨迹的m个多项式可以通过将其代价矩阵以块对角的方式串联起来进行联合优化:

$$J_{\mathrm{all}}=\begin{bmatrix}\boldsymbol{p}_i\\ \vdots\\ \boldsymbol{p}_m\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{Q}_1(t_{\mathrm{f}1}) & & \\ & \ddots & \\ & & \boldsymbol{Q}_m(t_{\mathrm{f}m})\end{bmatrix}\begin{bmatrix}\boldsymbol{p}_i\\ \vdots\\ \boldsymbol{p}_m\end{bmatrix}\tag{25}$$

各段轨迹端点上的约束即是关键点的约束, 对于关键点处的设定值约束, 就是使这些端点被固定在空间中已知的位置, 或者对端点处的速度、加速度等进行设定。总轨迹中第i段的设定值约束可以用多项式系数$\boldsymbol{p}_i$和端点导数之间的映射矩阵$\boldsymbol{A}_i$进行表示:

$$\boldsymbol{A}_i\boldsymbol{p}_i=\boldsymbol{d}_i\tag{26}$$

其中:

$$\boldsymbol{A}_i=\begin{bmatrix}\boldsymbol{A}_0\\ \boldsymbol{A}_{t_f}\end{bmatrix}_i,\ \boldsymbol{d}_i=\begin{bmatrix}\boldsymbol{d}_0\\ \boldsymbol{d}_{t_f}\end{bmatrix}_i \tag{27}$$

式中：$\boldsymbol{d}_i$为第i段轨迹的端点导数的设定值向量。通过构造这样的设定值约束，可对规划轨迹关键点处的状态量进行精准设定。

另一方面，考虑到轨迹规划的平滑性，希望能够实现位置量及其多阶导数的连续性，因此，在关键点处需要施加多阶连续性约束。这样的约束主要通过相邻2段轨迹之间的端点，即第i段轨迹的末尾与第$(i+1)$段轨迹开头的多阶导数匹配的形式来实现：

$$\boldsymbol{A}_{t_f,i}\boldsymbol{p}_i=\boldsymbol{A}_{0,i+1}\boldsymbol{p}_{i+1} \tag{28}$$

式(28)关键点处的等式约束可以整合为联合优化问题中线性等式约束集合的形式，表示为

$$\boldsymbol{A}_{\text{all}}\begin{bmatrix}\boldsymbol{p}_1\\ \boldsymbol{p}_2\\ \vdots\\ \boldsymbol{p}_m\end{bmatrix}=\begin{bmatrix}\boldsymbol{d}_1\\ \boldsymbol{d}_2\\ \vdots\\ \boldsymbol{d}_m\end{bmatrix} \tag{29}$$

下面介绍轨迹廊道不等式约束。廊道是对各段轨迹激进程度的限制，当某段轨迹的波动程度过大时，很可能已经超过了四旋翼物理条件下的跟踪范围，这显然是不可取的。在这种情况下，可对轨迹的波动范围进行约束，针对轨迹中位置量及其多阶导数设置合适的廊道，实现对轨迹趋势与轨迹状态量范围的需求设置。

图2为廊道约束的具体形式。首先定义第i段轨迹的端点为$\boldsymbol{p}_{i-1}=\left[x_{i-1},y_{i-1},z_{i-1}\right]^{\mathrm{T}}$和$\boldsymbol{p}_i=\left[x_i,y_i,z_i\right]^{\mathrm{T}}$，并定义$\boldsymbol{e}_i$为端点$\boldsymbol{p}_{i-1}$到$\boldsymbol{p}_i$的单位向量。则第$i$段轨迹上的轨迹点$\boldsymbol{p}(t)$距离端点连线的垂直向量$\boldsymbol{d}_i(t)$可表示为

$$\boldsymbol{d}_i(t)=\left[\boldsymbol{p}(t)-\boldsymbol{p}_{i-1}\right]-\left\{\left[\boldsymbol{p}(t)-\boldsymbol{p}_{i-1}\right]\boldsymbol{e}_i\right\}\boldsymbol{e}_i \tag{30}$$

定义第i段轨迹上的廊道宽度为δ_i，可以垂直向量$\boldsymbol{d}_i(t)$的无穷范数形式将廊道不等式约束表示为

$$\|\boldsymbol{d}_i(t)\|_\infty\leqslant\delta_i\qquad t_{i-1}\leqslant t\leqslant t_i \tag{31}$$

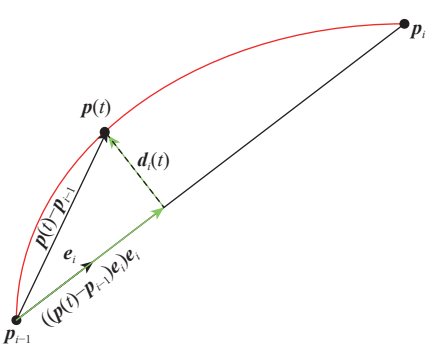


图2 廊道约束物理定义

Fig. 2 Physical definition of corridor constraints

该约束可通过在单段轨迹中引入n_{attach}个中间点来实现，表示为

$$\left|\boldsymbol{d}_{i,\mu}\left(t_{i-1}+\frac{j}{1+n_{\text{attach}}}(t_i-t_{i-1})\right)\right|\leqslant\delta_i$$
$$j=1,2,\cdots,n_{\text{attach}};\mu=x,y,z \tag{32}$$

式(32)表示垂直距离向量在选取的多个中间点时刻需要满足在廊道范围内，由于轨迹的连续性，只要选取合适的中间点，即可对单段轨迹的激进程度进行调整。将式(32)绝对值约束写作2个线性不等式约束，即可转换为二次规划问题中的约束式形式。

2.3 迭代优化的分段时间分配

在多段轨迹规划的过程中，由于代价函数中矩阵构建需要用到各段轨迹的预设时间，因此，各段轨迹的时间是预先进行分配的。尽管这样预先分配时间可能会使得规划的轨迹并不是最优解，但时间分配是可以不断进行修改的，从而对整体的轨迹规划过程进行优化。因此，可以在一开始设定分段时间初值，之后再不断迭代优化这些时间段。采用手动调整分段时间的方法进行优化，同时，为了实现规划轨迹的时间尽可能小，可将时间约束加入到原始的轨迹规划代价函数中。由于在规划过程中，多段轨迹的总时间并没有规定，因此，总时间是可以在轨迹优化的过程中变化的，这样也使轨迹优化能够在总时间最小与控制量平方的积分最小2个需求之间进行动态调整。式(33)给出了加入时间最小罚项后的代价函数表示形式：

$$J_{\text{time}}=\begin{bmatrix}\boldsymbol{p}_i\\ \vdots\\ \boldsymbol{p}_m\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{Q}_1(t_{\text{f}1})&&\\ &\ddots&\\ &&\boldsymbol{Q}_m(t_{\text{f}m})\end{bmatrix}\cdot\begin{bmatrix}\boldsymbol{p}_i\\ \vdots\\ \boldsymbol{p}_m\end{bmatrix}+k_{\text{time}}\sum_{i=1}^{m}t_{\text{f}i} \tag{33}$$

式中：k_{time}为时间约束的权值系数。新的代价函数能够迭代调整全段轨迹的总时间和各段轨迹的时间分配，实现最终轨迹的优化。

3 多参考状态量的四旋翼一体化跟踪控制

3.1 四旋翼无人机系统模型

四旋翼无人机具有4个独立的控制输入：拉力、俯仰力矩、滚转力矩和偏航力矩，其中拉力始终垂直于机身平面。在对四旋翼无人机建模时，可做如下假设：①四旋翼无人机是刚体；②质量和转动惯量不变；③四旋翼无人机几何中心与重心相

同; ④四旋翼无人机仅受重力和旋翼拉力。

图 3 为使用的世界坐标系及机体坐标系的建立, 两者均采用东-北-天坐标系。世界坐标系W由世界坐标系下的标准正交基$\{\boldsymbol{x}_W,\boldsymbol{y}_W,\boldsymbol{z}_W\}$组成, 机体坐标系$B$由世界坐标系下的标准正交基$\{\boldsymbol{x}_B,\boldsymbol{y}_B,\boldsymbol{z}_B\}$组成。机体系的原点与四旋翼无人机重心重合, 并且机体系与四旋翼无人机固连。

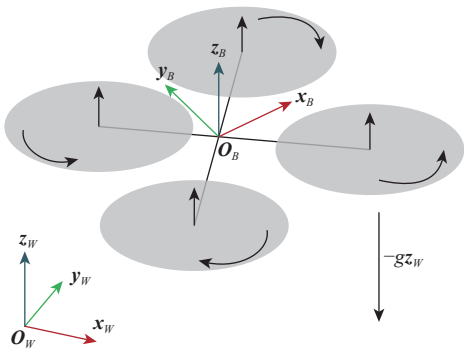


图 3　四旋翼无人机坐标系与系统模型

Fig. 3　Coordinate systems and system model of quadrotor UAVs

定义四旋翼无人机在$-\boldsymbol{z}_W$方向具有大小为g的重力加速度, 且质量为m, 转动惯量为J。四旋翼无人机重心位置表示为$\boldsymbol{p}=[x,y,z]^{\mathrm{T}}$, 位置的一阶导数速度表示为$v$, 位置的二阶导数加速度表示为$\boldsymbol{a}$, 位置的三阶导数加加速度表示为$\boldsymbol{j}$, 位置的四阶导数加加加速度表示为$\boldsymbol{s}$。四旋翼无人机的姿态信息由旋转矩阵$\boldsymbol{R}=[\boldsymbol{x}_B,\boldsymbol{y}_B,\boldsymbol{z}_B]$表示, 机体轴下的角速度表示为$\boldsymbol{\omega}$, 四旋翼所受旋翼拉力$f$作用于重心, 且拉力加速度$c=\dfrac{f}{m}$。

四旋翼无人机的位置运动学模型可表示为

$$\dot{\boldsymbol{p}}=\boldsymbol{v} \tag{34}$$

采用旋转矩阵表示姿态运动学模型为

$$\dot{\boldsymbol{R}}=\boldsymbol{R}\hat{\boldsymbol{\omega}} \tag{35}$$

其中, $\hat{\boldsymbol{\omega}}$为$\boldsymbol{\omega}$的斜对称矩阵, 表示为

$$\hat{\boldsymbol{\omega}}=\begin{bmatrix}0 & -\omega_z & \omega_y\\ \omega_z & 0 & -\omega_x\\ -\omega_y & \omega_x & 0\end{bmatrix} \tag{36}$$

四旋翼无人机的位置动力学模型可表示为

$$\dot{\boldsymbol{v}}=-g\boldsymbol{z}_W+c\boldsymbol{z}_B \tag{37}$$

姿态动力学模型可表示为

$$\boldsymbol{J}\dot{\boldsymbol{\omega}}=\boldsymbol{\tau}-\boldsymbol{\omega}\times\boldsymbol{J}\boldsymbol{\omega}-\boldsymbol{\tau}_{\mathrm{g}} \tag{38}$$

式中: $\boldsymbol{\tau}=[\tau_x,\tau_y,\tau_z]^{\mathrm{T}}$为旋翼产生的三轴力矩输入; $\boldsymbol{\tau}_{\mathrm{g}}$为陀螺力矩。

对于四旋翼无人机的控制分配模型, 由于选用X字形四旋翼, 旋翼产生的总拉力为

$$f=c_T(\varpi_1^2+\varpi_2^2+\varpi_3^2+\varpi_4^2) \tag{39}$$

旋翼产生的力矩为

$$\begin{cases}\tau_x=dc_T\left(\dfrac{\sqrt{2}}{2}\varpi_1^2-\dfrac{\sqrt{2}}{2}\varpi_2^2-\dfrac{\sqrt{2}}{2}\varpi_3^2+\dfrac{\sqrt{2}}{2}\varpi_4^2\right)\\ \tau_y=dc_T\left(\dfrac{\sqrt{2}}{2}\varpi_1^2+\dfrac{\sqrt{2}}{2}\varpi_2^2-\dfrac{\sqrt{2}}{2}\varpi_3^2-\dfrac{\sqrt{2}}{2}\varpi_4^2\right)\\ \tau_z=c_M\left(\varpi_1^2-\varpi_2^2+\varpi_3^2-\varpi_4^2\right)\end{cases} \tag{40}$$

式中: c_T为旋翼拉力系数; d为四旋翼臂长; c_M为旋翼力矩系数; $\varpi_i(i=1,2,3,4)$为旋翼转速。

后续控制方法将依据上述系统模型进行设计, 对上述模型进行变换与简化, 从而使控制律的执行更加方便与有效。

3.2　基于线性二次调节的反馈控制状态跟踪

在设计基于线性二次调节的反馈控制器之前, 需要对四旋翼非线性模型进行形式转换。四旋翼非线性模型中, 位置运动学模型及动力学模型保持不变, 对其中的向量展开表示。

对于姿态运动学及动力学模型, 采用拟合后的姿态控制器模型进行替代。四旋翼的姿态控制器采用 PID 控制器, 假设姿态控制器能够以一阶形式跟踪期望的滚转角和俯仰角, 设定期望偏航角为 0°, 并采用独立的偏航角姿态控制器进行控制保持。对底层姿态控制器的一阶近似使反馈控制器设计过程中能够考虑底层控制器的作用, 从而实现对参考状态量的精准跟踪。姿态控制器的一阶近似参数通过经典的系统辨识技术辨识。经过上述变换及拟合后, 四旋翼的非线性模型可表示为

$$\begin{cases}\dot{\boldsymbol{p}}(t)=\boldsymbol{v}(t)\\ \dot{\boldsymbol{v}}(t)=\boldsymbol{R}\begin{bmatrix}0\\0\\c\end{bmatrix}+\begin{bmatrix}0\\0\\-g\end{bmatrix}\\ \dot{\phi}(t)=\dfrac{1}{\tau_\phi}(K_\phi\phi_{\mathrm{d}}(t)-\phi(t))\\ \dot{\theta}(t)=\dfrac{1}{\tau_\theta}(K_\theta\theta_{\mathrm{d}}(t)-\theta(t))\end{cases} \tag{41}$$

式中: ϕ和θ分别为四旋翼的实际滚转角和实际俯仰角; ϕ_{d}和θ_{d}分别为四旋翼的期望滚转角和期望俯仰角; τ_ϕ和τ_θ分别为姿态控制器一阶系统的滚转时间常数和俯仰时间常数; K_ϕ和K_θ分别为姿态控制器一阶系统的滚转角控制增益和俯仰角控制增益。

对式(41)非线性模型进行线性化, 可得到如下的状态空间方程:

$$\dot{\boldsymbol{x}}(t)=\boldsymbol{A}_{\mathrm{c}}\boldsymbol{x}(t)+\boldsymbol{B}_{\mathrm{c}}\boldsymbol{u}(t)=$$

$$\begin{bmatrix}0&0&0&1&0&0&0&0\\0&0&0&0&1&0&0&0\\0&0&0&0&0&1&0&0\\0&0&0&0&0&0&0&g\\0&0&0&0&0&0&-g&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&\dfrac{1}{\tau_\phi}&0\\0&0&0&0&0&0&0&\dfrac{1}{\tau_\theta}\end{bmatrix}\boldsymbol{x}(t)+$$

$$\begin{bmatrix}0&0&0\\0&0&0\\0&0&0\\0&0&0\\0&0&0\\0&0&1\\\dfrac{K_\phi}{\tau_\phi}&0&0\\0&\dfrac{K_\theta}{\tau_\theta}&0\end{bmatrix}\boldsymbol{u}(t)\tag{42}$$

式中：状态量$\boldsymbol{x}(t)=\left[p_x,p_y,p_z,v_x,v_y,v_z,\phi,\theta\right]^{\mathrm{T}}$，控制量$\boldsymbol{u}(t)=[\phi_{\mathrm{d}},\theta_{\mathrm{d}},c]^{\mathrm{T}}$。对式(42)连续的状态空间方程进行离散化，得到离散状态空间方程的状态矩阵：

$$\boldsymbol{A}=\mathrm{e}^{\boldsymbol{A}_{\mathrm{c}}T_{\mathrm{s}}}\tag{43}$$

控制矩阵为

$$\boldsymbol{B}=\int_0^{T_{\mathrm{s}}}\mathrm{e}^{\boldsymbol{A}_{\mathrm{c}}\tau}\mathrm{d}\tau\boldsymbol{B}_{\mathrm{c}}\tag{44}$$

其中，T_{s}为采样时间。

在给定参考状态$\boldsymbol{x}_{\mathrm{ref},k}$、控制输入$\boldsymbol{u}_k(k=0,1,\cdots,N)$和2个正定权值矩阵$\boldsymbol{Q}$和$\boldsymbol{R}$后，可将有限时域内的二次型代价函数定义为

$$J_{\mathrm{lqr}}=\left(\sum_{k=0}^{N-1}(\boldsymbol{x}_k-\boldsymbol{x}_{\mathrm{ref},k})^{\mathrm{T}}\boldsymbol{Q}(\boldsymbol{x}_k-\boldsymbol{x}_{\mathrm{ref},k})+\boldsymbol{u}_k{}^{\mathrm{T}}\boldsymbol{R}\boldsymbol{u}_k\right)+(\boldsymbol{x}_N-\boldsymbol{x}_{\mathrm{ref},N})^{\mathrm{T}}\boldsymbol{P}(\boldsymbol{x}_N-\boldsymbol{x}_{\mathrm{ref},N})\tag{45}$$

其中，状态量和控制量需要满足以下约束：

$$\begin{cases}\boldsymbol{x}_{k+1}=\boldsymbol{A}\boldsymbol{x}_k+\boldsymbol{B}\boldsymbol{u}_k\\\boldsymbol{u}_k\in\boldsymbol{U}\\\boldsymbol{x}_0=\boldsymbol{x}(t_0)\end{cases}\tag{46}$$

式中：第1个式子表示四旋翼的离散状态空间方程；第2个式子表示四旋翼的控制输入$\boldsymbol{u}_k$需要满足设定的控制约束$\boldsymbol{U}$；第3个式子表示四旋翼的初始状态$\boldsymbol{x}_0$为当前时刻状态$\boldsymbol{x}(t_0)$。

通过求解黎卡缇方程，可得到终止代价函数的系数矩阵$\boldsymbol{P}$及反馈矩阵$\boldsymbol{K}$，黎卡缇方程为

$$\boldsymbol{P}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{A}-\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{B}\left(\boldsymbol{R}+\boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{B}\right)^{-1}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{A}+\boldsymbol{Q}\tag{47}$$

其中反馈矩阵$\boldsymbol{K}$可表示为

$$\boldsymbol{K}=\left(\boldsymbol{R}+\boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{B}\right)^{-1}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{A}\tag{48}$$

因此，反馈控制量为

$$\boldsymbol{u}_k^*=-\boldsymbol{K}\left(\boldsymbol{x}_k-\boldsymbol{x}_{\mathrm{ref},k}\right)\tag{49}$$

反馈控制量$\boldsymbol{u}_k^*$包含了期望滚转角、期望俯仰角及期望拉力加速度。基于线性二次调节的反馈控制状态跟踪如图4所示。首先，将离线规划的参考轨迹转换为由参考位置$\boldsymbol{p}_{\mathrm{ref}}$、参考速度$\boldsymbol{v}_{\mathrm{ref}}$、参考滚转角$\phi_{\mathrm{ref}}$和参考俯仰角$\theta_{\mathrm{ref}}$组成的参考状态量$\boldsymbol{x}_{\mathrm{ref},k}$。其次，LQR反馈控制器接收参考状态量，计算期望的反馈控制量。最后，底层PID姿态控制器接收反馈控制量与参考偏航角ψ，转换为期望四元数，并输出期望拉力与力矩。经过控制分配后得到旋翼期望转速并作用于四旋翼。通过调整权值矩阵$\boldsymbol{Q}$和$\boldsymbol{R}$，可实现对参考状态量的较好跟踪。

当然，也可将参考轨迹转换为参考状态量后，作为前馈控制量与反馈控制量$\boldsymbol{u}_k^*$融合后传递至姿态控制器。

3.3 基于非线性模型预测控制的前馈控制状态跟踪

本节使用的四旋翼非线性模型同样采用姿态控制器一阶近似模型替代了姿态运动学与动力学模型。该替代过程一方面使非线性模型预测控制使用的模型具有准确性，另一方面与直接使用全状态模型相比增加了控制的稳定性。非线性模型预测控制采用的四旋翼模型如式(41)所示。

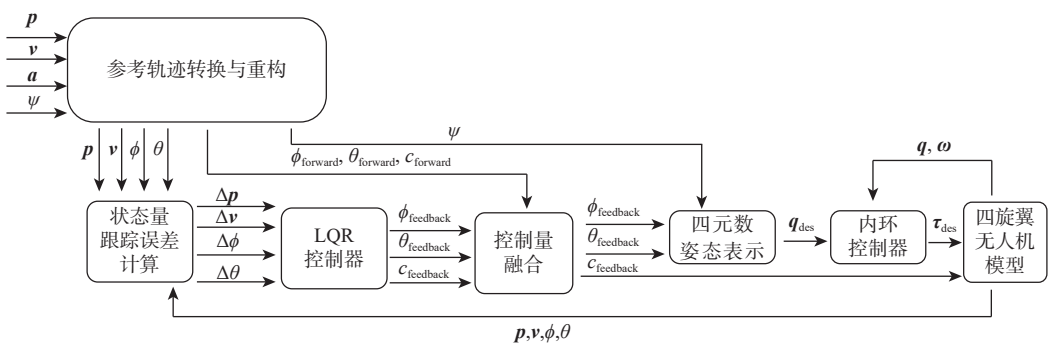


图4 基于线性二次调节的反馈控制状态跟踪

Fig. 4 Feedback control state tracking based on linear quadratic regulation

非线性模型预测控制利用了完整的四旋翼无人机动态模型, 在跟踪激进的轨迹时能够实现更好的性能。非线性模型预测控制的优化问题可表示为

$$\min_{\boldsymbol{U},\boldsymbol{X}} \int_{t=0}^{T} \|h(\boldsymbol{x}(t),\boldsymbol{u}(t))-\boldsymbol{y}_{\text{ref}}\|_{\boldsymbol{Q}}^2 \mathrm{d}t + \left\|h(\boldsymbol{x}(T))-\boldsymbol{y}_{N,\text{ref}}\right\|_{\boldsymbol{Q}_N}^2 \tag{50}$$

式中: 第 1 项为过程代价函数, 第 2 项为终止代价函数; $\boldsymbol{Q}$为状态量及控制量的权值矩阵; $\boldsymbol{Q}_N$为终止状态量的权值矩阵。

$$\begin{cases} \dot{\boldsymbol{x}} = f(\boldsymbol{x}(t),\boldsymbol{u}(t)) \\ \boldsymbol{u}(t) \in \boldsymbol{U} \\ \boldsymbol{x}(t) \in \boldsymbol{X} \\ \boldsymbol{x}(0) = \boldsymbol{x}(t_0) \end{cases} \tag{51}$$

式中: 第 1 个式子为新的四旋翼非线性模型; 第 2 个式子为四旋翼的控制输入$\boldsymbol{u}(t)$需要满足设定的控制约束$\boldsymbol{U}$; 第 3 个式子为四旋翼的状态量$\boldsymbol{x}(t)$需要满足设定的状态约束$\boldsymbol{X}$; 第 4 个式子为四旋翼的初始状态, $\boldsymbol{x}(0)$为当前时刻状态$\boldsymbol{x}(t_0)$。

采用多重打靶方法对四旋翼连续系统在具有时间区间$[t_k,t_{k+1}]$的时间网格$t_0,t_1,\cdots,t_N$上进行离散化, 同时也将不等式约束和控制输入离散到同一时间网格上。对每个区间求解边值问题, 同时也附加了连续性约束。由于四旋翼模型的非线性性质和附加的约束, 该优化问题可看作一个非线性规划问题 (nonlinear programming, NLP)。采用序列二次规划 (sequential quadratic programming, SQP) 方法求解 NLP。

基于非线性模型预测控制的前馈控制器将代价函数中的过程代价函数和终止代价函数具体化, 四旋翼的状态跟踪控制优化问题可表示为

$$\begin{aligned} \min_{\boldsymbol{U},\boldsymbol{X}} \int_{t=0}^{T} & \left((\boldsymbol{x}(t)-\boldsymbol{x}_{\text{ref}}(t))^{\mathrm{T}}\boldsymbol{Q}_{\boldsymbol{x}}(\boldsymbol{x}(t)-\boldsymbol{x}_{\text{ref}}(t))+\right. \\ & (\boldsymbol{u}(t)-\boldsymbol{u}_{\text{ref}}(t))^{\mathrm{T}}\boldsymbol{R}_{\boldsymbol{u}}(\boldsymbol{u}(t)-\boldsymbol{u}_{\text{ref}}(t))\Big)\mathrm{d}t+ \\ & \left((\boldsymbol{x}(T)-\boldsymbol{x}_{\text{ref}}(T))^{\mathrm{T}}\boldsymbol{P}(\boldsymbol{x}(T)-\boldsymbol{x}_{\text{ref}}(T))\right) \end{aligned} \tag{52}$$

式中: 过程代价函数包含了实际状态量和参考状态量之差的平方项、实际控制量和参考控制量之差的平方项; 终止代价函数为实际终止状态量与参考终止状态量之差的平方; $\boldsymbol{Q}_{\boldsymbol{x}}$、$\boldsymbol{R}_{\boldsymbol{u}}$和$\boldsymbol{P}$分别为对应的权值矩阵; T为预测时域。

采用 ACADO 工具箱生成非线性模型预测控制求解器, 该工具箱能够为一般的最优控制问题生成非常快速的自定义 C 代码求解器。对于四旋翼的非线性模型预测控制问题, ACADO 工具箱采用高斯-牛顿海森近似实时迭代求解。通过多重打靶法对连续时间模型离散化。得到的稀疏二次规划问题可以被压缩传递给二次规划求解器 qpOASES, 通过活动集法进行求解。

利用 ACADO 工具箱提供的语法将四旋翼系统动态模型、约束与变量、代价函数描述清楚, 对求解器的相关求解参数合理设置, 即可实时求解基于非线性模型预测控制的前馈控制量。

基于非线性模型预测控制的前馈控制状态跟踪如图 5 所示。首先, 将离线规划的参考轨迹转换为由参考位置$\boldsymbol{p}_{\text{ref}}$、参考速度$\boldsymbol{v}_{\text{ref}}$、参考滚转角$\phi_{\text{ref}}$和参考俯仰角$\theta_{\text{ref}}$组成的参考状态量$\boldsymbol{x}_{\text{ref}}$。其次, NMPC 前馈控制器接收参考状态量, 通过生成的求解器实时求解控制时域内的前馈控制量。最后, 底层 PID 姿态控制器接收前馈控制量与参考偏航角ψ, 转换为期望四元数, 并输出期望拉力与力矩。经过控制分配后得到旋翼期望转速并作用于四旋翼。

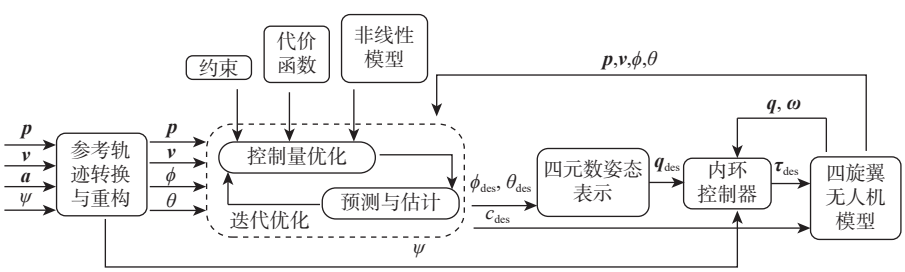


图 5　基于非线性模型预测控制的前馈控制状态跟踪

Fig. 5　Feed-forward control state tracking based on nonlinear model predictive control

3.4　前馈反馈融合的多状态一体化跟踪控制

本文方法是基于线性二次调节的反馈控制方法与基于非线性模型预测控制的前馈控制方法的融合过程。

在控制律设计时加入前馈项能够提升跟踪的实时性和跟踪精度。相比直接使用参考轨迹状态量作为前馈项可能导致的控制过程发散的问题, 以非线性模型预测控制求解的控制量作为前馈项, 能够在提升实时性、准确性的同时保证控制的稳定性。

另一方面，由于姿态控制器的一阶近似模型为线性模型，线性二次调节反馈控制器可直接求解反馈矩阵，实现状态反馈中极点的最优配置。

本文方法的反馈控制项保证了状态跟踪的稳定性，前馈控制项在提升跟踪精度和实时性的同时，也保证了使用前馈项时的稳定性。二者的融合过程能够实现对参考状态量稳定、高实时性、高精度的跟踪过程。

非线性模型预测控制中，设定预测采样时间为$T_{\mathrm{s,nmpc}}$，即控制量更新周期；预测步长为N_{nmpc}，预测周期即为$T_{\mathrm{s,nmpc}}N_{\mathrm{nmpc}}$。而在反馈控制跟踪的线性二次调节中，设定控制采样时间$T_{\mathrm{s,lqr}}=T_{\mathrm{s,nmpc}}$，这样，线性二次调节控制方法求解的控制序列中各项在控制周期上与$T_{\mathrm{s,nmpc}}$相等，从而实现二者的控制量更新周期的同步。

对于同一给定参考轨迹，对轨迹状态量进行高密度离散化，形成参考轨迹状态序列。线性二次调节中的当前参考量即为参考轨迹序列的当前项，非线性模型预测控制中的末端参考量为参考轨迹序列的当前项，将当前状态量到末端参考量之间进行线性拟合后的离散量作为预测时域内的参考状态量序列。由于轨迹状态量的离散化程度较大，且非线性模型预测控制中的线性拟合进一步增加了参考状态量序列的离散密度，因此，非线性模型预测控制与线性二次调节在控制时域内的参考状态量相差较小，能够实现对轨迹参考序列中同一项的同步跟踪控制。图6为LQR与NMPC在控制时间轴上的比较。

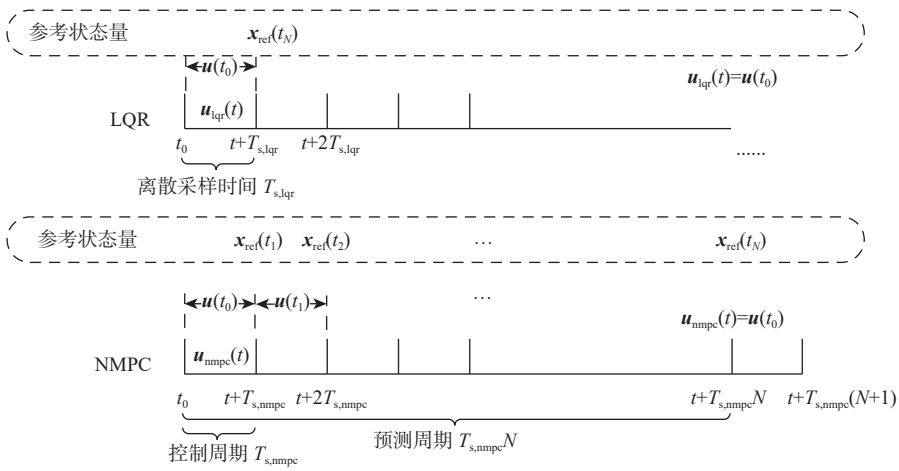


图6 LQR与NMPC的控制时间轴比较

Fig. 6 Control timeline comparison between LQR and NMPC

融合线性二次调节方法与非线性模型预测控制方法后，本文方法的最终控制量$\boldsymbol{u}_{\mathrm{final}}$可表示为

$$\boldsymbol{u}_{\mathrm{final}}=\lambda_{\mathrm{lqr}}\boldsymbol{u}_{\mathrm{lqr}}+\lambda_{\mathrm{nmpc}}\boldsymbol{u}_{\mathrm{nmpc}} \tag{53}$$

式中：λ_{lqr}和λ_{nmpc}分别为线性二次调节控制量和非线性模型预测控制量的权重系数；$\boldsymbol{u}_{\mathrm{lqr}}$和$\boldsymbol{u}_{\mathrm{nmpc}}$分别为线性二次调节控制量和非线性模型预测控制量。最终控制量包含了前馈控制量和反馈控制量2项，$\boldsymbol{u}_{\mathrm{lqr}}$对应了反馈项，$\boldsymbol{u}_{\mathrm{nmpc}}$对应了前馈项。图7为融合

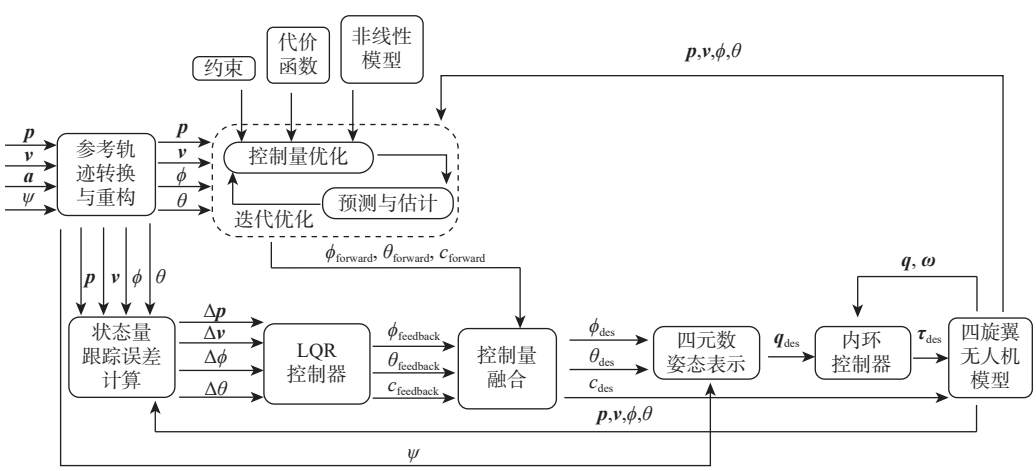


图7 本文方法的总体控制框架

Fig. 7 Overall control framework of proposed method

线性二次调节方法和非线性模型预测控制方法的本文方法的总体控制框架。

本文方法的目的是对包含位置、速度和姿态角的多参考状态量进行跟踪。本文的穿框任务对速度与姿态的跟踪提出了一定的要求, 特别是轨迹中特定位置处的姿态需要满足一定要求。在这样的约束条件下, 本文方法具有更大的优势, 能够实现对参考状态量稳定、高实时性、高精度的跟踪过程。

4　飞行试验与分析

本文方法验证平台为自组装的 F330 四旋翼无人机, 采用 F330 机架, 由妙算 Mainfold 2、Pixhawk v5+飞控、电池、optitrack 定位设备组成。选用的电机为 TMOTOR F80 Pro, 单个电机的最大拉力可达 2.1 kg, 能够实现四旋翼的大推重比。平台的结构如图 8 所示。表 1 为四旋翼无人机的参数。

本文的飞行试验基于 PX4_Firmware 固件的软件在环仿真(SITL)和硬件在环仿真(HITL)。在基于 ROS 操作系统的 Gazebo 仿真平台上进行仿真测试后, 移植软件程序进行飞行试验。本文的一体化跟踪控制器中的姿态环采用的是调参后的 PX4_Firmware 姿态控制器。



图 8　四旋翼无人机验证平台

Fig. 8　Quadcopter UAV verification platform

表 1　四旋翼无人机参数

Table 1　Quadcopter UAV parameters

型号	轴距/mm	电机型号	桨叶型号	总质量/kg
F330	330	TMOTOR F80 Pro	5055三叶桨	1.2

本文的穿框飞行试验是基于离线规划的穿框轨迹, 在线转换为参考状态量, 并采用实时运行的本文方法。最后也将本文方法和常规的前馈 PID 控制方法进行了对比分析。

首先, 对穿越 60°滚转角窄框的轨迹进行规划。轨迹中的控制量选取为加速度量的二阶导数加加加速度, 状态量选取为位置、速度、加速度和加加速度, 初始化的轨迹表达式为 6 阶多项式。表 2 为 60°滚转角穿框轨迹规划过程中的约束设置。

表 2　60°滚转角穿框轨迹规划的约束设置

Table 2　Constraint settings of frame trajectory planning when crossing a 60° roll angle gap

初始点 位置量/m	终止点 位置量/m	穿越 滚转角/(°)	穿越点 速度/(m·s^{-1})	穿越点 位置量/m	穿越点 加速度量/(m·s^{-2})	穿越前缓冲点 z轴位置/m	穿越后缓冲点 z轴位置/m	轨迹段数	轨迹段 时间分配/s
(0,0,1)	330	60	(2.5,0,0)	(3,1,1.5)	(0,−8.313 8,−5)	1.2	1.2	4	4.3,0.7,0.7,4.3

为了实现穿越倾斜窄框的轨迹规划, 可以在规划中设置多个关键点。本文设置了 5 个关键点, 包括起始点、终止点、穿越点、穿越前缓冲点和穿越后缓冲点。其中起始点和终止点均为位置点固定, 位置点多阶导数均为 0。穿越点处的位置和加速度需要与窄框所在位置及倾斜角保持一致。对于穿越前后的缓冲点, 位置点中 z 轴可设置稍低于穿越点, 得到更加平滑的轨迹。穿越前后的缓冲点可以提供短时间内较为平缓的加速度变化过程, 从而提升穿越点处加速度的可达性。

设置初始点为(0,0,1), 终止点为(6,0,1)。穿越点处的滚转角可设置为 10°～70°, 这可以通过设置穿越点的加速度来完成。图 9 和图 10 为 60°滚转角穿框轨迹的各轴位置量及各阶导数曲线和最终轨迹曲线。

设置穿越点位置为(3,1,1.5), 穿越点速度为 2.5 m/s, 穿越前缓冲点和穿越后缓冲点 z 轴均为 1.2 m, 调整轨迹总时间为 10 s, 轨迹中各段的时间分配为 4.3, 0.7, 0.7, 4.3 s。可看到此时轨迹的位置量和速度量较为平滑, 同时加速度量也满足四旋翼物理约束, 规划的轨迹理论可行。

随后, 采用本文方法进行穿越 60°滚转角窄框的飞行试验。表 3 为一体化控制方法的主要参数。

图 11 为本文方法下穿越 60°滚转角窄框时各轴的期望位置和实际位置曲线。重点关注四旋翼到达预设窄框中心位置附近时的位置量和姿态量。

图 12 为穿越 60°滚转角窄框时各轴的实际位置和实际滚转角。当四旋翼滚转角达到 59.55°, 与期望角度仅差 0.45°时, x 轴坐标为 2.885 m, y 轴坐标为 0.627 m, z 轴坐标为 1.435 m。

经过计算, 在达到窄框滚转角时, 实际位置、

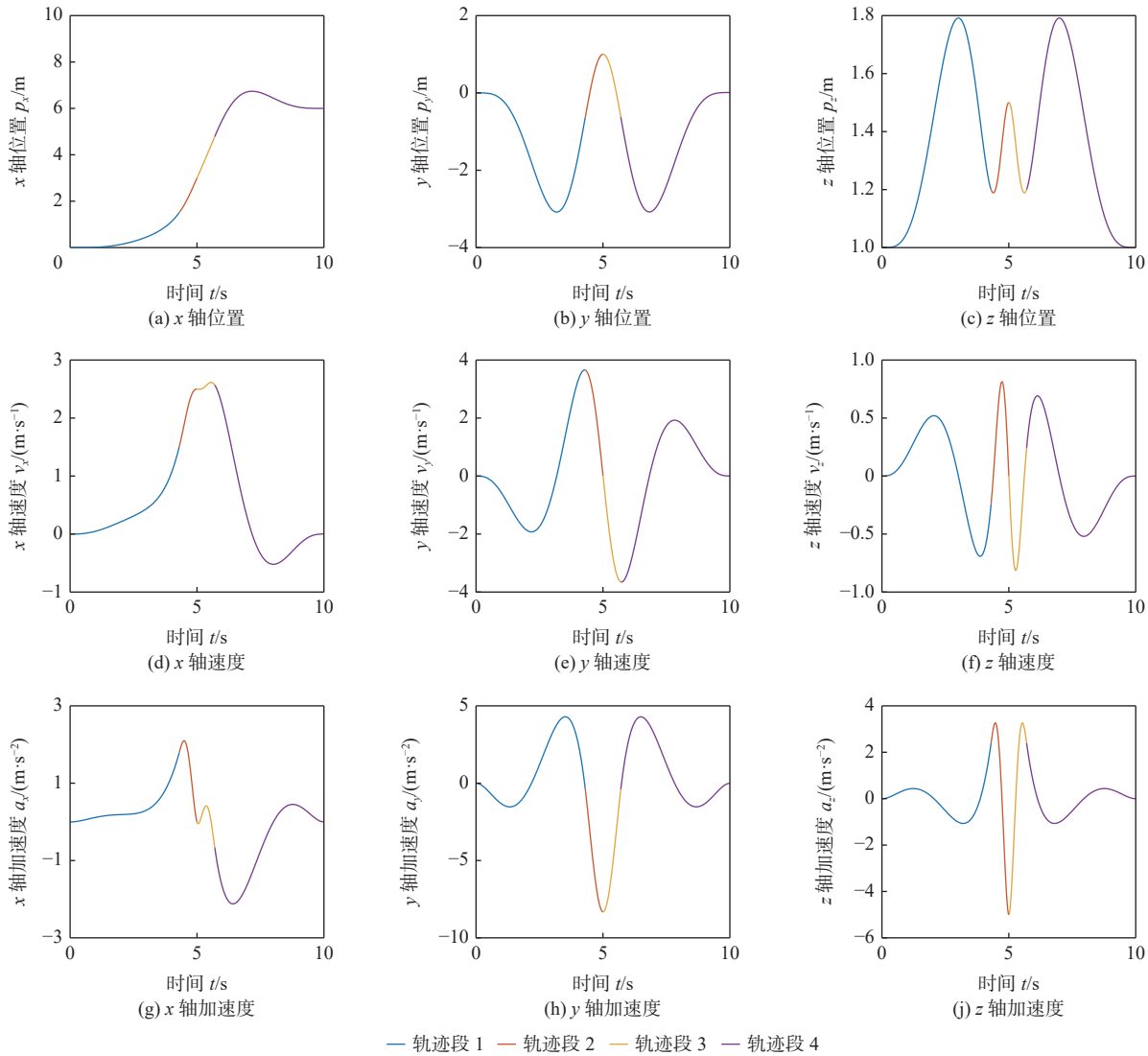


图 9 穿越 60°滚转角窄框时各坐标轴的位置量及各阶导数曲线

Fig. 9 Three-axis positions and derivative curves when crossing 60° roll narrow gap

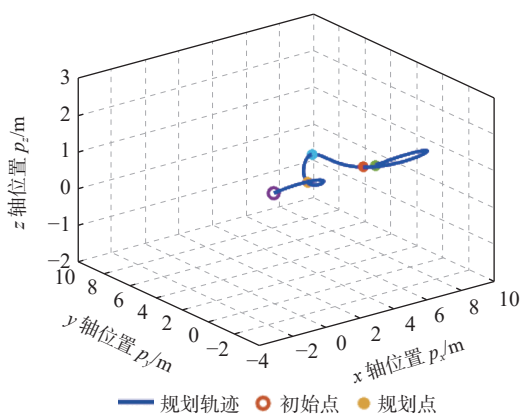


图 10 穿越 60°滚转角窄框的轨迹及关键点

Fig. 10 Trajectory and key points when crossing 60° roll narrow gap

滚转角和期望位置、滚转角的误差分别为：x 轴误差为 0.115 m，y 轴误差为 0.373 m，z 轴误差为 0.065 m，滚转角误差为 0.45°。飞行试验中使用的 Pixhawk v5+飞控采用扩张卡尔曼滤波器（EKF）对内置传感器的测量值进行滤波，从而获取可靠的状态估计值。因此，实际位置、姿态等测量值数据曲线较为平滑。

在实际的 10 多次飞行试验中，四旋翼均能成功穿越 2.5 倍其机体尺寸的窄框。四旋翼到达窄框附近时，几乎与框边缘平行的滚转角使其具备穿框的前提条件，2.5 m/s 的穿越速度保证了整个飞机通过窄框。图 13 为四旋翼穿越 60°窄框的连续过程。

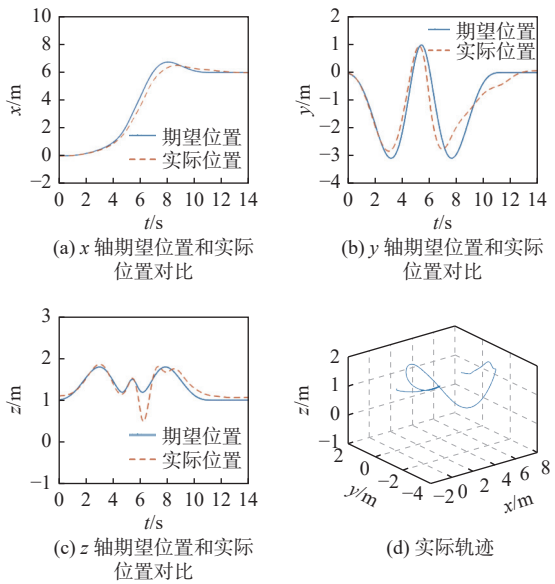
最后，采用前馈 PID 控制方法进行 30°、40°和 60°滚转角穿框轨迹的跟踪飞行试验，并与本文方法对比。其中，穿越速度均为 2.5 m/s。前馈 PID 控制方法的前馈项由参考轨迹直接设定。表 4 为不同滚转角穿框轨迹下前馈 PID 控制方法和本文方法的实际穿越点误差、实际滚转角的对比结果。

对比 2 种方法下实际穿越点的误差和实际滚转角，本文方法相比前馈 PID 控制方法在穿框轨迹

表 3　本文方法主要参数

Table 3　Main parameters of proposed method

位置量权值向量	速度量权值向量	俯仰角权值系数	滚转角权值系数	期望俯仰角权值系数	期望滚转角权值系数	前馈控制项权值系数	反馈控制项权值系数	控制频率/Hz
[300,500,300]	[30,50,20]	50	300	20	100	0.4	0.65	100



(a) x 轴期望位置和实际位置对比　(b) y 轴期望位置和实际位置对比

(c) z 轴期望位置和实际位置对比　(d) 实际轨迹

图 11　穿越 60°滚转角窄框时各轴的期望位置和实际位置曲线

Fig. 11　Three-axis expected and actual positions when crossing 60° roll narrow gap

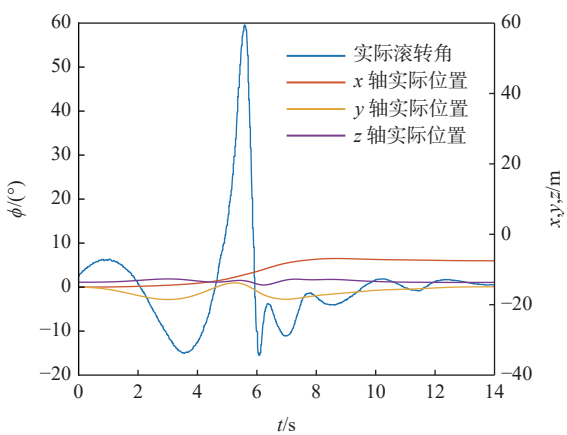


图 12　穿越 60°滚转角窄框时各坐标轴实际位置和实际滚转角

Fig. 12　Three-axis actual positions and actual roll angle when crossing 60° roll narrow gap



图 13　四旋翼穿越 60 °滚转角窄框的连续过程

Fig. 13　Continuous process of a quadrotor crossing 60 ° roll narrow gap

表 4　不同滚转角穿框轨迹下前馈 PID 控制方法和本文方法的实际穿越点对比

Table 4　Comparison of actual crossing points between feed-forward PID control method and proposed method under trajectories with different roll angles

方法	期望滚转角/(°)	实际角度/(°)	x轴误差/m	y轴误差/m	z轴误差/m
前馈PID控制方法	30	28.32	0.767	0.038	0.210
	40	33.68	0.647	0.107	0.231
	60	53.63	0.726	0.499	0.373
本文方法	30	31.31	0.229	0.015	0.007
	40	40.46	0.223	0.125	0.076
	60	59.55	0.115	0.373	0.065

跟踪方面性能更加优越。其中, 跟踪 30°滚转角穿框轨迹时 z 轴误差降低了 96.7%, 且相比前馈 PID 控制方法更容易在穿框轨迹跟踪飞行试验中达到期望滚转角。

5　结　论

1) 本文方法能够实现不同滚转角穿框轨迹的成功规划, 在满足约束设置的同时保证了轨迹的平滑性和连续性;

2) 包含一阶近似姿态控制器的四旋翼无人机非线性模型在用于一体化控制器设计时, 实现了较好的跟踪效果;

3) 本文方法能够实现穿越 60 °大滚转角窄框的飞行过程, 在达到期望滚转角的同时, z 轴跟踪误差仅为 0.065 m;

4) 本文方法相比常规的前馈 PID 控制方法在穿框轨迹跟踪方面性能更加优越, 其中, 三轴误差最大降低了 96.7%, 且更容易达到期望滚转角。

未来可从以下方面进行改善: 考虑四旋翼动态模型的高机动轨迹规划; 飞行验证场景的完善, 增加穿越多个框或俯仰角、滚转角组合框的飞行场景。

参考文献（References）

[1] FAESSLER M, FONTANA F, FORSTER C, et al. Autonomous, vision-based flight and live dense 3D mapping with a quadrotor micro aerial vehicle[J]. Journal of Field Robotics, 2016, 33(4): 431-450.

[2] CASTELLANO G, CASTIELLO C, MENCAR C, et al. Crowd detection in aerial images using spatial graphs and fully-convolutional neural networks[J]. IEEE Access, 2020, 8: 64534-64544.

[3] CUSTERS B. The future of drone use[M]. Hague: TMC Asser Press, 2016: 3-20.
[4] MARTINEZ C, SAMPEDRO C, CHAUHAN A, et al. Towards autonomous detection and tracking of electric towers for aerial power line inspection[C]// 2014 International Conference on Unmanned Aircraft Systems. Piscataway: IEEE Press, 2014: 284-295.
[5] PASCHALL S, ROSE J. Fast, lightweight autonomy through an unknown cluttered environment[C]// 2017 IEEE Aerospace Conference. Piscataway: IEEE Press, 2017: 1-8.
[6] FALANGA D, MUEGGLER E, FAESSLER M, et al. Aggressive quadrotor flight through narrow gaps with onboard sensing and computing using active vision[C]// 2017 IEEE International Conference on Robotics and Automation. Piscataway: IEEE Press, 2017: 5774-5781.
[7] ZHOU B Y, GAO F, WANG L Q, et al. Robust and efficient quadrotor trajectory generation for fast autonomous flight[J]. IEEE Robotics and Automation Letters, 2019, 4(4): 3529-3536.
[8] ZHOU B Y, PAN J, GAO F, et al. RAPTOR: Robust and perception-aware trajectory replanning for quadrotor fast flight[J]. IEEE Transactions on Robotics, 2021, 37(6): 1992-2009.
[9] GAO F, WU W, GAO W L, et al. Flying on point clouds: online trajectory generation and autonomous navigation for quadrotors in cluttered environments[J]. Journal of Field Robotics, 2019, 36(4): 710-733.
[10] QUAN L, HAN L X, ZHOU B Y, et al. Survey of UAV motion planning[J]. IET Cyber-Systems and Robotics, 2020, 2(1): 14-21.
[11] FAESSLER M, FALANGA D, SCARAMUZZA D. Thrust mixing, saturation, and body-rate control for accurate aggressive quadrotor flight[J]. IEEE Robotics and Automation Letters, 2017, 2(2): 476-482.
[12] FOEHN P, SCARAMUZZA D. Onboard state dependent LQR for agile quadrotors[C]// 2018 IEEE International Conference on Robotics and Automation. Piscataway: IEEE Press, 2018: 6566-6572.
[13] FAESSLER M, FRANCHI A, SCARAMUZZA D. Differential flatness of quadrotor dynamics subject to rotor drag for accurate tracking of high-speed trajectories[J]. IEEE Robotics and Automation Letters, 2018, 3(2): 620-626.
[14] FALANGA D, FOEHN P, LU P, et al. PAMPC: Perception-aware model predictive control for quadrotors[C]// 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway: IEEE Press, 2019: 1-8.
[15] SUN S, ROMERO A, FOEHN P, et al. A comparative study of nonlinear MPC and differential-flatness-based control for quadrotor agile flight[J]. IEEE Transactions on Robotics, 2022, 38(6): 3357-3373.
[16] HANOVER D, FOEHN P, SUN S H, et al. Performance, precision, and payloads: Adaptive nonlinear MPC for quadrotors[J]. IEEE Robotics and Automation Letters, 2022, 7(2): 690-697.
[17] KOSTADINOV D, SCARAMUZZA D. Online weight-adaptive nonlinear model predictive control[C]// 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway: IEEE Press, 2021: 1180-1185.
[18] MELLINGER D, KUMAR V. Minimum snap trajectory generation and control for quadrotors[C]// 2011 IEEE International Conference on Robotics and Automation. Piscataway: IEEE Press, 2011: 2520-2525.

Integrated control method for quadrotors' aggressive trajectory tracking under multiple constraints

WANG Yingxun, LI Xin, CAI Zhihao, ZHAO Jiang*

(School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China)

Abstract: The increasing demand for high-dynamic flight of quadrotors has made it an increasingly popular research topic. In order to solve the state tracking control problem of aggressive trajectories when quadrotors undertake activities such as navigating the cracks in the ruins and the gaps in the forest, this work develops an integrated control strategy based on model predictive control. This technique incorporates integrated tracking control of numerous reference states as well as aggressive trajectory planning under multiple limitations. Flight tests have verified the superior performance of the proposed control method in this paper compared with the feed-forward PID control method in tracking the planned aggressive trajectories. In-flight tests, quadrotors successfully crossed the narrow gap of 60° roll angle, and their actual roll angle reached a large angle of 60°, while the z-axis error is only 0.065 m.

Keywords: multiple constraints; integrated control; model predictive control; trajectory tracking; trajectory planning

Received: 2022-04-02; **Accepted**: 2022-06-12; **Published Online**: 2022-06-23 16:13
URL: link.cnki.net/urlid/11.2625.V.20220622.1758.004
* **Corresponding author**. E-mail: jzhao@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0174

基于虚拟协商与滚动时域优化的分布式协同制导策略

杨秀霞*，姜子劼，张毅，王聪，于浩
(海军航空大学，烟台 264001)

摘　　要：针对目标机动条件下的多无人机协同制导问题，提出基于虚拟协商机制与滚动时域优化的分布式协同制导策略。为同时改善时间协调与导引控制效果，在滚动时域优化框架下设计局部协调变量与目标函数，构建有限时域内的多约束协同制导优化模型；为增强制导系统对机动目标的鲁棒性，设计基于非线性自回归神经网络的干扰观测器；针对无人机之间的控制输入耦合，提出虚拟协商机制，提出虚拟协商机制，实现同步决策、排除内部“分歧”，并利用 Nesterov 动量的均方根传播算法设计制导指令生成策略；分别进行数字仿真与半实物仿真实验。结果表明：所提策略可在线预测与优化，应对各类干扰，提升协同制导效果，并会在实际任务中得到进一步的应用与检验。

关　键　词：协同制导；滚动时域优化；神经网络；虚拟协商；均方根传播算法
中图分类号：TJ765.3
文献标志码：A　　**文章编号**：1001-5965(2024)01-0061-16

21世纪以来，无人机(unmanned aerial vehicle, UAV)逐步成为现代战争中的重要力量。然而，随着专用探测技术、电子压制手段等反无人机作战措施的应用与发展，无人机作战面临空前挑战。如何在日益复杂的战场环境下发挥无人机集群作战优势，是当前军事领域的重要研究方向。

多无人机协同制导是引导各无人机同时实施拦截的控制技术，是防止目标逃逸、提升拦截成功率的重要手段。近年来，随着自杀式无人机[1-2]、“忠诚僚机”[3-4]等作战样式相继提出，国内外学者对于多无人机协同制导问题开展了广泛研究。相关成果可按照通信方式、协调方式及控制方式进行划分。

无人机间通信方式分为集中式和分布式2种。集中式通信指至少有一架无人机可与其他全部无人机进行信息交互；分布式通信指所有无人机仅能与其相邻无人机进行信息交互。Zhao 和 Zhou[5]分析了无人机可以获得飞行时间的全局信息与仅能从相邻无人机获取信息2种情况，并基于比例导引法分别设计了集中式协调算法与分散式协调算法；彭琛等[6]从工程实践角度出发，指出了集中式控制结构具有鲁棒性差、通信代价高等固有缺陷；杨剑影等[7]指出，采取分布式通信能够有效避免通信延迟、数据丢包等情况的发生，并且当某个无人机节点出现故障时，也不会由此引发剩余无人机控制失效。因此，分布式协同制导具有较强的鲁棒性，以及更广泛的研究与应用价值。

协调方式的差异来自于协调变量的选择。目前，普遍采用的协调变量包括剩余攻击时间与相对距离[7]。采取剩余攻击时间作为协调变量的制导方法[8-12]大多采用双层协同制导控制架构[8]，即以某种导引律作为底层导引控制，以包含剩余攻击时间误差反馈的控制律作为上层协调控制。张帅等[9]基于有限时间滑模方法与一致性理论分别设计了视线法向与视线方向的加速度指令，使各无人机以期望落角实施拦截，且实现了拦截时间的一致收

收稿日期：2022-03-21；录用日期：2022-06-10；网络出版时间：2022-06-20 10:06
网络出版地址：link.cnki.net/urlid/11.2625.V.20220617.0845.001
基金项目：山东省自然科学基金 (ZR2020MF090)
*通信作者. E-mail：yangxiuxia@126.com

引用格式：杨秀霞，姜子劼，张毅，等. 基于虚拟协商与滚动时域优化的分布式协同制导策略[J]. 北京航空航天大学学报，2024，50(1)：61-76.
YANG X X，JIANG Z J，ZHANG Y，et al. Distributed cooperative guidance strategy based on virtual negotiation and rolling horizon optimization[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(1)：61-76 (in Chinese).

敛；Cho等[11]在制导指令中引入超前角的正连续非线性函数作为附加项，实现了对各无人机碰撞时间的有效控制。上述协调方式的有效性已被广泛证实，但由于在实现方式上依赖于对加速度指令的几何分解[9-10]（否则该类协调方式难以在控制输入饱和条件下同时确保攻击时间与导引误差的精确可控），其控制效果仍具有提升空间。采用相对距离作为协调变量的制导方式[13-15]需要动态地规划参考运动状态，引导无人机保持跟随。邹丽等[13]基于“领-从”结构设计了协同制导律，跟随者根据领导者的方位与速度修正自身姿态，从而实现协同。杨剑影等[7]指出，由于过多地关注一致性误差，该类方法难以确保无人机以足够的精度命中目标，这也是上述制导方式尚未解决的本质性困难。

从控制方式来说，目前，大多数协同制导方法通过预先设计的控制律计算当前控制动作，因此，难以有效应对目标机动等无法预期的各类干扰，而滚动时域优化（rolling horizon control, RHC）通过在线求解有限时域内的开环最优控制获得控制动作[16]，其可以利用预测与反馈，在各类干扰下判断不同控制策略的优劣。因此，对于攻防对抗激烈、状态变换迅速的协同制导问题而言，基于RHC方法设计控制策略能够有效提高系统鲁棒性。此外，目前多数关于协同制导的研究工作均未涉及无人机个体行为的相互影响，以及控制动作选择不当可能引发的群体内部“分歧”。在对无人机群体行为机制进行研究时，段海滨和邱华鑫[17]提出航向协商折中规则，解决了航向信息冲突条件下无人机自主飞行轨迹选择的问题；为消解多智能体任务分配过程中的竞标冲突，文献[18-19]提出了一致性协商机制。以上工作对于研究无人机个体间的协作协商方式具有借鉴意义。

基于此，本文在RHC框架下设计了基于虚拟协商（virtual negotiation, VN）机制的多无人机分布式协同制导策略，能够在目标机动条件下，引导各无人机同时、精准拦截。

1 问题描述与模型建立

1.1 问题描述与模型假设

本文主要研究多个无人机对单个目标的协同制导问题，主要性能判据为全局脱靶量与攻击时间误差，在研究过程中做出如下假设：

假设1 无人机控制系统具有理想动态特性，控制指令传递、执行机构响应无延迟。

假设2 无人机与目标在飞行过程中均可视为质点。

假设3 目标飞行器的位置与速度信息可实时测量。

假设4 制导过程中，通信噪声及延迟对网络连通性的影响忽略不计。

1.2 多无人机协同制导模型

考虑由N架无人机组成的无人机编队，无人机间的通信拓扑关系用有向图$G=(V,E,\boldsymbol{A})$表示。其中，$V=\{v_{\mathrm{UAV}1},v_{\mathrm{UAV}2},\cdots,v_{\mathrm{UAV}N}\}$为无人机集合，$E\subseteq V\times V$为边集，邻接矩阵$\boldsymbol{A}=[a_{ij}]\in\mathbf{R}^{N\times N}$表示无人机间有无通信关系，若边$(v_{\mathrm{UAV}i},v_{\mathrm{UAV}j})\in E$，则表示无人机$i$可以接收来自无人机$j$的信息，称无人机$j$为无人机$i$的相邻无人机或邻节点，此时$a_{ij}=1$，反之$a_{ij}=0$。

无人机i与目标在飞行过程中遵循如下动力学方程：

$$\begin{cases}\dot{x}_{\mathrm{M}i}=v_{\mathrm{M}i}\cos\theta_{\mathrm{M}i}\cos\psi_{\mathrm{M}i}\\ \dot{y}_{\mathrm{M}i}=v_{\mathrm{M}i}\sin\theta_{\mathrm{M}i}\\ \dot{z}_{\mathrm{M}i}=-v_{\mathrm{M}i}\cos\theta_{\mathrm{M}i}\sin\psi_{\mathrm{M}i}\end{cases}\tag{1}$$

$$\begin{cases}\dot{x}_{\mathrm{T}}=v_{\mathrm{T}}\cos\theta_{\mathrm{T}}\cos\psi_{\mathrm{T}}\\ \dot{y}_{\mathrm{T}}=v_{\mathrm{T}}\sin\theta_{\mathrm{T}}\\ \dot{z}_{\mathrm{T}}=-v_{\mathrm{T}}\cos\theta_{\mathrm{T}}\sin\psi_{\mathrm{T}}\end{cases}\tag{2}$$

式中：$(x_{\mathrm{M}i},y_{\mathrm{M}i},z_{\mathrm{M}i})$和$(x_{\mathrm{T}},y_{\mathrm{T}},z_{\mathrm{T}})$分别为无人机$i(i=1,2,\cdots,N)$和目标在惯性坐标系中的坐标；$\theta_{\mathrm{M}i}$和$\theta_{\mathrm{T}}$分别为无人机$i$和目标的航迹倾斜角；$\psi_{\mathrm{M}i}$和$\psi_{\mathrm{T}}$分别为无人机$i$和目标的航迹偏转角；$v_{\mathrm{M}i}$和$v_{\mathrm{T}}$分别为无人机$i$和目标的空速。

无人机i和目标的三维运动示意图如图1所示。

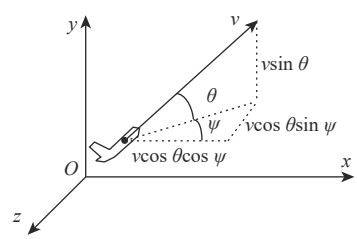


图1 无人机i和目标的三维运动示意图

Fig. 1 Three-dimensional motion diagram of UAVi and target

设控制量：

$$\boldsymbol{u}_i=\begin{bmatrix}u_{1i}\\u_{2i}\\u_{3i}\end{bmatrix}=\begin{bmatrix}\dot{\theta}_{\mathrm{M}i}\\\dot{\psi}_{\mathrm{M}i}\\\dot{v}_{\mathrm{M}i}\end{bmatrix}\tag{3}$$

状态量：

$$\boldsymbol{x}_i=\begin{bmatrix}x_{1i}\\x_{2i}\\x_{3i}\\x_{4i}\\x_{5i}\\x_{6i}\end{bmatrix}=\begin{bmatrix}x_{\mathrm{M}i}-x_{\mathrm{T}}\\y_{\mathrm{M}i}-y_{\mathrm{T}}\\z_{\mathrm{M}i}-z_{\mathrm{T}}\\\dot{x}_{\mathrm{M}i}-\dot{x}_{\mathrm{T}}\\\dot{y}_{\mathrm{M}i}-\dot{y}_{\mathrm{T}}\\\dot{z}_{\mathrm{M}i}-\dot{z}_{\mathrm{T}}\end{bmatrix}\tag{4}$$

建立如下制导系统状态方程：

$$\dot{\boldsymbol{x}}_i = \boldsymbol{A}\boldsymbol{x}_i + \boldsymbol{B}_{1i}a_{1i} + \boldsymbol{B}_{2i}a_{2i} + \boldsymbol{B}_{3i}a_{3i} - \boldsymbol{C}_1\dot{\theta}_{\mathrm{T}} - \boldsymbol{C}_2\dot{\psi}_{\mathrm{T}} - \boldsymbol{C}_3\dot{v}_{\mathrm{T}} \tag{5}$$

式中: $\boldsymbol{A}=\begin{bmatrix}0&0&0&1&0&0\\0&0&0&0&1&0\\0&0&0&0&0&1\\0&0&0&0&0&0\\0&0&0&0&0&0\\0&0&0&0&0&0\end{bmatrix}$; $\boldsymbol{B}_{1i}=\begin{bmatrix}0\\0\\0\\-v_{\mathrm{M}i}\sin\theta_{\mathrm{M}i}\cos\psi_{\mathrm{M}i}\\v_{\mathrm{M}i}\cos\theta_{\mathrm{M}i}\\v_{\mathrm{M}i}\sin\theta_{\mathrm{M}i}\sin\psi_{\mathrm{M}i}\end{bmatrix}$;

$\boldsymbol{B}_{2i}=\begin{bmatrix}0\\0\\0\\-v_{\mathrm{M}i}\cos\theta_{\mathrm{M}i}\sin\psi_{\mathrm{M}i}\\0\\-v_{\mathrm{M}i}\cos\theta_{\mathrm{M}i}\cos\psi_{\mathrm{M}i}\end{bmatrix}$; $\boldsymbol{B}_{3i}=\begin{bmatrix}0\\0\\0\\\cos\theta_{\mathrm{M}i}\cos\psi_{\mathrm{M}i}\\\sin\theta_{\mathrm{M}i}\\-\cos\theta_{\mathrm{M}i}\sin\psi_{\mathrm{M}i}\end{bmatrix}$;

$\boldsymbol{C}_1=\begin{bmatrix}0\\0\\0\\-v_{\mathrm{T}}\sin\theta_{\mathrm{T}}\cos\psi_{\mathrm{T}}\\v_{\mathrm{T}}\cos\theta_{\mathrm{T}}\\v_{\mathrm{T}}\sin\theta_{\mathrm{T}}\sin\psi_{\mathrm{T}}\end{bmatrix}$; $C_2=\begin{bmatrix}0\\0\\0\\-v_{\mathrm{T}}\cos\theta_{\mathrm{T}}\sin\psi_{\mathrm{T}}\\0\\-v_{\mathrm{T}}\cos\theta_{\mathrm{T}}\cos\psi_{\mathrm{T}}\end{bmatrix}$;

$\boldsymbol{C}_3=\begin{bmatrix}0\\0\\0\\\cos\theta_{\mathrm{T}}\cos\psi_{\mathrm{T}}\\\sin\theta_{\mathrm{T}}\\-\cos\theta_{\mathrm{T}}\sin\psi_{\mathrm{T}}\end{bmatrix}$。

设 $t_{\mathrm{f}i}$ 为无人机 i 的制导终止时刻, 则其命中误差可以描述为

$$J_{1i}(t_{\mathrm{f}i}) = \sqrt{x_{1i}^2(t_{\mathrm{f}i}) + x_{2i}^2(t_{\mathrm{f}i}) + x_{3i}^2(t_{\mathrm{f}i})} \tag{6}$$

一致性误差描述为

$$J_{2i}(t_{\mathrm{f}i}) = \left| t_{\mathrm{f}i} - \left(\sum_{j=1}^{N} a_{ij}t_{\mathrm{f}j} \Big/ \gamma_i \right) \right| \tag{7}$$

式中: $\gamma_i = \sum_{j=1}^{N} a_{ij}$ 为无人机 i 的节点入度。

对于 $\forall v_{\mathrm{UAV}i} \in V$, 当 $J_{1i}(t_{\mathrm{f}i}) < \varepsilon$ 且有 $J_{2i}(t_{\mathrm{f}i}) < \varepsilon$ 时, 称各无人机完成了一致性命中, ε 为较小正数。

目前, 多数研究采取如式 (7) 所示的一致性误差设计协调变量。但是, 在目标机动条件下, 对于 $t_{\mathrm{f}i}$ 的估算难度大大提升。对此, 本文在 RHC 框架下设计协同制导策略, 利用实时反馈对估计误差进行校正, 从而适应具有高动态特性的制导系统的实际变化。

2　RHC 框架下目标函数设计与系统干扰估计

2.1　RHC 框架

RHC 方法是一种基于模型的闭环优化控制策略, 其将控制时域等间隔(记为 Δt)地分割为若干采样时刻。在每个时刻 t, 根据实际测量信息与过往系统状态, 对未来一段时间(即预测时域 T_{P}, $T_{\mathrm{P}} = k_{\mathrm{P}}\Delta t$, $k_{\mathrm{P}} \in \mathbf{N}^+$)内的外部干扰进行估计, 通过求解开环优化问题得到该段时间受控系统的最优控制输入序列 $\boldsymbol{u}^*(t), \boldsymbol{u}^*(t+\Delta t), \cdots, \boldsymbol{u}^*(t+(k_{\mathrm{P}}-1)\Delta t)$, 并将所得输入序列的第 1 个分量 $\boldsymbol{u}^*(t)$ 作用于实际系统。当时间步进到下一采样时刻 $t+\Delta t$, 再以新测得的系统状态作为初始条件进行下一轮预测和优化, 并重复该过程, 直到控制结束。

在 RHC 框架下, 求解多无人机协同制导问题就是求解每一个预测时域的开环最优控制。因此, 目标函数的设计是否合理、对系统干扰的预测是否准确决定了最终控制效果。

2.2　协同制导目标函数设计

基于剩余攻击时间协同的制导方法大多采取在底层导引律上叠加协调控制分量的设计形式, 在无人机控制输入饱和的条件下, 这种方式难以同时确保攻击时间与导引误差的精确可控。对此, 本节基于 RHC 的一般性假设, 构建有限时域协同制导目标函数, 对无人机导引控制与时间协调策略进行一体化设计。

引入 RHC 的一般性假设:

假设 5　预测时域以外 ($\tau > t + T_{\mathrm{P}}$), 控制量与系统干扰为 0。

在该假设下, 当 $\tau > t + T_{\mathrm{P}}$, 无人机 i 与目标均可视为零控状态, 此后双方相对距离的平方与时间 τ 的函数关系如式 (8) 所示。根据二次函数极值定理, 得无人机与目标之间的相对距离 $r_i^2(\tau, \boldsymbol{x}_i(t+T_{\mathrm{P}}))$ 在式 (9) 处取极小值, 该点即为假设 5 条件下无人机 i 在预测时域终点处的剩余攻击时间, 记为 $t_{\mathrm{go}i}(\boldsymbol{x}_i(t+T_{\mathrm{P}}))$。

$$\begin{aligned} r_i^2(\tau, \boldsymbol{x}_i(t+T_{\mathrm{P}})) = {} & [x_{1i}(t+T_{\mathrm{P}}) + \tau x_{4i}(t+T_{\mathrm{P}})]^2 + \\ & [x_{2i}(t+T_{\mathrm{P}}) + \tau x_{5i}(t+T_{\mathrm{P}})]^2 + \\ & [x_{3i}(t+T_{\mathrm{P}}) + \tau x_{6i}(t+T_{\mathrm{P}})]^2 \end{aligned} \tag{8}$$

$$\begin{aligned} \tau = {} & -\frac{x_{1i}(t+T_{\mathrm{P}})\,x_{4i}(t+T_{\mathrm{P}})}{x_{4i}^2(t+T_{\mathrm{P}}) + x_{5i}^2(t+T_{\mathrm{P}}) + x_{6i}^2(t+T_{\mathrm{P}})} + \\ & \frac{x_{2i}(t+T_{\mathrm{P}})\,x_{5i}(t+T_{\mathrm{P}})}{x_{4i}^2(t+T_{\mathrm{P}}) + x_{5i}^2(t+T_{\mathrm{P}}) + x_{6i}^2(t+T_{\mathrm{P}})} + \\ & \frac{x_{3i}(t+T_{\mathrm{P}})\,x_{6i}(t+T_{\mathrm{P}})}{x_{4i}^2(t+T_{\mathrm{P}}) + x_{5i}^2(t+T_{\mathrm{P}}) + x_{6i}^2(t+T_{\mathrm{P}})} \end{aligned} \tag{9}$$

根据无人机 i 的通信拓扑关系, 定义局部协调变量平均剩余攻击时间 $\bar{t}_{\mathrm{go}i}(t+T_{\mathrm{P}})$ 为

$$\bar{t}_{\mathrm{go}i}(t+T_{\mathrm{P}}) = \left(\sum_{j=1}^{N} a_{ij} t_{\mathrm{go}j}\left(\boldsymbol{x}_j(t+T_{\mathrm{P}})\right) \right) \Big/ \gamma_i \tag{10}$$

据此定义无人机 i 在预测时域 $[t, t+T_{\mathrm{P}}]$ 内的目标函数为

$$\phi_i(\boldsymbol{x}_i(t+T_\mathrm{P}))=\left[x_{1i}(t+T_\mathrm{P})+\bar{t}_{\mathrm{go}i}(t+T_\mathrm{P})x_{4i}(t+T_\mathrm{P})\right]^2+\left[x_{2i}(t+T_\mathrm{P})+\bar{t}_{\mathrm{go}i}(t+T_\mathrm{P})x_{5i}(t+T_\mathrm{P})\right]^2+\left[x_{3i}(t+T_\mathrm{P})+\bar{t}_{\mathrm{go}i}(t+T_\mathrm{P})x_{6i}(t+T_\mathrm{P})\right]^2 \tag{11}$$

其物理意义为：若无人机 i 与目标均以预测时域终点处的瞬时状态飞行，经过 $\bar{t}_{\mathrm{go}i}(t+T_\mathrm{P})$ 后，双方相对距离的平方。如此设计目标函数的原因如下：

1）在到达时刻 $t_{\mathrm{f}i}$ 前，总会存在某个时刻 $t+T_\mathrm{P}<t_{\mathrm{f}i}$，使得 $t_{\mathrm{f}i}-(t+T_\mathrm{P})\leqslant\Delta t$，此时剩余时间已不足以令无人机 i 再次施加控制，因此，$t+T_\mathrm{P}$ 即为实际停控时刻。由于双方此后均无任何控制输入，$t_{\mathrm{go}i}(\boldsymbol{x}_i(t+T_\mathrm{P}))$ 即为该条件下无人机 i 的实际剩余攻击时间，即

$$t_{\mathrm{f}i}=t+T_\mathrm{P}+t_{\mathrm{go}i}(\boldsymbol{x}_i(t+T_\mathrm{P})) \tag{12}$$

则在 $t_{\mathrm{f}i}$ 无法事先确定的情况下，若某时刻 $t+T_\mathrm{P}$ 满足 $t_{\mathrm{go}i}(\boldsymbol{x}_i(t+T_\mathrm{P}))<\Delta t$，则 $t_{\mathrm{f}i}$ 可根据式 (12) 确定，并且 $r_i^2(\tau,\boldsymbol{x}_i(t+T_\mathrm{P}))$ 和 $t_{\mathrm{go}i}(\boldsymbol{x}_i(t+T_\mathrm{P}))$ 分别为此后双方的真实相对距离和无人机 i 的实际剩余攻击时间。

2）当对于 $\forall v_{\mathrm{UAV}i}\in V$，$t_{\mathrm{go}i}(\boldsymbol{x}_i(t+T_\mathrm{P}))<\Delta t$ 时，若通过优化目标函数使得 $\phi_i(\boldsymbol{x}_i(t+T_\mathrm{P}))<\varepsilon$，由于 $r_i^2(\tau,\boldsymbol{x}_i(t+T_\mathrm{P}))$ 在 $\tau=t_{\mathrm{go}i}(\boldsymbol{x}_i(t+T_\mathrm{P}))$ 处取极小值，并根据式 (12)，可以得到制导终点处的命中误差满足：

$$J_{1i}^2(t_{\mathrm{f}i})=r_i^2\left(t_{\mathrm{go}i}(\boldsymbol{x}_i(t+T_\mathrm{P})),\boldsymbol{x}_i(t+T_\mathrm{P})\right)\leqslant r_i^2\left(\bar{t}_{\mathrm{go}i}(t+T_\mathrm{P}),\boldsymbol{x}_i(t+T_\mathrm{P})\right)=\phi_i(\boldsymbol{x}_i(t+T_\mathrm{P}))<\varepsilon \tag{13}$$

为使式(13)描述更加简洁，证明过程略去时间变量 $(t+T_\mathrm{P})$。

根据二次方程求根公式，当 $\phi_i(\boldsymbol{x}_i)<\varepsilon$ 时，由式 (9) 和式 (11) 可得

$$\left|\bar{t}_{\mathrm{go}i}-t_{\mathrm{go}i}(\boldsymbol{x}_i)\right|<\left[(x_{1i}x_{4i}+x_{2i}x_{5i}+x_{3i}x_{6i})^2-\left(x_{4i}^2+x_{5i}^2+x_{6i}^2\right)\left(x_{1i}^2+x_{2i}^2+x_{3i}^2-\varepsilon\right)\right]^{1/2}\Big/\left(x_{4i}^2+x_{5i}^2+x_{6i}^2\right)$$

进一步地，

$$\left[\bar{t}_{\mathrm{go}i}-t_{\mathrm{go}i}(\boldsymbol{x}_i)\right]^2<(x_{1i}x_{4i}+x_{2i}x_{5i}+x_{3i}x_{6i})^2\Big/\left(x_{4i}^2+x_{5i}^2+x_{6i}^2\right)^2-\left(x_{4i}^2+x_{5i}^2+x_{6i}^2\right)\left(x_{1i}^2+x_{2i}^2+x_{3i}^2-\varepsilon\right)\Big/\left(x_{4i}^2+x_{5i}^2+x_{6i}^2\right)^2=\varepsilon\Big/\left(x_{4i}^2+x_{5i}^2+x_{6i}^2\right)-\left[(x_{1i}x_{5i}-x_{2i}x_{4i})^2+(x_{1i}x_{6i}-x_{3i}x_{4i})^2+(x_{2i}x_{6i}-x_{3i}x_{5i})^2\right]\Big/\left(x_{4i}^2+x_{5i}^2+x_{6i}^2\right)^2<\varepsilon\Big/\left(x_{4i}^2+x_{5i}^2+x_{6i}^2\right)$$

由此得到制导终点处的一致性误差满足：

$$J_{2i}(t_{\mathrm{f}i})=\left|\bar{t}_{\mathrm{go}i}(t+T_\mathrm{P})-t_{\mathrm{go}i}(\boldsymbol{x}_i(t+T_\mathrm{P}))\right|<\frac{\sqrt{\varepsilon}}{\sqrt{x_{4i}^2(t+T_\mathrm{P})+x_{5i}^2(t+T_\mathrm{P})+x_{6i}^2(t+T_\mathrm{P})}} \tag{14}$$

式中：$\sqrt{x_{4i}^2(t+T_\mathrm{P})+x_{5i}^2(t+T_\mathrm{P})+x_{6i}^2(t+T_\mathrm{P})}$ 为非零常数。因此，当优化目标函数使 $\forall v_{\mathrm{UAV}i}\in V$，$\phi_i(\boldsymbol{x}_i(t+T_\mathrm{P}))<\varepsilon$，即可确保各无人机一致性命中。

考虑无人机动力学特性，构建有限时域内的多无人机协同制导优化模型如下：

$$\min J_i=\phi_i(\boldsymbol{x}_i(t+T_\mathrm{P})) \tag{15}$$

s.t.

$$|u_{1i}(\tau)|\leqslant u_{1i\max} \tag{16}$$

$$|u_{2i}(\tau)|\leqslant u_{2i\max} \tag{17}$$

$$|u_{3i}(\tau)|\leqslant u_{3i\max} \tag{18}$$

$$v_{\mathrm{M}i\min}\leqslant v_{\mathrm{M}i}(\tau)\leqslant v_{\mathrm{M}i\max} \tag{19}$$

$$t\leqslant\tau\leqslant t+T_\mathrm{P} \tag{20}$$

式 (16)～式 (18) 为有界控制输入约束；式 (19) 为无人机空速约束；式 (20) 为有限时域约束。至此，本文在 RHC 框架下完成了协同制导目标函数设计及优化模型构建。

2.3 基于自回归神经网络的系统干扰观测器

在 2.2 节所构建的模型中，相邻无人机的控制输入、目标飞行器的机动指令构成了系统干扰。对于此类干扰，传统的控制方法采取在控制回路中串入滤波器的方式设计干扰观测器，而多数滤波算法仅能对当前时刻的系统干扰进行估计，难以在 RHC 框架下有效完成多步预测工作。近年来，随着人工神经网络的发展，机器学习逐渐成为时间序列预测的主流方法[20]，并且衍生出长短期记忆神经网络、卷积神经网络等网络模型，但由于结构复杂，训练用时长，这类模型很难用于对算法实时性要求较高的在线估计任务，相较于此类复杂网络模型，非线性自回归神经网络（nonlinear auto regressive neural network, NARnet）具有结构简单、易于实现等特点，可以对具有时间趋势特征的数据进行在线预测，因此，本文采用 NARnet 设计干扰观测器。

NARnet 源于非线性自回归（nonlinear auto regression, NAR）模型，该模型认为变量在时刻 t 的取值是其特定时间步长内历史状态的非线性函数，即

$$y(t)=f(y(t-\Delta t),y(t-2\Delta t),\cdots,y(t-p\Delta t)) \tag{21}$$

式中：y 为任意时间变量；$p\in\mathbf{N}^+$ 为特定时间步长。

NARnet 可对 NAR 模型进行辨识，得到神经网络模型如下：

$$\hat{y}(t)=g(y(t-\Delta t),y(t-2\Delta t),\cdots,y(t-p\Delta t)|\boldsymbol{\omega}) \tag{22}$$

式中：$\boldsymbol{\omega}$ 为向量化的网络参数。

利用辨识得到的网络模型, 在闭环反馈模式下, 以新的估计值更新网络输入, NARnet 即可实现多步递归估计:

$$\hat{y}(t+k\Delta t)=g\left(\hat{y}(t+k\Delta t-\Delta t),\hat{y}(t+k\Delta t-2\Delta t),\cdots,\right.$$
$$\left.\hat{y}(t+k\Delta t-p\Delta t)|\boldsymbol{\omega}\right)\qquad l=0,1,\cdots,l_{\mathrm{P}}-1 \tag{23}$$

为提高 NARnet 在线预测效率, 对网络结构与训练过程进行如下设计: ①简化网络结构, 以单隐含层网络构建 NARnet; ②限定训练集规模, 仅保留特定数量的历史状态; ③采用 Levenberg-Marquardt(L-M)算法进行网络训练, 提升训练速度。

设当前时刻t, 定义(行)向量$\boldsymbol{y}\in\mathbf{R}^m$用于存储变量$y$的时间序列数据。NARnet 预测过程如下:

步骤 1　更新/初始化向量$\boldsymbol{y}$。若$t=0$, 初始化向量$\boldsymbol{y}=\boldsymbol{0}$; 若$t>1$, 以新的测量值$y(t-\Delta t)$更新向量$\boldsymbol{y}$, 即

$$\boldsymbol{y}\leftarrow\boldsymbol{y}\begin{bmatrix}0 & \cdots & 0 & 0\\ 1 & & & 0\\ & \ddots & & \vdots\\ & & 1 & 0\end{bmatrix}+\begin{bmatrix}0\\ \vdots\\ 0\\ y(t-\Delta t)\end{bmatrix}^{\mathrm{T}} \tag{24}$$

步骤 2　对$\boldsymbol{y}$进行归一化处理。

$$\boldsymbol{y}'=\begin{cases}\dfrac{\boldsymbol{y}-y_{\min}\boldsymbol{l}_{1\times m}}{y_{\max}-y_{\min}} & y_{\max}\neq y_{\min}\\ \boldsymbol{l}_{1\times m} & \text{其他}\end{cases} \tag{25}$$

式中: $y_{\max}$和$y_{\min}$分别为向量$\boldsymbol{y}$中的最大值和最小值; $\boldsymbol{l}_{1\times m}$为$m$维全 1(行)向量。

设$\boldsymbol{y}'=\left[y'_1,y'_2,\cdots,y'_m\right]$, 则网络输入层为

$$\boldsymbol{S}_{\mathrm{in}}=\left[\boldsymbol{s}_{\mathrm{in}\,1},\boldsymbol{s}_{\mathrm{in}\,2},\cdots,\boldsymbol{s}_{\mathrm{in}\,m-p}\right]^{\mathrm{T}} \tag{26}$$

式中: $\boldsymbol{s}_{\mathrm{in}\,k}=\left[y'_k,y'_{k+1},\cdots,y'_{k+p-1}\right]$, $k=1,2,\cdots,m-p$。

网络期望输出为

$$\boldsymbol{s}_{\mathrm{out}}=\left[y'_{p+1},y'_{p+2},\cdots,y'_m\right]^{\mathrm{T}} \tag{27}$$

步骤 3　随机初始化网络参数。随机初始化网络权值$\boldsymbol{W}_1\in\mathbf{R}^{n_{\mathrm{h}}\times p}$、$\boldsymbol{W}_2\in\mathbf{R}^{1\times n_{\mathrm{h}}}$与网络阈值$\boldsymbol{b}_1\in\mathbf{R}^{n_{\mathrm{h}}\times 1}$、$b_2$。其中, n_{h}为隐含层神经元数量。

采用 S 型生长曲线函数 Sigmoid 作为网络激活函数, 则网络前馈过程可描述为

$$\boldsymbol{n}_{\mathrm{out}}=\boldsymbol{W}_2\sigma\left(\boldsymbol{W}_1\boldsymbol{S}_{\mathrm{in}}^{\mathrm{T}}+\boldsymbol{b}_1\boldsymbol{l}_{1\times(m-p)}\right)+b_2\boldsymbol{l}_{1\times(m-p)} \tag{28}$$

式中:

$$\sigma(\boldsymbol{\Lambda})=\left[\frac{1}{1+\mathrm{e}^{-\lambda_{ij}}}\right]$$

其中: $\boldsymbol{\Lambda}=\left[\lambda_{ij}\right]$为任意维数矩阵; $\boldsymbol{n}_{\mathrm{out}}$为网络输出层。

将式 (28) 改写为抽象函数形式:

$$\boldsymbol{n}_{\mathrm{out}}=g\left(\boldsymbol{S}_{\mathrm{in}}^{\mathrm{T}}|\boldsymbol{\omega}\right) \tag{29}$$

网络误差函数采用均方误差函数:

$$L(\boldsymbol{e})=\mathrm{tr}\left(\boldsymbol{e}^{\mathrm{T}}\boldsymbol{e}\right)/2 \tag{30}$$

式中: 误差向量$\boldsymbol{e}=\boldsymbol{s}_{\mathrm{out}}^{\mathrm{T}}-\boldsymbol{n}_{\mathrm{out}}$表示网络期望输出与实际输出的差值。

步骤 4　基于 L-M 算法的网络训练。网络训练即通过误差信息的反向传播调整网络参数, 使误差函数最小的过程, 其实质上是求解最小二乘问题$\min L(\boldsymbol{e})=\mathrm{tr}\left(\boldsymbol{e}^{\mathrm{T}}\boldsymbol{e}\right)/2$。L-M 算法是一类非线性最小二乘算法, 其同时具有梯度法与 Gauss-Newton 法的算法特性。通过引入阻尼项与 Jacobian 矩阵, L-M 算法可以显著改善误差函数的下降速度, 从而克服梯度法收敛过缓及 Gauss-Newton 法易错失最值点等缺陷。因此, 采用 L-M 算法执行网络训练。训练过程描述如下:

步骤 4.1　定义最大训练次数k_0与期望误差ε_0; 初始化训练次数$k'=0$与阻尼因子$\mu=\mu_0$。

步骤 4.2　计算网络误差$L(\boldsymbol{e})$, 若$L(\boldsymbol{e})<\varepsilon_0$, 则终止训练, 输出网络参数, 否则执行步骤 4.3～步骤 4.5。

步骤 4.3　计算误差向量$\boldsymbol{e}$关于参数向量$\boldsymbol{\omega}$的 Jacobian 矩阵, 记为$\boldsymbol{J}_{e|\omega}$。

步骤 4.4　更新网络参数:

$$\boldsymbol{\omega}\leftarrow\boldsymbol{\omega}-\left(\boldsymbol{J}_{e|\omega}^{\mathrm{T}}\boldsymbol{J}_{e|\omega}+\mu\boldsymbol{I}\right)^{-1}\boldsymbol{J}_{e|\omega}^{\mathrm{T}}\boldsymbol{e}^{\mathrm{T}} \tag{31}$$

式中: $\boldsymbol{I}$为单位矩阵。

步骤 4.5　更新训练次数$k'\leftarrow k'+1$, 若$k'=k_0$, 则终止训练, 输出网络参数, 否则执行步骤 4.2。

步骤 5　多步递归估计。利用训练得到的网络模型, 按如下过程进行递归估计:

步骤 5.1　令$k'=1$, 将(归一化处理后的)最近p步测量值$\boldsymbol{y}_{\mathrm{in}}=\left[y'_{m-p+1},y'_{m-p+2},\cdots,y'_m\right]^{\mathrm{T}}$输入网络, 计算网络输出$\hat{y}'_1=g\left(\boldsymbol{y}_{\mathrm{in}}|\boldsymbol{\omega}\right)$。

步骤 5.2　当$k'<k_{\mathrm{P}}$时, 执行步骤 5.3～步骤 5.4, 否则执行步骤 5.5。

步骤 5.3　以估计值$\hat{y}'_k$更新向量$\boldsymbol{y}_{\mathrm{in}}$:

$$\boldsymbol{y}_{\mathrm{in}}=\left[y'_{m-p+k'+1},\cdots,y'_m,\hat{y}'_1,\cdots,\hat{y}'_{k'}\right]^{\mathrm{T}} \tag{32}$$

步骤 5.4　将$\boldsymbol{y}_{\mathrm{in}}$输入网络, 计算网络输出:

$$\hat{y}'_{k+1}=g\left(\boldsymbol{y}_{\mathrm{in}}|\boldsymbol{\omega}\right) \tag{33}$$

并且令$k\leftarrow k+1$, 返回步骤 5.2。

步骤 5.5　对$\hat{\boldsymbol{y}}'=\left[\hat{y}'_1,\hat{y}'_2,\cdots,\hat{y}'_{k_{\mathrm{P}}}\right]$进行反归一化处理, 得到$\hat{\boldsymbol{y}}=\left[\hat{y}_1,\hat{y}_2,\cdots,\hat{y}_{k_{\mathrm{P}}}\right]$, 以及$\hat{y}(t+k\Delta t)=\hat{y}_{k+1}$, $k'=0,1,\cdots,k_{\mathrm{P}}-1$。NARnet 预测完成。

采用上述NARnet设计观测器，利用部分历史数据进行少量训练，即可对预测时域内的系统干扰进行在线估计。本文观测对象包括目标飞行器的航迹倾斜角速度$\dot{\theta}_{\mathrm{T}}(\tau)$、航迹偏转角速度$\dot{\psi}_{\mathrm{T}}(\tau)$及加速度$\dot{v}_{\mathrm{T}}(\tau)$，$\tau=t,t+\Delta t,\cdots,t+(k_{\mathrm{P}}-1)\Delta t$。

此外，由于前一采样时刻(以及较早时刻)的$\dot{\theta}_{\mathrm{T}}$、$\dot{\psi}_{\mathrm{T}}$及$\dot{v}_{\mathrm{T}}$无法直接测量获得，假设上述变量在每一采样时刻保持瞬时恒定，则利用测量信息θ_{T}、ψ_{T}及v_{T}即可对其进行计算。例如：$\dot{\theta}_{\mathrm{T}}(t-\Delta t)=(\theta_{\mathrm{T}}(t)-\theta_{\mathrm{T}}(t-\Delta t))/\Delta t$。

至此，完成了基于NARnet的系统干扰观测器的设计。

3 基于VN机制的制导指令生成策略

3.1 VN机制

在2.2节构建的协同制导模型中，相邻无人机的控制输入会通过局部协调变量影响无人机控制动作的选择。为避免由此类控制输入耦合现象引起系统局部或整体失调，研究者提出基于领导者-跟随者的编队控制策略[13]及基于等级关系的群体行为机制[17]。文献[13]利用领导者的状态信息设计全局协调量，消除了跟随者之间的控制输入耦合，但在全局协调信息缺失的条件下，跟随者无法利用编队中其他成员的信息保持跟随。文献[17]利用群体中的等级关系，使各无人机按等级顺序产生控制动作，但等待前序无人机产生控制动作的时间会随等级增加而累加，在种群规模庞大时，难以维持系统整体的协同高效。

针对上述问题，受多智能体协商思想的启发，本文设计了基于VN机制的协同制导指令生成策略。通过模拟与相邻无人机的协商交互过程，在预测相邻无人机控制输入的同时，实现无人机自身控制动作的优化。在该过程中，无人机以其相邻无人机作为协商对象，完成双方控制输入的交替更新，从而使双方目标函数逐步下降。由于一方控制动作的改变会以协调变量反应在另一方的目标函数中，当一方目标函数值不再下降时，即可认为双方不再具有协商空间，协商结束。产生的控制动作不仅可以支持无人机独立完成制导决策，还尽可能避免了无人机之间产生“分歧”，有利于促进系统整体协调一致。另外，为提高控制动作的在线优化效率，采用Nesterov动量的均方根传播(Nesterov momentum-root mean square propagation, NM-RMSP)算法更新控制输入，对于非凸优化问题具有较好的性能。

与文献[13,17]所提方法相比，基于VN机制的协同制导策略适用于完全分布式编队结构，各无人机可根据自身状态与相邻无人机信息产生控制动作，无需等待其他无人机完成决策，从而保证了决策的时效性。此外，VN机制避免了双方多频次的真实协商交互，减轻了无人机通信负担。

3.2 基于VN与NM-RMSP算法的制导指令生成策略

本节以无人机i为例，简述基于VN机制的制导指令生成过程。无人机i在采样时刻t接收的相邻无人机信息包括位置信息$x_j(t)$、$y_j(t)$、$z_j(t)$，速度信息$\theta_j(t)$、$\psi_j(t)$、$v_j(t)$，节点度γ_j及前一时刻控制输入$\boldsymbol{u}_j(t-\Delta t)$。

3.2.1 确立“协商”对象

在虚拟协商过程中，无人机i无法获知相邻无人机的其他邻节点信息，由此产生了如下2种情况：①无人机i无法判断相邻无人机之间是否存在通信关系；②无人机i无法根据式(10)计算相邻无人机的局部协调变量值。

对此，无人机i在虚拟协商过程中，对相邻无人机的通信协调情况做如下假设：

假设6 相邻无人机之间不进行通信协调。

假设7 相邻无人机与其邻节点(除无人机i以外)能够时刻保持剩余攻击时间的协调一致。

根据假设6和假设7，可对相邻无人机的平均剩余攻击时间进行估计：

$$\hat{\bar{t}}_{\mathrm{go}j}=\left[t_{\mathrm{go}i}(\boldsymbol{x}_i)+(\gamma_j-1)\,t_{\mathrm{go}j}(\boldsymbol{x}_j)\right]/\gamma_j \tag{34}$$

观察$\hat{\bar{t}}_{\mathrm{go}j}$的形式可知，本文在虚拟协商过程中消除了相邻无人机之间的控制输入耦合，于是可以将相邻无人机整体作为无人机i的协商对象。在此基础上，定义新的目标函数：

$$\Phi_i=\sum_{j=1}^{N}\hat{\phi}_j(\boldsymbol{x}_j) \tag{35}$$

式中：$\hat{\phi}_j(\boldsymbol{x}_j)$为基于$\hat{\bar{t}}_{\mathrm{go}j}$计算的邻节点目标函数。

式(35)对多个协商对象进行了整合，形成了无内部冲突的利益群体，避免了多方利益交织而难以协商一致的情况。

3.2.2 基于邻域搜索的控制输入初始化策略

邻域搜索是一类通过搜索当前解的邻域来获得更优解的优化方法。为提升初始控制输入的质量，将领域搜索应用到RHC框架中，即以每一时刻的控制动作作为邻域中心，从该邻域内产生下一时刻初始控制输入序列。该初始化策略具有以下作用：①如果下一时刻的最优控制输入出现在该邻域内，则可显著提高优化速度；②从该邻域出

发, 在有限回合内进行搜索, 避免下一时刻控制动作的大幅改变, 从而减轻控制执行机构的响应负担。

基于该策略, 根据式 (36) 随机产生无人机 i 在当前时刻t的 3 个控制分量的初始序列$\boldsymbol{u}_{1i}^0$、$\boldsymbol{u}_{2i}^0$与$\boldsymbol{u}_{3i}^0$:

$$\begin{cases} \underline{u}_{\lambda i} = \max\left(u_{\lambda i}(t-\Delta t)-\delta, u_{\lambda i\,\min}\right) \\ \bar{u}_{\lambda i} = \min\left(u_{\lambda i}(t-\Delta t)+\delta, u_{\lambda i\,\max}\right) \\ u_{\lambda i}^0(\tau) = \mathrm{rand}\left[\underline{u}_{\lambda i}, \bar{u}_{\lambda i}\right] \\ \boldsymbol{u}_{\lambda i}^0 = \left[u_{\lambda i}^0(t), u_{\lambda i}^0(t+\Delta t), \cdots, u_{\lambda i}^0[t+(k_{\mathrm{P}}-1)\Delta t]\right]^{\mathrm{T}} \\ \quad \tau = t, t+\Delta t, \cdots, t+(k_{\mathrm{P}}-1)\Delta t;\ \ \lambda = 1,2,3 \end{cases} \tag{36}$$

式中: δ为领域半径。

参照式 (36), 产生相邻无人机的 3 个控制分量的初始序列$\boldsymbol{u}_{1j}^0$、$\boldsymbol{u}_{2j}^0$与$\boldsymbol{u}_{3j}^0$, 并将所有相邻无人机的初始序列组成矩阵: $\boldsymbol{U}_{1i}^0 \in \mathbf{R}^{k_{\mathrm{P}}\times\gamma_i}, \boldsymbol{U}_{2i}^0 \in \mathbf{R}^{k_{\mathrm{P}}\times\gamma_i}, \boldsymbol{U}_{3i}^0 \in \mathbf{R}^{k_{\mathrm{P}}\times\gamma_i}$。

3.2.3　基于 NM-RMSP 算法的控制输入更新

为改善目标函数在下降过程中出现振荡现象, 提升收敛速度, 采用 NM-RMSP 算法更新控制输入。RMSP 是梯度法的改进算法, 其利用过往更新过程中的平方梯度调节变量的更新速度, 解决了目标函数在更新过程中振荡幅度过大的问题。Nesterov 动量则在计算梯度前, 对变量进行临时更新, 再根据当前情况进行修正, 对改善振荡同样具有显著作用。因此, 使用 NM-RMSP 算法更新控制输入可以提高目标函数收敛速度, 确保双方快速协商一致。基于 NM-RMSP 算法的控制输入更新过程描述如下:

在更新开始前, 定义全局学习率κ、衰减速率ρ及动量系数α, 并初始化更新速度$\boldsymbol{v}_{\lambda i}^0 = \boldsymbol{0}_{k_{\mathrm{P}}\times 1}$、累积(平方)梯度$\boldsymbol{d}_{\lambda i}^0 = \boldsymbol{0}_{k_{\mathrm{P}}\times 1}$($\lambda = 1,2,3$)。更新开始, 设当前协商次数为$k$($k \geqslant 1$), 计算临时更新:

$$\tilde{\boldsymbol{u}}_{\lambda i}^{k-1} = \boldsymbol{u}_{\lambda i}^{k-1} + \alpha \boldsymbol{v}_{\lambda i}^{k-1} \tag{37}$$

根据式 (11)、式 (9)、式 (5), 利用链式法则计算目标函数值$\phi_i^{k-1}(\boldsymbol{x}_i(t+T_{\mathrm{P}}))$关于$\tilde{\boldsymbol{u}}_{\lambda i}^{k-1}$的梯度:

$$\boldsymbol{g}_{\lambda i}^{k-1} = \nabla_{\tilde{\boldsymbol{u}}_{\lambda i}^{k-1}} \phi_i^{k-1}(\boldsymbol{x}_i(t+T_{\mathrm{P}})) \tag{38}$$

计算累积梯度:

$$\boldsymbol{d}_{\lambda i}^k = \rho \boldsymbol{d}_{\lambda i}^{k-1} + (1-\rho)\, \boldsymbol{g}_{\lambda i}^{k-1} \odot \boldsymbol{g}_{\lambda i}^{k-1} \tag{39}$$

式中: "$\odot$" 为矩阵或向量对应元素相乘的运算符。

计算更新速度:

$$\boldsymbol{v}_{\lambda i}^k = \alpha \boldsymbol{v}_{\lambda i}^{k-1} - \frac{\kappa}{\sqrt{\boldsymbol{d}_{\lambda i}^k}} \odot \boldsymbol{g}_{\lambda i}^{k-1} \tag{40}$$

更新控制输入:

$$\boldsymbol{u}_{\lambda i}^k = \boldsymbol{u}_{\lambda i}^{k-1} + \boldsymbol{v}_{\lambda i}^k \tag{41}$$

相邻无人机控制输入(矩阵)的更新可参照式 (37)～式 (41)。

3.2.4　虚拟协商过程

基于上述工作, 将无人机 i 在采样时刻t的虚拟协商过程描述如下:

步骤 1　初始化协商次数$k=0$、无效协商次数$K=0$; 生成协商双方的初始控制输入$\boldsymbol{u}_{1i}^0$、$\boldsymbol{u}_{2i}^0$、$\boldsymbol{u}_{3i}^0$及$\boldsymbol{U}_{1i}^0$、$\boldsymbol{U}_{2i}^0$、$\boldsymbol{U}_{3i}^0$; 初始化 NM-RMSP 算法参数。

步骤 2　以各无人机当前状态、控制输入及外部干扰观测值, 由t向$t+T_{\mathrm{P}}$积分, 计算$\boldsymbol{x}_i(t+T_{\mathrm{P}})$、$\boldsymbol{x}_j(t+T_{\mathrm{P}})$。

步骤 3　根据式 (9)～式 (11)、式 (34)、式 (35), 计算$\phi_i^0(\boldsymbol{x}_i(t+T_{\mathrm{P}}))$及$\Phi_i^0(t+T_{\mathrm{P}})$。

步骤 4　更新协商次数$k \leftarrow k+1$, 若k大于最大协商次数K_0, 则协商结束, 输出$\boldsymbol{u}_{1i}^k$、$\boldsymbol{u}_{2i}^k$、$\boldsymbol{u}_{3i}^k$, 否则执行步骤 5～步骤 11。

步骤 5　重复步骤 2, 根据式 (9)、式 (10) 计算$t_{\mathrm{go}j}\left(\boldsymbol{x}_j(t+T_{\mathrm{P}})\right)$及$\bar{t}_{\mathrm{go}i}(t+T_{\mathrm{P}})$。

步骤 6　基于 NM-RMSP 算法更新控制输入, 得到$\boldsymbol{u}_{1i}^k$、$\boldsymbol{u}_{2i}^k$、$\boldsymbol{u}_{3i}^k$。

步骤 7　重复步骤 2, 根据式 (9)、式 (34) 计算$t_{\mathrm{go}i}(\boldsymbol{x}_i(t+T_{\mathrm{P}}))$及$\hat{\bar{t}}_{\mathrm{go}j}(t+T_{\mathrm{P}})$。

步骤 8　基于 NM-RMSP 算法更新控制输入, 得到$\boldsymbol{U}_{1i}^k$、$\boldsymbol{U}_{2i}^k$、$\boldsymbol{U}_{3i}^k$。

步骤 9　重复步骤 2 和步骤 3, 计算$\phi_i^k(\boldsymbol{x}_i(t+T_{\mathrm{P}}))$及$\Phi_i^k(t+T_{\mathrm{P}})$。

步骤 10　第k次协商结束, 比较$\phi_i^k(\boldsymbol{x}_i(t+T_{\mathrm{P}}))$与$\phi_i^{k-1}(\boldsymbol{x}_i(t+T_{\mathrm{P}}))$, $\Phi_i^k(t+T_{\mathrm{P}})$与$\Phi_i^{k-1}(t+T_{\mathrm{P}})$, 若$\left(\phi_i^{k-1}(\boldsymbol{x}_i\cdot(t+T_{\mathrm{P}})) - \phi_i^k(\boldsymbol{x}_i(t+T_{\mathrm{P}}))\right) \big/ \phi_i^{k-1}(\boldsymbol{x}_i(t+T_{\mathrm{P}})) < \beta_0$　或$\left(\Phi_i^{k-1}(t+T_{\mathrm{P}}) - \Phi_i^k(t+T_{\mathrm{P}})\right) \big/ \Phi_i^{k-1}(t+T_{\mathrm{P}}) < \beta_0$, 则 $K \leftarrow K+1$, 表明双方本次协商无效, 否则$K \leftarrow 0$, 其中, β_0为较小正数。

步骤 11　若K大于最大无效协商次数K_0, 可以认为双方目标函数无明显下降趋势, 则协商结束, 输出$\boldsymbol{u}_{1i}^k$、$\boldsymbol{u}_{2i}^k$、$\boldsymbol{u}_{3i}^k$, 否则执行步骤 4。

分别取$\boldsymbol{u}_{1i}^k$、$\boldsymbol{u}_{2i}^k$、$\boldsymbol{u}_{3i}^k$的第 1 个分量$u_{1i}^k(t)$、$u_{2i}^k(t)$、$u_{3i}^k(t)$, 则$u_{1i}^k(t)$、$u_{2i}^k(t)$、$u_{3i}^k(t)$即为无人机 i 当前时刻的制导指令: $\boldsymbol{u}_i(t) = \begin{bmatrix} u_{1i}^k(t) \\ u_{2i}^k(t) \\ u_{3i}^k(t) \end{bmatrix}$。

至此, 完成了基于 VN 机制的制导指令生成策略的设计。为直观描述基于 VN-RHC 的多无人机分布式协同制导过程, 建立系统流程, 如图 2 所示。

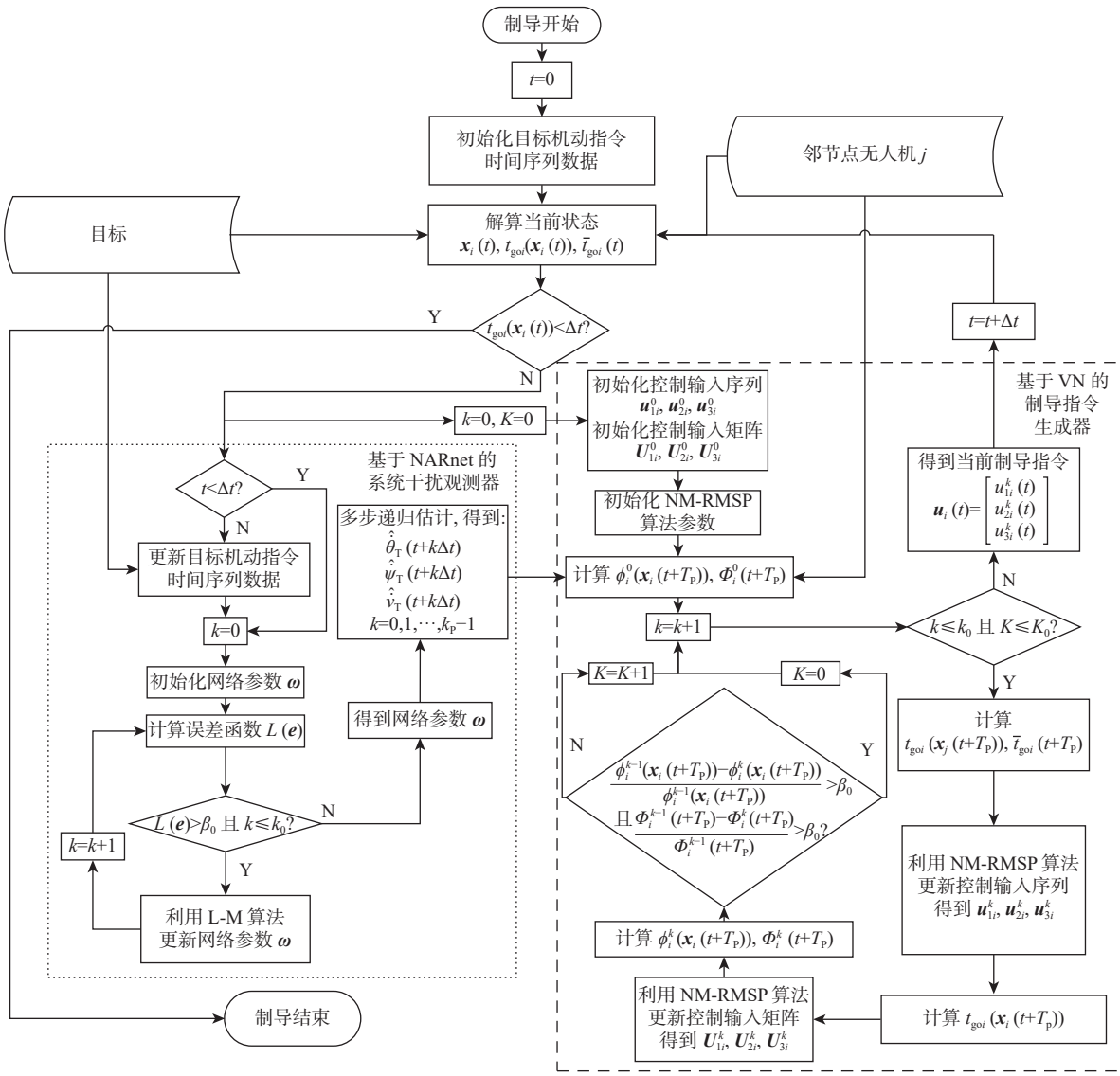


图 2 基于 RHC 与 VN 的分布式协同制导流程

Fig. 2 Flow of distributed cooperative guidance based on RHC and VN

4 仿真实验

为验证 VN-RHC 制导策略的有效性，分别进行数字仿真实验与半实物仿真实验。

4.1 数字仿真实验

在数字仿真实验中，利用 3 种目标机动形式检验 VN-RHC 制导策略的实际效果，并选择基于增广比例导引律和一致性理论（augmented proportional navigation-consensus control, APN-CC）的分布式协同制导律[21]进行对比实验。3 种目标机动形式包括：①纵向随机常值机动、侧向随机正弦机动；②纵向随机方波机动、侧向随机常值机动；③纵向随机正弦机动、侧向随机方波机动。

假设制导过程中，目标恒速飞行。选取全局制导终点下的脱靶量与攻击时间误差作为协同制导性能指标，其中，全局制导终点 $\bar{t}_f=\sum_{i=1}^{N}t_{fi}/N$。目标初始状态如表 1 所示，各无人机及目标性能参数如表 2 所示，无人机编队的通信拓扑结构及初始状态如图 3 所示。

取采样间隔 $\Delta t=0.1$ s，预测时域 $T_P=0.5$ s，仿真得到 2 种制导策略下无人机与目标三维运动轨迹，如图 4 所示，无人机与目标相对距离变化情况如图 5 所示，各无人机剩余攻击时间与平均剩余攻击时间之差的变化情况如图 6 所示，在 VN-RHC 协同制导策略下，各无人机制导指令的变化情况如图 7 所示。

表 1 目标初始状态

Table 1 Target initial state

x/m	y/m	z/m	航迹倾斜角/rad	航迹偏转角/rad	空速/(m·s^{-1})
1 500	800	1 500	0	3π/4	20

表 2　无人机及目标性能参数

Table 2　Performance parameters of UAVs and target

无人机/目标	最大航迹倾斜角速度/(rad·s^{-1})	最大航迹偏转角速度/(rad·s^{-1})	最大加速度/(m·s^{-2})	最大空速/(m·s^{-1})	最小空速/(m·s^{-1})
无人机1	0.05	0.05	1	30	20
无人机2	0.05	0.05	1	30	20
无人机3	0.05	0.05	1	30	20
无人机4	0.05	0.05	1	30	20
目标	0.05	0.05			

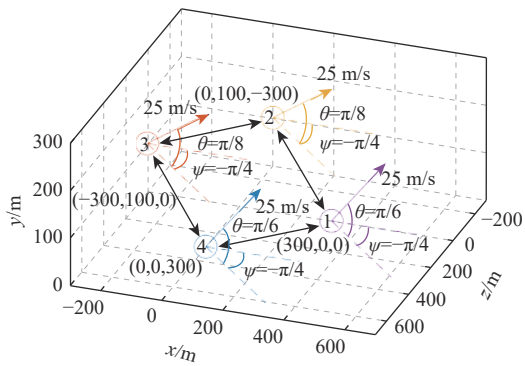


图 3　各无人机通信拓扑结构及初始状态示意图

Fig. 3　Schematic diagram of communication topology and initial state of each UAV

由图 4 可以看出, 在 2 种制导策略下, 各无人机均能完成协同拦截, 且飞行轨迹平滑。由图 5 可以看出, 在全局制导终点处, VN-RHC 制导策略针对 3 种目标机动形式均能获得更小的脱靶量。由图 6 可以看出, 在 APN-CC 制导策略下, 各无人机攻击时间误差虽能较快收敛, 但 VN-RHC 制导策略最终获得了更好的收敛效果。

在 2 种协同制导策略下, 针对 3 种目标机动形式的总脱靶量与总攻击时间误差如表 3 所示。

对比结果表明, VN-RHC 协同制导策略提升了无人机导引精度, 取得了更好的时间协调效果。另外, 由图 7 可以看出, 对采取上述 3 种机动形式的目标进行拦截时, 各无人机制导指令在短时间内均未发生较大变化, 说明 VN-RHC 制导策略有效抑制了制导过程中飞行角度与速度的大幅变动, 会改善执行机构响应延迟等实际控制问题。

为进一步检验 VN 制导策略与 NARnet 干扰观测器的计算效率, 分别取最大训练次数为 50, 最大协商次数为 50, 在上述 3 种工况下仿真得到各时刻 NARnet 训练次数, 如图 8 所示, 各无人机协商次数如图 9 所示。在采样时刻t对$t+k\Delta t(0\leqslant k<k_{\mathrm{P}}=5)$时刻的系统干扰进行预测, 预测误差如图 10 所示。

由图 5 与图 10 可知, 经过训练的 NARnet 干扰观测器可以准确预测有限步长内的系统干扰, 并最终提升各无人机的制导精度。由图 8 与图 10 可知, 对于常值机动与方波机动, NARnet 可以在较短时间内对目标机动指令的改变做出响应, 对于正弦机动, NARnet 虽然需要多次训练趋近目标机动指令的变化趋势, 但总体训练耗时较短(见表 4), 说明通过对网络结构与训练过程的合理优化, 提升了 NARnet 的预测效率。由图 9 可知, 各无人机仅需少数几次协商, 即可产生当前控制动作, 说明基于 NM-RMSP 算法的更新策略与基于邻域搜索的初始化策略提升了协商速度, 缩短了无人机决策耗时(见表 4)。由图 5 和图 6 可知, VN 机制促进了系统整体的决策进程, 且提升了各无人机的协同拦截效果。

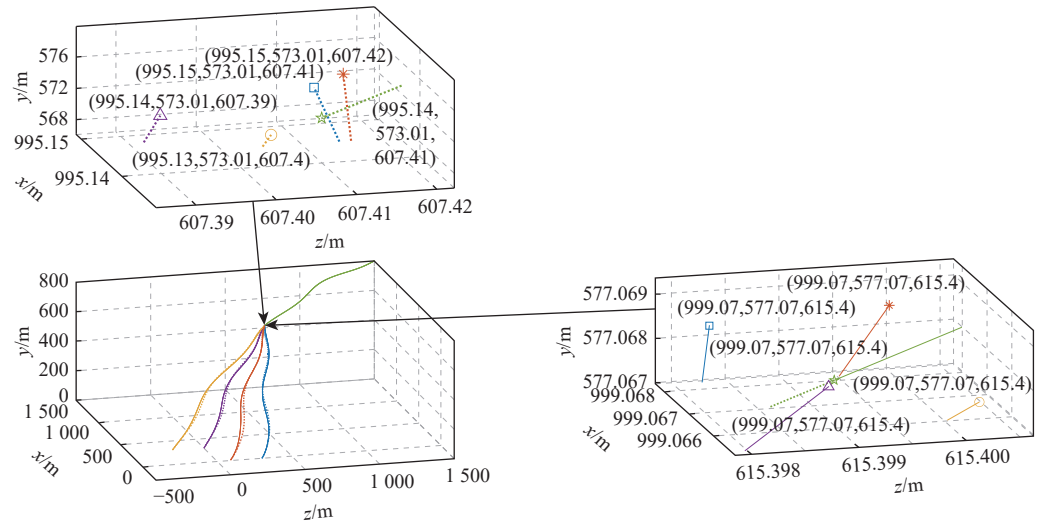
针对上述 3 种工况, 选定不同的算法参数(如 NARnet 最大单步训练次数、无人机最大单步协商次数等)进行多次仿真, 得到算法用时与制导效果的相关情况, 如表 4 所示。仿真硬件参数为: Intel(R) Core(TM) i7-10750H CPU @ 2.60 GHz, 16 GB RAM。

由表 4 可知, 在多次仿真实验中, 无人机平均单步决策时长维持在 0.02 s 以内, 表明本文设计的 VN-RHC 制导策略提升了制导指令在线优化速度, 可以满足算法实时性的要求; 此外, 最大单步协商次数对于制导效果与算法用时的影响更为显著, 在上述仿真环境中, 最大协商次数选择 50 最佳, 而对于初速较大、航程较远的无人机制导问题及其他任务设定, 具体参数取值则有待进一步探究。以上数字仿真实验结果表明, VN-RHC 制导策略可以通过高效地在线预测与优化, 有效引导无人机实施协同拦截。

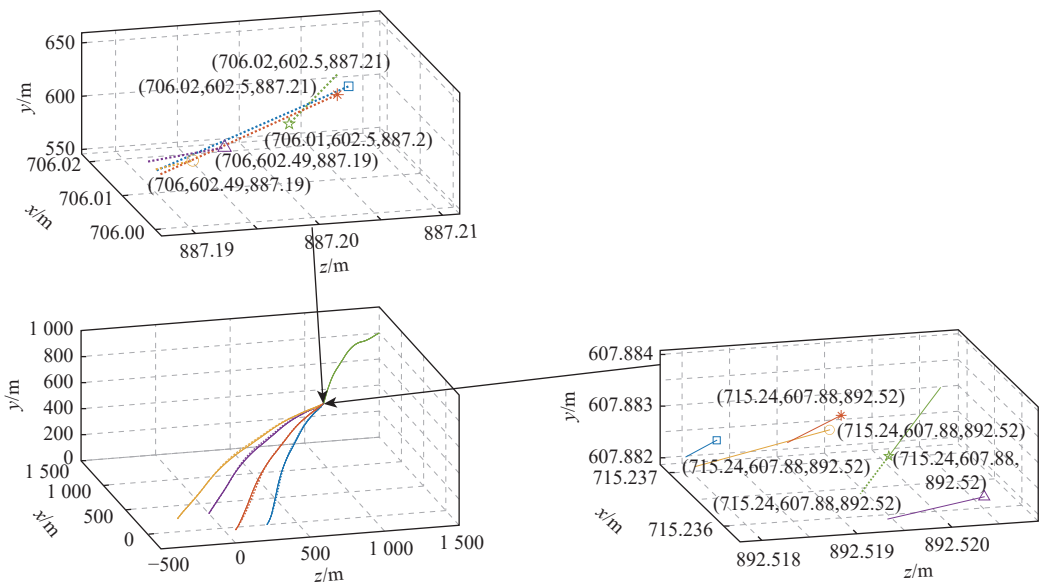
4.2　半实物仿真实验

为检验 VN-RHC 协同制导策略在真实场景中的表现, 基于 RflySim 实验平台开展半实物仿真实验。该平台闭环仿真构架如图 11 所示。将 Pixhawk 自动驾驶仪串入仿真回路, 仿真系统构成如图 12 所示。其中, Pixhawk 可根据 VN-RHC 制导策略解算的制导指令产生 PWM 控制信号, CopterSim 则根据控制信号运行无人机运动动态模型, 对无人机位姿进行解算。计算过程考虑了空气阻力、旋转阻尼及无人机机身结构等物理参数, 因而可以模拟无人机飞行的真实状态。在通信实现方面, 该平台利用 UDP 网络进行虚拟组网, 因此, 通过增加 UDP 网络的输出链路分支即可实现无人机节点间的分布式通信。

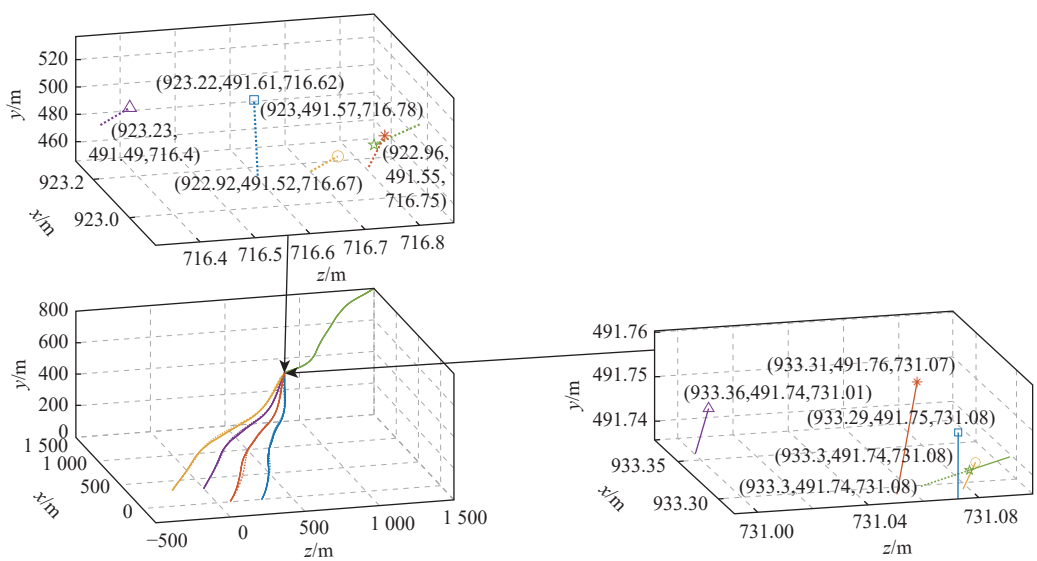
基于该半实物仿真系统, 本文设定的实验场景为 3 架小型无人机对目标实施合围拦截, 各无人机



(a) 目标纵向随机常值机动、侧向随机正弦机动



(b) 目标纵向随机方波机动、侧向随机常值机动

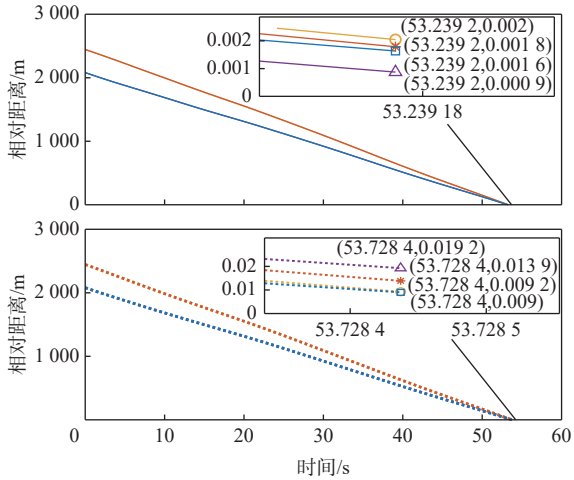


(c) 目标纵向随机正弦机动、侧向随机方波机动

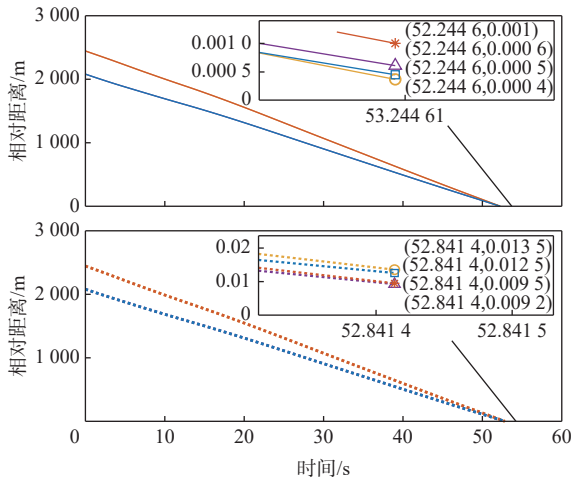
无人机 1, VN-RHC　无人机 2, VN-RHC　无人机 3, VN-RHC　无人机 4, VN-RHC　无人机 1, APN-CC
无人机 2, APN-CC　无人机 3, APN-CC　无人机 4, APN-CC　目标

图 4　无人机-目标三维运动轨迹

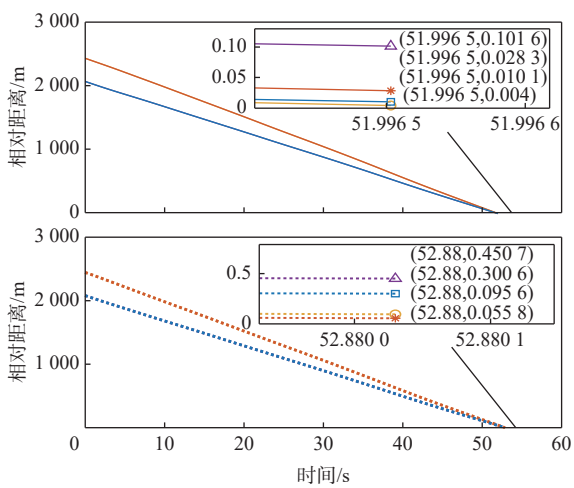
Fig. 4　Three-dimensional motion trajectory of UAVs and target



(a) 目标纵向随机常值机动、侧向随机正弦机动



(b) 目标纵向随机方波机动、侧向随机常值机动



(c) 目标纵向随机正弦机动、侧向随机方波机动

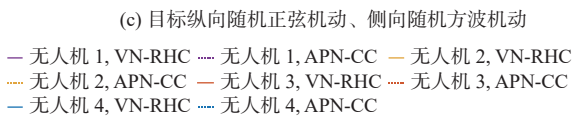
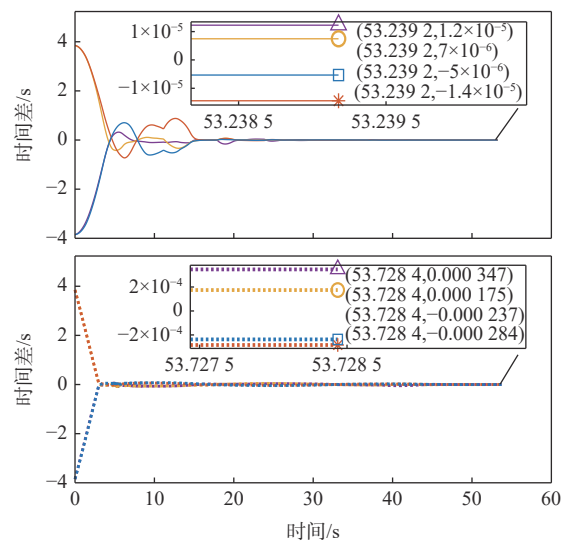


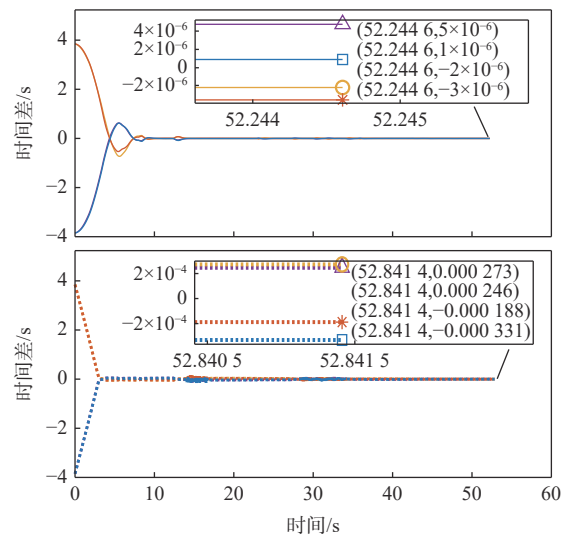
图 5　无人机-目标相对距离变化情况

Fig. 5　Variation of UAV-target relative distance

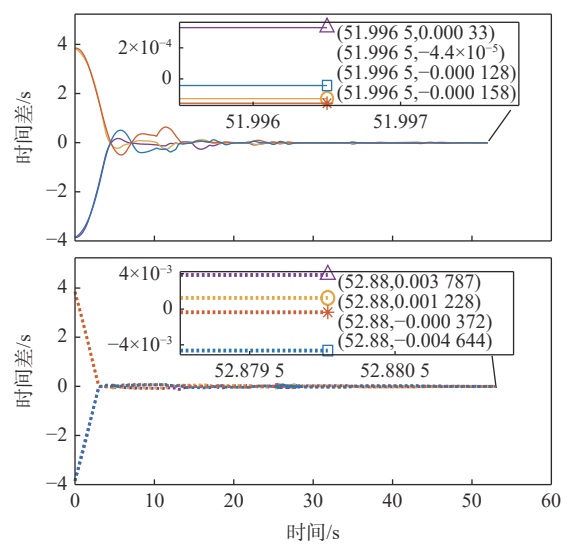
与目标的初始状态及无人机间通信拓扑结构如图 13 所示, 其中, 载具 1 为目标, 载具 2 为无人机 3, 载具 3 为无人机 1, 载具 4 为无人机 2。



(a) 目标纵向随机常值机动、侧向随机正弦机动



(b) 目标纵向随机方波机动、侧向随机常值机动



(c) 目标纵向随机正弦机动、侧向随机方波机动

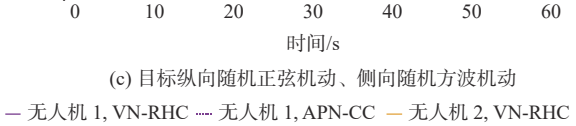
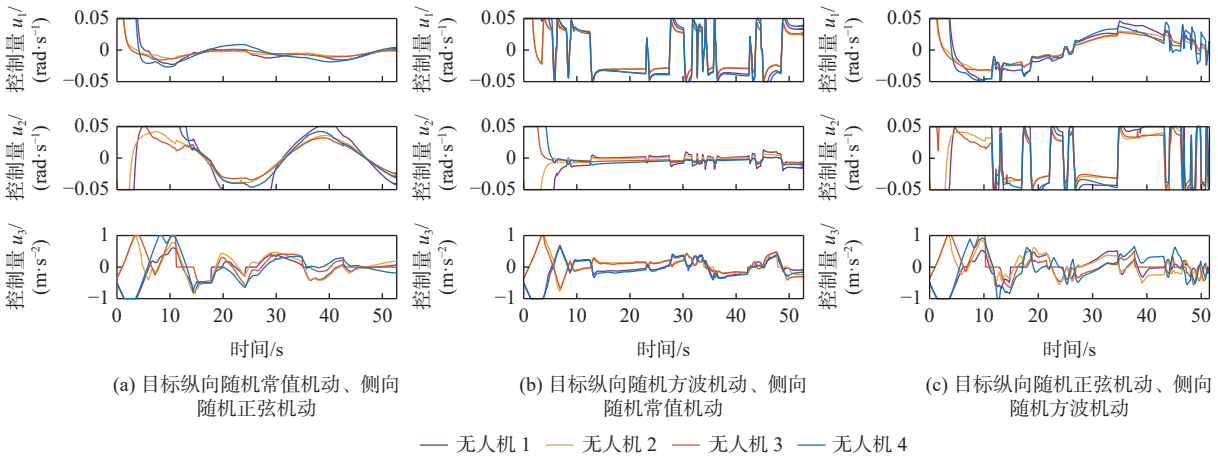


图 6　剩余攻击时间与平均剩余攻击时间之差变化情况

Fig. 6　Variation of difference between time to go and average



(a) 目标纵向随机常值机动、侧向随机正弦机动　(b) 目标纵向随机方波机动、侧向随机常值机动　(c) 目标纵向随机正弦机动、侧向随机方波机动

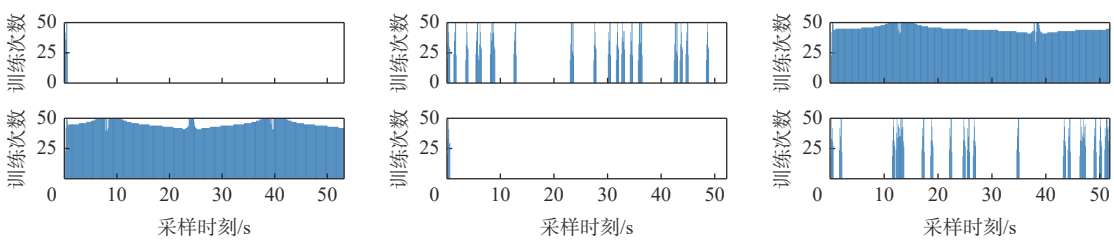
图 7　VN-RHC 协同制导策略下无人机制导指令变化情况

Fig. 7　Variation of UAV guidance command under VN-RHC cooperative guidance strategy

表 3　2 种协同制导策略的制导效果

Table 3　Guidance performance of two cooperative guidance strategies

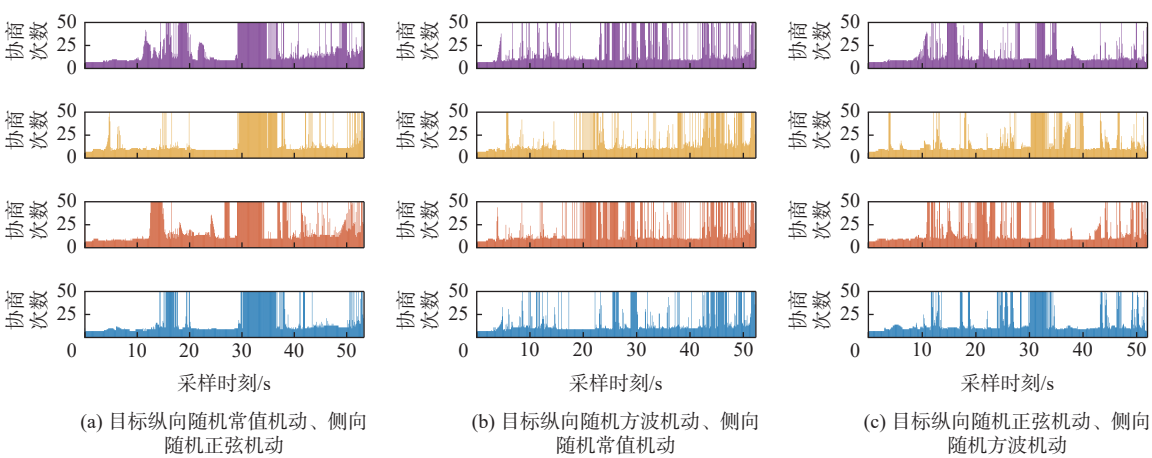
制导策略	总脱靶量/m			总攻击时间误差/s		
	纵向随机常值机动、侧向随机正弦机动	纵向随机方波机动、侧向随机常值机动	纵向随机正弦机动、侧向随机方波机动	纵向随机常值机动、侧向随机正弦机动	纵向随机方波机动、侧向随机常值机动	纵向随机正弦机动、侧向随机方波机动
VN-RHC	0.006 3	0.002 4	0.143 9	3.936×10^{-5}	1.127×10^{-5}	6.593×10^{-4}
APN-CC	0.051 3	0.044 7	0.902 7	0.001	0.001	0.01



(a) 目标纵向随机常值机动、侧向随机正弦机动　(b) 目标纵向随机方波机动、侧向随机常值机动　(c) 目标纵向随机正弦机动、侧向随机方波机动

图 8　各时刻 NARnet 训练次数

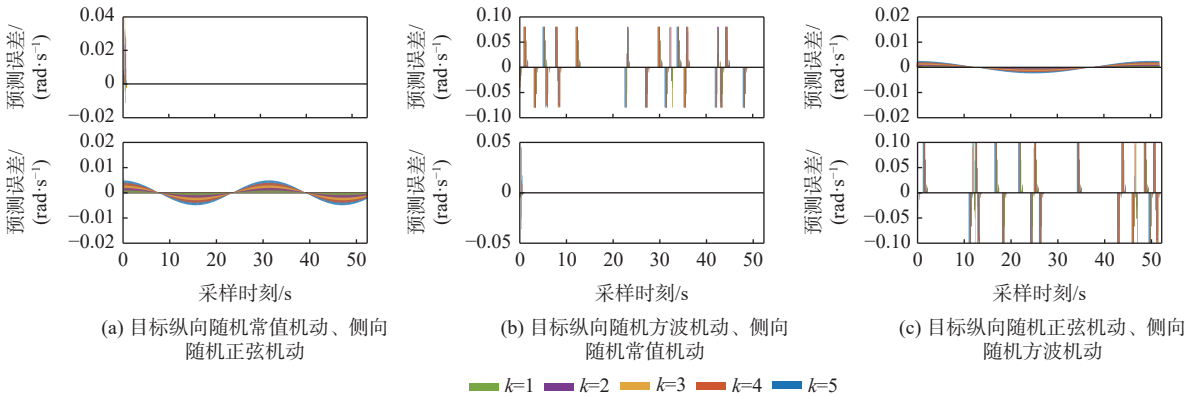
Fig. 8　NARnet training times at each sampling time



(a) 目标纵向随机常值机动、侧向随机正弦机动　(b) 目标纵向随机方波机动、侧向随机常值机动　(c) 目标纵向随机正弦机动、侧向随机方波机动

图 9　各时刻无人机协商次数

Fig. 9　UAV negotiation times at each sampling time



(a) 目标纵向随机常值机动、侧向随机正弦机动　(b) 目标纵向随机方波机动、侧向随机常值机动　(c) 目标纵向随机正弦机动、侧向随机方波机动

图 10　各时刻 NARnet 预测误差

Fig. 10　NARnet prediction error at each sampling time

表 4　不同算法参数下的仿真结果

Table 4　Simulation results under different algorithm parameters

仿真次序	NARnet最大单步训练次数	无人机最大单步协商次数	平均脱靶量/m	平均攻击时间误差/s	平均单步训练次数	平均单步协商次数	无人机平均单步决策用时/s	计算总时长/s
1	100	100	0.014 3	6.193×10^{-5}	19	28	0.017 8	112
2	100	50	0.013 3	6.386×10^{-5}	19	21	0.015 1	95.046 9
3	100	20	0.022 0	2.432×10^{-4}	19	13	0.012 5	79.265 6
4	50	100	0.015 1	7.838×10^{-5}	18	28	0.018 1	114.078 1
5	50	50	0.012 7	5.917×10^{-5}	18	20	0.014 9	93.859 4
6	50	20	0.019 3	1.750×10^{-4}	18	13	0.012 1	76.437 5

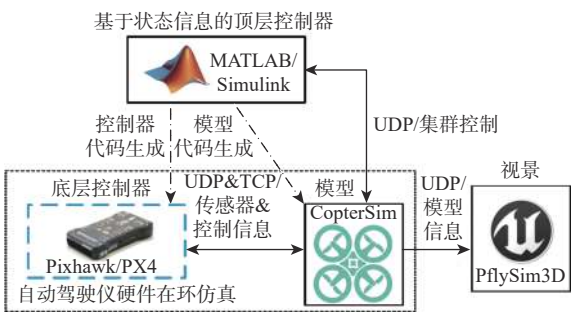


图 11　RflySim 闭环仿真构架

Fig. 11　RflySim closed loop simulation framework

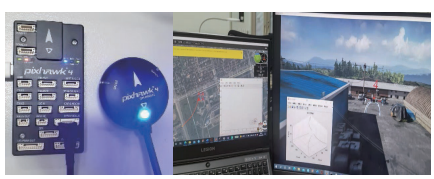


图 12　半实物仿真系统组成

Fig. 12　Components of hardware in loop simulation system

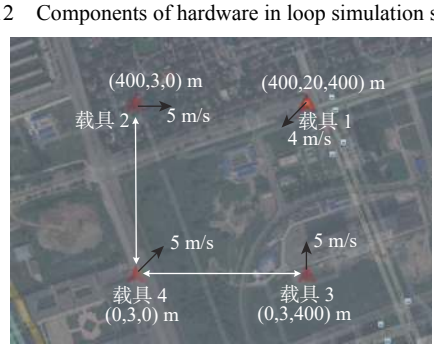


图 13　无人机-目标初始状态及无人机通信拓扑结构

Fig. 13　Initial state of UAVs and target and communication topology of UAVs

设定系统采样间隔为 1/30 s, 预测时域为 0.5 s, 记录无人机状态信息。耗时 58.3 s, 拦截完成, 仿真过程如图 14 和图 15 所示。图 14 为无人机与目标的三维运动轨迹及 RflySim3D 视景下无人机包围与拦截目标的瞬时画面; 图 15(a)～图 15(c) 分别为各无人机与目标的相对距离、各无人机剩余攻击时间及攻击时间误差的变化情况; 图 15(d) 和图 15(e) 分别为各无人机 3 轴向速度及 3 个制导指令分量的变化情况。

由图 14 可知, 在半实物仿真系统中, VN-RHC

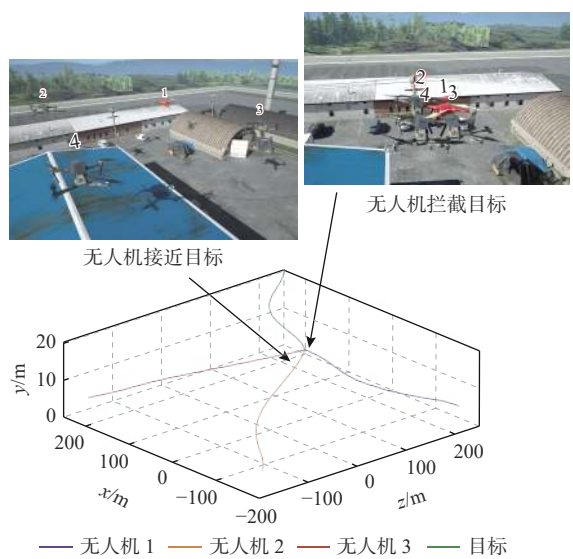
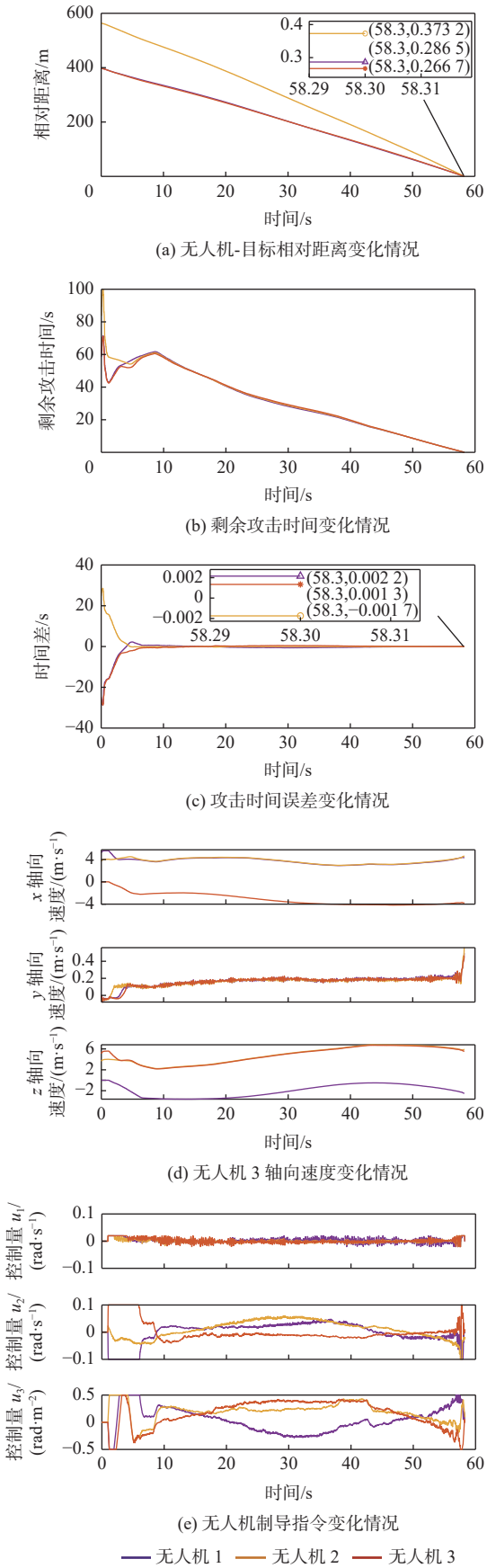


图 14　半实物仿真实验无人机、目标三维运动轨迹

Fig. 14　Three-dimensional motion trajectory of UAVs and target in hardware in loop simulation

制导策略仍然可以引导各无人机完成拦截，但由于受到机体结构的限制，无人机无法在具有真实物理碰撞效果的实验软件中取得更小的制导误差，因此，图 15(a) 中标明的数据为实际发生碰撞时双方 GPS 方位的相对距离。由图 15(d) 和图 15(e) 可知，由于随机扰动与未建模动态未得到有效表达，系统实际输出与期望输出存在偏差，从而引发了制导指令与无人机速度的抖动，这一现象符合无人机实际运动特征，并且说明 VN-RHC 制导策略可以在具有真实物理效果的半实物仿真系统中有效应对此类情况。上述半实物仿真实验结果表明，本文设计的 VN-RHC 协同制导策略在实际任务场景与仿真条件下具有较强的适应能力，并有望在更广泛的任务设定与更真实的实验环境中得到进一步检验。



(a) 无人机-目标相对距离变化情况

(b) 剩余攻击时间变化情况

(c) 攻击时间误差变化情况

(d) 无人机 3 轴向速度变化情况

(e) 无人机制导指令变化情况

图 15 半实物仿真结果

Fig. 15 Results of hardware in loop simulation

5 结 论

1）为增强制导系统动态适应能力，在 RHC 框架下，通过在线求解开环优化问题获得协同制导指令；为同时确保时间协同与导引控制精度，设计目标函数对命中误差与一致性误差进行统一描述。

2）为提升制导策略鲁棒性，利用单隐含层 NARnet 对目标机动指令进行在线估计，并使用 L-M 算法进行网络训练，提升预测效率。

3）针对无人机之间控制输入耦合问题，设计基于 VN 机制的协同制导指令生成策略，在分布式通信架构下，实现多无人机同步、自主决策。利用 NM-RMSP 算法进行更新，提升无人机协商速度。

实验结果表明，本文设计的分布式协同制导策略对各类机动目标均取得了较好的制导效果，对于减缓协同制导过程中的通信负担、提升系统鲁棒性及制导指令在线计算效率具有显著意义。

下一步将对复杂障碍环境下多无人机分布式协同制导问题展开研究，并对研究成果进行进一步地检验。

参考文献（References）

[1] 陈清阳, 辛宏博, 王玉杰, 等. 一种多机协同打击的快速航迹规划方法[J]. 北京航空航天大学学报, 2022, 48(7): 1145-1153.
CHEN Q Y, XIN H B, WANG Y J, et al. A rapid path planning method for multiple UAVs to cooperative strike[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(7): 1145-1153(in Chinese).

[2] 王晓光, 章卫国, 刘洋. 伴飞诱饵干扰下的自杀式无人机攻击策略[J]. 航空学报, 2015, 36(9): 3137-3146.
WANG X G, ZHANG W G, LIU Y. Suicide drones' attack strategy on the condition of escort free-flight decoys influence[J]. Acta Aeronautica et Astronautica Sinica, 2015, 36(9): 3137-3146(in Chinese).

[3] 吴捷, 胡盛华, 乔莎莎, 等. "忠诚僚机"式有人/无人机协同作战

概念与任务管理技术研究[J]. 航空电子技术, 2021, 52(2): 27-31.
WU J, HU S H, QIAO S S, et al. Research on combat concept and mission management technique of “loyal wingman” manned/unmanned aerial vehicle collaborative[J]. Avionics Technology, 2021, 52(2): 27-31(in Chinese).

[4] 田磊, 赵启伦, 董希旺, 等. 拒止环境下基于“忠诚僚机”的护航策略[J]. 北京航空航天大学学报, 2021, 47(5): 1058-1067.
TIAN L, ZHAO Q L, DONG X W, et al. Escort strategy based on loyal wingman in denial environment[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(5): 1058-1067(in Chinese).

[5] ZHAO J, ZHOU R. Unified approach to cooperative guidance laws against stationary and maneuvering targets[J]. Nonlinear Dynamics, 2015, 81(4): 1635-1647.

[6] 彭琛, 刘星, 吴森堂, 等. 多弹分布式协同末制导时间一致性研究[J]. 控制与决策, 2010, 25(10): 1557-1561.
PENG C, LIU X, WU S T, et al. Consensus problems in distributed cooperative terminal guidance time of multi-missiles[J]. Control and Decision, 2010, 25(10): 1557-1561(in Chinese).

[7] 杨剑影, 周佳玲, 魏小倩. 多导弹攻击高机动目标的分布式协同制导关键技术[J]. 航空兵器, 2017, 24(3): 3-12.
YANG J Y, ZHOU J L, WEI X Q. Key technologies of distributed cooperative guidance and control method for multiple missiles attacking the maneuvering target[J]. Aero Weaponry, 2017, 24(3): 3-12 (in Chinese).

[8] 赵世钰, 周锐. 基于协调变量的多导弹协同制导[J]. 航空学报, 2008, 29(6): 1605-1611.
ZHAO S Y, ZHOU R. Multi-missile cooperative guidance using coordination variables[J]. Acta Aeronautica et Astronautica Sinica, 2008, 29(6): 1605-1611(in Chinese).

[9] 张帅, 宋天莉, 焦巍, 等. 带有拦截时间约束的协同制导方法[J]. 北京航空航天大学学报, 2023, 49(8): 1956-1963.
ZHANG S, SONG T L, JIAO W, et al. Cooperative guidance method with interception time constraint[J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49(8): 1956-1963(in Chinese).

[10] 田野, 蔡远利, 邓逸凡. 一种带时间协同和角度约束的多导弹三维协同制导律[J]. 控制理论与应用, 2022, 39(5): 788-798.
TIAN Y, CAI Y L, DENG Y F. A 3D cooperative guidance law for multiple missiles with line-of-sight angle constraint[J]. Control Theory & Applications, 2022, 39(5): 788-798(in Chinese).

[11] CHO D, KIM H J, TAHK M J. Nonsingular sliding mode guidance for impact time control[J]. Journal of Guidance, Control, and Dynamics, 2016, 39(1): 61-68.

[12] ZHANG Y A, WANG X L, WU H L. Impact time control guidance law with field of view constraint[J]. Aerospace Science and Technology, 2014, 39: 361-369.

[13] 邹丽, 丁全心, 周锐. 异构多导弹网络化分布式协同制导方法[J]. 北京航空航天大学学报, 2010, 36(12): 1432-1435.
ZOU L, DING Q X, ZHOU R. Distributed cooperative guidance for multiple heterogeneous networked missiles[J]. Journal of Beijing University of Aeronautics and Astronautics, 2010, 36(12): 1432-1435(in Chinese).

[14] ZHAO S Y, ZHOU R. Cooperative guidance for multimissile salvo attack[J]. Chinese Journal of Aeronautics, 2008, 21(6): 533-539.

[15] 邹丽, 周锐, 赵世钰, 等. 多导弹编队齐射攻击分散化协同制导方法[J]. 航空学报, 2011, 32(2): 281-290.
ZOU L, ZHOU R, ZHAO S Y, et al. Decentralized cooperative guidance for multiple missile groups in salvo attack[J]. Acta Aeronautica et Astronautica Sinica, 2011, 32(2): 281-290(in Chinese).

[16] 王晓芳, 刘冬责, 郑艺裕, 等. 一种基于模型预测控制技术的多导弹协同作战制导方法: CN105202972A[P]. 2017-01-18.
WANG X F, LIU D Z, ZHENG Y Y, et al. Multi-missile cooperative engagement guidance method based on model predictive control technique: CN105202972A[P]. 2017-01-18(in Chinese).

[17] 段海滨, 邱华鑫. 基于群体智能的无人机集群自主控制[M]. 北京: 科学出版社, 2018: 168-171.
DUAN H B, QIU H X. Unmanned aerial vehicle swarm autonomous control based on swarm intelligence[M]. Beijing: Science Press, 2018: 168-171(in Chinese).

[18] CASADO E J M, SCARLATTI D, ESTEBAN-CAMPILLO D, et al. Network of unmanned vehicles: US20140074339[P]. 2014-03-13.

[19] JOHNSON L, CHOI H L, HOW J P. The hybrid information and plan consensus algorithm with imperfect situational awareness[C]// Proceedings of the Springer Tracts in Advanced Robotics. Berlin: Springer, 2016: 221-233.

[20] 杨海民, 潘志松, 白玮. 时间序列预测方法综述[J]. 计算机科学, 2019, 46(1): 21-28.
YANG H M, PAN Z S, BAI W. Review of time series prediction methods[J]. Computer Science, 2019, 46(1): 21-28(in Chinese).

[21] 孙雪娇, 周锐, 吴江, 等. 攻击机动目标的多导弹分布式协同制导律[J]. 北京航空航天大学学报, 2013, 39(10): 1403-1407.
SUN X J, ZHOU R, WU J, et al. Distributed cooperative guidance law for multiple missiles attacking maneuver target[J]. Journal of Beijing University of Aeronautics and Astronautics, 2013, 39(10): 1403-1407(in Chinese).

Distributed cooperative guidance strategy based on virtual negotiation and rolling horizon optimization

YANG Xiuxia*，JIANG Zijie，ZHANG Yi，WANG Cong，YU Hao

(Naval Aviation University，Yantai 264001，China)

Abstract：The multi-UAV cooperative guidance under target maneuvering is studied, and a distributed cooperative guidance strategy based on virtual negotiation mechanism and rolling horizon optimization is proposed. To improve the effect of time coordination and guidance control at the same time, the local coordination variables and objective function are designed under the rolling horizon optimization framework. The multi-constraint cooperative guidance optimization model is then constructed in the finite time domain. To enhance the robustness of the guidance system to maneuvering targets, a disturbance observer based on nonlinear autoregressive neural network is designed. To address the control input coupling between UAVs, a virtual negotiation mechanism is proposed, realizing synchronous decision-making and eliminating internal “differences”. The guidance command generation strategy is also designed by using the root mean square propagation algorithm of Nesterov momentum. The experiments of digital simulation and hardware in the loop simulation are carried out. The results show that the proposed guidance strategy can deal with various system interferences and improve the effectiveness of the cooperative guidance through online prediction and optimization, thus expected to be further applied and tested in practical projects.

Keywords：cooperative guidance；rolling horizon optimization；neural network；virtual negotiation；root mean square propagation algorithm

Received：2022-03-21；**Accepted**：2022-06-10；**Published Online**：2022-06-20 10：06

URL：link.cnki.net/urlid/11.2625.V.20220617.0845.001

Foundation item：Shadong Provincial Natural Science Foundation (ZR2020MF090)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0186

基于QAR数据的民机高高原进近着陆风险评估方法

陈农田[1,*]，满永政[2]，李俊辉[2]

(1. 中国民用航空飞行学院 航空工程学院，广汉 618307；2. 中国民用航空飞行学院 民航安全工程学院，广汉 618307)

摘 要：民机高高原进近着陆是高原飞行高风险阶段。为有效实施高高原进近着陆风险识别和等级判据，提出基于熵权可变模糊识别的长短时记忆网络与深度神经网络(LSTM-DNN)相融合的深度学习风险评估方法。基于快速存取记录器(QAR)记录的高高原飞行数据，借鉴民机飞行品质监控（FOQA）咨询通告和行业QAR监控标准，结合指标重要度分析与Delphi专家调查，提取着陆时航向变化大、航迹低、610～305 m进近时下降率大、接地时垂直加速度及153～15 m进近时下降率大5个关键监控项目作为民机高高原进近着陆风险评估指标。为克服评估指标权重主观性偏差，应用熵权法确定评估指标权重，基于可变模糊识别方法构建风险等级隶属函数，建立基于LSTM-DNN的民机高高原进近着陆风险评估模型。以成都—拉萨进近着陆航段为例，提取QAR数据，对该风险评估模型进行训练与测试，并与Logistic多元回归、支持向量机(SVM)等评估方法进行比较，结果表明：所提方法平均准确率达到94.18%，最高可达94.79%，验证了方法的客观有效性。

关 键 词：高高原飞行；快速存取记录器数据；熵权；可变模糊识别；LSTM-DNN深度学习；风险评估
中图分类号：V328.5
文献标志码：A **文章编号**：1001-5965(2024)01-0077-09

根据《高原机场运行》定义，高高原机场是指海拔高度在2 438 m及以上的机场[1]。在全球范围内，高高原机场有46个，其中，中国有19个，约占全球总数量的41%，主要位于西藏、甘肃、四川、青海、湖北和云南6个省份。高高原民航安全运输对保障高原民生经济发挥了重要作用。高高原机场具有海拔高、气压低、空气稀薄、地形复杂、日照强烈且不均衡等特点，使得飞机在高高原飞行操纵难度加大，因此，对高高原运行飞机性能和机组操纵提出更高的要求。民机进近着陆是飞行过程中最复杂、最频繁，也是飞行事故率最高的飞行阶段[2]。在民机进近着陆过程中，可供飞行员处理时间极短，飞行员精神需高度集中。通过对近10年的民航不安全事件数据统计，除无法按飞行阶段分类的447起事故征候，在总数为927起事故征候当中，进近着陆阶段227起，约占24%，位居第一[3]。由于高高原地区通常存在地形复杂、净空条件差等特点，飞机进行进近着陆时常遭遇低能见度、低空风切变、大风等复杂天气。高高原地区地形及运行环境影响使得民机高高原机场终端运行阶段(含进近着陆)运行风险更加突出。国际民航组织(International Civil Aviation Organization, ICAO)在颁布的《安全管理手册》中要求现代的民航安全管控需以数据驱动的风险管理为核心[4]。快速存取记录器(quick access recorder, QAR)数据在民航事故调查、民机故障诊断方面具有广泛应用，已成为航空公司开展针对性飞行训练、提高航班运行品质和安全水平的重要依据。因此，开展基于QAR数据的民机高高原进近着陆阶段风险识别与管控研究具有现实意义。

近年来，国内外学者在针对民机高高原进近着陆飞行员操纵行为及QAR数据飞行品质监控应用

收稿日期：2022-03-25；**录用日期**：2022-06-14；**网络出版时间**：2022-06-24 14:53
网络出版地址：link.cnki.net/urlid/11.2625.V.20220623.1334.003
基金项目：国家自然科学基金民航联合基金重点项目(U2033202)；四川省科技厅重点研发项目(2022YFG0213)
*通信作者. E-mail：chennongtian@hotmail.com

引用格式：陈农田，满永政，李俊辉. 基于QAR数据的民机高高原进近着陆风险评估方法[J]. 北京航空航天大学学报，2024，50(1)：77-85.
CHEN N T，MAN Y Z，LI J H. Risk assessment method for civil aircraft approach and landing at high plateau based on QAR data[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(1)：77-85 (in Chinese).

方面开展了系列研究。Xu等[5]根据QAR数据空间结构的变化给出故障信息，定位到一定的故障模式。Kluever[6]运用最优控制理论建模，利用频域分析提出评估高阶手动控制动力系统操作质量的方法。文献[7]通过分析QAR飞行数据和影响稳定进近因素，得到评估进近着陆阶段安全性的聚类系数。霍纬纲等[8]采用模糊C均值对QAR数据进行抽取和模糊预处理，利用模糊自适应聚类方法有效诊断超限事件。祁明亮等[9]利用某机型的QAR超限事件建立数学模型，并以黄金分割法为基础设计方法，为飞行员模拟机针对性训练提供指导。Wang等[10]并于风险评估原则开发了着陆操作绩效评估模型，并基于QAR数据选取着陆参数作为评估该模型飞行员着陆操作性能的指标，建立飞行着陆运行性能评价系统。鲁志东等[11]构建了民机飞行数据仿真实验平台，运用随机模拟和相关性分析方法解析能量状态参数偏离诱发进近着陆风险的影响规律，根据仿真结果给出能量状态安全边界。徐浩军等[12]通过飞行安全参数风险度叠加仿真计算操纵空间飞行风险拓扑云图，得到飞行场景下飞行安全谱和飞行风险概率整个操纵空间，构建飞行安全操纵空间，引导驾驶员正确操纵。徐骏驰和孙有朝[13]针对高高原机场特点，提出对运输类飞机在高高原机场运行的适航验证方法。

国内外学者在民机进近着陆风险与分析及QAR数据飞行品质监控应用方面取得了系列成果，特别是飞行员进近着陆操纵行为评价方面，但是在基于QAR飞行数据的高高原民机进近着陆风险评估方面研究较少。本文基于QAR飞行数据，借鉴民机飞行品质监控（flight operations quality assurance, FOQA）咨询通告，根据高高原民机进近着陆运行特点，提出民机高高原进近着陆风险评估关键参数指标，为克服指标权重的主观性问题，应用熵权法确定评估指标权重。依据民机机型高高原运行状态监控标准，利用可变模糊识别方法建立风险等级隶属函数，提出基于长短时记忆网络（long short term memory, LSTM）与深度神经网络（deep neural network, DNN）相融合的民机高高原进近着陆风险评估方法，为有效实施高高原民机进近着陆安全风险管控提供参考。

1 民机进近着陆风险评估指标确定

高高原民机进近着陆风险关键监控项目的选取直接影响评估结果。选取的关键监控项目能够直接反映民机进近着陆阶段的构型状态，需要具备可获取性、可衡量性和监控性。目前，QAR数据主要用于飞行过程性能可视化、适航维修、飞行品质监控等方面[14-16]。本文通过借鉴民机FOQA监控项目规范指标[14]，如表1所示。表中，ILS为仪表着陆系统，IVV为瞬时垂直速度，ALL为飞行着陆灯。根据A320飞机QAR飞行数据遴选民机在高高原进近着陆阶段风险评估的指标，综合借鉴进近着陆操纵安全行为评价相关研究[17-19]，结合飞行管理人员和飞行技术人员的意见及高高原实际运行情况，利用Delphi法对进近着陆监控项目重要性进行打分，最后一次咨询最终确定5个主要监测指标：着陆时航向变化大、航迹低、610～305 m进近时下降率大、接地时垂直加速度及153～15 m进近时下降率大。

表1 空客系列飞机飞FOQA监控项目规范指标（部分）[14]

Table 1 Specification indicators of FOQA project for airbus series aircraft (partial)[14]

监控项目	监控参数	监控点
下降率大	IVV, ALL	610～305 m(含) 305～152 m(含) 152～15 m(含)
进近滚转角大	滚转角, ALL	457～152 m(含) 152～61 m(含) 61～15 m(含)
着陆滚转角大	滚转角, ALL	15 m至所有机轮接地
低高速使用减速板	减速板, ALL	使用减速板
进近速度小	空速, ALL	305～15 m(含)
进近速度大	空速, ALL	152～15 m(含)
着陆速度大	空速, ALL	15 m以下
ILS下滑道偏离	下滑道偏离, ALL	305 m以下
ILS航向道偏离	航向道偏离, ALL	305 m以下
低空大速度	空速, ALL	762 m以下

2 基于熵权可变模糊识别的风险等级隶属函数

2.1 熵权法

为克服指标权重的主观性问题，采用熵权法确定各指标权重。由于QAR数据各指标趋向和量纲都不同，为便于各指标比较，原始数据需要进行规范化处理，一般将其压缩在[0,1]之间，因此，要对指标进行标准化处理。对于越大越好数值指标采用式(1)，而越小越好指标采用式(2)。

步骤1 数据无量纲化处理。

$$r_{ij}=\frac{x_{ij}-x_j^{\min}}{x_j^{\max}-x_j^{\min}} \tag{1}$$

$$r_{ij}=1-\frac{x_{ij}-x_j^{\min}}{x_j^{\max}-x_j^{\min}} \tag{2}$$

式中: x_{ij} 为 QAR 数据; r_{ij} 为第 x_{ij} 的规范化数据; $x_j^{\min}$ 为指标特征矩阵 $\boldsymbol{X}$ 中第 j 个指标所对应的最小值; $x_j^{\max}$ 为指标特征矩阵 $\boldsymbol{X}$ 中第 j 个指标所对应的最大值。

根据标准化后的各评价指标数值计算各指标权重。

步骤 2　计算熵权值。

$$E_i = -\frac{1}{\ln m}\sum_{j=1}^{m} f_{ij}\ln f_{ij} \tag{3}$$

式中: m 为评价样本的数量; f_{ij} 为数据预处理。

$$f_{ij} = r_{ij}\Big/\sum_{i=1}^{m} r_{ij} \tag{4}$$

步骤 3　计算熵权向量 $\boldsymbol{w}$。

$$\begin{cases} \boldsymbol{w} = [w_1, w_2, \cdots, w_m]^{\mathrm{T}} \\ w_i = (1-E_i)\Big/\left(n - \sum_{i=1}^{n} E_i\right) \end{cases} \qquad i = 1,2,\cdots,m \tag{5}$$

式中: w_i 为指标 i 的权重。

从式(5)中可以得出, 若某个评价指标的信息熵越小, 表示某个指标变异程度越大, 提供信息越多, 在综合评价中起的作用越大, 反之亦然。该方法利用指标信息量大小确定权重, 熵权法具有客观性[20]。

2.2　可变模糊识别方法

文献 [20] 建立了模糊集理论, 文献 [20-23] 将模糊集理论引入水文水资源系统研究, 创建了模糊水文水资源学, 目前在诸多领域具有广泛应用。本文基于该理论对 QAR 飞行数据进行分析, 根据高高原民机运行状态监控标准, 建立高高原民机进近着陆风险等级隶属函数。计算步骤如下:

1) 设有 m 个评价样本, n 个评价指标, 构建样本特征值向量: $\boldsymbol{X}=[x_1, x_2, \cdots, x_n]$。设样本评价等级 $h=1,2,\cdots,c$, 其中, 1 为最优等级, c 为最差等级。则等级评价标准特征值矩阵为

$$\boldsymbol{I}_{ab} = \begin{bmatrix} [a_{11},b_{11}] & [a_{12},b_{12}] & \cdots & [a_{1c},b_{1c}] \\ [a_{21},b_{21}] & [a_{22},b_{22}] & \cdots & [a_{2c},b_{2c}] \\ \vdots & \vdots & & \vdots \\ [a_{m1},b_{m1}] & [a_{m2},b_{m2}] & \cdots & [a_{mc},b_{mc}] \end{bmatrix} = ([a,b]_{ih}) \tag{6}$$

2) 对数据进行预处理:

$$f_{ij} = \left(1+r_{ij}\right)\Big/\sum_{i=1}^{n}\left(1+r_{ij}\right)\quad i=1,2,\cdots,n;\ j=1,2,\cdots,m \tag{7}$$

3) 将特征值 x_{ij} 与级别 h 指标 i 的吸引域区间的中间值 M_{ih} 进行比较, 若 $x_{ij}<M_{ih}$, 可得相对隶属度为

$$\boldsymbol{u}_{\boldsymbol{A}}(x_{ij})_h = \begin{cases} 0.5\left(1+\dfrac{x_{ij}-a_{ih}}{M_{ih}-a_{ih}}\right)\boldsymbol{F} & x_{ij}\in[a_{ih},M_{ih}] \\ 0.5\left(1-\dfrac{x_{ij}-a_{ih}}{c_{ih}-a_{ih}}\right)\boldsymbol{F} & x_{ij}\in[c_{ih},a_{ih}] \end{cases} \tag{8}$$

式中: $\boldsymbol{F}$ 为单位矩阵。

若 $x_{ij}>M_{ih}$, 得相对隶属度为

$$\boldsymbol{u}_{\boldsymbol{A}}(x_{ij})_h = \begin{cases} 0.5\left(1+\dfrac{x_{ij}-b_{ih}}{M_{ih}-a_{ih}}\right)\boldsymbol{F} & x_{ij}\in[M_{ih},b_{ih}] \\ 0.5\left(1-\dfrac{x_{ij}-b_{ih}}{d_{ih}-a_{ih}}\right)\boldsymbol{F} & x_{ij}\in[b_{ih},d_{ih}] \end{cases} \tag{9}$$

4) c_{ih} 和 d_{ih} 可根据矩阵 $\boldsymbol{I}_{ab}$ 中各级指标标准值区间两侧相邻区间的上限和下限确定。

计算综合相对隶属度 ${}_j\boldsymbol{U}_h$ 为

$${}_j\boldsymbol{U}_h = \left\{\boldsymbol{E} + \left[\frac{\sum_{i=1}^{m}\left[w_i\left(\boldsymbol{E}-\boldsymbol{u}_{\boldsymbol{A}}\left(x_{ij}\right)_h\right]^p}{\sum_{i=1}^{m}\left[w_i\boldsymbol{u}_{\boldsymbol{A}}\left(x_{ij}\right)_h\right]^p}\right]^{\frac{a}{p}}\right\}^{-1} \tag{10}$$

式中: a 为准则参数; $\boldsymbol{E}$ 为单位矩阵; p 为距离参数; $\boldsymbol{A}$ 为训练集样本矩阵。其中, a 和 p 有 4 种组合方式: 当 p=1, a=1 时, 为模糊综合评价模型; 当 p=1, a=2 时, 为理想模型; 当 p=2, a=1 时, 为神经元激励函数模型(S 型函数); 当 p=2, a=2 时, 为模糊评价优选模型。由此可见, 4 种参数组合构成可变模糊集。

对 $\boldsymbol{u}_{\boldsymbol{A}}(x_{ij})_h$ 进行归一化处理:

$${}_j\boldsymbol{u}_h = \boldsymbol{u}_{\boldsymbol{A}}(x_{ij})_h\Big/\sum_{h=1}^{c}\boldsymbol{u}_{\boldsymbol{A}}(x_{ij})_h \tag{11}$$

5) 计算应用级别特征值模型, 根据 H_j 及其平均值对样本进行综合评价:

$$H_j = [1,2,\cdots,c]\cdot{}_j\boldsymbol{u}_h \tag{12}$$

结合判断准则($\bar{H}$为级别特征值 H 的平均值):

$$\begin{cases} 1.0 \leqslant \bar{H} \leqslant 1.5 & \text{归属于1级} \\ h-0.5 < \bar{H} \leqslant h & \text{归属于}h\text{级,偏}(h-1)\text{级} \\ & (h=2,3,\cdots,c-1) \\ h < \bar{H} \leqslant h+0.5 & \text{归属于}h\text{级,偏}(h+1)\text{级} \\ & (h=2,3,\cdots,c-1) \\ c-0.5 \leqslant \bar{H} \leqslant c & c\text{级} \end{cases} \tag{13}$$

3　本文方法

3.1　本文方法的结构

本文方法由 3 部分组成, 分别为数据预处理、网络训练及输出模型分类。首先, 数据经过处理构成时间序列; 然后, 将数据输入到 LSTM 模型中, 获得风险要素在未来的变化趋势, 将获得的未来变化趋势输入到时间序列中, 同理继续操作, 可获得未

来多个时刻的趋势，将所有趋势组合起来，作为 DNN 的输入；最后，根据 Softmax 函数进行分类预测，其分析框架流程如图 1 所示。

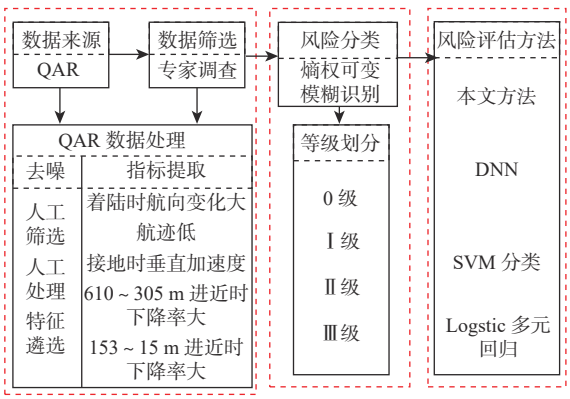


图 1 分析框架流程

Fig. 1 Flow chart of analysis framework

3.2 深度神经网络

DNN 是在多层神经网络中，倒数第 1 层与倒数第 2 层之间神经元相互连接。DNN 属于一种经典网络，DNN 中的每个运算节点都代表一个神经元，层层之间神经元相互连接，但是同一层神经元不连接。本文方法采用 3 层的 DNN，包括输入层、隐藏层和输出层，如图 2 所示。相邻 2 层之间，上一层的输出作为下一层的输入，每个节点包括激活函数、偏执量和权重[24-26]。

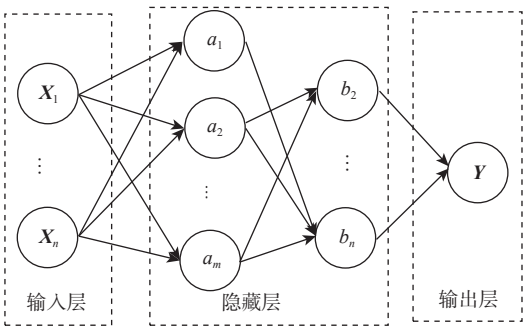


图 2 深度神经网络

Fig. 2 Deep neural network

DNN 包括向前传播和反向传播。DNN 可以拟合任何函数，其表达能力和非线性拟合能力强。前向传播的目的是生产预报值。

$$a_n^{(m)} = \sum_{i=1}^{N} a_i^{(m-1)} w_i^{(m-1)} = \text{activate}\left(a_n^{(m)}\right) \tag{14}$$

式中：N 为上一层的单元个数；activate 为激活函数。

中间层的激励函数 ReLU[27] 为

$$\text{ReLU}(z) = \max(0, z) \tag{15}$$

输出层的激活函数 sigmoid 为

$$\text{sigmoid}(z) = \frac{1}{1+\text{e}^{-z}} \tag{16}$$

后向传播的目的是取最优的过程。在后向传播中，前向传播所得预测结果和真实结果进行对比，优化权重系数，最终减小预测值和真实值差距。本文实验训练超参数中，学习率设置为 0.001，权值衰减率设置为 1×10^{-6}。

3.3 长短时记忆网络

LSTM 是一种特殊的循环神经网络（recurrent neural network, RNN）。与普通循环神经网络相比，LSTM 增加记忆功能，并且隐藏层信息只来源于当前输入和上一刻的隐藏层信息，因此，其只能处理一定的短期依赖，无法处理长期依赖，这就产生了梯度爆炸或梯度消失的问题。LSTM 在循环神经网络基础上引入细胞状态，根据细胞状态决定保留或遗忘。LSTM 解决了循环神经网络存在的长期依赖问题。从微观上看，LSTM 使用了遗忘门、输入门、输出门 3 种门来保持和控制信息，选择性地记忆反馈误差函数随梯度下降的修正参数，从而实现时间上的记忆功能，并防止梯度消失。LSTM 单元内部的计算式为

$$\begin{cases} \boldsymbol{f}_t = \sigma\left(\boldsymbol{W}_f * [\boldsymbol{n}_{t-1}, \boldsymbol{x}_t] + \boldsymbol{b}_f\right) \\ \boldsymbol{i}_t = \sigma(\boldsymbol{W}_i * [\boldsymbol{n}_{t-1}, \boldsymbol{x}_t] + \boldsymbol{b}_i) \\ \tilde{\boldsymbol{C}}_t = \tanh(\boldsymbol{W}_C * [\boldsymbol{n}_{t-1}, \boldsymbol{x}_t] + \boldsymbol{b}_C) \\ \boldsymbol{C}_t = \boldsymbol{f}_t * \boldsymbol{C}_{t-1} + \boldsymbol{i}_t * \tilde{\boldsymbol{C}}_t \\ \boldsymbol{O}_t = \sigma(\boldsymbol{W}_O * [\boldsymbol{n}_{t-1}, \boldsymbol{x}_t] + \boldsymbol{b}_O) \end{cases} \tag{17}$$

式中：$\boldsymbol{i}$、$\boldsymbol{O}$ 和 $\boldsymbol{f}$ 分别为输入门、输出门和遗忘门；$\boldsymbol{W}$ 和 $\boldsymbol{b}$ 分别为连接权重和偏置；*为卷积运算；$\boldsymbol{C}_t$ 和 $\tilde{\boldsymbol{C}}$分别为更新的状态和候选状态；$\boldsymbol{x}$ 为输入；$\boldsymbol{n}$ 为输出；t 为时间；σ 为激活函数。

3.4 LSTM-DNN 融合方法

DNN 与 LSTM 都是深度学习的主流方法，相对而言，DNN 更适合进行数据预测，提取数据的局部特征，并组合抽象成高层特征。而 LSTM 更适合进行时间扩展，其具有长期记忆功能，适于处理时间序列。在 QAR 数据的特征提取中，不仅需要考虑不同参数之间的空间联系，还需要考虑数据在时间维度上的变化。因此，本文将 DNN 与 LSTM 这 2 种方法相融合，提出一种 LSTM-DNN 融合方法，使方法具有时间和预测准确度的特征表达能力。预测过程具体为：①数据预处理，对数据进行归一化操作，统一不同特征的量纲。对归一化后的数据进行降维，消除数据集中的冗余信息，并生成针对 LSTM 方法的时间序列数据。②利用时间序列数据对 LSTM 方法进行训练。训练达到最大迭代次数后，输出时间预测模型。③利用第②步得到的数据集对 DNN 方法进行训练。训练达到最大迭代次数后，输出最优的分类方法。④根据分类方法进行数据预测。

4　实例分析与验证

4.1　实例分析流程

对采集的民机高高原进近着陆 QAR 数据进行预处理, 选择并提取基于 QAR 数据参数的风险评估特征指标。基于特征选择和提取结果, 构建熵权可变模糊识别风险等级隶属函数, 建立民机高高原进近着陆风险 LSTM-DNN 模型深度学习风险评估方法, 并进行训练与测试。将评价结果与 Logistic 多元回归、支持向量机(support vector machines, SVM)等方法比较, 准确率变化曲线如图 3 所示。

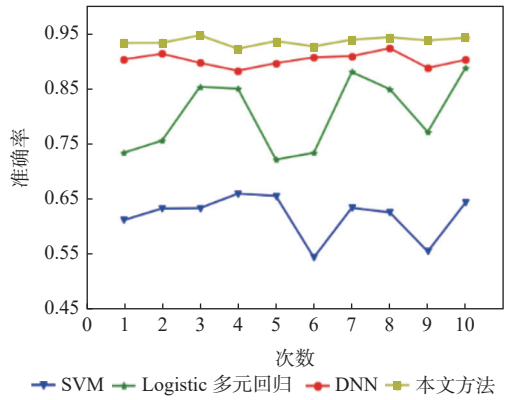


图 3　准确率变化对比

Fig. 3　Comparison of accuracy change

4.2　基于熵权可变模糊识别的风险等级判据

4.2.1　风险等级划分

借鉴民机 FOQA 咨询通告[14], 结合某航空公司高高原运行监控标准, 将空客 A320 机型高高原进近着陆风险划分为 4 个等级(本文定义 0 级为无风险等级), 根据等级定量划分方法确定风险等级, 如表 2 所示。

表 2　FOQA 风险评价指标分级标准

Table 2　Grading standard of FOQA risk evaluation index

风险等级	着陆时航向变化大/(°)	153～15 m进近时下降率大/(m·min^{-1})	航迹低/(°)	610～305 m进近时下降率大/(m·min^{-1})	接地时垂直加速度/g
0级	<4	≥−700	≤2.2	≥−1 200	≤1.5
Ⅰ级	[4,5]	[−800,−700)	(2.2,2.4]	[−1 500,−1 200)	(1.5,1.6]
Ⅱ级	(5,6)	(−900,−800)	(2.4,2.6)	(−1 800,−1 500)	(1.6,1.75)
Ⅲ级	≥6	≤−900	≥2.6	≤−1 800	≥1.75

4.2.2　基于熵权可变模糊识别的风险计算

根据 4.2.1 节所述提取风险等级评价关键指标, 结合熵权法, 运用可变模糊理论, 对飞机进近着陆不同等级进行计算划分, 详细计算过程如式(18)～式(20)所示(其他各样本计算方法相同, 本文不一一列出)。以提取的第 1 组数据为例, 可得特定的指标特征值向量 $\boldsymbol{X}$=[1.7,−600,0.09,−992,1]。

根据表 2 所示风险评价指标分级标准, 可得风险评价指标标准值区间矩阵为

$$\boldsymbol{I}_{ab}=\begin{bmatrix}[0,4] & [4,5] & [5,6] & [6,10]\\ [-700,400] & [-800,-700] & [-900,-800] & [-900,-1\,100]\\ [0,2.2] & [2.2,2.4] & [2.4,2.6] & [2.6,5]\\ [-1\,200,200] & [-1\,500,-1\,200] & [-1\,500,-1\,800] & [-1\,800,-2\,000]\\ [0,1.5] & [1.5,1.6] & [1.6,1.7] & [1.7,3]\end{bmatrix} \tag{18}$$

根据式(7)计算 M_{ih} 值, 得到隶属度为 1 的点值矩阵, 判断指标数据 x_{ij} 落在 M_{ih} 值的左端还是右端, 结合式(8)和式(9), 计算样本指标对评价级别的相对隶属度, 隶属度矩阵为

$$\boldsymbol{u}_A(x_{11})_1=\begin{bmatrix}0.775 & 0.09 & 0 & 0\\ 0.186 & 0 & 0 & 0\\ 0.502\,5 & 0.005 & 0 & 0\\ 0.68 & 0.32 & 0 & 0\end{bmatrix} \tag{19}$$

依据 2.1 节给出的熵权法确定指标的权重为

$$\boldsymbol{w}=[0.222\,5,0.148\,2,0.203\,6,0.213\,8,0.211\,9]^{\mathrm{T}} \tag{20}$$

根据式(10)求得参数 a 和 p 的 4 种组合下样本 j 对级别 h 的综合相对隶属度, 再结合式(11)和式(12)求得级别特征值 H。根据式(13)得出 $\bar{H}$=1.169 80≤1.5, 因此, 第 1 组数据的评价等级为 0 级, 如表 3 所示。

依据熵权可变模糊识别方法, 可以计算出其他样本组风险等级, 如表 4 所示。

4.3　基于 LSTM-DNN 深度学习的风险评估

自然语言处理一般采用精准度、准确率、召回率作为常用评判模型优劣性标准, 而准确率在多标

表3 4种参数组合下综合相对隶属度

Table 3 Comprehensive relative membership degree under combination of four parameters

参数组合	综合相对隶属度				H
	0级	Ⅰ级	Ⅱ级	Ⅲ级	
a=1 p=1	0.893 5	0.106 5	0	0	1.106 5
a=1 p=2	0.768 0	0.232 0	0	0	1.232 0
a=2 p=1	0.720 7	0.279 3	0	0	1.279 3
a=2 p=2	0.938 5	0.061 4	0	0	1.061 4

签预测模型结果评判时最为广泛使用，因此，本文采用准确率判断不同预测模型的效果。将本文方法与DNN方法、Logistic多元回归方法、SVM方法进行比较。随机选取50条作为测试集，分别对不同方法进行10次实验，在测试集上计算准确率。如图3所示，本文方法准确率比DNN方法、Logistic多元回归方法、SVM方法准确率高。

选取DNN方法作为评估预测方法进行测试和验证。为更直观反映本文方法的优劣，实验选取500个数据，其中，90%为测试集测试模型，10%为验证集以调整模型测试参数。图4为测试集与验证集在100个测试轮次下的准确率变化曲线。

表4 模糊评价识别结果

Table 4 Fuzzy evaluation and recognition results

样本组	着陆时航向变化大/(°)	153～15 m进近时下降率大/(m·min^{-1})	航迹低/(°)	610～305 m进近时下降率大/(m·min^{-1})	接地时垂直加速度/g	等级
1	1.8	−800	0.01	−912	0.96	0级
2	1.8	−832	0.08	−976	1.0	0级
3	4	−832	5	−1 316	1.15	Ⅰ级
4	9	−848	7	−1 558	0.5	Ⅱ级
5	0.6	−864	0.1	−800	0.98	Ⅰ级
⋮	⋮	⋮	⋮	⋮	⋮	⋮
499	2	−656	0	−999	2.80	Ⅱ级
500	8	−672	5	−988	3.78	Ⅲ级

由图4可知，随着测试轮次的增加，测试集的准确率趋于稳定，说明DNN方法收敛。测试集识别结果如图5所示。本文方法平均准确率达到94.18%，最高可达94.79%。

分别用1、3、5层隐藏层的本文方法进行预测，输入参数为：[1.2,−600,0.07,−1 024,0.99]。如表5所示。计算得出训练集为0级，且隐藏层数为3时，预测值为0.002 01，由此可得隐藏层数为3时最佳。

4.4 实验验证

提取某段成都—拉萨进近着陆航段A320的QAR飞行数据，隐藏层数设置为3，运用本文方法处理得到风险评估结果，如表6所示。

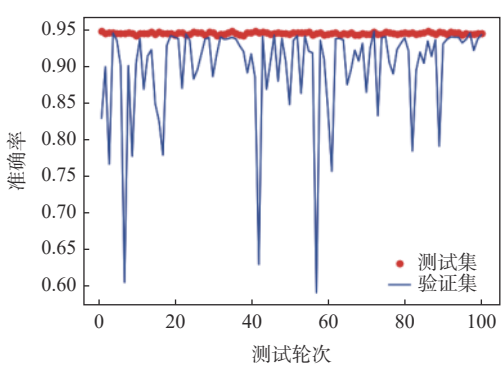


图4 测试集与验证集准确率变化

Fig. 4 Accuracy change of trial set and verification set

由表6可得，本次高高原民机进近着陆过程中，第99个样本风险等级为Ⅱ级，其接地时垂直加

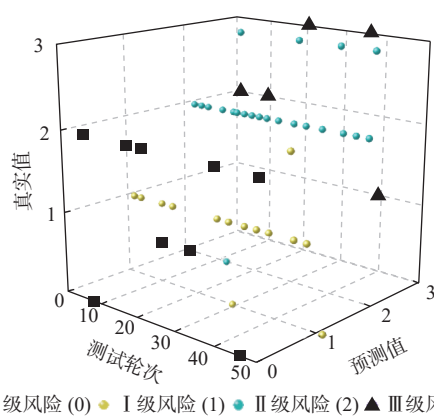


图5 本文方法风险等级识别结果

Fig. 5 Results of risk level recognition of proposed method

表5 不同隐藏层DNN预测值

Table 5 Predicted values of DNN with different hidden layers

隐藏层数	预测值
1	0.007 01
3	0.002 01
5	0.007 94

速度为 2.44g>1.75g, 接地时垂直加速度指标是导致风险等级为Ⅱ级的主要原因, 可能会诱发重着陆不安全事件发生。根据航向指标, 可得高高原低空风切变可能是导致航向偏离进而接地时垂直加速度过大的主要原因, 其他时间阶段为 0 级。根据最大值原则, 可以评定该高高原进近着陆总体风险等级为Ⅱ级风险, 需要加强对高高原低空风切变气象条件探测和预警, 提升飞行员面对突发状况时飞行应急处置能力和操纵技能。由该航段实际报告可知, 落地时存在重着陆现象。通过与实际安全信息分析报告相比较, 评估结论与实际相一致, 证明本文方法的客观有效性。

表 6　成都—拉萨进近着陆航段 QAR 数据 (部分)

Table 6　QAR data of Chengdu−Lhasa approach and landing segment (partial)

样本数量	着陆时航向变化大/(°)	153～15 m进近时下降率大/(m·min^{-1})	航迹低/(°)	610～305 m进近时下降率大/(m·min^{-1})	接地时垂直加速度/g	等级
1	0.9	−763	0.11	−896	0.99	0 级
2	0.6	−729	0.11	−880	1.01	0 级
⋮	⋮	⋮	⋮	⋮	⋮	⋮
50	1.6	−700	8	−832	0.96	0 级
51	2	−649	4	−848	1.01	Ⅰ 级
⋮	⋮	⋮	⋮	⋮	⋮	⋮
99	1	−832	4	−730	2.44	Ⅱ 级
100	1.7	−630	−0.67	−726	1.01	0 级

5　结　论

1) 本文建立了着陆时航向变化大、航迹低、610～305 m 进近时下降率大、接地时垂直加速度及 153～15 m 进近时下降率大 5 个民机高高原进近着陆风险评价关键监控项目。

2) 结合熵权与可变模糊识别方法建立风险评价关键指标风险隶属函数, 得出相对隶属度, 从而确定该组 QAR 数据对应的风险等级。

3) 运用本文方法进行实验获取损失变化特征, 根据各方法对比验证了本文方法可用性。通过比较不同隐藏层的 DNN 确定最优隐藏层数。

4) 根据 QAR 数据评估民机高高原进近着陆风险水平, 平均准确率为 94.18%, 验证了本文方法在民机高高原进近着陆识别和评估的客观有效性。

参考文献（References）

[1] 中国民用航空局飞行标准司. 高原机场运行: AC-121-FS-2015-21R1[S]. 北京: 中国民用航空局, 2015.
Flight Standards Department of Civil Aviation Administration of China. Plateau airport operation: AC-121-FS-2015-21R1[S]. Beijing: Civil Aviation Administration of China, 2015(in Chinese).

[2] 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).

[3] 罗婷婷, 孙瑞山. 基于 Petri 网的进近操作管理模型研究[J]. 中国安全生产科学技术, 2013, 9(11): 137-141.
LUO T T, SUN R S. Study on management system model for approach based on Petri net[J]. Journal of Safety Science and Technology, 2013, 9(11): 137-141(in Chinese).

[4] International Civil Aviation Organization. Safety management manual: Doc 9859[S]. Montreal: International Civil Aviation Organization, 2018.

[5] XU Y J, HOU W K, LI W Z, et al. Aero-engine gas-path fault diagnosis based on spatial structural characteristics of QAR data[C]// Proceedings of the Annual Reliability and Maintainability Symposium. Piscataway: IEEE Press, 2018: 1-7.

[6] KLUEVER C A. Unpowered approach and landing guidance with normal acceleration limitations[J]. Journal of Guidance, Control, and Dynamics, 2007, 30(3): 882-885.

[7] 周长春, 胡栋栋. 基于灰色聚类方法的航空公司飞机进近着陆阶段安全性评估[J]. 中国安全生产科学技术, 2012, 8(7): 99-102.
ZHOU C C, HU D D. Safety assessment of aircraft during approach-landing stage based on grey clustering method[J]. Journal of Safety Science and Technology, 2012, 8(7): 99-102(in Chinese).

[8] 高小霞, 霍纬纲, 冯兴杰. 基于模糊关联分类器的民机超限事件诊断方法[J]. 北京航空航天大学学报, 2014, 40(10): 1366-1371.
GAO X X, HUO W G, FENG X J. Civil aircraft's exceedance event diagnosis method based on fuzzy associative classifier[J]. Journal of Beijing University of Aeronautics and Astronautics, 2014, 40(10): 1366-1371(in Chinese).

[9] 祁明亮, 邵雪焱, 池宏. QAR 超限事件飞行操作风险诊断方法[J]. 北京航空航天大学学报, 2011, 37(10): 1207-1210.
QI M L, SHAO X Y, CHI H. Flight operations risk diagnosis method on quick-access-record exceedance[J]. Journal of Beijing University of Aeronautics and Astronautics, 2011, 37(10): 1207-1210(in Chinese).

[10] WANG L, ZHANG J Y, DONG C T, et al. A method of applying flight data to evaluate landing operation performance[J]. Ergonomics, 2019, 62(2): 171-180.

[11] 鲁志东, 张曙光, 戴闰志, 等. 大型民机进近着陆段异常能量风险判据[J]. 航空学报, 2021, 42(6): 624132.

LU Z D, ZHANG S G, DAI R Z, et al. Abnormal energy risk criteria of large civil airplanes in approach and landing[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(6): 624132(in Chinese).

[12] 李哲, 徐浩军, 薛源, 等. 基于风险预测的飞行安全操纵空间构建方法[J]. 北京航空航天大学学报, 2018, 44(9): 1839-1849.
LI Z, XU H J, XUE Y, et al. Construction method of flight safety manipulation space based on risk prediction[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(9): 1839-1849(in Chinese).

[13] 徐骏驰, 孙有朝. 运输类飞机高高原机场运行的适航验证方法[J]. 航空计算技术, 2018, 48(2): 129-134.
XU J C, SUN Y C. Airworthiness requirement of transportation category aircraft operation on high altitude airports[J]. Aeronautical Computing Technique, 2018, 48(2): 129-134(in Chinese).

[14] 中国民用航空局飞行标准司. 飞行品质监控 (FOQA) 实施与管理: AC-121/135-FS-2012-45R1[S]. 北京: 中国民用航空局, 2012.
Flight Standards Department of Civil Aviation Administration of China. Implementation and management of flight operation quality assurance(FOQA): AC-121/135-FS-2012-45R1[S]. Beijing: Civil Aviation Administration of China, 2012(in Chinese).

[15] CUI L F, ZHANG C Q, ZHANG Q Z, et al. A method for aero-engine gas path anomaly detection based on Markov transition field and multi-LSTM[J]. Aerospace, 2021, 8(12): 374.

[16] ATASOY V E, CETEK C. Enhanced cruise range prediction for narrow-body turbofan commercial aircraft based on QAR data[J]. The Aeronautical Journal, 2021, 125(1286): 672-701.

[17] 张鹏, 杨涛, 刘亚楠, 等. 基于 CNN-LSTM 的 QAR 数据特征提取与预测[J]. 计算机应用研究, 2019, 36(10): 2958-2961.
ZHANG P, YANG T, LIU Y N, et al. Feature extraction and prediction of QAR data based on CNN-LSTM[J]. Application Research of Computers, 2019, 36(10): 2958-2961(in Chinese).

[18] 刘柳. 基于 QAR 数据的着陆阶段飞行风险研究[D]. 重庆: 重庆大学, 2018.
LIU L. Research on flight risk in landing using QAR data[D]. Chongqing: Chongqing University, 2018(in Chinese).

[19] 郑磊, 池宏, 邵雪焱. 基于 QAR 数据的飞行操作模式及其风险分析[J]. 中国管理科学, 2017, 25(10): 109-118.
ZHENG L, CHI H, SHAO X Y. Pattern recognition and risk analysis for flight operations[J]. Chinese Journal of Management Science, 2017, 25(10): 109-118(in Chinese).

[20] 李俊辰. 基于机器学习的起飞、进近阶段飞行成绩评估[D]. 天津: 中国民航大学, 2020.
LI J C. Flight performance evaluation during takeoff and approach based on machinelearning[D]. Tianjin: Civil Aviation University of China, 2020(in Chinese).

[21] KONDO S, KIDERA S, KIRIMOTO T. Target detection algorithm using independent component analysis for pulse Doppler radar[J]. IEICE Communications Express, 2014, 3(7): 211-216.

[22] 陈勇刚, 熊升华, 贺强, 等. 通用航空机队设备可靠性动态识别模型[J]. 航空学报, 2018, 39(3): 172-183.
CHEN Y G, XIONG S H, HE Q, et al. Dynamic recognition method for reliability of general aviation fleet equipment[J]. Acta Aeronautica et Astronautica Sinica, 2018, 39(3): 172-183(in Chinese).

[23] 陈守煜. 可变模糊集理论与模型及其应用[M]. 大连: 大连理工大学出版社, 2009.
CHEN S Y. Theory and model of variable fuzzy sets and its application[M]. Dalians: Dalian University of Technology Press, 2009(in Chinese).

[24] KUMAR R K, RAO K S. A novel approach to unsupervised pattern discovery in speech using convolutional neural network[J]. Computer Speech & Language, 2022, 71: 101259.

[25] PRAYUDA M F. Classification of sad emotions and depression through images using convolutional neural network (CNN)[J]. Jurnal Informatika Universitas Pamulang, 2021, 6(1): 90.

[26] 王奕惟, 莫李平, 王奕首, 等. 基于全航段 QAR 数据和卷积神经网络的航空发动机状态辨识[J]. 航空动力学报, 2021, 36(7): 1556-1563.
WANG Y W, MO L P, WANG Y S, et al. Aero-engine status identification based on full-segment QAR data and convolutional neural network[J]. Journal of Aerospace Power, 2021, 36(7): 1556-1563(in Chinese).

[27] 王帅, 黄海鸿, 韩刚, 等. 基于 PCA 与 GA-BP 神经网络的磁记忆信号定量评价[J]. 电子测量与仪器学报, 2018, 32(10): 190-196.
WANG S, HUANG H H, HAN G, et al. Quantitative evaluation of magnetic memory signal based on PCA & GA-BP neural network [J]. Journal of Electronic Measurement and Instrumentation, 2018, 32(10): 190-196(in Chinese).

Risk assessment method for civil aircraft approach and landing at high plateau based on QAR data

CHEN Nongtian[1,*], MAN Yongzheng[2], LI Junhui[2]

(1. College of Aviation Engineering, Civil Aviation Flight University of China, Guanghan 618307, China;
2. College of Civil Aviation Safety Engineering, Civil Aviation Flight University of China, Guanghan 618307, China)

Abstract: The high plateau approach and landing of civil aircraft is a high-risk stage of high plateau flight. To effectively identify the risk and its grade of this approach and landing, a long short term memory-deep neural network (LSTM-DNN) deep learning risk assessment method is proposed based on the variable fuzzy identification of entropy weights. This method utilizes high-altitude flight data recorded by the quick access recorder (QAR), referencing the advisory notices from the flight operations quality assurance (FOQA) of civil aviation as well as the industry QAR monitoring standards. The method combines indicator importance analysis with Delphi expert surveys to extract five key monitoring items for civil aviation high-altitude approach and landing risk assessment, including significant changes in heading during landing, low trajectory, large descent rate during the 610−305 m approach, touchdown vertical acceleration during landing, and high descent rate during the 153−15 m approach. To overcome the subjective bias of the evaluation index weight, the entropy weight method is then used to determine the evaluation index weight, with the risk level membership function constructed based on the variable fuzzy identification method. Finally, the LSTM-DNN risk assessment model for civil aircraft approach and landing at high plateau is established. Taking the Chengdu−Lhasa approach and landing segment as an example, this study extracted the QAR data to train and test the risk assessment model, and compared the results with those of the evaluation methods such as Logistic multiple regression and support vector machines (SVM). The results show that the recognition rate of the proposed method reaches 94.18% on average with the highest being 94.79%, verifying the effectiveness of the method.

Keywords: high plateau flight; quick access recorder data; entropy weight; variable fuzzy identification; LSTM-DNN deep learning; risk assessment

Received: 2022-03-25; **Accepted**: 2022-06-14; **Published Online**: 2022-06-24 14:53

URL: link.cnki.net/urlid/11.2625.V.20220623.1334.003

Foundation items: National Natural Science Foundation of China Civil Aviation Joint Fund Key Project (U2033202); Sichuan Provincial Science and Technology Department Key R & D Program (2022YFG0213)

* **Corresponding author**. E-mail: chennongtian@hotmail.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0211

基于协同滤波轨迹预测的机动目标RTPN拦截制导律

李继广[1, 2, 3, *]，陈欣[2]，董彦非[1]，屈高敏[1]，赵成功[1]，张阿龙[1]

(1. 西安航空学院 飞行器学院，西安 710077；2. 南京航空航天大学 自动化学院，南京 211106；
3. 西北机电工程研究院，咸阳 712000)

摘 要：针对当前空中威胁目标拦截的实际需求，结合拦截器本身的机动能力，基于全覆盖协同策略，提出一种协同探测的现实真比例导引律（RTPN）制导拦截方法。所提方法解决了传统RTPN方法未考虑拦截器饱和过载限制及对任意机动目标捕获区域的确定问题。此外，针对拦截过程中对目标运动轨迹测量误差及协同探测数据丢包所引起的数据融合精度和鲁棒性问题，提出一种分布式协同滤波算法；针对数据传输和拦截器本身动力学响应延迟等问题，提出一种航迹预测算法。仿真结果验证所提方法能够有效解决饱和过载下的捕获区域确定及动力学延迟问题，及协同探测数据融合中数据丢包所引起的鲁棒性和精度问题。

关 键 词：协同拦截；机动目标；现实真比例导引律；过载限制；滤波算法；航迹预测；数据融合

中图分类号：E926.4; V249.3

文献标志码：A **文章编号**：1001-5965(2024)01-0086-11

目前，各种飞行器成为战场、工业、生活常见的设备之一。面对普遍、大量的空中目标，常规拦截手段不仅效率低下，而且成本高昂。如何反制这些空中目标成为当前亟待解决的工程难题之一。

对于实际作战场景下的拦截来说，拦截成本之所以居高不下主要原因有拦截效率低下及拦截器、探测设备成本高昂。这些问题最终又都可归结为拦截导引算法问题。为了解决拦截作战面临的难题，许多学者提出多种解决方案：文献[1]提出一种信息不完全情况下的微分对策制导方法；文献[2]应用有限时间滑模控制和一致性理论设计了具有拦截时间约束的协同制导律；文献[3]设计了一种动能拦截器末制导律；文献[4]采用扩展比例导引律协同突防的问题；文献[5]提出一种机动目标拦截制导方法；文献[6]设计了一种终端角度约束快速终端滑模制导律；文献[7]研究了具有过载约束弱机动拦截器对强动能力目标的协同拦截问题；文献[8]分析了滑模变结构制导律末制导律的捕获区问题；文献[9]研究了多飞行器协同拦截机动目标的分布式协同围捕制导问题；文献[10]设计了一种自适应积分滑模制导律；文献[11]设计了一种广义相对偏置比例制导律；文献[12]设计了导弹逆轨拦截高速机动目标的三维最优制导律。

这些研究成果很大程度上提高了拦截空中威胁的成功率。但是从降低成本和工程实践来考虑，多拦截器协同拦截和比例导引(proportional navigation, PN)方法依然是提高拦截效率、降低成本最有效的方法[13-14]。

收稿日期：2022-04-02；**录用日期**：2022-09-09；**网络出版时间**：2022-10-17 18:10
网络出版地址：link.cnki.net/urlid/11.2625.V.20221017.1350.002
基金项目：陕西省自然科学青年基金项目(2021JQ-858)；陕西省自然科学基金重点项目(2022JZ-37)；陕西高等教育教学改革研究项目(21BY163)；青年科学基金(62103441)；校级科研项目(2021KY0210)
* **通信作者**. E-mail：912646963@qq.com

引用格式：*李继广，陈欣，董彦非，等. 基于协同滤波轨迹预测的机动目标RTPN拦截制导律[J]. 北京航空航天大学学报，2024，50(1)：86-96.*
LI J G，CHEN X，DONG Y F，et al. RTPN interception guidance law for maneuvering target based on collaborative filtering trajectory prediction[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(1)：86-96 (in Chinese).

真比例导引 (true proportional navigation, TPN) 是以视线为参考的制导方法, 现实真比例导引律 (realistic true proportional navigation, RTPN) 是 TPN 方法的工程实现形式[15-20]。RTPN 在数学上容易处理和分析, 文献 [21-24] 对该方法进行了深入的研究。文献 [25] 利用李雅普诺夫类方法证明了三维 RTPN 加速度的上界; 文献 [26] 分析了 RTPN 捕获区域、可移动距离、可移动时间等特性; 文献 [27-28] 证明了 RTPN 具有 3D 空间和 2D 空间同样的拦截捕获能力; 文献 [29] 提出变结构 RTPN 制导律; 文献 [30-31] 设计了一种针对机动目标的 RTPN 制导律。

这些研究成果极大的丰富了 RTPN 方法的内涵和使用场景。但这些研究成果所研究的捕获区域未考虑机动目标时所生成的指令加速度的上界, 及动能拦截器的过载饱和限制, 因而在工程应用上存在着使用限制。

同时, 为了降低高精度探测器带来的成本问题, 多拦截器协同探测方案是常用的手段之一。在拦截成本和拦截效率双重指标下, 多拦截器协同拦截面临着协同探测和协同拦截策略两大难题。多拦截器协同探测需要多种数据进行有效的融合才能达到必须的精度要求。然而, 在实际应用中, 数据融合过程常常面临着数据丢包的问题。如何保证在数据丢包情况下的融合精度, 是协同探测的关键。在多拦截器协同策略上, 最常用的策略是应用一致性原理有限时间控制理论加以解决[32]。该协同策略很好的解决了拦截效率问题, 但是却需要消耗多枚拦截器, 存在拦截成本较高的问题。

针对普遍存在的空中来袭目标进行拦截, 其本质是一种交会制导拦截问题。针对拦截的实际场景, 结合低成本拦截器具体的机动能力限制, 本文分析了多拦截器协同拦截策略, 分析了任意机动目标具有加速度上限和拦截器饱和过载限制下, 任意机动目标能够有效进行拦截的捕获区域, 并针对实际系统具有的测量误差及拦截器本身动力学响应延迟情况提出了协同探测的目标轨迹预测算法, 最后针对协同探测的数据丢包问题, 提出一种协同滤波抗数据丢包的数据融合算法。

1　全覆盖协同拦截策略分析

对于拦截任务而言, 核心需求是确保目标被有效拦截。这就要求拦截器有足够的速度和过载限制下的机动区域。本文把拦截器在捕获能力和机动能力限制下能够进行有效拦截的区域称为捕获区域, 记为A_{F}。对于有一定机动能力的拦截目标, 同样存在着一个机动逃脱的逃逸区域, 记为A_{E}。因此, 只有在拦截器的捕获域A_{F}大于拦截目标的逃逸区域A_{E}时, 才能实现有效的拦截。

对于低成本的拦截器来说, 其捕获域有效, 往往不能对目标的逃逸区域有效覆盖, 这就需要多个拦截器进行协同拦截。基于全覆盖思想, 有 2 种覆盖方式: ①以多枚拦截器均匀的覆盖目标逃逸区域; ②一枚拦截器捕获域的中心针对目标, 其他拦截器对称分布在该拦截器两旁, 以实现对中心拦截器捕获域不能覆盖的区域进行补盲。很明显, 第 1 种覆盖方式在拦截器的总数量上较少。但是第 2 种思路也存在明显的优势: 相较于第 1 种策略多枚拦截器命中目标, 只有一枚拦截器命中目标。在考虑拦截器可回收的前提下, 具有成本方面的优势; 相较于第 1 种覆盖方式需要复杂的协同方式和制导方式, 第 2 种覆盖方式可以采用工程上常用的比例导引及其改进方法, 具有方便实现、贴合工程应用的优势。

因此, 本文采用第 2 种协同拦截方案。从该协同拦截策略可知, 为了实现协同拦截, 首先要确定单个拦截器本身的捕获域A_{F}。拦截器的捕获域主要由拦截器的速度、初始位置和过载上限等自身固有性能因素所决定, 同时受到制导律本身特性的影响。速度、过载上限等属于拦截器的固有属性, 本文不做讨论, 本文主要从制导律的角度进行拦截器可达域的确定。

在上述研究的基础上, 研究多枚拦截器的分配算法, 以达到确定多拦截器中末制导交班阵位、多拦截器各自所覆盖的目标机动区域目的。

为了方便表述, 做如下定义: a_{T}为目标的机动能力; $\tilde{a}_{\mathrm{T}}=a_{\mathrm{T}}+\varepsilon$为目标的瞬时逃逸能力, $|\varepsilon|$值越大, 拦截器进行有效拦截所需要的末端机动过载越大。根据目标自身的机动能力, 其存在一个末端机动区域, 记为U_{T}。对于拦截器而言, 每一枚拦截器同样存在着一个可以有效进行拦截的捕获区域, 对于该区域表示为$U_{\mathrm{M},i}$。则协同拦截的全覆盖策略如下: ①编号为 i 拦截器的捕获区域中心正对目标, 使得该拦截器的捕获区域中心和拦截器的机动区域中心重合; ②编号为 i–1 和 i+1 的拦截器在编号为 i 的拦截器对称分布, 并且捕获区域连接却不重合; ③类似于以上策略, 分别布置 i–2 和 i+2 号拦截器, 直到 $U_{\mathrm{T}}\in\cdots\cap U_{\mathrm{M},i-2}\cap U_{\mathrm{M},i-1}\cap U_{\mathrm{M},i}\cap U_{\mathrm{M},i+1}\cap U_{\mathrm{M},i+2}\cap\cdots$, 即实现了捕获区域的全覆盖。

2　协同拦截机动目标制导律

2.1　RTPN 制导律

在介绍制导律之前, 首先建立拦截器和拦截目

标间的相对运动方程。该拦截过程在空间范围上较小、时间上较短，因此，拦截过程不考虑地球曲率和旋转的影响。同时，将拦截器和拦截目标都做质点，以视线 (LOS) 方向和垂直于视线方向建立坐标系 (e_r,e_θ)，如图 1 所示。r为拦截器和拦截目标相对距离；$v_r=\dot{r}$为拦截器和拦截目标接近速度；$a_r=\ddot{r}$为弹目沿视线方向的相对加速度；q为视线角；$\dot{q}$为视线角速率；$\ddot{q}$为视线角加速度；a_l为拦截器加速度；a_m为拦截目标加速度；$v_\theta=r\dot{q}$为拦截器和拦截目标垂直相对速度。

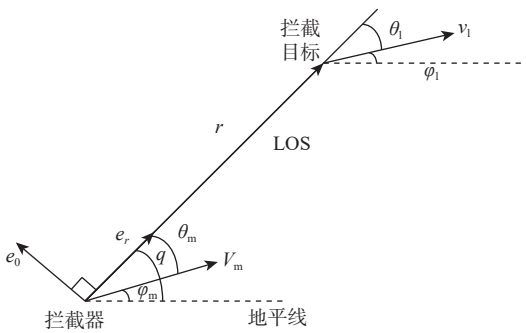


图 1 拦截相对运动坐标系

Fig. 1 Intercept relative motion coordinate system

在该坐标系下，拦截器和拦截目标间相对运动方程可表示为

$$\begin{cases}\ddot{r}-r\dot{q}^2=a_{lr}-a_{mr}\\ r\ddot{q}+2\dot{r}\dot{q}=a_{l\theta}-a_{m\theta}\end{cases}\tag{1}$$

结合实际情况可知，拦截器和拦截目标相对速度较大，且拦截目标沿视线方向的机动对突防所起的作用较小。因此，在设计拦截制导律时仅考虑拦截目标垂直于视线方向的机动加速度。若拦截目标的机动能力满足：

$$\begin{cases}a_{lr}=0\\ a_{l\theta}=a_l(t)\\ |a_l(t)|\leqslant\alpha\end{cases}\tag{2}$$

式中：α为目标机动加速度的上限，为一定制常数，在实际的拦截任务中，该值可以根据拦截器允许拦截目标物的最大值进行确定；t为时间。

RTPN 制导律为加速度垂直于弹目视线，与视线转率成正比，即

$$a_{mr}=0\tag{3}$$

$$a_{m\theta}=N|\dot{r}|\dot{q}\tag{4}$$

式中：N为比例导引系数。

2.2 制导指令加速度上限

在分析制导指令加速度上限之前，首先介绍有效拦截的评判标准-捕获区域。捕获区域指的是目标和拦截器的相对运动状态有关的状态空间。当拦截器和拦截目标相对运动状态进入该空间，则其相对运动最终必然满足如式 (5) 所示的拦截有效约束条件，也就意味着将在某个最终时刻实现对目标拦截捕获。拦截有效约束条件为

$$\begin{cases}0\leqslant r(t_f)<r_{miss}\\ \dot{r}(t_f)<\dot{r}_{Imp}=v_{rImp}<0\end{cases}\tag{5}$$

式中：r_{miss}为容许脱靶量；v_{rImp}为容许碰撞速度。

对于该捕获区域，文献 [9] 得出如下结论：式 (1) 所示的制导系统，对使用如式 (3) 和式 (4) 所示 RTP 的动能拦截器，针对式 (2) 所示的机动目标，如果拦截过程中拦截器和拦截目标接近速度满足：

$$\dot{r}(t)<0\qquad t\in[0,t_0]\tag{6}$$

式中：$t_0\in[0,+\infty]$为常数；比例导引系数满足：

$$N>2\tag{7}$$

则式 (8) 成立：

$$|\dot{r}(t)\dot{q}(t)|\leqslant\delta=\max\left\{|\dot{r}(0)\dot{q}(0)|,\frac{\alpha}{N-2}\right\}\qquad\forall t\in[0,t_0]\tag{8}$$

且 RTPN 制导指令加速度满足：

$$|a_{m\theta}|\leqslant N\delta\qquad\forall t\in[0,t_0]\tag{9}$$

式中：δ为常数。

由该结论可得，当 $N>2$ 时，对任意机动目标，RTPN 指令加速度都不大于 $N/(N-2)$倍的目标最大机动过载。因此，应用该结论及所提前预知的目标最大机动能力水平（即α的值），可合理选择拦截器最大机动过载。

2.3 制导律捕获区分析

2.2 节制导指令加速度上限结论是在拦截目标机动能力有限的条件下得到的结论，即式 (2) 所述的约束条件。但是，式 (5) 所述的目标捕获定义在拦截器和拦截目标相对距离方面的要求过于严苛，而对拦截器和拦截目标接近速度的要求又过于宽泛，在工程应用上缺乏具体的指导意义。为了解决该问题，本文在式 (3) 所示的目标捕获定义下，不考虑拦截器饱和过载限制，RTPN 制导律对机动加速度满足式 (2) 的任意机动目标的捕获区域范围。

本文首先分析不考虑拦截器饱和过载时 RTPN 对任意机动目标的捕获区域的问题。有如下定理。

定理 1 由式 (1) 所确定的制导系统，采用式 (3) 和式 (4) 所示的 RTPN 动能拦截制导律，对式 (2) 确定的机动目标，如果初始相对状态满足条件：

$$\begin{cases}\dot{r}(0)<0\\ \dot{r}^2(0)\left(1-\dfrac{\dot{r}_{Imp}^4}{\dot{r}^4(0)}\right)>2r^2(0)\dot{q}^2(0)\left(1-\dfrac{r_{miss}^2}{r^2(0)}\right)\end{cases}\tag{10}$$

且 RTPN 的比例导引系数满足条件：

$$N > 2 + \alpha \sqrt{\frac{2\left[r^2(0) - r_{\text{miss}}^2\right]}{\dot{r}^4(0) - \dot{r}_{\text{Imp}}^4}} \tag{11}$$

则拦截有效条件为

$$\dot{r}(t) < \dot{r}_{\text{Imp}} \qquad r(t) > r_{\text{miss}};\ \forall t \in [0, +\infty) \tag{12}$$

则必存在一个终端时刻 $t_{\text{f}} \in [0, +\infty)$, 使得式 (3) 成立, 实现对目标的有效拦截。

对于定理 1, 证明如下。

证明　假设式 (12) 不成立, 由于 r 和 $\dot{r}$ 是连续的, 必然存在常数 $t_1 \in (0, +\infty)$ 使得:

$$\dot{r}(t_1) = \dot{r}_{\text{Imp}} \qquad r(t_1) > r_{\text{miss}} \tag{13}$$

由于 $\dot{r}$ 是连续的, 根据式 (10) 和式 (13), 可得

$$\dot{r}(t) < \dot{r}_{\text{Imp}} < 0 \qquad r(t) > r_{\text{miss}};\ \forall t \in [0, t_1) \tag{14}$$

由式 (11) 和式 (14), 根据定理 1 可得

$$|\dot{r}(t)\dot{q}(t)| \leqslant \delta \qquad \forall t \in [0, t_1) \tag{15}$$

将式 (15) 代入式 (1), 当 $a_{tr} = 0$、$a_{mr} = 0$ 时, 有

$$\dot{r}^2\ddot{r} = r(\dot{r}\dot{q})^2 \leqslant r\delta^2 \qquad \forall t \in [0, t_1) \tag{16}$$

由于 $\ddot{r} = \mathrm{d}\dot{r}/\mathrm{d}t = \dot{r}\mathrm{d}\dot{r}/\mathrm{d}r$, 式 (16) 可表示为

$$\dot{r}^3 \frac{\mathrm{d}\dot{r}}{\mathrm{d}r} \leqslant r\delta^2 \qquad \forall t \in [0, t_1) \tag{17}$$

对式 (17) 积分可得

$$\dot{r}^4(t) \geqslant \dot{r}_0^4 + 2\delta^2(r^2(t) - r_0^2) = f(r) \qquad \forall t \in [0, t_1) \tag{18}$$

由于 r、$\dot{r}$、$f(r)$ 是连续的, 则式 (14) 和式 (18) 存在常数 $t_2 \in [0, t_1)$ 使得 $f(r(t_2)) = \dot{r}_{\text{Imp}}^4$ 成立, 即

$$r^2(t_2) = r^2(0) + \frac{\dot{r}_{\text{Imp}}^4 - \dot{r}^4(0)}{2\delta^2} \tag{19}$$

由于 $r(t_2)$ 满足有效拦截条件, 则有 $r^2(t_2) > r_{\text{miss}}^2$, 进而可得

$$\delta^2 > \frac{\dot{r}^4(0) - \dot{r}_{\text{Imp}}^4}{2(r^2(0) - r_{\text{miss}}^2)} \Rightarrow \delta > \sqrt{\frac{\dot{r}^4(0) - \dot{r}_{\text{Imp}}^4}{2(r^2(0) - r_{\text{miss}}^2)}} \tag{20}$$

如下分 2 步论述。

1) 若式 (21) 成立:

$$|\dot{r}(0)\dot{q}(0)| \leqslant \frac{\alpha}{N-2} \tag{21}$$

此时, $\delta = \dfrac{\alpha}{N-2}$, 式 (20) 可表示为

$$N < 2 + \alpha \sqrt{\frac{2(r^2(0) - r_{\text{miss}}^2)}{\dot{r}^4(0) - \dot{r}_{\text{Imp}}^4}} \tag{22}$$

对比式 (22) 和式 (11) 可知, 两式存在明显的矛盾。因此, 假设不成立, 则式 (12) 成立, 且从式 (21) 与式 (22) 可得出式 (10)。

2) 若式 (21) 不成立, 即

$$|\dot{r}(0)\dot{q}(0)| > \frac{\alpha}{N-2} \tag{23}$$

此时, $\delta = |\dot{r}(0)\dot{q}(0)|$, 式 (20) 可表示为

$$\dot{r}^2(0)\left(1 - \frac{\dot{r}_{\text{Imp}}^4}{\dot{r}^4(0)}\right) < 2r^2(0)\dot{q}^2(0)\left(1 - \frac{\dot{r}_{\text{miss}}^2}{r^2(0)}\right) \tag{24}$$

对比式 (24) 和式 (10) 可知, 两式存在明显的矛盾。因此, 假设不成立, 则式 (12) 成立, 且从式 (23) 与式 (10) 可得出式 (11)。　**证毕**

定理 2　实际给出在不考虑拦截器饱和和机动过载时 RTPN 制导律对机动目标的有效捕获区域, 及所对应的比例导引系数 N。若动能拦截器的机动过载具有上限限制时, RTPN 制导律对机动目标的捕获区域则需根据定理 3 加以确定。

定理 3　式 (1) 所确定的制导系统, 采用式 (5) 所示 RTPN 动能拦截器对式 (2) 所确定的机动目标进行拦截, 若拦截器和拦截目标初始相对运动状态满足:

$$|\dot{r}(0)\dot{q}(0)| \leqslant \frac{a_{\max} - \alpha}{2} \tag{25}$$

$$\begin{cases} v_r(0) = \dot{r}(0) < v_{r\text{Imp}} = \dot{r}_{\text{Imp}} < 0 \\ \left(v_r^4(0) - v_{r\text{Imp}}^4\right) > \dfrac{\left(r^2(0) - r_{\text{miss}}^2\right)(a_{\max} - \alpha)^2}{2} \end{cases} \tag{26}$$

且 RTPN 的比例导引系数 N 由式 (27) 确定:

$$\frac{2a_{\max}}{a_{\max} - \alpha} \leqslant N \leqslant \frac{a_{\max}}{|\dot{r}(0)\dot{q}(0)|} \tag{27}$$

则式 (12) 成立, 且在拦截过程中有

$$|a_{m\theta}(t)| \leqslant a_{\max} \qquad \forall t \in [0, +\infty) \tag{28}$$

证明过程如下。

引理　式 (1) 所确定的制导系统, 采用式 (3) 和式 (4) 所示 RTPN 动能拦截器对式 (2) 所示的机动目标进行拦截, 若式 (6) 成立, 且弹目初始相对状态满足:

$$|\dot{r}(0)\dot{q}(0)| \leqslant \frac{a_{\max} - \alpha}{2} \tag{29}$$

且比例导引系数满足:

$$\frac{2a_{\max}}{a_{\max} - \alpha} \leqslant N \leqslant \frac{a_{\max}}{|\dot{r}(0)\dot{q}(0)|} \tag{30}$$

则下式成立:

$$|a_{m\theta}(t)| \leqslant a_{\max} \qquad \forall t \in [0, t_0] \tag{31}$$

由引理可知, 若式 (29) 成立, 则式 (32) 成立, 而比例导引系数可由式 (30) 来确定。

$$\frac{2a_{\max}}{a_{\max} - \alpha} \leqslant 2 + \frac{a}{|\dot{r}(0)\dot{q}(0)|} \leqslant \frac{a_{\max}}{|\dot{r}(0)\dot{q}(0)|} \tag{32}$$

由式 (30) 和式 (31) 可得式 (10) 和式 (11), 则定理 3 成立, 即式 (12) 成立。设 $t_0 = +\infty$, 由式 (30)、式 (31)、式 (12) 成立表明, 引理成立, 即 $t_0 = +\infty$, 式 (31) 成立, 因此, 式 (28) 得证。

由定理3可知,在2D空间平面上,当比例导引系数N由式(30)确定时,RTPN制导律对机动目标的捕获区域为式(27)和式(29)所确定的区域。当拦截相对运动状态在拦截过程中的任意时刻进入该捕获区域后,采用式(30)所确定的RTPN比例导引系数N,则可实现对加速度满足式(2)的任意机动目标进行拦截。

3 分布式卡尔曼滤波轨迹预测

定理2和定理3证明过程忽略了拦截器、地面引导系统测量误差及拦截器本身动力学响应延迟情况。实际上,导引测量误差和动力学延迟是客观存在的。这些因素的存在会使得拦截成功率大为降低。

就整个拦截体系来说,该拦截系统仅具有几个发射装置和简易的三坐标雷达。为了解决导引测量误差和动力学延迟问题,本文采用多机协同测量,拦截目标轨迹预测的方案加以解决。

同时在协同探测的过程中,系统中的某些探测数据来源节点可能因为通信链路、飞行角度、电磁环境等因素而引起数据丢包。为了解决数据丢包情况下,数据融合的鲁棒性和精度,本文采用了分布式卡尔曼滤波算法。

3.1 预测运动模型

对于拦截目标的高程信息,可用三坐标雷达直接提供,而空间方位信息,由拦截器导引装置提供。假设被拦截目标是一个受控系统,拦截器导引装置所测得拦截目标位置状态信息为$\boldsymbol{x}_y(k)=[\,s(k)\quad\theta(k)\,]^{\mathrm{T}}$,其中,$s(k)$、$\theta(k)$分别为拦截器探测到拦截目标的速度、方位信息。同时该拦截目标位置状态信息可表示为平面位置坐标形式$\boldsymbol{x}_y(k)=[\,\boldsymbol{x}(k)\quad\boldsymbol{y}(k)\,]^{\mathrm{T}}$。2种状态方程之间的关系为

$$\begin{aligned}\boldsymbol{x}(k+1)&=\boldsymbol{x}(k)+t_{\mathrm{s}}\boldsymbol{s}(k)\cos\boldsymbol{\theta}(k)\\ \boldsymbol{y}(k+1)&=\boldsymbol{y}(k)+t_{\mathrm{s}}\boldsymbol{s}(k)\sin\boldsymbol{\theta}(k)\end{aligned}\tag{33}$$

若拦截系统中有n个拦截器协同拦截,则第n个拦截器所具有的运动方程可转化为

$$\boldsymbol{x}_n(k+1)=\boldsymbol{x}_n(k)+f_1[\boldsymbol{x}_y(k)]\tag{34}$$

3.2 分布式卡尔曼滤波融合算法

为了解决数据丢包情况下,数据融合的鲁棒性和精度不足的问题,本文采用分布式卡尔曼滤波算法加以解决。分布式卡尔曼滤波融合算法要求在每个采样时刻将式(34)进行线性化。因此,在工作点$\hat{\boldsymbol{x}}_y(k)=[\,\hat{\boldsymbol{s}}(k)\quad\hat{\boldsymbol{\theta}}(k)\,]^{\mathrm{T}}$处进行一阶泰勒近似展开,并代入式(34),可得

$$\boldsymbol{x}(k+1)=\boldsymbol{x}(k)+\boldsymbol{B}(k)\boldsymbol{x}_y(k)+\boldsymbol{\omega}(k)\tag{35}$$

式中:$\boldsymbol{B}(k)=\begin{bmatrix}t_{\mathrm{s}}\cos\hat{\boldsymbol{\theta}}(k) & -t_{\mathrm{s}}\hat{\boldsymbol{s}}(k)\sin\hat{\boldsymbol{\theta}}(k)\\ t_{\mathrm{s}}\sin\hat{\boldsymbol{\theta}}(k) & -t_{\mathrm{s}}\hat{\boldsymbol{s}}(k)\cos\hat{\boldsymbol{\theta}}(k)\end{bmatrix}$,$\boldsymbol{\omega}(k)=\begin{bmatrix}\omega_x(k) & \omega_y(k)\end{bmatrix}^{\mathrm{T}}$为线性化产生的误差。

编号为i的拦截器机载导引装置量测方程为

$$\boldsymbol{y}_i(k)=\boldsymbol{H}_i\boldsymbol{x}_i(k)+\boldsymbol{v}_i(k)\tag{36}$$

式中:$\boldsymbol{y}_i(k)$、$\boldsymbol{H}_i$和$\boldsymbol{v}_i(k)$分别为量测向量、观测矩阵和传感器噪声;$\boldsymbol{v}_i(k)=\begin{bmatrix}v_{xi}(k) & v_{yi}(k)\end{bmatrix}^{\mathrm{T}}$,其协方差矩阵为$\boldsymbol{V}$。

若有n架拦截器协同对同一拦截目标进行测量,则量测方程可扩展为

$$\begin{bmatrix}y_{i1}(k)\\ y_{i2}(k)\\ \vdots\\ y_{in}(k)\end{bmatrix}=\begin{bmatrix}H_{i1}(k)\\ H_{i2}(k)\\ \vdots\\ H_{in}(k)\end{bmatrix}\boldsymbol{x}(k)+\begin{bmatrix}v_{i1}(k)\\ v_{i2}(k)\\ \vdots\\ v_{in}(k)\end{bmatrix}\tag{37}$$

得出状态空间模型动态系统:

$$\boldsymbol{x}_{k+1}^i=\boldsymbol{A}_k^i\boldsymbol{x}_k^i+\boldsymbol{B}_k^i\boldsymbol{\omega}_k^i\tag{38}$$

$$\boldsymbol{z}_k^i=\boldsymbol{C}_k^i\boldsymbol{x}_k+\boldsymbol{v}_k^i\tag{39}$$

$$\boldsymbol{x}_{k+1}=\boldsymbol{A}_k\boldsymbol{x}_k+\boldsymbol{B}_k\boldsymbol{\omega}_k\tag{40}$$

$$\boldsymbol{z}_k=\boldsymbol{C}_k\boldsymbol{x}_k+\boldsymbol{v}_k\tag{41}$$

假设$\boldsymbol{\omega}(t)$和$\boldsymbol{v}(t)$是零均值,方差阵为$\boldsymbol{Q}$和$\boldsymbol{R}$的不相关白噪声:

$$\boldsymbol{E}\left\{\boldsymbol{\omega}_k^2\right\}=\boldsymbol{0}$$

$$\boldsymbol{E}\left\{\boldsymbol{v}_k^2\right\}=\boldsymbol{0}$$

$$\boldsymbol{E}\left\{\boldsymbol{\omega}_k\,\boldsymbol{\omega}_k^{i^2}\right\}=\boldsymbol{Q}_k$$

$$\boldsymbol{E}\left\{\boldsymbol{v}_k\,\boldsymbol{v}_k^{i^2}\right\}=\boldsymbol{R}_k$$

$$\boldsymbol{E}\left\{\boldsymbol{\omega}_k\,\boldsymbol{v}_k^{i^2}\right\}=\boldsymbol{S}_k$$

对于式(38)~式(41)描述的系统,根据递推射影公式,有如下递推关系:

$$\hat{\boldsymbol{x}}_{k+1|k+1}=\hat{\boldsymbol{x}}_{k+1|k}+\boldsymbol{K}_{k+1}\boldsymbol{\varepsilon}_{k+1}\tag{42}$$

$$\boldsymbol{K}_{k+1}=\boldsymbol{E}[\boldsymbol{x}_{k+1}\boldsymbol{\varepsilon}_{k+1}^{\mathrm{T}}]\boldsymbol{E}[\boldsymbol{\varepsilon}_{k+1}\boldsymbol{\varepsilon}^{\mathrm{T}}{}_{k+1}]^{-1}\tag{43}$$

式中:$\boldsymbol{K}_{k+1}$为Kalman滤波器增益。

对式(40)两边取射影可得

$$\hat{\boldsymbol{x}}_{k+1|k}=\boldsymbol{A}_k\hat{\boldsymbol{x}}_{k|k}+\boldsymbol{B}_k\mathrm{proj}(\boldsymbol{\omega}_k|y_1,y_2,\cdots,y_t)\tag{44}$$

由式(40)迭代可得

$$\boldsymbol{x}_k\in\boldsymbol{L}(\boldsymbol{\omega}_{k-1},\cdots,\boldsymbol{\omega}(0),\boldsymbol{x}(0))\tag{45}$$

同理,应用于式(41)有

$$\boldsymbol{z}\in\boldsymbol{L}(\boldsymbol{v}_k,\boldsymbol{\omega}_{k-1},\cdots,\boldsymbol{\omega}(0),\boldsymbol{x}(0))\tag{46}$$

则有

$$\boldsymbol{L}(z_1,z_2,\cdots,z_k)\subset\boldsymbol{L}(\boldsymbol{v}_k,\cdots,\boldsymbol{v}_1,\boldsymbol{\omega}_{k-1},\cdots,\boldsymbol{\omega}(0),\boldsymbol{x}(0)) \tag{47}$$

由式 (47) 得

$$\boldsymbol{\omega}_k\perp\boldsymbol{L}(z_1,z_2,\cdots,z_k) \tag{48}$$

应用影射公式和$\boldsymbol{E}\left\{\boldsymbol{\omega}_k^2\right\}=0$得

$$\mathrm{proj}(\boldsymbol{\omega}_k|z_1,z_2,\cdots,z_k)=0 \tag{49}$$

则式 (44) 可表示为

$$\hat{\boldsymbol{x}}_{k+1|k}=\boldsymbol{A}_k\hat{\boldsymbol{x}}_{k|k} \tag{50}$$

同理对式 (41) 两边取射影可得

$$\hat{\boldsymbol{z}}_{k|k-1}=\boldsymbol{C}_k\hat{\boldsymbol{x}}_{k|k-1} \tag{51}$$

由式 (48) 可知:

$$\boldsymbol{v}_{k+1}\perp\boldsymbol{L}(z_1,z_2,\cdots,z_k) \tag{52}$$

有

$$\mathrm{proj}(\boldsymbol{v}_{k+1}|z_1,z_2,\cdots,z_k)=0 \tag{53}$$

可得

$$\hat{\boldsymbol{z}}_{k+1|k}=\boldsymbol{C}_k\hat{\boldsymbol{x}}_{k+1|k} \tag{54}$$

从而得到新息序列$\boldsymbol{\varepsilon}_{k+1}$为

$$\begin{aligned}\boldsymbol{\varepsilon}_{k+1}&=\boldsymbol{z}_{k+1}-\widehat{\boldsymbol{z}}_{k+1|k}=a_{k+1}\boldsymbol{C}_{k+1}\boldsymbol{x}_{k+1}+\\&a_{k+1}\boldsymbol{\nu}_{k+1}+(1-a_{k+1})\boldsymbol{C}_{k+1}\widehat{\boldsymbol{x}}_{k+1|k}-\boldsymbol{C}_{k+1}\widehat{\boldsymbol{x}}_{k+1|k}=\\&a_{k+1}\boldsymbol{C}_{k+1}\boldsymbol{x}_{k+1}+a_{k+1}\boldsymbol{\nu}_{k+1}-a_{k+1}\boldsymbol{C}_{k+1}\widehat{\boldsymbol{x}}_{k+1|k}\end{aligned} \tag{55}$$

滤波估值误差及方差阵为

$$\tilde{\boldsymbol{x}}_{k|k}=\boldsymbol{x}_k-\hat{\boldsymbol{x}}_{k|k} \tag{56}$$

$$\tilde{\boldsymbol{x}}_{k+1|k}=\boldsymbol{x}_{k+1}-\hat{\boldsymbol{x}}_{k+1|k} \tag{57}$$

$$\boldsymbol{P}_{k|k}=\boldsymbol{E}\left[\tilde{\boldsymbol{x}}_{k|k}\tilde{\boldsymbol{x}}_{k|k}^{\mathrm{T}}\right] \tag{58}$$

$$\boldsymbol{P}_{k+1|k}=\boldsymbol{E}\left[\tilde{\boldsymbol{x}}_{k+1|k}\tilde{\boldsymbol{x}}_{k+1|k}^{\mathrm{T}}\right] \tag{59}$$

将式 (56)代入式 (55) 得

$$\begin{aligned}\boldsymbol{\varepsilon}_{k+1}&=a_{k+1}\boldsymbol{C}_{k+1}\boldsymbol{x}_{k+1}+a_{k+1}\boldsymbol{\nu}_{k+1}-a_{k+1}\boldsymbol{C}_{k+1}\widehat{\boldsymbol{x}}_{k+1|k}=\\&a_{k+1}\boldsymbol{C}_{k+1}\boldsymbol{x}_{k+1}-a_{k+1}\boldsymbol{C}_{k+1}\widehat{\boldsymbol{x}}_{k+1|k}+a_{k+1}\boldsymbol{\nu}_{k+1}=\\&a_{k+1}\boldsymbol{C}_{k+1}\left(\boldsymbol{x}_{k+1}-\widehat{\boldsymbol{x}}_{k+1|k}\right)+a_{k+1}\boldsymbol{\nu}_{k+1}=\\&a_{k+1}\boldsymbol{C}_{k+1}\tilde{\boldsymbol{x}}_{k+1|k}+a_{k+1}\boldsymbol{\nu}_{k+1}\end{aligned} \tag{60}$$

将式 (53) 代入式 (57) 可得

$$\begin{aligned}\tilde{\boldsymbol{x}}_{k+1|k}&=\boldsymbol{x}_{k+1}-\widehat{\boldsymbol{x}}_{k+1|k}=\boldsymbol{A}_k\boldsymbol{x}_k+\boldsymbol{B}_k\boldsymbol{w}_k-\boldsymbol{A}_k\widehat{\boldsymbol{x}}_{k|k}=\\&\boldsymbol{A}_k\left(\boldsymbol{x}_k-\widehat{\boldsymbol{x}}_{k|k}\right)+\boldsymbol{B}_k\boldsymbol{w}_k=\boldsymbol{A}_k\tilde{\boldsymbol{x}}_{k|k}+\boldsymbol{B}_k\boldsymbol{w}_k\end{aligned} \tag{61}$$

由式 (42) 得

$$\begin{aligned}\tilde{\boldsymbol{x}}_{k+1|k+1}&=\boldsymbol{x}_{k+1}-\hat{\boldsymbol{x}}_{k+1|k+1}=\boldsymbol{x}_{k+1}-\hat{\boldsymbol{x}}_{k+1|k}-\boldsymbol{K}_{k+1}\boldsymbol{\varepsilon}_{k+1}=\\&\tilde{\boldsymbol{x}}_{k+1|k}-\boldsymbol{K}_{k+1}\boldsymbol{\varepsilon}_{k+1}\end{aligned} \tag{62}$$

将式 (60) 代入式 (62) 得

$$\tilde{\boldsymbol{x}}_{k+1|k+1}=\tilde{\boldsymbol{x}}_{k+1|k}-\boldsymbol{K}_{k+1}\left(a_{k+1}\boldsymbol{C}_{k+1}\tilde{\boldsymbol{x}}_{k+1|k}\right)-\boldsymbol{K}_{k+1}\left(a_{k+1}\boldsymbol{\nu}_{k+1}\right) \tag{63}$$

由式 (59) 可得

$$\begin{aligned}\boldsymbol{P}_{k+1|k}&=\boldsymbol{E}\left[\tilde{\boldsymbol{x}}_{k+1|k}\tilde{\boldsymbol{x}}_{k+1|k}^{\mathrm{T}}\right]=\\&\boldsymbol{E}\left[\left(\boldsymbol{A}_k\tilde{\boldsymbol{x}}_{k|k}+\boldsymbol{B}_k\boldsymbol{\omega}_k\right)\left(\boldsymbol{A}_k\tilde{\boldsymbol{x}}_{k|k}+\boldsymbol{B}_k\boldsymbol{\omega}_k\right)^{\mathrm{T}}\right]=\\&\boldsymbol{E}\left[\left(\boldsymbol{A}_k\tilde{\boldsymbol{x}}_{k|k}\right)\left(\boldsymbol{A}_k\tilde{\boldsymbol{x}}_{k|k}\right)^{\mathrm{T}}\right]+\boldsymbol{E}\left[\left(\boldsymbol{A}_k\tilde{\boldsymbol{x}}_{k|k}\right)(\boldsymbol{B}_k\boldsymbol{\omega}_k)^{\mathrm{T}}\right]+\\&\boldsymbol{E}\left[(\boldsymbol{B}_k\boldsymbol{\omega}_k)\left(\boldsymbol{A}_k\tilde{\boldsymbol{x}}_{k|k}\right)^{\mathrm{T}}\right]+\boldsymbol{E}\left[(\boldsymbol{B}_k\boldsymbol{\omega}_k)(\boldsymbol{B}_k\boldsymbol{\omega}_k)^{\mathrm{T}}\right]=\\&\boldsymbol{A}_k\boldsymbol{P}_{k|k}\boldsymbol{A}_k^{\mathrm{T}}+\boldsymbol{B}_k\boldsymbol{Q}\boldsymbol{B}_k^{\mathrm{T}}\end{aligned} \tag{64}$$

对于$\boldsymbol{P}_{k+1|k+1}$, 由式 (59) 及式 (63) 得

$$\begin{aligned}\boldsymbol{P}_{k+1|k+1}&=\boldsymbol{E}\left[\left(\tilde{\boldsymbol{x}}_{k+1|k}-\boldsymbol{K}_{k+1}\left(a_{k+1}\boldsymbol{C}_{k+1}\tilde{\boldsymbol{x}}_{k+1|k}+a_{k+1}\boldsymbol{\nu}_{k+1}\right)\right)\times\right.\\&\left.\left(\tilde{\boldsymbol{x}}_{k+1|k}-\boldsymbol{K}_{k+1}\left(a_{k+1}\boldsymbol{C}_{k+1}\tilde{\boldsymbol{x}}_{k+1|k}+a_{k+1}\boldsymbol{\nu}_{k+1}\right)\right)^{\mathrm{T}}\right]=\\&\boldsymbol{E}\left[\left(\tilde{\boldsymbol{x}}_{k+1|k}\right)\left(\tilde{\boldsymbol{x}}_{k+1|k}\right)^{\mathrm{T}}\right]+\lambda\boldsymbol{E}\left[\left(\tilde{\boldsymbol{x}}_{k+1|k}\right)\left(\boldsymbol{K}_{k+1}\boldsymbol{C}_{k+1}\tilde{\boldsymbol{x}}_{k+1|k}\right)^{\mathrm{T}}\right]-\\&\lambda\boldsymbol{E}\left[\left(\boldsymbol{K}_{k+1}\boldsymbol{C}_{k+1}\tilde{\boldsymbol{x}}_{k+1|k}\right)\left(\tilde{\boldsymbol{x}}_{k+1|k}\right)^{\mathrm{T}}\right]-\\&\lambda\boldsymbol{E}\left[\left(\boldsymbol{K}_{k+1}\boldsymbol{C}_{k+1}\tilde{\boldsymbol{x}}_{k+1|k}\right)\left(\boldsymbol{K}_{k+1}\boldsymbol{C}_{k+1}\tilde{\boldsymbol{x}}_{k+1|k}\right)^{\mathrm{T}}\right]-\\&\lambda\boldsymbol{E}\left[(\boldsymbol{K}_{k+1}\boldsymbol{\nu}_{k+1})(\boldsymbol{K}_{k+1}\boldsymbol{\nu}_{k+1})^{\mathrm{T}}\right]=\\&\boldsymbol{P}_{k+1|k}-\lambda\boldsymbol{P}_{k+1|k}(\boldsymbol{K}_{k+1}\boldsymbol{C}_{k+1})^{\mathrm{T}}-\lambda\boldsymbol{K}_{k+1}\boldsymbol{C}_{k+1}\boldsymbol{P}_{k+1|k}+\\&\lambda\boldsymbol{K}_{k+1}\boldsymbol{C}_{k+1}\boldsymbol{P}_{k+1|k}(\boldsymbol{K}_{k+1}\boldsymbol{C}_{k+1})^{\mathrm{T}}+\lambda\boldsymbol{K}_{k+1}\boldsymbol{R}_{k+1}\boldsymbol{K}_{k+1}^{\mathrm{T}}\end{aligned} \tag{65}$$

有

$$\begin{aligned}\boldsymbol{E}\left[\boldsymbol{\varepsilon}_{k+1}\boldsymbol{\varepsilon}_{k+1}^{\mathrm{T}}\right]&=\boldsymbol{E}\left[\left(a_{k+1}\boldsymbol{C}_{k+1}\tilde{\boldsymbol{x}}_{k+1|k}+a_{k+1}\boldsymbol{\nu}_{k+1}\right)\times\right.\\&\left.\left(a_{k+1}\boldsymbol{C}_{k+1}\tilde{\boldsymbol{x}}_{k+1|k}+a_{k+1}\boldsymbol{\nu}_{k+1}\right)^{\mathrm{T}}\right]=\\&\lambda\boldsymbol{E}\left[\left(\boldsymbol{C}_{k+1}\tilde{\boldsymbol{x}}_{k+1|k}\right)\left(\boldsymbol{C}_{k+1}\tilde{\boldsymbol{x}}_{k+1|k}\right)^{\mathrm{T}}\right]+\\&\lambda\boldsymbol{E}\left[(a_{k+1}\boldsymbol{\nu}_{k+1})(a_{k+1}\boldsymbol{\nu}_{k+1})^{\mathrm{T}}\right]=\\&\lambda\left(\boldsymbol{C}_{k+1}\boldsymbol{P}_{k+1|k}\boldsymbol{C}_{k+1}^{\mathrm{T}}+\boldsymbol{R}_{k+1}\right)\end{aligned} \tag{66}$$

由式 (57) 可得

$$\begin{aligned}\boldsymbol{E}\left[\boldsymbol{x}_{k+1}\boldsymbol{\varepsilon}_{k+1}^{\mathrm{T}}\right]&=\boldsymbol{E}\left[\left(\tilde{\boldsymbol{x}}_{k+1|k}+\hat{\boldsymbol{x}}_{k+1|k}\right)\times\right.\\&\left.\left(a_{k+1}\boldsymbol{C}_{k+1}\tilde{\boldsymbol{x}}_{k+1|k}+a_{k+1}\boldsymbol{\nu}_{k+1}\right)^{\mathrm{T}}\right]=\\&\lambda\boldsymbol{E}\left[\left(\tilde{\boldsymbol{x}}_{k+1|k}\right)\left(\boldsymbol{C}_{k+1}\tilde{\boldsymbol{x}}_{k+1|k}\right)^{\mathrm{T}}\right]=\lambda\boldsymbol{P}_{k+1|k}\boldsymbol{C}_{k+1}^{\mathrm{T}}\end{aligned} \tag{67}$$

对于$\boldsymbol{K}_{k+1}$, 将式 (66) 和式 (67) 代入式 (43) 得

$$\begin{aligned}\boldsymbol{K}_{k+1}&=\lambda\boldsymbol{P}_{k+1|k}\boldsymbol{C}_{k+1}^{\mathrm{T}}\left(\lambda\left(\boldsymbol{C}_{k+1}\boldsymbol{P}_{k+1|k}\boldsymbol{C}_{k+1}^{\mathrm{T}}+\boldsymbol{R}_{k+1}\right)\right)^{-1}=\\&\boldsymbol{P}_{k+1|k}\boldsymbol{C}_{k+1}^{\mathrm{T}}\left(\boldsymbol{C}_{k+1}\boldsymbol{P}_{k+1|k}\boldsymbol{C}_{k+1}^{\mathrm{T}}+\boldsymbol{R}_{k+1}\right)^{-1}\end{aligned} \tag{68}$$

由K_{k+1}简化式 (65) 得

$$\begin{aligned}\boldsymbol{P}_{k+1|k+1}&=\boldsymbol{P}_{k+1|k}-\lambda\boldsymbol{K}_{k+1}\boldsymbol{C}_{k+1}\boldsymbol{P}_{k+1|k}-\lambda\boldsymbol{P}_{k+1|k}(\boldsymbol{K}_{k+1}\boldsymbol{C}_{k+1})^{\mathrm{T}}+\\&\lambda\boldsymbol{K}_{k+1}\left(\boldsymbol{C}_{k+1}\boldsymbol{P}_{k+1k}\boldsymbol{C}_{k+1}^{\mathrm{T}}+\boldsymbol{R}_{k+1}\right)\boldsymbol{K}_{k+1}^{\mathrm{T}}=\\&\boldsymbol{P}_{k+1|k}-\lambda\boldsymbol{K}_{k+1}\boldsymbol{C}_{k+1}\boldsymbol{P}_{k+1k}-\lambda\boldsymbol{P}_{k+1k}(\boldsymbol{K}_{k+1}\boldsymbol{C}_{k+1})^{\mathrm{T}}+\\&\lambda\boldsymbol{P}_{k+1k}\boldsymbol{C}_{k+1}^{\mathrm{T}}\boldsymbol{K}_{k+1}^{\mathrm{T}}=\boldsymbol{P}_{k+1k}-\lambda\boldsymbol{K}_{k+1}\boldsymbol{C}_{k+1}\boldsymbol{P}_{k+1k}\end{aligned} \tag{69}$$

最终得到式 (70)～式 (75) 所示系统的递推 Kalman 滤波器为

$$\hat{\boldsymbol{x}}_{k+1|k+1}=\hat{\boldsymbol{x}}_{k+1|k}+\boldsymbol{K}_{k+1}\left[\boldsymbol{C}_{k+1}\boldsymbol{P}_{k+1|k}+\boldsymbol{v}_{k+1}\right] \tag{70}$$

$$\hat{\boldsymbol{x}}_{k+1|k}=\boldsymbol{A}_k\hat{\boldsymbol{x}}_{k|k} \tag{71}$$

$$\boldsymbol{\rho}_{k+1|k}=\boldsymbol{A}_k\boldsymbol{P}_{k|k}-\boldsymbol{B}_k\boldsymbol{\omega}_k \tag{72}$$

$$\boldsymbol{\rho}_{k+1|k+1}=\boldsymbol{P}_{k+1|k}-\boldsymbol{K}_{k+1}\boldsymbol{C}_{k+1}\boldsymbol{P}_{k+1|k}-\boldsymbol{K}_{k+1}\boldsymbol{v}_{k+1} \tag{73}$$

$$\boldsymbol{P}_{k+1|k}=\boldsymbol{A}_k\boldsymbol{P}_{k|k}\boldsymbol{A}_k^{\mathrm{T}}-\boldsymbol{A}_k\boldsymbol{K}_k\boldsymbol{S}_k^{\mathrm{T}}\boldsymbol{B}_k^{\mathrm{T}}-\boldsymbol{B}_k\boldsymbol{S}_k\boldsymbol{K}_k^{\mathrm{T}}\boldsymbol{A}_k^{\mathrm{T}}+\boldsymbol{B}_k\boldsymbol{Q}_k\boldsymbol{B}_k^{\mathrm{T}} \tag{74}$$

$$\begin{aligned}\boldsymbol{P}_{k+1|k+1}=&\boldsymbol{P}_{k+1|k}-\boldsymbol{K}_{k+1}\boldsymbol{C}_{k+1}\boldsymbol{P}_{k+1|k}-\boldsymbol{P}_{k+1|k}\boldsymbol{C}_{k+1}^{\mathrm{T}}\boldsymbol{K}_{k+1}^{\mathrm{T}}+\\&\boldsymbol{K}_{k+1}\boldsymbol{C}_{k+1}\boldsymbol{P}_{k+1|k}\boldsymbol{C}_{k+1}^{\mathrm{T}}\boldsymbol{K}_{k+1}^{\mathrm{T}}+\boldsymbol{K}_{k+1}\boldsymbol{R}_{k+1}\boldsymbol{K}_{k+1}^{\mathrm{T}}\end{aligned} \tag{75}$$

3.3 轨迹预测

为了表述方便，将式 (37) 表述为

$$\boldsymbol{y}_n(k)=\boldsymbol{H}_n\boldsymbol{x}(k)+\boldsymbol{v}_n(k) \tag{76}$$

若拦截器系统动力学延迟时间为常数τ，则需要根据当前测量的数据预测τ之后拦截目标所处的位置。于是将系统的动力学延迟问题转化为对拦截目标的航迹预测问题。

对于每个预测步长为τ的预测问题,为得到误差协方差最小的最优解：

$$\boldsymbol{S}^2(k+\tau|k)=\mathrm{var}[\boldsymbol{x}(k+\tau)-\boldsymbol{x}(k+\tau|k)] \tag{77}$$

式中：$\boldsymbol{x}(k+\tau)$和$\boldsymbol{x}(k+\tau|k)$分别为拦截目标的测量位置和预测值；var 为协方差运算。

$\boldsymbol{x}(k+\tau)$可以通过下式求得:

$$\begin{cases}\boldsymbol{x}(k+\tau+1)=g[\boldsymbol{x}(k+\tau|k)]+\boldsymbol{\omega}(k+\tau)\\ \boldsymbol{y}_i(k+\tau)=\boldsymbol{H}_i[(k+\tau)]\boldsymbol{x}(k+\tau)\end{cases} \tag{78}$$

式中：$\boldsymbol{H}_i$为观测矩阵；g为递推运算。

选择$x(k+\tau|k)$使式 (35) 能够产生最优的轨迹预测值，即

$$\boldsymbol{x}(k+\tau|k)=\begin{bmatrix}\boldsymbol{H}_i(k)\hat{\boldsymbol{x}}(k+\tau|k)\\ \boldsymbol{H}_i(k)\boldsymbol{g}[\hat{\boldsymbol{x}}(k+\tau|k)]\\ \vdots\\ \boldsymbol{H}_i(k)\boldsymbol{g}\{\cdots\boldsymbol{g}[\hat{\boldsymbol{x}}(k+\tau|k)]\cdots\}\end{bmatrix} \tag{79}$$

式中：$\hat{\boldsymbol{x}}(k+\tau|k)$为拦截目标的估计值。

4 仿真验证

4.1 分布式卡尔曼滤波融合算法验证

目标拦截需要精确探测目标的运动轨迹。本文采用 3 架拦截器协同探测的方式进行协同探测，并用本文的分布式卡尔曼滤波融合算法进行数据处理。机载探测设备采样周期为 10 Hz。仿真结果如图 2 和图 3 所示。从图 2 和图 3 可知，相较于参考轨迹，3 架无人机单独测量的结果都存在一定的误差。采用本文协同融合算法将 3 架无人机探测的数据融合之后，融合预测得到的轨迹和参考轨迹具有较高的一致性。仿真结果显示，本文融合算法可以有效的提高测量精度，消除噪声干扰。

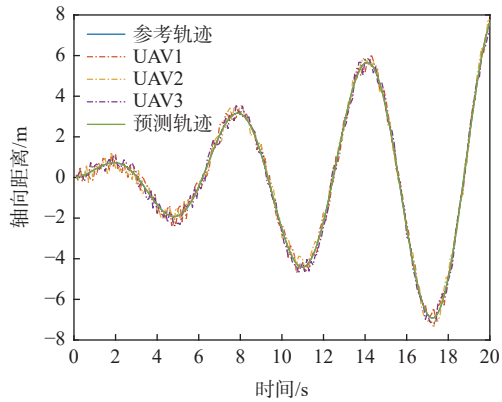


图 2 轴向位置探测和预测效果

Fig. 2 Axial position detection and prediction effect

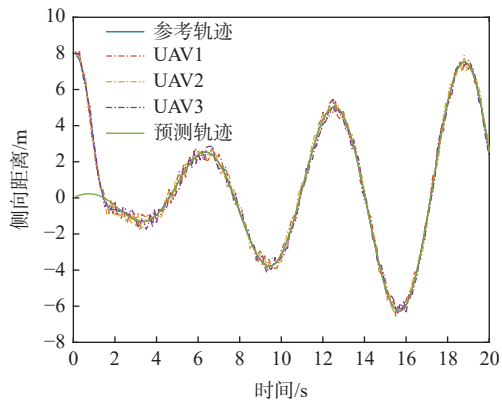


图 3 侧向位置探测和预测效果

Fig. 3 Lateral position detection and prediction effect

在实际拦截过程中，协同拦截的拦截器探测数据可能存在缺失。本文融合算法的优势之一是可以满足数据丢包情况下融合结果的鲁棒性和精度。在数据丢包 10% 和 20% 的情况下，数据融合结果如图 4 所示。

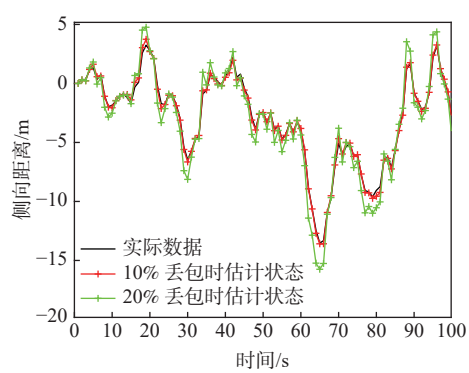


图 4 具有数据丢包情况下的融合结果

Fig. 4 Fusion results with data packet loss

从仿真结果可知, 在数据丢包的情况下, 融合后的结果会有所偏差, 且随着丢包量增大而增大。但是融合结果仍然和实际结果相吻合, 融合结果的有效性得到了保证。仿真结果证明本文融合算法在丢包情况下的有效性。

4.2　延迟轨迹预测

在实际过程中, 数据的传输存在一定的延迟。成功的拦截需要精确的探测到当前目标的位置信息。本文通过轨迹预测算法来解决数据传输延迟的问题。若系统延迟时间常数为 0.2 s, 采用本文预测算法仿真结果如图 5 所示。

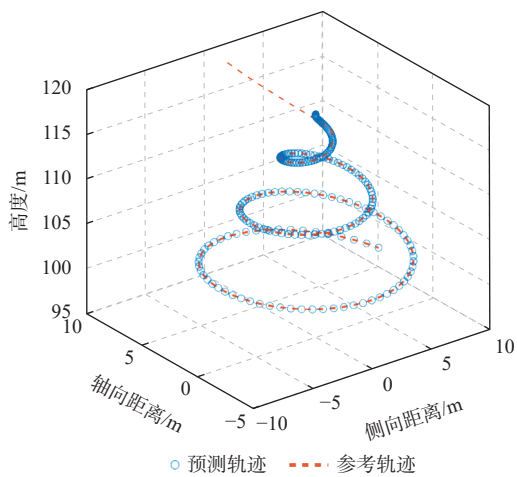


图 5　拦截目标飞行轨迹预测

Fig. 5　Trajectory prediction for intercepting target

从图 5 可知, 本文预测算法在系统延迟 0.2 s 的情况下, 可以精确的预测障碍物的飞行轨迹, 使得预测得到的飞行轨迹和参考轨迹基本重合。证明本文预测算法可以有效对拦截目标飞行航迹进行预测, 消除了系统延迟的影响。

4.3　RTPN 制导拦截律验证

验证拦截器捕获区域大于目标机动区域时的拦截情况。本文拦截制导律是基于拦截器和拦截目标在同一高度上进行推导的, 为验证拦截效果, 本文在同一高度下进行拦截仿真验证。就拦截器的机动能力而言, 其机动能力如下: 法向机动能力不大于 3g, 径向机动能力不大于 1g; 拦截目标限制过载为 2.5g。当比例系数为 0.5、1.0、2.0、2.5 时的仿真效果如图 6～图 8 所示。

从仿真效果图 6 可知, 拦截器成功对拦截目标进行有效拦截, 成功摧毁目标。从图 7 可知, 拦截过程中, 拦截器法向过载平稳, 无剧烈波动。从图 8 可知, 拦截过程中, 拦截器径向过载逐渐收敛为 0, 实现等速平稳拦截目标。

现实中, 拦截目标的运动是空间中三维运动。

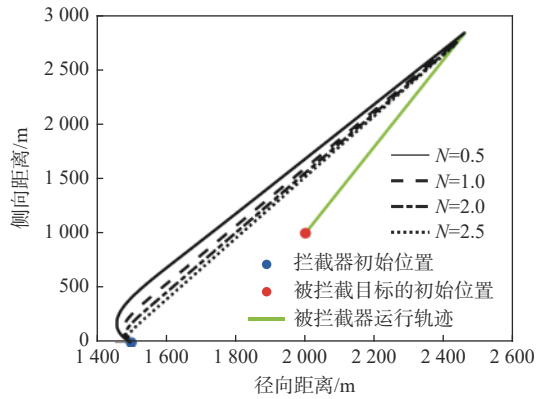


图 6　拦截器和拦截目标拦截运动轨迹

Fig. 6　Motion trajectory of Interceptor and interceptor target interceptor

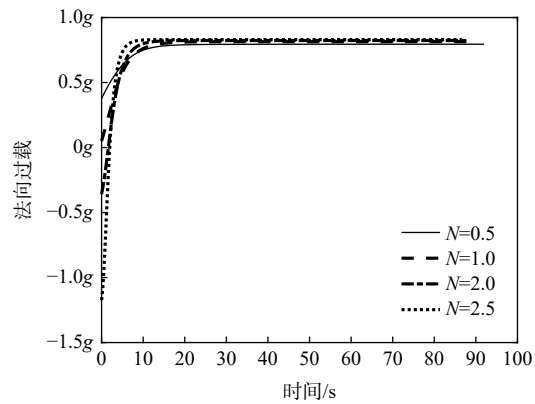


图 7　拦截器拦截过程法向过载

Fig. 7　Normal overload in intercepting process of Interceptor

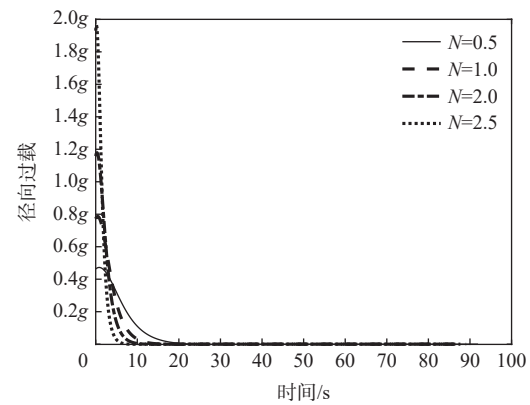


图 8　拦截器拦截过程径向过载

Fig. 8　Radial overload in intercepting process of Interceptor

为验证对具有三维空间机动能力拦截目标的拦截效果, 本文进行了三维环境下的拦截仿真验证。仿真效果如图 9～图 11 所示。

从仿真效果图 9 可知, 拦截器成功对具有三维机动能力的拦截目标进行有效拦截, 并成功摧毁目标。从图 10 可知, 拦截过程中, 拦截器法向过载平稳, 无剧烈波动。从图 11 可知, 拦截过程中, 拦截

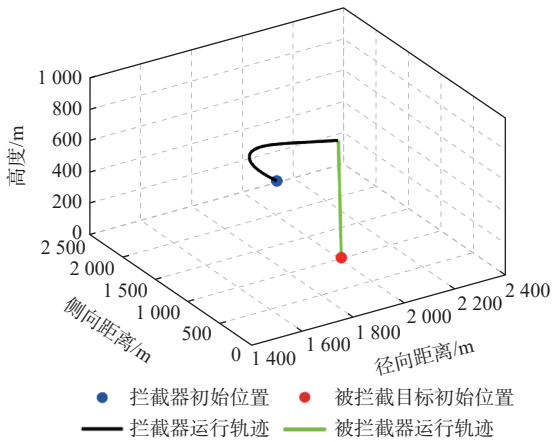


图9 拦截器和拦截目标三维拦截运动轨迹

Fig. 9 3D intercept motion trajectory of Interceptor and intercepting target

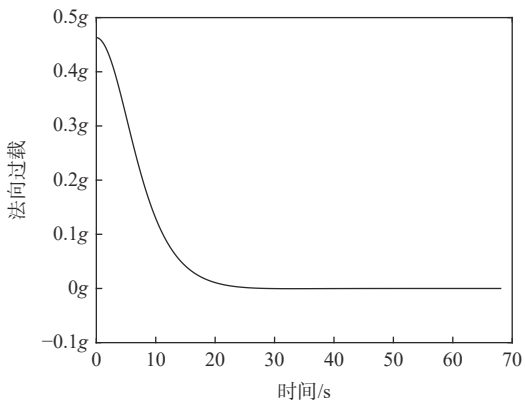


图10 拦截器三维拦截过程法向过载

Fig. 10 Normal overloading of Interceptor during 3-D interception process

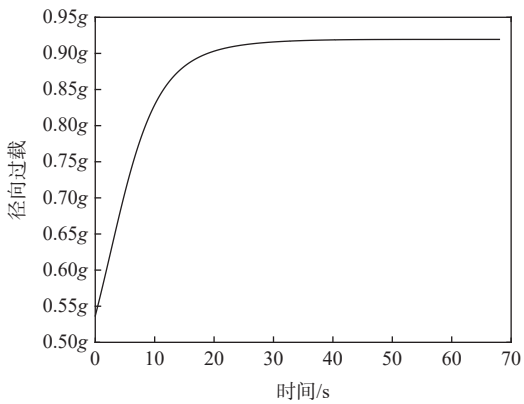


图11 拦截器三维拦截过程径向过载

Fig. 11 Radial overload of Interceptor during 3-D interception process

器径向过载逐渐收敛为0,实现了等速平稳拦截的目标。

从仿真结果可知,本文方法有效地解决了RTPN制导方法的饱和过载限制、有效捕获问题和测量误差及拦截器本身动力学响应延迟等问题。

4.4 全覆盖策略验证

当拦截器捕获区域小于目标机动区域时,单一拦截器就不能进行有效拦截。此时,就有必要通过全覆盖方法进行有效拦截。本文采用5架拦截器进行协同拦截。其中,拦截器3捕捉区域中心正对目标,拦截器1～拦截器5在拦截器3周围对称部署协同拦截。当目标通过机动进行突防时,协同拦截结果如图12所示。

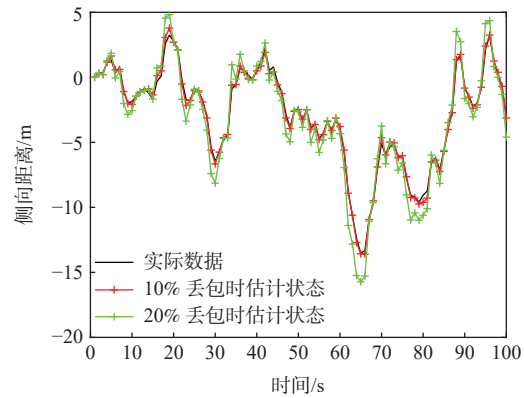


图12 全覆盖协同拦截

Fig. 12 Full coverage cooperative interception

从仿真结果可知,当目标的机动能力超过中心拦截器3的捕获能力时,相邻的拦截器2通过有效的覆盖捕捉区域成功捕获目标,并进行有效拦截。仿真结果显示,全覆盖系统策略是有效的。

5 结 论

1)本文分析了在拦截器饱和过载限制下,RTPN对任意机动目标能够有效进行拦截的捕获区域。有效地解决了RTPN制导方法的饱和过载限制,有效捕获问题。

2)针对实际系统具有的测量误差及拦截器本身动力学响应延迟情况,提出协同探测的目标轨迹预测算法,本文预测算法有效的解决了系统数传输延迟和动力学延迟问题,以及协同探测数据融合中数据丢包所引起的鲁棒性和精度问题。同时,得到比例系数范围及有效捕获区域,拦截器可以有效的拦截机动目标。

3)针对协同探测的数据丢包问题,提出一种协同滤波抗数据丢包的数据融合算法。

参考文献(References)

[1] 程涛, 周浩, 董晓飞, 等. 多飞行器突防打击一体化微分对策制导律设计[J]. 北京航空航天大学学报, 2022, 48(5): 898-909.
CHENG T, ZHOU H, DONG X F, et al. Differential game guidance law design for integration of penetration and strike of multiple flight vehicles[J]. Journal of Beijing University of Aeronautics and

Astronautics, 2022, 48(5): 898-909 (in Chinese).

[2] 张帅, 宋天莉, 焦巍, 等. 带有拦截时间约束的协同制导方法[J]. 北京航空航天大学学报, 2023, 49(8): 1956-1963.
ZHANG S, SONG T L, JIAO W, et al. Cooperative guidance method with interception time constraint[J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49(8): 1956-1963 (in Chinese).

[3] 杨彪, 荆武兴, 李涧青, 等. 动能拦截器运动伪装末制导律设计[J]. 宇航学报, 2020, 41(1): 91-100.
YANG B, JING W X, LI J Q, et al. Motion camouflage terminal guidance law of kinetic energy interceptor[J]. Journal of Astronautics, 2020, 41(1): 91-100 (in Chinese).

[4] 王少博, 郭杨, 王仕成, 等. 带有引诱角色的多飞行器协同最优制导方法[J]. 航空学报, 2020, 41(2): 323402.
WANG S B, GUO Y, WANG S C, et al. Cooperative optimal guidance method for multi-aircraft with luring role[J]. Acta Aeronautica et Astronautica Sinica, 2020, 41(2): 323402 (in Chinese).

[5] 王亚宁, 王辉, 林德福, 等. 基于虚拟视角约束的机动目标拦截制导方法[J]. 航空学报, 2022, 43(1): 324799.
WANG Y N, WANG H, LIN D F, et al. Guidance method for maneuvering target interception based on virtual look angle constraint[J]. Acta Aeronautica et Astronautica Sinica, 2022, 43(1): 324799 (in Chinese).

[6] 景亮, 张忠阳, 崔乃刚, 等. 固定时间收敛扰动观测终端滑模制导律设计[J]. 系统工程与电子技术, 2019, 41(8): 1820-1826.
JING L, ZHANG Z Y, CUI N G, et al. Fixed-time disturbance observer based terminal sliding mode guidance law[J]. Systems Engineering and Electronics, 2019, 41(8): 1820-1826 (in Chinese).

[7] 肖惟, 于江龙, 董希旺, 等. 过载约束下的大机动目标协同拦截[J]. 航空学报, 2020, 41(S1): 723777.
XIAO W, YU J L, DONG X W, et al. Cooperative interception against highly maneuvering target with acceleration constraints[J]. Acta Aeronautica et Astronautica Sinica, 2020, 41(S1): 723777 (in Chinese).

[8] 李万礼, 李炯, 雷虎民, 等. 基于滑模变结构制导律的捕获区分析[J]. 系统工程与电子技术, 2021, 43(11): 3321-3329.
LI W L, LI J, LEI H M, et al. Analysis of capture region based on sliding mode variable structure guidance law[J]. Systems Engineering and Electronics, 2021, 43(11): 3321-3329 (in Chinese).

[9] 于江龙, 董希旺, 李清东, 等. 拦截机动目标的分布式协同围捕制导方法[J]. 航空学报, 2022, 43(9): 513-533.
YU J L, DONG X W, LI Q D, et al. Distributed cooperative encirclement hunting guidance method for intercepting the maneuvering target[J]. Acta Aeronautica et Astronautica Sinica, 2022, 43(9): 513-533 (in Chinese).

[10] 黄景帅, 张洪波, 汤国建, 等. 拦截大气层内机动目标的自适应积分滑模制导律[J]. 宇航学报, 2019, 40(1): 51-60.
HUANG J S, ZHANG H B, TANG G J, et al. Adaptive integral sliding-mode guidance law for intercepting endo-atmospheric maneuvering targets[J]. Journal of Astronautics, 2019, 40(1): 51-60 (in Chinese).

[11] 王荣刚, 唐硕. 拦截高速运动目标广义相对偏置比例制导律[J]. 西北工业大学学报, 2019, 37(4): 682-690.
WANG R G, TANG S. Intercepting higher-speed targets using generalized relative biased proportional navigation[J]. Journal of Northwestern Polytechnical University, 2019, 37(4): 682-690(in Chinese).

[12] 熊少锋, 魏明英, 赵明元, 等. 逆轨拦截机动目标的三维最优制导律[J]. 宇航学报, 2020, 41(1): 80-90.
XIONG S F, WEI M Y, ZHAO M Y, et al. Three dimensional optimal guidance law against maneuvering targets for head-on engagement[J]. Journal of Astronautics, 2020, 41(1): 80-90 (in Chinese).

[13] 白志会, 黎克波, 苏文山, 等. 现实真比例导引拦截任意机动目标捕获区域[J]. 航空学报, 2020, 41(8): 323947.
BAI Z H, LI K B, SU W S, et al. Capture region of RTPN guidance law against arbitrarily maneuvering targets[J]. Acta Aeronautica et Astronautica Sinica, 2020, 41(8): 323947 (in Chinese).

[14] 马萌晨, 宋申民. 拦截机动目标的多导弹协同制导律[J]. 航空兵器, 2021, 28(6): 19-27.
MA M C, SONG S M. Multi missile cooperative guidance law for intercepting maneuvering targets[J]. Aviation Weapons, 2021, 28(6): 19-27(in Chinese).

[15] 周觐, 雷虎民, 侯峰, 等. 拦截高速目标的比例与反比例导引捕获区分析[J]. 宇航学报, 2018, 39(9): 1003-1012.
ZHOU J, LEI H M, HOU F, et al. Capture region analysis of proportional navigation and retro-proportional navigation guidance for hypersonic target interception[J]. Journal of Astronautics, 2018, 39(9): 1003-1012 (in Chinese).

[16] LI K B, SU W S, CHEN L. Performance analysis of realistic true proportional navigation against maneuvering targets using Lyapunov-like approach[J]. Aerospace Science and Technology, 2017, 69: 333-341.

[17] LI K B, ZHANG T T, CHEN L. Ideal proportional navigation for exoatmospheric interception[J]. Chinese Journal of Aeronautics, 2013, 26(4): 976-985.

[18] 王婷, 周军. 三维理想比例导引律的捕获区域分析[J]. 西北工业大学学报, 2007, 25(1): 83-86.
WANG T, ZHOU J. Is capture region of 3D ideal proportional navigation (IPN) also the biggest?[J]. Journal of Northwestern Polytechnical University, 2007, 25(1): 83-86 (in Chinese).

[19] TYAN F. Capture region of a GIPN guidance law for missile and target with bounded maneuverability[J]. IEEE Transactions on Aerospace and Electronic Systems, 2011, 47(1): 201-213.

[20] 黎克波, 陈磊, 张翼. 真比例导引律的降维分析方法[J]. 国防科技大学学报, 2012, 34(3): 1-5.
LI K B, CHEN L, ZHANG Y. Dimension-reduction method of true proportional navigation guidance law[J]. Journal of National University of Defense Technology, 2012, 34(3): 1-5 (in Chinese).

[21] LIU Y H, LI K B, CHEN L, et al. Novel augmented proportional navigation guidance law for mid-range autonomous rendezvous[J]. Acta Astronautica, 2019, 162: 526-535.

[22] LI K B, SHIN H S, TSOURDOS A. Capturability of a sliding-mode guidance law with finite-time convergence[J]. IEEE Transactions on Aerospace and Electronic Systems, 2019, 56(3): 2312-2325.

[23] GARAI T, MUKHOPADHYAY S, GHOSE D. Approximate closed-form solutions of realistic true proportional navigation guidance using the Adomian decomposition method[J]. Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering, 2009, 223(3): 189-199.

[24] LI K B, SHIN H S, TSOURDOS A, et al. Capturability of 3D PPN against lower-speed maneuvering target for homing phase[J]. IEEE

Transactions on Aerospace and Electronic Systems, 2020, 56(1): 711-722.

[25] LI K B, SHIN H S, TSOURDOS A, et al. Performance of 3-D PPN against arbitrarily maneuvering target for homing phase[J]. IEEE Transactions on Aerospace and Electronic Systems, 2020, 56(5): 3878-3891.

[26] LI K B, BAI Z H, SHIN H S, et al. Capturability of 3D RTPN guidance law against true-arbitrarily maneuvering target with maneuverability limitation[J]. Chinese Journal of Aeronautics, 2022, 35(7): 75-90.

[27] YANG C D, YANG C C. Analytical solution of three-dimensional realistic true proportional navigation[J]. Journal of Guidance, Control, and Dynamics, 1996, 19(3): 569-577.

[28] MOON J, KIM K, KIM Y. Design of missile guidance law via variable structure control[J]. Journal of Guidance, Control, and Dynamics, 2001, 24(4): 659-664.

[29] 李桂英, 于志刚, 张扬. 带有角度约束的机动目标拦截协同制导律[J]. 系统工程与电子技术, 2019, 41(3): 626-635.
LI G Y, YU Z G, ZHANG Y. Cooperative guidance law with angle constraint to intercept maneuvering target[J]. Systems Engineering and Electronics, 2019, 41(3): 626-635 (in Chinese).

[30] 张帅, 郭杨, 王仕成, 等. 考虑探测效能的有限时间协同制导方法[J]. 兵工学报, 2019, 40(9): 1849-1859.
ZHANG S, GUO Y, WANG S C, et al. A finite-time cooperative guidance method considering cooperative detection efficiency[J]. Acta Armamentarii, 2019, 40(9): 1849-1859 (in Chinese).

[31] LIU Y. Study and simulation to terminal guidance of kinetic energy interceptor on space[J]. Journal of Projectiles, Rockets, Missiles and Guidance, 2006, 8(4): 25-34.

[32] MOON J, KIM Y. Design of missile guidance law via variable structure control[C]// Proceedings of the AIAA Guidance, Navigation, and Control Conference and Exhibit. Reston: AIAA, 2000: 4068.

RTPN interception guidance law for maneuvering target based on collaborative filtering trajectory prediction

LI Jiguang[1, 2, 3, *], CHEN Xin[2], DONG Yanfei[1], QU Gaomin[1], ZHAO Chenggong[1], ZHANG Along[1]

(1. College of Aircraft, Xi'an Aeronautical University, Xi'an 710077, China;
2. College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China;
3. Northwest Institute of Mechanical and Electrical Engineering, Xianyang 712000, China)

Abstract: According to the actual requirements of the current air threat target interception, combined with the interceptor's own maneuverability and based on the full coverage cooperative strategy, a realistic true proportional navigation law (RTPN) guided interception method for collaborative detection is designed. It solves the problem that the traditional RTPN method does not consider the interceptor saturation overload limit and the determination of the acquisition area of any maneuvering target. A distributed cooperative filtering technique is suggested to address the accuracy and robustness of data fusion caused by measurement inaccuracy of target trajectory and packet loss of cooperative detection data during interception. Aiming at the problems of data transmission and the delay of interceptor dynamic response, a track prediction algorithm is proposed. The results of the simulations demonstrate that the suggested approach may successfully address issues with acquisition region identification, dynamic latency under saturation overload, and resilience and accuracy issues brought on by data packet loss in cooperative detection data fusion.

Keywords: collaborative interception; maneuvering target; realistic true proportional navigation law; overload limit; filtering algorithm; track prediction; data fusion

Received: 2022-04-02; **Accepted**: 2022-09-09; **Published Online**: 2022-10-17 18:10

URL: link.cnki.net/urlid/11.2625.V.20221017.1350.002

Foundation items: Shaanxi Natural Science Youth Fund Project (2021JQ-858); Shaanxi Natural Science Fund Key Project (2022JZ-37); Shaanxi Higher Education Teaching Reform Research Project (21BY163); Youth Science Fund (62103441); University-level Research Project (2021KY0210)

* **Corresponding author**. E-mail: 912646963@qq.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0264

无人机集群分组编队自适应跟踪-包含控制

张毅，于浩，杨秀霞*，姜子劼，王晨蕾
(海军航空大学，烟台 264001)

摘　　要：针对无人机(UVA)集群分组编队跟踪-包含控制问题，建立包括分组协同、跟踪-包含的分层协同控制架构。为突出集群系统的顶层规划能力，提出基于队形参数组的分组协同控制策略，克服了现有多编队控制方法难以实现队形变换的缺陷；在考虑组间通信耦合的基础上，设计自适应分组编队跟踪-包含控制律，克服了控制输入对Laplacian矩阵特征值的依赖，并给出集群系统实现期望队形的可行性约束；对控制输入引入饱和约束，并进一步推导得到输入饱和约束下系统稳定的充要条件，避免了集群的大幅机动行为。研究结果表明：所提策略能够实现无人机集群分组编队的跟踪-包含控制。

关　键　词：分组编队；跟踪-包含控制；协同控制；输入饱和；队形变换
中图分类号：V249
文献标志码：A　　**文章编号**：1001-5965(2024)01-0097-13

随着无人系统相关技术的发展，无人机(unmanned aerial vehicles, UAV)集群协同控制问题受到国内外学者的广泛关注[1-3]，并产生众多密切相关的研究分支，如编队控制[4-6]、编队包含控制[7]等。

编队控制是指通过设计控制律，使多UAV系统在有限时间内形成并保持期望的编队队形。而在实际任务环境中，UAV集群通常存在“功能异构”的情况，如部分UAV配备探测能力强的机载传感器且具有高性能的计算系统，能够根据战场环境和任务的变化，规划出一条安全高效的飞行航迹，另一部分UAV携带攻击设备，由于机载能力的限制难以配备其他传感设备，对外部环境的感知相对较弱。在此场景下出现了编队包含控制问题：通过设计集群控制协议，使编队领导者(探测UAV)形成期望编队队形的同时，跟随者(攻击UAV)进入由编队领导者建立的飞行凸包内，后者可为前者提供安全稳定的飞行环境，使跟随者专注于执行打击任务。

研究编队包含控制问题的多数文献中，将编队中心视为静止态，赵学远等[8-9]假设多UAV系统存在多个领导者，分别在切换拓扑和联合通信拓扑条件下设计了分布式控制器，并采用矩阵分解法将包含问题转化为低阶系统的渐近稳定问题，实现了对静态中心的包含控制；Han等[10]利用智能体的邻接位置和速度信息，构造了具有时变时延的编队包含控制协议，并给出系统稳定的充分条件。然而，对于UAV这类任务性较强的集群系统，在运动过程中往往需要跟踪参考轨迹或目标航迹，由此便催生了较为复杂的编队跟踪-包含控制问题，此时传统的静态包含控制算法并不满足实际的应用需求。对此，魏志强等[11]针对异构集群的跟踪-合围问题，基于鲁棒自适应、分布式观测器和一致性控制理论，设计切换拓扑下的集群分层控制策略，实现了对无人机-无人车(unmanned aerial vehicles-unmanned ground vehicles, UAV-UGV)空地异构系统的分布式控制；Hua等[12]研究了未知跟踪-领导者输入情况下的编队跟踪-包含控制问题，通过设计控制协议

收稿日期：2022-04-20；录用日期：2022-06-20；网络出版时间：2022-07-06 08:42
网络出版地址：link.cnki.net/urlid/11.2625.V.20220705.0908.002
基金项目：山东省自然科学基金(ZR2020MF090)
*通信作者. E-mail：yangxiuxia@126.com

引用格式：张毅，于浩，杨秀霞，等. 无人机集群分组编队自适应跟踪-包含控制[J]. 北京航空航天大学学报，2024，50(1)：97-109.
ZHANG Y，YU H，YANG X X，et al. Adaptive group formation tracking-containment control for UAV swarm[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(1)：97-109 (in Chinese).

避免了控制输入的大幅抖动；李蒙[13]基于滑模控制方法设计编队跟踪-包含控制协议，给出智能体系统在固定时间内收敛的充分条件。

针对单一编队的跟踪-包含控制问题，上述文献给出了可行方案，然而大规模集群系统在执行任务时，可能会面临多目标的情况，仅靠单一编队难以实现对分散目标的合围跟踪，此时需要对集群进行分组，通过组间的协同完成围捕任务[14]，可显著提高作战效率。目前对于分组编队包含控制的研究相对较少，Lu等[15]在文献[12]的基础上研究了分组编队的跟踪-包含控制问题，并给出了确定控制参数的具体算法；Hu等[16]以多目标监控为任务背景，研究了多智能体系统的包含控制问题，并将其应用到多机器人系统；Han等[17]研究了通信时延条件下的分组编队包含控制问题，提出基于变量代换的方法获取反馈矩阵，并给出设计控制协议的算法流程。

现有的研究成果并未考虑集群系统子编队间的队形变换问题，使得上述方案在具体任务背景下应用较为受限。针对上述问题，本文研究了UAV集群分组编队的跟踪-包含控制问题，提出基于队形参数组的分组协同控制策略，使集群在完成分组编队包含控制的同时能够实现编队间的队形变换，并有效避免集群的大幅机动行为，更加符合工程实际。

1 预备知识与问题描述

1.1 图论基础

UAV之间的通信关系可由有向图$G=(V,E,\boldsymbol{A})$来表示，其中$V=\{v_1,v_2,\cdots,v_M\}$为UAV的集合，$E\subseteq\{(v_i,v_j);\ v_i,v_j\in V;i\neq j\}$为通信边的集合，$\boldsymbol{A}=[a_{ij}]\in\mathbf{R}^{M\times M}$为有向图对应的邻接矩阵。令$e_{ij}=(v_i,v_j)$表示从$v_i$到$v_j$的通信边，若$e_{ij}\in E$，则$a_{ij}\neq 0$；反之，$a_{ij}=0$。$v_i$邻接UAV的集合表示为$N_i=\{v_j\in V;\ e_{ji}\in E\}$。定义$v_i$的入度为$d_{\text{in}}(v_i)=\sum_{i=1}^{M}a_{ij}(i=1,2,\cdots,M)$，则图$G$的入度矩阵为$\boldsymbol{D}=\text{diag}\{d_{\text{in}}(v_1),d_{\text{in}}(v_2),\cdots,d_{\text{in}}(v_M)\}$。图$G$的Laplacian矩阵$\boldsymbol{L}\in\mathbf{R}^{M\times M}$定义为$\boldsymbol{L}=\boldsymbol{D}-\boldsymbol{A}$。

1.2 问题描述

图1为UAV集群分组编队示意图。UAV集群系统由出发地到目的地对敌方多处目标执行侦察任务，为此将系统分为多个子编队，包括虚拟领导者、跟踪-领导者、编队领导者和跟随者。图中虚拟领导者预先规划整体的运动轨迹并为分组编队间的协同提供相位基准；跟踪-领导者在协同控制律的作用下实现相位协同，即实现子编队间的协同，从而为所属子编队规划飞行轨迹；编队领导者装载传感设备执行目标探测任务，跟随者配备攻击装置执行打击任务，通过控制算法的设计使编队领导者形成多个期望的飞行凸包，并使跟随者的运动状态收敛到所属编队的凸包内，据此可以保证跟随者时刻处于编队领导者形成的安全凸包内，继而实现集群系统穿越未知区域，确保了UAV集群飞行的安全。

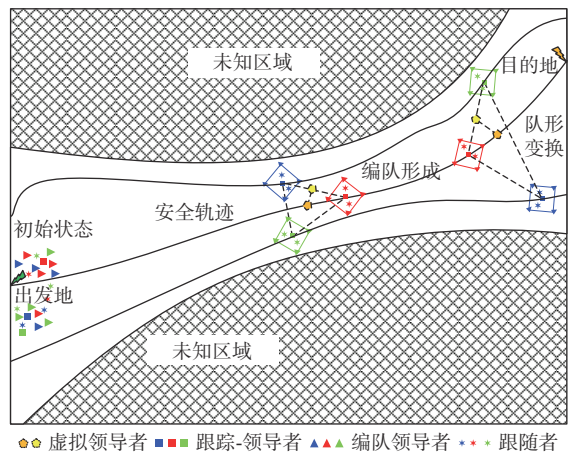


图1 UAV集群分组编队示意图

Fig. 1 Schematic diagram of UAV swarm group formation

系统中UAV的运动模型表示为

$$\dot{\boldsymbol{x}}_k(t)=\boldsymbol{A}\boldsymbol{x}_k(t)+\boldsymbol{B}\boldsymbol{u}_k(t) \tag{1}$$

式中：$\boldsymbol{x}_k(t)=\begin{bmatrix}\boldsymbol{p}_k^{\mathrm{T}}(t)\\ \boldsymbol{v}_k^{\mathrm{T}}(t)\end{bmatrix}$；$\boldsymbol{A}=\begin{bmatrix}0&1\\0&0\end{bmatrix}\otimes\boldsymbol{I}_n$；$\boldsymbol{B}=\begin{bmatrix}0\\1\end{bmatrix}\otimes\boldsymbol{I}_n$，$\otimes$表示克罗内克积；$\boldsymbol{p}_k(t)\in\mathbf{R}^n$、$\boldsymbol{v}_k(t)\in\mathbf{R}^n$和$\boldsymbol{u}_k(t)\in\mathbf{R}^m$分别为UAV的位置、速度和控制输入。

通过下标k区分不同类型的UAV：①$k=i$。$i=1,2,\cdots,N$表示编队领导者；$i=N+1,N+2,\cdots,N+M$表示跟随者；②$k=\mathrm{T}\bar{i}$，$\bar{i}=1,2,\cdots,p$表示跟踪-领导者$\text{UAV}_{\mathrm{T}\bar{i}}$，$p$为编队个数；③$k=a,b$表示虚拟领导者。

假设1 集群系统的通信拓扑满足：①编队领导者之间可进行信息交互（同组或异组），且各组领导者中均存在有向生成树；②跟随者之间可进行信息交互（同组或异组）且对于每个跟随者，在其所在分组内存在以编队领导者为根节点的有向生成树；③跟踪-领导者之间存在以虚拟领导者为根节点的有向生成树。

据此，集群系统通信拓扑对应的Laplacian矩阵可表述为

$$\boldsymbol{L}=\begin{bmatrix}\boldsymbol{L}_{\mathrm{T}}&\boldsymbol{0}_{p\times N}&\boldsymbol{0}_{p\times M}\\ \boldsymbol{L}_{\mathrm{EL}}&\boldsymbol{L}_{\mathrm{E}}&\boldsymbol{0}_{N\times M}\\ \boldsymbol{0}_{M\times p}&\boldsymbol{L}_{\mathrm{FE}}&\boldsymbol{L}_{\mathrm{F}}\end{bmatrix}$$

式中：$\boldsymbol{L}_{\mathrm{T}}\in\mathbf{R}^{p\times p}$为跟踪-领导者之间的信息交互；$\boldsymbol{L}_{\mathrm{EL}}\in\mathbf{R}^{N\times p}$为跟踪-领导者与编队领导者之间的信息交互；$\boldsymbol{L}_{\mathrm{E}}\in\mathbf{R}^{N\times N}$为编队领导者之间的信息交互；

$\boldsymbol{L}_{\mathrm{F}} \in \mathbf{R}^{M\times M}$为跟随者之间的信息交互, $\boldsymbol{L}_{\mathrm{FE}} \in \mathbf{R}^{M\times N}$为编队领导者与跟随者之间的信息交互。

进一步可将分块矩阵$\boldsymbol{L}_{\mathrm{E}}$、$\boldsymbol{L}_{\mathrm{F}}$分解为

$$\boldsymbol{L}_{\mathrm{E}}=\begin{bmatrix} \boldsymbol{L}_{\mathrm{E}1} & \mathbf{0} & \cdots & \mathbf{0} \\ \boldsymbol{L}_{\mathrm{E}21} & \boldsymbol{L}_{\mathrm{E}2} & \cdots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{L}_{\mathrm{E}p1} & \boldsymbol{L}_{\mathrm{E}p2} & \cdots & \boldsymbol{L}_{\mathrm{E}p} \end{bmatrix}$$

$$\boldsymbol{L}_{\mathrm{F}}=\begin{bmatrix} \boldsymbol{L}_{\mathrm{F}1} & \boldsymbol{L}_{\mathrm{F}12} & \cdots & \boldsymbol{L}_{\mathrm{F}1p} \\ \boldsymbol{L}_{\mathrm{F}21} & \boldsymbol{L}_{\mathrm{F}2} & \cdots & \boldsymbol{L}_{\mathrm{F}2p} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{L}_{\mathrm{F}p1} & \boldsymbol{L}_{\mathrm{F}p2} & \cdots & \boldsymbol{L}_{\mathrm{F}p} \end{bmatrix}$$

式中: 对角元素(矩阵)$\boldsymbol{L}_{\mathrm{E}\bar{i}}$、$\boldsymbol{L}_{\mathrm{F}\bar{i}}$表示组内的信息交互; 非对角元素(矩阵)$\boldsymbol{L}_{\mathrm{E}\bar{i}\bar{j}}$、$\boldsymbol{L}_{\mathrm{F}\bar{i}\bar{j}}$表示组间的信息交互。

假设 2　$\boldsymbol{L}_{\mathrm{E}\bar{i}\bar{j}}$、$\boldsymbol{L}_{\mathrm{F}\bar{i}\bar{j}}$的行和为 0, 即组间的信息交互满足入度平衡原则。

为方便描述, 分别令$\tilde{n}_{\bar{i}}$和$\bar{n}_{\bar{i}}$表示第$\bar{i}$分组中的编队领导者和跟随者个数, 标号的和分别为$\tilde{\upsilon}_{\bar{i}}=\sum_{j=0}^{\bar{i}-1}\tilde{n}_j$和$\bar{\upsilon}_{\bar{i}}=N+\sum_{j=0}^{\bar{i}-1}\bar{n}_j$, 且$\tilde{n}_0=\bar{n}_0=0$; 跟踪-领导者、编队领导者、跟随者的集合分别表示为V_{T}、V_{E}和V_{F}, 对于分组$\bar{i}$三者可表示为$V_{\cdot\bar{i}}$, 且有$\bigcup_{\bar{i}=1}^{p}V_{\cdot\bar{i}}=V_{\cdot}$。

2　集群控制策略设计及稳定性分析

结合如图 1 所示的任务背景, 本文给出如图 2 所示的 UAV 集群分组编队控制架构。

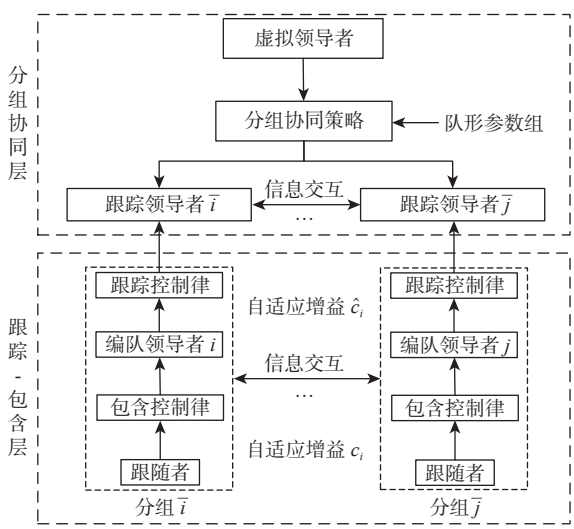


图 2　UAV 集群分组编队控制架构

Fig. 2　Architecture of UAV swarm group formation control

分组协同层通过设计基于队形参数组的控制协议, 使各跟踪-领导者 (子编队中心) 以虚拟领导者UAV_a为中心、UAV_b为基准实现相位协同; 跟踪-包含层利用自适应一致性理论并基于相对状态信息设计跟踪-包含控制协议, 使编队领导者形成期望子编队队形的同时, 跟随者收敛到其形成的凸包内, 避免了对大规模集群系统 Laplacian 矩阵特征值的计算, 降低了算法难度。

2.1　分组协同层控制律设计

分组协同策略如同编队间的纽带, 促使各编队相互协作完成任务, 相较于传统分组编队控制方法而言, 本节设计基于队形参数组的分组协同策略, 突出集群系统的顶层规划能力, 在确保跟踪-领导者满足相位协同的基础上实现分组编队间的队形变换。

为满足上述控制需求, 设计包含平移、旋转和缩放 3 种基本编队运动的编队队形参数组:

$$\boldsymbol{Z}(N)=\left\{\boldsymbol{p}_a(t),\boldsymbol{p}_b(t),\boldsymbol{C}_{\mathrm{T}1}(t),\cdots,\boldsymbol{C}_{\mathrm{T}p}(t)\right\} \tag{2}$$

$$\boldsymbol{C}_{\mathrm{T}\bar{i}}(t)=r_{\mathrm{T}\bar{i}}(t)\boldsymbol{R}_{\mathrm{T}\bar{i}}(t) \qquad \bar{i}\in\{1,2,\cdots,p\} \tag{3}$$

式中: $\boldsymbol{p}_a(t)$为 UAV 集群编队的中心, 其数值的改变对应了编队的平移运动; $r_{\mathrm{T}\bar{i}}(t)\in\mathbf{R}$为 UAV 集群编队队形的缩放参数, 表示跟踪-领导者到 UAV 集群编队中心$\boldsymbol{p}_a(t)$距离的大小, 对应编队的缩放运动; $\boldsymbol{R}_{\mathrm{T}\bar{i}}(t)=\begin{bmatrix}\cos\theta_{\mathrm{T}\bar{i}} & -\sin\theta_{\mathrm{T}\bar{i}}\\ \sin\theta_{\mathrm{T}\bar{i}} & \cos\theta_{\mathrm{T}\bar{i}}\end{bmatrix}$表示以$\boldsymbol{p}_b(t)$为基准的队形旋转参数, 决定了 UAV 集群编队的旋转运动。

如图 3 所示, 为确保 UAV 集群编队间的相位协同, 定义$\boldsymbol{p}_b(t)=[\cos(\theta_b(t))\ \ \sin(\theta_b(t))]^{\mathrm{T}}$作为相位基准, 在进行旋转运动时, 跟踪领导者只需旋转$\theta_b(t)$即可保持与动态基准的夹角保持不变 (即$\theta_i'=\theta_i$), 相较于传统定轴静基准x而言, 该基准的设立可确保 UAV 集群编队进行旋转运动前后, UAV 集群编队队形描述与初始的队形描述一致, 便于后续运动状态的解算。

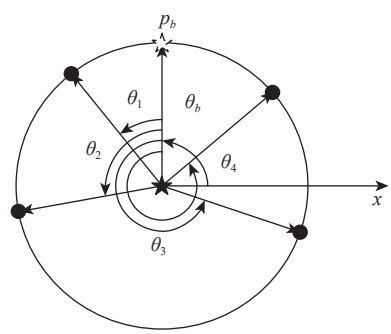


图 3　编队队形描述示意图

Fig. 3　Schematic diagram of formation description

结合参数组的定义, 各跟踪-领导者$\mathrm{UAV}_{\mathrm{T}\bar{i}}$形成的期望队形状态可描述为

$$\lim_{t\to\infty}\boldsymbol{p}_{\mathrm{T}\bar{i}}(t)=\boldsymbol{p}_a(t)+\boldsymbol{C}_{\mathrm{T}\bar{i}}(t)\boldsymbol{p}_b(t) \tag{4}$$

即跟踪-领导者$\mathrm{UAV}_{\mathrm{T}\bar{i}}$到达以$\boldsymbol{p}_a(t)$为中心, 以$\boldsymbol{p}_b(t)$为基准的期望位置。

根据式 (4), 定义位置和速度误差:

$$\tilde{\boldsymbol{p}}_{\mathrm{T}\bar{i}}(t)=\boldsymbol{p}_{\mathrm{T}\bar{i}}(t)-\boldsymbol{p}_a(t)-\boldsymbol{C}_{\mathrm{T}\bar{i}}(t)\boldsymbol{p}_b(t) \tag{5}$$

$$\bar{\boldsymbol{v}}_{\mathrm{T}\bar{i}}(t)=\boldsymbol{v}_{\mathrm{T}\bar{i}}(t)-\boldsymbol{v}_a(t)-\boldsymbol{C}_{\mathrm{T}\bar{i}}(t)\boldsymbol{v}_b(t)-\dot{\boldsymbol{C}}_{\mathrm{T}\bar{i}}(t)\boldsymbol{p}_b(t) \tag{6}$$

定义 1 给定跟踪-领导者$\mathrm{UAV}_{\mathrm{T}\bar{i}}$任意初始状态，若满足：

$$\begin{cases}\lim\limits_{t\to\infty}\left\|\bar{\boldsymbol{p}}_{\mathrm{T}\bar{i}}(t)-\bar{\boldsymbol{p}}_{\mathrm{T}\bar{j}}(t)\right\|\to 0\\ \lim\limits_{t\to\infty}\left\|\bar{\boldsymbol{v}}_{\mathrm{T}\bar{i}}(t)-\bar{\boldsymbol{v}}_{\mathrm{T}\bar{j}}(t)\right\|\to 0\end{cases}$$

则称多跟踪-领导者形成期望队形并保持。式中，$\bar{i},\bar{j}\in\{1,2,\cdots,p\}$且$\bar{i}\neq\bar{j}$。

基于邻接误差设计编队控制律：

$$\begin{cases}\boldsymbol{u}_{\mathrm{T}\bar{i}}=\boldsymbol{\kappa}_{\mathrm{T}\bar{i}}-\sum\limits_{\bar{j}=1}^{p}a_{\bar{i}\bar{j}}\left[k_1\left(\bar{\boldsymbol{v}}_{\mathrm{T}\bar{i}}-\bar{\boldsymbol{v}}_{\mathrm{T}\bar{j}}\right)+k_2\left(\bar{\boldsymbol{p}}_{\mathrm{T}\bar{i}}-\bar{\boldsymbol{p}}_{\mathrm{T}\bar{j}}\right)\right]-\\ \qquad k_1a_{\bar{i}b}\bar{\boldsymbol{v}}_{\mathrm{T}\bar{i}}-k_2a_{\bar{i}b}\bar{\boldsymbol{p}}_{\mathrm{T}\bar{i}}\\ \boldsymbol{\kappa}_{\mathrm{T}\bar{i}}=\ddot{\boldsymbol{C}}_{\mathrm{T}\bar{i}}\boldsymbol{p}_b+2\dot{\boldsymbol{C}}_{\mathrm{T}\bar{i}}\boldsymbol{v}_b+\boldsymbol{C}_{\mathrm{T}\bar{i}}\boldsymbol{u}_b+\boldsymbol{u}_a\end{cases} \tag{7}$$

式中：k_1、k_2为控制增益；$a_{\bar{i}b}$为跟踪-领导者$\mathrm{UAV}_{\mathrm{T}\bar{i}}$与虚拟领导者之间的通信权重。

令 $\bar{\boldsymbol{p}}_{\mathrm{T}}(t)=\left[\bar{\boldsymbol{p}}_{\mathrm{T}1}^{\mathrm{T}}(t),\bar{\boldsymbol{p}}_{\mathrm{T}2}^{\mathrm{T}}(t),\cdots,\bar{\boldsymbol{p}}_{\mathrm{T}p}^{\mathrm{T}}(t)\right]^{\mathrm{T}}$，$\bar{\boldsymbol{v}}_{\mathrm{T}}(t)=\left[\bar{\boldsymbol{v}}_{\mathrm{T}1}^{\mathrm{T}}(t),\bar{\boldsymbol{v}}_{\mathrm{T}2}^{\mathrm{T}}(t),\cdots,\bar{\boldsymbol{v}}_{\mathrm{T}p}^{\mathrm{T}}(t)\right]^{\mathrm{T}}$，可得闭环误差方程为

$$\dot{\boldsymbol{\phi}}=\boldsymbol{M}\boldsymbol{\phi} \tag{8}$$

式中：

$$\boldsymbol{M}=\begin{bmatrix}\boldsymbol{0} & \boldsymbol{I}_{nN}\\ -k_2(\boldsymbol{H}_b\otimes\boldsymbol{I}_n) & -k_1(\boldsymbol{H}_b\otimes\boldsymbol{I}_n)\end{bmatrix}$$

$$\boldsymbol{\phi}=\left(\bar{\boldsymbol{p}}^{\mathrm{T}},\bar{\boldsymbol{v}}^{\mathrm{T}}\right)^{\mathrm{T}},\ \boldsymbol{H}_b=\boldsymbol{L}_L+\boldsymbol{\Upsilon}^b$$

$$\boldsymbol{\Upsilon}^b=\mathrm{diag}\left(a_{1b},a_{2b},\cdots,a_{pb}\right)$$

引理 1[18-19] 在假设 1 成立的前提下，具有如下结论：①$\boldsymbol{H}_b$是非奇异的，且其所有特征根λ_b均有正实部；②$\boldsymbol{H}_b^{-1}$为非负的；③存在正定对角矩阵$\boldsymbol{S}=\mathrm{diag}\left\{s_1,s_2,\cdots,s_p\right\}$，使得$\boldsymbol{S}\boldsymbol{H}_b+(\boldsymbol{H}_b\boldsymbol{S})^{\mathrm{T}}$为正定对角矩阵，其中$\boldsymbol{S}=\left(\mathrm{diag}(\boldsymbol{d})\right)^{-1}$，$\boldsymbol{d}=\left[d_1,d_2,\cdots,d_p\right]^{\mathrm{T}}=\boldsymbol{H}_b^{-1}\boldsymbol{1}_p$（$\boldsymbol{1}_p$为全为 1 的$p$维列向量）。

引理 2[20] 考虑二次多项式：

$$f(\lambda)=\lambda^2+(a_1+ib_1)\lambda+a_2+ib_2 \tag{9}$$

当且仅当满足：

$$\begin{cases}a_1>0\\ a_1^2a_2+a_1b_1b_2-b_2^2>0\end{cases} \tag{10}$$

$f(\lambda)=0$的特征根均具有负实部。

定理 1 在假设 1 成立的基础上，若满足：

$$\frac{k_1^2}{k_2}>\max_{\bar{i}\in\{1,2,\cdots,p\}}\left\{\frac{\left(\mathrm{Im}\left(\lambda_{\bar{i}}^b\right)\right)^2}{\left|\lambda_{\bar{i}}^b\right|^2\mathrm{Re}\left(\lambda_{\bar{i}}^b\right)}\right\} \tag{11}$$

则跟踪-领导者$\mathrm{UAV}_{\mathrm{T}\bar{i}}$在控制律作用下能够形成期望编队队形。式中，$\lambda_{\bar{i}}^b$为$\boldsymbol{H}_b$的特征值。

证明：要使跟踪-领导者$\mathrm{UAV}_{\mathrm{T}\bar{i}}$形成期望队形，则只需证明式 (8) 稳定，式 (8) 的特征多项式可表示为

$$\left|\lambda\boldsymbol{I}_{2nN}-\boldsymbol{M}\right|=\left|\begin{bmatrix}\lambda\boldsymbol{I}_{nN} & -\boldsymbol{I}_{nN}\\ k_2(\boldsymbol{H}_b\otimes\boldsymbol{I}_n) & \lambda\boldsymbol{I}_{nN}+k_1(\boldsymbol{H}_b\otimes\boldsymbol{I}_n)\end{bmatrix}\right|=\lambda^2+k_1\lambda_{\bar{i}}^b\lambda+k_2\lambda_{\bar{i}}^b \tag{12}$$

要证$|\boldsymbol{\phi}(t)|\to 0$，仅需证多项式 (12) 的特征值均具有负实部即可，则需满足：

$$\begin{cases}\mathrm{Re}\left(\lambda_{\bar{i}}^b\right)>0\\ k_1^2k_2\left(\mathrm{Re}\left(\lambda_{\bar{i}}^b\right)\right)^3+k_1^2k_2\mathrm{Re}\left(\lambda_{\bar{i}}^b\right)\left(\mathrm{Im}\left(\lambda_{\bar{i}}^b\right)\right)^2>\\ \qquad k_2^2\left(\mathrm{Im}\left(\lambda_{\bar{i}}^b\right)\right)^2\end{cases} \tag{13}$$

由引理 2 可知，$\mathrm{Re}\left(\lambda_{\bar{i}}^b\right)>0$，若满足式 (11)，则矩阵$\boldsymbol{M}$的特征值均具有负实部，即$|\boldsymbol{\phi}(t)|\to 0$，系统可形成期望队形。 **证毕**

需要说明的是，控制增益k_1、k_2的选取不仅与系统稳定性有关，还与系统收敛速度息息相关，在满足式 (11) 的基础上，调整两者取值并确保k_1^2/k_2的值不变，会提高编队生成的收敛速度。

考虑$\mathrm{UAV}_{\mathrm{T}\bar{i}}$的实际运动约束，对上述控制输入做限幅处理，引入非线性饱和函数：

$$\mathrm{sat}(y)=\begin{cases}u & y>u\\ y & -u\leqslant y\leqslant u\\ -u & y<-u\end{cases} \tag{14}$$

设$\boldsymbol{u}_{\mathrm{i_s}}(t)=\left[\boldsymbol{u}_{1\mathrm{s}}(t),\boldsymbol{u}_{2\mathrm{s}}(t),\cdots,\boldsymbol{u}_{p\mathrm{s}}(t)\right]^{\mathrm{T}}$表示输入饱和约束下的控制量。

引入满足如下定义的时变参数$0<\delta_i(t)\leqslant 1$：

$$\delta_i(t)=\begin{cases}1 & \|\boldsymbol{u}_{\mathrm{T}\bar{i}}(t)\|_\infty\leqslant u_{\max}\\ \dfrac{u_{\max}}{\|\boldsymbol{u}_{\mathrm{T}\bar{i}}(t)\|_\infty} & \|\boldsymbol{u}_{\mathrm{T}\bar{i}}(t)\|_\infty>u_{\max}\end{cases} \tag{15}$$

经变换可得到：

$$\delta_i(t)\|\boldsymbol{u}_{\mathrm{T}\bar{i}}(t)\|_\infty=\begin{cases}\|\boldsymbol{u}_{\mathrm{T}\bar{i}}(t)\|_\infty & \|\boldsymbol{u}_{\mathrm{T}\bar{i}}(t)\|_\infty\leqslant u_{\max}\\ u_{\max} & \|\boldsymbol{u}_{\mathrm{T}\bar{i}}(t)\|_\infty>u_{\max}\end{cases} \tag{16}$$

进而推知：

$$\frac{\delta_i(t)\|\boldsymbol{u}_{\mathrm{T}\bar{i}}(t)\|_\infty}{u_{\max}}\leqslant 1 \tag{17}$$

结合式 (17) 和sat(·)定义可得：$\boldsymbol{u}_{i_s}(t)=\delta_i(t)\boldsymbol{u}_{\mathrm{T}\bar{i}}(t)$。

对于跟踪-领导者$U_{\mathrm{T}\bar{i}}$的控制输入式 (7)，令$\boldsymbol{u}_{\mathrm{T}}(t)=\left[\boldsymbol{u}_{\mathrm{T}1}(t),\boldsymbol{u}_{\mathrm{T}2}(t),\cdots,\boldsymbol{u}_{\mathrm{T}p}(t)\right]^{\mathrm{T}}$，必然存在对角矩阵$\boldsymbol{Q}(t)$，使如下等式成立：

$$\boldsymbol{u}_s(t)=\mathrm{sat}(\boldsymbol{Q}(t)\boldsymbol{u}(t))=\boldsymbol{Q}(t)\boldsymbol{u}(t) \tag{18}$$

式中：$\boldsymbol{Q}(t)=\mathrm{diag}\left\{\delta_1\cdot I_n,\delta_2\cdot I_n,\cdots,\delta_p\cdot I_n\right\}$。由式 (18) 可以看出，引入饱和约束的控制量等价于原控制输入左乘对角矩阵。

引理 3[21] 若$\boldsymbol{A}$和$\boldsymbol{B}$为半正定的 Hermite 矩阵，$\lambda_{\max}(\boldsymbol{A})$和$\lambda_{\min}(\boldsymbol{A})$分别表示$\boldsymbol{A}$的最大和最小特征值且$\lambda_i(\boldsymbol{B})$表示$\boldsymbol{B}$的特征值。则有如下结论：

$$\lambda_{\min}(\boldsymbol{A})\lambda_i(\boldsymbol{B})\leqslant\lambda_i(\boldsymbol{A}\boldsymbol{B})\leqslant\lambda_{\max}(\boldsymbol{A})\lambda_i(\boldsymbol{B})$$

结合式 (18), 系统误差状态方程 (8) 可重新表示为

$$\dot{\boldsymbol{\phi}} = \boldsymbol{M}_{\mathrm{s}}\boldsymbol{\phi} + \boldsymbol{N}_{\mathrm{s}} \tag{19}$$

式中:

$$\boldsymbol{M}_{\mathrm{s}} = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{I}_{nN} \\ -k_2\boldsymbol{Q}(t)(\boldsymbol{H}_b\otimes\boldsymbol{I}_n) & -k_1\boldsymbol{Q}(t)(\boldsymbol{H}_b\otimes\boldsymbol{I}_n) \end{bmatrix}$$

$$\boldsymbol{N}_{\mathrm{s}} = \begin{bmatrix} \boldsymbol{0} \\ (\boldsymbol{\kappa}_{\bar{i}}(t)\otimes\boldsymbol{I}_n) - \boldsymbol{Q}(t)(\boldsymbol{\kappa}_{\bar{i}}(t)\otimes\boldsymbol{I}_n) \end{bmatrix}$$

令 $\boldsymbol{\chi}(t) = \begin{bmatrix} \boldsymbol{I}_{np} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Q}(t) \end{bmatrix}$, 则 $\boldsymbol{M}_{\mathrm{s}} = \boldsymbol{\chi}(t)\boldsymbol{M}$。根据 $\boldsymbol{\chi}(t)$ 定义可知, $\boldsymbol{\chi}(t)$ 为正定矩阵, 且分别有 $\lambda_{\max}(\boldsymbol{\chi}) = 1$ 和 $\lambda_{\min}(\boldsymbol{\chi}) = \delta, 0 < \delta \leqslant 1$。

根据引理 3 有:

$$\delta\lambda_i(-\boldsymbol{M}) \leqslant \lambda_i(-\boldsymbol{M}\boldsymbol{\chi}) \leqslant \lambda_i(-\boldsymbol{M}) \tag{20}$$

变换可得

$$\lambda_i(\boldsymbol{M}) \leqslant \lambda_i(\boldsymbol{M}_{\mathrm{s}}) \leqslant \delta\lambda_i(\boldsymbol{M}) \tag{21}$$

由定理 1 可知, 无约束系统在式 (11) 条件下特征值均具有负实部, 结合上述推导过程, 若式 (11) 成立, 则输入饱和约束下系统的特征值仍均具有负实部, 即系统是稳定的。然而引入饱和约束后, 系统矩阵特征值的幅值变小, 根据频域理论推知, 闭环极点离虚轴距离变近。系统的收敛速度减慢, 此时需要合理选取控制增益 k_1、k_2 调节系统的收敛速度。

2.2　跟踪-包含层控制律设计

本节的控制目标有 2 个: ①控制编队领导者以期望的编队队形实现对中心领导者的跟踪; ②控制跟随者进入所属分组中编队领导者形成的飞行凸包内。为此给出如下定义:

定义 1　给定有界集合 $\rho_1 \in \mathbf{R}^{2n}$, 对于任意初始状态 $\boldsymbol{x}_i(0) \in \rho_1$, $(i = 1,2,\cdots,N)$, 若分组 $\bar{i}$ 中编队领导者的状态满足:

$$\lim_{t\to\infty}\left(\boldsymbol{x}_i(t) - \boldsymbol{h}_i(t) - \left([1]_{1\times\tilde{n}_i}\otimes\boldsymbol{x}_{\mathrm{T}\bar{i}}(t)\right)\right) = \boldsymbol{0} \quad \bar{i} = 1,2,\cdots,p \tag{22}$$

则称编队领导者实现了以期望队形 $\boldsymbol{h}_i(t)$ 跟踪 $U_{\mathrm{T}\bar{i}}$。

定义 2　给定有界集合 $\rho_2 \in \mathbf{R}^{2n}$, 对于任意初始状态 $\boldsymbol{x}_i(0) \in \rho_2$, $(i = N+1, N+2,\cdots,N+M)$, 若跟随者的状态满足:

$$\lim_{t\to\infty}\left(\boldsymbol{x}_i(t) - \sum_{j=\tilde{v}_{\bar{i}}+1}^{\tilde{v}_{\bar{i}}+\tilde{n}_{\bar{i}}}\alpha_{ij}\boldsymbol{x}_j(t)\right) = \boldsymbol{0} \tag{23}$$

则称分组编队实现包含控制。式中, $\sum_{j=\tilde{v}_{\bar{i}}+1}^{\tilde{v}_{\bar{i}}+\tilde{n}_{\bar{i}}}\alpha_{ij} = 1$。

需要说明的是, 定义 2 保证了跟随者的运动状态(位置和速度)收敛至编队领导者形成的飞行凸包内, 不仅在空间位置上确保飞行环境的安全性, 还使得跟随者的瞬时速度限制在特定范围内, 有效降低了能量损耗和外界障碍对跟随者的威胁。

定义 3　若定义 1 和定义 2 同时成立, 则称 UAV集群系统实现分组编队跟踪-包含控制。

引理 4　根据假设 1 和 $\boldsymbol{L}_{\mathrm{E}}$、$\boldsymbol{L}_{\mathrm{F}}$ 的定义, $\boldsymbol{L}_{\mathrm{E}}$ 和 $\boldsymbol{L}_{\mathrm{F}}$ 中每个对角分块子矩阵的特征值均具有正实部。

引理 5[15]　存在实对角矩阵 $\boldsymbol{\Delta}_1$、$\boldsymbol{\Delta}_2$、$\boldsymbol{G}$ 和 $\boldsymbol{\Xi}$ 使如下不等式成立:

$$\boldsymbol{\Lambda}_1 = \boldsymbol{\Delta}_1\boldsymbol{G}\boldsymbol{L}_{\mathrm{E}} + \boldsymbol{L}_{\mathrm{E}}^{\mathrm{T}}\boldsymbol{\Delta}_1\boldsymbol{G} > \boldsymbol{0} \tag{24}$$

$$\boldsymbol{\Lambda}_2 = \boldsymbol{\Delta}_2\boldsymbol{\Xi}\boldsymbol{L}_{\mathrm{F}} + \boldsymbol{L}_{\mathrm{F}}^{\mathrm{T}}\boldsymbol{\Delta}_2\boldsymbol{\Xi} > \boldsymbol{0} \tag{25}$$

引理 6[15] 若假设 2 成立, 则有:

1) 对于分组 $\bar{i}$, 有:

$$-(\boldsymbol{L}_{\mathrm{E}\bar{i}})^{-1}\boldsymbol{L}_{\mathrm{EL}\bar{i}} = [1]_{1\times\tilde{n}_{\bar{i}}} \tag{26}$$

2) 由 $\boldsymbol{L}_{\mathrm{FE}\bar{i}}[1]_{1\times\tilde{n}_{\bar{i}}} + \boldsymbol{L}_{\mathrm{F}\bar{i}}\boldsymbol{1}_{\tilde{n}_{\bar{i}}} = \boldsymbol{0}$, 可得

$$-\boldsymbol{L}_{\mathrm{F}}^{-1}\boldsymbol{L}_{\mathrm{FE}} = \mathrm{diag}\left\{-\boldsymbol{L}_{\mathrm{F}1}^{-1}\boldsymbol{L}_{\mathrm{FE}1}, -\boldsymbol{L}_{\mathrm{F}2}^{-1}\boldsymbol{L}_{\mathrm{FE}2}, \cdots, -\boldsymbol{L}_{\mathrm{F}p}^{-1}\boldsymbol{L}_{\mathrm{FE}p}\right\} = \mathrm{diag}\left\{\boldsymbol{\tau}_1, \boldsymbol{\tau}_2, \cdots, \boldsymbol{\tau}_p\right\} \tag{27}$$

为方便表述, 后文公式中省略变量 t。由定义 1, 对于第 i 个编队领导者定义编队跟踪误差:

$$\boldsymbol{\xi}_i = \sum_{j=p+1}^{p+N} a_{(p+i)j}\left((\boldsymbol{x}_i - \boldsymbol{h}_i) - (\boldsymbol{x}_j - \boldsymbol{h}_j)\right) + a_{(p+i)\bar{i}}(\boldsymbol{x}_i - \boldsymbol{h}_i - \boldsymbol{x}_{\mathrm{T}\bar{i}}) \tag{28}$$

式中: $i \in \{1,2,\cdots,N\}$, $\bar{i} \in \{1,2,\cdots,p\}$, $a_{(p+i)j}$ 为编队领导者之间的通信权重, $a_{(p+i)\bar{i}}$ 为跟踪-领导者 $\mathrm{UAV}_{\mathrm{T}\bar{i}}$ 与编队领导者之间的通信权重。故式 (28) 中, 第 1 项为期望编队生成误差, 第 2 项为轨迹跟踪误差。

由定义 2, 对于第 i 个跟随者定义包含误差:

$$\boldsymbol{\zeta}_i = \sum_{j=1}^{N+M} a_{(p+i)(p+j)}(\boldsymbol{x}_i - \boldsymbol{x}_j) \tag{29}$$

式中: $i \in \{N+1, N+2,\cdots,N+M\}$, $a_{(p+i)(p+j)}$ 既包括编队领导者与跟随者之间的通信权重, 也包括跟随者之间的通信权重。

受文献 [22] 启发, 设计分组编队跟踪-包含控制律如下:

$$\begin{cases} \boldsymbol{u}_i = (c_i + \varepsilon_i)\boldsymbol{K}_1\boldsymbol{\xi}_i + \boldsymbol{w}_i - \sigma_1\boldsymbol{f}(\boldsymbol{\xi}_i) \\ \dot{c}_i = \boldsymbol{\xi}_i^{\mathrm{T}}\boldsymbol{\Psi}\boldsymbol{\xi}_i \qquad i \in \{1,2,\cdots,N\} \end{cases} \tag{30}$$

$$\begin{cases} \boldsymbol{u}_i = (\hat{c}_i + \hat{\varepsilon}_i)\boldsymbol{K}_2\boldsymbol{\zeta}_i - \sigma_2\boldsymbol{f}(\boldsymbol{\zeta}_i) \\ \dot{\hat{c}}_i = \boldsymbol{\zeta}_i^{\mathrm{T}}\boldsymbol{\Psi}\boldsymbol{\zeta}_i \qquad i \in \{N+1, N+2,\cdots,N+M\} \end{cases} \tag{31}$$

式中: $\boldsymbol{w}_i$ 为待确定的编队队形辅助函数; $\boldsymbol{K}_1$、$\boldsymbol{K}_2$ 为待求的反馈增益矩阵; $\sigma_1 > 0$、$\sigma_2 > 0$ 且为常数;

$\hat{c}_i(t)$和$c_i(t)$分别为编队领导者和跟随者的自适应增益；ε_i、$\hat{\varepsilon}_i$分别为关于$\boldsymbol{\xi}_i$、$\boldsymbol{\zeta}_i$的光滑函数；$\boldsymbol{f}(\cdot)$为待确定的非线性函数，$\boldsymbol{\Psi}$为待确定参数矩阵。

为将系统化为紧凑形式，令

$$\boldsymbol{x}_{\mathrm{T}}=\left[\boldsymbol{x}_{\mathrm{T}1}^{\mathrm{T}},\boldsymbol{x}_{\mathrm{T}2}^{\mathrm{T}},\cdots,\boldsymbol{x}_{\mathrm{T}p}^{\mathrm{T}}\right]^{\mathrm{T}},\ \boldsymbol{x}_{\mathrm{E}}=\left[\boldsymbol{x}_1^{\mathrm{T}},\boldsymbol{x}_2^{\mathrm{T}},\cdots,\boldsymbol{x}_N^{\mathrm{T}}\right]^{\mathrm{T}},$$
$$\boldsymbol{x}_{\mathrm{F}}=\left[\boldsymbol{x}_{N+1}^{\mathrm{T}},\boldsymbol{x}_{N+2}^{\mathrm{T}},\cdots,\boldsymbol{x}_{N+M}^{\mathrm{T}}\right]^{\mathrm{T}},\ \boldsymbol{h}=\left[\boldsymbol{h}_1^{\mathrm{T}},\boldsymbol{h}_2^{\mathrm{T}},\cdots,\boldsymbol{h}_N^{\mathrm{T}}\right]^{\mathrm{T}},$$
$$\boldsymbol{\varpi}=\left[\boldsymbol{w}_1^{\mathrm{T}},\boldsymbol{w}_2^{\mathrm{T}},\cdots,\boldsymbol{w}_N^{\mathrm{T}}\right]^{\mathrm{T}},$$
$$\boldsymbol{F}(\boldsymbol{\xi})=\left[\boldsymbol{f}^{\mathrm{T}}(\boldsymbol{\xi}_1),\boldsymbol{f}^{\mathrm{T}}(\boldsymbol{\xi}_2),\cdots,\boldsymbol{f}^{\mathrm{T}}(\boldsymbol{\xi}_N)\right]^{\mathrm{T}},$$
$$\boldsymbol{F}(\boldsymbol{\zeta})=\left[\boldsymbol{f}^{\mathrm{T}}(\boldsymbol{\zeta}_1),\boldsymbol{f}^{\mathrm{T}}(\boldsymbol{\zeta}_2),\cdots,\boldsymbol{f}^{\mathrm{T}}(\boldsymbol{\zeta}_M)\right]^{\mathrm{T}}。$$

将式(30)和式(31)代入系统方程并引入输入饱和约束，整理可得

$$\dot{\boldsymbol{x}}_{\mathrm{E}}=(\boldsymbol{I}_N\otimes\boldsymbol{A})\boldsymbol{x}_{\mathrm{E}}+(\boldsymbol{I}_N\otimes\boldsymbol{B})\,\mathrm{sat}(\boldsymbol{u}_{\mathrm{E}}) \tag{32}$$

$$\dot{\boldsymbol{x}}_{\mathrm{F}}=(\boldsymbol{I}_M\otimes\boldsymbol{A})\boldsymbol{x}_{\mathrm{F}}+(\boldsymbol{I}_M\otimes\boldsymbol{B})\,\mathrm{sat}(\boldsymbol{u}_{\mathrm{F}}) \tag{33}$$

令$\boldsymbol{\xi}=\left[\boldsymbol{\xi}_1^{\mathrm{T}},\boldsymbol{\xi}_2^{\mathrm{T}},\cdots,\boldsymbol{\xi}_N^{\mathrm{T}}\right]^{\mathrm{T}}$，$\boldsymbol{\zeta}=\left[\boldsymbol{\zeta}_1^{\mathrm{T}},\boldsymbol{\zeta}_2^{\mathrm{T}},\cdots,\boldsymbol{\zeta}_M^{\mathrm{T}}\right]^{\mathrm{T}}$得到分组编队闭环误差方程：

$$\boldsymbol{\xi}=(\boldsymbol{L}_{\mathrm{E}}\otimes\boldsymbol{I}_n)(\boldsymbol{x}_{\mathrm{E}}-\boldsymbol{h})+(\boldsymbol{L}_{\mathrm{EL}}\otimes\boldsymbol{I}_n)\boldsymbol{x}_{\mathrm{T}} \tag{34}$$

$$\boldsymbol{\zeta}=(\boldsymbol{L}_{\mathrm{F}}\otimes\boldsymbol{I}_n)\boldsymbol{x}_{\mathrm{F}}+(\boldsymbol{L}_{\mathrm{FE}}\otimes\boldsymbol{I}_n)\boldsymbol{x}_{\mathrm{E}} \tag{35}$$

求导可得

$$\begin{aligned}\dot{\boldsymbol{\xi}}=&[(\boldsymbol{I}_N\otimes\boldsymbol{A})+(\boldsymbol{c}+\boldsymbol{\varepsilon})\boldsymbol{L}_{\mathrm{E}}\otimes\boldsymbol{B}\boldsymbol{K}_1]\boldsymbol{\xi}+\\&(\boldsymbol{L}_{\mathrm{E}}\otimes\boldsymbol{A})\boldsymbol{h}+(\boldsymbol{L}_{\mathrm{EL}}\otimes\boldsymbol{B})\boldsymbol{u}_{\mathrm{T}}+(\boldsymbol{L}_{\mathrm{E}}\otimes\boldsymbol{B})\boldsymbol{\varpi}-\\&(\boldsymbol{L}_{\mathrm{E}}\otimes\boldsymbol{I}_n)\dot{\boldsymbol{h}}-\sigma_1(\boldsymbol{L}_{\mathrm{E}}\otimes\boldsymbol{B})\boldsymbol{F}(\boldsymbol{\xi})\end{aligned} \tag{36}$$

$$\begin{aligned}\dot{\boldsymbol{\zeta}}=&[(\boldsymbol{I}_M\otimes\boldsymbol{A})+(\hat{\boldsymbol{c}}+\hat{\boldsymbol{\varepsilon}})\boldsymbol{L}_{\mathrm{F}}\otimes\boldsymbol{B}\boldsymbol{K}_2]\boldsymbol{\zeta}+\\&[(\boldsymbol{c}+\boldsymbol{\varepsilon})\boldsymbol{L}_{\mathrm{FE}}\otimes\boldsymbol{B}\boldsymbol{K}_1]\boldsymbol{\xi}-\sigma_2(\boldsymbol{L}_{\mathrm{F}}\otimes\boldsymbol{B})\boldsymbol{F}(\boldsymbol{\zeta})+\\&(\boldsymbol{L}_{\mathrm{FE}}\otimes\boldsymbol{B})\boldsymbol{\varpi}-\sigma_1(\boldsymbol{L}_{\mathrm{FE}}\otimes\boldsymbol{B})\boldsymbol{F}(\boldsymbol{\xi})\end{aligned} \tag{37}$$

结合上述分析过程，下面给出集群系统在所提出的控制协议作用下，实现分组编队跟踪-包含控制所需满足的条件。

引理7[23] 若$(\boldsymbol{A},\boldsymbol{B})$是可控的，且$\boldsymbol{A}$的特征值均在左半平面，则对于每个$\Theta\in(0,1]$，存在唯一的正定矩阵$\boldsymbol{P}(\Theta)>\boldsymbol{0}$满足Riccati方程：

$$\boldsymbol{P}(\Theta)\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}(\Theta)-\boldsymbol{P}(\Theta)\boldsymbol{B}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}(\Theta)+\Theta\boldsymbol{I}=\boldsymbol{0}$$

进一步可得$\lim\limits_{\Theta\to0}\boldsymbol{P}(\Theta)=\boldsymbol{0}$。

定理2 在假设1和假设2成立的基础上，若如下条件成立：

1）期望队形函数$\boldsymbol{h}$满足可行性条件：

$$\boldsymbol{A}\boldsymbol{h}+\boldsymbol{B}\boldsymbol{\varpi}-\dot{\boldsymbol{h}}=\boldsymbol{0} \tag{38}$$

2）对于给定的控制参数σ_1、σ_2满足：

$$\begin{cases}\sigma_1\geqslant u_{\max}\\ \sigma_2\geqslant\max\{\omega_1,\omega_2,\cdots,\omega_N\}+\sigma_1\end{cases} \tag{39}$$

3）$\boldsymbol{K}_1=\boldsymbol{K}_2=-\boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}(\Theta)$，$\varepsilon_i=\boldsymbol{\xi}_i^{\mathrm{T}}\boldsymbol{P}(\Theta)\boldsymbol{\xi}_i$，$\hat{\varepsilon}_i=\boldsymbol{\zeta}_i^{\mathrm{T}}\cdot\boldsymbol{P}(\Theta)\boldsymbol{\zeta}_i$且$\boldsymbol{\Psi}=\boldsymbol{P}(\Theta)\boldsymbol{B}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}(\Theta)$。

4）非线性函数$\boldsymbol{f}(\cdot)$定义为

$$\boldsymbol{f}(\boldsymbol{\vartheta})=\begin{cases}\boldsymbol{0} & \|\boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}(\Theta)\boldsymbol{\vartheta}\|=0\\ \dfrac{\boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}(\Theta)\boldsymbol{\vartheta}}{\|\boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}(\Theta)\boldsymbol{\vartheta}\|} & \|\boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}(\Theta)\boldsymbol{\vartheta}\|\neq0\end{cases} \tag{40}$$

则UAV集群系统在控制律的作用下能够实现分组编队跟踪-包含控制。

证明： 首先证明编队领导者在控制律作用下能够以期望队形实现对$U_{\mathrm{T}i}$的跟踪。考虑如下Lyapunov函数：

$$V_1=\frac{1}{2}\sum_{i=1}^{N}\bar{g}_i(2c_i+\varepsilon_i)\varepsilon_i+\frac{1}{2}\sum_{i=1}^{N}\bar{g}_i(c_i-\beta_1)^2 \tag{41}$$

式中：$\boldsymbol{\Delta}_1\boldsymbol{G}'=\mathrm{diag}\{\bar{g}_1,\bar{g}_2,\cdots,\bar{g}_N\}$为正定对角矩阵，且$\bar{g}_i>0$；$\beta_1>0$为常数；对于$\forall i\in E$，$\dot{c}_i(t)\geqslant0$，且当$t\to\infty$时，有$\dot{c}_i(t)=0$。因此，$V_1$是正定的。

令常数$\partial_1>0$，满足如下不等式：

$$\partial_1\geqslant\sup_{\Theta\in(0,1],\|\boldsymbol{h}_i(t)\|\leqslant\omega,\boldsymbol{x}_i(0)\in\rho_1}V_1(0)$$

因为$\|\boldsymbol{h}_i(t)\|$，ρ_1是有界的且$\lim\limits_{\Theta\to0}\boldsymbol{P}(\Theta)=\boldsymbol{0}$并结合$\dot{c}_i$的定义可知$\partial_1>0$是存在的。

令$\boldsymbol{D}_v(\partial_1)=\left\{\boldsymbol{\xi}\in\mathbf{R}^{2Nn}:V_1\leqslant\partial_1\right\}$，$\Theta_*\in(0,1]$且对于$\Theta\in(0,\Theta_*]$，满足：$\|\boldsymbol{u}_i(t)\|_\infty\leqslant\Delta$。$\Theta_*$的存在性可由$\lim\limits_{\Theta\to0}\boldsymbol{P}(\Theta)=\boldsymbol{0}$证明。结合低增益反馈控制理论，式(32)可转化为

$$\begin{aligned}\dot{\boldsymbol{x}}_{\mathrm{E}}=&[\boldsymbol{I}_N\otimes\boldsymbol{A}+(\boldsymbol{c}+\boldsymbol{\varepsilon})\boldsymbol{L}_{\mathrm{E}}\otimes\boldsymbol{B}\boldsymbol{K}_1]\boldsymbol{x}_{\mathrm{E}}-\\&[(\boldsymbol{c}+\boldsymbol{\varepsilon})\boldsymbol{L}_{\mathrm{E}}\otimes\boldsymbol{B}\boldsymbol{K}_1]\boldsymbol{h}+[(\boldsymbol{c}+\boldsymbol{\varepsilon})\boldsymbol{L}_{\mathrm{EL}}\otimes\boldsymbol{B}\boldsymbol{K}_1]\boldsymbol{x}_{\mathrm{T}}+\\&(\boldsymbol{I}_N\otimes\boldsymbol{B})\boldsymbol{\varpi}-\sigma_1(\boldsymbol{I}_N\otimes\boldsymbol{B})\boldsymbol{F}(\boldsymbol{\xi})\end{aligned} \tag{42}$$

对V_1求导可得：

$$\dot{V}_1=\sum_{i=1}^{N}2\bar{g}_i(c_i+\varepsilon_i)\boldsymbol{\xi}_i^{\mathrm{T}}\boldsymbol{P}(\Theta)\dot{\boldsymbol{\xi}}_i+\sum_{i=1}^{N}\bar{g}_i(\varepsilon_i+c_i-\beta_1)\dot{c}_i \tag{43}$$

其中，

$$\sum_{i=1}^{N}\bar{g}_i(\varepsilon_i+c_i-\beta_1)\dot{c}_i=\boldsymbol{\xi}^{\mathrm{T}}\left[(\boldsymbol{c}+\boldsymbol{\varepsilon}-\beta_1\boldsymbol{I})\boldsymbol{\Delta}_1\boldsymbol{G}'\otimes\boldsymbol{\Psi}\right]\boldsymbol{\xi}$$

$$\begin{aligned}\sum_{i=1}^{N}2\bar{g}_i(c_i+\varepsilon_i)\boldsymbol{\xi}_i^{\mathrm{T}}\boldsymbol{P}(\Theta)\dot{\boldsymbol{\xi}}_i=&\\&\boldsymbol{\xi}^{\mathrm{T}}\left[(\boldsymbol{c}+\boldsymbol{\varepsilon})\boldsymbol{\Delta}_1\boldsymbol{G}'\otimes\left(\boldsymbol{P}(\Theta)\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}(\Theta)\right)-\right.\\&\left.\left(\boldsymbol{\Delta}_1\boldsymbol{G}\boldsymbol{L}_{\mathrm{E}}+\boldsymbol{L}_{\mathrm{E}}^{\mathrm{T}}\boldsymbol{\Delta}_1\boldsymbol{G}'\right)(\boldsymbol{c}+\boldsymbol{\varepsilon})^2\otimes\boldsymbol{\Psi}\right]\boldsymbol{\xi}+\\&2\boldsymbol{\xi}^{\mathrm{T}}[(\boldsymbol{c}+\boldsymbol{\varepsilon})\boldsymbol{\Delta}_1\boldsymbol{G}'\boldsymbol{L}_{\mathrm{E}}\otimes\boldsymbol{P}(\Theta)\boldsymbol{A}]\boldsymbol{h}-\\&2\boldsymbol{\xi}^{\mathrm{T}}[(\boldsymbol{c}+\boldsymbol{\varepsilon})\boldsymbol{\Delta}_1\boldsymbol{G}'\boldsymbol{L}_{\mathrm{E}}\otimes\boldsymbol{P}(\Theta)]\dot{\boldsymbol{h}}+\\&2\boldsymbol{\xi}^{\mathrm{T}}[(\boldsymbol{c}+\boldsymbol{\varepsilon})\boldsymbol{\Delta}_1\boldsymbol{G}'\boldsymbol{L}_{\mathrm{E}}\otimes\boldsymbol{P}(\Theta)\boldsymbol{B}]\boldsymbol{\varpi}+\\&2\boldsymbol{\xi}^{\mathrm{T}}[(\boldsymbol{c}+\boldsymbol{\varepsilon})\boldsymbol{\Delta}_1\boldsymbol{G}'\boldsymbol{L}_{\mathrm{EL}}\otimes\boldsymbol{P}(\Theta)\boldsymbol{B}]\boldsymbol{u}_{\mathrm{T}}-\\&2\sigma_1\boldsymbol{\xi}^{\mathrm{T}}[(\boldsymbol{c}+\boldsymbol{\varepsilon})\boldsymbol{\Delta}_1\boldsymbol{G}'\boldsymbol{L}_{\mathrm{E}}\otimes\boldsymbol{P}(\Theta)\boldsymbol{B}]\boldsymbol{F}(\boldsymbol{\xi})\end{aligned} \tag{44}$$

设$\boldsymbol{\Delta}_1\boldsymbol{G}'\boldsymbol{L}_{\mathrm{E}}+\boldsymbol{L}_{\mathrm{E}}^{\mathrm{T}}\boldsymbol{\Delta}_1\boldsymbol{G}'$的最小正特征值为$\lambda_1^{\min}$，则有

$$\sum_{i=1}^{N}2\bar{g}_i(c_i+\varepsilon_i)\boldsymbol{\xi}_i^{\mathrm{T}}\boldsymbol{P}(\Theta)\dot{\boldsymbol{\xi}}_i\leqslant$$
$$\boldsymbol{\xi}^{\mathrm{T}}\left[(\boldsymbol{c}+\boldsymbol{\varepsilon})\boldsymbol{\Lambda}_1\boldsymbol{G}'\otimes\left(\boldsymbol{P}(\Theta)\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}(\Theta)\right)-\lambda_1^{\min}(\boldsymbol{c}+\boldsymbol{\varepsilon})^2\otimes\boldsymbol{\Psi}\right]\boldsymbol{\xi}+$$
$$2\boldsymbol{\xi}^{\mathrm{T}}[(\boldsymbol{c}+\boldsymbol{\varepsilon})\boldsymbol{\Lambda}_1\boldsymbol{G}'\boldsymbol{L}_{\mathrm{E}}\otimes\boldsymbol{P}(\Theta)\boldsymbol{A}]\boldsymbol{h}-$$
$$2\boldsymbol{\xi}^{\mathrm{T}}[(\boldsymbol{c}+\boldsymbol{\varepsilon})\boldsymbol{\Lambda}_1\boldsymbol{G}'\boldsymbol{L}_{\mathrm{E}}\otimes\boldsymbol{P}(\Theta)]\dot{\boldsymbol{h}}+$$
$$2\boldsymbol{\xi}^{\mathrm{T}}[(\boldsymbol{c}+\boldsymbol{\varepsilon})\boldsymbol{\Lambda}_1\boldsymbol{G}'\boldsymbol{L}_{\mathrm{E}}\otimes\boldsymbol{P}(\Theta)\boldsymbol{B}]\boldsymbol{\varpi}+$$
$$2\boldsymbol{\xi}^{\mathrm{T}}[(\boldsymbol{c}+\boldsymbol{\varepsilon})\boldsymbol{\Lambda}_1\boldsymbol{G}'\boldsymbol{L}_{\mathrm{EL}}\otimes\boldsymbol{P}(\Theta)\boldsymbol{B}]\boldsymbol{u}_{\mathrm{T}}-$$
$$2\sigma_1\boldsymbol{\xi}^{\mathrm{T}}[(\boldsymbol{c}+\boldsymbol{\varepsilon})\boldsymbol{\Lambda}_1\boldsymbol{G}'\boldsymbol{L}_{\mathrm{E}}\otimes\boldsymbol{P}(\Theta)\boldsymbol{B}]\boldsymbol{F}(\boldsymbol{\xi})\tag{45}$$

若编队函数可行性条件式 (38) 成立, 则

$$(\boldsymbol{I}_N\otimes\boldsymbol{A})\boldsymbol{h}+(\boldsymbol{I}_N\otimes\boldsymbol{B})\boldsymbol{\varpi}-(\boldsymbol{I}_N\otimes\boldsymbol{I}_n)\dot{\boldsymbol{h}}=\boldsymbol{0}\tag{46}$$

两边同乘$(\boldsymbol{c}+\boldsymbol{\varepsilon})\boldsymbol{\Lambda}_1\boldsymbol{G}'\boldsymbol{L}_{\mathrm{E}}\otimes\boldsymbol{P}(\Theta)$可得

$$[(\boldsymbol{c}+\boldsymbol{\varepsilon})\boldsymbol{\Lambda}_1\boldsymbol{G}'\boldsymbol{L}_{\mathrm{E}}\otimes\boldsymbol{P}(\Theta)\boldsymbol{A}]\boldsymbol{h}-$$
$$[(\boldsymbol{c}+\boldsymbol{\varepsilon})\boldsymbol{\Lambda}_1\boldsymbol{G}'\boldsymbol{L}_{\mathrm{E}}\otimes\boldsymbol{P}(\Theta)]\dot{\boldsymbol{h}}+$$
$$[(\boldsymbol{c}+\boldsymbol{\varepsilon})\boldsymbol{\Lambda}_1\boldsymbol{G}'\boldsymbol{L}_{\mathrm{E}}\otimes\boldsymbol{P}(\Theta)\boldsymbol{B}]\boldsymbol{\varpi}=\boldsymbol{0}\tag{47}$$

由$\boldsymbol{f}(\cdot)$定义推知:

$$\boldsymbol{\xi}_i^{\mathrm{T}}\boldsymbol{P}(\Theta)\boldsymbol{B}\boldsymbol{f}(\boldsymbol{\xi}_i)=\left\|\boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}(\Theta)\boldsymbol{\xi}_i\right\|\tag{48}$$

根据 Cauchy-Schwarz 不等式有:

$$\boldsymbol{\xi}_i^{\mathrm{T}}\boldsymbol{P}(\Theta)\boldsymbol{B}\boldsymbol{f}\left(\boldsymbol{\xi}_j\right)\leqslant\left\|\boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}(\Theta)\boldsymbol{\xi}_i\right\|\left\|f\left(\boldsymbol{\xi}_j\right)\right\|\leqslant\left\|\boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}(\Theta)\boldsymbol{\xi}_i\right\|\qquad\forall i\neq j\tag{49}$$

则

$$-2\sigma_1\boldsymbol{\xi}^{\mathrm{T}}[(\boldsymbol{c}+\boldsymbol{\varepsilon})\boldsymbol{\Lambda}_1\boldsymbol{G}'\boldsymbol{L}_{\mathrm{E}}\otimes\boldsymbol{P}(\Theta)\boldsymbol{B}]\boldsymbol{F}(\boldsymbol{\xi})=$$
$$-2\sigma_1\sum_{i=1}^{N}(c_i+\varepsilon_i)\bar{g}_i\boldsymbol{\xi}_i^{\mathrm{T}}\boldsymbol{P}(\Theta)\boldsymbol{B}\sum_{k=1}^{p}a_{(p+i)k}\boldsymbol{f}(\boldsymbol{\xi}_i)-$$
$$2\sigma_1\sum_{i=1}^{N}(c_i+\varepsilon_i)\bar{g}_i\boldsymbol{\xi}_i^{\mathrm{T}}\boldsymbol{P}(\Theta)\boldsymbol{B}\sum_{j=1}^{N}a_{(p+i)(p+j)}\cdot$$
$$\left(\boldsymbol{f}(\boldsymbol{\xi}_i)-\boldsymbol{f}\left(\boldsymbol{\xi}_j\right)\right)\leqslant-2\sigma_1\sum_{i=1}^{N}(c_i+\varepsilon_i)\bar{g}_i\sum_{k=1}^{p}a_{(p+i)k}\left\|\boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}(\Theta)\boldsymbol{\xi}_i\right\|\tag{50}$$

$$2\boldsymbol{\xi}^{\mathrm{T}}[(\boldsymbol{c}+\boldsymbol{\varepsilon})\boldsymbol{\Lambda}_1\boldsymbol{G}'\boldsymbol{L}_{\mathrm{EL}}\otimes\boldsymbol{P}(\Theta)\boldsymbol{B}]\boldsymbol{u}_{\mathrm{T}}=$$
$$2\sum_{i=1}^{N}(c_i+\varepsilon_i)\bar{g}_i\sum_{k=1}^{p}a_{(p+i)k}\boldsymbol{\xi}_i^{\mathrm{T}}\boldsymbol{P}(\Theta)\boldsymbol{B}\boldsymbol{u}_{\mathrm{T}k}\leqslant$$
$$2\sum_{i=1}^{N}(c_i+\varepsilon_i)\bar{g}_i\sum_{k=1}^{p}a_{(p+i)k}\left\|\boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}(\Theta)\boldsymbol{\xi}_i\right\|\|\boldsymbol{u}_{\mathrm{T}k}\|\leqslant$$
$$2u_{\max}\sum_{i=1}^{N}(c_i+\varepsilon_i)\bar{g}_i\sum_{k=1}^{p}a_{(p+i)k}\left\|\boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}(\Theta)\boldsymbol{\xi}_i\right\|\tag{51}$$

令$\sigma_1\geqslant u_{\max}$, 综上可得

$$\dot{V}_1\leqslant\boldsymbol{\xi}^{\mathrm{T}}\left[(\boldsymbol{c}+\boldsymbol{\varepsilon})\boldsymbol{\Lambda}_1\boldsymbol{G}'\otimes\left(\boldsymbol{P}(\Theta)\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}(\Theta)+\boldsymbol{\Psi}\right)-\left(\lambda_1^{\min}(\boldsymbol{c}+\boldsymbol{\varepsilon})^2+\beta_1\boldsymbol{\Lambda}_1\boldsymbol{G}'\right)\otimes\boldsymbol{\Psi}\right]\boldsymbol{\xi}\tag{52}$$

其中, $\left(\lambda_1^{\min}(\boldsymbol{c}+\boldsymbol{\varepsilon})^2+\beta_1\boldsymbol{\Lambda}_1\boldsymbol{G}'\right)\otimes\boldsymbol{\Psi}$可化简为

$$-\boldsymbol{\xi}^{\mathrm{T}}\left[\left(\lambda_1^{\min}(\boldsymbol{c}+\boldsymbol{\varepsilon})^2+\beta_1\boldsymbol{\Lambda}_1\boldsymbol{G}'\right)\otimes\boldsymbol{\Psi}\right]\boldsymbol{\xi}\leqslant-2\boldsymbol{\xi}^{\mathrm{T}}\left[\sqrt{\lambda_1^{\min}\beta_1(\boldsymbol{c}+\boldsymbol{\varepsilon})^2\boldsymbol{\Lambda}_1\boldsymbol{G}'}\otimes\boldsymbol{\Psi}\right]\boldsymbol{\xi}\tag{53}$$

取$\beta_1\geqslant\dfrac{\max_{i\in E}\bar{g}_i}{\lambda_1^{\min}}$, 进而可得

$$\dot{V}_1\leqslant\boldsymbol{\xi}^{\mathrm{T}}\left[(\boldsymbol{c}+\boldsymbol{\varepsilon})\boldsymbol{\Lambda}_1\boldsymbol{G}'\otimes\left(\boldsymbol{P}(\Theta)\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}(\Theta)-\boldsymbol{\Psi}\right)\right]\boldsymbol{\xi}\tag{54}$$

利用变量代换$\boldsymbol{\Upsilon}_1=\left(\sqrt{(\boldsymbol{c}+\boldsymbol{\varepsilon})\boldsymbol{\Lambda}_1\boldsymbol{G}'}\otimes\boldsymbol{I}\right)\boldsymbol{\xi}$, 可将式 (54) 化简为

$$\dot{V}_1\leqslant\boldsymbol{\Upsilon}_1^{\mathrm{T}}[\boldsymbol{I}_N\otimes\left(\boldsymbol{P}(\Theta)\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}(\Theta)-\boldsymbol{P}(\Theta)\boldsymbol{B}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}(\Theta)\right)]\boldsymbol{\Upsilon}_1\tag{55}$$

显然, 对于给定的参数$\Theta>0$, 当$\boldsymbol{P}(\Theta)\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}(\Theta)-\boldsymbol{P}(\Theta)\boldsymbol{B}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}(\Theta)=-\Theta\boldsymbol{I}<\boldsymbol{0}$时, $\dot{V}_1\leqslant0$。当且仅当$\boldsymbol{\Upsilon}_1=\boldsymbol{0}$时, $\dot{V}_1=0$。根据 LaSalle 不变定理, 跟踪误差方程 (34) 是渐近稳定的, 则$\lim\limits_{t\to\infty}\boldsymbol{\xi}(t)=\boldsymbol{0}$, 即

$$\lim_{t\to\infty}[(\boldsymbol{L}_{\mathrm{E}}\otimes\boldsymbol{I}_n)(\boldsymbol{x}_{\mathrm{E}}-\boldsymbol{h})+(\boldsymbol{L}_{\mathrm{EL}}\otimes\boldsymbol{I}_n)\boldsymbol{x}_{\mathrm{T}}]=\boldsymbol{0}$$

并转换为

$$\lim_{t\to\infty}\left[\boldsymbol{x}_{\mathrm{E}}-\boldsymbol{h}+\left(\boldsymbol{L}_{\mathrm{E}}^{-1}\boldsymbol{L}_{\mathrm{EL}}\otimes\boldsymbol{I}_n\right)\boldsymbol{x}_{\mathrm{T}}\right]=\boldsymbol{0}\tag{56}$$

则对于分组$\bar{i}\in\{1,2,\cdots,p\}$, 由式 (26) 可得

$$\lim_{t\to\infty}\left(\boldsymbol{x}_i-\boldsymbol{h}_i-[1]_{1\times\bar{n}_i}\otimes\boldsymbol{x}_{\mathrm{T}\bar{i}}\right)=\boldsymbol{0}\tag{57}$$

由此可知, 编队领导者$i\in\{1,2,\cdots,N\},i\in\bar{i}$实现了以期望队形跟踪$U_{\mathrm{T}\bar{i}}$。

下面证明跟随者在控制律作用下收敛至领导者形成的凸包内。考虑如下 Lyapunov 函数:

$$V_2=\frac{1}{2}\sum_{i=1}^{M}\bar{\varphi}_i(2\hat{c}_i+\hat{\varepsilon}_i)\hat{\varepsilon}_i+\frac{1}{2}\sum_{i=1}^{M}\bar{\varphi}_i(\hat{c}_i-\beta_2)^2\tag{58}$$

式中: $\boldsymbol{\Lambda}_2\boldsymbol{\Xi}=\mathrm{diag}\{\bar{\varphi}_1,\bar{\varphi}_2,\cdots,\bar{\varphi}_M\}$为正定对角矩阵, 且$\bar{\varphi}_i>0$; $\beta_2>0$为常数; $\dot{\hat{c}}_i(t)\geqslant0$, 且当$t\to\infty$时, 有$\dot{\hat{c}}_i(t)=0$。因此, V_2是正定的。

类似的, 可将式 (33) 转化为

$$\dot{\boldsymbol{x}}_{\mathrm{F}}=[\boldsymbol{I}_M\otimes\boldsymbol{A}+(\hat{\boldsymbol{c}}+\hat{\boldsymbol{\varepsilon}})\boldsymbol{L}_{\mathrm{F}}\otimes\boldsymbol{B}\boldsymbol{K}_2]\boldsymbol{x}_{\mathrm{F}}+[(\hat{\boldsymbol{c}}+\hat{\boldsymbol{\varepsilon}})\boldsymbol{L}_{\mathrm{FE}}\otimes\boldsymbol{B}\boldsymbol{K}_2]\boldsymbol{x}_{\mathrm{E}}-\sigma_2(\boldsymbol{I}_M\otimes\boldsymbol{B})\boldsymbol{F}(\boldsymbol{\zeta})\tag{59}$$

对V_2求导可得

$$\dot{V}_2=\sum_{i=1}^{M}2\bar{\varphi}_i(\hat{c}_i+\hat{\varepsilon}_i)\boldsymbol{\zeta}_i^{\mathrm{T}}\boldsymbol{P}(\Theta)\dot{\boldsymbol{\zeta}}_i+\sum_{i=1}^{M}\bar{\varphi}_i(\hat{\varepsilon}_i+\hat{c}_i-\beta_2)\dot{\hat{c}}_i\tag{60}$$

其中:

$$\sum_{i=1}^{M}\bar{\varphi}_i(\hat{\varepsilon}_i+\hat{c}_i-\beta_2)\dot{\hat{c}}_i=\boldsymbol{\zeta}^{\mathrm{T}}\left[(\hat{\boldsymbol{c}}+\hat{\boldsymbol{\varepsilon}}-\beta_2\boldsymbol{I})\boldsymbol{\Lambda}_2\boldsymbol{\Xi}\otimes\boldsymbol{\Psi}\right]\boldsymbol{\zeta}$$

$$\sum_{i=1}^{M}2\bar{\varphi}_i(\hat{c}_i+\hat{\varepsilon}_i)\boldsymbol{\zeta}_i^{\mathrm{T}}\boldsymbol{P}(\Theta)\dot{\boldsymbol{\zeta}}_i=$$
$$\boldsymbol{\zeta}^{\mathrm{T}}\left[(\hat{\boldsymbol{c}}+\hat{\boldsymbol{\varepsilon}})\boldsymbol{\Lambda}_2\boldsymbol{\Xi}\otimes\left(\boldsymbol{P}(\Theta)\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}(\Theta)\right)-\right.$$
$$\left.(\hat{\boldsymbol{c}}+\hat{\boldsymbol{\varepsilon}})\left(\boldsymbol{\Lambda}_2\boldsymbol{\Xi}\boldsymbol{L}_{\mathrm{FE}}+\boldsymbol{L}_{\mathrm{FE}}^{\mathrm{T}}\boldsymbol{\Lambda}_2\boldsymbol{\Xi}\right)(\hat{\boldsymbol{c}}+\hat{\boldsymbol{\varepsilon}})\otimes\boldsymbol{\Psi}\right]\boldsymbol{\zeta}+$$
$$2\boldsymbol{\zeta}^{\mathrm{T}}[(\hat{\boldsymbol{c}}+\hat{\boldsymbol{\varepsilon}})\boldsymbol{\Lambda}_2\boldsymbol{\Xi}\boldsymbol{L}_{\mathrm{FE}}(\boldsymbol{c}+\boldsymbol{\varepsilon})\otimes\boldsymbol{P}(\Theta)\boldsymbol{B}\boldsymbol{K}_2]\boldsymbol{\xi}-$$
$$2\sigma_1\boldsymbol{\zeta}^{\mathrm{T}}[(\hat{\boldsymbol{c}}+\hat{\boldsymbol{\varepsilon}})\boldsymbol{\Lambda}_2\boldsymbol{\Xi}\boldsymbol{L}_{\mathrm{FE}}\otimes\boldsymbol{P}(\Theta)\boldsymbol{B}]\boldsymbol{F}(\boldsymbol{\xi})-$$
$$2\sigma_2\boldsymbol{\zeta}^{\mathrm{T}}[(\hat{\boldsymbol{c}}+\hat{\boldsymbol{\varepsilon}})\boldsymbol{\Lambda}_2\boldsymbol{\Xi}\boldsymbol{L}_{\mathrm{F}}\otimes\boldsymbol{P}(\Theta)\boldsymbol{B}]\boldsymbol{F}(\boldsymbol{\zeta})+$$
$$2\boldsymbol{\zeta}^{\mathrm{T}}[(\hat{\boldsymbol{c}}+\hat{\boldsymbol{\varepsilon}})\boldsymbol{\Lambda}_2\boldsymbol{\Xi}\boldsymbol{L}_{\mathrm{FE}}\otimes\boldsymbol{P}(\Theta)\boldsymbol{B}]\boldsymbol{\varpi}\tag{61}$$

令$\lambda_2^{\min}$为$\boldsymbol{\Lambda}_2\boldsymbol{\Xi}\boldsymbol{L}_{\mathrm{FE}}+\boldsymbol{L}_{\mathrm{FE}}^{\mathrm{T}}\boldsymbol{\Lambda}_2\boldsymbol{\Xi}$的最小正特征值，则有

$$\sum_{i=1}^{M}2\bar{\varphi}_i(\hat{c}_i+\hat{\varepsilon}_i)\boldsymbol{\zeta}_i^{\mathrm{T}}\boldsymbol{P}(\Theta)\dot{\boldsymbol{\zeta}}_i\leqslant$$
$$\boldsymbol{\zeta}^{\mathrm{T}}\left[(\hat{\boldsymbol{c}}+\hat{\boldsymbol{\varepsilon}})\boldsymbol{\Lambda}_2\boldsymbol{\Xi}\otimes\left(\boldsymbol{P}(\Theta)\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}(\Theta)\right)-\right.$$
$$\left.\lambda_2^{\min}(\hat{\boldsymbol{c}}+\hat{\boldsymbol{\varepsilon}})^2\otimes\boldsymbol{\Psi}\right]\boldsymbol{\zeta}+$$
$$2\boldsymbol{\zeta}^{\mathrm{T}}[(\hat{\boldsymbol{c}}+\hat{\boldsymbol{\varepsilon}})\boldsymbol{\Lambda}_2\boldsymbol{\Xi}\boldsymbol{L}_{\mathrm{FE}}(\boldsymbol{c}+\boldsymbol{\varepsilon})\otimes\boldsymbol{P}(\Theta)\boldsymbol{B}\boldsymbol{K}_2]\boldsymbol{\xi}-$$
$$2\sigma_2\boldsymbol{\zeta}^{\mathrm{T}}[(\hat{\boldsymbol{c}}+\hat{\boldsymbol{\varepsilon}})\boldsymbol{\Lambda}_2\boldsymbol{\Xi}\boldsymbol{L}_{\mathrm{F}}\otimes\boldsymbol{P}(\Theta)\boldsymbol{B}]\boldsymbol{F}(\boldsymbol{\zeta})+$$
$$2\boldsymbol{\zeta}^{\mathrm{T}}[(\hat{\boldsymbol{c}}+\hat{\boldsymbol{\varepsilon}})\boldsymbol{\Lambda}_2\boldsymbol{\Xi}\boldsymbol{L}_{\mathrm{FE}}\otimes\boldsymbol{P}(\Theta)\boldsymbol{B}](\boldsymbol{\varpi}-\sigma_1\boldsymbol{F}(\boldsymbol{\xi}))\tag{62}$$

下面分别对式(62)的各项进行简化：

$$-2\sigma_2\boldsymbol{\zeta}^{\mathrm{T}}[(\hat{\boldsymbol{c}}+\hat{\boldsymbol{\varepsilon}})\boldsymbol{\Lambda}_2\boldsymbol{\Xi}\boldsymbol{L}_{\mathrm{F}}\otimes\boldsymbol{P}(\Theta)\boldsymbol{B}]\boldsymbol{F}(\boldsymbol{\zeta})\leqslant$$
$$-2\sigma_2\sum_{i=1}^{M}(\hat{c}_i+\hat{\varepsilon}_i)\bar{\varphi}_i\left\|\boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}(\Theta)\boldsymbol{\zeta}_i\right\|\sum_{j=1}^{N}a_{(p+N+i)(p+j)}\tag{63}$$

令$\boldsymbol{\eta}_i=\boldsymbol{w}_i-\sigma_1\boldsymbol{f}(\boldsymbol{\xi}_i)$，$\boldsymbol{\eta}=\left[\eta_1,\eta_2,\cdots,\eta_N\right]^{\mathrm{T}}$，考虑到期望编队函数是有界的，设$\|\boldsymbol{w}_i\|\leqslant\omega_i$，$i=1,2,\cdots,N$，又因为$\|\boldsymbol{f}(\boldsymbol{\xi}_i)\|\leqslant1$，可得$\left\|\boldsymbol{\eta}_i\right\|\leqslant\omega_i+\sigma_1$，则

$$2\boldsymbol{\zeta}^{\mathrm{T}}[(\hat{\boldsymbol{c}}+\hat{\boldsymbol{\varepsilon}})\boldsymbol{\Lambda}_2\boldsymbol{\Xi}\boldsymbol{L}_{\mathrm{FE}}\otimes\boldsymbol{P}(\Theta)\boldsymbol{B}](\boldsymbol{\varpi}-\sigma_1\boldsymbol{F}(\boldsymbol{\xi}))=$$
$$2\sum_{i=1}^{M}(\hat{c}_i+\hat{\varepsilon}_i)\bar{\varphi}_i\boldsymbol{\zeta}_i^{\mathrm{T}}\boldsymbol{P}(\Theta)\boldsymbol{B}\sum_{j=1}^{N}a_{(p+N+i)(p+j)}\boldsymbol{\eta}_i\leqslant$$
$$2\sum_{i=1}^{M}(\hat{c}_i+\hat{\varepsilon}_i)\bar{\varphi}_i\left\|\boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}(\Theta)\boldsymbol{\zeta}_i\right\|\sum_{j=1}^{N}a_{(p+N+i)(p+j)}\left\|\boldsymbol{\eta}_i\right\|\leqslant$$
$$2\bar{\mu}\sum_{i=1}^{M}(\hat{c}_i+\hat{\varepsilon}_i)\bar{\varphi}_i\left\|\boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}(\Theta)\boldsymbol{\zeta}_i\right\|\sum_{j=1}^{N}a_{(p+N+i)(p+j)}\tag{64}$$

式中：$\bar{\mu}=\max\{\omega_1,\omega_2,\cdots,\omega_N\}+\sigma_1$。

由2.1节证明可知，编队领导者可形成期望编队并跟踪上UAV$_{0\bar{i}}$，即有$\lim\limits_{t\to\infty}\boldsymbol{\xi}(t)=0$，推导可得$V_2$的简化形式为

$$\dot{V}_2\leqslant\boldsymbol{\zeta}^{\mathrm{T}}\left[(\hat{\boldsymbol{c}}+\hat{\boldsymbol{\varepsilon}})\boldsymbol{\Lambda}_2\boldsymbol{\Xi}\otimes\left(\boldsymbol{P}(\Theta)\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}(\Theta)+\boldsymbol{\Psi}\right)-\right.$$
$$\left.\left(\lambda_2^{\min}(\hat{\boldsymbol{c}}+\hat{\boldsymbol{\varepsilon}})^2+\beta_2\boldsymbol{\Lambda}_2\boldsymbol{\Xi}\right)\otimes\boldsymbol{\Psi}\right]\boldsymbol{\zeta}\tag{65}$$

其中，$\left(\lambda_1^{\min}(\boldsymbol{c}+\boldsymbol{\varepsilon})^2+\beta_1\boldsymbol{\Lambda}_1\boldsymbol{G}\right)\otimes\boldsymbol{\Psi}$可化简为

$$-\boldsymbol{\zeta}^{\mathrm{T}}\left[\left(\lambda_2^{\min}(\hat{\boldsymbol{c}}+\hat{\boldsymbol{\varepsilon}})^2+\beta_2\boldsymbol{\Lambda}_2\boldsymbol{\Xi}\right)\otimes\boldsymbol{\Psi}\right]\boldsymbol{\zeta}\leqslant$$
$$-2\boldsymbol{\zeta}^{\mathrm{T}}\left[\sqrt{\lambda_2^{\min}\beta_2(\hat{\boldsymbol{c}}+\hat{\boldsymbol{\varepsilon}})^2\boldsymbol{\Lambda}_2\boldsymbol{\Xi}}\otimes\boldsymbol{\Psi}\right]\boldsymbol{\zeta}\tag{66}$$

取$\beta_2\geqslant\dfrac{\max\bar{\varphi}_i}{\lambda_2^{\min}}$，进而可得

$$\dot{V}_2\leqslant\boldsymbol{\zeta}^{\mathrm{T}}\left[(\hat{\boldsymbol{c}}+\hat{\boldsymbol{\varepsilon}})\boldsymbol{\Lambda}_2\boldsymbol{\Xi}\otimes\left(\boldsymbol{P}(\Theta)\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}(\Theta)-\boldsymbol{\Psi}\right)\right]\boldsymbol{\zeta}\tag{67}$$

利用变量代换$\boldsymbol{\Upsilon}_2=\left(\sqrt{(\hat{\boldsymbol{c}}+\hat{\boldsymbol{\varepsilon}})\boldsymbol{\Lambda}_2\boldsymbol{\Xi}}\otimes\boldsymbol{I}\right)\boldsymbol{\zeta}$，可将式(67)化简为

$$\dot{V}_2\leqslant\boldsymbol{\Upsilon}_2^{\mathrm{T}}\left[\boldsymbol{I}_M\otimes\left(\boldsymbol{P}(\Theta)\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}(\Theta)-\right.\right.$$
$$\left.\left.\boldsymbol{P}(\Theta)\boldsymbol{B}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}(\Theta)\right)\right]\boldsymbol{\Upsilon}_2\tag{68}$$

显然有$\dot{V}_2\leqslant0$。当且仅当$\boldsymbol{\Upsilon}_2=\boldsymbol{0}$时，$\dot{V}_2=0$。利用LaSalle不变定理，包含误差方程(35)是渐近稳定的，则$\lim\limits_{t\to\infty}\boldsymbol{\zeta}(t)=\boldsymbol{0}$，即

$$\lim_{t\to\infty}[(\boldsymbol{L}_{\mathrm{F}}\otimes\boldsymbol{I}_n)\boldsymbol{x}_{\mathrm{F}}+(\boldsymbol{L}_{\mathrm{FE}}\otimes\boldsymbol{I}_n)\boldsymbol{x}_{\mathrm{E}}]=\boldsymbol{0}$$

并转化为

$$\lim_{t\to\infty}\left[\boldsymbol{x}_{\mathrm{F}}+\left(\boldsymbol{L}_{\mathrm{F}}^{-1}\boldsymbol{L}_{\mathrm{FE}}\otimes\boldsymbol{I}_n\right)\boldsymbol{x}_{\mathrm{E}}\right]=\boldsymbol{0}\tag{69}$$

则对于分组$\bar{i}\in\{1,2,\cdots,p\}$，由式(27)可得

$$\lim_{t\to\infty}\left(\boldsymbol{x}_i-\boldsymbol{\tau}_i\boldsymbol{x}_j\right)=\boldsymbol{0}\tag{70}$$

由此可知，跟随者$i\in\{N+1,N+2,\cdots,N+M\}$，$i\in\bar{i}$收敛至编队领导者$j\in\{1,2,\cdots,N\}$，$j\in\bar{i}$形成的凸包内。**证毕**

结合本节分析过程，可知：①集群系统要想实现分组编队队形，期望编队函数需确保与UAV系统运动学模型是兼容的；②当$p=1$时，集群仅包括一个分组，无需考虑分组间的信息交互，此时研究的问题退化为单一编队的跟踪-包含控制问题，即本文是对传统编队包含控制问题的拓展。

3 仿真与分析

本节对所设计的分组编队跟踪-包含控制协议进行验证。考虑由12架编队领导者、6架跟随者和3架跟踪-领导者$U_{\mathrm{T}\bar{i}}$组成的集群系统，并将系统分为3个分组：$V_{\mathrm{E}1}=\{1,2,3,4\}$，$V_{\mathrm{F}1}=\{13,14\}$；$V_{\mathrm{E}2}=\{5,6,7,8\}$，$V_{\mathrm{F}2}=\{15,16\}$；$V_{\mathrm{E}3}=\{9,10,11,12\}$，$V_{\mathrm{F}3}=\{17,18\}$。子群之间的通信拓扑结构如图4所示。

其中虚线箭头表示异组编队领导者之间和跟随着之间的通信权重，其余UAV之间的通信权重均为1，不同子编队之间的通信权重满足假设2描述的入度平衡假设。

3.1 分组编队协同控制仿真

根据拟领导者和跟踪-领导者$U_{\mathrm{T}\bar{i}}$间的通信拓扑，求得H_b的特征值均大于0，由式(11)可知只需满足$k_1^2/k_2>0$即可。对$U_{0\bar{i}}$的初始运动状态做如下设定：$\boldsymbol{x}_{\mathrm{T}1}=[-25,2,-25,3,5,2]^{\mathrm{T}}$；$\boldsymbol{x}_{\mathrm{T}2}=[-30,-2,30,0,5,1]^{\mathrm{T}}$；

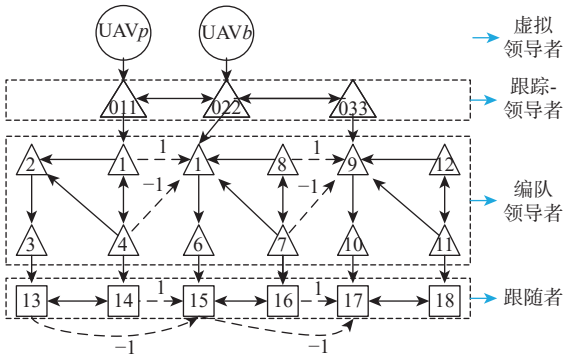


图 4　通信拓扑结构

Fig. 4　Communication topology diagram

$\boldsymbol{x}_{\mathrm{T}3}=[20,3,-20,2,5,2]^{\mathrm{T}}$, 设定控制增益 $k_1=2,k_2=1$, 输入阈值 $u_{\max}=5$。

为了验证分组协同控制算法的有效性, 设当 $t\in[0,20)$ s时, 队形缩放因子 $\boldsymbol{r}_{\mathrm{T}}=[40,30,30]$, 旋转因子 $\boldsymbol{R}_{\bar{i}}$ 中 $\boldsymbol{\theta}_{\mathrm{T}}=\left[\dfrac{\pi}{3},\dfrac{2\pi}{3},0\right]$; 当 $t\in[20,40)$ s时, 缩放因子 $\boldsymbol{r}_{\mathrm{T}}=[20,40,40]$, 旋转因子 $\boldsymbol{R}_{\bar{i}}$ 中 $\boldsymbol{\theta}_{\mathrm{T}}=\left[\dfrac{-2\pi}{3},0,\dfrac{\pi}{3}\right]$。

如图 5 所示, 跟踪-领导者在各时段内均能按照期望队形参数运动且飞行轨迹平滑, 实现由“三角编队”到“倒三角编队”的变换; 如图 6 和图 7 所示, 在 $t=10$ s左右, 误差收敛为 0, $t=20$ s时, 误差曲线出现波动, 原因在于此时编队期望队形发生瞬时变化, 误差曲线陡增, 随后在控制协议式 (7) 的作用下误差继续收敛至 0; 如图 8 所示, 饱和约束下的控制律在 $[-5,5]$ 的范围内抖动并收敛至稳态值。图 5～图 8 结果证实, 本文设计的分组编队协同控制律满足编队队形变换的要求。

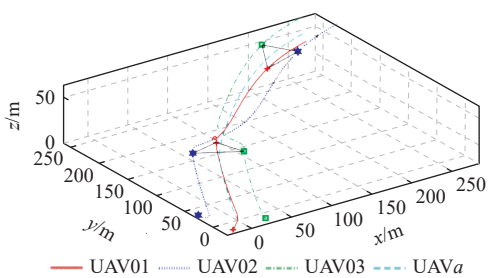


图 5　跟踪-领导者航迹规划

Fig. 5　Tracking-leader path planning

3.2　跟踪-包含控制仿真

表征时变队形的期望编队函数设计如下:

$$\boldsymbol{h}_i(t)=\begin{bmatrix}10\sin\left(0.3t+\left(i-4\bar{i}+3\right)\pi/2\right)\\3\cos\left(0.3t+\left(i-4\bar{i}+3\right)\pi/2\right)\\10\sin\left(0.3t+\left(i-4\bar{i}+3\right)\pi/2\right)\\3\cos\left(0.3t+\left(i-4\bar{i}+3\right)\pi/2\right)\\10\cos\left(0.3t+\left(i-4\bar{i}+3\right)\pi/2\right)\\-3\sin\left(0.3t+\left(i-4\bar{i}+3\right)\pi/2\right)\end{bmatrix}$$

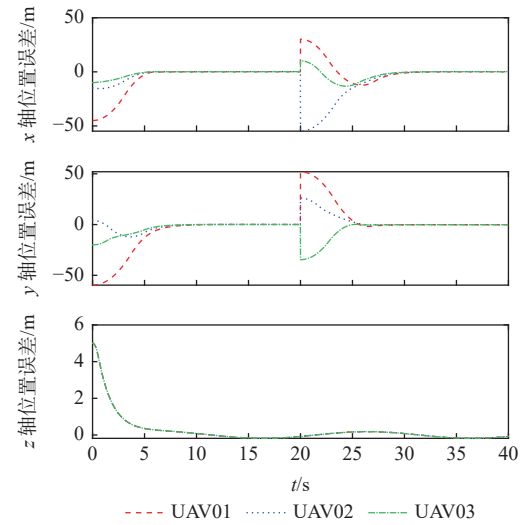


图 6　跟踪-领导者位置误差

Fig. 6　Position errors of tracking-leaders

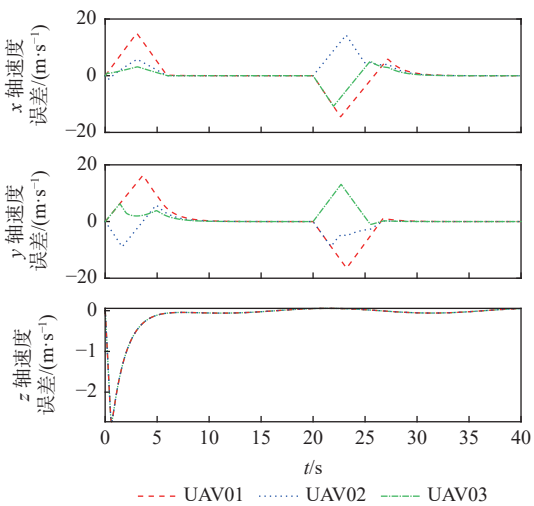


图 7　跟踪-领导者速度误差

Fig. 7　Velocity errors of tracking-leaders

其中, $\bar{i}\in\{1,2,3\}$。

选取 $\Theta=0.01$, 求解 Riccati 方程得到: $\boldsymbol{P}(\Theta)=\begin{bmatrix}0.045\,8 & 0.100\,0\\0.100\,0 & 0.458\,3\end{bmatrix}\otimes\boldsymbol{I}_3$, 据此求得反馈增益矩阵: $\boldsymbol{K}_1=\boldsymbol{K}_2=\begin{bmatrix}-0.100\,0 & 0.458\,3\end{bmatrix}\otimes\boldsymbol{I}_3$

$$\boldsymbol{\Psi}=\begin{bmatrix}0.010\,0 & 0.045\,8\\0.045\,8 & 0.210\,0\end{bmatrix}\otimes\boldsymbol{I}_3$$

根据式 (39), 取 $\sigma_1=8,\sigma_2=19$, 非线性函数 $\boldsymbol{f}(\cdot)$ 由式 (40) 构造。结合分组编队协同策略, 得到 UAV 集群分组编队仿真结果如图 9 所示。

“正方形”表示跟踪-领导者, “三角形”表示编队领导者, “六角星”表示跟随者, $t=0$ s表示各 UAV 的初始状态; $t=19$ s时编队领导者在控制律作用下形成期望编队队形, 且跟随者收敛至领导者形成的飞行凸包内; $t=35$ s时队形参数组已发生变

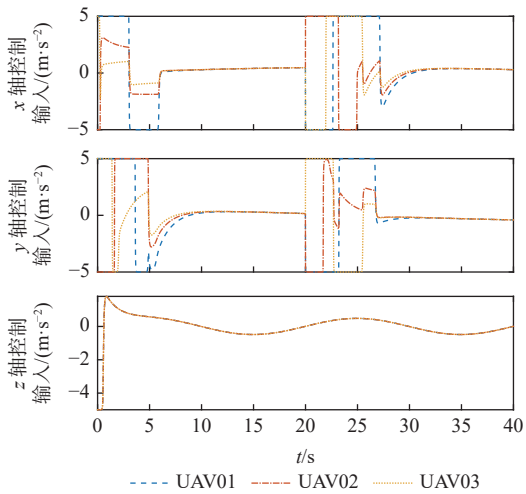


图 8 饱和约束下的控制输入

Fig. 8 Control input under saturation constraint

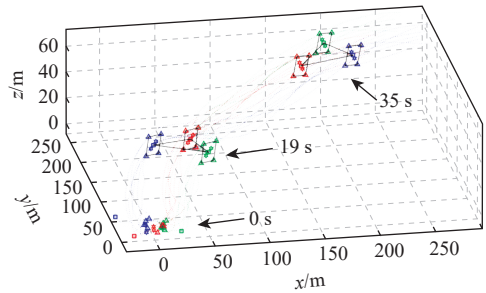


图 9 分组编队运动轨迹

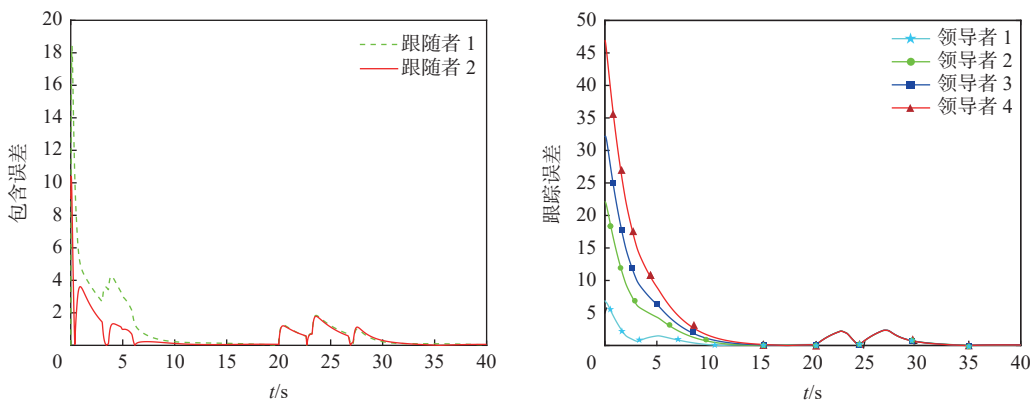
Fig. 9 Trajectory of group formation

化，集群系统实现分组队形变换的同时，编队领导者和跟随者形成期望包含队形。

如图 10～图 12 所示，各分组编队领导者在t = 15 s左右形成期望飞行凸包并跟踪上跟踪-领导者，跟踪误差收敛为 0，由于参数组的变化出现抖动，随后在控制律的作用下收敛为 0；包含误差在t = 15 s左右收敛为 0，而在20～30 s之间，伴随队形参数组的变化，期望集群队形处于调整过程，包含误差出现波动后收敛为 0。

图 13～图 14 描述了基于误差的自适应增益随时间收敛至稳态值，其变化趋势与误差曲线保持一致。

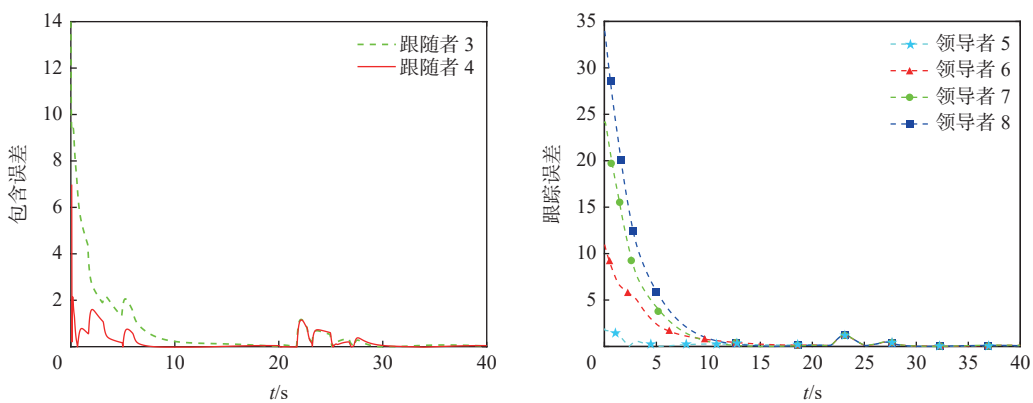
图 9～图 14 仿真结果表明，所设计的基于参数组的分组编队跟踪-包含控制策略能够实现编队间的队形变换，并实现多编队的包含控制。



(a) 包含误差 (b) 跟踪误差

图 10 分组 1 跟踪-包含误差

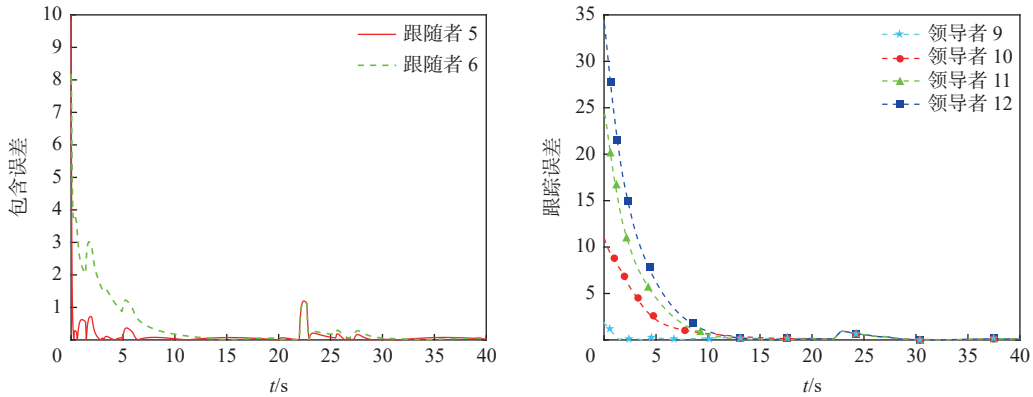
Fig. 10 Tracking-containment errors of subgroup 1



(a) 包含误差 (b) 跟踪误差

图 11 分组 2 跟踪-包含误差

Fig. 11 Tracking-containment errors of subgroup 2



(a) 包含误差　　(b) 跟踪误差

图 12　分组 3 跟踪-包含误差

Fig. 12　Tracking-containment errors of subgroup 3

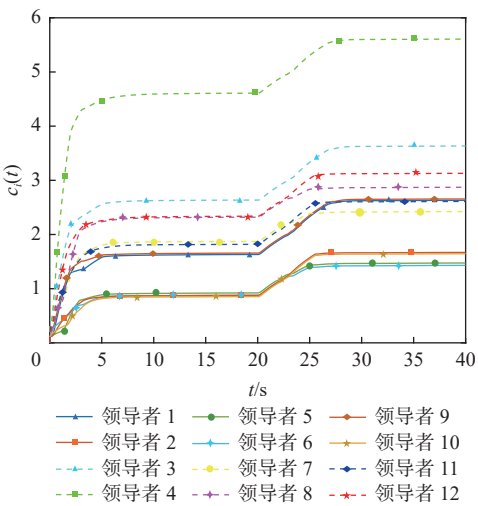


图 13　自适应增益 $c_i(t)$

Fig. 13　Adaptive gain $c_i(t)$

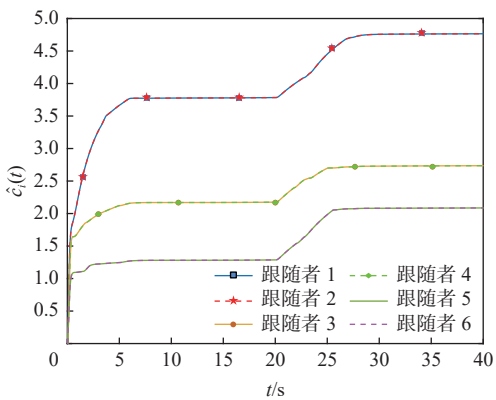


图 14　自适应增益 $\hat{c}_i(t)$

Fig. 14　Adaptive gain $\hat{c}_i(t)$

4　结　论

1) 本文建立了分层分组的集群控制架构, 为实现大规模 UAV 集群的分组协同控制提供了可行性方案;

2) 提出基于队形参数组的分组协同控制策略, 突出了集群系统的顶层规划能力, 实现了集群系统分组编队间的灵活队形变换;

3) 所设计的自适应跟踪-包含控制律不依赖通信拓扑的信息, 即底层控制架构是分布式的;

下一步将对异构集群系统分组编队的合围问题进行深入研究, 并对研究成果做进一步检验。

参考文献（References）

[1] 马雨微. 无人机集群作战系统的群智能任务分配方法研究[D]. 哈尔滨: 哈尔滨工程大学, 2021.
MA Y W. Research on swarm intelligence task allocation method for UAV swarms combat system[D]. Harbin: Harbin Engineering University, 2021 (in Chinese).

[2] 乔琛. 面向无人机集群的协同编队控制算法与仿真实现[D]. 成都: 电子科技大学, 2021.
QIAO C. Collaborative formation control algorithm and simulation for UAV swarm[D]. Chengdu: University of Electronic Science and Technology of China, 2021 (in Chinese).

[3] 吴傲, 杨任农, 梁晓龙, 等. 基于信息素决策的无人机集群协同搜索算法[J]. 北京航空航天大学学报, 2021, 47(4): 814-827.
WU A, YANG R N, LIANG X L, et al. Cooperative search algorithm based on pheromone decision for UAV swarm[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(4): 814-827 (in Chinese).

[4] ALI Z A, HAN Z G. Multi-unmanned aerial vehicle swarm formation control using hybrid strategy[J]. Transactions of the Institute of Measurement and Control, 2021, 43(12): 2689-2701.

[5] SASKA M, HERT D, BACA T, et al. Formation control of unmanned micro aerial vehicles for straitened environments[J]. Autonomous Robots, 2020, 44(6): 991-1008.

[6] 杨明月, 寿莹鑫, 唐勇, 等. 多四旋翼无人机编队保持与避碰控制[J]. 航空学报, 2022, 43(S1): 89-99.
YANG M Y, SHOU Y X, TANG Y, et al. Multi-Quadrotor UAVs formation maintaining and collision avoidance control[J]. Acta Aeronautica et Astronautica Sinica, 2022, 43(S1): 89-99(in Chinese).

[7] 韩涛. 多智能体系统的编队与包含控制问题研究[D]. 武汉: 华中科技大学, 2017.

HAN T. Research on formation and containment control problems of multi-agent systems[D]. Wuhan: Huazhong University of Science and Technology, 2017 (in Chinese).

[8] 赵学远, 周绍磊, 王帅磊, 等. 切换拓扑条件下的多无人机系统编队包含控制[J]. 计算机科学, 2020, 47(S1): 577-582.
ZHAO X Y, ZHOU S L, WANG S L, et al. Formation containment control of multi-UAV system under switching topology[J]. Computer Science, 2020, 47(S1): 577-582 (in Chinese).

[9] 肖支才, 赵学远, 周绍磊, 等. 联合连通拓扑条件下多无人机系统编队包含控制[J]. 电光与控制, 2020, 27(3): 27-32.
XIAO Z C, ZHAO X Y, ZHOU S L, et al. Formation containment control of multi-UAV system under joint connected topology[J]. Electronics Optics & Control, 2020, 27(3): 27-32 (in Chinese).

[10] HAN L, DONG X W, LI Q D, et al. Formation-containment control for second-order multi-agent systems with time-varying delays[J]. Neurocomputing, 2016, 218: 439-447.

[11] 魏志强, 翁哲鸣, 化永朝, 等. 切换拓扑下异构无人集群编队-合围跟踪控制[J]. 航空学报, 2023, 44(2): 258-273.
WEI Z Q, WENG Z M, HUA Y Z, et al. Formation-containment tracking control for heterogeneous unmanned swarm systems with switching topologies[J]. Acta Aeronautica et Astronautica Sinica, 2023, 44(2): 258-273(in Chinese).

[12] HUA Y Z, DONG X W, HAN L, et al. Formation-containment tracking for general linear multi-agent systems with a tracking-leader of unknown control input[J]. Systems & Control Letters, 2018, 122: 67-76.

[13] 李蒙. 基于固定时间收敛的多智能体系统编队包含控制研究[D]. 武汉: 湖北工业大学, 2021.
LI M. Research on formation-containment control of multi-agent systems based on fixed-time convergence[D]. Wuhan: Hubei University of Technology, 2021 (in Chinese).

[14] 田磊, 赵启伦, 董希旺, 等. 异构多智能体系统分组输出时变编队跟踪控制[J]. 航空学报, 2020, 41(7): 323727.
TIAN L, ZHAO Q L, DONG X W, et al. Time-varying output group formation tracking for heterogeneous multi-agent systems[J]. Acta Aeronautica et Astronautica Sinica, 2020, 41(7): 323727 (in Chinese).

[15] LU Y Z, DONG X W, LI Q D, et al. Time-varying group formation-containment tracking control for general linear multiagent systems with unknown inputs[J]. IEEE Transactions on Cybernetics, 2022, 52(10): 11055-11067.

[16] HU J Y, BHOWMICK P, LANZON A. Distributed adaptive time-varying group formation tracking for multiagent systems with multiple leaders on directed graphs[J]. IEEE Transactions on Control of Network Systems, 2019, 7(1): 140-150.

[17] HAN L, XIE Y X, LI X D, et al. Time-varying group formation tracking control for second-order multi-agent systems with communication delays and multiple leaders[J]. Journal of the Franklin Institute, 2020, 357(14): 9761-9780.

[18] MENG Z Y, LIN Z L, REN W. Robust cooperative tracking for multiple non-identical second-order nonlinear systems[J]. Automatica, 2013, 49(8): 2363-2372.

[19] REN W, CAO Y C. Distributed coordination of multi-agent networks: Emergent problems, models, and issues[M]. London: Springer London, 2011.

[20] ZAHREDDINE Z. An extension of the routh array for the asymptotic stability of a system of differential equations with complex coefficients[J]. Applicable Analysis, 1993, 49: 61-72.

[21] CURTIN T, BELLINGHAM J, CATIPOVIC J, et al. Autonomous oceanographic sampling networks[J]. Oceanography, 1993, 6(3): 86-94.

[22] HU J Y, BHOWMICK P, LANZON A. Two-layer distributed formation-containment control strategy for linear swarm systems: Algorithm and experiments[J]. International Journal of Robust and Nonlinear Control, 2020, 30: 6433-6453.

[23] LIN Z L. Low gain feedback[M]. London: Springer London, 1999.

Adaptive group formation tracking-containment control for UAV swarm

ZHANG Yi，YU Hao，YANG Xiuxia*，JIANG Zijie，WANG Chenlei

(Naval Aviation University，Yantai 264001，China)

Abstract: To address the group formation tracking-containment control problem of unmanned aerial vehicles (UAV) swarm, a hierarchical cooperative control architecture including group cooperation and tracking-containment is established. In order to highlight the top-level planning ability of swarm system, a group cooperative control strategy based on formation parameter group is proposed, which overcomes the defect that the existing multi formation control methods are difficult to realize formation transformation. On the basis of considering the communication coupling between groups, an adaptive group formation tracking-containment control law is designed to overcome the dependence of the control input on the eigenvalues of the Laplacian matrix, and the feasibility constraint of the expected formation is given. The saturation constraint is introduced into the control input, and the necessary and sufficient conditions for the stability of the system under the input saturation constraint are further derived to avoid the large-scale maneuver of the swarm. The results show that the proposed strategy can realize the group formation tracking-containment control of UAV swarm.

Keywords: group formation; tracking-containment control; collaborative control; input saturation; formation transformation

Received：2022-04-20；**Accepted**：2022-06-20；**Published Online**：2022-07-06 08：42

URL：link.cnki.net/urlid/11.2625.V.20220705.0908.002

Foundation item：National Natural Science Foundation of Shandong Province (ZR2020MF090)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0193

恶劣天气下机场离场航班运行韧性评估及恢复

王兴隆，赵俊妮，王进*
(中国民航大学 空中交通管理学院，天津 300300)

摘 要：为保障恶劣天气下机场整体性能，科学评估机场离场航班运行韧性，提高航班恢复能力，从而有效缓解天气的影响。给出机场离场航班运行定义，从机场离场航班运行系统性能出发，分析航班离场延误时间、离场总延误时间、离场航班正常率和机场离场航班运行系统综合韧性指数4个指标，对系统在恶劣天气条件下的韧性变化进行评估；提出机场离场航班运行系统性能恢复策略，利用遗传算法对离场延误航班进行优化排序；以2012年北京首都国际机场“721”特大暴雨事件为实例进行数据分析，得到暴雨影响下首都机场的性能指标和韧性指数，对比分析机场离场航班运行系统性能及韧性水平变化。研究结果表明：受暴雨影响，机场离场航班运行系统综合韧性指数由0.457 3下降到0.062 8，暴雨减小后上升到0.222 3；在进行航班优化排序后，离场航班总延误时间减少了24.85%，优化后机场性能恢复速度提升了13.89%，优化后的机场离场航班运行系统最小韧性指数提升了13.38%，系统性能优先恢复至初始状态，表明所提恢复策略有效。

关 键 词：恶劣天气；离场航班运行；韧性模型；综合韧性指数；恢复策略
中图分类号：U8; V355.2
文献标志码：A **文章编号**：1001-5965(2024)01-0110-12

恶劣天气是造成机场容量下降、大量离场航班同时延误、取消的重要原因。恶劣天气发生时，离场航班正常性急剧降低、运输量减少，随着天气消退航班逐步恢复。单一脆弱性或鲁棒性指标已无法准确描述上述动态全过程，为研究恶劣天气从发生到消散过程下机场离场航班运行性能的变化规律，离场航班韧性研究应需而生。韧性一般是指系统受到外部事件扰动时有效抵抗、适应并从影响中恢复的能力，交通系统韧性的概念及其量化方法由Murray-Tuite首次提出，交通系统的韧性包含适应性、快速恢复性等10个方面[1]。Gao等开发了一套预测系统韧性的分析工具，利用车辆的GPS数据量化交通系统的韧性，提出密度、异质性和对称性3个影响系统韧性的关键结构性因素，韧性是可以变好和变差的[2]。Ip和Wang[3]基于对脆弱性和韧性的评价，提出一种交通网络结构优化模型，对中国铁路网络脆弱性和韧性进行评估和分析[3]。Lordan等对全球航空运输网络进行建模，并通过对基于不同标准选择的某些机场进行模拟攻击评估鲁棒性，提出一种可用于检测关键机场的自适应选择策略[4]。王德龙和王超峰运用复杂网络理论构建机场网络级联失效抗毁性评估模型，通过对不同机场节点进行蓄意攻击分析网络节点失效率，进而衡量机场网络抗毁性[5]。但以上方法研究重点聚焦在干扰事件发生后系统保持正常运行的概率或遭受不利影响的程度，忽略了系统偏离正常服务水平后如何快速恢复到稳定状态问题。

为了提升干扰事件发生后系统的恢复速度，D’Lima和Medda使用均值恢复随机模型中的均值回归参数表示韧性，量化系统恢复速度，并将该模型应用

收稿日期：2022-03-28；**录用日期**：2022-07-10；**网络出版时间**：2022-07-19 14:26
网络出版地址：link.cnki.net/urlid/11.2625.V.20220718.1820.002
基金项目：国家重点研发计划(2020YFB1600101)；国家自然科学基金(62173332)；天津市教育委员会自然科学重点基金(2020ZD01)
* **通信作者**. E-mail：jwang@cauc.edu.cn

引用格式：王兴隆，赵俊妮，王进. 恶劣天气下机场离场航班运行韧性评估及恢复[J]. 北京航空航天大学学报，2024，50(1)：110-121.
WANG X L，ZHAO J N，WANG J. Resilience assessment and recovery of airport departure flights under severe weather[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(1)：110-121 (in Chinese).

于伦敦地铁的案例[6]。Sun 和 Zeng 以航空网络为背景, 构建目标节点恢复模型, 设计了一种融合局部介数中心性和局部紧密中心性综合指标混合恢复方法 (hybrid recovery method, HRM)[7]。Nan 和 Sansavini 根据系统受到干扰后不同阶段的状态性能, 确定了吸收、适应和恢复 3 种基本复原能力, 提出综合系统韧性量化的集成度量, 和用于表示基础设施系统的故障行为进行建模, 通过混合多层模型分析电力供应 3 个子系统之间的耦合关系[8]。王兴隆和苗尚飞通过运用复杂网络理论构建扇区网络模型, 定义网络基本特征参数, 分析结构特性,提出扇区网络韧性概念, 并采用定量评估方法对其进行度量[9]。Janić 发现大规模破坏性事件会造成机场、航路容量下降, 进而提出航班延误指标衡量航空网络受损程度, 定义韧性概念为给定条件下研究对象的运行水平, 利用美国遭遇飓风事例验证方法可行性[10]。潘维煌通过机场、航路、管制扇区建立空中交通相依网络, 分析研究受干扰后系统脆弱性表现规律, 提出脆弱性改进策略[11]。周语和邵荃研究不确定因素对航班恢复过程中地面服务保障时间的影响, 构建以最小化延误成本和保证航班公平性为优化目标的航班恢复调度模型, 运用模拟退火算法对航班数据进行仿真[12]。张启钱等以最小航班总延误为目标, 引入滚动时域控制策略, 建立了航班动态排序模型, 并分别采用滚动时域控制策略下的遗传算法和现有的先到先服务算法求解, 证明前者有效[13]。虽然韧性评估和恢复都已经被考虑进机场离场航班运行问题中, 并取得了一定成果, 但评估中大多只考虑单一指标衡量系统性能, 并且仅考虑理想天气状态下的离场航班运行是不够的, 因而特殊天气逐渐成为需要综合考虑的另一因素。

在研究恶劣天气对民航运输系统的影响方面, Inniss 和 Ball 提出单参数机场到达容量分布(arrival capacity distribution, ACD)模型, 考虑天气的随机性, 使用概率容量预测的方法分析不确定天气对机场进离场航班延误的影响[14]。Klein 等提出了一种新的预测模型, 该模型不仅可以预测机场航班延误, 还可以比较不同天气对机场航班延误的影响程度[15]。Kicinger 等分析了几种恶劣天气如低云、低能见度以及雷暴期间机场容量的减少程度, 以天气条件作为约束, 分析其对美国 35 个机场管理的影响[16]。王时敏描述了对恶劣天气的界定标准, 以及典型气象条件对航班延误的影响, 并建立了单一航段受恶劣天气影响的航班延误预测模型[17]。尉炜等基于复杂网络理论, 将天气类型转换为机场容量分布, 研究了天气对航班延误的影响[18]。有关特殊天气对航班影响的研究已有一定的研究成果, 但大多数都是采用仿真天气的方法, 对不确定的天气因素影响下的离场航班运行进行评估与预测, 而很少将实际天气的影响程度进行量化分析。

本文以航班运行为出发点, 评估恶劣天气下机场离场航班运行系统的韧性, 并提出系统受到干扰后对延误航班的有效恢复策略, 对机场离场航班在恶劣天气下及时恢复到稳定状态, 具有重要现实意义。

1　恶劣天气下机场离场航班运行韧性评估

恶劣天气是指会对航班运行效率、航班运行安全造成不利影响的航空气象条件, 如降水、雷暴和风切变等, 影响范围内的机场航班会发生延误甚至取消, 航班运行效率会明显下降, 乘客、机场和航空公司会受到不同方面不同程度的影响。

离场航班运行是一个连续动态过程, 其影响因素种类、形式、数目繁多, 运行性能随条件不断改变, 分析受到恶劣天气等干扰后的离场航班运行韧性对于提高机场航班运行效率至关重要。

1.1　机场离场航班运行韧性评估

由于小时离场航班正常率与航班的延误和取消有关, 考虑单一指标无法准确评估机场离场航班运行性能。由此, 基于机场航班延误和正常率 2 个方面建立 3 个系统性能指标来研究机场系统性能及其变化, 并提出综合韧性指数, 用于评估外界干扰下机场离场航班运行系统韧性。3 个机场离场航班运行性能韧性指标如下。

1) 航班离场延误时间。给定时间段内机场航班离场延误的计算方法不唯一。本文参考文献 [19] 采用以下公式来计算给定时间段 $[t_1,t_2]$ 内航班 f 的离场延误 t_{d}^f:

$$t_{\mathrm{d}}^f=\begin{cases}t_{\mathrm{a}}^f-t_{\mathrm{s}}^f & t_{\mathrm{s}}^f\geqslant t_1,\ t_{\mathrm{a}}^f\leqslant t_2\\ t_{\mathrm{a}}^f-t_1 & t_1\leqslant t_{\mathrm{a}}^f\leqslant t_2\\ t_2-t_{\mathrm{s}}^f & t_1\leqslant t_{\mathrm{s}}^f\leqslant t_2\\ t_2-t_1 & t_{\mathrm{s}}^f\leqslant t_1,\ t_{\mathrm{a}}^f\geqslant t_2\\ 0 & \text{其他}\end{cases}\tag{1}$$

式中: f 为一离场航班; t_{a}^f 为该航班的实际离场时间; t_{s}^f 为该航班的计划离场时间; t_{d}^f 为该航班的延误时间, $t_{\mathrm{a}}^f\geqslant t_{\mathrm{s}}^f$。所以, 机场在给定时间段 $[t_1$-$t_2]$ 的航班离场延误公式为

$$T_{\mathrm{d}}=\sum_{f\in F}t_{\mathrm{d}}^f\tag{2}$$

式中: F 为机场离场航班的集合。航班离场延误时间越小, 代表航班受到恶劣天气影响越小, 系统运

行性能越好。

2）离场总延误时间。离场总延误指机场离场航班遭受恶劣天气影响后造成延误的时间总和。如图1所示，蓝线代表机场离场航班的累计需求，根据时间可得机场离场航班需求量，其斜率表示机场离场航班计划需求率；橙线代表机场航班的实际累计离场航班数，根据时间可得机场实际离场航班数量，其斜率表示机场吞吐率。在初始正常状态下，2条曲线近乎重叠，表明此状态下机场离场航班运行系统供需平衡。从t_1时刻开始，受某种干扰影响，机场实际离场航班减少，实际供应率（吞吐率）降低至低于计划需求率，从而导致供需失衡。t_2时刻，系统开始恢复，并且供应率在短时间内超过需求率。最终，实际的累计离场数量曲线与离场的累积需求曲线再次重叠，供需平衡，系统恢复稳定状态。

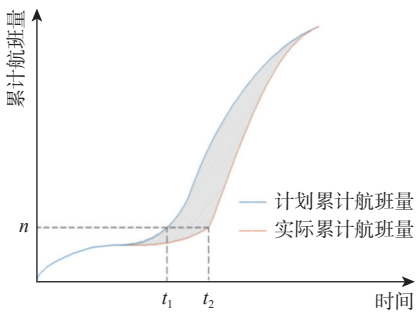


图1 时刻t计划累计航班量与实际累计航班量对比

Fig. 1 Comparison of planned cumulative flight volume and actual cumulative flight volume at time t

若给定一累计航班量n，2条曲线分别对应2个时刻t_1和t_2，t_2–t_1即为n架航班离场总延误。因此，离场总延误D_{SWD}计算公式为

$$D_{\mathrm{SWD}}=\int_{T_s}^{T_e}[S(t)-A(t)]\,\mathrm{d}t \tag{3}$$

式中：$S(t)$为机场离场航班的需求量；$A(t)$为机场实际离场航班数量；T_s为航班时刻表的初始时间；T_e为最后一架离场航班实际起飞时刻。离场总延误指标越小，代表实际供应率与计划需求率越接近，运行性能越稳定。

3）离场航班正常率。正常离场航班与该时间段内计划离场航班总数之比为该时间段的离场航班正常率。

所以，给定时间段$[T_1,T_2]$内，离场航班正常率：

$$P_{T_1,T_2}=\frac{N_{\mathrm{ontime},T_1,T_2}}{N_{\mathrm{schedule},T_1,T_2}} \tag{4}$$

式中：$N_{\mathrm{schedule},T_1,T_2}$表示$[T_1,T_2]$时间段内的计划离场航班总数；$N_{\mathrm{ontime},T_1,T_2}$表示计划离场时间在$[T_1,T_2]$时间段内且离场延误小于30 min的航班总数。单位时间内的正常率越高，表示系统运行效率越高，系统性能越好，准点率的取值范围为[0,1]。

1.2 恶劣天气对机场离场航班运行的影响

机场在遭受干扰后，离场航班运行性能会产生明显波动，其系统韧性过程可大体分为4个阶段：初始稳定阶段、性能下降阶段、性能恢复阶段和新的稳定阶段，如图2所示。

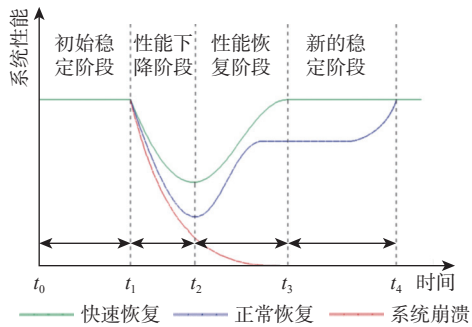


图2 外界干扰下系统韧性变化图

Fig. 2 Change diagram of system resilience under external disturbances

纵轴为随时间变化的系统性能，表征系统性能的指标直接反映系统韧性的变化。恶劣天气下离场航班运行系统的韧性与航班的延误和取消有关，而小时离场航班正常率直接反映了航班的延误和取消。因此，本文选取机场小时离场航班正常率来描述机场离场航班运行系统性能。

图2中t_0～t_4分别为干扰发生前某特定时刻、干扰发生时刻、系统性能下降到最低值时刻、系统性能恢复到某一新的稳定状态时刻、新的稳定状态阶段某特定时刻。

在外界干扰下，韧性的变化过程有3种可能。第1种（图中绿线），干扰发生后，性能随时间先下降后上升，最终恢复到初始稳定状态；第2种（图中蓝线），干扰对系统的影响消失后，韧性先恢复到新的稳定状态，并随时间最终达到初始稳定状态。第3种（图中红线），系统在遭受干扰后性能随时间持续降低，最终崩溃。

韧性过程较为复杂，需要综合的指标评价体系量化系统的韧性。综合韧性指数能够定量地分析不同阶段航班运行韧性，如图3所示，系统受外界干扰影响下的性能变化过程由蓝线描述。

F_0为系统性能的期望值，等于初始稳定阶段的系统性能。F为系统实际性能。外界干扰发生前，系统以初始稳定状态运行。外界干扰发生后，系统性能产生明显波动，随时间先下降后上升，先恢复到新的稳定状态，最终恢复到初始稳定状态。

由文献[20]提出韧性计算公式：

$$\phi(t)=\frac{F(t)-F(t_2)}{F(t_1)-F(t_2)}\qquad t_2\leqslant t\leqslant t_4 \tag{5}$$

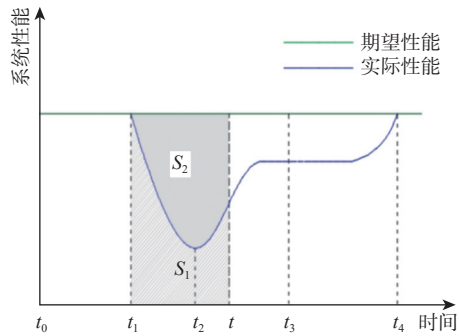


图 3　系统韧性示意图

Fig. 3　System resilience diagram

由式(5)计算得到的韧性指数，为系统性能恢复阶段的性能恢复量与系统遭受干扰性能下降阶段的性能损失量的比值。但该指数只能代表性能恢复阶段的恢复能力，无法代表系统性能损失阶段的抵抗能力，且该指数只能粗略计算某一时刻的系统性能，造成的误差影响较大。

通过上述分析，本文提出一种综合韧性指数来衡量恶劣天气影响下机场离场航班运行系统的性能。由此，系统在时刻 t 的韧性：

$$R(t)=\frac{\int_{t_1}^{t}F(t)\,\mathrm{d}t}{F_0(t)(t-t_1)}\qquad t_1\leqslant t\leqslant t_4 \tag{6}$$

式中：$F_0(t)$为时刻 t 离场航班运行的期望性能；F 为离场航班运行初始稳定状态的性能，令 $F_0(t)=F$；$\int_{t_1}^{t}F(t)\,\mathrm{d}t$表示从恶劣天气发生时刻 t_1 到时刻 t 系统性能的实际累计值；$F_0(t)(t-t_1)$表示从恶劣天气发生时刻到时刻 t 系统的期望累计性能。

$R(t)$表示为从时刻 t_1 到时刻 t，曲线 $F(t)$与横轴间的面积 S_1 与曲线 $F_0(t)$与横轴之间的面积 S_2 的比值。

2　恶劣天气下机场离场航班运行韧性恢复

本文考虑在恶劣天气影响时，通过气象条件确定可用离场时间，对等待离场航班进行重新排序，使离场航班在满足飞机正常起降条件时，能有序实施离场，达到离场航班总延误最小，同时可以有效加快机场离场航班运行系统性能的恢复，使系统性能尽早恢复到稳定状态。

2.1　模型建立

2.1.1　模型假设

为了解决离场航班优化排序问题，做如下假设：

1）机场实施进离场分离运行模式，进离场航班之间互不影响。

2）每个航班在所考虑时间范围内只执行唯一一次离场任务。

3）容量受限时实施流量控制。

4）忽略进离场航路点的容量对其影响。

2.1.2　模型元素

1）变量。i 为机场离场航班流($i=1,2,3,\cdots,m$)，即共要对 m 个航班进行优化排序。

2）权重系数。在航班调度中，有 2 种原则：一种是先到先服务原则，另一种是优先级原则。本文为减小离场总延误，考虑航班机型及延误成本，得到航班优先级，根据优先级对离场航班排序进行优化，保证机场和航空公司运行效益。参考文献 [21]，航空器的运营成本与其最大载客量有关，延误成本包括指挥费、停场费、维修费、起降费、服务费，与飞机起飞质量直接相关。根据不同机型延误成本确定了不同航班优先级 P_{r}^{i}，如表 1 所示。

表 1　不同机型延误成本及航班优先级

Table 1　Delay cost and flight priority of different aircraft types

机型	最大起飞质量/kg	延误成本/(元·h^{-1})	优先级
轻型机(L)	<7 000	208	0.3
中型机(M)	7 000～136 000	2 916	4
重型机(H)	>136 000	4 167	5.7

3）目标函数。本文模型以研究时间段内离场航班总延误最小为目标，所以目标函数表示为

$$f(x)=\min\sum_{i=1}^{n}\left(t_{\mathrm{y}}^{i}-t_{\mathrm{s}}^{i}\right)P_{\mathrm{r}}^{i} \tag{7}$$

式中：t_{y} 为优化后的航班离场时刻；t_{s} 为航班计划离场时刻；P_{r}^{i}为航班 i 的优先级。

4）约束条件。①某一时刻离场流量应不大于该时刻机场可用最大离场容量：$\sum x_i^t\leqslant D_t,\forall t\in T$，其中，$x_i^t$表示航班时刻优化问题的决策变量($i=1,2,3,\cdots,m$)，$D_t$表示给定的机场在时刻 t 的离场容量，T表示航班可用离场时刻集合；②优化后的航班离场时刻不能提前于计划起飞时刻：$t_{\mathrm{y}}^{i}\geqslant t_{\mathrm{s}}^{i}$；③离场航班 j、k 应满足放行间隔，M 为足够大的正整数，具体数值可根据优化时段确定：$C_i\geqslant C_k+S_{ki}-\left(t_{\mathrm{s}}^{k}+M-t_{\mathrm{s}}^{i}+S_{ki}\right)P_{\mathrm{r}}^{i}$，其中，$C_i$ 表示航班 i 实际离场时间；S_{ki} 表示航班 k 与 i 的间隔。当航班 i 在航班 j 之前时等式成立，否则为 P_{r}^{k}。

2.2　基于遗传算法的模型求解

2.2.1　遗传算法简介

遗传算法是一种通过模拟自然进化过程搜索最优解的算法。其主要特点是可以只使用目标函数和相应适应度函数，不需要其他辅助信息；优化结果与初始条件无关，算法独立于求解域；采用概

率化的寻优算法，不需要确定的规则就能自动获取和指导优化的搜索空间，自适应地调整搜索方向。本文选取遭受恶劣天气影响最大时间段的所有航班，所获取数据大多为机场航班运行层面。由于只考虑恢复策略的有效性，而遗传算法具有鲁棒性，更加适用于检验对离场航班重新优化排序是否有利于提高机场离场航班运行系统的恢复能力，故以遗传算法求解。

2.2.2 遗传算法的应用

1）参数编码的设计。本文采用实数编码方法，直接用实数表示基因，容易理解且不需要解码过程，但容易过早收敛，从而陷入局部最优。但此处只是检验对离场航班重新优化排序是否能提高机场离场航班运行系统的恢复能力，是否能作为一种机场在恶劣天气影响下的恢复策略，所以不必求全局最优。

2）适应度函数的设计。适应度函数可以评价个体优劣，其选取直接影响到遗传算法的收敛速度及能否找到最优解。适应度函数表示为

$$F(x)=\frac{1}{\sum_{i=1}^{n}\left(t_{\mathrm{y}}^{i}-t_{\mathrm{s}}^{i}\right)P_{\mathrm{r}}^{i}} \tag{8}$$

3）遗传算子的设计。

① 选择。本文采用锦标赛法进行选择，取 N 为种群中个体数量的百分之十。

② 交叉。交叉分3个步骤：

步骤 1 在种群中随机挑选2个个体。

步骤 2 确定所挑选的2个个体的交叉概率，与交叉率进行比较，若不大于交叉率，执行步骤3，进行交叉，否则，返回步骤1。

步骤 3 随机确定交叉位置：将随机选择的2个个体交换位置处的基因，产生2个新的个体，检验产生的2个新个体是否满足约束条件，将满足的新个体放入下一代种群中，不满足的筛去，重复以上操作，直至新的种群规模达到原来的种群规模。

③变异。变异分4个步骤：

步骤 1 在种群中随机挑选1个个体。

步骤 2 确定所挑选个体的变异概率，与变异率进行比较，若不大于变异率，执行步骤3，进行变异，否则，返回步骤1。

步骤 3 随机确定变异位置，任意选择一个基因将随机选择的个体该位置处的基因替换掉，产生一个新的个体，检验产生的新个体是否满足约束条件，若满足，将新个体放回原来的种群中，替换随机选择的个体，若不满足，保持随机选择的个体不变。

步骤 4 重复以上步骤，达到迭代次数。遗传算法的流程如图4所示。

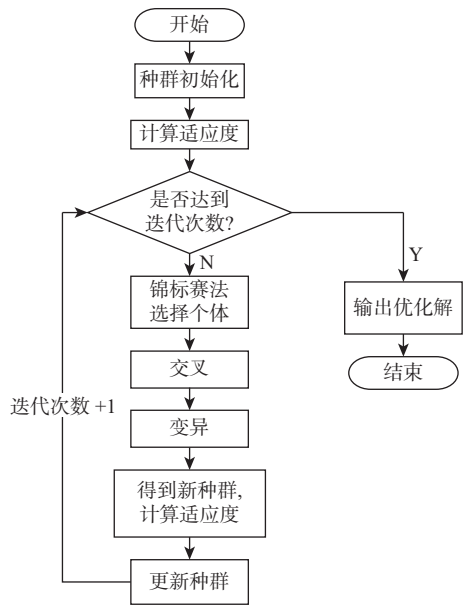


图4 遗传算法流程

Fig. 4 Genetic algorithm flow

3 实例分析

本文实例选取2012年7月21日北京特大暴雨（以下简称“721”特大暴雨）为背景，选用2012年7月20日至7月22日这3天的北京首都国际机场的航班数据。

3.1 离场航班韧性评估

3.1.1 航班小时离场延误

利用式(1)和式(2)计算得到2012年7月20日至22日这3天的北京首都国际机场小时离场航班延误，如图5所示。

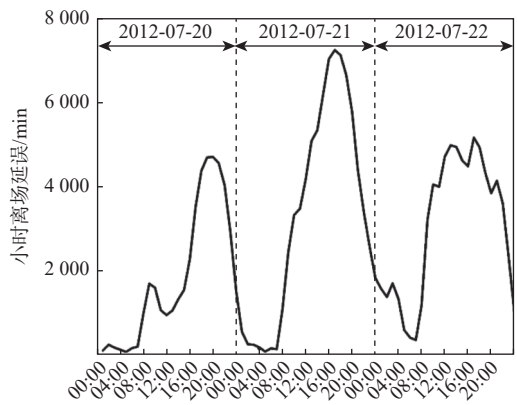


图5 航班小时离场延误

Fig. 5 Flight hours departure delay

如图6所示：每天03:00—05:00机场计划小时离场航班量最低；航班量高峰期为上午07:00—09:00，07:00达到峰值后逐渐减少；21日16:00航班小时离场延误达到最大，此时航班取消量达到峰值。2条

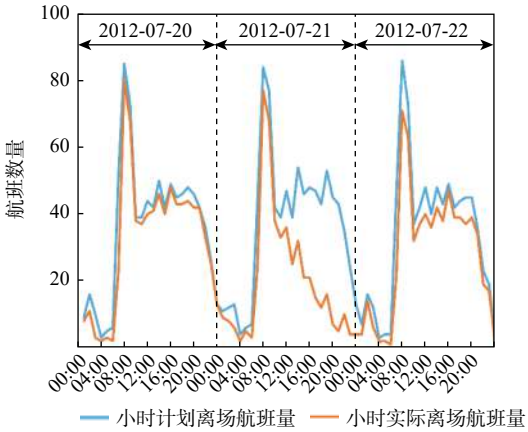


图 6　小时计划离场航班与实际离场航班对比

Fig. 6　Hourly scheduled departure flight and actual departure flight comparison

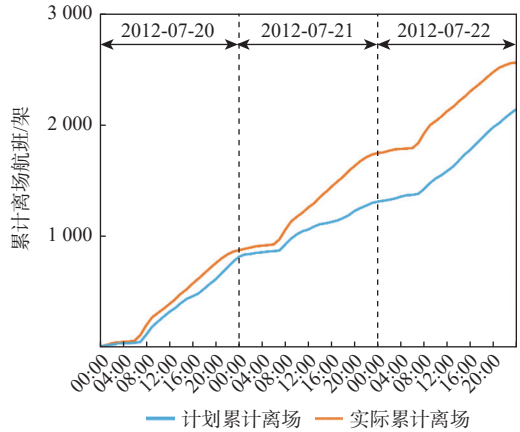


图 7　计划累计离场航班与实际累计离场航班对比

Fig. 7　Comparison of planned cumulative departure flights and actual cumulative departure flights

曲线对比可知：20 日曲线差值较少；而 21 日 2 条曲线相差较大；22 日，2 条曲线差值逐渐减少。表明 21 日特大暴雨导致很多航班取消，22 日机场离场航班运行系统性能逐渐恢复。

结合图 5 和图 6 可以得出：

1）正常情况下，03:00—05:00 航班量最低，机场容量远能满足机场流量需求，延误较少；05:00 之后，机场流量持续增加直至超过机场容量，导致机场延误增大；05:00 之后航班开始明显增多，但 06:00 机场小时延误才开始明显增大，说明延误具有延时性。所以，虽然 07:00—09:00 机场离场航班需求最大，09:00—20:00 机场离场需求一直处于中等水平，但由于离场延误的延时性，15:00—20:00 时机场离场延误最为严重；20:00 至第 2 天 03:00，离场需求降低，机场容量再次满足机场流量需求，延误逐渐减小。

2）22 日和 23 日的机场离场航班小时延误在 08:00—20:00，明显大于 21 日的机场离场航班小时延误，表明恶劣天气会使机场航班延误增大；其中 22 日 16:00 左右，机场离场航班小时延误最大，表明北京首都机场此时的延误情况最为严重，机场离场航班运行系统性能最差。

3.1.2　系统离场总延误

恶劣天气影响下的 3 天内，机场的离场需求和供给，如图 7 所示，因数据量不足，未得到图 2 所示标准曲线，但走势与图 2 相符。

由图 7 可知，20 日 00:00 至 21 日 08:00 左右，机场未遭受特大暴雨期间，机场累计离场需求和机场实际累计离场量基本相等；在 21 日 08:00 以后，由于特大暴雨的影响，系统性能下降，机场实际吞吐能力低于机场累计需求量，供低于需，供需不平衡；在 22 日凌晨暴雨减小后，系统性能逐渐恢复，机场实际吞吐能力缓慢回升。

由曲线走势可以推出：在该时间轴之后的某个时间点，机场实际吞吐能力将恢复到与机场累计需求相平衡的状态，机场性能重新达到新的稳定状态。将以上数据带入式(3)求得此次特大暴雨导致北京首都国际机场在这 3 天产生的离场总延误 D_{SWD}=18 812 min。

3.1.3　日离场航班正常率

根据航班正常率的定义和式(4)，通过编程，求取 3 天的离场航班正常率。

编程求解过程：

1）利用 xlsread 函数将数据导入 MATLAB。

2）利用 for 循环、if 条件语句、mod 取余函数、floor 向下取整函数，将日时分形式的时刻转换为以 2012 年 7 月 20 日 00:00 时开始计时的 min 形式的时刻，运用代码可得到从 2012 年 7 月 20 日 00:00 开始计时，以 min 为单位的每个航班的计划离场时间集合 ss 和实际离场时间集合 rr。

3）为便于之后统计 3 天的离场航班正常率，利用 for 循环和 if 语句将所有航班架次根据计划离场时间分为 3 类：jiaci20、jiaci21、jiaci22，分别代表 20 日、21 日、22 日 3 天的航班。

4）利用 for 循环和 if 语句，分别统计 20 日、21 日、22 日 3 天内运行正常的离场航班记为 z20、z21、z22；再次利用 for 循环和 if 语句，可筛选出 20 日、21 日、22 日 3 天分别被取消的航班数，记为 cancel20、cancel21、cancel22。

5）用一天中实际离场时间与计划离场时间计算得到当天实际正常运行的航班数量，再根据当天离场航班总数，即可得到当天离场航班正常率。

最终，得出 20 日、21 日、22 日 3 天的离场航班正常率分别为 40.65%、14.38%、16.16%。20 日北京

首都国际机场未受暴雨影响阶段，日航班离场正常率为40.65%，21日遭受暴雨影响，日航班离场正常率下降到14.38%，在22日凌晨暴雨停止，系统性能逐渐恢复，因此，22日航班离场正常率略有增长，为16.16%。预计随着时间增长，机场日航班离场正常率能逐渐恢复初始稳定值。

3.1.4 机场离场航班运行系统综合韧性指数

北京首都国际机场位于北京市顺义区，由参考文献[22]得到7月21日19:00至22日03:00顺义区的小时降水量，如图8所示，其中小时降水量大于或等于8.1 mm/h为大雨，在2.6～8.0 mm/h之间为中雨，小于或等于2.5 mm/h为小雨。顺义区小时降水量在21日20:00开始逐渐减小，在22日00:00—01:00之间减小到最低值，之后又逐渐增加，01:00—02:00之间小时降水量达到新高峰，02:00后降水量再次减小，降水逐渐停止。航班起降取决于机场能见度和跑道视程等条件，由于无法得到暴雨影响期间机场的能见度和跑道视程及机场实际小时降水量，所以选取机场所在顺义区小时降水量作为参考。

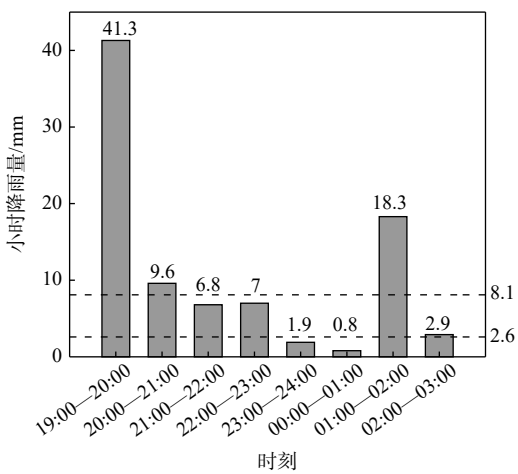


图8 21日19:00至22日03:00顺义区小时降雨量

Fig. 8 Hourly rainfall in Shunyi District from 19:00 on the 21st to 03:00 on the 22nd

利用2012年7月20日23:00至22日11:00北京首都国际机场的离场航班数据，根据民航航班正常统计办法规定，计算得到每小时离场航班正常率，代表该小时机场离场航班运行系统的性能F。

实际情况中，离场航班正常率仍会因除天气以外的其他因素影响而波动，如跑道容量等。所以即使是在机场未受暴雨影响阶段，系统性能稳定但仍会波动，且小时离场航班正常率很难达到100%。在暴雨影响阶段，离场航班正常率可能也会受到除暴雨以外的因素影响。综上，提出假设：机场离场航班运行系统受暴雨影响阶段，影响离场航班正常率的因素除暴雨外，其他均与系统未受暴雨影响阶段影响离场航班正常率的因素一样，且影响程度均相同。7月20日机场未受暴雨影响，所以，以20日小时离场航班正常率的平均值40.24%，作为21日及22日受暴雨影响阶段机场离场航班运行系统性能的期望值，即$F_0=0.402\,4$，则$F(t)$与$F_0(t)$之间的差值均由“721”特大暴雨对系统的影响导致。由系统期望性能$F_0(t)$和系统实际性能$F(t)$画出暴雨影响下首都机场离场航班运行系统性能变化图，如图9所示。

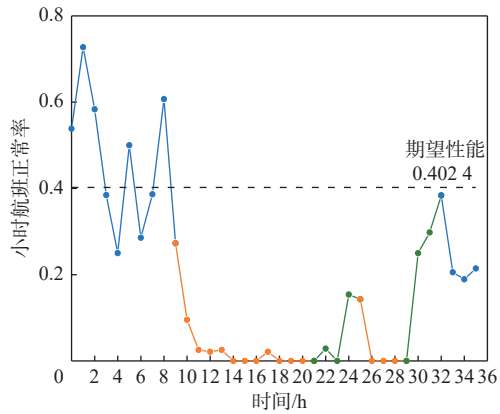


图9 首都机场离场航班运行系统性能变化

Fig. 9 Performance change of flight operation system of Capital Airport Departure

图9横坐标为采样时间20日23:00至22日11:00，共36 h，其中0表示20日23:00至21日00:00，1表示21日00:00至01:00等；纵坐标为机场离场航班运行系统的性能$P(t)$，用机场小时离场航班正常率表示。

分析图9可知，“721”特大暴雨下北京首都机场离场航班运行系统性能变化大体分为6个阶段：

1）初始稳定阶段($0\ \text{h}\leqslant t<8\ \text{h}$)：系统未受到恶劣天气干扰，性能稳定波动。

2）性能下降阶段($8\ \text{h}\leqslant t<21\ \text{h}$)：“721”特大暴雨对首都机场产生了影响，系统性能开始下降。

3）性能初次恢复阶段($21\ \text{h}\leqslant t<25\ \text{h}$)：由图8可知21日20:00至22日01:00顺义区小时降水量明显减少，推测天气应有所好转，对应图9中($21\ \text{h}\leqslant t<25\ \text{h}$)(即21日20:00至22日00:00)系统性能有少量恢复。

4）性能再次下降阶段($25\ \text{h}\leqslant t<29\ \text{h}$)：从22日01:00—02:00该小时降水量明显增长，天气再次恶化，对应图中($25\ \text{h}\leqslant t<29\ \text{h}$)(即22日00:00—04:00)系统性能再次下降到最小值。22日02:00—03:00顺义区降水量再次减小至2.9 mm，但$t>29\ \text{h}$(即22日04:00)开始系统性能才再次恢复，可能是因为雨量虽减小但能见度较低，无法满足飞机正常起降，也

有可能是因为决策延迟, 未能即刻执行飞机起降等。

5) 性能再次恢复阶段(29 h ⩽ t < 32 h): 暴雨对机场的影响减小, 性能开始逐渐恢复。

6) 系统性能再次稳定阶段(32 h ⩽ t < 35 h): 32 h～35 h 时间段内, 性能达到新的稳定波动状态。

由于无法得知具体暴雨开始影响首都机场航班的正常运行时间, 便根据历史航班数据中每架航班的实际起飞时间与计划起飞时间得到航班延误时间, 如图 10 所示, 由图可明显看出大概从第 140 个数据开始, 航班延误逐渐增大, 第 140 个数据计划起飞时刻为 21 日 08:00, 所以机场离场航班大概从 21 日 08:00 左右开始受到暴雨影响, 离场航班延误逐渐增加。因此, 将 21 日 08:00 定为机场离场航班运行开始遭受恶劣天气干扰的时刻, 按图 9 横轴时间表示方式, t_1 = 9 h。因此, 本文从 21 日 08:00 开始每隔一小时计算一次从系统受干扰开始(即从t_1 = 9 h 开始)到此刻t的系统韧性。

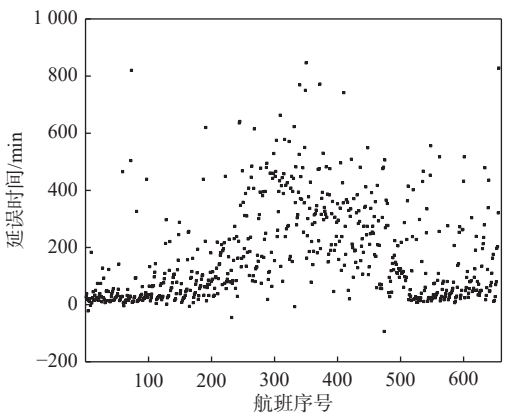


图 10　离场航班延误时间

Fig. 10　Delay of departing flights

在此基础上, 采用式(6)计算得到暴雨影响下北京首都国际机场的机场离场航班运行系统综合韧性指数变化曲线, 如图 11 所示。

图 11 中横坐标表示从 21 日 08:00 至 22 日 11:00, 每一小时记一次。纵轴为横轴所对应小时内机场离场航班运行系统的韧性指数。由图可知, 系统的韧性指数从 21 日 08:00—09:00 的 0.457 3 下降到 21 日 21:00—22:00 的 0.062 8(最小系统韧性指数), 再上升到 22 日 10:00—11:00 的 0.222 3, 表明系统韧性指数也会因恶劣天气的影响先下降后上升, 在系统因暴雨影响性能下降阶段, 韧性指数也有所下降, 在暴雨影响减小, 系统性能恢复阶段, 机场离场航班运行系统韧性也在逐渐恢复。

3.2　机场离场航班运行恢复

由图 8 和图 9 分析可知, 机场降雨量先减后增, 至 22 日 02:00 降雨量允许航班正常起降, 机场离场航班运行性能逐渐恢复。根据中国民用航空气象定义的降水强度划分标准将 21 日 20:00 至 22 日 12:00, 共 16 h, 按每 5 min 一个时间片, 划分为 3 类时间段, 如表 3 所示。

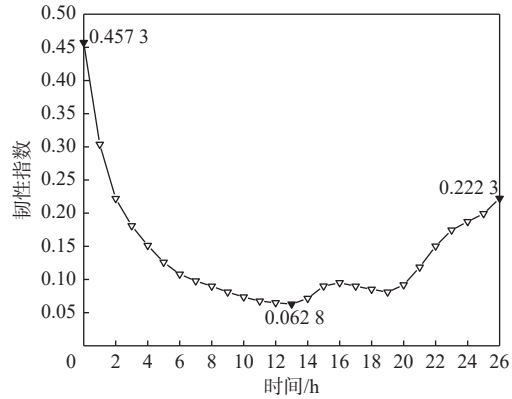


图 11　韧性变化

Fig. 11　Resilience change

表 3　时间分段表

Table 3　Time interval schedule

降雨量	时间段
大雨	[20:00,21:00], [00:55,02:05]
中雨	[21:05,23:20], [00:40,00:50], [02:10,02:30]
小雨	[23:25,00:35], [02:35,12:00]

根据降水强度, 假设大雨情况下机场航班不能实施起降, 中雨情况下机场离场容量为 1 架次, 小雨情况下机场离场容量为 3 架次。

选取实际离场时间为 2012 年 7 月 21 日 21:05 及以后, 计划离场时间为 2012 年 7 月 22 日 10:00 以前的计划离场航班 299 架次 $(m = 299$ 架次$)$, 以离场航班总延误最小为目标, 对航班进行优化排序, 每个航班分配一个可用离场时刻。部分航班数据如表 4 所示。

选取 299 架次航班的实际离场时间, 最后一班飞机实际离场时间为 23 日 11: 51, 为增强仿真的真实性, 选取此次优化排序可分配离场时刻 T 为 [21:05, 23:20], [00:40,00:50], [02:10,02:30], [23:25,00:35], [02:35,12;00], 每 5 min 一个离场时刻, 则共有 165 个可用离场时刻。在中雨时间段内有 36 个可用离场时刻, 每个可用离场时刻可安排 1 架离场航班起飞; 在小雨时间段内有 129 个可用离场时刻, 每个可用离场时刻可安排 3 架离场航班起飞。

本文设种群个体数n = 200, 锦标赛法从种群中随机选择N = 20个个体, 最大迭代次数为 100, 交叉率c = 0.8, 变异率v = 0.1。最终输出优化后的航班起飞时刻表如表 5 所示。

表4 优化前航班时刻表（部分）

Table 4 Flight schedule before optimization (partial)

航班编号	计划起飞时刻	实际起飞时刻	延误时间/min	航班编号	计划起飞时刻	实际起飞时刻	延误时间/min
1	10:50	21:05	615	11	13:25	01:55	750
2	11:45	21:20	575	12	13:25	03:31	846
3	11:55	22:58	663	13	14:00	21:54	474
4	12:05	21:43	578	14	14:05	22:00	475
5	12:30	22:01	571	15	14:05	02:57	772
6	12:50	23:13	623	16	14:15	23:05	530
7	13:00	21:46	526	17	14:45	22:47	482
8	13:00	01:49	769	18	14:50	21:17	387
9	13:15	21:14	479	19	14:55	21:19	384
10	13:20	22:30	550	20	14:55	22:51	476

表5 优化后离场航班时刻表（部分）

Table 5 Optimized departure flight schedule (partial)

航班编号	计划起飞时刻	实际起飞时刻	延误时间/min	航班编号	计划起飞时刻	实际起飞时刻	延误时间/min
1	10:50	21:50	660	11	13:25	22:45	560
2	11:45	21:35	590	12	13:25	03:00	815
3	11:55	22:00	605	13	14:00	22:30	510
4	12:05	21:55	590	14	14:05	22:50	525
5	12:30	22:05	575	15	14:05	22:35	510
6	12:50	22:10	560	16	14:15	23:00	525
7	13:00	22:25	565	17	14:45	23:10	505
8	13:00	22:20	560	18	14:50	23:05	495
9	13:15	22:15	540	19	14:55	23:20	505
10	13:20	22:40	560	20	14:55	03:10	735

将优化后的实际离场时间2012年7月21日21:05及以后，计划离场时间2012年7月22日10:00以前的计划离场航班299架次与优化前的离场航班时刻进行替换，利用式(1)和式(2)计算，得到优化后2012年7月20至22日这3天的北京首都国际机场小时离场航班延误，如图12所示。

由图12可知，由于优化后的航班数量有限，对比图5的曲线差异集中在横轴中间区域，最大小时离场延误时间与7月20日、7月22日差距显著减少。20日00:00至21日08:00左右，机场未遭受特大暴雨；在21日08:00以后，由于特大暴雨的影响，机场离场航班无法正常起飞，图中优化后的延误时间对比图5显著减少；在22日凌晨暴雨减小后，对机场离场航班延误影响降低，小时离场延误也相应减小。

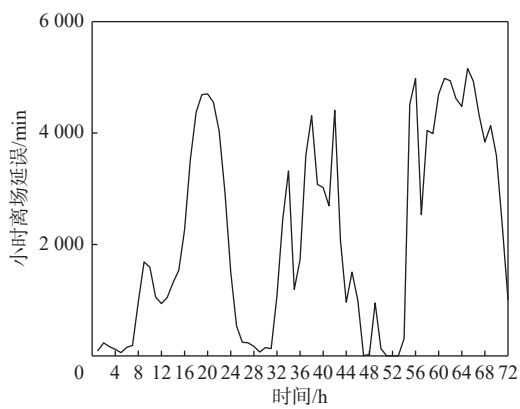


图12 优化后航班小时离场延误

Fig. 12 Optimized flight hour departure delay

优化排序前的离场航班离场总延误为58 137 min，优化排序后为43 690 min，总延误时间减少了24.85%。可以明显看出，对离场航班优化排序可以有效地减少离场航班总延误，提高机场离场航班运行系统恢复能力，使机场快速恢复到稳定状态。

利用优化后数据，求得优化后小时离场航班正常率，得到优化后系统性能变化图，如图13所示。对比图9可见在22日02:00左右便基本恢复到新的稳定状态，优化后机场离场航班运行系统性能恢复速度提高13.89%。

为了分析优化前后系统综合韧性指数变化趋势，本文利用优化排序后航班数据从21日08:00开始每隔一小时计算一次从系统受干扰(即从$t_1 = 9$ h开始)到此刻t的系统韧性，得到优化后北京首都国际机场的机场离场航班运行系统综合韧性指数变

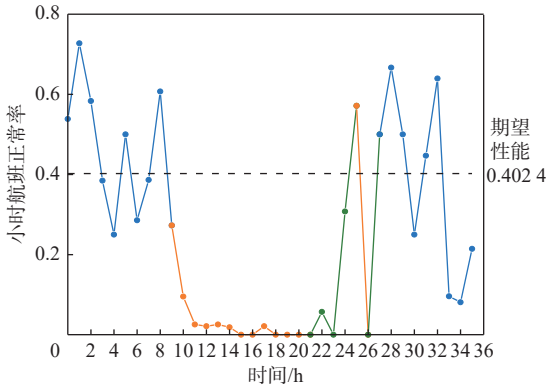


图 13　优化后性能变化

Fig. 13　Change in performance after optimization

化曲线, 如图 14 所示。

由图 14 可知, 系统的韧性指数从 21 日 8:00—9:00 的 0.457 3 下降到 21 日 21:00—22:00 的 0.071 2 (最小系统韧性指数), 再上升到 22 日 10:00—11:00 的 0.450 9。分析图 11 与图 14 可知, 对机场离场航班进行优化排序, 系统遭受恶劣天气后的最小韧性指数由 0.062 8 上升至 0.071 2, 在相同暴雨影响条件下, 优化后的机场离场航班运行系统韧性恢复更迅速, 且更接近于系统原有韧性水平, 表明系统性能可以在遭受干扰后, 快速恢复至初始状态, 因此, 可直观看出进行航班优化排序这一策略对恢复机场离场航班运行具有有效性。

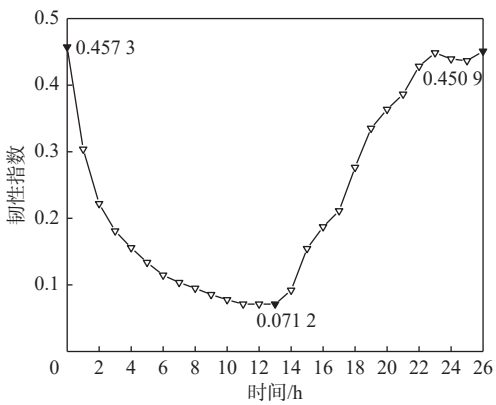


图 14　优化后韧性变化

Fig. 14　Change in resilience after optimization

4　结　论

1) 引入韧性的概念, 以航班作为研究对象, 从机场离场航班运行特性出发, 描述了机场离场航班运行系统的定义; 提出机场离场航班运行性能韧性指标, 评估系统韧性。

2) 利用遗传算法对在恶劣天气下的可用离场时间段内延误航班及该时间段内计划起飞的航班进行优化排序, 对“721”特大暴雨下北京首都机场实例进行分析, 求解说明采用遗传算法对航班进行重新优化排序可以作为一种有效的机场离场航班运行恢复策略。

3) 本文研究假设进离场分离, 未考虑进场航班, 因此, 综合考虑进离场航班运行, 进一步验证航班优化排序策略对提高系统恢复阶段性能的有效性, 是下一步研究的重点。

参考文献 (References)

[1] MURRAY-TUITE P M. A comparison of transportation network resilience under simulated system optimum and user equilibrium conditions[C]// Proceedings of the 2006 Winter Simulation Conference. Piscataway: IEEE Press, 2007: 1398-1405.

[2] GAO J X, BARZEL B, BARABÁSI A L. Universal resilience patterns in complex networks[J]. Nature, 2016, 530(7590): 307-312.

[3] IP W H, WANG D W. Resilience and friability of transportation networks: Evaluation, analysis and optimization[J]. IEEE Systems Journal, 2011, 5(2): 189-198.

[4] LORDAN O, SALLAN J M, SIMO P, et al. Robustness of the air transport network[J]. Transportation Research Part E, 2014, 68: 155-163.

[5] 王德龙, 王超峰. 基于蓄意攻击下的民用机场网络级联失效抗毁性分析[J]. 交通运输工程与信息学报, 2020, 18(3): 172-178.
WANG D L, WANG C F. Analysis of cascading failure and resistance of network in civil airports based on deliberate attacks[J]. Journal of Traffic and Transportation Engineering and Information, 2020, 18(3): 172-178(in Chinese).

[6] D'LIMA M, MEDDA F. A new measure of resilience: An application to the London Underground[J]. Transportation Research Part A:Policy and Practice, 2015, 81: 35-46.

[7] SUN W M, ZENG A. Target recovery in complex networks[J]. The European Physical Journal B, 2017, 90(1): 10.

[8] NAN C , SANSAVINI G. A quantitative method for assessing resilience of interdependent infrastructures[J]. Reliability Engineering & System Safety, 2017, 157: 35-53

[9] 王兴隆, 苗尚飞. 空域扇区网络结构特性分析及韧性评估[J]. 北京航空航天大学学报, 2021, 47(5): 904-911.
WANG X L, MIAO S F. Structural characteristics analysis and resilience assessment of airspace sector network[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(5): 904-911(in Chinese).

[10] JANIĆ M. Modelling the resilience, friability and costs of an air transport network affected by a large-scale disruptive event[J]. Transportation Research Part A:Policy and Practice, 2015, 71: 1-16.

[11] 潘维煌. 基于相依网络理论的空中交通系统脆弱性研究[D]. 天津: 中国民航大学, 2019.
PAN W H. Research on the vulnerability of air traffic system based on interdependent network theory [D]. Tianjin: Civil Aviation University of China, 2019(in Chinese).

[12] 周语, 邵荃. 基于不确定因素扰动的机场大面积航班恢复规划[J]. 科学技术与工程, 2018, 18(16): 300-305.
ZHOU Y, SHAO Q. Airport large-scale flight recovery planning based on uncertainty disturbance[J]. Science Technology and Engineering, 2018, 18(16): 300-305(in Chinese).

[13] 张启钱, 胡明华, 施赛锋, 等. 多跑道航班起降调度优化算法[J]. 交通运输工程学报, 2012, 12(6): 63-68.
ZHANG Q Q, HU M H, SHI S F, et al. Optimization algorithm of flight takeoff and landing on multi-runways[J]. Journal of Traffic and Transportation Engineering, 2012, 12(6): 63-68(in Chinese).

[14] INNISS T R, BALL M O. Estimating one-parameter airport arrival capacity distributions for air traffic flow management[J]. Air Traffic Control Quarterly, 2004, 12(3): 223-251.

[15] KLEIN A, CRAUN C, LEE R S. Airport delay prediction using weather-impacted traffic index (WITI) model[C]// 29th Digital Avionics Systems Conference. Piscataway: IEEE Press, 2010: 1-13.

[16] KICINGER R, SABHNANI G, KRISHNA S, et al. Comparison of the impacts of airport terminal/surface weather hazards[C]// Proceedings of the AIAA Guidance, Navigation, and Control Conference. Reston: AIAA, 2011: 6385.

[17] 王时敏. 恶劣天气对航班延误影响的初步量化研究[D]. 南京: 南京航空航天大学, 2017.
WANG S M. Research on the impact of severe weather on flight delay[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2017(in Chinese).

[18] 尉炜, 邵荃, 向红奕, 等. 基于复杂网络的恶劣天气对航班延误影响的研究[J]. 航空计算技术, 2018, 48(2): 48-51.
WEI W, SHAO Q, XIANG H Y, et al. Research on impact of bad weather based on complex network on flight delay[J]. Aeronautical Computing Technique, 2018, 48(2): 48-51(in Chinese).

[19] WANG Y J, ZHAN J M, XU X H, et al. Measuring the resilience of an airport network[J]. Chinese Journal of Aeronautics, 2019, 32(12): 2694-2705.

[20] HENRY D, EMMANUEL RAMIREZ-MARQUEZ J. Generic metrics and quantitative approaches for system resilience as a function of time[J]. Reliability Engineering & System Safety, 2012, 99: 114-122.

[21] 王楠, 戴福青, 齐雁楠. 基于-跑道容量的航班恢复优化模型[J]. 科学技术与工程, 2020, 20(15): 6279-6285.
WANG N, DAI F Q, QI Y N. Flight recovery optimization model based on runway capacity[J]. Science Technology and Engineering, 2020, 20(15): 6279-6285(in Chinese).

[22] 段伯隆, 张文龙, 刘海文, 等. 北京“7.21”特大暴雨过程暖区降水和锋面降水的时空分布特征[J]. 暴雨灾害, 2017, 36(2): 108-117.
DUAN B L, ZHANG W L, LIU H W, et al. The spatial and temporal distributions of warm sector rainfall and frontal rainfall for the torrential rain event in Beijing on 21 July 2012[J]. Torrential Rain and Disasters, 2017, 36(2): 108-117(in Chinese).

Resilience assessment and recovery of airport departure flights under severe weather

WANG Xinglong，ZHAO Junni，WANG Jin*

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

Abstract：In order to ensure the overall performance of the airport under severe weather, scientifically evaluate the resilience of the airport's flight operations, improve flight recovery capabilities, and alleviate the impact of severe weather effectively. This article first gives the definition of airport departure flight operation. Starting from the performance of the airport departure flight operation system, it analyzes flight departure delay time, total departure delay time, departure flight normality rate and airport departure flight operation system comprehensive resilience index four indicators to evaluate the resilience changes of the system under severe weather conditions.It is important to present airport departure flight operating system's performance recovery method, to employ a genetic algorithm to optimize the order of the delayed departure flights. Finally, this article takes the “721” heavy rain event in Beijing Capital International Airport in 2012 as an example to analyze the data, obtains the performance index and resilience index of the Capital Airport under the influence of heavy rain, and compare and analyze the changes in airport departure flight operating system performance and resilience level. The results indicate that under the influence of heavy rain, the comprehensive resilience index of the airport departure flight operation system decreased from 0.457 3 to 0.062 8, and increased to 0.222 3 after the rainstorm decreased. The delay time is reduced by 24.85%, the airport performance recovery speed is increased by 13.89% after optimization, and the minimum resilience index of the optimized airport departure flight operation system is increased by 13.38%, the system performance is given priority to restore to the initial state, indicating the effectiveness of the proposed recovery strategy.

Keywords：severe weather；departure flight operation；resilience model；composite resilience index；recovery strategy

Received：2022-03-28；**Accepted**：2022-07-10；**Published Online**：2022-07-19 14：26

URL：link.cnki.net/urlid/11.2625.V.20220718.1820.002

Foundation items：National Key R & D Program of China (2020YFB1600101); National Natural Science Foundation of China (62173332); Natural Science Key project of Tianjin Municipal Education Commission (2020ZD01)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0230

基于图对比学习的稳健交通流量预测

刘伟*，贾素玲
(北京航空航天大学 经济管理学院，北京 100191)

摘　　要：作为智能交通系统的核心技术，稳健的交通流量预测是一个长期存在但具有挑战的任务。不论是工业界还是学术界，现有模型需要大量训练数据、易受数据噪声影响而表现不稳健，是限制该领域发展与实际应用落地的重要因素。而在学术界，图对比学习可以通过数据增强与对比损失来降低数据需求量，同时提升模型抵抗数据噪声的能力。提出一种交通流量图对比学习(TFGCL)框架，用于稳健的交通流量预测。TFGCL框架有3个创新点：针对交通流量图(TFG)数据的独特时空特性，TFGCL框架从时间和空间2个角度出发，提出3种TFG数据增强方法。针对TFG数据中语义相似的假负样本，提出一个过滤策略使TFGCL框架能够免受其干扰，从而学习到高质量的表征。TFGCL框架通过联合交通流量预测任务和图对比学习任务进行同时训练。在3个真实交通数据集上与8个基线模型进行对比实验，结果表明：TFGCL框架的预测性能更为稳健，较最优基线模型最高提升6.24%，TFGCL框架的稳健性尤其体现在数据缺失较为明显的数据集和长时交通流量预测任务中。

关　键　词：交通流量预测；交通流量图；数据噪声；图对比学习；预测稳健性
中图分类号：TP399
文献标志码：A　　**文章编号**：1001-5965(2024)01-0122-12

交通流量预测是智能交通系统(intelligent transportation system, ITS)的一个基本问题，对交通管理和公共安全具有重要意义。一个稳健的交通流量预测模型对于现实世界中的很多ITS应用都非常有益。例如，准确的车流量预测可以支撑交通管理部门更好地理解和管理交通拥堵，准确的人流量预测可以帮助大型活动组织者维持人群安全。

交通流量预测旨在基于历史交通流量观测数据，设计预测模型，对未来的交通流量趋势进行准确的预测。这是一个典型的时空数据预测问题，挑战主要来自于时间和空间2个方面：①时间，某个道路历史不同时刻的流量会对该道路未来不同时刻的流量产生影响；②空间，道路之间通过路网连接，不同道路的交通状况会产生相互影响。因此，有效捕捉时空相关性对解决交通流量预测问题至关重要。

得益于深度学习技术的发展，现在的交通流量预测模型可以充分挖掘交通流量数据的依赖关系，包括利用循环神经网络(recurrent neural networks, RNN)[1-2]或时域卷积网络(temporal convolutional networks, TCN)[3]捕捉时间依赖，利用卷积神经网络(convolutional neural networks, CNN)[4]或图神经网络(graph neural networks, GNN)[5-6]捕捉空间依赖，从而预测未来的交通流量。然而，已有的交通流量预测模型通常需要大规模的数据支撑才能取得较好的模型效果，并且模型效果很容易受到数据噪声的干扰[7]，也就是说，基于深度学习的交通流量预测模型的稳健性依赖于数据质量。与此同时，交通预测领域的公开数据集规模通常较小，例如只有几个月，这大大限制了最终可用于训练的样本量；另一方面，交通流量传感器的读数通常存在误差，甚至有时会因为信号中断产生数据丢失。因此，现

收稿日期：2022-04-07；录用日期：2022-05-14；网络出版时间：2022-05-31 14:22
网络出版地址：link.cnki.net/urlid/11.2625.V.20220530.2025.004
*通信作者. E-mail：wayne@buaa.edu.cn

引用格式：刘伟，贾素玲. 基于图对比学习的稳健交通流量预测[J]. 北京航空航天大学学报，2024，50(1)：122-133.
LIU W，JIA S L. Robust traffic flow prediction based on graph contrastive learning[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(1)：122-133 (in Chinese).

实中的交通数据质量较低。为了缓解交通流量预测中数据质量与模型稳健性之间的矛盾, 一个直观的想法是对交通流量数据进行数据增强, 一方面可以扩展训练样本量, 让模型获得足够的训练数据, 另一方面可以通过引入对比学习损失, 让模型能够学到数据中的扰动不变性, 从而产生噪声不敏感的、高质量的交通流量数据表征, 用于稳健的交通流量预测。

目前交通对比学习方面的研究较少, 由于交通流量数据通常定义在交通路网上, 而交通路网可以抽象为图结构 (graph structure), 因此, 本文可以借鉴图对比学习[8] 方法进行交通对比学习。现有的图对比学习方法主要针对静态图数据(即不随时间变化的图数据), 其中不乏一些在无监督任务上取得了卓越性能的研究[8-10]。这些方法的共同理念是对于每个输入图, 采用随机数据增强方法生成 2 个相似视图, 然后输入到图编码器中得到 2 个视图的高级图表征, 最后在表征空间中, 最大限度地拉近具有类似语义信息的图表征 (正样本对), 同时最小化具有不相关语义信息的图表征 (负样本对), 从而让模型学习到数据中的扰动不变性。通常情况下, 对于一个输入图 (也称为锚点), 正样本对就是通过图数据增强生成的该锚点的 2 个视图, 负样本对则由该锚点的视图与其他图的视图构成。通过这种数据增强与对比学习的方式, 可以得到更加稳健的图表征, 用于多种下游任务, 例如节点分类, 链接预测等。

然而, 现有的图对比学习方法无法直接应用于随时间和空间变化的交通流量数据中, 主要有以下几点原因:

1) 2 个阶段的训练适用性差。图对比学习通常包含 2 个阶段, 首先利用对比损失预训练一个表征编码器, 然后利用编码器输出的表征做输入训练线性回归模型做预测, 或者在预训练的编码器后加一个未训练的解码器, 最后微调整个模型进行预测。然而, 文献 [8] 表明, 这种方法比较适用于图分类或图节点分类任务, 在预测任务上表现并不好。

2) 数据增强未充分研究。数据增强的好坏直接决定了对比学习是否有效[11], 好的数据增强可以帮助模型学习到稳健的数据表征, 排除数据噪声的影响。然而, 专门针对随时间和空间变化的交通流量数据增强技术还未被充分研究。

3) 语义相似样本被忽略。在图对比学习方法中, 通常把同一批次不同样本的数据增强结果看作是负样本对。这忽略了交通流量数据中的语义相似性, 例如交通流量的邻近性和周期性[12]。而具有语义相似的交通流量数据样本不应该被当作负样本对, 在表征空间中推远他们之间的距离。

为了解决训练适用性数据噪声语义相似样本等挑战, 本文提出一种交通流量图对比学习 (traffic flow graph contrastive learning, TFGCL) 框架。TFGCL 框架针对交通流量数据独特的时空特性, 对现有的图对比学习范式进行了三大调整: ①本文将原始的交通流量预测损失和对比学习损失相耦合, 以联合训练的方式来提高模型的性能; ②本文从空间域和时间域 2 个角度对交通流量数据进行扰动, 达到数据增强的目的, 从而构建样本对用于图对比学习; ③本文设计了一个策略用于过滤假负样本对 (语义相似度高的负样本对), 避免 TFGCL 框架被假负样本误导。

1　相关工作

历经多年不断的研究和实践, 交通流量预测任务已经取得了重大的进展和丰富的成果, 主要分为 2 个方面: 基于物理模型的方法和数据驱动的方法。基于物理模型的方法通常基于某一理论假设 (例如运动波理论[13]、车辆跟驰理论[14]、排队论[15]), 将交通系统表示为耦合的微分方程组, 然后通过微分方程驱动的交通系统仿真实现交通流量预测。这种方法保证了仿真结果可以表示全域上的交通动态变化, 而不仅仅是基于历史观测数据的拟合。但是, 这样的模型通常对交通系统做了很强的假设, 导致模型可能无法捕捉到现实世界中复杂的人类行为及其他不确定因素[16]。

数据驱动方法在早期主要利用时间序列模型进行交通流量预测, 包括差分整合移动平均自回归 (autoregressive integrated moving average, ARIMA) 模型和向量自回归 (vector auto-regression, VAR)[17] 模型。这些方法通常基于平稳性假设, 因此, 很难捕捉到动态、复杂的交通流量数据中的非线性相关性。近些年来, 得益于深度学习强大的表征能力, 基于深度学习的交通预测广受关注, 并且已经形成了相对标准的模式。对于时间相关性的建模, 常用的方法有 RNN、一维 CNN(1D-CNN) 及 TCN。由于 RNN 在序列数据建模方面的强大能力, 其常被用于交通数据的时间相关性建模, 如时间图卷积网络 (temporal graph convolutional network, TGCN)[18]、交通图卷积长短期记忆神经网络 (traffic graph convolutional long short-term memory neural network, TGCLSTM)[19]、扩散卷积递归神经网络 (diffusion convolutional recurrent neural network, DCRNN)[1]、自适应图卷积递归网络 (adaptive graph convolutional recurrent network,

AGCRN)[2] 等。RNN 在建模中依赖于前一个时间步骤的结果，而 1D-CNN 不受此限制，因此可以实现并行计算，在时间相关性建模方面效率更高，相关研究包括时空图卷积网络 (spatio-temporal graph convolutional networks, ST-GCN)[20]、基于注意力机制的时空图卷积网络 (attention based spatial-temporal graph convolutional networks, ASTGCN)[21] 等。有些模型引入了 TCN，其每个隐藏层的感受野可以指数级地扩张，这样扩张的时间卷积模块可以捕捉更长的时间相关性，相关研究如 Graph-WaveNet[3]、基于图神经网络的多元时间序列预测 (multivariate time series forecasting with graph neural networks, MTGNN)[22] 等。对于空间相关性的建模，常用的方法有 CNN 和 GNN。由于传统的 CNN 模型只能对欧式空间中的数据进行建模，研究人员一般将城市地图按照经纬度划分为标准的网格结构，从而将不同区域的交通状况转换成“图像”形式，然后用 CNN 来学习不同区域之间的空间相关性，如时空残差网络 (spatial-temporal residual network, ST-ResNet)[12]、时空动态网络 (spatial-temporal dynamic network, STDN)[23] 等。但是，真实的城市交通网络 (如路网、地铁网) 更类似于图的结构，而 GNN 适合这种结构，所以目前的研究大多采用 GNN 直接对交通数据进行建模。GNN 模型的一个关键问题是图邻接矩阵的构建，现有研究包括根据路网结构预定义邻接矩阵的方法，如 DCRNN[1]、STGCN[20]、TGCN[18]，以及从数据中自适应学习空间相关性来定义邻接矩阵的方法，如 Graph-WaveNet[3]、MTGNN[22]、层次图卷积网络 (hierarchical graph convolution networks, HGCN)[24] 等。

基于深度学习的交通预测模型通常需要大量且高质量的数据做支撑，然而，在实际应用中，由于传感器采集误差等外部因素，通常导致采集到的数据集中存在很多缺失值和异常值[25]，而且能够获取的数据量通常也是有限的，比如几天或者几个月。因此，仍然需要构建稳健的时空预测模型对这类低质量数据进行良好的预测。本文针对存在误差和缺失的低质量交通流量预测问题展开研究，以加强学术界在这方面的研究。

2 问题定义

2.1 交通流量图

交通流量图 (traffic flow graph, TFG) 描述了一段时间内交通路网 (或传感器网络) 上的交通流量状况，如图 1 所示。其中，交通路网由图 $G=(V,E,\boldsymbol{A})$ 定义，V 为节点集合，每个节点代表一条路段，且

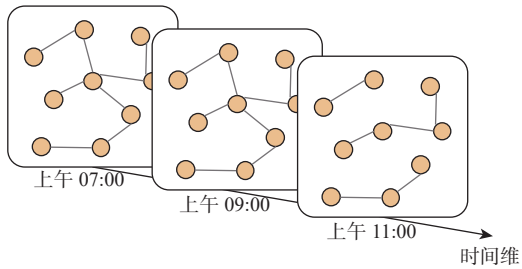


图 1 TFG 数据示例

Fig. 1 Illustration of traffic flow graph

$|V|=N$ 为路段总数，E 为边的集合，描述节点之间的连通性，$\boldsymbol{A}\in\mathbf{R}^{N\times N}$ 为图 G 邻接矩阵；每一时刻 t 的交通流量状况就是定义在图 G 上的特征矩阵 $\boldsymbol{X}^t\in\mathbf{R}^{N\times F}$，$F$ 为特征维度，除了流量之外通常还包含其他的辅助信息 (如天气等)，一段时间内的交通流量状况为 $\boldsymbol{X}^{(t-S):t}\in\mathbf{R}^{S\times N\times F}$。那么交通流量图可以写为 $G=[\boldsymbol{X}^{(t-S):t};G]$，其时间相关性体现在 $\boldsymbol{X}^{(t-S):t}$ 中，而空间相关性由 G 表示。

2.2 交通流量预测

交通流量预测任务就是学习一个函数 f，基于历史 S 个时刻的 TFG 数据预测未来 T 个时刻的交通流量变化，形式化为

$$[\boldsymbol{X}^{(t-S):t};G]\xrightarrow{f(\cdot)}\boldsymbol{Y}^{t:(t+T)} \tag{1}$$

式中：$\boldsymbol{X}^{(t-S):t}\in\mathbf{R}^{S\times N\times F}$ 包含了时刻 $[t-S,t)$ 的输入特征；$\boldsymbol{Y}^{t:(t+T)}\in\mathbf{R}^{T\times N\times F'}$ 为未来 T 步的预测结果。

3 交通流量图对比学习框架

本文提出一个 TFGCL 框架，引入图对比学习来提升交通流量预测模型的稳健性。TFGCL 框架通过图对比学习使得模型学到的交通流量数据表征既能够用于区分不同的交通流量模式，有助于提升模型的性能，又能够抵御数据噪声的干扰，能够增强模型的稳健性。

3.1 图对比学习

图对比学习的目标是学习一个图编码器，其能够有效地从输入数据中抽取图/节点表征。一个典型的图对比学习框架 (如 GraphCL[8]) 的工作原理如下：对于每个输入图，采用随机数据增强方法生成 2 个相似视图，然后输入到图编码器和图读出函数中，分别得到 2 个视图的高级图表征，接着用一个非线性变换 (称为投影头) 将图表征映射到新的隐空间，在该空间中计算对比损失。

模型训练的时候，同一批次会有 M 个图作为输入进行上述操作，因此得到 $2M$ 个图表征。不妨用 $\boldsymbol{z}_{n,i}$、$\boldsymbol{z}_{n,j}$ 表示同一批次内第 n 个图的 2 个相似视图产生的图表征，用 $\phi(\boldsymbol{z}_{n,i},\boldsymbol{z}_{n,j})=\boldsymbol{z}_{n,i}^{\mathrm{T}}\boldsymbol{z}_{n,j}/\left(\|\boldsymbol{z}_{n,i}\|\,\|\boldsymbol{z}_{n,j}\|\right)$ 表示其之间的余弦相似度。那么，GraphCL 中的对比

损失可以写为

$$\mathcal{L}_n = -\ln \frac{\exp(\phi(z_{n,i}, z_{n,j})/\tau)}{\sum_{n'=1, n'\neq n}^{M} \exp(\phi(z_{n,i}, z_{n',j})/\tau)} \tag{2}$$

式中: τ为温度参数, 同一批次内的第n个样本共有$2M-1$个负样本。

为了将预训练的模型用于下游任务, 例如图分类, 通常在图编码器之后, 使用交叉熵损失训练一个线性分类器, 该图编码器可以被冻结 (对应线性评估) 或解冻 (对应微调), 此外, 在模型推理时一般会舍弃投影头。

3.2　交通流量图对比学习

本文 TFGCL 框架如图 2 所示, 其中G'所缺失的边和蓝色的结点表示对 TFG 数据中图结构和特征的扰动。首先, 利用 TFG 数据增强模块 (详见第 3.2.1 节) 将输入数据$G=[\boldsymbol{X}^{(t-S):t}; G]$转化为其相似视图$G'$, 然后, 利用时空编码器将$G$和$G'$映射到表征空间得到$\boldsymbol{H}, \boldsymbol{H}' \in \mathbf{R}^{S\times N\times D}$, 其中$D$为表征空间的维度。表征被输入到 2 个分支中:

1) 预测分支。该分支将表征$\boldsymbol{H}$输入到时空解码器中, 用于未来的预测交通流量, 预测结果$\hat{\boldsymbol{Y}}^{t:(t+T)}$将用于与真实值$\boldsymbol{Y}^{t:(t+T)}$计算交通流量预测损失。

2) 对比分支。该分支以$\boldsymbol{H}$和$\boldsymbol{H}'$作为输入, 利用平均函数在时间和空间维度做池化, 得到表征$\boldsymbol{g}, \boldsymbol{g}' \in \mathbf{R}^D$,利用映射头将图表征映射到新空间中得到$\boldsymbol{z}, \boldsymbol{z}' \in \mathbf{R}^D$, 其中, 映射头由 2 个线性变换组成, 第

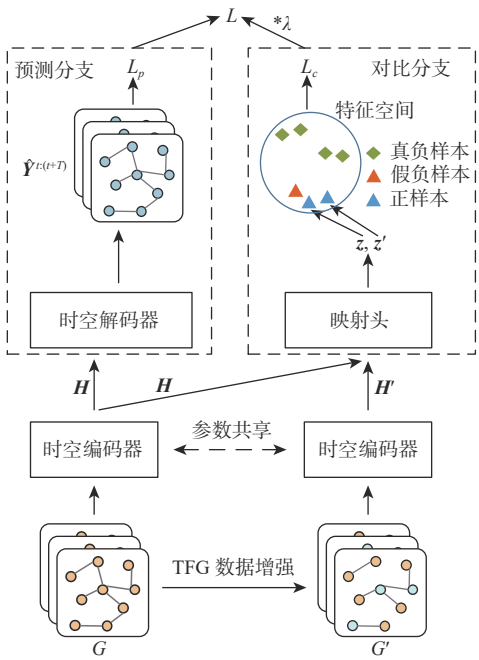


图 2　TFGCL 框架

Fig. 2　TFGCL architecture

1 个线性变换之后带有线性整流 (rectified linear unit, ReLU) 激活函数, 通过对比损失函数来最大化正样本对 ($\boldsymbol{z}$和$\boldsymbol{z}'$) 之间的相似度, 最小化负样本对 (过滤掉假负样本后) 之间的相似度。得到两部分损失之后, 通过对 L_p 和 L_c 两部分损失进行加权求和得到最终的损失 L。

3.2.1　交通流量数据增强

数据增强是对比学习的重要组成部分。其主要负责构建语义上相似的样本, 影响着所学表征的质量[11]。目前针对图的数据增强技术主要考虑空间关系 (例如边扰动、子图采样), 但是没有考虑 TFG 数据独有的时间关系。因此, 本文提出 3 种专门针对 TFG 数据的增强方法, 分别从图结构 (边)、图特征 (节点)、时间域来对输入数据进行扰动, 使得模型学到的表征能够抵御数据噪声的干扰。注意在数据增强的时候只关注 TFG 数据中的交通流量信息, 其他辅助信息不做考虑, 因此本节$\boldsymbol{X}^{(t-S):t} \in \mathbf{R}^{S\times N}$。下面对数据增强方法给出具体的描述:

1) 边遮挡。图数据增强中的边扰动是指对无权图的边进行删除或增加。然而, TFG 数据中的图是带权重的, 在增加边的时候很难去给定一个好的权重, 这使得传统的边扰动方法难以完全迁移到 TFG 数据中。因此, 本文对边扰动进行修正来满足 TFG 数据, 即通过遮挡邻接矩阵$\boldsymbol{A}$中的一部分带权边来扰动图结构。增强后的结果$\boldsymbol{A}'$表示为

$$\boldsymbol{A}'_{ij} = \begin{cases} \boldsymbol{A}_{ij} & M_{ij} > \alpha_1 \\ \boldsymbol{0} & M_{ij} \leqslant \alpha_1 \end{cases} \tag{3}$$

式中: M采样自均匀分布$U(0,1)$; α_1为可调节的参数, 表示边遮挡比例, 例如$\alpha_1=0.1$表示会有 10% 的边会被遮挡。针对同一批次的所有样本, 本文采用同样的增强邻接矩阵$\boldsymbol{A}'$。

2) 节点特征遮挡。受到环境因素的影响, TFG 数据采集过程中通常会出现数据丢失的现象。为了增强模型的抵御数据丢失现象的能力, 本文通过对 TFG 数据的节点特征进行随机遮挡来模拟该现象。通过节点特征遮挡得到的增强 TFG 数据$\boldsymbol{R}_{ij}^{(t-S):t}$表示为

$$\boldsymbol{R}_{ij}^{(t-S):t} = \begin{cases} \boldsymbol{X}_{ij}^{(t-S):t} & M_{ij} > \alpha_2 \\ \boldsymbol{0} & M_{ij} \leqslant \alpha_2 \end{cases} \tag{4}$$

式中: $M \sim U(0,1)$; α_2为可调节的参数, 表示节点遮挡比例。

3) 时间插值。TFG 数据天然地随时间连续变化, 但是负责采集数据的传感器只能给出离散的值 (例如 5 min 间隔的读数)。受此启发, 本文沿着时间轴移动数据, 以利用 2 个连续时间步之间的中间状态。具体实现为

$$\boldsymbol{R}^{(t-S):t}=\beta\boldsymbol{X}^{(t-S):t}+(1-\beta)\boldsymbol{X}^{(t-S+1):(t+1)} \tag{5}$$

式中：$\beta\sim U(\alpha_3,1)$，每一个样本都有其β；α_3为可调节的参数，表示在插值结果的组成中原始数据所占的比例不少于α_3，例如$\alpha_3=0.5$表示每一个插值结果都有不少于0.5的成分来自原始数据，其余的来自下一时刻的数据。这种数据增强方式类似于混淆增强的思想[26]，区别在于本文在连续的时间域上进行加权平均来保证插值的精度。

注意，节点特征遮挡和时间插值都是对节点特征进行干扰。为了保证插值结果的可靠性，实验中先对节点特征进行插值，再对其进行特征遮挡。

3.2.2 时空编/解码器

由于TFG数据具有复杂的时间和空间依赖，本文考虑同时对时间依赖和空间依赖进行建模，保证时空编码器能够有效提取时空依赖。为此，本文提出一个融合图卷积网络(graph convolutional networks, GCN)[27]和Transformer[28]的时空编码器，在保证Transformer建模动态时间信息的同时，引入GCN捕捉空间依赖，如图3所示，GCN为图卷积网络，Concat代表在时间维度上拼接，Att表示注意力机制。由于每个时间步的TFG数据都分别使用了不同的GCN，因此，可以捕捉到不同时间步动态的空间依赖性。

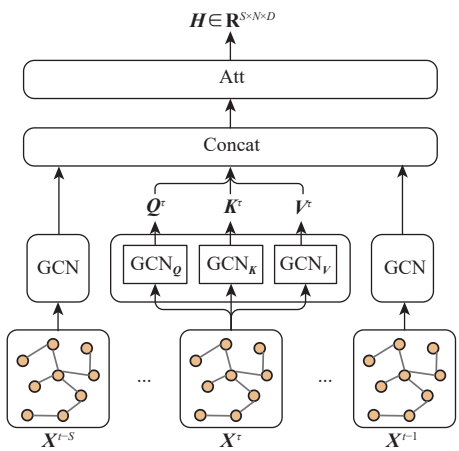


图3 时空编码器。

Fig. 3 Spatio-temporal Encoder

Transformer中最重要的结构是自注意力机制，而自注意力机制的输入：查询($\boldsymbol{Q}$)、键值($\boldsymbol{K}$)和值($\boldsymbol{V}$)，通常由输入数据$\boldsymbol{X}$经过线性变换得到，线性变换不能很好地捕捉TFG数据中的空间依赖，因此，本文提出使用GCN替换线性变换。具体来说，针对每一时刻的TFG数据$\boldsymbol{X}^t\in\mathbf{R}^{N\times F}$，都有：

$$\begin{cases}\boldsymbol{Q}^t=\tilde{\boldsymbol{A}}\boldsymbol{X}^t\boldsymbol{W}_{\boldsymbol{Q}}^t\\ \boldsymbol{K}^t=\tilde{\boldsymbol{A}}\boldsymbol{X}^t\boldsymbol{W}_{\boldsymbol{K}}^t\\ \boldsymbol{V}^t=\tilde{\boldsymbol{A}}\boldsymbol{X}^t\boldsymbol{W}_{\boldsymbol{V}}^t\end{cases} \tag{6}$$

式中：$\tilde{\boldsymbol{A}}\in\mathbf{R}^{N\times N}$为带自环的正则化邻接矩阵；$\boldsymbol{Q}^t,\boldsymbol{K}^t,\boldsymbol{V}^t\in\mathbf{R}^{N\times D}$为GCN的输出结果；$\boldsymbol{W}_{\boldsymbol{Q}}^t,\boldsymbol{W}_{\boldsymbol{K}}^t,\boldsymbol{W}_{\boldsymbol{V}}^t\in\mathbf{R}^{F\times D}$为GCN中待学习的参数矩阵。具体来说，$\tilde{\boldsymbol{A}}=\boldsymbol{I}_n+\boldsymbol{D}^{-1/2}\boldsymbol{A}\boldsymbol{D}^{-1/2}$，$\boldsymbol{I}_n$为单位矩阵，$\boldsymbol{A}$为交通流量图的邻接矩阵；$\boldsymbol{D}=\text{Diag}(\text{rowsum}(\boldsymbol{A}))$为邻接矩阵$\boldsymbol{A}$的度矩阵，rowsum表示行求和，Diag表示构造对角阵。从公式可以看出，针对某时刻的交通流量图$\boldsymbol{X}^t\in\mathbf{R}^{N\times F}$，首先对其右乘一个参数矩阵进行特征的线性变换，并对其左乘了一个带自环的正则化邻接矩阵，使得流量图中每个节点都接收了来自一阶邻居的信息，从而捕捉了局部的空间相关性，当堆叠多个这样的模块之后，就可以捕捉更长距离的空间相关性。为了便于加入时间维度后的计算，将GCN输出结果的维度合并为$d_{\text{model}}=N\times D$，即$\boldsymbol{Q}^t,\boldsymbol{K}^t,\boldsymbol{V}^t\in\mathbf{R}^{d_{\text{model}}}$。

将输入数据$\boldsymbol{X}^{(t-S):t}$分别通过GCN得到$\boldsymbol{Q}^{(t-S):t}$，$\boldsymbol{K}^{(t-S):t},\boldsymbol{V}^{(t-S):t}\in\mathbf{R}^{S\times d_{\text{model}}}$，用于计算注意力系数：

$$\text{Att}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V})=\text{softmax}\left(\frac{\boldsymbol{Q}\boldsymbol{K}^{\mathrm{T}}}{\sqrt{d_{\text{model}}}}\right)\boldsymbol{V} \tag{7}$$

式中：$\text{Att}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V})\in\mathbf{R}^{S\times d_{\text{model}}}$，为了简化公式，本文省略了3个输入的上标。最后，将表征的后二维展开，得到图表征$\boldsymbol{H}^{(t-S):t}\in\mathbf{R}^{S\times N\times D}$，用于交通预测任务和对比学习任务。为了简化书写，省略其上标，用$\boldsymbol{H}$表示学到的数据表征，即$\boldsymbol{H}=\boldsymbol{H}^{(t-S):t}$。

在原始的Transformer中，解码器输出时间步长度与输入时间步保持一致，然后提取键值($\boldsymbol{K}$)和值($\boldsymbol{V}$)用于解码器来完成序列-序列(Seq2Seq)的任务，如机器翻译。由于交通预测任务也是一个标准的Seq2Seq任务，所以时空编码器和时空解码器的交互过程与原始版本的Transformer一致。

3.2.3 双任务联合训练

为了保证交通流量预测的准确性和稳健性，本文对流量预测任务和对比学习任务同时优化。流量预测分支使用平均绝对值误差(mean absolute error, MAE)作为损失函数，定义为

$$\mathcal{L}_p=\frac{1}{M}\sum_{i=1}^{M}\left|\hat{\boldsymbol{Y}}_i^{t:(t+T)}-\boldsymbol{Y}_i^{t:(t+T)}\right| \tag{8}$$

式中：M为一个批次数据内的样本总数。

在对比学习分支中，传统损失函数没有考虑假负样本的情况，这不利于对比学习模型有效学习数据表征。假负样本，即与正样本相似度高的正样本被错标为负样本。这在TFG数据中很常见，因为TFG数据中每个时刻的样本都与其他时刻的样本有关系，例如邻近性和周期性[12]。举例来说，早上07:00和07:30的交通状况变化不会很大，周一早上07:00和周二早上07:00的交通状况也较为相

似, 这周和上周相同时间的交通状况也会呈现出类似的现象。因此, 这些语义相似的样本应该具有相似的表征, 在表征空间中被拉近, 而不是像传统方法一样被当作负样本推远。

为此, 本文提出一个简单的策略用于筛掉假负样本。该策略利用“当天时间”这个输入信息筛掉假负样本, 不将其纳入损失函数的计算当中。具体来说, 设定一个可调节的假负样本阈值α_4, 用t_i代表输入数据的采样时间, 对于锚点数据及其采样时间t, 筛掉假负样本之后的负样本集合定义为

$$\ddot{Z}=\{\ddot{z}_i \mid |t-t_i|>\alpha_4\} \tag{9}$$

于是, 可以得到对比分支的损失函数为

$$L_c=\frac{1}{M}\sum_{i=1}^{M}-\ln\frac{\exp(\phi(z_i,z_i')/\tau)}{\sum\limits_{z_j\in\ddot{Z}}\exp(\phi(z_i,z_j)/\tau)} \tag{10}$$

式中: $\phi(\cdot,\cdot)$为余弦相似度。如果设定$\alpha_3=0$, 那么公式 (10) 就会退化为公式 (2) 中的传统对比学习损失。

最终的损失函数定义如下:

$$L=L_p+\lambda L_c \tag{11}$$

式中: λ用于在模型的 2 个分支之间进行权衡。

4　实验结果与分析

4.1　实验平台

本文所用实验平台的操作系统为 Ubuntu 16.04, 配置一个 64 GB 内存、2 个 CPU (Intel Xeon E5-2 680 v2) 和 1 个 GPU (NVIDIA GeForce GTX 1 080 Ti)。实验使用 Python 3.6.10 和机器学习库 PyTorch 1.4.0 来构建模型, 并利用 CUDA 9.2 加速模型训练。

4.2　数据集

为了评估本文 TFGCL 框架, 本文在 3 个真实交通数据集上进行了实验, 包括 METR-LA[1]、PEMS04 和 PEMS08[29]。

METR-LA 数据集记录了美国洛杉矶高速公路上 207 个传感器的交通流量数据。PEMS04 和 PEMS08 记录了美国加州 2 个不同地区的交通数据, 这些数据由加州交通局测量系统 (Caltrans Performance Measurement System, PEMS) 收集[30], 原始数据中包含采样率为 30 s 的加州高速公路数据, 包含交通流量、平均速度和平均占用率数据, 本文着重对交通流量进行实验分析, 首先将交通流量数据以 5 min 的时间窗口进行汇总, 每天产生 288 个数据点, 然后使用 Z-score 对数据进行标准化, 并按照 7∶1∶2 的比例划分为训练、验证和测试集。

此外, 为了构建多步预测任务, 本文用 12 步的历史数据来预测未来的 12 步。邻接矩阵由路网距离构造[31], 具体为

$$A_{ij}=\begin{cases}\exp\left(-\dfrac{\mathrm{dist}(v_i,v_j)^2}{\sigma^2}\right) & \mathrm{dist}(v_i,v_j)\leqslant k\\ 0 & \text{其他}\end{cases} \tag{12}$$

式中: A_{ij}为邻接矩阵的元素; $\mathrm{dist}(v_i,v_j)$为传感器v_i到传感器v_j之间的路网距离; σ为所有距离的标准差; k为阈值。所有数据集的统计信息如表 1 所示。

表 1　数据统计信息

Table 1　Statistical information on data

数据集	节点个数	边个数	时间步个数	数据缺失率/%
METR-LA	207	1 515	34 272	8.11
PEMS04	307	340	16 992	3.18
PEMS08	170	295	17 856	0.70

4.3　基准模型与评估指标

本文在 3 个评估指标下对比了 8 个基准模型, 评估指标包括平均绝对误差 (mean absolute error, MAE)、均方误差 (root mean square error, RMSE) 和平均绝对百分比误差 (mean absolute percentage error, MAPE), 对比的基准模型如下:

1) 传统模型: 历史平均 (historical average, HA) 模型。将交通流按季节建模, 用历史季节的加权平均值预测。

2) ARIMA 模型: 一种经典的时间序列预测模型。

3) 时空深度学习模型: 门控循环单元 (gated recurrent unit, GRU) 模型。RNN 的变种, 以 GRU 作为隐藏单元的编码器-解码器框架, 适用于时间序列预测。

4) 扩散卷积循环神经网络 (diffusion convolution recurrent neural network, DCRNN)[1] 模型: 用扩散图卷积替换 GRU 中的矩阵相乘以捕捉空间依赖关系, 并构建编码器-解码器框架用于时空预测。

5) 时空图卷积网络 (spatial-temporal graphconvolutional network, STGCN)[20] 模型: 基于卷积网络的模型, 结合图卷积和一维卷积来分别捕捉空间和时间的相关关系。

6) 图小波网络 (graph wavenet, GWNet)[3] 模型: 在图卷积中引入自适应邻接矩阵捕捉动态的空间依赖, 并使用一维因果卷积捕捉时间相关性。

7) AGCRN[2] 模型: 借用了图的思想, 通过自适应模块增强了传统的图卷积, 并将其组合到 RNN 中以捕捉时空相关性。

8) 多变量时间序列图神经网络 (multivariate time series forecasting with graph neural networks,

MTGNN)[22]模型：在学习空间表征的同时优化图结构以学到更好的空间表征，再结合TCN提取时空表征进行多维时间序列预测。

4.4 实现细节与超参数

文中涉及的所有机器学习基准模型都用Pytorch实现。在TFGCL框架中，对于预测分支，本文堆叠了2层改进后的Transformer分别作为时空编码器和时空解码器，其中时空表征的维度$D=128$，本文还用4个注意力头来稳定模型学习效果，其实现机制与原始版本的Transformer一致。最后，改进Transformer的前向传播中隐藏单元数为256。对于对比学习分支，映射头的线性变化中第1层的输出维度为256，对比损失中的温度参数$\tau=0.1$，本文从$\alpha_1=\{0.05,0.1,0.15,0.2,0.25\}$，$\alpha_2=\{0.01,0.03,0.05,0.07,0.09\}$，$\alpha_3=\{0.1,0.3,0.5,0.7,0.9\}$，以及$\alpha_4=\{0\ \text{min},30\ \text{min},60\ \text{min},120\ \text{min}\}$中根据模型在验证集上的效果来选定数据增强的阈值。模型损失函数中的平衡系数通过验证集的效果从$\lambda=\{0.01,0.05,0.1,0.5,1.0\}$中确定。本文采用随机梯度下降优化器，模型的最大训练轮数为100，批次大小为128。对于所有的实验，本文在不同的随机种子上重复5次，并报告了评估指标的平均值。

4.5 测试结果分析

表2为不同模型的实验结果。从表2中可以发现：

表2 不同模型在数据集METR-LA、PEMS04和PEMS08上的性能比较

Table 2 Performance comparison of different models on the METR-LA, PEMS04 and PEMS08 datasets

数据集	模型	MAE			MAPE			RMSE		
		15 min	30 min	60 min	15 min	30 min	60 min	15 min	30 min	60 min
METR-LA	HA	4.145	4.145	4.145	0.129	0.129	0.129	7.772	7.772	7.772
	ARIMA	3.853	4.925	6.915	0.093	0.137	0.174	6.475	9.608	14.539
	GRU	3.879	3.971	4.101	0.109	0.113	0.120	7.468	7.746	8.231
	DCRNN	2.709	3.125	3.599	0.070	0.085	0.105	5.254	6.386	7.620
	STGCN	2.793	3.167	3.540	0.073	0.087	0.102	5.455	6.514	7.525
	GWNet	2.804	3.206	3.587	0.074	0.091	0.108	5.491	6.530	7.421
	AGCRN	2.836	3.209	3.574	0.075	0.089	0.102	5.507	6.535	7.424
	MTGNN	2.660	3.038	3.467	0.069	0.083	0.099	5.114	6.132	7.217
	TFGCL框架	**2.517**	**2.827**	**3.245**	**0.066**	**0.079**	**0.093**	**4.875**	**5.806**	**6.763**
PEMS04	HA	24.520	24.520	24.520	0.167	0.167	0.167	39.838	39.838	39.838
	ARIMA	21.822	24.681	28.869	0.146	0.170	0.221	35.521	41.218	53.266
	GRU	22.441	22.506	22.583	0.155	0.156	0.157	36.286	36.342	36.447
	DCRNN	19.581	21.467	24.864	0.133	0.147	0.175	31.125	34.067	39.228
	STGCN	19.149	20.133	21.567	0.133	0.140	0.150	30.301	31.886	34.200
	GWNet	17.692	18.574	19.956	0.125	0.131	0.149	28.516	29.888	31.848
	AGCRN	18.132	18.834	19.851	0.124	0.130	0.139	29.221	30.464	31.965
	MTGNN	17.929	18.760	20.135	0.133	0.140	0.150	28.837	30.296	32.510
	TFGCL框架	**17.377**	**18.093**	**19.290**	**0.122**	**0.127**	**0.135**	**28.271**	**29.080**	**30.668**
PEMS08	HA	21.193	21.193	21.193	0.138	0.138	0.138	36.644	36.644	36.644
	ARIMA	20.992	24.998	28.953	0.140	0.168	0.179	36.570	39.545	48.436
	GRU	19.992	20.126	20.461	0.124	0.125	0.127	32.276	32.569	33.200
	DCRNN	15.139	16.619	19.345	0.098	0.107	0.125	23.476	25.982	30.058
	STGCN	15.166	16.188	17.819	0.100	0.106	0.116	23.615	25.401	27.818
	GWNet	13.486	14.349	15.672	0.091	0.093	0.108	21.615	23.375	25.855
	AGCRN	14.146	14.962	16.427	0.092	0.097	0.121	22.241	24.055	26.557
	MTGNN	14.001	14.883	16.583	0.097	0.101	0.116	21.968	23.624	26.128
	TFGCL框架	**13.277**	**14.057**	**15.221**	**0.090**	**0.091**	**0.105**	**21.280**	**22.900**	**25.112**

注：其中最优结果用粗体表示，次优结果用下划线表示。

1）TFGCL框架在不同数据缺失率的数据集上表现均为最优。数据缺失率在3个数据集上存在差异（见表1），本文TFGCL框架在所有数据集上表现均为最优，但次优的模型却不同如图4所示，模型性能排序越小越好，例如在PEMS08上是GWNet，在METR-LA上却是MTGNN，同时GWNet在METR-

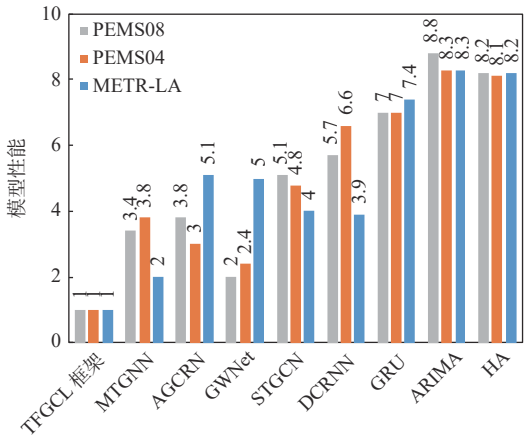


图 4　模型性能在不同数据集的平均排序对比

Fig. 4　Comparison of model performances of average ranking on different datasets

LA 上表现排名第 5。这是因为 TFGCL 框架在训练过程中通过数据增强和对比学习已经适应了数据扰动的情况, 即使在数据缺失的情况下也能学到高质量的表征, 从而做出更稳健的、受数据扰动影响更小的预测。

2) TFGCL 框架在缺失率更大的数据集上提升更多。数据缺失率在不同数据集的排序为 METR-LA>PEMS04>PEMS08(见表 1), 而如表 3 所示 TFGCL 框架相较于最优基线模型的提升排序却是 METR-LA>PEMS04>PEMS08。这种现象一方面说明已有的交通流量预测模型对于数据缺失较为敏感, 另一方面也验证了本文引入的对比学习策略的确能够有效应对这种状况。

表 3　TFGCL 框架较最优基线模型的提升

Table 3　The improvement of TFGCL compared to optimal baselines model

数据集	MAE	MAPE	RMSE	平均得分
METR-LA	6.24	4.54	5.42	5.40
PEMS04	2.57	2.03	2.42	2.34
PEMS08	2.15	2.12	2.15	2.14

3) TFGCL 框架在长时预测上的提升更大。如图 5 所示, TFGCL 框架在长时预测上相较于最优基线模型的提升通常更大, 因为 TFGCL 框架在短时预测和长时预测上的性能平衡地更好, 所以针对不同长度的时间步的预测更加稳健, 而其他基线模型在长时预测时效果损失地较快、也较为明显。

本节 3 个发现说明了本文 TFGCL 框架在真实数据集的交通流量预测问题上表现地更为稳健, 不论是不同数据缺失率还是不同预测时长, TFGCL 框架相较于基线模型都有明显的提升。

4.6　模型稳健性分析

为进一步评估 TFGCL 框架的稳健性, 设计实验如下: 对标准化后的训练数据随机添加高斯噪声, 以交通流量图中每个节点为单位产生数据扰动, 其中, 高斯噪声服从标准正太分布, 扰动比例为 10%、20%、30%、40%、50%。此外, 对被扰动的数据按照 10% 的比例构造数据缺失, 相当于产生更大程度的扰动, 本文采用数据遮挡 (置 0) 实现数据缺失。以不同模型的多步预测性能取平均值作为泛化性能指标, 在同等数据扰动比例的情况下, 模型的泛化性能越好, 则代表模型越稳健。本文仅给出以 RMSE 作为评估指标不同模型在 PEMS08 数据集上的实验结果, 因为其他评估指标和其他数据集上的结果与之类似。

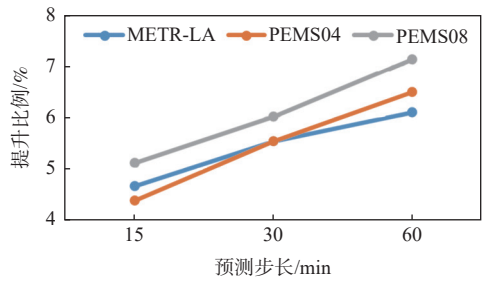


图 5　TFGCL 框架在不同预测步长下较最优基线模型的提升

Fig. 5　Improvement of TFGCL compared to optimal baseline model on different prediction horizon

实验结果如图 6 所示, 从中有 3 个点发现: ① 随着数据扰动率逐渐越大, 模型性能普遍出现下滑趋势, 因为数据扰动会给模型带来干扰导致性能下降; ② 时空深度学习模型性能下滑速度比传统模型快, 这说明时空深度学习模型容易受到数据扰动影响; ③ TFGCL 框架的性能下滑在所有时空深度学习模型中最小, 也就是 TFGCL 框架的稳健性较强, 因为本文专门设计了数据增强模块对输入数据进行扰动, 并利用对比学习损失函数赋予模型抵抗这种干扰的能力, 从而达到稳健的交通流量预测。

4.7　模型消融实验

本文提出 3 种数据增强方法和 1 个假负样本过

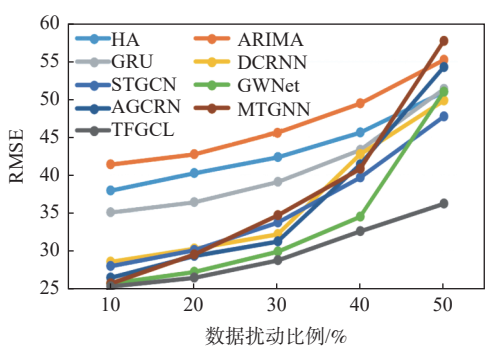


图 6　不同模型在 PEMS08 数据集上的稳健性比较

Fig. 6　Robustness comparison of different models on PEMS08 dataset

滤策略来赋予 TFGCL 框架抵抗数据噪声干扰的能力，提升模型的稳健性。下面对这些模块的有效性进行验证。为此，引入 TFGCL 框架的 4 个变种模型，分别是：① w/o EM，代表没有边遮挡模块的 TFGCL 框架；② w/o NM，代表没有节点属性遮挡模块的 TFGCL 框架；③ w/o TI，代表没有时间插值模块的 TFGCL 框架；④ w/o FN，代表没有假负样本过滤策略的 TFGCL 框架。由于实验结果类似，这里只给模型以 RMSE 为评估指标在验证集上的平均结果。

实验结果如图 7 所示。首先，w/o NM 与 w/o FN 相较于 TFGCL 框架的性能下降较大，验证了节点属性遮挡和假负样本过滤策略的有效性。节点属性遮挡会将所选节点一段时间内的属性全部置 0，模拟了交通流量数据采集过程中通常会出现的数据丢失现象，这直接要求模型必须在部分数据缺失的情况下做出准确的预测，因此，该数据增强方法的有效性最强。假负样本过滤策略排除掉了那些语义相似却被认为是负样本对的样本，避免了模型受到错误的指导，朝着错误的方向优化参数，因为交通流量数据中邻近性和周期性的存在使得传统的负样本对生成方法很容易产生较多的假负样本对，因此该策略的有效性较强，这也说明了时空数据对比学习中排除假负样本的重要性。

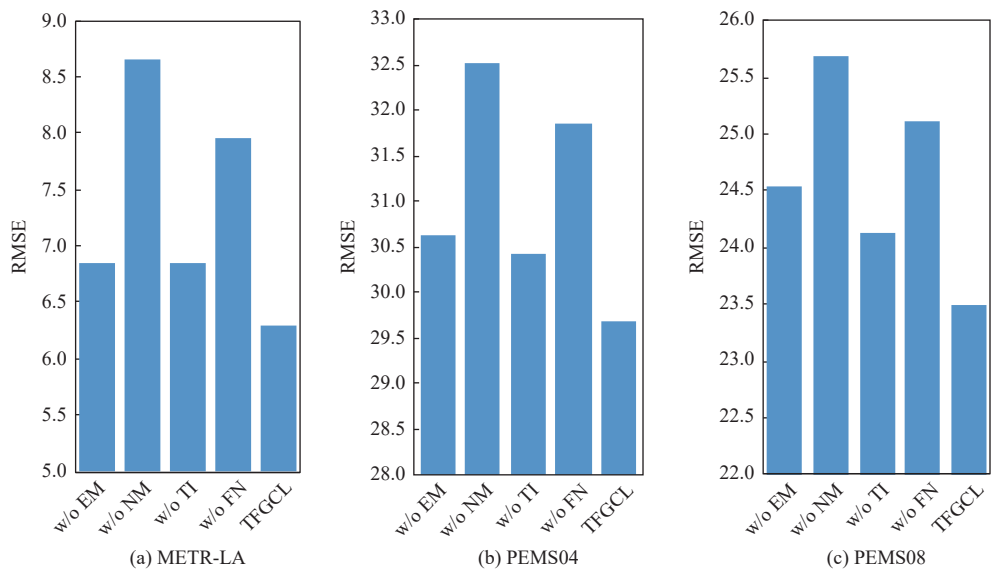


图 7　不同数据增强方法与策略的消融实验

Fig. 7　Ablation study on different data augmentation methods and strategies

其次，w/o EM 与 w/o TI 相较于 TFGCL 框架的性能下降较小，说明边遮挡和时间插值的有效性较弱。因为边遮挡是基于全局随机采样方法的，可能会遮挡掉一些权重较低的对模型来说不重要的边，因此局限了该方法的有效性，观察边遮挡在 3 个数据集上的有效性变化可以发现：在边较多的数据集中，边遮挡的有效性越弱，因为其中有大量低权重的边，而在边较少的数据集中，边遮挡的有效性逐渐提升，因为有较大的概率遮挡掉重要的边，迫使模型在没有改变的情况下也能做预测，从而使得模型的预测更稳健。时间插值相当于一种借助相邻数据进行数据补全方法，该方法相当于对数据进行平滑操作，并不会引入很强的数据增强结构，因此其有效性较为有限。

总的来说，这些变种模型性能排序为：w/o NM < w/o FN < w/o EM < w/o TI，也就是说，所提出数据增强方法和策略的有效性排名为：节点属性遮挡 > 假负样本过滤 > 边遮挡 > 时间插值。

4.8　计算效率对比

本文对比了不同模型在所有数据集上的运行效率。为了公平起见，主要对比性能较优的时空深度学习方法。计算了这些模型每个轮次 (epoch) 的平均训练时长 (train) 以及在验证集上的测试消耗时长 (eval)，以此作为运行效率进行对比，实验结果如表 4 所示。

从表 4 可以发现所有模型的训练时间基本上与测试时间成正比，训练时间越长的模型测试时间通常越长；TFGCL 框架的运行效率总体优于其他模型，因为 TFGCL 框架采用了效率较高的 Transformer 来捕捉全局的时间相关性，而基于 RNN 架构的 DCRNN、AGCRN 需要迭代计算、耗时较久，基于 CNN 的 STGCN、GWNet 等在捕捉全局相关性时需要叠加较多的卷积层导致效率有所下降。虽然 GWNet 在 PEMS08 数据集上的评估时长与本文模型 TFGCL 框架接近，但 TFGCL 框架的训练时长较短，且在该数据集上，TFGCL 框架的性能优于 GWNet

表 4　不同模型的运行效率比较

Table 4　Operational efficiency comparison of different models　　s

模型	train			eval		
	METR_LA	PEMS04	PEMS08	METR_LA	PEMS04	PEMS08
DCRNN	382.9	274.9	262.3	28.6	21.2	19.2
STGCN	40.3	8.7	20.3	4.7	3.8	2
GWNet	33.7	25.4	15	1.2	1	0.5
AGCRN	37.4	25	18.6	2.1	1.4	1
MTGNN	34	24.4	15.7	1.5	1.2	0.7
TFGCL框架	24.2	17.4	11.6	0.8	0.6	0.4

(见表 2)。

综合表 2 与表 4 来看, TFGCL 框架不仅能够取得稳健的交通预测性能, 而且可以达到较低的运行耗时。因此, 在实际应用中, TFGCL 框架的实用性更强。

5　结　论

1) TFGCL 框架针对 TFG 数据的独特时空特性, 对现有图对比学习方法进行改进使其能够赋能交通流量预测问题。

2) 分别从时间和空间 2 个方面设计了 3 种 TFG 数据增强方法用于对比学习, 并提出 1 个过滤策略用于过滤对比学习过程中的假负样本, 这 2 个方面的结合使得模型能够学到高质量的表征, 用于更稳健的交通流量预测。

3) 使用 3 个真实交通数据集, 与 8 种模型进行比较, 发现 TFGCL 框架的性能更加稳健, 尤其是在数据缺失较为明显的数据集上和长时交通流量预测任务中。

为使本文 TFGCL 框架能处理各种类型的时空预测任务, 仍需要优化数据预处理过程和实验过程的各项参数。

参考文献 (References)

[1] LI Y G, YU R, SHAHABI C, et al. Diffusion convolutional recurrent neural network: Data-driven traffic forecasting[EB/OL]. (2018-02-22)[2021-12-20]. https://arxiv.org/abs/1707.01926.pdf.

[2] BAI L, YAO L N, LI C, et al. Adaptive graph convolutional recurrent network for traffic forecasting[C]//Proceedings of the 34th International Conference on Neural Information Processing Systems. New York: ACM, 2020: 17804-17815.

[3] WU Z H, PAN S R, LONG G D, et al. Graph wavenet for deep spatial-temporal graph modeling[C]//Proceedings of the 28th International Joint Conference on Artificial Intelligence. New York: ACM, 2019: 1907-1913.

[4] TANG X F, YAO H X, SUN Y W, et al. Joint modeling of local and global temporal dynamics for multivariate time series forecasting with missing values[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(4): 5956-5963.

[5] 李佩钰. 一种基于小波和神经网络的短时交通流量预测[J]. 计算机技术与发展, 2020, 30(1): 135-139.
Li P Y. Short-term traffic flow prediction based on wavelet and neural network[J]. Computer Technology and Development, 2020, 30(1): 135-139(in Chinese).

[6] TIAN C Y, CHAN W K V. Spatial-temporal attention wavenet: A deep learning framework for traffic prediction considering spatial-temporal dependencies[J]. IET Intelligent Transport Systems, 2021, 15(4): 549-561.

[7] 冯宁, 郭晟楠, 宋超, 等. 面向交通流量预测的多组件时空图卷积网络[J]. 软件学报, 2019, 30(3): 759-769.
FENG N, GUO S N, SONG C, et al. Multi-component spatial-temporal graph convolution networks for traffic flow forecasting[J]. Journal of Software, 2019, 30(3): 759-769(in Chinese).

[8] YOU Y N, CHEN T L, SUI Y D, et al. Graph contrastive learning with augmentations[EB/OL]. (2021-04-03)[2021-12-21]. https://arxiv.org/abs/2010.13902.pdf.

[9] XU D K, CHENG W, LUO D S, et al. InfoGCL: Information-aware graph contrastive learning[EB/OL]. (2021-10-28)[2021-12-23]. https://arxiv.org/abs/2110.15438.pdf.

[10] ZHU Y Q, XU Y C, YU F, et al. Graph contrastive learning with adaptive augmentation[C]// Proceedings of the Web Conference 2021. New York: ACM, 2021: 2069-2080.

[11] CHEN T, KORNBLITH S, NOROUZI M, et al. A simple framework for contrastive learning of visual representations[C]// Proceedings of the 37th International Conference on Machine Learning. New York: ACM, 2020: 1597-1607.

[12] ZHANG J B, ZHENG Y, QI D K. Deep spatio-temporal residual networks for citywide crowd flows prediction[C]// Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence. New York: ACM, 2017: 1655-1661.

[13] DAGANZO C F, LAVAL J A. Moving bottlenecks: A numerical method that converges in flows[J]. Transportation Research Part B:Methodological, 2005, 39(9): 855-863.

[14] 田钧方, 朱陈强, 贾宁, 等. 基于轨迹数据的车辆跟驰行为分析与建模综述[J]. 交通运输系统工程与信息, 2021, 21(5): 148-159.
TIAN J F, ZHU C Q, JIA N, et al. Review of car-following behavior analysis and modeling based on trajectory data[J]. Journal of Transportation Systems Engineering and Information Technology, 2021, 21(5): 148-159 (in Chinese).

[15] SAEED N, LARSEN O I. Application of queuing methodology to analyze congestion: A case study of the Manila International Con-

tainer Terminal, Philippines[J]. Case Studies on Transport Policy, 2016, 4(2): 143-149.

[16] MO Z B, SHI R Y, DI X A. A physics-informed deep learning paradigm for car-following models[J]. Transportation Research Part C:Emerging Technologies, 2021, 130: 103240.

[17] JORG B, HAMILTON J D. Time series analysis[J]. Contemporary Sociology, 1995.

[18] ZHAO L, SONG Y J, ZHANG C, et al. T-GCN: A temporal graph convolutional network for traffic prediction[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 21(9): 3848-3858.

[19] CUI Z Y, HENRICKSON K, KE R M, et al. Traffic graph convolutional recurrent neural network: A deep learning framework for network-scale traffic learning and forecasting[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 21(11): 4883-4894.

[20] YU B, YIN H T, ZHU Z X. Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting[C]// Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence. California: International Joint Conferences on Artificial Intelligence Organization, 2018: 3634-3640.

[21] GUO S N, LIN Y F, FENG N, et al. Attention based spatial-temporal graph convolutional networks for traffic flow forecasting[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33(1): 922-929.

[22] WU Z H, PAN S R, LONG G D, et al. Connecting the dots: Multivariate time series forecasting with graph neural networks[C]// Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: ACM, 2020: 753-763.

[23] YAO H X, TANG X F, WEI H A, et al. Revisiting spatial-temporal similarity: A deep learning framework for traffic prediction[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33(1): 5668-5675.

[24] GUO K, HU Y L, SUN Y F, et al. Hierarchical graph convolution network for traffic forecasting[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(1): 151-159.

[25] YI X, ZHENG Y, ZHANG J, et al. ST-MVL: Filling missing values in Geo-sensory time series data[C]//Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence. New York: AAAI Press, 2016: 2704-2710.

[26] ZHANG H Y, CISSE M, DAUPHIN Y N, et al. Mixup: Beyond empirical risk minimization[EB/OL]. (2018-04-27)[2021-12-25]. https://arxiv.org/abs/1710.09412.pdf.

[27] KIPF T N, WELLING M. Semi-supervised classification with graph convolutional networks[EB/OL]. (2017-02-22)[2021-12-26]. https://arxiv.org/abs/1609.02907.pdf.

[28] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]// Proceedings of the 31st International Conference on Neural Information Processing Systems. New York: ACM, 2017: 6000-6010.

[29] SONG C, LIN Y F, GUO S N, et al. Spatial-temporal synchronous graph convolutional networks: A new framework for spatial-temporal network data forecasting[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(1): 914-921.

[30] CHEN C, PETTY K, SKABARDONIS A, et al. Freeway performance measurement system: Mining loop detector data[J]. Transportation Research Record:Journal of the Transportation Research Board, 2001, 1748(1): 96-102.

[31] SHUMAN D I, NARANG S K, FROSSARD P, et al. The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains[J]. IEEE Signal Processing Magazine, 2013, 30(3): 83-98.

Robust traffic flow prediction based on graph contrastive learning

LIU Wei*, JIA Suling

(School of Economics and Management, Beihang University, Beijing 100191, China)

Abstract: Robust traffic flow prediction, as the core technology of Intelligent Transportation Systems, is a long-standing but challenging task. The fact that current models need a lot of training data and are susceptible to noise disturbance is a major factor that is restricting the growth of this subject. In academia, graph contrastive learning can alleviate the data-demanding issue and improve the model's ability to resist data noise through data augmentation and contrastive learning. Therefore, this paper proposes a Traffic Flow prediction framework that incorporates graph contrast learning (TFGCL) for robust traffic flow prediction. The framework has three contributions: First of all, given the unique spatio-temporal characteristics of traffic flow graph (TFG) data, TFGCL proposes three TFG data augmentation methods from the perspective of time and space. Secondly, in order to learn high-quality representations, this work also suggests a filtering method to shield the model from harsh negative samples with identical semantics. Finally, TFGCL jointly trains the traffic flow prediction task and the graph contrastive learning task. Extensive experiments with 8 baselines on 3 real traffic datasets show that the prediction performance of the TFGCL framework is more robust (an improvement of 6.24% compared to the best baseline), especially in datasets with obvious data missing and long-term traffic flow forecasting tasks.

Keywords: traffic flow prediction; traffic flow graph; noise perturbation; graph contrastive learning; robustness of prediction

Received: 2022-04-07; **Accepted**: 2022-05-14; **Published Online**: 2022-05-31 14:22

URL: link.cnki.net/urlid/11.2625.V.20220530.2025.004

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0234

星载SAR天线展开机构装配精度确信可靠性建模

李云[1]，张湘霄[2]，孙富强[2,*]，李晓阳[2]，韩建超[1]
(1. 北京卫星制造厂有限公司，北京 100094；2. 北京航空航天大学 可靠性与系统工程学院，北京 100191)

摘 要：星载合成孔径雷达（SAR）天线展开机构是空间桁架结构，其杆系和铰链的空间位置相互耦合，在装配过程中主要依赖于人工装调，装配精度难以准确控制，装配可靠性直接影响星载SAR天线的服役性能。为此，基于可靠性科学原理，提出星载SAR天线装配精度确信可靠性建模方法。考虑天线装配过程中影响装配精度的内因和外因，确定表征天线展开机构装配精度的性能指标及阈值，建立性能裕量模型。开展多源不确定性分析与量化，构建星载SAR天线展开机构装配精度确信可靠性模型。以某型号星载SAR天线展开机构为研究对象，进行可靠性敏感度分析和验证。研究结果表明：所提方法可为星载SAR天线展开机构装调提供理论指导，有效提高装配精度和效率。

关 键 词：确信可靠性；天线展开机构；装配精度；性能裕量；不确定性
中图分类号：TB114.3
文献标志码：A **文章编号**：1001-5965(2024)01-0134-10

星载合成孔径雷达(synthetic aperture radar，SAR)天线是卫星的重要组成部分，也是直接执行卫星功能的重要物理平台，在通信、对地观测、军事侦察等众多领域具有广泛应用[1]。星载SAR天线展开机构是一种典型的空间可展桁架机构，其结构尺寸大且较为复杂，杆系等组件之间多用铰链联结，且杆系和铰链的空间位置相互耦合、相互影响。目前，SAR天线展开机构的装配主要依赖于人工装调，装配精度难以准确控制，导致装配周期长且费时费力，严重影响卫星研制进度。此外，装配过程还受到装配工艺、工装、人员操作等各种不确定性因素的影响，导致装配性能降低，最终影响星载SAR天线在轨展开的指向精度与性能[2]。为保证装配精度，工程人员经常在装配过程中投入大量精力和时间进行装调。然而，由于缺少有效的理论指导，航天企业目前主要采用测量、调整、测量的纯经验盲调模式[3]。因此，需要对星载SAR天线展开机构的装配精度进行可靠性建模，分析影响装配精度的主要因素，为星载SAR天线高精度装配提供理论指导。

装配精度可靠性是评价装配质量高低的重要指标，通常定义为：产品在规定的装配条件下，实际精度满足设计要求的概率[4]。近年来，学者们针对天线展开机构装配过程中的各类误差源与装配精度可靠性之间的关系进行了分析研究。Szekely等[5]考虑天线展开机构的杆件结构参数存在随机误差，利用随机有限元法对星载SAR天线展开机构的可靠性进行预测。张建国等[6]考虑尺寸误差与环境因素的影响，建立了展开机构的非概率可靠性分析模型。刘良[7]假设天线展开机构杆长误差与铰链间隙误差分别服从正态分布和均匀分布，基于运动学原理建立其装配精度可靠性模型。吴建云等[8]通过Adams软件建立展开机构的动力学模型，通过蒙特卡罗仿真分析方法计算了铰链锁定随

收稿日期：2022-04-12；**录用日期**：2022-05-06；**网络出版时间**：2022-05-23 15:54
网络出版地址：link.cnki.net/urlid/11.2625.V.20220523.1000.001
基金项目：可靠性与环境工程技术重点实验室稳定支持项目 (WDZC20220101)
*通信作者. E-mail：sunfuqiang@buaa.edu.cn

引用格式：李云，张湘霄，孙富强，等. 星载SAR天线展开机构装配精度确信可靠性建模 [J]. 北京航空航天大学学报，2024，50（1）：134-143. LI Y，ZHANG X X，SUN F Q，et al. Belief reliability modeling for assembly accuracy of spaceborne SAR antenna deployable mechanism [J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50（1）：134-143（in Chinese）.

机误差对装配精度误差的影响。杨玉龙和关富玲[9]利用随机有限元法建立天线展开机构平面度与杆长加工误差的随机数据模型。丁建中和王春洁[10]采用矩阵法分块建模的方法建立了铰链间隙与展开机构平面度的分析模型。余德文等[3]考虑天线展开机构的装配过程及原理, 从几何角度关系出发, 建立了星载 SAR 天线机构装配精度与杆长误差及铰链安装位置误差的关系模型。Greene[11]采用有限元法建立了天线展开机构杆长随机误差与其装配精度的关系模型。Mobrem[12]采用有限元法分析了天线结构单个构件制造公差的随机分布对天线展开机构精度影响。Hedgepeth[13]认为天线结构的随机制造误差是引起其精度变化的主要原因, 并建立了其与天线表面精度变化的理论模型。总的来说, 目前大部分研究针对天线展开机构的设计、制造误差分析等随机不确定性分析其对可靠性的影响, 尚未综合考虑装配过程中测量、操作误差等人因认知不确定性和制造误差等随机不确定性对装配精度可靠性的综合影响, 实现可靠性的分析与量化。

针对上述问题, 本文将确信可靠性理论[14]引入到装配精度分析中, 基于可靠性科学原理, 提出星载 SAR 天线展开机构装配精度确信可靠性建模与分析方法, 解决装配精度难以准确控制的问题, 提高天线展开机构装配质量和效率。首先, 建立星载 SAR 天线展开机构装配精度的理论计算模型, 并从性能裕量的角度, 引入装配精度的阈值, 建立星载 SAR 天线展开机构装配精度性能裕量模型。其次, 对裕量模型开展多源不确定性分析与量化, 考虑 3 个主要的不确定性来源: 间隙配合误差、测量设备误差及人因操作误差。然后, 构建星载 SAR 天线展开机构装配精度确信可靠性模型。最后, 以具体型号的星载 SAR 天线展开机构为研究对象进行方法验证, 通过开展可靠性敏感度分析, 为星载 SAR 天线展开机构定量装调提供理论指导。

1　性能裕量建模

1.1　天线展开机构装配过程

星载 SAR 天线展开机构主要由星体支撑架、内框架组件、外框架组件、桁架杆、支撑杆组件、90°铰链、180°铰链等组成, 如图 1 所示。天线内框架组件通过 2 个 90°锁定铰链与星体连接, 天线外板通过 2 个 180°锁定铰链与天线内板连接, 天线内外板依靠支撑杆增强与星体的连接刚度。

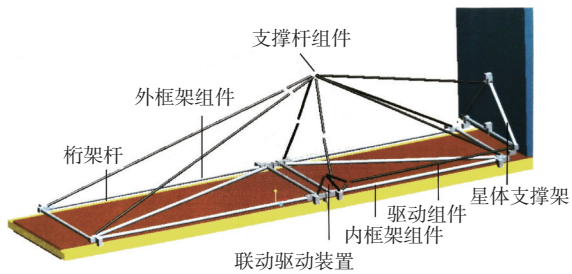


图 1　星载 SAR 天线展开机构组成

Fig. 1　Composition of space borne SAR antenna deployable mechanism

星载 SAR 天线展开机构的装配工艺流程如图 2 所示。具体过程为: 模拟天线框架就位后, 操作人员根据设计图纸, 借助测量仪器和安装设备定位并锁定天线面板上铰链和连接各杆件的空间铰链位置; 将展开机构框架、桁架和星体支撑架等杆件安装在各铰链之间, 并完成联动驱动机构等安装; 安装完成后, 将天线与模拟墙对接, 测试天线面板的装配精度和性能。装配精度达标后, 将天线与模拟墙、模拟板分解, 完成与真实框架的对接与转运等工作。

在测试装配精度环节中, 当所有铰链锁定在理想位置、杆系保持绝对刚性的理想状态下, 天线展开机构可以得到理想的装配精度。而在实际装配过程中, 由于加工误差、装配误差和测量误差的影响, 造成杆系内部存在装配应力, 无法达到规定装

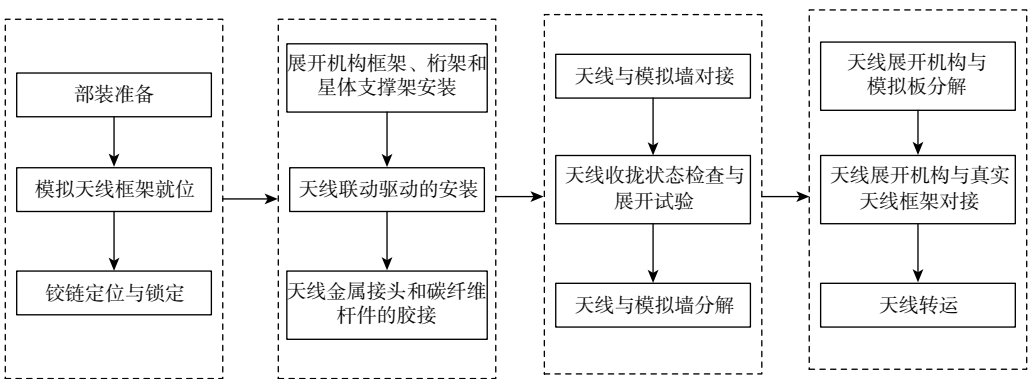


图 2　星载 SAR 天线展开机构装配工艺流程

Fig. 2　Assembly process of spaceborne SAR antenna deployable mechanism

配精度，操作人员将进行多次调试，才能满足装配要求。由于人工装调主要依赖经验，无法精确控制需要调整的零部件及数值，导致装调过程十分繁琐和复杂。

由于星载SAR天线在装配过程中性能退化趋势并不明显，不考虑其确定性退化规律。首先，在星载SAR天线功能性能分析的基础上，基于裕量可靠原理，结合性能阈值建立裕量模型；然后，针对裕量模型从内在属性与外界条件2个角度开展不确定性分析，并基于不确定理论量化不确定性，构建星载SAR天线装配过程的确信可靠性模型。

1.2 装配精度性能裕量模型

由于SAR天线展开机构内外板间隙较小，铰链锁定后两板结构刚度较大，且装配过程中先使用模拟框架（假板）进行安装，在整个装调过程中，两板原有的平面度保持不变，因此，在建模过程中，将其平面度相等，并将连接各杆件（O_1A_1、O_1A_2、O_1A_3和O_1A_4等）中间铰链造成的安装误差直接体现在杆长误差上，简化后的天线展开机构如图3所示。

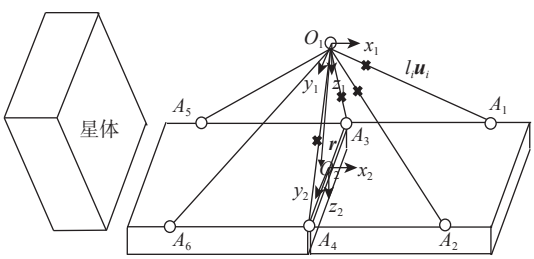


图3 天线展开机构简化示意图

Fig. 3 Simplified schematic diagram of antenna deployable mechanism

天线展开机构的工作原理为：在点A_6处设置电机驱动组件，作为主动驱动源，在O_1A_1杆和O_1A_2杆中间铰链分别安装储能弹簧，作为被动驱动源，使得展开机构具有确定的展开运动，在O_1A_3、O_1A_4杆处布置锁定机构，天线完全展开后实现自动锁死。当天线在轨展开后，需要由支撑机构来保证其展开精度，而支撑机构的支撑位置精度则由锁定机构的铰链来确定。装配过程中，由于装配精度不足，造成支撑机构、锁定铰链等安装位置误差，最终导致天线展开精度降低。展开精度包括天线的指向精度和平面精度，与天线展开机构的变形程度有关。因此，在装配过程中，可用天线展开状态下的变形程度表征装配精度，本文将针对天线装配精度进行性能裕量建模。

从几何关系出发，建立天线展开机构的装配精度模型，表示天线装配精度与各支撑杆件、铰链误差之间的关系。具体过程如下[3]：以天线面板几何中心O_2为坐标原点建立坐标系$O_2x_2y_2z_2$，其中，x_2轴方向沿天线面板的长边方向，正向由铰链A_6指向A_2，y_2轴方向沿天线面板的短边方向，正向由铰链A_1指向A_2，z_2轴正向由右手法则确定。以支撑杆交汇中心O_1为坐标原点建立坐标系$O_1x_1y_1z_1$，其中，x_1轴、y_1轴、z_1轴分别与x_2轴、y_2轴、z_2轴平行，且方向一致，如图3所示。

基于星载SAR天线展开机构的结构，从几何角度关系出发，由坐标原点O_1、O_2和天线铰链位置形成6个封闭环，其中，第i个封闭环O_1-O_2-A_i-O_1的杆长误差Δl_i为

$$\Delta l_i = \left[\boldsymbol{u}_i^{\mathrm{T}} \quad (\boldsymbol{R}\boldsymbol{p}_i \times \boldsymbol{u}_i)^{\mathrm{T}}\right][\Delta\boldsymbol{r} \quad \Delta\boldsymbol{\theta}]^{\mathrm{T}} + \left[\boldsymbol{u}_i^{\mathrm{T}}\boldsymbol{R}\right]\left[\Delta\boldsymbol{p}_i\right] \tag{1}$$

式中：$\boldsymbol{u}_i$为支撑杆交汇中心O_1指向铰链A_i的单位矢量，$\boldsymbol{u}_i=(x_i, y_i, z_i)$，$i$=1,2,⋯,6；$\boldsymbol{p}_i$为铰链$A_i$在坐标系$O_2x_2y_2z_2$中的坐标；$\boldsymbol{R}$为坐标系$O_2x_2y_2z_2$到$O_1x_1y_1z_1$的齐次变换矩阵；$\Delta\boldsymbol{\theta}=[\Delta\theta_x, \Delta\theta_y, \Delta\theta_z]$为天线展开机构面板在坐标系$O_1x_1y_1z_1$中的姿态误差；$\Delta\boldsymbol{r}=[\Delta r_x, \Delta r_y, \Delta r_z]$为天线展开机构面板在坐标系$O_1x_1y_1z_1$中的位置误差；$\Delta\boldsymbol{p}_i$为铰链$A_i$在坐标系$O_2x_2y_2z_2$中的安装位置误差，$i$=1,2,⋯,6。

6个封闭环O_1-O_2-A_i-O_1均满足式(1)，将其完全组合可得到

$$\Delta\boldsymbol{L} = \boldsymbol{J}_v\Delta\boldsymbol{V} + \boldsymbol{J}_p\Delta\boldsymbol{P} \tag{2}$$

式中：$\Delta\boldsymbol{V}=[\Delta\boldsymbol{r}, \Delta\boldsymbol{\theta}]^{\mathrm{T}}$；$\Delta\boldsymbol{L}=[\Delta l_1, \Delta l_2, \Delta l_3, \Delta l_4, \Delta l_5, \Delta l_6]^{\mathrm{T}}$分别为6根杆的长度误差；$\Delta\boldsymbol{P}=[\Delta\boldsymbol{p}_1, \Delta\boldsymbol{p}_2, \Delta\boldsymbol{p}_3, \Delta\boldsymbol{p}_4, \Delta\boldsymbol{p}_5, \Delta\boldsymbol{p}_6]^{\mathrm{T}}$分别为6个铰链的安装位置误差；$\boldsymbol{J}_v$、$\boldsymbol{J}_p$为计算天线展开机构装配精度的传递矩阵，与星载SAR天线展开机构的支撑结构间的几何关系相关，具体表达为

$$\boldsymbol{J}_v = \begin{bmatrix} x_1 & y_1 & z_1 & \boldsymbol{R}\boldsymbol{p}_1\cdot x_1 & \boldsymbol{R}\boldsymbol{p}_1\cdot y_1 & \boldsymbol{R}\boldsymbol{p}_1\cdot z_1 \\ x_2 & y_2 & z_2 & \boldsymbol{R}\boldsymbol{p}_2\cdot x_2 & \boldsymbol{R}\boldsymbol{p}_1\cdot y_1 & \boldsymbol{R}\boldsymbol{p}_2\cdot z_2 \\ x_3 & y_3 & z_3 & \boldsymbol{R}\boldsymbol{p}_3\cdot x_3 & \boldsymbol{R}\boldsymbol{p}_3\cdot y_3 & \boldsymbol{R}\boldsymbol{p}_3\cdot z_3 \\ x_3 & y_4 & z_4 & \boldsymbol{R}\boldsymbol{p}_4\cdot x_4 & \boldsymbol{R}\boldsymbol{p}_4\cdot y_4 & \boldsymbol{R}\boldsymbol{p}_4\cdot z_4 \\ x_5 & y_5 & z_5 & \boldsymbol{R}\boldsymbol{p}_5\cdot x_5 & \boldsymbol{R}\boldsymbol{p}_5\cdot y_5 & \boldsymbol{R}\boldsymbol{p}_5\cdot z_5 \\ x_6 & y_6 & z_6 & \boldsymbol{R}\boldsymbol{p}_6\cdot x_6 & \boldsymbol{R}\boldsymbol{p}_6\cdot y_6 & \boldsymbol{R}\boldsymbol{p}_6\cdot z_6 \end{bmatrix} \tag{3}$$

$$\boldsymbol{J}_b = \begin{bmatrix} \boldsymbol{u}_1^{\mathrm{T}}\boldsymbol{R} & & & & & \\ & \boldsymbol{u}_2^{\mathrm{T}}\boldsymbol{R} & & & & \\ & & \boldsymbol{u}_3^{\mathrm{T}}\boldsymbol{R} & & & \\ & & & \boldsymbol{u}_4^{\mathrm{T}}\boldsymbol{R} & & \\ & & & & \boldsymbol{u}_5^{\mathrm{T}}\boldsymbol{R} & \\ & & & & & \boldsymbol{u}_6^{\mathrm{T}}\boldsymbol{R} \end{bmatrix} \tag{4}$$

由式(2)可推得天线装配精度$\Delta\boldsymbol{V}$在坐标系$O_1x_1y_1z_1$中与6个杆长误差、6个铰链安装位置误差的关系为

$$\Delta\boldsymbol{V} = \left[\Delta\boldsymbol{r} \quad \Delta\boldsymbol{\theta}\right]^{\mathrm{T}} = \left[\boldsymbol{J}_v^{-1} \quad \boldsymbol{J}_v^{-1}\boldsymbol{J}_b\right]\left[\Delta\boldsymbol{L} \quad \Delta\boldsymbol{P}\right]^{\mathrm{T}} = \boldsymbol{J}_v^{-1}[\Delta\boldsymbol{L}]^{\mathrm{T}} + \boldsymbol{J}_v^{-1}\boldsymbol{J}_b[\Delta\boldsymbol{P}]^{\mathrm{T}} = \boldsymbol{J}_v^{-1}\left([\Delta\boldsymbol{L}]^{\mathrm{T}} + \boldsymbol{J}_b[\Delta\boldsymbol{P}]^{\mathrm{T}}\right) \tag{5}$$

将式 (3)、式 (4) 代入式 (5), 可得到天线展开机构装配精度的具体表达形式为

$$\Delta \boldsymbol{V}=\boldsymbol{J}_v^{-1}\begin{bmatrix}\boldsymbol{u}_1^{\mathrm{T}}\boldsymbol{R}\Delta\boldsymbol{p}_1+\Delta l_1\\ \boldsymbol{u}_2^{\mathrm{T}}\boldsymbol{R}\Delta\boldsymbol{p}_2+\Delta l_2\\ \boldsymbol{u}_3^{\mathrm{T}}\boldsymbol{R}\Delta\boldsymbol{p}_3+\Delta l_3\\ \boldsymbol{u}_4^{\mathrm{T}}\boldsymbol{R}\Delta\boldsymbol{p}_4+\Delta l_4\\ \boldsymbol{u}_5^{\mathrm{T}}\boldsymbol{R}\Delta\boldsymbol{p}_5+\Delta l_5\\ \boldsymbol{u}_6^{\mathrm{T}}\boldsymbol{R}\Delta\boldsymbol{p}_6+\Delta l_6\end{bmatrix}=\begin{bmatrix}\Delta r_x\\ \Delta r_y\\ \Delta r_z\\ \Delta\theta_x\\ \Delta\theta_y\\ \Delta\theta_z\end{bmatrix} \tag{6}$$

已知 $\boldsymbol{J}_v^{-1}$ 为 6×6 矩阵, 则天线展开机构装配精度 $\Delta\boldsymbol{V}$ 为 6×1 矩阵, 分别代表天线展开机构在坐标系 $O_1x_1y_1z_1$ 上的位置误差与姿态误差。在实际装配过程中, 天线面板的平面度误差可以较好地衡量天线展开机构发生变形的程度, 因此,选择天线面板平面度误差作为衡量天线装配精度的主要指标。天线面板的平面度误差 ΔT 指被测平面相对于基准平面在 z 轴位置上的变动量。因此, 平面度误差 ΔT 可表示为

$$\Delta T=\Delta r_z \tag{7}$$

式中: Δr_z 为天线面板在 z_1 轴方向上的变动量, 与杆长误差 $\Delta\boldsymbol{L}$ 和铰链安装位置误差 $\Delta\boldsymbol{P}$ 相关。

根据确信可靠性理论, 产品性能裕量为其关键性能参数与阈值间的距离, 决定着产品的可靠程度。星载 SAR 天线展开机构装配过程中, 天线面板的平面度误差为关键性能参数, 其阈值由设计公差和装配工艺要求确定, 记为 ΔT_{th}。当天线面板的平面度误差高于 ΔT_{th} 时, 天线展开机构的装配精度较低, 无法完成精确的展开动作, 可能导致整个卫星无法工作, 此时, 判定天线展开机构装配不满足要求, 需重新进行装调。因此, 可以建立天线展开机构装配精度的性能裕量模型为

$$M_{T,\mathrm{P}}=\Delta T_{\mathrm{th}}-\Delta T \tag{8}$$

式中: $M_{T,\mathrm{P}}$ 为天线面板平面度误差的裕量。

2　不确定性分析与量化

根据星载 SAR 天线展开机构装配精度的性能裕量模型可知, 不确定性来源可分为: ①内部因素, 铰链配合间隙引起的杆件长度误差; ②外界条件, 人机环耦合因素造成的安装位置误差。

2.1　内部因素不确定性

在星载 SAR 天线展开机构装配过程中, 由于铰链定位精度存在误差, 使得杆件与铰链轴套形成间隙配合误差, 导致桁架杆系装配完成后杆件与铰链轴套间隙不一致, 发生部分接触, 对结构整体产生内应力。铰链间隙如图 4 所示。

在完成天线翻转过程后, 内应力释放会造成相邻杆件承受拉力或压力作用, 杆件长度因此发生变化, 进而导致天线面板平面度变化, 影响天线展开精度, 且展开机构中各铰链的间隙配合误差可视为相互独立的变量。因此, 可将铰链的间隙配合误差转化为杆件长度误差, 转化方法为[15]

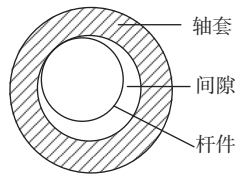


图 4　铰链间隙示意图

Fig. 4　Hinge clearance diagram

$$\Delta l_i=\Delta c_{i,1}+\Delta c_{i,2} \tag{9}$$

式中: $\Delta c_{i,1}$、$\Delta c_{i,2}$ 分别为第 i 个杆件两端铰链的间隙。一般将铰链间隙视为随机变量, 认为其服从均值为 $\mu_{\Delta c_i}$、标准差为 $\sigma_{\Delta c_i}$ 的正态分布, 可将其不确定性量化为

$$\Delta c_i\sim N\left(\mu_{\Delta c_i},\sigma_{\Delta c_i}^2\right) \tag{10}$$

因此, 在装配过程中, 杆件长度误差 Δl_i 也为随机变量, 认为其服从均值为 $\mu_{\Delta l_i}$、标准差为 $\sigma_{\Delta l_i}$ 的正态分布。可将杆长误差的不确定性量化为

$$\Delta l_i\sim N\left(\mu_{\Delta l_i},\sigma_{\Delta l_i}^2\right) \tag{11}$$

式中: $\mu_{\Delta l_i}=\mu_{\Delta c_{i,1}}+\mu_{\Delta c_{i,2}}$; $\sigma_{\Delta l_i}^2=\sigma_{\Delta c_{i,1}}^2+\sigma_{\Delta c_{i,2}}^2$。

2.2　外界条件不确定性

在星载 SAR 天线展开机构的装配过程中, 由于人因失误、设备精度等因素的影响, 导致测量铰链安装位置的过程中存在不确定性, 进而造成铰链安装位置误差的不确定性。

1) 测量误差中的设备因素, 主要反映安装铰链过程中激光跟踪仪与电子经纬仪设备精度造成的误差。由于天线装配过程是在恒温恒湿环境中进行, 本文不考虑环境扰动对铰链安装误差的影响。2 种设备误差的不确定性量化过程如下:

① 由于激光跟踪仪自身的精度具有随机性, 假定其造成的铰链安装位置误差服从正态分布, 且在空间 x_2 轴、y_2 轴、z_2 轴 3 个方向误差的不确定性相同, 因此, 下文均以 x_2 轴方向铰链的不确定性为例, 将其误差的不确定性量化为

$$\Delta p_{i_e1}\sim N\left(\mu_{\Delta p_{i_e1}},\sigma_{\Delta p_{i_e1}}^2\right) \tag{12}$$

式中: $\mu_{\Delta p_{i_e1}}$ 为由激光跟踪仪自身精度造成的铰链安装位置误差的均值; $\sigma_{\Delta p_{i_e1}}$ 为该误差的标准差。

② 电子经纬仪自身的测量精度具有随机性, 假设其造成的误差服从正态分布, 可将误差的不确定性量化为

$$\Delta p_{i_e2} \sim N\left(\mu_{\Delta p_{i_e2}}, \sigma^2_{\Delta p_{i_e2}}\right) \tag{13}$$

式中：$\mu_{\Delta p_{i_e2}}$为由电子经纬仪精度造成的铰链安装位置误差的均值；$\sigma_{\Delta p_{i_e2}}$为该误差的标准差。

2）在实际装配测量过程中，由于操作人员相关知识水平与操作熟练程度有限，导致作业过程中出现人因失误，造成铰链的安装位置误差。

① 在使用激光跟踪仪进行测量的过程中，由于反射器需要人工手动放置，操作人员摆放反射器时会出现一定误差，进而导致测量误差[16]。该误差本质上是由于人对事物的认知能力不足造成的，属于认知不确定性问题。因此，采用正态不确定分布对这类认知不确定性进行量化：

$$\Delta p_{i_r1} \sim N\left(\mu_{\Delta p_{i_r1}}, \sigma^2_{\Delta p_{i_r1}}\right) \tag{14}$$

式中：$\mu_{\Delta p_{i_r1}}$为激光跟踪仪测量过程中由于操作人员认知不确定性造成的铰链安装位置误差的均值；$\sigma_{\Delta p_{i_r1}}$为标准差，可以通过科学实践获取更多数据，从而降低认知不确定性。

② 在使用电子经纬仪测量目标位置时，由于相互瞄准、整平需要人工操作，并采用望远镜放大观测目标，而操作人员人眼对目标点存在视觉上的瞄准误差，导致测量出现误差[17]。同样，采用正态不确定分布对操作人员使用电子经纬仪测量过程中因认知不确定性导致的铰链安装位置误差进行量化：

$$\Delta p_{i_r2} \sim N\left(\mu_{\Delta p_{i_r2}}, \sigma^2_{\Delta p_{i_r2}}\right) \tag{15}$$

式中：$\mu_{\Delta p_{i_r2}}$为电子经纬仪测量过程中由于操作人员认知不确定性造成的铰链安装位置误差的均值；$\sigma_{\Delta p_{i_r2}}$为标准差。

综上，可得到间隙配合误差、测量设备误差和人因操作误差的不确定性量化结果。根据式(7)及不确定性的分析与量化结果可知，天线展开机构的平面度误差受杆长误差Δl_i、设备精度造成的铰链安装位置误差Δp_{i_e1}和Δp_{i_e2}、操作人员失误造成的铰链安装位置误差Δp_{i_r1}和Δp_{i_r2}影响，可表示为

$$\Delta T = f\left(\Delta l_i, \Delta p_{i_e1}, \Delta p_{i_e2}, \Delta p_{i_r1}, \Delta p_{i_r2}\right) \tag{16}$$

式中：f表示平面度误差ΔT与杆长误差、铰链安装位置误差的函数关系。

3 确信可靠性建模

确信可靠性理论[14]认为，在实际应用中，可以采用概率论、不确定理论[18]和机会理论[19]对产品可靠度进行度量。其中，概率论可用于描述随机（固有）不确定性；不确定理论用于描述认知（主观）不确定性；机会理论则可用于描述随机与认知混合不确定性。对于存在随机与认知混合不确定性的系统，其性能参数可称为不确定随机变量。

星载SAR天线展开机构的性能裕量m为不确定随机变量，则确信可靠度是系统的性能裕量大于0的机会，即

$$R_{\mathrm{B}} = \mathrm{Ch}\{\tilde{m} > 0\} \tag{17}$$

式中：$\mathrm{Ch}(\cdot)$为机会测度。

星载SAR天线展开机构的确信可靠度可计算为其性能裕量大于0的机会。由于考虑随机和认知混合不确定性的影响，采用机会理论对确信可靠性进行度量，即

$$R_{\mathrm{B}} = \mathrm{Ch}\left(\tilde{m}_{T,\mathrm{P}} > 0\right) = \mathrm{Ch}((\Delta T_{\mathrm{th}} - \Delta T) > 0) = 1 - \Phi(\Delta T_{\mathrm{th}} - \Delta T) \tag{18}$$

为量化SAR天线展开机构装配可靠的机会测度，需给出装配精度所服从的机会分布，Liu[20]给出了机会空间上的运算法则。

根据运算法则，由于装配精度ΔT与杆长误差和铰链安装位置误差有关，受到随机和认知混合不确定性的影响，ΔT为不确定随机变量。式(16)中的不确定性因素Δl_i、Δp_{i_e1}、Δp_{i_e2}为相互独立的随机变量，概率分布分别为Ψ_1、Ψ_2、Ψ_3；Δp_{i_r1}、Δp_{i_r2}为相互独立的不确定变量，不确定分布分别为Υ_1、Υ_2。若f为可测函数，则不确定随机变量ΔT的机会分布为

$$\Phi(\Delta T) = \int_{R^m} F(\Delta T_{\mathrm{th}}; b_1, b_2, b_3)\mathrm{d}\Psi_1(b_1)\mathrm{d}\Psi_2(b_2)\mathrm{d}\Psi_3(b_3) \tag{19}$$

式中：$F(\Delta T_{\mathrm{th}}; b_1, b_2, b_3) = M\{f(b_1, b_2, b_3, \tau_1, \tau_2) \leqslant \Delta T_{\mathrm{th}}\}$表示随机变量取值固定为$b_1$、$b_2$、$b_3$时不确定变量$f(b_1, b_2, b_3, \tau_1, \tau_2)$通过$\Upsilon_1$、$\Upsilon_2$决定的不确定分布，$\tau_1$、$\tau_2$为相互独立的不确定变量。

在该定理中，若$f(b_1, b_2, b_3, \tau_1, \tau_2)$为连续函数，且对于$\tau_1$、$\tau_2$是单调递增的，则机会分布$\Phi(\Delta T)$中的$F(\Delta T_{\mathrm{th}}; b_1, b_2, b_3)$是方程(20)的根：

$$f\left(b_1, b_2, b_3, \Upsilon_1^{-1}(\alpha), \Upsilon_2^{-1}(\alpha)\right) = \Delta T_{\mathrm{th}} \tag{20}$$

式中：$\Upsilon_1^{-1}(\alpha)$、$\Upsilon_2^{-1}(\alpha)$分别为Δp_{i_r1}、Δp_{i_r2}的不确定逆分布，表达为

$$\Upsilon_1^{-1}(\alpha) = \mu_{\Delta p_{i_r1}} + \frac{\sigma_{\Delta p_{i_r1}}\sqrt{3}}{\pi}\ln\left(\frac{\alpha}{1-\alpha}\right) \tag{21}$$

$$\Upsilon_2^{-1}(\alpha) = \mu_{\Delta p_{i_r2}} + \frac{\sigma_{\Delta p_{i_r2}}\sqrt{3}}{\pi}\ln\left(\frac{\alpha}{1-\alpha}\right) \tag{22}$$

其中：α为信度。

但在建立的装配精度性能裕量模型中，不确定变量和随机变量的数量较多，因此，很难利用解析方法求解。为此，采用文献[21]中提出的不确定-

随机仿真方法计算星载 SAR 天线装配过程的确信可靠度。不确定-随机仿真方法主要分为内层不确定仿真与外层随机仿真。外层随机仿真采用蒙特卡罗仿真方法求取具有随机不确定性参数的均值, 即 y_1、y_2、y_3; 得到均值后进行内层不确定仿真, 仿真原理为: 利用不确定逆分布和信度的关系 (即式 (21)、式 (22)) 对由外层随机仿真得到的随机不确定性参数均值和不确定变量的不确定逆分布所组成的相关方程求解, 得到式 (20) 的根 α, 具体流程如下:

1) 根据 Δl_i、Δp_{i_e1}、Δp_{i_e2} 的概率分布 Ψ_1、Ψ_2、Ψ_3 随机生成一组样本 $\boldsymbol{v}_k=\left(\Delta l_i^{(k)},\Delta p_{i_e1}^{(k)},\Delta p_{i_e2}^{(k)}\right)$。

2) 求解方程 $f\left(\Delta l_i^{(k)},\Delta p_{i_e1}^{(k)},\Delta p_{i_e2}^{(k)},\Upsilon_1^{-1}(\alpha),\Upsilon_2^{-1}(\alpha)\right)=\Delta T_{\mathrm{th}}$ 的根 α, 并置 $C_k=\alpha$。

3) 重复步骤 1 和步骤 2 共 N 次, 得到 $C_1,C_2,\cdots,C_N$。

4) 输出 $R_{\mathrm{B}}=\sum_{k=1}^{N}C_k/N$。

4　案　例

4.1　分析对象

使用某型号星载 SAR 天线作为案例研究对象, 其主要尺寸参数如表 1 所示, 装配工艺要求规定星载 SAR 天线展开后的平面度阈值 ΔT_{th} 为 0.2 mm。

表 1　某型号星载 SAR 天线尺寸参数

Table 1　Dimension parameters of an spaceborne SAR antenna of a certain model　mm

参数	数值
A_1A_2	1 653.73
A_1A_3	2 162.35
A_3A_5	1 802.87
OC	2 147.51
OA_1	2 335.99
OA_3	883.80
OA_5	2 007.84
O_1O_2	901.89

对于间隙配合误差, 根据文献 [15] 提供的铰链安装间隙试验数据, 确定星载 SAR 天线铰链安装间隙均值 $\mu_{\Delta c_i}$=0.12 mm, 标准差 $\sigma_{\Delta c_i}$=0.04 mm, 因此, 根据式 (9) 可计算得到杆长误差 $\mu_{\Delta l_i}$=0.24 mm, $\sigma_{\Delta l_i}$=0.08 mm。

对于测量设备误差, 天线装配过程中使用 Leica AT960 型激光跟踪仪与 T3000 型电子经纬仪测量铰链位置。激光跟踪仪标称精度为 0.015 mm[22], 即 $\mu_{\Delta p_{i_e1}}$=0 mm, $\sigma_{\Delta p_{i_e1}}$=0.015 mm。电子经纬仪的测量精度会随测距的增加而减小, T3000 型电子经纬仪精度为 0.05 mm[23], 即 $\mu_{\Delta p_{i_e2}}$=0 mm, $\sigma_{\Delta p_{i_e2}}$=0.05 mm。

装配操作人员在测量过程中往往会造成人因失误, 导致铰链安装位置出现误差, 通过统计分析历史测量数据可以得到操作人员在使用激光跟踪仪与电子经纬仪时所造成的测量误差, 如表 2 所示。

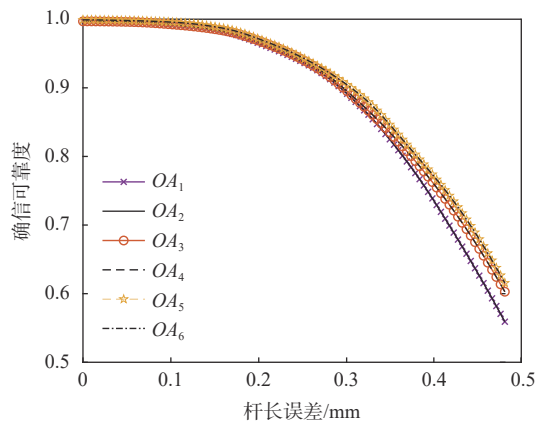
表 2　操作人员造成的测量误差

Table 2　Measurement errors caused by operators

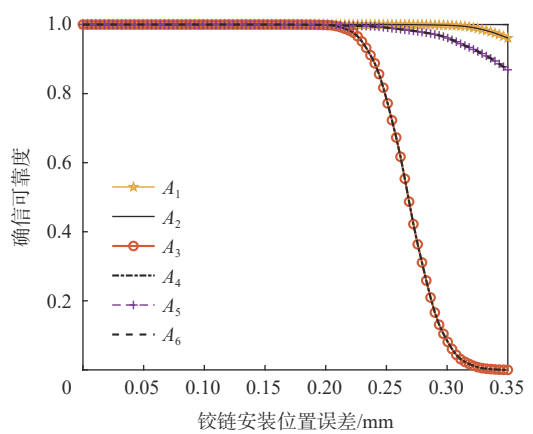
$\mu_{\Delta p_{i_r1}}$/mm	$\mu_{\Delta p_{i_r2}}$/mm	$\sigma^2_{\Delta p_{i_r1}}$/mm	$\sigma^2_{\Delta p_{i_r2}}$/mm
0	0	0.022^2	0.030^2

4.2　确信可靠性评估

由于星载 SAR 天线在装配过程中性能退化趋势并不明显, 不考虑星载 SAR 天线的确信可靠度随时间的变化关系。在确信可靠度评估中, 根据不确定性参数估计结果和确信可靠度方法流程, 分别计算确信可靠度随各杆长误差和铰链安装位置误差的变化关系, 得到确信可靠度评估结果, 如图 5 所示。可以看出, 由于杆件 OA_1 与 OA_2、OA_3 与 OA_4、OA_5 与 OA_6, 铰链 A_1 与 A_2、A_3 与 A_4、A_5 与 A_6 完



(a) 确信可靠度随杆长误差变化曲线



(b) 确信可靠度随铰链安装位置误差变化曲线

图 5　确信可靠度随杆长误差和铰链安装误差的变化曲线

Fig. 5　Curves of belief reliability varies with rod length error and hinge installation position error

全对称，其对装配精度的影响相同，且杆件 OA_1 与 OA_2、铰链 A_3 与 A_4 对天线展开机构的可靠性影响较大。

4.3 敏感度分析

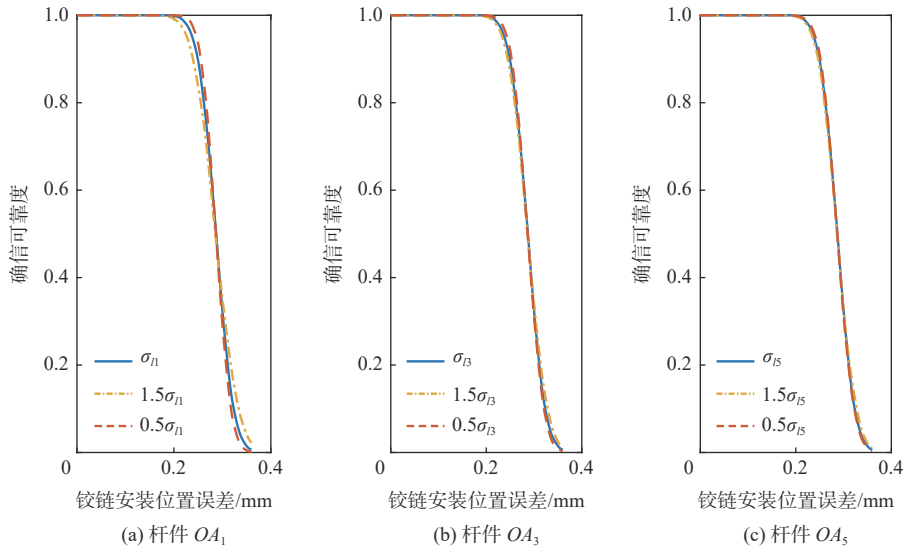
对杆长误差、铰链安装位置误差等影响装配精度的零部件参数进行敏感度分析，比较不同零部件参数对装配精度的影响大小，确定敏感度较大的零部件参数，在天线展开机构装配过程中进行重点控制。

由于杆件 OA_1 与 OA_2、OA_3 与 OA_4、OA_5 与 OA_6 完全对称，选择杆件 OA_1、OA_3、OA_5 进行敏感度分析。首先，调整各杆长误差的标准差，分别绘制0.5倍、1倍、1.5倍标准差的确信可靠度变化曲线，如图6所示。

与杆件相同，铰链 A_1 与 A_2、A_3 与 A_4、A_5 与 A_6 也完全对称，因此，选择铰链 A_1、A_3、A_5 进行敏感度分析，将其标准差分别调整为0.5倍、1倍、1.5倍，绘制确信可靠度变化曲线，如图7所示。

为进一步分析各参数的敏感度，分别选取铰链安装位置误差与杆长误差为0.25 mm的确信可靠度作为基准，保证其他参数不变的情况下，针对待分析的参数分别计算0.5倍参数标准差、1倍参数标准差和1.5倍参数标准差时铰链安装位置误差与杆长误差为0.25 mm的确信可靠度，并将计算结果除以所有参数为1倍标准差下的确信可靠度，得到参数敏感度对比结果如图8所示。

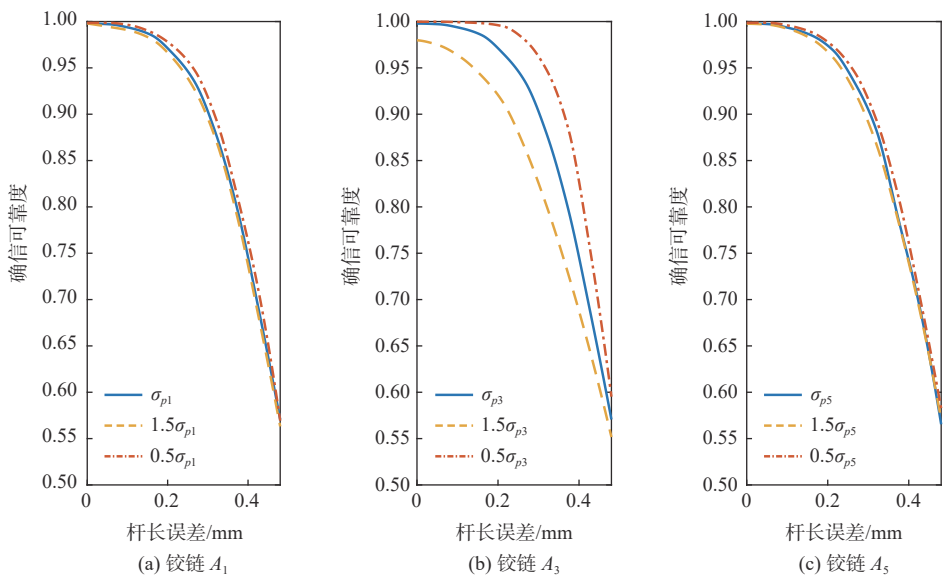
从图8中可以看出，杆件 OA_1 和铰链 A_3 对星载SAR天线装配可靠度影响最大。由于杆件 OA_1 和



(a) 杆件 OA_1　(b) 杆件 OA_3　(c) 杆件 OA_5

图6　不同杆长误差随铰链安装位置误差变化的确信可靠度曲线

Fig. 6　Belief reliability curves of manufacturing errors of different rod parts changing with errors of hinge installation position



(a) 铰链 A_1　(b) 铰链 A_3　(c) 铰链 A_5

图7　不同铰链安装位置误差随杆长误差变化的确信可靠度曲线

Fig. 7　Belief reliability curves of different hinge installation position errors changing with rod length error

OA_2、铰链 A_3 和 A_4 完全对称, 杆件 OA_1 和 OA_2、铰链 A_3 和 A_4 对星载 SAR 天线装配可靠度影响最大, 即敏感度较高, 在装调过程中需要重点控制。图 9 为利用仿真分析方法绘制的杆件 OA_1 长度误差-铰链 A_3 安装位置误差-确信可靠度变化曲面。从三维曲面中可以看出, 当铰链 A_3 安装位置误差控制在 0.1 mm 内、杆件 OA_1 两侧的铰链间隙控制在 0.3 mm 内时, 星载 SAR 装配过程的确信可靠度能够保持在 0.8 以上。

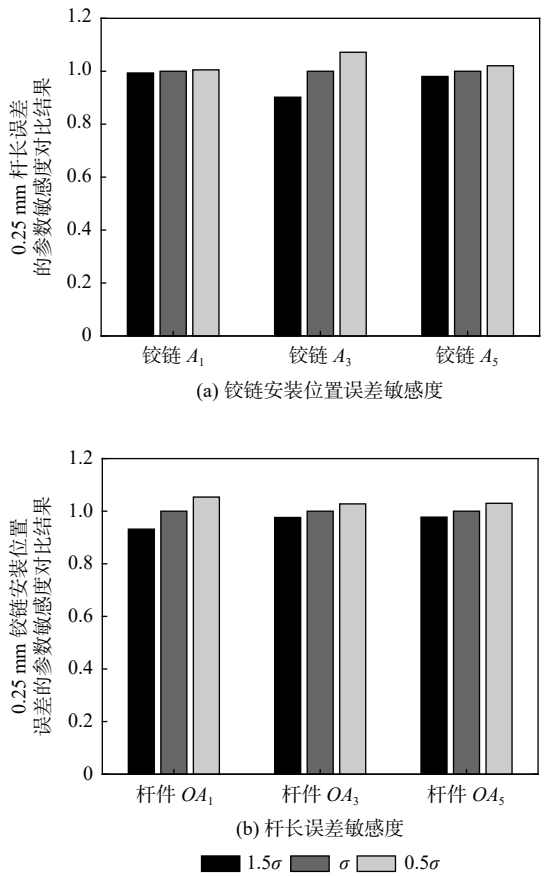


图 8　参数敏感度分析

Fig. 8　Parameter sensitivity analysis

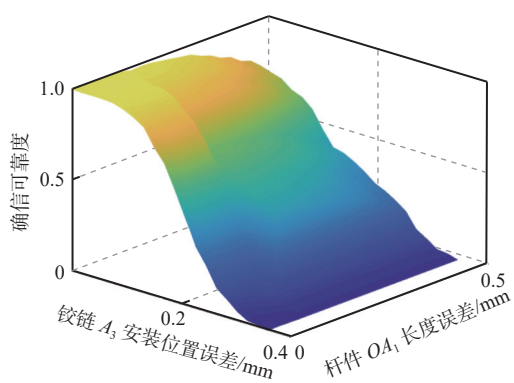


图 9　杆件 OA_1 长度误差-铰链 A_3 安装位置误差-确信可靠度变化曲面

Fig. 9　Variation surface of rod OA_1 length error - hinge A_3 installation position error - belief reliability

综上, 可确定敏感度较高的零部件为杆件 OA_1、OA_2 和铰链 A_3、A_4。因此, 在实际装调过程中, 通过控制杆件 OA_1、OA_2 两边的铰链安装间隙及铰链 A_3、A_4 安装位置误差(如在铰链 A_3、A_4 安装过程中提高测量仪器精度), 可有效提高 SAR 天线装配的可靠性, 为装调过程提供理论指导, 解决目前天线装配过程中传统方法反复装调难以达到展开精度要求的问题。

5　结　论

基于可靠性科学原理和确信可靠性理论, 对星载 SAR 天线装配可靠性开展研究, 提出了星载 SAR 天线装配确信可靠性建模与分析方法, 并以具体型号产品为例验证了方法的有效性。

1) 基于可靠性科学原理和确信可靠性理论, 构建星载 SAR 天线装配精度性能裕量模型, 并对装配精度性能裕量模型中的多源不确定性进行分析和量化, 开展了星载 SAR 天线装配精度确信可靠性建模。

2) 在星载 SAR 天线装配精度裕量建模中, 考虑了装配过程中内部与外在因素对其精度的影响, 包括杆长误差、铰链安装位置误差等, 基于几何原理构建了考虑多种影响因素的星载 SAR 天线装配精度模型, 并考虑了装配精度与其阈值之间的距离, 构建了星载 SAR 天线装配精度性能裕量模型。

3) 通过星载 SAR 天线的具体案例验证了本文方法的可行性。结果表明, 本文方法可以为提高天线装配精度的可靠性提供指导, 在敏感度较高的铰链安装过程中, 提高测量仪器精度、控制铰链安装间隙等有助于提高 SAR 天线装配的可靠性。

星载 SAR 天线装配精度可靠性建模方法能够从理论上辨识装配过程中加工误差、装配误差、测量误差及不确定性对可靠性的影响, 并指导 SAR 天线展开机构的装调过程, 可以为其他空间机构的装配精度可靠性研究提供有效的理论方法。

参考文献 (References)

[1] MOREIRA A, PRATS-IRAOLA P, YOUNIS M, et al. A tutorial on synthetic aperture radar[J]. IEEE Geoscience and Remote Sensing Magazine, 2013, 1(1): 6-43.

[2] WANG R, LOFFELD O, NIES H, et al. Focus FMCW SAR data using the wavenumber domain algorithm[J]. IEEE Transactions on Geoscience and Remote Sensing, 2010, 48(4): 2109-2118.

[3] 余德文, 赵强强, 陈飞飞, 等. 利用结构等效转换的空间可展机构装配误差建模与灵敏度分析[J]. 西安交通大学学报, 2019, 53(4): 15-23.
YU D W, ZHAO Q Q, CHEN F F, et al. Assembly error modeling and sensitivity analysis for spatial deployable mechanisms with

structural equivalent transformation[J]. Journal of Xi'an Jiaotong University, 2019, 53(4): 15-23(in Chinese).

[4] 苏春, 黄漪. 装配误差传递建模及其精度可靠性评估[J]. 中国机械工程, 2017, 28(19): 2359-2364.
SU C, HUANG Y. Modeling for assembly error propagations and assessments of assembly precision reliability[J]. China Mechanical Engineering, 2017, 28(19): 2359-2364(in Chinese).

[5] SZEKELY G S, TEICHERT W H, BRENNER C E, et al. Practical procedures for reliability estimation of spacecraft structures and their components[J]. AIAA Journal, 1998, 36(8): 1509-1515.

[6] 张建国, 陈建军, 段宝岩, 等. 基于非概率模型的星载天线展开机构可靠性分析[J]. 西安电子科技大学学报, 2006, 33(5): 739-744.
ZHANG J G, CHEN J J, DUAN B Y, et al. Reliability analysis of the deployment mechanism of a large satellite antenna based on the non-probabilistic model[J]. Journal of Xidian University, 2006, 33(5): 739-744(in Chinese) .

[7] 刘良. 一种空间折展机构的设计与分析[D]. 南宁: 广西大学, 2020: 1-9.
LIU L. Design and analysis of a space deployable and foldable mechanism[D]. Nanning: Guangxi University, 2020: 1-9(in Chinese) .

[8] 吴建云, 王春洁, 汪瀚. 基于蒙特卡洛法的卫星天线板展开精度分析[J]. 航天返回与遥感, 2013, 34(6): 89-94.
WU J Y, WANG C J, WANG H. Accuracy analysis of satellite antenna plate deployment based on Monte Carlo method[J]. Spacecraft Recovery & Remote Sensing, 2013, 34(6): 89-94(in Chinese).

[9] 杨玉龙, 关富玲. 可展桁架天线形面精度理论分析[J]. 空间科学学报, 2009, 29(5): 529-533.
YANG Y L, GUAN F L. Theoretical analysis of surface error for deployable truss antenna[J]. Chinese Journal of Space Science, 2009, 29(5): 529-533(in Chinese).

[10] 丁建中, 王春洁. 含铰链间隙板式卫星天线展开精度分析[J]. 北京航空航天大学学报, 2016, 42(12): 2625-2631.
DING J Z, WANG C J. Deployment accuracy analysis of planar satellite antenna with joint clearances[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(12): 2625-2631(in Chinese).

[11] GREENE W H. Effects of random member length errors on the accuracy and internal loads of truss antennas[J]. Journal of Spacecraft & Rockets, 2015, 22(5): 554-559.

[12] MOBREM M. Methods of analyzing surface accuracy of large antenna structures due to manufacturing tolerances: AIAA 2023-1453[R]. Reston: AIAA, 2003.

[13] HEDGEPETH J M. Influence of fabrication tolerances on the surface accuracy of largeantenna structures[J]. AIAA Journal, 1982, 20(5): 680-686.

[14] KANG R. Belief reliability theory and methodology[M]. Berlin: Springer, 2021: 60-71.

[15] 孙中超, 喻天翔, 崔卫民, 等. 铰链间隙对连杆机构运动精度的重要性测度分析[J]. 中国机械工程, 2014, 25(21): 2874-2879.
SUN Z C, YU T X, CUI W M, et al. Importance measure analysis of revolute joint clearance about motion accuracy of linkage mechanism[J]. China Mechanical Engineering, 2014, 25(21): 2874-2879 (in Chinese).

[16] NGUYEN T T, AMTHOR A, AMENT C. High precision laser tracker system for contactless position measurement[C]//Proceedings of the IEEE International Conference on Control System, Computing and Engineering. Piscataway: IEEE Press, 2012: 97-102.

[17] 刘建新. 电子经纬仪测角原理及其在航天器检测中的应用[J]. 航天器环境工程, 2007, 24(2): 116-120 .
LIU J X. Measuring principle of electronic theodolite and its applications in spacecraft inspection[J]. Spacecraft Environment Engineering, 2007, 24(2): 116-120(in Chinese) .

[18] LIU B D. Uncertainty theory[M]. Berlin: Springer, 2010: 387-394.

[19] LIU Y H. Uncertain random variables: A mixture of uncertainty and randomness[J]. Soft Computing, 2013, 17(4): 625-634.

[20] LIU Y H. Uncertain random programming with applications[J]. Fuzzy Optimization and Decision Making, 2013, 12(2): 153-169.

[21] 张清源. 不确定随机系统确信可靠性度量与分析[D]. 北京: 北京航空航天大学, 2020: 77-83.
ZHANG Q Y. Certitude reliability measurement and analysis for uncertain stochastic systems[D]. Beijing: Beihang University, 2020: 77-83(in Chinese) .

[22] 孙海丽, 姚连璧, 周跃寅, 等. 激光跟踪仪测量精度分析[J]. 大地测量与地球动力学, 2015, 35(1): 177-181.
SUN H L, YAO L B, ZHOU Y Y, et al. Analysis of measurement accuracy of laser tracker[J]. Journal of Geodesy and Geodynamics, 2015, 35(1): 177-181(in Chinese).

[23] 钱关旻, 袁榕, 袁祥伟, 等. 电子经纬仪一测回竖直角测角标准偏差测量结果不确定度评定[J]. 中国计量, 2019(8): 108-109.
QIAN G M, YUAN R, YUAN X W, et al. Uncertainty in measurement results of standard deviation of angle measurement in vertical angle by electronic theodolite[J]. China Metrology, 2019(8): 108-109 (in Chinese) .

Belief reliability modeling for assembly accuracy of spaceborne SAR antenna deployable mechanism

LI Yun[1]，ZHANG Xiangxiao[2]，SUN Fuqiang[2,*]，LI Xiaoyang[2]，HAN Jianchao[1]

(1. Beijing Spacecraft Manufacturing Co.，Ltd.，Beijing 100094，China；
2. School of Reliability and Systems Engineering，Beihang University，Beijing 100191，China)

Abstract: The deployable mechanism for a spaceborne synthetic aperture radar (SAR) antenna is a space truss structure, with the rod system and hinge coupled in space. Assembling accuracy is difficult to control precisely during the assembly process, and assembly reliability directly impacts the spaceborne SAR antenna's service performance. Therefore, based on the principle of reliability science, a precision assurance reliability modeling method for spaceborne SAR antenna ground assembly is proposed in this paper. First, the performance index and threshold indicating the assembly accuracy of the antenna deployable mechanism were established, taking into account both internal and external elements that have an impact on the assembly accuracy in the antenna assembly process. Then, the multi-source uncertainty analysis and quantification were carried out to construct the assembly precision certainty reliability model of the spaceborne SAR antenna deployable mechanism. Finally, the reliability sensitivity analysis and verification of a spaceborne SAR antenna deployable mechanism are carried out. The results show that the proposed method can provide theoretical guidance for the mounting and adjusting of spaceborne SAR antenna deployable mechanism, and effectively improve the assembly accuracy and efficiency.

Keywords: belief reliability; antenna deployable mechanism; assembly accuracy; performance margin; uncertainty

Received：2022-04-12；**Accepted**：2022-05-06；**Published Online**：2022-05-23 15：54

URL：link.cnki.net/urlid/11.2625.V.20220523.1000.001

Foundation item：Stable Supporting Fund of Science & Technology on Reliability & Environmental Engineering Key Laboratory (WDZC20220101)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0201

基于动态视觉传感器的无人机目标检测与避障

蔡志浩，陈文军，赵江，王英勋*
(北京航空航天大学 自动化科学与电气工程学院，北京 100191)

摘　　要：针对无人机在动态环境中感知动态目标与躲避高速动态障碍物，提出了基于动态视觉传感器的目标检测与避障算法。设计了滤波方法和运动补偿算法，滤除事件流中背景噪声、热点噪声及由相机自身运动产生的冗余事件；设计了一种融合事件图像和RGB图像的动态目标融合检测算法，保证检测的可靠性。根据检测结果对目标运动轨迹进行估计，结合障碍物运动特点和无人机动力学约束改进速度障碍法躲避动态障碍物。大量仿真试验、手持试验及飞行试验验证了所提算法的可行性。

关 键 词：事件相机；事件滤波；运动补偿；融合检测；速度障碍法
中图分类号：V249.3; TP249
文献标志码：A　　**文章编号**：1001-5965(2024)01-0144-10

无人机因其小巧的结构和多变的操作在复杂场景及困难任务中有着广泛的应用。但是，在动态环境中，感知高速运动物体、躲避高速动态障碍物仍然是需要攻破的难题。由于无人机承重能力和能源供应能力有限，携带如雷达、声呐、激光等大质量、高功耗的传感器会大大减少飞机飞行时间，以相机为主要传感器的感知方法已经成为一种主要的发展趋势[1]。

当前，主流相机感知方法还存在许多限制。例如，RGB相机是一种常见的基于帧的相机，在场景光照不足及相机高速运动的情况下，会产生运动模糊。Falanga等[2]指出在躲避相对速度较大的障碍物时，需减小感知延迟，表明所采用的传感器具有更高的时间分辨率。

事件相机是一种新型的仿生视觉传感器[3]，如动态视觉传感器(dynamic vision sensor, DVS)。事件相机与RGB相机不同，其输出像素级亮度变化，不是标准的强度帧。每个事件$\boldsymbol{e}_i$表示为：$\boldsymbol{e}_i=[x,y,t,p]$，$x$和$y$分别为触发的事件在图像坐标系的位置，$t$为事件触发时刻，$p$为事件的极性[4]。事件相机输出的是以微秒时间分辨率的事件流$\boldsymbol{E}_{\text{stream}}$[5]，相当于以每秒数千帧的速度运行的传统视觉传感器，但其数据量小得多。由于具有局部处理功能，事件相机整体功率、数据存储和计算性能要求也大大降低，并且动态范围增加了数个数量级。事件相机不受运动模糊的影响，且具有非常高的时间动态范围，使得其能在高速运动期间或以高动态范围为特征的场景中提供可靠的视觉信息。

但是，事件相机通常存在感光阈值造成的噪声、背景噪声[6]、热点噪声[7]。需要根据噪声特点进行滤波，以便后续事件的处理。单个事件无任何意义，需要对事件进行数据关联，可通过事件可视化进行。事件的可视化方法常用的有累积固定时间窗口[8]、累积固定事件数量[8]和基于生命周期[9]的方法。基于生命周期[9]的方法可有效抑制加速梯度引起的运动模糊，但计算量较大，不便于部署在无人机机载设备上。

运动补偿[10]指在图像中不仅保留事件的位置，还构建运动模型，假设事件由场景中移动的物体边缘触发产生，通过时间信息对事件进行2D变换，使

收稿日期：2022-04-01；录用日期：2022-05-06；网络出版时间：2022-05-26 14:44
网络出版地址：link.cnki.net/urlid/11.2625.V.20220526.0927.001
*通信作者. E-mail：wangyx@buaa.edu.cn

引用格式：蔡志浩，陈文军，赵江，等. 基于动态视觉传感器的无人机目标检测与避障[J]. 北京航空航天大学学报，2024，50(1)：144-153. CAI Z H，CHEN W J，ZHAO J，et al. Object detection and obstacle avoidance based on dynamic vision sensor for UAV[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(1)：144-153 (in Chinese).

图像的对比度达到最大，从而得到边缘锐化的图像。文献 [10] 使用光流的方法获取环境中深度信息，对事件图像进行运动补偿，文献 [11] 通过深度相机直接获取深度信息进行运动补偿，减小了计算量。

文献 [2] 指出，机器人在高动态环境中飞行时，其与障碍物之间的相对速度往往很大，相对速度越大，感知延迟（指感知环境并处理捕获的数据生成控制指令所需的时间）的作用就越关键。Falanga 等[12] 在感知延迟问题的基础上，使用改进的人工势场法计算飞机每个事件步长的排斥力，躲避了相对速度约为 10 m/s 的动态障碍物。Sanket 等[13] 提出了一种基于事件相机与深度学习的动态避障网络 EV-DodgeNet。文献 [11] 通过预测动态障碍物的运动轨迹躲避了 7～12 m/s 的动态障碍物。

当前，国内外学者针对动态视觉传感器的目标检测与避障算法做出了一定的探索，但还存在改进空间：①在算法能够在无人机机载设备上实时运行的前提下，进一步提高目标检测的稳定性[11-13]。②简化无人机避障控制，减少无人机避障控制指令[12]。为改进以上问题，本文研究了基于动态视觉传感器的无人机目标检测与避障算法，主要创新点如下：

1）设计了一种针对事件流的事件滤波方法，可以滤除背景噪声和热点噪声，并保持事件的异步性。

2）提出了一种融合事件图像和 RGB 图像的动态目标融合检测算法，在高动态场景中进行了仅依靠 RGB 相机和事件相机检测试验对比，体现了融合检测算法的鲁棒性。

3）根据无人机动力学模型和视场角约束改进了速度障碍（velocity obstacle, VO）法，求解出动态场景下的最优期望速度，并对比了文献 [12] 采用的人工势场法，在高速度动态障碍物时，该方法具有较稳定的效果。

4）搭载飞机平台进行实际飞行验证，成功躲避高速运动障碍物，验证了算法的可行性与在飞机平台上的实时性。

1　算法整体框架

本文主要针对高动态场景中，动态障碍物对四旋翼无人机飞行安全造成威胁的情况，躲避动态障碍物，即实现图 1 所示的躲避高速运动障碍物的场景。

为实现图 1 所示的任务场景，设计了图 2 所示的算法框架，主要包括事件预处理、高动态场景目标检测和动态避障 3 部分。

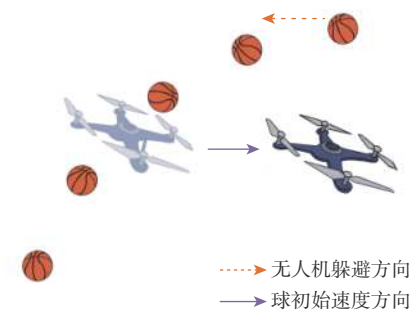


图 1　任务场景

Fig. 1　Task scenario

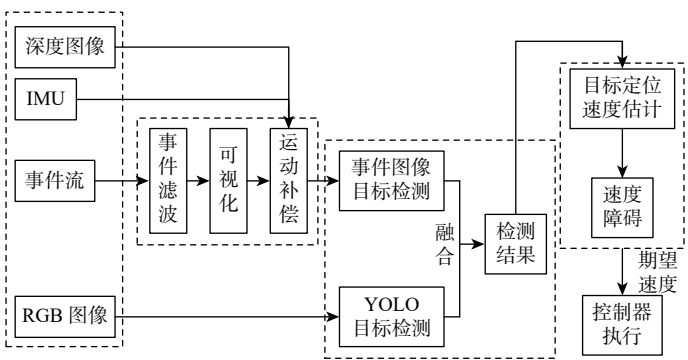


图 2　本文算法框架

Fig. 2　Framework of proposed algorithm

1）事件预处理。本文设计的滤波方法对事件流进行处理，滤除事件的背景噪声和热点噪声。采用事件数量和固定时间双阈值的形式可视化事件，生成事件图像。根据 IMU、深度信息及无人机的里程计信息进行运动补偿，优化事件图像。

2）高动态场景目标检测。对事件图像进行动态目标分割，获取检测结果，RGB 相机的图像使用 YOLO-tiny 获取检测结果，通过置信度进行融合，获得最终检测结果。

3）动态避障。为提高障碍物位置估计的精度，融合检测结果和深度图像获得深度信息，并通过最小围椭球体体积（minimum volume enclosing ellipsoid, MVEE）将障碍物拟合为椭球，获得相机坐标系下的位置。通过 Kalman 滤波估计障碍物运动速度。根

据任务场景特点改进速度障碍法，在优化问题中，考虑了无人机的视场角约束和动力学模型，获取期望速度躲避动态障碍物。

2 事件流滤波及可视化算法

2.1 事件滤波

在一定的时空内，背景噪声事件数量小于某一阈值的概率较高，由环境中物体运动或光强变化引起的真实事件数大于该阈值的概率较高。一个事件是真实事件或噪声事件的概率需要一段事件的事件流来计算。对于待处理的事件$\boldsymbol{e}_0(x_0,y_0,t_0,p_0)$，使用二维高斯核：

$$G(d,t)=\frac{1}{2\pi\sigma_1\sigma_2}\mathrm{e}^{-\frac{1}{2}\left[\frac{(\Delta d_i)^2}{\sigma_1^2}+\frac{(\Delta t_i)^2}{\sigma_2^2}\right]} \tag{1}$$

式中：$\Delta d_i=\sqrt{(x_i-x_0)^2+(y_i-y_0)^2}$为该事件与其他事件的空间距离；$\Delta t_i=|t_i-t_0|$为该事件与其他事件的时间戳距离。

将事件流用二维高斯核进行卷积，得到该事件的真实事件概率：

$$P_{e_0}=\boldsymbol{E}_{\mathrm{stream}}(d,t)\otimes\boldsymbol{G}(d,t) \tag{2}$$

$$P_{e_0}=\sum_{i=1}^{N}\frac{1}{2\pi\sigma_1\sigma_2}\mathrm{e}^{-\frac{1}{2}\left[\frac{(\Delta d_i)^2}{\sigma_1^2}+\frac{(\Delta t_i)^2}{\sigma_2^2}\right]} \tag{3}$$

式中：“⊗”表示卷积运算；σ_1和σ_2分别为事件间的空间距离Δd和时间距离Δt的标准差，与事件相机分辨率和事件流的时间间隔相关。

如式(4)所示，当事件概率小于阈值λ时，认为该事件为背景事件；当事件概率大于或等于阈值λ时，认为该事件为真实事件：

$$e_i=\begin{cases}\{x_i,y_i,t_i,p_i\} & P_{e_i}\geqslant\lambda\\ \mathrm{null} & P_{e_i}<\lambda\end{cases} \tag{4}$$

2.2 事件图像可视化算法

本文使用事件数据数量及时间阈值双限制的方式进行累计得到事件图像。如式(5)所示，产生的事件图像既可以保证高时间分辨率，又可以使事件中具有足够的信息以进一步处理。

$$T_{i,j}=\begin{cases}\displaystyle\int_{t_{\mathrm{start}}}^{t_{\mathrm{end}}}E_{\mathrm{stream}}(t_{\mathrm{cur}})\mathrm{d}t_{\mathrm{cur}} & \displaystyle\sum_{\Delta t}(e_i)<N\\ \displaystyle\sum_{i=0}^{N}e_i & \displaystyle\sum_{\Delta t}(e_i)\geqslant N\end{cases} \tag{5}$$

式中：t_{start}为开始时间；t_{cur}为当前时间；t_{end}为结束时间。

2.3 运动补偿

将从t_0时刻开始某个很短时间间隔Δt的事件存储为事件流$\boldsymbol{E}_{t_0}$，利用IMU数据计算Δt内，世界坐标系下的平均角速度$\bar{\omega}$和相机坐标系到世界坐标系的旋转矩阵${}^{W}\boldsymbol{R}_C$。

旋转补偿的主要目的是滤除由相机自身旋转所产生的事件。根据获得的平均角速度$\bar{\omega}$和罗德里格斯旋转公式：

$$\boldsymbol{R}_e=\cos\theta\boldsymbol{I}+(1-\cos\theta)\boldsymbol{n}\boldsymbol{n}^{\mathrm{T}}+\sin\theta\boldsymbol{n}^{\wedge} \tag{6}$$

式中：“$^{\wedge}$”为向量到反对称矩阵的转换符；$\boldsymbol{R}_e$为需求得的事件相机旋转矩阵；$\boldsymbol{n}$为一个单位长度的旋转轴；旋转角度$\theta=\bar{\omega}(t-t_0)$，$t$为当前事件的时间戳。

由于事件数量较多，选择每一毫秒计算一次相机旋转矩阵。针对该时间段的事件进行旋转补偿：$(x,y,t,p)\rightarrow(x_{\mathrm{r}},y_{\mathrm{r}},t_0,p)$，即保持事件的极性$p$不变，将像素坐标$(x,y)$转换为旋转后的像素坐标$(x_{\mathrm{r}},y_{\mathrm{r}})$，并且将时间戳归一化到$t_0$时刻。运动补偿后的事件仍然是在像素坐标系下，要将事件通过事件相机的内参矩阵转换到相机坐标系下，操作如下：

$$z_c\begin{bmatrix}x\\y\\1\end{bmatrix}=\begin{bmatrix}f&0&0&0\\0&f&0&0\\0&0&1&0\end{bmatrix}\begin{bmatrix}x_c\\y_c\\z_c\\1\end{bmatrix} \tag{7}$$

式中：f为相机焦距。

通过相机坐标系和世界坐标系的平移矩阵${}^{W}\boldsymbol{T}_E$，将事件由相机坐标系$\boldsymbol{X}_E$转换到世界坐标系下，再乘以平移补偿矩阵，将事件重新转移到相机坐标系下，完成平移补偿。平移补偿过程表示为

$$\boldsymbol{X}'_E={}^{W}\boldsymbol{T}_E^{-1}\boldsymbol{T}_{t-t_0}{}^{W}\boldsymbol{T}_E\boldsymbol{X}_E \tag{8}$$

式中：$\boldsymbol{X}'_E$为平移运动补偿后事件在相机坐标系下的坐标；${}^{W}\boldsymbol{T}_E^{-1}$为由世界坐标系转换到相机坐标系的平移矩阵；$\boldsymbol{T}_{t-t_0}$为平移补偿矩阵，表示从$t_0$时刻开始到$t$时间段内，相机在世界坐标系下的平移。

$$\boldsymbol{T}_{t-t_0}=\begin{bmatrix}\boldsymbol{I}&\boldsymbol{v}(t-t_0)\\\boldsymbol{0}^{\mathrm{T}}&1\end{bmatrix} \tag{9}$$

式中：$\boldsymbol{v}$为无人机里程计获得的线速度信息。

3 融合事件图像与RGB图像的动态目标检测

3.1 基于事件图像的检测

运动补偿后，需要对事件图像进行处理，通过均值滤波去除椒盐噪声。利用动态障碍物具有更高的平均时间戳的特性，检测动态障碍物。通过形态学操作将事件分成不同对象。每个像素点$T_{i,j}$与评价分数$\rho(i,j)\in[-1,1]$相关联，可定量表示每个动态障碍物与背景的区别，用评价分数ρ区分动态障碍物O_{k}和背景B_{g}。

$$\rho(i,j)=\left(T_{i,j}-M\right)/\Delta t \tag{10}$$

$$B_{\mathrm{g}}=\{i,j|\rho(i,j)\leqslant\lambda\}$$
$$O=\{i,j|\rho(i,j)>\lambda\} \tag{11}$$

式中：M为该事件图像的平均时间戳；$O=O_1\cup\cdots O_n$为动态障碍物集合，n为独立动态障碍物的个数；λ为预定义的每个动态障碍物的最小置信值。为检测独立的动态障碍物，利用简单的形态学操作（碰撞）对事件图像像素进行分组。针对每个动态障碍物生成一个检测框，包括中心点坐标$C^{O_k}=(x,y)$和边长$L^{O_k}=(w,h)$，w和h分别为检测框的宽和高。

3.2　融合检测

由于事件相机工作原理的限制，其不能较好地检测静态目标。当相机或目标高速运动时，基于事件相机动态目标检测的效果会优于基于普通相机的一步法机动目标检测效果；而在相机或目标运动非常缓慢时，基于普通相机的一步法机动目标检测效果会优于基于事件相机的动态目标检测效果。为实现高速运动和缓慢运动时都具有良好的检测效果，本文提出根据置信度将事件相机的检测结果和普通相机的检测结果进行融合。

如图 3 所示，基于 RGB 相机的目标检测采用一步法的检测算法（FD），使用了 YOLO-tiny[14] 的网络框架，检测流程更加简化，检测速度大幅提升，适用于无人机高速目标检测，其输出包括置信度P_{frame}和检测框。基于事件相机的动态目标检测（event-based detection, ED），由事件相机输出事件流，经事件滤波，进行运动补偿，通过 2.1 节所述方法完成动态目标检测。

融合前，需解决 2 种检测通道的时间同步问题。通常，RGB 相机的输出帧速度较低，约为 25 Hz（时间间隔为 40 ms），因此，FD 的频率也约为 25 Hz。由于基于事件数据数量事件图像表示方法的频率不确定，在融合检测时，ED 采用时间阈值的方法，一般取 20 ms 或 10 ms。当 ED 取 20 ms 时，FD 与 ED 的第 1 帧对齐，当 ED 取 10 ms 时，FD 与 ED 的第 2 帧对齐，如图 4 中红色框所示。

由于 ED 的检测过程无置信度的输出，并且 RGB 图像的目标特征更丰富，本文认为在无人机与目标相对速度较小时，RGB 图像的检测效果更可靠，相应得到的置信度P_{frame}越高；在无人机与目标相对速度较大时，RGB 图像会出现运动模糊的情况，相应的置信度P_{frame}越小，ED 的检测结果更可靠。融合策略表示如下：

$$B_{\mathrm{box}}=\begin{cases}B_{\text{box-frame}} & P_{\mathrm{frame}}\geqslant P_{\mathrm{threshold}}\\ B_{\text{box-event}} & P_{\mathrm{frame}}<P_{\mathrm{threshold}}\end{cases} \tag{12}$$

式中：B_{box}表示检测框；$P_{\mathrm{threshold}}$为设计参数，表示置信度阈值，本文设为 0.5。当$P_{\mathrm{frame}}\geqslant P_{\mathrm{threshold}}$时，将

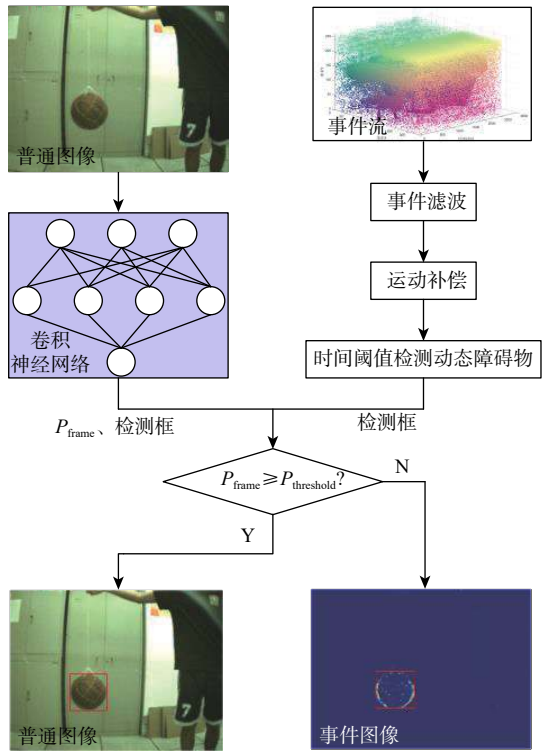


图 3　融合检测算法

Fig. 3　Fusion detection algorithm

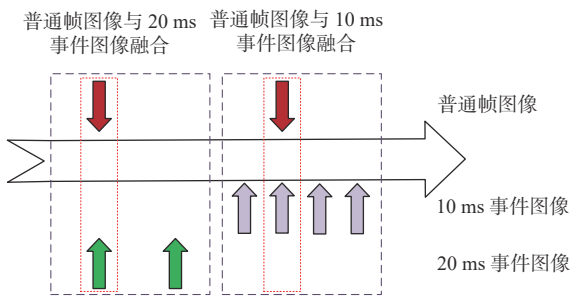


图 4　RGB 图像和事件图像时间线

Fig. 4　Timeline of RGB frame-image and event-image

FD 检测结果作为融合检测的输出；当$P_{\mathrm{frame}}<P_{\mathrm{threshold}}$时，将 ED 检测结果作为融合检测的输出。

4　基于速度障碍法的动态避障算法

4.1　动态障碍物的位置与速度估计

若已知当地重力加速度$\boldsymbol{g}^W$，并且假设球在过程中自由落体运动，忽略球在过程中受到的空气阻力，则根据初始位置$\boldsymbol{p}_{b_{i_0}}^W$和初始速度$\boldsymbol{v}_{b_{i_0}}^W$可估计出球的运动轨迹。因此，球的运动轨迹表示为

$$\boldsymbol{p}_{b_{i_0}}^W+\boldsymbol{v}_{b_{i_0}}^W\Delta t+\frac{1}{2}\boldsymbol{g}^W\Delta t^2=\boldsymbol{p}_{e_t}^W+d_{b_t}\boldsymbol{R}_{e_t}^W\boldsymbol{u}_{b_t} \tag{13}$$

式中：$\Delta t=t-t_0$为第一次检测到球到当前时间的时间差；$\boldsymbol{p}_{e_t}^W$为t时刻事件相机坐标原点在世界坐标中的位置；$\boldsymbol{R}_{e_t}^W$为 t 时刻事件相机坐标系相对于世界坐标系的旋转矩阵；d_{b_t}为t时刻球到相机的深度信息；

$\boldsymbol{u}_{b_i}$为t时刻像素坐标系下球的坐标。

1）深度估计。动态障碍物位置的确定需要深度信息，依靠单目相机测量深度一般需要目标的真实大小或连续2帧及2帧之间的平移与旋转运动信息，通常存在较大的扰动，因此，本文提出一种融合检测结果与深度相机的深度估计方法。

根据相机内外参可得事件相机像素到深度相机像素点的映射关系，如下：

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix}_D = \boldsymbol{K}_D{}^E\boldsymbol{R}_D^{-1}\boldsymbol{K}_E^{-1}\left[\frac{z_{CE}}{z_{CD}}\begin{bmatrix} u \\ v \\ 1 \end{bmatrix}_E - \boldsymbol{K}_E{}^E\boldsymbol{T}_D\right] \tag{14}$$

式中：下标E表示事件相机参数；下标D表示深度相机参数；z_{CE}为事件相机坐标系下深度信息；z_{CD}为深度相机坐标系下深度信息；$\boldsymbol{K}$、${}^E\boldsymbol{R}_D$和${}^E\boldsymbol{T}_D$分别为相机内参矩阵、事件相机坐标系相对于深度相机坐标系的旋转矩阵和平移矩阵。

如图5所示，C_D和C_E为事件相机和深度相机坐标系，事件相机和深度相机上下放置，由刚体设备固定，事件相机相对深度相机的旋转矩阵${}^E\boldsymbol{R}_D$和平移矩阵${}^E\boldsymbol{T}_D$可通过相机间外参标定确定。

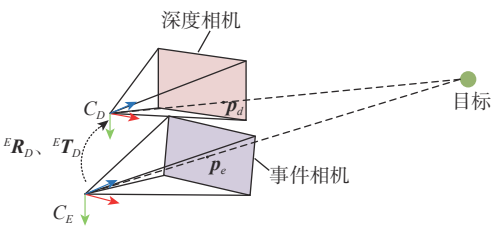


图5 融合测距原理

Fig. 5 Principle of fusion ranging

2）速度估计。确定位置信息后，需对速度做出估计，使用Kalman滤波[15]，因为对障碍物位置信息的确定是有噪声的，尤其是障碍物距离无人机较远时，而且速度障碍法对速度的要求十分准确，需要确定生成的速度障碍，同时对速度做出较为准确的估计。将障碍物的位置信息作为输入量，引入时间延迟（Kalman滤波是低通滤波器，通常时间延迟低于0.3 ms），以障碍物的位置和速度作为输出量。

3）动态障碍物表示。用椭球表示障碍物，因为其能很好地表示位置和大小的期望高斯误差[16]。从成像平面的事件图像数据构建世界参考系中的障碍物椭球。用迭代法在6个投影点周围拟合出最小体积椭球。

$$\begin{aligned} &\operatorname{mininum}_{\boldsymbol{A}} \ \operatorname{In}\det\left(\boldsymbol{A}^{-1}\right) \\ &\text{s.t.} \begin{cases} (\boldsymbol{p}_i-\boldsymbol{c})^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{p}_i-\boldsymbol{c}) \leqslant 1 & i=1,\cdots,m \\ \boldsymbol{A}>0 \end{cases} \end{aligned} \tag{15}$$

式中：$\boldsymbol{A}$为$n\times n$的椭圆方程的中心形式的矩阵，即满足$(\boldsymbol{x}-\boldsymbol{c})^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{x}-\boldsymbol{c})=1$；$\boldsymbol{c}$为$n$维向量，表示椭圆的中心；$\boldsymbol{p}_i$为矩阵$\boldsymbol{P}$的第$i$个向量，矩阵$\boldsymbol{P}$为包含$m$个点的$n$维数据矩阵。

上述算法直接针对累积的事件图像检测框内的事件进行拟合，根据拟合结果和深度信息、深度统计的误差，将障碍物在相机坐标系下进行表示。

算法1 球的拟合与表示算法。

输入： 事件图像、检测框、深度信息。

输出： 相机坐标系的坐标、椭球的各轴长。

根据检测和事件图像，确定数据矩阵$\boldsymbol{P}=\{p_1, p_2, \cdots, p_m\}$。

求解优化函数式(15)。

根据相机内参与深度信息确定障碍物在相机坐标系的坐标与椭球的轴长。

end

根据拟合结果，确定椭球的2个轴长，再结合深度信息和深度信息的误差δ_d，确定椭球另外一个轴长（即$2\delta_d$），根据式(7)获得动态障碍物在相机坐标系下的表示，即动态障碍物与相对的相对位置。

4.2 考虑无人机动力学约束的速度障碍法

传统的速度障碍法[17]仅考虑了当前控制周期内可实现的速度集合进行的规划，未充分考虑碰撞前无人机与障碍物间的距离和时间的影响，未考虑无人机动力学约束下可达速度空间的合理选择问题，因此，需要对传统的速度障碍法进行改进。

1）传统速度障碍法改进。如图6所示，对于运动的障碍物B，针对避障前的检测，考虑到碰撞前无人机和障碍物之间的距离和时间约束，改变障碍物的半径R_B，如下：

$$\hat{R}_B = (h(\lambda_t)+h(\lambda_d))R_B \tag{16}$$

式中：$\hat{R}_B$为更新后的障碍物半径；$\lambda_t = t_c/T$，$\lambda_d = L_{A,B}/L_{\text{safe}}$，$t_c = L_{A,B}/V_{A,B}$，$t_c$为碰撞时间，$L_{A,B}$为无人机和障碍物之间的距离，$V_{A,B}$为无人机和障碍物之间的相对速度，$L_{\text{safe}}$为无人机飞行过程中躲避动态障碍物的安全距离；$h(^*)$为权重函数，当$h(\lambda_t)+h(\lambda_d)>1$时，使$h(\lambda_t)+h(\lambda_d)=1$，保证障碍物半径不被过度放大。

根据文献[2]，为使无人机在高动态环境中对任务响应较快，尽量使其不改变纵向速度，即无人机前进方向速度大小不发生改变。如图7所示，确定安全距离L_{safe}。图中，$\boldsymbol{e}_1$方向为纵向运动方向，即速度$\boldsymbol{V}_A$不可改变方向，s为无人机的最大检测距离。感知延迟τ_{delay}表示障碍物进入感知区域的时间与无人机开始执行避障动作时刻之间的间隔。安全距离L_{safe}即为在$\boldsymbol{e}_2$方向完成横向运动的时间及感知

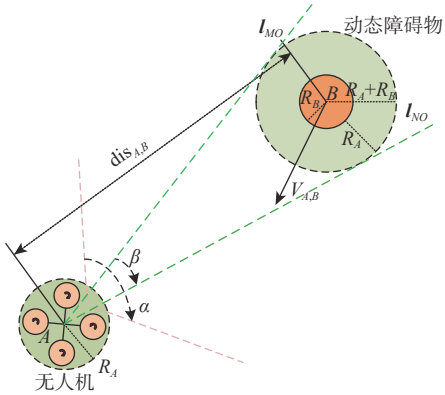


图 6　动态避障原理

Fig. 6　Dynamic obstacle avoidance

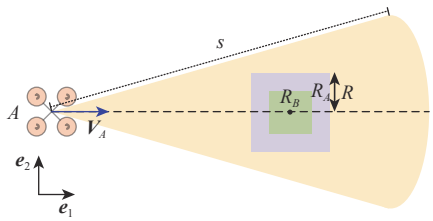


图 7　无人机动力学约束的避障模型

Fig. 7　Obstacle avoidance model of UAV with dynamic constraints

延迟时间段内，无人机在$\boldsymbol{e}_1$方向上最少飞行的距离。

$$L_{\mathrm{safe}}=\begin{cases} l & l<s \\ s & l\geqslant s \end{cases} \tag{17}$$

式中：$l=|\boldsymbol{V}_A|(t+\tau_{\mathrm{delay}})$为无人机在$\boldsymbol{e}_1$方向上最少飞行的距离，$t=\sqrt{2a_2(R_A+R_B)}$为在$\boldsymbol{e}_2$方向完成横向运动的时间，$R_A$为无人机的安全半径，$a_2$为$\boldsymbol{e}_2$方向上无人机的加速度。

针对无人机动力学约束下的可达速度空间，根据实际的无人机模型，确定无人机在Δt范围内，能够达到的速度变化和角度变化。在一个控制周期ΔT内，速度变化的最大值和最小值分别为$\Delta V_{A\max}$和$\Delta V_{A\min}$，角度变化的最大值和最小值为$\Delta\alpha_{A\max}$和$\Delta\alpha_{A\min}$。

2）最优避障策略设计。

策略 1　沿$\boldsymbol{l}_{MO}$或$\boldsymbol{l}_{NO}$。为保证无人机安全飞行，可使期望相对速度沿着相对碰撞堆的边缘，即此刻的无人机期望速度为$\boldsymbol{v}_{\max}(\boldsymbol{l}_{MO})$或$\boldsymbol{v}_{\max}(\boldsymbol{l}_{NO})$，且保证使得无人机速度改变最小。

策略 2　采用优化策略。在无人机高动态飞行中，应尽快响应任务，在保证无人机飞行安全的前提下，使得无人机的速度改变最小，偏离期望速度最小。设计的优化目标函数为$J(\Delta\boldsymbol{V})$，其中，障碍物速度$\boldsymbol{V}_B$已抵消，$\boldsymbol{V}_{\mathrm{target}}$为无人机的期望速度（朝向目标点方向），$\boldsymbol{V}_{\mathrm{cur}}$为无人机当前的飞行速度。

$$\begin{aligned} &\text{mininum} && \left[\boldsymbol{V}_{\mathrm{target}}-(\boldsymbol{V}_{\mathrm{cur}}+\Delta\boldsymbol{V})\right]^2+\mu\Delta\boldsymbol{V}^2 \\ &\text{s.t.} && \begin{cases} \Delta\boldsymbol{P}\dfrac{\boldsymbol{V}_{\mathrm{cur}}+\Delta\boldsymbol{V}}{|\boldsymbol{V}_{\mathrm{cur}}+\Delta\boldsymbol{V}|}+(R_A+R_B)^2-\Delta\boldsymbol{P}^2\leqslant 0 \\ -\tau\boldsymbol{g}\leqslant\Delta\boldsymbol{V}\leqslant\tau\boldsymbol{g} \end{cases} \end{aligned} \tag{18}$$

约束条件如式 (18) 所示。其中，第 1 个约束表示相对速度的方向要在锥面$\boldsymbol{l}_{MO}A\boldsymbol{l}_{NO}$外，$\Delta\boldsymbol{P}\dfrac{\boldsymbol{V}_{\mathrm{cur}}+\Delta\boldsymbol{V}}{|\boldsymbol{V}_{\mathrm{cur}}+\Delta\boldsymbol{V}|}$表示相对位置在相对速度方向的投影，第 1 个约束等号成立时，表示此时的速度方向恰好落于射线$\boldsymbol{l}_{MO}$或$\boldsymbol{l}_{NO}$上；第 2 个约束表示无人机速度的变化受无人机动力学约束，τ为无人机推重比，$\boldsymbol{g}$为1×3矢量，表示当地重力加速大小的向量。

5　试验与分析

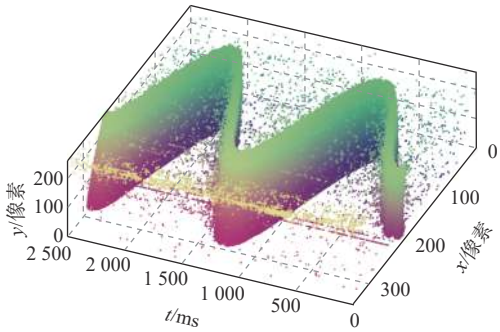
本文提出的融合事件图像和 RGB 图像的融合检测算法具有一定鲁棒性，根据无人机动力学约束改进的速度障碍法可躲避高速运动的动态障碍物。为验证算法的有效性和性能，进行了仿真试验和手持试验，并设计了无人机平台进行飞行试验。

5.1　仿真试验与手持试验

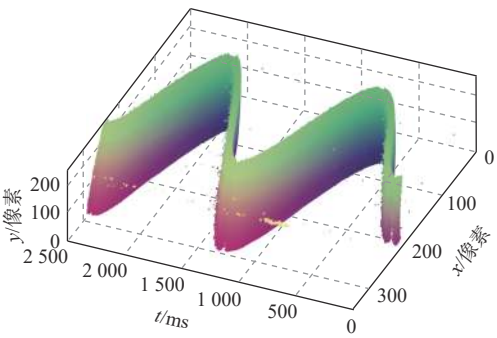
1）动态障碍物检测与定位。根据 2.1 节提出的事件流滤波方法，使用 DAVIS 346 Color 事件相机，记录碳素杆左右运动产生的事件。如图 8 所示，滤波前，时空图中沿时间轴一直存在的事件为热点噪声，无规则散落在时空中的事件为背景噪声。本文滤波算法可以较好地滤除背景噪声和热点噪声。当前没有统一的滤波评价标准，因为无法分清某一事件为真实事件还是噪声事件，时空图的方式能较直观地反映滤波效果，保留了有关联的事件和事件的异步性。

使用 DAVIS 346 color 相机，以累积事件数量 15 000 和时间间隔 20 ms 为例，对比运动补偿前后的图像，如图 9 所示。

在目标由运动到静止的过程中，关闭照明再打开，环境的光照强度由 145 lx 到 15 lx，再变为 145 lx。使用 DAVIS 346 color 相机，记录 RGB 图像和事件流进行目标检测。试验整个过程为 17 s，RGB 图像的频率为 25 Hz，事件图像的频率为 50 Hz，融合检测的设计参数$P_{\mathrm{threshold}}$取 0.6。试验结果如图 10 所示，图中融合结果的检测框颜色表示该结果来源不同（红色表示来源于事件图像检测，蓝色表示来源于 RGB 图像，绿色表示同时来源于两者），检测结果对比如表 1 所示（统计时以 RGB 图像频率为基准，总计 431 帧，其中，黑暗环境占 104 帧，目标静止时间占 32 帧）。融合检测结果相比于仅依靠



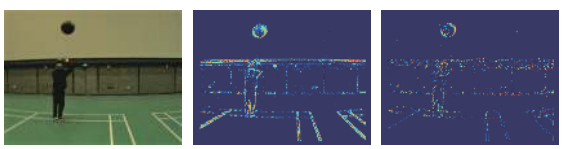
(a) 滤波前



(b) 滤波后

图 8 事件滤波时空图

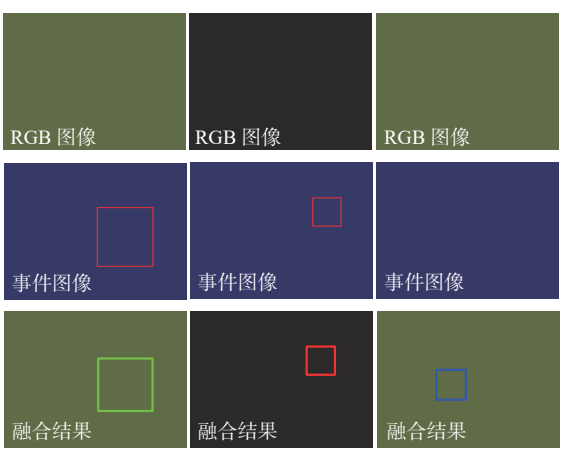
Fig. 8 Space-time diagram of event-filter



(a) RGB 图像 (b) 运动补偿前累积事件图像 (c) 运动补偿后事件图像

图 9 运动补偿效果

Fig. 9 Result of motion compensation



(a) 检测结果 1 (b) 检测结果 2 (c) 检测结果 3

图 10 融合检测结果

Fig. 10 Result of fusion detection

RGB 图像检测，检测率提高了 28.4%，相比于仅依靠事件图像检测结果，检测率提高了 9.1%。

根据检测结果的检测框，将其长宽进行缩放（试验中缩放为初始的 1/2），并在新的检测框中利用直方图进行统计，得出目标的深度信息。图 11 为目标抛物运动时融合测距试验。图 12 为目标在平行光轴运动时，PnP、基于针孔相机模型测距和融合测距结果对比，以 Optitrack 记录为参考真值，统计结果如表 2 所示，融合测距结果的最大误差和平均误差均为最小。

表 1 检测结果对比

Table 1 Comparison of detection results

数据来源	检测帧数	检测率/%
RGB图像	293	67.9
事件图像	376	87.2
融合结果	415	96.3

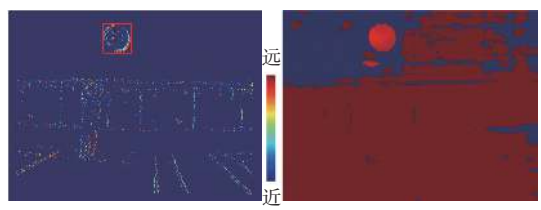


图 11 融合测距试验

Fig. 11 Fusion ranging test

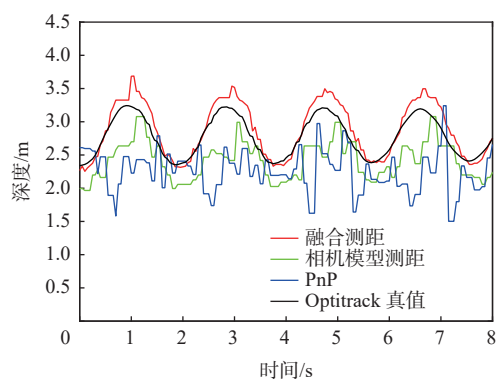


图 12 测距结果对比

Fig. 12 Comparison of ranging results

表 2 测距结果对比

Table 2 Comparison of ranging results m

方法类型	最大误差	平均误差
PnP	1.523 3	0.619 3
相机模型测距	0.826 1	0.351 3
融合测距	0.477 0	0.136 8

2）避障仿真。按照真实试验飞机平台参数设计，试验参数如表 3 所示。考虑视场角和传感器检测距离，本文设计的无人机高动态避障算法如算法 2 所示。

算法 2 无人机高动态避障算法。

输入： 无人机运动状态$\boldsymbol{S}_A(t)$、障碍物运动状态$\boldsymbol{S}_B(t)$。

输出： 下一时刻的目标速度$\boldsymbol{V}_B$。

If $\mathrm{dis}_{A,B} < s$ and $< (\boldsymbol{P}_B - \boldsymbol{P}_A, \boldsymbol{V}_A) < \alpha$ then

确定绝对碰撞堆ACC与相对速度$V_{A,B}$

If $V_{A,B} \in ACC$ then

由式 (18) 求得期望速度变化$\Delta \boldsymbol{V}$

$\boldsymbol{V}_{\mathrm{cur}} \leftarrow \boldsymbol{V}_{\mathrm{cur}} + \Delta \boldsymbol{V}$

End if

End if

End

表 3　避障仿真参数

Table 3　Obstacle avoidance simulation parameter

参数	数值
无人机质量m_A/kg	1.2
无人机推重比τ	6
无人机安全半径R_A/cm	30
无人机最大速度$\lvert V_{A\max}\rvert$/(m·s^{-1})	5
无人机视角α/(°)	120
无人机最大检测距离s/m	4
障碍物半径R_B/cm	12.3
障碍物最大速度$\lvert V_{B\max}\rvert$/(m·s^{-1})	12

在二维仿真中, 设置初始位置为(0 m,0 m), 目标位置为(0 m,15 m), 在(0 m,8 m)处抛掷动态障碍物, 尽量不改变图 7 所示的$\boldsymbol{e}_1$方向的速度, 以飞行时间为代价, 进行避障仿真, 并对比了文献 [12] 的人工势场法。仿真结果如图 13 所示, 结果对比如表 4 所示。

根据仿真结果, 本文算法在躲避动态障碍物时的路径优于文献 [12] 的避障算法, 且不考虑斥力和吸引力的设计问题及无人机的偏航角控制问题。

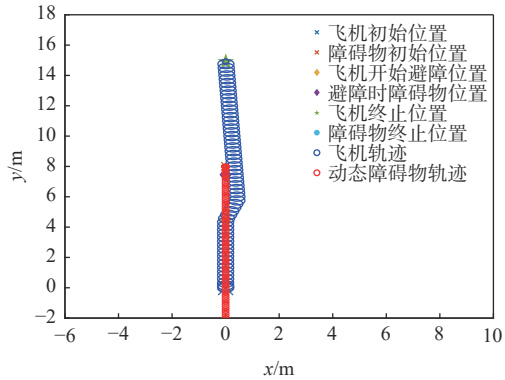
在三维仿真中, 设置初始位置为(0,0,0) m、目标位置为(10,15,2) m, 并在沿途设置 2 个障碍物, 障碍物水平速度为 9.6 m/s, 做自由落体运动, 如图 14 所示。

5.2　实际飞行试验

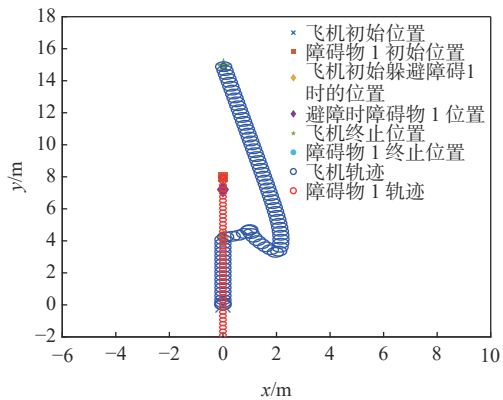
1) 试验系统架构。如图 15 所示, 无人机采用轴距为 330 mm 的四旋翼无人机, 搭载 CUAV V5+ 飞控、DJI Manifold-2G 计算平台、DAVIS 346 color 事件相机、Intel RealSense D435i 深度相机及主动捕捉系统定位球, 用于记录真实数据。软件平台基于机器人操作系统(robot operating system, ROS)进行各模块间的信息通信。

2) 试验结果。由于无人机搭载的 DJI Manifold-2G 平台计算有限, 进行飞行试验时, 未采用优化的避障策略, 而采用最优避障策略, 即策略 1。

如图 16 所示, 进行实际飞行试验时, 无人机悬停, 向无人机抛掷篮球, 篮球的水平速度为



(a) 速度障碍法 (使用优化)



(b)人工势场法

图 13　9.6 m/s动态障碍物仿真结果

Fig. 13　simulation results of 9.6 m/s dynamic obstacle

表 4　避障仿真结果对比

Table 4　Comparison of obstacle avoidance simulation results

障碍物速度/(m·s^{-1})	飞行时间/ms	
	文献[12]方法	本文方法
2.4	3 274	3 082
9.6	3 956	3 086

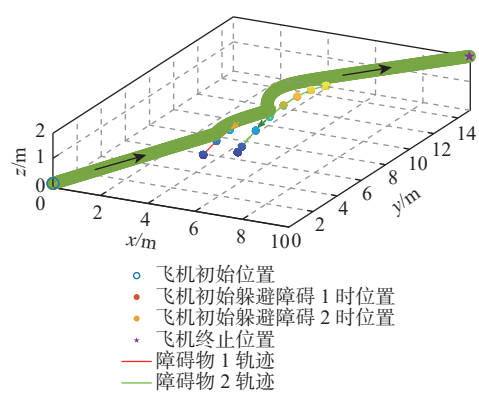


图 14　3D 避障仿真

Fig. 14　3D obstacle avoidance simulation

7.018 m/s, 该过程的速度变化曲线与滚转角裱花曲线分别如图 17 和图 18 所示。过程中, 无人机

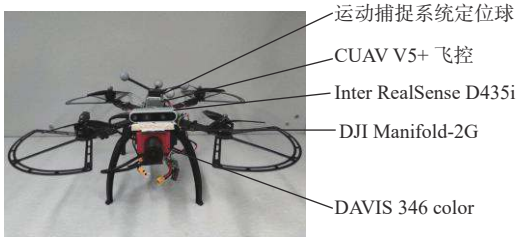


图 15 试验硬件平台

Fig. 15 Test hardware platform



图 16 实际飞行试验

Fig. 16 Real world experience

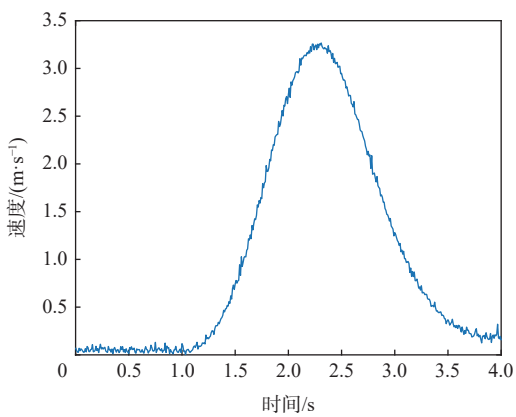


图 17 无人机速度变化曲线

Fig. 17 UAV velocity change curve

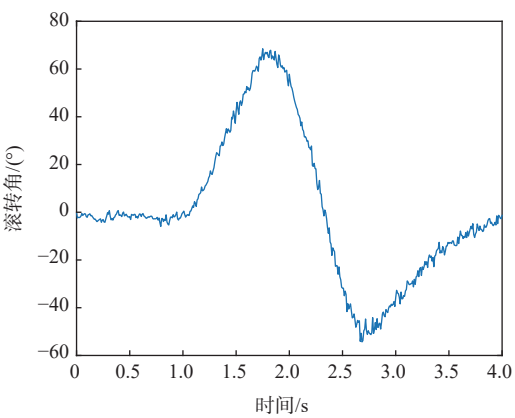


图 18 无人机滚转角变化曲线

Fig. 18 UAV roll angle change curve

最大速度变化为 3.267 m/s, 表明能够有效地躲避动态障碍物。

6 结 论

本文设计的基于动态视觉传感器无人机目标检测与避障算法能够较好地在高动态场景下完成无人机的目标检测与躲避速度较高的动态障碍物。

1) 事件滤波方法能够有效地滤除热点噪声和背景噪声, 并保持事件的异步性。

2) 融合检测算法在动态场景中优于仅依靠RGB 相机的目标检测和仅依靠事件相机的目标检测, 具有一定的鲁棒性。

3) 改进的速度障碍法能有效地躲避高速障碍物, 在时间代价上优于基于人工势场的避障算法。

但依旧有很大的改进空间: ①直接使用事件进行目标检测, 利用事件的异步性完成目标检测; ②设计端到端的系统框架, 减少各模块之间通信造成的感知延迟。

参考文献 (References)

[1] 王亭亭, 蔡志浩, 王英勋. 无人机室内视觉/惯导组合导航方法[J]. 北京航空航天大学学报, 2018, 44(1): 176-186.
WANG T T, CAI Z H, WANG Y X. Integrated indoor vision/inertial navigation method of UAVs in indoor environment[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(1): 176-186(in Chinese).

[2] FALANGA D, KIM S, SCARAMUZZA D. How fast is too fast? The role of perception latency in high-speed sense and avoid[J]. IEEE Robotics and Automation Letters, 2019, 4(2): 1884-1891.

[3] DELBRUCK T. Neuromorophic vision sensing and processing[C]// Proceedings of the 46th European Solid State Device Research Conference. Piscataway: IEEE Press, 2016: 7-14.

[4] BRANDLI C, BERNER R, YANG M H, et al. A 240 × 180 130 dB 3 μs latency global shutter spatiotemporal vision sensor[J]. IEEE Journal of Solid-State Circuits, 2014, 49(10): 2333-2341.

[5] GALLEGO G, DELBRÜCK T, ORCHARD G M, et al. Event-based vision: A survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 44(1): 154-180.

[6] KHODAMORADI A, KASTNER R. $O(N)$-space spatiotemporal filter for reducing noise in neuromorphic vision sensors[J]. IEEE Transactions on Emerging Topics in Computing, 2021, 9(1): 15-23.

[7] FENG Y, LV H Y, LIU H L, et al. Event density based denoising method for dynamic vision sensor[J]. Applied Sciences, 2020, 10(6): 2024.

[8] 闫昌达, 王霞, 左一凡, 等. 基于事件相机的可视化及降噪算法[J]. 北京航空航天大学学报, 2021, 47(2): 342-350.
YAN C D, WANG X, ZUO Y F, et al. Visualization and noise reduction algorithm based on event camera[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(2): 342-350(in Chinese).

[9] MUEGGLER E, FORSTER C, BAUMLI N, et al. Lifetime estimation of events from dynamic vision sensors[C]//Proceedings of the IEEE International Conference on Robotics and Automation. Piscat-

away: IEEE Press, 2015: 4874-4881.

[10] MITROKHIN A, FERMÜLLER C, PARAMESHWARA C, et al. Event-based moving object detection and tracking[C]//Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway: IEEE Press, 2019: 1-9.

[11] HE B T, LI H J, WU S Y, et al. FAST-dynamic-vision: Detection and tracking dynamic objects with event and depth sensing[C]//Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway: IEEE Press, 2021: 3071-3078.

[12] FALANGA D, KLEBER K, SCARAMUZZA D. Dynamic obstacle avoidance for quadrotors with event cameras[J]. Science Robotics, 2020, 5(40): 1-14.

[13] SANKET N J, PARAMESHWARA C M, SINGH C D, et al. EVDodgeNet: Deep dynamic obstacle dodging with event cameras[C]//Proceedings of the IEEE International Conference on Robotics and Automation. Piscataway: IEEE Press, 2020: 10651-10657.

[14] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 779-788.

[15] JULIER S J, UHLMANN J K. Unscented filtering and nonlinear estimation[J]. Proceedings of the IEEE, 2004, 92(3): 401-422.

[16] MOSHTAGH N. Minimum volume enclosing ellipsoids[J]. Convex Optimization, 2005, 111: 1-9.

[17] VAN DEN BERG J, LIN M C, MANOCHA D. Reciprocal velocity obstacles for real-time multi-agent navigation[C]//Proceedings of the IEEE International Conference on Robotics and Automation. Piscataway: IEEE Press, 2008: 1928-1935.

Object detection and obstacle avoidance based on dynamic vision sensor for UAV

CAI Zhihao，CHEN Wenjun，ZHAO Jiang，WANG Yingxun*

(School of Automation Science and Electrical Engineering，Beihang University，Beijing 100191，China)

Abstract：UAVs face a significant problem while trying to avoid moving objects while in flight. In order to detect and avoid high-speed dynamic obstacles in a dynamic environment, an algorithm for target detection and obstacle avoidance based on a dynamic vision sensor was designed. Firstly, we propose an event filter method and to filter the background noise, hot noise, and the method preserves the asynchrony of events. The motion compensation algorithm is designed to filter redundant events caused by the camera's own motion in the event stream. For dynamic object detection, a fusion detection algorithm of event image and RGB image is designed, it has higher robustness in a highly dynamic environment. Finally, to avoid dynamic obstacles combined with the features of obstacle movement and UAV dynamic restrictions, finally, we enhanced the velocity obstacle method and estimated the target trajectory in accordance with the detection results. A large number of simulation tests, hand-held tests, and flight tests are carried out to verify the feasibility of the algorithm.

Keywords：event-based camera; event filter; motion compensation; fusion detection; velocity obstacle method

Received：2022-04-01；**Accepted**：2022-05-06；**Published Online**：2022-05-26 14：44

URL：link.cnki.net/urlid/11.2625.V.20220526.0927.001

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0202

板翅式换热器传递函数动态模型及参数确定方法

梁兴壮[1,2,*]，黄志远[2]，艾凤明[2]，袁振伟[2]，汪箭[1]

(1. 中国科学技术大学 火灾科学国家重点实验室，合肥 230026；2. 中国航空工业集团公司沈阳飞机设计研究所，沈阳 110035)

摘 要：飞机综合环控/热管理系统迫切需要建立快捷高效的换热器动态仿真模型，以满足先进控制系统设计需求。基于此，针对板翅式换热器动态模型，提出一种包含2个延迟环节和4个一阶惯性环节的传递函数矩阵形式，并从换热器机理模型出发，利用拉普拉斯变换推导模型中4个时间常数的计算公式，提出基于换热器效率的传热热阻计算关系式的辨识方法，解决了主要模型参数设置难题。以板翅式换热器为研究对象，在Simulink仿真平台搭建其传递函数动态模型，并与AMESim仿真平台搭建的机理模型进行对比分析，结果显示：2种模型所得空气和冷却水出口温度分别随入口温度和质量流量阶跃变化的动态响应曲线吻合很好，其中，4种工况空气和冷却水出口温度最大稳态偏差分别为0.034 ℃和0.029 ℃，当冷却水入口质量流量阶跃变化时，空气出口温度动态响应相对偏差最大，为9.27%，当空气入口质量流量阶跃变化时，冷却水出口温度动态响应相对偏差最大，为7.03%。

关 键 词：换热器；动态模型；传递函数矩阵；时间常数；传热热阻

中图分类号：V245.3

文献标志码：A **文章编号**：1001-5965(2024)01-0154-09

现代飞机综合环控/热管理系统包括空气循环制冷、蒸发循环制冷、电子设备液冷及燃油热管理等多个循环子系统，各子系统之间通过换热器进行热量交换[1]。换热器的换热性能随进口流体参数或热载荷变化的动态响应时间较长，多个换热器动态特性相互耦合作用使飞机综合环控/热管理系统长时间处于复杂的非平衡状态。为此，迫切需要建立快捷高效的换热器动态仿真模型，以满足该系统优化设计和基于模型的自适应控制需求[2]。

换热器传统动态仿真模型采用集总参数思想，基于以质量、动量和能量守恒等偏微分方程组表示的换热机理进行建立，具体可分为集总参数、分布参数和分段集总参数3种[3]。其中，集总参数动态仿真模型通常把整体换热器的某个参数视为一个节点进行简化求解计算[4-5]；分布参数动态仿真模型通常沿换热器流体流动方向进行分段，将主要参数随空间的偏微分项离散化，进行求解[6]；分段集总参数动态仿真模型通常将换热器分为若干段，每段按照集总参数法求解[7]。上述3种仿真模型涉及偏微分或微分方程组求解，不适应复杂系统优化设计或先进控制系统设计。

近年来，随着各行各业对系统优化的迫切需求，以及基于模型的优化控制技术的不断发展，关于换热器动态特性及动态仿真模型吸引了众多科研工作者进行研究[8-9]，主要包括采用实验和CFD数值分析方法进一步研究换热器动态响应特性[10-13]，以及在传统机理仿真模型基础上提出3类简化建模方法，即节点热网络简化方法[14-16]、基于拉普拉斯变换的近似一阶惯性的传递函数方法[17-19]和基于输入输出统计数据的代理模型方法[20-22]。

研究结果表明，节点热网络仿真模型求解速度

收稿日期：2022-04-01；录用日期：2022-04-23；网络出版时间：2022-05-10 17:33
网络出版地址：link.cnki.net/urlid/11.2625.V.20220509.1528.010
*通信作者. E-mail：liangxingzhuang@126.com

引用格式：梁兴壮，黄志远，艾凤明，等. 板翅式换热器传递函数动态模型及参数确定方法[J]. 北京航空航天大学学报，2024，50(1)：154-162. LIANG X Z，HUANG Z Y，AI F M，et al. Dynamic model by transfer function and parameter determination method of plate fin heat exchanger[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(1)：154-162 (in Chinese).

较慢，需要合理选择节点数和时间步长[15]。传递函数及代理模型求解速度快，与换热器动态特性吻合较好，可用于先进控制算法研究[19-20]，但不同学者给出的模型具体表达方式有一定差别，如时间常数等模型参数需要结合工程经验公式或大量的试验数据才能确定，给模型的实际推广使用带来很大难度。

本文针对板翅式换热器传热过程，提出了一种由 2 个延迟环节和 4 个一阶惯性环节组成的、以传递函数表示的简化动态仿真模型，研究了模型中稳态增益系数、响应时间常数及传热热阻等关键参数的确定方法，并在 Simulink 仿真平台搭建仿真模型。以某空-液板翅式换热器为例，与 AMESim 仿真平台搭建机理模型进行对比分析，结果显示，冷热流体出口温度随入口温度和流量阶跃变化的动态响应过程吻合较好，满足复杂系统优化设计或先进控制系统设计需求。

1　机理模型

换热器机理模型一般采用模块化和集总参数建模方法。首先，根据逆流或叉流换热方式将换热器划分为多个换热模块，每个模块内 2 种流体及传热壁面三者之间流动传热过程如图 1 所示。假设换热器内部流体流动是一维的，忽略壁面的传热热阻和轴向热传导，且与外界无热量交换。

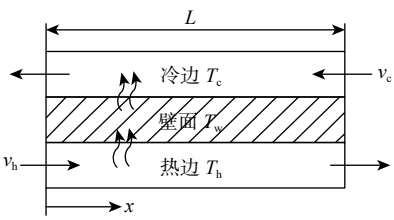


图 1　换热器单个模块内流动传热过程示意图

Fig. 1　Schematic diagram of flow and heat transfer process in a single module of a heat exchanger

冷热流体和传热壁面 3 个节点的能量守恒方程分别为

$$L\left(\frac{q_m c_p}{v}\right)_c\frac{\partial T_c}{\partial \tau}-L\left(q_m c_p\right)_c\frac{\partial T_c}{\partial x}=(\eta_o\alpha A)_c(T_w-T_c) \tag{1}$$

$$L\left(\frac{q_m c_p}{v}\right)_h\frac{\partial T_h}{\partial \tau}+L\left(q_m c_p\right)_h\frac{\partial T_h}{\partial x}=(\eta_o\alpha A)_h(T_w-T_h) \tag{2}$$

$$\left(Mc_p\right)_w\frac{\partial T_w}{\partial \tau}=(\eta_o\alpha A)_h(T_h-T_w)+(\eta_o\alpha A)_c(T_c-T_w) \tag{3}$$

式中：T 为温度；q_m 为质量流量；c_p 为比热容；v 为流速；L 为流动长度；η_o 为翅片效率；α 为传热系数；A 为传热面积；M 为传热结构质量；τ 为时间；x 为流动方向；下标 h 表示热边流体，c 表示冷边流体，w 表示换热壁面。

2　简化动态仿真模型

2.1　动态仿真模型

换热器中，冷热流体及传热结构具有储热能力，同时彼此之间存在传热阻力，为此，冷热流体间热量传输必经过储放热过程，并在温度的动态响应上表现出热惯性。

一阶热惯性环节的微分方程表达式为

$$\tau_0\frac{\mathrm{d}y(\tau)}{\mathrm{d}\tau}+y(\tau)=Ku(\tau) \tag{4}$$

式中：τ_0 为时间常数；K 为稳态增益；$y(\tau)$ 为出口温度；$u(\tau)$ 为入口温度。

式(4)对时间进行拉普拉斯变换，可整理成传递函数为

$$G(s)=\frac{Y(s)}{U(s)}=\frac{K}{\tau_0 s+1} \tag{5}$$

因为板翅式换热器的传热壁面非常薄，所以对于空-液换热器或液-液换热器而言，常常可以将传热结构热容全部或平均分配到液体工质中[23]。这样，冷热流体之间可以简化为一个传热热阻，进而使用一阶惯性环节建立其动态响应关系。

同时，考虑到换热器内流体热容和流速影响，流体入口温度发生阶跃变化后，需要经过一段时间才能到达出口。为此，同一侧流体出口温度随入口温度的响应特性需要在一阶惯性环节基础上增加一个纯延迟环节。

为此，参考板翅式换热器冷热流体传热机理，建立其简化动态仿真模型，如图 2 所示。该仿真模型可以描述换热器冷热流体出口温度的动态响应特性，主要由 4 个一阶惯性环节和 2 个纯延迟环节组成，具体包含 4 个稳态增益 K_1～K_4、4 个动态时间常数 τ_1～τ_4、2 个纯延迟时间常数 τ_{h_delay} 和 τ_{c_delay}。

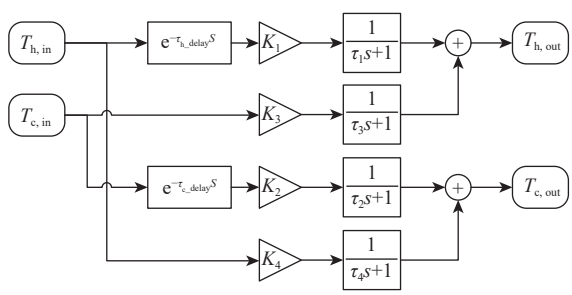


图 2　换热器简化动态仿真模型

Fig. 2　Simplified dynamic simulation model of heat exchanger

2.2 稳态增益系数确定方法

换热器简化动态仿真模型中,稳态增益系数决定了冷热流体出口温度平衡状态值,可以利用换热器换热效率进行确定。

换热器冷热流体进出口参数满足能量守恒,即

$$\left(q_m c_p\right)_{\mathrm{h}}\left(T_{\mathrm{h,in}}-T_{\mathrm{h,out}}\right)=\left(q_m c_p\right)_{\mathrm{c}}\left(T_{\mathrm{c,out}}-T_{\mathrm{c,in}}\right) \tag{6}$$

式中:下标 in 表示入口,out 表示出口。

换热器效率 η 为实际传热量与理论最大传热量之比,即

$$\eta=\frac{\left(q_m c_p\right)_{\mathrm{h}}\left(T_{\mathrm{h,in}}-T_{\mathrm{h,out}}\right)}{\left(q_m c_p\right)_{\min}\left(T_{\mathrm{h,in}}-T_{\mathrm{c,in}}\right)} \tag{7}$$

式中:下标 min 表示冷热流体热容较小值。

根据换热器冷热流体进出口能量守恒方程式(4)和换热效率定义式(5),可得

$$\begin{cases}T_{\mathrm{h,out}}=\left(1-\eta\dfrac{\left(q_m c_p\right)_{\min}}{\left(q_m c_p\right)_{\mathrm{h}}}\right)T_{\mathrm{h,in}}+\eta\dfrac{\left(q_m c_p\right)_{\min}}{\left(q_m c_p\right)_{\mathrm{h}}}T_{\mathrm{c,in}}\\ T_{\mathrm{c,out}}=\left(1-\eta\dfrac{\left(q_m c_p\right)_{\min}}{\left(q_m c_p\right)_{\mathrm{c}}}\right)T_{\mathrm{c,in}}+\eta\dfrac{\left(q_m c_p\right)_{\min}}{\left(q_m c_p\right)_{\mathrm{c}}}T_{\mathrm{h,in}}\end{cases} \tag{8}$$

则 4 个稳态增益系数计算公式为

$$\begin{cases}K_1=1-\eta\left(q_m c_p\right)_{\min}/\left(q_m c_p\right)_{\mathrm{h}}\\ K_2=1-\eta\left(q_m c_p\right)_{\min}/\left(q_m c_p\right)_{\mathrm{c}}\\ K_3=\eta\left(q_m c_p\right)_{\min}/\left(q_m c_p\right)_{\mathrm{h}}\\ K_4=\eta\left(q_m c_p\right)_{\min}/\left(q_m c_p\right)_{\mathrm{c}}\end{cases} \tag{9}$$

2.3 响应时间常数确定方法

从换热器机理模型的特征方程式(1)~式(3)出发,通过拉普拉斯变换和反变换推导其动态模型中 4 个时间常数,推导过程如图 3 所示。

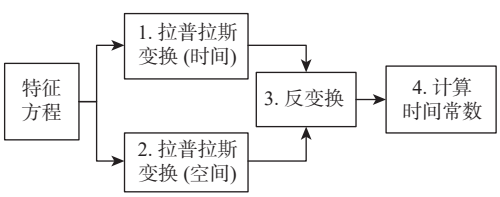


图 3 时间常数推导过程

Fig. 3 Derivation process of time constant

2.3.1 时间项拉普拉斯变换

用记号 $\tilde{T}$ 表示温度对时间导数项拉普拉斯变换。首先,对式(3)中壁面温度进行拉普拉斯变换,得

$$\tilde{T}_{\mathrm{w}}(s)=\frac{1}{\phi_{\mathrm{w}}s+1}\left(\frac{(\eta_{\mathrm{o}}\alpha A)_{\mathrm{h}}\tilde{T}_{\mathrm{h}}(s)}{(\eta_{\mathrm{o}}\alpha A)_{\mathrm{h}}+(\eta_{\mathrm{o}}\alpha A)_{\mathrm{c}}}+\frac{(\eta_{\mathrm{o}}\alpha A)_{\mathrm{c}}\tilde{T}_{\mathrm{c}}(s)}{(\eta_{\mathrm{o}}\alpha A)_{\mathrm{h}}+(\eta_{\mathrm{o}}\alpha A)_{\mathrm{c}}}\right) \tag{10}$$

然后,将式(1)和式(2)分别对时间进行拉普拉斯变换,并将式(10)代入,得

$$UA\frac{\tilde{T}_{\mathrm{c}}(l,s)}{\phi_{\mathrm{w}}s+1}=\tilde{T}_{\mathrm{h}}(l,s)\left(s\left(\frac{L}{v}q_m c_p\right)_{\mathrm{h}}+UA\frac{\phi_{\mathrm{c}}s+1}{\phi_{\mathrm{w}}s+1}\right)+\left(q_m c_p\right)_{\mathrm{h}}\frac{\mathrm{d}\tilde{T}_{\mathrm{h}}(l,s)}{\mathrm{d}l} \tag{11}$$

$$UA\frac{\tilde{T}_{\mathrm{h}}(l,s)}{\phi_{\mathrm{w}}s+1}=\tilde{T}_{\mathrm{c}}(l,s)\left(s\left(\frac{L}{v}q_m c_p\right)_{\mathrm{c}}+UA\frac{\phi_{\mathrm{h}}s+1}{\phi_{\mathrm{w}}s+1}\right)-\left(q_m c_p\right)_{\mathrm{c}}\frac{\mathrm{d}\tilde{T}_{\mathrm{c}}(l,s)}{\mathrm{d}l} \tag{12}$$

式中:$\phi_{\mathrm{w}}=M_{\mathrm{w}}c_{p,\mathrm{w}}/\left((\eta_{\mathrm{o}}\alpha A)_{\mathrm{h}}+(\eta_{\mathrm{o}}\alpha A)_{\mathrm{c}}\right)$;$\phi_{\mathrm{h}}=M_{\mathrm{w}}c_{p,\mathrm{w}}/(\eta_{\mathrm{o}}\alpha A)_{\mathrm{h}}$;$\phi_{\mathrm{c}}=M_{\mathrm{w}}c_{p,\mathrm{w}}/(\eta_{\mathrm{o}}\alpha A)_{\mathrm{c}}$;$l=x/L$;$(UA)^{-1}=(\eta_{\mathrm{o}}\alpha A)_{\mathrm{h}}^{-1}+(\eta_{\mathrm{o}}\alpha A)_{\mathrm{c}}^{-1}$。

2.3.2 空间项拉普拉斯变换

设空间项 l 对应的复变量为 n,对式(11)和式(12)分别进行空间拉普拉斯变换,得

$$(n+f_{\mathrm{h}}(s))\tilde{T}_{\mathrm{h}}(n,s)=\tilde{T}_{\mathrm{h}}(0,s)+\frac{UA}{\left(q_m c_p\right)_{\mathrm{h}}}\cdot\frac{\tilde{T}_{\mathrm{c}}(n,s)}{\phi_{\mathrm{w}}s+1} \tag{13}$$

$$(n-f_{\mathrm{c}}(s))\tilde{T}_{\mathrm{c}}(n,s)=\tilde{T}_{\mathrm{c}}(0,s)-\frac{UA}{\left(q_m c_p\right)_{\mathrm{c}}}\cdot\frac{\tilde{T}_{\mathrm{h}}(n,s)}{\phi_{\mathrm{w}}s+1} \tag{14}$$

式中:$f_{\mathrm{h}}(s)=sL/v_{\mathrm{h}}+\dfrac{UA}{\left(q_m c_p\right)_{\mathrm{h}}}\cdot\dfrac{\phi_{\mathrm{c}}s+1}{\phi_{\mathrm{w}}s+1}$;$f_{\mathrm{c}}(s)=sL/v_{\mathrm{c}}+\dfrac{UA}{\left(q_m c_p\right)_{\mathrm{c}}}\cdot\dfrac{\phi_{\mathrm{h}}s+1}{\phi_{\mathrm{w}}s+1}$。

整理式(13)和式(14),得

$$\tilde{T}_{\mathrm{c}}(n,s)=\frac{(n+f_{\mathrm{h}}(s))\tilde{T}_{\mathrm{c}}(0,s)}{(n+g_1)^2-g_3^2}-\frac{UA\cdot\tilde{T}_{\mathrm{h}}(0,s)}{\left(q_m c_p\right)_{\mathrm{c}}(\phi_{\mathrm{w}}s+1)\left[(n+g_1)^2-g_3^2\right]} \tag{15}$$

式中:$g_1=(f_{\mathrm{h}}(s)-f_{\mathrm{c}}(s))/2$;$g_2=(f_{\mathrm{h}}(s)+f_{\mathrm{c}}(s))/2$;$g_3^2=g_2^2-(UA)^2/\left[\left(q_m c_p\right)_{\mathrm{c}}\left(q_m c_p\right)_{\mathrm{h}}(\phi_{\mathrm{w}}s+1)^2\right]$。

2.3.3 拉普拉斯反变换

对式(15)空间项进行拉普拉斯反变换,得

$$\tilde{T}_{\mathrm{c}}(l,s)=\mathrm{e}^{-g_1 l}\tilde{T}_{\mathrm{c}}(0,s)\left[\cosh(g_3 l)+\frac{(f_{\mathrm{h}}-g_1)\sinh(g_3 l)}{g_3}\right]-\frac{UA}{\left(q_m c_p\right)_{\mathrm{c}}}\cdot\frac{\mathrm{e}^{-g_1 l}}{\phi_{\mathrm{w}}s+1}\cdot\frac{\sinh(g_3 l)}{g_3}\tilde{T}_{\mathrm{h}}(0,s) \tag{16}$$

式中:l=1 时表示冷侧入口,l=0 时表示冷侧出口。

则冷侧出口温度表达式为

$$T_{\mathrm{c,out}}=\tilde{T}_{\mathrm{c}}(0,s)=\frac{\mathrm{e}^{g_1}\tilde{T}_{\mathrm{c}}(1,s)}{\cosh g_3+\dfrac{g_2\sinh g_3}{g_3}}+\frac{UA}{\left(q_m c_p\right)_{\mathrm{c}}}\cdot\frac{\tilde{T}_{\mathrm{h}}(0,s)\sinh g_3}{(\phi_{\mathrm{w}}s+1)g_3\left(\cosh g_3+\dfrac{g_2\sinh g_3}{g_3}\right)} \tag{17}$$

2.3.4 响应时间常数计算

由式(17)可得

$$\frac{T_{c,out}}{T_{c,in}}=\frac{\tilde{T}_c(0,s)}{\tilde{T}_c(1,s)}=\frac{e^{g_1}}{\cosh g_3+g_2\dfrac{\sinh g_3}{g_3}} \tag{18}$$

$$\frac{T_{c,out}}{T_{h,in}}=\frac{\tilde{T}_c(0,s)}{\tilde{T}_h(0,s)}=\frac{UA}{(q_mc_p)_c(\phi_w s+1)(g_3\coth g_3+g_2)} \tag{19}$$

令$h_2(s)=e^{-g_1}\left[\cosh g_3+(g_2\sinh g_3)/g_3\right]$，对其在$s$=0v处进行麦克劳林展开，并忽略二阶以上项，则式(18)可以简化为

$$\frac{T_{c,out}}{T_{c,in}}=\frac{1}{h_2(s)}=\frac{1}{h_2(0)\left[1+s\dfrac{h_2'(0)}{h_2(0)}\right]}=\frac{(h_2(0))^{-1}}{\dfrac{h_2'(0)}{h_2(0)}s+1} \tag{20}$$

结合图 2 所示换热器一阶简化动态模型可知，冷边流体入口温度到出口温度的一阶响应时间常数为

$$\tau_2=\frac{h_2'(0)}{h_2(0)} \tag{21}$$

同理，可推导出另外 3 个时间常数 τ_1、τ_3、τ_4，如下：

$$\begin{cases}\tau_1=g_{10}'+\dfrac{g_{30}g_{30}'-g_{20}'+\dfrac{g_{20}g_{30}'}{g_{30}}-g_{20}g_{30}'\coth g_{30}}{g_{30}\coth g_{30}-g_{20}}\\ \tau_2=-g_{10}'+\dfrac{g_{30}g_{30}'+g_{20}'-\dfrac{g_{20}g_{30}'}{g_{30}}+g_{20}g_{30}'\coth g_{30}}{g_{30}\coth g_{30}+g_{20}}\\ \tau_3=\phi_w+\dfrac{g_{20}'+g_{30}'\coth g_{30}-g_{30}g_{30}'\operatorname{csch}^2 g_{30}}{g_{30}\coth g_{30}+g_{20}}\\ \tau_4=\tau_3\end{cases} \tag{22}$$

式中：$g_{10}'=g_1'(0)$；$g_{20}=g_2(0)$；$g_{20}'=g_2'(0)$；$g_{30}=g_3(0)$；$g_{30}'=g_3'(0)$。

2.4　模型参数计算流程

板翅式换热器简化动态仿真模型参数具体计算流程如图 4 所示。其中，稳态增益系数根据换热器效率利用式(9)进行计算；动态响应时间常数由式(22)进行计算，主要由换热器内冷热流体、芯体热容量等固定参数及冷热流体质量流量、传热热阻等状态参数决定；同一种流体出口温度随入口温度变化的纯延迟环节时间常数可由换热器长度与流速之比来确定。

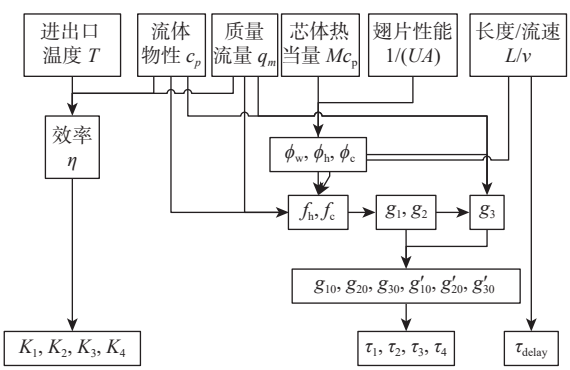


图 4　换热器模型参数计算流程

Fig. 4　Calculation process of heat exchanger model parameters

为此，该仿真模型输入参数主要包括冷热流体的入口温度、质量流量及物性参数，以及换热器芯体质量、传热热阻、翅片结构参数。

3　传热热阻辨识方法

传热热阻是换热器动态仿真模型换热效率和时间常数确定的关键参数。目前，针对不同板翅结构的传热因子经验计算公式的研究成果较少，且受加工过程影响，换热器实际传热热阻往往与根据传热因子经验公式计算结果存在较大偏差。为此，利用换热器稳态性能试验数据，反向辨识其传热热阻经验公式，可以有效提高传热热阻的准确性。

3.1　简化关系式

对于板翅式换热器，传热 j 因子和传热系数之间的关系可以表示为

$$\alpha=j\frac{q_m}{A_d}c_pPr^{2/3} \tag{23}$$

式中：Pr 为普朗特数；A_d 为流动截面积；传热 j 因子通常可以表示为 $j=C\cdot Re^n$，C 和 n 为系数，雷诺数 Re 可表示为 $Re=q_md/(A_d\mu)$，d 为翅片当量直径，μ 为动力黏度。

因此，对特定板翅式换热器，其传热系数与质量流量间关系式可简化为

$$\alpha=k_1q_m^{n+1} \tag{24}$$

式中：k_1 为待定系数。

另外，板翅式换热器翅片表面效率计算式为

$$\eta_{fin}=\frac{\tanh(m'h')}{m'h'} \tag{25}$$

将其展开并忽略高阶项，可得

$$\eta_{fin}\approx\frac{m'h'-(m'h')^3/3}{m'h'}=1-\frac{1}{3}(m'h')^2 \tag{26}$$

式中：翅片参数 $m'=\sqrt{2\alpha/(\lambda_{fin}\delta_{fin})}$，$\lambda_{fin}$ 为翅片导热系数，δ_{fin} 为翅片厚度；h'为翅片参数。

翅片效率 η_o 可以由翅片表面效率 η_{fin}、一次传热面积 A_1(隔板面积)和二次传热面积 A_2(翅片表面积)计算得到，即

$$\eta_o=\frac{A_1+A_2\eta_{fin}}{A_1+A_2} \tag{27}$$

进而由式(24)、式(26)和式(27)可得

$$\eta_o\alpha A=\left(k_1q_m^{n+1}\right)A\left\{A_1+A_2\left[1-\frac{1}{3}\left(h'\sqrt{\frac{2\alpha}{\lambda_{fin}\delta_{fin}}}\right)^2\right]\right\}\Big/(A_1+A_2)=k_2q_m^{n+1}+k_3q_m^{2(n+1)} \tag{28}$$

则整个换热器的总传热热阻与冷热流体质量流量间计算关系式为

$$\frac{1}{UA}=\frac{1}{(\eta_o\alpha A)_h}+\frac{1}{(\eta_o\alpha A)_c}=\frac{1}{k_{h1}q_{m,h}^{n_h+1}+k_{h2}q_{m,h}^{2(n_h+1)}}+\frac{1}{k_{c1}q_{m,c}^{n_c+1}+k_{c2}q_{m,c}^{2(n_c+1)}} \tag{29}$$

式中：k_{h1}、k_{h2}、k_{c1}、k_{c2}、n_h 和 n_c 为待定系数。

3.2 辨识方法

换热器传热单元数 U_{NTU} 定义为

$$U_{NTU}=\frac{UA}{(q_mc_p)_{min}} \tag{30}$$

对于逆流式换热器，U_{NTU} 与效率间有如下关系：

$$U_{NTU}=\frac{1}{1-C^*}\ln\left(\frac{1-\eta C^*}{1-\eta}\right) \tag{31}$$

式中：$C^*=(q_mc_p)_{min}/(q_mc_p)_{max}$ 为冷热两侧的热容量之比。

利用不同质量流量下换热器的效率数据，即可建立换热器传热热阻与不同质量流量间的样本数据集合 $\left(q_{m,h}^{(i)},q_{m,c}^{(i)},(1/UA)^{(i)}\right)$。

以式(29)为传热热阻辨识表达式，以冷热流体质量流量为自变量，以传热热阻 $1/UA$ 为因变量，引入一定辨识方法确定待定系数 k_{h1}、k_{h2}、k_{c1}、k_{c2}、n_h、n_c 的值，即可获得传热热阻计算表达式，具体过程如图 5 所示。

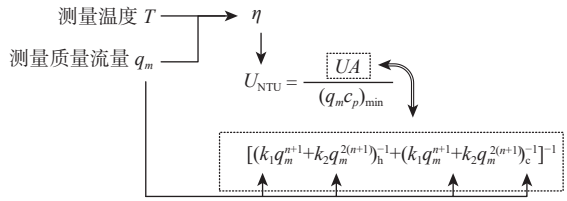


图 5 传热热阻计算过程示意图

Fig. 5 Schematic diagram of thermal resistance calculation process

4 传热热阻辨识及动态模型验证

以某空气-冷却水逆流板翅式换热器为研究对象，以 AMESim 仿真平台动态机理模型仿真结果为标准，对第 3 节传热热阻辨识方法及第 2 节简化动态仿真模型进行验证。

4.1 研究对象

某空气-冷却水逆流板翅式换热器主要结构参数如表 1 所示。该换热器热边工质为空气，冷边工质为水，翅片形式为锯齿形，材质为铝，长度为 600 mm，宽度为 300 mm。

表 1 换热器基本结构参数

Table 1 Basic parameters of heat exchanger

冷/热边	翅片间距/mm	翅片高度/mm	翅片厚度/mm	层数	隔板厚度/mm	面板厚度/mm	封条宽度/mm
热边	1.3	2.5	0.15	10	0.6	2	6
冷边	1.3	2.5	0.15	11	0.6	2	6

4.2 AMESim 标准机理模型

为提高模型的准确性，将换热器按照流体流动方向平均划分为 5 个节点[24]，在 AMESim 仿真平台搭建其机理仿真模型，如图 6 所示。其中，每个节点仿真模型主要由芯体和冷热流体流动传热 3 个仿真模块组成，分别对应式(1)～式(3) 3 个能量守恒方程的求解。

芯体仿真模块两端输入端口分别与冷热流体流动传热模块相连，主要利用冷热流体质量流量、比热容、温度和传热热阻，计算冷热流体换热量及芯体温度。而两侧冷热流体流动传热模块则利用该换热量及芯体温度，计算出口温度变化情况。

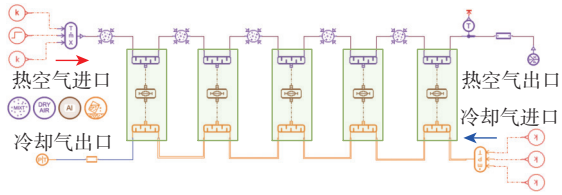


图 6 换热器五节点 AMESim 标准机理仿真模型

Fig. 6 Five-node standard mechanism model of heat exchanger in AMESim software

换热器两侧工质分别设定为空气和水，根据换热器基本结构参数计算每个节点 3 个仿真模块的输入参数，如表 2 所示。其中，芯体模块的质量参数及两侧流体流动传热模块的对流换热面积和容积分别为整个换热器的 1/5，对流换热量按给定传热因子经验公式进行计算。

4.3 传热热阻辨识

根据换热器设计工况，选取热侧空气质量流量范围为 1.00～3.50 kg/s，冷侧冷却水质量流量范围为 0.80～1.20 kg/s。利用 AMESim 仿真平台 Study Manager 工具箱进行拉丁超立方试验工况设计，并进行 200 次稳态仿真试验工况，获得换热器效率曲面，如图 7 所示。

表 2　AMESim 换热器仿真模块参数设置

Table 2　Parameter settings of heat exchanger module in AMESim software

模块	参数	数值
热侧流体流动传热模块	流动系数	0.6
	自由流动横截面积/mm^2	6 000
	换热特征长度/mm	1.5
	对流换热面积/m^2	9.304 6×1/5
	翅片面积与总面积之比	0.975 0
	翅片效率	0.962 8
	换热器体积容积/ L	3.7×1/5
	对流传热计算方式	Colburn j因子
	j因子表达式$j=f(Re)$	$0.144\,7Re^{-0.368}$
芯体模块	初始温度/℃	20
	流动形式	逆流
	芯体质量/kg	9.306 9×1/5
冷侧流体流动传热模块	流速计算	等效为孔板
	自由流横截面积/mm^2	6 600
	自由流动管径/mm	1.5
	最大流动系数	0.99
	临界雷诺数	1000 $\sqrt{2}$
	换热器特征长度/mm	1.5
	对流换热面积/m^2	10.235 1×1/5
	换热器内部容积/L	4.1×1/5
	对流传热计算方式	Colburn j因子
	j因子表达式$j=f(Re)$	$0.472\,2Re^{-0.536}$

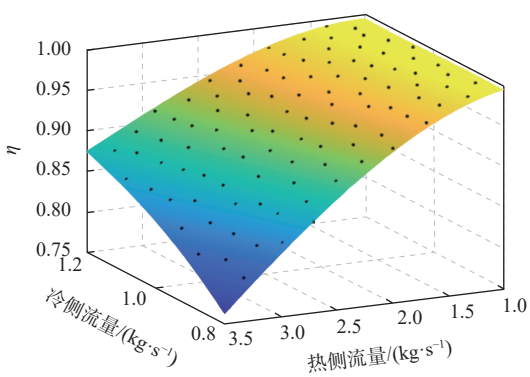


图 7　换热器效率曲面

Fig. 7　Efficiency map of heat exchanger

取上述换热器效率计算结果的 65% 作为样本数据, 根据图 5 所示辨识步骤, 利用 MATLAB 中的曲线拟合工具箱, 采用 Trust-Region-Reflective 算法对传热热阻计算关系式(29)中的待定系数进行拟合辨识。辨识误差平方和为 $1.256\,3\times10^{-5}$, 确定系数 R^2 为 0.999 9, 均方根标准差为 3.183×10^{-4}, 满足拟合精度要求。具体拟合公式为

$$\frac{1}{UA}=\frac{1}{1.557q_{m,\mathrm{h}}^{1-0.351\,6}-0.003\,5q_{m,\mathrm{h}}^{(1-0.351\,6)\times2}}+\frac{1}{12.40q_{m,\mathrm{c}}^{1-0.552\,5}-1.298q_{m,\mathrm{c}}^{(1-0.552\,5)\times2}} \tag{32}$$

进一步针对 35% 的非样本工况, 利用拟合公式获得换热器传热热阻与换热效率换算标准值对比, 如图 8 所示。可以看出, 两者能达到很好的吻合精度, 最大相对偏差为 0.95%, 满足工程设计精度要求, 也为动态仿真模型计算的准确性奠定了基础。

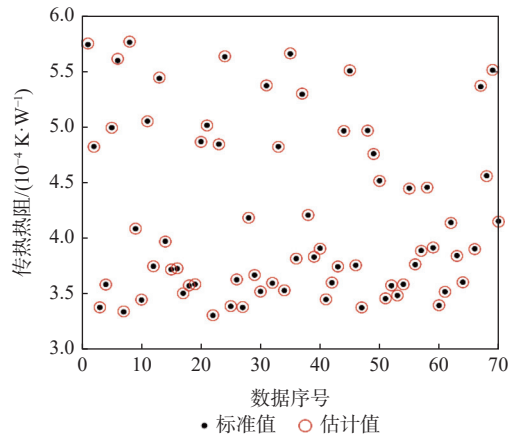


图 8　非样本工况下传热热阻拟合结果与标准对比

Fig. 8　Comparison of identification result of heat exchange resitance at non sample points

4.4　简化动态模型验证

为方便与测控系统联合仿真, 基于 Simulink 仿真平台, 利用可变传递函数仿真模块搭建换热器简化动态仿真模型, 如图 9 所示。图中, 稳态增益系数、响应时间常数及纯延迟时间等模型参数根据图 4 所示流程编写 M 函数进行计算, 传热热阻采用拟合辨识关系式(32)。

4 种仿真计算工况如表 3 所示。先达到某平衡状态, 再在第 120 s 时将空气或冷却水的入口温度或质量流量分别发生一定阶跃。

利用 AMESim 机理模型和 Simulink 简化动态仿真模型进行仿真计算, 获得 4 种工况下空气及冷却水出口温度动态响应对比曲线, 如图 10 和表 4 所示。

动态过程相对误差 ε_r 采用绝对误差积分平均值与阶跃前后稳态变化值之比来表示, 具体定义如下:

$$\varepsilon_\mathrm{r}=\frac{\int_{t_1=120}^{t_2=140}|e|\,\mathrm{d}t}{\Delta T(t_2-t_1)} \tag{33}$$

式中: e 为绝对偏差; ΔT 为阶跃前后稳态温度变化值。

从图 10 和表 4 可以看出:

1) 4 种工况下, 2 种仿真模型稳态计算结果吻

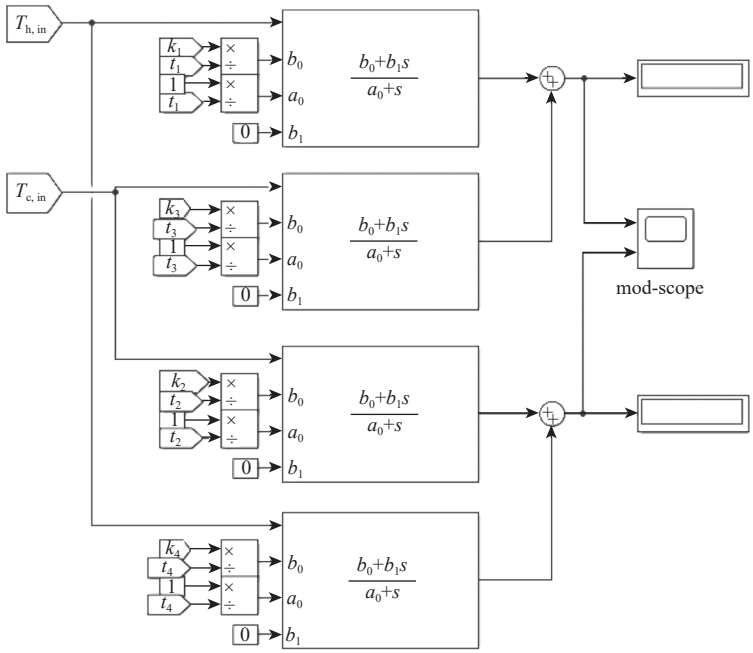


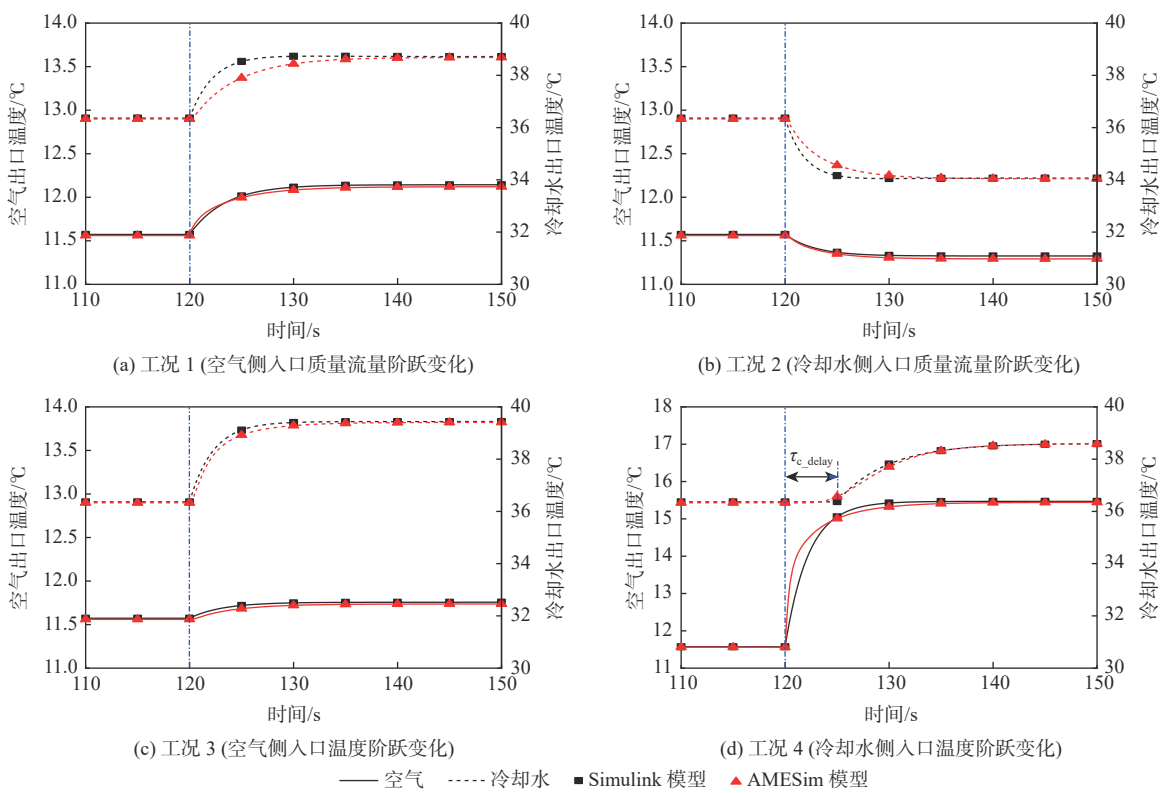
图 9 换热器 Simulink 简化动态仿真模型

Fig. 9 Simplified dynamic simulation module of heat exchanger in Simulink software

表 3 仿真工况参数

Table 3 Simulation condition parameters

工况	质量流量/(kg·s^{-1})		入口温度/℃	
	空气侧	冷却水侧	空气侧	冷却水侧
平衡	1.5	0.8	70	10
工况1	阶跃至1.65	0.8	70	10
工况2	1.5	阶跃至0.88	70	10
工况3	1.5	0.8	阶跃至77	10
工况4	1.5	0.8	70	阶跃至14



(a) 工况 1 (空气侧入口质量流量阶跃变化)

(b) 工况 2 (冷却水侧入口质量流量阶跃变化)

(c) 工况 3 (空气侧入口温度阶跃变化)

(d) 工况 4 (冷却水侧入口温度阶跃变化)

图 10 2 种仿真模型所得出口温度动态响应曲线

Fig. 10 Dynamic response curves of outlet temperature obtained by two simulation models

表 4　2 种仿真模型所得出口温度对比

Table 4　Comparison of outlet temperature obtained by two simulation models

工况	标准模型阶跃变化/℃		稳态偏差/℃		动态最大偏差/%		动态相对偏差/%	
	空气出口	冷却水出口	空气出口	冷却水出口	空气出口	冷却水出口	空气出口	冷却水出口
工况1	0.561	2.360	0.023	0.027	0.049	0.270	4.91	7.03
工况2	0.269	2.288	0.034	0.026	0.034	0.101	9.27	2.13
工况3	0.176	3.080	0.020	0.029	0.039	0.310	8.56	4.11
工况4	3.882	2.264	0.026	0.012	0.997	0.197	4.08	2.34

合很好。空气及冷却水出口温度最大稳态偏差分别为 0.034 ℃ 和 0.029 ℃, 主要来源于换热效率插值计算误差。说明该模型使用稳态换热效率确定增益系数, 能够很好地保证稳态结果的准确性。

2) 2 种仿真模型的动态响应过程也吻合得较好, 验证了简化动态仿真模型中时间常数确定方法的正确性。其中, 工况 2 冷却水入口质量流量阶跃变化时, 空气出口温度动态响应相对偏差最大为 9.27%; 工况 1 空气入口质量流量阶跃变化时, 冷却水出口温度动态响应相对偏差最大为 7.03%。偏差主要来源包括换热量采用代数平均温差进行计算, 时间常数的推导过程忽略了高阶项。

3) 2 种仿真模型计算结果中, 冷却水出口温度随入口温度阶跃变化的动态响应均向后延迟了 5 s, 空气侧由于流速较快延迟不明显, 进而验证了简化动态仿真模型中, 根据换热器长度与流速之比确定纯延迟时间的正确性。

5　结　论

本文围绕飞机综合环控/热管理系统中典型板翅式换热器动态仿真模型建模方法, 得到如下结论:

1) 提出一种由 2 个延迟环节和 4 个一阶惯性环节组成的传递函数简化模型, 计算效率高, 可用于流量控制系统联合仿真和优化设计。

2) 基于换热器机理模型拉普拉斯变换和反变换推导, 给出了简化动态仿真模型中时间常数的计算公式, 同时给出了基于换热器效率进行传热热阻计算公式的辨识方法, 解决了换热器动态仿真模型中 2 个主要模型参数的设置问题。

3) 以 AMESim 仿真平台搭建的某空-液板翅式换热器详细动态模型仿真结果为标准, 对在 Simulink 仿真平台搭建的新型动态仿真模型进行验证, 空气及冷却水出口温度随两侧入口温度和质量流量阶跃变化的动态响应过程及稳态结果均能很好地吻合, 满足工程精度要求。

参考文献 (References)

[1] WANG J X, LI Y Z, LIU X D, et al. Recent active thermal management technologies for the development of energy-optimized aerospace vehicles in China[J]. Chinese Journal of Aeronautics, 2021, 34(2): 1-27.

[2] DEPPEN T O, HEY J E, ALLEYNE A G, et al. A model predictive framework for thermal management of aircraft[C]//Proceedings of the ASME 8th Annual Dynamic Systems and Control Conference. New York: ASME, 2016.

[3] ROETZEL W, XUAN Y. Dynamic behaviour of heat exchangers [M]. Southampton: WIT Press/Computational Mechanics Publications, 1999.

[4] 冷伟, 房德山, 徐治皋, 等. 对单相换热器集总参数模型动态初始负偏移的机理分析[J]. 热能动力工程, 2001, 16(3): 287-289.
LENG W, FANG D S, XU Z G, et al. An analysis of the mechanism governing the dynamic and initial negative deviation of a lumped parameter model for a single-phase heat exchanger[J]. Journal of Engineering for Thermal Energy and Power, 2001, 16(3): 287-289(in Chinese).

[5] 宋俊虎, 袁修干, 林贵平. 某种叉流板翅式热交换器动态性能的计算[J]. 北京航空航天大学学报, 1999, 25(5): 558-560.
SONG J X, YUAN X G, LIN G P. Transient performances calculation of a gas-to-gas crossflow heat exchanger[J]. Journal of Beijing University of Aeronautics and Astronautics, 1999, 25(5): 558-560 (in Chinese).

[6] 解增忠, 张俊峰, 罗雄麟. 原油换热网络的动态建模与仿真[C]//第九届全国化学工艺年会. 北京: 中国石化出版社, 2005: 1198-1206.
XIE Z Z, ZHANG J F, LUO X L. Dynamic modeling and simulation of crude oil heat exchanger networks[C]//Proceedings of the 9th National Chemical Technology Annual Meeting. Beijing: China Petrochemical Press, 2005: 1198-1206(in Chinese).

[7] ORAVEC J, BAKOŠOVÁ M, TRAFCZYNSKI M, et al. Robust model predictive control and PID control of shell-and-tube heat exchangers[J]. Energy, 2018, 159: 1-10.

[8] CHEN M H, SUN X D, CHRISTENSEN R N, et al. Dynamic behavior of a high-temperature printed circuit heat exchanger: Numerical modeling and experimental investigation[J]. Applied Thermal Engineering, 2018, 135: 246-256.

[9] HEY J E, HODSON S L, YAZAWA K, et al. Experimental characterization of dynamic heat exchanger behavior[J]. International Journal of Heat and Mass Transfer, 2018, 121: 933-942.

[10] SANGI R, MÜLLER D. Dynamic modelling and simulation of a slinky-coil horizontal ground heat exchanger using Modelica[J]. Journal of Building Engineering, 2018, 16: 159-168.

[11] NASH A L, BADITHELA A, JAIN N. Dynamic modeling of a sensible thermal energy storage tank with an immersed coil heat exchanger under three operation modes[J]. Applied Energy, 2017,

195: 877-889.

[12] AGARWAL A, B M O, PITSO I. Numerical analysis and performance enhancement of compact heat exchanger using computational fluid dynamics[J]. Journal of Engineering Research, 2021, 9: 15503.

[13] GONG M J, PENG M J, ZHU H S. Research of multiple refined degree simulating and modeling for high pressure feed water heat exchanger in nuclear power plant[J]. Applied Thermal Engineering, 2018, 140: 190-207.

[14] FRATCZAK M, NOWAK P, CZECZOT J, et al. Simplified dynamical input-output modeling of plate heat exchangers-case study[J]. Applied Thermal Engineering, 2016, 98: 880-893.

[15] WANG Y R, YOU S J, ZHENG W D, et al. State space model and robust control of plate heat exchanger for dynamic performance improvement[J]. Applied Thermal Engineering, 2018, 128: 1588-1604.

[16] ROMIE F E. Transient response of gas-to-gas crossflow heat exchangers with neither gas mixed[J]. Journal of Heat Transfer, 1983, 105(3): 563-570.

[17] WHALLEY R, EBRAHIMI K M. Heat exchanger dynamic analysis [J]. Applied Mathematical Modelling, 2018, 62: 38-50.

[18] GVOZDENAC D. Analytical solution of dynamic response of heat exchanger[M]//MITROVIĆ J. Heat exchangers: basics design applications. London: Intechopen, 2012: 53-78.

[19] AL-DAWERY S K, ALRAHAWI A M, AL-ZOBAI K M. Dynamic modeling and control of plate heat exchanger[J]. International Journal of Heat and Mass Transfer, 2012, 55(23-24): 6873-6880.

[20] ZAHID K, PATEL R, MUJTABA I. Development of a dynamic fouling model for a heat exchanger[J]. Chemical Engineering Transactions, 2016, 52: 1135-1140.

[21] YIN J, JENSEN M K. Analytic model for transient heat exchanger response[J]. International Journal of Heat and Mass Transfer, 2003, 46(17): 3255-3264.

[22] GAO T Y, GEER J, SAMMAKIA B. Development and verification of compact transient heat exchanger models using transient effectiveness methodologies[J]. International Journal of Heat and Mass Transfer, 2015, 87: 265-278.

[23] PEARSON J. EASY5 user guide[Z]. Chicago: Boeing Company, 2005: 265-278.

[24] 张一鸣, 袁宁, 陈玉春. 回热循环微燃机中换热器计算模型与仿真研究[J]. 燃气轮机技术, 2012, 25(1): 38-43.
ZHANG Y M, YUAN N, CHEN Y C. Studies on calculation model and simulation of the heat exchanger in micro-turbine[J]. Gas Turbine Technology, 2012, 25(1): 38-43(in Chinese).

Dynamic model by transfer function and parameter determination method of plate fin heat exchanger

LIANG Xingzhuang[1, 2, *], HUANG Zhiyuan[2], AI Fengming[2], YUAN Zhenwei[2], WANG Jian[1]

(1. State Key Laboratory of Fire Science, University of Science and Technology of China, Hefei 230026, China;
2. AVIC Shenyang Aircraft Design and Research Institute, Shenyang 110035, China)

Abstract: A fast and efficient dynamic simulation model of heat exchangers is urgently needed to meet the requirements of the advanced control system design for the integrated environmental control and thermal management system of aircraft. In this paper, a dynamic model by transfer function matrix, composed of two delay links and four first-order inertia links, is proposed for plate-fin heat exchangers. The identification method for the calculation formulas of thermal resistance based on the efficiency of the heat exchanger is given, which solves the issue of setting two key parameters in the model. The calculation formulas of four-time constants in the simplified model are derived according to the mechanism of the heat exchanger by using the Laplace transform. Taking an air-liquid plate-fin heat exchanger as the research object, the new dynamic simulation model is built in Simulink software, and compared with the mechanism model built in AMESim software. The results show that both models have similar dynamic responses of the outlet temperature at both flow sides with step change of inlet temperature or flow rate. Under four working conditions, the highest steady-state variations of the air and cooling water outlet temperatures are 0.034 ℃ and 0.029 ℃, respectively. The maximum dynamic relative deviation of airflow outlet temperature is 9.27% with a step change of the cooling water flow rate. The maximum dynamic relative deviation of cooling water flow outlet temperature is 7.03% with a step change of the airflow rate.

Keywords: heat exchanger; dynamic model; transfer function matrix; time constant; thermal resistance

Received: 2022-04-01; **Accepted**: 2022-04-23; **Published Online**: 2022-05-10 17:33
URL: link.cnki.net/urlid/11.2625.V.20220509.1528.010

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0252

电动静液作动器的非线性变阻尼积分滑模控制

杨荣荣[1,*]，张玲[1]，赵家黎[1]，付永领[2]，张朋[3]

(1. 兰州理工大学 机电工程学院，兰州 730050；2. 北京航空航天大学 机械工程及自动化学院，北京 100191；
3. 北京精密机电控制设备研究所，北京 100076)

摘　　要：对于电动静液作动器(EHA)，传统滑模控制器存在加速度信息难以获取，参数不易整定和控制信号抖振等问题，从而造成控制器很难应用于实际。针对以上问题，利用奇异摄动理论对EHA数学模型进行合理的降阶，从而使控制器设计避免了使用加速度信息。在此基础上，利用降阶模型设计了一种新型非线性变阻尼积分滑模控制器(NSMC)，该控制器可根据位置控制误差实现系统阻尼比由欠阻尼到过阻尼的自适应调节，能有效提高位置阶跃调节性能。设计了一种基于滤波器的不确定项估计器对EHA中存在的参数不确定性和外部扰动进行实时估计并补偿。滑模面积分项的引入和不确定项估计器的使用，一方面使控制器中无需使用切换函数，实现了EHA的无抖振滑模控制，另一方面使系统整个动态过程完全表现为滑动模态，从而可根据EHA控制指标直接整定滑模面参数，大大简化了参数整定过程。同时利用Lyapunov稳定性理论对整个闭环系统和滑模面的稳定性进行了详细分析。分别与PI控制器、传统滑模控制器(SMC)和传统变阻尼滑模控制器(DVSMC)进行了详细的仿真分析比较，仿真结果表明NSMC能有效提高EHA位置跟踪性能和增强对参数不确定性和外部扰动的鲁棒性。

关 键 词：电动静液作动器；奇异摄动；变阻尼；积分滑模控制；不确定项估计器
中图分类号：TH137
文献标志码：A　　**文章编号**：1001-5965(2024)01-0163-10

电动静液作动器(electro-hydrostatic actuator, EHA)具有集成度高、功重比大、可靠性高、效率高和安装维护性好等优点，是未来多电和全电飞机机载作动系统的重要组成部分，其作动性能直接影响飞机的安全和机动性能[1-3]。但EHA是一种机电液耦合的高阶复杂非线性系统，且系统中存在大量参数不确定性和外部负载扰动，因此要实现EHA的高性能位置控制往往比较困难。

为此，已有多种先进控制策略被用于提升和改进EHA的鲁棒性和位置作动性能，如串级比例积分微分(proportional integral and derivative, PID)控制[4-5]、自适应反步控制[6]、自适应鲁棒控制[7]、误差符号函数积分鲁棒控制[8]、自抗扰控制[9]、H_∞控制[10]和滑模控制[11-13]等。上述方法中，滑模控制是一种成熟且系统化的设计方法，适用于线性和非线性系统，对于系统参数不确定性和外部扰动都具有很好的鲁棒性，且设计步骤简单明确，有严格的证明过程和理论支撑，因此得到了广泛应用。文献[14]提出了一种基于变结构滤波器的复合滑模控制器，其中变结构滤波器用于估计未知参数和非可测状态，而滑模控制器用于补偿EHA系统中的各种不确定性。文献[15]利用线性二次型方法设计滑模面，并用滑模控制器处理EHA中非线性摩擦。文献[16]设计了一种非线性扰动观测器，用以实时估计

收稿日期：2022-04-18；录用日期：2022-06-10；网络出版时间：2022-06-14 11:09
网络出版地址：link.cnki.net/urlid/11.2625.V.20220613.1814.004
基金项目：甘肃省科技计划项目(20JR10RA175)
*通信作者. E-mail：yr13236@163.com

引用格式：杨荣荣，张玲，赵家黎，等. 电动静液作动器的非线性变阻尼积分滑模控制[J]. 北京航空航天大学学报，2024，50(1)：163-172.
YANG R R，ZHANG L，ZHAO J L，et al. Nonlinear variable damping integral sliding mode control for electro-hydrostatic actuator[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(1)：163-172 (in Chinese).

EHA中的不确定性，在一定程度上减小了抖振。文献[17]提出了一种抗积分饱和的自适应滑模控制器，该方法实现了切换函数sign的增益自适应调节，但并未考虑增益自适应的过估计问题。文献[18]利用最优理论设计滑模面，采用模糊规则调节切换增益。文献[19]针对EHA提出了一种全新的滑模控制架构，该方法将EHA的压力和电机泵的动态模型看作一个整体设计滑模面，从而有效的降低了控制器设计复杂度，同时通过设计新型滑模观测器和趋近律提高鲁棒性和抑制抖振，但该趋近律过于复杂，需整定的参数较多，难以应用于实际。为了进一步提高EHA的位置阶跃性能，文献[20]提出了一种变阻尼滑模控制策略，该方法在到达滑模面后可实现系统阻尼比的自适应调节，能在缩小调节时间的同时实现无超调调节。文献[21-22]在此基础上，融合自适应控制的思想，进一步扩展了该方法的应用。但文献[21-22]中控制器设计需要作动加速度信息，而实际中加速度是不可量测的，通常通过对位置进行二次微分获取，但考虑到位置传感器的量测噪声，实际上该方法很难获取有效的加速度信息，相反还会引起控制信号的强烈振荡，甚至造成系统失稳。而且文献[21-22]也未给出变阻尼滑模面的稳定性证明，控制器中仍然使用了sign函数，抖振问题依然存在。

针对以上问题，本文在文献[20-21]的基础上提出了一种新型非线性变阻尼积分位置滑模控制器，该方法具有以下优点：①利用奇异摄动理论对EHA模型进行降阶处理，滑模面设计无需加速度信息；②滑模面引入变阻尼非线性积分项，积分项的引入一方面使控制器对常值低频不确定性具有强鲁棒性，另一方面通过设定积分初值可使滑模变量在初始时刻便处于滑模面，控制器具有全局鲁棒性；③非线性变阻尼项的引入可实现系统阻尼比的自适应调节，在缩小调节时间的同时实现了EHA位置阶跃响应的无超调调节；④设计了一种整定参数仅有一个的不确定项估计器，对EHA中的不确定性实时估计并补偿，从而滑模控制器中无需使用切换函数sign，抖振问题得到有效解决；⑤滑模面积分项和不确定项估计器的引入，使系统响应与滑模变量的到达阶段无关，系统响应完全表现为滑动模态，因此滑模面参数可根据EHA控制性能指标直接确定，避免了传统滑模控制中滑模面参数整定的盲目性，使控制器更易被应用于实际。

1 EHA数学模型

EHA的基本原理如图1所示，主要由集成式电液泵(integrated electrohydraulic pump，IEHP)、阀、蓄能器、双出杆液压缸、传感器件和驱动控制器等组成。其中IEHP是一种新型的液压动力单元，其将永磁同步电动机与柱塞式液压泵高度融合，两者共转子、共壳体，相较于传统的电机泵组，在体积、噪声和效率等方面都具有相当明显优势[23]。

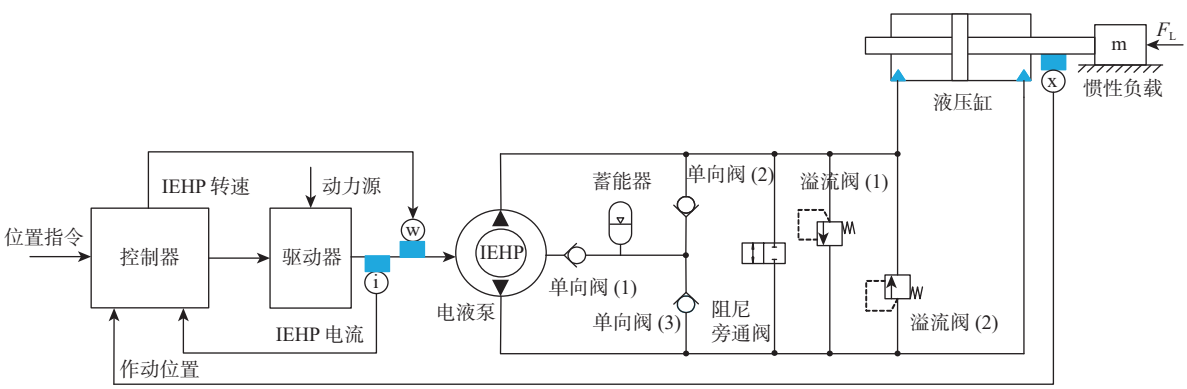


图1 EHA原理图

Fig. 1 Schematic diagram of EHA

1.1 IEHP电压和运动方程

尽管IEHP定转子之间为油隙而非气隙，但油隙并不改变转子永磁体产生的磁密形状，因此仍然可以假设磁密形状正弦状，定子绕组反电动势波形也为正弦，转子永磁体为表贴式。旋转dq坐标系中的电压方程可写为

$$\begin{cases} u_d = R_h i_d + L_d \dot{i}_d - p\omega_m L_q i_q \\ u_q = R_h i_q + L_q \dot{i}_q + p\omega_m (L_d i_d + \psi_f) \end{cases} \tag{1}$$

式中：u_d和u_q分别为d和q轴电压分量；i_d和i_q分别为d和q轴电流分量；L_d和L_q分别为d和q轴等效电感；R_h为定子电阻；ψ_f转子永磁体磁链；p为极对数；ω_m为机械角速度。

IEHP电磁转矩T_e可表示为

$$T_e = 1.5 p \psi_f i_q \tag{2}$$

IEHP定转子之间为油隙，因此需要考虑转子

旋转时与油液产生的摩擦转矩, 根据 Pertoff 定理[24], 摩擦转矩 T_{of} 为

$$T_{of}=2\pi\mu Lr^3\omega_m/\kappa=B_{og}\omega_m \tag{3}$$

式中: B_{og} 为油隙摩擦系数; μ 为液压油液的动力黏度; L 为 IEHP 转子长度; r 为 IEHP 转子半径; κ 为油隙厚度。

IEHP 三大运动副产生的摩擦转矩: $T_f=B_m\omega_m$, 其中 B_m 为黏性摩擦系数, 故 IEHP 运动方程式可写为

$$J_p\dot{\omega}_m=T_e-B_{tp}\omega_m-D_p(p_a-p_b) \tag{4}$$

式中: J_p 为 IEHP 转动惯量; $B_{tp}=B_{og}+B_m$ 为总黏性摩擦系数; D_p 为 IEHP 排量; p_a 和 p_b 分别为 IEHP 的出油口和回油口压力。

1.2　IEHP 流量方程

IEHP 流量方程为

$$\begin{cases}Q_a=D_p\omega_m-C_{ip}(p_a-p_b)-V_a\beta_e^{-1}\dot{p}_a-Q_{opa}\\ Q_b=D_p\omega_m-C_{ip}(p_a-p_b)+V_b\beta_e^{-1}\dot{p}_b+Q_{opb}\end{cases} \tag{5}$$

式中: Q_a 和 Q_b 分别为 IEHP 的出油口和进油口流量; C_{ip} 为 IEHP 的内泄漏流量系数; Q_{opa} 和 Q_{opb} 分别为 IEHP 两腔外泄漏流量; β_e 为油液弹性模量; V_a 和 V_b 分别为 IEHP 出油腔和回油腔体积。

1.3　液压缸流量方程

忽略液压缸外泄漏流量, 液压缸流量方程可写为

$$\begin{cases}Q_1=A\dot{x}+\beta_e^{-1}(V_{10}+Ax)\dot{p}_1+C_{ic}(p_1-p_2)\\ Q_2=A\dot{x}-\beta_e^{-1}(V_{20}-Ax)\dot{p}_2+C_{ic}(p_1-p_2)\end{cases} \tag{6}$$

式中: Q_1 和 Q_2 分别为液压缸两腔的流入和流出流量; A 为液压缸活塞有效作用面积; x 为液压缸活塞杆位移; p_1 和 p_2 分别为液压缸两腔压力; C_{ic} 为液压缸内泄漏系数; V_{10} 和 V_{20} 分别为液压缸两侧密闭容腔容积。

1.4　液压缸运动方程

假设活塞杆和负载为刚性连接, 则液压缸运动方程可表示为

$$m_t\ddot{x}=A(p_1-p_2)-B_t\dot{x}-F_f-F_L \tag{7}$$

式中: m_t 为活塞、活塞杆和负载总质量; B_t 为液压缸和负载总黏性摩擦系数; F_f 为其他非线性摩擦力; F_L 为施加在活塞杆上的外负载力。

1.5　压力动态方程

根据流量连续性特性, 有关系式:

$$\begin{cases}Q_a=Q_1+Q_{c1}-Q_{r1}\\ Q_b=Q_2+Q_{c2}-Q_{r2}\end{cases} \tag{8}$$

式中: Q_{c1} 和 Q_{c2} 分别为通过 2 个单向阀的流量; Q_{r1} 和 Q_{r2} 分别为通过 2 个溢流阀的流量。

忽略 IEHP 到液压缸管路的压力损失, 则有 $p_a=p_1, p_b=p_2$。此外 V_a 和 V_b 满足 $V_a=V_{10}+Ax, V_b=V_{20}-Ax$。同时考虑到液压缸为对称缸, 因此有近似关系式 $\dot{p}_1\approx-\dot{p}_2$, 并且假设初始时活塞处于液压缸中位, 则 $V_{10}\approx V_{20}=V_0$, V_0 为液压缸左右两腔的初始容积。再考虑式 (6) 和式 (8) 可得

$$\dot{p}_L=\beta_eV_0^{-1}(D\omega_m-A\dot{x}-C_tp_L+Q_{un}) \tag{9}$$

式中: 负载压力 $p_L=p_1-p_2$; 单向阀、溢流阀流量损耗及 IEHP 等效外泄漏量 $Q_{un}=(Q_{r1}+Q_{c2}-Q_{c1}-Q_{r2})/2-\left(Q_{opa}-Q_{opb}\right)/2$; 液压缸和 IEHP 总泄漏系数 $C_t=C_{ip}+C_{ic}$。

选取状态变量 $\boldsymbol{X}=\begin{bmatrix}x_1 & x_2 & x_3 & x_4 & x_5 & x_6\end{bmatrix}^T=\begin{bmatrix}x & \dot{x} & p_L & \omega_m & i_q & i_d\end{bmatrix}^T$, 则系统状态方程可写为

$$\begin{cases}\dot{x}_1=x_2\\ \dot{x}_2=m_t^{-1}[Ax_3-B_tx_2+d_{Lm}]\\ \dot{x}_3=\beta_eV_0^{-1}\left(D_px_4-Ax_2-C_tx_3+Q_{un}\right)\\ \dot{x}_4=J_p^{-1}\left(3/2\cdot p\psi_fx_5-B_{tp}x_4+d_{Lp}\right)\\ \dot{x}_5=L_q^{-1}\left[-R_hx_5-px_4(L_dx_6+\psi_f)+u_q\right]\\ \dot{x}_6=L_d^{-1}\left(-R_hx_6+px_4L_qx_5+u_d\right)\end{cases} \tag{10}$$

式中: $d_{Lm}=-F_f-F_L$; $d_{Lp}=-D_pp_L$。

2　基于奇异摄动理论的降阶模型

将 u_d 和 u_q 看作 EHA 最终控制输入量, 由状态方程可以看出 EHA 为 6 阶 2 输入系统。为了便于控制器设计, 利用串级控制的思想, 将状态变量 x、$\dot{x}$ 和 p_L 构成 EHA 位置子系统, 状态变量 ω_m 看作 IEHP 转速子系统, 而 i_d 和 i_q 分别构成 IEHP 的 d 轴和 q 轴电流子系统。对于位置子系统, 如果直接选择状态变量 x、$\dot{x}$ 和 p_L 设计滑模控制器, 由于 x_2 动态方程中的 d_{Lm} 为非匹配型不确定项, 而传统滑模控制器对非匹配型不确定项不具备鲁棒性。常用的方法是通过坐标变换, 将状态变量转换为 x、$\dot{x}$ 和 $\ddot{x}$ 设计滑模控制器, 这样尽管可以将 d_{Lm} 转换为匹配型, 但此时控制器设计需要加速度 $\ddot{x}$ 信息, 而实际中 $\ddot{x}$ 很难准确获得。

重写电静液作动系统流量连续性方程 (8) 为

$$\dot{p}_L=\beta_eV_0^{-1}(D_p\omega_m-A\dot{x}-C_tp_L+Q_{un}) \tag{11}$$

注意到油液弹性模量 β_e 是一个数值非常大的物理量, 通常为 700～1 500 MPa, 因此有 $0<\beta_e^{-1}\ll1$, $0<\beta_e^{-1}V_0\ll1$, β_e^{-1} 或 $\beta_e^{-1}V_0$ 均可作为奇异摄动参数, 本文选奇异摄动参数 $\varepsilon_s=\beta_e^{-1}V_0$。定义 $\boldsymbol{X}_s=\begin{bmatrix}x & \dot{x}\end{bmatrix}^T$, $Z_s=p_L$, 将 EHA 压力动态方程和活塞杆运动方程写为奇异摄动标准形式:

$$\begin{cases}\ddot{x}=m_t^{-1}Ap_L-m_t^{-1}B_t\dot{x}+m_t^{-1}d_{Lm}(t)\\ \varepsilon_s\dot{p}_L=D_p\omega_m-A\dot{x}-C_tp_L+Q_{un}\end{cases} \tag{12}$$

令$\varepsilon_s=0$，可求得p_L的准稳态量$\bar{p}_L$为

$$\bar{p}_L=C_t^{-1}(D_p\omega_m-A\dot{x}+Q_{un}) \tag{13}$$

定义$y_\tau=p_L-\bar{p}_L$，则

$$\begin{aligned}\varepsilon_s\dot{y}_\tau&=\varepsilon_s\dot{p}_L-\varepsilon_s\dot{\bar{p}}_L=\\&D_p\omega_m-A\dot{x}-C_tp_L+Q_{un}-\varepsilon_s\dot{p}_L=\\&C_t\bar{p}_L-C_tp_L-\varepsilon_s\dot{p}_L=-C_ty_\tau-\varepsilon_s\dot{p}_L\end{aligned} \tag{14}$$

设$\varepsilon_s=0$，用变量t表示时间变量，同时定义新的时间尺度变量$\tau_s=t/\varepsilon_s$，将式(14)转换到新的时间尺度τ_s框架下，则边界层模型方程为

$$\mathrm{d}y_\tau/\mathrm{d}\tau_s=-C_ty_\tau \tag{15}$$

显然，边界层模型关于平衡点$y_\tau=0$是渐进稳定的。将准稳态$\bar{p}_L$代入式(12)得降阶模型为

$$\begin{aligned}\ddot{x}_p&=m_t^{-1}A\bar{p}_L-m_t^{-1}B_t\dot{x}_p+m_t^{-1}d_{Lm}(t)=\\&g\omega_m+\varphi\dot{x}_p+h=\\&g_0\omega_m+\varphi_0\dot{x}_p+\Delta g\omega_m+\Delta\varphi\dot{x}_p+h=\\&g_0\omega_m+\varphi_0\dot{x}_p+D\end{aligned} \tag{16}$$

式中：x_p为模型降阶后的位置；$\dot{x}_p$为降阶后的速度；$\ddot{x}_p$为降阶后的加速度；$g=g_0+\Delta g=AD_p(m_tC_t)^{-1}$；$\varphi=\varphi_0+\Delta\varphi=-\left[A^2(m_tC_t)^{-1}+B_tm_t^{-1}\right]$；$h=A(m_tC_t)^{-1}Q_{un}+m_t^{-1}d_{Lm}$；$D=\Delta g\omega_m+\Delta\varphi\dot{x}_p+h$为等效的总不确定性；$g_0$和$\varphi_0$分别为$g$和$\varphi$的标称量；$\Delta g$和$\Delta\varphi$分别为$g$和$\varphi$的摄动量。

由Tikhonov定理[25]可知，当ε_s足够小时，x_p、$\dot{x}_p$和$\bar{p}_L$将无限接近于原始状态x、$\dot{x}$和p_L。由式(16)可以看出位置子系统模型由原来的3阶降为2阶，且降阶模型与参数β_e和V_0无关。很明显，位置子系统经过模型降阶处理后，非匹配不确定项d_{Lm}转换为匹配型，从而大大简化了控制器的设计过程，同时滑模控制器设计无需加速度$\ddot{x}$信息，因此控制器更易被应用于实践。

3 非线性变阻尼积分滑模控制器

本文设计目标是为EHA位置子系统设计一种新型非线性变阻尼积分滑模控制器(nonlinear variable damping integral sliding mode controller, NSMC)，在系统存在大量系统不确定性时，作动位置x仍能准确地跟踪给定位置指令x_d。对于位置阶跃响应，在实现无超调调节的同时进一步减小调节时间。此外尽可能的简化控制器的设计复杂度和参数整定难度，从而使其更易应用于实际。IEHP转速和电流子系统仍采用传统比例积分(proportional and integral, PI)控制器。

3.1 非线性变阻尼积分滑模控制器设计

将IEHP转速ω_m看作位置子系统的控制量，即IEHP转速子系统的给定量。定义EHA位置跟踪误差：$e=x_d-x_p$。利用位置子系统降阶模型(16)，设计积分滑模面为

$$s=\dot{e}+\int\left[2\omega_nf(e)\dot{e}+\omega_n^2e\right]\mathrm{d}t \tag{17}$$

式中：s为滑模变量；$f(e)=(\xi_\infty-\xi_0)/\left(1+\delta e^2\right)+\xi_0$为变阻尼函数，$\xi_0$为初态阻尼比，$\xi_\infty$为终态阻尼比；$\omega_n$为固有频率；$\delta$为变阻尼敏感系数。

对s关于时间t求导可得

$$\begin{aligned}\dot{s}&=\ddot{e}+2\omega_nf(e)\dot{e}+\omega_n^2e=\\&\ddot{x}_d-\ddot{x}_p+2\omega_nf(e)\dot{e}+\omega_n^2e=\\&\ddot{x}_d-g_0\omega_m-\varphi_0\dot{x}_p-D+2\omega_nf(e)\dot{e}+\omega_n^2e\end{aligned} \tag{18}$$

设计控制量ω_m为

$$\omega_m=g_0^{-1}\left(\ddot{x}_d-\varphi_0\dot{x}_p+2\omega_nf(e)\dot{e}+\omega_n^2e\right)+g_0^{-1}u_L \tag{19}$$

式中：u_L为不确定性补偿控制分量。

将式(19)代入$\dot{s}$可得

$$\dot{s}=-u_L-D \tag{20}$$

3.2 不确定项估计器设计

本节将通过设计不确定项估计器对总不确定性D进行实时估计并补偿以提高系统的抗扰动能力。对式(20)两边同时进行滤波处理可得：

$$\dot{s}_f=-u_{Lf}-D_f \tag{21}$$

式中：s_f、u_{Lf}和D_f分别满足$\tau\dot{s}_f+s_f=s$，初值满足$s_f(0)=0$；$\tau\dot{u}_{Lf}+u_{Lf}=u_L,u_{Lf}(0)=0$；$\tau\dot{D}_f+D_f=D,D_f(0)=0$；$\tau$为滤波时间常数。

设计不确定性D的估计值$\hat{D}$为

$$\hat{D}=-u_{Lf}-(s-s_f)/\tau \tag{22}$$

定义估计误差：$\tilde{D}=D-\hat{D}$，其动态方程为

$$\begin{aligned}\dot{\tilde{D}}&=\dot{\hat{D}}-\dot{D}=-\dot{u}_{Lf}-(\dot{s}-\dot{s}_f)/\tau-\dot{D}=\\&-(u_L-u_{Lf})/\tau-(-u_L-D+u_{Lf}+D_f)/\tau-\dot{D}=\\&(D_f-D)/\tau-\dot{D}=-\tilde{D}/\tau-\dot{D}\end{aligned} \tag{23}$$

假设$\dot{D}$满足$\left|\dot{D}\right|<D_L$。注意到不确定性$D=\Delta g\omega_m+\Delta\varphi\dot{x}_p+h$，考虑到$\Delta g$和$\Delta\varphi$均为慢变摄动量，故$\dot{D}=\Delta g\dot{\omega}_m+\Delta\varphi\ddot{x}_p+\dot{h}$，其中，IEHP的角加速度$\dot{\omega}_m$和EHA的作动加速度$\ddot{x}$在际中必是有限的，此外，未知参数摄动值$\Delta g$和$\Delta\varphi$也为有限值，而$\dot{h}$与外负载变化量$\dot{F}_L$有关，而实际中EHA所承受的外负载也是在一定范围内变化，因此$\dot{F}_L$也有界，故上述假设是合理的。

求解式(23)所示微分方程，可得

$$\tilde{D}=\exp(-t/\tau)\tilde{D}(0)-\tau\dot{D}\left[1-\exp(-t/\tau)\right]\leqslant\tau|D_L| \tag{24}$$

由式(24)可见，当$t\to+\infty$时，$\tilde{D}\to-\tau\dot{D}\leqslant\tau\left|\dot{D}\right|$，而$\dot{D}$满足：$\left|\dot{D}\right|<D_L$，因此$\tilde{D}\leqslant\tau D_L$。而且$\tau$越小，$\tilde{D}$越

小, 即$\hat{D}$对D的估计越准确。

设计$u_{\mathrm{L}}=-\hat{D}$, 并将其代入$\dot{s}$可得

$$\dot{s}=-u_{\mathrm{L}}-D=\hat{D}-D=-\tilde{D} \tag{25}$$

传统滑模控制器中引入切换函数 sign 的目的有 2 个: ①确保滑模变量s由初值$s(0)$能够到达滑模面; ②补偿系统中的不确定性。切换增益越大, 到达滑模面的速度越快, 对不确定性的鲁棒性也越强, 但同时抖振现象也越严重。本文设计的滑模面式 (17) 中引入了积分项, 如果取积分项的初值为$-\dot{e}(0)$, 就可使$s(0)=0$, 这意味着滑模变量在初始时刻便在滑模面上, 而且积分项使系统对于常值和低频不确定性本身就具有强鲁棒性。而不确定项估计器又实现了对不确定性的实时准确估计和补偿, 因此控制器中无需使用 sign 函数, 从而从根本上避免了抖振现象。为了进一步抑制估计误差$\tilde{D}$, u_{L}中加入鲁棒项ks, 其中k为大于零的常数。故u_{L}最终可修正为

$$u_{\mathrm{L}}=-\hat{D}+ks \tag{26}$$

最终控制量ω_{m}为

$$\omega_{\mathrm{m}}=g_0^{-1}\left(\ddot{x}_d-\varphi_0\dot{x}_{\mathrm{p}}+2\omega_{\mathrm{n}}f(e)\dot{e}+\omega_{\mathrm{n}}^2e\right)+g_0^{-1}\left(-\hat{D}+ks\right) \tag{27}$$

当滑模变量s到达滑模面$s=0$后, 对其关于时间t求导, 可以得到滑动模态为

$$\ddot{e}+2\omega_{\mathrm{n}}f(e)\dot{e}+\omega_{\mathrm{n}}^2e=0 \tag{28}$$

由式(28)可见, 滑动模态为典型的二阶系统, 而二阶系统的响应特性由阻尼比$f(e)$和固有频率ω_{n}决定。

传统滑模控制中滑模变量s的动态过程分为到达阶段和滑动模态 2 个阶段, 而本文设计的控制器由于滑模面积分项和不确定项估计器的引入, 使滑模变量的整个动态过程省去了到达阶段, 最终仅表现为滑动模态, 继而误差e的动态过程也几乎完全由滑动模态决定。此时便可以根据 EHA 的位置控制性能指标 (如超调、调节时间和带宽等) 直接确定$f(e)$和ω_{n}, 从而避免了滑模面参数整定的盲目性, 使其更易被应用于实际。同时, 由$f(e)$表达式可以看出: 当$|e|\to+\infty$时, $f(e)\to\xi_0$; 当$|e|\to0$时, $f(e)\to\xi_\infty$, 这意味着随着滑动模态的进行, 阻尼比可在ξ_0和ξ_∞间自适应变化, 这样便可使 EHA 位置阶跃响应在加快响应速度的同时实现无超调调节, 从而取得比定阻尼比更优的位置阶跃调节性能。

4　系统稳定性分析

选择 Lyapunov 函数$V_1=0.5s^2+0.5\tilde{D}^2$, 对其关于时间$t$求导可得

$$\begin{aligned}\dot{V}_1=s\dot{s}+\tilde{D}\dot{\tilde{D}}=&-ks^2-s\tilde{D}+\dot{D}\tilde{D}-\tilde{D}^2/\tau\leqslant\\&-ks^2+0.5s^2+0.5\tilde{D}^2+0.5D_{\mathrm{L}}{}^2+0.5\tilde{D}^2-\tilde{D}^2/\tau=\\&-(k-0.5)s^2-(1/\tau-1)\tilde{D}^2+0.5D_{\mathrm{L}}{}^2=\\&-\beta V_1+0.5D_{\mathrm{L}}{}^2\end{aligned} \tag{29}$$

式中: $0<\tau<1$; $k>0.5$; $\beta=\min\{k-1,1/\tau-1\}$。

求解式 (29) 所示微分不等式, 可得: $V_1(t)\leqslant\exp(-\beta t)V_1(0)+0.5D_{\mathrm{L}}^2/\beta$, 继而可得

$$\begin{cases}|s|\leqslant\sqrt{2\exp(-\beta t)V_1(0)+D_{\mathrm{L}}{}^2/\beta}\\|\tilde{D}|\leqslant\sqrt{2\exp(-\beta t)V_1(0)+D_{\mathrm{L}}{}^2/\beta}\end{cases} \tag{30}$$

由式(30)可见, s和$\tilde{D}$均有界, 且当$t\to+\infty$时, $|s|\to\sqrt{D_{\mathrm{L}}^2/\beta}$, $|\tilde{D}|\to\sqrt{D_{\mathrm{L}}^2/\beta}$。

滑模面的稳定性分以下 2 种情况证明。

1) 如果$\dot{D}=0$, 即不确定性D为慢变扰动, 由式 (24) 明显可知$\tilde{D}$指数收敛, 则到达滑模面后满足: $\ddot{e}+2\omega_{\mathrm{n}}f(e)\dot{e}+\omega_{\mathrm{n}}^2e=0$。选择 Lyapunov 函数$V_1=0.5\omega_{\mathrm{n}}^2e^2+0.5\dot{e}^2$, 对其关于时间$t$求导:

$$\begin{aligned}\dot{V}_2=\omega_{\mathrm{n}}^2e\dot{e}+\dot{e}\ddot{e}=&\omega_{\mathrm{n}}^2e\dot{e}-2\omega_{\mathrm{n}}f(e)\dot{e}^2-\omega_{\mathrm{n}}^2e\dot{e}=\\&-2\omega_{\mathrm{n}}f(e)\dot{e}^2\leqslant0\end{aligned} \tag{31}$$

当$\dot{V}_2\equiv0$时, 可得$e=\dot{e}=0$, 根据 LaSalle 不变集原理可知e和$\dot{e}$渐近稳定。

2) 如果$\dot{D}\neq0$, 且满足$|\dot{D}|<D_{\mathrm{L}}$, 则到达滑模面后满足: $\ddot{e}+2\omega_{\mathrm{n}}f(e)\dot{e}+\omega_{\mathrm{n}}^2e=\eta$, 其中$\eta\leqslant\sqrt{D_{\mathrm{L}}^2/\beta}$, 将其改写为

$$\ddot{e}+2\omega_{\mathrm{n}}f(e)\dot{e}+(\omega_{\mathrm{n}}^2-\eta/e)e=0 \tag{32}$$

当$\omega_{\mathrm{n}}^2-\eta/e>0$时, 取$\gamma=\omega_{\mathrm{n}}^2-\eta/e+\sigma$, $\sigma>0$。选择 Lyapunov 函数$V_3=0.5(\gamma-\sigma)e^2+0.5\dot{e}^2$, 对其关于时间$t$求导可得

$$\begin{aligned}\dot{V}_3=&(\gamma-\sigma)e\dot{e}+\dot{e}\ddot{e}=\\&(\gamma-\sigma)e\dot{e}-2\omega_{\mathrm{n}}f(e)\dot{e}^2-(\omega_{\mathrm{n}}^2-\eta/e)e\dot{e}=\\&-2\omega_{\mathrm{n}}f(e)\dot{e}^2\leqslant0\end{aligned} \tag{33}$$

可见当满足条件$\omega_{\mathrm{n}}^2-\eta/e>0$时, $\dot{V}_3\leqslant0$, 即e和$\dot{e}$均有界, 且当$t\to+\infty$时, $|e|\leqslant\eta/\omega_{\mathrm{n}}^2$。

5　参数整定

所设计控制器共有 6 个参数需要整定, 其中ω_{n}可先根据控制指标确定控制带宽, 然后选择ω_{n}为控制带宽的 2～3 倍。变阻尼敏感系数δ越大, 变阻尼越迟顿, 变阻尼效果越不明显, 但更易产生超调, 系统响应越接近阻尼比固定为ξ_0时的位置阶跃响应。δ越小, 变阻尼越灵敏, 变阻尼效果越明显, 但响应速度越慢, 越接近阻尼比固定为ξ_∞时的

位置阶跃响应。初始阻尼比ξ_0选择时应确保系统初始阶段处于欠阻尼状态，因此选择范围为$0<\xi_0<1$，但ξ_0过小，会造成误差e的收敛速度过快，这样可能会产生超调和振荡，而ξ_0过大会造成e的收敛速度过慢，从而影响响应速度。终态阻尼比ξ_∞选择时需确保在滑动终态阶段系统处在临界阻尼或过阻尼状态，因此其选择范围为$\xi_\infty \geqslant 1$，但ξ_∞选择过大，同样会造成响应速度变慢，尤其在误差e接近平衡点时表现更为明显。因此一般初选$\xi_0=0.1$，$\xi_\infty=1$，然后再根据最终位置阶跃响应结果进行调整。尽管$\dot{D}$有界，但其上界值D_{L}实际中很难准确获取，而且D_{L}的大小会直接影响最终位置控制精度和不确定项估计器时间常数τ和鲁棒增益k的选择，因为τ越小，对D_{L}的抑制能力越强，意味着对D的估计越准确，k越大，对估计误差$\tilde{D}$的抑制效果越好。通常可初选$\tau=1/(3\sim5)\omega_{\mathrm{n}}, k=0$，然后再根据位置控制精度逐渐减小$\tau$和增大$k$，直至满足控制精度要求。但需要注意过小的$\tau$和过大的$k$都会造成控制信号对位置传感器的量测噪声变得更为敏感，甚至会激发系统未建模动态，造成系统振荡，因此实际应用时需折中考虑。

6 仿真验证分析

为了验证本文设计控制器的有效性，将其与PI控制器、传统滑模控制器(SMC)[20]和文献[21]中的变阻尼滑模控制器(damping variable sliding mode controller, DVSMC)进行分析比较。仿真相关参数如表1所示。

由文献[20]可知，SMC的控制律为

$$\omega_{\mathrm{m}}=g_{01}^{-1}(\dddot{x}_{\mathrm{d}}-a_3\ddot{x}-a_2\dot{x}+c_2\ddot{e}+c_1\dot{e})+\varepsilon\mathrm{sign}(s) \tag{34}$$

式中：$a_2=-\beta_{\mathrm{e}}C_{\mathrm{t}}(m_{\mathrm{t}}V_0)^{-1}-A^2\beta_{\mathrm{e}}(m_{\mathrm{t}}V_0)^{-1}$；$a_3=-\beta_{\mathrm{e}}C_{\mathrm{t}}(m_{\mathrm{t}}V_0)^{-1}-B_{\mathrm{t}}m_{\mathrm{t}}^{-1}$；$g_{01}=A\beta_{\mathrm{e}}D_{\mathrm{p}}(m_{\mathrm{t}}V_0)^{-1}$；$\varepsilon>0$。为了便于比较，将参数$c_1$和$c_2$写为阻尼比$\xi$和固有频率$\omega_{\mathrm{n}}$形式：$c_1=\omega_{\mathrm{n}}^2$，$c_2=2\xi\omega_{\mathrm{n}}$。

由文献[21]可知，DVSMC的控制律为

$$\omega_{\mathrm{m}}=g_{01}^{-1}\left[\dddot{x}_{\mathrm{d}}-a_3\ddot{x}-a_2\dot{x}+2\omega_{\mathrm{n}}(f_1(e)\dot{e})'+\omega_{\mathrm{n}}^2\dot{e}\right]+\varepsilon\mathrm{sign}(s) \tag{35}$$

式中：$f_1(e)=\xi_\infty/\left(1+\delta e^2\right)+\xi_0$。

PI控制器参数：$K_{\mathrm{p}}=1\times10^4$，$K_{\mathrm{i}}=7\times10^5$。SMC控制器仿真参数：$\omega_{\mathrm{n}}=10\pi$ rad/s，$\xi=1.2$，$\varepsilon=10$。

DVSMC控制器仿真参数：$\omega_{\mathrm{n}}=10\pi$ rad/s，$\varepsilon=10$。注意到当$|e|\to+\infty$时，$f(e)\to\xi_0$，当$|e|\to0$时，$f(e)\to\xi_0+\xi_\infty$，$\delta=20$。为了比较的公平性，取$\xi_0=0.1$，$\xi_\infty=1.1$，保证终态阻尼比为$\xi_0+\xi_\infty=1.2$。

NSMC控制器仿真参数：$\omega_{\mathrm{n}}=10\pi$ rad/s，$\xi_0=0.1$，$\xi_\infty=1.2$，$\delta=20$，$\tau=1/5\omega_{\mathrm{n}}=0.006$。为了公平起见，$k=\varepsilon=10$。

表1 仿真参数

Table 1 Simulation parameters

参数	数值
液压缸活塞直径/m	0.066
活塞杆直径/m	0.045
活塞行程/m	0.2
液压缸和IEHP总泄漏系数/(10^{-11} m$^3\cdot$s$^{-1}\cdot$Pa^{-1})	2.1
油液弹性模量/(10^8 N$\cdot$m^{-2})	6.86
液压缸初始总体积/(10^{-4} m^3)	3
液压缸和负载总黏性摩擦系数/(N$\cdot$m$^{-1}\cdot$s)	1 100
活塞、活塞杆和负载质量/(10^{-6} kg)	60
IEHP排量/(m$^{-3}\cdot$rad^{-1})	1.59
IEHP总粘性摩擦系数/(N$\cdot$m$\cdot$rad$^{-1}\cdot$s)	0.002
IEHP定子电阻/Ω	0.2
IEHP d、q轴等效电感/H	0.001 87
IEHP 极对数	2
IEHP 永磁体磁链/Wb	0.25
IEHP 转动惯量/kg$\cdot$m^2	0.001 2
IEHP最高转速/(rad$\cdot$s^{-1})	314

ω_{m}限幅值为IEHP的最高转速314 rad/s。IEHP转速环和d、q轴电流环均采用PI控制器，其中转速环带宽取50 Hz，电流环带宽取200 Hz。转速环PI控制器参数分别为$K_{\omega\mathrm{p}}=0.502\,4$，$K_{\omega\mathrm{i}}=0.837\,3$。$d$和$q$轴电流环PI控制器参数相同为：$K_{\mathrm{cp}}=23.487\,2$，$K_{\mathrm{ci}}=1\,380$。

系统仿真框图如图2所示。仿真在以下3种情形下进行：

1）位置指令为小阶跃信号$x_{\mathrm{d}}=0.01$ m。为了便于比较，此时EHA无外负载力。图3为采用4种方法时位置阶跃响应对比曲线。由图中可以看出尽管PI响应速度较快，但出现了约21%的超调，而其余3种方法均实现了无超调调节。由于DVSMC和NSMC的变阻尼机制，两者调节时间均要快于定阻尼SMC。但应注意到，由于DVSMC的滑模变量初值并不为零，只有在经历了到达阶段后才进行阻尼的自适应调节，而NSMC由于滑模面积分和估计器的引入，从t=0时刻就开始了变阻尼调节，因此，NSMC的变阻尼调节效果要比DVSMC更明显，从而造成其调节时间快于DVSMC。

为了进一步说明NSMC的变阻尼效果，固定NSMC的阻尼比，即分别取$f(e)=0.1$和1.2，位置阶跃响应对比曲线如图4所示。从图4可以看出NSMC的响应处于阻尼比为0.1和1.2两者曲线中间，且在初始阶段NSMC与阻尼比为0.1时的曲线重合，随着误差e逐渐减小，阻尼逐渐增大，最终实现无超

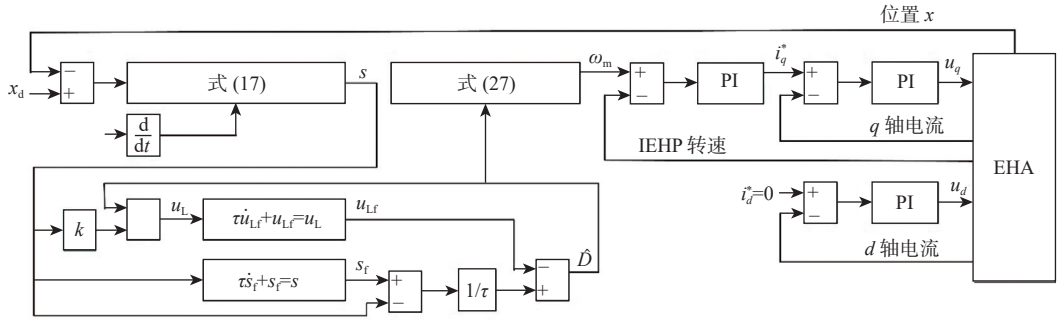


图 2　系统仿真框图

Fig. 2　System simulation block diagram

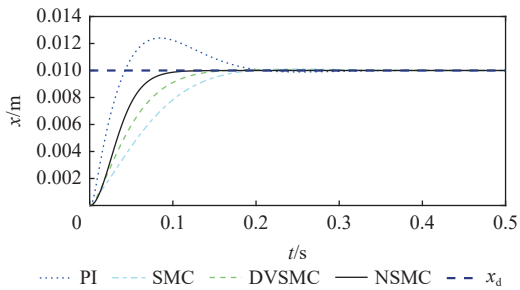


图 3　x_d = 0.01 m时位置阶跃响应对比曲线

Fig. 3　Position step response comparison curves with x_d = 0.01 m

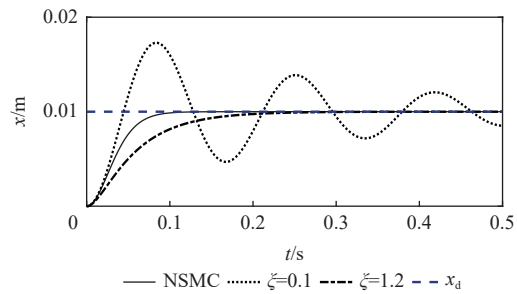


图 4　x_d = 0.01 m时 NSMC 变阻尼与定阻尼位置阶跃响应对比曲线

Fig. 4　Position step response comparison curves under NSMC variable damping and constant damping with x_d = 0.01 m

调调节。

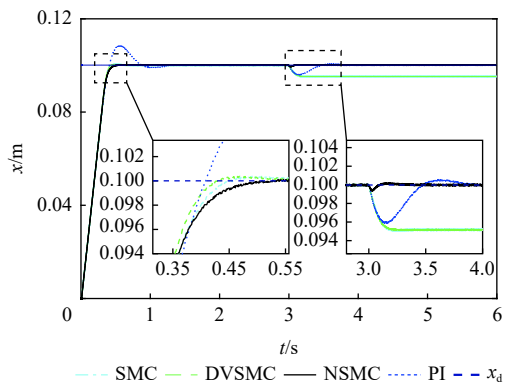
2）位置指令为大阶跃信号 x_d = 0.1 m。由于NSMC 滑模面中引入积分项, 因此当目标位置为大阶跃信号时, 必须对由于ω_m限幅产生的积分饱和进行处理。为了比较公平, PI 和 NSMC 均采用相同的抗积分饱和法: 后向反馈法, 反馈系数k_v = 1。同时, 为了比较 4 种方法的抗扰动能力和噪声敏感性, 位置信号 x 加入均匀分布且幅值为1×10^{-4} m的噪声信号。初始阶段 EHA 空载, 在 t = 3 s时突加 1 000 N 的外负载力。位置阶跃性能对比曲线如图 5 所示。从图 5(a) 可以看出, 在初始阶段 4 种方法的控制信号ω_m均达到了限幅值 314 rad/s, 因此图中的位置响应曲线是重合的。而当位置接近指令值x_d时, 尽管 PI 由于抗饱和措施并未产生严重的积分饱和现象, 但是仍然产生了越 8.3% 的超调, NSMC 由于抗积分饱和措施其响应速度虽然略慢于 SMC 和 DVSMC, 但仍然实现了无超调调节, 而 SMC 和 DVSMC 产生了约 0.5% 的超调。在 t = 3 s施加外负载时, PI 产生了约 0.004 m 的位置降, 位置恢复时间约为 1 s, NSMC 产生了约 0.000 6 m 的位置降, 位置恢复时间约为 0.3 s, 而 SMC 和 DVSMC 则产生了约 0.005 m 的稳态误差, NSMC 的抗扰动能力明显优于其他三者。尽管可以通过增大 PI 控制器的积分增益增大 PI 的抗扰能力, 但这会进一步增大超调。SMC 和 DVSMC 的抗扰能力主要通过切换函数 sign 抑制, 切换函数增益越大, 抗扰能力越强。而 SMC 和 DVSMC 控制律中的切换增益此时为ε = 10, 由于该值此时小于扰动值, 不满足可达性条件, 造成滑模变量 s 只能收敛到滑模面的一个邻域内, 故而造成位置误差 e 存在稳态误差。而 NSMC 由于采用估计器实时估计扰动并补偿, 因此极大地提高了其抗扰动能力, 可以计算此时的扰动量为$D=-F_L/m_t=-16.7$, 从图 5(c) 看出$\hat{D}$实现了对D准确估计。

从图 5(b) 看出 SMC 和 DVSMC 的控制信号ω_m相较于 NSMC 出现了较为严重的抖振现象, 原因在于一方面ω_m中使用了 sign 函数, 另一方面是由于ω_m中的加速度信号产生的噪声放大效应。而为了消除 SMC 和 DVSMC 由于扰动产生的稳态误差, 就需要增大切换增益ε以满足可达性条件, 实际上可以计算此时扰动量为$-F_L\beta_e C_t/m_t V_0=-3\,811$, 因此$\varepsilon$至少应取为 3 811, 但这会加剧$\omega_m$的抖振, 即使将 sign 函数替换为滑模控制中为了消除抖振最常用的饱和函数, 仍会由于加速度的噪声放大效应使得ω_m中依然存在抖振, 因此 SMC 和 DVSMC 的抖振问题很难解决, 难以应用于实际。而 NSMC 在模型降阶后, 控制器无需加速度信息, 避免了加速度产生的噪声放大问题, 另外滑模面积分项和不确定项估计器的使用, 控制器中无需使用 sign 函数, 因

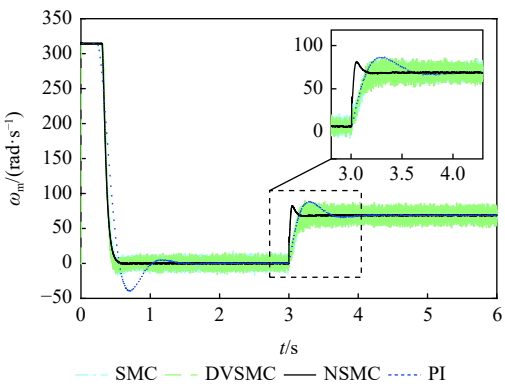
此从根本上解决了抖振问题，从图5(b)也可以看出NSMC控制信号ω_m十分平滑，大大降低了其应用难度。

3）位置指令为时变信号$x_d=0.002\sin(4\pi t)\cdot[1-\exp(-t)]$ m，且系统存在复杂参数不确定性和外部扰动构成的复杂时变扰动。

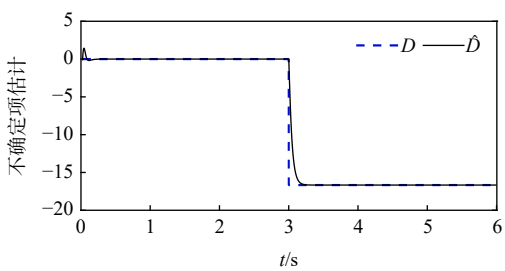
假设从$t=0$ s开始施加外负载$F_L=1\,000\cdot\sin(2\pi t)$ N。负载质量m_t、泄漏系数C_t和液压缸和负载总黏性摩擦系数B_t在$t=2$ s时分别变为原来值的3倍，当$t=4$ s时变为原来值的0.5倍。NSMC和PI的控制器参数不变。为了提高SMC和DVSMC的抗扰动能力，将其切换增益增大为1 000，但这样会产生非常大的抖振，为了便于进行比较，将sign函数用饱和函数sat替代，其余控制参数不变。



(a) 位置阶跃响应对比曲线



(b) 控制信号ω_m对比曲线

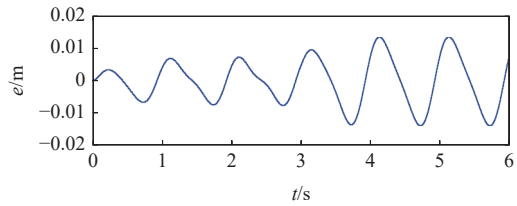


(c) 不确定项估计器扰动估计结果

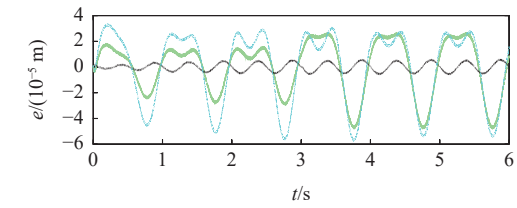
图5 $x_d=0.1$ m，1 000 N阶跃外负载力时位置阶跃性能对比曲线

Fig. 5 Position step performance comparison curves under 1 000 N step external load force with $x_d=0.1$ m

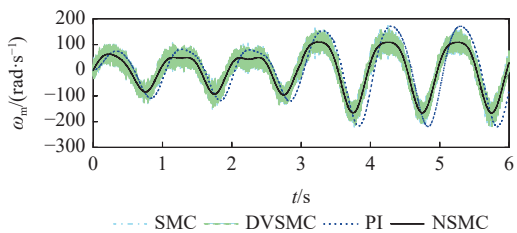
图6为位置跟踪性能对比曲线，由图6(a)和(b)可以看出，PI的位置跟踪误差最大，SMC和DVSMC次之，NSMC的跟踪误差最小，说明此时NSMC对复杂不确定性的鲁棒性显著强于其他三者，这主要得益于不确定项估计器$\hat{D}$对D的估计补偿，从图6(d)和(e)也可以看出，$\hat{D}$实现了对D的准确估计。而且从图6(c)可以看出，NSMC的控制信号ω_m要比SMC和DVSMC的光滑的多，SMC和DVSMC均产生了较大的抖动，由于此时用饱和函



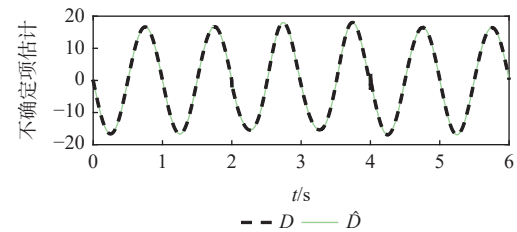
(a) PI位置跟踪误差



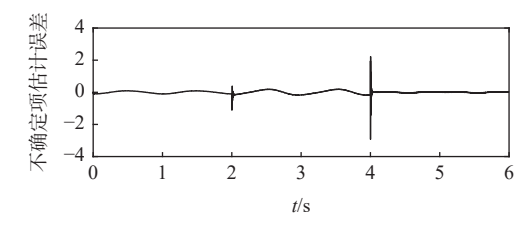
(b) SMC、DVSMC和NSMC位置跟踪误差对比曲线



(c) 控制信号ω_m对比曲线



(d) 不确定项估计器扰动估计结果



(e) 不确定项估计器扰动估计误差

图6 x_d为时变信号，存在复杂时变扰动时位置跟踪性能对比曲线

Fig. 6 Position step performance comparison curves under complex time-varying disturbance when x_d is time-varying sinusoidal signal

数替代了 sign 函数, 因此抖动主要来源于加速度带来的噪声放大。

7　结　论

本文针对 EHA 设计了一种新型非线性变阻尼积分滑模控制器, 该控制器具有结构简单、参数易整定和控制信号无抖振等优点。通过仿真分析对比, 可以得到以下结论:

1) 无论是小阶跃和大阶跃位置指令信号, 利用 NSMC 的阻尼自适应机制, 位置阶跃响应的调节时间均小于其余 3 种控制器, 且同时能实现无超调调节。而对于时变位置指令, NSMC 的位置跟踪误差也要小于其他 3 种方法。

2) 设计的基于滤波器的不确定项估计器, 整定参数仅有一个, 无论是对于阶跃型外负载扰动还是由参数不确定性和时变外负载构成的复杂扰动, 估计器均实现准确估计, 从而确保 NSMC 的抗扰性能优于其余三者。

3) 由于模型降阶不确定项估计器处理、积分项和不确定项估计器的使用, NSMC 避免了加速度和 sign 函数的使用, 从而使 NSMC 的控制信号比 SMC 和 DVSMC 要平滑的多, 去抖效果明显, 实现了无抖振滑模控制。

4) NSMC 的结构形式要比 DVSMC 更为简洁, 无需对非线性变阻尼函数求导, 滑模面参数可根据 EHA 控制指标直接整定, 相较于 SMC 和 DVSMC 更易被应用于实际。

参考文献 (References)

[1] 付永领, 韩旭, 杨荣荣, 等. 电动静液作动器设计方法综述[J]. 北京航空航天大学学报, 2017, 43(10): 1939-1952.
FU Y L, HAN X, YANG R R, et al. Review on design method of electro-hydrostatic actuator[J]. Journal of Beijing University of Aeronautics and Astronautics, 2017, 43(10): 1939-1952 (in Chinese).

[2] MARÉ J C, FU J. Review on signal-by-wire and power-by-wire actuation for more electric aircraft[J]. Chinese Journal of Aeronautics, 2017, 30(3): 857-870.

[3] ALLE N, HIREMATH S S, MAKARAM S, et al. Review on electro hydrostatic actuator for flight control[J]. International Journal of Fluid Power, 2016, 17(2): 125-145.

[4] 张英臣, 祁晓野. 基于 PID 和滑模控制的电动静液作动器的研究[J]. 机床与液压, 2017, 45(21): 122-126.
ZHANG Y C, QI X Y. Study of electro-hydraulic actuator based on PID and sliding mode control[J]. Machine Tool & Hydraulics, 2017, 45(21): 122-126 (in Chinese).

[5] EL SAYED M A, HABIBI S. Inner-loop control for electro-hydraulic actuation systems[J]. Journal of Dynamic Systems, Measurement, and Control, 2012, 134(1): 014501.

[6] AHN K K, NAM D N C, JIN M L. Adaptive backstepping control of an electrohydraulic actuator[J]. IEEE/ASME Transactions on Mechatronics, 2014, 19(3): 987-995.

[7] SUN W C, GAO H J, YAO B. Adaptive robust vibration control of full-car active suspensions with electrohydraulic actuators[J]. IEEE Transactions on Control Systems Technology, 2013, 21(6): 2417-2422.

[8] GE Y W, YANG X W, DENG W X, et al. RISE-based composite adaptive control of electro-hydrostatic actuator with asymptotic stability[J]. Machines, 2021, 9(9): 181.

[9] WANG M K, WANG Y, FU Y L, et al. Experimental investigation of an electro-hydrostatic actuator based on the novel active compensation method[J]. IEEE Access, 2020, 8: 170635-170649.

[10] ZHANG H, LIU X T, WANG J M, et al. Robust H_∞ sliding mode control with pole placement for a fluid power electrohydraulic actuator (EHA) system[J]. The International Journal of Advanced Manufacturing Technology, 2014, 73(5-8): 1095-1104.

[11] 郭新平, 汪成文, 刘华, 等. 基于扩张状态观测器的泵控电液伺服系统滑模控制[J]. 北京航空航天大学学报, 2020, 46(6): 1159-1168.
GUO X P, WANG C W, LIU H, et al. Extended-state-observer based sliding mode control for pump-controlled electro-hydraulic servo system[J]. Journal of Beijing University of Aeronautics and Astronautics, 2020, 46(6): 1159-1168 (in Chinese).

[12] 李小虎, 王孙安, 张育林, 等. 基于反馈变结构控制电动静液作动器的研究[J]. 机床与液压, 2008, 36(12): 71-74.
LI X H, WANG S A, ZHANG Y L, et al. Study on sliding mode control based on feedback for electro-hydrostatic actuator[J]. Machine Tool & Hydraulics, 2008, 36(12): 71-74 (in Chinese).

[13] LIN Y, SHI Y, BURTON R. Modeling and robust discrete-time sliding-mode control design for a fluid power electrohydraulic actuator (EHA) system[J]. IEEE/ASME Transactions on Mechatronics, 2013, 18(1): 1-10.

[14] WANG S, BURTON R, HABIBI S. Sliding mode controller and filter applied to an electrohydraulic actuator system[J]. Journal of Dynamic Systems, Measurement, and Control, 2011, 133(2): 024504.

[15] WANG S, HABIBI S, BURTON R, et al. Sliding mode control for a model of an electrohydraulic actuator system with discontinuous nonlinear friction[C]//2006 American Control Conference. Piscataway: IEEE Press, 2006, 1-2: 780.

[16] ALEMU A E, FU Y L. Sliding mode control of electro-hydrostatic actuator based on extended state observer[C]//2017 29th Chinese Control and Decision Conference (CCDC). Piscataway: IEEE Press, 2017: 758-763.

[17] LEE L M, PARK P H , KIM K J S. Design and experimental evaluation of a robust position controller for an electro- hydrostatic actuator using adaptive anti-windup sliding mode scheme[J]. The Scientific World Journal, 2013, 2013: 590708.

[18] 张振, 李海军, 诸德放. EHA 反馈线性化最优滑模面双模糊滑模控制[J]. 北京航空航天大学学报, 2016, 42(7): 1398-1405.
ZHANG Z, LI H J, ZHU D F. Double fuzzy sliding mode control for EHA based on feedback linearization optimal sliding surface[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(7): 1398-1405 (in Chinese).

[19] YANG R R, FU Y L, ZHANG L, et al. A novel sliding mode control framework for electrohydrostatic position actuation system[J]. Mathematical Problems in Engineering, 2018, 2018: 7159891.

[20] 杨荣荣. 基于新型电液泵的电静液作动系统控制策略研究[D]. 北

京: 北京航空航天大学, 2019: 76-77.
YANG R R. Research on control strategy for electro-hydrostatic actuation system based on a novel integrated electrohydraulic pump[D]. Beijing: Beijing University of Aeronautics and Astronautics, 2019: 76-77 (in Chinese).
[21] 王明康, 付永领, 赵江澳, 等. 电动静液作动器的新型变阻尼级联滑模控制[J]. 北京航空航天大学学报, 2021, 47(8): 1612-1618.
WANG M K, FU Y L, ZHAO J A, et al. Novel damping-variable sliding mode cascade control for electro-hydrostatic actuator[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(8): 1612-1618 (in Chinese).
[22] WANG M K, WANG Y, YANG R R, et al. A sliding mode control strategy for an ElectroHydrostatic actuator with damping variable sliding surface[J]. Actuators, 2020, 10(1): 3.
[23] 付永领, 李祝锋, 薛晶, 等. 一种带转速反馈的永磁轴向柱塞式电液泵: CN203423578U[P]. 2015-08-26.
FU Y L, LI Z F, XUE J, et al. Permanent magnet radial plunger type electro-hydraulic pump with rotating speed feedback: CN203423578U[P]. 2015-08-26 (in Chinese).
[24] PONOMAREV P, POLIKARPOVA M, HEINIKAINEN O, et al. Design of integrated electro-hydraulic power unit for hybrid mobile working machines[C]//Proceedings of the 14th European Conference on Power Electronics and Applications (EPE2011). Piscataway: IEEE Press, 2011: 1-10.
[25] KHALIL H K. Nonlinear systems[M]. 3rded. Upper Saddle River, NJ: Prentice Hall, 2002: 423-430.

Nonlinear variable damping integral sliding mode control for electro-hydrostatic actuator

YANG Rongrong[1,*], ZHANG Ling[1], ZHAO Jiali[1], FU Yongling[2], ZHANG Peng[3]

(1. School of Mechanical and Electrical Engineering, Lanzhou University of Technology, Lanzhou 730050, China;
2. School of Mechanical Engineering and Automation, Beihang University, Beijing 100191, China;
3. Beijing Research Institute of Precise Mechatronic and Controls, Beijing 100076, China)

Abstract: For electro-hydrostatic actuator (EHA), the traditional sliding mode controller has many defects, such as control signal chattering, difficulty in acceleration information acquirement and controller parameters tuning, which make the controller difficult to be applied in practice. To solve the above problems, the reduced order mathematical model of EHA is obtained reasonably by using singular perturbation theory, which result in avoiding the use of acceleration information. On this basis, a novel nonlinear variable damping integral sliding mode controller (NSMC) is synthesized by employing the reduced order model. NSMC can adaptively adjust the system damping ratio from underdamping to overdamping according to the position control error, and can effectively improve the position step performance. Besides, filter-based uncertainty estimator is designed to estimate and compensate the parameter uncertainties and external disturbances of EHA in real time. Due to the introduction of sliding mode surface integral term and uncertainty estimator, switching function is unnecessary, on the one hand, the free-chattering sliding control can be achieved. On the other hand, the whole dynamic process is determined directly by the dynamic of sliding mode surface, so that sliding mode surface parameters can be tuned directly according to EHA's control performance index, which greatly simplifies the parameters tuning process. Meanwhile, the stability of the whole closed-loop system and sliding mode surface is proved via Lyapunov stability theory. Finally, by comparing respectively with PI, conventional sliding mode (SMC) and damping variable sliding mode controller (DVSMC), the simulation results indicated that NSMC can effectively improve EHA position tracking performance and enhance the robustness of parameter uncertainty and external disturbance.

Keywords: electro-hydrostatic actuator; singular perturbation; damping variable; integral sliding mode control; uncertainty estimator

Received: 2022-04-18; **Accepted**: 2022-06-10; **Published Online**: 2022-06-14 11:09
URL: link.cnki.net/urlid/11.2625.V.20220613.1814.004
Foundation item: Gansu Provincial Science and Technology Plan Project (20JR10RA175)
* **Corresponding author**. E-mail: yr13236@163.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0213

过冷大水滴动力学行为对翼型结冰的影响分析

刘飞宇[1]，邓甜[1, 2, *]

(1. 中国民航大学 中欧航空工程师学院，天津 300300；2. 中国空气动力研究与发展中心 结冰与防除冰重点实验室，绵阳 621000)

摘 要：过冷大水滴（SLD）是极端危险的飞行环境之一。因大粒径水滴独特的动力学行为变形破碎、飞溅反弹，传统结冰计算方法难以准确地反映SLD结冰情况。采用Navier-Stokes方法求解流场、Euler方法计算水滴撞击、Shallow Water模型模拟结冰，并与NASA实验结果进行了对比验证方法可信。结果表明：SLD动力学行为对结冰和冰形影响较大。其中，变形破碎改变了水滴运动轨迹和撞击范围，降低了水滴撞击极限，导致上下结冰极限减小2.83 %、2.13 %；飞溅降低了驻点附近水滴收集率，导致前缘积冰量减少8.09 %；反弹显著降低了水滴撞击极限，导致上下结冰极限减小30.69 %、20.01 %；撞击后飞溅反弹二次水滴再入流场使得上下结冰极限增加6.14 %、3.71 %。同时，与干净翼型相比带冰翼型空气动力学性能严重退化，在相同迎角下，升力更小、阻力更大、气动效率更低。

关 键 词：过冷大水滴；动力学行为；翼型结冰；冰形几何特征；气动特性

中图分类号：V211.41+2

文献标志码：A **文章编号**：1001-5965(2024)01-0173-14

过冷大水滴 (supercooled large droplets, SLD) 通常指直径大于40 μm, 在冻雨环境中甚至能达到2 000 μm的过冷水滴[1]。SLD会在气动力作用下发生变形, 当变形足够大时发生破碎。SLD与壁面发生碰撞后, 部分水滴会发生飞溅或反弹, 产生的二次液滴又将再次注入流场, 可能将与下游壁面发生多次碰撞。

Villedieu等[2]基于拉格朗日法, 通过计算单个水滴粒子的运动方程来跟踪该粒子。Honsek等[3]推导了水滴-壁面相互作用过程的欧拉模型。Bourgault等[4]提出一种含水滴的欧拉模型来替代传统的拉格朗日粒子跟踪方法, 给出了合适的液滴方程边界条件。Iuliano等[5]基于欧拉法框架解释了SLD动力学行为的特殊性质。

Leal[6]提出了一个基于液滴偏心和振动畸变的拓展阻力模型, 解释SLD条件下气动阻力的增加。Pilch和Erdman[7]基于临界韦伯数、破碎时间和速度预测稳定液滴碎片的最大尺寸。Hsiang和Faeth[8]根据韦伯数和奥内佐格数将SLD的变形与破碎机制划分为6个阶段。O'Rourke和Amsden[9]建立了泰勒类比破碎 (Taylor analogy breakup, TAB) 模型计算液滴破碎。Clark[10]建立了二维直线液滴振动模型。Ibrahim等[11]建立了液滴变形与破碎 (drop deformation and breakup, DDB) 模型, 适用于纯拉伸流动中液滴的剪切型变形。

Bai和Gosman[12]确定了SLD撞击壁面4种可能机制：黏附、反弹、铺展与飞溅。Mundo等[13-15]提出了一种基于实验数据的液滴碰壁预测模型。Wright等[16-17]基于对Mundo模型的修正来确定SLD飞溅或反弹的概率。Norde等[18]研究了SLD撞击壁面时水膜存在的影响, 提出了一个由沉积模型组成的飞溅模型。Bilodeau等[19]通过改变各表面边

收稿日期：2022-04-02；录用日期：2022-06-06；网络出版时间：2022-09-15 08:01
网络出版地址：link.cnki.net/urlid/11.2625.V.20220913.1845.012
基金项目：国家自然科学基金 (U1933110)；结冰与防除冰重点实验室开放课题 (IADL20200305)；天津市教委科研计划项目 (2020KJ036)
* 通信作者. E-mail：t-deng@cauc.edu.cn

引用格式：刘飞宇，邓甜. 过冷大水滴动力学行为对翼型结冰的影响分析 [J]. 北京航空航天大学学报，2024，50（1）：173-186.
LIU F Y，DENG T. Influence of dynamic behavior of supercooled large droplets on airfoil icing[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50（1）：173-186（in Chinese）.

界条件来跟踪由SLD飞溅、反弹行为产生的二次液滴。Trujillo等[20]通过实验和计算研究了SLD撞击平面时,撞击角度对飞溅/反弹阈值的影响。Rutkowski等[16]确定了能够再次撞击翼型表面的飞溅二次液滴需要具有的逃逸速度。

易贤等[21]采用数值计算的手段,通过模拟SLD运动过程中的变形、破碎与飞溅现象,分别研究了水滴直径、流场速度及绕流物体尺寸变化对水滴动力学行为及其效应的影响。张辰等[22]针对SLD成冰模拟中破碎模型的不足,根据对破碎机理的认识提出了多因素破碎(multifactor-effected breakup, MEB)判定模型。任靖豪等[23]基于欧拉-拉格朗日数值模拟框架,发展了一套适合求解三维多体模型水滴撞击特性的计算方法,采用粒子统计与面积比率计算相结合的方式,实现了水滴收集率的高效精确预测。周志宏等[24]发展了一种水滴撞击特性的高效计算方法,采用目标扩散追踪方法对水滴所在网格单元进行快速计算,并插值得到该点处的流场信息,逐个求解水滴运动方程得到各水滴的运动轨迹,从而确定水滴撞击极限和收集系数等撞击特性参数。Glenn研究中心[25]通过大量实验建立了SLD冰积累的冰形数据库。

综上所述,研究人员从理论和实验方面研究了SLD的动力学特性与结冰机理,创造了很多有意义的方法与模型。但当前国内外针对SLD环境下的航空器结冰问题的研究还不是很充分,尤其是SLD宏观结冰数值模拟方面,公开文献中的研究还比较有限,对SLD条件下的结冰过程、冰形特点和气动特性影响还缺乏深入的认识。本文使用雷诺平均Navier-Stokes方法计算空气流场、欧拉方法计算SLD粒子轨迹,采用准稳态方法研究持续结冰引起的翼型几何形状变化及其对空气、水滴流场的影响,分析了SLD不同动力学行为对翼型结冰的影响。

1 理论与方法

1.1 SLD运动轨迹模型

欧拉方法将分散的液相视为连续流体,能够预测与气动解相同节点处的液滴体积分数和速度分量,且避免了拉格朗日方法中密集粒子跟踪过程的插值,在复杂模型计算上具有较大优势[26]。

欧拉两相流方法以空间中的位置为研究对象,通过建立和求解控制体积中的质量和动量方程,给出水滴表观密度和速度的分布。应用与空气流场相似的解析方法,可得到水滴流场的连续性方程与动量方程如下[4]:

$$\frac{\partial \alpha}{\partial t}+\nabla\cdot(\alpha \boldsymbol{V}_{\mathrm{d}})=0 \tag{1}$$

$$\frac{\partial(\alpha \boldsymbol{V}_{\mathrm{d}})}{\partial t}+\nabla(\alpha \boldsymbol{V}_{\mathrm{d}}\otimes \boldsymbol{V}_{\mathrm{d}})=\frac{C_D Re_{\mathrm{d}}}{24K}\alpha(\boldsymbol{V}_{\mathrm{a}}-\boldsymbol{V}_{\mathrm{d}})+\alpha\left(1-\frac{\rho_{\mathrm{a}}}{\rho_{\mathrm{d}}}\right)\frac{1}{Fr^2} \tag{2}$$

式中:α为液态水体积分数;t为时间;$\boldsymbol{V}_{\mathrm{d}}$为液滴速度矢量;$\boldsymbol{V}_{\mathrm{a}}$为空气速度矢量;$\rho_{\mathrm{a}}$与$\rho_{\mathrm{d}}$分别为空气与液滴的密度;$C_D$为阻力系数;$Re_{\mathrm{d}}$为液滴雷诺数;$K$为惯性参数;$Fr$为局部Froude数。$K$和$Fr$分别具体表示为

$$K=\frac{\rho_{\mathrm{d}} d^2 V_{\mathrm{a},\infty}}{18L\mu_{\mathrm{a}}} \tag{3}$$

$$Fr=\frac{V_{\mathrm{a},\infty}}{\sqrt{Lg}} \tag{4}$$

其中:d为液滴直径;$V_{\mathrm{a},\infty}$为无穷远处空气速度;L为特征长度;μ_{a}为空气动力黏度;g为重力加速度。球形液滴的阻力系数C_D由半经验公式(5)得到:

$$C_D=\begin{cases}(24/Re_{\mathrm{d}})\left(1+0.15Re_{\mathrm{d}}^{0.687}\right) & Re_{\mathrm{d}}\leqslant 1\,300\\ 0.4 & Re_{\mathrm{d}}>1\,300\end{cases} \tag{5}$$

在韦伯数We低于13时,液滴的阻力系数C_D介于球形液滴阻力系数$C_{D,\mathrm{sphere}}$和圆盘状液滴阻力系数$C_{D,\mathrm{disk}}$之间:

$$C_D=fC_{D,\mathrm{sphere}}+(1-f)C_{D,\mathrm{disk}} \tag{6}$$

式中:比例系数f表示为

$$f=1-E^2 \tag{7}$$

其中:中间参数E由Hsiang和Faeth[8,27]给出的经验公式计算得到:

$$E=\left(1+0.07\sqrt{We}\right)^{-3} \tag{8}$$

通过求解控制体积中的质量和动量方程,最终可以得到液滴在撞击表面的质量与速度分布,从而求得局部水滴收集率作为结冰模型的输入参数。由于1.2~1.3节的特殊物理现象,为了求解SLD的运动和撞击特性,仍然需要对模型进行修正。

1.2 SLD变形与破碎模型

在气动力作用下,SLD达到某一临界条件后将会发生变形。不均匀的压力将会在液滴上产生表面波,而表面张力又试图将液滴固定在一起,液滴的形状开始从球状变形到圆盘状,液滴的阻力系数也逐渐增大。当表面张力终于无法再将液滴固定在一起时,大液滴将破碎成若干小液滴,液滴的轨迹将发生显著的变化,如图1所

示, δy 表示初始水滴流分布宽度; δs 表示撞击时水滴流分布宽度。

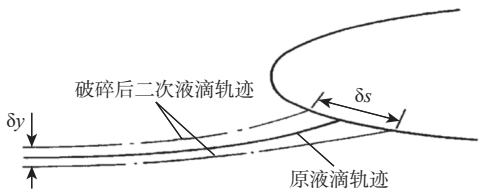


图 1　SLD 破碎导致的液滴轨迹变化

Fig. 1　Change of droplet trajectory caused by SLD breakup

用来描述液滴破碎概率的无量纲数是韦伯数, Pilch 和 Erdman[7] 定义了一个基于临界韦伯数We_{critical}的液滴破碎相关性公式, 必须超过We_{critical}液滴才会发生破碎:

$$We_{\text{critical}} = 12\left(1 + 1.077Oh^{1.6}\right) \tag{9}$$

式中: 奥内佐格数 Oh 是一个将黏性力、惯性力以及表面张力联系起来的一个无量纲数:

$$Oh = \frac{\mu_{\text{d}}}{\sqrt{d\rho_{\text{d}}\sigma_{\text{d}}}} \tag{10}$$

其中: μ_{d}为液滴动力黏度; σ_{d}为液滴表面张力系数, 再结合液滴直径d与液滴的密度ρ_{d}可以求得We_{critical}的具体数值, 若有

$$We_{\text{critical}} \leqslant We = \frac{\rho_{\text{d}} V_{\text{d}}^2 L}{\sigma_{\text{d}}} \tag{11}$$

则液滴将突破表面张力的束缚破碎成若干小液滴, 破碎后二次液滴粒子直径将由 (12) 式确定:

$$d_{\text{s}} = 6.2d_{\text{i}}\left(\frac{\rho_{\text{d}}}{\rho_{\text{a}}}\right)^{1/4} Re_{\text{d}}^{-1/2} \tag{12}$$

式中: d_{i}为破碎前液滴粒子直径; d_{s}为破碎后二次液滴粒子直径。

1.3　SLD 飞溅与反弹模型

SLD 撞击壁面后可能会发生飞溅或反弹, 如图 2 所示, V_{i} 和 V_{s} 分别表示原液滴和二次液滴的速度; $\boldsymbol{n}$ 和 $\boldsymbol{t}$ 分别表示法向和切向。Mundo 等[15] 定义了一个参数 k_{M} 来确定 SLD 撞击壁面后飞溅或反弹的概率:

$$k_{\text{M}} = k_{\text{C}}^{5/8} \tag{13}$$

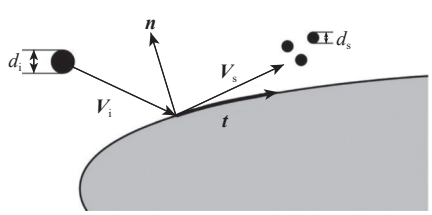


图 2　液滴撞击壁面示意图

Fig. 2　Diagram of droplet impact on wall

式中: k_{C} 是由韦伯数和奥内佐格数定义的 Cossali 参数, 即

$$k_{\text{C}} = WeOh^{-2/5} \tag{14}$$

Wright-Potapczuk 模型根据 k_{M} 的修正量 k_{Y} 来确定飞溅或回弹的概率:

$$k_{\text{Y}} = \left(\frac{L_{\text{WC,i}}}{\rho_{\text{d}}}\right)^{-1/8} k_{\text{M}}^{1/2} \tag{15}$$

式中: $L_{\text{WC,i}}$ 表示初始液态水含量。

随后在 k_{Y} 的基础上, 通过考虑水滴撞击接触角 θ_{i} 定义了 Wright 参数 k_{W}, 即

$$k_{\text{W}} = \frac{0.86k_{\text{Y}}}{\sin\theta_{\text{i}}^{5/4}} \tag{16}$$

通过 We、Oh 与 $L_{\text{WC,i}}$、初始接触角 θ_{i} 可以求得 Wright 参数 k_{W} 的具体数值。在 $k_{\text{W}} > 200$的情况下, 液滴撞击壁面将发生飞溅, 飞溅二次液滴与原液滴直径之比、液态水含量之比、速度 V 之比分别为

$$\frac{d_{\text{s}}}{d_{\text{i}}} = 8.72\text{e}^{-0.028k_{\text{Y}}} \tag{17}$$

$$\frac{L_{\text{WC,s}}}{L_{\text{WC,i}}} = 0.7(1-\sin\theta_{\text{i}})\left[1-\text{e}^{-0.009\,202\,6(k_{\text{W}}-200)}\right] \tag{18}$$

$$\begin{cases} \dfrac{V_{n,\text{s}}}{V_{n,\text{i}}} = 0.30 - 0.002\theta_{\text{i}} \\ \dfrac{V_{t,\text{s}}}{V_{t,\text{i}}} = 1.075 - 0.002\,5\theta_{\text{i}} \end{cases} \tag{19}$$

在$k_{\text{W}} < 200$的情况下, 液滴撞击壁面将发生反弹, 反弹二次液滴与原液滴各项参数之比为

$$\begin{cases} \dfrac{d_{\text{s}}}{d_{\text{i}}} = 1 \\ \dfrac{L_{\text{WC,s}}}{L_{\text{WC,i}}} = 1 \\ \dfrac{V_{n,\text{s}}}{V_{n,\text{i}}} = 1 \\ \dfrac{V_{t,\text{s}}}{V_{t,\text{i}}} = 1 \end{cases} \tag{20}$$

式 (17)～式(20)中: 下标 i 表示原液滴; 下标 s 表示撞击后产生的二次液滴; 下标n表示法向; 下标t表示切向。

1.4　结冰模型

在空气流场、水滴流场及水滴撞击特性已知的基础上, 对热力学系统进行传质与传热计算, 可以获得水滴撞击表面的水膜高度与冻结分数, 再进行时间积分可以得到水滴撞击表面各个点的积冰高度, 随后再进行面积分便可勾勒出整体的积冰形状。

在结冰预测研究中, 使用 Shallow Water 结冰模

型(SWIM)进行结冰模拟,该模型在热源与结冰沉积方面与经典的Messinger结冰模型有相似之处,但水膜流动受三维偏微分方程(partial differential equation, PDE)控制,在结构化和非结构化表面网格上采用稳定的有限体积格式离散。SWIM由质量和能量平衡的2个偏微分方程与动量系统的代数方程组成。壁面薄膜由气流剪切应力和重力或旋转体积力驱动。模型方程概述如下:

质量守恒方程

$$\rho_{\mathrm{w}}\left[\frac{\partial h_{\mathrm{f}}}{\partial t}+\nabla\cdot(\boldsymbol{u}h_{\mathrm{f}})\right]=\dot{m}_{\beta}-\dot{m}_{\mathrm{ice}}-\dot{m}_{\mathrm{evap}} \tag{21}$$

式中:ρ_{w}为水的密度;h_{f}为水膜高度;$\boldsymbol{u}$为水膜速度;$\dot{m}_{\beta}$、$\dot{m}_{\mathrm{ice}}$和$\dot{m}_{\mathrm{evap}}$分别为液滴撞击、固化和蒸发的质量通量。

动量方程

$$\boldsymbol{u}=\frac{h_{\mathrm{f}}}{2\mu_{\mathrm{w}}}\boldsymbol{\tau}-\frac{h_{\mathrm{f}}^{2}}{3\mu_{\mathrm{w}}}\rho_{\mathrm{w}}\boldsymbol{a} \tag{22}$$

式中:$\boldsymbol{\tau}$为表面剪切应力矢量;μ_{w}为水动力黏度;$\boldsymbol{a}$为水膜受到的累积加速度(离心力、科里奥利力、重力)。

能量方程

$$\begin{aligned}\rho_{\mathrm{w}}\left[\frac{\partial(h_{\mathrm{f}}C_{\mathrm{w}}T)}{\partial t}+\nabla\cdot(\boldsymbol{u}h_{\mathrm{f}}C_{\mathrm{w}}T)\right]=&\\ \left[C_{\mathrm{w}}T_{\mathrm{d}}+\frac{\|\boldsymbol{u}_{\mathrm{d}}\|^{2}}{2}\right]\dot{m}_{\beta}+(L_{\mathrm{f}}-C_{\mathrm{ice}}T)\dot{m}_{\mathrm{ice}}-&\\ \dot{Q}_{\mathrm{evap}}+\dot{Q}_{\mathrm{h}}+\dot{Q}_{\mathrm{cond}}&\end{aligned} \tag{23}$$

式中:T为取决于雾凇冰或明冰状态的水膜或冰的温度,单位为℃;C_{w}为水的比热容;C_{ice}为冰的比热容;$\boldsymbol{u}_{\mathrm{d}}$为初始撞击的水滴的速度矢量;$T_{\mathrm{d}}$为初始撞击的水滴的温度;$L_{\mathrm{f}}$为水的冻结潜热;$\dot{Q}_{\mathrm{evap}}$、$\dot{Q}_{\mathrm{h}}$和$\dot{Q}_{\mathrm{cond}}$为分别来自于蒸发、对流和传导的热通量。

进行方程联立求解,可将其简化为关于T和冻结分数n_{F}的方程,其中n_{F}可以表示为

$$n_{\mathrm{F}}=\frac{C_{\mathrm{w}}}{L_{\mathrm{f}}}\left(\phi+\frac{\varepsilon}{b}\right) \tag{24}$$

式中:ϕ为水滴能量传递参数;ε为空气能量传递参数;b为相对热因子。

求解得到T和冻结分数n_{F}之后,再结合水滴撞击求解得到的结果,积分可得整个翼面上的结冰质量和厚度。

2 数值模拟方法与验证

2.1 准稳态结冰模拟方法

非定常的三维积冰模拟计算成本非常高,即使用最新的计算机,计算所需资源也令人望而却步[28-29]。而如果将冰积累的非稳态演变过程细分为数个准稳态冰积累过程,这个问题就容易得到解决。

准稳态结冰过程的每个循环为4步,如图3所示。

步骤1 空气流场求解,根据如图4所示的翼型结冰模拟关键参数,如无穷远处的流场空气静温T_{∞}、流场空气静压P_{∞}、来流速度V_{∞}以及迎角α,确定空气-壁面剪切应力张量$\boldsymbol{\tau}_{\mathrm{wall}}$、对流热通量$\dot{Q}_{\mathrm{h}}$、翼型周围空气流速$u_{\mathrm{a}}$与空气密度$\rho_{\mathrm{a}}$。

步骤2 SLD壁面撞击计算,根据如图4所示的翼型结冰模拟关键参数,如液态水含量L_{WC}、液滴中值体积直径M_{VD}等,确定翼型表面水滴收集率β的分布与翼型周围水滴运动速度u_{d}的分布。

步骤3 系统热力学分析,计算水滴撞击表面水回流量及局部结冰速率。

步骤4 对带冰翼型进行网格重构,用作下一循环。重复该过程,直至达到所需的结冰时间。

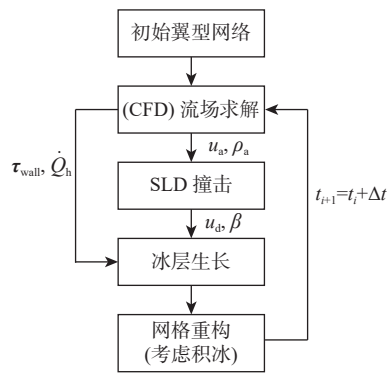


图3 准稳态结冰模拟流程图

Fig. 3 Quasi-steady-state icing simulation flow chart

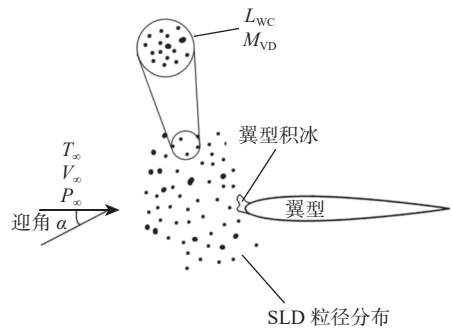


图4 翼型结冰模拟关键参数

Fig. 4 Key parameters for airfoil icing simulation

2.2 计算域与网格划分

计算所采用的模型为NACA0012翼型,弦长0.533 4 m。计算域及网格划分情况如图5所示。

采用非结构网格,网格数量约40万,调整壁面第一层网格厚度使得壁面处$Y^{+}=1$,远场距离壁面约10倍弦长。进行网格无关性验证,静压为101 325 Pa,

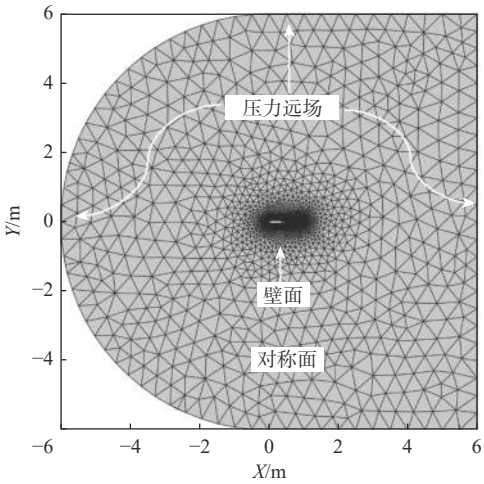
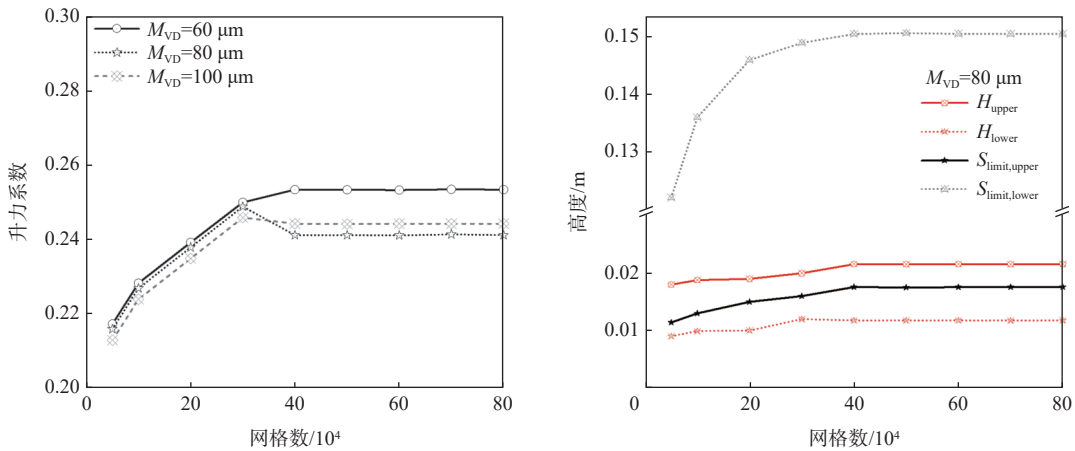


图 5　计算域及网格划分

Fig. 5　Computational domain and meshing

真空速为 200 n mile/h(1 n mile/h=1.852 km/h), 迎角为 3.8°, 温度为−10.79 ℃, L_{WC}=1 g/m^3, M_{VD} 分别为 60 μm、80 μm、100 μm(单分布), 结冰时间为 210 s 的翼型结冰情况如图 6(a) 所示。可见, 当网格数量达到 40 万后, 网格数量增加对翼型升力系数 C_L 没有明显影响。考察 M_{VD}=80 μm, 网格数量对翼型前缘冰形特征参数[30](见图 7), 包括上下冰角高度 H_{upper}、H_{lower}, 上下结冰极限 $S_{limit,upper}$、$S_{limit,lower}$ 的影响, 如图 6(b) 所示。可以看出, 网格数量达到 40 万后网格数量增加对冰形参数计算精度提升没有明显贡献。基于计算资源最优化的考虑, 采用 40 万网格进行后续的计算。此外, 在图 7 中 θ_{upper} 与 θ_{lower} 分别表示上下冰角角度; $H_{stagnation}$ 表示驻点结冰高度。



(a) 网格无关性验证-C_L　　(b) 网格无关性验证-冰形参数

图 6　网格无关性验证

Fig. 6　Grid independence verification

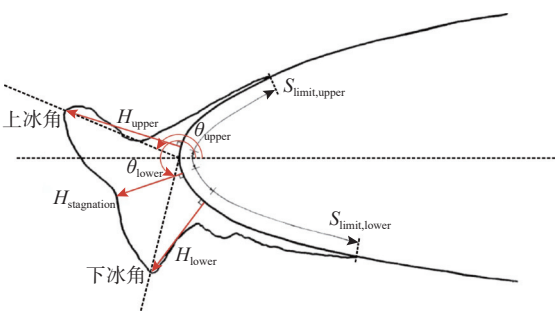


图 7　翼型前缘冰形特征参数

Fig. 7　Ice shape characteristic parameters of airfoil leading edge

2.3　方法验证

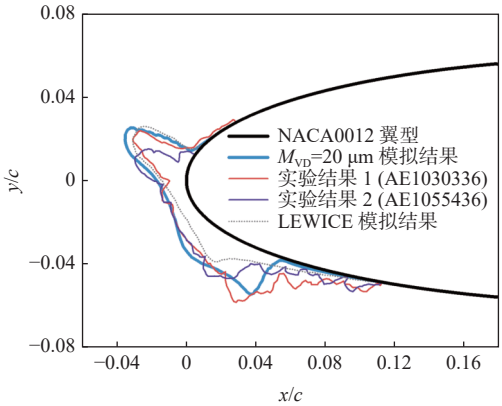
为了验证上述翼型结冰模型与方法的准确性, 选取了如表 1 所示的 NASA Glenn 研究中心结冰风洞 (icing research tunnel, IRT) 的实验工况进行数值模拟, 其中, 结冰表面为 0.533 4 m(算例 1)、0.914 4 m(算例 2、算例 3)NACA0012 翼型, 水滴场中水滴尺寸分布遵循 Langmuir-D 分布模型。

将数值模拟得到的冰形结果与实验结果进行对比, 结果如图 8 所示, c 为翼型弦长。从图 8 中可以看出, 虽然计算结果与实验结果之间存在一定的误差, 但总体接近。将计算误差列于表 2, 从表中可以看出, 最大误差<10%, 证明了本文的结冰预测模型对 M_{VD}<40 μm 的过冷小水滴和 M_{VD}>40 μm 的 SLD 都具有可靠性。

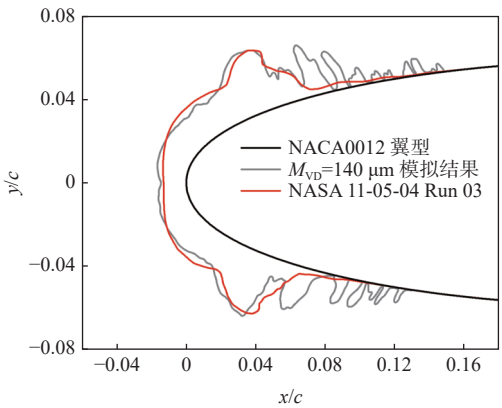
表 1　验证算例工况

Table 1　Working conditions of examples to be verified

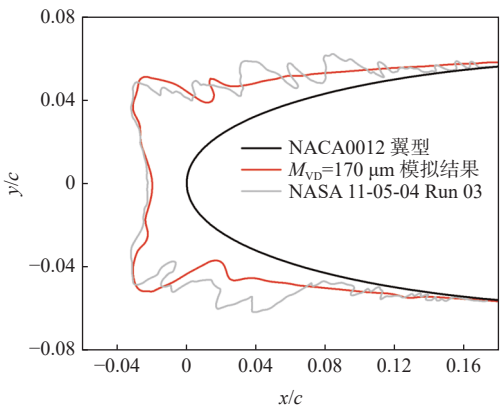
算例	来流速度/(n mile·h^{-1})	静温/℃	迎角/(°)	L_{WC}/(g·m^{-3})	M_{VD}/μm	结冰时间/s
1	200	−10.79	3.8	1.00	20	231
2	100	−9	0	1.17	140	14×60
3	100	−18	0	1.46	170	11×60



(a) 算例 1



(b) 算例 2



(c) 算例 3

图 8 模拟-实验冰形结果对比

Fig. 8 Comparison of simulation-experiment ice shape results

表 2 翼型前缘冰形特征参数模拟误差

Table 2 Simulation error of ice shape characteristic parameters of airfoil leading edge %

算例	特征参数模拟误差				
	H_{upper}	H_{lower}	$H_{stagnation}$	θ_{upper}	θ_{lower}
1	8.7	2.7	4.2	0.7	3.0
2	−2.3	−9.7	−8.7	2.5	−1.9
3	−3.2	1.9	−2.7	−2.1	7.6

3 数值模拟结果分析

3.1 SLD 动力学行为影响分析

在静压为 101 325 Pa, 静温为−10.79 ℃, 来流速度为 200 n mile/h, 迎角为 3.8°, L_{WC}=1 g/m^3, 水滴直径为 500 μm 的条件下, 使用 Pilch & Erdman 模型计算 SLD 的变形与破碎, 使用 Wright-Potapczuk 模型计算 SLD 的飞溅与反弹, 并考虑 SLD 撞击壁面后飞溅反弹产生的二次液滴再次回到流场中对结冰的影响(注: 为书写简便, 称此考虑二次液滴的模型为“再入模型”)。

3.1.1 SLD 变形与破碎的影响

大水滴在流场中运动时, 周围压力分布不均匀, 气流在某一区域加速而在另外一区域减速, 这导致了气动力的产生。气动力使大水滴发生变形, 当变形到一定程度, 表面张力与黏性力将不再能维持大水滴的整体性, 原本的大液滴将破碎成数个较小的液滴。破碎后产生的二次液滴常常不再遵循其原有的运动大小和方向, 使得流场中原有的 L_{WC} 与水滴速度分布发生了局部改变, 如图 9 所示: 在加入 SLD 变形破碎模型之后, 翼型上缘局部 L_{WC}、x 方向水滴速度分量 V_1 显著减少, y 方向水滴速度分量 V_2 显著增加; 翼型尾缘局部 x 方向水滴速度分量显著增加。

选择翼型弦长中点平面, 对加入 SLD 变形破碎模型前后, L_{WC}、水滴速度分布变化进行定量分析, 结果如图 10 所示。可以看出, SLD 变形破碎模型对模拟结果产生了显著的影响, 在相同位置, L_{WC} 最多相差 9.76%; X 方向水滴速度最多相差 5.24%; Y 方向水滴速度最多相差 11.6%。

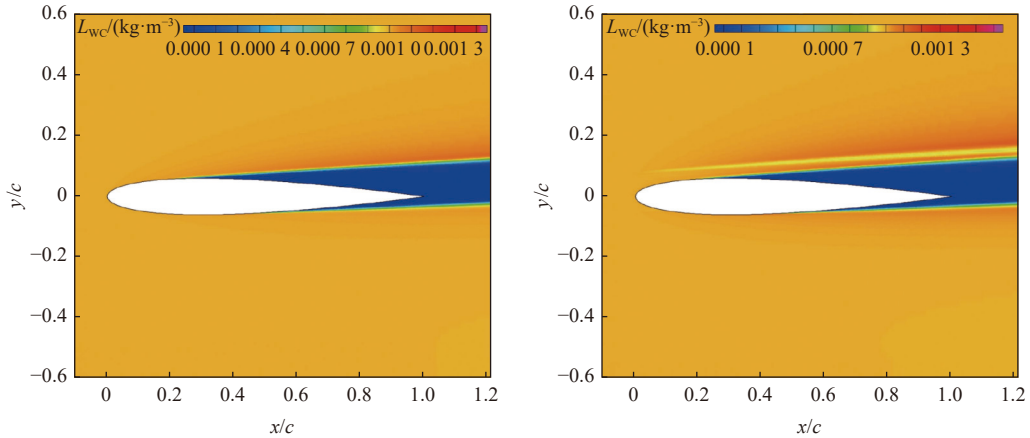
3.1.2 SLD 飞溅与反弹的影响

水滴在壁面与远场处单位面积上的质量通量比值称为水滴收集率 β, 图 11 比较了不同模型下水滴收集率 β 在翼型表面的变化。图中, s 为在 xOy 平面中, 翼型表面某一点沿着翼型表面到前缘的曲线长度, 正值表示点在翼型下缘, 负值表示点在翼型上缘; 无量纲数 s/c 用来定义点在翼型上的位置。

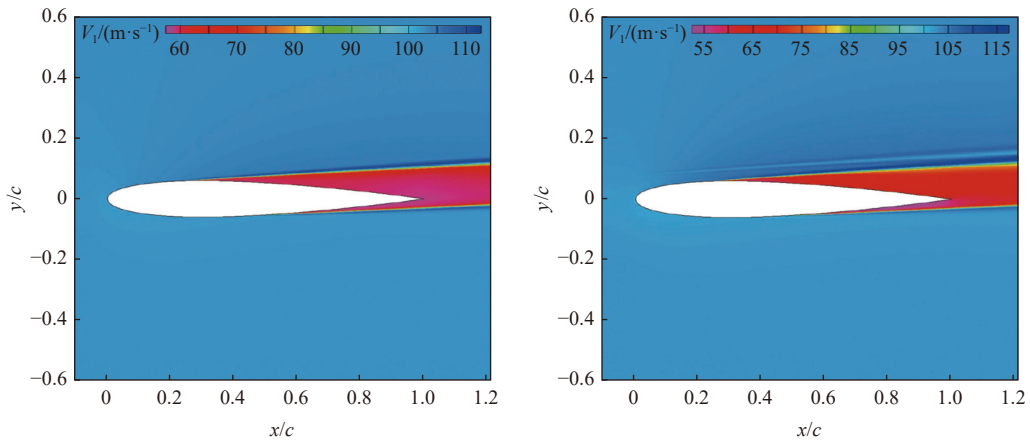
SLD 的飞溅行为显著降低了最大水滴收集率 β_{max}, 降低率达 15.56%; SLD 的反弹行为显著降低了翼型上下缘水滴撞击极限, 分别下降 24.36% 和 26.55%。

从图 11 中也可以看出, 飞溅反弹二次液滴再注入流场, 对翼型表面的水滴收集率 β 也产生了较大的影响。进一步比较有无飞溅反弹二次液滴对模拟结果的影响, 如图 12 所示。

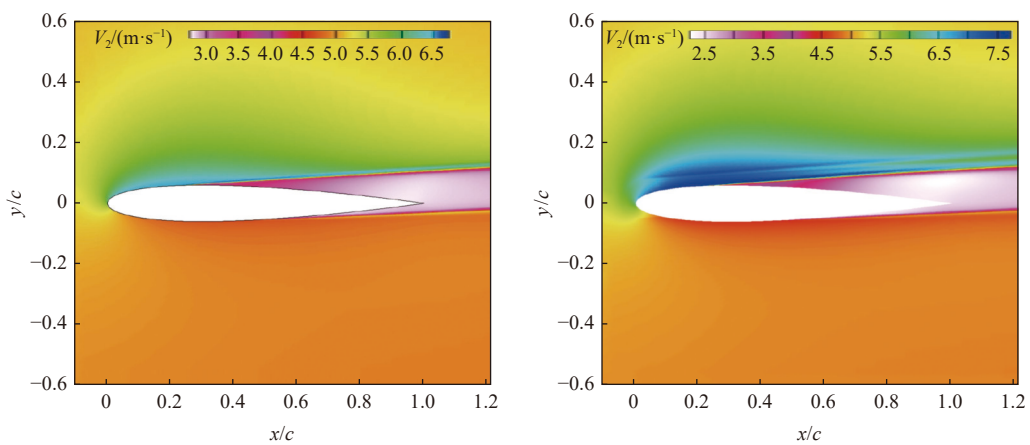
可以看出, 碰撞后二次液滴再注入流场, 显著增加了翼型附近局部的 L_{WC}、液滴速度。进一步比较翼型前后缘水滴平均直径的分布, 如图 13 所示, 不同模型下曲线的差异证明了这些较小的碰撞二



(a) L_{WC} 分布



(b) x 方向水滴速度分布



(c) y 方向水滴速度分布

图 9　无模型(左)与变形破碎模型(右)云图比较

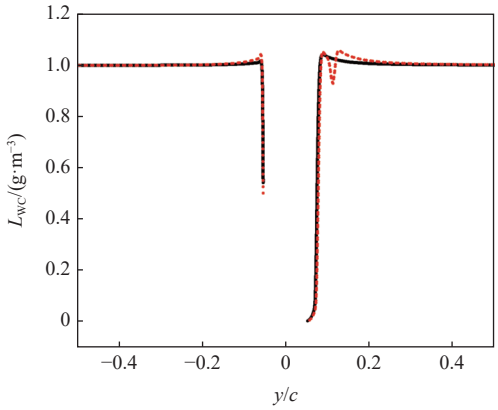
Fig. 9　Comparison of cloud images of no model (left) and deformed breakup model (right)

次液滴的存在, 水滴平均直径值在前缘、后缘局部最大降低了 50%、60%

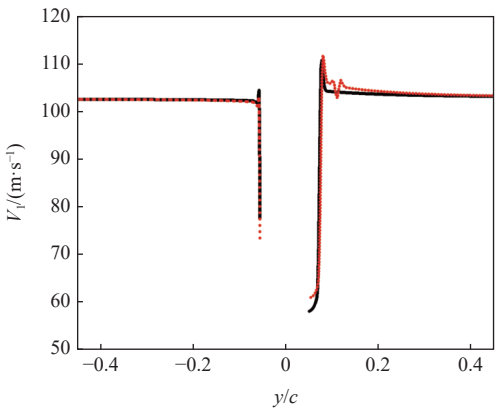
SLD 撞击壁面飞溅反弹产生的二次液滴再注入流场对翼型附近的 L_{WC}、液滴速度、液滴直径分布都产生了明显的影响。这些二次液滴随着气流继续向下游运动, 可能与下游部件发生多次撞击, 从而增加撞击部位的局部液滴收集率和液态水含量。因此, 如果不考虑该模型, 那么 SLDs 与壁面碰撞后产生的二次液滴将不会被引入到计算域中, 这将会对结冰模拟带来误差。

3.1.3　SLD 不同动力学行为对翼型前缘积冰冰形的影响

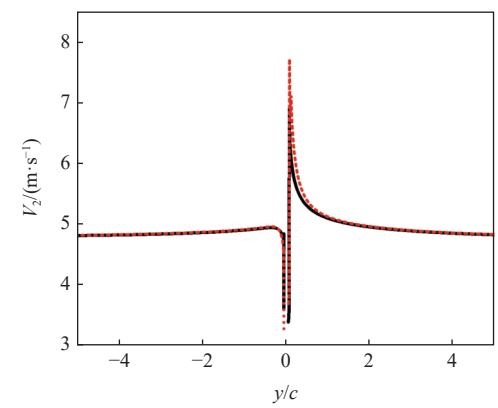
经过 240 s 结冰之后, 不同 SLD 动力学行为对积冰冰形产生了不同的影响, 如图 14 所示。可以看出, 虽然不同模型条件下生成的冰形大体相似, 但不同动力学行为造成的几何细节差异十分明显。进一步比较翼型前缘冰形特征参数, 如表 3 所示, 发现 SLD 的变形破碎行为使得上下结冰极限



(a) L_{WC} 分布



(b) x 方向水滴速度分布



(c) y 方向水滴速度分布

无模型 变形破碎模型

图 10 翼型弦长中点平面 L_{WC} 和水滴速度分布比较

Fig. 10 Comparison of L_{WC} and droplet velocity distribution in midpoint plane of airfoil chord length

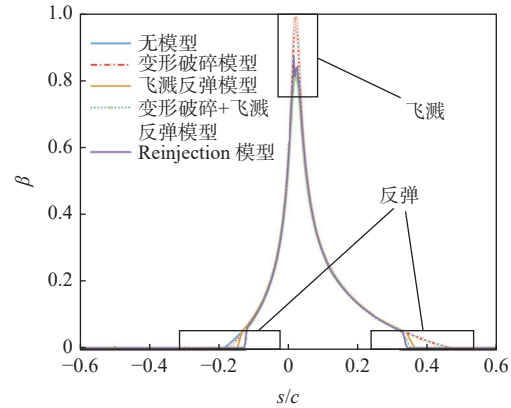


图 11 不同模型下水滴收集率 β 在翼型表面的变化

Fig. 11 Variation of droplet collection efficiency β on airfoil surface under different models

$S_{limit,upper}$、$S_{limit,lower}$ 分别降低了 2.83%、2.13%; SLD 的飞溅行为显著降低了位于翼型前缘的最大水滴收集率 β_{max}, 导致上冰角高度 H_{upper} 减小了 4.33%, 前缘积冰量减少 8.09%; SLD 的反弹行为显著降低了翼型上下缘水滴撞击极限, 导致下冰角角度 θ_{lower} 减少了 0.57%, 上下结冰极限 $S_{limit,upper}$、$S_{limit,lower}$ 分别降低了 30.69%、21.01%; 飞溅反弹产生的二次水滴再注入流场改变了上冰角下游部分羽毛状冰的形状和位置, 并使上下结冰极限 $S_{limit,upper}$、$S_{limit,lower}$ 分别增加了 6.14%、3.71%。

3.2 SLD 粒径效应影响分析

3.2.1 对冰形的影响

在考虑 SLD 变形破碎、飞溅反弹以及飞溅反弹二次水滴再注入流场的情况下, 对粒径分别为 20、60、80、100、120、140、200、300、400、500、1 000 μm(单尺度)的过冷水滴在 0.533 4 m NACA0012 翼型表面结冰现象进行模拟, 探究 SLD 粒径大小对翼型表面积冰冰形的影响, 计算工况如表 4 所示。

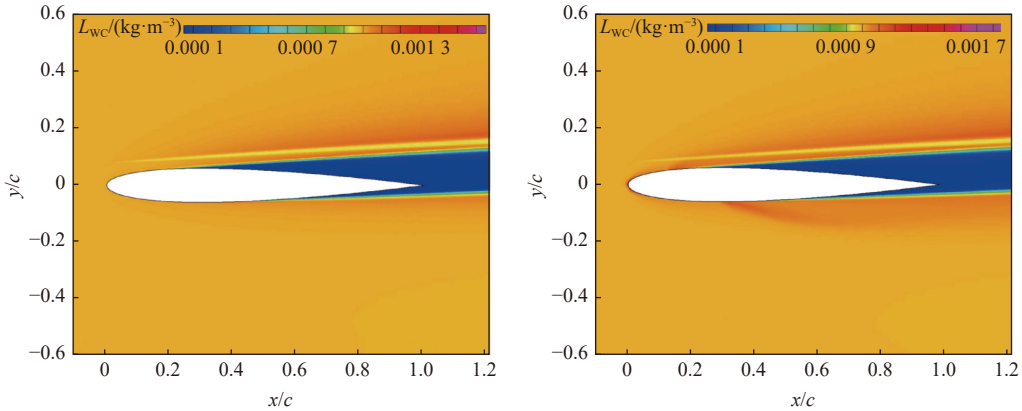
计算的翼型前缘积冰冰形如图 15 所示。按照粒径逐渐增大的顺序, 计算各个算例相较于前一算例的翼型前缘冰形特征参数变化, 结果如表 5 所示。

从表 5 中可以看出: 在 20～1 000 μm 范围内随着过冷水滴粒径的增大, 上冰角角度 θ_{upper} 先减小后增大; 下冰角角度 θ_{lower} 先增大后减小; 上下结冰极限 $S_{limit,upper}$、$S_{limit,lower}$ 一直在增大。这是由于水滴粒径越大, 撞击壁面的范围也就越大。此外, 当过冷水滴粒径增加到一定极限时, 撞击壁面后将不能产生足够的热交换使之完全冻结, 水滴中一部分液态水将会在空气剪应力、压力梯度、液体表面张力、离心力或重力的作用下产生溢流形成流动的水膜, 在水滴撞击区域外被冻结, 如图 16 所示。

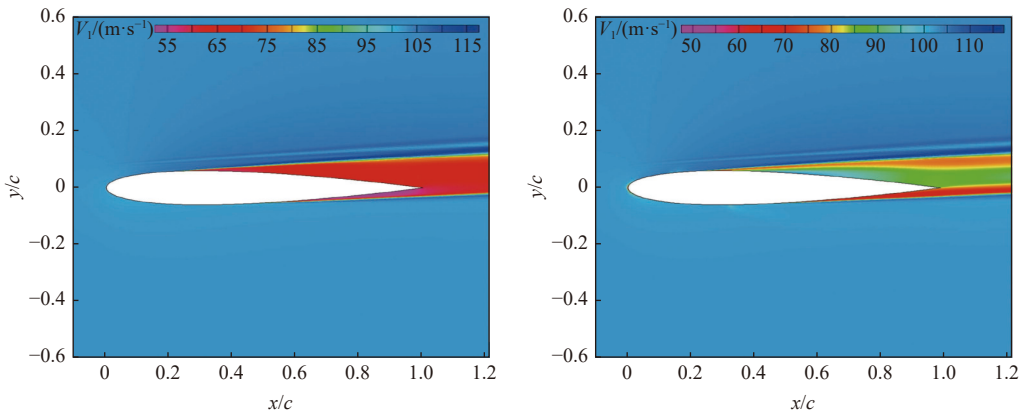
3.2.2 对气动特性的影响

在 3.2.1 节计算结果中, 选取粒径为 100 μm、300 μm 和 500 μm 的 SLD 翼型结冰模拟结果进行进一步分析, 探究 SLD 粒径大小对翼型气动性能的影响。比较这 3 个算例中, 翼型升力系数 C_L、阻力系数 C_D 在持续结冰过程中随时间的变化情况, 如图 17 所示。比较经过 210 s 持续结冰之后, 这 3 个算例中带冰翼型在相同工况下, 翼型升力系数 C_L、阻力系数 C_D 随迎角的变化, 如图 18 所示。

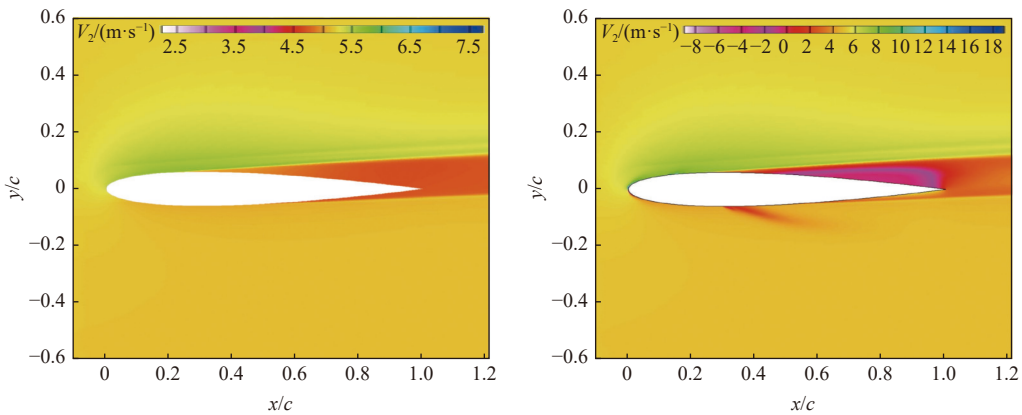
升力系数 C_L、阻力系数 C_D 与翼型几何特征密



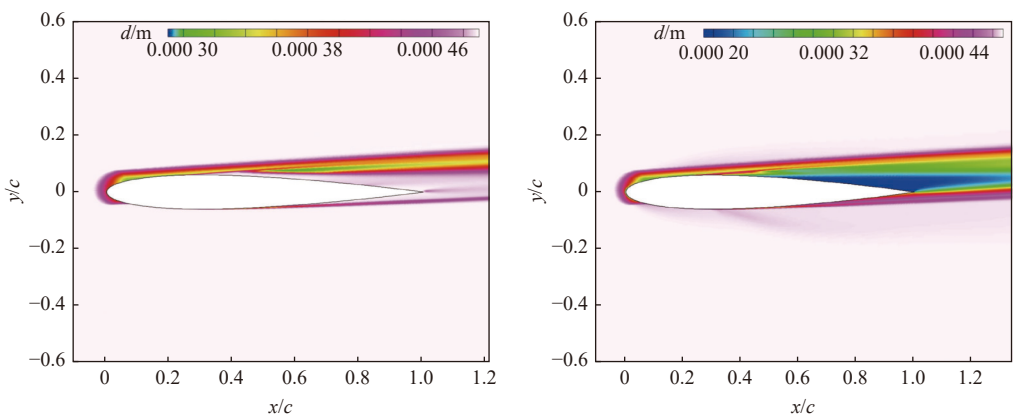
(a) L_{WC} 分布



(b) x 方向水滴速度分布



(c) y 方向水滴速度分布



(d) 水滴直径分布

图 12　变形破碎+飞溅反弹模型(左)与再入模型(右)云图比较

Fig. 12　Comparison of deformation and breakup + splashing and bouncing model (left) cloud images and reinjection model (right)

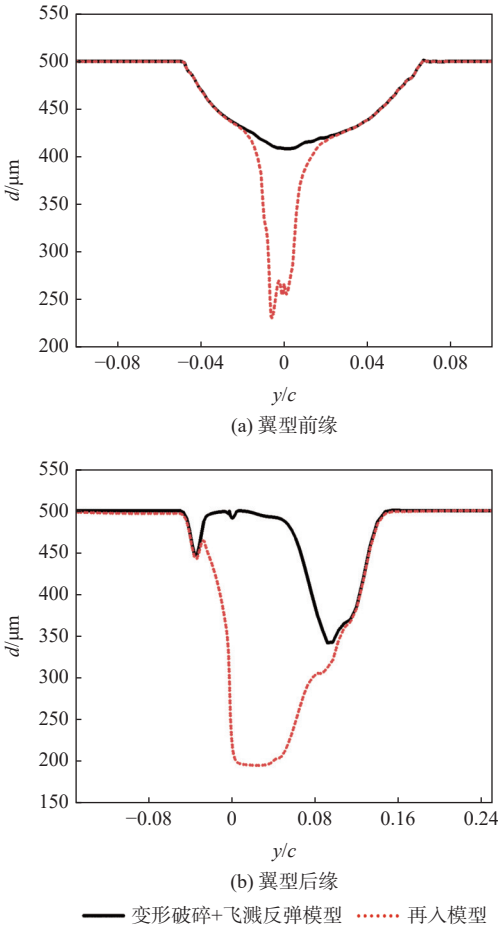


图 13 水滴平均直径分布比较

Fig. 13 Comparison of average diameter distribution of drop

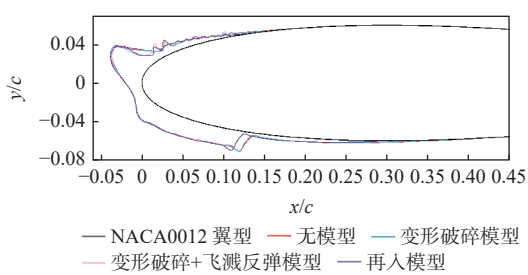


图 14 不同模型下翼型前缘积冰冰形

Fig. 14 Icing shape of airfoil leading edge ice accumulationunder different models

表 3 不同动力学行为对翼型前缘冰形特征参数的影响

Table 3 Effects of different dynamic behaviors on ice shape characteristic parameters at airfoil leading edge %

模型	特征参数变化					
	H_{upper}	H_{lower}	θ_{upper}	θ_{lower}	$S_{limit,upper}$	$S_{limit,lower}$
变形破碎	−0.73	0.09	−0.31	0.01	−2.83	−2.13
飞溅反弹	−4.33	3.52	1.41	−0.57	−30.69	−21.01
再入	2.97	−2.14	−0.18	0.03	6.14	3.71

切相关。上述 3 个算例积冰冰形虽然在细节上差异明显，但宏观来看带冰翼型几何形状总体相似。

表 4 数值模拟工况

Table 4 Working conditions of numerical simulation

来流速度/(n mile·h^{-1})	静温/℃	静压/Pa	迎角/(°)	L_{WC}/(g·m^{-3})	结冰时间/s
200	−10.79	101 325	3.8	1.00	210

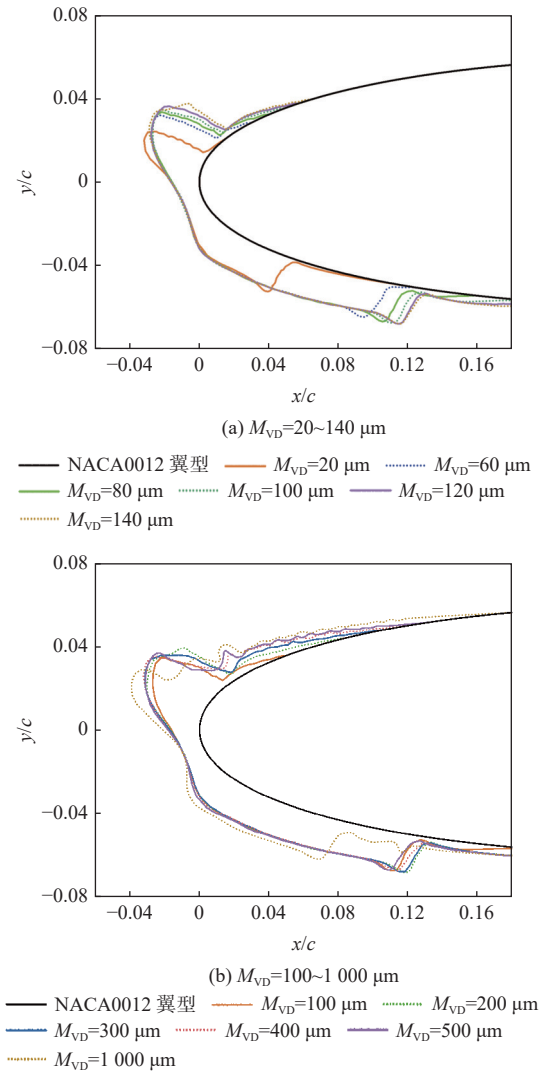


图 15 M_{VD}=20～1 000 μm SLD 翼型前缘积冰冰形

Fig. 15 M_{VD}=20～1 000 μm SLD ice shape at wing leading edge

因此在上述 3 个算例中，翼型升阻力系数随时间、迎角变化的数值、趋势虽略有差异但总体相近。经过 210 s 持续结冰之后，3 个算例翼型升力系数 C_L 分别减少 38.0%、38.9% 和 40.2%；阻力系数 C_D 分别增加 119.0%、107.1% 和 106.2%。

相比较之下，压力系数 C_P 在翼型表面的分布更为直观反映了积冰的几何特征变化，如图 19 所示。

吸力峰值(suction peak)是翼型上的最低气压点或 C_P 分布图上的最高点，通常位于翼型前缘之后，如图 20 所示。从压力系数 C_P 分布图中可以看出，由于上冰角的出现，3 种工况下的带冰翼型较干净

表 5　翼型前缘冰形特征参数的变化

Table 5　Variation of ice shape characteristic parameters at the airfoil leading edge

M_{VD}/μm	ΔH_{upper}%	ΔH_{lower}%	$\Delta\theta_{upper}$%	$\Delta\theta_{lower}$%	$\Delta S_{limit,upper}$%	$\Delta S_{limit,lower}$%
60	−1.52	−5.82	−10.35	5.99	287.84	6.36
80	1.22	0.63	−1.51	0.64	54.31	13.82
100	0.16	−0.59	−1.21	0.29	32.44	11.83
120	−0.76	−0.33	−2.35	0.16	14.56	5.41
140	4.33	−0.52	4.07	0.03	16.41	5.43
200	1.84	−0.33	−0.76	0.19	32.99	10.73
300	4.05	−0.12	1.01	−0.13	26.13	9.08
400	3.19	−1.53	2.53	−0.06	12.57	3.85
500	−1.60	0.93	0.98	−0.18	27.23	5.25
1 000	12.53	17.04	13.25	−3.81	67.04	12.13

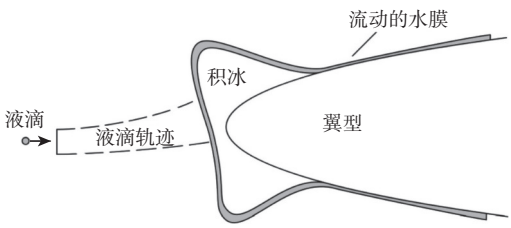
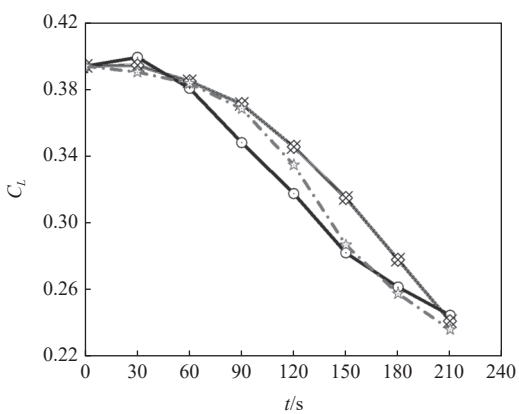
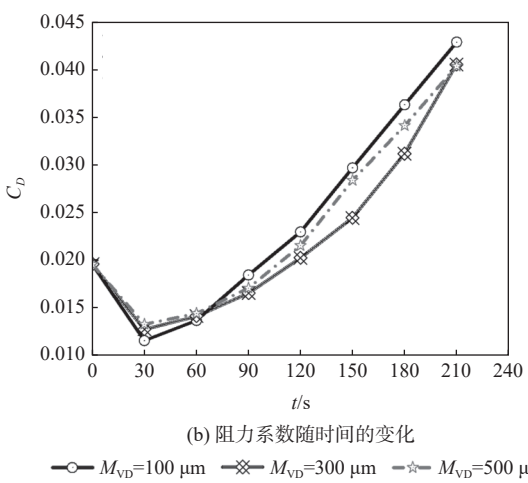


图 16　水膜溢流现象

Fig. 16　Water film runback phenomenon



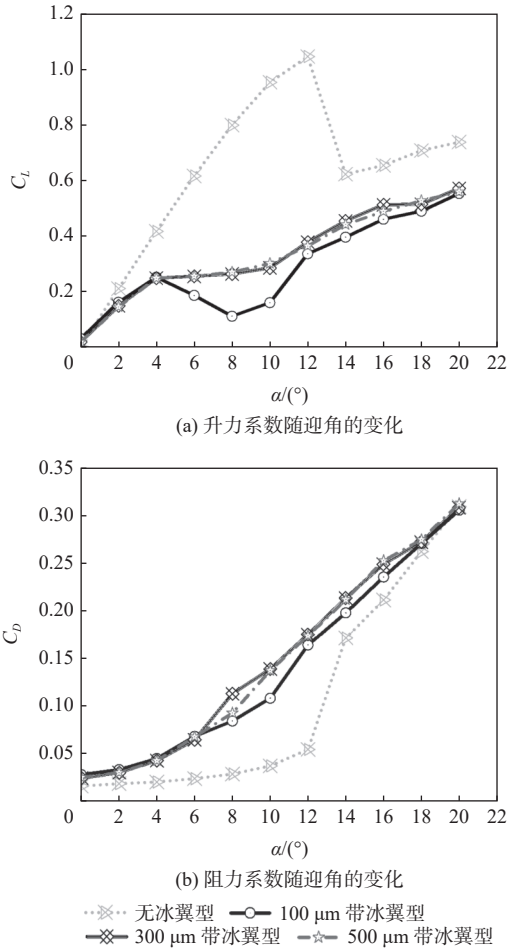
(a) 升力系数随时间的变化



(b) 阻力系数随时间的变化

图 17　翼型升阻力系数随时间变化

Fig. 17　Variation of lift coefficient and drag coefficient of airfoil with time



(a) 升力系数随迎角的变化

(b) 阻力系数随迎角的变化

图 18　带冰翼型升阻力系数随迎角变化

Fig. 18　Variation of lift coefficient and drag coefficient of icing airfoil with angle of attack

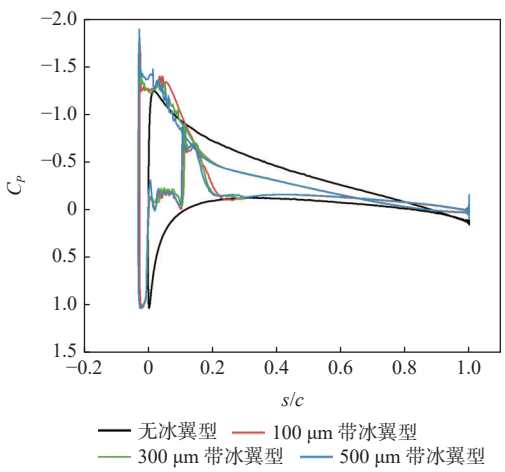


图 19　压力系数 C_P 在翼型表面的变化

Fig. 19　Variation of pressure coefficient C_P on airfoil surface

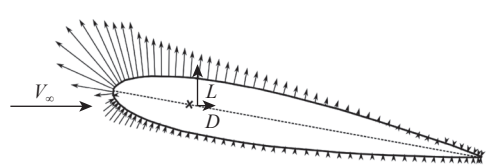
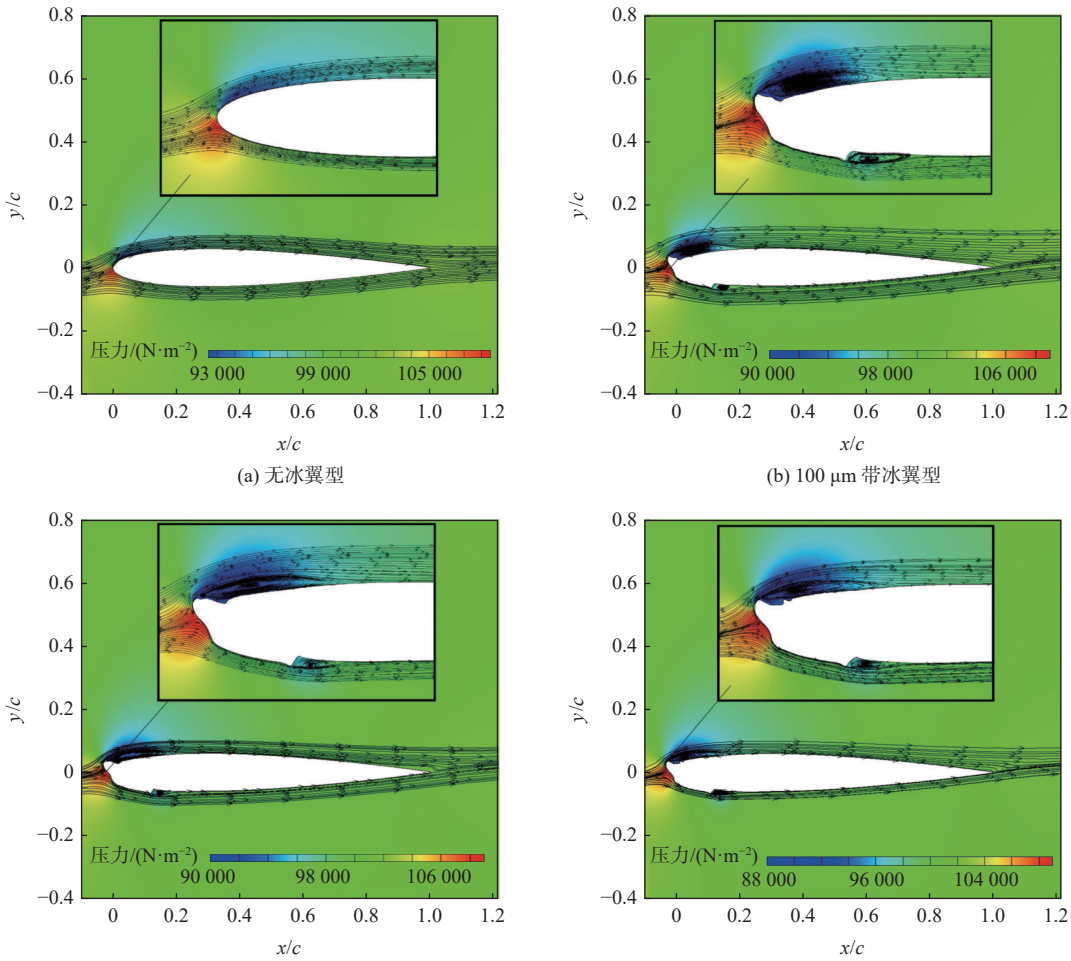


图 20　吸力峰值示意图

Fig. 20　Diagram of suction peak

翼型，吸力峰值分别增加37.85%、40.27%和51.62%。其次，由于积冰的生长造成了驻点位置的前移，3种工况下的带冰翼型较干净翼型，位于驻点的压力系数峰值分别前移了0.021 4c、0.025 5c、0.028 4c。此外，从尾缘处的压力系数分布变化可以看出，冰粗糙度在空气动力学中起着重要作用，带冰翼型与干净翼型表面粗糙度的分布差异影响了下游的边界层，从而导致了不同的尾缘分离行为[31]。

翼型周围对压力及流线分布进行分析，发现翼型表面积冰破坏了原本的流线型气动外形，在上下冰角下游区域产生涡流形成了明显的流动分离，并对翼型周围的压力分布产生了明显的影响，如图21所示。H_{upper}越大，逆压力梯度越大，流动分离越明显，吸力峰值也就越大。



(a) 无冰翼型　(b) 100 μm带冰翼型

(c) 300 μm带冰翼型　(d) 500 μm带冰翼型

图21　翼型周围压力及流线分布

Fig. 21　Pressure and streamline distribution around airfoil

4　结　论

本文针对SLD撞击翼型表面结冰现象给出了三维准稳态数值仿真方法，探究了SLD主要动力学行为：变形与破碎、飞溅与反弹、撞击二次液滴再入流场对结冰的影响，比较了SLD粒径大小对积冰冰形和气动特性退化的影响。通过分析计算结果，得到如下结论：

1）基于Pilch & Erdman变形破碎模型，对SLD变形破碎过程进行了模拟，探究了其对SLD碰撞特性与成冰过程的影响。分析发现SLD变形破碎行为改变了原有的水滴运动轨迹和撞击范围，使水滴的撞击极限和总的水滴收集效率减小，从而导致了翼型上下结冰极限以及总结冰量减小。

2）基于Wright-Potapczuk飞溅反弹模型，对SLD撞击壁面过程进行了模拟，探究SLD飞溅反弹行为对水滴碰撞特性与成冰过程的影响。分析发现SLD飞溅行为会降低前缘水滴收集率，导致前缘积冰量减小；SLD反弹行为会降低翼型上下缘水滴撞击极限，从而导致上下结冰极限减小。

3）飞溅反弹产生的二次水滴再注入流场对下游部位的结冰产生了影响，改变了上冰角下游部分羽毛状冰的形状和位置，并使上下结冰极限增加。

4）水滴尺寸的增加导致水滴具有更大的惯性，

偏离气流流线的能力也越大, 更多水滴可以与翼型表面碰撞, 撞击到壁面的位置也越远, 这带来了更大的水滴撞击极限, 进而导致结冰厚度和结冰范围的增加。

5) 与干净翼型相比, 带冰翼型空气动力学特性严重退化, 在相同迎角下, 升力更小、阻力更大、气动效率更低。

参考文献（References）

[1] FEDERAL AVIATION ADMINISTRATION. Airplane and engine certification requirements in supercooled large drop, mixed phase, and ice crystal icing conditions: RIN 2120–AJ34[S]. Washington, D. C. : Federal Aviation Administration (FAA), DOT, 2010.

[2] VILLEDIEU P, TRONTIN P, GUFFOND D, et al. SLD Lagrangian modeling and capability assessment in the frame of ONERA 3D icing suite[C]//4th AIAA Atmospheric and Space Environments Conference. Reston: AIAA 2012: 3132.

[3] HONSEK R, HABASHI W G, AUBE M S. Eulerian modeling of in-flight icing due to supercooled large droplets[J]. Journal of Aircraft, 2008, 45(4): 1290-1296.

[4] BOURGAULT Y, HABASHI W G, DOMPIERRE J, et al. A finite element method study of Eulerian droplets impingement models[J]. International Journal for Numerical Methods in Fluids, 2015, 29(4): 429-449.

[5] IULIANO E, MINGIONE G, PETROSINO F, et al. Eulerian modeling of SLD physics towards more realistic aircraft icing simulation[C]// AIAA Atmospheric and Space Environments Conference. Reston: AIAA, 2010: 7676.

[6] LEAL L G. Bubbles, drops and particles[J]. International Journal of Multiphase Flow, 1979, 5(3): 229-230.

[7] PILCH M, ERDMAN C A. Use of breakup time data and velocity history data to predict the maximum size of stable fragments for acceleration-induced breakup of a liquid drop[J]. International Journal of Multiphase Flow, 1987, 13(6): 741-757.

[8] HSIANG L P, FAETH G M. Near-limit drop deformation and secondary breakup[J]. Int Multiphase Flow, 1992, 18(5): 635-652.

[9] O'ROURKE P J, AMSDEN A A. The TAB method for numerical calculation of spray droplet breakup: No. 872089[R]. Toronto: SAE Technical Paper, 1987: 148-161.

[10] CLARK M M. Drop breakup in a turbulent flow-I. Conceptual and modeling considerations[J]. Chemical Engineering Science, 1988, 43(3): 671-679.

[11] IBRAHIM E A, YANG H Q, PRZEKWAS A J. Modeling of spray droplets deformation and breakup[J]. Journal of Propulsion and Power, 1993, 9(4): 651-654.

[12] BAI C, GOSMAN A D. Development of methodology for spray impingement simulation: No. 950283[R]. Detroit: SAE Transactions, 1995: 550-568.

[13] MUNDO C, SOMMERFELD M, TROPEA C. Droplet-wall collisions: Experimental studies of the deformation and breakup process[J]. International Journal of Multiphase Flow, 1995, 21(2): 151-173.

[14] MUNDO C, TROPEA C, SOMMERFELD M. Numerical and experimental investigation of spray characteristics in the vicinity of a rigid wall[J]. Experimental Thermal & Fluid Science, 1997, 15(3): 80-86.

[15] MUNDO C, SOMMERFELD M, TROPEA C. On the modeling of liquid sprays impinging on surfaces[J]. Atomization & Sprays, 1998, 8(6): 625-652.

[16] RUTKOWSKI A, WRIGHT W, POTAPCZUK M. Numerical study of droplet splashing and re-impingement[C]//41st Aerospace Sciences Meeting and Exhibit. Reston: AIAA 2003: 388.

[17] WRIGHT W. Further refinement of the LEWICE SLD model[C]// 44th AIAA Aerospace Sciences Meeting and Exhibit. Reston: AIAA, 2006: 464.

[18] NORDE E, HOSPERS J M, VAN DER WEIDE E, et al. Splashing model for impact of supercooled large droplets on a thin liquid film[C]//52nd Aerospace Sciences Meeting. Reston: AIAA, 2014: 738.

[19] BILODEAU D R, HABASHI W, BARUZZI G, et al. An Eulerian re-impingement model of splashing and bouncing supercooled large droplets[C]// 5th AIAA Atmospheric and Space Environments Conference. Reston: AIAA, 2013: 3058.

[20] TRUJILLO M F, MATHEWS W S, LEE C F, et al. Modelling and experiment of impingement and atomization of a liquid spray on a wall[J]. International Journal of Engine Research, 2000, 1(1): 87-105.

[21] 易贤, 李维浩, 王庄宇等. 过冷大水滴结冰过程中动力学行为的影响因素[C]//第十届全国流体力学学术会议论文摘要集. 杭州: 中国力学学会流体力学专业委员会, 2018, 1: 357.
Yi Xian, Li Weihao, Wang Zhuangyu et al. Influence factors of dynamic behavior of supercooled large water droplets during icing process [C]//Abstracts of the 10th National Conference on Fluid Mechanics.Hangzhou : Fluid Mechanics Committee of Chinese Society of Mechanics, 2018, 1 : 357(in Chinese).

[22] 张辰, 孔维梁, 刘洪. 大粒径过冷水滴结冰模拟破碎模型研究[J]. 空气动力学学报, 2013, 31(2): 144-150.
ZHANG C, KONG W L, LIU H. Research on simulation crushing model of large particle size supercooled water droplets icing[J]. Journal of Aerodynamics, 2013, 31(2): 144-150(in Chinese).

[23] 任靖豪, 王强, 刘宇, 等. 大型商用运输机机翼增升构型水滴撞击特性计算[J]. 空气动力学学报, 2021, 39(1): 52-58.
REN J H, WANG Q, LIU Y, et al. Calculation of droplet impact characteristics of wing lift configuration for large commercial transport aircraft[J]. Journal of Aerodynamics, 2021, 39(1): 52-58 (in Chinese).

[24] 周志宏, 易贤, 桂业伟, 等. 水滴撞击特性的高效计算方法[J]. 空气动力学学报, 2014, 32(5): 712-716.
ZHOU Z H, YI X, GUI Y W, et al. Efficient method for calculating droplet impact characteristics[J]. Journal of Aerodynamics, 2014, 32(5): 712-716(in Chinese).

[25] POTAPCZUK M G, TSAO J C, KING-STEEN L C. Bimodal SLD ice accretion on a NACA 0012 airfoil model[C]//9th AIAA Atmospheric and Space Environments Conference. Reston: AIAA, 2017: 4478.

[26] SUN F, HAN S, QIAN G, et al. A novel multi-objective optimization model for minute-in-trail strategy based on sector workload restriction[C]//2015 3rd International Conference on Artificial Intelli-

gence, Modelling and Simulation (AIMS). Piscataway: IEEE Press, 2015: 65-70.

[27] HSIANG L P, FAETH G M. Drop deformation and breakup due to shock wave and steady disturbances[J]. International Journal of Multiphase Flow, 1995, 21(4): 545-560.

[28] OZCER I, SWITCHENKO D, BARUZZI G S, et al. Multi-shot icing simulations with automatic re-meshing: No. 2019-01-1956[R]. Montreal: SAE Technical Paper, 2019: 956-972.

[29] FOULADI H, ALIAGA C N, HABASHI W G. Quasi-unsteady icing simulation of an oscillating airfoil[C]//7th AIAA Atmospheric and Space Environments Conference. Reston: AIAA, 2015: 3020.

[30] WRIGHT W, CHUNG J. Correlation between geometric similarity of ice shapes and the resulting aerodynamic performance degradation-A preliminary investigation using WIND[C]//38th Aerospace Sciences Meeting and Exhibit. Reston: AIAA , 2000: 97.

[31] BRAGG M, BROEREN A, ADDY H, et al. Airfoil ice-accretion aerodynamic simulation[C]//45th AIAA Aerospace Sciences Meeting and Exhibit. Reston: AIAA, 2007: 85.

Influence of dynamic behavior of supercooled large droplets on airfoil icing

LIU Feiyu[1], DENG Tian[1, 2, *]

(1. Sino-European Institute of Aviation Engineering, Civil Aviation University of China, Tianjin 300300, China;
2. Key Laboratory of Icing and Anti/De-Icing, China Aerodynamics Research and Development Center, Mianyang 621000, China)

Abstract: Supercooled large droplets (SLD) represent one of the most hazardous flight conditions. The unique dynamic behavior of large droplets, including deformation, fragmentation, splash, and rebound, poses challenges for accurately assessing SLD icing using traditional icing calculation methods. In this study, the Navier-Stokes method was employed to solve the flow field, the Euler method was used to calculate droplet impact, and the Shallow Water model was utilized to simulate ice accretion. The credibility of the proposed methodology was verified by comparing the results with NASA experimental data. The findings demonstrate that the dynamic behavior of SLD significantly influences icing and ice formation. Specifically, deformation and fragmentation alter the trajectory and impact range of droplets, reducing the droplet impact limit, resulting in a 2.83% and 2.13% decrease in upper and lower icing limits, respectively. Splashing reduces the collection efficiency of droplets near the stagnation point, resulting in an 8.09% reduction in leading-edge ice accretion. Rebound considerably lowers the droplet impact limit, leading to a 30.69% and 20.01% decrease in upper and lower icing limits, respectively. Moreover, the re-entry of secondary droplets into the flow field following rebound increases the upper and lower icing limits by 6.14% and 3.71%, respectively. Furthermore, the aerodynamic performance of the ice-contaminated airfoil significantly deteriorates compared to a clean airfoil. At the same angle of attack, the lift decreases, the drag increases, and the aerodynamic efficiency decreases.

Keywords: supercooled large droplets; dynamic behavior; airfoil icing; geometric characteristics of ice shape; aerodynamic characteristics

Received: 2022-04-02; **Accepted**: 2022-06-06; **Published Online**: 2022-09-15 08:01

URL: link.cnki.net/urlid/11.2625.V.20220913.1845.012

Foundation items: National Natural Science Foundation of China (U1933110); Open Project of Key Laboratory for Icing and Anti-icing (IADL20200305); Scientific Research Project of Tianjin Education Commission (2020KJ036)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0240

升空平台相对测量误差对定位精度的影响及定位算法

欧阳晓凤[1]，曾芳玲[1,*]，吕大千[1]，董天宝[1]，韩宜静[2]

(1. 国防科技大学 电子对抗学院，合肥 230037；2. 宇航动力学国家重点实验室，西安 710043)

摘 要：针对精准协同任务下的分布式升空平台高精度相对定位需求，基于超宽带（UWB）测距传感器研究了相对测距与定位模型，对UWB模块在非理想情况下存在的误差因素进行了推导分析，包括天线相位中心误差、测距刷新率和悬停稳定度。在此基础上对噪声成分进行了实测分析与滤波处理：通过Allan方差辨识UWB传感器噪声，并将有色噪声模型及估计参数引入到定位算法量测更新中，通过改进扩展卡尔曼滤波算法实现了分布式升空平台相对定位。仿真结果表明：在相对测量噪声服从非高斯分布情况下，所提算法相较于传统算法位置估计精度提高约23.43%。与传统算法相比，所提算法可减少测量色噪声影响，提高卫星导航拒止环境下的相对定位精度。

关 键 词：升空平台；相对定位；误差分析；超宽带传感器；Allan方差

中图分类号：TN953.7

文献标志码：A **文章编号**：1001-5965(2024)01-0187-11

准确的位置信息是分布式升空平台协同、空时频同步遂行任务的基础。美国国防高级研究计划局 (Defense Advanced Research Projects Agency, DARPA) 关于精确电子战系统的公告[1]中指出，电磁信号到达目标区域后的空间协同一致性影响电磁信号的空间聚焦效果，对于1 GHz以上的信号级协同，平台相对定位精度需达到厘米级别[2-3]。现有的实时动态（real time kinematic, RTK）、精密单点定位（precise point positioning, PPP）技术可在卫星导航系统的支撑下实现厘米级定位，但该类技术对全球定位系统 (global positioning system, GPS) 等卫星导航技术的依赖过于严重，在复杂地形环境和干扰环境下，难以完成精准的定位与导航服务。

考虑到分布式升空平台组成的编队中各平台间距较小，相对定位与导航传感器相对于绝对导航传感器能够实现更高的定位精准度、更强的抗干扰能力以及更低廉的成本价格，使得相对定位与导航技术已经成为近年来研究热点之一。现有卫星导航拒止条件下的相对定位技术研究大多聚焦于视觉信息辅助方式[3-5]，但视觉导航图像处理量巨大、实时性较差，且易受光线等环境因素的影响，无法满足无人机、系留气球等升空平台在复杂环境下的厘米级定位需求。因此，研究兼顾高精度和高灵活性的相对定位算法，是解决拒止条件下多平台协同问题的关键手段。

相对定位算法从升空平台控制方式上可分为集中式和分布式：集中式控制方式需要大量信息交互，平台数量规模受限，在通信带宽和距离受限的情况下难以获得理想的定位精度；分布式控制方式通过节点之间的局部协作，实现对多个升空平台的全局协作定位。目前，多源融合是实现分布式架构下升空平台相对定位的关键手段，通过多平台传感器之间的信息交互与协同，可弥补单平台惯性传感器件的误差，提高定位精度[6]。Koppanyi[7]、Johnson[8]和Dierenbach[9]等先后提出在室内、森林等卫星导航不可用环境下，通过GPS与无线定位融合方式对

收稿日期：2022-04-12；录用日期：2022-05-27；网络出版时间：2022-06-16 11:38
网络出版地址：link.cnki.net/urlid/11.2625.V.20220615.1935.001
*通信作者. E-mail：zella@ustc.edu

引用格式：欧阳晓凤，曾芳玲，吕大千，等. 升空平台相对测量误差对定位精度的影响及定位算法 [J]. 北京航空航天大学学报，2024，50（1）：187-197. OUYANG X F，ZENG F L，LYU D，et al. Positioning accuracy and localization algorithm with relative measurement errors in blast-off platforms [J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50（1）：187-197（in Chinese）.

“低小慢”平台提供连续定位信息。文献 [10] 则主要针对无人机编队过程中的相对导航问题，提出一种基于超宽带（ultra wide band, UWB）的相对差分定位算法，在卫星可见数目较少时定位精度显著提升，但在完全拒止情况下直接采用 UWB 定位结果精度仍然较差。文献 [11] 针对 UWB 传感器存在的测量误差进行了测算，并在预设车辆轨迹前提下提出了基于二维插值模型的运动补偿方法，有效提高了 UWB 定位精度，但本文算法难以推广到升空平台中。文献 [12-13] 主要对 UWB 通信信道进行优化，以满足精度、功耗和通信的鲁棒性需求，通过物理层参数调整将测距误差降低约 20%，但该方法的动态适应性较差。

为解决动态条件下 UWB 外部传感器与平台惯导等内部导航传感器的融合滤波，传感器量测噪声协方差参数估计的准确性对定位精度至关重要。目前，定位领域中对各类传感器测量色噪声的自适应估计算法主要通过引入遗忘因子调节量测误差[14-15]，但多源协同定位状态维度较高，算法稳定性和收敛性难以保证。Allan 方差[16-17] 已成为表征惯性传感器、卫星导航部件测量误差的一种常用方法，是一种直接快速确定非高斯噪声效应特性的时域分析技术，可与卡尔曼滤波、粒子滤波等滤波技术配合使用[18]，但需要调整 Kalman 算法以适应非白噪声[19]。

本文结合理论模型与实测数据，定量分析了 UWB 传感器应用于分布式升空平台相对定位场景下的测量性能、误差因素、噪声成分及参数，并基于 UWB/IMU（inertial measurement unit）组合架构给出了分布式升空平台厘米级相对定位算法和仿真结果。

1 相对定位原理与测量模型

1.1 相对定位原理

在卫星导航拒止环境下，惯性导航系统因具有自主、隐蔽及短航时高精度的优良特性，是分布式平台定位与导航的最佳选择。然而，惯性导航器件存在定位误差随时间而积累的缺陷，因此需要其他外部测量信息的辅助。外部传感器通过升空平台之间的通信交互，可通过相对距离测量计算邻近节点的位置，弥补内部惯性器件的误差，具体原理如图 1 所示。

图 1 为利用 Stage 智能体仿真软件[20] 模拟的 20 个分布式升空平台动力学模型，各个平台搭载同质内部传感器和外感传感器，内部传感器测量自运动、感知位置、姿态和方位的变化，外部传感器识别和测量与其他平台的相对距离和位置信息。如

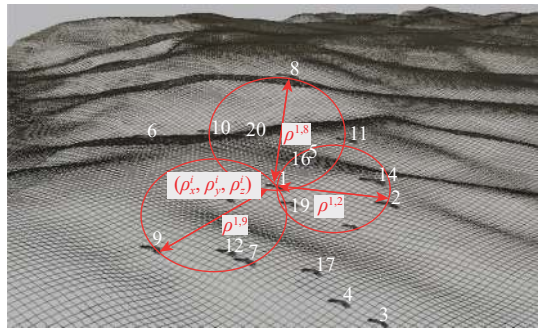


图 1 分布式升空平台相对定位原理

Fig. 1 Simplified relative motion model of distributed blast-off platform

图 1 所示，分布式升空平台一般是在一定区域内紧密编队，在高精度定位需求下可以采用无线测距定位技术，假设平台 i 实时接收回复平台 j 发送的测距信号，距离测量方程为

$$h^{ij}=\sqrt{\left(p_x^j-p_x^i\right)^2+\left(p_y^j-p_y^i\right)^2+\left(p_z^j-p_z^i\right)^2}\qquad i\neq j \tag{1}$$

式中：h^{ij} 为平台 i 和 j 几何距离理论值；$\left(p_x^i,p_y^i,p_z^i\right)$ 和 $\left(p_x^j,p_y^j,p_z^j\right)$ 分别为平台 i 和平台 j 的位置坐标。各平台将其姿态、相对位置坐标等信息传送到待定位平台接收模块，假设平台 i 位于未知坐标点 $\left(\hat{p}_x^i,\hat{p}_y^i,\hat{p}_z^i\right)$，邻近平台的位置估计值分别为 $\left(\hat{p}_x^j,\hat{p}_y^j,\hat{p}_z^j\right)$，2 个平台之间的距离测量值为 $\rho^{i,j}$，$j=1,2,\cdots,K$，K 为平台数量。通过式（2）的定位方程即可得到待定位成员的位置：

$$\sum_{j=1}^{K}\left(\rho^{i,j}-\sqrt{\left(\hat{p}_x^j-\hat{p}_x^i\right)^2+\left(\hat{p}_y^j-\hat{p}_y^i\right)^2+\left(\hat{p}_z^j-\hat{p}_z^i\right)^2}\right)=0 \tag{2}$$

理论上，式 (2) 在 $K\geqslant 4$ 的情况下都是可解的，但实际上由于传感器测量中的噪声和误差导致距离估计结果出现误差，式（2）右侧是非零的。因此，需要通过最小化误差函数来实现对未知成员的定位。考虑到 UWB 信号采用持续时间极短的窄脉冲，对时间、空间的分辨力都很强，在理想环境下可达到厘米级定位精度，因此本文以 UWB 模块为升空平台外部测量传感器，通过理论建模与硬件实测相结合的方式分析分布式平台间的相对定位误差，并提出误差修正与定位算法。

1.2 UWB 测量模型与性能分析

UWB 测距模块通常采用到达时间（time of arrival, TOA）、到达时间差（time difference of arrival, TDOA）和飞行时间 (time of flight, TOF) 测距技术[21-23]，其中 TOF 技术在有效距离、抗干扰能力、硬件实现和测距误差等方面的综合表现更具优势，且采用双向

测距 (two-way, TW) 算法不需要进行节点间的数据同步, TW-TOF 测量原理如图 2 所示。模块 1 发出的脉冲信号经过 TOF 后由模块 2 接收, 模块 2 将同时接收到直达信号和多径信号, 利用前沿检测 (leading edge detection, LED) 技术识别脉冲前沿。经过精确、恒定的转发时间 (turn around time, TAT) 后发送响应脉冲。模块 1 在 TOF 之后接收该脉冲及多径信号, 调整 LED 偏移量和 TAT 后完成整个对话时间 (conversation time, CT) 的测量。

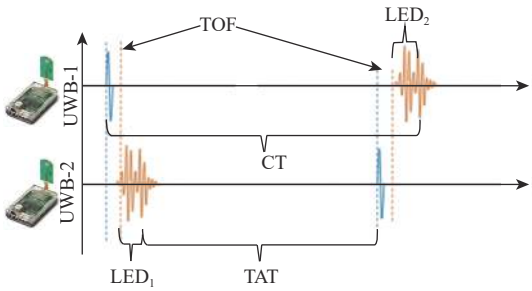


图 2　TW-TOF 测量原理及信号波形

Fig. 2　Measuring principle and signal waveform of TW-TOF

通过上述方法, 计算得到 k 时刻的相对距离 $\rho_k^{i,j}$ 表达式为

$$\rho_k^{i,j}=\frac{1}{2}(t_{\mathrm{CT}}(k)-t_{\mathrm{TAT}}(k)-t_{\mathrm{LED1}}(k)-t_{\mathrm{LED2}}(k))\cdot c+\zeta \quad (3)$$

式中: c 为电磁波传播速度; t_{CT} 为测量完整对话时间; t_{TAT} 为信号转发时间; t_{LED1} 为模块 1 检测到的 LED 时延; t_{LED2} 为模块 2 检测到的 LED 时延; ζ 为天线、温度和信道干扰等因素造成的测量噪声, 一般被建模为与距离无关的白噪声 $N(b_0,\sigma)$, 在实际应用之前可根据模块出厂参数对偏差项 b_0 和标准差 σ 进行校准。

2　相对测量误差分析

2.1　天线相位中心

分布式升空平台为完成协同任务, 需要获取任务载荷天线相位中心的位置, 而利用 UWB 模块测得的是测量天线相位中心位置, 两者之间的固定偏差可以事先测定, 但升空平台的运动状态变化会导致天线相位中心发生偏移, 其对相对距离测量和相对定位的影响无法事先测得和消除, 本节对这一误差进行分析和估计。

1.1 节中给出了基于 UWB 传感器的观测方程, 解算某一升空平台位置可对式 (2) 进行观测方程线性化, 然后利用最小二乘法估计节点位置。不失一般性, 设此时平台 K 的位置 $\hat{\boldsymbol{X}}^K=\left[\hat{p}_x^K,\hat{p}_y^K,\hat{p}_z^K\right]$ 待解算, 将式 (2) 构成的方程组中各观测方程与第 K−1 个方程作差, 并设 $\Delta\boldsymbol{\varepsilon}$ 为误差向量, 由此可将观测方程变换为式(4)矩阵形式:

$$\Delta\boldsymbol{\varepsilon}=\boldsymbol{A}\hat{\boldsymbol{X}}^K-\delta\boldsymbol{\rho} \quad (4)$$

式中,

$$\delta\boldsymbol{\rho}=\begin{bmatrix}\left(\rho^{K,K-1}\right)^2-\left(\rho^{K,1}\right)^2+\left(\hat{p}_x^1\right)^2-\left(\hat{p}_x^{K-1}\right)^2+\left(\hat{p}_y^1\right)^2-\left(\hat{p}_y^{K-1}\right)^2+\left(\hat{p}_z^1\right)^2-\left(\hat{p}_z^{K-1}\right)^2\\ \left(\rho^{K,K-1}\right)^2-\left(\rho^{K,2}\right)^2+\left(\hat{p}_x^2\right)^2-\left(\hat{p}_x^{K-1}\right)^2+\left(\hat{p}_y^2\right)^2-\left(\hat{p}_y^{K-1}\right)^2+\left(\hat{p}_z^2\right)^2-\left(\hat{p}_z^{K-1}\right)^2\\ \vdots\\ \left(\rho^{K,K-1}\right)^2-\left(\rho^{K,K-2}\right)^2+\left(\hat{p}_x^{K-2}\right)^2-\left(\hat{p}_x^{K-1}\right)^2+\left(\hat{p}_y^{K-2}\right)^2-\left(\hat{p}_y^{K-1}\right)^2+\left(\hat{p}_z^{K-2}\right)^2-\left(\hat{p}_z^{K-1}\right)^2\end{bmatrix}$$

$$\boldsymbol{A}=-2\begin{bmatrix}\left(\hat{p}_x^1-\hat{p}_x^{K-1}\right)\left(\hat{p}_y^1-\hat{p}_y^{K-1}\right)\left(\hat{p}_z^1-\hat{p}_z^{K-1}\right)\\ \left(\hat{p}_x^2-\hat{p}_x^{K-1}\right)\left(\hat{p}_y^2-\hat{p}_y^{K-1}\right)\left(\hat{p}_z^2-\hat{p}_z^{K-1}\right)\\ \vdots\\ \left(\hat{p}_x^{K-2}-\hat{p}_x^{K-1}\right)\left(\hat{p}_y^{K-2}-\hat{p}_y^{K-1}\right)\left(\hat{p}_z^{K-2}-\hat{p}_z^{K-1}\right)\end{bmatrix}$$

构造误差平方函数 $\Delta\boldsymbol{\varepsilon}^{\mathrm{T}}\Delta\boldsymbol{\varepsilon}$ 并利用加权最小二乘法[24]得到最优位置估计结果:

$$\hat{\boldsymbol{X}}^K=(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{W}\delta\boldsymbol{\rho} \quad (5)$$

式中: $\boldsymbol{W}$ 为加权矩阵。为分析定位解算过程中天线相位中心误差的影响, 此处在图 1 基础上以 2 个升空平台相对关系为例, 推导定位误差与测距误差两者的关系。

图 3 所示双平台测量天线空间关系, 为描述升空平台及其测量天线的空间关系, 以平台 K 所在位置为坐标原点建立坐标系。设 2 个升空平台配置的 UWB 测量天线与水平面呈垂直角度安装, 天线相位中心在水平方向和垂直方向夹角分别为 $\varphi^{K,i}$ 和 $\theta^{K,i}$, 则相对距离测量值 $\hat{\rho}^{K,i}$ 满足:

$$\partial\hat{\rho}^{K,i}/\partial\boldsymbol{X}=(-\sin\theta^{K,i}\cos\varphi^{K,i},-\sin\theta^{K,i}\sin\varphi^{K,i},-\cos\theta^{K,i})\triangleq(\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3) \quad (6)$$

式中: 方向矢量 $\boldsymbol{e}_1$、$\boldsymbol{e}_2$ 和 $\boldsymbol{e}_3$ 满足 $\left|\boldsymbol{e}_1^2+\boldsymbol{e}_2^2+\boldsymbol{e}_2^3\right|^2=1$。假设平台 K 天线相位中心偏移量为 $\delta\boldsymbol{\rho}\left(\theta^{K,i},\varphi^{K,i}\right)$, 当

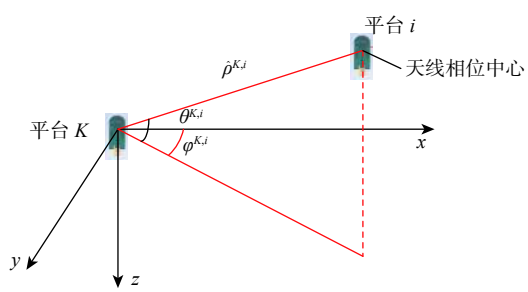


图 3　双平台测量天线空间关系

Fig. 3　Geometric relationship between two-platform measuring antennas

各平台观测量无相关性且权重相等时，结合式(6)可知，定位误差$\delta\boldsymbol{X}\left(\theta^{K,i},\varphi^{K,i}\right)$的表达式为

$$\delta\boldsymbol{X}\left(\theta^{K,i},\varphi^{K,i}\right)=\boldsymbol{N}^{-1}\boldsymbol{A}^{\mathrm{T}}\delta\boldsymbol{\rho}\left(\theta^{K,i},\varphi^{K,i}\right) \tag{7}$$

$$\boldsymbol{N}=\begin{bmatrix}[\boldsymbol{e}_1^2] & [\boldsymbol{e}_1\boldsymbol{e}_2] & [\boldsymbol{e}_1\boldsymbol{e}_3]\\ [\boldsymbol{e}_2\boldsymbol{e}_1] & [\boldsymbol{e}_2^2] & [\boldsymbol{e}_2\boldsymbol{e}_3]\\ [\boldsymbol{e}_3\boldsymbol{e}_1] & [\boldsymbol{e}_3\boldsymbol{e}_2] & [\boldsymbol{e}_3^2]\end{bmatrix} \tag{8}$$

对于在平台K周围呈随机分布的多个升空平台，设$\theta^{K,i}\in[0,\theta_{\max}]$，$\theta_{\max}$为分布式平台天线相位中心垂直方向最大值，$\theta^{K,i}$的概率密度函数用$f\left(\theta^{K,i}\right)$表示，所有观测数据之和$[\cdots]$可等效为积分运算$\int_0^{\theta_{\max}}\int_{\pi/2-\alpha}^{\pi/2+\alpha}\cdots f\left(\theta^{K,i}\right)\sin\theta^{K,i}\mathrm{d}\theta^{K,i}\mathrm{d}\varphi^{K,i}$，俯仰角偏差量$\alpha\in[0,\pi]$。

$$\boldsymbol{N}=\begin{bmatrix}\left(\alpha-\frac{1}{2}\sin(2\alpha)\right)\int_0^{\theta_{\max}}f(x)\sin^3x\mathrm{d}x & 0 & 0\\ 0 & \left(\alpha+\frac{1}{2}\sin(2\alpha)\right)\int_0^{\theta_{\max}}f(x)\sin^3x\mathrm{d}x & \sin\alpha\int_0^{\theta_{\max}}f(x)\sin^2x\cos x\mathrm{d}x\\ 0 & 0 & 2\alpha\int_0^{\theta_{\max}}f(x)\sin x\cos^2x\mathrm{d}x\end{bmatrix} \tag{9}$$

$$\boldsymbol{A}^{\mathrm{T}}\delta\boldsymbol{\rho}\left[[\boldsymbol{e}_1\delta\boldsymbol{\rho}]\quad[\boldsymbol{e}_2\delta\boldsymbol{\rho}]\quad[\boldsymbol{e}_3\delta\boldsymbol{\rho}]\right]^{\mathrm{T}}=-\int_0^{\theta_{\max}}\delta\boldsymbol{\rho}\left(\theta^{K,i},\varphi^{K,i}\right)f\left(\theta^{K,i}\right)\begin{bmatrix}\sin^2\left(\theta^{K,i}\right)\cos\left(\varphi^{K,i}\right)\\ \sin^2\left(\theta^{K,i}\right)\sin\left(\varphi^{K,i}\right)\\ \sin\left(\theta^{K,i}\right)\cos\left(\theta^{K,i}\right)\end{bmatrix}\mathrm{d}\theta^{K,i} \tag{10}$$

将式(9)和式(10)代入式(8)可知，当$\alpha=\pi$时，天线相位中心偏移并不影响水平方向的定位精度，只影响高程定位精度。此时，分布式升空平台在高程方向上的分布范围$\theta_{\max}$越小，天线相位中心偏移带来的误差因素越小，当各个平台在同一高度面分布时该项误差可忽略不计。当$\alpha\neq\pi$时，$\delta\boldsymbol{\rho}$取决于天线相位误差函数，本文以UWB微带天线[25]为例分析其误差函数：

$$\delta\boldsymbol{\rho}\left(\theta^{K,i},\varphi^{K,i}\right)=\boldsymbol{A}\cdot\sin\theta^{K,i}\cdot\cos\left(\varphi^{K,i}-a_0\right) \tag{11}$$

式中：a_0为天线方位角。当所有平台UWB天线俯仰和方位角相同时，将式(11)代入式(8)可将定位误差简化为

$$\delta\boldsymbol{X}^K=\boldsymbol{A}\begin{bmatrix}\cos a_0\\ \sin a_0\\ 0\end{bmatrix} \tag{12}$$

由式(12)可知，天线相位中心误差偏移影响水平定位精度，而对高程测量无影响，且与其他平台位置无关。因此，对于有厘米级定位需求的升空平台，各平台应尽可能悬停或系留于同一高度面，并在所有平台配置相同的已标校天线，来降低该项误差对定位的影响。

2.2 测距刷新率

理想情况下，升空平台在悬停或系留状态下执行任务时绝对静止，定位精度与刷新率无关；然而在实际情况下，受到空气气流等环境影响，升空平台仍存在微小运动，当测距和定位刷新率不足时，相对速度$\Delta\boldsymbol{v}_k^{ij}$和速度测量误差$\delta\boldsymbol{v}_k$将导致定位精度下降，从而对协同任务效率造成影响。下面以升空平台i和j为例，分析UWB模块测距刷新率F对平台定位精度的影响。

当测距刷新率为F，$\boldsymbol{X}_k^i$和$\boldsymbol{X}_k^j$分别为任意时刻k平台i和平台j的位置坐标；T表示每次测量更新的时间间隔，则k和$k+1$时刻平台i对平台j的相对位置关系可表示为

$$\begin{cases}\boldsymbol{X}_{k+1}^i-\boldsymbol{X}_{k+1}^j=\boldsymbol{X}_k^i-\boldsymbol{X}_k^j+T\Delta\boldsymbol{v}_k^{ij}\\ \left(\rho_k^{ij}\right)^2=\left\|\boldsymbol{X}_k^i-\boldsymbol{X}_k^j\right\|^2\end{cases} \tag{13}$$

考虑测量噪声时，估计得到的相对位置关系可表示为

$$\hat{\boldsymbol{X}}_{k+1}^i-\hat{\boldsymbol{X}}_{k+1}^j=\hat{\boldsymbol{X}}_k^i-\hat{\boldsymbol{X}}_k^j+T\left(\Delta\boldsymbol{v}_k^{ij}+\delta\boldsymbol{v}_k\right)+T\left(\Delta\boldsymbol{v}_k^{ij}+\delta\boldsymbol{v}_k\right)\times\left[\rho_k^{ij}\dot{\rho}_k^{ij}-\left(\Delta\boldsymbol{v}_k^{ij}+\delta\boldsymbol{v}_k\right)^{\mathrm{T}}\left(\hat{\boldsymbol{X}}_k^i-\hat{\boldsymbol{X}}_k^j\right)\right] \tag{14}$$

其中：$T\geqslant\frac{1}{F}$表示位置测量更新时间，当各个平台之间可同时测距时$T\doteq\frac{1}{F}$。令k时刻间隔平台i对平台j的相对位置估计误差为$\tilde{\boldsymbol{X}}_k^{ij}=\left(\boldsymbol{X}_k^i-\boldsymbol{X}_k^j\right)-\left(\hat{\boldsymbol{X}}_k^i-\hat{\boldsymbol{X}}_k^j\right)$，并对式(13)中第2式两边求微分可得到关系式$\rho_k^{ij}\cdot\dot{\rho}_k^{ij}=\left(\Delta\boldsymbol{v}_k^{ij}\right)^{\mathrm{T}}\left(\boldsymbol{X}_k^i-\boldsymbol{X}_k^j\right)$，此时误差变化过程可表示为

$$\tilde{\boldsymbol{X}}_{k+1}^{ij}=\left[\boldsymbol{I}-T(\Delta\boldsymbol{v}_k^{ij}+\delta\boldsymbol{v}_k)(\Delta\boldsymbol{v}_k^{ij}+\delta\boldsymbol{v}_k)^{\mathrm{T}}\right]\tilde{\boldsymbol{X}}_k^{ij}+T\left[\delta\boldsymbol{v}_k-(\Delta\boldsymbol{v}_k^{ij}+\delta\boldsymbol{v}_k)\delta\boldsymbol{v}_k^{\mathrm{T}}\tilde{\boldsymbol{X}}_k^{ij}\right]\triangleq\boldsymbol{D}_k^{ij}\tilde{\boldsymbol{X}}_k^{ij}+\delta\boldsymbol{v}_k^i \tag{15}$$

式中：$\boldsymbol{I}$为单元矩阵。定义李雅普诺夫函数$J\left(\tilde{\boldsymbol{X}}_k^{ij}\right)=\left(\tilde{\boldsymbol{X}}_k^{ij}\right)^{\mathrm{T}}\tilde{\boldsymbol{X}}_k^{ij}$[26]，基于此对相对位置误差上限进行分析推导如下：

$$\Delta J_k^{k+1}=J\left(\tilde{\boldsymbol{X}}_{k+1}^{ij}\right)-J\left(\tilde{\boldsymbol{X}}_k^{ij}\right)=\left(\tilde{\boldsymbol{X}}_{k+1}^{ij}\right)^{\mathrm{T}}\left(\tilde{\boldsymbol{X}}_{k+1}^{ij}\right)-\left(\tilde{\boldsymbol{X}}_k^{ij}\right)^{\mathrm{T}}\left(\tilde{\boldsymbol{X}}_k^{ij}\right)=\left(T^2\left\|\Delta\boldsymbol{v}_k^{ij}+\delta\boldsymbol{v}_k\right\|^2-2T\right)\left(\tilde{\boldsymbol{X}}_k^{ij}\right)^{\mathrm{T}}\cdot\left(\Delta\boldsymbol{v}_k^{ij}+\delta\boldsymbol{v}_k\right)\left(\Delta\boldsymbol{v}_k^{ij}+\delta\boldsymbol{v}_k\right)^{\mathrm{T}}\left(\tilde{\boldsymbol{X}}_k^{ij}\right)+2(\delta v_k)^{\mathrm{T}}\boldsymbol{D}_k^{ij}\tilde{\boldsymbol{X}}_k^{ij}+\left\|\delta\boldsymbol{v}_k^i\right\|^2 \tag{16}$$

假设升空平台间距最大值、速度测量误差最大值、悬停相对速度最大值分别为$\bar{\rho}$、$\bar{\delta}$和$\bar{v}$, 速度测量误差$\|\delta\boldsymbol{v}_k\|\leqslant\bar{\delta}$, 悬停状态下的相对速度小量满足$\left\|\Delta\boldsymbol{v}_k^{ij}\right\|\leqslant\bar{v}$且各方向分量线性独立, 即$\exists\underline{v}>0$, s.t. $\left(\Delta\boldsymbol{v}_k^{ij}+\boldsymbol{\varepsilon}_k\right)\left(\Delta\boldsymbol{v}_k^{ij}+\boldsymbol{\varepsilon}_k\right)^{\mathrm{T}}\geqslant\underline{v}\boldsymbol{I}$, 其中, $\underline{v}$表示相对速度小量的上界, 一般无风或微风状态下悬停下无人机相对速度为 0.1～1 m/s, 则有

$$\left\|\delta\boldsymbol{v}_k^i\right\|\leqslant T\bar{\delta}[1+(2\bar{v}+\bar{\delta})\bar{\rho}] \tag{17}$$

$$2(\delta\boldsymbol{v}_k)^{\mathrm{T}}\boldsymbol{D}_k^{ij}\tilde{\boldsymbol{X}}_k^{ij}\leqslant 2T\bar{\delta}\left[1+(2\bar{v}+\bar{\delta})\bar{\rho}\right]\left\|\tilde{\boldsymbol{X}}_k^{ij}\right\| \tag{18}$$

$$\left(T^2\left\|\Delta\boldsymbol{v}_k^{ij}+\delta\boldsymbol{v}_k\right\|^2-2T\right)\left(\tilde{\boldsymbol{X}}_k^{ij}\right)^{\mathrm{T}}\left(\Delta\boldsymbol{v}_k^{ij}+\delta\boldsymbol{v}_k\right)\left(\Delta\boldsymbol{v}_k^{ij}+\delta\boldsymbol{v}_k\right)^{\mathrm{T}}\cdot\left(\tilde{\boldsymbol{X}}_k^{ij}\right)\leqslant\left[T^2\left(2\bar{v}+\bar{\delta}\right)^2-2T\right]\left(\tilde{\boldsymbol{X}}_k^{ij}\right)^{\mathrm{T}}\cdot\left(\Delta\boldsymbol{v}_k^{ij}+\delta\boldsymbol{v}_k\right)\left(\Delta\boldsymbol{v}_k^{ij}+\delta\boldsymbol{v}_k\right)^{\mathrm{T}}\left(\tilde{\boldsymbol{X}}_k^{ij}\right)\leqslant\underline{v}\left[T^2\left(2\bar{v}+\bar{\delta}\right)^2-2T\right]\left\|\tilde{\boldsymbol{X}}_k^{ij}\right\|^2 \tag{19}$$

代入式 (16) 可得

$$\Delta J_k^{k+1}\leqslant\underline{v}\left[T^2(2\bar{v}+\bar{\delta})^2-2T\right]\left\|\tilde{\boldsymbol{X}}_k^{ij}\right\|^2+2T\bar{\delta}\left[1+\left(2\bar{v}+\bar{\delta}\right)\bar{\rho}\right]\left\|\tilde{\boldsymbol{X}}_k^{ij}\right\|+\left(T\bar{\delta}\left[1+\left(2\bar{v}+\bar{\delta}\right)\bar{\rho}\right]\right)^2 \tag{20}$$

解上述一元二次不等式, 可知当定位刷新时间$0<T<\dfrac{1}{\left(2\bar{v}+\bar{\delta}\right)^2}$时, 相对位置估计误差存在上界:

$$\left\|\tilde{\boldsymbol{X}}_k^{ij}\right\|\leqslant\frac{\bar{\delta}\left[1+\left(2\bar{v}+\bar{\delta}\right)\bar{\rho}\right]}{\underline{v}\left[2-T\left(2\bar{v}+\bar{\delta}\right)^2\right]}\cdot\left(1+\sqrt{1+\underline{v}T\left[2-T\left(2\bar{v}+\bar{\delta}\right)^2\right]}\right) \tag{21}$$

图 4 给出了相对位置误差与刷新率和相对速度的关系, 其中升空平台间距最大 500 m, 测速误差

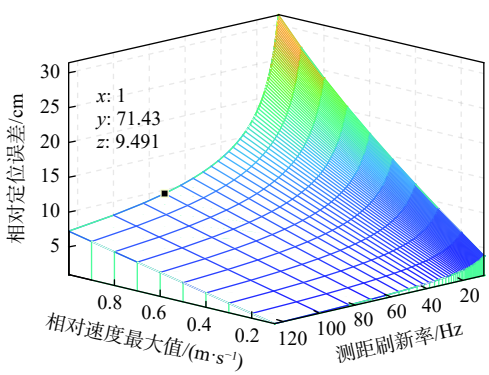


图 4　位置估计误差与测距刷新率、相对速度的关系

Fig. 4　Relationship of position estimation error with ranging refresh rate and relative velocity

$\bar{\delta}=0.1$ m/s[27], 测距刷新率F取 8～125 Hz, 相对速度取 0.1～1 m/s。如图 4 所示, 当两机每秒相对移动 1 m 时, 若要将测距刷新率对定位误差的影响控制在厘米量级, F>70 Hz, 且刷新率越高, 造成的相对定位误差越小。因此, 本文在后续测距和定位仿真实验中均设 UWB 传感器测距刷新率为最大值。

2.3　悬停稳定度

由于受到空中气流、机载飞控精度等诸多因素影响, 升空平台自身的不稳定性使得升空平台相对距离存在时变性, 导致传感器测量出现误差。由于实际升空状态下的平台悬停稳定度影响难以量化分析, 本文通过实测方式对相对测量性能进行评估。以四旋翼无人机为例, 本文测试采用视觉比对方法, 利用影像采集单元对升空悬停状态下的无人机按预设频率逐帧抓取影像, 并实时传输至计算机进行像素点对比, 通过比较两帧之间的无人机偏移量得到稳定度数据 (具体测试方法见文献 [28])。同时, 本文利用全站仪对 UWB 模块相对距离测量值进行比对, 提供具有毫米级精度的相对距离验核数据, 但由于激光验核方式在距离增大时难以对准, 本次水平方向测量基线为 20 m, 测试场景如图 5 所示。



图 5　悬停稳定度测试场景

Fig. 5　Test scenario of hovering stability

本次测试先后在微风 (风力 2～3 级) 和无风 (风力 0～1 级) 情况下进行, 共采集 30 组测试数据。测试中的机载 UWB 模块、固定 UWB 模块和移动影像采集模块位于同一高度面, 测量天线采用了 2.1 节微带天线, 测量数据通过 USB 接口输出至计算机进行存储和统计分析, 测试结果如图 6 所示。图中黑色曲线及标注表示通过图像比对得到的无人机悬停稳定度值, 蓝色曲线及标注则为 UWB 距离测量误差值。

从测试结果可以看出悬停稳定度与相对距离测量误差存在相关性: 对于第 1～15 组测试, 由于风力等环境影响较大, 无人机悬停误差最大值为 18.5 cm, UWB 模块测距偏差呈现大幅波动, 统计可得前 15 组测试中的相对测距偏差均值约为 6.78 cm; 对于第 16～30 组, 无人机悬停状态稳定, 此时测距偏差均值降低至−3.89 cm, 与静态 UWB 测量性能相当。

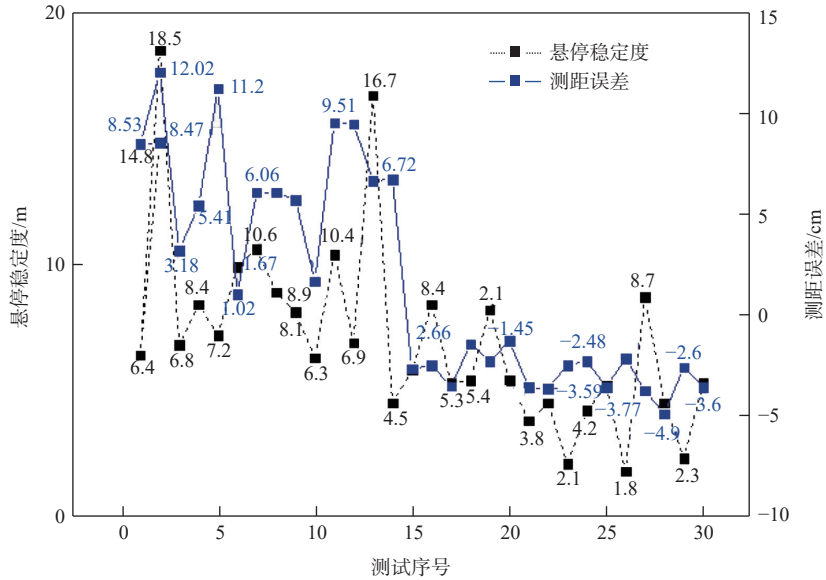


图 6 悬停稳定度对 UWB 测距误差的影响

Fig. 6 Effect of hovering stability on UWB ranging bias

综上所述，对于处于微动态下的无人机等升空平台，平台间相对距离测量值往往包含复杂的测量噪声，定位误差与天线相位中心偏移、测距刷新率、相对速度及悬停稳定性有关，且无法先验确定。因此，下一节将基于实测数据研究测量噪声在线分析和估计方法，改善相对测距和定位性能。

3 测量噪声 Allan 方差分析与建模

3.1 测量噪声分析

本节主要基于 UWB 模块实测数据，对其测量噪声进行分析和建模，提出基于 Allan 方差的噪声在线分析和参数估计算法，从而在后续滤波算法中实时计算并递归更新噪声参数，实现分布式升空平台的测距和定位精度的提升。Allan 方差将噪声分析并划分为量化噪声 (Q)、白噪声 (WN)、闪烁噪声 (FN)、随机游走噪声 (RW) 和高斯马尔可夫过程 (GM) 等。进行 Allan 方差计算前，先按时间簇将采样时刻k之前的所有数据分组，每组含$\bar{m}$个采样数据，共分成$\lfloor k/\bar{m}\rfloor$组，并计算每组数据的平均值$\bar{\mu}$，标准 Allan 方差计算式为[16]

$$\sigma_{\mu}^{2}(k)=\frac{1}{2(\lfloor k/\bar{m}\rfloor-1)}\sum_{i=1}^{\lfloor k/\bar{m}\rfloor-1}(\bar{\mu}_{i+1}-\bar{\mu}_{i})^{2} \tag{22}$$

为提高计算效率和稳定性，本文采用交叠式 Allan 方差。首先将采样间隔为τ_0的序列分为一组，设$\bar{n}$为该组数据起始编号，并计算每组数据的平均值$\bar{\mu}_n(\bar{m}\tau_0)\triangleq\frac{1}{\bar{m}}\sum_{k=\bar{n}}^{\bar{n}+\bar{m}+1}\mu(k)$，则交叠式 Allan 方差的计算式及其迭代如下：

$$\sigma_{\mu}^{2}(k)=\frac{1}{2(k-2\bar{m}+1)}\sum_{n=1}^{k-2\bar{m}+1}(\bar{\mu}_{\bar{m}+\bar{n}}-\bar{\mu}_{\bar{n}})^{2}=\left(1-\frac{1}{k-2\bar{m}+1}\right)\sigma_{\mu}^{2}(k-1)+\frac{1}{2(k-2\bar{m}+1)}\cdot(\bar{\mu}_{k-\bar{m}+\bar{n}}-\bar{\mu}_{k-2\bar{m}+1}) \tag{23}$$

图 7 给出了 2.3 节中测量数据的 Allan 方差分析结果，其中 UWB 传感器搭载在图 5 中升空悬停的旋翼无人机上，UWB 模块测距刷新率为 125 Hz，单次采集时长 1 000 s，其他参数设置同 2.3 节。

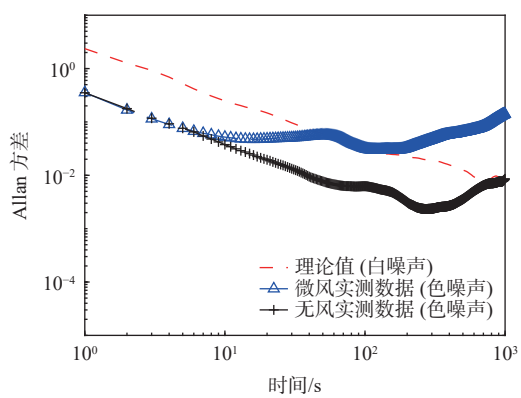


图 7 UWB 实测数据 Allan 方差分析 (与白噪声理论值对比)

Fig. 7 Allan plots of analysis UWB measurements (compared with theoretical values under white noise)

由实测数据的 Allan 方差分析可知，悬停状态下的 UWB 数据处于微动态，测量数据中包含有色噪声，如图 7 中带有色噪声标识的 Allan 方差曲线所示，而白噪声标识的 Allan 方差曲线为根据 UWB 理论模型得到的仿真数据进行 Allan 方差分析后得到。对比图 7 与 Allan 方差样本模型[16]可知，实际测量中的距离测量结果包含具有时间相关性的闪烁噪声，尤其在微风、稳定性较差情况下测量值中

还包含明显的随机游走噪声, 而理想情况下仅考虑了白噪声。因此, 在实际测量环境中, 应将式 (3) 中的噪声项修正为与时间有关的噪声项$\zeta(k)$, 需对其进行分析建模以降低色噪声影响。

3.2　UWB 测量噪声建模

根据文献 [29-30] 可知, 通过将单位白噪声通过预先设计的成形滤波器, 可以产生具有特定功率谱密度(power spectral density, PSD)的随机误差。反之, 当已知随机误差的功率谱密度, 可以得到成形滤波器的不同形式及其随机误差的微分方程描述。

1) 闪烁噪声。闪烁噪声源自易受随机闪烁影响的电子线路和其他通道的闪烁噪声, 由于其低频特性, 表现为数据中的零偏波动。闪烁噪声的功率谱密度和 Allan 方差分别为

$$S_B(\tilde{f})=\frac{B^2}{2\pi\tilde{f}} \tag{24}$$

$$\sigma_B^2=\frac{2B^2\ln 2}{\pi} \tag{25}$$

式中: B为零偏不稳定性系数; $\tilde{f}$为频率。

2) 量化噪声。量化噪声是通过以数字形式对模拟信号进行编码而引入到模拟信号中的误差之一。该噪声是由被采样点的实际幅度与模数转换器的位分辨率之间的微小差异引起的[31]。对于无限长相关时间, 量化噪声 PSD 可以表示为

$$S_Q(\tilde{f})=\frac{4Q_z^2}{\bar{m}\tau_0}\sin^2(\pi\tilde{f}\bar{m}\tau_0)\approx(2\pi\tilde{f})^2Q_z^2\bar{m}\tau_0 \quad \tilde{f}<\frac{1}{2\bar{m}\tau_0} \tag{26}$$

式中: Q_z为量化噪声系数, 与器件的量化位数有关; τ_0为数模转换时的采样时间间隔。实际上, 在有限相关时间内, A/D 转换过程产生的量化噪声在任何采样瞬间的误差可以看作是不相关的, 所以量化噪声可近似为白噪声序列。根据噪声方差与其本征随机过程的功率谱密度对应关系可得量化噪声的 Allan 方差分别为

$$\sigma_Q^2=\frac{3Q_z^2}{(\bar{m}\tau_0)^2} \tag{27}$$

3) 随机游走。随机游走噪声源于相关时间比采样时间短得多的高频噪声项, 在近距离、静态测量中可忽略, 但随着相对距离增大这种噪声的影响无可避免。设W为随机游走系数, 随机游走噪声的功率谱、Allan 方差可分别表示为

$$S_W(\tilde{f})=\frac{W^2}{(2\pi\tilde{f})^2} \tag{28}$$

$$\sigma_W^2=\frac{W^2\bar{m}\tau_0}{3} \tag{29}$$

综上所述, 通过对各类色噪声的建模可将 UWB 测量方差表示为$\sigma_\mu^2=\sigma_N^2+\sigma_B^2+\sigma_W^2$, 其中$\sigma_N^2$包含了近似为白噪声序列的量化噪声$\sigma_Q^2$。下面结合测试数据和交叠式 Allan 方差计算方法, 给出实测 UWB 数据的噪声参数 (见表 1, 其中 LOS 为 2 个 UWB 模块无遮挡情况, NLOS 为存在遮挡情况), 以及估计的σ_μ^2与真实数据的比对结果, 如图 8 所示。

表 1　基于 UWB 实测数据的 Allan 方差随机模型参数

Table 1　Parameters of Allan variance random model based on measured UWB data

距离/cm	σ_N	B	W	σ_μ
2 500 (LOS)	0.351	0.721	0.543	2.63
3 000 (LOS)	0.812	0.564	0.471	4.10
3 500 (LOS)	0.637	0.356	0.738	2.85
4 000(LOS)	0.545	0.417	0.791	3.96
4 500 (LOS)	0.592	0.458	0.427	2.78
5 000 (LOS)	0.679	0.673	0.501	3.26
5 000(NLOS)	0.609	0.232	1.569	4.76

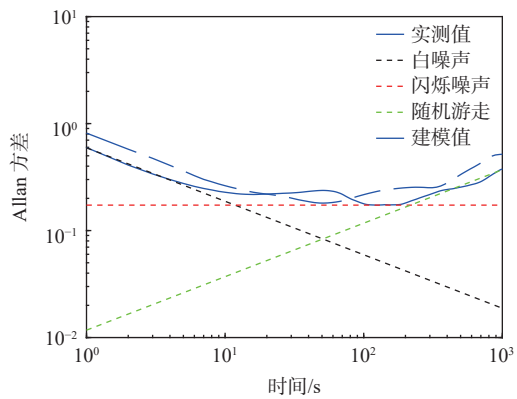


图 8　实测数据与模拟数据的 Allan 方差分析

Fig. 8　Allan variance analysis of data measurements and simulation

图 8 中蓝色实线表示实测相对距离的 Allan 方差, 黑色虚线代表斜率为−1/2 的白色噪声, 红色虚线代表斜率为 0 的闪烁噪声, 绿色虚线代表斜率为+1/2 的随机游走噪声, 蓝色点划线为模拟数据的 Allan 方差。比较实测数据和模拟数据可知, 采用 Allan 方差估计可以很好的逼近实际测量噪声方差, 且 Allan 方差与传统方差具有等价性, 利用递推 Allan 方差计算, 有利于在与内部传感器数据融合中迭代更新量测噪声协方差估计值, 从而自适应地调整噪声参数、提高定位精度。值得说明的是, 由于 Allan 方差估计精度与样本数据长度有关, 为保证升空平台在协同任务下的测量误差估计精度, 应连续地测量和存储 UWB 测量数据, 并尽量提高测距刷新率, 当测量数据采集时长较短或平台相对距离变化较剧烈时, Allan 方差估计精度将大幅下降。

4 色噪声下基于改进扩展卡尔曼滤波的相对定位算法

4.1 含测量色噪声的改进扩展卡尔曼滤波算法

扩展卡尔曼滤波 (extended Kalman filter, EKF) 是一种结合噪声测量来估计系统状态的有效方法，在无人机等非线性动力学系统中有着广泛应用。考虑到 UWB 相对测量包含的非零均值、非高斯时变噪声，需对传统 EKF 算法进行改进，提出含有色测量噪声下的分布式升空平台相对定位算法。

对于k时刻，设任一升空平台状态模型为$\boldsymbol{x}_i$，$\boldsymbol{u}_i$为平台控制矢量，高斯过程噪声$\boldsymbol{\varepsilon}^k \sim N\left(\boldsymbol{0},\boldsymbol{Q}^k\right)$下平台$i^{\text{th}}$的状态方程可写为

$$\boldsymbol{x}_i^k = f\left(\hat{\boldsymbol{x}}_i^{k-1},\boldsymbol{u}_i^k\right)+\boldsymbol{\varepsilon}^k \tag{30}$$

EKF 作为一种典型的非线性滤波器，可将多种传感器提供的信息进行有效融合，已广泛应用于组合导航系统中。一般地，可将平台i^{th}在k时刻的测量方程写为

$$\boldsymbol{y}_i^k = \boldsymbol{H}^k\boldsymbol{x}_i^k+\boldsymbol{R}_i^k \tag{31}$$

式中：$\boldsymbol{H}^k$为 Jacobian 矩阵，其具体表达取决于 IMU 传感器测量方式；$\boldsymbol{y}_i^k$为传感器对运动平台的速度、姿态等测量值；$\boldsymbol{R}_i^k$为测量噪声，对于本文所研究的实际传感器模块中存在的有色噪声，其测量误差是比较复杂的非平稳过程，可利用$\boldsymbol{R}_i^k$协方差矩阵和 Allan 方差的等价性进行误差的在线估计与迭代更新。基于上面的状态与测量方程，在$\left[\hat{\boldsymbol{x}}_i^{k-1}\right]^+$处一阶 Taylor 展开可得色噪声下的 EKF 算法步骤如下：

步骤 1 预测。

$$\left[\boldsymbol{x}_i^k\right]^- = f\left(\left[\hat{\boldsymbol{x}}_i^{k-1}\right]^+,\boldsymbol{u}_i^k\right) \tag{32}$$

$$\left[\boldsymbol{P}_i^k\right]^- = \boldsymbol{F}_i^k\left[\boldsymbol{P}_i^{k-1}\right]^+\left(\boldsymbol{F}_i^k\right)^{\mathrm{T}}+\boldsymbol{Q}^k \tag{33}$$

式中：$\boldsymbol{F}_i^k = \left.\dfrac{\partial f(\hat{\boldsymbol{x}}_i,\boldsymbol{u}_i)}{\partial \boldsymbol{x}_i}\right|_{\left[\hat{\boldsymbol{x}}_i^{k-1}\right]^+}$；$\left(\hat{\boldsymbol{x}}_i^{k-1}\right)^+$表示$\hat{\boldsymbol{x}}_i^{k-1}$的后验估计；$\left[\boldsymbol{P}_i^k\right]^-$和$\left[\boldsymbol{P}_i^k\right]^+$分别表示误差协方差矩阵$\boldsymbol{P}_i^k$的先验和后验估计。

步骤 2 更新。

$$\tilde{\boldsymbol{z}}_i^k = \boldsymbol{y}_i^k-\boldsymbol{H}_i^k\left[\hat{\boldsymbol{x}}_i^k\right]^- \tag{34}$$

$$\boldsymbol{S}_i^k = \boldsymbol{\Omega}_i^k\left[\boldsymbol{P}_i^{k-1}\right]^+\left(\boldsymbol{\Omega}_i^k\right)^{\mathrm{T}}+\boldsymbol{\delta}_i^k \tag{35}$$

$$\boldsymbol{K}_i^k = \left[\boldsymbol{P}_i^k\right]^-\left(\boldsymbol{\Omega}_i^k\right)^{\mathrm{T}}\left(\boldsymbol{S}_i^k\right)^{-1} \tag{36}$$

$$\left[\hat{\boldsymbol{x}}_i^k\right]^+ = \left[\hat{\boldsymbol{x}}_i^k\right]^- + \boldsymbol{K}_i^k\tilde{\boldsymbol{z}}_i^k \tag{37}$$

$$\left[\boldsymbol{P}_i^k\right]^+ = \left(\boldsymbol{I}-\boldsymbol{K}_i^k\right)\left[\boldsymbol{P}_i^{k-1}\right]^+ \tag{38}$$

其中初始时刻$\left[\boldsymbol{P}_i^0\right]^+ = E\left[\left(\left[\hat{\boldsymbol{x}}_i^0\right]^+-\hat{\boldsymbol{x}}_i^0\right)\left(\left[\hat{\boldsymbol{x}}_i^0\right]^+-\hat{\boldsymbol{x}}_i^0\right)^{\mathrm{T}}\right]$；$\boldsymbol{\Omega}_i^k = \boldsymbol{H}_i^k+\left.\dfrac{\partial[\delta\boldsymbol{X}^i\ \boldsymbol{0}]^{\mathrm{T}}}{\partial\boldsymbol{x}}\right|_{\left[\hat{\boldsymbol{x}}_i^k\right]^-}$，$\boldsymbol{0}$表示全零矢量，其长度取决于当前进行滤波的测量值个数。

4.2 算法复杂度分析

下面对含测量色噪声的改进 EKF 算法复杂度进行分析，运算复杂度一方面在一定程度上代表着计算所需资源的大小，另一方面也是衡量定位算法可行性的重要参考指标。每秒所执行的等效浮点运算次数(floating-point operations, FLOPs) 是分析运算时间复杂度的有效方法之一[32]，该方法通过对算法中矩阵加减、乘法、求逆等浮点操作的数量进行定量表征。

设升空平台导航系统状态向量$\boldsymbol{x}_i$为n维，量测向量$\boldsymbol{y}_i$为l维，根据 4.1 节所给算法步骤可知，式 (32) 状态方程先验估计$\left[\boldsymbol{x}_i^k\right]^-$需要$2n^2-n$次 FLOPs，式 (33) 协方差先验估计$\left[\boldsymbol{P}_i^k\right]^-$需$4n^3-n^2$次 FLOPs；式 (34) 测量方程预测需$2nl$次 FLOPs，式 (35) 协方差预测需$3n^2l+3n^2+2l$次 FLOPs；后验估计中式 (36) 的卡尔曼滤波增益估计需$l^3+3n^2+3n^2l-2nl$次 FLOPs，式 (37) 状态估计更新需$2nl+l$次 FLOPs，式 (38) 协方差后验估计更新需要$2nl^2-nl+2nl^2$次 FLOPs。由此可得本文所提算法的总体时间复杂度为$4n^3+(6l+1)n^2+\left(10l^2+l-1\right)n+l^3+l$，而对于典型 EKF 算法，总体复杂度可表示为$4n^3+6n^2l+6nl^2+n^2-n+l^3$[33]，复杂度的增加主要体现在矩阵求导运算和方差计算上，由于实际的 Jacobian 矩阵运算复杂度取决于非线性函数，与系统状态和测量维数有关，图 9 分析了不同测量维度下 EKF 算法和改进算法复杂度的变化曲线。

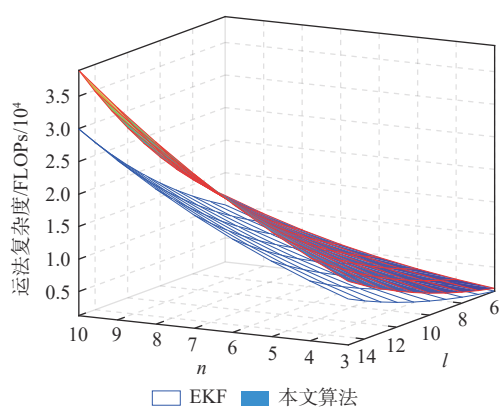


图 9 算法复杂度分析

Fig. 9 Complexity analysis of proposed algorithm

由仿真结果可知，当状态估计维数$n=3\sim10$，$l=6\sim15$时，本文算法的复杂度较标准 EKF 算法有所增加，当$n=10$，$l=15$时，运算复杂度增加量最大

为 17.9%, 对定位实时性的影响不大。

4.3　仿真算例

本节利用前述定位模型、传感器测量模型及噪声参数, 在分布式升空平台的动态相对定位仿真中构造贴近于实际环境的相对测量噪声, 并在滤波过程补偿噪声引起的定位精度损失。假设升空平台在停留仍存在约为 0.1～1 m/s 的运动, 并设该运动轨迹呈现图 10 所示的随机、缓慢位移。升空平台由 IMU 器件测量自运动, 平台间通过收发 UWB 信号进行动态相对测距和定位, 这里利用了第 3 节中对 UWB 测量噪声的分析和建模结果, 其中噪声参数设置为 $B=0.07$, $W=0.05$ 和 $\sigma_N^2=0.06$, 本次仿真时长 1 000 s, 刷新率 125 Hz, 平台数量 $N=10$, 其他仿真参数如表 2 所示。

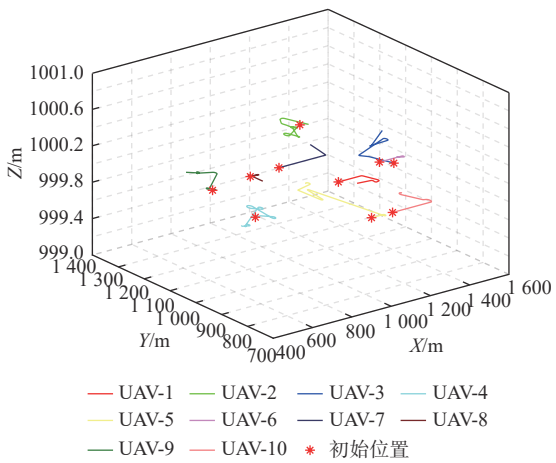


图 10　升空平台运动特征仿真

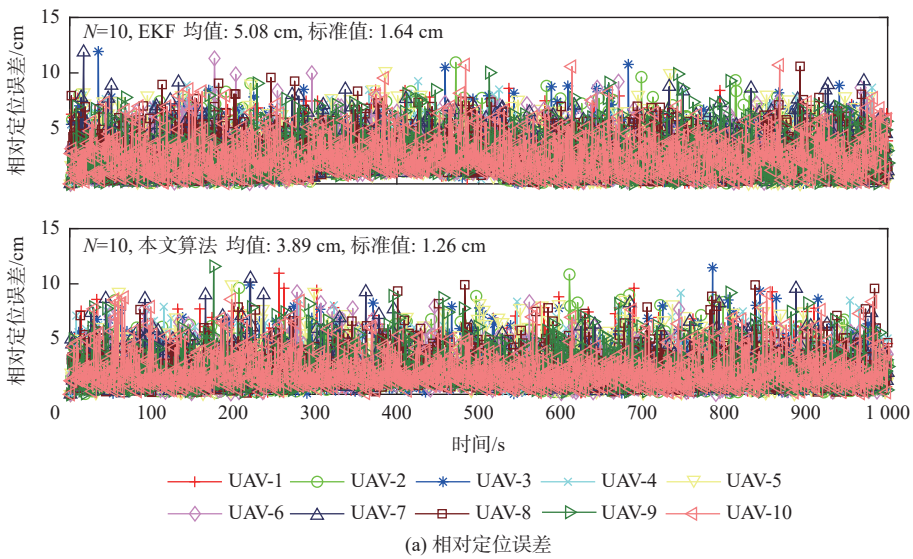
Fig. 10　Simulation of motion characteristics of blast-off platform

表 2　仿真参数设置

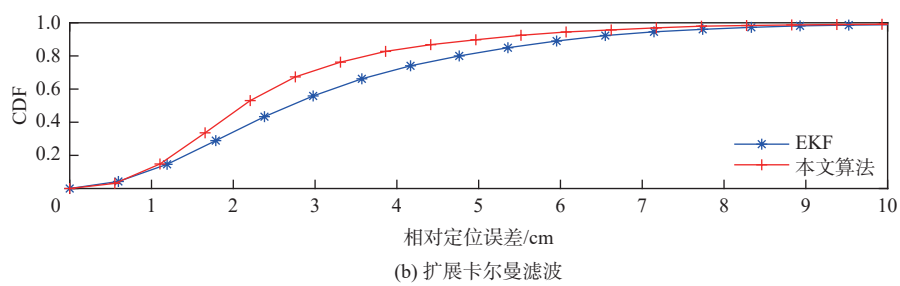
Table 2　Simulation parameters

参 数	数 值
陀螺仪测量误差/ (°)	$N(0.001,0.001)$
加速度计测量误差/(m·s^{-2})	$N(0.001,0.003)$
磁力计测量误差/(°)	$N(0.02,0.05)$
四元数运动模型噪声协方差	0.01
位置过程噪声协方差/m	1
角速率过程噪声协方差/(rad·s^{-1})	10^{-6}
线加速度过程噪声协方差/(rad·s^{-2})	10^{-6}

图 11 给出了采用本文算法 滤除升空平台 UWB 模块色噪声时的定位误差, 并与传统 EKF 算法定位性能进行比较。仿真结果表明, 在有色噪声干扰的情况下, EKF 位置估计位置在 200～600 s 之间波动剧烈, 从而产生了较大的误差, 而本文算法充分考虑并滤除了测量色噪声, 算法性能在升空平台存在动态情况下更加稳定, 从而产生更高精度的位置估计结果。从误差的累积分布函数 (cumulative distri-



(a) 相对定位误差



(b) 扩展卡尔曼滤波

图 11　分布式升空平台定位误差分析

Fig. 11　Position error analysis of distributed blast-off platforms

bution function, CDF) 结果可以看出, 在置信度为 95% 的情况下, 该算法的估计误差在 6 cm 以内。通过计算平均偏差统计, 本文算法的标准偏差约为 1.26 cm, 而 EKF 算法的标准偏差约为 1.64 cm, 2 种不同滤波算法的平均位置误差分别为 3.89 cm 和 5.08 cm, 位置估计误差降低约 23.43%。

5 结 论

1) 当升空平台 UWB 传感器测量数据包含有色噪声, 本文算法能够在滤波过程对噪声参数进行在线估计和调整, 同时补偿传感器噪声引起的定位精度损失。

2) 本文算法相较于传统算法的位置估计精度提高了约 23%, 可实现非理想情况下分布式升空平台的相对定位, 提升平台协同能力和任务完成能力。

3) 所提相对定位算法可在地面锚节点或其他方式的辅助下, 进一步通过距离和角度约束消除分布式升空平台整体旋转不确定性和平移不确定性, 获得绝对位置校正。

本文算法有望弥补传统算法仅考虑理想噪声测量环境, 解决真实环境中多无人机近距离、高精度协同定位的需求。在后续研究工作中, 将进一步研究论证 GNSS 拒止条件下的机载小型定位模块和集群飞行实验, 并通过升空测试在无人机平台上验证定位算法。

参考文献 (References)

[1] FOX N. Precision electronic warfare (PREW): BAA 09-65 [R]: Arlington Strategic Technology Office of the US Dedense Advanced Research Projects Agency, 2009: 5-32.

[2] 李磊, 王彤, 蒋琪. 美国 CODE 项目推进分布式协同作战发展[J]. 无人系统技术, 2018, 1(3): 59-66.
LI L, WANG T, JIANG Q. DARPA CODE program promotes the development of distributed collaborative operations[J]. Unmanned Systems Technology, 2018, 1(3): 59-66 (in Chinese).

[3] 臧会凯, 雷欢, 但晓东, 等. 分布式雷达相参发射原理与性能分析[J]. 电子与信息学报, 2015, 37(8): 1801-1807.
ZANG H K, LEI H, DAN X D, et al. Theory and performance analysis of coherent transmission for distributed radars[J]. Journal of Electronics & Information Technology, 2015, 37(8): 1801-1807(in Chinese).

[4] GOEL S, KEALY A, GIKAS V, et al. Cooperative localization of unmanned aerial vehicles using GNSS, MEMS inertial, and UWB sensors[J]. Journal of Surveying Engineering, 2017, 143(4): 04017007.

[5] 王均晖, 孙蕊, 程琦, 等. 无人机组合导航直接法与间接法滤波方式比较[J]. 北京航空航天大学学报, 2020, 46(11): 2156-2167.
WANG J H, SUN R, CHENG Q, et al. Comparison of direct and indirect filtering modes for UAV integrated navigation[J]. Journal of Beijing University of Aeronautics and Astronautics, 2020, 46(11): 2156-2167(in Chinese).

[6] LIU C, PRIOR S D , TEACY W L, et al. Computationally efficient visual-inertial sensor fusion for global positioning system-denied navigation on a small quadrotor[J]. Advances in Mechanical Engineering, 2016, 8(3): 1687-1698.

[7] KOPPANYI Z, TOTH C K. INDOOR Ultra-wide band nerwork adjustment using maximum likelihood estimation[J]. ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences, 2014, 2(1): 31-35.

[8] JOHNSON J, DEWBERRY B. Ultra-wideband aiding of GPS for quick deployment of anchors in a GPS-denied ad-hoc sensor tracking and communication system[C]//Proceedings of the 24th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS 2011). Washington: The Institute of Navigation, 2011: 3959-3966.

[9] DIERENBACH K, OSTROWSKI S, JOZKOW G, et al. UWB for navigation in GNSS compromised environments[C]//Proceedings of the 28th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS+ 2015). Washington: The Institute of Navigation, 2015: 2380-2389.

[10] 熊骏, 熊智, 于永军, 等. UWB 辅助的无人机近距离编队相对导航方法[C]// 中国惯性技术学会高端前沿专题学术会议论文集. 2017: 254-263.
XIONG J, XIONG Z, YU Y J, et al. Relative navigation algorithm for UAV close-formation flight aided by UWB[C]// Proceedings of the High-end Frontier Academic Conference of China Inertial Technology Society. 2017: 254-263 (in Chinese).

[11] QUERALTA J P, AlMANSA C M, SCHIANO F, et al. Uwb-based system for UAV localization in gnss-denied environments: Characterization and dataset[C]//2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Piscataway: IEEE Press, 2020: 4521-4528.

[12] SHULE W, ALMANSA C M, QUERALTA J P, et al. Uwb-based localization for multi-uav systems and collaborative heterogeneous multi-robot systems[J]. Procedia Computer Science, 2020, 175: 357-364.

[13] GROBWINDHAGER B, BOANO C A, RATH M, et al. Enabling runtime adaptation of physical layer settings for dependable UWB communications[C]//2018 IEEE 19th International Symposium on" A World of Wireless, Mobile and Multimedia Networks". Piscataway, IEEE Press, 2018: 1-11.

[14] DING W D, WANG J L, RIZOS C, et al. Improving adaptive Kalman estimation in GPS/INS integration[J]. Journal of Navigation, 2007, 60(3): 517-529.

[15] LIU R, LIU F, LIU C, et al. Modified Sage-Husa adaptive Kalman filter-based SINS/DVL integrated navigation system for AUV[J]. Journal of Sensors, 2021, 2021: 9992041.

[16] BOARD I S. IEEE standard definitions of physical quantities for fundamental frequency and time metrology-random instabilities: 1139-2008[S]. Piscataway: IEEE Press, 2008: 5-36.

[17] VIG J R, FERRE-PIKAL E S, CAMPARO J C, et al. IEEE standard definitions of physical quantities for fundamental frequency and time metrology-random instabilities[S]. Piscataway: IEEE, 1999: 1139-1999.

[18] HUANG Y, HUA L, WANG Q, et al. Visible light positioning and navigation using noise measurement and mitigation[J]. IEEE Transactions on Vehicular Technology, 2019, 68(11): 11094-11106.

[19] CHANG G. On Kalman filter for linear system with colored meas-

urement noise[J]. Journal of Geodesy, 2014, 88(12): 1163-1170.

[20] DEMARCO K , SQUIRES E , DAY M , et al. Simulating collaborative robots in a massive multi-agent game environment (SCRIMMAGE)[C]//Springer Proceedings in Advanced Robotics. Switzerland, AJ: Springer Nature, 2019: 283-291.

[21] PALA S , JAYAN S , KURUP D G. An accurate UWB based localization system using modified leading edge detection algorithm[J]. Ad Hoc Networks, 2020, 97: 102017

[22] 李荣冰, 王念曾, 刘建业, 等. 面向相对导航的 UWB 测距误差估计与补偿方法[J]. 仪器仪表学报, 2019, 40(5): 28-35.
LI R B, WANG N C, LIU J Y, et al. UWB ranging error estimation and compensation method for relative navigation[J]. Chinese Journal of Scientific Instrument, 2019, 40(5): 28-35 (in Chinese).

[23] MIHAI I V, TAMAS R, SHARAIHA A. An UWB physical optics approach for Fresnel-zone RCS measurements on a complex target at non-normal incidence[J]. Sensors, 2019, 19(24): 5454-5470.

[24] GUO K, QIU Z, MIAO C, et al. Ultra-wideband-based localization for quadcopter navigation[J]. Unmanned Systems, 2016, 4(1): 23-34.

[25] ZHOU J M. A design of micro-strip antenna in the RFID system and UWB system[J]. Recent Advances in Electrical & Electronic Engineering, 2016, 9(2): 119-123.

[26] 崔正阳, 王勇. 无人机固定时间路径跟踪容错制导控制[J]. 北京航空航天大学学报, 2021, 47(8): 1619-1627
CUI Z Y, WANG Y. Fault-tolerant fixed-time path following guidance control of UAV[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(8): 1619-1627(in Chinese).

[27] 匡银虎, 张虹波. 多旋翼无人飞行器悬停姿态精确控制仿真研究[J]. 计算机仿真, 2018, 35(3): 34-37.
KUANG Y H, ZHANG H B. Simulation study on accurate control of hovering attitude of multi-rotor unmanned aerial vehicle[J]. Compute Simulation, 2018, 35(3): 34-37 (in Chinese).

[28] 张奇. 一种飞行器悬停功能稳定性测试方法和系统: CN106742051A[P]. 2017-05-31.
ZHANG Q. Aircraft hovering function stability testing method and system: CN106742051A[P]. 2017-05-31(in Chinese).

[29] MCGUIRE K, DE CROON G, DE WAGTER C, et al. Efficient optical flow and stereo vision for velocity estimation and obstacle avoidance on an autonomous pocket drone[J]. IEEE Robotics and Automation Letters, 2017, 2(2): 1070-1076.

[30] 耿梦梦, 任元, 樊亚洪, 等. MSCSG 随机误差测试及误差源分析[J]. 北京航空航天大学学报, 2021, 47(8): 1697-1704.
GENG M M, REN Y, FAN Y H, et al. Random error test and error source analysis of MSCSG[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(8): 1697-1704(in Chinese).

[31] BROWN R G, HWANG P Y C. Introduction to random signals and applied Kalman filtering with Matlab exercises and solutions[M]. Hoboken: John Wiley & Sons, Inc., 1996: 87-125.

[32] HUNGER R. Floating point operations in matrix-vector calculus[M]. Munich: Munich University of Technology, Inst. for Circuit Theory and Signal Processing, 2005.

[33] 张召友. 非线性 Bayesian 滤波及其在 SINS/GPS 紧耦合导航中的应用研究[D]. 哈尔滨: 哈尔滨工程大学, 2014.
ZHANG S Y. Research on nonlinear Bayesian filtering and its application for SINS/GPS tightly coupled navigation[D]. Harbin :Harbin Engineering University, 2014(in Chinese).

Positioning accuracy and localization algorithm with relative measurement errors in blast-off platforms

OUYANG Xiaofeng[1], ZENG Fangling[1,*], LYU Daqian[1], DONG Tianbao[1], HAN Yijing[2]

(1. College of Electronic Engineering, National University of Defense Technology, Hefei 230037, China;
2. The State Key Laboratory of Astronautic Dynamics, Xi'an 710043, China)

Abstract: In response to the high-precision relative positioning requirements for precise collaborative operations involving distributed blast-off platforms, this research proposes a relative ranging and positioning model based on ultra wide band (UWB) sensors. It analyzes the error factors and noise components of UWB modules under non-ideal conditions with antenna phase center errors, ranging refresh rates, and hover stability. To improve the overall performance of ranging and positioning, the noise component in UWB measurements is identified and analyzed by Allan variance, and the colored noise model along with estimated parameters are integrated into the measurement update of the positioning algorithm. The relative positioning of distributed blast-off platforms is then obtained through improving the expanded Kalman filter. Simulation experiments show that when the UWB measurement noise follows a non-Gaussian distribution, the position estimation accuracy of the proposed algorithm is 23.43% higher than that of the traditional EKF algorithm. Compared to traditional algorithms, the proposed algorithm reduces the influence of measurement color noise, thereby increasing relative positioning accuracy in satellite navigation-denied environments.

Keywords: blast-off platform; relative positioning; error analysis; ultra wide band sensor; Allan variance

Received: 2022-04-12; **Accepted**: 2022-05-27; **Published Online**: 2022-06-16 11:38

URL: link.cnki.net/urlid/11.2625.V.20220615.1935.001

* **Corresponding author**. E-mail: zella@ustc.edu

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0259

纤维排布对CFRP感应加热固化成型过程影响表征

付天宇[1,*]，杨宁[2]，古云飞[3]，许家忠[4]，蒋悦[3]，李斌[5]

(1. 常州信息职业技术学院 智能装备学院，常州 213164；2. 山西大同大学 计算机与网络工程学院，大同 037009；
3. 哈尔滨理工大学 机械动力工程学院，哈尔滨 150080；4. 哈尔滨理工大学 自动化学院，哈尔滨 150080；
5. 重庆市特种设备检测研究院，重庆 401120)

摘 要：在采用电磁感应加热方式实现碳纤维复合材料（CFRP）加热固化成型的过程中，如何准确表征材料各组成部分的属性参数状态对研究材料加热过程中温度、固化度及应力场分布至关重要。根据纤维实际铺层方式对材料各组成部分进行独立表征，以及碳纤维壳体的等效电-磁-热特性建立了CFRP感应加热有限元细观分析模型，将纤维织构和树脂基体单独建立，实现CFRP在感应加热过程中其内部纤维织构中物理场的变化规律以及对材料整体升温变化规律进行系统分析。通过该模型计算了材料温度、固化度及应力场分布的变化规律，揭示了碳纤维排布对材料感应加热效果的影响，并根据感应加热实验验证了该模型可以准确表征各物理场的分布状态。为研究CFRP在感应加热中各物理场变化规律提供了有效的计算模型和分析方法。

关 键 词：碳纤维；感应加热；温度场；固化度；应力场
中图分类号：V258+.3; TB332
文献标志码：A **文章编号**：1001-5965(2024)01-0198-10

随着碳纤维复合材料(carbon fiber reinforced polymer, CFRP)在人们日常生活和工业生产中的应用越加广泛[1-2]，对材料的固化成型方式、能源使用率及污染物的排放等提出了新的要求[3-4]，其中感应加热技术具有加热效率高、能耗低、设备占地小及非接触等众多优势[5]，将感应加热技术与CFRP的加热固化成型相结合，可以大幅提高材料的升温速率、减少环境污染、降低生产成本、减少环境中杂质的渗入等[6-7]。而CFRP在固化过程中其温度场、固化速率及应力场变化规律均受到纤维排布方式和体积分数的直接影响，因此当采用感应加热方式实现CFRP的固化成型过程中，如何表征纤维排布方式及体积分数，是探究材料属性状态变化的关键因素[8-9]。

CFRP固化成型与电磁加热技术的联合应用，受到了行业内学者们的广泛关注，并对CFRP在感应加热过程中的升温机理及温度场状态进行了较为深入的研究[10-12]。Bruce等[13]将可以预测不同CFRP结构的升温模型输入到有限元热传导分析中，从而实现为层压板平面提供了预测材料内部瞬态热分布有限元模型及分析方法。Mitschang等[14]对碳纤维层合板感应加热过程能够形成闭合回路的材料中横截面积、电阻率及材料长度等参数对材料内部发热量的影响进行研究。Banerjee等[15]设计了多谐振Halbach阵列的线圈绕组排布方式，使线圈中所产生的交变磁场均集中在线圈绕组上表面且均匀分布，增加CFRP壳体温度场的均匀性，提高能源使用效率。Fredrik等[16]将单向层表示为均匀和各向异性的域，电导率和热导率用张量表示，用于计算感应加热过程中基于单向层的交叉层

收稿日期：2022-04-19；**录用日期**：2022-06-20；**网络出版时间**：2022-07-25 12:57
网络出版地址：link.cnki.net/urlid/11.2625.V.20220722.1531.002
基金项目：国家自然科学基金 (52303031)；常州市应用基础研究计划 (CJ20220035)；碳纤维压力容器成型技术研究实验室 (KYPT202203Z)
* **通信作者**. E-mail：827263230@qq.com

*引用格式：付天宇，杨宁，古云飞，等. 纤维排布对CFRP感应加热固化成型过程影响表征 [J]. 北京航空航天大学学报，2024，50（1）：198-207.
FU T Y，YANG N，GU Y F，et al. Characterization of influence of fiber arrangement on CFRP induction heating curing process [J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50（1）：198-207 （in Chinese）.*

CFRP 板中的温度分布和加热功率分布。学者们主要通过层合模型研究 CFRP 物理场的变化规律[17-19], 未能直观表征纤维排布方式与材料温度场、固化度及应力场分布的内在关联。

因此, 为揭示纤维排布方式与加热过程中各参数分布状态的映射关系。本文采用了细观建模方法, 将 CFRP 中纤维织构及树脂基体进行单独表示, 并根据 CFRP 的等效电-磁-热特性形成感应加热有限元细观分析模型。计算了不同纤维排布方式下的 CFRP 温度场、固化度及应力场的分布规律, 揭示内部碳纤维对各物理场状态所产生的作用和影响。通过 CFRP 感应加热实验, 采集了加热过程中材料温度场、升温速率及应力变化曲线, 与模拟计算结果对比分析可以证明本文所建立的有限元模型可以实现不同纤维排布方式对加热过程中各参数影响的准确表征。

1　有限元模型建立

1.1　几何模型建立

由于 CFRP 在感应加热过程中需要纤维束存在有效闭合回路, 因此本文选取了平纹编织、单向均匀交织及单向非均匀交织 3 种纤维排布方式进行研究, 图 1 为 3 种纤维铺层方式的几何模型, 类型 1 为平纹编织结构, 类型 2 为单向均匀交织结构, 类型 3 为单向非均匀交织结构在 X 方向上铺层数量较多。

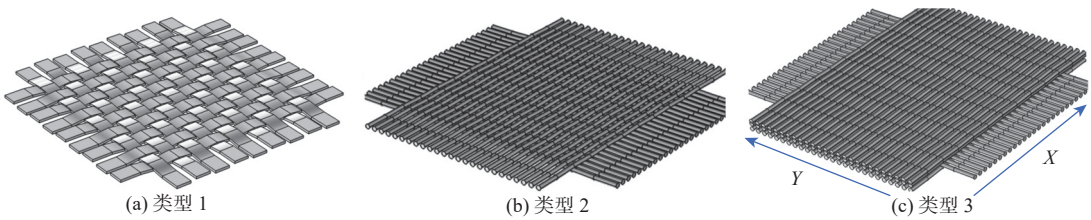


图 1　纤维铺层方式的几何模型

Fig. 1　Geometric model of fiber layup method

模型中网格划分质量直接影响了物理场计算的准确性, 而在 CFRP 感应加热过程中, 纤维束节点处物理场的变化, 对材料升温过程中的准确性影响最为明显。因此, 网格划分需要更加密集以反映出电势降的变化程度, 保证感应加热过程中物理场计算的准确性, 本文网格尺寸选择为 0.05 mm。图 2 为平纹编织结构的网格划分质量直方图, 可以看出网格质量主要集中在 0.5～0.8 之间。图 3 为 3 种纤维铺层方式及节点处的网格划分质量图, 可以看出各纤维织构的节点处网格质量均高于 0.6, 可以准确反映出物理场的变化情况。

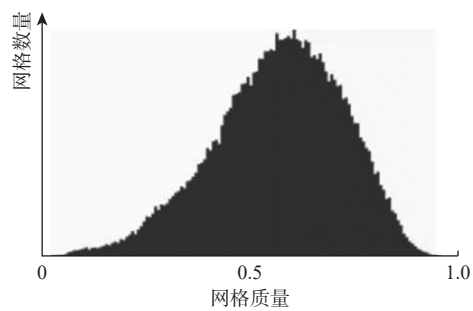


图 2　材料网格划分质量直方图

Fig. 2　Material meshing histogram

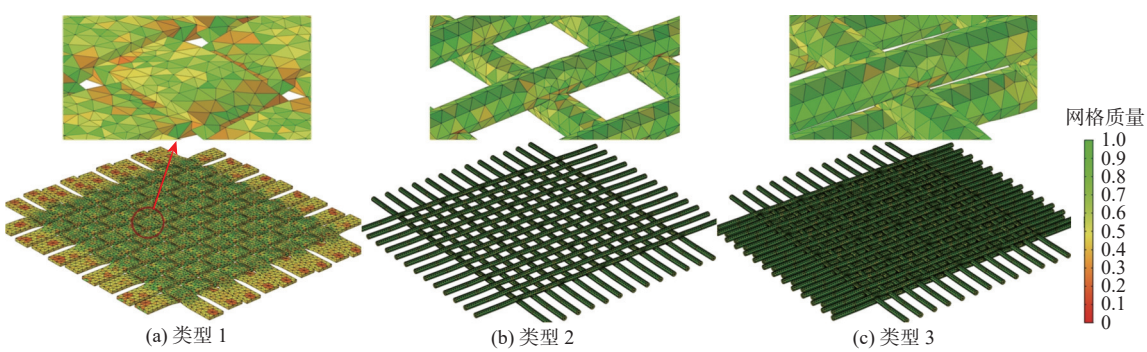


图 3　网格划分质量图

Fig. 3　Meshing quality map

1.2　物理参数设定

模型中各 CFRP 的尺寸均为 60 mm×60 mm×5 mm, 感应线圈尺寸为直径为 30 mm, 长度为 20 mm, 匝数为 50, 电流大小为 16 A, 频率大小为 13 kHz, 外界环境温度为 20 ℃, 线圈与碳纤维表面距离为 2.5 mm, 在研究过程中 CFRP 平板中的树脂基体

及碳纤维型号均相同，仅改变纤维的铺层方式。表1为几何模型中各材料赋值的参数，其参数的获取均由CFRP生产商家及COMSOL材料库提供。

表1 材料相关仿真参数

Table 1 Material related simulation parameter

材料	导热系数/(W·(m·K)$^{-1}$)	比热容/(J·(kg·K)$^{-1}$)	密度/(kg·m^{-3})	电导率/(S·m^{-1})	相对介电常数	相对渗透率	弹性模量/GPa	热膨胀系数/K^{-1}
空气				0	1	1		
碳纤维	50	850	1 500	6.4×10^{4}		1	160	-0.6×10^{-6}
树脂	0.2	1 100	1 200	1×10^{-2}	3.2	1	1	6×10^{-5}
线圈			8 960	6×10^{7}	1	1		

2 理论模型建立

2.1 纤维织构涡流及发热量

交变磁场在碳纤维织构中产生感应电场 $\boldsymbol{E}$，V/m，以频率的形式可以表达为

$$\boldsymbol{E}=-\mathrm{j}\omega\boldsymbol{A} \tag{1}$$

式中：$\mathrm{j}^2=-1$；$\omega=2\pi f$为角频率，rad/s，f为磁场频率，Hz；$\boldsymbol{A}$为磁矢量势，Wb/m。

纤维织构中产生的总电流总 $\boldsymbol{J}_{\mathrm{t}}$ 和由传导电流 $\boldsymbol{J}_{\mathrm{i}}=\sigma\boldsymbol{E}$和位移电流$\boldsymbol{J}_{\mathrm{d}}=\mathrm{j}\omega\boldsymbol{D}$组成：

$$\boldsymbol{J}_{\mathrm{t}}=\boldsymbol{J}_{\mathrm{i}}+\boldsymbol{J}_{\mathrm{d}}=\sigma\boldsymbol{E}+\mathrm{j}\omega\boldsymbol{D} \tag{2}$$

式中：σ为电导率，S/m。

纤维织构中的涡旋电流与其自身电阻所产生的焦耳热作为CFRP的加热热源，其中热源 $Q_{\mathrm{rh,i}}$ 单位为W/m^3 可以表示为

$$Q_{\mathrm{rh,i}}=\boldsymbol{J}_{\mathrm{t}}\boldsymbol{E} \tag{3}$$

2.2 材料热传导原理

当电源激活后，由纤维束中的感应电流所产生的热量转移到了周围纤维束、树脂及环境空气中，纤维部分的热传导可以通过傅里叶热传导定律来表示：

$$\rho C_p\frac{\partial T}{\partial t}+\nabla(-k\nabla T)=Q \tag{4}$$

式中：ρ为材料的密度，kg/m^3；C_p为材料的比热容，J/(kg·K)；k为材料的导热系数，W/(m·K)；Q为热源发出的热量，W/m^3；T为温度，K；t为时间，s；∇为哈密顿算子；∇T为温度梯度，K/m。

由于环氧树脂为相变材料，在固化过程中由液态转变为固态发生放热现象，在树脂中发生相变时的热传导公式为

$$\rho C_p\mu\cdot\nabla T+\nabla(-k\nabla T)=Q \tag{5}$$

式中：μ为树脂在加热时的流动速度，m/s。

2.3 固化及应力

双参数固化模型，由Sestak-Berggren最早提出，由于自催化反应诱导期的存在，反应速率需经过一定时间才能达到最大值，一般固化度为30%～40%时固化速率达到最大值，模型如下：

$$\frac{\mathrm{d}\alpha}{\mathrm{d}t}=k\alpha^m(1-\alpha)^n \tag{6}$$

式中：m和n为反应级数；α为固化度。

CFRP中的残余应力主要考虑由热膨胀和化学收缩引起的。

$$\{\sigma\}=[E_0](\{\varepsilon_{\mathrm{tot}}\}-\{\varepsilon_{\mathrm{tc}}\})+\{\sigma_0\} \tag{7}$$

$$\{\varepsilon_{\mathrm{tc}}\}=\{\varepsilon_{\mathrm{th}}\}+\{\varepsilon_{\mathrm{ch}}\}=\{\alpha_1\Delta T\}+\{\Delta\varepsilon_{\mathrm{r}}\Delta\alpha\} \tag{8}$$

式中：σ和σ_0分别为内应力和初始应力；E_0为弹性模量；$\varepsilon_{\mathrm{tot}}$、$\varepsilon_{\mathrm{tc}}$、$\varepsilon_{\mathrm{th}}$和$\varepsilon_{\mathrm{ch}}$分别为总应变、自由热化学应变、热膨胀应变和化学收缩应变，其中自由热化学应变是化学收缩应变和应变的叠加；$\Delta\varepsilon_{\mathrm{r}}$为化学收缩系数；$\Delta\alpha$为固化度变化量。

树脂的弹性模量 E_{m} 可以表示为

$$E_{\mathrm{m}}=\left(1-\frac{\alpha-\alpha_{\mathrm{gel}}}{1-\alpha_{\mathrm{gel}}}\right)E_{\mathrm{m}}^{0}+\frac{\alpha-\alpha_{\mathrm{gel}}}{1-\alpha_{\mathrm{gel}}}E_{\mathrm{m}}^{\infty} \tag{9}$$

式中：E_{m}^{0}和E_{m}^{∞}分别为固化树脂和完全固化树脂的弹性模量；α_{gel}为凝胶点树脂的固化度。

2.4 材料表面换热

在感应加热过程中，当物体的温度高于外界环境后，加热物体升温后表面与周围环境进行热交换，其中与空气之间的交换最多，因此需要考虑空气的对流热通量：

$$-k\nabla T=h(T_{\mathrm{amb}}-T) \tag{10}$$

式中：h为由边界类型和表面性质所确定的传热系数，W/(m^2·K)；T_{amb}为外部空气温度；T为被加热材料表面温度。

当温度逐渐升高时，需要考虑到材料所产生的热辐射，扩散表面向各个方向上均匀的释放辐射强度，材料表面辐射强度为

$$-k\nabla T=\varepsilon\gamma\left(T_{\mathrm{amb}}^4-T^4\right) \tag{11}$$

式中：ε为材料表面的辐射率；γ为玻尔兹曼常数。

图4为CFRP平板在感应加热过程中的相关原理，可以看出纤维织构中的发热量主要由于材

料内部的纤维束之间形成了有效的闭合回路, 从而产生的焦耳热使其整体升温。图中, R_f 为纤维电阻, ρ_f 为纤维密度, l_f 为纤维长度, A_f 为纤维截面积, ΔV_f 为纤维直接加热电压降, R_{jd} 为介电加热中的电阻部分, C_{jd} 为电容部分, ΔV_j 为介电加热电压降。

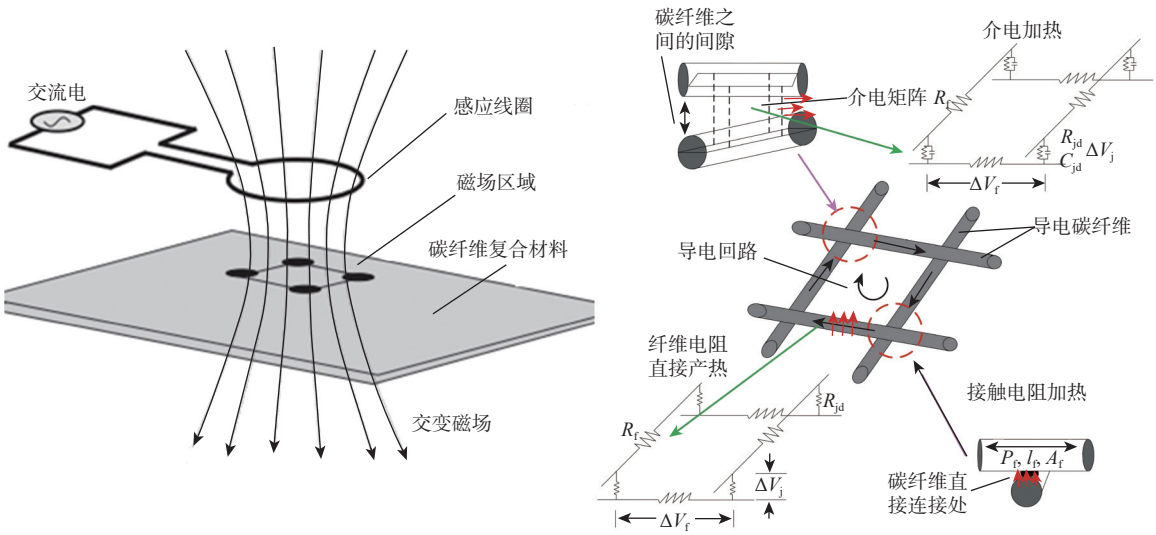


图 4　CFRP 感应加热原理

Fig. 4　Principle of CFRP induction heating

3　感应加热升温机理

图 5 为不同纤维编织结构的 CFRP 在感应加热过程中涡旋电流分布规律, 从图 5(b)和图 5(c)中可以看出平纹编织和单向均匀交织结构的涡旋电流主要集中在线圈外圆所对应的纤维织构中, 在 CFRP 感应加热的实际生产中只需要考虑纤维束轴向上所产生的涡旋电流即可, 图 5(d)为纤维束及节点处的电流走向, 可以看出涡流主要沿纤维束走向流动, 仅在纤维束拐角区域存在杂散的电流分布, 电流通过纤维束节点后沿交叠纤维束持续流动。而图 5(e)所示的类型 3 在 y 方向上涡旋电流分布更加密集, 这是由于 y 方向上的纤维含量较少, 涡旋电流形成闭合回路后在该方向上存在了电流集中的现象。

图 6 为 3 种类型 CFRP 纤维织构中的发热量分

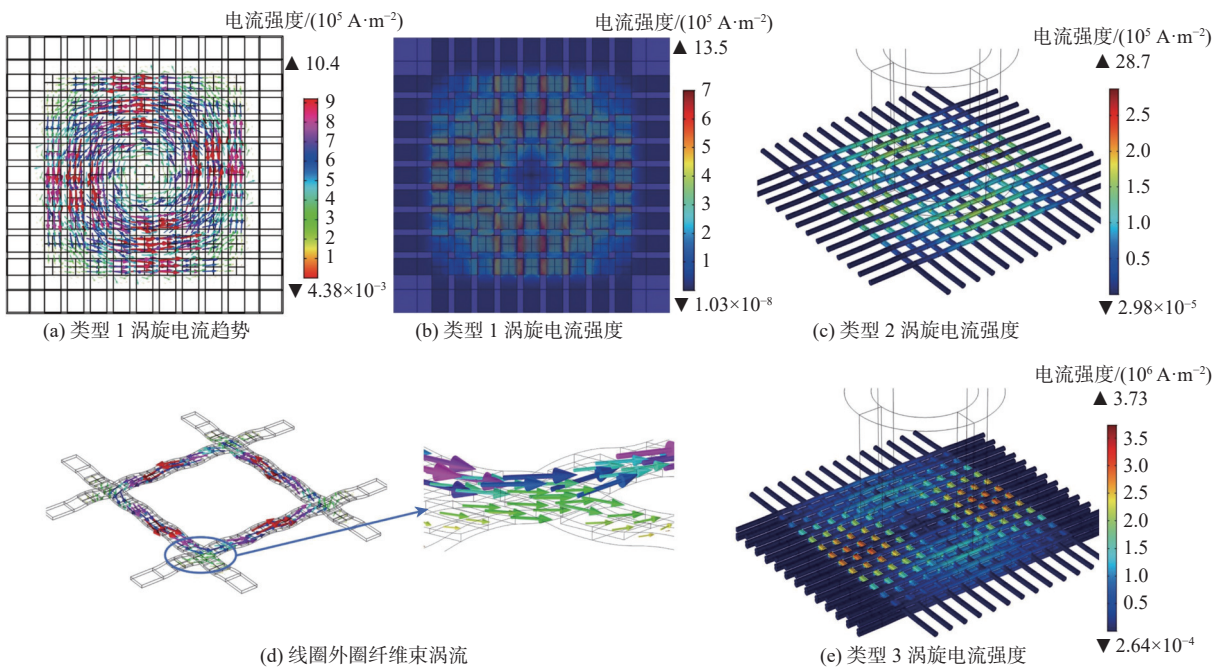


图 5　纤维编织结构中涡旋电流分布

Fig. 5　Eddy current distribution in fiber braided structures

布情况，CFRP中的发热量是由纤维织构中所产生的涡旋电流与其自身电阻作用所产生的焦耳热，因此发热量与涡旋电流的分布区域基本一致。可以看出类型1和类型2的发热量分布基本相同，均集中在线圈所对应的外圆纤维织构区域且发热量在外圆纤维束中的分布相对均匀，而类型3中所产生的发热量主要集中在沿y轴上的纤维束中，沿x方向上纤维束中的发热量明显减少，也说明了纤维排布方式不同对材料内部参数分布状态有着明显影响。

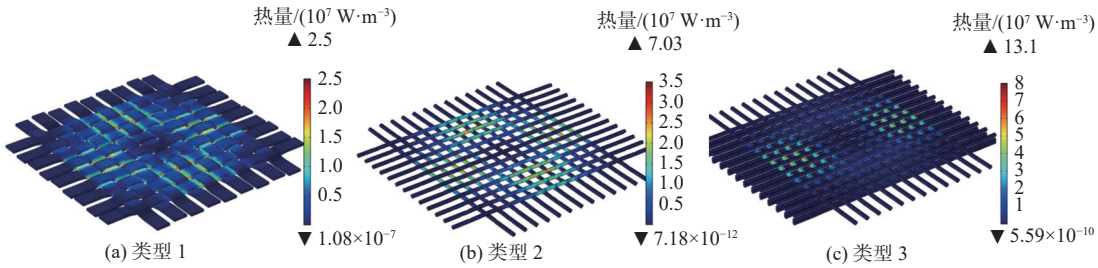


图6 纤维织构中的发热量分布

Fig. 6 Heat distribution in fiber texture

4 材料加热状态分析

图7为3种编织结构的CFRP加热至稳态时温度场分布状态，可以看出类型1和类型2的温度场分布状态基本相同，温度场集中在线圈所对应材料区域，呈十字型向外分布，其中类型1的温度数值更高。而类型3的温度场集中在纤维织构的中心区域并呈现长条形分布，这是由于该类型中x方向上的纤维束排布更多其热量的传导较y方向更加迅速，并在左右2个温区的共同作用下，造成了温度场主要沿x方向分布的现象。

CFRP在交变磁场的作用下所产生的焦耳热为其固化成型提供了有效热量，图8为3种类型的CFRP在感应加热温度数值基本相同的条件下，达到30 min左右时的固化度分布情况，从图中可以看出类型1和类型2编织结构的CFRP固化度分布状态基本相同，而类型2的固化度最高值明显高于类型1，且类型2在沿着x、y方向上的固化度更加均匀，由于纤维铺层方式的原因在四角处的固化梯度下降较为明显。而类型3受到其温度场分布状态的影响，沿着x方向上的树脂固化程度更好且材料整体固化效果高于类型1和类型2，说明此时单向

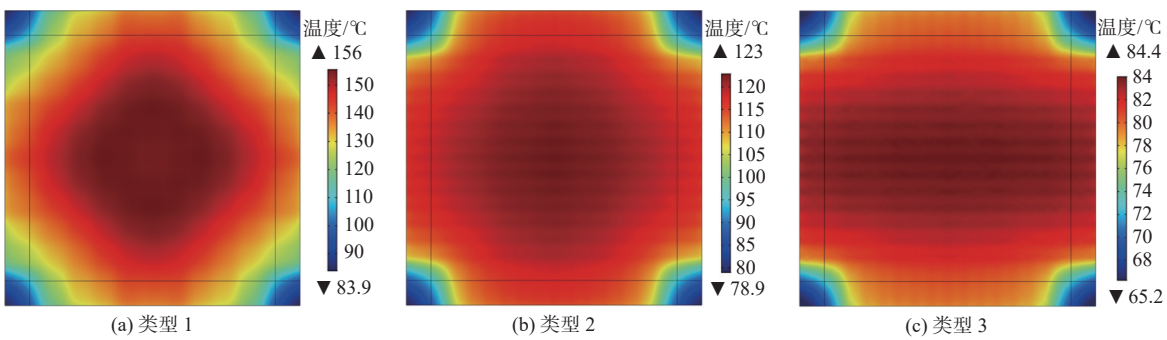


图7 各类型CFRP加热至稳态时温度场分布状态

Fig. 7 Distribution of temperature field when each type of CFRP is heated to a steady state

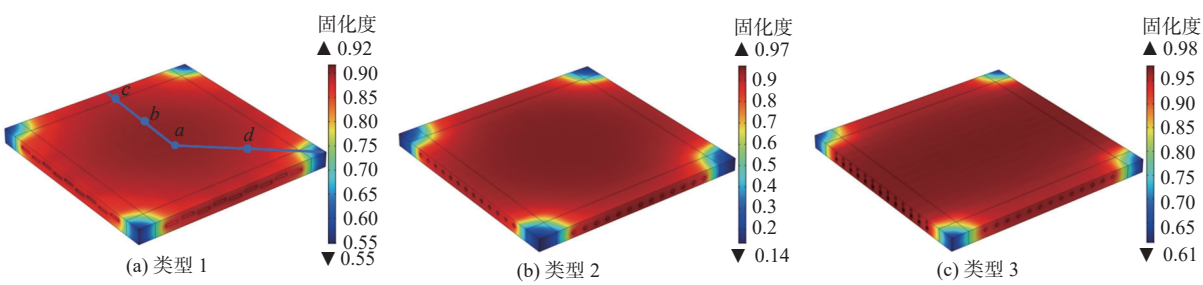


图8 各类型CFRP固化度分布

Fig. 8 Distribution of curing degree of various types of CFRP

铺层结构的固化效果高于平纹编织结构, 且材料内部的纤维含量对固化效果起到明显作用。

图 9 为 3 种编织结构的 CFRP 加热过程中升温速率及固化度变化曲线, 其采集点位在材料 y 方向及对角线上共 4 处如图 8 所示, 从数值采集曲线可以看出, 3 种类型的材料中心区域温度及固化度上升速率均最快。类型 1 中各点的升温速率和固化度上升速率较为接近, 说明该结构的温度场及固化度场均匀性较高, 类型 2 在加热过程中温度场及固化度场均匀性较类型 1 有所下降, 而材料的固化度上升速率明显提高, 类型 3 中的温度及固化度上升速率较快且在短时间内迅速提高, 其中该固化度 S 型曲线明显左移。从图 9 中可以看出升温速率与固化度变化曲线存在相互匹配关系, 且固化度曲线相对升温速率曲线存在滞后现象, 说明了纤维体积含量对升温速率及固化度上升起到促进作用。

在 CFRP 加热固化至稳态后其内部应力场分布情况如图 10 所示, 为与实际测试情况相对比, 在模拟计算时设置材料底面为固定平面并调整 3 种类型的温度场数值均在 120 ℃ 左右。从图 10 中类型 1 可以看出平纹编织织构在应力场的作用下形变较为明显, 尤其在织构边缘处纤维束产生了交替浮动的形变, 在纤维织构的四角处的应力集中更加明显; 类型 2 单向均匀交织结构所产生的应力交平纹编织结构更高, 织构边缘处的纤维束形变量基本相同且变形方向一致; 类型 3 单向非均匀交织结构沿 y 方向纤维束的应力明显高于 x 方向且形变量较大, 该铺层结构的应力明显高于前 2 种铺层结构, 说明了纤维铺层均匀性影响了材料应力的分布状态, 同时纤维束交点处应力更为集中。

从图 10 整体应力分布及形变规律看到, 类型 1 和类型 2 的材料整体形变量及形变状态基本相同均向上凸起, 尤其在材料的四周边缘位置产生的形变量更加明显, 也可以看出材料整体的应力场分布及数值大小基本相同。类型 3 的材料整体形变主要沿着 y 方向产生形变, 沿着 x 方向上的形变量较少, 可以看出材料中心区域的整体形变减少, 在纤维含量较多的方向可以控制材料的形变程度。

图 11 为各类型材料加热过程中应力变化曲线, 应力变化采集点如图 7 所示, 采集位置为纤维与树脂基体的交界处, 从图 10 可以看出 3 种类型的 CFRP 其应力分布和上升过程均不相同, 类型 1 中

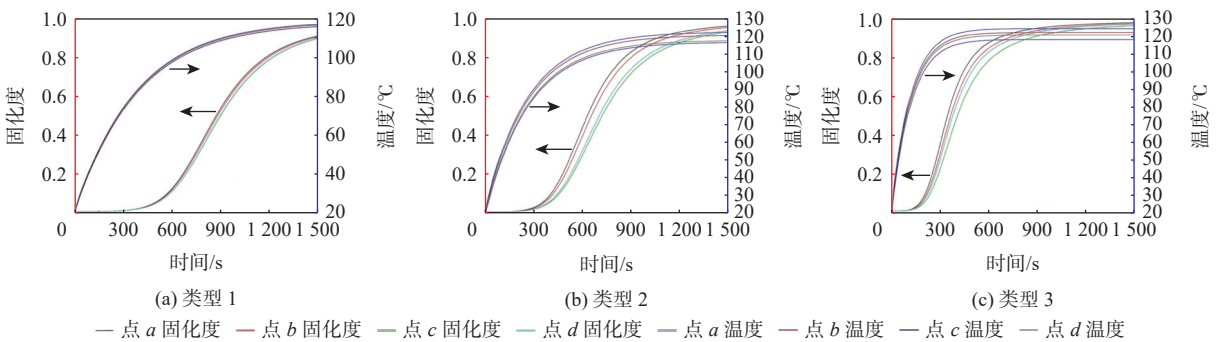


图 9　CFRP 加热过程中升温速率及固化度变化曲线

Fig. 9　Variation curves of heating rate and curing degree during CFRP heating process

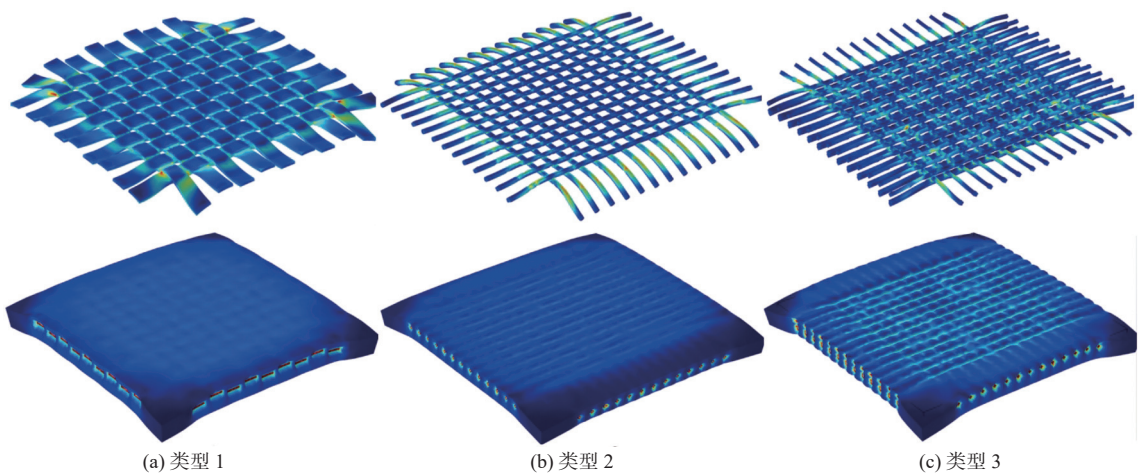
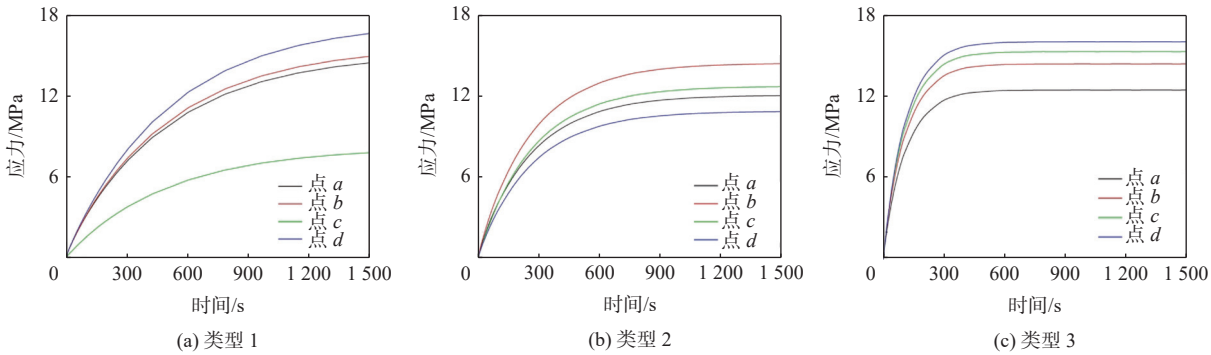


图 10　各纤维织构及材料整体的应力场分布

Fig. 10　Stress field distribution of each fiber texture and material as a whole



(a) 类型 1　(b) 类型 2　(c) 类型 3

图 11　各类型材料加热过程中应力变化曲线

Fig. 11　Stress change curves of various types of materials during heating

点 *d* 处的应力大小及曲线上升速率最高，说明该区域温度场均匀性较差，点 *a* 和点 *b* 的应力曲线变化较为接近，此时两点之间的温度梯度相差较小且固化速率相对一致，而点 *c* 处的应力明显低于其他位置。类型 2 整体的应力曲线明显更小且曲线相对接近，说明类型 2 在温度上升及固化效果上更加均匀，此时在线圈对应的 *b* 点处应力最高。类型 3 中点 *d* 处的应力明显更高，而中心区域的应力最低，从中心区域到边界应力大小逐渐上升。通过对 3 种类型的应力曲线变化分析，说明了 CFRP 内部的纤维铺层方式对加热过程中应力场的分布存在明显影响。

5　实验验证

图 12 为感应加热 CFRP 实验平台，包括了热成像仪、感应加热线圈、升降平台、可调控电源及其他加热和检测设备。

从图 12 可以看出 CFRP 在加热过程中材料预浸布放置在支撑座上表，感应线圈放在支撑座下侧，且支撑座表面为非金属材质，以降低加热过程中线圈的能量损耗。



图 12　感应加热实验平台

Fig. 12　Induction heating experimental platform

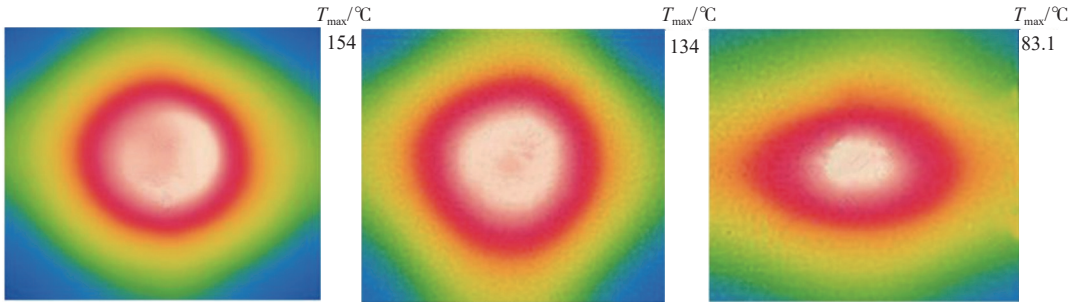
为验证本文所建立的有限元模型对 CFRP 感应加热各物理场状态分析的准确性，进行了 CFRP 感应加热温度场及应力分析实验，在实验过程中各线圈直径、匝数、间距、电流及频率等相关参数均相等。图 13 为 3 种结构的 CFRP 在感应加热 20 min 左右表面温度场分布情况，可以看出类型 1 和类型 2 的温度场分布状态基本接近，但是类型 1 在加热时所获得的温度数值更高，且温度场分布的十字型更加规整。类型 3 的温度场则主要沿着 *X* 方向呈现长条形状进行分布，其温度场数值最低，与图 5 所示的 3 种结构的 CFRP 感应加热温度场仿真结果基本相同，说明本文所建立的有限元细观模型可以实现 CFRP 感应加热温度场分布规律的准确验证。本文使用静态电阻式应变仪对 CFRP 固化过程中的应变进行单层多点数据采集，应变片设置于第 1 和第 2 层中间，对比模型预测应力数据与实验应力数据，验证本文所建立的有限元模型对应力预测的准确性。图 14 为平纹编织 CFRP 感应加热应力实验采集装置，图 14(a)为平纹编织材料结构和线圈尺寸，与仿真模型结构等比例扩大，应力采集点位置与模拟采集位置基本相符，此时感应加热实验中线圈直径为 90 mm，匝数为 45，电流大小 16 A，频率大小为 13 kHz，外界环境温度 17～20℃，CFRP 预浸布尺寸为 120 mm×120 mm，图 14(b)为瞬态应力采集装置，可实现材料加热过程中应力变化瞬时表征和数据采集。

通过实验测量获得各点应变随时间变化的数据，代入式 (12) 中即得到 CFRP 感应加热固化过程中个点应力变化曲线。

$$\sigma_t = E_t \varepsilon_t \tag{12}$$

式中：σ_t为总应力；ε_t为应变；E_t为材料的瞬时弹性模量。

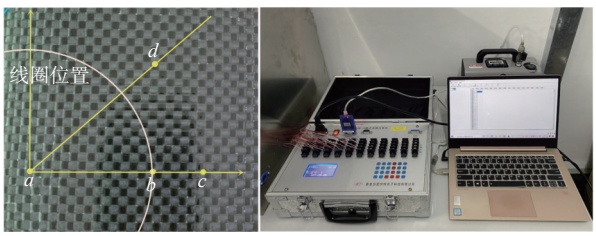
图 15 为平纹编织 CFRP 在感应加热实验中各



(a) 类型 1　　(b) 类型 2　　(c) 类型 3

图 13　感应加热实验材料温度场分布状态

Fig. 13　Distribution of temperature field of materials in induction heating experiment



(a) 平纹编织 CFRP　　(b) 瞬态应力采集装置

图 14　感应加热应力采集装置

Fig. 14　Induction heating stress acquisition device

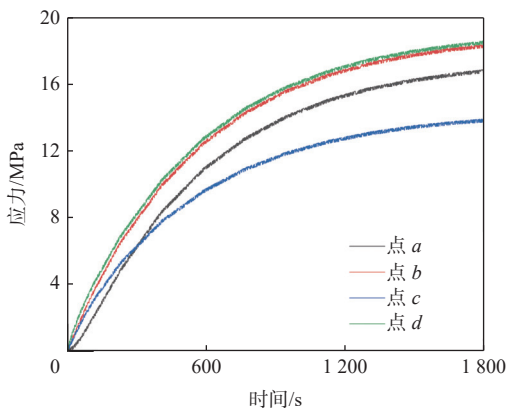
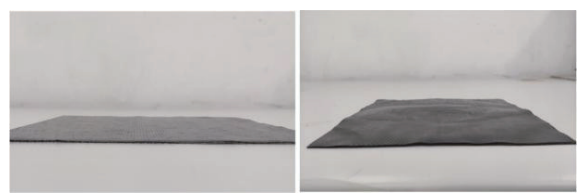


图 15　平纹编织 CFRP 感应加热实验中各点应力变化曲线

Fig. 15　Variation curve of stress at each point in induction heating experiment of plain weave CFRP

点应力变化曲线, 从图中可以看出各点应力的上升过程与模拟计算结果基本相符, 但是点 *a* 和点 *c* 的曲线上升过程存在交叉, 这是由于实验中线圈尺寸较大点位选取时 *c* 点距离线圈边缘更近, 导致点 *c* 处初始升温和固化较快使得此时应力数值上升较快, 随着加热时间延长点 *a* 处的温度较点 *c* 处更加集中。实验中各点位置应力数值均小于仿真结果其偏差越 10% 左右, 证明了本文所建立的有限元模型可以准确表征材料在感应加热过程中的应力变化情况, 然而由于传感器在交变磁场的作用下也导致了各点采集的应力曲线波动较高。

图 16 为实验用平纹编织 CFRP 预浸布铺层及采用感应加热固化后的材料状态, 可以看出固化后材料四边的形变较为明显如图 16(b) 所示, 材料中心区域出现相应褶皱, 与模拟计算中纤维织构的形变效果基本相同, 然而感应加热升温速率较快, 在材料平面温度场尚未达到均匀时, 产生局部固化导致中心区域形变与计算结果略有偏差, 因此后续研究可设计磁场分布及线圈结构提高材料升温过程中的温度场均匀性, 以改善材料固化成型质量。



(a) 预浸布铺层　　(b) 加热固化后状态

图 16　预浸布铺层及固化后的材料状态

Fig. 16　Prepreg layup and cured material condition

6　结　论

1) 文本根据 CFRP 内部纤维束的实际铺层结构及等效电-磁-热特性, 采用了 CFRP 感应加热有限元细观建模方法, 模拟计算了不同铺层方式的 CFRP 在交变磁场作用下的涡旋电流、发热量及温度场的时空变化规律, 其中涡旋电流和发热量的分布与纤维排布方式直接相关, 而纤维在铺层向上的体积分数提高了该方向的热传导效率, 导致材料温度场分布状态的变化。

2) 在材料参数及材料加热温度最高值均相同的条件下, 类型 3 结构的 CFRP 升温及固化度上升速率最快, 类型 2 平纹编织结构的 CFRP 各点处的温度及固化度分布均匀性最高, 并且材料的温速率与固化度变化曲线存在相互匹配关系, 固化度曲线相对升温速率曲线存在滞后现象, 其中纤维体积含量也对升温速率及固化度上升起到了相应的促进作用。

3) 文本所建立的有限元细观模型可以有效反

映出材料在感应加热过程中的应力场分布及型变情况。根据各类型CFRP的应力场分析可以看出纤维铺方式对其应力场分布产生较为明显的影响，材料应力场的变化规律受到了温度场的影响，单向均匀交织的CFRP在感应加热过程中所产生的应力场最小。通过感应加热实验验证了本文所建立的模型对各物理参数表征的准确性，为CFRP感应加热相关研究提供了有效的计算模型和分析方法。

参考文献（References）

[1] NING F, CONG W, QIU J, et al. Additive manufac-turing of carbon fiber reinforced thermoplastic composites using fused deposition modeling[J]. Composites Part B Engineering, 2015, 80: 369-378.

[2] SAMBRUNO A, Banon F, SALGUERO J, et al. Study of milling of low thickness thermoplastic carbon fiber composites in function of tool geometry and cutting conditions[J]. The International Journal of Advanced Manufacturing Technology, 2021, 114(7-8): 2515-2526.

[3] ABLIZ D, YUGANG D. Curing methods for advanced polymer composites-A review[J]. Polymers & Polymer Composites, 2013, 21(6): 101-106.

[4] SUN X F, MA S N, LIU H W. Research on properties of carbon fiber/epoxy resin composite material by microwave curing[J]. Rare Metal Materials and Engineering, 2012, 41: 422-424.

[5] MENANA H, FELIACHI M. An integro-differential model for 3-D eddy current computation in carbon fiber reinforced polymer composites[J]. IEEE Transactions on Magnetics, 2011, 47(4): 756-763.

[6] BAYERL T, DUHOVIC M, MITSCHANG P, et al. The heating of polymer composites by electromagnetic induction-A review[J]. Composites Part A, 2014, 57: 27-40.

[7] JUN C, JINHAO Q, TOSHIYUKI T. Numerical analysis of correlation between fibre orientation and eddy current testing signals of carbon-fibre reinforced polymer composites[J]. International Journal of Applied Electromagnetics and Mechanics, 2012, 39(1-4): 251-259.

[8] 付天宇, 许家忠, 赵辉, 等. CFRP感应加热线圈中心区域温度场[J]. 复合材料学报, 2021, 38(10): 3314-3322.
FU T Y, XU J Z, ZHAO H, et al. Temperature field in the central area of CFRP induction heating coil[J]. Acta Materiae Compositae Sinica, 2021, 38(10): 3314-3322(in Chinese) .

[9] FU T Y, XU J Z, HUI Z. Analysis of induction heating temperature field of plain weave CFRP based on finite element meso model[J]. Applied Composite Materials, 2021, 28(31): 149-163.

[10] DAS T K, GHOSH P, DAS N C. Preparation, development, outcomes, and application versatility of carbon fiber-based polymer composites: a review[J]. Advanced Composites and Hybrid Materials, 2019, 2(2): 214-233.

[11] LUNDSTRÖM F, FROGNER K, ANDERSSON M. A method for inductive measurement of equivalent electrical conductivity in thin non-consolidated multilayer carbon fiber fabrics[J]. Composites Part B Engineering, 2017, 140: 204-213.

[12] LUNDSTRÖM F, FROGNER K, WIBERG O, et al. Induction heating of carbon fiber composites: Investigation of electric and thermal properties[J]. International Journal of Applied Electromagnetics and Mechanics, 2017, 53: S21-30.

[13] BRUCE K F, ROY L M, JOHN W G. A local theory of heating in cross-ply carbon fiber thermoplastic composites by magnetic induction[J]. Polymer Engineering & Science, 1992, 32(5): 357-369.

[14] MITSCHANG P, RUDOLF R, NEITZEL M. Continuous induction welding process, modelling and realisation[J]. Journal of Thermoplastic Composite Materials, 2002, 15(2): 127-153.

[15] BANERJEE A, AVESTRUZ A T, SURAKITBOVORN K, et al. Uniform single-sided induction heating using multiphase, multi-resonant halbach windings [C]//Applied Power Electronics Conference & Exposition. Piscataway: IEEE Press, 2014: 844-851.

[16] FREDRIK L, KENNETH F, MATS A. A numerical model to analyse the temperature distribution in cross-ply CFRP during induction heating[J]. Composites Part B:Engineering, 2020, 202: 108419.

[17] TZENG J T, HSIEH K T. Electromagnetic analysis of composite structures subjected to transient magnetic fields[J]. Journal of Composite Materials, 2019, 54(6): 745-752.

[18] RICCIO A, RUSSO A, RAIMONDO A, et al. A numerical/ experimental study on the induction heating of adhesives for composite materials bonding[J]. Materials Today Communications, 2018, 15: 203-213.

[19] LIU C W , QU C Y, HAN L, et al. Preparation of carbon fiber-reinforced polyimide composites via in situ induction heating [J]. High Performance Polymers, 2017, 29(9): 263-273.

Characterization of influence of fiber arrangement on CFRP induction heating curing process

FU Tianyu[1,*], YANG Ning[2], GU Yunfei[3], XU Jiazhong[4], JIANG Yue[3], LI Bin[5]

(1. Department of Intelligent Equipment, Changzhou College of Information Technology, Changzhou 213164, China;
2. School of Computer and Network Engineering, Shanxi Datong University, Datong 037009, China;
3. School of Mechanical and Power Engineering, Harbin University of Science and Technology, Harbin 150080, China;
4. School of Automation, Harbin University of Science and Technology, Harbin 150080, China;
5. Chongqing Special Equipment Testing and Research Institute, Chongqing 401120, China)

Abstract: In the process of using electromagnetic induction heating to achieve the curing and shaping of carbon fiber reinforced Polymer (CFRP) materials, accurately characterizing the property parameters of the material's various components is of utmost importance in studying temperature, degree of curing, and stress field distribution during the heating process. We have established a finite element microanalysis model for CFRP induction heating based on the actual fiber layup method for independently characterizing the various components of the material and the equivalent electro-magnetic-thermal properties of the carbon fiber shell. This model separates the fiber texture and resin matrix, allowing for a systematic analysis of the changes in physical fields within the internal fiber structure of CFRP during the induction heating process and their effects on the overall temperature variation of the material. Through this model, we have computed the variations in material temperature, degree of curing, and stress field distribution, revealing the influence of carbon fiber arrangement on the effectiveness of induction heating. Furthermore, we have validated the model's ability to accurately represent the distribution states of various physical fields through induction heating experiments. This model provides an effective computational framework and analytical approach for studying the changes in various physical fields during CFRP induction heating.

Keywords: carbon fiber; induction heating; temperature field; degree of cure; stress field

Received: 2022-04-19; **Accepted**: 2022-06-20; **Published Online**: 2022-07-25 12:57

URL: link.cnki.net/urlid/11.2625.V.20220722.1531.002

Foundation items: National Natural Science Foundation of China (52303031); Changzhou Applied Basic Research Plan (CJ20220035); Research Laboratory of Carbon Fiber Pressure Vessel Forming Technology (KYPT202203Z)

* **Corresponding author**. E-mail: 827263230@qq.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0221

三维自由弯曲过程中临界偏距对管材成形精度影响

范鑫[1,2]，程宗辉[1,2,3]，李少鑫[3]，舒送[1,2]，程诚[2,3,*]，郭训忠[3]

(1. 安徽省航空设备测控与逆向工程实验室，芜湖 241007；2. 国营芜湖机械厂，芜湖 241007；
3. 南京航空航天大学 材料科学与技术学院，南京 210016)

摘　　要：三维自由弯曲成形过程中弯管处于少约束状态，弯管弯曲半径的大小及轴线形状取决于相应的弯曲模偏距值。为了准确成形出预设弯管的形状，提高弯管的自由弯曲成形精度，基于弯曲模临界偏距 U_e 建立了弯曲模偏距 U 和弯管弯曲半径 R（U-R）关系数学模型。采用有限元模拟与实验相结合的方法，研究了不同摩擦系数及材料参数条件下临界偏距的演化规律及其对弯管成形精度的影响。研究结果表明：引入临界偏距的 U-R 关系拟合结果更加符合实验结果。随着摩擦系数的增加，弯管塑性变形程度增加，临界偏距的数值降低，弯管弯曲半径减小。相比于 SS304 不锈钢弯管，相同工艺参数下 6061 铝合金弯管的弯曲半径增大，临界偏距降低。

关　键　词：管材；三维自由弯曲；弹性-塑性变形；摩擦系数；材料参数
中图分类号：TG386
文献标志码：A　　**文章编号**：1001-5965(2024)01-0208-08

复杂形状金属弯曲管件因其高强度、高精度、轻质量和低能耗等特点被广泛应用于航空航天、核能装备、舰船、汽车及能源输送等工业领域[1-2]。目前，金属管件的弯曲变形主要是通过绕弯、推弯、滚弯和拉弯等传统的弯曲成形技术来实现。然而，由于轴线形状及成形工装的限制，上述传统弯曲成形技术难以实现空间多弯、轴线曲率连续变化的复杂弯曲构件一次整体成形。三维自由弯曲成形技术通过伺服联动控制技术控制弯曲模的运动轨迹，以实现复杂形状管材、线材和型材等构件的弯曲成形[3]。

在三维自由弯曲成形过程中，位于弯曲模和导向机构之间的弯管处于少约束状态，弯曲构件的成形质量受材料属性及工艺参数的影响变化非常显著。近些年来，相关学者就不同材料参数、管模间隙和弯曲角速度等因素对管材自由弯曲成形质量的影响规律开展了大量研究。程旋[4]探索了屈服强度和弹性模量2个材料参数对矩形管材自由弯曲成形的影响规律，研究表明随着弹性模量的增加，成形管材弯曲半径逐渐变小；而随着屈服强度的增加成形管材弯曲半径逐渐变大。郝用兴等[5]研究了管模间隙及弯曲模圆角参数对管材自由弯曲成形质量的影响，研究表明管模间隙的增大，将导致管材的弯曲半径逐渐增大；同时还发现弯曲模圆角半径较小时，管材截面畸变严重，随着圆角半径的增大，椭圆率下降，成形质量好。李秋等[6]通过建立管材三维自由弯曲成形有限元模型分析探讨了弯曲角速度对管材弯曲变形行为的影响，研究表明弯曲角速度越大，管材的内外侧壁厚变化越大，越容易发生拉裂和起皱等成形缺陷。

金属管材三维自由弯曲成形过程中，管材的弯曲弧段会经历从弹性变形到塑性变形的转变[7]。目前，很少有学者对金属管材弯曲成形过程弹性-塑性变形临界点进行研究，都是基于塑性成形理论分

收稿日期：2022-04-06；录用日期：2022-07-22；网络出版时间：2022-10-18 11:14
网络出版地址：link.cnki.net/urlid/11.2625.V.20221017.1624.008
基金项目：国家自然科学基金 (52105360,52175328,U1937206)；数字制造装备与技术国家重点实验室开放课题 (DMETKF2021004)；江苏省工业感知及智能制造装备工程研究中心开放基金 (2021-04)；江苏省“双创博士”项目 (JSSCBS20210157)
* 通信作者. E-mail：c_cheng@nuaa.edu.cn

引用格式：范鑫，程宗辉，李少鑫，等. 三维自由弯曲过程中临界偏距对管材成形精度影响 [J]. 北京航空航天大学学报，2024，50（1）：208-215. FAN X，CHENG Z H，LI S X，et al. Effect of critical eccentricity on forming accuracy of tubes in 3D free bending process [J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50（1）：208-215（in Chinese）.

析金属管材弯曲成形过程中的失稳破坏现象进行弯曲成形极限的相关研究。Sun 等[8] 提出引导因子和动态参考点优化非支配排序遗传算法(non-dominated sorting genetic algorithms, NSGA), 通过有限元模拟和弯曲实验, 实现了管材弯曲回弹的数据集构建和精度验证。Yang 等[9] 采用差分加热法对钛合金管进行弯曲成形, 研究得出整体加热可打破管材的弯曲极限, 抑制弯曲过程中的起皱现象。戴莉等[10] 分析了材料参数波动对金属管材弯曲成形过程的失稳起皱影响规律, 研究发现随着屈服强度的增大或弹性模量、硬化指数的减小, 管材的起皱趋势增大, 泊松比和强度系数对管材起皱趋势的影响较小。然而, 在金属管材自由弯曲成形过程中, 弹性-塑性变形临界点的确定对于准确确定弯管最大弯曲半径和轴线形状, 提高弯曲管材成形质量具有重要意义。

本文从三维自由弯曲原理出发, 首先, 分析管材成形过程中的应力应变分布情况, 揭示管材自由弯曲成形机理; 其次, 通过有限元模拟和自由弯曲成形实验方法, 确定弹性-塑性变形临界点对应的临界偏距 U_e; 最后, 基于临界偏距 U_e 建立管材弯曲模偏距 U 和弯管弯曲半径 R(U-R)关系数学模型, 并分析摩擦系数和材料参数对临界偏距 U_e 的影响规律。

1　自由弯曲成形管材变形机理分析

1.1　自由弯曲原理分析

金属管材三维自由弯曲成形过程的原理示意图如图 1 所示。当弯曲模在 xOy 平面内移动时, 处于弯曲模与导向模之间的弯曲弧段始终受到推进机构、弯曲模施加的作用力 F_1、F_2 的影响, 形成弯曲力矩 M 从而使管材发生弯曲变形。其中弯曲模中心点与导向模前端沿 z 轴方向的距离称为 A, 弯曲模中心点与管材中心轴线在 xOy 平面内的投影距离称为偏距 U。由于弯曲模可以在 xOy 平面内沿任意方向移动来改变偏距 U 的数值大小, 因此, 管材的弯曲方向可以在 0°～360°之间进行调整, 同时管材的弯曲半径也能够基于偏距 U 的数值大小而进行相应的变化。弯曲力矩 M 与作用力 F_1 和 F_2 之间的关系可以描述为

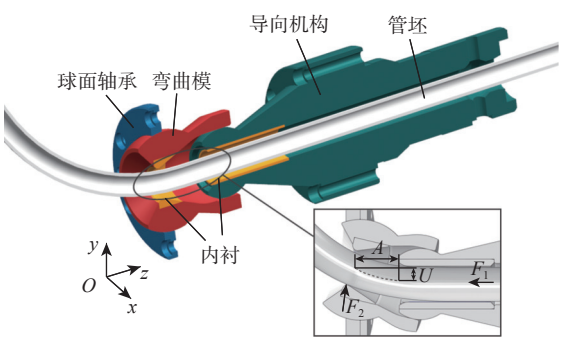


图 1　三维自由弯曲成形原理示意图

Fig. 1　Schematic diagram of 3D free bending principle

$$M = F_1 U + F_2 A \tag{1}$$

1.2　管材应力状态解析

三维自由弯曲成形过程, 处于弯曲模和导向模间的弯管应力状态和几何参数如图 2 所示, O 为管材半径的圆心点, d_0 为管材直径, σ_θ、σ_ρ、σ_φ分别为切向应力、径向应力、周向应力, ε_θ、ε_ρ、ε_φ分别为切向应变、径向应变、周向应变, R_i、R_o分别为弧段最内侧、最外侧曲率半径, ρ为弧段任意一点处曲率半径, R_0、R_ξ分别为原始中性层、应变中性层。其中, 弯管圆弧内侧处于压应力状态, 弯管圆弧外侧处于拉应力状态。

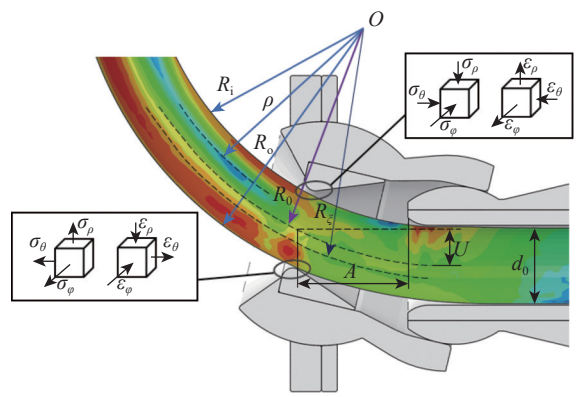


图 2　三维自由弯曲弯管应力状态和几何参数示意图

Fig. 2　Schematic diagram of stress state and geometric parameters of a 3D free bending bent tube

根据 Hencky 应力应变关系[11] 和管材体积不可压缩假设 ($\varepsilon_\rho+\varepsilon_\theta+\varepsilon_\varphi=0$), 弯管径向力平衡方程和横截面周向应力可以表示为

$$\frac{\mathrm{d}\sigma_\rho}{\mathrm{d}\rho}+\frac{\sigma_\rho-\sigma_\theta}{\rho}=0 \tag{2}$$

$$\sigma_\varphi=\frac{\sigma_\theta+\sigma_\rho}{2} \tag{3}$$

基于 von Mises 屈服准则[12] 和管材的周向应变 ($\varepsilon_\varphi=0$)的假设, 弯管的等效应力$\bar{\sigma}$和等效应变$\bar{\varepsilon}$分别可以简化为

$$\bar{\sigma}=\frac{\sqrt{3}}{2}\left|\sigma_\theta-\sigma_\rho\right| \tag{4}$$

$$\bar{\varepsilon}=\frac{2}{\sqrt{3}}\left|\varepsilon_\theta\right| \tag{5}$$

结合 Hollomon 方程[13] 和轴向推力引起的切向压应力σ_N拟合弯管的流动应力-应变关系:

$$\bar{\sigma}=\frac{\sqrt{3}}{2}\left|\sigma_\theta-\sigma_\rho\right|+\sigma_N \tag{6}$$

考虑到自由弯曲过程中,在轴向推力的作用下弯管瞬时应变中性层呈现由弯管内侧向外侧移动的趋势,弯管切向应变可以表示为

$$\varepsilon_\theta = \ln\frac{\rho}{R_\xi} \tag{7}$$

结合式(3)~式(7),弯管切向应力和径向应力之间的关系可以推导为

$$\sigma_\theta - \sigma_\rho = \pm K\left(\frac{2}{\sqrt{3}}\right)^{n+1}\left(\pm\ln\frac{\rho}{R_\xi}\right)^n - \frac{2}{\sqrt{3}}\sigma_N \tag{8}$$

式中:K为强度系数;n为硬化指数。当$\rho \geqslant R_\xi$时,取“+”;当$\rho < R_\xi$时,取“−”。

将式(8)代入式(2)中,可以获得以下关系:

$$\mathrm{d}\sigma_\rho = \left[\pm K\left(\frac{2}{\sqrt{3}}\right)^{n+1}\left(\pm\ln\frac{\rho}{R_\xi}\right)^n - \frac{2}{\sqrt{3}}\sigma_N\right]\frac{\mathrm{d}\rho}{\rho} \tag{9}$$

由于弯管的外拱面和内凹面处的径向应力为0,根据计算的积分常数,弯管内侧和外侧的径向应力可以分别表述为

$$\sigma_{\rho\mathrm{i}} = \left(\frac{2}{\sqrt{3}}\right)^{n+1}\frac{K}{n+1}\left(-\ln\frac{R_0-R_\mathrm{i}}{R_\xi}\right)^{n+1} + \frac{2}{\sqrt{3}}\sigma_N\ln\frac{R_\mathrm{i}}{R_0} \tag{10}$$

$$\sigma_{\rho\mathrm{o}} = \left(\frac{2}{\sqrt{3}}\right)^{n+1}\frac{K}{n+1}\left(-\ln\frac{R_0+R_\mathrm{o}}{R_\xi}\right)^{n+1} + \frac{2}{\sqrt{3}}\sigma_N\ln\frac{R_\mathrm{o}}{R_0} \tag{11}$$

将式(10)和式(11)代入式(8),弯管外侧和内侧的切向应力可推导为

$$\sigma_{\theta\mathrm{i}} = \left(\frac{2}{\sqrt{3}}\right)^{n+1}\frac{K}{n+1}\left[\left(-\ln\frac{R_0-R_\mathrm{i}}{R_\xi}\right)^{n+1} - \left(-\ln\frac{\rho}{R_\xi}\right)^{n+1}\right] + \frac{2}{\sqrt{3}}\sigma_N\ln\frac{R_\mathrm{i}}{R_0} \tag{12}$$

$$\sigma_{\theta\mathrm{o}} = \left(\frac{2}{\sqrt{3}}\right)^{n+1}\frac{K}{n+1}\left[\left(-\ln\frac{R_0+R_\mathrm{o}}{R_\xi}\right)^{n+1} + \left(-\ln\frac{\rho}{R_\xi}\right)^{n+1}\right] + \frac{2}{\sqrt{3}}\sigma_N\ln\frac{R_\mathrm{o}}{R_0} \tag{13}$$

结合式(3)、式(8)、式(12)和式(13),弯管外侧和内侧的等效应力可推导为

$$\overline{\sigma_\mathrm{i}} = \left(\frac{2}{\sqrt{3}}\right)^{n+1}\frac{K}{2}\left[\frac{2}{n+1}\left(-\ln\frac{R_0-R_\mathrm{i}}{R_\xi}\right)^{n+1} - \left(-\ln\frac{\rho}{R_\xi}\right)^{n}\right] + \frac{2}{\sqrt{3}}\sigma_N\ln\frac{R_\mathrm{i}}{R_0} \tag{14}$$

$$\overline{\sigma_\mathrm{o}} = \left(\frac{2}{\sqrt{3}}\right)^{n+1}\frac{K}{2}\left[\frac{2}{n+1}\left(-\ln\frac{R_0+R_\mathrm{o}}{R_\xi}\right)^{n+1} + \left(-\ln\frac{\rho}{R_\xi}\right)^{n}\right] + \frac{2}{\sqrt{3}}\sigma_N\ln\frac{R_\mathrm{o}}{R_0} \tag{15}$$

根据式(14)和式(15)的推导过程可以发现,在轴向推力的作用下,弯管内侧的应力要高于弯管外侧,因此,在材料拉-压变形屈服应力对称性的假设下,弯管内侧要先于弯管外侧发生塑性变形,并且其塑性变形行为也更加明显。

2 自由弯曲成形弯曲模临界偏距的确定

2.1 临界偏距的定义

在金属管材三维自由弯曲过程中,弯管内侧和外侧区域都要经历从弹性变形向塑性变形的转变过程[14]。管材弯曲弧段的形成是由材料发生塑性变形而引起的,因此,金属管材弯曲弧段的形成存在一个弹性-塑性变形的临界点,其对应于弯曲模的临界偏距U_e。图3为平面螺旋典型弯曲构件弯曲模偏距U与弯管弯曲半径R的映射原理。由图3可知,当弯曲模的偏距U在U_e范围内变化时,金属管材不会发生塑性弯曲变形,不能形成弯曲圆弧段;而当弯曲模偏距$U>U_\mathrm{e}$时,随着偏距U的增大,相应的弯管曲率$1/R$也越来越大。

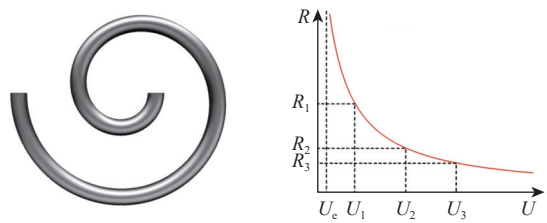


图3 平面螺旋典型弯曲构件的U-R映射原理图

Fig. 3 Schematic mapping of dimension characteristics and U-R relationship of planar involute components

考虑到自由弯曲过程中金属管材弯曲半径的大小取决于相应弯曲模偏距U的数值,同时由弯曲模的移动所引起的弯管轴线过渡变化也会对金属管材的最终弯曲形状产生影响。因此,准确确定临界偏距U_e对获得自由弯曲成形弯管U-R理论关系及确定弯曲模的运动轨迹具有重要意义。

2.2 三维自由弯曲有限元模型建立

本文以SS304不锈钢管材为研究对象,具体尺寸:直径为20 mm,壁厚为1 mm。为了准确进行三维自由弯曲成形模拟,需要获得SS304管材基本材料的力学性能参数。以GBT 228-2010为标准,通过室温单向拉伸实验获得SS304管材流动应力-应变曲线。金属管件的自由弯曲成形为冷加工,加工硬化现象明显,本文采用Hollomon硬化方程($\sigma = K\varepsilon^n$,其中ε为应变)来拟合管材的弹性变形阶段,结果如图4所示。由图4可知,实验数据与拟合结果高度相似,SS304在弯曲变形过程中的应力应变状态可真实反映出来,拟合获得的本构模型和SS304力

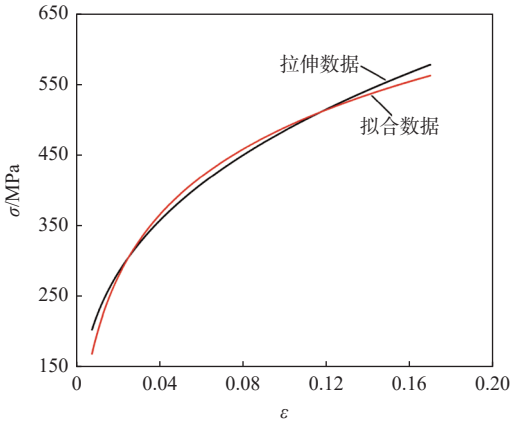


图 4　SS304 管材流动应力-应变曲线

Fig. 4　True stress-strain curves of SS304 tube

学性能参数如式(16)和表 1 所示。

$$\sigma = 1\,040.884\varepsilon^{0.082} \tag{16}$$

表 1　SS304 管材力学性能参数

Table 1　Mechanics performance parameters of SS304 tube

弹性模量 E/GPa	延伸率 Δ/%	屈服强度 $\sigma_{0.2}$/MPa	抗拉强度 σ_b/MPa	强度系数 K/MPa	硬化指数 n
194	15.72	235.42	265.76	1 040.88	0.33

为准确模拟管材的自由弯曲变形行为, 基于 ABAQUS/Explicit 和 Standard 模块分别进行 SS304 管材的弯曲变形及回弹行为分析。三维自由弯曲数值模拟的有限元模型如图 5 所示, 包括球面轴承、弯曲模、导向机构、压紧机构及管材等装配组件。其中, 除管材外所有装配组件均设置为刚性体, 管材设置为可变形壳体, 并将管材和其他各装配组件按 S4R、C3D8R 单元类型进行网格划分。实验所用弯曲模具与导向模具的材料为 Cr12MOV, 结合文献 [15-16] 关于 Cr12MOV和不锈钢的摩擦研究, 确定有限元模拟的摩擦系数 f=0.1, 有限元模拟的工艺参数设置如表 2 所示, 其中 D 为直径, t 为壁厚。

为了准确确定 SS304 管材自由弯曲成形过程中的弯曲模临界偏距 U_e, 分别进行偏距 U 为 0.2, 0.3, 0.4, 0.5, 1.0 mm 的自由弯曲数值模拟。图 6 不

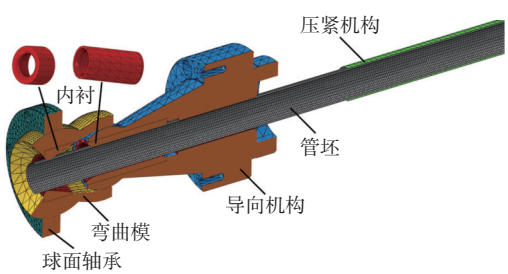


图 5　三维自由弯曲数值模拟有限元模型

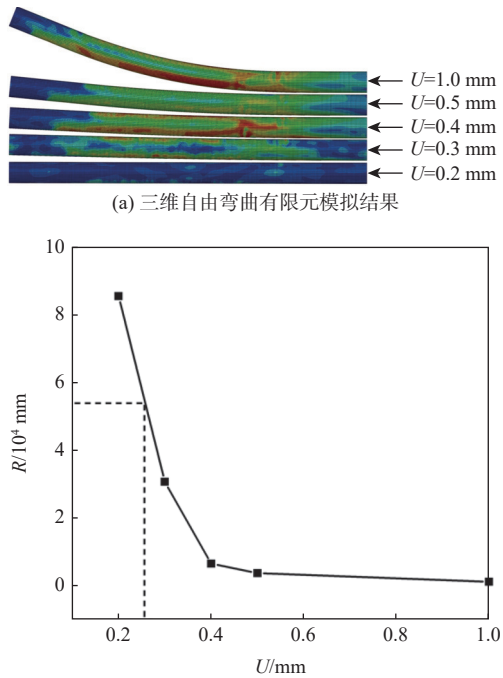
Fig. 5　Finite element model for numerical simulation of 3D free bending

表 2　SS304 管材自由弯曲有限元模拟工艺参数设置

Table 2　Setting parameters of free bending finite element simulation of SS304 tube

管材尺寸(D×t)/(mm×mm)	推进速度 v/(mm·s^{-1})	管模间隙 Δc/mm	摩擦系数 f
20×1	15	0.1	0.1

同弯曲模偏距条件下所获得的自由弯曲有限元模拟结果及相应的 U-R 实验关系。由图 6 可知, 随着偏距 U 的增大, 弯曲半径 R 逐渐减小, 弯曲模偏距 U=0.2 mm 时弯曲半径接近于无限大, 且弯曲模偏距 U=0.2 mm 和 U=0.3 mm 时弯曲半径急剧缩小; 结合 U-R 关系的演化趋势, 可以初步确定弯曲模临界偏距 U_e=0.25 mm, 并且在该临界偏距条件下, 弯管的弯曲半径达到 52 000 mm。



(a) 三维自由弯曲有限元模拟结果

(b) 三维自由弯曲有限元模拟 U-R 关系

图 6　三维自由弯曲有限元模拟结果以及相应的 U-R 关系

Fig. 6　Finite element simulation results and corresponding U-R relationship of 3D free bending

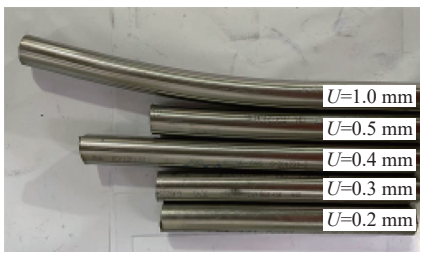
2.3　三维自由弯曲成形实验

为验证所建立数值模拟有限元模型的可靠性, 基于南京航空航天大学研制的三轴自由弯曲成形装备, 以表 2 工艺参数为基础开展 SS304 不锈钢管材在不同弯曲模偏距 U 下的三维自由弯曲成形实验。实验装备如图 7 所示, 球面轴承于 xOy 平面的运动可带动弯曲模绕导向机构作旋转运动, 以实现弯曲模空间上的位置变化, 推进机构在 z 轴上的运动则保证实验管材的匀速进给。通过伺服联动控制系统对 x、y、z 这 3 轴的运动统一调配, 管材能够实现空间上的复杂形状弯曲成形。图 8 为不同弯曲模偏距条件下所获得的自由弯曲成形实验弯管

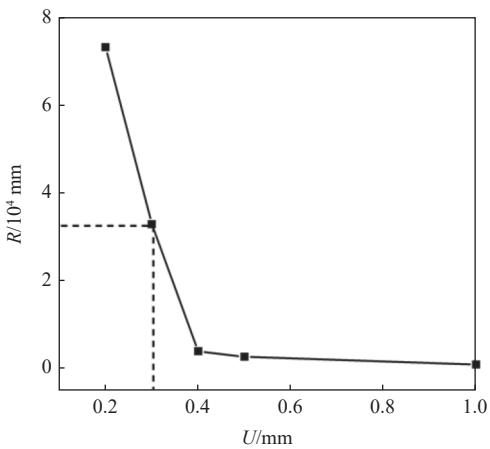


图7 三维自由弯曲成形装备

Fig. 7 3D free bending molding equipment



(a) 实验弯管



(b) 实验 *U-R* 关系

图8 三维自由弯曲成形实验弯管以及相应的 *U-R* 关系

Fig. 8 Experimental bends and corresponding *U-R* relationship of 3D free bending

及相应的 *U-R* 实验关系。可以发现，弯曲半径 *R* 随着弯曲模偏距 *U* 的减小而迅速增大。根据弯曲半径 *R* 与弯曲模偏距 *U* 的演化趋势，可以总结出，自由弯曲成形实验获得的临界偏距 U_e=0.3 mm，自由弯曲实验结果与数值模拟结果比较接近。

3 结果和讨论

3.1 基于临界偏距 U_e 的自由弯曲 *U-R* 关系确定

三维自由弯曲成形过程中，弯管处于少约束状态，此时弯曲模的空间位姿将直接影响管材的弯曲半径和轴线形状。为了准确成形出预设形状的弯管，针对具有不同弯曲半径的弧段需要首先确定其对应的弯曲模偏距 *U* 的数值大小，因此，建立准确的 *U-R* 关系数学模型是进行三维自由弯曲成形工艺分析的基础。本文进行了 *U* 为 2, 4, 6, 8, 10 mm 5种偏距条件下的自由弯曲成形实验，获得了相应的弯管弯曲半径，并将 *U-R* 实验数据总结，如图9所示。由图9可知，随着弯曲模偏距 *U* 的减小，金属管材的弯曲半径 *R* 逐渐增大，说明管材的变形程度逐渐变小。结合 Guo 等[17]对弯曲模偏距 *U* 和弯曲半径 *R* 的研究，采用式(17)来拟合上述实验 *U-R* 关系数据：

$$R=\frac{k}{U^2}+c \tag{17}$$

式中：*k* 为拟合系数；*c* 为常数项。

由于金属管材弯曲半径的增大趋势随着偏距 *U* 的减小越发明显，因此，为了准确描述上述 *U-R* 关系的变化规律，引入弯曲模临界偏距 U_e，即 *U-R* 关系曲线变化的极限偏距值。本文采用式(18)来重新拟合实验 *U-R* 关系数据：

$$R=\frac{k}{(U-U_e)^2}+c \tag{18}$$

式中：U_e 取值 0.3。式(17)和式(18)的拟合结果如图9、表3所示。通过对比可以发现，式(18)拟合结果更加接近实验结果，并且能够合理描述金属管材弯曲半径随着弯曲模偏距的减小快速变大的趋势。

为了验证基于临界偏距 U_e 的自由弯曲成形 *U-R* 关系的准确性，针对一典型连续变曲率构件进行自由弯曲成形实验，典型构件的三维数模如图10(a)所示。基于式(17)和式(18)所确定的 *U-R* 关系数学模型计算不同轴线位置所对应的弯曲模偏距 *U*，并进行相应的自由弯曲成形实验。图10(b)为不同 *U-R* 关系数学模型所对应的实验弯管。可以发

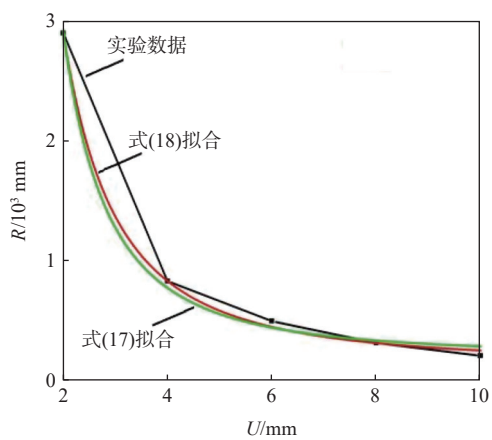


图9 三维自由弯曲成形 *U-R* 关系曲线及拟合结果

Fig. 9 *U-R* curves and fitting results of 3D free bending molding

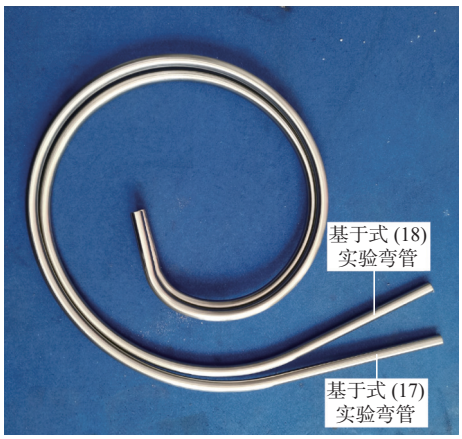
表3 *U-R* 关系拟合结果

Table 3 The fitting results of *U-R* relationship

公式	*k*	*c*	R^2(COD)
式(17)	7 834.610 ± 225.709	202.246 ± 35.888	0.997 52
式(18)	11 044.577 ± 185.952	141.952 ± 885	0.999 15



(a) 几何数模



(b) 基于式 (17) 和式 (18) 的实验弯管

图 10　连续变曲率构件自由弯曲成形实验

Fig. 10　Experiment on free bending forming of planar involute components

现, 当弯管弯曲半径较小时, 基于式(17)和式(18)获得实验弯管形状几乎一致, 并与理论模型比较接近。而随着弯管弯曲半径的逐渐增大, 实验弯管与理论弯管的误差增大, 但由式(18)获得的实验弯管构件轴线形状更加接近于理论数模。上述现象表明, 随着弯曲半径 R 的增加, 式(18)可以更加准确描述 SS304 不锈钢管材弯曲半径 R 随弯曲模偏距 U 的演化趋势, 特别是在大弯曲半径条件下。

3.2　摩擦系数对弯曲模临界偏距的影响

三维自由弯曲过程中, 管材始终与弯曲模和导向模处于动态接触状态, 因此, 模具与管材之间的摩擦行为会影响管材自由弯曲变形, 进而导致弯管的弯曲半径发生不同程度的变化。为了分析摩擦效应对管材自由弯曲过程中弯曲模临界偏距 U_e 的影响, 通过有限元数值模拟方法分析不同摩擦系数(f=0.02、0.05、0.1、0.2)条件下 SS304 不锈钢管材的自由弯曲变形行为。图 11 为 4 种摩擦系数条件下 SS304 不锈钢弯管的弯曲模临界偏距 U_e 的变化情况。由图 11 可知, 随着摩擦系数的增加, 临界偏距 U_e 呈现出下降的趋势, 即在相同的弯曲模偏距条件下, 摩擦系数的增加会导致弯管的弯曲半径减小。

为了解释上述现象, 图 12 为弯曲模偏距 U=4, 摩擦系数分别为 0.02 和 0.2 情况下弯管塑性变形过

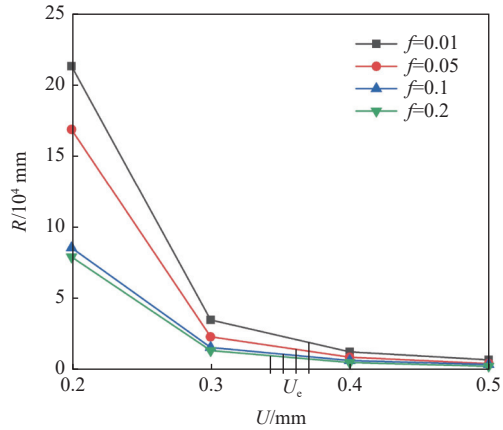
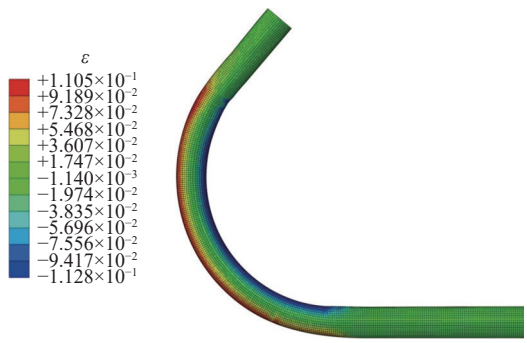
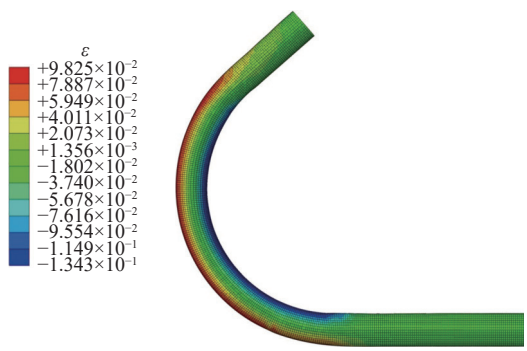


图 11　不同摩擦系数条件下临界偏距 U_e 演化情况

Fig. 11　Evolution of critical eccentricity under different friction coefficients



(a) f=0.02



(b) f=0.2

图 12　自由弯曲成形弯管应变分布情况

Fig. 12　Strain distribution of 3D free-bend-forming bends

程中的应变分布情况。由图 12 可知, 管材的应变绝对值由应变中性层向两侧逐渐增加, 反映出管材内外侧的应力由中性层向两侧逐渐增加的趋势, 符合式 (14) 和式 (15) 的应力分布规律, 进一步证明有限元模型的可靠性。在弯曲成形过程中, 管材受到轴向推力的附加作用, 弯管内侧的等效应变要大于弯管外侧。在弯曲模偏距条件下, 随着摩擦系数的增加, 管材进给所受摩擦阻力增大, 轴向推力进一步增加, 管材弯曲力矩随之增加, 进而导致弯管整体的塑性变形程度增加, 相应的弯管弯曲半径减小, 弯曲模临界偏距 U_e 减小。

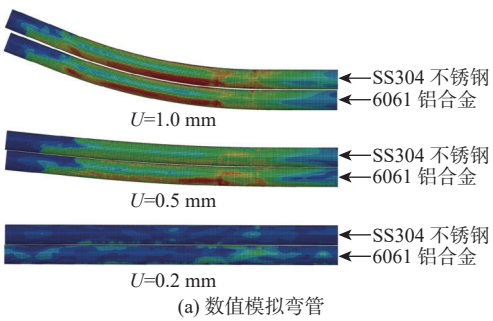
3.3 材料参数对弯曲模临界偏距的影响

三维自由弯曲过程中，管材处于少约束状态，此时材料参数的微小变化会对管材的塑性变形行为和回弹演化规律产生很大影响，进而造成弯管成形轴线尺寸精度的降低。为了分析金属管材的材料力学性能参数对弯曲模临界偏距 U_e 的影响，通过有限元数值模拟方法分析6061铝合金管材自由弯曲过程中的塑性变形行为。有限元模拟的工艺参数设置如表4所示。铝合金管材的材料参数中屈服强度为57.7 MPa，弹性模量为66.9 GPa[18]。图13为6061铝合金管和SS304不锈钢管的自由弯曲数值模拟结果对比，可看出铝合金管的临界偏距 U_e=0.23 mm，且铝合金管的弯曲半径明显大于不锈钢管。这是因为铝合金的屈服强度和弹性模量比不锈钢小，由此导致铝合金弯管自由弯曲过程中回弹行为要大于不锈钢弯管，弯曲半径更大，进而使铝合金弯管数值模拟获得的临界偏距 U_e 小于不锈钢弯管。

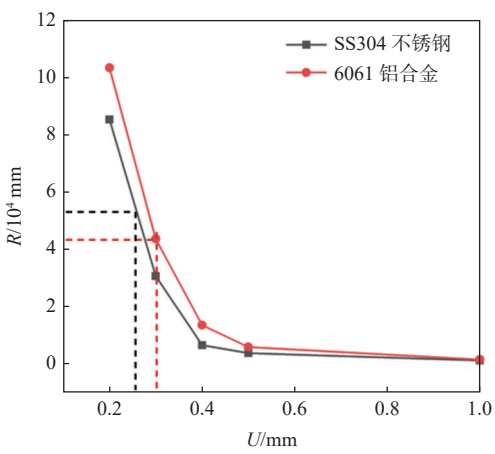
表4 6061管材自由弯曲有限元模拟工艺参数设置

Table 4 Setting parameters of free bending finite element simulation of 6061 tube

管材尺寸($D\times t$)/(mm×mm)	推进速度 v/(mm·s^{-1})	管模间隙 Δc/mm	摩擦系数 f
20×1	15	0.1	0.1



(a) 数值模拟弯管



(b) 临界偏距 U_e

图13 不同材料参数条件下的临界偏距模拟情况对比

Fig. 13 Simulation results of critical eccentricity under different material parameters

4 结 论

1）随着弯管弯曲半径的增大，相较于原始 U-R 关系拟合公式，引入临界偏距 U_e 后 U-R 关系拟合结果更加接近实验数据，准确表达出弯管弯曲半径 R 随弯曲模偏距 U 的演化趋势。

2）随着弯曲模与管材间摩擦系数的增大，材料的塑性流动难度增加，轴向推力进一步增大，导致弯管整体的塑性变形程度增加，弯曲半径减小，临界偏距 U_e 的数值降低。

3）相同弯曲参数条件下，6061铝合金弯管屈服强度和弹性模量低于SS304不锈钢弯管导致6061铝合金弯管的回弹行为大于SS304不锈钢弯管，进一步造成弯曲半径的增大，临界偏距 U_e 的降低。

参考文献（References）

[1] 陶杰, 刘红兵. 金属管件冷成形技术研究进展[J]. 机械制造与自动化, 2009, 38(4): 1-5.
TAO J, LIU H B. Reviewing of cold forming technology for T-shape tube and elbow[J]. Machine Building & Automation, 2009, 38(4): 1-5(in Chinese).

[2] GUO X Z, XIONG H, LI H, et al. Forming characteristics of tube-bending with small bending radii based on a new spherical connection[J]. International Journal of Machine Tools and Manufacture, 2018, 133: 72-84.

[3] MURATA M, KUBOKI T. CNC tube forming method for manufacturing flexibly and 3-Dimensionally bent tubes[J]. Springer Berlin Heidelberg, 2015, 4(28): 363-368.

[4] 程旋. 矩形截面空心构件三维自由弯曲成形机理及缺陷调控研究[D]. 南京: 南京航空航天大学, 2020.
CHENG X. Forming mechanism and defect control of hollow components with rectangular section manufactured by 3D free bending[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2020(in Chinese).

[5] 郝用兴, 张少华, 刘亚辉. 基于数值模拟的管材三维自由弯曲成形规律研究[J]. 制造自动化, 2021, 43(11): 101-104.
HAO Y X, ZHANG S H, LIU Y H. Research on three dimensional free bending forming law of tube based on numerical simulation[J]. Manufacturing Automation, 2021, 43(11): 101-104(in Chinese).

[6] 李秋, 王华军, 田梦芸, 等. 管材三维无芯弯曲过程有限元模拟[J]. 模具工业, 2013, 39(9): 33-36.
LI Q, WANG H J, TIAN M Y, et al. Finite element simulation of 3D bending process of tube[J]. Die & Mould Industry, 2013, 39(9): 33-36(in Chinese).

[7] 鄂大辛, 宁汝新, 古涛. 管材弯曲过程中的弹塑性变形分析[J]. 兵工学报, 2009, 30(10): 1353-1356.
E D X, NING R X, GU T. Analysis on the elastic-plastic deformation during tube-bending process[J]. Acta Armanmentarii, 2009, 30(10): 1353-1356(in Chinese).

[8] SUN C, WANG Z L, ZHANG S Y, et al. Toward axial accuracy prediction and optimization of metal tube bending forming: A nov-

el GRU-integrated Pb-NSGA-Ⅲ optimization framework[J]. Engineering Applications of Artificial Intelligence, 2022, 114: 105193.

[9] YANG H, LI H, MA J, et al. Breaking bending limit of difficult-to-form titanium tubes by differential heating-based reconstruction of neutral layer shifting[J]. International Journal of Machine Tools and Manufacture, 2021, 166: 103742.

[10] 戴莉, 方军, 程璐, 等. 材料参数对高强不锈钢管数控绕弯成形失稳起皱的影响[J]. 精密成形工程, 2017, 9(1): 91-95.
DAI L, FANG J, CHENG L, et al. Effects of material parameters on wrinkling of high-strength stainless steel tube in numerical control potary draw bending[J]. Journal of Netshape Forming Engineering, 2017, 9(1): 91-95(in Chinese).

[11] 鄂大辛, 周大军. 金属管材弯曲理论及成形缺陷分析[M]. 北京: 北京理工大学出版社, 2016.
E D X, ZHOU D J. Metal tube bending: theory and forming defects analysis[M]. Beijing: Beijing Institute of Technology Press, 2016(in Chinese).

[12] 白春玉, 齐丕骞, 牟让科, 等. 基于经典 Von Mises 应力的多轴等效应力修正方法研究[J]. 振动与冲击, 2015, 34(23): 166-170.
BAI C Y, QI P Q, MU R K, et al. Multiaxial equivalent stress amendment algorithm based on classical von mises stress[J]. Journal of Vibration and Shock, 2015, 34(23): 166-170(in Chinese).

[13] CHEN J, SHU W, LI J. Constitutive model of Q345 steel at different intermediate strain rates[J]. International Journal of Steel Structures, 2017, 17(1): 127-137.

[14] GUO X Z, XIONG H. Numerical simulation and experimental study on mechanism and characteristics of tube free-bending forming process[J]. Procedia Manufacturing, 2018, 15: 836-843.

[15] WANG J, ZHOU J, ZHU S S, et al. Friction properties of groove texture on $Cr_{12}MoV$ surface[J]. Journal of Central South University, 2017, 24(2): 303-310.

[16] PALANIKUMAR P, GNANASEKARAN N, SUBRAHMANYA K, et al. Effect of sliding speed and rise in temperature at the contact interface on coefficient of friction during full sliding of SS304[J]. Materials Today:Proceedings, 2020, 27: 1996-1999.

[17] GUO X Z, MA Y N, CHEN W L, et al. Simulation and experimental research of the free bending process of a spatial tube[J]. Journal of Materials Processing Technology, 2018, 255: 137-149.

[18] 马福业. 6061 铝合金管材成形性能评价及参数优化研究[D]. 南京: 南京航空航天大学, 2017.
MA F Y. Formability evaluation and parameters optimization of 6061 aluminum alloy tubes[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2017(in Chinese).

Effect of critical eccentricity on forming accuracy of tubes in 3D free bending process

FAN Xin[1, 2], CHENG Zonghui[1, 2, 3], LI Shaoxin[3], SHU Song[1, 2], CHENG Cheng[2, 3, *], GUO Xunzhong[3]

(1. Anhui Province Aviation Equipment Testing and Control and Reverse Engineering Laboratory, Wuhu 241007, China;
2. State Owned Wuhu Machinery Factory, Wuhu 241007, China;
3. College of Materials Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

Abstract: The tube is less limited during 3D free bending, and the eccentricity of the bending die determines the radius of the bend and the form of the bending axis. In order to accurately process the shape of the tube and improve the forming accuracy during the free bending process, a new free bending mechanical model was established for critical eccentricity. The influence of different friction coefficients and material parameters on the change of critical eccentricity were studied by using the finite element simulation and free bending tests. The findings have demonstrated that the experimental findings and the fitting outcomes of the *U*-*R* relationship with critical eccentricity are in a good agreement. With the increase of friction coefficient, the degree of plastic deformation of the bending tube increases, and the value of critical eccentricity and the bending radius are decreasing. Compared with SS304 tubes, the bending radius of the 6061 Al tube increases and the critical eccentricity decreases at the same process parameters.

Keywords: tube; 3D free bending; elastic-plastic deformation; friction coefficient; material parameters

Received: 2022-04-06; **Accepted**: 2022-07-22; **Published Online**: 2022-10-18 11:14

URL: link.cnki.net/urlid/11.2625.V.20221017.1624.008

Foundation items: National Natural Science Foundation of China (52105360,52175328,U1937206); Opening Project of State Key Lab of Digital Manufacturing Equipment and Technology (DMETKF2021004); Open Fund of Jiangsu Industrial Perception and Intelligent Manufacturing Equipment Engineering Research Center (2021-04); Jiangsu Shuangchuang Talent Program (JSSCBS20210157)

* **Corresponding author**. E-mail: c_cheng@nuaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0246

广义多边形凸包弹性线递支模拟算法

崔钰萍，李子涵，郑国磊*
(北京航空航天大学 机械工程及自动化学院，北京 100191)

摘 要：针对诸多领域涉及的平面弯曲图形的凸包计算，提出弹性线递支模拟算法用于计算简单闭广义多边形的弹性包络线。所提算法基于物理模型，通过判断各支点是否受力平衡来判别其是否为弹性包络线上的平衡支点，并据此分别进行前进、回弹和跳跃等操作，直至计算出所有平衡支点进而求出其弹性包络线。3种典型的简单闭广义多边形的对比测算表明：所提算法可实时稳健地求解平面任意简单闭广义多边形的弹性包络线，具有高效性和普遍适用性。

关 键 词：广义多边形；曲边形；凸包；弹性包络线；支撑线
中图分类号：TP301.6
文献标志码：A **文章编号**：1001-5965(2024)01-0216-08

凸包计算是计算几何的基本问题之一，在现代工业和信息技术诸多领域(如计算机辅助设计、计算机图形学、运筹学等)中均有涉及，碰撞检测、模式识别、形状分析和图像处理等也皆可归结为此类问题。点集的凸包(包含点集的最小凸集)求解算法大致可分为图遍历法和增量法，经典的Graham扫描算法[1]和Jarvis步进算法[2]皆将输入点作为图中临时边的顶点并检查其凹凸性来更新凸包，此为图遍历法；Preparata和Hong[3]分治算法和快速凸包算法[4]皆由初始凸包开始，通过检查剩余点是否属于当前凸包来更新凸包，此为增量法。简单多边形由于自身的有序性，其凸包求解算法更为精巧。Sklansky[5]率先提出简单多边形凸包的线性时间算法，尽管该算法在部分情况下会因自交导致无效，但其形式简洁且易于实现。Mccallum和Avis[6]提出逻辑缜密的线性效率算法，几年后Lee[7]提出空间复杂度更低的算法，Melkman[8]也提出时间复杂度为$O(n)$的算法，但因实现较为复杂致使应用受到局限。Chen[9]的陈氏算法基于Sklansky算法并通过增加Reject操作避免其自交，吴中海等[10]改进了陈氏算法中Reject条件判断存在的失误及初始点选取不完备的问题。孔宪庶和蔡洪学[11]运用Graham扫描技术和Lee[7]的双动线检测算法提出时间和空间复杂度均为线性的算法；但王丽青等[12]通过反例指出其删除操作的局限性并采用折半查找相关边来加以修正；张林[13]引入分治思想并删除内边，提高了王丽青等[12]改进后算法的时间效率。近期中国其他专家学者立足生产实际，亦对凸包算法进行了大量有益的探索[14-16]。

近年来凸包在前沿科技中的应用层出不穷，例如将凸包应用于计算机视觉[17-18]和路径规划[19-20]等。然而目前的应用大多基于有限点集的凸包算法，对于实际应用中存在的弯曲物体诸如路径规划中的圆形障碍物等类的凸包问题研究较少。国外针对平面特定的弯曲图形(如二次圆弧[21]、三次样条曲线[22])的凸包求解提出一些理论算法，但这些算法大多基于复杂的曲线计算(如求曲线公切线)，导致算法代价高、稳健性差，且不具备实时性能[23]。为统一求解直边和曲边图形的凸包，本文通过将数学模型转化为物理模型，提出弹性线递支模拟算法，用于求解简单闭广义多边形的弹性包络线，该算法运算简单、鲁棒性好且具有实时性，对于推动

收稿日期：2022-04-15；录用日期：2022-06-05；网络出版时间：2022-06-13 16:41
网络出版地址：link.cnki.net/urlid/11.2625.V.20220613.1350.002
*通信作者. E-mail：zhengguolei@buaa.edu.cn

引用格式：崔钰萍，李子涵，郑国磊. 广义多边形凸包弹性线递支模拟算法[J]. 北京航空航天大学学报，2024，50(1)：216-223.
CUI Y P，LI Z H，ZHENG G L. Computing convex hull of a generic polygon with simulation of progressive support for an elastic line[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(1)：216-223 (in Chinese).

相关领域的技术进步具有积极意义。

1　广义多边形及其弹性包络线

1.1　定　义

本文将包含圆弧或自由曲线段等曲边的多边形称为广义多边形。相关定义如下。

定义 1　设$C=(c_0,c_1,\cdots,c_l)$为欧氏空间R^2中的一组无限直线和/或自由曲线(简称自然线), $V=(v_0,v_1,\cdots,v_n)$为位于C上的一个点序列, 且满足v_i和$v_{i+1}(i=0,1,\cdots,n-1)$为某条自然线$c_j(j=0,1,\cdots,l)$上一条线段$e_i$的 2 个端点, 记为$e_i(v_i,v_{i+1},c_j)$, 称$e_i$为定义在$c_j$上的一条边, 则将边序列$E=(e_i|i=0,1,\cdots,n-1)$称为一个广义多边形$P$, 记为$P=(V,E)$。$e_i(i=0,1,\cdots,n-1)$的连接顺序代表$E$和$P$的方向。若$v_0=v_n$, 则称$P$为闭广义多边形; 否则, 称$P$为开广义多边形。对于 2 个广义多边形$P_1$和$P_2$, 若二者间所有边均重合, 则称$P_1$与$P_2$相等, 记为$P_1=P_2$。

定义 2　若一个多边形P同时满足如下 2 个条件, 则称其为简单广义多边形: ①除$v_0=v_n$成立外, $v_i\neq v_j(i,j=0,1,\cdots,n,i\neq j)$, 即内部顶点均不重合; ②除顶点$v_i$和$v_{i+1}$外, e_i与e_j $(i,j=0,1,\cdots,n-1,i\neq j)$不存在其他共点。

下文若无特别说明, 多边形$P(V_P,E_P)$均指简单闭广义多边形(既是简单广义多边形又是闭广义多边形), 其顶点V_P和边E_P序列为$V(v_P^0,v_P^1,\cdots,v_P^{n_P})$和$E(e_P^0,e_P^1,\cdots,e_P^{n_P})$, 其中$v_P^0$与$v_P^{n_P}$不重合, 为边$e_P^{n_P}$的终点和始点及$P$的首末顶点。

定义 3　设$P=(V_P,E_P)$为简单闭广义多边形, 其中$V=(v_P^0,v_P^1,\cdots,v_P^{n_P})$, $E=(e_P^0,e_P^1,\cdots,e_P^{n_P})$, 沿$P$的方向, 将其所在的欧氏空间$R^2$分割为左右两侧区域, 将其中的左侧区域称为$P$的广义多边形域, 记为$A(P)$。若$A(P)$的面积为有限值, 则称$A(P)$为有限域; 否则, 称$A(P)$为无限域。

定义 4　设$P=(V_P,E_P)$为一闭广义多边形, 其中顶点序$V_P=(v_P^i|i=0,1,\cdots,n_P)$, 边序$E_P=(e_P^i|i=0,1,\cdots,n_P)$, 其周长记为$|P|$, L为一条周长小于$|P|$的有向环形弹性线。假设$A(P)$为一块刚性板, 用$A(P)$支撑$L(A(P)\leqslant A(L))$, 若自然状态下L发生弹性变形后为紧紧围绕着$A(P)$的广义多边形$Q=(V_Q,E_Q)$且Q与P方向一致, 则称Q为P的弹性包络线, 记为$Q=\mathrm{conv}(P)$。显然, Q是P的一个凸包, 弹性包络边示意图如图 1 所示。

通过模拟弹性线的物理变形现象来将多边形凸包定义为其弹性包络线, 不但从物理学角度揭示了凸包的物理本质, 更重要的是为求解广义多边形的凸包提供了一种全新的思路。本文第 2 节拟详细介绍这种思路及其实现算法。

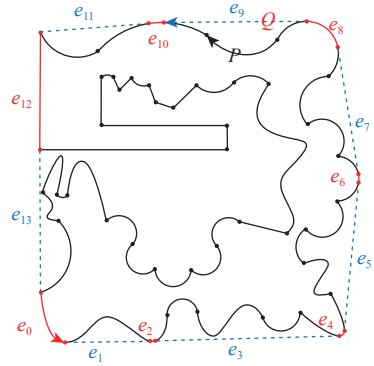


图 1　弹性包络边示意图

Fig. 1　Schematic diagram of elastic envelope edge

1.2　弹性包络线与多边形关系

本文弹性包络线Q的边可分为支撑边e_s和回弹边e_r这 2 种类型: ①支撑边e_s, 重合于P上某条边$e_P^i(e_P^i\in E_P)$, 如图 1 所示, e_s或为直边或为曲边, 取决于e_P^i所属自然线的类型; ②回弹边e_r, 不重合于P上任一条边, 故为直边, 如图 1 虚线段所示。

弹性包络线Q的一个顶点v_Q或重合于P上的一个顶点$v_P^i(v_P^i\in V_P)$, 或不重合于P上任意顶点但位于P的某条边e_P^i内。

综上弹性包络线具有如下性质。

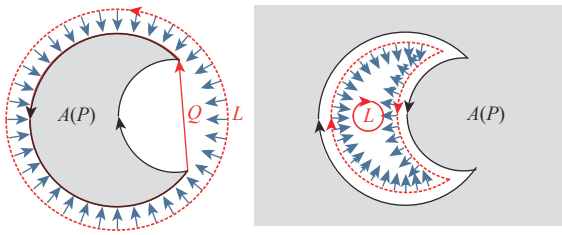
性质 1　若$Q=(V_Q,E_Q)$为$P=(V_P,E_P)$的弹性包络线, 其中$V_P=(v_P^i|i=0,1,\cdots,n_P)$, $E_P=(e_P^i|i=0,1,\cdots,n_P)$, $V_Q=(v_Q^i|i=0,1,\cdots,n_Q)$, $E_Q=(e_Q^i|i=0,1,\cdots,n_Q)$, 则对于$\forall j(j=0,1,\cdots,n_Q)$, $\exists i(i=0,1,\cdots,n_P)$满足: ①$v_Q^j=v_P^i$或$v_Q^j\in e_P^i$; ②若$e_Q^j$不为直线段则$e_Q^j\subset e_P^i$。

定理 1　简单闭广义多边形的弹性包络线也是简单闭广义多边形。

1.3　弹性包络线存在性

定理 2　对于闭广义多边形P, 若$A(P)$为有限域(P方向为逆时针方向), 则P存在弹性包络线; 否则, 亦即$A(P)$为无限域(P方向为顺时针方向), 则P不存在弹性包络线。

证明　由定义 4 可知, $A(P)\leqslant A(L)$, 结合定义 3 知L或在P的右侧区域或与P重合。当P为逆时针方向时, 其左侧区域封闭, 即$A(P)$的面积为有限值, $A(P)$为有限域。此时$A(P)$支撑L的示意图如图 2(a)所示。当L自然收缩包络$A(P)$时将受力形变为紧密围绕$A(P)$的广义多边形Q, Q即为P的弹性包络线。反之, 当P为顺时针方向时, 其左侧区域开放, $A(P)$为无限域。此时$A(P)$支撑L的示意图如图 2(b)所示。当L自然收缩时将因缺乏支撑而处于不贴合$A(P)$的原始松弛状态, 故$A(P)$为无限域时P不存在弹性包络线。　**证毕**



(a) $A(P)$ 为有限域 (b) $A(P)$ 为无限域

图 2 环形弹性线收缩示意图

Fig. 2 Schematic diagram of contraction of a circular elastic line

性质 2 若两多边形相等，则其弹性包络线也相等。

2 本文算法

2.1 本文算法流程

根据前述弹性包络线的定义及其性质，本文算法的核心思想可概括为：模拟弹性线从多边形起点沿边序前进，逐步获取使其受力发生形变的支撑点序，直至弹性线闭合且不松弛，该算法称为弹性线递支模拟算法。

本文算法总体流程如图 3 所示，其中主要步骤为：①平衡判断，根据 3 点位置判断弹性包络线在中间点处是否受力平衡；②前进判断，判断算法是否可继续向前推进，并更新前进后的点；③相离判断，直线段与多边形相离是回弹的前提；④回弹，当 3 点不平衡时向平衡支点回弹；⑤跳跃，若不可回弹则根据规则进行跳跃以简化算法；⑥求解包络线，根据平衡支点集构造弹性包络线的点序和边序。具体步骤如下。

步骤 1 已知 $P(V_P,E_P)$ 且 $A(P)$ 为有限域，其中 $V_P=(v_P^i|i=0,1,\cdots,n_P)$，$E_P=(e_P^i|i=0,1,\cdots,n_P)$。点序 V_S 为空。从 v_P^0 开始，顺次前进获得初始 p_i、p_j 和 p_k 这 3 点，并将 p_i 加入 V_S，$V_S=(v_S^0)$。

步骤 2 判断 p_j 是否为平衡支点：若是，转步骤 3；否则，转步骤 4。

步骤 3 将 p_j 加入 V_S，$V_S=(v_S^0,v_S^1,\cdots,v_S^m)$，执行步骤 5 获得新的点 p_{new}。若 p_{new} 为空转步骤 8；否则令 $p_i=p_j$，$p_j=p_k$，$p_k=p_{\text{new}}$ 后转步骤 2。

步骤 4 计算并判断 $\overline{p_ip_k}$ 与 P 关系：若相离，转步骤 6；否则，计算后转步骤 7。

步骤 5 若对任意 $v_S^i\in V_S(i\neq m)$ 满足 $v_S^m\neq v_S^i$，若 $p_k\notin V_S$，则依据动态步长沿 p_k 所在边前进获得新的点 p_{new}；若 $p_k\in V_S$ 且 $p_k\neq v_S^m$，则依照 V_S 按顺序前进获得 $p_{\text{new}}(p_{\text{new}}\in V_S)$；若 $p_k\in V_S$ 且 $p_k=v_S^m$，则 p_k 不可前进，p_{new} 为空。若存在 $v_S^i\in V_S(i\neq m)$ 满足 $v_S^m=v_S^i$，则 p_k 不可前进，p_{new} 为空，并删除 V_S 中 v_S^i 之前的点。返回 p_{new}。

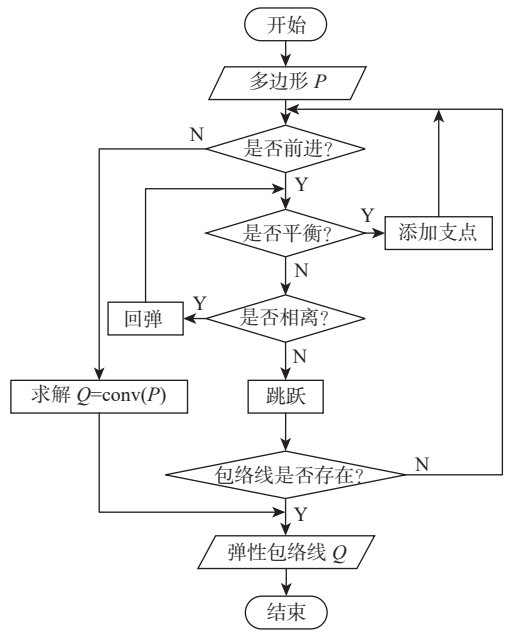


图 3 本文算法流程

Fig. 3 Flow of the proposed algorithm

步骤 6 判断 $p_i=v_S^0$ 是否成立，若 $p_i=v_S^0$ 说明已回弹至起始点，执行步骤 5 获得 p_{new}，若 p_{new} 为空转步骤 8；否则令 $p_j=p_k$，$p_k=p_{\text{new}}$ 后转步骤 2。若 $p_i\neq v_S^0$，则进行回弹，即令 $p_j=p_i$，$p_i=v_S^{m-1}$ 并将 v_S^m 从 V_S 中删除，转步骤 2。

步骤 7 根据 p_i、p_k 求新多边形 P' 的弹性包络线 $\text{conv}(P')$，若 $\text{conv}(P')$ 不存在则按规则跳跃后转步骤 6，若 $\text{conv}(P')$ 存在则 $\text{conv}(P)=\text{conv}(P')$ 并结束运算。

步骤 8 根据点序 $V_S=(v_S^0,v_S^1,\cdots,v_S^m)$ 和多边形 $P(V_P,E_P)$，求弹性包络线的点序 V_Q 和边序 E_Q，构造弹性包络线 Q。

2.2 平衡判断

本文算法的关键步骤即寻找 P 上支撑弹性线使其发生形变的点，由物理模型可知，弹性包络线上的点处处受力平衡，因此，支撑点即为使弹性线受力平衡的点。本文对弹性线受力模型进行分析，提出判断弹性线在多边形上的某点是否受力平衡的简易方法。

定理 3 设 p_i、p_j 和 p_k 为均匀轻质弹性线 l 上的相邻前后 3 点且贴合于广义多边形 P，$\boldsymbol{F}_1$、$\boldsymbol{F}_2$ 为 p_i、p_k 对 p_j 的拉力，$\boldsymbol{F}_N$ 为 p_j 所受的 P 给予的支撑力。若 $\langle\boldsymbol{F}_1,\boldsymbol{F}_2\rangle+\langle\boldsymbol{F}_1,\boldsymbol{F}_N\rangle+\langle\boldsymbol{F}_2,\boldsymbol{F}_N\rangle=2\pi$ 或 $\langle\boldsymbol{F}_1,\boldsymbol{F}_2\rangle=\pi$，则弹性线在 p_j 点受力平衡。

定理 4 设 p_i、p_j 和 p_k 为均匀轻质弹性线 l 上的相邻前后 3 点且贴合于广义多边形 P，l 左侧为 $A(P)$，如图 4 所示。若 $\boldsymbol{F}_1$ 顺时针旋转到 $\boldsymbol{F}_2$ 的角度 $\theta\leqslant\pi$，则弹性线在 p_j 点受力平衡。

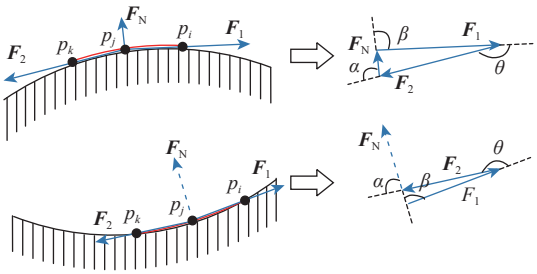


图 4　受力平衡与非平衡向量示意图

Fig. 4　Vector diagram of force equilibrium and force non-equilibrium

证明　若$\boldsymbol{F}_1$顺时针旋转到$\boldsymbol{F}_2$的角度$\theta=\pi$，则$\langle\boldsymbol{F}_1,\boldsymbol{F}_2\rangle=\pi$，满足定理 3。若$\boldsymbol{F}_1$顺时针旋转到$\boldsymbol{F}_2$的角度$\theta<\pi$，设$\overrightarrow{p_ip_j}$顺时针旋转到$\boldsymbol{F}_N$的角度为$\theta_1$，$\overrightarrow{p_jp_k}$顺时针旋转到$\boldsymbol{F}_N$的角度为$\theta_2$，如图 5 所示。若$\boldsymbol{F}_N$存在，则其方向与$\boldsymbol{F}_1$、$\boldsymbol{F}_2$的角平分线平行并指向域外，又因$l$左侧为$A(P)$，因而$\theta_1<\pi$且$\theta_2<\pi$。因为$\boldsymbol{F}_1$与$\overrightarrow{p_ip_j}$方向相反，$\boldsymbol{F}_2$与$\overrightarrow{p_jp_k}$方向相同，所以$\boldsymbol{F}_1$顺时针旋转到$\boldsymbol{F}_N$的角度为$\pi+\theta_1(\pi<\pi+\theta_1<2\pi)$；$\boldsymbol{F}_2$顺时针旋转到$\boldsymbol{F}_N$的角度为$\theta_2(\theta_2<\pi)$。故$\theta=(\pi+\theta_1)-\theta_2$，$\langle\boldsymbol{F}_1,\boldsymbol{F}_N\rangle=\pi-\theta_1$，$\langle\boldsymbol{F}_2,\boldsymbol{F}_N\rangle=\theta_2$。因$\theta<\pi$，则$\langle\boldsymbol{F}_1,\boldsymbol{F}_2\rangle=(\pi+\theta_1)-\theta_2$，故：

$$\langle\boldsymbol{F}_1,\boldsymbol{F}_2\rangle+\langle\boldsymbol{F}_1,\boldsymbol{F}_N\rangle+\langle\boldsymbol{F}_2,\boldsymbol{F}_N\rangle=[(\pi+\theta_1)-\theta_2]+(\pi-\theta_1)+\theta_2=2\pi$$

满足定理 3，弹性线在p_j点受力平衡。　**证毕**

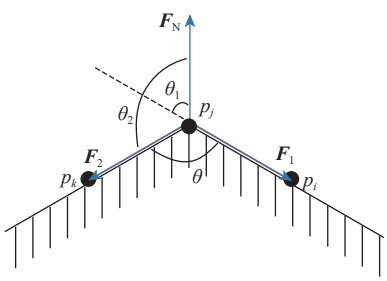


图 5　平衡支点角度示意图

Fig. 5　Schematic diagram of angle of balance pivot point

定义 5　设p_i、p_j和p_k为均匀轻质弹性线l上的相邻前后 3 点且贴合于广义多边形P，若弹性线在p_j受力平衡，则称p_j在P上的贴合点为平衡支点，记为$p_j=B_{i,k}$；否则称为非平衡支点，记为$p_j\neq B_{i,k}$。

2.3　回弹原理

当弹性线贴合多边形但受力不平衡时，弹性线会在其内部应力的作用下回弹为不再处处贴合多边形的直线段，最终组成弹性包络线的回弹段。当$p_j\neq B_{i,k}$时，回弹运算为：舍弃当前非平衡支点p_j，使前序平衡支点p_i为新的p_j点。因p_i、p_j和p_k为均匀轻质弹性线l上的相邻前后 3 点，故回弹前需通过位置关系计算保证p_i与p_k间无其他支点，即$\overline{p_ip_k}$与P相离(不相交)。满足$\overline{p_ip_k}$与P相离直接进行回弹运算；否则不可进行回弹，执行跳跃计算。

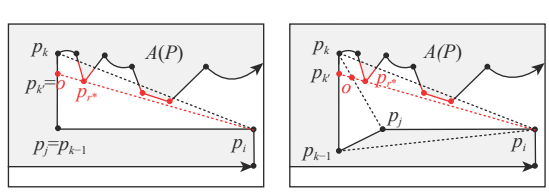
2.3.1　位置关系计算

定义 6　设$\overline{p_ip_k}$为一直线段，P为一个广义多边形。若$\overline{p_ip_k}$与P存在除p_i和p_k以外的共点(即重合点)，且该共点在P上的邻域所具有的 2 个边界点分别位于$\overline{p_ip_k}$的两侧，则$\overline{p_ip_k}$与P的关系为交叉，记为$\overline{p_ip_k}\cap P=\otimes$，共点为叉点。若$\overline{p_ip_k}$与$P$不交叉且存在除$p_i$和$p_k$以外的共点，则$\overline{p_ip_k}$与$P$的关系为接触，记为$\overline{p_ip_k}\cap P=\odot$，共点为触点。若$\overline{p_ip_k}$与$P$不存在共点，则$\overline{p_ip_k}$与$P$的关系为相离，记为$\overline{p_ip_k}\cap P=\varnothing$。

位置关系计算首先根据定义 6 判断$\overline{p_ip_k}$与P的关系，若$\overline{p_ip_k}\cap P=\otimes$或$\overline{p_ip_k}\cap P=\odot$，说明$p_i$与$p_k$间存在新的支撑段。下一步需计算新支撑段的支撑点：当$\overline{p_ip_k}\cap P=\odot$时，触点即为支撑点；当$\overline{p_ip_k}\cap P=\otimes$时，则需将$p_k$点往其前序点退后直至$\overline{p_ip_k}\cap P=\odot$，此时的触点为支撑点。

定理 5　设p_i、p_j、p_k和p_{k-1}为多边形P上的 4 个点，p_{k-1}为p_k的前序点。已知当$p_{k-1}=p_jp_{k-1}=p_j$时，$p_j\neq B_{i,k}$；当$p_{k-1}\neq p_j$时，$p_j\neq B_{i,k}$，$p_j=B_{i,k-1}$，$p_{k-1}\neq B_{j,k}$。若$\overline{p_ip_k}\cap P=\otimes$，则$p_{k-1}$到$p_k$所经边$e_{i_k}$上必有一点$p_{k'}$使得$\overline{p_ip_{k'}}\cap P=\odot$。

证明　若$p_{k-1}=p_j$，易知$\overline{p_ip_j}$&$\overline{p_jp_k}\cap P=\varnothing$。设$E_\Delta$为$P$在$\Delta p_ip_jp_k$内的边集合，因$\overline{p_ip_k}\cap P=\otimes$，故可设$p_{r*}=\arg\min\limits_{p_r\in E_\Delta}\dfrac{\overrightarrow{p_ip_k}\cdot\overrightarrow{p_ip_r}}{\|\overrightarrow{p_ip_k}\|\cdot\|\overrightarrow{p_ip_r}\|}$。延长$p_ip_{r*}$至与$\overline{p_jp_k}$相交，记交点为$o$，如图 6(a)所示。则$\overline{p_{r*}o}$与$P$无叉点，因为若存在叉点，则该叉点在$P$上邻域的边界点之一$p_{r'}$满足$\dfrac{\overrightarrow{p_ip_k}\cdot\overrightarrow{p_ip_{r'}}}{\|\overrightarrow{p_ip_k}\|\cdot\|\overrightarrow{p_ip_{r'}}\|}<\dfrac{\overrightarrow{p_ip_k}\cdot\overrightarrow{p_ip_{r*}}}{\|\overrightarrow{p_ip_k}\|\cdot\|\overrightarrow{p_ip_{r*}}\|}$，与$p_{r*}$定义矛盾。同理$\overline{p_ip_{r*}}$与$P$无叉点。因此，$\overline{p_io}\cap P=\odot$，又因$o\in\overline{p_{k-1}p_k}$，故$p_{k'}=o$，定理 5 得证。若$p_{k-1}\neq p_j$，易知$\overline{p_jp_{k-1}}$、$\overline{p_jp_k}$&$\overline{p_{k-1}p_k}\cap P=\varnothing$。同上可知$\overline{p_io}\cap P=\odot$，$o\in\overline{p_jp_k}$。因$p_j\neq B_{i,k}$，$p_j=B_{i,k-1}$，$p_{k-1}\neq B_{j,k}$，故$p_j$点为$\Delta p_ip_{k-1}p_k$内点，则延长$p_io$必与$\overline{p_{k-1}p_k}$相交，记交点为$p_{k'}$，如图 6(b)所示。因$\overline{p_jp_{k-1}}$、$\overline{p_jp_k}$&$\overline{p_{k-1}p_k}\cap P=\varnothing$，则$\overline{op_{k'}}\cap P=\varnothing$。故$\overline{p_ip_{k'}}\cap P=\odot$，定理 5 得证。　**证毕**



(a) $p_{k-1}=p_j$　　(b) $p_{k-1}\neq p_j$

图 6　触点求解示意图

Fig. 6　Schematic diagram of contact point solution

综上，位置关系计算的主要步骤如下。

步骤 1 已知多边形P和其上p_i、p_k这2点，根据定义6判断是否$\overline{p_ip_k}\cap P=\otimes$，若$\overline{p_ip_k}\cap P=\otimes$，转步骤2；否则转步骤3。

步骤 2 求p_k前序点p_{k-1}，依据定理5采用二分法将p_k沿边向p_{k-1}退后直至$\overline{p_ip_k}\cap P=\odot$，转步骤3。

步骤 3 若$\overline{p_ip_k}\cap P=\odot$，求出此时$\overline{p_ip_k}$与$P$的触点，转步骤4；否则，即$\overline{p_ip_k}\cap P=\varnothing$，转步骤5。

步骤 4 进行跳跃计算。

步骤 5 进行回弹运算。

2.3.2 跳跃计算

当$p_j\neq B_{i,k}$且$\overline{p_ip_k}\cap P=\odot$时，$\overline{p_ip_k}$与多边形$P$的触点$(v_{t0},v_{t1},\cdots,v_{tm})$为$p_i$与$p_k$间的新支撑点。因算法结束条件为支撑点重复，故前进轨迹中存在由线段$\overline{v_{tm}p_k}$和从p_k到v_{tm}的边组成的环路。弹性包络线为简单闭广义多边形，因此，该环路要么无弹性包络线，算法可略过环路的计算进行跳跃；要么其弹性包络线即为多边形的弹性包络线。故可通过跳跃计算避开环路以减少计算量。根据触点的位置，将跳跃分为前跳跃、后跳跃和前后跳跃，具体分类方法如规则1所示。

规则 1 已知点p_k及$\overline{p_ip_k}$与$P=(V,E)$的所有触点$(v_{t0},v_{t1},\cdots,v_{tm})$，其中$V=(v_0,v_1,\cdots v_n)$，$E=(e_0,e_1,\cdots,e_n)$。若$(v_{t0},v_{t1},\cdots,v_{tm})$皆为$p_k$前序点，则进行前跳跃；若$(v_{t0},v_{t1},\cdots,v_{tm})$皆为$p_k$后序点，则进行后跳跃；否则进行前后跳跃。

若触点皆为p_k前序点，设p_k点位于P的边为$e_{i_k}(0\leqslant i_k\leqslant n)$上，则$P$存在$p_k$、$v_{i_k+1}$、$\cdots$、$v_n$、$v_0$、$v_1$、$\cdots$、$v_{t0}$、$v_{t1}$、$\cdots$、$v_{tm}$形成的环，如图7所示。在该环中，$v_{t0}$、$v_{t1}$、$\cdots$、$v_{tm}$位于同一直线上，因此，该环可简化为由$p_k$、$v_{i_k+1}$、$\cdots$、$v_n$、$v_0$、$v_1$、$\cdots$、$v_{t0}$形成的环。

前跳跃的主要步骤如下。

步骤 1 求点p_k和首末触点v_{t0}、v_{tm}在P上的所在边e_{i_k}和e_{i_0}、$e_{i_m}(0\leqslant i_0\leqslant i_m\leqslant n)$。

步骤 2 求$V'=(p_k,v_{i_k+1},\cdots,v_n,v_0,\cdots,v_{i_0},v_{t0})$，$E'=(e_{i'_k},e_{i_k+1},\cdots,e_n,e_0,\cdots,e_{i'_0},e_{tk})$的多边形$P'$的弹性包络线$\mathrm{conv}(P')$，其中$e_{i'_k}\subset e_{i_k}$，$e_{i'_0}\subset e_{i_0}$，$e_{tk}=\overline{v_{t0}p_k}$。若$\mathrm{conv}(P')$不存在，转步骤3；否则$P$的弹性包络线$\mathrm{conv}(P)=\mathrm{conv}(P')$，输出$\mathrm{conv}(P)$后结束计算。

步骤 3 向前进行跳跃，即直接略去e_{i_k}后的所有边，使$p_k=v_{tm}$后进行回弹运算。

与前跳跃类似，后跳跃的主要步骤如下。

步骤 1 求点p_k和首末触点v_{t0}、v_{tm}在P上的所在边e_{i_k}和e_{i_0}、$e_{i_m}(0\leqslant i_0\leqslant i_m\leqslant n)$。

步骤 2 求$P'=(V',E')$的弹性包络线$\mathrm{conv}(P')$，

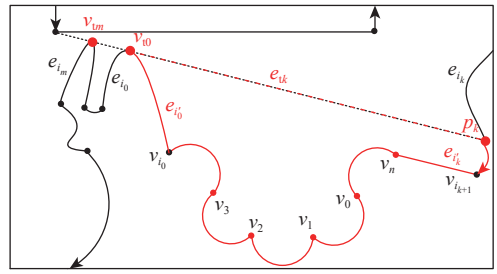


图7 前跳跃示意图

Fig. 7 Forward jump diagram

$V'=(p_k,v_{i_k+1},\cdots,v_{i_0},v_{t0})$，$E'=(e_{i'_k},e_{i_k+1},\cdots,e_{i'_0},e_{tk})$，其中$e_{i'_k}\subset e_{i_k}$，$e_{i'_0}\subset e_{i_0}$，$e_{tk}=\overline{v_{t0}p_k}$。若$\mathrm{conv}(P')$不存在，转步骤3；否则$P$的弹性包络线$\mathrm{conv}(P)=\mathrm{conv}(P')$，输出$\mathrm{conv}(P)$后结束计算。

步骤 3 向后进行跳跃，即直接略去e_{i_m}前的边，使$p_k=v_{tm}$后进行回弹运算。

前后跳跃的主要步骤如下。

步骤 1 求点p_k在P上的所在边e_{i_k}和$(v_{t0},v_{t1},\cdots,v_{tm})$中最靠近$p_k$的触点$v_{ts}$、$v_{te}$及其所在边$e_{i_s}$、$e_{i_e}$。

步骤 2 求$P'=(V',E')$，$V'=(p_k,v_{i_k+1},\cdots,v_{i_e},v_{te})$，$E'=(e_{i'_k},e_{i_k+1},\cdots,e_{i'_e},e_{tk})$的弹性包络线$\mathrm{conv}(P')$，其中$e_{i'_k}\subset e_{i_k}$，$e_{i'_e}\subset e_{i_e}$，$e_{tk}=\overline{v_{te}p_k}$。若$\mathrm{conv}(P')$不存在，转步骤3；否则$P$的弹性包络线$\mathrm{conv}(P)=\mathrm{conv}(P')$，输出$\mathrm{conv}(P)$后结束计算。

步骤 3 向前进行跳跃，即直接略去e_{i_k}后的所有边，使$p_k=v_{ts}$后进行回弹运算。

3 实验及测试

本文算法的关键步骤皆经过严格的数学证明，理论上可适用于任意简单闭广义多边形。时间效率上，本文算法的平衡判断、前进判断、回弹和求解包络线步骤均只需要线性时间，仅不平衡点的位置关系计算时间复杂度稍高。在所有的线段与多边形均相离等苛刻条件下，不执行跳跃步骤，其消耗时间为$O(mn)$，其中n为计算点数量，m为n中的非平衡支点。而对于平衡点较多或非平衡点可进行跳跃的多边形，本文算法的时间复杂度远小于$O(n^2)$。

由于目前对自由曲线的凸包求解算法多数基于复杂的曲线计算，与本文算法思想不同，故无法直接比较。因此，为验证本文算法的有效性及评估算法的执行效率，在配置为Intel Core i5-8250U 1.6 GHz处理器和8 GB内存的计算机上用C++实现本文算法，求解如图8所示的3个典型多边形。



(a) 全凸多边形　(b) 内折多边形　(c) 外凹多边形

图 8　3 种典型多边形的弹性包络线

Fig. 8　Elastic envelopes of three typical polygons

图 8(a) 为全凸广义多边形, 其本身即为弹性包络线; 图 8(b) 内凹区域复杂, 在内部进行多次折拐; 图 8(c) 的凹区域较为简单, 仅在外边界进行凹凸交替。本文算法前进步长设为动态步长, 当取不同步长使得n为不同数值时, 各算例多次运行时间均值如表 1 所示。

表 1　广义多边形弹性包络线计算时间

Table 1　Calculation time of elastic envelope of generic polygon

点数量	全凸多边形 计算时间/ms	内折多边形 计算时间/ms	外凹多边形 计算时间/ms
30	0.078 5	0.126 8	0.169 0
60	0.125 4	0.259 2	0.360 6
90	0.176 5	0.413 5	0.756 0
120	0.228 0	0.629 8	1.212 0
150	0.283 4	0.926 0	1.805 0
180	0.338 2	1.194 2	2.604 0
240	0.454 0	1.838 8	4.076 0

采用幂函数进行曲线拟合的结果如图 9 所示。由图可知, 不同形状的广义多边形求解效率差别较大, 平衡点占据比例大的多边形(见图 8(a))运行效率接近线性, 内凹区域复杂的多边形(见图 8(b))由于可进行跳跃其运行效率高于仅简单内凹不可进行跳跃的多边形(见图 8(c)), 实测 3 个算例的运

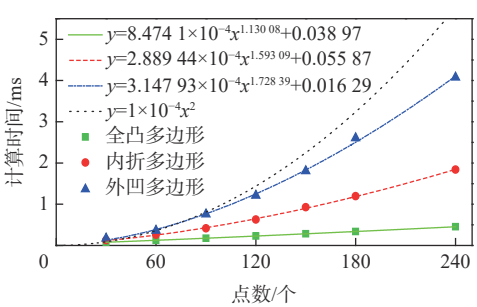
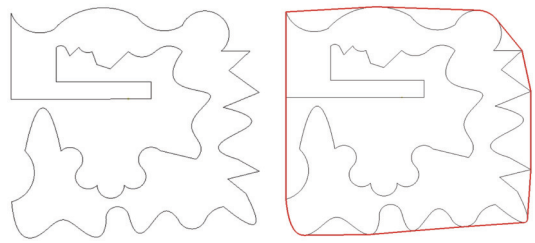


图 9　计算时间拟合结果

Fig. 9　Fitting result of calculating time

算时间复杂度均小于 $O(n^2)$。

对于更加复杂的广义多边形的弹性包络线求解, 本文算法也能稳健解决。例如, 图 10(a) 所示的简单闭广义多边形, 当取不同步长使得n分别为 248、1 168 和 2 318 时, 求解结果如图 10(b) 所示, 运行时间均值分别为 0.004 9 s、0.099 5 s 和 0.350 9 s。



(a) 广义多边形　(b) 弹性包络线

图 10　复杂广义多边形示意图

Fig. 10　Complex generalized polygon schematics

4　结　论

1) 从 3 种典型简单闭广义多边形弹性包络线计算结果可知, 本文算法对凸区域较多和凹区域复杂的平面图形的凸包计算适用, 且具有较高的计算效率; 对较为复杂的广义多边形的凸包计算仍然成立, 但算法的计算效率与步长的选取相关, 采用动态步长可提高效率。

2) 本文算法提出利用物理模型计算凸包的新思路, 构造了弹性线在张力、支持力下稳定的计算条件, 并将其转换为可计算的数学方法, 实现实时稳健地求解平面上任意自由闭曲线的凸包。

本文算法尚有一定的改进空间, 如可通过删除不相关边等预处理减少边的计算数量从而提高计算效率; 也可借助于解析方法使动态步长的选取更加灵活。如何通过引介上述方法进一步提高本文算法的效率, 将是后续工作的研究重点。

参考文献 (References)

[1] GRAHAM R L. An efficient algorithm for determining the convex hull of a finite planar set[J]. Information Processing Letters, 1972, 1(4): 132-133.

[2] JARVIS R A. On the identification of the convex hull of a finite set of points in the plane[J]. Information Processing Letters, 1973, 2(1): 18-21.

[3] PREPARATA F P, HONG S J. Convex hulls of finite sets of points in two and three dimensions[J]. Association for Computing Machinery, 1977, 20(2): 87-93.

[4] LINH N K, AN P T, HOAI T V. A fast and efficient algorithm for determining the connected orthogonal convex hulls[J]. Applied Mathematics and Computation, 2022, 429: 127183.

[5] SKLANSKY J. Measuring concavity on a rectangular mosaic[J].

IEEE Transactions on Computers, 1972, 100(12): 1355-1364.

[6] MCCALLUM D, AVIS D. A linear algorithm for finding the convex hull of a simple polygon[J]. Information Processing Letters, 1979, 9(5): 201-206.

[7] LEE D T. On finding the convex hull of a simple polygon[J]. International Journal of Computer & Information Sciences, 1983, 12(2): 87-98.

[8] MELKMAN A A. On-line construction of the convex hull of a simple polyline[J]. Information Processing Letters, 1987, 25(1): 11-12.

[9] CHEN C L. Computing the convex hull of a simple polygon[J]. Pattern Recognition, 1989, 22(5): 561-565.

[10] 吴中海, 叶澄清, 潘云鹤. 一个改进的简单多边形凸包算法[J]. 计算机辅助设计与图形学学报, 1997, 9(1): 9-13.
WU Z H, YE C Q, PAN Y H. An improved algorithm of convex hull computing[J]. Journal of Computer-Aided Design & Computer Graphics, 1997, 9(1): 9-13(in Chinese).

[11] 孔宪庶, 蔡洪学. 简单多边形凸包的双动线检测算法[J]. 计算机学报, 1994, 17(8): 596-600.
KONG X S, CAI H X. An algorithm for finding the convex hull of a simple polygon using active double line testing[J]. Chinese Journal of Computers, 1994, 17(8): 596-600(in Chinese).

[12] 王丽青, 陈正阳, 陈树强, 等. 一个改进的简单多边形凸包算法[J]. 计算机工程, 2007, 33(3): 200-201.
WANG L Q, CHEN Z Y, CHEN S Q, et al. An improved algorithm of simple polygon convex hull[J]. Computer Engineering, 2007, 33(3): 200-201(in Chinese).

[13] 张林. 改进的多边形凸包算法[J]. 长春工业大学学报 (自然科学版), 2013, 34(5): 560-563.
ZHANG L. An improved algorithm of polygon convex hull[J]. Journal of Changchun University of Technology(Natural Science Edition), 2013, 34(5): 560-563(in Chinese).

[14] 刘子尧. 基于夹角的凸包算法改进[J]. 软件导刊, 2018, 17(4): 94-96.
LIU Z Y. Improvement of convex hull algorithm based on included angle[J]. Software Guide, 2018, 17(4): 94-96(in Chinese).

[15] WANG C, ZHOU R G. A quantum search algorithm of two-dimensional convex hull[J]. Communications in Theoretical Physics, 2021, 73(11): 115102.

[16] 杨宏伟. 平面散乱点云凸包快速求解算法[J]. 机械研究与应用, 2021, 34(2): 55-56.
YANG H W. Accelerating algorithm for convex hull construction of 2d scattered point cloud[J]. Mechanical Research & Application, 2021, 34(2): 55-56 (in Chinese).

[17] 林晓, 刘祖祥, 郑晓妹, 等. 基于凸包改进的流行排序显著性检测[J]. 计算机辅助设计与图形学学报, 2019, 31(5): 761-770.
LIN X, LIU Z X, ZHENG X M, et al. Saliency detection based on improved manifold ranking via convex hull[J]. Journal of Computer-Aided Design & Computer Graphics, 2019, 31(5): 761-770(in Chinese).

[18] CHEN M M, ZHENG D, RENG X Q. Research on gesture recognition technology based on convex hull algorithm[J]. Scientific Journal of Intelligent Systems Research, 2021, 3(7): 192-199.

[19] 武伟璐. 基于快速凸包算法的 4D 航迹规划研究[D]. 天津: 中国民航大学, 2020.
WU W L. Research on 4D trajectory planning based on quickhull algorithm[D]. Tianjin: Civil Aviation University of China, 2020(in Chinese).

[20] WU Y, GUPTA A, KURZEJA K, et al. Chocc: Convex hull of cospherical circles and applications to lattices[J]. Computer-Aided Design, 2020, 129: 102903.

[21] NGUYEN L K, SONG C, RYU J, et al. QuickhullDisk: A faster convex hull algorithm for disks[J]. Applied Mathematics and Computation, 2019, 363: 124626.

[22] ELBER G, KIM M S, HEO H S. The convex hull of rational plane curves[J]. Graphical Models, 2001, 63(3): 151-162.

[23] KIM Y J, LEE J, KIM M S, et al. Efficient convex hull computation for planar freeform curves[J]. Computers & Graphics, 2011, 35(3): 698-705.

Computing convex hull of a generic polygon with simulation of progressive support for an elastic line

CUI Yuping，LI Zihan，ZHENG Guolei*

(School of Mechanical Engineering and Automation，Beihang University，Beijing 100191，China)

Abstract：The computation of the convex hull of the Jordan curve has found widespread application in recent years. A simulation of progressive support for an elastic line approach was suggested in this paper to determine the elastic envelope of a straightforward closed generic polygon. Based on the physical model, this algorithm could determine whether a point was a balanced fulcrum on the elastic envelope line by judging whether it was balanced by force. According to these findings, the algorithm performed different operations such as forward, spring-back, and jump respectively, until all the balanced support points were selected and the elastic envelope was eventually generated. The contrastive analysis of three typical generic polygons demonstrates that the proposed algorithm can solve the elastic envelope of arbitrary simple closed generic polygons synchronously, and it is robust, efficient, and universally applicable.

Keywords：generic polygon; curved shape; convex hull; elastic envelop; supporting line

Received：2022-04-15；**Accepted**：2022-06-05；**Published Online**：2022-06-13 16：41

URL：link.cnki.net/urlid/11.2625.V.20220613.1350.002

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0189

飞机地面除冰保障过程动态预测

李彪[1,2]，邢志伟[2,*]，王立文[1]
(1. 中国民航大学 航空工程学院，天津 300300；2. 中国民航大学 电子信息与自动化学院，天津 300300)

摘 要：针对冰雪气象下除冰保障过程的精细化管理及预测精度低下的问题，提出了一种基于时空关联动态贝叶斯网络的飞机地面除冰保障过程动态预测方法。在系统性分析除冰保障过程的基础上，设计了面向离港除冰队列的时空关联节点判别方法，基于K最近邻算法简化关联节点并构建了变结构动态贝叶斯网络模型。基于核注意力机制的除冰保障节点先验概率密度估计方法，结合条件概率更新结果构建了面向不同状态的飞机地面除冰保障过程动态预测方法。实验结果表明：所提方法在考虑除冰保障演化不确定性的基础上实现了各节点的动态预测，平均绝对误差为2.34 min，整体精度相比静态贝叶斯网络方法最大提升8.66%，为地面除冰运行战术组织及决策控制提供了有效依据。

关 键 词：航空运输；动态预测；时空关联动态贝叶斯网；核注意力机制；地面除冰保障过程
中图分类号：V351.392; U8
文献标志码：A **文章编号**：1001-5965(2024)01-0224-10

安全是四型机场建设的核心要义，冰雪气象下的飞机地面除冰保障过程作为影响中高纬度机场协同运行的重要场景之一，其态势的精准预测成为了特殊环境下机场安全和性能提升的热点和难点。目前，国内外学者针对包括飞机除冰等不利条件在内的机场运行保障过程预测方法开展了一系列研究。文献[1]面向机场协同决策(airport collaborative decision making, A-CDM)架构，阐述了除冰保障运行过程的关键里程碑节点及协同管理体系，但未明确区分除冰模式及作业类型。文献[2]针对日常过站航班保障过程节点的演化及推理设计了基于动态贝叶斯网络的动态预测方法，但并未考虑除冰保障等异常环境的适用性。文献[3]基于长短期记忆(long short-term memory, LSTM)神经网络建立了数据驱动的航班延误预测模型，同时分析了包括除冰在内的不利气象对结果产生的制约。文献[4]评估了雷暴天气对航班保障计划的影响，得到了基于鲁棒开环最优控制的目标函数，通过NavSim软件预测计划的变更及其演化规律。文献[5]利用图信号及其频谱表征了机场保障网络的时空分布，基于异常值识别并预测了各子空间的延误情况，研究了包括冰雪条件在内的系统级机场延误网络交互关系。文献[6]建立了基于多性能参数的机场网络延误状态迁移模型，并采用贝叶斯估计等方法提升模型预测精度及泛化能力，但缺乏除冰运行模式的转换机制。文献[7]结合机场构型及运行保障环节，差异计量离港网络波及效应并预测平均滑出时间，但其稳定性低，无法适应飞机地面除冰保障过程的动态预测。

上述方法存在除冰保障过程针对性弱、预测精度低、稳定性差的不足之处，基于此，本文提出了一种基于时空关联贝叶斯网络的飞机地面除冰保障过程动态预测方法，构建飞机地面除冰保障时空关联动态贝叶斯网络(spatial-temporal correlation dynamic

收稿日期：2022-03-25；**录用日期**：2022-07-15；**网络出版时间**：2022-07-22 17:29
网络出版地址：link.cnki.net/urlid/11.2625.V.20220721.1719.003
基金项目：国家重点研发计划(2018YFB1601200)
*通信作者. E-mail：cauc_xzw@163.com

引用格式：李彪，邢志伟，王立文. 飞机地面除冰保障过程动态预测[J]. 北京航空航天大学学报，2024，50(1)：224-233.
LI B，XING Z W，WANG L W. Dynamic prediction for aircraft ground deicing operation process[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(1)：224-233 (in Chinese).

Bayesian network, ST-DBN) 模型以提升鲁棒性[8]。在此基础上, 设计基于核注意力机制 (kernel attention mechanism, K-Attention) 的先验概率动态更新方法, 重点关注网络模型的预测精度和适应能力, 并基于实际数据验证了方法的合理性和有效性。

1　除冰保障过程 ST-DBN 模型

1.1　除冰保障过程分析

飞机地面除冰是指冰雪气象环境下采用喷洒除冰液的方式去除机翼、机身、尾翼等飞机关键表面积冰 (霜) 以保障飞机起飞和飞行安全的一系列过程。根据不同除冰模式, 可分为机位与集中除冰 (包括关车和慢车 2 种), 其中, 作业类型又分为双车和四车, 其作业实况如图 1 所示。

结合机场协同决策关键里程碑节点演化情况及国内枢纽机场除冰保障作业和运行流程, 可将飞机地面除冰保障过程视为不利冰雪条件启动除冰作业时离港保障里程碑节点的总称, 包括进入机场运控离港除冰队列到离港起飞的所有核心节点。其中, 机位和集中除冰保障过程的时空顺序存在一定差异, 基本情况如表 1 所示。



图 1　飞机地面除冰实况

Fig. 1　Actual situation of aircraft ground deicing

表 1　不同模式的除冰保障过程

Table 1　Deicing operation process of different modes

模式	里程碑节点	A-CDM代码	空间状态
机位除冰	进入除冰队列, 允许登机	PPT	机位
	登机结束	EPT	机位
	撤轮挡	OBT	机位
	除冰开始	CZT	机位
	除冰结束	EZT	机位
	推出滑行	SAT	滑行道
	离港起飞	TOT	跑道端头
集中除冰	进入除冰队列, 允许登机	PPT	机位
	登机结束	EPT	机位
	撤轮挡/推出	SAT	机位
	到达除冰等待点	WZT	滑行道
	除冰开始	CZT	除冰位
	除冰结束	EZT	除冰位
	离港起飞	TOT	跑道端头

注: 若代码前缀为A, 表示实际时间, 即该节点已发生或正在发生; 若代码为E, 表示估计时间, 即该节点未发生。

为后续方便处理, 将机场飞行区按照空间构型分类编码, 采用三位整型变量表示除冰保障过程具体空间状态, 其中, 机位、滑行道、除冰位、跑道端头的百位数取值分别为 1～4, 后 2 位由各区位构型实际差异决定。由机场除冰运行手册及相关安全标准, 对地面除冰保障过程做如下理想化处理: ①各区域停机坪到除冰坪或起飞跑道端头的滑行路径相对固定, 不考虑飞机离港滑行路径的动态优化; ②机场启动离港除冰后, 所有飞机须除冰后放行起飞, 忽略始发不除等特殊情况; ③假设除冰车、加液站等除冰资源充足并按离港队列开展保障作业, 其调度问题及所产生的等待时间不在本文研究范畴内。

综上, 飞机地面除冰保障过程可视为一个带有时空关联的拓扑结构, 如图 2 所示。图中: 各节点用$X_{(j,k)}^{i}$表示, 涵义是长度为N的离港除冰队列中第$i(i=1,2,\cdots,N)$架飞机地面除冰保障过程中所在空间状态为j(结合表 1 与机场实际构型确定) 的第k个节点 ($k=1,2,\cdots,7$), 通过构建以第i架飞机除冰保障过程为主干的拓扑结构, 判别其他过程$l(l=1,2,\cdots,N;\ l\neq i)$内节点和主干各节点的关联关系, 由此设计面向除冰保障过程的网络模型及动态预测方法[9]。

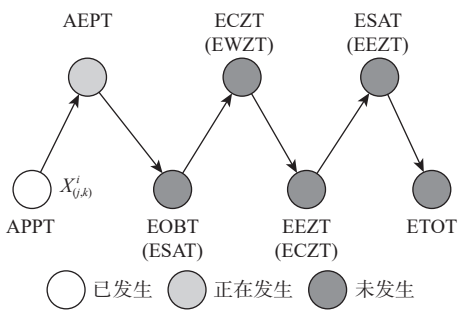


图 2　飞机地面除冰保障过程拓扑结构

Fig. 2　Topological structure of aircraft ground deicing operation process

1.2　ST-DBN 模型

动态贝叶斯网络作为复杂因果推理的重要建模工具之一, 已被广泛应用于系统辨识、过程决策、事件预测等方面[10-12]。为构建除冰保障过程的贝叶斯网络模型, 本文提出了面向除冰队列排序的时空关联节点判别方法。针对队列长度为N的第i架飞机的除冰保障过程, 各节点和其他过程$l(l=1,2,\cdots,N;\ l\neq i)$中节点$X_{(j',k')}^{l}$的空间关联属性$S_k(i,l)$为

$$S_k(i,l)=\delta_s|j-j'| \tag{1}$$

$$\delta_s=\begin{cases}0 & s_i^k(j)-s_l^{k'}(j')\neq 0\\ 1 & s_i^k(j)-s_l^{k'}(j')=0\end{cases} \tag{2}$$

式中：δ_s为除冰保障节点空间关联决策变量，用于判别不同除冰保障过程的2个节点是否属于同一空间区域；j'为第l架次非主干的飞机地面除冰保障过程k'节点的空间状态；$s_i^k(j)$、$s_l^{k'}(j')$分别为机场实际区位编号经百位高斯取整函数转化前后的节点空间状态。

同理，时间关联属性$T_k(i,l)$为

$$T_k(i,l)=|X_{(j,k)}^i(t)-X_{(j',k')}^l(t)| \tag{3}$$

式中：$X_{(j,k)}^i(t)$和$X_{(j',k')}^l(t)$为除冰保障节点的时间状态，由于不同模式下整体时空逻辑存在差异，节点顺序也可能表现出一定偏差[13]。

基于拉普拉斯标准分布函数将时空关联属性叠加并归一化，即除冰保障过程时空关联判定因子$\alpha_k(i,l)$为

$$\alpha_k(i,l)=\frac{1}{2}\left[\exp(-|S_k(i,l)-\gamma_S|)+\exp(-|T_k(i,l)-\gamma_T|)\right] \tag{4}$$

式中：γ_S为除冰保障过程空间关联期望容限，与当前主干节点的空间演化状态相关；γ_T为时间关联期望容限，结合机场除冰运行实际取5 min。

分别计算除冰队列中各架次保障过程不同节点与当前主干过程中所有节点的时空关联判定因子$\alpha_k(i,l)$，若大于0.5，则节点之间存在关联关系[14]。同时，为提升除冰保障过程节点预测结果的更新速率，基于K-最近邻算法简化主干过程的关联节点[15]，基本流程如图3所示。其中，$d(k,z)$为各关联节点z与k节点的时空距离，即

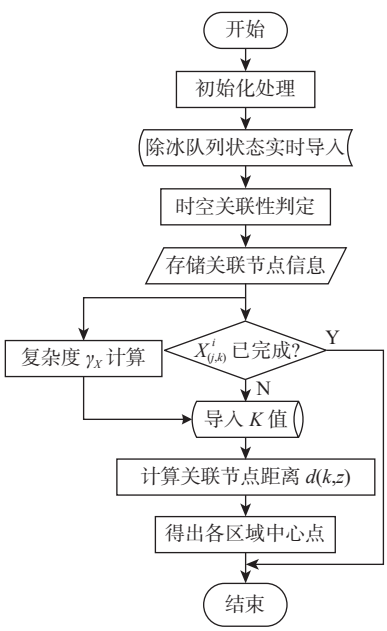


图3 关联节点简化算法流程

Fig. 3 Flow chart of simplified algorithm for associated nodes

$$d(k,z)=\sqrt{(t_k-t_z)^2+(s'(k)-s'(z))^2} \tag{5}$$

式中：t_k和t_z分别为节点实际时间和预测时间；$s'(k)$和$s'(z)$分别为经十位高斯取整函数转化后的实际空间状态和预测空间状态。

节点关联复杂度γ_X由离港除冰队列飞机数量N和关联节点数量N_k共同决定，计算式为

$$\gamma_X=\sum_{k=1}^{7}|N-N_k|/(7N) \tag{6}$$

最终，面向过程间相关性得到以第i架飞机除冰保障过程为主干的变结构ST-DBN模型[16]，定义如下。

定义1 面向飞机地面除冰保障过程的ST-DBN模型可定义为$\left(B^0,B_{\rightarrow}^{\tilde{T}},B_{\rightarrow}^{2\tilde{T}},\cdots,B_{\rightarrow}^{\lambda\tilde{T}},\cdots\right)$，其中，$B^0$为进入离港除冰队列后的初始贝叶斯网络，$B_{\rightarrow}^{\lambda\tilde{T}}$为$\lambda$时间片内的贝叶斯网络，$\tilde{T}$为时间片更新周期。相邻时间片间各除冰保障节点的条件概率分布$P\left\{X_{(j,k)}^i\left(\lambda\tilde{T}\right)|X_{(j,k)}^i\left[(\lambda-1)\tilde{T}\right]\right\}$为

$$P\left\{X_{(j,k)}^i(\lambda\tilde{T})|X_{(j,k)}^i\left[(\lambda-1)\tilde{T}\right]\right\}=P\left\{X_{(j,k)}^i\left(\lambda\tilde{T}\right)\right\}\sum_{m=1}^{n_z}\alpha_{z_m}(\lambda\tilde{T})P_a\left\{z_m\left[(\lambda-1)\tilde{T}\right]\right\} \tag{7}$$

式中：$P\{X_{(j,k)}^i\left(\lambda\tilde{T}\right)\}$为网络主干节点的先验概率；$P_a\left\{z_m\left[(\lambda-1)\tilde{T}\right]\right\}$为关联父节点$z_m$产生的条件概率；$\alpha_{z_m}\left(\lambda\tilde{T}\right)$为关联节点的归一化加权；$n_z$为$\lambda\tilde{T}$时间片内关联父节点数量。

$$\alpha_{z_m}(\lambda\tilde{T})=d(k,z_m)\Big/\sum_{m=1}^{n_z}d(k,z_m) \tag{8}$$

由于飞机除冰保障过程时空演化特性存在一定的相似，且节点间的关联关系随时间动态演化，构建的ST-DBN模型存在如下性质：

1）相对平稳性。随着时间的演化及关联节点的更新，$B_{\rightarrow}^{\lambda\tilde{T}}$网络结构及参数可能发生变化，除冰保障过程贝叶斯网络所有的根节点均已发生且不包含回环[17]。

2）一阶马尔可夫性。各除冰保障节点之间的边界由时空关联性判定得出，且位于同一时间片或相邻时间片，不存在跨越或突变的情况。

结合ST-DBN模型的性质，本文设计了面向变结构加权贝叶斯网络的条件概率推理方法，网络主干过程各关联父节点的条件概率$P_a(z_m)$为

$$P_a(z_m)=P(z_m)\Big/P(Z^{i'}) \tag{9}$$

式中：$P(z_m)$为z_m节点发生的概率；$Z^{i'}$为z_m节点在离港除冰队列中所处$i'=1,2,\cdots,N$；$i'\neq i$过程的所有

祖先节点集合, 其元素定义为$z_h^{i'}$, 由此先验概率$P\left(Z^{i'}\right)$可基于链式法则解得, 即

$$P\left(Z^{i'}\right)=\prod_{h=1}^{m_z-1}P\left(z_h^{i'}|z_{h+1}^{i'}\right) \tag{10}$$

其中: m_z为当前状态祖先节点的数量; $P\left(z_h^{i'}|z_{h+1}^{i'}\right)$为除冰保障过程前后父子节点的条件概率。

通过式 (7)～式 (10) 可得飞机地面除冰保障过程 ST-DBN 模型的概率推理过程, 导出 ST-DBN 主干过程节点的条件概率密度模型, 由条件概率的演化可动态预测飞机地面除冰保障过程。

2　除冰保障过程动态预测

2.1　核注意力机制

注意力机制核心思想源于人类视觉的注意力, 广泛应用于计算机视觉、深度学习、自然语言处理等领域[18]。结合飞机地面除冰保障过程的队列属性演化情况, 设计了基于核注意力机制的除冰保障过程节点先验概率密度估计方法。核注意力机制的总体工作原理如图 4 所示。由 ST-DBN 模型的主干除冰保障过程态势确定动态属性及核函数更新的数据库, 同时结合注意力机制基本流程及状态转换特性确定各除冰保障节点先验概率核函数的配置参数, 从而支撑面向 ST-DBN 模型动态演化的概率推理。

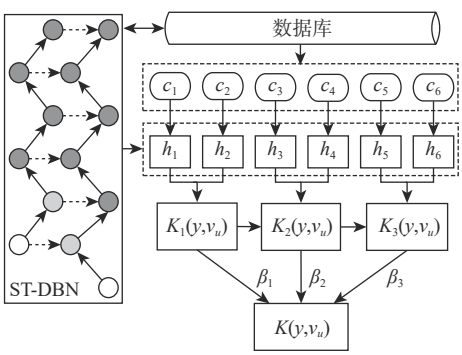


图 4　核注意力机制工作原理

Fig. 4　Mechanism of kernel attention

根据除冰队列演化情况及除冰保障过程运行规范的核心内容, 构建除冰保障过程 ST-DBN 模型的动态属性集合为$C=\{c_w|w=1,2,3,4,5,6\}$。其中, c_1为除冰队列容量, 由上时段机场除冰系统表现得出; c_2为机型, 分为大型、中型、小型机; c_3为平均准点率, 表示当前时段所有进港航班到达情况; c_4为除冰模式, 包括慢车和关车 2 种; c_5为起飞容限, 指放行起飞目标值和预测值间的偏差; c_6为作业模式, 分为双车和四车。

$$c_1=(N^*-N_{\lambda-1})/N^* \tag{11}$$

$$c_3=\frac{1}{N}\sum_{i=1}^{N}(T^*-|a_i-e_i|)/T^* \tag{12}$$

$$c_5=|r_i-w_i|\big/T^* \tag{13}$$

式中: N^*为当前除冰资源配置下单位时段内允许进入除冰队列的最大飞机数; $N_{\lambda-1}$为上一时段完成除冰作业飞机数; T^*为机场航班运行最大容差, 一般取 30 min; a_i和e_i分别为第i架飞机实际和预计到达时间; r_i和w_i分别为第i架飞机预计和目标起飞时间。

结合机场除冰运行保障规范及标准流程, 并由图 4 可知, ST-DBN 模型的属性集合C决定各核函数的配置参数[19], 核注意力机制的状态转换实现如下:

$$h_w=\left\{1+\exp[-(c_wN(c_w)/N^*)]\right\}^{-1} \tag{14}$$

$$\beta_u=\exp(N(c_u))\Big/\sum_{w=1}^{3}\exp(N(c_w)) \tag{15}$$

$$v_u=\tanh\left(\sum_{p=2u-1}^{2u}c_wh_p+e_u\right) \tag{16}$$

$$K(y,v_u)=\sum_{u=1}^{3}\beta_uK_u(y,v_u) \tag{17}$$

式中: h_w为各属性元素输入对应注意力的隐藏状态; $N(c_u)$为动态属性下的样本数量; β_u为归一化后的信息权重; v_u为注意力表征机制, 表示核函数带宽; e_u为当前状态的偏移量, 可由节点预测值与目标的平均相对误差表示; $K(y,v_u)$为注意力状态转换后的核函数; $K_u(y,v_u)$为各模块的核函数, 由历史数据分析结果分别选择高斯、指数、柯西核函数更新先验概率[20], 即

$$K_1(y,v_u)=\frac{1}{\sqrt{2\pi}}\exp\left(\frac{y^2}{v_u}\right) \tag{18}$$

$$K_2(y,v_u)=\exp(y/v_u) \tag{19}$$

$$K_3(y,v_u)=\left(1+y^2/v_u\right)^{-1} \tag{20}$$

综上, 可实时计算各节点先验概率值。

2.2　动态预测方法

根据飞机地面除冰保障过程的 ST-DBN 模型及面向离港除冰队列注意力机制的核概率密度方法, 本文设计了一种面向除冰保障过程动态演化的预测方法, 具体步骤如下:

步骤 1　初始化处理。基于“冰雪气象相似日”思想导入机场除冰保障过程历史数据[21-23], 经数据转换形成包括第i个飞机地面除冰保障过程所有属性及节点时空数据的样本空间Ω。

步骤 2　获取当前离港除冰队列信息。依据所

预测除冰保障过程各节点的时空状态判别队列中其他过程节点的关联性，并基于K-最近邻算法简化关联节点，得到当前状态下的ST-DBN模型。

步骤3 估计各节点的先验概率。结合ST-DBN模型与核注意力机制下的动态属性集合得到各除冰保障过程节点的概率密度核函数，面向样本空间Ω实时构建对应节点先验概率密度估计模型$\hat{f}(y)$，即

$$\hat{f}(y)=\frac{1}{N_0 v_u}\sum_{m'=1}^{N_0}K(y-y_{m'}) \tag{21}$$

式中：N_0为Ω的样本数量；$y_{m'}$为样本实际值。

步骤4 条件概率推理。面向ST-DBN模型的构型推导地面除冰保障过程各节点的条件概率模型，以条件概率最大值的对应时间作为节点预测值。

步骤5 过程演化及更新。随时间更新ST-DBN模型的构型，重复步骤2～步骤4，从而实时更新除冰保障节点预测时间，直至第i架飞机完成起飞。

3 实验与分析

验证数据来源于国内某枢纽机场除冰指挥系统中的共享运行数据，具体情况如表2和表3所示。表中：IBT表示上轮挡。以该机场2020年冬春航季历史数据开展飞机地面除冰保障过程的动态预测，对机场飞行区除冰保障空间区位重新编码，结合其构型及其分布确定关联节点简化K值为4。

表2 除冰保障过程属性数据样例

Table 2 Data sample of aircraft deicing operation process attributes

属性	状态或数值
出港日期	2020-12-12
出港航班号	CA1805
机型	B738
出港机位	140
起飞跑道	36R
准点率/%	60
除冰模式	双车
作业模式	关车
除冰容量/(架次·h^{-1})	27
除冰类型	机位

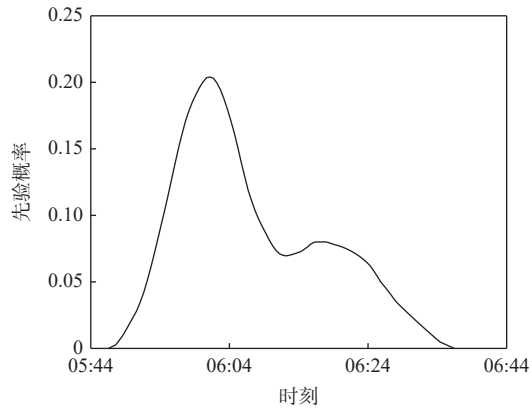
表3 除冰保障过程数据样例

Table 3 Data sample of aircraft deicing operation process

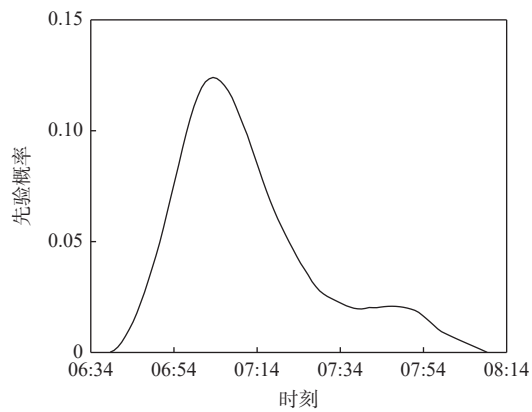
除冰保障过程节点	时刻
IBT	05:24
PPT	06:06
EPT	06:26
OBT	06:51
CZT	06:50
EZT	07:00
SAT	07:07
TOT	07:17

3.1 除冰保障过程仿真实验

参照“冰雪气象相似日”特征提取历史数据，针对所预测的除冰保障主干过程各节点条件概率密度核函数的演化情况，构建对应的样本空间，并基于核注意力机制得到各除冰保障过程节点的先验概率模型。其中，登机结束和推出滑行节点的先验概率模型如图5所示。



(a) 登机结束 (EPT)



(b) 推出滑行 (SAT)

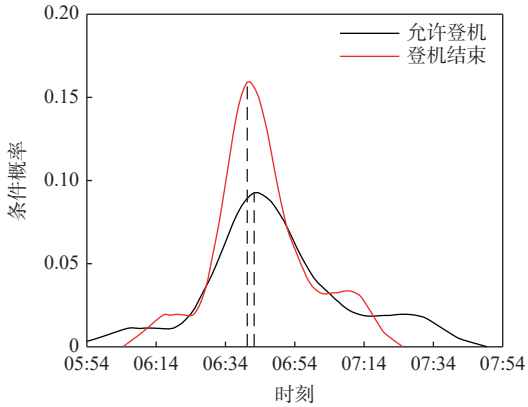
图5 关键除冰保障过程节点先验概率模型

Fig. 5 Prior probability model of key deicing operation process node

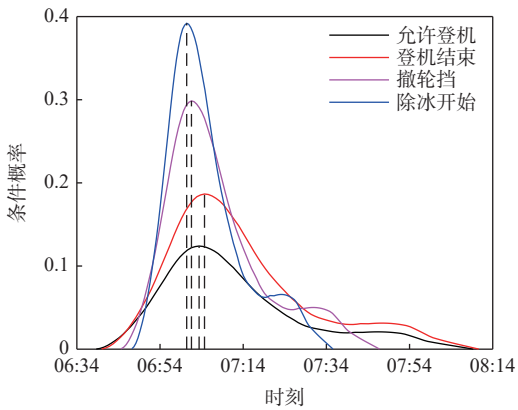
基于ST-DBN模型开展除冰保障节点的概率推理，并以条件概率最大值作为各节点的预测结果。随时间演化更新预测过程的关联节点及ST-DBN构型，并实时开展除冰保障过程节点的概率推理。其中，撤轮挡和除冰结束节点的条件概率演化过程如图6所示。

以各除冰保障节点与上轮挡节点时刻的差值表示预测结果，如表4所示。

由于ST-DBN模型生成过程及核注意力机制下的概率密度估计考虑了除冰保障过程的外部影响，同时除冰类型不同导致节点时空顺序存在一定差距，分别选择30个机位和集中除冰的保障过程



(a) 撤轮挡 (OBT)



(b) 除冰结束 (EZT)

图 6　除冰保障节点条件概率模型演化过程

Fig. 6　Evolution process for conditional probability model of deicing operation node

表 4　单除冰保障过程预测结果

Table 4　Prediction results of single deicing operation process

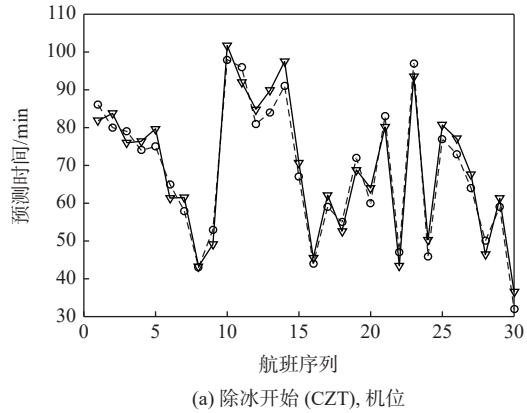
除冰保障节点	初始预测时间/min	最终预测时间/min	实际时间/min
PPT	40.22	40.22	42
EPT	62.11	60.31	58
OBT	78.28	76.32	77
CZT	88.94	87.94	86
EZT	99.37	96.41	96
SAT	98.92	105.81	103
TOT	108.83	111.68	113

展开预测, 其中, 开始除冰和离港起飞节点预测结果如图 7 所示。

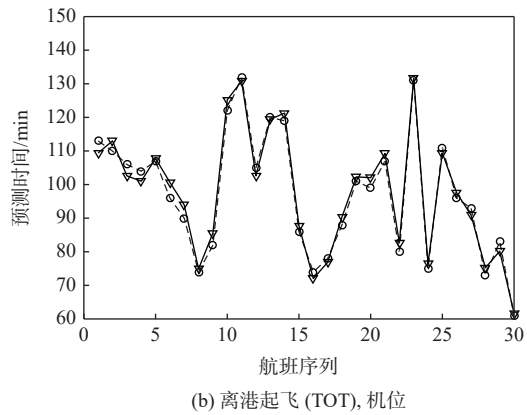
3.2　实验结果分析

首先, 针对飞机地面单除冰保障过程开展结果分析, 即主干除冰保障过程各节点在实际动态预测中内部演化情况, 分别计算单除冰保障过程各节点的平均绝对百分误差 MAPE 和平均绝对误差 MAE, 即

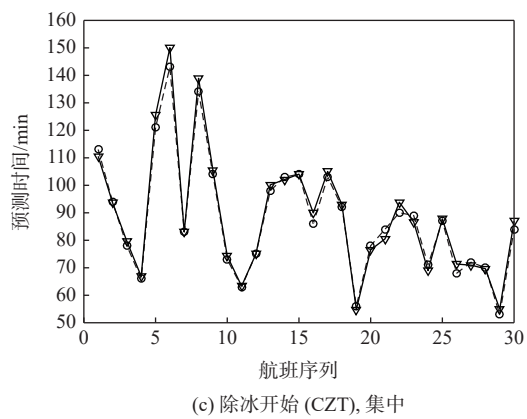
$$e_{\mathrm{MAPE}}=\frac{1}{M_k}\sum_{k_0=1}^{M_k}\left|\frac{Y_i^k-\hat{Y}_i^k(k_0)}{Y_i^k}\times 100\%\right| \tag{22}$$



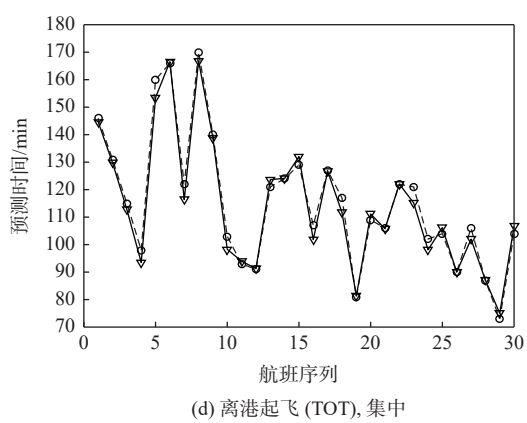
(a) 除冰开始 (CZT), 机位



(b) 离港起飞 (TOT), 机位



(c) 除冰开始 (CZT), 集中



(d) 离港起飞 (TOT), 集中

实际值　预测值

图 7　不同除冰保障过程预测结果

Fig. 7　Prediction results of different deicing operation processes

$$e_{\mathrm{MAE}}=\frac{1}{M_k}\sum_{k_0=1}^{M_k}\left|Y_i^k-\hat{Y}_i^k(k_0)\right| \tag{23}$$

式中：Y_i^k为第i架次第k个除冰保障节点的真实值；$\hat{Y}_i^k(k_0)$为该除冰保障节点的第k_0次预测结果；M_k为预测次数。

计算结果如表5所示。可知，所有节点的平均绝对误差控制在5 min内，且均值为2.34 min；平均绝对百分误差的均值为3.03%，标准差为1.42%，低于标准线10.0%，表明结果能够满足除冰保障过程动态预测需求，满足机场除冰战术决策所规定的最高容限。

表5 单除冰保障过程预测结果误差分析

Table 5 Error analysis of prediction results for single deicing operation process

除冰保障节点	平均绝对百分误差/%	平均绝对误差/min
PPT	4.24	1.78
EPT	5.34	3.21
OBT	0.86	0.67
CZT	2.61	2.25
EZT	2.47	2.37
SAT	2.85	2.94
TOT	2.81	3.17

同时，为分析该预测方法的鲁棒性，以不同模式的除冰保障过程为主干开展动态预测的验证，即讨论预测方法对多除冰保障过程的适应能力，计算多除冰保障过程各节点综合的平均相对误差MRE和均方根误差RMSE，即

$$e_{\mathrm{MRE}}=\sum_{k=1}^{7}\left|Y_i^k-\bar{\hat{Y}}_i^k\right|/\bar{\hat{Y}}_i^k \tag{24}$$

$$e_{\mathrm{RMSE}}=\sqrt{\frac{1}{7}\sum_{k=1}^{7}\left(Y_i^k-\bar{\hat{Y}}_i^k\right)^2} \tag{25}$$

$$\bar{\hat{Y}}_i^k=\frac{1}{M_k}\sum_{k_0=1}^{M_k}\hat{Y}_i^k(k_0) \tag{26}$$

式中：$\bar{\hat{Y}}_i^k$为该除冰保障节点所有预测结果的均值。

具体计算结果如图8所示。可知，2种模式的平均相对误差均小于0.5，慢车模式的预测效果优于关车模式，表明模型具有相对稳定性；同时，均方根误差远小于0.1 min，满足预测模型精确性要求。

3.3 对比分析

对于除冰保障微观演化过程的先验概率更新方式，采用高斯核概率密度估计与直方图方式动态预测除冰保障过程，与加入核注意力机制的本文方法进行比较，最终预测结果如图9所示。可知，本文方法单除冰保障过程节点的平均绝对误差比其他方法最大降低4.85 min，整体除冰保障过程，平均绝对误差的均值比直方图和高斯核方法分别降低2.37 min和2.45 min，表明预测结果的稳定性优于其他方法。

为验证基于ST-DBN模型的飞机地面除冰保障过程预测方法的稳定性和精确性，选择最能反映多除冰保障过程性能的离港起飞节点开展预测分析，并选择图神经网络(graph neural network, GNN)模型和静态贝叶斯网络(static Bayesian network, SBN)

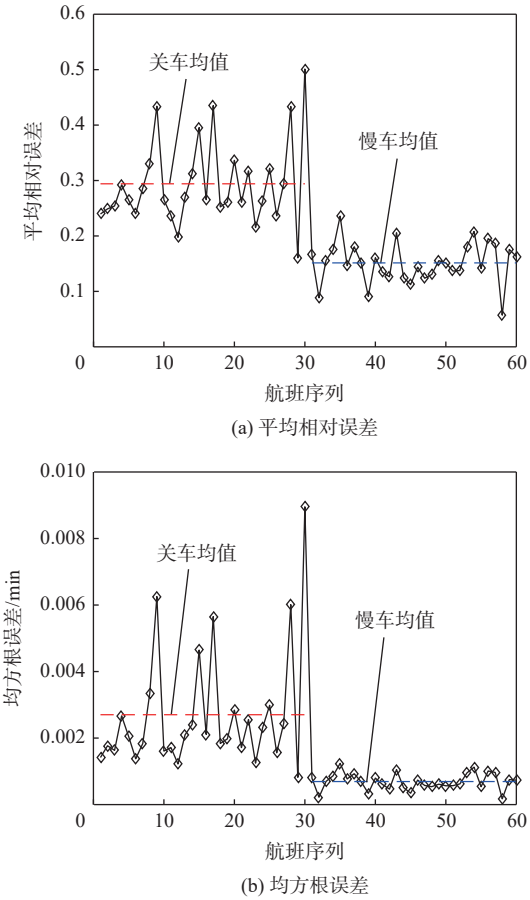


图8 多除冰保障过程预测结果误差分析

Fig. 8 Error analysis of prediction results for multiple deicing operation process

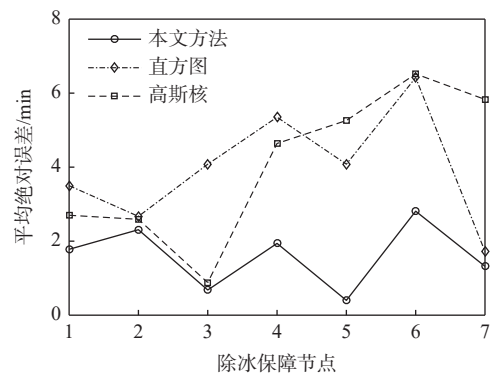


图9 不同概率更新方式预测结果的平均绝对误差对比

Fig. 9 Comparison of mean absolute error of prediction results for different probability updating methods

与本文方法的预测结果作对比, 最终的预测结果如图 10 所示。通过计算分析可知, ST-DBN 模型的平均精度为 87.79%, 高于 GNN 和 SBN 的 81.67% 和 79.13%。

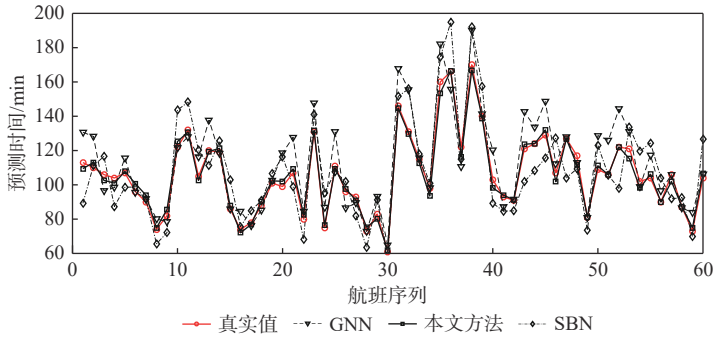


图 10　不同方法的离港起飞节点预测结果对比

Fig. 10　Comparison of departure & take-off node prediction results for different methods

分析 ST-DBN 模型简化前后的预测精度及速率情况, 计算简化前后除冰保障过程动态贝叶斯网络的预测精度及速率可知, 基于 K 最近邻算法简化后的平均预测精度比简化前最大提升 0.79%, 避免了关联节点引发的过拟合; 简化后的预测速率与简化前相比最大提升 19.76%, 在最大程度保证总体精度的基础上满足了机场除冰运行保障决策需求。

同时, 分别计算 3 种方法的希尔不等系数 e_{TIC}, 即

$$e_{\mathrm{TIC}}=\frac{\sqrt{\dfrac{1}{7M'}\sum_{i=1}^{M'}\sum_{k=1}^{7}\left(Y_i^k-\bar{\hat{Y}}_i^k\right)^2}}{\sqrt{\dfrac{1}{7M'}\sum_{i=1}^{M'}\sum_{k=1}^{7}\left(Y_i^k\right)^2}+\sqrt{\dfrac{1}{7M'}\sum_{i=1}^{M'}\sum_{k=1}^{7}\left(\bar{\hat{Y}}_i^k\right)^2}} \tag{27}$$

式中: M'为各方法除冰保障预测次数。

计算可知, ST-DBN 模型的希尔不等系数为 0.013, 比 GNN 和 SBN 分别低 0.04 和 0.048, 表明本文方法的精度优于其他方法。综上可得, 本文所提除冰保障过程预测方法能够满足不利气象条件运行战术决策信息共享的基本要求。

4　结　论

本文基于变结构 ST-DBN 模型预测飞机地面除冰保障过程, 并面向机场实际运行数据开展实验验证, 得到结论如下:

1) 时空关联判定决定贝叶斯网络模型的基本结构, 结合 K-最近邻算法简化, 为 ST-DBN 模型提升了整体的预测精度。

2) 队列容量、机型、准点率、除冰模式、起飞容限、作业模式等 6 个因素限制了除冰保障过程的演化进程, 因此将其作为核注意力机制下先验概率密度估计的输入。

3) 飞机地面除冰保障过程的预测精度随时间演化发生改变, 需根据运行态势调节注意力以提升概率推理能力。

冰雪不利条件下的除冰保障过程受环境、资源、运行模式等多重约束, 影响其预测精度的原因也相对复杂。本文方法仅考虑了外部属性所造成的不确定性, 因此尚存在不足之处。后续将从环境演化关联及除冰资源微观运行感知层面提升除冰保障过程的预测能力。

参考文献 (References)

[1] 鲍帆, 蒋伟煜. 基于机场协同决策 (A-CDM) 的除冰管理研究[C]//第一届空中交通管理系统技术学术年会. 南京: 中国指挥与控制学会, 2018: 330-334.
BAO F, JIANG W Y. Research on deicing management based on airport CDM[C]//The First Annual Conference on Air Traffic Management System Technology. Nanjing: Chinese Institute of Command and Control, 2018: 330-334(in Chinese).

[2] 王立文, 李彪, 邢志伟, 等. 过站航班地面保障过程动态预测[J]. 北京航空航天大学学报, 2021, 47(6): 1095-1104.
WANG L W, LI B, XING Z W, et al. Dynamic prediction of ground support process for transit flight[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(6): 1095-1104(in Chinese).

[3] GUI G, LIU F, SUN J L, et al. Flight delay prediction based on aviation big data and machine learning[J]. IEEE Transactions on Vehicular Technology, 2020, 69(1): 140-150.

[4] GARCIA-HERAS J, SOLER M, GONZALEZ-ARRIBAS D, et al. Robust flight planning impact assessment considering convective phenomena[J]. Transportation Research Part C: Emerging Technologies, 2021, 123: 102968.

[5] LI M Z, KARTHIK G, HAMSA B. Graph signal processing techniques for analyzing aviation disruptions[J]. Transportation Science, 2021, 16(3): 1-22.

[6] 王春政, 胡明华, 杨磊, 等. 基于 Agent 模型的机场网络延误预测[J]. 航空学报, 2021, 42(7): 324604.
WANG C Z, HU M H, YANG L, et al. Airport network delay pre-

diction based on Agent model[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(7): 324604(in Chinese).

[7] BADRINATH S, BALAKRSHNAN H, MA J, et al. Comparative analysis of departure metering at United States and European airports[J]. AIAA Journal of Air Transportation, 2020, 28(3): 93-104.

[8] SALEH Y, SAMI E, FRANCOIS M. Computational fluid dynamics investigation of transient effects of aircraft ground deicing jets[J]. Journal of Thermophysics and Heat Transfer, 2019, 33(1): 117-127.

[9] 张政, 陈艳艳, 梁天闻. 基于网约车数据的城市区域出行时空特征识别与预测研究[J]. 交通运输系统工程与信息, 2020, 20(3): 89-94.
ZHANG Z, CHEN Y Y, LIANG T W. Regional travel demand mining and forecasting using car-hailing order records[J]. Journal of Transportation Systems Engineering and Information Technology, 2020, 20(3): 89-94(in Chinese).

[10] 邢志伟, 刘洪恩, 李彪, 等. 基于时空关联网络的机场机位运行过程建模[J]. 系统工程与电子技术, 2021, 43(3): 722-730.
XING Z W, LIU H E, LI B, et al. Modelling for airport gate operation process based on relational spatio-temporal network[J]. System Engineering and Electronics, 2021, 43(3): 722-730(in Chinese).

[11] FRIDERKOS O, OLIVE M, BARANGER E, et al. A non-intrusive space-time interpolation from compact stiefel manifolds of parametrized rigid-viscoplastic FEM problems[J]. Computational Mechanics, 2021, 68(4): 861-883.

[12] TEHRANI A F, YEH H G, KWON S C. BER performance of space-time parallel ICI cancellation of OFDM in MIMO power line communications[J]. IEEE Systems Journal, 2021, 15(2): 1742-1752.

[13] HAN X, HSIEH C, KO S. Spatial modeling approach for dynamic network formation and interactions[J]. Journal of Business and Economic Statistics, 2021, 39(1): 120-135.

[14] MONDO G, RODRIGUEZ M, CLARAMUNT C, et al. Modeling consistency of spatio-temporal graphs[J]. Data & Knowledge Engineering, 2013, 84(3): 59-80.

[15] 杨晓玲, 冯山, 袁钟. 基于相对距离的反 k 近邻树离群点检测[J]. 电子学报, 2020, 48(5): 937-945.
YANG X L, FENG S, YUAN Z. Outlier detection based on reversed k-nearest neighborhood MST of relative distance measure[J]. Acta Electronica Sinica, 2020, 48(5): 937-945(in Chinese).

[16] XU B, SUN F C. Composite intelligent learning control of strict-feedback systems with disturbance[J]. IEEE Transactions on Cybernetics, 2018, 48(2): 730-741.

[17] 娄文忠, 赵悦岑, 冯恒振, 等. 基于贝叶斯网络的 MEMS 安全系统可靠性分析[J]. 北京理工大学学报, 2021, 41(9): 952-960.
LOU W Z, ZHAO Y C, FENG H Z, et al. Reliability analysis on MEMS S&A device based on Bayesian network[J]. Transactions of Beijing Institute of Technology, 2021, 41(9): 952-960(in Chinese).

[18] 廖华年, 徐新. 基于注意力机制的跨分辨率行人重识别[J]. 北京航空航天大学学报, 2021, 47(3): 605-612.
LIAO H N, XU X. Cross-resolution person re-identification based on attention mechanism[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(3): 605-612(in Chinese).

[19] KARDAKIS S, PERIKOS I, GRIVOKOPOULOU F, et al. Examining attention mechanisms in deep learning models for sentiment analysis[J]. Applied Sciences, 2021, 11(9): 3883.

[20] YAN J, PENG Z, YIN H, et al. Trajectory prediction for intelligent vehicles using spatial-attention mechanism[J]. IET Intelligent Transport Systems, 2020, 14(4): 1855-1863.

[21] LAFHRISSI F, DOUZI S, DOUZI K, et al. IDS-attention: An efficient algorithm for intrusion detection systems using attention mechanism[J]. Journal of Big Data, 2021, 8(1): 1-21.

[22] PRECUP S A, GELLERT A, MATEI A, et al. Towards an assembly support system with dynamic Bayesian network[J]. Applied Sciences, 2022, 12(3): 985.

[23] 沈琳, 于劲松, 唐荻音, 等. 图模型与学习算法结合的贝叶斯网络自动建模[J]. 北京航空航天大学学报, 2016, 42(7): 1486-1493.
SHEN L, YU J S, TANG D Y, et al. Automatic learning of Bayesian network structure using graph model and learning algorithm[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(7): 1486-1493(in Chinese).

Dynamic prediction for aircraft ground deicing operation process

LI Biao[1, 2], XING Zhiwei[2, *], WANG Liwen[1]

(1. College of Aeronautical Engineering, Civil Aviation University of China, Tianjin 300300, China;
2. College of Electronic Information and Automation, Civil Aviation University of China, Tianjin 300300, China)

Abstract: Aiming at the problem of fine management and low prediction accuracy of deicing operation process under ice and snow weather, a prediction method for aircraft ground deicing operation process based on the temporal and spatial correlation dynamic Bayesian network is proposed. A spatial-temporal correlation node identification method for departure deicing queue is developed based on a systematic analysis of the deicing operation process. The correlation node is then simplified using the K-nearest neighbor algorithm, and a dynamic Bayesian network model with variable structure is created. A priori probability density estimation method for deicing operation nodes based on kernel attention mechanism is studied. Combined with the conditional probability updating results, a dynamic prediction method for aircraft ground deicing support process for different states is constructed. A dynamic prediction approach for the aircraft ground deicing support process is built using the conditional probability updating findings in combination. The average absolute error is 2.34 min, and the whole accuracy is increased by 8.66% compared with static Bayesian network method, which can provide an effective decision-making basis for the tactical organization and control of ground deicing operations.

Keywords: air transportation; dynamic prediction; spatial-temporal correlation dynamic Bayesian network; kernel attention mechanism; ground deicing operation process

Received: 2022-03-25; **Accepted**: 2022-07-15; **Published Online**: 2022-07-22 17:29
URL: link.cnki.net/urlid/11.2625.V.20220721.1719.003
Foundation item: National Key R & D Program of China (2018YFB1601200)
* **Corresponding author**. E-mail: cauc_xzw@163.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0209

自抗扰控制在陀螺电机稳速控制中的应用

关伊哲，冯明*
(北京科技大学 机械工程学院，北京 100083)

摘 要：陀螺电机的转速稳定性影响着陀螺仪性能，为提高永磁同步陀螺电机(PMSGM)的稳速控制精度，提出速度环采用自抗扰控制(ADRC)，滤波环节采用基于模型自适应滤波算法的方法。将ADRC应用于陀螺电机矢量控制系统中，其速度环替代传统比例积分(PI)控制，减小速度环输出波动，从而提高陀螺电机的稳速精度；采用基于模型自适应滤波算法代替传统低通滤波算法，无需相位延时补偿并使估计得到的反电动势波形更为平滑，提高转子位置角估计的准确性和稳速控制精度。因此，设计开发PMSGM矢量控制器，仿真及实验结果表明：所提方法相较于传统控制方法，有效地提高了陀螺电机的稳速控制精度。

关 键 词：永磁同步陀螺电机；矢量控制；自抗扰控制；自适应滤波；稳速精度
中图分类号：TM301.2
文献标志码：A **文章编号**：1001-5965(2024)01-0234-09

陀螺电机作为陀螺仪的核心元件，其稳速控制精度影响着陀螺仪的工作性能[1]。在永磁同步陀螺电机(permanent magnet synchronons gyro motor, PMSGM)矢量控制系统中，一般采用速度-电流双闭环结构，速度环输出的稳定性决定了电流环输入的稳定性，进而影响了电机的稳速控制精度。当稳速控制精度要求不高时，速度环通常采用比例积分(proportional integral, PI)控制，其结构简单，参数易调[2]。传统PI控制器是线性控制器，直接对误差取值进行线性加权和计算，存在取值波动较大、快速性与超调量矛盾等问题[3]。而实际应用中的PMSGM电机是非线性控制系统，当稳速控制精度要求较高时，传统PI控制的性能便不能满足需求。

针对PI控制的缺点，文献[4]提出了自抗扰控制(active disturbance rejection control, ADRC)技术，利用误差反馈控制率思想来消除实际信号与参考信号间的误差，实时补偿系统内外扰动，使系统稳态误差大大减小。由于其具有诸多优势，ADRC已被应用于不同的电机控制领域，文献[5]将ADRC应用于异步电机系统，设计了4个一阶ADRC器，提高了电机控制精度及鲁棒性，该方法主要适用于异步电机系统。文献[6]将ADRC应用于永磁同步电机(permanent magnet synchronons motor, PMSM)位置环，证实了该控制方法的优越性，不过该方法主要针对有位置传感器的电机。文献[7]将ADRC应用于无位置传感器控制中，提高了PMSM在低速区域的动态与稳态性能。文献[8]将ADRC应用于电动车用PMSM，提高了电机的动态响应性能，但并没有进行稳态性能分析。

在无位置传感器矢量控制系统中，转子位置角的估计准确性也影响着电机稳速精度。文献[9]提出采用低通滤波后反电动势作为滑膜观测器(sliding mode observation, SMO)反馈信号的方法，削弱了系统抖振，增加了转子位置角估计精度，但该方法存在反馈信号的幅值和相位发生变化的问题。文献[10-12]采用软件锁相环(phase locked loop,

收稿日期：2022-04-02；**录用日期**：2022-05-31；**网络出版时间**：2022-06-24 13:57
网络出版地址：link.cnki.net/urlid/11.2625.V.20220623.1109.002
基金项目：军委后勤保障部重点项目(BX219C014)
* **通信作者**. E-mail：mingfeng@me.ustb.edu.cn

引用格式：*关伊哲，冯明. 自抗扰控制在陀螺电机稳速控制中的应用[J]. 北京航空航天大学学报，2024，50(1)：234-242.*
GUAN Y Z, FENG M. Application of active disturbance rejection control in gyro motor steady speed control[J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50(1): 234-242 (in Chinese).

PLL)算法提高了转子位置角的估计精度，不过该算法也引入了低通滤波环节，存在相位延时补偿问题。

本文为了解决传统 PI 控制下速度环输出即电流环输入波动较大的问题，以及 SMO 中使用低通滤波器带来的相位延时等问题，提出在速度环采用 ADRC，滤波环节采用基于模型自适应滤波算法的方法。在速度环引入 ADRC，使得电机在稳定状态下速度环输出波动更小；并且在滑模观测器中构建反电动势观测方程，采用基于观测模型的自适应滤波算法代替传统低通滤波算法，避免了相位延时补偿并使滤波后的反电动势波形更为平滑、转子位置角的估计更为准确。最后，将本文方法应用于陀螺电机矢量控制系统中，使陀螺电机获得较高的稳速控制精度。

1　速度环自抗扰控制设计

ADRC 结构由三部分组成，分别为跟踪微分器（tracking differentiator, TD）、扩张状态观测器（extended state observer, ESO）及非线性状态误差反馈（nonlinear state error feedback, NLSEF）[13]。TD 对输入信号进行平滑处理，并得到其微分值[14]；ESO 对系统所受干扰进行预测[15]；NLSEF 环节在减小系统超调量的同时保证系统的响应速度，实现控制量的精确输出。其中，NLSEF 具有与 PI 控制相似的结构，是针对 PI 控制中线性组合进行的改进，使得控制精度更高。

ADRC 的结构框图如图 1 所示。其中，v为 ADRC 输入，u为 ADRC 输出的总控制器，y为被控系统输出，e 为跟踪误差，z 为估计值，b 为扰动补偿系数。

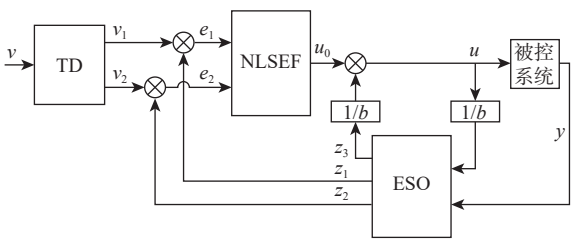


图 1　ADRC 结构框图

Fig. 1　ADRC structure block diagram

下面给出 ADRC 三部分各自的方程。

TD 的方程为

$$\begin{cases} v_1(k+1)=v_1(k)+hv_2(k) \\ v_2(k+1)=v_2(k)+hu_{\mathrm{h}} \\ u_{\mathrm{h}}=\mathrm{fhan}(v_1(k)-v,v_2(k),r,h_0) \end{cases} \tag{1}$$

式中：$v_1(k)$和$v_2(k)$分别为k时刻输入信号的跟踪信号和微分信号；h为控制系统离散化周期；u_{h}为最速控制下的最优控制量；$\mathrm{fhan}(v_1(k)-v,v_2(k),r,h_0)$为最速综合函数，满足式(2)～式(4)。

$$\mathrm{fhan}(x_1,x_2,r,h_0)=\begin{cases} -r\mathrm{sgn}(\alpha) & |\alpha|>d \\ -r\dfrac{\alpha}{d} & |\alpha|\leqslant d \end{cases} \tag{2}$$

$$\alpha=\begin{cases} x_2+\dfrac{\sqrt{d^2+8r|y|}-d}{2}\mathrm{sgn}(y) & |y|>d_0 \\ x_2+\dfrac{y}{h_0} & |y|\leqslant d_0 \end{cases} \tag{3}$$

$$\begin{cases} d=rh_0 \\ d_0=h_0d \\ y=x_1+h_0x_2 \end{cases} \tag{4}$$

式中：α为非线性因子；r为跟踪响应系数；h_0为滤波系数；$\mathrm{sgn}(\cdot)$为符号函数。

NLSEF 的方程为

$$\begin{cases} e_1=v_1(k)-z_1(k) \\ e_2=v_2(k)-z_2(k) \\ u_0=\beta_1f_{\mathrm{a}}(e_1,\alpha_1,\delta)+\beta_2f_{\mathrm{a}}(e_2,\alpha_2,\delta) \\ u(k)=u_0-\dfrac{z_3(k)}{b} \end{cases} \tag{5}$$

式中：e_1和e_2分别为目标跟踪误差和目标导数跟踪误差；β_1和β_2分别为e_1和e_2的非线性增益；α_1和α_2分别为e_1和e_2的非线性因子；u_0为 NLSEF 输出的控制量；$f_{\mathrm{a}}(\varepsilon,\alpha,\delta)$为非线性函数，表达式为

$$f_{\mathrm{a}}(\varepsilon,\alpha,\delta)=\begin{cases} |\varepsilon|^{\alpha}\mathrm{sgn}(\varepsilon) & |\varepsilon|>\delta \\ \dfrac{\varepsilon}{\delta^{1-\alpha}} & |\varepsilon|\leqslant\delta \end{cases} \tag{6}$$

式中：ε为估计误差；δ为非线性系数，其值大于 0。

ESO 的方程为

$$\begin{cases} \varepsilon=z_1-y(k) \\ z_1(k+1)=z_1(k)+h\left[z_2(k)-\beta_{01}\varepsilon\right] \\ z_2(k+1)=z_2(k)+h\left[z_3(k)-\beta_{02}f_{\mathrm{a}}(\varepsilon,\alpha_1,\delta)+bu(k)\right] \\ z_3(k+1)=z_3(k)-h\beta_{03}f_{\mathrm{a}}(\varepsilon,\alpha_2,\delta) \end{cases} \tag{7}$$

式中：z_1和z_2分别为y的估计值和估计值的微分信号；z_3为系统扰动的估计值；β_{01}、β_{02}、β_{03}为 ESO 的增益；u为 ADRC 输出的总控制量。

矢量控制系统总体框图如图 2 所示。将 ADRC 应用于陀螺电机矢量控制系统速度环，图 2 中的电

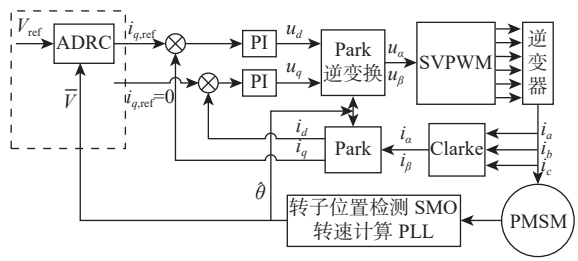


图 2　矢量控制系统总体框图

Fig. 2　General block diagram of vector control system

机转速设定值V_{ref}为ADRC的输入，电流环q轴设定指令$i_{q,\mathrm{ref}}$为ADRC的输出，SVPWM为空间矢量脉宽调制。

2 基于模型自适应滤波算法

在SMO估算转子位置角时，由于开关切换信号包含高频噪声，通常对估测得到的反电动势进行低通滤波，带来了相位延时的问题；同时低通滤波后的反电动势波形不够平滑，导致PLL算法求得的转子位置角不够精确。

因此，本文采用基于模型自适应滤波算法代替传统低通滤波算法，构建反电动势估计自适应模型，使SMO输出的反电动势曲线更为平滑，以达到滤波效果。该方法不仅无需相位延时补偿，滤波后的反电动势波形也更为平滑，有利于提高转子位置角估计精度。

对于三相对称的表贴式PMSM在α-β轴系下数学模型为

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} i_\alpha \\ i_\beta \end{bmatrix} = -\frac{1}{L_{\mathrm{z}}-L_{\mathrm{h}}}\left(R_{\mathrm{s}}\begin{bmatrix} i_\alpha \\ i_\beta \end{bmatrix} + \begin{bmatrix} u_\alpha \\ u_\beta \end{bmatrix} - \frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} \varphi_\alpha \\ \varphi_\beta \end{bmatrix}\right) \tag{8}$$

式中：i_α和i_β分别为电流在α和β轴的分量；R_{s}为定子相电阻；L_{z}为三相每相的自感；L_{h}为任意两相间互感；u_α和u_β分别为电压在α和β轴的分量；φ_α和φ_β分别为转子永磁体磁链在α和β轴的分量，可表示为

$$\begin{cases} \varphi_\alpha = \varphi_{\mathrm{f}}\cos\theta_{\mathrm{r}} \\ \varphi_\beta = \varphi_{\mathrm{f}}\sin\theta_{\mathrm{r}} \end{cases} \tag{9}$$

式中：φ_{f}为永磁体磁链；θ_{r}为转子位置角。

对式(9)求导可得

$$\begin{cases} \dfrac{\mathrm{d}\varphi_\alpha}{\mathrm{d}t} = -\varphi_{\mathrm{f}}\omega_{\mathrm{r}}\sin\theta_{\mathrm{r}} \\ \dfrac{\mathrm{d}\varphi_\beta}{\mathrm{d}t} = \varphi_{\mathrm{f}}\omega_{\mathrm{r}}\cos\theta_{\mathrm{r}} \end{cases} \tag{10}$$

式中：ω_{r}为转子旋转角速度。

磁链对时间的求导即为感应电动势，因此有

$$\begin{cases} e_\alpha = -\varphi_{\mathrm{f}}\omega_{\mathrm{r}}\sin\theta_{\mathrm{r}} \\ e_\beta = \varphi_{\mathrm{f}}\omega_{\mathrm{r}}\cos\theta_{\mathrm{r}} \end{cases} \tag{11}$$

式中：e_α和e_β分别为反电动势在α和β轴的分量。

对式(11)求导可得

$$\begin{cases} \dfrac{\mathrm{d}e_\alpha}{\mathrm{d}t} = -\dfrac{\mathrm{d}\omega_{\mathrm{r}}}{\mathrm{d}t}\varphi_{\mathrm{f}}\sin\theta_{\mathrm{r}} - \omega_{\mathrm{r}}e_\beta \\ \dfrac{\mathrm{d}e_\beta}{\mathrm{d}t} = \dfrac{\mathrm{d}\omega_{\mathrm{r}}}{\mathrm{d}t}\varphi_{\mathrm{f}}\cos\theta_{\mathrm{r}} + \omega_{\mathrm{r}}e_\alpha \end{cases} \tag{12}$$

在高速PMSGM控制中，通常一个控制周期的时间很短是1/10^{-6} s级，因此，将一个控制周期内的转速看作不变，故式(12)可简化得到反电动势数学模型为

$$\begin{cases} \dfrac{\mathrm{d}e_\alpha}{\mathrm{d}t} = -\omega_{\mathrm{r}}e_\beta \\ \dfrac{\mathrm{d}e_\beta}{\mathrm{d}t} = \omega_{\mathrm{r}}e_\alpha \end{cases} \tag{13}$$

建立反电动势估计自适应律为

$$\begin{cases} \dfrac{\mathrm{d}\hat{e}_\alpha}{\mathrm{d}t} = -\hat{\omega}_{\mathrm{r}}\hat{e}_\beta - l(\hat{e}_\alpha - e_\alpha) \\ \dfrac{\mathrm{d}\hat{e}_\beta}{\mathrm{d}t} = \hat{\omega}_{\mathrm{r}}\hat{e}_\alpha - l(\hat{e}_\beta - e_\beta) \\ \dfrac{\mathrm{d}\tilde{\omega}}{\mathrm{d}t} = \hat{e}_\beta\tilde{e}_\alpha - \hat{e}_\alpha\tilde{e}_\beta \end{cases} \tag{14}$$

式中：$\hat{e}_\alpha$和$\hat{e}_\beta$分别为α和β反电动势观测值；$\tilde{e}_\alpha$和$\tilde{e}_\beta$分别为α和β反电动势误差值；$\hat{\omega}_{\mathrm{r}}$为转速观测值；$\tilde{\omega}$为转速误差值；$l$为正常数，表示观测器增益。

为了证明所设计的自适应律稳定，将式(14)减去式(13)得到反电动势的误差方程为

$$\begin{cases} \dfrac{\mathrm{d}\tilde{e}_\alpha}{\mathrm{d}t} = -\tilde{\omega}_{\mathrm{r}}\hat{e}_\beta - \omega_{\mathrm{r}}\tilde{e}_\beta - l\tilde{e}_\alpha \\ \dfrac{\mathrm{d}\tilde{e}_\beta}{\mathrm{d}t} = \tilde{\omega}_{\mathrm{r}}\hat{e}_\alpha + \omega_{\mathrm{r}}\tilde{e}_\alpha - l\tilde{e}_\beta \end{cases} \tag{15}$$

再进行李雅普诺夫稳定性判断，李雅普诺夫函数V为

$$V = \frac{1}{2}\left(\tilde{e}_\alpha^2 + \tilde{e}_\beta^2 + \tilde{\omega}_{\mathrm{r}}^2\right) \tag{16}$$

$$\frac{\mathrm{d}V}{\mathrm{d}t} = -l\left(\tilde{e}_\alpha^2 + \tilde{e}_\beta^2\right) + \tilde{\omega}_{\mathrm{r}}\left[\frac{\mathrm{d}\tilde{\omega}_{\mathrm{r}}}{\mathrm{d}t} - \left(\hat{e}_\beta\tilde{e}_\alpha - \hat{e}_\alpha\tilde{e}_\beta\right)\right] \tag{17}$$

由于式(17)等号右边第1项恒负，且第2项为0，因此，该多项式是负定的，即构建的反电动势估计自适应律是稳定的。利用构建的观测方程对SMO模块观测得到的反电动势进行滤波，再由PLL模块求得转速与转子位置角，具体转子位置信息提取过程如图3所示。

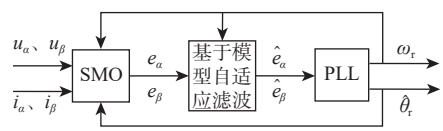


图3 转子位置信息提取过程

Fig. 3 Rotor position information extraction process

3 仿 真

为了验证本文方法在提高稳速控制精度方面的有效性，在MATLAB/Simulink中搭建了陀螺电机控制系统仿真模型，如图4所示。其中，n为转速，下标set为设定值，下标fed为反馈值，θ为转子位置角，V_{DC}为电机电压，n_θ为电机实际转子位置角，n_{SMO}和θ_{SMO}分别为SMO估计的电机转速和转子位置角，PWM_Out为脉宽调制输出，T_{m}为负载转

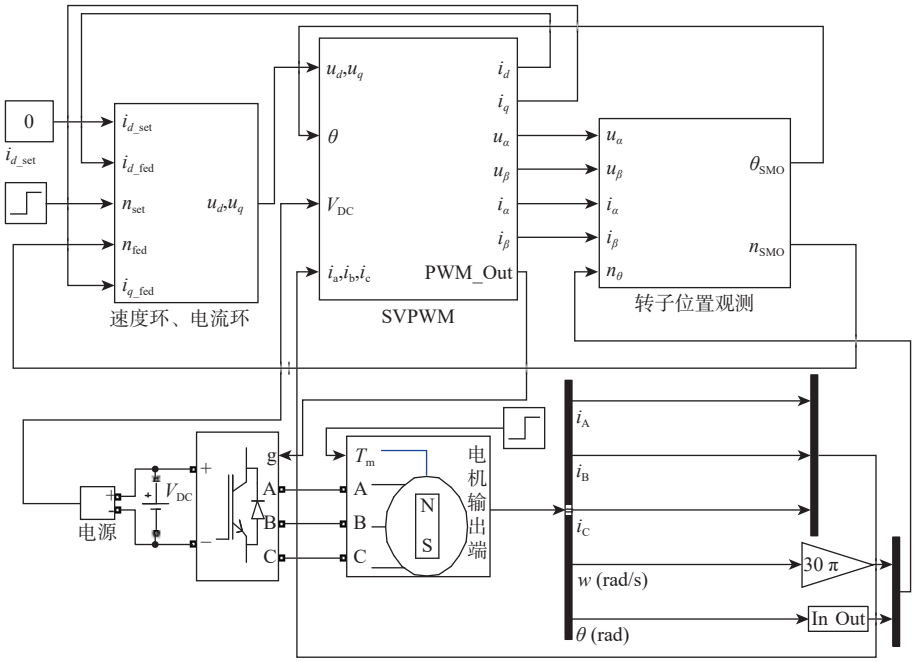


图 4　电机控制系统仿真模型

Fig. 4　Simulation model of motor control system

矩。分别在速度环模块中搭建了 ADRC 与 PI 控制模型进行仿真对比，在转子位置观测模块中搭建了基于模型自适应滤波算法与低通滤波算法模型进行仿真对比。其中电机模块参数与陀螺电机一致，如表 1 所示。

表 1　陀螺电机参数

Table 1　Gyro motor parameters

额定电压/v	相电阻/Ω	相电感/μH	转动惯量/(kg·m²)	反电动势常数/(V·s·rad⁻¹)	极对数/对
24	3.56	360	0.000 02	0.011 55	6

3.1　自抗扰控制与 PI 控制对比

设置目标转速为 6 000 r/min，除速度环分别采用 2 种控制方法以外，其余模块设置相同，ADRC 与 PI 控制对比的仿真模型如图 5 所示。

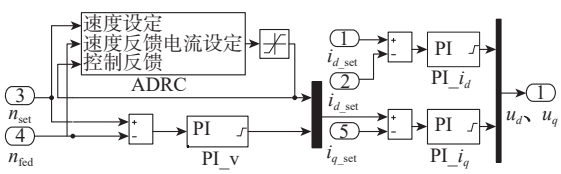


图 5　ADRC 与 PI 控制对比仿真模型

Fig. 5　Simulation model of ADRC and PI control comparison

本节介绍 ADRC 与 PI 控制的参数设定过程。首先，确定 PI 控制器参数，先进行比例环节与积分环节的粗调，使得电机可以在 0.07 s 内达到目标转速，接着，进行参数的细调，使仿真到达 0.4 s 后 PI 的稳态控制效果达到最优，最终确定控制系数 $k_p = 25$、$k_i = 0.3$。其次，确定 ADRC 的主要参数，为了能够直观且准确的对比出 2 种方法的控制效果，使 NLSEF 模块中β_1、β_2与k_p、k_i数值相同，即 2 种方法的控制量输出环节增益系数相同。随后，进行 TD 模块与 ESO 模块参数粗调，同样使得电机可以在 0.07 s 内达到目标转速，保证快速性一致。最后，进行 TD 模块与 ESO 模块参数细调，使仿真到达 0.4 s 后获得最优稳态控制效果，ADRC 器参数如表 2 所示。

表 2　自抗扰控制器参数

Table 2　ADRC parameters

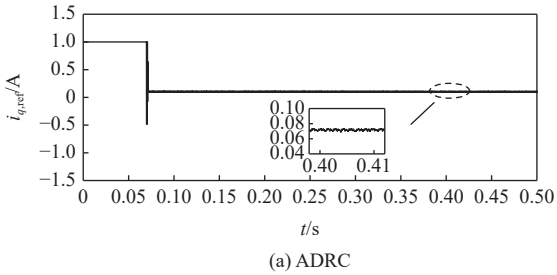
参数	数值	参数	数值
h	0.000 05	β_1	25
h_0	0.000 05	β_2	0.3
r	100 000	β_{01}	25
α_1	1.0	β_{02}	3 000
α_2	0.5	β_{03}	200

本文仅考虑电机处于稳定运行状态下的控制效果。电机稳定运行时，ADRC 与 PI 控制速度环输出波形如图 6 所示。由图可知，相比于 PI 控制，ADRC 下速度环输出波动更小。

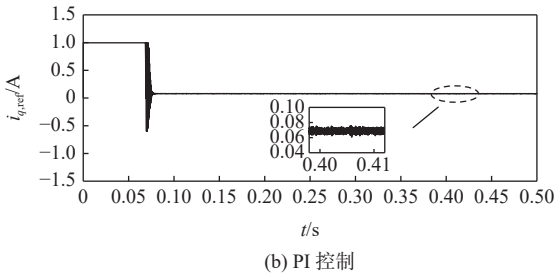
ADRC 与 PI 控制下电机稳定运行时转速波形如图 7 所示。由图可知，相比于 PI 控制，ADRC 下电机转速波动更小，且稳速精度更高。

3.2　基于模型自适应滤波算法与低通滤波算法对比

设置目标转速为 6 000 r/min，保证除滤波环节以外其余模块设置相同，基于模型自适应滤波算法



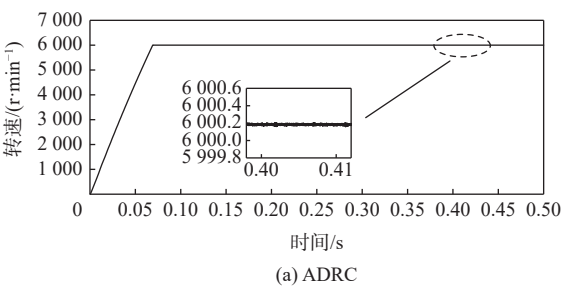
(a) ADRC



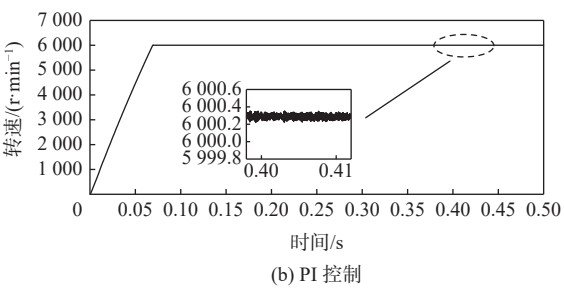
(b) PI 控制

图 6 不同控制方法下速度环输出波动

Fig. 6 Velocity loop output fluctuation under different control methods



(a) ADRC



(b) PI 控制

图 7 不同控制方法下电机转速波动

Fig. 7 Motor speed fluctuation under different control methods

仿真模型如图 8 所示，低通滤波算法仿真模型如图 9 所示，其中，K 为滤波系数，w 为电机角速度，w_c 为采用低通滤波算法后的角速度。

对比 2 种算法下反电动势滤波效果及转子位置角估计效果，如图 10 和图 11 所示。由图 10 可知，相比于低通滤波算法，经过基于模型自适应滤波算法后反电动势波形更为平滑且无相位延时产生。

2 种不同滤波算法下，PLL 模块输出的转子位置角与电机实际转子位置角对比情况如图 11 所示。由图可知，基于模型自适应滤波算法下转子位置角的估计更为准确。

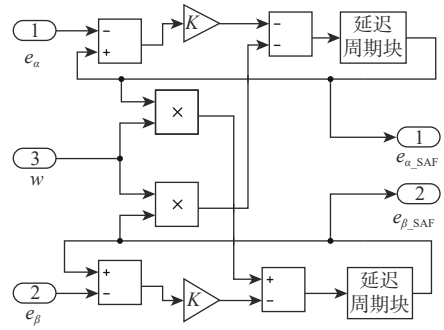


图 8 基于模型自适应滤波算法仿真模型

Fig. 8 Simulation model based on model adaptive filtering algorithm

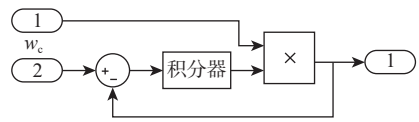


图 9 低通滤波算法仿真模型

Fig. 9 Simulation model of low pass filtering algorithm

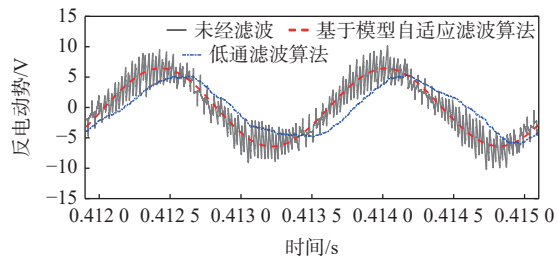


图 10 反电动势波形对比

Fig. 10 Comparison of inverse emf waveforms

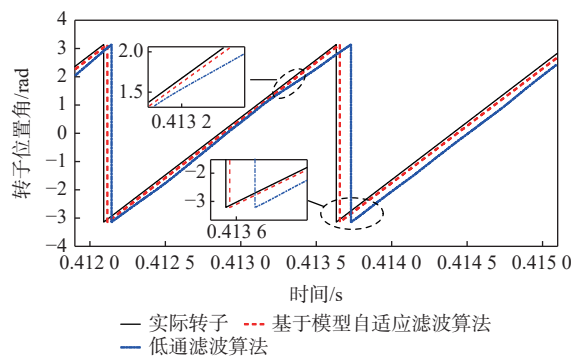


图 11 转子位置角对比

Fig. 11 Comparison of rotor position angles

4 实验分析

为验证基于模型自适应滤波算法与 ADRC 在提高稳速控制精度方面的有效性，搭建转速测试实验平台。实验系统框图如图 12 所示，实验装置如图 13 所示，实验中使用到的设备包括：B&K Type2981 型光电传感器，其电压为 20 V，电流为 3～20 mA；NI6215 型采集卡，其频率为 10 MHz；DLM2054 型示波器，其通道为 4，频率为 500 MHz。实验中所

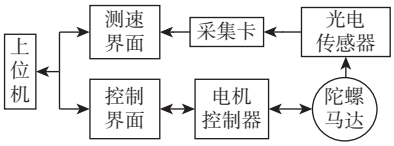
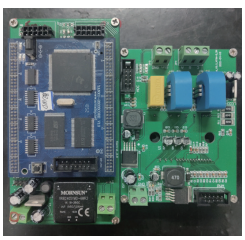
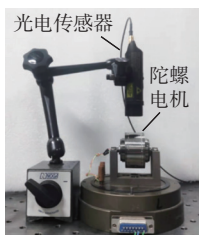


图 12　实验系统框图

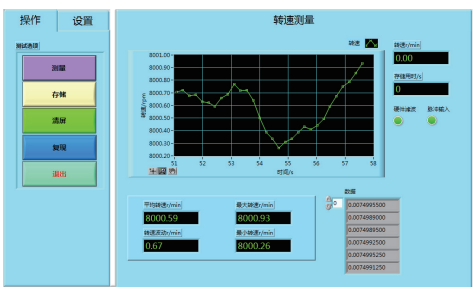
Fig. 12　Experimental system block diagram



(a) 电机控制器



(b) 永磁同步陀螺电机



(c) 测速界面

图 13　实验装置

Fig. 13　Experimental devices

用转速测试系统精度为 $1/10^5$ 级, 可以满足对 $1/10^4$ 级转速精度的测量。

4.1　自抗扰控制与 PI 控制对比实验

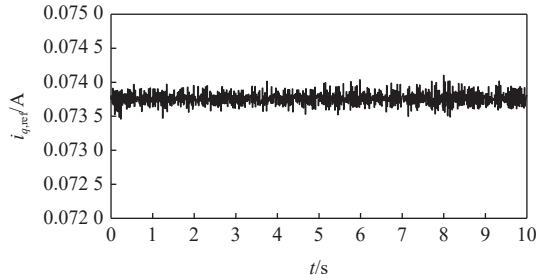
为验证 ADRC 在提高稳速控制精度方面的有效性, 采集 2 种控制方法下速度环输出值、电机转速波动情况进行对比分析。实验过程中, 设定电机目标转速为 6 000 r/min, 控制器参数选取原则与 3.1 节仿真参数选取原则一致, 由控制器记录速度环输出值, 转速测试系统记录电机转速值。

不同控制方法下速度环输出波动和电机转速波形分别如图 14 和图 15 所示。由图可知, ADRC 下速度环输出波动更小, 电机转速波动更小且稳速精度更高。

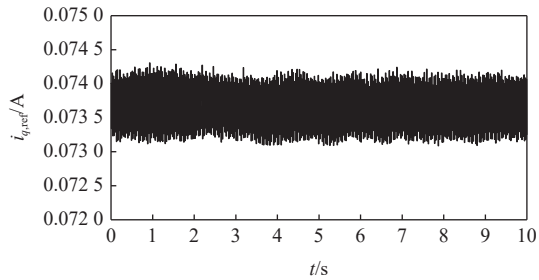
4.2　基于模型自适应滤波算法与低通滤波算法对比实验

为验证基于模型自适应滤波算法在提升稳速控制精度方面的有效性, 采集 2 种滤波算法下反电动势波形、转子位置角波形及电机转速波动情况进行对比分析。实验过程中, 设定目标转速为 6 000 r/min, 由控制器记录滤波后反电动势值与转子位置角值, 转速测试系统记录电机转速值。

滤波后反电动势波形对比如图 16 所示。转子位置角波形对比如图 17 所示。采用 2 种滤波算法



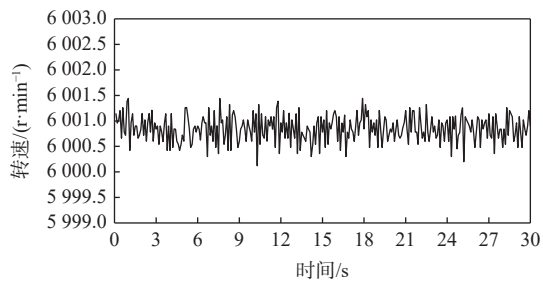
(a) ADRC



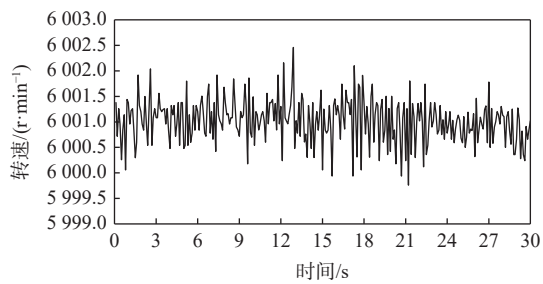
(b) PI 控制

图 14　不同控制方法下速度环输出波动

Fig. 14　Velocity loop output fluctuation under different control methods



(a) ADRC



(b) PI 控制

图 15　不同控制方法下电机转速波形

Fig. 15　Motor speed waveform under different control methods

时电机转速波形如图 18 所示。由图 16～图 18 可知, 采用基于模型自适应滤波算法可以使反电动势波形更为平滑, 估算的转子位置角抖动更小, 电机转速波动更小。

4.3　稳速控制实验

对比本文方法与传统方法的稳速控制效果。传统方法为速度环采用 PI 控制、滤波模块采用低通滤波算法的方法。

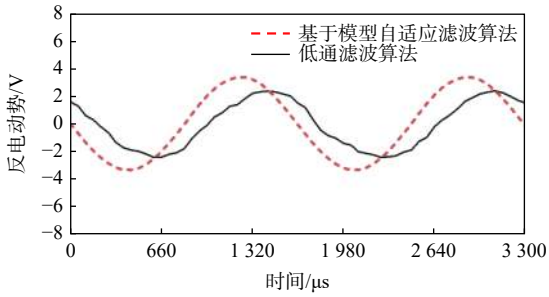


图 16 滤波后反电动势波形

Fig. 16 Filtered back emf waveform

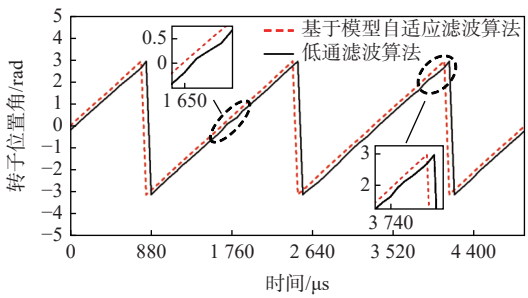
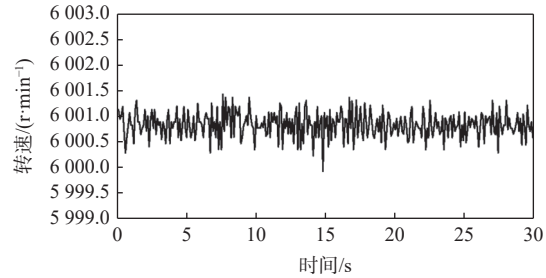


图 17 转子位置角波形

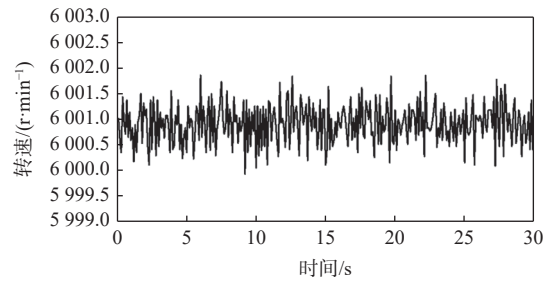
Fig. 17 Rotor position angle waveform

电机运行转速分别为 6 000, 8 000, 10 000 r/min 时进行实验，通过光电转速传感器采集 30 s 内不同控制方式下电机转速信息。为了更好体现稳速控制精度，定义转速波动率与稳速精度 2 个指标，计算式分别为

$$\sigma = \frac{n_{\max} - n_{\min}}{2n_{\text{avg}}} \tag{18}$$



(a) 基于模型自适应滤波算法



(b) 低通滤波下算法

图 18 不同滤波算法下电机转速波形

Fig. 18 Motor speed waveform under different filtering algorithms

$$\eta = \frac{|n_{\text{avg}} - n_{\text{set}}|}{n_{\text{set}}} \times 100\% \tag{19}$$

式中：σ为转速波动率；η为稳速精度；$n_{\max}$、$n_{\min}$和n_{avg}分别为 30 s 内转速最大值、最小值和平均值；n_{set}为转速设定值即目标转速。

转速测试计算结果如表 3 所示。

表 3 转速测试结果对比

Table 3 Comparison of speed test results

转速/(r·min⁻¹)	$n_{\max}$		$n_{\min}$		n_{avg}		σ/‰		η/‰	
	本文方法	传统方法	本文方法	传统方法	本文方法	传统方法	本文方法	传统方法	本文方法	传统方法
6 000	6 001.62	6 001.62	6 001.62	5 999.76	6 000.84	6 001.08	0.14	0.22	0.14	0.18
8 000	8 002.45	8 003.41	7 999.79	7 999.04	8 001.24	8 001.48	0.17	0.27	0.16	0.19
10 000	10 003.50	10 005.67	9 999.83	9 999.67	10 001.68	10 002.38	0.18	0.3	0.17	0.24

由表 3 可知，陀螺电机在不同转速下，采用本文方法时，电机的转速波动率与稳速精度均小于传统方法，有效的提高了稳速控制精度且均达到了 $1/10^4$ 级。

5 结 论

1）本文在速度环引入了 ADRC 代替 PI 控制，降低了速度环的输出波动，提高了稳速控制精度。

2）本文在滑模观测器中引入了基于模型自适应滤波算法代替低通滤波算法，无需相位延时补偿并增加了转子位置角估计的准确性，提高了稳速控制精度。

3）基于上述 2 种算法开发了永磁同步陀螺电机矢量控制系统，通过仿真及实验验证了本文方法相较于传统方法可有效地提高陀螺电机的稳速控制精度。

参考文献（References）

[1] 赵美蓉, 张鑫鑫, 黄银国, 等. 陀螺电机转速稳定性对寻北影响及其稳速控制[J]. 纳米技术与精密工程, 2016, 14(2): 81-86.
ZHAO M R, ZHANG X X, HUANG Y G, et al. Effect of stability of rotary speed of gyro motor on north seeking and its speed control[J]. Nanotechnology and Precision Engineering, 2016, 14(2): 81-86(in Chinese).

[2] SU Y X, ZHENG C H, DUAN B Y. Automatic disturbances rejec-

tion controller for precise motion control of permanent-magnet synchronous motors[J]. IEEE Transactions on Industrial Electronics, 2005, 52(3): 814-823.

[3] 符慧. 永磁同步电机转速环 PI 控制策略研究 [D]. 南京: 南京航空航天大学, 2017: 21-22.
FU H. Research on PI control strategy of permanent magnet synchronous motor speed loop [D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2017: 21-22(in Chinese).

[4] 韩京清. 自抗扰控制器及其应用[J]. 控制与决策, 1998, 13(1): 19-23.
HAN J Q. Active disturbance rejection controller and its application[J]. Control and Decision, 1998, 13(1): 19-23(in Chinese).

[5] 贺虎成, 孙磊, 张玉峰, 等. 基于矢量控制的异步电机自抗扰控制[J]. 电机与控制学报, 2019, 23(4): 120-125.
HE H C, SUN L, ZHANG Y F, et al. Asynchronous motor active disturbance rejection control based on vector control[J]. Electric Machines and Control, 2019, 23(4): 120-125(in Chinese).

[6] GUO B L, BACHA S, ALAMIR M. A review on adrc based pmsm control designs [C]//IECON 2017-43rd Annual Conference of the IEEE Industrial Electronics Society. Piscataway: IEEE Press, 2017: 1747-1753.

[7] HAN Y, LI H, LI W T. Research on PMSM sensorless system based on ADRC-PBC strategy [C]//2016 Chinese Control and Decision Conference. Piscataway: IEEE Press, 2016: 3186-3191.

[8] 刘春光, 张征, 陈路明, 等. 电动车用永磁同步电机转速自抗扰控制[J]. 火力与指挥控制, 2020, 45(11): 71-77.
LIU C G, ZHANG Z, CHEN L M, et al. Rotating speed active disturbance rejection control for permanent magnet synchronous motor in electric vehicles[J]. Fire Control & Command Control, 2020, 45(11): 71-77(in Chinese).

[9] 孙佃升. 改进滑模观测器的永磁同步电机无传感器控制[J]. 电力电子技术, 2018, 52(2): 54-56.
SUN D S. Sensorless control of PMSM using improved sliding mode observer[J]. Power Electronics, 2018, 52(2): 54-56(in Chinese).

[10] HUANG J C, LI H S, XU Q H, et al. Sensorless vector control of PMSM using sliding mode observer and fractional-order phase-locked loop[C]//Proceedings of the 31st Chinese Control Conference. Piscataway: IEEE Press, 2012: 4513-4518.

[11] 李冉, 龙雪涛, 陈辉. 基于锁相环的永磁同步电机无传感器控制[J]. 电气传动, 2013, 43(8): 8-12.
LI R, LONG X T, CHEN H. Sensorless control of permanent magnet synchronous motor based on phase-locked loop[J]. Electric Drive, 2013, 43(8): 8-12(in Chinese).

[12] 郑昌陆, 胡月波, 耿宇翔. 基于锁相环原理的永磁同步电机无传感器控制方法[J]. 电机与控制应用, 2015, 42(3): 24-28.
ZHENG C L, HU Y B, GENG Y X. Sensorless control method of permanent magnet synchronous motor based on principle of phase-locked loop[J]. Electric Machines & Control Application, 2015, 42(3): 24-28(in Chinese).

[13] 韩京清. 从 PID 技术到“自抗扰控制”技术[J]. 控制工程, 2002, 9(3): 13-18.
HAN J Q. From PID technique to active disturbances rejection control technique[J]. Basic Automation, 2002, 9(3): 13-18(in Chinese).

[14] 武利强, 林浩, 韩京清. 跟踪微分器滤波性能研究[J]. 系统仿真学报, 2004, 16(4): 651-652.
WU L Q, LIN H, HAN J Q. Study of tracking differentiator on filtering[J]. Acta Simulata Systematica Sinica, 2004, 16(4): 651-652(in Chinese).

[15] REN C, LI X H, YANG X B, et al. Extended state observer-based sliding mode control of an omnidirectional mobile robot with friction compensation[J]. IEEE Transactions on Industrial Electronics, 2019, 66(12): 9480-9489.

Application of active disturbance rejection control in gyro motor steady speed control

GUAN Yizhe，FENG Ming*

(School of Mechanical Engineering，University of Science and Technology Beijing，Beijing 100083，China)

Abstract：The performance of the gyroscope is impacted by the steadiness of the gyro motor's rotating speed. To improve the stability control precision of the permanent magnet synchronons gyro motor (PMSGM), propose the method use of the active disturbance rejection control (ADRC) in the speed loop and a model-based adptive filtering algorithm in the filtering process. the ADRC was proposed to replace the traditional proportional integral (PI) control in the speed loop of the gyro motor vector control system to reduce the output jitter of the speed loop and improve the stability precision of gyro motor. While this was going on, a model-based adaptive filtering algorithm was put out to take the place of the conventional low-pass filtering algorithm. This algorithm does away with phase delay correction and produces a smoother, more precise waveform for the predicted back electromotive force. So the accuracy of rotor position angle estimation and steady speed control were improved. Therefore, the PMSGM vector controller was designed. The simulation and experimental results showed that the proposed method was better than traditional control methods, and it could effectively improve the speed control precision of gyro motor.

Keywords：permanent magnet synchronons gyro motor; vector control; active disturbance rejection control; adaptive filtering; steady speed precision

Received：2022-04-02；**Accepted**：2022-05-31；**Published Online**：2022-06-24 13：57

URL：link.cnki.net/urlid/11.2625.V.20220623.1109.002

Foundation item：Key Projects of the Logistics Support Department of the Central Military Commission (BX219C014)

* **Corresponding author**. E-mail：mingfeng@me.ustb.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0219

空间可展开天线铰链热变形分析与试验验证

杨智杰[1]，张从发[2]，赵瑞杰[1]，王春洁[3]，赵军鹏[1,*]

(1. 北京航空航天大学 机械工程及自动化学院，北京 100191；2. 北京空间飞行器总体设计部，北京 100094；3. 北京航空航天大学虚拟现实技术与系统国家重点实验室，北京 100191)

摘　　要：在轨热变形是影响空间可展开天线精度的关键因素之一，因此，空间可展开天线结构的在轨热变形分析具有十分重要的意义。随着空间探测任务要求的不断提高，铰链热变形对空间可展开天线指向精度的影响已经不可忽略。以某空间可展开天线的根部铰链和臂间铰链为研究对象，建立其有限元模型，计算铰链主弯方向和侧弯方向在温度每变化 10 ℃ 时相对角度的变化。基于高低温箱和电子经纬仪构建铰链精度测试系统，开展铰链热变形试验。试验结果表明：根部铰链和臂间铰链在主弯方向和侧弯方向呈现纯结构热变形特性，仿真结果与试验分析结果基本一致，验证有限元模型和试验分析方法的正确性。根据铰链主弯和侧弯方向角度相对变化，评估热变形对铰链变形的影响。该建模和试验分析方法可为同类型空间可展开天线设计、分析和优化及类似可展开机构精度影响因素的分析提供参考。

关　键　词：可展开天线；有限元分析；铰链；热变形；高低温环境

中图分类号：V443.4

文献标志码：A　　**文章编号**：1001-5965(2024)01-0243-07

随着空间科学等研究的不断深入，空间探测任务要求的不断提高，对伸展臂[1-3]、卫星天线[4-6]等可展开机构的精度要求也越来越高，对其精度的研究已成为当前的研究热点之一。由于空间可展开天线的尺寸一般较大，对热载荷的敏感性高、精度的稳定性差，在空间可展开天线精度的各种影响因素中，在轨热变形的影响至关重要[7]。

目前，空间可展开天线热变形研究方面，国内外学者已经做了大量工作。朱敏波等[8]运用有限元方法，进行了某星载大型展开天线温度场和热变形的数值计算，并研究了基本结构参数对天线形状精度的影响。Mahaney 和 Thornton [9]分析了四面体桁架天线的热变形，并研究了交叉和平行杆件相互遮挡对杆件内温度分布的影响。杨仲[10]通过控制有限元网格质量，利用统一的有限元模型进行了热变形分析和电磁场辐射分析。武聪魁等[11]基于弹性热力学理论及有限元理论，研究了非均匀温度场对计及与不计及金属铰链的环形可展天线形面精度和张力比的影响规律。张惠峰和关富玲[12]采用热-结构耦合分析方法，计算了不同时刻下可展桁架天线在轨热温度场和热变形。刘飞[13]基于有限元理论，提出了一种天线反射面单元划分技术，并开发了用于补偿天线反射面温度变形的软件。赵军忠等[14]建立了某星载天线的有限元模型，研究了由于天线热变形导致的波束指向变化及偏差。吴楠[15]建立了碳纤维夹层天线面板有限元模型，以天线面板的热变形残差为优化目标，进行了天线结构参数优化。

以上研究中，忽略了铰链热变形对天线变形的影响，无法满足一些特殊构型的天线热变形分

收稿日期：2022-04-06；**录用日期**：2022-05-14；**网络出版时间**：2022-06-07 13:07
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220606.1804.005.html
基金项目：国家自然科学基金 (51635002)
* **通信作者**. E-mail：zhaojunpeng@buaa.edu.cn

引用格式：杨智杰，张从发，赵瑞杰，等. 空间可展开天线铰链热变形分析与试验验证 [J]. 北京航空航天大学学报，2024，50（1）：243-249.
YANG Z J，ZHANG C F，ZHAO R J，et al. Thermal deformation analysis and experimental verification of spatial deployable antenna hinge[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50（1）：243-249（in Chinese）.

析的准确性要求。故本文基于有限元理论，以某空间可展开天线铰链为研究对象，在分析空间可展开天线及其铰链组成、工作原理和结构特点的基础上，研究其有限元建模方法并建立相应的有限元模型；基于高低温箱和电子经纬仪建立铰链精度测试系统，运用仿真分析和试验验证相结合的方法[16]，研究铰链热变形特性，并根据铰链在主弯方向和侧弯方向温度每变化 10 ℃ 时相对角度的变化，评估铰链热变形对根部铰链和臂间铰链变形的影响。

1 某空间可展开天线工作原理

空间可展开天线的主要功能是获取与传递信息，其指向精度对接收与发送信号、数据传输和目标跟踪等都有重要影响。图 1 为某空间可展开天线，主要由 1 个根部铰链、2 个臂间铰链、3 根复合材料臂杆、电缆、电缆托、绳索联动装置和压紧释放装置等组成。其中，臂间铰链、根部铰链分别通过螺纹与复合材料臂杆连接，绳索联动装置保证空间可展开天线展开的同步性。模型中坐标系以空间可展开天线压紧释放装置方向为 x 轴，沿臂杆方向为 y 轴，其指向展开方向，z 轴与 x、y 轴成右手直角坐标系，为铰链转轴方向。

当卫星处于发射阶段时，空间可展开天线呈收拢状态，压紧释放装置通过压紧座，将复合材料臂杆压紧，如图 1(a)所示；当卫星入轨后，根据展开指令要求，压紧释放装置解锁，天线铰链在联动绳索的作用下同步展开，天线臂杆向外伸展，如图 1(b)所示；当卫星处于在轨运行阶段时，空间可展开天线完全展开到位，最终通过铰链中的锁定杆锁定，保证锁定后的铰链不会发生回弹，如图 1(c)所示。

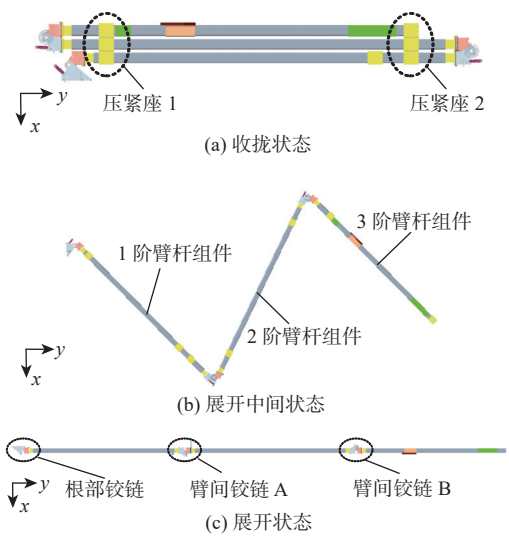


图 1 某空间可展开天线

Fig. 1 A spatially deployable antenna

2 天线铰链热变形分析

2.1 有限元建模方法

根部铰链和臂间铰链主要包含公铰、母铰、涡卷弹簧、转轴和锁定杆等部分，如图 2 所示。为减少因材料热膨胀系数不匹配产生的热变形，公铰和母铰均采用铝合金材料。锁定杆采用钛合金材料，以满足锁定瞬时的冲击承载要求。公铰和母铰套在转轴上，且可以相对于转轴转动。在接收到天线展开指令后，铰链在绳索联动装置的作用下，绕转轴同步展开，当天线展开到位后，公铰螺栓面和母铰接触，锁定杆锁定，铰链完全锁紧。

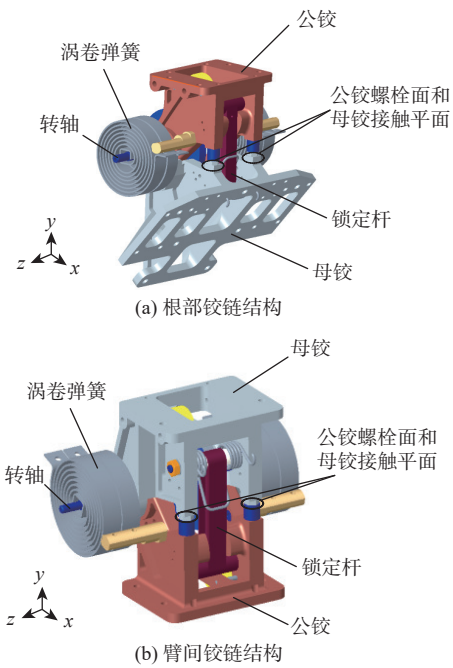


图 2 根部铰链和臂间铰链结构

Fig. 2 Root hinge structure and interarm hinge structure

为了进行铰链热变形分析，建立空间可展开天线有限元模型，如图 3 所示。

复合材料臂杆使用壳单元模拟，根部铰链和臂间铰链采用实体单元进行离散，且分别与复合材料臂杆使用壳-实体耦合方法连接，电缆、电缆托和压紧释放装置等效为质量点。锁定杆与公铰接触面之间建立接触，公铰和母铰的螺栓接触面之间建立接触。根据实践经验，接触属性中摩擦系数取

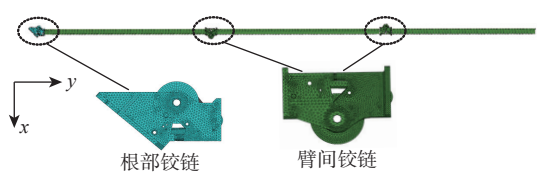


图 3 某空间可展开天线有限元模型

Fig. 3 Finite element model of a space deployable antenna

0.3。公铰与转轴、母铰与转轴之间的转动采用弹簧阻尼单元模拟, 单元刚度分别取铰链模型整体刚度的一半, 铰链整体刚度参数如表 1 所示, 其中 R_x、R_y、R_z 分别为铰链绕 x、y、z 轴的转动刚度。锁定杆与母铰之间的转动及联动轮的转动采用连接器单元模拟, 单元刚度取 10 Nmm/rad。

表 1　铰链刚度参数

Table 1　Hinge stiffness parameters

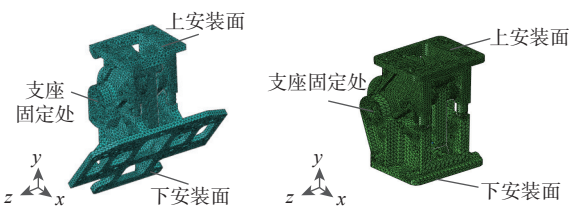
关节	R_x/(10^3 Nm·rad^{-1})	R_y/(10^3 Nm·rad^{-1})	R_z/(10^3 Nm·rad^{-1})	x向平移刚度/(10^6 N·m^{-1})	y向平移刚度/(10^5 N·m^{-1})	z向平移刚度/(10^5 N·m^{-1})
根部铰链	5	5	5	1	5	5
臂间铰链	5	5	5	1	5	5

2.2　铰链热变形分析

基于本文铰链有限元模型, 对根部铰链和臂间铰链进行热变形分析, 计算铰链在热载荷作用下的变形情况。

在铰链热变形分析中, 边界条件为在铰链支座固定处施加固定约束, 如图 4 所示。

空间可展开天线铰链在高低温环境中, 各部分由于升温或降温的不均匀, 会产生不同程度的热膨胀, 导致铰链热变形和热应力的产生, 从而对空间可展开天线的指向精度造成影响。



(a) 根部铰链　　(b) 臂间铰链

图 4　铰链边界条件设置

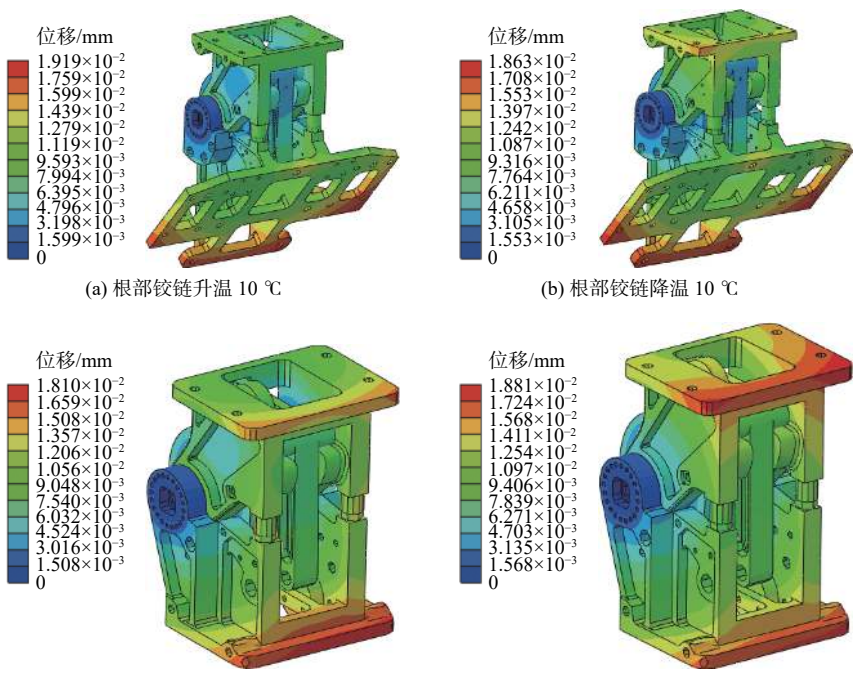
Fig. 4　Hinge boundary condition setting

假设初始温度为 25 ℃, 分别计算铰链每升、降温 10 ℃ 时的热变形, 铰链热变形云图如图 5 所示。由图可知, 在热载荷作用下, 根部铰链最大热变形分别为 0.019 19, 0.018 63 mm, 臂间铰链最大热变形分别为 0.018 10, 0.018 81 mm。升高和降低相同温度后, 铰链热变形量有所差别, 主要是由于接触非线性等因素所导致。

2.3　热变形后铰链指向计算

铰链主弯方向和侧弯方向的热变形对空间可展开天线指向精度影响最为显著。根部铰链和臂间铰链主弯方向变形即为安装面坐标系绕 z 轴的转动, 侧弯方向变形即为安装面坐标系绕 x 轴的转动, 铰链安装面坐标系转角示意如图 6 所示。

利用有限元模型计算出升、降温 10 ℃ 铰链热变形结果, 通过最小二乘法分别对变形后铰链上下安装面的所有单元节点进行拟合[17], 计算安装面法线向量。再根据铰链安装面法线向量, 计算升、降温 10 ℃ 铰链安装面主弯和侧弯方向转角, 计算式



(a) 根部铰链升温 10 ℃　　(b) 根部铰链降温 10 ℃

(c) 臂间铰链升温 10 ℃　　(d) 臂间铰链降温 10 ℃

图 5　铰链热变形云图

Fig. 5　Hinge thermal deformation cloud image

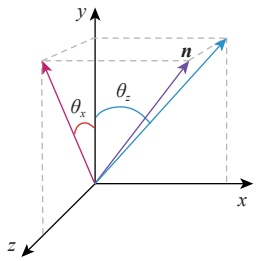


图6 法线向量转角示意图

Fig. 6 Schematic diagram of normal vector angle

分别为

$$\cos\theta_z=\frac{n_y}{\sqrt{{n_x}^2+{n_y}^2}} \tag{1}$$

$$\cos\theta_x=\frac{n_y}{\sqrt{{n_y}^2+{n_z}^2}} \tag{2}$$

式中：θ_z、θ_x分别为升温10℃或降温10℃铰链安装面主弯方向、侧弯方向转角，rad；$\boldsymbol{n}$为铰链安装面法线向量，$\boldsymbol{n}=[\ n_x\quad n_y\quad n_z\ ]^{\mathrm{T}}$。

计算铰链上下两安装面之间主弯和侧弯方向在温度每变化10℃时相对角度的变化值，计算式分别为

$$\alpha=\frac{(\theta_{z\mathrm{UH}}-\theta_{z\mathrm{DH}})+(\theta_{z\mathrm{UC}}-\theta_{z\mathrm{DC}})}{2} \tag{3}$$

$$\beta=\frac{(\theta_{x\mathrm{UH}}-\theta_{x\mathrm{DH}})+(\theta_{x\mathrm{UC}}-\theta_{x\mathrm{DC}})}{2} \tag{4}$$

式中：α和β分别为10 ℃温度变化导致的铰链两安装面之间主弯和侧弯方向相对角度变化值；$\theta_{z\mathrm{UH}}$、$\theta_{z\mathrm{DH}}$分别为升温10 ℃时铰链上、下安装面主弯方向转角；$\theta_{z\mathrm{UC}}$、$\theta_{z\mathrm{DC}}$分别为降温10 ℃时铰链上、下安装面主弯方向转角；$\theta_{x\mathrm{UH}}$、$\theta_{x\mathrm{DH}}$分别为升温10 ℃时铰链上、下安装面侧弯方向转角；$\theta_{x\mathrm{UC}}$、$\theta_{x\mathrm{DC}}$分别为降温10 ℃时铰链上、下安装面侧弯方向转角。

根据式(3)和式(4)，可解得根部铰链两安装面之间主弯方向和侧弯方向在温度每变化10℃时相对角度的变化值分别为0.005 4°和0.004 1°，臂间铰链两安装面之间主弯方向、侧弯方向在温度每变化10℃时相对角度的变化值分别为0.004 4°和0.001 0°。

3 天线铰链热变形试验验证

3.1 铰链热变形试验

为获得铰链在轨时的实际热变形特性与结果，基于高低温箱和电子经纬仪构建了铰链精度测试系统，如图7所示。该系统主要由1个高低温箱和4个电子经纬仪组成，电子经纬仪两两互瞄，透过高低温箱玻璃对箱内立方镜进行准直测量，记录偏航角和俯仰角，经准直解算后即可得到立方镜的坐标系。

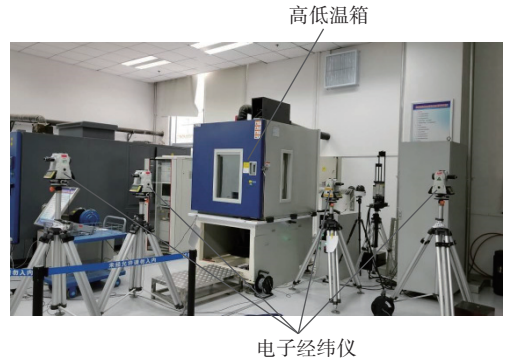


图7 铰链精度测试系统

Fig. 7 Hinge accuracy test system

高低温箱中单个温度循环试验流程，如图8所示。

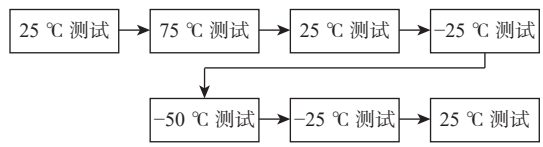


图8 单个温度循环试验流程

Fig. 8 Flow chart of single temperature cycle test

试验过程中，根部铰链和臂间铰链安装状态为：锁定状态下铰链的转轴一端通过螺钉固定在底部支撑座上，另一端处于自由状态，将公铰转接板和母铰转接板分别与公铰和母铰法兰面通过螺栓连接，在2个转接板上分别可靠地固定1个立方镜。臂间铰链试验安装状态如图9所示。此种安装状态，有效控制了地面重力对铰链变形结果的影响。否则，因在公铰、母铰上附加了转接板，可能给试验结果带来影响。

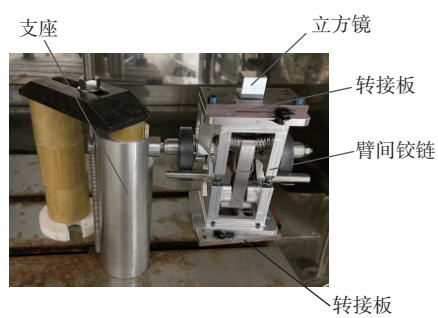


图9 臂间铰链试验安装状态

Fig. 9 Experiment installation state of interarm hinge

按照温度循环试验流程进行3个循环测试，在每个温度循环测试过程中，待铰链温度测点到达目标温度后，高低温箱内温度稳定并保温30 min后进行精度测量，在每个目标温度处测量并计算2个立方镜坐标系的相对位置关系。

根部铰链主弯方向相对角度变化情况，如图10所示。由图可知，3个循环中铰链2个安装面之间相对角度变化趋势一致，随着温度升高而变小，在

最高温度为 75 ℃ 处达到最小为 359.690 7°; 随着温度降低而增大, 在最低温度为−50 ℃ 处达到最大为 359.760 7°。说明根部铰链在主弯方向呈现纯结构热变形特性, 故可计算得到 2 个安装面之间在温度每变化 10 ℃ 时相对角度的变化值为 0.004 2°。

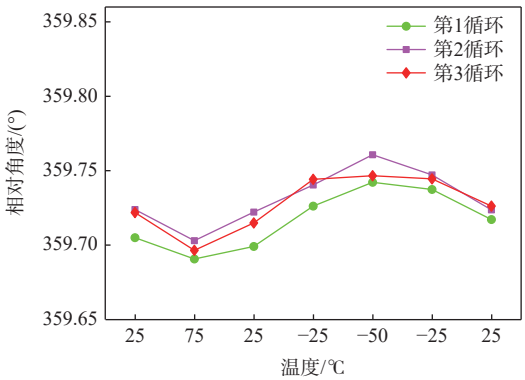


图 10　根部铰链主弯方向相对角度变化

Fig. 10　Relative angle changes of main bend direction of root hinge

根部铰链侧弯方向相对角度变化情况, 如图 11 所示。由图可知, 经历 1 个高低温循环之后, 第 2 和第 3 循环中铰链 2 个安装面之间相对角度变化趋势一致, 随着温度升高而变小, 在最高温度为 75 ℃ 处达到最小为 359.020 6°; 随着温度降低而增大, 在最低温度−50 ℃ 处达到最大, 为 359.065 9°。说明根部铰链在侧弯方向呈现一定的结构热变形特性, 故以第 2 和第 3 循环为准, 计算得到 2 个安装面之间在温度每变化 10 ℃ 时相对角度的变化值为 0.002 6°。

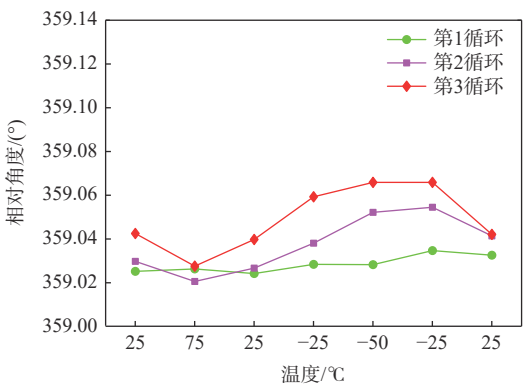


图 11　根部铰链侧弯方向相对角度变化

Fig. 11　Relative angle changes of lateral bend direction of root hinge

臂间铰链主弯方向相对角度变化情况, 如图 12 所示, 已剔除第 1 循环中温度为−50 ℃ 时由于测量失误出现的异常点。由图 12 可知, 3 个循环中铰链 2 个安装面之间相对角度变化趋势一致, 随着温度升高而变小, 在最高温度 75 ℃ 处达到最小为 359.835 7°; 随着温度降低而增大, 在最低温度−50 ℃ 处达到最大为 359.908 4°。说明臂间铰链在主弯方向呈现纯结构热变形特性, 故可计算得到 2 个安装面之间在温度每变化 10 ℃ 时相对角度的变化值为 0.004 6°。

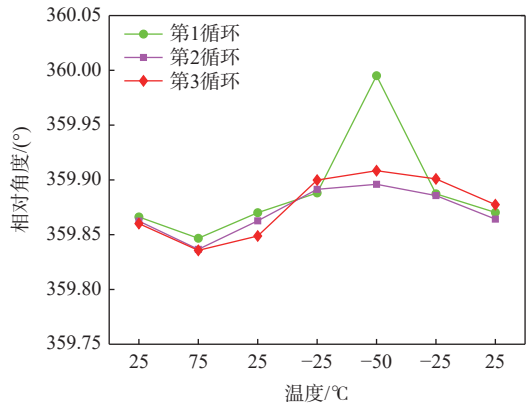


图 12　臂间铰链主弯方向相对角度变化

Fig. 12　Relative angle changes of main bend direction of interarm hinge

臂间铰链侧弯方向相对角度变化情况, 如图 13 所示, 已剔除第 1 循环中温度为−50 ℃ 时由于测量失误出现的异常点。由图 13 可知, 臂间铰链在侧弯方向也呈现一定的结构热变形特性, 故以第 2 和第 3 循环为准, 计算得到 2 个安装面之间在温度每变化 10 ℃ 时相对角度的变化值为 0.000 6°。

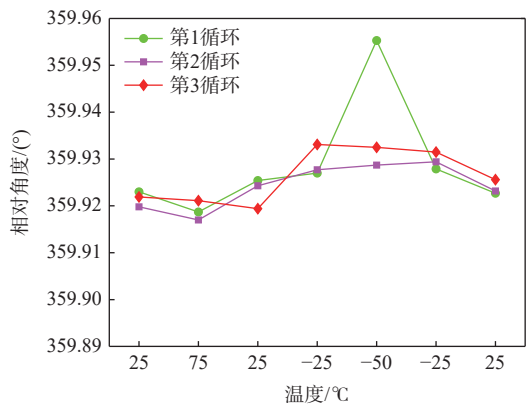


图 13　臂间铰链侧弯方向相对角度变化

Fig. 13　Relative angle changes of lateral bend direction of interarm hinge

由铰链热变形试验可知, 根部铰链和臂间铰链在主弯方向和侧弯方向呈现纯结构热变形特性, 可以通过温度变化量进行仿真分析和计算。

3.2　铰链热变形分析与试验验证

根部铰链与臂间铰链热变形分析与试验结果对比, 如表 2 所示。由表可知, 试验值与分析值基本一致, 有效验证了有限元模型和试验方法的正确性; 天线铰链主弯方向转动角度大于侧弯方向, 说明铰链主弯方向热变形对铰链变形影响更大。

表2 铰链热变形分析与试验结果的转动角度

Table 2 Rotation angle of hinge thermal deformation analysis and test results (°)

结果类型	根部铰链主弯方向	根部铰链侧弯方向	臂间铰链主弯方向	臂间铰链侧弯方向
分析值	0.005 4	0.004 1	0.004 4	0.001 0
试验值	0.004 2	0.002 6	0.004 6	0.000 6

4 结 论

1）针对空间可展开天线结构的在轨热变形问题，提出了某空间可展开天线根部铰链和臂间铰链的有限元建模和分析方法，通过与铰链热变形试验对比，验证了有限元模型和试验方法的正确性，该建模与试验方法可为同类型空间可展开天线的设计、分析和优化提供参考。

2）基于高低温箱和电子经纬仪构建了铰链精度测试系统，运用经纬仪透过高低温箱上的石英玻璃进行高低温环境中锁定状态下铰链热变形试验。试验结果表明：铰链在主弯方向和侧弯方向呈现纯结构热变形特性。

参考文献（References）

[1] 郭宏伟, 邓宗全, 刘荣强. 空间索杆铰接式伸展臂参数设计与精度测量[J]. 光学精密工程, 2010, 18(5): 1105-1111.
GUO H W, DENG Z Q, LIU R Q. Parameter design and precision measurement of space cable-strut deployable articulated mast[J]. Optics and Precision Engineering, 2010, 18(5): 1105-1111(in Chinese).

[2] 胡明, 邓国兵, 陈文华, 等. 空间索杆铰接式伸展臂故障树分析[J]. 空间科学学报, 2012, 32(4): 598-604.
HU M, DENG G B, CHEN W H, et al. Fault tree analysis of the space cable-strut deployable articulated mast[J]. Chinese Journal of Space Science, 2012, 32(4): 598-604(in Chinese).

[3] 高明星. 可重复展收三棱柱伸展臂精度分析与设计研究 [D]. 哈尔滨: 哈尔滨工业大学, 2021: 17-34.
GAO M X. Precision analysis and design research on the extendable arm of the triangular prism with repeatable expansion and contraction [D]. Harbin: Harbin Institute of Technology, 2021: 17-34(in Chinese).

[4] MEI G J, XIANG B B, WANG N, et al. Correction method of antenna pointing error caused by the main reflector deformation[J]. Chinese Astronomy and Astrophysics, 2021, 45: 236-251.

[5] SAIAH B D S, ELHABIB B, ALI K, et al. Analysis of the communication links between the AlSat-1b satellite and the ground station: The impact of the auto tracking system on antenna pointing accuracy[J]. International Journal of Satellite Communications and Networking, 2021, 39(5): 486-499.

[6] WANG M, WANG P A. Assembling accuracy analysis of antenna array sub-module based on stream-of-inaccuracy theory[J]. Mechanical Engineering and Technology, 2020, 9(6): 628-637.

[7] 张惠峰. 空间可展天线精度测量、热分析、可靠性分析及间隙影响研究 [D]. 杭州: 浙江大学, 2010: 80-105.
ZHANG H F. Precision measurement, thermal analysis, reliability analysis and gap effect research of space deployable antenna [D]. Hangzhou: Zhejiang University, 2010: 80-105(in Chinese).

[8] 朱敏波, 曹峰云, 刘明治, 等. 星载大型可展开天线太空辐射热变形计算[J]. 西安电子科技大学学报, 2004, 31(1): 28-31.
ZHU M B, CAO F Y, LIU M Z, et al. The calculation of thermal radiaiton distortion forlarge deployable antennas in satellites[J]. Journal of Xidian University, 2004, 31(1): 28-31(in Chinese).

[9] MAHANEY J, THORNTON E A. Self-shadowing effects on the thermal-structural response of orbiting trusses[J]. Journal of Spacecraft and Rockets, 1987, 24(4): 342-348.

[10] 杨仲. 反射面天线热变形对其方向图的影响[J]. 电子机械工程, 2015, 31(5): 19-22.
YANG Z. Effects of thermal deformation on the far-field pattern of reflector antenna[J]. Electro-Mechanical Engineering, 2015, 31(5): 19-22(in Chinese).

[11] 武聪魁, 何柏岩, 袁鹏飞. 计及金属铰链的环形可展天线热-结构分析[J]. 工程设计学报, 2020, 27(3): 349-356.
WU C K, HE B Y, YUAN P F. Thermal-structural analysis of hoop deployable antenna with metal hinges[J]. Chinese Journal of Engineering Design, 2020, 27(3): 349-356(in Chinese).

[12] 张惠峰, 关富玲. 可展桁架天线热-结构耦合分析[J]. 浙江大学学报 (工学版), 2010, 44(12): 2320-2325.
ZHANG H F, GUAN F L. Thermal-structural coupling analysis of deployable truss antenna[J]. Journal of Zhejiang University (Engineering Science), 2010, 44(12): 2320-2325(in Chinese).

[13] 刘飞. 反射面天线热变形分析与补偿研究 [D]. 西安: 西安电子科技大学, 2014: 39-51.
LIU F. Study on thermal deformation analysis and compensation of reflector antenna [D]. Xi'an: Xidian University, 2014: 39-51(in Chinese).

[14] 赵军忠, 刘越东, 刘晓勇. 某星载天线热变形对跟踪指向精度的影响仿真分析[J]. 遥测遥控, 2013, 34(5): 1-4.
ZHAO J Z, LIU Y D, LIU X Y. Analysis of influences of thermal distortion on track pointing precision of a kind of satellite antenna[J]. Journal of Telemetry, Tracking and Command, 2013, 34(5): 1-4(inChinese).

[15] 吴楠. 高精度碳纤维夹层天线面板热变形特性研究 [D]. 南京: 南京航空航天大学, 2020: 22-37.
WU N. Study on thermal deformation characteristics of high precision carbon fiber sandwich antenna panel [D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2020: 22-37(in Chinese).

[16] 刘国青, 阮剑华, 罗文波, 等. 航天器高稳定结构热变形分析与试验验证方法研究[J]. 航天器工程, 2014, 23(2): 64-70.
LIU G Q, RUAN J H, LUO W B, et al. Research on thermal deformation analysis and test verification method for spacecraft high-stability structure[J]. Spacecraft Engineering, 2014, 23(2): 64-70(in Chinese).

[17] LIU T D, WANG C J, WANG B J. Flatness error evaluation of the satellite antenna[J]. International Journal of Future Computer and Communication, 2015, 4(4): 262-265.

Thermal deformation analysis and experimental verification of spatial deployable antenna hinge

YANG Zhijie[1]，ZHANG Congfa[2]，ZHAO Ruijie[1]，WANG Chunjie[3]，ZHAO Junpeng[1,*]

(1. School of Mechanical Engineering and Automation，Beihang University，Beijing 100191，China；
2. Beijing Institute of Spacecraft System Engineering，Beijing 100094，China；
3. State Key Laboratory of Virtual Reality and Systems，Beihang University，Beijing 100191，China)

Abstract: The in-orbit thermal deformation analysis of space-deployable antennae is very important because it is one of the major factors impacting the accuracy of these antennae.With the increasing requirements of space exploration missions, the influence of hinge thermal deformation on the pointing accuracy of space-deployable antennae cannot be ignored. In order to calculate the relative angle changes of the main bending direction and side bending direction of the hinges with each temperature change of 10 ℃, a finite element model of the root and arm hinges of a spatial deployable antennae was established. The hinge precision test system was constructed based on the high and low-temperature chamber and electronic theodolite, and the hinge thermal deformation test was carried out. The experimental results show that the root hinge and the interarm hinge exhibit pure structural thermal deformation characteristics in the main and side bending directions. The simulation results are basically consistent with the experimental results, which verify the validity of the finite element model and experimental analysis method. According to the relative Angle changes of the main and side bending directions of the hinge, the effect of thermal deformation on hinge deformation was assessed. The modeling and experimental analysis method can provide a reference for the design, analysis, and optimization of deployable antennas in the same space and the analysis of the factors affecting the accuracy of similar deployable mechanisms.

Keywords: deployable antenna; finite element analysis; hinge; thermal deformation; high and low temperature environment

Received：2022-04-06；**Accepted**：2022-05-14；**Published Online**：2022-06-07 13：07
URL：kns.cnki.net/kcms/detail/11.2625.V.20220606.1804.005.html
Foundation item：National Natural Science Foundation of China (51635002)
* **Corresponding author**. E-mail：zhaojunpeng@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0222

拓扑优化相变热沉在不同重力场的性能

王仕淞，殷健宝，邢玉明*，侯煦，王子贤
(北京航空航天大学 航空科学与工程学院，北京 100191)

摘　　要：为强化相变热沉的性能，基于带惩罚的固体各向同性材料(SIMP)法对山梨糖醇相变热沉进行传热拓扑优化设计。对拓扑优化的相变热沉(Ⅰ)和直肋式的相变热沉(Ⅱ)进行定常重力(0～20g)和周期性重力条件的数值模拟研究。利用无量纲数对比二者的热性能，研究结果表明：热沉Ⅰ的吸热和均热性能优于热沉Ⅱ；在相同重力环境下，以80 ℃为目标，热沉Ⅰ温控时间比热沉Ⅱ平均多26.8%；微低重力下热沉内部近乎导热，略逊于常规重力下热性能；超重力强化的自然对流可以显著提升热沉性能，10g条件相比常规重力的温控时间增加8.94%，而相同Ra^*下，周期性重力对热沉Ⅰ的性能有抑制作用。研究结果对飞行器载相变热沉的地面细化设计有一定指导意义。

关　键　词：相变材料；山梨糖醇；拓扑优化；热控性能；无量纲数
中图分类号：V221.8; V221.92; V228.3; V245.343
文献标志码：A　　**文章编号**：1001-5965(2024)01-0250-10

相变材料可以在近乎等温范围内吸收/释放大量潜热，因此，基于相变材料的储热技术被广泛用于太阳能光热发电[1]、低温制冷[2]、建筑暖通[3]、电动汽车热管理[4]及军事[5]等领域。近期的研究表明，相变热控技术在高热流密度设备热管理领域仍然不可或缺[6-8]。

有机相变材料和低熔点合金均可用于电子设备热管理，但低熔点合金密度大[9]，阻碍了其应用于空间和质量有限的飞行器。有机相变材料具有相容性好、密度小、来源广泛等优点，因此，在弹载设备[10]与卫星热控方面[11]均有应用，但是导热系数低[12]会影响其快速吸热。因此，众多学者对相变热沉的传热强化进行了广泛研究。

赵亮等[13]利用双精度模型模拟了铜泡沫/硬脂酸/石墨烯纳米片(GnP)复合相变热沉的热控性能，结果表明98%泡沫铜+3% GnP的硬脂酸复合相变材料温度梯度最小，且芯片温度最低，性能优于含95%泡沫铜及5% GnP的复合相变材料。Wang等[14]针对高碳醇/泡沫石墨复合相变热沉进行实验研究，实验结果显示泡沫石墨的强化导热作用主要发生在潜热吸收阶段，以90 ℃为控温目标，复合相变热沉的热控时间比纯高碳醇热沉的长24%。除了高导热介质，相变热沉中增加肋片结构也是强化传热的有效方式。Baby和Balaji[15]针对含针肋的正二十烷/铝相变热沉进行大量实验测试，还利用神经网络和遗传算法确定了相变热沉的最佳肋参数。Desai等[16]研究了不同形状的针肋对相变热沉强化传热的影响，研究结果表明三棱柱肋较对比肋有最大的传热面积因而热沉具有最低的“极限临界温度”。

近年来，随着增材制造技术[17]的协同发展，拓扑优化结构突破加工的限制，能够更加广泛用于强化传热[18]。文献[19-20]对发光二极管(light emitting diode, LED)设备的自然对流散热肋结构进行拓扑优化设计，并进行数值模拟和实验研究，研究表明具有复杂结构的拓扑优化热沉的散热性能好于直

收稿日期：2022-04-06；录用日期：2022-04-30；网络出版时间：2022-05-19 11:22
网络出版地址：link.cnki.net/urlid/11.2625.V.20220517.1831.004
基金项目：航空科学基金(20172851018)
*通信作者. E-mail：xym505@163.com

引用格式：王仕淞，殷健宝，邢玉明，等. 拓扑优化相变热沉在不同重力场的性能[J]. 北京航空航天大学学报，2024，50(1)：250-259.
WANG S S，YIN J B，XING Y M，et al. Topological optimzation of phase change heat sink performance in different gravity fields[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(1)：250-259 (in Chinese).

肋热沉和点阵结构热沉。文献 [21-24] 对管翅式相变储能换热器的翅结构进行拓扑优化, 优化结果指出考虑导热和液态相变材料自然对流对拓扑优化的构型有影响, 考虑自然对流的拓扑构型在能量存储速率更具优势。然而, 用于热管理的拓扑优化相变热沉的研究仍然较少。

除了导热系数, 对于飞行器载相变热沉, 机械振动、超重、微重力及电磁场等外场环境对热沉的热控性能也有影响。罗叶刚等[25] 研究机械振动环境对硬脂酸相变热沉的热控性能影响, 实验结果指出振动带来的强迫对流能够有效强化相变组件的热控性能, 并且相对于振动频率, 振幅变化对影响结果的扰动较小。文献 [26-27] 从抛物线飞行实验出发, 结合数值模拟, 研究微重力下相变材料的自由表面马兰戈尼对流对热沉的融化过程有显著影响。文献 [28-29] 也分别对相变热沉在超重和微低重力环境下的性能进行数值模拟研究, 结果表明重力引起的自然对流是影响热沉性能的关键。重力环境影响了飞行器相变热控模块的设计, 但不同重力环境对于拓扑优化相变热沉的影响还未知。

本文针对飞行器载设备相变热沉, 采用带惩罚的固体各向同性材料 (solid isotropic materials with penalization, SIMP) 法对山梨糖醇/铝的热沉进行二维传热结构的拓扑优化, 在定常和随时间正弦变化的重力条件下进行数值模拟研究, 并将其与相同体积分数的直肋热沉作对比。本文还量化了重力环境对拓扑优化相变热沉的性能影响, 为飞行器载相变热沉的地面细化设计提供一定的指导。

1　相变热沉的设计

本文针对飞行器载设备的热条件设计了 2 种相变热沉: ①相变拓扑优化的热沉(热沉Ⅰ), ②含直肋的热沉(热沉Ⅱ)。为了便于确定热沉参数, 更加贴近实际热控模块, 本文采用文献 [30] 中的部分设计参数。导热强化材料采用铝, 肋体积分数 a=30%; 根据文献 [30], 相变材料选择山梨糖醇, 因为山梨糖醇相变温度与电子设备极限结温匹配, 糖醇类相变潜热值高, 适合用作电子设备热控吸热材料, 但在实际应用中需要对糖醇进行改性, 以增强热沉的可靠性。二者的材料物性见表 1。

考虑到飞行器飞行环境的复杂性, 鉴于以下原因, 本文对热控组件的优化设计计算中忽略自然对流: ①如导弹的飞行重力环境是变化的, 空间探测器往往处于失重和低重力状态, 因此, 只考虑最基础的使用状态, 即零重力; ②在相同零重力的基础上设计更有利于后续不同工况的横向比较; ③液态相变材料的自然对流增加了问题的非线性, 即使在二维探索设计阶段的计算量非常大, 在之后的三维工程设计中的计算量将难以接受。

表 1　相变热沉的材料物性

Table 1　Material properties of heat sinks

材料	参数	数值
山梨糖醇[30-31]	熔点T_m/℃	75
	比热容c_p (25 ℃)/ [J·(kg·K)$^{-1}$]	2 050
	比热容c_p (100 ℃)/ [J·(kg·K)$^{-1}$]	2 990
	密度ρ/(kg·m^{-3})	1 490
	导热系数k/[W·(m·K)$^{-1}$]	0.59
	相变潜热L/(J·kg^{-1})	177 000
	热膨胀系数β/T^{-1}	0.000 1
	动力黏度μ/(Pa·s)	0.003 3
铝	比热容c_{pal} (25 ℃)/[J·(kg·K)$^{-1}$]	871
	密度ρ_{al}/(kg·m^{-3})	2 719
	导热系数k_{al}/[W·(m·K)$^{-1}$]	238

1.1　控制方程

相变传热采用焓法, 其控制方程为

$$\rho\frac{\partial h}{\partial t}=\nabla\cdot(k\nabla T) \tag{1}$$

$$h=c_p(T-T_{ref})+L \tag{2}$$

式中: ρ 为材料密度; h 为焓值; k 为材料的导热系数; c_p为材料的比热容; T_{ref} 为参考温度; L 为相变潜热; t 为时间。

1.2　材料物性插值

拓扑优化设计采用 SIMP 方法[23,32] 进行材料物性的插值, 计算式为

$$k(x)=k_{PCM}+\gamma^p(k_{TCE}-k_{PCM}) \tag{3}$$

$$\rho(x)=\rho_{PCM}+\gamma(\rho_{TCE}-\rho_{PCM}) \tag{4}$$

$$c_p(x)=c_{p,PCM}+\gamma\left(c_{p,TCE}-c_{p,PCM}\right) \tag{5}$$

式中: 下标 PCM、TCE 分别为相变材料、导热强化材料; γ 为材料的伪密度, 导热强化区域 γ=1, 相变材料区域 γ=0; p 为惩罚因子; x 为材料单元位置, 在设计域内通过式 (1) 对 2 种材料在有限元节点上进行插值。

将由式 (1) 和式 (2) 可以得到相变潜热带来的热源项, 插值形式为

$$Q(x)=-(1-\gamma)\rho\frac{\partial L}{\partial t}=-(1-\gamma)\rho L\frac{\partial\varphi}{\partial t} \tag{6}$$

式中: φ为相变材料的液相分数; $Q(x)$为插值后的设计域内热源。

1.3　优化问题的表达

优化问题的数学表达式为

Minimize

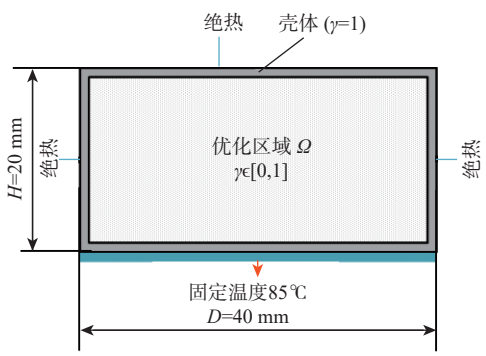
$$F(\gamma)=\frac{1}{n}\int_{\Omega}(T-T_{\mathrm{av}})^2\mathrm{d}\Omega$$

$$\text{s.t.}\quad \frac{\int\gamma\mathrm{d}\Omega}{\int\mathrm{d}\Omega}\leqslant a,\ 0\leqslant\gamma\leqslant 1 \tag{7}$$

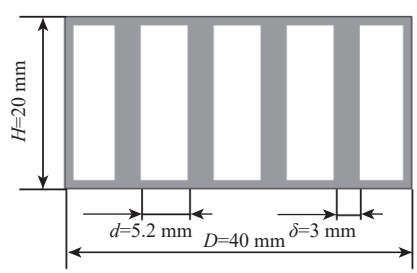
式中：Ω为设计域；T_{av}为域内平均温度；F为目标函数，即温度方差；a为导热强化材料的体积分数。

1.4 优化设计的边界条件

图1为热沉Ⅰ的拓扑优化设计域和热沉Ⅱ的结构示意图。热沉Ⅰ的导热拓扑优化设计域示意如图1(a)所示。热沉尺寸为40 mm×20 mm(长度D×宽度H)，浅灰色区域为优化设计的区域，优化区域内存在式(6)对应的热源。深灰色区域为非优化区域，即保持热沉壁厚为1 mm。热沉底面设定为以维持热设备的恒温85 ℃，其他3个表面为绝热边界。热沉Ⅱ起对照作用，其内部肋的体积分数a=30%，与热沉Ⅰ相同，热沉Ⅱ的结构示意图如图1(b)所示，热沉Ⅱ包含4支肋结构。其中，d为翅片间距，δ为肋片厚度。



(a) 热沉Ⅰ的拓扑优化设计域示意图



(b) 热沉Ⅱ的结构示意图

图1 热沉Ⅰ的拓扑优化设计域和热沉Ⅱ的结构示意图

Fig. 1 Topological optimization design domain of heat sink Ⅰ and structural schematic diagram of heat sink Ⅱ

2 相变数值模拟

2.1 物理模型

物理模型与热沉拓扑优化设计过程对应，初始温度为70 ℃，热沉的底部模拟芯片产生均匀的20 W的热流，其他三面均设为绝热边界来简化计算和分析。

2.2 数学模型

2.2.1 模型假设

采用熔多孔介质法模拟相变热沉吸热融化的过程。为降低模拟的复杂性，数值模型中还有以下合理假设：①液态相变材料的流动为层流；②除相变材料密度ρ外，其他物性为常数，但密度仅在Boussinesq假设项与温度呈线性关系，因而忽略材料的体积膨胀；③材料均为各向同性。

2.2.2 无量纲控制方程

采用无量纲的控制方程便于对比研究不同热沉的性能，本文基于x-y-z直角坐标系，利用以下无量纲数

$$\begin{cases} x^*=\dfrac{x}{l_{\mathrm{c}}},\quad y^*=\dfrac{y}{l_{\mathrm{c}}} \\ u^*=\dfrac{u}{V},\quad v^*=\dfrac{v}{V} \\ Fo=\dfrac{\alpha t}{l_{\mathrm{c}}^2} \\ \theta=\dfrac{T_{\mathrm{s}}-T_{\mathrm{m}}}{T_{\mathrm{m}}-T_{\mathrm{i}}},\quad p^*=\dfrac{p-p_0}{\rho V^2} \end{cases} \tag{8}$$

式中：l_{c}为特征长度，本文中取l_{c}=H；T_{s}为热源温度；Fo为傅里叶数；α为热扩散系数；T_{i}为初始温度；u和u^*分别为x方向的速度和无量纲速度；v和v^*分别为y方向的速度和无量纲速度；p和p^*分别为压力和无量纲压力；V为速度。

得到连续性方程为

$$\frac{\partial u^*}{\partial x^*}+\frac{\partial v^*}{\partial y^*}=0 \tag{9}$$

动量方程为

$$\begin{aligned}&\frac{\partial u^*}{\partial Fo}+u^*\frac{\partial u^*}{\partial x^*}+v^*\frac{\partial u^*}{\partial y^*}=-\frac{\partial p^*}{\partial x^*}+Pr\left(\frac{\partial^2 u^*}{\partial x^*}+\frac{\partial^2 u^*}{\partial y^*}\right)+\\&PrRa\theta+S_x\left(\frac{\partial v^*}{\partial Fo}+u^*\frac{\partial v^*}{\partial x^*}+v^*\frac{\partial v^*}{\partial y^*}\right)=\\&-\frac{\partial p^*}{\partial y^*}+Pr\left(\frac{\partial^2 v^*}{\partial x^*}+\frac{\partial^2 v^*}{\partial y^{*2}}\right)+PrRa\theta+S_y\end{aligned} \tag{10}$$

能量方程为

$$\begin{aligned}&\frac{\partial h^*}{\partial Fo}+\frac{\partial(u^*h^*)}{\partial x^*}+\frac{\partial(v^*h^*)}{\partial y^*}=\\&\frac{\partial^2\theta}{\partial x^*}+\frac{\partial^2\theta}{\partial y^*}+\frac{\partial^2\theta}{\partial z^*}+\frac{1}{Ste}V^*\cdot\nabla\varphi\end{aligned} \tag{11}$$

式中：Ste为斯蒂芬数；V^*为无量纲速度；Pr为普朗特数；h^*为无量纲比焓，与φ的关系为

$$h^*=\theta+\frac{\varphi}{Ste} \tag{12}$$

S_x和S_y为相变引起的动量汇项，表达式为

$$S_x=-A(\varphi)u^*S_y=-A(\varphi)v^* \tag{13}$$

式中: φ和 $A(\varphi)$ 的表达式和相关参数的取值, 可参见文献 [33]。

式 (10) 和式 (11) 引入了 4 个无量纲数, 利用无量纲数分析可以减少变量, 降低问题的复杂程度。本文所用到的无量纲数有: 表示无量纲时间的傅里叶数 Fo, 表示热源温度的无量纲温度 θ, 表示自然对流强度的瑞利数 Ra, 本文采用修正的瑞利数[14] Ra^*; 引入修正的斯蒂芬数[29] Ste^*; $Fo{\cdot}Ste^*$表示热沉的加热时间。

铝壳体和肋结构内导热控制方程为

$$\rho_{\mathrm{al}} c_{\mathrm{p_{al}}} \frac{\partial T}{\partial t} = k_{\mathrm{al}} \left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} \right) \tag{14}$$

在相变材料与铝结构的界面有边界条件为

$$\begin{cases} -k \left(\dfrac{\partial T_{\mathrm{PCM}}}{\partial n} \right) = -k_{\mathrm{al}} \dfrac{\partial T_{\mathrm{al}}}{\partial n} \\ T_{\mathrm{PCM}} = T_{\mathrm{al}} \end{cases} \tag{15}$$

式中: 下标 al 表示材料铝。

2.3　网格和时间无关性验证

分别选择网格尺寸为 0.125, 0.250, 0.500 mm 进行数值计算, 得到相变材料平均液相分数φ随时间变化如图 2 所示。由图可知, 网格尺寸为 0.250 mm 的液相分数曲线与 0.500, 0.125 mm 之间误差分别为 11.7%、0.9%, 因此, 选择网格尺寸为 0.250 mm 的网格进行计算。采用网格尺寸为 0.250 mm 网格在不同时间步长计算的液相分数随时间变化曲线如图 3 所示。由图可知, 时间步长从 0.01 s 变化到 0.50 s, 曲线之间最大误差不超过 3%, 由此选择 0.05 s 作为模拟时间步长。

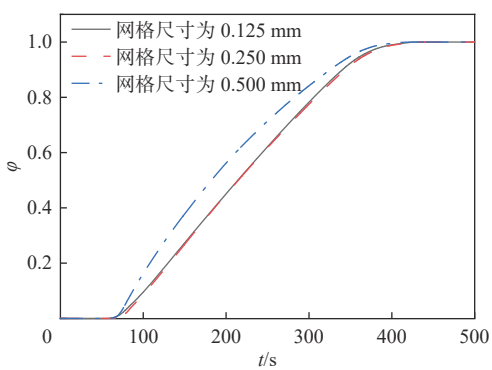


图 2　不同网格尺寸液相分数随时间变化曲线

Fig. 2　Time varying curves of liquid phase fraction with different grid sizes

2.4　数值方法

融化过程的数值计算在 ANSYS/FLUENT 中进行, 关于数值模型的有效性验证及相关的参数设置, 详见文献 [13,33-34]。

优化过程在 COMSOL/Multiphysics 中进行, 设定肋结构的体积分数 a=0.5, 初始材料的伪密度 γ 自动及惩罚因子 p=5。瞬态计算时间步长为 0.01 s, 设定相变材料液相分数为 100% 是瞬态计算的停止条件; 采用全局收敛的移动渐近线法 (global convergent method of moving asymptotes, GCMMA) 进行优化更新计算材料伪密度, 令最大总迭代步数为 1 000, 优化终止残差为 10^{-6}, 采取双曲正切投影得到清晰的结构边界。利用 1.2 节的物性插值方法和 1.4 节的优化边界条件, 得到树状的拓扑优化热沉, 如图 4 所示。

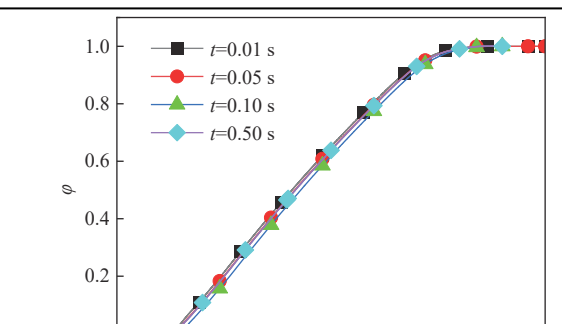


图 3　不同时间步长液相分数随时间变化曲线

Fig. 3　Time varying curves of liquid phase fraction with different time steps



图 4　热沉 I 树状拓扑优化结果示意图

Fig. 4　Schematic of tree-like topological optimization result for heat sink I

3　结果讨论

优化后的热沉中间有 2 支“主干肋”, 两侧还各有一干支结构, 因而热沉 II 设计有 4 支肋结构。拓扑优化结构呈树状或叶脉状, 可以更好的将热源发热量萃取到相变材料中, 实现温度的均匀分布, 达到散热控温的目的。

在零重力 (w=0g, w 为重力)、微低重力 (w=0.01g, 0.1g)、常规重力 (w=1g)、超重力 (w=5g, 10g, 15g, 20g) 和周期性重力 ($g\sin(2\pi t/10)$ 、$g\sin(2\pi t/5)$、$10g\sin(2\pi t/10)$) 条件下分别进行了数值模拟。利用 2.2.2 节中无量纲温度 θ、液相体积分数φ、无量纲时间 Fo、修正的斯蒂芬数 Ste^*、修正的瑞利数 Ra^*等参数对热沉在不同重力场的控温、融化及速度场特性进行对比分析。

3.1 恒定重力

3.1.1 温控和融化特性

树状拓扑优化热沉 θ 与 $Fo{\cdot}Ste^*$ 的关系如图 5 所示，其物理意义是热沉的热控温度水平与热沉吸收热源的发热量的关系。由图可知，当 θ<0.17 时，热沉处于显热阶段，由于吸热能力不足导致热源表面温度快速上升；当 θ<0.7 时，重力对热沉Ⅰ的控温性能没有影响，因为此阶段属于显热和融化前期阶段，热沉中的导热占主导。当 0.7≤θ≤3.7 之间时，重力对于热沉Ⅰ的吸热控温能力有显著的影响。随着重力加速度的增加，热沉对热源温度上升的抑制变强。液态山梨糖醇的自然对流传热在此阶段占主导作用，而驱动自然对流的浮升力由重力加速度决定，因此，在较大的 Ra^* 情况下，相变材料的潜热吸收速率也更大。图 6 可以印证这一点，曲线表示液态相变材料体积分数 φ 随加热过程的变化。由图 6 可知，在梨糖醇融化过半后，各融化曲线变化速率有明显区别，高 Ra^* 条件下融化速率更快，因而可以迅速吸热，延缓芯片温度上升。

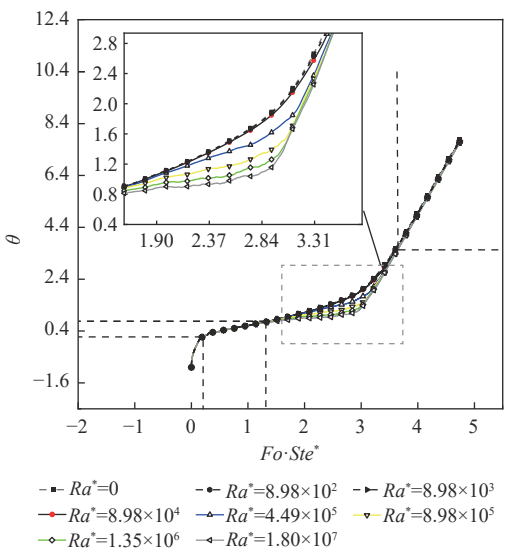


图 5 热沉Ⅰ，θ 与 $Fo{\cdot}Ste^*$ 的关系

Fig. 5 Heat sink Ⅰ, θ relationship with $Fo{\cdot}Ste^*$

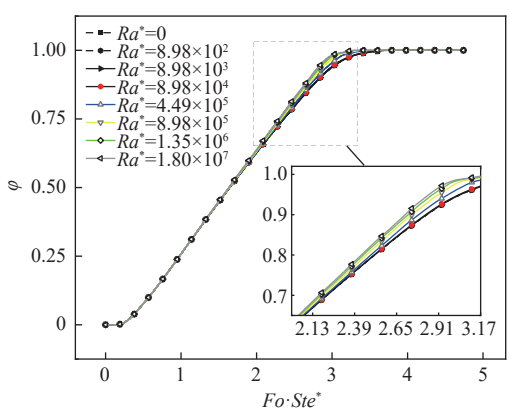


图 6 热沉Ⅰ液相体积分数 φ 与 $Fo{\cdot}Ste^*$ 的关系

Fig. 6 Heat sink Ⅰ liquid phase volume fraction φ relationship with $Fo{\cdot}Ste^*$

在图 5 和图 6 中，零重力和微低重力对应的曲线基本重合，此时 Ra^* 较小，热沉内部自然对流很弱，整体的传热方式近乎导热。另外，常规重力环境 (Ra^*=8.98×10⁴) 下热沉的控温性能仅略强于微低重力，这是因为热沉在重力方向尺度不大，也会使得自然对流强度有限，不能利用自然对流明显提升温控性能。

因此，充分利用超重带来的后融化阶段相变热沉的性能的提升，可以有效提升热沉对热源的控温时间。

3.1.2 热性能的对比

为对比研究热沉Ⅰ与热沉Ⅱ的热控特性，分别选取了零重力 (w=0g, Ra^*=0)、低重力 (w=0.1g, Ra^*=8.98×10³)、常规重力 (w=1g, Ra^*=8.98×10⁴) 和超重力 (w=10g, Ra^*=8.98×10⁵) 为基本工况进行研究。3 种重力场的热沉Ⅰ与热沉Ⅱ的无量纲温度曲线如图 7 所示。由图可知，零重力条件下，当 $Fo{\cdot}Ste^*$>1 时，热沉Ⅰ的曲线在热沉Ⅱ的下方，说明拓扑优化肋对于热沉在潜热吸收阶段的纯导热作用有显著的增强。重力场对于热沉Ⅰ和热沉Ⅱ的影响是一致的，高重力加速度引起的自然对流在 2 种热沉融化过程都起到关键作用。相同 Ra^* 下，热沉Ⅰ的控温能力明显好于热沉Ⅱ，说明拓扑优化使得热沉的快速吸热和均温能力增强。图 7 的 2 个红色方框中，可以观察到热沉Ⅰ和热沉Ⅱ的对流效应开始占据主导至作用消失对应的无量纲温度 θ 基本一致，仅是加热时间不同，所以温度是决定热沉内部液态相变材料对流作用的主导因素，因此，这也间接指出，选择与控温目标匹配的相变材料对于有效延长控温时长是很重要的。

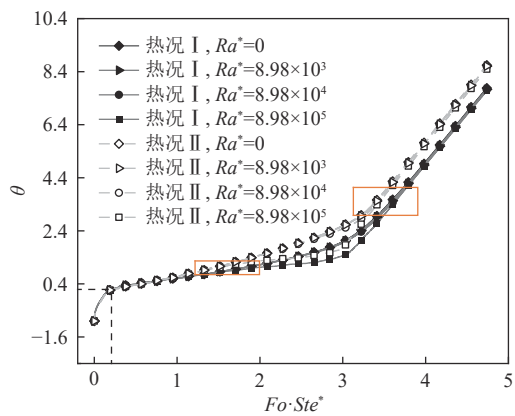


图 7 热沉Ⅰ与热沉Ⅱ的 θ 与 $Fo{\cdot}Ste^*$ 关系

Fig. 7 θ of heat sinks Ⅰ and Ⅱ relationship with $Fo{\cdot}Ste^*$

在 3 个时刻 (t=200 s, $Fo{\cdot}Ste^*$=1.90; t=250 s、$Fo{\cdot}Ste^*$=2.37; t=300 s, $Fo{\cdot}Ste^*$=2.84)2 种热沉的平均温度 θ_{av} 及对于应热源的温度 θ，如表 2 所示。由表

可知, 在所有的工况中, 热沉 Ⅰ 的 θ_{av} 和 θ 都要明显小于热沉 Ⅱ。这可以再次直接证明: 经由拓扑优化的热沉 Ⅰ 的内部均温能力和外部控温能力都要好于热沉 Ⅱ。当然, 均温能力和控温能力是耦合的, 具有强均温能力的热控组件也是其控温能力提升的前提, 从实质上, 导热系数是控温组件的决定性参数, 而拓扑优化的结果就是增大了热沉模块的等效导热系数。

表 2　重力场中不同时刻热沉 Ⅰ 和 Ⅱ 的温度

Table 2　Temperature of heat sinks Ⅰ and Ⅱ at different time in gravity field

Ra^*	热沉	$Fo{\cdot}Ste^*$=1.90		$Fo{\cdot}Ste^*$=2.37		$Fo{\cdot}Ste^*$=2.84	
		θ_{av}	θ	θ_{av}	θ	θ_{av}	θ
0	Ⅰ	**0.715**	**1.038**	**1.014**	**1.341**	**1.429**	**1.754**
	Ⅱ	1.155	1.362	1.596	1.807	2.132	2.346
8.98×10^3	Ⅰ	**0.711**	**1.034**	**1.001**	**1.335**	**1.417**	**1.743**
	Ⅱ	1.149	1.356	1.588	1.798	2.120	2.332
8.98×10^4	Ⅰ	**0.707**	**1.030**	**1.000**	**1.325**	**1.402**	**1.720**
	Ⅱ	1.147	1.350	1.579	1.783	2.072	2.263
8.98×10^5	Ⅰ	**0.654**	**0.961**	**0.825**	**1.108**	**1.043**	**1.311**
	Ⅱ	1.031	1.201	1.200	1.327	1.034	1.544

注: 表中黑色加粗为拓扑优化热沉对应数据。

考虑到热沉在飞行器中的应用, 以弹载设备温控模块为例, 设定控温目标不超过 80 ℃(θ=1), 控温时间不少于 3 min($Fo{\cdot}Ste^*$=1.71)。热沉 Ⅰ 和热沉 Ⅱ 的控温时长关系如图 8 所示。由图可知, 10 倍的重力加速度对于热沉 Ⅰ 的延长最多为 8.94%; 而对热沉 Ⅱ 的热控时间延长作用较弱仅为 4.94%。热沉 Ⅰ 在设定的控温目标限制下的控温时长远大于热沉 Ⅱ, 控温时间平均延长 26.8%。由图 8 数据还可知, 随着 Ra^*的增加, 热沉 Ⅰ 的热控时间延长率分别为 0.36%、0.36%、8.95%; 而热沉 Ⅱ 的分别为 0.32%、0.23%、4.94%。相同重力下, 热沉 Ⅰ 的热控时间延长率均大于热沉 Ⅱ, 说明热沉 Ⅰ 的优化结构还对相变材料自然对流的传热强度有正向作用。

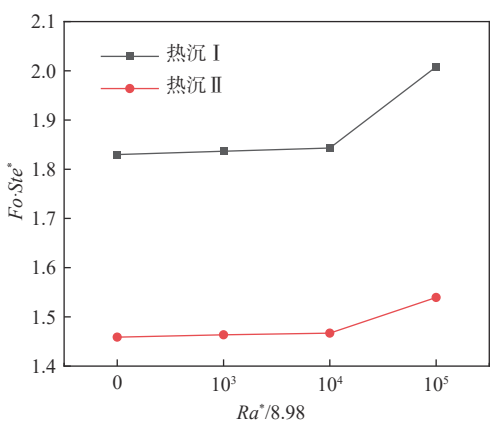


图 8　以 θ=1 为目标热沉 Ⅰ 与热沉 Ⅱ 的控温时长

Fig. 8　Temperature control time of heat sinks Ⅰ and Ⅱ at θ=1

此外, 热沉 Ⅰ 的 $Fo{\cdot}Ste^*$均大于 1.71, 说明该热沉可以将设备表面温度在 3 min 内有效控制在 80 ℃以下, 然而热沉 Ⅱ 却不能完成此任务。综上所述, 拓扑优化技术可以适用于飞行器相变热沉设计, 相变热控模块的细化设计是必要的。目前, 先进的增材制造技术飞速发展, 拓扑优化技术用于散热设备可以用有限的质量约束实现更高效的散热。进一步, 拓扑优化应用于散热设备不仅应考虑到组件的热性能, 还应综合考虑力学性能, 以便于应用于实际飞行器[35]。

3.1.3　相变速度场分析

无量纲速度 V^*的计算式为

$$V^* = \frac{\sqrt{u^2+v^2}}{V_{\varphi=1}} \tag{16}$$

式中: $V_{\varphi=1}$为相变材料完全融化后稳定的速度大小。

热沉 Ⅰ 和热沉 Ⅱ 中液相材料无量纲速度 V^*的大小平均值随 $Fo{\cdot}Ste^*$的变化关系, 如图 9 所示。由图可知, 热沉 Ⅰ 和热沉 Ⅱ 对应的 V^*随 $Fo{\cdot}Ste^*$的变化规律大致相同, 在融化阶段($Fo{\cdot}Ste^*$约为 1～3), V^*逐

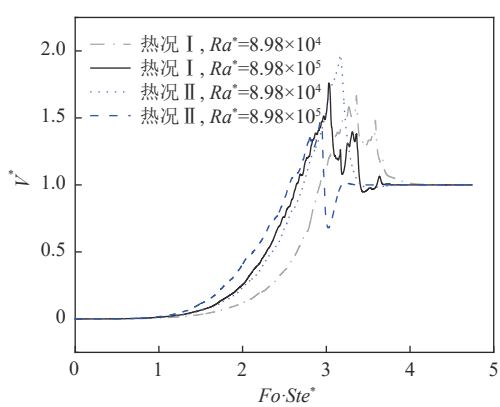


图 9　热沉 Ⅰ 与热沉 Ⅱ 的 V^*与 $Fo{\cdot}Ste^*$关系

Fig. 9　V^* of heat sinks Ⅰ and Ⅱ relationship with $Fo{\cdot}Ste^*$

渐平滑增加，融化结束至相变材料流场稳定（$Fo{\cdot}Ste^*$约为3～4），V^*存在剧烈的波动，此阶段由图6可知相变材料即将完全融化，这说明固相相变材料的消失会使流场进一步发展，产生短暂的波动，转向液相显热吸收阶段。

重力对于V^*的变化有显著影响，在Ra^*=8.98×10^{5}时，热沉Ⅰ和热沉Ⅱ的V^*的增大速率要大于Ra^*=8.98×10^{4}条件下，表明重力的增大可以增加相变材料整体的自然对流速度，进而增强融化阶段的换热强度。然而在速度波动阶段重力的影响没有上述规律，但是热沉Ⅰ的波动情况要更加明显，其曲线的峰要多于热沉Ⅱ，这是热沉Ⅰ的中复杂结构导致的。

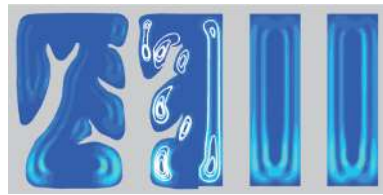
在不同时刻（t=250 s，$Fo{\cdot}Ste^*$=2.37；t=350 s，$Fo{\cdot}Ste^*$=3.32；t=450 s，$Fo{\cdot}Ste^*$=4.27）情况下，2种热沉在不同Ra^*下流场的变化，如图10所示。由图可知，在不同时刻，10g的重力加速度对应的热沉速度更大，这与图9所示的速度平均分布情况一致。热沉Ⅰ的流线展示出拓扑结构的分叉处存在涡，这些涡增加了热沉Ⅰ流场的复杂性；而热沉Ⅱ只在每2支肋之间形成2个长条形的涡。

热沉Ⅰ和Ⅱ的融化情况如图11所示。由图可知，对比图11(a)和11(c)或11(b)和11(d)可以印证图6中的曲线关系，更大的重力引起的自然对流可以加快相变材料的融化；横向对比11(b)或11(d)，相同同条件下，热沉Ⅱ的融化进程快于热沉Ⅰ。结合图10和图11可知，拓扑优化的热沉Ⅰ存在众多利于局部换热的涡，使其对于全局的均温性强于热沉Ⅱ，正因如此，热沉Ⅰ无局部过热没有像热沉Ⅱ更快速的融化，因而热沉Ⅰ有足够的潜热来实现更长时间的控温。

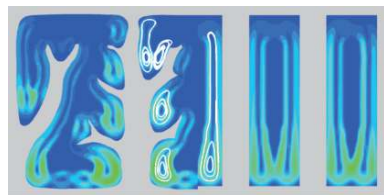
3.2 周期性重力

除了恒定重力场，周期性变化的重力场也会对相变热沉的对流产生影响，其效果是增强或抑制尚待研究。研究显示，超声振动[36]和机械振动[25,37-38]对相变材料的传热有一定强化作用。基于周期性重力场，周期与飞机“荷兰滚”模态周期相当（10，5 s），研究该范围内重力的幅值和频率变化对热沉Ⅰ性能的影响。本文另外引入无量纲频率$\omega^*=\omega l_c^2/\alpha$，其中$\omega$为重力加速度变化的频率。

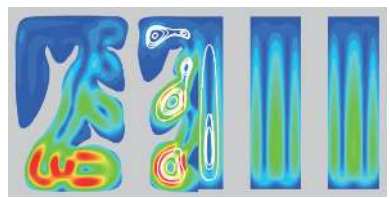
周期性重力场和定常重力场下热沉Ⅰ的θ与$Fo{\cdot}Ste^*$的关系，如图12所示。由图可知，ω^*的出现没有改变液相对流占主导的温度区间，依然在0.17<θ<3.7的范围。文献[38-39]采用了增强率R_v来量化波动的加速度引起的增强效应，R_v定义为波动的加速度幅值与重力加速度大小之比，在本文



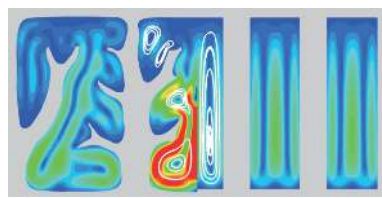
(a) 热沉Ⅰ，Ra^*=8.98×10^4，t=250 s，$Fo{\cdot}Ste^*$=2.37，w=1g



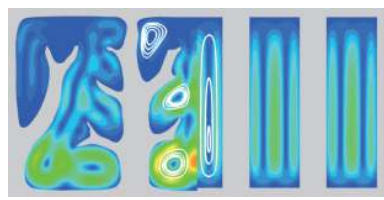
(b) 热沉Ⅱ，Ra^*=8.98×10^5，t=250 s，$Fo{\cdot}Ste^*$=2.37，w=10g



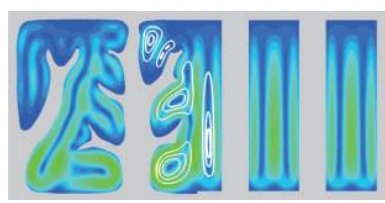
(c) 热沉Ⅰ，Ra^*=8.98×10^4，t=350 s，$Fo{\cdot}Ste^*$=3.32，w=1g



(d) 热沉Ⅱ，Ra^*=8.98×10^5，t=350 s，$Fo{\cdot}Ste^*$=3.32，w=1g



(e) 热沉Ⅰ，Ra^*=8.98×10^4，t=450 s，$Fo{\cdot}Ste^*$=4.27，w=1g



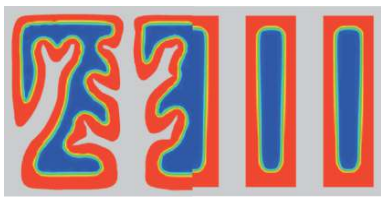
(f) 热沉Ⅱ，Ra^*=8.98×10^5，t=450 s，$Fo{\cdot}Ste^*$=4.27，w=10g



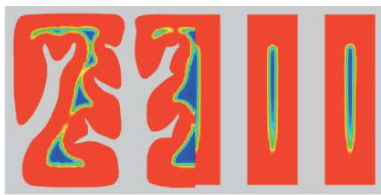
图10 热沉Ⅰ与热沉Ⅱ在不同Ra^*下流场的变化

Fig. 10 Varying of flow for field heat sinks Ⅰ and Ⅱ at different Ra^*

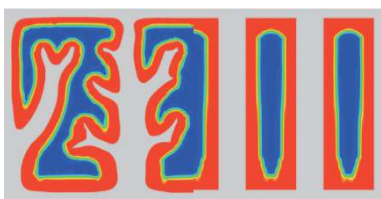
中，当Ra^*=8.98×10^3、w=0.1g，Ra^*=8.98×10^4、w=1g，Ra^*=8.98×10^5、w=10g时，R_v=−1，0，9。显然，增大Ra^*则R_v增大，热沉的传热被强化，这一点通过对比曲线②和③，④和⑤及⑥和⑦亦可得到。但是周期性重力（$\omega^*\neq0$）的增强效果没有定常重力（$\omega^*=0$）效果明显。这是因为周期性重力的等效的重力加速度的



(a) 热沉Ⅰ, Ra^*=8.98×10⁴, t=200 s, $Fo{\cdot}Ste^*$=1.90, w=1g



(b) 热沉Ⅱ, Ra^*=8.98×10⁴, t=300 s, $Fo{\cdot}Ste^*$=2.84, w=1g



(c) 热沉Ⅰ, Ra^*=8.98×10⁵, t=200 s, $Fo{\cdot}Ste^*$=1.90, w=10g



(d) 热沉Ⅱ, Ra^*=8.98×10⁵, t=300 s, $Fo{\cdot}Ste^*$=2.84, w=10g



图 11　热沉Ⅰ与热沉Ⅱ在不同 Ra^*下的融化进展

Fig. 11　Varying of melting progress for heat sinks Ⅰ and Ⅱ at different Ra^*

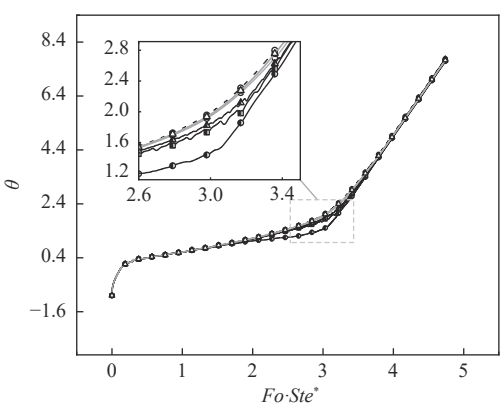


图 12　周期性重力场中热沉Ⅰ的 θ 与 $Fo{\cdot}Ste^*$关系

Fig. 12　θ of heat sink Ⅰ in periodic gravity field relationship with $Fo{\cdot}Ste^*$

矢量平均值为 0, 当 $\omega^*\neq0$ 时, 融化阶段热沉内部的总体传热速率被削弱。

对比图中②、④和⑥这 3 条曲线可以发现, 在相同 Ra^*下, ω^*的增加并未对热沉控温曲线带来明显改变, 然而由局部放大图可以发现, ④和⑥这 2 条曲线几乎重合但略高于曲线②, 同样, 当 Ra^*=8.98×10⁵, 对比曲线③、⑤和⑦可以发现, 虽然⑤和⑦产生明显的波动, 但 ω^*的增大会使 θ 曲线上移, 即热沉控温能力下降。低频重力的周期与热沉内部流动特征尺寸和速度确定的流动时间尺度是相近的, 因此, 较低的 ω^*反而抑制了热沉内部的对流的发展, 造成等效导热的下降。另外, 本文中的 Ra^*还不够大, 进一步通过增大尺度提高 Ra^*或许可以加强速度场的扰动, 由此改善周期性重力环境中热沉的内部传热效果。

4　结　论

1) 以温度方差为目标函数, 得到树状肋拓扑优化热沉, 其吸热能力和均温能力好于直肋热沉, 因此, 能实现更长的温控时间。

2) 以 80 ℃ 为控温目标, 拓扑优化热沉的热控时间均超过 3 min, 且比直肋热沉平均延长 26.8%。

3) 固态相变材料的导热可被拓扑优化结构强化, 并且液态相变材料还可在拓扑优化热沉的分支之间形成涡, 有利潜热的快速储存。

4) 由于自然对流, 重力的增加可以提高热沉的控温能力, 主要体现在融化的后半阶段; 在完全融化后, 重力对热沉控温能力的影响消失。w=10g 条件下拓扑优化热沉控温时间较 w=1g 延长了 8.94%。微低重力范围热沉的控温性能几乎不变, 然而相同 Ra^*下周期性的重力对热沉性能有抑制作用。

参考文献（References）

[1] KUMAR A, TIWARI A K, SAID Z. A comprehensive review analysis on advances of evacuated tube solar collector using nanofluids and PCM[J]. Sustainable Energy Technologies and Assessments, 2021, 47: 101417.

[2] SELVNES H, ALLOUCHE Y, MANESCU R I, et al. Review on cold thermal energy storage applied to refrigeration systems using phase change materials[J]. Thermal Science and Engineering Progress, 2021, 22: 100807.

[3] DA CUNHA S R L, DE AGUIAR J L B. Phase change materials and energy efficiency of buildings: A review of knowledge[J]. Journal of Energy Storage, 2020, 27: 101083.

[4] THAKUR A K, PRABAKARAN R, ELKADEEM M, et al. A state of art review and future viewpoint on advance cooling techniques for lithium-ion battery system of electric vehicles[J]. Journal of Energy Storage, 2020, 32: 101771.

[5] LYU J, LIU Z W, WU X H, et al. Nanofibrous kevlar aerogel films and their phase-change composites for highly efficient infrared stealth[J]. ACS Nano, 2019, 13(2): 2236-2245.

[6] YANG T Y, BRAUN P V, MILJKOVIC N, et al. Phase change material heat sink for transient cooling of high-power devices[J]. International Journal of Heat and Mass Transfer, 2021, 170: 121033.

[7] SHAMBERGER P J, BRUNO N M. Review of metallic phase change materials for high heat flux transient thermal management applications[J]. Applied Energy, 2020, 258: 113955.

[8] MATHEW J, KRISHNAN S. A review on transient thermal management of electronic devices[J]. Journal of Electronic Packaging, 2022, 144(1): 010801.

[9] ZHAO L, XING Y M, LIU X. Experimental investigation on the thermal management performance of heat sink using low melting point alloy as phase change material[J]. Renewable Energy, 2020, 146: 1578-1587.

[10] 何智航. 热管 PCM 热控装置设计及性能研究[J]. 热科学与技术, 2018, 17(1): 80-86.
HE Z H. Design and performance study on heat pipe/PCM composite thermal control device[J]. Journal of Thermal Science and Technology, 2018, 17(1): 80-86(in Chinese).

[11] 宁献文, 李劲东, 王玉莹, 等. 中国航天器新型热控系统构建进展评述[J]. 航空学报, 2019, 40(7): 022874.
NING X W, LI J D, WANG Y Y, et al. Review on construction of new spacecraft thermal control system in China[J]. Acta Aeronautica et Astronautica Sinica, 2019, 40(7): 022874(in Chinese).

[12] GE H S, LI H Y, MEI S F, et al. Low melting point liquid metal as a new class of phase change material: An emerging frontier in energy area[J]. Renewable and Sustainable Energy Reviews, 2013, 21: 331-346.

[13] 赵亮, 邢玉明, 刘鑫, 等. 基于硬脂酸复合相变材料的被动热沉性能[J]. 北京航空航天大学学报, 2019, 45(5): 970-979.
ZHAO L, XING Y M, LIU X, et al. Performance of a passive heat sink using stearic acid based composite as phase change material[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(5): 970-979(in Chinese).

[14] WANG S S, XING Y M, HAO Z L, et al. Experimental study on the thermal performance of PCMS based heat sink using higher alcohol/graphite foam[J]. Applied Thermal Engineering, 2021, 198: 117452.

[15] BABY R, BALAJI C. Thermal optimization of PCM based pin fin heat sinks: An experimental study[J]. Applied Thermal Engineering, 2013, 54(1): 65-77.

[16] DESAI A N, GUNJAL A, SINGH V K. Numerical investigations of fin efficacy for phase change material (PCM) based thermal control module[J]. International Journal of Heat and Mass Transfer, 2020, 147: 118855.

[17] 江洪, 刘敬仪. 3D 打印在航空航天领域中的应用初探[J]. 新材料产业, 2019(2): 21-24.
JIANG H, LIU J Y. Application of 3D printing in aerospace field[J]. Advanced Materials Industry, 2019(2): 21-24(in Chinese).

[18] DBOUK T. A Review about the engineering design of optimal heat transfer systems using topology optimization[J]. Applied Thermal Engineering, 2017, 112: 841-854.

[19] ALEXANDERSEN J, SIGMUND O, MEYER K E, et al. Design of passive coolers for light-emitting diode lamps using topology optimisation[J]. International Journal of Heat and Mass Transfer, 2018, 122: 138-149.

[20] LAZAROV B S, SIGMUND O, MEYER K E, et al. Experimental validation of additively manufactured optimized shapes for passive cooling[J]. Applied Energy, 2018, 226: 330-339.

[21] PIZZOLATO A, SHARMA A, MAUTE K, et al. Design of effective fins for fast PCM melting and solidification in shell-and-tube latent heat thermal energy storage through topology optimization[J]. Applied Energy, 2017, 208: 210-227.

[22] PIZZOLATO A, SHARMA A, GE R H, et al. Maximization of performance in multi-tube latent heat storage-optimization of fins topology, effect of materials selection and flow arrangements[J]. Energy, 2020, 203: 114797.

[23] 游吟, 赵耀, 赵长颖, 等. 相变储热单元内肋片结构的拓扑优化[J]. 科学通报, 2019, 64(11): 1191-1199.
YOU Y, ZHAO Y, ZHAO C Y, et al. The topology optimization of the fin structure in latent heat storage[J]. Chinese Science Bulletin, 2019, 64(11): 1191-1199(in Chinese).

[24] TIAN Y, LIU X L, XU Q, et al. Bionic topology optimization of fins for rapid latent heat thermal energy storage[J]. Applied Thermal Engineering, 2021, 194: 117104.

[25] 罗叶刚, 邢玉明, 赵亮, 等. 振动环境对相变组件热控性能影响的实验[J]. 航空动力学报, 2019, 34(3): 592-599.
LUO Y G, XING Y M, ZHAO L, et al. Experiment on the thermal management performance of phase change material under vibration environment[J]. Journal of Aerospace Power, 2019, 34(3): 592-599(in Chinese).

[26] EZQUERRO J M, BELLO A, SALGADO SÁNCHEZ P, et al. The thermocapillary effects in phase change materials in microgravity experiment: Design, preparation and execution of a parabolic flight experiment[J]. Acta Astronautica, 2019, 162: 185-196.

[27] SALGADO SÁNCHEZ P, EZQUERRO J M, FERNÁNDEZ J, et al. Thermocapillary effects during the melting of phase change materials in microgravity: Heat transport enhancement[J]. International Journal of Heat and Mass Transfer, 2020, 163: 120478.

[28] XU Y, ZHU Z H, LI S, et al. Numerical investigation on melting process of a phase change material under supergravity[J]. Journal of Thermal Science and Engineering Applications, 2021, 13(2): 021014.

[29] KANSARA K, SINGH V K, PATEL R, et al. Numerical investigations of phase change material (PCM) based thermal control module (TCM) under the influence of low gravity environment[J]. International Journal of Heat and Mass Transfer, 2021, 167: 120811.

[30] IRADUKUNDA A C, VARGAS A, HUITINK D, et al. Transient thermal performance using phase change material integrated topology optimized heat sinks[J]. Applied Thermal Engineering, 2020, 179: 115723.

[31] BOLMATENKOV D N, YAGOFAROV M I, SOKOLOV A A, et al. The heat capacities and fusion thermochemistry of sugar alcohols between 298.15 K and T_m: The study of *D*-sorbitol, *D*-mannitol and *myo*-inositol[J]. Journal of Molecular Liquids, 2021, 330: 115545.

[32] HO J Y, SEE Y S, LEONG K C, et al. An experimental investigation of a PCM-based heat sink enhanced with a topology-optimized tree-like structure[J]. Energy Conversion and Management, 2021, 245: 114608.

[33] HOU X, XING Y M, HAO Z L. Multi-objective optimization of a

composite phase change material-based heat sink under non-uniform discrete heating[J]. Applied Thermal Engineering, 2021, 197: 117435.

[34] 侯煦, 邢玉明, 郝兆龙, 等. 高碳醇/膨胀石墨复合相变热沉多目标优化[J]. 北京航空航天大学学报, 2021, 47(9): 1866-1873.
HOU X, XING Y M, HAO Z L, et al. Multi-objective optimization of a high alcohol/expanded graphite composite PCM based heat sink[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(9): 1866-1873(in Chinese).

[35] ZHOU H, ZHANG X Y, ZENG H Z, et al. Lightweight structure of a phase-change thermal controller based on lattice cells manufactured by SLM[J]. Chinese Journal of Aeronautics, 2019, 32(7): 1727-1732.

[36] OH Y K, PARK S H, CHO Y I. A study of the effect of ultrasonic vibrations on phase-change heat transfer[J]. International Journal of Heat and Mass Transfer, 2002, 45(23): 4631-4641.

[37] LAN C W. Effects of axial vibration on vertical zone-melting processing[J]. International Journal of Heat and Mass Transfer, 2000, 43(11): 1987-1997.

[38] VADASZ J J, MEYER J P, GOVENDER S, et al. Experimental study of vibration effects on heat transfer during solidification of paraffin in a spherical shell[J]. Experimental Heat Transfer, 2016, 29(3): 285-298.

[39] BARDAN G, MOJTABI A. On the Horton-Rogers-Lapwood convective instability with vertical vibration: Onset of convection[J]. Physics of Fluids, 2000, 12(11): 2723-2731.

Topological optimzation of phase change heat sink performance in different gravity fields

WANG Shisong，YIN Jianbao，XING Yuming*，HOU Xu，WANG Zixian

(School of Aeronautic Science and Engineering，Beihang University，Beijing 100191，China)

Abstract：To enhance the performance of phase change heat sink based heat sinks, the topology optimization of a sorbitol/aluminum phase change heat sink based heat sink is carried out using solid isotropic material with penalization (SIMP) method. Numerical investigations on phase change heat sink with topologically optimized fins (I) and straight fins (II) under the constant gravity (0～20*g*) and periodic gravity are presented. Dimensionless numbers are used to compare the thermal performance of the two heat sinks. The results show that heat sink I performs better than heat sink II. Under the same gravity environment with 80 ℃ as the goal, the temperature control time of the heat sink I is extended by up to 26.8% on average. Under the microgravity and low gravity, heat conduction dominates the heat sinks, of which the thermal performance is slightly inferior to that under conventional gravity. The natural convection of the liquid PCM driven by supergravity significantly enhances the heat transfer. For heat sink I, the temperature control time of 10*g* is 8.94% higher than that of conventional gravity, and the periodic gravity has an inhibitory effect at the same Ra^*. Research findings provide guidance for detailing the design of aircraft phase change heat sinks .

Keywords：phase change material；sorbitol；topological optimization；thermal control performance；dimensionaless number

Received：2022-04-06；**Accepted**：2022-04-30；**Published Online**：2022-05-19 11：22

URL：link.cnki.net/urlid/11.2625.V.20220517.1831.004

Foundation item：Aeronautical Science Foundation of China (20172851018)

* **Corresponding author**. E-mail：xym505@163.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0229

基于预设性能的飞机全电刹车系统滑模控制

谢明军[1]，段京奇[1]，马文冉[2]，贾玉红[2,*]
(1. 北京航空航天大学 大型飞机高级人才培训班，北京 100191；2. 北京航空航天大学 航空科学与工程学院，北京 100191)

摘 要：针对飞机全电防滑刹车系统具有较强的非线性特征，以及机电作动器（EMA）中存在的干扰不利于系统稳定性，提出一种有限时间预设性能反演滑模控制方法。在合理简化的基础上，建立含滑移率子系统和EMA子系统的飞机全电防滑刹车系统的数学模型，并引入有限时间预设性能函数；利用预设性能反演控制方法设计滑移率控制器，获得参考刹车压力控制率，保证滑移率跟踪误差在有限时间内收敛到预设范围内；为跟踪参考刹车压力，利用非奇异终端滑模控制方法设计EMA控制器，针对EMA中存在的干扰，设计扩张状态观测器估计，并在控制器中进行补偿，提高控制器的鲁棒性和控制精度；通过干跑道和冰跑道2种情况下的数值仿真验证所提方法控制效果。

关 键 词：防滑刹车系统；有限时间；预设性能；扩张状态观测器；机电作动器；滑移率
中图分类号：V221.3; TB553
文献标志码：A **文章编号**：1001-5965(2024)01-0260-08

防滑刹车系统作为飞机不可或缺的子系统，对于飞机航行最为关键的起降安全起着至关重要的作用。传统的液压刹车系统存在可维护性和可维修性较差、液压油容易泄露等缺点，而飞机全电刹车系统采用机电作动器(electromechanical actuator, EMA)，其体积小、易维护，且不存在油液泄露的危险，动态性能比传统的液压刹车系统更好，是下一代刹车系统的发展趋势[1-2]。

飞机的安全起飞和着陆离不开稳定和高效的防滑刹车控制策略，基于滑移率的控制策略是目前国际上研究的主要方向，滑移率控制的实质是指将实际滑移率保持在其最佳值，从而使结合系数达到最大值[3-4]。在现有的关于防滑刹车滑移率控制的文献中，已经提出多种控制方法，如模糊控制[5]、滑模控制[6]和神经网络控制[7]等，这些方法主要集中在防滑刹车系统的可控性上，在跟踪设定的最佳滑移率方面取得了一定的成果，但在稳定性和瞬态性能方面却贡献较少，也不能预先获得期望的控制行为。文献[8]基于障碍Lyapunov函数方法设计的控制器实现了对滑移率误差上界的约束，保证了刹车过程中的侧向稳定性，但对于收敛速度和稳态误差等性能指标考虑较少。文献[9]提出的预设性能控制方法通过约束跟踪误差的上下界保证收敛速度和超调量满足预先设定的条件，但不能让跟踪误差在有限时间内收敛到预设的约束区域内。文献[10-12]在文献[9]的基础上，设计了有限时间预设性能函数，可以实现在有限时间内的跟踪期望值。然而，关于具有预设性能的最佳滑移率有限时间跟踪控制的研究鲜有报道，此外，上述研究未考虑对机电作动系统中存在的干扰进行补偿。

针对现有文献的不足，本文先建立合理简化的全电飞机数学模型，并将其分为2部分：滑移率子系统和EMA子系统[13]。针对滑移率子系统，引入有限时间性能函数，对滑移率跟踪误差进行转换，基于Lyapunov理论设计反演控制器，得到刹车压

收稿日期：2022-04-06；录用日期：2022-06-03；网络出版时间：2022-06-10 08:17
网络出版地址：link.cnki.net/urlid/11.2625.V.20220608.1739.001
*通信作者. E-mail：jiayuhong@buaa.edu.cn

引用格式：谢明军，段京奇，马文冉，等. 基于预设性能的飞机全电刹车系统滑模控制[J]. 北京航空航天大学学报，2024，50(1)：260-267.
XIE M J，DUAN J Q，MA W R，et al. Sliding mode control for electric braking systems of aircraft based on prescribed performance[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(1)：260-267 (in Chinese).

力控制率, 使滑移率在瞬态和稳态性能上都能得到相应的约束, 并在有限时间内收敛到预设范围内。针对 EMA 子系统, 采用非奇异终端滑模控制, 保证参考压力能在有限的时间内被跟踪, 通过设计扩张状态观测器[14] 对 EMA 系统中的干扰进行估计与补偿, 提升系统的抗干扰能力。

1　数学模型

1.1　飞机地面模型

在建立飞机防滑刹车模型时, 做出以下合理假设: 飞机在地面滑跑过程中不受侧风的影响; 飞机前轮的摩擦系数恒定; 左右两侧受刹机轮制动效果一致。

基于上述假设, 在地面滑跑过程中, 飞机的受力分析如图 1 所示。

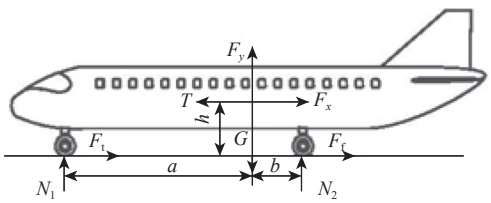


图 1　飞机地面滑跑受力分析

Fig. 1　Analysis of force on aircraft during ground running

根据牛顿第二运动定律推导的飞机纵向运动方程、垂直方向平衡方程和质心力矩平衡方程, 可得

$$\begin{cases} m\dot{V}_x = T - F_x - n_1F_t - n_2F_f \\ mg - F_y - n_1N_1 - n_2N_2 = 0 \\ n_1N_1a - n_2N_2b - n_2F_fh - n_1F_th = 0 \end{cases} \tag{1}$$

式中: m为飞机质量; V_x为飞机的纵向滑跑速度; T为发动机推力; F_x为空气阻力; F_t和F_f分别为前机轮和主机轮所受到的地面结合力; F_y为飞机升力; N_1和N_2分别为前机轮和主机轮所受到的支持力; n_1和n_2分别为前机轮和主轮机轮个数; g为重力加速度; a和 b 分别为飞机重心到前轮中心和主轮中心的水平距离; h为飞机中心高度。其中, T、F_x、F_y、F_t、F_f的计算式分别为

$$\begin{cases} T = T_0 + k_vV_x \\ F_x = \dfrac{1}{2}JC_xSV_x^2 \\ F_y = \dfrac{1}{2}JC_ySV_x^2 \\ F_t = \mu_fN_1 \\ F_f = \mu N_2 \end{cases} \tag{2}$$

式中: k_v为发动机推力系数; J为空气密度; C_x和C_y分别为空气阻力和飞机升力系数; S为飞机机翼面积; μ_f为不受刹的前机轮与跑道的结合系数; μ为受刹的主机轮与地面间的结合系数。

1.2　飞机机轮动力学模型

在刹车过程中, 刹车力矩作用在飞机的主机轮上, 主机轮的受力分析如图 2 所示, 其中, T_b 为刹车力矩。

刹车时主机轮的运动方程可描述为

$$I\dot{\omega} = F_fR_g - k_bp \tag{3}$$

式中: I为单个机轮的转动惯量; $\dot{\omega}$为飞机受刹机轮的角加速度; R_g为刹车半径; k_b为力矩转化系数; p为刹车压力。

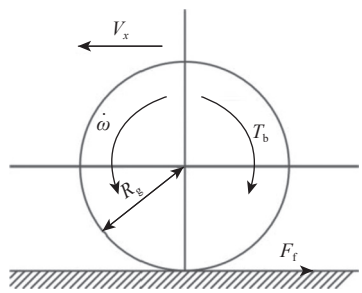


图 2　飞机机轮受力分析

Fig. 2　Analysis of force on aircraft wheel

1.3　滑移率模型

在主机轮受刹过程中, 主机轮的速度始终小于或等于飞机的纵向速度, 可将滑移率λ定义为

$$\lambda = (V_x - \omega r)/V_x \tag{4}$$

对式(4)求导可得

$$\dot{\lambda} = \frac{(1-\lambda)\dot{V}_x - \dot{\omega}R_g}{V_x} \tag{5}$$

将式(1)～式(3)式代入式(5), 整理后可得滑移率导数为

$$\dot{\lambda} = -\frac{\mu N_2R_g^2}{IV_x} + \frac{1-\lambda}{mV_x}(T - F_x - n_2\mu N_2 - n_1\mu_fN_1) + \frac{k_bR_g}{V_xI}p \tag{6}$$

为了便于控制器的设计, 式 (6) 可表示为

$$\dot{\lambda} = f(\lambda) + \frac{k_bR_g}{IV_x}p \tag{7}$$

式中: $f(\lambda) = -\dfrac{\mu N_2R_g^2}{IV_x} + \dfrac{1-\lambda}{mV_x}(T - F_x - n_2\mu_2N_2 - n_1\mu_1N_1)$

结合系数μ随飞机滑跑状态及地面条件的变化呈现非线性变化, 本文采用 Pacejka 结合系数模型[15] 描述为

$$\mu(\lambda) = D\sin(C\arctan(B\lambda)) \tag{8}$$

式中: D、C和B分别为峰值因子、刚度因子和曲线形状因子。

通过改变D、C、B的值, 可以建立不同跑道状态下的模型。3 种不同跑道状态下 Pacejka 模型参数如表 1 所示, 对应的μ-λ曲线如图 3 所示。

表1 Pacejka模型参数

Table 1 Pacejka model parameters

跑道状态	D	C	B
干跑道	0.8	1.534 4	14.032 6
湿跑道	0.4	2.019 2	8.209 8
冰跑道	0.2	2.087 5	7.201 8

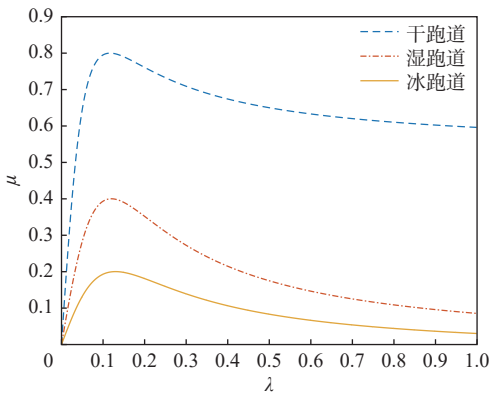


图3 μ-λ曲线

Fig. 3 μ-λ curves

1.4 机电作动器模型

EMA是一个复杂的机电系统，其主要组成部分包括无刷直流电机、减速齿轮、滚珠丝杠、压紧盘和刹车盘等[16]。其工作原理可描述为：由无刷直流电机通过减速齿轮驱动滚珠丝杠，压紧盘受到丝杠的驱动进而对刹车盘加压，从而产生刹车力矩，考虑到摩擦、齿隙等非线性特性影响，很难得到一个可以准确描述EMA动态行为的数学模型[17]，因此，将模型中参数摄动、齿隙和摩擦引起的扰动叠加视为复合干扰d，建立机电作动器的数学模型为

$$\dot{p}=\frac{C_{\mathrm{b}}L_0}{2\pi}\omega_{\mathrm{m}} \tag{9}$$

$$\dot{\omega}_{\mathrm{m}}=-\frac{L_0}{2\pi J_{\mathrm{m}}}p-\frac{B_{\mathrm{v}}}{J_{\mathrm{m}}}\omega_{\mathrm{m}}+\frac{k_{\mathrm{T}}}{J_{\mathrm{m}}}i+d \tag{10}$$

式中：C_{b}为刹车盘刚度系数；L_0为滚珠丝杠行程；ω_{m}为电机转速；J_{m}为转子转动惯量；k_{T}为电机转矩常数；i为电枢电流；B_{v}为黏滞阻尼系数；d为模型中参数摄动、齿隙和摩擦引起的扰动叠加形成的复合干扰，表示为

$$d=f_{\mathrm{p}}+f_{\mathrm{d}}+\Delta f$$

式中：f_{p}为齿轮传动间隙、联轴器间隙等传动部件在运行过程中周期性振荡引起的扰动；f_{d}为摩擦引起的扰动；Δf为参数摄动引起的扰动，$\Delta f=\Delta_1 p+\Delta_2\omega_{\mathrm{m}}$，$\Delta_1$、$\Delta_2$分别为由刹车压力$p$、电机转速$\omega_{\mathrm{m}}$参数摄动引起的扰动系数。

1.5 飞机全电刹车系统整体模型

联立式(7)、式(9)和式(10)，令$x_1=\lambda$，$x_2=p$，$x_3=\omega_{\mathrm{m}}$，飞机全电刹车系统整体模型可表示为

$$\begin{cases}\dot{x}_1=f(x_1)+\dfrac{k_{\mathrm{b}}R_{\mathrm{g}}}{IV_x}x_2\\ \dot{x}_2=\dfrac{C_{\mathrm{b}}L_0}{2\pi}x_3\\ \dot{x}_3=-\dfrac{L_0}{2\pi J_{\mathrm{m}}}x_2-\dfrac{B_{\mathrm{v}}}{J_{\mathrm{m}}}x_3+\dfrac{k_{\mathrm{T}}}{J_{\mathrm{m}}}i+d\\ y=\lambda\end{cases} \tag{11}$$

式中：y为系统输出。

为简化数学模型，做出以下合理假设：干扰d有界，即满足$|d|\leqslant D$；飞机的纵向滑跑速度V_x、刹车压力x_2、电机转速x_3和电枢电流i均可测[8]。

1.6 有限时间预设性能函数

针对飞机全电刹车系统，有限时间预设性能控制是指将滑移率误差$e(t)$在有限时间内收敛到一个预先设定的可调节边界范围内，并且期望最大超调量小于设计值，给出有限时间性能函数定义。

定义 针对光滑连续函数$\varphi(t)$：$\mathbf{R}_+\to\mathbf{R}_+$，需满足以下3个条件：

1）$\varphi(t)>0$且$\dot{\varphi}(t)\leqslant 0$

2）$\lim\limits_{t\to T_{\mathrm{f}}}\varphi(t)=\varphi_\infty>0$

3）当$t>T_{\mathrm{f}}$时，$\varphi(t)=\varphi_\infty$

对于防滑刹车系统，有限时间性能函数选择为[10]

$$\varphi(t)=\begin{cases}(\varphi_0-\varphi_\infty)(1-t/T_{\mathrm{f}})^{l+1}+\varphi_\infty & 0\leqslant t\leqslant T_{\mathrm{f}}\\ \varphi_\infty & t>T_{\mathrm{f}}\end{cases} \tag{12}$$

式中：$l\geqslant 1$；φ_0、φ_∞为正数；T_{f}为性能函数收敛的期望时间。

预设性能过程可以描述为

$$\underline{e}(t)<e(t)<\bar{e}(t) \tag{13}$$

式中：$\underline{e}(t)=-\delta_{\min}\varphi(t)$，$\bar{e}(t)=\delta_{\max}\varphi(t)$，$\delta_{\min}$和$\delta_{\max}$为大于0的设计参数，且初始值满足$\underline{e}(0)<e(0)<\bar{e}(0)$。

预设性能函数历时曲线如图4所示。

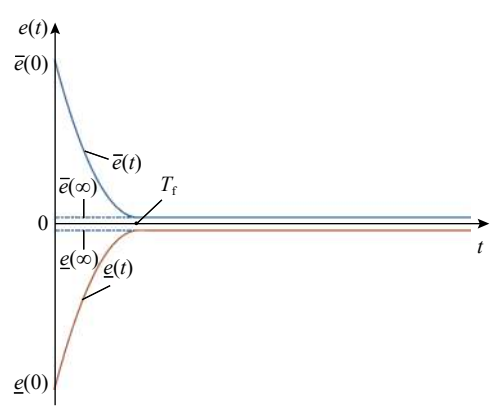


图4 预设性能函数历时曲线

Fig. 4 Duration curves of prescribed performance functions

2　防滑刹车控制器设计

2.1　滑移率控制器设计

防滑刹车系统的控制目标是保证实际滑转率x_1能够跟踪期望滑移率λ_{d}, 故定义滑移率误差为$e=x_1-\lambda_{\mathrm{d}}$, 则误差导数为

$$\dot{e}=\dot{x}_1=f(x_1)+\frac{k_{\mathrm{b}}R_{\mathrm{g}}}{IV_x}x_2 \tag{14}$$

采用预设性能反演控制方法, 能使滑移率误差在暂态和稳态性能上都能得到相应的约束, 满足规定的性能。针对式(13)直接处理难度较大的问题, 构造非线性变化为

$$\varepsilon=f(e,\bar{e},\underline{e})=\frac{e}{(\bar{e}-e)(e-\underline{e})} \tag{15}$$

式中: ε为转换后的误差变量

分析可知式(15)满足:

1) 当且仅当$e(t)=0$时, $\varepsilon(t)=0$

2) $\lim\limits_{e\to\bar{e}}\varepsilon(t)=+\infty\qquad\lim\limits_{e\to\underline{e}}\varepsilon(t)=+\infty$

ε与原误差e相关, 且误差e越大, ε也越大; 根据性能函数单调递减的性质和转换后的误差变量所满足的条件, 在满足$\underline{e}(0)<e(0)<\bar{e}(0)$的初始条件下, 若误差变量收敛, 则原误差将收敛并满足式(13)所要求的预设性能目标。

针对本文研究的防滑刹车系统, 选取李雅普诺夫函数为

$$V_1=\frac{1}{2}\varepsilon^2 \tag{16}$$

将式(15)带入V_1中并求导, 整理可得

$$\begin{aligned}\dot{V}_1=&\frac{\left[f(x_1)+\dfrac{k_{\mathrm{b}}R_{\mathrm{g}}}{IV_x}x_2\right](\bar{e}-e)(e-\underline{e})e}{(\bar{e}-e)^3(e-\underline{e})^3}-\\&\frac{e^2(e-\underline{e})\left[\dot{\bar{e}}-f(x_1)-\dfrac{k_{\mathrm{b}}R_{\mathrm{g}}}{IV_x}x_2\right]}{(\bar{e}-e)^3(e-\underline{e})^3}+\\&\frac{e^2(\bar{e}-e)\left[f(x_1)+\dfrac{k_{\mathrm{b}}R_{\mathrm{g}}}{IV_x}x_2-\dot{\underline{e}}\right]}{(\bar{e}-e)^3(e-\underline{e})^3}\end{aligned} \tag{17}$$

设计参考刹车压力虚拟控制量为

$$\begin{aligned}x_2^{\ *}=&\frac{k_1e(\bar{e}-e)(e-\underline{e})V_xI}{k_{\mathrm{b}}R_{\mathrm{g}}\left[\bar{e}(e-\underline{e})-e(\bar{e}-e)\right]}+\\&\frac{e\dot{\bar{e}}(e-\underline{e})V_xI}{k_{\mathrm{b}}R_{\mathrm{g}}\left[\bar{e}(e-\underline{e})-e(\bar{e}-e)\right]}+\\&\frac{e\dot{\underline{e}}(\bar{e}-e)V_xI}{k_{\mathrm{b}}R_{\mathrm{g}}\left[\bar{e}(e-\underline{e})-e(\bar{e}-e)\right]}-\frac{f(x_1)V_xI}{k_{\mathrm{b}}R_{\mathrm{g}}}\end{aligned} \tag{18}$$

式中: k_1为大于 0 的设计参数, 则式(17)可化简为

$$\dot{V}_1=-\frac{k_1e^2}{(\bar{e}-e)^2(e-\underline{e})^2}\leqslant 0 \tag{19}$$

根据分析可得, 李雅普诺夫函数V_1正定, $\dot{V}_1$半负定, 故采用预设性能反演控制方法设计的滑移率控制器能在有限时间内收敛。

2.2　EMA 控制器设计

在建立了滑移率控制器后, 需要有一个 EMA 控制器来实现相应的制动压力控制。本文采用非奇异终端滑模控制方法设计电枢电流的控制率i, 以便p在有限的时间内跟踪给定的压力$x_2^{\ *}$, 先定义压力误差变量为$e_p=x_2-x_2^{\ *}$, 则误差变量导数和二阶导数分别为

$$\dot{e}_p=\frac{C_{\mathrm{b}}L_0}{2\pi}x_3-\dot{x}_2^* \tag{20}$$

$$\ddot{e}_p=\frac{C_{\mathrm{b}}L_0}{2\pi}\left(-\frac{L_0}{2\pi J_{\mathrm{m}}}x_2-\frac{B_{\mathrm{v}}}{J_{\mathrm{m}}}x_3+\frac{k_{\mathrm{T}}}{J_{\mathrm{m}}}i+d\right)-\ddot{x}_2^* \tag{21}$$

设计非奇异终端滑模模态[18]

$$s=e_p+\frac{1}{\beta}\dot{e}_p^{\frac{q}{n}} \tag{22}$$

式中: $\beta>0$; q、n为正奇数, $1<q/n<2$。

对式(22)求导可得

$$\begin{aligned}\dot{s}=&\dot{e}_p+\frac{C_{\mathrm{b}}L_0n}{2\pi\beta q}\dot{e}_p^{\frac{q}{n}-1}\left(-\frac{L_0}{2\pi J_{\mathrm{m}}}x_2-\frac{B_{\mathrm{v}}}{J_{\mathrm{m}}}x_3\right)+\\&\frac{n}{\beta q}\dot{e}_p^{\frac{q}{n}-1}\left(\frac{C_{\mathrm{b}}L_0k_{\mathrm{T}}}{2\pi J_{\mathrm{m}}}i+\frac{C_{\mathrm{b}}L_0}{2\pi}d-\ddot{x}_2^*\right)\end{aligned} \tag{23}$$

得到等效控制项为

$$i_{\mathrm{eq}}=\left(\frac{\beta q}{p}\dot{e}_p^{-\frac{q}{n}}+\ddot{x}_2^*\right)\frac{2\pi J_{\mathrm{m}}}{C_{\mathrm{b}}L_0k_{\mathrm{T}}}+\frac{L_0}{2\pi k_{\mathrm{T}}}x_2+\frac{B_{\mathrm{v}}}{J_{\mathrm{m}}}x_3 \tag{24}$$

设计切换控制项为

$$i_{\mathrm{n}}=\left[-(D+\eta)\mathrm{sgn}(s)-k_2s\right]\frac{J_{\mathrm{m}}}{k_{\mathrm{T}}} \tag{25}$$

式中: η和k_2为大于 0 的设计参数; i为电枢电流的控制率, 计算式为

$$\begin{aligned}i=i_{\mathrm{eq}}+i_{\mathrm{n}}=&\left(\frac{\beta q}{p}\dot{e}_p^{-\frac{q}{n}}+\ddot{x}_2^*\right)\frac{2\pi J_{\mathrm{m}}}{C_{\mathrm{b}}L_0k_{\mathrm{T}}}+\frac{L_0}{2\pi k_{\mathrm{T}}}x_2+\\&\frac{B_{\mathrm{v}}}{k_{\mathrm{T}}}x_3+\frac{J_{\mathrm{m}}}{k_{\mathrm{T}}}\left[-(D+\eta)\mathrm{sgn}(s)-k_2s\right]\end{aligned} \tag{26}$$

将式(26)中的i带入式(23)中, 求得

$$\dot{s}=\frac{C_{\mathrm{b}}L_0n}{2\pi\beta q}\dot{e}_p^{\frac{q}{n}-1}\left[-D\mathrm{sgn}(s)-\eta\mathrm{sgn}(s)-k_2s+d\right] \tag{27}$$

构造李雅普诺夫函数为

$$V_2=V_1+\frac{1}{2}s^2 \tag{28}$$

对V_2求导可得

$$\begin{aligned}\dot{V}_2=&\dot{V}_1+s\dot{s}\leqslant\\&\dot{V}_1-\frac{C_{\mathrm{b}}L_0n}{2\pi\beta q}\dot{e}_p^{\frac{q}{n}-1}\left(\eta|s|+k_2s^2\right)\leqslant 0\end{aligned} \tag{29}$$

分析可知, 李雅普诺夫函数V_2正定, $\dot{V}_2$半负定, 系统稳定。

为提高控制器对干扰d的适应性，在式(26)的基础上，设计扩张状态观测器估计EMA系统干扰的实时作用值，并在反馈中给予补偿。

将干扰d作为EMA系统的一个状态进行扩张，可得

$$\begin{cases}\dot{x}_2=\dfrac{C_{\mathrm{b}}L_0}{2\pi}x_3\\ \dot{x}_3=-\dfrac{L_0}{2\pi J_{\mathrm{m}}}x_2-\dfrac{B_{\mathrm{v}}}{J_{\mathrm{m}}}x_3+x_4+\dfrac{k_{\mathrm{T}}}{J_{\mathrm{m}}}i\\ \dot{x}_4=\dot{d}\end{cases}\tag{30}$$

设计扩张状态观测器为

$$\begin{cases}e_2=x_2-\hat{x}_2\\ \dot{\hat{x}}_2=\dfrac{C_{\mathrm{b}}L_0}{2\pi}\hat{x}_3+\beta_1e_2\\ \dot{\hat{x}}_3=-\dfrac{L_0}{2\pi J_{\mathrm{m}}}\hat{x}_2-\dfrac{B_{\mathrm{v}}}{J_{\mathrm{m}}}\hat{x}_3+\hat{x}_4+\dfrac{k_{\mathrm{T}}}{J_{\mathrm{m}}}i+\beta_2e_2\\ \dot{\hat{x}}_4=\beta_3e_2\end{cases}\tag{31}$$

式中：e_2为压力观测误差；$\hat{x}_2$、$\hat{x}_3$、$\hat{x}_4$分别为x_2、x_3、x_4的观测值；β_1、β_2、β_3为扩张状态观测器的可调增益，其与观测器带宽ω_0存在以下关系$\beta_1=3\omega_0$、$\beta_2=3\omega_0^2$、$\beta_3=\omega_0^3$。其中，增大ω_0值将使得ESO对误差e_2更加敏感，即对输出的变化响应更加迅速，但过大的带宽会扩大噪声的影响，引起观测器输出值的高频抖动。

通过设计扩张状态观测器，得到干扰d的估计值$\hat{x}_4$，将其代入式(26)中，可得

$$\begin{aligned}i=&\left(\frac{\beta q}{p}\dot{e}_p^{-\frac{q}{n}}+\dot{x}_2^*\right)\frac{2\pi J_{\mathrm{m}}}{C_{\mathrm{b}}L_0k_{\mathrm{T}}}+\frac{L_0}{2\pi k_{\mathrm{T}}}x_2+\\ &\frac{B_{\mathrm{v}}}{k_{\mathrm{T}}}x_3-\frac{J_{\mathrm{m}}}{k_{\mathrm{T}}}\hat{x}_4-\frac{J_{\mathrm{m}}}{k_{\mathrm{T}}}\eta\mathrm{sgn}(s)-\frac{k_2J_{\mathrm{m}}}{k_{\mathrm{T}}}s\end{aligned}\tag{32}$$

建立滑移率控制器和基于扩张状态观测器的EMA控制器后，得到飞机防滑刹车系统的完整控制结构如图5所示。

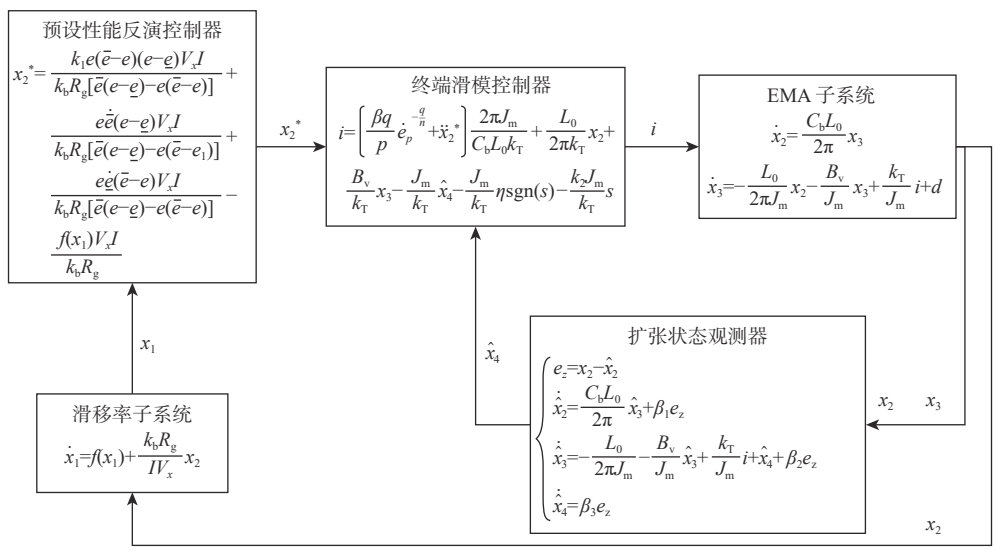


图5 全电飞机防滑刹车系统控制结构

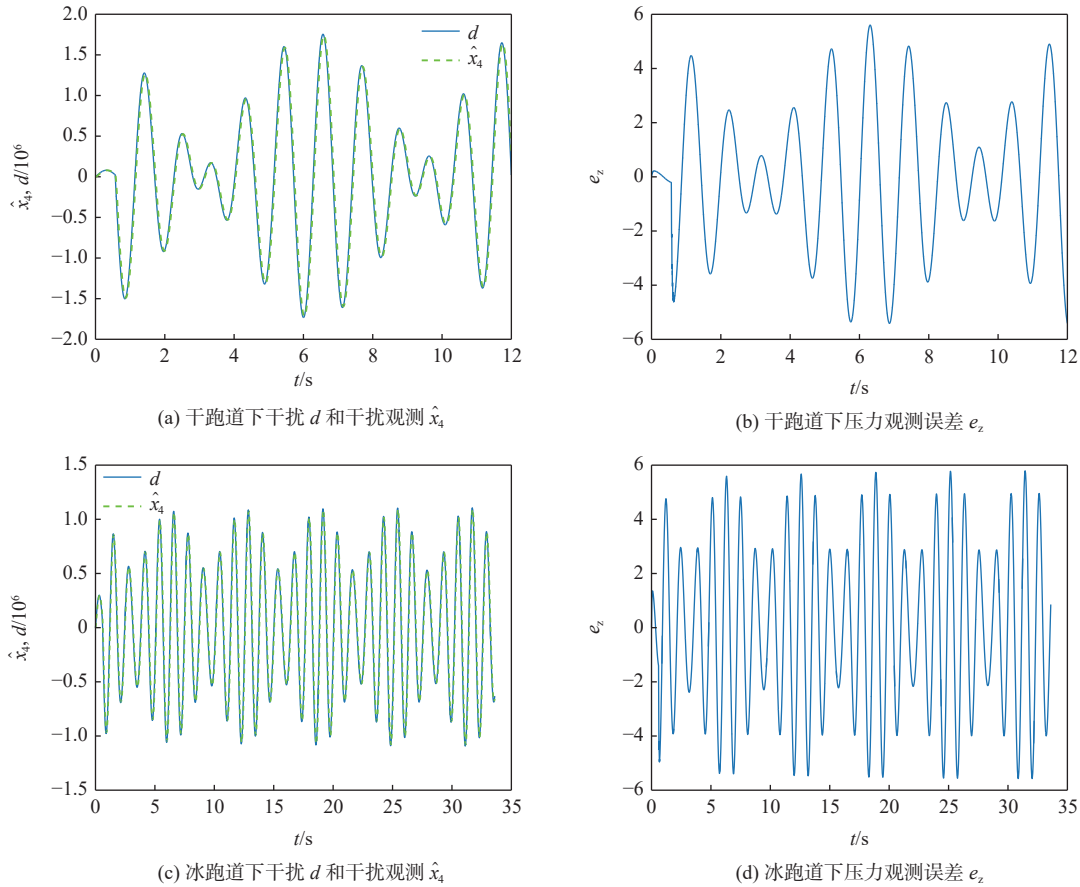
Fig. 5 Structure of electric antiskid braking system of aircraft

3 仿真分析

为验证本文设计的防滑刹车控制率的可靠性，使用MATLAB软件在跑道状态良好的干沥青跑道和摩擦系数较低的冰跑道中进行仿真，根据Pacejka模型计算出干沥青跑道的期望滑移率为0.117，冰跑道的期望滑移率为0.129，设置仿真步长为0.001 s。

飞机初始速度为65 m/s，机轮初始角速度为130 rad/s，系统模型部分参数分别为：m=145 47 kg，a=1.84 m，b=10.08 m，h=5 m，R_{g}=0.5 m，在干跑道中控制器和扩张状态观测器参数设定为：$l=1$，$\varphi_0=0.6$，$\varphi_\infty=0.005$，$T_{\mathrm{f}}=0.3$，$\delta_{\min}=1$，$\delta_{\max}=1$，$k_1=1$，$\beta=1.2$，$q=19$，$n=11$，$\eta=0.01$，$k_2=30$，$\omega_0=120$；在冰跑道中$\beta=2.5$，$k_2=20$，$\omega_0=100$，其他参数与干沥青跑道保持一致。将EMA子系统中的干扰设定为：$f_{\mathrm{p}}=800\,000\sin(5t)$，$\Delta f=0.1x_2\sin(5t)+0.2x_3\sin(6t)$，$f_{\mathrm{d}}=7\,500\tanh(20x_3)+\tanh(20x_3)-\tanh(2x_3)$，为防止初始扰动过大产生冲击作用，对初始扰动进行了光滑处理。

不同跑道情况下干扰观测仿真结果如图6所示。由图可知，扩张状态观测器能够实现对干扰d的稳定跟踪，从而有效的对EMA控制器进行补偿，故压力观测误差较小，除初始时刻外，压力观测误差控制在[−6,6]内。能够实现对干扰的稳定跟踪，且观测误差曲线和扰动观测曲线光滑，并未出



(a) 干跑道下干扰 d 和干扰观测 $\hat{x}_4$　　(b) 干跑道下压力观测误差 e_z

(c) 冰跑道下干扰 d 和干扰观测 $\hat{x}_4$　　(d) 冰跑道下压力观测误差 e_z

图 6　干扰观测仿真结果

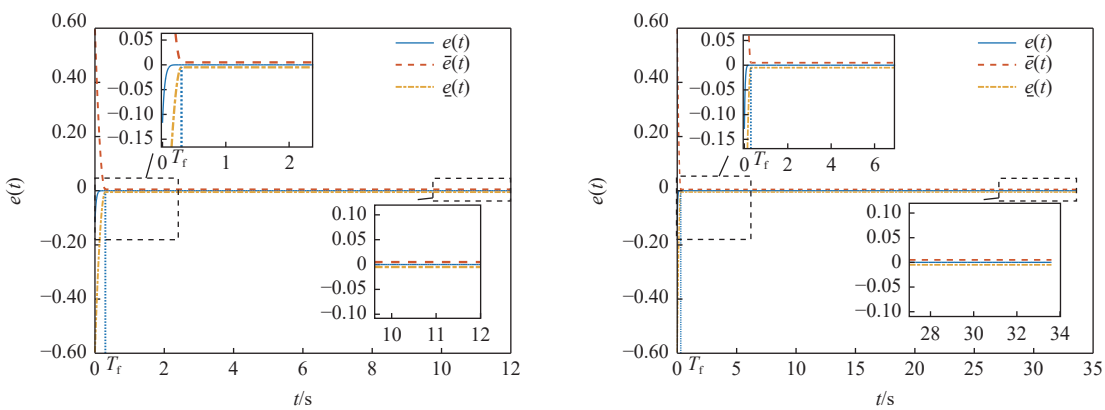
Fig. 6　Simulation results of disturbance observation

现高频抖动现象, 表明ω_0的取值合适。

飞机在不同跑道滑跑中的滑移率跟踪误差如图 7 所示。由图可知, 飞机在滑行过程中, 滑移率跟踪误差能在预先设定的 0.3 s 内收敛到预先设定的稳态误差 [−0.005,0.005] 内, 收敛速度快, 且跟踪误差一直处于式(13)定义的误差界限$\underline{e}(t) < e(t) < \bar{e}(t)$中, 无明显超调。

不同跑道情况下的飞机速度 V_x 和机轮速度 wR_x 仿真对比曲线如图 8 所示。由图可知, 除初始时机轮速度下降较快外, 其他时刻飞机速度和机轮速度下降都比较平缓, 未出现较大的波动。

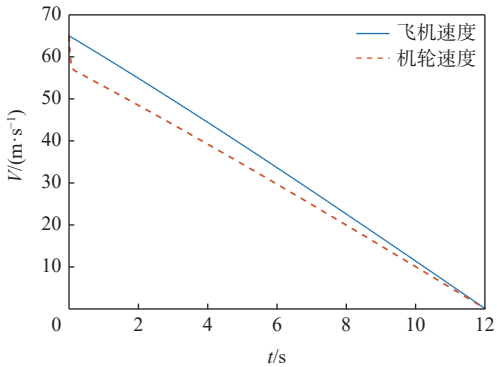
不同跑道情况下的滑移率如图 9 所示。由图可知, 在滑跑过程中, 在干跑道上, 飞机在 0.3 s 内稳定在最佳滑移率 0.117; 在冰跑道上, 飞机在 0.3 s 内稳定在最佳滑移率 0.129, 且在低速时也能保持相应的最佳滑移率, 此时机轮与地面的结合系数最大, 从而获得最大结合力, 提高刹车效率。



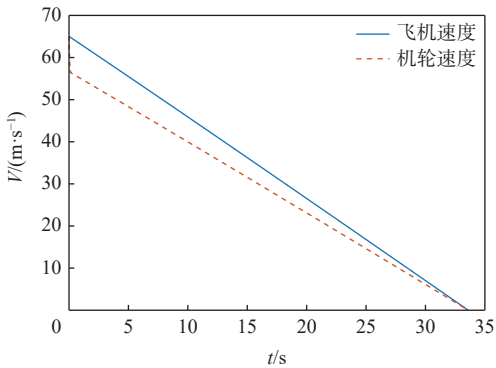
(a) 干跑道　　(b) 冰跑道

图 7　不同跑道下滑移率误差仿真结果

Fig. 7　Simulation results of slip ratio error in different runways



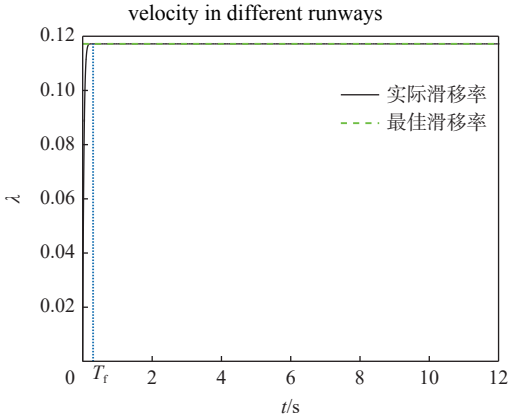
(a) 干跑道



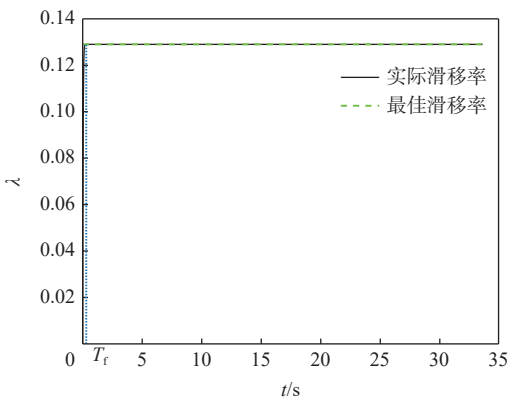
(b) 冰跑道

图 8 不同跑道下飞机速度和机轮速度仿真结果

Fig. 8 Simulation results of aircraft velocity and wheel velocity in different runways



(a) 干跑道



(b) 冰跑道

图 9 不同跑道下滑移率仿真结果

Fig. 9 Simulation results of slip ratio in different runways

4 结 论

1）本文方法能使滑移率跟踪误差在预先设定的 0.3 s 内收敛到预先设定的范围 [−0.005,0.005]，即在 0.3 s 内达到最佳滑移率，收敛速度快，且不出现震荡现象，超调量小，稳定性好。

2）本文方法能适应不同的跑道状态，在跑道状态良好的干跑道、摩擦系数较低的冰跑道滑行和飞机处于低速段滑行时，都能保持最佳滑移率状态，刹车效率得到提高，且在不同跑道时调整控制器的参数较为简单。

3）考虑到 EMA 作动器的非线性特性的影响，采用基于扩张状态观测器的非奇异终端滑模控制方法设计电枢电流控制率，能准确估计扰动，并有效跟踪参考刹车压力。

参考文献（References）

[1] 李繁飙, 黄培铭, 阳春华, 等. 基于非线性干扰观测器的飞机全电刹车系统滑模控制设计[J]. 自动化学报, 2021, 47(11): 2557-2569.
LI F B, HUANG P P, YANG C H, et al. Sliding mode control design of aircraft electric brake system based on nonlinear disturbance observer[J]. Acta Automatica Sinica, 2021, 47(11): 2557-2569(in Chinese).

[2] 付龙飞. 飞机防滑刹车系统的非线性控制技术应用研究 [D]. 西安: 西北工业大学, 2017.
FU L F. Research on application of nonlinear control technology in aircraft anti-skid braking system [D]. Xi'an: Northwestern Polytechnical University, 2017(in Chinese).

[3] 陈晓雷, 林辉, 李兵强, 等. 飞机全电刹车系统滑移率约束反演滑模控制[J]. 上海交通大学学报, 2015, 49(12): 1855-1861.
CHEN X L, LIN H, LI B Q, et al. Backstepping sliding mode control for aircraft electric braking systems with slip ratio constraint[J]. Journal of Shanghai Jiao Tong University, 2015, 49(12): 1855-1861(in Chinese).

[4] 付龙飞, 田广来, 王鹏, 等. 飞机防滑刹车系统滑移率自适应滑模控制研究[J]. 西北工业大学学报, 2015, 33(5): 770-774.
FU L F, TIAN G L, WANG P, et al. Adaptive sliding mode control strategy for the aircraft antiskid braking system based on slip ratio[J]. Journal of Northwestern Polytechnical University, 2015, 33(5): 770-774(in Chinese).

[5] 韩伟健. 飞机全电刹车机电作动系统的建模与控制方法研究 [D]. 西安: 西北工业大学, 2015.
HAN W J. Research on modeling and control method of electro-mechanical actuating system of aircraft electric brake [D]. Xi'an: Northwestern Polytechnical University, 2015(in Chinese).

[6] 杜成林. 飞机防滑刹车系统变结构控制研究 [D]. 沈阳: 东北大学, 2008.
DU C L. Research on variable structure control of aircraft anti-skid brake system [D]. Shenyang: Northeastern University, 2008(in Chinese).

[7] 田广来, 谢利理, 岳开宪, 等. 飞机防滑刹车系统的最佳滑移率式控制方法研究[J]. 航空学报, 2005, 26(4): 461-464.

TIAN G L, XIE L L, YUE K X, et al. Study on optimal control method of an aircraft anti-skid braking system based on slip-ratio[J]. Acta Aeronautica et Astronautica Sinica, 2005, 26(4): 461-464(in Chinese).

[8] 何恒, 吴瑞祥. 改进的 BP 神经网络在飞机防滑刹车系统的应用[J]. 北京航空航天大学学报, 2004, 30(6): 561-564.
HE H, WU R X. Improved BP neural network in design of aircraft antiskid braking system[J]. Journal of Beijing University of Aeronautics and Astronautics, 2004, 30(6): 561-564(in Chinese).

[9] BECHLIOULIS C P, ROVITHAKIS G A. Robust adaptive control of feedback linearizable mimo nonlinear systems with prescribed performance[J]. IEEE Transactions on Automatic Control, 2008, 53(9): 2090-2099.

[10] TAN J, DONG Y F, SHAO P Y, et al. Anti-saturation adaptive fault-tolerant control with fixed-time prescribed performance for UAV under AOA asymmetric constraint[J]. Aerospace Science and Technology, 2022, 120: 107264.

[11] ZHUANG M L, TAN L G, LI K H, et al. Fixed-time formation control for spacecraft with prescribed performance guarantee under input saturation[J]. Aerospace Science and Technology, 2021, 119: 107176.

[12] ZHOU T, LIU C, LIU X, et al. Finite-time prescribed performance adaptive fuzzy control for unknown nonlinear systems[J]. Fuzzy Sets and Systems, 2021, 402: 16-34.

[13] ZHANG X, LIN H. Backstepping fuzzy sliding mode control for the antiskid braking system of unmanned aerial vehicles[J]. Electronics, 2020, 9(10): 1731.

[14] GAO Z Q. Scaling and bandwidth-parameterization based controller tuning [C]//Proceedings of the 2003 American Control Conference. Piscataway: IEEE Press, 2003: 4989-4996.

[15] 孙辉, 闫建国, 屈耀红. 输入输出受限的无人机防滑刹车系统容错控制[J]. 北京航空航天大学学报, 2017, 43(6): 1132-1140.
SUN H, YAN J G, QU Y H. Fault-tolerant control of UAV anti-skid braking system with input and output constraints[J]. Journal of Beijing University of Aeronautics and Astronautics, 2017, 43(6): 1132-1140(in Chinese).

[16] 陈晓雷, 林辉, 马冬麒. 无人机全电式自主刹车系统滑模极值搜索控制[J]. 控制理论与应用, 2015, 32(11): 1439-1448.
CHEN X L, LIN H, MA D L. Sliding-mode extremum-seeking control for all-electric active braking system in unmanned aerial vehicle[J]. Control Theory & Applications, 2015, 32(11): 1439-1448(in Chinese).

[17] LIANG B, ZHU Y Q, LI Y R, et al. Adaptive nonsingular fast terminal sliding mode control for braking systems with electro-mechanical actuators based on radial basis function[J]. Energies, 2017, 10(10): 1637.

[18] 冯勇, 鲍晟, 余星火. 非奇异终端滑模控制系统的设计方法[J]. 控制与决策, 2002, 17(2): 194-198.
FENG Y, BAO S, YU X H. Design method of non-singular terminal sliding mode control systems[J]. Control and Decision, 2002, 17(2): 194-198(in Chinese).

Sliding mode control for electric braking systems of aircraft based on prescribed performance

XIE Mingjun[1], DUAN Jingqi[1], MA Wenran[2], JIA Yuhong[2,*]

(1. Advanced Training Center of Large Aircraft, Beihang University, Beijing 100191, China;
2. School of Aeronautic Science and Engineering, Beihang University, Beijing 100191, China)

Abstract: The aircraft's electric antiskid braking system has strong nonlinear characteristics, and the interference in the electromechanical actuator (EMA) is not conducive to system stability, a sliding mode backstepping control method based on finite-time prescribed performance control is proposed. An electric antiskid braking system model with slip rate subsystem and EMA subsystem is established based on reasonable simplification with the finite time convergence prescribed performance function introduced. Based on the backstepping design, a finite-time prescribed performance control algorithm is designed to generate reference braking pressure for the slip rate controller, which can limit the slip tracking error in the predefined bounds within a finite time by appropriately choosing design parameters. Using a control algorithm for the nonsingular terminal sliding mode, the EMA controller is designed. To improve EMA subsystem robustness and precision, an extended state observer is designed, which can estimate and compensate the disturbances. Simulation results verity the control effects of the proposed method for dry and ice runways.

Keywords: antiskid braking system; finite time; prescribed performance; extended state observer; electromechanical actuator; slip ratio

Received: 2022-04-06; **Accepted**: 2022-06-03; **Published Online**: 2022-06-10 08:17
URL: link.cnki.net/urlid/11.2625.V.20220608.1739.001

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0173

基于平滑矩阵集优化的双基地MIMO雷达相干目标DOD和DOA联合估计

游致远[1,2]，胡国平[1,*]，周豪[1]，郑桂妹[1]

(1. 空军工程大学 防空反导学院，西安 710051；2. 空军工程大学 研究生院，西安 710051)

摘 要：针对在进行相干目标测向时出现的协方差矩阵秩亏现象，双基地多输入多输出(MIMO)雷达探测相干目标会出现精度差、角度分辨率低等问题，提出一种基于分集平滑优化对相干目标波离方向角(DOD)和波达方向角(DOA)进行联合估计的算法。所提算法能通过分集平滑降维协方差矩阵，构造更有效的平滑矩阵集对相干目标进行联合角度估计。将发射阵列和接收阵列进行分集平滑并构建发射子矢量，计算其协方差矩阵；将每个发射子矢量的协方差矩阵进行平滑重构而后进行加权处理得到前向平滑矩阵；对前向平滑矩阵进行共轭翻转等变换后，得到前后向平滑矩阵，再对其进行奇异值(SVD)分解；利用降维多重信号分类(RD-MUSIC)算法对信号进行空间谱估计并实现DOD和DOA自动配对。所提算法有效提高双基地MIMO雷达对相干目标的测向性能，并通过实验证明所提算法的可行性。

关 键 词：双基地雷达；MIMO雷达；相干目标；分集平滑；角度联合估计

中图分类号：TN953^{+}.5

文献标志码：A **文章编号**：1001-5965(2024)01-0268-08

随着多输入多输出(multiple input multiple outout, MIMO)通信理论不断发展和成熟，Rabideau和Parker[1]于2004年提出MIMO雷达这个概念。MIMO雷达因为有着波形分集增益[2-4]、更佳的参数分辨能力和扩展虚拟阵列的优点被众多学者提及和研究。在研究信号处理的过程中波达方向角(direction of arrival, DOA)估计[5-12]是近年来与MIMO雷达联系紧密的课题方向。在对相干目标进行DOA估计时，若使用常规的子空间类算法，会导致信号子空间的矢量位移至噪声子空间中，从而使满秩的协方差矩阵发生秩亏现象，无法准确估计相干信源数目。学者们为了解决这一问题，提出2种算法，分别是非降维类处理算法和降维类处理算法，前者主要包含有频域平滑算法[13]、Toeplitz方法[14]，后者主要包含有经典的空间平滑算法[15]和矩阵重构类算法[13]，都可以有效解决相干目标测向这一问题。文献[16]提出一种基于匹配追踪信号分解和$\ell_{1,2}$-MUSIC算法的混合算法，在提高测角分辨率的同时也提升了阵列对相干信源的鲁棒性。

在双基地MIMO雷达[17-18]相干目标测向这一领域，近年来不断有学者提出新的研究思路和研究方法。文献[19]提出一种经典的基于双基地MIMO雷达单快拍下的重构满秩Toeplitz矩阵方法，可以有效的解决相干信源造成秩亏的问题，并得到了目标的DOD和DOA，但该方法存在复杂度大的缺点。文献[20]在文献[19]的基础上，提出一种复数域转化方法，将复数域转为实数域进行计算，并利用复杂度更小的旋转不变技术估计信号参数

收稿日期：2022-03-21；**录用日期**：2022-06-20；**网络出版时间**：2022-06-24 16:53
网络出版地址：link.cnki.net/urlid/11.2625.V.20220623.1720.005
基金项目：国家自然科学基金(62071476)
* **通信作者**. E-mail：hgp6068@163.com

引用格式：*游致远，胡国平，周豪，等. 基于平滑矩阵集优化的双基地MIMO雷达相干目标DOD和DOA联合估计[J]. 北京航空航天大学学报，2024，50(1)：268-275. YOU Z Y，HU G P，ZHOU H，et al. Joint DOA and DOD estimation of bistatic MIMO radar coherent targets based on smoothing matrix sets optimization[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(1)：268-275 (in Chinese).*

(estimating signal parameter via rotational invariance techniques, ESPRIT)算法进行空间谱估计, 有效地降低了算法复杂度。文献 [21] 提出一种针对双基地 MIMO 雷达的空间差分平滑算法, 此算法是将传统的空间平滑改进成了非对称空间平滑, 平滑方式变成了基于阵列旋转不变式的发射分集平滑, 得到了测向效果更好的目标波离方向角(direction of departure, DOD)和 DOA, 然而这种算法的缺点是相干信源估计数目受到限制。文献 [22] 提出一种联合发射阵列和接收阵列的分集平滑算法, 该算法可以获得更多的协方差矩阵进行空间平滑, 得到更精确的目标 DOD 和 DOA, 但该算法没有利用每个协方差矩阵的全部信息, 使得信息利用率较低。文献 [23] 在文献 [22] 的基础上变换了平滑方式, 较多地利用协方差矩阵信息, 但是算法计算量过大。本文在文献 [22-23] 的基础上提出了一种基于分集平滑优化对相干目标 DOD 和 DOA 进行联合估计的算法, 通过二维协方差矩阵的降维处理、重组协方差矩阵中的子矩阵块, 重新构建了平滑矩阵集, 与文献 [22] 比较, 本文算法充分地利用了更多的协方差矩阵信息, 提高了信息利用率, 能够得到更加准确的目标 DOD 和 DOA。通过仿真实验, 验证本文算法分别在角度分辨率、空间谱估计性能和测角精度 3 个方面均优于文献 [21] 和文献 [22] 提出的算法, 更加精准有效地估计出相干目标的 DOD 和 DOA。

1　双基地均匀线阵 MIMO 雷达模型

双基地[22-24]均匀线阵 MIMO 雷达模型由阵列数为 M 的发射阵列和阵列数为 N 的接收阵元组成, 如图 1 所示, 阵元间距 $d=\lambda/2$, λ 为信号波长。发射阵列和接收阵列的位置集合分别可表示为

$$S_{\mathrm{t}}=\{0,d,2d,\cdots,(M-1)d\} \tag{1}$$

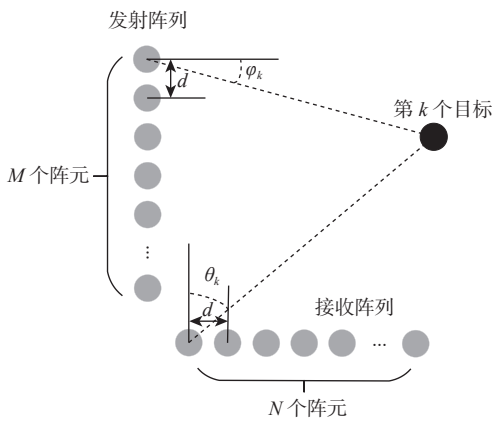


图 1　双基地均匀线阵 MIMO 雷达结构

Fig. 1　Structure of bistatic uniform linear array MIMO radar

$$S_{\mathrm{r}}=\{0,d,2d,\cdots,(N-1)d\} \tag{2}$$

若已知空域中有 K 个相干目标, 目标的方位信息可以用 2 组参数来表示, 分别为: φ_k、θ_k, k=1, 2, ⋯, K。其中 φ_k 表示第 k 个目标的 DOD, θ_k 表示第 k 个目标的 DOA。则经过匹配滤波后的回波信号模型为

$$\begin{aligned}\boldsymbol{x}(t)&=\left[\boldsymbol{a}_{\mathrm{t}}(\varphi)\otimes\boldsymbol{a}_{\mathrm{r}}(\theta_1),\boldsymbol{a}_{\mathrm{t}}(\varphi)\otimes\boldsymbol{a}_{\mathrm{r}}(\theta_2),\cdots,\boldsymbol{a}_{\mathrm{t}}(\varphi_K)\otimes\boldsymbol{a}_{\mathrm{r}}(\theta_K)\right]\\&\boldsymbol{s}(t)+\boldsymbol{n}(t)=(\boldsymbol{A}_{\mathrm{t}}\circ\boldsymbol{A}_{\mathrm{t}})\boldsymbol{s}(t)+\boldsymbol{n}(t)=\boldsymbol{A}\boldsymbol{s}(t)+\boldsymbol{n}(t)\end{aligned} \tag{3}$$

式中: $\boldsymbol{s}(t)=[s_1(t),s_2(t),\cdots,s_K(t)]^{\mathrm{T}}$, $s_k(t)$为第 k 个目标的发射系数; $\boldsymbol{n}(t)$ 为高斯白噪声矢量; ∘和⊗分别表示 Khatri-Rao 积和 Kronecker 积; $\boldsymbol{A}=\boldsymbol{A}_{\mathrm{t}}\circ\boldsymbol{A}_{\mathrm{r}}$表示回波信号模型的导向矩阵, $\boldsymbol{A}_{\mathrm{t}}$和$\boldsymbol{A}_{\mathrm{r}}$表示为

$$\boldsymbol{A}_{\mathrm{t}}=\left[\boldsymbol{a}_{\mathrm{t}}(\varphi_1),\boldsymbol{a}_{\mathrm{t}}(\varphi_2),\cdots,\boldsymbol{a}_{\mathrm{t}}(\varphi_K)\right] \tag{4}$$

$$\boldsymbol{A}_{\mathrm{r}}=[\boldsymbol{a}_{\mathrm{r}}(\theta_1),\boldsymbol{a}_{\mathrm{r}}(\theta_2),\cdots,\boldsymbol{a}_{\mathrm{r}}(\theta_K)] \tag{5}$$

式中, $\boldsymbol{a}_{\mathrm{t}}(\varphi_k)$为双基地 MIMO 雷达发射阵列的第 k 个目标的导向矢量, $\boldsymbol{a}_{\mathrm{r}}(\theta_k)$为双基地 MIMO 雷达接收阵列的第 k 个目标的导向矢量, $\boldsymbol{a}_{\mathrm{t}}(\varphi_k)$、$\boldsymbol{a}_{\mathrm{r}}(\theta_k)$也可表示为

$$\boldsymbol{a}_{\mathrm{t}}(\varphi_k)=[1,\mathrm{e}^{-\mathrm{j}\pi\sin\varphi_k},\mathrm{e}^{-\mathrm{j}2\pi\sin\varphi_k},\cdots,\mathrm{e}^{-\mathrm{j}(M-1)\pi\sin\varphi_k}] \tag{6}$$

$$\boldsymbol{a}_{\mathrm{r}}(\theta_k)=[1,\mathrm{e}^{-\mathrm{j}\pi\sin\theta_k},\mathrm{e}^{-\mathrm{j}2\pi\sin\theta_k},\cdots,\mathrm{e}^{-\mathrm{j}(N-1)\pi\sin\theta_k}] \tag{7}$$

$\boldsymbol{x}(t)$是一个 $MN\times K$维的矩阵, 若将 $\boldsymbol{x}(t)$划分为 N 个子矩阵, 每一个子矩阵是$M\times K$维。

若假设 $\boldsymbol{\Phi}_{\mathrm{r}}=\mathrm{diag}\left[\mathrm{e}^{-\mathrm{j}\pi\sin\theta_1},\mathrm{e}^{-\mathrm{j}\pi\sin\theta_2},\cdots,\mathrm{e}^{-\mathrm{j}\pi\sin\theta_K}\right]$, 则式 (3) 可表示为

$$\begin{aligned}\boldsymbol{x}(t)=&[\boldsymbol{x}_1(t),\boldsymbol{x}_2(t),\cdots,\boldsymbol{x}_N(t)]^{\mathrm{T}}=\\&\left[\boldsymbol{A}_{\mathrm{t}},\boldsymbol{A}_{\mathrm{t}}\boldsymbol{\Phi}_{\mathrm{r}},\cdots,\boldsymbol{A}_{\mathrm{t}}\boldsymbol{\Phi}_{\mathrm{r}}^{N-1}\right]^{\mathrm{T}}\boldsymbol{s}(t)+\boldsymbol{n}(t)\end{aligned} \tag{8}$$

2　基于优化平滑矩阵集的前后向分集平滑算法

2.1　分集平滑算法

在对相干目标进行测向的过程中, 传统的子空间类算法因协方差矩阵秩亏而失效。有学者针对双基地 MIMO 雷达探测相干目标这个问题提出发射接收分集平滑算法[22-23]。但该算法所利用的平滑矩阵方式有限, 并且文献 [23] 的平滑算法计算复杂度较高, 所以本文改进了平滑接收子集矩阵模型, 提出一种基于分集平滑优化对相干目标 DOD 和 DOA 进行联合估计的算法。

先将 M 个发射阵元划分为Q个发射子阵, 每个发射子阵由 P_{t}个阵元组成。对式 (8) 中的信号进行发射分集平滑, 可得到Q个发射子矢量, 其中第q_{t}个发射子矢量可表示为

$$\boldsymbol{x}_{q_t}(t)=\left[\boldsymbol{x}_{q_t1},\boldsymbol{x}_{q_t2},\cdots,\boldsymbol{x}_{q_tN}\right]^{\mathrm{T}}\cdot\left[\boldsymbol{A}_{P_t}\boldsymbol{\Phi}_{\mathrm{t}}^{q_t-1},\boldsymbol{A}_{P_t}\boldsymbol{\Phi}_{\mathrm{t}}^{q_t-1}\boldsymbol{\Phi}_{\mathrm{r}},\cdots,\boldsymbol{A}_{P_t}\boldsymbol{\Phi}_{\mathrm{t}}^{q_t-1}\boldsymbol{\Phi}_{\mathrm{t}}^{N-1}\right]^{\mathrm{T}}\cdot\boldsymbol{s}(t)+\boldsymbol{n}_{q_t}(t)=\left(\boldsymbol{A}_{P_t}\circ\boldsymbol{A}_{\mathrm{r}}\right)\boldsymbol{\Phi}_{\mathrm{t}}^{\theta_{q_t}-1}\boldsymbol{s}(t)+\boldsymbol{n}_{q_t}(t)\tag{9}$$

式中：$q_t=1,2,\cdots,Q_t$；$Q_t=M-P_t+1$；$\boldsymbol{A}_{P_t}$为$\boldsymbol{A}_t$的前P_t行构成的子矩阵；$\boldsymbol{n}_{q_t}(t)$为对应的噪声矢量；$\boldsymbol{\Phi}_t=\mathrm{diag}\left[\mathrm{e}^{-\mathrm{j}\pi\sin\varphi_1},\mathrm{e}^{-\mathrm{j}\pi\sin\varphi_2},\cdots,\mathrm{e}^{-\mathrm{j}\pi\sin\varphi_k}\right]$。

同理，可将N个接收阵元划分为Q个接收子阵，每个接收子阵由P_r个阵元组成，对第q_t个发射子矢量展开分集平滑可以得到第q_tq_r个接收子矢量，可表示为

$$\boldsymbol{x}_{q_tq_r}(t)=[\boldsymbol{x}_{q_tq_r},\boldsymbol{x}_{q_t(q_r+1)},\cdots,\boldsymbol{x}_{q_t(q_r+P_r-1)}]^{\mathrm{T}}=[\boldsymbol{A}_{P_t}\boldsymbol{\Phi}_{\mathrm{t}}^{q_t-1}\boldsymbol{\Phi}_{\mathrm{r}}^{q_r-1},\boldsymbol{A}_{P_t}\boldsymbol{\Phi}_{\mathrm{t}}^{q_t-1}\boldsymbol{\Phi}_{\mathrm{r}}^{q_r},\cdots,\boldsymbol{A}_{P_t}\boldsymbol{\Phi}_{\mathrm{t}}^{q_t-1}\boldsymbol{\Phi}_{\mathrm{r}}^{q_r+P_r-2}]^{\mathrm{T}}\cdot\boldsymbol{s}(t)+\boldsymbol{n}_{q_tq_r}(t)=(\boldsymbol{A}_{P_t}\circ\boldsymbol{A}_{P_r})\boldsymbol{\Phi}_{\mathrm{t}}^{q_t-1}\boldsymbol{\Phi}_{\mathrm{r}}^{q_r-1}\boldsymbol{s}(t)+\boldsymbol{n}_{q_tq_r}(t)\tag{10}$$

式中：$q_r=1,2,\cdots,Q_r$，$Q_r=N-P_r+1$；$\boldsymbol{A}_{P_r}$为$\boldsymbol{A}_t$的前P_r行构成的子矩阵。

由式(9)和式(10)可得，Q_r个接收子矢量构成1个发射子矢量，每个接收子矢量就是相对应的发射接收平滑矩阵窗，若以第1个发射子矢量为例，即包含Q_r个接收子矢量的$\boldsymbol{x}_1(t)$，可得到一个$NQ_t\times NQ_t$维的协方差子矩阵为

$$\boldsymbol{R}_1=(\boldsymbol{A}_{P_t}\circ\boldsymbol{A}_{\mathrm{r}})\boldsymbol{\Phi}_{\mathrm{t}}^{q_t-1}\boldsymbol{R}_{\mathrm{s}}\boldsymbol{\Phi}_{\mathrm{t}}^{1-q_t}\left(\boldsymbol{A}_{P_t}\circ\boldsymbol{A}_{\mathrm{r}}\right)^{\mathrm{H}}+\sigma^2\boldsymbol{I}_{P_tN}\tag{11}$$

式中：$\boldsymbol{R}_{\mathrm{s}}=E\left[\boldsymbol{s}(t)\boldsymbol{s}^{\mathrm{H}}(t)\right]$为目标协方差矩阵；$\sigma^2$为噪声的方差；上标H表示共轭转置。

2.2 接收子阵集的平滑优化

若以第1个发射子矢量的协方差子阵$\boldsymbol{R}_1$为例，如图2所示，$\boldsymbol{R}_1$为一个$NQ_t\times NQ_t$维的矩阵。

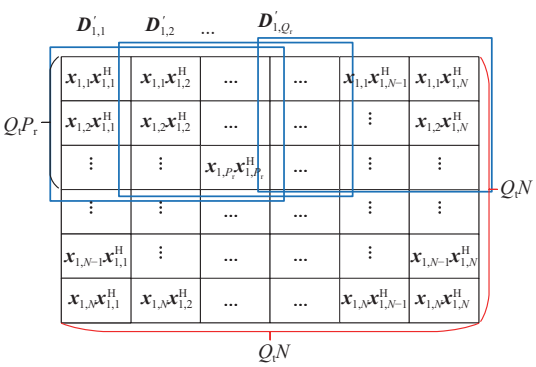


图2 协方差子矩阵平滑示意图

Fig. 2 Covariance submatrix smoothing diagram

因为N个接收阵元划分为Q_r个接收子阵，每个接收子阵由P_r个阵元组成，所以可以将$\boldsymbol{R}_1$的m行至$m+P_r-1$行的子矩阵设为

$$\boldsymbol{D}_{q_t,m}=\left[\left[\boldsymbol{x}_{1,m}\boldsymbol{x}_{1,1}^{\mathrm{H}},\boldsymbol{x}_{1,m}\boldsymbol{x}_{1,2}^{\mathrm{H}},\cdots,\boldsymbol{x}_{1,m}\boldsymbol{x}_{1,N}^{\mathrm{H}}\right],\left[\boldsymbol{x}_{1,m+1}\boldsymbol{x}_{1,1}^{\mathrm{H}},\boldsymbol{x}_{1,m+1}\boldsymbol{x}_{1,2}^{\mathrm{H}},\cdots,\boldsymbol{x}_{1,m+1}\boldsymbol{x}_{1,N}^{\mathrm{H}}\right],\cdots,\left[\boldsymbol{x}_{1,m+P_r}\boldsymbol{x}_{1,1}^{\mathrm{H}},\boldsymbol{x}_{1,m+P_r}\boldsymbol{x}_{1,2}^{\mathrm{H}},\cdots,\boldsymbol{x}_{1,m+P_t}\boldsymbol{x}_{1,N}^{\mathrm{H}}\right]\right]\tag{12}$$

式中：$m\in[1,Q_r]$；$\boldsymbol{R}_1$中有Q_r个$\boldsymbol{D}_{q_t,m}$。

若以$\boldsymbol{R}_1$中第1个子矩阵$\boldsymbol{D}_{q_t,1}$为例，可以对其进行空间平滑处理，如图2所示，将$\boldsymbol{D}_{q_t,1}$分成Q_r个平滑子阵，按平滑顺序可组成新的矩阵为

$$\boldsymbol{R}_{q_t,1}=\left[\boldsymbol{D}'_{1,1},\boldsymbol{D}'_{1,2},\cdots,\boldsymbol{D}'_{1,Q_r}\right]\tag{13}$$

式中：$\boldsymbol{D}'_{m,n}$表示$\boldsymbol{R}_1$中的第m个子矩阵$\boldsymbol{D}_{q_t,m}$中的第n个平滑子阵，其中$m\in[1,Q_r]$，$n\in[1,Q_r]$，$\boldsymbol{D}'_{m,n}$可表示为

$$\boldsymbol{D}'_{m,n}=\left[\left[\boldsymbol{x}_{1,m}\boldsymbol{x}_{1,n}^{\mathrm{H}},\boldsymbol{x}_{1,m}\boldsymbol{x}_{1,n+1}^{\mathrm{H}},\cdots,\boldsymbol{x}_{1,m}\boldsymbol{x}_{1,n+p_t-1}^{\mathrm{H}}\right],\left[\boldsymbol{x}_{1,m+1}\boldsymbol{x}_{1,n}^{\mathrm{H}},\boldsymbol{x}_{1,m+1}\boldsymbol{x}_{1,n+1}^{\mathrm{H}},\cdots\boldsymbol{x}_{1,m+1}\boldsymbol{x}_{1,n+p_t-1}^{\mathrm{H}}\right],\cdots,\left[\boldsymbol{x}_{1,m+p_t-1}\boldsymbol{x}_{1,n}^{\mathrm{H}},\boldsymbol{x}_{1,m+p_t-1}\boldsymbol{x}_{1,n+1}^{\mathrm{H}},\cdots,\boldsymbol{x}_{1,m+p_t-1}\boldsymbol{x}_{1,n+p_r-1}^{\mathrm{H}}\right]\right]\tag{14}$$

然后对Q_r个新的组合矩阵进行加权处理，有效利用更多的协方差矩阵信息，得到第q_t个发射子矢量的平滑矩阵$\boldsymbol{D}_{q_t}$，可表示为

$$\boldsymbol{D}_{q_t}=\frac{1}{Q_r}\sum_{m=1}^{Q_r}\boldsymbol{R}_{q_t,m}\tag{15}$$

将Q_t个发射子矢量进行平滑，根据前后向空间平滑算法变换可得前向平滑矩阵$\boldsymbol{D}_f$为

$$\boldsymbol{D}_{\mathrm{f}}=\frac{1}{Q_t}\sum_{q_t=1}^{Q_t}\boldsymbol{D}_{q_t}\tag{16}$$

若当$p_tp_r>K$，$Q_tQ_r>K$，$Q_t>1$，$Q_r>1$时，$\boldsymbol{D}_f$的秩为可检测的最大相干目标数。

2.3 前后向分集平滑算法

虽然优化平滑矩阵集后的前向平滑也可以检测相干信源目标，并取得一定的效果，但和前后向平滑算法相比，后者可以提升更大的测角精度和降低信源相关性。

若前向空间平滑表示按照正向平滑的子阵列，那么后向空间平滑即表示关于正向平滑子阵列对称分布的虚拟阵列，可以有效扩展阵列孔径，尽量避免经典空间平滑算法中的阵列孔径减小的缺点。因为前向平滑子阵列和后向平滑子阵列在接收信号的协方差矩阵中表现是相位相反、位置共轭，所以后向平滑可以等效于前向平滑的共轭加翻转。

根据文献[22]可知，式(19)中的$\boldsymbol{A}_{P_t}\circ\boldsymbol{A}_{P_r}$的性质如下：

$$\boldsymbol{A}_{P_t}\circ\boldsymbol{A}_{P_r}=\frac{\boldsymbol{J}_{P_tP_r}(\boldsymbol{A}_{P_t}\circ\boldsymbol{A}_{P_r})^*}{\boldsymbol{\Phi}_{\mathrm{r}}^{P_r-1}\boldsymbol{\Phi}_{\mathrm{t}}^{P_t-1}}=\frac{(\boldsymbol{J}_{P_r}\boldsymbol{B}_{P_r}^*)\circ(\boldsymbol{J}_{P_t}\boldsymbol{B}_{P_t}^*)}{\boldsymbol{\Phi}_{\mathrm{r}}^{P_r-1}\boldsymbol{\Phi}_{\mathrm{t}}^{P_t-1}}\tag{17}$$

式中，$\boldsymbol{J}_{P_r}$为阶数为P_r的单位逆矩阵；$\boldsymbol{J}_{P_t}$为阶数为P_t的单位逆矩阵，$\boldsymbol{J}_{P_tP_r}$表示$\boldsymbol{J}_{P_t}\otimes\boldsymbol{J}_{P_r}$；上标“*”表示共轭。

$$\boldsymbol{J}_{Q_tQ_rP_t}=\mathrm{blkdiag}\left[J_{P_t},J_{P_t},\cdots,J_{P_t}\right]_{Q_tQ_r}\tag{18}$$

所以平滑矩阵集优化后的后向平滑矩阵 $\boldsymbol{D}_{\mathrm{b}}$ 为

$$\boldsymbol{D}_{\mathrm{b}}=\frac{1}{Q_{\mathrm{t}}}\sum_{q_{\mathrm{t}}=1}^{Q_{\mathrm{t}}}\boldsymbol{J}_{P_{\mathrm{t}}P_{\mathrm{t}}}\boldsymbol{D}_{q_{\mathrm{t}}}^{*}\boldsymbol{J}_{Q_{\mathrm{t}}Q_{\mathrm{t}}P_{\mathrm{t}}} \tag{19}$$

根据式 (16) 和式 (19) 可建立平滑矩阵集优化后的前后向平滑矩阵为

$$\boldsymbol{D}_{\mathrm{fb}}=\frac{1}{2}(\boldsymbol{D}_{\mathrm{f}}+\boldsymbol{D}_{\mathrm{b}}) \tag{20}$$

对 $\boldsymbol{D}_{\mathrm{fb}}$ 进行奇异值(singular value decomposition, SVD)分解得到信号子空间 $\boldsymbol{E}_{\mathrm{n}}$, 而后进行空间谱联合估计时, 传统的方法会使用二维多重信号分类(two-dimension multiple signal classification, 2D-MUSIC)算法或 ESPRIT 算法, 使用 2D-MUSIC 算法可以得到更精确的估计结果, 但是会使得计算复杂度更高; 使用 ESPRIT 算法可以减小搜索所需的数据量, 从而大幅度减小计算复杂度, 但是估计精度会有所下降。

本文选择了一种适用于空间谱联合估计的降维多重信号分类(reduced dimension multiple signal classification, RD-MUSIC)算法[24], 该算法不仅可以避免 2D-MUSIC 算法的高复杂度, 而且在估计性能方面远优于 ESPRIT 算法, 在计算复杂度方面仅略高于 ESPRIT 算法。RD-MUSIC 算法在进行 DOD 和 DOA 联合估计时, 可以生成特征值矩阵去连接 DOD 维度和 DOA 维度, 达到自动排序的效果, 从而实现目标 DOD 和 DOA 的自动配对。

RD-MUSIC 算法的谱函数为

$$f_{\mathrm{music}}=\frac{1}{\left[\hat{\boldsymbol{a}}_{\mathrm{t}}(\varphi)\otimes\hat{\boldsymbol{a}}_{\mathrm{r}}(\theta)\right]^{\mathrm{H}}\boldsymbol{E}_{\mathrm{n}}\boldsymbol{E}_{\mathrm{n}}^{\mathrm{H}}\left[\hat{\boldsymbol{a}}_{\mathrm{t}}(\varphi)\otimes\hat{\boldsymbol{a}}_{\mathrm{r}}(\theta)\right]} \tag{21}$$

由文献 [24] 可知, $\hat{\boldsymbol{a}}_{\mathrm{t}}(\theta)\otimes\hat{\boldsymbol{a}}_{\mathrm{t}}(\theta)$ 表示为

$$\hat{\boldsymbol{a}}_{\mathrm{t}}(\varphi)\otimes\hat{\boldsymbol{a}}_{\mathrm{r}}(\theta)=\left[\hat{\boldsymbol{a}}_{\mathrm{t}}(\varphi)\otimes\boldsymbol{I}_{N}\right]\hat{\boldsymbol{a}}_{\mathrm{r}}(\theta) \tag{22}$$

式中, $\boldsymbol{I}_{N}$ 表示 $N\times N$ 的单位矩阵, 因此, 式 (21) 可以转化为

$$f_{\mathrm{music}}=\frac{1}{\hat{\boldsymbol{a}}_{\mathrm{r}}(\theta)^{\mathrm{H}}\left[\hat{\boldsymbol{a}}_{\mathrm{t}}(\varphi)\otimes\boldsymbol{I}_{N}\right]^{\mathrm{H}}\boldsymbol{E}_{\mathrm{n}}\boldsymbol{E}_{\mathrm{n}}^{\mathrm{H}}\left[\hat{\boldsymbol{a}}_{\mathrm{t}}(\varphi)\otimes\boldsymbol{I}_{N}\right]\hat{\boldsymbol{a}}_{\mathrm{r}}(\theta)} \tag{23}$$

再将式 (23) 中的一个分式定义为

$$\boldsymbol{V}(\varphi)=\left[\hat{\boldsymbol{a}}_{\mathrm{t}}(\varphi)\otimes\boldsymbol{I}_{N}\right]^{\mathrm{H}}\boldsymbol{E}_{\mathrm{n}}\boldsymbol{E}_{\mathrm{n}}^{\mathrm{H}}\left[\hat{\boldsymbol{a}}_{\mathrm{t}}(\varphi)\otimes\boldsymbol{I}_{N}\right] \tag{24}$$

所以式 (21) 可表示为

$$f_{\mathrm{music}}=\frac{1}{\hat{\boldsymbol{a}}_{\mathrm{r}}(\theta)^{\mathrm{H}}\boldsymbol{V}(\varphi)\hat{\boldsymbol{a}}_{\mathrm{r}}(\theta)} \tag{25}$$

则第 k 个相干目标的 DOA 可表示为

$$\hat{\theta}_{k}=\arg\min\frac{1}{\boldsymbol{e}_{1}^{\mathrm{T}}\boldsymbol{V}(\varphi_{k})^{-1}\boldsymbol{e}_{1}}=\arg\max\left(\boldsymbol{e}_{1}^{\mathrm{T}}\boldsymbol{V}(\varphi_{k})^{-1}\boldsymbol{e}_{1}\right) \tag{26}$$

式中: $\boldsymbol{e}_{1}=[1,0,\cdots,0]^{\mathrm{T}}$。

同理通过 $\hat{\boldsymbol{a}}_{\mathrm{t}}(\varphi)\otimes\hat{\boldsymbol{a}}_{\mathrm{r}}(\theta)=[\boldsymbol{I}_{M}\otimes\hat{\boldsymbol{a}}_{\mathrm{r}}(\theta)]\hat{\boldsymbol{a}}_{\mathrm{t}}(\varphi)$ 替换到式 (24) 中也可得到目标的 DOD, 其中, $\boldsymbol{I}_{M}$ 表示 $M\times M$ 的单位矩阵。

2.4　算法步骤

平滑矩阵集优化算法具体如下:

输入　$\boldsymbol{x}(t)=\boldsymbol{A}\boldsymbol{s}(t)+\boldsymbol{n}(t)$　　t=1,2,$\cdots$,L;

输出　$\{\hat{\varphi}_{k},\hat{\theta}_{k}\}$　　$k=1,2,\cdots,K$;

步骤 1　将发射阵列和接收阵列进行分集平滑, 并构建第 q_{t} 个发射子矢量 $\boldsymbol{x}_{q_{\mathrm{t}}}$, 计算 $\boldsymbol{x}_{q_{\mathrm{t}}}$ 的协方差矩阵 $\boldsymbol{R}_{q_{\mathrm{t}}}$。

步骤 2　将 $\boldsymbol{R}_{q_{\mathrm{t}}}$ 进行平滑处理, 利用 Q_{r} 个平滑矩阵加权处理得到第 q_{t} 个发射子矢量的分集平滑矩阵 $\boldsymbol{D}_{q_{\mathrm{t}}}$。

步骤 3　通过构建 Q_{t} 个发射子矢量的分集平滑矩阵可以得到前向平滑矩阵 $\boldsymbol{D}_{\mathrm{f}}$。

步骤 4　将 $\boldsymbol{D}_{\mathrm{f}}$ 进行变换得前后向平滑矩阵 $\boldsymbol{D}_{\mathrm{fb}}$, 对其进行 SVD 分解得到信号子空间 $\boldsymbol{E}_{\mathrm{n}}$。

步骤 5　使用 RD-MUSIC 算法进行空间谱估计, 使得目标的 DOD 和 DOA 实现自动匹配, 并得到 $\{\hat{\theta}_{k},\hat{\varphi}_{k}\}$。

3　仿真实验和结果分析

3.1　场景设定

假设双基地 MIMO 雷达的发射阵元数 M=9, 接收阵元数 N=9, 发射子阵数为 $Q_{\mathrm{t}}=5$,每个发射子阵由 5 个阵元组成, 接收子阵数为 $Q_{\mathrm{r}}=5$,每个接收子阵由 5 个阵元组成, 测试的信源均为全相干信号。针对相干目标进行 200 次 DOD 和 DOA 联合估计的蒙特卡雷仿真实验的均方根误差(root mean squared error, RMSE)计算表达式为

$$\sigma_{\mathrm{RMSE}}=\sqrt{\frac{1}{2\times200K}\sum_{i=1}^{200}\sum_{k=1}^{K}\left[\left(\hat{\varphi}_{k}^{i}-\varphi_{k}\right)^{2}+\left(\hat{\theta}_{k}^{i}-\theta_{k}\right)^{2}\right]} \tag{27}$$

式中: $\hat{\varphi}_{k}^{i}$ 和 $\hat{\theta}_{k}^{i}$ 分别为第 i 次实验中的 φ_{k} 和 θ_{k} 的角度估计值。

本节通过对比本文算法、文献 [22] 提出的发射接收分集平滑(transmit receive diversity smoothing, TRDS)算法和文献 [21] 提出的非对称空间差分平滑(asymmetric spatial difference smoothing, ASDS)算法的角度分辨率性能、空间谱性能和均方误差大小。

3.2　目标 DOD 和 DOA 估计的角度分辨率比较

本节实验通过对比 3 种算法对 2 个全相干目标角度相近情况下的角度分辨性能来验证 3 种算法的角度分辨性能。

假设信噪比 (signal to noise ratio, SNR) 的值 r_{SNR}

为10 dB, 快拍数为300, 估计目标为2个, 其角度信息为: $(\varphi_1,\theta_1)=(27°,27°)$, $(\varphi_2,\theta_2)=(29°,29°)$。图3为不同平滑算法角度分辨率比较。由图可知, 本文算法因为使用更有效的平滑矩阵集, 所以能很好地对2个相近目标实现超分辨估计, TRDS算法[22]可以分辨出有2个目标存在, 但是分辨性能较差, ASDS算法[21]在相同条件下, 无法分辨有2个目标的存在, 分辨性能不如前2种算法。实验表明: 本文算法具有较好的分辨性能。

3.3 目标DOD和DOA估计的空间谱比较

本实验通过比较3种算法的空间谱, 验证3种算法的空间谱估计性能。

假设 r_{SNR} 为10 dB, 快拍数为300, 估计的5个相干目标角度信息为: $(\varphi_1,\theta_1)=(5°,5°)$, $(\varphi_2,\theta_2)=(15°,15°)$, $(\varphi_3,\theta_3)=(25°,25°)$, $(\varphi_4,\theta_4)=(35°,35°)$, $(\varphi_5,\theta_5)=(45°,45°)$。图4为不同平滑算法的归一化空间谱对比。从图可知, 本文算法因其使用更有效的平滑矩阵集, 获得的更多的数据量, 能够较为准确地估计出5个全相干目标的方位, TRDS算法[22]可以获得5个谱峰相对较弱的空间谱, 而ASDS算法[21]仅能估计出3个目标的方位。实验表明: 本文算法在空间谱估计效果方面优于其他2种算法, 并且此时

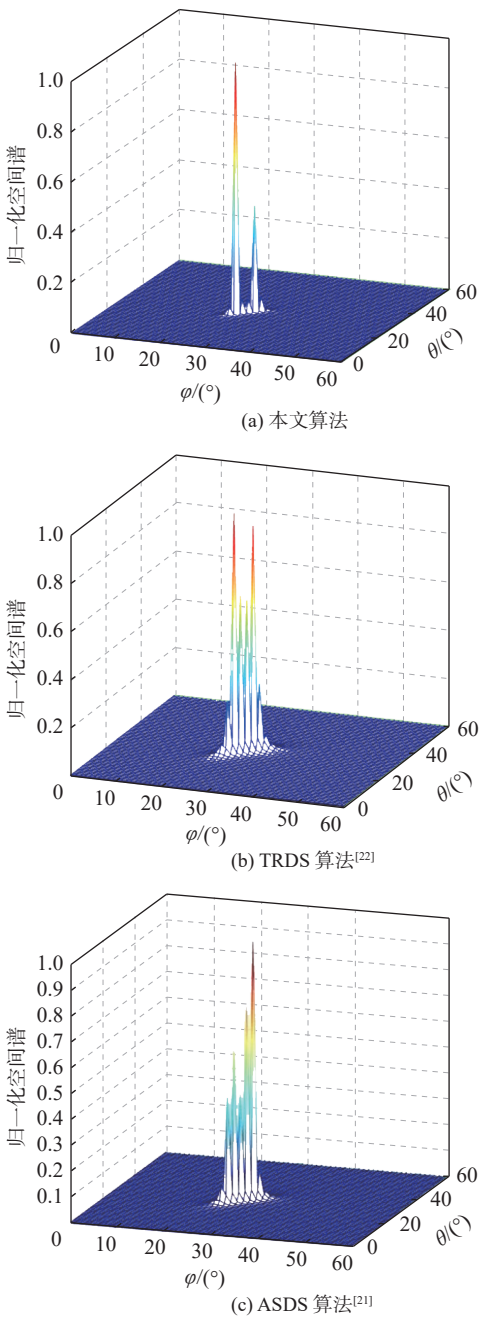


图3 不同平滑算法角度分辨率比较

Fig. 3 Comparison of angular resolution of different smoothing algorithms

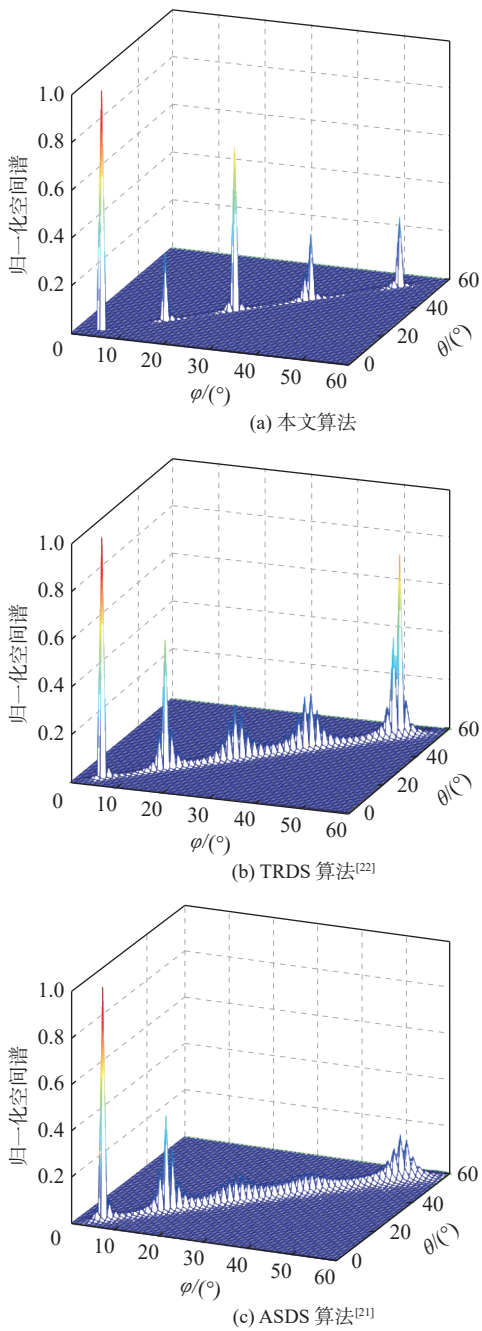


图4 不同平滑算法的归一化空间谱对比

Fig. 4 Comparison of angular resolution of different smoothing algorithms

$\boldsymbol{D}_t$的秩为 5，表示可检测的最大相干目标数为 5，在其他 2 种算法中，皆存在一定的自由度损失情况。

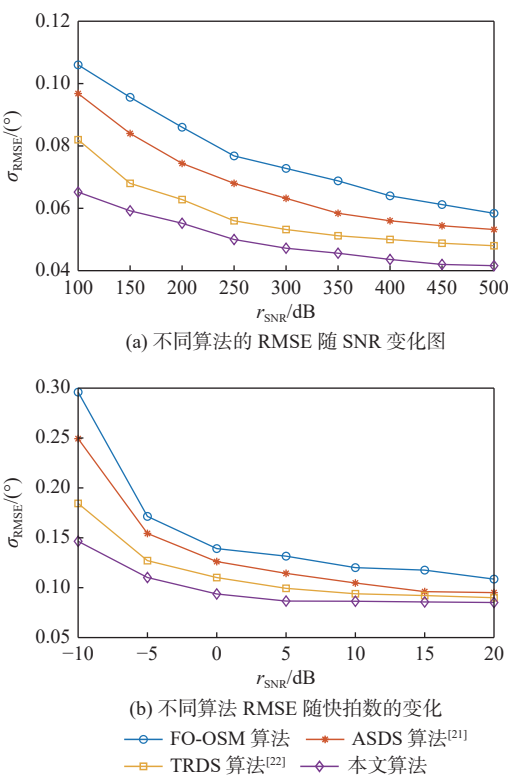
3.4　目标 DOD 和 DOA 估计的均方根误差

本节实验为比较本文算法与 TRDS 算法[22]、ASDS 算法[21]和前向平滑的优化平滑矩阵集（forward optimized smoothing matrix，FO-OSM）算法的均方误差。

假设空间中有 3 个相干目标，其角度信息为 $(\varphi_1,\theta_1)=(5°,5°),(\varphi_2,\theta_2)=(25°,25°),(\varphi_5,\theta_5)=(45°,45°)$，对目标角度的搜索步长为 0.1°。图 5 为不同算法 RMSE 随 SNR 和快拍数的变化。图 5(a) 表示 4 种算法的测角 RMSE 随 SNR 变化图，在实验中，快拍数设置为 200，SNR 设置为从−10 dB 到 20 dB，间隔为 5 dB。图 5(b) 表示 4 种算法的测角 RMSE 随快拍数变化图，SNR 设置为 10 dB，快拍数设置为从 100 次到 500 次，间隔 50 次。由图可知，4 种算法均能够有效地实现对相干目标方位的估计，本文算法因其构造的平滑集利用了更多的协方差矩阵信息，所以在测角精度方面优于其他 3 种算法，其中 FO-OSM 算法是本文算法的前向平滑算法，其估计效果较弱于其他 3 种算法。

3.5　目标 DOD 和 DOA 估计的复杂度

根据 3.2～3.4 节可知，本文算法在角度分辨率性能和估计精度方面优于其他几种算法。在时间复杂度方面，进行 200 次蒙特卡罗仿真实验时，不同算法的时间复杂度对比如图 6 所示，其中本文算法表示配合 RD-MUSIC 算法，本文算法 (2D) 表示本文算法配合 2D-MUSIC 算法。本次进行仿真实验的计算机配置为：Intel(R) Core(TM) i7-7 700HQ CPU @ 2.80 GHz 2.80 GHz 的中央处理器、16.0 GB 的运行内存、NVIDIA GeForce GTX 1050 Ti 4 GB 的显卡。



(a) 不同算法的 RMSE 随 SNR 变化图

(b) 不同算法 RMSE 随快拍数的变化

图 5　不同算法的 RMSE 随 SNR 和快拍数变化

Fig. 5　Variation of RMSE with SNR and snapshot number of different algorithms

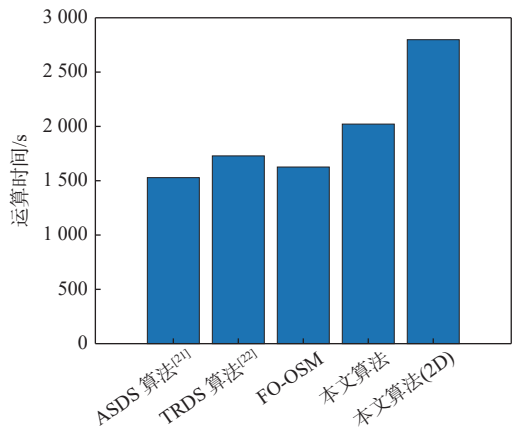


图 6　不同算法的时间复杂度对比图

Fig. 6　Time complexity comparison of different algorithms

由此结合算法的计算复杂度和时间复杂度可得出结论：本文算法可通过少量增加运算量实现角度估计性能的提升，其体现在优化平滑矩阵集上，并且相对于使用 2D-MUSIC 算法进行空间谱估计，使用 RD-MUSIC 算法能够有效减小时间复杂度。

4　结　论

1）本文算法组成了效率更高的平滑矩阵集，使得目标测向精度和目标测向分辨率提高。

2）本文算法将复杂的二维参数估计的协方差矩阵进行降维处理，使得平滑子矩阵数和平滑次数减小，和传统的双基地 MIMO 雷达联合估计相比，有效减小了计算复杂度。

3）本文算法在自由度方面仍存在不足，对相干目标测向时，估计目标数受到了矩阵秩亏的影响，所以不能对大量目标进行测向。

参考文献（References）

[1] RABIDEAU D J, PARKER P. Ubiquitous MIMO multifunction digital array radar[C]//The 37th Asilomar Conference on Signals, Systems and Computers. Piscataway: IEEE Press, 2004: 1057-1064.

[2] LUO K, MANIKAS A. Superresolution multitarget parameter estimation in MIMO radar[J]. IEEE Transactions on Geoscience and Remote Sensing, 2013, 51(6): 3683-3693.

[3] GONG P C, SHAO Z H, TU G P, et al. Transmit beampattern design based on convex optimization for MIMO radar systems[J]. Signal Processing, 2014, 94(6): 195-201.

[4] LI L. Cramer-Rao bound for parameter estimation in narrowband bistatic MIMO radar[J]. Applied Mechanics and Materials, 2014, 556-562: 5034-5037.

[5] SCHMIDT R. Multiple emitter location and signal parameter estimation[J]. IEEE Transactions on Antennas and Propagation, 1986, 34(3): 276-280.

[6] RAO B D, HARI K V S. Performance analysis of root-music[J]. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1989, 37(12): 1939-1949.

[7] KUNDA D. Modified MUSIC algorithm for estimating DOA of signals[J]. Signal Processing, 1996, 48(1): 85-90.

[8] 何子述, 黄振兴, 向敬成. 修正 MUSIC 算法对相干信号源的 DOA 估计性能[J]. 通信学报, 2000, 21(10): 14-17.
HE Z S, HUANG Z X, XIANG J C. The performance of DOA estimation for correlated signals by modified MUSIC algorithm[J]. Journal of Communications, 2000, 21(10): 14-17(in Chinese).

[9] GURBUZ A C, MCCLELLAN J H, CEVHER V. A compressive beamming method[C]//2008 IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway: IEEE Press, 2008: 2617-2620.

[10] LI J, CONAN J, PIERRE S. Joint estimation of channel parameters for MIMO communication systems[C]//2005 2nd International Symposium on Wireless Communication Systems. Piscataway: IEEE Press, 2005.

[11] YAN H, LI J, LIAO G. Multitarget identification and localization using bistatic mimo radar systems[J]. EURASIP Journal on Advances in Signal Processing, 2008(1): 1-8.

[12] DUOFANG C, BAIXIAO C, GUODONG Q. Angle estimation using ESPRIT in MIMO radar[J]. Electronics Letters, 2008, 44(12): 770.

[13] WANG H, KAVEH M. On the performance of signal-subspace processing: Part II: Coherent wide-band systems[J]. IEEE Transactions on Acoustics, Speech and Signal Processing, 1987, 35(11): 1583-1591.

[14] CHEN Y M, LEE J H, YEH C C, et al. Bearing estimation without calibration for randomly perturbed arrays[J]. IEEE Transactions on Singnal Processing, 1991, 39(1): 194-197.

[15] SHAN T J, KAILATH T. Adaptive beamforming for coherent signals and interference[J]. IEEE Transactions on Acoustics, Speech, and Signal Processing , 1985, 33(3): 527-536.

[16] MAHNAZ A P, SEDIGHEH G. M$\ell_{1,2}$-MUSIC algorithm for DOA estimation of coherent sources[J]. IET Signal Processing, 2017, 11(4): 429-436.

[17] 游致远, 胡国平, 周豪, 等. 基于阵列内插的双基地互质 MIMO 雷达角度联合估计方法[J]. 现代雷达, 2022, 44(5): 45-50.
YOU Z Y, HU G P, ZHOU H, et al. Joint estimation method of angles for bistatic coprime MIMO radar target based on array interpolation[J]. Modern Radar, 2022, 44(5): 45-50(in Chinese).

[18] 游致远, 胡国平, 周豪. 基于冗余阵元优化的双基地嵌套 MIMO 雷达目标 DOD 和 DOA 联合估计方法[J/OL]. 系统工程与电子技术. (2022-1-8)[2022-03-29]. http://kns.cnki.net/kcms/detail/11.2422.tn.20220328.1045.012.html.
YOU Z Y, HU G P, ZHOU H. Bistatic nested MIMO radar based on redundant element optimization joint estimation method of target DOD and DOA[J/OL]. Journal of Systems Engineering and Electronics. (2022-1-8)[2022-03-29]. http://kns.cnki.net/kcms/detail/11.2422.tn.20220328.1045.012.html(in Chinese).

[19] 张小飞, 陈晨. MIMO 雷达中一种降维 MUSIC 的相干角度估计算法[J]. 沈阳航空航天大学学报, 2012, 29(3): 64-70.
ZHANG X F, CHEN C. A reduced dimension-MUSIC bused coherent angle estimation for MIMO radar[J]. Journal of Shenyang University of Aeronautics and Astronautics, 2012, 29(3): 64-70(in Chinese).

[20] 梁浩, 崔琛, 伍波, 等. 双基地 MIMO 雷达相干目标角度估计算法[J]. 系统工程与电子技术, 2014, 36(6): 1068-1074.
LIANG H, CUI C, WU B, et al. Angle estimation of coherent target for bistatic MIMO radar[J]. Systems Engineering and Electronics, 2014, 36(6): 1068-1074(in Chinese).

[21] HONG S, WAN X R, KE H Y, et al. Spatial difference smoothing for coherent sources location in MIMO radar[J]. Signal Processing, 2015, 109: 69-83.

[22] ZHANG W, LIU W, WANG J, et al. Joint transmission and reception diversity smoothing for direction finding of coherent targets in MIMO radar[J]. IEEE Journal of Selected Topics in Signal Processing, 2014, 8(1): 115-124.

[23] SHI J P, HU G P, ZHANG X F, et al. Smoothing matrix set-based MIMO radar coherent source localisation[J]. International Journal of Electronics, 2018, 105(8): 1345-1357.

[24] BAIDOO E, HU J R, ZENG B, et al. Joint DOD and DOA estimation using tensor reconstruction based sparse representation approach for bistatic MIMO radar with unknown noise effect[J]. Signal Processing, 2021, 182: 107912.

Joint DOA and DOD estimation of bistatic MIMO radar coherent targets based on smoothing matrix sets optimization

YOU Zhiyuan[1, 2]，HU Guoping[1, *]，ZHOU Hao[1]，ZHENG Guimei[1]

(1. Air And Missile Defense College，Air Force Engineering University，Xi'an 710051，China；
2. Graduate College，Air Force Engineering University，Xi'an 710051，China)

Abstract：Due to the rank-deficient phenomenon of the covariance matrix in coherent targets estimation, a algorithm based on diversity smoothing optimization for joint estimation of coherent target wave direction of departure (DOD) angle and direction of arrival (DOA) is proposed to solve the problems of poor accuracy and low angular resolution of coherent target detection by bistatic multiple input multiple output (MIMO) radar. The proposed algorithm can construct a more effective smooth matrix sets for joint angle estimation of coherent targets through diversity smoothing and reducing the covariance matrix. Perform diversity smoothing on the transmit array and receive array, construct transmit sub-vectors, and calculate their covariance matrix. Smooth and reconstruct the covariance matrix of each transmit sub-vector and then perform weighting to obtain a forward smoothing matrix. The forward smoothing matrix is transformed by conjugate flipping and other transformations, and the forward and backward smoothing matrix is obtained, on which the singular value decomposition (SVD) is then performed. The spatial spectrum of the signal is estimated by using the reduced dimension multiple signal classification (RD-MUSIC) algorithm, and the automatic pairing of DOD and DOA is realized. The proposed algorithm effectively improves the direction performance of bistatic MIMO radar for coherent targets, and the feasibility of the proposed algorithm is proved by experiments.

Keywords：bistatic radar; MIMO radar; coherent target; diversity smoothing; joint angle estimation

Received：2022-03-21；**Accepted**：2022-06-20；**Published Online**：2022-06-24 16：53
URL：link.cnki.net/urlid/11.2625.V.20220623.1720.005
Foundation item：National Natural Science Foundation of China (62071476)
* **Corresponding author**. E-mail：hgp6068@163.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0177

基于飞机状态的备件动态规划技术

康子旭[1,2,3]，周栋[1,2,3]，李会欣[4,*]，郭子玥[1,2,3]，陈承璋[1,2,3]，宋子骋[1,2,3]

(1. 北京航空航天大学 可靠性与系统工程学院，北京 100191；2. 虚拟现实技术与系统国家重点实验室，北京 100191；
3. 可靠性与环境工程技术国防科技重点实验室，北京 100191；4. 河北航空有限公司，石家庄 050802)

摘 要：伴随民航业的竞争日趋激烈，航空公司迫切需要控制成本以提升效益。民机备件既是直接影响飞行安全的关键因素，备件成本又在航空公司可控成本中占比最高。因此，在保证飞行安全的前提下做好备件规划，对航空公司有至关重要的意义。根据民机备件的安全级别、故障修复时限等对备件分级设置保障率水平，采用符合备件故障特性的马尔可夫生灭过程，建立满足给定保障率水平的备件量计算模型，并在此基础上建立以保障成本最低为目标函数，满足保障率水平为约束函数的备件量动态规划模型。基于飞机状态分析历史故障数据、季节性、日利用率、机队规模对备件保障的影响，在所建模型中将故障率季节性差异、日利用率和飞机停场损失淡旺季差异进行评估以减小计算结果与实际需求的偏差。以H航空公司ERJ-190机型显示组件为例，对所建模型进行应用验证，计算结果与H航空公司运营实际相吻合，证明所建模型可为航空公司备件规划提供技术方法支持。

关 键 词：备件；飞机状态；马尔可夫生灭过程；经济性；动态规划
中图分类号：V37; V352
文献标志码：A **文章编号**：1001-5965(2024)01-0276-10

随着中国民航业进入蓬勃发展时期，各航空公司之间的竞争日趋激烈，盈利愈加困难，迫切需要探讨如何在保证安全的前提下，控制运营成本，提升经济效益。航空公司的可控成本主要是维修成本，而备件成本在维修成本中占比最高，约60%[1]，在运营投入和流动资金占用也最大。因此，做好备件规划工作，对航空公司降低成本提升效益起至关重要的作用。

民机备件品类繁多，生产制造标准高，通用性差，备件管理是一项复杂的工作。目前在民航业，航空公司备件管理工作中普遍存在备件思维简单粗放，员工对合适的库存量缺乏正确的认识，以及已有备件量计算模型与实际运行需求偏差较大等问题。大部分航空公司在备件管理方面出现的主要问题亟待解决。

许多专家学者针对备件库存量进行了大量研究，并且取得一定的研究成果。吕川等[2]应用泊松分布理论分别为循环件、非循环件和消耗件建立了实用的备件需求量计算模型。赵洪军等[3]利用航材需求服从泊松分布的特征，创建相应的优化模型以减少投入的成本，发挥出更为良好的经济性。瞿红春和姜柏松[4]根据民机备件保障率的定义构造出一种新的可靠性模型，提出应优先考虑故障率高的部件，并采用Barlow优化算法得出比飞机制造商推荐值低的备件量。李正为等[5]提出在对航材需求展开预测的基础上，利用泊松分布与正态分布来对备件量予以优化求解，从而获取更为科学的预测。何亚群等[6]提出了一种以飞机可用度为评价指标来确定可修件需求量的分析模型，并给出模型的优化求解方法。马麟等[7]针对部件服从任意寿

收稿日期：2022-03-22；录用日期：2022-05-06；网络出版时间：2022-05-27 11:10
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220525.1933.004.html
*通信作者. E-mail：zy2114112@buaa.edu.cn

引用格式：康子旭，周栋，李会欣，等. 基于飞机状态的备件动态规划技术[J]. 北京航空航天大学学报，2024，50(1)：276-285.
KANG Z X，ZHOU D，LI H X，et al. Dynamic planning technology of civil spare parts based on aircraft state[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(1)：276-285 (in Chinese).

命分布的多部件系统，提出了基于蒙特卡罗仿真-边际效益分析的备件配置优化算法。

备件量计算模型的研究可分为统计模型、分析模型[8]。统计模型以大量历史数据为基础，对备件需求分析，得出备件需求的统计规律，同时，结合部件的相关信息、管理者的经验等，对统计规律采用回归分析[9]、指数平滑[10]、概率统计理论[11-12]、贝叶斯法[13-14]、人工神经网络[15]等方法，建立备件量计算模型。分析模型通常利用工程理论与可靠性理论方法，将系统视为冷储备系统，在此基础上，对备件的工作、维修及订购过程进行分析，最终确定备件需求量，基于数学规划的备件量模型以备件的保障费用、可用度、可靠度、等待时间为目标约束条件[16-18]。

备件库存量的相关研究在多个方面进行了探索，但仍有不足：①着眼于基于寿命分布理论或采用某种分析方法构建备件量计算模型，没有充分考虑航空公司飞机运营状态对备件需求的影响；②基于经济性的计算中，备件成本仅涵盖了采购成本，未考虑飞机停场损失成本，或者考虑了损失成本而没有考虑淡旺季的差异。

针对以上问题，本文结合实际航空备件管理工作，采用符合备件维修特点的马尔可夫生灭过程建立备件量计算模型，并充分考虑飞机状态对备件需求的影响，在飞机备件满足给定的保障率的前提下，建立民机备件量动态规划模型，并基于经济性对备件进行了动态规划。

1　飞机状态对备件保障的影响

民机备件保障应考虑多因素影响，其中飞机状态为备件保障的一个重要影响因素。对于备件保障，有关飞机状态的影响主要包括以下影响因素：历史故障数据、季节、日利用率及机队规模等。

1.1　历史故障数据

历史故障数据包括平均故障间隔时间、故障原因、修理周期、年利用率、修理费用等，飞机部件的历史故障数据对于备件需求预测具有极其重要的意义，可以排除因地域、航线、技术、管理等带来的差异。平均故障间隔时间可决定可修系统的故障率，修理周期和年利用率决定可修系统的修复率，修理费用可为经济性提供参考依据。

1.2　季节

受季节影响，备件在不同月份发生故障次数有所差异。如在四季分明的地区，冬季寒冷，夏季温热，气温年较差较为悬殊。如对某航空公司 B-737NG 和 ERJ-190 机型备件进行分析，对备件的故障时间按月份进行统计，发现备件在不同的月份发生故障次数差异较大，如图 1 所示。

由图可知备件的故障呈现出季节性，以上备件的春夏季和秋冬季故障次数如图 2 所示。

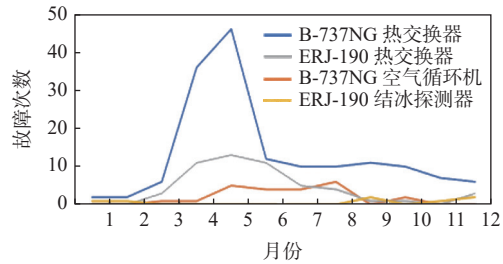


图 1　备件故障月份分布

Fig. 1　Month distribution of spare parts failure

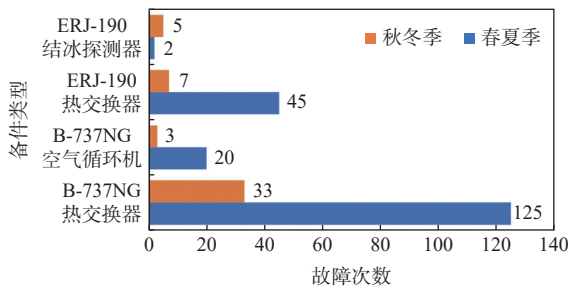


图 2　备件故障季节分布

Fig. 2　Seasonal distribution of spare parts failures

其中 4 月～9 月为春夏季，10 月～来年 3 月为秋冬季。

通过对备件故障时间的统计与比较，B-737NG 和 ERJ-190 机型分别在春夏季和秋冬季故障高发备件见表 1。可见春夏季柳絮、沙尘、雷雨较多，且天气高温湿热，空调系统、防雨系统、电子部件故障率较高；秋冬季冰、雪、霜较多，且天气寒冷，液压、燃油、防冰等系统部件故障率较高。

表 1　季节性故障高发备件列表

Table 1　Spare parts list with high incidence of seasonal failures

B-737NG 春夏季	ERJ-190 春夏季	B-737NG 秋冬季	ERJ-190 秋冬季
空气循环机	空气循环机	前缘缝翼作动筒	结冰探测器
压力调节关断活门	热交换器	发动机热防冰活门	等待提示的反馈
热交换器	总温探头	燃油泵	低限活门
温度控制活门	高压级活门	安定面配平马达	扭矩马达
雨刷马达	风扇活门	冲压门作动器	燃油管理组件
风挡组件	压力调节 关断活门	燃油增压泵	结冰探测器
总温探头	显示组件	起动活门	辅助动力装置 发电机
气象雷达收发机	自动刹车电门	液压机械组件	发动机防冰活门
惯性导航仪	数据管理系统	风挡加温 控制面板	发动机驱动泵
空中防撞系统 计算机	前电子舱风扇	发动机驱动泵	应急停留刹车活门

综上所述，本文的备件量计算模型中，针对以上季节性高发故障部件，平均故障间隔将采用季节性数据，而非全年平均数据。

1.3 日利用率

航空公司多以旅客运输为主业，由于旺季的飞机日利用率高于淡季，并且旺季客座率高，飞机因缺件导致的停场损失也较高，所以在备件保障中需要考虑淡旺季带来的影响。每年的春节前15天～春节后25天，大批外出务工人员返乡，加之春节长假旅游，此阶段为春运出行高峰期。每年7～8月，大、中、小学生放假，带动旅游热潮，此阶段为2个月的暑运高峰期。另外，五一劳动节、国庆节长假期间，国内旅客出行需求旺盛。除这些特殊的时期外，其他时间一般是出行淡季。

为使淡旺季的影响量化地体现到备件量计算模型中，据某航空公司对淡旺季的日利用率进行的统计得知，旺季飞机日利用率约为年平均日利用率的1.10倍，淡季的飞机日利用率约为年平均日利用率的0.90倍，故引入平均日利用率修正因子为

$$\theta=\begin{cases}1.10 & \beta=1\\ 0.90 & \beta=0\end{cases} \tag{1}$$

式中：$\beta=1$表示旺季，$\beta=0$表示淡季。

由于旺季日利用率高，停场导致飞行小时的损失增加，产生的飞机停场损失也相应增加，并结合上述航空公司对淡旺季日利用率统计结果得知：在航班客座率高的旺季，备件短缺造成的飞机停场损失要高于淡季，旺季的飞机停场损失约为淡季飞机停场损失的1.22倍。故引入飞机停场损失修正因子为

$$\psi=\begin{cases}1.22 & \beta=1\\ 1.00 & \beta=0\end{cases} \tag{2}$$

1.4 机队规模

国内诸多航空公司，按照机队规模分类，可以分为大型、中大型、中型、中小型、小型航空公司。机队规模对备件的影响是显而易见的，机队规模越大对备件的需求越多，平均备件保障成本越来越低。本文所述的机队规模属于中小型。在本文的备件量计算模型中有机队规模这一参数，机队规模发生变化时，将新参数代入模型计算即可。

2 基于马尔可夫生灭过程的备件量计算模型

民用飞机使用过程中，维修是保证使用可靠性的重要因素。在航空公司实际运营中，某一周转件故障后，通常需要立即拆下并安装其备件，拆下件进行维修，修复后作为备件。此类周转件所处的系统能够视作为一些部件或一个、多个修理设备组建的可修系统。

研究可修系统的主要数学工具是随机过程理论，部件的故障伴随着系统状态转移，复杂电子部件的寿命服从指数分布，通常其修复时间也服从指数分布，可采用马尔可夫过程进行研究。另外，马尔可夫生灭过程是可修系统备件预测的重要理论，其具有解析性好、计算方便等特性[19]。故本文应用该理论对可修系统进行描述。

2.1 问题描述及假设

某一复杂电子部件为某一机型的部件，机队含s架该机型，每架飞机上装有该部件m台，正常运营时有n台部件在工作，机队年利用率η_{AUR}，单位飞行小时h_{FH}，该部件历次故障的自上次修理使用时间t_{TSR}和修理周期T_{RTAT}可获得，用λ、μ分别表示部件的故障率和修复率，设该部件购置备件量为k，求满足保障率α的最小k值。

假设：①复杂电子部件在出现故障后，立即用库存备件将其更换并将部件修复如新，且更换时间很短，可以忽略不计；②备件在库存储期间不会失效也不会劣化；③维修资源充足，修理周期不受在修件数量的影响；④保障率只受备件量的影响，不考虑人为差错等的影响；⑤在某一时刻，最多只有1个部件出现故障，不会出现2个部件同时故障的情况。

2.2 模型中参数的确定

2.2.1 故障率

故障率λ由平均故障间隔时间t_{MTBF}确定，而t_{MTBF}取决于航空公司飞机状态。在工程实际中，对t_{MTBF}的近似估算主要源于数据和概率论知识，通过点估计可得到最有可能取得的t_{MTBF}值。t_{MTBF}的点估计计算式[20]为

$$t_{\mathrm{MTBF}}=\frac{1}{r}\left[\sum_{i=1}^{r}T_i+(l-r)T\right]^{[20]}$$

式中：当已知某部件故障日期、自修后使用时间等信息，可将部件故障日期数l看作样本量；T为设定的总时间；r为在时间T内发生故障的次数；在时间T内发生的第i次故障对应的自上次修理后的使用时间即为发生该故障前的运行时间T_i。经查询可知该部件设计鉴定t_{MTBF}为T_0，即部件能平均工作T_0。对于现场失效数据的统计分析，可取故障平均发生在$T/2$时刻[21]，故可取设定的总时间为$T=2T_0$。

因此，结合设计鉴定t_{MTBF}指标和t_{MTBF}点估计计算式可得到较为可信的t_{MTBF}值。后可求得该部件基于当前飞机状态的故障率为

$$\lambda=\frac{1}{t_{\mathrm{MTBF}}}$$

2.2.2　修复率

航空公司监控t_{TSR}的单位为飞行小时 (h_{FH}), 故障率的单位为h_{FH}^{-1}, 修复率μ也同样采用单位h_{FH}^{-1}, 因修理周期T_{RTAT}的单位为日历日, 通过年利用率做单位换算, 即

$$\mu=\frac{1}{t_{MTTR}}=\frac{365}{T_{RTAT}\eta_{AUR}} \tag{3}$$

式中: t_{MTTR}为平均修复时间。

考虑旺季与淡季飞机日利用率的差别, 加入本文所述的修正因子θ, 则备件修复率为

$$\mu=\frac{1}{t_{MTTR}}=\frac{365}{T_{RTAT}\eta_{AUR}\theta} \tag{4}$$

2.2.3　保障率

目前部分航空公司将各备件保障率α统一设定为 0.9, 没有考虑部件安全级别的区分。按安全级别, 备件可分为不可放行 (NO GO) 件、条件放行 (GO IF) 件及可放行 (GO) 件。NO GO 件指影响航空器适航性的基本设备。GO IF 件即航空公司最低设备清单 (minimum equipment list, MEL) 中所列的备件, 可以办理故障保留, 按照修复时限要求又可分为 4 类, A 类: 没有规定标准的期限, 其必须要根据最低设备清单规定的限制内, 落实好相应的修复工作; B 类: 其必须在连续 3 个日历日, 即 72 h 之内落实好相应的修复工作; C 类: 其必须在连续的 10 个日历日, 即 240 h 之内落实好相应的修复工作, 但并没有将发现当日包含进去; D 类: 其必须在连续的 120 个日历日, 即 2 880 h 之内落实好相应的修复工作。当备件为 NO GO 件, 其备件保障需求最高, 对应备件保障率水平最高, 当备件为 A 类 GO IF 件, 其备件保障需求较高, 对应备件保障率水平较高, B 类 GO IF 件次之, D 类 GO IF 件保障率水平为 GO IF 件中最低; 在 GO 件中, 为满足乘客需求, 与客舱服务相关的备件应予以重视, 而与客舱服务无关的 GO 件的保障率水平最低。

由于保障率的设定目的是进行安全与效益的权衡, 对于不同航空公司, 应根据自己的经营理念与实际情况的不同, 结合航空公司运营经验设定数值, 并在实际运营过程中不断完善修改备件保障率设定值, 使之符合理想运营情况。现针对对不同类型备件的保障率水平进行的多层级设定, 给出基于安全级别分类与修复时限的多层级备件保障率要求的参考设定值, 如表 2 所示。航空公司应结合实际运营经验与经营理念, 设定定制化的备件保障率水平并在运营中不断完善修改, 以符合实际理想经营状态。GO 件主要指的是影响程度较为轻微或乘坐者们的舒适性有所降低, 然而并不会对安全性带来负面影响的备件。

表 2　备件保障率设置

Table 2　Spare parts guarantee rate setting

备件类型	α
NO GO件	0.97
GO IF件, 修复时限A	0.95
GO IF件, 修复时限B	0.93
GO IF件, 修复时限C	0.92
GO IF件, 修复时限D	0.90
GO件, 与客舱服务相关	0.88
GO件, 与客舱服务无关	0.85

2.3　模型的建立

通过部件的故障数量i来表示系统的状态, 考虑到系统的前$n+k+1$个状态, 也就是第 0, 1, …, $k+n$个状态。系统各状态对应的工作部件与备件数如下。

状态 0: 工作部件数为n, 库存备件数为k。

状态 1: 工作部件数为n, 库存备件数为$k-1$。

⋮

状态k: 工作部件数为n, 库存备件数为 0。

状态$k+1$: 工作部件数为$n-1$, 库存备件数为 0。

⋮

状态$n+k-1$: 工作部件数为 1, 库存备件数为 0。

状态$n+k$: 工作部件数为 0, 库存备件数为 0。

详细的结合生灭过程的定义不难发现, 这一系统能够归为马尔可夫生灭过程。

对于状态$i(0\leqslant i\leqslant k)$, 单位时间系统的总故障率为$n\lambda$, 其中, λ为单个部件故障率; 对于状态$i(k+1\leqslant i\leqslant k+n)$, 单位时间系统的总故障率为$(n+k-i)\lambda$; 对于状态$i$, 系统的总修复率为$i\mu$。系统的转移状态图如图 3 所示。

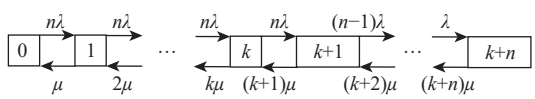


图 3　马尔可夫生灭过程系统状态转移图

Fig. 3　State transfer diagram of Markovian birth-death process system

对于状态 0, 处于Δt时间内, 从该状态起, 总故障概率近似为$n\lambda\Delta t$, 状态并未转移的概率为$[1-n\lambda\Delta t]$; 处于$t+\Delta t$时刻的时候, 该系统并未改变, 仍是该状态的总故障概率; 从该状态转移, 直至状态 1 的概率为$P_0 n\lambda\Delta t$。处于Δt时间以内, 从第 2 种状态转移到第 1 种状态的概率近似为$\mu\Delta t$; 处于$t+\Delta t$时刻的时候, 系统从第 2 种状态转移到第 1 种状态的概率为$P_1\mu\Delta t$, 其中, P_0为起始时刻稳定状态概率, P_1为第 2 种状态稳定时的概率。可得到

$$P_0(t+\Delta t)=P_0(t)(1-n\lambda\Delta t)+P_1(t)\mu\Delta t \tag{5}$$

$$\frac{[P_0(t+\Delta t)-P_0(t)]}{\Delta t}=-P_0(t)n\lambda+P_1(t)\mu \tag{6}$$

令$\Delta t \to 0$，则

$$\frac{\mathrm{d}P_0(t)}{\mathrm{d}t} = -P_0(t)n\lambda + P_1(t)\mu \tag{7}$$

同理可得$\frac{\mathrm{d}P_1(t)}{\mathrm{d}t}, \frac{\mathrm{d}P_2(t)}{\mathrm{d}t}, \cdots, \frac{\mathrm{d}P_{n+k}(t)}{\mathrm{d}t}$。

由微分方程，可得$P_0(t), P_1(t), \cdots, P_{n+k}(t)$，并且$\sum_{i=0}^{n+k} P_i = 1$。

若某部件装机数量及使用量比较大，则微分方程的维数增加，求解会变得困难。系统状态有限，必然存在平稳分布，而且本文所关心的也是部件稳定工作的情况。这意味着处于极限情况的时候，系统在完全稳态以后，P_i为常数，其微分为0。

对于状态0，有

$$P_1 = \frac{n\lambda}{\mu}P_0$$

对于状态1，有

$$P_2 = \frac{n^2}{2!}\left(\frac{\lambda}{\mu}\right)^2 P_0$$

⋮

对于状态k有

$$P_{k+1} = \frac{n^{k+1}}{(k+1)!}\left(\frac{\lambda}{\mu}\right)^{k+1} P_0$$

对于状态$k+1$有

$$P_{k+2} = \frac{n^k n(n-1)}{(k+2)!}\left(\frac{\lambda}{\mu}\right)^{k+2} P_0 = \frac{n^k n!}{(k+2)!(n-2)!}\left(\frac{\lambda}{\mu}\right)^{k+2} P_0$$

⋮

对于状态$k+n$有

$$P_{k+n} = \frac{n^k n!}{(k+n)!}\left(\frac{\lambda}{\mu}\right)^{k+n} P_0$$

综上，系统各状态的概率为

$$\begin{cases} P_i = \dfrac{n^i}{i!}\left(\dfrac{\lambda}{\mu}\right)^i P_0 & 0 \leqslant i \leqslant k \\ P_i = \dfrac{n^k n!}{i!(n+k-i)!}\left(\dfrac{\lambda}{\mu}\right)^i P_0 & k+1 \leqslant i \leqslant k+n \end{cases} \tag{8}$$

根据全概率公式推导可得

$$P_0 = \left[\sum_{i=0}^{k}\frac{n^i}{i!}\left(\frac{\lambda}{\mu}\right)^i + \sum_{i=k+1}^{k+n}\frac{n^k n!}{i!(n+k-i)!}\left(\frac{\lambda}{\mu}\right)^i\right]^{-1} \tag{9}$$

所以系统的可靠度为

$$\sum_{j=0}^{k} p_j = \sum_{j=0}^{k}\frac{n^j}{j!}\left(\frac{\lambda}{\mu}\right)^j P_0 = \sum_{j=0}^{k}\frac{n^j}{j!}\left(\frac{\lambda}{\mu}\right)^j \cdot \left[\sum_{i=0}^{k}\frac{n^i}{i!}\left(\frac{\lambda}{\mu}\right)^i + \sum_{i=k+1}^{k+n}\frac{n^k n!}{i!(n+k-i)!}\left(\frac{\lambda}{\mu}\right)^i\right]^{-1} \tag{10}$$

结合式(10)，不难求出给定的可靠度下对应的数量最少的备件，与此同时，亦能够得到系统的可靠度。

对于给定的保障率α，按照式(11)可求得的k值即为满足保障率α的最小备件量，求解计算流程图如图4所示。具体来说，计算编程时，在通过使用

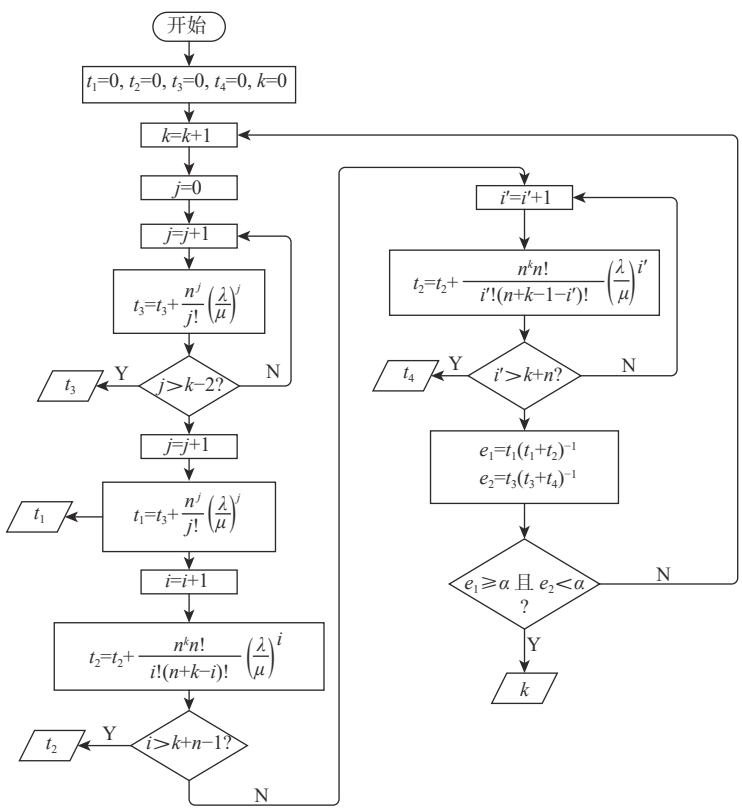


图4 基于马尔可夫生灭过程备件量计算流程

Fig. 4 Flow for calculating the amount of spare parts based on Markovian birth and death process

MATLAB 软件编码建立函数后, 输入式(11)的变量k及参数n、λ、μ, 由软件计算得到不同备件量所对应的保障率, 后根据给定的保障率条件得到满足要求的最小备件量k, 且对于给定的保障率, 按照式(11)有且仅有唯一的最小备件量k满足要求, 故该值即为所求结果。

$$\begin{cases} \sum_{j=0}^{k}\frac{n^j}{j!}\left(\frac{\lambda}{\mu}\right)^j\cdot \left[\sum_{i=0}^{k}\frac{n^i}{i!}\left(\frac{\lambda}{\mu}\right)^i+\sum_{i=k+1}^{k+n}\frac{n^k n!}{i!(n+k-i)!}\left(\frac{\lambda}{\mu}\right)^i\right]^{-1}\geqslant\alpha \\ \sum_{j=0}^{k-1}\frac{n^j}{j!}\left(\frac{\lambda}{\mu}\right)^j\cdot \left[\sum_{i=0}^{k-1}\frac{n^i}{i!}\left(\frac{\lambda}{\mu}\right)^i+\sum_{i=k}^{k+n-1}\frac{n^k n!}{i!(n+k-1-i)!}\left(\frac{\lambda}{\mu}\right)^i\right]^{-1}<\alpha \end{cases} \tag{11}$$

3　基于经济性备件量动态规划模型

3.1　假　设

为构建基于经济性备件量的动态模型, 本文作出如下假设: ①备件仓储成本很低可以忽略不计; ②备件缺件时可以及时租赁到器材; ③只有 NO GO 件的短缺导致飞机停场损失, 任务其他备件短缺不会导致飞机停场; ④租赁备件运输成本很低可以忽略不计。

3.2　参数的确定

C_{T}为总保障成本, 本文模型的目标是总保障成本最低。C_{P}为采购成本, 即备件单价P_{u}与备件量k的乘积。C_{S}为缺件成本, 即发生缺件后租赁备件的费用, 等于缺件数、概率、修复周期、日租金的乘积, 即

$$C_{\mathrm{S}}=\sum_{j=k+1}^{k+n}(j-k)P_j\frac{R_{\mathrm{D}}}{j\mu} \tag{12}$$

其中, R_{D}为备件日租金。根据已知航空公司与互援单位签署的互援协议, 租赁可用部件 (非全新) 基准价格为目录价格的 72%, 承租方需付的日租金为基准价格的 0.45%, 即

$$R_{\mathrm{D}}=P_{\mathrm{u}}\times 72\%\times 0.45\% \tag{13}$$

C_{O}为停场损失, 当 NO GO 件出现缺件情况时, 会导致飞机停场继而产生停场损失, 对于其他类备件该项费用以 0 计。旺季航班客座率高, 备件短缺造成的飞机停场损失也要高于淡季, 在计算时乘以修正因子ψ, 并根据所保障的时间段确定飞机停场损失修正因子ψ的取值。

3.3　模型的建立

目标是总保障成本C_{T}最低, 约束条件为满足给定保障率α, 模型如下:

$$\begin{cases} \min\ C_{\mathrm{T}}=C_{\mathrm{P}}+C_{\mathrm{s}}+C_{\mathrm{O}} \\ \text{s.t.}\ \sum_{i=0}^{k}P_i\geqslant\alpha \end{cases} \tag{14}$$

式 (14) 展开即

$$\begin{cases} \min\quad C_{\mathrm{T}}=kP_{\mathrm{u}}+\sum_{j=k+1}^{k+n}(j-k)\frac{n^k n!}{j!(n+k-j)!}\left(\frac{\lambda}{\mu}\right)^j\cdot \left[\sum_{i=0}^{k}\frac{n^i}{i!}\left(\frac{\lambda}{\mu}\right)^i+\sum_{i=k+1}^{k+n}\frac{n^k n!}{i!(n+k-i)!}\left(\frac{\lambda}{\mu}\right)^i\right]^{-1}\cdot \left(\frac{R_D}{j\mu}+C_{\mathrm{O}}\right) \\ \text{s.t.}\quad \sum_{j=0}^{k}\frac{n^j}{j!}\left(\frac{\lambda}{\mu}\right)^j\cdot \left[\sum_{i=0}^{k}\frac{n^i}{i!}\left(\frac{\lambda}{\mu}\right)^i+\sum_{i=k+1}^{k+n}\frac{n^k n!}{i!(n+k-i)!}\left(\frac{\lambda}{\mu}\right)^i\right]^{-1}\geqslant\alpha \end{cases}$$

总成本的计算流程如图 5 所示。可求得不同备件量所对应的保障成本, 继而得到在满足保障率要求下使保障成本最小k值, 即为最佳备件量。

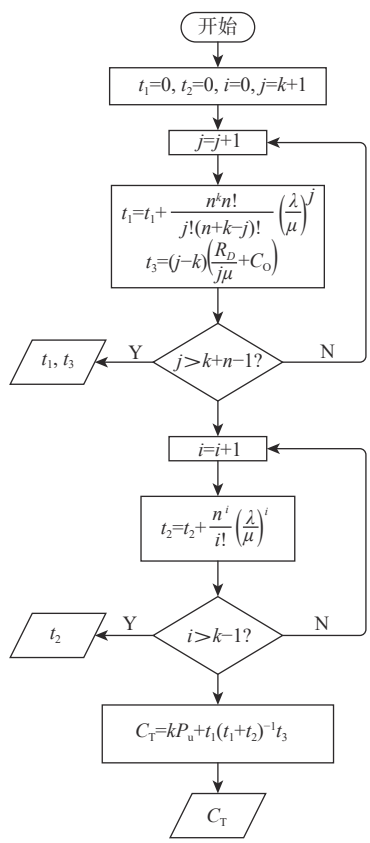


图 5　备件总保障成本计算流程

Fig. 5　Flow for calculating the total guaranteed cost of spare parts

4　算例分析

以 H 航空公司的 ERJ-190 机型显示组件为例,

应用本文模型进行备件量的计算与优化。

4.1 机型显示组件备件量计算

ERJ-190机型显示组件寿命服从指数分布，每架飞机装有显示组件5台，其设计鉴定t_{MTBF}为10 000 h_{FH}，该机型机队规模为6架，机队年利用率为3 000 h_{FH}，历次故障数据见表3，显示组件为GO IF件，修复时限为A类，器材单价为146 000 $，需确定暑运期间该显示组件的备件量。

表3 显示组件历史故障数据

Table 3 Display of component history failure data

故障日期	t_{TSR}/h_{FH}	T_{RTAT}/d	修理后入库日期
2018/9/11	273.58	47	2018/10/28
2018/7/2	14 126.06	43	2018/8/14
2018/5/26	7 462.16	53	2018/7/18
2018/5/2	10 515.37	41	2018/6/12
2018/4/2	3 392.15	39	2018/5/11
2018/3/22	3 424.39	54	2018/5/15
2017/9/18	16 227.26	41	2017/10/29
2017/5/19	13 358.58	88	2017/8/15
2016/12/20	3 668.06	41	2017/1/30
2016/12/5	14 256.33	40	2017/1/14
2016/9/17	10 188.50	39	2016/10/26
2016/9/9	351.26	50	2016/10/29
2016/8/27	46.20	23	2016/9/19
2016/8/21	3 104.11	33	2016/9/23
2016/7/25	7 537.74	44	2016/9/7

按故障发生日期，将以上故障数据分春夏季和秋冬季(4～9月为春夏季，10月～来年3月为秋冬季)，汇总可得，显示组件在春夏季的故障次数为13次，t_{MTBF}为6 923.64 h_{FH}；显示组件在秋冬季的故障次数为2次，t_{MTBF}为8 962.20 h_{FH}。

由本节历史故障数据可知，显示组件的故障集中发生在春夏季，春夏季故障间隔时间较短，故障率相对较高，考虑满足暑运期间的需求，故使用春夏季的故障率进行备件量计算，即故障率

$$\lambda = \frac{1}{t_{MTBF}} = 1.44\times10^{-4}$$

平均修理周期$T_{RTAT}=45$ d，同时考虑暑运期间属于旺季，将日利用率修正因子加入计算，则修复率为

$$\mu = \frac{1}{t_{MTTR}} = \frac{365}{T_{RTAT}\eta_{AUR}\theta} = 2.46\times10^{-3}$$

由于显示组件属于修复时限为A类的GO IF件，参照表2保障率α设置为0.95。工作的部件数量为

$$n = sm = 30$$

以上数据代入基于马尔可夫生灭过程备件量计算模型，使用MATLAB软件计算结果可得到表4。

表4 显示组件不同备件量对应保障率

Table 4 Guarantee rates for different spare parts quantities for display components

k	α
0	0.181 5
1	0.485 5
2	0.747 3
3	0.900 0
4	0.967 2
5	0.990 8
6	0.997 8
7	0.999 5

由计算结果可知，为显示组件购置4台备件在旺季可满足保障率$\alpha\geqslant0.95$的要求。

4.2 ERJ-190机型显示组件基于经济性动态规划

对于H航空公司而言，ERJ-190飞机显示组件为GO IF件，可以办理故障保留，不考虑停场损失，即$C_O=0$；单价P_u=146 000 $。

日租金

$$R_D = p_u\times72\%\times0.45\%$$

代入本文模型中，并使用MATLAB软件计算结果可得表5。

表5 显示组件不同备件量对应总保障成本

Table 5 Total cost of guarantee for different spare parts for display components

k	C_T/$
0	138 330
1	194 070
2	307 000
3	442 100
4	585 000
5	730 200
6	876 000
7	1 022 000

由4.1节计算可知，满足约束条件的k值为$k\geqslant4$，结合本节结果可知，在备件量为4时既满足保障率要求，又使得保障成本最小。

实际上，H航空公司ERJ-190显示组件的当前备件数量正是4台。备件利用率可以反映出实际备件数量的合理性。由修理后入库日期与故障日期可推知每次故障时更换备件后的备件利用率在0.25～1.00之间，且每次故障更换备件后的平均备件利用率为0.48，故备件量为4可在保证工作的显

示组件不缺件的情况下, 实际使用的备件数较多。由此可知, ERJ-190 显示组件实际备件量为 4 时保障状态较为理想。本文的计算模型与航空公司运营实际相吻合。

4.3　对比验证

1) 用传统可修件指数寿命备件量计算模型[22]计算的最低备件量为 8, 较本文计算模型结果多出一倍。由此表明, 在保证飞行安全情况下, 本文计算模型可显著降低备件储备量, 节约备件成本。

2) 在 4.1 节中的备件量计算中若不考虑保障的时间段, 则修复率:

$$\mu=\frac{1}{t_{\mathrm{MTTR}}}=\frac{365}{T_{\mathrm{RTAT}}\eta_{\mathrm{AUR}}}=2.7\times10^{-3}$$

替换备件量计算模型中的μ值, 得到计算结果如表 6 所示。

表 6　未考虑旺季对日利用率影响的不同备件量保障率

Table 6　Different spare parts volume guarantee rates without considering impact of peak season on daily utilization rates

k	α
0	0.210 4
1	0.533 1
2	0.787 2
3	0.922 4
4	0.976 6
5	0.994 0
6	0.998 7
7	0.999 7

不考虑旺季日利用率的影响后, 相同备件量对应的保障率水平发生了变化, 假如要求的保障率水平为 0.92, 则不考虑季节影响计算得备件量为 3, 而由表 4 可知, 备件量 3 在旺季只能达到 0.9 的保障率, 满足 0.92 保障率需要的最小备件量为 4 台。即当不考虑旺季日利用率的影响后, 相同备件量对应的保障率水平发生变化, 计算出的备件量在旺季有可能未到达所给定的保障率水平。

3) 若未考虑保障率水平的细化分级, 将备件保障率统一设定为 0.9, 则由表 4 可知, 保障率为 0.9 的显示组件最低备件量为 3, 这与实际较为理想的保障状态不符。由此表明对备件分级设置保障率水平可有效减小理论值与实际需求的偏差, 使备件规划方案更贴合实际。

4) 若在备件量计算模型中不对故障率季节性差异加以考虑, 则由平均故障间隔采用全年平均数据可求得故障率为

$$\lambda=\frac{1}{t_{\mathrm{MTBF}}}=1.39\times10^{-4}$$

替换备件量计算模型中的λ值得到计算结果如表 7 所示。

表 7　未考虑季节性差异的不同备件量对应保障率

Table 7　Guarantee rates corresponding to different spare parts volumes without taking into account seasonal differences

k	α
0	0.192 2
1	0.503 7
2	0.763 0
3	0.909 1
4	0.971 1
5	0.992 2
6	0.998 2
7	0.999 6

不考虑故障率季节性差异这一因素将会导致同一备件量对应保障率水平变化, 对比表 4 与表 7 可知, 当不考虑故障率季节性差异后, 每一备件量对应保障率均有所提升, 因此, 未考虑该因素计算出的相对较小的备件量在暑运期间有可能未能达到保障率要求。

5　结　论

1) 通过分析基于飞机状态的历史故障数据、季节性、日利用率、机队规模对备件保障的影响, 考虑故障率季节性差异与日利用率淡旺季差异的备件量计算模型提高了备件需求预测精度。

2) 基于民机备件安全级别的分类, 以及 GO IF 件故障修复时限的不同要求, 对备件保障率的设置进行了多层级设定, 使备件保障规划更加符合实际情况。

3) 对基于符合备件维修特点的马尔可夫生灭过程的备件量计算模型及基于成本分析的备件量动态规划模型进行应用与验证。通过比较计算结果与航空公司实际运营需求, 证明了本文计算模型可在确保保障率的前提下有效控制经济成本, 提升航空公司效益。

此外, 在实际工程中影响备件规划的因素很多, 还需对本文计算模型作出改进。由于以故障率和修复率为参数构建转移概率矩阵的马尔可夫生灭过程理论在考虑运行环境与人的因素等时存在局限性, 因此, 在未来研究中, 需要对不确定背景下的基于马尔可夫生灭过程的备件规划模型进行研究。同时, 本文计算模型未考虑修理周期受在修件数对预测准确性的影响, 在未来工作中需要进一步完善改进。

参考文献（References）

[1] 杜燕. A航空公司航材管理与成本优化研究[D]. 上海: 上海交通大学, 2015: 4.
DU Y. Research on air material management and cost optimization of A airlines[D]. Shanghai: Shanghai Jiao Tong University, 2015: 4 (in Chinese).

[2] 吕川, 赵宇, 张坚. 飞机备件需求量的确定方法[J]. 北京航空航天大学学报, 1995, 21(4): 130-133.
LYU C, ZHAO Y, ZHANG J, et al. Method for determining the sparesdemand of airliner[J]. Journal of Beijing University of Aeronautics and Astronautics, 1995, 21(4): 130-133(in Chinese).

[3] 赵洪军, 齐校锋, 刘海翔. 基于泊松分布的航材库存决策优化研究[J]. 长沙航空职业技术学院学报, 2009, 9(1): 40-43.
ZHAO H J, QI X F, LIU H X. Research on optimization of air material inventory based on Poisson distribution[J]. Journal of Changsha Aeronautical Vocational and Technical College, 2009, 9(1): 40-43(in Chinese).

[4] 瞿红春, 姜柏松. 基于航材系统保障率的备件优化模型的研究[J]. 航空制造技术, 2004, 47(9): 79-82.
QU H C, JIANG B S. Study on the aeronautical material system reliability based optimization model of spare parts[J]. Aeronautical Manufacturing Technology, 2004, 47(9): 79-82(in Chinese).

[5] 李正为, 黄俊, 倪现存, 等. 面向航空公司的飞机首批航材计划研究[J]. 航空维修与工程, 2006(3): 52-54.
LI Z W, HUANG J, NI X C, et al. The study of first spare part plan of the plane which faces the airline company[J]. Aviation Maintenance & Engineering, 2006(3): 52-54(in Chinese).

[6] 何亚群, 谭学峰, 金福禄. 基于可用度的飞机可修件需求分析[J]. 系统工程与电子技术, 2004, 26(3): 848-849.
HE Y Q, TAN X F, JIN F L. Demand analysis of aircraft repairable spares based on availability[J]. Systems Engineering and Electronics, 2004, 26(3): 848-849(in Chinese).

[7] 龙军, 康锐, 马麟, 等. 任意寿命分布的多部件系统备件配置优化算法[J]. 北京航空航天大学学报, 2007, 33(6): 698-700.
LONG J, KANG R, MA L, et al. Algorithm for spares optimization of system with multi-items with different life distributions[J]. Journal of Beijing University of Aeronautics and Astronautics, 2007, 33(6): 698-700(in Chinese).

[8] 张澳夫. 基于 PHM 的备件配置问题研究[D]. 北京: 北京理工大学, 2014: 10-12.
ZHANG A F. Research on the problem of spare parts based on Prognostics & Health management[D]. Beijing: Beijing Institute of Technology, 2014: 10-12(in Chinese).

[9] 马秀红, 宋建社, 董晟飞. 基于回归分析的备件故障率预测模型[J]. 计算机仿真, 2003, 20(11): 6-8.
MA X H, SONG J S, DONG S F. Model of spare parts failure rate based on linear regression[J]. Computer Simulation, 2003, 20(11): 6-8(in Chinese).

[10] 毕义明, 王汉功. 武器装备损耗备件预测模型研究[J]. 装备指挥技术学院学报, 2004, 15(1): 1-4.
BI Y M, WANG H G. Research on the consumption standard model of weapon's equipment spare[J]. Journal of Institute of Command and Technology, 2004, 15(1): 1-4(in Chinese).

[11] 徐晓燕. 一种基于需求特性分类的备件库存管理方法及其实证研究[J]. 系统工程理论与实践, 2006, 26(2): 62-67.
XU X Y. A demand classification based approach of inventory management for spare-parts and its application[J]. Systems Engineering Theory & Practice, 2006, 26(2): 62-67(in Chinese).

[12] 赵宇, 伏洪勇, 张坚, 等. 航空电子设备备件需求量分析[J]. 系统工程与电子技术, 2002, 24(3): 1-3.
ZHAO Y, FU H Y, ZHANG J, et al. Demand analysis of the spare parts of avionic devices[J]. Systems Engineering and Electronics, 2002, 24(3): 1-3(in Chinese).

[13] ARONIS K P, MAGOU I, DEKKER R, et al. Inventory control of spare parts using a Bayesian approach: A case study[J]. European Journal of Operational Research, 2004, 154(3): 730-739.

[14] KAMATH K R, PAKKALA T P M. A Bayesian approach to a dynamic inventory model under an unknown demand distribution[J]. Computers & Operations Research, 2002, 29(4): 403-422.

[15] 王进才, 范英飙. 基于人工神经网络建立武器系统备件预测模型[J]. 质量与可靠性, 2010(5): 13-17.
WANG J C, FAN Y B. Establishment of spare parts prediction model of weapon system based on artificial neural network[J]. Quality and Reliability, 2010(5): 13-17(in Chinese).

[16] 董博超, 宋保维, 梁庆卫, 等. 基于马尔可夫模型的备件配置优化方法[J]. 系统工程, 2011, 29(9): 124-126.
DONG B C, SONG B W, LIANG Q W, et al. The method of spares configuration optimization based on Markov model[J]. Systems Engineering, 2011, 29(9): 124-126(in Chinese).

[17] 毛德军, 李庆民, 张志华. 以装备可用度为中心的保障方案优化方法[J]. 兵工学报, 2011, 32(5): 636-640.
MAO D J, LI Q M, ZHANG Z H. An optimizing method for availability-centric equipment support plan[J]. Acta Armamentarii, 2011, 32(5): 636-640(in Chinese).

[18] 王乃超, 康锐. 多约束条件下备件库存优化模型及分解算法[J]. 兵工学报, 2009, 30(2): 247-251.
WANG N C, KANG R. An optimization model for inventory spares under multi-constraints and its decomposition algorithm[J]. Acta Armamentarii, 2009, 30(2): 247-251(in Chinese).

[19] 陈顶, 方志耕, 刘思峰. 可修排队系统备件灰色生灭预测模型[J]. 系统工程理论与实践, 2020, 40(5): 1326-1338.
CHEN D, FANG Z G, LIU S F. Grey BD prediction model for spare parts of repairable queuing system[J]. Systems Engineering-Theory & Practice, 2020, 40(5): 1326-1338(in Chinese).

[20] 王昊. 国产数控车床可靠性评估方法研究[D]. 沈阳: 东北大学, 2018: 134-135.
WANG H. Research on reliability assessment methods of domestic CNC lathes[D]. Shenyang: Northeastern University, 2018: 134-135(in Chinese).

[21] 凌刚. 继电保护装置平均故障间隔时间指标的探讨[J]. 低压电器, 2012(20): 34-39.
LING G. Study on the mean time between failure of relay protection equipment[J]. Low Voltage Apparatus, 2012(20): 34-39(in Chinese).

[22] 叶心太, 余仁波. 装备保障中备件需求量确定方法研究[J]. 舰船电子工程, 2010, 30(5): 155-158.
YE X T, YU R B. Research on method of definited spare part requirement during equipment support[J]. Ship Electronic Engineering, 2010, 30(5): 155-158(in Chinese).

Dynamic planning technology of civil spare parts based on aircraft state

KANG Zixu[1, 2, 3], ZHOU Dong[1, 2, 3], LI Huixin[4, *], GUO Ziyue[1, 2, 3], CHEN Chengzhang[1, 2, 3], SONG Zicheng[1, 2, 3]

(1. School of Reliability and Systems Engineering, Beihang University, Beijing 100191, China;
2. State Key Laboratory of Virtual Reality Technology and Systems, Beijing 100191, China;
3. State Key Defense Science and Technology Laboratory on Reliability and Environmental Engineering, Beijing 100191, China;
4. Hebei Airlines Ltd., Shijiazhuang 050802, China)

Abstract: Civil aircraft spare parts affect flight safety and the cost of spare parts accounts for the highest proportion of airlines' controllable costs. It is significant to make spare parts planning on the premise of ensuring flight safety. A spare parts quantity model is built based on the Markov birth and death process. Supply rates for spare parts are determined according to the spare parts' security level and repair time limit. A spare parts dynamic programming model is established aiming at minimum cost while meeting a certain supply rate. Based on aircraft status, the effect of historical failure data, seasonality, daily utilization rate, and fleet size on spare parts coverage was examined. To reduce the discrepancy between the proposed model's calculated results and the actual demand, the seasonal differences in failure rate, daily utilization rate, and seasonal differences in aircraft parking losses were evaluated. The proposed model is verified by the example of the ERJ-190 Display Unit. Results are consistent with the actual operation of the airline, which proves that the proposed model can provide technical support for spare parts planning.

Keywords: spare parts; aircraft state; Markov birth and death process; economy; dynamic planning

Received: 2022-03-22; **Accepted**: 2022-05-06; **Published Online**: 2022-05-27 11:10
URL: kns.cnki.net/kcms/detail/11.2625.V.20220525.1933.004.html
* **Corresponding author**. E-mail: zy2114112@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0195

融合边界处理机制的学习型麻雀搜索算法

王子恺[1]，黄学雨[2,*]，朱东林[1]，闫少强[3]，李权[1]，郭伟[1]

(1. 江西理工大学 信息工程学院，赣州 341000；2. 江西理工大学 软件工程学院，南昌 330013；
3. 火箭军工程大学 作战保障学院，西安 710025)

摘 要：为改善麻雀搜索算法（SSA）初始化阶段种群分布不充分，寻优过程中容易受到局部最优解干扰的不足，提出融合边界处理机制的学习型麻雀搜索算法(HSSA)。使用Piecewise map初始化种群，提高种群的分散程度；使用排序配对学习与竞争学习策略分别更新跟随者和警戒者，确保各代的最优解信息能够引导下一代的位置更新；自适应的警戒者数量使得警戒者作用被强调，提供灵活的应变机制；根据不同阶段的寻优特点制定多策略边界处理机制，保留住种群数量的同时，为超出边界的个体提供更加合理的搜索位置。经过12个基准函数的仿真实验，并借助消融实验、Wilcoxon秩和检验等证明了HSSA在收敛速度上的稳定性和寻优的高效性。

关 键 词：麻雀搜索算法；Piecewise map；排序配对学习；竞争学习；多策略边界处理
中图分类号：TP301.6
文献标志码：A **文章编号**：1001-5965(2024)01-0286-13

各领域发展迅猛，随着复杂的优化问题不断涌现，传统的优化方案已经无法在时间和精度上同时满足技术发展所需求的高效率。研究者们根据自然界中生物的行为，探索出了一系列群智能优化算法。常见的群智能优化算法包括粒子群算法(particle swarm optimization algorithm, PSO)[1-2]、蚁群(ant colony system, ACS)算法[3]、人工蜂群算法(artificial bee colony algorithm, ABC)[4]、灰狼优化(grey wolf optimizer, GWO)算法[5]、鲸鱼优化算法(whale optimization algorithm, WOA)[6]等。相比于传统的优化方法，群智能算法具有参数少、运行时间短、实验操作简单等优点。在当今的非线性和多模态的寻优问题中呈现出了优良的寻优能力和可行性。

2020年，Xue和Shen[7]提出了一种新的群智能优化麻雀搜索算法(sparrow search algorithm, SSA)。该算法实现的难度较低，自组织能力较好，且在一个完整的捕食流程中展现出了较好的全局勘探和局部开采能力。但是，寻优速度过快也使得该算法在寻找最优解的过程中容易过快地收敛，提前进入到一个停滞的状态。针对这些问题，研究者们做了大量的尝试。吕鑫等[8]结合Tent混沌序列和高斯变异策略提出了混沌SSA(chaos sparrow search optimization algorithm, CSSA)，对局部个体的混沌扰动使得在解决陷入局部最优的问题上受益良多，减少算法的限制，避免不必要的停滞。毛清华和张强[9]针对算法的迭代后期局部极值的早熟问题提出了一种基于柯西变异和反向学习的改进SSA(improved sparrow algorithm combining Cauchy mutation and opposition-based learning, ISSA)，先是运用sin混沌初始化种群，然后分别将自适应权重、柯西变异和反向学习策略融入在发现者和加入者更新阶段。不仅增加了新解，还增强了算法跳出局部空间的能力。Zhu等[10]提出将透镜反向学习策略用在发现者的位置上更新之后，又在加入者的更新阶段融合

收稿日期：2022-03-28；录用日期：2022-04-29；网络出版时间：2022-05-10 08:41
网络出版地址：link.cnki.net/urlid/11.2625.V.20220507.2247.001
基金项目：国家重点研发计划(2020YFB1713700)
*通信作者. E-mail：1137753107@qq.com

*引用格式：王子恺，黄学雨，朱东林，等. 融合边界处理机制的学习型麻雀搜索算法[J]. 北京航空航天大学学报，2024，50(1)：286-298.
WANG Z K，HUANG X Y，ZHU D L，et al. Learning sparrow search algorithm of hybrids boundary processing mechanisms[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(1)：286-298 (in Chinese).*

了改进的正弦余弦机制, 进一步提升算法的搜索空间深度。Zhou 等[11] 尝试在 SSA 的侦察者更新阶段加入交叉和变异操作, 并结合受控段相位值的改进方法使之更好地应用在嘈杂的自然环境中。Yan 等[12] 将加入者阶段再细分成全局搜索和局部搜索 2 个阶段, 分别加入可变螺旋因子和改进迭代搜索策略并使用边界外个体数量的变化验证改进策略的效果。此外, 引入了逐维透镜学习和搜索范围更加灵活的动态边界。

SSA 引起了学者们很大的研究兴趣, 就结果来看, 在原始算法的改进上仍有较大发掘潜力, 且现有的改进存在一定缺点。比如, 随机初始化方式过于简单, 难以在多样性上得到保证。边界处理机制采用了直接无差异置于边界或是随机的初始位置的形式, 这些机制都欠缺合理性。上述不足不仅会损失掉 SSA 中种群的多样性, 也会在很大程度上影响收敛的效果。正如没有免费午餐定理描述的: 没有任何一种元启发式算法能够适用于所有的优化问题[13]。本文提出融合边界处理机制的学习型 SSA(简称 HSSA)来克服这些缺点, 创新点如下:

1) 通过引入 Piecewise map 提高了种群的分散程度, 增加了多样性。

2) 提出组合的学习机制, 引入排序配对学习和竞争学习策略。排序配对学习能够从容应对跟随者更新后的无序场面, 而竞争学习的使用强调了每一次迭代中的较优解。这样做既确保大部分较优解能够进入下一次迭代, 又能够以此对差解进行学习处理, 进而既保留了竞争力, 也提高了避免过早停滞的能力。

3) 提出自适应警戒者机制, 通过在各个阶段控制警戒者的比例从而增大警戒作用, 以此监督整个种群的各种行为。

4) 提出全新的多策略边界处理机制, 在保留运动方向的同时结合了当前最优位置、当前位置等信息, 避免了多样性的丧失, 也使得边界处理机制更加合理化。

1　麻雀搜索算法

种群中麻雀发现了危险便发出警报, 得到的报警值和安全阈值进行比较, 根据不同的结果按照不同计算式进行更新:

$$\boldsymbol{X}_i^{t+1}=\begin{cases}\boldsymbol{X}_i^t\exp\left(\dfrac{-i}{\alpha T}\right) & R_2<S_{\mathrm{T}}\\ \boldsymbol{X}_i^t+Q\boldsymbol{L} & R_2\geqslant S_{\mathrm{T}}\end{cases}\tag{1}$$

式中: $X_{i,j}^t$为第 t 代, 第 i 个麻雀在第 j 维中的位置, $\boldsymbol{X}_i^t=[X_{i,1}^t,X_{i,2}^t,\cdots,X_{i,J}^t]$; T 为最大迭代次数; t 为当前迭代次数; α 为符合$(0,1]$的均匀数; Q 为服从标准正态分布的随机数; $\boldsymbol{L}$ 为全 1 的矩阵; 报警值$R_2\in[0,1]$, 安全值$S_{\mathrm{T}}\in[0.5,1]$。

加入者位置更新如下:

$$\boldsymbol{X}_i^{t+1}=\begin{cases}Q\exp\left(\dfrac{\boldsymbol{X}_{\mathrm{w}}^t-\boldsymbol{X}_i^t}{i^2}\right) & i>N/2\\ \boldsymbol{X}_{\mathrm{P}}^{t+1}+\left|\boldsymbol{X}_i^t-X_{\mathrm{P}}^{t+1}\right|\boldsymbol{A}^+\cdot\boldsymbol{L} & \text{其他}\end{cases}\tag{2}$$

式中: $\boldsymbol{X}_{\mathrm{w}}^t$为当前全局最差位置; N 为麻雀总数; $\boldsymbol{A}$ 为一个 $1\times d$的矩阵, 元素随机赋值−1 或 1; $\boldsymbol{A}^+$满足 $\boldsymbol{A}^+=\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}})^{-1}$。

当$i>N/2$时, 表示该加入者的适应度值低, 无权争取食物, 必须去往新的地点找寻觅食机会; $i\leqslant N/2$时, 表示该加入者竞争到在最优位置附近进行觅食的机会。

和发现者一样, 通常将总数的 10%～20% 的个体作为警戒者, 时刻监视并在危险时提醒整个种群进行反哺食行为。其位置更新如下:

$$\boldsymbol{X}_i^{t+1}=\begin{cases}\boldsymbol{X}_i^t+\beta\left|\boldsymbol{X}_i^t-\boldsymbol{X}_b^t\right| & f_i\neq f_{\mathrm{g}}\\ \boldsymbol{X}_i^t+K\dfrac{\boldsymbol{X}_i^t-\boldsymbol{X}_{\mathrm{w}}^t}{(f_i-f_{\mathrm{w}})+\varepsilon} & f_i=f_{\mathrm{g}}\end{cases}\tag{3}$$

式中: β 为服从标准正态分布的随机数; K 为$[-1,1]$内的随机数; ε 为无穷小的常数, 避免分母为 0 的情况; f_i、f_{g}和f_{w}分别为当前麻雀的适应度、全局最优和全局最差的适应度值。

2　本文算法

2.1　Piecewise map 初始化种群

Li 和 Yorke[14] 较早提出混沌的概念, 来描述自然界中特殊的运动形式。在现代群智能优化算法中, 研究者们尝试将各种混沌序列用于多种群智能算法的初始化种群操作, 在增加了理论依据的同时, 获得了优良的收敛结果。

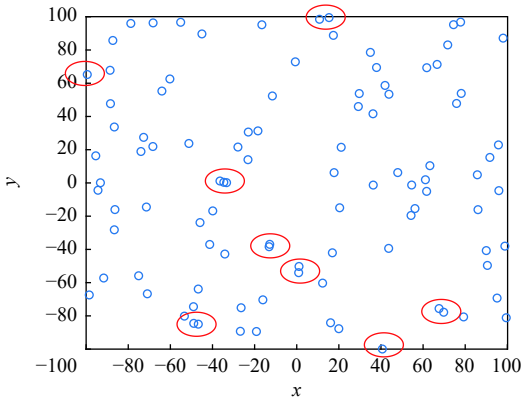
为了改善 SSA 通过后续寻优方式获得的解的质量, 一个优良的初始种群分布是整个算法的基础。本文采用分段映射 Piecewise map 来对 HSSA 的初始化阶段进行混沌变换。Piecewise map 作为分段线性映射的一种, 和 Liebovitch map、Bernoulli shift map 一样, 其对分布在不同位置的初始值设置了不同的变换式。具体的数学表达式为

$$X_{n+1}=\begin{cases}\dfrac{X_n}{d} & 0\leqslant X_n<d\\ \dfrac{X_n-d}{0.5-d} & d\leqslant X_n<0.5\\ \dfrac{1-d-X_n}{0.5-d} & 0.5\leqslant X_n<1-d\\ \dfrac{1-X_n}{d} & 1-d\leqslant X_n<1\end{cases}\tag{4}$$

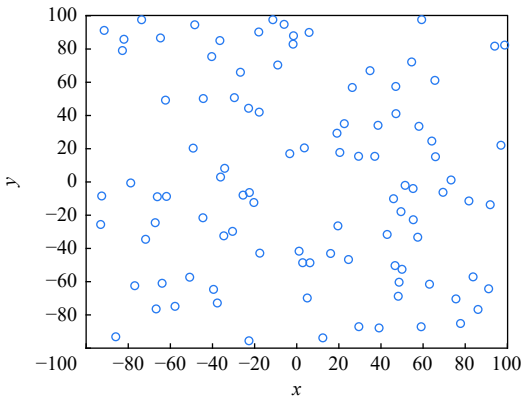
显然, d 为一个控制参数, 用于决定四段式的分

段范围使其之间无重叠的部分。实验证明，$d=0.3$ 时效果最佳，故本文实验中 d 均取0.3。

图 1 为种群初始分布。由图可知，映射后的初始化分布中群体的分散程度更高，且在边界上的个体和重合的个体数量更少（见图 1 中椭圆部分）。在初始化阶段分布的广度较大能够保证种群的多样性，且减少局部最优的吸引。



(a) 随机初始化



(b) Piecewise map 初始化

图 1 初始种群分布

Fig. 1 Distribution of initial population

2.2 自适应警戒者策略

警戒者作为 SSA 中不可或缺的一环，扮演着突发事件的角色，代表了种群习性中的预警能力。在原始的 SSA 中，警戒者数量在种群初始化前设置为 10%～20% 之间的定值，用于监察整个种群的动向。但是，固定不变的警戒者数量并不能够最大化其在整个过程中的作用。实际上，在觅食行为的初期，全局搜索的要求更高，需要更多的警戒者去尝试竞争最优位置，迫使领头的个体能够提高自身的勘探能力；而随着迭代的进行，少量的警戒者既能够保留一定的跳出机制，又不会频繁的干扰对于最优解的开发。因此，本文对警戒者的数量占全体数量的比例 P_{SD} 进行动态的改变，最大化警戒者的作用，表达式为

$$P_{SD}=20\frac{e^{a}}{e^{a}+1} \tag{5}$$

式中：$a=\dfrac{T}{t}-1$。

通过式(5)可知，P_{SD} 初期占据整个种群的 20%，随着迭代的进行，逐渐减小到 10% 附近，有效增大了警戒者机制在全局的发挥。

2.3 多策略学习机制

学习机制是模拟日常生活中知识传递算法提出的一种信息传递方法，将预设的一些位置或速度信息赋予其他的个体或向量，从而对该个体或向量进行变换，最知名的学习机制是反向学习[15]。而随着研究的深入，一些结合物理学知识的新型学习机制也逐渐进入人们的视野，如透镜学习[16]、精英反向学习、一般反向学习[17] 等。本文将介绍排序配对学习（ranking paired learning, RPL）策略和竞争学习策略 2 种优秀的学习策略并组合使用，最大程度上体现算法的性能优势。

2.3.1 排序配对学习策略

通常，在日常学习中，研究者们会分成小组进行学习交流。在寻优问题中亦是如此。整个种群中，适应度较差的麻雀个体需要根据特定的原则向适应度高的个体进行学习，在学习交流中提升自己觅食的能力，从而获得更多参与竞争的机会。Deng 等[18] 就配对学习的策略进行研究，并提出了适应度差最大化原则。研究表明，将最优和最差的学生进行配对学习不能够实现个体之间的最佳辅导效率。而如果将学生的能力进行排序，并分成前半部分和后半部分，再分别由后半部分的学生对应向前面部分的同序号学生进行学习，这样可以实现辅导效率的最大化[18]。在 SSA 中，麻雀类比学生，而适应度值就代表成绩。基于上述原则，提出了一种全新的学习策略 RPL 策略。其基本思想描述如下：将种群中的全部的 M 个体按照适应度大小进行排序，平均分成 2 个自然组，组内个体位置占优的组称为范例组 G_E，个体位置记为 $\boldsymbol{X}_i^E$；个体位置差的组称为学习组 G_L，个体位置称为 $\boldsymbol{X}_i^L$，两组均有 $M/2$ 个个体。图 2 为 RPL 的示意图。通过上述描述可知，RPL 应满足：

$$G_E\cap G_L=\varnothing \tag{6}$$

$$\begin{cases}\forall \boldsymbol{X}_i^E\in G_E\\ \forall \boldsymbol{X}_i^L\in G_L\end{cases}\qquad f\left(\boldsymbol{X}_i^E\right)<f\left(\boldsymbol{X}_i^L\right) \tag{7}$$

$$\begin{cases}f\left(\boldsymbol{X}_i^E\right)<f\left(\boldsymbol{X}_{i+1}^E\right)\\ f\left(\boldsymbol{X}_i^L\right)<f\left(\boldsymbol{X}_{i+1}^L\right)\end{cases} \tag{8}$$

式中：$i=1,2,\cdots,M/2$。

排序完成之后，G_L 组按照如下方案开始向

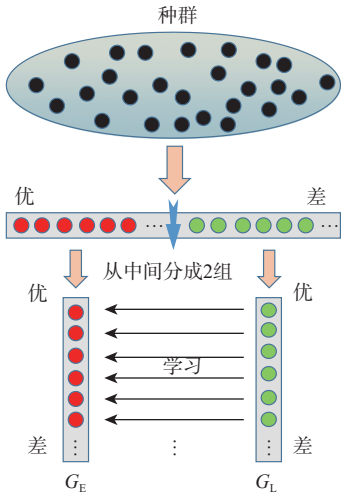


图 2　RPL 示意图

Fig. 2　Schematic diagram of RPL

G_E 组进行配对学习:

$$\boldsymbol{X}_1^{\mathrm{L}} \to \boldsymbol{X}_1^{\mathrm{E}} \cdots \boldsymbol{X}_{M/2}^{\mathrm{L}} \to \boldsymbol{X}_{M/2}^{\mathrm{E}} \tag{9}$$

RPL 之所以被提出是为了将位置较差的个体通过学习机制往好的方向移动, 并在不改变优势群体适应度的前提下使得整个麻雀种群都能够整体成长。

文献 [19] 模拟现实生活中人们喜欢效仿明星行为这一现象提出一种社会模拟优化算法(social mimic optimization algorithm, SMO), 同时, 提出了一种模仿的方法。该方法中通过定义模仿者同被模仿者之间的差值, 并以该差值进行模拟操作。将模拟步骤同上述 RPL 进行结合, 定义参数 K 表示学习者 $\boldsymbol{X}_i^{\mathrm{L}}$ 与被学习者 $\boldsymbol{X}_i^{\mathrm{E}}$ 的差值, 改进后的 K 满足:

$$K = \begin{cases} \dfrac{f\left(\boldsymbol{X}_i^{\mathrm{E}}\right) - f\left(\boldsymbol{X}_i^{\mathrm{L}}\right)}{f\left(\boldsymbol{X}_i^{\mathrm{L}}\right)} & K \neq 0 \\ \cos\left(\dfrac{\pi(T-t)}{2T}\right) & K = 0 \end{cases} \tag{10}$$

当 K 的值不为 0, 满足式(10)的上面式子; 反之, K 按照式(10)下面的式子得到一个(0,1)的随机值。$\boldsymbol{X}_i^{\mathrm{L}}$ 的学习计算式满足:

$$\boldsymbol{X'}_i^{\mathrm{L}} = (K+1)\,\boldsymbol{X}_i^{\mathrm{L}} \tag{11}$$

在跟随者阶段将学习后得到的解和之前的解进行比较, 通过贪婪原则保留更优的位置。贪婪选择的数学表达式如下:

$$\boldsymbol{X}_i = \begin{cases} \boldsymbol{X'}_i^{\mathrm{L}} & f\left(\boldsymbol{X'}_i^{\mathrm{L}}\right) \leqslant f\left(\boldsymbol{X}_i^{\mathrm{L}}\right) \\ \boldsymbol{X}_i^{\mathrm{L}} & f\left(\boldsymbol{X'}_i^{\mathrm{L}}\right) > f\left(\boldsymbol{X}_i^{\mathrm{L}}\right) \end{cases} \tag{12}$$

2.3.2　竞争学习策略

SSA 之所以寻优速度快, 离不开完整的搜索机制。为在进入下一次循环之前能够对解的质量有所保证, 保持较优解的存在及将给予较差解的学习机会可以很好地解决较优解丢失的问题, 同时保证下一次输入的精度。在该思想的基础上提出了一种竞争学习的策略, 其思想可描述为: 在每一次迭代中, 将该阶段种群中的所有个体都进行两两组队。本文设置个体数目为 M, 则拥有 $M/2$ 个小组。对于每个小组中的 2 个个体, 都进行适应度的比较。将适应度好的直接保存并进入下一次迭代, 而对于适应度差的个体, 则要采取相应的学习机制(见式(13))来更新自己的位置:

$$X'^{t}_{i,j} = X'^{t}_{i,j} + \lambda\left(X'_{\mathrm{b},j} - X'^{t}_{i,j}\right) \tag{13}$$

式中: λ 为服从正态分布的随机数; $X'_{\mathrm{b},j}$ 为当前迭代中竞争学习策略开始前排序时的最优解。每一次竞争学习之后, 半数的个体位置会得到更新并以新值传递至下一次迭代中[20]。图 3 为竞争学习的示意图。

由图 3 可知, 每一组竞争中的优胜者都会将自己的信息作为参考提供给竞争失败者进行学习。关于学习方法, 文献 [11] 提出对阶段中的最优个体进行学习, 文献 [17] 认为不仅是最优个体, 一次迭代后的最差位置信息同样具有更新的意义。基于对上述文献的研究, 本文认为 2 种策略都不能够充分地利用种群中的有用信息, 使得迭代的过程中最优位置的信息被过分强调了。

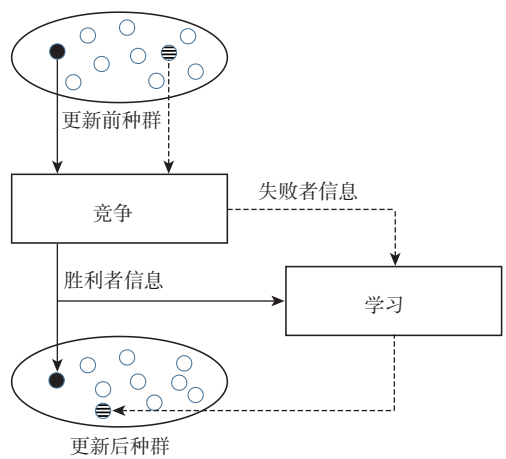


图 3　竞争学习示意图

Fig. 3　Schematic diagram of competitive learning

因此, 本文结合竞争学习的策略提出了将随机匹配的 2 个个体进行比较之后再更新较差的个体, 胜出者直接晋级下一次迭代的方法。这样不仅不会使得较优的个体位置丢失, 最大程度地运用了优秀个体的信息。同时, 随机选择竞争个体的机制也保留了群智能算法的随机性这一特点, 便于验证算法的鲁棒性。

2.4　多策略边界处理

边界处理策略是一种保存种群多样性的重要

方法之一，对于超出边界的个体位置进行重新定义使得有效位置之外的群体重新回到有效序列。纵观整个SSA，发现者更新阶段的核心目的是在全局范围内进行寻优，而跟随者阶段则是根据之前寻到的位置信息判断其是否为找寻的全局最优值。原始的SSA直接将超出的个体无差别更新在既定的边界上的方法会造成边界局部聚集，进而会造成多样性的缺失，为接下来的优化带来恶劣影响。整个SSA分3个阶段，发现者、跟随者(加入者)、警戒者。因此，鉴于SSA中3个阶段的差异性，本文提出了一种全新的多策略边界处理方法，根据不同更新阶段的目的进行针对性处理。

在发现者阶段，整个种群应当致力于寻找新的食物，当前最优即是潜在的最合适觅食点。但是觅食的初期伴随着最优位置的不确定性，故个体的边界处理应当保留一定的随机性。同理，在警戒者更新之后，紧接着就会是下一次迭代的开始，初始化的种群越分散，相应阶段的策略开展获取最优解的可能性就越大。随机性较强的处理方法赋予了种群的一定几率的跳出局部最优的能力。该阶段处理方法如下：

$$\boldsymbol{X}_i' = \boldsymbol{u}_{\mathrm{b}} - (\boldsymbol{u}_{\mathrm{b}} - \boldsymbol{l}_{\mathrm{b}})\mathrm{rand} \qquad \boldsymbol{X} > \boldsymbol{u}_{\mathrm{b}} \text{或} \boldsymbol{X} < \boldsymbol{l}_{\mathrm{b}} \tag{14}$$

式中：$\boldsymbol{u}_{\mathrm{b}}$和$\boldsymbol{l}_{\mathrm{b}}$分别为搜索范围的上界和下界。

文献[12]证明加入者更新之后，超出边界的潜在数量达到3个更新阶段后最多的数量[12]。在原来的跟随者机制中，适应度较好的跟随者已经向当前全局最优附近移动，保证了潜在最优区域的搜索。因此，该阶段超出边界的个体处理方式也应当围绕当前最优位置$\boldsymbol{X}_{\mathrm{b}}$展开。

考虑到原来的SSA中，跟随者阶段的更新操作会产生2种情况，其一维示意图如图4所示。图4(a)的情况是个体更新之后的位置超出上边界，图4(b)则是更新后的位置超出下边界。为了使处理后的麻雀个体重回既定边界内，又能够充分运用到几个重要位置(当前全局最优$\boldsymbol{X}_{\mathrm{b}}$、更新前的位置$\boldsymbol{X}_i$、照式(4)更新后的跟随者位置$\boldsymbol{X}_i$、函数的上下界$\boldsymbol{u}_{\mathrm{b}}$和$\boldsymbol{l}_{\mathrm{b}}$)的信息。同时，为了和原始跟随者的更新方向保持一致(如图4中箭头所示)，按照超出边界的2种不同情况提出如下表达式。

超出上边界：

$$\boldsymbol{X}_i'' = \boldsymbol{X}_{\mathrm{b}} + |\boldsymbol{u}_{\mathrm{b}} - \boldsymbol{X}_{\mathrm{b}}| \circ \frac{|\boldsymbol{X}_i - \boldsymbol{u}_{\mathrm{b}}|}{|\boldsymbol{X}_i' - \boldsymbol{X}_i|} \tag{15}$$

超出下边界：

$$\boldsymbol{X}_i'' = \boldsymbol{X}_{\mathrm{b}} - |\boldsymbol{l}_{\mathrm{b}} - \boldsymbol{X}_{\mathrm{b}}| \circ \frac{|\boldsymbol{X}_i - \boldsymbol{l}_{\mathrm{b}}|}{|\boldsymbol{X}_i' - \boldsymbol{X}_i|} \tag{16}$$

式中："∘"表示哈达玛积。

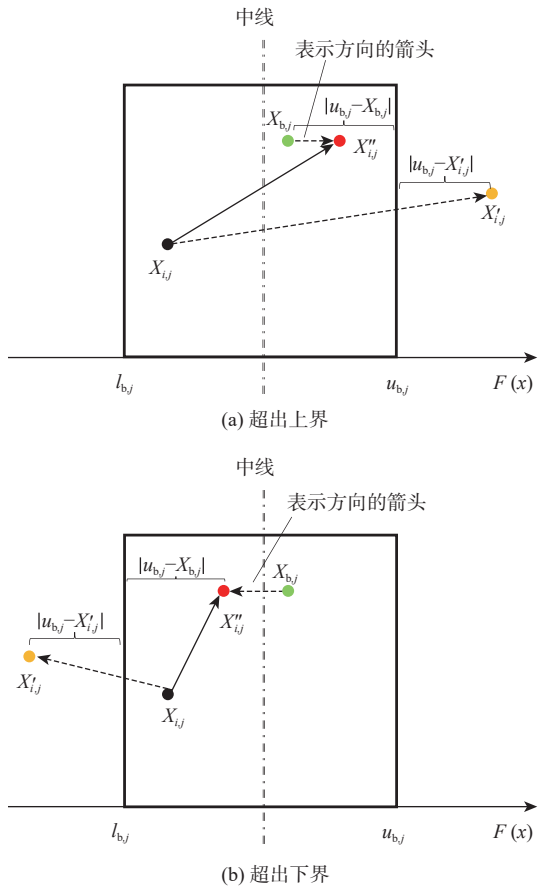


图4 跟随者阶段边界示意图

Fig. 4 Schematic diagram of follower phase boundary processing

不难看出，$|\boldsymbol{l}_{\mathrm{b}} - \boldsymbol{X}_{\mathrm{b}}|$和$|\boldsymbol{u}_{\mathrm{b}} - \boldsymbol{X}_{\mathrm{b}}|$控制了步长的单位距离，而后面的分式控制了单位数。两者的结合实现了对于位置更新的步长控制，能够保证处理后的位置在预设的范围内。

2.5 HSSA流程

HSSA利用Piecewise map初始化种群，在跟随者和警戒者的位置更新完成后分别加入RPL和竞争学习，最大化每一次迭代中较优部分解的信息并以此来更新差解，并保证将最优解群传递至下一代中。此外，自适应警戒者的处理使得警戒者群体的重要性被强调。同时，根据不同阶段的更新制定多策略边界处理机制，使得种群分布更合理，并保证了种群的多样性。具体实现步骤如下。

步骤1 设置初始化种群及各参数，包括种群规模N，发现者比例P_{PD}、警戒者比例P_{SD}，目标函数的维度设置为D，初始化种群分布区域的上下界分别设定为$\boldsymbol{u}_{\mathrm{b}}$、$\boldsymbol{l}_{\mathrm{b}}$，最大迭代次数$T$，报警阈值$\delta$，求解精度$\varepsilon$。

步骤2 使用式(4)中的Piecewise map初始化种群。

步骤3 计算每只麻雀的适应度f_i，选出当前

最优适应度f_b和其所对应的位置$\boldsymbol{X}_b$, 以及当前最劣适应度f_w和其对应的位置$\boldsymbol{X}_w$。

步骤 4　根据设置的比例P_{PD}随机选取适应度最优的N_b个麻雀作为发现者, 剩下的为加入者, 并根据式 (1) 更新发现者的位置。按照式 (14) 进行边界处理。

步骤 5　确定跟随者群体, 并根据式 (2) 更新加入者的位置。

步骤 6　对当前种群按照适应度值再次排序, 并根据式 (10)～式 (12) 再次更新种群。按式 (15) 和式 (16) 进行边界处理。

步骤 7　从种群中按照比例P_{SD}随机生成N_s个警戒者, 并根据式 (5) 对P_{SD}进行更新。同时, 根据式 (3) 和式 (13) 进行警戒者位置更新操作, 按照式 (14) 进行边界处理。

步骤 8　根据麻雀种群当前的状态, 更新整个种群在整个觅食过程中的最优位置$\boldsymbol{X}_b$, 最佳适应度值f_b, 以及最差位置$\boldsymbol{X}_w$和最劣适应度值f_w。

步骤 9　判断迭代是否结束。若算法达到最大迭代次数或求解精度达到设定值, 则判定循环结束, 并输出寻优结果; 否则返回步骤 3 继续下一次迭代操作, 且当前迭代次数$t' = t+1$。

2.6　复杂度分析

时间复杂度是一个判断算法性能和计算运行代价的重要指标之一。从宏观和微观 2 个角度分析 HSSA 是否增加了时间复杂度。首先, 在宏观角度, 假设算法的最大迭代次数为 T, 维度为 D, 种群规模为 M。那么, 根据智能优化算法的时间复杂度计算式可知, SSA 的时间复杂度为$O_1 = O(TDM)$。而对于 HSSA, 结构上并没有发生改变, 所以, 其时间复杂度$O_2 = O(TDM)$。显然, 在宏观上时间复杂度没有增加。另一方面, 从微观角度可以得到, 基于排序配对学习和竞争学习的时间复杂度O_3、O_4分别为$O_3 = O(TDM)$、$O_4 = O(TDM)$; 边界处理和初始化阶段并未增加时间复杂度。综上可知, 在微观角度, HSSA 时间复杂度增加了$O_3 + O_4 = O(TDM)$, 但是各步骤的增加都没有引起数量级的变化, 总的时间复杂度依然是$O(TDM)$。

综上可知, 不管是宏观还是微观的角度, 时间复杂度都没有发生变化。可见, HSSA 并没有增加消耗, 进而证明其的可行性。

3　仿真实验和结果分析

为了验证 HSSA 的性能, 测试其在收敛速度和求解精度上的表现, 本文选取了 12 个常用的基准函数进行验证, 并进行 2 组实验。所选基准函数的理论最优值都是 0。第 1 组将 HSSA 和最新的 5 个 SSA 及其变体进行比较, 这些变体分别是文献 [17] 提出的多策略学习型 SSA(简称 IHSSA)、文献 [10] 中的学习型 SSA(learning sparrow search algorithm, LSSA)、文献 [21] 提出的增强型 SSA(enhanced sparrow search algorithm, ESSA)和文献 [12] 中所述的基于迭代局部搜索的 SSA(简称 YSSA), 旨在判断 HSSA 是否能够超过其他变体的性能。第 2 组实验将 HSSA 和经典的算法 PSO、GWO、WOA、教与学优化(teaching learning based optimization, TLBO)算法、蝴蝶优化算法(butterfly optimization algorithm, BOA)进行比较, 旨在测试麻雀算法的寻优能力。具体的参数设置如表 1 所示, 基准函数的相关信息如表 2 所示。为了公平起见, 每个算法的种群数量都设置为 100, 最大迭代次数为 500。此外, 为了避免实验数据的偶然性, 各个算法独立运行 30 次, 计算其平均值、标准差和最优值, 并对最优的数据进行加粗处理。其中, 第 1 组实验整理在表 3 中, 第 2 组整理在表 4 中。本文的实验环境均为 Intel(R) Core(TM) i5-9 300H CPU @ 2.40GHz. 并采用 MATLAB2019a 进行仿真。

表 1　算法参数

Table 1　Algorithm parameter

算法	参数	数值
SSA	P_{PD}	0.2
	P_{SD}	0.6
	δ	0.6
IHSSA[17]	P_{PD}	0.2
	P_{SD}	0.2
LSSA[10]	P_{PD}	0.2
	P_{SD}	0.2
ESSA[21]	P_{PD}	0.2
	P_{SD}	0.2
YSSA[12]	P_{PD}	0.2
	P_{SD}	0.2
HSSA	P_{PD}	0.2
	δ	0.8
	d	0.3
PSO	c_1	1.494 45
	c_2	1.494 45
GWO	a_{max}	2
	a_{min}	0
BOA	a	0.1
	b	0.025
	c	0.1
	p	0.8

表2 基准函数

Table 2 Benchmark function

函数	维度	搜索区域
$F_1(x)=\sum_{i=1}^{n}x_i^2$	30	[−100,100]
$F_2(x)=\sum_{i=1}^{n}\lvert x_i\rvert+\prod_{i=1}^{n}\lvert x_i\rvert$	30	[−10,10]
$F_3(x)=\sum_{i=1}^{n}\left(\sum_{j=1}^{i}x_j\right)^2$	30	[−100,100]
$F_4(x)=\max_i\{\lvert x_i\rvert,1\leqslant i\leqslant n\}$	30	[−100,100]
$F_5(x)=\sum_{i=1}^{n-1}\left[100(x_{i+1}-x_i^2)^2+(x_i-1)^2\right]$	30	[−30,30]
$F_6(x)=\sum_{i=1}^{n}(x_i+0.5)^2$	30	[−100,100]
$F_7(x)=\sum_{i=1}^{n}ix_i^4+\mathrm{rand}(0,1)$	30	[−1.28,1.28]
$F_8(x)=\sum_{i=1}^{n}\left[x_i^2-10\cos(2\pi x_i)+10\right]$	30	[−5.12,5.12]
$F_9(x)=-20\exp\left(-0.2\sqrt{\frac{1}{2}\sum_{i=1}^{n}x_i^2}\right)-\exp\left(\frac{1}{n}\sum_{i=1}^{n}\cos(2\pi x_i)\right)+20+\mathrm{e}$	30	[−32,32]
$F_{10}(x)=\frac{1}{4\,000}\sum_{i=1}^{n}x_i^2-\prod_{i=1}^{n}\cos\left(\frac{x_i}{\sqrt{i}}\right)+1$	30	[−600,600]
$F_{11}(x)=\frac{\pi}{n}\left\{10\sin(\pi y_1)+\sum_{i=1}^{n-1}(y_i-1)^2\left[1+10\sin^2(\pi y_{i+1})\right]+(y_n-1)^2\right\}+\sum_{i=1}^{n}u(x_i,10,100,4)$，$y_i=1+\frac{x_i+1}{4}$，$u(x_i,a,k,m)=\begin{cases}k(x_i-a)^m & x_i>a\\0 & -a<x_i<a\\k(-x_i-a)^m & x_i<-a\end{cases}$	30	[−50,50]
$F_{12}(x)=\sum_{i=1}^{n}\left(z_i^2-10\cos(2\pi z_i)+10\right)$，$Z_i=\begin{cases}y_i & \lvert y_i\rvert<0.5\\\frac{\mathrm{round}(2y_i)}{2} & \lvert y_i\rvert\leqslant 0.5\end{cases}$	10	$[-10^{10},10^{10}]$

对于单峰函数F_1～F_4，HSSA和LSSA[17]、YSSA[12]一样，全都找到了理论最优值0，且性能远超其他的3种SSA变体。对于F_5和F_6，HSSA性能不如ESSA[21]和YSSA，但是和其他3种变体相比各指标都提升了1个量级，稳定性更出色。对于F_7，HSSA表现最优，寻到的最优解最接近理论最优值。相比之下，5个经典算法在7个单峰函数上都不能找到理论最优解，且落后很远。可见基础算法的劣势明显，改进的意义巨大。对于多峰函数F_8～F_{10}和定维函数F_{12}，SSA及其变体表现相当，在3项指标上看不出差距。但相比于经典算法，可见多个量级的提升。总的来说，HSSA较其他SSA及其变体相比，稳定性能更好，在7个函数中表现最佳。通过对比近一年内学术论文中提出的变体，使得本文具有较好的研究意义。而相比于经典算法，HSSA在9个测试函数中高居第一，其他3个函数表现第2，亦可证明其优势明显。同时，为了更直观地展示算法在各个函数上的表现，图5展示了6个SSA变体在12个测试函数的收敛曲线。

从图5中可以看出，HSSA和其他变体一样，寻优的速度很快。但是，考虑到各个算法改进的策略和侧重点不同，没有一个算法能够作为稳定的最好存在。F_1～F_4中，YSSA[12]和LSSA[10]表现最佳，HSSA排名第3，但是在F_5、F_6、F_{11}中，这2个算法表现很差，但HSSA依然保持可观的优势。对于F_7～F_{10}、F_{12}函数，6个变体都展示出了以SSA为基础的算法的收敛速度快的特点，HSSA表现稳定，稳居第2、第3位。

此外，为了更好地证明本文提及的逐个策略都在寻优过程中发挥了重要作用，本文将SSA加入Piecewise map初始化的麻雀算法（简称ISSA1）、加入排序配对学习的麻雀算法（简称ISSA2）、加入自适应警戒者策略和竞争学习的麻雀算法（简称ISSA3）、结合初始化和排序学习的麻雀算法（简称

表 3　HSSA 与 SSA 变体测试对比

Table 3　Comparison of test HSSA with SSA variants

函数	F_1			F_2			F_3			F_4		
	平均值	标准差	最优值	平均值	标准差	最优值	平均值	标准差	最优值	平均值	标准差	最优值
SSA	0	0	0	6.62×10^{-227}	0	0	1.85×10^{-254}	0	0	8.07×10^{-241}	0	0
IHSSA[17]	0	0	0	7.91×10^{-163}	0	5.34×10^{-164}	8.96×10^{-264}	0	0	4.48×10^{-163}	0	0
LSSA[10]	0	0	0	**0**	**0**	**0**	0	0	0	**0**	**0**	**0**
ESSA[21]	1.54×10^{-172}	0	0	6.54×10^{-95}	3.52×10^{-94}	0	1.57×10^{-121}	8.45×10^{-121}	0	8.26×10^{-79}	4.45×10^{-78}	0
YSSA[12]	0	0	0	**0**	**0**	**0**	0	0	0	**0**	**0**	**0**
HSSA	**0**	**0**	**0**	9.28×10^{-251}	0	0	**0**	**0**	**0**	2.73×10^{-265}	0	0

函数	F_5			F_6			F_7			F_8		
	平均值	标准差	最优值	平均值	标准差	最优值	平均值	标准差	最优值	平均值	标准差	最优值
SSA	1.88×10^{-5}	3.17×10^{-5}	1.90×10^{-8}	5.73×10^{-10}	2.25×10^{-9}	5.12×10^{-14}	1.54×10^{-4}	1.09×10^{-4}	4.65×10^{-6}	0	0	0
IHSSA[17]	7.72×10^{-7}	3.54×10^{-6}	8.74×10^{-12}	6.57×10^{-11}	1.16×10^{-10}	2.76×10^{-13}	1.99×10^{-4}	1.46×10^{-4}	9.25×10^{-6}	0	0	0
LSSA[10]	4.67×10^{-5}	1.96×10^{-4}	0	8.10×10^{-10}	2.28×10^{-9}	3.55×10^{-12}	8.23×10^{-5}	6.02×10^{-5}	5.65×10^{-6}	0	0	0
ESSA[21]	2.21×10^{-6}	7.70×10^{-6}	2.45×10^{-12}	$\mathbf{9.36\times10^{-13}}$	$\mathbf{1.21\times10^{-12}}$	1.03×10^{-14}	1.70×10^{-4}	1.17×10^{-4}	2.98×10^{-5}	0	0	0
YSSA[12]	**0**	**0**	**0**	9.58×10^{-6}	1.38×10^{-5}	$\mathbf{7.30\times10^{-19}}$	6.84×10^{-5}	5.03×10^{-5}	1.34×10^{-5}	0	0	0
HSSA	2.25×10^{-6}	5.36×10^{-6}	1.19×10^{-12}	4.26×10^{-10}	8.55×10^{-10}	9.79×10^{-15}	1.34×10^{-4}	$\mathbf{8.51\times10^{-5}}$	$\mathbf{1.65\times10^{-6}}$	**0**	**0**	**0**

函数	F_9			F_{10}			F_{11}			F_{12}		
	平均值	标准差	最优值	平均值	标准差	最优值	平均值	标准差	最优值	平均值	标准差	最优值
SSA	8.88×10^{-16}	9.86×10^{-32}	8.88×10^{-16}	0	0	0	2.23×10^{-11}	6.80×10^{-11}	1.52×10^{-15}	0	0	0
IHSSA[17]	8.88×10^{-16}	9.86×10^{-32}	8.88×10^{-16}	0	0	0	4.23×10^{-12}	7.02×10^{-12}	2.85×10^{-14}	0	0	0
LSSA[10]	8.88×10^{-16}	9.86×10^{-32}	8.88×10^{-16}	0	0	0	7.92×10^{-11}	1.87×10^{-10}	1.25×10^{-13}	0	0	0
ESSA[21]	8.88×10^{-16}	9.86×10^{-32}	8.88×10^{-16}	0	0	0	$\mathbf{1.17\times10^{-13}}$	$\mathbf{2.75\times10^{-13}}$	2.92×10^{-17}	0	0	0
YSSA[12]	8.88×10^{-16}	9.86×10^{-32}	8.88×10^{-16}	0	0	0	7.93×10^{-7}	1.29×10^{-6}	$\mathbf{1.40\times10^{-17}}$	0	0	0
HSSA	$\mathbf{8.88\times10^{-16}}$	$\mathbf{9.86\times10^{-32}}$	$\mathbf{8.88\times10^{-16}}$	**0**	**0**	**0**	1.90×10^{-11}	2.11×10^{-11}	1.76×10^{-16}	**0**	**0**	**0**

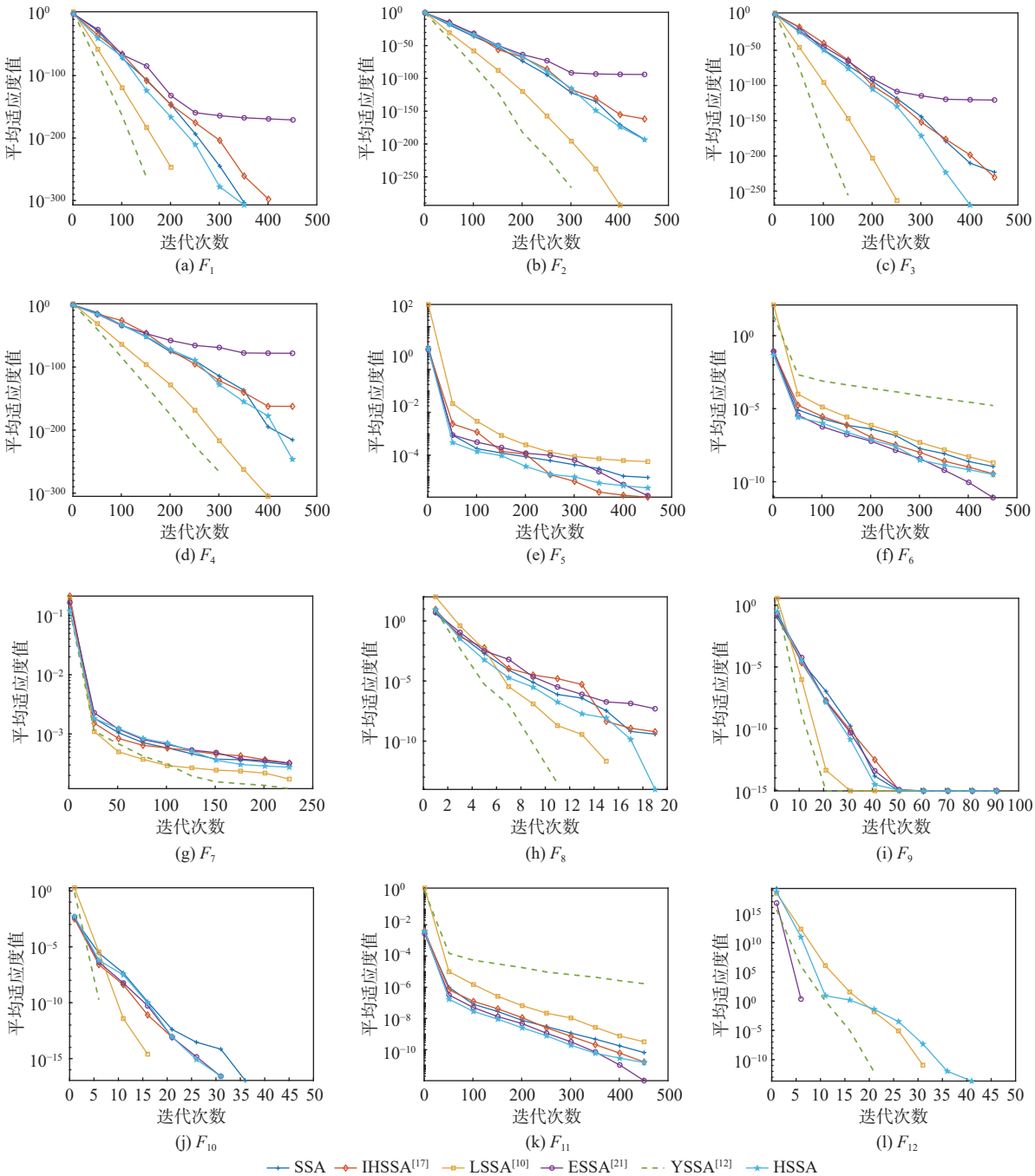
表 4　HSSA 与经典算法测试对比

Table 4　Comparison of test HSSA with classical algorithms

函数	F_1			F_2			F_3			F_4		
	平均值	标准差	最优值	平均值	标准差	最优值	平均值	标准差	最优值	平均值	标准差	最优值
PSO	7.85×10^{-11}	1.72×10^{-10}	4.53×10^{-12}	5.37×10^{-5}	1.25×10^{-4}	1.27×10^{-6}	5.87	3.40	2.17	2.10×10^{-1}	1.25×10^{-1}	4.70×10^{-2}
WOA	5.82×10^{-96}	1.63×10^{-95}	2.07×10^{-105}	1.67×10^{-57}	7.79×10^{-57}	2.40×10^{-63}	1.28×10^{4}	7.55×10^{3}	1.02×10^{3}	2.22×10^{1}	2.73×10^{1}	2.94×10^{-2}
GWO	7.47×10^{-41}	8.36×10^{-41}	2.40×10^{-42}	5.44×10^{-24}	3.42×10^{-24}	1.26×10^{-24}	1.03×10^{-11}	2.11×10^{-11}	1.92×10^{-14}	2.24×10^{-10}	2.92×10^{-10}	4.35×10^{-11}
BOA	9.77×10^{-12}	7.45×10^{-13}	8.57×10^{-12}	8.65×10^{-6}	2.65×10^{-5}	3.54×10^{-10}	8.89×10^{-12}	9.75×10^{-13}	6.52×10^{-12}	5.11×10^{-9}	2.67×10^{-10}	4.52×10^{-9}
TLBO	2.64×10^{-85}	3.31×10^{-85}	3.22×10^{-86}	1.16×10^{-42}	5.50×10^{-43}	4.67×10^{-43}	1.69×10^{-15}	3.61×10^{-15}	8.21×10^{-17}	2.15×10^{-34}	7.83×10^{-35}	9.95×10^{-35}
HSSA	**0**	**0**	**0**	$\mathbf{9.28\times10^{-251}}$	**0**	**0**	**0**	**0**	**0**	$\mathbf{2.73\times10^{-265}}$	**0**	**0**

函数	F_5			F_6			F_7			F_8		
	平均值	标准差	最优值	平均值	标准差	最优值	平均值	标准差	最优值	平均值	标准差	最优值
PSO	3.65×10^{1}	3.46	1.09	6.46×10^{-11}	9.90×10^{-11}	2.19×10^{-12}	1.65×10^{-2}	6.24×10^{-2}	7.95×10^{-2}	5.14×10^{1}	1.43×10^{1}	2.39×10^{1}
WOA	2.68×10^{1}	0.20	2.63×10^{1}	4.78×10^{-3}	2.50×10^{-3}	1.56×10^{-3}	1.13×10^{-3}	1.23×10^{-3}	7.51×10^{-5}	0	0	0
GWO	2.63×10^{1}	7.04×10^{-1}	2.52×10^{1}	2.50×10^{-1}	2.15×10^{-1}	2.20×10^{-5}	4.74×10^{-4}	2.71×10^{-4}	1.30×10^{-4}	4.00×10^{-1}	1.25	0
BOA	2.89×10^{1}	2.52×10^{-2}	2.89×10^{1}	5.25	5.34	3.67	$\mathbf{2.08\times10^{-5}}$	$\mathbf{2.19\times10^{-5}}$	$\mathbf{1.95\times10^{-7}}$	1.00×10^{-11}	9.53×10^{-13}	8.13×10^{-12}
TLBO	1.90×10^{1}	1.04	1.72×10^{1}	$\mathbf{4.30\times10^{-16}}$	$\mathbf{6.62\times10^{-16}}$	$\mathbf{1.02\times10^{-17}}$	8.53×10^{-4}	3.02×10^{-4}	4.72×10^{-4}	6.60	5.16	0
HSSA	$\mathbf{2.25\times10^{-6}}$	$\mathbf{5.36\times10^{-6}}$	$\mathbf{1.19\times10^{-12}}$	4.26×10^{-10}	8.55×10^{-10}	9.79×10^{-15}	1.34×10^{-4}	8.51×10^{-5}	1.65×10^{-6}	**0**	**0**	**0**

函数	F_9			F_{10}			F_{11}			F_{12}		
	平均值	标准差	最优值	平均值	标准差	最优值	平均值	标准差	最优值	平均值	标准差	最优值
PSO	1.04×10^{-1}	3.93×10^{-1}	7.48×10^{-7}	1.14×10^{-2}	1.45×10^{-2}	3.30×10^{-11}	4.49×10^{-2}	8.74×10^{-2}	5.87×10^{-14}	1.16×10^{20}	2.95×10^{19}	5.87×10^{19}
WOA	3.97×10^{-15}	2.71×10^{-15}	8.88×10^{-16}	4.21×10^{-3}	1.29×10^{-2}	0	1.91×10^{-1}	1.02	1.84×10^{-4}	5.92×10^{-17}	3.19×10^{-16}	0
GWO	2.75×10^{-14}	2.71×10^{-15}	2.22×10^{-14}	3.82×10^{-3}	7.78×10^{-3}	0	1.67×10^{-2}	9.03×10^{-3}	2.46×10^{-6}	2.10	3.01	0
BOA	4.45×10^{-9}	2.24×10^{-10}	3.89×10^{-9}	1.27×10^{-11}	7.39×10^{-13}	1.12×10^{-11}	7.15×10^{-1}	1.14×10^{-1}	5.00×10^{-1}	1.59×10^{7}	3.90×10^{7}	4.65×10^{4}
TLBO	5.98×10^{-15}	1.76×10^{-15}	4.44×10^{-15}	0	0	0	$\mathbf{1.05\times10^{-17}}$	$\mathbf{1.46\times10^{-17}}$	$\mathbf{6.56\times10^{-20}}$	5.11	1.03	3.09
HSSA	$\mathbf{8.88\times10^{-16}}$	$\mathbf{9.86\times10^{-32}}$	$\mathbf{8.88\times10^{-16}}$	**0**	**0**	**0**	1.90×10^{-11}	2.11×10^{-11}	1.76×10^{-16}	**0**	**0**	**0**



注: 对于 F_7～F_{10}、F_{12}, 6 个 SSA 变体早早地收敛, 再展示 500 代的收敛图不直观, 因此, 只展示到收敛至最优值附近的代数。

图 5 测试函数收敛曲线

Fig. 5 Convergence curve of test functions

ISSA4)、结合初始化和竞争学习的麻雀算法(简称ISSA5)、结合排序学习和竞争学习的麻雀算法(简称ISSA6)和结合所有策略的 HSSA 进行比较, 测试后的结果统计在表 5 中, 最优的数据加粗显示。

从表 5 中可以看出, 在 12 个测试函数中, HSSA 在 9 个函数上表现最佳。而相比于 SSA 中的数据, 各改进后的数据都有提升, 特别是在 F_5～F_7, 每一个 ISSA 都较 SSA 有所提升, 在其他的函数上, 寻优精度相当, F_1～F_4、F_8、F_{10}、F_{12} 寻到了理论最优值, 表现出色。综上所述, 本文各策略的改进效果很好, 但所有策略的组合性能最佳。

仅根据算法的平均值和标准差不能够全面的评价算法的性能, 本文采用 Wilcoxon 秩和检验对 HSSA 进行验证, 判断其与其他的 SSA 变体及经典算法是否具有较大的差异性。表 6 为显著水平 p 为 0.05 下的秩和检验结果。当 p<0.05 时, 可认为 2 种算法之间存在显著性差异; p>0.05 时, 则表明 2 种算法之间的差异性不明显; 表中空白处表示两组数据差异过小, 可以忽略。由表 6 可知, HSSA 与各算法均存在显著性差异, 仅个别算法存在性能

表 5　HSSA 与各改进策略消融实验对比

Table 5　Comparion of HSSA with ablation experiments for each improved strategy

函数	F_1			F_2			F_3			F_4		
	平均值	标准差	最优值	平均值	标准差	最优值	平均值	标准差	最优值	平均值	标准差	最优值
SSA	0	0	0	6.62×10^{-227}	0	0	1.86×10^{-254}	0	0	8.07×10^{-241}	0	0
ISSA1	0	0	0	2.12×10^{-189}	0	0	0	0	0	8.44×10^{-252}	0	0
ISSA2	0	0	0	1.71×10^{-198}	0	0	0	0	0	3.17×10^{-201}	0	0
ISSA3	0	0	0	**1.10×10^{-256}**	**0**	**0**	0	0	0	0	0	0
ISSA4	0	0	0	8.17×10^{-201}	0	0	0	0	0	5.46×10^{-195}	0	0
ISSA5	0	0	0	1.50×10^{-210}	0	0	3.72×10^{-237}	0	0	8.45×10^{-249}	0	0
ISSA6	0	0	0	3.67×10^{-225}	0	0	2.45×10^{-306}	0	0	1.01×10^{-262}	0	0
HSSA	**0**	**0**	**0**	9.28×10^{-251}	0	0	**0**	**0**	**0**	**2.70×10^{-265}**	**0**	**0**

函数	F_5			F_6			F_7			F_8		
	平均值	标准差	最优值	平均值	标准差	最优值	平均值	标准差	最优值	平均值	标准差	最优值
SSA	1.88×10^{-5}	3.17×10^{-5}	1.90×10^{-8}	5.73×10^{-10}	2.25×10^{-9}	5.12×10^{-14}	1.54×10^{-4}	1.09×10^{-4}	4.60×10^{-6}	0	0	0
ISSA1	1.86×10^{-5}	5.01×10^{-5}	1.20×10^{-9}	2.03×10^{-10}	4.54×10^{-10}	1.50×10^{-13}	1.23×10^{-4}	1.03×10^{-4}	1.21×10^{-5}	0	0	0
ISSA2	1.61×10^{-5}	3.30×10^{-5}	7.78×10^{-10}	7.68×10^{-11}	1.30×10^{-10}	4.92×10^{-15}	1.54×10^{-4}	1.46×10^{-4}	7.77×10^{-6}	0	0	0
ISSA3	5.04×10^{-6}	8.48×10^{-6}	2.46×10^{-10}	4.82×10^{-10}	9.39×10^{-10}	6.24×10^{-13}	1.28×10^{-4}	1.02×10^{-4}	1.92×10^{-6}	0	0	0
ISSA4	4.20×10^{-6}	4.76×10^{-6}	3.38×10^{-9}	9.93×10^{-11}	2.32×10^{-10}	1.15×10^{-13}	1.29×10^{-4}	8.72×10^{-5}	5.31×10^{-6}	0	0	0
ISSA5	1.37×10^{-6}	1.99×10^{-6}	1.57×10^{-9}	**3.65×10^{-11}**	**5.76×10^{-11}**	1.14×10^{-14}	1.58×10^{-4}	9.62×10^{-5}	6.21×10^{-6}	0	0	0
ISSA6	1.61×10^{-6}	3.81×10^{-6}	1.13×10^{-10}	8.31×10^{-11}	1.57×10^{-10}	**6.33×10^{-16}**	9.96×10^{-5}	8.64×10^{-5}	1.77×10^{-6}	0	0	0
HSSA	**2.25×10^{-6}**	**5.36×10^{-6}**	**1.19×10^{-12}**	4.26×10^{-10}	8.55×10^{-10}	9.79×10^{-15}	1.34×10^{-4}	**8.51×10^{-5}**	**1.65×10^{-6}**	**0**	**0**	**0**

函数	F_9			F_{10}			F_{11}			F_{12}		
	平均值	标准差	最优值	平均值	标准差	最优值	平均值	标准差	最优值	平均值	标准差	最优值
SSA	8.88×10^{-16}	9.86×10^{-32}	8.88×10^{-16}	0	0	0	2.23×10^{-11}	6.80×10^{-11}	1.52×10^{-15}	0	0	0
ISSA1	8.88×10^{-16}	9.86×10^{-32}	8.88×10^{-16}	0	0	0	7.17×10^{-12}	1.91×10^{-11}	4.34×10^{-15}	0	0	0
ISSA2	8.88×10^{-16}	9.86×10^{-32}	8.8×10^{-16}	0	0	0	1.54×10^{-11}	4.68×10^{-11}	9.30×10^{-16}	0	0	0
ISSA3	8.88×10^{-16}	9.86×10^{-32}	8.88×10^{-16}	0	0	0	5.23×10^{-11}	1.26×10^{-10}	1.14×10^{-15}	0	0	0
ISSA4	8.88×10^{-16}	9.86×10^{-32}	8.88×10^{-16}	0	0	0	5.26×10^{-12}	1.43×10^{-11}	3.45×10^{-16}	0	0	0
ISSA5	8.88×10^{-16}	9.86×10^{-32}	8.88×10^{-16}	0	0	0	7.27×10^{-12}	2.20×10^{-11}	4.78×10^{-15}	0	0	0
ISSA6	8.88×10^{-16}	9.86×10^{-32}	8.88×10^{-16}	0	0	0	**3.36×10^{-12}**	**7.56×10^{-12}**	6.77×10^{-16}	0	0	0
HSSA	**8.88×10^{-16}**	**9.86×10^{-32}**	**8.88×10^{-16}**	**0**	**0**	**0**	1.90×10^{-11}	2.11×10^{-11}	**1.76×10^{-16}**	**0**	**0**	**0**

表 6　HSSA 与各算法的秩和检验结果

Table 6　Results based on wilcoxon rank test for HSSA with each algorithm

算法	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}
PSO	1.2×10^{-12}	1.4×10^{-11}	1.2×10^{-12}	6.5×10^{-12}	3.0×10^{-11}	2.2×10^{-2}	3.0×10^{-11}	1.2×10^{-12}	1.2×10^{-12}	1.2×10^{-12}	1.2×10^{-2}	1.2×10^{-12}
GWO	1.2×10^{-12}	1.4×10^{-11}	1.2×10^{-12}	6.5×10^{-12}	3.0×10^{-11}	3.0×10^{-11}	1.4×10^{-7}	1.2×10^{-5}	3.8×10^{-13}	5.6×10^{-3}	3.0×10^{-11}	1.5×10^{-4}
WOA	1.2×10^{-12}	1.4×10^{-11}	1.2×10^{-12}	6.5×10^{-12}	3.0×10^{-11}	3.0×10^{-11}	1.5×10^{-5}		2.9×10^{-7}	8.2×10^{-2}	3.0×10^{-11}	3.3×10^{-1}
TLBO	1.2×10^{-12}	1.4×10^{-11}	1.2×10^{-12}	6.5×10^{-12}	3.0×10^{-11}	3.0×10^{-11}	3.0×10^{-11}	1.7×10^{-11}	4.5×10^{-13}		3.0×10^{-11}	1.2×10^{-12}
BOA	1.2×10^{-12}	1.4×10^{-11}	1.2×10^{-12}	6.5×10^{-12}	3.0×10^{-11}	3.0×10^{-11}	2.0×10^{-9}	1.2×10^{-12}	1.2×10^{-12}	1.2×10^{-12}	3.0×10^{-11}	1.2×10^{-12}
SSA		3.1×10^{-4}	1.6×10^{-1}	8.2×10^{-1}	7.7×10^{-2}	1.6×10^{-2}	5.4×10^{-1}				3.9×10^{-1}	
IHSSA[17]		1.4×10^{-11}	2.2×10^{-2}	3.5×10^{-11}	3.0×10^{-5}	7.3×10^{-3}	6.1×10^{-1}				1.3×10^{-1}	
LSSA[10]		3.1×10^{-4}		1.1×10^{-2}	1.5×10^{-1}	1.1×10^{-1}	5.9×10^{-4}				1.9×10^{-4}	
ESSA[21]	1.3×10^{-5}	4.3×10^{-6}	3.5×10^{-7}	1.4×10^{-7}	1.3×10^{-2}	1.3×10^{-10}	7.8×10^{-1}				3.8×10^{-9}	
YSSA[12]		3.1×10^{-4}		1.1×10^{-2}	1.2×10^{-12}	5.6×10^{-10}	7.7×10^{-6}				4.6×10^{-9}	

相当的情况，进一步验证了HSSA算法的独特性。

考虑到表3和表4中的测试函数数据统计仅能看出不同算法在同一测试函数上的性能表现。为了对上述12个函数的综合排名进行更直观地展示，将记录的11种算法的30次仿真实验得到的各测试函数的平均值和标准差采用Friedman检验的综合排名记录在表7中。

由表7可知，HSSA和YSSA[12]以平均排名3.60共享排名第1。和经典的基础算法相比，SSA及其变体排名明显更好，这再次证明SSA的性能好，具有不错的研究前景。而HSSA较其他各SSA变体提升明显，足以说明HSSA具有较强的综合搜索性能。

表7 Friedman统计结果

Table 7 Friedman statistical results

算法	PSO	GWO	WOA	TLBO	BOA	SSA	IHSSA[17]	LSSA[10]	ESSA[21]	YSSA[12]	HSSA
综合排名	10.00	8.79	8.54	6.63	8.63	4.19	3.81	3.94	4.27	3.60	3.60

4 结 论

1）HSSA的寻优性能提升显著。例如在和SSA变体的比较中，12个测试函数中有7个排名第1；在和经典算法的比较中，12个测试函数中的有9个都高居第1，且在数量级上有显著的提升。Friedman检验后的综合排名亦可以体现这一点。

2）HSSA各阶段从保护多样性出发，获得了较好的寻优精度。初始化阶段利用Piecewise map使得分布更加均匀；自适应警戒者、混合学习策略、边界处理在一定程度上减少多样性的丧失，较好地维持了种群的寻优能力。

3）HSSA充分运用了重要位置的信息。例如RPL中运用配对后的优胜者来更新较差位置；竞争学习运用到了当前最优；边界处理机制也体现了这一点。

4）HSSA各部分的有机组合是提升寻优性能的根本保证，全新的多策略边界处理机制使得该过程得到的种群分布更加合理。消融实验与秩和检验都证明了其有效性。

总的来说，HSSA具有较好的寻优性能，但是和最新提出的SSA变体相比，仍存在不足之处。下一步的计划是尝试发掘更优的策略以提高寻优性能，在更多的函数上测试其性能，并通过在工程优化问题、路径规划、数据分类及预测等领域进一步验证改进的可行性和实用性。

参考文献（References）

[1] EBERHART R, KENNEDY J. A new optimizer using particle swarm theory[C]//MHS'95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science. Piscataway: IEEE Press, 2002: 39-43.

[2] KENNEDY J, EBERHART R. Particle swarm optimization [C]//Proceedings of ICNN'95-International Conference on Neural Networks. Piscataway: IEEE Press, 2002: 1942-1948.

[3] DORIGO M, GAMBARDELLA L M. Ant colony system: A cooperative learning approach to the traveling salesman problem[J]. IEEE Transactions on Evolutionary Computation, 1997, 1(1): 53-66.

[4] KARABOGA D. An idea based on honey bee swarm for numerical optimization, Technical Report-TR06[R]. Kayseri: Erciyes University, 2005.

[5] MIRJALILI S, MIRJALILI S M, LEWIS A. Grey wolf optimizer[J]. Advances in Engineering Software, 2014, 69: 46-61.

[6] MIRJALILI S, LEWIS A. The whale optimization algorithm[J]. Advances in Engineering Software, 2016, 95: 51-67.

[7] XUE J, SHEN B. A novel swarm intelligence optimization approach: Sparrow search algorithm[J]. Systems Science & Control Engineering, 2020, 8(1): 22-34.

[8] 吕鑫, 慕晓冬, 张钧, 等. 混沌麻雀搜索优化算法[J]. 北京航空航天大学学报, 2021, 47(8): 1712-1720.
LYU X, MU X D, ZHANG J, et al. Chaos sparrow search optimization algorithm[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(8): 1712-1720(in Chinese).

[9] 毛清华, 张强. 融合柯西变异和反向学习的改进麻雀算法[J]. 计算机科学与探索, 2021, 15(6): 1155-1164.
MAO Q H, ZHANG Q. Improved sparrow algorithm combining Cauchy mutation and opposition-based learning[J]. Journal of Frontiers of Computer Science and Technology, 2021, 15(6): 1155-1164(in Chinese).

[10] OUYANG C T, ZHU D L, WANG F Q. A learning sparrow search algorithm[J]. Computational Intelligence and Neuroscience, 2021, 2021: 1-23.

[11] ZHOU S H, XIE H, ZHANG C C, et al. Wavefront-shaping focusing based on a modified sparrow search algorithm[J]. Optik, 2021, 244: 167516.

[12] YAN S Q, YANG P, ZHU D L, et al. Improved sparrow search algorithm based on iterative local search[J]. Computational Intelligence and Neuroscience, 2021, 2021: 1-31.

[13] JOYCE T, HERRMANN J M. A review of no free lunch theorems, and their implications for metaheuristic optimization[J]. Nature-inspired Algorithms and Applied Optimization, 2018, 744: 27-51.

[14] LI T Y, YORKE J A. The theory of chaotic attractors[M]. Berlin: Springer, 2004: 77-84.

[15] TIZHOOSH H R. Opposition-based learning: A new scheme for machine intelligence[C]//International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce. Piscataway: IEEE Press, 2006: 695-701.

[16] OUYANG C T, ZHU D L, QIU Y X. Lens learning sparrow search algorithm[J]. Mathematical Problems in Engineering, 2021, 2021: 1-17.

[17] WANG Z K, HUANG X Y, ZHU D L. A multistrategy-integrated learning sparrow search algorithm and optimization of engineering problems[J]. Computational Intelligence and Neuroscience, 2022, 2022: 1-21.

[18] DENG H B, PENG L Z, ZHANG H B, et al. Ranking-based biased learning swarm optimizer for large-scale optimization[J]. Information Sciences, 2019, 493: 120-137.

[19] BALOCHIAN S, BALOOCHIAN H. Social mimic optimization algorithm and engineering applications[J]. Expert Systems with Applications, 2019, 134: 178-191.

[20] CHENG R, JIN Y C. A competitive swarm optimizer for large scale optimization[J]. IEEE Transactions on Cybernetics, 2015, 45(2): 191-204.

[21] 王振东, 汪嘉宝, 李大海. 一种增强型麻雀搜索算法的无线传感器网络覆盖优化研究[J]. 传感技术学报, 2021, 34(6): 818-828.
WANG Z D, WANG J B, LI D H. Study on WSN optimization coverage of an enhanced sparrow search algorithm[J]. Chinese Journal of Sensors and Actuators, 2021, 34(6): 818-828(in Chinese).

Learning sparrow search algorithm of hybrids boundary processing mechanisms

WANG Zikai[1]，HUANG Xueyu[2,*]，ZHU Donglin[1]，YAN Shaoqiang[3]，LI Quan[1]，GUO Wei[1]

(1. School of Information Engineering，Jiangxi University of Science and Technology，Ganzhou 341000，China；
2. School of Software Engineering，Jiangxi University of Science and Technology，Nanchang 330013，China；
3. School of Combat Support，Rocket Military Engineering University，Xi'an 710025，China)

Abstract：A learning sparrow search algorithm called HSSA that combines a boundary processing mechanism is proposed in order to address the sparrow search algorithm's (SSA) insufficient population distribution during the initialization stage and the optimization process' insufficient interference from local optimal solutions. Using the Piecewise map initializes the population, which improves the population distribution. In order for the position updates of the following generation to be guided by the optimal solution data of each generation, the follower and vigilante are updated individually using the sorted pairing learning and competitive learning procedures. According to the optimization characteristics of different stages, a multi-strategy boundary processing mechanism is formulated. While preserving the population size, it provides a more reasonable search location for individuals beyond the boundary. After 12 simulation experiments of reference functions, the stability of HSSA in convergence speed and the efficiency of optimization are proved by means of ablation experiment and Wilcoxon rank sum test.

Keywords：sparrow search algorithm；Piecewise map；ranking paired learning；competitive learning；multi-strategy boundary processing mechanisms

Received：2022-03-28；**Accepted**：2022-04-29；**Published Online**：2022-05-10 08：41

URL：link.cnki.net/urlid/11.2625.V.20220507.2247.001

Foundation item：National Key R & D Program of China (2020YFB1713700)

* **Corresponding author**. E-mail：1137753107@qq.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0231

LS-SVM 和抗差估计的 GNSS/INS 紧组合欺骗检测算法

柯晔[1,2]，吕志伟[1,*]，周玟龙[1]，邓旭[1]，商向永[1,3]，武文博[1]

(1. 战略支援部队信息工程大学 地理空间信息学院，郑州 450001；2. 解放军 31618 部队，福州 350000；
3. 光电对抗测试评估技术重点实验室，洛阳 471000)

摘 要：针对传统欺骗检测算法对斜率较小的斜坡式欺骗检测时间过长、虚警率和漏检率偏高等问题，提出一种最小二乘支持向量机（LS-SVM）和抗差估计的全球卫星导航系统（GNSS）和惯性导航系统（INS）紧组合欺骗检测算法。所提算法通过抗差自适应调整增益矩阵，有效削弱欺骗对新息的影响，将抗差优化的训练数据集经 LS-SVM 回归得到的预测新息来代替滤波器中的欺骗新息，从而进一步提高对斜率较小的斜坡式欺骗检测处理能力。仿真结果表明：在检测欺骗值为 0.1 m/s 的斜坡式欺骗时，所提算法与传统算法相比，检测时间缩短 26.65%，虚警率降低 40.63%，定位精度提高 72.72%。所提算法具有检测快、虚警率低的优势，适用于 GNSS/INS 紧组合导航用户的斜坡式欺骗检测。

关 键 词：最小二乘支持向量机；抗差估计；GNSS/INS 紧组合；欺骗检测；斜坡式
中图分类号：U666.1; V249; TN954+.1
文献标志码：A **文章编号**：1001-5965(2024)01-0299-09

全球卫星导航系统 (global navigation satellite system, GNSS) 和惯性导航系统 (inertial navigation system, INS) 的紧组合是以 GNSS 伪距、伪距率测量值和 INS 对应的预测值作差并输入组合导航滤波进行解算，而后输出导航信息。GNSS 具有全球性、全天候、高精度、连续性和实时性的优点[1]，但其易受欺骗干扰，欺骗干扰是指干扰源生成与真实信号类似的欺骗信号或转发真实信号用以欺骗接收机，从而解算出错误的导航信息[2]；而 INS 具有自主性强、隐蔽性好、短时精度高、抗干扰能力强的优点，但其导航误差随时间积累。因此，GNSS 与 INS 进行优势互补，能够增加系统可靠性和冗余度。在 GNSS/INS 紧组合导航系统中，GNSS 欺骗影响 Kalman 滤波器解算错误的新息，导致生成错误的导航数据；同时错误的数据通过闭合环路反馈校正机制影响 INS 预测值，最终输出组合导航欺骗信息[3]。因此，对 GNSS/INS 紧组合导航系统进行实时性和准确性欺骗检测是非常必要的。

GNSS/INS 紧组合欺骗检测算法主要是基于新息向量作为检验统计量[4]，采用二元假设检验的算法。常用的算法有 χ^2 检验法[5] 和自主完好性监测外推 (autonomous integrity monitored extrapolation, AIME) 法[6]，χ^2 检验法类似于“快照法”，如：多子集分离 (multiple solution separation, MSS) 法[7]，即对当前历元的新息检验统计量进行实时检测，适合阶跃式欺骗检测；而 AIME 法类似于“连续法”，需设置合理“检测窗口”求平均新息检验统计量，适合斜坡式欺骗检测[8]。然而，欺骗的影响使得 GNSS 输入观测量有误差影响，从而影响 Kalman 滤波器输出的新息检验统计量，导致其对欺骗敏感度降低，

收稿日期：2022-04-12；**录用日期**：2022-07-25；**网络出版时间**：2022-07-27 09:41
网络出版地址：link.cnki.net/urlid/11.2625.V.20220726.1759.001
基金项目：国家自然科学基金 (42174036)；地理信息工程国家重点实验室基金 (SKLGIE2020-Z-2-1)
* **通信作者**. E-mail：lvzhiwei@sina.com

引用格式：柯晔，吕志伟，周玟龙，等. LS-SVM 和抗差估计的 GNSS/INS 紧组合欺骗检测算法 [J]. 北京航空航天大学学报，2024，50（1）：299-307.
KE Y，LYU Z W，ZHOU W L，et al. Tightly-coupled GNSS/INS spoofing detection algorithm for LS-SVM and robust estimation [J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50（1）：299-307（in Chinese）.

引起检测时间过长，虚警率和漏检率偏高等问题[9]。针对上述问题，文献 [10] 和文献 [11] 分别对新息和新息速率检验统计量进行优化，进一步提高了欺骗检测效率和检测性能。

支持向量机(support vector machine, SVM)是一种基于统计学和结构风险最小化的机器学习方法[12]。该方法适用于小样本、高维度和非线性的数学问题，具有较好的预测效果。最小二乘 SVM(least squares SVM, LS-SVM) 是 SVM 的改进算法，具有运算效率高、收敛速度快的特点，在紧组合导航系统中检测斜坡式欺骗具有广泛应用价值。刘海颖等[13]针对各种斜率的斜坡式欺骗进行仿真，分别施加欺骗值为 2, 1, 0.2, 0.1 m/s 的斜坡式欺骗，用 AIME 法检测时间分别为 22, 32, 75, 90 s，结果表明斜率越小的斜坡式欺骗，检测时间越长。钟丽娜等[14]针对组合导航斜率较小的斜坡式欺骗检测延时过长的问题，设计 LS-SVM 与 AIME 结合改进的算法，对比传统 AIME 法，对欺骗值为 0.25 m/s 的斜坡式欺骗检测时间分别为 25.4 s 和 32.3 s，检测时间缩短了 6.9 s，检测延时减小了 21.4%。文献 [9] 设计了 LS-SVM 增强的 AIME 的算法，对欺骗值为 0.1 m/s 的欺骗检测时间为 40.70 s，同时又设置了 10、50 和 100 的“检测窗口”进行对比，检测时间分别为 40.70, 37.80, 37.78 s，可以看出 50 与 100 的“检测窗口”进行对比，并没有缩短检测时间。但这 3 种算法，只进行了检测时间的对比，没有进行虚警率和漏检率的对比。张闯等[15]设计 LS-SVM 辅助的缓变检测与容错法，利用 LS-SVM 预测的新息代替原始新息自适应调整增益矩阵，提升估计精度和检测灵敏度，对欺骗值为 0.6 m/s 和 0.8 m/s 的欺骗进行仿真，相比 χ^2 检验法，其漏检率分别降低了 22% 和 16%，且其维持在 0.245。基于上述基础，本文提出一种 LS-SVM 和抗差估计的 GNSS 和 INS 紧组合欺骗检测算法，进一步研究斜率较小的斜坡式欺骗检测。

1 基于抗差估计的 AIME 紧组合欺骗检测算法

1.1 AIME 欺骗检测算法

AIME 欺骗检测算法是通过检验新息向量是否服从卡方分布来判断 GNSS/INS 紧组合系统是否受到了欺骗[16]。GNSS/INS 紧组合是以 GNSS 和 INS 的伪距和伪距率作差作为观测量 $\boldsymbol{Z}_k$ 输入组合导航滤波器，采用 Kalman 滤波器对系统状态进行预测更新，由此 k 时刻的新息向量 $\boldsymbol{r}_k$ 及其协方差矩阵 $\boldsymbol{V}_k$ 分别为

$$\boldsymbol{r}_k = \boldsymbol{Z}_k - \boldsymbol{H}_k \hat{\boldsymbol{X}}_k^- \tag{1}$$

$$\boldsymbol{V}_k = \boldsymbol{H}_k \boldsymbol{P}_k^- \boldsymbol{H}_k^{\mathrm{T}} + \boldsymbol{R}_k \tag{2}$$

式中：$\boldsymbol{H}_k$、$\hat{\boldsymbol{X}}_k^-$、$\boldsymbol{P}_k^-$ 和 $\boldsymbol{R}_k$ 分别为观测矩阵、状态估计矩阵、预测状态协方差矩阵和观测噪声协方差矩阵。基于新息及其协方差矩阵计算的检验统计量[6]为

$$S_{\mathrm{avg}}^2 = \boldsymbol{r}_{\mathrm{avg}}^{\mathrm{T}} \boldsymbol{V}_{\mathrm{avg}}^{-1} \boldsymbol{r}_{\mathrm{avg}} \tag{3}$$

式中：S_{avg}^2、$\boldsymbol{r}_{\mathrm{avg}}$ 和 $\boldsymbol{V}_{\mathrm{avg}}^{-1}$ 分别为 AIME 的检验统计量、新息的加权平均值和协方差矩阵均值的倒数，表达式为

$$\boldsymbol{r}_{\mathrm{avg}} = (\boldsymbol{V}_{\mathrm{avg}}^{-1})^{-1} \sum_{k=1}^{l} \boldsymbol{V}_k^{-1} \boldsymbol{r}_k \tag{4}$$

$$\boldsymbol{V}_{\mathrm{avg}}^{-1} = \sum_{k=1}^{l} \boldsymbol{V}_k^{-1} \tag{5}$$

式中：l 为检测窗口。当 GNSS/INS 紧组合不存在欺骗时，S_{avg}^2 服从自由度为 n 的中心化 χ^2 分布，当存在欺骗时，S_{avg}^2 服从非中心化参数为 σ 的非中心 χ^2 分布。故原假设 H_0 为 $S_{\mathrm{avg}}^2 \sim \chi^2(0,n)$，备择假设 H_1 为 $S_{\mathrm{avg}}^2 \sim \chi^2(\sigma,n)$。设虚警率为 α，检测阈值为 T_{d}，则有 $P\{S_{\mathrm{avg}}^2 \geqslant T_{\mathrm{d}}\} = \alpha$。则欺骗判断标准为

$$\begin{cases} S_{\mathrm{avg}}^2 \geqslant T_{\mathrm{d}} \\ S_{\mathrm{avg}}^2 < T_{\mathrm{d}} \end{cases} \tag{6}$$

式中：上方表示有欺骗，下方表示无欺骗。

1.2 基于抗差估计的 AIME 欺骗检测算法

AIME 欺骗检测算法对于斜率较大的斜坡式欺骗具有较好的检测效果，但对于斜率较小的斜坡式欺骗(如欺骗值为 0.1m/s)检测效果较差。在实际要求较高的应用中，对斜率较小的斜坡式欺骗进行检测时，会出现检测时间过长或难以检测，产生漏检现象；同时，闭环校正难以抑制正常通道的新息偏离正常值，导致虚警现象发生。在 GNSS/INS 紧组合导航系统中，新息可以反应系统传播和观测量的误差。然而，组合导航滤波器在预测和更新阶段中，GNSS 欺骗对新息具有一定的影响，当新息受欺骗时，称为欺骗新息，即真实新息并不等于欺骗新息[9]。文献 [9] 和 [11] 分析了 GNSS 欺骗对新息的影响，表明欺骗新息的值一般低于真实新息的值，其减少的增量为 $\boldsymbol{H}_{k+1}\boldsymbol{\Phi}_k\boldsymbol{K}_k\Delta\boldsymbol{Z}_k$，其中，$\boldsymbol{\Phi}_k$、$\boldsymbol{K}_k$ 和 $\Delta\boldsymbol{Z}_k$ 分别为转移矩阵、增益矩阵和欺骗幅度。通过滤波循环更新，导致新息的增量减少逐渐增大，从而降低了欺骗检测性能。

故通过引进抗差估计算法，可以很好的削弱欺骗对新息的影响，降低漏检率和虚警率。因此，使用 IGG-III 等价权函数[17]，定义 $W_{\mathrm{avg}}^{2,i}$ 为抗差估计中的

新息检验统计量, 服从自由度为 1 的χ^2分布, 上标i为通道数($i=1,2,\cdots,n$, n为可见卫星数, 即卫星通道数), 则构造三段函数p_i判断标准:

$$p_i=\begin{cases}1 & W_{\text{avg}}^{2,i}\leqslant T_{\text{d1}}\\ \dfrac{T_{\text{d1}}}{W_{\text{avg}}^{2,i}}\left\{\dfrac{T_{\text{d2}}-W_{\text{avg}}^{2,i}}{T_{\text{d2}}-T_{\text{d1}}}\right\}^2 & T_{\text{d1}}<W_{\text{avg}}^{2,i}<T_{\text{d2}}\\ 0 & T_{\text{d2}}\leqslant W_{\text{avg}}^{2,i}\end{cases}\tag{7}$$

式中: T_{d1}和T_{d2}是常数, 根据蒙特卡罗仿真实验, 并结合新息检验统计量的统计特性选择对应的抗差门限范围, 令$T_{\text{d2}}=3.0\sim8.0$, $T_{\text{d1}}=0.5T_{\text{d2}}$。由式(7)得出, 当$W_{\text{avg}}^{2,i}\leqslant T_{\text{d1}}$时, 表明第$i$个通道不存在欺骗, 故权重为 1; 当$T_{\text{d2}}\leqslant W_{\text{avg}}^{2,i}$时, 表明存在欺骗, 故权重为 0; 当$T_{\text{d1}}<W_{\text{avg}}^{2,i}<T_{\text{d2}}$时, 表明可能存在欺骗, 且该通道的检验量值低于正常计算值, 故权重小于 1, 降低欺骗对新息的影响, 采用自适应调整等价权的抗差估计进行处理。

定义等价权矩阵$\bar{\boldsymbol{P}}_k$, 通过调整$\bar{\boldsymbol{P}}_k$进一步对增益矩阵进行调整[18], 具体推导如下:

$$\bar{P}_k^i=R_k^i/p_i\tag{8}$$

$$\hat{\boldsymbol{K}}_k=\boldsymbol{P}_k^-\boldsymbol{H}_k^{\text{T}}(\boldsymbol{H}_k\boldsymbol{P}_k^-\boldsymbol{H}_k^{\text{T}}+\bar{\boldsymbol{P}}_k)^{-1}\tag{9}$$

式中: $\bar{P}_k^i$和R_k^i各自代表$\bar{\boldsymbol{P}}_k$和$\boldsymbol{R}_k$的第i个通道测量值。

2　基于 LS-SVM 和抗差估计的 AIME 紧组合欺骗检测算法

2.1　基于 LS-SVM 的 AIME 欺骗检测算法

LS-SVM 是 SVM 的重新表述, 与神经网络等算法相比, LS-SVM 更多应用在学习和回归算法中, 其将二次规划问题转化为求解线性方程的问题。

设训练集表示为$D=\{(x_k,y_k)|k=1,2,\cdots,N\}$, x_k, y_k和N分别为输入集, 输出集和训练样本数, 非线性回归函数[19]:

$$f(\boldsymbol{x})=\boldsymbol{w}\cdot\boldsymbol{\varphi}(x_k)+b\text{[10]}\tag{10}$$

式中: $\boldsymbol{w}$和b分别为权向量和偏差量, $\boldsymbol{\varphi}(x_k)$为特征图。由此 LS-SVM 转化为

$$\begin{cases}\min\ J(\boldsymbol{w},\boldsymbol{e})=\dfrac{1}{2}\|\boldsymbol{w}\|^2+\dfrac{1}{2}\gamma\displaystyle\sum_{k=1}^{N}e_k^2\\ \text{s.t.}\quad y_k=\boldsymbol{w}\cdot\boldsymbol{\varphi}(x_k)+b+e_k\end{cases}\tag{11}$$

式中: γ和e_k代表惩罚和松弛因子。为解决优化问题, 定义拉格朗日函数寻找最小解:

$$L(\boldsymbol{w},b,\boldsymbol{e},\alpha_{\text{L}})=J(\boldsymbol{w},\boldsymbol{e})-\sum_{k=1}^{N}\alpha_k(\boldsymbol{w}\cdot\boldsymbol{\varphi}(x_k)+b+e_k-y_k)\tag{12}$$

式中: α_{L}为拉格朗日因子。

对式(12)中的因变量$\boldsymbol{w}$、b、α_k和e_k求偏导数并取零, 可得

$$\begin{bmatrix}0 & \boldsymbol{I}^{\text{T}}\\ 1 & \boldsymbol{\Omega}+\gamma^{-1}\boldsymbol{I}\end{bmatrix}\begin{bmatrix}b\\ \alpha_{\text{L}}\end{bmatrix}=\begin{bmatrix}\boldsymbol{0}\\ \boldsymbol{y}\end{bmatrix}\tag{13}$$

式中: $\boldsymbol{\Omega}=\boldsymbol{\varphi}(x_i)^{\text{T}}\boldsymbol{\varphi}(x_j)=\boldsymbol{K}(x_i,x_j)$。故推导出回归函数为

$$f(\boldsymbol{x})=\boldsymbol{w}\cdot\boldsymbol{\varphi}(x_k)+b=\sum_{k=1}^{N}\alpha_k(\boldsymbol{\varphi}(x_k)\cdot\boldsymbol{\varphi}(\boldsymbol{x}))+b=\sum_{k=1}^{N}\alpha_kK(x_k,\boldsymbol{x})+b\tag{14}$$

因此, 在 GNSS/INS 紧组合系统有欺骗时, 回归函数用观测量$\bar{\boldsymbol{Z}}_k$作为输入, 来预测输出新息向量$\bar{\boldsymbol{r}}_k$, 故式(14)转化为

$$\bar{\boldsymbol{r}}_k=\boldsymbol{w}\cdot\boldsymbol{\varphi}(\bar{\boldsymbol{Z}}_k)+b\tag{15}$$

故, 不同于文献 [9] 和 [14] 的 LS-SVM 检测算法, 本文基于 LS-SVM 的 AIME 欺骗检测算法是在欺骗情况下提出的, 定义其检验统计量为

$$\bar{S}_{\text{avg}}^2=\bar{\boldsymbol{r}}_{\text{avg}}^{\text{T}}\bar{\boldsymbol{V}}_{\text{avg}}^{-1}\bar{\boldsymbol{r}}_{\text{avg}}\tag{16}$$

式中: $\bar{S}_{\text{avg}}^2$、$\bar{\boldsymbol{r}}_{\text{avg}}$和$\bar{\boldsymbol{V}}_{\text{avg}}^{-1}$分别为 LS-SVM 的 AIME 的检验统计量、新息的加权平均值和协方差矩阵均值的倒数, 另外:

$$\bar{\boldsymbol{r}}_{\text{avg}}=\left(\bar{\boldsymbol{V}}_{\text{avg}}^{-1}\right)^{-1}\sum_{k=1}^{l}\bar{\boldsymbol{V}}_k^{-1}\bar{\boldsymbol{r}}_k\tag{17}$$

$$\bar{\boldsymbol{V}}_{\text{avg}}^{-1}=\sum_{k=1}^{l}\bar{\boldsymbol{V}}_k^{-1}\tag{18}$$

2.2　基于 LS-SVM 和抗差估计的 AIME 欺骗检测算法

为了削弱欺骗对新息的影响, 进一步缩短斜率较小的欺骗检测时间, 降低漏检率和虚警率, 本文提出了一种基于 LS-SVM 和抗差估计的 AIME 欺骗检测算法, 引入抗差估计对欺骗新息进行初步处理, 将抗差处理的新息和受 GNSS 欺骗的观测量作为 LS-SVM 训练输入量, 训练参数采用径向基函数(radial basis function-kernel, RBF-Kernel), 其他参数通过参数优化函数获得[20], 通过 LS-SVM 回归的预测新息来代替组合导航滤波器中的欺骗新息, 使得欺骗产生的误差更多地反应在预测新息中。本文算法的具体实施流程如图 1 所示, 主要步骤如下。

步骤 1　抗差估计

在 GNSS 欺骗情况下, 计算欺骗新息和新息检验统计量, 对新息检验统计量进行抗差处理, 得到优化的新息, 作为下一步 LS-SVM 训练输入集。

1) 计算初步检验统计量。在k时刻的欺骗情况

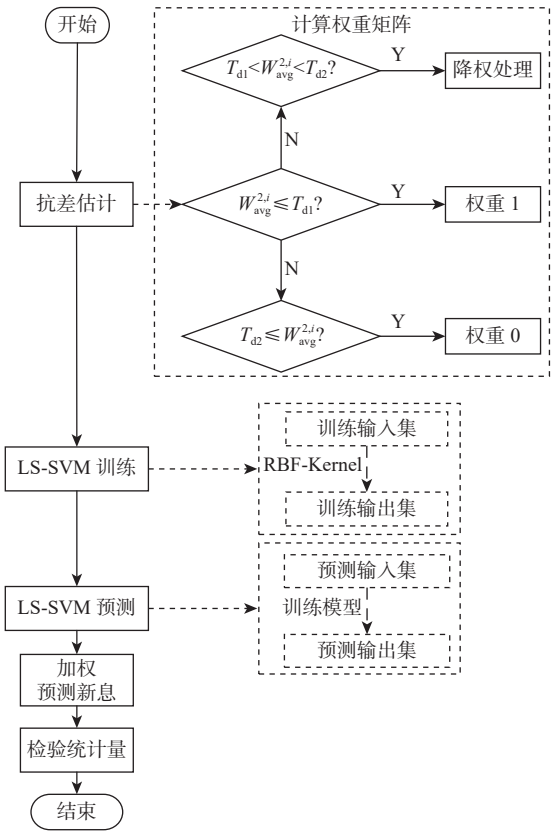


图 1 本文算法的流程

Fig. 1 Flow chart of proposed algorithm

下，计算新息检验统计量。

2）计算权重矩阵。将$W_{avg}^{2,i}$代入式 (7) 计算权重矩阵$\boldsymbol{p}$，再由式 (9) 计算新增益矩阵$\hat{\boldsymbol{K}}_k$。

3）保留训练数据集。当k时刻滤波循环结束时，保留观测量$\boldsymbol{Z}_k$和抗差优化的新息$\boldsymbol{r}_k$，然后进行$k+1$时刻循环。当整个循环结束时，保留整个 LS-SVM 训练数据集，观测量$\boldsymbol{Z}$和新息$\boldsymbol{r}$。

步骤 2 LS-SVM 训练

通过步骤 1 经抗差保留的$\boldsymbol{Z}$和$\boldsymbol{r}$作为训练集进行训练，得到训练模型。

1）训练输入。将$\boldsymbol{Z}$作为训练输入。

2）训练输出。将$\boldsymbol{r}$作为训练输出。

3）训练参数。训练参数选择 LS-SVM 的核函数是 RBF-Kernel，其他参数通过参数优化函数获得[19]。通过训练数据集和参数设置可得训练模型，为下一步回归预测作准备。

步骤 3 LS-SVM 预测

1）预测输入集。在欺骗下 GNSS 观测量受影响，则此时 GNSS 观测量和 INS 的预测观测量作差的观测量为$\bar{\boldsymbol{Z}}_k$，将其作为预测输入集。

2）预测输出集。利用步骤 2 得到的训练模型，由式 (15) 计算预测输出$\bar{\boldsymbol{r}}_k$。

步骤 4 计算加权预测新息

由式 (19) 和 (20) 计算加权预测新息$\tilde{\boldsymbol{r}}_{avg}$：

$$\tilde{\boldsymbol{r}}_{avg}=\left(\tilde{\boldsymbol{V}}_{avg}^{-1}\right)^{-1}\sum_{k=1}^{l}\bar{\boldsymbol{V}}_k^{-1}\bar{\boldsymbol{r}}_k \tag{19}$$

$$\tilde{\boldsymbol{V}}_{avg}^{-1}=\sum_{k=1}^{l}\bar{\boldsymbol{V}}_k^{-1} \tag{20}$$

式中：$\tilde{\boldsymbol{V}}_{avg}$和$\bar{\boldsymbol{V}}_k$分别为$\tilde{\boldsymbol{r}}_{avg}$和$\bar{\boldsymbol{r}}_k$的协方差矩阵。

步骤 5 计算检验统计量

基于 LS-SVM 和抗差估计的 AIME 欺骗检测算法计算归一化检验统计量$\tilde{S}_{avg}^2$为

$$\tilde{S}_{avg}^2=\tilde{\boldsymbol{r}}_{avg}^{\mathrm{T}}\tilde{\boldsymbol{V}}_{avg}^{-1}\tilde{\boldsymbol{r}}_{avg} \tag{21}$$

步骤 6 欺骗检测

$$\begin{cases}\tilde{S}_{avg}^2\geqslant T_d\\ \tilde{S}_{avg}^2<T_d\end{cases} \tag{22}$$

式中：上方表示有欺骗，下方表示无欺骗。

3 仿真实验与对比分析

为了验证本文算法（简称 M4）的有效性，引入其他 3 种算法进行对比分析，①AIME 欺骗检测算法（简称 M1）；②基于抗差估计的 AIME 的欺骗检测算法（简称 M2）；③在欺骗情况下的基于 LS-SVM 的 AIME 欺骗检测算法（简称 M3）。

针对 GNSS/INS 紧组合导航系统欺骗检测仿真实验，设计 3 个实验场景如表 1 所示。①对 1 路通道施加不同斜率的斜坡式欺骗时，对比 M1 检测情况；②对 1 路通道施加欺骗值为 0.1 m/s 的斜坡式欺骗时，对比 M1 与 M2 的检测情况；③对随机 2 路通道施加欺骗值为 0.1 m/s 斜坡式欺骗时，对比 M2、M3 和 M4 的检测情况及位置误差和均方根（root mean square error, RMSE）情况。

表 1 欺骗场景设置

Table 1 Spoofing scenario settings

实验序号	欺骗值/(m·s^{-1})	通道	持续时间/s
1	0.5,0.4,0.3,0.2,0.1	1	350～550
2	0.1	1	350～550
3	0.1	1,6	350～550

3.1 仿真条件

仿真设置 GNSS 星座模型假定圆轨道，且卫星均匀分布在 6 个轨道面上。无 GNSS 信号遮挡、衰减、干扰或反射等，GNSS 和惯性测量单元 (inertial measurement unit, IMU) 的更新频率分别为 1 Hz 和 100 Hz，IMU 参数设置如表 2 所示。虚警率P_{fa}为1×10^{-5}[21]，检测窗口l为 10，M1、M2、M3 和 M4 计算所得阈值均为 19.51。仿真轨迹使用 MATLAB 模拟飞机轨迹[22]，飞机的导航轨迹如图 2 所示，其中，实线箭头表示飞行轨迹的方向。

表 2　IMU 仿真参数设置

Table 2　IMU simulation parameter settings

加速度计随机噪声/$\left(\mathrm{mg}\cdot\sqrt{\mathrm{Hz}}^{-1}\right)$	加速度计零偏/mg	陀螺仪随机噪声/$(\sqrt{\mathrm{h}})$	陀螺仪零偏/$((°)\cdot\mathrm{h}^{-1})$
20	(30,−45,26)	0.002	(−0.000 9,0.001 3, −0.000 8)

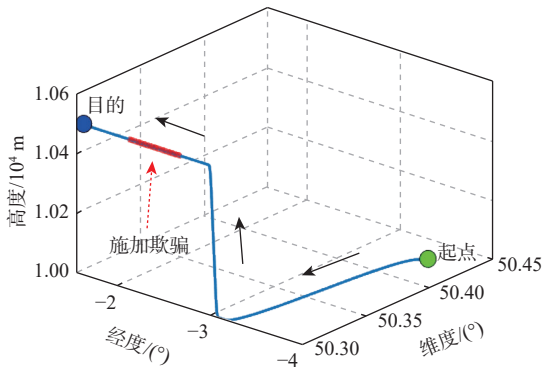


图 2　飞行轨迹

Fig. 2　Flight trajectory

3.2　仿真结果与分析

3.2.1　实验 1

通过实验 1 对通道 1 施加不同程度的斜坡式欺骗, 对比 M1 的检测能力。对通道 1 分别施加欺骗值为 0.5, 0.4, 0.3, 0.2, 0.1 m/s 的斜坡式欺骗, 欺骗持续时间为 350～550 s, M1 不同斜率的对比仿真结果如图 3 所示, 其中, "T" 为阈值, 由数据表明: 施加欺骗值为 0.5, 0.4, 0.3 m/s 的斜坡式欺骗时, 都是在时间为 370 s 时被检测出来, 但速率越大, 其检测值越大, 越容易超过阈值; 而施加欺骗值为 0.2, 0.1 m/s 的斜坡式欺骗时, 分别于 380 s 和 410 s 被检测出来, 说明斜率越小, 检测时间越长(见图中黑色圆点)。由此可得, 针对不同斜率的斜坡式欺骗, M1 检测时间和检验统计量不同。

3.2.2　实验 2

通过实验 2 对通道 1 施加欺骗值为 0.1 m/s 的斜坡式的欺骗, 对比 M1 与 M2 的检测能力, 验证抗差估计降低了正常通道的虚警率。

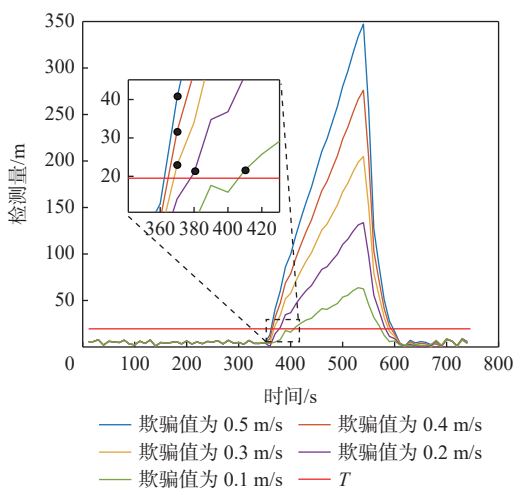
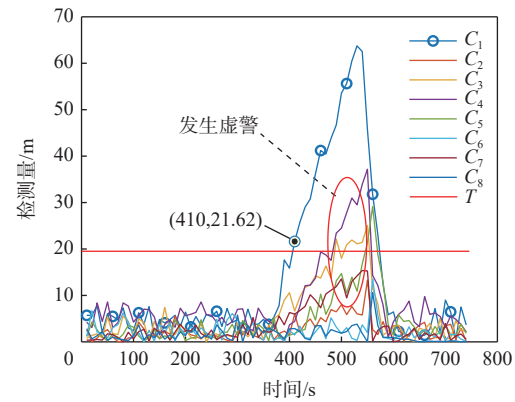


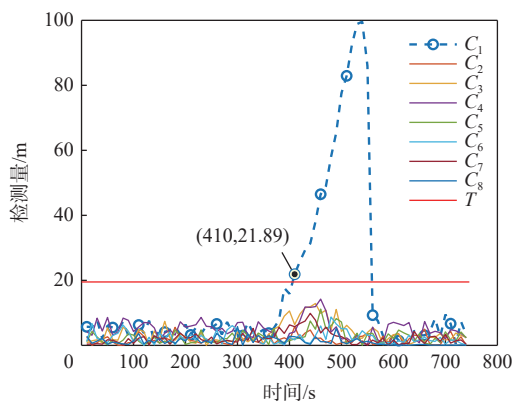
图 3　M1 不同斜率的仿真对比

Fig. 3　Simulation comparison of different slopes of M1

对通道 1 施加欺骗值为 0.1 m/s 的斜坡式欺骗, M1 与 M2 的斜坡式欺骗检测仿真结果如图 4 所示, 其中, "C" 为通道, 即 "C_1、C_2 和 C_3" 分别为通道 1、通道 2 和通道 3, 其余以此类推。图 4 中数据表明: ①图 4(a) 是 M1 不同通道的虚警仿真图, 显示了通道 1 的检测时长为 60 s, 检验统计量为 21.62 m, 而通道 3、通道 4 和通道 5 均受到了不同程序的欺骗影响, 导致其新息偏离正常值, 出现虚警情况。②图 4(b) 是 M2 不同通道的虚警仿真图, 显示了通道 1 检测时长为 60 s, 检验统计量为 21.89 m, 缩短检测时间未缩短, 但其余通道的新息值未受欺骗影响, 未出现虚警情况, 证明抗差估计有效削弱并抑制了欺骗对新息的影响。



(a) M1 仿真



(b) M2 仿真

图 4　M1 与 M2 仿真对比

Fig. 4　Simulation comparison of M1 and M2

为了进一步说明 M2 的优势, 表 3 为实验 2 的 100 次循环的蒙特卡罗仿真情况。其中, "*" 表示施加欺骗的通道。

对通道 1 施加欺骗值为 0.1m/s 的欺骗时, 结合图 4 和表 3 数据表明: ①M1 与 M2 的检测时长均为 60 s, 对比两者检测时间不变。②对于漏检率, M1 与 M2 的通道 1 均为 0。对于虚警率, M1 的通

表3 实验2蒙特卡罗仿真结果

Table 3 Monte Carlo simulation results of exp.2 %

算法	C_1*漏检率	虚警率						
		C_2	C_3	C_4	C_5	C_6	C_7	C_8
M1	0	8	82	100	50	0	31	0
M2	0	1	3	21	0	0	0	0

道2、通道3、通道4、通道5、通道6、通道7和通道8分别为8%、82%、100%、50%、0、31%和0，而M2对应分别为1%、3%、21%、0、0、0和0，M2比M1的虚警率分别降低了7%、79%、79%、50%、0、31%和0，平均降低为30.75%，最大降低为79%。由此可得，M2优于M1，且M2的抗差估计效果能够很好的削弱欺骗对新息的影响，降低了正常通道的虚警率，但M2的检测时间没有发生缩短，只是起到了抑制或削弱欺骗对新息影响的作用。

3.2.3 实验3

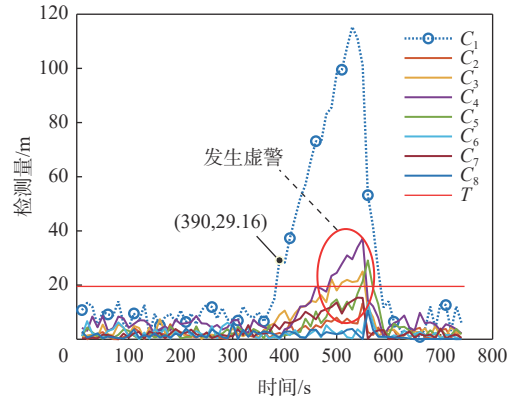
针对实验2的检测时间未缩短问题，通过实验3设置2组实验，分别对通道1和通道6施加欺骗值为0.1 m/s的斜坡式欺骗，对比M2、M3和M4的检测和位置误差情况，验证M4继承了M2的抗差估计效果且M4缩短了检测时间并提升了检测性能。

第1组实验对通道1施加欺骗值为0.1 m/s的斜坡式欺骗，M3和M4仿真不同通道虚警情况如图5所示，结合图4(b)和图5可得：①M3和M4对通道1的检测时长均为40 s，其检测量分别为29.16，28.90 m，对比M2的检测时间缩短了20 s，说明M3和M4在加入LS-SVM算法后，缩短了检测时间，提升了检测效率；②对比M3和M4的虚警情况，M3中正常通道发生虚警现象，而M4未出现虚警情况，说明M4继承了M2的抗差估计效果。

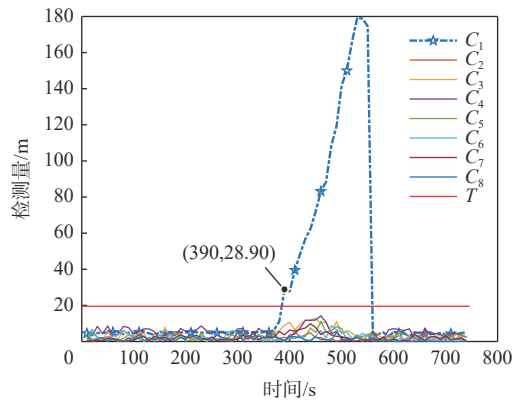
第2组实验对通道6施加欺骗值为0.1 m/s的斜坡式欺骗，对比M2、M3和M4仿真检测情况，验证随机选择的通道6，M4也继承了M2的抗差估计，且M4优于M2和M3的检测性能。由图6数据表明：M2、M3和M4的检测时长分别为50，40，40 s，说明AIME经过LS-SVM训练，得到回归的预测新息，能够缩短检测时间，但M3和M4的检验统计量分别为23.02，26.62 m，说明抗差估计对欺骗检测有一定影响。

为了进一步说明M4的优势，实验3的100次循环的蒙特卡罗仿真情况如表4所示。M2，M3和M4对通道6的位置误差和均方根对比分别如图7和表5所示。

结合图5～图7与表4和表5数据表明：①对于通道1，M4比M2检测时间缩短了33.3%，对于通道6，检测时间缩短了20%；则M4比M2平均检测时间缩短了26.65%。②对于漏检率，M2、M3和M4的通道1、通道6均为0。对于虚警率，M2、M3和M4的平均虚警率分别为3.5%、41.3%和0.67%，



(a) M3不同通道的虚警率仿真



(b) M4不同通道的虚警率仿真

图5 M3与M4的仿真对比

Fig. 5 Simulation comparison of M3 and M4

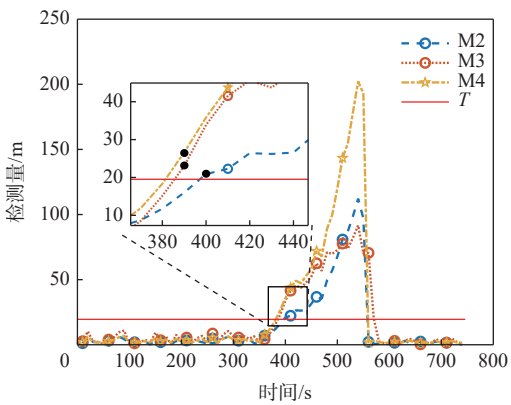


图6 M2，M3和M4对通道6的仿真对比

Fig. 6 Simulation comparison of M2, M3 and M4 on C_6

表4 实验3蒙特卡罗仿真结果

Table 4 Monte Carlo simulation results of exp.3 %

算法	漏检率		虚警率					
	C_1*	C_6*	C_2	C_3	C_4	C_5	C_7	C_8
M2	0	0	2	3	13	0	0	3
M3	0	0	7	69	100	43	29	0
M4	0	0	0	2	2	0	0	0

M4 比 M3 降低了 40.63%; 由此可得, M4 继承了 M2 的抗差估计, 降低了虚警率, 同时经 LS-SVM 训练, 得出的回归预测新息检验统计量, 从而缩短了检测时间。③图 7 为 M2, M3 和 M4 对通道 6 的位置误差仿真对比, 其中, “No spoofing” 代表无欺骗的位置误差, 在 350～550 s 加入欺骗, 3 种算法的北向、东向和地向位置误差开始出现波动, 其中 M4 与 “No spoofing” 的变化曲线较为接近, 说明定位精度较好。表 5 所示均方根对比、“No spoofing”、M2、M3 和 M4 的 3 个方向平均的均方根分别为 0.4, 1.15, 2.2, 0.66 m, 说明 M4 与 “No spoofing” 的精度较为接近, 对比 M4 和 M3, 精度提高了 72.72%。

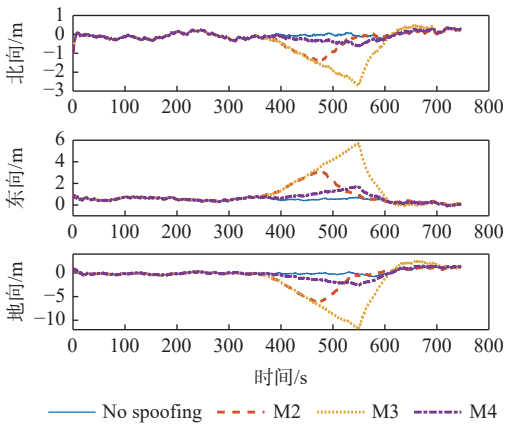


图 7　M2, M3 和 M4 对通道 6 的位置误差仿真对比

Fig. 7　Simulation comparison of position errors of M2, M3 and M4 on C_6

表 5　通道 6 的位置误差和均方根对比

Table 5　Position error and RMSE comparison of C_6

算法	误差	北向/m	东向/m	地向/m
No spoofing	最大值	0.99	0.95	1.46
	均值	0.04	0.48	0.07
	均方根	0.16	0.52	0.51
M2	最大值	1.39	3.08	6.11
	均值	0.21	0.83	0.69
	均方根	0.44	1.1	1.92
M3	最大值	2.70	5.71	11.98
	均值	0.42	1.23	1.59
	均方根	0.86	1.90	3.84
M4	最大值	0.99	1.70	2.63
	均值	0.11	0.63	0.26
	均方根	0.25	0.73	0.99

4　结　论

1) 介绍了传统欺骗检测 AIME 及 LS-SVM 的基本原理, 比较分析了传统检测算法的对检测斜率较小的斜坡式欺骗具有局限性。

2) 在 GNSS 受欺骗情况下, 通过抗差估计自适应调整增益矩阵, 从而自适应调整新息检验统计量的权重, 有效地削弱了欺骗对新息的影响。而后, 将抗差优化的新息和观测量作为 LS-SVM 训练数据, 通过 LS-SVM 回归的预测新息代替滤波器中的欺骗新息, 使得欺骗产生的误差更多地反应在预测新息中。进一步提高对斜率较小的斜坡式欺骗的检测处理能力, 有效提升检测灵敏度。

3) 在处理欺骗值为 0.1 m/s 的斜坡式欺骗时, 本文算法与传统算法相比, 检测时间缩短了 26.65%, 虚警率降低了 40.63%, 定位精度提高了 72.72%。证明本文算法具有检测时间短、检测性能好的优势, 适用于 GNSS/INS 紧组合导航系统用户的斜坡式欺骗检测。

为使本文算法能有效地针对各种场景下的车载和机载导航用户欺骗检测, 未来可开展工作: ①增加实测数据的算法实验验证; ②考虑欺骗识别与抑制技术; ③考虑多通道的欺骗检测技术。④运用方差膨胀法设置更加合理的抗差门限值。

参考文献 (References)

[1] 边少锋, 胡彦逢, 纪兵. GNSS 欺骗防护技术国内外研究现状及展望[J]. 中国科学:信息科学, 2017, 47(3): 275-287.
BIAN S F, HU Y F, JI B. Research status and prospect of GNSS anti-spoofing technology[J]. Scientia Sinica (Informationis), 2017, 47(3): 275-287(in Chinese).

[2] LIU Y, LI S H, FU Q W, et al. Analysis of Kalman filter innovation-based GNSS spoofing detection method for INS/GNSS integrated navigation system[J]. IEEE Sensors Journal, 2019, 19(13): 5167-5178.

[3] 张超, 吕志伟, 张伦东, 等. 基于新息速率抗差估计的 INS/GNSS 组合导航系统欺骗检测算法[J]. 中国惯性技术学报, 2021, 29(3): 328-333.
ZHANG C, LYU Z W, ZHANG L D, et al. A spoofing detection algorithm for INS/GNSS integrated navigation system based on innovation rate and robust estimation[J]. Journal of Chinese Inertial Technology, 2021, 29(3): 328-333(in Chinese).

[4] ZHANG L Y, ZHAO H B, SUN C, et al. Enhanced GNSS spoofing detector via multiple-epoch inertial navigation sensor prediction in a tightly-coupled system[J]. IEEE Sensors Journal, 2022, 22(9): 8633-8647.

[5] BRUMBACK B, SRINATH M. A Chi-square test for fault-detection in Kalman filters[J]. IEEE Transactions on Automatic Control, 1987, 32(6): 552-554.

[6] LIU H Y, ZHENG G, WANG H N, et al. Research on integrity monitoring for integrated GNSS/SINS system[C]//The 2010 IEEE International Conference on Information and Automation. Piscataway: IEEE Press, 2010: 1990-1995.

[7] LI X, FANG K, LI X, et al. Fault identification method of GNSS/INS integrated navigation system based on the fusion of chi-square test and multiple solution separation algorithm[C]//China Satellite Navigation Conference Proceedings. Berlin: Springer,

2021: 558-569.

[8] LIU Y, LI S H, FU Q W, et al. Impact assessment of GNSS spoofing attacks on INS/GNSS integrated navigation system[J]. Sensors, 2018, 18(5): 1433.

[9] ZHONG L N, LIU J Y, LI R B, et al. Approach for detecting soft faults in GPS/INS integrated navigation based on LS-SVM and AIME[J]. Journal of Navigation, 2017, 70(3): 561-579.

[10] 柯晔, 吕志伟, 周玟龙, 等. GNSS/INS 紧组合的新息优化抗差估计欺骗检测算法[J]. 中国惯性技术学报, 2022, 30(2): 272-280.
KE Y, LYU Z W, ZHOU W L, et al. Innovation optimal robust estimation spoofing detection algorithm of tightly coupled GNSS/INS integration[J]. Journal of Chinese Inertial Technology, 2022, 30(2): 272-280(in Chinese).

[11] KE Y, LYU Z W, ZHANG C, et al. Tightly coupled GNSS/INS integration spoofing detection algorithm based on innovation rate optimization and robust estimation[J]. IEEE Access, 2022, 23(7): 72444-72457.

[12] 戴海发, 卞鸿巍, 马恒, 等. 基于一类 SVM 的综合导航系统信息故障检测方法[J]. 中国惯性技术学报, 2017, 25(4): 555-560.
DAI H F, BIAN H W, MA H, et al. Information fault detection for integrated navigation systems using one-class support vector machine[J]. Journal of Chinese Inertial Technology, 2017, 25(4): 555-560(in Chinese).

[13] 刘海颖, 冯成涛, 王惠南. 一种惯性辅助卫星导航系统及其完好性检测方法[J]. 宇航学报, 2011, 32(4): 775-780.
LIU H Y, FENG C T, WANG H N. Method of inertial aided satellite navigation and its integrity monitoring[J]. Journal of Astronautics, 2011, 32(4): 775-780(in Chinese).

[14] 钟丽娜, 刘建业, 李荣冰, 等. 基于 LSSVM 的惯性/卫星紧组合导航系统缓变故障检测方法[J]. 中国惯性技术学报, 2012, 20(6): 754-758.
ZHONG L N, LIU J Y, LI R B, et al. Approach for detection of slowly growing errors in INS/GNSS tightly-coupled system based on LS-SVM[J]. Journal of Chinese Inertial Technology, 2012, 20(6): 754-758(in Chinese).

[15] 张闯, 赵修斌, 庞春雷, 等. LS-SVM 辅助的小幅值及缓变故障检测与容错方法[J]. 中国惯性技术学报, 2019, 27(3): 415-420.
ZHANG C, ZHAO X B, PANG C L, et al. LS-SVM assisted fault detection and tolerance method for small-amplitude fault and gradual fault[J]. Journal of Chinese Inertial Technology, 2019, 27(3): 415-420(in Chinese).

[16] WANG S Z, ZHAN X Q, ZHAI Y W, et al. Fault detection and exclusion for tightly coupled GNSS/INS system considering fault in state prediction[J]. Sensors, 2020, 20(3): 590-611.

[17] YANG Y X, HE H B, XU G C. Adaptively robust filtering for kinematic geodetic positioning[J]. Journal of Geodesy, 2001, 75(2): 109-116.

[18] ZHANG C A, ZHAO X B, PANG C L, et al. Improved fault detection method based on robust estimation and sliding window test for INS/GNSS integration[J]. Journal of Navigation, 2020, 73(4): 776-796.

[19] MEHRKANOON S, SUYKENS J A K. LS-SVM approximate solution to linear time varying descriptor systems[J]. Automatica, 2012, 48(10): 2502-2511.

[20] YING Z, KEONG K C. Fast leave-one-out evaluation and improvement on inference for LS-SVMs[C]//Proceedings of the 17th International Conference on Pattern Recognition. Piscataway: IEEE Press, 2004: 494-497.

[21] SUN R, ZHANG W Y, ZHENG J Z, et al. GNSS/INS integration with integrity monitoring for UAV No-fly zone management[J]. Remote Sensing, 2020, 12(3): 524.

[22] GROVES P D. Principles of GNSS, inertial, and multisensor integrated navigation systems[J]. IEEE Aerospace and Electronic Systems Magazine, 2015, 30(2): 26-27.

Tightly-coupled GNSS/INS spoofing detection algorithm for LS-SVM and robust estimation

KE Ye[1, 2], LYU Zhiwei[1, *], ZHOU Wenlong[1], DENG Xu[1], SHANG Xiangyong[1, 3], WU Wenbo[1]

(1. School of Geospatial Information, PLA Strategic Support Force Information Engineering University, Zhengzhou 450001, China;
2. 31618 of the PLA, Fuzhou 350000, China;
3. Key Laboratory of Electro-Optical Countermeasures Test & Evaluation Technology, Luoyang 471000, China)

Abstract: The traditional spoofing detection algorithm suffers from a prolonged time of detecting ramp spoofing with small slopes, a high false alarm rate, and a high missed detection rate. Therefore, this study proposes a spoofing detection algorithm with tightly-coupled global navigation satellite system (GNSS) and inertial navigation system (INS) based on least squares-support vector machine (LS-SVM) and robust estimation. The algorithm effectively mitigates the influence of spoofing on innovation by adaptively adjusting the gain matrix with robustness. It then replaces the spoofing innovation in the filter with the forecasted innovation obtained by LS-SVM regression of the training data set optimized with robustness, thus further improving the detection and processing ability of ramp spoofing with small slopes. Simulation results show that when detecting 0.1 m/s ramp spoofing, the proposed algorithm can shorten the detection time by 26.65%, reduce the false alarm rate by 40.63% and improve the positioning accuracy by 72.72%, compared with the traditional algorithm. The proposed algorithm has the advantages of fast detection and low false alarm rate, suitable for ramp spoofing detection of tightly integrated GNSS/INS navigation users.

Keywords: least squares-support vector machine; robust estimation; tightly-coupled GNSS/INS; spoofing detection; ramped

Received: 2022-04-12; **Accepted**: 2022-07-25; **Published Online**: 2022-07-27 09:41

URL: link.cnki.net/urlid/11.2625.V.20220726.1759.001

Foundation items: National Natural Science Foundation of China (42174036); Foundation of State Key Laboratory of Geo-Information Engineering (SKLGIE2020-Z-2-1)

* **Corresponding author**. E-mail: lvzhiwei@sina.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0232

月基装备行走运动学分析与连续步态规划

何天宇[1]，董洋[1]，檀傈锰[2]，王岩[2]，赵军鹏[1,*]，王春洁[1,3]

(1. 北京航空航天大学 机械工程及自动化学院，北京 100191；2. 北京空间飞行器总体设计部，北京 100094；
3. 北京航空航天大学 虚拟现实技术与系统国家重点实验室，北京 100191)

摘　　要：针对缓冲行走功能一体的月面探测器构型设计困难、步态分析复杂等问题，提出了一种基于相同支链构型的四足月基装备(LBE)设计分析方法与连续步态规划方案。设计复合功能着陆腿机构及整机多状态位姿；构建LBE在行走状态下的单腿运动学模型，对其进行正逆运动学求解，分析单腿与整机工作空间和运动性能；对LBE整机进行水平月面的连续步态规划，并设计单步足端轨迹，通过动力学仿真，验证LBE行走步态稳定性。仿真结果表明：LBE沿前进方向连续稳定运行，质心竖直方向浮动仅占整器高度的0.24%，俯仰角最大摆动幅度为0.34°，翻滚角最大摆动幅度为0.27°。

关　键　词：月基装备；构型设计；运动学分析；可操作度；步态规划
中图分类号：V476.3; TH122
文献标志码：A　　**文章编号**：1001-5965(2024)01-0308-09

月基装备(lunar-based equipment, LBE)作为执行复杂月面任务的多功能一体式机器人，是构建无人月球基地的基础，其基本功能包括着陆缓冲、行走搬运与姿态调整。LBE研究难点之一是需要同时考虑着陆缓冲和行走搬运两过程的工作原理，并处理好不同工作条件之间的状态切换。

在各国已发射的一系列天体探测器中，探测任务通常由着陆器和巡视器搭配完成[1]。若着陆器具备一定的移动和调姿能力，可自行前往存放区域，则简化了月球基地配置的操作内容，同时能够有效减小发射质量和包络空间，大幅降低成本[2-3]。洛克希德马丁公司主要进行着陆器的概念设计和总体方案设计，其提出了可移动月球着陆器设想[4]。美国喷气推进实验室提出了模块化可增压移动式着陆器，为长时间的深空探测任务和月球基地建立提供了依据[5]。梁鲁等[6]提出了可移动式月球着陆器概念，分析了其在增强探测灵活性和月球基地建造、运营等任务中的优势。香港大学提出了月面移动机器人概念Mobitat，着陆后机器人收拢反推力发动机，展开移动系统，进行巡视探测[7]。林荣富等[8-10]基于李群理论提出了一种新的机器人构型综合方法，列出了多种用于腿式移动着陆器的新型腿结构，并综合考虑承载能力、发射空间和能耗等因素，提出了着陆与行走功能分离的四腿式移动机器人。Han等[11-13]提出了一种可重构腿式移动着陆器，并提出了基于多状态协同优化的系统优化设计策略，该设计具有良好的模式切换和功能操作兼容性。刘志[14]和杨建中[15]等分别提出了缓冲行走一体探测器设计方案，均在着陆缓冲机构基础上增加丝杠驱动模块，实现探测器着陆后的调姿与行走。佟振鸣[16]提出了一种四足可移动式着陆探测机器人构型设计方案，并建立了运动学正反解数学模型，设计了最优步态。高峰等[17]将着陆器和巡视器融为一体，设计出可在月面反复着陆和行走的飞跃器。蔡旭[18]发明了一种四自由度可重复着陆缓冲的串并混联腿足机构。贾山等[19]提出了一种缓冲

收稿日期：2022-04-12；**录用日期**：2022-06-12；**网络出版时间**：2022-06-22 09:09
网络出版地址：link.cnki.net/urlid/11.2625.V.20220621.1028.002
基金项目：国家自然科学基金(U2037602)
*通信作者. E-mail：zhaojunpeng@buaa.edu.cn

引用格式：何天宇，董洋，檀傈锰，等. 月基装备行走运动学分析与连续步态规划[J]. 北京航空航天大学学报，2024，50(1)：308-316.
HE T Y，DONG Y，TAN L M，et al. Kinematic analysis and continuous gait planning of lunar-based equipment in walking state[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(1)：308-316 (in Chinese).

行走一体化六足着陆器, 并设计了行走与转弯步态, 推导了足端迈步轨迹与各主动驱动关节伸缩轨迹之间的对应关系。周壮壮[20]研究了行走式着陆探测器在复位机构故障状态下的行走能力。

上述研究集中于 LBE 的概念方案设计、构型综合、运动学分析等方面, 从构型设计到运动学建模分析再到步态规划的系统设计分析研究较少。在现有可移动式着陆器的构型中, 一部分过于复杂, 运动副和连杆数目较多, 一部分将驱动部件和缓冲部件设置于缓冲腿中, 增大了着陆腿的详细设计难度。对此, 本文提出一种基于相同支链构型的着陆腿并联机构设计方法, 可在对现有着陆腿较小改动下实现行走功能。在此基础上, 完成了机构行走状态下的正逆运动学建模, 求解机构的工作空间与灵巧度分布, 并对 LBE 行走过程进行步态规划和驱动映射求解, 通过虚拟样机仿真验证运动学模型的正确性, 实现 LBE 行走功能的系统设计与分析。

1　LBE 构型设计

缓冲行走一体式 LBE 由主体和 4 套缓冲行走一体式着陆腿组成。通过控制四腿的精准作动, 满足 LBE 从发射到着陆缓冲再到行走调姿的多状态功能需求。

1.1　复合功能单腿构型设计

为实现行走功能, 每套着陆腿机构需要至少 3 个自由度, 且需满足末端工作空间范围足够大、腿末端运动灵活。并联腿机构相较串联腿机构更紧凑、稳定, 精度更高, 动力学性能较好。因此, 本文选择三自由度的并联机构实现着陆腿的行走功能。

如图 1 所示, 为最大程度保留着陆缓冲功能, 缓冲行走一体的 LBE 着陆腿构型在主支柱、辅助支柱与机体连接处各增加一个驱动臂, 各驱动臂根部采用转动副驱动, 保证着陆腿末端足垫处的三自由度运动, 满足并联机构要求。LBE 单腿采用 RU-2RUS 构型, 将主支柱支链末端球铰包含在该支链内, 使 4 条支链构型相同, 便于支链运动学分析。3 条支链的转动副 R 和胡克副 U 之间为驱动臂, 胡克副 U 和球副 S 之间为缓冲腿。主驱动臂和缓冲腿对应的支链为具有三自由度的主支链, 自身有 1 个驱动; 左右辅驱动臂和缓冲腿对应 2 条副支链, 通过末端球铰与主支柱相连, 提供另外 2 个驱动。

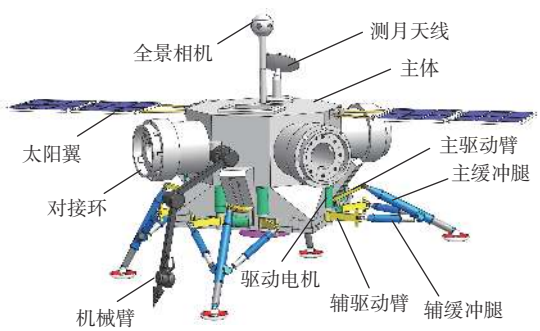


图 1　缓冲行走一体式 LBE 构型

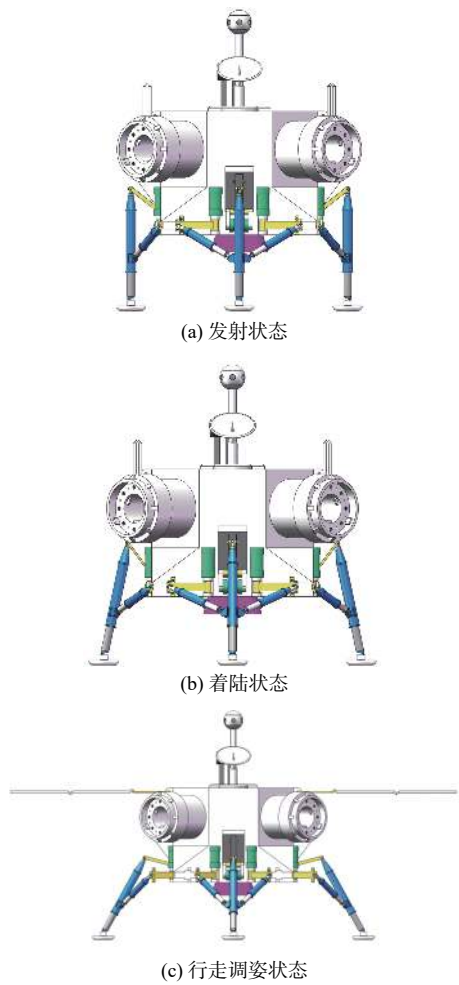
Fig. 1　Buffering and walking integrated LBE configuration

1.2　多状态位姿设计

如图 2 所示, 缓冲行走一体式 LBE 可分为 3 种工作状态。

图 2(a) 为发射状态, 各着陆腿机构的辅助支柱驱动臂通过锁定机构与机体锁定。主支柱驱动臂根部转动副调整到合适角度, 使得整体外包络圆柱直径最小。

图 2(b) 为着陆状态, 各着陆腿机构辅助支柱驱动臂仍为锁定状态, 主支柱驱动臂通过电机调整到与机体贴紧状态。缓冲过程中, 着陆腿重构为 UP-2UPS 并联机构。缓冲时, 主支柱内部缓冲元件压缩, 主支柱驱动臂将剩余冲击载荷传递给机体; 辅助支柱可能出现拉伸或压缩 2 种状态, 通过缓冲元



(a) 发射状态

(b) 着陆状态

(c) 行走调姿状态

图 2　LBE 工作状态

Fig. 2　Working status of LBE

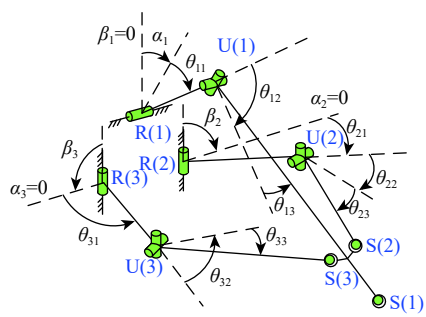
件吸能后，残余冲击载荷通过锁定机构传递给主体，避免对各驱动电机造成冲击。缓冲过程结束后，各腿支柱通过内部复位机构复原至未缓冲的原长状态。

图2(c)为行走调姿状态，各着陆腿机构辅助支柱驱动臂锁定机构解锁，各腿具有3个驱动臂转动副提供的3个自由度。通过调整驱动角度，将足垫调整到工作空间的合适位置，而后根据步态与足端轨迹规划驱动迈步，实现行走功能。实现调姿功能时，四足全部处于支撑相状态，通过协调改变4个着陆腿机构的12个驱动角度，实现机体空间6个自由度的运动。

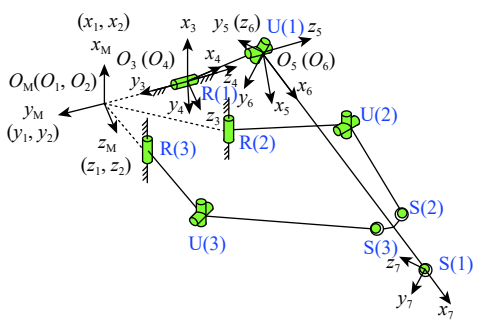
2 着陆腿机构行走运动学分析

2.1 单支链运动学模型

由于单腿的3条支链构型相同，可先构建单支链运动学模型，再通过支链间几何关系构建单腿的运动学模型。$\{O_{\mathrm{M}}\}$为LBE全局坐标系，对于任意支链$i(i=1,2,3)$，可根据支链RUS构型获得如图3(a)所示标注参数。β_i为转动副R(i)到z_{M}垂线偏离$x_{\mathrm{M}}O_{\mathrm{M}}z_{\mathrm{M}}$平面的夹角；转动副R($i$)处转角为$\alpha_i+\theta_{i1}$，其中，$\alpha_i$为支链$i$驱动臂初始状态与$x_{\mathrm{M}}O_{\mathrm{M}}y_{\mathrm{M}}$平面的夹角，$\theta_{i1}$为R($i$)处驱动转角；胡克副U($i$)处转角为$\theta_{i2}$与$\theta_{i3}$；$r_i$为转动副R($i$)到$z_{\mathrm{M}}$距离；$h_i$为转动副R($i$)到$x_{\mathrm{M}}O_{\mathrm{M}}y_{\mathrm{M}}$距离；$l_{i1}$为支链$i$驱动臂长度；$l_{i2}$为支链$i$缓冲腿长度。如图3(b)所示，以主支柱支链为例定义支链坐标系，构建支链运动学模型。



(a) 单腿参数定义



(b) 单腿主支柱支链坐标系

图3 单腿构型简图

Fig. 3 Configuration sketch of single leg

由改进DH参数法建立支链i的连杆参数，如表1所示。由连杆参数可得到支链i对应的坐标变换矩阵${}^{0}_{6}\boldsymbol{T}_i$。

表1 LBE单腿单支链改进DH参数

Table 1 Improved DH parameter of single chain of single leg for LBE

连杆k	转角ϕ_{k-1}	偏置a_{k-1}	转角φ_k	偏置d_k
1	0	0	β_i	0
2	0	r_i	0	h_i
3	π/2	0	$\alpha_i+\theta_{i1}$	0
4	0	l_{i1}	θ_{i2}	0
5	−π/2	0	θ_{i3}	0
6	0	l_{i2}	0	0

2.2 正运动学求解

根据坐标变换矩阵可以得出任意点相对于球副S(i)坐标系位置与其相对于全局O_{M}坐标系位置之间的关系，如下：

$$\begin{bmatrix}x\\y\\z\\1\end{bmatrix}={}^{0}_{6}\boldsymbol{T}_i\begin{bmatrix}{}^{\mathrm{S}(i)}x\\{}^{\mathrm{S}(i)}y\\{}^{\mathrm{S}(i)}z\\1\end{bmatrix}\tag{1}$$

可根据3条支链末端球铰位置关系列如下方程：

$$\begin{cases}\begin{bmatrix}x_{\mathrm{S}(2)}\\y_{\mathrm{S}(2)}\\z_{\mathrm{S}(2)}\\1\end{bmatrix}={}^{0}_{6}\boldsymbol{T}_1\begin{bmatrix}-(1-\mu)l_{12}\\-d_y\\d_z\\1\end{bmatrix}={}^{0}_{6}\boldsymbol{T}_2\begin{bmatrix}0\\0\\0\\1\end{bmatrix}\\\begin{bmatrix}x_{\mathrm{S}(3)}\\y_{\mathrm{S}(3)}\\z_{\mathrm{S}(3)}\\1\end{bmatrix}={}^{0}_{6}\boldsymbol{T}_1\begin{bmatrix}-(1-\mu)l_{12}\\d_y\\d_z\\1\end{bmatrix}={}^{0}_{6}\boldsymbol{T}_3\begin{bmatrix}0\\0\\0\\1\end{bmatrix}\end{cases}\tag{2}$$

式中：μ为主辅支柱连接点位置距主支柱胡克铰距离与主支柱缓冲腿长度比值；d_y和d_z分别为接头实际位置距主辅支柱连接点的侧移量和后移量。

在设计参数α_i、β_i、r_i、h_i、l_{i1}、$l_{i2}(i=1,2,3)$确定的条件下，给定驱动参数θ_{11}、θ_{21}、θ_{31}，可求解方程得到未知量θ_{12}、θ_{22}、θ_{32}、θ_{13}、θ_{23}、θ_{33}，从而正向求解出单腿末端球铰S_1的位置。

2.3 逆运动学求解

根据式(1)，可进一步得到支链i末端球副相对于球副S(i)坐标系位置与其相对于全局O_{M}坐标系位置之间的关系，如下：

$$\begin{bmatrix}x_{\mathrm{S}(i)}\\y_{\mathrm{S}(i)}\\z_{\mathrm{S}(i)}\\1\end{bmatrix}={}^{0}_{6}\boldsymbol{T}_i\begin{bmatrix}{}^{\mathrm{S}(i)}x_{\mathrm{S}(i)}\\{}^{\mathrm{S}(i)}y_{\mathrm{S}(i)}\\{}^{\mathrm{S}(i)}z_{\mathrm{S}(i)}\\1\end{bmatrix}={}^{0}_{6}\boldsymbol{T}_i\begin{bmatrix}0\\0\\0\\1\end{bmatrix}=\begin{bmatrix}t_{xi}\\t_{yi}\\t_{zi}\\1\end{bmatrix}\tag{3}$$

式中：

$$\begin{cases} t_{xi}=\cos\beta_i[l_{i2}\cos(\alpha_i+\theta_{i1})\cos\theta_{i2}\cos\theta_{i3}- \\ \quad l_{i2}\sin(\alpha_i+\theta_{i1})\sin\theta_{i2}\cos\theta_{i3}+ \\ \quad l_{i1}\cos(\alpha_i+\theta_{i1})+r_i]-l_{i2}\sin\beta_i\sin\theta_{i3} \\ t_{yi}=\sin\beta_i[l_{i2}\cos(\alpha_i+\theta_{i1})\cos\theta_{i2}\cos\theta_{i3}- \\ \quad l_{i2}\sin(\alpha_i+\theta_{i1})\sin\theta_{i2}\cos\theta_{i3}+ \\ \quad l_{i1}\cos(\alpha_i+\theta_{i1})+r_i]+l_{i2}\cos\beta_i\sin\theta_{i3} \\ t_{zi}=l_{i2}\sin(\alpha_i+\theta_{i1})\cos\theta_{i2}\cos\theta_{i3}- \\ \quad l_{i2}\cos(\alpha_i+\theta_{i1})\sin\theta_{i2}\cos\theta_{i3}+ \\ \quad l_{i1}\sin(\alpha_i+\theta_{i1})+h_i \end{cases} \tag{4}$$

求解得到

$$\begin{cases} \theta_{i1}=\arccos\left((l_{i2}^2-l_{i1}^2-x_{\mathrm{S}(i)}^2-y_{\mathrm{S}(i)}^2-z_{\mathrm{S}(i)}^2-r_i^2- \right. \\ \quad h_i^2+2(x_{\mathrm{S}(i)}\cos\beta_i+y_{\mathrm{S}(i)}\sin\beta_i)r_i+2z_{\mathrm{S}(i)}h_i)/ \\ \quad \left.\left(2l_{i1}\sqrt{(r_i-(x_{\mathrm{S}(i)}\cos\beta_i+y_{\mathrm{S}(i)}\sin\beta_i))^2+(h_i-z_{\mathrm{S}(i)})^2}\right)\right)- \\ \quad \alpha_i-\gamma_i \\ \theta_{i3}=\arcsin((y_{\mathrm{S}(i)}\cos\beta_i-x_{\mathrm{S}(i)}\sin\beta_i)/l_{i2}) \\ \theta_{i2}=\arccos(((z_{\mathrm{S}(i)}-h_i)\sin(\alpha_i+\theta_{i1})+ \\ \quad (x_{\mathrm{S}(i)}\cos\beta_i+y_{\mathrm{S}(i)}\sin\beta_i-r_i)\cos(\alpha_i+\theta_{i1})-l_{i1})/ \\ \quad (l_{i2}\cos\theta_{i3})) \end{cases} \tag{5}$$

式中:

$$\gamma_i=\arctan\left(\frac{h_i-z_{\mathrm{S}(i)}}{r_i-(x_{\mathrm{S}(i)}\cos\beta_i+y_{\mathrm{S}(i)}\sin\beta_i)}\right) \tag{6}$$

末端球副 S_1 在全局坐标系下坐标为已知条件, 则可根据式 (5) 求出主支柱支链的驱动参数 θ_{11} 和从动角度 θ_{12}、θ_{13}, 同时可得出变换矩阵 ${}^{0}_{6}\boldsymbol{T}_1$。根据式 (2) 可求出球副 S(2) 与 S(3) 在全局坐标系下的坐标。再根据式 (5) 可求出 2 个辅助支柱支链的驱动参数 θ_{21}、θ_{31} 及从动角度 θ_{22}、θ_{32}、θ_{23}、θ_{33}, 完成机构逆运动学求解。

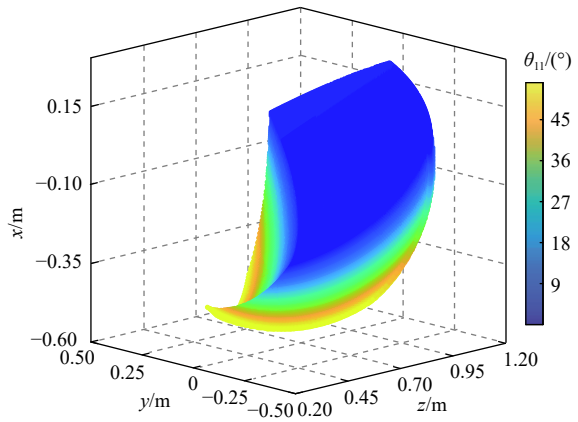
2.4　工作空间与可操作度

设定缓冲行走一体式 LBE 的 $+z$ 方向腿各支链构型参数如表 2 所示, 其他构型参数为: μ=0.6, d_y=0.05 m, d_z=0.02 m。各着陆腿机构驱动范围分别为: $\theta_{11}\in[0,53.13]°$, $\theta_{21}\in[0,100]°$, $\theta_{31}\in[0,100]°$。根据上述参数, 结合正向运动学, 得到着陆腿末端足垫球铰相对全局坐标系 $\{O_{\mathrm{M}}\}$ 的工作空间, 如图 4 所示。图 4(a) 为 $+z$ 方向单腿工作空间三维图, 图 4(b) 为整器行走工作空间三维图, 图 4(c) 为整器行走工作空间在 yOz 平面的投影。

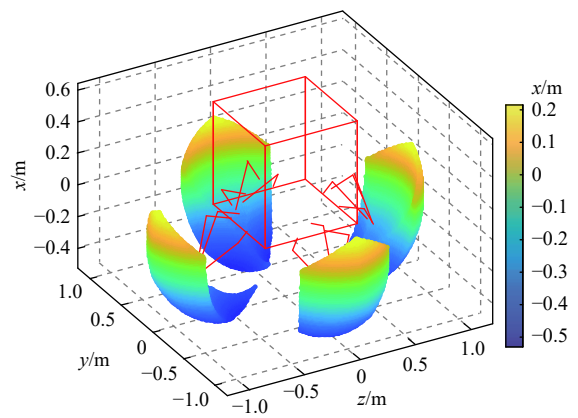
表 2　支链构型参数

Table 2　Configuration parameters of branched chain

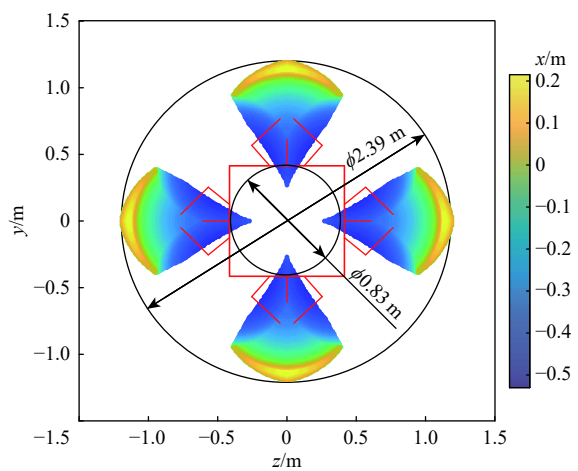
支链	α/(°)	β/(°)	r/m	h/m	l_1/m	l_2/m
主支链	36.87	0	0.07	0.415	0.2	0.6
左支链	0	−90	0.125	0.415	0.2	0.35
右支链	0	90	0.125	0.415	0.2	0.35



(a) 单腿工作空间三维图



(b) 四腿工作空间三维图



(c) 四腿工作空间 y-z 轴投影

图 4　工作空间示意图

Fig. 4　Workspace diagram

由图 4(a) 可知, 着陆腿末端工作空间各方向跨距均较大, 分布合理。其中, x 轴方向跨距为 −0.52～0.22 m, y 轴方向跨距为 −0.39～0.39 m, z 轴方向跨距为 0.26～1.20 m。x 取值 −0.4～−0.3 m 的范围内, 对应的 z 向、y 向跨距均较大, 则以着陆状态下着陆腿末端点 (−0.316,0,0.783) m 的附近位置作为起始位置, 运动范围均较大。因此, 当前构型工作空间可保证从着陆状态切换到行走初始状态过程中, 各驱动角度调整量较小, 构型合理。

当固定主支链驱动角度时，着陆腿末端工作空间为曲面。图4(a)中，工作空间各点颜色变化反映主支链驱动角的变化对着陆腿末端工作曲面的影响。可以看出，各主支链驱动角对应的末端工作曲面形状相似，驱动角越大，对z向空间利用率越大，对x向空间利用率越小。因此，选用较大主支链驱动角，可提高单步跨距，适应平坦地形；选用较小主支链驱动角，可提高单步步高，适应越障地形。

图4(b)和图4(c)中，工作空间各点颜色变化反映x坐标高度。由图4(c)可看出，工作空间从LBE主体向外辐射，距离中心越远的工作空间对应的x坐标越高，且工作空间最外圈包络圆直径约为主体跨距3倍，整器行走规划可利用范围较大。

在求解LBE工作空间的基础上，通过旋量法求解单腿机构运动雅可比矩阵，进一步分析单腿工作空间中任意点位置的运动灵巧度。

以足端球铰为原点，各坐标轴方向与全局坐标系相同建立参考系，如图5所示构建LBE单腿的旋量系。

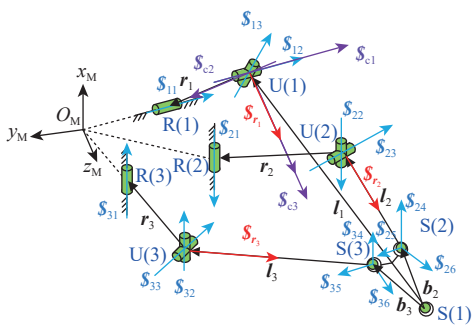


图5 单腿旋量系

Fig. 5 Single-legged spinning gauge system

图5中：$\boldsymbol{b}_2$、$\boldsymbol{b}_3$、$\boldsymbol{l}_1$、$\boldsymbol{l}_2$、$\boldsymbol{l}_3$、$\boldsymbol{r}_1$、$\boldsymbol{r}_2$、$\boldsymbol{r}_3$为支链内相邻运动副连接矢量，随单腿姿态调整而变化。$\$_{11}$、$\$_{21}$、$\$_{31}$为驱动位置的3个主动副旋量，表达式如下：

$$\begin{cases} \$_{11}=\begin{bmatrix} [0 \quad -1 \quad 0]^{\mathrm{T}} \\ (\boldsymbol{l}_1+\boldsymbol{r}_1)\times[0 \quad -1 \quad 0]^{\mathrm{T}} \end{bmatrix} \\ \$_{21}=\begin{bmatrix} [-1 \quad 0 \quad 0]^{\mathrm{T}} \\ (\boldsymbol{b}_2+\boldsymbol{l}_2+\boldsymbol{r}_2)\times[-1 \quad 0 \quad 0]^{\mathrm{T}} \end{bmatrix} \\ \$_{31}=\begin{bmatrix} [1 \quad 0 \quad 0]^{\mathrm{T}} \\ (\boldsymbol{b}_3+\boldsymbol{l}_3+\boldsymbol{r}_3)\times[1 \quad 0 \quad 0]^{\mathrm{T}} \end{bmatrix} \end{cases} \tag{7}$$

$\$_{12}$、$\$_{22}$、$\$_{32}$、$\$_{13}$、$\$_{23}$、$\$_{33}$、$\$_{24}$、$\$_{34}$、$\$_{25}$、$\$_{35}$、$\$_{26}$、$\$_{36}$为消极副对应的运动副旋量。通过互易旋量理论将三支链的消极副旋量消除，得到对应的互易旋量$\$_{r_1}$、$\$_{r_2}$、$\$_{r_3}$，表达式如下：

$$\begin{cases} \$_{r_1}=\begin{bmatrix} \dfrac{[x_{S_1}-x_{U_1} \quad 0 \quad z_{S_1}-z_{U_1}]^{\mathrm{T}}}{|[x_{S_1}-x_{U_1} \quad 0 \quad z_{S_1}-z_{U_1}]^{\mathrm{T}}|} \\ \boldsymbol{l}_1\times\dfrac{[x_{S_1}-x_{U_1} \quad 0 \quad z_{S_1}-z_{U_1}]^{\mathrm{T}}}{|[x_{S_1}-x_{U_1} \quad 0 \quad z_{S_1}-z_{U_1}]^{\mathrm{T}}|} \end{bmatrix} \\ \$_{r_2}=\begin{bmatrix} \dfrac{\boldsymbol{l}_2}{|\boldsymbol{l}_2|} \\ \boldsymbol{b}_2\times\dfrac{\boldsymbol{l}_2}{|\boldsymbol{l}_2|} \end{bmatrix} \\ \$_{r_3}=\begin{bmatrix} \dfrac{\boldsymbol{l}_3}{|\boldsymbol{l}_3|} \\ \boldsymbol{b}_3\times\dfrac{\boldsymbol{l}_3}{|\boldsymbol{l}_3|} \end{bmatrix} \end{cases} \tag{8}$$

由于支链1仅有3个自由度，等效于空间约束了另外3个自由度，可以得到支链1另外3个由等效约束提供的互易旋量$\$_{c1}$、$\$_{c2}$、$\$_{c3}$。

$$\begin{cases} \$_{c1}=\begin{bmatrix} [0 \quad 1 \quad 0]^{\mathrm{T}} \\ \boldsymbol{l}_1\times[0 \quad 1 \quad 0]^{\mathrm{T}} \end{bmatrix} \\ \$_{c2}=\begin{bmatrix} \dfrac{\boldsymbol{r}_1}{|\boldsymbol{r}_1|} \\ \boldsymbol{l}_1\times\dfrac{\boldsymbol{r}_1}{|\boldsymbol{r}_1|} \end{bmatrix} \\ \$_{c3}=\begin{bmatrix} [0 \quad 0 \quad 0]^{\mathrm{T}} \\ \dfrac{[x_{S_1}-x_{U_1} \quad 0 \quad z_{S_1}-z_{U_1}]^{\mathrm{T}}}{|[x_{S_1}-x_{U_1} \quad 0 \quad z_{S_1}-z_{U_1}]^{\mathrm{T}}|} \end{bmatrix} \end{cases} \tag{9}$$

LBE单腿并联机构仅有3个自由度，为少自由度并联机构，其雅可比矩阵由约束雅可比矩阵和驱动雅可比矩阵组成[21]。因此，由式(10)可计算出拼接雅可比矩阵逆阵，从而得到LBE单腿的速度雅可比矩阵，满足式(11)给定的足端运动速度与驱动速度关系。

$$\boldsymbol{J}^{-1}=\begin{bmatrix} \boldsymbol{J}_{\mathrm{a}} \\ \boldsymbol{J}_{\mathrm{c}} \end{bmatrix}=\begin{bmatrix} \$_{r_1}^{\mathrm{T}}\boldsymbol{\Delta}/\$_{r_1}^{\mathrm{T}}\boldsymbol{\Delta}\$_{11} \\ \$_{r_2}^{\mathrm{T}}\boldsymbol{\Delta}/\$_{r_2}^{\mathrm{T}}\boldsymbol{\Delta}\$_{21} \\ \$_{r_3}^{\mathrm{T}}\boldsymbol{\Delta}/\$_{r_3}^{\mathrm{T}}\boldsymbol{\Delta}\$_{31} \\ \$_{c1}^{\mathrm{T}}\boldsymbol{\Delta} \\ \$_{c2}^{\mathrm{T}}\boldsymbol{\Delta} \\ \$_{c3}^{\mathrm{T}}\boldsymbol{\Delta} \end{bmatrix} \tag{10}$$

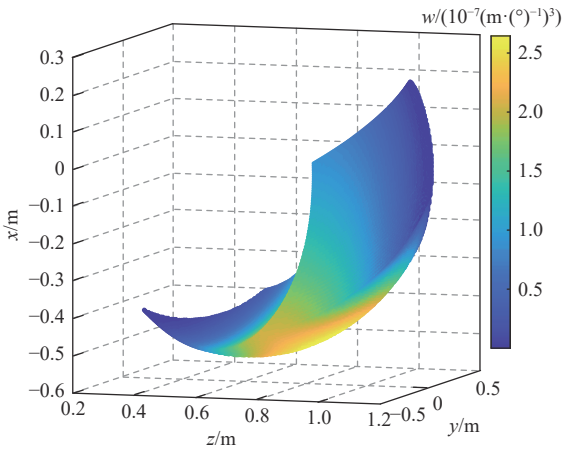
$$\begin{bmatrix} \boldsymbol{\omega}_{\mathrm{P}} \\ \boldsymbol{v}_{\mathrm{P}} \end{bmatrix}=\begin{bmatrix} \boldsymbol{J}_{11} & \boldsymbol{J}_{12} \\ \boldsymbol{J}_{21} & \boldsymbol{J}_{22} \end{bmatrix}\begin{bmatrix} \dot{\theta}_{11} & \dot{\theta}_{21} & \dot{\theta}_{31} & 0 & 0 & 0 \end{bmatrix}^{\mathrm{T}} \tag{11}$$

在LBE单腿运动学研究中，更关注足端平动速度和驱动的关系，即$\boldsymbol{J}_{21}$。因此，可根据式(12)计算出足端在工作空间中各位置的移动可操作度w，找寻运动灵巧性较好的工作空间。由结果绘制各点

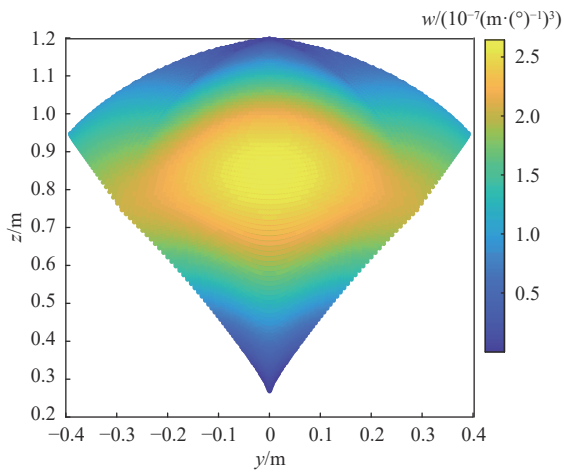
单腿可操作度, 如图 6 所示。

$$w = \sqrt{\det(\boldsymbol{J}_{21}^{\mathrm{T}}\boldsymbol{J}_{21})} = |\det(\boldsymbol{J}_{21})| \tag{12}$$

由图 6(a) 可知, x>−0.3 m 的部分和 z<0.6 m 的部分可操作度较小, 即足端位于对应位置时, 驱动变化导致的足端移动较小, 且该位置趋近于机构的奇异位置, 应避免足端落入该区域。由图 6(b) 可知, 可操作度较大部分集中在各主支链驱动角对应的末端工作曲面中心位置, 因此, 左右辅助支链的驱动搭配应避免同时位于边界的极限情况。



(a) 可操作度三维图



(b) 可操作度 y-z 投影

图 6　单腿可操作度

Fig. 6　Operability of single leg

3　LBE 连续步态规划

3.1　四足连续步态规划

缓冲行走一体式 LBE 的行走功能应保证月基装备主体在移动过程中匀速运动, 且尽量避免左右晃动与上下浮动, 从而保证月基装备上其余载荷的稳定转运。若考虑月面为理想平整地形, 可规划 LBE 的周期性连续步态, 避免机体间歇性运动, 保证良好的运动连续性及较快的行走速度, 达到平坦地形匀速平稳转运的目的。

由于 LBE 4 套着陆腿机构完全一致, 4 个方向的步态规划相同。忽略腿质量对整器质心的影响, 当 4 条腿中单腿抬起, 另外 3 足连接线组成稳定边界三角形, 定义在运动方向上质心距离稳定边界三角形最小距离为稳定裕度。为保证 LBE 整器不翻倒, 给定极限稳定裕度 s。定义单腿完成一次迈腿时, 月基装备主体相对于地面的移动距离为单腿步距 a。定义单腿完成一次迈腿时, 足端相对于地面的移动距离为单腿跨距 d。

图 7 为选定水平向左方向为前进方向的步态规划示意图。按照图 7(a) 对各迈步腿编号, 迈步顺序为 1-2-3-4。

初始状态如图 7(a) 所示; 当 LBE 向前运动时, 四腿处于支撑相且相对主体向后摆动, 使主体相对地面移动一个极限稳定裕度 s, 为 1 腿迈步作准备, 如图 7(b) 所示; 1 腿和 2 腿顺次沿前进方向迈步, 各移动一个单腿跨距 d, 同时主体相对地面移动 2 个单腿步距 a, 如图 7(c) 和 7(d) 所示; 四腿处于支撑相且相对主体向后摆动, 使主体相对地面移动 2 个极

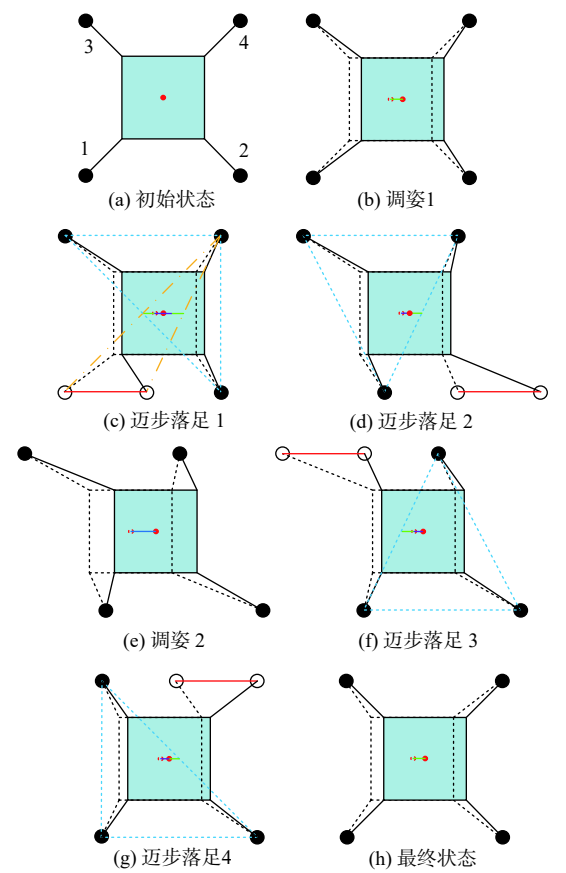


图 7　步态示意图

Fig. 7　Gait diagram

限稳定裕度 2s, 为 3 腿迈步做准备, 如图 7(e) 所示; 3 腿和 4 腿顺次沿前进方向迈步, 如图 7(f) 和 7(g) 所示; 最终四腿处于支撑相向后摆动, 使 LBE 回到初始状态, 如图 7(h) 所示, 完成向前运动的一个步态周期。

结合图 7(b)～(d), 可根据几何关系确定各步态参数之间的关系, 如下:

$$d = 4s + 4a \tag{13}$$

3.2 足端曲线规划

由于迈步顺序 1-2-3-4 与迈步顺序 3-4-1-2 的步态模式对称, 单向行进时, 每条腿有 2 种步序, 对应 2 种足端轨迹曲线。由于 LBE 可以向 4 个方向任意行进, 每条腿共有 8 种足端轨迹曲线。足端轨迹曲线由 2 部分组成, 一部分为该腿作为支撑相时足端相对主体的平动, 另一部分为该腿作为摆动相时足端从迈步到落足的整个运动过程。绘制单腿 4 方向足端相对主体的运动范围示意图, 如图 8 所示。图中, 弧线示意摆动相运动足端轨迹, 直线示意支撑相运动足端轨迹。各足端轨迹轮廓相同, 仅足端的起始点和运动方向不同, 因此可统一规划。

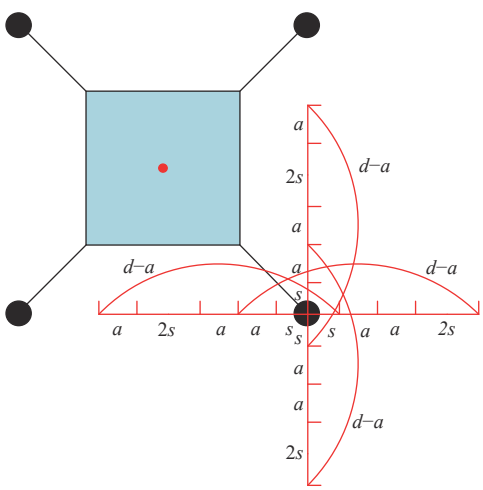
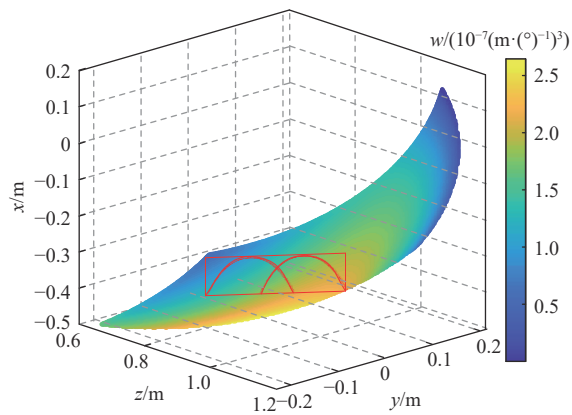


图 8 单腿四向足端相对主体运动范围

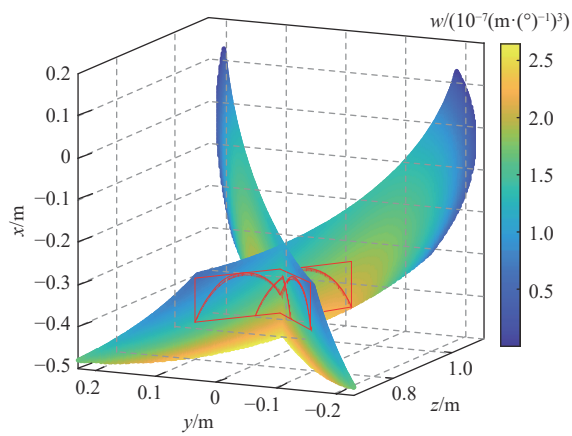
Fig. 8 Four direction motion range of foot pad relative to main body

支撑相运动足端轨迹总长度为 $4s+3a$, 足端在该段轨迹上运动速度与主体匀速运动速度相同, 该过程总时长为 $(4s+3a)/v$; 摆动相运动过程中, 主体运动一个单步跨距 a, 为保证主体整个行走过程中速度一致, 该过程总时长为 a/v。摆动相足端轨迹按照三次样条曲线插值确定[20], 为保证 LBE 行进间具有一定跨沟越障能力, 当摆动相足端运动到轨迹线在水平面投影全长的 25%、50%、75% 时, 足端纵向分别运动到步高 p 的 80%、100%、80%。

以+z 方向腿为例, 给定初始运动的起始点位置 y 坐标为 0 m, z 坐标为 0.9 m, 设定极限稳定裕度 s 为 0.011 8 m, 单步跨距 a 为 0.035 m。则可在工作空间内按四向足端运动范围搜索出足端轨迹的边界包络矩形, 计算出足端运动步高并插值得到摆动相轨迹曲线。绘制工作空间足端轨迹及边界包络矩形, 如图 9 所示。图中, 工作空间各点颜色表示该处可操作度, 包络矩形内可操作度较大, 轨迹规划较合理。已知各腿随时间变化足端轨迹后, 可根据 2.3 节逆运动学求解出主辅支柱支链的 3 个驱动角度随时间变化曲线。



(a) 单侧足端轨迹



(b) 双侧足端轨迹

图 9 工作空间中的足端轨迹

Fig. 9 Foot pad trajectory in workspace

3.3 仿真校验

针对提出的缓冲行走一体式 LBE, 在多体动力学软件中构建虚拟样机, 完成了连续步态仿真, 如图 10 所示。图 10 示为一个步态周期内调姿、各腿迈步落足过程, 与 3.1 节所述步态规划对应。由仿真结果可知, LBE 主体沿前进方向近似匀速直线运动, 行走速度约为 0.014 m/s; LBE 整体质心沿竖直方向上的位移, 在静止状态下整器质心竖直位置附近呈周期性波动, 振幅 0.002 5 m, 占主体高度的 0.24%, 该波动主要由各腿质心相对

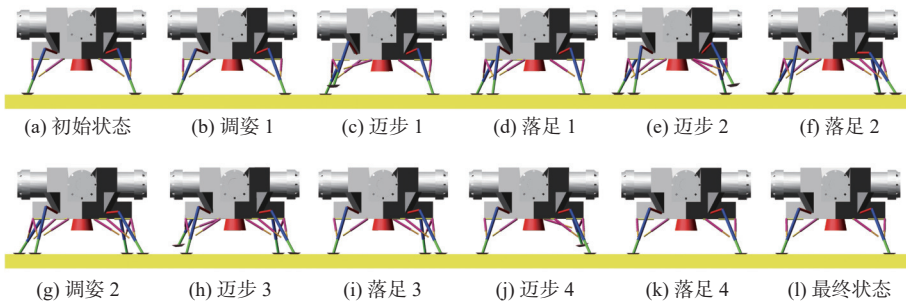


图 10　步态仿真图

Fig. 10　Gait simulation diagram

主体质心位置改变及足垫落足冲击造成, 对整器影响较小; LBE 在行走过程中, 月基装备主体的俯仰角和翻滚角均维持在0°左右, 俯仰角最大摆动幅度为0.34°, 翻滚角最大摆动幅度为0.27°。以上结果表明, LBE 行走过程平稳, 质心浮动和整器偏移均较小。

4　结　论

1) 本文以 LBE 为研究对象, 提出了一种新型的基于相同支链构型的缓冲行走一体式着陆腿机构设计方法, 并对多状态位姿进行设计。该构型采用了 RU-2RUS 三自由度并联机构, 在对现有着陆腿较小改动下实现行走功能, 实现了 4 套着陆腿机构在 LBE 发射状态、着陆状态、行走调姿状态的位姿切换与稳定工作。

2) 运用改进 DH 参数法建立了行走状态下单支链运动学模型, 在正向运动学和逆向运动学求解的基础上, 计算了单腿和整器的工作空间。着陆腿单腿各方向跨距均较大, 分布合理。整器工作空间最外圈包络圆直径约为主体跨距 3 倍, 整器行走规划可利用范围较大。

3) 运用旋量法求解了单腿机构运动雅可比矩阵, 进一步分析单腿工作空间中任意点位置的运动可操作度。当前构型坐标系下足端 $x<-0.3$ m 且 $z>0.6$ m 的部分可操作度较大, 且位于各主支链驱动角对应的末端工作曲面中心位置, 便于足端轨迹规划。

4) 设计了 LBE 在理想月面平整地形行走状态下的周期性连续步态, 在工作空间内规划了足端轨迹曲线, 并运用多体动力学软件对 LBE 行走过程进行了仿真。整器竖直方向振幅 0.002 5 m, 占主体高度的 0.24%, 俯仰角和翻滚角均维持在 0°左右, 表明缓冲行走一体式着陆腿机构运动稳定, 行走过程对整器质心影响较小, 行走状态稳定, 无较大起伏与摇摆。

本文对缓冲行走一体式 LBE 进行了系统的设计与分析, 完成了从构型设计、运动学分析、步态规划到虚拟样机验证。本文研究成果可为该类型 LBE 的设计研制与分析提供思路。

参考文献（References）

[1] 孙俊凯, 孙泽洲, 辛鹏飞, 等. 深空着陆探测足式机器人发展综述[J]. 中国机械工程, 2021, 32(15): 1765-1775.
SUN J K, SUN Z Z, XIN P F, et al. Review on development of legged robots for deep space landing exploration[J]. China Mechanical Engineering, 2021, 32(15): 1765-1775(in Chinese).

[2] 张志贤, 梁鲁, 果琳丽, 等. 轮腿式可移动载人月面着陆器概念设想[J]. 载人航天, 2016, 22(2): 202-209.
ZHANG Z X, LIANG L, GUO L L, et al. Conceptual design of manned lunar lander with wheel-legged mobile system[J]. Manned Spaceflight, 2016, 22(2): 202-209(in Chinese).

[3] 路达, 刘金国, 高海波. 星球表面着陆巡视一体化探测机器人研究进展[J]. 航空学报, 2021, 42(1): 100-116.
LU D, LIU J G, GAO H B. Integrated exploration robots for planetary surface landing and patrolling: A review[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(1): 100-116(in Chinese).

[4] BIRCKENSTAEDT B, HOPKINS J, KUTTER B, et al. Lunar lander configurations incorporating accessibility, mobility, and centaur cryogenic propulsion experience: AIAA 2006-7284[R]. Reston: AIAA, 2006.

[5] HOWE A S. A modular habitation system for human planetary and space exploration[C]//Proceedings of the 45th International Conference Environmental Systems. Emmaus: ICES, 2015.

[6] 梁鲁, 张志贤, 果琳丽, 等. 可移动式月球着陆器在载人月球探测活动中的任务分析[J]. 载人航天, 2015, 21(5): 472-478.
LIANG L, ZHAGN Z X, GUO L L, et al. Task analysis of mobile lunar lander in crewed lunar exploration missions[J]. Manned Spaceflight, 2015, 21(5): 472-478(in Chinese).

[7] HUNG A, HOWE A S. A kit-of-parts approach to pressure vessels for planetary surface construction: AIAA 2003-6281[R]. Reston: AIAA, 2003.

[8] LIN R F, GUO W Z. Novel design of a family of legged mobile lander[C]//Proceedings of the Intelligent Robotics and Applications. Berlin: Springer, 2017: 261-272.

[9] LIN R F, GUO W Z, Chen X B, et al. Type synthesis of legged mobile landers with one passive limb using the singularity property[J]. Robotica, 2018, 36(12): 1-21.

[10] LIN R F, GUO W Z. Novel design of a family of legged mobile landers based on decoupled landing and walking functions[J]. Journal of Mechanical Science and Technology, 2020, 34(9):

3815-3822.

[11] HAN Y C, GUO W Z, PENG Z K, et al. Dimensional synthesis of the reconfigurable legged mobile lander with multi-mode and complex mechanism topology[J]. Mechanism and Machine Theory, 2020, 155: 104097.

[12] HAN Y C, GUO W Z. A computable framework to efficiently design both current and robotic legged landers for extraterrestrial exploration mission[J]. Journal of Mechanical Design, 2022, 144(1): 014501.

[13] HAN Y C, GUO W Z, ZHAO D H, et al. Multi-mode unified modeling and operation capability synergistic evaluation for the reconfigurable legged mobile lander[J]. Mechanism and Machine Theory, 2022, 171: 104714.

[14] 刘志, 周杰, 靳宗向, 等. 一种可行走的月球软着陆机构: CN104943875B[P]. 2016-11-23.
LIU Z, ZHOU J, JIN Z X, et al. A walking lunar soft landing mechanism: CN104943875B[P]. 2016-11-23(in Chinese).

[15] 杨建中, 高峰, 李新立, 等. 一种具有着陆缓冲功能的行走机器人: CN105127975B[P]. 2017-04-19.
YANG J Z, GAO F, LI X L, et al. A kind of walking and buffering robot with landing function: CN105127975B[P]. 2017-04-19 (in Chinese).

[16] 佟振鸣. 移动式着陆探测机器人构型设计与行走规划研究[D]. 上海: 上海交通大学, 2016: 15-49.
TONG Z M. On configuration design and gait planning of mobile exploration lander[D]. Shanghai: Shanghai Jiaotong University, 2016: 15-49(in Chinese).

[17] 高峰, 尹科, 孙乔, 等. 探月足式飞跃机器人设计与控制[J]. 飞控与探测, 2020, 3(4): 1-7.
GAO F, YIN K, SUN Q, et al. Design and control of legged leaping robot in lunar exploration[J]. Flight Control & Detection, 2020, 3(4): 1-7(in Chinese).

[18] 蔡旭. 轮腿式可重复着陆巡视机器人的机构设计与运动规划[D]. 上海: 上海交通大学, 2020: 23-33.
CAI X. Mechanism design and motion planning of a wheel-legged repetitive landing inspection robot[D]. Shanghai: Shanghai Jiaotong University, 2020: 23-33(in Chinese).

[19] 贾山, 周向华, 陈金宝, 等. 缓冲/行走一体化着陆器运动学研究与步态规划[J]. 宇航学报, 2021, 42(4): 467-476.
JIA S, ZHOU X H, CHEN J B, et al. Kinematics research and gait planning of buffering/walking integrated lander[J]. Journal of Astronautics, 2021, 42(4): 467-476(in Chinese).

[20] 周壮壮. 行走式着陆探测器复位机构故障状态下行走能力研究[D]. 哈尔滨: 哈尔滨工业大学, 2020: 13-20.
ZHOU Z Z. Research on walking ability under the failure of reset mechanism of walking landing probe[D]. Harbin: Harbin Institute of Technology, 2020: 13-20(in Chinese).

[21] JOSHI S A, TSAI L W. Jacobian analysis of limited-DOF parallel manipulators[J]. Journal of Mechanical Design, 2002, 124(2): 254-258.

Kinematic analysis and continuous gait planning of lunar-based equipment in walking state

HE Tianyu[1], DONG Yang[1], TAN Limeng[2], WANG Yan[2], ZHAO Junpeng[1,*], WANG Chunjie[1,3]

(1. School of Mechanical Engineering and Automation, Beihang University, Beijing 100191, China;
2. Beijing Institute of Spacecraft System Engineering, Beijing 100094, China;
3. State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing 100191, China)

Abstract: A design and analysis approach and continuous gait planning scheme of quadruped lunar-based equipment (LBE) based on the same chain configuration are suggested in order to address the issues of challenging configuration design and complex gait analysis of lunar probes with buffer and walking functions. The multi-functional landing leg mechanism and the multi-state pose of the whole machine are designed. Then, the single-leg kinematics model of LBE is constructed. Its forward and inverse kinematics are solved. The workspace and motion performance of the single leg and the whole machine in a walking state are analyzed. Moreover, the continuous gait planning for the LBE on the horizontal lunar surface is carried out, and the one-step trajectory of the foot pad is designed. The motion stability of the landing leg mechanism and the walking gait stability of the LBE are confirmed through the virtual prototype model simulation. The simulation results show that the LBE operates continuously and stably along the forward direction, the vertical floating of the centroid accounts for only 0.24% of the height of the whole device, and the maximum swing amplitude of pitch angle and roll angle are 0.34° and 0.27°.

Keywords: lunar-based equipment; configuration design; kinematic analysis; manipulability; gait planning

Received: 2022-04-12; **Accepted**: 2022-06-12; **Published Online**: 2022-06-22 09:09

URL: link.cnki.net/urlid/11.2625.V.20220621.1028.002

Foundation item: National Natural Science Foundation of China (U2037602)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0233

基于领域概念图的航天新闻自动摘要模型

黄浩宁[1,2]，陈志敏[1,*]，徐聪[1,2]，张晓燕[3]

(1. 中国科学院 国家空间科学中心，北京 100190；2. 中国科学院大学，北京 100049；3. 国家无线电监测中心，北京 100037)

摘　　要：互联网海量的航天新闻中隐含着大量航天情报信息，对其进行理解与压缩是提高后续情报分析效率的基础。然而通用的自动摘要算法往往会忽略很多航天领域关键信息，且有监督自动摘要算法需要对领域文本进行大量的数据标注，费时费力。因此，提出一种基于领域概念图的无监督自动摘要(DCG-TextRank)模型，利用领域术语辅助引导图排序，提高模型对领域文本的理解力。该模型分3个模块：领域概念图生成、图权重初始化、图排序及语义筛选。根据句向量相似度和领域术语库，将文本转换为包含句子节点和领域术语节点的领域概念图；根据航天新闻文本特征初始化领域概念图权值；采用TextRank模型对句子进行排序，并在语义筛选模块通过图节点聚类及设置摘要语义保留度的方法改进TextRank的输出，充分保留文本的多语义信息并降低冗余。所提模型具有领域可移植性，且实验结果表明：在航天新闻数据集中，所提模型相比传统TextRank模型性能提升了14.97%，相比有监督抽取式文本摘要模型BertSum和MatchSum性能提升了4.37%～12.97%。

关 键 词：自动文本摘要；领域概念图；预训练语言模型；图排序算法；图节点聚类

中图分类号：TP391.1

文献标志码：A　　**文章编号**：1001-5965(2024)01-0317-11

随着航天工业的快速发展，世界各国对太空资源的争夺已从传统技术层面延伸至政治、外交、军事领域，各国航天机构纷纷开展航天情报收集与分析工作，在如此激烈的竞争下，情报分析效率显得尤为重要。互联网海量的航天新闻中隐含着大量航天情报信息，对这些非结构化文本数据进行压缩与筛选，提取出有价值的信息是情报收集的重要一环，也是后续情报分析的基础。目前针对特定领域的自动化文本信息抽取已有不少相关研究工作[1-5]，但是这些工作都是针对领域命名实体识别的，针对文本数据的摘要和筛选工作研究很少。在海量数据下，自动摘要技术通过对原文进行理解、处理与压缩，自动生成包含原文重要信息的短篇幅摘要，可以有效解决航天新闻文本信息量爆炸的问题，提高后续情报分析效率。因此，本文主要研究航天新闻的摘要抽取。

自动文本摘要方法主要分2种：抽取式摘要和生成式摘要[6]。前者通过在原文中挑选符合主旨的句子进行组合形成摘要，后者则通过对原文进行理解，重新组织语言，通过“复述”的形式简要表达原文主旨。其中生成式摘要主要采用Seq2Seq模型[7-9]，编码器学习文本的语义信息，将文本表示为一个向量，解码器则从文本向量中提取信息并加工，生成目标摘要，这种方式更接近人的思维方式。但是循环神经网络(recurrent neural network, RNN)的“长句依赖”问题也导致生成式摘要存在不连贯、不准确、不稳定和不可控的问题[10]，且对低频词不友好。而在航天情报收集与处理中，摘要的准确性直接影响后续情报分析工作，且关键信息往往集中在那些领域相关的低频词上，如航天发射目的、运载火箭、发射场地等。为保证摘要的准确性和可控性，本文采用抽取式文本摘要方法。

收稿日期：2022-04-12；**录用日期**：2022-06-25；**网络出版时间**：2022-08-26 16:50
网络出版地址：link.cnki.net/urlid/11.2625.V.20220826.1046.004
基金项目：国家自然科学基金(91738101)；国家重点研发计划(2020YFB1807900)
*通信作者. E-mail：chenzm@nssc.ac.cn

引用格式：黄浩宁，陈志敏，徐聪，等. 基于领域概念图的航天新闻自动摘要模型[J]. 北京航空航天大学学报，2024，50(1)：317-327.
HUANG H N，CHEN Z M，XU C，et al. Automatic summarization model of aerospace news based on domain concept graph[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(1)：317-327 (in Chinese).

有监督摘要抽取主要通过机器学习或神经网络方法学习句子的特征及句子与标注的映射关系[11-12]，已经取得较好的效果，但是航天领域面临着缺乏标注数据集的难点，主流的有监督学习方法不适用于航天领域。无监督摘要抽取相比来说成本低，基于具体实例进行特征学习，避免过拟合，可移植性更高[13]，更适用于缺乏标注数据的航天领域摘要抽取。

以航天发射新闻“A Soyuz rocket launched 34 small broadband satellites for OneWeb Thursday from the Baikonur Cosmodrome in Kazakhstan…”（https://spacenews.com/arianespace-launches-first-large-batch-of-oneweb-satellites-on-soyuz-rocket/）为例，高频单词“launch”总共出现18次，而发射公司“OneWeb”、发射场“the Baikonur Cosmodrome”和运载火箭“Soyuz rocket”这些包含关键情报信息的词总共出现16次。单词“launch”出现的句子大多在介绍发射任务的意义与作用，而新闻主旨为卫星发射的详情（包括发射公司、发射场、运载火箭等）。与一般文本摘要任务不同，在航天新闻摘要抽取中，领域关键术语可以用来辅助增强句子间的关联关系，对包含低频航天关键术语的句子，应给予更多的关注。而通用的摘要抽取模型往往将所有词汇同等看待，或者关注文章中的高频词汇。针对航天新闻文本信息量爆炸、通用摘要模型往往会忽略很多航天领域关键信息的问题，在缺乏领域标注文本数据集的情况下，本文提出DCG-TextRank模型，本文模型无需进行大量的文本数据标注，避免了领域文本标注的高成本问题。在图排序中引入领域术语，构建领域概念图来表征文本内部结构，使得摘要尽可能保留文本中的领域关键信息。对传统TextRank图初始化方式进行改进，在考虑句向量相似度的基础上，融合领域关键词信息和句子位置信息。通过对句子节点进行语义聚类和筛选，弥补TextRank输出摘要语义单一和冗余的缺点。构建了航天新闻自动摘要评估数据集SpaceNews。在SpaceNews数据集上，相比传统无监督自动摘要算法，本文模型在性能上具有较大的提升。本文模型为无监督学习，具有领域可移植性。

1 相关研究

无监督抽取式文本摘要通常有3个过程：句子评分、语义筛选、按序输出，常用的方法有：基于句子特征的方法[14-16]、基于主题聚类的方法[17-19]和基于图排序的方法[20-21]。基于句子特征的方法主要考虑使用句子位置、关键词词频、标题相关性等因素对句子的重要程度进行评估，将每个句子看作孤立的单元，未考虑到句子的相关性，很难学习到篇章语义信息。基于主题聚类的方法则是通过主题建模和句子聚类的方法，对句子进行语义聚类和筛选，保留文章的多语义特性。基于图排序的方法如LexRank[20]、TextRank[21]等将文本表示为句子关系图，然后采用图排序方法对句子打分，充分考虑到图的全局语义信息，是目前应用最广泛的无监督摘要抽取方式。

TextRank是一种经典的图排序文本摘要模型，将句子表示为节点，用单词共现情况初始化节点间连边权重，借鉴PageRank的思想，对句子进行排序打分，最后直接选取得分最高的几个句子作为摘要输出。在TextRank模型基础上，一些研究将基于特征的方法融入图排序中，改进其性能。叶雷等[22]将句子间的新闻要素共现信息融入到句子相关性计算中，实现汉越双语新闻自动摘要；余珊珊等[23]通过设置相关规则，将文档篇章结构融入到句子相似度计算中，对标题、句子位置、段落结构等额外信息进行定量化计算，并基于此调整句子节点间的相关性权重；Wan等[24]在构图时不仅考虑到句子与句子间的关系，还考虑到了句子与单词、单词与单词间的关系。文献[25-26]采用word2vec进行词向量的表示学习，并引入关键词覆盖率、标题相似度等信息，优化句子相关性权重；李峰等[27]通过关键词指导关键句子的摘取，提高新闻摘要质量；文献[28]采用基于Transformer的双向编码表示（bidirectional encoder representations from transformers, BERT）进行句子的表示学习，用句向量余弦相似度表示句子节点相关性；文献[25,28]对图排序得到的候选句子做了冗余处理，认为向量相似度超过一定阈值的句子间存在信息冗余。

上述方法存在一定的相似性，都只是通过融入句子位置、标题相关性、向量相似度等额外信息对TextRank中句子节点间的边权重进行调整与改进，仍存在一系列问题：①在进行候选句子筛选时，仅仅依据图排序打分值挑选候选句子，使得候选句子大多属于同一个语义类别，没有考虑到文章的多语义性，从而丢失了重要语义信息。②没有考虑到不同领域文本的差异性，不适用于特定领域的摘要抽取。③在去冗余时，都是采用设置句子相似度阈值的方法，这依赖于句子编码模型的性能。

本文将上述3种摘要抽取方法融合，在图排序的基础上，考虑句子的特征和主题，并加入领域术语和关键词辅助句子排序，使得图排序模型能够对全文产生更深入的理解。

2 DCG-TextRank自动摘要模型

2.1 模型框架

本文提出DCG-TextRank模型。如图1所示，

本文模型分 3 个模块: 领域概念图生成模块、图权重初始化模块、图排序及语义筛选模块。

在领域概念图生成模块, 根据句向量相似度和术语共现情况, 构建包含句子节点和术语节点的文本领域概念图, 以此来表征文本内部结构。在图权重初始化模块, 根据文本特点设计句子与句子间、术语与术语间、句子与术语间节点关系权重。图排序模块采用 TextRank 模型, 并通过节点聚类方法改进 TextRank 的输出, 充分保留文本语义信息并消除冗余。

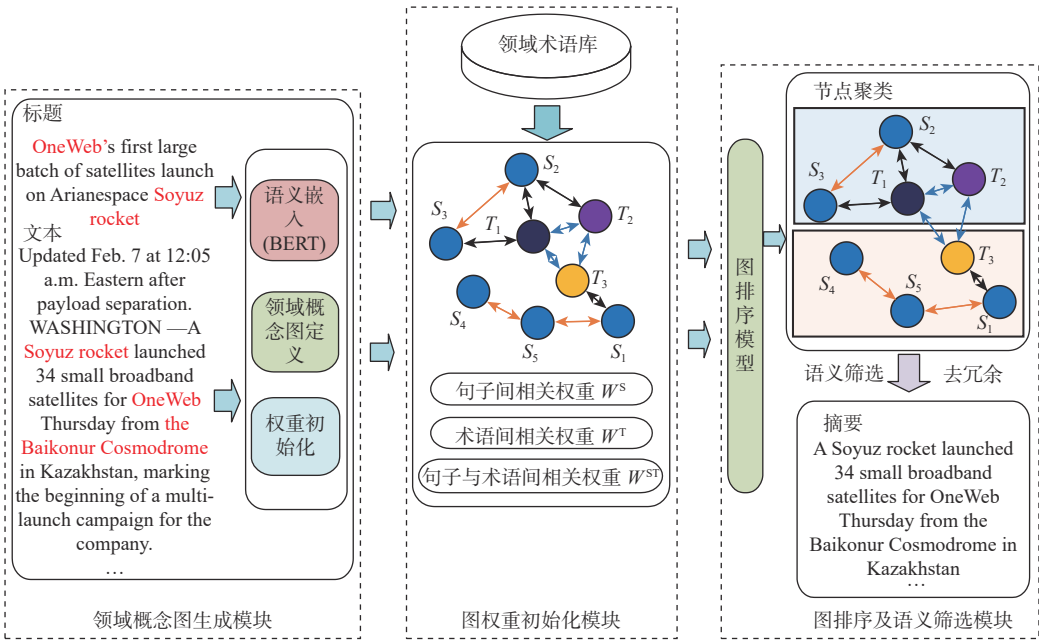


图 1　DCG-TextRank 模型框架

Fig. 1　Model framework of DCG-TextRank

2.2　领域概念图生成

2.2.1　句向量相似度计算

在传统的 TextRank 中, 同等看待每个句子, 将句子共享单词数作为句子间相关性权重。但是经研究表明, 句子间的向量相似度可以表征句子间的语义相关性[22,29]; 传统的词、句嵌入模型 word2vec、doc2vec 等在进行编码时, 不能很好地解决一词多义问题, 忽略了很多上下文信息。BERT 是 2018 年 Goole 提出的语言预训练模型[30], 通过堆叠多层双向 Transformer[31] 的 encoder 进行特征提取, 能够很好地捕捉句子中的双向关系, 获得文本中丰富的语义信息。

本文参考 BERT 预训练阶段的自监督学习思想, 在 SpaceNews 数据集上, 采用句子连续性预测方法对 BERT 进行微调, 使 BERT 学习到航天领域文本的特点。

在构建微调 BERT 所需的语料时, 对于 SpaceNews 数据集中的每个句子对正例 $I=\{s_1,s_2\}$, 随机挑选数据集中句子 s' 替换 s_2, 得到一个相应的负例 $B=\{s_1,s'\}$, 使得正负例比例为 1∶1。

BERT 微调模型如图 2 所示, 对于每个句子对, 用 [SEP] 将 2 个句子拼接, 得到 {[CLS],w_{11},w_{12},⋯, [SEP],w_{21},w_{22},⋯,[SEP]}, 作为 BERT 的输入, 将 BERT 最后一层 [CLS] 对应的输出看作整个输入的语义编码, 并添加一个全连接层, 对句子是否连续进行二分类, 采用交叉熵损失函数:

$$\text{Loss}=-\sum_i (y_i \ln L_{i+}+(1-y_i)\ln L_{i-}) \tag{1}$$

式中: y_i 为第 i 个样本的标签, 正例为 1, 负例为 0; L_{i+} 为第 i 个样本预测为正例的概率; L_{i-} 为第 i 个样本预测为负例的概率。

得到 BERT 编码的句子向量后, 计算句子间的初始相关性矩阵 $\boldsymbol{A}$, 如式 (2) 所示, 采用余弦相似

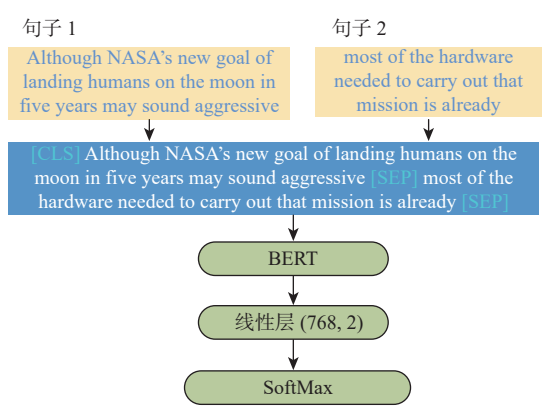


图 2　BERT 微调模型

Fig. 2　BERT fine-tuning model

度，$\boldsymbol{e}=[\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n]$为采用BERT进行编码的句子向量，$n$为句子个数。

$$A_{i,j}=\max\left(0,\cos\left(\boldsymbol{e}_i,\boldsymbol{e}_j\right)\right) \tag{2}$$

2.2.2 构建领域概念图

领域术语是指包含本领域关键信息的词汇集合。本文以航天领域为例，构建领域术语库：首先通过ITU数据库和网络爬虫获取部分领域术语，并采用命名实体识别和关键词抽取任务对其进行扩充，最终得到航天领域术语库，具体如表1所示。

表1 航天术语库

Table 1 Glossary of aerospace terms

术语类型	个数	示例
政府空间机构	78	UK Space Agency
组织机构	29	Defense Department
私营航天公司	122	SpaceX
火箭/卫星/卫星网络	5 933	StarLink
航天中心/发射场	70	Andoya Spaceport
其他关键词	450	satellite

为了对出现领域专业术语的句子给予更高的注意力，在传统文本图中加入术语节点，基于领域术语库构建文本领域概念图，其定义为

$$G=\{S,T,R_{\mathrm{SS}},R_{\mathrm{ST}},R_{\mathrm{TT}}\} \tag{3}$$

式中：G为领域概念图；S和T分别为句子节点集合和术语节点集合；图中包含3种节点间关系：句子与句子间关系R_{SS}、句子与术语间关系R_{ST}、术语与术语间关系R_{TT}。

领域概念图构建流程如算法1所示，对于一个输入文本，首先通过关键词检测抽取出文本中出现的领域术语集合Ω，并初始化领域概念图中节点：将Ω中的每个术语当作一个节点，对于第i个术语，记作T_i，同样将文本中所有句子看作一个节点，第j个句子记作S_j。接着初始化图中连边：①认为出现在标题中的术语往往是文章中最关键的术语，且这些术语往往存在较大的关联关系，为此，连接所有同时出现在标题中的术语节点；②将所有句子节点与其中出现的术语节点进行连接，这样同时出现在一个句子中的术语将会通过句子节点进行间接语义关联，术语共现次数越多，相关性也越大；同时，若2个句子共享术语个数越多，2个句子间相关性也越大；③对于所有句子对，若句向量相似度大于0，则说明句子间有较大相关性，建立连接。

算法1 领域概念图构建算法

输入 领域术语库D，文本标题H，文本内容$c=\{s_1,s_2,\cdots,s_n\}$，n为句子个数。

输出 文本领域概念图G

1: 根据术语库D检测文本中出现的术语集合：$t=\{t_1,t_2,\cdots,t_m\}$

2: 初始化G中节点：T_i为文本中编号为i的术语，S_j为第j个句子

3: if 标题H同时包含t_j和t_k then:

4: link $T_j\leftrightarrow T_k$

5: for s_i in c do

6: for t_j in t do

7: if s_i 包含 t_j then:

8: link $S_i\leftrightarrow T_j$

9: for $i\leftarrow 1$ to n do

10: for $j\leftarrow 1$ to n do

11: if $A_{i,j}>0$ then

12: link $S_i\leftrightarrow S_j$

13: return G

2.3 图权重初始化

得到领域概念图G后，对其进行边权重的初始化，得到最终的文本领域概念图邻接矩阵$\boldsymbol{G}'$。

2.3.1 句子间权重初始化

影响句子重要程度的因素有很多，除了句子间语义相似度外，在书写规范的长文本中，句子的位置也会影响句子的重要程度[32]，比如在新闻文本中，文章的开头往往起到文章内容的概括作用，包含文章的大部分信息，这些句子应该予以更多的重视；除此之外，对于航天领域文本，包含领域关键词的句子往往应给予更高程度的重视。根据上述2点，结合排序模型采用边权重对节点重要程度进行分流传播的特点，在句子余弦相似度基础上，融入关键词覆盖度和位置信息加权，调整句子相关性权值。

首先计算句子的额外信息加权值，如式(4)～式(6)所示，其中i为句子在文中的绝对位置，$\boldsymbol{w}^1$为关键词覆盖度向量，$w_i^1\in[0,1]$，$K(S_i)$为句子S_i包含关键词个数；$\boldsymbol{w}^2$为句子位置信息加权向量，$w_i^2\in[0,1]$，n为文章中句子总数，句子越靠前则位置信息加权w_i^2越大。$\boldsymbol{w}^1$和$\boldsymbol{w}^2$共同构成句子额外信息加权向量$\boldsymbol{w}$。

$$w_i^1=\frac{K(S_i)}{\max(K(S_i))} \tag{4}$$

$$w_i^2=1-\left(\frac{i}{n}\right)^2 \tag{5}$$

$$\boldsymbol{w}=\boldsymbol{w}^1+\boldsymbol{w}^2 \tag{6}$$

最后用$\boldsymbol{w}$对句子节点间权重进行调整，得到最终的句子相关度矩阵$\boldsymbol{W}^1$：

$$W_{i,j}^1=\begin{cases}A_{i,j}+w_j & A_{i,j}>0\\ 0 & A_{i,j}\leqslant 0\end{cases} \tag{7}$$

2.3.2　句子与术语间权重初始化

定义句子与术语间关系权重为: 术语在句子中出现的比例。如式 (8) 所示, $F_{i,j}$ 为术语 T_j 出现在句子 S_i 中的次数, $\boldsymbol{W}^2$ 为句子与术语间的相关性矩阵, $W^2_{i,j}$ 表示句子 S_i 与术语 T_j 间边的权重。

$$W^2_{i,j}=\begin{cases}\dfrac{F_{i,j}}{\sum\limits_{T_k\in S_i}F_{i,k}} & T_j\in S_i\\ 0 & \text{其他}\end{cases} \tag{8}$$

2.3.3　术语间权重初始化

根据式 (8), 部分术语已经通过相关句子进行间接关联, 本文定义术语间直接相关度权重为术语在标题中共现情况, 如式 (9) 所示, 若术语 T_i 和 T_j 同时出现在标题中, 则认为其直接相关, 其中 $F_{H,j}$ 为术语 T_j 在标题中出现次数, $\boldsymbol{W}^4$ 为术语相关性矩阵。

$$W^4_{i,j}=\begin{cases}\dfrac{F_{H,j}}{\sum\limits_{T_k\in H}F_{H,k}} & T_i,T_j\in H\\ 0 & \text{其他}\end{cases} \tag{9}$$

最终得到领域概念图邻接矩阵 $\boldsymbol{G}'$, 如式 (10) 所示, $\boldsymbol{G}'$ 由 4 个分块矩阵 $\boldsymbol{W}^1$、$\boldsymbol{W}^2$、$\boldsymbol{W}^3$、$\boldsymbol{W}^4$ 组成, 其中 $\boldsymbol{W}^2$ 为 $\boldsymbol{W}^3$ 的转置, n 为文章句子数, m 为文章中出现的术语个数。

$$\boldsymbol{G}'=\begin{bmatrix}\begin{bmatrix}0 & \cdots & W^1_{1,n}\\ \vdots & & \vdots\\ W^1_{n,1} & \cdots & 0\end{bmatrix} & \begin{bmatrix}W^2_{1,1} & \cdots & W^2_{1,m}\\ \vdots & & \vdots\\ W^2_{n,1} & \cdots & W^2_{n,m}\end{bmatrix}\\ \begin{bmatrix}W^3_{1,1} & \cdots & W^3_{1,n}\\ \vdots & & \vdots\\ W^3_{m,1} & \cdots & W^3_{m,n}\end{bmatrix} & \begin{bmatrix}0 & \cdots & W^4_{1,m}\\ \vdots & & \vdots\\ W^4_{m,1} & \cdots & 0\end{bmatrix}\end{bmatrix} \tag{10}$$

2.4　图排序及语义筛选

2.4.1　图排序

在对领域概念图进行初始化后, 采用 TextRank 模型对图中句子和术语进行排序, 得到各个节点的最终重要性得分, 计算式为

$$M(v_i)=1-d+d\sum_{v_j\in I(v_i)}\frac{G'_{ji}}{\sum\limits_{v_k\in O(v_j)}G'_{jk}}M(v_j) \tag{11}$$

式中: d 为阻尼因子, 设置为 0.85; v_i 为图中节点, 经过多轮迭代收敛后, $M(v_i)$ 即为节点 v_i 的重要性得分; $I(v_i)$ 为与节点 v_i 相连的所有入边节点; $O(v_j)$ 为与节点 v_j 相连的所有出边节点。

2.4.2　语义筛选

根据 TextRank 排序模型的特点, 得分排名 Top-k 的句子虽然涵盖了文章的大部分语义信息, 但是这些句子往往语义比较相似, 这使得直接选取得分 Top-k 的句子会造成摘要信息的冗余及原文信息量的损失。为充分保留文本信息并消除冗余, 排序结束后, 按照句子语义相关性对句子节点进行聚类和筛选。

1) 图节点聚类划分。图节点聚类的算法有很多, 比如 k-means、Louvain[33] 算法等。其中 Louvain 算法无需提前定义类别个数, 就能得到层次完整的不同分辨率下的图节点聚类。考虑到文本涵盖主题个数不确定的特点, 本文采用 Louvain 算法进行图节点的聚类。

2) 确定摘要长度上限。对于每篇文章, 输出摘要长度上限 $N_{\max}$ 如式 (12) 所示, 其中 N 为文章单词个数, r 为原文摘取比例, 是可调节参数。

$$N_{\max}=rN \tag{12}$$

3) 确定类别内部摘取文本长度。经过节点聚类后, 每个类别内部的句子语义相近, 类别个数即为文章语义类别总数。如式 (13) 所示, C_i 为第 i 个类别, 将类别总得分看作其代表语义信息的重要程度, 那么每个类别内部摘取文本长度 N_i 应与类别总得分成正比。

$$N_i=N_{\max}\frac{\sum\limits_{v_j\in C_i}M(v_j)}{\sum\limits_{v_k\in G}M(v_k)} \tag{13}$$

4) 过滤无关信息。文本中往往包含一些与文本主旨无关的句子, 例如“以上就是今日报道”等。对每个语义类别都输出其代表句子, 虽然能够最大程度上保留原文中所有语义信息, 但是也可能保留与主旨无关的句子。为此, 设置输出句子得分上限值, 降低输出与主旨无关句子的概率, 如式 (14) 所示, 若所选句子总得分达到 P 值, 停止对句子的摘取, 其中 γ 为原文语义保留程度, $\gamma\in(0,1]$。

$$P=\gamma\sum_i M(S_i) \tag{14}$$

5) 系统摘要输出。最后对所有摘取句子按照在原文中位置排序输出。

3　实验结果与分析

3.1　数据集介绍

由于目前还没有公开的航天领域文本摘要数据集, 本文从 spacenews 网站 (https://spacenews.com/section/launch-section/) 通过网络爬虫获取了 2014 年至今的航天发射新闻, 并对文本进行清洗、标注, 构建航天文本摘要数据集 SpaceNews, 共包含 1 000 篇航天发射新闻, 全部为英文, 句子总数为 31 161, 句子平均长度为 19.81 个单词, 摘要平均长度为 27.21 个单词。

3.2 评价方法

评价摘要的优劣不仅要考虑信息的覆盖率，还应考虑系统摘要所反馈信息的准确率，若系统摘要包含了参考摘要中的大部分信息，还包含了更多的无关信息，那么也不能称之为好的摘要。ROUGE是文本自动摘要任务中最常用的一种评价指标，根据系统摘要与参考摘要间 n 元组（n-gram）共现信息来评价生成摘要的质量，本文将ROUGE_1、ROUGE_2、ROUGE_L作为评价指标，同时综合考虑 n-gram 的召回率和准确率，采用F1-Measure对实验结果进行评价：

$$R_n=\frac{\sum_{s\in R}\sum_{b\in s}Q_{\text{match}}(b)}{\sum_{s\in R}\sum_{b\in s}Q(b)} \tag{15}$$

$$P_n=\frac{\sum_{s\in R}\sum_{b\in s}Q_{\text{match}}(b)}{\sum_{s\in A}\sum_{b\in s}Q(b)} \tag{16}$$

$$F_n=\frac{2R_nP_n}{R_n+P_n} \tag{17}$$

$$F_{\text{average}}=\frac{F_1+F_2+F_{\text{L}}}{3} \tag{18}$$

式中：P_n、R_n和F_n分别为系统摘要中 n 元组准确率、召回率和F1-Measure；b 表示 n 元组（n-gram）；F_{average} 为3种 n-gram 的F1-Measure均值；F_1 为1元组F1-Measure；F_2 为2元组F1-Measure；F_{L} 为最长公共子序列F1-Measure；R 为参考摘要；A 为模型产生的系统摘要；$Q_{\text{match}}(b)$为 A 与 R 匹配到的 n-gram 数量。

3.3 实验设计

实验环境为windows 10，64位，硬件环境为：NVIDIA Quadro RTX 3000 GPU，Intel(R) Core(TM) i7-6700 CPU @ 3.40 GHz 3.41 GHz，软件环境为：Pycharm，Python 3.7，使用到的第三方库有：torch、transformers、networkx、gensim、community、sklearn等。

本文在SpaceNews数据集上做了4组实验：①DCG-TextRank与传统无监督摘要抽取模型进行对比；②语义筛选模块调优；③DCG-TextRank与主流有监督文本摘要模型进行对比；④实例分析各模型在航天新闻上的摘要抽取结果。

3.4 实验1

3.4.1 对比模型

1）TextRank[21]：将单词共现个数作为句子间的相关度，采用图排序算法PageRank对句子排序。

2）TFIDF + TextRank：采用词频-逆文档频率（term frequency-inverse document frequency，TFIDF）算法抽取文章关键词，将共享关键词个数作为句子间相关度。

3）doc2vec+TextRank：使用doc2vec模型学习句向量，采用句向量余弦相似度作为句子相关度。

4）BERT(未微调)+TextRank：将BERT作为句子编码器，并采用BERT句向量相似度作为TextRank中句向量相关度。

5）BERT+TextRank：在SpaceNews数据集上对BERT进行微调并实验。

3.4.2 实验细节

1）BERT模型细节：使用SpaceNews数据集，在开源BERT预训练模型bert-base-uncased上微调，模型共12层，注意力头（attention_heads）个数为12，最大序列长度为512个单词，隐藏层维度为768，隐层采用gelu激活函数。微调时学习率设置为 1×10^{-5}，共训练10个epoch。

2）doc2vec参数设置：采用句向量的分布记忆模型（distributed memory model of paragraph vectors，PV-DM），句向量维度与BERT相同，均为768，预测上下文最大距离为3，初始学习率为0.025，共训练10个epoch，

3）TFIDF设置：在航天新闻文本中，领域关键词一般不超过5个，TFIDF参数topK=5。

4）TextRank迭代次数设置：为了实验的公平性，TextRank最大迭代次数均为200。

3.4.3 实验结果对比分析

第1组实验将本文模型与传统无监督模型TextRank等进行对比，验证本文模型DCG-TextRank在特定领域自动摘要的有效性，结果对比如表2所示所有模型原文抽取比例 r 均设置为20%，DCG-TextRank语义筛选模块语义保留度 γ=20%。其中，DCG-TextRank-S1：仅加入BERT模块和句子位置权重信息。DCG-TextRank-S2：上述基础上加入领域术语和关键词信息，构建领域概念图。DCG-TextRank：本文模型，在上述模型基础上加入语义筛选模块。

由表2可知，①与传统无监督模型TextRank相比，加入doc2vec和BERT句子语义编码的TextRank在数据集上的 F_{average} 分别提升了3.46%和4.06%，说明加入句子编码模块能够有效提升TextRank的性能，且BERT要优于doc2vec。TFIDF+TextRank相比其他传统模型取得了最好的效果，这是因为航天领域文本摘要更关注领域关键词汇信息。②本文模型在SpaceNews数据集上效果最佳，F_1、F_2、F_{L} 分别达到37.97%、26.43%、36.03%，说明本文模型能够生成更高质量的航天领域文本摘要。③微调后的BERT+TextRank模型 F_{average} 提高了0.4%，虽然微调后性能提升较为微弱，但也说明微调后BERT

学习到了航天领域文本的特点。④比较 DCG-TextRank-S1 和仅加入 BERT 的 TextRank, 可以看出, 融入句子位置信息能够有效提高图排序模型性能, F_{average} 提高了 2.46%。⑤DCG-TextRank-S2 相比 DCG-TextRank-S1, F_{average} 提高了 1.97%, 说明在文本图中加入领域术语能够有效引导图排序, 使得模型更加关注那些包含领域关键信息的句子, 提高模型性能。⑥在加入语义筛选模块后, DCG-TextRank 模型 F_{average} 提升了 3.09%, 说明经过节点聚类和语义筛选后, 候选句子间冗余降低, 与主旨无关的信息减少, 所提语义筛选方法能够捕获更精细、全面的语义信息, 有效提升候选摘要的质量。

表 2　本文模型与传统无监督模型实验结果对比

Table 2　Comparison of experimental results between the proposed model and traditional unsupervised models　%

模型	F_1	F_2	F_{L}	F_{average}
TextRank	26.98	13.44	24.06	21.50
TFIDF+TextRank	31.32	19.29	29.86	26.82
doc2vec+TextRank	30.05	16.95	27.87	24.96
BERT(未微调)+ TextRank	30.71	17.43	28.54	25.56
BERT+TextRank	31.08	17.96	28.84	25.96
DCG-TextRank-S1	33.12	20.90	31.25	28.42
DCG-TextRank-S2	34.95	23.10	33.12	30.39
DCG-TextRank	37.97	26.43	36.03	33.48

3.5　实验 2

第 2 组实验研究了不同摘取比例 r 和语义保留度 γ 对 DCG-TextRank 模型性能的影响, 确定最优参数, 实验评价标准为 F_{average} 得分。

首先固定 γ=20%, 研究 r 对摘要质量的影响。一般的文本摘要长度不应超过原文的 20%, 实验研究了 r 为 1%～20% 时, 生成摘要质量的变化, 步长为 1%。如图 3 所示, 随着 r 的增大, 摘要质量先增后减, 当 r=14% 时, 模型性能最优, 此时 F_{average}=34.50%, 摘要平均长度为 51.506 个单词。这是因为当摘要过短时, 会丢失很多关键信息, 而当摘要过长时, 又会导致信息冗余并引入多余的无用信息, 所以当输出摘要长度为原文的 14% 时, F_{average} 最高。

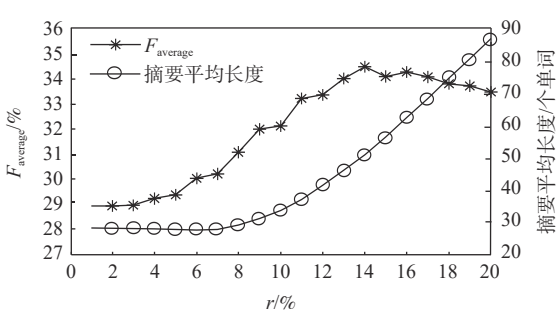


图 3　参数 r 对模型性能的影响

Fig. 3　Impact of parameter r on model performance

接着固定 r=14%,对比不同 γ 取值下的模型性能。如图 4 所示, 当 γ=6% 时, 模型性能最佳, 这是因为得分总值 6% 的句子已经涵盖了文章的大部分信息, 随着 γ 逐渐增长, 输出句子的冗余度也逐渐增大, 从而导致算法性能降低。当 γ 达到 18% 后, 模型性能达到稳定, F_{average} 和摘要平均长度不再变化。这说明在未采用 γ 进行信息过滤前, 从原文摘取句子的得分总值大多为 18% 左右, 从而导致在 γ 大于 18% 时, 未能起到过滤作用。

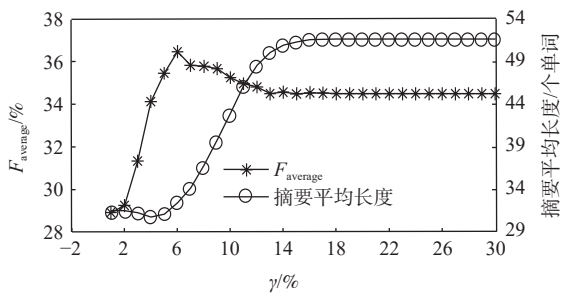


图 4　参数 γ 对模型性能的影响

Fig. 4　Impact of parameter γ on model performance

最终的最佳参数组合为: r=14%, γ=6%, 此时 F_{average}=36.47%, 摘要平均长度为 32.371 个单词。

3.6　实验 3

3.6.1　对比模型

1) 抽取式模型。① BertSum + Classifier[34]: 采用 BERT 作为句子编码层, 摘要判断层采用线性分类器对句子进行打分。② BertSum + Transformer[34]: 在 BertSum 后增加 Transformer 模块, 提取文档级特征。③ BertSum + RNN[34]: 采用 RNN 提取文档级特征。④ MatchSum[35]: 首先采用 BertSum 提取句子级候选摘要句, 对其进行组合形成文档级候选摘要, 然后通过基于 BERT 的孪生网络对原文和候选摘要进行编码, 并采用余弦相似度作为候选摘要得分。

2) 生成式模型。① Seq2Seq+Attention: 编码器和解码器分别采用双向和单向长短时记忆(long short-term memory, LSTM)网络, 并融入注意力机制。② Pointer-generator[36]: 将指针网络加入 Seq2Seq+Attention 模型中, 在保留模型生成新词能力的同时, 通过指针网络从原文复制单词, 提高信息准确性。③ Pointer-generator+Coverage [36]: 在 Pointer-generator 的基础上加入 Coverage 机制, 在 decoder 端加入注意力重复惩罚机制, 有效防止生成重复文本。

3.6.2　实验结果对比分析

第 3 组实验选择了 7 个有监督文本摘要的模型, 分别在 SpaceNews 数据集(训练集、测试集比例为 6:4)上进行实验, 与 DCG-TextRank 进行对比, 如

表3所示。实验中词向量维度均设置为768, DCG-TextRank参数设置为: r=14%, γ=6%。

由表3可知, ①在缺乏大量训练数据集情况下, 抽取式模型MatchSum和BertSum在SpaceNews数据集上性能均低于本文模型, 说明在缺乏大量领域标注数据集的情况下, 本文模型具有一定的优势。②本文模型在$F_{average}$指标上优于3个生成式模型。其中基本Seq2Seq模型在SpaceNews数据集上效果欠佳, 这是因为Seq2Seq模型需要大量训练数据来学习文本的特点, 不适用于缺乏标注数据集的航天新闻领域。Pointer-generator和Pointer-generator+Coverage虽然相比Seq2Seq性能有很大的提升, 但是在小训练样本下, 其仍存在生成式模型不连贯、不稳定、单词重复的缺陷, 摘要对比实例如表4所示, 因此, 不适合应用于对信息准确性要求严格的航天新闻领域。③加入Coverage机制的Pointer-generator模型F_2指标远低于本文模型, 但是F_1和F_L值却高于本文模型, 通过对其生成摘要进行分析, 认为原因是: 在摘要生成中, 指针网络倾向于从原文中复制单词, 而在加入Coverage机制后, 一定程度上避免了同一位置词汇的重复生成, 从而使得生成摘要与参考摘要间存在较长的非连续公共子序列, 所以F_1和F_L值较高; 但由于其生成摘要连贯性较差, 共现2元组较少, 所以F_2值较低。如表4所示, 本文模型在摘要的可读性、连贯性及关键信息含量上要优于Pointer-generator +Coverage模型。

表3 本文模型与主流有监督模型实验结果对比

Table 3 Comparison of experimental results between proposed model and mainstream supervised models %

模型	F_1	F_2	F_L	$F_{average}$
BertSum + Classifier	31.71	14.60	24.18	23.50
BertSum+RNN	36.25	25.20	33.74	31.73
BertSum+Transformer	38.09	24.78	33.42	32.10
MatchSum	37.73	25.05	32.54	31.77
Seq2Seq+Attention	13.69	5.35	11.94	10.32
Pointer-generator	38.11	9.76	36.78	28.22
Pointer-generator+Coverage	45.67	21.61	44.17	37.15
DCG-TextRank	41.17	30.06	38.18	36.47

3.7 实验4

从数据集中选取标题为“Atlas 5 launches NASA asteroid sample return mission”的航天发射新闻, 部分模型生成的摘要结果如表4所示, 参考摘要与关键信息如表5所示, 其中加黑单词为文本中的重要

表4 摘要对比实例

Table 4 Example of summary comparison

模型	摘要
TFIDF+TextRank	The spacecraft will deposit the samples collected by TAGSAM into a sample return canister based on the same design used for the Stardust comet dust **sample return mission**. Collection of the sample will mark the end of the science phase of the mission as the spacecraft moves to a safe distance from the asteroid.
BERT+TextRank	**OSIRIS-REx** is the third mission in **NASA's** New Frontiers program of medium-sized planetary missions, with an estimated cost of $800 million plus launch and operations. A few days before Juno's arrival, **NASA** formally extended the mission for New Horizons, the first New Frontiers mission, which flew past the dwarf planet Pluto in July 2015.
BertSum+Transformer	An **Atlas 5** successfully launched a **NASA** mission to **visit a near Earth asteroid and return samples of it to Earth** Sept. 8. The United Launch Alliance **Atlas 5** 411 lifted off at 7:05 p.m. Eastern from Space Launch Complex 41 at Cape Canaveral, Florida.
MatchSum	An **Atlas 5** successfully launched a **NASA** mission to **visit a near Earth asteroid and return samples of it to Earth** Sept. 8. TAGSAM is designed to collect at least 60 grams of material, but project officials are confident it can collect much more, perhaps up to 2 kilograms. The spacecraft will deposit the samples collected by TAGSAM into a sample return canister based on the same design used for the Stardust comet dust **sample return mission**.
Pointer-generator	Atlas atlas successfully launched a **nasa** mission mission visit a earth earth earth earth of of to to earth …(后面有较长重复单词, 未列出)
Pointer-generator+Coverage	**Atlas 5** launched launched a **nasa nasa** mission to a a earth earth and and of of to to earth earth sept sept sept sept …(后面有较长重复单词, 未列出)
DCG-TextRank	An **Atlas 5** successfully launched a **NASA** mission to **visit a near Earth asteroid and return samples of it to Earth** Sept.8. **OSIRIS-REx** is the third mission in **NASA's** New Frontiers program of **medium-sized planetary missions**, with an estimated cost of $800 million plus launch and operations.

表5 参考摘要与关键信息

Table 5 Reference summary and key information

参考摘要	关键信息
An **Atlas 5** successfully launched a **NASA** mission to **visit a near Earth asteroid and return samples of it to Earth** Sept.8. **OSIRIS-REx** is the third mission in **NASA**'s New Frontiers program of **medium-sized planetary missions**, with an estimated cost of $800 million plus launch and operations. **NASA** is also preparing for the competition for **the fourth New Frontiers mission**.	运载火箭: **Atlas 5** 机构: **NASA** 本次发射任务: **visit a near Earth asteroid and return samples of it to Earth** 任务系列: **medium-sized planetary missions** 第三次任务: **OSIRIS-Rex** 计划: **the fourth New Frontiers mission**

情报信息。可以看出,生成式模型在摘要的准确性和可读性上效果欠佳。相比其他抽取式文本摘要模型,本文模型对新闻中的航天事件表述更加完整,且在关键信息的捕获上有很好的效果,针对此航天发射新闻,除了美国国家航空航天局(National Aeronautics and Space Administration, NASA)的下一步发射计划外,其他关键信息均在摘要中有所反馈。

4　结　论

1) 相比传统图排序模型,利用领域概念图引导句子排序,使得包含领域关键信息的句子得到更高的重视;

2) 综合考虑句子位置、领域关键词覆盖度等句子特征,辅助学习文章的全局语义信息;

3) 采用图节点聚类及设置摘要语义保留度的方法改进 TextRank 的输出,充分保留文本的多语义信息并降低冗余。

4) 在 SpaceNews 数据集上,本文模型相比传统 TextRank 模型性能提高了 14.97%,相比有监督抽取式文本摘要模型 BertSum 和 MatchSum 性能提升了 4.37%～12.97%。

5) 本文模型采用自监督学习方法对 BERT 进行微调,并结合无监督图排序方法进行文本自动摘要,无需进行文本数据的标注,对于一个新的领域,只需提供领域术语库即可,可移植性强。

参考文献（References）

[1] 冯鸾鸾, 李军辉, 李培峰, 等. 面向国防科技领域的技术和术语语料库构建方法[J]. 中文信息学报, 2020, 34(8): 41-50.
FENG L L, LI J H, LI P F, et al. Constructing a technology and terminology corpus oriented national defense science[J]. Journal of Chinese Information Processing, 2020, 34(8): 41-50(in Chinese).

[2] MAURYA P, JAFARI O, THATTE B, et al. Building a comprehensive NER model for satellite domain[J]. SN Computer Science, 2022, 3(3): 1-8.

[3] JAFARI O, NAGARKAR P, THATTE B, et al. SatelliteNER: An effective named entity recognition model for the satellite domain[C]//12th International Conference on Knowledge Management and Information Systems. Beijing: SCITEPRESS, 2020: 100-107.

[4] LU Y, YANG R, JIANG X, et al. A military named entity recognition method based on pre-training language model and BiLSTM-CRF[J]. Journal of Physics: Conference Series, 2020, 1693(1): 012161

[5] 高翔, 张金登, 许潇, 等. 基于 LSTM-CRF 的军事动向文本实体识别方法[J]. 指挥信息系统与技术, 2020, 11(6): 91-95
GAO X, ZHANG J D, XU X, et al. Military trend text entity recognition method based on LSTM-CRF[J]. Command Information System and Technology, 2020, 11(6): 91-95(in Chinese).

[6] SHINDE M, MHATRE D, MARWAL G. Techniques and research in text summarization—A survey[C]//2021 International Conference on Advance Computing and Innovative Technologies in Engineering. Piscataway: IEEE Press, 2021: 260-263.

[7] ZHENG C, ZHANG K, WANG H J, et al. Enhanced Seq2Seq autoencoder via contrastive learning for abstractive text summarization[C]//2021 IEEE International Conference on Big Data. Piscataway: IEEE Press, 2021: 1764-1771.

[8] PRASAD C, KALLIMANI J S, HAREKAL D, et al. Automatic text summarization model using Seq2Seq technique[C]//2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud). Piscataway: IEEE Press, 2020: 599-604.

[9] SHI T, KENESHLOO Y, RAMAKRISHNAN N, et al. Neural abstractive text summarization with sequence-to-sequence models[J]. ACM Transactions on Data Science, 2021, 2(1): 1-37.

[10] WANG Q, REN J. Summary-aware attention for social media short text abstractive summarization[J]. Neurocomputing, 2021, 425: 290-299.

[11] MIAO W, ZHANG G, BAI Y, et al. Improving accuracy of key information acquisition for social media text summarization[C]//2019 IEEE International Conferences on Ubiquitous Computing & Communications(IUCC) and Data Science and Computational Intelligence(DSCI) and Smart Computing, Networking and Services. Piscataway: IEEE Press, 2019: 408-415.

[12] NALLAPATI R, FEIFEI Z, Bowen Z. Summarunner: A recurrent neural network based sequence model for extractive summarization of documents[C]//Thirty-first AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2017: 3075-3081.

[13] WANG Y, TANG S, ZHU F, et al. Revisiting the transferability of supervised pretraining: An MLP perspective[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2021: 9173-9183.

[14] VANDERWENDE L, SUZUKI H, BROCKETT C, et al. Beyond SumBasic: Task-focused summarization with sentence simplification and lexical expansion[J]. Information Processing & Management, 2007, 43(6): 1606-1618.

[15] CAO Z, WEI F, DONG L, et al. Ranking with recursive neural networks and its application to multi-document summarization[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2015, 29(1): 2153-2159.

[16] HONG K, CONROY J M, FAVRE B, et al. A repository of state of the art and competitive baseline summaries for generic news summarization[C]//Proceedings of the LREC. Marseille: LREC, 2014: 1608−1616.

[17] LIU N, LU Y, TANG X J, et al. Multi-document summarization algorithm based on significance sentences[C]//2016 Chinese Control and Decision Conference. Piscataway: IEEE Press, 2016: 3847-3852.

[18] WU R S, LIU K, WANG H L. An evolutionary summarization system based on local-global topic relationship[J]. Journal of Chinese Information Processing, 2018, 32(9): 75-83.

[19] LITVAK M, VANETIK N, LIU C, et al. Improving summarization quality with topic modeling[C]//Proceedings of the 2015 Workshop on Topic Models: Post-Processing and Applications. New York: ACM, 2015: 39-47.

[20] ERKAN G, RADEV D R. LexRank: Graph-based lexical centrality as salience in text summarization[J]. Journal of Artificial Intelligence Research, 2004, 22(1): 457-479.

[21] MIHALCEA R, TARAU P. TextRank: Bringing order into text[C]//Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing. Barcelona: EMNLP, 2004: 404-411.

[22] 叶雷, 余正涛, 高盛祥, 等. 多特征融合的汉越双语新闻摘要方法[J]. 中文信息学报, 2018, 32(12): 84-91.
YE L, YU Z T, GAO S X, et al. A bilingual news summarization in Chinese and Vietnamese based on multiple features[J]. Journal of Chinese Information Processing, 2018, 32(12): 84-91(in Chinese).

[23] 余珊珊, 苏锦钿, 李鹏飞. 基于改进的 TextRank 的自动摘要提取方法[J]. 计算机科学, 2016, 43(6): 240-247.
YU S S, SU J D, LI P F. Improved TextRank-based method for automatic summarization[J]. Computer Science, 2016, 43(6): 240-247(in Chinese).

[24] WAN X, YANG J, XIAO J. Towards an iterative reinforcement approach for simultaneous document summarization and keyword extraction[C]//Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics. Stroudsburg: ACL, 2007: 552-559.

[25] 汪旭祥, 韩斌, 高瑞, 等. 基于改进 TextRank 的文本摘要自动提取[J]. 计算机应用与软件, 2021, 38(6): 155-160.
WANG X X, HAN B, GAO R, et al. Automatic extraction of text summarization based on improved TextRank[J]. Computer Applications and Software, 2021, 38(6): 155-160(in Chinese).

[26] 黄波, 刘传才. 基于加权 TextRank 的中文自动文本摘要[J]. 计算机应用研究, 2020, 37(2): 407-410.
HUANG B, LIU C C. Chinese automatic text summarization based on weighted TextRank[J]. Application Research of Computers, 2020, 37(2): 407-410(in Chinese).

[27] 李峰, 黄金柱, 李舟军, 等. 使用关键词扩展的新闻文本自动摘要方法[J]. 计算机科学与探索, 2016, 10(3): 372-380.
LI F, HUANG J Z, LI Z J, et al. Automatic summarization method of news texts using keywords expansion[J]. Journal of Frontiers of Computer Science and Technology, 2016, 10(3): 372-380(in Chinese).

[28] 方萍, 徐宁. 基于 BERT 双向预训练的图模型摘要抽取算法[J]. 计算机应用研究, 2021, 38(9): 2657-2661.
FANG P, XU N. Graph model summary extraction algorithm based on BERT bidirectional pretraining[J]. Application Research of Computers, 2021, 38(9): 2657-2661(in Chinese).

[29] FERREIRA R, LINS R D, FREITAS F, et al. A new sentence similarity method based on a three-layer sentence representation[C]//2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technologies. Piscataway: IEEE Press , 2014: 110-117.

[30] DEVLIN J, CHANG M W, LEE K, et al. BERT: Pre-training of deep bidirectional transformers for language understanding[C]//Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Stroudsburg: ACL, 2019: 4171–4186.

[31] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Advances in Neural Information Processing Systems. New York: ACM, 2017: 5998-6008.

[32] LIN C Y, HOVY E. Identifying topics by position[C]//Fifth Conference on Applied Natural Language Processing. New York: ACM, 1997: 283-290.

[33] SEIFIKAR M, FARZI S, BARATI M. C-blondel: An efficient Louvain-based dynamic community detection algorithm[J]. IEEE Transactions on Computational Social Systems, 2020, 7(2): 308-318.

[34] LIU Y, LAPATA M. Text summarization with pretrained encoders[C]//Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing. Stroudsburg: ACL, 2019: 3721–3731.

[35] ZHONG M, LIU P, CHEN Y, et al. Extractive summarization as text matching[C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. Stroudsburg: ACL, 2020: 6197-6208.

[36] SEE A, LIU P J, MANNING C D. Get to the point: summarization with pointer-generator networks[C]//Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics. Stroudsburg: ACL, 2017: 1073-1083.

Automatic summarization model of aerospace news based on domain concept graph

HUANG Haoning[1, 2], CHEN Zhimin[1, *], XU Cong[1, 2], ZHANG Xiaoyan[3]

(1. National Space Science Center, Chinese Academy of Sciences, Beijing 100190, China;
2. University of Chinese Academy of Sciences, Beijing 100049, China;
3. State Radio Monitoring Center, Beijing 100037, China)

Abstract: The effectiveness of subsequent intelligence analysis can be increased by comprehending and compressing the vast amount of aerospace information that is hidden in the Internet's aerospace news. However the general automatic summarization algorithms tend to ignore many domain key Information, and the existing supervised automatic summarization algorithms need to annotate a lot of data in the domain text. It is time-consuming and laborious. Therefore, we proposed an unsupervised automatic summarization model TextRank based on domain concept graph (DCG-TextRank). It is based on a domain concept graph, which uses domain terms to help guide graph ordering and improve the model's understanding of domain text. The model has three modules: domain concept graph generation, graph weight initialization, graph sorting and semantic filtering. Transform the text into domain concept graph containing sentence nodes and domain term nodes according to sentence vector similarity and domain term database. Initialize the domain concept graph weight according to the features of aerospace news text. Use the TextRank algorithm to sort the sentences, and in the semantic filtering module, the output of TextRank is improved by clustering the graph nodes and setting the semantic retention of the abstract, which fully preserves the semantic Information of text and reduces redundancy. The proposed model is domain portable, and experimental findings indicate that in the aerospace news dataset, the proposed model performs 14.97% better than the conventional TextRank model and 4.37%～12.97% better than the supervised extraction text summary models BertSum and MatchSum.

Keywords: automatic text summarization; domain concept graph; Pre-trained language model; graph sorting algorithm; graph node clustering

Received: 2022-04-12; **Accepted**: 2022-06-25; **Published Online**: 2022-08-26 16:50

URL: link.cnki.net/urlid/11.2625.V.20220826.1046.004

Foundation items: National Natural Science Foundation of China (91738101); National Key R & D Program of China (2020YFB1807900)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0243

基于隐含非线性维纳退化过程的剩余寿命预测

杨家鑫[1]，唐圣金[1,*]，李良[1]，孙晓艳[2]，祁帅[3]，司小胜[1]

(1. 火箭军工程大学 导弹工程学院，西安 710025；2. 火箭军工程大学 作战保障学院，西安 710025；
3. 火箭军装备部驻郑州军代表室，郑州 450000)

摘　　要：准确的剩余寿命预测有助于提高系统的可靠性安全性，并降低系统全生命周期的经济成本。工程应用中，由于不确定测量的影响，随机退化系统的非线性退化特征一般处于隐含状态。针对现有隐含尺度变换非线性维纳退化过程主要应用于退化建模和寿命分布估计的问题，提出一种针对隐含尺度变换非线性维纳退化过程的剩余寿命预测方法。建立同时考虑测量误差和尺度变换非线性维纳过程的退化模型，并通过Kalman滤波方法根据设备现场退化数据进行参数在线更新。推导出考虑模型参数在线更新的剩余寿命概率密度函数和累积分布函数解析表达式。基于历史退化数据，提出一种针对隐含尺度变换非线性维纳退化过程模型未知参数的极大似然无偏估计方法。通过仿真退化数据和实际涡扇发动机数据进行实验验证。实验结果表明：同时考虑测量不确定性和非线性退化特征的剩余寿命预测方法具有更高的预测精度。

关　键　词：剩余寿命；非线性；尺度变换；不确定测量；维纳过程；Kalman滤波
中图分类号：TB114.3
文献标志码：A　　**文章编号**：1001-5965(2024)01-0328-13

近年来，随着航空航天和武器系统等高新技术领域的快速发展，设备的可靠性安全性要求越来越高。如果设备出现故障，将可能导致系统停机，甚至造成难以想象的后果[1]。工程实践表明，预测与健康管理(prognoses and health management, PHM)技术有助于降低系统故障风险和提高设备可靠性[2-3]。在PHM技术中，准确预测剩余寿命是做出正确维护决策的关键。因此，剩余寿命预测成为可靠性研究领域的热点问题[4-5]。

在当前的剩余寿命预测方法中，数据驱动方法通常不需要建立设备复杂的退化机理模型，具有较强的泛化能力[6]。其主要分为人工智能方法和统计数据驱动方法2类。其中，人工智能方法利用人工智能算法拟合同类设备的特征退化数据，从而获取设备的退化路径，实现剩余寿命预测[7]。虽然该方法不需要建立具体数学模型，应用场景广泛，但为了保证预测精度，往往需要大量退化数据，会导致较高的经济成本和较长的计算时间。统计数据驱动方法根据可用的观测数据和统计模型，以概率的方式估计设备的剩余寿命，具有较好的数学性质，便于观测比较，所以在剩余寿命预测领域应用广泛[8]。值得注意的是，设备的退化过程容易受到外界环境和自身制造工艺的影响，退化存在不确定性。为解决这个问题，通常使用概率密度函数(probability density functions, PDF)来描述剩余寿命[9]。在统计数据驱动方法中，随机过程方法具有较好描述设备退化随机性的能力，因此，统计数据驱动方法以随机过程建模为主。最常用的建模方法为：伽马过程[10]、逆高斯过程[11]和维纳过程[2-4, 6, 12]。其中，伽马过程和逆高斯过程只适用于单调的退化系统，存在建模

收稿日期：2022-04-15；**录用日期**：2022-05-13；**网络出版时间**：2022-05-30 16:43
网络出版地址：link.cnki.net/urlid/11.2625.V.20220527.1918.001
基金项目：国家自然科学基金(61703410,61873175,61873273,61773386,61922089)；陕西省自然科学基础研究计划(2022JM-376)
*通信作者. E-mail：tangshengjin27@126.com

*引用格式：杨家鑫，唐圣金，李良，等. 基于隐含非线性维纳退化过程的剩余寿命预测[J]. 北京航空航天大学学报，2024，50(1)：328-340.
YANG J X，TANG S J，LI L，et al. Remaining useful life prediction based on implicit nonlinear Wiener degradation process[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(1)：328-340 (in Chinese).*

的局限性, 而维纳过程适用于单调和非单调的退化系统, 因此, 维纳过程模型在近些年被广泛使用[13]。

使用维纳过程的剩余寿命预测基本原理是根据设备历史退化数据估计表示设备共性特征的模型未知参数, 再结合被测设备的实时退化状态与统计理论预测设备距离当前时刻何时失效。一般情况下, 由于测量仪器、操作人员技术和环境等因素, 测量的退化数据容易受到污染, 不能直接反映真实的退化状态, 在剩余寿命预测时需要考虑测量误差的影响[14]。Peng 和 Tseng[15] 建立了同时考虑随机效应、退化过程不确定性和测量误差 3 层不确定性的线性退化模型。在此基础上, Si 等[16] 针对包含这种 3 层不确定性的线性退化模型, 基于 Kalman 滤波得出退化速度和退化状态的联合后验分布, 并得到相应的剩余寿命 PDF 解析表达式。然而, 在工程应用中, 存在大量线性模型无法有效跟踪预测的非线性退化过程。针对该问题, 一些学者进一步展开了非线性维纳过程建模研究。Feng 等[17] 考虑退化过程不确定性和测量误差, 推导了非线性退化场合的剩余寿命分布, 郑建飞等[4] 进一步建立了同时考虑 3 层不确定性的非线性维纳过程一般化模型, 并得到了剩余寿命 PDF 的解析表达式。可以看出, 针对一般隐含维纳过程退化模型的在线剩余寿命预测方法研究已经较为成熟。

上述直接建模的非线性退化过程, 其动态过程并不存在非线性退化特性。另一些学者认为非线性退化过程的动态特性也存在非线性特征, 并使用尺度变换维纳过程模型进行退化建模和寿命估计[18-23]。该模型最早是由 Whitmore 和 Schenkelberg[18] 提出, 并基于维纳过程模型首次到达失效阈值时间服从逆高斯分布的理论, 推导寿命分布的累积分布函数(cumulative distribution function, CDF)。Ye 等[19] 进一步通过正态分布描述随机效应, 得到了考虑随机效应的寿命分布估计。而后, Pan 等[20] 基于全概率定律推导了考虑测量误差的寿命分布。基于删节正态分布描述随机效应的隐含尺度变换非线性维纳退化过程的寿命分布可参考文献 [21]。文献 [18-21] 主要是研究隐含尺度变换非线性维纳过程的退化建模和寿命分布估计问题, 并没有研究剩余寿命预测问题。在剩余寿命预测的应用场合, Ye 等[22] 针对直接估计当前退化状态复杂的问题, 提出一种递归估计退化率和当前退化状态联合分布的算法, 然而并没有得出剩余寿命分布的解析表达式, 只能通过数值算法计算剩余寿命分布。Huang 等[23] 针对尺度变换非线性维纳退化过程, 得出考虑正态分布随机效应的剩余寿命密度函数解析表达式, 但并没有考虑测量误差的影响。综上所述, 针对隐含尺度变换非线性维纳退化过程, 现有文献主要研究退化建模和寿命分布估计问题, 在剩余寿命预测方面研究较少, 基于尺度变换的隐含维纳退化过程模型的剩余寿命问题还需要进一步研究。

离线参数估计是剩余寿命预测的前提。目前离线参数估计方法种类繁多, 例如, 传统的极大似然估计(maximum likelihood estimation, MLE)方法[15], 通过迭代次数逐渐逼近真实参数值的期望最大化算法[12] 及文献 [24-25] 针对 MLE 方法估计漂移系数方差可能为负的问题而提出的 2 步 MLE 方法等。对于隐含尺度变换非线性维纳退化过程, 考虑随机效应退化模型的参数估计方法主要是传统 MLE 方法[19] 和期望最大化算法[20-21]。为了提高参数估计精度, Tang 等[26] 推导了考虑随机效应线性维纳退化过程参数估计的解析解, 研究了参数估计结果的性质, 并提出线性退化过程、隐含线性退化过程和非线性退化过程的无偏参数估计方法, 取得了令人满意的结果。Han 等[27] 进一步将该方法推广到非线性隐含退化过程, 并提出相应的无偏参数估计方法。然而, 该方法尚未应用于带测量误差的隐含尺度变换非线性维纳退化过程。

综上分析, 在同时考虑测量误差和非线性影响的情况下, 基于隐含非线性尺度变换维纳退化过程的模型参数估计和剩余寿命预测在理论上还有待完善。本文首先建立隐含非线性退化过程的状态空间模型, 然后采用 Kalman 滤波方法在线更新模型参数, 进一步推导考虑状态估计的剩余寿命分布的 PDF 和 CDF 表达式。接着, 提出针对隐含尺度变换非线性维纳退化过程的无偏参数估计方法。最后, 通过仿真数据和涡扇发动机数据分析比较不同退化模型的预测效果。

1　基于隐含尺度变换非线性维纳退化过程的剩余寿命预测

1.1　退化建模

在实际应用中, 随机退化设备的非线性退化特征广泛存在, 基本线性退化模型难以有效跟踪非线性特征, 本节基于尺度变换的非线性维纳退化模型进行剩余寿命预测。令 $X(t)$ 表示被测设备在 t 时刻的真实退化量, 则其退化模型具体表达式可以表示为

$$X(t)=\lambda\Lambda(t;\theta)+\sigma_B B(\Lambda(t;\theta)) \tag{1}$$

式中：$\Lambda(t;\theta)$为以θ为特征参数的单调连续非线性函数，典型的非线性函数有2种，分别为$\Lambda(t;\theta)=t^{\theta}$和$\Lambda(t;\theta)=\exp(\theta t)-1$；$\sigma_B$为扩散系数，表示退化过程的动态特性，因式(1)中$\Lambda(t;\theta)\neq t$，所以$B(\Lambda(t;\theta))$为非标准布朗运动；$\lambda$为漂移系数，表征设备退化速度，通常假设$\lambda$服从均值为$\mu_\lambda$，方差为$\sigma_\lambda^2$的正态分布，即$\lambda\sim N(\mu_\lambda,\sigma_\lambda^2)$，以表示不同设备间的个体差异，且假设$\lambda$与$B(\Lambda(t;\theta))$相互独立。

测量误差会增加预测的不确定性，导致剩余寿命的预测结果偏离真实值，进一步造成操作人员做出错误的维护决策。因此，需要考虑测量误差对预测精度的影响，假设任意时刻的测量误差ε独立同分布，都服从均值为0，方差为σ_ε^2的正态分布，即$\varepsilon\sim N(0,\sigma_\varepsilon^2)$，且与$\lambda$和$B(\Lambda(t;\theta))$相互独立。则退化模型表达式为

$$Y(t)=\lambda\Lambda(t;\theta)+\sigma_B B(\Lambda(t;\theta))+\varepsilon \tag{2}$$

1.2 基于Kalman滤波的参数在线更新

如式(2)所示，模型中的漂移系数λ用于描述不同设备之间的差异，为了适应被预测设备的个体退化特征，需要对被评估设备进行参数更新。本节将Si等[16]针对线性退化系统建立的状态空间模型进一步衍生到非线性退化系统，再利用Kalman滤波方法根据被测设备的现场退化数据在线更新漂移系数λ。

为了准确识别设备当前退化状态，以离散时间方程表示状态方程和观测方程。基于式(1)和式(2)建立的退化模型，当检测到被测设备在t_k时刻的退化值$\boldsymbol{Y}_{0:k}$时，t_k时刻的状态空间方程可表示为

$$\begin{cases} x_k=x_{k-1}+\lambda_{k-1}(\Lambda(t_k;\theta)-\Lambda(t_{k-1};\theta))+v_k \\ \lambda_k=\lambda_{k-1} \\ y_k=x_k+\varepsilon_k \end{cases} \tag{3}$$

式中：x_k为t_k时刻的真实退化量；λ_k为t_k时刻的漂移系数；y_k为t_k时刻的退化测量值；$v_k=\sigma_B B(\Lambda(t_k;\theta))-\sigma_B B(\Lambda(t_{k-1};\theta))$；$\varepsilon_k$为$t_k$时刻的测量误差；$\{v_k\}_{k\geqslant 1}$和$\{\varepsilon_k\}_{k\geqslant 1}$独立同分布。根据维纳过程的性质，可得$v_k=\sigma_B B(\Delta\Lambda(t_k;\theta))$，则$v_k\sim N(0,\sigma_B^2\Delta\Lambda(t_k;\theta))$，其中，$\Delta\Lambda(t_k;\theta)=\Lambda(t_k;\theta)-\Lambda(t_{k-1};\theta)$。

由于测量误差的影响，真实退化状态x_k和漂移系数λ_k都是隐含状态。则可进一步将式(3)状态空间方程表示为

$$\begin{cases} \boldsymbol{z}_k=\boldsymbol{A}_k\boldsymbol{z}_{k-1}+\boldsymbol{\eta}_k \\ y_k=\boldsymbol{C}\boldsymbol{z}_k+\varepsilon_k \end{cases} \tag{4}$$

式中：$\boldsymbol{z}_k\in\mathbf{R}^{2\times 1}$；$\boldsymbol{A}_k\in\mathbf{R}^{2\times 2}$；$\boldsymbol{\eta}_k\in\mathbf{R}^{2\times 1}$；$\boldsymbol{C}\in\mathbf{R}^{1\times 2}$；$\boldsymbol{\eta}_k\sim N(0,\boldsymbol{Q}_k)$，$\boldsymbol{Q}_k\in\mathbf{R}^{2\times 2}$。具体数值为

$$\boldsymbol{z}_k=\begin{bmatrix} x_k \\ \lambda_k \end{bmatrix};\ \boldsymbol{A}_k=\begin{bmatrix} 1 & \Lambda(t_k;\theta)-\Lambda(t_{k-1};\theta) \\ 0 & 1 \end{bmatrix};$$

$$\boldsymbol{\eta}_k=\begin{bmatrix} v_k \\ 0 \end{bmatrix};\ \boldsymbol{C}=\begin{bmatrix} 1 \\ 0 \end{bmatrix}^{\mathrm{T}};\ \boldsymbol{Q}_k=\begin{bmatrix} \sigma_B^2\Delta\Lambda(t_k;\theta) & 0 \\ 0 & 0 \end{bmatrix}。$$

为便于后续公式推导，首先，定义变量$\boldsymbol{z}_k$在t_k时刻的期望和方差分别为

$$\hat{\boldsymbol{z}}_{k|k}=\begin{bmatrix} \hat{x}_{k|k} \\ \hat{\lambda}_{k|k} \end{bmatrix}=E(\boldsymbol{z}_k|\boldsymbol{Y}_{0:k}) \tag{5}$$

$$\boldsymbol{P}_{k|k}=\begin{bmatrix} \kappa_{x,k}^2 & \kappa_{x\lambda,k}^2 \\ \kappa_{x\lambda,k}^2 & \kappa_{\lambda,k}^2 \end{bmatrix}=\mathrm{cov}(\boldsymbol{z}_k|\boldsymbol{Y}_{0:k}) \tag{6}$$

式中：$\hat{x}_{k|k}=E(x_k|\boldsymbol{Y}_{0:k})$；$\hat{\lambda}_{k|k}=E(\lambda_k|\boldsymbol{Y}_{0:k})$；$\kappa_{x,k}^2=\mathrm{var}(x_k|\boldsymbol{Y}_{0:k})$；$\kappa_{\lambda,k}^2=\mathrm{var}(\lambda_k|\boldsymbol{Y}_{0:k})$；$\kappa_{x\lambda,k}^2=\mathrm{cov}(x_k\lambda_k|\boldsymbol{Y}_{0:k})$。

进一步，将变量$\boldsymbol{z}_k$的一步预测期望和方差分别定义为

$$\hat{\boldsymbol{z}}_{k|k-1}=\begin{bmatrix} \hat{x}_{k|k-1} \\ \hat{\lambda}_{k|k-1} \end{bmatrix}=E(\boldsymbol{z}_k|\boldsymbol{Y}_{0:k-1}) \tag{7}$$

$$\boldsymbol{P}_{k|k-1}=\begin{bmatrix} \kappa_{x,k-1}^2 & \kappa_{x\lambda,k-1}^2 \\ \kappa_{x\lambda,k-1}^2 & \kappa_{\lambda,k-1}^2 \end{bmatrix}=\mathrm{cov}(\boldsymbol{z}_k|\boldsymbol{Y}_{0:k-1}) \tag{8}$$

根据以上定义，Kalman滤波的预测和更新可具体表示为

预测：

$$\begin{cases} \hat{\boldsymbol{z}}_{k|k-1}=\boldsymbol{A}_k\hat{\boldsymbol{z}}_{k-1|k-1} \\ \boldsymbol{P}_{k|k-1}=\boldsymbol{A}_k\boldsymbol{P}_{k-1|k-1}\boldsymbol{A}_k^{\mathrm{T}}+\boldsymbol{Q}_k \end{cases}$$

更新：

$$\begin{cases} \hat{\boldsymbol{z}}_{k|k}=\hat{\boldsymbol{z}}_{k|\boldsymbol{k}-1}+\boldsymbol{K}(k)\left(y_k-\boldsymbol{C}\hat{\boldsymbol{z}}_{k|k-1}\right) \\ \boldsymbol{P}_{k|k}=\boldsymbol{P}_{k|k-1}-\boldsymbol{K}(k)\boldsymbol{C}\boldsymbol{P}_{k|k-1} \\ \boldsymbol{K}(k)=\boldsymbol{P}_{k|k-1}\boldsymbol{C}^{\mathrm{T}}\left[\boldsymbol{C}\boldsymbol{P}_{k|k-1}\boldsymbol{C}^{\mathrm{T}}+\sigma_\varepsilon^2\right]^{-1} \end{cases}$$

式中：初始值为$\hat{\boldsymbol{z}}_{0|0}=\begin{bmatrix} 0 \\ \mu_\lambda \end{bmatrix}$；$\boldsymbol{P}_{0|0}=\begin{bmatrix} 0 & 0 \\ 0 & \sigma_\lambda^2 \end{bmatrix}$；$\boldsymbol{K}(k)$为增益矩阵。

基于式(4)，根据Kalman滤波的高斯性质和贝叶斯原理，$\boldsymbol{z}_k|\boldsymbol{Y}_{0:k-1}$服从多维正态分布，即$\boldsymbol{z}_k|\boldsymbol{Y}_{0:k}\sim N(\hat{\boldsymbol{z}}_{k|k},\boldsymbol{P}_{k|k})$。则有

$$\begin{cases} \lambda_k|\boldsymbol{Y}_{0:k}\sim N\left(\hat{\lambda}_{k|k},\kappa_{\lambda,k}^2\right) \\ x_k|\boldsymbol{Y}_{0:k}\sim N\left(\hat{x}_{k|k},\kappa_{x,k}^2\right) \\ x_k|\lambda_k,\boldsymbol{Y}_{0:k}\sim N(\mu_{x_k|\lambda,k},\sigma_{x_k|\lambda,k}^2) \end{cases} \tag{9}$$

式中：

$$\begin{cases} \mu_{x_k|\lambda,k}=\hat{x}_{k|k}+\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}\left(\lambda_k-\hat{\lambda}_{k|k}\right) \\ \sigma_{x_k|\lambda,k}^2=\kappa_{x,k}^2\left(1-\rho_k^2\right) \\ \rho_k=\kappa_{x\lambda,k}^2/(\kappa_{\lambda,k}\kappa_{x,k}) \end{cases} \tag{10}$$

1.3 在线剩余寿命预测

剩余寿命定义为从当前时刻到设备出现故障

时刻的时间长度。当检测到现场退化数据$\boldsymbol{Y}_{0:k}$时(实际退化状态为$\boldsymbol{X}_{0:k}$), 在$t \geqslant t_k$时刻的实际退化过程可表示为

$$X(t) = x_k + \lambda(\Lambda(t;\theta) - \Lambda(t_k;\theta)) + \sigma_B B(\Lambda(t;\theta) - \Lambda(t_k;\theta)) \tag{11}$$

令$l_k = t - t_k\,(l_k \geqslant 0)$表示$t_k$时刻的剩余寿命, 则退化过程$\{X(t), t \geqslant t_k\}$可以转化为退化过程$\{Z(l_k), l_k \geqslant 0\}$:

$$\begin{aligned} Z(l_k) &= X(l_k + t_k) - X(t_k) = \\ &\lambda(\Lambda(t;\theta) - \Lambda(t_k;\theta)) + \sigma_B B(\Lambda(t;\theta) - \Lambda(t_k;\theta)) \end{aligned} \tag{12}$$

式中: $Z(0) = 0$。

设备随着使用时间的增加逐渐老化, 当退化量达到临界点时, 则无法正常运行。为了能够在设备出现故障之前做出有效维护, 通常定义设备的退化量首次达到某一阈值时失效。基于首达时间的概念, 设备在t_k时刻的剩余寿命可以转化为退化量$\{X(l_k), l_k \geqslant 0\}$首次穿过阈值$w_k = w - x_k$的时间[6, 24], 其中, w_k为t_k时刻的失效阈值, w为t时刻的失效阈值。相应地, 剩余寿命可以定义为

$$\begin{aligned} L_k &= \inf\{l_k : X(l_k + t_k) \geqslant w | \boldsymbol{X}_{0:k}\} = \\ &\inf\{l_k : Z(l_k) \geqslant w - x_k | \boldsymbol{X}_{0:k}\} \end{aligned} \tag{13}$$

基于式 (1) 退化模型, 设备真实退化过程剩余寿命的 PDF 为$f_{L_k|\lambda_k,x_k,\boldsymbol{Y}_{0:k}}(l_k|\lambda_k, x_k, \boldsymbol{Y}_{0:k})$和 CDF 为$F_{L_k|\lambda_k,x_k,\boldsymbol{Y}_{0:k}}(l_k|\lambda_k, x_k, \boldsymbol{Y}_{0:k})$分别如式(14)和式(15)所示, 详细推导过程参考文献 [23]。

$$f_{L_k|\lambda_k,x_k,\boldsymbol{Y}_{0:k}}(l_k|\lambda_k, x_k, \boldsymbol{Y}_{0:k}) = \frac{w - x_k}{\sqrt{2\pi\sigma_B^2\zeta_{l_k}^3}} \exp\left(-\frac{(w - x_k - \lambda_k\zeta_{l_k})^2}{2\sigma_B^2\zeta_{l_k}}\right)\frac{\mathrm{d}\zeta_{l_k}}{\mathrm{d}l_k} \tag{14}$$

$$F_{L_k|\lambda_k,x_k,\boldsymbol{Y}_{0:k}}(l_k|\lambda_k, x_k, \boldsymbol{Y}_{0:k}) = 1 - \Phi\left(\frac{w - x_k - \lambda_k\zeta_{l_k}}{\sqrt{\sigma_B^2\zeta_{l_k}}}\right) + \exp\left(\frac{2\lambda_k(w - x_k)}{\sigma_B^2}\right)\Phi\left(\frac{-(w - x_k) - \lambda_k\zeta_{l_k}}{\sqrt{\sigma_B^2\zeta_{l_k}}}\right) \tag{15}$$

式中: $\zeta_{l_k} = \Lambda(t_k + l_k;\theta) - \Lambda(t_k;\theta)$。

由于测量误差的影响, 真实退化量x_k为隐含状态。然而, 式 (14) 和式 (15) 的使用条件为x_k已知, 故该式不能用来预测不确定测量退化系统的剩余寿命。此处, 因 Kalman 滤波方法能够根据现场退化数据获得接近真实的退化状态x_k和具有状态特征的模型参数, 于是根据式 (14) 和式 (15) 可以进一步推导得出剩余寿命的 PDF 和 CDF 表达式。在此之前, 先给出引理 1 和引理 2。

引理 1[16]　如果$Z \sim N(u, \sigma^2)$, $w_1, w_2, A, B \in \mathbf{R}$, $G \in \mathbf{R}^+$, 则式(16)成立。

$$E_Z\left[(w_1 - AZ)\exp\left(-\frac{(w_2 - BZ)^2}{2G}\right)\right] = \sqrt{\frac{G}{B^2\sigma^2 + G}} \cdot \left(w_1 - A\frac{Bw_2\sigma^2 + \mu G}{B^2\sigma^2 + G}\right) \cdot \exp\left(-\frac{(w_2 - B\mu)^2}{2(B^2\sigma^2 + G)}\right) \tag{16}$$

引理 2[28]　如果$Z \sim N(u, \sigma^2)$, $A, B \in \mathbf{R}$, $G \in \mathbf{R}^+$, 则式(17)成立。

$$E_Z[\exp(AZ)\Phi(G + BZ)] = \exp\left(Au + \frac{A^2}{2}\sigma^2\right) \cdot \Phi\left(\frac{G + Bu + AB\sigma^2}{\sqrt{1 + B\sigma^2}}\right) \tag{17}$$

当$A = 0$时, 可得

$$E_Z[\Phi(G + BZ)] = \Phi\left(\frac{G + Bu}{\sqrt{1 + B^2\sigma^2}}\right)。$$

基于引理 1、引理 2 和参数的在线更新结果, 可得剩余寿命如定理 1 所示。

定理 1　对于式 (2) 所示的隐含尺度变换非线性维纳过程模型, 定义设备的失效阈值为w, 当检测到被测设备在t_k时刻的现场退化数据$\boldsymbol{Y}_{0:k}$时, 其剩余寿命的 PDF 和 CDF 表达式分别为

$$f_{L_k|\boldsymbol{Y}_{1:k}}(l_k|\boldsymbol{Y}_{0:k}) = \frac{w_k\sigma_B^2\left(B_k^2\kappa_{\lambda,k}^2 + C_k\right) - A_k\left(B_k w_k\kappa_{\lambda,k}^2 + \hat{\lambda}_{k|k}C_k\right)}{C_k\sqrt{2\pi\left(B_k^2\kappa_{\lambda,k}^2 + C_k\right)^3}} \cdot \exp\left(-\frac{\left(w_k - B_k\hat{\lambda}_{k|k}\right)^2}{2\left(B_k^2\kappa_{\lambda,k}^2 + C_k\right)}\right)\frac{\mathrm{d}\zeta_{l_k}}{\mathrm{d}l_k} \tag{18}$$

$$F_{L_k|\boldsymbol{Y}_{1:k}}(l_k|\boldsymbol{Y}_{0:k}) = 1 - \Phi\left(\frac{w - \hat{x}_{k|k} - \zeta_{l_k}\hat{\lambda}_{k|k}}{\sqrt{C_k + B_k^2\kappa_{\lambda,k}^2}}\right) + \frac{1}{\sqrt{1 - 2B_{1,k}\kappa_{\lambda,k}^2}} \cdot \exp\left(\frac{2A_{1,k}\hat{\lambda}_{k|k} + A_{1,k}^2\kappa_{\lambda,k}^2 - 2B_{1,k}\hat{\lambda}_{k|k}^2}{2\left(1 - 2B_{1,k}\kappa_{\lambda,k}^2\right)}\right) \cdot \Phi\left(\frac{C_{1,k} + D_k\hat{\lambda}_{k|k} + A_{1,k}D_k\kappa_{\lambda,k}^2 - 2B_{1,k}C_{1,k}\kappa_{\lambda,k}^2}{\sqrt{\left(1 - 2B_{1,k}\kappa_{\lambda,k}^2\right)^2 + D_k^2\kappa_{\lambda,k}^2\left(1 - 2B_{1,k}\kappa_{\lambda,k}^2\right)}}\right) \tag{19}$$

式中: $w_k = w - \hat{x}_{k|k} + \rho_k\frac{\kappa_{x,k}}{\kappa_{\lambda,k}}\hat{\lambda}_{k|k}$; $A_k = \sigma_B^2\rho_k\frac{\kappa_{x,k}}{\kappa_{\lambda,k}} - \sigma_{x_k|\lambda,k}^2$; $B_k = \zeta_{l_k} + \rho_k\frac{\kappa_{x,k}}{\kappa_{\lambda,k}}$; $C_k = \kappa_{x,k}^2\left(1 - \rho_k^2\right) + \sigma_B^2\zeta_{l_k}$; $A_{1,k} = \frac{2\left(w - \hat{x}_{k|k} + \rho_k\frac{\kappa_{x,k}}{\kappa_{\lambda,k}}\hat{\lambda}_{k|k}\right)}{\sigma_B^2}$; $B_{1,k} = \frac{2\left(w - \sigma_{x_k|\lambda,k}^2 - \rho_k\frac{\kappa_{x,k}}{\kappa_{\lambda,k}}\sigma_B^2\right)}{\sigma_B^4}$; $C_{1,k} = \frac{-w + \hat{x}_{k|k} - \rho_k\frac{\kappa_{x,k}}{\kappa_{\lambda,k}}\hat{\lambda}_{k|k}}{\sqrt{\sigma_B^2\zeta_{l_k} + \sigma_{x_k|\lambda,k}^2}}$; $D_k = \frac{\rho_k\frac{\kappa_{x,k}}{\kappa_{\lambda,k}} - \zeta_{l_k} - \frac{2}{\sigma_B^2}\sigma_{x_k|\lambda,k}^2}{\sqrt{\sigma_B^2\zeta_{l_k} + \sigma_{x_k|\lambda,k}^2}}$。

具体推导过程见附录 A。

2　离线参数估计

模型参数表征同类设备的整体退化特征, 模型参数离线估计的准确性对后验参数的在线更新至

关重要。本节提出针对隐含尺度变换非线性维纳过程模型的无偏参数估计方法。由式(2)可知,需要估计的模型未知参数为$\boldsymbol{\Theta}=\{u_\lambda,\sigma_\lambda^2,\sigma_B^2,\theta,\sigma_\varepsilon^2\}$。

2.1 无偏参数估计方法

假设共有n个测试设备,每个设备在相同的时刻$t_1,t_2,\cdots,t_m$测量退化状态,则第i个设备在t_j时刻的退化量可以表示为

$$y_{i,j}=Y_i(t_j)=\lambda_i\Lambda(t_j;\theta)+\sigma_B B\left(\Lambda(t_j;\theta)\right)+\varepsilon_{i,j} \tag{20}$$

式中:$1\leqslant i\leqslant n$,$1\leqslant j\leqslant m$,λ_i表示第i个设备的退化速率。

令$\Delta y_{i,j}=Y_i(t_j)-Y_i(t_{j-1})$,则$\Delta y_{i,j}$可表示为

$$\Delta y_{i,j}=Y_i(t_j)-Y_i(t_{j-1})=\lambda_i\Delta\nu_j+\sigma_B B(\Delta\nu_j)+\varepsilon_{i,j}-\varepsilon_{i,j-1} \tag{21}$$

式中:$\Delta\nu_j=\Lambda(t_j;\theta)-\Lambda(t_{j-1};\theta)$。令$\Delta\boldsymbol{\nu}=(\Delta\nu_1,\Delta\nu_2,\cdots,\Delta\nu_m)'$,$\Delta\boldsymbol{t}_j=(\Delta t_1,\Delta t_2,\cdots,\Delta t_m)'$,$\Delta\boldsymbol{y}_i=(\Delta y_{i,1},\Delta y_{i,2},\cdots,\Delta y_{i,m})'$和$\boldsymbol{Y}=(\Delta\boldsymbol{y}_1',\Delta\boldsymbol{y}_2',\cdots,\Delta\boldsymbol{y}_n')'$,则有$\Delta\boldsymbol{y}_i\sim N(u_\lambda\Delta\boldsymbol{\nu},\boldsymbol{\Sigma})$,$\boldsymbol{\Sigma}=\sigma_\lambda^2\Delta\boldsymbol{\nu}\Delta\boldsymbol{\nu}'+\sigma_B^2\boldsymbol{\Omega}+\sigma_\varepsilon^2\boldsymbol{P}$。此处,$\boldsymbol{\Omega}$和$\boldsymbol{P}$都为$m$阶矩阵,$\boldsymbol{\Omega}=\mathrm{diag}(\Delta\nu_1,\Delta\nu_2,\cdots,\Delta\nu_m)$,

$$\boldsymbol{P}=\begin{bmatrix}1&-1&&&\\-1&2&-1&&\\&-1&2&\ddots&\\&&\ddots&\ddots&-1\\&&&-1&2\end{bmatrix}_{m\times m}。$$

Tang等[26]给出基本线性、非线性和隐含线性维纳过程模型的无偏参数估计方法,Han等[27]进一步给出隐含非线性维纳过程模型的无偏参数估计方法,根据文献[26-27],令$\boldsymbol{D}=\boldsymbol{\Omega}+\phi\boldsymbol{P}$,$\phi=\sigma_\varepsilon^2/\sigma_B^2$,则可以得到隐含尺度变换维纳过程模型的离线参数的无偏估计表达式为

$$\hat{\mu}_\lambda\left(\hat{\theta},\hat{\phi}\right)=\frac{1}{n}\sum_{i=1}^{n}\frac{\Delta\boldsymbol{y}'\boldsymbol{D}^{-1}\Delta\boldsymbol{\nu}}{\Delta\boldsymbol{\nu}'\boldsymbol{D}^{-1}\Delta\boldsymbol{\nu}} \tag{22}$$

$$\hat{\sigma}_\lambda^2\left(\hat{\theta},\hat{\phi}\right)=\frac{1}{n-1}\sum_{i=1}^{n}\left(\mu_\lambda-\hat{\lambda}_i\right)^2-\frac{1}{n}\sum_{i=1}^{n}\frac{(\Delta\boldsymbol{y}_i-\hat{\lambda}_i\Delta\boldsymbol{\nu})'\boldsymbol{D}^{-1}(\Delta\boldsymbol{y}_i-\hat{\lambda}_i\Delta\boldsymbol{\nu})}{(m-1)\Delta\boldsymbol{\nu}'\boldsymbol{D}^{-1}\Delta\boldsymbol{\nu}} \tag{23}$$

$$\hat{\sigma}_B^2\left(\hat{\theta},\hat{\phi}\right)=\frac{1}{n}\sum_{i=1}^{n}\frac{1}{m-1}\left(\Delta\boldsymbol{y}_i-\hat{\lambda}_i\Delta\boldsymbol{\nu}\right)'\boldsymbol{D}^{-1}\left(\Delta\boldsymbol{y}_i-\hat{\lambda}_i\Delta\boldsymbol{\nu}\right) \tag{24}$$

式中:$\hat{\lambda}_i=\Delta\boldsymbol{y}_i'\boldsymbol{D}^{-1}\Delta\boldsymbol{\nu}/(\Delta\boldsymbol{\nu}'\boldsymbol{D}^{-1}\Delta\boldsymbol{\nu})$。隐含尺度变换维纳过程模型的参数估计推导过程与一般维纳过程模型相同,具体推导过程可参考文献[26-27]。

通过MATLAB函数“FMINSEARCH”最大化式(25),可以得出参数θ和ϕ的估计值,将其代入式(22)~式(24),可得u_λ、σ_λ^2和σ_B^2的值,进一步可得到参数σ_ε^2。

$$\ln L(\theta,\phi|\boldsymbol{Y})=-\sum_{i=1}^{n}\frac{m}{2}\ln(2\pi)-\sum_{i=1}^{n}\ln|\boldsymbol{D}|-\sum_{i=1}^{n}\frac{m}{2}-\sum_{i=1}^{n}\frac{m-1}{2}\ln\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\left(\Delta\boldsymbol{y}_i-\hat{\lambda}_i\Delta\boldsymbol{\nu}\right)'\boldsymbol{D}^{-1}\left(\Delta\boldsymbol{y}_i-\hat{\lambda}_i\Delta\boldsymbol{\nu}\right)}{m-1}\right)-\sum_{i=1}^{n}\frac{1}{2}\ln\left(\frac{1}{n}\sum_{i=1}^{n}\Delta\boldsymbol{\nu}'\boldsymbol{D}^{-1}\Delta\boldsymbol{\nu}\left(\hat{\mu}_\lambda-\frac{\Delta\boldsymbol{y}'_i\boldsymbol{D}^{-1}\Delta\boldsymbol{\nu}}{\Delta\boldsymbol{\nu}'\boldsymbol{D}^{-1}\Delta\boldsymbol{\nu}}\right)^2\right) \tag{25}$$

2.2 经验无偏参数估计方法

在工程应用中,由于设备之间存在个体差异,每个设备的检测时刻和次数可能不同。例如,本文使用的涡扇发动机退化数据每组检测次数就不同。对于上述的退化数据,常用的参数估计方法无法直接得到漂移系数方差和扩散系数的剖面解析表达式。因此,本文针对隐含尺度变换非线性维纳退化过程模型提出一种经验无偏参数估计方法。

假设共有n个测试设备,第i个设备的测量时间为$t_{i,1},t_{i,2},\cdots,t_{i,m_i}$,其中$1\leqslant i\leqslant n$,$1\leqslant j\leqslant m_i$,$m_i$表示第$i$个设备的测量次数。令第$i$个设备的时间间隔为$\Delta\boldsymbol{\nu}_i=(\Delta\nu_{i,1},\Delta\nu_{i,2},\cdots,\Delta\nu_{i,m_i})'$,其中$\Delta\nu_{i,j}=\Lambda(t_{i,j};\theta)-\Lambda(t_{i,j-1};\theta)$,则将式(22)~式(24)中的$m$和$\Delta\boldsymbol{\nu}$替换为$m_i$和$\Delta\boldsymbol{\nu}_i$可得经验无偏参数估计表达式,具体如下:

$$\hat{\mu}_\lambda\left(\hat{\theta},\hat{\phi}\right)=\frac{1}{n}\sum_{i=1}^{n}\frac{\Delta\boldsymbol{y}'_i\boldsymbol{D}_i^{-1}\Delta\boldsymbol{\nu}_i}{\Delta\boldsymbol{\nu}'_i\boldsymbol{D}_i^{-1}\Delta\boldsymbol{\nu}_i} \tag{26}$$

$$\hat{\sigma}_\lambda^2\left(\hat{\theta},\hat{\phi}\right)=\frac{1}{n-1}\sum_{i=1}^{n}\left(\mu_\lambda-\hat{\lambda}_i\right)^2-\frac{1}{n}\sum_{i=1}^{n}\frac{\left(\Delta\boldsymbol{y}_i-\hat{\lambda}_i\Delta\boldsymbol{\nu}_i\right)'\boldsymbol{D}_i^{-1}\left(\Delta\boldsymbol{y}_i-\hat{\lambda}_i\Delta\boldsymbol{\nu}_i\right)}{(m_i-1)\Delta\boldsymbol{\nu}'_i\boldsymbol{D}_i^{-1}\Delta\boldsymbol{\nu}_i} \tag{27}$$

$$\hat{\sigma}_B^2\left(\hat{\theta},\hat{\phi}\right)=\frac{1}{n}\sum_{i=1}^{n}\frac{1}{m_i-1}\left(\Delta\boldsymbol{y}_i-\hat{\lambda}_i\Delta\boldsymbol{\nu}_i\right)'\boldsymbol{D}_i^{-1}\left(\Delta\boldsymbol{y}_i-\hat{\lambda}_i\Delta\boldsymbol{\nu}_i\right) \tag{28}$$

式中:$\hat{\lambda}_i=\Delta\boldsymbol{y}'_i\boldsymbol{D}_i^{-1}\Delta\boldsymbol{\nu}_i/\left(\Delta\boldsymbol{\nu}'_i\boldsymbol{D}_i^{-1}\Delta\boldsymbol{\nu}_i\right)$。其他参数估计流程可参考2.1节所述的等间隔情况下的参数估计方法。

3 实 验

本节首先使用仿真方法模拟非线性退化过程,分析设备在以式(2)为退化路径情况下不同模型的预测效果。然后,通过实际涡扇发动机数据来验证式(2)的有效性和实用性。为简便起见,定义式(2)同时考虑非线性和测量误差的维纳过程模型为M_0,式(1)仅考虑非线性的维纳过程模型为M_1来研究不考虑测量误差对预测精度的影响。此外,定义将式(2)中非线性系数设成$\theta=1$的线性模型为M_2

来研究不考虑非线性对预测精度的影响。为更直观地展示 3 种模型预测效果的差异, 进一步计算各模型在每一个测量点的均方误差(mean squared error, MSE)。如下:

$$M_{\mathrm{SE}}=\int_0^{\infty}(l_k+t_k-T)^2 f_{L_k|\boldsymbol{Y}_{0:k}}(l_k|\boldsymbol{Y}_{0:k})\mathrm{d}l_k \tag{29}$$

式中: M_{SE}为均方误差; T为被测设备的实际失效时间; l_k为在t_k时刻的剩余寿命; $f_{L_k|\boldsymbol{Y}_{0:k}}(l_k|\boldsymbol{Y}_{0:k})$为在$t_k$时刻估计的剩余寿命 PDF。

3.1　仿真实验

假设模型参数$u_\lambda=1$, $\sigma_\lambda^2=0.062\,5$, $\sigma_B^2=0.49$, $\theta=1.5$, $\sigma_\varepsilon^2=9$, 样本组数$N=11$, 测量时间间隔$\Delta t=1$ h。仿真退化路径如图 1 所示。

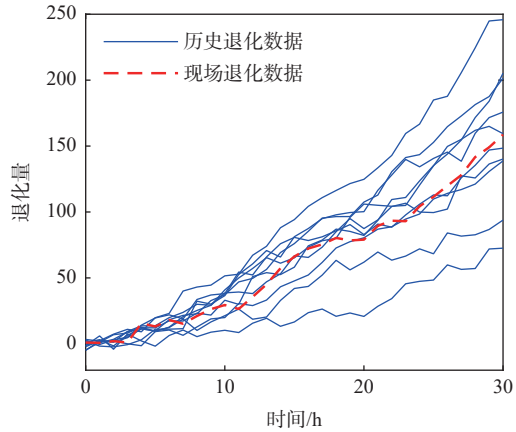


图 1　仿真退化路径

Fig. 1　Simulated degradation path

上述提到的 3 种退化模型, 基于历史退化数据, 使用无偏参数估计方法获得的离线参数估计结果如表 1 所示。其中参数u_λ、σ_λ^2、σ_B^2、θ、σ_ε^2的真实值分别为 1、0.062 5、0.49、1.5、9。从表 1 可知, M_0估计的离线参数更接近真实值, M_1次之, M_2最差。通常表示模型不确定性的σ_B^2参数数值越小, 预测的结果不确定性越小。表 1 中, 当不考虑测量误差影响时, M_1估计的σ_λ^2和σ_B^2分别是M_0估计值的 1.4 倍和 2.5 倍。当不考虑非线性时, σ_λ^2和σ_B^2分别是M_0估计值的 15 倍和 6.8 倍。这说明不考虑非线性和测量误差都会使预测结果偏离真实值, 且一般情况下, 前者对预测精度影响更大。因此, M_0模型具有更好的预测效果。

表 1　仿真数据离线参数估计

Table 1　Offline parameter estimation of simulation data

模型	u_λ	σ_λ^2	σ_B^2	θ	σ_ε^2
M_0	1.242 6	0.132 34	4.483 2	1.426 2	7.480 2
M_1	1.702 92	0.193 597	11.299 1	1.333 12	0
M_2	5.285 21	2.013 32	30.492 4	1	3.355 57

在获取离线参数估计后, 本文使用 Kalman 滤波方法在线更新模型后验参数。图 2 为 3 种模型在不同时刻的漂移系数变化曲线。由图 2 可知, 在整个退化周期内, M_0的漂移系数始终比其他 2 种模型更接近真实值。除此之外, M_0和M_1的漂移系数具有相似的变化趋势, 逐渐向设定的模型参数值接近, M_2则是波动大, 偏离真实值较远, 这进一步说明了M_0的预测效果优势明显。

图 3 为设备在使用一段时间后几个典型监测点剩余寿命预测的 PDF 分布。为方便比较 3 种模型之间的差异, 画出被测设备在第 26 h 的剩余寿命分布曲线, 如图 4 所示。从图 3 和图 4 可以看出,

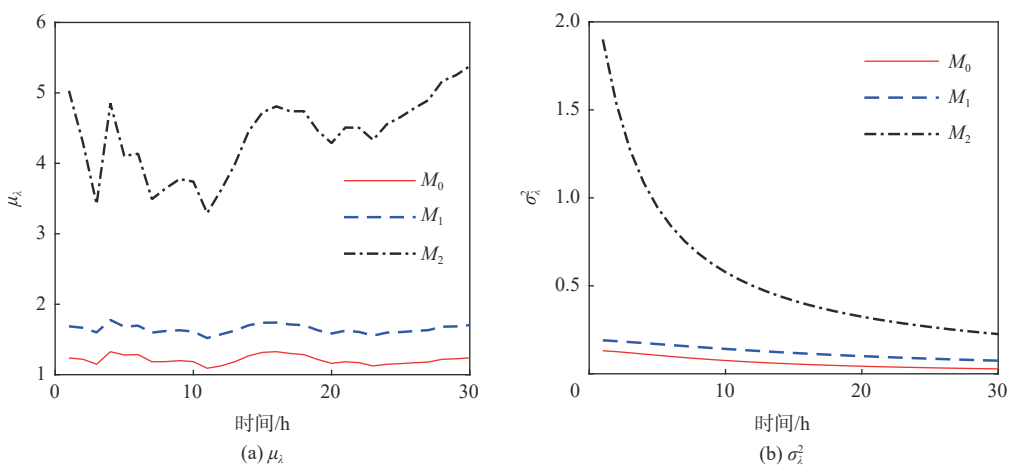


图 2　M_0、M_1和M_2在不同时刻的漂移系数估计

Fig. 2　Estimation of drift coefficients of M_0, M_1 and M_2 at different times

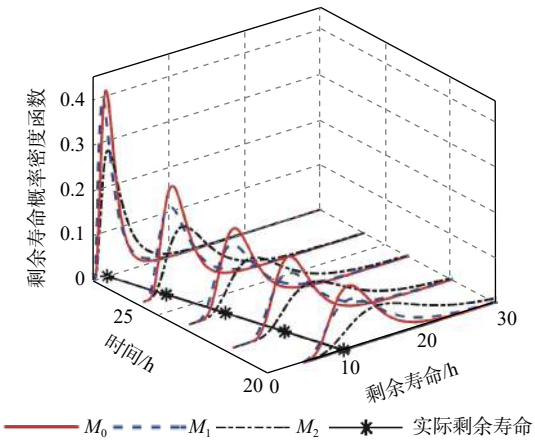


图 3 估计的剩余寿命分布

Fig. 3 Estimated distribution of remaining useful life

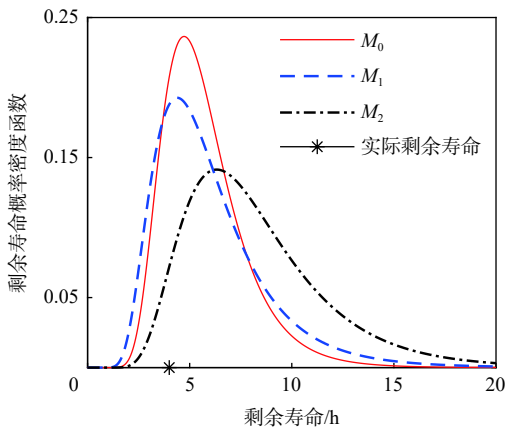


图 4 26 h 的剩余寿命分布

Fig. 4 Remaining useful life distribution at 26 h

3 种模型剩余寿命分布曲线都能够覆盖真实剩余寿命,随着检测的被测设备现场退化数据越多,分布曲线越集中于真实剩余寿命值,且分布图形M_0始终高于M_1和M_2,M_1高于M_2,这不仅说明了剩余寿命的预测精度与获取的现场退化数据量成正比,同时也说明了M_0估计的剩余寿命不确定性最小,预测精度最高,M_1次之。

图 5 为估计的剩余寿命 MSE,由图可知,M_0与M_1的 MSE 曲线具有相同的变化趋势,且M_1曲线高于M_0。这是因为在测量误差相对较小的情况下,测量误差虽然会增大预测的不确定性,导致M_1的 MSE 偏大,但是其对模型跟踪非线性退化路径的效果影响较小。同样的,上述漂移系数的在线更新曲线也解释了该问题。然而,M_2的 MSE 在整个退化周期均远大于M_0和M_1。造成该现象的原因是线性模型难以实现对非线性退化路径的有效跟踪。因此,在进行剩余寿命预测时,针对退化系统选用合适的退化模型十分重要。

注 1 与上述步骤相同,分别以M_1和M_2的模型为退化路径生成仿真数据,并使用M_0、M_1和M_2模型进行 RUL 预测,可以得到结论:当非线性退化系统的测量误差非常小或不存在测量误差时,M_0和M_1的预测效果十分接近,优于M_2。当带测量误差的退化系统的非线性退化特征不明显或退化系统为线性退化特征时,M_0和M_2的预测效果十分接近,优于M_1。总的来说,M_0的灵活性更强,能够适用于更多的退化系统。

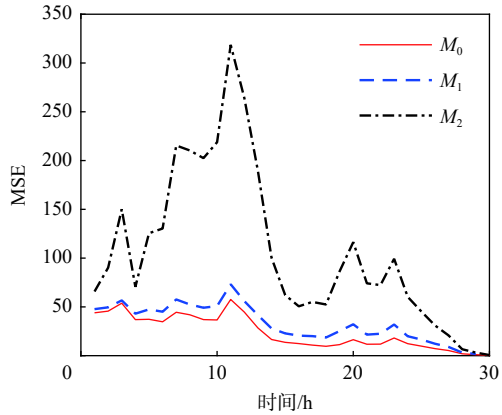


图 5 估计的剩余寿命 MSE

Fig. 5 MSE of estimated remaining useful life

3.2 实例分析

为进一步验证本文模型的实用性,本文使用美国航空航天局发布的涡扇发动机数据集,该数据集由 218 个相同且独立的单元退化数据组成。其中,每个单元均从初始运行到失效,且有 21 个传感器检测不同的退化特征,详细的数据介绍参考文献 [29]。由于 21 个传感器退化数据中,一些传感器没有明显的退化趋势,如果将其用于 RUL 预测分析,可能会降低预测的准确性,根据 Le 等[29] 建议,使用可用传感器的子集有助于提高 RUL 估计的准确性,因此,本文选用 7 个退化特征比较明显的传感器采集的退化数据,并使用文献 [29] 中提出的主成分分析方法处理数据。为便于说明本文模型的预测效果且不失一般性,选用数据集中的 11 组~20 组退化数据,并对数据初始值进行简单处理(该步骤不改变原有退化特性),退化路径如图 6 所示。

从图 6 中可以看出,发动机退化数据波动大,具有明显的非线性和不确定测量特征。本文定义数据的前 9 组为历史退化数据,第 10 组数据为被测设备现场退化数据,使用无偏参数估计方法估计模型参数。由于直接使用该估计方法,M_1和M_2的离线参数估计结果中$\sigma_\lambda^2<0$,不符合实际。因此,使用 Tang 等[24] 提出的 2 步 MLE 方法估计M_1和M_2的模型未知参数,以保证σ_λ^2不为负。3 种模型参数估计结果如表 2 所示。从表 2 中可以看出,M_0和M_2

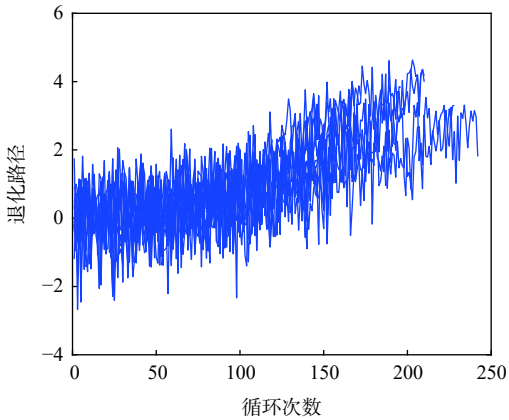


图 6　涡扇发动机部分单元退化路径

Fig. 6　Partial unit degradation path of turbofan engine

表 2　涡扇发动机数据离线参数估计

Table 2　Offline parameter estimation of turbofan engine data

模型	u_λ	σ_λ^2	σ_B^2	θ	σ_ε^2
M_0	$4.491\,6\times10^{-6}$	$8.624\,4\times10^{-13}$	$2.941\,3\times10^{-6}$	2.589 5	0.589 19
M_1	0.038 712	0.000 238 19	2.54	0.836 98	0
M_2	0.016 695	$3.266\,8\times10^{-5}$	0.011 918	1	0.535 02

的模型参数σ_ε接近整个周期退化量的 20%, 表明了数据测量误差相对较大。除此之外, M_0的σ_λ^2和σ_B^2均小于其他 2 种模型, 这说明M_0不确定性最小, 预测效果最好。值得注意的是, 由于数据测量误差相对较大, M_1在估计参数时未考虑测量误差的影响, 于是误差的不确定性被扩散系数所淹没, 导致了σ_B^2偏大和非线性系数$\theta<1$。

图 7 为 3 种模型在一些不同时刻的剩余寿命分布。为了更清晰地比较 3 种模型的预测效果, 分别画出被测设备在第 130 次和第 140 次循环的剩余寿命分布图, 如图 8 所示。从图 7 和图 8 可以看出, M_0和M_2的剩余寿命分布均覆盖了实际剩余寿命, 且M_0的剩余寿命分布曲线比M_2高, 围绕实际剩余寿命值更加紧凑。随着检测到的被测设备现场退化数据越多, 2 个模型剩余寿命的 PDF 曲线逐渐变尖。然而, M_1的剩余寿命分布曲线虽然比M_0和M_2的更高, 但其在整个退化周期内始终偏向于一侧, 不能很好地覆盖实际剩余寿命值。这是由于M_1未考虑测量误差的影响, 导致离线参数估计的结果没能有效地跟踪退化路径。同时, σ_B^2和θ作为固定参数, 不随检测数据的增加而更新, 这使得M_1在整个周期内始终按照错误的退化特征进行剩余寿命预测。因此, M_1的预测效果比其他 2 种模型差。

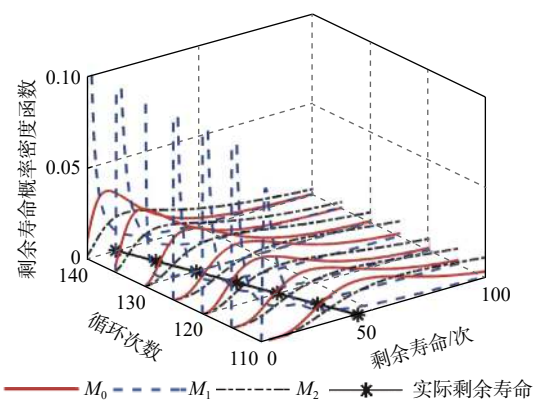
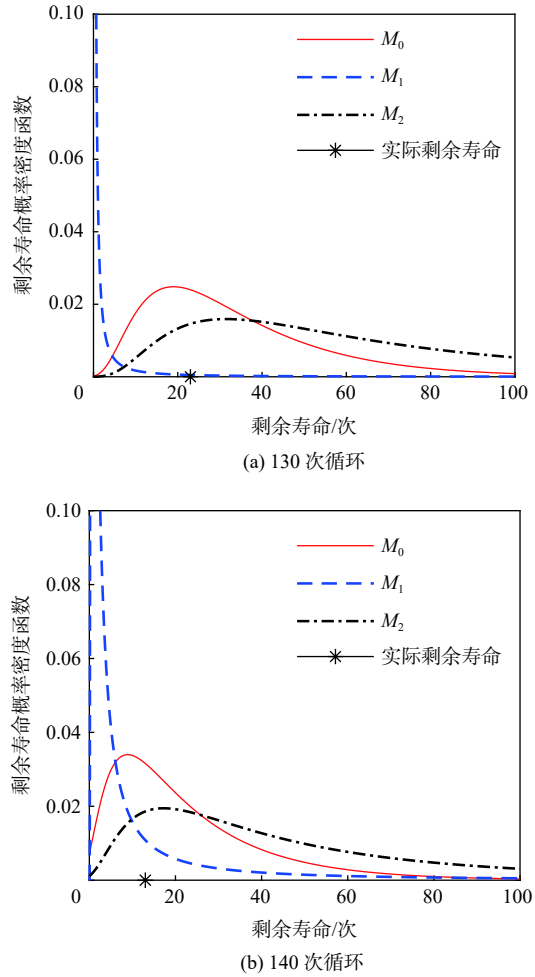


图 7　不同循环次数的估计剩余寿命分布

Fig. 7　Estimated remaining useful life distribution at different cycles



(a) 130 次循环

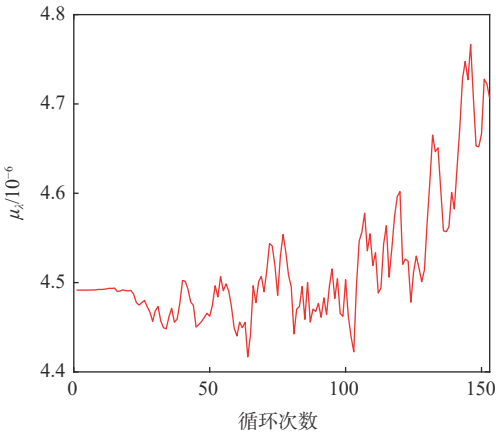
(b) 140 次循环

图 8　不同循环次数的寿命分布

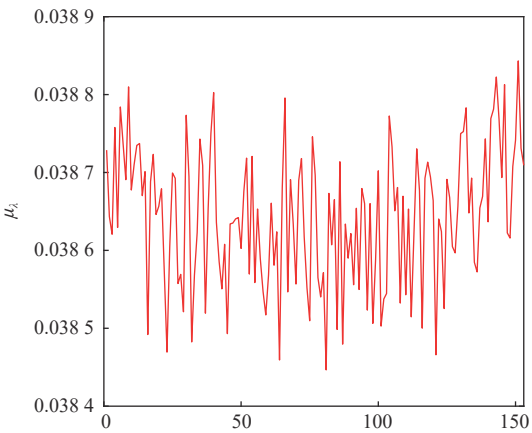
Fig. 8　Remaining useful life distributions at different cycles

图 9 为 3 种模型的退化速率均值u_λ在整个周期在线更新的结果。从图 9 中可以看出, M_1的u_λ曲线波动最大。这主要是因为M_1没有考虑测量误差, 无法对现场退化数据进行实时滤波, 始终追踪着带误差的退化轨迹。相反, M_0和M_2得到了经过滤波的现场退化数据, u_λ曲线较为平稳。

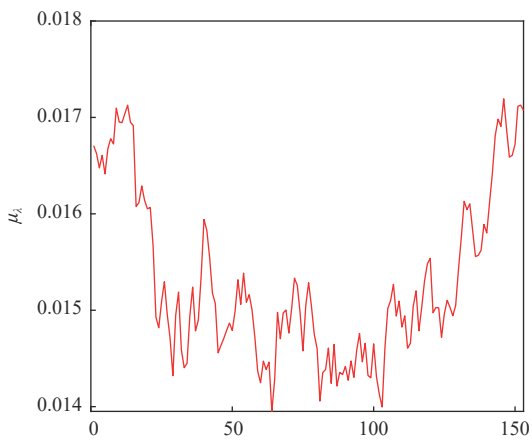
从 MSE 的角度比较 3 种模型的预测精度, 如图 10 所示。图中, M_0的 MSE 在整个退化过程中均小于M_1和M_2, M_1的平均 MSE 分别是M_0和



(a) M_0



(b) M_1



(c) M_2

图 9 M_0、M_1和M_2在不同时刻的u_λ后验参数估计

Fig. 9 Posteriori parameter estimation of u_λ for M_0, M_1 and M_2 at different times

M_2的 8 倍和 3.8 倍，说明M_0的预测精最高，M_1最差。值得注意的是，M_1的 MSE 曲线不仅高于其他 2 种模型，而且存在较大波动。原因是M_1在线更新后的λ波动大，进而导致预测的剩余寿命也存在较大波动。总的来说，对带有测量误差的非线性退化系统，本文使用的模型能够有效地预测剩余寿命。

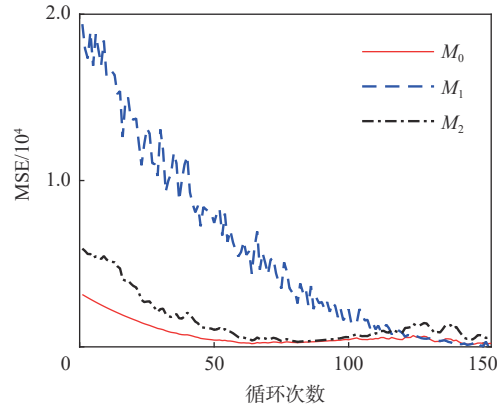


图 10 不同循环次数的估计剩余寿命 MSE

Fig. 10 MSE of estimated remaining useful life at different cycles

4 结 论

1）本文针对同时具有非线性和不确定测量的随机退化设备，使用隐含尺度变换的维纳退化过程模型进行剩余寿命预测。实验结果表明，同时考虑测量不确定性和非线性退化特征的剩余寿命预测方法具有更高的预测精度。

2）由于设备的固有性能和测量工具的退化等因素，测量误差可能会随着退化时间呈现一定规律的变化。后续，将继续深入研究测量误差与退化时间之间的联系，进一步提高设备的剩余寿命预测精度。

参考文献（References）

[1] 李天梅, 司小胜, 刘翔, 等. 大数据下数模联动的随机退化设备剩余寿命预测技术[J]. 自动化学报, 2022, 48(9): 2119-2141.
LI T M, SI X S, LIU X, et al. Data-model interactive remaining useful life prediction technologies for stochastic degrading devices with big data[J]. Acta Automatica Sinica, 2022, 48(9): 2119-2141(in Chinese) .

[2] 王凤飞, 唐圣金, 孙晓艳, 等. 考虑随机效应的多源信息融合剩余寿命预测 [J/OL]. 北京航空航天大学学报, 2022, (2021-03-01) [2022-04-03]. https://link.cnki.net/urlid/11.2625.V.20220228.1906.003.html. DOI:10.13700/j.bh.1001-5965.2021.0782.
WANG F F, TANG S J, SUN X Y, et al. Remaining useful life prediction based on multi source information with considering random effects [J/OL]. Journal of Beijing University of Aeronautics and Astronautics, 2022, (2021-03-01)[2022-04-03]. https://link.cnki.net/urlid/11.2625.V.20220228.1906.003.html. DOI:10.13700/j.bh.1001-5965.2021.0782(in Chinese) .

[3] WANG Z, CHEN Y, CAI Z, et al. Methods for predicting the remaining useful life of equipment in consideration of the random failure threshold[J]. Journal of Systems Engineering and Electronics, 2020, 31(2): 415-431.

[4] 郑建飞, 胡昌华, 司小胜, 等. 考虑不确定测量和个体差异的非线性随机退化系统剩余寿命估计[J]. 自动化学报, 2017, 43(2): 259-270.
ZHENG J F, HU C H, SI X S, et al. Remaining useful life estima-

tion for nonlinear stochastic degrading systems with uncertain measurement and unit-to-unit variability[J]. Acta Automatica Sinica, 2017, 43(2): 259-270(in Chinese).

[5] ZHANG Z, SI X, HU C, et al. Degradation data analysis and remaining useful life estimation: A review on Wiener-process-based methods[J]. European Journal of Operational Research, 2018, 271(3): 775-796.

[6] 万昌豪, 刘志国, 唐圣金, 等. 基于不完美先验信息的随机系数回归模型剩余寿命预测方法[J]. 北京航空航天大学学报, 2021, 47(12): 2542-2551.
WAN C H, LIU Z G, TANG S J, et al. Remaining useful life prediction method based on the nonlinear random coefficient regression model with imperfect prior information[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(12): 2542-2551 (in Chinese).

[7] WANG D, YANG F, TSUI K L, et al. Remaining useful life prediction of lithium-ion batteries based on spherical cubature particle filter[J]. IEEE Transactions on Instrumentation and Measurement, 2016, 65(6): 1282-1291.

[8] SI X S, WANG W, HU C H, et al. Remaining useful life estimation – A review on the statistical data driven approaches[J]. European Journal of Operational Research, 2011, 213(1): 1-14.

[9] SI X S, WANG W, HU C H, et al. A Wiener-process-based degradation model with a recursive filter algorithm for remaining useful life estimation[J]. Mechanical Systems and Signal Processing, 2013, 35(1-2): 219-237.

[10] LIN C P, LING M H, CABRERA J, et al. Prognostics for lithium-ion batteries using a two-phase gamma degradation process model[J]. Reliability Engineering & System Safety, 2021, 214: 107797.

[11] HAO S, YANG J, BERENGUER C. Degradation analysis based on an extended inverse Gaussian process model with skew-normal random effects and measurement errors[J]. Reliability Engineering & System Safety, 2019, 189: 261-270.

[12] TANG S, XU X, YU C, et al. Remaining useful life prediction with fusing failure time data and field degradation data with random effects[J]. IEEE Access, 2020, 8: 11964-11978.

[13] LIU D, WANG S, ZHANG C. Reliability estimation by fusing multiple-source information based on evidential variable and Wiener process[J]. Computers & Industrial Engineering, 2021, 162: 107745.

[14] WHITMORE G A. Estimating degradation by a Wiener diffusion process subject to measurement error[J]. Lifetime Data Analysis, 1995, 1(3): 307-319.

[15] PENG C Y, TSENG S T. Mis-specification analysis of linear degradation models[J]. IEEE Transactions on Reliability, 2009, 58(3): 444-455.

[16] SI X S, WANG W, HU C H, et al. Estimating remaining useful life with three-source variability in degradation modeling[J]. IEEE Transactions on Reliability, 2014, 63(1): 167-190.

[17] FENG L, WANG H, SI X, et al. A state-space-based prognostic model for hidden and age-dependent nonlinear degradation process[J]. IEEE Transactions on Automation Science and Engineering, 2013, 10(4): 1072-1086.

[18] WHITMORE G A, SCHENKELBERG F. Modelling accelerated degradation data using wiener diffusion with a time scale transformation[J]. Lifetime Data Analysis, 1997, 3(1): 27-45.

[19] YE Z S, WANG Y, TSUI K L, et al. Degradation data analysis using Wiener processes with measurement errors[J]. IEEE Transactions on Reliability, 2013, 62(4): 772-780.

[20] PAN D, WEI Y, FANG H, et al. A reliability estimation approach via Wiener degradation model with measurement errors[J]. Applied Mathematics and Computation, 2018, 320: 131-141.

[21] PAN D, LU S, LIU Y, et al. Degradation data analysis using a Wiener degradation model with three-source uncertainties[J]. IEEE Access, 2019, 7: 37896-37907.

[22] YE Z, CHEN N, TSUI K L. A Bayesian approach to condition monitoring with imperfect inspections[J]. Quality and Reliability Engineering International, 2015, 31(3): 513-522.

[23] HUANG Z, XU Z, WANG W, et al. Remaining useful life prediction for a nonlinear heterogeneous Wiener process model with an adaptive drift[J]. IEEE Transactions on Reliability, 2015, 64(2): 687-700.

[24] TANG S, YU C, WANG X, et al. Remaining useful life prediction of lithium-ion batteries based on the Wiener process with measurement error[J]. Energies, 2014, 7(2): 520-547.

[25] 唐圣金, 郭晓松, 周召发, 等. 步进应力加速退化试验的建模与剩余寿命估计[J]. 机械工程学报, 2014, 50(16): 33-40.
TANG S J, GUO X S, ZHOU Z F, et al. Step stress accelerated degradation process modeling and remaining useful life estimation[J]. Journal of Mechanical Engineering, 2014, 50(16): 33-40 (in Chinese) .

[26] TANG S, WANG F, SUN X, et al. Unbiased parameters estimation and mis-specification analysis of Wiener process-based degradation model with random effects[J]. Applied Mathematical Modelling, 2022, 109: 134-160.

[27] HAN Y, MA C, TANG S, et al. Residual life estimation of lithium-ion batteries based on nonlinear Wiener process with measurement error[J]. Journal of Risk and Reliability, 2023, 237(1): 133-151.

[28] SI X S, WANG W, CHEN M Y, et al. A degradation path-dependent approach for remaining useful life estimation with an exact and closed-form solution[J]. European Journal of Operational Research, 2013, 226(1): 53-66.

[29] LE SON K, FOULADIRAD M, BARROS A, et al. Remaining useful life estimation based on stochastic deterioration models: A comparative study[J]. Reliability Engineering & System Safety, 2013, 112: 165-175.

附录 A

式 (14) 和式 (15) 假设被测设备的真实退化状态 x_k 和漂移系数 λ_k 已知, 然而, 对于带测量误差的退化系统, x_k 和 λ_k 为隐含状态, 因此, 式(14)和式(15)不能直接用于 RUL 预测。考虑到使用 Kalman 滤波方法在线更新模型参数可以得到 x_k 和 λ_k 的估计状态, 于是将其代入式 (14) 和式 (15) 可以进一步得到隐含非线性维纳退化过程模型考虑模型参数在线更新的 RUL 的 PDF 和 CDF。

1) PDF 推导过程。

① 考虑到 $x_k\left|\lambda_k, \boldsymbol{Y}_{0:k} \sim N\left(u_{x_k\mid\lambda,k}, \sigma^2_{x_k\mid\lambda,k}\right)\right.$, 基于引理 1, 则式 (14) 可表示为

$$E_{x_k|\lambda_k,\boldsymbol{Y}_{0:k}}\left(f_{L_k|\lambda_k,x_k,\boldsymbol{Y}_{0:k}}(l_k|\lambda_k,x_k,\boldsymbol{Y}_{0:k})\right)=$$
$$E_{x_k|\lambda_k,\boldsymbol{Y}_{0:k}}\left[(w_1-AZ)\exp\left(-\frac{(w_2-BZ)^2}{2G}\right)\frac{\mathrm{d}\zeta_{l_k}}{\mathrm{d}l_k}\right]=$$
$$\sqrt{\frac{\sigma_B^2\zeta_{l_k}}{\sigma_{x_k|\lambda,k}^2+G}}\left(w_1-A\frac{w_2\sigma_{x_k|\lambda,k}^2+\mu_{x_k|\lambda,k}G}{\sigma_{x_k|\lambda,k}^2+G}\right)\cdot$$
$$\exp\left(-\frac{(w_2-\mu_{x_k|\lambda,k})^2}{2(\sigma_{x_k|\lambda,k}^2+G)}\right)\frac{\mathrm{d}\zeta_{l_k}}{\mathrm{d}l_k} \tag{A1}$$

式中：$w_1=\dfrac{w}{\sqrt{2\pi\sigma_B^2{\zeta_{l_k}}^3}}$；$w_2=w-\lambda_k\zeta_{l_k}$；$A=\dfrac{1}{\sqrt{2\pi\sigma_B^2{\zeta_{l_k}}^3}}$；$B=1$；$G=\sigma_B^2\zeta_{l_k}$。

将$\mu_{x_k|\lambda,k}=\hat{x}_{k|k}+\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}\left(\lambda_k-\hat{\lambda}_{k|k}\right)$代入可转化为式(A2)。

② 对于$\lambda_k\left|\boldsymbol{Y}_{0:k}\sim N\left(\hat{\lambda}_{k|k},\kappa_{\theta,k}^2\right)\right.$，基于引理1，式(14)可以进一步表示为式(A3)。

PDF推导完毕。

$$f_{L_k|\boldsymbol{Y}_{0:k}}(l_k|\boldsymbol{Y}_{0:k})=E_{x_k|\lambda_k,\boldsymbol{Y}_{0:k}}\left(f_{L_k|\lambda_k,x_k,\boldsymbol{Y}_{0:k}}(l_k|\lambda_k,x_k,\boldsymbol{Y}_{0:k})\right)=\sqrt{\frac{\sigma_B^2\zeta_{l_k}}{\kappa_{x,k}^2\left(1-\rho_k^2\right)+G}}\cdot$$
$$\left[w_1-A\frac{w_2\kappa_{x,k}^2\left(1-\rho_k^2\right)+\left(\hat{x}_{k|k}+\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}\lambda_k-\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}\hat{\lambda}_{k|k}\right)G}{\kappa_{x,k}^2(1-\rho_k^2)+G}\right]\cdot\exp\left[-\frac{\left[w_2-\left(\hat{x}_{k|k}+\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}\lambda_k-\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}\hat{\lambda}_{k|k}\right)\right]^2}{2\left[\kappa_{x,k}^2\left(1-\rho_k^2\right)+G\right]}\right]\frac{\mathrm{d}\zeta_{l_k}}{\mathrm{d}l_k}=$$
$$\sqrt{\frac{\sigma_B^2\zeta_{l_k}}{\kappa_{x,k}^2\left(1-\rho_k^2\right)+G}}\left[w_1-A\frac{w\kappa_{x,k}^2+\hat{x}_{k|k}G-\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}\hat{\lambda}_{k|k}G-w\rho_k^2\kappa_{x,k}^2-\left(\zeta_{l_k}\kappa_{x,k}^2-\zeta_{l_k}\rho_k^2\kappa_{x,k}^2-\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}G\right)\lambda_k}{\kappa_{x,k}^2(1-\rho_k^2)+G}\right]\cdot$$
$$\exp\left(-\frac{\left[w-\hat{x}_{k|k}+\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}\hat{\lambda}_{k|k}-\left(\zeta_{l_k}+\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}\right)\lambda_k\right]^2}{2\left[\kappa_{x,k}^2(1-\rho_k^2)+G\right]}\right)\frac{\mathrm{d}\zeta_{l_k}}{\mathrm{d}l_k}=\sqrt{\frac{\sigma_B^2\zeta_{l_k}}{\kappa_{x,k}^2(1-\rho_k^2)+G}}\left(w_{1,k}-A_{1,k}\lambda_k\right)\cdot\exp\left(-\frac{(w_k-B_k\lambda_k)^2}{2C_k}\right)\frac{\mathrm{d}\zeta_{l_k}}{\mathrm{d}l_k} \tag{A2}$$

式中：$w_{1,k}=w_1-A\left[\dfrac{w\kappa_{x,k}^2+\hat{x}_{k|k}G-\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}\hat{\lambda}_{k|k}G-w\rho_k^2\kappa_{x,k}^2}{\kappa_{x,k}^2(1-\rho_k^2)+G}\right]$；$A_{1,k}=-\dfrac{A\left(\zeta_{l_k}\kappa_{x,k}^2-\zeta_{l_k}\rho_k^2\kappa_{x,k}^2-\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}G\right)}{\kappa_{x,k}^2(1-\rho_k^2)+G}$；$C_k=\kappa_{x,k}^2\left(1-\rho_k^2\right)+G$；

$w_k=w-\hat{x}_{k|k}+\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}\hat{\lambda}_{k|k}$；$B_k=\zeta_{l_k}+\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}$。

$$E_{\lambda_k|\boldsymbol{Y}_{0:k}}\left(f_{L_k|\boldsymbol{Y}_{0:k}}(l_k|\boldsymbol{Y}_{0:k})\right)=\sqrt{\frac{\sigma_B^2\zeta_{l_k}}{\kappa_{x,k}^2\left(1-\rho_k^2\right)+G}}\sqrt{\frac{C_k}{B_k^2\kappa_{\lambda,k}^2+C_k}}\left(w_{1,k}-A_{1,k}\frac{B_kw_k\kappa_{\lambda,k}^2+\hat{\lambda}_{k|k}C_k}{B_k^2\kappa_{\lambda,k}^2+C_k}\right)\cdot\exp\left(-\frac{(w_k-B_k\hat{\lambda}_{k|k})^2}{2(B_k^2\kappa_{\lambda,k}^2+C_k)}\right)\frac{\mathrm{d}\zeta_{l_k}}{\mathrm{d}l_k}=$$
$$\sqrt{\frac{\sigma_B^2\zeta_{l_k}}{B_k^2\kappa_{\lambda,k}^2+C_k}}\left(w_{1,k}-A_{1,k}\frac{B_kw_k\kappa_{\lambda,k}^2+\hat{\lambda}_{k|k}C_k}{B_k^2\kappa_{\lambda,k}^2+C_k}\right)\cdot\exp\left(-\frac{(w-\hat{x}_{k|k}-\zeta_{l_k}\hat{\lambda}_{k|k})^2}{2(B_k^2\kappa_{\lambda,k}^2+C_k)}\right)\frac{\mathrm{d}\zeta_{l_k}}{\mathrm{d}l_k}=$$
$$\left\{\frac{wC_k+(w-\hat{x}_{k|k}+\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}\hat{\lambda}_{k|k})\sigma_B^2\zeta_{l_k}-\left[\kappa_{x,k}^2(1-\rho_k^2)+\sigma_B^2\zeta_{l_k}\right]w}{C_k\sqrt{2\pi\sigma_B^2{\zeta_{l_k}}^3}}+\frac{\zeta_{l_k}\kappa_{x,k}^2-\zeta_{l_k}\rho_k^2\kappa_{x,k}^2-\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}\sigma_B^2\zeta_{l_k}}{C_k\sqrt{2\pi\sigma_B^2{\zeta_{l_k}}^3}}\cdot\frac{B_kw_k\kappa_{\lambda,k}^2+\hat{\lambda}_{k|k}C_k}{B_k^2\kappa_{\lambda,k}^2+C_k}\right\}\cdot$$
$$\sqrt{\frac{\sigma_B^2\zeta_{l_k}}{B_k^2\kappa_{\lambda,k}^2+C_k}}\cdot\exp\left(-\frac{\left(w-\hat{x}_{k|k}-\zeta_{l_k}\hat{\lambda}_{k|k}\right)^2}{2\left(B_k^2\kappa_{\lambda,k}^2+C_k\right)}\right)\frac{\mathrm{d}\zeta_{l_k}}{\mathrm{d}l_k}=\frac{w_k\sigma_B^2\left(B_k^2\kappa_k^2+C_k\right)-A_k\left(B_kw_k\kappa_k^2+\hat{\lambda}_{k|k}C_k\right)}{C_k\sqrt{2\pi\left(B_k^2\kappa_{\lambda,k}^2+C_k\right)^3}}\cdot\exp\left(-\frac{(w_k-B_k\hat{\lambda}_{k|k})^2}{2(B_k^2\kappa_{\lambda,k}^2+C_k)}\right)\frac{\mathrm{d}\zeta_{l_k}}{\mathrm{d}l_k} \tag{A3}$$

式中：$A_k=\sigma_B^2\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}-\sigma_{x_k|\lambda,k}^2$。

2）CDF推导过程。

① 考虑到$x_k|\lambda_k,\boldsymbol{Y}_{0:k}\sim N\left(u_{x_k|\lambda,k},\sigma_{x_k|\lambda,k}^2\right)$，基于引理2和全概率公式，式(15)可以转化为式(A4)。

将$\mu_{x_k|\lambda,k}$代入，式(A4)可转化为式(A5)。

$$E_{x_k|\lambda_k,\boldsymbol{Y}_{0:k}}\left(F_{L_k|\lambda_k,x_k,\boldsymbol{Y}_{0:k}}(l_k|\lambda_k,x_k,\boldsymbol{Y}_{0:k})\right)=1-\Phi\left(\frac{w-\mu_{x_k|\lambda,k}-\lambda_k\zeta_{l_k}}{\sqrt{\sigma_B^2\zeta_{l_k}+\sigma_{x_k|\lambda,k}^2}}\right)+\exp\left(\frac{2\lambda_k\left(w-\mu_{x_k|\lambda,k}\right)}{\sigma_B^2}+\frac{2\lambda_k^2\left(w-\sigma_{x_k|\lambda,k}^2\right)}{\sigma_B^4}\right)\cdot$$
$$\Phi\left(\frac{-w+\mu_{x_k|\lambda,k}-\lambda_k\zeta_{l_k}-\dfrac{2\lambda_k}{\sigma_B^2}\sigma_{x_k|\lambda,k}^2}{\sqrt{\sigma_B^2\zeta_{l_k}+\sigma_{x_k|\lambda,k}^2}}\right)\tag{A4}$$

$$E_{x_k|\lambda_k,\boldsymbol{Y}_{0:k}}\left(F_{L_k|\lambda_k,x_k,\boldsymbol{Y}_{0:k}}(l_k|\lambda_k,x_k,\boldsymbol{Y}_{0:k})\right)=1-\Phi\left(\frac{w-\hat{x}_{k|k}+\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}\hat{\lambda}_{k|k}-\left(\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}+\zeta_{l_k}\right)\lambda_k}{\sqrt{\sigma_B^2\zeta_{l_k}+\sigma_{x_k|\lambda,k}^2}}\right)+$$
$$\exp\left(\frac{2\lambda_k\left(w-\hat{x}_{k|k}+\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}\hat{\lambda}_{k|k}\right)}{\sigma_B^2}+\frac{2\lambda_k^2\left(w-\sigma_{x_k|\lambda,k}^2-\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}\sigma_B^2\right)}{\sigma_B^4}\right)\cdot$$
$$\Phi\left(\frac{-w+\hat{x}_{k|k}-\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}\hat{\lambda}_{k|k}+\lambda_k\left(\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}-\zeta_{l_k}-\dfrac{2}{\sigma_B^2}\sigma_{x_k|\lambda,k}^2\right)}{\sqrt{\sigma_B^2\zeta_{l_k}+\sigma_{x_k|\lambda,k}^2}}\right)=$$
$$1-\Phi\left(C_{2,k}+D_{1,k}\lambda_k\right)+\exp\left(A_{1,k}\lambda_k+B_{1,k}\lambda_k^2\right)\Phi\left(C_{1,k}+\lambda_k D_k\right)\tag{A5}$$

$$E_{\lambda_k|\boldsymbol{Y}_{0:k}}\left(\Phi\left(C_{2,k}+D_{1,k}\lambda_k\right)\right)=\Phi\left(\frac{C_{2,k}+D_{1,k}\hat{\lambda}_{k|k}}{\sqrt{1+D_{1,k}^2\kappa_{\lambda,k}^2}}\right)=\Phi\left(\frac{w-\hat{x}_{k|k}+\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}\hat{\lambda}_{k|k}-\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}\hat{\lambda}_{k|k}-\zeta_{l_k}\hat{\lambda}_{k|k}}{\sqrt{\sigma_B^2\zeta_{l_k}+\sigma_{x_k|\lambda,k}^2+\left(\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}+\zeta_{l_k}\right)^2\kappa_{\lambda,k}^2}}\right)=$$
$$\Phi\left(\frac{w-\hat{x}_{k|k}-\zeta_{l_k}\hat{\lambda}_{k|k}}{\sqrt{\sigma_B^2\zeta_{l_k}+\sigma_{x_k|\lambda,k}^2+\left(\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}+\zeta_{l_k}\right)^2\kappa_{\lambda,k}^2}}\right)=\Phi\left(\frac{w-\hat{x}_{k|k}-\zeta_{l_k}\hat{\lambda}_{k|k}}{\sqrt{C_k+B_k^2\kappa_{\lambda,k}^2}}\right)\tag{A6}$$

$$E_{\lambda_k|\boldsymbol{Y}_{0:k}}\left(\exp\left(A_{1,k}\lambda_k+B_{1,k}\lambda_k^2\right)\Phi\left(C_{1,k}+\lambda_k D_k\right)\right)=\frac{1}{\sqrt{1-2B_{1,k}\kappa_{\lambda,k}^2}}\exp\left(\frac{2A_{1,k}\hat{\lambda}_{k|k}+A_{1,k}^2\kappa_{\lambda,k}^2-2B_{1,k}\hat{\lambda}_{k|k}^2}{2\left(1-2B_{1,k}\kappa_{\lambda,k}^2\right)}\right)\cdot$$
$$\Phi\left(\frac{C_{1,k}+D_k\hat{\lambda}_{k|k}+A_{1,k}D_k\kappa_{\lambda,k}^2-2B_{1,k}C_{1,k}\kappa_{\lambda,k}^2}{\sqrt{\left(1-2B_{1,k}\kappa_{\lambda,k}^2\right)^2+D_k^2\kappa_{\lambda,k}^2\left(1-2B_{1,k}\kappa_{\lambda,k}^2\right)}}\right)\tag{A7}$$

$$F_{L_k|\boldsymbol{Y}_{0:k}}(l_k|\boldsymbol{Y}_{0:k})=1-\Phi\left(\frac{w-\hat{x}_{k|k}-\zeta_{l_k}\hat{\lambda}_{k|k}}{\sqrt{C_k+B_k^2\kappa_{\lambda,k}^2}}\right)+\frac{1}{\sqrt{1-2B_{1,k}\kappa_{\lambda,k}^2}}\exp\left(\frac{2A_{1,k}\hat{\lambda}_{k|k}+A_{1,k}^2\kappa_{\lambda,k}^2-2B_{1,k}\hat{\lambda}_{k|k}^2}{2\left(1-2B_{1,k}\kappa_{\lambda,k}^2\right)}\right)\cdot$$
$$\Phi\left(\frac{C_{1,k}+D_k\hat{\lambda}_{k|k}+A_{1,k}D_k\kappa_{\lambda,k}^2-2B_{1,k}C_{1,k}\kappa_{\lambda,k}^2}{\sqrt{\left(1-2B_{1,k}\kappa_{\lambda,k}^2\right)^2+D_k^2\kappa_{\lambda,k}^2\left(1-2B_{1,k}\kappa_{\lambda,k}^2\right)}}\right)\tag{A8}$$

式中: $C_{2,k}=\dfrac{w-\hat{x}_{k|k}+\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}\hat{\lambda}_{k|k}}{\sqrt{\sigma_B^2\zeta_{l_k}+\sigma_{x_k|\lambda,k}^2}}$; $D_{1,k}=-\dfrac{\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}+\zeta_{l_k}}{\sqrt{\sigma_B^2\zeta_{l_k}+\sigma_{x_k|\lambda,k}^2}}$; $A_{1,k}=\dfrac{2\left(w-\hat{x}_{k|k}+\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}\hat{\lambda}_{k|k}\right)}{\sigma_B^2}$; $B_{1,k}=\dfrac{2\left(w-\sigma_{x_k|\lambda,k}^2-\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}\sigma_B^2\right)}{\sigma_B^4}$;
$C_{1,k}=\dfrac{-w+\hat{x}_{k|k}-\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}\hat{\lambda}_{k|k}}{\sqrt{\sigma_B^2\zeta_{l_k}+\sigma_{x_k|\lambda,k}^2}}$; $D_k=\dfrac{\rho_k\dfrac{\kappa_{x,k}}{\kappa_{\lambda,k}}-\zeta_{l_k}-\dfrac{2}{\sigma_B^2}\sigma_{x_k|\lambda,k}^2}{\sqrt{\sigma_B^2\zeta_{l_k}+\sigma_{x_k|\lambda,k}^2}}$。

② 考虑到 λ_k 的随机性, 将 $1-\Phi\left(C_{2,k}+D_{1,k}\lambda_k\right)+\exp\left(A_{1,k}\lambda_k+B_{1,k}\lambda_k^2\right)\Phi\left(C_{1,k}+\lambda_k D_k\right)$ 分为 2 部分, 基于引理 2 可得: 第 1 部分推导如式 (A6) 所示; 第 2 部分推导如式 (A7) 所示; 综合 2 部分可得式 (A8)。

CDF 推导完毕。

Remaining useful life prediction based on implicit nonlinear Wiener degradation process

YANG Jiaxin[1], TANG Shengjin[1,*], LI Liang[1], SUN Xiaoyan[2], QI Shuai[3], SI Xiaosheng[1]

(1. Missile Engineering College, Rocket Force University of Engineering, Xi'an 710025, China;
2. Operational Support College, Rocket Force University of Engineering, Xi'an 710025, China;
3. Military Representative Office of Rocket Force Equipment Department in Zhengzhou, Zhengzhou 450000, China)

Abstract: Accurate remaining useful life prediction helps to improve the reliability and safety of the system and reduce the economic cost of the whole life cycle of the system. In engineering applications, due to the influence of uncertain measurement, the nonlinear degradation characteristics of stochastic degradation systems are in an implicit state. Since lifetime distribution estimation and degradation modeling are now the two main applications for the implicit scale transformation nonlinear Wiener degradation process, a remaining useful life prediction approach for this process is provided in this study. The parameters are updated online in accordance with the field degradation data of the equipment by the Kalman filtering approach, after the degradation model based on the nonlinear Wiener process is constructed and takes into account both measurement errors and nonlinear deterioration through scale transformation. The analytical expression of the probability density function and cumulative distribution function of remaining useful life considering online updating of model parameters are derived. Then, based on the historical degradation data, a maximum likelihood unbiased estimation method for the unknown parameters of the implicit scale transformation nonlinear Wiener degradation process model is proposed. The simulation degradation data and actual turbofan engine data are used for experimental verification. The experimental results show that the remaining useful life prediction method considering both measurement uncertainty and nonlinear degradation characteristics obtains higher prediction accuracy.

Keywords: remaining useful life; nonlinear; scale transformation; uncertain measurement; Wiener process; Kalman filtering

Received: 2022-04-15; **Accepted**: 2022-05-13; **Published Online**: 2022-05-30 16:43

URL: link.cnki.net/urlid/11.2625.V.20220527.1918.001

Foundation items: National Natural Science Foundation of China (61703410,61873175,61873273,61773386,61922089); Natural Science Basic Research Program of Shaanxi (2022JM-376)

* **Corresponding author**. E-mail: tangshengjin27@126.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0244

爆炸冲击波作用后变形机翼模态数值模拟

肖良丰，周兰伟*，李向东
(南京理工大学 机械工程学院，南京 210094)

摘　　要：为研究典型机翼在爆炸冲击波作用下的毁伤效应及其损伤后的结构动力学特性，基于有限元方法，研究了爆炸当量、爆炸方位及爆炸距离等对典型机翼损伤程度的影响，并分析了机翼结构变形程度与模态的关系。研究结果表明：冲击波强度和作用位置的变化对机翼结构的损伤及模态频率会产生不同程度的影响；随着冲击波强度的增加，机翼结构越早产生变形，对应的各阶模态频率下降越大，其中，二阶频率最大减少了15.02%；爆炸点位于机翼中心正上方时，机翼的变形最大；与无损伤机翼模态固有频率相比，爆炸冲击波作用在机翼中心位置时，各阶模态频率减小幅度最大，减少幅度为8.29%～15.02%。

关　键　词：爆炸冲击波；冲击损伤；机翼；模态；数值仿真
中图分类号：V214.1+1; TP391.9
文献标志码：A　　**文章编号**：1001-5965(2024)01-0341-09

战斗飞机最重要的要求之一就是其在战斗中的生存能力。在作战行动中，飞机可能会遭受战斗损伤，从而改变其结构动力学特性或其他性能，最终导致飞机飞行能力下降[1]。

在爆炸冲击波对机翼的损伤研究方面，Dehart和Basdekas[2]对飞机结构的结构响应进行了早期预测，推导了相关分析公式，根据弹性动态响应设计给定了爆炸强度的机身结构。针对机身结构整体响应的试验观察和有限元分析，具有代表性的研究是Settles等[3]所做的工作，其利用高速冲击波成像技术可视化飞机机身内的冲击载荷。Simmons和Schleyer[4]通过试验和有限元技术评估了飞机结构在冲击载荷作用下使用传统铆接技术和激光焊接技术的铝合金面板的响应和失效模式。Dacko和Toczyski[5]基于LS-DYNA软件分析了C4炸药爆炸产生的压力载荷作用在机身不同位置时机身的损伤，研究表明，机身在相对较小装药爆炸下只产生较小变形，但加固系统的结构构件发生严重损坏，在所有载荷条件下，爆炸装药下方的舱室地板均受到不同程度的破坏。程帅等[6]开展了爆炸试验，获得了飞机典型结构表面的反射超压历程、加筋结构应变、位移等响应数据。韩璐和韩庆[7]采用数值计算方法分析了冲击波超压对飞机结构的损伤。尹文明和任勇生[8]基于ANSYS软件分析了复合材料铺层方式对机翼固有频率、相对位移和相对应力的影响。Carrera等[9]采用数值计算方法研究了不同爆炸初始位置和炸药量下，爆炸冲击波对飞行状态下飞机的毁伤，研究结果能够用于评估不同初始爆炸位置和不同装药量下飞机的载荷。张媛等[10]采用数值模拟方法研究了直升机旋翼在战斗部爆炸形成的冲击波作用下的损伤模式。张宇等[11]基于LS-DYNA有限元软件研究了炸药当量、爆炸间距及爆炸方位等对典型飞机机翼在爆炸冲击波毁伤下的毁伤效应。冯晓伟等[12]从飞机结构强度设计角度出发，研究了飞机目标在大当量爆炸冲击波作用下的毁伤效应评估方法，通过有限元计算分析，获得了典型飞机目标机翼在均布冲击波下的失效判据。

对于机翼模态方面的研究，Vazquez等[13]通过有限元分析了非定常流动条件下变形机翼的模态频率，其低阶模态频率与冲击试验结果吻合较好。

收稿日期：2022-04-15；**录用日期**：2022-09-23；**网络出版时间**：2022-10-10 08:45
网络出版地址：link.cnki.net/urlid/11.2625.V.20221008.1724.009
*通信作者. E-mail：lwzhou@njust.edu.cn

引用格式：肖良丰，周兰伟，李向东. 爆炸冲击波作用后变形机翼模态数值模拟[J]. 北京航空航天大学学报，2024，50(1)：341-349.
XIAO L F，ZHOU L W，LI X D. Numerical simulation of deformed airfoil modal after blast shock wave[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(1)：341-349 (in Chinese).

Gasparetto 等[14]使用 ANSYS APDL 对飞机机翼的完整结构进行了建模，分析了机翼结构模态，数值结果与试验固有频率相比平均相差只有 5.8%。Masini 和 Timme[15]采用正交和动态模态分解方法，模拟了冲击波在翼尖附近波动引起的机翼振动。Banerjee[16]采用动力刚度法研究了大展弦比飞机的模态特性，获得了两翼飞机和四翼运输机的前 5 阶模态振型。Neu 等[17]采用随机子空间识别和频域分解 2 种工作模态分析方法，研究了跨声速条件下受气动激励机翼的模态参数，研究表明，随机子空间识别方法能够检测出一定频率范围内所有的弯曲和扭转模态。Kerschen 等[18]研究了全尺寸飞机的非线性模态及其相应的振动频率。Gloth 和 Sinapius[19]在检测飞机结构非线性行为中提出了扫描正弦激励测试的概念，用于模态测试试验。Arquier 等[20]提出了使用时间积分周期法和模态表示法计算无阻尼非线性机械系统的非线性振型。

目前，国内外研究主要集中在对机翼结构的纯爆炸冲击毁伤效果及纯模态方面，在爆炸冲击波作用后变形机翼的模态变化规律方面，相关研究较少。

机翼作为飞机的主要部件之一，不仅影响着飞机的气动性能，还承受着来自外界的各种载荷。在考虑影响飞机正常飞行的因素过程中，机翼是飞机暴露面积最大的部位。因此，研究冲击波对机翼的毁伤效应及其结构动力学特性变化对于提高飞机的作战生存能力具有重要意义。本文以典型机翼结构为研究对象，采用有限元方法，建立爆炸冲击波作用下机翼结构损伤变形模型，分析爆炸当量、爆炸方位及爆炸距离等对典型机翼结构损伤变形的影响，并分析机翼结构遭受爆炸冲击损伤变形后的模态频率的变化规律。

1 有限元建模

1.1 数值计算模型

计算选用的典型机翼结构如图 1 所示，其弦长 c=1 637.7 mm，展长 b=4 800 mm；机翼主要由铝合金蒙皮、长桁支撑梁、支撑梁凸缘、翼肋等结构部件组成。对机翼的组成结构进行分析，可知长桁支撑梁、翼肋和蒙皮是机翼的主要承力件，故将机翼结构等效为由上下端面和内部支撑结构组成的简化机翼结构。机翼蒙皮和内部支撑结构材料使用 2024 铝合金材料，蒙皮厚度为 1.0 mm，内部支撑结构厚度为 5.0 mm。

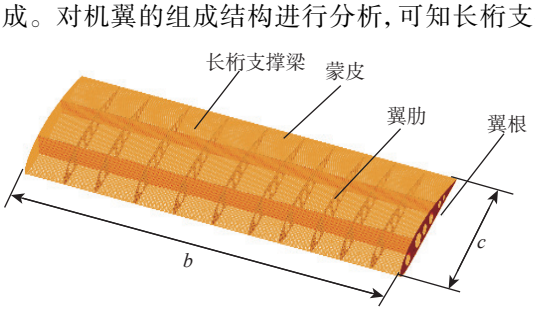


图 1 机翼结构

Fig. 1 Wing structure

建立计算模型时，机翼蒙皮和内部支撑结构均采用壳单元的形式进行模拟，划分网格时，将网格尺寸设置为 2.0 mm，根据机翼结构参数给定其对应的厚度，并在机翼根部位置设置固支约束。爆炸中心位于机翼上方，其距机翼蒙皮上端面的距离为 S，距翼根的距离为 L，位于弦长的中间位置。爆炸中心与机翼的位置关系如图 2 所示。

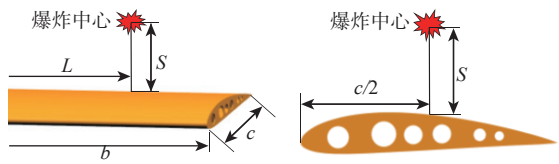


图 2 爆炸中心与翼展和弦长的关系

Fig. 2 Explosion center versus wingspan and chord length

1.2 材料参数

数值计算时，采用 Euler 方法计算冲击波对机翼的损伤。爆炸冲击波由球形 TNT 装药爆炸产生，通过 LS-DYNA 软件中的*LOAD-BLAST 关键字来施加冲击波载荷。该关键字选取的工况为理想的空气爆炸，用于施加常规炸药爆炸产生的压力载荷。在该关键字中，定义球形 TNT 炸药的质量，根据装药与目标表面的位置关系确定爆炸中心与目标的 x、y、z 坐标点，起爆时间保持默认，爆炸类型设置为空气中爆炸，爆炸压力失效时间保持默认。输入参数后，即可计算出作用在目标上的爆炸压力载荷。由于机翼在损伤过程中变形较大，计算中采用 Johnson-Cook 本构模型来模拟飞机机翼在爆炸冲击波作用下的损伤变形效果。其结构材料参数如表 1[21]所示。

表 1 机翼材料参数[21]

Table 1 Wing material parameters[21]

参数	数值
密度ρ/(g·cm^{-3})	2.78
弹性模量E/GPa	73.083
泊松比μ	0.33
本构参数n	0.73
本构参数C	0.008 3
本构参数m	1.7
本构参数T_m	775.00
本构参数T_r	300.00
本构参数A/MPa	369.00
本构参数B/MPa	648.00
失效参数D_1, D_2, D_3, D_4, D_5	0.13, 0.13, −1.5, 0.011, 0

2　机翼损伤变形计算及分析

根据第 1 节建立的典型机翼结构的爆炸冲击波损伤数值模型, 模拟机翼在无限空间爆炸冲击波作用下的损伤变形过程, 包括冲击波的传播过程, 计算时间为 20 ms, 为进一步进行模态分析研究提供原始模型。

2.1　爆炸当量

为研究不同冲击波强度对机翼损伤的影响, 计算分析 TNT 当量质量 W 分别为 0.05, 0.10, 0.15, 0.20 kg 时机翼的损伤情况。其中, 爆炸点位于机翼中心正上方, 其爆炸中心距离机翼上表面蒙皮 1.0 m, 爆炸点位置示意如图 2 所示。

机翼结构变形位移云图如图 3 所示。可以看出, 随着 TNT 当量质量的增加, 机翼蒙皮发生变形的面积逐渐增大, 内部支撑结构发生不同程度的弯曲变形。随着变形损伤面积的扩展, 将影响机翼蒙皮的气动能力, 然而机翼结构并未失效。当 TNT 当量质量超过 0.20 kg 后, 机翼结构的弯曲变形从翼梢向机翼中心区域延伸, 内部支撑结构对蒙皮的支撑能力下降, 使蒙皮更易发生损伤失效。

沿翼展方向机翼上表面蒙皮典型节点的变形量曲线如图 4 所示。分析可知, W<0.10 kg 时, 机翼上表面蒙皮仅发生轻微变形, 机翼整体发生较小的弯曲变形; 随着 TNT 当量质量的进一步增加, 蒙皮变形量急剧增加, 爆炸点正下方机翼蒙皮沿翼肋两边产生较大的凹陷变形, 同时伴随机翼整体结构产生较大的弯曲变形。

图 5 为机翼上表面蒙皮典型节点随时间的变形关系曲线。可知, TNT 当量质量越大, 机翼结构

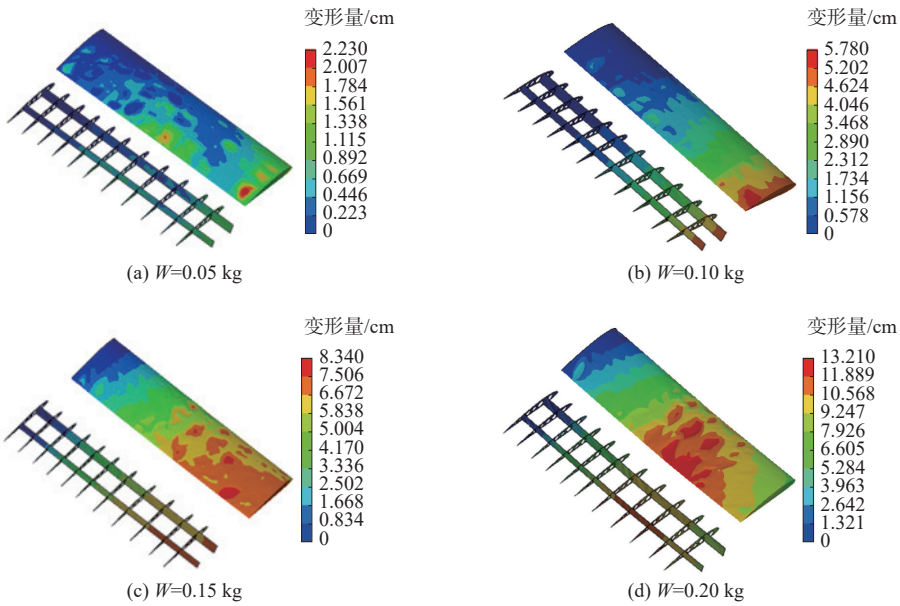


图 3　不同强度冲击波作用后机翼结构变形云图

Fig. 3　Deformation cloud diagram of wing structure after shock waves of different intensities

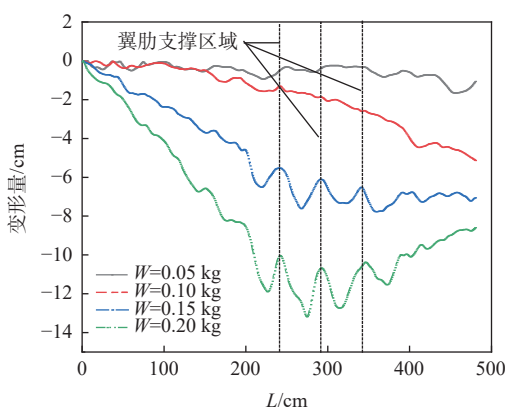


图 4　变形量与 TNT 当量质量 W、距离 L 的关系

Fig. 4　Deformation versus TNT equivalent mass W and distance L

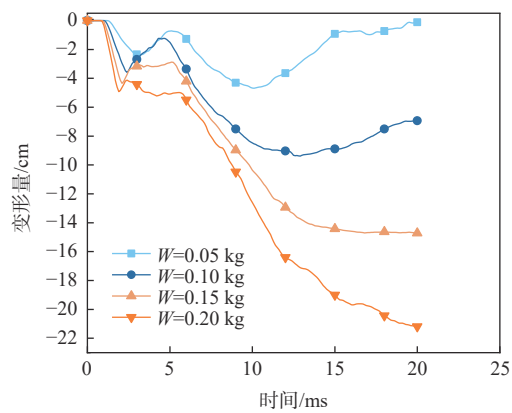


图 5　爆炸点正下方机翼蒙皮节点变形

Fig. 5　Deformation of wing skin node directly below explosion point

越早发生变形。说明TNT当量质量的增加会使其爆炸产生的冲击波的速度增加，从而越早作用于机翼上，同时更强的冲击波也会造成机翼有更大的变形。

2.2 爆炸方位

为研究不同爆炸位置对机翼的损伤情况，设置爆炸中心位置分别为 L/b=0, L/b=0.25, L/b=0.5, L/b=0.75 和 L/b=1.0, L/b 为翼根和爆炸中心间距与翼展长度之比。爆炸中心位于机翼正上方 1.0 m, W=0.20 kg。

从图6可以看到，当爆炸中心位于 L/b=0 时，机翼蒙皮仅发生轻微变形，且最大变形发生在靠翼梢后缘的区域；爆炸中心位于 L/b=0.25 时，最大变形位于爆炸点正下方区域附近；爆炸中心处于机翼 L/b=0.5 与 L/b=1.0 之间时，机翼蒙皮最大变形发生在翼梢区域。爆炸中心越靠近机翼中心区域，爆炸冲击波对机翼蒙皮的影响越大，损伤程度越严重。但在所有情况下，机翼蒙皮及内部支撑结构都只是发生不同程度的变形，机翼整体结构完整，未失效。

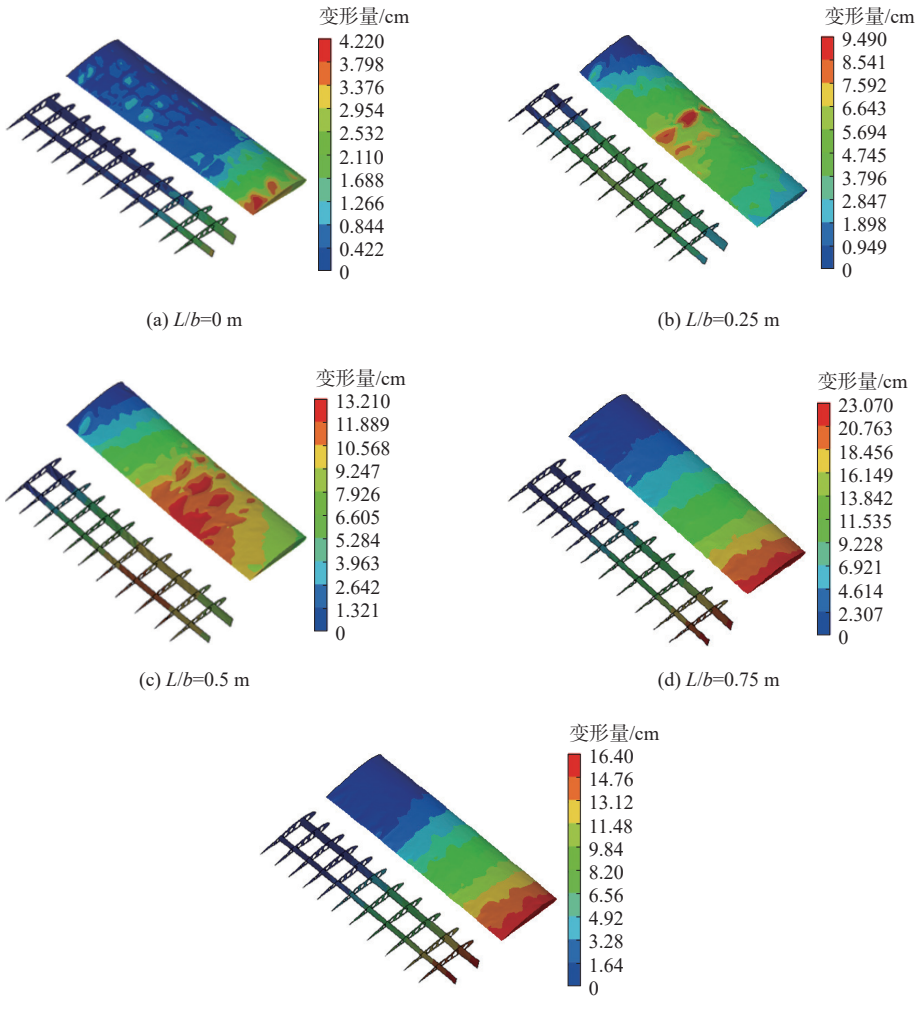


图6 冲击波作用在不同位置时机翼结构变形云图

Fig. 6 Deformation cloud diagram of wing structure when shock wave acts on different positions

图7为爆炸中心在机翼不同位置正上方时沿翼展方向机翼上表面蒙皮典型节点变形量曲线。可以看出，在所有情况下，机翼蒙皮都产生了变形，但并未破裂失效。当爆炸中心位于 L/b=0.25 和 L/b=0.75 之间正上方时，正下方区域蒙皮产生较大的凹陷变形，如图中虚线所示位置为翼肋支撑区域，由于翼肋的限制，蒙皮变形受到了一定程度的抑制。爆炸冲击波作用中心位于机翼 L/b=0.75 以后的区域时，机翼蒙皮不再出现凹凸变形，而是整体弯曲变形，在此之后产生的变形变化趋于一致。

图8为机翼上表面蒙皮典型节点的变形量随时间变化曲线。可以看到，当爆炸中心位于 L/b=0 正上方时，由于机翼与机身之间的固支关系，变形位移最小。爆炸中心位于 L/b=0.5 及以后时，结构产生的变形程度更大，蒙皮节点位移变化趋势一致。

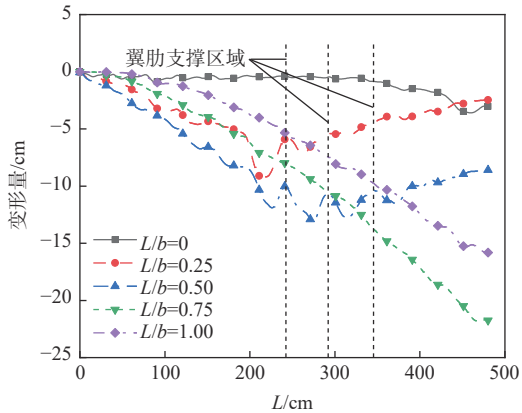


图 7　变形量与爆炸点 L/b、L 的关系

Fig. 7　Deformation versus blow-up point L/b and L

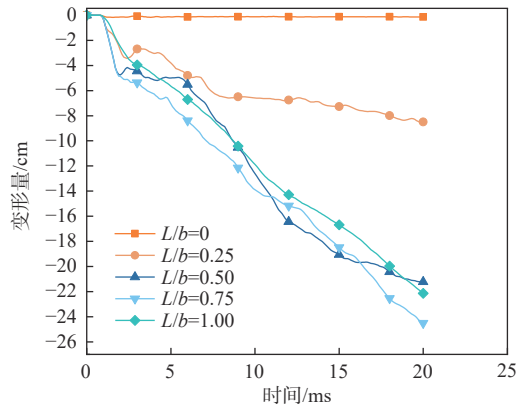


图 8　机翼上表面翼型节点的变形

Fig. 8　Deformation of typical nodes on upper surface of wing

2.3　爆炸距离

为研究不同爆炸距离对机翼结构的损伤, 设置爆炸中心分别距离机翼上表面蒙皮为 1.0, 1.5, 2.0, 2.5, 3.0 m, W=0.20 kg。

根据图 9 可知, 当爆炸距离为 1.0 m 时, 在爆炸中心正下方机翼蒙皮发生大面积变形, 机翼结构整体发生弯曲变形, 机翼内部支撑结构会限制机翼蒙皮的进一步变形; 当爆炸距离为 1.5 m 时, 机翼结构的最大变形量开始向翼梢附近区域集中, 此时最大变形量与爆炸距离为 1.0 m 时相比减少了 42.1%; 当爆炸距离增加到 2.0 m 以后, 蒙皮最大变形出现在翼梢附近区域, 最大变形量相应减少了 53.97%、55.15% 和 56.50%。由于翼根位置固支, 在遭受冲击波作用后, 机翼整体发生弯曲变形。

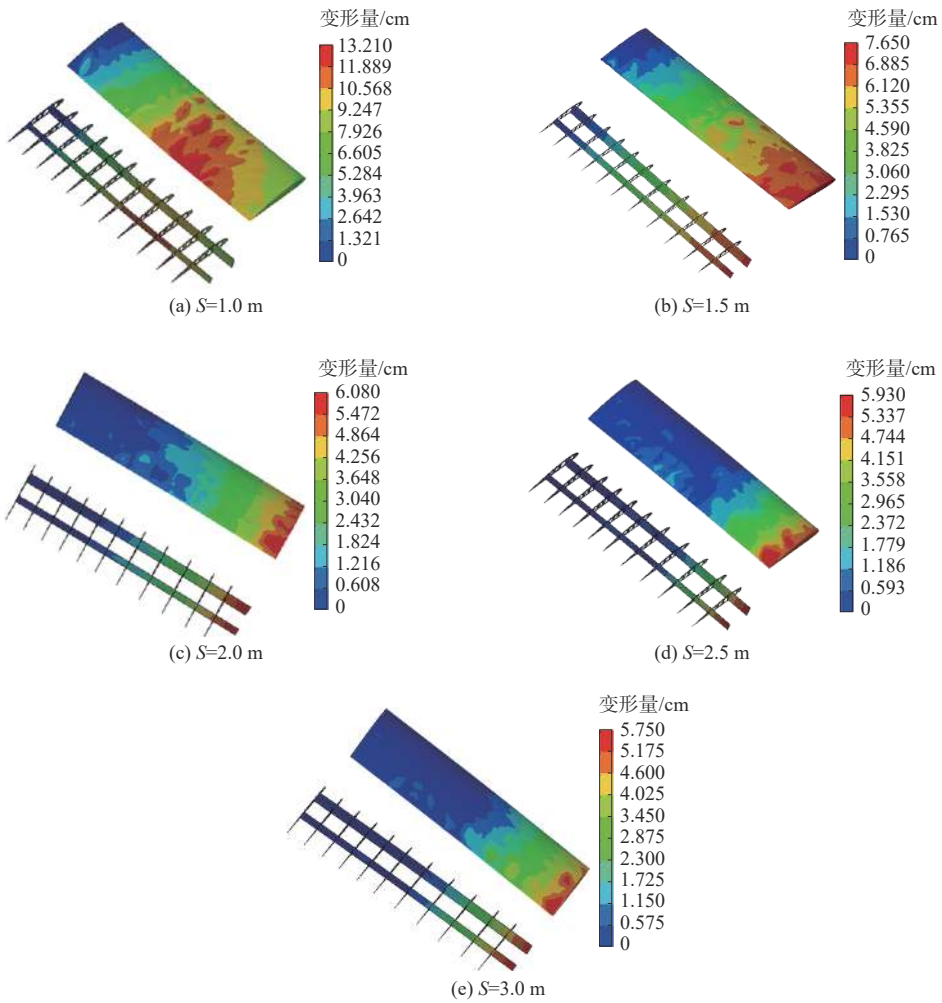


图 9　不同爆炸距离时机翼结构变形云图

Fig. 9　Cloud diagram of wing structure deformation at different explosion distances

图 10 为沿翼展方向机翼上表面蒙皮典型节点的变形曲线。可以看到，当爆炸距离为 1.0 m 时，爆炸中心正下方区域蒙皮发生大面积的凹陷变形，机翼结构整体产生较大的弯曲变形，翼梢处变形量达到 8.37 cm；随着爆炸距离的增大，蒙皮变形量逐渐减少，机翼结构仅出现较小的弯曲变形；爆炸间距越小，机翼结构越早出现变形，且变形程度越大；当爆炸距离超过 2.0 m 以后，蒙皮及机翼整体结构变形的变化趋于一致。

图 11 为机翼上表面蒙皮典型节点变形量随时间变化曲线。可以得到，爆炸距离越小，冲击波强度越大，机翼结构越早产生变形。爆炸距离为 1.0 m 和 1.5 m 时，随着冲击波作用时间的增加，蒙皮节点位移呈逐渐增大的趋势；爆炸距离为 2.0 m、2.5 m 和 3.0 m 的情况下，随着冲击波作用时间的增加，蒙皮节点位移遵循先增大后减小的趋势。

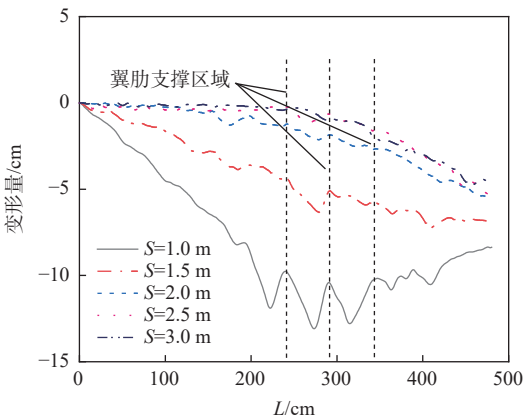


图 10 变形量与 S、L 的关系

Fig. 10 Deformation versus S and L

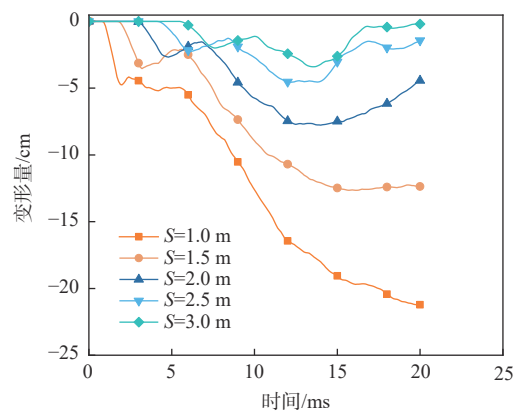


图 11 机翼蒙皮典型节点变形量-时间曲线

Fig. 11 Typical node deformation-time curves of wing skin

3 机翼损伤变形对模态的影响

基于第 2 节建立的模态分析模型，研究爆炸冲击波作用后机翼的模态变化规律。

3.1 机翼模态分析

表 2 为无冲击损伤机翼对应的前三阶模态固有频率，图 12 为无冲击损伤机翼结构的模态振型响应位移云图。分析可知，机翼结构的振型随频率的变大而逐渐变得更加复杂；一阶和二阶模态振型的最大位移主要集中在 L/b=0.25 到 L/b=1.0 的翼展区域，且沿翼展方向呈增大趋势；三阶振动最大位移主要集中在 L/b=0.25 到 L/b=0.6 的翼展后缘区域。

由于篇幅限制，只给出部分遭受冲击损伤变形机翼模态振型图。图 13 为 W=0.20 kg，爆炸中心点

表 2 无冲击损伤机翼固有频率

Table 2 Natural frequency of wing without impact damage

模态	固有频率/Hz
1	87.15
2	414.36
3	449.89

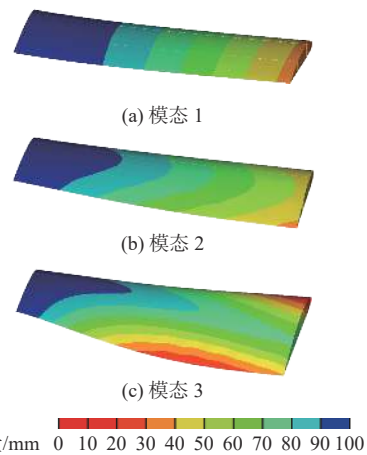


图 12 无冲击损伤条件下机翼前三阶模态响应

Fig. 12 The first third-order modal response of wing without impact damage conditions

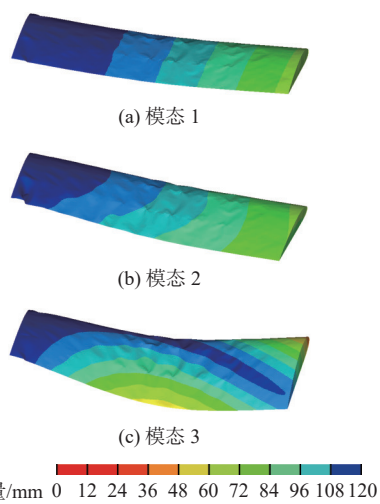


图 13 冲击损伤条件下机翼前三阶模态响应

Fig. 13 The first third-order modal response of wing under impact damage conditions

位于机翼蒙皮上表面 1.0 m 位置时, 受爆炸冲击损伤变形后机翼前三阶模态振型响应云图。可知, 爆炸冲击作用下, 机翼结构出现损伤变形, 其各阶模态发生不同程度的变化, 最大位移集中区域也发生了明显的改变, 一阶和二阶振型的最大位移主要发生在 L/b=0.5 到 L/b=1.0 区域; 三阶模态的最大位移主要集中在 L/b=0.6 以后的翼展后缘区域。

3.2　冲击波强度的影响

为量化损伤效应, 通过从受损变形机翼的相同模态频率中减去未受损机翼的模态频率, 并使用百分比来表示存在损伤变形而产生的模态频率差异。损伤变形与未损伤机翼的模态频率相对偏差定义如下:

$$\Delta f = f_{\mathrm{d}} - f_{\mathrm{un}} \tag{1}$$

$$r = \frac{\Delta f}{f_{\mathrm{un}}} \times 100\% \tag{2}$$

式中: Δf为频率受损变形机翼和未受损机翼模态频率之间的差值; f_{d} 为受损变形机翼的模态频率; f_{un} 为未受损机翼的固有模态频率; r 为频率相对偏差。

本文模型选择在不同强度冲击波作用后的机翼, 该机翼遭受不同强度的爆炸冲击波作用后结构产生不同程度的变形, 但并未失效。图 14 为该状态下机翼前三阶模态相对偏差变化趋势。可以看出, 随着 TNT 当量质量的增加, 机翼结构变形程度增加, 各阶模态频率相对偏差也呈增大的趋势。可知, 当 TNT 当量质量小于 0.10 kg 时, 各阶模态频率相对偏差变化较慢; TNT 当量质量大于 0.10 kg 后, 各阶模态频率相对偏差变化速度加快。二阶模态在 TNT 当量质量大于 0.10 kg 后频率相对偏差值增加速度最快, 在 TNT 当量质量为 0.20 kg 时, 频率相对偏差达到了−15.02%; 不同 TNT 当量质量产生的冲击波对机翼结构的二阶模态频率影响较大, 对三阶模态频率影响相对较小。

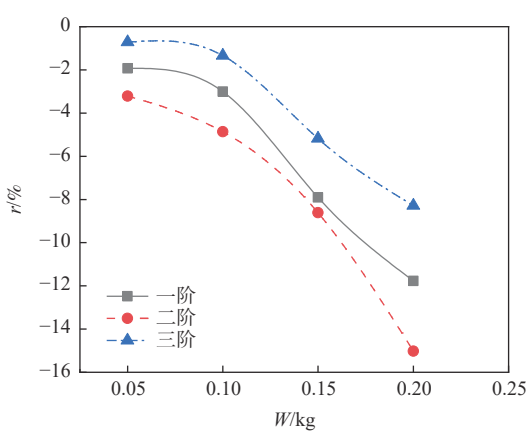


图 14　频率相对偏差与 W 的关系

Fig. 14　Relative frequency deviation versus W

在不同 TNT 当量质量爆炸冲击波作用下, 机翼前三阶模态频率都有不同程度的下降, 频率变化与 TNT 当量质量呈负相关。依据各阶模态的变化量, 可以估计对应的冲击波强度大小, 同时估计出机翼结构的整体变形量。

3.3　爆炸冲击波作用位置的影响

为考察爆炸方位因素对机翼结构模态固有频率的影响, 以不同位置遭受爆炸冲击损伤变形后的机翼为计算模型, 分析机翼受冲击损伤变形后前三阶模态固有频率的变化规律。图 15 为同一强度的冲击波作用在机翼不同位置处变形机翼各阶模态频率相对偏差曲线。可以看出, 与无损伤变形的机翼相比, 有损伤变形机翼的各阶模态频率都有不同程度的减小, 且变化趋势基本一致。各阶模态频率在 L/b=0.5 位置处相对偏差最大, 分别达到了−11.77%、−15.02% 和−8.29%。爆炸冲击波作用在机翼翼根处造成的各阶模态频率变化较小, 分别只有−3.61%、−3.93% 和−2.24%。而当爆炸冲击波作用在翼梢位置时, 三阶模态频率相对偏差最小, 仅为−1.70%。作用在机翼结构不同位置处的爆炸冲击波对机翼二阶模态频率影响最大, 最大频率相对偏差达 15.02%, 对机翼的三阶模态频率影响最小。

从图 15 中可以看出, 随着爆炸中心的变化, 爆炸冲击波造成的机翼变形对二阶模态频率的影响始终是最大的。爆炸点在 L/b=0 到 L/b=0.25 之间正上方时, 一阶和二阶模态频率下降幅度非常接近, 直到 L/b=0.5 位置处时, 各阶频率下降值达到最大, 之后各阶模态下降幅值逐渐减小。依据各阶模态的变化规律, 可以预估爆炸冲击波作用中心在翼展方向所处的位置。

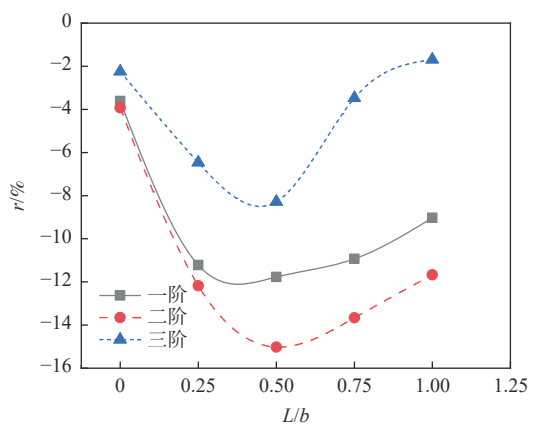


图 15　频率相对偏差与爆炸点 L/b 的关系

Fig. 15　Relative frequency deviation versus blow-up point L/b

3.4　爆炸冲击波作用距离的影响

为研究爆炸距离因素对机翼模态固有频率的影响, 采用遭受爆炸冲击波冲击损伤变形后的机翼

作为模态分析模型，分析不同爆炸距离造成的机翼损伤变形对模态的影响规律。图16为同一TNT当量质量的炸药在不同距离下爆炸对机翼造成损伤变形后各阶模态频率的相对偏差曲线。可以看出，随着爆炸距离的增大，各阶频率相对偏差量减小。当爆炸距离为1.0 m时，爆炸冲击波造成的机翼结构损伤变形使各阶模态都产生了较大的偏差，其中，对二阶模态频率影响最大，频率相对偏差达到了15.02%；随着爆炸距离的增加，一阶和三阶模态频率相对偏差逐渐减小，而对于二阶模态，频率相对偏差在爆炸距离为2.0 m以前逐渐减少，之后呈先增大后再减少的变化趋势。

从图16中可以看出，随着爆炸距离的变化，对机翼结构三阶模态频率的影响最小，变化过程较为平缓；而对于二阶模态频率影响最大，且变化过程较为复杂。依据各阶模态频率的变化规律，可以预估爆炸中心距离机翼的距离范围，进而确定出机翼结构的变形程度。

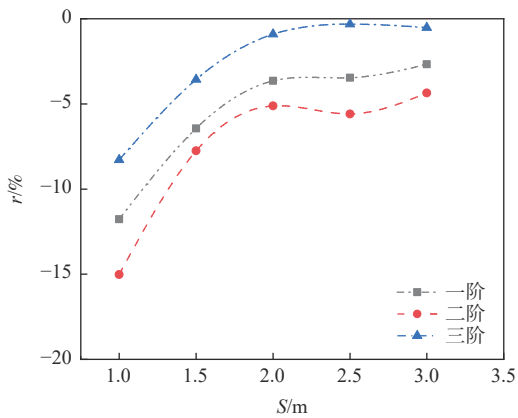


图16 频率相对偏差与S的关系

Fig. 16 Relative frequency deviation versus S

4 结 论

通过对典型机翼进行爆炸冲击波有限元计算，并对遭受冲击波损伤变形后的机翼进行模态分析，得出如下结论：

1）冲击波强度、爆炸位置、爆炸距离的变化对机翼结构的损伤及模态频率都会产生不同程度的影响。

2）随着冲击波强度的增加，机翼结构越早产生变形，对应的各阶模态频率下降也越大，其中，TNT当量质量为0.20 kg时，二阶频率最大减少了15.02%。

3）随着爆炸方位的变化，爆炸点位于机翼中心正上方时，冲击波造成机翼的变形最大。与无损伤机翼模态固有频率相比，冲击波作用在机翼中心位置时，各阶模态频率减小幅度最大，减少幅度为8.29%～15.02%。

4）随着爆炸距离的增加，作用在机翼上的冲击波强度迅速下降，机翼表面变形程度大幅下降。冲击损伤对各阶模态频率的影响逐渐减小，其中，距离增加到3.0 m时，各阶模态频率减小幅度只有2.24%～3.93%。

参考文献（References）

[1] BALL R E. The fundamentals of aircraft combat survivability: Analysis and design[M]. Reston: AIAA, 1985.

[2] DEHART R, BASDEKAS N. Blast response of a typical aircraft fuselage structure: AD275080[R]. Washington, D.C.: Bureau of Naval Weapons, 1962: 1-95.

[3] SETTLES G S, BENWOOD J R, GATTO J A. Optical shock wave imaging for aviation security[C]//Proceedings of the Fluids Engineering Division Summer Meeting. New York: ASME, 2003, 36975: 113-118.

[4] SIMMONS M C, SCHLEYER G K. Pulse pressure loading of aircraft structural panels[J]. Thin-Walled Structures, 2006, 44(5): 496-506.

[5] DACKO A, TOCZYSKI J. Structural response of a blast loaded fuselage[J]. Journal of KONES, 2010, 17(1): 101-109.

[6] 程帅, 刘文祥, 童念雪, 等. 爆炸载荷下飞机典型加筋结构毁伤特性[J]. 爆炸与冲击, 2021, 41(1): 88-95.
CHENG S, LIU W X, TONG N X, et al. Damage characteristics of aircraft typical stiffened under explosive load[J]. Explosion and Shock Waves, 2021, 41(1): 88-95(in Chinese).

[7] 韩璐, 韩庆. 飞机在模拟混合破片威力场打击下的易损性计算[J]. 航空工程进展, 2014, 5(4): 455-462.
HAN L, HAN Q. Vulnerability calculation of aircraft under simulated mixed fragment power field[J]. Advances in Aeronautical Science and Engineering, 2014, 5(4): 455-462(in Chinese).

[8] 尹文明, 任勇生. 基于ANSYS的复合材料机翼的模态分析[C]//第十七届玻璃/复合材料学术年会. 北京: 中国硅酸盐学会, 2008: 19-21.
YIN W M, REN Y S. Modal analysis of composite wing based on ANSYS[C]//Proceedings of the Academic Annual Conference on Fiberglass/Composite Materials. Beijing: Chinese Ceramic Society, 2008: 19-21(in Chinese).

[9] CARRERA E, MORELLO V, VALLETTA D, et al. Simulation of shock wave impact due to an explosion on a flying flexible aircraft[J]. Combustion, Explosion, and Shock Waves, 2007, 43(6): 732-740.

[10] 张媛, 胡传辉, 刘刚, 等. 冲击波和破片对直升机旋翼的联合毁伤研究[J]. 弹箭与制导学报, 2013, 33(4): 109-112.
ZHANG Y, HU C H, LIU G, et al. Joint damage of helicopter rotor by the shock wave and fragment[J]. Journal of Projectiles, Rockets, Missiles and Guidance, 2013, 33(4): 109-112(in Chinese).

[11] 张宇, 王彬文, 白春玉. 典型飞机机翼结构爆炸冲击毁伤数值分析研究[J]. 机械设计与制造, 2022, 41(9): 1468-1474.
ZHANG Y, WANG B W, BAI C Y. Research on numerical analysis of explosion impact damage of typical aircraft wing structure[J]. Mechanical Design and Manufacturing, 2022, 41(9): 1468-1474(in

Chinese).

[12] 冯晓伟, 卢永刚, 李永泽. 飞机目标在爆炸冲击波作用下的毁伤效应评估方法[J]. 高压物理学报, 2019, 33(4): 124-128.
FENG X W, LU Y G, LI Y Z. Evaluation method of damage effect of aircraft target under blast shock wave[J]. Chinese Journal of High Pressure Physics, 2019, 33(4): 124-128(in Chinese).

[13] VAZQUEZ C G, JOSHI K, BHATTACHARYA S, et al. Modal analysis of a morphing wing for unsteady flow conditions: AIAA 2020-2009[R]. Reston: AIAA, 2020.

[14] GASPARETTO V E L, MACHADO M, CARNEIRO S H S. Experimental modal analysis of an aircraft wing prototype for SAE aerodesign competition[J]. DYNA, 2020, 87(214): 100-110.

[15] MASINI L, TIMME S. Scale-resolving simulation of shock buffet onset physics on a civil aircraft wing[C]//Proceedings of the Applied Aerodynamics Conference. Reston: AIAA, 2018, 19: 391-409.

[16] BANERJEE J R. Modal analysis of sailplane and transport aircraft wings using the dynamic stiffness method[J]. Journal of Physics, 2016, 15: 473-486.

[17] NEU E, JANSEN F, KHATIB A, et al. Operational modal analysis of a wing excited by transonic flow[J]. Aerospace Science and Technology, 2016, 49: 73-79.

[18] KERSCHEN G, PEETERS M, GOLINVAL J C, et al. Nonlinear modal analysis of a full-scale aircraft[J]. Journal of Aircraft, 2013, 50(5): 1409-1419.

[19] GLOTH G, SINAPIUS M. Influence and characterization of weak non-linearities in swept-sine modal testing[J]. Aerospace Science and Technology, 2004, 8(2): 111-120.

[20] ARQUIER R, BELLIZZI S, BOUC R, et al. Two methods for the computation of nonlinear modes of vibrating systems at large amplitudes[J]. Computers & Structures, 2006, 84(24): 1565-1576.

[21] BUYUK M, KURTARAN H, MARZOUGUI D, et al. Automated design of threats and shields under hypervelocity impacts by using successive optimization methodology[J]. International Journal of Impact Engineering, 2008, 35(12): 1449-1458.

Numerical simulation of deformed airfoil modal after blast shock wave

XIAO Liangfeng，ZHOU Lanwei*，LI Xiangdong

(School of Mechanical Engineering，Nanjing University of Science and Technology，Nanjing 210094，China)

Abstract: The effects of explosive mass, blast azimuth, and blast distance on the degree of damage to typical airfoils were investigated using the finite element method in order to analyze the damaging effect of typical airfoils under the action of blast shock wave and its structural dynamics following damage. The relationship between the deformation and the model of the wing structure was analyzed. The findings indicate that the wing structure damage and the modal frequency will be affected to varying degrees by changes in shock wave strength and action position. The changes in the shock wave intensity and action position will have different degrees of influence on the damage to the wing structure and the modal frequency. As the shock wave intensity increases, the earlier the wing structure deforms, and the deformation becomes larger the corresponding modal frequencies of all orders also drop more quickly, of which the second-order frequency decreases by 15.02% at most. When the explosion point was located directly above the center of the wing, the deformation of the wing was greatest. When the shock wave acts on the center of the wing, with the modal natural frequency of the undamaged wing, the modal frequency of each order showed the largest reduction, ranging from 8.29% to 15.02%.

Keywords: blast shock wave; impact damage; wing; modal; numerical simulation

Received：2022-04-15；**Accepted**：2022-09-23；**Published Online**：2022-10-10 08：45

URL：link.cnki.net/urlid/11.2625.V.20221008.1724.009

* **Corresponding author**. E-mail：lwzhou@njust.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0249

2195-T8 铝锂合金疲劳多裂纹融合试验与仿真研究

庞闯[1]，刘德俊[1]，田干[1,*]，郭伟[2]，刘德博[3]

(1. 火箭军工程大学 导弹工程学院，西安 710025；2. 北京航天试验技术研究所，北京 100048；
3. 北京宇航系统工程研究所，北京 100076)

摘　　要：铝锂合金作为航空航天广泛应用的合金材料，其疲劳断裂行为的研究对结构安全性评价具有重要意义。以第三代铝锂合金2195-T8为研究对象，通过恒幅拉-拉疲劳试验和有限元方法对2195-T8铝锂合金疲劳裂纹扩展行为进行试验与仿真研究。基于断面显微测量与观察，在仿真模型中引入多个初始裂纹，模拟多裂纹的融合扩展过程，获得多裂纹独自扩展、交融时扩展和融合后扩展的规律。结果表明：裂纹融合前，在疲劳循环载荷作用下，裂纹尖端应力强度因子总体上不断增大，塑性区域体积增加区域平缓；当裂纹相互融合时，裂纹面处应力强度因子瞬时增大，远高于其余裂尖数值大小；随着裂纹进一步融合，尖端应力强度因子数值趋于平稳；裂纹完全融合后，到达裂纹快速扩展阶段，塑性区域体积与扩展步数呈正比增加，扩展速率呈现先快后慢的规律，裂纹面交汇融合成新的椭圆形状裂纹面。

关　键　词：疲劳断裂；铝锂合金；裂纹融合；有限元；应力强度因子
中图分类号：TG147
文献标志码：A　　**文章编号**：1001-5965(2024)01-0350-09

铝锂合金具有低密度、高比强度、高比刚度和高断裂韧性等独特优势，是目前航空航天、军事领域和民用工业领域认为可全面替代传统铝合金的最具竞争力的合金材料[1-2]，但其在制造、安装或服役过程中会不可避免地产生裂纹缺陷，导致结构材料的服役能力下降等问题[3]。因此，对铝锂合金的抗疲劳性能和断裂行为进行研究极为重要。

目前，学者们对铝合金疲劳断裂行为从多方面进行了研究。在多裂纹的疲劳扩展研究中，魏国前等[4]对焊接接头焊趾部位裂纹扩展进行分析，验证了多裂纹的扩展具有明显的不同步特性和裂纹最终有向半椭圆形状演变的趋势；吴鹏飞等[5]在高周疲劳载荷下对铝合金进行了裂纹扩展试验，利用仿真软件得到了复合疲劳裂纹开裂方向的数值分析，且复合疲劳断口表现为脆性断裂；李如俊等[6]研究了深海潜水器壳体表面扰动裂纹对主裂纹的影响，表明扰动裂纹的存在和位置均会对主裂纹的应力强度因子产生影响，应力强度因子是裂纹疲劳扩展中的主要断裂参量[7]；针对贮箱存在的缺陷，赵颖等[8]通过仿真在应力强度值最大区域引入裂纹，对不同情况下的应力强度因子进行分析，得到了各类强度因子变化规律；在此基础上，徐利辉[9]、Amiri[10]和汪必升[11]等分别利用模态叠加法、损伤力学方法和扩展有限元方法研究了动态应力强度因子的变化。

当前，针对疲劳多裂纹扩展过程中应力强度因子变化规律的研究仍然广受关注，国内外对疲劳裂纹的探究大多集中在裂纹形貌、裂纹面断口和应力强度因子的分析上，但对动态多裂纹融合扩展过程中的尖端断裂参量、扩展过程中的机理和裂纹面形

收稿日期：2022-04-14；录用日期：2022-06-20；网络出版时间：2022-06-24 15:59
网络出版地址：link.cnki.net/urlid/11.2625.V.20220623.1342.004
基金项目：国家自然科学基金 (52075541)
*通信作者. E-mail：tiangan_2012@163.com

引用格式：庞闯，刘德俊，田干，等. 2195-T8铝锂合金疲劳多裂纹融合试验与仿真研究[J]. 北京航空航天大学学报，2024，50(1)：350-358.
PANG C，LIU D J，TIAN G，et al. Experimental and simulation study on fatigue multi crack fusion of 2195-T8 Al-Li alloy[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(1)：350-358 (in Chinese).

貌形成进行分析的相关研究还有待深入和扩展, 因此, 对动态多裂纹融合试验与仿真的研究不仅仅能为解决金属断裂问题提供参考, 更能加强对铝锂合金结构强度的评估。

本文通过恒幅拉-拉疲劳试验和等效裂纹模型方法对 2195-T8 铝锂合金疲劳多裂纹扩展行为进行试验与仿真研究, 从疲劳裂纹断口裂纹融合现象进行分析, 对断口处多裂纹融合现象进行仿真验证, 不仅分析了疲劳多裂纹融合扩展中动态应力强度因子的变化规律, 还对裂纹尖端处应力和塑性区域等方面展开了研究, 对 2195-T8 铝锂合金的应用及稳定性评价具有一定的工程价值。

1　相关理论

1.1　裂纹扩展

金属疲劳破坏过程通常分为 3 个阶段[12]: 裂纹萌生、裂纹稳定扩展和裂纹失稳扩展。图 1 为 2195-T8 铝锂合金裂纹扩展曲线。裂纹扩展速率 da/dN 为交变荷载作用下每次循环对应的裂纹扩展量, 在裂纹尖端应力场控制下, 应力强度因子 K 是表征裂纹尖端断裂水平的重要指标。

最初, 裂纹主要存在于应力集中的金属表面, 对于裂纹萌生阶段 Ⅰ, 其裂纹扩展速率 da/dN 基本小于 10^{-4} mm/cycle。当 ΔK 接近门槛值 ΔK_{th} 时, 裂纹开始扩展, 并且裂纹扩展速率 da/dN 与材料的微观结构和应力水平有关; 阶段 Ⅱ 为裂纹的稳定扩展区域, 其速率在 10^{-4}～10^{-3} mm/cycle 之间; 阶段 Ⅲ 是高速率裂纹扩展失稳区, 该区域裂纹扩展速率急剧增加, 裂纹尺寸迅速增大至结构断裂[13-14]。

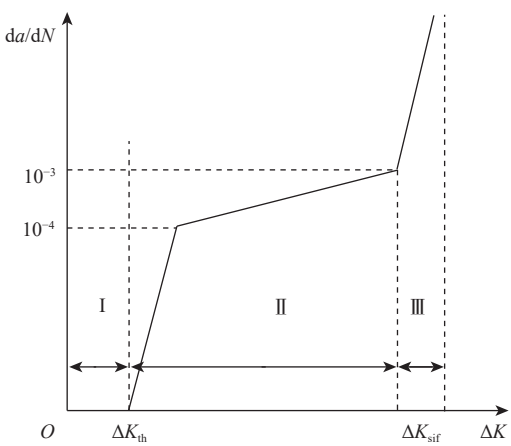


图 1　2195-T8 铝锂合金裂纹扩展速率曲线

Fig. 1　Crack growth rate curve of 2195-T8 Al-Li alloy

1.2　裂纹融合准则

疲劳多裂纹萌生后, 随着裂纹前沿进一步扩展, 相邻裂纹的裂纹前沿逐渐靠拢并融合成一条新的裂纹前沿, 如图 2 所示。在多裂纹的扩展仿真中, 要考虑多裂纹的融合准则, 需要采用裂纹扩展速率 da/dN 描述裂纹扩展的行为。

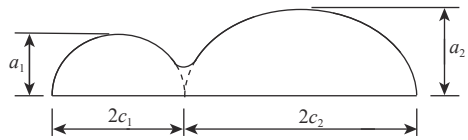


图 2　裂纹融合准则示意图

Fig. 2　Schematic diagram of crack fusion criterion

采用 Paris 公式作为裂纹扩展速率的模型, 在裂纹体网格模型中插入裂纹, 将裂纹前沿线离散为点, 将这些离散点定义为裂纹扩展的源点, 指定裂纹前沿应力强度因子中值点扩展距离为 $\Delta a_{\mathrm{median}}$, 计算所有前沿点的局部扩展距离, 得到新的裂纹前沿点, 再采用多项式曲线拟合方法, 即可得到新的裂纹前沿。Δa_i 的计算表达式为[15-16]

$$\Delta a_i = \Delta a_{\mathrm{median}}(K_i/K_{\mathrm{median}})^n \tag{1}$$

式中: Δa_i 为裂纹前沿节点 i 的裂纹增量; $\Delta a_{\mathrm{median}}$ 为上一个裂纹前沿到下一个裂纹前沿的裂纹增量; K_i 为裂纹前沿节点 i 的应力强度因子幅值; K_{median} 为整个裂纹前沿强度因子幅度的平均值; n 为自定义参数, 一般选取与 Paris 公式中相同的 m 值。

2　试样恒幅拉-拉疲劳试验

2.1　裂纹扩展试验

为得到 2195-T8 铝锂合金裂纹扩展的性能参数, 本文进行了仿真研究, 在国家标准条件下进行铝锂合金的裂纹扩展试验, 试验形状及尺寸如图 3(a)所示, 试验状态如图 3(b)所示。

经过应力比 R 为 0.1 的疲劳裂纹扩展试验, 得到 2195-T8 铝锂合金的 da/dN-ΔK 数据, 图 4 中给出了 2195-T8 铝锂合金裂纹扩展速率 da/dN 与应力强度因子幅值 ΔK 数值曲线。根据 Paris 公式, 通过拟合 da/dN-ΔK 试验数据, 可以得到铝锂合金在 Paris 区的疲劳裂纹扩展速率曲线方程, 其公式如下:

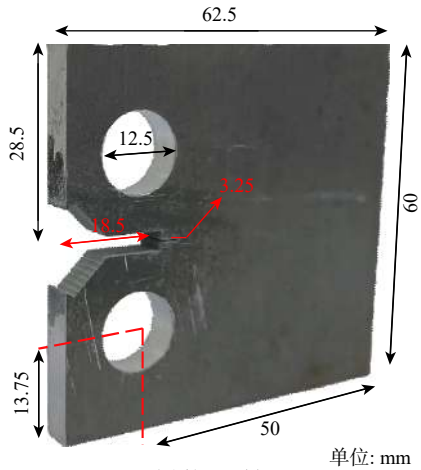
$$\mathrm{d}a/\mathrm{d}N = C(\Delta K)^n = 1.313\times10^{-7}(\Delta K)^{2.802} \tag{2}$$

式中: C、n 为描述疲劳裂纹扩展性能的基本参数。

多裂纹融合行为一般发生在裂纹稳步扩展阶段, 可基于 Paris 公式对裂纹融合扩展进行计算。裂纹扩展试验对本文疲劳多裂纹融合扩展仿真过程具有支撑作用, 可提高仿真的可靠性, 使仿真有据可循, 对 2195-T8 铝锂合金疲劳试验中试样出现多裂纹融合扩展行为进行研究和分析, 从而得出铝锂合金的扩展特性和规律。

2.2　疲劳试验

本文以材料为 2 195-T8 铝锂合金、长为 112 mm、



(a) 裂纹扩展试样



(b) 裂纹扩展试验

图 3 裂纹扩展试样与试验

Fig. 3 Crack growth specimens and tests

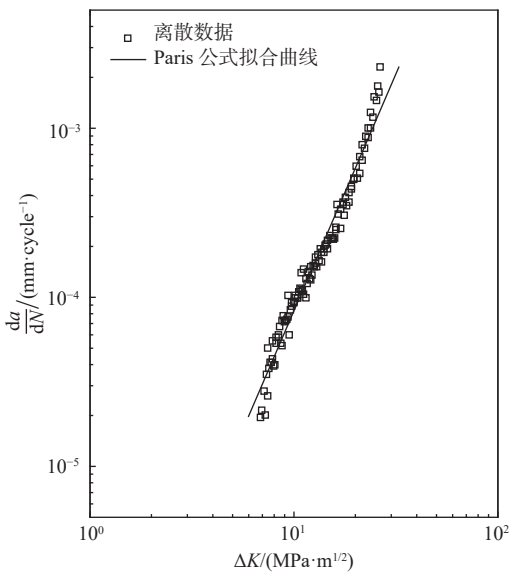


图 4 2195-T8 铝锂合金 da/dN-ΔK 曲线(室温)

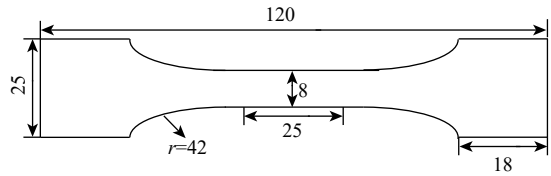
Fig. 4 da/dN-ΔK curve of 2195-T8 Al-Li alloy (room temperature)

宽为 25 mm、厚度为 4.5 mm 的狗骨状薄片试样为试验对象，试样形状及尺寸如图 5 所示，其基本力学性能如表 1 所示。

在 INSTRON8801 疲劳试验机上进行 2195-T8 铝锂合金试样的疲劳试验，疲劳试验样机如图 6(a)



(a) 试样形状



(b) 试样尺寸

图 5 疲劳试验试样形状与尺寸

Fig. 5 Shape and size of fatigue test specimens

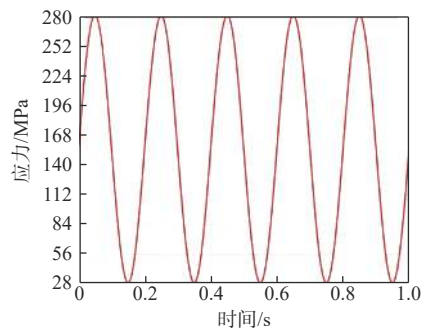
表 1 2195-T8 铝锂合金的基本力学性能

Table 1 Basic mechanical properties of 2195-T8 AL-Li alloy

抗拉强度 σ_b/MPa	屈服强度 σ_s/MPa	拉伸率 δ/ %	弹性模量 E/GPa
609.90	583.33	11.4	72.29



(a) 样机



(b) 加载频率

图 6 伺服试验机

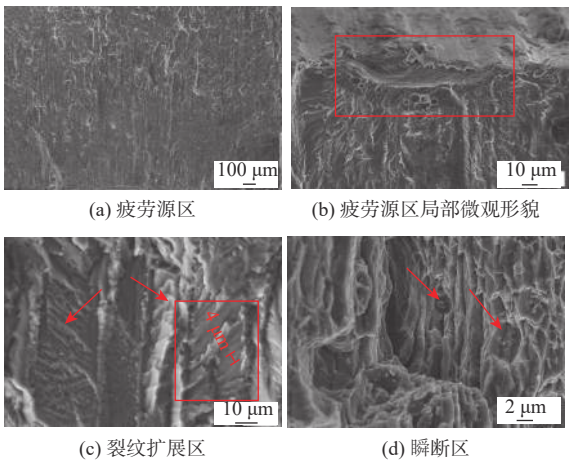
Fig. 6 Servo testing machine

所示，伺服试验机拉伸参数如图 6(b) 所示，试验加载频率为 5 Hz，采用轴向拉-拉加载方式，应力比 R=0.1，正弦波加载。疲劳试验完成后，利用扫描电镜对断面进行形貌观察分析。

2.3 断口分析

试样在疲劳载荷连续作用下最终断裂。试样在工作段发生断裂的断口如图 7 和图 8 所示。铝锂合金的外在表面缺陷、形状与内在的缺陷、晶界强度等在疲劳应力幅值作用下，内部会出现裂纹萌生、扩展融合过程[17]。

图 7 为扫描电镜不同放大倍数下的试样疲劳断口形貌。图 7(a)为疲劳源区照片，可以看出，裂纹呈散射性向周围扩展出去。图 7(b)为疲劳源区放大后的微观形貌，图中有明显的沟槽和放射状条纹。从图 7(a)和图 7(b)中可以看出，在裂纹扩展前期，试样断口形貌存在经典疲劳条纹特征。由于金属内部在疲劳载荷作用下，晶体之间出现位错运

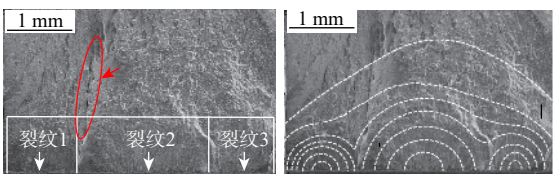


(a) 疲劳源区　(b) 疲劳源区局部微观形貌

(c) 裂纹扩展区　(d) 瞬断区

图 7　扫描电镜下的试样疲劳断口形貌

Fig. 7　SEM micrograph for fatigue fracture surface of specimens



(a) 宏观断口形貌　(b) 裂纹海滩条带

图 8　疲劳拉伸试样断口形貌

Fig. 8　Fracture morphology of fatigue tensile specimen

动, 裂纹在晶体上扩展, 导致位错塞积现象, 表面缺陷和内部缺陷在其影响下, 出现裂纹扩展[18], 萌生后裂纹向应力集中的方向继续扩展, 从而形成放射状条纹。

图 7(c)为裂纹扩展区形貌, 可以看到明显的疲劳辉纹, 且疲劳辉纹是与裂纹扩展方向垂直略呈弯曲并相互平行的沟槽状花样。疲劳裂纹萌生结束后, 裂纹在相邻几个晶粒间进行扩展, 但由于晶粒之间力的作用, 导致裂纹向垂直于疲劳载荷方向进行扩展, 完成多裂纹融合过程。

当多裂纹完成融合扩展到临界值后, 会出现瞬间断裂的现象, 即疲劳裂纹扩展的瞬断阶段, 在该过程中, 裂纹尖端出现极大的应力集中, 裂纹扩展速率急剧增大。图 7(d)为扫描电镜下的试样疲劳断口瞬断区形貌图, 可以看到典型的韧性断裂形貌(韧窝和空洞)。

图 8(a)为试样宏观断裂图, 图中, 河流状条纹标志裂纹的扩展方向, 裂纹路径总体上垂直于加载轴, 河流状条纹中的亮色条带是由于裂纹在不同晶粒之间的扩展引起的。如图 8(a)中箭头所示, 可以看到, 裂纹 1～裂纹 3 之间存在两裂纹相互融合时由相互作用产生的“脊”, 抑制了裂纹之间的相互融合, “脊”是由裂纹相遇时因裂纹扩展方向、扩展能力不同而产生应力集中导致的。当裂纹间的“脊”超过承受极限, 达到裂纹合并的临界条件时, 可将 2 个表面裂纹表征为 1 个大裂纹[19], 出现裂纹的融合现象。

图 8(b)为 2195-T8 铝锂合金疲劳断口图, 图中反映出试样侧向存在 3 个清晰的半椭圆形裂纹, 大小各不相同。从断裂路径和扩展趋势可以看出, 主裂纹是裂纹 2, 次裂纹是裂纹 1 和裂纹 3, 3 个裂纹都沿着板厚垂直方向进行裂纹扩展。在扩展演化的过程中, 具有明显的裂纹融合行为, 可以看出, 裂纹的融合过程具有阶段性, 由于裂纹 2、裂纹 3 的距离相对较小, 作用在裂纹尖端的应力相对较大, 因此先发生融合。在融合过程中, 由于疲劳载荷作用, 裂纹 2 和裂纹 3 有继续向前扩展的趋势, 裂纹之间相互靠近, 当裂纹 2 的右侧裂纹前沿与裂纹 3 的左侧裂纹前沿相遇, 根据裂纹合并准则, 2 条裂纹前沿形成 1 条有继续向前扩展趋势的共面新裂纹。同理, 融合后的裂纹进一步合并裂纹 1, 形成 1 条不规则形状的裂纹前沿, 在疲劳载荷作用下, 继续扩展成 1 条具有椭圆形状的裂纹前沿。最终形成的总裂纹继续沿着裂纹扩展方向进行扩展, 而后形成具有半椭圆形状的海滩条带, 直至试样发生断裂。

3　多裂纹融合扩展仿真

3.1　仿真过程

为分析疲劳裂纹在扩展过程中的动态尖端应力强度因子的变化规律, 以标准疲劳试样为研究对象, 按照实际尺寸, 在 ABAQUS 中建立无裂试样的网格模型, 采用各向同性的线弹性材料, 弹性模量设置为 72.29 GPa, 泊松比为 0.28, 试样一端为固定约束, 另一端施加轴向载荷, 设置应力比为 0.1, 应力幅度为 28～280 MPa, 由于四面体网格自适应性强, 网格类型为 4 节点的 C3D4 单元, 单元数为 56 082, 其完整试样网格如图 9 所示。



图 9　完整试样网格

Fig. 9　Complete sample grid

由于裂纹扩展过程是非线性的, 扩展过程中网格需要不断重划分, 为减少整体网格的规模, 提高计算分析效率, 将完整网格模型进行子模型分割, 截取试样发生断裂的工作段作为裂纹扩展区域进行剖分, 如图 10 所示。

根据试样断口特征和裂纹测量结果, 将 3 种不同初始裂纹插入剖分后子模型的侧面, 其位置依照试样断口形貌设定。设置裂纹 1(A=0.3, B=0.2)、裂

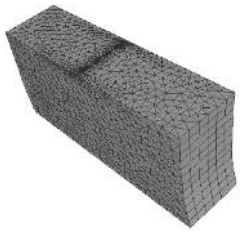


图 10 子模型

Fig. 10 Sub model

纹 2(A=0.5, B=0.4)、裂纹 3(A=0.2, B=0.4), A 和 B 分别为椭圆裂纹的长轴和短轴。不同初始长深比的裂纹 1～裂纹 3 插入后,试样重新划分网格模型如图 11 中第 1 步所示。

基于 Paris 裂纹扩展模型,根据 2.1 节裂纹扩展试验得到的参数, C 和 n 分别取为 1.313×10^{-7} 和 2.802,经过 64 个裂纹扩展步,完成 2195-T8 铝锂合金的疲劳裂纹扩展仿真(含初始裂纹)。在裂纹扩展过程中,裂纹面在各个扩展步中的变化如图 11 所示。扩展步 1～扩展步 5 是裂纹融合前独自扩展阶段;扩展步 6～扩展步 22 是裂纹融合阶段,主裂纹 2 先与裂纹 3 融合后,再合并裂纹 1,形成一条主裂纹;扩展步 23～扩展步 31 中,裂纹面融合后,整体呈现半椭圆型,裂纹扩展趋势沿表面长度方向强于其深度方向,属于裂纹疲劳扩展的稳定扩展阶段;扩展步 31 后,裂纹的扩展方向主要沿深度方向进行,属于裂纹的快速断裂阶段,再经过 33 个扩展步最终断裂,表明 2195-T8 铝锂合金裂纹融合行为会造成裂纹扩展过程中的快速断裂现象。

3.2 仿真结果

在疲劳载荷作用下,疲劳裂纹扩展路径与试验结果的比较如图 12 所示。可以看出,仿真与实际铝锂合金试样在疲劳载荷作用下发生断裂,其裂纹的扩展趋势基本一致,表明疲劳裂纹扩展融合仿真符合试验结果,能够进一步对裂纹扩展过程中裂纹尖端断裂参数变化与扩展过程进行分析。

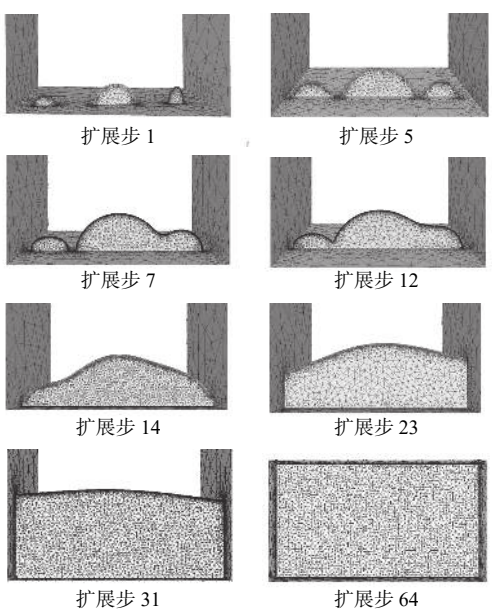
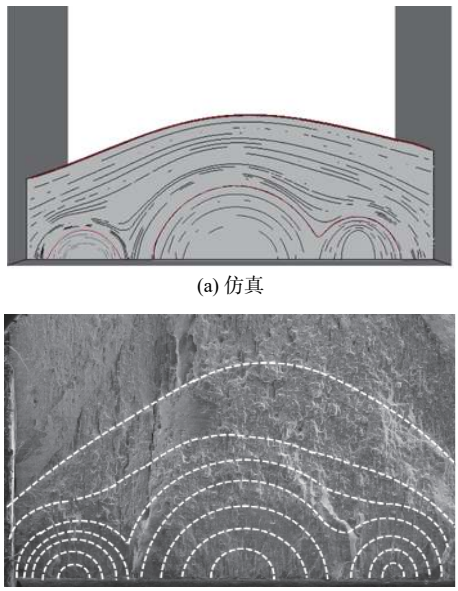


图 11 裂纹面扩展演变

Fig. 11 Crack surface propagation and evolution



(a) 仿真

(b) 试验

图 12 仿真与试验裂纹扩展

Fig. 12 Simulation and experimental crack propagation

3.3 结果分析

3.3.1 裂纹尖端应力强度因子分析

应力强度因子 K 反映了裂纹尖端的应力场,表示裂纹尖端处应力奇异点的强弱程度,是裂纹扩展和断裂的主要控制参量。随着裂纹前沿尖端的不断演变,各前沿点 K 及其分布也不断发生变化[20]。

为更好地分析裂纹在扩展过程中相互融合时应力强度因子随之变化的规律,将仿真结果绘制的各条裂纹前沿分为 3 个阶段,如图 13 所示,对多裂纹扩展融合仿真结果数据进行提取分析,研究裂纹在各个扩展阶段的应力强度因子数值变化情况。

图 13(a)为 3 条裂纹在相遇前疲劳载荷作用下应力强度因子与裂纹尖端相对距离的分布曲线。在裂纹扩展的前 5 个扩展步中,裂纹前沿的应力强度因子随着扩展步的进行不断增大,基本呈现“两端高中间底缓”的形态。

图 13(b)为裂纹 2 与裂纹 3 融合开始到 3 条裂纹完全融合过程中的应力强度因子与裂纹尖端相对距离的分布曲线,可以看出,当裂纹 1～裂纹 3 开始融合的瞬间,裂纹最深处的应力强度因子增长缓慢,而裂纹融合处的应力强度因子急剧增大。随着裂纹的进一步扩展,融合处的应力强度因子降低,与裂纹前沿上的各点应力强度因子处于相对平稳

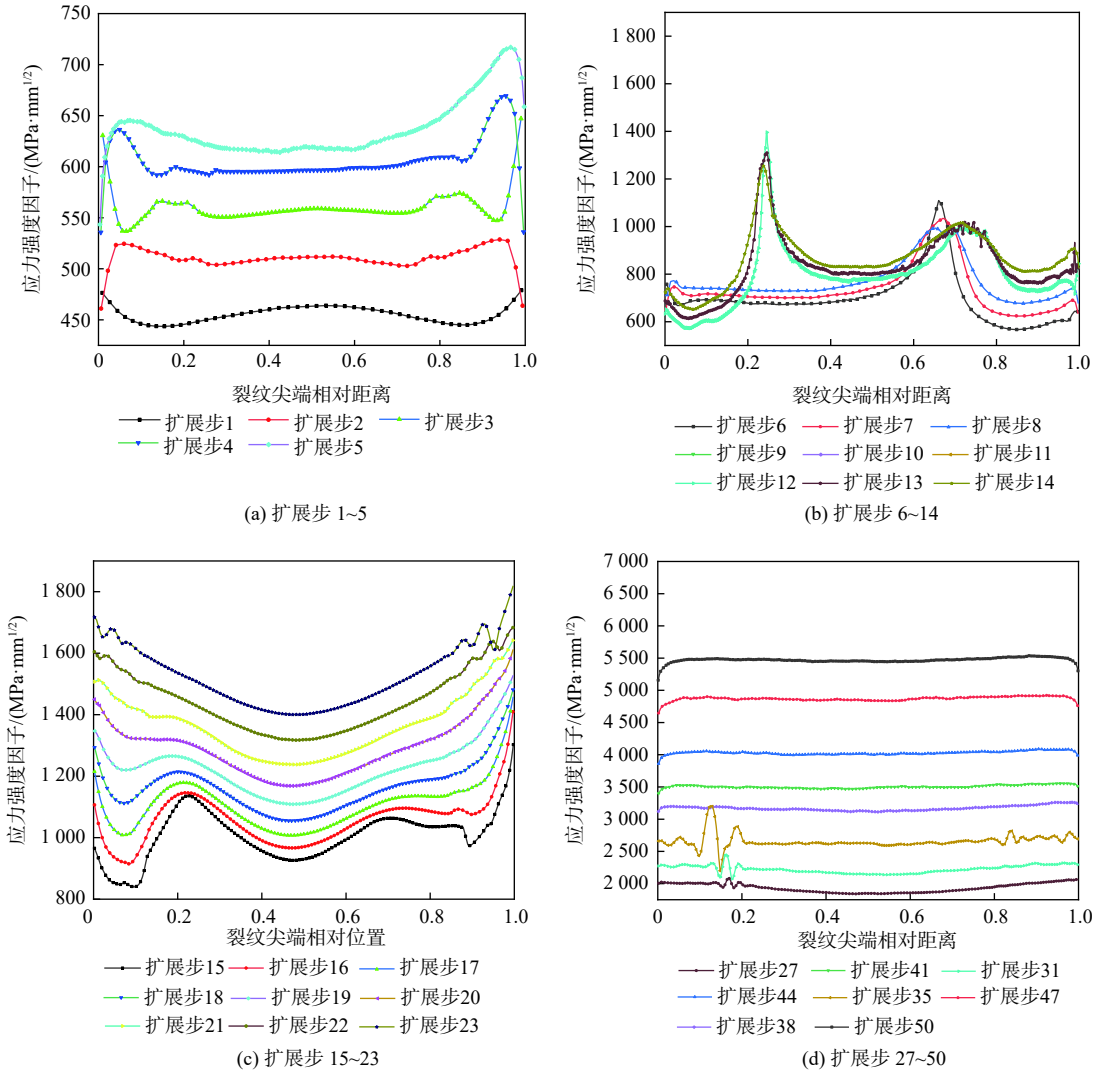


图 13　裂纹尖端应力强度因子

Fig. 13　Stress intensity factor at crack tip

的数值。

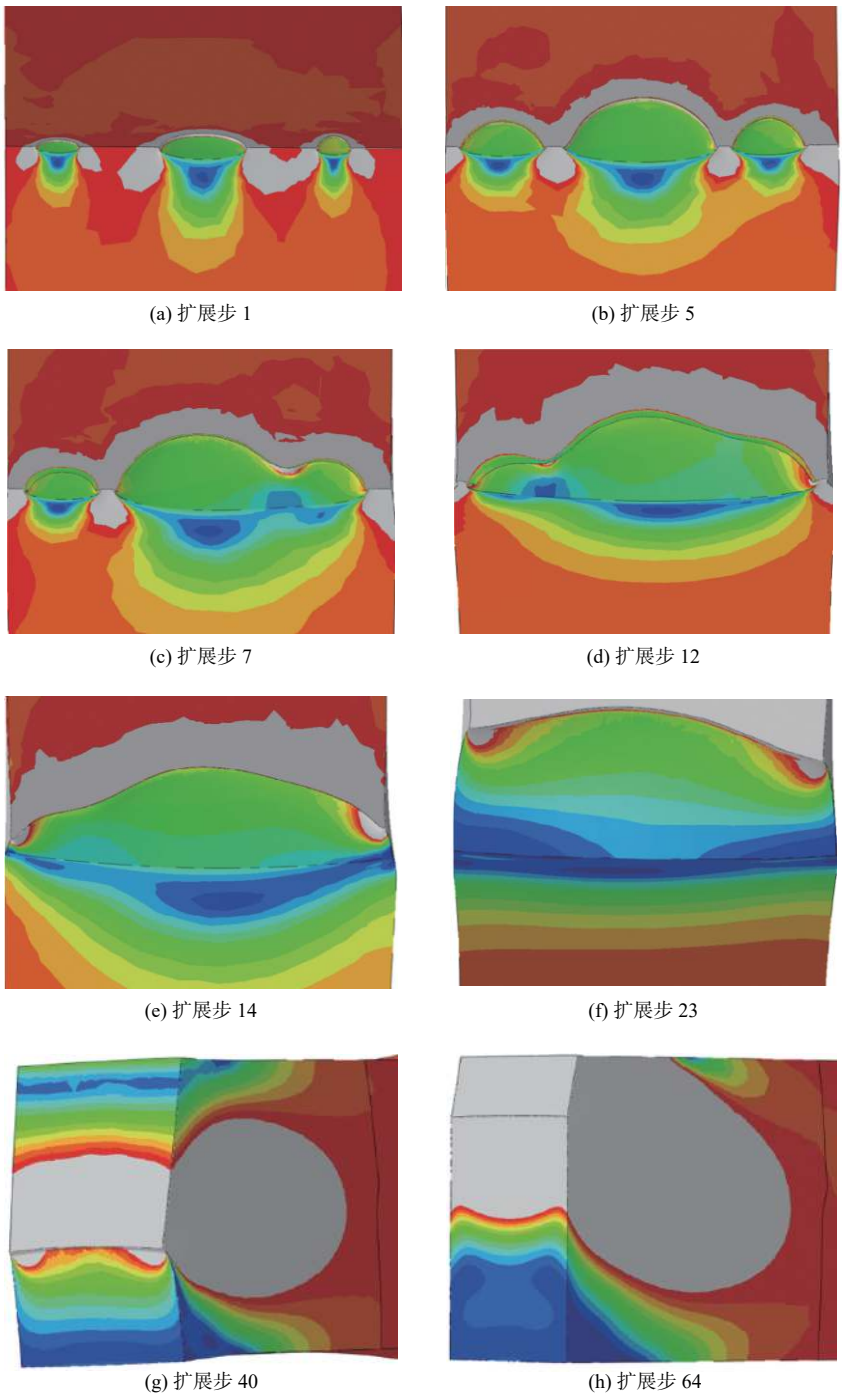
图 13(c)为多裂纹融合后的稳定扩展阶段, 可以看出, 随着融合裂纹的扩展, 应力强度因子的变化趋势基本呈现“两端高、中间低”的形态。这是因为主裂纹在融合完两侧次裂纹后, 在疲劳载荷作用下, 将次裂纹的裂纹前沿扩展到与主裂纹相同的裂纹前沿, 才能维持主裂纹继续向前扩展。在融合次裂纹的过程中, 裂纹 1 和裂纹 3 位置最深处具有较小的应力强度因子, 随着裂纹的进一步扩展, 主裂纹从融合后类似“山峰”形态转变为椭圆形态, 裂纹尖端数值通过不断的拟合变化形成新的裂纹应力强度因子, 裂纹前沿的应力强度因子曲线趋势基本呈凹形分布。

图 13(d)为裂纹融合后的快速断裂阶段, 可以看出, 裂纹融合后的应力强度因子变化曲线在整体上已经类似单个裂纹扩展时的状态, 呈现“两端高、中间低”的趋势, 并且应力强度因子随裂纹扩展的进行而迅速提高, 裂纹的扩展方向主要沿深度方向进行, 直至试样断裂。

3.3.2　裂纹尖端应力场分析

图 14 为预制缺陷处裂纹扩展过程中应力场演变云图, 图中, 灰色部分代表应力数值超过铝锂合金屈服强度的区域。可以看出, 沿裂纹尖端区域应力强度逐渐增加, 且表面裂纹两侧的应力数值最高, 所占高应力区域比例相对最高, 说明由于裂纹尖端两侧强烈的应力集中, 直接导致裂纹沿长度方向的扩展能力高; 随着扩展步的增加, 尖端达到屈服区域的体积大小也随之增加。在第 1 步与第 5 步扩展步应力云图中(见图 14(a)、(b)), 裂纹长度刚超过预制缺陷长度, 三裂纹开始融合前的准备, 能够看出高应力位置沿着裂尖形状分布, 长度方向两侧仍是高应力聚集区域, 而随着裂尖深度方向的扩展, 其应力分布逐渐呈现平整化。

图 14(c)、(d) 展现了三裂纹逐步开始融合, 到最终演化成单裂纹前缘的应力云图。从第 6 步到第



(a) 扩展步 1　(b) 扩展步 5

(c) 扩展步 7　(d) 扩展步 12

(e) 扩展步 14　(f) 扩展步 23

(g) 扩展步 40　(h) 扩展步 64

图 14　裂纹扩展步中应力云图

Fig. 14　Stress nephogram in crack propagation step

13 步应力演化中能够看出，融合后沿裂纹尖端应力更高，屈服区域极大增加，融合位置处出现高应力、超屈服区域。裂纹融合前，塑性区域体积增加区域平缓，裂纹融合后，塑性区域体积急剧增加，其扩展速率呈现先快后慢的规律。

图 14(e)、(h) 展现了三裂纹完全融合后的应力云图。从第 14 步到第 23 步应力演化中，塑性区域持续增长，其增长速度呈现先快后慢的状态，说明裂纹融合扩展后会立即引起更大范围的高应力，并随裂纹扩展的进行而迅速提高；在这种高应力集中和循环疲劳荷载的连续作用下，导致了 2195-T8 铝锂合金最终快速断裂的现象。

图 15 为疲劳裂纹扩展过程中超过塑性强度的单元体积变化过程。裂纹还未完全融合时，即扩展步 12 之前，塑性单元体积近乎相同，仅有微弱增长；当主裂纹 2 分别融合 2 条次裂纹后，塑性单元体积持续增长，其增长速度呈现先快后慢的状态，说明融合裂纹扩展会立即引起更大范围的高应力，并随裂纹扩展的进行而迅速提高；在扩展后期，塑性体积增加缓慢，代表疲劳裂纹扩展的进程已达到

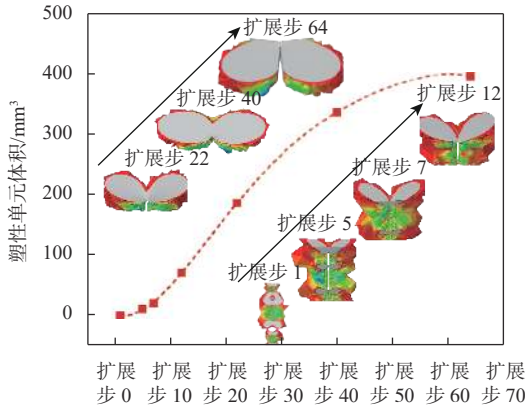


图 15　塑性单元体积变化

Fig. 15　Volume change of plastic elements

了疲劳寿命的末期, 即接近瞬时断裂阶段, 此时, 塑性单元体积大小已逐渐饱和, 使得塑性单元体积在扩展末期呈现缓慢增长的状态。

4　结　论

通过 2195-T8 铝锂合金恒幅拉-拉疲劳试验和多裂纹扩展融合仿真, 分析了疲劳载荷下多裂纹扩展的整个过程, 为 2195-T8 铝锂合金疲劳断裂评估和典型行为研究提供了参考。主要结论如下:

1) 多裂纹扩展初期, 次裂纹会与主裂纹裂纹进行融合, 且无论融合后的裂纹是何形状, 最终融合后的裂纹前沿都会具有椭圆形状。

2) 当 2 个椭圆裂纹相互融合时, 两裂纹在接触点处与其他各处相比, 应力强度因子瞬间增大, 在扩展后期, 裂纹前沿的应力强度因子趋于相近数值。

3) 裂纹融合扩展后会立即引起更大范围的高应力, 裂纹扩展的方向向两侧进行蔓延, 当裂纹前沿接近椭圆曲线趋势时, 才会沿着深度方向进行扩展。

4) 塑性区域体积随疲劳裂纹扩展步数增加而增加。裂纹融合前, 塑性区域体积增加区域平缓, 裂纹融合后, 塑性区域体积急剧增加, 其扩展速率呈现先快后慢的规律。

参考文献（References）

[1] XU Y, WANG X J, YAN Z T, et al. Corrosion properties of light-weight and high-strength 2195 Al-Li alloy[J]. Chinese Journal of Aeronautics, 2011, 24(5): 681-686.

[2] 尹嘉明, 马鹏程, 陈永来, 等. 铝锂合金的发展及在运载火箭贮箱的应用[J]. 兵器材料科学与工程, 2020, 43(3): 137-141.
YIN J M, MA P C, CHEN Y L, et al. The development of Al-Li alloys and its application in launch vehicle tank[J]. Ordnance Material Science and Engineering, 2020, 43(3): 137-141(in Chinese).

[3] 梁准, 张颖晗, 高原, 等. 低温下 2219 铝合金贮箱焊接缺陷的试验与仿真分析[J]. 强度与环境, 2021, 48(5): 52-57.
LIANG Z, ZHANG Y H, GAO Y, et al. Experiment and simulation analysis for weld defected 2219 aluminum alloy tank at low temperature[J]. Structure & Environment Engineering, 2021, 48(5): 52-57 (in Chinese).

[4] 魏国前, 陈斯雯, 余茜, 等. 焊趾多裂纹的试验与仿真分析[J]. 焊接学报, 2019, 40(11): 75-81.
WEI G Q, CHEN S W, YU X, et al. Experimental and simulation study on multiple cracks of weld toe[J]. Transactions of the China Welding Institution, 2019, 40(11): 75-81(in Chinese).

[5] 吴鹏飞, 杨邦成, 岳仕航. AA7075-T6 Ⅰ-Ⅱ复合疲劳裂纹扩展研究[J]. 宇航材料工艺, 2022, 52(1): 82-88.
WU P F, YANG B C, YUE S H. Experimental study on Ⅰ-Ⅱ mixed-mode fatigue crack growth of AA7075-T6[J]. Aerospace Materials & Technology, 2022, 52(1): 82-88(in Chinese).

[6] LI R J, ZHU Y M, FANG W J, et al. Multi-crack interaction and influence on the spherical pressure hull for a deep-sea manned submersible[J]. Journal of Marine Science and Application, 2021, 20(3): 491-503.

[7] 白树伟, 姜楠, 樊俊玲, 等. 疲劳裂纹扩展前缘形貌分析及数值模拟[J]. 科学技术与工程, 2019, 19(1): 256-260.
BAI S W, JIANG N, FAN J L, et al. Morphology analysis and numerical simulation of fatigue crack propagation leading edge[J]. Science Technology and Engineering, 2019, 19(1): 256-260(in Chinese).

[8] 赵颖, 黄美, 杨梦灵, 等. 推进剂贮箱金属焊接件疲劳裂纹对应力强度因子影响的数值分析[J]. 航天器环境工程, 2019, 36(3): 247-251.
ZHAO Y, HUANG M, YANG M L, et al. Numerical analysis of fatigue crack vs. stress impact factor of metal welded joint for satellite propellant tank[J]. Spacecraft Environment Engineering, 2019, 36(3): 247-251(in Chinese).

[9] 徐利辉, 马蒙, 刘维宁, 等. 基于模态叠加法的裂纹板动态应力强度因子计算[C]//第 29 届全国结构工程学术会议. 北京:《工程力学》杂志社, 2020: 295-303.
XU L H, MA M, LIU W N, et al. Calculation of dynamic stress intensity factor of cracked plate based on modal superposition method [C]//Proceedings of the 29th National Structural Engineering Academic Conference. Beijing: Engineering Mechanics Magazine, 2020: 295-303(in Chinese).

[10] AMIRI M, ARCARI A, AIROLDI L, et al. A continuum damage mechanics model for pit-to-crack transition in AA2024-T3[J]. Corrosion Science, 2015, 98: 678-687.

[11] 汪必升, 李毅波, 廖雅诗, 等. 基于扩展有限元模型的动态应力强度因子计算[J]. 中国机械工程, 2019, 30(11): 1294-1301.
WANG B S, LI Y B, LIAO Y S, et al. Calculation of dynamic stress intensity factors based on XFEM model[J]. China Mechanical Engineering, 2019, 30(11): 1294-1301(in Chinese).

[12] CHEN M X, LI Y B, XIA L Y, et al. Effects of pre-rolling on mechanical properties and fatigue crack growth rate of 2195 Al-Li alloy[J]. Journal of Central South University, 2022, 29(3): 836-847.

[13] 郦正能. 应用断裂力学[M]. 北京: 北京航空航天大学出版社, 2012: 111-112.
LI Z N. Applied fracture mechanics[M]. Beihang: Beihang University Press, 2012: 111-112(in Chinese).

[14] 岳应娟, 李宁, 陈飞, 等. 埋藏裂纹应力强度因子的有限元分

析[J]. 武汉科技大学学报, 2015, 38(2): 125-128.
YUE Y J, LI N, CHEN F, et al. Finite element analysis of stress intensity factor of embedding cracks[J]. Journal of Wuhan University of Science and Technology, 2015, 38(2): 125-128(in Chinese).

[15] NEWMAN J C JR, RAJU I S. Stress-intensity factors for internal surface cracks in cylindrical pressure vessels[J]. Journal of Pressure Vessel Technology, 1980, 102(4): 342-346.

[16] 刘城志, 岳应娟, 陈飞. 一种基于FRANC3D的高压气瓶安全性评定的新方法[J]. 铸造技术, 2012, 33(6): 685-687.
LIU C Z, YUE Y J, CHEN F. New method to assess safety of high-pressure cylinder based on FRANC3D[J]. Foundry Technology, 2012, 33(6): 685-687(in Chinese).

[17] 艾书民, 于明, 成晓鸣, 等. 基于Franc3D软件的三维裂纹扩展分析与应用[J]. 机械强度, 2018, 40(1): 251-254.
AI S M, YU M, CHENG X M, et al. Analysis and application of three-dimensional crack growth based on Franc3D[J]. Journal of Mechanical Strength, 2018, 40(1): 251-254(in Chinese).

[18] 蹇海根, 姜锋, 文康, 等. 不同应力下7B04铝合金的疲劳断口[J]. 中南大学学报(自然科学版), 2010, 41(1): 132-137.
JIAN H G, JIANG F, WEN K, et al. Fatigue fracture of 7B04 aluminum alloy under different stresses[J]. Journal of Central South University (Science and Technology), 2010, 41(1): 132-137(in Chinese).

[19] 占勇, 温建锋, 涂善东, 等. 基于疲劳寿命的共面多裂纹合并准则研究[J]. 压力容器, 2016, 33(2): 1-9.
ZHAN Y, WEN J F, TU S D, et al. Combination rules for multiple coplanar cracks based on fatigue life[J]. Pressure Vessel Technology, 2016, 33(2): 1-9(in Chinese).

[20] 吴圣川, 李存海, 张文, 等. 金属材料疲劳裂纹扩展机制及模型的研究进展[J]. 固体力学学报, 2019, 40(6): 489-538.
WU S C, LI C H, ZHANG W, et al. Research progress on fatigue crack propagation mechanisms and models in metal materials[J]. Journal of Solid Mechanics, 2019, 40(6): 489-538(in Chinese).

Experimental and simulation study on fatigue multi crack fusion of 2195-T8 Al-Li alloy

PANG Chuang[1], LIU Dejun[1], TIAN Gan[1,*], GUO Wei[2], LIU Debo[3]

(1. School of Missile Engineering, Rocket Force Engineering University, Xi'an 710025, China;
2. Beijing Institute of Aerospace Experiment Technology, Beijing 100048, China;
3. Beijing Institute of Astronautic Systems Engineering, Beijing 100076, China)

Abstract: As an alloy material widely used in aerospace, the study of fatigue fracture behavior of aluminum-lithium alloy is of great significance to the evaluation of structural safety. The third generation Al-Li alloy 2195-T8 is taken as the research object, and the fatigue crack propagation behavior of 2195-T8 Al-Li alloy is tested and simulated by constant amplitude tensile fatigue test and finite element method. Based on the microscopic measurement and observation of the cross-section, multiple initial cracks were introduced into the simulation model, and then the fusion and propagation process of multiple cracks was simulated. The results show that before crack fusion, under the action of fatigue cyclic load, the value of stress intensity factor at the crack tip increases continuously, and the volume of the plastic region increases gradually; when the cracks fuse with each other, the stress intensity factor at the crack surface increases instantaneously, the stress intensity facto is much higher than the value of the rest of the crack tip. With the further fusion of the crack, the value of the stress intensity factor at the tip tends to be stable; after the crack is completely fused, it reaches the stage of rapid crack expansion, the volume of the plastic region increases in direct proportion to the expansion step, and the growth rate is first fast, then slow; finally, the crack surfaces meet and merge into a new elliptical shape crack surface.

Keywords: fatigue fracture; Al-Li alloys; crack fusion; finite element; stress intensity factor

Received: 2022-04-14; **Accepted**: 2022-06-20; **Published Online**: 2022-06-24 15:59
URL: link.cnki.net/urlid/11.2625.V.20220623.1342.004
Foundation item: National Natural Science Foundation of China (52075541)
* **Corresponding author**. E-mail: tiangan_2012@163.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 405,407 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

《北京航空航天大学学报》
第五届编辑委员会

主　任（主　编）：赵沁平

（以下按姓氏笔画为序）

副主任（副主编）：丁希仑　王少萍　孙志梅　李秋实　李焕喜　杨嘉陵
苗俊刚　相　艳　徐立军　钱德沛　曹晋滨

编　委：马殿富　王　琪　王　聪　邓小燕　王青云　王荣明　刘　宇
刘　红　江　洁　刘　强　闫　鹏　朱天乐　刘铁钢　齐铂金
陈万春　邹正平　苏东林　杨世春　沈成平　邱志平　宋知人
杨树斌　张晓林　杨晓奕　杨继萍　李惠峰　吴新开　张瑞丰
杨照华　宋凝芳　周　锐　林宇震　林贵平　战　强　姚仰平
胡庆雷　赵秋红　段海滨　赵巍胜　席　平　郭　宏　徐　洁
徐世杰　郭洪波　康　锐　翟　锦　熊华钢

北京航空航天大学学报

Beijing Hangkong Hangtian Daxue Xuebao

（原《北京航空学院学报》）

（月刊　1956年创刊）

第50卷 第1期　2024年1月

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

JOURNAL OF BEIJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS (JBUAA)

(Monthly, Started in 1956)

Vol.50 No.1　January 2024

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

ISSN 1001-5965

中国标准连续出版物号：ISSN 1001-5965 / CN 11-2625/V　　国内定价：50.00元／期