

ISSN 1001-5965 CN 11-2625/V

北京航空航天大學

JOURNAL OF BEIJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS

飞行器旋转翼折展过程动稳定性研究进展 甘文彪 左振杰 向锦武 赵忠良 蔡军 马上

4 | 2024 Vol.50

北京航空航天大学学报

第50卷 第4期 (总第374期) 2024年4月

目 次

飞行器旋转翼折展过程动稳定性研究进展…… 甘文彪, 左振杰, 向锦武, 赵忠良, 蔡军, 马上(1053) 机场航班时刻资源管理研究进展 ………………………………………………………… 王艳军,水笑雨,王梦尹(1065) 引入 ESO 的永磁超环面电机无模型预测电流控制 ……………………………………………………………刘 欣, 王 正 阳, 王 晓远(1085) 跨级融合门控自适应网络用于视网膜血管分割……………梁礼明,余洁,陈鑫,雷坤,周珑颂(1097) 基于时频分析的机场跑道平整度评价方法 …………… 齐麟, 杨帅, 解镇州, 金天昱, 黄信(1110) 基于 ECSDNN 的航空安全事件风险等级预测 ………………………………………… 冯 霞 , 桑 潇 , 左 海 超 (1117) 30% HNO3 对 2195 铝锂合金预腐蚀疲劳特性研究 ……… 刘德俊, 田干, 李玉龙, 张炜, 张有宏 (1129) 基于 AESL-GA 的 BN 球磨机滚动轴承故障诊断方法 ………… 王进花,汤国栋,曹洁,李亚洁(1138) 基于装卸顺序的中型机多航段协同配载优化 ……… 赵向领、李云飞、王治宇、徐吉辉、李鹏飞(1147) 北斗三号仅用作追踪的机载设备适航要求分析 ………… 马振洋,周中华,张帆,王鹏,柯炳清(1162) 基于数值延拓的日月综合借力 DRO 入轨策略 ………………………………………………………………… 张 晨 (1176) 空中交通 CPS 结构特性及韧性评估 ……………………………………………… 王兴隆,魏奕雯,贺敏(1187) 基于模型预测控制算法的精确着舰控制方法 ……………万兵,苏析超,汪节,韩维,李常久(1197) 双余度机电作动系统动态力均衡控制方法 ………………… 孙晓哲, 吴江, 石林轩, 杨建忠(1208) 改进反激拓扑模式下远端稳压电源控制方案 … 张帅,韩小文,李儒鹏,程鹏,李正天,王玉姣(1229) 典型立式轻工业建筑结构抗震性能实例研究 … 吴斌,杨悦,姚仰平,王凇晗,姜雪薇,许立言(1240) 嫦娥五号发动机降落羽流扬尘特性研究……张海燕,李思新,王鹢,李存惠,张小平,王卫东(1251) 无参数自适应罚函数的高效代理模型优化设计方法 …………… 张伟, 高正红, 王超, 夏露(1262) RP-3 航油火焰三维温度场和碳烟浓度场测量 ·································· 左冰姆, 周宾, 戴明露(1273) 基于马蹄涡前缘吸气控制的压气机叶栅流动机理………………唐耀璇,刘艳明,安宇飞,孙运政(1282) 可变弯度翼身融合布局气动特性分析与设计 ………… 王雨桐, 蓝庆生, 周铸, 杨体浩, 宋超(1292) 基于图卷积网络的表格隶属关系抽取…………………………………………………………张字童,李启元,刘树衎(1308) 基于伪距残差和新息的 GNSS/IMU 抗差自适应定位算法 …………………………………… 刘 正 午 , 孙 蕊 , 蒋 磊 (1316) 基于均匀圆阵矢量传感器的 DOA 和极化参数联合估计 ……………………………………………………

考虑推力意图的航空器连续爬升垂直剖面预测 杜卓铭,张军峰,苗洪连,王斌(1347)
基于贝叶斯网络强化学习的复杂装备维修排故策略生成
·····································
基于 anchor-free 的光学遥感舰船关重部位检测算法 张冬冬, 王春平, 付强(1365)
面向无人驾驶车辆的行车安全场模型构建方法谢楚安,任羿,杨德真,冯强,孙博,王自力(1375)
碳化材料的烧蚀性能参数辨识方法 程莫瀚,李文光,王志(1384)
复杂多方向威胁下的导弹预警雷达优化部署方法

期刊基本参数: CN 11-2625/V * 1956 * m * A4 * 374 * zh * P * ¥ 50.00 * 400 * 35 * 2024-04

(编辑张嵘孙芳苏磊 下欢欢王茜李亚泰)

JOURNAL OF BEIJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS

Vol. 50 No. 4 (Sum 374) April 2024

CONTENTS

Research progress on dynamic stability of rotating variant wing opening and closing process for aircraft	
GAN Wenbiao, ZUO Zhenjie, XIANG Jinwu, ZHAO Zhongliang, CAI Jun, MA Shang	(1053)
Research progress on airport slot allocation	
WANG Yanjun, SHUI Xiaoyu, WANG Mengyin	(1065)
Ultrasonic longitudinal torsional and low frequency torsional compound vibration tapping experiment	
HAN Fengqi, ZHANG Deyuan	(1077)
Model-free predictive current control for permanent magnet toroidal motor with extended state observer	
LIU Xin, WANG Zhengyang, WANG Xiaoyuan	(1085)
Cross-level fusion gated adaptive network for retinal vessel segmentation	
LIANG Liming, YU Jie, CHEN Xin, LEI Kun, ZHOU Longsong	(1097)
Evaluation method for roughness of airport runway based on joint time-frequency analysis	
	(1110)
Prediction of aviation safety event risk level based on ensemble cost-sensitive deep neural network	
FENG Xia, SANG Xiao, ZUO Haichao	(1117)
Research on pre-corrosion fatigue properties of 2195 Al-Li alloys in 30% HNO ₃	
LIU Dejun, TIAN Gan, LI Yulong, ZHANG Wei, ZHANG Youhong	(1129)
Fault diagnosis method of BN ball mill rolling bearing based on AESL-GA	
WANG Jinhua, TANG Guodong, CAO Jie, LI Yajie	(1138)
Cooperating loading balance optimization for medium-sized aircraft with multiple flight legs based on loading and unloading sequence	
ZHAO Xiangling, LI Yunfei, WANG Zhiyu, XU Jihui, LI Pengfei	(1147)
Analysis of airworthiness requirements of Beidou-3 airborne equipment only used for tracking	
MA Zhenyang, ZHOU Zhonghua, ZHANG Fan, WANG Peng, KE Bingqing	(1162)
Low-energy transfer from Earth into DRO with hybrid gravity assist and numerical continuation	
······ ZHANG Chen	(1176)
Structural characteristics and resilience evaluation of air traffic CPS	
WANG Xinglong, WEI Yiwen, HE Min	(1187)
A precise landing control method based on model predictive control algorithm	
WAN Bing, SU Xichao, WANG Jie, HAN Wei, LI Changjiu	(1197)
Dynamic force equalization for dual redundancy electro-mechanical actuation system	
SUN Xiaozhe, WU Jiang, SHI Linxuan, YANG Jianzhong	(1208)
Robust analysis of hydrodynamic performance under variable rotation speeds	
	(1219)
Improved remote regulated power supply control scheme in improved flyback converter	
	(1229)
Case study on seismic behavior of typical multistory light industrial building structures	
WU Bin, YANG Yue, YAO Yangping, WANG Songhan, JIANG Xuewei, XU Liyan	(1240)

Nozzle plume erosion property on lunar dust in Chang'E-5 mission	
ZHANG Haiyan, LI Sixin, WANG Yi, LI Cunhui, ZHANG Xiaoping, WANG Weidong	(1251)
Efficient surrogate-based aerodynamic optimization with parameter-free adaptive penalty function	
	(1262)
Measurement of three-dimensional temperature and soot volume fraction for RP-3 jet fuel flame	
ZUO Bingxian, ZHOU Bin, DAI Minglu	(1273)
Flow mechanism of horseshoe vortex suction control for compressor cascade	
	(1282)
Design and aerodynamic analysis of blended wing body with variable camber technology	
WANG Yutong, LAN Qingsheng, ZHOU Zhu, YANG Tihao, SONG Chao	(1292)
Tabular subordination relation extraction based on graph convolutional networks	
	(1308)
Robust adaptive position algorithm for GNSS/IMU based on pseudorange residual and innovation	
LIU Zhengwu, SUN Rui, JIANG Lei	(1316)
Joint estimation of DOA and polarization parameters based on uniform circle array with vector sensor	
SHI Chunpeng, HE Huafeng, HE Yaomin, LI Zhen, YAN Shaoqiang, ZHANG Xiaoyu	(1325)
Construction and application of fault knowledge graph for aero-engine lubrication system	
WU Chuang, ZHANG Liang, TANG Xilang, CUI Lijie, XIE Xiaoyue	(1336)
Aircraft vertical profile prediction for continuous climb based on thrust intention	
DU Zhuoming, ZHANG Junfeng, MIAO Honglian, WANG Bin	(1347)
Complex equipment troubleshooting strategy generation based on Bayesian networks and reinforcement learning	
LIU Baoding, YU Jinsong, HAN Danyang, TANG Diyin, LI Xin	(1354)
Ship's critical part detection algorithm based on anchor-free in optical remote sensing	
	(1365)
Construction method of driving safety field model for unmanned vehicles	
XIE Chu'an, REN Yi, YANG Dezhen, FENG Qiang, SUN Bo, WANG Zili	(1375)
Identification method of ablation performance parameters of carbonized materials	
······ CHENG Mohan, LI Wenguang, WANG Zhi	(1384)
Deployment optimization method for missile early warning radar under complex and multi-directional missile threats	
LIU Wei, LIU Changyun, GUO Xiangke, FAN Liangyou, HE Sheng, LAN Hao	(1392)
Study of tensile properties of laminates containing microvascular channels with different diameters	
RAN Guangling, MA Yihao, AN Ziqian, ZHAO Dafang, GUO Xin, CHENG Xiaoquan	(1405)
Multi-strategy fusion improved adaptive mayfly algorithm	
JIANG Yufei, XU Xianze, XU Fengqiu, GAO Bo	(1416)

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

DOI: 10.13700/j.bh.1001-5965.2022.0469

飞行器旋转翼折展过程动稳定性研究进展

甘文彪^{1,*}, 左振杰², 向锦武², 赵忠良³, 蔡军⁴, 马上³

(1. 北京航空航天大学 无人系统研究院,北京 100191; 2. 北京航空航天大学 航空科学与工程学院,北京 100191;

3. 中国空气动力研究与发展中心高速空气动力研究所,绵阳 621000; 4. 湖南云箭集团有限公司,长沙 410100)

摘 要:旋转翼是一种可绕固定轴旋转变体的机翼,广泛应用于新型无人机、巡飞弹、航空炸弹等飞行器,其折展过程中的动稳定特性是决定旋转翼飞行器设计成败的关键基础问题。基于此,梳理了近年来飞行器旋转翼折展过程动稳定机理的研究进展。介绍了旋转翼发展历程及其面临的折展动稳定关键问题;从非定常气动数值模拟、非定常动态特性数值模拟、CFD/RBD一体化耦合数值模拟3个层次阐述了折展过程动稳定数值模拟进展;介绍了旋转翼折展扰动下的非线性动力学建模及动稳定性分析的现状;分析了旋转翼折展动稳定性机理的风洞试验验证情况;总结了旋转翼折展过程动稳定研究面临的科学问题,并提出了可行的研究方向。

关键 词:非定常流;旋转翼;气动运动耦合;动力学机理;动态分析

中图分类号: V211; V212

文献标志码: A 文章编号: 1001-5965(2024)04-1053-12

旋转翼是一种绕固定轴旋转变后掠的机翼,具 有激波减阻、宽速域飞行、结构紧凑、高可靠变体 等优点,其设计理念广泛应用于新型无人机、远程 巡飞弹、便携式导弹、航空制导炸弹等飞行器。然 而,在旋转翼飞行器发射展开和折叠回收(简称折 展,主要是新型蜂群无人机具有回收需求)过程中, 飞行器存在复杂的翼身流动干扰和气动/运动耦合 问题, 蕴藏多种扰动因素影响的稳定性动力学机 理。在飞行器旋转翼折展机动飞行时,存在大气湍 流、非对称转捩和底部干扰等因素。同时,由于机 翼后掠角快速持续变化,加剧了转捩、分离、旋涡 等复杂现象的动态演化,使流动呈现强烈的非定常 特性,给飞行器带来运动耦合下的气动效应。此 时,气动力和力矩呈现非线性特征,且随着机翼折 展,气动参数发生剧烈变化。非对称的气动力矩使 得飞行器出现滚转运动,致使横航向气动力矩显著 增大,纵向压心移动引起俯仰力矩变化,同时受自 身转动惯量的影响,飞行器极易发生滚转运动与俯 仰或偏航运动的叠加,形成气动和运动耦合。这种 耦合现象使旋转翼飞行器的横向运动快速失稳,航 向和纵向稳定性减弱,严重影响旋转翼飞行器的运 动稳定性,甚至产生某些颠覆性后果,造成飞行失 控,危及飞行安全。因此,深入研究飞行器旋转翼 折展过程的复杂流动和气动特性,探索其稳定性机 理,对发展旋转翼飞行器具有重要意义。

目前,飞行器运动稳定性分析通常基于线化小 扰动理论,结合动导数建立气动力和运动模型。针 对旋转翼折展过程气动外形快速变化出现的具有 明显非线性特征的气动/运动耦合问题,传统动导数 概念不再成立,且难以准确表征其运动稳定性,而 应该采用广义的动稳定性参数来描述。事实上,分 析旋转翼折展过程的动稳定性,飞行器运动可描述 为纵横向耦合的旋转翼高频快速折展运动和全机 低频六自由度运动。

针对这类旋转翼折展高频/全机低频的耦合运动情况,本文开展计算流体力学和刚体动力学(compu-

收稿日期: 2022-06-11; 录用日期: 2022-08-19; 网络出版时间: 2022-09-22 19:50 网络出版地址: link.cnki.net/urlid/11.2625.V.20220921.1157.001

基金项目:国家自然科学基金(U2141249,11902018,U2141252);航空科学基金(2019ZA051001)

*通信作者. E-mail: ganhope@buaa.edu.cn

引用格式: 甘文彪, 左振杰, 向锦武, 等. 飞行器旋转翼折展过程动稳定性研究进展 [J]. 北京航空航天大学学报, 2024, 50 (4): 1053-1064. GAN W B, ZUO Z J, XIANG J W, et al. Research progress on dynamic stability of rotating variant wing opening and closing process for aircraft [J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50 (4): 1053-1064 (in Chinese). tational fluid dynamics/rigid body dynamic, CFD/ RBD)一体化数值模拟,分析旋转翼折展过程的流 动机理,并将气动力负载的影响表示为瞬时运动状 态参数及其导数的函数;在此基础上,通过非定常 气动降阶和扰动参数建模分析,揭示折展过程气动 力特性及其受扰运动的稳定性机理,并进一步探析 产生扰动的流动微观特征及其背后的动力学机 理。数值模拟和建模分析的合理性往往通过典型 状态的风洞对比试验来实现流动机理与运动稳定 性机理验证,因此,本文从飞行器旋转翼发展历程 及其折展动稳定问题入手,重点阐述与旋转翼折展 过程动稳定相关的 CFD/RBD一体化耦合数值模 拟、动稳定建模分析和风洞对比试验验证等研究现 状,对折展过程动稳定科学问题形成总结,为研究 提出可行的方向。

1 旋转翼发展历程及折展动稳定问题

旋转翼飞行器布局概念在20世纪40年代提 出,此后设计了多种旋转翼飞机概念方案。20世 纪 70 年代, Jones^[1] 对旋转翼布局飞行器的阻力进 行了深入细致的研究,根据高速空气动力学理论提 出高速飞行时,兼顾左右对称性和前后对称性的体 积分布阻力最小,揭示了旋转翼布局相比于普通后 掠翼布局具有优越的减阻特性^[2];同时, Holdaway 也通过跨声速的风洞试验验证了斜置翼具有优越 于后掠翼的减阻特性^[2]。在这些研究的基础上, 1979年, NASA 开展了旋转翼飞行器 AD-1 计划, AD-1飞机及其旋转翼折展示意如图1所示,该飞 行器的机翼仅采用一套中央变后掠机构,结构质量 更轻,后掠角变化范围达到 0°~60°,当后掠角为零 时,可以保证良好的起飞特性及低速飞行时的升力 特性,高速飞行时,则通过改变旋转翼的旋转角来 降低飞行阻力,从而改善飞行器的飞行品质,有效 扩大飞行器的飞行包线。AD-1 计划成功验证了旋 转翼飞机低速飞行的可行性,但在试飞中发现,当 左侧机翼向前偏至某角度时,机翼负载陡增,变形 严重,并导致飞行品质急剧恶化,最终计划中止。 AD-1是旋转翼飞行器发展史上一个重要的里程 碑,为斜置旋转翼折展研究提供了宝贵的数据和经 验,但该计划并没有将旋转翼变体飞行器的真正优 势完全体现出来。

近年来,随着无人飞行器研发的进步,旋转翼 技术得到了快速发展,并越来越多地应用于新型蜂 群无人机、远程巡飞弹和航空灵巧炸弹等飞行器。 典型的,如美国"蝎子"轻型灵巧制导炸弹和"小精 灵"蜂群无人机计划,典型旋转翼飞行器示意如图 2



图 1 AD-1 飞机及其旋转翼折展示意图 Fig. 1 Schematic diagram of folding and spreading of its rotating variant wing the AD-1aircraft as well as the rotating variant wing



(a) "蝎子"发射前旋转翼折叠和发射后展开



(b)"小精灵"发射前旋转翼折叠和发射后展开



(c)"小精灵"回收前旋转翼展开和回收折叠

图 2 典型旋转翼飞行器示意图 Fig. 2 Schematic diagram of typical rotary wing aircraft

所示。"蝎子"采用上单弹翼和X尾翼布局,其弹翼 在弹体上方可旋转折展;"蝎子"挂在发射架或装 在发射筒中时,弹翼和弹体平行;高速发射后,"蝎 子"的弹翼快速旋转至与弹体呈90°夹角,大大提高 了射程,能够实现灵巧制导高亚声速和超声速攻 击。"小精灵"蜂群无人机采用下单弹翼和X尾翼 布局,其弹翼在弹体下方可旋转折展,弹体上方具 有回收对接装置,目前,"小精灵"蜂群无人机计划 的研究还处于发射和回收的飞行试验验证阶段。

综合旋转翼的发展历程,旋转翼折展稳定飞行 可概括为3类情况:①以典型模态切换特征的斜置 旋转翼折展和展开过程;②以高亚声速和超声速飞 行为主的旋转翼展开过程;③以发射回收为特征的 亚声速旋转翼展开和折叠过程。

针对折展稳定飞行情况,旋转翼折展过程主要 面临气动机理不明、动力学建模方法有待深化、试 验方法有待发展等问题。 1)旋转翼的折展过程,必然带来飞行器质量分布的持续变化和绕流流动的剧烈变化,从而产生固定翼飞行器所不具备的机翼旋转运动而带来的部件之间的气动干扰,以及旋转变形过程中由于飞行器左右不对称而带来的纵/横向耦合干扰,进而影响飞行器整体的运动稳定性,同时,这些特有的现象还会随着旋转翼折展速率等参数的不同而产生变化,需要综合风洞试验、数值计算和建模仿真等多种方法,揭示其中机理。

2)旋转翼折展过程,飞行器存在大气湍流、非 对称转捩和底部干扰等因素,同时,机翼后掠角快 速持续变化,加剧了转捩、分离、旋涡等复杂现象 的演化,这会形成复杂飞行扰动,使飞行器气动/运 动耦合非线性动稳定特性;而针对具有强物理背景 的飞行扰动和非线性动稳定分析,目前并不具备通 用的建模方法。

3)目前,常规风洞试验、模型及其部件是相对静止的,难以真实模拟旋转翼折展过程中机翼运动带来的非定常复杂绕流流动,获取的气动特性不真实,针对旋转翼折展过程的实际模拟需求,需要发展静动态结合的试验模拟方法。

2 旋转翼折展非定常数值模拟现状

旋转翼折展为快速旋转运动过程,飞行器具有 复杂非定常特征:①流动包含多尺度的旋涡运动, 旋转翼后掠角快速持续变化加剧了旋涡运动与破 裂;②流场是多频甚至耦合的,往往具有不同频率 的振荡运动及耦合运动,如旋转翼快速折展非对称 前缘涡和机身高频分离涡的耦合;③流动是非定常 演化的,随折展参数变化,由周期振荡解逐渐演变 为多周期、准周期或混沌运动。

针对旋转翼折展过程开展数值模拟,是评估分 析飞行器非定常动态特性及其机理的基础,其研究 现状体现在3个层次:以基本气动特性和流动特征 分析为主的非定常气动数值模拟、以变可信度动稳 定参数评估为主的动态特性数值模拟、以气动和动 态模拟为基础的耦合旋转翼折展的 CFD/RBD 一体 化耦合数值模拟。

2.1 非定常气动数值模拟

针对飞行器旋转翼折展过程进行气动分析,其 核心问题是如何高效高精度地实现非定常气动数 值模拟,非定常气动数值模拟的关键技术有时间推 进方法和湍流模拟。

在时间推进方面,双时间步方法是目前时间精 度计算非定常问题的最广泛、最有效的方法。该方 法一经提出就备受关注,并迅速用于非定常数值模 拟。双时间步方法在冻结的真实时刻上引入类牛顿迭代的子迭代过程,通过子迭代过程弥补近似处 理带来的时间精度损失,达到非定常计算时间精度 要求。双时间步方法原理简单且易编程实现,物理 时间步长的选取不受稳定性条件的限制,可依据物 理问题的精度要求来确定,在子迭代过程中可采用 各种定常计算的加速收敛技术;同时,双时间步适 合引入预估校正方法,在旋转翼折展时气动/运动耦 合求解方面具有较好的潜力。

在非定常湍流模拟方面,依据精确度可分为直 接数值模拟 (direct numerical simulation, DNS)、大涡 模拟 (large eddy simulation, LES)、湍流模式理论雷 诺平均数值模拟 (Reynolds average Navier-Stockes, RANS) 及介于 RANS 和 LES 之间的混合方法, 如脱 体涡模拟 (detached-eddy simulation, DES)^[3]。相比传 统 RANS 方法, DES 类方法的一大优势是对大分离 问题特别有效,且网格需求量又远远小于 LES 方 法,给工程实际问题提供了解决之道,得到了广泛 的应用。但针对旋转翼折展这类复杂高动态且计 算量过大的流动问题, DES 类方法难以兼顾动网格 技术插值精度需求,无法获得足够精细的转捩、激 波等流动特征,目前应用最多的仍是非定常 RANS 方法。大量的工程实用表明,发展一种对任何流动 都具有高精度模拟能力的湍流模型几乎是不可能 的,但针对旋转翼折展过程发生分离、转捩和激波 等的特殊流动,从应用最广的湍流模型着手改进仍 是可能的。因此,近年来,从湍流模型着手,进行分 离流和可压缩性修正³³,并引入具有改进的转捩输 运方程,综合改善高频湍流转捩和低频脱体涡模拟 能力对提高旋转翼折展过程湍流模拟精度和效率 具有较好的参考价值。以 SST 湍流模型为例的转 捩改进如下:

$$F_{\text{length}} = \min \left((0.1 \exp \left(0.022 \overline{Re}_{\omega} + 12 \right) + 0.45, 300 \right)$$
(1)

$$Re_{\theta t} = 803.73(Tu + 0.606\ 7)^{-1.027}F_{\lambda,K}$$
(2)

$$F_{\lambda,K} = \begin{cases} 1 + F_{\lambda} e^{-Tu/3} \lambda \leq 0\\ 1 + F_{K} \left(1 - e^{-2Tu/3} \right) + 0.556 \cdot \left(1 - e^{-23.9\lambda} \right) e^{-Tu/3};\\ \lambda > 0 \end{cases}$$

$$\begin{cases} \frac{\partial \overline{Re}_{\theta t}}{\partial t} + u_j \frac{\partial \overline{Re}_{\theta t}}{\partial x_j} = P_{\theta t} + \frac{\partial}{\partial x_j} \left[\left(v + \frac{v_t}{\sigma_{\gamma}} \right) \frac{\partial \overline{Re}_{\theta t}}{\partial x_j} \right] \\ \frac{\partial \gamma}{\partial t} + u_j \frac{\partial \gamma}{\partial x_j} = P_{\gamma} - E_{\gamma} + \frac{\partial}{\partial x_j} \left[\left(v + \frac{v_t}{\sigma_{\gamma}} \right) \frac{\partial \gamma}{\partial x_j} \right] \end{cases}$$
(4)

式中: *F*_{length}为修正的经验函数关系式, 通过其控制转捩区长度; *Re_{th}为当地动量*雷诺数, 采用 Keerati

给定的函数来确定^[4]; \overline{Re}_{θ_t} 为转捩动量雷诺数; γ 为 间歇因子; Tu 为湍流度; P_{θ_t} 为转捩动量雷诺数输运 方程产生项; P_{γ} 为间歇因子输运方程产生项; E_{γ} 为 间歇因子输运方程破坏项; u_j 为速度; $v \pi v_c$ 分别为 层流和湍流运动粘性系数; $\sigma_{\theta_t} \pi \sigma_{\gamma}$ 分别为捩动量 雷诺数和间歇因子输运的扩散项常数。

2.2 非定常动态特性数值模拟

旋转翼折展是一个动态过程,面临复杂的动态 气动力和动稳定性问题,飞行器的气动力会随着飞 行姿态的改变而呈现非定常、非线性的变化,气动 力不仅取决于当前飞行姿态及其变化率,还依赖于 流动变化的历史过程,气动力系数的变化也不具有 导数特征,传统上基于小扰动理论提出的动导数概 念不适合表征飞行器的动稳定性,可采用动稳定性 参数或组合参数来描述飞行器的非线性动态特 性。但由于气动建模是动稳定性分析必不可少的 基础,而气动建模往往以气动力系数和导数作为基 本参数,因此,动导数在气动建模的参数方面也是 不可或缺的,同时还需关注动导数的非线性或其相 关的组合参数。

目前,获取动稳定参数的主要方法有工程估 算、数值模拟、风洞试验和飞行试验。工程估算主 要是经验、半经验方法,亚声速以细长体理论、升 力面理论为主^[5],超声速、高超声速主要是基于牛 顿理论或活塞理论发展而来的方法^[6-8]。工程估算 具有简单、高效的特点,并且结果具有一定精度^[9], 但其对复杂外形适应能力较差且不能预估黏性效 应的影响。

近十几年来,采用非定常数值模拟强迫振动 或自由振动求解动导数成为主流,国内外涌现了大 量该方面的研究工作^[9-14]。其中,任玉新等^[10]、刘 伟^[11]、袁先旭^[12]等在中国较早采用非定常 CFD 方 法模拟强迫振动求解动导数,数值模拟了尖锥、钝 锥、导弹、弹道外形等典型外形的强迫振动流场, 并辨识其静、动稳定性参数。范晶晶等^[15]尝试了 有翼导弹外形大攻角下动导数的模拟。郭东等^[16] 应用网格速度法模拟飞行器俯仰振荡求解动导 数。此外,还有部分研究将时间谱方法应用于非定 常流场的求解,采用频域方法模拟具有明显周期特 征的强迫振动过程,辨识静、动稳定性参数^[17]。

综合来看,绝大部分数值计算的是三通道阻尼 导数,其与旋转翼折展相关的重要阻尼参数(包括 交叉导数、洗流时差导数及马格努斯力矩系数导数 等)较少涉及。相比主阻尼导数,其他数的计算难 度较大,不确定度较高。由于这些动稳定参数的量 值较小^[18],需要数值模拟不仅能捕捉旋转翼折展过 程的流动非对称现象,还要能够精准计算表面摩擦 阻力。目前,对于旋转翼折展这种刚体运动状态, 难以从理论上给出动稳定参数计算不确定度的估 计方法,动稳定参数的描述和计算工作还需加大研 究力度。

2.3 CFD/RBD一体化耦合数值模拟

在旋转翼折展过程中,飞行器的空气动力与旋转翼折叠和展开、全机机动运动是密不可分的, CFD/RBD一体化耦合数值模拟成为飞行器动态研究的必然选择。CFD/RBD一体化耦合数值模拟技术常用于预测飞行器动态特性^[19],可以提供接近实际的飞行运动轨迹、姿态、气动参数及丰富的流场信息^[20],分析飞行动稳定性、极限环振动参数(频率和振幅)及分岔参数^[21],直观反映飞行品质。与强迫振动法单纯获得动导数相比,一体化方法可以进行物理机理及动力学机制方面的研究^[22-23],还可以进行多自由度的计算,有效弥合了纵横向的多自由度耦合效应。

CFD/RBD一体化耦合数值模拟涉及流体控制 方程和运动模型方程的耦合求解。在数值模拟过 程中,通过求解流体控制方程可以获得气动力、力 矩,积分运动模型方程可以获得位置、姿态和速 度、加速度,二者交替提供参数,互为输入和输出, 逐步沿时间推进,从而模拟整个运动过程。CFD/ RBD 耦合形式有松耦合和紧耦合 2 种。紧耦合算 法理论上流动和运动的误差同时收敛,可消除流动 与运动求解过程中的时间滞后,但需要刚体运动方 程参与内迭代,为达到二阶时间精度,需要反复迭 代多次并频繁生产动态网格,求解计算量较大,动 态网格鲁棒性较低,目前实际应用较少。松耦合算 法交替求解流动控制方程和刚体动力学方程,简单 且易于实现,但由于2个子系统积分时间不同步 (两者存在一个时间步长的延迟),为保证整个系统 的稳定性和计算准确性,往往需要采用较小的时间 步长。对于气动/运动的紧耦合并行计算,考虑到需 要在双时间步的内迭代步中对流动和运动控制方 程进行同步迭代^[24],二者数据交换频繁,同时计算 方法一经确定后很少改动,因此,直接在原始 CFD 代码中添加运动求解代码,并进行整体编译,可以 有效提高计算效率[25]。图3给出了气动/运动耦合 并行计算流程,图4为该流程对应计算的网格重叠 和流场示意图。

目前,国内外结合重叠网格或非结构重构等动 网格技术,CFD/RBD一体化耦合数值模拟广泛应 用于飞行器抛罩、抛壳、抛弹、抛副油箱、飞行器 级间分离、助推分离、子母弹抛撒等多体运动问





Fig. 3 Flow chart of aerodynamic/motion coupling parallel computing^[24]



图 4 基于网格重叠的计算网格和流场示意图 Fig. 4 Schematic diagram of calculation grid and flow field diagram based on grid overlap

题。在面向旋转翼折展的非对称大幅度复杂动态 问题时,相比弹性网格、非结构重构等方法,基于重 叠网格的动网格技术具有更高的精度优势。然而, 重叠网格技术在鲁棒性、效率和精度等方面仍有进 一步提升的需求,特别是对分布式装配技术的研 究,明显滞后于 CFD 并行求解器,需要发展高效鲁 棒、自动化程度高、与并行求解器性能匹配的重叠 网格分布式装配算法。

3 旋转翼折展稳定性动力学建模分析

在旋转翼折展过程中,飞行器受扰运动,基于 数值模拟的数据参数,需要开展折展扰动下的非线 性动力学基本建模、扰动参数建模及受扰全过程的 动态稳定性分析研究。

3.1 旋转翼折展过程的非定常气动建模

针对旋转翼折展气动/运动强耦合可能带来的 不稳定、强非线性问题,非定常气动建模是研究旋 转翼飞行器动态特性的重要支撑技术。

传统的气动建模由早期预估瞬时运动参数变 化和载荷的线性关系发展而来^[26],采用以稳定性导 数为基础的线化模型,难以对非定常气动力特性进 行严谨表达。为此,Tobak^[27]提出了非线性阶跃响 应的非定常气动建模方法,从数学上对气动力进行 了严密且通用的表达,但由于建模采用积分形式, 与微分形式的运动方程耦合困难,难以用于分析飞 行品质。为改进这些不足,Murphy等^[28]指出气动 力系数对阶跃输入的响应呈指数衰减规律,发展出 新的通用积分模型,并用于战斗机非定常气动力的 预测,取得了较好的成效。为进一步克服积分模型 的使用困难, Goman 等^[29] 发展了一种新微分形式的 非定常气动力模型——状态空间模型,该模型建模 过程引入流场内部状态变量(如流动分离点位置), 将气动导数表达成内部状态变量的函数,成功用于 多种飞行器大迎角非定常气动特性的预测,其与传 统线性模型形式类似,可用作飞行品质分析。在此 基础之上,状态空间模型经过进一步推广,使气动 载荷与物理流动现象关联,提出了时间常数,完成 了三角翼横侧向的非定常气动力建模,模型与试验 结果吻合很好^[30-36]。为提高状态空间模型效率.汪 清等^[30]将非线性非定常效应的气动力增量表达为 一阶微分方程,提出了改进的非定常气动力状态空 间模型,并成功应用于F-18等飞机的大迎角非定常 气动力分析。与此同时,基于流场特征的新分析方 法得到了快速发展,主要包括本征正交分解 (proper orthogonal decomposition, POD) 方法^[31-33] 和谐波平 衡 (harmonic balance, HB) 方法^[34-36]。基于流体计算 模型的输入响应获得正交基,模型的状态量可视为 该组正交基的线性叠加,通过增加模型阶数更准确 地模拟实际流场运动特性。这类方法通常在流场 分析方面具有优势,但由于考虑整个流场状态特 性,模型较为复杂,阶数较高,不便用于全机复杂动 态分析。

近年来,随着智能科学和多学科交叉融合的迅速发展,非定常气动力建模领域大量引入了智能算法,如 Volterra 级数方法^[37-38]、线性状态空间模型^[39-40]、自回归移动平均 (auto regressive-moving-average, ARMA)模型^[41]、基于人工神经网络的替代模型^[42-43]、径向基函数 (radial basis function, RBF) 插值^[44-45]和Kriging模型^[46-47]、支持向量机 (support vector machine, SVM) 方法^[48]、极限学习机 (extreme learning machine, ELM^[49])等。依据这些算法的非定常气动力建模并不关注内部函数关系,将气动力参数和状态量之间的关系看作"黑箱",通过训练高质量的样本数据

来建立非定常气动力模型。因此,这类建模方法可称为系统辨识方法。系统辨识方法虽然能精确地预测样本值范围内的结果,但其外插能力一般,当系统参数较多时,计算量是极其巨大的,甚至可能导致无法辨识。相比传统方法,这些智能算法模型以"样本数据"为根基,并不关注流场特性,容易缺乏明确的物理背景支撑。

综合以上非定常气动建模现状,针对旋转翼折 展时的全机动态特性,系统辨识方法具有更好的预 测精度,但需要优化、提高效率,并结合动力学仿真 分析来强化其物理背景。

3.2 旋转翼折展过程的扰动参数建模

在旋转翼折展过程中,除宏观的全机动态特性 之外,微观的复杂流动带来了飞行参数的扰动,使 飞行器存在气动建模物理背景弱、扰动影响和机理 不明的问题。因此,旋转翼折展过程的扰动参数建 模分析越来越受到关注。现阶段,主要研究扰动参 数建模^[50-51]及流动机理分析^[52]、基于智能算法的随 机型扰动参数建模^[53-55]。

开展扰动参数建模及流动机理分析,将深度探 析产生扰动的流动微观特征及其背后的动力学机 理。比较有前景的方法是采用长短时记忆 (long short-term memory, LSTM) 神经网络建立流动特征 与飞行扰动参数关系模型^[53-54],提取转捩位置、湍 流区大小、激波位置、分离区位置等流动特征作为 输入变量,输出变量为飞行器扰动幅度、频率、姿 态角及其变化速率等飞行扰动参数,利用深度学习 建立输入和输出之间的非线性数学关系。基于 CFD 高精度求解器计算获得包含输入变量和输出 变量的训练样本、测试样本,基于训练样本确定深 度学习参数,基于测试样本评价模型在训练过程中 的表现及作为最优模型的选择标准,建立流动特征 与飞行扰动参数关系的 LSTM 神经网络模型。基 于深度学习的流动特征与飞行扰动参数建模分析 过程如图 5 所示。



图 5 基于深度学习的流动特征与扰动参数建模过程

Fig. 5 Flow characteristics and disturbance parameter modeling process based on deep learning

在扰动幅度、频率、姿态角及其变化速率等参数影响下,旋转翼折展动力学模型受扰变化,针对扰动参数建模,重点在于开展气动降阶模型修正。 基于修正后的动力学模型,分析评估扰动参数对飞 行运动稳定性的影响机理。

在气动降阶模型修正时,典型的方式是采用极 限学习机,开展随机型气动降阶建模,并基于扰动 参数完成降阶模型修正^[55]。气动降阶模型由原理 的自变量转化为随机自变量,输出变量仍为气动导 数,建立随机型气动降阶模型;基于扰动幅度、频 率、姿态角及其变化速率,开展基于扰动参数的降 阶模型修正。建模全过程如图 6 所示。

考虑扰动幅度、频率、姿态角度变化,随机型 气动降阶模型中的参数表达为扰动幅度、频率的多 项式函数,针对这些函数关系式的参量需要进一步 辨识,现阶段辨识主要采用优化算法,常用的有粒 子群、遗传算法等。

3.3 旋转翼折展动稳定性分析

旋转翼折展是飞行器变体的过程,其稳定性分 析问题始终是旋转翼飞行器研究的焦点,是关系飞



图 6 基于极限学习机的随机型气动建模过程

Fig. 6 Process of stochastic aerodynamic modeling based on limit learning machine

行器是否可控、如何控制的基础性问题。相比固定 翼飞行器,变体飞行器存在折叠、旋转等变形过程, 该过程中流动特性复杂,存在非对称转捩干扰、激 波诱导的边界层分离、旋涡的运动与破裂等复杂现 象^[9],飞行器的本体动力学特性也会随着飞行器的 变形而改变,两者相互作用会诱发严重的气动/运动 非线性耦合干扰,由此引发的稳定性问题受到国内 外长期持续关注,主要是在气动和扰动参数建模的 基础上,针对飞行器变体的非线性动稳定性分析及 控制器设计进行了大量研究。

乐挺等^[56-59]推导了Z型翼飞行器纵向运动动 力学方程,采用 CFD 方法通过准定常分析了 Z 型 翼折叠过程的动力学特性。Henry^[60]和殷明^[61]等基 于 Jacobi 线化方法,针对变翼展变体飞行器进行了 线性参变系统的建模,分析了变形过程中系统的变 化特性。Henry^[62]基于小扰动线性化方程,研究了 伸展翼变形过程机翼伸缩对飞行器滚动、螺旋和荷 兰滚模态的影响。金鼎等^[63]分析了折叠机翼变形 过程飞行器纵向运动的稳定性。Beaverstock^[64]分 析了变体飞行器的动态性能,并优化了飞行器的折 叠翼翼型和飞行控制系统。Seigler等[65-66]基于慢 变系统理论, 通过构造参数依赖的 Lyapunov 函数, 分析了变展长、变后掠和变尾翼速率对系统瞬态稳 定性的影响,给出了变体过程满足一致稳定的变体 速率上界。Shi和 Wan^[67]建立了显含变体速率的飞 行器动力学方程,通过求解方程特征值,得到了系统 稳定的变体速度上界。张杰、吴森堂[68]和李文成[69] 将变展长、变后掠角的变体飞行器简化为多刚体系 统,采用 Kane 方法建立动力学模型,分析了机翼变 体运动对飞行器动力学的附加影响。杨贯通等^[70] 采用 Newton-Euler 方法建立了变后掠变展长飞行器 的五刚体动力学模型,推导了变体产生的附加气动 力和力矩,基于纵向解耦简化和准定常假设,分析 了不同变体形式和速度下飞行器的纵向动态响应。

旋转翼飞行器折展过程相比传统变体飞行器 具有变形模式相对简单、模态间耦合较大等特点。 传统变体飞行器的动态性能研究方法虽在旋转翼 折展过程的研究中具有一定的指导意义,但由于旋 转翼的折展过程是一个强非线性系统,飞行器的动 态特性会随着旋转翼折展角度的不同存在很大差 异。现有的动态响应研究方法多将整个折展过程 看成一个慢时变系统进行研究,不能完全反映多个 阶段的响应特性,对后续控制系统的设计提出更大 挑战。针对旋转翼折展过程的强非线性特点,类比 多模型自适应控制理论,需要研究基于多模型的旋 转翼飞行器动态响应与模态分析技术,实现多个飞 行模态下的响应分析。

4 旋转翼折展风洞试验现状

现阶段,风洞试验仍是旋转翼折展过程动稳定 机理验证最主要的研究手段,能够与数值模拟相互 校核对折展典型状态的动稳定性进行分析。国内 外的风洞试验主要针对固定翼航空航天飞行器,建 立了较为成熟的试验技术,主要包括研究基本气动 特性的测力、测压等各类静态试验,获得模型在固 定马赫数、雷诺数、迎角等试验状态下的气动力特 性、舵面效率、纵横向静稳定性及表面离散点的压 力分布等基本气动特性。

针对旋转翼折展进行试验研究,直接相关的有 飞行器全机变体和旋转翼折展2类试验。在飞行 器全机变体方面,长期以来,主要针对对称翼布局 飞行器开展翼面折展有效性验证试验。利用高速 摄影手段获得飞行器对称翼的折展过程影像,利用 风洞流场条件检验不同雷诺数、马赫数、迎角等飞 行状态下作动机构的结构合理性和有效性。在旋 转翼折展研究方面,集中于低速研究,而面向高低 速旋转翼折展过程的复杂流动机理、动态气动特性 试验、动态失稳特性试验,国内外有一定的基础储备。

在飞行器动态气动特性试验研究方面,以小扰 动线化理论建立的强迫振动动导数试验技术为主, 具有代表性的动导数试验设备包括美国 AEDC-PWT 4T 和 16T(16S)风洞中的大迎角、大载荷俯 仰、偏航和滚转动导数试验设备,可以测量模型的 俯仰、偏航、滚转阻尼导数、交叉导数和交叉耦合 导数。贺中等^[71]建立了覆盖 0.6 m 量级、1 m 量级 到 2 m 量级的细长体亚跨超声速动导数试验技术。

在动态失稳特性试验研究方面,适用于低速试 验具有代表性的有:美国 Bihrle 应用研究所的多轴 试验系统 MAT,德国 DNW-NWB 的六自由度动态 试验系统 MPM,英国 Bristol 大学的六自由度动态 试验装置^[72],俄罗斯 TsAGI 的俯仰、滚转和偏航三 自由度动态试验系统,史志伟等^[73]在1m非定常风 洞中建立的五自由度动态试验装置,邓学蓥等^[74] 在 D4 低速风洞中建立的大振幅振荡试验装置;在 高速动态失稳特性研究方面,赵忠良等^[75-76]建立了 自由滚转试验技术和大角度俯仰滚转单双自由度 强迫振动试验技术,陶洋等^[77] 对翼身组合体、三角 翼、窄条翼导弹、小展弦比飞翼等各类飞行器开展 了试验研究,获得了相关布局飞行器的滚转失稳特 性和气动力动态迟滞特性,动态试验风洞和结果如 图 7 所示,其中, *o*为滚转角, *a*为攻角。

基于现有研究基础,飞行器旋转翼折展过程的





风洞试验主要研究包括不同旋转翼偏角的气动力 风洞考核验证试验、典型状态的表面流谱显示试 验、纵横向动稳定参数试验、旋转翼动态折展过程 的自由滚转试验、旋转翼折展过程的动态测力试 验。试验研究有望对旋转翼折展过程实现计算验 证、气动力分析、典型状态的流动机理和稳定性动 力学机理验证。

5 结 论

针对飞行器旋转翼折展过程的复杂气动现象 及其诱导的动稳定性问题,国内外主要采用数值模 拟、建模仿真分析与风洞试验相结合的技术方法, 重点研究如何获取旋转翼在不同折展状态下,飞行 器的复杂绕流流动和气动特性,以及纵/横向耦合稳 定性特性;力求揭示复杂气动现象产生机理和纵/横 向耦合运动作用机理;以期厘清旋转翼折展过程 中,不同扰动状态下的运动稳定性动力学特性和机 理。可为更广范围的旋转翼飞行器创新设计提供 基础理论和技术支撑,研究的主要结论和发展方向 如下:

1)旋转翼折展过程中,翼面快速旋转和全机低频振动耦合运动过程复杂,针对数值模拟分离/转捩/激波求解精度差、运动耦合动网格插值误差大、非

定常计算效率低等问题;需要改进 RANS 方法的转 捩、分离和激波模拟能力,进而发展基于重叠网格 的 CFD/RBD 一体化紧耦合数值模拟方法,分析旋 转翼折展复杂动态气动/运动耦合机理。

2)针对旋转翼折展过程湍流、转捩、机动动作等因素引起的复杂飞行扰动,需要强化扰动建模背后的流动机理;针对旋转翼气动/运动耦合动稳定参数,需兼顾精度、效率和物理背景;发展基于极限学习机和深度学习等系统辨识类建模方法,将有利于高效、高精度地实现旋转翼扰动下的非线性气动力及其扰动参数建模。

3)旋转翼折展过程中,飞行器质量分布的持续 变化和绕流流动的剧烈变化,从而产生的机翼旋转 运动而带来的部件之间的气动干扰,以及旋转变形 过程中由于飞行器左右不对称而带来的纵/横向耦 合干扰,都会影响飞行器整体的运动稳定性;这些 问题的解决需要结合数值模拟和随机建模分析,深 入评估典型状态的运动稳定性及其扰动参数的气 动和动力学机理。

4)目前,常规的风洞试验、模型及其部件是相 对静止的,难以真实模拟旋转翼折展过程中机翼运 动而带来非定常复杂绕流流动,存在获取的气动特 性不真实的问题;在现有试验基础之上,可以在2.4 m 以上量级的风洞中发展六自由度的流动和气动特 性风洞试验方法,可以从典型流油试验、自由滚转 试验、折展过程动态测力试验角度实现流动特征和 动态稳定性机理验证。

参考文献(References)

- [1] JONES R T. Reduction of wave drag by antisymmetric arrangement of wings and bodies[J]. AIAA Journal, 1972, 10(2): 171-176.
- [2] 郝南松, 刘昀, 王进, 等. 斜置平板的低速风洞实验研究[J]. 空气动力学学报, 2013, 31(3): 344-349.
 HAO N S, LIU Y, WANG J, et al. Wind tunnel investigation of a

low speed oblique plate[J]. Acta Aerodynamica Sinica, 2013, 31(3): 344-349(in Chinese).

- [3] 阎超, 屈峰, 赵雅甜, 等. 航空航天 CFD 物理模型和计算方法的 述评与挑战[J]. 空气动力学学报, 2020, 38(5): 829-857.
 YAN C, QU F, ZHAO Y T, et al. Review of development and challenges for physical modeling and numerical scheme of CFD in aeronautics and astronautics[J]. Acta Aerodynamica Sinica, 2020, 38(5): 829-857(in Chinese).
- [4] GAN W B, ZHANG X C, MA T L, et al. Robust design and analysis of a conformal expansion nozzle with inverse-design idea[J]. Chinese Journal of Aeronautics, 2018, 31(1): 79-88.
- [5] 严恒元.飞行器气动特性分析与工程计算[M].西安:西北工业大学出版社,1990:92-113.

YAN H Y. Aerodynamic characteristics analysis and engineering calculation of aircraft[M]. Xi'an: Northwestern Polytechnical Uni-

versity Press, 1990: 92-113(in Chinese).

- [6] 刘溢狼,张伟伟,田八林,等. 一种超音速高超音速动导数的高效 计算方法[J]. 西北工业大学学报, 2013, 31(5): 824-828.
 LIU Y L, ZHANG W W, TIAN B L, et al. Effectively calculating supersonic and hypersonic dynamic derivatives[J]. Journal of Northwestern Polytechnical University, 2013, 31(5): 824-828(in Chinese).
- [7] STALNAKER J. Rapid computation of dynamic stability derivatives[C]//Proceedings of the 42nd AIAA Aerospace Sciences Meeting and Exhibit. Reston: AIAA, 2004: 210.
- [8] MOORE F G, MOORE L Y. 2009 version of the aeroprediction code: AP09[J]. Journal of Spacecraft and Rockets, 2008, 45(4): 677-690.
- [9] 刘绪,刘伟,柴振霞,等.飞行器动态稳定性参数计算方法研究进展[J]. 航空学报, 2016, 37(8): 2348-2369.
 LIU X, LIU W, CHAI Z X, et al. Research progress of numerical

method of dynamic stability derivatives of aircraft[J]. Acta Aeronautica et Astronautica Sinica, 2016, 37(8): 2348-2369(in Chinese).

- [10] 任玉新, 刘秋生, 沈孟育. 飞行器动态稳定性参数的数值计算方法[J]. 空气动力学学报, 1996, 14(2): 117-126.
 REN Y X, LIU Q S, SHEN M Y. Numerical calculation method of dynamic stability parameters of aircraft[J]. Acta Aerodynamica Sinica, 1996, 14(2): 117-126(in Chinese).
- [11] 刘伟. 细长机翼摇滚机理的非线性动力学分析及数值模拟方法 研究[D]. 长沙: 国防科学技术大学, 2004: 114-122.
 LIU W. Nonlinear dynamics analysis for mechanism of slender wing rock and study of numerical simulation method[D]. Changsha: National University of Defense Technology, 2004: 114-122(in Chinese).
- [12] 袁先旭,张涵信,谢昱飞. 基于 CFD 方法的俯仰静、动导数数值 计算[J]. 空气动力学学报, 2005, 23(4): 458-463.
 YUAN X X, ZHANG H X, XIE Y F. The pitching static/dynamic derivatives computation based on CFD methods[J]. Acta Aerodynamica Sinica, 2005, 23(4): 458-463(in Chinese).
- [13] 孙涛, 高正红, 黄江涛. 基于 CFD 的动导数计算与减缩频率影响 分析[J]. 飞行力学, 2011, 29(4): 15-18.
 SUN T, GAO Z H, HUANG J T. Identify of aircraft dynamic derivatives based on CFD technology and analysis of reduce frequency[J]. Flight Dynamics, 2011, 29(4): 15-18(in Chinese).
- [14] MCGOWAN G Z, KURZEN M J, NANCE R P, et al. High fidelity approaches for pitch damping prediction at high angles of attack[J]. Journal of Spacecraft and Rockets, 2014, 51(5): 1474-1484.
- [15] 范晶晶, 阎超, 李跃军. 飞行器大迎角下俯仰静、动导数的数值计算[J]. 航空学报, 2009, 30(10): 1846-1850.
 FAN J J, YAN C, LI Y J. Computation of vehicle pitching static and dynamic derivatives at high angles of attack[J]. Acta Aeronautica et Astronautica Sinica, 2009, 30(10): 1846-1850(in Chinese).
- [16] 郭东, 徐敏, 陈士橹. 基于网格速度法的非定常流场模拟和动导数计算[J]. 西北工业大学学报, 2012, 30(5): 784-788. GUO D, XU M, CHEN S L. An effective computation method based on field velocity approach for unsteady flow simulation and obtaining dynamic derivatives[J]. Journal of Northwestern Polytechnical University, 2012, 30(5): 784-788(in Chinese).
- [17] MURMAN S. A reduced-frequency approach for calculating dynamic derivatives[C]//Proceedings of the 43rd AIAA Aerospace Sciences Meeting and Exhibit. Reston: AIAA, 2005: 840.
- [18] 赵瑞. 飞行器动态气动特性与仿真技术[M]. 北京: 北京理工大学

出版社, 2019.

ZHAO R. Dynamic aerodynamic characteristics and numerical simulation method of flight vehicles[M]. Beijing: Beijing Institute of Technology Press, 2019(in Chinese).

- [19] ZHANG H X, ZHANG Z, YUAN X X, et al. Physical analysis and numerical simulation for the dynamic behaviour of vehicles in pitching oscillations or rocking motions[J]. Science in China Series E:Technological Sciences, 2007, 50(4): 385-401.
- [20] 李锋,杨云军,崔尔杰,等.飞行器自激振荡的流动物理与动力学 机制[J].空气动力学学报,2009,27(1):106-113.
 LI F, YANG Y J, CUI E J, et al. Flow physics and dynamics mechanism of self-oscillation of an flight vehicle[J]. Acta Aerodynamica Sinica, 2009, 27(1): 106-113(in Chinese).
- [21] 李跃军, 闫超. 非定常流动计算的混合时间推进方法研究[J]. 空 气动力学学报, 2007, 25(4): 483-487.
 LI Y J, YAN C. A hybrid time stepping scheme for calculating unsteady flows[J]. Acta Aerodynamica Sinica, 2007, 25(4): 483-487 (in Chinese).
- [22] 赵海洋. 返回舱动态稳定性物理机理分析及被动/主动控制方法 研究[D]. 长沙: 国防科学技术大学, 2007: 63-72.
 ZHAO H Y. Study on the mechanism and passive/active control methods of the dynamic stability of the reentry capsules[D]. Changsha: National University of Defense Technology, 2007: 63-72(in Chinese).
- [23] 杨小亮.飞行器多自由度耦合摇滚运动数值模拟研究[D].长沙: 国防科学技术大学, 2012: 42-46.
 YANG X L. Numerical investigation of aircraft rock in multiple de-

grees of freedom[D]. Changsha: National University of Defense Technology, 2012: 42-46(in Chinese).

- [24] 王晓冰. 高机动飞行器气动/运动耦合特性数值模拟研究[D]. 绵阳: 中国空气动力研究与发展中心, 2015: 24-29.
 WANG X B. Numerical investigation of aerodynamics and flight dynamics coupling for flight vehicle with high maneuverability[D].
 Mianyang: China Aerodynamics Research and Development Center, 2015: 24-29(in Chinese).
- [25] 马戎. 基于动态混合网格的气动/运动耦合一体化计算方法研究
 [D]. 绵阳: 中国空气动力研究与发展中心, 2015: 59-62.
 MA R. Numerical methods for aerodynamic/kinematic coupling problems based on dynamic hybrid grids[D]. Mianyang: China Aerodynamics Research and Development Center, 2015: 59-62(in Chinese).
- [26] 米百刚. 基于 CFD 的动导数计算及非线性气动力建模技术[D]. 西安: 西北工业大学, 2018: 7-10.

MI B G. Calculating dynamic derivative and modeling nonlinear unsteady aerodynamics based on CFD[D]. Xi'an: Northwestern Polytechnical University, 2018: 7-10(in Chinese).

- [27] TOBAK M. On the use of the indicial function concept in the analysis of unsteady motion of wings and wing-tail combinations: NACAR-1188[R]. Washington, D. C. : NACAR, 1954.
- [28] MURPHY P, KLEIN V, FRINK N. Unsteady aerodynamic modeling in roll for the NASA generic transport model[C]//Proceedings of the AIAA Atmospheric Flight Mechanics Conference. Reston: AI-AA, 2012: 4652.
- [29] GOMAN M G, KHRABROV A N, USOLTSEY S P. Unsteady aerodynamic model for large amplitude maneuvers and its parameter

identification[C]//Proceedings of the 11th IFAC Symposium on System Identification. Kitakyushu: IFAC, 1997: 399-404.

- [30] 汪清, 钱炜祺, 丁娣. 飞机大迎角非定常气动力建模研究进展[J].
 航空学报, 2016, 37(8): 2331-2347.
 WANG Q, QIAN W Q, DING D. A review of unsteady aerodynamic modeling of aircrafts at high angles of attack[J]. Acta Aeronautica et Astronautica Sinica, 2016, 37(8): 2331-2347(in Chinese).
- [31] BERAN P S, LUCIA D J, PETTIT C L. Reduced-order modelling of limit-cycle oscillation for aeroelastic systems[J]. Journal of Fluids and Structures, 2004, 19(5): 575-590.
- [32] DOWELL E H, THOMAS J P, HALL K C. Transonic limit cycle oscillation analysis using reduced order aerodynamic models[J]. Journal of Fluids and Structures, 2004, 19(1): 17-27.
- [33] ZHOU Q, CHEN G, DA RONCH A, et al. Reduced order unsteady aerodynamic model of a rigid aerofoil in gust encounters[J]. Aerospace Science and Technology, 2017, 63: 203-213.
- [34] LIU L, DOWELL E H, THOMAS J P. A high dimensional harmonic balance approach for an aeroelastic airfoil with cubic restoring forces[J]. Journal of Fluids and Structures, 2007, 23(3): 351-363.
- [35] THOMAS J P, DOWELL E H, HALL K C. Modeling viscous transonic limit cycle oscillation behavior using a harmonic balance approach[J]. Journal of Aircraft, 2004, 41(6): 1266-1274.
- [36] EKICI K, KIELB R E, HALL K C. The effect of aerodynamic asymmetries on turbomachinery flutter[J]. Journal of Fluids and Structures, 2013, 36: 1-17.
- [37] SILVA W. Identification of nonlinear aeroelastic systems based on the Volterra theory: Progress and opportunities[J]. Nonlinear Dynamics, 2005, 39(1): 25-62.
- [38] BALAJEWICZ M, NITZSCHE F, FESZTY D. Application of multi-input Volterra theory to nonlinear multi-degree-of-freedom aerodynamic systems[J]. AIAA Journal, 2010, 48(1): 56-62.
- [39] SILVA W. Recent enhancements to the development of CFD-based aeroelastic reduced-order models[C]//Proceedings of the 48th AIAA/ ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference. Reston: AIAA, 2007.
- [40] SILVA W A, BARTELS R E. Development of reduced-order models for aeroelastic analysis and flutter prediction using the CFL3D v6.0 code[J]. Journal of Fluids and Structures, 2004, 19(6): 729-745.
- [41] COWAN T J, ARENA A S JR, GUPTA K K. Accelerating computational fluid dynamics based aeroelastic predictions using system identification[J]. Journal of Aircraft, 2001, 38(1): 81-87.
- [42] ZHANG W W, WANG B B, YE Z Y, et al. Efficient method for limit cycle flutter analysis by nonlinear aerodynamic reduced-order models[J]. AIAA Journal, 2012, 50(5): 1019-1028.
- [43] MARQUES F D, ANDERSON J. Identification and prediction of unsteady transonic aerodynamic loads by multi-layer functionals[J]. Journal of Fluids and Structures, 2001, 15(1): 83-106.
- [44] HUANG G B, SARATCHANDRAN P, SUNDARARAJAN N. A generalized growing and pruning RBF (GGAP-RBF) neural network for function approximation[J]. IEEE Transactions on Neural Networks, 2005, 16(1): 57-67.
- [45] MACKMAN T J, ALLEN C B, GHOREYSHI M, et al. Comparison of adaptive sampling methods for generation of surrogate aerodynamic models[J]. AIAA Journal, 2013, 51(4): 797-808.

- [46] LIU H J, HU H Y, ZHAO Y H, et al. Efficient reduced-order modeling of unsteady aerodynamics robust to flight parameter variations[J]. Journal of Fluids and Structures, 2014, 49: 728-741.
- [47] LI W C, JIN D P, ZHAO Y H. Efficient nonlinear reduced-order modeling for synthetic-jet-based control at high angle of attack[J]. Aerospace Science and Technology, 2017, 62: 98-107.
- [48] CHEN G, SUN J, MAO W T, et al. Limit cycle oscillation control for transonic aeroelastic systems based on support vector machine reduced order model[J]. Transactions of the Japan Society for Aeronautical and Space Sciences, 2013, 56(1): 8-14.
- [49] GU J, ZOU Q Y, DENG C H, et al. A novel robust online extreme learning machine for the non-gaussian noise[J]. Chinese Journal of Electronics, 2023, 32(1): 130-139.
- [50] 姜泽翔,杨立本,王栋,等.倾转双旋翼无人机风场扰动下的建模 和控制器设计[J].飞行力学,2021,39(5):38-43.
 JIANG Z X, YANG L B, WANG D, et al. Modeling and controller design of tilting twin-rotor UAV under wind field disturbance[J]. Flight Dynamics, 2021, 39(5):38-43(in Chinese).
- [51] 高振兴,顾宏斌.复杂大气扰动下大型飞机飞行动力学建模研究
 [J].系统仿真学报,2009,21(17):5556-5561.
 GAO Z X, GU H B. Research on modeling of flight dynamics for large aircraft in complex atmospheric disturbance[J]. Journal of System Simulation, 2009, 21(17): 5556-5561(in Chinese).
- [52] 陆昌根, 赵玲慧, 沈露予. 局部扰动对平板边界层流动稳定性影响的研究[J]. 空气动力学学报, 2012, 30(1): 63-67. LU C G, ZHAO L H, SHEN L Y. Numerical study on the effect of local disturbance on hydrodynamic stability of plate boundary layer[J]. Acta Aerodynamica Sinica, 2012, 30(1): 63-67(in Chinese).
- [53] 何磊,张显才,钱炜祺,等.基于长短时记忆神经网络的非定常气动力建模方法[J].飞行力学,2021,39(5):8-12.
 HE L, ZHANG X C, QIAN W Q, et al. Unsteady aerodynamics modeling method based on long short-term memory neural network[J]. Flight Dynamics, 2021, 39(5): 8-12(in Chinese).
- [54] 武频, 孙俊五, 封卫兵. 基于自编码器和 LSTM 的模型降阶方法
 [J]. 空气动力学学报, 2021, 39(1): 73-81.
 WU P, SUN J W, FENG W B. Reduced order model based on autoencoder and long short-term memory network[J]. Acta Aerodynamica Sinica, 2021, 39(1): 73-81(in Chinese).
- [55] 吕永玺. 先进战斗机大迎角建模和控制关键技术研究[D]. 西安: 西北工业大学, 2018. LYUYX. Research on the key technologies of modeling and control of the advanced fighter at high angle of attack[D]. Xi'an: North-
- western Polytechnical University, 2018(in Chinese). [56] 乐挺, 王立新, 艾俊强. 变体飞机设计的主要关键技术[J]. 飞行力 学, 2009, 27(5): 6-10.

YUE T, WANG L X, AI J Q. Key technologies in morphing aircraft design[J]. Flight Dynamics, 2009, 27(5): 6-10(in Chinese).

- [57] YUE T, WANG L X, AI J Q. Longitudinal linear parameter varying modeling and simulation of morphing aircraft[J]. Journal of Aircraft, 2013, 50(6): 1673-1681.
- [58] YUE T, WANG L X, AI J Q. Gain self-scheduled H_x control for morphing aircraft in the wing transition process based on an LPV model[J]. Chinese Journal of Aeronautics, 2013, 26(4): 909-917.
- [59] 乐挺, 王立新, 艾俊强. Z 型翼变体飞机的纵向多体动力学特性 [J]. 航空学报, 2010, 31(4): 679-686.

YUE T, WANG L X, AI J Q. Longitudinal multi-body dynamic characteristics of Z-wing morphing aircraft[J]. Acta Aeronautica et Astronautica Sinica, 2010, 31(4): 679-686(in Chinese).

- [60] HENRY J, PINES D. A mathematical model for roll dynamics by use of a morphing-span wing[C]//Proceedings of the 48th AIAA/ ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference. Reston: AIAA, 2007.
- [61] 殷明, 陆宇平, 何真. 变体飞行器 LPV 建模与鲁棒增益调度控制
 [J]. 南京航空航天大学学报, 2013, 45(2): 202-208.
 YIN M, LU Y P, HE Z. LPV modeling and robust gain scheduling control of morphing aircraft[J]. Journal of Nanjing University of Aeronautics & Astronautics, 2013, 45(2): 202-208(in Chinese).
- [62] HENRY J J. Roll control for UAVs by use of a variable-span morphing wing[D]. Annapolis: University of Maryland, 2005: 54-63.
- [63] 金鼎,张炜,艾俊强.折叠机翼变体飞机纵向操纵性与稳定性研究[J].飞行力学,2011,29(1):5-8.
 JIN D, ZHANG W, AI J Q. Study on longitudinal maneuverability and stability of folding wing morphing aircraft[J]. Flight Dynamics, 2011, 29(1): 5-8(in Chinese).
- [64] BEAVERSTOCK C S, FINCHAM J, FRISWELL M I, et al. Effect of symmetric & asymmetric span morphing on flight dynamics[C]//Proceedings of the AIAA Atmospheric Flight Mechanics Conference. Reston: AIAA, 2014.
- [65] SEIGLEI T M. Dynamics and control of morphing aircraft[D]. Blacksburg: Virginia Polytechnic Institute and State University, 2005: 73-105.
- [66] SEIGLER T M, NEAL D A. Analysis of transition stability for morphing aircraft[J]. Journal of Guidance, Control, and Dynamics, 2009, 32(6): 1947-1954.
- [67] SHI R Q, WAN W Y. Analysis of flight dynamics for large-scale morphing aircraft[J]. Aircraft Engineering and Aerospace Technology, 2015, 87(1): 38-44.
- [68] 张杰, 吴森堂. 一种变体飞行器的动力学建模与动态特性分析[J]. 北京航空航天大学学报, 2015, 41(1): 58-64. ZHANG J, WU S T. Dynamic modeling for a morphing aircraft and dynamic characteristics analysis[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(1): 58-64(in Chinese).
- [69] 李文成. 变体飞行器动力学建模与稳定性分析及控制[D]. 南京: 南京航空航天大学, 2017.

LI W C. Dynamic modeling, stability analysis and control of variant aircraft[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2017(in Chinese).

[70] 杨贯通, 唐胜景, 赵林东, 等. 变后掠变展长飞行器动力学建模与

动态响应分析[J]. 兵工学报, 2014, 35(1): 102-107. YANG G T, TANG S J, ZHAO L D, et al. Dynamic modeling and response of a morphing UAV with variable sweep and variable

span[J]. Acta Armamentarii, 2014, 35(1): 102-107(in Chinese).

- [71] 贺中, 吴军强, 蒋卫民, 等. 细长体大迎角非对称流动的高速 PIV 风洞试验研究[J]. 空气动力学学报, 2014, 32(3): 295-299.
 HE Z, WU J Q, JIANG W M, et al. Study on asymmetric flow over slender body at high angles of attack via particle image velocimetry test in high speed wind tunnel[J]. Acta Aerodynamica Sinica, 2014, 32(3): 295-299(in Chinese).
- [72] ARAUJO-ESTRADA S A, LOWENBERG M H, NEILD S, et al. Evaluation of aircraft model upset behaviour using wind tunnel manoeuvre rig[C]//Proceedings of the AIAA Atmospheric Flight Mechanics Conference. Reston: AIAA, 2015.
- [73] 史志伟, 耿玺, 程克明, 等. 一种适用于复杂风洞动态试验的新型 机构[C]//中国力学大会-2015, 北京: 中国力学学会, 2015: 244.
 SHI Z W, GENG X, CHENG K M, et al. A new mechanism for dynamic test in complex wind tunnel[C]//Proceedings of the Chinese Congress of Theoretical and Applied Mechanics 2015. Beijing: The Chinese Society of Theoretical and Applied Mechanics, 2015: 244(in Chinese).
- [74] 邓学蓥,石伟,王延奎,等.两类非对称涡流动所诱导的摇滚运动
 [J]. 气体物理, 2016, 1(1): 13-24.
 DENG X Y, SHI W, WANG Y K, et al. Wing rock motions induced by two kinds of asymmetric vortices flows[J]. Physics of Gases, 2016, 1(1): 13-24(in Chinese).
- [75] 赵忠良,杨海泳,马上,等. 某典型飞行器模型俯仰/滚转两自由度 耦合动态气动特性[J]. 航空学报, 2018, 39(12): 107-116.
 ZHAO Z L, YANG H Y, MA S, et al. Unsteady aerodynamic characteristics of two-degree-of-freedom pitch/roll coupled motion for a typical vehicle model[J]. Acta Aeronautica et Astronautica Sinica, 2018, 39(12): 107-116(in Chinese).
- [76] 陶洋, 赵忠良, 李浩, 等. 80°/65°双三角翼滚转稳定特性预测研究
 [J]. 实验流体力学, 2013, 27(6): 43-46.
 TAO Y, ZHAO Z L, LI H, et al. Study on prediction of rolling stability characteristics of 80/65 double delta wings[J]. Journal of Experiments in Fluid Mechanics, 2013, 27(6): 43-46(in Chinese).
- [77] 陶洋, 赵忠良, 王红彪, 等. 前体涡诱导机翼摇滚扰动控制高速风 洞试验研究[J]. 实验流体力学, 2014, 28(1): 21-25. TAO Y, ZHAO Z L, WANG H B, et al. Flow control investigation on wing rock induced by forebody vortex at high speed wind tunnel[J]. Journal of Experiments in Fluid Mechanics, 2014, 28(1): 21-25(in Chinese).

Research progress on dynamic stability of rotating variant wing opening and closing process for aircraft

GAN Wenbiao^{1,*}, ZUO Zhenjie², XIANG Jinwu², ZHAO Zhongliang³, CAI Jun⁴, MA Shang³

(1. Institute of Unmanned System, Beihang University, Beijing 100191, China;

2. School of Aeronautic Science and Engineering, Beihang University, Beijing 100191, China;

3. High Speed Aerodynamics Institute, China Aerodynamics Research and Development Center, Mianyang 621000, China;

4. Hunan Vanguard Group Co., Ltd., Changsha 410100, China)

Abstract: One type of wings that may spin along a fixed axis is the rotating variant wing. It is frequently utilized in the domains of aviation bombs, cruise missiles, new concept unmanned aerial vehicles, and other aircraft. Its dynamics stability mechanism of opening and closing process is a key fundamental question for the design of rotating variant wing aircraft. This paper summarizes the research progress on dynamics stability of aircraft rotating variant wing opening and closing process in recent years. Firstly, the development history of the rotating variant wing and the key issues of dynamic stability in the opening and closing process are introduced. Secondly, the numerical simulation of dynamic stability in the opening and closing process is expounded from three levels: unsteady aerodynamic numerical simulation, dynamic characteristic simulation, and CFD/RBD integrated coupling simulation progress. The introduction covers the present state of dynamic stability analysis and nonlinear dynamic modeling under the rotating variant wing disturbance. After that, the analysis is done on the wind tunnel test results confirming the rotating variant wing stability. Ultimately, the conclusion is made that scientific problems faced by the research on the dynamic stability of the rotating variant wing during opening and closing are presented, and the feasible research directions are proposed.

Keywords: unsteady flow; rotating variant wing; coupling of aerodynamic and kinematic; dynamic mechanism; dynamic analysis

Received: 2022-06-11; Accepted: 2022-08-19; Published Online: 2022-09-22 19:50 URL: link.cnki.net/urlid/11.2625.V.20220921.1157.001

Foundation items: National Natural Science Foundation of China (U2141249,11902018,U2141252); Aeronautical Science Foundation of China (2019ZA051001)

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn DOI: 10.13700/j.bh.1001-5965.2022.0425

机场航班时刻资源管理研究进展

王艳军*,水笑雨,王梦尹

(南京航空航天大学 民航学院,南京 211106)

摘 要:繁忙机场都会面临严重的交通拥堵和航班延误问题。解决该问题的一个有效手段 是对机场航班时刻资源进行优化管理。全面介绍了航班时刻资源管理相关的研究进展,系统回顾了 机场公布容量设置、机场航班时刻配置的主要方法和手段。从公布容量确定、单机场、机场网络层 面和机场群航班时刻配置及航班时刻配置技术复杂性等方面,得出机场航班时刻资源优化配置的关 键技术问题。展望了机场航班时刻资源管理的主要方向,并对下一步研究提出建议。

关键 词:机场容量;时刻分配;容量需求管理;战略流量管理;资源优化

中图分类号: V355

文献标志码: A 文章编号: 1001-5965(2024)04-1065-12

大多数繁忙机场都面临着严重的交通拥堵和 航班延误问题,航班时刻需求远远大于机场所能供 给的数量,因此,需要进行容量和需求管理。航班 时刻是"航空器在指定日期和时间,为抵离某个机 场而使用相关基础设施与服务的权利",是航空运 输系统中最宝贵的资源[1]。航空公司和其他航空器 运营人(以下统称"航空承运人")基于所分配的航 班时刻制定航班计划,为客户提供航空运输服务。 虽然目前仍没有一个公认的方法评估时刻的价值^[2], 但从中国最近试行的时刻拍卖及伦敦希思罗机场 航班时刻在航空公司之间的转卖,即可窥见航班时 刻价值之巨大,繁忙机场的一个航班时刻价值上千 万美元。由于物理设施资源和运行程序等因素的 限制,机场航班时刻资源有限。航班时刻资源优化 配置能够快速地提高容量利用率,尽可能满足航空 公司航班时刻需求,具有重要的理论意义和应用价 值,受到越来越多的关注^[3-4]。本文对机场航班时刻 资源管理的相关研究工作进行了综述,梳理了航班 时刻资源管理研究的最新进展和面临的挑战,为未 来的航班时刻资源管理研究和应用提供参考。

机场航班时刻资源管理基本概念及 方式

1.1 航班时刻协调机场

国际航空运输协会(International Air Transportation Association, IATA)根据机场的供需关系将机场分为3个级别:航班时刻协调机场、航班时刻 辅协调机场和非航班时刻协调机场^[5]。其中,航班 时刻协调机场是指满足以下4个条件的机场:①在 相应的时段内,对机场物理设施使用的需求远远超 过了该机场的容量;②在短期内不可能通过对机场 的物理设施进行改扩建的方式来满足需求;③机场 无法通过自主的航班时刻调整来解决供需不匹配 的问题;④需要一个航班时刻配置的过程,在航班 时刻配置期间必须通过一位时刻协调员来为所有 航空承运人分配进场和离场时刻。

由于航空运输需求具有较强的季节性特征,有 些机场在夏秋(冬春)航季为航班时刻协调机场,在 冬春(夏秋)航季为非航班时刻协调机场。根据 2022年1月5日数据,2022夏秋(冬春)航季世界上 共有198(170)个航班时刻协调机场和153(153)个

收稿日期: 2022-05-28; 录用日期: 2022-07-06; 网络出版时间: 2022-07-19 14:46 网络出版地址: link.cnki.net/urlid/11.2625.V.20220718.1808.001

基金项目: 国家自然科学基金 (U2033203)

*通信作者. E-mail: ywang@nuaa.edu.cn

引用格式: 王艳军, 水笑雨, 王梦尹. 机场航班时刻资源管理研究进展 [J]. 北京航空航天大学学报, 2024, 50 (4): 1065-1076. WANG Y J, SHUI X Y, WANG M Y. Research progress on airport slot allocation [J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50 (4): 1065-1076 (in Chinese).

2024 年

航班时刻辅协调机场。中国共有22个航班时刻协 调机场,是世界上航班时刻协调机场最多的国家。

虽然航班时刻协调机场数量不足全球机场总量的 6%,但是这些机场为世界上近 50% 的旅客提供运输服务。因此,合理配置航班时刻具有重要的现实意义和应用价值。

1.2 航班时刻配置办法和程序

1.2.1 航班时刻管理准则

航班时刻配置的方法可分为行政分配和市场 化配置。由于对航班时刻的价值和归属问题存在 较大的争议,大多数国家都采用行政化分配的方 法,即由管理者或第3方根据一定的规则与航空公 司协调,对机场航班时刻进行配置。IATA 在定期 更新的《Worldwide Slot Guidelines》(WSG)中提出一 系列的规则为协调机场和辅协调机场解决时刻分 配问题⁶⁶。2020年6月, IATA、国际机场理事会 (Airports Council International, ACI)和世界机场协 调人小组 (Worldwide Airport Coordinators Group, WWACG) 共同发表了第1版《Worldwide Airport Slot Guidelines》(WASG)^[5]。各个国家和地区依据 WSG/WASG 规则,提出适用于本国或本地区的航 班时刻配置方法和程序。如2018年4月中国开始 实施的《民航航班时刻管理办法》用以指导管理机 场的时刻配置^[1]: 欧盟国家目前依然采用的方法主 要是基于1993年发布的航班时刻管理办法(即 Regulation 95/93)^[2]。这些办法实施的目的是保证 航空承运人公平参与航空市场竞争,提高航班时刻 资源高效配置和有效使用等。

WSG 早在上世纪就已被提出,虽然之后经过 不断的修订,但仍然存在一些问题。这些规章和办 法仅为航班时刻配置和管理提供了指导性意见,在 航班时刻分配过程中依然存在一些技术问题和机 制问题亟待解决。2020年,美国麻省理工学院的 Odoni对 Regulation 95/93 航班时刻规章进行了深刻 的评述,指出欧盟在航班时刻配置过程中存在的关 键问题,并提出相关的解决建议^[7]。本文重点对航 班时刻行政分配中面临的难点重点问题进行分析, 为航班时刻管理者和研究者提供参考。

1.2.2 航班时刻配置关键概念

1) 航班系列。航班时刻配置中一个关键的概

念,指"航班配置特定运营日(周一至周日的某一 日)的航班时刻,应当占用特定运营日相同的时间; 若无可能,尽量配置在大致相同的时间"^[1]。WASG 要求航班时刻系列最小长度为5,即至少连续5周 在同一时间(或近似时间)内的时刻。如表1所示, 3个航空公司申请某一机场(机场1)的起飞时刻 (起飞时刻1)。其中,航班AL123申请下一航季从 2020年12月8日至2021年2月6日期间,于每周 二、周四和周六早上08:00起飞;航班BL223申请 下一个航季中每天一班的时刻;而航班CL001申请 下一航季从2020年10月26日至2020年11月 29日,每周一、周三、周五、周日执行的从机场1起 飞至机场4的航班时刻。由此可见,航空公司申请 时刻系列的多样性和时刻管理规则的复杂性将为 优化航班时刻配置大大增加难度。

2)历史优先权。历史优先权是指航空承运人 在上一个同航季某一时刻系列的实际执行率达到 了 80%及以上(中国还要求执行的航段要覆盖超 过 2/3 航季),航空公司在本航季将有优先权继续使 用该时刻系列,即所谓的祖父权利。该航班时刻系 列被称为"历史时刻"。

执行率为 80% 的规定通常被称为"不使用就 放弃"规则。因此,一些航空公司为了能够在下一 个同航季继续使用该航班时刻,经常会运营一些载 客率非常低的航班。

3)时刻申请的优先级。航空承运人所提出的时刻申请可分为4类:历史时刻、历史时刻调整、新进航空承运人申请及其他时刻申请。最后一版的WSG(第10版)⁶⁰和中国的时刻管理办法^[1]中,航班时刻配置的优先级从高到低的顺序为历史时刻、历史时刻调整申请、新进航空承运人申请和其他时刻申请。

根据最新的 WASG^[5] 规定, 历史时刻和不影响 时刻协调参数的历史时刻调整(如仅改变航班号) 具有第1优先权, 即所谓的祖父权利。这些航班时 刻统称为"未改变的历史时刻"。在分配完未改变 的历史时刻之后, 剩余的航班时刻(包含新增的时 刻)都被放入时刻池中进行统一配置。

新进航空承运人申请、非新进航空承运人申请 和影响到时刻协调参数的历史时刻调整享有同等

表1 航空公司申请机场航班时刻系列

 Table 1
 An example of series of slot request

航班号	机型	运营日	起飞机场1	起飞时刻1	落地时刻1	目的地机场	开始日期	结束日期
AL123	320	周二、周四、周六	机场1	08:00	10:50	机场2	2020-12-01	2021-02-06
BL223	320	周一至周日	机场1	07:10	09:45	机场3	2020-10-25	2021-03-27
CL001	320	周一、周三、周五、周日	机场1	14:00	19:30	机场4	2020-10-26	2020-11-29

的优先权。其中,新进航空承运人是指在某一机场 特定运营日内持有的航班时刻数量少于规定的航 班时刻数量(如7个航班时刻)的航空公司。只有 航空公司才具有新进航空承运人资格。按照规定, 时刻池中的50%时刻应分配给新进航空承运人。

4)航班时刻协调参数。航班时刻协调参数或 机场公布容量给出了机场单位时间(通常为1h或 15 min)能够分配给航空承运人的时刻数量。一些 机场或将其进一步细分为机场的进场容量、离场容 量和机场总容量,或可分为国际航站楼容量、申根 地区航站楼容量等。

航班时刻协调参数实际上设置了机场可分配 时刻总数的上限。

5) 航班时刻配置流程。航班时刻配置的基本 流程是,时刻协调员或相关机构确定机场的航班时 刻协调参数,航空承运人提出时刻申请,然后由时 刻协调员根据航班时刻协调参数和时刻申请,按照 一定的规则为航空承运人配置航班时刻,如图1所 示。优化配置的航班时刻应尽量满足航空承运人 的需求,不超出航班时刻协调参数的限制。





Fig. 1 Basic process of slot allocation

无论是 WASG^[5] 还是中国民用航空局实施的 时刻管理办法^[1], 航班时刻资源管理涉及到的 2 个 关键环节分别是航班时刻协调参数的确定和航班 时刻初始分配。航班时刻初始分配根据时刻配置 的相关规则将时刻分配给航空承运人。以下将对 机场公布容量确定方法和时刻配置优化模型与算 法进行详细综述。

2 机场公布容量及其确定方法

航班时刻配置的首要基础是确定机场的公布 容量(或称为时刻容量、协调参数等)。公布容量构 成了现有航班时刻分配方法中的基本概念,表示基 于最大吞吐量的容量管理度量。其规定了每单位 时间在航班时刻协调机场可用于分配的航班时刻 总数,并且通常使用每个协调时间间隔内的计划起 降架次表示。为了更加精细化的管理,公布容量还 可根据不同时间间隔内可用的离场架次、进场架次 和/或总起降架次表示。确定公布容量至关重要,因 为公布容量决定了航班时刻配置过程中的供给侧, 即机场一共有多少航班时刻可以供航空公司使用。

2.1 典型机场容量的定义及评估方法

机场容量的一个通用定义是在给定运行环境 和管制规则等条件下,单位时间内(通常是1h或 15 min)机场能提供服务的飞行架次。合理确定机 场公布容量的难点之一在于准确评估机场的运行 容量。公布容量取决于机场的运行容量。理论上, 每个航班时刻协调机场的时刻协调员,在给定的假 设和属性下,考虑各种容量决定因素进行综合评估 研究才能确定公布容量^[8-11]。机场容量可进一步划 分为陆侧容量和空侧容量^[12]。陆侧容量考虑机场 航站楼内基础设施布局、值机柜台分布等,是机场 能够提供服务的最大旅客数量。空侧容量考虑机 场空侧结构布局,包括跑道系统、滑行道系统和停机位/坪系统,是机场在单位时间内能够为航空器提供服务的架次^[13]。通常,机场的空侧容量是机场容

量的瓶颈。根据容量的应用场景和主要考虑因素, 机场容量具有不同的定义和评估方式。表2为几 种典型的机场容量定义及评估方法。

Table 2	Widely used	l airport capacity	and its	s measuring	methods
---------	-------------	--------------------	---------	-------------	---------

容量类型	定义	容量的 时间尺度	是否考虑 航班延误	主要评估方法	应用场景
理论容量	在给定机场构型、运行规则和机队组成情况 下,机场在单位时间内能够提供的服务架次	a, h	否	类似历史数据/电子表 格/排队论模型/仿真	战略规划
最大/饱和容量	在持续交通服务请求下,每单位时间内能够实 现的预期飞行架次	h, 15min	是	排队论模型/仿真	战术运行及运行后分析
实际容量	考虑随机达到航班的排队特征,机场实际运行 的容量,通常比最大容量低10%~20%	h, 15 min	是	仿真	预战术/战术
持续容量	在连续几个小时内可以实现的小时容量。最大 容量的运行通常不能超过1~2 h	h	否	仿真	预战术/战术
结构容量	在宏观阶段用于确定机场容量的基线,在长期 预测中故意不考虑一些不可靠或可变因素。通 常在2~5 a前评估结构容量	a, 季节	否	历史数据/查询表/电子 表格/仿真/容量包络线	早期的空中交通流量容 量管理
计划容量	原则上在IATA航季开始之前,或至少在实际运行前18个月计算的机场容量	d, h, 15 min	否	查表/电子表格/仿真/容 量包络线	用于网络战略阶段和初 始时刻分配
运行容量	在运行前一天评估的机场容量,目的是结合最 新的运行信息更新机场容量	h, 15 min	否	仿真	网络预战术和战术阶段
公布容量	时刻协调机场在单位时间(通常为1h或15min) 能够分配给航空承运人的时刻数量	h, 15 min	否	结合运行容量之后综合 分析	时刻配置

只有当交通需求大于容量资源供给,容量才会成为瓶颈,出现航班延误。部分学者或机构在定义机场容量时,认为需要考虑航班延误,基于航班延误设定阈值来确定机场容量^[12-14]。例如,美国关于机场容量的定义(除去持续容量)和国际机场协会关于容量的定义都是基于给定的延误水平确定容量^[14]。然而,需要指出的是,延误阈值的确定对于各个机场并不相同。例如,旅客对大型枢纽机场的延误容忍度可能会高,但是对小机场的航班延误容忍度则较低。同时,由于空中等待成本要高于地面等待成本,进场延误阈值要低于离场延误阈值。图2为几种典型容量之间及与航班延误之间的关系。

欧洲航空安全组织并不建议在容量分析时引 入延误。因为机场容量仅与需求有关,与具体的航 班时刻表无关。在容量评估时引入延误,默认了基





于机场容量的航班时刻表预测。但是,航班时刻表 和机场容量应该是相互独立的^[15]。

根据机场容量使用的场景或时期,可以采用不同的容量评估方法。常用方法可分为4类:数学建模、历史运行数据挖掘、计算机仿真和基于管制员工作负荷^[16-17]。建立数学模型能够快速估算出机场的空侧容量,但缺点是计算结果的精度不高^[9];计算机仿真方法考虑机场物理结构、运行规则和航空器性能等因素,建立高度逼真的仿真模型来模拟机场运行,根据仿真后的航班确定机场容量。常见的机场仿真软件包括AirTop^[18]、TAAM^[19]和SIMMOD^[20]等。计算机仿真可以高度逼真的模拟机场的运行,但缺点是耗时、软件价格高;仿真结果依赖于模型的输入。当运行条件或计划变化时,需重新建立仿真模型进行计算。

需要注意的是,机场容量并不是一个简单的数 值,而是关于容量率的范围。机场的容量包络线是 以进场率和离场率为横纵坐标的一个近似的凸多 边形,其定义了在不同运行条件下可能实现的最大 容量,如图3所示。Gilbo最早通过定义帕累托边 界来确定机场的最大容量^[21]。机场的实际运行容 量受到众多因素影响,包括天气、跑道运行方式、 机型比例等^[22]。世界上主要国家或地区均发布了 相关报告或指导手册,规定了机场容量评估及管理 遵循的原则、程序和所用评估技术等,为机场容量 评估提供了初步的管理方法和流程^[15,23]。

2.2 公布容量的确定

"公布容量"并非一个快速可测的"数量",而





匹配)^[28]。当前确定航班时刻协调机场公布容量的 方法受到了很多批评,存在的主要问题如下:①航 班时刻管理机构有充足的空间进行说明和调整; ②根据仪表飞行规则容量的百分比设定极低的容 量水平;③确定公布容量的过程是经验性的或临时 的;④很少与利益相关者沟通^[9]。

公布容量的确定所存在的问题主要集中在 2个方面:①缺乏协同一致的解读、确定和管理公 布容量的方法;②公布容量设定的容量水平过高或 过低。公布容量还需考虑容量利用率和服务质量 (航班延误或准点率)之间的均衡^[29-30]。如果公布容 量设置过高,而实际运行容量也较高,那么容量利 用率也会很高;但当实际运行容量低于公布容量 时,将会造成航班延误。如果公布容量设置较低, 虽然保证了较低的航班延误和较高的准点率,但会 造成容量资源浪费。同时,也会导致过多的拒绝或 更改航空公司申请的航班时刻,间接地干预了航空 公司航班时刻设置的偏好。Barnhart等指出未来研 究重点之一,是使用分析建模来正确地确定和分配 机场容量,同时研究过度分配航班时刻和过低分配 时刻的权衡和影响^[30]。

2.3 机场公布容量确定所面临的挑战

基于使用先进方法和工具设置公布容量所面 临的挑战总结如下。

 1)公布容量的设置必须要考虑所观测到的机 场运行的所有情景(例如天气良好/恶劣等)。仅仅 关注恶劣天气的保守公布容量将会造成不必要的 过低的航班时刻表;仅关注天气良好下的公布容量 2)公布容量的设置应该尽可能的按照高精度 来设置。例如,不同的跑道、航站楼、机坪设置单 独的公布容量,从而使得时刻需求和机场运行能力 得到更好的匹配。

3)当前主要的公布容量设置方法是为机场一 天的高峰时段设置一个"平滑"的公布容量。例 如,从早上06:00到晚上10:00机场的公布容量为 80架次/h。然而,设置公布容量时应该考虑小时的 不同。按小时设置公布容量的做法正在被越来越 多的欧洲机场所采用(如阿姆斯特丹机场、伦敦希 思罗机场和巴黎戴高乐机场)。

此外,应适当设置容量低谷以在实际运行过程 中吸收航班延误。例如,在爱尔兰都柏林机场和日 本东京成田机场,在公布时刻协调参数时会在 06:00或14:00设置一个容量低谷,以吸收航班累积 延误。

4)设置公布容量时还应考虑需求的时段模式, 同时考虑交通需求和容量。航空公司和机场可能 会倾向于一定程度的高延误来换取更多的时刻。 公布容量和延误水平之间的均衡应该深入的探讨。

因此,公布容量设置为一个合理的数值,既能 实现既定的服务水平,同时也考虑了机场利用率和 延误之间的非线性关系。如何合理的设置容量曲 线是一个值得深入研究的问题。

3 基于行政手段的时刻配置

在机场公布容量确定之后,航空公司将会提交 时刻申请。航班时刻管理部门将会根据相关的规 则进行时刻配置。Zografos 等在 2017 年和 2018 年 对航班时刻分配问题相关研究进行了综述[31-32],指 出机场航班时刻分配是一个极具挑战性的研究方 向。航班时刻分配和最佳公布容量水平的设定是 密切相关的,都是优化机场稀缺资源分配和使用的 核心。航班时刻分配的主要难点之一在于如何考 虑多重因素的限制,为整个航季合理有效地分配航 班时刻系列。时刻配置的主要方法可分为基于市 场手段的航班时刻分配和基于行政手段的时刻分 配。基于市场驱动的时刻配置主要从时刻拍 卖[33-34]和交易[35-42]等方面在理论上探讨解决机场 拥堵的方法。一些学者认为基于拥堵的收费策略 是对所有机场都适应的时刻分配策略^[42]。在航班 时刻配置过程中引入市场机制,在一定程度上可以 促进时刻的高效利用,但时刻权属、时刻定价等关 键问题尚未解决。由于目前尚未有任何国家采用 基于市场的航班时刻配置方法,因此,本文仅针对

基于行政手段的航班时刻配置研究进行综述。表 3 为时刻配置模型和算法研究的相关文献。

3.1 问题描述

机场航班时刻分配问题旨在将航空公司在机 场所申请的运行(着陆或起飞)分配到给定分配时 段内的协调时间间隔,其主要目标是最小化基于延 误的成本。延误成本可以表示为运行延误或调配 延误。运行延误通常以航班预计进场/离场延误和/ 或乘客总延误表示。"调配延误"是基于距离的度 量,用于表示所请求时刻与所分配时刻之间的差异 (即"被移动的步数")。除了考虑分配效率,时刻 配置的目标还可以为公平与公正、获取与竞争及环 境目标等。时刻分配问题的主要约束包括机场公 布容量、现有时刻分配框架下的优先级、航空公司 和机场的运营要求(例如航班周转时间、航班连 接)、流量管理约束(如空域扇区的容量)及时刻分

表 3 机场航班时刻配置相关研究文献

Table 3	Research literature on airport slot allocation
---------	--

文献	优化目标	约束条件	时刻配置尺度	求解算法/规划模型
文献[3]	①长系列偏移的最小化;②全季节系列偏移的最小化;③最大限度地减少长途航班的偏移;④最小化国际航班偏移;⑤最小化本地 飞机时隙请求偏移;⑥最大偏移最小;⑦偏移时隙数量最小座椅偏移最小;⑧最小化协调机场航班偏移	①时刻唯一性约束; ②机场容量约束; ③航 班周转时间约束	单机场,整个航季	启发式算法、基于 ϵ 约束的方法、多准 则决策分析、 TOPSIS法、PSAM
文献[4]	①总偏移最小化;②最大偏移最小化;③绝对公平的最大偏离(MDA)最小化	①时刻唯一性约束;②机场容量约束;③航 班周转时间约束	单机场,整个航季	TOSAM 启发式算法
文献[28]	总偏移最小	①以时段和天为单位的机场容量约束;②航 班周转时间约束	单机场,整个航季	线性松弛算法
文献[31]	运行效率、时刻分配效率、时刻资源利用、 环境影响、公平性	①机场公布容量约束;②现有航班时刻分配 框架下的优先级(如,历史航班时刻持有 量);③航空公司和机场的运行要求(例如, 周转时间、地面等待、航班连接性); ④ATFM空中交通流量管理约束(如,空域 扇区的容量);⑤其他用于时刻请求调度的 标准。	单机场,整个航季 机场网络,多天/整个 航季	公布容量建模与战略 航班时刻调度模型
文献[32]	①总偏移最小化;②最大航班偏移量最小化	①时刻的唯一性约束;②机场容量约束; ③航班周转时间约束	单机场,整个航季	整数线性规划模型 ——SAM-II
文献[43]	①总偏移最小化;②总偏移平方的最小化	①时刻唯一性约束;②机场容量约束;③航 班周转时间约束	单机场,整个航季	大邻域搜索技术 (LNS);目标可行泵 算法(FP)
文献[44]	①总偏移最小化;②最大偏移最小;③偏移 时隙数量最小	①时刻的唯一性约束;②机场容量约束; ③航班周转时间约束	单机场,整个航季	2阶段mat启发式算法
文献[45]	①总偏移最小化;②最大偏移量最小	①时刻唯一性约束;②机场容量约束;③航 班飞行时间约束;④航班周转时间约束	单机场,整个航季	大规模邻域搜索/建 设性启发式算法和改 进启发式算法
文献[46]	①准时性能最大化; ②单个航班最大偏移量 最小化/时刻表偏移的加权总和最小化③航 空公司间的公平性最大化	①时刻唯一性约束;②航班计划轮挡时间; ③航班最大偏移量约束;④航班周转时间约 束;⑤15 min航班进场和离场容量约束	单机场,一天	结合随机排队模型和 容量综合利用动态规 划模型的机场拥堵 模型
文献[47]	①航空公司间公平的最大化;②总偏移最小 化	①时刻唯一性约束;②机场容量约束;③航 班周转时间约束;④航空公司公平性约束	单机场,整个航季	混合整数规划模型、 <i>ε</i> 约束法
文献[48]	①总偏移最小化;②最小化与绝对公平的最 大偏差;③最小化与平均公平性(MMR)的 最大偏差;④最小化基尼指数;	①时刻唯一性约束;②机场容量约束;③航 班飞行时间约束;④航班周转时间约束	单机场,整个航季	混合整数规划模型、 <i>ε</i> 约束方法
文献[49]	总偏移最小	①机场容量限制;②空域扇区容量限制	机场网络,一天	迭代局部搜索算法
文献[50]	总偏移和运行延误加权总和的最小	①时刻唯一性约束;②机场容量约束;③航 班飞行时间约束;④航班周转时间约束	机场网络,整个航季	随机规划,并行算法
文献[51]	①空域用户的总成本最小;②偏移时隙数量 最小	①时刻的唯一性约束;②机场谷重约束; ③航班对前后航班一致约束;④请求对的 最小、最大间隔时间约束;⑤航班周转时间 约束	机场网络,多天	整数线性规划模型 ——SOSTA
文献[51]	①总偏移最小化;②最大偏移最小;③被拒 绝的时刻总数最小;④被偏置的时刻总数最 小	①时刻唯一性约束;②机场容量约束;③航 班飞行时间约束;④航班周转时间约束	单机场,整个航季	启发式算法/基于优 先级的多目标舱位分 配模型(PSAM)
文献[52]	①总偏移最小化;②绝对公平的最大偏离 (MDA)最小化;③偏离航司偏好偏移的偏 移量最小化;	①时刻唯一性约束;②机场容量约束;③航 班周转时间约束	单机场,整个航季	混合整数规划模型、 <i>ε</i> 约束方法
文献[53]	总偏移最小化	①时刻唯一性约束;②机场容量约束;③航 班周转时间约束	单机场,整个航季	分支切断方法、整数 规划模型

1071

配额外标准等。在过去的十几年中,越来越多的学 者关注时刻分配问题。在现有文献中,单机场、机 场网络层面和机场群的时刻分配都有相关研究 开展。

3.2 单机场航班时刻分配

单机场航班时刻分配是在机场容量和机场/航空公司运营限制下,给定航班时刻分配策略,将机场时刻分配给各个航空公司的过程,本质上属于资源受限的分配问题。对于资源受限分配问题的统一表示法和分类方案,可以参考 Brucker 等的工作^[54]。

Pritsker 等提出在现实情况下广泛适用的一般 性资源受限调度问题的规划建模方法^[55]。在此基 础上, Zografos 等于 2012 年构建了单机场航班时刻 优化问题的整数线性规划模型^[28]。该模型在 IATA WSG^[6]和 EU 95/93^[2]航班时刻管理规则的基础上, 考虑了机场容量和航班周转时间的限制,引入时刻 请求的优先级组概念,以计划延误最小为目标进行 分层求解。模型通过 3 个协调机场的实际航班数 据进行验证,结果表明能够大大降低航空公司请求 时刻和实际分配航班时刻之间的偏移。

航班时刻配置模型逐步由单目标优化问题向 多目标优化方向转变[43]。优化的目标并不仅仅局 限于总的时刻偏移最小化,还可以包括最小化最大 时刻偏移量等。Zografos 等权衡了时刻分配效率和 可接受性,构造了2个双目标时刻协调模型^[32]。第 1个规划模型考虑最小化总计划偏移和最大可接受 时刻偏移量,第2个模型考虑最小化总计划偏移和 分配在可接受时间窗之外的时刻请求的数量。结 果表明牺牲少量航班时刻协调效率就可有效提高 航空公司对协调航班时刻的可接受度。2018年, Ribeiro 等考虑了 IATA 时刻分配规则^[44], 建立了航 班时刻分配优化模型,并通过模型优化结果与航班 时刻实际配置结果对比,证明了该模型能够更大限 度的满足航空公司的需求、充分利用机场公布容 量。他们构建了一个多目标字典最小化的航班时 刻分配模型^[45]。该模型还考虑 IATA 指南中的一系 列复杂的优先级和规定,最后模型以字典最小化方 式而不是分层次法求解。此外,还进一步探讨了如 何调整当前的 IATA 时刻分配规则,以此提高航班 时刻协调机场的时刻分配效率和效益^[46]。

随着研究的进展,越来越多的学者开始关注航 班时刻配置中的公平性。Zografos和Jiang^[47-48]引 入了公平性,提出一个双目标模型研究计划偏移和 公平性的权衡问题,并使用*ε*约束方法求解。针对 一个拥挤机场,模型分别进行了分层和非分层求 解。非分层情况下求解模型,忽略了优先级组,所 有的时刻将同时进行分配。分析表明,在2种情况 下计划偏移的很小"牺牲"的增加,将会使公平性 显著提高。

航班时刻配置实际工作中还面临航空公司需 求的不确定性及时刻申请取消等问题。在对欧洲 机场航班时刻配置研究时, Odoni 发现在欧洲 32 个 国家的所有航班时刻协调机场,大约有10%的航班 时刻在初始阶段分配给航空公司之后并未被使用 (约为100万个航班时刻)^[7],造成航班时刻资源的 严重浪费。目前航班时刻协调员面临的最大挑战 是需要在各种容量限制下满足航空公司的需求。 容量限制的数量和种类的增加被认为是航班时刻 分配复杂性的主要来源之一。最早通过数学建模 的航班时刻配置工作仅考虑了机场跑道容量和飞 机周转时间限制。后来,一些机场的停机坪和/或机 场航站楼的容量可能也成为限制因素,相关模型已 扩展到能够捕获航站楼和停机坪的容量限制。在 欧洲,大部分机场必须同时为申根和非申根航班提 供服务。因此,机场能够处理国际旅客数量的能力 可能会限制可以安排的国际航班时刻数量。

单机场航班时刻配置过程中的难点和所面临 的挑战总结如下。

 1)单机场航班时刻配置的主要难点是时刻配置需要满足各种约束限制。这些约束来自机场的 公布容量、航班连接、排班规律性/时刻表规律性和 其他技术限制。目前时刻协调员面临的最大挑战 是在各种容量限制下满足航空公司的需求。

2) 要充分考虑复杂而多样的目标和约束,考虑 其在航班时刻配置模型中的合理性和有效性。目 前单机场航班时刻配置模型逐步由单目标优化向 多目标优化转变,容量约束的数量和种类也大大增 加,成为航班时刻分配复杂性的主要来源之一。 Ribeiro 等对 2015 年夏航季葡萄牙 3 个机场的时刻 配置进行了优化,包括2个大中型机场:马德拉机 场(年起降25000架次)和波尔图机场(年起降 85000架次), 及一个大型机场里斯本机场(年起降 约200000架次)[45]。当在考虑跑道容量限制的基 础上添加航站楼容量限制时,模型的约束数量从 3358397个增加到3691247个。继续考虑停机坪 和航站楼容量时,约束的数量高达3872687个,计 算复杂度显著增加。使用 CPLEX 求解器, 经过 7天时间计算也无法找到最优解。因此,需要启发 式算法和更多新颖的方法来解决这个大规模优化 问题。

3)在遵循时刻分配规则的基础上,更加满足实际的航班时刻管理工作需求,充分应对航班时刻配

置实际工作中航空公司需求的不确定性及时刻申 请取消等问题,减少航班时刻资源的浪费。这也是 航班时刻管理部门需要解决的重要问题。

3.3 机场网络层面航班时刻分配

机场网络层面航班时刻分配模型克服了单机 场模型的主要缺陷,即忽略了出发机场和目的地机 场航班时刻协调配置的连续性和一致性,从而导致 航班时刻分配实际效果低下。Castelli 等提出一个 机场网络层面的航班时刻分配模型,旨在最小化因 始发地和目的地机场分配时刻不一致导致的航空 公司申请航班时刻的偏移^[35,49]。该模型基于 Bertsimas 等的空中交通流量管理模型^[36],考虑了航 班飞行所经过的扇区的容量限制。Corolli等进一 步考虑了网络层面时刻分配时机场容量的不确定 性问题,提出了2阶段的随机规划算法^[50]。网络层 面的时刻优化问题面临的一个关键难点是模型的 求解。Pellegrini 等^[51]提出了启发式算法在可接受 时间水平内求得网络层面时刻分配的满意解。上 述研究仅关注一天的时刻配置。Benlic^[56]考虑了机 场容量和航路扇区容量的限制,提出一个2阶段的 智能优化算法,对欧洲机场网络一个航季的时刻进 行了优化。然而,该研究并未考虑航空公司时刻申 请的历史优先权等因素,仅通过模拟生成的时刻申 请数据进行验证。

网络层面的时刻配置研究并未考虑航班时刻 配置过程中的规则限制。已有算法在求解大规模 机场航班时刻配置时未能考虑历史时刻、新进航空 承运人等约束,存在求解效率低甚至无法求到最优 解等问题。

机场网络层面航班时刻分配过程中的现有难 点和所面临的挑战如下。

 1)机场网络层面的航班时刻配置应当充分 考虑航班时刻配置过程中的规则限制,这是一个目 前尚未解决的技术难点。在机场网络运行层面是 否应当考虑空域容量约束,也是一个值得探讨的问题。

2)为机场网络层面航班时刻配置优化问题开 发及时有效的求解方法,将是一个巨大的挑战。网 络层面航班时刻配置模型的约束数量和计算复杂 度显著增加,已有算法可能无法求出最优解。

3)由于机场所属国家或地区不同,如果对机场 网络中所有的机场航班时刻进行统一配置,还会面 临一些政策等方面的问题。

因此,现阶段实施机场网络层面航班时刻统一 配置的条件尚未成熟。

3.4 机场群时刻配置

近年来随着机场群的建设发展涌现出一个新

的问题是机场群的航班时刻资源协同配置。早期 Hansen 和 Weidner 定义多机场系统的边界为50 km^[57],而 Bonnefoy 和 Hansman 随后将其扩大至50 mi(约 80.5 km)^[58]。虽然对机场群并无一个统一的定义,但机场群一般是指存在于城市群周围的机场。城市群是城市发展到一定阶段的空间组织形式,是在地域上集中分布的若干城市和特大城市集聚而成的庞大的、多核心、多层次的城市集团,是生产要素、空间资源和流通市场一体化优化的对象空间^[59]。机场群并非区域内多个机场的简单集合,更是以协同运行和差异化发展为主要特征的多机场体系。

自 2009 年以来,众多学者针对多机场系统的 运行开展了深入的研究[60-68]。文献 [61-62, 68-70] 对 亚特兰大、洛杉矶、纽约和迈阿密4个多机场系 统,从机场构型相关性、空域耦合程度、交通流的 交互及环境天气等限制因素进行了分析和对比,为 NextGen 计划下多机场系统的运行概念提供理论基 础。文献 [60,69] 建立了基于系统论的框架体系对 多机场系统的运行相关性进行研究,对3个多机场 系统建立了网络拓扑结构,用延误对服务运行的相 关性进行分析。在多机场系统中,纽约多机场系 统、旧金山湾区多机场系统一直都是重要研究对 象^[69], 而研究出发点包括空域运行^[70-71]、交通流模 式识别[67, 72-75]、离场放行[76]、容量提升等方面[66-77]。 2015年之后,多机场研究的重点转向多机场一体化 运行^[78-80]。2020年, Yoo 等利用美国航空航天局开 发的交通管理仿真工具,对纽约机场群的2种不同 运行情景进行仿真分析,验证了在纽约机场群内一 体化需求管理的效益^[81]。国内针对多机场研究最 早始于 2009 年, Wang 等对上海多机场的离场航班 排序问题进行了探讨^[82]。随后,相关学者针对京津 冀、长三角和珠三角的多机场系统运行开展了相关 研究^[83]。目前多机场研究的重点主要是机场群内 交通模式特征识别、空域运行效率和进离场战术优 化等,鲜有针对机场群内需求管理和航班时刻配置 的研究。

由于机场群内机场之间的距离较近,各个机场 的进离场航班之间存在相互影响,使得所有机场很 难都在最优的配置下运行。这意味着对一个机场 进行资源分配和航班时刻优化时,不仅仅需要考虑 当前机场的进离场航班之间的权衡,同时还需要考 虑多机场系统中其他机场的进离场容量的权衡。 深入理解机场容量的权衡及他们与其他外部因素 (如单机场和多机场的运行条件,机场布局)的相关 性,对于高效利用机场容量非常重要。高效利用机 场群时刻资源面临如下挑战。

合理的分工定位、优化的航线网络结构是机场群高效运行的前提和基础。机场群内多个枢纽机场、干线机场和支线机场并存,机场功能定位存在差异,机场发展目标和交通需求不同。已有多机场描述性分析工作并未考虑这些重要因素,机场群运行(交通需求)模式识别结果对提升机场群时刻资源管理效率作用甚微。如何综合机场群整体及个体发展需求,优化机场群内交通流,对于提高机场群航班时刻配置效果十分重要。

2)科学合理确定机场群公布容量是一个巨大的挑战。单机场稳健性公布容量确定方法已经是一个尚未解决的难题。在充分考虑机场群内机场功能定位、机场群运行交互等因素的同时,来确定机场和机场群的公布容量,鲜有相关工作开展,这是一个极具挑战的课题。

3) 机场群航班时刻资源优化配置应考虑目标 和约束条件的多样性及不确定性,这是一个需要攻 克的技术难题。机场群航班时刻资源管理应考虑 不同航空公司在协调机场、辅协调机场和非协调机 场的时刻需求,综合机场、空域容量和运行等方面 限制,权衡机场及航空公司的利益,为未来的6个 月分配对应的时刻。无论是优化建模还是应用求 解都是一个不小的挑战。

4 结束语

合理配置机场航班时刻资源关系到航空运输 系统运行的效率和效益。世界上大多数的繁忙机 场主要采用基于行政手段的时刻配置方法。本文 首先对机场航班时刻资源管理中的关键问题进行 梳理,主要包括航班时刻配置关键概念、航班时刻 配置办法和程序、机场公布容量及其确定方法、公 布容量确定所存在的问题和所面临的挑战等,然后 从单机场航班时刻分配、机场网络层面时刻分配和 机场群时刻配置3个方面,重点回顾了基于行政分 配的航班时刻配置手段和方法。航班时刻配置的 结果直接关系到国家的经济命脉,对整个交通运输 体系具有举足轻重的作用。因此,基于市场化的航 班时刻配置方法难以在航班时刻初级市场配置中 得以应用。市场化手段唯一可发挥作用的地方是 航班时刻的次级市场配置。当前的航班时刻管理 办法允许时刻在航空公司之间流动,但是缺乏明确 合理的监管机制。一个关键的问题是不完全竞争 与不完全法规之间的平衡。随着航空运输业和其 他运输方式的发展,科学高效管理繁忙机场的航班 时刻资源将会面临新一轮的持续挑战。

参考文献(References)

 [1] 中国民用航空局. 民航航班时刻管理办法: 民航发[2018]1号[EB/ OL]. (2018-01-05)[2022-04-25]. https://www.caac.gov.cn/XXGK/ XXGK/ZFGW/201802/t20180224_49532.html.

Civil Aviation Administration of China. Civil aviation flight slot management measures: CAAC [2018] No. 1[EB/OL]. (2018-01-05)[2022-04-25]. https://www.caac.gov.cn/XXGK/XXGK/ZFGW/ 201802/(20180224_49532.html.

- UNION E. Council Regulation (EEC) No 95/93 of 18 January 1993 on common rules for the allocation of slots at Community airports [EB/OL]. (1993-01-18)[2022-04-25]. https://www.eumonitor.eu/ 9353000/1/j9vvik7m1c3gyxp/vhckkmr4ymww.
- [3] JORGE D, AN-TUNES RIBEIRO N, PAIS ANTUNES A. Towards a decision-support tool for airport slot allocation: Application to Guarulhos (Sao Paulo, Brazil)[J]. Journal of Air Transport Management, 2021, 93: 102048.
- [4] KATSIGIANNIS F A, ZOGRAFOS K G, FAIRBROTHER J. Modelling and solving the airport slot-scheduling problem with multiobjective, multi-level considerations[J]. Transportation Research Part C:Emerging Technologies, 2021, 124: 102914.
- [5] Airports Council International, International Air Transport Association, Worldwide Airport Coordinators Group. Worldwide Airport Slot Guidelines (WASG)[R]. Geneva: IATA, 2020.
- [6] IATA. Worldwide Slot Guidelines (WSG) [R]. Geneva: IATA, 2011.
- [7] ODONI A R. A review of certain aspects of the slot allocation process at Level 3 airports under regulation 95/93: ICAT-2020-09 [R].
 Cambridge: MIT International Center for Air Transportation, 2020.
- [8] NEUFVILLE R, ODONI A R, BELOBABA P, et al. Airport systems planning, design, and management [M]. 2nd ed. New York: McGraw Hill, 2013: 89-95.
- [9] JACQUILLAT A, ODONI A R. A roadmap toward airport demand and capacity management[J]. Transportation Research Part A:Policy and Practice, 2018, 114: 168-185.
- [10] BALL M, BARNHART C, NEMHAUSER G, et al. Chapter 1 air transportation: Irregular operations and control[M]. Amsterdam: Elsevier, 2007: 1-67.
- [11] STRAKA P, MARTIN P, VIMARD V, et al. Capacity assessment at Dublin Airport for the purpose of setting slot coordination parameters: P2410D008 [R]. Hampshire: Askhelio: 2017.
- [12] TAY G, KARPSTEIN R, HORNUNG M. Airport capacity assessment of the air transportation network in selected global regions [C]// Proceedings of the Aviation Technology, Integration, and Operations Conference. Reston: AIAA, 2018: 4234.
- [13] MORISSET T. Comparing capacities and delays at major European and American airports[D]. Cambridge: Massachusetts Institute of Technology, 2011.
- TRANSSOLUTIONS, CONSULTING F, RAKAS J, et al. Defining and measuring aircraft delay and airport capacity thresholds[M].
 Washington, D. C. : Transportation Research Board, 2014.
- [15] O'FLYNN, S. Airport capacity assessment metho-dology: 1.1[R].

Brussels: EUROCONTROL, 2016.

[16] 胡明华.空中交通流量管理理论与方法[M].北京:科学出版社,2010.

HU M H. Theory and method of air traffic flow management[M]. Beijing: Science Press, 2010 (in Chinese).

- [17] RAMAMOORTHY K, HUNTER G. Simulation-based airport capacity estimation[C]// Proceedings of the Aviation Technology, Integration, and Operations Conference. Reston: AIAA, 2013: 4354.
- [18] DHIR V. Estimation of runway capacity at Changi Airport using AirTOP and Monte Carlo analysis[D]. Singapore: Nanyang Technological University, 2015.
- [19] BAZARGAN M, FLEMING K, SUBRAMANIAN P. A simulation study to investigate runway capacity using TAAM[C]// Proceedings of the Winter Simulation Conference. Piscataway: IEEE Press, 2003: 1235-1243.
- [20] SCHINWALD C, PLÖTNER K O, HORNUNG M. Using airport fast-time simulation models to increase the quality of airport capacity utilization studies[C]// Proceedings of the AIAA Modeling and Simulation Technologies Conference. Reston: AIAA, 2016: 421.
- [21] GILBO E P. Airport capacity: Representation, estimation, optimization[J]. IEEE Transactions on Control Systems Technology, 1993, 1(3): 144-154.
- [22] KICINGER R, CHEN J T, STEINER M, et al. Probabilistic airport capacity prediction incorporating weather forecast uncertainty[C]// Proceedings of the AIAA Guidance, Navigation, and Control Conference. Reston: AIAA, 2014: 1465.
- [23] GENTRY J, DUFFY K, SWEDISH W J. Airport capacity profiles: F055-L11-014[R]. Washington, D.C.: FAA, 2014.
- [24] 许辰澄. 基于历史数据的机场系统运行耦合特性研究[D]. 南京: 南京航空航天大学, 2020: 39-67.

XU C C. Research on coupling characteristics of airport system operation based on historical data[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2020: 39-67(in Chinese).

- [25] RAILSBACK P, SHERRY L. Survey of rationales and methods for determining declared airport capacity[C]// Transportation Research Board 85th Annual Meeting. Wangshington, D.C.: TRB, 2006: 01029356.
- [26] STAMATOPOULOS M A, ZOGRAFOS K G, ODONI A R. A decision support system for airport strategic planning[J]. Transportation Research Part C:Emerging Technologies, 2004, 12(2): 91-117.
- [27] KÖSTERS D. Study on the usage of declared capacity at major German airports: 2007-01-18[R]. Aachen: RWTH Aachen University, 2007.
- [28] ZOGRAFOS K G, SALOURAS Y, MADAS M A. Dealing with the efficient allocation of scarce resources at congested airports[J]. Transportation Research Part C:Emerging Technologies, 2012, 21(1): 244-256.
- [29] SWAROOP P, ZOU B, BALL M O, et al. Do more US airports need slot controls? A welfare based approach to determine slot levels[J]. Transportation Research Part B:Methodological, 2012, 46(9): 1239-1259.
- [30] BARNHART C, FEARING D, ODONI A, et al. Demand and capacity management in air transportation[J]. EURO Journal on Trans-

portation and Logistics, 2012, 1(1-2): 135-155.

- [31] ZOGRAFOS K G, MADAS M A, ANDROUTSOPOULOS K N. Increasing airport capacity utilisation through optimum slot scheduling: Review of current developments and identification of future needs[J]. Journal of Scheduling, 2017, 20(1): 3-24.
- [32] ZOGRAFOS K G, ANDROUTSOPOULOS K N, MADAS M A. Minding the gap: Optimizing airport schedule displacement and acceptability[J]. Transportation Research Part A:Policy and Practice, 2018, 114: 203-221.
- [33] BALL M O, BERARDINO F, HANSEN M. The use of auctions for allocating airport access rights[J]. Transportation Research Part A:Policy and Practice, 2018, 114: 186-202.
- [34] 丁蓉蓉. 航班时刻资源拍卖与协同分配问题研究[D]. 南京: 南京 航空航天大学, 2018: 37-112.
 DING R R. Research on auction and co-allocation of flight slot resource[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2018: 37-112(in Chinese).
- [35] CASTELLI L, PELLEGRINI P, PESENTI R. Airport slot allocation in Europe: Economic efficiency and fairness[J]. International Journal of Revenue Management, 2012, 6(1/2): 28.
- [36] BERTSIMAS D, LULLI G, ODONI A. An integer optimization approach to large-scale air traffic flow management[J]. Operations Research, 2011, 59(1): 211-227.
- [37] SHENG D, LI Z C, FU X W. Modeling the effects of airline slot hoarding behavior under the grandfather rights with use-it-or-lose-it rule[J]. Transportation Research Part E:Logistics and Transportation Review, 2019, 122: 48-61.
- [38] MIRANDA V A P, OLIVEIRA A V M. Airport slots and the internalization of congestion by airlines: An empirical model of integrated flight disruption management in Brazil[J]. Transportation Research Part A:Policy and Practice, 2018, 116: 201-219.
- [39] AVENALI A, D'ALFONSO T, LEPORELLI C, et al. An incentive pricing mechanism for efficient airport slot allocation in Europe[J]. Journal of Air Transport Management, 2015, 42: 27-36.
- [40] FUKUI H. An empirical analysis of airport slot trading in the United States[J]. Transportation Research Part B:Methodological, 2010, 44(3): 330-357.
- [41] BASSO L J, ZHANG A M. Pricing vs. slot policies when airport profits matter[J]. Transportation Research Part B:Methodological, 2010, 44(3): 381-391.
- [42] VERHOEF E T. Congestion pricing, slot sales and slot trading in aviation[J]. Transportation Research Part B:Methodological, 2010, 44(3): 320-329.
- [43] ANDROUTSOPOULOS K N, MANOUSAKIS E G, MADAS M A. Modeling and solving a bi-objective airport slot scheduling problem[J]. European Journal of Operational Research, 2020, 284(1): 135-151.
- [44] RIBEIRO N A, JACQUILLAT A, ANTUNES A P, et al. An optimization approach for airport slot allocation under IATA guidelines[J]. Transportation Research Part B:Methodological, 2018, 112: 132-156.
- [45] RIBEIRO N A, JACQUILLAT A, ANTUNES A P. A large-scale neighborhood search approach to airport slot allocation[J]. Trans-

portation Science, 2019, 53(6): 1772-1797.

- [46] RIBEIRO N A, JACQUILLAT A, ANTUNES A P, et al. Improving slot allocation at Level 3 airports[J]. Transportation Research Part A:Policy and Practice, 2019, 127: 32-54.
- [47] ZOGRAFOS K G, JIANG Y. A Bi-objective efficiency-fairness model for scheduling slots at congested airports[J]. Transportation Research Part C:Emerging Technologies, 2019, 102: 336-350.
- [48] JIANG Y, ZOGRAFOS K G. A decision making framework for incorporating fairness in allocating slots at capacity-constrained airports[J]. Transportation Research Part C:Emerging Technologies, 2021, 126: 103039.
- [49] PELLEGRINI P, CASTELLI L, PESENTI R. Secondary trading of airport slots as a combinatorial exchange[J]. Transportation Research Part E:Logistics and Transportation Review, 2012, 48(5): 1009-1022.
- [50] COROLLI L, LULLI G, NTAIMO L. The time slot allocation problem under uncertain capacity[J]. Transportation Research Part C:Emerging Technologies, 2014, 46: 16-29.
- [51] PELLEGRINI P, BOLIĆ T, CASTELLI L, et al. SOSTA: An effective model for the Simultaneous Optimisation of Airport SloT Allocation[J]. Transportation Research Part E:Logistics and Transportation Review, 2017, 99: 34-53.
- [52] FAIRBROTHER J, ZOGRAFOS KG, GLAZEBROOK KD. A slotscheduling mechanism at congested airports that incorporates efficiency, fairness, and airline preferences[J]. Transportation Science, 2019, 54(1): 115-138.
- [53] FAIRBROTHER J, ZOGRAFOS KG. Optimal scheduling of slots with season segmentation[J]. European Journal of Operationtal Research, 2021, 291(3): 961-982.
- [54] BRUCKER P, DREXL A, MÖHRING R, et al. Resource-constrained project scheduling: Notation, classification, models, and methods[J]. European Journal of Operational Research, 1999, 112(1): 3-41.
- [55] PRITSKER A, WAITERS LJ, WOLFE PM. Multiproject scheduling with limited resources: A zero-one programming approach[J]. Management Science, 1969, 16(1): 93-108.
- [56] BENLIC U. Heuristic search for allocation of slots at network level[J]. Transportation Research Part C:Emerging Technologies, 2018, 86: 488-509.
- [57] HANSEN M, WEIDNER T. Multiple airport systems in the United States: Current status and future prospects[J]. Transportation Research Record Journal of the Transportation Research Board, 1995: 8-17.
- [58] BONNEFOY P A, HANSMAN R J. Emergence of secondary airports and dynamics of regional airport systems in the United States[J]. Massachusetts Institute of Technology, 2006.
- [59] 傅志寰,陆化普.城市群交通一体化:理论研究与案例分析[M].
 北京:人民交通出版社, 2016.
 FU Z H, LU H P. Urban agglomeration transportation integration:

Theory and case study[M]. Beijing: China Communications Press, 2016 (in Chinese).

[60] AYYALASOMAYAJULA S, DELAURENTIS D. Developing strategies for improved management of airport metroplex resources[C]// Proceedings of the 9th AIAA Aviation Technology, Integration, and Operations Conference. Reston: AIAA, 2009: 7036.

- [61] REN L L, CLARKE J P, SCHLEICHER D, et al. Contrast and comparison of metroplex operations: An air traffic management study of Atlanta, los angeles, New York, and Miami[C]// Proceedings of the 9th AIAA Aviation Technology, Integration, and Operations Conference. Reston: AIAA, 2009: 7134.
- [62] TIMAR S, SCHLEICHER D, SARAF A, et al. Metroplex demand analysis and applications[C]// Proceedings of the AIAA Modeling and Simulation Technologies Conference. Reston: AIAA, 2010: 7611.
- [63] CAPOZZI B, ATKINS S, CHOI S. Towards optimal routing and scheduling of metroplex operations[C]// Proceedings of the 9th AI-AA Aviation Technology, Integration, and Operations Conference. Reston: AIAA, 2009: 7037.
- [64] DONALDSON A, HANSMAN R J. Capacity improvement potential for the New York metroplex system[C]// Proceedings of the 10th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference. Reston: AIAA, 2010: 9285.
- [65] VINAY V, AYYALASOMAYAJULA S, DELAURENTIS D. A metroplex capacity enhancement strategy through coordinated use of runway dependencies[C]// Proceedings of the 11th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference. Reston: AIAA, 2011: 7067.
- [66] CAPOZZI B, ATKINS S. A hybrid optimization approach to air traffic management for metroplex operations[C]// Proceedings of the 10th AIAA Aviation Technology, Integration, and Operations Conference. Reston: AIAA, 2010: 9062.
- [67] AMARO CARMONA M A, SAEZ NIETO F, VERDONK GAL-LEGO C E. A data-driven methodology for characterization of a terminal manoeuvring area in multi-airport systems[J]. Transportation Research Part C:Emerging Technologies, 2020, 111: 185-209.
- [68] SCHLEICHER D, CLARKE J P, SARAF A, et al. A concept of operations for a NextGen metroplex scheduling concept[C]// Proceedings of the 11th AIAA Aviation Technology, Integration, and Operations Conference. Reston: AIAA, 2011: 7065.
- [69] MARZUOLI A C, FERON E, HANSEN M, et al. Multimodality in a metroplex environment: A case study in the San Francisco Bay Area[C]// Proceedings of the AIAA Infotech @ Aerospace. Reston: AIAA, 2015: 1571.
- [70] CLARKE J P, REN L L, MCCLAIN E, et al. Evaluating concepts for metroplex operations[C]// Proceedings of the 10th AIAA Aviation Technology, Integration, and Operations Conference. Reston: AIAA, 2010: 9249.
- [71] WIELAND F, TYAGI A, KUMAR V, et al. METROSIM: A metroplex-wide route planning and airport scheduling tool[C]// Proceedings of the 14th AIAA Aviation Technology, Integration, and Operations Conference. Reston: AIAA, 2014: 2162.
- [72] MURÇA M C R, HANSMAN R J. Identification, characterization, and prediction of traffic flow patterns in multi-airport systems[J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 20(5): 1683-1696.

[73] SIDIROPOULOS S, MAJUMDAR A, HAN K, et al. Identifying

significant traffic flow patterns in Multi-Airport Systems Terminal Manoeuvring Areas under uncertainty[C]// Proceedings of the 16th AIAA Aviation Technology, Integration, and Operations Conference. Reston: AIAA, 2016: 3162.

- [74] ZHANG Y, TANG H L, KNORR D, et al. Measuring the benefits of NextGen metroplex in convective weather: case study of north texas metroplex[C]// Proceedings of the Aviation Technology, Integration, and Operations Conference. Reston: AIAA, 2018: 4235.
- [75] MURÇA M C R, HANSMAN R J, LI L S, et al. Flight trajectory data analytics for characterization of air traffic flows: A comparative analysis of terminal area operations between New York, Hong Kong and Sao Paulo[J]. Transportation Research Part C:Emerging Technologies, 2018, 97: 324-347.
- [76] CHEVALLEY E, PARKE B, KRAUT J, et al. Scheduling and delivering aircraft to departure fixes in the NY metroplex with controller-managed spacing tools[C]// Proceedings of the 15th AIAA Aviation Technology, Integration, and Operations Conference. Reston: AIAA, 2015: 2428.
- [77] IDRIS H R, SHEN N, SARAF A P, et al. Queue buffer sizing for efficient and robust integrated departure scheduling[C]// Proceedings of the 16th AIAA Aviation Technology, Integration, and Operations Conference. Reston: AIAA, 2016: 3901.
- [78] IDRIS H. Queuing analysis of interdependencies between multi-airport system operations[C]// Proceedings of the 9th AIAA Aviation

Technology, Integration, and Operations Conference. Reston: AI-AA, 2009: 7039.

- [79] XUE M, ZELINSKI S, MULFINGER D G. Uncertainty analysis of integrated departures and arrivals: A los angeles case study[C]// Proceedings of the Aviation Technology, Integration, and Operations Conference. Reston: AIAA, 2013: 4347.
- [80] XUE M, ZELINSKI S. Optimization of integrated departures and arrivals under uncertainty[C]// Proceedings of the Aviation Technology, Integration, and Operations Conference. Reston: AIAA, 2013: 4322.
- [81] YOO H S, EVANS A D, KULKARNI D, et al. Benefit assessment of the integrated demand management concept for multiple New York metroplex airports[C]// Proceedings of the AIAA Scitech 2020 Forum. Reston: AIAA, 2020: 1400.
- [82] WANG Y J, HU M H, SUI D, et al. Departure scheduling in a multi-airport system[C]//Proceedings of the 8th USA/Europe Air Traffic Management Research and Development Seminar. California: ATM, 2009.
- [83] 彭瑛. 多机场群运行评估与优化研究[D]. 南京: 南京航空航天大学, 2014: 36-90.

PENG Y. Evaluation and optimazation of metroplex operation[D].Nanjing: Nanjing University of Aeronautics and Astronautics, 2014: 36-90(in Chinese).

Research progress on airport slot allocation

WANG Yanjun^{*}, SHUI Xiaoyu, WANG Mengyin

(College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China)

Abstract: Most of the busiest airports airports face severe congestion and flight delays. Optimal allocating airport slots is one of the most effective ways in airport demand management. This paper systematically reviews the progress of research in airport slots allocation, as well as the main techniques in setting airport declared capacity and allocating slots. The major challenging issues of stochastic optimization were identified in such areas as the determination of airport declared capacity, the slot allocation in an airport, in an airport network and in the multi-airport system, and the technical complexity. Main research directions in slot allocation are further discussed and suggested.

Keywords: airport capacity; slot allocation; demand and capacity management; strategic air traffic flow management; stochastic optimization

Received: 2022-05-28; Accepted: 2022-07-06; Published Online: 2022-07-19 14:46 URL: link.cnki.net/urlid/11.2625.V.20220718.1808.001

* Corresponding author. E-mail: ywang@nuaa.edu.cn

Foundation item: National Natural Science Foundation of China (U2033203)

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

DOI: 10.13700/j.bh.1001-5965.2022.0453

超声纵扭-低频扭转复合振动攻丝实验

韩凤起1,张德远2,*

(1. 河北工业职业技术大学智能制造学院,石家庄 050091; 2. 北京航空航天大学 机械工程及自动化学院,北京 100191)

摘 要: 针对钛合金低频扭转振动攻丝效率偏低的问题,提出了超声纵扭-低频扭转复合振 动攻丝方法。进行了超声纵扭振动攻丝实验、不同切削液的复合振功攻丝实验和变参数复合振动攻 丝实验,研究复合振动攻丝的工艺效果及参数变化的影响规律。实验结果表明:超声纵扭振动增强 了切削液的润滑效果,降低了攻丝扭矩,与普通连续攻丝相比,M3螺纹孔的超声纵扭振动攻丝可 降低攻丝扭矩约18%;在所选参数范围内,与低频扭转振动攻丝相比,复合振动攻丝最大可降低攻 丝扭矩约30%;复合振动攻丝降低攻丝扭矩的效果,随净切削量、主轴转速的增大和回退量的减小 而降低。

关 键 词: 振动攻丝; 超声纵扭振动; 低频扭转振动; 攻丝扭矩; 切削液润滑 中图分类号: TG62

中国万天与: 1002

文献标志码: A 文章编号: 1001-5965(2024)04-1077-08

钛及钛合金是重要的结构金属,因其密度低, 比强度和比刚度高,耐腐蚀性能、低温性能、抗疲 劳和蠕变性能良好,具有很大的发展潜力和应用前 景,在航空航天、兵器、汽车工业、医疗等领域的应 用日益广泛^[1-3]。

螺纹连接是一种常见且重要的机械连接方式。由于内螺纹的加工受刀具刚性、排屑条件和刀 具直径等因素的限制,加工难度比外螺纹大得多, 是最难加工的曲面类型之一^[4]。钛合金作为典型的 难加工材料,小螺纹孔的攻丝尤其困难,表现为攻 丝扭矩大,丝锥磨损快、易折断。

钛合金弹性模量小、屈强比大,导致已加工表 面回弹量大,攻丝过程中,丝锥后刀面受到严重的 挤压和摩擦,产生较大的摩擦扭矩,致使攻丝扭矩 较大;同时,钛合金切削变形系数和导热系数小,攻 丝过程中,切削刃和摩擦界面温度高^[5-6],较高的切 削温度加剧了丝锥磨损,而丝锥的磨损进一步导致 已加工表面的回弹增大,以及对后刀面摩擦和挤压 的加剧,形成恶性循环。

振动切削是一种脉冲切削,改变了工具与被加 工材料之间的空间-时间存在条件,从而改变了切 削机理,达到了增强切削液润滑效果、降低切削力 和切削温度、提高加工质量和刀具使用寿命等目 的^[7-8]。振动切削应用于攻丝加工,也获得了良好的 效果。

低频振动攻丝方式相对容易实现,开展的研究 较多且比较成熟。张德远和陈鼎昌^[5]建立了重复 切削理论,分析了攻丝力矩的减小与重复切削次数 和切削量的对应关系,解释了低频扭转振动攻丝降 低攻丝扭矩的基本原理。在此基础上,北京航空航 天大学的研究团队^[9-12]研制了伺服电机式振动攻丝 系统,并进行了不锈钢、钛合金、高温合金、复合材 料等难加工材料的攻丝实验,研究结果表明,低频 扭转振动攻丝能够显著降低攻丝扭矩,提高丝锥使 用寿命和螺纹精度。韩荣第^[13]、殷宝麟^[14]等研究 了净切削时间比、频率和振幅与攻丝扭矩的关系,

收稿日期: 2022-06-02; 录用日期: 2022-07-09; 网络出版时间: 2022-08-11 13:28 网络出版地址: link.cnki.net/urlid/11.2625.V.20220810.0922.003

引用格式: 韩凤起,张德远,超声纵扭-低频扭转复合振动攻丝实验 [J].北京航空航天大学学报,2024,50(4):1077-1084. HAN F Q, ZHANG D Y. Ultrasonic longitudinal torsional and low frequency torsional compound vibration tapping experiment [J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50(4): 1077-1084 (in Chinese).

基金项目:中央引导地方科技发展资金 (216Z1004G);河北省教育厅科学技术研究项目 (ZD2020331);河北工业职业技术大学博士基金 (zk202101)

^{*}通信作者. E-mail: zhangdy@ buaa.edu.cn

结果表明,应力波促使微裂纹产生和扩展,降低攻 丝扭矩,且净切削时间比对攻丝扭矩影响最大。王 贵成等^[15]研究了不同振动方式对攻丝效果的影响, 研究表明,"进—停—退—停—进—停"的振动方式 具有良好的断屑和润滑效果,有利于提高丝锥的寿命。

超声振动加工技术凭借其特有的加工优势,已成为机械加工领域重要的发展方向^[16]。近年来,已有学者开始进行超声振动攻丝技术的探索。Kuo^[17]进行了钛合金纵向超声振动攻丝实验,研究表明,与普通攻丝相比,M3、M3.5 螺纹孔的超声振动攻丝纽矩分别降低了 28% 和 8%。Tsao 和 Kuo^[18]进行了不同涂层丝锥的超声振动攻丝实验,结果表明,超声振动攻丝能够提高螺纹质量。Pawar 等^[19]进行了不同振动形式的超声振动攻丝实验,发现纵-扭振动比纵向振动具有更好的降低攻丝扭矩的效果。

低频扭转振动攻丝易于实现,操作简单,在实际应用中取得了较好的效果,但由于需要较小的净切削量和较多的重复切削次数才能获得较好的降低攻丝扭矩的效果,导致加工效率相对较低,限制了其广泛应用。如果将超声振动和低频扭转振动同时应用于振动攻丝,将能够同时发挥2种振动方式的优势,获得更好的加工效果。

本文在钛合金 TC4 上进行了超声纵扭-低频扭 转复合振动攻丝实验,并对复合振动攻丝与低频扭 转振动攻丝的效果进行了对比,研究了低频扭转振 动参数对复合振动攻丝的影响,为今后的研究和应 用提供依据。

1 实 验

1.1 超声振动方式的选择

振动方向的选择是超声振动加工应用的一个 重要问题。文献 [7] 的研究结果证明, 虽然各种振 动方向都能够改善切削状况,但当振动方向沿主切 削力方向时,加工效果最好。本文采用带有螺旋槽 的变幅杆,将压电陶瓷的纵向振动转换为具有扭转 和纵向运动分量的右旋方向的纵-扭复合振动方 式,以期得到良好的攻丝效果。不同规格螺纹的升 角不同,针对每一规格螺纹制造振动方向与螺纹升 角相同的纵扭换能器难度较大。同时,鉴于钛合金 材料弹性模量小、已加工表面回弹量大的特点,在 纵扭振动方向和螺纹升角不同的情况下,更有利于 攻丝过程中在丝锥切削刀齿侧后刀面和已加工表 面之间形成间隙,一方面减少接触面积,另一方面 更有利于切削液的渗透。另外,考虑到小直径丝锥 的扭转刚性较弱,采取较大的扭转振动振幅,有利 于在刀齿前刀面和切屑之间形成更大间隙,增加切 削液的润滑效果。因此,实验中的纵扭振动应采用 尽量大的扭转和纵向振动分量。

1.2 实验系统及基本实验条件

本文的实验系统如图 1 所示, 所采用的超声纵 扭振动换能器是在文献 [10] 中换能器的基础上改 进的, 结构如图 2 所示。





在 400 倍的工具显微镜下观察,当输入激励频 率为 23.5 kHz 时,换能器纵扭振动的扭转振动分量 相对较大。在保证换能器和驱动电源正常工作的 情况下,采用尽量高的输入电压,使换能器输出较 大的振幅。当输入电压为 45 V(驱动电源最高工作 电压 50 V)、输入电流为 3 A 时,换能器端部的扭转 振动分量约 4 μm,纵向振动分量约 5.5 μm,合成的 沿右旋螺旋方向振动的振幅约 7 μm。在此状态下, 向换能器端部洒上一点水,会立刻腾起一团水雾, 说明换能器工作状态良好。实验中,保持换能器工 作频率、输入电压和电流不变,使其在各实验中的 工作状态一致。工件加工成小圆柱,被抱紧在超声 换能器头部的孔内。

丝锥的低频扭转振动由工控机通过伺服控制 器控制低频扭转振动攻丝机主轴实现,参数包括净 切削量*L*_T、回退量*L*_G和主轴转速*V*,可在复合振动 攻丝实验中根据具体实验目的进行设置。 在低频扭转振动攻丝机的工作台上安装一个 浮动工作台,该浮动工作台的台面受较小的力就可 以进行平动,并可以传递扭矩,实现攻丝过程中丝 锥和工件底孔的自动对正,以消除对中误差对攻丝 扭矩测量结果的影响。在浮动工作台上安装 Kistler 9257B 测力仪,超声纵扭振动换能器通过支座固定 在测力仪上。攻丝过程中,测力仪的力信号通过电 荷放大器经 A/D 卡输入到数据采集计算机中。

实验采用 M3 高速钢全磨制 H2 精度机用直槽 丝锥。工件材料为退火状态钛合金 TC4, 洛氏硬度 为 33。在工件上预先加工直径 2.5 mm 的通孔作为 底孔。

2 实验结果和讨论

2.1 超声纵扭振动攻丝与普通连续攻丝的对比

为考查超声纵扭振动攻丝降低攻丝扭矩的效 果,进行了超声纵扭振动攻丝和普通连续攻丝的对 比实验。实验中,低频扭转振动攻丝机主轴转速为 5 r/min。先关闭超声纵扭振动换能器电源进行普 通连续攻丝实验,再更换工件,打开超声纵扭振动 换能器电源进行超声纵扭振动攻丝实验,以此交替 进行,攻丝过程中,采用蓖麻油进行润滑。为防止 丝锥折断,2种方式均在攻丝时间达到100s左右结 束攻丝并退回丝锥,测得的2种方式攻丝扭矩对比 如图3所示,超声纵扭振动攻丝的攻丝扭矩比普通 连续攻丝降低约18%。





由图 3 可见,虽然超声纵扭振动攻丝相比普通 连续攻丝能够明显降低攻丝扭矩,但攻丝扭矩仍然 随攻丝深度的增加而不断增大,最终很可能导致丝 锥折断。因此,仅通过超声纵扭振动达不到理想的 工艺效果。

虽然超声纵扭振动降低攻丝扭矩的效果不如 低频扭转振动显著,但如将二者复合在一起,就可 以在低频扭转振动的基础上进一步降低攻丝扭矩, 从而提高攻丝效率。

2.2 超声纵扭振动的减摩及切削液润滑效果

根据已有的研究结果,超声振动产生的工艺效 果,很大一部分原因是改善了切削液的润滑。为验 证超声纵扭振动对切削液润滑效果的影响,分别选 用水溶性乳化液和蓖麻油作为切削液进行超声纵 扭-低频扭转复合振动攻丝实验。

实验过程中,超声电源和低频扭转振动攻丝机 先同时工作,进行复合振动攻丝,一段时间后关闭 超声电源,仅进行低频扭转振动攻丝,再过一段时 间又打开超声电源,进行复合振动攻丝,如此交替 进行,在此过程中始终用切削液进行充分润滑。

以水溶性乳化液为切削液的实验中,复合振动 攻丝过程中状态良好,但关闭超声电源后,低频扭 转振动攻丝过程中发出刺耳的噪声,再打开超声电 源后噪声消失,采集到的攻丝扭矩如图4所示。可 以看出,复合振动攻丝的攻丝扭矩很小,而且平稳, 但仅低频扭转振动攻丝时攻丝扭矩很大,并且剧烈 波动,这一特点在如图5所示的放大后的攻丝扭矩 对比波形中显示得非常清楚。



图 4 水溶性乳化液为切削液的攻丝扭矩波形对比

Fig. 4 Comparison of tapping torque waveform with watersoluble emulsion as cutting fluid





由于钛合金弹性模量小,导致回弹较大的已加 工表面对丝锥后刀面产生剧烈的摩擦。虽然低频 扭转振动攻丝能够通过重复切削显著降低已加工 表面的最终回弹量,但如图 6 所示的重复切削模 型,在每个振动周期的净切削(*L*_T)阶段的切削过程 及其附近区域的重复切削过程中已加工表面的回 弹量还是较大,与后刀面发生剧烈摩擦,虽然水溶 性乳化液的冷却效果较好,但是润滑效果却相对较 差,因此,低频扭转振动攻丝中由于丝锥后刀面与 工件剧烈摩擦而发出强烈的噪声。在低频扭转振 动中增加超声纵扭振动后,上述摩擦产生的噪声消 失,说明超声纵扭振动明显改善了后刀面的摩擦 状况。



Fig. 6 Repeated cutting model of low frequency torsional vibration tapping

能够降低摩擦系数是超声振动特有的现象。 在相互接触的2个物体中,如果其中一个发生超声 振动,接触界面上的质点在一个周期中受力大小和 方向都会发生变化,导致质点有产生旋转运动的趋 势,与普通情况相比,摩擦状态发生了变化,同时, 振动减小了摩擦的有效面积,并且振动力减小了接 触的正压力,上述因素都有利于减小摩擦^[20]。

本文采用的超声纵扭振动换能器存在纵向振动,纵向振动的横向分离减摩作用降低了丝锥前后 刀面与切屑和已加工表面的摩擦。同时,在工件振动过程中,丝锥侧后刀面和工件的接触界面周期性的分离,出现一定的间隙,切削液在泵吸作用下进入间隙而形成润滑油膜^[21],而且超声振动会产生辐射压,也促使切削液在辐射压的作用下进入侧后刀 面和工件的间隙内形成润滑膜^[20,22],降低丝锥侧后 刀面的摩擦系数。虽然超声纵扭振动过程中丝锥 切削刀齿的顶后刀面与工件不分离,但由于振动存 在一定的扭转振动分量,顶后刀面和工件之间存在 相对往复运动,而且低频扭转振动减轻了已加工表 面与后刀面的挤压,增强了超声振动对切削液的泵 吸作用,改善了切削液的润滑效果。与上述后刀面 的摩擦状况类似,超声纵扭振动换能器存在的扭转 振动分量使丝锥前刀面和切屑之间也存在周期性 的分离,促使切削液更容易渗透到前刀面和切屑之 间,降低了刀-屑摩擦系数。因此,在上述实验中, 当在低频扭转振动攻丝中复合超声纵扭振动后,丝 锥前后刀面的摩擦状况有了明显的改善,减小了摩 擦力,降低了攻丝扭矩。

虽然水溶性乳化液黏性较小,更容易发挥超声振动改善润滑的作用,但由于钛合金攻丝中丝锥后 刀面摩擦严重,一般都采用润滑性较好的植物油作 为切削液,而且图4和图5所示的水溶性乳化液的 低频扭转振动攻丝扭矩波形很乱,不便于数据的测 量,不利于不同参数下复合振动与低频扭转振动的 攻丝效果对比。因此,本文采用蓖麻油作为切削液 进行实验,得到的复合振动攻丝和低频扭转振动攻 丝的攻丝扭矩对比如图7所示。可以看出,采用蓖 麻油后,由于润滑性能增强,低频扭转振动攻丝中 攻丝扭矩也比较稳定,便于实验规律的研究。因 此,后续实验中均采用蓖麻油作为切削液。



图 7 蓖麻油为切削液的攻丝扭矩波形对比

Fig. 7 Comparison of tapping torque waveforms with castor oil as cutting fluid

2.3 工艺参数对复合振动攻丝的影响

2.3.1 基本实验方法

本文进行了不同参数下复合振动攻丝和低频 扭转振动攻丝的对比实验,通过分析不同参数下二 者的攻丝扭矩变化情况,研究低频扭转振动参数对 复合振动攻丝的影响规律。

由于攻丝过程中从开始攻丝到丝锥切削锥部 分完全攻入孔内的期间攻丝扭矩是逐渐增加的,为 保证各参数下基本实验条件的一致性,预先在工件 上攻出约3mm深的螺纹,使得开始实验时切削锥 部分已全部进入孔中。同时,每一组参数采用同一 支丝锥。实验时,先根据选择的参数进行复合振动 攻丝,攻丝深度约0.5mm后关闭超声电源,仅保持 低频扭转振动攻丝,再攻约 0.5 mm 深后停止攻 丝。在此方法下得到的攻丝扭矩特点如图 7 所 示。进行若干次实验,以2种振动攻丝方式下的峰 值攻丝扭矩和回退摩擦扭矩的平均值进行比较。

2.3.2 净切削量的影响

在主轴转速V = 400 r/min和回退量 $L_{G} = 40^{\circ}$ 不 变的情况下,净切削量L_T对2种攻丝方式峰值攻丝 扭矩和回退摩擦扭矩的影响如图8所示,复合振动 攻丝的峰值攻丝扭矩和回退摩擦扭矩相对于低频 扭转振动攻丝的降低比例如图9所示。



5 0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0 5.5 6.0 6.5 净切削量/(°) 图 9 净切削量对复合振动攻丝效果的影响 Fig. 9 Effect of net cutting amount on compound vibration tapping

10

从实验结果可以看出,随着净切削量L_T增大, 2种攻丝方式下的峰值攻丝扭矩和回退摩擦扭矩均 增大,而复合振动攻丝降低攻丝扭矩的效果逐渐降 低。当净切削量为0.5°时,相对于低频扭转振动攻 丝,复合振动攻丝峰值扭矩可降低约30%,而当净 切削量为6°时,复合振动攻丝扭矩仅能降低约 10%

实验中发现,在低频扭转振动攻丝方式下,当

净切削量L_T大于2°后,攻丝过程中会发出因丝锥后 刀面和工件摩擦而产生的噪声,净切削量L_T越大, 噪声越强烈。在复合振动攻丝方式下,只有当净切 削量L_T达到6°后才会有较明显的噪声。这一方面 说明超声纵扭振动增强了切削液的润滑效果,另一 方面也说明超声纵扭振动增强切削液润滑效果的 作用随净切削量的增加而降低。

由于在回退量不变的情况下,重复切削次数随 净切削量的增大而减少,导致已加工表面的回弹量 增加,增大了对丝锥后刀面的挤压和摩擦,而且随 着净切削量的增大,丝锥前后刀面的高接触应力区 长度增加,减弱了超声纵扭振动对切削液的泵吸作 用。因此,为充分发挥超声纵扭振动增强切削液润 滑效果的作用,不宜选择过大的净切削量。

2.3.3 回退量的影响

在主轴转速V = 400 r/min和净切削量 $L_T = 1^{\circ}$ 不 变的情况下,回退量LG对2种攻丝方式攻丝扭矩的 影响如图 10 和图 11 所示。实验结果表明,复合振 动攻丝的峰值攻丝扭矩和回退摩擦扭矩相对于低





2024年

频扭转振动攻丝的降低比例随回退量的增加而 增大。

这是因为在净切削量不变的情况下,随着回退 量的增大,重复切削次数增加,已加工表面对丝锥 后刀面的挤压和摩擦减轻,超声纵扭振动增大切削 液润滑作用的效果得到增强。

2.3.4 主轴转速的影响

在切削量L_T = 2°和回退量L_G = 40°不变的条件 下,主轴转速V对攻丝扭矩的影响如图 12 和图 13 所示。实验结果显示,复合振动攻丝降低攻丝扭矩 的效果随主轴转速的提高而逐渐降低,原因是蓖麻 油的渗透能力相对较弱,随着主轴转速的提高,蓖 麻油在超声纵扭振动的泵吸作用下向丝锥前后刀 面的间隙渗透的时间减少,降低了超声纵扭振动增 强切削液润滑效果的作用。



Fig. 12 Effect of spindle speed on tapping torque





2.3.5 复合振动攻丝效果的综合分析

实验结果表明,在低频扭转振动攻丝的基础上 复合超声纵扭振动,能够有效降低攻丝扭矩,因此, 采用复合振动攻丝方式可为低频扭转振动选择较 大的净切削量、较小的回退量和较高的主轴转速, 从而提高攻丝效率。例如,参考图 8 所示数据,其 他参数不变的情况下,净切削量为 4°的复合振动攻 丝的峰值攻丝扭矩与净切削量为 2°的低频扭转振 动攻丝的峰值攻丝扭矩比较接近,即在保证峰值攻 丝扭矩与净切削量为 2°的低频扭转振动攻丝基本 相同的情况下,采用复合振动攻丝可将攻丝效率提 高接近 1 倍。

与低频扭转振动攻丝相比,由于复合振动攻丝 降低了攻丝扭矩,进一步降低了丝锥钝化对刃口钝 圆和后刀面磨损的敏感性,同时,超声纵扭振动使 丝锥后刀面与工件的瞬时接触面积减小^[17,20],并增 强了切削液的润滑效果,有利于降低后刀面磨损。 因此,复合振功攻丝更有利于提高丝锥使用寿命, 并且切削液润滑效果的增强有助于提高螺纹的表 面质量。

将以复合振动攻丝方式加工的螺纹孔剖开,在 工具显微镜下观察,可见螺纹牙型规整,表面光亮 无明显划痕,如图 14 所示。使用螺纹塞规对加工 的螺纹孔进行检验,所有螺纹孔通规能够顺利通过 且间隙很小,止规旋入 0.5~1 扣,符合合格标准(通 规旋合通过,止规旋合量不超过 2 个螺距,即为合 格),说明采用合适的复合振动攻丝参数可获得良 好的螺纹质量。

需要注意的是,在复合振动攻丝中,通过低频 扭转振动必要的重复切削次数即可基本去除已加 工表面的回弹,如果重复切削次数过多,超声纵扭 振动的振幅过大,有可能造成螺纹超差。因此,当 被加工材料、丝锥直径等加工条件不同时,低频扭 转振动、超声纵扭振动的参数如何进行优化选择, 需要进行进一步的深入系统研究。



图 14 螺纹照片 Fig. 14 Photo of thread

3 结 论

本文针对钛合金 TC4 进行了超声纵扭-低频扭 转复合振动攻丝实验,分析了超声纵扭振动增强切 削液润滑、降低攻丝扭矩的效果及原因,以及工艺 参数对复合振动攻丝的影响规律,结论如下: 1) 超声纵扭振动应用于攻丝加工, 增强了切削 液的润滑效果, 降低了攻丝扭矩。

2) 在超声纵扭-低频扭转复合振动攻丝中, 在 超声纵扭振动参数固定的情况下, 低频扭转振动采 用较小的净切削量、较大的回退量和较低的主轴转 速, 可以获得较大的重复切削次数, 有效减轻已加 工表面和丝锥后刀面的挤压和摩擦, 增强超声纵扭 振动对切削液的泵吸作用, 并增加切削液的渗透时 间, 从而获得更好的切削液的润滑效果, 更大程度 的降低攻丝扭矩。

3) 与低频扭转振动攻丝相比, 超声纵扭-低频 扭转复合振动攻丝可以在保持良好加工状态的前 提下, 显著提高加工效率。

有必要进一步深入系统研究超声纵扭振动参数对工艺效果的影响,确定不同加工条件下(如不同加工对象和螺纹直径等)的复合振动参数优化选择理论和方法,促进超声纵扭-低频扭转复合振动 攻丝技术的推广应用。

参考文献(References)

[1] 李毅, 赵永庆, 曾卫东. 航空钛合金的应用及发展趋势[J]. 材料导报, 2020, 34(增刊 1): 280-282.

LI Y, ZHAO Y Q, ZENG W D. Application and development trend of aviation titanium alloy[J]. Materials Reports, 2020, 34(Sup1): 280-282(in Chinese).

- [2] 原国森, 兖利鹏, 韩艳艳. 钛合金的应用进展[J]. 热加工工艺, 2017, 46(4): 13-16.
 YUAN G S, YAN L P, HAN Y Y. Application progress of titanium alloy[J]. Hot Working Technology, 2017, 46(4): 13-16(in Chinese).
- [3] 刘彬, 刘延斌, 杨鑫, 等. TITANIUM 2008: 国际钛工业、制备技术 与应用的发展现状[J]. 粉末冶金材料科学与工程, 2009, 14(2): 67-73.

LIU B, LIU Y B, YANG X, et al. TITANIUM 2008: Development of international titanium industry, preparation technology and applications[J]. Materials Science and Engineering of Powder Metallurgy, 2009, 14(2): 67-73(in Chinese).

[4] 陈志同. 扭转振动机构学及振动攻丝技术研究[D]. 北京: 北京航空航天大学, 2001: 1-9.

CHEN Z T. Research on torsional vibration mechanism and vibration tapping technology[D]. Beijing: Beihang University, 2001: 1-9 (in Chinese).

- [5] ZHANG D Y, CHEN D C. Relief-face friction in vibration tapping
 [J]. International Journal of Mechanical Sciences, 1998, 40(12): 1209-1222.
- [6] 张利军, 申伟. 钛合金材料攻丝技术的开发[J]. 工具技术, 2015, 49(2): 55-57.

ZHANG L J, SHEN W. Research on tapping technology of titanium alloy material[J]. Tool Engineering, 2015, 49(2): 55-57(in Chinese).

[7] 限部淳一郎. 精密加工: 振动切削 (基础与应用)[M]. 韩一昆, 薛万夫, 孙祥根, 等, 译. 北京: 机械工业出版社, 1985: 19-147.
 KUMABE J. Precision machining: Vibration cutting (foundation

and application)[M]. HAN Y K, XUE W F, SUN X G, et al, translated. Beijing: China Machine Press, 1985: 19-147(in Chinese).

- [8] 李祥林,薛万夫,张日升.振动切削及其在机械加工中的应用
 [M].北京:北京科学技术出版社, 1985: 5-15.
 LI X L, XUE W F, ZHANG R S. Vibration cutting and its application in machining[M]. Beijing: Beijing Science and Technology Press, 1985: 5-15(in Chinese).
- [9] 李光军,张德远. 难加工材料的低频振动攻丝研究[J]. 中国机械
 工程, 2004, 15(6): 494-497.
 LI G J, ZHANG D Y. Investigation of low-frequency vibration tap-

ping in diffcult-cutting matierals[J]. China Mechanical Engineering, 2004, 15(6): 494-497(in Chinese).

[10] 李光军. 振动攻丝及其监控技术研究[D]. 北京: 北京航空航天大学, 2004: 49-100.
 LIG J. Research on vibration tapping and its monitoring techno-

logy[D]. Beijing: Beihang University, 2004: 49-100(in Chinese). [11] 韩凤起. 难加工材料低频扭转振动攻丝理论和工艺的研究[D]. 北

京:北京航空航天大学, 2008: 13-108. HAN F Q. Study on theory and technology of vibration tapping of difficult-to-cutmaterials[D].Beijing:BeihangUniversity,2008:13-108 (in Chinese)

- [12] 刘洋, 韩凤起, 卢慧敏, 等. Al₂O₃p/Al 复合材料振动攻丝螺纹质量研究[J]. 航空精密制造技术, 2008, 44(6): 48-50. LIU Y, HAN F Q, LU H M, et al. Study on thread quality of low-frequency torsional vibration tapping on composite Al₂O₃p/Al[J]. Aviation Precision Manufacturing Technology, 2008, 44(6): 48-50 (in Chinese).
- [13] 韩荣第, 殷宝麟, 崔永鹏. 难加工材料小孔振动攻丝试验研究[J]. 哈尔滨工业大学学报, 2008, 40(1): 50-53.
 HAN R D, YIN B L, CUI Y P. Investigation of small-hole vibration tapping in difficult-to-cut materials[J]. Journal of Harbin Institute of Technology, 2008, 40(1): 50-53(in Chinese).
- [14] YIN B L, HAN R D. Investigation of the torque characteristics in vibration tapping of hardened steel[J]. International Journal of Machine Tools and Manufacture, 2006, 46(6): 623-630.
- [15] 王贵成,朱金辉,王磊,等.振动方式对攻丝加工效果影响的实验研究[J].机械科学与技术,2008,27(7):857-860.
 WANG G C, ZHU J H, WANG L, et al. Experimental study of the effect of vibration mode on vibratory tapping[J]. Mechanical Science and Technology for Aerospace Engineering, 2008, 27(7): 857-860(in Chinese).
- [16] 张德远,刘逸航,耿大喜,等. 超声加工技术的研究进展[J]. 电加 工与模具, 2019(5): 1-10. ZHANG D Y, LIU Y H, GENG D X, et al. The research progress of ultrasonic machining technology[J]. Electromachining & Mould, 2019(5): 1-10(in Chinese).
- [17] KUO K L. Experimental investigation of ultrasonic vibrationassisted tapping[J]. Journal of Materials Processing Technology, 2007, 192-193: 306-311.
- [18] TSAO C C, KUO K L. Ultrasonic-assisted vibration tapping using taps with different coatings[J]. Transactions of Nonferrous Metals Society of China, 2012, 22: s764-s768.
- [19] PAWAR S, PATIL S, PAWAR P, et al. Effect of axial and torsional vibrations on tapping performance[J]. Applied Mechanics and Materials, 2016, 826: 88-92.
[20] 彭太江,杨树臣,杨志刚,等.超声波的减摩特性[J].吉林大学学报(工学版),2006,36(增刊2):88-90.
 PENG T J, YANG S C, YANG Z G, et al. Friction-reducing charac-

teristics of ultrasonic wave[J]. Journal of Jilin University (Engineering and Technology Edition), 2006, 36(Sup2): 88-90(in Chinese).

[21] 吴博达,常颖,杨志刚,等.超声振动减摩性能的实验研究及理论 分析[J]. 中国机械工程, 2004, 15(9): 813-815.

WU B D, CHANG Y, YANG Z G, et al. Experimental study and

theoretical analysis on anti-friction capability of ultrasonic vibration[J]. China Mechanical Engineering, 2004, 15(9): 813-815(in Chinese).

[22] 常颖,彭太江,阚君武,等. 超声振动对摩擦系数影响的试验研究
 [J]. 压电与声光, 2003, 25(6): 511-513.
 CHANG Y, PENG T J, KAN J W, et al. Experiment study on the influence on friction-factor by ultrasonic vibration[J]. Piezoelectrics & Acoustooptics, 2003, 25(6): 511-513(in Chinese).

Ultrasonic longitudinal torsional and low frequency torsional compound vibration tapping experiment

HAN Fengqi¹, ZHANG Deyuan^{2,*}

Intelligent Manufacturing Institute, Hebei Vocational University of Industry and Technology, Shijiazhuang 050091, China;
 School of Mechanical Engineering and Automation, Beihang University, Beijing 100191, China)

Abstract: In view of the low efficiency of low frequency torsional vibration tapping of titanium alloy, the ultrasonic longitudinal torsional and low frequency torsional compound vibration tapping method was proposed. To investigate the process effect of compound vibration tapping and the influence rule of parameter change on it, three experiments were conducted in succession: the ultrasonic longitudinal torsional vibration tapping experiment, the compound vibration tapping experiment with different cutting fluids, and the compound vibration tapping experiment with variable parameters. The experimental results show that ultrasonic longitudinal torsional vibration enhances the lubricating effect of the cutting fluid and reduces the tapping torque. Compared with ordinary continuous tapping, the ultrasonic longitudinal torsional vibration tapping to low frequency torsional vibration tapping, the tapping torque by about 18%. Comparing compound vibration tapping to low frequency torsional vibration tapping, the tapping torque decreases with the chosen parameter range. The effect of compound vibration tapping on reducing the tapping torque decreases with the increase of the net cutting amount, the spindle speed, and the decrease of the backward amount.

Keywords: vibration tapping; ultrasonic longitudinal torsional vibration; low frequency torsional vibration; tapping torque; cutting fluid lubrication

Received: 2022-06-02; Accepted: 2022-07-09; Published Online: 2022-08-11 13:28 URL: link.cnki.net/urlid/11.2625.V.20220810.0922.003

Foundation items: Central Funding Project for Local Science and Technology Development (216Z1004G); Funded by Science and Technology Project of Hebei Education Department (ZD2020331); Doctoral Fund of Hebei Vocational University of Industry and Technology (zk202101)

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn DOI: 10.13700/j.bh.1001-5965.2022.0495

引入 ESO 的永磁超环面电机无模型预测电流控制

刘欣^{1,2,*},王正阳¹,王晓远²

(1. 天津工业大学 天津市现代机电装备技术重点实验室, 天津 300387; 2. 天津大学 电气自动化与信息工程学院, 天津 300072)

摘 要:为改善无差拍预测电流控制 (DPCC) 对永磁超环面电机系统参数的依赖性,研究 了引入扩张状态观测器 (ESO) 的永磁超环面电机无模型预测电流控制 (MFPCC-ESO) 策略。根据永 磁超环面电机的复合转子结构,引入自转运动影响系数与磁势系数,在旋转坐标系下建立该电机的 时变数学模型;利用永磁超环面电机系统的输入与输出,建立该电机具有时变比例因子的超局部模 型,同时引入 ESO 对超局部模型的干扰部分进行实时估计,并利用朱利稳定判据证明了 ESO 的稳 定性;结合延时补偿的 DPCC 预测得到参考电压矢量,从而实现永磁超环面电机的 MFPCC-ESO 策 略。对参数匹配和失配下永磁超环面电机 MFPCC-ESO 策略与 DPCC 策略进行对比分析,仿真结果 表明: MFPCC-ESO 策略下的永磁超环面电机具有优越的动态和稳态性能及强鲁棒性,同时该控制 策略还能有效降低永磁超环面电机的输出波动。

关 键 词:超环面电机;超局部模型;无模型;预测电流控制;扩张状态观测器

中图分类号: TM351

文献标志码: A 文章编号: 1001-5965(2024)04-1085-12

永磁电机以其体积小、结构简单、功率密度高 及可靠性高等优点被广泛应用于电动汽车、航空航 天等高性能的电机驱动领域^[1-4]。为进一步扩展永 磁电机的应用场合,国内外学者正致力于具有空间 结构的永磁电机的研究与开发。文献 [5] 提出了一 种考虑空间设计的横向磁通永磁电机,并对该电机 的结构参数与结构强度进行了研究。文献 [6] 开发 了一种永磁球型电机,同时提供了该空间电机的磁 场建模方法。永磁超环面电机是一种新型的永磁 空间电机,其将永磁电机与行星蜗杆驱动技术进行 有机结合,具有结构紧凑、效率高、输出转矩大及 承载能力强等优点[7-8]。永磁超环面电机以其独特 的空间结构与功能集成特点,在航空航天、军事等 领域有着广阔的应用前景⁹⁹。目前,主要针对永磁 超环面电机的结构优化及电磁特性进行研究[10-12]. 为实现该电机的高性能输出,有必要进一步对其控

制策略进行研究。

近年来,随着微处理器的快速发展,模型预测 电流控制 (model predictive current control, MPCC) 以其原理简单、动态响应快、非线性约束处理方便 的优点成为了典型的最优控制算法,在永磁电机驱 动领域中发挥着重要作用^[13-14]。文献 [15] 提出了一 种适用于永磁同步电机的 MPCC 策略,通过设计新 型比例积分形式的成本函数,获得了快速的瞬态控 制性能。文献 [16] 提出了一种改进的永磁同步电 机 MPCC 方案,基于电流轨迹圆选择最优电压矢 量,有效降低电流脉动与转矩脉动。然而, MPCC 算法的优越性能建立在确定的电机参数上,如电 感、电阻和磁链,实际工程中,多变的工况会导致参 数失配,使 MPCC 算法产生预测误差,进而削弱控 制性能^[17-18]。近年来,为降低参数失配对预测控制 带来的影响,无模型预测电流控制 (model-free

*通信作者. E-mail: liuxin@tiangong.edu.cn

收稿日期: 2022-06-16;录用日期: 2022-07-25;网络出版时间: 2022-08-11 15:26 网络出版地址: link.cnki.net/urlid/11.2625.V.20220810.0919.002

基金项目:国家自然科学基金(51875408);天津市研究生科研创新项目(2021YJSB229)

引用格式: 刘欣, 王正阳, 王晓远, 引入 ESO 的永磁超环面电机无模型预测电流控制 [J]. 北京航空航天大学学报, 2024, 50 (4): 1085-1096. LIU X, WANG Z Y, WANG X Y. Model-free predictive current control for permanent magnet toroidal motor with extended state observer[J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50 (4): 1085-1096 (in Chinese).

predictive current control, MFPCC)成为了研究热 点。文献 [19]针对永磁同步电机提出一种连续电 压矢量 MFPCC 方案,根据电压矢量的解耦特性对 电压矢量进行优化,提高了参数失配时的稳态控制 精度。文献 [20]提出了永磁同步电机的改进 MFPCC 策略,通过设计先进的电流梯度更新机制, 避免参数失配带来的影响,并提高了电机的稳态性 能。文献 [21]提出了基于三电平逆变器供电的永 磁电机 MFPCC 方案,采用电流差分更新技术与候 选向量选择算法,增强了参数鲁棒性,并降低了电 流谐波含量。文献 [22]针对永磁同步电机提出有 限集容错 MFPCC 方案,通过将无模型容错预测控 制器与滑模观测器进行结合,在失磁故障和参数摄 动下获得了优越的暂稳态性能及强鲁棒性。

本文基于永磁超环面电机的特殊空间结构,分 析其时变电感参数,建立时变非线性数学模型。为 提高预测电流控制对永磁超环面电机参数变化的 鲁棒性,提出永磁超环面电机带扩张状态观测器的 无模型预测电流控制(model-free predictive current control with extended state observer, MFPCC-ESO)策 略。该策略在时变数学模型的基础上,将与永磁超 环面电机参数相关的部分视为系统的干扰量,建立 超局部模型。同时,设计电流扩张状态观测器对超 局部模型的干扰部分进行在线估计,降低该模型对 永磁超环面电机参数依赖的同时,保证超局部模型 与永磁超环面电机实际模型的一致性。结合无差 拍预测电流控制(deadbeat predictive current control, DPCC)预测得到未来时刻的参考电压矢量,通过空 间矢量脉宽调制以控制逆变器运行。通过对永磁 超环面电机 MFPCC-ESO 策略的仿真分析, 验证该 策略的可行性和有效性。

1 永磁超环面电机结构原理

永磁超环面电机的结构简图如图1所示。永 磁超环面电机主要由蜗杆内定子、行星轮、行星架 转子、环面外定子组成。蜗杆内定子由不同尺寸的 硅钢片叠压而成,其表面均匀分布螺旋状的电枢 槽,电枢槽内放置电枢绕组。环面外定子由多个 NS极相间分布的空间螺旋永磁梁组成。行星轮与 行星架转子固联,行星轮的圆周上均匀分布永磁 齿,并与蜗杆内定子间形成电磁内啮合,与环面外 定子间形成电磁外啮合。

当蜗杆内定子电枢绕组通入三相交流电时,其 表面会产生变截面空间螺旋的旋转磁场。行星轮 受到蜗杆内定子空间磁场力的驱动,产生自转的同 时公转。行星轮的公转会驱动行星架转子转动,实 现永磁超环面电机的动力输出。同时,在行星轮的



图 1 永磁超环面电机的结构简图

Fig. 1 Structural diagram of permanent magnet toroidal motor

运转过程中,其电磁外啮合产生的磁场力有利于保 证永磁超环面电机的平稳运行。

2 永磁超环面电机数学模型

2.1 电感参数分析

由于永磁超环面电机结构的特殊性,行星轮在 空间磁场力驱动下公转与自转。行星轮的复合运 动对永磁超环面电机磁路中的磁阻有重要影响,从 而影响磁链对应的电感参数。图2为行星轮磁齿 的运动轨迹。可以看出,只考虑行星轮公转时,行 星轮磁齿在空间的运动轨迹为圆形,结合行星轮自 转,行星轮磁齿在空间的运动轨迹为环形螺旋。考 虑行星轮公转与自转运动的关系,引入自转运动的 影响因素,可得到蜗杆内定子绕组的电感。



图 2 行星轮磁齿的运动轨迹

Fig. 2 Motion track of magnetic teeth on planet

根据转化机构法,对永磁超环面电机施加一个 反方向的行星架转子机械角速度 ω_{lm},得到行星轮 自转角速度与环面外定子旋转角速度间的比值为

$$i_{23}^{1} = \frac{\omega_{2m} - \omega_{1m}}{\omega_{3m} - \omega_{1m}} = 1 - \frac{\omega_{2m}}{\omega_{1m}} = -\frac{z_{3}}{z_{2}}$$
(1)

式中: ω_{2m} 和 ω_{3m} 分别为行星轮自转角速度和环面 外定子旋转角速度; z_2 和 z_3 分别为行星轮和环面外 定子的磁齿数。

由式(1)得到行星轮自转运动的影响系数为

$$k_2 = \frac{\theta_2}{\theta_1} = \frac{(z_2/2)\omega_{2m}}{(m/2)\omega_{1m}} = \frac{z_2 + z_3}{m}$$
(2)

式中: θ₁ 和 θ₂ 分别为行星架转子输出电角度和行星 轮自转电角度; m 为行星轮个数。

蜗杆内定子具有特殊的变截面结构,与其结构参数相比,行星轮磁齿位置角度的变化对蜗杆内定子绕组电感的平均值与二次谐波幅值的影响可以被忽略。引入自转运动的影响系数,可以得到A相蜗杆内定子绕组的自感及AB相蜗杆内定子绕组间互感分别为

$$L_{AA(\theta_1)} = L'_{s0} - L'_{s2}\cos(2\theta_1) - L''_{s2}\cos(2k_2\theta_1)$$
(3)
$$M_{AB(\theta_1)} = -\frac{1}{2}L'_{s0} - L'_{s2}\cos(2(\theta_1 + \pi/6)) -$$

$$L_{s2}' \cos(2(k_2\theta_1 + \pi/6)) \tag{4}$$

式中: L'so为蜗杆内定子绕组自感平均值; L's2和L's2分 别为与公转和自转有关的蜗杆内定子绕组自感二 次谐波幅值。

为分析行星轮和蜗杆内定子绕组间的互感,将 行星轮上永磁齿等效为励磁绕组。基于等效电流 法,向励磁绕组通入等效励磁电流,从而产生磁动 势。等效电流 $I_{\rm f}$ 在公转和自转方向的分量分别表 示为 $I_{\rm fl}$ 和 $I_{\rm fl}$,如图 3 所示。公转平面磁动势 $F_{\rm fl}$ 与 自转平面磁动势 $F_{\rm fl}$ 分别由蜗杆包角 $\varphi_{\rm v}$ 内的当量永 磁齿和单个永磁齿产生,由此引入磁势系数 $k_{\rm l}$ 表示 两平面内磁动势的关系为: $k_{\rm l}F_{\rm fl}=F_{\rm fl}$,其中, $k_{\rm l}$ 为包 角内当量磁齿数的倒数, $k_{\rm l}=2\pi/(z_2\varphi_{\rm v})$ 。通过引入磁 势系数,得到行星轮永磁齿和 A 相蜗杆内定子绕组 之间的互感为

 $M_{fA(\theta_1)} = L_{AAd}(k_1 \cos(k_2 \theta_1 - 2\pi/3) + \cos(\theta_1 - 2\pi/3)) \quad (5)$

式中: L_{AAd}为 A 相蜗杆内定子绕组轴线与直轴重合时磁链对应的电感。通过相位关系可获得行星轮 永磁齿与其他两相绕组间的互感。



图 3 等效电流示意图 Fig. 3 Schematic diagram of equivalent current

2.2 数学模型

基于蜗杆内定子绕组的电感方程,通过坐标变换,得到永磁超环面电机的直轴与交轴电感为

$$\begin{cases} L_{d(\theta_1)} = \frac{3}{2}L'_{s0} + \frac{1}{2}L'_{s2} + \frac{1}{2}L''_{s2}\cos(2c\theta_1) \\ L_{q(\theta_1)} = \frac{3}{2}L'_{s0} - \frac{1}{2}L'_{s2} - \frac{1}{2}L''_{s2}\cos(2c\theta_1) \end{cases}$$
(6)

式中: $L_{d(\theta_1)}$ 和 $L_{q(\theta_1)}$ 分别为永磁超环面电机的直轴和 交轴电感; $c=k_2-1$ 。

通过坐标变换,得到永磁超环面电机在旋转坐 标系下的非线性数学模型,其磁链方程为

$$\begin{cases} \psi_{d(\theta_1)} = L_{d(\theta_1)} i_d + \psi_f(k_1 \cos(c\theta_1) + 1) \\ \psi_{q(\theta_1)} = L_{q(\theta_1)} i_q + \psi_f k_1 \sin(c\theta_1) \end{cases}$$
(7)

式中: $\psi_{d(\theta_1)}$ 和 $\psi_{q(\theta_1)}$ 分别为直轴和交轴磁链; i_d 和 i_q 分 别为直轴和交轴电流; ψ_f 为永磁体磁链, $\psi_f = L_{AAd}I_f$ 。

为简化计算,忽略对电压影响小的量,得到永 磁超环面电机在旋转坐标系下的电压方程为

$$\begin{cases} u_d = Ri_d - \psi_f \omega_1 k_s(\theta_1) + L_{d(\theta_1)} \frac{\mathrm{d}i_d}{\mathrm{d}t} - \omega_1 L_{q(\theta_1)} i_q \\ u_q = Ri_q + \psi_f \omega_1 (1 + k_c(\theta_1)) + L_{q(\theta_1)} \frac{\mathrm{d}i_q}{\mathrm{d}t} + \omega_1 L_{d(\theta_1)} i_d \end{cases}$$

$$(8)$$

式中: $k_s(\theta_1)=k_1k_2\sin(c\theta_1)$; $k_c(\theta_1)=k_1k_2\cos(c\theta_1)$; u_d 和 u_q 分 别为直轴和交轴电压; ω_1 为行星架转子的输出电角 速度; R为绕组的相电阻。

对永磁超环面电机采用 *i*_d=0 的控制策略,旋转 坐标系下的转矩方程与运动方程可分别表示为

$$T_{e(\theta_1)} = \frac{3}{2} n_{p1} \psi_f i_q (1 + k_e(\theta_1))$$
(9)

$$J\frac{\mathrm{d}\omega_{\mathrm{lm}}}{\mathrm{d}t} = T_{\mathrm{e}(\theta_{\mathrm{l}})} - T_{\mathrm{L}} - B\omega_{\mathrm{lm}} \tag{10}$$

式中: $T_{e(\theta_1)}$ 为旋转坐标系下的转矩; n_{p1} 为行星轮对数; T_L 为负载转矩; J和 B分别为转动惯量和摩擦系数。

3 无差拍模型预测电流控制

通过前向欧拉近似法对式(8)进行离散化处理,可得到永磁超环面电机在k时刻的直交轴电压为

$$\begin{cases} u_{d}^{k} = Ri_{d}^{k} - \psi_{f}\omega_{1}^{k}k_{s}(\theta_{1}) + L_{d(\theta_{1})}^{k}\frac{i_{d}^{k+1} - i_{d}^{k}}{T_{s}} - \omega_{1}^{k}L_{q(\theta_{1})}^{k}i_{q}^{k} \\ u_{q}^{k} = Ri_{q}^{k} + \psi_{f}\omega_{1}^{k}(1 + k_{c}(\theta_{1})) + L_{q(\theta_{1})}^{k}\frac{i_{q}^{k+1} - i_{q}^{k}}{T_{s}} + \omega_{1}^{k}L_{d(\theta_{1})}^{k}i_{d}^{k} \end{cases}$$

$$(11)$$

式中: $i_{d,q}^{k}$ 和 $i_{d,q}^{k+1}$ 分别为k时刻和k+1时刻的直交轴 蜗杆内定子电流分量; $u_{d,q}^{k}$ 和 $L_{d,q(\theta_{1})}^{k}$ 分别为k时刻的 直交轴电压分量和电感分量; T_{s} 为采样周期。

基于无差拍 MPCC 原理, 令 k+1 时刻的直交轴 蜗杆内定子电流与直交轴参考电流i^{*}_{d,q}相等, 从而实 现蜗杆内定子电流的无差拍跟踪。此时, 通过式 (11) 得到的直交轴电压为 k 时刻的直交轴参考电压。

在无差拍 MPCC 策略中,输出电压矢量与调节 电流之间存在控制延时,电流控制效果会被削弱。 为补偿延时问题,令 k+2 时刻的蜗杆内定子电流与

参考值相等, 计算 k+1 时刻的直交轴参考电压为

$$\begin{cases}
u_d^* = Ri_d^{k+1} - \psi_f \omega_1^{k+1} k_s(\theta_1) + L_{d(\theta_1)}^{k+1} \frac{i_d^* - i_d^{k+1}}{T_s} - \\
\omega_1^{k+1} L_{q(\theta_1)}^{k+1} i_q^{k+1} \\
u_q^* = Ri_q^{k+1} + \psi_f \omega_1^{k+1} (1 + k_c(\theta_1)) + L_{q(\theta_1)}^{k+1} \frac{i_q^* - i_q^{k+1}}{T_s} + \\
\omega_1^{k+1} L_{d(\theta_1)}^{k+1} i_d^{k+1}
\end{cases}$$
(12)

式中: u_a^t 、 u_q^* 分别为第 k+1 周期内的直、交轴参考电 压; $L_{d,q(\theta_1)}^{k+1}$ 时刻的直交轴电感分量。由于电磁 时间常数远小于机械时间常数,且采样周期足够 短,可认为 $\omega_1^k \approx \omega_1^{k+1}$ 。然而在 k 时刻 $i_{d,q}^{k+1}$ 是未知量, 由式 (11) 可得到 k+1 时刻的蜗杆内定子电流预测 值为

$$\begin{cases} i_{d}^{k+1} = i_{d}^{k} + \frac{T_{s}}{L_{d(\theta_{1})}^{k}} (u_{d}^{k} - Ri_{d}^{k} + \psi_{f} \omega_{1}^{k} k_{s}(\theta_{1}) + \omega_{1}^{k} L_{q(\theta_{1})}^{k} i_{q}^{k}) \\ i_{q}^{k+1} = i_{q}^{k} + \frac{T_{s}}{L_{q(\theta_{1})}^{k}} (u_{q}^{k} - Ri_{q}^{k} - \psi_{f} \omega_{1}^{k} (1 + k_{c}(\theta_{1})) - \omega_{1}^{k} L_{d(\theta_{1})}^{k} i_{d}^{k}) \end{cases}$$

$$(13)$$

在永磁超环面电机的运行过程中,为避免直交 轴参考电压幅值超过逆变器直流母线电压最大值 $U_{dc}/\sqrt{3}$,对直交轴参考电压进行如下限幅处理:

$$u_{d,q}^{**} = \begin{cases} \frac{U_{dc}}{\sqrt{3}u_{s}^{*}}u_{d,q}^{*} & u_{s}^{*} > U_{dc}/\sqrt{3} \\ u_{d,q}^{*} & u_{s}^{*} \le U_{dc}/\sqrt{3} \end{cases}$$
(14)

式中: $u_{d,q}^{**}$ 为限幅后的直交轴参考电压; U_{dc} 为直流 母线电压; u_s^* 为参考电压, $u_s^* = \sqrt{(u_d^*)^2 + (u_q^*)^2} \circ u_{d,q}^{**}$ 通过反 Park 变换与空间矢量脉宽调制(space vector pulse width modulation, SVPWM)环节, 生成脉冲信 号作用于逆变器, 进而控制永磁超环面电机。

4 引入 ESO 的无模型预测电流控制

永磁超环面电机 DPCC 策略的实现需要获取 准确的电机参数,包括磁链、电感和电阻。为降低 预测电流控制对永磁超环面电机参数的依赖,设计 引入 ESO 的永磁超环面电机 MFPCC 策略。

4.1 超局部模型建立

根据无模型控制原理,在永磁超环面电机电压 方程的基础上,以直交轴电压与直交轴电流分别作 为永磁超环面电机系统的输入变量与输出变量,其 他与电机参数相关的部分视为系统的干扰量,建立 永磁超环面电机的一阶超局部模型为

$$\begin{bmatrix} \frac{di_d}{dt} \\ \frac{di_q}{dt} \end{bmatrix} = \begin{bmatrix} \alpha_{d(\theta_1)} & 0 \\ 0 & \alpha_{q(\theta_1)} \end{bmatrix} \begin{bmatrix} u_d \\ u_q \end{bmatrix} + \begin{bmatrix} F_d \\ F_q \end{bmatrix}$$
(15)

式中: $\alpha_{d(\theta_1)}$ 、 $\alpha_{q(\theta_1)}$ 为设计的永磁超环面电机系统的

比例因子, $\alpha_{d(\theta_1)} = 1/L_{d(\theta_1)}$, $\alpha_{q(\theta_1)} = 1/L_{q(\theta_1)}$; F_d 、 F_q 分别 为永磁超环面电机系统的直、交轴干扰量, 表示为

$$\boldsymbol{F}_{dq} = \boldsymbol{A}\boldsymbol{i}_{dq} + (\boldsymbol{B} - \boldsymbol{B}_1)\boldsymbol{u}_{dq} + \boldsymbol{C}$$
(16)

$$\vec{\mathbf{x}} \div : \boldsymbol{B} = \begin{bmatrix} \frac{1}{L_{d(\theta_1)}} & 0\\ 0 & \frac{1}{L_{q(\theta_1)}} \end{bmatrix}; \boldsymbol{B}_1 = \begin{bmatrix} \frac{1}{L'_{d(\theta_1)}} & 0\\ 0 & \frac{1}{L'_{q(\theta_1)}} \end{bmatrix};$$
$$\boldsymbol{A} = \begin{bmatrix} \frac{-R}{L_{d(\theta_1)}} & \frac{\omega_1 L_{q(\theta_1)}}{L_{d(\theta_1)}} \\ \frac{-\omega_1 L_{d(\theta_1)}}{L_{q(\theta_1)}} & \frac{-R}{L_{q(\theta_1)}} \end{bmatrix}; \quad \boldsymbol{C} = \begin{bmatrix} \frac{\psi_1 \omega_1 k_s(\theta_1)}{L_{d(\theta_1)}} \\ \frac{-\psi_1 \omega_1(1+k_c(\theta_1))}{L'_{q(\theta_1)}} \end{bmatrix};$$
$$\boldsymbol{i}_{dq} = \begin{bmatrix} i_d \\ i_q \end{bmatrix}; \quad \boldsymbol{u}_{dq} = \begin{bmatrix} u_d \\ u_q \end{bmatrix}; \quad \boldsymbol{F}_{dq} = \begin{bmatrix} F_d \\ F_q \end{bmatrix}; \quad \boldsymbol{L}'_{d(\theta_1)} \cdot L'_{q(\theta_1)} \not\supset \dot{\mathcal{I}} \dot{\mathcal{K}}$$

磁超环面电机的实际电感值,当实际电感值为标称 电感值时,输出电压部分的干扰为零。

4.2 扩张状态观测器设计

由式(16)可知,参数变化会直接影响超局部模型的干扰部分。为降低参数变化对超局部模型的影响,设计超环面 ESO 估计直交轴干扰量。基于永磁超环面电机的超局部模型,建立永磁超环面电机 ESO 方程为

$$\begin{cases} \boldsymbol{e}_{\mathrm{rr}} = \hat{\boldsymbol{i}}_{dq} - \boldsymbol{i}_{dq} \\ \dot{\boldsymbol{i}}_{dq} = \boldsymbol{\alpha}_{dq(\theta_1)} \boldsymbol{u}_{dq} + \hat{\boldsymbol{F}}_{dq} - \beta_1 \boldsymbol{I} \boldsymbol{e}_{\mathrm{rr}} \\ \dot{\boldsymbol{F}}_{dq} = -\beta_2 \boldsymbol{I} \boldsymbol{e}_{\mathrm{rr}} \end{cases}$$
(17)

式中:
$$\boldsymbol{\alpha}_{dq(\theta_1)} = \begin{bmatrix} \alpha_{d(\theta_1)} & 0\\ 0 & \alpha_{q(\theta_1)} \end{bmatrix}$$
; $\hat{\boldsymbol{i}}_{dq} = \begin{bmatrix} \dot{\boldsymbol{i}}_d\\ \hat{\boldsymbol{i}}_q \end{bmatrix}$, $\hat{\boldsymbol{i}}_d \setminus \hat{\boldsymbol{i}}_q$ 分別

为直、交轴蜗杆内定子电流估计值; $\hat{F}_{dq} = \begin{bmatrix} \hat{F}_d \\ \hat{F}_q \end{bmatrix}$,

 \hat{F}_{a} 、 \hat{F}_{q} 分别为直、交轴干扰估计值; I为单位矩阵; β_{1} 、 β_{2} 为 ESO 的误差反馈增益值。

将式(17)进行拉普拉斯变换可得

$$\begin{cases} \hat{i}_{d}(s) = \frac{\alpha_{d(\theta_{1})}u_{d}s + (\beta_{2} + \beta_{1}s)i_{d}}{s^{2} + \beta_{1}s + \beta_{2}} \\ \hat{i}_{q}(s) = \frac{\alpha_{q(\theta_{1})}u_{q}s + (\beta_{2} + \beta_{1}s)i_{q}}{s^{2} + \beta_{1}s + \beta_{2}} \end{cases}$$
(18)

由式(18)可知,永磁超环面电机 ESO 的特征方 程为

$$s^2 + \beta_1 s + \beta_2 = 0 \tag{19}$$

为使永磁超环面电机 ESO 稳定, 令 ESO 的期望 极点为- ω_0 ,则期望特征方程为[$s - (-\omega_0)$][$s - (-\omega_0)$] = 0,联立式 (19),可得 ESO 的误差反馈增益值为

$$\begin{cases} \beta_1 = 2\omega_0 \\ \beta_2 = \omega_0^2 \end{cases}$$
 (20)

式中: ω_0 为永磁超环面电机 ESO 的带宽, 合理设计 带宽能够使永磁超环面电机 ESO 具有良好的观测 性能。

采用前向欧拉近似法,对式(17)进行离散化处 理可得

$$\begin{cases} \boldsymbol{e}_{\mathrm{rr}}^{k} = \hat{\boldsymbol{i}}_{dq}^{k} - \boldsymbol{i}_{dq}^{k} \\ \hat{\boldsymbol{i}}_{dq}^{k+1} = \hat{\boldsymbol{i}}_{dq}^{k} + (\boldsymbol{\alpha}_{dq(\theta_{1})}^{k} \boldsymbol{u}_{dq}^{k} + \hat{\boldsymbol{F}}_{dq}^{k} - \beta_{1} \boldsymbol{I} \boldsymbol{e}_{\mathrm{rr}}^{k}) \boldsymbol{T}_{\mathrm{s}} \\ \hat{\boldsymbol{F}}_{dq}^{k+1} = \hat{\boldsymbol{F}}_{dq}^{k} - \beta_{2} \boldsymbol{T}_{\mathrm{s}} \boldsymbol{I} \boldsymbol{e}_{\mathrm{rr}}^{k} \end{cases}$$
(21)

由式 (21) 可得, 永磁超环面电机 ESO 的传递函 数为

$$\begin{cases} G_1(z) = \frac{\hat{i}_d(z)}{i_d(z)} = \frac{\beta_2 T_s^2 + \beta_1 T_s(z-1)}{(z-1)^2 + \beta_1 T_s(z-1) + \beta_2 T_s^2} \\ G_2(z) = \frac{\hat{i}_q(z)}{i_q(z)} = \frac{\beta_2 T_s^2 + \beta_1 T_s(z-1)}{(z-1)^2 + \beta_1 T_s(z-1) + \beta_2 T_s^2} \end{cases}$$
(22)

根据式 (22) 可得离散化的 ESO 特征方程为

$$(z-1)^2 + \beta_1 T_s(z-1) + \beta_2 T_s^2 = 0$$
 (23)

根据朱利稳定判据,得到离散化的 ESO 稳定条 件如下:

$$\begin{cases} \left| 1 - \beta_1 T_s + \beta_2 T_s^2 \right| < 1 \\ \left| 2 - \beta_1 T_s \right| < 2 - \beta_1 T_s + \beta_2 T_s^2 \end{cases}$$
(24)

联立式 (20) 与式 (24), 得到 ESO 稳定条件为

 $0 < \omega_0 < 2/T_s$ (25)

当ω₀在式 (25) 的取值范围内,能够保证 ESO

特征方程的根位于 z 平面的单位圆中, 从而保证离 散化的永磁超环面电机 ESO 的稳定性。

4.3 超局部无模型电流预测

采用前向欧拉近似法,对永磁超环面电机的一 阶超局部模型进行离散化处理可得

$$\boldsymbol{i}_{dq}^{k+1} = \boldsymbol{i}_{dq}^{k} + (\boldsymbol{a}_{dq(\theta_{1})}^{k}\boldsymbol{u}_{dq}^{k} + \boldsymbol{F}_{dq}^{k})T_{s}$$
(26)

式中: F_{dq}^{k} 由 \hat{F}_{dq}^{k} 代替, \hat{F}_{dq}^{k} 通过永磁超环面电机 ESO观测得到。

采用 DPCC 策略,并考虑延时补偿,使 k+2 时刻 的直交轴蜗杆内定子电流与直交轴参考电流相等, 可得 k+1 时刻的直交轴参考电压为

$$\boldsymbol{u}_{dq}^{\text{ref}} = \frac{\boldsymbol{i}_{dq}^{*} - \boldsymbol{i}_{dq}^{k+1}}{\boldsymbol{\alpha}_{dq(\theta_{1})}^{k+1} T_{\text{s}}} - \frac{\boldsymbol{\hat{F}}_{dq}^{k+1}}{\boldsymbol{\alpha}_{dq(\theta_{1})}^{k+1}}$$
(27)

式中: $i_{dq}^* = \begin{bmatrix} i_d^* \\ i_q^* \end{bmatrix}$; $u_{dq}^{\text{ref}} = \begin{bmatrix} u_d^{\text{ref}} \\ u_q^{\text{ref}} \end{bmatrix}$, $u_d^{\text{ref}} \pi u_q^{\text{ref}} \beta$ 别为 k+1时刻的直轴和交轴参考电压。

与 DPCC 策略类似, MFPCC-ESO 策略的实现 需对输出参考电压进行限幅处理,反 Park 变换与 SVPWM 调制,通过逆变器控制永磁超环面电机运 行。永磁超环面电机 MFPCC-ESO 系统结构框图如 图 4 所示,该控制系统框图内的逆变器、SVPWM 模块等采用 MATLAB/Simulink 数据库中的模块。



Fig. 4 Block diagram of MFPCC-ESO system for permanent magnet toroidal motor

5 仿真分析

为验证本文提出的 MFPCC-ESO 策略的有效 性,对永磁超环面电机 DPCC 策略和 MFPCC-ESO 策略进行对比研究。2 种控制策略的采样周期设置 为 T_s =5×10⁻⁵ s;转速环 PI 参数设置相同;逆变器的 直流母线电压设置为 U_{dc} =311 V。ESO 带宽设置为 ω_0 =4 200。用于仿真分析的永磁超环面电机参数如 表 1 所示。表中: p_n 为额定功率, U_N 为额定电压, f_N 为额定频率。

Table 1	Permanent magnet	toroidal motor parameter
	4 ¥4	粉店

圭 1

参数	数值
R/Ω	3.8
$n_{\rm p1}$	4
Z_3	28
Z_2	12
k_1	0.3
k_2	5
$p_{ m n}/{ m kW}$	1.5
$U_{ m dc}/{ m V}$	311
$B/(N \cdot m \cdot s)$	0.001
$J(\text{kg}\cdot\text{m}^2)$	0.001 5
$L_{ m s0}^\prime/{ m H}$	0.001 2
L'_{s2}/H	0.000 4
$L_{\rm s2}^{\prime\prime}/{ m H}$	0.003
$\psi_{\rm f}/{ m Wb}$	0.26
$U_{ m N}/{ m V}$	250
$f_{ m N}/{ m Hz}$	100

5.1 参数匹配时动稳态性能分析

图 5 为 DPCC 策略与 MFPCC-ESO 策略下永磁 超环面电机的三相蜗杆内定子电流波形与 A 相蜗 杆内定子电流频谱,给定转速1000 r/min,负载为 4 N·m。可知, DPCC 策略与 MFPCC-ESO 策略下的 蜗杆内定子电流波形存在周期性的小波动,这是由 于永磁超环面电机的时变参数所引起的,同时每相 蜗杆内定子电流的波形存在差异,与永磁超环面电 机蜗杆内定子的特殊变截面结构及其上电枢绕组 的布线方式有关。与 DPCC 策略相比, MFPCC-ESO 策略下的蜗杆内定子电流波动相对较小,电流波形 更平滑。同时, DPCC 策略与 MFPCC-ESO 策略下 的A相蜗杆内定子电流总谐波畸变率 THD 分别为 18.32% 与 13.54%, MFPCC-ESO 策略下的蜗杆内定 子电流 THD 值相对降低 26.09%。因此,本文提出 的 MFPCC-ESO 策略能有效降低蜗杆内定子电流波 动,较好地抑制电流谐波。

图 6 为稳态条件下采用 2 种策略的永磁超环面 电机的转速、转矩和直交轴电流波形。图中: n 为 转速。可以看出, DPCC 策略与 MFPCC-ESO 策略 下的输出转速、转矩和直交轴电流在稳态阶段均存 在周期性的波动。同时, 2 种控制策略下的直轴电 流波动顺序存在差异, 这是由于在"*i*_a=0"的控制影 响下直轴电流的上下波动顺序是瞬时的调整过 程。与 DPCC 策略相比, 采用 MFPCC-ESO 策略的 永磁超环面电机输出波动有所降低, 其转速、转 矩、交轴、直轴电流波动率分别降低 22.2%、23.3%、 28.3%、15.6%。因此, 进一步说明了 MFPCC-ESO



图 5 三相蜗杆内定子电流波形与频谱

Fig. 5 Three-phase worm inner stator current waveforms and harmonic spectra



Fig. 6 Output waveforms at steady-state conditions

策略能有效降低永磁超环面电机的输出波动,提高 电机的稳态性能。 图 7为采用 DPCC 策略与 MFPCC-ESO 策略 的永磁超环面电机的转速阶跃响应。永磁超环面 电机空载运行,在 0.12 s 时转速从 1000 r/min 升至 1200 r/min。仿真结果表明,2 种控制策略下转速、 转矩和直交轴电流通过短暂的调整阶段能够达到 稳态。但与 DPCC 策略相比,MFPCC-ESO 策略下 的转速、转矩和直交轴电流具有较短的动态调整时 间,转速与转矩超调量较小,仅交轴电流的超调量 稍大,因此,在永磁超环面电机的变速运行阶段, MFPCC-ESO 策略具有比 DPCC 策略优越的动态 性能。

图 8 为采用 DPCC 策略与 MFPCC-ESO 策略的 永磁超环面电机的负载阶跃响应。永磁超环面电 机转速恒定为1000 r/min,在0.12 s 时负载从0 N·m 突变至6 N·m。通过图 8 可以看出,在突加负载后 的瞬态阶段,2 种控制策略下转矩与直轴电流变化 幅度相似,转速与交轴电流变化幅度相同且较小, 抗干扰能力强。MFPCC-ESO 策略下,转速、转矩、 交轴电流的调整时间为 50 ms,直轴电流经过 8 ms





达到稳态,与 DPCC 策略相比,相应调整时间分别减少了 3.8% 和 46.7%。因此, MFPCC-ESO 策略具有比 DPCC 策略更快的动态响应。

为验证永磁超环面电机 ESO 的性能,负载阶跃 下 ESO 对干扰量的估计结果与电流估计误差结果 如图 9 所示。可以看出,估计的直交轴干扰量存在 与时变参数密切相关的周期性波动,且能在负载变 化瞬间产生相应变化,为 MFPCC-ESO 策略提供了 预测干扰值。同时,负载阶跃变化对电流估计误差 影响较小,交轴与直轴电流的估计误差在瞬态与稳 态阶段始终维持在零附近。因此,永磁超环面电机 ESO 具有良好的观测性能。

5.2 参数失配时动稳态性能分析

为验证本文所提 MFPCC-ESO 策略的优越性, 进一步研究控制器参数与永磁超环面电机参数不 匹配时采用 2 种控制策略时永磁超环面电机的输 出性能。

直轴电感参数失配前后采用 DPCC 策略与 MFPCC-ESO 策略时永磁超环面电机的稳态波形如 图 10 和图 11 所示。此时,给定转速 1 000 r/min,负载为 4 N·m,图 10 中,0.125 s 时控制器中电感参数 改变为其标称值的 1.5 倍,图 11 中,0.125 s 时控制 器中电感参数改变为其标称值的 0.5 倍。







图 11 电感失配前后的稳态波形 (L_{d(θ1}) → 0.5L_{d(θ1}))
 Fig. 11 Steady state waveforms before and after inductance mismatch (L_{d(θ1}) → 0.5L_{d(θ1}))

从图 10 可以看出, DPCC 策略依赖永磁超环面 电机的直轴电感参数, 当直轴电感发生失配时, 转 速变化较小, 转矩、交轴与直轴电流脉动明显增 大。与 DPCC 策略相比, MFPCC-ESO 策略下的转 速、转矩、直交轴电流在直轴电感失配前后没有显 著变化, 同时, MFPCC-ESO 策略下永磁超环面电机 具有更好的电流跟踪性能。因此, 在直轴电感参数 失配条件下, MFPCC-ESO 策略的稳态性能好, 参数 鲁棒性强。

由图 11 可以看出, 直轴电感失配导致 DPCC 策略下的转速、转矩与直交轴电流均出现不同程度的调节过程, 转速存在明显的失调, 且交轴电流出现明显的跟踪误差。与 DPCC 策略相比, MFPCC-ESO 策略下的输出受参数失配的影响较小, 转速的失调量较少, 交轴电流跟踪误差较小。进一步证明了 MFPCC-ESO 策略具有优越的稳态性能和很强的直轴电感参数鲁棒性。

为进一步分析交轴电感、电阻及磁链参数失配 时采用 DPCC 策略与 MFPCC-ESO 策略的永磁超环 面电机稳态性能,对控制器参数分别变化为标称值 的 25%、50%、75%、100%、125%、150%、175% 时 的交轴电流进行分析。通过标准差公式计算得到 不同参数失配时 2 种控制策略下的交轴电流脉动 (见图 12),其交轴电流标准差为

$$\sigma_q = \sqrt{\frac{1}{N} \sum_{x=1}^{N} (i_q(x) - i_{q_a \mathbf{v}})^2}$$
(28)

式中: N 为采样个数; i_q(x)和 i_{q_av}分别为交轴电流的 采样数据和采样平均值。





由图 12 可以看出,当直交轴电感、电阻和磁链存在不同程度失配时,DPCC策略下,交轴电流脉动值均大于本文所提的 MFPCC-ESO策略。从整体看,相同参数失配对 DPCC策略下交轴电流脉动的影响显著大于 MFPCC-ESO策略。同时,电阻与磁链失配不会对 MFPCC-ESO策略产生影响。因此,进一步验证了在参数失配条件下,MFPCC-ESO策略具有比 DPCC策略优越的稳态性能与强鲁棒性。

为分析参数失配条件下采用 DPCC 策略与 MFPCC-ESO 策略的永磁超环面电机动态性能,研 究参数失配时 2 种控制策略下永磁超环面电机的 转速阶跃响应,如图 13 所示。直交轴电感、电阻与 磁链均为标称值的 1.5 倍,永磁超环面电机空载运



Fig. 13 Speed step responses under parameters mismatch

行,在 0.12 s 时转速从 1 000 r/min 升至 1 200 r/min。

由仿真结果可知,与 DPCC 策略相比,参数失 配时,采用 MFPCC-ESO 策略的转速、转矩及直交 轴电流具有较短的动态调整时间与较小的超调 量。因此, MFPCC-ESO 策略具有比 DPCC 策略优 越的动态性能。同时,由于 DPCC 策略涉及更多的 电机参数,随着多个参数发生失配,转矩与直交轴 电流存在更高的脉动。MFPCC-ESO 策略下,永磁 超环面电机输出稳定且波动幅值小,表现出很强的 参数鲁棒性。

图 14 为参数失配条件下采用 DPCC 策略与 MFPCC-ESO 策略时永磁超环面电机的负载阶跃响 应。直交轴电感、电阻与磁链均为标称值的 1.5倍,永磁超环面电机转速恒定为1000 r/min,在 0.12 s时负载从0 N·m 突变至6 N·m。图 14表明, 与 DPCC 策略相比,参数失配时,采用 MFPCC-ESO 策略的转速、转矩及直交轴电流经过短暂的动态调 整,能更快地达到稳态值。同时, MFPCC-ESO 策略 下的交轴电流超调量小于 DPCC 策略。因此,在负 载阶跃变化时, MFPCC-ESO 策略具有比 DPCC 策 略优越的动态性能。

为研究参数失配时永磁超环面电机 ESO 的性能,给出参数失配时 ESO 对直交轴干扰量估计结果 及电流估计误差结果,如图 15 所示。

由图 15 可以看出,参数失配时估计的直交轴



图 14 参数失配下的负载阶跃响应





图 15 参数失配下 ESO 的估计结果



干扰量依然存在由时变参数引起的周期性波动,其 在负载阶跃变化时能迅速响应,达到实时为预测模 型提供干扰估计值的目的。参数失配未引起直交 轴电流估计误差的增大,其数值大小在零附近,说 明永磁超环面电机 ESO 具有有效的稳定性和良好 的观测性。

6 结 论

本文从永磁超环面电机的结构原理特色出发, 引入自转运动影响系数与磁势系数,推导出与此系 数相关的时变电感参数与电机模型。针对传统 DPCC策略存在的参数依赖性问题,结合无模型控 制与延时补偿的 DPCC算法,研究了引入 ESO 的永 磁超环面电机 MFPCC策略。通过仿真实验对 MFPCC-ESO策略与 DPCC策略进行对比分析,得 到如下结论:

1) 在参数匹配情况下,与 DPCC 策略相比, MFPCC-ESO 策略在转速阶跃与负载阶跃阶段具有 更快的动态响应特性,同时该策略能有效降低永磁 超环面电机的输出波动和电流 THD 值,其稳态性 能更优越。

2)相比于 DPCC 策略, MFPCC-ESO 策略在永 磁超环面电机发生不同程度参数失配时的动稳态 性能更优越,同时,该控制策略对参数的鲁棒性 更强。

3) 在参数匹配和失配条件下, 永磁超环面电机 ESO 能对直交轴干扰量及电流进行有效估计, 具有良好的稳定性与观测性。

参考文献(References)

- [1] MAO J K, LI H M, YANG L G, et al. Non-cascaded model-free predictive speed control of SMPMSM drive system[J]. IEEE Transactions on Energy Conversion, 2022, 37(1): 153-162.
- [2] 匡晓霖,徐金全,黄春蓉,等.六相永磁同步电机驱动控制方式[J]. 北京航空航天大学学报,2019,45(7):1361-1369.
 KUANG X L, XU J Q, HUANG C R, et al. Drive-control modes of six-phase PMSM[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(7): 1361-1369(in Chinese).
- [3] SUN C, SUN D, CHEN W H, et al. Improved model predictive control with new cost function for hybrid-inverter open-winding PMSM system based on energy storage model[J]. IEEE Transactions on Power Electronics, 2021, 36(9): 10705-10715.
- [4] 苏光靖,李红梅,李争,等. 永磁同步直线电机无模型电流控制[J].
 电工技术学报, 2021, 36(15): 3182-3190.
 SU G J, LI H M, LI Z, et al. Research on model-free current control of permanent magnet synchronous linear motor[J]. Transactions of China Electrotechnical Society, 2021, 36(15): 3182-3190(in Chinese).
- [5] ZHANG W J, XU Y L, ZHOU G X. Research on a novel transverse flux permanent magnet motor with hybrid stator core and

disk-type rotor for industrial robot applications[J]. IEEE Transactions on Industrial Electronics, 2021, 68(11): 11223-11233.

- [6] ZHOU S L, LI G L, WANG Q J, et al. Geometrical equivalence principle based modeling and analysis for monolayer halbach array spherical motor with cubic permanent magnets[J]. IEEE Transactions on Energy Conversion, 2021, 36(4): 3241-3250.
- [7] 洪玫,姚立纲. 基于齿面网格的超环面行星蜗杆传动系统实体建模[J]. 中国机械工程, 2014, 25(7): 867-872.
 HONG M, YAO L G. Entity modeling of toroidal drive based on tooth surface grid[J]. China Mechanical Engineering, 2014, 25(7): 867-872(in Chinese).
- [8] 郝秀红,朱学军,许立忠. 机电集成超环面传动系统参数振动研究[J]. 振动与冲击, 2013, 32(22): 113-118.
 HAO X H, ZHU X J, XU L Z. Parametric vibration of an electromechanical integrated toroidal drive[J]. Journal of Vibration and Shock, 2013, 32(22): 113-118(in Chinese).
- [9] LIU X, LI D, ZUO L. Modeling and control for an integrated permanent magnet toroidal motor drive with nonlinear electromagnetic parameters[J]. Applied Mathematical Modelling, 2021, 89: 154-170.
- [10] LIU X, WANG H D. Analytical calculation and analysis of air gap magnetic field for electromechanical integrated toroidal drive[J]. Advances in Mechanical Engineering, 2019, 11(12): 168781401989592.
- [11] 刘欣, 许立忠, 聂岭. 新型超环面混合励磁电机的结构及特性分析[J]. 中国电机工程学报, 2015, 35(20): 5335-5343.
 LIU X, XU L Z, NIE L. Structure and characteristic analysis of a novel toroidal motor with hybrid excitation[J]. Proceedings of the CSEE, 2015, 35(20): 5335-5343(in Chinese).
- [12] XU L Z, LI R, SONG W T. Permanent magnetic toroidal drive with half stator[J]. Advances in Mechanical Engineering, 2017, 9(1): 168781401668631.
- [13] 秦艳忠, 阎彦, 陈炜, 等. 永磁同步电机参数误差补偿-三矢量模型 预测电流控制[J]. 电工技术学报, 2020, 35(2): 255-265. QIN Y Z, YAN Y, CHEN W, et al. Three-vector model predictive current control strategy for permanent magnet synchronous motor drives with parameter error compensation[J]. Transactions of China Electrotechnical Society, 2020, 35(2): 255-265(in Chinese).
- [14] ZHANG Y C, JIANG H, YANG H T. Model predictive control of

PMSM drives based on general discrete space vector modulation[J]. IEEE Transactions on Energy Conversion, 2021, 36(2): 1300-1307.

- [15] LIU X C, ZHOU L B, WANG J, et al. Robust predictive current control of permanent-magnet synchronous motors with newly designed cost function[J]. IEEE Transactions on Power Electronics, 2020, 35(10): 10778-10788.
- [16] SUN X D, WU M K, LEI G, et al. An improved model predictive current control for PMSM drives based on current track circle[J]. IEEE Transactions on Industrial Electronics, 2021, 68(5): 3782-3793.
- [17] 史涔溦, 解正宵, 陈卓易, 等. 永磁同步电机无参数超局部模型预 测控制[J]. 电机与控制学报, 2021, 25(8): 1-8. SHI C W, XIE Z X, CHEN Z Y, et al. Model-free predictive control based on ultra-local model for permanent magnet synchronous machines[J]. Electric Machines and Control, 2021, 25(8): 1-8(in Chinese).
- [18] NIU S X, LUO Y X, FU W N, et al. Robust model predictive control for a three-phase PMSM motor with improved control precision[J]. IEEE Transactions on Industrial Electronics, 2021, 68(1): 838-849.
- [19] ZHOU Y N, LI H M, LIU R D, et al. Continuous voltage vector model-free predictive current control of surface mounted permanent magnet synchronous motor[J]. IEEE Transactions on Energy Conversion, 2019, 34(2): 899-908.
- [20] MA C W, LI H Y, YAO X L, et al. An improved model-free predictive current control with advanced current gradient updating mechanism[J]. IEEE Transactions on Industrial Electronics, 2021, 68(12): 11968-11979.
- [21] YU F, ZHOU C H, LIU X, et al. Model-free predictive current control for three-level inverter-fed IPMSM with an improved current difference updating technique[J]. IEEE Transactions on Energy Conversion, 2021, 36(4): 3334-3343.
- [22] 赵凯辉,周瑞睿,冷傲杰,等. 一种永磁同步电机的有限集无模型 容错预测控制算法[J]. 电工技术学报, 2021, 36(1): 27-38. ZHAO K H, ZHOU R R, LENG A J, et al. Finite control set modelfree fault-tolerant predictive control for permanent magnet synchronous motor[J]. Transactions of China Electrotechnical Society, 2021, 36(1): 27-38(in Chinese).

Model-free predictive current control for permanent magnet toroidal motor with extended state observer

LIU Xin^{1, 2, *}, WANG Zhengyang¹, WANG Xiaoyuan²

Tianjin Key Laboratory of Modern Electromechanical Equipment Technology, Tiangong University, Tianjin 300387, China;
 School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China)

Abstract: The model-free predictive current control method for the toroidal motor with extended state observer (MFPCC-ESO) was investigated in order to reduce the dependence of deadbeat predictive current control (DPCC) on the parameters of a permanent magnet toroidal motor system. According to the composite rotor structure of a toroidal motor, the rotation motion influence coefficient and magnetomotive force coefficient were introduced. The rotating coordinate system was used to build the time-varying mathematical model of the toroidal motor. Then the ultra-local model with a time-varying scaling factor was established by using the input and output of the toroidal motor system. Meanwhile, ESO was introduced to estimate the interference part of the ultra-local model in real time, and the stability of ESO was proved by using the Jury criterion. Combining with the delay compensation DPCC, the reference voltage vector was obtained. The MFPCC-ESO for the toroidal motor was further realized. The MFPCC-ESO strategy and DPCC strategy were compared and analyzed for toroidal motor under parameter matching and mismatch. The simulation results show that the toroidal motor with the MFPCC-ESO strategy has superior dynamic performance, steady-state performance, and strong robustness. Meanwhile, the proposed control strategy can also reduce the output fluctuation of the toroidal motor effectively.

Keywords: toroidal motor; ultra-local model; model-free; predictive current control; extended state observer

Foundation items: National Natural Science Foundation of China (51875408); Tianjin Research Innovation Project for Postgraduate Students (2021YJSB229)

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

DOI: 10.13700/j.bh.1001-5965.2022.0410

跨级融合门控自适应网络用于视网膜血管分割

梁礼明*,余洁,陈鑫,雷坤,周珑颂

(江西理工大学电气工程与自动化学院,赣州 341000)

摘 要: 针对现有多数算法对浅层特征提取不足,导致分割结果中血管边界模糊、毛细血 管欠分割且包含噪声等问题,提出一种跨级融合门控自适应网络。该网络中的密集门控通道变换模 块,通过促进通道之间的竞争或协同关系充分提取浅层特征信息,避免浅层粗粒度特征信息丢失; 通过跨层次融合模块捕获各层跨维度交互信息,有效聚合多尺度上下文特征;采用双自适应特征融 合方法有效引导相邻层次特征融合,抑制噪声。在公共数据集 DRIVE、CHASEDB1 和 STARE 上进 行验证,结果表明:所提网络准确率分别为 0.965 2、0.966 8 和 0.969 5, F₁ 值分别为 0.854 4、0.815 2 和 0.841 2,在多个指标上均处于较高水平,优于现有先进算法。

关 键 词:视网膜血管分割;密集门控通道变换;跨层次融合模块;双自适应特征融合;三 重注意力模块

中图分类号: TP391

文献标志码: A 文章编号: 1001-5965(2024)04-1097-13

视网膜在高血压、糖尿病视网膜病变、动脉硬 化等眼科疾病的诊断中发挥重要作用,视网膜血管 的形态信息(如厚度、曲率、密度等)可作为疾病检 测和诊断的重要指标。然而,目前的临床实践中通 常采用人工目测的方法获取形态学信息,费时费力 且主观性强,因此,从眼底图像中自动准确地分割 视网膜血管是一项非常有必要的研究课题。

早期研究的重点在于利用手工制作特征来分 割视网膜血管的经典机器学习算法,尽管在某种特 定情况下取得了较好的分割性能,但手工制作特征 不足以代表视网膜血管的复杂语义,因此,在具有 许多复杂病例的数据集中表现较差^[1-4]。近年来,许 多基于全卷积网络的视网膜血管分割算法被提出, 主要基于对称的U型结构,通过不断堆积卷积层和 下采样来逐步提取上下文特征。文献[5]提出一种 由注意力机制捕获全局信息并将其置于特征融合 中增强特征的卷积神经网络;文献[6]将补丁学习 策略运用于 Dense U-Net,提升网络分割性能;文献 [7] 将多尺度输入层和密集块引入 U-Net 中,使网络 能够利用更丰富的空间上下文信息;文献 [8] 通过 引入残差网络避免血管结构信息丢失,并将空洞卷 积置于网络底端扩大感受野;文献 [9] 采用与文献 [7] 相似的多尺度输入结构,并引入多特征融合模 块提取多尺度血管特征,提高网络对细小血管分割 能力。尽管 U 型结构实现了可观的性能,但仍然不 足以应对视网膜血管分割的挑战,一般来说存在以 下问题:①网络通常缺乏有效提取多尺度上下文信 息的能力;②每个阶段的跳跃连接直接结合局部信 息,引入了不相关的背景噪声,使网络很难将视网膜 血管与周围渗出物和噪声区分开,尤其是微小血管。

针对以上问题,本文提出一种跨级融合门控自适应网络(cross-level fusion gated adaptive network, CFGA-Net)用于视网膜血管分割。①提出密集门控通道转换(dense gated channel transformation, DECT)

收稿日期: 2022-05-21; 录用日期: 2022-08-05; 网络出版时间: 2022-08-23 11:03 网络出版地址: link.cnki.net/urlid/11.2625.V.20220823.0958.002

*通信作者. E-mail: 9119890012@ jxust.edu.cn

基金项目:国家自然科学基金 (51365017,61463018);江西省自然科学基金面上项目 (20192BAB205084);江西省研究生创新专项资金 (YC2021-S585);江西省教育厅科学技术项目 (GJJ170491,GJJ2200848)

引用格式: 梁礼明, 余洁, 陈鑫, 等. 跨级融合门控自适应网络用于视网膜血管分割 [J]. 北京航空航天大学学报, 2024, 50 (4): 1097-1109. LIANG L M, YU J, CHEN X, et al. Cross-level fusion gated adaptive network for retinal vessel segmentation [J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50 (4): 1097-1109 (in Chinese).

模块,建立通道之间的竞争或合作关系,提升特征 复用率,改善深层网络的鲁棒性。②提出跨层次融 合模块(cross-level fusion module, CFM),置于网络 底端,捕获跨维度交互作用,计算注意权值,提取隐 藏的上下文信息,并聚合多尺度特征,提高网络处 理血管尺度与形态变化较大数据的能力。③提出 双自适应特征融合(dual adaptive feature fusion, DAFF)模块,自适应结合语义信息与空间信息,抑 制不相关的背景噪声,有效引导相邻层次特征融 合,以捕获更多可区分的语义信息。

1 本文网络

1.1 密集门控通道变换模块

密集卷积是文献 [10] 提出的一种结构,其将本 层输出与前面所有层输出连接作为下层输入,提升 特征复用率,但单个卷积层只对特征图中空间位置 的邻近局部上下文起作用,引起局部歧义。为解决 该问题,本文引入门控通道变换 (gated channel transformation, GCT) 模块增强卷积层。GCT 模块 结构如图 1 所示,其采用规范化组件来建立通道之 间的竞争或合作关系。由于规范化操作是无参数 的,为使 GCT 模块具有可学习性,设计了全局上下 文嵌入算子和门控自适应算子,前者用于嵌入全局 上下文,并控制归一化前每个通道的权重,后者根 据归一化输出对输入特征通道进行智能调整。假 设 $x \in \mathbf{R}^{C \times H \times W}$ 为特征图中的激活特征,*C* 为特征图 中的通道数,*W* 为特征图的宽度,*H* 为特征图的高 度,GCT 模块对其进行如下变换:

 $\hat{\boldsymbol{x}} = F(\boldsymbol{x} | \boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\beta}) \qquad \boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\beta} \in \mathbf{R}^C$ (1)

式中: α 、 γ 、 β 为可训练参数, α 有助于嵌入输出的 自适应性, γ 、 β 用于控制激活门限。



Fig. 1 GCT module

GCT 模块变换由以下 3 部分组成:

 1)全局上下文嵌入算子。具有大感受野的信息能有效避免由小感受野的信息引起的局部混淆,因此,设计全局上下文嵌入算子用于汇聚每个通道的全局上下文信息,该算子能利用卷积层小感受野之外的全局上下文信息。给定嵌入权重α = [α₁,...,α_c,...],c ≤ C, 全局上下文嵌入算子在通道 c上的定义为

$$\boldsymbol{s}_{c} = \boldsymbol{\alpha}_{c} \|\boldsymbol{x}_{c}\|_{2} = \boldsymbol{\alpha}_{c} \left\{ \left[\sum_{i=1}^{H} \sum_{j=1}^{W} \left(\boldsymbol{x}_{c}^{i,j} \right)^{2} \right] + \varepsilon \right\}^{\frac{1}{2}}$$
(2)

式中: ε为极小的常量,以避免在零点处求导。

与多数注意力机制不同的是,GCT模块并不使 用全局平均池化 (global average pooling,GAP)进行 通道上下文信息的汇聚,因为GAP在某些极端情 况下会失效。例如,当压缩和激励 (squeeze-andexcitation,SE)注意力模块位于实例规范化 (instance normalization,IN) 层之后时,IN 层将每个特征通道 的均值固化,导致GAP 的输出对于任何输入都是 恒定的。为避免该问题,GCT模块采用 *l*₂范数代替 GAP。同时,提出采用可训练参数*α*。控制不同通道 的权重。特别是当α_c近似于0时,通道 c 将不会被 纳入到通道规范化中,即α_c使 GCT 模块能学习到 一个通道相对其他通道是独立的极限情况。

2) 通道归一化。归一化方法可以建立通道之 间的竞争关系,具有轻量和稳定的性能,与局部响 应归一化 (local response normalization, LRN) 层类 似, GCT 模块采用 l_2 范数做跨通道归一化。设通道 $s = [s_1, \dots, s_c, \dots]$,则跨通道归一化公式为

$$\hat{\mathbf{s}}_{c} = \frac{\sqrt{c}\mathbf{s}_{c}}{\|\mathbf{s}\|_{2}} = \sqrt{c}\mathbf{s}_{c} / \left(\sum_{c=1}^{C} \mathbf{s}_{c}^{2} + \varepsilon\right)^{\frac{1}{2}}$$
(3)

式中: \sqrt{c} 用于归一化 \hat{s}_c 的尺度,避免当C较大时 \hat{s}_c 的尺度过小。

3) 门控自适应算子。GCT 模块采用门控自适 应机制来适应原有特征, 使其在训练过程中既促进 通 道 间 的 竞 争, 也 促 进 合 作 。 令 选 通 权 值 $\gamma =$ $[\gamma_1, \dots, \gamma_c, \dots]$ 和 选 通 偏 置 $\beta = [\beta_1, \dots, \beta_c, \dots]$, 设 置 如下选通函数:

$$\hat{\boldsymbol{x}}_c = \boldsymbol{x}_c \left[1 + \tanh(\boldsymbol{\gamma}_c \hat{\boldsymbol{s}}_c + \boldsymbol{\beta}_c) \right]$$
(4)

每个原始通道 x_c 的权重由其对应的门调整,即 1+tanh($\gamma_c \hat{s}_c + \beta_c$)。由于通道归一化是无参数的,设 计可训练的权值γ和偏移量β来学习如何对门控通 道的激活进行智能控制。通过结合归一化方法和 门控机制,GCT模块能够促进通道之间的竞争和协 同关系。当通道的门控权值γ。被正激活时,GCT模 块取得了类似于 LRN 层的通道竞争关系。当门控 权值γ。被负激活时,GCT模块鼓励该通道与其他通 道协作。同时,当选通权值和偏置为零时,该选通 函数允许原始特征传递到下一层,描述如下:

$$\hat{\mathbf{x}} = F(\mathbf{x}|\boldsymbol{\alpha}, \mathbf{0}, \mathbf{0}) = \mathbf{x} \tag{5}$$

本文结合 GCT 模块与密集卷积设计 DGCT 模

块用于特征提取,如图 2 所示。DGCT 模块主要包含 4 层特征提取层,每个特征提取层先经过 GCT 模块鼓励浅层次通道之间的合作关系,增强深层次 通道间的竞争关系,提升网络的泛化能力,再进行 降维操作,减少参数量后输入卷积层学习提取特征,最终将本层输入与输出融合作为下层输入。第 *i* 层的输入和输出为

$$\boldsymbol{X}_{i-\text{in}} = \operatorname{cat}(\boldsymbol{X}_{i-1-\text{in}}, \boldsymbol{X}_{i-1-\text{out}}) \tag{6}$$

$$\boldsymbol{X}_{i-\text{out}} = \text{Conv}(\text{GCT}(\boldsymbol{X}_{i-\text{in}})) \tag{7}$$

式中: Conv 表示卷积操作; cat 表示特征融合。



图 2 DGCT 模块 Fig. 2 DGCT module

1.2 三重注意力机制

经过多年的计算机视觉研究,越来越多的卷积 神经网络结构都使用通道注意力和空间注意力或 结合两者提高性能。传统通道注意力的计算方法通 常采用 GAP 将输入张量分解为每个通道一个像素, 导致空间信息大量损失,通道和空间维度之间的相互 依赖性不存在。卷积块注意力模块^[11](convolutional block attention module, CBAM)将空间注意力引入通 道注意力,以弥补空间信息的损失,但通道注意和 空间注意是相互分离计算的,未考虑两者之间的任 何关系。本文引入一种轻量级但有效的三重注意 力机制 (triple attention mechanism, TAM),通过旋转 操作和残差变换来建立输入张量 (*C*, *H*)、(*C*, *W*) 和 (*H*, *W*) 维间的依赖关系,捕获跨维度交互作用来计 算注意力权值,结构如图 3 所示。 具体操作为:给定一个输入张量 $x \in \mathbb{R}^{C \times H \times W}$,将 其传递给 TAM 中的 3 个分支。在第 1 个分支中, 建立高度和通道维度间的相互作用。首先,将输入 x沿 H轴逆时针旋转 90°,得到旋转张量 $\hat{x}_1 \in \mathbb{R}^{W \times H \times C}$, 通过 Z-Pool 层拼接该维度上的平均池化和最大池 化特征,将张量 \hat{x}_1 的第 0 维减少到 2,即 $\hat{x}_1^* \in \mathbb{R}^{2 \times H \times C}$, 使该层能够保留实际张量的丰富特征,同时缩小其 深度使进一步计算变得轻量级;然后,将 \hat{x}_1^* 经过卷 积层和 BN 层后通过 Sigmoid 激活函数生成注意力 权值作用于 \hat{x}_1 ;最后,沿 H 轴顺时针旋转 90°以维持 输出 x 原始输入形状。在第 2 个分支中,将 x沿着 W 轴逆时针旋转 90°得到旋转张量 $\hat{x}_2 \in \mathbb{R}^{H \times C \times W}$,通 过 Z-Pool 层简化为 $\hat{x}_2^* \in \mathbb{R}^{2 \times C \times W}$ 后生成注意力权值作 用于 \hat{x}_1 ,沿 W 轴顺时针旋转 90°以维持x原始输入 形状。在第 3 个分支中,输入张量x的通道数经过



Fig. 3 Triple attention mechanism

Z-Pool 层减少为 2,得到缩减张量 $\hat{x}_3 \in \mathbb{R}^{2 \times H \times W}$,生成 注意力权值校准x。平均聚合 3 个分支中生成的细 化张量,得到最终输出y,即

$$y = \frac{1}{3} \left(\widehat{\mathbf{x}}_1 \sigma(C_1(\widehat{\mathbf{x}}_1^*)) + \overline{\mathbf{x}}_2 \sigma(C_2(\widehat{\mathbf{x}}_2^*)) + \mathbf{x} \sigma(C_3(\widehat{\mathbf{x}}_3)) \right) \quad (8)$$

式中: σ 为 Sigmoid 激活函数; C_1 、 C_2 和 C_3 为标准卷 积层。

设输入特征图为($B, C_{in}, W_{in}, H_{in}$),输出特征图为 ($B, C_{out}, W_{out}, H_{out}$), Liner(C_{in}, C_{out} , bias = False)为全连 接层, B为批次,表示在一个批次中处理的样本数 量。在计算通道注意力时,瓶颈处使用的缩减比为r, 参数量为 $C_{in} \times C_{out}$; 2D 卷积Conv2d(C_{out}, C_{in} ,kernel = K,bias = False)参数量为 $K \times K \times C_{in} \times C_{out}$ 。同时,计 算了不同注意力机制在 ResNet-50 上的开销,方法 为:在每个块中添加注意层,将r固定为 16, CBAM 和 TAM 的 k 值固定为 7。如表 1 所示, TAM 开销 成本远低于 CBAM、全局上下文和 SE 注意力机制。

1.3 跨层次融合模块

为避免信息丢失并充分利用多尺度特征,对不同编码器的空间语义特征进行深度上下文特征提取,本文基于 TAM 提出 CFM,其结构如图 4 所示。由于不同层的特征对任务的贡献不同,且具有更高分辨率的低层特征比高层特征需要更多的计算资

表1 不同注意力机制的参数复杂度和开销比较

Table 1 comparison of parameter complexity and overhead of various attentional mechanisms

注意力机制	参数量	存储开销/10 ⁶
SE	$2C^{2}/r$	2.514
CBAM	$2C^{2}/r+2k^{2}$	2.532
全局上下文	$2C^{2}/r+C$	2.548
TAM(本文)	$6k^2$	0.004 8

源,但对模型性能贡献较小,首先,通过最大池化操 作将第1~3层编码器输出映射到与第4层编码器 输出相同的特征空间,即

$$\hat{\boldsymbol{x}}_{\text{En-in}}^{i} = \text{pool}(\boldsymbol{x}_{\text{En-in}}^{i}) \qquad i = 1, 2, 3 \tag{9}$$

其次,通过1×1卷积对其降维得到特征图特征 x后,将x分别输入并行的空洞卷积层和TAM中, 通过空洞卷积层提取多尺度特征,其中第1~4层 的卷积核大小均为3×3,空洞率分别设置为1、2、 3、5,TAM捕获跨维度交互作用计算注意权值校准 多尺度特征信息,经过Sigmoid函数将其权重重新 标定得到x',将特征图x与x'相乘获得本层输出。 最后,将多层次特征融合相互补充,得到一个全面 的特征表达输入解码层。



Fig. 4 Cross-level fusion module

1.4 双自适应特征融合模块

在原始 U-Net 中, 编码器和解码器路径是通过 跳跃式直接连接的, 但两者之间存在一定语义差 距, 不同通道特征图对目标区域的恢复有不同贡 献, 但直接连接将引入噪声。为自动调整不同特征 映射的贡献、弱化由于跳过连接而产生的语义空 白,文献 [12] 在跳跃连接处引入自适应特征融合模 块对血管特征进行结构性补充,但对浅层次特征图 中的噪声过滤不完全,易产生嘈杂的空间特征图。 因此,本文提出 DAFF 模块,将编码部分的低级特 征与解码部分经上采样的高级特征结构信息相互 补充融合,动态权衡血管与背景噪声的权重比,得 (11)

到深层与浅层特征的综合映射。如图 5 所示, DAFF 模块的输入为编码部分的低级特征映射 x_{en}和解码 部分经上采样的高级特征映射 x_{de}, 生成 2 个新特征 x'_{en}和 x'_{de}, 其计算公式如下:

$$\boldsymbol{x}_{en}' = \operatorname{cat}(\boldsymbol{\rho} \circ \boldsymbol{x}_{en}, (1 - \boldsymbol{\rho}) \circ \boldsymbol{x}_{de})$$
(10)

 $x'_{de} = \operatorname{cat}(\boldsymbol{\rho} \circ x_{de}, (1-\boldsymbol{\rho}) \circ \boldsymbol{x}_{en})$

式中: "o" 表示哈达玛积; p为融合系数矩阵, 计算

公式为

$$\boldsymbol{\rho} = S\left(\varphi(R(\varphi_{\rm en}\boldsymbol{x}_{\rm en} + \varphi_{\rm de}\boldsymbol{x}_{\rm de}))\right) \tag{12}$$

其中: φ 、 φ_{en} 、 φ_{de} 为线性变换操作; R和 S分别为 Relu 和 Sigmoid 激活函数。

最终, DAFF 模块的输出为

$$\boldsymbol{x}_{\text{dein}} = \boldsymbol{x}_{\text{de}}' + \boldsymbol{x}_{\text{en}}' \tag{13}$$



图 5 双自适应特征融合模块 Fig. 5 Dual adaptive feature fusion module

1.5 网络结构

本文提出的 CFGA-Net 主要由 5 部分组成:数据预处理、U型结构、DGCT 模块、DAFF 模块和 CFM,整体结构如图 6 所示。原始视网膜图像经过 预处理后裁剪为 1 像素×64 像素×64 像素的随机块 输入 U 型网络, 网络对随机块进行通道扩张,得到 尺寸为 48 像素×64 像素×64 像素的特征图,再对这 些特征图进行编码解码操作,本文网络编码解码层 采用 DGCT 模块取代普通卷积层,以充分提取血管 特征,鼓励浅层次通道之间的合作关系,增强更深 层次通道间的竞争关系。连接部分分别为 CFM 和 DAFF 模块, CFM 置于网络底端, 将不同编码器空 间的语义特征融合, 以辅助原始编码器进行深度上 下文特征提取, 而 DAFF 模块置于跳跃连接处, 将 低层特征图和高层特征图自适应结合语义信息和 空间信息, 抑制低层不相关的背景噪声, 保留高层 更详细的局部语义信息, 解码器最后一层为带有 Softmax 激活函数的 1×1 卷积, 输出结果为 1 像素× 64 像素×64 像素局部分割图, 最终将局部分割结果 拼接得到血管分割结果。





2 实验结果与对比分析

2.1 数据集与预处理

本文在DRIVE、STARE、CHASEDB1 等3个数据

集上对模型视网膜血管分割性能进行评估。DRIVE 数据集由 40 张分辨率为 565 像素×584 像素的视网 膜图像构成,官方已将其平均划分为训练集和测试 集; STARE 数据集包含 20 张分辨率为 700 像素× 605 像素的视网膜图像; CHASEDB1 数据集包含 28 张分辨率为 999 像素×960 像素的儿童视网膜图 像。由于数据集 CHASEDB1 和 STARE 未划分特定 的训练集和测试集,为公平起见,将 CHASEDB1 数 据集前 20 张用于训练,后 8 张用于测试,而 STARE 数据集采用留一法进行交叉验证。

视网膜图像中存在血管与背景对比度较低和 光照不均匀现象,为得到更好的分割结果,本文对 图像进行如下预处理操作:首先,进行 RGB 三通道 加权 (rgb three channel weighting, TCW)。将图像分 离为 RGB 三通道特征图,绿色通道中,血管与背景 对比度高能较好显示血管结构,但仅提取绿色通道 将损失其他通道中的血管特征,因此,提出通道加

$$\boldsymbol{g} = \eta \boldsymbol{g}_{\mathrm{R}} + \kappa \boldsymbol{g}_{\mathrm{G}} + \iota \boldsymbol{g}_{\mathrm{B}} \tag{14}$$

式中:**g**为彩色图像的三维张量;η、κ和ι分别为 RGB三通道融合权重系数。

然后,进行限制对比度的自适应直方图均衡化 (contrast limited adaptive histogram equalization, CLA-HE),进一步提高血管和背景的对比度。最后,进行 自适应伽马变换,提升暗部细节并降低伪影干扰。 预处理各阶段图像如图7所示。由于训练数据较 少,为得到稳定可靠的结果,减少过拟合现象,本文 对训练集数据进行如下方式扩充:①随机旋转操 作,将数据扩充10倍;②采用大小为64像素×64像 素的窗口将图像随机截取为局部图像块。





(b) TCW





(d) 自适应伽马变换

(a) 原始图像

图 7 预处理各阶段效果 Fig. 7 Effect of each stage of preprocessing

2.2 实验设置

本文实验环境基于 AMD R5 5600X CPU 32 GB 内存,显卡为 NVIDA Geforce RTX 3060Ti GPU 8 GB,采用 Ubuntu16.04 操作系统。模型在 pytorch 1.7.1 框架上进行实验, batch size 为 16, epoch 为 100,采用交叉熵损失函数, Adam 优化器, 初始学习 率为 0.001,学习率变化为每 5 轮降低 0.5 倍, 若连 续 7 次迭代未取得最优, 则停止训练。

2.3 评估指标

为定量评估算法分割性能,本文采用灵敏度 S_e、特异性S_p、准确率A_{cc}、F₁分数、ROC 曲线下和 PR 曲线下的 AUC 面积作为客观评价指标,各具体 计算式分别为

$$S_{\rm e} = \frac{T_{\rm P}}{T_{\rm P} + F_{\rm N}} \tag{15}$$

$$S_{\rm p} = \frac{T_{\rm N}}{T_{\rm N} + F_{\rm P}} \tag{16}$$

$$A_{\rm cc} = \frac{T_{\rm P} + T_{\rm N}}{T_{\rm P} + F_{\rm P} + T_{\rm N} + F_{\rm N}}$$
(17)

$$F_{1} = \frac{2T_{\rm P}}{2T_{\rm P} + F_{\rm P} + F_{\rm N}} \tag{18}$$

式中: T_{P} 、 T_{N} 、 F_{P} 、 F_{N} 分别为真阳性、真阴性、假阳性、假阴性像素样本。

2.4 实验结果分析

为更直观地体现本文算法在视网膜血管上的 分割性能,在公开数据集 DRIVE、CHASEDB1和 STARE上进行验证,并与文献 [13-15] 算法进行对 比,为公平比较,文献 [13-15] 采用与本文相同的实 验环境。文献 [13] 为U-Net 算法,文献 [14] 在文献 [13] 的基础上增加注意力机制自动学习参数来调整权 重的 Attention U-Net,文献 [15] 在文献 [14] 的基础 上提出 SE-Resnet 结构代替普通卷积进行编码解 码,一定程度上提升了网络对细小血管的分割能力。 2.4.1 分割结果客观分析

不同算法在 3 个公开数据集上的平均性能指标对比如表 2 所示。可以看出,本文算法在所有指标上基本达到领先。本文算法 CHASEDB1 数据集上所有指标均为最优;虽然文献 [15] 算法的灵敏度在 DRIVE 数据集上略高本文算法 0.002 5,文献 [14] 算法的特异性在 STARE 数据集上略高本文算法 0.001 2,但对于医学图像分割,F₁值更为重要,其包含灵敏度信息的同时还考虑准确率,F₁值越高,表

1.0

明算法所识别的血管像素越齐全和准确。本文算 法的 F₁ 值在 3 个数据集上相比其他算法平均高出 0.006 7、0.009 5 和 0.020 7, 说明其能够精准识别出 更多的血管像素。

图 8 和图 9 展示了 ROC 曲线和 PR 曲线。由于 视网膜中细小血管占比较小,模型对该类样本的分 割性能在这 2 个指标中难以体现,但本文算法在数 据集 DRIVE 和 CHASEDB1 上 AUC(ROC) 和 AUC (PR)值相比其他算法平均高出 0.0017、0.0024 和 0.0064、0.0124,说明本文算法有效提升了细小血 管分割的准确率,减少了误分割现象。可以看出, 本文算法整体性能最优。

2.4.2 分割结果主观分析

图 10 为本文算法与文献 [13-15] 算法在不同数据集上的整体分割结果。其中,第1~3 行分别为数据集 DRIVE、CHASEDB1 和 STARE 中健康视网膜图像,第4~6 行分别为数据集 DRIVE、CHASEDB1和 STARE 中病变视网膜图像。

观察图 10 第 1~3 行可知,由于视盘区域附近 光照不均, 文献 [13-14] 算法均出现不同程度的主 血管断裂现象,虽然文献[15]算法与本文算法很好 地避免了该现象,但文献[15]算法与其他算法均在 视网膜边界处出现了血管断裂现象,且有不同程度 的细小血管误分割现象,而本文算法可对不同血管 进行区分并精准分割,这是因为文献[13-14]算法 仅采用普通卷积层进行编码解码,其特征提取能力 较低, 而文献 [15] 算法将残差学习与 SE 注意力机 制相结合提出 SE-Resnet 结构增强特征提取能力, 但对细小血管特征提取能力不强。本文算法利用 DGCT 模块进行特征提取,在增强特征提取能力的 同时,通过门控自适应机制适应原有特征,降低光 照不均的影响。从图 10 第 4、6 行可知,由于眼底 图像血管与背景对比度低,其他算法均出现血管欠 分割现象,这是因为高层特征不足以描述细节信 息,低层特征中背景混乱和语义歧义,文献 [13] 算 法直接将低层特征与高层特征拼接将引入背景噪

	表 2	不同数	据集上的	平均性能	指标评估	结果	
		-			_		-

数据集	算法	F_1	$A_{\rm cc}$	$S_{\rm e}$	$S_{ m p}$	AUC(ROC)	AUC(PR)
	U-Net ^[13]	0.8474	0.963 1	0.830 2	0.981 2	0.985 5	0.931 0
	Attention U-Net ^[14]	0.844 9	0.962 7	0.828 7	0.9814	0.985 2	0.929 9
DRIVE	ASR-UNet ^[15]	0.8507	0.964 1	0.835 1	0.982 1	0.986 2	0.933 6
	CFGA-Net(本文)	0.854 4	0.965 2	0.832 6	0.983 7	0.987 3	0.937 9
	U-Net ^[13]	0.800 2	0.964 1	0.793 4	0.981 1	0.983 4	0.884 8
	Attention U-Net ^[14]	0.8109	0.965 8	0.809 0	0.981 3	0.984 5	0.894 2
CHASEDBI	ASR-UNet ^[15]	0.806 1	0.965 0	0.804 7	0.980 9	0.984 8	0.891 6
	CFGA-Net(本文)	0.815 2	0.966 8	0.809 7	0.982 4	0.986 6	0.902 6
	U-Net ^[13]	0.8192	0.964 4	0.759 1	0.988 8	0.984 3	0.915 8
GTADE	Attention U-Net ^[14]	0.8184	0.964 6	0.752 2	0.9898	0.983 6	0.916 1
STARE	ASR-UNet ^[15]	0.823 9	0.965 3	0.764 9	0.989 1	0.985 3	0.920 1
	CFGA-Net(本文)	0.841 2	0.969 5	0.802 3	0.988 6	0.989 9	0.934 7



图 8 不同算法在 DRIVE 数据集上 ROC 曲线与 PR 曲线对比 Fig. 8 Comparison of ROC curves and PR curves of different algorithms on DRIVE dataset



图 9 不同算法在 CHASSEDB1 数据集上 ROC 曲线与 PR 曲线对比

Fig. 9 Comparison of ROC curves and PR curves of different algorithms on CHASSEDB1 dataset



图 10 不同算法分割结果

Fig. 10 Segmentation results of different algorithms

声, 文献 [14-15] 算法虽然引入注意力门(attention gate, AG)模块将低层特征用于细化高层特征, 使注

意力系数凸显局部区域的特征,但对于血管边缘信息提取不足,相较而言,本文算法采用 DAFF 模块

1105

对低级特征与高级特征结构信息相互补充融合,动态权衡血管与背景噪声的权重比,能够抑制背景噪声保留更多的血管局部特征与边缘信息,较好地避免了血管欠分割现象。观察图10第5行可知,在病变区域,文献[13-15]算法均将色素上皮改变区域误分割为血管,且出现不同程度的主血管断裂现象,而本文算法加入CFM,在提取血管多尺度信息的同时,捕获跨维度相互作用来计算注意力权值,增强血管特征权重,能够避免病变区域干扰,分割出更多血管目不易断裂。

为更为细致地对分割结果进行分析,图11详

细展示了U-Net^[13]、AttentionU-Net^[14]、ASR-UNet^[15]和本文算法 CFGA-Net 在数据集 DRIVE、CHASEDB1上视网膜边界、视盘、病变区域和大血管交叉区域局部分割细节。可以看出,本文算法在视网膜边界光照不均区域或视盘周围伪影区域,能够识别低对比度下的细小血管,精准分割出主血管和细小血管;在多条血管交叉或紧靠的情况下,提取的血管不易出现断裂或分割不足现象,同时血管轮廓清晰连通性好;在病变区域,本文算法能够识别硬性渗出物,避免生成伪影,总体上分割结果与标签基本一致。



图 11 分割结果细节对比

Fig. 11 Details comparison of segmentation results

2.4.3 算法对比

为验证本文算法的先进性,从灵敏度、特异性 和准确率3个指标将本文算法与近年来先进算法 进行定量对比,对比结果如表3所示。

文献 [21] 提出金字塔尺度聚合模块用于编码 解码, 以聚合当前、较高和较低层特征, 并分别对编

表 3 不同算法在数据集 DRIVE、CHASEDB1 和 STARE 上的性能指标对比

 Table 3
 Comparison of performance indicators of different algorithms in DRIVE, CHASEDB1 and STARE datasets

笛社	Se				$S_{\rm p}$		$A_{ m cc}$		
舟伍	DRIVE	CHASEDB1	STARE	DRIVE	CHASEDB1	STARE	DRIVE	CHASEDB1	STARE
文献[16]	0.765 3	0.763 3	0.758 1	0.981 8	0.980 9	0.9846	0.954 2	0.961 0	0.961 2
文献[17]	0.763 2	0.781 5	0.742 3				0.953 6	0.958 7	0.960 3
文献[18]	0.763 1	0.764 1	0.773 5	0.982 0	0.9806	0.985 7	0.953 8	0.960 7	0.963 8
文献[19]	0.791 8	0.645 7	0.802 1	0.970 8	0.965 3	0.9561	0.957 7	0.934 0	0.944 5
文献[20]	0.794 1	0.8176	0.7598	0.979 8	0.9704	0.9878	0.955 8	0.960 8	0.964 0
文献[21]	0.821 3	0.803 5		0.9807	0.978 7		0.961 5	0.963 9	
文献[22]	0.735 2	0.727 9	0.726 5	0.977 5	0.965 8	0.975 9	0.948 0	0.945 2	0.9548
文献[23]	0.835 3	0.8176	0.7946	0.975 1	0.977 6	0.982 1	0.957 9	0.963 2	0.962 6
文献[24]	0.812 5	0.801 2	0.8078	0.976 3	0.973 0	0.972 1	0.961 0	0.957 8	0.9586
CFGA-Net(本文)	0.832 6	0.809 7	0.802 3	0.983 7	0.982 4	0.9886	0.965 2	0.9668	0.969 5

码器和解码器中的金字塔尺度聚合模块进行金字 塔输入增强和深度金字塔监督2种优化,以改善细 小血管特征提取,但2个数据集指标均低于本文 算法,其中,灵敏度低 0.0113 和 0.0062,特异性低 0.0030和0.0037,准确率低0.0037和0.0029,说明 DGCT 模块对细小血管特征提取能力强于金字塔 尺度聚合模块。文献 [24] 提出将视网膜图像分解 为2个子区域,用其自身的增强滤波器进行处理提 取独特特征,并采用多尺度三棒滤波对细小血管进 行检测,使用改进的 3D 块匹配对图像进行去噪处 理,其在 STARE 数据集中灵敏度取得最高,但特异 性与准确率比本文算法低 0.0165 和 0.0109。文献 [20] 提出 AA-UNet, 将边界框回归到视网膜区域 生成注意力掩码,并将其作为加权函数与模型中的 差分特征图相乘,使模型更加关注血管区域,并 在底部采用密集空洞卷积模块以重用特征, 使血管 的细节更加突出,其在 CHASSEDB1 数据集中灵 敏度比本文算法高 0.0079, 但其在数据集 DRIVE、 STARE 上灵敏度比本文算法低 0.0385 和 0.0425, 表明其提取多尺度血管能力不如 CFM, 泛化性能 较低。文献 [23] 提出一种多任务分割网络, 分别从

眼底图像中精确分割出粗血管和细血管,并设计 新型损失函数解决粗细血管比例不均衡问题,通过 融合网络将多任务分割网络输出融合,使其在数 据集 DRIVE和 CHASEDB1上灵敏度达到最高,分 别比本文算法高 0.0027和 0.0079,而本文算法采 用 DAFF 模块与 CFM 能很好地适应粗细血管比例 与背景不均衡,增强多尺度特征提取能力,对细 小血管识别准确率较高,使本文算法特异性和准 确率分别比献 [23] 高 0.0086、0.0048和 0.0073、 0.0036。综合分析可知,本文算法具有较好的泛化 能力,能更加精准地分割不同大小的血管,鲁棒性 较强。

2.4.4 模型改进前后对比

为验证 CFGA-Net 中各模块的有效性, 在公共 数据集 DRIVE、CHASEDB1和 STARE上进行消融 实验, 结果如表 4 所示。表中: 模型 1 为原始 U-Net 网络, 模型 2 将模型 1 中普通卷积层替换为 DGCT 模块, 模型 3 在模型 2 的基础上加入 AFF 模块, 模 型 4 在模型 2 的基础上加入 DAFF 模块, 模型 5 在 模型 4 的基础上加入将 TAM 替换为 SE 注意力的 CFM, 模型 6 为本文算法。

表 4 算法改进前后结果对比

構刊	F_1		$A_{ m cc}$		Se		$S_{ m p}$			AUC					
侠望	DRIVE	CHASEDB1	STARE	DRIVE	CHASEDB1	STARE	DRIVE	CHASEDB1	STARE	DRIVE	CHASEDB1	STARE	DRIVE	CHASEDB1	STARE
1	0.8474	0.800 2	0.8192	0.963 1	0.964 1	0.9644	0.830 2	0.793 4	0.759 1	0.981 2	0.981 1	0.9888	0.985 5	0.983 4	0.984 3
2	0.852 7	0.813 9	0.824 3	0.964 8	0.966 1	0.9659	0.832 3	0.8193	0.781 0	0.983 3	0.9807	0.9873	0.986 7	0.986 1	0.987 0
3	0.853 0	0.812 2	0.824 1	0.964 7	0.965 9	0.9652	0.8296	0.813 4	0.777 0	0.982 6	0.981 1	0.9879	0.982 2	0.985 4	0.986 0
4	0.852 8	0.813 6	0.8276	0.964 9	0.9660	0.966 5	0.829 7	0.8223	0.786 5	0.983 9	0.9803	0.9874	0.987 0	0.986 0	0.987 5
5	0.853 0	0.813 2	0.833 1	0.965 0	0.966 1	0.9672	0.829 0	0.813 8	0.783 9	0.984 0	0.9813	0.9886	0.985 5	0.985 9	0.988 8
6	0.8544	0.815 2	0.841 2	0.965 2	0.9668	0.969 5	0.832 6	0.809 7	0.802 3	0.983 7	0.9824	0.9886	0.987 3	0.9866	0.989 9

由表4可知,模型2采用本文所设计的DGCT 模块,能充分提取更详细的血管特征,分割出更多 血管,相比模型1,除在数据集CHASEDB1、STARE 上特异性指标下降,其他指标均提升;模型3加入 AFF模块,多数指标基本下降;而模型4加入DAFF 模块,多数指标高于模型3;说明本文所提DAFF模 块增强血管特征权重、降低背景噪声能力比AFF 模块更优,模型5加入SE注意力引导的CFM,在数 据集DRIVE和CHASEDB1上的指标并未有明显变 化,在STARE数据集上,S。指标明显下降,其他指 标均小幅度提升,而模型6加入由三重注意力引导 的CFM,可以获取更大的感受野和更丰富的多尺度 血管信息,增强血管辨别能力,在STARE数据集 上,所有指标均明显提升,在 DRIVE 数据集上,除 S_p 略微下降,其他指标均有不同幅度提升,说明血 管识别能力有所增强,而由于 CHSDEDB1 数据集 视网膜图像尺寸较大,血管像素占比较少,灵敏度 未见明显提升,但 F_1 值、 A_{ee} 、AUC 值明显提升。综 合可知本文所提模块的有效性和算法设计的合理性。

为进一步证明本文提出的分割图像网络与模块的优势,对各部分组合的实验分割图进行比较。在图 12 中,通过局部放大并比较相应的组合分割图像可以看出,本文提出的 DGCT 模块能防止大血管交叉处断裂现象,DAFF 模块有效抑制病灶 区域特征,精准分割细小血管;结合所有模块能区分病灶区域精准分割粗细血管,对硬性渗出物处

梁礼明,等:跨级融合门控自适应网络用于视网膜血管分割

(a) 原图 (b) 标签 (c) U-Net[13] (d) DGCT (e) DGCT+DAFF (f) CFGA-Net 图 12 模型改进分割图像结果对比



分割骨折血管有一定修复效果,使网络性能达到 最高。

3 结 论

本文提出一种新的视网膜血管分割算法 CFGA-Net, 主要结论如下:

 1)设计密集门控通道变换模块增强特征信息 的提取与传递,缓解了分割边界模糊、细小血管分 割不足问题。

2)提出跨层次融合模块提取编码器各层特征
 图中多尺度信息,动态捕捉跨维度相互作用,充分
 利用图像的三维信息,提高细小血管分割能力。

3) 设计双自适应特征融合模块使网络深、浅

层信息传递更为顺畅,在各层信息融合过程中充分 利用更丰富的特征信息,大大降低分割图像中的噪 声,提高分割精度。

4)本文算法在3个公共数据集上进行了一系列的实验,结果表明,提出的CFGA-Net算法整体性能优于现有先进算法,消融实验证明了本文提出的3个创新模块的有效性。

参考文献(References)

[1] 贾洪,郑楚君,李灿标,等. 基于局部线结构约束的 FCM 聚类视 网膜血管分割[J]. 光学学报, 2020, 40(9): 0910001.
 JIA H, ZHENG C J, LI C B, et al. Retinal blood vessel segmentation based on fuzzy C-means clustering according to the local line structural constraints[J]. Acta Optica Sinica, 2020, 40(9): 0910001

(in Chinese).

 [2] 周阳. 基于匹配滤波引导局部特征空间仿射传播聚类的视网膜 血管分割[J]. 信息通信, 2019, 32(7): 35-39.
 ZHOU Y. Retinal vascular segmentation based on affine propagation clustering in local feature space guided by matched filtering[J].

Information & Communications, 2019, 32(7): 35-39(in Chinese).

- [3] DASH J, BHOI N. Retinal blood vessel extraction using morphological operators and Kirsch's template[C]//Proceedings of the Soft Computing and Signal Processing. Berlin: Springer, 2019: 603-611.
- [4] 汪维华, 张景中, 吴文渊. 改进的形态学与 Otsu 相结合的视网膜 血管分割[J]. 计算机应用研究, 2019, 36(7): 2228-2231.
 WANG W H, ZHANG J Z, WU W Y. New approach to segment retinal vessel using morphology and Otsu[J]. Application Research of Computers, 2019, 36(7): 2228-2231(in Chinese).
- [5] LI X, JIANG Y C, LI M L, et al. Lightweight attention convolutional neural network for retinal vessel image segmentation[J]. IEEE Transactions on Industrial Informatics, 2021, 17(3): 1958-1967.
- [6] WANG C, ZHAO Z Y, REN Q Q, et al. Dense U-Net based on patchbased learning for retinal vessel segmentation[J]. Entropy, 2019, 21(2): 168.
- [7] YUE K J, ZOU B J, CHEN Z L, et al. Retinal vessel segmentation using dense U-Net with multiscale inputs[J]. Journal of Medical Imaging, 2019, 6(3): 034004.
- [8] 胡扬涛, 裴洋, 林川, 等. 空洞残差 U 型网络用于视网膜血管分割
 [J]. 计算机工程与应用, 2021, 57(7): 185-191.
 HU Y T, PEI Y, LIN C, et al. Atrous residual U-Net for retinal vessel segmentation[J]. Computer Engineering and Applications, 2021, 57(7): 185-191(in Chinese).
- [9] 易三莉,陈建亭,贺建峰. 基于多路径输入和多尺度特征融合的 视网膜血管分割[J]. 光电子·激光, 2021, 32(7): 735-741. YI S L, CHEN J T, HE J F. Retinal vascular segmentation based on multi-scale input and multi-scale feature fusion[J]. Journal of Optoelectronics·Laser, 2021, 32(7): 735-741(in Chinese).
- [10] HUANG G, LIUZ, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 2261-2269.
- [11] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision. Berlin: Springer, 2018: 3-19.
- [12] 姜大光,李明鸣,陈羽中,等. 骨架图引导的级联视网膜血管分割 网络[J]. 工程科学学报, 2021, 43(9): 1244-1252.
 JIANG D G, LI M M, CHEN Y Z, et al. Cascaded retinal vessel segmentation network guided by a skeleton map[J]. Chinese Journal of Engineering, 2021, 43(9): 1244-1252(in Chinese).

- [13] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation[C]//Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention. Berlin: Springer, 2015: 234-241.
- [14] OKTAY O, SCHLEMPER J, LE FOLGOC L, et al. Attention U-Net: Learning where to look for the pancreas[EB/OL]. (2018-05-20)[2022-05-01]. http://arxiv.org/abs/1804.03999.
- [15] 易三莉,陈建亭,贺建峰. ASR-UNet: 一种基于注意力机制改进的视网膜血管分割算法[J]. 山东大学学报 (理学版), 2021, 56(9):
 13-20.
 YI S L, CHEN J T, HE J F. ASR-UNet: An improved retinal ve-

ssels segmentation algorithm based on attention mechanism[J]. Journal of Shandong University (Natural Science), 2021, 56(9): 13-20 (in Chinese).

- [16] YAN Z Q, YANG X, CHENG K T. Joint segment-level and pixelwise losses for deep learning based retinal vessel segmentation [J]. IEEE Transactions on Biomedical Engineering, 2018, 65(9): 1912-1923.
- [17] LIN Y, ZHANG H G, HU G. Automatic retinal vessel segmentation via deeply supervised and smoothly regularized network[J]. IEEE Access, 2018, 7: 57717-57724.
- [18] YAN Z Q, YANG X, CHENG K T. A three-stage deep learning model for accurate retinal vessel segmentation[J]. IEEE Journal of Biomedical and Health Informatics, 2019, 23(4): 1427-1436.
- [19] SANTOS J C M D, CARRIJO G A, CARDOSO C D F D S, et al. Fundus image quality enhancement for blood vessel detection via a neural network using CLAHE and Wiener filter[J]. Research on Biomedical Engineering, 2020, 36(2): 107-119.
- [20] LV Y, MA H, LI J N, et al. Attention guided U-Net with atrous convolution for accurate retinal vessels segmentation[J]. IEEE Access, 2020, 8: 32826-32839.
- [21] ZHANG J W, ZHANG Y C, XU X W. Pyramid U-Net for retinal vessel segmentation[C]//Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway: IEEE Press, 2021: 1125-1129.
- [22] TCHINDA B S, TCHIOTSOP D, NOUBOM M, et al. Retinal blood vessels segmentation using classical edge detection filters and the neural network[J]. Informatics in Medicine Unlocked, 2021, 23: 100521.
- [23] YANG L, WANG H X, ZENG Q S, et al. A hybrid deep segmentation network for fundus vessels via deep-learning framework[J]. Neurocomputing, 2021, 448: 168-178.
- [24] KHAN T M, KHAN M A U, REHMAN N U, et al. Width-wise vessel bifurcation for improved retinal vessel segmentation[J]. Biomedical Signal Processing and Control, 2022, 71: 103169.

Cross-level fusion gated adaptive network for retinal vessel segmentation

LIANG Liming^{*}, YU Jie, CHEN Xin, LEI Kun, ZHOU Longsong

(School of Electrical Engineering and Automation, Jiangxi University of Science and Technology, Ganzhou 341000, China)

Abstract: To address the insufficient shallow feature extraction of most existing algorithms, which results in noise, blurred vascular boundary and capillary under segmentation, a cross-level fusion gated adaptive network is proposed. Firstly, shallow feature information is fully extracted by the dense gated channel transformation module in the network with promotion of competition or cooperation of channels to avoid the loss of shallow coarse-grained feature information. Secondly, cross-dimensional interaction information of each layer is captured by cross-level fusion module to effectively aggregate multi-scale context features. Thirdly, dual adaptive feature fusion method is used to guide the feature fusion of adjacent layers effectively and suppress noise. The validation was performed on public data sets DRIVE, CHASEDB1 and STARE, and the accuracy rates were 0.9652, 0.9668 and 0.9695 respectively; the F_1 values were 0.8544, 0.8152 and 0.8412 respectively. The results show that the proposed network is at a high level in many indexes, and is superior to the existing advanced algorithms.

Keywords: retinal vessel segmentation; dense gated channel transformation; cross-level fusion module; dual adaptive feature fusion; triple attention module

URL: link.cnki.net/urlid/11.2625.V.20220823.0958.002

Received: 2022-05-21; Accepted: 2022-08-05; Published Online: 2022-08-23 11:03

Foundation items: National Natural Science Foundation of China (51365017,61463018);General Project of Natural Science Foundation of Jiangxi, China (20192BAB205084); Jiangxi Province Graduate Innovation Special Fund (YC2021-S585); Science and Technology Project of Jiangxi Provincial Department of Education (GJJ170491,GJJ2200848)

^{*} Corresponding author. E-mail: 9119890012@ jxust.edu.cn

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

DOI: 10.13700/j.bh.1001-5965.2022.0459

基于时频分析的机场跑道平整度评价方法

齐麟*,杨帅,解镇州,金天昱,黄信

(中国民航大学交通科学与工程学院,天津 300300)

摘 要:机场跑道平整度是道面评价的重要内容,目前常用的平整度评价方法无法确定不 平整在跑道上的分布情况,也无法针对不同机型飞机的滑跑安全性进行分类评价。基于此,提出了 一种基于时频分析的跑道平整度评价方法。基于飞机三自由度动力模型,采用参数化分析方法对飞 机以不同速度在具有不同振幅、不同波长的不平整跑道上滑跑时飞机重心处竖向加速度均方根分布 特征进行统计分析;反算飞机重心处竖向加速度均方根为 0.25g 和 0.4g 时对应的跑道不平整信号的 振幅、波长与滑跑速度,建立基于跑道不平整信号频域信息与飞机滑跑速度的飞机滑跑安全振动判 断曲面;采用 S 变换方法对机场跑道高程进行时频分析,得到跑道不同位置处不平整信号的波长与 振幅分布曲面;分析飞机在跑道上不同位置的滑跑速度,以坐标位置代替滑跑速度,形成基于跑道 不平整信号频域信息与跑道位置的飞机滑跑安全振动判断曲面;与 S 变换获得的曲面叠加,得到基 于时频分析的跑道平整度评价曲面。与现有评价方法相比,所提方法可以对跑道具体位置的平整度 情况进行评价,并能够根据不同的机型对跑道平整度分级评价。

关键词:平整度;时频分析;机场跑道;评价方法;安全振动判断准则

中图分类号: V351.11

文献标志码: A 文章编号: 1001-5965(2024)04-1110-07

跑道的平整度是机场道面管理的重要内容,飞 机起降安全受到道面状况显著影响^[1]。Kanazawa等^[2] 使用飞机模拟器在构造的跑道上进行滑行模拟,建 立了机场跑道平整度与飞行员在振动下的感受二 者之间的关系;Calautti等^[3]通过研究提出了评价机 场跑道平整度状况的指标。中国机场跑道的平整 度评价指标主要采用国际平整度指数 IRI^[4-6]。IRI 是基于 1/4 车辆模型单位行驶距离车辆簧载质量和 非簧载质量的累积相对位移提出的公路平整度评 价指标^[7-8]。Múčka^[9]分析了机场跑道的实测数据, 对直尺指数与刚性和柔性道面 IRI 之间的相互关系 进行了分析研究;程国勇等^[10]认为 IRI 与飞机的尺寸 不协调,根据 Chen 和 Chou^[11]的研究,使用 APRas 进行分析,对比飞机在滑跑过程中的敏感波长和 IRI 的敏感波长,显示二者具有较大的区别。中国 引入了波音平整度指数 BBI 作为机场跑道平整度 评价指标。BBI 由美国波音公司提出,在美国大部 分机场得到应用^[12]。Emery 等^[13] 研究表明, BBI 的 评级结果太保守; 而且 BBI 不能准确描述单个波长 粗糙度及多个波长粗糙度^[14]。

现有机场跑道平整度评价方法不能考虑跑道 连续隆起的情况,也没有考虑飞机的滑跑速度,不 能得到跑道具体位置的不平整分布情况。当飞机 滑跑速度较高时,采用 IRI 评价跑道平整度是不合 理的^[15]。此外,不同机型的飞机差异性较大,目前的 评价方法并未考虑机型对跑道平整度评价的影响^[16]。

本文提出一种基于时频分析的跑道平整度评价方法。基于飞机三自由度动力模型,以跑道不平整激励波长、振幅及飞机滑跑速度为3个变化参数,建立飞机滑跑安全振动判断曲面。实测跑道进

收稿日期:2022-06-08;录用日期:2022-08-19;网络出版时间:2022-09-1413:55 网络出版地址:link.enki.net/urlid/11.2625.V.20220913.2006.015

基金项目: 国家重点研发计划 (2021YFB2600500); 中央高校基本科研业务费专项资金 (3122019104)

^{*}通信作者. E-mail: qilin1208@vip.163.com

引用格式: 齐麟, 杨帅, 解镇州, 等. 基于时频分析的机场跑道平整度评价方法 [J]. 北京航空航天大学学报, 2024, 50 (4): 1110-1116. QIL, YANG S, XIE Z Z, et al. Evaluation method for roughness of airport runway based on joint time-frequency analysis [J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50 (4): 1110-1116 (in Chinese).

行 S 变换,得到随跑道位置的波长与振幅分布曲面,与飞机滑跑安全振动判断曲面叠加,得到基于时频分析的机场跑道平整度评价方法,能够对跑道 具体位置进行平整度评价,并根据机型对跑道平整 度进行分级评价。

1 跑道不平整高程模拟

谐波叠加算法是一种离散化数值模拟道面的 方法,其基本思想是采用大量具有随机相位的正弦 或余弦之和来表示随机过程或随机现象^[17]。

对机场跑道平整度准确评价,需要通过真实的 道面信息计算飞机振动响应,并通过控制跑道纵断 面波长、振幅来构建跑道。在跑道上任取一点*l*_i, 都有一个随机变量*x*(*l*_i),其平均值为0。采用谐波 叠加算法构造跑道不平整度*x*(*l*)的三角级数表达 式为

$$x(l) = \sum_{i=1}^{N} a_i \sin\left(2\pi l n_i + \theta_i\right) \tag{1}$$

式中: l为跑道水平距离, m; n_i 为第 i个区间中心频率; a_i 为高斯独立随机变量, 平均值为 0, 标准差为 σ_i ; θ_i 为[0, 2π]范围内的均布独立随机变量。

首先,控制波长 λ 的大小,对于空间频率有 λ = 1/n。将空间频率范围(n_1,n_2) N等分,如下:

$$\Delta n = \left(\frac{1}{\lambda_2} - \frac{1}{\lambda_1}\right) / N \tag{2}$$

$$n_i = n_1 + \Delta n \left(i - \frac{1}{2} \right)$$
 $i = 1, 2, \cdots, N$ (3)

式中: Δn为N等分后区间宽度。

其次,控制振幅*a*_i的大小,将空间频率*N*等分后,给每个区间赋予对应的振幅,将所需振幅范围 (*a*₁,*a*₂)进行取值,*a*_i取(*a*₁/*N*, *a*₂/*N*)中的随机值,再 进行谐波叠加。

$$x(l) = \sum_{i=1}^{N} a_i \sin\left(2\pi l \left(\frac{1}{\lambda_1} + \left(i - \frac{1}{2}\right) \left(\frac{1}{\lambda_2} - \frac{1}{\lambda_1}\right) / N\right) + \theta_i\right)$$
(4)

式(4)可构建具有不同参数的道面断面高程曲 线。由采样定理 $\Delta l \leq 1/10n_2$, $\Delta l = v\Delta t$, v 为飞机在通 过空间频率 n 的不平整跑道时滑跑速度。因此, 时 间的最低采样频率应满足:

$$\Delta t \leqslant \frac{1}{10\nu n_2} \tag{5}$$

2 飞机滑跑安全振动判断准则

飞机的振动在水平方向表现较少,且相对于其 余方向基本可以忽略,因此,本文忽略侧倾转动。 假设飞机只存在竖向运动和俯仰转动,三自由度飞 机模型如图1所示^[18]。图中: *M*_p和 *m* 分别为簧载 质量和非簧载质量; *K*、*k* 分别为对应的刚度系数; *C*、*c* 分别为对应的阻尼系数; 下标 f、l、r 分别代表 前、左后、右后起落架。

本文选取某C类主流客机进行研究,飞机参数 如表1所示。



图 1 飞机动力学模型^[18]

Fig. 1 Aircraft dynamics model^[18]

表1 C类主流客机参数取值

Table 1 Parameter values of class C mainstream

passenger aircraft

参数	数值
$M_{\rm p}/{ m kg}$	59 033
$m_{\rm f}/{ m kg}$	390
$m_{\rm l}/{ m kg}$	888
$m_{ m r}/{ m kg}$	888
$K_{\rm f}/({\rm kN}\cdot{\rm m}^{-1})$	110
$K_{\rm l}/({\rm kN}\cdot{\rm m}^{-1})$	614
$K_{\rm r}/({\rm kN}\cdot{\rm m}^{-1})$	614
$C_{\rm f'}({\rm kN}\cdot({\rm m}\cdot{\rm s}^{-1})^{-1})$	143
$C_{l}/(\mathrm{kN}\cdot(\mathrm{m}\cdot\mathrm{s}^{-1})^{-1})$	625
$C_r/(\mathrm{kN}\cdot(\mathrm{m}\cdot\mathrm{s}^{-1})^{-1})$	625
$k_{\rm f}/({\rm MN}\cdot{ m m}^{-1})$	4
$k_{\rm l}/({\rm MN}\cdot{ m m}^{-1})$	4
$k_{\rm r}/({\rm MN}\cdot{ m m}^{-1})$	4
$c_{\rm f'}({\rm kN}\cdot{\rm (m}\cdot{\rm s}^{-1})^{-1})$	4
$c_{\rm l}/({\rm kN}\cdot({\rm m}\cdot{\rm s}^{-1})^{-1})$	4
$c_r/(kN \cdot (m \cdot s^{-1})^{-1})$	4

采用参数化分析方法,对飞机在不同振幅、波 长、滑跑速度等3个参数作用下,以某C类主流客 机为例,输入不同频率、振幅的跑道高程信息,将重 心处竖向加速度均方根输出。根据Debord^[19]的研 究结果显示,能够较为明显地引起飞机动力学响应 的波长处在小于120m的范围内。因此,设跑道波

2024年

长的变化范围为 0~120 m, 步长为 5 m, 跑道振幅 的变化范围为 0~0.1 m, 步长为 0.005 m。飞机在跑 道上滑行, 当速度接近起飞速度时, 取其代表速度 200 km/h(即 55.56 m/s)。飞机需要由滑行道进入跑 道, 此时滑跑速度较低, 取代表速度为 40 km/h (即 11.11 m/s)。在 11~56 m/s 速度下, 对飞机动力 学模型进行参数化分析。

本文定义飞机重心处竖向加速度均方根A_r 作为飞机在跑道不平整激励下的振动响应^[20],如下:

$$A_{\rm r} = \sqrt{\frac{1}{N} \sum_{j=1}^{N} A_j^2}$$
 (6)

式中: A_j为第 j个样本点飞机重心处的竖向加速度; N为总样本点数。

设飞机滑跑速度为 30 m/s, 在同一波长下, 飞机 重心处竖向加速度均方根在不同振幅作用下的分 布如图 2 所示。



amplitude at the same wavelength

由图 2 可以看出, 飞机模型以 30 m/s 在不平整 跑道上模拟滑跑时, 波长不变, 其振幅与重心处竖 向加速度均方根成正比, 且 10~15 m 波长对应的 振幅较其他波长对应的振幅产生的加速度均方根 更大。

在 0.025~0.03 m 不平整的振幅激励下,飞机重 心处竖向加速度均方根在不同波长作用下的分布 如图 3 所示。可知,振幅不变时,波长和飞机重心 处竖向加速度均方根无明显线性关系,故不可以函





数关系来直接表示波长和振幅。波长不变时,振幅 越大,飞机重心处竖向加速度均方根的值越大。可 通过在固定的波长范围内找到重心处竖向加速度 均方根不同安全等级对应的振幅,绘制基于频域特 征的飞机滑跑安全振动判断曲线。

机场新建的跑道,其竖向加速度最大不应超过 0.25g,旧跑道则不应超过 0.4g(波音公司制定的疲 劳标准),与管理规范中对机场跑道平整度状况分 为好、中、差 3 个等级相对应。飞机重心处竖向加 速度均方根作为不平整激励下振动响应,对跑道平 整度状况分级有如下定义。评价好:0~0.25g,评价 中:0.25g~0.4g,评价差:>0.4g。因此,根据机场跑 道平整度分级标准,找到同一波长区间内重心处竖 向加速度均方根为 0.25g 和 0.4g 所对应的振幅值, 重心处竖向加速度均方根为 0~0.25g 范围内评价 为好,重心处竖向加速度均方根为 0.25g~0.4g 范 围内评价为中,重心处竖向加速度均方根为 0.4g 以 上评价为差。基于频域特征的飞机滑跑安全振动 判断曲线如图 4 所示。



图 4 基于频域特征的飞机滑跑安全振动判断曲线 Fig. 4 Aircraft taxiing safety vibration judgment curves based on frequency domain features

当飞机以其他速度在跑道上滑跑时,同样需要 利用飞机动力学模型,在不同参数的跑道上,以不 同速度模拟滑跑并拟合出不同速度下基于频域特 征的飞机滑跑安全振动判断曲线。因此,应在 11~56 m/s速度下,建立基于跑道不平整信号频域 信息与飞机滑跑速度的飞机滑跑安全振动判断曲 面,如图 5 所示。

如图 5 所示,该曲面可以作为 C 类主流机型的 飞机滑跑安全振动判断准则。在跑道不平整激励 下,飞机的振动响应位于红色曲面以下,即重心处 竖向加速度均方根为 0~0.25g 范围内,评价为 "好";位于红色曲面和蓝色曲面之间,即重心处竖 向加速度均方根为 0.25g~0.4g 范围内,评价为 "中";位于蓝色曲面以上,即重心处竖向加速度均 方根为 0.4g 以上,评价为 "差"。该评价曲面考虑 了飞机在不同位置的滑跑速度不同,并且能够对不 同机型进行分级评价。对于其他机型的滑跑安全 振动判断准则,可以采用相同的方法根据机型建立 飞机动力学模型,进行参数化分析,得到对应的飞 机滑跑安全振动判断准则。



图 5 飞机滑跑安全振动判断曲面 Fig. 5 Judgment surface of aircraft taxiing safety vibration

3 机场跑道平整度评价方法

3.1 基于 S 变换的跑道高程时频分析

S 变换能根据频率的变化自适应地调整分析时 宽,并提供直观的时间频率特征,无需选择窗口函 数域分析尺度,对于频率的空间分布,能够有较好 的反映。连续空间信号一维变换如式(7)所示:

$$S(\tau,n) = \int_{-\infty}^{+\infty} h(x) \frac{|n|}{\sqrt{2\pi}} e^{-\frac{(\tau-x)^2 n^2}{2}} e^{-i2\pi nx} dx \qquad (7)$$

式中: $S(\tau,n)$ 为经过变换后的时频谱矩阵, τ 为时间, 控制窗口函数在时间轴上位置, n为频率; h(x)为分 析信号; $|n|/\sqrt{2\pi}e^{-\frac{(\tau-x)^2n^2}{2}}$ 为高斯窗函数。

跑道高程信息作为离散空间,对其进行S变换

时,需要采用离散形式。傅里叶变换频移定理应用于信号*h*(*x*)的S变换如式(8)所示:

$$S(\tau,n) = \int_{-\infty}^{+\infty} h(\alpha+n) \mathrm{e}^{-\frac{2\pi^2 \alpha^2}{n^2}} \mathrm{e}^{\mathrm{i}2\pi\alpha\tau} \mathrm{d}\alpha \qquad n \neq 0 \qquad (8)$$

式中: $h(mdx)(m = 0, 1, 2, \dots, N - 1)$ 为对应h(x)的离散 空间信号; dx为采样点的间距。

将频率转化为波长进行波长-振幅-位置三维时 频分析,图 6 为 3 000 m 跑道的高程曲线。

利用实测跑道进行 S 变换,得到跑道随位置的 波长与振幅分布曲面,如图 7 所示。



图 6 实测 3000 m 跑道高程曲线





Fig. 7 Wavelength and amplitude distribution surfaces at different plane positions

3.2 机场跑道平整度评价方法

飞机实际滑跑过程中不会以恒定速度匀速滑 行,在变速情况下评价跑道尤为重要,将飞机在3000m 的跑道上以11 m/s 匀加速至56 m/s 的速度滑跑。 根据第2节获得的飞机滑跑安全振动判断准则,分 析飞机在跑道上不同位置滑跑速度,以坐标位置代 替滑跑速度,形成基于跑道不平整信号频域信息和 跑道位置的飞机滑跑安全振动判断曲面,与S变换 得到的跑道不同位置波长与振幅分布曲面叠加,即 可得到机场跑道平整度的评价方法。采用本文方 法,考虑不同位置飞机的滑跑速度不同,并且能够 对跑道的具体位置平整度的状况进行评价,进而为 机场跑道的管理者提供准确的平整度信息,如图 8 所示。

图 8 中, 青色曲面表示经 S 变换后的跑道三维

2024 年

信息, 红色和蓝色曲面表示飞机滑跑安全振动判断 曲面。位于红色曲面以下, 即重心处竖向加速度均 方根为 0~0.25g 范围内, 评价为"好", 此安全区间 内飞机振动响应较小, 飞机运行安全; 位于红色曲 面和蓝色曲面之间, 即重心处竖向加速度均方根为 0.25g~0.4g 范围内, 评价为"中", 此范围安全区间 会给飞机上的乘客带来不适及驾驶舱内仪表的严 重干扰, 必须立即修复跑道, 但无需关闭受影响的 跑道; 位于蓝色曲面以上, 即重心处竖向加速度均 方根为 0.4g 以上, 评价为"差", 此阶段严重影响飞 机滑行, 应立即关闭受影响的跑道, 必须进行维修 以使跑道恢复到正常水平。由图 8 可知, 该实测跑 道应在1500~1600m、2200~2400m、2700~2800m 位置处对跑道进行修复。

根据 FAA 实测的跑道高程,将跑道高程输入 模拟飞行器模拟 C 类主流机型,采用本文方法建立 跑道平整度评价曲面,如图 9 所示。

测试中,飞行员根据振动的颠簸程度对跑道平 整度进行打分评价,如表2所示^[21]。



图 8 基于时频分析的跑道平整度评价曲面 Fig. 8 Evaluation surfaces of runway flatness based on







	Table 2	Pi	ilot rating ^[21]	
飞行员编号	打分		飞行员编号	打分
1	5.4		18	4.5
2	3		19	2.2
3	5.5		20	6.6
4	6.5		21	4.5
5	3		22	3.3
6	2.5		23	3.3
7	5		24	4.8
8	5.3		25	5.6
9	6.4		26	4.8
10	2.6		27	6.3
11	2.9		28	3
12	1.2		29	2.6
13	5.5		30	4.8
14	4.4		31	2.7
15	4.6		32	3.3
16	6.3		33	2.9
17	6.9			

飞行员评分^[21]

表 2

评分 0 为无法通行, 0~1 为很差, 1~3 为差, 3~5 为一般, 5~7 为好, 7~9 为很好, 9~10 为完 美。飞行员评分结果平均分为 4.31, 表示跑道平整 度状况一般, 对应的跑道平整度评价曲面也有部分 位于蓝色曲面以上, 即重心处竖向加速度均方根为 0.4g 以上, 此部分评价为"差", 验证了本文方法的 可靠性。

本文方法通过 S 变换对跑道进行时频分析,得 出跑道的频域信息,再与飞机滑跑安全振动判断曲 面叠加,得出了基于时频分析的跑道平整度评价方 法。该方法能够判断跑道具体位置的平整度情况, 并且能够根据机型的不同对跑道平整度进行分级 评价。对于其他机型,可以建立新的飞机动力学模 型,参数化分析后得到该机型的滑跑安全振动判断 准则,再与经过 S 变换得到的跑道不同位置波长和 振幅分布曲面叠加,得到适用于该机型的跑道平整 度的评价方法。

4 结 论

1) 基于飞机三自由度动力模型,采用参数化分 析方法,对飞机以不同速度在具有不同振幅、不同 波长的不平整跑道上滑跑时飞机重心处竖向加速 度均方根分布特征进行统计分析。反算飞机重心 处竖向加速度均方根为 0.25g 与 0.4g 时,对应的跑 道不平整信号的振幅、波长与滑跑速度,建立基于 跑道不平整信号频域信息与飞机滑跑速度的飞机 滑跑安全振动判断准则。 2)采用S变换方法对机场跑道高程进行时频 分析,得到跑道不同位置处不平整信号的波长与振 幅分布曲面。

3)分析飞机在跑道上不同位置滑跑速度,以坐标位置代替滑跑速度,形成基于跑道不平整信号频域信息与跑道位置的飞机滑跑安全振动判断曲面,与跑道不同位置处不平整信号的波长与振幅分布曲面叠加,即得到了基于时频分析的跑道平整度评价曲面,可以对跑道具体位置的平整度情况进行评价,并能够根据不同的机型对跑道平整度进行分级评价。

参考文献(References)

- [1] 蔡靖,李岳,宗一鸣,等. 湿滑跑道飞机着陆轮胎-水膜-道面相互 作用[J]. 北京航空航天大学学报, 2017, 43(12): 2382-2391.
 CAI J, LI Y, ZONG Y M, et al. Aircraft tire-water film-pavement interaction on wet pavement in landing[J]. Journal of Beijing University of Aeronautics and Astronautics, 2017, 43(12): 2382-2391 (in Chinese).
- [2] KANAZAWA H, SU K, NOGUCHI T, et al. Evaluation of airport runway pavement based on pilots' subjective judgement[J]. International Journal of Pavement Engineering, 2010, 11(3): 189-195.
- [3] CALAUTTI J, MURRELL S D, GERARDI T. Roughness assessment in pavement management at New York metropolitan area airports[C]//Proceedings of the FAA Worldwide Airport Technology Transfer Conference. Washington, D. C. : FAA, 2004: 1-13.
- 【4】 凌建明,刘诗福, 袁捷, 等. 机场道面平整度评价指标的相关性分析[J]. 公路交通科技, 2020, 37(3): 17-23.
 LING J M, LIU S F, YUAN J, et al. Analysis on correlation of evaluation indicators of airport pavement roughness[J]. Journal of Highway and Transportation Research and Development, 2020, 37(3): 17-23(in Chinese).
- [5] 吴庆雄, 陈宝春, 奚灵智. 路面平整度 PSD 和 IRI 评价方法比较
 [J]. 交通运输工程学报, 2008, 8(1): 36-41.
 WU Q X, CHEN B C, XI L Z. Comparison of PSD method and IRI method for road roughness evaluation[J]. Journal of Traffic and Transportation Engineering, 2008, 8(1): 36-41(in Chinese).
- [6] 张献民,陈新春,李少波. 基于国际平整度指数 IRI 的飞机动载系数分析[J]. 南京航空航天大学学报 2016, 48(1): 136-142. ZHANG X M, CHEN X C, LI S B. Aircraft dynamic load coefficient based on international roughness index[J]. Journal of Nanjing University of Aeronautics & Astronautics, 2016, 48(1): 136-142(in Chinese).
- [7] SIVAKUMAR S, HARAN A P. Mathematical model and vibration analysis of aircraft with active landing gears[J]. Journal of Vibration and Control, 2015, 21(2): 229-245.
- [8] 董忠红, 吕彭民. 高等级路面上的车辆动载荷[J]. 长安大学学报 (自然科学版), 2010, 30(1): 95-99.
 DONG Z H, LU P M. Dynamic load of vehicle on high-class pavement[J]. Journal of Chang'an University (Natural Science Edition), 2010, 30(1): 95-99(in Chinese).

- [9] MÚČKA P. Relationship between international roughness index and straightedge index[J]. Journal of Transportation Engineering, 2012, 138(9): 1099-1112.
- [10] 程国勇, 郭稳厚, 雷亚伟. 基于飞机全尺寸模型的机场道面平整 度评价理论研究[J]. 公路工程, 2016, 41(4): 1-5. CHENG G Y, GUO W H, LEI Y W. Research on the airport pavement roughness evaluation based on full aircraft model[J]. Highway Engineering, 2016, 41(4): 1-5(in Chinese).
- [11] CHEN Y H, CHOU C P. Effects of airport pavement-profile wavelength on aircraft vertical responses[J]. Transportation Research Record: Journal of the Transportation Research Board, 2004, 1889(1): 83-93.
- [12] FAA. Guidelines and procedures for measuring airfield pavement roughness: Advisorycircular: 150/5380-9[R]. Washington, D.C.:FAA, 2009.
- [13] EMERY S, HEFER A, HORAK E. Roughness of runways and significance of appropriate specifications and measurement[C]//Proceedings of the 11th Conference of Asphalt Pavements in Southern Africa. Sun City: CAPSA Group, 2015: 16-19.
- [14] 刘春梅.考虑长轴距效应的道面平整度对飞机滑行振动的影响
 [D].天津:中国民航大学, 2020: 6-10.
 LIU C M. The influence of pavement roughness considering longwheelbase effect on taxing aircraft vibration[D]. Tianjin: Civil Aviation University of China, 2020: 6-10(in Chinese).
- [15] 凌建明,刘诗福,袁捷,等.采用 IRI 评价机场道面平整度的适用 性[J]. 交通运输工程学报, 2017, 17(1): 20-27.
 LING J M, LIU S F, YUAN J, et al. Applicability of IRI based evaluation of airport pavement roughness[J]. Journal of Traffic and Transportation Engineering, 2017, 17(1): 20-27(in Chinese).
- [16] 郭稳厚. 与机型相关的机场道面相对平整度分析理论研究[D]. 天津: 中国民航大学, 2015: 52-61.
 GUO W H. The research on the theory of pavement relatively roughness related to aircraft type[D]. Tianjin: Civil Aviation University of China, 2015: 52-61(in Chinese).
- [17] LOPRENCIPE G, ZOCCALI P. Comparison of methods for evaluating airport pavement roughness[J]. International Journal of Pavement Engineering, 2019, 20(7): 782-791.
- [18] 凌建明,刘诗福,袁捷,等.不平整激励下机场道面和公路路面平 整度评价综合分析[J].同济大学学报(自然科学版),2017,45(4): 519-526.

LING J M, LIU S F, YUAN J, et al. Comprehensive analysis of pavement roughness evaluation for airport and road with different roughness excitation[J]. Journal of Tongji University (Natural Science), 2017, 45(4): 519-526(in Chinese).

- [19] DEBORD K J. Runway roughness measurement, quantification and application--The Boeing method: D6-81746[R]. Chicago: Boeing, 1995.
- [20] ZHANG H L, YANG W Q. Evaluation method of pavement roughness based on human-vehicle-road interaction[C]//Proceedings of the Integrated Transportation Systems: Green, Intelligent, Reliable. Reston: American Society of Civil Engineers, 2010: 3541-3551.
- [21] FAA. Boeing 737-800 final surface roughness study data collection: DOT/FAA/TC-18/8[R]. Washington, D. C. : FAA, 2017.

Evaluation method for roughness of airport runway based on joint time-frequency analysis

QI Lin*, YANG Shuai, XIE Zhenzhou, JIN Tianyu, HUANG Xin

(School of Transportation Science and Engineering, Civil Aviation University of China, Tianjin 300300, China)

Abstract: Airport runway roughness is an important content of airport pavement evaluation. At present, the commonly used roughness evaluation methods can't determine the distribution of unevenness on the runway, and can't classify and evaluate the taxiing safety of different types of aircraft. In this paper, a runway roughness evaluation method based on joint time-frequency analysis is proposed. The parametric analysis method is used to statistically analyze the root-mean-square distribution characteristics of vertical acceleration at the aircraft's center of gravity when the aircraft runs at different speeds on uneven runways with different amplitudes and wavelengths. This analysis is based on the three-degree-of-freedom dynamic model of the aircraft. The amplitude, wavelength of the uneven signal of the runway, and taxiing speed corresponding to the vertical acceleration root-mean-square of the center of gravity of the aircraft as 0.25g and 0.4g is calculated back, and the judgment surface of the aircraft taxiing safety vibration based on the frequency domain information of the uneven signal of therunway is established. In order to determine the wavelength and amplitude distribution surfaces of unequal signals at various points along the runway, a joint time-frequency analysis is conducted using the S transform method on the elevation of the airport runway. This paper analyzes the taxiing speed of the aircraft at different positions on the runway, and replaces the taxiing speed with the coordinate position, forming a taxiing safety vibration surface based on the frequency domain information of the uneven signal and the position of the runway. The runway roughness evaluation surface based on joint timefrequency analysis is obtained by superimposing with the surface obtained by the S transform. In contrast to other evaluation techniques, it is capable of assessing the degree of roughness at a given runway location and rating the roughness of the runway based on the kind of aircraft.

Keywords: roughness; time-frequency analysis; airport runway; evaluation method; safety vibration judgment criterion

Received: 2022-06-08; Accepted: 2022-08-19; Published Online: 2022-09-14 13:55 URL: link.cnki.net/urlid/11.2625.V.20220913.2006.015

Foundation items: National Key R & D Program of China (2021YFB2600500); The Fundamental Research Funds for the Central Universities (3122019104)

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn **DOI**: 10.13700/j.bh.1001-5965.2022.0473

基于 ECSDNN 的航空安全事件风险等级预测

冯霞^{1,2},桑潇^{1,2},左海超^{1,2,*}

(1. 中国民航大学 计算机科学与技术学院,天津 300300; 2. 民航智慧机场理论与系统重点实验室,天津 300300)

摘 要:航空安全事件风险等级预测是主动风险管理的重要手段。考虑海量航空安全事件数据呈现的高维复杂、类不平衡等特性,提出一种基于集成代价敏感深度神经网络(ECSDNN)的航空安全事件风险等级预测方法。采用分类型属性嵌入特征编码和数值型属性拼接的方法实现航空安全事件数据的特征表示;综合考虑错分比例和固定代价设计代价敏感矩阵和代价敏感损失函数,构建基于代价敏感深度神经网络(CSDNN)的基分类器模型;采用硬投票方法,集成多个参数不同、性能各异的基分类器,构建航空安全事件风险等级预测模型。在航空安全事件报告系统(ASRS)数据集上的实验结果表明:相比基准算法,所提 ECSDNN 模型的预测准确率提升了4.51%;相比单个 CSDNN 基分类器,所提 ECSDNN 模型的预测准确率提升了3.17%。验证了基于 ECSDNN 的航空安全事件风险等级预测方法的有效性。

关 键 词: 航空安全; 风险等级预测; 嵌入特征编码; 代价敏感; 深度神经网络; 集成学习 中图分类号: V328; TP391

文献标志码: A 文章编号: 1001-5965(2024)04-1117-12

随着航空出行需求的快速增长,空域中飞机数 量必将大幅增加,航空安全态势将更趋复杂多变, 给航空安全风险管理带来了极大的技术挑战。基 于历史航空安全事件数据构建事件风险等级预测 模型,量化分析航空安全事件的风险严重程度,有 助于在发生具体航空安全事件时,根据事件关键属 性提前预判事件可能的风险程度,识别未来安全态 势演变,是风险管理人员量化风险、设置优先级并 实施资源配置的重要决策依据,也是主动安全风险 管理的重要技术手段。

长期以来, 航空安全事件风险严重程度分析是 一件复杂困难但非常重要的工作, 主要依靠领域专 家人工完成, 专业素养要求高且繁琐耗力。随着计 算技术的进步, 航空界开始采用专门的信息系统记 录已经发生的各类型航空安全事件, 并逐步积累了 大量的航空安全事件报告数据, 其中, 包含了专业 人员对已发生事件的诊断分析。同时, 人工智能、

数据挖掘等自动分析技术不断发展,使得研究人员 基于现有航空安全事件数据集开展安全风险自动 识别研究成为可能。国外围绕航空安全事件自动 分析的研究主要集中在致因因素分析[1-4]、主题查 找[5-7]、异常检测分析[8-10]等。国内航空领域相关研 究较少,主要采用关联规则^[11]、内容分析^[12]、文本 聚类[13]等方法对航空安全事件进行挖掘分析。其 他交通领域的相关研究更多集中在基于历史发生 的事件、事故数据预测未来事件、事故风险发生的 概率[14-16],基于运行状态、道路属性、天气等因素研 判安全风险[17],以及事件或事故发生时主要致因因 素研判[18]等。由于事件风险预测本身的挑战及事 件数据集呈现的风险类别不平衡性,有关航空安全 事件风险自动预测的研究还很有限, Srinivasan 等^[19] 基于美国国家运输安全委员会(National Transportation Safety Board, NTSB)航空安全报告数据集,选取 描述事件概要的非结构化属性.采用长短期记忆

收稿日期: 2022-06-11;录用日期: 2022-09-11;网络出版时间: 2022-09-29 11:27 网络出版地址: link.cnki.net/urlid/11.2625.V.20220928.1033.001

基金项目:国家自然科学基金民航联合基金重点项目 (U2333206);国家重点研发计划 (2021YFF0603902);中央高校基本科研业务费专项资金 (3122021063);天津市研究生科研创新项目 (2021YJSO2S14)

*通信作者. E-mail: hczuo@cauc.edu.cn

引用格式: 冯霞, 桑潇, 左海超. 基于 ECSDNN 的航空安全事件风险等级预测 [J]. 北京航空航天大学学报, 2024, 50(4): 1117-1128. FENG X, SANG X, ZUO H C. Prediction of aviation safety event risk level based on ensemble cost-sensitive deep neural network [J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50(4): 1117-1128 (in Chinese).

(long short-term memory, LSTM)方法, 分别构建多 个二分类模型,识别事件类型、事件损害程度、事 件有无死亡等,但未直接进行事件风险类别识别。 Alkhamisi 和 Mehmood^[20] 基于航空安全事件报告的 6 类结构 化 属 性, 采 用 循 环 神 经 网 络 (recurrent neural network, RNN)构建了事件风险等级预测模 型,将风险等级划分为高、中、低3类,粒度较粗。 倪晓梅等^[21]借助卷积神经网络(convolutional neural network, CNN)探索了航空安全事件报告中非结构 化属性事件概要与事件风险等级之间的关系,并构 建了风险等级预测模型,但未利用事件报告结构化 属性中蕴含的丰富信息。Zhang 和 Mahadevan^[22] 选 择航空安全事件报告中29个典型结构化属性,采 用深度神经网络(deep neural network, DNN)构建事 件风险程度预测模型,但忽视了飞行高度、相对位 置等和风险类型高度关联的重要属性,仅通过数据 采样解决数据集存在的类不平衡问题,模型预测性 能有限。总体来讲,现有基于航空安全事件报告数 据集进行航空安全风险预测的研究仍处于前期探 索阶段。

DNN 是近年来人工智能领域的研究热点之一^[23],在处理海量高维数据方面具有学习能力强、 覆盖范围广、可移植性好等优势,在计算机视觉、 自然语言理解、高维稀疏数据挖掘等领域都有成功 应用^[24]。但 DNN 方法在解决数据不平衡问题时性 能有限,代价敏感学习通过为不同类别样本提供不 同权重改善机器学习模型性能,可有效缓解分类任 务中的数据不平衡问题^[25]。基于此,本文结合 DNN 和代价敏感学习,研究提出一种基于集成代 价敏感深度神经网络(ensemble cost-sensitive deep neural network, ECSDNN)的航空安全事件风险等级 预测方法。

1)考虑航空安全事件报告系统(aviation safety reporting system, ASRS)航空安全事件数据分类型 属性占主导、取值多样且值间有关联等特性,提出 一种结合嵌入特征编码和拼接的事件报告特征表 示方法。

2)考虑数据类不平衡性及不同类错分敏感性 不同,提出一种综合动态更新的错分比例和固定代 价的代价敏感损失函数计算方法,结合 DNN,构建 航空安全事件风险等级预测基分类器模型代价敏 感深度神经网络(cost-sensitive deep neural network, CSDNN)。

3)为进一步提升航空安全事件风险等级预测 性能,采用硬投票方法,集成了多个参数、超参数和 基准损失函数各不相同的基分类器,构建了航空安 全事件风险等级集成预测模型。

1 航空安全事件风险等级标注

收集记录并分析已经发生的各类型航空安全 事件,是航空界安全信息管理的主流做法。航空安 全报告系统是专门收集、分析和共享航空安全信 息,提供航空安全告警服务的航空安全信息系 统^[26]。本文选择全球最大也最有影响力的 ASRS 作 为实验数据集。

1.1 ASRS 数据集

ASRS 是自愿、保密、非惩罚性的航空安全报 告系统^[27], 创建于 20 世纪 70 年代, 接收来自飞行 员、管制员、调度员、机组人员、维修技术人员、操 作员及其他工作人员提交的航空安全事件报告, 描 述和记录了发生的不安全事件、危险状况、经验教 训等, 旨在帮助防止未来犯同样的错误。ASRS 数 据集经匿名化处理后进行全球范围数据共享。

ASRS 数据集中,每份报告记录了事件发生时间、地点、环境、飞机等多组属性值,表1为ASRS 报告的部分属性及其取值样例。可以看出,ASRS

表 1 ASRS 报告示例 Table 1 ASRS report example

属性组	属性	属性值			
मने देख	日期	202007			
印1日	当地时间	1801—2400			
	参考地点	ZMP.ARTCC			
地点	州参考	NM			
	AGL高度	5 181.6 m			
	飞行条件	VMC			
环境	天气元素/能见度	10			
	光	Daylight			
	航空交通管制	Center ZMP			
	飞机操作者	Air Carrier			
	创建模型名称	Commercial Fixed Wing			
	人员数量	2			
飞机	操作部分	Part 121			
	飞行计划	IFR			
	任务	Passenger			
	飞行阶段	Climb			
	空域	Class E ZMP			
	飞机部件	Pressurization Control System			
组成	参考	Х			
	问题	Improperly Operated			
	人员所在位置	Facility ZMP.ARTCC			
	报告组织	Government			
人物	功能模块	Enroute; Trainee			
	资质	Air Traffic Control Developmental			
	人为因素	Situational Awareness			
		ATC Issue All Types; Airspace			
	星堂	Violation All Types;			
	40.16	Deviation/Discrepancy-Procedural			
事件		Published Material/Policy			
	探测者	Person Air Traffic Control			
	场合	In-flight			
	结果	General None Reported/Taken			
	影响因素/情况	Environment-Non Weather Related;			
评估		Human Factors; Airspace Structure			
	主要问题	Human Factors			
	-				

注:信息来源于https://asrs.arc.nasa.gov/search/database.html。

报告包括多个属性组,每组属性又包含多个属性。

1.2 ASRS 航空安全事件风险等级标注

ASRS 数据集中,事件结果属性与事件风险严 重程度高度相关,其由安全专家人工分析给出,共 有 36 种不同取值,分别代表事件的不同结果,暗含 着事件的风险程度。国际民航组织 Doc9859^[28]将 事件风险等级分为高风险、中风险和低风险 3 种, 本文在此基础上,结合专家意见,借鉴文献 [22],将 风险等级划分为高风险、中高风险、中风险、中低 风险和低风险 5 种风险等级,每种事件结果对应一 个风险类别,同一风险类别可包含多种类型的事件 结果,具体划分对照如表 2 所示。

考虑 ASRS 数据集中存在一个事件有多个事件 结果的现象,本文的做法是:将事件结果对应的最 高风险类别确定为相应事件的风险类别,如表 3 所 示。表中,某事件共有 3 个事件结果,分别对应 3 种风险类别,本文标注最高风险类别 4 为该事件 风险类别。

风险等级	事件结果
低风险(0)	涉及一般警察及保安;空中交通管制部门提供了协助;飞机设备问题消失;机组人员返回登机口/机组人员FLC遵从自动化建议; 一般没有报告被采取
中低风险(1)	通用维修行动/通用航班取消延迟;通用放行拒绝飞机未被接受;机组人员请求空中交通管制协助澄清;机组人员凌驾于自动控制/机场飞机自动化凌驾于机组人员/机组人员FLC凌驾于自动控制;机组人员退出被侵入的空域/安全着陆/返回安全区域/返回 起飞机场
中风险(2)	一般工作被拒绝;机组人员重新归位;机组人员转移/执行返航错过进场;空中交通管制签发了新的许可;机组人员克服设备问题/拒绝起飞/采取规避行动
中高风险(3)	一般撤离;机组人员重新控制飞机;管制空中交通管制发出通告;机组人员紧急降落;管制空中交通管制发出警报
高风险(4)	一般紧急状态;一般身体受伤/丧失行动能;机组人员飞行停飞;航空管制分离交通管制;飞机损坏

表 3 风险等级界定

Table 3 Risk level definition

风险等级
高风险(4)
中高风险(3)
中风险(2)

1.3 ASRS 航空安全事件数据特点

ASRS 数据集积累的海量丰富事件数据为本文 研究事件风险等级预测提供了很好的数据基础,但 其固有的数据海量、属性高维、分类型属性占主导及 风险类型不平衡性等特点也带来了很多挑战,如下:

1)数据海量。ASRS数据集中,包含飞行员、 管制员等各类人员提交的发生在各个运行阶段和 不同位置的海量复杂事件报告(包括一般报告、 ATC报告和维修报告等),过去45年,ASRS数据集 已收集汇总了超过170万份安全报告^[29]。

2)属性高维。ASRS数据集中,每个事件报告 都包含多个属性组,如时间、地点、环境、飞机、人 物、事件等,每个属性组又包含多个具体属性列。

3) 分类型属性占主导。ASRS 数据集中, 分类 型属性占主导, 此类属性不仅取值离散且可能的值 也非常多, 如人物属性组下的连接故障属性, 其取 值共包括 4000 多种, 而且不同值之间存在一定程 度的关联性。

4)风险类型分布不平衡。本文从 ASRS 数据 集中选取 2010 年 1 月至 2021 年 6 月的所有事件报 告,经预处理后,按照1.2节所述方法对每个事件的 风险等级(类别)进行标注。图1给出了不同风险 等级包含的事件报告数量。可知,ASRS数据集中 的事件数据存在类不平衡性。





2 本文方法

如图 2 所示, 航空安全事件风险等级预测模型 ECSDNN 构建框架包括 3 部分: 特征工程、基分类 器和集成分类器。①特征工程。首先, 采用 1.2 节 所述方法对 ASRS 数据集中的事件数据进行风险 等级标注; 然后, 对事件属性进行预处理并遴选预




测用属性;最后,对预测用分类型属性进行特征嵌入,再与预测用数值型属性特征进行拼接,实现特征编码。②基分类器。采用 DNN 作为模型网络结构,采用动态更新的错分比例和固定代价相结合的方式构建代价敏感矩阵,得到事件风险等级预测基分类器模型 CSDNN。③集成分类器。采用硬投票的集成方式,集成多个参数各异、性能不同的基分类器,形成事件风险等级预测集成模型 ECSDNN。

2.1 特征工程

2.1.1 特征选择

如第1节所述, ASRS 数据集具有属性高维特 点。分析数据集发现, 有些属性(如经纬度等)几乎 所有事件的取值都为空, 本文直接去除了这些多数 事件取值为空的属性列。

对于取值不为空的属性列,考虑属性间冗余特 性及属性与风险类型的相关性等,本文首先选取了 与风险类别相关的属性组:时间、地点、环境、人 物、飞机、组成、事件和评估,然后从每组属性中, 考虑属性内涵及其与风险等级潜在的相关性,选择 用于风险等级预测建模的具体属性。按此原则,本 文共选择 50 个预测用属性作为模型输入,具体如 表4所示。

2.1.2 特征编码

本文选取的预测用属性可分为分类型(如天气 等)和数值型(如高度等)2种。

分类型属性取值具有以下特性:①离散,其值 不连续;②多,具体到某一属性,可能有数千种不同 取值;③关联,具体到某一属性,其各种可能的取值 不是完全相互独立的,而是有一定的相关性。例 如,天气的取值晴天、雨天、阴天等,各个不同值之 间的距离并不相同。针对分类型属性特点,本文借

表 4	航空安全事件预测用屋性	
1.4	加上又工事口以加加商口	

Table 4	Attributes	for	aviation	safety	event	nrediction
	Attributes	101	aviation	saluty	CVCIII	րուսուստո

事件属性组	事件属性
地点	当地时间、州参考、相对角度、相对距离、AGL高度、 MSL高度
飞机1	航空交通管制、飞机操作者、创建模型名称、人员数量、 操作部分、飞行计划、任务、飞行阶段、飞行路线、领 空、导航
飞机2	航空交通管制、飞机操作者、创建模型名称
人物1	人员所在位置、在飞机上的位置、报告组织、功能模块、 资质、人为因素、沟通中断
人物2	人员所在位置、在飞机上的位置、报告组织、功能模块、 资质、人为因素、沟通中断
环境	飞机状态、天气元素\能见度、工作环境因素、光、最高飞 行限度
组成	飞机部件、参考、制造商、问题
时间	当地时间
事件	异常、探测者、是否有乘客参与、场合
评估	影响因素\情况、主要问题

鉴词向量编码思想,采用嵌入特征编码方式,对属 性的每个可能取值进行编码,将属性值映射到向量 空间并加入梯度更新,使原本离散、多且有关联的 属性值转换为指定维度的嵌入空间表示,并在嵌入 空间中可以有效衡量不同属性值的距离关系。嵌 入空间示意如图 3 所示。

从图 3 可以看出,嵌入特征编码实质是找到分 类型属性τ在嵌入空间的向量表示函数值 |z|,如 式(1)所示:

(1)

$z = \text{embedding}(\tau)$

若事件数据的预测用属性共s个,其中,分类型 属性k个,数值型属性s-k个。记分类型属性 $\tau_i(1 \le i \le k)$ 特征编码为 $z_i(1 \le i \le k)$,则对所有分类 (2)

型属性进行特征拼接,得到分类型属性特征表示, 图 4 所示。 记为**Z**,如式(2)所示:

 $\mathbf{Z} = \mathbf{z}_1 \oplus \cdots \oplus \mathbf{z}_k$

对于数值型属性 $\tau_i(k+1 \le i \le s)$,记其数值特征为 $z_i(k+1 \le i \le s)$,将其与分类型属性的嵌入特征表示进行拼接,得到事件数据的特征表示,记为Z',如式 (3)所示:

$$\mathbf{Z}' = \mathbf{z}_1 \oplus \dots \oplus \mathbf{z}_k \oplus \mathbf{z}_{k+1} \oplus \dots \oplus \mathbf{z}_s \tag{3}$$

将拼接后的属性特征 Z'送入神经网络, 经多轮 迭代不断优化, 得到最优映射函数 embedding。分 类型属性嵌入特征编码和拼接的具体过程如



图 4 嵌入向量图示

Fig. 4 Embedding vectors

通过对分类型属性进行嵌入特征编码,一方面 有效解决了此类复杂属性特征表示问题,另一方面 也在一定程度上降低了参数量及模型复杂度。假 设某分类型属性共有σ(σ>0)个不同值,则经过编 码后,特征维度为其算数平方根的向下取整,记为 σ',大幅度减少了随后分类模型训练的时间复杂度。

$$\sigma' = |\sqrt{\sigma}| \tag{4}$$

2.2 基分类器 CSDNN

基分类器 CSDNN 的设计结合了 DNN 强大的 高维数据建模能力和代价敏感学习解决类不平衡 问题的能力。

2.2.1 CSDNN 的网络结构

基本的 CSDNN 网络结构如图 5 所示,包含输入层、隐藏层和输出层,层与层之间全连接。

具体到单个神经元来讲,其线性关系Y如下:

$$Y = \sum_{i=1}^{m} w_i x_i + b \tag{5}$$

式中: *m*为神经元个数; *w*_i为权重; *x*_i为输入; *b*为偏置。 考虑ReLU非负区间的梯度为常数, 不存在梯



度消失问题,能使模型的收敛维持在一个稳定状态,因此,选取ReLU作为隐藏层神经元激活函数。 考虑Softmax函数能将多个神经元的输出映射为 (0,1)内的概率分布,因此,选取Softmax函数作为输 出层神经元激活函数。对于整体的网络结构,若定 义第*1*-1层有*m*个神经元,则对于最后一层,即*1*层 的第*j*个神经元的输出值*Y*_j如下:

$$Y_j = a_j^l = \text{Softmax}(z_j^l) = \text{Softmax}\left(\sum_{k=1}^m w_{jk}^l a_k^{l-1} + b_j^l\right)$$
(6)

嵌入空间

图 3 嵌入空间图示

 $z = \text{embedding}(\tau)$

 τ_1

2.2.2 代价敏感矩阵和损失函数

代价敏感学习是一种追求最低分类错误率的 学习算法,通过定义不同类型的代价,度量机器学 习模型在建模过程中的各类损失,从而获得满足最 小期望损失准则的学习模型。在代价敏感学习中, 代价敏感矩阵可由领域专家根据经验事先设定或 采用参数学习的方法获得^[30]。

1) 基于错分比例动态更新的代价敏感矩阵。 基于错分比例构建代价敏感矩阵,其主要思想是: 根据网络每轮迭代之后的混淆矩阵,动态更新代价 敏感矩阵。模型动态训练过程中,每一轮学习都会 生成一个混淆矩阵 $M_{nxn} = [(m_{ij})(0 \le i \le n-1, 0 \le j \le n-1)],该矩阵列代表预测类别,列的总和表示预测$ 为该类别的样本总数,行代表实际类别,行的总和表示该类别实际样本总数,n为风险等级类别数量, $<math>m_{ij}$ 表示第*i*行*j*列的元素,其值是实际为*i*类别模型 预测为*j*类别的样本数。

记动态错分比例矩阵为 $M'_{nxn} = [(m'_{ij})(0 \le i \le n-1, 0 \le j \le n-1)], m'_{ij}为第i行j列的元素, 表示实际为i类别被预测为j类别的错分样本在所有错分样本中的占比, 计算如式 (7) 所示:$

$$m'_{ij} = \frac{m_{ij}}{\sum_{i=0}^{n-1} \sum_{j=0}^{n-1} m_{ij}} \qquad i \neq j$$
(7)

基于错分比例动态更新的代价敏感矩阵记为 C₁, 计算如式 (8) 所示:

$$c_{1ij} = \delta(m'_{ij})^{\alpha} \qquad 0 \leq i \leq n-1, 0 \leq j \leq n-1, \alpha \in \{1, 2\}$$
(8)

式中: c_{1ij} 为 C_1 矩阵元素; α 为错分比例指数, 用于调整错分惩罚程度; δ 为错分比例系数, 用于调整代价取值的数量级。

2)固定代价矩阵。固定代价矩阵的设计,一要 考虑类不平衡导致的模型向多数类偏移问题,二要 考虑不同风险类别对错分敏感度不同。记固定代 价矩阵为 $S_{nxn} = [(s_{ij})(0 \le i \le n - 1, 0 \le j \le n - 1)], s_{ij}$ 表 示第i类样本错分为第j类样本的应受到的固定代 价惩罚,用其类别编号差的绝对值进行衡量,如 式(9)所示:

$$s_{ij} = |i - j|$$
 $0 \le i \le n - 1, 0 \le j \le n - 1$ (9)

对固定代价矩阵的优化,从3个方面进行: ①考虑到某类数据错分代价过大时会影响模型的 学习速度和收敛速度,最终降低模型预测精度,对 矩阵进行了归一化处理;②考虑不同风险等级错分 敏感度不同,即针对某一具体事件,将其错分为更 低风险等级或更高风险等级,其敏感程度相差很大,将矩阵上三角的元素值置为0;③为防止模型陷入局部最优无法收敛,对矩阵元素值取二次幂,以避免过惩罚。

优化后的固定代价矩阵记为 $C_2 = [(s'_{ij})(0 \le i \le n-1, 0 \le j \le n-1)]$ 。 s'_{ii} 的计算如式 (10) 所示:

$$s_{ij}' = \begin{cases} 0 & i \le j \\ \left(\frac{s_{ij}}{s_{\max}}\right)^2 & 0 \le i \le n - 1, 0 \le j \le n - 1 \end{cases}$$
(10)

最终,综合考虑错分比例和固定代价的代价敏 感矩阵*C*计算如式(11)所示:

$$\boldsymbol{C} = \beta_1 \boldsymbol{C}_1 + \beta_2 \boldsymbol{C}_2 \tag{11}$$

式中: β_1 、 β_2 为权重系数, 且 β_1 , $\beta_2 \in \{0, 0.5, 1\}$, 可调整 动态错分矩阵和固定矩阵的权重占比。

3)代价敏感损失函数。考虑代价敏感学习的 DNN,其网络损失函数L_{es}计算如式(12)所示:

$$L_{\rm cs} = \eta c_{ij} L_{\rm base} \tag{12}$$

式中: c_{ij} 为C矩阵元素; L_{base} 为基准损失函数; η 为代 价敏感程度系数, $\eta \in \{0, 10, 20, 30\}$, 若 $\eta = 0$, 则模型 退化为不考虑代价敏感损失的 DNN。

为构建更具多样性的基分类器,本文分别采用3种不同基准损失函数L_{base},训练不同的CSDNN基分类器。

 ① 交叉熵(cross entropy, CE)损失函数, 计算 如式(13)所示:

$$E = \begin{cases} -\ln p & y = 1\\ -\ln(1-p) & \ddagger \ell \ell \end{cases}$$
(13)

式中: *p*为模型预测为实际标签的概率,取值范围为 (0,1); *y* = 1表示模型预测值与样本标签值一致。

为方便计算简化交叉熵函数表示方式,定义一 个关于*p*的函数*p*_t^[31]为

$$p_t = \begin{cases} p & y = 1\\ 1 - p & \ddagger \psi \end{cases}$$
(14)

则交叉熵可表示为

$$E(p, y) = E(p_t) = -\ln p_t$$
 (15)

② Focal Loss(FL)^[31] 损失函数,其是对交叉熵 损失函数的一种改进,可以将重心聚焦在更难识别 的样本上。计算如式 (16) 所示:

$$F_l = \alpha_t (1 - p_t)^{\gamma} E \tag{16}$$

在交叉熵损失函数基础上,引入权重因子α,和 调节因子γ,其中,α,取值范围为[0,1],γ取值范围为 [0,5]。 ③标签平滑(label smoothing, LS)^[32]损失函数, 其是一种基于交叉熵损失函数进行正则化策略后 的函数。标签平滑能够防止模型在训练过程中过 于相信真实标签,从而提高模型在未知数据上的泛 化能力,其计算如式(17)所示:

$$L_{\rm s} = \begin{cases} (1-\varepsilon)E & y=1\\ \varepsilon E & \ddagger \psi \end{cases}$$
(17)

式中: ε为很小的超参数。

2.2.3 CSDNN 训练算法

综合前述分析,基分类器 CSDNN 模型训练算 法可描述如下。

输入:训练样本集*D*₁、测试样本集*D*₂及固定代 价矩阵*C*₂。

输出:基分类器 CSDNN 模型。

步骤1 分别对训练样本集中各分类型属性进 行特征嵌入编码,拼接数值型的属性得到事件特征 表示作为网络输入。

步骤2 采用固定代价矩阵C₂,在训练样本集 D₁上进行训练,并根据初始代价损失函数L_{es}进行 神经网络的梯度更新,w为权重,w'为更新后的权 重,b为偏量,b'为更新后的偏置,如式(18)所示:

$$w' = w - \frac{\partial L_{\rm cs}}{\partial w} \tag{18}$$

$$b' = b - \frac{\partial L_{\rm cs}}{\partial b} \tag{19}$$

步骤3 在测试样本集*D*2上进行测试,根据在 *D*2上计算出的错分比例矩阵*M*_{n×n},计算得到基于动态错分比例的代价敏感矩阵*C*₁。

步骤 4 根据代价敏感矩阵 C_2 和 C_1 ,计算最终的代价敏感矩阵 $C = \beta_1 C_1 + \beta_2 C_2$ 。

步骤 5 结合代价敏感矩阵 C, 在训练样本集 D₁上进行第 2 轮训练, 并通过新的 L_{es}进行新一轮 神经网络的梯度更新。

步骤 6 重复步骤 3~步骤 5,根据上一轮在测试样本集 *D*₂上的错分比例矩阵,得到新的代价敏感矩阵 *C*,直到满足设置训练轮数,此时,得到的经过训练后的网络模型为基分类器 CSDNN。

2.3 集成分类器 ECSDNN

为进一步提升模型性能,设计集成分类器 ECSDNN,考虑各基分类器 CSDNN 性能差异不大, 集成策略采用投票法。投票法是一种遵循少数服 从多数原则的集成学习模型,通过多个模型的集成 降低方差,从而提高模型的鲁棒性,又由于基模型 能预测出清晰的类别标签,本文采用硬投票的集成 策略,硬投票指预测结果是所有投票结果最多出现 的类。

ECSDNN集成过程如图 6 所示。记u_i为将事件 预测为i风险等级的分类器数量,选取数值最大的 u_i对应的风险等级i作为最终输出的风险等级,若存 在相同的u_i取值,则随机输出其中之一对应的i值。



图 6 集成模型 ECSDNN 构建 Fig. 6 Ensemble model ECSDNN construction

3 实 验

3.1 实验数据

选取 ASRS 数据集中 2010—2020 年所有航空 安全事件数据共 56 610 条,作为实验数据。考虑数 据类别不平衡问题,本文以最高数据量类别为基 准,对其他较少数据量类别进行重采样,共得到 90 140 条数据,定义该数据集为 φ ,训练集 φ_1 和测试集 φ_2 按 照 9:1 的比例划分,并满足 $\varphi_1 \cup \varphi_2 = \varphi, \varphi_1 \cap \varphi_2 = \emptyset$ 。

3.2 实验基准算法

实验基准算法包括随机森林(random forest, RF)、K最邻近(K-nearest neighbor, KNN)、朴素贝 叶斯分类器(naive Bayes classifier, NBC)、决策树分 类(decision tree classification, DTC)、逻辑回归分类

2024年

器(logistics regression classifier, LRC)、RNN $_{\circ}$

3.3 评价指标

i=0

本文评价指标包括精准率的宏平均值、召回率的宏平均值、F₁分值的宏平均值、精准率的加权平均值、召回率的加权平均值、F₁分值的加权平均值、准确率。计算公式如式 (20)~式 (26) 所示:

$$P_{\text{Macro}} = \frac{\sum_{j=0}^{n-1} R_{\text{TP}_j}}{\sum_{j=0}^{n-1} (R_{\text{TP}_j} + R_{\text{FP}_j})}$$
(20)

$$R_{\text{Macro}} = \frac{\sum_{j=0}^{n-1} R_{\text{TP}_j}}{\sum_{j=0}^{n-1} (R_{\text{TP}_j} + R_{\text{FN}_j})}$$
(21)

$$F_{1_Macro} = \frac{\sum_{j=0}^{n-1} 2R_{TP_j}}{\sum_{j=0}^{n-1} (22)}$$

$$\sum_{j=0}^{n-1} \left(2R_{\mathrm{TP}_j} + R_{\mathrm{FP}_j} + R_{\mathrm{TN}_j} \right)$$

$$P_{\text{Weighted}} = \sum_{j=0}^{n-1} \left(\frac{R_{\text{TP}_j}}{R_{\text{TP}_j} + R_{\text{FP}_j}} \cdot \frac{o_j}{o_{\text{all}}} \right)$$
(23)

$$R_{\text{Weighted}} = \sum_{j=0}^{n-1} \left(\frac{R_{\text{TP}_{j}}}{R_{\text{TP}_{j}} + R_{\text{FN}_{j}}} \cdot \frac{o_{j}}{o_{\text{all}}} \right)$$
(24)

$$F_{1_\text{Weighted}} = \sum_{j=0}^{n-1} \left(\frac{2R_{\text{TP}_j}}{2R_{\text{TP}_j} + R_{\text{FP}_j} + R_{\text{TN}_j}} \cdot \frac{o_j}{o_{\text{all}}} \right) \quad (25)$$

$$A_{\rm CC} = \frac{\sum_{j=0}^{n-1} (R_{\rm TP_j} + R_{\rm TN_j})}{\sum_{j=0}^{n-1} (R_{\rm TP_j} + R_{\rm FN_j} + R_{\rm FP_j} + R_{\rm TN_j})} \times 100\% \quad (26)$$

式中:n为风险等级数量;R_{TP}为真正;R_{FP}为假正; R_{TN}为真负;R_{FN}为假负;o为支持度;o_j为数据中实 际为*j*类的支持数量;o_{all}为所有类别的支持数量。

3.4 实验参数设置

训练基分类器 CSDNN 共 8 个, 每个 CSDNN 设 置 1 个 embedding 输入层、3 个隐藏层和 1 个输出 层, 3 个隐藏层神经元个数分别为 929、200 和 50, 层与层之间采用全连接方式, Batch_size 为 64, 学习 率为 0.001, 通过实验设置基分类器消融时η为 10, 其他实验参数设置如表 5 所示。

3.5 实验结果分析

3.5.1 对比实验

为验证本文方法的有效性,采用相同数据集划

表 5 基分类器参数设置 Table 5 Base classifier parameter settings

模型	损失函数	epoch	η	β_1	β_2	α	δ
CSDNN1	LS	200	10	0	1	2	10
CSDNN2	FL	200	10	0	1	2	10
CSDNN3	FL	200	10	0	1	1	1
CSDNN4	CE	200	30	0.5	0.5	1	1
CSDNN5	CE	200	20	0.5	0.5	1	1
CSDNN6	CE	200	10	0.5	0.5	1	1
CSDNN7	CE	100	10	0	1	1	1
CSDNN8	CE	100	10	0	1	1	1

分,分别用本文方法和基准算法进行实验,实验结果如表6所示。

由表 6 可知,本文模型 ECSDNN 的准确率达到 了 84.80%,取得较好的预测性能。ECSDNN 在其他 各项指标,如精准率的宏平均值、召回率的宏平均 值、F₁分值的宏平均值、精准率的加权平均值、召 回率的加权平均值、F₁分值的加权平均值,都表现 出了比基准算法好的性能,分别为 0.847 8、0.851 7、 0.849 3、0.845 8、0.848 0、0.846 4,表明 ECSDNN 的 预测性能比较稳定。ECSDNN 模型的实验混淆矩 阵及不同事件风险等级的预测准确率如表 7 所示。

从表7可看出,混淆矩阵对角线上数据值显著 高于其他部分,特别是对于较高风险等级类别,其 非对角线中的数值相对更小。根据表7中的准确 率可看出,模型在不同事件风险等级数据上的预测 性能存在差异,总体来讲,在低风险、中低风险、中 高风险和高风险这4个事件类别上的表现更好,在 中风险事件类别上相对较差。这是因为:与许多文 献划分3种风险类别相比,本文划分了5种风险类 别,数据标注难度更大。另外,现实中的许多事件 就是介于中风险与中高风险或中风险与中低风险 之间,不同安全专家的划分结论也不一样,对于自 动分类模型来讲,挑战就更大。

图 7 给出了真阳性率和假阳性率的变化关系 ROC曲线。可以看出,本文模型的各个风险类别 的 ROC 曲线都靠近左上角,且各个类别的 AUC 值 均大于 0.90。

3.5.2 消融实验

为验证基分类器 CSDNN 模型的合理性,本文 对基分类器进行消融实验。选取相同数据集,以相 同划分方式,设置相同训练轮数 70 轮。CSDNN 基 分类器的η取值为 10, DNN 分类器的η取值为 0,其 余参数均相同,实验结果如表 8 所示。

表 6 不同算法对比实验结果 Table 6 Comparison of experimental results with different algorithms

模型	P_{Macro}	$R_{ m Macro}$	F_{1_Macro}	$P_{\rm Weighted}$	$R_{\rm Weighted}$	$F_{1_Weighted}$	$A_{ m CC}/\%$
RF	0.808 3	0.8007	0.802 3	0.816 1	0.802 9	0.8074	80.29
KNN	0.582 7	0.571 0	0.5717	0.604 8	0.577 0	0.5863	57.70
NBC	0.579 3	0.5769	0.574 5	0.590 3	0.575 5	0.579 1	57.55
DTC	0.803 3	0.793 5	0.794 1	0.8187	0.797 1	0.804 2	79.71
LRC	0.681 4	0.6761	0.676 8	0.688 1	0.677 3	0.6807	67.73
RNN	0.7894	0.795 2	0.7897	0.7874	0.790 2	0.786 1	79.02
ECSDNN	0.847 8	0.8517	0.849 3	0.845 8	0.848 0	0.8464	84.80

表 7 不同风险等级混淆矩阵结果及准确率

Table 7 Confusion matrix results and accuracy of different risk levels

风险等级 -	混淆矩阵值						
	低风险	中低风险	中风险	中高风险	高风险	$A_{\rm CC}/\%$	
低风险	1 541	31	160	15	31	86.67	
中低风险	16	1 619	53	9	36	93.42	
中风险	209	120	1 345	140	135	69.01	
中高风险	26	14	128	1 549	39	88.21	
高风险	40	32	85	43	1 544	88.53	



Fig. 7 ROC curves

由表 8 可以看出,在各个性能评价指标上,本 文所构建的基分类器 CSDNN 都明显优于 DNN 模型,其中,在准确率评价指标上提升了 1.1%。

为验证集成分类器 ECSDNN 模型的有效性, 对 ECSDNN 进行了消融实验,表9给出了集成模型 ECSDNN 和8个基分类器的实验结果对比。

从表9可看出, ECSDNN 集成模型预测性能明显优于单个基分类器 CSDNN。其中,准确率指标提升了3.17%,其他指标也都有明显优势,说明本文所提 ECSDNN 模型适用性更强,能一定程度解决航空安全风险等级预测问题。

表 8 CSDNN 基分类器消融实验结果

Table 8 CSDNN base classifier ablation experiment result

模型	P _{Macro}	R _{Macro}	F_{1_Macro}	$P_{\rm Weighted}$	$R_{\rm Weighted}$	$F_{1_Weighted}$	$A_{\rm CC}/\%$
DNN	0.795 9	0.804 8	0.797 6	0.793 9	0.7991	0.793 5	79.91
CSDNN	0.807 8	0.815 5	0.808 7	0.806 2	0.8101	0.805 0	81.01

表9 消融实验结果

Table 9 Ablation experiment results

模型	P_{Macro}	R _{Macro}	F_{1_Macro}	P_{Weighted}	$R_{\rm Weighted}$	$F_{1_Weighted}$	$A_{\rm CC}$ /%
CSDNN1	0.8144	0.821 1	0.816 5	0.812 0	0.8163	0.812 8	81.63
CSDNN2	0.809 9	0.8163	0.8117	0.808 0	0.8117	0.808 4	81.17
CSDNN3	0.807 5	0.814 2	0.809 1	0.805 6	0.8094	0.805 7	80.94
CSDNN4	0.807 5	0.815 9	0.809 6	0.805 4	0.810 5	0.805 8	81.05
CSDNN5	0.8102	0.8164	0.812 2	0.807 8	0.8117	0.808 7	81.17
CSDNN6	0.814 0	0.820 1	0.8159	0.8119	0.8156	0.812 6	81.56
CSDNN7	0.8106	0.8177	0.812 8	0.808 3	0.812 8	0.809 2	81.28
CSDNN8	0.808 5	0.815 8	0.8106	0.8064	0.810 8	0.807 0	81.08
ECSDNN	0.847 8	0.8517	0.849 3	0.845 8	0.848 0	0.8464	84.80

为验证选取属性的合理性,首先,采用极端梯 度增强算法(extreme gradient boosting, XGBoost)计 算每个属性的特征重要性,属性重要性排序如 图 8 所示。其次,去掉其中最不重要的 10 个属性, 用剩余 40 个预测属性和本文方法进行其余参数均 相同的对比实验,实验结果如表 10 所示。可以看 出,当减少预测用属性数量时,模型的各项性能指 标均下降,表明了本文属性子集选择的合理有效性。



Fig. 8 Importance ranking of features

表 10 属性合理性实验结果

Table 10 Attribute rationality test results

属性数量	$P_{\rm Macro}$	R _{Macro}	$F_{1_{\rm Macro}}$	$P_{\rm Weighted}$	$R_{\rm Weighted}$	$F_{\rm 1_Weighted}$	$A_{\rm CC}$ /%
40	0.786 0	0.795 0	0.7878	0.784 0	0.7893	0.783 8	78.93
50	0.8172	0.823 2	0.818 9	0.815 1	0.8187	0.8158	81.87

4 结 论

1)本文提出一种基于 ECSDNN 的航空安全事 件风险等级预测方法,在 ASRS 数据集上,本文方 法优于经典的基准模型,预测能力较之最优基准模 型提升了 4.51%,有明显提升; ECSDNN 模型较之 单个 CSDNN 模型预测能力提升了 3.17%;本文方 法取得较好的预测结果,准确率达到 84% 以上。通 过实验证了本文方法的有效性。

2)本文研究成果可以为航空安全系统的主动 风险管理提供决策依据。对于航空安全事件报告 数据集中的新增事件报告,可以自动生成风险等 级,简化航空安全专家复杂繁琐的分析过程;当实 际发生具体航空安全事件时,有助于迅速形成事件 风险程度的预判,为事件的风险应对、资源配置提 供有力的决策支持。

航空安全事件更新速度快,本文所提航空安全 事件风险等级预测方法中代价敏感部分,未来可结 合实时更新的航空安全事件进行对代价敏感矩阵 的优化,以适应航空安全事件风险等级预测的实时 性需求。

参考文献(References)

- [1] HU X, WU J, HE J R. Textual indicator extraction from aviation accident reports: AIAA 2019-2939[R]. Reston: AIAA, 2019.
- [2] ROBINSON S. Multi-label classification of contributing causal factors in self-reported safety narratives[J]. Safety, 2018, 4(3): 30.
- [3] ROSE R L, PURANIK T G, MAVRIS D N. Natural language processing based method for clustering and analysis of aviation safety narratives[J]. Aerospace, 2020, 7(10): 143.
- [4] SUBRAMANIAN S V, RAO A H. Deep-learning based time series forecasting of go-around incidents in the national airspace system: AIAA 2018-0424[R]. Reston: AIAA, 2018.
- [5] PARADIS C, KAZMAN R, DAVIES M, et al. Augmenting topic finding in the NASA aviation safety reporting system using topic modeling: AIAA 2021-1981[R]. Reston: AIAA, 2021.
- [6] TANGUY L, TULECHKI N, URIELI A, et al. Natural language processing for aviation safety reports: From classification to interactive analysis[J]. Computers in Industry, 2016, 78: 80-95.
- [7] KUHN K D. Using structural topic modeling to identify latent topics and trends in aviation incident reports[J]. Transportation Research Part C:Emerging Technologies, 2018, 87: 105-122.
- [8] YAN W L, ZHOU J H. Early fault detection of aircraft components using flight sensor data[C]//Proceedings of the IEEE International Conference on Emerging Technologies and Factory Automation. Piscataway: IEEE Press, 2018: 1337-1342.
- [9] JANAKIRAMAN V M, NIELSEN D. Anomaly detection in aviation data using extreme learning machines[C]//Proceedings of the International Joint Conference on Neural Networks. Piscataway: IEEE Press, 2016: 1993-2000.

- [10] LIU Y F, LV J H, MA S L. A real time anomaly detection method based on variable *n*-gram for flight data[C]//Proceedings of the IEEE International Conference on High Performance Computing and Communications, IEEE International Conference on Smart City, IEEE International Conference on Data Science and Systems. Piscataway: IEEE Press, 2018: 370-376.
- [11] 冯霞,李娟娟, 闫冠男. 关联规则挖掘在航空安全报告分析中的 应用[J]. 计算机工程与设计, 2011, 32(1): 218-220.
 FENG X, LI J J, YAN G N. Application of association rules mining

in aviation safety reports analysis[J]. Computer Engineering and Design, 2011, 32(1): 218-220(in Chinese).

- [12] 刘俊杰, 李华明, 梁文娟, 等. 基于内容分析法的航空安全自愿报告信息分析[J]. 中国安全科学学报, 2012, 22(4): 90-96.
 LIU J J, LI H M, LIANG W J, et al. Analysis of aviation safety confidential reports based on content analysis method[J]. China Safety Science Journal, 2012, 22(4): 90-96(in Chinese).
- [13] 刘俊杰, 杜尹岚, 闫慧娟. Python 环境下的航空安全报告信息分 析方法[J]. 科学技术与工程, 2021, 21(10): 4278-4283. LIU J J, DU Y L, YAN H J. The analysis method of aviation safety reporting information based on Python[J]. Science Technology and Engineering, 2021, 21(10): 4278-4283(in Chinese).
- [14] 宁静, 佘红艳, 赵东, 等. 一种路网级交通事故风险预测方法[J].
 北京邮电大学学报, 2022, 45(2): 72-78.
 NING J, SHE H Y, ZHAO D, et al. A road-level traffic accident risk prediction method[J]. Journal of Beijing University of Posts and Telecommunications, 2022, 45(2): 72-78(in Chinese).
- [15] 柳本民,廖岩枫,涂辉招,等.基于模拟实验的低等级公路车辆过 弯风险预测模型[J].同济大学学报(自然科学版), 2021, 49(4): 499-506.

LIU B M, LIAO Y F, TU H Z, et al. Risk prediction model of vehicle driving in small radius curves based on simulation experiment[J]. Journal of Tongji University (Natural Science), 2021, 49(4): 499-506(in Chinese).

- [16] 赵海涛,程慧玲,丁仪,等.基于深度学习的车联边缘网络交通事 故风险预测算法研究[J].电子与信息学报,2020,42(1):50-57. ZHAO H T, CHENG H L, DING Y, et al. Research on traffic accident risk prediction algorithm of edge Internet of vehicles based on deep learning[J]. Journal of Electronics & Information Technology, 2020, 42(1): 50-57(in Chinese).
- [17] 赵晓华, 亓航, 姚莹, 等. 基于可解释机器学习框架的快速路立交出口风险预测及致因解析[J]. 东南大学学报 (自然科学版), 2022, 52(1): 152-161.
 ZHAO X H, QI H, YAO Y, et al. Risk prediction and causation ana-

lysis of expressway interchange exits based on interpretable machine learning framework[J]. Journal of Southeast University (Natural Science Edition), 2022, 52(1): 152-161(in Chinese).

[18] 曾小清,林海香,王奕曾,等.基于事故数据的轨道交通运行安全 风险辨识方法[J].同济大学学报(自然科学版),2022,50(3):418-424. ZENG X Q, LIN H X, WANG Y Z, et al. Safety risk identification of rail transit signaling system based on accident data[J]. Journal of TongjiUniversity(NaturalScience),2022,50(3):418-424(inChinese).

- [19] SRINIVASAN P, NAGARAJAN V, MAHADEVAN S. Mining and classifying aviation accident reports: AIAA 2019-2938[R]. Reston: AIAA, 2019.
- [20] ALKHAMISI A O, MEHMOOD R. An ensemble machine and deep learning model for risk prediction in aviation systems[C]//Proceedings of the 6th Conference on Data Science and Machine Learning Applications. Piscataway: IEEE Press, 2020: 54-59.
- [21] 倪晓梅, 王华伟, 熊明兰, 等. 基于文本挖掘的民航事件风险评估
 [J]. 湖南大学学报 (自然科学版), 2022, 49(6): 73-79.
 NI X M, WANG H W, XIONG M L, et al. Civil aviation incident risk assessment based on text mining[J]. Journal of Hunan University (Natural Sciences), 2022, 49(6): 73-79(in Chinese).
- [22] ZHANG X G, MAHADEVAN S. Ensemble machine learning models for aviation incident risk prediction[J]. Decision Support Systems, 2019, 116: 48-63.
- [23] FANG Y. Feature selection, deep neural network and trend prediction[J]. Journal of Shanghai Jiaotong University (Science), 2018, 23(2): 297-307.
- [24] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [25] 吴雨茜, 王俊丽, 杨丽, 等. 代价敏感深度学习方法研究综述[J]. 计算机科学, 2019, 46(5): 1-12.
 WU Y X, WANG J L, YANG L, et al. Survey on cost-sensitive deep learning methods[J]. Computer Science, 2019, 46(5): 1-12(in Chinese).
- [26] RUSKIN K J, CORVIN C, RICE S, et al. Alarms, alerts, and warnings in air traffic control: An analysis of reports from the aviation safety reporting system[J]. Transportation Research Interdisciplinary Perspectives, 2021, 12: 100502.
- [27] 刘梦娜. 基于文本挖掘的航空安全事故报告致因因素分析和风险预测[D]. 合肥: 安徽建筑大学, 2019. LIU M N. Analysis of influencing factors and risk prediction of aviation safety accident report based on text mining[D]. Hefei: Anhui Jianzhu University, 2019(in Chinese).
- [28] International Civil Aviation Organization. Safety managementmanual (SMM)[EB/OL]. [2022-08-28]. https://www.icao.int/NACC/Documents/Meetings/2014/SSPSMSANT/Doc9859.pdf#search=doc9859.
- [29] CARMONA M. What is the NASA ASRS?[EB/OL]. [2022-04-11]. https://asrs.arc.nasa.gov/uassafety.html.
- [30] 万建武,杨明.代价敏感学习方法综述[J].软件学报,2020,31(1):
 113-136.
 WAN J W, YANG M. Survey on cost-sensitive learning method[J].

Journal of Software, 2020, 31(1): 113-136(in Chinese).

- [31] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2017: 2999-3007.
- [32] MÜLLER R, KORNBLITH S, HINTON G. When does label smoothing help?[C]//Proceedings of the 33rd International Conference on Neural Information Processing Systems. New York: ACM, 2019: 4694-4703.

Prediction of aviation safety event risk level based on ensemble cost-sensitive deep neural network

FENG Xia^{1, 2}, SANG Xiao^{1, 2}, ZUO Haichao^{1, 2, *}

College of Computer Science and Technology, Civil Aviation University of China, Tianjin 300300, China;
 Key Laboratory of Smart Airport Theory and System, CAAC, Tianjin 300300, China)

Abstract: One key component of active risk management is the prediction of aviation safety event risk levels. Considering the characteristics of high-dimensional complexity and class imbalance presented by massive aviation safety event data, this paper proposes an aviation safety event risk level prediction method based on an ensemble cost-sensitive deep neural network (ECSDNN). First, the feature representation of aviation safety event data is realized by using the method of splicing type attribute embedding coding and numerical attribute; secondly, a cost-sensitive matrix and a cost-sensitive loss function are designed comprehensively considering the misclassification ratio and fixed cost, and a base classifier model based on a cost-sensitive deep neural network (CSDNN) is constructed; finally, an ensemble prediction model of aviation safety event risk level ECSDNN is created by integrating various base classifiers with varying parameters and performances using the hard voting approach. The experimental results on the aviation safety reporting system (ASRS) dataset demonstrate that the prediction accuracy of the ECSDNN model is improved by 3.17% when compared with the single CSDNN base classifier and by 4.51% when compared with the optimal prediction ability of the benchmark algorithm. The effectiveness of the ensemble cost-sensitive deep neural network method for aviation safety event risk level prediction is verified.

Keywords: aviation safety; risk level prediction; embedded feature encoding; cost-sensitivity; deep neural network; ensemble learning

URL: link.cnki.net/urlid/11.2625.V.20220928.1033.001

Received: 2022-06-11; Accepted: 2022-09-11; Published Online: 2022-09-29 11:27

Foundation items: Key Project of Civil Aviation Joint Fund of National Natural Science Foundation of China (U2333206); National Key R & D Program of China (2021YFF0603902); The Fundamental Research Funds for the Central Universities (3122021063); Tianjin Research Innovation Project for Postgraduate Students (2021YJSO2S14)

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn DOI: 10.13700/j.bh.1001-5965.2022.0445

30% HNO, 对 2195 铝锂合金预腐蚀疲劳特性研究

刘德俊1,田干1,*,李玉龙2,张炜1,张有宏1

(1. 火箭军工程大学 导弹工程学院,西安 710025; 2. 中国运载火箭技术研究院 北京 100076)

摘 要:模拟液体导弹贮箱材料和使用特点,进行 2195-T8 铝锂合金在 30% 硝酸溶液中预腐蚀不同时间后的疲劳试验,利用扫描电子显微镜 (SEM)、电子背散射衍射仪 (EBSD) 及透射电子显微镜 (TEM)等显微表征方法,研究了预腐蚀对疲劳性能的影响。结果表明:预腐蚀表面形貌以晶间腐蚀和点蚀为主,在轧制方向上形成条带状腐蚀链,其产生与长条形晶界和金属间粒子连续腐蚀作用相关;经预腐蚀后,铝锂合金试样在 210,280,330 MPa 中等应力幅下的预腐蚀疲劳寿命较无腐蚀试样有所增加,16h腐蚀试样寿命提高最为明显,其原因是腐蚀通过将合金表面微裂纹、孔洞等缺陷消除、增加表面粗糙度等方式,降低了应力集中程度,进而延缓了疲劳裂纹萌生。

关键 词:铝锂合金;推进剂贮箱;预腐蚀;疲劳特性;裂纹萌生

中图分类号: V45; TG178

文献标志码: A 文章编号: 1001-5965(2024)04-1129-09

铝锂合金以其高强度、高比模量、低密度的突 出物理性能在航空航天、军事领域得到了广泛应 用^[1-3]。作为"十四五"重点研究的合金材料,铝锂 合金在火箭、导弹等轻量化建设中发挥着重要作 用^[4-6]。2195 铝锂合金是液体导弹、火箭推进剂贮 箱潜在的结构材料,在长时间贮存过程中,合金经 受腐蚀作用,其机械性能将不同程度下降,对其预 腐蚀疲劳特性的研究将对服役年限的延长具有十 分重要的意义。

近年来,国内外学者关于铝锂合金的抗腐蚀能 力及疲劳特性开展了广泛研究。在腐蚀机理研究 方面,Ma等^[7]比较了合金中富铜及贫铜金属间粒 子腐蚀能力,发现富铜粒子含Li元素更多,抗腐蚀 能力更强;文献[8]对比了不同热处理工艺2198 铝 锂合金抗腐蚀能力,证实了T3工艺易产生晶间腐 蚀,但不存在剥落腐蚀;Liu等^[9]研究了不同粗糙度 铝锂合金的腐蚀特性,认为粗糙表面因金属化合物 存在导致腐蚀性能退化;此外,Lei等^[10]针对不同 Li含量铝锂合金点蚀损伤进行了原位观测,试验表 明高 Li 含量大大缩短了点蚀萌生及快速扩展进程,加速了点蚀形成。在疲劳断裂方面, Zhong 等^[11]研究了 2A97 铝锂合金显微组织与疲劳裂纹萌生及早期扩展行为关系,发现晶粒取向差、位错与滑移带相互作用造成了裂纹分叉和偏转;此外,铝锂合金表面富铜粒子^[12]、位错增殖与塞积^[13]、驻留滑移带^[14]被证实是裂纹萌生机制,并影响裂纹扩展路径行为。以上研究对不同牌号的铝锂合金单一腐蚀和疲劳性能展开了分析,但是鲜有对 2195 铝锂合金预腐蚀疲劳性能的报道。

液体推进剂主要成分为 N₂O₄,在长期贮存环境 中由于含有水分,二者通过水解反应转化为 HNO₃, 进而侵蚀贮箱合金材料,文献 [15-16]发现 30% HNO₃反应活性最强,可用于模拟加速腐蚀状态。 因此,本文采用 30% HNO₃溶液进行腐蚀环境 对 2195-T8 铝锂合金预腐蚀,然后进行疲劳性能 测试。结合扫描电子显微镜(scanning electron microscopy, SEM)、透射电子显微镜(transmission electron microscope, TEM)及电子背散射衍射仪

*通信作者. E-mail: tiangan_2012@163.com

收稿日期:2022-05-31;录用日期:2022-07-01;网络出版时间:2022-07-1916:42 网络出版地址:link.cnki.net/urlid/11.2625.V.20220719.0854.001

基金项目:国家自然科学基金 (52272446,52075541);陕西省自然科学基金 (2022JM-243)

引用格式:刘德俊,田干,李玉龙,等. 30% HNO₃ 对 2195 铅锂合金预腐蚀疲劳特性研究 [J]. 北京航空航天大学学报, 2024, 50 (4): 1129-1137. LIUDJ, TIANG, LIYL, et al. Research on pre-corrosion fatigue properties of 2195 Al-Li alloys in 30% HNO₃[J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50 (4): 1129-1137 (in Chinese).

(electron backscatter detector, EBSD)等显微手段,分 析预腐蚀对 2195-T8 铝锂合金疲劳寿命的影响,为 液体导弹长贮的相容性和安全性研究提供参考。

1 试验方法

试验所用材料是由西南铝业提供的 4.5mm 厚 2195-T8 铝锂合金板材,基本力学性能如表 1 所示, 其成分组成质量分数为: Li 1.0%, Cu 4.0 %, Si 0.03%, Fe 0.05%, Mn 0.05%, Mg 0.4%, Zn 0.01%, Ti 0.02%, Zr 0.14%, Ag 0.4%, Al 余量。

表 1 2195-T8 铝锂合金力学性能 Table 1 Mechanical properties of 2195-T8 Al-Li alloys

弹性模量 E/GPa	屈服强度 σ_s /MPa	极限强度 $\sigma_{\rm b}/{ m MPa}$	延伸率/%
72.90	583.33	609.90	11.4

参照 GB/T 26077—2010^[17],将板材线切割成如 图 1 所示的狗骨状试样,共 60 件。试样在磨抛机 上从 400 目打磨至 2 000 目(1 mm=14 目),并将抛 光后试样进行超声清洗、干燥,利用 HITACHI SU8010 SEM 与 Talos 200S TEM 对初始试样表面进 行背散射电子(back-scaterred electron, BSE)及 EBSD 显微观察与晶粒取向分析。

如图 2(a)所示,试样夹持端用聚四氟乙烯模具 密封后浸泡在 30% HNO3 至 8,16,24 h,每组预腐 蚀试样数为 15。经清洗、干燥、SEM 观察腐蚀形 貌后,按 GB3075-82 标准^[18],在 INSTRON_8801 电 液伺服疲劳试验机上、室温下进行疲劳试验(见



图 1 试样尺寸及形状 Fig. 1 Specimen size and shape



(a) 预腐蚀试验

(b) 疲劳加载

图 2 试验状态 Fig. 2 Test state 图 2(b)),加载频率为 5 Hz,输入波形为正弦波,应 力比 *R*=0.1,应力峰值分别为 140,210,280,330, 440 MPa,每个应力水平下的试件数为 3,然后采用 SEM 进行断面分析。

2 试验结果

2.1 显微组织

图 3 为 2195-T8 铝锂合金表面形貌。在轧制方向上,合金晶粒呈长条状,且间歇分布着如方框所示带状粒子簇。如图 3(b)所示,在粒子簇带中,粒子随机分布在晶粒内部或晶界上,而相邻条带之间的粒子数量少、体积小且分布随机,图 3(c)中元素能谱表明粒子中存在含量较高的 Fe、Cu 元素(见图 3(b)黑色箭头指向)。此外,表面上还观察到白色箭头指向的微观缺陷,这些缺陷的存在使合金表面不连续,造成应力集中,容易导致裂纹萌生,从而劣化合金机械性能。

图 4 为 2195-T8 铝锂合金 EBSD 结果。在图 4(a)、 图 4(b)中,尺寸较大晶粒内部无明显颜色梯度,内



(a) 表面形貌



(b) 图(a) 中方框放大1 000倍



图 3 2195-T8 铝锂合金 BSE 形貌及元素能谱结果 Fig. 3 BSE micrographs and EDS results of 2195-T8 Al-Li alloys

部形变位错得到消除,晶粒间取向差很小,属于亚 晶结构;此外,合金中还存在细小的再结晶晶粒,从 局部取向图中可见形成再结晶区域有一定局部取 向差。由于部分再结晶晶粒位于大亚晶内部,使得 该部分亚晶内具有取向差。图 4(c)为沿迹线的取 向差,明显看出经过晶界时取向差出现峰值突变, 均值在 30°以上,说明合金中晶界属于高角度晶界。



图 5(a)为三亚晶界交界处 TEM 图,可以看到 在晶内针状的 Al₂CuLi(*T*₁相)均匀分布,同时还分 散着球状的 Al₃Li / Al₃ Zr(δ′/β′相)。在晶界处,未 观察到无析出带,但是有二次相嵌入在晶界上(如 箭头所示),沿 A-B 路径 Cu 强度分布如图 5(b)所 示,可见 Cu 在二次相处达到峰值。

2.2 预腐蚀形貌

图 6 为 2195-T8 铝锂合金分别经历 8, 16, 24 h 后的腐蚀形貌。从图 6(a)、图 6(c)、图 6(e)可以看



(a) 二亚丽乔文乔处TEM图



图 5 2195-T8 铝锂合金 TEM 结果 Fig. 5 TEM results of 2195-T8 Al-Li alloys

出,随着腐蚀时间的增加,腐蚀程度逐渐加深。在 轧制方向上,形成了一条条腐蚀链(虚线框所示), 腐蚀链上腐蚀程度高于非链上的位置,且腐蚀链之 间距离随腐蚀时间延长而缩短。图 6(b)、图 6(d) 及图 6(f)为腐蚀 8,16,24 h的局部形貌放大图,从 中可看出腐蚀形貌由点蚀(细箭头所示)及晶间腐 蚀(宽箭头所示)构成。经 8 h腐蚀后,合金表面以 点蚀为主,伴随着晶间腐蚀。然而,腐蚀 16 h和 24 h 后,合金表面晶间腐蚀程度加剧,再结晶小晶粒之 间因晶界腐蚀而间隙变宽;点蚀则主要出现在大尺 寸长晶粒内部,且长晶粒晶界因侵蚀而不再连续。 另外,如图中短箭头所示,可见一些边缘粗糙,由被 腐蚀的晶界围成的坑,这是由于含 Fe、Cu 等重金属 粒子周围基体或晶界被腐蚀后,使粒子脱落形成的 空位。

2.3 预腐蚀形貌

通过拉-拉疲劳试验,得到 60 个 2195-T8 铝锂 合金试样预腐蚀 8,16,24 h 及无腐蚀试样在不同 应力下的疲劳寿命,结果采用式(1)拟合,且按照 式(2)和式(3)分别计算出试样的中值寿命与标准 差^[19],其结果如表 2 所示。

$$S^m N = C$$

(1)



(d)图(c)实线放大图

(e) 24 h腐蚀形貌 图 6 预腐蚀不同时间后合金表面形貌

(f)图(e)实线放大图

Fig. 6 Alloy surface morphology after different durations of pre-corrosion

表 2 不同预腐蚀时间下疲劳寿命

Table 2 Fatigue cycles under different pre-corrosion durations

마누 안국 제	广大和历	· 古井 吉 人 小 与	AL NE	
ŀij ¤j/h	<u>述</u> 刀/MPa	波穷寿命/伙	N ₅₀ / (X	σ
	140	867 558, 345 447, 587 392	560 449	0.201
	210	45 680, 55 923, 110 945	65 687	0.202
0	280	35 504, 35 196, 31 748	34 106	0.027
	330	25 535, 19 844, 19 068	21 299	0.068
	440	9 027, 9 630, 9 322	9 323	0.014
	140	274 086, 624 517, 266 834	357 457	0.209
	210	92 995, 74 407, 70 273	78 536	0.064
8	280	28 399, 37 898, 32 243	32 618	0.063
	330	24 911, 23 303, 20 549	22 849	0.042
	440	9 996, 10 176, 10 012	10 086	0.004
	140	543 995, 476 059, 533 021	516 813	0.031
	210	92 055, 74 507, 78 209	81 252	0.048
16	280	49 749, 53 557, 42 351	48 323	0.052
	330	20 723, 18 132, 26 496	21 513	0.083
	440	9 032, 11 481, 13 426	11 166	0.087
	140	254 628, 217 528, 293 901	253 439	0.065
	210	84 543, 95 573, 110 622	96 328	0.058
24	280	50 236, 49 204, 49 803	49 746	0.005
	330	24 163, 27 778, 21 368	24 296	0.057
	440	7 611, 7 926, 6 660	5 988	0.039

式中: S^m为应力水平, m为常数; N为疲劳寿命; C为常数。

$$\overline{X} = \lg N_{50} = \frac{1}{n} \sum_{i=1}^{n} \lg N_i$$
(2)

$$\sigma = \sqrt{\frac{n \sum_{i=1}^{n} \lg^{2} N_{i} - \left(\sum_{i=1}^{n} \lg N_{i}\right)^{2}}{n(n-1)}}$$
(3)

式中: **X**为各应力幅中值对数疲劳寿命; *n* 为每个应 力幅下的试样个数; *N_i* 对应第*i* 个试样的疲劳寿命; *N_{s0}* 为具有 50% 存活率的疲劳寿命; σ 为子样方差。

如图 7 所示,分别为不同铝锂合金疲劳性能与 拉伸性能对比和不同预腐蚀时间下 2195-T8 铝锂合 金的 S-N曲线。从图 7(a)可看出 2195-T8 铝锂合金 疲劳强度低于其他牌号铝锂合金,然而其静拉伸强 度却达到 607 MPa 明显高于其他牌号铝锂合金,说 刘德俊,等:30% HNO,对2195 铝锂合金预腐蚀疲劳特性研究



Fig. 7 Mechanical properties comparison among various Al-Li alloys and pre-corrosion fatigue results

明 2195 铝锂合金作为运载火箭或液体导弹贮箱材 料,为承载大吨量推进剂其静强度要求很高,因此, 其疲劳强度只需要满足一般要求即可。然而,在推 进剂腐蚀损伤作用下,2195-T8 铝锂合金机械强度 会随之退化,由于在长贮过程中,贮箱要不断进行 打压试验并监测,因此,2195-T8 铝锂合金预腐蚀后 疲劳强度的退化规律要进一步探究。从 S-N曲线 看出,对无腐蚀试样,高荷载下合金每周次循环过 程中的累积损伤相较低荷载下更多,导致试样寿命 趋于一致,分散性较小,而低应力幅下,试样损伤形 式多样化进而使寿命分散化。从不同时间预腐蚀 试样 S-N曲线可看出,相较于无腐蚀试样,预腐蚀 试样在各个应力下寿命分散性降低。预腐蚀 8 h 试 样的 S-N曲线与无腐蚀的曲线相近,而 16 h 和 24 h

第4期

的 S-N 曲线在 210,280,330 MPa 中等应力范围内 的寿命高于无腐蚀试样,在 140 MPa 低应力和 440 MPa 高应力下的寿命则接近或低于无腐蚀寿 命。此外,预腐蚀 24 h 的试样 S-N 曲线随应力幅值 呈明显的线性下降。腐蚀后的试样对中等应力范 围内的疲劳不敏感,甚至增加疲劳寿命,这与合金 腐蚀特征紧密联系。

2.4 断口形貌

不同预腐蚀时长试样的疲劳断口形貌如图 8 所示。图 8(a)~图 8(d)为各类预腐蚀试样疲劳裂 纹源区形貌,对无腐蚀试样(见图 8(a)),疲劳裂纹 萌生于合金表面缺陷边缘处,边缘处同时存在微小 孔洞及通过滑移挤入/挤出形成的滑移平面(如箭头 所示),裂纹从孔洞及滑移带局部应力集中处萌生,



Fig. 8 Micrography of fatigue fracture surface of 2195-T8 Al-Li alloys under different pre-corrosion durations

且萌生的位置环绕缺陷边缘;在裂纹扩展初始阶段,裂纹沿着晶粒主滑移系方向以纯剪切形式扩展,由于晶粒取向及扩展阻力的不同,疲劳条纹出现在不同高度的晶面上,垂直于加载方向扩展。经腐蚀后的试样,疲劳裂纹从连片的腐蚀坑群中萌生且未观察到滑移带特征,裂纹源较未腐蚀试样数目更多、位置分布更广。同时,可明显看到位于不同高度的解理面疲劳裂纹扩展早期在不同晶粒上形成的垂直于加载方向的疲劳条纹。

图 8(e)~图 8(h)为不同腐蚀时长试样的稳定 裂纹扩展区,此时,应力强度较高,裂纹尖端塑性增 强,裂纹扩展沿 2 个滑移系统同时或交替进行,进 而形成疲劳辉纹。从图中看出,预腐蚀与未腐蚀试 样在稳定扩展区断口形貌类似,疲劳辉纹在长条形 晶粒内沿着晶界方向延长,且裂纹扩展方向基本一 致。但是,在不同晶粒内,疲劳辉纹间距及长度不 同,且疲劳辉纹止于晶界处,这是由于晶粒取向的 不同会使晶粒对裂纹扩展的阻力存在差异,晶界对 裂纹在相邻晶粒中的连续扩展形成阻碍^[20]。

当疲劳裂纹扩展到一定程度,形成瞬时断裂 区。如图 8(i)~图 8(1)所示,不同腐蚀时长试样的 瞬时断裂区基本相似,能看到断口有明显的分层特 征和韧性脊,具有方向性,且韧性脊上分布着大小 不一的韧窝,属于以塑性断裂为主的沿晶分层断裂。

3 分析与讨论

3.1 腐蚀损伤机理分析

对 2195-T8 铝锂合金,其在轧制及应变过程中 会产生晶体塑性变形,引起位错运动,由 EBSD 结 果可知,2195-T8 铝锂合金中晶粒间取向差较小,但 晶界角度很大,这会导致位错在晶界塞积,*T*₁相倾 向于在晶界附近析出,使晶界具有优先腐蚀特性; 此外,晶界上分布着富铜相,其铜含量高于附近基 体,在局部电化学进程中,充当阴极作用促进基体 腐蚀,而且文献 [21] 还阐释了高角度晶界吉布斯自 由能更高,使晶界活性进一步增强。因此,HNO₃腐 蚀后在试件表面看到了成片的晶间腐蚀。

如图 4 所示,合金表面晶粒沿轧制方向呈长条状,晶界被拉长。由于晶界腐蚀活性高,腐蚀沿长条形晶界更易进行,从而促进了条带状腐蚀链的形成^[22]。此外,沿轧制方向上的含 Fe、Cu 等金属元素的金属间粒子对腐蚀链的形成也起到了一定作用。由于高含量重金属元素的存在,粒子的电化学腐蚀电位高于基体,使基体成为阳极而被侵蚀^[23]。单个粒子腐蚀能力有限,导致腐蚀具有局限性,但是多粒子组成的粒子簇则大大增强了腐蚀侵蚀能

力,这是因为粒子簇体积更大,使基体与之腐蚀电 位差更高,增强了腐蚀能力^[24-25]。另外,文献[26] 表 明,金属间粒子的相互作用会使腐蚀具有连续性, 且有向基体内部腐蚀的趋势,造成粒子周边基体因 被腐蚀而呈沟槽状,从而孤立出金属间粒子,致使 粒子容易松动,这解释了清洗后的腐蚀表面上粗糙 轮廓的腐蚀孔洞的形成原因,是由金属间粒子在清 洗过程中从表面脱落导致的。因此,在长条晶界的 晶间腐蚀和轧制方向上金属间粒子簇连续腐蚀的 共同作用下,沿轧制方向产生了条带状的腐蚀链。

3.2 预腐蚀对疲劳寿命的影响

从 S-N曲线看出,30% HNO3 预腐蚀后的试样 在210,280,330 MPa中等应力范围内的寿命较无 腐蚀试样寿命均有不同程度的增加,特别是预腐蚀 16 h试样,其寿命增加最多,这与腐蚀造成疲劳寿 命下降的结论有所不同。

疲劳寿命由裂纹萌生寿命和扩展寿命组成,其 中裂纹萌生寿命占总体寿命 90% 以上,而无腐蚀试 样与预腐蚀试样之间差别主要在于疲劳源之间的 差别,因此疲劳源的研究能够阐明预腐蚀试样在中 等应力范围内寿命增加的原因^[27]。

图 9 为中等应力条件下各类预腐蚀试样的疲 劳裂纹萌生区,从图中看出无腐蚀试样裂纹萌生于 合金表面缺陷,预腐蚀试样裂纹则从腐蚀坑群中萌 生,因为这些位置强度薄弱、应力集中程度高^[28]。 对无腐蚀试样,表面缺陷两侧为打磨后平整的平 面,而预腐蚀试样表面都是相邻毗连的腐蚀坑,表 面粗糙度、应力集中程度的差异导致了预腐蚀试样 疲劳寿命对中等应力的不敏感。

图 10 为无腐蚀试样表面经腐蚀的演化示意 图,合金经 30% HNO,腐蚀后,其表面产生了深度 不一、相互连接的腐蚀坑,且腐蚀后表面看上去较 无腐蚀的粗糙,这是因为晶间腐蚀和点蚀在合金表 面广泛分布[29-30]。图 3 中显示了腐蚀前合金表面存 在部分微裂纹、划痕及孔洞等缺陷,通过均匀腐蚀 恰好消除了这些表面缺陷。文献 [31] 表明表面粗 糙度对应力集中有很大影响,表面多缺陷的存在会 降低单缺陷的应力集中系数,而且缺陷间距越小, 应力集中系数降低越明显。如图 10(b)所示,预腐 蚀试样表面点蚀产生的腐蚀坑彼此相连,形成粗糙 表面,降低了原表面缺陷与腐蚀表面的间距,一定 程度上降低了原未腐蚀表面缺陷处应力集中,从而 增加了疲劳裂纹萌生寿命,造成预腐蚀试样寿命在 中等应力范围内得到增加。由于低应力幅下(140 MPa) 损伤累积时间更长,高应力幅下(440 MPa)单次损 伤程度更大,导致预腐蚀对合金造成的初始损伤被



(a) 无腐蚀



(b) 预腐蚀8 h



(c) 预腐蚀16 h



(d) 预腐蚀24 h

图 9 中等应力下各预腐蚀试样裂纹萌生区

Fig. 9 Crack initiation zones of various pre-corrosion samples under median stress levels



图 10 腐蚀形貌演变示意图

Fig. 10 Schematic diagram of corrosion morphology

放大,使无腐蚀试样的寿命较预腐蚀试样更长。由 图 6 可知,16 h 预腐蚀疲劳极限较 8 h 和 24 h 高的 主要原因是其腐蚀形成的腐蚀形貌比 8 h 更均匀, 降低表面缺陷应力集中程度更高;由图 8 和图 9 可 知,24h预腐蚀试样腐蚀程度最高,蚀坑侵入基体 内部更深,增大了应力集中,使裂纹更易从腐蚀坑 处萌生。

4 结 论

1) 2195-T8 铝锂合金经 30% HNO₃ 腐蚀后的表面以晶间腐蚀和点蚀为主,随腐蚀时间延长,腐蚀程度加重,沿轧制方向形成条带状的腐蚀链,其产生与长条形晶界和金属间粒子连续腐蚀作用相关。

 2)预腐蚀 8,16,24 h 试样疲劳寿命在 210, 280,330 MPa 中等应力范围内寿命较无腐蚀增长。

3) 应力集中程度不同导致预腐蚀 8,16,24 h 试样在中等应力范围内疲劳寿命不敏感,2195-T8 铝锂合金经 30% HNO3 腐蚀消除了合金原表面上 划痕、微裂纹等缺陷,表面粗糙度增加,但是降低了 应力集中,延缓了疲劳裂纹萌生。

参考文献(References)

- [1] XU Y, WANG X J, YAN Z T, et al. Corrosion properties of lightweight and high-strength 2195 Al-Li alloy[J]. Chinese Journal of Aeronautics, 2011, 24(5): 681-686.
- [2] LIN Y, ZHENG Z Q, LI S C, et al. Microstructures and properties of 2099 Al-Li alloy[J]. Materials Characterization, 2013, 84: 88-99.
- [3] GOEBEL J, GHIDINI T, GRAHAM A J. Stress-corrosion cracking characterisation of the advanced aerospace Al-Li 2099-T86 alloy[J]. Materials Science and Engineering:A, 2016, 673: 16-23.
- [4] ABD EL-ATY A, XU Y, GUO X Z, et al. Strengthening mechanisms, deformation behavior, and anisotropic mechanical properties of Al-Li alloys: A review[J]. Journal of Advanced Research, 2018, 10: 49-67.
- [5] WANG X H, WANG J H, YUE X, et al. Effect of aging treatment on the exfoliation corrosion and stress corrosion cracking behaviors of 2195 Al–Li alloy[J]. Materials & Design, 2015, 67: 596-605.
- [6] CHARALAMPIDOU C, DIETZEL W, ZHELUDKEVICH M, et al. Corrosion-induced mechanical properties degradation of Al-Cu-Li (2198-T351) aluminum alloy and the role of side-surface cracks[J]. Corrosion Science, 2021, 183: 109330.
- [7] MA Y, ZHOU X, HUANG W, et al. Localized corrosion in AA2099-T83 aluminum–lithium alloy: The role of intermetallic particles[J]. Materials Chemistry and Physics, 2015, 161: 201-210.
- [8] DE SOUSA ARAUJO J V, MILAGRE M X, FERREIRA R O, et al. Exfoliation and intergranular corrosion resistance of the 2198 Al-Cu-Li alloy with different thermomechanical treatments[J]. Materials and Corrosion, 2020, 71(12): 1957-1970.
- [9] LIU J H, ZHAO K, YU M, et al. Effect of surface abrasion on pitting corrosion of Al-Li alloy[J]. Corrosion Science, 2018, 138: 75-84.
- [10] LEI X W, SAATCHI A, GHANBARI E, et al. Studies on pitting corrosion of Al-Cu-Li alloys part I: Effect of Li addition by microstructural, electrochemical, In-situ, and pit depth analysis[J]. Materials, 2019, 12(10): 1600.

- [11] ZHONG J, ZHONG S, ZHENG Z Q, et al. Fatigue crack initiation and early propagation behavior of 2A97 Al–Li alloy[J]. Transactions of Nonferrous Metals Society of China, 2014, 24(2): 303-309.
- [12] 薛喜丽,郑子樵,胡芳,等.时效制度对 2A97 铝锂合金疲劳裂纹 扩展速率的影响[J].稀有金属材料与工程,2016,45(12):3319-3324.

XUE X L, ZHENG Z Q, HU F, et al. Effect of aging conditions on the fatigue crack propagation rate of 2A97 aluminum-lithium alloy[J]. Rare Metal Materials and Engineering, 2016, 45(12): 3319-3324 (in Chinese).

- [13] 许罗鹏,曹小建,李久楷,等. 铝锂合金 2198-T8 高周疲劳性能及 其裂纹萌生机理[J]. 稀有金属材料与工程, 2017, 46(1): 83-89.
 XU L P, CAO X J, LI J K, et al. High cycle fatigue properties and crack initiation mechanisms of Al-Li 2198-T8 alloy[J]. Rare Metal Materials and Engineering, 2017, 46(1): 83-89 (in Chinese).
- [14] 范雪松,郑子樵,张龙,等. 2397 合金高周疲劳性能及裂纹萌生扩展行为[J]. 稀有金属材料与工程, 2017, 46(5): 1327-1333.
 FAN X S, ZHENG Z Q, ZHANG L, et al. High-cycle fatigue properties and crack initiation and propagation behavior of 2397 alloy[J]. Rare Metal Materials and Engineering, 2017, 46(5): 1327-1333 (in Chinese).
- [15] TIAN G, JIN G F, ZHANG W, et al. Investigation on electrochemical corrosion characteristics of 2A14 aluminum alloy in nitric acid[J]. Surface Review and Letters, 2017, 24(1): 1850016.
- [16] LIU D J, TIAN G, JIN G F, et al. Characterization of localized corrosion pathways in 2195-T8 Al–Li alloys exposed to acidic solution[J]. Defence Technology, 2023, 25: 152-165.
- [17] 刘涛,高怡斐,董莉,等.金属材料疲劳试验轴向应变控制方法:
 GB/T 26077—2010[S].北京:中国国家标准化管理委员会, 2011:
 5-6.

LIU T, GAO Y F, DONG L, et al. Metallic materials-fatigue testing-axial strain-controlled method: GB/T 26077—2010[S]. Beijing: Standardization Administration of China, 2011: 5-6.

- [18] 高舜之,何荣年.金属轴向疲劳试验方法:GB 3075-82[S].北京: 冶金部钢铁研究总院, 1982: 7-9.
 GAO S Z, HE R N. Method of axial force controlled fatigue testing of metals: GB 3075-82[S]. Beijing:Iron and Steel Research Institute of the Ministry of Metallurgy, 1982: 7-9.
- [19] 刘轩,刘慧丛,李卫平,等. 7075 铝合金在不同温度盐水环境中的 腐蚀疲劳行为[J]. 航空学报, 2014, 35(10): 2850-2856.
 LIU X, LIU H C, LI W P, et al. Corrosion fatigue behavior of 7075 aluminum alloy in saline water environment at different temperatures[J]. Acta Aeronautica et Astronautica Sinica, 2014, 35(10): 2850-2856 (in Chinese).
- [20] LUO C, ALBU S P, ZHOU X R, et al. Continuous and discontinuous localized corrosion of a 2xxx aluminum-copper-lithium alloy in sodium chloride solution[J]. Journal of Alloys and Compounds, 2016, 658: 61-70.
- [21] 李劲风,郑子樵,李世晨,等. 2195 铝-锂合金晶间腐蚀及剥蚀行 为研究[J]. 材料科学与工程学报, 2004, 22(5): 640-643.

LI J F, ZHENG Z Q, LI S C, et al. Study on intergranular corrosion and exfoliation corrosion behaviors of 2195 Al-Li alloy[J]. Journal of Materials Science and Engineering, 2004, 22(5): 640-643 (in Chinese).

[22] 孔祥,郑子樵,李劲风,等.时效制度对 2099 铝锂合金力学和应 力腐蚀性能的影响[J].稀有金属材料与工程, 2016, 45(12): 3271-3277.

KONG X, ZHENG Z Q, LI J F, et al. Effects of aging treatment on mechanical and stress corrosion properties of 2099 alloy[J]. Rare Metal Materials and Engineering, 2016, 45(12): 3271-3277 (in Chinese).

- [23] ZHANG X, ZHOU X, HASHIMOTO T, et al. The influence of grain structure on the corrosion behaviour of 2A97-T3 Al-Cu-Li alloy[J]. Corrosion Science, 2017, 116: 14-21.
- [24] HUGHES A E, BOAG A, GLENN A M, et al. Corrosion of AA2024-T3 Part II: Co-operative corrosion[J]. Corrosion Science, 2011, 53(1): 27-39.
- [25] LI M C, SEYEUX A, WIAME F, et al. Insights on the Al-Cu-Fe-Mn intermetallic particles induced pitting corrosion of Al-Cu-Li alloy[J]. Corrosion Science, 2020, 176: 109040.
- [26] SONG H P, LIU C C, ZHANG H, et al. In-situ SEM study of fatigue micro-crack initiation and propagation behavior in pre-corroded AA7075-T7651[J]. International Journal of Fatigue, 2020, 137: 105655.
- [27] MA M Y, ZHANG J Y, YI D Q, et al. Investigation of high-cycle fatigue and fatigue crack propagation characteristic in 5083-O aluminum alloy[J]. International Journal of Fatigue, 2019, 126: 357-368.
- [28] 马少华,回丽,周松,等. 腐蚀环境对预腐蚀铝合金腐蚀疲劳性能的影响[J]. 材料工程, 2015, 43(2): 91-95.
 MA S H, HUI L, ZHOU S, et al. Influence of corrosion environments on corrosion fatigue property of pre-corroded aluminum alloy[J]. Journal of Materials Engineering, 2015, 43(2): 91-95 (in Chinese).
- [29] 周松, 谢里阳, 回丽, 等. 航空铝合金预腐蚀疲劳寿命退化规律[J].
 东北大学学报 (自然科学版), 2016, 37(7): 969-973.
 ZHOU S, XIE L Y, HUI L, et al. Fatigue life degenerating rule of pre-corroded aviation aluminum alloy[J]. Journal of Northeastern University (Natural Science), 2016, 37(7): 969-973 (in Chinese).
- [30] 佘玲娟, 郑子樵, 钟申, 等. 6156-T62 铝合金的高周疲劳性能研究
 [J]. 稀有金属材料与工程, 2012, 41(7): 1201-1205.
 SHE L J, ZHENG Z Q, ZHONG S, et al. High cycle fatigue performance of 6156-T62 aluminum alloy[J]. Rare Metal Materials and Engineering, 2012, 41(7): 1201-1205 (in Chinese).
- [31] 章刚, 刘军, 刘永寿, 等. 表面粗糙度对表面应力集中系数和疲劳 寿命影响分析[J]. 机械强度, 2010, 32(1): 110-115.
 ZHANG G, LIU J, LIU Y S, et al. Effect of roughness on surface stress concentration factor and fatigue life[J]. Journal of Mechanical Strength, 2010, 32(1): 110-115 (in Chinese).

Research on pre-corrosion fatigue properties of 2195 Al-Li alloys in 30% HNO₃

LIU Dejun¹, TIAN Gan^{1,*}, LI Yulong², ZHANG Wei¹, ZHANG Youhong¹

School of Missile Engineering, Rocket Force University of Engineering, Xi'an 710025, China;
 China Academy of Launch Vehicle Technology, Beijing 100076, China)

Abstract: To simulate the material and application characteristics of liquid-missile propellant tanks, fatigue tests were carried out after pre-corrosion in 30% HNO₃ for different hours. Effects of pre-corrosion on fatigue properties were investigated by scanning electron microscopy (SEM), transmission electron microscope (TEM), and electron backscatter detector (EBSD) methods. The results show that the surface morphology after corrosion is mainly composed of intergranular corrosion and pitting. Stripped corrosion chains are produced along the rolling direction, which are related with corroded long grain boundaries and intermetallic particles. The fatigue limit values of pre-corroded specimens are lower than those of the uncorroded specimens. The fatigue limit of 16 h corroded specimens is the highest, while the limit of 24 h corroded specimens declines to zero. The fatigue cycles under median stress amplitudes of 210 MPa, 280 MPa and 330 MPa, higher than those of uncorroded specimens, and the increase of 16 h corroded specimens is the most obvious. The enhancement of fatigue cycles attributes to the elimination of microcracks, pores and the improvement of surface roughness through pre-corrosion, which reduces the stress concentration and retards the crack initiation.

Keywords: Al-Li alloys; propellant tank; pre-corrosion; fatigue properties; crack initiation

Received: 2022-05-31; Accepted: 2022-07-01; Published Online: 2022-07-19 16:42 URL: link.cnki.net/urlid/11.2625.V.20220719.0854.001

Foundation items: National Natural Science Foundation of China (52272446,52075541); Shaanxi Natural Science Foundation (2022JM-243)

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn DOI: 10.13700/j.bh.1001-5965.2022.0428

基于 AESL-GA 的 BN 球磨机滚动轴承故障诊断方法

王进花^{1,*},汤国栋¹,曹洁^{1,2},李亚洁¹

(1. 兰州理工大学电气工程与信息工程学院,兰州 730050; 2. 甘肃省制造信息工程研究中心,兰州 730050)

摘 要: 针对基于知识的贝叶斯网络 (BN) 构建方法存在不完全和不精确的缺点,提出一种基于知识引导和数据挖掘的 BN 结构构建方法。针对单一信号故障诊断结果不精确的问题和故障信息中存在的不确定性问题,将电流信号与振动信号融合建立 BN 的特征节点,分别提取 2 种信号的故障特征参数,利用区分度指标法进行特征筛选,将其作为 BN 结构特征层的节点。将专家知识构建的初始 BN 结构结合自适应精英结构遗传算法 (AESL-GA) 进行结构优化,通过自适应限制进化过程中的搜索空间,减少自由参数的数量,提高其全局搜索能力,得到最优 BN 结构。通过MQY5585 溢流型球磨机滚动轴承实测数据和 Paderborn University 轴承数据集对所提方法进行验证,结果证明了所提方法的有效性。

关 键 词: 贝叶斯网络; 故障诊断; 自适应精英结构遗传算法; 滚动轴承; 信号融合 中图分类号: TP277; TD453

文献标志码: A 文章编号: 1001-5965(2024)04-1138-09

冶金行业作为中国实现全面工业化过程中的 一个重要组成部分,对国家的工业经济发展有着重 要的意义。由于冶金设备常年工作在较为恶劣的 环境中,设备经常发生各类故障导致停机,严重时甚 至会发生安全事故。因此,对设备进行故障诊断变 得尤为重要^[1-2]。

在冶金选矿过程中,使用球磨机进行物料研磨 是整个冶炼过程中尤为重要的一环,其中,球磨机 滚动轴承的运行稳定性直接影响着冶炼厂的经济 效益,因此,对滚动轴承建立科学有效的故障诊断 机制是极为必要的。随着智能诊断技术的快速发 展,基于信号处理和深度学习结合的诊断方法获得 了广泛研究。Ben Ali等^[3]针对滚动轴承振动信号 的非平稳、非线性特点,提出基于经验模态分解 (empirical mode decomposition, EMD)能量熵的特征 提取方法,将选择的特征用于训练人工神经网络 (artificial neural network,ANN)对轴承故障进行分 类。Zhang等^[4]提出一种基于全卷积神经网络的模 型,在频域中挖掘故障信息,实现了故障分类。Xu 等^[5]通过连续小波变换将时域振动信号转换为二 维灰度图像,利用包含局部和全局信息的多级特征 来进行轴承的故障诊断。以上方法均采用振动信 号进行故障诊断研究,然而球磨机滚动轴承发生故 障时,其原因可能是多方面的,并且故障征兆与故 障之间存在耦合性和不确定性,单一信号源不能完 全反映所有故障,且实际工程中球磨机滚动轴承的 故障数据难以获取,故障样本少,无法训练获得精 确的深度学习模型。

贝叶斯网络 (Bayesian network, BN) 是解决复杂 系统不确定性诊断推理的有效方法之一^[6-8],其能够 结合多种渠道的信息进行推理诊断。近年来, BN 被越来越多地应用于故障诊断领域中。Sahu 和 Palei^[9]基于专家经验及设备日志记录对重型土方 机械牵引系统建立原因、症状和故障之间的因果关 系,利用故障推理进行诊断。Wang 等^[10]提出一种 基于 BN 的柴油机喷射系统故障诊断方法,根据文

*通信作者. E-mail: wjh0615@lut.edu.cn

引用格式: 王进花, 汤国栋, 曹洁, 等. 基于 AESL-GA 的 BN 球磨机滚动轴承故障诊断方法 [J]. 北京航空航天大学学报, 2024, 50 (4): 1138-1146. WANG J H, TANG G D, CAO J, et al. Fault diagnosis method of BN ball mill rolling bearing based on AESL-GA[J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50 (4): 1138-1146 (in Chinese).

收稿日期:2022-05-28;录用日期:2022-07-10;网络出版时间:2022-09-05 10:50 网络出版地址: link.cnki.net/urlid/11.2625.V.20220905.0837.001

基金项目:国家重点研发计划(2020YFB1713600);国家自然科学基金(62063020);甘肃省自然科学基金(20JR5RA463)

献和领域专家的实践经验,使用 BN 中的节点来参数化各类故障和症状。Xiao 等^[11]构造用于牵引变 压器故障诊断的 BN,并对 4 种常见类型的故障进 行诊断,表现出较好的诊断效果。同时, BN 还被应 用于风力发电机、燃油系统等领域^[12-15]。目前,实 际工程应用案例中多采用专家知识构造 BN 结构, 由于专家知识存在的主观性和故障信息的不完善 等缺点,会导致模型与实际系统存在偏差,而用于 结构学习的算法,存在收敛过慢和容易陷入局部最 优的缺点。

基于以上分析,本文将专家知识和历史数据的 建模方法相结合,提出一种多特征融合的 BN 故障 诊断模型,采用电流和振动 2 种信号进行特征参数 的提取,基于专家知识构建初始模型,引入自适应精 英结构遗传算法 (adaptive elite-based structure learner using genetic algorithm, AESL-GA) 实现网络结构的 优化,并通过 BN 推理完成故障诊断。

1 贝叶斯网络

BN 是一种基于概率计算的描述因果关系的推 理模型^[16],其基本组成为:表示变量和变量之间因 果关系的有向无环图(directed acyclic graph, DAG) 及表示变量之间概率参数的条件概率表(conditional probability table, CPT)。图 1 为一个简单的 BN 模型 拓扑结构, DAG 由节点 $X = \{X_1, X_2, X_3, X_4\}$ 和有向连 接边组成。CPT 表示该节点相对于其父节点的条 件概率分布。BN 能够明确表示复杂系统中各因素 间的依赖关系,并进行诊断推理。



图 1 BN 模型简略图 Fig. 1 Brief diagram of BN model

BN 的诊断是一个基于先验知识计算后验概率的推理过程^[17]。贝叶斯推理的基本公式如下:

$$P(A) = \sum_{i=1}^{n} P(B_i) P(A|B_i)$$
(1)

$$P(B_i|A) = \frac{P(B_i)P(A|B_i)}{\sum_{j=1}^{n} P(B_j)P(A|B_j)}$$
(2)

式中: P(A)为事件 A 发生的先验概率; P(A|B_i)为事件 B_i发生的前提下事件 A 发生的条件概率。基于式 (1)和式 (2),结合先验概率及新获得的证据,可计算求解目标的后验概率。

BN故障诊断模型建立的难点在于网络结构的

构建,目前广泛应用的结构构建方法有专家知识^[18] 和基于数据的方法^[19-22],但专家知识的不完善和主 观性会影响最终模型的准确度,而基于数据的方法 在结构学习时以随机生成的结构作为初始结构进 行学习,会降低结构学习的效率。因此,本文利用 专家经验建立初始结构,通过限制搜索范围提升收 敛速度,以此来提高 BN 结构学习的效率。

2 特征提取

在故障诊断中,通常采用振动信号作为诊断信 息源,但振动信号容易受到噪声和环境干扰,并且 难以精确识别部分电气故障^[23]。当机械设备发生 故障时,电机转子运动过程中的气隙磁场会发生变 化,导致磁通密度发生改变,进而影响三相电流。 因此,采用振动信号和电流信号结合进行诊断将更 为精确。对电流和振动信号分别进行时域、频域特 征提取,时域信号能够直观反映信号的变化状态, 对不同类型故障可做出准确判断,分析频域信号成 分的组成和大小,可判别设备初期微弱故障^[24]。针 对球磨机及其运行工况,分别对电流信号和振动信 号选取时域、频域中的10个特征参数。

设所采集信号序列为 $\{x_i\}(i = 1, 2, \dots, N), N$ 为采 样点数,平均值为 $\bar{x},$ 标准差为 σ 。定义 $\{x_i\}$ 中极大值 为 $\{x_{mj}\}(j = 1, 2, \dots, N_m), N_m$ 为 $\{x_i\}$ 中极大值总数, $\{x_{mj}\}$ 的平均值为 \bar{x}_m ,标准差为 σ_m 。定义 $\{x_i\}$ 中极小 值为 $\{x_{kn}\}(n = 1, 2, \dots, N_k), N_k$ 为 $\{x_i\}$ 中极小值总数, $\{x_{kn}\}$ 的平均值为 \bar{x}_k ,标准差为 σ_k 。设 $\{f_i\}(i = 1, 2, \dots, I)$ 为频率序列, I为采样频率的一半, $F(f_i)$ 为 f_i 的频 谱值。拟提取特征参数 $S_1 \sim S_{10}$ 如下:

$$S_{1} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} x_{i}^{2}}$$
(3)

$$S_2 = \frac{1}{N_{\rm m}} \sum_{j=1}^{N_{\rm m}} \left| x_{\rm mj} \right| \tag{4}$$

$$S_{3} = \left| \frac{1}{N\sigma^{3}} \sum_{i=1}^{N} (x_{i} - \overline{x})^{3} \right|$$
(5)

$$S_4 = \frac{1}{\sigma^4} \sum_{i=1}^{N} (x_i - \bar{x})^4$$
 (6)

$$S_{5} = \left| \frac{1}{\sigma_{m}^{4}} \sum_{j=1}^{N_{m}} (x_{mj} - \bar{x}_{m})^{4} \right|$$
(7)

$$S_{6} = \sqrt{\sum_{i=1}^{l} f_{i}^{4} F(f_{i}) / \sum_{i=1}^{l} f_{i}^{2} F(f_{i})}$$
(8)

$$S_{7} = \frac{\sum_{i=1}^{I} f_{i}^{2} F(f_{i})}{\sqrt{\sum_{i=1}^{I} F(f_{i}) \sum_{i=1}^{I} f^{4} F(f_{i})}}$$
(9)

$$S_8 = \sum_{i=1}^{l} F(f_i)$$
 (10)

$$S_{9} = \sqrt{\sum_{i=1}^{I} F^{2}(f_{i})}$$
(11)

$$S_{10} = \left| \frac{1}{\sigma_{k}^{4}} \sum_{n=1}^{N_{k}} (x_{kn} - \overline{x}_{k})^{4} \right|$$
(12)

为防止出现特征冗余导致 BN 建模难度增大, 影响最终诊断精度的情况,本文采用区分度指标对 提取的特征量进行筛选,得到对故障变化敏感的特征。区分度指标^[25]D定义为

$$D = \frac{|\mu_1 - \mu_2|}{\sqrt{\sigma_1^2 + \sigma_2^2}}$$
(13)

式中: μ_1 、 μ_2 分别为某一特征量对应 2 种状态下的 平均值; σ_1 、 σ_2 为对应的标准差。D 值越大, 表明 该特征参数对这 2 种状态转化的识别越灵敏。

本文球磨机共有3种运行工况C₁~C₃,对电流 信号和振动信号中提取的10个特征参数进行区分 度指标计算,振动信号特征参数D值结果如表1所 示,电流信号特征参数D值结果如表2所示。

根据表 1 和表 2 所计算出的特征参数 D 值, 最 终选取 D 值较大的振动信号特征参数中 S_1 、 S_2 、 S_6 、 S_7 作为 BN 特征层中的 P_1 、 P_2 、 P_3 、 P_4 节点; 电流信号 特征参数中的 S_1 、 S_6 作为 BN 特征层中的 P_5 、 P_6 节点。

|--|

运行工况	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	<i>S</i> ₉	S ₁₀
C_1	4.82	3.21	1.77	1.69	1.46	3.44	4.95	1.38	2.15	1.58
C_2	3.53	2.37	1.21	1.22	1.80	3.27	3.62	1.44	2.04	2.25
C_3	2.51	2.89	2.07	2.73	0.65	2.11	2.20	2.04	1.33	1.31

表 2 电流信号特征参数 D 值

 Table 2
 Characteristic parameter D value of current signal

运行工况	S_1	S_2	S ₃	S_4	<i>S</i> ₅	S_6	<i>S</i> ₇	S_8	S_9	S_{10}
C_1	3.65	2.16	1.87	1.22	1.18	3.31	2.06	1.42	1.14	1.78
C_2	2.85	1.69	1.59	2.06	2.25	2.56	1.71	2.16	2.17	1.96
C_3	2.14	0.31	2.76	1.14	0.97	2.94	1.95	1.97	1.56	0.65

基于专家知识和 AESL-GA 的 BN 故障诊断模型

3.1 BN 故障诊断模型节点选取和初始结构确定

基于 BN 的球磨机滚动轴承的故障诊断中要综 合考虑设备故障的类型、位置、征兆及设备的运行 工况。故障诊断模型是否符合该研究对象的工作 机制和故障机理严重影响诊断模型的正确性和最 终的诊断精度。用于故障诊断的 BN 结构一般为 3 层结构,即运行工况层、故障层、故障征兆层。运 行工况层节点为设备的运行状况,故障层节点为设 备发生的故障种类,故障征兆层节点为筛选出的特 征参数。图 2 为金川公司 MQY5585 溢流型球磨机 滚动轴承通过专家知识构建的初始 BN 结构。图 中: *C_i*为球磨机运行工况, *P_i*为第 2 节中筛选的电 流和振动信号特征层节点, *F_i*为故障类型,包括正 常状态、内圈故障、外圈故障。



冬	2	基于专家知识的球磨机滚动轴承初始 BN 结构
Fig. 2	Ι	nitial BN structure of ball mill rolling bearing based on
		expert knowledge

3.2 AESL-GA

目前,常用构建 BN 结构的算法有 K2 评分算 法、PC 算法等。但这类算法在结构搜索时以全局 节点为搜索范围,导致网络结构收敛缓慢。AESL-GA^[26]以遗传算法为基础,改进了遗传算法中节点 之间因果关系构建时父节点的限制,并且在每一代 中,收集适应度高于预定阈值的个体作为精英集 合,通过这种约束提高了全局搜索能力,同时提升 了 BN 结构学习效率。

假设一个 BN 故障诊断模型中有 n 个特征节点 和 m 个故障节点,则可将 BN 结构表示为矩阵 $R = [r_{i,j}]$ (i = 1, 2, ..., n; j = 1, 2, ..., m),如式 (14) 所 示。如果症状节点 S_i 与故障节点 F_j 存在因果关系, $r_{i,j} = 1$,否则 $r_{i,j} = 0$ 。

$$\boldsymbol{R} = \begin{bmatrix} 0 & r_{1,2} & \cdots & r_{1,m} \\ r_{2,1} & 0 & \cdots & r_{2,m} \\ \vdots & \vdots & & \vdots \\ r_{n,1} & r_{n,2} & \cdots & r_{n,m} \end{bmatrix}$$
(14)

根据式 (14),将矩阵转换为一维字符串 { $r_{1,1},...,r_{1,m},r_{2,1},...,r_{2,m},...,r_{n,1},...,r_{n,m}$ },该字符串 被编码为初始种群中个体的基因。对于具有n个 特征节点和m个故障节点的 BN,一个基因的长度 为n+m。在进行结构优化的过程中执行选择、交 叉和变异操作:①选择操作,本文选择适合用于 BN 结构选择的轮盘赌方法进行选择操作,以此来 保证更优的 BN 结构被选择;②交叉操作,交叉过 程采取 2 个个体交叉,即 DAG 中有向边的交叉操 作;③变异操作,根据 BN 结构,结合变异率如式(15) 所示,随机选择一个位置(i, j),使 $r_{i,j}$ 由 1 变 0 或由 0 变 1(即对网络进行加边、减边、反向操作)。

$$\mu_{k,j} = \left[\varepsilon + (1-\varepsilon)\left(1 - \frac{f(x_k)}{f_{\max}}\right)\right] \left[\varepsilon + (1-\varepsilon)\left(1 - w_{i,j}\right)\right]$$
(15)

式中: $\mu_{k,j}$ 为每个在 j位点的个体 x_k 发生突变的概率; $w_{i,j}$ 为等位基因 i 在精英集合中位置 j 出现的概率; $f(x_k)$ 为当前个体的适应度; f_{max} 为最佳个体的适应度; ε 为一个很小的正数, 以避免出现零概率。 网络结构的评估中, 采用了一种混合评分函数^[27], 其由 2 部分组成, 如式 (16) 所示, 第1 部分位 BDeu 评分函数, 第 2 部分为惩罚项。

$$f_{\max}(G|D) = f_{\text{BDeu}}(G|D) + k \sum \lg \left[f\left(r_{i,j}, q_{i,j}\right) \right] \quad (16)$$

式中:k为衡量专家知识贡献大小的系数,可通过多次实验确定; $q_{i,j}$ 为 $r_{i,j}$ 所对应的概率专家知识矩阵Q中的元素。

本文适应度函数采用式 (17) 所示的 BDeu 评分 函数:

$$f_{\text{BDeu}}(G|D) = \lg P(G) + \sum_{i=1}^{n+m} \sum_{j=1}^{q_i} \left\{ \lg \left[\frac{\Gamma(\alpha/q_i)}{\Gamma(\alpha/q_i + N_{ij})} \right] + \sum_{k=1}^{r_i} \lg \left[\frac{\Gamma(\alpha/(r_iq_i) + N_{ijk})}{\Gamma(\alpha/(r_iq_i))} \right] \right\}$$
(17)

式中: Γ 为 Gamma 函数; r_i 为节点 X_i 可能的状态数;

 q_i 为 X_i 的父节点可能的联合赋值个数,对于一个症状节点, q_i 为其父故障节点可能的联合分配的数量,对于故障节点, q_i 为0; N_{ijk} 为符合 X_i 取第k个状态,其父结点取第j个联合赋值条件的数据样本数; N_{ij} 可以用 $\sum_{k} N_{ijk}$ 计算; α 为一个表示等效样本大小的超参数;P(G)为当前 BN 结构G的先验概率。通过式(17)所示函数计算个体适应度,收集适应度值高于预定义阈值的个体作为精英集合,并且将最优秀的个体传播到下一代种群。

在进行结构学习时, 父节点的数量增大, 会增加结构学习的计算负担。因此, 可以将最大父节点数 (maximum number of parents, MP)限制为高于但接近或等于网络中实际最大父节点数。设 $M' = \{M_i^t\}, i = 1, 2, \cdots, n$ 为第 t代种群的 MP 集合, 式 (18)表示了 M_i' 是如何通过 M_i^{i-1} 和一个时变参考阈值 H_i' 动态限制的, 其中, 安全阈值构成了第 t代任何节点 i 的父节点集的参考大小, 在设置 H_i' 时, 可认为精英个体的父辈群体被认为比其他个体更接近 真实的群体。在任何情况下, 动态阈值 M_i' 都不允许 超过最大阈值 M_{max} 。

$$\begin{cases} H_{i}^{t} = \left| \bigcup_{x \in \xi^{t}} P_{ax,i}^{t} \right| + 1 \\ M_{i}^{t} = \begin{cases} M_{i}^{t-1} - 1 & M_{i}^{t-1} > H_{i}^{t} \\ M_{i}^{t-1} + 1 & M_{i}^{t-1} < H_{i}^{t}, M_{i}^{t-1} < M_{\max} \\ M_{i}^{t-1} & \\ \end{bmatrix} \end{cases}$$
(18)

式中: *P*^{*i*}_{ax,t}为第 *t* 代个体节点的第 *i* 个节点的父集; *ξ*^{*i*}为第 *t* 代精英集合。

本文使用一种概率专家知识,以此减少专家知识的主观不确定性带来的影响,并将其代替 AESL-GA 中随机生成的初始矩阵,以此减少结构学习的时间,提高结构寻优的效率。概率专家知识矩阵 $Q = [q_{i,j}](i = 1, 2, ..., n; j = 1, 2, ..., m)表示初始 BN 结构中特征参数与故障类型之间存在因果关系的可能性大小。确定 BN 结构的基本策略是:<math>q_{i,j}$ 越大, $r_{i,j}$ 为1的可能性越高,即专家建议连接i和j。 $r_{i,j}$ 的大小根据式(19)确定:

$$r_{i,j,\text{prior}} = \begin{cases} 1 & q_{i,j} \ge a_{i,j} \\ 0 & q_{i,j} < a_{i,j} \end{cases}$$
(19)

式中: ai,j为规定数。

在融合专家知识的 BN 结构学习中, 根据概率 专家知识, 在结构学习的过程中需遵循: 增加概率 专家知识中可能性较大的有向边被添加的概率; 降 低概率专家知识中可能性较大的有向边被删除的 概率; 对无概率专家知识的节点之间按照原始 BDeu 评分函数进行处理。

3.3 算法流程

基于专家知识和 AESL-GA 的 BN 结构学习步骤如下:

步骤1 对目标系统的故障机理进行分析,确 定节点变量及网络层数,并收集研究对象故障诊断 相关领域的专家知识。

步骤 2 根据概率专家知识矩阵 Q,结合历史数据通过 AESL-GA 进行结构优化,构建无环 BN 结构。

步骤 3 通过式 (16) 混合评分函数进行 BN 结构评分,并判断是否满足终止条件,即是否达到连续 N 代评分值最高。如果满足终止条件,则可以输出该最优结构;如果不满足则进行步骤 4。

步骤 4 通过自适应方法,对群体进行选择、 交叉、变异处理,检查是否存在环形结构,并修正存 在的环形结构,产生新种群,并继续进行步骤 3。

图 3 为 BN 故障诊断模型构建流程。





4 实验结果与分析

本文采用金川公司投入运行的球磨机实测数据和由 Lessmeier 等^[28]提供的德国 Paderborn University 轴承数据集分别进行实验验证,证明本

文方法的有效性和泛化能力。

4.1 实训数据

实验采用金川公司生产车间运行的球磨机滚 动轴承实测数据。根据球磨机研磨介质运动规律 及不同磨矿阶段各物料运动规律,将球磨机分为 3种不同工况:①粗磨,转速为78%~84%的临界转 速;②细磨,转速为72%~78%的临界转速;③超细 磨,转速为68%~72%的临界转速。对3种工况时 球磨机滚动轴承的振动信号和电流信号进行采集, 采样频率为64 kHz,按照设备历史故障类型将其故 障状态分为:内圈故障(IF)、外圈故障(OF)、正常状 态(Normal)3种。取每种故障下振动信号和电流信 号各100个样本,每个样本信号点为1000,共600个 样本,其中,训练样本400个,测试样本200个。利 用本文方法,其 BN 故障诊断模型结构如图4 所示。

对该模型进行推理测试,结果如表3所示。 3次测试都表明最大后验概率的推理结果是正确的。



图 4 基于本文方法的球磨机 BN 故障诊断模型 Fig. 4 BN fault diagnosis model of ball mill based on the proposed method

表 3 推理测试结果

	14	ble 5 Reasoning te	st results	
测试 序号	测试故障	与故障相关的 异常征兆	故障类型	后验概率%
1	内圈	$P_3(高), P_4(高), P_5(低)$	内圈故障	94.25
2	外圈	$P_1(高), P_4(低), P_6(高)$	外圈故障	92.18
3	正常	$P_1(高), P_2(高)$	正常状态	97.42

通过本文方法构建的球磨机滚动轴承 BN 故障 诊断模型由 3 种工况、6 个特征量、3 种运行故障状 态构成,其故障识别率如表 4 所示。

由故障识别结果可知,3种运行工况下故障识 别率均达到97%以上。采用混淆矩阵的方法可视 化工况 C₃下的诊断结果如图 5 所示。混淆矩阵的 纵坐标表示真实的故障类别,本文坐标表示预测的 故障类别,主对角线上为每种故障状态下的诊断 精度。

在同等条件下,分别与文献 [29] 中的卷积神经 网络 (convolutional neural network, CNN)、长短期记

表 4 基于本文方法的球磨机数据故障识别率

I able 4	Fault re	ecognition rate of ball r	nill data based on the		
proposed method					
Τi	况	故障类型	识别率%		
		正常状态	98.43		
C_1	1	内圈故障	97.72		
		外圈故障	98.55		
		正常状态	98.56		
С	2	内圈故障	97.83		
		外圈故障	98.43		
		正常状态	98.66		
C	3	内圈故障	98.17		
		外圈故障	98.50		



Fig. 5 Confusion matrix under ball mill data

忆网络(long short-term memory networks, LSTM)、 仅使用振动信号的贝叶斯网络(vibration signal-Bayesian network, V-BN)及K2-BN方法进行对比实 验,结果如图 6 所示。实验结果表明,仅使用振动 信号的 V-BN诊断模型的识别率要远低于本文方 法,将电流信号中提取到的故障特征量加入网络, 提高了网络的故障识别能力。由于 K2 算法的输入





Fig. 6 Comparison of diagnostic accuracy and running time of different models under ball mill data

对节点顺序的要求极高, K2-BN 方法的诊断效果较差, 且结构学习所需的时间较长。LSTM 方法的诊断准确率仅达到 92.48%, CNN 方法的诊断识别率相较于 V-BN 方法有所提高, 但依旧低于本文方法, 表明了本文方法的优越性。

4.2 Paderborn University 轴承数据集

Paderborn University 轴承数据集提供了轴承 2种损伤部位和3种损伤方式。大多数轴承故障诊 断研究将故障划分为内圈故障、外圈故障和正常状态3种类型,本文将内圈故障、外圈故障按不同损 伤方式细化,故障描述如表5所示。选择同一损伤 程度下的故障数据样本,每种故障分别在4种工况 下测得,每组数据由2种电流信号和1种振动信号 组成,所有信号都以相同的采样频率采样(F_s= 64kHz)。取每种故障各工况下的振动和电流各100个 样本,每个样本包含1000个采样点,共1800组样 本,其中,训练样本1400组,测试样本400组。

表 5 数据集故障描述 Table 5 Dataset fault description

- ***	
故障节点标签	故障类型
F_0	正常(Nor)
F_1	外圈钻孔(OD)
F_2	内圈电雕刻(IE),外圈电雕刻(OE)
F_3	内圈放电沟槽 (IEDM), 外圈放电沟槽(OEDM)
F_4	外圈塑化(OP)
F_5	内圈点蚀(IFP),外圈点蚀(OFP)

首先,对该数据集中无用的错误数据进行剔除,由于空缺数据较少,采用均值填充法进行空缺数据补缺。然后,分别对每种工况下的电流、振动信号进行特征提取。最后,对提取的多个时域、频域特征进行区分度指标识别,得到最敏感的9个特征,经离散化后作为故障征兆节点和故障位置节点。表6为特征筛选结果。同时,构建基于专家知识的 Paderborn University 轴承数据集初始 BN 结构,如图7所示。

基于上述数据处理结果,利用本文方法,构建 Paderborn University 轴承数据集下的 BN 故障诊断 模型,如图 8 所示。分别对 4 种工况下的故障进行 诊断,其识别率如表 7 所示。

采用混淆矩阵的方法可视化工况 C₄下的分类 结果如图 9 所示。结果表明,本文方法对滚动轴承 各类故障均有良好的诊断效果,在该工况下的故障 识别率都达到了 90% 以上,其中, Nor 与 OD 故障 的识别率达到了 97% 以上。

将不同运行工况下的诊断结果进行对比,如

表 6 特征量筛选结果					
Table 6	Feature selection results				
故障节点标签	筛选出的特征				
F_0	$S_1, S_3, S_4, S_5, S_8, S_9$				
F_1	S_1, S_2, S_3				
F_2	S_3, S_4, S_9				
F_3	S_1, S_4, S_6				
F_4	S_1, S_4, S_7, S_8				
F_5	S_5, S_7, S_8, S_9				



图 7 基于专家知识的 Paderborn University 轴承数据集初始 BN 结构

Fig. 7 Paderborn University bearing dataset Initial BN structure of based on expert knowledge



图 8 基于本文方法在 Paderborn University 数据集下的 BN 故障诊断模型

Fig. 8 BN fault diagnostic model under Paderborn University dataset based on the proposed method

图 10 所示。在不同运行工况下,各类故障的诊断 效果不尽相同,其中,工况4下的诊断识别率要优 于其他3种,由此可见,机械设备的运行工况不同 会对诊断结果造成一定误差,单一工况下的故障诊 表7 Paderborn University数据集下4种工况条件时

的故障识别率 Table 7 Fault recognition rate under 4 work

able 7	Fault r	ecognition	rate under	4 working	
condit	tions in 1	Paderborn	University	dataset	%

			v		
故障类型	C_1	C_2	C_3	C_4	
正常	97.86	97.73	97.77	97.83	
外圈钻孔	95.88	96.18	96.43	97.44	
内圈电雕刻	96.85	96.28	96.79	96.94	
外圈电雕刻	96.89	96.46	96.83	96.78	
内圈放电沟槽	96.54	95.93	96.87	96.44	
外圈放电沟槽	96.58	96.63	96.40	96.72	
外圈塑化	87.38	87.13	90.92	91.83	
内圈点蚀	89.68	88.39	88.57	90.44	
外圈点蚀	89.78	89.58	89.65	91.83	



图 9 基于 Paderborn University 数据集下的混淆矩阵





图 10 不同工况下各类故障诊断识别率对比

Fig. 10 Comparison of various fault diagnosis accuracy rates under different working conditions

断会降低故障诊断的效果。

为进一步验证本文方法的有效性,在同等条件 下分别与文献 [29] 中的 CNN、LSTM、V-BN 及 K2-BN 方法进行对比实验,结果如图 11 所示。由实验 结果可知,本文方法优于另外 4 种方法,进一步证



图 11 Paderborn University 数据集下不同模型的 诊断精度及运行时间

Fig. 11 Diagnostic accuracy and running time of different models under Paderborn University dataset

明了本文方法的有效性,表明本文方法可有效应用 于球磨机滚动轴承故障诊断。

5 结 论

1)本文方法结合专家知识与历史运行数据,通 过 AESL-GA 确定了 BN 结构,有效解决了由于故 障信息耦合性造成的故障诊断模型构建困难的 问题。

2)由于振动信号对部分故障的识别不够灵敏, 分别对振动和电流信号进行特征提取,将其用于 BN 故障诊断模型特征层的构建中。

3)通过 Paderborn University 轴承数据集和金 川公司球磨机实测数据对本文方法进行验证,并与 CNN 方法、LSTM 方法、K2-BN 方法和 V-BN 方法 进行对比,结果显示,本文方法有更高的诊断识别 率,可以有效诊断耦合多特征因素的故障类型,实 现球磨机滚动轴承的高效故障诊断。

参考文献(References)

- [1] DAI X W, GAO Z W. From model, signal to knowledge: A datadriven perspective of fault detection and diagnosis[J]. IEEE Transactions on Industrial Informatics, 2013, 9(4): 2226-2238.
- BARTLETT L M, HURDLE E E, KELLY E M. Integrated system fault diagnostics utilising digraph and fault tree-based approaches
 [J]. Reliability Engineering & System Safety, 2009, 94(6): 1107-1115.
- [3] BEN ALI J, FNAIECH N, SAIDI L, et al. Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals[J]. Applied Acoustics, 2015, 89(9): 16-27.
- [4] ZHANG W D, ZHANG F, CHEN W, et al. Fault state recognition of rolling bearing based fully convolutional network[J]. Computing in Science & Engineering, 2019, 21(5): 55-63.
- [5] XU G W, LIU M, JIANG Z F, et al. Bearing fault diagnosis method based on deep convolutional neural network and random forest ensemble learning[J]. Sensors, 2019, 19(5): 1088.
- [6] PEARL J. Bayesian networks: A model of self-activated memory for evidential reasoning[C]//Proceedings of the 7th Conference of the Cognitive Science Society. Irvine: University of California, 1985: 15-17.
- [7] CAI B P, LIU H L, XIE M. A real-time fault diagnosis methodology of complex systems using object-oriented Bayesian networks
 [J]. Mechanical Systems and Signal Processing, 2016, 80(1): 31-44.
- [8] 鞠萍华, 柯磊, 冉琰, 等. 基于动作单元的机电产品故障溯源诊断 方法[J]. 湖南大学学报 (自然科学版), 2020, 47(2): 60-66.
 JU P H, KE L, RAN Y, et al. Method of fault root causes tracing analysis for electro-mechanical products based on action unit[J].
 Journal of Hunan University (Natural Sciences), 2020, 47(2): 60-66(in Chinese).
- [9] SAHU A R, PALEI S K. Real-time fault diagnosis of HEMM using Bayesian network: A case study on drag system of dragline[J]. Engineering Failure Analysis, 2020, 118(12): 104917.

- [10] WANG J X, WANG Z W, STETSYUK V, et al. Exploiting Bayesian networks for fault isolation: A diagnostic case study of diesel fuel injection system[J]. ISA Transactions, 2019, 86(5): 276-286.
- [11] XIAO Y, PAN W G, GUO X M, et al. Fault diagnosis of traction transformer based on Bayesian network[J]. Energies, 2020, 13(18): 4966.
- [12] CAI B P, LIU Y H, FAN Q, et al. Multi-source information fusion based fault diagnosis of ground-source heat pump using Bayesian network[J]. Applied Energy, 2014, 114(2): 1-9.
- [13] CHEN B, XIE L, LI Y Z, et al. Acoustical damage detection of wind turbine yaw system using Bayesian network[J]. Renewable Energy, 2020, 160(7): 1364-1372.
- [14] 朱彦沛, 邢宁哲, 纪雨形, 等. 基于交互式主动探测的电力综合数 据网故障定位算法[J]. 电力系统自动化, 2017, 41(4): 35-40. ZHU Y P, XING N Z, JI Y T, et al. Fault location algorithm of integrated data network for power system based on interactive active detection[J]. Automation of Electric Power Systems, 2017, 41(4): 35-40(in Chinese).
- [15] 王金鑫, 王忠巍, 马修真, 等. 柴油机燃油系统多故障的解耦与诊断技术[J]. 控制与决策, 2019, 34(10): 2249-2255.
 WANG J X, WANG Z W, MA X Z, et al. Decoupling and diagnosis of multi-fault of diesel engine fuel system[J]. Control and Decision, 2019, 34(10): 2249-2255(in Chinese).
- [16] CAI B P, HUANG L, XIE M. Bayesian networks in fault diagnosis[J]. IEEE Transactions on Industrial Informatics, 2017, 13(5): 2227-2240.
- [17] 刘浩然, 马明, 李世昭, 等. 一种应用于故障诊断中的高效推理算法[J]. 控制与决策, 2015, 30(11): 2033-2040.
 LIU H R, MA M, LI S Z, et al. An effective inference algorithm for fault diagnosis[J]. Control and Decision, 2015, 30(11): 2033-2040(in Chinese).
- [18] 王双成, 唐海燕, 刘喜华. 用于风险管理的贝叶斯网络学习[J]. 控 制与决策, 2007, 22(5): 569-572.
 WANG S C, TANG H Y, LIU X H. Learning Bayesian networks in risk management[J]. Control and Decision, 2007, 22(5): 569-572(in Chinese).
- [19] 王海羽,刘浩然,张力悦,等. 基于节点块序列约束的局部贝叶斯 网络结构搜索算法[J]. 自动化学报, 2020, 46(6): 1210-1219.
 WANG H Y, LIU H R, ZHANG L Y, et al. Local Bayesian network structure searching using constraint of node chunk sequence
 [J]. Acta Automatica Sinica, 2020, 46(6): 1210-1219(in Chinese).
- [20] 刘浩然,孙美婷,李雷,等. 基于蚁群节点寻优的贝叶斯网络结构 算法研究[J]. 仪器仪表学报, 2017, 38(1): 143-150.
 LIU H R, SUN M T, LI L, et al. Study on Bayesian network structure learning algorithm based on ant colony node order optimization[J]. Chinese Journal of Scientific Instrument, 2017, 38(1): 143-150(in Chinese).
- [21] HUANG W C, KOU X Y, ZHANG Y, et al. Operational failure analysis of high-speed electric multiple units: A Bayesian network-K2 algorithm-expectation maximization approach[J]. Reliability Engineering & System Safety, 2021, 205(2): 107250.
- [22] SPIRTES P, GLYMOUR C, SCHEINES R. Causation, prediction, and search[M]. Cambridge: The MIT Press, 2001.
- [23] LIU M K, TRAN M Q, WENG P Y. Fusion of vibration and cur-

rent signatures for the fault diagnosis of induction machines[J]. Shock and Vibration, 2019, 2019: 7176482.

[24] 赵志宏. 基于振动信号的机械故障特征提取与诊断研究[D]. 北京: 北京交通大学, 2012.
 ZHAO Z H. Research on vibration signal based machinery fault fea-

ture extraction and diagnosis[D]. Beijing: Beijing Jiaotong University, 2012 (in Chinese).

- [25] LI K, CHEN P, WANG S M. An intelligent diagnosis method for rotating machinery using least squares mapping and a fuzzy neural network[J]. Sensors, 2012, 12(5): 5919-5939.
- [26] CONTALDI C, VAFAEE F, NELSON P C. Bayesian network hybrid learning using an elite-guided genetic algorithm[J]. Artificial Intelligence Review, 2019, 52(1): 245-272.

- [27] LI T T, ZHAO Y, ZHANG C B, et al. A knowledge-guided and data-driven method for building HVAC systems fault diagnosis[J]. Building and Environment, 2021, 198(5): 107850.
- [28] LESSMEIER C, KIMOTHO J K, ZIMMER D, et al. Condition monitoring of bearing damage in electromechanical drive systems by using motor current signals of electric motors: A benchmark data set for data-driven classification[C]//Proceedings of the European Conference of the Prognostics and Health Management Society. New York: PHM Society, 2016: 1-17.
- [29] HOANG D T, KANG H J. A motor current signal-based bearing fault diagnosis using deep learning and information fusion[J]. IEEE Transactions on Instrumentation and Measurement, 2020, 69(6): 3325-3333.

Fault diagnosis method of BN ball mill rolling bearing based on AESL-GA

WANG Jinhua^{1,*}, TANG Guodong¹, CAO Jie^{1, 2}, LI Yajie¹

College of Electrical and Information Engineering, Lanzhou University of Technology, Lanzhou 730050, China;
 Gansu Manufacturing Information Engineering Research Center, Lanzhou 730050, China)

Abstract: To address the imperfect and imprecise shortcomings of the knowledge-based Bayesian network (BN) construction method, this paper proposes a BN structure construction method based on knowledge guidance and data mining. Firstly, aiming at the problem of inaccurate fault diagnosis results of a single signal and the uncertainty in fault information, the current signal and the vibration signal are fused to establish the characteristic nodes of the BN. The fault characteristic parameters of the two kinds of signals are extracted respectively, and the feature selection is carried out by the distinguish index method, which is used as the node of the feature layer of the BN structure. Secondly, the initial BN structure constructed by expert knowledge is combined with the adaptive elite-based structure learner using genetic algorithm (AESL-GA) to optimize the structure. By adaptively restricting the search space in the evolution process, reducing the number of free parameters and improving its global search ability, we obtain optimal BN structure. Finally, the method is verified by the measured data of the ball mill rolling bearing of Jinchuan Company and the data set of Paderborn University, which proves the effectiveness of the proposed method.

Keywords: Bayesian network; fault diagnosis; adaptive elite-based structure learner using genetic algorithm; rolling bearing; signal fusion

URL: link.cnki.net/urlid/11.2625.V.20220905.0837.001

Received: 2022-05-28; Accepted: 2022-07-10; Published Online: 2022-09-05 10:50

Foundation items: National Key R & D Program of China (2020YFB1713600); National Natural Science Foundation of China (62063020); Gansu Natural Science Foundation (20JR5RA463)

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

DOI: 10.13700/j.bh.1001-5965.2022.0439

基于装卸顺序的中型机多航段协同配载优化

赵向领^{1,*},李云飞²,王治宇³,徐吉辉²,李鹏飞⁴

(1. 中国民航大学 空中交通管理学院,天津 300300; 2. 空军工程大学 装备管理与无人机工程学院,西安 710000;

3. 珠海翔翼航空技术有限公司, 珠海 519030; 4. 天津大学 管理与经济学部, 天津 300072)

摘 要:多航段、多经停的航空货运航班,飞行过程中需要合理控制飞机的重心(CG)位置,保持平衡状态,中转站装卸操作过程中需要避免额外装卸操作,减少时间和劳力浪费,因此,研究多航段航班的协同配载问题具有重要现实意义。根据航空器自身特点与集装器(ULD)装卸顺序,结合各航段飞机装载平衡与过站机场ULD装卸操作要求,通过协调分配各航段ULD和散货位置,建立了中型机联程航班多目标整数线性规划模型。模型考虑了ULD和散货的质量、体积及与舱位的匹配限制,航空器限重、舱位限重、舱位及区域累积限重、上下舱联合限重和CG位置的平衡限制,以及前后航段ULD和散货的连接性约束。根据造成中间机场额外装卸操作的2种原因,基于装卸顺序,提出装卸优化模型,通过ULD的舱内平移,优化了CG,减少了装卸次数。以B757-200F机型为例,在两装一卸、一装两卸和两装两卸3种场景下,采用商用求解器Gurobi,分别针对3种不同目标函数组合进行求解和分析对比。实验表明:所提模型可以有效协调过站货物的额外装卸次数,优化前后两航段的CG位置。

关 键 词: 航空运输; 航空货运; 多航段; 协同配载; 载重平衡; 多目标; 整数线性规划 中图分类号: U8; V271.9; O221

文献标志码: A 文章编号: 1001-5965(2024)04-1147-15

航空货运配载是将不同类型的集装器(unit load device, ULD)合理分配到飞机舱位内,在保障飞行 安全的前提下,实现装载最大化和重心 (center of gravity, CG)偏离量最小。正确的 ULD 装载计划不 仅可以保证飞机安全,还可以降低燃油消耗,以及 航空公司的成本和二氧化碳排放。配载不合理易 造成严重后果,历史上有众多因配载不合理而造成 的航空事故。

然而,当前航空公司的配载,特别是货运配载 还处于计算机辅助配载阶段,许多配载方案需要人 工调配获得,使得一个合理的货运配载方案需要专 业的技术人员进行长时间、高强度的脑力劳动。

Feng 等^[1]总结当前航空货物运行的研究时指 出,飞机的配载问题是航空运输中最基本的 NP 难 组合优化问题之一。早期的配载研究主要以单航 段为主,可以分为启发式算法和混合整数线性规划 方法。

在 启 发 式 算 法 求 解 装 载 配 平 问 题 方 面, Amiouny 等^[2]分析了飞机一维装载平衡问题, 证明 为 NP 难问题, 考虑了装载集装箱的种类、位置和 装载优先次序, 对货物装载规划模型的安全要求提 出了有用的见解; Wodziak 和 Fadel^[3]将飞机装载平

收稿日期:2022-05-29;录用日期:2022-08-05;网络出版时间:2022-08-2616:27 网络出版地址:link.cnki.net/urlid/11.2625.V.20220826.1430.005

基金项目:国家自然科学基金 (71802141,52074309,U2333217);南航载重平衡管理系统提升 (KC22204);中央高校基本科研业务费专项资金 (3122020049);中国民航大学研究生科研创新资助项目 (2021YJS060)

*通信作者.E-mail: zxl-llx@163.com

引用格式:赵向领,李云飞,王治宇,等.基于装卸顺序的中型机多航役协同配载优化 [J].北京航空航天大学学报,2024,50(4):1147-1161. ZHAO X L, LI Y F, WANG Z Y, et al. Cooperating loading balance optimization for medium-sized aircraft with multiple flight legs based on loading and unloading sequence [J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50(4): 1147-1161 (in Chinese).

衡问题看作装箱问题,将飞机当作箱子,ULD当作 待装载货物,运用 Amiouny 等^[2]的思想,使用遗传 算法进行求解: Mathur^[4] 针对一维飞机装载平衡问 题提出了一种贪婪启发式算法,其目标是将给定长 度和质量的小箱子放入 ULD 中, 尽可能使该区域 的重心达到既定目标; Heidelberg 等⁵¹认为飞机上 的货物装载是二维装箱问题,建立的模型包括纵向 平衡、横向平衡,以及货物在地板上接触点的最大 质量、货物所在地板区域内可承受的最大质量等约 束; Dahmani 和 Krichen^[6-7]提出了多目标粒子群优 化算法,将一组货物分配到最小数量的ULD中,再 将其放置到飞机特定的货舱内,模型考虑了优化业 载总质量和优先级;谷润平等¹⁸针对货机主货舱以 装载量最大和重心偏移量最小为目标,建立了双目 标混合整数线性规划模型,考虑飞机货舱的重心包 线限制、质量限制、载荷限制、结构限制等,并使用 改进的遗传算法求解;赵向领和杜有权¹⁹建立了民 航客改货飞机的载重配平模型,用于解决航空器配 载问题,模型以运输业载量最大和重心位置最优为 目标,并运用遗传算法进行求解。

在(混合)整数线性规划方法方面, Brosh^[10]研 究了关于飞机货舱内货物分配和寻找最优装载布 局问题,提出关于飞机货舱内货物分配的整数线性 规划,该模型是对规则散装货物一次性批量装载, 目标是在给定体积、质量、重心约束条件下,使运 输货物利润最大化(即业载量最大); Thomas 等[11] 针对联邦快递公司所使用的轻型飞机建立了一种 整数线性规划模型; Mongeau 和 Bes^[12] 通过将集装 箱装载到飞机上对燃料消耗和安全要求两者之间 进行权衡,提出一种整数线性规划模型,目标是在 重心处于安全范围内且飞机的业载量最大,模型约 束主要包含预定义位置类型限制,各货舱容积限 制、质量限制、多种 ULD 混装限制等,该模型只估 算了粗略的重心,求解时间相对较长;Kaluzny和 Shaw^[13]针对军用飞机的货物装载平衡问题建立了 一种混合整数线性规划模型,目标是最小化重心偏 移量和最大化运输货物的价值,然而除了空间和重 心约束,未考虑飞机结构和安全约束,如地板强度、 累积负载等,这些约束在商业航空货运中起关键作 用; Limbourg 等^[14]提出了将 ULD 分配到装载位置 的混合整数规划模型,考虑了飞机的纵向和横向平 衡限制,以及ULD的组合装载约束、累积装载约 束,目标是使飞机的惯性矩和飞机结构应力及燃油 消耗最小化,以提高飞机的稳定性和节省成本,但 该模型无法解决 ULD 多于预定义位置的问题: Verstichel^[15] 和 Vancroonenburg^[16] 等建立了一种用

于解决飞机质量和平衡问题的混合整数线性规划 模型,目标是使飞机重心偏离目标值最小和所装载 货物价值最大,该模型保证了飞机结构的完整和稳 定,货物和机组人员的安全,以及飞机安全高效的 装卸货物,求解速度较快,但重心约束未使用飞机 重心包线;Zhao等^[17]提出一种基于不同状态下实 际重心变化的重心包线约束,并对重心进行了精确 描述,进而设计了线性化策略来简化非线性的重心 包线,在考虑位置、质量、平衡等约束条件下,建立 了总业载最大、重心偏差最小的混合整数规划模 型,使用商业求解器可在几秒内得到最优解,并且 在所有测试实例中保持重心在指定的重心包线范 围内。

随着航空运输业的快速发展,在单航段直达运 输模式的基础上出现了多航段联程运输。多航段 联程运输模式的飞机在中途站过站时间较短(最短 为40 min),这段时间内需要完成货物的卸载、配 载、再装载及其他保障工作,时间仓促。始发站的 配载方案直接决定中转站的装卸效率,如果 ULD 装载位置不合理, 会导致装卸过程存在大量可 以避免的额外装卸操作,占用大量时间,消耗大量 人力物力。对专业配载员来说,获得多航段的协调 配载的可行方案已有很大难度,更不用说优化配 载。合理装载能够较好地处理过站作业,减少装卸 操作数量,提高效率,缩短周转时间,降低机场操作 的人工成本。因此,装载优化工作具有重要意义。 然而,多航段的运输模式通常受到通程业载和过站 业载的限制, ULD 的装卸效率也有一定影响, 使得 各航段的载重平衡与过站操作效率存在"此消彼 长"的矛盾;如何协调过站机场中间操作的效率与 前后两航段的载重平衡,使多航段整体达到较好的 经济效益,是配载研究中的一个难点。

在多航段配载方面, Larsen 和 Mikkelsen^[18] 最早 提出多航段配载问题,指出货物装载计划的困难会 随货舱分舱的大小和航段数量的增加而增大。 Lurkin 和 Schyns^[19]建立了一种用于解决多航段飞 行的飞机配载问题的混合整数线性规划模型,该模 型讨论集装箱数量少于飞机舱位预留的位置,可以 将所有给定的 ULD 装载,目标是最小化由重心引 起的燃油成本和中转机场重新装载 ULD 的操作成 本之和,考虑了飞机的稳定性约束,以及飞机的每 个舱位质量限制、组合质量限制、累积质量限制和 货运约束(如 ULD 类别、卸载顺序等),但是,中间 机场装卸操作中,为减少约束条件,把留在飞机上 的过站 ULD 限制在原来位置,不允许位置调整,失 去为后一航段优化重心的机会。 Brandt^[20]建立了一种多航段货运配载问题的混 合线性整数规划模型,目标是最大化装载量,飞行 重心偏离指定重心最小和最小化中间机场不必要 的 ULD 装卸操作,虽然考虑了多种现实操作的限 制条件,但是位置质量限制、累积质量限制和上下 舱质量联合限制被泛化为统一的质量限制约束,这 与实际运行相差较远,且在中间机场操作中并未考 虑被留在货舱的 ULD 的位置变化。

综上所述, 配载问题是航空运输中的一个基本 的 NP 难问题, 当前针对航空配载方面的研究主要 集中在单航段的配载问题上, 多航段协同配载问题 研究甚少。但在当前多航段的配载研究中, 缺少在 减少中间机场额外装卸操作的同时, 通过对过站 ULD 的位置调整, 优化后一航段重心的方案。这一 点是非常必要的, 特别是后一航段距离较长时, 较 合理的重心可以节约较多燃油、降低碳排放和环境 污染。这是本文的研究动机之一。

上述文献主要针对诸如 B747F、B777F、A330 等重型航空货运飞机的配载研究。重型机主货舱 和下货舱都是 ULD, 均划分了预定义位置, 重型机 的配载将多种 ULD 分配到飞机上的不同预定义舱 位位置, 属于广义的指派问题。

国内诸如顺丰、圆通等公司的货运运输主要 以 B757-200F、B737-800、A320及相关的客改货等 中型运输机为主。中型机与重型机在配载方面存 在较大区别,主要表现为:中型机腹舱装散货,一般 有4个腹舱。将散货分配到不同的腹舱,同时需考 虑腹舱的容积、质量限制,具有多背包问题或多箱 装箱问题的特点。

航空配载涉及的 ULD 装箱问题的研究主要 有: Chen 等^[21]建立了将不规则物体装载到三维 ULD 中的 0-1 混合整数线性规划模型; Liu 等^[22]使 用多目标粒子群优化算法研究多目标二维装箱问 题; Yan 等^[23]提出一种非线性的混合整数规划模 型,用于解决 ULD 装载计划问题; Li 等^[24]提出一种 折中的大规模搜索邻域启发式模拟退火算法,将货 物分配到 ULD 内; Tang^[25]提出一种用于解决随机 需求下的空运货物装载问题的方法,将情景分解和 遗传算法技术相结合; Paquay 等^[26-29]考虑含有不同 特殊物品约束要求的三维多箱装箱问题,目标是最 大程度地减少未使用的体积。

这些装箱问题不需要考虑飞机配载相关的约束问题。但是,对于中型机而言,将散货装入腹舱本身是装箱问题,而该装箱问题又是整个配载问题的一部分。本文研究的问题正是国内货运运输的中型机的腹舱装散货的配载问题。

本文的创新点主要有:①研究多航段配载问题,减少中间机场额外装卸操作,并且通过调整留 在机舱内过站 ULD 的位置,实现对后一航段飞行 重心的优化。②研究中型机的货运配载问题,将腹 舱装散货的装箱问题与主货舱装 ULD 的指派问题 相结合,建立配载的组合优化模型。

本文建立了中型机多航段协同配载问题的混 合整数线性规划模型,优化中间过站机场装卸操作 和前后两航段的重心位置,考虑了飞机对 ULD 和 散货的质量、体积、高度、位置匹配等限制,飞机的 重心平衡、上下货舱质量联合限制,以及前后两航 段等限制条件,在减少 ULD 操作次数的同时,通过 平移两航段中 ULD 的位置减小重心偏差。选择单 通道 B757-200F 机型对模型进行求解和验证。

1 问题描述

基于现场操作经验,在联程航班中,为节约航 班过站过程中的货物装卸时间,便于过站装卸操 作,过站的货物应远离舱门,到站的货物靠近舱门 放置。图1为多航段运输飞机转运示意图,ULD按 照运输目的地的不同可以分为直达运输和联程 运输。



Fig. 1 Schematic diagram of aircraft transfer

直达运输指 ULD 只经过一站即卸载的运输, 即从机场 A 装载到机场 B 卸载,用集合 U₁ 表示;或 者从机场 B 装载到机场 C 卸载,用集合 U₂ 表示;为 方便中间机场的装卸,减少没必要的倒箱装卸操 作,此类 ULD 应尽量装载到舱门附近。联程运输 指 ULD 直接从起始机场 A 运送到目的机场 C,无 需在中间机场 B 卸载,用集合 U₃表示,装载时应远 离舱门。据此可知,ULD 随飞机时间越长,装载位 置应离舱门越远。

多航段协同配载中, ULD 的分配具有堆栈问题 的"先进后出, 后进先出"特点。先装入飞机的 ULD 需要推到货舱的内部, 距离门口较远, 后装入 的 ULD 距离门口较近, 会挡住先装入 ULD 的卸载 通道; 卸载时, 需要先卸载距门口较近的 ULD, 清理 通道后再卸载内部 ULD。

多航段的协同配载中,专家经验法是:在起飞机场优先分配需要联程运输 U, 的 ULD, 装载到远

离舱门的位置;后分配 U₁和 U₂的 ULD,装载到靠 近舱门附近,便于中间机场卸载和装入。

为方便表述装卸顺序,以 B757-200F 机型为例, 定义机舱舱位如图 2 所示。B757-200F 运输机在结 构上有 1 个主货舱、4 个下货舱。主货舱一般预定 义 15 个 M 型 ULD 位置,下货舱装载散货。用 *C_j* 表示飞机的第*j* 个主货舱舱位。以装载顺序为目标 定义 ULD 的装卸顺序, *C_j* 越大,装载优先级越高, 越优先装载,越最后卸载。



图 2 B757-200F 货舱侧视图 Fig. 2 Side view of B757-200F cargo compartment

如果 U_1 的一个 ULD 位于位置 C_8 , 在中间机场 需要卸载, 舱门在 C_1 位置, 卸载时, 如果 C_1 、 $C_3 \sim C_7$ 的位置有 U_3 的 ULD, 则需将 U_3 的 ULD 先卸载, 之后重新装回, 造成了额外的装卸操作, 应减少这 种不必要的装卸操作。同理, 在装载 U_2 时, U_3 也可 能会阻碍 U_2 的装载, 此时也会造成额外的卸载。

如果 U₃的 ULD 在前后两航段发生调舱,即 2个 ULD 舱位位置发生变换,由于 ULD 在飞机机 舱内无法完成位置变换,需要将发生位置变换的 ULD 和阻碍发生位置变换的 ULD 卸载,之后重新 装回,同样也造成了额外的装卸操作。

为保证作业过程中飞机安全稳定,飞机机头不 会翘起,经停站卸货过程中必须时刻调整过站板箱 位置,保证飞机重心处于安全范围内,杜绝发生飞 机沉尾现象。B757-200F 装载平衡手册中关于中转 站装卸 ULD 装载顺序要求为: $C_{15} \rightarrow C_{14} \rightarrow C_{13} \rightarrow C_{12} \rightarrow$ $C_{11} \rightarrow C_{10} \rightarrow C_9 \rightarrow C_8 \rightarrow C_7 \rightarrow C_6 \rightarrow C_5 \rightarrow C_4 \rightarrow C_3 \rightarrow C_2 \rightarrow C_1;$ 卸 载顺序要求为: $C_1 \rightarrow C_2 \rightarrow C_3 \rightarrow C_4 \rightarrow C_5 \rightarrow C_6 \rightarrow C_7 \rightarrow C_8 \rightarrow$ $C_9 \rightarrow C_{10} \rightarrow C_{11} \rightarrow C_{12} \rightarrow C_{13} \rightarrow C_{14} \rightarrow C_{15}$ 。

根据 ULD 到达目的地的不同和装卸顺序要 求,可以设计 ULD 理想的装载位置,如图 2 所示。 为减少不必要的操作,假设 10 个 U_3 的 ULD 装载到 远离 舱门 的 位 置 (见图 2 灰 色 区 域),即 C_{15} 与 $C_6 \sim C_{14}$ 位置;需要在中转站卸载/装载的 5 个 U_1 或 U_2 的 ULD 装载到靠近舱门的位置 (见图 2 白色区 域),即 $C_2, C_1, C_3 \sim C_3,$ 以方便中转站卸载/装载。

此外,还需考虑散货装卸顺序限制等问题,方 便集合为 B₁、B₂的散货在中转站卸载和装载。用 D_h表示飞机的第 h 个下货舱,由于下货舱舱门口分 别在 D₁和 D₄号下货舱中,在中转站卸载和装载的 散货尽量装载到 D₁和 D₄号下货舱,过站的散货尽 量装载到 D_2 和 D_3 号下货舱,这样方便散货在中转 站卸载,并保证飞机前部分质量或后部分质量不会 偏差太大。对此本文引入下货舱对应的优先级,如 图 2 所示, D_1 和 D_4 舱对应 1, D_2 和 D_3 舱对应 0, 分 别用 $\beta(1)=\beta(4)=1, \beta(2)=\beta(3)=0$ 表示。

基于上述分析,本文研究的中型机多航段的协同配载问题,需要在起飞机场、过站机场对 ULD 和 散货进行合理的舱位分配,保障过站前后每个飞行 航段的载重平衡,且减少中间过站机场额外装卸操 作。因此, U_3 的 ULD 在每个航段都需被尽量装载 到远离舱门的位置, U_1 和 U_2 尽量装载到舱门附近; 散货 B_3 尽量装载到 D_2 、 D_3 号下货舱,且保持航段 前后舱位不变,散货 B_1 和 B_2 尽量装载到含有舱门 的 D_1 、 D_4 号下货舱;同时需要考虑 ULD 的质量、高 度,散货的质量、体积,飞机的运行平衡、上下货舱 的联合限制等,以优化整个飞行任务。

2 模型建立

2.1 模型假设

①ULD 中 装 载 的 货 物 (质 量)均 匀 分 布, ULD 的重心在其几何中心位置;②不考虑下货舱散 货的摆放方式,且重心在该舱的预定义位置;③各 航段起飞燃油和航程燃油已知。

2.2 优化模型

多航段协同配载优化同单航段配载相比,在考 虑上下货舱每个舱位的基本约束和联合约束的同 时,还需要考虑中间机场对货物操作的问题。

2.2.1 主货舱对 ULD 的限制

航段 l 中,每个 ULD 最多装载到一个舱位中:

$$\sum_{j \in N_{\mathsf{P}}} x_{ijl} \leq 1 \tag{1}$$

式中: $N_{\rm P}$ 为飞机主货舱舱位数集合; x_{yl} 为主货舱 ULD 分配的决策变量, 如果在航段 l, 第 i 个待装载 ULD 被分配到第 j 个主货舱舱位位置, 则决策变量 为 1, 否则为 0, 其中, $i \in N_{\rm U}$, $j \in N_{\rm P}$, $l \in L$; $N_{\rm U}$ 为待装 载 ULD 集合; L 为飞行航段数集合。

航段1中,每个舱位最多装载一个ULD:

$$\sum_{i \in N_{\rm U}} x_{ijl} \le 1 \tag{2}$$

式中: $N_{\rm H}$ 为飞机下货舱舱位集合; y_{bhl} 为下货舱散货 分配的决策变量,如果在航段l,第b个待装载的散 货被分配到第h个下货舱,则决策变量为1,否则 为0,其中, $b \in N_{\rm B}$, $h \in N_{\rm H}$, $l \in L$; $N_{\rm B}$ 为待装载散货集合。

设 W_i为第 i个待装载 ULD 的质量, W_j为第 j个舱位的最大限制载荷,则在任意航段 l 所装载 的 ULD 质量不能超过所在舱位的最大质量:

$$\sum_{i \in N_{U}} W_{i} x_{ijl} \leq W_{j} \tag{3}$$

设*H_i*为第*i*个待装载 ULD 的高度,*H_j*为第*j*个 舱位的最大高度,则任意航段*l*上所装载的 ULD 的高度不超过所在舱位的最大高度:

$$\sum_{i \in \mathcal{N}_{i}} H_i x_{ijl} \leqslant H_j \tag{4}$$

设 *C*_p 为累积限重的组合位置, *C*_{POS} 为主货舱 组合位置的集合。对于 B757-200F(见图 2), *C*_{POS}= {(*C*₁₅, *C*₂, *C*₁, *C*₃, *C*₄), (*C*₅, *C*₆, *C*₇, *C*₈, *C*₉), (*C*₁₀, *C*₁₁, *C*₁₂, *C*₁₃, *C*₁₄)}。则舱位累积限重,即在任意航段 *l* 上, 所 装入组合舱位 *C*_p 的 ULD 质量之和不超过所规定 的最大质量限制 *W*_{Cp}:

$$\sum_{j \in C_{\rm P}} \sum_{i \in N_{\rm U}} W_i x_{ijl} \leq W_{C_{\rm P}} \qquad \forall C_{\rm P} \in C_{\rm POS}$$
(5)

2.2.2 下货舱对散货的限制

航段1中,每个散货最多装载到一个下货舱中:

$$\sum_{h \in N_{\rm H}} y_{bhl} \le 1 \tag{6}$$

式中: $N_{\rm H}$ 为飞机下货舱舱位集合; y_{bhl} 为下货舱散货 分配的决策变量,在航段l,第b个待装载的散货被 分配到第h个下货舱。

设 w_b为第 b个待装载散货的质量, W_b为第 h个下货舱所能承受的最大质量,则代表航段 l中, 每个下货舱所能装载的最大质量受到该舱能够承 受的最大载荷限制:

$$\sum_{b \in N_{P}} w_{b} y_{bhl} \leqslant W_{h} \tag{7}$$

式中:N_B为待装载散货集合。

设v_b为第b个待装载散货的体积,V_b为第h个 下货舱所能容纳的最大容积,则任意航段l上,每个 下货舱所能容纳的最大容积限制为

$$\sum_{b \in N_{\rm B}} v_b y_{bhl} \leqslant V_h \tag{8}$$

设 W_{h_1,h_2} 为 h_1 号下货舱和 h_2 号下货舱所能承受的最大累积限制质量, C_H 为下货舱组合舱集合,对于 B757-200F(见图 2), C_H ={ $(D_1,D_2), (D_3,D_4)$ }。则下货舱的累积限重,即在任意航段l上,所装入组合舱位 C_H 的散货质量之和不超过所规定的最大质量限制 W_{h_1,h_2} :

$$\sum_{b_1 \in N_{\rm B}} w_{b_1} y_{b_1 h_1 l} + \sum_{b_2 \in N_{\rm B}} w_{b_2} y_{b_2 h_2 l} \leqslant W_{h_1, h_2} \qquad \forall (h_1, h_2) \in C_{\rm H}$$
(9)

2.2.3 飞机载重与平衡限制

式(10)~式(12)为飞机在不同质量 W、任意航 段 l下的运行平衡要求。在整个飞行过程中,需要 时刻掌握飞机的重心位置。通常,飞机的重心必须 限定在飞机的重心包线之内,飞机的重心包线是由 飞机的指数和质量围成的一个重心范围。飞机的 指数是对飞机各种力矩的缩放和坐标转换。

$$I_{\text{TOW},l}^{\text{FWD}} \leq I_{\text{TOW},l} \leq I_{\text{TOW},l}^{\text{AFT}}$$
(10)

$$I_{ZFW,l}^{FWD} \leq I_{ZFW,l} \leq I_{ZFW,l}^{AFT}$$
(11)

$$I_{\text{LDW},l}^{\text{FWD}} \leq I_{\text{LDW},l} \leq I_{\text{LDW},l}^{\text{AFT}}$$
(12)

式中: $I_{*,l}^{\text{FWD}}$ 和 I_{*l}^{AFT} 分别为在航段 l, 起飞质量 (takeoff weight, TOW)、无油质量 (zero fuel weight, ZFW)、着 陆质量 (landing weight, LDW)下的重心包线前限和 后限指数; $I_{W,l}$ 为飞机在不同质量 W、不同航段 l下 配载结果的重心所对应的指数表达形式,其计算方 式为

$$I_{W,l} = I_{\text{OEW}} + F_{W,l} + \sum_{i \in N_{\text{D}}} \sum_{j \in N_{\text{P}}} \frac{W_i x_{ijl} (B_j - C_{\text{D}})}{\varepsilon} + \sum_{b \in N_{\text{B}}} \sum_{h \in N_{\text{H}}} \frac{w_b y_{bhl} (B_h - C_{\text{D}})}{\varepsilon}$$
(13)

式中: I_{OEW} 为空机质量 (operational empty weight, OEW)指数; $F_{W,I}$ 为飞机在不同质量 W、不同航段 l下燃油质量所对应的燃油指数; B_j 为飞机第j个舱 位所对应的力臂; B_h 为飞机第h个下货舱所对应的 力臂; C_p 为飞机的基准平衡力臂; ε 为缩小系数。

为确保飞行安全,避免超载与空载,提高航空 运输经济效益,需要对飞机各航段的业载做限制。 任意航段1上飞机的业载限制为

$$\sum_{i \in N_{\rm U}} \sum_{j \in N_{\rm P}} W_i x_{ijl} + \sum_{b \in N_{\rm B}} \sum_{h \in N_{\rm H}} w_b y_{bhl} \leq W_{\rm MPL,l}$$
(14)

式中: W_MPL, 为最大业载, 计算式为

$$W_{\text{MPL},l} = \min \left\{ W_{\text{MZFW}} - W_{\text{OEW}}, W_{\text{MTOW}} - W_{\text{OEW}} - W_{\text{TOF},l}, \\ W_{\text{MLDW}} - W_{\text{OEW}} - W_{\text{RFW},l} \right\}$$
(15)

式中: W_{MZFW} 、 W_{MTOW} 、 W_{MLDW} 分别为最大无油质量、最大起飞质量、最大着陆质量; W_{OFW} 为空机质量; $W_{TOF,I}$ 为航段 I下的起飞质量所对应的起飞油量; $W_{RFW,I}$ 为航段 I下的着陆质量所对应的备用油量。

式(16)和式(17)为飞机上下舱联合累积质量 限制,即在任意航段 l上,所装入组合主货舱 $C_{L,i}(C_{R,i})与 D_1和 D_2(D_3和 D_4)下货舱的货物质量$ 之和不超过所在区域所规定的最大质量限制 $<math>W_{C_{A}}(W_{C_{A}})$:

$$\sum_{j \in C_{\mathrm{L},i}} \left(\sum_{i \in N_{\mathrm{U}}} W_i x_{ijl} + \sum_{b_1 \in N_{\mathrm{B}}} \lambda_{D_{1j}} w_{b_1} y_{b_1 D_1 l} + \sum_{b_2 \in N_{\mathrm{B}}} \beta_{D_{2j}} w_{b_2} y_{b_2 D_2 l} \right) \leqslant W_{C_{\mathrm{L},i}} \quad \forall C_{\mathrm{L},i} \subseteq C_{Z\mathrm{L}} \quad (16)$$

$$\sum_{j \in C_{\mathrm{R},i}} \left(\sum_{i \in N_{\mathrm{U}}} W_i x_{ijl} + \sum_{b_4 \in N_{\mathrm{B}}} \lambda_{D_{4j}} w_{b_4} y_{b_4 D_4 l} + \sum_{b_3 \in N_{\mathrm{B}}} \beta_{D_{3j}} w_{b_3} y_{b_3 D_3 l} \right) \leqslant W_{C_{\mathrm{R},i}} \quad \forall C_{\mathrm{R},i} \subseteq C_{Z\mathrm{R}} \quad (17)$$

式中: $\lambda_{D_{e_i}}$ 、 $\beta_{D_{e_j}}$ 为表 1 中第 D_* 个下货舱在第 j 个区 域所对应的质量系数; $C_{L_i}(C_{R_i})$ 为上下舱联合累积 质量限制的主货舱组合位置; $C_{ZL}(C_{ZR})$ 为每个区域 所构成的主货舱累积组合舱位集合。

对于 B757-200F 上下舱联合限重,根据机型结构将货舱分为 15 个区域,其具体计算公式如表 1 所示, C_{ZL}和 C_{ZR}分别为

$$C_{ZL} = \left\{ (C_{15}), (C_{15}, C_2), (C_{15}, C_2, C_1), (C_{15}, C_2, C_1, C_3), \\ (C_{15}, C_2, C_1, C_3, C_4), (C_{15}, C_2, C_1, C_3, C_4, C_5), \\ (C_{15}, C_2, C_1, C_3, C_4, C_5, C_6) \right\}$$

 $C_{ZR} = \left\{ (C_{14}), (C_{14}, C_{13}), (C_{14}, C_{13}, C_{12}), (C_{14}, C_{13}, C_{12}, C_{11}), \\ (C_{14}, C_{13}, C_{12}, C_{11}, C_{10}), (C_{14}, C_{13}, C_{12}, C_{11}, C_{10}, C_{9}), \\ (C_{14}, C_{13}, C_{12}, C_{11}, C_{10}, C_{9}, C_{8}), \\ (C_{14}, C_{13}, C_{12}, C_{11}, C_{10}, C_{9}, C_{8}, C_{7}) \right\}$

X	间	实际区间联合载量计算公式	区间限重/kg
	$C_{\rm L1}$	C ₁₅	2 716
	$C_{\rm L2}$	$C_{\rm L1} + C_2 + D_1 \times 10\%$	4 360
	C_{L3}	$C_{\rm L2}$ + C_1 + D_1 ×80%	7 871
$C_{\rm ZL}$	$C_{\rm L4}$	$C_{L3}+C_3+D_1\times10\%+D_2\times40\%$	11 557
	C_{L5}	$C_{\rm L4} + C_4 + D_2 \times 40\%$	14 043
	$C_{\rm L6}$	$C_{\rm L5}$ + C_5 + D_2 ×20%	15 893
	$C_{\rm L7}$	$C_{\rm L6} + C_{6}$	17 712
	$C_{\rm R8}$	$C_{\mathrm{R7}}+C_{7}$	23 748
	C_{R7}	$C_{ m R6} + C_8$	21 909
C _{ZR}	$C_{ m R6}$	$C_{\rm R5} + C_9 + D_3 \times 20\%$	20 102
	C_{R5}	$C_{\rm R4} + C_{10} + D_3 \times 50\%$	18 257
	C_{R4}	$C_{R3} + C_{11} + D_3 \times 30\% + D_4 \times 10\%$	15 733
	$C_{\rm R3}$	$C_{\text{R2}}+C_{12}+D_4 \times 40\%$	11 969
	C_{R2}	$C_{\rm R1} + C_{13} + D_4 \times 30\%$	8 2 3 9
	$C_{\rm R1}$	$C_{14}\!\!+\!\!D_4\!\!\times\!\!20\%$	3 476

表 1 上下舱联合区间限重 Table 1 Combined weight limit of upper and lower cabins

2.2.4 中间机场的额外装卸操作

保证同一个 U_3 的ULD在航段1和航段2中都 被装载:

$$\sum_{j\in N_{\mathsf{P}}} x_{ij1} = \sum_{j\in N_{\mathsf{P}}} x_{ij2} \qquad \forall i \in U_3$$
(18)

使得联程的散货 B₃ 在 2 个航段中保持下货舱位置 不变:

$$y_{bh1} = y_{bh2} \qquad \forall b \in B_3, \forall h \in N_{\rm H}$$
 (19)

在各航段使得 U_3 的ULD i_1 在 i_2 的后面:

$$\begin{cases} S_{i_{1},i_{2}}^{l} \ge x_{i_{1}j_{1}l} + x_{i_{2}j_{2}l} - 1 \\ S_{i_{1},i_{2}}^{l} \le x_{i_{1}j_{1}l} \\ S_{i_{1},i_{2}}^{l} \le x_{i_{2}j_{2}l} \end{cases} \quad \forall i_{1}, i_{2} \in U_{3}, \forall j_{1}, j_{2} \in N_{P}, j_{1} > j_{2} \end{cases}$$

$$(20)$$

式中: S_{i_1,i_2}^l 为辅助决策变量,如果在航段 l属于 U_3 的 ULD i_1 在 i_2 在后面,则为 1,否则为 0,其中 i_1 , $i_2 \in U_3$, $l \in L_o$ 本文默认指向机头方向为前,机尾为 后。该辅助决策变量的作用是判断分配到不同航 段飞机舱位上的 2 个联程 ULD i_1 和 i_2 顺序是否保 持一直,即 S_{i_1,i_2}^1 是否与 S_{i_1,i_2}^2 相同,如果不相同则说明 前后两个航段 ULD i_1 和 i_2 发生了顺序变化,这需要 在中间机场进行额外装卸倒舱调序,增加装卸操作。 2.2.5 目标函数

过站散货如果在中间机场发生调舱,需要工作 人员在货舱内寻找需要调舱的货物,难以实现,浪 费大量时间,因此,本文所建模型考虑了将到达相 同目的地的散货尽量装载到相同的货舱内作为目 标函数之一,过站的散货在2个航段中保持舱位不 变作为约束条件之一。

使得2个航段上的业载量最大,目标函数为

$$\max \sum_{l \in L} \left(\sum_{i \in N_{U}} \sum_{j \in N_{P}} W_{i} x_{ijl} + \sum_{b \in N_{B}} \sum_{h \in N_{H}} w_{b} y_{bhl} \right)$$
(21)

使得2个航段上的重心偏移量最小,目标函数为

$$\min \sum_{l \in L} \left| P_{\text{CGT},l} - P_{\text{CGTOW},l} \right|$$
 (22)

式中: P_{CGT,} 为航段 l 起飞质量的目标重心; P_{CGTOW}, 为航段 l 起飞质量的重心。

在航空运输中,一般会用平均空气动力弦 (mean aerodynamic chord, MAC)来表示飞机重心的 位置,假设在航段 / 上飞机重心为 $B_{CG,I}$, B_{LEMAC} 为从 基准到平均空气动力弦前缘的长度, B_{MAC} 为平均空 气动力弦长,则此时的重心 $P_{CGTOW,I}$ 可表示为

$$P_{\text{CGTOW},J} = \frac{B_{\text{CG},J} - B_{\text{LEMAC}}}{B_{\text{MAC}}} \times 100\%$$
 (23)

式中: *B*_{CG}, 为力臂形式表示的飞机重心, 飞机重心 的确定综合运用了合力矩定理, 即系统合力在作用 点产生的力矩等于各分力在其作用点的力矩之 和。在某个航段中, 产生力矩的有空机质量 *W*_{OEW}、 燃油质量 *W*_{TOF},分配到飞机各舱位 *j* 的 ULD 质量及

(

各下货舱 h 的散货质量,设其距离基准点的距离分 别为 B_{OEW} 、 B_{TOE} 、 B_i 、 B_b ,则飞机的重心 B_{CGI} 为

$$B_{\text{CG},l} = \left(W_{\text{OEW}} B_{\text{OEW}} + W_{\text{TOF}} B_{\text{TOF}} + \sum_{i \in N_{\text{U}}} \sum_{j \in N_{\text{P}}} W_i B_j x_{ijl} + \sum_{b \in N_{\text{B}}} \sum_{h \in N_{\text{H}}} w_b B_h y_{bhl} \right) / \left(W_{\text{OEW}} + W_{\text{TOF}} + \sum_{i \in N_{\text{U}}} \sum_{j \in N_{\text{P}}} W_i x_{ijl} + \sum_{b \in N_{\text{B}}} \sum_{h \in N_{\text{H}}} w_b y_{bhl} \right)$$
(24)

式中:分母为飞机起飞质量:分子为当前质量对应 力矩。

使U1、U2尽量装载到舱门附近,目标函数为

$$\min \sum_{i \in U_1} \sum_{j \in N_P} j x_{ij1} + \sum_{i \in U_2} \sum_{j \in N_P} j x_{ij2}$$
(25)

为方便中间机场卸载和装载散货,应使得相同 目的地散货尽量装载到相同的散货舱内,目标函 数为

$$\max\left(\sum_{b_{1}\in B_{1}}\sum_{h\in N_{H}}\beta(h)y_{b_{1}h1} + \sum_{b_{2}\in B_{2}}\sum_{h\in N_{H}}\beta(h)y_{b_{2}h2} - \sum_{l\in L}\sum_{b_{3}\in B_{3}}\sum_{h\in N_{H}}\beta(h)y_{b_{3}hl}\right)$$
(26)

为避免 U₂的 ULD 发生顺序调整,造成额外装 卸操作,目标函数为

$$\min \sum_{i_1 \in U_3} \sum_{i_2 \in U_3} \left| S_{i_1, i_2}^1 - S_{i_1, i_2}^2 \right| \qquad i_1 \neq i_2$$
 (27)

实验与分析 3

目前, 配载被认为是 NP 难问题, 尚无多项式求 解方法,随着待装载 ULD 数量的增加,主货舱装载 ULD的复杂度呈增长趋势。启发式算法很难在大 规模规划问题上求得最优解。因此,本文采用 Gurobi 求解器的分支剪界方法。在求解算例的过程中, Gap 值逐渐减小为 0, 即多航段协同配载整数规划 模型达到全局最优解。

为验证模型有效性,针对 B757-200F 机型,采集 多个航班数据,进行算例分析与验证。表2为 B757-200F 基本参数,表3为主货舱最大舱位载荷、 最大高度限制及分区载荷,表4为下货舱舱位质 量、体积及分区载荷限制,表5为所需油量指数, F为燃油量,I为对应指数。测试算例计算机条件 处理器为第8代英特尔酷睿 i5, 运行内存8GB。软 件运行环境操作系统为 Win10, Anaconda3 Spyder Gurobi 9.5.1

表 2 B757-200F 基本参数

$W_{\rm OEW}/{ m kg}$	$W_{\rm MTOW}/{\rm kg}$	$W_{ m MZFW}/ m kg$	$B_{\rm MAC}/{ m m}$	$C_{\rm D}/{\rm m}$	I _{OEW}	$W_{\rm MLDW}/{ m kg}$	$P_{\rm CGT, l}/\rm MAC$	$B_{\rm LEMAC}/{\rm m}$	З
52 7 52	108 862	83 460	5.07	26.36	31.3	95 254	23	25.19	70 000

Table 3 Main cargo space constraints										
舱位	力臂/m	最大载荷/kg	最大高度/m	分区载荷						
C_{15}	9.91	2 716	2							
C_2	12.17	2 948	2	СССС和						
C_1	14.43	2 948	2	C ₁₅ 、C ₂ 、C ₁ 、C ₃ 和 C ₄ 质量和						
C_3	16.69	2 948	2	不超过18 000 kg						
C_4	18.95	2 948	2							
C_5	21.21	2 948	2							
C_6	23.47	2 948	2							
C_7	25.73	4 264	2	<i>C</i> ₅~ <i>C</i> ₅质量和 不超过24 000 kg						
C_8	27.99	4 264	2	-1702.221000 Kg						
C_9	30.25	2 948	2							
C_{10}	32.51	2 948	2							
C_{11}	34.77	2 948	2							
C_{12}	37.03	2 948	2	C ₁₀ ~C ₁₄ 质量和 不超过29 000 kg						
C_{13}	29.29	2 948	2	, <u>, , , , , , , , , , , , , , , , , , </u>						
C_{14}	41.55	2 948	1.95							

表3 主货舱舱位约束

表4 下货舱舱位约束

Table 4 Lower cargo hold space constraints

下货舱舱位	舱位限重/kg	舱位累积限重/kg	力臂/m	可用体积/m ³
D_1	2 496	4 672	13.58	4.9
D_2	4 672	4 672	18.64	13.9
D_3	3 773	7 393	32.73	14.2
D_4	5 606	7 393	38.29	16.7

表 5 油量指数

		Table 5 I	Fuel index		
F/kg	Ι	F/kg	Ι	F/kg	Ι
500	+0.1	1 000	+0.2	1 500	+0.3
2000	+0.4	2 500	+0.5	3 000	+0.7
3 500	+0.8	4 000	+1.0	4 500	+1.2
5 000	+1.4	5 500	+1.6	6 000	+1.8
7 000	+2.3	7 500	+2.6	8 000	+2.9
8 500	+3.2	9 000	+3.6	9 500	+4.0
10 000	+4.5	10 500	+5.0	11 000	+5.6
11 500	+6.2	12 500	+7.6	15 000	+7.5
15 500	+7.1	16 000	+6.7	17 000	+5.9
18 000	+5.1	19 000	+4.3	20 000	+3.5
24 000	+0.6	26 000	-0.8	26 500	-1.2
27 000	-1.6	27 500	-1.9		

3.1 算例设计

根据不同机场运输特性和两航段长度,本文设 置了两装一卸、一装两卸、两装两卸3种测试场 景。两装一卸为飞机在起飞机场和中间机场装载, 只在目的机场卸载,中间机场不卸货,一般航段 1较短,航段2较长,常见于国内-国内-国际运输。 一装两卸为货物从起飞机场装载后,发向2个目标 机场,中间机场和目的机场卸载货物,中间机场不 装载,一般航段1距离较长,航段2距离较短,常见 于国际-国内-国内。两装两卸为多机场相互运输, 飞机在起飞机场装载货物运往中间机场和目的机 场卸载,在中间机场还装载部分货物到目的机场, 一般2段航距相差不大,常用于国内运输。

表 6 为本文算例设计中货物计划装载清单。

表中: "<"、"="、">"3 类符号分别表示每种场景 下待装载 ULD 数量小于、等于、大于飞机舱位数 的情况;每个航段的燃油量不相同,通过燃油量和 飞机三大质量计算得到飞机每个航段最大业载 W_{MPL1}和 W_{MPL2}。

由于多目标之间的相互冲突,在实际操作中可能涉及到取舍。不同航空公司应根据公司战略和业务实际情况在目标函数选择上做出调整。例如,要求过站快,航段2为短途运行的航班,一般更偏重于减少中间机场ULD装卸次数;若中转时间宽裕,航段2时间长,偏重于重心优化可更节约燃油。

在后续实验中,基于目标函数,本文讨论了综 清单。 合优化(考虑所有目标)、偏好 ULD 装卸次数优化 表6 货物装载清单

Table 6 Cargo loading list

			-			-								-
测试	ULD 却榵	算例	$U_1/$	$U_2/$	$U_3/$	$B_1/$	$B_2/$	$B_3/$	W _{TOF,1} /	W _{RFW,1} /	W _{MPL,1} /	W _{TOF,2} /	W _{RFW,2} /	W _{MPL,2} /
初泉	风侠	<u></u> 1-1	(kg(+))	4 634(3)	(kg(+))	(Kg(+))	763(50)	2.221(150)	<u>к</u> g 7 000	<u>к</u> д 500	<u>kg</u> 30.708	<u>к</u> g 15.000	2 500	к <u>g</u> 30 708
	<	1-2	0	6 054(5)	13 971(9)	0	1 028(70)	1 996(140)	8 000	1 000	30 708	16 000	2 000	30 708
		1-3	0	11 271(7)	9 599(6)	0	1 355(90)	1 184(80)	7 500	1 000	30 708	15 500	2 500	30 708
		1-4	0	15 618(10)	6 848(5)	0	2 361(150)	1.521(100)	8 500	1 500	30 708	16 000	2 000	30 708
两装 一卸		1-5	0	18 655(12)	4 588(3)	0	2 752(180)	2 353(150)	8 000	1 000	30 708	17 000	3 000	30 708
	=	1-6	0	14 071(9)	7 952(6)	0	2 400(160)	3 078(200)	9 000	1 000	30 708	15 500	2 000	30 708
		1-7	0	4 568(4)	18 306(11)	0	3 080(200)	3 629(240)	9 000	1 500	30 708	24 000	4 000	30 708
		1-8	0	30 540(20(10))	7 220(5)	0	2 166(150)	1 635(100)	8 500	1 000	30 708	26 500	5 000	29 610
	>	1-9	0	25 170(18(6))	13 178(9)	0	1 545(100)	2 043(130)	9 500	2 000	30 708	27 500	5 500	28 6 10
		1-10	0	10 869(7)	13 134(10(8))	0	1 513(100)	2 324(150)	8 000	500	30 708	26 000	5 000	30 1 10
		2_1	2 672(2)	0	18 155(11)	1 398(80)	0	2 801(160)	16,000	3 000	30 708	7.000	500	30 708
	2	2-1	6 111(4)	0	11 424(8)	2 062(120)	0	2 604(150)	15 000	2 500	30 708	7 500	1 000	30 708
		2-3	13 034(9)	0	6 604(4)	2 617(150)	0	2 100(120)	15 500	2 000	30 708	8 000	1 000	30 708
		2-4	10 438(6)	0	14 426(9)	1 685(100)	0	2 665(150)	17 000	3 000	30 708	8 500	1 500	30 708
壮		2-5	11 036(8)	0	9 902(7)	2 275(130)	0	2 023(110)	26 500	5 000	29 610	9 000	2 000	30 708
兩卸	=	2-6	11 742(9)	0	8 548(6)	2 560(150)	0	1 824(100)	27 500	6 000	28 610	9 500	2 000	30 708
		2-7	17 643(11)	0	6 349(4)	2 891(160)	0	2 118(120)	27 000	5 500	29 1 10	9 500	1 000	30 708
		2-8	29 946(20(9))	0	10 087(6)	3 540(200)	0	1 757(100)	18 000	3 500	30 708	10 000	2 000	30 708
	>	2-9	22 490(15(8))	0	10 603(7)	2 773(160)	0	2 144(120)	19 000	4 000	30 708	11 000	2 000	30 708
		2-10	10 863(7)	0	18 568(13(8))	2 118(120)	0	2 522(140)	20 000	4 500	30 708	11 500	2 000	30 708
		3-1	2 587(2)	8 018(5)	15 230(10)	60/947	1 499(100)	1 729(120)	10 500	2 500	30 708	10 500	3 000	30 708
	<	3-2	11 246(7)	11 627(8)	10 205(7)	1 834(120)	2 330(140)	1 757(120)	11 000	3 000	30 708	9 500	1 500	30 708
		3-3	12 059(8)	7 074(5)	10 593(7)	2 023(140)	1 490(100)	1 804(120)	10 500	3 000	30 708	10 000	2 000	30 708
		3-4	3 078(2)	2 836(2)	21 057(13)	1 513(60)	2 271(80)	4 242(160)	9 500	1 500	30 708	11 000	3 000	30 708
两装		3-5	8 074(5)	6 568(5)	14 278(10)	1 851(100)	1 587(100)	1 711(100)	10 000	2 000	30 708	10 000	2 000	30 708
两卸	=	3-6	13 753(9)	10 177(6)	9 400(6)	4 415(150)	3 032(110)	2 964(110)	11 000	3 000	30 708	12 500	2 500	30 708
		3-7	19 430(12)	16 475(12)	3 827(3)	5 059(180)	5 028(180)	2 242(80)	10 500	3 000	30 708	9 500	1 500	30 708
		3-8	16 679(10(6))	16 320(10(6))	13 056(9)	2 830(160)	2 888(160)	1 749(100)	12 500	2 500	30 708	10 000	2 000	30 708
	>	3-9	7 563(5)	11 576(8(5))	23 684(15(10))	2 778(100)	2 765(100)	2 844(100)	9 500	1 500	30 708	10 000	2 000	30 708
		3-10	10 022(7)	12 273(9(7))	17 413(13(8))	2 732(120)	2 689(120)	3 102(140)	10 000	2 000	30 708	11 000	3 000	30 708

注:表中数据形式Z(X(Y))和Z(X)中,X表示各场景下待装ULD或散货数量,Z表示待装质量,Y为当前航空器可装载ULD数量,仅当X>Y时表示为X(Y)。

(去掉式 (22))和偏好重心优化(去掉式 (20)、式 (25) 和式 (27))3种优化情况。

3.2 算例分析

3.2.1 业载量

表 7 为各航段所装载 ULD 和散货的分配结果。表中: U 和 B 分别为所装载 ULD 和散货的数量; W_{PL} 为航段业载量。

由于具有相同目标函数式(21),3个不同优化 情况在每个航段所装载业载量达到了最大且相同, ULD 和散货数量相同。

另外,在3种不同目标函数的组合下,除算例 3-5外,其他算例中过站的散货都装载到了 D₂、 D₃号下货舱,直达的散货装载到了 D₁、D₄号下货 舱,2类散货没有出现混装的情况。

表 7 各航段业载分配

 Table 7
 Payload distribution of each leg

3.2.2 重心偏差

表 8 为 3 种不同目标函数下的重心偏差计算结 果。表中: |ΔC_{MAC,}| = |P_{CGT,} - P_{CGTOW,}|为各航段的起 飞重心偏差。图 3 为 3 个场景 2 个航段的重心偏差 雷达图,所有雷达图中均使用对数函数lg(|ΔC_{MAC}|+2) 将结果转化放大。

结合表3可知,偏好ULD装卸次数优化重心偏 差较大,平均分别为4.049%MAC和3.136%MAC, 方差为5.458和4.829,重心偏差较大,容易超出重 心包线。综合优化稍逊于偏好重心优化,2个模型 重心偏差均较小,最大平均为0.693%MAC。在实 际操作中,重心偏差相差2%MAC是可接受的。

3.2.3 中间机场额外装卸次数

表 9 为在 3 种不同优化模型下的中间机场对 表 8 重心偏差

Table 8 Center of gravity deviation

						-	$ \Delta C_{\text{MAC}} /\%\text{MAC}$		$ \Delta C_{\text{MAC}} /\%\text{MAC}$		$\Delta C_{\rm MAC}$ //%MAC		
算例	<i>U</i> /个		<i>B</i> /件		$W_{\mathrm{PL},l}/\mathrm{kg}$		算例	(综合优化)		(偏好装卸)		(偏好重心)	
序号	<i>l</i> =1	<i>l</i> =2	<i>l</i> =1	<i>l</i> =2	<i>l</i> =1	<i>l</i> =2	予 号	<i>l</i> =1	<i>l</i> =2	<i>l</i> =1	<i>l</i> =2	<i>l</i> =1	<i>l</i> =2
1-1	10	13	150	220	17 313	22 710	1-1	0.894	0.759	4.960	7.336	0.439	0.372
1-2	9	14	140	210	15 967	23 049	1-2	0.898	0.748	2.165	0.282	0.441	0.366
1-3	6	13	80	170	10 783	23 389	1-3	0.394	0.745	0.843	1.741	0.496	0.324
1-4	5	15	100	250	8 369	26 348	1-4	0.512	0.708	0.079	0.852	0.579	0.663
1-5	3	15	150	330	6 941	28 348	1-5	1.538	0.878	2.658	1.090	1.538	1.056
1-6	6	15	200	360	11 030	27 501	1-6	0.975	0.404	10.176	8.362	0.480	0.351
1-7	11	15	240	440	21 935	29 583	1-7	1.652	1.271	7.025	3.508	0.403	0.314
1-8	5	15	100	250	8 855	28 217	1-8	0.312	0.663	3.927	1.517	0.484	0.311
1-9	9	15	130	230	15 221	27 23 1	1-9	0.889	0.637	9.953	5.567	0.437	0.311
1-10	8	15	150	250	13 303	25 685	1-10	0.028	0.517	5.711	3.374	0.458	0.320
2-1	13	11	240	160	25 026	20 956	2-1	0.514	0.832	4.045	3.528	0.398	0.375
2-2	12	8	270	150	22 201	14 028	2-2	0.146	0.897	1.665	4.501	0.410	0.413
2-3	13	4	270	120	24 355	8 704	2-3	0.363	0.489	2.986	4.616	0.773	0.994
2-4	15	9	250	150	29 214	17 091	2-4	0.370	0.910	3.557	2.669	0.339	0.432
2-5	15	7	240	110	25 236	11 925	2-5	0.183	0.491	3.370	0.193	0.655	0.936
2-6	15	6	250	100	24 674	10 372	2-6	0.210	0.949	6.829	0.700	0.602	0.135
2-7	15	4	280	120	29 001	8 467	2-7	0.629	0.976	4.772	8.388	0.581	0.453
2-8	15	6	300	100	30 704	11 844	2-8	0.675	0.924	3.432	4.481	0.644	0.911
2-9	15	7	280	120	29 199	12 747	2-9	0.011	0.795	3.701	2.530	0.332	0.443
2-10	15	8	260	140	28 638	15 657	2-10	0.450	0.850	3.975	2.854	0.607	0.853
3-1	12	15	180	220	20 493	26 476	3-1	0.852	0.388	2.502	1.894	0.403	0.375
3-2	14	15	240	260	25 042	25 919	3-2	0.811	0.805	3.963	3.991	0.379	0.382
3-3	15	12	260	220	26 479	20 961	3-3	0.766	0.822	3.922	2.374	0.375	0.403
3-4	15	15	220	240	29 890	30 406	3-4	0.396	0.388	4.117	0.671	0.364	0.357
3-5	15	15	200	200	25 914	24 144	3-5	0.802	0.791	3.968	4.049	0.380	0.388
3-6	15	15	260	220	30 442	25 573	3-6	0.198	0.631	3.111	0.708	0.357	0.370
3-7	15	15	260	260	30 558	27 572	3-7	0.358	0.375	1.619	2.887	0.358	0.375
3-8	15	15	260	260	28 637	29 910	3-8	0.737	0.741	7.375	5.180	0.358	0.363
3-9	15	15	200	200	30 312	30 703	3-9	0.363	0.360	0.773	1.454	0.363	0.360
3-10	15	15	260	260	27 153	27 103	3-10	0.179	0.055	4.302	2.786	0.179	0.054




U₃的 ULD 的额外操作结果。表中: E_{unload} 为中间机 场对 ULD 的额外装卸次数。图 4 和图 5 分别为中 间机场额外装卸次数分布图和模型求解时间雷达 图。图 4 中,将额外装卸次数分为 5 组,分别为 [0,2]、[3,5]、[6,8]、[9,11]、[12,14],再按照装卸次数 绘制额外装卸次数分布图。

结合表9可知,偏好重心优化模型的额外装卸 次数最多,平均为7.5,综合优化和偏好ULD装卸 次数优化额外装卸次数平均分别为3.5和3.6。综 合优化和偏好重心优化在求解时间上相对较长,平 均分别为8.448 s和5.809 s,偏好ULD装卸次数优 化所用时间最少,平均为0.856 s。

由图 4 可知,综合优化和偏好 ULD 装卸次数优

	表9	额外装卸次	数
Table 9	Nun	nber of extra	operations

算例		$E_{\rm unload}$			时间/s	
序号	综合优化	偏好装卸	偏好重心	综合优化	偏好装卸	偏好重心
1-1	0	0	10	8.39	0.48	8.46
1-2	9	8	9	0.39	0.45	5.21
1-3	2	6	6	0.33	0.25	0.12
1-4	4	4	5	0.28	0.22	0.19
1-5	0	0.27	0	3	0.45	0.19
1-6	0	0	6	4.05	0.64	10.75
1-7	0	0	11	1.57	0.63	9.48
1-8	3	1	5	0.58	0.5	0.34
1-9	0	0	9	7.98	0.46	8.07
1-10	8	8	9	1.27	0.91	3.46
2-1	10	11	11	0.63	0.56	3.98
2-2	4	7	7	0.41	0.41	3.72
2-3	0	0	4	0.57	0.29	2.79
2-4	8	8	9	0.58	0.62	5.02
2-5	6	7	7	0.47	0.44	1.61
2-6	6	5	6	0.27	0.23	0.08
2-7	4	4	4	0.34	0.2	0.12
2-8	6	4	6	1.19	0.37	0.1
2-9	6	6	7	0.35	0.21	7.81
2-10	7	8	8	100.75	0.94	6.69
3-1	1	1	10	1.32	0.74	3.87
3-2	7	7	7	5.05	2.45	1.73
3-3	7	7	7	2.79	1	15.53
3-4	0	0	13	0.94	0.81	3.4
3-5	1	0	10	9.17	0.69	13.44
3-6	6	6	6	0.44	3.04	1.78
3-7	0	0	3	10.4	2.4	8.05
3-8	0	0	9	7.81	1.48	5.9
3-9	0	0	10	76.79	1.97	38.45
3-10	0	0	8	7.87	2.03	3.94

化均有 14 组算例额外装卸次数在 0~2, 而偏好重 心却为 0 组; 综合优化和偏好 ULD 装卸次数优化 分别有 2 组和 1 组算例在 9~11 次, 而偏好重心却 为 11 组, 并且有 1 组算例在 12~14 次, 该组算例额 外装卸次数达到了 13。由图 5 可知, 偏好 ULD 装 卸次数优化求解所用时间相对较少且集中, 而综合 优化算例 2-10、3-9 所用时间较长, 导致雷达图呈分 散状态, 偏好重心优化相对次之。

3.2.4 舱内移动 ULD

1)两装一卸算例分析。图 6~图 8 为场景两装 一卸中的算例 1-1 在不同目标函数组合下所计算绘 制的 ULD 和散货装载量的分配示意图。图中:将 每次配载结果的重心标注到飞机机身上,如图 6(a) 所示,重心为 24%MAC,与最优重心相差 1%MAC。



图 4 中间机场额外装卸次数

Fig. 4 Number of extra operations at intermediate airports







Fig. 6 Test 1-1 comprehensive optimization stowage results





图 7 算例 1-1 偏好装卸配载结果

Fig. 7 Test 1-1 preferred loading and unloading stowage results





由图 6(a)和图 6(b)分配结果可知, U₃ 的 ULD (灰色区域)在 2 个航段间发生舱位位置的变化, 但 额外装卸次数依然为 0(见表 9)。这是因为: 在 式 (20)和式 (27)的作用下, U₃ 的 ULD 按照模型分 配结果在前后两航段发生了位置平移挪动, 如 3~10号 ULD 分别向飞机内部发生了平移, 为中间 机场再次装载 U₂ 的 ULD 空出门口附近的位置。 这种平移并无前后次序的调整, 因此, 不需要卸载 和再装载, 减少了额外装卸操作。

同时,在航段1,因为装载布局松散,分布比较 均匀,所以可获得更好的重心效果。如果在航段 1就按照平移后的结果进行配载,配载结果如 图 6(c)所示,重心偏差却高达12.285%MAC,接近 重心包线所规定的边界范围,这是不可接受的。因 此,前后2个航段中U₃的ULD可以前后移动,增加 了配载灵活性,不仅可以减少中间机场的额外装卸 操作,还可以适当优化重心,相对于文献[19-20],这 是本文的一个重要创新点。

同理,图7(a)和图7(b)的分配结果显示额外装 卸次数为0;由于只考虑重心偏差,在图8(a)和图8(b) 中,2个航段中U₃的ULD并未保持舱位位置顺序 不变,需在中间机场进行调舱,U₁、U₂的ULD也未 分配到舱门口,额外装卸次数为10,即需要全部卸下再装上。假设每次卸载需要3min,装载需要5min,则10次操作需要额外80min,对于要求快速过站的航班来说,占用了本可以节省的多数时间,而且还增加了场面事故发生的概率和劳动力成本。

2)一装两卸案例分析。图 9~图 11 为场景一 装两卸中的算例 2-3 在不同目标函数组合下所计算 绘制的 ULD 和散货装载量的分配示意图。

由图 9(a)和图 10(a)的航段 1 配载结果中可 知,由于装载顺序目标的作用 (式 (25)), U₃ 的 ULD 全部装载到机舱内部,且没有阻挡在中间机场卸载 的 U₁ 的 ULD,不会造成额外装卸操作;在中间机场 卸载 U₁后,留在机舱内的 U₃ 通过位置移动实现航 段 2 的重心进一步优化,图 9(b)的航段 2 重心偏差 为 0.489%MAC,如果不移动(见图 9(c)),重心将为 39.731%MAC,超出重心包线,这在实际中是不可取 的;图 10(b)的航段 2 重心偏差为 4.616%MAC,如



图 9 算例 2-3 综合优化配载结果







Fig. 10 Test 2-3 preferred loading and unloading stowage results





图 11 算例 2-3 偏好重心配载结果

Fig. 11 Test 2-3 preferred center of gravity stowage results

果不移动,重心将为26.104%MAC,体现了次序约 束的优势。

在图 11 中,由于只考虑重心偏差,2个航段中 过站的 ULD 并未保持舱位位置顺序不变,需在中 间机场进行调舱,中转的 ULD 也未分配到舱门口, 因此,需要 4 次额外装卸操作。

3)两装两卸案例分析。图 12~图 14 为场景两 装两卸中的算例 3-8 在不同目标函数组合下所计算 绘制的 ULD 和散货装载量的分配示意图。

与一装两卸相同,在图 12(a)和图 13(a)的航段 1 的配载结果中,由于装载顺序目标函数的作用, U₃ 的 ULD 全部装载到舱内部,未造成额外装卸



(b) 航段2

图 13 算例 3-8 偏好装卸配载结果

Fig. 13 Test 3-8 preferred loading and unloading stowage results



图 14 算例 3-8 偏好重心配载结果

Fig. 14 Test 3-8 preferred center of gravity stowage results

操作。由于可选 ULD 的数量大于机舱预定义的位置,在式 (21) 的作用下,将全部装满,U₃ 没有剩余空间可以移动调整。U₂的分配调整可以适当优化 航段 2 的重心位置。

由图 14(a)和图 14(b)的分配结果可知,由于只 考虑重心偏差,2个航段中过站的 ULD 并未保持舱 位位置顺序不变,需在中间机场进行调舱,中转的 ULD 也未分配到舱门口,因此,额外装卸次数为9。

综上,综合优化在中间机场额外装卸次数上或 重心偏差上接近于偏好 ULD 装卸次数优化或偏好 重心优化,效果较好;装载顺序的目标使得 U₃的 ULD 尽可能向机舱内部分配,减少了在中间机场阻 挡 U₁卸载和 U₂装载的几率;顺序约束及其目标函 数可以在航班有空余位置时,有效地通过在机舱内 前后平移 ULD 来减少额外装卸操作,并优化重心, 效果较为明显,属于本文的创新。

4 结 论

 本文建立了主货舱装载 ULD、腹舱装散货 的中型机多航段协同配载优化模型。

2)考虑常见的一装两卸、两装一卸和两装两 卸3种实际运输情况,分别从装载量、起飞重心偏 离量、中间机场 ULD 额外装卸次数及模型求解时 间多个方面,分析验证了模型的有效性。

3) 基于装载顺序的多航段配载优化,能够使得 过站 ULD 尽可能向机舱内部分配,减少了在中间 机场阻挡装卸 ULD 的概率,从而有效减少中间机 场的额外操作。

4) 对前后两航段过站 ULD 的顺序约束及其目标函数, 松弛了早期研究中固定位置约束, 为两航段的 ULD 分配提供了灵活性, 可以在航班有空余位置时, 通过在机舱内前后平移优化重心位置, 不进行位置顺序调整, 减少额外装卸次数, 效果明显。

随着中国航空货运的崛起和航空公司中型机

客改货情况的增多,以及军事航空运输的应急需求 的增多,未来将基于机型特点和运输实际条件做进 一步的研究,如客改货中不拆座椅的轻改、主货舱 散装、不规则特种货物运输、装箱与配载整体优 化、多机型多机集装集配等问题,同步开展运输任 务装载风险辨识与防控的策略研究。

参考文献(References)

- [1] FENG B, LI Y Z, SHEN Z J M. Air cargo operations: Literature review and comparison with practices[J]. Transportation Research Part C:Emerging Technologies, 2015, 56: 263-280.
- [2] AMIOUNY S V, BARTHOLDI J J, VANDE VATE J H, et al. Balanced loading[J]. Operations Research, 1992, 40(2): 238-246.
- [3] WODZIAK J R, FADEL G M. Packing and optimizing the center of gravity location using a genetic algorithm[J]. Journal of Computers in Industry, 1994, 11: 2-14.
- [4] MATHUR K. An integer-programming-based heuristic for the balanced loading problem[J]. Operations Research Letters, 1998, 22(1): 19-25.
- [5] HEIDELBERG K R, PARNELL G S, AMES J E. Automated air load planning[J]. Naval Research Logistics, 1998, 45(8): 751-768.
- [6] DAHMANI N, KRICHEN S. On solving the bi-objective aircraft cargo loading problem[C]//Proceedings of the 5th International Conference on Modeling, Simulation and Applied Optimization. Piscataway: IEEE Press, 2013: 1-6.
- [7] KRICHEN S, DAHMANI N. Solving a load balancing problem with a multi-objective particle swarm optimisation approach: Application to aircraft cargo transportation[J]. International Journal of Operational Research, 2016, 27(1/2): 62.
- [8] 谷润平, 贾旭颖, 赵向领, 等. 民航货机装载优化准确建模仿真研究[J]. 计算机仿真, 2019, 36(3): 20-26. GU R P, JIA X Y, ZHAO X L, et al. Research on loading, optimization and accurate modeling and simulation of civil aviation cargo aircraft[J]. Computer Simulation, 2019, 36(3): 20-26(in Chinese).
- [9] 赵向领, 杜有权. 基于遗传算法的民用航空器配载问题[J]. 中国 科技论文, 2021, 16(8): 849-854.
 ZHAO X L, DU Y Q. Civil aircraft stowage based on genetic algorithm[J]. China Science Paper, 2021, 16(8): 849-854(in Chinese).
- [10] BROSH I. Optimal cargo allocation on board a plane: A sequential linear programming approach[J]. European Journal of Operational Research, 1981, 8(1): 40-46.
- [11] THOMAS C, CAMPBELL K, HINES G, et al. Airbus packing at federal express[J]. Interfaces, 1998, 28(4): 21-30.
- [12] MONGEAU M, BES C. Optimization of aircraft container loading[J]. IEEE Transactions on Aerospace and Electronic Systems, 2003, 39(1): 140-150.
- [13] KALUZNY B L, SHAW R H A D. Optimal aircraft load balancing[J]. International Transactions in Operational Research, 2009, 16(6): 767-787.
- [14] LIMBOURG S, SCHYNS M, LAPORTE G. Automatic aircraft cargo load planning[J]. Journal of the Operational Research Society, 2012, 63(9): 1271-1283.
- [15] VERSTICHEL J, VANCROONENBURG W, SOUFFRIAU W, et

al. A mixed integer programming approach to the aircraft weight and balance problem[J]. Procedia Social and Behavioral Sciences, 2011, 20: 1051-1059.

- [16] VANCROONENBURG W, VERSTICHEL J, TAVERNIER K, et al. Automatic air cargo selection and weight balancing: A mixed integer programming approach[J]. Transportation Research Part E:Logistics and Transportation Review, 2014, 65: 70-83.
- [17] ZHAO X L, YUAN Y, DONG Y, et al. Optimization approach to the aircraft weight and balance problem with the centre of gravity envelope constraints[J]. IET Intelligent Transport Systems, 2021, 15(10): 1269-1286.
- [18] LARSEN O, MIKKELSEN G. An interactive system for the loading of cargo aircraft[J]. European Journal of Operational Research, 1980, 4(6): 367-373.
- [19] LURKIN V, SCHYNS M. The airline container loading problem with pickup and delivery[J]. European Journal of Operational Research, 2015, 244(3): 955-965.
- [20] BRANDT F. The air cargo load planning problem[D]. Karlsruher: Karlsruher Instituts für Technologie, 2017: 61-70.
- [21] CHEN C S, LEE S M, SHEN Q S. An analytical model for the container loading problem[J]. European Journal of Operational Research, 1995, 80(1): 68-76.
- [22] LIU D S, TAN K C, HUANG S Y, et al. On solving multiobjective bin packing problems using evolutionary particle swarm optimization[J]. European Journal of Operational Research, 2008, 190(2): 357-382.
- [23] YAN S Y, SHIH Y L, SHIAO F Y. Optimal cargo container load-

ing plans under stochastic demands for air express carriers[J]. Transportation Research Part E:Logistics and Transportation Review, 2008, 44(3): 555-575.

- [24] LI Y Z, TAO Y, WANG F. A compromised large-scale neighborhood search heuristic for capacitated air cargo loading planning[J]. European Journal of Operational Research, 2009, 199(2): 553-560.
- [25] TANG C H. A scenario decomposition-genetic algorithm method for solving stochastic air cargo container loading problems[J]. Transportation Research Part E:Logistics and Transportation Review, 2011, 47(4): 520-531.
- [26] PAQUAY C, SCHYNS M, LIMBOURG S. Three dimensional bin packing problem applied to air cargo[C]//Proceedings of the 4th International Conference on Information Systems, Logistics and Supply Chain. Quebec: Scientific Congresses and Symposiums, 2012: 1-6.
- [27] PAQUAY C, SCHYNS M, LIMBOURG S. A mixed integer programming formulation for the three-dimensional bin packing problem deriving from an air cargo application[J]. International Transactions in Operational Research, 2016, 23(1-2): 187-213.
- [28] PAQUAY C, LIMBOURG S, SCHYNS M. A tailored two-phase constructive heuristic for the three-dimensional multiple bin size bin packing problem with transportation constraints[J]. European Journal of Operational Research, 2018, 267(1): 52-64.
- [29] PAQUAY C, LIMBOURG S, SCHYNS M, et al. MIP-based constructive heuristics for the three-dimensional bin packing problem with transportation constraints[J]. International Journal of Production Research, 2018, 56(4): 1581-1592.

Cooperating loading balance optimization for medium-sized aircraft with multiple flight legs based on loading and unloading sequence

ZHAO Xiangling^{1,*}, LI Yunfei², WANG Zhiyu³, XU Jihui², LI Pengfei⁴

(1. College of Air Traffic Management, Civil Aviation University of China, Tianjin 300300, China;

2. Equipment Management and Unmanned Aerial Vehicle Engineering College, Air Force Engineering University, Xi'an 710000, China;

3. Zhuhai Xiang Yi Aviation Technology Co., Ltd., Zhuhai 519030, China;

4. College of Management and Economics, Tianjin University, Tianjin 300072, China)

Abstract: For air cargo flights with multiple legs and multiple stops, it is necessary to reasonably control the center of gravity (CG) position of the aircraft, and maintain a balanced state during the flight; during the loading and unloading operation at intermediate airports, it is important to avoid extra loading and unloading operations, so as to reduce waste of time and labor. Therefore, studying the cooperating loading balance problem of multiple flight legs flights was of great practical significance. In this paper, according to the aircraft's characteristics and the sequence of unit load device (ULD) loading and unloading, and combined with the loading and unloading operation requirements for the aircraft loading balance in each flight leg and ULD in the over-station airport, a multi-objective integer linear programming model for connecting flights is constructed by cooperating and distributing the positions of ULD and bulk cargo in each flight leg. The model takes into account the weight, volume and matching restrictions of the ULD and bulk cargo. It considered the constraints of aircraft weight limitation, cabin weight limitation, cabin and regional cumulative weight limitation, upper and lower cabins combined weight limitation, and the CG location balance limit. Additionally, the model considers the connectivity constraints of ULD and bulk cargoes in the front and rear legs. According to the two reasons of the extra loading and unloading operation of the intermediate airports, and based on the loading and unloading sequence, the loading and unloading optimization model is established, the CG is optimized and the number of loading and unloading is reduced through the translation of the ULD in the cabin. Taking the B757-200F model as an example, the commercial solver Gurobi was used for the three scenarios of two loadings and one unloading, one loading and two unloadings and two loadings and two unloadings. For three different combinations of objective functions are solved, analyzed and compared. Experimental results show that the model can effectively cooperate and optimize additional loading and unloading operations and CG location of the front and rear legs.

Keywords: air transportation; air cargo; multiple flight legs; cooperating loading balance; weight and balance; multi-objective; integer linear programming

URL: link.cnki.net/urlid/11.2625.V.20220826.1430.005

Received: 2022-05-29; Accepted: 2022-08-05; Published Online: 2022-08-26 16:27

Foundation items: National Natural Science Foundation of China (71802141,52074309,U2333217); Improvement of Loading Balance Management System for China Southern Airlines (KC22204); The Fundamental Research Funds for the Central Universities (3122020049); Graduate Research and Innovation Funding Project of Civil Aviation University of China (2021YJS060)

^{*} Corresponding author. E-mail: zxl-llx@163.com

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

DOI: 10.13700/j.bh.1001-5965.2022.0452

北斗三号仅用作追踪的机载设备适航要求分析

马振洋^{1,2},周中华³,张帆^{1,2,*},王鹏^{1,2},柯炳清³

(1. 中国民航大学 安全科学与工程学院,天津 300300; 2. 中国民航大学 民航航空器适航审定技术重点实验室,天津 300300;
 3. 航天恒星科技有限公司,北京 100094)

摘 要:随着北斗三号卫星导航系统服务正式开通,北斗二号相关适航标准已无法满足当前机载设备应用场景需求,严重制约了北斗在民航领域的应用进程,急需开展北斗三号相应标准的研究与制定。针对北斗在民用航空器中应用第1阶段"仅用作航空器追踪",对北斗二号与北斗三号 开展差异性分析,提出标准修订总体方案,综合考虑适航安全性、机载设备应用场景,结合工业实践,根据设备构型提出指标修订方案,形成审定要素及符合性方法建议,为设备研发与适航取证提供标准支持。仿真分析了设备天线增益、卫星可见性等关键参数,验证了所修订技术指标的正确性。

关键词:技术标准规定;北斗三号;航空器追踪;定位;短报文;适航

中图分类号: V243.1

文献标志码: A 文章编号: 1001-5965(2024)04-1162-14

MH370事件的发生引起了全球民航界对航空 器追踪监控的重点关注。航空器追踪指由航空承 运人按标准的时间间隔,针对每架飞行中的航空器 在地面记录并更新航空器 4D 位置信息的过程^[1]。 国际航空运输协会(International Air Transport Association, IATA)在 MH370事件发生后于 2014年 4月成立航空器追踪任务组,旨在研究使用现有技 术在短期内解决航空器追踪问题,并对各种方案进 行审查以制定航空器追踪性能标准建议。2015年 11月,国际民航组织(International Civil Aviation Organization, ICAO)理事会通过了《国际民航公约》 附件 6 第 I 部分的第 39次修订,制定了例行航空 器追踪规范,并强制要求航空承运人在 2018年 11月 8 日前,通过自动报告实现其在海洋区域运行 至少每 15 min 航空器位置 4D/15 追踪^[2]。

美国联邦航空管理局 (Federal Aviation Administration, FAA) 发布的咨询通告给出了提供监测手段机载设备的适航批准方法^[3]。中国民用航空局于 2015 年成立中国民航航空器追踪监控体系建

设领导小组,提出"建成中国民航航空器追踪监控体系"的总体目标。2016年8月,中国民用航空局明确要求对于中国情报区以内运行的航空器,航空承运人可通过空管单位的现有监视手段实现航空器追踪^[1]。

为推动航空器运行追踪持续监控,2019年8月 14日,中国民用航空局正式颁布了基于北斗二号的 《仅用作航空器追踪的北斗卫星导航系统(BDS)机 载设备》(CTSO-2C604)^[4],对北斗机载设备所包含 的卫星无线电导航服务(radio navigation satellite service, RNSS)定位单元及天线、卫星无线电测定 服务(radio determination satellite service, RDSS)通信 单元及天线,提出了最低性能要求及测试方法,解 决了民航对北斗机载设备的无标准可依的紧迫需 求问题,推动了以北斗机载设备为代表的国产设备 在民航领域的应用^[5]。中国民用航空局牵头组织了 北斗在运输类飞机的应用示范项目,围绕"1个地 面数据处理系统+20 架运输飞机加改装北斗机载设 备+示范飞行试验"目标,以CTSO-2C604标准^[4]作

*通信作者. E-mail: zhangfan621@foxmail.com

收稿日期: 2022-05-31; 录用日期: 2022-06-11; 网络出版时间: 2022-06-22 15:41 网络出版地址: link.enki.net/urlid/11.2625.V.20220621.1954.005

基金项目:国家自然科学基金(U2133203);中国民用航空局民航安全能力项目(2020-142)

引用格式: 马振洋,周中华,张帆,等.北斗三号仅用作追踪的机载设备适航要求分析 [J].北京航空航天大学学报,2024,50(4):1162-1175. MAZY, ZHOUZH, ZHANGF, et al. Analysis of airworthiness requirements of Beidou-3 airborne equipment only used for tracking [J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50(4): 1162-1175 (in Chinese).

为审定基础,开展相应北斗机载设备适航审查活动,国内已有3家供应商的设备取得中国民用航空技术标准规定项目批准书(China Civil Aviation Technical Standard Order Authorization, CTSOA);同时,以波音B737NG、空客A320s系列2个主力机型为应用载体,开展北斗设备装机方案研究与适航符合性验证,完成了2个机型各10架机的北斗机载设备装机,改装方案获得了局方补充型号合格证(supplementary type certificate, STC)并投入运营,在亚太区域首次实现了对运输飞机的独立、自主、可控的安全监视与追踪,1 min 间隔下的短报文通信成功率约94%,具备独立的航班安全监视与追踪能力^[6]。

2020年7月31日,北斗三号全球卫星导航系 统正式开通。北斗系统在轨运行服务卫星共 45颗,包括15颗北斗二号卫星和30颗北斗三号卫 星,联合为用户提供7种服务。2020年11月,在 ICAO导航系统专家组第6次全体会议上,中国民 用航空局空管行业管理办公室和中国卫星导航系 统管理办公室共同组织的北斗ICAO标准化工作团 队成功推进北斗三号系统全部性能指标完成专家 技术验证^[7],标志着将北斗系统正式写入ICAO标 准,这个最核心、最主要工作已经完成。

当前,北斗卫星导航系统(BeiDou Navigation Satellite System, BDS)处于过渡阶段,北斗二号运控 系统相关工作交接给北斗三号运控系统承担,北斗 二号空间段卫星不再补充建设,北斗二号 RNSS 和 RDSS 将随着北斗二号卫星寿命结束而终止。目 前,基于北斗二号的 CTSO-2C604 标准^[4]已在航空 工业领域得到充分工程应用与实践,但仍存在一定 局限性。由于标准仅支持北斗二号体制,仅通过 B1I 频点提供 RNSS, 而随着北斗 ICAO 标准化工作 的推进, B1I、B1C和B2a已正式写入ICAO标准和 建议措施 (standards and recommended practices, SARPs)^[8]中,因此,CTSO-2C604标准^[4]难以与BDS 的国际标准相接轨,限制了其在国内航空领域的应 用和发展。此外,标准仅支持亚太地区 RDSS,在当 前北斗三号已经建成并正式提供服务下,无法为全 球 RDSS 提供最低性能要求及测试方法依据, 仅支 持北斗二号的北斗机载设备将无法继续为用户提 供期望的功能与服务,影响了北斗在民航领域的应 用进程。因此,基于北斗三号的 CTSO-2C604a 标准^[9] 于 2022 年 3 月正式发布。针对上述问题,本文分 析了该标准最低性能指标要求及测试方法,并探索 基于北斗三号的机载设备适航符合性验证路径,研 究关键技术指标仿真与计算方法,为建成基于自主

知识产权的航空器追踪监控体系奠定基础,进一步 推进北斗适航标准在国际上的应用与引领。

1 北斗二号与北斗三号差异性分析

1.1 服务类型差异分析

根据 BDS"三步走"发展战略,中国于 2012 年 建成北斗二号系统,向亚太地区提供定位、测速、 授时和短报文通信服务;2020 年建成北斗三号系 统,为全球用户提供定位导航授时、全球短报文通 信和国际搜救服务,同时可为中国及周边地区用户 提供星基增强、地基增强、精密单点定位和区域短 报文通信等服务^[10]。

针对仅用作航空器追踪的北斗机载设备提供 RNSS和RDSS两大服务。北斗二号和北斗三号为 民航提供开放的系统服务差异如表1所示。针对 RNSS, BDS可以向位于地表及其以上1000km 空间的全球用户提供免费RNSS,由北斗二号和北 斗三号星座联合提供^[11]。针对RDSS,北斗三号区 域短报文通信服务(regional short message communication, RSMC),服务容量已提高到10⁷次/h,接收 机发射功率降低到1~3W,单次通信能力1000个 汉字;此外,开通了全球短报文通信服务(global short message communication, GSMC),单次通信能 力40个汉字^[12]。

1.2 频点差异分析

1) RNSS。北斗二号 RNSS 仅提供 B1I 信号频 点。北斗三号提供 5 个信号频点, 如表 2 所示。

北斗三号 RNSS 模式分为单频和双频 2 种模 式。单频信号推荐使用 BII 或 BIC 信号, 双频信号 推荐使用 BII 和 B3I 组合信号或 BIC 和 B2a 组合 信号^[14]。对于双频服务,由于其主要是在单频服务 精度无法满足要求时,利用空间相关性,通过双频 来消除或降低卫星星历误差及电离层、对流层延时 等误差来提高服务性能。因此,针对"仅用作航空 器追踪"仅采用单频服务即可满足定位精度要求。 同时,中国民用航空局北斗机载设备应用示范项目 采用的设备均为单频服务,飞行数据表明单频可以 满足航空器定位追踪需求。因此, RNSS 采用单频 模式,沿用基于北斗二号的 CTSO-2C604 标准^[4] 中 BII 信号,并纳入 BIC 信号作为定位频点。

2) RDSS。北斗二号仅向中国及周边地区提供 RSMC 服务,采用上行 Lf0 发射频点,下行 S11 接收 频点;北斗三号 RSMC 采用上行 Lf1、Lf2 发射频 点,下行 S2C 接收频点,GSMC 采用上行 Lf4 发射 频点,下行 GSMC-B2b 接收频点。表 3 总结了基于 北斗二号和北斗三号的仅用作航空器追踪的北斗 机载设备所采用的空间信号。

表1 北斗二号与北斗三号系统提供服务差异分析

Table 1 Analysis of differences in service provided by Beidou-2 and Beidou-3 systems

肥友米刑	1	信号频点	卫星		
瓜牙矢型	北斗二号	北斗三号	北斗二号	北斗三号	
PNISS(今球苏国)		B1I, B3I		3GEO+3IGSO+24MEO	
KN55(主承径图)		B1C, B2a, B2b		3IGSO+24MEO	
GSMC(全球范围)		L:上行		14MEO:上行	
(15)((主环范围)		GSMC-B2b:下行		3IGSO+24MEO:下行	
PSMC(中国及周边地区)	L:上行	L: 上行	2050	2050	
KSMC(中国及间边地区)	S:下行	S: 下行	SGEO	3GEO	
DNCC(中国及国为地区)	DU	B1I, B3I	2CEO - 2ICSO - 24MEO	3GEO+3IGSO+24MEO	
KNS5(平国及问辺地区)	вП	B1C, B2a, B2b	30E0+, $31030+$, $24ME0$	3IGSO+24MEO	

注:中国及周边地区指东经75°~135°,北纬10°~55°的区域^[13];MEO、GEO、IGSO分别为卫星绕地轨道中的中地球轨道、地球静止轨道及倾斜地球同步轨道。

表 2 北斗三号 RNSS 信号频点

Table 2 Signal frequency points of Beidou-3 RNSS

信号频点	中心频率/MHz	带宽/MHz	调制方式
B1C	1 575.42	32.736	B1C_data: BOC (1,1)
			B1C_pilot: QMBOC(6,1,4/33)
B2a	1 176.45	20.46	BPSK (10)
B2b	1 207.14	20.46	BPSK (10)
B1I	1 561.098	4.092	BPSK
B3I	1 268.52	20.46	BPSK

注:QMBOC、BPSK分别为正交复用二进制偏移载波调制及二进制相移键控调制。

表 3 信号频点			
Table 3 Signal frequency points			
服务类型	北斗二号	北斗三号	
DNCC (今戒志国)		B1I: (1 561.98±2.046) MHz	
KN35(主环范围)		B11: (1 561.98±2.046) MHz B1C: (1 575.42±16.368) MHz 上行Lf4: (1 624.524±1.637 6) MHz	
CSMC(令球茲国)		上行Lf4: (1 624.524±1.637 6) MHz	
USMC(主体范围)		下行GSMC-B2b: (1 207.14±10.23) MHz	
DSMC(中国及国法地区)	上行L: (1615.68±4.08) MHz	上行:Lf1: (1614.26±4.08) MHz, Lf2: (1618.34±4.08) MHz	
KSMC(中国及周辺地区)	下行S: (2 491.75±4.08) MHz	下行S: (2 491.75±8.16) MHz	
DNCC(由国及国进地区)	D1L (1561.08+2.046) MIL	B1I: (1 561.98±2.046) MHz	
KINO3(平西汉间辺地区)	D11: (1 301.96±2.040) MHZ	B1C: (1 575.42±16.368) MHz	

2 基于北斗三号的标准修订总体方案

2.1 CTSO-2C604a 标准总体框架

基于北斗三号的仅用作航空器追踪的北斗机 载设备提供两大服务: RNSS 定位服务和 RDSS 短 报文通信服务。标准总体框架如图 1 所示。

RNSS 定位服务包含 RNSS 定位天线和 RNSS 定位单元。

RNSS 定位天线。包含有源和无源定位天线,支持 BII或 BIC 频点。无源部分标准修订认可 RTCA DO-228^[15];有源部分标准修订认可 RTCA DO-301^[16]。

2) RNSS 定位单元。定位单元参照 BD420011-2015^[17]。

RDSS 短报文通信服务包含 RDSS 通信天线和 RDSS 通信单元,涵盖全球和区域短报文服务。 1) RDSS 通信天线。下行接收天线:①区域短 报文接收天线包含下行 S 频点有源/无源接收天线, 参照 CTSO-2C604^[4]确定最低性能标准,参照 BD420004-2015^[18]确定测试方法;②全球短报文接 收天线为 B2b 频点有源/无源天线,参照 CTSO-2C604^[4]确定最低性能标准,参照 BD420004-2015^[18] 确定测试方法。上行 L 发射天线:①区域短报文发 射天线分为Lf1或Lf2 频点有源/无源天线,参照 CTSO-2C604^[4]确定最低性能标准,参照 BD420004-2015^[18] 确定测试方法;②全球短报文发射天线分为Lf4 频 点有源/无源天线,参照 CTSO-2C604^[4]确定最低性 能标准,参照 BD420004-2015^[18]确定测试方法。

2) RDSS 通信单元。区域和全球短报文服务 通信单元均参照 BD420007-2015^[19] 的组成、一般要 求、功能要求和性能要求等标准制定维度开展北斗 三号 CTSO 指标和测试方法等标准制定工作。



图 1 CTSO-2C604a 标准^[9] 总体框架 Fig. 1 Overall framework of CTSO-2C604a standard^[9]

2.2 设备构型

根据 2.1 节北斗三号服务与频点分析结论,基 于北斗三号的仅用作航空器追踪的北斗机载设备 类别划分如表 4 所示。

表4 设备类别

E and a sta

Table 4 Equipment category				
服务范围	设备	型别	应用范围	
			仅支持区域服务,仅支持B1I频点	
区域	R	RC	仅支持区域服务,仅支持B1C频点	
		RIC	仅支持区域服务,支持B11、B1C频点	
			仅支持全球服务,仅支持B1I频点	
全球	G	GC	仅支持全球服务,仅支持B1C频点	
		GIC	仅支持全球服务,支持B11、B1C频点	
		RGI	支持区域+全球, 仅支持B1I频点	
区域+全球	RG	RGC	支持区域+全球, 仅支持B1C频点	
		RGIC	支持区域+全球,支持B1I、B1C频点	

设备分类取决于系统提供的服务范围及 RNSS能够接收的北斗公开服务信号。R类仅支持 区域服务;G类仅支持全球服务;RG类既支持区域 服务也支持全球服务。I是 RNSS 仅能接收 B11 信 号的 BDS 机载设备;C是 RNSS 仅能接收 B1C 信号 的 BDS 机载设备;IC是 RNSS 既能接收 B11 信号也 能接收 B1C 频点的 BDS 机载设备。其中,RG类设 备在东经 75°~135°、北纬 10°~55°区域优先使用 RSMC,超出该范围使用GSMC。

3 标准修订及关键参数仿真分析

3.1 标准修订总体情况

CTSO-2C604a标准^[9]保持原CTSO-2C604标准^[4] 基本框架不变,仍以附录形式分别对设备4个单元 提出最低性能要求,并提供建议的测试方法。由于 RNSS定位单元在功能上无更改,附录B与CTSO-2C604标准^[4]保持一致,无修订。其他单元根据新 增频点、服务范围及北斗三号系统特性,结合工业 实践开展修订。为保证标准过渡一致性,修订后的 标准能够兼容北斗三号和北斗二号的特征。总体 修订情况如图2所示。

3.2 RNSS 定位天线

3.2.1 RNSS 无源定位天线指标修订分析

1) B1I 频点无源定位天线。RNSS B1I 频点无 源定位天线要求, 经差异性分析, 所有要求均与 CTSO-2C604 附录 A1.1^[4]保持一致, 因此, 附录 A1.1 未进行修订。

2) B1C 频点无源定位天线。基于北斗二号的 CTSO-2C604^[4] 不含 B1C 频点,因此,该附录内容主 要基于 RTCA DO-228^[15] 标准,并结合 CTSO-2C604^[4] 工业实践进行修订。修订内容除 2.2.1.1 节工作频 率、2.4.1.9 节试验频率与 RTCA DO-228^[15] 标准保



图 2 CTSO-2C604a 标准^[9] 修订总体情况

Fig. 2 General revision of CTSO-2C604a standards^[9]

持一致外,其他内容与 CTSO-2C604 附录 A2.1^[4]保 持一致。

3) 支持 B1I、B1C 频点的无源定位天线。对于 同时支持机载卫星定位功能的 B1I 和 B1C 频点无 源天线,应分别满足 B1I 频点无源天线(CTSO-2C604 附录 A1.1^[4])和 B1C 频点无源天线(CTSO-2C604 附录 A2.1^[4])的要求。

3.2.2 RNSS 有源定位天线指标修订分析

CTSO-2C604 附录 A1.2^[4] 是 RNSS 定位天线在 B1I 频点的最低性能要求及测试方法。CTSO-2C604 附录 A2.2^[4] 是 RNSS 定位天线在 B1C 频点的最低 性能要求及测试方法。基于北斗二号的 CTSO-2C604^[4] 不含 B1C 频点,该附录内容主要基于 RTCA DO-301^[16] 标准,并结合 CTSO-2C604^[4] 进行修订。 修订内容及原因如表 5 所示。

3.2.3 天线增益仿真与卫星可见性分析验证

目前,标准采用的 BII 频点由北斗二号和北斗 三号共同提供,BII 频点机载设备已在运输类飞机 广泛应用,且得到认可。考虑到 BIC 频点为新增频 点,且 BIC 中心频点与 GPS L1 中心频点相同,美国 联邦航空管理局 (Federal Aviation Administration, FAA)通过 RTCA DO-228^[15]和 RTCA DO-301^[16]相 关 TSO 标准,认可了 L1 频点,因此,综合考虑飞行 安全及社会各界的紧迫需求,通过仿真分析验证单 天线和多天线下 B1C 频点的天线增益。

1) 落地功率及灵敏度指标分析。

①条件 1: 根据文献 [20], 当卫星仰角大于 5°, 在地球表面附近接收机右旋圆极化天线为 0 dBi 增 益(或线性极化天线为 3 dBi 增益)时, 卫星发射的 导航信号到达接收机天线输出端的最小功率电平 为-159 dBW(MEO 卫星)/-161 dBW(IGSO 卫星)。

②条件2:根据文献[21],卫星发射的BII信号到 达接收机天线输出端的最小功率电平为-163 dBW。

③条件 3: 根据文献 [9], RNSS 定位天线 B1C 频点增益在卫星仰角 5°时为-7.5 dBic, 经过 RNSS 定位天线接收后, 卫星发射的导航信号 B1C 频点到 达天线输出端的最小功率电平为-166.5 dBW(MEO 卫星)/-168.5 dBW(IGSO 卫星)。

④条件 4: 根据文献 [9, 22], RNSS 单元捕获灵
 敏度为-163 dBW, 跟踪灵敏度为-168 dBW。

2) B1C 频点天线增益仿真。

图 3 为单天线和多天线情况下 B1C 频点天线 增益仿真结果。不同仰角下的单、多天线增益分别 如表 6 和表 7 所示。结果表明,单天线下 5°仰角 时,天线增益能够达到约-4.5 dBi,满足 RTCA DO-301^[16]大于-5.5 dBi 的要求,但考虑多天线叠加后,

1167

表 5 RNSS 有源定位天线标准修订分析

Table 5 Standard revision analysis of RNSS active positioning antenna

修改项	CTSO-2C604a ^[9] 修改	修改原因	修改范围
2.2.3.1节 天线单元方向图	删除0°仰角场型比指标	①参考AC 20-138A ^[3] ,由于5°以下仰角存在地面干扰、多径效应,如果5° 仰角以下卫星参与解算,会导致定位误差增大,因此,5°仰角以下的卫星 不会参与定位解算;②由于北斗短报文设备是多频天线,在增加B2b频点 后,会增大实现0°仰角相对增益难度,此外该指标应用场景为进近阶段 导航场景,标准旨在解决定位追踪问题,因此,从工业实践角度可不考核 该指标	①B1I (CTSO-2C604 附录A1.2 ^[4]); ②B1C (CTSO-2C604 附录A2.2 ^[4])
2.2.5节 G/T比	将不小于-32.6 dB/K改为不 小于-33.6 dB/K。删除"在 所有方位角、频率范围 (1575.42±2) MHz和整个环 境温度范围内,5°仰角 G/T比不应小于 -31.6 dB/K"	①GPS的L1信号与B1C信号频谱有差异,B1C频点在(1575.42±2) MHz 无信号,不需要考核;②RNSS定位天线G/T值取决于天线的增益和噪声 两大因素,考虑到多天线设计制约并结合工程实践经验,标准将天线增 益在RTCA DO-301 ^[16] 基础上降低2 dB,噪声水平基本相当,因此,RNSS 定位天线G/T值也相应改为不小于-33.6 dB/K	B1C (CTSO-2C604 附录A2.2 ^[4])
2.2.6.1节 带宽内最小增益	修改增益测量频率为 (1575.42±7.5) MHz	①RTCA DO-301 ^[16] 是根据GPS L1信号特性制定的,GPS的L1信号采用 BPSK调制,主信号带宽为±1.023 MHz。在RTCA DO-301 ^[16] 中,其参数考 核带宽主要针对该主信号,其考核带宽为±2 MHz。而B1C信号的信号频 谱与GPS的L1信号完全不同,B1C导航信号在设计时为了不对GPS L1信 号构成干扰,其频谱设计为分裂型,避开了GPS L1信号的主频点,从而导 致B1C信号在(1575.42±2) MHz范围内功率很弱,因此,在针对B1C的产 品标准中,其参数考核带宽不应再局限于±2 MHz。②根据《北斗卫星导 航系统空间信号接口控制文件公开服务信号B1C(1.0版)》,已明确B1C 信号采用BOC(1,1)+BOC(6,1)的调制方式 ^[20] ,根据BOC调制频谱特性,可 计算出其主信号带宽为±(6.138+1.023) MHz=±7.161 MHz,因此,可 把±7.5 MHz作为采用B1C信号的指标考核带宽	B1C (CTSO-2C604 附录A2.2 ⁽⁴⁾)
2.2.6.2节 有源子组件增益	修改增益频率为(1575.42 ±7.5) MHz带宽内≥26.5 dB	同2.2.6.1节	B1C (CTSO-2C604 附录A2.2 ^[4])
2.2.6.3节 带内增益压缩点	删除"在1000~ 1315 MHz之间,最小限值 应为+23 dBm"	B1C和B1I频响在1000~1315 MHz频点与B2b工作频率冲突	①B1I (CTSO-2C604 附录A1.2 ^[4]); ②B1C (CTSO-2C604 附录A2.2 ^[4])
2.2.11.2节 群时延与视线角	删除	①根据AC 20-138A ^[3] ,差分群时延需求为针对高精度精密进近场景,如 LPV和GLS。本标准为定位追踪,无需满足该场景指标。②测试环境要 求非常高,测试成本高,多频段天线难以实现该需求,导致研发费用高	①B1I (CTSO-2C604 附录A1.2 ^[4]); ②B1C (CTSO-2C604 附录A2.2 ^[4])
2.2.12节 直流电源接口	删除	①在机载GPS导航标准体系中,导航天线标准和导航接收机标准是独立 的,需要考虑导航天线和导航接收机的互换性,电源接口参数一致性是 实现互换的基础条件,因此RTCA DO-301 ¹⁶ 中存在该条需求。但CTSO- 2C604标准 ^[4] 是把接收机和天线作为系统来设计的,该指标为系统指标, 因此不需要对设备之间接口做约束。限定天线电源接口电压会限定设 备的具体实现方案。②由于定位追踪设备还需要发射信号,不可能在 4.5 V电压输入时实现大功率发射。如果单独对L的发射信号电压做约 束,会存在接口电压参数相矛盾的问题	①B1I (CTSO-2C604 附录A1.2 ^[4]); ②B1C (CTSO-2C604 附录A2.2 ^[4])
2.4.1.9节 试验频率	将"最低17个频率下"改为 "最低5个频率下";将"最 低17个频率下"改为"最低 9个频率下"	RTCA DO-301 ^[16] 中测试带宽为32 MHz, 一共测试17个点, 间隔是2 MHz。 B1I频段一共只有4 MHz带宽, 测试5个点即可满足B1C频段一共只有 15 MHz带宽, 测试9个点即可: 1 575.42 MHz、(1 575.42±2) MHz、 (1 575.42±4) MHz、(1 575.42±6) MHz、(1 575.42±7.5) MHz	①B11 (CTSO-2C604 附录A1.2 ⁽⁴⁾); ②B1C (CTSO-2C604 附录A2.2 ⁽⁴⁾)
2.4.2.2.6节 差分群时延要求	删除"A.2.2.11.2节的要求"	该指标和2.2.11.2节群时延指标相对应,群时延指标删除,则应删除本条 中A.2.2.11.2节	①B1I (CTSO-2C604 附录A1.2 ^[4]); ②B1C (CTSO-2C604 附录A2.2 ^[4])
2.4.2.6.1节 视轴增益压缩点 与试验	删除2.4.2.6.1节表2-5中 试验1	该指标与2.2.6.3节带内增益压缩点指标 "1 000~1 315 MHz对应", 删除 该内容, 则此条试验1删除	①B1I (CTSO-2C604附录 A1.2 ⁽⁴⁾); ②B1C



Fig. 3 Gain simulation of B1C frequency point antenna (right-handed circular polarization, frequency 1.575 GHz)

表 6 B1C 频点个同似角ト	的甲大线增益
-----------------	--------

Table 6	Single-antenna gain of B1C frequency point at
	different elevation angles

增益点	仰角/(°)	增益/dBi
m1	0	5.253 1
m2	-60	-0.607 3
m3	-70	-2.284 4
m4	-75	-3.155 1
m5	-80	-4.042 7
m6	-85	-4.452 9

表 7 B1C 频点不同仰角下的多天线增益

 Table 7
 Multi-antenna gain of B1C frequency point at

 different elevation angles

unter ent elevation angles			
增益点	仰角/(°)	增益/dBi	
m1	0	3.911 5	
m2	-60	-1.528 5	
m3	-70	-3.4194	
m4	-75	-4.405 2	
m5	-80	-5.398 4	
m6	-85	-6.3872	

B1C 频点 5°仰角下天线增益降为约-6.4 dBi。

结合天线增益仿真结果、卫星发射的导航信号的落地功率(条件1~条件3)和 RNSS 单元的灵敏 度指标(条件4), RNSS 单元的捕获上升卫星仰角 屏蔽角将变为约20°, 而下降卫星仰角屏蔽角针对 MEO卫星可以小于5°, IGSO卫星可以达到约5°。 受此影响, 需进一步针对 RNSS 单元在全球范围内 的可见星情况进行仿真。

3) 卫星可见性仿真。

对全球范围内地面附近位置目标的卫星可见

性进行仿真,只用北斗三号卫星,MEO 卫星和 IGSO 卫星播发 B1C 频点,设置仰角屏蔽角为 20°,地球半 径为6 370 km, IGSO 卫星轨道高度为 35 786 km, MEO 卫星轨道高度为 21 528 km。仿真时间为 2021 年 10 月 8 日 0 点至 2021 年 10 月 9 日 0 点。

为检查全球范围内 RNSS 单元的卫星可见性, 将全球按 1°×1°网格进行划分,对每个网格 24 h内 可见卫星个数进行统计。BDS MEO 1~24 号和 IGSO 1~3 号卫星在全球区域范围内最好累积覆盖 结果如图 4 所示,可见星数范围为 7~11 颗。



图 4 全天最好累积覆盖结果



BDS MEO 1~24 号和 IGSO 1~3 号卫星在全球区域范围内最差累积覆盖结果如图 5 所示,可见星数范围为4~8 颗。

BDS MEO 1~24 号和 IGSO 1~3 号卫星在全球 区域范围内最差累积低于 4 颗卫星情况进行检查, 全球不存在可见星数低于 4 颗的位置, 如图 6 所示。

从仿真分析可知, 在仰角屏蔽角设置为 20°且 按照最差可见星情况下进行统计, RNSS 单元在全 第4期



北斗卫星覆盖数 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14

图 5 全天最差累积覆盖结果

Fig. 5 Worst cumulative coverage results of the whole day





球范围内的可见星数仍大于4颗,可满足航空器追踪定位使用。根据 CTSO-2C604a 标准^[9], RNSS 单元仅上升卫星仰角屏蔽角约为20°,下降卫星仰角屏蔽角仍可达到或优于约5°,因此, RNSS 单元可见星情况将优于以上仿真情况,可满足航空器追踪定位使用。在亚太区域,由于 BDSIGSO 卫星设计,可见星情况将会更优。

3.3 RDSS 通信单元

3.3.1 RDSS 通信单元指标修订分析

由于北斗三号系统提供全球 RDSS,因此,对于 RDSS 通信单元,主要增加了 GSMC 指标要求,同 时结合北斗二号区域 RDSS 服务工业实践,对于部 分指标进行修订,如表 8 所示。

3.3.2 关键参数计算与分析

接收灵敏度指标规定了接收机可以接收到的 并仍能正常工作的最低信号强度。标准增加了 GSMC接收灵敏度指标,本节开展接收灵敏度指标 计算,为指标制定提供依据。同时,为保证接收机 能接收到并且成功解调通信电文,对通信传输链路 余量进行分析与评估,从而为整个系统的可行性和 实用性提供保证。

输入参数设置如表9所示。RSMC服务根据 BDS信号体制采用LDPC编码,GSMC服务采用 Turbo编码,分别计算接收端的解调门限(每比特能 量对噪声功率谱密度比值 *E*_b/*N*₀)。接收机噪声温度 按照典型多级接收系统考虑计算。*R*_b为 MEO 和 GEO 专用段通信速率。BD-3 GEO 接收天线增益 按照 60°仰角到 90°仰角增益单调增加考虑,因此, 计算中以最差情况 60°仰角增益值代入计算,保证 指标保守性。BD-3 MEO 接收天线增益只考虑轴 向,使用轴向增益代入计算。AD 量化误差计算得 0.60 dB。

根据式(1),分别计算 MEO 与 GEO 的接收灵 敏度,计算结果与 CTSO-2C604a^[9]要求的结果进行 对比,如表 10 所示。

$$S = E_{\rm b}/N_0 + kTR_{\rm b} + L - G_{\rm A} \tag{1}$$

式中: $T 为 T_a$ 和 T_t 之和, T_t 为多级放大接收电路噪声温度。

 $T_{t} = T_{0} + T_{1}/G_{0} + T_{2}/(G_{0}G_{1}) + T_{3}/(G_{0}G_{1}G_{2})$ (2) 式中: $T_{i}(i=0, 1, 2, 3)$ 为第 *i* 级放大电路的噪声温度; $G_{i}(i=0, 1, 2)$ 为第 *i* 级放大电路增益。

根据《北斗卫星导航系统空间信号接口控制文件—公开服务信号 B2b(1.0版)》^[23],当卫星仰角大于 5°,在地球表面附近的接收机右旋圆极化天线为0dBi增益(或线性极化天线为3dBi增益)时,卫星发射的导航信号到达接收机天线输出端的最小功率电平为-160dBW,考虑加严后指标确定为-159.8dBW。经计算,接收灵敏度能够达到-171.46dBW,可以满足-159.8dBW的指标要求。

进一步计算链路余量,得到 BD-3 GEO 链路余 量为 2.65 dB, BD-3MEO 链路余量为 11.66 dB,满足 RDSS 通信所需的设计余量。考虑到 RSMC 和 GSMC 设备入网测试要求, RDSS 通信单元需要在 更低仰角满足误码率要求, RSMC 入网测试要求在 天线仰角参数方面更为严格,而 GSMC 入网测试在 接收信号功率值、天线仰角等参数方面都会更为严 格,链路设计余量可用于补偿设备入网测试的加严 要求,当链路设计余量不够时,还需要结合天线的 结构空间设计、有源电路设计等再对 RDSS 通信天 线进行再优化。

3.4 RDSS 通信天线

3.4.1 RDSS 通信天线指标修订分析

RDSS 通信天线增加了 GSMC-B2b 接收频点, 上行 L 发射频点与北斗二号有较大区别。针对上 述问题,指标新增与修订情况如表 11 所示。

3.4.2 关键参数计算与分析

针对短报文 GSMC 和 RSMC, RDSS 通信天线 具有 S 频点、B2b 频点和 L 频点;针对导航定位服 务, RNSS 定位天线具有 BII、B1C 频点。为降低设 备改装难度并提高集成度,采用多天线一体化设 计。考虑带来的性能损失,针对单天线和多天线 2 种场景对关键参数进行计算与仿真。

天线 S 频点增益仿真结果如图 7 所示。不同 仰角下的单、多天线增益分别如表 12 和表 13 所

表 8 RDSS 通信单元标准修订分析

Table 8 Standard revision analysis of RDSS communication unit

修改项	CTSO-2C604a ^[9] 修改	修改原因	修改范围
增加:标识	北斗RDSS单元需采用物理标识方式对北斗三号短报文用户终端序列 是进行工产标识	北斗三号短报文用户终端的唯 一身份识别,用于短报文人网	①RSMC(CTSO- 2C604附录C1 ^[4]); ②CSMC(CTSO
	与近行正确你你	认证和北斗卡申请	2C604附录C2 ^[4])
4.3.1节自检与 初始化功能	将 "b) 北斗RDSS单元开机或重新捕获卫星信号后,根据需要,自动或 手动(由用户进行设置)发送一个特定格式的查询申请,查询未接收的 信息"修订为 "b) 北斗RDSS单元开机或卫星信号恢复后,能够跟踪卫	北斗三号取消查询功能,根据 星座状态适应性修改	①RSMC(CTSO- 2C604附录C1 ^[4]); ②GSMC(CTSO-
	星信号"		2C604附录C2 ^[4])
			①RSMC(CTSO-
432节状态检测功能	删除"抑制状态"	北斗三号取消抑制功能,根据	2C604附录C1 ^[4]);
1.5.2 1. 108.12.03 5418		星座状态适应性修改	2)GSMC(CTSO-
			2C604附录C2 ^[4])
4.4.1节接收 灵敏度	增加 "GSMC:接收灵敏度(方位角0°~360°,仰角90°,误码率不大于 1×10 ⁻⁵):不大于-159.8 dBW"	GSMC系统指标中基带处理能 力要求为-157.8 dBW,考虑天 线法向增益,指标加严为 -159.8 dBW	GSMC(CTSO- 2C604附录C2 ^[4])
			①RSMC(CTSO-
4.4.2节接收	①修改4.4.2节"RSMC:北斗RDSS单元接收信道数不小于8";②增加	根据北斗管理机构相关要求修	2C604附录C1 ^[4]);
通道数	"GSMC服务:北斗RDSS单元接收信道数不小于12"	改	2)GSMC(CTSO-
		<u> </u>	2C604附录C2 ^[4])
	①删除4.4.3节中"具备指挥功能的北斗RDSS单元从加电开机至捕获	①系统指标不做要求,由产品	<pre>①RSMC(CTSO-</pre>
4.4.3节首次	北斗GEO卫星RDSS信号并解调出信息所需时间应不大于10s";②将	规范来限定;②GSMC的	2C604附录C1 ^[4]);
捕获时间	捕获北斗GEO卫星KDSS信号 改为 捕获北斗MEO卫星KDSS信 只" 副除在在2节中"目发指挥功能的也计DDSS单元计加中工机石墙	RDSS信亏进过MEO接收,杀	2)GSMC(CTSO-
	亏,删除4.4.5 P中 具备循件切能的几半KDSS毕儿从加电开机主捕	沉信你小悯安尔,田厂而观氾 李阻宁	2C604附录C2 ^[4])
444节重捕菇	从电子GEO卫星RD35旧子开辟两田旧志历而时间型小尺于103	系统指标不做要求 由产品规	
·····································	删除	苏来限定 (DRSMC(CTSO-
	no. i v A	北斗系统会自动调整终端发送2	2C604附录C1 ^[4]);
4.4.8节切放输出切率	删除	功率,设备内部接口不做限定(2)GSMC(CTSO-
4.4.9节发射信号载波	删论	此指标为传统通信方式的延 2	2C604附录C2 ^[4])
相位调制偏差	加115、	续,对北斗三号不再沿用适用	
4.4.10节发射信号 频率准确度	删除	系统指标不做要求,由产品规 范来限定	RSMC(CTSO- 2C604附录C1 ^[4])
5.4.1节自检与初始化 功能测试	将"b)在实际卫星信号下,北斗RDSS单元的数据端口与计算机相连接,将其设置为关机状态,设置另一台终端设备向其发送报文通信,开机后检查是否收到终端发送的报文通信"改为"b)北斗RDSS单元的数据端口与计算机相连接,确认卫星信号播发正常,加电开机后检查北斗RDSS单元能够跟踪卫星信号"	北斗二号具有上电查询功能, 可符合"自检与初始化功能测 试"原要求,北斗三号上电查 询功能取消,根据星座状态适 应性修改	①RSMC(CTSO- 2C604附录C1 ^[4]); ②GSMC(CTSO- 2C604附录C2 ^[4])
		根据《北斗卫星导航系统发展	(I)RSMC(CTSO-
5.4.3 节RDSS业务 服务功能测试	将"检查各项功能是否正常"改为"R类设备仅通过GEO卫星完成, G类设备仅通过MEO卫星完成,RG类设备分别通过GEO卫星和 MEO卫星完成,检查各项功能是否正常"	报告(4.0版)》 ¹⁰ 中表"北斗系 统计划提供的服务类型",不 同服务类型、不同信号频段卫 星播发手段不同	2C604附录C1 ^[4]); ②GSMC(CTSO- 2C604附录C2 ^[4])
	①将"测试次数按照仰角10°、30°、50°、75°,共测4次。样机在不同仰 角下接收测试信号。单次测试采集的电文总和为10 ⁶ ,将北斗RDSS单	星播发手段不同	
5.5.1节接收 灵敏度测试	元接收的出站信息与信号源播发的原始信息进行比较,统计误码率, 其应满足4.4.1节要求"改为"RSMC服务:专用段速率为24 kbps,测试 次数按照仰角60°、70°、80°、90°,共测4次。样机在不同仰角下接收 测试信号。单次测试采集的电文总和为10 ⁶ ,将北斗RDSS单元接收的 出站信息与信号源播发的原始信息进行比较,统计误码率,其应满足 C.4.4.1节要求。S载波电平应设置为-153.8 dBW"。②增加 "GSMC服务:将北斗RDSS单元和模拟信号源连接,模拟信号源播发 B2b频点卫星模拟信号(到天线有源组件入口的信号功率符合 C.4.4.1节要求),测试仰角为90°。单次测试采集的电文总和为10 ⁶ ,将 北斗RDSS单元接收的出站信息与信号源播发的原始信息进行比较, 统计误码率,其应满足C.4.4.1节要求。B2b载波电平应设置为-159.8 dBW"	①与灵敏度指标对应,根据星座状态适应性修改(指标计算过程见3.3.2节);②北斗三号增加GSMC服务,根据星座状态适应性修改(指标计算过程见3.3.2节),与灵敏度指标对应	①RSMC(CTSO- 2C604附录C1 ^[4]); ②GSMC(CTSO- 2C604附录C2 ^[4])
5.5.2节接收 通道数测试	增加"GSMC服务:测试系统播发12个波束信号,测试信号到达被测单 元天线有源组件入口功率为-159.8 dBW。通过串口检测被测单元信 息接收情况,判断并记录被测单元接收通道数,其数值应满足 C.4.4.2节要求"	北斗三号增加GSMC,根据星 座状态适应性修改	GSMC(CTSO- 2C604附录C2 ^[4])
5.5.3节~5.5.10节	删除	系统指标不做要求,由产品规 范来限定	①RSMC(CTSO- 2C604附录C1 ^[4]);

表 9 RDSS 关键参数设置

Table 9 RDSS key parameter setting

E_{b}	N_0	玻尔兹曼	天线	工作	接收机	接收电路	BD-3 MEO	BD-3 GEO	BD-3 GEO(60°)	BD-3 MEO(90°)	AD
BD-3 MEO	BD-3 GEO	常数	噪声温度	温度	总噪声	总噪声	速率	速率	接收天线增益	接收天线增益	量化误差
LDPC	Turbo	<i>k</i> /dB	$T_{\rm a}$ /°C	$T_0/^{\circ}$ C	温度T/℃	温度 T_t /℃	$R_{\rm b}/{\rm bps}$	$R_{\rm b}/{\rm bps}$	$G_{\rm A}/{ m dB}$	$G_{\rm A}/{ m dB}$	L/dB
3.00	3.60	1.3800×10^{-23}	-173.15	16.85	145.16	245.16	1 000.00	24 000.00	3.00	3.60	0.60

表 10 接收灵敏度指标验证

Table 10 Verification of receiving sensitivity indexes

MEO用户终端接	转收功率/dBW	GEO用户终端接收功率/dBW		
指标要求	计算结果	指标要求	计算结果	
≤ −159.80	-171.46	≤ −153.80	-156.45	

表 11 RDSS 通信天线标准修订分析

Table 11 Standard revision analysis of RDSS communication antenna

修改项	CTSO-2C604a ^[9] 修改	修改原因	修改范围
附录4 A4.1 法向	增加 "B2b频点法向极化增益应不小于2	增加GSMC B2b频点接收天线法向极化增益	B2b频点
极化增益	dBic"	要求	(CTSO-2C604附录D2.1 ^[4])
附录4 A4.1 10°仰角	增加 "B2b频点10°仰角平均极化增益应不	增加GSMC B2b频点接收天线10°仰角平均	B2b频点
平均极化增益	小于-9 dBic"	极化增益要求,与S频点要求一致	(CTSO-2C604附录D2.1 ^[4])
	①C题上接收工作版变为(2.401.75)		①S频点
附寻4 4 4 1 工作 标志	15 妙点按收入线的工作频举为(2491.75± 9.16) MU ₂ ⑦ 撞加"P2b 短点接收手线的工	①北斗三号S频点拓宽; ②增加GSMC B2b频	(CTSO-2C604附录D1.1 ^[4]);
附求4 A4.1 工作频平	6.10) MHZ; ②增加 B20频点按收入线的上 作频率 ±(1 207 14±10 22) MHz "	点工作频率	②B2b频点
	作频率为(1207.14±10.25) MHZ。		(CTSO-2C604附录D2.1 ^[4])
			①S频点
附寻4 4 4 1 供由 特州	副除	边冬山郭北岳 天伊西式	(CTSO-2C604附录D1.1 ^[4]);
附求4 A4.1 洪电付住	加州东	以田内即泪你, 小似女水	②B2b频点
			(CTSO-2C604附录D2.1 ^[4])
	①将工作频率改为"Lfl (1614.26±4.08)	①北北二县1 辆占安射王建工作辆索与北北	①Lf1、Lf2频点
附寻4 44 9 工作频率	MHz; Lf2 (1 618.34±4.08) MHz"; ②增加	①北平二与L频点及别入线工作频率可北平 二早不同。⑦增加CSMCI 频点发射于线工	(CTSO-2C604附录D1.2 ^[4]);
附水4A4.2工作频率	"GSMC:L频点发射天线工作频率为Lf4	二 5 不问; ② 审加USIME L频点及别入线上 作频素	②Lf4频点
	(1 624.524±1.637 6) MHz"	11-频平	(CTSO-2C604附录D2.2 ^[4])
			①Lf1、Lf2频点
附录4 44 2 供由特性	叫除	设久内郊护标 不做更求	(CTSO-2C604附录D1.2 ^[4]);
附水4A4.2 洪电付住	加加水	议审内印指你, 小阪安永	②Lf4频点
			(CTSO-2C604附录D2.2 ^[4])
	①RSMC修订为"天线等效全向辐射功率		①141 149555
	(EIRP)不大于10 dBW(方位角0°~360°, 仰	①北斗系统指标中RSMC终端发射功率为	(CTSO_2C604附录D1 2 ^[4]).
附录4 A4.2发射EIRP	角10°~70°)";②增加"GSMC:天线等效全	3~10 W; ②GSMC终端发射功率为10~15	②I f4 新占
	向辐射功率(EIRP)不大于15 dBW(方位角	W	(CTSO-2C604附录D2 2 ^[4])
	0°~360°, 仰角10°~70°)"		(0100-2000-4013(02.2))



Fig. 7 Gain simulation of S-frequency point antenna (right-handed circular polarization, frequency 2.492 GHz)

示。单天线及多天线下 60°仰角时,天线增益约为 3 dBic。单天线及多天线下法向增益大于 3.7 dBic。 基于接收灵敏度链路计算结果,为满足灵敏度 -153.8 dBW 要求,天线 S 频点法向极化增益不小 于 3 dBic 方可满足接收链路约 3 dB 的设计余量要 求,此指标与 RDSS 单元接收灵敏度指标共同对设 备构成约束。

表 12 S 频点不同仰角下的单天线增益

 Table 12
 Single-antenna gain of S-frequency point at different elevation angles

增益点	仰角/(°)	增益/dBic
m1	0	5.274 4
m2	-60	0.551 5
m3	-70	-0.942 9
m4	-75	-1.311 0
m5	-80	-2.415 5
m6	-85	-3.615 7

天线 B2b 频点增益仿真结果如图 8 所示。不同仰角下的单、多天线增益分别如表 14 和表 15 所示。单天线及多天线下轴向 90°仰角时,天线增益约为 3.6 dBic。B2b 频点法向极化增益指标参考 S 频点,并考虑 B2b 频点频率低于 S 频点,且 B2b 频点信号落地功率高于 S 频点,放宽 1 dB 后得到。

表 13 S 频点不同仰角下的多天线增益 Table 13 Multi-antenna gain of S-frequency point at different elevation angles

		-
增益点	仰角/(°)	增益/dBic
m1	0	3.714 7
m2	-60	0.054 2
m3	-70	-1.698 7
m4	-75	-2.633 3
m5	-80	-4.113 9
m6	-85	-5.299 0



图 8 B2b 频点天线增益仿真 (右旋圆标化, 频率为 1.207 GHz)

表 14 B2b 频点不同仰角下的单天线增益

 Table 14
 Single-antenna gain of B2b frequency point at

表 15 B2b 频点不同仰角下的多天线增益 Table 15 Multi-antenna gain of B2b frequency point at

Ċ	lifferent elevation an	gles	different elevation angles			
增益点	仰角/(°)	增益/dBic	增益点	仰角/(°)	增益/dBic	
m1	0	4.986 5	m1	0	3.652 3	
m2	-60	-0.049 9	m2	-60	-1.978 7	
m3	-70	-1.714 9	m3	-70	-3.393 7	
m4	-75	-2.540 3	m4	-75	-4.379 9	
m5	-80	-3.389 5	m5	-80	-5.408 7	
m6	-85	-3.993 7	m6	-85	-6.481 1	

Fig. 8 Gain simulation of B2b frequency point antenna (right-handed circular polarization, frequency 1.207 GHz)

4 审定要素及符合性验证方法

根据基于北斗二号的 CTSO-2C604^[4] 机载设备 CTSOA 适 航 取 证 经 验,基于北斗 三号 的 CTSO-2C604a^[9] 符合性方法(method of compliance, MOC), MC0 为符合性声明, MOC1 为说明性文件, MOC2

为分析/计算, MOC3 为安全评估, MOC4 为实验室 试验, MOC5 为地面试验, MOC6 为飞行试验, MOC7 为航空器检查, MOC8 为模拟器试验, MOC9 为设 备合格鉴定。建议以 MOC4 为主, 结合 MOC0 和 MOC1 开展符合性验证工作。审定要素和建议的 符合性方法如表 16 所示。

表 16 审定要素及符合性方法

Table 16	Certification	elements	and MOC
rable ro	certification	cicinciits	and moc

项目	审定要素	MOC	
	偏离	MOC0	
。真正明书	设备类别、功能、失效状态类别、环境鉴定、软/硬件鉴定、标记、手册、持续适航文件、铭牌图纸、部件 清单、质量系统说明、材料和工艺规范清单等	MOC1	
一败女小	失效状态类别	MOC3	
	功能、环境鉴定、软/硬件鉴定、安装程序和限制	MOC4	
	标记、铭牌	MOC7	
完位于线(无源于线)	适航性、预期功能、阻燃性、工作频率、驻波比和阻抗、天线增益、轴比、极化		
定位天线(有源天线)	适航性、预期功能、阻燃性、工作频率、驻波比和阻抗、天线单元方向图与无源元件增益、极化和轴比、 G/T比、总转换增益与增益压缩、输出负载稳定性、带内增益频率响应、烧毁保护、脉冲功率饱和恢复时 间、群时延	MOC1 MOC4	
定位单元	适航性、预期功能、阻燃性、设备构成、定位、精度、首次定位时间、灵敏度、位置更新率、位置分辨力	MOC1 MOC4	
通信单元	适航性、预期功能、阻燃性、电压驻波比、极化特性与法向轴比、法向极化增益、10°仰角极化增益不圆度、10°仰角平均极化增益、工作频率、噪声系数	MOC1 MOC4	
通信天线	适航性、预期功能、阻燃性、工作频率、极化方式及法向轴比、电压驻波比、法向极化增益、10°仰角极化 增益不圆度、10°仰角平均极化增益、发射EIRP	MOC1 MOC4	

针对 MOC4 中环境试验,由于定位、通信单元 一般安装于舱内,定位和通信天线一般安装于蒙皮 外,应根据安装位置的不同,开展相应环境鉴定试 验项目,环境试验项目如表 17 所示。

表 17 设备环境试验项目 Table 17 Equipment environmental test event

参考DO-160章节 ^[24]	试验项目	定位和 通信单元	定位和 通信手线
4	泪庙知三庙	通用平九	一一一一一
4	個反相回反	N	N
5	温度受化	N	N
6	湿度	\checkmark	\checkmark
7	运行冲击和坠撞安全	\checkmark	\checkmark
8	振动	\checkmark	\checkmark
9	爆炸大气	Х	Х
10	防水	\checkmark	\checkmark
11	流体敏感性	Х	\checkmark
12	砂尘	Х	\checkmark
13	霉菌	Х	\checkmark
14	盐雾	Х	\checkmark
15	磁效应	\checkmark	\checkmark
16	电源输入	\checkmark	Х
17	电压尖峰	\checkmark	Х
18	音频敏感性	\checkmark	Х
19	感应信号敏感性	\checkmark	\checkmark
20	射频敏感性	\checkmark	\checkmark
21	射频能量发射	\checkmark	\checkmark
22	雷电感应瞬态敏感性	\checkmark	\checkmark
23	雷电直接效应	Х	\checkmark
24	结冰	Х	\checkmark
25	静电放电	\checkmark	\checkmark
26	防火及可燃性	Х	\checkmark

注:"√"为需要进行该项试验;"X"为不需进行该项试验。

中国民用航空局在审查过程中,应根据产品规 范的设备类别明确适用的 CTSO 最低性能标准内 容,并参照标准主要修改内容重点开展符合性进行 适航审查;同时, RG 类设备在中国及周边区域涉 及 RSMC 和 GSMC 优先级判定,服务切换设置 0.1° 的切换空间,在此空间内完成切换均可视为符合标 准要求。选择 B1C 频点及 GSMC 的申请人,建议 在局方审查前完成样机研制,并做好符合性验证摸 底;由于北斗短报文服务相关内容暂未固化,建议 2 种北斗用户卡均支持,并尽快完成设备短报文入 网测试,固化设计状态。

5 结 论

本文以 CTSO-2C604 标准为基础,针对仅用作 航空器追踪的北斗机载设备中 RNSS 定位天线、 RNSS 定位单元、RDSS 通信单元、RDSS 通信天线 对其进行修订,提出 CTSO-2C604a 适航要求体系, 并针对上述修订开展仿真,验证关键指标正确性。 在此基础上明确审定要素,并提出符合性验证重点 工作建议,为北斗机载设备研发单位提供研制指导 与适航取证指南。

后续将进一步推进标准应用场景探索,引导工 业方将产品从单一设备向集成式系统扩展,为北斗 民航应用后续阶段以及更高安全等级设备研发奠 定基础,同时推动标准的国际化衔接与引领。

参考文献(References)

- [1] 中国民用航空局. 航空承运人航空器追踪监控实施指南: AC-121-fs-06-127[S]. 北京: 中国民用航空局, 2016.
 Civil Aviation Administration of China. Implementation guide for aircraft tracking and monitoring of air carriers: AC-121-fs-06-127
 [S]. Beijing: Civil Aviation Administration of China, 2016(in Chinese).
- [2] 中国民用航空局.中国民航航空器追踪监控体系建设实施路线 图[S].北京:中国民用航空局,2017.

Civil Aviation Administration of China. Roadmap for the construction and implementation of China civil aviation aircraft tracking and monitoring system[S]. Beijing: Civil Aviation Administration of China, 2017(in Chinese).

- [3] FAA. Airworthiness approval of global navigation satellite system (GNSS) equipment: AC 20-138A[S]. Washington, D. C.: FAA, 2010.
- [4] 中国民用航空局. 仅用作航空器追踪的北斗卫星导航系统 (BDS) 机载设备: CTSO-2C604[S]. 北京: 中国民用航空局, 2019.
 Civil Aviation Administration of China. Beidou satellite navigation system (BDS) airborne equipment used for aircraft tracking only: CTSO-2C604[S]. Beijing: Civil Aviation Administration of China, 2019(in Chinese).
- [5] 马赞,周中华,王鹏,等. 仅用作航空器追踪的北斗机载设备适航 要求分析[J]. 航空学报, 2019, 40(11): 323155.
 MA Z, ZHOU Z H, WANG P, et al. Analysis of airworthiness re-

quirements of BDS airborne equipment for aircraft tracking only[J]. Acta Aeronautica et Astronautica Sinica, 2019, 40(11): 323155(in Chinese).

- [6] 汪万维, 吴仁彪, 孙可, 等. 基于北斗的亚太航班安全监视与追踪
 [J]. 安全与环境学报, 2022, 22(6): 3259-3268.
 WANG W W, WU R B, SUN K, et al. Asia-Pacific flight safety surveillance and tracking based on BeiDou[J]. Journal of Safety and Environment, 2022, 22(6): 3259-3268(in Chinese).
- [7] 陈颖, 卢鋆, 刘成, 等. 北斗卫星导航系统国际民航标准重点问题 研究与指标验证[J]. 天文研究与技术, 2022, 19(5): 447-457.
 CHEN Y, LU J, LIU C, et al. Research and assessment on key issues of ICAO SARPs for BeiDou navigation satellite system[J]. Astronomical Research & Technology, 2022, 19(5): 447-457 (in Chinese).
- [8] ICAO. International standards and recommended practices annex 10 aeronautical telecommunications: Volume I[S]. Montreal: ICAO2018.
- [9] 中国民用航空局. 仅用作航空器追踪的北斗卫星导航系 (BDS) 机载设备: CTSO-2C604a[S]. 北京: 中国民用航空局, 2022. Civil Aviation Administration of China. Beidou satellite navigation system (BDS) airborne equipment used for aircraft tracking only: CTSO-2C604a[S]. Beijing: Civil Aviation Administration of China, 2022(in Chinese).
- [10] 中国卫星导航系统管理办公室.北斗卫星导航系统发展报告 (4.0版)[S].北京:中国卫星导航系统管理办公室, 2019.
 China Satellite Navigation Office. BeiDou navigation satellite system development report (version 4.0)[S]. Beijing: China Satellite Navigation Office, 2019(in Chinese).
- [11] YANG Y X, XU Y Y, LI J L, et al. Progress and performance eva-

luation of BeiDou global navigation satellite system: Data analysis based on BDS-3 demonstration system[J]. Science China Earth Sciences, 2018, 61(5): 614-624.

- [12] LI G, GUO S R, LV J, et al. Introduction to global short message communication service of BeiDou-3 navigation satellite system[J]. Advances in Space Research, 2021, 67(5): 1701-1708.
- [13] 中国卫星导航系统管理办公室.北斗卫星导航系统公开服务性 能规范 3.0 版: BDS-OS-PS-3.0[S].北京:中国卫星导航系统管理 办公室, 2021.

China Satellite Navigation Office. BeiDou navigation satellite system open service performance standard (version 3.0): BDS-OS-PS-3.0 [S]. Beijing: China Satellite Navigation Office, 2021 (in Chinese).

- [14] YANG Y X, GAO W G, GUO S R, et al. Introduction to BeiDou-3 navigation satellite system[J]. Navigation, 2019, 66(1): 7-18.
- [15] RTCA. Minimum operational performance standards for global navigation satellite systems (GNSS) airborne antenna equipment: RT-CA DO-228[S]. Washington, D. C.: RTCA, 1995.
- [16] RTCA. Minimum operational performance standards for global navigation satellite systems (GNSS) airborne active antenna equipment for the L1 frequency band: RTCA DO-301[S]. Washington, D. C.: RTCA, 2006.
- [17] 中国卫星导航系统管理办公室.北斗/全球卫星导航系统 (GNSS)定位设备通用规范: BD420011-2015[S].北京:中国卫星 导航系统管理办公室, 2015.

China Satellite Navigation Office. General specification for BeiDou/global navigation satellite systems (GNSS) positioning devices: BD420011-2015[S]. Beijing: China Satellite Navigation Office, 2015(in Chinese).

- [18] 中国卫星导航系统管理办公室.北斗/全球卫星导航系统导航型 天线性能要求及测试方法: BD420004-2015[S].北京:中国卫星 导航系统管理办公室, 2015.
 China Satellite Navigation Office. Performance requirements and test method for BeiDou/global navigation satellite systems(GNSS) navigation antenna: BD420004-2015[S]. Beijing: China Satellite Navigation Office, 2015(in Chinese).
- [19] 中国卫星导航系统管理办公室.北斗用户终端 RDSS 单元性能 要求及测试方法: BD420007-2015[S].北京:中国卫星导航系统 管理办公室, 2015.

China Satellite Navigation Office. Performance requirements and test method for BDS RDSS unit: BD420007-2015[S]. Beijing: China Satellite Navigation Office, 2015(in Chinese).

- [20] 中国卫星导航系统管理办公室.北斗卫星导航系统空间信号接口 控制 文件公开服务信号 B1C(1.0版): BDS-SIS-ICD-B1C-1.0[S].北京:中国卫星导航系统管理办公室,2017.
 China Satellite Navigation Office. BeiDou navigation satellite system signal in space interface control document open service signal B1C (version 1.0): BDS-SIS-ICD-B1C-1.0[S]. Beijing: China Satellite Navigation Office, 2017(in Chinese).
- [21] 中国卫星导航系统管理办公室.北斗卫星导航系统空间信号接 口控制文件公开服务信号 B1I(3.0版): BDS-SIS-ICD-B1I-3.0 [S] 北京:中国卫星导航系统管理办公室, 2019.

China Satellite Navigation Office. BeiDou navigation satellite system signal in space interface control document open service signal B1I (version 3.0): BDS-SIS-ICD-B1I-3.0[S]. Beijing: China Satellite Navigation Office, 2019(in Chinese).

- [22] 国家市场监督管理总局,国家标准化管理委员会.北斗/全球卫星导航系统: GB/T 42577—2023[S].北京:中国标准出版社, 2023. State Administration for Market Regulation, Standardization Administration of the People's Republic of China. Definitions and descriptions of BeiDou/global navigation satellite system satellite parameters for high precision applications: GB/T 42577—2023[S]. Beijing: Standards Press of China, 2023(in Chinese).
- [23] 中国卫星导航系统管理办公室.北斗卫星导航系统空间信号接 口控制文件公开服务信号 B2b(1.0版): BDS-SIS-ICD-B2b-1.0

[S] 北京: 中国卫星导航系统管理办公室, 2020.

China Satellite Navigation Office. BeiDou navigation satellite system signal in space interface control document open service signal B2b (version 1.0): BDS-SIS-ICD-B2b-1.0[S]. Beijing: China Satellite Navigation Office, 2020(in Chinese).

[24] Radio Technical Commission for Aeronautics. Environmental conditions and test procedures for airborne equipment: RCTA/DO-160-G-2010[S]. Washington, D.C.: RTCA Program Management Committee, 2010.

Analysis of airworthiness requirements of Beidou-3 airborne equipment only used for tracking

MA Zhenyang^{1, 2}, ZHOU Zhonghua³, ZHANG Fan^{1, 2, *}, WANG Peng^{1, 2}, KE Bingqing³

(1. College of Safety Science and Engineering, Civil Aviation University of China, Tianjin 300300, China;

2. Key Laboratory of Civil Aircraft Airworthiness Technology, Civil Aviation University of China, Tianjin 300300, China;

3. Space Star Technology Co., Ltd., Beijing 100094, China)

Abstract: With the opening of Beidou-3 navigation satellite system service, the relevant airworthiness standard of Beidou-2 cannot meet the requirements of the current application scenarios of airborne equipment, which seriously restricts the application process of Beidou in the field of civil aviation. It is urgent to carry out the research and formulation of the corresponding standard of Beidou-3. The primary goal of Beidou's implementation in civil aircraft is to be "used for aircraft tracking only". In order to achieve this, the distinction between Beidou-2 and Beidou-3 is examined, and a proposed modification plan is presented. In order to provide standard support for equipment development and airworthiness certification, the index revision scheme is put forward according to the equipment configuration, and the approval elements and compliance method suggestions are formed. This is done by taking into account the airworthiness safety and the application scenario of airborne equipment, combined with industrial practice. The key parameters such as equipment antenna gain and satellite visibility are simulated and analyzed to verify the correctness of the revised technical indexes.

Keywords: technical standard order; Beidou-3; aircraft tracking; positioning; short messages; airworthiness

Received: 2022-05-31; Accepted: 2022-06-11; Published Online: 2022-06-22 15:41 URL: link.cnki.net/urlid/11.2625.V.20220621.1954.005

Foundation items: National Natural Science Foundation of China (U2133203); Civil Aviation Safety Capability Project of CAAC (2020-142)

^{*} Corresponding author. E-mail: zhangfan621@foxmail.com

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

DOI: 10.13700/j.bh.1001-5965.2022.0494

基于数值延拓的日月综合借力 DRO 入轨策略

张晨*

(中国科学院 空间应用工程与技术中心 太空应用重点实验室, 北京 100094)

摘 要: 远距离逆行轨道 (DRO) 是地月空间中的一类周期轨道,这类轨道具有长期稳定、 入轨能量低的特点,可作为未来载人月球和载人火星任务的中转站。对于地球至 DRO 的两脉冲入 轨任务,采用日月综合借力 (即同时使用弱稳定边界 (WSB) 和月球借力 (LGA)) 可以最大化入轨质 量,但是这类轨道对初值非常敏感。使用日月综合借力拓展 DRO 入轨脉冲包络,改进构造方法和 提供解析梯度大幅提高多步打靶收敛率,提出 2 层伪弧长延拓方法进一步降低任务总脉冲。数值仿 真采用共振比为 2:1 的 DRO,脉冲最低解采用"LGA+WSB+2LGA"的飞行模式,飞行时间为 123 天,近地轨道发射脉冲为 3.125 km/s, DRO 入轨脉冲仅为 19.7 m/s。

关键词:N体问题;远距离逆行轨道;弱稳定边界;月球借力;数值延拓

中图分类号: V412.41

文献标志码: A 文章编号: 1001-5965(2024)04-1176-11

地月空间蕴藏着巨大的经济和战略价值,如月 球资源开采、广域信息服务、低成本进入深空等, 种类繁多的三体周期轨道将在地月空间开发利用 中发挥重要的作用^[1-2]。远距离逆行轨道(distant retrograde orbits, DRO)是地月系统中一类适合长期 稳定停泊的轨道族,自从上世纪 60 年代 Hénon^[3] 在 三体周期轨道分类时发现该轨道族, DRO 已经被研 究了几十年[48]。随着近年来美国多个地月空间任 务的推动, DRO 重新得到学术和工程界的关注。例 如,小行星重定向任务(asteroid redirect mission, ARM)^[9] 计划从小行星上抓取一块巨石并拖至 DRO, 由于 NASA 工作重点的转移, 2017 年该任务被迫取 消。"阿尔忒弥斯(Artemis)"的首颗卫星将在 DRO 短暂停泊后返回地球^[10]。即将建设的月球轨 道站(Lunar Orbital Platform-Gateway, LOP-G)也可 在必要的时候转移至 DRO。此外还有学者提出将 DRO 轨道作为火星探测的中转站^[11],并从脉冲消耗 等角度分析了这种方案带来的效益。

对于各类采用 DRO 的深空探测任务, 如何降低 DRO 入轨代价成为学者们研究的重要问题,根据借 力模式的不同,地球低轨(low Earth orbit, LEO)至 DRO 的转移轨道可分为4类: 直接转移、月球借力 (Lunar gravity assist, LGA)、太阳借力(本文将弱稳 定边界(weak stability boundary, WSB)转移称为"太 阳借力")和日月综合借力。Weltch等^[12]在圆型限 制性三体问题(circular restricted three-body problem, CR3BP)探讨了直接转移和LGA(含近月脉冲)的 DRO 入轨策略,并分析了 DRO 入轨相位对速度脉 冲的影响。文献 [13] 在 CR3BP 模型下计算了 DRO 与 LEO、近直线晕轨道(near rectilinear halo orbit, NRHO)、L₄和L,附近周期轨道之间的两脉冲转移 轨道,获得了任务时间和速度脉冲的帕累托前沿。 Dawn 等^[10] 面向 "Artemis-1" 猎户座飞船的飞行任 务,设计了采用单次 LGA 的 DRO 入轨和离轨策 略。直接转移和 LGA 的任务时间短(4~20 天)但 是速度脉冲高(200~600 m/s),适合载人任务使

收稿日期:2022-06-16;录用日期:2022-08-19;网络出版时间:2022-09-0115:05 网络出版地址: link.cnki.net/urlid/11.2625.V.20220901.1058.001

*通信作者. E-mail: chenzhang@csu.ac.cn

引用格式: 张晨. 基于数值延拓的日月综合借力 DRO 入轨策略 [J]. 北京航空航天大学学报, 2024, 50 (4): 1176-1186. ZHANG C. Low-energy transfer from Earth into DRO with hybrid gravity assist and numerical continuation [J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50 (4): 1176-1186 (in Chinese).

基金项目:中国科学院空间科学战略性先导科技专项 (XDA30040400); 航天飞行动力学国家级重点实验室基金 (6142210200302); 中国科学院 青促会创新人才项目 (292022000030)

用。对于飞行时间不敏感的无人货运任务,一些学 者提出可采用 WSB 或日月综合借力^[14-20] 降低任务 脉冲。文献 [21-22] 利用 L₁ 和 L₂ 近不变流形设计 了 WSB 转移的 DRO 入轨策略, 流形拼接的主要问 题是难以获得复杂的日月综合借力飞行轨道,因而 发射脉冲和入轨脉冲都较高(地球发射脉冲约3.20 km/s, DRO 入轨脉冲约 89 m/s^[21])。Zhang 等^[23]使用 "直接转录+局部优化" 探索了 DRO 至不同近地轨 道的转移轨道,仿真结果中包含了复杂的日月综合 借力飞行模式,但是求解数量庞大的非线性规划问 题导致计算效率较低,此外彼此孤立的局部收敛解 也不便于理解转移轨道随飞行参数的演化过程。 Scheuerle 等^[24]采用"多步打靶+数值延拓"研究了 地球至 DRO 的转移轨道, 这类方法具有更高的收 敛解利用效率,仿真表明从简单的"WSB"轨道出 发可以延拓得到敏感的"LGA+WSB"转移轨道,共 振比 3:1 的 DRO 入轨脉冲需要约 68 m/s. 但是受限 于问题构造方法及数值延拓策略,该研究并没有充 分挖掘日月综合借力的潜力。

本文针对数值最为敏感,但同时借力效果最佳的日月综合借力 DRO 入轨问题展开研究。通过改进多步打靶问题构造方法和提供解析梯度大幅提高了算法收敛率,同时提出 2 层伪弧长延拓策略进一步降低任务总脉冲,脉冲最低解的近地轨道发射脉冲为 3.125 km/s,共振比 2:1 的 DRO 入轨脉冲仅为 19.7 m/s。本文还将日月综合借力轨道进一步细分成 4 种类型,其中"LGA+WSB+NLGA"转移方式具有最低的任务总脉冲。

1 动力学模型和状态转移矩阵

对于日-地-月四体系统,地球和月球的偏心率 分别为 0.016 7 和 0.054 9, 黄白交角仅为 5°。使用 平面双圆限制性四体问题(planar bicircular restricted four-body problem, PBCR4BP)能够很好的刻画四体 动力学的主要特征。图1为地月旋转系下 PBCR4BP 示意图,定义地球、月球、太阳、卫星分别表示为 P_1 、 P_2 、 P_3 和 P_{sc} ,使用地-月旋转坐标系描述卫星的 运动,其中坐标系原点位于地-月质心, x轴沿地-月 连线方向由地球指向月球, y轴垂直于 x轴。将地-月质量之和、地-月之间平均距离、地球和月球绕其 公共质心旋转的平均角速度作为单位量,定义µ 表示地-月质量参数,则地球和月球质量分别为 $1-\mu 和 \mu$, 位置分别为 $[-\mu, 0]^{T}$ 和 $[1-\mu, 0]^{T}$ 。在此 旋转坐标系中,太阳的质量记作ms,太阳到地-月质 心的平均距离为 ρ_{o} ,太阳绕地-月系质心以角速度 ω_{o} 顺时针旋转,假设在 t_0 =0时刻太阳位于x轴上,则 太阳相位角 $\theta_s = \omega_s t$,太阳的位置为 $[\rho_s \cos \theta_s, \rho_s \sin \theta_s]^T$, PBCR4BP的动力学模型如下:

$$\begin{bmatrix} \ddot{x} \\ \ddot{y} \end{bmatrix} = \begin{bmatrix} x + 2\dot{y} - \frac{(1-\mu)(x+\mu)}{r_1^3} - \frac{\mu(x-1+\mu)}{r_2^3} + a_{s,x} \\ y - 2\dot{x} - \frac{(1-\mu)y}{r_1^3} - \frac{\mu y}{r_2^3} + a_{s,y} \end{bmatrix}$$
(1)

式中: [a_{s,x},a_{s,y}]为太阳引力加速度,表示为

$$\begin{cases} a_{s,x} = -\frac{m_s \left(x - \rho_s \cos \theta_s\right)}{r_3^3} - \frac{m_s \cos \theta_s}{\rho_s^2} \\ a_{s,y} = -\frac{m_s \left(y - \rho_s \sin \theta_s\right)}{r_3^3} - \frac{m_s \sin \theta_s}{\rho_s^2} \end{cases}$$
(2)

其中, r₁、r₂和r₃分别表示卫星到地球、月球和太阳 的距离, 可表示为

$$r_1 = \left[(x+\mu)^2 + y^2 \right]^{1/2} \tag{3}$$

$$r_2 = \left[(x + \mu - 1)^2 + y^2 \right]^{1/2} \tag{4}$$

$$r_3 = \left[\left(x - \rho_s \cos \theta_s \right)^2 + \left(y - \rho_s \sin \theta_s \right)^2 \right]^{1/2}$$
 (5)



图 1 平面双圆限制性四体模型

Fig. 1 Planar bicircular restricted four-body model

在轨道设计和修正时,经常用到状态转移矩阵 这一概念。定义 $\varphi(x_0;t)$ 为以 x_0 为初值,从时间 t_0 运 动到t所得到的解,系统的状态转移矩阵表示为

$$\boldsymbol{\Phi}(t_0,t) = \frac{\partial \boldsymbol{\varphi}(\boldsymbol{x}_0;t)}{\partial \boldsymbol{x}_0} \tag{6}$$

状态转移矩阵在一阶程度上刻画了初始状态 改变量对终端状态改变量所产生的影响,其动力学 方程表示为

$$\dot{\boldsymbol{\Phi}}(t_0,t) = \frac{\partial \boldsymbol{f}(\boldsymbol{x})}{\partial \boldsymbol{x}} \boldsymbol{\Phi}(t_0,t) = \boldsymbol{A} \boldsymbol{\Phi}(t_0,t)$$
(7)

$$\boldsymbol{\Phi}(t_0, t_0) = \frac{\partial \boldsymbol{\varphi}(\boldsymbol{x}_0; t_0)}{\partial \boldsymbol{x}_0} = \boldsymbol{I}$$
(8)

式(7)中矩阵A的推导参见附录A。y为状态变 量x和状态转移矩阵Φ构成的列向量,其动力学方 程表示为

$$\dot{\mathbf{y}} = \mathbf{F}(\mathbf{y}) \rightarrow \begin{bmatrix} \dot{\mathbf{x}} \\ \operatorname{vec}\left(\dot{\boldsymbol{\Phi}}\right) \end{bmatrix} = \begin{bmatrix} f(\mathbf{x}) \\ \operatorname{vec}\left(A\boldsymbol{\Phi}\right) \end{bmatrix}$$
(9)

其中,"vec"为将矩阵变为列向量的算子。以 $x(t_0) = x_0 和 \boldsymbol{\Phi}(t_0, t_0) = I_{4\times 4}$ 作为积分初值,对式(9) 在时间区间 $[t_0, t]$ 进行数值积分即可得到 $\boldsymbol{\Phi}(t_0, t)$ 。

2 DRO 和月球轨道站

2.1 DRO 族

DRO 是圆型限制性三体问题中一类特殊的周期轨道族,在地月旋转系下围绕月球呈顺时针(逆行)运动。图 2 展示了通过微分修正与数值延拓得到的 DRO 族,LU 为归一化距离单位(参见附录表 C1),不同的颜色用于区分轨道周期。共振比指月球公转周期与卫星轨道周期之比,具有典型共振比的周期轨道可保证卫星与地球、月球具有周期性的几何关系,本文在数值仿真中用到了共振比 2:1的 DRO(轨道周期约为14 天)。





对于给定周期的 DRO, 定义 \mathbf{x}_{dro}^{0} 为 DRO 沿 y轴 正向穿过地月连线的状态, P_{dro} 为 DRO 周期, σ 为 相位因子,则 DRO 上任意一点的状态 $\mathbf{x}_{dro}(\sigma)$ 表 示为

$$\begin{cases} \boldsymbol{x}_{dro}(\sigma) = \boldsymbol{\varphi} \left(\boldsymbol{x}_{dro}^{0}, t_{0} = 0; t \right) \\ t = t_{0} + \sigma P_{dro} \quad \sigma \in [0, 1) \end{cases}$$
(10)

式中: *φ*表示以*x*⁰_{dro}为初始状态, 从*t*₀到*t*积分动力学 方程所得到的解。

2.2 日月综合借力转移轨道

对于 LEO 至各类 月 心轨 道 (如 近 月 轨 道、 DRO、NRHO 等)的转移轨道, 采用 WSB 转移方式 可以有效降低入轨脉冲。图 3 展示了 WSB 转移方 式示意图, 卫星初始位于近地圆轨道, 施加出发脉 冲后首先抵达距地球大约 3~5 倍地月距的 WSB。 在这里卫星持续受到太阳引力的影响并不断改变 轨道能量, 图中黑色箭头代表引力加速度方向, 如 果卫星的远地点恰好位于日地旋转系的 2/4 象限, 则太阳的影响等效为一个脉冲,将后半段轨道的近 地点从 LEO 无动力/自由抬升至月球轨道附近,进 而降低了各类月心轨道的入轨脉冲。WSB 方式的 缺点在于需要施加较大的发射脉冲(200 km 近地圆 轨道需要 3.19~3.20 km/s 发射脉冲)才能够抵达 WSB 区域。



为最大化卫星入轨质量,可以从 2 个方面改进 WSB 转移方式,首先,可以在地球发射段添加 LGA 以形成"LGA+WSB"转移轨道,月球的作用等 效为一个脉冲,将卫星远地点从 400 000 km 无动力 /自由抬升至 WSB。其次,可以在卫星返回地月空 间后再添加若干次 LGA(简写为 NLGA),进一步降 低入轨脉冲。综上,"LGA+WSB+NLGA"方式综合 利用了日月引力的影响,能够最大化 DRO 入轨质 量(见图 4),但由于 WSB 和 NLGA 相互耦合,这类 转移轨道对初值非常敏感。



2.3 DRO 月球轨道站

图 5 展示了 DRO 月球轨道站支持的载人月球 和火星任务示意图。其中 LEO、月球低轨(low Lunar orbit, LLO)和火星低轨(low Mars orbit, LMO) 都位于引力势阱的谷底,从谷底爬出来就进入了一 望无垠的"引力平原",在 DRO 部署月球轨道站有 3 方面优势:首先,DRO 具有较高的轨道能级且接 近"引力平原",除了能支持载人月球任务还能够 方便进入深空;其次,DRO 稳定但稳定裕度有限, 维持成本和入轨/离轨成本都相对较低;最后,DRO 位于地月之间,对地月空间其他轨道都具有较好的 可达性。



图 5 月球轨道站支持的载人月球和火星任务示意图

Fig. 5 Schematic diagram of cislunar orbital station supporting manned missions to the Moon and Mars

3 多步打靶修正和数值延拓

3.1 多步打靶修正

多体系统复杂的动力学环境使得转移轨道对 初值非常敏感,多步打靶技术可以扩大收敛域并提 高算法的鲁棒性。图6展示了多步打靶示意图。 整条轨道被N个离散点分割为等间距的N-1条轨 道段,多步打靶变量表示为

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{z}_1, \boldsymbol{z}_2, \cdots, \boldsymbol{z}_N, \boldsymbol{\tau}_1, \boldsymbol{\tau}_N \end{bmatrix}^{\mathrm{T}}$$
(11)

式中: z为离散点的打靶变量; τ_1 为任务起始时刻; τ_N 为任务结束时刻。





定义x为离散点的状态变量,对于通常情况,所 有离散点的打靶变量都等于状态变量:

$$\mathbf{z}_n = \mathbf{x}_n^0 = \left[x_n^0, y_n^0, \dot{x}_n^0, \dot{y}_n^0 \right]^{\mathrm{T}}$$
 $(n = 1, 2, \cdots, N)$ (12)

为提高算法的收敛性,设置第1个离散点的打 靶变量为

$$\boldsymbol{z}_1 = \boldsymbol{e}_1 = \begin{bmatrix} \boldsymbol{R}, \alpha, \beta \end{bmatrix}^{\mathrm{T}}$$
(13)

式中: **R**为近地点地心距; α为相位角; β为速度脉冲 比。假设卫星沿切向从地球停泊轨道出发,则第 1个离散点的状态变量表示为

$$\mathbf{x}_{1}^{0}(\mathbf{e}_{1}) = \begin{bmatrix} R\cos\alpha - \mu \\ R\sin\alpha \\ (-\beta\sqrt{(1-\mu)/R} + R)\sin\alpha \\ (\beta\sqrt{(1-\mu)/R} - R)\cos\alpha \end{bmatrix}$$
(14)

第2,3,…,N-1个离散点的打靶变量表示为

$$z_n = x_n^0$$
 (n = 2, 3, ..., N - 1) (15)

第N个离散点的打靶变量表示为

$$\boldsymbol{z}_N = \boldsymbol{e}_N = [\boldsymbol{\sigma}, \Delta \boldsymbol{V}, \boldsymbol{\omega}]^{\mathrm{T}}$$
(16)

式中: σ 为入轨时刻的 DRO 相位; ΔV 为入轨脉冲; ω 为脉冲相位角(定义 v_{dro} 为入轨时刻 DRO 轨道的 速度矢量, ω 为 v_{dro} 沿顺时针至 ΔV 转过的角度)。假 设入轨时刻 DRO 的状态表示为 $x_{dro}(\sigma)$,则第N个离 散点的状态变量为

$$\boldsymbol{x}_{N}^{0}\left(\boldsymbol{e}_{N}\right) = \begin{bmatrix} x_{dro} \\ y_{dro} \\ \dot{x}_{dro} - \Delta V \frac{\dot{x}_{dro} \cos \omega - \dot{y}_{dro} \sin \omega}{\sqrt{\dot{x}_{dro}^{2} + \dot{y}_{dro}^{2}}} \\ \dot{y}_{dro} - \Delta V \frac{\dot{x}_{dro} \sin \omega + \dot{y}_{dro} \cos \omega}{\sqrt{\dot{x}_{dro}^{2} + \dot{y}_{dro}^{2}}} \end{bmatrix}$$
(17)

各段轨道的状态连接约束略有不同,其中第 1段轨道的状态连接约束为

 $\zeta_{1} = \mathbf{x}_{1}^{f} \left(\mathbf{x}_{1}^{0} \left(\mathbf{e}_{1} \right), \tau_{1}, \tau_{2} \right) - \mathbf{x}_{2}^{0} = 0$ (18)

第2,3,…,N-2段轨道的状态连接约束为

$$ζN-1 = xfN-1(x0N-1, τN-1, τN) - x0N(eN) = 0$$
(20)
筆 1 个 **离** 散 占 的 功 界 幼 **~~σ~~** 为

$$\Psi_1 = R - r_{\rm leo}^* = 0 \tag{21}$$

式中: *R* 为卫星地心距, *r*^{*}_{leo}为 LEO 轨道地心距。 综上, 多步打靶约束 *F*(*X*)表示为

$$\boldsymbol{F}(\boldsymbol{X}) = \left[\boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2, \cdots, \boldsymbol{\zeta}_{N-1}, \boldsymbol{\psi}_1\right]^{\mathrm{T}} = \boldsymbol{0}$$
 (22)

定义*DF*(X)为多步打靶约束*F*(X)的雅可比矩 阵,其详细推导参见附录 B。

多步打靶问题共包含4N个打靶变量和4N-3 个等式约束。当得到雅克比矩阵DF(X),使用最小 二乘更新方程对打靶变量进行修正,收敛后即可得 到转移轨道,更新方程表示为

$$\boldsymbol{X}_{k+1} = \boldsymbol{X}_{k} - \boldsymbol{D}\boldsymbol{F}\left(\boldsymbol{X}_{k}\right)^{\mathrm{T}} \left(\boldsymbol{D}\boldsymbol{F}\left(\boldsymbol{X}_{k}\right) \cdot \boldsymbol{D}\boldsymbol{F}\left(\boldsymbol{X}_{k}\right)^{\mathrm{T}}\right)^{-1} \boldsymbol{F}\left(\boldsymbol{X}_{k}\right)$$
(23)

多步打靶的仿真参数(参见附录C)。多步打靶 收敛后, LEO 发射脉冲表示为

$$\Delta v_{\rm dep} = (|\beta| - 1) \sqrt{\frac{1 - \mu}{R}}$$
 (24)

DRO 入轨脉冲表示为

 $\Delta v_{\rm arr} = \Delta V \tag{25}$

任务总脉冲_{Δvall}表示为

$$\Delta v_{\text{all}} = (|\beta| - 1) \sqrt{\frac{1 - \mu}{R}} + \Delta V \qquad (26)$$

需要说明的是, 三维 DRO 入轨问题的推导与 之类似, 在此不再赘述。

3.2 伪弧长延拓

最小二乘迭代得到的是在解空间内相互孤立 的离散收敛解,为了提高收敛解的利用效率并进一 步降低任务总脉冲,可以在离散解附近进行局部 延拓。

最简单的数值延拓方法是参数增量法(见 图 7),通过引入一个具有明确物理含义的辅助参数 λ (如设置为 DRO 入轨时间 $\lambda = \tau_N$),以 $\Delta\lambda$ 为步长逐 步增加辅助参数 λ ,在每一个增量步内,采用最小二 乘更新方程进行打靶,其初值取为前一增量步的 解,只要 $\Delta\lambda$ 足够小,就能满足局部收敛条件。参数 增量法的缺点是 λ 可能非单调,如果雅可比矩阵在 λ 转向点附近是奇异或者病态的,参数增量法将会 失效。



Fig. 7 Parametric incremental continuation

另一种数值延拓方法为伪弧长延拓法,如图 8 所示。该方法要求打靶变量比打靶约束多1维,计



Fig. 8 Pseudo-arc continuation

算步骤分成预测和校正2步。假设第k步打靶变量的初值为X_k,打靶变量下一步的预测值为

$$\boldsymbol{X}_{k+1} = \boldsymbol{X}_k + \Delta \boldsymbol{s} \boldsymbol{d} \tag{27}$$

式中: Δ*s*为解曲线的延拓弧长,通常设置为常数。 *d*_k为雅可比矩阵的零空间向量,表示为

$$\boldsymbol{d}_{k} = \pm \operatorname{Null}\left(\boldsymbol{DF}\left(\boldsymbol{X}_{k}\right)\right) \tag{28}$$

在引入延拓参数Δs的同时,也增加了一个打靶 约束,使得打靶变量和约束数相等,校正步表示为

$$\bar{F}(X_{k+1}) = \begin{bmatrix} F(X_{k+1}) \\ (X_{k+1} - X_k)^{\mathrm{T}} d_k - \Delta s \end{bmatrix} = \mathbf{0}$$
 (29)

式(29)存在唯一解,迭代收敛后打靶变量将 沿着解曲线滑动弧长Δs。校正步的雅可比矩阵 表示为

$$G\bar{F}(X_{k+1}) = \begin{bmatrix} GF(X_{k+1}) \\ d_k^{\mathrm{T}} \end{bmatrix} = \mathbf{0}$$
 (30)

需要说明的是, 伪弧长延拓并不直接对辅助参数λ进行延拓, 通过引入延拓弧长Δs, 有效避免了转向点附近雅可比矩阵奇异的问题, 因而比参数增量 法的稳定性和收敛性更强。

对于平面 DRO 入轨问题, 多步打靶问题包含 4N打靶变量和4N-3等式约束, 还需要增加 2 个等 式约束才能使用伪弧长延拓, 本文仅对以下 2 种情 况进行讨论:

1) 延拓 DRO 人轨脉冲。固定 LEO 发射时间 $\tau_1 = \tau_1^*$ 和速度脉冲比 $\beta = \beta^*$,发射相位自由,通过伪 弧长延拓最小化 DRO 入轨脉冲 Δv_{arr} 。 第4期

$$\bar{F}_{1}\left(\boldsymbol{X}_{k+1}\right) = \begin{bmatrix} \boldsymbol{F}\left(\boldsymbol{X}_{k+1}\right) \\ \left(\boldsymbol{X}_{k+1} - \boldsymbol{X}_{k}\right)^{\mathrm{T}} \boldsymbol{d}_{k} - \Delta \boldsymbol{s} \\ \boldsymbol{\tau}_{1} - \boldsymbol{\tau}_{1}^{*} \\ \boldsymbol{\beta} - \boldsymbol{\beta}^{*} \end{bmatrix} = \boldsymbol{0} \quad (31)$$

此时打靶变量数和打靶约束数相等,更新方程 表示为

$$\boldsymbol{X}_{k+1} = \boldsymbol{X}_k - \boldsymbol{D}\boldsymbol{\bar{F}}_1 \left(\boldsymbol{X}_k \right)^{-1} \boldsymbol{\bar{F}}_1 \left(\boldsymbol{X}_k \right)$$
(32)

2)延拓 LEO 发射脉冲。固定 DRO 入轨时间 $\tau_N = \tau_N^*$ 和入轨脉冲 $\Delta V = \Delta V^*$, DRO 入轨相位自由, 通过伪弧长延拓最小化 LEO 发射脉冲 Δv_{dep} 。

$$\bar{F}_{2}(X_{k+1}) = \begin{bmatrix} F(X_{k+1}) \\ (X_{k+1} - X_{k})^{\mathrm{T}} d_{k} - \Delta s \\ \tau_{N} - \tau_{N}^{*} \\ \Delta V - \Delta V^{*} \end{bmatrix} = \mathbf{0} \quad (33)$$

此时打靶变量数和打靶约束数相等,更新方程 表示为

$$\boldsymbol{X}_{k+1} = \boldsymbol{X}_{k} - \boldsymbol{D}\boldsymbol{\bar{F}}_{2}\left(\boldsymbol{X}_{k}\right)^{-1}\boldsymbol{\bar{F}}_{2}\left(\boldsymbol{X}_{k}\right)$$
(34)

4 数值仿真

本节首先通过多步打靶得到离散的局部收敛 解,再使用伪弧长延拓进一步降低任务总脉冲。多 步打靶的计算流程如下:

1) 随机生成太阳相位角 $\theta_{s}(\tau_{N})$ 和 DRO 入轨参数 $[\sigma, \Delta V, \omega]^{T}$,取值范围参见附录 C,获得 DRO 入轨时间 τ_{N} 和入轨状态 \mathbf{x}_{N}^{0} (见式 (17))。

2) 从*x*⁰⁰出发逆向积分动力学模型, 使得卫星先 抵达 WSB 再抵达近地点, 为了提高计算效率仅保 留近地点小于 50 000 km 的解。

3) 使用多步打靶修正转移轨道(见式 (22) 和 式 (23))。

根据以上流程,随机构造10⁶个 DRO 入轨状态,从 DRO 逆向积分先抵达 WSB 再抵达近地点的有3658条轨道(仅占初值的0.36%),计算时间为226 s。接着使用多步打靶修正转移轨道,表1展示了使用解析梯度对收敛率的改进效果,与有限差分相比,在计算时间类似的情况下,解析梯度将多步打靶收敛率提高了约7.6倍。上述仿真采用 MATLAB编程语言和并行计算,计算平台为64 核高性能工作站,多步打靶最大迭代数为20,积分器为龙格库塔7/8 阶变步长方法,编译为 MEX 提高积分速度,积分精度为1⁻¹⁰。在数值积分过程中,如果卫星地心距小于地球半径,或者月心距小于月球半径,则终止计算。

表1 有限差分和解析梯度的对比

 Table 1
 Comparison between finite difference and

名称	收敛数	收敛率/%	计算时间/s
有限差分	309	8.44	18.01
解析梯度	2 361	64.54	20.24

图 9 展示了多步打靶收敛解的局部解空间,横 坐标为 LEO 发射脉冲,纵坐标为 DRO 入轨脉冲, 根据借力方式的不同,图 9 可分为 4 种转移方式, 红色圆圈为"WSB",蓝色方块为"LGA+WSB",绿 色圆点为"WSB+NLGA",黑色十字为"LGA+WSB+ NLGA",其中 LGA 的判定标准为近月点位于月球 影响球内,WSB 的判定标准为远地点位于 3~5 倍 地月距。通过分析可得到以下结论:

1) 在 "WSB" 之前引入单次 LGA 能够将发射 脉冲从 3.185 km/s 降低至 3.125 km/s,节省约 60 m/s。

 在"WSB"之后引入 NLGA 能够将 DRO 入 轨脉冲从 40 m/s 降低至 20 m/s, 节省约 20 m/s。

 3)与"WSB"方案相比,"LGA+WSB+NLGA" 方案能够节省约80m/s的任务总脉冲。



对于复杂的"LGA+WSB+NLGA"转移方式,敏 感的 WSB 和 NLGA 相互耦合,使得图 9 左下区域 (低总脉冲区域)的收敛解非常稀疏,仅增加初值也 很难大幅提升这一区域的收敛解数量。针对这一 问题,本文提出双层伪弧长法进一步探索收敛解附 近的局部解空间,计算流程如下:

1)固定 DRO 入轨脉冲,分别沿正负 2 个方向 延拓 LEO 发射脉冲(见式 (31)),保存 LEO 发射脉 冲最低的解作为下一层延拓的初值。

2)固定 LEO 发射脉冲,分别沿正负 2 个方向 延拓 DRO 入轨脉冲(见式 (33)),保存 DRO 入轨脉 冲最低的解。

需要说明的是,由于事先无法知道正确的延拓 方向,因而每层延拓都沿正负2个方向分别进行 200次。整个延拓过程耗时 416 s, 图 10 展示了延 拓后各区间解数量的变化,可以发现在数值延拓 后,低发射脉冲和低入轨脉冲的收敛解在数量上都 获得了大幅提升。图 11 展示了数值延拓后的解空 间,其中横轴为飞行时间,纵轴分别为发射脉冲、入 轨脉冲和任务总脉冲,黑色菱形为帕累托前沿,红 色五角星为总脉冲最低的解。延拓后大量收敛解 的 LEO 发射脉冲都位于 3.125 km/s 附近(图 9 则基 本没有),此外当飞行时间为100天左后时,就已经 出现入轨脉冲仅为20m/s左右的解。

延拓前 600 延拓后 轨道数 400 200 3.14 3.12 3.16 3.18 3.20 3.22 LEO发射脉冲/(km·s⁻¹) (a) 数值延拓后各LEO发射脉冲区间轨道数的变化 300 延拓前 延拓后 轨道数 200 100 0 0.04 0.02 0.06 0.08 0.10 DRO入轨脉冲/(km·s⁻¹) (b) 数值延拓后各DRO入轨脉冲区间轨道数的变化

下面进一步展示图 11 中总脉冲最低解(红色

图 10 数值延拓后各区间轨道数的变化





Fig. 11 Solution space after numerical continuation

五角星)的延拓过程。图 12 展示了第1 层延拓的 飞行轨道,其中灰色细实线为延拓轨道,蓝色粗实 线为 LEO 发射脉冲最低的解。图 13 展示了第 2 层 延拓的飞行轨道,红色粗实线为DRO 入轨脉冲最 低的解。图 14 展示了伪弧长延拓的相空间变化轨 迹,其中x轴为LEO发射脉冲,y轴为DRO入轨脉 冲,z轴为飞行总时间,图中"蓝色圆圈"为第1层 延拓初值,"红色方块"为第1层延拓最优解(也是 第2层延拓初值),"红色五角星"为第2层延拓最 优解。可以发现第1层延拓中转移轨道从蓝色圆 圈滑动到红色方块(LEO发射脉冲逐渐降低),这个 过程与图 12 中地月转移段逐渐隆起(近月点逐渐 降低)的过程相对应。在第2层延拓中转移轨道从 红色方块出发分别沿2个方向探索,这个过程与图13 中入轨点在 DRO 轨道上滑动的过程相对应。最终 转移轨道采用了"LGA+WSB+2LGA"的飞行方式。 表2对比了延拓前后转移轨道的参数,可以发



图 12 第1层延拓和最优解

Fig. 12 The first level continuation and optimal solution







图 14 伪弧长延拓的相空间变化轨迹 Fig. 14 Phase space path with pseudo-arc continuation

现通过数值延拓 LEO 的发射脉冲降低了 53.2 m/s, DRO 人轨脉冲降低了 21.2 m/s,飞行时间仅增加了 3.35 天,其中 LU、VU 和 TU 分别为归一化的距离、 速度和时间单位,参见附录表 C1。

表 2 数值延拓前后轨道对比 Table 2 Comparison of orbits before and after numerical

aantinuation

	continuation	
任务名称	延拓前	延拓后
LEO发射时间/TU	-31.944 945 682 3	-32.992 737 946 4
DRO入轨时间/TU	-4.422 676 457 9	-4.699 448 716 6
LEO华针位墨/III	[-0.025 926 418 9,	[-0.020 484 378 0,
LEU及别位且/LU	-0.010 152 072 7]	-0.014 946 105 5]
LEO发射速度/VU	[6.340 104 171 1,	[9.288 605 287 2,
	-8.603 190 754 6]	-5.179 229 375 6]
DRO相位因子 σ	0.071 769 660 2	0.976 618 137 7
LEO发射脉冲/(km·s ⁻¹)	3.178 4	3.125 2
DRO入轨脉冲/(km·s ⁻¹)	0.040 9	0.019 7
任务总脉冲/(km·s ⁻¹)	3.219 4	3.144 9
任务总时间/d	119.669 9	123.022 4

最后对比 LEO 至 GEO、LLO 和 DRO 的入轨质 量,从入轨质量的角度解释 DRO 更适合作为未来 远深空探测前哨站的原因。假设采用长征 5 号运 载火箭, LEO 入轨质量 2.5×10⁴ kg,火箭和卫星的比 冲都为 300 s,表 3 对比了 3 类任务的速度脉冲和剩 余质量。

首先,对比 GEO 和 DRO 的人轨质量,假设采用 霍曼变轨,LEO 至 GEO 的入轨质量仅为 6.571×10³ kg, 尽管 DRO 远离地球,DRO 却可以达到 GEO 入轨质 量的 1.30 倍。

表 3 3 种任务场景对比

Table 3 Comparison between 3 mission scenarios

任务名称	发射脉冲/(km·s ⁻¹)	入轨脉冲/(km·s ⁻¹)	入轨质量/10 ³ kg
LEO至GEO	2.454	1.477	6.571
LEO至LLO (综合借力)	3.125	0.650	6.929
LEO至DRO (综合借力)	3.125	0.019	8.584

其次,对比 LLO 和 DRO 的人轨质量,假设都采 用日月综合借力,由于 LLO 位于月球引力势阱的 谷底,LLO 人轨脉冲仅能从直接转移的约 800 m/s 降低至约 650 m/s^[16]。DRO 入轨脉冲则能从直接 转移的约 600 m/s^[12]降低至约 20 m/s(仅为 LLO 入 轨脉冲的 2.5% 至 3%),在飞行时间都为 120 天左 右的条件下,DRO 入轨质量可以达到 LLO 的约 1.23 倍。

5 结 论

1) 针对 LEO 至 DRO 的低能转移问题, 通过推导解析梯度, 与有限差分法相比提高了 7.6 倍的收敛率, 上述改进同样适用于 LEO 至其他地月空间周期轨道的低能转移问题。

2)针对敏感的日月综合借力转移方式,提出双 层伪弧长延拓方法,有效提高了多步打靶收敛解的 利用效率,同时进一步降低了任务总脉冲。

3) "LGA+WSB+NLGA"方案仅需 3.125 km/s 的 LEO 发射脉冲和约 20 m/s 的 DRO 入轨脉冲,使 得 DRO 最大入轨质量可以达到 GEO 的 1.30 倍,或 是 LLO 的 1.23 倍。

参考文献(References)

- WHITLEY R, MARTINEZ R. Options for staging orbits in cislunar space[C] //Proceedings of the IEEE Aerospace Conference.
 Piscataway: IEEE Press, 2016: 1-9.
- [2] 曾豪, 李朝玉, 彭坤, 等. 地月空间 NRHO 与 DRO 在月球探测中 的应用研究[J]. 宇航学报, 2020, 41(7): 910-919.
 ZENG H, LI Z Y, PENG K, et al. Research on application of Earth-Moon NRHO and DRO for lunar exploration[J]. Journal of Astronautics, 2020, 41(7): 910-919 (in Chinese).
- [3] HÉNON M. Numerical exploration of the restricted problem[J]. Astronomy and Astrophysics, 1969(1): 223-238.
- [4] DEMEYER J, GURFIL P. Transfer to distant retrograde orbits using manifold theory[J]. Journal of Guidance, Control, and Dynamics, 2007, 30(5): 1261-1267.
- [5] MINGOTTI G, TOPPUTO F, BERNELLI-ZAZZERA F. Transfers to distant periodic orbits around the Moon via their invariant manifolds[J]. Acta Astronautica, 2012(79): 20-32.
- [6] BEZROUK C, PARKER J S. Long term evolution of distant retrograde orbits in the Earth-Moon system[J]. Astrophysics and Space Science, 2017, 362(9): 176.
- [7] LARA M. Design of distant retrograde orbits based on a higher order analytical solutiont[J]. Acta Astronautica, 2019(161): 562-578.
- [8] 吴小婧, 曾凌川, 巩应奎. DRO 计算及其在地月系中的摄动力研究[J]. 北京航空航天大学学报, 2020, 46(5): 883-892.
 WU X J, ZENG L C, GONG Y K. DRO computation and its per-

WU X J, ZENG L C, GONG Y K. DRO computation and its perturbative force in the Earth-Moon system[J]. Journal of Beijing University of Aeronautics and Astronautics, 2020, 46(5): 883-892 (in Chinese).

- [9] STRANGE N, LANDAU D, MCELRATH T, et al. Overview of mission design for NASA asteroid redirect robotic mission concept[C]//Proceedings of the 33rd International Electric Propulsion Conference. Washington, D. C. : NASA, 2013, 1-13.
- [10] DAWN T F, GUTKOWSKI J, BATCHA A, et al. Trajectory design considerations for exploration mission 1[C] // Proceedings of the Space Flight Mechanics Meeting. Reston: AIAA, 2018: 968.
- [11] CONTE D, DI CARLO M, HO K, et al. Earth-Mars transfers through Moon Distant Retrograde Orbits[J]. Acta Astronautica, 2018(143): 372-379.
- [12] WELCH C M, PARKER J S, BUXTON C. Mission considerations for transfers to a distant retrograde orbit[J]. The Journal of the Astronautical Sciences, 2015, 62(2): 101-124.
- [13] CAPDEVILA L R, HOWELL K C. A transfer network linking Earth, Moon, and the triangular libration point regions in the Earth-Moon system[J]. Advances in Space Research, 2018, 62(7): 1826-1852.
- [14] BELBRUNO E A, MILLER J K. Sun-perturbed Earth-to-moon transfers with ballistic capture[J]. Journal of Guidance Control Dynamics, 1993, 16(4): 770-775.
- [15] YAGASAKI K. Sun-perturbed Earth-to-Moon transfers with low energy and moderate flight time[J]. Celestial Mechanics and Dynamical Astronomy, 2004, 90(3): 197-212.
- [16] TOPPUTO F. On optimal two-impulse Earth-Moon transfers in a four-body model[J]. Celestial Mechanics and Dynamical Astronomy, 2013, 117(3): 279-313.
- [17] OSHIMA K, TOPPUTO F, YANAO T. Low-energy transfers to the Moon with long transfer time[J]. Celestial Mechanics and Dynamical Astronomy, 2019, 131(1): 4.
- [18] PARKER J S, ERSON R L. Low-energy lunar trajectory design[M]. Pasadena: John Wiley & Sons, 2014: 1-381.
- [19] ZHANG Z, HOU X. Transfer orbits to the Earth–Moon triangular libration points[J]. Advances in Space Research, 2015, 55(12): 2899-2913.
- [20] PARRISH N L, KAYSER E, UDUPA S, et al. Survey of ballistic lunar transfers to near rectilinear halo orbit[C]//Proceedings of the AAS/AIAA Astrodynamics Specialist Conference. Reston:AIAA, 2019: 1-20.
- [21] XU M, XU S J. Exploration of distant retrograde orbits around Moon[J]. Acta Astronautica, 2009, 65(5-6): 853-860.
- [22] TAN M H, ZHANG K, LV M B, et al. Transfer to long term distant retrograde orbits around the Moon[J]. Acta Astronautica, 2014(98): 50-63.
- [23] ZHANG R K, WANG Y, ZHANG C, et al. The transfers from lunar DROs to Earth orbits via optimization in the four body problem[J]. Astrophysics and Space Science, 2021, 366(6): 49.
- [24] SCHEUERLE S T, MCCARTHY B P, HOWELL K C. Construction of ballistic lunar transfers leveraging dynamical systems techniques[C]//Proceedings of the AAS/AIAA Astrodynamics Specialist Conference. Reston: AIAA, 2020: 1-20.

附录A: 动力学模型的雅可比矩阵

式 (7) 中A表示为

$$\boldsymbol{A} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ A_{31} & A_{32} & 0 & 2 \\ A_{41} & A_{42} & -2 & 0 \end{bmatrix}$$
(A1)

$$\begin{cases} A_{31} = 1 - \frac{1 - \mu}{r_1^3} - \frac{\mu}{r_2^3} - \frac{m_s}{r_3^3} + \frac{3(1 - \mu)(x + \mu)^2}{r_1^5} + \\ \frac{3\mu(x + \mu - 1)^2}{r_2^5} + \frac{3m_s\left(x - \rho_s \cos \theta_s\right)^2}{r_3^5} \\ A_{41} = A_{32} = \frac{3(1 - \mu)(x + \mu)y}{r_1^5} + \frac{3\mu(x + \mu - 1)y}{r_2^5} + \\ \frac{3m_s\left(x - \rho_s \cos \theta_s\right)\left(y - \rho_s \sin \theta_s\right)}{r_3^5} \\ A_{42} = 1 - \frac{1 - \mu}{r_1^3} - \frac{\mu}{r_2^3} - \frac{m_s}{r_3^3} + \\ \frac{3(1 - \mu)y^2}{r_1^5} + \frac{3\mu y^2}{r_2^5} + \frac{3m_s\left(y - \rho_s \sin \theta_s\right)^2}{r_3^5} \end{cases}$$
(A2)

附录 B: 打靶约束的雅可比矩阵

图 B1 展示了打靶约束的雅可比矩阵DF(X) 其中, **Φ**_n为状态转移矩阵, 此外:

$$\frac{\partial \zeta_1}{\partial \boldsymbol{e}_1} = \frac{\partial \boldsymbol{x}_1^I}{\partial \boldsymbol{x}_1^0} \cdot \frac{\partial \boldsymbol{x}_1^0}{\partial \boldsymbol{e}_1} = \boldsymbol{\Phi}_1 \frac{\partial \boldsymbol{x}_1^0}{\partial \boldsymbol{e}_1}$$
(B1)

$$\frac{\partial \boldsymbol{\zeta}_{N-1}}{\partial \boldsymbol{e}_{N}} = \frac{\partial \boldsymbol{\zeta}_{N-1}}{\partial \boldsymbol{x}_{N}^{0}} \cdot \frac{\partial \boldsymbol{x}_{N}^{0}}{\partial \boldsymbol{e}_{N}} = -\left[\frac{\partial \boldsymbol{x}_{N}^{0}}{\partial \boldsymbol{x}_{dro}} \dot{\boldsymbol{x}}_{dro} \boldsymbol{P}_{dro}, \frac{\partial \boldsymbol{x}_{N}^{0}}{\partial \Delta V}, \frac{\partial \boldsymbol{x}_{N}^{0}}{\partial \varphi}\right]$$
(B2)

$$\frac{\partial \Psi_1}{\partial e_1} = [1,0,0] \tag{B3}$$

进一步推导 $\partial \zeta_n / \partial \tau_1 \pi \partial \zeta_n / \partial \tau_N$, 第*n*段轨道内, 单轨道变分示意图如图 B2 所示, $x_n^f (x_n^0, \tau_n, \tau_{n+1})$ 的 变分表示为

$$\delta \mathbf{x}_{n}^{f} = \boldsymbol{\Phi}_{n} \delta \mathbf{x}_{n}^{0} - \boldsymbol{\Phi}_{n} \dot{\mathbf{x}}_{n}^{0} \delta \tau_{n} + \dot{\mathbf{x}}_{n}^{f} \delta \tau_{n+1}$$
(B4)
通过式 (B4) 可得

离散点打靶变量 初始/终端时刻 x_{2}^{0} x_{3}^{0} x_{N-1}^0 e_1 e_N τ_1 τ_N $\partial \zeta_1$ $\partial \zeta_1$ $\partial \zeta_1$ -*I* ζ_1 ∂e_1 $\partial \tau_1$ $\overline{\partial} \tau_N$ $\partial \zeta_2$ $\partial \zeta_2$ ζ_2 Φ, -1 状态连接约束 $\partial \tau_1$ $\partial \tau_N$ • : • $\partial \zeta_{N-1}$ $\partial \zeta_{N-}$ $\partial \zeta_N$ $\boldsymbol{\Phi}_{N-1}$ ζ_{N-1} $\partial \tau$ $\partial \tau_1$ $\partial \tau_{\Lambda}$ 边界约束 $\partial \Psi$ Ψ_1 de 图 B1 打靶约束的雅可比矩阵示意图

Fig. B1 Jacobian of constraints



图 B2 单轨道段变分示意图



$$\begin{cases} \frac{\partial \boldsymbol{x}_{n}^{f}}{\partial \boldsymbol{x}_{n}^{0}} = \boldsymbol{\varPhi}_{n} \\ \frac{\partial \boldsymbol{x}_{n}^{f}}{\partial \boldsymbol{\tau}_{n}} = -\boldsymbol{\varPhi}_{n} \dot{\boldsymbol{x}}_{n}^{0} \qquad (B5) \\ \frac{\partial \boldsymbol{x}_{n}^{f}}{\partial \boldsymbol{\tau}_{n+1}} = \dot{\boldsymbol{x}}_{n}^{f} \\ \frac{\partial \mathbf{x}_{n}^{f}}{\partial \boldsymbol{\tau}_{n+1}} = \dot{\boldsymbol{x}}_{n}^{f} \\ \text{In } \mp \boldsymbol{\tau}_{n} \pi \boldsymbol{\tau}_{n+1} \overleftarrow{\boldsymbol{x}} \overrightarrow{\boldsymbol{\pi}} \overleftarrow{\boldsymbol{y}} \\ \begin{cases} \boldsymbol{\tau}_{n} = \boldsymbol{\tau}_{1} + \frac{n-1}{N-1} \left(\boldsymbol{\tau}_{n} - \boldsymbol{\tau}_{1}\right) \\ \boldsymbol{\tau}_{n+1} = \boldsymbol{\tau}_{1} + \frac{n}{N-1} \left(\boldsymbol{\tau}_{n} - \boldsymbol{\tau}_{1}\right) \end{cases} \qquad (B6)$$

式 (B6) 对 τ_1 和 τ_N 的梯度表示为

$$\begin{pmatrix}
\frac{\partial \tau_n}{\partial \tau_1} = 1 - \frac{n-1}{N-1} \\
\frac{\partial \tau_{n+1}}{\partial \tau_1} = 1 - \frac{n}{N-1} \\
\frac{\partial \tau_n}{\partial \tau_N} = \frac{n-1}{N-1} \\
\frac{\partial \tau_{n+1}}{\partial \tau_N} = \frac{n}{N-1}
\end{cases}$$
(B7)

则 $\partial \zeta_n / \partial \tau_1 \pi \partial \zeta_n / \partial \tau_N$ 分别表示为

$$\frac{\partial \boldsymbol{\zeta}_n}{\partial \tau_1} = \frac{\partial \boldsymbol{x}_n^f}{\partial \tau_n} \cdot \frac{\partial \tau_n}{\partial \tau_1} + \frac{\partial \boldsymbol{x}_n^f}{\partial \tau_{n+1}} \cdot \frac{\partial \tau_{n+1}}{\partial \tau_1} = -\boldsymbol{\varPhi}_n \dot{\boldsymbol{x}}_n^0 \left(1 - \frac{n-1}{N-1}\right) + \dot{\boldsymbol{x}}_n^f \left(1 - \frac{n}{N-1}\right)$$
(B8)

$$\frac{\partial \boldsymbol{\zeta}_{n}}{\partial \tau_{N}} = \frac{\partial \boldsymbol{x}_{n}^{f}}{\partial \tau_{n}} \cdot \frac{\partial \tau_{n}}{\partial \tau_{N}} + \frac{\partial \boldsymbol{x}_{n}^{f}}{\partial \tau_{n+1}} \cdot \frac{\partial \tau_{n+1}}{\partial \tau_{N}} = -\boldsymbol{\varPhi}_{n} \dot{\boldsymbol{x}}_{n}^{0} \left(\frac{n-1}{N-1}\right) + \dot{\boldsymbol{x}}_{n}^{f} \left(\frac{n}{N-1}\right)$$
(B9)

附录 C: 仿真参数

表 C1 列出了数值仿真参数。

表 C1 数值仿真参数

Table C1 Numerical simulation parameters

名称	数值		
太阳质量常数 $\mu_s/(km^3 \cdot s^{-2})$	132 712 440 017.987 0		
地球质量常数 $\mu_e/(km^3 \cdot s^{-2})$	398 600.432 896 939 2		
月球质量常数 $\mu_m/(km^3 \cdot s^{-2})$	4 902.800 582 147 764		
重力加速度 g ₀ /(m·s ⁻²)	9.806 65		
地球平均半径 Re/km	6 378.137		
月球平均半径 R _m /km	1 738		
LEO地心距约束 r [*] _{leo} /km	6 578.137		
月球影响球半径 r _{soi,m} /km	66 100		
归一化距离/LU	1		
归一化时间/TU	1		
归一化速度/VU	1		
归一化地月质量参数 µ	0.012 151		
归一化太阳质量 m _s	3.28901e5		
归一化太阳距离 ρ_s	388.811		
归一化太阳角速度 ω_s	-0.925 19		
多步打靶离散点 n	20		
伪弧长延拓步长 Δs	0.1		
DRO入轨太阳相位角 $\theta_{s}(\tau_{N})$	[0,2π)		
DRO入轨相位因子 σ	[0,1)		
DRO入轨脉冲模 ΔV/(m·s ⁻¹)	[10,100]		
DRO入轨脉冲方向角 φ /rad	[0,2π)		
2:1共振DRO位置初值r ⁰ /LU	[0.808 936 204 186,0,]		
2:1共振DRO速度初值v ⁰ _{dro} /VU	[0,0.515 632 164 331]		

注: 1 LU=384 400 km, 1 TU=4.348 113 05 d, 1 VU=1.023 232 81 km/s。

Low-energy transfer from Earth into DRO with hybrid gravity assist and numerical continuation

ZHANG Chen*

(Key Laboratory of Space Utilization, Technology and Engineering Center for Space Utilization, Chinese Academy of Sciences, Beijing 100094, China)

Abstract: Distant retrograde orbits (DRO) are well-known trajectory types in cislunar space, such orbits have long-term stability and low insertion cost. In cislunar space, DRO are well-known trajectory types with minimal insertion costs and long-term stability. A cislunar station deployed on DRO might be expected to deliver a crew to the moon or Mars for exploration missions in the future. For low-energy transfer from Earth into DRO, the maximum delivery mass can be achieved by utilizing a weak stability boundary (WSB) and multiple lunar gravity assist (LGA) simultaneously, but this kind of transfer is very sensitive to initial values. A novel two-level pseudo-arc continuation method was proposed to explore local solution space, and this paper aims to improve both computational and transfer efficiency when leveraging hybrid gravity assist in cislunar space. Additionally, a modified problem description with an analytical gradient is used to improve multiple shooting efficiency under a bicircular restricted four-body problem. In the numerical simulation, the minimum cost solution shows "LGA+WSB+2LGA", where the time of flight is 123 days, the LEO launching cost is 3.125 km/s and the 2:1 DRO insertion maneuver only needs 19.7 m/s.

Keywords: N-body problem; distant retrograde orbits; weak stability boundary; Lunar gravity assist; numerical continuation

URL: link.cnki.net/urlid/11.2625.V.20220901.1058.001

Received: 2022-06-16; Accepted: 2022-08-19; Published Online: 2022-09-01 15:05

Foundation items: Strategic Priority Research Program on Space Science, the Chinese Academy of Sciences (XDA30040400); Foundation of State Key Laboratory of Space Flight Dynamics Lab (6142210200302); Youth Innovation Promotion Association, Chinese Academy of Sciences (292022000030)

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

DOI: 10.13700/j.bh.1001-5965.2022.0313

空中交通 CPS 结构特性及韧性评估

王兴隆*,魏奕雯,贺敏

(中国民航大学 空中交通管理学院,天津 300300)

摘 要: 针对空中交通体系结构复杂、耦合性强、脆弱性高的特点,为缓解因受扰导致的 级联失效现象,对其结构韧性进行研究。建立了空中交通信息物理系统 (CPS) 模型,提出同层度与 层间度、同层介数与层间介数等指标,分析级联失效过程;提出空中交通 CPS 韧性的概念,并采用 定量评估方法对受损能力和恢复能力进行度量;对比不同扰动-恢复策略下的空中交通 CPS 表现, 制定最佳恢复策略,提高受扰时的韧性。以华东地区空中交通 CPS 为例进行分析,结果表明:空中 交通 CPS 网络度分布服从幂率分布,介数服从指数分布;在基于介数度值扰动下,采用介数恢复策 略可以有效提高受扰时空中交通 CPS 的韧性。

关键词:信息物理系统;级联失效;扇区网络;航路网络;空中交通 CPS 韧性中图分类号:U8;V355.2
 文献标志码:A
 文章编号:1001-5965(2024)04-1187-10

空中交通信息物理系统(cyber physical system, CPS)是由信息系统与物理系统相互耦合形成的多 维复杂的信息物理系统^[1]。近年来, CPS 的理论和 应用逐渐成为交通运输领域研究的热点。原豪男 和郭戈利用 CPS 使得车辆能够协同调度以降低成 本^[2];徐春婕等将 CPS 应用于铁路客运站智能机房 监控平台,基于智能服务模型实现故障分析预警^[3]; Levshun 等应用安全 CPS 改进了铁路设施基础模 型^[4]; Chen 等利用车辆 CPS 对城市环境进行了短期 交通预测^[5]。

空中交通深度融合空中导航系统和地面监视 系统,实现信息识别处理与运行态势感知,符合 CPS的结构特性。对于空中交通 CPS,国内外学者 对此研究相对较少。Chen和 Zhang通过集成 Modelica 模型和架构分析设计语言来指定和建模 航空 CPS^[6]; Wang和He 建立了华东空域 CPS,对其 影响力节点进行了识别,并对系统抗毁性进行了分 析^[7]; Jeong 和 Park 基于 CPS 对空管系统中的物理 信息网进行了高度集成^[8]; Wang 和 Zhao 等将 CPS 各模块进行耦合,实现了单元级网络系统^[9]; Roy 和 Sridhar 基于 CPS 建立了空管系统评价模型,用于评 估网络发生故障后时空域系统流量的引响^[10]。以 上研究从 CPS 的角度对空中交通系统进行叙述,为 研究空中交通 CPS 奠定了基础。

由于空中交通系统的复杂性和耦合性的增加^[11], 使其在面对扰动时更加脆弱。当空中交通网络受 到扰动时,部分节点的失效通过网络中的依赖关系 扩散,从而引发级联失效。近年来,高经东通过分 析空域扇区网络级联失效过程并建立模型,提出 2种优化策略^[12];Wu等建立了级联失效传播模型, 并结合风险评估的方法对网络出现级联失效时的 鲁棒性进行分析^[13];刘明学等研究了空中交通 CPS 的级联失效现象,并提出多种策略以缓解级联失效 带来的不利后果^[14]。

为缓解级联失效的影响,"韧性"这一概念在研 究交通系统受到破坏时的恢复能力方面有着较好

*通信作者. E-mail: xl-wang@cauc.edu.cn

引用格式: 王兴隆, 魏奕雯, 贺敏. 空中交通 CPS 结构特性及韧性评估 [J]. 北京航空航天大学学报, 2024, 50(4): 1187-1196.

WANG X L, WEI Y W, HE M. Structural characteristics and resilience evaluation of air traffic CPS[J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50 (4) : 1187-1196 (in Chinese).

收稿日期: 2022-05-03; 录用日期: 2022-07-16; 网络出版时间: 2022-07-19 14:29 网络出版地址: link.enki.net/urlid/11.2625.V.20220718.1835.003

基金项目:国家自然科学基金(62173332);国家重点研发计划(2020YFB1600101)

的应用^[15]。韧性是指系统在外界干扰下通过减少 初始负面影响,自适应调节使其适应扰动,并最终 从扰动中恢复。杨琦等从公共交通网络韧性出发, 阐述了复杂网络在公共交通韧性领域的研究现 状^[16];范瀚文等对海上运输关键节点安全韧性因素 进行识别,引入熵权法对指标赋权进行分析^[17];马 书红等以关中平原城市群为例,分析交通网络应对 攻击时的韧性变化特征^[18];石宗北等通过研究空中 交通网络下的子图结构在外界扰动下的变化规律, 评估其结构韧性^[19]; Miao 等结合复杂网络理论, 对 中国航空网络模型进行脆弱性及韧性分析,构建更 高的系统韧性^[20]。

上述研究用韧性指标对交通网络进行评估及 优化,但大部分仅从单层网络入手,只能考虑到系 统内部的变化,忽略了网络的多维耦合特征。为更 好地把握空中交通的整体性,本文建立了空中交通 CPS 模型,将物理网(航路网)和信息网(扇区网)相 互耦合,对相关韧性指标进行分析,研究空中交通 CPS 在度值、介数扰动及恢复下的韧性变化规律, 为优化空中交通 CPS、提高安全效率提供参考。

1 空中交通 CPS 网络

1.1 空中交通 CPS 构建

CPS 是综合计算、网络和物理环境的多维复杂 系统,已有学者通过 CPS 对具有耦合关系的系统进行 建模,并进行相关研究^[21]。根据复杂网络理论^[22], 结合现有空中交通规则^[23]及管制扇区与航路之间 的关系,可将空中交通系统拓扑为相互耦合的2层 网络,即空中交通 CPS,该系统是指将计算、网络、 信息元素和物理元素彼此融合交织,目的是促进信 息物理元素的互动和数据敏感的自动决策,为空中 交通系统结构特性及韧性研究提供新的思路。

空中交通 CPS 由信息网与物理网耦合而成。 其中,空中交通 CPS 结构如图 1(a) 所示,信息网为 扇区网,以管制扇区几何中心为节点,以扇区间航 班移交关系为边;物理网为航路网,以航路点为节 点,以2个航路点之间的航段为边。

扇区网与航路网之间存在着一定的耦合关系, 具体结构如图 1(b)所示。其中,航路网为物理网, 承载着空中交通流的运行,包括航空器在航路网中 沿航段的飞行;扇区网为信息网,承载着信息流的 传递,包括气象因素、管制指挥、交通态势等在信 息网中的传递。在实际运行中,空中交通 CPS 通过 地空数据链,将空中交通信息进行感知后与通信相 互融合。扇区管制员接收气象监视台、雷达、ADS-B等设备监测到的数据,对其进行分析处理,并据



Fig. 1 Schematic diagram of air traffic CPS

此发送指令至扇区内航空器;飞行员通过反馈相应 指令,改变当前航空器飞行速度、高度、位置等,实 现空域内多架航空器的高效、安全运行。

为方便研究, 对空中交通 CPS 扇区网做以下 假设:

 4)每个管制扇区只负责管辖在其内运行的 航班。

2) 合并空域中不同高度的扇区, 视为统一 高度。

3) 将扇区网视为无向网络。

对空中交通 CPS 航路网做以下假设:

 1)将机场、导航台和规定报告点等视作航路 点,当导航台与机场位置重合时,仅考虑机场点。

2) 不考虑空域中的临时航线及国际航线, 忽略 部分孤立边界航路点。

3)将航路网视为无向网络。

将空中交通 CPS 表示为集合 $G = (G^{c}, G^{p}, E),$ $G^{c} = \{V^{c}, E^{c}\}$ 为扇区网, $G^{p} = \{V^{p}, E^{p}\}$ 为航路网, $V^{c} = \{V_{i}, i = 1, 2, \dots, n\}$ 和 $V^{p} = \{V_{i}, i = 1, 2, \dots, m\}$ 分别为 扇区网和航路网中的节点集合, $E^{c} = \{e_{ij} \neq 0 | i, j = 1, 2, \dots, m\}$ 分别为扇区 网和航路网中连边的集合, n和m分别为扇区网和 航路网中节点个数。用邻接矩阵 α (维度为 $n \times n$)和 β (维度为 $m \times m$)表示扇区网和航路网中的节点层 内连接关系,若节点*i*与*j*之间存在连边,则 α_{ij} =1或 β_{ij} =1;否则, α_{ij} =0或 β_{ij} =0。 $E = \{\gamma_{ij} \neq 0 | i \in V^{C}, j \in V^{P}\}$ 表示扇区网中节点与航路网中节点的连边,即层间 连接关系,若扇区*i*和航路点*j*相连,则 γ_{ij} =1,否则 γ_{ij} =0。

1.2 空中交通 CPS 特征参数

空中交通 CPS 的扇区网和航路网在复杂网络的基础上具有特有的领域特征,通过定义以下指标,将空域扇区特性通过网络层面进行分析,充分反映空中交通 CPS 的结构特性。

 1) 度。度是复杂网络常用指标之一,其含义为 与某一节点相连的节点个数,计算公式如下:

$$K_i = \sum_{i=1}^{n} \alpha_{ij} \tag{1}$$

式中: n为网络中的节点总数; α_{ij} 为节点i和j的连接 情况, 当i和j相连, 则 α_{ij} =1, 反之则 α_{ij} =0。

在空中交通 CPS 中, 节点的度可以分为同层度 与层间度。同层度表示与一个节点相连的同一层 网络中的节点个数, 对于航路网, 节点同层度越大, 与该航路点存在航段连接关系的航路点越多, 其在 网络中的重要性越强; 对于扇区网, 节点同层度越 大, 则该扇区与越多扇区之间存在航班移交关系。 层间度表示与一个节点相连的不同层网络的节点 个数, 航路点的层间度表示对其进行控制指挥的扇 区个数, 管制扇区的层间度表示其范围内所辖航路 点个数。节点的度计算公式如下:

$$K_i = K_{i1} + K_{i2}$$
 (2)

式中: Ki1为节点i的同层度; Ki2为节点i的层间度。

度累积分布*P*(*K*)表示网络中度不小于*K*的节 点数占网络节点总数的比例,计算公式如下:

$$P(K) = \sum_{K'=K}^{\infty} P_{K'} \tag{3}$$

式中: *P_{K'}*为度分布, 表示空域扇区网络中度为*K'*的 节点数占总节点数的比例。

2)介数。节点的介数表示网络中所有最短路
 径中经过该节点的数量比例,计算公式为

$$B_i = \sum_{\substack{i \neq k}} \frac{n_{jk}(i)}{n_{jk}} \tag{4}$$

式中: n_{jk}为节点 j和节点 k之间最短路径数; n_{jk}(i)为 节点 j和节点 k之间最短路径中经过扇区 i的数目。

对于空中交通 CPS,其介数同样分为同层介数 与层间介数。同层介数表示同一层网络中所有最 短路径中经过该节点的路径数目占所有最短路径 总数的比例。对于航路网,若任意2个航路点之间 的多条最短路径都经过某个航路点,则证明该节点 同层介数较大,对应该节点在网络中的信息传输能 力越强;对于扇区网,若多条最短路径交汇于介数 较大的节点,使得该点成为扇区网的中心,则证明 该节点同层介数较大,该扇区对整个扇区网来说具 有更强的信息传输能力。

层间介数表示与一个节点相连的不同层网络 的节点介数。航路点的层间介数表示其经过扇区 的路径数目占所有扇区最短路径数目的比例,扇区 网的层间介数表示其范围内所辖航路点经过扇区 的路径数目占总航路点经过扇区的路径数目的比例。 节点的介数计算公式如下:

$$B_i = B_{i1} + B_{i2} \tag{5}$$

式中: B_{il}为节点i的同层介数; B_{i2}为节点i的层间 介数。

介数累积分布 P(B)表示网络中介数不小于 B节点数占网络节点总数的比例,计算公式如下:

$$P(B) = \sum_{B'=B}^{\infty} P_{B'} \tag{6}$$

式中: P_B,为介数分布,表示空域扇区网络中介数为 B'的节点数占总节点数的比例。

1.3 空中交通 CPS 的级联失效过程

在空中交通 CPS 中,扇区网和航路网间存在着 一定的对应关系,即每个管制扇区对扇区网内的多 个航路点进行指挥。当技术故障、恶劣天气等特殊 事件发生时,受影响范围会从单条航段逐渐蔓延至 整个扇区或多个扇区,甚至扩散到整个空域系统, 引起空中交通 CPS 级联失效,从而使空域内航班大 量延误或取消,造成严重后果。

一旦发生级联失效,空中交通 CPS 可能迅速崩 溃直至瘫痪,如图 2 所示,主要经历以下 4 个阶段:



图 2 空中交通 CPS 级联失效示意图



1) 正常运行。航路网及扇区网均正常运行。

2)开始失效。网络中部分节点由于特殊事件 发生受到扰动而失效,为保证网络持续运行,若航 路网内某一航路点失效,则将失效航路点内的航班 分配至相邻航路点;若扇区网内某一管制扇区失 效,则将其管制任务分配至相邻扇区,与其相连的 航路点由于空中交通 CPS 的耦合特征而失效。

3) 失效扩散。接收了相邻节点分配的管制任

务或航班后,部分管制扇区和航路点因功能受到影响而失效,继续将其管制任务和航班分配至相邻节点;部分节点由于脱离了最大连通子图,与其他节点失去联系而失效,级联失效在网络中的影响范围进一步增大。

 4)网络恢复。当网络严重受损几乎瘫痪后,由 于管制员指挥疏通,交通流降低,网络逐渐恢复 正常。

在实际运行过程中,由于级联失效的扩散影响, 可能导致网络迅速瘫痪,使得空中交通系统受损。 因此,对空中交通 CPS 采用韧性评估方法对其进行 度量,对比分析不同扰动-恢复过程下的韧性变化, 探寻空中交通 CPS 在外界干扰下的最佳恢复策略, 从而提高网络韧性,以及运行效率和安全性。

2 空中交通 CPS 韧性评估

2.1 韧性受损及恢复过程

为缓解级联失效带来的破坏,引用韧性这一概 念研究空中交通 CPS 受到破坏时的恢复能力。将 空中交通 CPS 的韧性定义为:在外界干扰下,网络 能够有效应对风险扰动,降低性能损失,并通过自 适应调节快速恢复到原有状态或达到一定的可接 受状态,从而保证系统正常运行的能力。当系统面 临扰动时,根据 Nan 和 Sansarini^[24]提出的评估韧性 模型,可将其受损恢复过程分为图 3 所示的 4 个阶 段,每个阶段分别对应不同的能力。其中,纵坐标 *L(t)*表示当前系统的性能水平,表征性能水平的指 标有网络连通度、延误架次、延误时间等。本文主 要考虑空中交通 CPS 的连通性。由于指标存在一 定差异,对其进行归一化处理:





$$v_i = \frac{u_i - u_{\min}}{u_{\max} - u_{\min}}$$
 $i = 1, 2, 3, \cdots$ (7)

式中: v_i为经归一化处理后的第i个指标; u_i为性能 水平指标中第i个指标; u_{min}和u_{max}分别为性能水平 指标中的最小值和最大值。

L值经归一化处理后,取值范围在[0,1]之间变

动,系统受扰瘫痪后,L值逐渐减至0,系统处于稳定状态时,L值为1。

第1阶段($t_0 < t < t_d$)是系统从初始稳定阶段到 扰动事件发生前的预防阶段,该阶段可以反映出系 统的预测能力,即系统在发生扰动前阻止或避免发 生扰动的能力。第2阶段($t_{i} < t < t_{i}$)是系统受到扰 动后造成失效后的阶段;在该阶段中,系统在ta时刻 (A点)受到扰动,扰动会对系统产生初步损伤 (B点), 而在损伤持续传播后, 扰动会使系统性能 逐步下降直至在t,到达最低值(C点), B 点到 C 点 的曲线是由系统的级联失效所导致。由图2可以 看出,第2阶段的周期远小于其他阶段,因此,通常 将 BC 两点之间的过程看作瞬时过程。通过上述表 述,该阶段侧重于系统对扰动事件发生后的吸收能 力。第3阶段($t_r < t < t_{ns}$)是系统从最大受损值(C点) 逐渐恢复到初始状态(即D点)下的阶段,该阶段侧 重于系统的适应及恢复能力,在扰动事件发生后, 可以及时恢复,以便系统得到快速改善。第4阶段 $(t > t_{ns})$ 是系统达到新的稳定状态(D点),此时系统 的性能水平会处于一个新的稳定状态,由于系统在 恢复过程中可能受到改动,致使恢复后的性能水平 可能不能完全等于初始时的性能水平,在此不做 考虑。

2.2 韧性评估指标

在空中交通 CPS 网络中,可以用以下指标衡量 评估系统在破坏阶段中吸收扰动的能力。

1) 鲁棒性。鲁棒性反映空中交通 CPS 受到扰 动时, 仍能保持网络稳定、正常运行的能力^[25]。其 定义如下:

$$R = \min\{L(t)\}(t_{\rm d} < t < t_{\rm ns})$$
(8)

式中: *L*(*t*)为离散函数,其值随时间变化; *t*_{ns}为系统 到达新的稳定状态的时刻; *t*_d为系统受到外界干扰 开始下降时刻; *R*可以衡量当空中交通 CPS 受扰时 对系统产生的最大的影响。

2)破坏速率 *P*_{DP}和恢复速率 *P*_{RP}。破坏速率反 映空中交通 CPS 网络受扰后性能损失的快慢程度, 破坏阶段(disruptive phase, DP)曲线的平均斜率即 为该值的大小,其定义如下:

$$P_{\rm DP} = \frac{1}{K_{\rm DP}} \left| \sum_{i=1}^{K_{\rm DP}} \frac{L(t_i) - L(t_i - \Delta t)}{\Delta t} \right| \tag{9}$$

 $\Delta t = t_{\rm r} - t_{\rm d}$

式中: K_{DP} 为扰动阶段($t_d < t_i < t_r$)检测到的线段数 量, i为该阶段的第i条线段; $L(t_i)$ 为该阶段第i条线 段上L的值; Δt 为扰动阶段时间变化。

恢复速率表示空中交通 CPS 网络扰动消失后

(10)

(12)

的恢复快慢程度,即系统在恢复阶段(recovery phase, RP)的平均斜率,计算公式如下:

$$P_{\rm RP} = \frac{1}{K_{\rm RP}} \left| \sum_{i=1}^{K_{\rm RP}} \frac{L(t_i) - L(t_i - \Delta t)}{\Delta t} \right|$$
(11)

 $\Delta t = t_{\rm ns} - t_r$

式中: K_{RP} 为恢复阶段($t_r < t_i < t_{ns}$)检测到的线段数 量; Δt 为恢复阶段时间的变化。

3) 性能损失 F。反映空中交通 CPS 受扰后引 起的扰动图像与扰动发生前后L(t)曲线围成的区域 的面积, 其定义如下:

$$F_{\rm DP} = \int_{t_{\rm d}}^{t_{\rm r}} (L(t_0) - L(t)) dt$$
 (13)

式中: t₀为 CPS 网络处于初始状态的时刻, L(t₀)为 CPS 网络未受干扰时的初始水平。

4)单位时间性能损失 *T*。当空中交通 CPS 网络受外界干扰导致性能损失时,由于破坏事件所引起的相关状态可能会影响系统受扰时刻,引入单位时间性能损失。在扰动阶段,其定义如下:

$$T_{\rm DP} = \frac{1}{t_{\rm r} - t_{\rm d}} \int_{t_{\rm d}}^{t_{\rm r}} (L(t_0) - L(t)) {\rm d}t \qquad (14)$$

式中: t_r为系统受到最大损伤时刻; t_d为系统受到外 界干扰开始下降时刻。

在恢复阶段,其定义为

$$T_{\rm RP} = \frac{1}{t_{\rm ns} - t_{\rm r}} \int_{t_{\rm r}}^{t_{\rm ns}} [L(t_0) - L(t)] dt$$
 (15)

式中:tns为系统结束扰动时刻。

5)恢复能力 A。当空中交通 CPS 网络经历扰 动和恢复后,其新的稳定状态下的性能水平或高或 低于初始性能,定义其恢复能力如下:

$$A = \left| \frac{L(t_{\rm ns}) - L(t_{\rm r})}{L(t_0) - L(t_{\rm r})} \right|$$
(16)

式中: $L(t_r)$ 为 CPS 网络受到最大损伤时的性能水 平; $L(t_n)$ 为 CPS 网络结束扰动时的性能水平。

为更加全面度量韧性过程,定义综合韧性指标 N表示空中交通 CPS 在受到扰动时的韧性,计算公 式如下:

$$N = f(R, P_{\rm DP}, P_{\rm RP}, T, A) = R \frac{P_{\rm RP}}{P_{\rm DP}} T^{-1} A$$
 (17)

这一指标能够反映 CPS 网络受扰后的系统韧性 水平,当扰动事件发生后,若系统能够有效减小由扰 动带来的偏差,则认为系统有着较好的韧性水平。

2.3 空中交通 CPS 韧性评估过程

空中交通 CPS 运行过程较为复杂,在特殊事件中受到外界影响时将导致网络受损,使整个网络运行性能下降。为降低由于外界干扰导致的扇区网络性能下降,通过对失效节点采用适当的恢

复策略使得网络性能得到恢复,从而达到初始正 常水平。采用不同失效-恢复策略对空中交通 CPS 的韧性进行评估,主要分为4个阶段,其流程如图4 所示。



Fig. 4 Flow chart of resilience assessment

1)预防阶段。开始时,空中交通 CPS 处于正 常运行状态。

2) 扰动阶段。采用度值、介数 2 种形式对网络 进行扰动, 节点受到扰动会因崩溃而失效。对于扇 区网, 设备故障、恶劣天气会导致网络失效; 对于航 路网, 军事演练、交通态势冲突会导致网络失效。 将所有节点按度值从大到小排序, 先攻击度值较大 的节点, 依次按一定时间间隔, 使得空中交通 CPS 失效 10% 的节点, 将失效节点从网络中移除, 再将由于脱离最大连通子图而与网络失去联系的 节点也删除, 直至正常运行节点数降低至 50%。

3)恢复阶段。根据度值、介数对空中交通 CPS进行恢复。根据节点度值大小,从大到小排 序,按10%的速率先恢复度值较大的节点。航路网 和扇区网的层内和层间连接关系能够重新建立,空 中交通 CPS性能逐渐稳定。

4)恢复终止。节点数量恢复到100%,空中交
(18)

通 CPS 恢复至新的稳定状态,可以重新正常运行。 在韧性评估过程中,采用最大连通子图表征网 络的连通性能,计算公式如下:

G = g'/g

式中:g为扇区网和航路网在受到扰动前的节点数; g'为受到攻击后扇区网或航路网最大连通子图中 的节点数。

3 实证分析

本文选取华东地区空中交通管理局所辖空域, 建立空中交通 CPS 模型。图 5(a)为华东地区空中 交通 CPS 航路网,包含 171 个航路点和 262 条边; 图 5(b)为华东地区空中交通 CPS 扇区网,由 33 个 管制扇区和 69 条边构成;2 层网络相互耦合形成复 杂的多维空中交通 CPS 模型。

3.1 空中交通 CPS 网络特性分析

华东地区空中交通 CPS 的度累积分布如图 6 所示。度值是表示节点重要性的关键指标,当某一 节点的度值较大时,表明该节点周围有较多节点与 之相连,大多为网络的中心节点,对网络影响较 大。其中,最大度值为 24,最小度值为 1。在图 6 中, 空中交通 CPS 的度累积分布在双对数坐标下服从 双段幂率分布,随着节点度值的增加,扇区数量开 始递减,当节点度值大于平均度时,扇区数量将会 呈现快速下降的趋势。度分布规律较不均匀,具有 复杂网络所特有的无标度特征。

华东地区空中交通 CPS 的介数累积分布如 图 7 所示。网络介数分布呈现指数分布, 拟合于 *P*(*B*) = 0.872 5e^{-40.27B}, *R*² = 0.965 5。随着扇区节点 度值的增加, 扇区数量逐渐减少至 0。介数大的 扇区通常处于整个扇区网的中心位置, 很大程度上 影响其他扇区的正常运行情况; 介数小的扇区通 常处于整个扇区网的边缘位置, 虽然对其他扇区正 常运行影响不是很大, 但是在整个扇区网中保证其 正常运行也起着重要的作用, 从而保证航班的有效 运行。

3.2 空中交通 CPS 网络韧性性能分析

根据 2.3 节中的扰动-恢复规则,对华东地区空 中交通 CPS 进行扰动和恢复对比,并运用 2.2 节中 韧性相关指标,对该过程中的韧性进行度量。

图 8 为不同扰动及恢复策略下扇区网最大连 通子图相对值的变化。图中纵轴数字反映网络性 能指标值的变化,越接近于 1,网络性能越好,越接 近于 0,网络性能越差。可以看出,介数扰动对网络 性能的影响明显强于度值扰动。在 t=2 之前,网络 并未受到干扰,处于预防阶段,网络性能水平保持



在初始稳定水平。在 t=2 时,网络开始受到外界干扰,性能开始下降,并在 t=7 时下降到最低值。

在扰动阶段,对比图 8(a)、图 8(b)和图 8(c)、 图 8(d),当 t=4时,基于度值扰动的扇区网性能下 降至 0.670 5,在同时刻下,基于介数扰动的扇区网 性能下降至 0.398 8。当 t=7时,性能达到最低值, 其中度值扰动为 0.09,介数扰动为 0.11。可以看 出,介数扰动对扇区网破坏速度较快,但度值扰动 对扇区网破坏较为严重。因此,在实际运行中应先





保护介数较大的扇区正常运行。

在恢复阶段,对比图 8(a)、图 8(d)和图 8(b)、 图 8(c),当 t=9时,基于介数扰动下的介数恢复上 升至 0.734 1,在同时刻下,基于介数扰动下的度值 恢复上升至 0.672 4。因此,在扰动发生后,介数恢 复效果更好,能以更快的速度使其免受外界干扰, 从而保证扇区网的正常运行。

图 9 为不同扰动及恢复策略下航路网最大连 通子图相对值的变化。在扰动阶段,对比图 9(a)、 图 9(b)和图 9(c)、图 9(d),当 *t*=4 时,基于度值扰





动的航路网网性能下降至 0.415 2, 在同时刻下, 基 于介数扰动下的航路网性能下降至 0.508 8。当 t=7时,性能达到最低值,其中度值扰动为 0.029, 介 数扰动为 0.076。可以看出, 度值扰动对航路网破 坏速度较快, 也较为严重。因此, 在实际运行中, 应 先保护度值较大的航路正常运行。

在恢复阶段,对比图 9(a)和图 9(b),当 t=9时, 在介数恢复下航路网性能上升至 0.573 1,此时度值 恢复上升至 0.415 2,因此,在扰动发生后,介数恢复 速度较快,恢复性能较好。对比图 9(c)和图 9



Fig. 8 Changes in maximal connected subgraph of sector network



图 9 航路网最大连通子图变化

Fig. 9 Changes in maximal connected subgraph of air route network

(d),在 t=9时,度值恢复速度较快,但结合韧性值 N,总体效果略差于介数恢复。综上,介数恢复能以 更快速度使其免受外界干扰,从而保证航路网正常 运行。

图 10 为华东地区空中交通 CPS 在不同扰动方 式下的网络性能损失。可以看出,度值扰动对航路 网的影响和破坏最大,使网络性能损失最多;而介 数扰动使扇区网的性能损失最大。

通过对图 8、图 9 分析计算,得出华东地区空中 交通 CPS 在不同扰动方式下的综合韧性评价值,如 图 11 所示。可以看出,对于扇区网,介数扰动下综 合韧性值较小,表现出较差的韧性;对于航路网,度





Fig. 11 General resilience of air traffic CPS

1195

值扰动下综合韧性值较小,表现出较差的韧性。韧 性越差,表示易在外界干扰下趋于瘫痪状态。因 此,在实际运行状况下,应先对度值较大的的扇区 网和介数较大的航路网进行保护,避免级联失效引 起的网络整体性能快速降低。

在恢复阶段,由图9可以看出,在 =7 时,网络性能开始恢复,并于 =12 时逐渐恢复至初始稳定状态。对比图11 中2组柱状图,在度值和介数扰动下,介数恢复在航路网和扇区网中恢复效果均优于度值恢复,因此,在实际运行过程中,在扰动发生后采取介数恢复能够快速恢复网络状态,使其正常运行。

4 结 论

1)构建空中交通 CPS 网络模型,分析扇区网 与航路网之间的耦合关系,在复杂网络的基础上定 义相关特征参数,分析在网络耦合依赖影响下的 CPS 网络级联失效过程。

2)结合韧性概念,分析空中交通 CPS 受扰动 后其性能下降和恢复的阶段性过程,定义综合韧性 值,并采用不同扰动-恢复策略对网络韧性进行 评估。

3)以华东地区空域为例进行实证分析,结果表明,华东地区空中交通 CPS 度服从双段幂率分布, 介数服从指数分布;在扰动阶段,介数扰动对扇区 网影响较大,而度值扰动对航路网破坏更严重;在 恢复阶段,不同扰动方式下介数恢复效果均最好, 使扇区网和航路网最快恢复至新的稳定状态。

在后期的研究中,将考虑在级联失效情况影响 下,空中交通 CPS 网络该如何进一步提升其韧性及 应对特殊事件的能力。

参考文献(References)

[1] HUSY, LIUCL, PIURIV. 信息物理系统专题简介[J]. 中国科学:信息科学, 2022, 52(2): 376.
 HUSY, LIUCL, PIURIV. Introduction to information physical

systems[J]. Science in China:Information Science, 2022, 52(2): 376(in Chinese).

- [2] 原豪男, 郭戈. 交通信息物理系统中的车辆协同运行优化调度[J]. 自动化学报, 2019, 45(1): 143-152.
 YUAN H N, GUO G. Vehicle cooperative optimization scheduling in transportation cyber physical systems[J]. Acta Automatica Sinica, 2019, 45(1): 143-152(in Chinese).
- [3] 徐春婕,陈瑞凤,贺晓玲,等. 基于 CPS 的铁路客运站智能机房监 控平台[J]. 计算机工程与设计, 2022, 43(2): 596-601.
 XU C J, CHEN R F, HE X L, et al. Intelligent computer room monitoring platform based on CPS for railway passenger station computer room[J]. Computer Engineering and Design, 2022, 43(2): 596-601(in Chinese).
- [4] LEVSHUN D, KOTENKO I, CHECHULIN A. The application of

the methodology for secure cyber-physical systems design to improve the semi-natural model of the railway infrastructure[J]. Microprocessors and Microsystems, 2021, 87: 103482.

- [5] CHEN C, LIU X M, QIU T, et al. A short-term traffic prediction model in the vehicular cyber-physical systems[J]. Future Generation Computer Systems, 2020, 105: 894-903.
- [6] CHEN W Q, ZHANG L C. Physical and cyber convergence approach to design future complex aviation cyber physical systems [C]//Proceedings of the 6th IEEE International Conference on Software Engineering and Service Science. Piscataway: IEEE Press, 2015: 541-544.
- [7] WANG X L, HE M. Influential node ranking and invulnerability of air traffic cyber physical system[J]. Transactions of Nanjing University of Aeronautics and Astronautics, 2021, 38(2): 288-297.
- [8] JEONG Y S, PARK J H. Adaptive network-based fuzzy inference model on CPS for large scale, intelligent and cooperative surveillance[J]. Computing, 2013, 95(10): 977-992.
- [9] WANG H J, ZHAO H T, ZHANG J, et al. Survey on unmanned aerial vehicle networks: A cyber physical system perspective[J]. IEEE Communications Surveys & Tutorials, 2020, 22(2): 1027-1070.
- [10] ROY S, SRIDHAR B. Cyber-threat assessment for the air traffic management system: A network controls approach[C]//Proceedings of the 16th AIAA Aviation Technology, Integration, and Operations Conference. Reston: AIAA, 2016: 4354.
- [11] 郭九霞. 新一代民航运输系统安全韧性理论与方法研究[D]. 成都: 电子科技大学, 2021: 1-5.
 GUO J X. Research on the theory and method of safety and toughness of new generation civil aviation transportation system[D].
 Chengdu: University of Electronic Science and Technology of China, 2021: 1-5(in Chinese).
- [12] 高经东. 基于复杂网络的空域扇区网络分析及抗毁性研究[D]. 天津: 中国民航大学, 2018: 27-32.
 GAO J D. Analysis and invulnerability research of sectors network based on complex network[D]. Tianjin: Civil Aviation University of China, 2018: 27-32(in Chinese).
- [13] WU Y P, CHEN Z L, ZHAO X D, et al. Robust analysis of cascading failures in complex networks[J]. Physica A:Statistical Mechanics and Its Applications, 2021, 583: 156320.
- [14] 王兴隆, 贺敏, 刘明学. 空中交通 CPS 级联失效与缓解策略[J]. 北 京航空航天大学学报, 2021, 47(12): 2426-2433.
 WANG X L, HE M, LIU M X. Air traffic CPS cascading failure and mitigation strategy[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(12): 2426-2433(in Chinese).
- [15] HOSSEINI S, BARKER K, RAMIREZ-MARQUEZ J E. A review of definitions and measures of system resilience[J]. Reliability Engineering & System Safety, 2016, 145: 4761.
- [16] 杨琦,张雅妮,周雨晴,等. 复杂网络理论及其在公共交通韧性领域的应用综述[J]. 中国公路学报, 2022, 35(4): 215-229. YANG Q, ZHANG Y N, ZHOU Y Q, et al. A review of complex network theory and its application in the resilience of public transportation systems[J]. China Journal of Highway and Transport, 2022, 35(4): 215-229(in Chinese).
- [17] 范瀚文,常征,王聪.海上运输通道关键节点安全韧性影响因素 及评价[J]. 上海海事大学学报, 2022, 43(2): 40-46.
 FAN H W, CHANG Z, WANG C. Influencing factors and evaluation of safety resilience of key nodes in maritime transportation

channels[J]. Journal of Shanghai Maritime University, 2022, 43(2): 40-46(in Chinese).

[18] 马书红, 武亚俊, 陈西芳. 城市群多模式交通网络结构韧性分析: 以关中平原城市群为例[J]. 清华大学学报 (自然科学版), 2022, 62(7): 1228-1235.

MA S H, WU Y J, CHEN X F. Structural resilience of multimodal transportation networks in urban agglomerations: A case study of the Guanzhong Plain urban agglomeration network[J]. Journal of Tsinghua University (Science and Technology), 2022, 62(7): 1228-1235(in Chinese).

- [19] 王兴隆, 石宗北, 陈仔燕. 空中交通网络模体识别及子图结构韧 性评估[J]. 航空学报, 2021, 42(7): 324715.
 WANG X L, SHI Z B, CHEN Z Y. Air traffic network motif recognition and sub-graph structure resilience evaluation[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(7): 324715(in Chinese).
- [20] WANG X L, MIAO S F, TANG J Q. Vulnerability and resilience analysis of the air traffic control sector network in China[J]. Sustainability, 2020, 12(9): 3749.
- [21] REN L, LIAO H, CASTILLO-EFFEN M, et al. Transformation of mission-critical applications in aviation to cyber-physical systems [M]//SONG H B, RAWAT D B, JESCHKE S, et al. Cyber-physical

systems: Foundations, principles, and applications. Amsterdam: Elsevier, 2017: 339-362.

- [22] 徐开俊,肖成坤,杨泳,等. 基于复杂网络理论的中国城市航空网 络有向加权分析[J]. 科学技术与工程, 2021, 21(36): 15669-15673. XU K J, XIAO C K, YANG Y, et al. Directed weighted analysis of Chinese urban aviation network based on complex network theory[J]. Science Technology and Engineering, 2021, 21(36): 15669-15673(in Chinese).
- [23] 中国民用航空局. 民用航空空中交通管理规则: CCAR-93TM-R5[S]. 北京: 中国民用航空局, 2017.
 Civil Aviation Administration of China. Air traffic management rules for civil aviation: CCAR-93TM-R5[S]. Beijing: Civil Aviation Administration of China, 2017(in Chinese).
- [24] NAN C, SANSAVINI G. A quantitative method for assessing resilience of interdependent infrastructures[J]. Reliability Engineering & System Safety, 2017, 157: 35-53.
- [25] 徐野.复杂互联系统与网络鲁棒性研究[M].北京:电子工业出版 社,2015:37-40.

XU Y. Study of robustness in complex interconnected system and networks[M]. Beijing: Publishing House of Electronics Industry, 2015: 37-40(in Chinese).

Structural characteristics and resilience evaluation of air traffic CPS

WANG Xinglong^{*}, WEI Yiwen, HE Min

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

Abstract: An air traffic network is a multi-dimensional complex system, which becomes vulnerable due to the coupling relationship when disturbed. This research examines the structural resilience of the system to mitigate the cascading failures caused by disturbance. An air traffic cyber physical system (CPS) model is developed. The relevant indicators such as the intra-layer degree and inter-layer degree, and intra-layer betweenness and inter-layer betweenness are defined, and the cascading failure process is analyzed. The resilience of air traffic CPS is then proposed, and a quantitative assessment method is used to measure the system's disturbance response and recovery capabilities. The performance of the air traffic CPS is compared under different perturbation-recovery strategies to improve its resilience when perturbed. The results show that the air traffic CPS network degree follows a power law distribution while the betweenness follows an exponential distribution. For both perturbation methods, betweenness recovery leads to fastest recovery of air traffic CPS performance.

Keywords: cyber physical system; cascading failure; sector network; air route network; air traffic CPS resilience

Foundation items: National Natural Science Foundation of China (62173332); National key Research and Development Program of China (2020YFB1600101)

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

DOI: 10.13700/j.bh.1001-5965.2022.0383

基于模型预测控制算法的精确着舰控制方法

万兵,苏析超*,汪节,韩维,李常久

(海军航空大学,烟台 264001)

摘 要: 针对舰载机着舰过程中的航母运动、舰尾流扰动等影响,提出一种基于模型预测 控制-线性二次型高斯 (MPC-LQG) 算法的精确着舰控制方法,开展飞机纵向通道的精确着舰控制研 究。对飞机纵向动力学、舰尾流、航母运动、轨迹跟踪等,进行建模和分析。对 MPC-LQG 算法进 行融合,以实现轨迹跟踪控制,核心思想是:模型预测控制进行航母运动补偿;设计全维状态观 测,实现全状态反馈,以实现最优着舰控制。对不同的着舰情况和初始条件进行算法仿真,并与其 他算法进行仿真比较。仿真结果表明,所提方法的轨迹跟踪效果很好,高度偏差 0.1~0.2 m;相比 传统着舰控制,所提方法动态响应快、着舰精度高。

关 键 词: 模型预测控制-线性二次型高斯; 甲板运动; 引导指令; 着舰控制; 轨迹跟踪 中图分类号: V249; TP273

文献标志码: A 文章编号: 1001-5965(2024)04-1197-11

舰载机在进场末段截获并进入着舰进近窗(锁定下滑道和对准斜甲板中线)后进行下滑着舰飞行。假如航母运动与舰尾流影响不大,舰-机速度恒定情况下,着舰过程则为2个位动平台遭遇问题,该过程几乎无侧向运动且容易操控,而纵向运动则是在给定飞机配平的平衡点下由比例引导+反馈控制则可完成着舰^[1-3]。然而,实际中航母随机运动和舰尾流强烈扰动是不可避免的,因此,本文着重研究舰载机着舰纵向轨迹精确跟踪控制。

目前,着舰引导技术已经历了数次迭代发展: 由仪表引导(instrument landing system, ILS)到微波 引导(microwave landing system, MLS),再到基于 GPS的联合精确进近和着陆系统(joint precision approach and landing system, JPALS)^[3-4],以及舰载机 自主感知引导着舰^[5-6]。全自动着舰引导控制系统 (automatic carrier landing system, ACLS)包含着舰引 导和着舰控制 2 个部分,本文重点是对着舰控制方 法进行研究。

飞行控制技术的发展,从基于经典控制理论的

逐环闭合设计到现代控制技术的增益矩阵综合设 计,再到基于模型辨识、控制律组合设计等智能控 制方法的演化,整个控制理论越来越成熟^[7]。然而, 对于舰载机着舰控制问题,理想引导下滑道是明确 的,但航母动平台和舰尾流干扰是必须面对的问 题,一个是甲板运动补偿,另一个是干扰的抑制^[8]。 目前,着舰控制方面进行了大量的研究。

经典比例积分微分 (proportional integral derivative, PID) 控制技术方面, 文献 [9-10] 提出改进鸽群优化方法对着舰控制器的多环 PID 控制参数进行寻优。

现代控制技术方面,针对不确定性尾流影响, Zhu和Yang^[11]提出一种自适应滑模反演控制方法, 来抑制舰尾流对着舰飞行轨迹的影响,其干扰抑制 效果比较好。理论上滑模控制抗干扰能力强,但仿 真发现其控制抖振现象比较明显,不利于实际 应用。

针对航母运动补偿问题, Wang 等^[12]提出一种 带有附加项的模型预测控制(model predictive control,

收稿日期: 2022-05-18; 录用日期: 2022-07-17; 网络出版时间: 2022-08-09 09:24

网络出版地址: link.cnki.net/urlid/11.2625.V.20220808.1654.002

^{*}通信作者. E-mail: suxich@126.com

引用格式: 万兵,苏析超,汪节,等.基于模型预测控制算法的精确着舰控制方法 [J].北京航空航天大学学报,2024,50(4):1197-1207. WAN B, SUX C, WANG J, et al. A precise landing control method based on model predictive control algorithm [J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50(4):1197-1207 (in Chinese).

MPC)算法,该方法基于纵向着舰点的预测,建立着 舰风险模型,将甲板运动预测也纳入控制系统中, 实现航母运动补偿控制。

Koo 等^[13]设计了一种考虑航母甲板运动的 MPC 自动着舰控制系统,通过设计模型预测控制器,在一定时间范围内获得着舰控制的最优引导 律。在智能控制领域内,张杨等^[14-15]提出自适应反 演控制方法,解决干扰影响下模型参数不确定性的 控制问题。此外,自适应模糊控制、H_a鲁棒控制等 方法也用于舰尾流抑制^[16-17]。

改变作动器执行方式也有较多探索,如基于 多操纵面的控制分配,实现直接力/力矩控制着 舰^[18-19],通过直接舵面分配,将姿态运动和位置运 动解耦,提高着舰轨迹控制精度和系统响应速 度^[20-21]。此外,推力矢量控制技术也是提高"低动 压"下控制性能的有效方法。但是这些着舰控制 方法对飞机模型重构、操纵面控制分配提出新的 要求^[22]。

鉴于此,本文综合 MPC 算法和线性二次型高 斯(linear quadratic Gaussian, LQG)现代控制技术^[23-24], 开展甲板运动和舰尾流影响下的着舰控制研究。 MPC 算法能够将先验信息引入到当前的控制律设 计中,能够有效地处理轨迹跟踪过程中的控制饱和 问题。

LQG则是一种综合设计方法,基于卡尔曼滤波 的全维状态观测的 LQR 最优控制技术来实现。本 文对舰尾流扰动抑制,采用前馈控制方法进行补偿 抵消,原因是舰尾流是关于航母运动和自然风场的 函数,对于给定海况和航母运动,这种固连于航母 尾部的空间扰动风场可以用解析函数来描述。因 此,对舰尾流的前馈补偿,将其等效为随舰-机相对 位置的附加迎角扰动,将其作为扰动输入,故可以 在控制上进行抵消补偿。

1 舰载机着舰模型

1.1 飞机纵向运动模型

着舰运动为非线性模型,由上面分析,可将其 简化为围绕进入着舰窗口的配平状态下小扰动线 性模型,忽略横航向通道的耦合影响,其舰载机着 舰过程如图1所示。数学模型为

 $\begin{cases} \dot{\boldsymbol{x}} = \boldsymbol{A}_{\mathrm{M}}\boldsymbol{x} + \boldsymbol{B}_{\mathrm{M}}\boldsymbol{u} + \boldsymbol{E}_{\mathrm{M}}\alpha_{\mathrm{g}} \\ \boldsymbol{x} = [\Delta \nu/V_{0}, \Delta \alpha, \Delta \theta, \Delta q, \Delta h/V_{0}]^{\mathrm{T}} \\ \boldsymbol{u} = [\Delta \delta_{\mathrm{e}}, \Delta \delta_{\mathrm{Lef}}, \Delta \delta_{\mathrm{T}}]^{\mathrm{T}} \end{cases}$ (1)

式中: Δv 、 $\Delta \alpha$ 、 $\Delta \theta$ 、 Δq 和 Δh 分别表示舰载机相对 于配平状态点的增量的归一化速度、迎角、俯仰



图 1 由进近窗进入下滑着舰阶段

Fig. 1 Transition from approach phase into glide landing phase

角、俯仰角速度和高度; $\Delta \delta_{e}$ 、 $\Delta \delta_{Lef}$ 和 $\Delta \delta_{T}$ 分别为升 降舵偏角、前缘襟翼偏角和油门控制输入相对于标 称状态的偏差; α_{g} 为垂直风扰动引起的迎角增量, 也称为舰尾流对着舰影响。

1.2 甲板运动及理想着舰点

在典型海况下,航母在涌浪作用下会产生六自 由度运动,但比较重要的是纵摇、横摇、航向及垂 荡运动,其模型可描述为

 $\begin{cases} \theta_{\rm S} = [0.5\sin(0.6t) + 0.3\sin(0.63t) + 0.25]/57.3\\ \phi_{\rm S} = [2.5\sin(0.5t) + 3\sin(0.52t) + 0.5]/57.3\\ \psi_{\rm S} = [0.25\sin(0.7t) + 0.5\sin(0.1t)]/57.3\\ Z_{\rm S} = [4\sin(0.6t) + \sin(0.2t)] \times 0.304 \text{ 8} \end{cases}$ (2)

式中: θ_s、φ_s、ψ_s及Z_s分别为相对于航母重心的纵 摇角、横摇角、航向角及垂荡位移的变化情况。甲 板运动形式通常由白噪声信息来模拟,60 s内运动 信息如图 2 所示。

舰载机理想着舰点(ideal touchdown point, ITP) 位于着舰中跑道中线 2、3 道拦阻索中间,根据航母 运动及着舰点与重心相对关系,如图 3 所示, ITP 运 动描述如下:

$$\begin{cases} \Delta L_{\text{TTP/S}} = -Z_{\text{S}} \sin \theta_{\text{S}} \\ \Delta Y_{\text{TTP/S}} = Z_{\text{S}} \sin \phi_{\text{S}} \\ \Delta Z_{\text{TTP/S}} = Z_{\text{S}} - L_{\text{TTP/S}} \sin \theta_{\text{S}} - Y_{\text{TTP/S}} \sin \phi_{\text{S}} \end{cases}$$
(3)

式中: ΔL_{ITP/S}、ΔY_{ITP/S}、ΔZ_{ITP/S}分别为由于图 3 中甲 板运动影响,母舰重心与 ITP 纵向、横向、垂向偏 差变化情况,式中描述了甲板运动所引起的运动模 型。图 4 则对 ITP 随时间变化进行了模拟。

1.3 着舰引导轨迹设计

由于航母甲板运动,舰载机着舰过程的实际航 迹如图 5(a) 所示,其中 $V_{a/e}$ 、 $V_{s/e}$ 、 $V_{a/s}$ 分别为飞机、 航母相对大地坐标系的速度和飞机相对于航母的 速度。图 5(b) 中,A 点表示舰载机当前时刻所处的 位置, h_e 为舰载机当前时刻距离理想位置 A_0 的高度 偏差, B_0 为 ΔT 时刻后舰载机的期望位置,弧线 A_0B_0 为通过甲板运动预测获得的 ITP 轨迹,弧线 AB为弧线 A_0B_0 经平移后得到的参考轨迹。弧线



Fig. 2 Simulation of deck motion information



图 3 甲板运动信息的模拟 Fig. 3 Simulation of deck motion data

*AB*为机舰相对运动的理想下滑道(下滑角为γ₀,飞 机相对着舰点速度 $V_{a/itd}$ 、位置 $P_{a/itd}$)。弧线 A_0B_0 为飞机根据甲板运动预测的引导轨迹,其会随甲 板运动变化进行预测调整,作为飞机的轨迹跟踪 指令。

本文需要专门设计跟踪指令轨迹(引导轨迹),如图 5(b)中的红色弧线 *AB*₀。原因是如果直接采用弧线 *A*₀*B*₀作为引导轨迹,那么在控制初始阶段,高度指令与实际高度的偏差会很大(线段 *AA*₀),这么大的初始误差会让预测控制的指标函数过大,也让指标函数中初始响应的信息占比过大,也会让控制器的增益在初始阶段增益过大,稳

定性下降。引导轨迹(弧 AB_0)上的点满足 $h_t = h_e l_1/(l_1 + l_2)$ 关系式,其中, h_t 表示跟踪理想轨迹的高度 误差修正指令, h_e 表示因为甲板运动导致的理想 高度变化补偿指令。此时,引导轨迹中就同时包 含了理想轨迹的形状信息及当前高度误差的修正 信息。

为提高跟踪指令轨迹在不同条件下的适应能力,引入了调节因子 K_{α} 和 K_{β} ,设计跟踪指令轨迹 (弧线 AB_0)相对于 \overline{AB} 的高度差为

$$\Delta h = K_{\alpha}h_{\rm s} + K_{\beta}h_{\rm e}\frac{l_1}{l_1 + l_2} \tag{4}$$

当Ka和KB的取值均为1时,引导轨迹如图5中







的弧线AB₀所示。

1.4 舰尾流场模型

依据军用规范 MIL-F-8785C^[25], 舰尾流场可以 视作随机自由大气扰流、舰尾流稳态分量、周期性 分量及随机扰动分量 4 部分的合成, 即

$$\begin{cases} u_g = u_1 + u_2 + u_3 + u_4 \\ v_g = v_1 + v_4 \\ w_g = w_1 + w_2 + w_3 + w_4 \end{cases}$$
(5)

式中: ug、vg和wg分别为舰尾流的水平纵向分量、 水平横向分量和垂直分量;下标1、2、3、4分别表 示随机自由大气扰流,舰尾流的稳态分量,周期性 分量及随机扰动分量。



图 5 着舰引导轨迹设计示意图

Fig. 5 Schematic diagram of guidance trajectory design

2 MPC-LQG 轨迹精确跟踪算法

2.1 LQG 跟踪控制模型

假设系统状方程、测量方态程和输出方程为

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}_{u}\mathbf{u}(t) + \mathbf{B}_{w}\mathbf{w}(t) \\ \mathbf{y}(t) &= \mathbf{C}_{y}\mathbf{x}(t) + \mathbf{v}(t) \\ \mathbf{z}(t) &= \mathbf{C}_{z}\mathbf{x}(t) \\ \mathbf{B}_{w}\mathbf{w}(t) &= \mathbf{E}_{M}\alpha_{g} \end{aligned}$$
(6)

式中: x(t)为状态变量; u(t)为系统输入; w(t)为舰尾 流扰动 (见 1.4节); y(t)为测量输出; z(t)为系统输 出,表示舰载机飞行轨迹; ag为尾流变化等效迎角; v(t)表示测量噪声,为零均值高斯白噪声。这是由 于飞机反馈信号的量测噪声一般具有平稳随机性 质,与白噪声信号很相似。A, Bu, Cy, EM分别为状 态,输入,输出, 扰流转换矩阵。

设理想输出z_c(t)为飞机理想飞行轨迹,那么跟踪误差e(t)可表示为

$$\boldsymbol{e}(t) = \boldsymbol{z}_{\mathrm{c}}(t) - \boldsymbol{z}(t) \tag{7}$$

LQG 跟踪控制是最优控制的改进,方法是通过 搜索输入变量序列,使得性能指标函数J最小。为 确保系统在较少控制输入能量下能够稳定跟踪输 出,指标函数选取为

$$J = \frac{1}{2} \int_{t_0}^{t_f} [\boldsymbol{e}^{\mathrm{T}}(t)\boldsymbol{Q}\boldsymbol{e}(t) + \boldsymbol{u}^{\mathrm{T}}(t)\boldsymbol{R}\boldsymbol{u}(t)] \mathrm{d}t + \frac{1}{2}\boldsymbol{e}^{\mathrm{T}}(t_f)\boldsymbol{Q}_{\mathrm{f}}\boldsymbol{e}(t_f)$$
(8)

式中:Q、Qf和R为时变的加权矩阵,方便计算时间

片断范围内跟踪误差量e(t)、输入量u(t)和终端误差 $e(t_f)$ 的控制指标值;Q、 Q_f 为半正定矩阵;R为正定 矩阵; $e(t_f)$ 为终端跟踪误差。

由于*J*为一个满足系统状态方程的条件极值问题,对于此类问题求解,可引入待定的 Lagrange 乘 子向量 $\lambda^{T}(t) = [\lambda_{1}(t), \lambda_{2}(t), \cdots, \lambda_{n}(t)]$ 将*J*指标函数进行 松弛。可得到如下增广的性能指标问题:

$$\min J_A = \frac{1}{2} \int_{t_0}^{t_r} [\boldsymbol{e}^{\mathsf{T}}(t) \boldsymbol{Q} \boldsymbol{e}(t) + \boldsymbol{u}^{\mathsf{T}}(t) \boldsymbol{R} \boldsymbol{u}(t)] dt + \int_{t_0}^{t_r} \boldsymbol{\lambda}^{\mathsf{T}}(t) [\boldsymbol{A} \boldsymbol{x}(t) + \boldsymbol{B}_u \boldsymbol{u}(t) - \dot{\boldsymbol{x}}(t)] dt + \frac{1}{2} \boldsymbol{e}^{\mathsf{T}}(t) \boldsymbol{Q}_{\mathsf{I}} \boldsymbol{e}(t)$$

s.t.
$$\begin{cases} \mathbf{x}(t) = A\mathbf{x}(t) + B_{u}\mathbf{u}(t) + E_{M}\alpha_{g} \\ \mathbf{x}^{\min}(t) \leq \mathbf{x}(t) \leq \mathbf{x}^{\max}(t), \quad \mathbf{u}^{\min}(t) \leq \mathbf{u}(t) \leq \mathbf{u}^{\max}(t) \end{cases}$$
(9)

式中:LQ 跟踪控制输入u的求解为

$$\boldsymbol{u}(t) = -\boldsymbol{R}^{-1}\boldsymbol{B}_{\boldsymbol{u}}^{\mathrm{T}}[\boldsymbol{P}(t)\boldsymbol{x}(t) - \boldsymbol{g}(t)]$$
(10)

式中: *P*(*t*)、*g*(*t*)分别为式 (11)、式 (12) 的 Riccati 方 程的解,其终端条件由式 (13) 和式 (14) 表示:

$$\dot{\boldsymbol{P}}(t) = -\boldsymbol{P}(t)\boldsymbol{A} - \boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}(t) + \boldsymbol{P}(t)\boldsymbol{B}_{u}\boldsymbol{R}^{-1}\boldsymbol{B}_{u}^{\mathrm{T}}\boldsymbol{P}(t) - \boldsymbol{C}_{z}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{C}_{z}$$
(11)

$$\dot{\boldsymbol{g}}(t) = [\boldsymbol{P}(t)\boldsymbol{B}_{u}\boldsymbol{R}^{-1}\boldsymbol{B}_{u}^{\mathrm{T}} - \boldsymbol{A}^{\mathrm{T}}]\boldsymbol{g}(t) - \boldsymbol{C}_{z}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{z}_{\mathrm{c}}(t) \qquad (12)$$

$$\boldsymbol{P}(t_{\rm f}) = \boldsymbol{C}_z^{\rm T} \boldsymbol{Q}_{\rm f} \boldsymbol{C}_z \tag{13}$$

$$\boldsymbol{g}(t_{\rm f}) = \boldsymbol{C}_z^{\rm T} \boldsymbol{Q}_{\rm f} \boldsymbol{z}_{\rm c}(t) \tag{14}$$

综合式 (10)~式 (14),便可实现给定时间窗口 内的时变 LQG 跟踪控制。

2.2 MPC 预测流程

MPC 预测序列是基于甲板运动和舰尾流预测 信息实现的,如图 6 所示。







图 6 中符号及工作过程如下: $x_e(t)$ 为t时刻系统 期望的状态; $x^*(t)$ 为t时刻系统的实际状态; 控制 $\Pi(n)$,在区间 $\Lambda_n = [t_0 + n\delta, t_0 + n\delta + T]$ 内,由初始条 件 $x(t_0 + n\delta) = x^*(t_0 + n\delta)$ 及区间 Λ_n 内的期望轨迹 $x_e^n(t)$,采用式(1)来求解,经过控制 $\Pi(n)$ 计算可 获得 $[t_0 + n\delta, t_0 + n\delta + T]$ 时间段内的最优控制律。 控制量 $\Upsilon(n)$,以 $u_{opt}(t_0 + n\delta)$ 作为实际控制输入, $x^*(t_0 + n\delta)$ 为初始条件, 在 $[t_0 + n\delta, t_0 + (n+1)\delta]$ 时间 区间内采用式 (9)求解 $x^*(t_0 + (n+1)\delta)$; 控制量 $\Upsilon(n)$ 可实时模拟舰尾流及输出量测噪声等扰动对 着舰控制影响。

2.3 MPC-LQG 精确跟踪控制算法

1) LQG 求解策略。对于给定时间窗 A_n,时变 LQG 跟踪控制的关键是对 Riccati 方程求解,而其 求解效率和精度则直接影响算法的性能与控制精 度,且会降低后续 MPC 预测控制效果。本文采用 精细积分方法^[23]进行精确求解。

2) 基于自回归 (autoregression, AR) 预测算法进行甲板运动预测。着舰引导控制时,应提前获得甲板运动的预测信息。为提高算法效率,本文采用 p阶 AR 预测算法对甲板运动进行预测, AR 预测模型简写为^[3]

$$\xi_t = a_1 \xi_{t-1} + a_2 \xi_{t-2} + \dots + a_p \xi_{t-p} + \varsigma_t \tag{15}$$

式中: $\xi_{i-i}|i=0,1,2,\cdots,p$ 为待预测变量 ξ 的历史数据; $a_i|i=1,\cdots,p$ 为待定参数; ς_i 为白噪声。基于N个历史数据采用最小二乘法对待定参数 a_i 进行最优估计,以使目标函数最小化,记为 Z_{cost} :

$$\begin{cases} Z_{\text{cost}} = (\boldsymbol{\xi}_N - \boldsymbol{A}_N \boldsymbol{\hat{a}}_N)^{\mathrm{T}} (\boldsymbol{\xi}_N - \boldsymbol{A}_N \boldsymbol{\hat{a}}_N) \\ \boldsymbol{\hat{a}}_N = (\boldsymbol{A}_N^{\mathrm{T}} \boldsymbol{A}_N)^{-1} \boldsymbol{A}_N^{\mathrm{T}} \boldsymbol{\xi}_N \end{cases}$$
(16)

AR 预测算法通过式(17)对未来的l个数据进行 预测。

$$\hat{\xi}_{N+l} = \begin{cases} \sum_{j=1}^{p} \hat{a}_{j} \xi_{N+l-j} l = 1\\ \sum_{j=1}^{l} \hat{a}_{j} \xi_{N+l-j} + \sum_{j=l+1}^{p} \hat{a}_{j} \hat{\xi}_{N+l-j} l < l \le p \\ \sum_{j=1}^{p} \hat{a}_{j} \hat{\xi}_{N+l-j} l > p \end{cases}$$
(17)

3)终端条件判定。飞机在 LQG 控制下跟踪 MPC 预测的引导轨迹, 机舰下滑道如图 7 所示, 满 足飞机通过安全通过舰尾, 且在相对 ITP 位置近似 为 0 等条件时, 判定为轨迹终端条件:

$$\begin{cases} \left\| \boldsymbol{P}_{a/itd} \right\| \leq \varepsilon \quad \varepsilon > 0 \\ \boldsymbol{P}_{a/s}(x) > \boldsymbol{P}_{r/s}(x) \end{cases}$$
(18)



图 7 轨迹跟踪着舰终端示意图 Fig. 7 Schematic diagram of landing terminal for trajectory tracking

式中: $P_{a/itd}$ 为飞机相对 ITP 位置; $P_{a/s}$ 、 $P_{r/s}$ 分别为飞机、舰尾相对航母牵连坐标系 $O_s x_s y_s z_s$ 的位置; *(x)(如 $P_{a/s}(x)$)为位置在x方向数值, ε 为非常小量,其大小正比于仿真步长 Δt 。

4)精确跟踪算法步骤。在 MPC 预测时窗内, 由 AR 预测算法对甲板/ITP 运动进行在线估计,并 将 ITP 预测信息加入到着舰跟踪指令轨迹中,后采 用精细积分法求解该时窗内时变 LQG 问题的控制 输入u_{opt}(t₀+nδ),进行系统方程的下一步迭代计算, 得到下一时间的系统状态 $\mathbf{x}^*(t_0 + (n+1)\delta)$ 。

此时, MPC 预测时窗后移一个时间步, 以上面 系统状态作为初始状态, 重复上述 LQG 计算过程。 通过一步步时间窗移动, 直至满足轨迹终止条件结 束着舰精确跟踪过程。MPC-LQG 跟踪控制算法的 流程如图 8 所示。需要说明的是, 算法设计过程中, 仿真步长δ的选择应尽可能小, 否则跟踪控制算法 误差会很大, 但是步长选取过小, 算法效率下降导 致引起计算延时。因此, 迭代计算时应选择合适步长。



图 8 跟踪控制算法流程 Fig. 8 Flow of tracking control algorithm

3 仿真实验

3.1 初始条件设置

基本着舰参数: 航母速度为 15.4 m/s, 舰载机地 速为 70 m/s, 初始位置距 ITP 为 1386 m, 相对高度 h_0 =85 m, 基准下滑角 γ_c = -3.5°, 参考迎角 α_0 =8.3°, 可知参考高度 $h_c = h_0 + \int |V_{a/s}| \sin(\gamma_c) dt$ 。纵向运动 小扰动模型 A_M, B_M, E_M 参考文献 [12], 飞机作动器 的操控量的配平值及约束范围, 如表 1 所示。

实验环境为典型海况,甲板运动和舰尾流扰动 参照1.1节着舰模型。舰尾流扰动模型参考文献[3]

表1 操纵量约束情况

 Table 1
 Constraints of carrier aircraft control system

控制变量	变量范围/(°)	配平值/(°)	速率范围/((°)·s ⁻¹)	
δ_{e}	[-24, 10.5]	-11.9	40	
δ_{Lef}	[-3, 33]	+17.6	15	
δ_T	[0, 57.3]	14.6	32	

和军用规范 MIL-F-8785C^[25],着舰中飞机所受舰尾 流相对大地坐标系的模拟结果如图 9 所示,模拟结 果基于上面仿真基本参数在中度海况时完成。在 控制器跟踪仿真中,全程加载舰尾流扰动到飞机速 度,以此检验控制算法抑制舰尾流扰动影响。

3.2 实验结果与分析

基于 3.1 节初始条件,采用 MPC-LQG 精确跟 踪算法对着舰纵向控制进行仿真。

图 10 为仿真中 MPC 滚动预测的时间窗 A设定 为1.5 s, 迭代步长为0.01 s, 在着舰前15 s 引入甲板/ITP 运动的 AR 估计与补偿模块, 舰尾流及量测噪声也 嵌入到迭代算法。着舰跟踪控制效果如图 10 所 示, 图 10(a)为高度跟踪期望航迹情况, 其中绿线 为 ITP 运动情况, 在引入 ITP 运动信息后偏离基准 高度的变化情况。可以看出, 本文算法对 ITP 运动 的轨迹跟踪效果显著, 能够精准跟踪甲板运动, 可 以满足甲板运动的频率响应要求, 触舰时跟踪误差





小约为 0.1~0.2 m, 是跟踪的相角滞后导引的。如 果以该跟踪误差触舰, 着舰点与 ITP 的散布值 Δh_{track}/tan 3.5, 其范围为±1.63~±4.08, 显然该触舰 范围仍在第 2 道和第 3 道拦阻锁之间, 着舰效果显 著。同样, 引入甲板运动补偿时, 系统初始偏差较 大, 但控制算法能迅速跟踪上 ITP 运动补偿控制 指令。

着舰控制中,其他状态量、控制输入量的变量 如图 11 所示,从仿真结果看,控制输入量为小扰动 量基本没有突破操纵量约束边界,状态变量同样也 为小扰动量,其中除发动机推力输入Δδ_T在甲板运 动补偿后出现了边界输入。加入甲板运动补偿,使 得与原始跟踪指令发生较大偏差,系统想要快速跟 踪则需要较快的推力改变。从仿真结果看,本文算 法可精确地跟踪甲板运动时的引导轨迹。



3.3 不同条件的仿真分析

3.3.1 不同初始高度的仿真比较与分析

进入下滑道时,飞机与航母距离相同的情况 下,其距离为1065m,实验一开始便引入甲板运动 信息,研究不同初始高度的控制效果,其他参数设 置同3.1节。

仿真时全程引入甲板运动预测,其 MPC-LQG 轨迹跟踪控制效果如图 12 所示,可以看出,对于不 同初始高度条件,本文方法均可以较好地跟踪理想 引导轨迹,实现精准着舰控制。对于初始高度偏差 不太情况下, MPC-LQG 能够在 2~3 s 完成偏差修 正,然后精准跟踪甲板运动。从图中的跟踪情况 看,不同初始高度的跟踪曲线在修正后几乎贴合在 一块,即初始高度偏差对最终着舰跟踪控制几乎没 有影响。

3.3.2 不同方法的比较与分析

本文方法与传统 LQG、极点配置控制方法进 行着舰控制比较,仿真条件如 3.3.1 节,初始高度选 择为 60 m,实验同样直接加入甲板运动信息,2 种 方法的着舰轨迹和下滑道跟踪误差随时间变化情



Fig. 11 Changes in state and input quantities during landing process

况,如图 13 所示。从图 13 可以看出,增加运动信息预测的补偿控制比不增加运动预测的控制,指令

跟踪效果更好。从图 13 (a)可以看出,3种方法在 初始偏差、引入甲板运动信息后,表现出跟踪超调





Fig. 12 Landing trajectories and their deviation values at different initial heights





与振荡,显然 MPC 方法的响应速度、超调量均好于 未增加甲板运动预测的 LQG 控制、极点配置方 法。本文典型海况下甲板运动的主体角频率是 0.4~0.6 rad/s^[26],属于一个中低频指令,甲板运动信 息的加入可以理解为引入了一个阶跃输入信号。 从图 13(b)可看出,MPC 方法对高度跟踪误差的响 应情况除一开始有一个大的偏差外,总体上要优于 其他 2 种方法。其原因是控制律当中增加的运动 预测信息,可理解为在控制系统中引入了速度反 馈,预先考虑到下一步的指令信息。此外,到了着 舰末段,本文方法的跟踪误差约 0.15 m,即触舰点 在理想点后 2.4 m,处于第 2 道和第 3 道拦阻索范围 内,而 LQG 方法触舰点在理想点后 11.4 m,在第 3 道和第 4 道拦阻索之间,着舰品质不如本文方法。

造成上述现象的原因如下:① MPC-LQG 算法 加入了预测信息,所以,系统的整体带宽显著增加, 提高了系统反应的快速性和精确性;②MPC-LQG 算法采用了迭代优化的方法,充分考虑了当前和未 来一段时间的情况,而传统 LQG 算法在无限时间 内优化,极点配置在反馈控制过程中甲板随机运动 的跟踪效果不佳,相比这 2 种算法, MPC-LQG 算法 的抗干扰性和鲁棒性会明显提高。

本文航母运动预测是由 AR 预测算法完成的, 基于滚动更新方式通过历史数据的自回归方法预 测未来几秒内运动信息,预测信息较为精准,为着 舰控制器预先给出的控制律应该考虑的未知量,较 好地提高了对甲板运动的补偿效果。

4 结 论

本文基于 MPC-LQG 算法对舰载机着舰的纵向 运动进行了轨迹跟踪控制研究,着舰模型中考虑了 甲板运动、舰尾流影响,对某舰载机典型操纵模式 进行了着舰控制仿真研究,结论如下:

1)本文方法可较好地完成贴近真实环境下的 着舰轨迹跟踪控制, 触舰时跟踪高度偏低 0.1~0.2 m, 即触舰点在理想点之后 1.6~3.3 m, 处于第 2 道和 第 3 道拦阻索之间, 预测控制由于超前效应, 在着 舰末段整体轨迹是压低来控制的。这样的策略, 在 舰尾流幅度增大时, 可以较好的抑制上流风影响。

2) 基于 AR 预测算法对 ITP 运动预测结果可 靠,保证了着舰控制的引导律的精准。

3)不同条件的仿真表明,本文方法对初始高度 偏差的修正控制能力强,增加 MPC 预测控制策略, 比传统 LQG 算法控制的超调量小、响应速度快,在 着舰区轨迹跟踪的精度高。本文方法达到了着舰 标准的 III 类级别。

参考文献(References)

[1] 甄子洋, 王新华, 江驹, 等. 舰载机自动着舰引导与控制研究进展
 [J]. 航空学报, 2017, 38(2): 020435.

ZHEN Z Y, WANG X H, JIANG J, et al. Research progress in guidance and control of automatic carrier landing of carrier-based aircraft[J]. Acta Aeronautica et Astronautica Sinica, 2017, 38(2): 020435(in Chinese).

- [2] 段卓毅, 王伟, 耿建中, 等. 舰载机人工进场着舰精确轨迹控制技术[J]. 航空学报, 2019, 40(4): 622328.
 DUAN Z Y, WANG W, GENG J Z, et al. Precision trajectory manual control technologies for carrier-based aircraft approaching and landing[J]. Acta Aeronautica et Astronautica Sinica, 2019, 40(4): 622328(in Chinese).
- [3] 江驹, 王新华, 甄子洋, 等. 舰载机起飞着舰引导与控制[M]. 北京: 科学出版社, 2019.
 JIANG J, WANG X H, ZHEN Z Y, et al. Carrier aircraft takeoff and landing guidance and control[M]. Beijing: Science Press,

[4] MENG Y, WANG W, HAN H, et al. A visual/inertial integrated

- landing guidance method for UAV landing on the ship[J]. Aerospace Science and Technology, 2019, 85(2): 474-80.
- [5] WILKINSON C, FINDLAY D, NICHOLS J, et al. Shipboard aircraft simulation with ship-relative navigation sensor modeling[C]//Proceedings of the 54th AIAA Aerospace Sciences Meeting. Reston: AIAA, 2016: 1769.
- [6] DING Z, LI K, MENG Y, et al. FLIR/INS/RA Integrated Landing Guidance for Landing on Aircraft Carrier[J]. International Journal of Advanced Robotic Systems, 2015, 12(5): 1-9.
- [7] SEDLMAIR N, THEIS J, THIELECKE F. Flight testing automatic landing control for unmanned aircraft including curved approaches[J]. Journal of Guidance, Control, and Dynamics, 2021, 11: 1-36.
- [8] GUAN Z, LIU H, ZHENG Z, et al. Moving path following with integrated direct lift control for carrier landing[J]. Aerospace Science and Technology, 2022, 120: 107247.
- [9] DENG Y, DUAN H. Control parameter design for automatic carrier landing system via pigeon-inspired optimization[J]. Nonlinear Dynamics, 2016, 85(1): 97-106.
- [10] YANG Z, DUAN H, FAN Y, et al. Automatic carrier landing system multilayer parameter design based on cauchy mutation pigeoninspired optimization[J]. Aerospace Science and Technology, 2018, 79: 518-530.
- [11] ZHU Q, YANG Z. Design of air-wake rejection control for longitudinal automatic carrier landing cyber-physical system[J]. Computers & Electrical Engineering, 2020, 84: 106637.
- [12] WANG L, ZHANG Z, ZHU Q, et al. Longitudinal automatic carrier landing system guidance law using model predictive control with an additional landing risk term[J]. Proceedings of the Institution of Mechanical Engineers, 2019, 233(3): 1089-1105.

- [13] KOO S, KIM S, SUK J. Model predictive control for UAV automatic landing on moving carrier deck with heave motion[J]. IFAC, 2015, 48(5): 59-64.
- [14] 张杨, 吴文海, 汪杰, 等. 舰载无人机横侧向着舰控制律设计[J]. 航空学报, 2017, 38(S1): 721489.
 ZHANG Y, WU W H, WANG J. Design of carrier UAV lateral/ directional landing control law[J]. Acta Aeronautica et Astronautica Sinica, 2017, 38(S1): 721489(in Chinese).
- [15] 吴文海, 张杨, 胡云安, 等. 舰载机着舰非线性反演控制方法研究 进展[J]. 系统工程与电子技术, 2018, 40(7): 1578-1587.
 WU W H, ZHANG Y, HU Y A, et al. Research development in nonlinear backstepping control method of carrier-based aircraft landing[J]. Systems Engineering and Electronics, 2018, 40(7): 1578-1587(in Chinese).
- [16] SUBRAHMANYAM M B. H-infinity design of F/A-18A automatic carrier landing system[J]. Journal of Guidance, Control, and Dynamics, 1994, 17(1): 187-91.
- [17] HANSON C, SCHAEFER J, JOHNSON M. Design of low complexity model reference adaptive controllers[J]. NASA/TM, 2012, 215972.
- [18] LORENZETTI R C, NELSEN G L, JOHNSON R W. Direct lift control for approach and landing[J]. Journal of Aircraft, 1969, 6(3): 240-244.
- [19] NASTASI R, MARTORELLA P, HUFF R, et al. Carrier landing simulation results of precision flight path controllers in manual and automatic approach[C]//Proceedings of the 10th Atmospheric Flight Mechanics Conference. 1983: 2072.
- [20] DENHAM J W. Project MAGIC CARPET: Advanced Controls and Displays for Precision Carrier Landings[C]//Proceedings of the 54th AIAA Aerospace Sciences Meeting. Reston: AIAA, 2016, 1770.
- [21] 张志冰,张秀林,王家兴,等.一种基于多操纵面控制分配的 IDLC人工着舰精确控制方法[J]. 航空学报, 2021, 42(8): 142-157. ZHANG Z B, ZHANG X L, WANG J X, et al. An IDLC landing control method of carrier-based aircraft based on control allocation of multiple control surfaces[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(8): 142-157(in Chinese).
- [22] DURHAM W, BORDIGNON K, BECK R. Aircraft control allocation[M]. Chichester: John Wiley and Sons Ltd, 2016.
- [23] 高强, 谭述君, 钟万總. 精细积分方法研究综述[J]. 中国科学:技术 科学, 2016, 12: 5-16. GAO Q, TAN S J, ZHONG W X. A survey of the precise integration method[J]. Scientia Sinica Technologica, 2016, 12: 5-16(in Chinese).
- [24] STEVENS B, LEWIS F, JOHNSON E N. Aircraft control and simulation: Dynamics, controls design, and autonomous systems[M].3rh ed. Chichester: John Wiley and Sons Ltd, 2017.
- [25] Flying qualities of piloted aircraft:mil-std-1797. (Revision B Notice 1-Administrative)[S]. Washington D.C.: United States Department of Defense, 2012.
- [26] DURAND T S, WASICKO R J. Factors influencing glide path control in carrier landing[J]. Journal of Aircraft, 1971, 4(2): 146-158.

A precise landing control method based on model predictive control algorithm

WAN Bing, SU Xichao^{*}, WANG Jie, HAN Wei, LI Changjiu

(Naval Aviation University, Yantai 264001, China)

Abstract: This paper proposes a precise landing control method based on model predictive control-linear quadratic gaussian (MPC-LQG) to address the influences of carrier motion and wake turbulence during the landing process of carrier-based aircraft. The method is applied to research precise landing control in the longitudinal channel of the aircraft. The core idea of integrating model predictive control (MPC) with the linear quadratic gaussian (LQG) algorithm for trajectory tracking control is as follows: MPC is utilized for compensating the motion of the aircraft carrier; A full-dimensional state observer is designed to achieve full state feedback, thereby realizing optimal landing control. Finally, the simulation of the algorithm is carried out for a typical landing problem, and the landing control results with different initial conditions and algorithms are compared and analyzed. The simulation shows that the algorithm proposed has a good trajectory tracking effect, with the altitude deviation at ITP being 0.1~0.2m.

Keywords: model predictive control-linear quadratic Gaussian; deck motion; guidance command; carrier landing control; trajectory tracking

URL: link.cnki.net/urlid/11.2625.V.20220808.1654.002

Received: 2022-05-18; Accepted: 2022-07-17; Published Online: 2022-08-09 09:24

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

DOI: 10.13700/j.bh.1001-5965.2022.0466

双余度机电作动系统动态力均衡控制方法

孙晓哲1,吴江2,*,石林轩1,杨建忠1

(1. 中国民航大学 安全科学与工程学院,天津 300300; 2. 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: 飞机多电/全电化技术的发展使得余度机电作动系统广泛应用于飞控舵面作动系统 中,但其在主动/主动工作模式下由于作动器输出不同步而造成的力纷争问题仍需解决。针对该问 题,建立了系统完整的线性数学模型,分析了动态力纷争的来源和机理,研究并提出了基于速度和 加速度前馈补偿控制、基于力差值反馈的 PID 控制联合作用的动态力均衡控制方法,并对其均衡能 力及鲁棒性进行验证,结论得出所提方法不仅可以有效地减弱摩擦、间隙和指令延迟这3个因素造 成的动态力纷争,还对系统中各种参数的扰动有一定的鲁棒性。

关 键 词: 机电作动系统; 动态力纷争; 前馈补偿控制; 力差值反馈; 动态力均衡控制 中图分类号: V232

文献标志码: A 文章编号: 1001-5965(2024)04-1208-11

机电作动器(electro-mechanical actuator, EMA) 是飞机多电/全电化发展背景下的新应用趋势,这类 作动器消除了液压管道,显著提高了作动性能,降低 了系统的复杂性、重量和体积,并且有利于系统维修。

运输类飞机适航标准 25.671(c) 和 25.1309(b)(1)^[1] 款对飞控系统的安全性和可靠性提出极高的要求, 因此,主控舵面一般采用两余度或三余度作动器 (用于方向舵)并行驱动。然而,目前机电作动系统 大部分还只应用在辅助飞控舵面上,如波音787的 扰流板和水平安定面及 C919 的水平安定面^[2]。功 率驱动电子和电机技术的发展提高了 EMA 的功重 比,将EMA引入到主飞控作动系统具有很高的实 用价值。国内学者开展了针对主/主式双余度机电 作动系统的研究和试验工作,如 Actuation 2015、 Clean Sky2^[3-4]和 Corac-Genome^[5]等研究项目,致力 于提高 EMA 应用的技术成熟度并针对主/主式机 电作动系统积极开展飞行试验。主/主式 EMA 刚 性连接到同一舵面,由于传感器、作动器和舵面的 制造和安装误差积累等因素导致了余度作动器的 输出不一致,这就引起了同一舵面多个作动器间的

力纷争现象。力纷争会影响作动系统的性能,引起 飞机舵面结构疲劳或破坏、降低作动器有效载荷进 而影响飞行安全。适航局方在开展飞控系统适航 审查时力纷争是审查重点关注项目之一。

传统研究基本上把力纷争问题分为静态力纷 争和动态力纷争,并分别设计力均衡控制方法。相 对于传统液压作动器,EMA具有不同的物理特性, 如高刚度、非线性和强耦合性,因此,需开展针对主/ 主式机电作动系统力纷争问题的分析和力均衡方 法研究。

余度作动系统中力纷争的大小取决于各个作动器位置输出的偏差及连接刚度的大小, $\Delta F = k_{\text{stiffness}}\Delta x$ 为 EMA 中力纷争的决定式,本文假设 2个通道的连接刚度相等。因此,对于 EMA 这种高刚度的系统来说,需要更高精度的位置控制。

力纷争分为静态力纷争和动态力纷争^[6-7],静态 力纷争指系统到达稳态后各作动器输出力持续不 一致的现象,而动态力纷争指系统到达稳态之前的 瞬态过程中作动器输出力不一致的现象。静态力 纷争主要由 2 个作动器达到稳态时的位置差异导

收稿日期: 2022-06-08; 录用日期: 2022-10-10; 网络出版时间: 2022-11-15 14:53 网络出版地址: link.cnki.net/urlid/11.2625.V.20221114.1519.003

基金项目:中央高校基本科研业务费中国民航大学专项(3122020066)

*通信作者. E-mail: wujiang@buaa.edu.cn

引用格式: 孙晓哲, 吴江, 石林轩, 等. 双余度机电作动系统动态力均衡控制方法 [J]. 北京航空航天大学学报, 2024, 50 (4): 1208-1218. SUN X Z, WU J, SHI L X, et al. Dynamic force equalization for dual redundancy electro-mechanical actuation system [J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50 (4): 1208-1218 (in Chinese).

致,大部分都可以通过消除位置的静态误差来解 决^[3]。但是针对动态力纷争问题,就需要让余度作 动器同步运行,这意味着在控制时需要令余度作动 器的速度甚至加速度一致。文献 [3,8-9] 针对非相 似余度的混合作动系统的力纷争问题设计了相应 的力均衡控制方法,利用力差值的反馈信号来改变 作动器的输出位置,这对静态力纷争有良好的均衡 效果,然而对系统产生的动态力纷争峰值抑制效果 并不理想,存在小幅值的振荡,且快速性也有待提 高,显然针对静态力纷争的均衡技术并不足以解决 动态力纷争问题。文献 [10] 中研究得出 EMA 线性 二次型调节器(Linear quadratic regulator, LQR)的 4种引起力纷争的因素,分别是:位置传感器反馈参 数偏置、指令延迟、间隙、摩擦,其中,指令延迟、 间隙、摩擦3种因素主要造成的是动态力纷争。基 于这些因素, 文献 [11] 提出线性二次型调节器 (Linear quadratic regulator, LQR)反馈+神经网络前 馈的控制方法解决力纷争问题,但是其控制算法复 杂,均衡速度较慢。文献 [12-13] 基于混合式作动 系统的线性模型,设计了轨迹发生器+前馈补偿器、 力差值反馈的 PID 补偿器的动态力均衡控制策略, 并在非线性模型中完成了验证。当前研究针对主/ 主式机电作动系统力纷争均衡控制的研究较少[14-15]。

本文主要基于双余度机电作动系统线性模型 设计相应的动态力均衡控制方法,并通过仿真验证 其在摩擦、间隙和指令延迟这些动态力纷争关键影 响因素下的力均衡能力,通过蒙特卡罗方法验证均 衡控制器在系统不确定参数下的鲁棒性。

1 双余度 EMA 建模

1)作动器建模。系统由作动器电子控制装置、无刷直流电机、齿轮减速箱、滚珠丝杠和传感器等部分组成,无刷直流电机通过减速齿轮箱驱动 丝杠,由旋转运动转化为水平运动,EMA 结构如图 1 所示^[16]。

无刷直流电机传递函数模型对于研究和分析 其工作特性及控制器设计都是至关重要的。在定 子绕组两两导通的工作方式下,以a相绕组和b相



绕组导通为例,则有[17]

$$\begin{cases} i_{a} = -i_{b} = i \\ si_{a} = -si_{b} = i \end{cases}$$
(1)

式中: *i*_a和 *i*_b分别为 a 相和 b 相电流, *i* 为电机电流。 则 a、b 相电压为

$$u_{ab} = 2Ri + 2(L - M)si + (e_a - e_b)$$
(2)

式中: *u*_{ab}为 a、b 间相电压; *R*为绕组电阻; *L*和 *M*分 别为绕组的自感和互感; *e*_a和 *e*_b分别为 a 相和 b 相 绕组的反电动势。

本文不考虑换相暂态过程,则有反电动势 $e_a = -e_b$,且由于线电压 u_{ab} 等于直流母线电压 U_d ,则 式(2)可变为

$$U_{\rm d} = r_{\rm a}i + L_{\rm a}si + k_{\rm e}\omega \tag{3}$$

式中: U_d为直流母线电压; r_a绕组线电阻; L_a为绕组 等效线电感; k_e为反电动势系数; ω为电机角速度。

由此得到 EMA 的复频域方程:

$$\begin{cases} U_{d}(s) = r_{a}I(s) + L_{a}sI(s) + k_{e}\omega(s) \\ k_{t}I(s) - T_{L}(s) = J_{m}s\omega(s) + B_{m}\omega(s) \\ T_{L}(s) = F_{e}(s)\frac{1}{i_{g}} \cdot \frac{p}{2\pi} \\ sX(s) = \omega(s)\frac{1}{i_{g}} \cdot \frac{p}{2\pi} \end{cases}$$
(4)

式中: k_t 为电磁转矩系数; T_L 为电机轴的负载转矩; J_m 电机侧转动惯量; B_m 为阻尼系数; F_e 为滚珠丝杠 输出力; i_g 为齿轮减速比;p为滚珠丝杠导程;x为作 动器直线输出位移。

2) 舵面建模。2个 EMA 主/主式运行并联驱动 舵面,共同承担负载,其与舵面的连接处用连接刚 度来表示。该舵面模型体现了作动器通过与舵面 的刚性连接形成的强耦合特点。

$$\begin{cases} k_{L1}(x_1 - x_t) = F_1 \\ k_{L2}(x_2 - x_t) = F_2 \\ F_1 + F_2 - F_L = m_t s^2 x_t \end{cases}$$
(5)

式中: k_{L1}、k_{L2}为各作动器连接刚度; x₁、x₂为各作动器 输出位移; F₁、F₂为各作动器输出力; F_L为舵面的外 部空气负载; x₁为舵面位移; m₁为舵面等效质量。

3) 控制器。本文对 EMA 模拟实验平台进行系 统辨识获取系统参数^[18], 建立仿真模型并按照某商 用飞机的系统性能要求设计了控制器, 该系统采用 PID 三闭环控制, 对应的控制器分别是位置控制器 (automatic position regulator, APR)、转速控制器 (automatic speed regulator, ASR)和电流控制器 (automatic current regulator, ACR)。随后对设计的 控制器仿真验证。系统的性能需要同时满足位置 动态跟随和抗负载干扰 2 方面的性能要求。具体 要求如下^[19]:①动态跟随性能。在频率为 2 Hz 和幅 值为1mm的正弦位置指令下, 舵面幅值放大倍数 大于-3dB, 相位滞后小于45°。本文幅值选用1mm 是为了不让响应进入饱和非线性区域。②抗负载 性能。恒定空气动力载荷扰动下的飞行控制舵面 位置误差要能迅速消除, 并且没有静态误差。 如图 2(a)所示,系统在 2 Hz 的正弦位置指令 下,幅值放大倍数为-0.76 dB,相位延迟 23.6°。 图 2(b)中,1 s 时刻加入 4 000 N(系统最大额定负载)的阶跃负载后,系统在 0.2 s 内回到稳态位置, 并且没有稳态误差,由此验证了系统的性能。





4) 动态力纷争关键因素建模。 已有研究表明^[11], 引起 EMA 产生动态力纷争的 3 个因素分别 是: 摩擦、间隙和指令延迟, 其中摩擦和间隙也体现 了 EMA 非线性的特点, 因此, 本文对提出的动态力 均衡控制方法进行验证时, 将在其中一个通道引入 这些因素,以模拟动态力纷争现象。其中,摩擦模型 使用 Stribeck 模型^[20],间隙模型使用迟滞模型^[21]。 EMA 模型框图如图 3 所示。模型中使用的具体参 数如表 1 所示,表中的摩擦数据与间隙数据均由本 实验室 EMA 平台测得。



图 3 EMAs 完整模型框图

Fig. 3 EMAs complete model block diagram

表1	关键因素模型数值
	7 CAN NE X E

Table 1 Key factor model values

参数	数值
最大静摩擦力矩/(N·m)	0.25
库伦摩擦力矩/(N·m)	0.1
临界速度/(rad·s ⁻¹)	9.6
粘滞摩擦系数	0.001 37
间隙宽度/mm	0.4
延迟时间/ms	10

2 动态力纷争机理分析

研究动态力纷争产生的机理对设计动态力均 衡至关重要。对上述建立的线性部分模型进行简 化,简化后的等效系统框图如图4所示。图4中, G₁(s)、G₂(s)表示作动器的位置跟随性能, H₁(s)、 H₂(s)表示作动器的抗负载扰动性能。

由图可得



$$\begin{cases} x_{r}(s)G_{1}(s) - x_{1}(s) = F_{1}(s)H_{1}(s)G_{1}(s) \\ x_{r}(s)G_{2}(s) - x_{2}(s) = F_{2}(s)H_{2}(s)G_{2}(s) \\ & \text{that (5) } \text{ that (6) } \overrightarrow{\Pi} \cancel{H} \cancel{D} \cancel{M} \cancel{H} \gamma; \\ \gamma = F_{1} - F_{2} = \frac{G_{1}(s)x_{r} - x_{t}}{1 + k_{L1}H_{1}(s)G_{1}(s)}k_{L1} - \frac{G_{2}(s)x_{r} - x_{t}}{1 + k_{L2}H_{2}(s)G_{2}(s)}k_{L2}$$
(7)

由式 (5)~式 (7) 可得
$$\gamma = \frac{Q_1(s)x_r - Q_2(s)F_L}{Q_3(s)}$$
(8)

式中:

$$Q_1(s) = (1 + k_{L1}H_1G_1)(1 + k_{L2}H_2G_2)\{2k_{L1}k_{L2}(G_1 - G_2) + [(1 + k_{L2}H_2G_2)G_1k_{L1} - (1 + k_{L1}H_1G_1)G_2k_{L2}]m_1s^2\}$$

$$Q_2(s) = (1 + k_{L1}H_1G_1)(1 + k_{L2}H_2G_2) \cdot [k_{L2}(1 + k_{L1}H_1G_1) - k_{L1}(1 + k_{L2}H_2G_2)]$$

理想情况下,如果要消除力纷争,就要使得 2个系数项Q₁(s)和Q₂(s)为0,如式(9)所示:

$$\begin{cases} G_2 = G_1 \\ H_2 = \frac{k_{L2} - k_{L1}}{k_{L1}k_{L2}} \cdot \frac{1}{G_1} + H_1 \end{cases}$$
(9)

式(9)中第1个方程说明,若要消除动态力纷 争,2个作动器的动态位置跟随性能要一致。第 2个方程说明,要消除力纷争2个作动器的动态抗 负载扰动性能需要根据其连接刚度的值来设计,而 即使是在实际系统中,2个 EMA 的连接刚度也不 会有太大差别,并且 k_{L2} - $k_{L1} \ll k_{L1}k_{L2}$,因此,要消除 动态力纷争,只需使2个通道的动态位置跟随性能 和动态抗负载扰动性能尽可能的相似,即无论外部 负载存在与否,2个通道随位置指令作动时需要时 刻保持输出位置一致,消除 EMA 动态力纷争的问 题也就变成了如何实现作动器的实时同步问题。

3 动态力均衡控制方法

基于动态力纷争的主要来源和作用机理分析, 分别提出针对摩擦的前馈补偿控制方法,以及针对 间隙和指令延迟的反馈控制方法。

3.1 基于速度和加速度的前馈补偿控制

前馈控制可以提高系统对指令信号的响应速 度,减小跟踪误差的同时又不影响稳定性。根据上 述分析可知,消除动态力纷争的关键在于要控制 2个作动器的动态位移输出保持一致,为达成这一 目标,作动器速度和加速度也需要得到精准控制, 因为作动器速度和加速度是影响作动器动态行为 的关键。

3.1.1 轨迹生成器

本文设计了一个轨迹生成器,其将系统的初始 位置参考信号作为轨迹生成器的输入,输出位置、 速度、加速度的轨迹参考信号,并期望各作动器都 能跟随这3个参考信号,从而实现动态的同步。如 图5和图6所示,位置参考信号经过一个二阶滤波 器,得到位置参考*x*_u、速度参考*x*_u和加速度参考 *x*_u。在实际系统中,速度与加速度都受到作动器性 能的限制,因此,对其输出值进行限幅,其中速度幅 值为±1.5 m/s,加速度幅值为±2.5 m/s²。生成器中, 参考自然频率*w*_i和参考阻尼系数*ξ*_i需要根据系统初 始时的位置跟踪性能设定。

$$t_{\rm r} = (1 + 0.9\xi + 1.6\xi^2)/w_i \tag{10}$$

由于原系统 ξ =1,上升时间为 t_r =78 ms,根据式(10) 得到 w_i =44.8 rad/s。



图 5 二阶轨迹生成器理想模型







3.1.2 前馈补偿器

在 EMA 中, 电机的运行需要克服转子加减速 的惯性转矩、摩擦转矩及舵面负载折算到电机侧的 转矩。在这三者中, 摩擦为动态力纷争的主要影响 来源, 其可以通过前馈的方法来消除^[2-23]。因此, 设 计一个前馈补偿器, 该输入是轨迹生成器输出的各 个信号, 输出是各个转矩消耗项的相关补偿值, 使 得在反馈回路产生纠正作用前, 减少这些转矩消耗 对回路的影响。

考虑摩擦非线性后,电磁转矩 Tm如式 (11) 所示,根据图 3 所示框图可得需要的电流指令 Id如式 (12) 所示。

$$T_{\rm m} = k_{\rm t} \cdot I_{\rm m} = T_{\rm L} + J_{\rm m} \frac{\mathrm{d}\omega}{\mathrm{d}t} + T_{\rm f} \tag{11}$$

$$I_{\rm d} = (x_{\rm tr} - x)P_1 \frac{P_2 s + I_2}{s} - \omega \frac{P_2 s + I_2}{s}$$
(12)

式中: T_L 为负载折算到电机轴侧的力矩; T_f 为摩擦力矩; P_1 、 P_2 、 I_2 为三闭环的控制器参数。

电流环的频响远大于速度环,因此,需要的电流指令等于实际电流即 $I_d \approx I_m$ 。本文摩擦转矩通过基于 Stribeck 模型的估计进行前馈补偿,用关于速度的函数 $f(\dot{x}_t)$ 来表示。为使式 (12) 中 $x_t - x$ 项为 0,电流补偿值 I_d^* 表达如下:

$$I_{\rm d}^{*} = \frac{2\pi i}{p} \cdot \frac{P_2 s + I_2}{s} \dot{x}_{\rm tr} + \frac{f(\dot{x}_{\rm tr})}{k_{\rm t}} + \frac{2\pi i J_{\rm m}}{p k_{\rm t}} \ddot{x}_{\rm tr}$$
(13)

为避免发生正反馈,降低作动器的稳定性,忽 略外部负载转矩的影响。同时为使补偿值不受控 制器参数的影响,将式(13)的第1项前移到速度环 上,最终得到转速前馈的表达式为



Fig. 8 Verification of feedforward compensation control

为量化分析力均衡控制的效果,本文确定动态 力纷争的指标为额定负载力的 20%,所研究的系统 负载力为4 KN。

图 9为摩擦因素下的动态力均衡验证。由 图 9(a)可知,在位置阶跃指令下,摩擦因素的阻碍 作用使得作动器的位移输出滞后,因此,在指令初 期产生了动态力纷争的峰值,由于其在低速运动时 动静态摩擦转矩的跳变使得力纷争产生振荡现 象。加入前馈补偿后,由于系统响应变快,摩擦造 成的力纷争现象要比无前馈时提前一点,但是幅值 和振荡情况基本没有得到改善。当加入了摩擦补 偿项之后,幅值减小了 70% 左右,并且力矩抖动现 象也大大得到改善,系统很快地进入力均衡状态。 然而,在1s时刻加入阶跃负载后,3条曲线几乎重 合,力纷争现象基本没有改善,这是因为在前馈中, 为了避免正反馈,没有把负载考虑进来。图9(b) 中,在正弦指令下,加入了速度及加速度前馈补偿 和摩擦补偿后,动态力纷争减小了约75%,由此初 步证明了该方法的均衡效果。

3.2 基于力差值反馈的 PID 力均衡控制

除了摩擦,指令延迟和间隙因素也是动态力 纷争的主要来源,其均发生在控制的最外环即位 置环上,其中指令延迟无法通过前馈控制来消

加速度前馈和摩擦补偿的表达式为

$$I_{\rm d}^{*} = \frac{f(\dot{x}_{\rm tr})}{k_{\rm t}} + \frac{2\pi i}{p} \cdot \frac{J_{\rm m}}{k_{\rm t}} \ddot{x}_{\rm tr}$$
(15)

图7为前馈控制框图。





Fig. 7 Feedforward compensation control based on velocity and acceleration

对该控制方法进行单个 EMA 在无负载下的验证,如图 8 所示,在摩擦因素的存在下,无前馈补偿时,系统的响应滞后于参考轨迹,加入前馈控制后, 作动器的位移、速度和加速度都很好地跟随着轨迹 生成器生成的指令。本文前馈控制解决了力均衡 的关键问题,其使得作动器的运行按照设定的参考 轨迹来,因此,即使 2 个作动器有不同的摩擦影响, 也能很好地同步运行。





除,间隙虽然可以通过实验进行测量,但是传统 基于位置指令前馈补偿的方法与前面的轨迹生 成器存在冲突,同时也因为位置指令延迟因素的 存在而无法发挥作用。因此,将指令延迟和间隙 与其他一些不确定因素视为干扰,其只能用反馈 的办法来消除。进一步设计了基于力差值反馈 的 PID 力均衡控制方法,其是传统的基于力差值 反馈的 PI 力均衡控制方法的改进,弥补了快速性 上的不足^[12],该方法与前馈补偿的动态力均衡控 制联合作用,形成了最终的动态力均衡控制方 案,其控制框图如图 10 所示。

基于力差值反馈的力均衡控制的原理是将 2个输出力的差值信号 γ ,经过一个 PID 环节转化 为位置的补偿量,再引入到各个通道的位置反馈信 号中来,通过改变作动器的输出位移来达到力均 衡。其中积分项可减小静态力纷争,比例项可加快 力均衡的速度,减小动态均衡时间,微分项对消除 动态力均衡发挥了关键的作用。通过经典控制理 论校正方法和调参获得 PID 的最佳参数,分别为 $k_p = 1 \times 10^{-6} \text{ m/N} k_i = 1 \times 10^{-5} \text{ m/N} k_d = 1 \times 10^{-8} \text{ ms/N}$ 。 图 11 为分别对间隙因素、指令延迟和 3 种动态因 素复合影响下的力均衡仿真验证,仿真对比在有前 馈补偿的情况下,无反馈、传统 PI 反馈和 PID 反馈 控制的动态力均衡效果。

从图 11 中可以看到,前馈与反馈结合的力均 衡控制能有效地将力纷争减小到指标范围内 (800 N)。并且 PID 反馈比 PI 反馈的均衡效果更 好,力纷争峰值更小,时间更短,在阶跃指令条件下 响应更稳定。在摩擦、间隙和指令延迟的复合影响 下,在位置阶跃指令条件下,PI 控制使动态力纷争 降低了 90.7%,为动态力纷争指标的 68.2%; PID 控 制使动态力纷争降低了 94.1%,为动态力纷争指标 的 43.1%。位置正弦指令条件下,PI 控制使动态力 纷争减小了 91.6%,为动态力纷争指标的 42.5%; PID 控制使动态力纷争降低了 93.3%,为动态力纷 争指标的 33.9%,由此证明了该力均衡方案比传统 PI 反馈的动态力均衡效果更好。





Fig. 10 Dynamic force equalization control block diagram

3.3 鲁棒性验证

在实际的机电作动系统中,除了本文提到的 3种因素外,还存在着许多的非线性和不确定性因 素,这些因素在不同程度上影响着作动器的行为, 导致其产生动态力纷争。为了进一步验证本文提 出的动态力均衡技术的效果及鲁棒性,采用蒙特卡 罗方法对系统在各种不确定参数下的力纷争情况 进行仿真。将所建立的非线性模型中的参数分成



Fig. 11 Dynamic force equalization verification

4类,分别为电子类参数、电机参数、结构类参数及 负载力。根据每个参数的特性,定义其分布类型, 每个参数按照分布类型生成 500 个值。力纷争相 关指标如表 2 和图 12 所示。仿真分析的参数如表 3 所示。其中,由于力纷争大小与连接刚度成正比, 因此,设置了一个范围的刚度值进行仿真研究,以 此模拟高刚度下的力均衡效果。

将余度机电作动系统其中一个通道的参数按

1.0 1.2

1.2

时间/s

1.0 1.2

时间/s

1.4 1.6 1.8 2.0

1.4 1.6 1.8 2.0

0.8 1.0

时间/s

1.4 1.6 1.8 2.0

		表 2 仿真评估指标	6 000			
	Ta	able 2 Simulation evaluation indicators	² DffMax1			
-	序号	指标	4 000 -			
-	1	位置指令下的动态力纷争,负峰值DffMin1	-			
	2	位置指令下的动态力纷争,正峰值DffMax1	\mathbb{R}^{2000} 5 Sff1			
	3	外部负载指令下的动态力纷争,负峰值DffMin2				
	4	外部负载指令下的动态力纷争,正峰值DffMax2	1 DffMin1			
	5	位置指令下的静态力纷争Sffl	-2 000 -			
	6	外部负载指令下的静态力纷争Sff2	3 DffMin2			

照表3进行蒙特卡罗仿真,系统力纷争情况统计结 果如图 13 和图 14 所示。

图 13 中, DffMin1、DffMax1 行分别为位置指令







分布类型	参数的数学期望	数值	参数的方差	数值
正态	绕组电阻/Ω	1.77	绕组电阻/Ω ²	0.08
正态	绕组电感/H	0.012	绕组电感/H ²	0.006
正态	转动惯量/(kg·m ²)	1.1×10^{-4}	转动惯量/(kg ² ·m ⁴)	5.5×10^{-5}
正态	电磁转矩系数	0.067	电磁转矩系数	0.005
正态 反电势系数/(V·s·rad ⁻¹)		0.067	反电势系数/(V ² ·s ² ·rad ⁻²)	0.005
均匀	传感器反馈增益	-0.02(下限)	传感器反馈增益	0.02(上限)
均匀	指令信号延迟/s	0(下限)	指令信号延迟/s ²	0.04(上限)
均匀	传动间隙/m	0(下限)	传动间隙/m ²	4×10 ⁻⁴ (上限)
正态	库伦摩擦/(N·m)	0.10	库伦摩擦/(N ² ·m ²)	0.02
正态 最大静摩擦/(N·m)		0.25	最大静摩擦/(N ² ·m ²)	0.019
正态	Stribeck速度/(rad·s ⁻¹)	9.6	Stribeck速度/(rad ² ·s ⁻²)	0.48
正态	粘滞摩擦系数	1.37×10^{-3}	粘滞摩擦系数	2.6×10^{-4}
均匀	连接刚度/(N·m ⁻¹)	1.0×10^{8}	连接刚度/(N ² ·m ⁻²)	2.0×10^{8}
	外部负载/N	4 000	外部负载/N ²	4 000

表3

仿真参数

0 Sff1/N DffMax2/N DffMin2/N DffMax1/N DffMin1/N -200 -400 -600 600 400 200 0 0 -200-400-600 400 200 0 100 35 28.5 15 14 1.13 234 5.0 -3 0 \$ 20 14 10 28 -100100 Sff2/N 1.99 经 20 0 -1002+10x1. 0'0'0' 0.014 0.00 2+1081 - ⁸.0 ×+10~ 1.0+10.1 1.2+10 2+10. 9 0 0 0 0 Ŷ 0 0 õ 0.2 6 6. é, 0 $J_{\rm m}/({\rm N}\cdot{\rm m})$ $T_s/(N \cdot m)$ $L_{\rm a}/{\rm H}$ $b/(N \cdot m)$ 数量统计直方图 $T_c/(N \cdot m)$ $R_{\rm a}/\Omega$ $W_{s}/(rad \cdot s^{-1})$ d/s $k_{\rm l}/({\rm N}\cdot{\rm m}^{-1})$ 0/% h/m (比例)/% $(rad \cdot s^{-1})^{-1})$ 图 13 仿真结果

Fig. 13 Simulation results



下的动态力纷争负、正峰值的仿真结果,其中方框 内的点是力纷争幅值在100N以内的结果。DffMin2、 DffMax2行分别为外部负载指令下的动态力纷争 负、正峰值的仿真结果。Sff1、Sff2行分别为2个 阶段的静态力纷争,方框内的点是力纷争幅值在10N 以内的结果。

由图 13 可得,在模型参数的变化下,系统处于 非线性、高刚度的状态,但动态力纷争的峰值幅值 总体上都限制在 700 N以内,即额定负载力的 17.5%。静态力纷争限制在 150 N以下,即额定负 载力的 3.75%,都达到了指标要求。

由图 14 可知在位置指令下的动态力纷争负峰 值在-100~0 N的占比为 51.4%, 正峰值在 0~100 N 的占比为 46.6%。静态力纷争的幅值在 10 N 以下 的, 位置指令阶段为 78.2%, 抗负载干扰阶段为 67.4%。此外, 抗负载干扰阶段的动态力纷争也大 部分聚集在幅值较小的地方。因此, 该力均衡控制 技术的鲁棒性和均衡效果都较为良好。

4 结 论

本文针对双余度主/主式机电作动系统的动态 力纷争问题,建立了余度系统完整的动力学模型, 分析了动态力纷争的来源和作用机理,将非线性因 素视为干扰开展独立分析,解析各个非线性对力纷 争的影响规律,以开展针对性的均衡控制方法,研 究并提出基于速度和加速度前馈补偿控制和基于 力差值反馈的 PID 控制联合作用的动态力均衡控 制方法,并分析均衡控制效果。得到结论如下:

1)基于速度和加速度的前馈补偿控制可以使
 2个作动器同步运动,摩擦补偿项可以很大程度上减弱摩擦非线性带来的动态力纷争,使动态力纷争
 减小 70%;基于力差值反馈的 PID 控制可以有效减弱间隙和指令延迟带来的动态力纷争。

2)2种动态力均衡控制方法综合作用下,可有效消除由摩擦、间隙和指令延迟这3种因素带来的动态力纷争影响,使动态力纷争减小93%以上。

3)通过蒙特卡罗方法证明了本文方法在各种 系统参数扰动和高刚度下的均衡能力和鲁棒性,动 静态力纷争均在指标要求范围之内。

4)本文方法简单有效,有利于工程实现,同时 为进一步的力均衡方案的设计和验证提供了一种参考。

参考文献(References)

[1] 中国民用航空局.运输类飞机适航标准: CCAR-25-R4[S].北京: 中国民用航空局, 2011: 129-130.

Civil Aviation Administration of China. Airworthiness standard for transport aircraft: CCAR-25-R4[S]. Beijing: Civil Aviation Administration of China, 2011: 129-130.

[2] QIAO G, LIU G, SHI Z H, et al. A review of electromechanical actuators for more/all electric aircraft systems[J]. Journal of Mechanical Engineering Science, 2018, 232(22): 4128-4151.

- JENTINK H W. Exploitation of actuation2015 Pre-standardisation activities on power-by-wire NLR-TP-2016-162[R]. Amsterdam: NLR-Netherlands Aerospace Centre, 2016.
- [4] IANGRANDE P, GALASSINI A, PAPADOPOULOS S, et al. Considerations on the development of an electric drive for a secondary flight control electromechanical actuator[J]. IEEE Transactions on Industry Applications, 2019, 55(4): 3544-3554.
- [5] DERRIEN J C, SECURITE S D. Electromechanical actuator(ema) advanced technology for flight controls[C]//Proceedings of the 28th International Congress of The Aeronautical Sciences. Brisbane: ICAS Press, 2012: 1-10.
- [6] 张家盛, 刘波, 何晶晶. 飞机舵面余度作动系统力纷争均衡技术 研究[J]. 液压与气动, 2018(6): 99-104. ZHANG J S, LIU B, HE J J. Force equalization technology of redundant actuation system for aircraft's control surface[J]. Hydraulics & Pneumatics, 2018(6): 99-104(in Chinese).
- [7] 田亮,张岩山,王海维. 电传飞控作动系统建模与力纷争均衡研究[J]. 西北工业大学学报, 2020, 38(3): 643-648.
 TIAN L, ZHANG Y S, WANG H W. Fly-by-wire actuation system modeling and force fight equalization research[J]. Journal of Northwestern Polytechnical University, 2020, 38(3): 643-648(in Chi-
- nese). [8] 付永领,范殿梁,李祝锋. 非相似余度作动系统静态力均衡控制 策略[J]. 北京航空航天大学学报, 2014, 40(11): 1492-1499. FU Y L, FAN D L, LI Z F. Static force equalization for dissimilar redundant actuator system[J]. Journal of Beijing University of Aeronautics and Astronautics, 2014, 40(11): 1492-1499(in Chinese).
- [9] QI H, MARE J C, FU Y. Force equalization in hybrid actuation systems[C]//Proceedings of the 7th International Conference on Fluid Power Transmission and Control. Piscataway: IEEE Press, 2009: 342-347.
- [10] KOWALSKI R. Force fight in parallel-redundant electro-mechanical actuation systems[C]//Proceedings of More Electric Aircraft 2017. Piscataway: IEEE Press, 2017: 1-3.
- [11] KOWALSKI R, WINDELBERG J, LADNER R, et al. Force fight compensation for redundant electro-mechanical flight control actuators[C]//Proceedings of the 31st Congress of the International Council of the Aeronautical Sciences, Belo Horizonte: ICAS Press, 2018: 1-7.
- [12] 范殿梁, 付永领, 郭彦青, 等. 非相似余度作动系统动态力均衡控制策略[J]. 北京航空航天大学学报, 2015, 41(2): 234-240.
 FAN D L, FU Y L, GUO Y Q, et al. Dynamic force equalization for dissimilar redundant actuator system[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(2): 234-240(in Chinese).
- [13] WANG L J, MARÉ J C. A force equalization controller for active/active redundant actuation system involving servo-hydraulic and electro-mechanical technologies[J]. Journal of Aerospace En-

gineering, 2014, 228(10): 1768-1787.

- [14] UR REHMAN W, WANG X H, CHENG Y Q, et al. Motion synchronization for the SHA/EMA hybrid actuation system by using an optimization algorithm[J]. Automatika, 2021, 62(3-4): 503-512.
- [15] XUE Y, YAO Z Q. A way to mitigate force-fight oscillation based on pressure and position compensation for fly-by-wire flight control systems[J]. Transactions of the Japan Society for Aeronautical and Space Sciences, 2020, 63(1): 1-7.
- [16] 孙晓哲,杨珍书,杨建忠,等. 机电作动系统非指令振荡信号的故 障影响分析[J]. 北京航空航天大学学报, 2018, 44(7): 1419-1429. SUN X Z, YANG Z S, YANG J Z, et al. Failure effect analysis of uncommand oscillation signals in electromechanical actuation system[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(7): 1419-1429(in Chinese).
- [17] 付永领,齐海涛,王利剑,等. 混合作动系统的工作模式研究[J]. 航空学报, 2010, 31(6): 1177-1184.
 FUYL,QIHT, WANGLJ, et al. Research on operating modes in hybrid actuation systems[J]. Acta Aeronautica et Astronautica Sinica, 2010, 31(6): 1177-1184(in Chinese).
- [18] 白玉轩. 基于神经网络的飞控机电作动系统传感器故障检测研究[D]. 天津: 中国民航大学, 2020: 12-13.
 BAI Y X. Sensor fault detection of flight control electromechanical actuation system based on neutral network[D]. Tianjin: Civil Aviation University of China, 2020: 12-13(in Chinese).
- [19] WANG L. Force equalization for active/active redundant actuation system involving servo-hydraulic and electro-mechanical technologies[D]. Toulouse: INSA, 2012: 25-26.
- [20] 臧寿松. 机电系统中摩擦辨识与补偿方法的研究与实现[D]. 西安: 西安电子科技大学, 2012: 6-7. ZANG S S. Research and experiment of friction identification and compensation in mechatronic systems[D]. Xi'an: Xidian University, 2012: 6-7(in Chinese).
- [21] 兰远锋. 电动舵机伺服系统的间隙与摩擦补偿控制[D]. 北京: 北 京交通大学, 2016: 15-16.
 LAN Y F. Compensation control of electric servo systems with backlash and friction[D]. Beijing: Beijing Jiaotong University, 2016: 15-16(in Chinese).
- [22] 李亚飞. 伺服系统摩擦补偿与控制策略研究[D]. 南京: 南京航空 航天大学, 2016: 38-40.

LI Y F. Research on friction compensation and control strategy of servo system[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2016: 38-40(in Chinese).

[23] 袭著燕, 路长厚, 潘伟, 等. 带有摩擦前馈补偿的伺服控制器设计的研究[J]. 组合机床与自动化加工技术, 2006(12): 33-37.
XI Z Y, LU C H, PAN W, et al. Study on design of servo controller with friction feedforward compensation[J]. Modular Machine Tool & Automatic Manufacturing Technique, 2006(12): 33-37(in Chinese).

Dynamic force equalization for dual redundancy electro-mechanical actuation system

SUN Xiaozhe¹, WU Jiang^{2,*}, SHI Linxuan¹, YANG Jianzhong¹

(1. School of Safety Science and Engineering, Civil Aviation University of China, Tianjin 300300, China;

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

Abstract: Redundant electromechanical actuation systems are frequently utilized in flight control surface actuation systems as a result of advancements in more-electric/all-electric aircraft technology. However, a solution to the force fight brought on by actuator output asynchrony in active / active working mode is still pending. To solve this problem, a complete linear mathematical model of the system is established, the causes of dynamic force fight are analyzed, a dynamic force equalization control method based on the combined action of speed and acceleration feedforward compensation control and PID control based on force difference feedback is studied and proposed, and its equalization ability and robustness are verified. The dynamic force equalization is found to be robust against the disturbance of various system parameters and to be able to successfully minimize the dynamic force fight caused by friction, backlash, and command delay.

Keywords: electro-mechanical actuation system; dynamic force fight; feedforward compensation control; force difference feedback; dynamic force equalization

Foundation item: Fundamental Research Funds for the Central Universities Civil Aviation of University (3122020066)

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

DOI: 10.13700/j.bh.1001-5965.2022.0480

变速条件下流体动压效应稳健性分析

张泽斌*,景世钊,袁少朋,石明辉

(郑州大学 机械与动力工程学院,郑州 450001)

摘 要:以油膜作为中间介质,对于高速旋转机械传动系统的可靠性、稳定性的提升有显 著优势。而高速轴承-转子系统在加工制造和运行过程中,不可避免地存在影响系统承载性能的不确 定性因素,影响油膜形状,导致承载性能发生变化,实际工作性能偏离设计目标。以高速动压油膜 为研究对象,为掌握几何参数和转速对油膜承载性能波动的影响,进行稳健优化和分析,提取油膜 性能稳健优化结果的特征区域。针对传动系统的不同需求,采用计算流体力学方法求解不同转速下 的动压滑动轴承油膜压力场,进而求解其主要性能指标:承载力和摩擦功耗。建立研究目标的 Kriging 近似模型,并在样本点临近区域选择稳健目标计算的子空间,计算目标均值和方差。利用非 支配排序遗传算法(NSGA-II)求解不同目标组合的Pareto最优解集。结合自组织映射图(SOM) 方法进行相关性分析,提取设计目标与几何参数、转速之间的相关性特征,分析最优特征区域中偏 心率对目标稳健性的影响,最终确定稳健性好的特征区域,并选择个别结果进行仿真计算验证。结 果表明:所提出的最优性分析方法能够清晰展现出稳健最优性区域在设计空间中的分布情况,便于 降低几何参数和转速的波动对油膜性能的影响;所提方法能提升优化设计结果的可实现性,有效促 进理论设计与工程实际的衔接。

关 键 词: 稳健设计; Kriging 模型; Pareto 前沿; 自组织映射图; 滑动轴承; 最优性分析 中图分类号: TH117

文献标志码: A 文章编号: 1001-5965(2024)04-1219-10

滑动轴承借助流体动压效应形成动压油膜,动 压油膜相对于接近刚性的轴承滚子,在高速旋转机 械传动中有独特的优势,如油膜刚度高、摩擦系数 小、驱动功率低、磨损小、阻尼大等^[1-2]。作为产生 动压效应的关键,楔形油腔受到轴瓦结构等多因素 的影响^[3-4],如轴颈和轴瓦的加工表面的表面粗糙 度、表面形貌等影响几何参数的因素对于轴承实际 工作性能的研究起重要作用^[5-6]。在轴颈、轴瓦的 公称尺寸确定后,这些影响可直接反映在轴承性能 的波动上。此外,滑动轴承在转速较小的启动、停 车阶段往往出现轴颈轴瓦的摩擦磨损,甚至咬合失 效。而"启动-额定工况运转-停车"是轴承-转子系 统的必经阶段。传统的轴承设计方法往往仅考虑 额定工况下的性能优化,缺乏对升速和降速的阶段 滑动轴承综合性能的考量,设计结果不利于降低或 避免在非额定工况下的摩擦磨损。因此,考虑加工 制造、运转工况中出现的不确定性,对于高速旋转 机械传动有重要意义。

针对以上问题,本文对变速条件下,径向动压 滑动轴承的油膜性能,进行稳健性分析^[7]。选取承 载力和摩擦功耗为设计目标,采用偏心率来反映轴 颈和轴瓦的表面形貌、制造误差等常见的几何参数 不确定因素的综合影响。分析不同转速下的动压 滑动轴承力学性能的稳健性。采用基于"变量-目

收稿日期:2022-06-11;录用日期:2022-08-15;网络出版时间:2022-08-2916:13 网络出版地址:link.cnki.net/urlid/11.2625.V.20220829.1332.001

基金项目:国家自然科学基金(12272354);河南省自然科学基金(222300420547);河南省高等学校重点科研项目计划(20A460004)

^{*}通信作者. E-mail: zebin.zhang@zzu.edu.cn

引用格式: 张泽斌, 景世钊, 袁少朋, 等. 变速条件下流体动压效应稳健性分析 [J]. 北京航空航天大学学报, 2024, 50 (4): 1219-1228. ZHANG Z B, JING S Z, YUAN S P, et al. Robust analysis of hydrodynamic performance under variable rotation speeds [J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50 (4): 1219-1228 (in Chinese).

标"映射关系的高维数据表达方法,提取动压油膜 承载性能的稳健性特征区域,并基于高维数据信息 分析稳健性能最优区域。

稳健性指标通常通过随机模拟的过程来计算, 除了传统的 Monte-Carlo 模拟^[8],也可采用拉丁超立 方等采样方法分配数据点。随机展开也是近年来 常用的方法,比如采用混沌多项式计算稳健指标^[9-10]。 最优拉丁超立方试验设计方法采用正交性较好的 初始解,综合考虑正交性和均匀性^[11],结果更稳定, 效率高于基于随机点的 Monte-Carlo 模拟。本文采 用基于 Kriging 模型的内外层嵌套的稳健优化设计 方法,因此,选择最优拉丁超立方试验设计方法进 行随机模拟。使用近似模型来代替隐式功能函数, 在优化中调用近似模型,使用精确的模型可以大大 提高可靠性分析和优化效率,在这类方法中,如何 构建精确而有效的近似模型是此类方法的关键技术难 题^[12-14]。计算出来稳健性指标后,即可调用优化算 法进行稳健优化。

综上所述,本文结合试验设计方法,构建相应 的低成本、高精度的双层 Kriging 模型,内层以油膜 性能指标为目标函数,外层在油膜性能指标的基础 上进行稳健性指标的计算。并提出将折线图和自 组织映射2种方法用于高维结果数据的表达,清晰 反映出滑动油膜稳健特征区域。为稳健优化方法 在工程中的应用提供理论参考。

1 滑动轴承动压油膜优化模型介绍

二维流体动力学润滑方程-雷诺方程可表达为

 $\frac{\partial}{\partial x} \left(\frac{h^3}{\eta} \cdot \frac{\partial p}{\partial x} \right) + \frac{\partial}{\partial z} \left(\frac{h^3}{\eta} \cdot \frac{\partial p}{\partial z} \right) = 6\nu \frac{\partial h}{\partial x}$ (1)

式中:h为油膜的厚度,mm;v为润滑油黏度,Pa·s; p为进油处的压力,Pa。

由式(1)可知,油压的变化与润滑油的黏度、油膜的厚度的变化有关,利用式(1)可求出油膜中各 点的压力*p*,全部油膜压力之和即为油膜的承载力。

滑动轴承有高速化、重载化的发展趋势。转速的升高及载荷的增加容易使润滑油温度升高,而温 升将会导致润滑油的黏度降低。在高转速下达到 新的平衡之后,润滑油黏度的降低会引起新的偏心 率升高。在重载场合,偏心率升高带来油膜厚度的 减小,轴颈和轴瓦内表面之间接触的概率增加,从 而摩擦的概率也将会增大,随之摩擦功耗的增加, 将会直接导致滑动轴承的寿命降低。本文以降低 摩擦功耗,增加承载力为主要的研究目标。选取动 压滑动轴承的轴颈直径为 40 mm、宽度为 35 mm、 半径间隙为 0.05 mm,选择标准黏度为 0.048 Pa·s 的





为了说明本文设计方法,选择能够综合反映油 膜厚度、轴承尺寸、加工误差等几何参数变化的偏 心率作为研究油膜稳健性的变量。问题可描述为 max(*F*_c),min(*H*_f)

s.t.
$$\begin{cases} 0.1 \le \varepsilon \le 0.9 & (2) \\ 1\ 000 \le N \le 10\ 000 \end{cases}$$

式中: F_r 为承载力; H_f 为摩擦功耗, 可通过有限元法 求解压力边界条件和流量平衡条件下的雷诺方程 得到; ε 为偏心率, 参考初始设计的取值为 0.5; N为 转速, 参考设计的取值为 5 500 r/min。 F_r 和 H_f 共同 反映了油膜的性能指标。

偏心率对滑动轴承承载力和摩擦性功耗的影 响具体体现在:由于结构参数的变化或润滑油黏度 变化导致承载力均值和方差的变化的程度,即稳健 性指标,计算式分别为

$$\mu_{F_{\rm r}} = \frac{1}{N_{\rm SS}} \sum_{i=1}^{N_{\rm SS}} F_{\rm r_i} \tag{3}$$

$$\sigma_{F_{\rm r}}^2 = \frac{1}{N_{\rm SS} - 1} \sum_{i=1}^{N_{\rm SS}} (F_{\rm r_i} - \mu_{F_{\rm r}})^2 \tag{4}$$

式中: N_{ss} 为随机模拟过程所使用的样本点数; F_{ri} 为第i个样本点处的承载力值; μ_{Fi} 为随机模拟结果承载力均值; σ_{Fi}^2 为承载力的方差。

同样,摩擦功耗的稳健指标可表达为

$$\mu_{H_{\rm f}} = \frac{1}{N_{\rm SS}} \sum_{i=1}^{N_{\rm SS}} H_{\rm fi} \tag{5}$$

$$\sigma_{H_t}^2 = \frac{1}{N_{\rm SS} - 1} \sum_{i=1}^{N_{\rm SS}} (H_{\rm fi} - \mu_{H_t})^2 \tag{6}$$

式中: H_{fi} 为第i个样本点处的摩擦功耗取值; μ_{H_i} 为随机模拟结果摩擦功耗均值; $\sigma_{H_i}^2$ 为摩擦功耗的方

差。通过最大化承载力,最小化摩擦功耗,最小化 二者的方差进行稳健优化设计。

2 稳健建模及优化分析

本文的研究对象包含转速N和一个反映几何尺

寸参数的变量ε,均可看作随机变量来进行处理。 采用传统的近似模型方法即可建立满足稳健设计 要求的稳健模型。鉴于以上分析,综合最优拉丁超 立方试验设计方法和 Kriging 近似模型的方法,模 型的建立和稳健性优化分析流程如图 2 所示。



Fig. 2 Flowchart of robust optimization and analysis

稳健优化目标为承载力和摩擦功耗各自的均 值和方差共4个目标。本文基于一种双层模型,可 看作由内层到外层的2次建模过程。在此分别用 内层 Kriging 模型和外层稳健模型来表达,二者的 模型变量相同,目标函数不同。内层 Kriging 模型 的目标函数为承载力和摩擦功耗的响应值,对应的 第*i*个样本点可表达为[ε_i N_i F_{ri} H_{fi}]。外层稳 健模型的目标函数为承载力和摩擦功耗的4个稳 健 性 指 标,对应的 第*i*个 样 本 点 可 表 达 为 [ε_i N_i $\mu_{F_{ri}}$ $\sigma_{F_{ri}}^2$ μ_{H_a} $\sigma_{H_a}^2$]。建模方法均采用 Kriging 模型^[15]。具体建模过程如下:

步骤1 在由偏心率 ε 和转速N构成的二维变 量空间中,按照试验设计方法选择初始样本点*S*: [$(\varepsilon_1, N_1), \dots, (\varepsilon_i, N_i), \dots$];采用计算流体力学(computational fluid dynamics, CFD)仿真软件 Fluent 进行仿 真计算,获取各样本点处的目标响应值*Y*: [$(F_{r1}, H_{f1}), \dots, (F_{ri}, H_{fi}), \dots$];利用*S*、*Y*的数据建立内 层 Kriging 模型并进行交叉验证,如模型不满足精 度要求,采用序贯设计的方法进行加点^[16],直至模 型精度达到期望精度。

步骤 2 基于内层 Kriging 模型, 进行随机模拟 计算。在初始样本点集 S 中的每个样本点 5% 的邻 域范围, 选择设计空间的子空间, 在此子空间内进 行随机模拟的采样,样本点的响应值通过内层 Kriging模型获取,无需进行 CFD 仿真计算。对初 始样本集中的每个样本点均进行如上随机模拟,获 取各样本点邻域处的均值响应 μ_{F_t} 、 μ_{H_t} 和方差响应 $\sigma_{F_t}^2$ 、 $\sigma_{H_t}^2$ 。根据随机模拟的结果建立外层稳健模型 (μ_{F_t} , μ_{H_t} , $\sigma_{F_t}^2$, $\sigma_{H_t}^2$) = $f_{RO}(\varepsilon, N)$ 。

步骤3 基于外层稳健模型 *f*_{RO}进行稳健优化 设计,因包含4个目标,采用第2代非支配排序 遗传算法(non-dominated sorting genetic algorithm, NSGA-II)进行多目标优化^[17],并对优化结果进行分 析,提取油膜性能的稳健性敏感特征区域,为滑动 轴承的稳健设计起到铺垫的作用。

2.1 试验设计及近似模型建模

Kriging 模型基于概率统计理论,在已知样本数据基础上构建的近似模型或称预测模型^[18-19]。 Kriging 模型具有对非线性函数的良好近似能力,其表达式为

$$\hat{\mathbf{y}}(\mathbf{x}) = \mathbf{F}^{\mathrm{T}}(\mathbf{x})\boldsymbol{\beta} + \mathbf{z}(\mathbf{x})$$
(7)

式中:^ŷ为未知点的函数估计; β为回归系数; **F**^T(**x**) 为关于**x**的多项式,通常使用常数、一次或二次多 项式; **z**(**x**)为随机项决定了模型理论误差的分布 规律。

给定*m*个设计点*S* = $[x_1, x_2, \dots, x_m]^T$, *S* \in **R**^{*m*×*n*}和

响应 $\boldsymbol{Y} = [y_1, y_2, \dots, y_m]^T, \boldsymbol{Y} \in \mathbf{R}^{m \times 1}$,则任意待测点 x_i 的线性无偏估计值 $\hat{\boldsymbol{y}}(\boldsymbol{x})$ 为

$$\hat{y}(\boldsymbol{x}) = \boldsymbol{c}^{\mathrm{T}} \boldsymbol{Y} = \boldsymbol{F}^{\mathrm{T}}(\boldsymbol{x})\hat{\boldsymbol{\beta}} + \boldsymbol{r}^{\mathrm{T}}(\boldsymbol{x})\boldsymbol{R}^{-1}(\boldsymbol{Y} - \boldsymbol{F}\hat{\boldsymbol{\beta}})$$
(8)

式中: c为响应Y的系数向量; R为相关矩阵; $\hat{\beta}$ 为最 小二乘估计值; $r^{T}(x)$ 为样本点与预测点之间的相 关性。

在建立内层 Kriging 模型前,需要利用试验设 计确定样本点在设计空间中的分布情况。试验设 计的作用一般是通过有限的样本点尽可能的反映 设计空间的全局信息。样本点的选择是内层 Kriging 模型建模中至关重要的步骤,其分布情况直 接影响到模型的全局精度和收敛速度。考虑到问 题的维度数仅为2,CFD 仿真计算的计算量等因 素,初始模型的样本点集的选择采用部分因子试验 设计方法,选取2因子3水平共9个样本点,即滑 动轴承的9组偏心率和转速的组合数据。随机模 拟过程基于已经建立的内层 Kriging 模型,由于模 型预测的计算时间很短,因此,选择大量的随机计 算样本点。本文利用最优拉丁超立方获取随机模 拟所需的样本点,构造精度满足要求的外层稳健 模型。

2.2 模型验证

采用 K-fold 交叉验证方法来验证内层 Kriging 模型的精度^[20]。该方法将样本数据分为训练集和 验证集,训练集用于构建 Kriging 模型,而验证集则 可以用于检验相应模型的精度。同时,为保证验证 结果更具有客观性,所有样本数据轮流作为验证 集,并使用所有验证结果的均值作为最终检验标 准。具体步骤如下:

步骤1 将数据均匀分成*K*组,每组有*m/K*个 样本点,*m*为训练集与验证集的样本总数和。

步骤 2 将其中一组数据作为验证集*Q*_i,其余 *K*-1 组作为训练集。

步骤3 用 K-1 组训练集建立 Kriging 模型。

步骤 4 将验证集 Q_i 代入建立好的 Kriging 模型预测对应的响应,并记录模型响应值 \hat{y}_i 。

步骤5 重复步骤2,直到进行*K*组数据均有 且仅有一次作为验证集。

定义一个判断模型收敛的平均相对区间偏差 L,计算式为

$$L = \frac{1}{m} \left(\frac{\sum_{i=1}^{m} |y_i - \hat{y}_i|}{\max(\hat{y}_1, \hat{y}_2, \cdots, \hat{y}_m) - \min(\hat{y}_1, \hat{y}_2, \cdots, \hat{y}_m)} \right)$$
(9)

式中: y_i为每一个验证样本点的响应值,当L<0.1 时,则可以说明模型的精度符合收敛判定条件。

2.3 稳健建模

稳健模型旨在描述偏心率和转速的波动对设 计目标波动的影响,通过优化的方法降低油膜性能 指标对偏心率和转速波动的敏感性。本文基于确 定性模型建立稳健模型^[21-22]。为建立承载力和摩擦 功耗的稳健模型,需要进行随机模拟。在原设计样 本空间的9个点邻域范围的局部子空间按照最优 拉丁超立方原则采得1600个模拟点,为了使局部 子空间范围的确定更具有客观性,并使子空间与初 始样本点之间存在一定的联系,以式(10)确定子空 间的范围。

$$x_j \in U\left(x_i, \frac{1}{n}(b-a)\right)$$

 $i = 1, 2, \cdots, N_{\rm S}; j = 1, 2, \cdots, N_{\rm SS}$ (10)

式中: x_j为稳健建模所使用的样本点; N_s为初始样 本点的个数,本文取N_s = 9; a、b分别为初始样本空 间的上限、下限值。

式(10)表达了在初始样本点*x_i*以 ¹/_n(*b*-*a*)为邻 域边长的超立方子空间中,取*N*ss个随机模拟样本 点,稳健模型采样点的个数对稳健指标的计算结果 的影响未知,本文采用试算法,计算*N*ss = 250~ 7000点的均值与方差,发现随着*N*ss的增加均值和 方差呈下降趋势。*N*ss增加至1500后,稳健指标均 值和方差计算结果趋于稳定,考虑采样点过多会导 致计算量的时间及成本上升,因此,选择1600点来 进行承载力和摩擦功耗的稳健指标的计算。

2.4 动压油膜稳健性特征建模与优化

基于所得到的稳健模型,对本文动压滑动轴承 的油膜承载力和摩擦功耗的均值和方差进行优 化。稳健指标共4个,采用本文稳健优化方法,求 解承载力均值最大化,摩擦功耗均值最小化,二者 的方差都最小化的多目标稳健最优解集。采用 NSGA-II 求取 Pareto 最优解集,初始种群数目为 1000,变异概率为0.1,遗传代数上限定为100代^[17]。 对于不同的目标组合,所得到的最优解分布不同。 以下通过对不同目标组合的优化结果分析,阐明本 文的理论结论,旨在为滑动轴承的变转速条件下的 稳健优化设计提供理论依据。

首先,以油膜承载力均值最大化,摩擦功耗均 值最小化为目标,进行双目标优化设计,Pareto前沿 如图 3 所示。作为双目标优化的结果,Pareto 最优 解集包含多个设计结果,这些结果在目标空间中的 分布情况可以从 Pareto 前沿中反映。在众多结果 中选择一个方案,按照不同的标准将得到不同的结果,本文采用"单位承载力下的摩擦功耗"满意度曲线来辅助选择。



power consumption and carrying capacity

图 3 中,圆点为 NSGA-II 的第一代个体位置, 黑色离散点标示出承载力均值最大,摩擦功耗均值 最小的 Pareto 前沿。Pareto 最优解集中共有 105 个 非支配最优解。优化的目的是为设计提供理论依 据,最优解的分布反映了最优设计在设计空间中的 位置规律,通过对其进行分析,可以对设计起到指 导的作用。仅通过 Pareto 前沿图无法得知 Pareto 最 优解集在设计空间的具体分布。将设计变量和目 标函数独立分析,虽然可以得到各自在设计空间和 目标空间的分布情况,但设计变量和目标运过的联 系并不能反映出来。设计变量和目标函数变量一 共有 4 个,也不能直观地通过三维图像来表达。本 文采用"折线图"来形象化表征多目标设计结果之 间的联系,绘制步骤如下:

步骤1 将 Pareto 最优解集中的所有数据进行 归一化处理,将设计变量和目标函数的数值都统一 到 0~1的区间。设某变量(为 ε 、N、 μ_{F_r} 、 μ_{H_r} 之 一)x的取值范围为: $x_L \le x \le x_U$, x_L 、 x_U 分别为此变 量在整个设计空间中的下限、上限值。由于模型已 经确立,通过对设计空间进行网格化离散处理,很 容易找到 x_L 和 x_U 的近似值。

步骤 2 计算变量的归一化值 \tilde{x} 为 $\tilde{x} = \frac{x - x_L}{x_U - x_L}$ (11)

步骤 3 对 Pareto 前沿上的 105 个解, 均进行 如式 (11)处理, 得到 105 组归一化后的数据 ε 、N、 μ_{F_r} 、 μ_{H_r} ,将每组数据中的 4 个数值点用直 线段进行连接,可以得到 105 条折线,将这 105 条 折线绘制在一张图上,即可直观表达出 4 个变量之间的联系。

采用上述步骤,绘制了以承载力和摩擦功耗的

均值为优化目标的 Pareto 最优解集结果折线图,如图 4 所示。



四维数据折线图

Fig. 4 4D fold lines diagram of Pareto solutions of mean of friction power consumption and carrying capacity

由图 4可知, Pareto 解集中的设计点基本可 以分为2部分,第1部分的偏心率ε集中接近其 取值范围的极大值 0.9 处,转速N分布为3700~ 8200 r/min,对应的承载力和摩擦功耗均值基本随 着转速的变化呈线性变化,转速越高,承载力越高, 摩擦功耗也越高。值得指出的是,这个联系规律是 多目标解集揭示出来的最优解的规律,并非原设计 空间中设计变量和目标函数之间的联系规律。解 集中第2部分的转速N基本集中在2700 r/min,偏 心率分布则较"散",对应的承载力变化也较大,显 然承载力的变化源于偏心率的变化。此部分的最 优解均对应近乎最小的摩擦功耗。

通过以上分析,可得:从优化的角度,转速与摩 擦功耗相关性较强;承载力在偏心率基本固定的前 提下,同转速线性相关,但大偏心高转速的设计结 果并不在最优解集当中,说明转速对于承载力的影 响止于一定值;如果想获得最小的摩擦功耗,转速 不能太高。利用折线图分析,能很好地诠释 Pareto 解集的物理意义。

在众多最优解中选择一个设计方案,可以采用 优劣解距离法(technique for order preference by similarity to an ideal solution, TOPSIS)方法,基于归 一化后的数据矩阵,确定有限方案中的最优方案和 最劣方案,分别计算各评价对象与最优方案和最劣 方案间的距离,获得各评价对象与最优方案和最劣 方案间的距离,获得各评价对象与最优方案的相对 接近程度,以此作为评价优劣方案的依据^[23]。而对 于滑动轴承,常用单位承载力下的摩擦功耗最小作 为目标组合,反映在图中,即一系列过原点的直线, 直线斜率 $\alpha = \mu_{H_t}/\mu_{F_t}$ 即代表此目标。因此,本文以 $\alpha = \mu_{H_t}/\mu_{F_t}$ 构造满意度线,在 Pareto 最优解集中寻 找 α最小的方案,即满意度线同 Pareto 前沿的切 点。设计结果为ε=0.79, N=3 051 r/min,对应的承载 力均值和摩擦功耗均值分别为 4 209 N 和 218.7 W, 如图 3 所示。此设计命名为"方案 1"。

从稳健设计的角度,以承载力均值μ_F,和承载力 的方差σ_F,进行双目标优化设计。由于方差对应标 准差的平方,采用标准差并不会改变目标支配关 系,因此,为简便起见,采用μ_F,-σ_F,进行双目标优 化。优化结果直接用折线图进行表达,如图 5 所示。



图 5 承载力均值及其标准差最优解集的四维数据折线图 Fig. 5 4D fold lines diagram of Pareto solutions of mean and standard deviation of carrying capacity

由图 5 可知, 承载力均值及其标准差双目标 Pareto 解集清晰地分为 2 部分:①对应承载力均值 的最大化的解,②对应承载力波动最小化的解。 2 部分解分别对应偏心率ε的 2 个极限值,即极大 值 0.9 和极小值 0.1。而这 2 部分的解在转速上没 有表现出明显不同,转速N都在 8 200~9 800 r/min 范围之内,并且接近转速的上限值。作为μ_F-σ_F,双 目标 Pareto 最优解集的结果,表明高转速对于降低 承载力的波动有显著的效果。承载力的最小波动 对应的一定是高转速条件。而承载力与偏心率ε的 相关性更强。偏心率ε增加,承载力增加明显,过大 的偏心率不利于获取最小承载力波动的设计结果。

同样,也对摩擦功耗均值 μ_{H_t} 和摩擦功耗的标准 差 σ_{H_t} 进行双目标优化,结果如图6所示。

由图 6 可知, 摩擦功耗的稳健优化结果不同于 承载力的双目标稳健优化结果。Pareto 最优解集没 有明显的分类, 寻优算法均指向偏心率较小(0.3 $\leq \epsilon \leq 0.5$), 转速较低(1000 $\leq N \leq 2900$ r/min)的设计 区域, 标准差 σ_{H_t} 出现少部分的负值情况, 这是由于 所建立的是插值数学模型, 即使全部样本点处的 σ_{H_t} 都为正, 在外插区域, 仍然可能出现负值的插值 结果。此优化结果符合润滑理论分析, 要想获取低 摩擦功耗的结果, 转速要尽量低, 偏心率不能过 高。承载力的升高势必引起摩擦功耗的升高。



图 6 摩擦功耗均值及其标准差最优解集的四维数据折线图

Fig. 6 4D fold lines diagram of Pareto solutions of the mean and standard deviation of friction power consumption

仅对承载力和摩擦功耗的波动情况进行优化, 结果如图 7 所示。



图 7 承载力标准差和摩擦功耗标准差最优解集的四维数据 折线图

Fig. 7 4D fold lines diagram of Pareto solutions of standard deviation of friction power consumption and carrying capacity

不考虑承载力和摩擦功耗本身, 仅考虑二者的 波动, 进行双目标优化。由图 7 可知, 所得的结果 可以看出明显地分成 2 部分, 即承载力波动小的设 计和摩擦功耗的波动小的设计。这 2 部分设计结 果基本同 μ_{F_r} - σ_{F_r} 和 μ_{H_r} - σ_{H_r} 的双目标优化设计结果一 致。说明承载力和摩擦功耗的波动最优解几乎没 有交集, 同时降低 2 个目标的波动基本不可能, 至 少在目前的设计空间中是不现实的。

最后,为了尝试寻找所有稳健指标的最优解 集,对4个目标进行4目标优化,结果如图8 所示。

由图 8 不容易看出 Pareto 最优解集中的设计点 之间的关联性。根据以上分析,主要由于承载力和 摩擦功耗的波动无法寻得绝对最优解,导致转速 N和偏心率 *e*取值几乎分布在各自全部的取值范 围。此结果清楚显示出目标之间的制衡关系,目标 的增多,在现有设计空间中寻找绝对最优解的可能 性越小,只能获得对某个目标或某些目标组合的相



Fig. 8 6D fold lines diagram of Pareto solutions of four objectives

对更优解,即 Pareto 解集中的非支配排序非劣解。 由于摩擦功耗会导致轴瓦表面发热、磨损甚而"咬 死"。本文在 Pareto 解集中选择摩擦功耗均值最 小,承载力均值不小于 3 300 N 的设计结果。此方 案在折线图中以黑色加粗进行显示。由图 8 还可 知,这是一个大偏心、低转速的结果。由于偏心率 同承载力的波动相关性强, σ_{Fi}的值相对比较大,但 仍然是从 Pareto 最优解集中选择的最优解之一,不 同的标准对应不同的选择,此设计作为"方案 2"。

根据以上多目标优化过程,对优化结果在设计

变量-目标函数空间中的分布进行分析。变量数目 共有6个,鉴于人眼视觉限制,为对6维的参变量 进行更为直观的分析,需要借助高维可视化方法, 本文采用自组织映射图(self-organizing maps, SOM) 进行相关性分析^[24-25]。自组织映射方法可将多维数 据用与维度数相同个数的二维图像进行表达,每个 二维图像中的颜色可表达相应变量的幅值,每个设 计点在每个二维图像中都位于同一相对位置,每个 设计点中各设计变量取值通过其在每幅二维图中 对应位置的颜色来反映,便于观察设计变量和目标 函数变量之间的相关性。

为了标明稳健优化设计的成效,本文选择"全设计空间"范围,而不是 Pareto 最优解集的设计点,进行高维数据的 SOM 低维展开和分析。为了绘制SOM,首先对原设计空间用 4 000 个随机点进行离散化处理,4 个稳健目标的响应通过以上建立的稳健模型而直接计算得到,整体呈现为4 000×6的数据矩阵。同时,添加初始设计方案及稳健优化的2 个结果方案,对4 003×6的数据矩阵进行自组织聚类和重组,以颜色代表数据幅值的大小,得到如图9所示的六维数据SOM,并在其中标注出了最优设计结果。



Fig. 9 SOM visualization of 6D data and optimal design solutions

由图 9 可知,偏心率展开图与承载力展开图的 颜色有较为显著的对应性,说明这 2 个变量有较强 相关性;转速和摩擦功耗的颜色有一定的对应性, 说明转速对摩擦功耗影响较大。这些结论与以上 分析结论一致。在全设计空间来看,2 个方案均位 于低转速条件,相对较大的偏心率区域。说明无论 是从单位承载力下摩擦功耗的单目标优化,还是对 全部稳健目标进行4目标优化,设计结果所在的区 域基本一致。由于方案2的选择重点保证摩擦功 耗均值最小,因此,承载力的均值相对较低。为了 表1 多目标优化结果

验证优化结果,对2个设计方案进行CFD验算。目标响应采用的是初始Kriging模型而不是稳健模

型,因为稳健模型计算的是均值而不是单点响应, 结果如表1所示。

Table 1 Multi-objective optimization results						
设计方案	偏心率ε	转速N/(r·min ⁻¹)	F_r/N	$H_{\rm f}/{ m W}$	F _r 相对误差/%	H _f 相对误差/%
初始设计范围	[0.1,0.9]	[10 ³ ,10 ⁴]				
初始设计	0.50	5 500	2 444	486.3		
优化"方案1"	0.79	3 051	4 209	218.7	5.6	21.3
方案1-CFD结果	0.79	3 051	3 985	180.2		
优化"方案2"	0.67	3 045	3 257	195.1	1.3	9.8
方案2-CFD结果	0.67	3 045	3 300	177.8		

注:滑动轴承轴颈直径为40 mm,半径间隙为0.05 mm。

由表1可知,方案1的偏心率为0.79,转速为3051 r/min, CFD 计算结果承载力为3985.6 N,摩擦功耗为180.2 W,均比 Kriging 原始模型预测结果更小,相对误差分别为5.6%、21.3%。摩擦功耗误差较大,主要原因是此方案的设计点位于原始样本点的外插值区域,模型在此区域的精度较低。

方案 2 的 偏心 率为 0.67, 转速为 3 045 r/min, CFD 计算结果承载力为 3 300 N, 摩擦功耗为 177.8 W, Kriging 原始模型预测值分别为 3 257 N、195.15 W, 相对误差分别为 1.3%、9.8%。考虑到 CFD 计算误 差,模型在插值区域范围内的预测误差低于 10%, 可以满足对稳健设计起到指导的作用。

最终设计方案在一定程度上降低了摩擦功耗, 同时在不同的转速情况下,可以保持较好稳健性。 在滑动轴承的制造生产中由于加工工艺或其他因 素造成设计尺寸与实际设计不相符时,在一定范围 内还能够保证轴承的性能不会急剧下降,从而放宽 了轴承对制造与后续保养的要求,节约成本。

3 结 论

 本文以偏心率的变化反映几何参数的影响, 以转速反映转速条件的影响,采用稳健优化的方法 取得以上变量和油膜性能指标的映射关系。以承 载力和摩擦功耗的均值与方差作为目标函数,利用 Kriging 模型建立内层模型,并在采用交叉验证校核 模型的精度后,进一步建立外层稳健模型。

2) 基于 NSGA-II 多目标遗传算法进行 3 次双 目标优化,1 次 4 目标优化,得到 Pareto 最优解集。 从优化理论指导设计的角度,4 目标优化的结果证 实了双目标组合优化的必要性。再通过归一化折 线图对 Pareto 解集进行分析,获取油膜性能对参数 变化的响应机理。通过 SOM 对最优特征区域可视 化分析。此稳健优化方法不但表达出变量空间-目 标空间的内在联系,还清晰地揭示了高维最优解集 在空间中的分布情况,结合物理机理的分析,起到 了理论指导设计的优良效果。最终得到2个优化 设计结果,方案2所得结果与预测结果相近,进一 步验证了本文方法的可行性。

本文仅从稳健性特征的分析和提取方面,分析 了几何参数和转速对动压油膜稳健性的影响。研 究对象基于单一的滑动轴承结构,未考虑结构变化 情况下,滑动轴承的稳健优化设计。下一步的研究 计划将轴承几何参数尺寸看作设计变量,进一步拓 展设计空间和目标空间的范围。

参考文献(References)

- [1] 郭红, 岑少起, 张绍林. 圆柱、圆锥动静压滑动轴承设计 [M]. 郑州: 郑州大学出版社, 2013: 260-266.
 GUO H, CEN S Q, ZHANG S L. Design of cylindrical and conical hydrostatic sliding bearings[M]. Zhengzhou: Zhengzhou University Press, 2013: 260-266 (in Chinese).
- [2] 张直明. 滑动轴承的流体动力润滑理论 [M]. 北京:高等教育出版 社, 1986.
 ZUANC Z M. Hudraduramia lubriation theory of sliding basing

ZHANG Z M. Hydrodynamic lubrication theory of sliding bearing [M]. Beijing: Higher Education Press, 1986 (in Chinese).

- [3] 孟凡明, 隆涛, 高贵响, 等. 气穴对滑动轴承摩擦学性能影响的 CFD 分析[J]. 重庆大学学报, 2013, 36(7): 6-11.
 MENG F M, LONG T, GAO G X, et al. Study on the effect of cavitation on tribological performances of sliding bearing by CFD method[J]. Journal of Chongqing University, 2013, 36(7): 6-11 (in Chinese).
- 【4】 张泽斌, 李永, 陈荣尚, 等. 含气率对润滑油黏度的影响[J]. 河南 科技大学学报 (自然科学版), 2019, 40(2): 23-27.
 ZHANG Z B, LI Y, CHEN R S, et al. Effect of void fraction on viscosity of lubricating oil[J]. Journal of Henan University of Science and Technology (Natural Science), 2019, 40(2): 23-27 (in Chinese).
- [5] 郭红,夏伯乾,孙一休. 径向动静压浮环轴承-转子系统稳定性分析[J]. 振动与冲击, 2017, 36(5): 7-11.
 GUO H, XIA B Q, SUN Y X. Stability analysis of a journal floating ring hybrid bearing-rotor system[J]. Journal of Vibration and Shock, 2017, 36(5): 7-11 (in Chinese).
- [6] GUO H, YANG S, ZHANG S L, et al. Influence of temperature-vis-

cosity effect on ring-journal speed ratio and stability for a hydrodynamic floating ring bearing[J]. Industrial Lubrication and Tribology, 2019, 71(4): 540-547.

- [7] 陈立周. 稳健设计 [M]. 北京: 机械工业出版社, 2000. CHEN L Z. Robust design [M]. Beijing: China Machine Press, 2000 (in Chinese).
- [8] LEARY S, BHASKAR A, KEANE A. Optimal orthogonal-arraybased latin hypercubes[J]. Journal of Applied Statistics, 2003, 30(5): 585-598.
- [9] 赵轲, 高正红, 黄江涛, 等. 基于 PCE 方法的翼型不确定性分析及 稳健设计[J]. 力学学报, 2014, 46(1): 10-19.
 ZHAO K, GAO Z H, HUANG J T, et al. Uncertainty quantification and robust design of airfoil based on polynomial chaos technique[J].
 Chinese Journal of Theoretical and Applied Mechanics, 2014, 46(1): 10-19 (in Chinese).
- [10] 赵轲. 基于 CFD 的复杂气动优化与稳健设计方法研究[D]. 西安: 西北工业大学, 2015.

ZHAO K. Research on complex aerodynamic optimization and robust design method based on CFD[D]. Xi'an: Northwestern Polytechnical University, 2015 (in Chinese).

[11] FANG K T, LIN D K J, WINKER P, et al. Uniform design: Theory and application[J]. Technometrics, 2000, 42(3): 237-248.

[12] 尚宝平, 闫富宏, 刘高峰, 等. 基于近似模型的铣削工艺参数可靠 性设计优化[J]. 中国机械工程, 2019, 30(4): 480-485.
SHANG B P, YAN F H, LIU G F, et al. Reliability-based design optimization of milling process parameters using approximate model[J]. China Mechanical Engineering, 2019, 30(4): 480-485 (in Chinese).

- [13] 蒋琛, 邱浩波, 高亮. 随机不确定性下的可靠性设计优化研究进展[J]. 中国机械工程, 2020, 31(2): 190-205.
 JIANG C, QIU H B, GAO L. Research Progresses in reliability-based design optimization under aleatory uncertainties[J]. China Mechanical Engineering, 2020, 31(2): 190-205 (in Chinese).
- [14] ZHANG Z B, LI Y H. Exploration of anisotropic design space by using unified taylor-cokriging method[J]. Applied Mathematical Modelling, 2022, 110: 45-60.
- [15] 李小刚, 程锦, 刘振宇, 等. 基于双层更新 Kriging 模型的机械结构动态特性稳健优化设计[J]. 机械工程学报, 2014, 50(3): 165-173.

LI X G, CHENG J, LIU Z Y, et al. Robust optimization for dynamic characteristics of mechanical structures based on double renewal Kriging model[J]. Journal of Mechanical Engineering, 2014, 50(3): 165-173 (in Chinese).

- [16] 张泽斌, 张鹏飞, 郭红, 等. Kriging 序贯设计方法在滑动轴承优化 中的应用[J]. 哈尔滨工业大学学报, 2019, 51(7): 178-183. ZHANG Z B, ZHANG P F, GUO H, et al. Implementation of Kriging model based sequential design on the optimization of sliding bearing[J]. Journal of Harbin Institute of Technology, 2019, 51(7): 178-183 (in Chinese).
- [17] DEB K, AGRAWAL S, PRATAP A, et al. A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II [M]//Parallel Problem Solving from Nature PPSN VI. Berlin: Springer, 2000: 849-858.
- [18] LOPHAVEN S N, NIELSEN H B, SONDERGAARD J. DACE: A MATLAB kriging toolbox (version 2.0)[R]. Copenhagen: Technical University of Denmark, 2002.
- [19] HAN Z H, XU C Z, ZHANG L, et al. Efficient aerodynamic shape optimization using variable-fidelity surrogate models and multilevel computational grids[J]. Chinese Journal of Aeronautics, 2020, 33(1): 31-47.
- [20] GEISSER S. Bayesian estimation in multivariate analysis[J]. The Annals of Mathematical Statistics, 1965, 36(1): 150-159.
- [21] 冯泽彪, 汪建均. 考虑模型响应不确定性的稳健参数设计[J]. 控制与决策, 2019, 34(2): 233-242.
 FENG Z B, WANG J J. Multi-response robust parameter design based on uncertainty of model response[J]. Control and Decision, 2019, 34(2): 233-242 (in Chinese).
- [22] 顾晓光, 马义中, 汪建均, 等. 多元质量特性的满意参数设计[J]. 控制与决策, 2014, 29(6): 1064-1070.
 GU X G, MA Y Z, WANG J J, et al. Satisfactory parameter design for multivariate quality characteristics[J]. Control and Decision, 2014, 29(6): 1064-1070 (in Chinese).
- [23] HWANG C L, YOON K. Methods for multiple attribute decision making[M]. Multiple Attribute Decision Making, Berlin: Springer, 1981: 58-191.
- [24] KOHONEN T. The self-organizing map[J]. Proceedings of the IEEE, 1990, 78(9): 1464-1480.
- [25] 张泽斌, 张鹏飞, 李瑞珍. 基于自组织映射的高维优化参变量相 关性研究[J]. 西北工业大学学报, 2020, 38(3): 677-684. ZHANG Z B, ZHANG P F, LI R Z. SOM-based high-dimensional design spaces mapping for multi-objective optimization[J]. Journal of Northwestern Polytechnical University, 2020, 38(3): 677-684 (in Chinese).
Robust analysis of hydrodynamic performance under variable rotation speeds

ZHANG Zebin*, JING Shizhao, YUAN Shaopeng, SHI Minghui

(School of Mechanical and Power Engineering, Zhengzhou University, Zhengzhou 450001, China)

Abstract: Oil film, as the intermediate medium, offers substantial benefits for enhancing the stability and dependability of mechanical transmission systems with high rotation speeds. The high-speed bearing-rotor system is inevitably affected by environmental factors in manufacturing and operation. These factors affect the shape of the oil film, leading to changes in the performance of sliding bearing, and the actual working performance deviates from the design objectives. The high-speed dynamic pressure oil film was selected as the research object to grasp the influence of geometric parameters and rotation speeds on the fluctuation of oil film performance. Robust optimization and analysis were carried out and the robust optimization result of the oil film was extracted. The oil film pressure field of the hydrodynamic sliding bearing under various operating speeds is solved using the Computational Fluid Dynamics method in accordance with the various requirements of the transmission system. The performance properties: bearing capacity and friction power consumption are solved. The Kriging approximate model is established, and robust indices are calculated within the subspaces that are selected in the vicinity of the sample points. The non-dominated sorting genetic algorithm (NSGA-II) is used to solve the Pareto optimal solution of different objective combinations. The correlation analysis is carried out with the self-organizing maps (SOM), and the correlation characteristics between design objectives, geometric parameters, and rotation speeds are extracted. The robust optimal design is determined by analyzing the impact of the eccentricity on the robust indices in the optimized space. The results show that the proposed robust optimization method can clearly show the distribution of the robust optimality region in the design space and reduce the influence of geometrical parameters and rotation speeds on the oil film performance. The proposed method can improve the feasibility of design results and effectively promote the transition from a theoretical design to an engineering real-life practice.

Keywords: robust design; Kriging model; Pareto front; self-organizing maps; slide bearing; optimality analysis

Received: 2022-06-11; Accepted: 2022-08-15; Published Online: 2022-08-29 16:13 URL: link.cnki.net/urlid/11.2625.V.20220829.1332.001

Foundation items: National Natural Science Foundation of China (12272354); Natural Science Foundation of Henan Province (222300420547); Key Scientic Research Project of Colleges and Universities in Henan Province (20A460004)

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

DOI: 10.13700/j.bh.1001-5965.2022.0458

改进反激拓扑模式下远端稳压电源控制方案

张帅*,韩小文,李儒鹏,程鹏,李正天,王玉姣

(北京电子工程总体研究所,北京100854)

摘 要:为提高远端稳压电源的供电质量,最大限度的在高可靠性前提下保证发射设备输出电压快速、平稳地达到设定值,提出一种改进反激拓扑模式下远端稳压电源设计方式,并建立相应数学模型。设计了前置论域整定的模糊比例积分微分(PID)控制算法,通过论域整定模型将模糊PID参数直接映射到模糊规则下,消除试凑法偶然性的同时消除了量化因子对误差的放大作用,有效地避免了过度调整的现象;引用模糊逻辑来实时处理远端稳压电源的动态参量值,并根据反模糊映射函数将生成的量化值映射到控制元件。仿真环境模拟结果表明:与模糊 PID 控制相比,前置论域整定的模糊 PID 控制算法具有优良稳态性能,调节时间缩短 48.1%,响应时间缩短了 28.6%并降低到 1.4 ms;与此同时,该控制方案能够有效地抵抗突发性干扰,稳压输出时间缩短了 45.5%,响应时间缩短到 37.5%。实验结果表明:前置整定模糊 PID 控制算法可以运用到实际工业环境中,与工业环境下传统控制算法比,前置整定模糊 PID 控制算法可以大幅提高远端稳压电源的供电质量,具备优良的鲁棒性。

关 键 词: 模糊 PID 控制; 论域整定; 改进反激拓扑; 远端稳压电源; 量化因子 中图分类号: TP273.4

文献标志码: A 文章编号: 1001-5965(2024)04-1229-11

随着无人作战理念的兴起,越来越多的学者开 始关注地面发射设备的智能化研究^[1]。远端稳压电 源是驱动飞行器上各类用电负载工作的重要组件, 作为地面发射设备的重要组成部分,需具备实时调 压和远端稳压的功能,根本上解决低压大电流供电 输出时因发射平台下前级取力设备供电质量差、供 电电缆长、飞行器上负载变化大等造成输出端供电 质量差的问题^[2-3]。

对远端稳压电源控制方案研究已经比较深 人。陈朋等^[4]提出一种基于反激拓扑方法与电阻 电容(resistor capacitance, RC)隔离触发网络的短弧 脉冲氙灯驱动电源设计方法,从硬件角度阐释控制 方案可行性。冯丹等^[5]采用小信号理论详细地分 析了反激式开关电源的电流环和电压环,并推导出 系统模型的传递函数。文献 [6-8] 深入研究了反激 拓扑模式下硬件电路多路输出控制特性。王凤国等^[9] 基于 PI 控制器提出了一种远端供电电压补偿方 案,通过误差放大器输出调整脉冲宽度调制(pulse width modulation, PWM)控制器脉宽的方式,实时调 整负载端供电电压,为远端稳压电源控制方案的研 究提供了思路。孙凤等^[10]以仿真与实验相结合的 方式研究了电磁作动器的控制特性,分析得出模糊 比例积分微分(proportional-integral-derivative, PID) 控制算法具备更加优良动态性能。陈帅等^[11]针对 传统模糊 PID 控制算法 ΔK。规则库中K。参数在线修 正精准度不足的问题,提出ΔK。规则库优化改进方案。 赵天宇和苏庆宇^[12]针对无刷直流电机(brushless DC motor, BLDCM)精确调速的控制特点,设计出一 种变论域模糊 PID 控制器,有效提高了 BLDCM 控 制器的动态性能,为改进模糊 PID 控制器的设计提 供了思路。综上所述,针对远端稳压输出方面的问 题,在理论设计、实验研究和数值模拟方面开展了

收稿日期: 2022-06-08; 录用日期: 2022-08-15; 网络出版时间: 2022-09-30 10:48

网络出版地址: link.cnki.net/urlid/11.2625.V.20220927.1251.005

^{*}通信作者. E-mail: 2569119978@qq.com

引用格式: 张帅, 韩小文, 李儒鹏, 等. 改进反激拓扑模式下远端稳压电源控制方案 [J]. 北京航空航天大学学报, 2024, 50 (4): 1229-1239. ZHANG S, HAN X W, LI R P, et al. Improved remote regulated power supply control scheme in improved flyback converter [J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50 (4): 1229-1239 (in Chinese).

较多的研究,但针对特定背景下小信号模型的转化 方式与整体控制模型的建立仍然需要进一步解释 说明,对于远端输出过程中控制算法控制性能提高 也需要进一步研究。

基于上述研究现状,为了提高远端稳压电源的 供电质量,最大限度的保证飞行器负载端输出电压 快速、平稳地达到设定值。本文在反激拓扑模式基 础上设计了一种前置论域整定的模糊 PID 控制算 法,通过论域整定的方式将模糊 PID 参数直接映射 到模糊规则下,有效避免 PID 控制器的初值计算 时,试凑法偶然性和量化因子对误差的放大作用造 成论域集范围扩大的现象,使输入输出可以随着控 制需求变化更加精准的映射到模糊决策中,根本上 提高远端稳压电源的供电质量。

1 远端稳压电源工作原理及数学模型

1.1 改进反激拓扑模式下远端稳压电源工作原理

电源模块主要分为功率和控制2个部分,采用传 统的反激拓扑格式:电源模块的功率部分通过电源主 功率金属氧化物半导体(metal oxide semiconductor, MOS)管的开关工作状态,形成交流方波电压,经电 源主功率隔离变压器及电源输出整流滤波电路向 次传递能量和提供所需的输出电压;电源模块的控 制电路部分,使整个电源系统工作在闭环状态,调 节主功率 MOS 管占空比的大小,达到电源稳压输 出。电源模块的电路组成作用如下:输入滤波电 路、电源模块 PWM 控制电路供电部分、电流采样 电路、电源功率变换部分、输出整流滤波电路、电 源遥控开关控制电路、A/D 电压采样回路。

反激拓扑电路中,为了实时监测电压信号,一 般采用精密稳压源 TL431 和光耦配合形成电压反 馈信号,通过TL431和光耦的配合,不断地将输出 电压信号反馈给电源管理芯片,电源管理芯片通过 反馈的信号量,不断地对输出电压进行调整,进而 达到稳压的目的。在这种设计方式中,首先,线性 光耦作为一个隔离器件,在高频工作时不利于反馈 量的形成,其次,从安全和可靠性角度来看,光耦老 化时会发生漂移(如电流传输比(CTR)下降),并且 还容易受到外部扰动和温度等因素的影响。因此, 在改进反激拓扑模式下摒弃了传统的电源管理芯 片,利用浮点型数字信号处理器(digital signal processor, DSP)处理器处理芯片信息, 通过外置引 脚,增加了电源的实时状态信息输出能力。在电压 反馈环节,使用 A/D 采样机制获取实时电压值,不 再使用采用精密稳压源 TL431 和光耦配合形成电 压反馈信号对该电路进行电压反馈,可以实现对不 同输出电压值的不间断调整与检测,使电源模块稳 定的输出特定的电压值。改进反激拓扑模式下远 端稳压电源工作原理如图1所示。





Fig. 1 Working principle diagram of remote regulated power supply in improved flyback topology mode

1.2 改进反激拓扑模式下远端稳压电源数学模型

反激拓扑电路中,对电压和电流的反馈输出是 通过调节 MOS 管占空比实现的,为了检验控制算 法的调节能力,本文以电压反馈调节为例,远端稳 压电源简化电路模型可以简化成如图 2 所示。图 2 中, *G*_{mod} 为 PWM 调制等效传递函数, *G*_{pwr} 为反激变 换器等效传递函数, *G*_{LC} 为滤波电路传递函数。

在处理反激拓扑相关数学模型时。根据状态 不同对整个回路进行建模分析,开关管闭合时,原 边绕组在 Vi作用下,线圈内存储的能量上升;开关 管断开后,原边电流耦合到副边,能量输出到负载, 在这个过程中,开关器件和变压器是非线性部分, 为获取描述整个开关周期内电路的统一模型,作如 下几个假设:



图 2 远端稳压电源简化电路模型 Fig. 2 Remote regulated power supply simplified version control block diagram

1.

 1)不考虑变压器的漏感,对变压器和开关器件 做线性化处理,输入交流分量为0;

2) 假设系统处于理想状态, 电流传输比为 100%;

3) 忽略系统频率误差;

4)由于反馈回路主要实现电压采集,对控制回路影响不大,忽略输出电压采样隔离反馈网络的影响;

在假设的基础上,开关器件的平均化等效模型 如图 3 所示。





Fig. 3 Averaged equivalent model for switching devices

由于开关频率转换过快,在进行定量分析时,可 以根据图1所示输入输出端传递关系得出数学模型为

$$\begin{cases} l_{\rm in} = DI_{\rm out} \\ v_{\rm outp} = DV_{\rm inp} \end{cases}$$
(1)

式中:D为变换器的初始占空比设定值; V_{inp} 、 v_{outp} 分别为模型输入、输出侧电压; i_{in} 、 I_{out} 分别为模型输入、输出侧电流。由于启动过程和负载剧烈变化时瞬态值过大,所以在启动和交变负载变化场景时, 不适用该公式。

当电路到达稳态后,在直流分量上叠加上交流 分量为

$$\begin{cases} i_{\rm in} = (D + \tilde{d}) \left(I_{\rm out} + \tilde{i}_{\rm out} \right) \\ v_{\rm outp} = \left(D + \tilde{d} \right) \left(V_{\rm inp} + \tilde{v}_{\rm inp} \right) \end{cases}$$
(2)

式中: *ā*为变换器的初始占空比设定值交流分量; *ī*_{out}为模型输出侧电流交流分量; *v*_{inp}为模型输入侧 电压交流分量。

忽略小量,将其代替反激拓扑变换器中的变压 器和相关开关器件,得到如图4所示小信号模型。

可以得到小信号模型的对应关系。对整体电路做直流分析,去除该小信号中交流分量,同时短路电感器件,让内部电容断路(忽略寄生阻抗来自电感),可得

$$\begin{cases} \frac{V_{i}}{N} = V_{inp} - DV_{inp} \\ V_{0} = -\frac{D}{1 - D} \cdot \frac{V_{i}}{N} \\ I_{0} = -\frac{V_{i}}{NR} \cdot \frac{D}{1 - D} \\ I_{out} = \frac{V_{1}}{NR} \cdot \frac{D}{(1 - D)^{2}} \end{cases}$$
(3)



图 4 主电路等效小信号模型 Fig. 4 Equivalent small signal model of main circuit

式中: N为变换器的输出部分绕组比; V_i 为变换器 的初始输入电压值; V_0 、 I_0 分别为模型p点实际电 压、电流值; R为模型输出侧等效电阻量(与输出部 分对外接口有关)。

在对电路做交流分析,利用基尔霍夫定律可得

$$\begin{cases} -\tilde{v}_{\text{out}p} + \tilde{i}_{\text{out}}sL + \frac{\tilde{v}_{\text{out}p}}{D} = \frac{V_{\text{in}p}}{D}\tilde{d} \\ \tilde{v}_0 = \tilde{v}_{\text{out}p} - \tilde{i}_{\text{out}}sL \\ -\tilde{v}_{\text{out}p} + D\tilde{i}_{\text{out}} + I_{\text{out}}\tilde{d} + \tilde{i}_0 = \tilde{i}_{\text{out}} \end{cases}$$
(4)

式中: L 为变换器 LC 滤波电路电感值; s 为拉斯变换输出响应。

随着稳压范围确定占空比初始值 D 的大小也 随之确定, *d* 的大小与 v₀的传递函数关系也随之确 定, 理想状态下可得

$$G_{\rm mod}G_{\rm pwr} = \frac{\tilde{v}_0}{\tilde{d}} = (1 + sR_{\rm c}C)\left(1 - \frac{k_L sL}{R_L}\right) / 2sR_L \qquad (5)$$

式中: *R*。为变换器度外输出侧电阻等效值; *R*_L、*k*_L分 别为变换器的输出侧电感端电阻等效值、等效比例 系数; *C* 为电容性。

本设计中:忽略二极管压降,只考虑半波整流的影响。G_{1C}的计算式为

$$G_{\rm LC}(s) = \frac{V_{\rm out}}{V_{\rm in}} = \frac{1}{2(j^2 w^2 L C + 1)} \tag{6}$$

式中: *V*_{out}、*V*_{in}分别为输出节点、输入节点的电压 值; w 为频域分量。

令 *s*=j*w*, 传递函数 *G*_{LC} 为

$$G_{\rm LC}(s) = \frac{V_{\rm out}}{V_{\rm in}} = \frac{1}{2LCs^2 + 2} = \frac{1}{2 \times 10^{-5}s^2 + 2}$$
(7)

参考经验公式,可以得到电压输出与占空比输 入量之间的传递函数 G_s为

$$G_{s} = G_{LC}G_{mod}G_{pwr} = \frac{(1 + sR_{c}C)\left(1 - \frac{k_{L}sL}{R_{L}}\right)}{2sR_{L}} \cdot \frac{1}{2 \times 10^{-5}s^{2} + 2} = (8)$$
$$\frac{-7 \times 10^{4}s^{2} + 1.3s + 10^{4}}{7 \times 10^{-6}s^{3} + 0.1s^{2} + 0.7s + 10^{4}}$$

2 前置论域整定的模糊 PID 控制器 设计

前置论域整定是本次模糊 PID 控制器设计关键,在经典模糊 PID 算法的基础上,不借助比例因子和量化因子,直接将 K_p、K_i、K_d(3个变量分别为比例、积分、微分作用因子,作用因子不同,控制函数收敛速度、超调量和静态误差不同)作用到模糊 PID 控制器上,即将经过前置论域整定后模糊 PID 参数,直接映射到模糊规则中进行解算输出,可大大减少试凑法偶然性和量化因子对误差的放大,增加模糊 PID 控制精度。

模糊 PID 基于传统 PID 控制算法的基础上,利 用反馈误差实时修改 PID 控制参数来对抗系统内 部变化,从而实现对控制系统的精准控制。因此, 为进一步了解模糊 PID 控制输出特性,本文先了解 PID 控制,控制函数及其传递函数为^[13-15]

$$\begin{cases} u(t) = K_{\rm p}e(t) + K_{\rm i} \int_{0}^{t} e(t) dt + K_{\rm d} \frac{de(t)}{dt} \\ G_{s} = \frac{U(s)}{E(s)} = K_{\rm p} \left(1 + \frac{1}{T_{\rm i}s} + T_{\rm d}s \right) \end{cases}$$
(9)

式中: *T*_i、*T*_d分别为积分变换、微分变换值; *e*(*t*)为 *t* 时刻的系统误差。

将线性控制函数离散化,得到输入电压的增量式 PID 算法计算式为^[16]

$$\begin{cases} u(k) = u(k-1) + \Delta u(k) \\ u(k-1) = K_{p} \left\{ e(k-1) + K_{i} \sum_{j=0}^{k-1} e(j) + \\ K_{d}[e(k-1) - e(k-2)] \right\} \\ \Delta u(k) = K_{p}[e(k) - e(k-1)] + K_{i}e(k) + \\ K_{d}[e(k) - 2e(k-1)] + e(k-2) \end{cases}$$
(10)

本次实验控制原理模型如图 5 所示,模糊控制 方面采用两输入三输出的形式进行 PID 参数的实 时调节。图 5 中,输入量为系统误差 e 和误差变化 率 e,输出量为 PID 控制参量的调节值变化值。

经典模糊 PID 控制模型如图 6 所示。对比传 统模糊 PID 控制器可知^[17],前置论域整定的模糊 PID 控制在模糊推理部分增加前置整定环节,这就 促使模糊控制部分反馈到 PID 控制器的是整定后



图 5 前置论域整定的模糊 PID 控制原理模型

Fig. 5 Fuzzy PID control model diagram of pre-universal domain rectification





Fig. 6 Classical fuzzy PID control model diagram

的实际论域范围量,不再是 PID 变化量。

2.1 模糊化和模糊规则的确定

模糊化是模糊集的隶属函数映射下精确输入 值隶属度确定的过程。模糊规则采取 Mamdani标 准化设计规则^[18]。定义输入量系统误差 *e* 和输入 量误差变化率 *e*。均为 {-6,-5,-4,-3,-2,-1,0,1,2,3,4, 5,6},模糊子集为 {NB,NM,NS,ZE,PS,PM,PB},其中, N、P、B、M、S 分别为 Negative(负数)、Positive(正 数)、Big(大)、Medium(中)、Small(小)。

模糊子集的隶属函数常用三角函数^[19-20],如图 7 所示。建立模糊规则是设计模糊 PID 控制器的关 键环节^[21],为了获取优良的控制性能,根据电压输 出特性和控制要求,参考 PID 控制规律及专家经 验,建立模糊规则如表 1~表 3 所示。



图 7 输入量隶属函数

Fig. 7 Input membership function

表 1 比例因子 K_a 的模糊规则

Table 1Fuzzy control of proportional factor K_p

a(t)	$e_{\rm c}[{\rm d}e(t)/{\rm d}t]$										
e(l)	NB	NM	NS	ZE	PS	PM	PB				
NB	PB	PB	PM	PM	PS	ZE	ZE				
NM	PB	PB	PM	PS	PS	ZE	ZE				
NS	PM	PM	PM	PM	ZE	NS	NS				
ZE	PM	PM	PS	NS	NS	NM	NM				
PS	PS	PS	ZE	NS	NS	NM	NM				
PM	PS	ZE	NS	NM	NM	NM	NB				
PB	ZE	ZE	NM	NM	NM	NB	NB				

表 2 积分因子 K_i 的模糊规则

Table 2Fuzzy control of integral factor K_i

a(t)	$e_{\rm c}[de(t)/dt]$									
e(i)	NB	NM	NS	ZE	PS	PM	PB			
NB	NB	NB	NM	NM	NS	ZE	ZE			
NM	NB	NB	NM	NS	NS	ZE	ZE			
NS	NB	NM	NS	NS	ZE	PS	PS			
ZE	NM	NM	NS	ZE	PS	PM	PM			
PS	NM	NS	ZE	PS	PS	PM	PB			
PM	ZE	ZE	PS	PS	PM	PB	PB			
PB	ZE	ZE	PS	PM	PM	PB	PB			

表 3 微分因子 K_d 的模糊规则

Table 3Fuzzy control of derivative factor K_d

a(t)			e	$e_{\rm c}[{\rm d}e(t)/{\rm d}t]$]		
e(i) –	NB	NM	NS	ZE	PS	PM	PB
NB	PS	NS	NB	NB	NB	NM	PS
NM	PS	NS	NB	NB	NB	NM	PS
NS	ZE	NS	NS	NS	NS	NS	ZE
ZE	ZE	NS	NS	NS	NS	NS	ZE
PS	ZE	ZE	ZE	ZE	ZE	ZE	ZE
PM	PB	NS	PS	PS	PS	PS	PB
PB	PB	PM	PM	PM	PS	PS	PB

在设计模糊规则时,采取的格式为: if(*e* is NB) and (e_c is NB) then (K_p is PB) (K_i is NB) (K_d is PS)。

2.2 前置论域整定

在传统的模糊 PID 模糊化运算时,通过引入量 化因子的形式将 e 和 e_c输入变量的基本论域转化 为相应模糊论域,同样,经过模糊 PID 控制器处理 后,引入比例因子来使处理后变化量的模糊论域变 成实际控制对象的基本论域。在这个过程,需要利 用 PID 控制试凑法得到的 K_p、K_i、K_d基本参数,同 时需要根据作动平台的动态性能预测 PID 参数的 变化趋势。 这个过程不仅繁琐,而且在利用试凑法获取 PID 时,往往得到的是一个阈值范围,这个范围在 比例因子的作用下可能使整体动态性能下降^[22]。 因此,在获取基础 PID 参数之后,首先利用经典模 糊控制器设计方式获取 *e* 和 *e*_c范围,再对 SIMULINK 模拟器输出的数据进行分析,根据试凑法得到范围 校验,得出新的模糊论域。本文定义新论域的中心 点 *K*_{re}、*K*_{ic}、*K*_{dc} 的计算式为

$$K_{pc} = K_{p0} + f d_s k_p$$

$$K_{ic} = K_{i0} + f d_s k_i$$

$$K_{dc} = K_{d0} + f d_s k_d$$
(11)

式中: K_{p0} 、 K_{i0} 、 K_{d0} 分别为前置整定时由试凑法得到 的模糊 PID 初始参数; f为前置整定时单个稳定周 期各个量的采样频率; d_s 为 SIMOUT 所得数据积 分; k_p 、 k_i 、 k_d 为初选比例因子。

2.3 反模糊化

模糊推理所得出的结果不能直接映射到控制 电路,因此,在经过模糊推理后需要将参数 K_p、K_i、 K_a进行反模糊化处理,反模糊化就是将模糊规则定 义的控制量转化成执行量的过程,一般采用质心 法。假设 U 是电路模糊规则下非空的模糊集,在论域 范围的数值任一点离散为 n(n<7)个垂直分量,将这 n 个分量和其所对应的隶属度取加权平均,输出值为

$$K_x = \sum_{i=1}^{n} \mu_U(x_i) x_i / \sum_{i=1}^{n} \mu_U(x_i)$$
 (12)

式中: K_x 为 PID 参量的反模糊量精准值; x_i 模糊规则下映射的各个变量; $\mu_t(x_i)$ 为隶属度。

根据式(12)得出整定后 PID 参量的控制规则, 曲面图与矢量指示图如图 8 所示。



Fig. 8 Control parameter input and output relationship graphs

3 仿真分析

3.1 前置参数获取-传统模糊 PID 仿真分析

该控制算法的有效性,利用计算机搭建不同控制方案下远端稳压电源仿真环境,检验3种不同模式的控制方案实际控制效果,同时为实验搭建获取参数资料。此次仿真采用的传递函数为式(9),前置整定过程 SIMULINK 仿真模型如图 9 所示,其中, *K* 为 SIMULINK 自带模块形式,为比例因子 *K*_p。

由于 e 输入为 [-48, 48]V,变化率不小于 480 V/s, 输入和输出论域均为 [-6,6], 由式 (13) 计算量化因 子分别为 0.125 和 0.01。

$$Y = \frac{12}{b-a} \left(X - \frac{a+b}{2} \right) \tag{13}$$

式中:Y为量化因子;a、b为输入量范围。

通过试凑法初选 PID 控制参数分别为 40、 6 000、1。前置整定过程中,根据 PID 响应特性选 取比例因子为 K_n=40、K_i=1 200、K_d=0.1。

从主程序界面获取 SIMOUT 采集参数, 根据 式(11)可计算出论域中心点位置 K_{pc}=30、K_{ic}= 10 600、K_{dc}=2.2。根据论域中心点位置, 将比例因子 直接作用到论域, 得到论域范围为 [24, 36], [10 600, 11 200], [1.6, 2.8]。

3.2 前置整定模糊 PID 仿真结果分析

在反激拓扑电路中,电压和电流的反馈输出都 可以通过调节 MOS 管占空比实现,为简化实验,检 验控制算法的调节能力,本文以电压调节为例。区 别于实发流程多块热电池分步供电,在模拟训练 时,弹上(机载)设备统一使用外部稳压设备供电, 随模拟飞行进程演进各个设备相继进入工作状态, 加载运行会造成飞行器负载瞬时增加,此时稳压电 源会随负载的变化实时变化,为了模拟负载变化时 稳压电源的控制性能,仿真分析时在 30 ms 处手动 模拟加入 80 W 的功率扰动(应答机)。参考前置整 定流程,确定 PID 控制、模糊 PID 控制、前置整定 的模糊 PID 控制输入参量,前置整定模糊 PID 算法 SIMULINK 仿真模型如图 10 所示。

前置整定模糊 PID 算法仿真曲线如图 11 所示。由图 11 可知,这3 种方法都能实现指定电压 实时跟踪,但抗干扰能力和电压跟踪精度有所差距。

为分析跟踪性能,将仿真得到的该系统的动态 性能指标定量值进行汇总,如表4所示。由表4可 知,3种控制方法都在6.5 ms左右出现最大超调值, 前置参数整定的模糊 PID 控制下系统的最大超调 值仅为17.5%,远远小于之前的2种控制算法,具有 较好的动态稳定性。此外,3种不同控制方式上升 时间不同,前置参数整定模糊 PID 控制下系统的响 应时间仅为0.5 ms,表明在前置参数整定模糊 PID 控制下,系统可以快速到达设定值;与此同时, 3种控制方式调节时间相比,前置参数整定的模糊 PID 调节能力更强,在1.4 ms能满足稳压输出要 求,远小于国标对远端输出电压供电稳态范围值的 规定。

在 30 ms 加入负载扰动后, 3 种不同控制策略 下系统均出现不同规模的波动, 前置参数整定的模 糊 PID 控制算法的整体动态性能依旧明显小于其 他控制策略, 具有满足最小的超调量的同时能够在 0.5 ms 内快速响应, 且其在 1.2 ms 内迅速恢复稳压 输出。由此可知, 前置参数整定的模糊 PID 控制抗 干扰能力更强。因此, 前置整定的模糊 PID 控制抗 干扰能力更强。因此, 前置整定的模糊 PID 控制特 性优于模糊 PID 的控制特性和 PID 控制器控制特 性, 可以得出前置整定模糊 PID 在稳态性能方面具 有较小的过冲, 能够有效地保护电路内部元器件, 响应快且更加平稳, 适用于线型非时变系统和多数



图 9 前置整定过程 SIMULINK 仿真模型

Fig. 9 SIMULINK simulation model diagram of pre-tuning process









Table 4 P	re-tuned fuzzy	PID algorit	hm simulat	ion data
位直支式	坊制笠政	最大超调量/	上升时间/	调节时间/
仍其力式	江門來唱	%	ms	ms
	前置整定的 模糊PID控制	17.50	0.5	1.4
28 V稳定仿真	模糊PID控制	21.25	0.7	2.7
	PID控制	23.36	0.9	3.0
	前置整定的 模糊PID控制	4.80	0.5	1.2
系统扰动状态	模糊PID控制	4.90	0.8	2.2
	PID控制	6.20	1.0	3.0

6.20

1.0

3.0

前置整定模糊 PID 算法仿真数据 表 4

控制领域。

稳压电源输出验证 4

通过控制器控制反激拓扑结构下电源参数,将 前置反馈计算出论域代入控制算法,并将控制算法

集成映射到算法控制器上,如图12所示。

由图 12 可知,车控系统提供 V.,为 24 V 左右波 动,当稳压输出电源收到上位机输入远端稳压信号 后,软件会切换到远采模式,定时采集 VO REM ADC 管脚的 AD 采样值, 然后对采样值进行解析, 通过离散计算出实时误差和误差变化率。接着通 过模糊规则对输入的参数 e 和 e。进行模糊化规则 映射,将 PID 参数以模糊化的方式计算出来。最 后,将反模糊化计算数值代入增量式 PID 算法,将 计算所得数值代入控制系统,从而实现对稳压电源 的实时调整。

在设计编写控制器代码时,将仿真确定的量化 因子和比例因子作用到标准模糊 PID 论域,得到对 应系统的模糊 PID 控制算法后烧写到控制器上,接 着第一次上电,导出论域调整范围值,根据公式确 定论域中心点位置。根据论域中心点位置将比例 因子直接作用到论域,得到论域范围[24,36], [10 600, 11 200], [1.6, 2.8]。将此范围代入模糊 PID 编程算法。电源模块软件PID输出控制流程如 图 13 所示。

将更改后的代码烧写到控制器,在实验室进行 软件在线调试,将输出线路串入可调电阻后,设定 远端输出电压为 28 V,实时监测远采电压反馈结 果。基于以上分析,在一次电压设定值读取后,读 取 FLASH 内前置整定的模糊 PID 计算模型中实时 电压值,读取位置如图14所示。将2次调节过程 电压变化值导出,利用 ORIGIN 对上述数据绘图, 可得如图 15 所示输出电压变化趋势。

前 置 整 定 模 糊 PID 算 法 实 验 数 据 如 表 5 所 示。由表5可知,在前置整定的模糊 PID 控制算法 下远端稳压电源系统输出和仿真结果类似,在前置



Fig. 12 Schematic diagram of circuit connection of fuzzy PID control algorithm for pre-parameter tuning













整定的模糊 PID 控制方案作用下,远端稳压电源在 11 ms 左右第一次达到设定电压值,然后在控制器 微分、积分因子的作用下,经过 0.6 ms 的输出调整, 逐渐达到静态稳定状态,最终稳定输出。同模糊



图 15 远端输出电压变化趋势

Fig. 15 Remote output voltage change trend diagram

表 5 前置整定模糊 PID 算法实验数据

Table 5 Pre-tuned fuzzy PID algorithm simulation data

控制策略	最大超调量/%	上升时间/ms	调节时间/ms
前置整定的模糊PID控制	2.53	0.5	0.6
模糊PID控制	2.86	1.0	1.4

PID 控制算法相比,实时调节过程中,稳压输出调节时间缩短了 57.1%,达到预定电压值响应时间缩短了 50.0%。前置整定模糊 PID 控制算法可以根据不同的供电需求实时调节稳定输出不同的电压值; 且前置整定模糊 PID 在稳态性能方面具有较小的过冲,在满足远端稳压动态性能指标的同时,有效提升远端稳压电源的动态输出特性,最大程度的保护电路内部元器件。具备良好的稳态性能,频率响应快且更加平稳,适用于线型非时变系统和多数控制领域。

5 结 论

本文提出一种改进反激拓扑模式下远端稳压 电源的设计方法,根据控制电路计算出该远端稳压 电源数学模型。根据控制算法提出一种前置论域 整定的模糊 PID 控制算法,通过论域整定的方式将 模糊 PID 参数直接映射到模糊规则下,消除了试凑 法的偶然性和比例因子对误差的放大作用,有效地 避免了模糊论域过度调整的现象。

1)在低压大电流场景中,监测形式由模拟量转换成数字量是提高系统传输能力的一项重要举措。通过远端采样方式将输出远端电压值转化成数字量传输,将实时电压信号传输到控制器及人工交互界面,在保证高可靠性的同时增加了硬件系统的监测能力和传输能力。

2)在前置整定的模糊 PID 控制系统下,远端稳 压电源可以根据不同的供电需求实时调节并稳定 输出不同的电压值。与无前置整定的模糊 PID 控 制系统仿真结果相比,实时调节过程中,稳压输出 调节时间缩短了48.1%,达到预定电压值响应时间 缩短了28.6%;抵抗负载扰动方面,稳压输出时间缩 短了45.5%,达到预定电压值响应时间缩短了37.5%。 实验结果知:前置论域整定的模糊 PID 控制算法可 以大幅提高工业环境下此类远端稳压电源控制器 控制性能。

3)前置论域整定过程中,根据模糊控制规律设计的论域整定模型,可以极大提高改善模糊 PID 算法控制性能,该方法可以作为一种前置整定规律应用到相关非线性及时变系统模糊 PID 控制器设计中,对提高控制器动态响应特性有着重要意义。

参考文献(References)

 [1] 孙艳秋,吴庆军,张硕,等. 面向零窗口发射的全冗余一体化测发 控系统[J]. 导弹与航天运载技术, 2022(1): 53-58.
 SUN Y Q, WU Q J, ZHANG S, et al. Integrated measurement con-

trol and launch system with full redundancy for zero-window launching[J]. Missiles and Space Vehicles, 2022(1): 53-58 (in Chinese).

- [2] 孙志群,李强,袁卫,等.路面激励对火箭炮行进间发射控制的影响研究[J]. 兵器装备工程学报, 2021, 42(4): 37-41.
 SUN Z Q, LI Q, YUAN W, et al. Study on effect of road excitation on rocket launch control between travels[J]. Journal of Ordnance Equipment Engineering, 2021, 42(4): 37-41 (in Chinese).
- [3] 周逢道, 王爽, 赵心晖, 等. 高精度同步输出多功能发射控制技术
 [J]. 中南大学学报 (自然科学版), 2016, 47(8): 2636-2642.
 ZHOU F D, WANG S, ZHAO X H, et al. Multifunctional transmitter control technology with high-precision synchronization output[J].
 Journal of Central South University (Science and Technology), 2016, 47(8): 2636-2642 (in Chinese).
- [4] 陈朋,毛待春,陈思逾,等.基于反激拓扑与 RC 隔离触发网络的 短弧脉冲氙灯电源设计[J].强激光与粒子束,2021,33(3):123-131.

CHEN P, MAO D C, CHEN S Y, et al. Design of short-arc xenon flashlamp power supply based on flyback topology and RC isolation trigger network[J]. High Power Laser and Particle Beams, 2021, 33(3): 123-131 (in Chinese).

- [5] 冯丹,任宏滨,李伟伟,等.反激式开关电源控制系统小信号模型
 [J]. 探测与控制学报, 2015, 37(5): 89-93.
 FENG D, REN H B, LI W W, et al. Small signal model of flyback switching power supply control system[J]. Journal of Detection & Control, 2015, 37(5): 89-93 (in Chinese).
- [6] 唐潇,孙文杰,何明祖,等.双极性直线型变压器驱动源的研制[J]. 强激光与粒子束, 2021, 33(6): 36-43.
 TANG X, SUN W J, HE M Z, et al. A bipolar nanosecond pulse

source based on liner transformer driver[J]. High Power Laser and Particle Beams, 2021, 33(6): 36-43 (in Chinese).

[7] 袁义生, 赖立. 一种多模式电流馈 LCL 谐振变换器[J]. 中国电机 工程学报, 2020, 40(10): 3259-3270.

YUAN Y S, LAI L. A Multi-mode current-fed LCL resonant converter[J]. Proceedings of the CSEE, 2020, 40(10): 3259-3270 (in Chinese).
[8] 张峰,谢运祥,胡炎申,等.临界模式混合光伏微型逆变器的特性 分析[J].电工技术学报,2020,35(6):1290-1302.

ZHANG F, XIE Y X, HU Y S, et al. Characteristics analysis for a boundary conduction mode hybrid-type photovoltaic micro-inverter[J]. Transactions of China Electrotechnical Society, 2020, 35(6): 1290-1302 (in Chinese).

- [9] 王凤国,陈小雷,许宝立,等. 一种稳压电源远端供电电压补偿方法设计[J]. 电子器件, 2019, 42(4): 827-831.
 WANG F G, CHEN X L, XU B L, et al. Design of compensation method for remote power supply voltage of regulated power supply[J]. Journal of Electron Devices, 2019, 42(4): 827-831 (in Chinese).
- [10] 孙凤,张琪,徐方超,等.激光光路控制电磁作动器的模糊 PID 控制特性分析[J]. 兵工学报, 2019, 40(2): 430-441. SUN F, ZHANG Q, XU F C, et al. Analysis of fuzzy PID control characteristics on electromagnetic actuator for laser optical path control[J]. Acta Armamentarii, 2019, 40(2): 430-441 (in Chinese).
- [11] 陈帅, 汪成文, 张震阳, 等. 改进模糊 PID 方法及其在电液伺服控制中的应用[J]. 机电工程, 2021, 38(5): 559-565. CHEN S, WANG C W, ZHANG Z Y, et al. Improved fuzzy PID method and its application in electro-hydraulic servo control[J]. Journal of Mechanical & Electrical Engineering, 2021, 38(5): 559-565 (in Chinese).
- [12] 赵天宇, 苏庆宇. 基于变论域模糊控制的无刷直流电机转速问题
 [J]. 电机与控制应用, 2017, 44(11): 6-11.
 ZHAO T Y, SU Q Y. Brushless DC motor speed questions based on variable universe fuzzy control[J]. Electric Machines & Control Application, 2017, 44(11): 6-11 (in Chinese).
- [13] KARAHAN O. Design of optimal fractional order fuzzy PID controller based on cuckoo search algorithm for core power control in molten salt reactors[J]. Progress in Nuclear Energy, 2021, 139: 103868.
- [14] MITRA P, DEY C, MUDI R K. Fuzzy rule-based set point weighting for fuzzy PID controller[J]. SN Applied Sciences, 2021, 3: 651.
- [15] RANJBAR E, MENHAJ M B, SURATGAR A A, et al. Design of a fuzzy PID controller for a MEMS tunable capacitor for noise reduction in a voltage reference source[J]. SN Applied Sciences, 2021, 3: 609.
- [16] YU L H, SONG S X. Design of flow cytometer liquid circuit control system based on incremental PID algorithm[J]. Journal of Physics:Conference Series, 2020, 1633(1): 012001.
- [17] KHADANGA R K, KUMAR A, PANDA S. Application of interval type-2 fuzzy PID controller for frequency regulation of AC islanded microgrid using modified equilibrium optimization algorithm[J]. Arabian Journal for Science and Engineering, 2021, 46(10): 1-17.
- [18] WANG B, GENG H P, LI H, et al. Particle swarm optimizationbased fuzzy PID controller for stable control of active magnetic bearing system[J]. Journal of Physics:Conference Series, 2021, 1888(1): 012022.
- [19] KUMAR KHADANGA R, KUMAR A, PANDA S. Frequency control in hybrid distributed power systems via type-2 fuzzy PID controller[J]. IET Renewable Power Generation, 2021, 15(8): 1706-1723.

- [20] JIN X, CHEN K K, ZHAO Y, et al. Simulation of hydraulic transplanting robot control system based on fuzzy PID controller[J]. Measurement, 2020, 164: 108023.
- [21] SHI J Z, SONG Y. Mathematical Analysis of a simplified general type-2 fuzzy PID controller[J]. Mathematical Biosciences and En-

gineering, 2020, 17(6): 7994-8036.

[22] WANG R X, AN A M, WEN Y A, et al. Study on the influence of parallel fuzzy PID control on the regulating system of a bulb tubular turbine generator unit[J]. Journal of Electrical Engineering & Technology, 2021, 16(3): 1403-1414.

Improved remote regulated power supply control scheme in improved flyback converter

ZHANG Shuai^{*}, HAN Xiaowen, LI Rupeng, CHENG Peng, LI Zhengtian, WANG Yujiao

(Beijing Institute of Electronic System Engineering, Beijing 100854, China)

Abstract: In order to improve the power supply quality of the remote regulated power supply and ensure that the output voltage of the transmitter device reaches the set value quickly and smoothly under the premise of high reliability, a design method of the remote regulated power supply in the flyback topology mode is introduced, and the corresponding mathematical model is established. Here, a fuzzy proportional-integral-derivative (PID) control algorithm for pre-domain tuning is presented. It uses fuzzy logic to process the dynamic parameter value of the remote regulated power supply in real time, and it maps the generated quantitative value to the control element in accordance with the anti-fuzzy mapping function. This effectively avoids the phenomenon of over-adjustment and eliminates the accidentality of the trial method and the amplification effect of the scale factor on the error. According to simulation results, the fuzzy PID control algorithm with pre-domain tuning. The adjustment time, response time, and reduction to 1.4 ms are all shortened by 48.1%, 28.6%, and 37.5%, respectively, while the control scheme is also able to withstand abrupt interference with good effectiveness. The experimental results show that the pre-tuned fuzzy PID control algorithm in the industrial environment, and compared with the traditional control algorithm in the industrial environment, the pre-tuned fuzzy PID control algorithm can greatly improve the power supply quality of the remote regulated power supply and has excellent robustness.

Keywords: fuzzy PID control; universe of fuzzy sets; improved flyback converter; remote regulated output power supply; quantizer factor

URL: link.cnki.net/urlid/11.2625.V.20220927.1251.005

Received: 2022-06-08; Accepted: 2022-08-15; Published Online: 2022-09-30 10:48

^{*} Corresponding author. E-mail: 2569119978@qq.com

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

DOI: 10.13700/j.bh.1001-5965.2022.0405

典型立式轻工业建筑结构抗震性能实例研究

吴斌^{1,2},杨悦²,姚仰平²,王淞晗²,姜雪薇²,许立言^{2,*}

(1. 内蒙古工业大学 土木工程学院,呼和浩特 010051; 2. 北京航空航天大学 交通科学与工程学院,北京 100191)

摘 要: 立式轻工业建筑结构存在大跨度、高空重载及竖向质量分布不均匀等特点,为研 究该结构形式的振动响应特征,以某典型立式轻工业生产车间为背景,建立了精细化三维非线性计 算模型,通过有限元仿真技术对该结构形式开展抗震性能研究。结果表明:通过合理设计,立式轻 工业建筑结构能满足结构抗震设计要求,结构损伤呈由外及里的发展过程,塑性铰分布规律表现为 高层多,低层少,重载层多,非重载层少的特点,在抗震设计过程中,应将重载楼层视为薄弱层予 以加固,对外侧抗风柱应设置水平及垂直支撑,以延缓内部承载构件的损伤破坏,提高结构的整体 抗震性能。此外,当结构存在大质量比可变荷载时,应该考虑其质量变化对结构动力特性的影响; 当设备质量占比较小时,可忽略设备与结构之间的相互作用,抗震分析时可采用设备-结构固结模型 进行简化计算。

关键词: 立式轻工业建筑; 大跨度; 高空重载; 抗震性能; 时程分析

中图分类号: TU31; TU32; TU37

文献标志码: A 文章编号: 1001-5965(2024)04-1240-11

立式轻工业厂房结构形式多样,功能用途广 泛^[1-3],不仅能够满足现代生产和建设的需求^[4],提 升工业生产效率,同时还能有效解决国内工业用地 紧张等问题^[5],因此,随着中国经济的持续发展,立 式轻工业建筑将迎来更广阔的发展空间⁶⁰。然而, 由于多工种并存,生产工艺复杂,结构极不规则^[7], 立式轻工业建筑常处于大跨度、高空重载及竖向质 量不均匀等生产工况,与普通单层厂房相比,其抗 震性能要差,在地震荷载作用下受力状态更为复 杂,破坏后果也更为严重。20世纪90年代中期,美 国学者提出根据建筑物和结构重要性确定抗震性 能,使设计的工程结构在使用期间满足各种预定的 性能要求^[8-9]。文献 [10-12] 成果表明: 通过优化设 计采取必要的减震措施可以降低振动危害。对于 多层立式轻工业建筑的抗震设计还存在诸多问题 亟待解决。因此,必须对立式轻工业建筑结构的抗

震性能进行深入的研究和评估。

立式轻工业建筑结构内一般布置有大型的机 械设备、生产装置、加工和动力原材料及工业成品等^[13], 这些附属设施、工业原料及成品对整体结构的动力 特性有较大的影响。当动力耗材、工业原料及成品 等可变荷载发生改变时,结构在地震作用下的地震 响应会变得更加复杂。同时,大质量附属设施与主 体结构的相互作用对主体结构的动力特性也存在 较大影响^[1416],附属设施^[17]在地震作用下的动力响 应对质量比十分敏感^[18],当附属结构系统的振动频 率接近主体结构的频率时,会发生共振或多级共 振,对主体结构的动力反应产生动力放大作用^[19]。 因此,附属设施与结构的相互作用在抗震分析中同 样不容忽视。

针对上述问题,本文以某典型立式轻工业生产 车间为背景,建立了精细化三维非线性计算模型,

收稿日期: 2022-05-21; 录用日期: 2022-06-05; 网络出版时间: 2022-06-22 14:29 网络出版地址: link.cnki.net/urlid/11.2625.V.20220621.1457.004

基金项目:国家重点研发计划(2018YFC0705700);国家自然科学基金(51878018,51878379);博士科研启动金(DC2300001260)

*通信作者. E-mail: xuliyan@buaa.edu.cn

引用格式: 吴斌,杨悦,姚仰平,等.典型立式轻工业建筑结构抗震性能实例研究 [J]. 北京航空航天大学学报,2024,50(4):1240-1250. WUB, YANG Y, YAO Y P, et al. Case study on seismic behavior of typical multistory light industrial building structures [J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50(4):1240-1250 (in Chinese).

1241

分析了立式轻工业建筑结构的振动特性和损伤分 布及发展规律,研究了工业原料等可变荷载对结构 动力响应的影响,考虑了设备与结构之间的相互作 用。基于计算结果,在体系层面上提出立式轻工业 建筑的结构设计优化方案及实用抗震设计方法,为 立式工业建筑结构体系的进一步推广和应用奠定 基础。

1 实例工程概况

选取某多层电子元器件生产车间作为立式轻 工业厂房抗震分析实例。该生产车间所在地的抗 震设防烈度为7度,场地类型为Ⅲ类第一组,设计 基本地震加速度值为0.1g,特征周期为0.45 s。三 维结构如图1所示,其结构几何布置规则,建筑高 度为44.43 m,建筑平面约为450.9 m×190.2 m,主 跨度为15.6 m,最小跨度为7.2 m,最大跨度为 21.0 m。各楼层平面柱网布置几乎相同,如图2所 示,各层层高不同,第1~3层层高分别为6.5,15.5, 6.5 m,第4层为坡屋面,层高为13.0 ~ 15.93 m。

此厂房各层功能用途不同,第1层为技术下夹 层;第2层为加工车间,楼板布有清洗设备、真空蒸 镀设备、高温烘烤设备等;第3层同样为技术下夹 层;第4层为加工车间,布置有柔性基底涂膜及固 化设备、PECDV设备、溅射设备、清洗设备等。第 1、3层顶部楼板主要荷载分布如图3所示,由于生 产工艺流程不同,各功能区主要设备分布位置不



设备区1



图 1 某立式轻工业生产车间三维结构

Fig. 1 Three-dimension graph of a multistory light industrial production workshop







同,设备荷载及生产资料等可变荷载分布如表1 所示。

除顶层外各层楼板有均匀的圆形开洞,圆形洞 口直径为0.39m,无大开洞,第1层与第2层沿厂房 纵向方向设置有水平和垂直支撑。主体采用钢筋



图 3 第 1、3 层顶部楼板主要荷载分布

设备区2

Fig. 3 Major load distribution on top slab of 1 and 3 floors

表1	各层顶部楼板主要荷载分布

Table 1 Primary load distribution on top floors of each level

採目	设备荷载/(kN·m ⁻²)				生产资料等可变荷载/(kN·m ⁻²)			
街広	设备区1	设备区2	设备区3	其他	设备区1	设备区2	设备区3	其他
1	30	25	20	15	3	3	3	3
2				10				5
3	25	20		15	3	3		3
4				10				3

混凝土框架结构,第1层和第2层柱混凝土等级为 C50,第3层和第4层柱混凝土等级为C40,梁混凝 土等级均为C30,钢筋均为HRB335级钢筋,顶层屋 盖为钢结构,采用Q345级钢材。

2 计算模型

2.1 单元类型

在非线性时程分析过程中,主要借助 ABAQUS 通用型有限元软件对立式厂房进行精细化建模,其 中梁、柱及柱间支撑等杆系结构均采用梁单元进行 模拟^[20],梁柱中钢筋利用关键字"*rebar"进行添加, 在添加过程中保证拉压区钢筋截面面积不变。

2.2 本构模型

ABAQUS 作为一款通用型有限元分析软件,设置有子程序接口。为了更好地模拟结构的动力性能,本次弹塑性分析将采用 PQ-Fiber 子程序中的钢筋(USteel02)、混凝土(UConcrete02)材料本构来模拟梁柱的受力行为。

2.3 设备质量处理

生产设备等附属结构的质量能够改变主体结 构的动力响应,采用集中力的形式无法考虑质量对 结构动力特性的影响,因此,根据生产设备类型的 不同,将不同区域内设备荷载简化为均匀分布在该 区域的质量点,如图 3 所示,此种计算模型以下简 称设备-结构固结模型。

2.4 楼板设置

楼板的存在为梁提供了必要的支撑和约束作 用。轻工业生产车间楼板无大开洞,楼板与结构有 较好的整体性。因此,为充分考虑现有楼板对结构 的抗弯扭作用,依据刚性楼板假定^[21],在建模过程 中用刚性杆替代楼板的支撑和约束作用。

3 自振特性

轻工业生产车间的前 8 阶自振周期如表 2 所 示,结构前 4 阶振动模态如图 4 所示。由图 4 可 知,1 阶模态以 y 方向的平动为主,同时左右两端平 动幅度存在差异,表现出一定的扭转效应。第 2 阶 模态表现出很强的一致性,几何模型整个沿纵向 (x向)平移。第 3 阶模态为扭转变形,从俯视图可 知,结构发生明显的转动,转动轴为 z 方向。第 4 阶 模态为结构局部平动,主要表现在顶层结构坡屋面 两侧在纵向的局部平动。

表 2 生产线及电力厂房自振周期 Table 2 Natural period of vibration for production line and

		Ponte	P			
振型	自振周期/s	圆频率/Hz	振型	自振周期/s	圆频率/Hz	
1阶	0.898	1.113 6	5阶	0.521	1.919 4	
2阶	0.893	1.119 8	6阶	0.521	1.921 2	
3阶	0.843	1.186 2	7阶	0.499	2.003 3	
4阶	0.551	1.814 9	8阶	0.485	2.061 9	

 $\begin{array}{c} \overbrace{i_{1}} \\ \overbrace{i_{2}} \\ \overbrace{i_{2}} \\ i_{3} \\ i_{1} \\ i_{1} \\ i_{2} \\ i_{2} \\ i_{3} \\ i_{1} \\ i_{1} \\ i_{2} \\ i_{2}$

Fig. 4 First four vibration modes of structure

4 立式轻工业建筑抗震特性分析

4.1 地震波选取

综合考虑结构所处场地特征周期、抗震设防烈度,以及场地类别,选取2组天然波(Taft波和 El-Centro波(简称"El波"))和1组人工波进行结构的 弹塑性时程分析,为真实地模拟地震对结构的影响,所有地震波均采用三向输入,主波输入方向为 结构基本振型的平动方向(y方向)。规范反应谱和 所选波型主波反应谱的归一化曲线^[22]对比如图 5 所示。由图 5 可知,在轻工业生产车间的1阶自振 周期下,所选地震波的响应与规范响应基本一致, 选波合理,可进行弹塑性时程分析。

4.2 轻工业生产车间弹塑性时程分析

根据 GB 50011—2010^[23],该地区多遇地震、罕 遇地震对应的时程分析地震动加速度峰值分别为 0.16,1 m³。





Fig. 5 Comparison of response spectrum curves for light industry production workshops

4.2.1 位移响应

该生产车间左右两侧顶点分别定义为*A、B*点(见图 1),顶点*A*处的位移时程曲线如图 6 所示。 由图 6 可知,在多遇地震作用下,Taft 波和 El 波引起的结构位移响应较大,人工波引起的位移响应较小,3 种波中最大顶层位移为 12.13 mm,最大顶层位移率为 1/3 663;在罕遇地震作用下,地震波对结构的中后期振动响应影响显著,人工波引起的位移响应最大,其次是 El 波,最后为 Taft 波,3 种波中最大顶层位移为 177.64 mm,最大顶层位移率为 1/250。结构顶点*A、B*处的顶点位移时程曲线对比如图 7 所示。由图 7 可知,同时刻 2 处位移存在较大差异,表明结构存在较显著的扭转效应。

不同工况下结构最大响应如表 3 所示,结构最 大层间位移角包络图如图 8 所示。由表 3 可知,在 多遇地震作用下,结构第 2 层层间位移角最大,整 体位移角满足不能超过 1/800 的要求;在罕遇地震 作用下,薄弱层均集中在第 2 层和第 4 层,整体最 大层间位移角在 1/516~1/182 之间,满足规范对楼 层最大层间位移角不能超过 1/50 的要求^[23]。由图 8



图 6 轻工业生产车间地震下顶点 *A* 处位移时程曲线 Fig. 6 Displacement response curves at point *A* in light







可知,第2层和第4层位移角变化较为突出,该楼 层为主要的生产作业车间,相对于技术下夹层,其 质量变化较大,对地震作用较为敏感,因此,位移响

应变化较为明显,结构设计时应予以加强。

表 3 不同工况下结构最大响应

asnansas undar variaus anaratin

	Table 5	wiaximum su u	ictur ar responses	unuer various o	operating conditions		
工况	最大顶层 位移/mm	最大层间 位移/mm	1/最大层间 位移角	1/最大顶层 位角	1/规范最大层间 位移角 ^[23]	是否满足 规范要求 ^[23]	薄弱 楼层
Taft波(多遇地震)	12.09	5.78	2 684	3 674	800	是	2
El波(多遇地震)	12.13	5.79	2 675	3 663	800	是	2
人工波(多遇地震)	9.74	4.44	3 490	4 562	800	是	2
Taft波(罕遇地震)	57.42	30.88	516	774	50	是	4
El波(罕遇地震)	85.25	41.36	385	521	50	是	4
人工波(罕遇地震)	177.64	87.75	182	250	50	是	4

4.2.2 破坏特征及塑性铰分布

罕遇地震作用下结构各层塑性铰数量如表4

所示。由表4可知:3种地震波下,结构均表现出 底层梁柱出铰少,顶层塑性铰多的现象。由于顶



表 4 罕遇地震作用下结构各层塑性铰数量

Table 4 Number of plastic hinges at each level of structure under severe earthquake

神雪泣	1)	I.	2)	쿺	3)	层	4)		台 米ケ
地辰似	梁	柱	梁	柱	梁	柱	梁	柱	心奴
Taft波	0	0	1	0	0	0	259	396	656
El波	0	0	23	0	10	0	187	451	671
人工波	10	0	71	0	84	0	203	173	541

层为设备层,荷载高位布置对结构的动力性能影 响较大,因此,地震作用下该层出铰数量较多,破 坏较为严重。El波下第4层在10,25s的塑性铰 分布情况如图9所示。由图9可知,随着地震动 的持续进行,设备层的外侧抗风柱最先出铰,继 而内部梁柱也逐渐进入塑性,塑性铰的发展过程





图 9 轻工业生产车间在 El 波下第 4 层塑性铰分布 Fig. 9 Plastic hinge distribution on the 4th floor under earthquake effects of El wave

5 生产原料等质量变化对结构动力 特性的影响

在工业生产过程中,生产原材料的消耗、成品 的运输及临时堆载的改变等都会引起整体结构震 动特性发生变化。该轻工业生产车间中,清洗用 水、固化原料及加工原材料等可变荷载分布面积较 大,由该类情况所产生的活荷载达 1.25×10⁶ kN,占 总重量的 18%,在使用过程中,其质量的变化可能 对结构抗震产生不可忽视的影响,因此,本节对楼面按 0%P、100%P(P为可变荷载)2种工况进行罕遇地震作用下的振动特性分析。

3种地震波作用下顶点 A 处主波方向的位移 响应对比如图 10 所示。由图 10 可知,除人工波 外,在 Taft 波、El 波作用下,可变荷载减少后结构 顶点 A 处位移响应显著增强。El 波形罕遇地震 作用下结构塑性铰分布对比如图 11 所示。由图 11 可知,由于可变荷载减少后结构 A 侧位移响应相



图 10 3 种地震波罕遇地震作用下顶点 A 处位移时程 曲线对比





对剧烈,结构A侧第2层梁由不出铰转变为部分 出现塑性铰,结构损伤范围呈扩大趋势。3种波 作用下轻工业生产车间最大层间位移角如图12 所示。由图12可知,可变荷载减少后,各层层间 最大位移角呈增大趋势,各层相对位移增大。因 此,生产原料等可变荷载发生变化时,整体结构 振动特性会发生改变。当可变荷载质量相对较 大时,在对立式工业厂房进行抗震设计时,应考 虑动力耗材、加工原材料及临时堆载的质量变 化,其对结构动力响应、损伤分布及破坏程度有 较为明显的影响。





6 罕遇地震作用下设备-结构相互作用 对结构弹塑性响应的影响分析

在该轻工业生产车间中,第2、4 层设置有生产 设备,总设计重量约为5×10⁵ kN,汽机与结构的相 互作用可能会影响结构的振动特性与动力响应。 此前设备-结构固结模型(见2.3节),如图13(a)所 示,通过添加质量点将设备质量分散到周边的梁柱 上,该方法考虑了大型设备质量在动力分析中产生 的影响,但忽略了设备与结构之间的相互作用。因此,为考虑设备-结构相互作用对分析结果的影响,本文建立了设备-结构弹簧模型,如图 13(b)所示,对比分析设备-结构弹簧模型与设备-结构固结模型的差别。

如图 13(b) 所示, 该模型在汽机真实布置区域 构建刚性框架, 通过对节点赋予集中质量来模拟生 产设备, 并通过弹簧将设备与对应的基座底部横梁 的节点连接, 此模型能够较好地反映设备的质量和





刚度特征^[24],同时可以考虑设备-结构间的相互作用。计算中考虑真实的弹簧刚度与等效阻尼系数, 水平方向的弹簧刚度为1000 kN/mm,竖向弹簧刚 度为2000 kN/mm,等效阻尼系数 c 计算式为

$$c = 2\varsigma \sqrt{km} \tag{1}$$

式中:k为弹簧刚度;m为设备质量; ς 为阻尼比,本 文取 $\varsigma = 0.02$ 。

首先对设备-结构固结模型和设备-结构弹簧模型分别进行模态分析,振型对比如图 14 所示。频率、周期对比如表 5 所示。由图 14 可知,设备-结构固结模型与弹簧模型的模态分析结果基本一

致。前4阶振型依旧表现为y方向平动、x方向平 动、z方向扭转及坡屋面的局部平动。由表5可知, 设备与结构采用弹簧连接后,各阶振动频率和周期 基本与固结模型计算结果相同。主要是由于采用 弹簧连接方式并未改变整体结构的总质量,同时在 原结构基础上增设弹簧与设备相连,未能对厂房整 体刚度产生明显影响。

除模态分析外,本节同样选用该3种地震波形进行罕遇地震作用下的动力计算,对比设备-结构固结模型和设备-结构弹簧模型下顶点A处的位移时程曲线,如图15所示。由图15可知,设备-结构固结模型和设备-结构弹簧模型顶点A处的位移时程曲线变化规律基本一致,仅人工波下,结构顶点A最大位移有所增加。2种工况下各楼层层间位移角包络图对比如图16所示,由图16可知,包络图总体变化不大,在El波及人工波作用下,各层层间位移角略有减小但不明显。

罕遇地震作用下结构各层塑性铰数量如表 6 所示。对比表 4 和表 6 可以看出:设备-结构固结模 型中,3 种波作用下结构总塑性铰分别为 656,671, 541 个;设备-结构弹簧模型中,3 种地震波作用下结 构总塑性铰分别为 642,670,549 个,2 种工况下结 构塑性铰数量基本一致。由此可见,由于设备质量 占比较小(约为整体结构的 6.23%),故设备与结构 主体间的相互作用对结构整体动力响应、损伤分布 及破坏程度的影响并不明显。

选取与设备相连的梁柱及周边梁的损伤情况 进行分析,如图 17 所示,2 种工况下相同部位的塑 性铰和不同部位的塑性铰均已标出。由图 17 可 知,设备-结构固结模型下该处塑性铰明显多于弹 簧模型下的塑性铰。弹簧的使用起到一定的减震 耗能的作用,因此,设备荷载附近塑性铰减少。



通过设备-结构固结模型和设备-结构弹簧模型

Fig. 14 Comparison of modal analysis between equipment-structure fixed model and spring model

图 14 设备-结构固结模型与设备-结构弹簧模型振型对比

表 5 设备-结构固结模型与设备-结构弹簧模型频率、 周期对比

 Table 5
 Comparison of frequency and period between

 equipment-structure fixed model and spring model

振型 一	频率	<u>≪</u> /Hz	周期	期/s
	设备-结构 固结模型	设备-结构 弹簧模型	设备-结构 固结模型	设备-结构 弹簧模型
1阶	1.113 6	1.113 3	0.898	0.898
2阶	1.119 8	1.119 5	0.893	0.983
3阶	1.186 2	1.185 3	0.843	0.844
4阶	1.814 9	1.814 7	0.551	0.551
5阶	1.919 4	1.918 5	0.521	0.521
6阶	1.921 2	1.919 0	0.521	0.521
7阶	2.003 3	2.002 9	0.499	0.499
8阶	2.061 9	2.061 2	0.485	0.485











对比分析可知,当设备质量占比较小时,设备与结 构之间的相互作用对整体结构的动力响应影响不 大,此时设备-结构固结模型可以用于结构抗震分 析的简化计算,能较好的模拟出整体结构的动力响 应及损伤情况。采用弹簧连接设备与结构,局部振 动响应会发生改变,合理设置弹簧可以起到非传统 阻尼器的作用,能够减少对局部承载构件的损伤, 有利于优化承载构件的截面设计。

	Table 6	Number	of plastic hi	nges at each	level of strue	cture under	severe earthq	ıake	个
	1层		2层		3层		4层		台粉
地辰极	梁	柱	梁	柱	梁	柱	梁	柱	心奴
Taft波	0	0	1	0	0	0	245	396	642
El波	0	0	22	0	8	0	196	444	670
人工波	10	0	71	0	76	0	209	183	549



(b) 设备-结构弹簧模型

图 17 罕遇地震作用下某处与设备相连梁的塑性铰分布

Fig. 17 Distribution of plastic hinges in beams connected to equipment under rare seismic excitation

7 结 论

通过对典型立式轻工业建筑-元器件生产车间 进行抗震分析,分析了该类结构损伤分布及发展规 律,研究了生产资料质量变化及附属结构与主体之 间的相互作用的影响,对立式轻工业建筑结构地震 作用下的振动特性有了初步了解,可得到以下结论 和建议:

1)动力响应分析结果表明,该立式轻工业建筑 最大层间位移角包络图和各楼层层间位移角均能 满足《建筑结构抗震设计规范》的规定,这说明通过 合理设计,立式工业建筑在大跨及高空重载等多功 能需求下仍然可以满足结构抗震设计要求。

2)荷载高位布置对结构的动力性能影响较大。在罕遇地震作用下,各楼层塑性铰出现数量有明显差别,具体表现为高层多,低层少,载重层多,非载重层少。抗震设计中应将重载层视为薄弱层,建议加大梁截面宽度,增加腰筋,以增加重载梁截面的横向刚度和承载力。此外,塑性铰的发展过程是由外及里逐渐变化的,因此,建议对荷载高位布置层的外侧抗风柱设置水平及垂直支撑,提高外侧抗风柱的抗弯能力,同时建议对重载层柱脚钢筋进行加密,以增强其抗震能力。

3)大比重生产原料等可变荷载对整体结构抗 震性能有一定影响。随着生产原料的减少,结构动 力响应明显增大,结构损伤范围呈扩大趋势。因 此,在对立式工业厂房进行抗震设计时,应考虑动 力耗材、临时堆载及加工原材料的影响,其质量变 化对结构动力响应、损伤分布及破坏程度有较为明 显的影响。

4) 对于立式轻工业建筑,小质量比设备与结构 之间的相互作用对整体结构的振动特性影响较小, 在整体结构的动力分析中,小质量比设备可采用固 结模型进行简化研究,以减少建模及计算成本。此 外,根据研究对比可知,弹簧的使用能够减轻局部 构件的塑性损伤,设备与结构间合理设置弹簧阻尼 器可优化局部构件的动力响应,起到减震耗能的作 用,能够减少对局部承载构件的损伤,有利于优化 承载构件的截面设计。

参考文献(References)

- [1] 金仁超. 多层排架结构工业厂房在罕遇地震作用下的弹塑性反应分析 [D]. 南京: 南京理工大学, 2009.
 JIN R C. Elasto-plastic response analysis of multi-storey bent industrial workshop under rare earthquake[D]. Nanjing: Nanjing University of Science and Technology, 2009 (in Chinese).
- [2] 李智, 丁然, 聂鑫, 等. 标准化立式工业建筑楼面活荷载取值[J]. 建筑结构学报, 2023, 44(3): 268-278.
 LI Z, DING R, NIE X, et al. Determination of floor live loads on standardized multistory industrial buildings[J]. Journal of Building Structures, 2023, 44(3): 268-278 (in Chinese).
- [3] 王焕定, 王伟. 单层厂房弹塑性地震反应分析[J]. 哈尔滨建筑工 程学院学报, 1986(2): 86-103.
 WANG H D, WANG W. Analysis of the elasto-plastic seismic response of single-story mill buildings to ground motion[J]. Journal of Harbin University of Civil Engineering and Architecture, 1986(2): 86-103 (in Chinese).
- [4] 梁碧云. 新型多层框排架结构工业厂房的耗能减震研究[D]. 南京:南京理工大学, 2012.
 LIANG B Y. Study on energy consumption and shock absorption of new multi-storey frame-bent industrial workshop[D]. Nanjing: Nanjing University of Science and Technology, 2012 (in Chinese).
- [5] 王跃. 多层厂房结构体系的发展概况[J]. 工业建筑, 1993, 23(8): 3-6.

WANG Y. Developmental general situation of multistory factory structure system[J]. Industrial Construction, 1993, 23(8): 3-6 (in Chinese).

[6] 华建坤. 单跨多层大跨度工业建筑结构体系对比分析研究 [D]. 贵阳:贵州大学, 2016.

HUA J K. Comparative analysis and research on single-span multistorey long-span industrial building structure system[D]. Guiyang: Guizhou University, 2016 (in Chinese).

- [7] 全明研. 钢筋混凝土多层工业厂房的可靠性鉴定[J]. 工业建筑, 1993, 23(9): 31-34.
 QUAN M Y. Reliability appraisement of the reinforced concrete multistore industrial building[J]. Industrial Construction, 1993, 23(9): 31-34 (in Chinese).
- [8] MEDHEKAR M S, KENNEDY D J L. Displacement-based seis-

mic design of buildings-theory[J]. Engineering Structures, 2000, 22(3): 201-209.

- [9] PRIESTLEY M J N, KOWALSKY M J. Direct displacement-based seismic design of concrete buildings[J]. Bulletin of the New Zealand Society for Earthquake Engineering, 2000, 33(4): 421-444.
- [10] DE ANGELIS M, PERNO S, REGGIO A. Dynamic response and optimal design of structures with large mass ratio TMD[J]. Earthquake Engineering and Structural Dynamics, 2012, 41(1): 41-60.
- PAN D G, CHEN G D, WANG Z C. Suboptimal Rayleigh damping coefficients in seismic analysis of viscously-damped structures
 [J]. Earthquake Engineering and Engineering Vibration, 2014, 13(4): 653-670.
- [12] DISARNO L, ELNASHAI A S, NETHERCOT D A. Seismic response of stainless steel braced frames[J]. Journal of Constructional Steel Research, 2008, 64(7/8): 914-925.
- [13] 梁希强, 万叶青, 阮兵, 等. 按建筑功能区要求确定振动控制区域 划分方法的研究[J]. 建筑结构, 2020, 50(S1): 423-427.
 LIANG X Q, WAN Y Q, RUAN B, et al. Study on the method of determining the division of vibration control area according to the requirements of building functional area[J]. Building Structure, 2020, 50(S1): 423-427 (in Chinese).
- [14] CHAUDHURI S R. Simplified methods for the nonlinear seismic response evaluation of nonstructural components[D]. Irvine: University of California, 2005.
- [15] MEDINA R A, SANKARANARAYANAN R, KINGSTON K M. Floor response spectra for light components mounted on regular moment-resisting frame structures[J]. Engineering Structures, 2006, 28(14): 1927-1940.
- [16] CHAUDHURI S R, VILLAVERDE R. Effect of building nonlinearity on seismic response of nonstructural components: a parametric study[J]. Journal of Structural Engineering, 2008, 134(4): 661-670.

[17] 韩森, 王亮. 二次结构系统抗震研究的发展综述及展望[J]. 北京 建筑工程学院学报, 2002, 18(2): 30-34.
HAN M, WANG L. Review and expectation of the development of the seismic study on the primary-secondary systems[J]. Journal of Beijing Institute of Civil Engineering and Architecture, 2002, 18(2): 30-34 (in Chinese).

[18] 国巍, 李宏男. 多维地震作用下偏心结构楼板谱分析[J]. 工程力 学, 2008, 25(7): 125-132. GUO W, LI H N. Floor response spectrum of eccentric structure to two-dimensional earthquake[J]. Engineering Mechanics, 2008, 25(7): 125-132 (in Chinese).

[19] 常献辉. 剪力墙结构工业厂房的抗震性能研究[D]. 天津: 天津大学, 2012.

CHANG X H. Study on seismic performance of industrial workshop with shear wall structure[D]. Tianjin: Tianjin University, 2012 (in Chinese).

- [20] 聂建国, 陶慕轩. 采用纤维梁单元分析钢-混凝土组合结构地震反应的原理[J]. 建筑结构学报, 2011, 32(10): 1-10. NIE J G, TAO M X. Theory of seismic response analysis of steelconcrete composite structures using fiber beam elements[J]. Journal of Building Structures, 2011, 32(10): 1-10 (in Chinese).
- [21] 刘大海,杨翠如,钟锡根. 多层厂房差异平移—扭转耦联振动的 地震反应分析[J]. 土木工程学报, 1983, 16(4): 53-66. LIU D H, YANG C R, ZHONG X G. A differential translation-torsion coupling earthquake response analysis for multi-story factory buildings[J]. Journal of Civil Engineering, 1983, 16(4): 53-66 (in Chinese).
- [22] 邹胜斌. 工业建筑多层钢筋混凝土抗震框架的非线性动力反应 研究 [D]. 重庆: 重庆大学, 2001.
 ZOU S B. Study on nonlinear dynamic response of multi-storey reinforced concrete seismic frame in industrial buildings[D].
 Chongqing: Chongqing University, 2001 (in Chinese).
- [23] 中华人民共和国住房和城乡建设部,国家质量监督检验检疫总局.建筑抗震设计规范:GB 50011—2010[S].北京:中国建筑工业出版社,2010.

Ministry of Housing and Urban-Rural Development of the People's Republic of China, General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China. Code for seismic design of buildings: GB 50011—2010 [S]. Beijing: China Architecture & Building Press, 2010 (in Chinese).

[24] 丁然, 陶慕轩, 周萌, 等. 考虑楼板大开洞及大质量比附属设施的 结构抗震性能实例研究[J]. 工程力学, 2016, 33(3): 65-76.
DING R, TAO M X, ZHOU M, et al. Case study on structural seismic behavior considering large slab openings and auxiliary facilities with large mass ratios[J]. Engineering Mechanics, 2016, 33(3): 65-76 (in Chinese).

Case study on seismic behavior of typical multistory light industrial building structures

WU Bin^{1, 2}, YANG Yue², YAO Yangping², WANG Songhan², JIANG Xuewei², XU Liyan^{2,*}

(1. School of Civil Engineering, Inner Mongolia University of Technology, Hohhot 010051, China;

2. School of Transportation Science and Engineering, Beihang University, Beijing 100191, China)

Abstract: Vertical light industrial building structures exhibit characteristics such as large spans, high-altitude loads, and uneven vertical mass distribution. In order to investigate the vibration response characteristics of this structural form, this paper takes a typical vertical light industrial production workshop as a background and establishes a refined three-dimensional nonlinear computational model. Seismic performance research on this structural form is conducted through finite element simulation technology. The results show that the structure can meet the structural seismic requirements through reasonable design. The structural damage spreads from the outside to the inside. The plastic hinges are distributed more at high levels and less in low levels, and more on heavy load layers and less on non-heavy load layers. In the seismic design process, the heavily loaded floors should be considered as weak floors to be reinforced, and the horizontal and vertical supports should be set in the outer wind-resistant columns to delay the damage of the internal load-bearing members and improve the seismic performance of the structure. In addition, when a large mass ratio variable load exists on the structure, the effect of mass variation on the structural dynamic properties should be considered. When the mass of the equipment is smaller, the interaction between the equipment and the structure can be ignored, and the consolidation model for equipment and structure can be used for simplified calculation during seismic analysis.

Keywords: multistory light industrial building; long span; high altitude heavy load; seismic behavior; time history analysis

URL: link.cnki.net/urlid/11.2625.V.20220621.1457.004

Foundation items: National Key R & D Program of China (2018YFC0705700); National Natural Science Foundation of China (51878018,51878379); Doctoral Research Startup Fund (DC2300001260)

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

DOI: 10.13700/j.bh.1001-5965.2022.0447

嫦娥五号发动机降落羽流扬尘特性研究

张海燕^{1,2},李思新¹,王鹢²,李存惠²,张小平³,王卫东^{1,*}

(1. 西安电子科技大学 机电工程学院,西安 710071;2. 兰州空间技术物理研究所 真空技术与物理国家级重点实验室,

兰州 730000; 3. 澳门科技大学 月球与行星科学国家重点实验室,中国 澳门 999078)

摘 要: 在月球探测器着陆过程中,发动机羽流与月面相互作用后溅起的月尘是月面环境 危害的主要来源。本文以嫦娥五号任务测试数据作为仿真入口条件,采用计算流体动力学(CFD)两 阶段法建立了嫦娥五号任务中使用的喷管 1:1 模型和真空羽流扩散侵蚀模型,研究了喷管在不同降 落高度下的侵蚀速率,并计算了发动机距离月面高度为 0.5~2.0 m 范围时月尘颗粒的运动轨迹、扬 尘角和速度特性。结果表明,基于剪切应力得到的最大侵蚀速率为 8.83 kg/m²s,随着高度增加,侵 蚀速率降低,与嫦娥五号降落相机相同高度下的分析结果一致。粒径为 1、70 im 的月尘颗粒最大扬 尘高度分别为 0.72、0.36 m,最大速度分别为 2520、1010 m/s。不同粒径月尘的扬尘角范围为 1.44°~2.27°,计算的扬尘角与 Apollo 探月任务中的结果相近。

关键 词:月尘;羽流扬尘;嫦娥五号;计算流体动力学仿真;欧拉-拉格朗日方法

中图分类号: V411.6; V434.1

文献标志码: A 文章编号: 1001-5965(2024)04-1251-11

由于长期的太阳风注入、微流星体碰撞等原 因,月球表面覆盖着一层厚厚的广泛分布的风化 层。探测器降落过程中,发动机羽流喷射溅起的月 尘会引起散热器、光学器件效率降低甚至失效^[1-2], 导致机械结构卡死^[3],影响月尘采样及原位测量误 差等^[4]。因此,月尘是月面探测活动面临的最大风 险之一。Apollo 11 探月过程中,首次发现了羽流扬 尘现象^[5], Apollo 15 探月任务中羽流扬尘严重阻碍 了降落相机和航天员视线,使得探测器在无相机成 像和航天员引导的情况下着陆于坡度为 11°的月面 斜坡,严重影响了探测器安全着陆^[6]。因此,深入研 究羽流扬尘特性对提高未来月球探测器软着陆过 程的安全性具有重要意义。

月球表面无大气且低重力的环境使得发动机 羽流扬尘特性研究非常复杂,国内外研究学者先后 提出了地面实验、着陆图像分析和仿真研究3种方

法。地面实验研究过程中无法长时间保持高速羽 流的真空、低重力环境条件,无法在短时间内对不 同任务的整个羽流扬尘过程进行定量实验[7-8]。着 陆图像虽然可以客观记录扬尘过程中的月面侵蚀 特性,但有限的探月任务无法获得大量的数据,难 以支撑羽流扬尘特性的全阶段研究[6,9-11]。仿真研 究不受限于实验条件和数据数量,可以通过图像数 据分析结果,持续完善羽流扬尘模型,深入分析羽 流扬尘特性,成为了此类研究的主要手段。文献[12-13] 采取计算机流体动力学(computational fluid dynamics, CFD)和直接模拟蒙特卡罗(direct simulation Monte Carlo, DSMC)2种仿真手段,研究了 33 000 N 推力 下、距离月面不同高度月尘颗粒的扬尘速度和扬尘 角,发现2种计算方法得出的扬尘速度和扬尘角随 粒径的变化趋势一致,扬尘速度随颗粒粒径增大而 减小,扬尘角随颗粒粒径增大而增大。Morris 等^[14]

收稿日期:2022-05-31;录用日期:2022-09-03;网络出版时间:2022-09-1413:04 网络出版地址:link.cnki.net/urlid/11.2625.V.20220913.1828.011

基金项目:国家自然科学基金 (42004157);空间环境材料行为及评价技术国家重点实验室基金 (WDZC-HGD-2022-08);真空技术与物理国家 级重点实验室基金 (6142207200202);甘肃省自然科学基金 (21JR7RA743);甘肃省青年人才托举工程 (GXH20210611-05)

*通信作者. E-mail: wangwd@mail.xidian.edu.cn

引用格式:张海燕,李思新,王鹢,等. 嫦娥五号发动机降落羽流扬尘特性研究 [J].北京航空航天大学学报,2024,50(4):1251-1261. ZHANG HY, LISX, WANGY, et al. Nozzle plume erosion property on lunar dust in Chang'E-5 mission [J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50(4): 1251-1261 (in Chinese). 采用 DSMC 对发动机喷管内的羽流特性及月面真 空环境下羽流扩散特性进行了研究,获得了 Apollo 探月任务中发动机喷管内气体流动的特征参数, 研究得出 Apollo 探月任务中探测器着陆的起始 扬尘高度为15m,月尘颗粒的最大扬尘角为2.8°, 30 µm 颗粒在探测器距离月面 10 m 处的平均速度 为1700 m/s, 月面羽流侵蚀深度为2 cm。之后, 文 献 [15-16] 对月面月尘侵蚀质量速率进行了修正, 得到仿真计算与月面着陆分析结果的尺度系数为 2.5×10⁻⁹ kg/(m²·s)。Liever 等^[17] 提出了一种基于统 一流动求解器(UFS)的发动机羽流真空环境下的扬 尘特性研究方法,将此类研究的模拟能力扩展为两 相欧拉-拉格朗日方法,用于直接模拟月面尘埃的 侵蚀和坑化过程。He 等^[18] 基于 DSMC 研究了发动 机羽流作用下月尘颗粒的动力学特性,发现在羽流 气体密度和速度较大的区域,粒子速度可以被加速 到更大。Rahimi 等^[19] 使用 CFD 方法验证了 Apollo 探月任务中羽流气体在真空中的扩散侵蚀特性,并 与 DSMC 结果进行了比较, 验证了 CFD 方法对发 动机羽流扬尘效应研究的适用性,得出悬停高度不 会影响粒子最大速度随直径增加而减小的特性。

嫦娥五号作为中国首次月尘采样返回探月任 务,定量研究整个降落过程中羽流对月面的侵蚀特 性和月尘颗粒的动力学特性,有助于为后续探月任 务羽流月尘防护提供设计参数输入。但目前,国内 外只有对 Apollo 大推力喷管羽流扬尘特性的仿真 研究,还未有关于嫦娥五号小推力单喷管的羽流扬 尘特性分析,无法对这种特定任务中的月尘侵蚀扰 动特性进行研究。基于此,本研究采用 CFD 方法 对降落过程中距离月面较近区域进行了羽流扬尘 特性研究,研究嫦娥五号探测器着陆最后时刻的羽 流扬尘过程特性,为后续相关任务中敏感器件的月 尘防护提供可靠数据。

1 羽流扬尘模型

发动机羽流对月面的扰动侵蚀特性主要包括 3个阶段:发动机羽流气体喷管内流动、月面真空 环境下羽流扩散、月尘颗粒扬起运动。3个阶段均 采用 CFD 方法进行仿真模拟,使用有限体积法 (FVM)将计算区域划分为有限大小的子空间,通过 近似质量、动量和能量方程对羽流场进行计算。

1.1 发动机羽流气体喷管内流动模型

根据嫦娥五号发动机制导与导航系统的遥测数据可知, 嫦娥五号在最终降落羽流扬尘阶段的发动机推力约为 2 900 N^[20]。本研究第一阶段的发动 机羽流气体喷管内流动整个仿真阶段的发动机推 力输入均选择为 2 900 N, 模型选择压力入口和出 口, 入口压力为 0.25 MPa, 出口压力为 1.0 Pa, 喷管 气体介质选择水蒸气^[21-23], 仿真模型采用二维轴对称模型进行计算。

1.2 月面真空环境下羽流扩散模型

为了使仿真过程更加接近实际情况,月面真空 环境下的羽流扩散流动计算模型入口参数选取发 动机喷管出口的仿真计算结果,出口压力为 1.0 Pa, 模型采用以发动机喷管中心为对称轴的二维轴对 称模型,模型结构如图 1 所示,r 为发动机喷管的出 口半径。羽流气体选择可压缩的水蒸气。为了更 好地描述理想气体的物理特性,气体黏度选择萨瑟 兰黏性定律。



vacuum environment

1.3 月面黏性理论

月面侵蚀是指月面风化层在发动机羽流作用 下被吹起的过程。在此过程中风化层被吹起的速 率为侵蚀速率,被扰动的尺度被称为侵蚀半径。由 于真空的存在,使得月面风化层的侵蚀机制主要为 黏性侵蚀,其作用原理为风化层受到的剪切应力大 于自身的临界黏附强度。通过多余剪切应力,得出 质量侵蚀速率为^[19]

$$\frac{1}{2}au\dot{m} = \tau - \tau_{\rm c} \tag{1}$$

式中: a 为月尘颗粒可以从羽流气体中获得的速度 系数; m 为喷管出口羽流气体质量流量; u、r 分别 为羽流气体作用于月面风化层速度、剪切力; r。为 风化层可以承受的临界剪切力, 计算式为^[24]

$$\tau_{\rm c} = \tau_{\rm coh} + P_{\rm s} \tan \alpha \tag{2}$$

式中: τ_{coh} 为风化层内聚力; P_s为羽流作用在月面的 正压力; α 为风化层摩擦角。

根据风化层样品的力学特性研究^[24-25], 在风化 层表面的内聚力和摩擦角分别为 100 Pa 和 30°。月 尘颗粒特性 *a*、发动机羽流场特性和发动机距离月 面高度的综合参数ζ计算式分别为^[19]

$$a = \left[\frac{1}{2} + \sqrt{\left(\frac{1}{4} + \frac{1}{\zeta}\right)}\right]^{-1} \tag{3}$$

 $18\mu h$

$$\zeta = \frac{10\mu_{\rm c}n}{\rho_{\rm p}\sqrt{RT_{\rm c}(4+k_{\rm hyper})}} \times \left[\frac{1}{d_{\rm p}^2} + \frac{1}{d_{\rm p}} \times \frac{(4+k_{\rm hyper})C_{\rm d}}{72e\sqrt{2RT_{\rm c}}} \times \frac{F_{\rm Thrust}}{\mu_{\rm c}h^2}\right]$$
(4)

式中: μ_c 、 T_c 分别为喷管出口黏度、温度;h、 ρ_p 、 d_p 分别为喷管高度、颗粒密度、风化层颗粒直径; R为气体常数; C_d 为颗粒阻力系数, C_d =0.2; F_{Thrust} 为 发动机推力, F_{Thrust} =2 900 N; K_{hyper} 为发动机的高超 声速参数,计算式为

$$K_{\text{hyper}} = \gamma \ (\ \gamma - 1)Ma^2 \tag{5}$$

式中: y 为气体比热比; Ma 为喷管出口马赫数。

1.4 月尘颗粒扬起运动特性

沉积于月面风化层的月尘颗粒受到发动机羽 流扰动后会离开月面,以悬浮态的形式存在于月球 空间,造成着陆过程中的视线障碍。在月尘离开月 面时,月尘颗粒在发动机羽流场中会受到来自羽流 场的曳力、升力和月球重力的影响,在这3个力的 合力作用下,月尘颗粒会随羽流场运动。假设月尘 颗粒为光滑的球形颗粒,对于单个月尘颗粒的受力 F为

$$F = F_{\rm L} + F_{\rm D} + mg \tag{6}$$

式中:g为月球重力加速度,g=1.63 m/s²;F_L为颗粒 受到的升力;F_D为颗粒受到的曳力,单位质量上的 月尘颗粒曳力可以表示为

$$F_{\rm D} = \frac{18\mu}{\rho_{\rm n} {d_{\rm p}}^2} \cdot \frac{C_{\rm D} Re}{24}$$
(7)

式中: μ 为羽流气体黏度; $C_{\rm D}$ 为曳力系数;Re为羽流 流场雷诺数。根据嫦娥五号月尘采样分析数据可 知 $\rho_{\rm p}$ =3 195.2 kg/m³^[26]。雷诺数和曳力系数计算式 分别为

$$Re = \frac{\rho d_{\rm p} \left| u_{\rm p} - u_{\rm g} \right|}{\mu} \tag{8}$$

$$C_{\rm D} = a_1 + \frac{a_2}{Re} + \frac{a_3}{Re^2}$$
(9)

式中: u_p为月尘颗粒速度; u_g为羽流气体速度; a₁、 a₂和 a₃分别为曳力系数的相关计算常数,随着 Re的变化而变化^[27]。根据以上颗粒的受力,可以获 得发动机距离月面不同高度处的羽流场月尘运动 特性。

2 仿真结果及分析

2.1 发动机喷管内羽流气体流动特性及验证

嫦娥五号发动机喷管在 2 900 N 推力条件下, 羽流气体在喷管出口处的流动特性是引起月面尘 埃运动的直接因素。发动机内喷管流动特性云图 如图 2 所示。喷管入口处流体的流动速度低,在喷 管喉部,羽流受到其结构限制,被压缩形成压缩 波。经过喉部后,流体随着喷管结构的增大出现膨 胀波,膨胀过程中羽流内能转化成动能,导致喷管 内气体在出口处速度增大、温度降低。在喷管壁面 处存在速度和温度的边界层,此区域内的羽流处于 相对静止状态,使得速度降低、温度升高。因此,在 喷管内整个流动过程中,速度和温度的变化趋势 相反。

为了验证仿真结果的正确性,将喷管出口处参数代入推力F计算式

$$F = mv_{\rm e} - A_{\rm e}(P_{\rm e} - P_{\rm a}) = 0.93 \times 3\ 032 + \pi \times 0.401\ 5^2 \times (175.8 - 1) = 2\ 908.24\ {\rm N}$$
(10)

式中: v_e、P_e分别为喷管出口羽流气体速度、压力; A_e为喷管出口面积; P_a为环境压力。由式(10)计算 结果可知,喷管内气体流动结果符合 2900 N 推力 条件。

除了发动机推力验证外,还可将仿真结果得到 的出口截面羽流流动参数与地面实验测试结果进



Fig. 2 Cloud diagram plume gas flow contours in engine nozzle

流场模拟。

2.2 羽流真空扩散特性研究

将第一阶段喷管出口气体流动参数带入图1 模型中,通过数值仿真方法得到发动机距离月面

根据克努森数 Kn,可以判断流动连续性(Kn<

8

6

4 ½

2

0

0.5、1.0、1.5、2.0 m 时的羽流场流动特性,用于研究

0.1 时连续)。不同高度下羽流场的 Kn 云图如图 3 所示。Kn 均小于 0.1, 可以确认 CFD 软件适用于 2 900 N 推力条件下,发动机高度 2.0 m 以下的羽流

羽流气体在真空和低重力环境下的流动扩散。

行对比,如表1所示。由表1可知,马赫数、压力、 温度和速度等参数与地面实验结果相近,说明仿真 结果的可靠性。

表 1 发动机喷管出口流动参数仿真与实验结果对比 Table 1 Comparison of engine nozzle outlet flow parameters between simulation and ground test

方法	马赫数/Ma	压力/Pa	温度/K	速度/(m·s ⁻¹)
地面实验	5.22	170.02	744.69	3 218.56
仿真计算	4.93	175.83	802.32	3 032.12





不同喷管高度下羽流场的速度分布云图如 图 4 所示。喷管内的气体一旦离开喷管,在真空环 境下会迅速膨胀扩散,0.5m高度以上均形成了类 似碗状的冲击波流动。发动机喷管距离月面高度 越大,此区域的扩散越充分,碗状的冲击波形状更 明显,在其边缘有低速的边界层。在喷管的正下方 月面处出现了低速的压缩滞止区(云图中蓝色部 分),在此区域内羽流的速度远小于喷管出口下方 冲击波速度。羽流除了垂直向月面的扩散外,还包 括沿径向的扩散,最终达到充分流动扩散。由于壁 面的无滑移造成了速度的减小,羽流速度的最小值 出现在喷管外壁面和月球表面,形成了低速边界 层。喷管出口处的羽流速度最大,喷管正下方的滞 止区羽流速度最小,随后在充分扩散区域中心羽流 的速度较大。因此,可以通过速度得出羽流的流动 区域分布。

压力场是分析月面风化层受羽流扰动侵蚀位 置的重要指标。羽流在不同高度下的压力分布云 图如图 5 所示。压力场包括垂直于月面的射流和 沿着月面的径向扩散两部分,从喷管出口到正下方 月面,压力以垂直于月面为主导,到达月面后,以径 向扩散为主导。由射流引起的喷管正下方滞止区 压力最大,羽流速度最低,压力和密度最高,这是喷 管羽流强冲击月面的结果。此滞止区的边界随着 喷管高度的增加而增加,但压力值随着高度增加不 断减小,这是由于发动机喷管距离月面越高,羽流 在到达月面前扩散时间越多,扩散越充分,最终能 够到达月面的羽流气体分子比低高度时少,但引起 的月面风化层侵蚀区域范围越大。发动机喷管距 离月面越近,羽流气体离开喷管出口后扩散时间 少,主要以垂直月面扩散为主,径向扩散极少,大量 的羽流气体分子作用于月面后形成强的侵蚀效应, 月面的侵蚀区域范围较小。

不同发动机喷管高度下密度的径向分布如 图 6 所示。发动机喷管中心轴线对应的月面区域 密度值最高,随着径向距离的增加,密度值逐渐降 低。当发动机距离月面较近时,羽流气体会在垂直 月面方向膨胀,而距离月面较远时,羽流气体在垂 直方向和径向方向均会膨胀,但径向的膨胀更明 显。当发动机距离月面1.0 m以下时,在发动机喷 管中心轴线至发动机喷口半径区域的正下方出现 了密度的径向最大保持趋势,验证了发动机低高度





时的滞止区压力和密度的一致性,在此区域内羽流 气体的密度和速度较高,使得月尘颗粒可以迅速加 速。当发动机距离月面高度大于 1.0 m 时,喷管几 何尺寸对应的月面区域无高密度保持区域,说明在 1.0 m 以上距离后,羽流气体的径向扩散大于垂直 方向的扩散,使得月面侵蚀特性相对 1.0 m 以下高 度时弱,密度在月面的分布特性与图 5 中的压力云 图分布特性一致。

2.3 月面黏附侵蚀特性研究

发动机羽流引起的月面风化层的侵蚀半径和 侵蚀速率是评价不同推力、不同高度下羽流破坏力 的重要参数,计算结果同时可以验证仿真结果与实 际降落图像结果的差异。根据嫦娥五号任务中的 采样结果,着陆区域的风化层平均粒径为50 μm^[26], 故在计算侵蚀质量速率过程中选用此粒径的颗 粒。不同降落高度下的月尘质量侵蚀速率如图 7 所示。不同高度下获得的月尘质量侵蚀速率随着 探测器的高度降低而增大,在0.5 m时达到最大,且 此高度下的风化层侵蚀质量速率增加最大。侵蚀 速率的峰值点随着降落高度的降低逐渐向喷口中 心移动,当喷管距离月面高度为2.0m时,峰值点位 于距离喷管中心轴线0.8m的位置;而当喷管高度 为0.5m时,峰值点位于0.6m处,峰值点的移动反 应了羽流随着高度降低,扩散程度减小,此结果与 图4、图5的结果相同。通过分析嫦娥五号降落图 像计算结果,得出喷管高度为1.54m时的平均侵蚀 速率为1.496 kg/(m²·s),而仿真计算获得的1.5m高 度下的侵蚀速率为1.402 kg/(m²·s),与图像计算分析 结果非常相近,也验证了仿真结果的准确性。

2.4 月尘颗粒扬起特性研究

月面月尘颗粒被羽流侵蚀过程中的羽流扬尘 速度和高度取决于月尘颗粒所处的径向位置及自 身的尺寸,月尘粒径越大,越不容易被羽流扬起。 发动机在不同高度下,粒径为1μm月尘颗粒的运 动轨迹如图8所示。颗粒的初始位置(距离喷管中 心轴线的水平距离 r)为0.5 m,颗粒间距为0.5 m。 月尘颗粒初始位置距离喷管喷口越近,在发动机高 度相同时被扬起的高度越高。初始位置远离喷管



图 5 不同降落高度下羽流场压力云图

Fig. 5 The plume gas pressure contours at different engine altitudes



图 6 不同降落高度下月面羽流密度分布特性 Fig. 6 The density profile of lunar surface at different landing altitudes

时,月尘颗粒会沿着月面运动一段距离后被扩散的 羽流扬起。在整个流动区域边界处,距离月面高度 1.5 m时,初始位置越接近发动机喷管中心,被扬起 的高度越大,约为0.7 m。高度为0.5、1.0 m的羽流 气体由于没有充分的径向扩散时间,径向流动较 弱,使得远离发动机喷管中心的月尘颗粒没有被扬





起到很大高度(约0.5 m), 而高度为1.5、2.0 m 的羽 流具有较充分的径向扩散时间, 径向的流动较强, 颗粒会随羽流气体运动到距离月面较高的距离, 这 一现象与图 4~图 6 中的流动扩散现象符合。当发 动机距离月面高度为2.0 m 时, 初始位置为2.5、3.0 m 的月尘颗粒没有随羽流离开月球表面, 而是贴着月 面滑移了一段距离, 这是因为在2.0 m 高度时, 羽流



图 8 不同降落高度下粒径为 1 μm 颗粒运动轨迹 Fig. 8 The trajectories of 1 μm lunar dust particles at different engine altitudes

在真空环境中有了较充分扩散,在其边界处形成了 较强的压缩波,使得此区域的月尘颗粒受到压缩波 的影响无法离开月球表面。

不同喷管高度下粒径为70 µm 的月尘颗粒运 动轨迹如图9所示。由图9与图8对比可知,相同 发动机高度下,70 μm的月尘颗粒在羽流场边界处 的高度低于1μm颗粒的高度,且70μm月尘颗粒 的运动轨迹更接近直线。由图 9(a) 可知, 当高度为 0.5 m 时,初始位置为2.5、3.0 m 的月尘颗粒未被扬 起, 且在图 5(a) 中的相同位置存在明显的压力边界 层,使得此区域内的月尘颗粒受到极大的压力而无 法离开月面。图 9(a)和图 9(b)中月尘运动轨迹在 流场边界处有卷动, 而在图 9(c)和图 9(d)中未发 现,也可以说明这2个高度下流场的径向扩散强。 通过对比图 8 和图 9 可知,月尘粒径越小,越容易 受到羽流场的流动特性影响,月尘被扬起的高度也 越高。月尘粒径较大时,羽流场的压缩波不会使得 月尘颗粒出现卷动运动,而是以相对直线的运动轨 迹运动。嫦娥五号任务中对于1、70 µm 的月尘颗 粒最大扬起高度分别为 0.717、0.363 m, 70 μm 的月 尘颗粒最大扬起高度约为1 µm 的 1/2。

不同高度下 1 μm 月尘颗粒运动速度随月面径 向距离的变化如图 10 所示。由图 10(a) 和图 10(b) 可知, 1 μm 的月尘颗粒在发动机高度为 0.5、1.0 m 时的速度特性相似,从初始位置开始,速度先增大, 随后保持不变,在流场的出口区域速度继续增大。 由图 10(c)可知,在发动机高度为 1.5 m时,初始位 置为 0.5、1.0 m的月尘速度先增大,随后保持,最后 减小;初始位置为 1.5 m的月尘速度先以较小的速 度保持,随后随羽流运动增大后保持,在流场出口 处有增大的趋势。其保持较小速度状态是由于未 离开月面,在径向方向滑动或滚动所致,具体如 图 8(c)所示。由图 10(d)可知,在发动机高度为 2.0 m 时,初始位置为 1.5、2.0 m 的颗粒未被扬起,所以速 度也极小。在整个羽流侵蚀过程中,1 µm 的月尘颗 粒最大速度为 2 520 m/s。

70 μm 的月尘颗粒在不同高度下的速度随径向 距离的变化情况如图 11 所示。由图 11 与图 10 对 比可知,粒径为 70 μm 的月尘速度变化更加均匀, 没有出现速度尖点,在距离喷管较近的位置处快速 增大,随后速度增量减小,这是由于在喷管附近羽 流的密度较大,对颗粒的加速明显,随着羽流在径 向的扩散,羽流气体的密度衰减,对月尘颗粒的加 速作用逐渐减弱。初始位置在 0.5 m 时粒径为 70 μm 颗粒的最大速度约为 920 m/s。Lane 等通过 CFD 仿 真方法也研究了 Apollo 探月任务中月面羽流扬尘 速度特性,根据仿真计算得出探测器距离月面 2.0 m高度时, 70 μm 月尘颗粒在距离发动机喷管轴线







中心 4.0 m 位置处的颗粒速度约为 800 m/s。嫦娥 五号任务中相同高度处 70 μm 颗粒在距离发动机 喷管轴线中心 4.0 m 位置处的颗粒速度约为 780 m/s, 略低于 Apollo 探月任务中的颗粒速度。





除了月尘颗粒的运动速度和运动轨迹外,月尘 颗粒的羽流扬尘角也是评判月尘污染的一个重要 参数。嫦娥五号探测器在距离月面 2.0 m 以下范围 时月尘颗粒的最大羽流扬尘角如表2所示。距离 月面高度小于1.0m时,1µm的月尘颗粒羽流扬尘 角为 0.9°且小于 30、70、100 μm 月尘颗粒对应的高 度角。当探测器距离月面为 1.0~2.0 m 时, 1 μm 月



Table 2 Lunar dust elevation angle at different nozzla altitudas

nozze annudes					
高度/m	粒径/μm	扬尘角/(°)	平均扬尘角/(°)		
	1	0.90			
0.5	30	1.45	2.24		
0.5	70	2.24	2.24		
	100	2.63			
	1	0.49			
1.0	30	0.66	1.44		
1.0	70	1.44	1.44		
	100	1.84			
	1	3.17			
1.5	30	2.27	2.27		
1.5	70	2.27	2.27		
	100	2.25			
	1	2.82			
2.0	30	2.05	2.02		
2.0	70	2.02	2.02		
	100	2.02			

尘颗粒的羽流扬尘角大于其他粒径的月尘颗粒。

在 Apollo 探月任务中, 月尘采样返回样品的平 均粒径为 70 µm^[28],除 Apollo15 探月任务中得到的 扬尘角为 8.1°外,其他扬尘角的范围为 1.4°~2.6°, 如表3所示^[10]。月面坡度角的主要影响因素包括 月面坡度角及微地形(石块、凹坑、凸起等),其中 月面坡度角影响最大, Apollo 15 探月任务中的月尘 扬尘角之所以远大于其他任务,是因为 Apollo 15 探月任务探测器着陆地区的月面坡度角约为 11°,导致扬尘角增大。嫦娥五号任务中得到扬尘 角范围为 1.44°~2.27°, 与 Apollo 探月任务中月尘 扬尘角近似。

表 3 Apollo 和嫦娥五号探月任务扬尘角对比

Table 3 Elevation angle comparison in Apollo and Chang'E-5 lunar exploration missions

任务名称	扬尘角/(°)	任务名称	扬尘角/(°)
Apollo 11	2.6	Apollo 14	2.4
Apollo 15	8.1	Apollo 16	1.4
Apollo 17	2.0	嫦娥5号	1.44~2.27

结 论 3

月尘污染是月球探测任务中的关键影响因素, 也是最难克服的环境因素。本文研究了2900N发

动机推力条件下,羽流对月面侵蚀及月尘颗粒的动 力学特性。可以得出以下结论:

 建立了嫦娥五号探测器发动机仿真模型和 月面羽流侵蚀仿真模型,采用两阶段法对流场进行 了数值仿真。根据发动机喷管出口的仿真参数计 算出发动机推力约为2900N,验证了仿真的正确 性。同时通过仿真计算得出的侵蚀速率与降落图 像的计算结果相近,也验证了仿真的准确性。

2)将发动机喷管出口的羽流仿真参数代入月 面羽流侵蚀流动模型中,计算了发动机距离地面高 度 0.5、1.0、1.5、2.0 m时的羽流扩散流动特性。根 据羽流速度和压力云图可以得出在喷管出口处羽 流速度最大,羽流气体离开喷管后沿垂直月面方向 和平行月面方向流动扩散。发动机喷管距离月面 越低,喷管出口羽流的垂直月面方向扩散越强,平 行月面方向扩散越弱,导致月面侵蚀速率峰值向喷 管中心轴线移动。在喷管正下方形成羽流滞止区, 此区域内的月尘颗粒会沿着月面滚动或滑动,直到 滞止区结束后月尘颗粒会随羽流离开月面。

3) 发动机距离月面 1.0 m 以下时, 月尘颗粒在 径向运动过程中的高度和速度小于 1.5、2.0 m 高度 的。对于同一高度, 距离喷管中心轴线近的月尘颗 粒获得更大的速度和扬起高度。粒径为 1、70 μm 的月尘颗粒最大扬尘高度分别为 0.72、0.36 m, 最 大速度分别为 2 520、1 010 m/s。平均扬尘角范围 为 1.44°~2.27°, 与 Apollo 探月任务图像分析结果 非常接近。

本文重点针对嫦娥五号任务的羽流扬尘特性 进行了研究。未来中国载人登月中发动机推力为 80 000 N,比 Apollo 探月任务中使用的推力更大,降 落过程中会溅起更多月尘。极大影响航天员视线, 航天员出舱过程中容易吸入月尘造成呼吸道疾病, 黏附在宇航服表面的月尘会造成宇航服泄漏,极大 危害航天员生命安全。下一步着力研究载人探月 过程中羽流扬尘特性,为航天员在轨安全提供 保障。

致 谢 感谢中国国家航天局提供的嫦娥五 号降落图像。感谢甘肃省女科技工作者协会的项 目支持。感谢王永军高级工程师对论文撰写过程 中的指导及修改。

参考文献(References)

- [1] GAIER J R, HICKS M C, MISCONIN R M. Studies of simulated lunar dust on the properties of thermal-control surfaces[J]. Journal of Spacecraft and Rockets, 2013, 50(4): 848-852.
- [2] JOHNSON C L, DIETZ K L. Effects of the Lunar Environment on Optical Telescopes and Instruments [C]//Proceedings of the Space Astronomical Telescopes and Instruments. Orlando: International

Society for Optics and Photonics, 1991: 208-218.

- [3] PIRICH R, WEIR J, LEYBLE D, et al. Effects of the lunar environment on space vehicle surfaces [C]//2010 IEEE Long Island Systems, Applications and Technology Conferenc. Piscataway: IEEE, 2010: 1-6.
- [4] CAIN J R. Lunar dust: the hazard and astronaut exposure risks[J]. Earth, Moon, and Planets, 2010, 107(1): 107-125.
- [5] MORRIS A B, GOLDSTEIN D B, VARGHESE P L, et al. Plume impingement on a dusty lunar surface [C]//AIP Conference Proceedings. California: AIP, 2011: 1187-1192.
- [6] IMMER C, LANE J, METZGER P, et al. Apollo video photogrammetry estimation of plume impingement effects[J]. Icarus, 2011, 214(1): 46-52.
- [7] METZGER P T, IMMER C D, DONAHUE C M, et al. Jet-induced cratering of a granular surface with application to lunar spaceports[J]. Journal of Aerospace Engineering, 2009, 22(1): 24-32.
- [8] BÜHLER C A. Experimental investigation of lunar dust impact wear[J]. Wear, 2015, 342/343: 244-251.
- [9] DONOHUE C M, METZGER P T, IMMER C D. Empirical Scaling Laws of Rocket Exhaust Cratering [EB/OL]. [2021-04-12]. http://arxiv.org/abs/2104.05176.pdf.
- [10] LANE J E, METZGER P T. Estimation of Apollo lunar dust transport using optical extinction measurements[J]. Acta Geophysica, 2015, 63(2): 568-599.
- [11] METZGER P T, LANE J E, IMMER C D, et al. Scaling of erosion rate in subsonic jet experiments and Apollo lunar module landings [C]//Earth and Space 2010. Reston, VA: American Society of Civil Engineers, 2010: 191-207.
- [12] LANE J E, METZGER P T, CARLSON J W. Lunar dust particles blown by lander engine exhaust in rarefied and compressible flow [C]//Earth and Space 2010: Engineering, Science, Construction, and Operations in Challenging Environments. Reston, VA: American Society of Civil Engineers, 2010: 134-142.
- [13] LANE J E, METZGER P T, IMMER C D, et al. Lagrangian trajectory modeling of lunar dust particles [C]//Earth & Space 2008. Reston, VA: American Society of Civil Engineers, 2008: 1-9.
- [14] MORRIS A B, GOLDSTEIN D B, VARGHESE P L, et al. Plume impingement on a dusty lunar surface[J]. AIP Conference Proceedings, 2011, 1333(1): 1187-1192.
- [15] MORRIS A B, GOLDSTEIN D B, VARGHESE P L, et al. Modeling the interaction between a rocket plume, scoured regolith, and a plume deflection fence [C]//Earth and Space 2012: Engineering, Science, Construction, and Operations in Challenging Environments. Reston, VA: American Society of Civil Engineers, 2012: 189-198.
- [16] MORRIS A B, GOLDSTEIN D B, VARGHESE P L, et al. Approach for modeling rocket plume impingement and dust dispersal on the moon[J]. Journal of Spacecraft and Rockets, 2015, 52(2): 362-374.
- [17] LIEVER P, TOSH A, ARSLANBEKOV R, et al. Modeling of rocket plume impingement flow and debris transport in lunar environment [C]//Proceedings of the 50th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition. Reston, Virigina: AIAA, 2012: 800.
- [18] HE X Y, HE B J, CAI G B. Simulation of two-phase plume field of

liquid thruster[J]. Science China Technological Sciences, 2012, 55(6): 1739-1748.

- [19] RAHIMI A, EJTEHADI O, LEE K H, et al. Near-field plume-surface interaction and regolith erosion and dispersal during the lunar landing[J]. Acta Astronautica, 2020, 175: 308-326.
- [20] ZHANG H H, LI J, WANG Z G, et al. Guidance navigation and control for Chang'E-5 powered descent[J]. Space:Science and Technology, 2021, 2021: 9823609.
- [21] LI Y, REN D P, BO Z G, et al. Gas-particle two-way coupled method for simulating the interaction between a rocket plume and lunar dust[J]. Acta Astronautica, 2019, 157: 123-133.
- [22] BERG J J, GOLDSTEIN D B, VARGHESE P L, et al. DSMC simulation of europa water vapor plumes[J]. Icarus, 2016, 277: 370-380.
- [23] TSENG W L, LAI I L, IP W H, et al. The 3D direct simulation Monte Carlo study of Europa's gas plume[J]. Universe, 2022, 8(5):

261.

- [24] SLYUTA E N. Physical and mechanical properties of the lunar soil (a review)[J]. Solar System Research, 2014, 48(5): 330-353.
- [25] COLWELL J E, BATISTE S, HORÁNYI M, et al. Lunar surface: dust dynamics and regolith mechanics[J]. Reviews of Geophysics, 2007, 45(2): 1-26.
- [26] LI C L, HU H, YANG M F, et al. Characteristics of the lunar samples returned by the Chang'E-5 mission[J]. National Science Review, 2022, 9(2): 188.
- [27] MORSI S A, ALEXANDER A J. An investigation of particle trajectories in two-phase flow systems[J]. Journal of Fluid Mechanics, 1972, 55(2): 193.
- [28] MORRIS, A B. Simulation of rocket plume impingement and dust dispersal on the lunar surface[D]. Austin: University of Texas at Austin, 2012.

Nozzle plume erosion property on lunar dust in Chang'E-5 mission

ZHANG Haiyan^{1, 2}, LI Sixin¹, WANG Yi², LI Cunhui², ZHANG Xiaoping³, WANG Weidong^{1, *}

(1. School of Mechano-Electronic Engineering, Xidian University, Xi'an 710071, China;

Science and Technology on Vacuum Technology and Physics Laboratory, Lanzhou Institute of Physics, CAST, Lanzhou 730000, China;
 State Key Laboratory of Lunar and Planetary Sciences, Macau University of Science and Technology, Macau 999078, China)

Abstract: During a lander landing, the plume-lunar surface interaction induces lunar dust ejection, which is the main reason for lunar dust hazards. The study adopts computational fluid dynamics(CFD) method to build a one-toone nozzle model and vacuum plume flow and diffusion model, through which the lunar dust erosion mass is investigated, and the lunar dust trajectory, ejection angle and velocity are obtained when nozzle altitude is from 0.5 to 2.0 m. The results show that the maximum mass erosion rate is 8.83 kg/m2·s and this value decreases with nozzle altitude increase, which also is consistent with landing photo results in the Chang'E-5 mission. For lunar dust kinetic properties, the maximum velocity for 1 μ m and 70 μ m particles are 2 520 m/s and 1 010 m/s respectively, the maximum height for 1 μ m and 70 μ m particles are 0.72 m and 0.36 m respectively. The dust ejected angle ranges from 1.44° to 2.27°. The ejected angle calculated in Chang'E-5 mission is similar to that in Apollo mission.

Keywords: lunar dust; Plume ejection property; Chang'E-5 mission; computational fluid dynamics simulation; Euler-Lagrange method

Received: 2022-05-31; Accepted: 2022-09-03; Published Online: 2022-09-14 13:04 URL: link.cnki.net/urlid/11.2625.V.20220913.1828.011

Foundation items: National Natural Science Foundation of China (42004157); National Key Laboratory of Materials Behaviors and Evaluation Technoloty in Space Environments (WDZC-HGD-2022-08); National Key Laboratory of Science and Technology on Vacuum Technology and Physics, China (6142207200202); Natural Science Foundation of Gansu Province, China (21JR7RA743);The Youth Talent Promotion Project of Gansu Province (GXH20210611-05)

^{*} Corresponding author. E-mail: wangwd@mail.xidian.edu.cn

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

DOI: 10.13700/j.bh.1001-5965.2022.0451

无参数自适应罚函数的高效代理模型优化设计方法

张伟¹, 高正红^{1,2,*}, 王超³, 夏露^{1,2}

(1. 西北工业大学 航空学院,西安 710072; 2. 西北工业大学 翼型、叶栅空气动力学国家级重点实验室,西安 710072;

3. 北京电子工程总体研究所,北京 100854)

摘 要:在飞行器气动外形优化设计中,复杂约束条件导致设计空间可行域呈现不连续的 特征,且理想解大多靠近约束边界,传统高效代理模型方法难以适用。研究了参考点对优化设计的 影响,提出了一种考虑约束的参考点选择机制;对于最优解靠近边界的问题,罚函数法更加有效, 但惩罚因子的设置对于罚函数方法影响很大,不合适的惩罚因子反而会损害优化效率,分析了优化 过程中罚函数方法对惩罚因子的要求,提出了一种无参数自适应罚函数的代理模优化设计方法,引 入基于样本分析的惩罚项,结合归一化目标值和约束值,在优化过程中动态调整惩罚因子,使优化 能够尽可能地聚焦于可行域内,迅速收敛到最优解,实现样本的高效配置。通过带约束的函数算例 和翼型优化算例证实,所提方法可以大幅提高飞行器气动外形优化设计效率。

关键 词:优化设计;代理模型;参考点;气动设计;自适应罚函数

中图分类号: V211.3

文献标志码: A 文章编号: 1001-5965(2024)04-1262-11

基于 Kriging 模型的高效全局优化 (efficient global optimization, EGO)^[1-3]方法可以提高优化设计效率,因此在工程设计领域得到广泛应用^[4-7]。 EGO方法将优化搜索与代理模型加点重构相结合,根据 Kriging 模型对目标函数预测的标准差,以当前最优样本点为参考点,通过优化目标的改善期望 (expectation improvement, EI)函数以获取新的样本点^[8]。EGO方法同时考虑了代理模型的预测值以及预测的不确定性影响,加强和平衡了全局和局部 探索能力,因此具有更高效的全局优化能力^[9]。

飞行器气动外形设计问题往往面临诸多的约 束条件^[10-14],导致设计问题的可行域缩小且不连续, 给优化搜索带来困难,降低优化设计效率,甚至无 法获得理想的设计结果。因此,合理地处理目标与 约束,使优化聚焦于全局最优解附近的可行域是需 要解决的重要问题。 文献 [15-16] 基于约束函数的概率模型, 发展了 约束 EGO 方法, 通过将目标函数与约束函数相乘 的方式, 将多约束问题简化为无约束问题。Parr 等^[16] 应用约束 EGO 方法进行了考虑多约束的机翼外形 设计, 取得了理想的设计结果。但 Sasena 等^[15] 指 出, 随着优化的进行, 目标值的改善量逐渐减小, 此 时约束 EGO 方法中约束项的影响会逐渐增加乃至 占据主导, 这种改变会驱使优化远离约束边界位 置, 对于最优点靠近约束边界的问题, 约束 EGO 方 法难以快速找到理想设计结果。此外, 约束 EGO 方法在选择参考点时只考虑了目标值最优的样本, 却忽视了约束的影响, 导致了优化搜索过程无法集 中于可行域空间, 使优化设计效率下降。

针对这些问题, Basudhar 等^[17]利用支持向量机 (support vector machines, SVM)^[18]建立设计空间可 行域与设计变量之间的映射关系,将优化过程集中

ZHANG W, GAO Z H, WANG C, et al. Efficient surrogate-based aerodynamic optimization with parameter-free adaptive penalty function [J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50 (4) : 1262-1272 (in Chinese).

收稿日期: 2022-05-31; 录用日期: 2022-09-18; 网络出版时间: 2022-10-10 09:37

网络出版地址: link.cnki.net/urlid/11.2625.V.20221009.1124.011

基金项目: 翼型、叶栅空气动力学国家级重点实验室基金 (614220121020128)

^{*}通信作者. E-mail: zgao@nwpu.edu.cn

引用格式: 张伟,高正红,王超,等.无参数自适应罚函数的高效代理模型优化设计方法 [J]. 北京航空航天大学学报,2024,50(4):1262-1272.

于可行域内,提高优化搜索效率; Bagheri 等^[19] 通过 向 Kriging 模型引入噪声方差(noise variance)以及 在 EI 函数中进行参考点控制的方式,提高约束 EGO 方法对约束的处理能力,并在 NACA2412 翼 型气动优化设计中得到应用,提高优化设计效率。 此外,罚函数法也是一种理想的解决办法^[20-21]。罚 函数法通过引入惩罚因子,将约束函数转化为目标 的一部分^[21],将约束问题简化为无约束问题,在该 方法中,惩罚因子的选择对优化设计结果影响很 大,其设置大多源于设计者的先验知识[22-23]。但对 于复杂工程问题,先验知识的缺乏导致设计者通常 难以直接给出合适的惩罚因子。为此,自适应罚函 数方法应运而生,如王红涛等^[24]基于小生境微种群 遗传算法(micro generic algorithm, MGA)的锦标赛 选择因子原理,提出一种无惩罚因子的罚函数法, 并在跨音速翼型优化设计中取得了很好的效果;龙 腾等^[25]将引入自适应罚函数法将目标函数值和约 束值进行归一化处理,将约束问题转化为无约束问 题; Dong 等^[26] 将静态罚函数法与设计空间减缩技 术结合,使优化逐渐聚焦于可行域,以提高优化搜 索效率。以上方法虽然避免了惩罚因子设置的问 题,但引入了其他参数,同样为设计应用带来困难。

气动外形设计问题通常面临着几何、力矩等主 动约束,理想的设计结果大多位于可行域的边界 处,对于这类问题传统约束处理方法优化效率低 下。因此,本文针对传统约束 EGO 方法的参考点 选择问题,提出一种基于样本分析的参考点选择机 制,使优化过程集中于最优值附近的可行域,提高 优化设计效率。并针对罚函数法中惩罚因子的设 置问题,分析了优化各阶段目标与约束对于样本的 要求,提出一种基于无参数自适应罚函数的代理模 型优化方法。该方法综合考虑目标函数值和约束 的影响,选择参考点构建目标 EI 函数,结合对可行 解的评估,动态调整惩罚因子。通过带约束的函数 算例和美国航空航天学会 (AIAA) 气动优化设计讨 论组 (ADODG)^[27] 发布的标准翼型气动优化设计算 例以及飞翼布局翼型设计算例对本文提出方法进 行了验证。

1 约束 EGO 方法

Kriging 模型是一种基于统计理论的插值技 术, 给定样本集 $X = [x^1, x^2, \dots, x^n]^T$ 及其响应值 $Y = [y^1, y^2, \dots, y^n]^T$,则目标函数值与设计变量之间的真 实关系可以表示为

 $y(x) = g(x) + z(x) \tag{1}$

式中:g(x)为描述全局趋势回归模型;z(x)为静态随

机过程,用来描述真实样本与趋势函数之间的偏差。Kriging模型可以同时给出预测的均值和方差, 由此发展了高效的约束 EGO 方法。

对于考虑约束的一般性优化问题:

min
$$f(x)$$

s.t.
$$\begin{cases} g_j(x) \leq 0 & j = 1, 2, \cdots, k \\ x_i^{l} \leq x_i \leq x_i^{u} & i = 1, 2, \cdots, d \end{cases}$$
(2)

式中: f(x)为目标函数; g_j(x)为约束函数; x_i为设计 变量, xⁱ_i 和xⁱ分别为每个设计变量的上下边界, 所 有设计变量的边界所包含的区域构成整个设计空 间; d 为设计空间的维数; k 为约束的数量。

对于 Kriging 代理模型, 目标函数改善量的期 望 $E_t(x)$ 可通过式(3)计算:

$$E_{t}(x) = (y_{\min} - \hat{y})\Phi\left(\frac{y_{\min} - \hat{y}}{\sigma}\right) + \sigma\phi\left(\frac{y_{\min} - \hat{y}}{\sigma}\right) \quad (3)$$

式中: y_{min}为参考点,即当前最优样本点的目标函数 值; Φ为标准正态累积分布函数; φ为标准正态分布 概率密度函数; σ为预测的标准差; ŷ为代理模型预 测值。

针对约束函数分别建立 Kriging 模型, 将目标 EI 函数和约束概率函数相乘, 得到约束 EGO 函数 *C*{*E*_{*i*}(*x*)}, 约束优化问题简化为无约束问题。

$$C\{E_{i}(x)\} = E_{i}(x) \cdot \prod_{j=1}^{n} P[G_{j} \le 0]$$
(4)

由式(4)可以看出, $C{E_i(x)}$ 由目标函数改善量 $E_i(x)$ 和约束函数概率 $P[G_j \leq 0]$ 共同决定。若选择 的参考点未考虑约束, 当存在多个不连续可行域, 或可行域内的样本点目标函数值大于当前参考点 时, $E_i(x)$ 与 $P[G_j \leq 0]$ 会驱使算法进行随机搜索, 导 致 $C{E_i(x)}$ 函数失效。此外, $E_i(x)$ 项在靠近最优解 附近的区域时, 目标值改善量的减小会导致其失 效, 而 $P[G_j \leq 0]$ 项则会驱使算法远离可行域边界, 因此, 对于最优解靠边界的优化问题, 约束 EGO 方 法难以取得理想的结果。

气动外形设计问题往往需要考虑厚度、体积等 几何约束以及升力系数、力矩系数等气动约束,这 些约束大多为主动约束,即符合要求的理想解位于 约束边界处,由上述讨论可知传统约束 EGO 方法 难以适用于复杂气动设计问题,因此发展适用于气 动外形设计问题的约束处理方法是亟需解决的关 键问题。

2 参考点的选择对优化结果影响

由式(3)可知, EGO 方法目标函数值的提升依 赖于设计者选择的参考点以及 Kriging 模型提供 预测标准差。对于无约束优化问题, 一般选择当
前样本中具有最小目标函数值的样本作为参考 点。但针对约束优化问题,忽略约束的影响,仅考 虑目标函数值来选择参考点会导致无法搜索到最 优解。

本节利用二维约束 Branin 函数^[19] 优化算例研 究参考点的选择对优化效率的影响,函数表达式为

$$\min f(x) = \left(X_2 - \frac{5 \cdot 1}{4\pi^2} X_1^2 + \frac{5}{\pi} X_1 - 6\right)^2 + \left(10 - \frac{10}{8\pi}\right) \cos X_1 + \frac{X_1 + 35}{3}$$
$$X_1 = 15x_1 - 5, X_2 = 15x_2$$
$$\begin{cases}g(x) = -\left[\left(4 - 2 \cdot 1X_1^2 + \frac{X_1^4}{3}\right)X_1^2 + X_1X_2 + 6(4X_2^2 - 4)X_2^2 + 3\sin(6(1 - X_1)) + 3\sin(6(1 - X_2))\right] \\ 3\sin(6(1 - X_2))\right] \le 0$$
$$X_1 = 2x_1 - 1, X_2 = 2x_2 - 1 \tag{5}$$

式中: f(x)为原始 Branin 函数表达式,图 1 为示意 图,该函数在 $x_i \in [0,1]$, i = 1,2的设计空间内有 3 个 局部最优解。g(x)为约束函数,以Gomez#3 函数为 基础,添加了额外的正弦函数项以提高函数多峰 性,图 2 为约束函数示意图。



较之 Branin 函数,约束 Branin 函数解空间可 行域缩小,呈现不连续的特征;函数有2个局部最 优解和一个全局最优解。图3为约束 Branin 函数 示意图,其中,黑色实线内为可行域,实线外为不 可行域。红色圆点(0.9406,0.3171)为最优解位 置,最优值为12.0050。从图中可知,该函数有 3个不连续的可行域,最优解位于右下角可行域 边界处。

利用拉丁超立方设计(latin hypercube design,





Fig. 3 Contour of 2 dimensional constrained Branin function

LHD)方法生成2(*d*+1)=6个初始样本,其中*d*为设 计空间维数,构建 Kriging 模型。采用约束 EGO方 法进行优化,气动设计问题样本计算代价昂贵,优 化方法需要在有限次的迭代内获得最优解或者理 想解,因此,为了测试算法在有限次迭代步数内寻 得最优解的能力,算法终止条件有 2个:①目标函 数值与最优值误差小于 1%;②加点数达到 30次。 满足任一条件,算法停止。

图 4 和图 5 是选择目标函数最小值为约束 EGO 的参考点的优化结果。从图 4 可以看出,算法 在整个优化过程中并未找到最优解,且仅找到一个 满足约束的可行解。图 5 展示 *E_{1,max}*和 (*E₁×P₁*)_{max} 收敛历程, *P₁*表示约束函数,由图可知,在优化过程 中,约束 EGO 函数与 EGO 函数优化趋势基本相 同,但在优化后期, *E_{1,max}*和 (*E₁×P₁*)_{max} 变得非常小, 无法给优化方向进行有效的指导。

本文提出了考虑约束的参考点选择方法,表达 式为



式中: $x_{\text{feasibile}}$ 为可行域内的样本; $x_{\text{infeasibile}}$ 为不可行域 的解; $N_{x_{\text{feasibile}}}$ 为可行解的数量; N为总样本数; $f_{\text{max}} = \max(f(x)), g_{j,\text{max}} = \max(g_j(x)),$ 用于将样本的目标值 和约束值的归一化。

基于约束的满足与否,将其分成3类:①当样 本库内可行解数量超过总样本数量一半,则选择表 现最好的可行解作为参考点;②当仅存在不可行解 时,选择最靠近可行域边界的不可行解作为参考 点;当可行解数量少于总样本量一半时,将样本的 目标和约束值归一化处理,选择最好的样本点作为 参考点。该方法结合可行解的数量,自适应选择参 考点。在设计初期,样本库内不存在可行解时,选 择最靠近可行域边界的不可行解,使算法逼近全局 最优解所在的可行域,随着可行解的数增加,选 择表现最好的可行解作为参考点,提高优化设计 效率。

利用改进的约束 EGO方法对二维约束化 Branin 函数重新优化,优化参数设置与之前保持一 致。图 6 为改进约束 EGO方法目标函数收敛趋 势,从中可以看出,算法在优化过程中寻得 5 个可 行解,并在第 28 代得到全局最优解。从图 7 可知, 算法在优化过程中始终保持了 *E*_{*I*,max} 和 (*E*_{*I*}**P*_{*I*})max</sub>,对 优化方向起到正确的指导。图 8 为优化过程中参 考点目标值对比,可以看出,自第四次加点始,原始 约束 EGO方法的参考点的目标函数值比最优点更 小,已然无法为优化提供正确的指导,算法开始失 效,而考虑约束的参考点选择方法始终朝着满足约 束的目标值改善方向搜索,因此迅速收敛到了全局 最优解。



图 6 改进约束 EGO 方法目标函数优化趋势

Fig. 6 Improved constrained EGO method objective function optimisation trends



图 7 $E_{l_{j_{max}}}$ 和 $(E_l * P_l)_{max}$ 收敛历程 Fig. 7 Convergence of $E_{l_{max}}$ and $(E_l * P_l)_{max}$





reference points

3 无参数自适应罚函数的 EGO 优化 设计方法

对于最优解靠近约束边界的设计问题,通常基于 静态 罚 函数 的 EGO 方法 能够 比传统 约束 EGO 方法发挥更好的作用^[15]。然而,惩罚因子的选 择或依赖于设计者的先验知识,或需要反复尝试, 造成优化效率低下。鉴于此,本文提出一种无参数 自适应罚函数方法。

对于带约束优化问题式(2),罚函数法通过引 入惩罚因子将其转化为无约束问题:

$$\min F(x) = f(x) + \sum_{j=1}^{k} \lambda_j \max(g_j(x) - 0, 0)$$
 (7)

式中: F(x)为转换后的目标函数; λ_j 为每个约束对 应的惩罚因子, $j = 1, 2, \dots, k_o$

根据最小惩罚原则^[28-29],理想的惩罚项或惩罚 因子应该尽可能的小,同时刚好满足排除所有不可 行解的要求。

假设当前设计空间 $D \subseteq \mathbb{R}^d$, $f: D \to \mathbb{R}^d$ 为该空间 内任意函数, $g_{j,j=1,2,\cdots,k}: D \to \mathbb{R}^d$ 为设计空间内的约 束, 该函数在解空间存在唯一最优解 x^*_{feasible} , 并满足 $\{x^*_{\text{feasible}}\} < f(x_{\text{feasible}}): x_{\text{feasible}} \in D_{\text{feasible}}$, 其中 $D_{\text{feasible}} \subseteq \mathbb{R}^d$ 为可行域空间。此外,不可行域 $D_{\text{infeasible}} \subseteq$ \mathbb{R}^d 内存在样本点 $x^*_{\text{infeasible}}$, 该样本点满足 $\{x^*_{\text{infeasible}}\}$ $f(x^*_{\text{infeasible}}) < f(x_{\text{infeasible}}, 该样本点满足<math>\{x^*_{\text{infeasible}}\}$ $g_i(x^*_{\text{infeasible}}) < f(x_{\text{infeasible}}): x_{\text{infeasible}} \in D_{\text{infeasible}}\}$ 以及 $\{x^*_{\text{infeasible}}\}$ e -个极小的正实数 ε , 当每个惩罚项的惩罚系数 满足式(8)时,则其能够满足式(9), 使算法排除不 可行域内的不可行解, 快速获得得最优解。

可见,在优化初期,当不存在可行解或其数量 较少时,为了使算法快速找到可行解,应该设置较 大的惩罚系数以提高惩罚项对优化目标函数 F(x)的影响;而随着优化的进行,搜索到的可行解 逐渐逼近全局最优解x^{*}_{feasible},此时应提高样本目标 函数f的影响以提高优化收敛性和效率,因此应当 减小惩罚系数。

$$\lambda_{j} > \frac{f(x_{\text{feasible}}) + \varepsilon - f(x_{\text{infeasible}})}{g_{j}(x_{\text{infeasible}})} > \frac{f(x_{\text{feasible}}) + \varepsilon - f(x_{\text{infeasible}}^{*})}{g_{j}(x_{\text{infeasible}}^{*})}$$

$$(8)$$

$$x_{\text{feasible}}^{*} \in D_{\text{feasible}} , \quad x_{\text{infeasible}} \in D_{\text{infeasible}}$$

 $F(x_{\text{infeasible}}) > F(x_{\text{infeasible}}^*) > f(x_{\text{feasible}}^*)$ (9)

因此,本文提出了一种基于无参数自适应罚函 数方法的约束 EGO 方法,该方法结合设计过程中 可行解的数量变化,自适应调整惩罚因子。

首先,利用样本库内的目标函数和约束条件的 均值作为归一化因子,分别对目标函数和约束进行 归一化,避免人为设置惩罚因子,此时优化问题转 换为

min
$$F(x) = -\frac{E_i(x)}{f_{avg}} + \sum_{j=1}^k \frac{\max(g_j(x) - 0, 0)}{c_{j,avg}}$$
 (10)

式中: f_{avg} 为当前样本库内样本函数值的均值; $c_{javg} = \frac{1}{m} \sum_{i=1}^{m} \max(g_{i,j}(x_i) - 0, 0)$ 为样本库内所有样本的第 *i* 个约束条件的均值。

其次,由上述讨论可知,随着算法逼近可行域, 惩罚项应该逐渐减小,以提高优化效率,因此引入 式(11):

$$v(x) = \frac{N_{x_{\text{feasible}}}}{N} \sum_{j=1}^{k} \frac{\max(g_j(x) - 0, 0)}{g_{j,\text{avg}}}$$
(11)

在优化初期,通常不可行解数量占据较大比例,此时较大的v(x)使惩罚项充分发挥作用,随着可行解的数量逐渐增加,v(x)随之减小,使得目标项的作用更加凸显。

此时目标函数最终形式为

min
$$F(x) = -\frac{E_i(x)}{f_{avg}} + \sum_{j=1}^k \frac{\max(g_j(x) - 0, 0)}{g_{j,avg}} + v(x)$$

s.t. $x_i^l \le x_i \le x_i^u$ $i = 1, 2, \cdots, d$ (12)

利用二维约束 Branin 函数测试该方法的可靠性,优化设置与第2节保持一致。

由图 9 可知,自适应罚函数法在第 14 代寻得最 优解,明显优于传统约束 EGO 方法。结合样本分 析的动态罚函数策略,避免不连续的可行域空间导 致算法陷入局部最优,使其迅速收敛到最优解附近 的可行域,尤其对于最优解位于可行域边界的问 题,优化效率明显提高。



图 9 加点目标函数值 Fig. 9 Objective function values of infill samples

4 算例验证

4.1 函数测试算例

为了测试无参数自适应罚函数的约束 EGO 方法的效率和鲁棒性,本节选取 5 个经典非线性约束 函数 (G₁,G₃,G₅,G₆,G₂₄)^{30]}进行测试。这些函数维度 和约束特征各异,综合起来可以模拟诸多问题,函 数情况如表1所示。

Tab	le 1 I	Benchmark p	roblems for global optim	ization
函数	维度	约束数量	设计范围	最优解

表1 全局优化函数问题

G_1	13	9	$[0,1]^9 \times [0,100]^3 \times [0,1]$	-15
G_3	10	1	$[0, 10]^{10}$	-1.000 5
G_5	4	5	$[0, 1200]^2 \times [-0.55, 0.55]^2$	5 126.49
G_6	2	2	$[0, 10]^2$	-6 961.8
$G_{\rm 24}$	2	2	$[0,3] \times [0,4]$	-5.508

分别采用约束 EGO 方法(CEI)、基于无参数自 适应罚函数的约束 EGO 方法(adaptive penalty CEI, APCEI)和基于静态罚函数的约束 EGO 方法(penalty CEI, PCEI)进行优化对比。使用 LHD 方法在设计 空间内生成*d*+1个初始样本^[30],为了消除随机性, 每个问题重复 30次。每轮算法终止条件有 2个: ①目标函数值与最优值误差小于 1%; ②加点数达 到 200次。满足任一条件,算法停止。

表2展示了不同方法在30次重复测试寻到最优值时样本加点数的均值。">"表示重复测试存在未寻到最优解的情况。若30次皆寻到最优值,括号内为样本加点数的标准差;否则,括号内为未寻到最优解的次数。

可以看出,对于维度较高或者约束较多的 G₁、 G₃ 以及 G₅ 问题, CEI 方法无法在 200 次内寻到最 优解; PCEI 方法的优化结果比 CEI 更好,但无法保 证每次都收敛到最优解,鲁棒性较差; APCEI 方法 不但在所有重复测试中找到最优解,并且相较其余 方法,样本加点数更少,优化效率更高。对于维度 较低且约束较少的 G₆和 G₂₄问题,3种方法都可以快 速找到全局最优解,相较于 CEI和 PCEI 方法, APCEI 方法效率更高,鲁棒性更好。

	表 2 函数测试结	果
Table 2	Results of benchma	ark problem

函数	CEI	PCEI	APCEI
G_1	200(30)	>124.9(2)	115.1(17.48)
G_3	200(30)	>191.7(9)	175.3(9.6)
G_5	200(30)	53.6(1.75)	44.8(2.71)
G_6	74.6(3.55)	32.6(1.21)	14.3(0.99)
$G_{\rm 24}$	22.4(3.08)	13.5(1.11)	8.6(1.35)

4.2 翼型设计算例

为了进一步验证基于无参数自适应罚函数的 约束 EGO 方法对工程实际设问题的适用性,本节 以 RAE2822 翼型和飞翼布局翼型设计为例,并与 传统基于概率的约束 EGO 方法和静态罚函数的约 束 EGO 方法进行对比。气动设计算例采用基于多 块结构网格的 CFL3D 求解器进行气动数值计算, 湍流模型采用 SST 模型。

4.2.1 RAE2822 翼型设计算例

该算例是 ADODG 发布的第 2 个标准算例, 跨 音速黏性条件下 RAE2822 翼型减阻优化设计。算 例的设计状态为: Ma = 0.734, $C_l = 0.824$, $Re = 6.5 \times$ 10^5 , Ma为马赫数, C_l 为升力系数, Re为雷诺数。优 化目标为减小阻力系数, 约束条件为升力系数等于 0.824, 且低头力矩系数不大于 0.092, 同时翼型的面 积不减小。RAE2822 翼型优化的数学模型为

min C_d

s.t.
$$\begin{cases} C_l = 0.824 \\ |C_m| \le 0.092 \end{cases}$$
 (13)
$$A_{\text{Baseline}} \end{cases}$$

式中: *C*_m为翼型的力矩系数; *A*为翼型面积; *C*_d为阻力系数; *A*_{Baseline}为初始外形面积。

采用多块 C型网格进行气动特性计算, RAE2822 翼型网格的拓扑结构如图 10 所示, 网格 具体参数如表 3 所示。采用基于类函数和形函数 的参数化 (class function/shape function transformation, CST)方法^[31] 对翼型进行参数化, 翼型上下表 面各 10 个设计变量, 给定每个设计变量的扰动范 围, 翼型的设计空间如图 11 所示。

使用 LHD 方法在设计空间生成 100 个初始样本,分别采用 CEI、APCEI 和 PCEI 加点 200 次。结合设计者的经验,静态罚函数法的惩罚因子为1^[32]。此外,所有方法中都采用考虑约束的参考点机制选



图 10 RAE2822 计算网格

Fig. 10 Computational grid of RAE2822

表 3 RAE2822 翼型计算网格参数

Table 3 Parameter settings of computational grid of RAE2822

参数	数值
远场大小	50
网格规模	601×213
物面第一层网格距离	5×10^{-6}
物面法向网格增长率	1.13
翼型前缘网格距离	0.001
翼型后缘网格距离	0.001





Fig. 11 Initial design space of KAE2822 a

择参考点。

图 12 给出了不同方法的阻力系数收敛历程对 比。表4 给出了不同方法下翼型优化结果对比,无参 数自适应罚函数约束 EGO 方法的设计翼型阻力系 数最小,为111.6 counts(1 count=0.0001),基于静态罚 函数的约束 EGO 方法设计翼型次之,为112.1 counts, 约束 EGO 方法设计结果最差,为113.5 counts。

图 13 和图 14 分别给出了不同约束 EGO 方法 优化的翼型外形和压力分布对比优化的外形和压 力分布呈现了相似的变化趋势:优化翼型前缘变



图 12 不同方法阻力系数收敛历程对比

Fig. 12 Convergence of C_d for different methods

表 4 不同方法设计结果对比

Table 4 Comparison of design results for different methods

方法	C_d /counts	C_m	A
RAE2822	203.1	-0.0927	0.077 87
CEI	113.5	-0.0892	0.077 87
PCEI	112.1	-0.091 6	0.077 87
APCEI	111.6	-0.091 9	0.077 87

尖,最大厚度位置后移,后缘弯度有所增加;优化翼型的前缘吸力峰提高,后加载增强,压力恢复趋于 平缓,激波得到明显消除。这种现象与文献[32]结 果一致,说明各方法的优化趋势是一致且准确的。

该算例中最优解的力矩约束和几何约都接近 约束值,即最优值靠近可行域的边界,因此传统约 束 EGO 方法不能得到理想的结果;基于罚函数的 方法可以得到靠近约束边界的解,但静态罚函数的 惩罚因子设置依赖设计者以往的经验,难以在其他 设计问题直接应用;而无参数自适应罚函数约束 EGO 方法中惩罚因子在设计过程中自适应动态变 化,并且没有引入其他额外参数,降低了优化的难









Fig. 14 Pressure distribution of different optimized results

度,提高了优化设计效率。

4.2.2 飞翼布局翼型设计算例

飞翼布局飞机以其优越的气动特性,广泛应用 于运输机和无人机。由于水平尾翼的缺失,除了传 统翼型设计中的升阻特性要求外,飞翼布局翼型还 要求增加量化的抬头力矩以满足飞机纵向力矩配 平的需要。抬头力矩约束使得最优解位于可行域 边界位置,因此该算例可以更好地验证无参数自适 应罚函数约束 EGO 方法的能力。

本文以相对厚度为 18% 的对称翼型 NACA 65, 3-018 为初始翼型^[33],选取典型跨音速设计状态 *Ma*=0.70,*C_l*=0.25^[34]。设计目标为减阻,气动约束 为力矩系数不小于 0.03 及升力系数等于 0.25,几何 约束为翼型最大厚度不减小,设计模型如下所示:

(^{mil}	L_d	
{	$\begin{cases} C_l = 0.25 \\ C_l > 0.02 \end{cases}$	(14)
s.t.	$\begin{cases} C_m \ge 0.03 \end{cases}$	
l	$t/c_{\rm max} \ge 0.18$	

采用多块 C 型结构网格, NACA 65, 3-018 翼型 网格的拓扑结构如图 15 所示, 网格具体参数如 表 5 所示。同样采用 CST 参数化方法, 翼型上下 表面各 10 个设计变量, 共 20 个, 初始设计空间如 图 16 所示。

初始样本量 100 个,采用 CEI、APCEI 和 PCEI 分别加点 200 次,静态罚函数的惩罚因子为 1,所有 方法采用考虑约束的参考点选择机制。

从图 17 给出了目标函数收敛历程。表 6 给出 了不同约束 EGO 方法翼型优化结果对比,无参数 自适应罚函数约束 EGO 方法的设计翼型阻力系数 最小,为 86.5 counts;其次是基于静态罚函数的约 束 EGO 方法,其设计结果的阻力系数为 87.2 counts, 而约束 EGO 方法设计结果的阻力系数最大,为 87.9 counts。



图 15 NACA 65,3-018 计算网格 Fig. 15 Computational grid of NACA 65,3-018

表 5 NACA 65,3-018 翼型计算网格参数 Table 5 Parameter settings of computational grid of NACA 65,3-018

参数	数值
远场大小	50
网格规模	401×201
物面第一层网格距离	5×10^{-6}
物面法向网格增长率	1.12
翼型前缘网格距离	0.001
翼型后缘网格距离	0.000 5



图 16 NACA65, 3-018 翼型初始设计空间 Fig. 16 Initial design space of NACA65, 3-018 airfoil

图 18 和图 19 分别给出了不同约束 EGO 方法 优化的翼型外形和压力分布对比,相较于初始翼 型,优化翼型上表面趋于平坦,最大厚度位置后移, 前缘正弯度和后缘负弯度有所增加,呈现明显的前 缘正加载和后缘反加载特征;相较于初始翼型的强 激波,优化翼型的上下表面各出现一道弱激波,其 中 CEI 和 PCEI 得到的设计翼型上表面激波位置基 本重合,且相对靠近翼型前缘,而 APCEI 的激波位 置靠近翼型中部,且前缘吸力峰更高。

对于飞翼布局翼型算例,抬头力矩约束的满足



图 17 不同约束 EGO 方法 C_d 阻力系数收敛历程对比 Fig. 17 Convergence of C_d for different EGO methods

表 6 不同约束 EGO 方法设计结果对比 Table 6 Comparison of design results for different EGO methods

方法	C_d /counts	C_m	相对厚度
RAE2822	114.1	-0.009 8	0.18
CEI	87.9	0.0307	0.18
PCEI	87.2	0.0303	0.18
APCEI	86.5	0.030 0	0.18



图 18 RAE2822 翼型设计结果外形对比



由设计翼型的前缘正加载和后缘反加载共同配合 完成,而后缘反加载导致的负升力会造成翼型阻力 特性的损失,与设计目标相悖。设计翼型的抬头力 矩越大,其后缘反加载造成的负升力越多,导致其 阻力特性变差。因此,飞翼布局翼型算例其最优解 位于力矩约束边界处,此时 CEI 并不能取得理想的 结果;基于罚函数的约束 EGO 方法可更靠近约束 边界,但 PCEI 设置的静态惩罚因子存在后期收敛 缓慢问题,由图 17 可知; PCEI 通过自适应调整惩罚 因子,降低了优化难度,提高了优化设计效率。



图 19 NACA65, 3-018 翼型设计压力分布对比 Fig. 19 Pressure distribution of different optimized results for NACA65, 3-018 airfoil

5 结 论

本文对约束 EGO 方法的参考点的遴选及惩罚 因子对优化设计的影响进行了研究,提出了考虑约 束的参考点选择机制和基于无参数自适应罚函数 的约束 EGO 方法,通过一组非线性约束函数和 2个翼型设计算例对本文方法进行了验证。结论 如下:

 1)考虑约束的参考点选择机制结合约束信息, 综合考虑了样本库内可行解和不可行解的情况,分 别提出了不同的参考点选择机制,避免了因参考点 选择不合理造成 EI 函数失效的问题,提高了约束 EGO 算法的优化设计效率。

2)针对设计者难以直接给定合适的惩罚因子 的问题,本文提出了无参数自适应罚函数约束 EGO 方法,该方法对目标值和约束进行归一化处理,并 在优化过程中动态调整惩罚因子,使约束项在设 计过程中发挥作用的同时,不会对优化方向造成 误导。

3)本文针对无参数自适应罚函数约束 EGO, 首先进行了一组非线性约束函数算例测试,验证该 方法的效率和鲁棒性;通过 RAE2822 翼型和飞翼 布局翼型优化设计算例,验证了本文方法对于工程 实际问题的适用性,进一步证实了本文方法的可靠性。

参考文献(References)

- [1] QUEIPO N V, HAFTKA R T, SHYY W, et al. Surrogate-based analysis and optimization[J]. Progress in Aerospace Sciences, 2005, 41(1): 1-28.
- JONES D R. A taxonomy of global optimization methods based on response surfaces[J]. Journal of Global Optimization, 2001, 21(4): 345-383.

- [3] LOCATELLI M. Bayesian algorithms for one-dimensional global optimization[J]. Journal of Global Optimization, 1997, 10(1): 57-76.
- [4] WANG G G, SHAN S. Review of metamodeling techniques in support of engineering design optimization[J]. Journal of Mechanical Design, 2007, 129(4): 370-380.
- [5] XING J, LUO Y J, GAO Z H. A global optimization strategy based on the Kriging surrogate model and parallel computing[J]. Structural and Multidisciplinary Optimization, 2020, 62(1): 405-417.
- [6] LI Y H, SHI J J, CEN H, et al. A Kriging-based adaptive global optimization method with generalized expected improvement and its application in numerical simulation and crop evapotranspiration[J]. Agricultural Water Management, 2021, 245: 106623.
- [7] 夏露, 王丹. 基于 Kriging 自适应代理模型的气动优化方法[J]. 航空计算技术, 2013, 43(1): 13-17.

XIA L, WANG D. Aerodynamic optimization method based on Kriging adaptive surrogate model[J]. Aeronautical Computing Technique, 2013, 43(1): 13-17(in Chinese).

[8] 韩忠华. Kriging 模型及代理优化算法研究进展[J]. 航空学报, 2016, 37(11): 3197-3225.

HAN Z H. Kriging surrogate model and its application to design optimization: A review of recent progress[J]. Acta Aeronautica et Astronautica Sinica, 2016, 37(11): 3197-3225(in Chinese).

- [9] BHOSEKAR A, IERAPERTRITOU M. Advances in surrogate based modeling, feasibility analysis, and optimization: A review[J]. Computers & Chemical Engineering, 2018, 108(4): 250-267.
- [10] BONS N P, HE X L, MADER C A, et al. Multimodality in aerodynamic wing design optimization[J]. AIAA Journal, 2019, 57(3): 1004-1018.
- [11] LEIFSSON L, KO A, MASON W H, et al. Multidisciplinary design optimization of blended-wing-body transport aircraft with distributed propulsion[J]. Aerospace Science and Technology, 2013, 25(1): 16-28.
- [12] VISWANATH A, FORRESTER A I J, KEANE A J. Constrained design optimization using generative topographic mapping [J]. AIAA Journal, 2014, 52(5): 1010-1023.
- [13] ZHOU L, HUANG J T, GAO Z H, et al. Three-dimensional aerodynamic/stealth optimization based on adjoint sensitivity analysis for scattering problem[J]. AIAA Journal, 2020, 58(6): 2702-2715.
- [14] 黄江涛,高正红,余婧,等.大型民用飞机气动外形典型综合设计 方法[J]. 航空学报, 2019, 40(2): 522369.
 HUANG J T, GAO Z H, YU J, et al. A typical integrated design method for aerodynamic shape optimization of large civil aircraft
 [J]. Acta Aeronautica et Astronautica Sinica, 2019, 40(2): 522369(in Chinese).
- [15] SASENA M J, PAPALAMBROS P, GOOVAERTS P. Exploration of metamodeling sampling criteria for constrained global optimization[J]. Engineering Optimization, 2002, 34(3): 263-278.
- [16] PARR J M, KEANE A J, FORRESTER A I J, et al. Infill sampling criteria for surrogate-based optimization with constraint handling[J]. Engineering Optimization, 2012, 44(10): 1147-1166.
- [17] BASUDHAR A, DRIBUSCH C, LACAZE S, et al. Constrained efficient global optimization with support vector machines[J]. Structural and Multidisciplinary Optimization, 2012, 46(2): 201-221.
- [18] BRERETON R G, LLOYD G R. Support vector machines for classification and regression[J]. Analyst, 2010, 135(2): 230-267.

- [19] BAGHERI S, KONEN W, ALLMENDINGER R, et al. Constraint handling in efficient global optimization[C]// Proceedings of the Genetic and Evolutionary Computation Conference. New York: ACM, 2017, 1: 673-680.
- [20] KRAMER O. A review of constraint-handling techniques for evolution strategies[J]. Applied Computational Intelligence and Soft Computing, 2010, 2010: 3.
- [21] MEZURA-MONTES E, COELLO C A C. Constraint-handling in nature-inspired numerical optimization: past, present and future[J]. Swarm and Evolutionary Computation, 2011, 1(4): 173-194.
- [22] LIU B, GROUT V, NIKOLAEVA A. Efficient global optimization of actuator based on a surrogate model assisted hybrid algorithm[J]. IEEE Transactions on Industrial Electronics, 2018, 65(7): 5712-5721.
- [23] ZAEFFERER M, STORK J, FRIESE M, et al. Efficient global optimization for combinatorial problems[C]// Proceedings of the 2014 Annual Conference on Genetic and Evolutionary Computation. New York: ACM, 2014, 1: 871-878.
- [24] 王红涛, 竺晓程, 杜朝辉. 改进 EGO 算法在跨声速翼型气动优化 设计中的应用[J]. 上海交通大学学报, 2009, 43(11): 1832-1836.
 WANG H T, ZHU X C, DU Z H. Application of the improved EGO algorithm in transonic airfoil aerodynamic optimization design[J]. Journal of Shanghai Jiao Tong University, 2009, 43(11): 1832-1836(in Chinese).
- [25] 龙腾, 刘建, 陈余军, 等. 基于约束 EGO 的对地观测卫星多学科 设计优化[J]. 机械工程学报, 2018, 54(10): 133-142. LONG T, LIU J, CHEN Y J, et al. Multi-disciplinary design optimization of earth observation satellite based on constrained EGO[J]. Journal of Mechanical Engineering, 2018, 54(10): 133-142(in Chinese).
- [26] DONG H C, SONG B W, DONG Z M, et al. SCGOSR: Surrogatebased constrained global optimization using space reduction[J]. Applied Soft Computing, 2018, 65(c): 462-477.
- [27] NADARAJSH S. Aerodynamic design optimization: Drag minimization of the NACA 0012 in transonic inviscid flow[EB/OL]. (2017-09-15)[2022-05-31]. https://info.aiaa.org/tac/ASG/APATC/AeroDes ignOpt-DG/Tes.
- [28] KEANE A, FORRESTER A, SOBESTER A. Engineering design via surrogate modelling: A practical guide[M]. Reston: AIAA, 2008.
- [29] CASTILLO O, MELIN P. Genetic algorithms and simulated annealing[M]// Soft Computing and Fractal Theory for Intelligent Manufacturing. Heidelberg: Physica, 2003, 1: 93-125.
- [30] AKBARI H, KAZEROONI A. KASRA: A Kriging-based adaptive space reduction algorithm for global optimization of computationally expensive black-box constrained problems[J]. Applied Soft Computing, 2020, 90: 106154.
- [31] KULFAN B, BUSSOLETTI J. Fundamental parameteric geometry representations for aircraft component shapes[C]// 11th AIAA/ ISSMO Multidisciplinary Analysis and Optimization Conference. Reston: AIAA, 2006: 6948.
- [32] 王超,高正红,张伟,等. 自适应设计空间扩展的高效代理模型气动优化设计方法[J]. 航空学报, 2018, 39(7): 121745.
 WANG C, GAO Z H, ZHANG W, et al. Efficient surrogate-based aerodynamic design optimization method with adaptive design space expansion[J]. Acta Aeronautica et Astronautica Sinica, 2018,

39(7): 121745(in Chinese).

[33] SEARS W. Flying-wing airplanes-The XB-35/YB-49 program[C]// The evolution of aircraft wing design; Proceedings of the Symposium. Reston: AIAA, 1980: 3036.

[34] LIEBECK R H. Design of the blended-wing-body subsonic transport[J]. Journal of Aircraft, 2004, 41(1): 10-25.

Efficient surrogate-based aerodynamic optimization with parameter-free adaptive penalty function

ZHANG Wei¹, GAO Zhenghong^{1, 2, *}, WANG Chao³, XIA Lu^{1, 2}

(1. School of Aeronautics, Northwestern Polytechnical University, Xi'an 710072, China;

2. State Key Laboratory of Airfoil and Cascade Aerodynamics, Northwestern Polytechnical University, Xi'an 710072, China;

3. Beijing Institute of Electronic System Engineer, Beijing 100854, China)

Abstract: Complex constraints must be addressed in the aerodynamic optimizations. Distinct constraints not only influence the optimization outcomes but also significantly influence the optimization method's efficacy. This paper investigates the impact of reference points on optimization design results using the constrained efficient global optimization method (EGO) and suggests a mechanism for selecting reference points that takes constraints into account. Afterwards, for the problem of constraint processing, the constrained expected improvement (EI) method and the penalty function method are compared and found that the penalty function method can find a feasible solution that satisfies the constraints more quickly. However, in this process, the penalty factor has a great influence on the optimization efficiency, and inappropriate penalty factors will damage the optimization efficiency. Drawing on the aforementioned evaluation, this study suggests a constrained EGO technique utilizing an adaptive penalty function that is free of parameters. By normalizing the target value and the constraint value, the feasible solution with the smallest target value or the infeasible solution closest to the feasible region is selected as the reference point. The penalty factor is adaptively adjusted, so that the algorithm can search for the ideal solution sufficiently. This approach can significantly increase the optimization efficiency, as shown by the constrained test functions and airfoil design challenges.

Keywords: optimization design; surrogate model; reference point; aerodynamic design; adaptive penalty function

Received: 2022-05-31; Accepted: 2022-09-18; Published Online: 2022-10-10 09:37 URL: link.cnki.net/urlid/11.2625.V.20221009.1124.011

Foundation item: National Key Laboratory Foundation of Airfoil and Blade Grid Aerodynamics (614220121020128)

^{*} Corresponding author. E-mail: zgao@nwpu.edu.cn

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

DOI: 10.13700/j.bh.1001-5965.2022.0415

RP-3 航油火焰三维温度场和碳烟浓度场测量

左冰娴,周宾*,戴明露

(东南大学能源与环境学院,南京210096)

摘 要: RP-3 航空煤油是航空发动机的常用燃料之一,火焰温度及碳烟浓度分布作为反映 燃烧状态最直接的参数,是燃料优化配比的重要数据支撑。针对三维测量中,重建路径计算的准确 性受相机位姿影响的问题,提出了一种将三维辐射成像法与相机立体标定技术相结合的测量方法。 数值模拟结果表明,所提方法可以在提供较高的空间分辨率的同时保证测量结果的可靠性。利用标 定后的六相机测量系统,对 RP-3、空气预混火焰进行了实验测量,获得了其三维温度及碳烟浓度分 布。测量的空间分辨率为1mm×1mm,重建与热电偶测量温差的标准不确定度为 0.58 K,证明 了测量方法的精确性。

关键 词: RP-3 航空煤油; CCD 相机; 相机立体标定; 温度; 碳烟浓度
 中图分类号: V312⁺.1; TK311; TK314
 文献标志码: A 文章编号: 1001-5965(2024)04-1273-09

RP-3航空煤油是中国自主研发的航空煤油,目 前已广泛地应用于各类飞行器中^[1]。对其火焰温度 和碳烟浓度分布进行同步测量,可为燃料优化与航 空发动机设计提供数据支撑。辐射成像技术是一 种非侵入式测量技术,通过布置有限的探测设备即 可获取大量的投影信息,是目前常用于工业场合中 的测量技术^[2]。该方法通过对成像设备捕捉到的火 焰图像进行重建分析,便可从中提取出温度和碳烟 浓度等待测参数^[3]。

双色法是在工业测量中使用最广泛的方法。 Yan 和 Lou 以 R、G 双通道辐射强度作为投影量, 以谱带模型校正了波长 CCD 相应曲线对测量的干 扰,并给出了二维温度及碳烟浓度分布^[3],随后实验 测量了乙烯扩散火焰的二维温度与碳烟浓度分 布^[4],证明了双色法在燃烧诊断中的可靠性。在三 维的测量中,Worth 和 Dawson^[5]将相机架设在滑轨 上,利用单相机获取了不同角度下稳态轴对称火焰 的图像,实现了三维火焰的重构。Yan 等^[6] 通过在 相机前加设折射镜获取了6个角度的火焰图像,并 结合代数重建技术(algebra reconstruction technique, ART)算法重构了火焰的截面图像。Hossain^[7]耦合 相机和光纤传像束,利用层析成像原理,成功重建 了非轴对称火焰的三维温度分布。然而上述研究 对象多为轴对称火焰或假设了相机接收的为平行 光信号,并非真正的基于三维辐射成像的测量。

近年来,诸多学者对辐射成像技术的研究可分为以下2种:①提高图像光谱维度信息^[8-11],通过在成像设备前加设液晶可调滤波片,获取不同波长下的火焰图像,进一步提高了重建结果的精确性。然而文献[8]的模拟结果表明,当计算路径误差大于1%时,组分浓度的测量便会失真;②提高图像空间维度信息^[12-14],利用可以同时获取火焰图像的光强信息和路径信息的光场成像设备,达到了用单台成像设备获取火焰多空间信息的目的。但该方法对相机的偏移角有着较高的要求,如要保证重建结果的可信度,需要偏移角在1.875×10⁻⁴ rad 以内^[14]。由

收稿日期:2022-05-26;录用日期:2022-07-22;网络出版时间:2022-07-2814:33 网络出版地址:link.cnki.net/urlid/11.2625.V.20220727.1738.002

基金项目:国家重点研发计划(2017YFB0603204);国家自然科学基金(50976024,50906013)

*通信作者. E-mail: zhoubinde@seu.edu.cn

引用格式: 左冰娴,周宾,戴明露. RP-3 航油火焰三维温度场和碳烟浓度场测量 [J]. 北京航空航天大学学报, 2024, 50 (4): 1273-1281. ZUOBX, ZHOUB, DAIML. Measurement of three-dimensional temperature and soot volume fraction for RP-3 jet fuel flame [J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50 (4): 1273-1281 (in Chinese). 此可见,在三维测量中,无论采用何种测量系统,成 像设备的位姿状态决定了重建的可靠性,而设备的 位姿关系则受到外部参数的影响^[15]。如文献 [16] 所述,在粗糙和精细2种情况下,相机的平均校准 误差可分别低至 0.294 mm 和 0.027 mm。

本文提出了一种基于立体标定方法的三维辐 射成像技术。利用立体标定技术,通过构建相机的 旋转和平移矩阵,对相机的位置与姿态进行校正, 进而提高火焰重建的精确性。实验测量得到了空 气、RP-3预混火焰的三维温度和碳烟浓度分布。 重建温度与热电偶测量值的对比结果表明,本文提 出的测量方法在保证空间分辨率的同时,提高了测 量结果的精确性。

1 重建模型与实验方法

1.1 重建模型

相机成像原理和测量模型如图 1 所示, CCD 接收到的辐射能量为每条视线通过点的能量的累 加。将尺寸为 $L_f \times W_f \times H_f$ 的待测区域离散为 $a \times b \times c =$ N个长方体体元,每个体元的温度和碳烟浓度取该 体元的平均值。将相机的视场进行离散,每条离散 的射线将会以不同的路径穿过不同的体元序列,因 此在成像设备中呈现出不同的成像设备响应值。 图中, N_x 、 N_y 和 N_z 分别为沿x轴、y轴和z轴划分的体 元个数; L_{path} 和 ΔL_{path} 分别为离散射线穿过体元序列 的总长度和离散射线穿过某一体元的长度。





将相机视线离散为M条射线,以其中第m条射 线为例,并假设长度微元为p,那么根据投影定律, 图 1 中相机 CCD 感应到的视线 L_{path,m}的辐射强度

*I*_λ(*L*_{nath.m}),可表示为^[11]

$$I_{\lambda}(L_{\text{path},m}) = \int_{0}^{m_{1}} \kappa_{\lambda}(p) \cdot I_{\text{b},\lambda}(p) dp \approx \sum_{n=1}^{N} \kappa_{\lambda}(n) \cdot I_{\text{b},\lambda}(n) \cdot l_{\text{path},m}(n) = \sum_{n=1}^{N} l_{\text{path},m} \cdot H_{\lambda}(n)$$
(1)

式中: *m*₁为第*m*条射线穿过体元序列的长度; κ_λ为离 散体元中的碳烟吸收系数; *l*_{b,λ}为波长为λ处的单色 黑体辐射强度; *l*_{path,m}(*n*)为第*m*条路径穿过第*n*个体 元的长度。

将所有的路径联立后改写成矩阵形式,有

 $L_{M \times N} \cdot H_{N \times 1} = I_{M \times 1}$ (2) 式中:将积分问题转换为矩阵问题后,L为系数矩 阵,实际上等于相机视线穿过体元的路径长度;I为 投影矩阵,在实际测量中为相机采集到的单色辐射 强度;H为体元的辐射源项,是需要求解的量。

路径矩阵L的计算精度由相机的相对位姿确 定,因此获取相机的精确相对位姿对减少路径矩阵 的扰动有着重要的意义。根据相机立体标定技术, 可以通过平移矩阵和旋转矩阵对路径矩阵的起始 点及路径射线发射方向进行校正,校正点坐标和理 论点坐标关系可表示为

$$\begin{bmatrix} X_{c} \\ Y_{c} \\ Z_{c} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{1} \ \mathbf{R}_{2} \ \mathbf{R}_{3} \end{bmatrix} \cdot \begin{bmatrix} X_{0} \\ Y_{0} \\ Z_{0} \end{bmatrix} + \begin{bmatrix} T_{x} \\ T_{y} \\ T_{z} \end{bmatrix}$$
(3)

式中: X_c 、 Y_c 和 Z_c 为校正后相机的位置坐标; X_0 、 Y_0 和 Z_0 为相机理论测量坐标; T_x 、 T_y 和 T_z 为平移矩 阵的组成元素,其值是物理坐标系原点坐标在像素 坐标系中的位置; $[R_1 R_2 R_3]$ 为旋转矩阵,可表示为

$$\begin{cases} \mathbf{R}_{1} = \begin{bmatrix} \cos \beta \cos \theta \\ -\cos \alpha \sin \theta + \sin \alpha \sin \beta \cos \theta \\ \sin \alpha \sin \theta + \cos \alpha \sin \beta \cos \theta \end{bmatrix} \\ \mathbf{R}_{2} = \begin{bmatrix} \cos \beta \sin \theta \\ \cos \alpha \cos \theta + \sin \alpha \sin \beta \sin \theta \\ -\sin \alpha \cos \theta + \cos \alpha \sin \beta \sin \theta \end{bmatrix}$$
(4)
$$\mathbf{R}_{3} = \begin{bmatrix} -\sin \beta \\ \sin \alpha \cos \beta \\ \cos \alpha \cos \beta \end{bmatrix}$$

其中: α、β和θ分别为相机绕自身中心点所建立坐 标系中的x轴、y轴和z轴的旋转角度。

在路径矩阵被求解后,可以通过最小二乘法求 解 H,该方法可以表示为

$$\begin{cases} \min & \|Ax - b\|_2 \\ \text{s.t.} & l \le x \le u \end{cases}$$
(5)

式中: A为系数矩阵, 在本文中为L矩阵; b为结果向 量, 在本文中为I矩阵; l和u分别为H_a的最大最小值 边界。方程式 (2) 是一个不适定的反问题, 对于测 量误差十分敏感, 因此在总是存在有测量误差的现 实测量条件下, 需要特定算法来进行求解。本文采 用 Tikhonov 正则化算法, 将方程式 (5) 用三维 Tikhonov 正则表示, 有

式中: δ 为正则化参数,用于控制惩罚项和误差的相关性; δS_{3D} 为离散梯度矩阵,其表达式为

$$S_{3D,ij} = \begin{cases} 1 & i = j \\ -1/n_i & i \, \text{fl} \, \Re j \\ 0 & \text{I} \, \text{th} \end{cases}$$
(7)

在上述三维 Tikhonov 正则化表达式中, n_i根据 待求解体元所在的位置不同, 有不同的取值。当待 求解体元为中心体元时n_i = 6, 表示该体元有 6 个相 邻体元; 同样地, 当待求解体元分别为面体元、边体 元和角体元时, n_i的取值分别为 5、4 和 3。

当体元中辐射源项H_λ的信息被计算出来后,即 可从中提取所需要测量的温度信息及碳烟浓度信 息。对于任意体元内的辐射源项,根据其定义¹⁰, 可得

$$H_{\lambda}(n) = \kappa_{\lambda}(n) \cdot I_{\mathrm{b},\lambda}(n) \tag{8}$$

式中: *I*_{b,l}(*n*)为*n*号体元中火焰在波长为λ处的单色 黑体辐射强度,可由普朗克黑体定律计算:

$$I_{b,\lambda}(n) = c_1 / \{\lambda^5 [e^{c_2 / (\lambda T_n)} - 1]\}$$
(9)

其中: c_1 、 c_2 分别为普朗克第一、第二常量,分别取 为 3.742×10⁸ (W· μ m⁴)/m²和 1.439 μ m·K; T_n 为 n号 体元处的温度。根据气液燃料生成碳烟颗粒的性 质,其直径在 20~60 nm^[11],满足瑞利假设,因此其 碳烟吸收 κ_λ 系数近似表示:

$$\kappa_{\lambda}(n) = \frac{6\pi E(r)f_{\nu}(n)}{\lambda} \tag{10}$$

式中: *E*(*r*)为复折射率*r*的方程, *r* = *q*-*ik*, *q*和*k*为与 波长有关的常数,可以按文献 [9] 进行计算; *f*_v(*n*)为 体元中的碳烟体积浓度。

将式(8)~式(10)联立,可将方程最终描述为 如下形式:

$$H_{\lambda}(n) = \frac{6\pi E(r)c_1}{\lambda^6} \cdot \frac{f_{\nu}(n)}{e^{c_2/(\lambda T_n)} - 1}$$
(11)

当波长已知时,式(11)中仅有温度和碳烟浓度 2个未知数,若已知2个波长下的辐射源项,便可通 过求解方程来计算出该体元处的温度和碳烟浓度。

1.2 实验方法

图 2 展示了实验设备及测量系统布置方案。 燃烧平台煤油预混燃烧炉如图 2(a) 所示, RP-3 航 油和空气分别输入到煤油蒸发器和空气预热器中, 再经由内管和外管传输至混合室中混合,随后由喷 口喷出,并被点燃形成待测火焰。重建区域尺寸为 *L*_f×W_f×H_f = 25 mm×25 mm×40 mm。

多相机测量系统如图 2(b) 所示, 6 台相机与六 自由度位移台组合后安装在六边形支架上。相邻 相机间的夹角为 60°, 相机距燃烧器中轴线 130 mm。





高速数据采集器 (RedPitaya: STEMlab 125-10) 发出 的脉宽调制 (pulse width modulation, PWM) 信号,用 于同时触发 6 台相机,获取的图像经过交换机 (TP-Link: TL-SG1008VE), 通过以太网传输至电脑中储 存计算。

本文实验采用的相机为大恒彩色工业相机 (MER-132-43GC),可用于接收 R、G和 B 三个通道 的数据。相机前设置有焦距为 12 mm的镜头 (Computar M1214-VSW),并根据相机的量子响应效 率在镜头前加设 525 nm 和 610 nm 的双波滤波片, 使相机仅能获取 R、G 通道下的辐射信号。利用黑 体炉对相机的光电响应值进行标定,获取 R、G 值 与辐射强度值的对应关系。此外,利用经过温度修正^[17] 后的高精度 B 型热电偶对火焰场多点温度进行了 测量,用于验证实验方案及重建结果的有效性。

为了获取每个相机相对火焰的精确位姿关系, 对多相机系统进行了立体标定。首先在在待测区 域中轴线处悬挂带有标尺的重锤,通过调整六自由 度位移台,将相机的视场中心与重锤中心点对齐, 保证相机的视场中心与待测区域中轴线对齐。随 后,利用相机立体标定方法^[18-19],确定每台相机相对 理论位置与姿态的偏移量。

标定用的棋盘格标靶尺寸为 120 mm×120 mm, 棋盘格上交错布置有 15 个黑色正方形块,每个正 方形块的边长为 20 mm,制造公差为±0.005 mm。 对相邻两台相机依次进行标定,整个标定步骤需要 重复 6次,其中 1~5次标定可以得到 2~6号相机 相对 1号相机的相对位姿状态,第 6次标定用于验 证标定结果的准确性。当标定结果正确时,相机 6 相对于相机 1 位姿偏移量与第 6次标定结果得到 的校正量互为相反数。值得一提的是,每次标定仅 能获取相机与前一相机的相对位姿关系,因此需要 将相邻相机的旋转和平移量按相机次序进行传递, 方可得到 2~6号相机相对 1号相机的位姿校正矩 阵。校准后,总体的平均像素误差小于0.2像素。

以重建中心为原点,建立世界坐标系,重建中 心指向相机的方向和火焰喷射方向分别为*x*和*z*轴 正方向。表1为相机位姿标定的具体结果,α、β和 θ的正负号表示旋转方向,顺时针为正,逆时针为 负。对比理论值和实际值可知,相机位置偏移量为 2~5 mm,角度偏转在 0°~2.5°,说明仅通过手动调 整相机会使得相机与理论位置存在较大的几何偏 差,从而使得路径矩阵产生扰动并最终影响重建结 果的精确性。

表 1 相机位姿标定结果 Table 1 Camera position calibration results

			平移矩阵	=/mm					旋转	矩阵/(°)		
相机		r	ر	v	1	Z		α	1	в		θ
	理论值	实际值	理论值	实际值	理论值	实际值	理论值	实际值	理论值	实际值	理论值	实际值
相机1和2	113.596 5	109.263 2	-63.27	-63.3203	0	0.134 1	0	0.569 5	0	0	-60	-60.314 1
相机1和3	225.1667	221.2563	3.464 1	1.052 7	0	0.249 3	0	-0.292 9	0	0	-120	-119.078 2
相机1和4	223.1667	225.023 0	133.464 1	135.764 7	0	0.1187	0	0.5509	0	0	-180	-181.9560
相机1和5	109.583 3	114.1136	196.7321	199.143 3	0	0.082 5	0	0.372 1	0	0	-240	-238.979 1
相机1和6	-2.000 0	-0.415 2	130	128.223 1	0	0.2196	0	-0.5867	0	0	-300	-300.685 6

2 结果与讨论

2.1 相机位姿对重建结果的影响

在图 2 所示的实验条件下进行数值模拟,以双 峰非轴对称火焰为数值模拟对象,温度和碳烟浓度 的具体分布如图 3(a) 和图 3(b) 所示。重建区域被 离散为 25×25×40 总计 25 000 个体元,测量的空间 分辨率为 1 mm×1 mm。

定义相对误差用于评估重建的准确性,误差来 源包括理论模型误差、相机位姿和重建算法,其表 达式为

$$E_{\rm rel} = \left| \frac{\alpha_n^{\rm h} - \alpha_n^{\rm r}}{\alpha_n^{\rm r}} \right| \times 100\% \tag{12}$$

式中: $\alpha_n^h \pi \alpha_n^r$ 分别为温度或碳烟浓度在编号为n的体元中的假设值和重建值。

图 4 展示了相机位姿是否校正时各体元处的 重建误差。横轴为体元编号,纵轴为相应体元处的 相对重建误差。由图 4 可知,在相机位置和姿态均 被校正的情况下,温度及碳烟浓度的平均重建误差 分别为 2.26×10⁻⁵% 和 3.55×10⁻⁴%,重建结果良好。 此外,误差分布趋势与温度的分布趋势相同,温度 较高的地方,温度和碳烟浓度的重建误差也较大, 此时重建误差受理论模型误差的影响更大。其次, 由于在理论模型中忽略了碳烟自吸收的作用,因此 在相同的校正条件下,碳烟浓度的重建误差总体上 大于温度的重建误差,且相机位姿对碳烟浓度的影响相较于温度而言更高,这意味着碳烟浓度对相机 位姿的精确性更为敏感。

当相机的位置和姿态均不被校正时,温度和碳 烟浓度的平均重建误差分别达到34.08%和3.51×10⁴%, 火焰重构伪影严重,重建结果失真明显,表明了相 机位姿校正对保证测量结果有效性的重要性。值 得注意的是,仅校正位置时的温度与碳烟浓度的重 建误差小于仅校正姿态时的重建误差,说明了位置 偏移带来的影响要略高于相机角度偏摆所带来的 影响,而这可能是本次实验中,相机位置偏移量大 于角度偏摆量所造成的现象。

2.2 实验结果与讨论

在图 2 所示实验台上,利用经过位姿校正后的 多相机测量系统对 RP-3、空气预混火焰进行了拍 摄。航油流量为 120 mL/h,空气流量为 3.711 L/s, 6 台相机捕获的火焰图像如图 5 所示。对火焰图像 中无火焰区域进行了裁剪,裁剪后的图像尺寸为 450 像素×700 像素。设置相机的曝光时间为 200 μs, 确保图中不存在饱和像素点。

2.2.1 重建结果与讨论

温度与碳烟浓度的重建结果如图 6 所示。由 重建图像可知,火焰自低处向高处,温度分布整体 表现出由中间低两边高发展至中间高四周低的趋 势,而碳烟浓度分布则整体呈现出相反的趋势。由









Fig. 4 Effect of camera position correction on reconstruction results

于底层接近燃烧器,燃烧剧烈程度较低,该处重建 温度也较低,相应地在图像中显示的像素值也较 低。随着轴心高度的增加,燃烧程度逐渐加剧,温 度也随之逐渐增加。

图 7 展示了温度与碳烟浓度在不同层高处的

分布趋势。平均温度先随高度的上升而增加,在 21 层达到最大值1379.3 K,随后当高度的进一步增加,层平均重建温度反而下降。层最高温度则随着 高度的增加而上升,在第40 层时,取得整个重建区 域的最高温度1548.1 K。导致上述现象的原因是:



图 5 RP-3 火焰图像 Fig. 5 Images of RP-3 flame

在低层时,混合气体的温度处于着火临界点,临界 点上层处于火焰化学反应起始位置,导致燃烧不完 全,燃烧温度低。随着高度的增加,燃烧反应不断 加剧,燃料的燃烧更充分,温度也随之升高。随后, 火焰继续向上发展,火焰的最高温度持续增加,但 由于热扩散加剧,使得更高处的火焰温度的热量积 累下降,火焰边界处温度下降,层平均温度呈现下 降的趋势。

对于碳烟浓度而言,随着高度的增加,层平均 和最大碳烟浓度总体上表现为先下降后上升的趋势。在1~10层处,碳烟浓度呈现出下降趋势,随 后随着高度的增加,层平均碳烟浓度和最大碳烟浓 度也随之增加。碳烟的形成过程依次包括有碳烟 前驱体生成、颗粒形成、颗粒生长、颗粒凝聚凝结、新生颗粒转换为成熟颗粒和氧化阶段等上述 6个阶段^[20]。因此,火焰自下而上的发展过程,也是 碳烟颗粒逐步形成的过程。在1~10层碳烟发展 的过程中,由于火焰下半部分靠近燃烧器喷口,受 气流影响较大,在一定程度上扰乱了颗粒的团聚成 核过程,使得碳烟浓度呈现下降的趋势^[21]。此后, 随着氧化反应的进一步进行,碳烟颗粒逐渐发展成 型,颗粒数增加,碳烟浓度也随之上升。

2.2.2 不确定度分析

为了验证测量结果的准确性,利用热电偶在 20个测点对火焰温度进行了测量,2种方法在测点 处的测温值如图8所示。不同高度下,热电偶测温 与重建温度的层温度变化趋势保持一致即在高度 较高的区域,中心温度高而四周温度低;在较低的 区域,温度呈现出四周高而中心低,这种趋势与 图6的重建结果保持一致。

利用标准不确定度对测量的标准不确定度U进行了衡量,其计算式如下^[22]:



Fig. 6 Temperature reconstruction and soot volume fraction distribution



图 7 温度及碳烟浓度层重建分布趋势





图 8 重建温度与热电偶测量温度对比



$$U(D) = \sqrt{\frac{\sum_{i=1}^{\nu} \left(D_i - \overline{D}\right)^2}{\nu(\nu - 1)}}$$
(13)

式中: *D*_i为测点处重建温度与热电偶测温的温差; **D**为测量温差的平均值; v为测量次数,在本文中为 测点数,为20。通过立体标定校正后的测量整体标 准不确定度为0.58 K,主要由燃烧不均匀及环境对 火焰场的影响导致的燃空比波动、热电偶测量误 差、环境噪声引入多相机测量系统内的误差、理论 模型及重建算法误差以及相机光电响应标定误差 等因素引起。

此外, 热电偶和重建温差在高度和径向均为 10 mm 处取得最大值, 为 20.3 K, 此时测量的相对误 差为 1.51%。随着高度增加, 热电偶与重建温差减 小, 这种现象可能是高度高的区域温度也较高, 相 机接收到此处的辐射信号强, 信噪比高, 重建误差 相对较小所导致的。值得注意的是, 由于滤波片并 不能达到理论上的滤波效果, 相机获取的信号总是 包含在某一波段内的信号总值,因此相机接收到的 信号高于理论辐射强度值,也使得重建温度始终高 于热电偶测量值。

3 结 论

本文介绍了一种将相机立体标定技术与三维 辐射成像技术相结合的测量方法,可以在较高的空 间分辨率下实现火焰温度和碳烟浓度的高精确性 重建。由数值模拟和实验结果可得如下结论:

1) 当相机的位置与姿态均不被校正时, 温度和 碳烟浓度的平均重建误差分别为 34.08% 和 3.51× 10⁴%, 出现严重的失真现象。而使用立体标定方法 进行校正后, 二者的误差分别下降至 2.26×10⁻⁵% 和 3.55×10⁻⁴%, 表明了精确确认相机位姿在辐射成像 测量中的重要性。

2)随着火焰距喷口距离的增加,火焰层平均温 度先上升后下降,层平均碳烟浓度则先下降后上 升。这种趋势与碳氢燃料预混火焰发展及碳烟生 成规律相吻合。

3) 热电偶测量值和重建温度值在不同高度、不同径向处的分布趋势相同。2 种测量方法的最大温差为 20.3 K, 出现在低层温度较低的位置, 随着层高增加, 温度升高, 二者的温差逐渐减小, 证明了本文方法在燃烧诊断中的可靠性。

参考文献(References)

- [1] 贾洲侠, 徐国强, 邓宏武, 等. 亚临界压力下航空煤油 RP-3 动力 黏度测量[J]. 北京航空航天大学学报, 2014, 40(7): 934-938.
 JIA Z X, XU G Q, DENG H W, et al. Dynamic viscosity measurements of aviation hydrocarbon fuel RP-3 at sub-critical pressures[J].
 Journal of Beijing University of Aeronautics and Astronautics, 2014, 40(7): 934-938(in Chinese).
- [2] GRAUER S J, UNTERBERGER A, RITTLER A, et al. Instantaneous 3D flame imaging by background-oriented schlieren tomography [J]. Combustion and Flame, 2018, 196: 284-299.
- [3] YAN W J, LOU C. Two-dimensional distributions of temperature and soot volume fraction inversed from visible flame images[J]. Experimental Thermal and Fluid Science, 2013, 50: 229-233.
- [4] YAN W J, CHEN D M, YANG Z M, et al. Measurement of soot volume fraction and temperature for oxygen-enriched ethylene combustion based on flame image processing[J]. Energies, 2017, 10(6): 750.
- [5] WORTH N A, DAWSON J R. Tomographic reconstruction of OH* chemiluminescence in two interacting turbulent flames[J]. Measurement Science and Technology, 2013, 24(2): 024013.
- [6] LU G, YAN Y, CORNWELL S, et al. Impact of co-firing coal and biomass on flame characteristics and stability[J]. Fuel, 2008, 87(7): 1133-1140.
- [7] HOSSAIN M M, LU G, YAN Y. Optical fiber imaging based tomographic reconstruction of burner flames[J]. IEEE Transaction on In-

strumentation and Measurement, 2012, 61(5): 1417-1425.

- [8] 谢正超,王飞,张海丹,等.基于可见光-近红外多光谱图像的火焰 场参数重建[J]. 燃烧科学与技术, 2016, 22(6): 552-557. XIE Z C, WANG F, ZHANG H D, et al. Simultaneous reconstruction of three dimensional field parameters based on the visible and near-infrared multi-spectral flame images[J]. Journal of Combustion Science and Technology, 2016, 22(6): 552-557(in Chinese).
- [9] LIU H W, ZHENG S, ZHOU H C. Measurement of soot temperature and volume fraction of axisymmetric ethylene laminar flames using hyperspectral tomography[J]. IEEE Transactions on Instrumentation and Measurement, 2017, 66(2): 315-324.
- [10] LI M J, SUN K, HE Z. Integrated research of a multi-wavelength method in anisotropic scattering flame on soot temperature and radiative coefficient reconstruction[J]. Applied Optics, 2018, 57(21): 5899-5913.
- [11] 张海丹. 基于高光谱成像系统的火焰三维温度场和烟黑浓度场 重建研究[D]. 杭州: 浙江大学, 2016: 22-23. ZHANG H D. Three dimensional reconstruction of temperature and soot volume fraction distribution in flames based on hyperspectral imaging system[D]. Hangzhou: Zhejiang University, 2016: 22-23(in Chinese).
- [12] SUN J, HOSSAIN M M, XU C L, et al. Investigation of flame radiation sampling and temperature measurement through light field camera[J]. International Journal of Heat and Mass Transfer, 2018, 121: 1281-1296.
- [13] HUANG X, QI H, ZHANG X L, et al. Application of landweber method for three-dimensional temperature field reconstruction based on the light-field imaging technique[J]. Journal of Heat Transfer, 2018, 140(8): 082701.
- [14] 史景文,齐宏,孙安泰,等.光场相机的标定误差对三维温度场重 建的影响[J]. 燃烧科学与技术, 2022, 28(2): 220-228.

SHI J W, QI H, SUN A T, et al. Effect of calibration error of plenoptic camera on three-dimensional temperature field reconstruction [J]. Journal of Combustion Science and Technology, 2022, 28(2): 220-228(in Chinese).

- [15] KAHMEN O, ROFALLSKI R, LUHMANN T. Impact of stereo camera calibration to object accuracy in multimedia photogrammetry[J]. Remote Sensing, 2020, 12(12): 2057.
- [16] USAMENTIAGA R, GARCIA D F. Multi-camera calibration for accurate geometric measurements in industrial environments[J]. Measurement, 2019, 134: 345-358.
- [17] MCENALLY C S, KÖYLÜ Ü Ö, PFEFFERLE L D, et al. Soot volume fraction and temperature measurements in laminar nonpremixed flames using thermocouples[J]. Combustion and Flame, 1997, 109(4): 701-720.
- [18] FURUKAWA Y, PONCE J. Accurate camera calibration from multi-view stereo and bundle adjustment[J]. International Journal of Computer Vision, 2009, 84(3): 257-268.
- [19] GENOVESE K, CHI Y X, PAN B. Stereo-camera calibration for large-scale DIC measurements with active phase targets and planar mirrors[J]. Optics Express, 2019, 27(6): 9040.
- [20] SUN M S, GAN Z W, YANG Y Y. Numerical and experimental investigation of soot precursor and primary particle size of aviation fuel (RP-3) and n-dodecane in laminar flame[J]. Journal of the Energy Institute, 2021, 94: 49-62.
- [21] ABDALLA A O G, YING Y Y, JIANG B, et al. Comparative study on characteristics of soot from n-decane and RP-3 kerosene normal/inverse diffusion flames[J]. Journal of the Energy Institute, 2020, 93(1): 62-75.
- [22] WANG X R, DAI M L, YAN J, et al. Experimental investigation on the evaporation and micro-explosion mechanism of jatropha vegetable oil (JVO) droplets[J]. Fuel, 2019, 258: 115941.

Measurement of three-dimensional temperature and soot volume fraction for RP-3 jet fuel flame

ZUO Bingxian, ZHOU Bin*, DAI Minglu

(School of Energy and Environment, Southeast University, Nanjing 210096, China)

Abstract: RP-3 jet fuel is one of the frequently utilized fuels in aviation engines. Temperature and soot volume fraction distribution of flame are the vital data support for fuel optimization, which can directly indicate the state of combustion. The precision of the reconstruction path in three-dimensional measurement was influenced by camera position and attitude. A novel method that combines the radiation imaging system and the camera stereo calibration technology was proposed to address this problem. According to the results of numerical simulation, the method could not only guarantee the accuracy of measurement results, but also provide high spatial resolution. The experimental measurement of RP-3 and air premixed flame were carried out. The three-dimensional temperature and soot volume fraction distribution were also obtained by the calibrated six-camera system. The spatial resolution was $1 \text{ mm} \times 1 \text{ mm}$, and the uncertainty of temperature difference between reconstructed and thermocouple measured temperature was 0.58 K, demonstrating the accuracy of the method.

Keywords: RP-3 jet fuel; CCD camera; camera stereo calibration; temperature; soot volume fraction

Received: 2022-05-26; Accepted: 2022-07-22; Published Online: 2022-07-28 14:33 URL: link.cnki.net/urlid/11.2625.V.20220727.1738.002

Foundation items: National Key R&D Program of China (2017YFB0603204); National Natural Science Foundation of China (50976024,50906013)

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

DOI: 10.13700/j.bh.1001-5965.2022.0461

基于马蹄涡前缘吸气控制的压气机叶栅流动机理

唐耀璇,刘艳明*,安宇飞,孙运政

(北京理工大学 宇航学院,北京 100081)

摘 要:为探究从源头抑制角区分离的流动控制方法,采用数值模拟方法,以NACA65 叶栅为研究对象,利用叶栅前缘端壁吸气技术控制马蹄涡,结合拓扑分析对叶栅通道内的三维流动结构进行精确重构,以研究前缘端壁吸气控制马蹄涡进而改善叶栅通道流场结构的机理。结果表明:叶栅前缘端壁吸气技术可以有效推迟马蹄涡形成并削弱其强度,同时在吸气狭缝末端形成一对反旋流向涡对,在向下游发展过程中与通道涡相互作用;前缘端壁吸气通过控制马蹄涡,降低端壁边界层厚度,叶栅前缘通道涡发展受限,由回流组成的叶表分离被抑制;前缘及压力面侧吸气(EPS)直接作用于马蹄涡压力面分支,通道涡强度进一步被削弱,角区分离模式由闭式分离转为不完全闭式分离。最后,对比最优吸气系数下不同方案出口总压损失,发现当吸气量为进口质量流量的0.2%时,EPS方案出口截面总压损失降低5.8%;并且通过调整吸气系数,可以获得较好的变工况控制性能。

关键词:压气机叶栅;马蹄涡;角区分离;前缘吸气;拓扑分析

中图分类号: V231.3; TB553

文献标志码: A 文章编号: 1001-5965(2024)04-1282-10

角区分离是压气机中常见的端区流动现象,造 成压气机气动效率与稳定工作范围下降,是制约高 负荷压气机的主要因素之一^[1]。受叶栅/端壁结构 产生的逆压梯度及来流边界层中横向涡量的扭曲 和拉伸作用,压气机叶栅前缘边界层流体在分离点 附近向上卷起并沿展向拉伸形成马蹄涡^[2-3]。马蹄 涡经过前缘被分成压力面分支和吸力面分支,其中 马蹄涡压力面分支在向叶栅通道下游发展过程中 不断卷入边界层外部自由流,成为通道涡主要组成 部分,是产生角区分离的重要原因^[4],因此,控制马 蹄涡即可实现从源头抑制角区分离。

边界层吸气可通过削弱边界层厚度减小近壁 面低能流体卷起成涡的可能性,且具有良好的变工 况适应性。马蹄涡吸气控制研究最早基于翼身融 合体展开, Philips 等^[5]在翼型前缘开边长为 1.07T (*T* 为翼型最大厚度)的方形孔,通过实验验证前缘 吸气可以消除翼身融合体前缘的马蹄涡结构,同时 在孔的边缘会形成流向涡。

随后,马蹄涡吸气控制技术被引入内流叶轮机 中。Bloxham和 Bons^[6]认为涡轮通道分离起始于 马蹄涡产生处,并非通道涡遇到吸力面时,通过实 验验证涡轮叶片前缘端壁吸气能够降低总压损 失。Liesner 等^[7]以 NACA65-k48 叶栅为研究对象, 在马赫数 0.67、雷诺数 5.6×10⁵ 条件下进行实验及 数值模拟研究,根据马蹄涡压力面分支发展演化规 律在通道内布置非对称吸气槽,当吸气量为进口质 量流量的2.5%时,总压损失降低13%,叶栅效率提 高 1%。Guo 等^[8] 在压气机叶栅上游 13% 弦长处,沿 周向在整个通道端壁上开宽为1.67% 弦长的吸气槽, 通过实验验证发现通道涡对进口边界条件变化敏 感,当边界层吸气量达到进口质量流量的2.5%时,叶 栅通道总压损失减少约 1/4。Chen 等¹⁹对 Karsten 提出的方法进行改进,沿马蹄涡压力面分支发展路 径在端壁布置4段吸气槽,以抑制马蹄涡及通道涡

收稿日期: 2022-06-08; 录用日期: 2022-07-17; 网络出版时间: 2022-08-03 08:23

网络出版地址: link.cnki.net/urlid/11.2625.V.20220802.1058.001

^{*}通信作者. E-mail: liuym@bit.edu.cn

引用格式: 唐耀璇,刘艳明,安宇飞,等. 基于马蹄涡前缘吸气控制的压气机叶栅流动机理 [J]. 北京航空航天大学学报,2024,50(4): 1282-1291. TANG YX, LIU YM, AN YF, et al. Flow mechanism of horseshoe vortex suction control for compressor cascade [J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50(4): 1282-1291 (in Chinese).

对边界层内低能流体卷吸,通过数值模拟发现当吸 气量为进口质量流量的1%时,总压损失降低11.2%。 Zhang和 Chen^[10]通过实验研究近吸力面侧端壁非 定常吸气、吹气控制分离效果,发现前缘吸气、尾 缘吹气组合控制效果最好,分别推迟马蹄涡吸力面 分支、马蹄涡压力面分支形成。Liang等^[11]采用前缘 端壁及吸力面全展高组合吸气控制技术,通过数值 模拟发现端壁吸气通过抑制马蹄涡吸力面分支强 度降低叶栅损失,尤其在大冲角工况下表现良好。

以上研究表明前缘端壁吸气控制马蹄涡能够 抑制角区分离,降低通道内总压损失;但现有研究 尚未从马蹄涡绕流和拓扑结构的角度得出前缘吸 气的控制机理,因此,前缘吸气的马蹄涡控制效果 仍有较大提升空间。本文将结合拓扑分析方法,采 用前缘端壁吸气技术研究从源头控制马蹄涡进而 改善叶栅通道流动结构的机理,探索进一步提升控 制效果、抑制角区分离的方法和措施。

1 叶栅模型与数值方法

1.1 原型叶栅及吸气控制方案

本文以 NACA65 直列叶栅为研究对象,实验数 据来源于哈尔滨工业大学热力发动机实验室大尺 度低速风洞吹风实验^[12]。在叶栅通道模型中,x表 示展向、y表示节距方向、z表示流向。为探究端壁 不同位置吸气对压气机叶栅马蹄涡及角区分离的 控制作用,本文基于叶栅前缘曲率半径设计了3种 不同的吸气狭缝结构。文献 [13] 研究结果表明,在 流场中感受性较强的区域施加控制效果最好,因此 将狭缝布置在分离起始鞍点附近,取吸气狭缝宽度 0.2 mm, 距离叶栅前缘 1 mm; 仅考虑控制马蹄涡, 因此吸气狭缝沿流向布置在通道涡与马蹄涡吸力 面分支相遇前,即前缘及压力面侧吸气 (eading edge and pressure side, EPS)、前缘及吸力面侧吸气 (leading edge and suction side, ESS) 方案狭缝沿流向长度均 为15% 弦长, 前缘侧吸气 (leading edge suction, ES) 方案仅布置在叶栅前缘端壁,3种吸气狭缝布置方 案如<mark>图1</mark>所示。

1.2 网格及计算方法

网格采用 NUMECA 中的 AutoGrid5 模块自动 生成 O4H 型结构网格,吸气狭缝网格使用 IGG 模 块手动生成 O型、H型混合结构网格,壁面第 1 层 网格高度y⁺<1,为刻画叶栅前缘流场细节,对近端 壁及叶表的网格进行加密,最终确定叶栅通道网格 数约为 300 万,吸气狭缝网格数约为 30 万,原型叶 栅网格如图 2 所示。

数值计算使用商用软件 CFX 求解,选取剪切应 力输运 (shear-stress transport, SST) 湍流模型, Navier-



图 2 原型叶栅计算网格 Fig. 2 Computational grid for prototype blade

Stokes 方程组空间离散项、时间项均采用 High-Resolution 格式。计算边界条件与实验条件一致, 分别设置总压进口与静压出口,进口总温、总压分 别为 308 K、101 830 Pa,进口总压分布特性如图 3 所示,其中, $\bar{u} = u/u_{\infty}$,u为当地速度, u_{∞} 为自由流速 度, \bar{h} 为沿端壁法向无量纲高度;出口总温、静压分 别为 308 K、99 666 Pa^[12];通道两侧设置周期性边界 条件,叶片及端壁设置无滑移边界条件。吸气狭缝 网格与叶栅通道端壁网格之间采用 interface 连接, 通过调节吸气狭缝底部出口静压即可改变吸气 量。为尽可能减小吸气带来的附加影响,本文采用



图 3 叶栅进口附面层特性 Fig. 3 Boundary layer characteristic at blade cascade inlet surface

小流量吸气,保证压气机各项性能仍保持相当 水平。

图 4 为 0°、10°冲角下叶根处叶片型面静压升 系数与实验数据^[12]的对比结果,其中横坐标为用弦 长 C 无量纲化的流向位置,纵坐标 H_s为静压系数。 同时选择反映流场总体性能的参数进行校核,取距 叶片尾缘 46% 轴向弦长位置处节距平均能量损失 系数展向分布与实验结果进行对比,如图 5 所示, 其中 *ξ* 为能量损失系数,纵坐标为用叶栅展向高度 H 无量纲化的展向位置。可以看到数值结果与实 验数据基本吻合,说明本文采用的数值方法具有可 靠性。



图 4 叶高处叶表静压系数分布(x/H=4%)^[12]







2 计算结果与分析

本文将从前缘端壁吸气对前缘涡系及通道内 涡系结构的影响、前缘端壁吸气出口总压损失控制 效果及变工况性能等方面进行分析。

2.1 不同方案吸气流量的影响

由于各方案所开狭缝面积不同,本文首先研究 各方案不同吸气流量的影响,定义吸气系数 C_s为

$$C_{\rm S} = \frac{m_2}{m_1} \tag{1}$$

式中:m1和m2分别为进口质量流量和狭缝吸气流量。

图 6 为各方案 0°冲角时不同吸气系数下总压 损失系数的变化,其中 መ为总压损失系数,图中同 时给出了原型叶栅 (Ori) 0°冲角时的总压损失系 数。各方案最大吸气量均为吸气狭缝的堵塞流量, 图 6 并未给出更小的吸气量,吸气量更小时吸气狭 缝内流场不稳定。选择各方案中总压损失最小的 吸气系数进行流场结构分析及控制效果评估。



图 6 各方案不同吸气系数下总压损失系数分布

Fig. 6 Distribution of total pressure loss coefficients under different suction coefficient for each case

2.2 压气机内复杂流动结构分析

2.2.1 吸气对叶栅前缘涡系结构的影响

图 7 为不同方案叶栅前缘涡系结构示意图,涡 结构为基于 *Q* 准则^[14]的等值面 (*Q*=0.52),其中等值 面包络的涡结构内部条带为涡核位置,被来自上游 方向流体缠绕。叶栅前缘涡系由前缘角涡 (corner vortex, CV)、马蹄涡 (horseshoe vortex, HV)构成:边 界层流体受叶栅前缘压力梯度作用扭曲拉伸形成 马蹄涡,其拓扑形态为鞍点/分离点,垂直涡核的横 截面流态为不稳定螺旋点,从叶栅前缘向尾缘看, 马蹄涡吸力面分支 (horseshoe vortex suction side leg, HVS)旋向为逆时针,马蹄涡压力面分支 (horseshoe vortex pressure side leg, HVP)旋向为顺时针;近壁面 流体因叶栅/端壁固有角区沿物面切向攀升形成前 缘角涡,其拓扑形态为鞍点/附着点,横截面流态为

对比图 7 不同方案涡结构及涡核位置,结合 图 8 和图 9 不同方案马蹄涡压力面分支与吸力面 分支涡系结构图,取相应的叶高、节距(t)、弦长对 马蹄涡位置进行无量纲化,发现前缘端壁吸气对马 蹄涡位置及强度均有影响。图 8(b)、图 9(b)中马蹄











涡被明显抑制,表明仅在 ES 控制叶栅前缘马蹄涡效果最好;由图 7(c)、图 8(c)可知,在 EPS 马蹄涡压力面分支向叶片方向移动的距离为 0.8%*t*;由图 7(d)、图 9(d)可知,在 ESS 马蹄涡吸力面分支靠近叶片方

向移动的距离为 0.6%C, 同时在距前缘 2.8%C 处产 生流向涡。不同方案马蹄涡控制效果不同, 主要由 于马蹄涡吸力面分支涡轴方向与当地主流速度梯 度方向相反, 整个涡管持续受到周围流体压缩效应



Fig. 9 Suction side leg of horseshoe vortex system structure for different cases

强度减弱,该分支更容易被控制;根据文献 [15]可 知,相较于马蹄涡吸力面分支,马蹄涡压力面分支 强度更大,开缝面积相同的情况下马蹄涡压力面分 支侧吸气流量更大。

图 9(c)、图 9(d) 中存在 2 处流向涡量为负的位置, 根据涡核位置可知近叶片负涡量处为马蹄涡吸力面分支, 稍远处为吸气狭缝附近产生的流向涡, 两涡均与通道涡旋向相反。图 10(a) 可以进一步证明此涡的存在, 吸气狭缝左右边界当地速度矢量与其周围流体有很大区别, 这是由于吸气时狭缝周围 产生局部低压区, 外部高压流体会形成切向狭缝腔



室内壁的流向涡对。从叶栅前缘向尾缘方向看,狭 缝左侧流向涡 (SSV1) 旋向为顺时针,右侧流向涡 (SSV2) 旋向为逆时针。

2.2.2 前缘端壁吸气对叶栅通道内流动的影响

Dallmann^[16]根据物面上摩擦应力场中临界点 来判断拓扑类型,单通道无间隙叶栅拓扑结构中鞍 点与螺旋点数量关系式为

$$\sum N - \sum S = 0 \tag{2}$$

式中:N为结点,S为鞍点。

绘制施加控制前后吸力面与端壁极限流线图 谱及拓扑结构如图 11 所示。根据图 11(a)分析原 型叶栅拓扑结构及其对应三维分离流动:当流体输 运至叶栅通道内 40% 弦长处,叶片前缘势效应带来 的静压升使得边界层内低能流体聚集形成鞍点,其 附近两支极限流线交汇于此^[4],一支由通道涡输运 至吸力面诱导而成 (*S*₂-*N*₂),另一支由沿角区反向运 动的回流构成 (*S*₄-*N*₄);角区角涡 (*S*₃-*N*₃)出现在 50% 弦长处,沿流向在端壁、叶表分别演化形成分离 线、附着线;吸力面表面"龙卷风"形式分离涡由通 道涡诱导的角区闭式分离 (*S*₅-*N*₅)与回流造成的叶 表分离 (*S*₆-*N*₆) 2 部分组成;分离涡沿展向向叶中迁 移,形成脱落涡 (*S*₇-*N*₇)。

对比图 11 中各方案吸力面表面拓扑结构, ES 方案吸力面结点数量减少, EPS、ESS 方案中回流造 成的叶表分离 (*S*₆-*N*₆)范围减小,表明前缘端壁吸气 通过抑制回流来改善角区分离。前缘端壁吸气降 低进口边界层厚度,通道涡对进口边界层变化十分





敏感,叶栅前缘通道涡强度随之降低,通道内通流 面积增加,流体流速升高,回流区范围变小,角区回 流被抑制。

图 12 为叶栅通道内 10% 弦长处端壁边界层位 移厚度δ*、动量厚度δ**示意图,δ*、δ**均使用叶片 最大厚度T进行无量纲化。从图 12 中可以看出吸 气狭缝附近及近吸力面侧端壁边界层厚度降低。





吸气时狭缝附近流体与周围环境流体形成剪切层, 其包含的涡量在不稳定性的作用下卷起成涡环,其 包裹的流体受到卷吸作用向狭缝方向运动并进入 腔室,将前缘边界层内部分低能流体输运至抽吸狭 缝内,边界层厚度降低。

与原型相比, ES 方案边界层厚度几乎没有变 化,表明虽然 ES 控制前缘马蹄涡效果明显,但是小 流量吸气仅作用于叶栅前缘,无法对下游马蹄涡发 展产生影响; EPS 方案中吸力面侧端壁边界层厚度 明显降低,这是由于前缘端壁吸气通过调节端壁静 压分布,改善叶栅通道固有非对称流体环境,前缘 端壁压力梯度被削弱,横向二次流减弱,边界层厚 度降低; ESS 方案则主要降低狭缝附近边界层厚 度,对于较远位置处边界层厚度没有影响。

选择叶栅前缘边界层厚度变化明显的 EPS 方 案,给出图 13 近叶栅吸力面边界层厚度沿流向变 化示意图,进一步分析前缘端壁边界层厚度变化对 通道下游流体发展演化的影响。由图 13 可知在 50% 弦长处分离区迅速增长,叶栅通道内通流面积 减小,在 80% 弦长后基本趋于稳定;形状因子 (*H*₁₂) 为边界层位移厚度与边界层动量厚度之比,在 90% 弦长以后不再增加,呈稳定"凹"型分布,即回



1288



流(S₆-N₆)主要在90%弦长以后稳定发展。

对比图 11 中各方案端壁表面拓扑结构, ES、 ESS 方案端壁拓扑几乎没有变化, 而 EPS 方案端壁 拓扑结构变得更加复杂。定义当地气流偏转角 *a*^[17] 表达式为

 $\alpha = \arctan \frac{v_y}{v_z} \tag{3}$

式中: v,为当地周向速度; vz为当地轴向速度。

结合图 14(a) 叶栅前缘气流偏转角沿展向分 布, ES 方案有效抑制马蹄涡展向攀升, 这很好地解 释了 ES 方案控制马蹄涡效果好的原因; 但是 ES、 ESS 方案对改善通道横向二次流作用效果较弱, 因 此其端壁拓扑结构几乎没有变化。结合图 14(b) 近 端壁气流偏转角沿流向分布, 发现 EPS 方案吸气狭 缝附近气流偏转角减小, 即叶栅前缘端壁横向流动 减弱; 同时前缘端壁吸气引起的逆压梯度对马蹄涡 压力面分支施加了压缩效应, 马蹄涡压力面分支强 度降低, 通道涡由稳定状态转向不稳定状态进而被 削弱, 角区端壁处基本奇点变成一系列不稳定(二 阶)奇点, 分离涡变成许多小的耗散结构, 分离减 弱^[13], 这种效应从 5% 弦长一直持续至 50% 弦长处。 2.2.3 具有前缘端壁吸气控制的叶栅涡系结构

基于上述分析,构建如图 15 所示 EPS 方案叶 栅通道涡系结构。叶片及端壁表面拓扑结构如下: S_H、N_H分别为马蹄涡产生带来的前缘鞍点和尾缘 分离结点,S_T、N_T分别为叶栅尾缘鞍点和结点,S_C、 N_C分别为角区分离起始鞍点和结点,S_B、N_B分别为 压气机尾缘回流带来的分离鞍点和结点,S_M、N_M分 别为叶栅压力面脱落涡带来的分离鞍点和结点。 定义从叶栅前缘向尾缘方向看逆时针旋向为正 (+),顺时针旋向为负(-),涡系演化过程如下:

1) 马蹄涡 (HV) 在叶栅前缘形成并分为 2 个分



图 14 吸气控制对气流偏转角的影响

Fig. 14 Influence of suction control on flow deviation angle



图 15 具有前缘及压力面侧吸气的叶栅通道涡系演化示意图 Fig. 15 Schematic evolution of vortex flow pattern with suction at leading edge and pressure side

支,伴随振荡运动向下游延伸,马蹄涡压力面分支 (HVP,-)穿过通道向吸力面移动成为通道涡的主要 组成部分,在叶栅通道内约 40% 弦长处与对应的马 蹄涡吸力面分支 (HVS,+) 汇合。

2) 通道涡 (PV,-) 在向吸力面输运过程中通过

吸引主流强度增强。

3)吸气狭缝末端产生反旋涡对,其中SSV1(-) 与通道涡旋向相同,向下游输运过程中卷吸部分 PV边缘流体,SSV2(+)则在向下游输运过程中逐渐 耗散。

4) 角涡 (CVS,+) 的上洗作用与 PV 的下洗作用 在角区形成扭转运动。

5) 在逆压梯度和横向压力梯度作用下, PV 汇 集成展向涡,离开叶表时形成脱落涡 (CSV,+)。

2.3 前缘端壁吸气对出口截面总压损失的影响

2.3.1 0°工况控制效果

表1为距叶片尾缘46%轴向弦长位置处总压损失质量流量平均结果,与原型相比EPS方案出口截面总压损失降低5.8%,其他方案总压损失均升高。

表 1 不同吸气方案总压损失系数 Table 1 Total pressure loss coefficient in different

	•
notion	6000000000
	менатих
action	500mm 105

方案	$\varpi/\%$	$C_{\rm s}/\%$
Ori	5.49	
ES	5.70	0.005 8
EPS	5.17	0.207 0
ESS	5.75	0.016 2

图 16 为距叶片尾缘 46% 轴向弦长位置处总压损失节距平均结果,由图 16 可知 EPS 方案 *x/H*=0.1~0.25 区域总压损失降低,即 EPS 方案通过削弱通道涡降低角区分离损失。EPS 方案降低出口总压损失机理如下:





 EPS 方案压力面侧吸气有效控制马蹄涡压 力面分支强度进而削弱通道涡,流体在角区堆积减 少,叶栅通道通流面积增加,尾迹低速区与主流掺 混产生的损失降低。

2)吸气时吸气狭缝末端会伴随反旋流向涡对 的产生, EPS 方案中 SSV1 与通道涡旋向相同, SSV1 在向下游发展的过程中卷吸部分通道涡边缘流体, 诱导通道涡发展方向; 而 ESS 方案中 SSV2 与通道 涡相互作用, 两旋向相反旋涡在向下游输运过程中 发生强烈的剪切作用, 相应地带来额外的总压损失。

3) 分离模式变化对总压损失有影响,根据 Kutta-Joukowsky 定理,闭式分离将高能主流区和低能黏 滞流区分隔开来,所以在分离剪切层两侧的气流总 压不相等并伴随有总压损失,而开式分离不会产生 总压损失。结合图 11(c)可知, EPS 方案分离模式 为不完全闭式分离,相比于原型叶栅闭式分离状态 总压损失降低。

2.3.2 变工况控制性能

为分析前缘吸气的变工况控制性能,选择能够 降低叶栅出口总压损失的 EPS 方案,验证本文设计 的前缘端壁吸气方案的变工况性能。图 17 给出了 原型和 EPS 方案叶栅出口截面总压损失随冲角变 化示意图及总压损失系数降低百分比,吸气系数选 各工况下的最优方案。由图 17 可知,前缘端壁吸 气出口截面总压损失系数在各工况下均小于原型 叶栅。相比于原型叶栅,正冲角下总压损失降低程 度随冲角增大而增加,随冲角增大控制效果越发显 著;而在-6°冲角下总压损失降低程度较小。正冲 角工况下马蹄涡压力面分支提前向对应的吸力面 移动,前缘及压力面侧吸气通过改变吸气系数,能 更有效地抑制叶栅前缘端壁横向二次流,因而在大 冲角情况下降低总压损失效果更好。



Fig. 17 Distribution of total pressure loss coefficients under different incidence angle

3 结 论

本文采用数值方法,通过分析在前缘不同位置 小流量吸气叶栅前缘及通道内涡系结构、出口截面 总压损失变化情况及变工况性能,研究前缘端壁吸 气对马蹄涡及流动分离的控制机理,得到的相关结 论如下:

1) 前缘端壁吸气技术可以有效控制由边界层

发展而来的马蹄涡。其中前缘侧吸气 (ES) 抑制马 蹄涡展向攀升,有效控制叶栅前缘马蹄涡;前缘及 压力面侧吸气 (EPS)、前缘及吸力面侧吸气 (ESS) 方案分别控制马蹄涡压力面分支、马蹄涡吸力面分 支向下游的发展。

2) 在吸气狭缝末端会形成一对反旋流向涡对 (SSV1、SSV2), 在向下游发展过程中与通道涡相互 作用: EPS 方案中流向涡 SSV1 与通道涡旋向相同, 卷吸通道涡边缘流体; ESS 方案中流向涡 SSV2 与 通道涡旋向相反, 两涡相互剪切并产生总压损失。

3) 前缘端壁吸气通过控制马蹄涡从源头削弱 通道涡强度, 抑制由回流组成的叶表分离 (S₆-N₆)。 EPS 方案还通过削弱马蹄涡压力面分支, 使通道涡 由稳定状态转向不稳定状态, 角区端壁拓扑结构由 基本奇点变成一系列不稳定(二阶)奇点, 角区分离 模式由闭式分离变为不完全闭式分离, 分离进一步 被抑制。

4) EPS 方案可以降低叶栅出口总压损失,当吸 气量为进口质量流量的 0.2% 时,出口截面总压损 失降低 5.8%;在变工况条件下,通过调整吸气系数 可以取得较好的控制效果,且在正冲角下控制效果 更加明显。

参考文献(References)

[1] 吴艳辉, 王博, 付裕, 等. 轴流压气机角区分离的研究进展[J]. 航 空学报, 2017, 38(9): 107-128.

WU Y H, WANG B, FU Y, et al. Research progress of corner separation in axial-flow compressor[J]. Acta Aeronautica et Astronautica Sinica, 2017, 38(9): 107-128(in Chinese).

- [2] BAKER C J. Vortex flow around the bases of obstacles[D]. Cambridge : University of Cambridge, 1979.
- [3] ROBINSON S K. Coherent motions in the turbulent boundary layer[J]. Annual Review of Fluid Mechanics, 1991, 23(1): 601-639.
- WANG H, OLSON S J, GOLDSTEIN R J, et al. Flow visualization in a linear turbine cascade of high performance turbine blades[J]. Journal of Turbomachinery, 1997, 119(1): 1-8.
- [5] PHILIPS D B, CIMBALA J M, TREASTER A L. Suppression of the wing-body junction vortex by body surface suction[J]. Journal of Aircraft, 1992, 29(1): 118-122.

- [6] BLOXHAM M J, BONS J P. Leading-edge endwall suction and midspan blowing to reduce turbomachinery losses[J]. Journal of Propulsion and Power, 2010, 26(6): 1268-1275.
- [7] LIESNER K, MEYER R, LEMKE M, et al. On the efficiency of secondary flow suction in a compressor cascade[C]//Proceedings of the ASME Turbo Expo 2010: Power for Land, Sea, and Air. NewYork: ASME, 2010: 151-160.
- [8] GUO S, LU H, CHEN F, et al. Vortex control and aerodynamic performance improvement of a highly loaded compressor cascade via inlet boundary layer suction[J]. Experiments in Fluids, 2013, 54(7): 1570.
- [9] CHEN P, QIAO W, LIESNER K, et al. Effect of segment endwall boundary layer suction on compressor 3D corner separation[C]// Proceedings of the ASME Turbo Expo 2015: Turbine Technical Conference and Exposition. NewYork: ASME, 2015.
- [10] ZHANG H, CHEN S. A comparative experimental analysis of two unsteady flow control methods in a highly loaded compressor cascade[J]. Experiments in Fluids, 2020, 61(6): 132.
- [11] LIANG T, LIU B, SPENCE S. Effect of boundary layer suction on the corner separation in a highly loaded axial compressor cascade
 [J]. Journal of Turbomachinery, 2021, 143(6): 61002.
- [12] 赵桂杰. 弯掠扩压叶栅内附面层与二次流控制的研究[D]. 哈尔 滨: 哈尔滨工业大学, 2005.
 ZHAO G J. Study on boundary layer and secondary flow control in bowed-swept diffuser cascade[D]. Harbin: Harbin Institute of Technology, 2005(in Chinese).
- [13] TOBAK M, PEAKE D J. Topology of three-dimensional separated flows[J]. Annual Review of Fluid Mechanics, 1982, 14(1): 61-85.
- [14] 刘超群. Liutex-涡定义和第三代涡识别方法[J]. 空气动力学学报, 2020, 38(3): 413-431.
 LIU C Q, Liutex-third generation of vortex definition and identification methods[J]. Acta Aerodynamica Sinica, 2020, 38(3): 413-431(in Chinese).
- [15] HAO X, LE C, BAO L, et al. Numerical Investigations on Oscillating aspiration control in high-load compressor cascades[J]. Journal of Propulsion and Power, 2019, 35(4): 850-862.
- [16] DALLMANN U. Three-dimensional vortex structures and vorticity topology[J]. Fluid Dynamics Research, 1988, 3(1-4): 183-189.
- [17] 张华良. 采用叶片弯/掠及附面层抽吸控制扩压叶栅内涡结构的研究[D]. 哈尔滨: 哈尔滨工业大学, 2007. ZHANG H L. Investigation on application of dihedral/swept blade and boundary layer suction to control vortex configuration in compressor cascades[D]. Harbin: Harbin Institute of Technology, 2007(in Chinese).

Flow mechanism of horseshoe vortex suction control for compressor cascade

TANG Yaoxuan, LIU Yanming*, AN Yufei, SUN Yunzheng

(School of Aerospace Engineering, Beijing Institute of Technology, Beijing 100081, China)

Abstract: To explore the source flow control method for the corner separation, this paper takes the NACA65 cascade as the research object and applies the blade leading edge endwall suction technology to control the horseshoe vortex with the numerical simulation method. Combined with topology analysis, the three-dimensional flow field is accurately reconstructed and the control mechanism of the leading edge endwall suction is revealed to improve the cascade channel flow field performance. The results show that the leading edge endwall suction technology can effectively delay the formation of the horseshoe vortex and weaken its strength. Meanwhile, a pair of counter rotating vortices are formed at the end of the suction slit, which interacts with the passage vortex in the process of downstream development. By regulating horseshoe vortices, the suction lessens the thickness of the endwall boundary layer, inhibiting the formation of leading edge passage vortices and preventing surface separation caused by backflow. Because the leading edge and pressure side suction (EPS) directly acts on the pressure side leg of the horseshoe vortex, the passage vortex strength is further weakened, and the corner separation mode changes from closed separation to incomplete closed separation. Finally, the total pressure losses are compared at the outlet under the optimal suction coefficient. It is discovered that when the suction coefficient is 0.2%, the overall pressure loss at the EPS outflow section is decreased by 5.8%. At off-design situations, improved control performance can be attained by modifying the suction coefficient.

Keywords: compressor cascade; horseshoe vortex; corner separation; leading edge suction; topology analysis

Received: 2022-06-08; Accepted: 2022-07-17; Published Online: 2022-08-03 08:23

URL: link.cnki.net/urlid/11.2625.V.20220802.1058.001

^{*} Corresponding author. E-mail: liuym@bit.edu.cn

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

DOI: 10.13700/j.bh.1001-5965.2022.0493

可变弯度翼身融合布局气动特性分析与设计

王雨桐^{1,2}, 蓝庆生¹, 周铸^{1,*}, 杨体浩², 宋超¹

(1. 中国空气动力研究与发展中心 计算空气动力研究所, 绵阳 621000; 2. 西北工业大学 航空学院, 西安 710072)

摘 要: 变弯度机翼在提高常规布局客机气动特性方面有较大的潜力,但也会引起全机俯仰力矩变化,考虑翼身融合布局飞行器力臂短、配平阻力较大的特点,研究变弯度技术在翼身融合 布局飞行器上的减阻收益与配平惩罚。从工程实际出发,采用基于舵面偏转的方式实现后缘变弯度 并对比分析不同展向位置处舵面的配平能力;然后利用全局优化方法开展变弯度气动减阻优化设 计;最后对变弯度设计空间进行探索。结果表明:随着升力系数的改变,产生配平阻力最小的舵面 位置也会发生变化。当不考虑俯仰力矩配平约束时,采用变弯度技术至多可以获得4.62%的减阻收 益;在考虑俯仰力矩配平约束后,相比于采用中央体后缘舵面配平,采用变弯度技术至少能够减小 2.4×10⁻⁴的配平损失。不同升力系数下,变弯度的舵面偏转组合方式存在明显差异,小升力系数 下,多个舵面负偏的变弯度组合有利于减阻并增加抬头力矩;而大升力系数下则是通过多个舵面正 偏的组合实现减阻,但会导致低头力矩增加。基于多舵面组合偏转的变弯度减阻收益与力矩惩罚评 估,能为工程上设计可变弯度翼身融合布局飞行器提供参考。

关 键 词: 翼身融合; 后缘变弯度; 气动特性; 优化设计; 代理模型 中图分类号: V221.41⁺

文献标志码: A 文章编号: 1001-5965(2024)04-1292-16

为实现未来的超高效远程客机,迄今为止,各 种新概念气动布局大量涌现。其中翼身融合 (blended wing body, BWB)布局由于具有巡航效率 高、结构质量轻、燃油效率高、低噪声等特点得到 广泛关注^[1-3]。BWB 布局通常采用无尾布局形式, 机翼机身高度融合形成全升力面外形。在相同装 载要求下, BWB 布局的全机浸润面积更小从而减 少了摩擦阻力。Brown 和 Vos^[4]给出了亚声速民机 巡航效率与布局之间的关系,几种代表性 BWB 民 机方案的巡航效率因子已达 20 以上,显示出巨大的 性能优势和发展潜力。相对 B787 和 A350, 同客 座级的 BWB 客机巡航效率可提高 15%~20%^[5]。 Okonkwo 和 Smith^[6] 回顾了 BWB 的发展历史, 王刚 等^[5]介绍了 BWB 总体气动技术的研究进展。

机翼后缘变弯度的设计增强了机翼适应各种 飞行条件的能力,使工程师能在各工况下设计机翼 不同的后缘偏角以改善其气动性能,而不是一系列 飞行条件下的气动性能之间的折衷,是一种有效的 减小燃油消耗的技术^[7-8]。目前,机翼后缘变弯度技 术在常规布局上已有初步的工程应用。Boeing 和 NASA 共同开展了"可变弯度后缘襟翼"(variable camber continuous trailing edge flap, VCCTEF)项目研 究^[9],结果表明变弯度技术可以提高巡航效率、扩 宽抖振边界、增加机动性等性能优势。文献 [10-11] 也证明了 VCCTEF 的气动优势。此项技术被用于 Boeing757 的原型机 (generic transport model, GTM) 构型上,有效地减小了其巡航阻力^[12]。文献 [13-14] 在非常规布局桁架机翼上采用 VCCTEF 的研究,结 果表明在非常规布局上采用可变弯度后缘也可以 进一步提高飞机的整体性能。

将变弯度机翼与 BWB 布局飞行器相结合,可 以进一步提高 BWB 布局飞行器巡航阶段的气动

收稿日期: 2022-06-16; 录用日期: 2022-09-19; 网络出版时间: 2022-09-23 07:57

网络出版地址: link.cnki.net/urlid/11.2625.V.20220922.1003.001

^{*}通信作者. E-mail: f_yforever@126.com

引用格式: 王雨桐, 蓝庆生, 周铸, 等. 可变弯度翼身融合布局气动特性分析与设计 [J]. 北京航空航天大学学报, 2024, 50(4): 1292-1307. WANG Y T, LAN Q S, ZHOU Z, et al. Design and aerodynamic analysis of blended wing body with variable camber technology [J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50(4): 1292-1307 (in Chinese).

效率高,以实现超高效远程客机的目标。虽然 BWB 布局在气动性能上优势明显,但由于取消了 传统意义上的平尾和垂尾, BWB 布局飞行器的操 纵面主要布置在机身的后缘部分,其纵向力臂较常 规布局飞机短,因而导致了俯仰操纵效能大大降 低。配平相同俯仰力矩, BWB 布局将付出更大的 配平阻力惩罚。从工程实际出发,采用全翼展后缘 连续变弯度技术的设计变量较多,且对机翼变形材 料要求较高,在规律研究上具有指导意义,但是针 对大型客机,连续变弯技术的技术成熟度和可靠性 有待进一步提高,波音公司在 B787 上开发的后缘 可变弯度系统^[15]是通过操纵后缘舵面偏转实现变 弯度的, Reist 等^[16] 在常规布局公务机上开展了基 于舵面偏转的可变弯度研究,探究了减阻机理,并 指出综合考虑机翼、机身和其他升力表面之间的相 互作用对于设计变弯度机翼至关重要。何萌等[17] 在常规布局客机上开展了基于舵面偏转的变弯度 技术研究,研究结果表明变弯度技术会引起全机俯 仰力矩变化,带来的配平阻力对变弯度减阻收益影 响明显。

现有研究主要围绕基于常规布局飞机的后缘 变弯度对气动特性的影响研究,对于在 BWB 布局 飞行器上使用变弯度技术的研究较少。Mitridis 等^[18] 评估了在 BWB 布局上采用基于后缘襟翼偏转的变 弯度技术的可行性,结果表明相比于传统的襟翼偏 转,采用变弯度技术能获得更好气动性能。但整个 工作并未考虑舵面偏转对配平的影响,并且仅针对 展向固定位置处的舵面进行了对比研究,并未考虑 展向不同位置处舵面性能之间的差异,以及实际使 用中 BWB 布局飞行器后缘是多舵面配置的情况。 本文在考虑了 BWB 后缘多舵面配置的基础上,研 究了基于舵面偏转的变弯度技术对 BWB 布局飞行 器气动特性的影响,揭示了变弯度技术的减阻原 理。在考虑配平的情况下,量化变弯度技术对 BWB 布局飞行器带来的减阻收益与俯仰力矩惩罚,并探 究了舵面偏转组合方式与气动性能的关系。

1 优化方法

1.1 数值计算方法

本文使用基于雷诺平均 Navier-Stokes 方程的计 算流体力学 (computational fluid dynamics, CFD)流 场模拟方法,采用隐式时间推进方法、空间离散格 式为 JST 二阶中心格式,全湍流计算,湍流模型为 剪切应力传输 (shear stress transport, SST)。求解器 验证模型采用 ONERA-M6 机翼。利用商业软件 ICEM 划分多块结构网格,物面和对称面网格如 图 1 所示。网格量为 158 万,边界层第一层高度为 1×10⁻⁶,计算状态为 *Ma*=0.84, *a*=3.06°, *Re*= 11.72×10⁶。求解器的计算结果和实验数据进行对 比,如图2所示,图中实线为计算得到的截面压力 分布,圆点为实验数据,可以看出计算结果和实验 数据在各个机翼展向截面吻合均较好,说明采用的 CFD求解器满足优化设计的精度要求。



图 1 M6 机翼网格示意图 Fig. 1 Schematic diagram of M6 wing grid

1.2 优化算法

本文主要采用具备精英策略的非支配排序遗 传算法-II(non-dominated sorting genetic algorithm-II, NSGA-II)进行寻优。针对多目标问题适应度 难以直接确定的特点,该方法利用 NSGA 对整个种 群进行分层,得到 n 级非支配层,并为每一级非支 配层赋予相应的虚拟适应度值,级数越大,该适应 度值越小。NSGA-II 在 NSGA 的基础上做出如下 改进^[19]:

1)提出快速非支配排序法,降低了算法的计算 复杂度。

2)采用拥挤度和拥挤度比较算子,代替需要指 定共享半径的适应度共享策略,而且可以作为快速 排序后的同级比较中的胜出标准,保持了种群的多 样性。

3) 引入精英策略, 扩大采样空间。将父代种群 与其产生的子代种群组合, 共同竞争产生下一代种 群, 有利于保持父代中的优良个体进入下一代, 并 通过对种群中所有个体的分层存放, 使得最佳个体 不会丢失, 迅速提高种群水平。

NSGA-II 的基本思想为^[19]:通过随机模拟产生 规模为N的初始种群 P_i ,产生子代种群 Q_i ,并将上 述 2 个种群结合在一起形成规模为 2N的种群 R_i ; 合并父代种群与子代种群,进行快速非支配排序, 并对每个非支配层中的个体进行拥挤度计算,根据 非支配关系及个体的拥挤度选取合适的个体形成 新的父代种群 P_{i+1} ;通过遗传算法操作产生新的子 代种群 Q_{i+1} ,将 P_{i+1} 与 Q_{i+1} 合并形成新的种群 R_i ,重 复以上操作,直到满足结束条件为止。相应的程序 流程如图 3^[20]所示。











1.3 气动代理模型

遗传算法在优化过程中,通常需要计算上万的

种群,若全部调用高可信度 CFD 分析,会使得整个 优化过程耗费大量的计算资源。为降低计算成本, 提高优化效率,采用代理模型方法,基于一组采样 点,通过高保真的雷诺平均 Navier-Stokes 方程进行 评估,在此基础上生成气动代理模型。

常用的代理模型包括多项式响应面、径向基函数、Kriging模型、人工神经网络、支持向量回归、 多变量插值与回归、多项式混纯展开等方法。其中Kriging模型作为估计方差最小的无偏估计模型,在解决非线性程度较高的问题时较容易取得理想的拟合结果。考虑到变形过程中由于激波的影响可能存在的非线性特性,本文采用Kriging模型, 建立后缘连续变弯度与BWB构型气动特性之间较为精确的联系,研究变弯度技术在BWB布局上的减阻情况。

在构建代理模型时,初始样本点的采样通过拉 丁超立方(latin hypercube sampling, LHS)方法完 成。对于一个具有*m*维设计空间,*p*维输出空间的 代理模型建模问题,通过试验设计得到*n*个样本点 $X = [x_1, x_2, \dots, x_n], 其中 x_i = [x^1, x^2, \dots, x^m]^T \in \mathbb{R}^m$ 。对

1295

每个样本点进行数值分析得到对应的响应值 $Y = [y_1, y_2, ..., y_n]$,其中 $y_i = [y^1, y^2, ..., y^m]^T \in \mathbb{R}^p$ 。基 于 *X*和 *Y*构建该问题的 Kriging代理模型。 Kriging模型是一种插值模型,其插值结果定义为已 知样本函数响应值的线性加权,即

$$\hat{\mathbf{y}}(\mathbf{x}) = \sum_{i=1}^{n} \boldsymbol{\omega}^{(i)} \mathbf{y}_i \tag{1}$$

式中: **ω** = [**ω**⁽¹⁾, **ω**⁽²⁾, ..., **ω**⁽ⁿ⁾]是插值的加权系数。为 了计算加权系数, Kriging 模型引入统计学假设:将 未知函数看成是某个高斯稳态随机过程的具体实 现。对于 BWB 的空气动力学建模问题,该稳态随 机过程定义为

 $C_x(\boldsymbol{x}) = b + z(\boldsymbol{x}) \tag{2}$

式中: C_x 为任意的气动系数, $x = [\delta_0, \delta_1, \dots, \delta_5, Ma]^T$ 为输入向量; b为未知的全局趋势函数; z(x)为平稳 正态分布高斯随机过程的实现, 均值为 0, 方差和协 方差均不为 0。

1.4 优化设计系统

本文搭建的优化设计流程如图 4 所示,具体步骤如下:

1)试验设计通过拉丁超立方方法在设计参数
 空间内进行采样并获取设计参数初始样本。

2) 通过自由曲面变形(free form deform,FFD)参





数化方法实现 BWB 中央体和机翼后缘基于舵面偏转的变弯度。

3) 通过动网格技术自动生成变形后的网格, 通过 CFD 求解器获得初始样本点的气动力系数。

4) 基于初始样本点的输入(基于舵面偏转的变 弯角度)和输出(气动力系数),按照一定比例分配 训练集、测试集和验证集,训练代理模型。

5)验证代理模型精度,若精度满足要求,则采 用遗传算法 NSGA-II 在设计空间内寻优;若精度不 满足要求,则增加样本点,重新训练代理模型。

6)最后得到最优构型的变弯偏角和气动力 系数。

其中参数化模块采用的 FFD 参数化,作为目前 使用较为广泛的几何参数化方法,其基本思想为通 过对待变形物体的变形量进行参数化达到控制变 形目的。该方法避免了对几何外形本身进行参数 化拟合,大幅度降低了对初始构型的数据格式和布 局形式的依赖程度,并能够光滑描述曲线、曲面和 三维几何体的外形,方便地应用于整体和局部外形 设计。

2 基准构型设计分析

2.1 初始构型及后缘舵面划分

本文的 BWB 参考 Liou 等^[21]设计的 BWB 布局, 展长为 65.5 m, 机身长为 41 m, 参考面积为 849.51 m², 展弦比为 4.9, 本文定义 20% 以内为中央 体, 40% 以外为外翼段, 20%~40% 为过渡区。参考 Felder 等的研究^[22], 选择巡航设计点为:马赫数 Ma = 0.84, 升力系数 $C_L = 0.19$, 静稳定裕度为 7.83%。由于变弯度技术能够在巡航过程中通过改变后缘偏角,在非设计点状态下获得减阻收益, 从 而降低整个巡航段的平均阻力。考虑飞行过程中 随着燃油不断消耗,飞机的升力系数也会随时间发 生变化,本文选择 2 个非设计点状态,升力系数分别为 $C_L = 0.14 和 C_L = 0.24$ 。

考虑实际使用中 BWB 布局飞行器都是多舵 面配置的情况,参考文献 [22-23],后缘变弯度时使 用的舵面划分如图 5 所示,沿展向布置 6 个控制 舵面,包括中央体后缘 1 个舵面和机翼后缘 5 个 舵面。从中央体到翼尖,依次命名为舵 0~舵 5。 本文选取各舵面偏转范围均为 [-4°,+4°],规定下 偏为正,上偏为负。后缘变弯度采用定轴偏转方 式,通过 FFD 参数化方法实现, FFD 控制体如图 6 所示。

2.2 基准构型优化设计

采用离散伴随气动优化方法对 BWB 布局进行



图 5 舵面划分及剖面布置示意图





图 6 后缘偏转的 FFD 控制体

Fig. 6 FFD control volume for trailing-edge deflection

单点减阻优化,并实现构型自配平,将优化结果作 为后续研究变弯度技术的基准构型。

单点优化采用的 CFD 网格和 FFD 控制体如 图 7 所示, FFD 控制体沿展向设置 12 个控制剖面, 每个控制剖面上各有 19 个控制点分别控制上下表 面,采用各个 FFD 控制点的 Z 向位移和飞行器迎角 作为设计变量,共457 个设计变量。优化约束包括 翼剖面厚度约束和俯仰力矩约束。初始构型单目 标设计问题定义如下:

Min : C_D

	$\begin{pmatrix} C_L = 0.19 \\ C_L = 0 \end{pmatrix}$	(3)
5.1.	$\begin{array}{c} C_M = 0 \\ t \ge t_0 \end{array}$	

式中: C_L 、 C_D 、 C_M 分别为升力系数、阻力系数、俯仰力矩系数;t为当地厚度; t_0 为当地初始厚度。

优化后获得本文的基准构型,巡航状态 CFD 结



图 7 优化网格及 FFD 控制体 Fig. 7 Optimized grid and FFD control volume

果如表 1 所示。设计点处力矩自配平,巡航迎角 a为3.12°,阻力系数为9.04×10⁻³,升阻比为21。图 8 为设计点处机翼上表面压力云图及流线分布,图 9 为设计点处基准构型剖面压力系数,从图中可以看 到,中央体和翼身过渡区压力分布基本为无激波形 态,外翼段有弱激波。

表 1 基准构型气动力系数 Table 1 Aerodynamic coefficient of baseline configuration

α/(°)	C_L	C_D	C_M	L/D
2.58	0.14	0.007 95	0.004 0	17.62
3.12	0.19	0.009 04	0.000 1	21.00
3.63	0.24	0.011 48	-0.004 9	20.91



图 8 Ma = 0.84, C_L = 0.19基准构型表面压力及 流线图

Fig. 8 Surface pressure and streamline diagram of baseline configuration (Ma = 0.84, $C_L = 0.19$)

2.3 舵面配平能力分析

配平阻力对 BWB 布局飞行器气动性能影响显 著。本节研究不同位置舵面单独偏转进行俯仰力 矩配平时对阻力和俯仰力矩的改变情况,在此基础 上对比分析了不同升力系数下单舵面偏转所需的 配平阻力惩罚。

图 10 为 C_L = 0.14和 C_L = 0.24时各舵面单独偏转时阻力变化量与俯仰力矩变化量的关系,其中黑色虚线表示配平所需的俯仰力矩变化量。由于中央体后缘配平力矩最长,通常采用中央体后缘,即舵 0 进行配平^[2]。但根据图 10 所示, C_L = 0.14时采用舵 5 即翼尖处的后缘进行配平能带来最小的阻力惩罚。 C_L = 0.24时采用用舵 2 进行配平所带来的阻力惩罚最小。

表 2 为*C_L* = 0.14时通过舵 0 和舵 5 单独进行配 平的气动力结果。采用舵 0 配平时, 舵偏角为 4.91°, 阻力增加 2.2×10⁻⁴, 采用舵 5 配平时, 舵偏角 为 3.83°, 阻力增加 1.1×10⁻⁴。采用舵 5 配平时的阻 力惩罚是采用舵 0 配平时的一半。配平前后表面 压力、剖面压力、环量分布对比如图 11 所示, 相比







于基准构型, 舵 0 偏转配平时产生的配平阻力主要 是由于偏转使得载荷内移, 诱导阻力增加。相比于 舵 0 配 平, 采 用 舵 5 配 平 时, 虽 然 外 翼 段 展 向 91% 处截面显示激波强度略有增大,但由于展向载 荷更优, 诱导阻力的减少量更大, 最终导致整体上 阻力更小。需要注意的是, 虽然本文并未考虑变弯 度技术对横航向操纵特性的影响, 但舵 5 进行配平

 ΔC_M

(a) $C_l = 0.14$

实质上是副翼参与配平,当考虑横航向操纵特性后,采用舵5进行纵向配平,会增加舵面的使用压力。

 ΔC_M

(b) $C_l = 0.24$

訂增大,但由于展向载表 3 为 C_L = 0.24通过舵 0 和舵 2 进行配平的气[大,最终导致整体上动力结果。采用舵 0 配平时,舵偏角为-5.89°,阻力然本文并未考虑变弯增加 7.9×10⁻⁴,采用舵 2 配平时,舵偏角为-2.11°,阻影响,但舵 5 进行配平力增加 6.3×10⁻⁴。配平前后表面压力、剖面压力、表 2 C_L = 0.14时基准构型与单舵面配平构型气动力系数

Table 2 Aerodynamic coefficient of baseline configuration and single flap trim configuration ($C_L = 0.14$)

布局	舵 0/(°)	舵 1/(°)	舵 2/(°)	舵 3/(°)	舵 4/(°)	舵 5/(°)	α/(°)	C_D	C _M	$\Delta C_D / 10^{-4}$
基准构型	0	0	0	0	0	0	2.58	0.007 95	0.004 0	0
舵0 配平	4.91	0	0	0	0	0	2.41	0.008 17	0.000 3	2.2
舵5 配平	0	0	0	0	0	3.83	2.37	0.008 06	0.000 1	1.1



图 11 C_L = 0.14时基准构型与单舵面配平构型对比

Fig. 11 Comparison of baseline configuration and single flap trim configuration ($C_L = 0.14$)

环量分布对比如图 12 所示, 舵 0 配平时, 导致中央 体载荷内移,使得诱导阻力增加。采用舵2配平 时,如展向43%处截面显示,通过负偏增加弯度来 提供抬头力矩以供配平,但是导致了压差阻力增

加。与 C_L =0.14时一样,根据展向环量分布对比可 知, 与舵 0 配平相比, 舵 2 配平时展向环量分布更 接近椭圆形分布,主要是通过减小了诱导阻力来减 少配平所带来的阻力惩罚。

表 3	$C_L = 0.24$ 时基准构型与单舵面配平构型气动力系数
7X J	UL = 0.24的基准例至可半舵面配十例至飞幼刀示频

Table 3 Aerodynamic coefficient of baseline configuration and single flap trim configuration ($C_L = 0.24$)

布局	舵 0/(°)	舵 1/(°)	舵 2/(°)	舵 3/(°)	舵 4/(°)	舵 5/(°)	α/(°)	C_D	C_M	$\Delta C_{D} / 10^{-4}$
基准构型	0	0	0	0	0	0	3.63	0.011 48	-0.004 9	0
舵0 配平	-5.89	0	0	0	0	0	3.82	0.012 28	-0.0001	8
舵2 配平	0	0	-2.11	0	0	0	3.93	0.012 11	-0.000 1	6.3

1299





3 翼身融合布局飞行器变弯减阻优化 与分析

3.1 不考虑俯仰配平的变弯度技术减阻优化

本节定量研究变弯度技术对 BWB 布局飞行 器带来的气动减阻收益。当采用舵面组合偏转, 即 6 个舵可以同时自由偏转时,采用第 2 节所搭 建的优化设计系统,研究变弯度技术对 BWB 布 局飞行器带来的气动减阻收益优化问题可以表述为 $\begin{cases} \text{Min}: C_D \\ \text{s.t. } \Delta\delta \in [-4,4] \end{cases}$ (4)

式中: Δδ为舵面偏角。优化过程中首先采用拉丁超 立方试验设计方法生成 70个样本点,在给定的设 计状态下进行变弯度气动特性分析。在 70个样本 中随机选取 10个样本做测试集,代理模型绝对误 差如表 4 所示,其中升力系数误差与升力系数相比 可以忽略不计;阻力误差最大为 1.2×10⁻⁴,但平均误 差不到 0.7×10⁻⁴;代理模型的俯仰力矩平均误差为 0.000 3,满足要求。精度验证结果表明,训练所得
表 4 设计点 Kriging 代理模型精度验证结果

 Table 4
 Results of validation for Kriging model under

	design point	
气动性能	最大预测误差	平均预测误差
C_L	6.3×10 ⁻⁴	2.5×10^{-4}
C_D	1.2×10^{-4}	6.8×10^{-4}
C_M	7.6×10 ⁻⁴	2.9×10 ⁻⁴

的 Kriging 代理模型具有较高的预测精度,可以替 代高精度的 CFD 流场分析。

优化结果表示,设计状态CL=0.19时,基准构

型不变弯度时阻力最小,这是因为基准构型本身是 设计点处考虑俯仰力矩自配平约束的优化结果。针 对非设计点的变弯度优化结果如表5所示。小升 力系数下减阻 8×10⁻⁵(减阻 1.0%),此时俯仰力矩为 0.0072。大升力系数下,减阻 5.3×10⁻⁴(减阻 4.62%), 此时俯仰力矩为-0.0157。相比于小升力系数,大 升力系数下变弯度技术能带来更大的减阻收益,以 及更大的低头力矩。

图 13 和图 14 分别为变弯度前后机翼上激波 = 0.19时,基准构 强度对比、环量分布对比及剖面压力系数 C_p对 表 5 不考虑俯仰配平时变弯度减阻优化结果

Table 5 Results of variable camber drag reduction optimization without considering pitch trim



Fig. 13 Comparison of the baseline configuration and variable camber configuration without considering pitch trim ($C_L = 0.14$)



Fig. 14 Comparison of baseline configuration and variable camber configuration without considering pitch

trim ($C_L = 0.24$)

比。对于小升力系数的工况, 舵0、舵2、舵3负偏, 舵4 正偏。变弯度后外翼段后缘卸载、巡航迎角增 大, 引起载荷前移, 导致抬头力矩增加。同时外翼 段舵3、舵4处中后缘部分压力恢复平缓, 激波强 度略微减小。通过机翼展向环量分布对比可以看 出, 变形后载荷外移, 环量分布更理想, 诱导阻力 更小。对于大升力系数的工况, 舵0、舵1、舵3、舵4、舵5 正偏。变弯度技术使得载荷后移, 上翼面的压力 分布变得饱满, 压力恢复平缓, 激波强度降低, 波 阻减小。通过机翼展向环量分布对比仍可以看 出, 载荷外移, 环量分布更加靠近椭圆形环量分 布。变弯度后巡航迎角减小、后缘加载, 从而增 加了低头力矩。

3.2 考虑配平约束下变弯度技术减阻优化

考虑俯仰力矩配平约束的情况下,优化问题定

义如下:
Min:
$$C_D$$

s.t.
$$\begin{cases} C_L = 0.19 \\ \Delta \delta \in [-4, 4] \\ C_M = 0 \end{cases}$$
(5)

考虑配平时,设计点 C_L =0.19处仍是不偏舵时 阻力最小。针对不同升力系数进行变弯度减阻优 化,结果如表6所示。 C_L =0.14时,采用变弯度技 术实现配平与不变弯度相比阻力减小2×10⁻⁵,通过 变弯度技术基本实现了无力矩惩罚。在 C_L =0.24 时,采用变弯度技术实现配平与不变弯度相比阻力 增加5.1×10⁻⁴。通过表5和表6的对比可见,是否 考虑力矩配平对最佳偏角和减阻收益影响明显,在 进行 BWB 布局变弯度研究时必须考虑变弯度引起 全机俯仰力矩变化带来的配平阻力的影响。

	Та	ble 6 Resu	lts of variab	le camber dr	ag reductior	n optimizatio	on when co	onsidering pi	tch trim	
C_L	舵 0/(°)	舵 1/(°)	舵 2/(°)	舵 3/(°)	舵 4/(°)	舵 5/(°)	α/(°)	C_D	C_M	$\Delta C_D / 10^{-4}$
	0	0	0	0	0	0	2.58	0.007 95	0.004 0	0
0.14	4.91	0	0	0	0	0	2.41	0.008 17	0.000 3	2.2
	0.681	1.75	-0.501	-0.0914	0.422	1.65	2.45	0.007 92	0.000 4	-0.2
	0	0	0	0	0	0	3.63	0.011 48	-0.004 8	0
0.24	-5.89	0	0	0	0	0	3.82	0.012 28	-0.000 1	8
	-1.18	-2.47	0.496	-0.423	-0.528	-0.186	3.83	0.011 99	-0.000 3	5.1

表 6 考虑配平约束时变弯度减阻优化结果

从表 6 中还可以看出, 在非设计点处, 相比于 采用中央体后缘进行配平, 采用变弯度技术进行配 平所产生的配平惩罚更小。 $C_L = 0.14$ 时, 采用变弯 度技术实现配平可以减少 2.4×10^{-4} 的配平惩罚, 而 在 $C_L = 0.24$ 时, 可以减少 2.9×10^{-4} 的配平惩罚。图 15 和图 16 为 2 种方式配平时机翼上激波强度对比、 环量分布对比及剖面压力系数C_p对比。

采用2种方式配平时的迎角相差不大,*C*_L=0.14 时配平构型的迎角都小于基准构型不变弯度的情况,*C*_L=0.24时配平构型的迎角都大于基准构型不 变弯度的情况。从环量分布图中可以看出,变弯度 时,由于舵0处的偏转更小,导致小升力系数下,载



Fig. 15 Comparison of central body trailing edge trim configuration and variable camber trim configuration ($C_L = 0.14$)



Fig. 16 Comparison of central body trailing edge trim configuration and variable camber trim configuration ($C_L = 0.24$)

荷内移,大升力系数下,载荷外移,最终都使得展向 环量分布更贴近椭圆形环量分布,减小了诱导阻 力,从而实现更小的配平惩罚。

4 翼身融合布局飞行器变弯度设计 空间探索

为进一步探索变弯度时各舵面偏转对阻力和 俯仰力矩的影响,将阻力和俯仰力矩同时作为优化 目标,进行全局优化以获得 Pareto 解集,结果如 图 17 和图 18 所示,图中红色三角形为不变弯度的 基准构型的气动特性,红色正方形为采用变弯度技术配平时的气动特性。从图中可以看出,阻力和俯 仰力矩存在明显的支配关系。 $C_L = 0.14$ 时采用变 弯度技术的配平构型能产生不到 10⁻⁴的减阻, $C_L = 0.24$ 时采用变弯度技术的配平惩罚为 5×10⁻⁴ 左右,与第 3 节结果相同。 假设实际飞行过程中俯仰力矩 C_M在 [-0.01, 0.01] 之间变化。由图 17 可知, 小升力系数下, 阻力 变化在 3×10⁻⁴ 以内, 即在小升力系数下至多改变



图 17 $C_L = 0.14$ 多目标优化 pareto 前沿 Fig. 17 Multi-objective optimization pareto frontier ($C_L = 0.14$)





Fig. 18 Multi-objective optimization pareto frontier ($C_L = 0.24$)

3.8%的阻力。而大升力系数下,由图 18 可知,阻力 变化均在 10⁻³以上,即产生 8.7%以上的阻力变 化。可见变弯度技术在大升力系数下对阻力的影 响更强。

为了可视化揭示最优解集内,后缘变弯度舵偏 组合与阻力和俯仰力矩系数间的影响关系,引入数 学分析中的降维思想,对高维输入进行主成分分析 (principal component analysis,PCA),提取决定问题本 质的主要成分,将冗余目标剔除,在不失问题主特 征的前提下,将高维输入优化转化为低维输入优化 问题^[24-25]。

对 2 个非设计升力系数下得到的 pareto 前沿所 对应的各舵面偏角进行 PCA,结果如表 7 和表 8 所 示,其中百分数即为各主成分对应的贡献率。

 $C_L = 0.14$ 时,根据表7采用前3个主成分能保

Ta

留原始数据99%以上的特征。主成分1表示舵0、 舵1、舵3、舵5负偏的舵面组合,主成分2表示 舵 0、 舵 4 正偏, 舵 1 负偏的舵面组合, 主成分 3 表 示舵1负偏、舵2和舵5正偏的舵面组合。图19 为前3个主成分影响下气动阻力及俯仰力矩系数 变化情况,由图可知,当主成分1系数增加,主成分 2系数略微增加,主成分3系数减小时(即舵面组合 为舵0、舵1、舵2,舵3、舵5负偏,舵4正偏的情况 时)有利于减阳,但会导致抬头力矩的增加。参考 3.1 节 C_L = 0.14时变弯度减阻优化的结果,此时舵 面负偏有利于载荷外移,从而环量分布、减小诱导 阻力,并且舵3负偏,舵4正偏有利于外翼段中后 缘部分压力恢复平缓,使得激波强度减小。但舵面 负偏会导致后缘卸载、巡航迎角增大,引起载荷前 移,导致抬头力矩增加。若要进行配平,可以通过 减小主成分1系数、增加主成分2系数实现。

2024年

 Γ M III \mathcal{J} H III \mathcal{J} $C_L = 0.24$ H, R H K \mathbb{Z} 8 \mathcal{R} H f i 3 Λ 2 R \mathcal{L} R $\mathcal{C}_L = 0.24$ H, R H K \mathbb{Z} 8 \mathcal{R} H f i 3 Λ 2 R \mathcal{L} R \mathcal{L} \mathbb{Z} R \mathcal{R} \mathcal{S} \mathcal{S} \mathcal{R} \mathcal{R} \mathcal{R} R \mathcal{S} \mathcal{S} \mathcal{R} \mathcal{R} </tr

ble 7	Principal component	elements correspondence	$(C_I =$	0.14

主成分	舵 0/(°)	舵 1/(°)	舵 2/(°)	舵 3/(°)	舵 4/(°)	舵 5/(°)
主成分1(93.73%)	-0.426	-0.624	-0.172	-0.440	0.052	-0.451
主成分2(4.25%)	0.679	-0.524	-0.067	0.112	0.494	0.057
主成分3(1.05%)	-0.199	-0.399	0.556	-0.161	-0.115	0.673
主成分4(0.41%)	-0.090	-0.267	-0.729	0.276	-0.370	0.420
主成分5(0.32%)	-0.101	0.323	-0.353	-0.574	0.517	0.405

Table 8Principal component elements correspondence ($C_L = 0.14$)

	-			-		
主成分	舵 0/(°)	舵 1/(°)	舵 2/(°)	舵 3/(°)	舵 4/(°)	舵 5/(°)
主成分1(93.73%	<i>()</i> -0.426	-0.624	-0.172	-0.440	0.052	-0.451
主成分2(4.25%) 0.679	-0.524	-0.067	0.112	0.494	0.057
主成分3(1.05%) -0.199	-0.399	0.556	-0.161	-0.115	0.673
主成分4(0.41%) -0.090	-0.267	-0.729	0.276	-0.370	0.420
主成分5(0.32%) -0.101	0.323	-0.353	-0.574	0.517	0.405



图 19 CL = 0.14时 PCA 降维后不同组合舵偏下气动阻力及俯仰力矩系数变化云图

Fig. 19 Contour of aerodynamic drag and pitch moment coefficients under different combinations of flap deflection after PCA dimension reduction ($C_L = 0.14$)





Fig. 20 Contour of aerodynamic drag and pitch moment coefficients under different combinations of flap deflection after PCA dimension reduction ($C_L = 0.24$)

行配平,则可以通过减小主成分1系数,增加主成分3系数实现,即舵0、舵1、舵4负偏实现。

5 结 论

本文基于遗传算法和 kriging 代理模型构建了 全局优化框架, 研究了基于舵面偏转的后缘变弯度 技术对 BWB 布局飞行器气动性能的影响, 相关结 论如下:

1) 随着升力系数的变化, BWB 布局飞行器后

缘各舵面的配平能力也随之改变。虽然中央体后 缘舵面的力臂最长,但其配平能力并不一定是最优 的。小升力系数下,采用翼尖处的舵面进行配平阻 力损失最小;大升力系数下,则是采用后缘拐折处 的舵面进行配平阻力损失最小。

2)虽然 BWB 布局飞行器存在力臂短、配平损失大的固有问题,但采用变弯度技术仍能提高其在 巡航非设计点的气动性能,并带来更小的配平惩 罚。相比于小升力系数,变弯度技术在大升力系数

下会引起更剧烈的阻力变化。

3)不同巡航升力系数下,采用变弯度技术获得 减阻收益的机理和舵面偏转组合方式存在明显差 异。小升力系数下变弯度技术主要是通过舵0、舵1、 舵2, 舵3、舵5负偏, 舵4正偏的组合实现了诱导 阻力的减小并增加抬头力矩;大升力系数下则是采 用舵0、舵1, 舵4、舵5正偏的变弯度组合, 减少了 诱导阻力和激波阻力, 但会引起低头力矩的增加。

参考文献(References)

- [1] DEL ROSARIO R, FOLLEN G, WAHLS R, et al. Subsonic fixed wing project overview of technical challenges for energy efficient, environmentally compatible subsonic transport aircraft[C]//Proceedings of the 50th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition. Reston: AIAA, 2012: 9-12.
- [2] LYU Z, MARTINS J R R A. Aerodynamic design optimization studies of a blended-wing-body aircraft[J]. Journal of Aircraft, 2014, 51(5): 1604-1617.
- [3] REIST T A, ZINGG D W. Aerodynamic shape optimization of a blended-wing-body regional transport for a short range mission [C]//Proceedings of the 31st AIAA Applied Aerodynamics Conference. Reston: AIAA, 2013: 2414.
- [4] BROWN M, VOS R. Conceptual design and evaluation of blendedwing body aircraft[C]//Proceedings of the AIAA Aerospace Sciences Meeting. Reston: AIAA , 2018: 0522.
- [5] 王刚,张彬乾,张明辉,等. 翼身融合民机总体气动技术研究进展 与展望[J]. 航空学报, 2019, 40(9): 7-35.
 WANG G, ZHANG B Q, ZHANG M H, et al. Research progress and prospect for conceptual and aerodynamic technology of blended-wing-body civil aircraft[J]. Acta Aeronautica et Astronautica Sinica, 2019, 40(9): 7-35 (in Chinese).
- [6] OKONKWO P, SMITH H. Review of evolving trends in blended wing body aircraft design[J]. Progress in Aerospace Sciences, 2016, 82(1): 1-23.
- [7] MARQUES M, GAMBOA P, ANDRADE E. Design of a variable camber flap for minimum drag and improved energy efficienc y[C]//Proceedings of the 50th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference. Reston: AIAA, 2009: 2196.
- [8] MARTINS J R R A. Fuel burn reduction through wing morphing[J]. Green Aviation, 2016, 9(8): 73.
- [9] KAUL U K, NGUYEN N T. Drag optimization study of variable camber continuous trailing edge flap (VCCTEF) using OVERFLOW [C]//Proceedings of the 32nd AIAA applied aerodynamics conference. Reston: AIAA, 2014: 2444.
- [10] URNES J, NGUYEN N. A mission adaptive variable camber flap control system to optimize high lift and cruise lift to drag ratios of future n+3 transport aircraft[C]//Proceedings of the 51th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition. Reston: AIAA, 2013: 214.

- [11] TING E, DAO T, NGUYEN N T. Aerodynamic load analysis of a variable camber continuous trailing edge flap system on a flexible wing aircraft[C]//Proceedings of the 56th AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference. Reston: AIAA, 2015: 1839.
- [12] IPPOLITO C A, NGUYEN N T, TOTAH J, et al. Initial assessment of a variable-camber continuous trailing-edge flap system for dragreduction of non-flexible aircraft in steady-state cruise condition [C]//Proceedings of the AIAA Infotech@Aerospace Conference 2013. Reston: AIAA, 2013: 5143.
- [13] NGUYEN N T, XIONG J, SAGER J. Fuel-optimal trajectory optimization of mach 0.745 transonic truss-braced wing with variable camber continuous trailing edge flap[C]//Proceedings of the AIAA Aviation 2021 Forum. Reston: AIAA, 2021: 2575.
- [14] XIONG J, BARTELS R E, NGUYEN N T. Aeroelastic trim drag optimization of mach 0.745 transonic truss-braced wing aircraft with variable-camber continuous trailing-edge flap[C]//Proceedings of the AIAA Aviation 2021 Forum. Reston: AIAA, 2021: 2528.
- [15] NORRIS G. Boeing unveils plans for trailing edge variable camber on 787 to reduce drag, save weight[EB/OL]. (2006-06-12)[2006-06-12]. http://Tlightglobal.com/news/articles.2006.
- [16] THOMAS A R, KOO D, ZINGG D W. Aircraft cruise drag reduction through variable camber using existing control surfaces[J]. Journal of Aircraft, 2022: 1-10.
- [17] 何萌,杨体浩,白俊强,等.基于后缘襟翼偏转的大型客机变弯度 技术减阻收益[J]. 航空学报, 2020, 41(7): 165-180.
 HE M, YANG T H, BAI J Q, et al. Drag reduction benefits of variable camber technology of airliner based on trailing-edge flap deflection[J]. Acta Aeronautica et Astronautica Sinica, 2020, 41(7): 165-180(in Chinese).
- [18] MITRIDIS D, PAPANIKOLATOS N, PANAGIOTOU P, et al. Morphing technologies assessment on a tactical BWB UAV reference platform[C]//Proceedings of the AIAA Scitech 2022 Forum. Reston: AIAA, 2022: 1985.
- [19] DEB K, AGRAWAL S, PRATAP A, et al. A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA- II [C]//Proceedings of the International conference on parallel problem solving from nature PPSN VI: 6th International Conference. Berlin : Springer, 2000: 849-858.
- [20] 李莉. 基于遗传算法的多目标寻优策略的应用研究[D]. 无锡: 江 南大学, 2008.

LI L. Application research on multi-objectives optimization based on the genetic algorithm[D]. Wuxi: Jiangnan University , 2008.

- [21] LIOU M F, KIM H, LEE B, et al. Aerodynamic design of integrated propulsion-airframe configuration of a hybrid wing body aircraft[J]. Shock Waves, 2019, 29(8): 1043-1064.
- [22] FELDER J, KIM H, BROWN G, et al. An examination of the effect of boundary layer ingestion on turboelectric distributed propulsion systems[C]//Proceedings of the 49th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition. Reston: AIAA, 2011: 300.
- [23] CAI Y, XIE J, HARRISON E, et al. Assessment of longitudinal stability-and-control characteristics of hybrid wing body aircraft in

1307

conceptual design[C]//Proceedings of the AIAA AVIATION 2021 FORUM. Reston: AIAA, 2021: 2448.

[24] ABDI H, WILLIAMS L J. Principal component analysis[J]. Wiley Interdisciplinary Reviews: Computational Statistics, 2010, 2(4): 433-459.

[25] CARREIRA-PERPINAN M A. Continuous latent variable models for dimensionality reduction and sequential data reconstruction[D]. Sheffield : University of Sheffield, 2001.

Design and aerodynamic analysis of blended wing body with variable camber technology

WANG Yutong^{1, 2}, LAN Qingsheng¹, ZHOU Zhu^{1, *}, YANG Tihao², SONG Chao¹

Computational Aerodynamics Institute, China Aerodynamics Research and Development Center, Mianyang 621000, China;
 School of Aeronautics, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: The variable camber wing has great potential to improve the aerodynamic characteristics of conventional tube and wing design, but it can change the pitching moment of the whole aircraft. The drag reduction advantage and trim penalty of variable camber technology on a blended wing-body aircraft are explored, taking into account the features of short moment arm and substantial trim loss in the blended wing body. Considering the engineering practice, the variable camber of the trailing edge is realized based on flap deflection, and the trimming ability of the flap at different spanwise positions is compared and analyzed. Then, the global optimization method is used to carry out the optimization design of variable camber aerodynamic drag reduction; finally, the variable camber design space is explored. The results show that as the lift coefficient changes, the position of the flap that produces the least trim resistance also changes. Without the pitching moment trim constraint, at most 4.62% of the drag reduction benefit can be obtained by using the variable camber technology. When compared to trimming with the center body's trailing edge, variable camber technology can result in a trim loss of 2.4×10^{-4} when it comes to the pitching moment trim constraint. Under different lift coefficients, there are obvious differences in the deflection combination of flaps with variable camber. When the lift coefficient is small, the variable camber can decrease drag by combining the deflection of multiple flap negative deflections, which will increase the head-up moment. When the lift coefficient is large, however, the variable camber can decrease drag by combining the deflection of multiple flap positive deflections. The evaluation of variable camber drag reduction benefit and moment penalty based on the combined deflection of flaps can provide a reference for the engineering design of variable camber blended wing body aircraft.

Keywords: blended wing body; variable camber wing; aerodynamic performance; optimal design; surrogate model

URL: link.cnki.net/urlid/11.2625.V.20220922.1003.001

Received: 2022-06-16; Accepted: 2022-09-19; Published Online: 2022-09-23 07: 57

^{*} Corresponding author. E-mail: f_yforever@126.com

April 2024 Vol. 50 No. 4

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

DOI: 10.13700/j.bh.1001-5965.2022.0382

基于图卷积网络的表格隶属关系抽取

张宇童1,李启元1,刘树衎1,2,*

(1. 中国人民解放军海军工程大学电子工程学院,武汉 430033; 2. 东南大学 计算机科学与工程学院,南京 211102)

摘 要: 针对表格识别与分析领域中表内单元格间隶属关系抽取问题,定义表格隶属关系 抽取任务,结合表格与图结构的相似性,给出表内单元格的图表示方法,并提出一种基于图卷积网络(GCN)的表格隶属关系抽取模型。所提模型通过GCN对表内单元格及其邻近格进行特征的聚 合,预测单元格间是否存在隶属关系,实现关系抽取。为验证所提模型的有效性,标注中文表单 Rel-forms及英文表格 Rel-SciTSR 这 2 个数据集。通过实验,在上述 2 类数据集及联合数据集上 F₁分数分别达到 98.61%、96.55%、97.05%,验证所提模型在此 2 个数据集上的有效性,并分别分 析文本内容、坐标信息、单元格属性及格间相对方向等不同因素对隶属关系抽取实验结果的影响。

关键 词:表格分析;隶属关系;图表示;关系抽取;图卷积网络

中图分类号: TP183

文献标志码: A 文章编号: 1001-5965(2024)04-1308-08

表格识别与分析是指在数字文档(如 PDF 或图 像)中进行表格位置检测与结构分析等任务。由于 在数字文档格式类型中,表格的半结构化特性难以 被利用,为充分获取表格中的数据信息,实现表格 的自动化处理,需要对表格进行识别和结构分析研 究,其中,表格内单元格间隶属关系抽取是指在已 有表格数据的基础上,对表格内单元格间的蕴含关 联进行探索,预测单元格间存在的隶属关系,并依 据关系抽取数据,将表格转化为结构化数据的任 务。常见的表格可分为2类:①科学论文文档中大 量出现的表格;②信息登统记的表单。科技论文表 格包含了论文中的大量实验结果, 高效的杳询和理 解表中内容是十分必要的;常见表单通常涉及个人 信息的填写,因此,通过自动化实现表内单元格间 隶属关系抽取任务,可以对表单待填位置区域自动 识别、为表格数据自动化获取等任务提供支撑。由 于表格数据内容丰富、版式多样,实现表内单元格 间隶属关系抽取是一项富有挑战且亟需解决的

问题。

当前表格研究领域注重于表格的结构分析与 重建任务,未能够进一步探究表内各单元格间蕴含 的隶属关系。针对这一表格分析需求,本文主要工 作可概括如下。

1)提出一个基于图卷积网络 (graph convolutional network, GCN)^[1]的表格隶属关系抽取模型。 根据表格的特性,先将其转换为图结构,借助 GCN聚合每个单元格及其相邻格的特征。同时加 入单元格类别及方向特征,进一步提升各单元格间 隶属关系的预测能力。

2) 针对缺乏表格类数据单元格关系标注的问题,本文提供了2个隶属关系抽取的数据集,包括论文表格数据集 Rel-SciTSR 和中文表单数据集 Rel-forms,借助于上述数据集,进行了表单元格隶 属关系抽取实验,并取得了优异的结果,这也将为 后续的表格理解、表格智能填写等任务打下坚实的 基础。

收稿日期: 2022-05-18; 录用日期: 2022-08-12; 网络出版时间: 2022-09-14 12:56 网络出版地址: link.cnki.net/urlid/11.2625.V.20220915.1654.005

基金项目:湖北省自然科学基金(2018CFC800)

*通信作者. E-mail: liusk@seu.edu.cn

引用格式:张字童,李启元,刘树衎.基于图卷积网络的表格隶属关系抽取 [J].北京航空航天大学学报,2024,50(4):1308-1315. ZHANG Y T, LI Q Y, LIU S K. Tabular subordination relation extraction based on graph convolutional networks [J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50(4):1308-1315 (in Chinese). 3)考虑到本文模型的通用性,不仅在单一数据 集上进行实验,同时将2个数据集合并进行联合训 练,提升本文模型的泛化能力。探究多种特征对于 关系抽取效果的影响,首先,在纯文本模态进行实 验,其次,在单元格位置特征的基础上,添加属性特 征及相对方向特征,并对实验结果进行了对比与分 析,验证了本文模型的有效性。

1 相关工作

表格类数据蕴含丰富的内容和复杂的版式,是 电子文档的重要组成部分。针对表格识别与分析 领域中的相关任务,目前的研究可依据图结构的使 用划分为2类研究方法。

1.1 基于非图结构的方法

基于非图结构的方法主要针对于表格的行列 分布进行研究,忽视表格的图结构特性。文献 [2-4] 利用卷积神经网络进行行、列和单元格的分类识别, 实现表格的结构重建。但文献 [2-4] 未能实现端到 端操作,需借助后处理步骤。Raja等⁵³通过自上而 下进行单元格检测,将金字塔网络 (feature pyramid networks, FPN)⁶获取的特征与单元格特征拼接,通过 长短期记忆 (long short-term memory, LSTM) 网络^[7] 增强行列特征,自下而上构建表格结构,但其选用 Resnet101 作为骨干网络,参数量较大。Kong 等^[8] 根据梯度热图获得表格垂直线和水平线的像素,分 别找出行、列分割线子图,补足分割线长度并将两 图结合,通过连通域搜索算法确定单元格的位置, 但对于无框线表格难以处理。Qiao 等¹⁹ 较早提出 局部和全局金字塔掩码结构,将2种掩码依据重赋 分策略进行融合,通过单元格匹配-空单元格搜索-空单元格合并3步完成表格的重构过程。这是少 数能重构出空白单元格的结构分析方法,但后处理 过程较复杂。Long 等^[10] 提出了一种基于 CenterNet^[11] 的表格识别方法,通过学习相邻单元格的公共顶点 信息来重构表格结构,解决了自然场景图像中表格 的解析问题,但其仅适用于有线表格。虽然这些方 法在表格的结构分析任务上取得了不错的效果,但 处理方式更偏向于找寻分割线以达到重构目的,而 表格类数据由数量庞大的单元格构成,其本质是一 种图结构,运用图表示的方法会更容易理解表格。

1.2 基于图结构的方法

基于图结构的方法注重于利用表单元格间的 相关性,结合图结构处理方式进行研究。Chi等^[12] 以单元格尺寸和位置为输入特征,通过图注意力机 制预测各节点的邻接关系,重建表格结构。Xue等^[13] 定义单元格空间坐标预测分支与逻辑坐标预测分

支,建立一个端到端的表格结构分析模型,通过 Resnet50网络进行图像特征抽取后,以空间坐标预 测辅助逻辑坐标预测,但这种方式会使得分支依赖 性增强。Riba等^[14]选用图神经网络 (graph neural network, GNN)^[15] 模型进行发票的布局检测, 在节点 卷积中加入图边信息,增强了特征嵌入计算中的结 构知识。但其仅使用结构信息,忽略表中文字内 容,且只针对发票类数据。Qasim 等^[16] 设计一种既 能利用卷积神经网络进行图像特征提取,又能利用 GNN 进行顶点之间有效交互的结构,同时引入合 成数据集来填补大规模数据集的空白。但顶点特 征仅使用坐标信息,缺乏对单元格间信息的利用。 Li等^[17]将单元格位置、文本特征进行拼接,利用 GCN 进行聚合,再添加图像特征来判断单元格间的 水平垂直关系,从而实现表格结构的分析,同时还 提出了一个新的金融类中文表格数据集。但由于 使用最近邻算法构建邻接矩阵,这将会导致正负样 本数量不均衡。基于图结构的方法在表格的识别 与分析领域取得了优异的成绩,不仅利用了图像特 征,同时还将单元格的位置特征与文本特征加入其 中,探究表格的结构问题,但上述方法都未提及表 单元格间蕴含地隶属性质的探究,忽视了单元格间 在位置关联性之外还存在着内容关联性。

2 表格隶属关系抽取定义

表格类数据中存在2种类型的单元格,分别为 表标题格和表内容格。在标题格与内容格之间存 在着隶属关系,即内容格是隶属于对应的标题格, 其格中文本内容也将由其隶属的标题格所决定。 表格隶属关系的抽取结果如图1所示。

图 1(a)为论文表格数据的隶属关系抽取转换 过程,在论文表格数据中,将表格的第1行与第1 列视为该表格的标题格,其余单元格视为表格的内 容格,在单元格与内容格之间存在着隶属关系,即 内容格中的文本内容由标题格所决定。通过隶属 关系抽取,可以将内容格与标题格对应起来,抽取 相应文本内容以元组形式进行存储,便于管理和 查询。

图 1(b)为中文表单数据的隶属关系抽取转换 过程,其为待填写表单数据,即表单中的内容格均 为空白,而存在文本内容的是表单标题格,中文表 单中内容格与标题格间呈现相邻的位置关系,通常 为左右或上下结构,将标题格与对应的空白内容格 进行隶属关系抽取,以键值对的形式将表中数据进 行存储,其中空白格均以"待填"进行文本填充,实 现由表格图像到结构化数据的转换。 待埴



(b) 中文表单数据 图 1 2 类表格数据的隶属关系定义及抽取过程



本文主要针对论文表格及中文待填表单数据 进行了隶属关系抽取研究。依据单元格间的隶属 关系进行数据抽取.按表格数据类型以2类存储形 式(元组或键值对)进行数据管理,将其转化为结构 化数据,更好地管理和查询表格内容。

3 图卷积网络

由于图结构数据十分不规则,每个节点周围结 构都可能是独一无二的,这就使得传统的卷积神经 网络、循环神经网络都难以对其处理,因此,GCN 被提出。作为一种特征提取器,处理的对象是图结 构数据。通过利用 GCN 模块对图数据的处理能力 进行特征提取[18],可以完成节点分类、图分类、边 预测及图的嵌入表示等任务。GCN 的传播方式如 图2所示。

由图2可知,向GCN中输入一张图,经过若干 层 GCN 后, 节点聚合了相邻节点特征后由 X 变成 了 Z。Y为对立节点的标签,但其节点的连接关系 依旧不变,即GCN中的节点邻接矩阵是共享的,这 就为处理不规则的图结构数据提供了方法。



Fig. 2 Dissemination mode of GCN

3.1 表单元格的图表示方法

将原始表格类数据转换为表图,借助 GCN 为 特征聚合工具,应用于隶属关系抽取研究中。借助 GCN 进行图表示的具体工作流程如图 3 所示。



以图 3 为例,由原始表格数据到实现隶属关系 抽取及结构化数据转换可分为以下几个步骤。

步骤1 将原始表格类数据中所有单元格顺 序编号,构成 GCN 的输入顶点集 S = {s₁, s₂,…, s_N},其中,N为图中节点数量。连接所有标题-内容 格、随机连接内容-内容格和标题-标题格构建邻接 矩阵A。

根据邻接矩阵A绘制出相应的表图G。 步骤2 将单元格间的隶属关系作为边值在表图中进行标 注,以"1*"表示两单元格间存在隶属关系,"0*"表 示不存在。

步骤3 由于在实际传播中,邻接矩阵的构建 过程缺失自环信息,因此,在原始邻接矩阵中加入 节点本身的信息,即:

$$\hat{A} = A + I$$

w $(x_1, y_1) \cap_{(x_3, y_3)}$

表格

姓名

标题 籍贯

标题

第4期

(1)

式中:I为自环信息: \hat{A} 为加入自环后的邻接矩阵。 在图卷积层上进行操作,实现特征聚 步骤4 合,其具体过程可表示为

$$\boldsymbol{H}^{l} = f(\boldsymbol{H}^{l-1}, \hat{\boldsymbol{A}}) \tag{2}$$

式中: H^{l-1} 为上一卷积层的输出: $H^{(0)} = X$, 为节点的 输入特征。

每层的节点特征表示为其相邻节点特征和本 身特征的加权和,即

 $f(\boldsymbol{H}^{l}, \hat{\boldsymbol{A}}) = \sigma(\hat{\boldsymbol{D}}^{-\frac{1}{2}} \hat{\boldsymbol{A}} \hat{\boldsymbol{D}}^{-\frac{1}{2}} \boldsymbol{H}^{l} \boldsymbol{W}^{l})$ (3)

式中: $\hat{\boldsymbol{D}}^{-\frac{1}{2}} \hat{\boldsymbol{A}} \hat{\boldsymbol{D}}^{-\frac{1}{2}}$ 表示对 $\hat{\boldsymbol{A}}$ 进行归一化, 避免出现梯 度消失或梯度爆炸问题; σ 为 Sigmoid 激活函数; W^{l} 为可训练的权重矩阵。

步骤5 预测节点间的隶属关系,并依据隶属 关系进行内容抽取,实现结构化数据的转换。

t3, 13, h, h

标题:1

位置

 $\Theta_{(x_2, y_2)}$

内容

总体流程为:根据表格中单元格来定义表图 中相应节点的特征,以单元格属性(标题格或内容 格)为依据构建邻接矩阵。依据邻接矩阵建立一 个完整的表图,对矩阵中连接的节点间进行隶属 关系的判别,对于存在隶属关系的一对节点赋予 边权为1*,否则为0*(区别于邻接矩阵中的1、0, 后续实验中以1、0作为权值)。通过GCN将节点 与其邻近节点的特征相聚合, 增强节点特征的相 关性,进行格间隶属关系的判别,实现隶属关系抽 取。最终将抽取结果依据数据类别进行结构化 转换。

3.2 表格单元隶属关系抽取流程

针对引言所指的2类数据集,本文使用2类数 据集联合训练的方式增强模型的泛化能力,提出基 于 GCN 的表格隶属关系抽取模型, 如图 4 所示, w和h分别为图的宽度和高度。





图 4 本文模型 Fig. 4 The proposed model

先定义单元格的3种特征:位置特征、类别特 征和相对方向特征。其中节点i的位置特征定义为

 $X_{\text{pos}}^{i} = [x_{\min}, y_{\min}, x_{\max}, y_{\max}, x_{\text{center}}, y_{\text{center}}, c_{\text{weight}}, c_{\text{height}}]$ (4)

式中:各参数分别为第i个单元格左上坐标、右下 坐标、中心点坐标及单元格宽度和高度。

单元格类别特征以 one-hot 向量进行表示, 具 体定义为

 $X_{\text{class}}^i = [c_{\text{header}}, c_{\text{content}}]$ (5)

式中: c_{header} 表示该单元格为标题格; c_{content} 表示该单

元格为内容格。

方向特征以 one-hot 向量的形式进行表示,具 体定义为

$$\boldsymbol{X}_{rd}^{i} = [\boldsymbol{L}, \boldsymbol{R}, \boldsymbol{U}, \boldsymbol{D}] \tag{6}$$

式中:L、R、U、D分别为上、下、左、右四个方向。 边特征定义为

$$\boldsymbol{E}_i = \begin{bmatrix} \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{r} \end{bmatrix} \tag{7}$$

式中:x、y为节点索引;r为隶属关系标签。

按照边进行节点的特征扩充,即将所有相互连 接的节点在同类特征上进行节点间的特征拼接操 作,例如对节点*i*的位置特征扩充可以表示为 $X_{pos}^{i} = X_{pos}^{i} || X_{pos}^{j}$ (8) 式中: $X_{pos}^{i} \pi X_{pos}^{j}$ 分别为节点*i*和节点*j*的位置特征; "||"表示拼接操作。

将3类特征拼接,获得节点特征矩阵为

 $\boldsymbol{X} = \boldsymbol{X}_{\text{pos}} \| \boldsymbol{X}_{\text{class}} \| \boldsymbol{X}_{\text{rd}}$ (9)

最后,借助于 GCN 进行特征聚合操作,并以 log_Softmax 对节点关系进行分类,判别是否存在隶 属关系。

本文使用 log_Softmax+NLLLoss 函数作为模型 的损失函数来等价替换交叉熵损失函数。交叉熵 损失的具体计算式为

$$L = -\sum_{i=0}^{K-1} y_i \log_2 \hat{y}_i$$
 (10)

式中: yi为真实标签; ŷi为预测值; K为类别总数。

通过使用 log_Softmax+NLLLoss 函数,单独进行 log 运算,避免了交叉熵损失中 Softmax 函数可能导致数据上溢问题的出现。

本文模型将原始表格类数据中的单元格特性 (包括位置、属性、方向)作为表图节点特征,借助 于 GCN 进行单元格的特征聚合,对聚合后的节点 进行隶属关系的预测,找出所有存在隶属关系的单 元格,建立关联特性并抽取其内容进行结构 化存储。

4 实验结果与分析

4.1 数据集

针对现有数据集缺乏单元格隶属关系标注的问题,在公开数据集 SciTSR^[12]上随机抽取部分数据并进行关系标注,同时从互联网上收集了部分中 文表单数据,进行关系标注,构成了2类表格的关 系抽取数据集。

4.1.1 Rel-SciTSR

SciTSR^[12] 是一个大型的表格结构识别数据集, 包含 15 000 张 PDF 格式的表格。数据集共有 15 000 张图片,按照 8:2 的比例被划分为训练集和测试 集,同时也提供了只包含复杂表格的测试集 SciTSR-COMP。针对于 SciTSR^[12]数据集,从中随机抽取 1 000 张图像,将标注的节点对按照 8:2 划分为训练 集和测试集,其中训练集 65 000 条,测试集 16 000 条。 标注过程为:在标题格与标题格、标题格与内容 格、内容格与内容格间建立关系标签,规定表内容 格隶属于其对应的行列标题格,将其关系标签置为 1,其他连接关系标签置为 0。图 5 展示了部分实验 样例。

Name	YFCC100M (car)
Size (pixels)	502×480
Number of frames	640
Direct (D)	0.99 frames/s
Fast (A)	96.96 frames/s
Gain (A vs. D)	98

(a) 样例1

data set	tra	ining
	time	throughput
	(s)	(ent/s)
lighttpd	90.4	36 918.1
memcached	10.1	85 409.5
xlock	237.1	1 124.3
named	15 742.6	585.4

	Processing time (s)	
	HP	MP
Song 1	58	131
Song 2	171	345

(c) 样例3

	Offline ₁	Offline ₂
Camera source	1	2
Generated user	22	110
Positive pairs	11	55
Negative pairs	2,20	3,960
Sum of reposted	110	550

(d) 样例4

图 5 Rel-SciTSR 数据集样例

Fig. 5 Sample images from dataset of Rel-SciTSR

由于该数据集中表格图像源自科技论文,用于 存储各类实验结果,因此,表格版式多样,框线填充 规范不统一。

4.1.2 Rel-forms

借助于互联网的庞大资源,收集了 759 张中文 表单,标记节点对总数约为 38 000 条,其中训练集 31 000 条,测试集 7 000 条。标注单元格位置同时 与光学字符识别结果进行对齐,获得每张表单的单 元格坐标及相应文本内容。按照中文表单特性,具 有隶属关系的单元格通常是相邻的位置关系。因 此,随机选择与每个单元格具有邻接关系的单元格 进行关系标注。每个单元格与其隶属单元格的关 系标签置为 1,否则置为 0。图 6 为部分实验样 例。由图 6 可知,该数据集中包含常见的中文待填 表单,主要用于个人信息的登计,此类表格结构清 断且框线完整,含有标题格与空白格 2 类单元格。

4.2 实验设置

4.2.1 2类数据集联合训练

首先,在提出的2个数据集 Rel-SciTSR 和 Relforms上对表格类数据的单元格文本进行了探究, 在纯文本特征的基础上,证明文本特征对关系抽取



(b) 样例2

图 6 Rel-forms 数据集样例

Fig. 6 Sample images from dataset of Rel-forms

效果的影响。其次,分别探究本文模型基于单元格 位置、属性和方向关系的特征上,在2类数据上的 抽取效果。最后,考虑本文模型的泛化性,将2类 数据直接合并在3类特征上进行联合训练,探究本 文模型在2种数据类型上的普适性。

4.2.2 参数设置

本文中选择 Adam 作为优化器, 模型的 Batch



Size设置为16,并以准确率为指标对学习率L_r和训练轮数进行探究。

(1) 对学习率进行探究。由于联合数据集 Rel-SciTSR+Rel-forms 包含前 2 个数据集的所有样本,因此,选择数据集 Rel-SciTSR+Rel-forms 作为实验数据集。设置初始 L_r=0.0002,并以 0.0002 为增量进行递加,获得准确率变化趋势图如图 7 所示。最终选定 L_r=0.005 作为本文模型的学习率。



(2) 对本文模型训练训练轮数进行探究。在确 定学习率 L_r=0.005 后, 对本文模型训练轮数进行探 究, 在提出的 2 个数据集及联合数据集上进行实 验, 获得准确率及损失变化如图 8 所示。可以看出 在训练轮数接近 100 时, 本文模型的损失趋于稳 定, 本文模型达到收敛, 且准确率较高。



图 8 正确率及损失变化趋势 Fig. 8 Change trend of accuracy and loss

4.3 结果展示与分析

在表格中,文本内容是其重要组成部分,因此, 探究文本特征对隶属关系抽取的影响是十分必要 的。借助 bag of words 模型,对原始文本向量化后, 进行 embedding 操作,再放入门控循环单元(gated recurrent unit, GRU)模型中。最后,以每个单元格的文 本语义增强向量为输入,在 GCN 中进行特征聚合, 预测各单元格间的隶属关联性,具体实验结果如 表1所示。

由表1可知,由于论文表格中文本类别较少, 重复性较高,在仅考虑文本特征的基础上,难以取 得较好效果。而中文待填表单数据由于标题格文

表1 文本特征关系抽取结果

Table 1 Result of relation extraction in text feature

数据集	正确率/%
Rel-SciTSR	86.46
Rel-forms	91.76
Rel-SciTSR+ Rel-forms	88.93

本内容种类多,且其隶属格为待填写空白单元格,因此,中文表单数据的实验结果要高于论文表格。 在本文模型和提出的2个数据集上探究单元格 3类特征对隶属关系抽取的影响,以准确率P、召回 率 R 和 F₁分数 F₁为评估指标比较模型表现。具体 实验结果如表2所示,其中,Po表示位置特征, Cl表示单元格类别特征,Rd表示相对方向特征。 由表 2 可知, 无论是论文表格还是表单数据, 在单元格位置特征的基础上均可以取得良好的关 系抽取效果。在加入单元格类别特征及相对方向 特征后, 可以进一步对实验结果进行提升。这一提 升不仅体现在准确率及召回率, *F*₁分数作为综合评 价准确率与召回率的指标, 在上述 3 个数据集上分 别达到了 98.61%, 96.55% 和 97.05%, 验证了本文 模型的有效性。

通过对比在2种数据形式上的结果可以证明, 本文模型具有较好的泛化能力,能够适应所提供的 论文表格和中文表单数据,在关系抽取实验中表现 出良好的性能,这也将为下一步表格理解、表格智 能填写等下游任务打下良好的基础。

表 2 隶属关系抽取结果 Table 2 Result of subordination relation extraction

计行		Р			R			F_1	
特征	数据集①	数据集②	数据集③	数据集①	数据集②	数据集③	数据集①	数据集②	数据集③
Ро	97.51	93.59	95.05	95.19	93.96	95.56	96.34	93.77	95.30
Po+Cl	98.16	95.88	96.93	97.90	95.65	96.67	98.03	95.76	96.80
Po+Cl+Rd	98.82	96.49	97.18	98.40	96.61	96.93	98.61	96.55	97.05

注:①论文表格数据集Rel-SciTSR;②表单类数据集Rel-forms;③联合数据集Rel-SciTSR+ Rel-forms。

5 结 论

1) 针对目前表格识别与分析领域缺少对表内 单元格间蕴含隶属关系研究的问题,提出一种基于 GCN 的隶属关系抽取模型,并在提出的2类数据集 上均取得了优异的成绩。

2) 探究了本文模型的通用性, 合并 2 类数据并 在本文模型上进行联合训练, 实验结果证明本文模 型具有良好的泛化能力。

由于目前数据集规模较小且现有公共数据集 缺乏在此方面的标注,难以进行大规模训练。未来 将进一步完善数据类型,测试本文模型在其他表格 类数据集上的效果,进一步提升本文模型的泛化能力。

参考文献(References)

- [1] WANG H Y, CHENG Y H, PHILTP CHEN C L, et al. Semisupervised classification of hyperspectral image based on graph convolutional broad network[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2021, 14: 2995-3005.
- [2] ZOU Y J, MA J W. A deep semantic segmentation model for image-based table structure recognition[C]//Proceeclings of the 2020 15th IEEE International Conference on Signal Processing. Piscataway: IEEE Press, 2020: 274-280.
- [3] SCHREIBER S, AGNE S, WOLF I, et al. DeepDeSRT: Deep learn-

ing for detection and structure recognition of tables in document images[C]//Proceeclings of the 2017 14th IAPR International Conference on Document Analysis and Recognition. Piscataway: IEEE Press, 2017: 1162-1167.

- [4] SIDDIQUI S A, KHAN P I, DENGEL A, et al. Rethinking semantic segmentation for table structure recognition in documents[C]// Proceeclings of the 2019 International Conference on Document Analysis and Recognition. Piscataway: IEEE Press, 2019: 1397-1402.
- [5] RAJA S, MONDAL A, JAWAHAR C V. Table structure recognition using top-down and bottom-up cues[C]//Proceeclings of the European Conference on Computer Vision.Berlin: Springer, 2020: 70-86.
- [6] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 2117-2125.
- [7] GUO X L, ZHU S J, YANG Z W, et al. Consecutive missing data recovery method based on long-short term memory network[C]//Proceeclings of the 2021 3rd Asia Energy and Electrical Engineering Symposium. Piscataway: IEEE Press, 2021: 988-992.
- [8] KONG L J, BAO Y C, WANG Q W, et al. A gradient heatmap based table structure recognition[C]//Proceeclings of the 2021 13th International Conference on Machine Learning and Computing. New York: ACM, 2021: 456-463.
- [9] QIAO LS, LI ZS, CHENG Z, et al. LGPMA: Complicated table structure recognition with local and global pyramid mask align-

ment[C]//Proceeclings of the International Conference on Document Analysis and Recognition. Berlin: Springer, 2021: 99-114.

- [10] LONG R J, WANG W, XUE N, et al. Parsing table structures in the wild[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2021: 944-952.
- [11] ZHOU X Y, WANG D Q, KRHENBÜHL P. Objects as points[EB/OL]. (2019-04-25) [2022-03-24]. https://arxiv.org/ abs/1904.07850.
- [12] CHI Z W, HUANG H Y, XU H D, et al. Complicated table structurerecognition[EB/OL].(2019-08-28)[2022-03-10].https://arxiv.org/ abs/1908.04729.
- [13] XUE W Y, YU B S, WANG W, et al. TGRNet: A table graph reconstruction network for table structure recognition[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2021: 1295-1304.
- [14] RIBA P, DUTTA A, GOLDMANN L, et al. Table detection in invoice documents by graph neural networks[C]//Proceeclings of the

2019 International Conference on Document Analysis and Recognition . Piscataway: IEEE Press, 2019: 122-127.

- [15] SCARSELLI F, GORI M, TSOI A C, et al. Computational capabilities of graph neural networks[J]. IEEE Transactions on Neural Networks, 2009, 20(1): 81-102.
- [16] QASIM S R, MAHMOOD H, SHAFAIT F. Rethinking table recognition using graph neural networks[C]//Proceeclings of the 2019 International Conference on Document Analysis and Recognition. Piscataway: IEEE Press, 2019: 142-147.
- [17] LI Y R, HUANG Z, YAN J C, et al. GFTE: Graph-based financial table extraction[C]//Proceeclings of the International Conference on Pattern Recognition. Berlin: Springer, 2021: 644-658.
- [18] 郑海潇, 文斌. 基于图卷积网络的比特币非法交易识别方法[J].
 信息网络安全, 2021, 21(9): 74-79.
 ZHENG H X, WEN B. Bitcoin illegal transaction identification method based on graph convolutional network[J]. Netinfo Security, 2021, 21(9): 74-79(in Chinese).

Tabular subordination relation extraction based on graph convolutional networks

ZHANG Yutong¹, LI Qiyuan¹, LIU Shukan^{1, 2, *}

School of Electronic Engineering, Naval University of Engineering, PLA, Wuhan 430033, China;
 School of Computer Science and Engineering, Southeast University, Nanjing 211102, China)

Abstract: This study addresses the lack of research on extracting subordination relations between the cells in a table in the field of table recognition and analysis. It defines the table subordination relation extraction task. Based on the similarity between table and graph structures and the graph representation method for table cells, a subordination relation extraction model is proposed using graph convolutional networks(GCN). The proposed model predicts the subordination relation between cells by aggregating the features of table cells and their adjacent cells through the GCN, and then realizes relation extraction. To verify the effectiveness of the proposed model, two datasets, Chinese Rel-forms and English Rel-SciTSR, were annotated. The experiments achieve 98.61%, 96.55% and 97.05% F_1 scores on the above two datasets and their joint datasets respectively, thus verifying the effectiveness of the proposed model. The effects of different factors such as text content, coordinate information, cell attributes and relative orientation between cells of subordination relation extraction are also analyzed.

Keywords: tabular analysis; subordination relation; graph representation; relation extraction; graph convolutional network

Received: 2022-05-18; Accepted: 2022-08-12; Published Online: 2022-09-14 12:56 URL: link.cnki.net/urlid/11.2625.V.20220915.1654.005

Foundation item: Natural Science Foundation of Hubei Province (2018CFC800)

^{*} Corresponding author. E-mail: liusk@seu.edu.cn

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

DOI: 10.13700/j.bh.1001-5965.2022.0389

基于伪距残差和新息的 GNSS/IMU 抗差 自适应定位算法

刘正午,孙蕊*,蒋磊

(南京航空航天大学 民航学院,南京 211106)

摘 要:在全球导航卫星系统 (GNSS) 和惯性测量元件 (IMU) 组合导航系统中,抗差滤波 和自适应滤波常被用于提高组合导航的定位精度。但是抗差滤波和自适应滤波所适用的条件不同, 使用不当反而可能会降低组合导航的定位精度,针对此问题,提出基于伪距残差和新息的 GNSS/IMU 抗差自适应定位算法。所提算法基于伪距残差评估 GNSS 的定位质量,选择合适的滤波 算法进行 GNSS/IMU 组合导航解算。在长时间 GNSS 定位质量较差时,基于新息和伪距残差判断是 否 IMU 运动学推算误差大于 GNSS 观测值误差,从而根据判断的结果选择是否采用抗差因子。结 果表明:所提算法相对于扩展卡尔曼滤波算法在东、北和天方向上分别提高 36.05%、22.71% 和 56.22% 的定位精度。

关键词:伪距残差;新息;抗差滤波;自适应滤波;GNSS/IMU组合导航
 中图分类号:V324;P228
 文献标志码:A 文章编号:1001-5965(2024)04-1316-09

由于全球导航卫星系统 (global navigation satellite system, GNSS)和惯性测量元件 (inertial measurement units, IMU)具有天然的互补特性, GNSS/IMU组合导航被广泛用于提高城市环境下 车辆导航定位精度^[1-3]。但在组合导航中,一方面, GNSS 卫星信号常常会受到建筑物和树木的遮挡及 反射,无法为组合导航提供可靠、高质量的 GNSS 观测数据^[4-6];另一方面,低成本 IMU输出的数据也 可能存在粗差^[7]。2种情况都会严重降低 GNSS/ IMU组合导航的定位性能^[8-9]。对此,有学者提出抗 差滤波^[10-11]和自适应滤波^[12-13]来应对上述问题。 但是,抗差滤波和自适应滤波是 2种相反的策略, 其有各自的优点。自适应滤波是通过观测信息抑 制 IMU运动学推算误差,而抗差滤波是通过 IMU 运动学模型抑制 GNSS 观测值误差。而自适应滤 波和抗差滤波并不能区分误差的来源,若在使用自适应滤波时观测信息存在误差或使用抗差滤波时 IMU运动学模型存在误差,则组合导航的定位精度 可能不会得到改进,甚至可能会变得更差^[14-15]。因 此,Yang等^[16]提出抗差自适应滤波,从而能有效抵 抗 IMU运动学推算误差和 GNSS 观测值误差。但 在抗差自适应滤波中,构建自适应因子需要有多余 的观测值或可靠的观测信息^[17],且抗差自适应滤波 并不能区分误差的来源,从而可能会导致一些低精 度导航解^[18-19]。所以必须合理地选择滤波类型来分 配 GNSS 观测值和 IMU 运动学推算值在组合导航 定位中的贡献,从而抑制 IMU 运动学推算误差和 GNSS 观测值误差对 GNSS/IMU 组合导航定位结果 的影响。因此,高为广和陈谷仓^[20]提出了一种结合 自适应滤波和神经网络的 GNSS/INS 抗差组合导航

收稿日期: 2022-05-19; 录用日期: 2022-07-17; 网络出版时间: 2022-08-02 15:32 网络出版地址: link.cnki.net/urlid/11.2625.V.20220801.1844.003

*通信作者. E-mail: rui.sun@nuaa.edu.cn

基金项目:国家自然科学基金(42174025,41974033);工信部民用飞机专项科研项目(MJ-2020-S-03);江苏省自然科学基金(BK20211569);江苏省"六大人才高峰"项目(KTHY-014)

引用格式: 刘正午, 孙蕊, 蒋磊. 基于伪距残差和新息的 GNSS/IMU 抗差自适应定位算法 [J]. 北京航空航天大学学报, 2024, 50 (4): 1316-1324. LIU Z W, SUN R, JIANG L. Robust adaptive position algorithm for GNSS/IMU based on pseudorange residual and innovation [J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50 (4): 1316-1324 (in Chinese).

算法。谭兴龙等^[18]提出了一种支持向量机辅助选 择滤波类型的抗差自适应算法。Ning等^[19]提出了 一种最优径向基函数神经网络辅助选择滤波类型 的抗差自适应算法。上述算法能有效区分误差类 型,选择合适的滤波类型提高 GNSS/IMU 组合导航 的定位精度。但是需要使用大量的数据进行训练, 且算法也会影响导航定位的实时性。

针对上述问题,本文提出了基于伪距残差和新 息的 GNSS/IMU 抗差自适应定位算法。在长时间 GNSS 定位质量差时,基于新息和伪距残差判断 东、北、天方向上 GNSS 观测值定位误差与 IMU 运 动学推算定位误差的大小关系,从而决策在东、 北、天方向上是否继续使用抗差因子,使得组合导 航系统能得到 GNSS 观测值的及时修正。

1 GNSS/IMU 组合导航滤波模型

本文采用 GNSS/IMU 松组合导航模型,构建 输入卡尔曼滤波器的 21 维系统误差状态如式 (1) 所示:

$$\boldsymbol{X} = [(\delta \boldsymbol{r}_{\text{IMU}})^{\text{T}} (\delta \boldsymbol{\nu}_{\text{IMU}})^{\text{T}} (\delta \boldsymbol{\varPhi}_{\text{IMU}})^{\text{T}} (\boldsymbol{b}_{g})^{\text{T}} (\boldsymbol{b}_{a})^{\text{T}} (\boldsymbol{s}_{g})^{\text{T}} (\boldsymbol{s}_{a})^{\text{T}}]^{\text{T}}$$
(1)

式中: $\delta \mathbf{r}_{IMU}$ 、 $\delta \mathbf{v}_{IMU}$ 和 $\delta \boldsymbol{\Phi}_{IMU}$ 分别为 IMU 的三维位置 误差向量、三维速度误差向量和三维姿态误差向 量; \mathbf{b}_{g} 、 \mathbf{b}_{a} 、 s_{g} 和 s_{a} 分别表示陀螺仪三轴零偏、加速 度计三轴零偏、陀螺仪三轴比例因子和加速度计三 轴比例因子误差向量。抗差自适应卡尔曼滤波如 式 (2)~式 (7) 所示:

$$\widehat{X}_{k,k-1} = \boldsymbol{\Phi}_{k,k-1} \widehat{X}_{k-1} \tag{2}$$

$$\boldsymbol{\Sigma}_{k,k-1} = \boldsymbol{\Phi}_{k,k-1} \boldsymbol{\Sigma}_{k-1} \boldsymbol{\Phi}_{k,k-1}^{\mathrm{T}} + \boldsymbol{Q}_{k-1}$$
(3)

$$\boldsymbol{V}_{k} = \boldsymbol{Z}_{k} - \boldsymbol{H}_{k} \widehat{\boldsymbol{X}}_{k,k-1} \tag{4}$$

$$\boldsymbol{K}_{k} = \frac{1}{\alpha_{k}} \boldsymbol{\Sigma}_{k,k-1} \boldsymbol{H}_{k}^{\mathrm{T}} \left[\boldsymbol{H}_{k} \left(\frac{1}{\alpha_{k}} \boldsymbol{\Sigma}_{k,k-1} \right) \boldsymbol{H}_{k}^{\mathrm{T}} + \frac{\boldsymbol{R}_{k}}{\boldsymbol{\gamma}_{k}} \right]^{-1} \quad (5)$$

$$\widehat{\boldsymbol{X}}_{k} = \widehat{\boldsymbol{X}}_{k-1} + \boldsymbol{K}_{k} \boldsymbol{V}_{k} \tag{6}$$

$$\boldsymbol{\Sigma}_{k} = [\boldsymbol{I} - \boldsymbol{K}_{k} \boldsymbol{H}_{k}] \boldsymbol{\Sigma}_{k,k-1} [\boldsymbol{I} - \boldsymbol{K}_{k} \boldsymbol{H}_{k}]^{\mathrm{T}} + \boldsymbol{K}_{k} \frac{\boldsymbol{R}_{k}}{\boldsymbol{\gamma}_{k}} \boldsymbol{K}_{k}^{\mathrm{T}}$$
(7)

式中: $\hat{X}_{k,k-1}$ 为一步预测状态向量; $\Phi_{k,k-1}$ 为状态转移 矩阵; H_k 为观测矩阵; $\Sigma_{k,k-1}$ 为一步预测状态协方差 矩阵; Z_k 为观测向量,是 GNSS输出的位置与 IMU 推算的位置之差; V_k 为新息向量; K_k 为增益矩 阵; \hat{X}_k 为估计状态向量; Q_{k-1} 为动力学模型误差的 非负定方差矩阵; R_k 为观测噪声的对称正定方差矩 阵; Σ_k 为估计状态协方差矩阵; γ_k 和 α_k 分别为抗差 因子矩阵和自适应因子。

抗差因子计算式为

$$\gamma_k^x = \begin{cases} 1 & \left| \widetilde{V}_k^x \right| \le c \\ \frac{c}{\left| \widetilde{V}_k^x \right|} & \left| \widetilde{V}_k^x \right| > c \end{cases}$$
(8)

式中: γ_k^x 为 γ_k 的第 x 个对角线元素; c为常数, 通常 取 1~1.5; \tilde{V}_k^x 的计算式^[21] 为

$$\widetilde{V}_{k}^{x} = \frac{V_{k}^{x}}{\sqrt{\Sigma_{V_{k}}^{xx}}} \tag{9}$$

式中: V_k^x 为 V_k 的第x个元素; Σ_{V_k} 为新息向量的协方 差矩阵; $\Sigma_{V_k}^x$ 为 Σ_{V_k} 的第x个对角线元素。

误差判别统计量 $\Delta \overline{V}_k$ 和对应的自适应因子 α_k 计算式分别为

$$\Delta \overline{V}_{k} = \left(\frac{V_{k}^{\mathrm{T}} V_{k}}{\mathrm{tr}(\boldsymbol{\varSigma}_{k})}\right)^{\frac{1}{2}}$$
(10)

$$\alpha_{k} = \begin{cases} 1 & \Delta \overline{V}_{k} \leq c_{1} \\ \frac{c_{1}}{\Delta \overline{V}_{k}} & \Delta \overline{V}_{k} > c_{1} \end{cases}$$
(11)

式中: tr(·)表示求矩阵的迹; c_1 为常数,通常取 $1 \sim 1.5$ 。

2 本文算法框架

本文算法框架如图1所示,具体步骤如下。

1)判断 GNSS 的观测数据是否有冗余且空间 位置精度因子 (position dilution of precision, PDOP) 小于等于阈值 *T*_p(文中取为 30, 无单位)。若否, 则 直接选择抗差滤波进行解算, 因为数据无冗余则无 法计算伪距残差值, 而 PDOP 值异常很大时, 通常 是卫星的几何构型很差, 此时的 GNSS 定位结果很 不可靠。*G*_k = 1表示采用抗差因子, *A*_k = 0表示不采



Fig. 1 The proposed algorithm framework

用自适应因子;若是,则进入基于伪距残差的 GNSS 定位先验质量评估方法中评估 GNSS 的定位 质量。

2) 若评估为 GNSS 定位质量好,则 G_k=0, A_k=1,选择自适应滤波进行解算,当误差检验统计 量大于阈值时,采用自适应因子抑制 IMU运动学 推算误差;若评估为 GNSS 定位质量差,则进入基 于新息和伪距残差的抗差因子使用策略判断是否 在x方向(东、北和天方向)上使用抗差因子。

 3)通过选择的滤波类型和基于新息和伪距残 差的抗差因子使用策略的结果进行解算,得到最终 定位结果。

2.1 基于伪距残差的 GNSS 定位先验质量评估

想要合理地选择滤波的类型就需要评估 GNSS 的定位质量。传统上常用 PDOP 评估 GNSS 定位质 量,但是在城市环境中 PDOP 并不能有效地反映出 GNSS 的定位精度^[2-23]。而伪距残差常被用于判断 卫星信号是否受到遮挡的特征之一^[6, 24-25],其计算 式为

$$\boldsymbol{\eta} = \boldsymbol{\rho} - \boldsymbol{A} \cdot \boldsymbol{r} \tag{12}$$

式中: $\eta = [\eta_1, \eta_2, ..., \eta_n]^T$, η_i 为第 *i* 颗卫星的伪距残 差, *n*为卫星的数目; $\rho = [\rho_1, \rho_2, ..., \rho_n]^T$, ρ_i 为第 *i* 颗 卫星的伪距测量值与接受机初始位置反演的伪距 之差; *A*为 GNSS 观测矩阵; *r* = [δx , δy , δz , δt_1 , δt_2]^T为 最小二乘估计解矩阵;由于数据采用了全球定位系 统(global positioning system, GPS)和北斗卫星导航 系统(Beidou navigation satellite system, BDS),所以 *r*包含 GNSS 位置误差修正量(δx , δy , δz)、GPS 钟差 (δt_1) 和 BDS 钟差(δt_2)。

伪距残差绝对值越大的卫星受到建筑物反射 或遮挡的概率越大,也就是这颗卫星的伪距误差可 能越大。而受到建筑物反射或遮挡的卫星越多,则 GNSS 定位的精度会越差。所以可采用 *k* 时刻参与 GNSS 定位卫星伪距残差的绝对值的平均值(后续 称平均绝对伪距残差值)评估 *k* 时刻 GNSS 定位质 量的好坏,*k* 时刻平均绝对伪距残差值 *T_k*的计算 式为

$$T_k = \frac{\sum_{i=1}^n |\eta_i|}{n} \tag{13}$$

为验证平均绝对伪距残差值评估 GNSS 定位 误差的有效性,对文中2组实验(详见第3节)的数 据进行了分析。由于 GPS/BDS 三维定位误差与 PDOP 和卫星平均绝对伪距残差值具有不同的单 位,所以先将各组数据进行归一化处理,计算式为

$$y = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{14}$$

式中: *x*_{min}、*x*_{max}分别为该组数据中的最小值与最大 值。绘制了归一化之后的 GPS/BDS 三维定位误差 与 PDOP 和平均绝对伪距残差值 *T*_k的关系, 如图 2 和图 3 所示。





Fig. 2 Relationship between GPS/BDS 3D positioning error and PDOP as well as average absolute pseudorange residual in experiment 1





Fig. 3 Relationship between GPS/BDS 3D positioning error and PDOP as well as average absolute pseudorange residual in experiment 2

为进一步描述三者之间的关系,本文计算了 PDOP 和平均绝对伪距残差*T*_k与 GPS/BDS 三维定 位误差的斯皮尔曼等级相关系数(如表 1 和表 2 所 示),斯皮尔曼等级相关系数是用于评价 2 个变量 之间相关性的一种统计学方法,相关系数的绝对值 越大表示 2 个变量之间的相关关系越强,可参考 表 3 来描述斯皮尔曼等级相关系数的绝对值所对 应的相关程度^[25]。斯皮尔曼等级相关系数ρ计算式^[26] 如下式 (15) 所示:

$$\rho = 1 - \frac{6\sum_{i=1}^{N} d_i^2}{N(N^2 - 1)}$$
(15)

式中: N为元素个数; d_i为2组数据排序后的秩次 差值。

表1 实验1中斯皮尔曼等级相关系数

Table 1 Spearman rank correlation coefficients in

experiment 1

PDOP与GPS/BDS	平均绝对伪距残差与
三维定位误差	GPS/BDS三维定位误差
-0.03	0.58

表 2 实验 2 中斯皮尔曼等级相关系数

Table 2	Spearman rank	correlation	coefficients	in
---------	---------------	-------------	--------------	----

experiment 2

PDOP与GPS/BDS	平均绝对伪距残差与GPS/BDS
三维定位误差	三维定位误差
-0.24	0.91

表 3 斯皮尔曼等级相关系数相关程度[25]

 Table 3
 Correlation degree of Spearman rank

 correlation coefficient^[25]

ho	相关程度
0.00~0.19	很弱
0.20~0.39	弱
0.40~0.59	中等
0.60~0.79	强
0.80~1.0	很强

由图 2 和图 3 可知, 在城市环境中, 平均绝对 伪距残差值比 PDOP 更能真实地反映出 GNSS 定位 误差的变化。在大部分情况下, 随着 GNSS 定位误 差增大, 平均绝对伪距残差值也在增大。同时, 由 表 1 和表 2 可知, 在 2 组实验数据中, 平均绝对伪 距残差值与 GPS/BDS 三维定位误差的斯皮尔曼等 级相关系数为 0.58 和 0.91, 远大于 PDOP 与 GPS/BDS 三维定位误差的斯皮尔曼等级相关系数-0.03 和 -0.24。由表 3 可知, 平均绝对伪距残差值与 GPS/ BDS 定位误差为中等相关关系, 甚至有很强相关关 系。因此, 可采用平均绝对伪距残差值评估 GNSS 的定位质量。

通过收集开阔环境下 GNSS 定位质量较好时的数据,利用这段时间的平均绝对伪距残差值来设立一个阈值 T_y 。若 $T_k \leq T_y$,则认为 GNSS 定位质量可靠,可以使用自适应因子抑制 IMU运动学推算误差的影响;否则,认为 GNSS 定位质量差,则使用

抗差因子抑制 GNSS 观测误差的影响。阈值Ty的 计算式为

$$T_y = 3 \times \frac{\sum_{k=1}^{m} |T_k|}{m}$$
 (16)

式中: *T_k为 k* 时刻的平均绝对伪距残差值; *m* 为开 阔环境下收集的 GNSS 数据总时长。

2.2 基于新息和伪距残差的抗差因子使用策略

当 GNSS 定位质量长时间较差时,不能简单判 定就采用抗差滤波。因为在长时间 GNSS 定位质 量差时即使采用抗差滤波也仍可能会导致 GNSS/ IMU 组合导航的定位误差累积,从而可能存在 IMU 运动学推算定位误差会大于 GNSS 定位误差 的情况。若此时继续采用抗差因子,则会影响定位 质量相对较好的 GNSS 观测数据对 IMU运动学推 算位置的修正作用,从而产生比扩展卡尔曼滤波更 差的定位结果。

针对上述问题,本文提出了基于新息和伪距残 差的抗差因子使用策略来判断在长时间 GNSS 定 位质量差时是否继续采用抗差因子。因为在各方 向上 IMU 运动学推算定位误差大于 GNSS 定位误 差的时间不同,所以需要分不同的方向进行判断。 即判断是否 $B_k^x = 1$, B_k^x 默认为 0, $B_k^x = 1$ 的含义是判 定在 k 时刻 x 方向(东、北和天方向)上 GNSS 的定 位误差小于 IMU 推算的定位误差,故 x 方向不能继 续采用抗差因子,即让 x 方向上抗差因子 γ_k^x 为 1,计 算式为

$$\gamma_{k}^{x} = \begin{cases} 1 & \left| \tilde{V}_{k}^{x} \right| \leq c \vec{\mathfrak{g}} B_{k}^{x} = 1 \\ \frac{c}{\left| \tilde{V}_{k}^{x} \right|} & \left| \tilde{V}_{k}^{x} \right| > c \boxplus B_{k}^{x} = 0 \end{cases}$$
(17)

基于新息和伪距残差的抗差因子使用策略流 程如图 4 所示,具体步骤如下。

 判断是否 GNSS 定位质量在近 t(单位: s, 文 中取为5s)一直很差,导致近 t>(单位: s,下文同)都 是采用抗差滤波解算。若否,则直接结束,默认 B^x_k = 0;若是,则进一步判断是否 B^x_{k-1} = 1。

2) 若 $B_{k-1}^{x} \neq 1$ 且 $V_{k}^{x}V_{k-i}^{x} < 0$ 和 $T_{k} < \overline{T_{k}}$,则 $B_{k}^{x} = 1$; 否则,结束判断,默认 $B_{k}^{x} = 0$ 。 V_{k}^{x} 为 k 时刻 x 方向的 新息,即 GNSS 观测值计算的 x 方向的位置与 IMU 运动学推算的 x 方向的位置之差。 $V_{k}^{x}V_{k-i}^{x} < 0$ 表示 k 时刻 x 方向的新息与 k-i 时刻新息的正负符号都 相异, $i=1,2,\cdots,t$ 。所以可推测 k 时刻与 k-i 时刻在 x 方向上 GNSS 定位误差和 IMU 推算定位误差的 大小关系发生了变化,但不能判定在 x 方向上 GNSS 的定位误差是从比 IMU 推算的定位误差小变为 大,还是从大变为小。所以需要进一步判断, $\overline{T_{k}}$ 表



图 4 基于新息和伪距残差的抗差因子使用策略

Fig. 4 Application strategy of robust factor based on innovation and pseudorange residual

示 k-t 时刻至 k-1 时刻平均绝对伪距残差的平均 值,其计算式为

$$\overline{T_k} = \frac{\sum_{i=k-t}^{k-1} T_i}{t} \tag{18}$$

 $T_k < \overline{T}_k$ 表示 k 时刻平均绝对伪距残差值小于前 t 内的平均绝对伪距残差值的平均值。因为通常来说平均绝对伪距残差值 T_k 越小, GNSS 定位质量越好。所以判定 k 时刻 GNSS 定位质量优于前 t 窗口内 GNSS 定位质量,又因为 $V_k^x V_{k-i}^x < 0$,所以判断 k 时刻在 x 方向上 GNSS 的定位误差是从比 IMU 推算的定位误差大变为小,因此, x 方向上不能继续采用抗差因子降低 GNSS 观测值对最终定位结果的贡献, 即 $B_k^x = 1$ 。

3) 若 $B_{k-1}^x = 1 \pm V_k^x V_{k-1}^x > 0$, 则 $B_k^x = 1$; 否则, 结束 判断, 默认 $B_k^x = 0$ 。因为 $B_{k-1}^x = 1$ 表明 k-1 时刻 x 方 向判定 GNSS 的定位误差小于 IMU 推算的定位误 差。 $\pm V_k^x V_{k-1}^x > 0$ 表明 k 时刻 x 方向新息正负符号 与 k-1 时刻相同, 即判定 k 时刻 x 方向 GNSS 的定 位误差仍小于 IMU 推算的定位误差, 所以 $B_k^x = 1$ 。

3 实验结果及分析

为验证本文算法的有效性,进行了 2 组实验。 在实验 1 中,大约在 463 100~463 200 s 的 GPS 时之 间,约 100 s 左右, GNSS 定位质量持续较差(见图 2); 而在实验 2 采集数据的过程中,车辆在约 380 285 s 的 GPS 时停顿了约 300 s 的时间,从而使得 GNSS 定位质量持续较差的时间更长(见图 3)。实验设备 由惯导 STIM-300 和 GNSS 接收机 BDStar Navigation C520-AT 组成,其中惯导采样率为 125 Hz, GNSS 接 收机采样率为 1 Hz。阈值 *T*y计算为 1.05 m。参考 轨迹是由高精度惯导 HGuide N580 的数据经 Inertial Explorer 后处理得到。采用4种方法进行松组合导 航解算:

1) 扩展卡尔曼滤波算法;

2) 抗差扩展卡尔曼滤波算法;

3) 采用基于伪距残差的 GNSS 质量评估方法 判断采用的滤波类型(简称算法 1);

4) 本文算法。

3.1 实验1

实验1数据采集时间为2021年11月5日,采 集时长为1324s。参考行驶轨迹如图5所示。由 于地图更新存在延迟,不能很好地反映实验场景的 真实情况,故展示了行驶轨迹部分场景如图6所 示,可见数据采集地点为城市复杂环境,GNSS卫星 信号极易受到双边高楼和高架桥的遮挡。卫星数 目如图7所示,只有14s不满足数据冗余的条件。 GPS/BDS 三维定位误差与 PDOP 和平均绝对伪距 残差值的关系如图2所示,斯皮尔曼等级相关系数 如表1所示,实验结果如图8与图9和表4~表6所示。



图 5 实验 1 中参考行驶轨迹 Fig. 5 Reference driving track in experiment 1





(a) 场景1

(b) 场景2

图 6 行驶轨迹部分场景 Fig. 6 Partial scene of driving track

在图 9 中, 纵轴表示在该 GPS 时的 GPS/BDS 的三维定位误差。可见本文算法能在 GNSS 定位 质量好时选择自适应滤波进行解算, 而在 GNSS 定 位质量差的时选择抗差滤波进行解算。由表 4 可 知, 在实验 1 中, 选择自适应滤波进行组合导航解 算的时长为 911 s, 期间 GPS/BDS 三维定位误差平 均值为 4.22 m, 中位数为 3.88 m, 都远低于选用抗差



图 7 实验 1 中卫星数目





图 8 实验 1 中 4 种算法的定位误差



Fig. 8 Position errors of four algorithms in experiment 1

图 9 实验 1 中本文算法滤波类型选择结果与 GPS/BDS 三维 定位误差

Fig. 9 The proposed algorithm filter type selection results and GPS/BDS 3D positioning errors in experiment 1

滤波时的定位误差。

由图 8、表 5 和表 6 可知, 抗差扩展卡尔曼滤波 算法在东和北方向比扩展卡尔曼滤波算法定位精 度更差。这是因为在长时间 GNSS 观测误差出现 的前期, 抗差扩展卡尔曼滤波算法能抑制 GNSS 观 表 4 实验 1 中本文算法滤波类型选择结果

 Table 4
 Selection results of the proposed algorithm filter

type in Experiment 1

滤波类型	采用该滤波的 总时长/s	GPS/BDS三维定位 误差平均值/m	GPS/BDS三维定位 误差中位数/m
自适应滤波	911	4.22	3.88
抗差滤波	399	23.26	11.36

表 5 实验 1 中各算法均方根误差

Table 5 Root mean square error of each

algorithm in experiment 1

算法	东	北	天	2D	3D
扩展卡尔曼滤波算法	7.99	6.12	9.64	10.06	13.94
抗差扩展卡尔曼滤波算法	9.40	6.80	5.27	11.60	12.74
算法1	5.13	4.80	4.19	7.03	8.18
本文算法	5.11	4.73	4.22	6.96	8.14

表 6 实验 1 中各算法相对于扩展卡尔曼滤波提升率

 Table 6
 Improvement rate of each algorithm

relative to extended Kalman filter in

1	Experim	ent 1			%
算法	东	北	天	2D	3D
抗差扩展卡尔曼滤波算法 卡尔曼滤波算法	-17.65	-11.11	45.33	-15.27	8.57
算法1	35.79	21.57	56.54	30.20	41.30
本文算法	36.05	22.71	56.22	30.82	41.58

测误差的影响,但是随着长时间的累积,IMU运动 学推算值会受到 GNSS 观测误差的污染,且 IMU 推 算值误差本身就会随着时间进行累积。又当 GNSS 观测误差减小时,就容易导致 IMU运动学推算误 差超过 GNSS 观测值的误差,且此时新息仍然会很 大,若继续采用抗差因子反而会降低 GNSS/IMU 组 合导航的定位精度。由于车辆的运动特性,车辆位 置在东和北方向变化较大,而在天方向变化较小, 所以抗差扩展卡尔曼滤波在天方向精度总体有所 改善。

实验1证明,本文算法具有最高的定位精度, 方法2相对于扩展卡尔曼滤波算法在东、北、天、2D 和3D上分别提高了36.05%、22.71%、56.22%、30.82% 和41.58%的定位精度。而本文算法只比算法1的 精度略高,这是因为在选择合适的滤波类型之后, 当GNSS长时间定位质量差时,很少发生IMU推算 的定位误差比GNSS定位误差更大的情况,从而较 少采用到基于新息和伪距残差的抗差因子使用策 略来辅助判断东北天方向上抗差因子的使用。为 了验证本文算法相对于算法1的优势,进行了第 2组实验,实验2中GNSS定位质量较差的持续时 间更长,从而更容易发生在GNSS定位质量较差的 时候,IMU运动学推算的定位误差比GNSS定位误 差更大的情况。

m

3.2 实验2

实验 2 数据采集时间为 2021 年 11 月 11 日,采 集时长为 1 131 s,参考行驶轨迹如图 10 所示。卫 星数目如图 11 所示,只有 1 s 不满足数据冗余的条 件,GPS/BDS 三维定位误差与 PDOP 和平均绝对伪 距残差值的关系如图 3 所示,斯皮尔曼等级相关系 数如表 2 所示,实验结果如图 12、图 13 和表 7~表 9 所示。



图 10 实验 2 中参考行驶轨迹 Fig. 10 Reference driving track in experiment 2



Fig. 11 Number of satellites in experiment 2

由图 13 和表 7 可知,本文算法能合理地选择滤 波类型进行 GNSS/IMU 组合导航功能定位解算。 由图 12 和表 8、表 9 可知,本文算法仍具有最高的 精度。本文算法相对于扩展卡尔曼滤波算法在东、 北、天、2D 和 3D 上分别提高了 9.88%、2.71%、 36.09%、7.54% 和 27.95% 的定位精度。本文算法相 对于算法 1 在天方向上均方根误差从 11.14 m 减少 至 8.73 m(见表 8)。这是因为在 380 500 历元左右 时,由于长时间 GNSS 定位质量差,导致 IMU 运动 学推算误差受到污染,此时虽然 GNSS 定位质量较 差,但 IMU 推算的定位误差比 GNSS 定位质量较 差,但 IMU 推算的定位误差比 GNSS 定位误差更 大。而此时算法 1 会继续使用抗差因子降低 GNSS 观测值对最终定位结果的贡献,从而导致算法 1 比 扩展卡尔曼滤波算法的定位误差更大。而本文算



图 12 实验 2 中 4 种算法的定位误差





图 13 实验 2 中本文算法滤波类型选择结果与 GPS/BDS 三 维定位误差

Fig. 13 Selection results of the proposed algorithm filter type and GPS/BDS 3D positioning errors in experiment 2

表 7 实验 2 中本文算法滤波类型选择结果

 Table 7
 The proposed algorithm filter type selection results in experiment 2

2年2年米王1	采用该滤波的	GPS/BDS三维定位	GPS/BDS三维
<i>能似</i> 失型	总时长/s	误差平均值/m	定位误差中位数/m
自适应滤波	537	2.97	2.44
抗差滤波	593	17.82	11.36

表 8 实验 2 中各算法均方根误差 Table 8 Root mean square error of each

algorithm in experiment 2

m

算法	东	北	天	2D	3D
扩展卡尔曼滤波算法	6.48	4.42	13.66	7.84	15.75
抗差扩展卡尔曼滤波算法	7.72	4.82	12.60	9.10	15.54
算法1	6.10	4.20	11.14	7.40	13.37
本文算法	5.84	4.30	8.73	7.25	11.35

法采用基于新息和伪距残差的抗差因子使用策略 能判断 IMU 推算的定位误差与 GNSS 定位误差的 大小关系,从而及时地停止使用 x 方向的抗差因 子,即 γ_{k}^{x} = 1,从而让组合导航系统能及时得到定位

表 9 实验 2 中各算法相对于扩展卡尔曼滤波提升率 Table 9 Improvement rate of each algorithm relative to

extended Kalman filter in experiment 2					70
算法	东	北	天	2D	3D
抗差扩展卡尔曼滤波算法	-19.14	-9.05	7.76	-16.03	1.32
算法1	5.86	4.98	18.45	5.58	15.08
本文算法	9.88	2.71	36.09	7.54	27.95

质量相对更好的 GNSS 观测值的修正。通过实验 2,再一次验证了本文算法的最优性能,且进一步验 证了本文算法相对于算法1的优势。

4 结 论

1) 在城市环境中, 伪距残差比 PDOP 能更好地 反映出 GNSS 的真实定位误差, 伪距残差与 GNSS 定位误差具有中等甚至很强的正相关关系。

2)利用基于伪距残差的 GNSS 定位先验质量 评估算法能有效选择合适的滤波模型进行 GNSS/ IMU 组合导航定位解算,从而合理地分配 GNSS 观 测值和 IMU 运动学推算值对定位结果的贡献。

3) 基于新息和伪距残差的抗差因子使用策略 能在 GNSS 观测质量较差且 IMU 运动学推算受到 污染时,自主判别 x 方向上 GNSS 观测误差和 IMU 运动学推算误差的大小关系,从而能避免错误地使 用抗差因子,达到提高定位精度的效果。

参考文献(References)

- [1] YANG L, LI Y, WU Y L, et al. An enhanced MEMS-INS/GNSS integrated system with fault detection and exclusion capability for land vehicle navigation in urban areas[J]. GPS Solutions, 2014, 18(4): 593-603.
- [2] LI B X, CHEN G W, SI Y B, et al. GNSS/INS integration based on machine learning lightGBM model for vehicle navigation[J]. Applied Sciences, 2022, 12(11): 5565.
- [3] YUE S,CONG L, QIN H L, et al. A robust fusion methodology for MEMS-based land vehicle navigation in GNSS-Challenged environments[J]. IEEE Access, 2020, 8: 44087-44099.
- [4] LYU Z T, GAO Y. An SVM based weight scheme for improving kinematic GNSS positioning accuracy with low-cost GNSS receiver in urban environments[J]. Sensors, 2020, 20(24): 7265.
- [5] ZHANG G H, XU P H, XU H S, et al. Prediction on the urban GNSS measurement uncertainty based on deep learning networks with long short-term memory[J]. IEEE Sensors Journal, 2021, 21(18): 20563-20577.
- [6] SUN R, HSU L T, XUE D B, et al. GPS signal reception classification using adaptive Neuro-Fuzzy inference system[J]. Journal of Navigation, 2019, 72(3): 685-701.
- [7] WANG J, HAN H Z, MENG X L, et al. Robust wavelet-based inertial sensor error mitigation for tightly coupled GPS/BDS/INS integration during signal outages[J]. Survey Review, 2017, 49(357): 419-427.

- [8] CHANG G B. Robust Kalman filtering based on Mahalanobis distance as outlier judging criterion[J]. Journal of Geodesy, 2014, 88(4): 391-401.
- [9] ZHANG Q, NIU X J, SHI C. Impact assessment of various IMU error sources on the relative accuracy of the GNSS/INS systems[J]. IEEE Sensors Journal, 2020, 20(9): 5026-5038.
- [10] 杨元喜. 动态系统的抗差 Kalman 滤波[J]. 解放军测绘学院学报, 1997, 14(2): 79-84.

YANG Y X. Robust Kalman filter for dynamic systems[J]. Journal of the PLA Institute of Surveying and Mapping, 1997, 14(2): 79-84(in Chinese).

- [11] 王坚,刘超,高井祥,等. 基于抗差 EKF 的 GNSS/INS 紧组合算法 研究[J]. 武汉大学学报 (信息科学版), 2011, 36(5): 596-600.
 WANG J, LIU C, GAO J X, et al. GNSS/INS tightly coupled navigation model based on robust EKF[J]. Geomatics and Information Science of Wuhan University, 2011, 36(5): 596-600(in Chinese).
- [12] JIANG C, ZHANG S B, LI H, et al. Performance evaluation of the filters with adaptive factor and fading factor for GNSS/INS integrated systems[J]. GPS Solutions, 2021, 25(4): 130-141.
- [13] 吴富梅,杨元喜.基于小波阈值消噪自适应滤波的 GPS/INS 组合导航[J]. 测绘学报, 2007, 36(2): 124-128.
 WU F M, YANG Y X. GPS/INS integrated navigation by adaptive filtering based on wavelet threshold denoising[J]. Acta Geodaetica et Cartographica Sinica, 2007, 36(2): 124-128(in Chinese).
- GAO B B, HU G G, ZHONG Y M, et al. Cubature Kalman filter with both adaptability and robustness for tightly-coupled GNSS/ INS integration[J]. IEEE Sensors Journal, 2021, 21(13): 14997-15011.
- [15] CHANG G B. Kalman filter with both adaptivity and robustness[J]. Journal of Process Control, 2014, 24(3): 81-87.
- [16] YANG Y, HE H, XU G. Adaptively robust filtering for kinematic geodetic positioning[J]. Journal of Geodesy, 2001, 75(2-3): 109-116.
- [17] 杨元喜, 任夏, 许艳. 自适应抗差滤波理论及应用的主要进展[J].
 导航定位学报, 2013, 1(1): 9-15.
 YANG Y X, REN X, XU Y. Main progress of adaptively robust filter with applications in navigation[J]. Journal of Navigation and Positioning, 2013, 1(1): 9-15(in Chinese).
- [18] 谭兴龙, 王坚, 韩厚增. 支持向量回归辅助的 GPS/INS 组合导航 抗差自适应算法[J]. 测绘学报, 2014, 43(6): 590-597. TAN X L, WANG J, HAN H Z. SVR aided adaptive robust filtering algorithm for GPS/INS integrated navigation[J]. Acta Geodaetica et Cartographica Sinica, 2014, 43(6): 590-597(in Chinese).
- [19] NING Y P, WANG J, HAN H Z, et al. An optimal radial basis function neural network enhanced adaptive robust Kalman filter for GNSS/INS integrated systems in complex urban areas[J]. Sensors, 2018, 18(9): 3091.
- [20] 高为广,陈谷仓.结合自适应滤波和神经网络的 GNSS/INS 抗差 组合导航算法[J]. 武汉大学学报 (信息科学版), 2014, 39(11): 1323-1328.

GAO W G, CHEN G C. Integrated GNSS/INS navigation algorithms combining adaptive filter with neural network[J]. Geomatics and Information Science of Wuhan University, 2014, 39(11): 1323-1328(in Chinese).

- [21] ZHANG C, 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): 1-21.
- [22] ATIA M, WASLANDER S. Map-aided adaptive GNSS/IMU sensor fusion scheme for robust urban navigation[J]. Measurement, 2019, 131: 615-627.
- [23] WANG L, GROVES P, ZIEBART M. Multi-constellation GNSS performance evaluation for urban canyons using large virtual reality city models[J]. Journal of Navigation, 2012, 65(3): 459-476.
- [24] HSU L T, TOKURA H, KUBO N, et al. Multiple faulty GNSS measurement exclusion based on consistency check in urban canyons[J]. IEEE Sensors Journal, 2017, 17(6): 1909-1917.
- [25] SUN R, FU L X, WANG G Y, et al. Using dual-polarization GPS antenna with optimized adaptive neuro-fuzzy inference system to improve single point positioning accuracy in urban canyons[J]. Navigation, 2021, 68(1): 41-60.
- [26] BEST D, ROBERTS D. The upper tail probabilities of Spearman's Rho[J]. Applied Statistics, 1975, 24(3): 377-379.

Robust adaptive position algorithm for GNSS/IMU based on pseudorange residual and innovation

LIU Zhengwu, SUN Rui*, JIANG Lei

(College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China)

Abstract: Algorithms of robust filtering and adaptive filtering are commonly used to improve the positioning accuracy of the navigation system integrating global navigation satellite system (GNSS) and inertial measurement units (IMU). However, the conditions applicable to robust filtering and adaptive filtering are different, and improper use of the filter may reduce the positioning accuracy of the integrated system. To solve this problem, a robust adaptive position algorithm for GNSS/IMU is proposed based on pseudorange residual and innovation. The positioning quality of GNSS is evaluated based on pseudorange residuals. The appropriate filtering algorithm is selected to solve the GNSS/IMU integrated navigation. Innovation and pseudorange residuals are then used to determine whether the IMU kinematic estimation error is greater than the GNSS observation error in long-time low quality GNSS. The robust factor is used based on the determined results. Experimental results show that the positioning accuracy of the proposed algorithm is improved by 36.05%, 22.71%, and 56.22% in the east, north and up directions, respectively, compared with the results from the extended Kalman filter algorithm.

Keywords: pseudorange residual; innovation; robust filtering; adaptive filtering; GNSS/IMU integrated navigation

URL: link.cnki.net/urlid/11.2625.V.20220801.1844.003

Received: 2022-05-19; Accepted: 2022-07-17; Published Online: 2022-08-02 15:32

Foundation items: National Natural Science Foundation of China (42174025,41974033); Horizon 2020 EU-China aviation technology cooperation project, Greener Air Traffic Operations (MJ-2020-S-03); Natural Science Foundation of Jiangsu Province (BK20211569); Jiangsu provincial Six Talent Peaks Project (KTHY-014)

^{*} Corresponding author. E-mail: rui.sun@nuaa.edu.cn

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

DOI: 10.13700/j.bh.1001-5965.2022.0390

基于均匀圆阵矢量传感器的 DOA 和 极化参数联合估计

时春鹏¹,何华锋^{1,*},何耀民¹,李震¹,闫少强²,张孝宇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)

摘 要: 针对极化敏感阵列中非完备电磁矢量传感器多参数联合估计问题,提出一种基于 三维电磁矢量传感器平面圆阵模型的波达方向 (DOA) 和极化参数联合估计方法。利用阵列流型矩阵 的特性将四维谱函数进行解耦,严格证明 DOA 搜索与极化参数估计的不相关性,将四维谱搜索优 化为仅与 DOA 相关的二维谱搜索,提出 DOA 谱峰搜索二级比较策略,可使雷达自动识别谱峰坐 标,有效提高信号处理的实时性;提出一种基于精英反向学习和 Lévy 飞行的改进粒子群算法对极 化参数进行估计,提高极化参数的收敛性能。通过与同类算法进行仿真对比,结果表明:所提算法 在牺牲少量时间的条件下可提高估计精度。

关 键 词:参数估计;波达方向搜索策略;粒子群算法;精英反向学习;Lévy飞行;极化敏 感阵列

中图分类号: TN911.7

文献标志码: A 文章编号: 1001-5965(2024)04-1325-11

阵列信号处理是现代信号处理技术的重要组成部分,在雷达、声呐、地震勘测等领域有良好的应用前景和发展潜力。波达方向(direction of arrival, DOA)估计是阵列信号处理的一个重要分支,通过分布在不同空间位置的传感器阵列对来波信号进行接收采样,进而通过特定算法估计 DOA。随着电子对抗技术的发展,由于传统的标量传感器阵列不能获得完备的电场和磁场信息,其已无法满足被动测向系统的需要。

极化描述了电磁波矢量特性,是幅度、频率、 相位以外的重要参量,表征电场矢端在其传播界面 上的轨迹特性。充分利用电磁波的极化信息,有利 于提高阵列信号处理系统的分辨力和抗干扰能力, 从而在现代电子战中占据优势。因此,极化敏感阵 列(polarization sensitive array, PSA)受到了学者们的 广泛关注,以多重信号分类(multiple signal classification, MUSIC)算法^[1-2]和旋转不变性的信号参数估 计(estimation of signal parameters viarotational invariance techniques, ESPRIT)算法^[3-4]为代表的空域参数估计 算法已成功应用于极化敏感阵列。PSA通常由全 电磁矢量传感器(six-component electromagnetic vector sensor, SEVS)组成,每个 SEVS包含中心共点的 3个正交电偶极子和3个正交小磁环,可以获取入 射电磁波完备的六维信息,例如美国 Flame and Russell, Inc 公司制作的 SuperCART 电磁矢量传感 器。然而, SEVS 阵元通道间较为显著的互耦效应 及昂贵的造价,限定了其在工程中的应用。因此, 非完备电磁矢量传感器(incomplete electromagnetic vector sensor, IEVS)被广泛应用于阵列信号处理领 域,与 SEVS 相比, IEVS 利用更少的通道,在获得极 化信息的前提下显著降低了数据运算量^[5-6]。

文献 [7-9] 提出了极化 MUSIC 算法, 对传感器

收稿日期: 2022-05-19; 录用日期: 2022-08-19; 网络出版时间: 2022-09-14 12:54 网络出版地址: link.cnki.net/urlid/11.2625.V.20220913.1810.010 基金项目: 国家自然科学基金(61573366)

*通信作者.E-mail: hhf0903@163.com

引用格式: 时春鹏,何华锋,何耀民,等.基于均匀圆阵矢量传感器的 DOA 和极化参数联合估计 [J]. 北京航空航天大学学报, 2024, 50(4): 1325-1335. SHI C P, HE H F, HE Y M, et al. Joint estimation of DOA and polarization parameters based on uniform circle array with vector sensor [J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50(4): 1325-1335 (in Chinese). 接收数据进行联合 DOA 估计从而得到信号空域-极化域四维参数,但由于该类算法采用四维搜索的 模式导致计算复杂度较高。文献 [10-11] 研究了基 于 PSA 的长矢量(long vector, LV)接收数据模型 LV-MUSIC 算法, 在原有 MUSIC 算法的谱函数中 加入极化信息,通过对空域-极化联合谱函数的四 维搜索估计参数,但该算法中的高维搜索运算量巨 大,导致谱搜索时间较长。李京书和陶建武^[12]提出 利用四元数的强正交约束性对 DOA 和极化参数估 计的四元 MUSIC(quaternion MUSIC, Q-MUSIC)算 法,进而对数据模型进行降维处理,但四元数阵列 接收模型破坏了原有的信号矢量矩阵,仍需采用 LV数据模型对极化参数进行极化谱搜索。司伟建 等^[13]提出基于平面圆阵的降维长矢量 MUSIC (dimension-reduction long vector MUSIC, DL-MUSIC) 算法,通过天线接收电压矩阵和偏导数条件重构 DOA 谱函数后直接计算极化参数,并在来波数大 于阵元数时仍可有效估计信号参数。Hu 等^[14]提出 了一种新的阵列协方差矩阵更新方法,通过奇异值 分解重构噪声子空间,且能对相干信号实现较高精 度的估计。上述文献普遍是采用 DOA 二维谱搜索 获得信号的空域估计参数后,在人为输入 DOA 估 计值的前提下利用不同算法获取极化参数。若采 用人为识别的方法求取极化参数将会消耗大量时 间,不满足现代电子战中高速运载平台雷达对信号 处理实时性的要求。

智能优化算法由于具有解决非线性问题、全局寻优等优点,已有相关文献将智能优化算法成功应 用于 DOA 估计。黄志强等^[15]将遗传算法结合正交 变异算子和蜂王竞争机制引入圆阵的二维谱峰搜 索,提高了谱峰搜索的成功率。文献 [16-17]将蚁 群算法引入 MUSIC 谱估计中,提高了谱函数全局 寻优的能力,简化了计算过程。栾鹏程和吴谟^[18]结 合遗传算法和免疫算法保持了在低信噪比和小快 拍数情况下的估计精度。然而,上述文献普遍未考 虑信号的极化参数估计。

鉴于此,本文首先建立均匀圆阵非完备电磁矢 量传感器阵列(uniform circle array with incomplete electromagnetic vector sensor, UCA-IEVS)接收信号 模型,通过分解四维谱函数中的 DOA 参数与极化 参数,将四维谱搜索优化为仅与 DOA 参数与极化 参数,将四维谱搜索优化为仅与 DOA 参数相关的 二维谱搜索;而后,给出 DOA 参数搜索与极化参数 不相关的一致性证明。在估计 DOA 参数时,本文 提出 DOA 谱峰搜索二级比较策略,该策略能够使 雷达自动识别谱峰坐标并用于极化参数估计简化为 约束条件下的极化函数最优化问题,利用基于精英 反向学习和 Lévy 飞行的改进粒子群算法分别对每 个信号的极化函数寻优得出极化参数,该算法相比 于传统的粒子群(particle swarm optimization, PSO) 算法和鲸鱼算法(whale optimization algorithm, WOA) 具有更快的收敛速度;最后,通对比仿真实验,证明 本文算法有着较高的精度和较短的运算时间。

1 UCA-IEVS 接收信号模型

假设 IEVS 等角度均匀分布在高速运载平台雷 达上,每个 IEVS 包含 3 个相互正交的偶极子。圆 阵由 M 个阵元组成,半径为 R,接收信号为 K(K < M)个互不相关的远场窄带完全极化平面波,第 $k(1 \le k \le K)$ 个信号的方位角 $\theta_k \in [0,2\pi]$ 、俯仰角 $\varphi_k \in [0,\pi/2]$ 。入射波与圆阵中心点的示意图如图 1 所 示, $e_{\theta k}$ 、 $e_{\phi k}$ 为所在平面的一对标准正交基, $e_{\theta k}$ 平行 于 xOy 面, $e_{\phi k}$ 在 OPP'平面内且与 e_{rk} 垂直, $e_{\theta k}$ 、 $e_{\phi k}$ 和 e_{rk} 构成了右手直角坐标系。





以圆心 O 为参考点,令放置在 x 轴上的阵元为 第 1 个阵元,第 $m(1 \le m \le M)$ 个阵元位置坐标为 r_m , 第 k 个入射信号的单位矢量为 r_k ,则阵列对第 k 个 信号的空域导向矢量为

$$\boldsymbol{a}_{s}(\theta_{k},\varphi_{k}) = \left[e^{-j\varphi_{1}}, e^{-j\varphi_{2}}, \cdots, e^{-j\varphi_{m}} \right]^{T}$$
(1)

第 *m* 个阵元对第 *k* 个信号空间相位延迟 φ_{mk}表示为

$$\varphi_{mk} = \frac{2\pi}{\lambda_k} \boldsymbol{r}_k^{\mathrm{T}} \boldsymbol{r}_m \tag{2}$$

式中: λ_k 为第 k个信号的波长。对于完全极化波, 电场矢量 E_k 可利用 e_{ok} 和 e_{ok} 正交基表示为

$$\boldsymbol{E}_{k} = \cos \gamma_{k} \boldsymbol{e}_{\theta k} + \sin \gamma_{k} \mathrm{e}^{\mathrm{j} \eta_{k}} \boldsymbol{e}_{\varphi k} \tag{3}$$

式中: γ_k 、 η_k 被称为极化相位描述子,其正交分解示意如图 2 所示。







极化相位描述子中的极化幅角 $\gamma_k \in (0, \pi/2),$ 极 化相位差 $\eta_k \in [0, 2\pi],$ 具有如下性质:

$$\tan \gamma_k = \frac{|\mathbf{v}_{\varphi k}|}{|\mathbf{v}_{\theta k}|} \tag{4}$$

 $\eta_k = \arg(\mathbf{v}_{\varphi k}) - \arg(\mathbf{v}_{\theta k}) \tag{5}$

式中: $v_{\phi k}$ 、 $v_{\theta k}$ 分别为 $E_k \approx e_{\phi k}$ 、 $e_{\theta k}$ 上的投影。利用 坐标系转换关系^[19],将式(3)中球面坐标系的 E_k 转 换到空间直角坐标系 *Oxyz* 得

$$E_{k} = \left(-\sin\theta_{k}\cos\gamma_{k} + \cos\varphi_{k}\cos\theta_{k}\sin\gamma_{k}e^{j\eta_{k}}\right)e_{x} + \left(\cos\theta_{k}\cos\gamma_{k} + \cos\varphi_{k}\sin\theta_{k}\sin\gamma_{k}e^{j\eta_{k}}\right)e_{y} + \left(-\sin\varphi_{k}\sin\gamma_{k}e^{j\eta_{k}}\right)e_{z} = \left[\begin{array}{c} -\sin\theta_{k}\cos\varphi_{k}\cos\theta_{k}\\ \cos\theta_{k}\cos\varphi_{k}\sin\theta_{k}\\ 0 - \sin\varphi_{k}\end{array}\right] \left[\begin{array}{c} \cos\gamma_{k}\\ \sin\gamma_{k}e^{j\eta_{k}}\end{array}\right]$$
(6)

式中: e_x 、 e_y 、 e_z 表示空间直角坐标系 Oxyz 中 3 个与 坐标轴平行的标准正交基。在平面圆阵中,每个阵 元包含相互正交的 3 个电磁矢量传感器,即每个阵 元对第 k 个信号的极化域-空域导向矢量为

$$\mathbf{q}_{p}(\theta_{k},\varphi_{k},\gamma_{k},\eta_{k}) = \begin{bmatrix} -\sin\theta_{k}\cos\varphi_{k}\cos\theta_{k}\\ \cos\theta_{k}\cos\varphi_{k}\sin\theta_{k}\\ 0-\sin\varphi_{k} \end{bmatrix} \begin{bmatrix} \cos\gamma_{k}\\ \sin\gamma_{k}e^{j\eta_{k}} \end{bmatrix}$$
(7)

因此,阵列对第k个信号的接收导向矢量为

式中: ⊗表示克罗内克 (Kronecker)积。利用 Kronecker 积的性质^[20]将式 (8)改写为

$$\boldsymbol{a}(\theta_k,\varphi_k,\gamma_k,\eta_k) = \boldsymbol{J}(\theta_k,\varphi_k)\boldsymbol{h}(\gamma_k,\eta_k) \tag{9}$$

式中: $J(\theta_k, \varphi_k)$ 和 $h(\gamma_k, \eta_k)$ 分别为

$$\boldsymbol{J}(\theta_k,\varphi_k) = \boldsymbol{a}_{\mathrm{s}}(\theta_k,\varphi_k) \otimes \begin{bmatrix} -\sin\theta_k & \cos\varphi_k\cos\theta_k \\ \cos\theta_k & \cos\varphi_k\sin\theta_k \\ 0 & -\sin\varphi_k \end{bmatrix} (10)$$

$$\boldsymbol{h}(\boldsymbol{\gamma}_k, \boldsymbol{\eta}_k) = \begin{bmatrix} \cos \boldsymbol{\gamma}_k \\ \sin \boldsymbol{\gamma}_k e^{j\boldsymbol{\eta}_k} \end{bmatrix}$$
(11)

通过式 (8)~式 (11) 推导,将 *a*(θ_k,φ_k,γ_k,η_k) 分解 成仅包含空域参数的 *J*(θ_k,φ_k) 和仅包含极化参数的 *h*(γ_k,η_k) 这 2 部分。因此,针对 *K* 个信号入射到 *M* 个 阵元的 3*M*×1 维阵列接收数据矩阵可表示为

$$X(t) = \sum_{k=1}^{K} \boldsymbol{a}(\theta_k, \varphi_k, \gamma_k, \eta_k) s_k(t) + N(t) =$$

$$\sum_{k=1}^{K} \boldsymbol{J}(\theta_k, \varphi_k) \boldsymbol{h}(\gamma_k, \eta_k) s_k(t) + N(t) =$$

$$\boldsymbol{AS}(t) + N(t)$$
(12)

式中: $A=[a(\theta_1,\varphi_1,\gamma_1,\eta_1),a(\theta_2,\varphi_2,\gamma_2,\eta_2),...,a(\theta_K,\varphi_K,\gamma_K,\eta_K)]$ 为 $3M \times K$ 维阵列流型矩阵; $S(t)=[s_1(t),s_2(t),...,s_k(t)]^T$ 为 $K \times 1$ 维入射信号矢量; $N(t)=[n_1^{T}(t),n_2^{T}(t),...,n_M^{T}(t)]^T$ 为 $3M \times 1$ 维噪声矢量; 本文中假设噪声为 0 均值、 σ^2 方差的高斯白噪声。

2 入射信号的 DOA 估计

阵列接收数据矩阵X(t)的协方差矩阵为

$$\boldsymbol{R}_{\boldsymbol{X}} = \boldsymbol{E}[\boldsymbol{X}(t)\boldsymbol{X}^{\mathrm{H}}(t)] = \boldsymbol{A}\boldsymbol{R}_{\mathrm{S}}\boldsymbol{A}^{\mathrm{H}} + \boldsymbol{R}_{\mathrm{N}}$$
(13)

式中: R_s 和 R_N 分别为入射信号和噪声的协方差矩 阵; 上标 H 表示共轭转置。由于阵列只能接收有限 长度的信号数据,在实际被动测向系统中,通常利 用 R_x 的最大似然估计 \hat{R}_x 代替,根据最大似然理论得

$$\widehat{R}_{X} = \frac{1}{L} \sum_{t=1}^{L} X(t) X^{\mathrm{H}}(t)$$
(14)

式中:L为快拍采样数。对 \hat{R}_X 进行特征值分解有

$$\widehat{\boldsymbol{R}}_{\boldsymbol{X}} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{U}^{\mathrm{H}} = \boldsymbol{U}_{\mathrm{S}}\boldsymbol{\Sigma}_{\mathrm{S}}\boldsymbol{U}_{\mathrm{S}}^{\mathrm{H}} + \boldsymbol{U}_{\mathrm{N}}\boldsymbol{\Sigma}_{\mathrm{N}}\boldsymbol{U}_{\mathrm{N}}^{\mathrm{H}}$$
(15)

式中: $U=[U_1, U_2, \dots, U_{3M}]$ 为特征向量矩阵, 对角阵 $\Sigma=diag(\lambda_1, \lambda_2, \dots, \lambda_{3M})$ 中的特征值与特征向量——对 应; U_S 为 K个大特征值所对应特征向量张成的信 号子空间; U_N 为 3M-K个小特征值所对应特征向量 张成的噪声子空间; Σ_S, Σ_N 为 U_S, U_N 中特征向量对 应的特征值构成的对角阵; 利用子空间的正交性构 造四维谱函数为

$$P = \frac{1}{\boldsymbol{a}^{\mathrm{H}}(\theta,\varphi,\gamma,\eta)\boldsymbol{U}_{\mathrm{N}}\boldsymbol{U}_{\mathrm{N}}^{\mathrm{H}}\boldsymbol{a}(\theta,\varphi,\gamma,\eta)}$$
(16)

当 (θ,φ,γ,η) 为入射信号的真实参数 (θ_k,φ_k,γ_kη_k) 时,由子空间原理得

$$\boldsymbol{a}^{\mathrm{H}}(\theta_{k},\varphi_{k},\gamma_{k},\eta_{k})\boldsymbol{U}_{\mathrm{N}}\boldsymbol{U}_{\mathrm{N}}^{\mathrm{H}}\boldsymbol{a}(\theta_{k},\varphi_{k},\gamma_{k},\eta_{k}) = \boldsymbol{h}^{\mathrm{H}}(\gamma_{k},\eta_{k})\boldsymbol{J}^{\mathrm{H}}(\theta_{k},\varphi_{k})\boldsymbol{U}_{\mathrm{N}}\boldsymbol{U}_{\mathrm{N}}^{\mathrm{H}}\boldsymbol{J}(\theta_{k},\varphi_{k})\boldsymbol{h}(\gamma_{k},\eta_{k}) = \boldsymbol{0}$$

$$(17)$$

式中: (γ_k, η_k) 为极化参数; (θ_k, φ_k) 为 DOA 参数。由

于 $U_{N}U_{N}^{H}$ 仅与噪声子空间有关,不包含任何 DOA 或 极 化 参 数,故 $J^{H}(\theta_{k},\varphi_{k})U_{N}U_{N}^{H}J(\theta_{k},\varphi_{k})$ 仅 包 含 信 号 的 DOA 信息。

由式(9)可知,当 $\gamma_k \in (0,\pi/2)$ 时, $h(\gamma_k,\eta_k)\neq 0$,故 $J^{H}(\theta_k,\varphi_k)U_NU_N^H J(\theta_k,\varphi_k)$ 和 $a^{H}(\theta_k,\varphi_k,\gamma_k,\eta_k)U_NU_N^H a(\theta_k,\varphi_k,\gamma_k,\eta_k)$ 是完全等价的。下面给出 $J^{H}(\theta_k,\varphi_k)U_NU_N^H J(\theta_k,\varphi_k)=0$ 和 $a^{H}(\theta_k,\varphi_k,\gamma_k,\eta_k)U_NU_N^H a(\theta_k,\varphi_k,\gamma_k,\eta_k)=0$ 对 DOA 谱峰搜 索的一致性证明。

充分性。当(θ,φ)为入射信号真实 DOA 时,
 上述两式同时成立。

证明 当 (θ , φ) 为入射信号真实 DOA 时, 根据 MUSIC 算法子空间原理, $a(\theta_k,\varphi_k,\gamma_k,\eta_k)$ 必在阵列协方 差矩阵的信号子空间中, 此时 $a^{H}(\theta_k,\varphi_k,\gamma_k,\eta_k)U_NU_N^{H}$. $a(\theta_k,\varphi_k,\gamma_k,\eta_k)=h^{H}(\gamma_k,\eta_k)J^{H}(\theta_k,\varphi_k)U_NU_N^{H}J(\theta_k,\varphi_k)h(\gamma_k,\eta_k)=0$ 。 因为 $\gamma_k \in (0,\pi/2), h(\gamma_k,\eta_k)\neq 0$, 故必有 $J^{H}(\theta_k,\varphi_k)U_NU_N^{H}$. $J(\theta_k,\varphi_k)=0$ 成立。证毕。

2) 必要性。当 $J^{H}(\theta_{k},\varphi_{k})U_{N}U_{N}^{H}J(\theta_{k},\varphi_{k})=0$ 时, (θ,φ) 的取值与 $a^{H}(\theta_{k},\varphi_{k},\gamma_{k},\eta_{k})U_{N}U_{N}^{H}a(\theta_{k},\varphi_{k},\gamma_{k},\eta_{k})=0$ 得出的 值一致。

证明 假设 $(\theta_0,\varphi_0) \neq (\theta_k,\varphi_k)$,同时 $J^{H}(\theta_0,\varphi_0)U_NU_N^{H}$. $J(\theta_0,\varphi_0)=0$,此等式两边分别左乘 $h^{H}(\gamma_0,\eta_0)$ 和右 乘 $h(\gamma_0,\eta_0)$ 得 $h^{H}(\gamma_0,\eta_0)J^{H}(\theta_0,\varphi_0)U_NU_N^{H}J(\theta_0,\varphi_0)h(\gamma_0,\eta_0)=$ $a^{H}(\theta_0,\varphi_0,\gamma_0,\eta_0)U_NU_N^{H}a(\theta_0,\varphi_0,\gamma_0,\eta_0)=0$,根据MUSIC子 空间原理,只有在真实信号处 $a^{H}(\theta,\varphi,\gamma,\eta)U_NU_N^{H}$. $a(\theta,\varphi,\gamma,\eta)$ 才会为0,故 $a(\theta_0,\varphi_0,\gamma_0,\eta_0)$ 即为阵列对信号 的接收导向矢量,假设不成立,故 $(\theta_0,\varphi_0)=(\theta_k,\varphi_k)$ 。 证毕。

因此,利用 $J^{H}(\theta_{k},\varphi_{k})U_{N}U_{N}^{H}J(\theta_{k},\varphi_{k}) 和 a^{H}(\theta_{k},\varphi_{k},\gamma_{k},\eta_{k})$ · $U_{N}U_{N}^{H}a(\theta_{k},\varphi_{k},\gamma_{k},\eta_{k})$ 完全等价的性质,可基于式(17) 对二维参数 (θ,φ) 进行搜索得到极大值,从而得到 入射信号的 DOA 估计值。

$$f(\theta,\varphi) = \frac{1}{\det \left| \boldsymbol{J}^{\mathrm{H}}(\theta,\varphi) \boldsymbol{U}_{\mathrm{N}} \boldsymbol{U}_{\mathrm{N}}^{\mathrm{H}} \boldsymbol{J}(\theta,\varphi) \right|}$$
(18)

在实际被动测向系统中,接收到的信号个数 K往往大于 1,因此,DOA 的谱估计图往往呈现多 峰形式,并不能直接获取 K 个信号的谱峰坐标。若 采用人为输入谱峰坐标、再求取极化参数,该过程 将消耗大量时间,不满足信号处理实时性的要求。 因此,本文提出 DOA 谱峰搜索二级比较策略,目的 是帮助雷达自动识别谱峰坐标,而后将其用于极化 参数的估计。本文二级比较策略由 1、2 两级比较 器组成:第1级比较器选取整个谱函数的全局最大 值,输出最大值对应坐标,而后将谱函数最大值置 于 0;第 2 级比较器由迭代计数器、信号计数器和 比较器组成。假设谱峰搜索角度间隔为 n,二级比 较策略的步骤如下: 步骤1 令迭代计数器*i*和目标信号计数器*l*的 初始值均为1,选取谱函数 $f(\theta,\varphi)$ 的全局最大值 $f(\theta_1,\varphi_1)$,输出其对应的横纵坐标为信号的 DOA 估 $t(\hat{\theta}_1,\hat{\varphi}_1), \hat{+} \varphi f(\hat{\theta}_1,\hat{\varphi}_1) = 0;$

步骤 2 令 i = i + 1,若满足 $f(\theta, \varphi)$ 小于周围相 邻 4 个点的谱函数值,则判断该点所对应的谱函数 为谱峰, $f(\hat{\theta}_i, \hat{\varphi}_i) = 0$, l = l + 1,输出 $(\hat{\theta}_i, \hat{\varphi}_i)$;

步骤 3 若不满足 *f*(*θ*,*φ*)小于周围相邻 4 个点的谱函数值,则判断该点不是谱峰,执行步骤 2;

步骤 4 当 $l \ge K$ 时, 跳出第 2 级比较器, 结束循环。 最终得到 DOA 的估计值 ($\hat{\theta}_1, \hat{\varphi}_1$), ($\hat{\theta}_2, \hat{\varphi}_2$),…, ($\hat{\theta}_{\kappa}, \hat{\varphi}_{\kappa}$)。

DOA 谱峰搜索二级比较策略如图 3 所示。





3 入射信号的极化参数估计

通过式(17)进行谱峰搜索后,可获取入射信号的 DOA 参数。为详细表述极化参数估计,将四维联合谱函数重写为

$$P = \frac{1}{\boldsymbol{h}^{\mathrm{H}}(\boldsymbol{\gamma},\boldsymbol{\eta})\boldsymbol{J}^{\mathrm{H}}(\boldsymbol{\theta},\boldsymbol{\varphi})\boldsymbol{U}_{\mathrm{N}}\boldsymbol{U}_{\mathrm{N}}^{\mathrm{H}}\boldsymbol{J}(\boldsymbol{\theta},\boldsymbol{\varphi})\boldsymbol{h}(\boldsymbol{\gamma},\boldsymbol{\eta})}$$
(19)

2024年

由于 K个信号的 DOA 参数已完成估计,当 ($\theta_k, \varphi_k, \gamma, \eta$)取到信号极化参数的真实值 ($\theta_k, \varphi_k, \gamma_k, \eta_k$) 时,联合谱函数值相比其他任意非入射信号极化参 数所对应的值更大。因此,当 (γ, η) \neq (γ_k, η_k)时,第 k个入射信号的极化函数为

$$T_{k}(\gamma,\eta) = \boldsymbol{h}^{\mathrm{H}}(\gamma,\eta)\boldsymbol{J}^{\mathrm{H}}(\theta_{k},\varphi_{k})\boldsymbol{U}_{\mathrm{N}}\boldsymbol{U}_{\mathrm{N}}^{\mathrm{H}}\boldsymbol{J}(\theta_{k},\varphi_{k})\boldsymbol{h}(\gamma,\eta) >$$
$$\boldsymbol{h}^{\mathrm{H}}(\gamma_{k},\eta_{k})\boldsymbol{J}^{\mathrm{H}}(\theta_{k},\varphi_{k})\boldsymbol{U}_{\mathrm{N}}\boldsymbol{U}_{\mathrm{N}}^{\mathrm{H}}\boldsymbol{J}(\theta_{k},\varphi_{k})\boldsymbol{h}(\gamma_{k},\eta_{k})$$
$$(20)$$

此时,极化参数的估计问题可转换为

$$\begin{cases} \min T_k(\gamma, \eta) \\ \text{s.t } \gamma \in \left(0, \frac{\pi}{2}\right), \eta \in [0, 2\pi] \end{cases}$$
(21)

对于每一个确定的入射信号,当(γ , η)遍历各自 的定义域后,极化函数 $T_k(\gamma,\eta)$ 的最优解即为信号极 化参数的估计($\hat{\gamma}_k, \hat{\eta}_k$)。

关于优化问题的求解, PSO 算法是一种基于群体和进化的智能优化算法。与传统优化算法相比, PSO 算法具有更快的计算速度和更好的全局搜索能力。然而, 传统的 PSO 算法也存在一些问题, 如初始粒子分布较大的随机性容易导致早期局部收敛和后期震荡、收敛速度不能满足复杂优化工程的要求等。为有效解决这些问题, 本文提出一种基于精英反向学习和 Levy 飞行的改进粒子群算法。精英反向学习通过可行解中的精英个体构造其反向种群, 从当前和反向粒子种群构成的新种群中选择最优个体作为新一代个体, 进入下一次的迭代, 可有效增加粒子种群的多样性; Lévy 飞行策略能够使粒子在最优解附近进行小范围的搜索, 显著提升了搜索效率和收敛速度。

3.1 传统 PSO 算法

3.1.1 初始化粒子种群

在解空间中随机生成N个粒子:

$$\boldsymbol{x}_{ij} = \boldsymbol{L}_{\mathrm{b},j} + (\boldsymbol{U}_{\mathrm{b},j} - \boldsymbol{L}_{\mathrm{b},j}) \cdot \mathrm{rand}(0,1)$$
(22)

式中: **x**_{ij} 表示 N 个粒子的初始位置,其中 i=1,2,…,N, j=1,2,…,d, d 表示解空间的维数; **U**_{bj} 和 **L**_{bj} 分别为解 空间第 j 维取值范围的最小值和最大值。

3.1.2 粒子的迭代

为实现全局搜索和局部搜索之间的平衡,文 献 [21] 提出了带有惯性权重的粒子群算法,其速度 和位置的迭代过程表示为

$$\mathbf{v}_{ij}(t+1) = w\mathbf{v}_{ij}(t) + c_1 r_1(t) [\mathbf{p}_{ij}(t) - \mathbf{x}_{ij}(t)] + c_2 r_2(t) [\mathbf{p}_{gi}(t) - \mathbf{x}_{ij}(t)]$$
(23)

$$\mathbf{x}_{ij}(t+1) = \mathbf{x}_{ij}(t) + \mathbf{v}_{ij}(t+1)$$
(24)

式中: $|v_{ii}| \in [-v_{max}, v_{max}]$ 为粒子的速度, v_{max} 为常数;

w为惯性权重; c₁、c₂为学习因子; r₁、r₂为0到1之间的随机数; p_{ij}为第 i 个粒子的历史最优位置; p_{gi}为整个粒子种群搜索到的最优位置; t 为当前迭代次数。

为保证粒子在极值点附近拥有更为精确的搜 索性能,文献[22]提出的动态惯性权重线性递减策 略被广大学者接受并认可,表达式为

$$w = w_{\max} - \frac{(w_{\max} - w_{\min})t}{T_{\max}}$$
(25)

式中:w_{max}为最大惯性权重;w_{min}为最小惯性权重; T_{max}为最大迭代次数。

3.2 精英反向学习初始化粒子种群

在传统 PSO 算法中, 粒子的初始位置往往采用 随机生成的方式, 由于最优解的位置未知, 随机生成 的粒子可能不能覆盖最优解附近的空间导致收敛 速度较慢。本文利用精英反向学习对初始粒子种 群进行优化, 精英反向学习粒子位置如式^[23](26) 所示:

$$\overline{\boldsymbol{x}}_{ij} = \operatorname{rand}(0, 1) \cdot (\boldsymbol{L}_{\mathrm{b}, j} + \boldsymbol{U}_{\mathrm{b}, j}) - \boldsymbol{x}_{ij}$$
(26)

以式(19)的极化函数作为适应度函数,通过比 较初始化粒子种群和其反向学习粒子种群每个粒 子的适应度函数,筛选出 N 个适应度较优的粒子作 为算法的优化粒子种群^[24],即

$$\boldsymbol{x}_{ij} = \begin{cases} \boldsymbol{\bar{x}}_{ij} & T\left(\boldsymbol{\bar{x}}_{ij}\right) < T\left(\boldsymbol{x}_{ij}\right) \\ \boldsymbol{x}_{ij} & T\left(\boldsymbol{\bar{x}}_{ij}\right) \ge T\left(\boldsymbol{x}_{ij}\right) \end{cases}$$
(27)

3.3 Lévy 飞行策略

Lévy 飞行是服从 Lévy 分布的随机搜索方法, 是一种短距离的搜索与偶尔较长距离相结合的行 走方式。本文将 Lévy 飞行策略应用于 PSO 算法, 可实现在最优解附近小范围搜索,提高了搜索效 率,大幅降低了陷入局部最优的概率。Lévy 飞行的 位置更新表达式^[25] 为

$$\boldsymbol{x}(t+1) = \boldsymbol{x}(t) + \alpha L(\beta)\boldsymbol{x}(t)$$
(28)

式中: **x**(*t*)为第 *t* 代粒子种群的位置; α为步长缩放 因子; *L* 为优化步长系数, 服从 Lévy 飞行分布:

$$L \sim \frac{\beta \Gamma(\beta) \sin(\pi \beta/2)}{\pi} \cdot \frac{1}{s^{1+\beta}} \qquad s \ge s_0 \ge 0$$
 (29)

式中: β 为常数,通常取 1.5; $\Gamma(\beta)$ 为伽马函数; *s* 表达 式为

$$s = \frac{\mu}{|\nu^{1/\beta}|} \quad \mu \sim N(0, \sigma^2), \nu \sim N(0, 1)$$
 (30)

$$\sigma^{2} = \left\{ \frac{\Gamma(1+\beta)}{\beta\Gamma((1+\beta)/2)} \cdot \frac{\sin(\beta\pi/2)}{2^{(\beta-1)/2}} \right\}$$
(31)

3.4 本文算法估计极化参数的步骤

将本文算法引入寻找极化函数 *T_k*(γ,η) 全局最 优解的搜索过程,步骤如下: **步骤1** 设定粒子规模 N, 最大迭代次数 T_{max}, 可行解维度 d, 可行解边界 L_b、U_b, 粒子最大速度 v_{max}, 最大和最小惯性权重 w_{max}、w_{min} 等算法基本参数;

步骤2 在可行解范围内随机生成N个初始化 粒子并计算每一个粒子的适应度值;

步骤3 采用精英反向学习优化粒子种群,计 算优化后粒子种群的适应度值,筛选出N个适应度 较优的粒子作为算法的优化粒子种群,将适应度值 最优的粒子位置作为当前最优位置;

步骤4 根据式(23)和式(24)对每个粒子的速 度和位置进行更新,并计算粒子的适应度值;

步骤 5 利用式(28)对当前最优位置进行 Lévy 飞行策略更新,计算粒子的适应度值;

步骤 6 比较 Lévy 飞行策略优化前后粒子的 适应度值,保留适应度值较优的粒子位置;

步骤7 若当前迭代次数小于最大迭代次数, 返回第2步,否则结束算法,输出最优粒子位置以 及最优适应度值。

4 仿真实验

本文的仿真实验均是在装有 Inter® CoreTM 2Duo E8400@3.00 GHz和8.00 GB 的RAM, MATALB R2020a PC 机上完成的。仿真采用 10 个阵元组成 的极化平面圆阵, 10 阵元均匀地分布在同一圆周 上,圆阵半径 R 取入射信号的半波长,信号频率均 为 2 GHz,系统采样频率为 8 GHz,快拍采样数 L_s = 1 024。假设入射信号均为相互独立的远场窄带完 全极化波,信噪比 (signal-noise ratio, SNR) 为 10 dB, 信源数 K=3,入射信号的参数设置为 ($\theta_1, \varphi_1, \gamma_1, \eta_1$)= (300°,10°,0°,80°), ($\theta_2, \varphi_2, \gamma_2, \eta_2$)=(70°,30°,40°,60°), ($\theta_3, \varphi_3, \gamma_3, \eta_3$)= (120°,60°,75°,100°)。

4.1 入射信号的 DOA 估计

设定谱搜索间隔 n=0.5°, 图 4 和图 5 分别为在 极化参数未知的情况下本文算法的 DOA 的谱估计 图及俯视图。

在图 4 中,每个谱峰所对应的横纵坐标即为信









号的方位角与俯仰角,在信号极化参数未知的情况 下能够获得信号的 DOA 参数。由图 4 可知,本文 方法可清晰分辨出谱峰位置,且不存在"伪峰"。 进一步由图 5 可知,除了信号真实 DOA 附近有等 高线外,其余位置均未有等高线,说明谱峰仅在信 号真实 DOA 附近形成,且在真实 DOA 处形成的谱 峰较为尖锐。

在获得信号的 DOA 谱估计图后, 若采用传统的人为输入 DOA 估计($\hat{\theta}_k, \hat{\varphi}_k$), 再将其用于极化参数的估计, 对于每个信号, 此过程平均所需时间约为3 s。采用谱峰搜索二级比较策略舍弃人为操作的流程, 雷达可自动识别 DOA 估计图的所有极大值并将其用于每个信号极化参数的估计, 所需时间仅为3 ms 左右。此外, 由表 1 可知, 随着信源数 k 越大其优越性体现的越为明显, 极大地减少了算法运行时间。

表1	1 不同信源数下识别谱峰所需时间
Table 1	Time required to identify spectral peaks with

	different number	ers of sources s		
信源数	所需时间			
	人为输入	采用DOA二级比较策略		
1	约3	2×10 ⁻³		
2	约6	7.4×10 ⁻³		
3	约9	7.6×10 ⁻³		
4	约12	8.2×10 ⁻³		
5	约15	18.4×10 ⁻³		

4.2 入射信号的极化参数估计

在获取所有信号的 DOA 估计($\hat{\theta}_k, \hat{\varphi}_k$)后, 对每个 信号的极化参数进行估计。图 6 为 3 个信号的极 化函数三维空间图, 为使三维图清晰可视, 使 η 由 大到小排列。由图 5~图 7 可知, $T_k(y,\eta)$ 仅存在一 个局部最优点即全局最优点, 可证明本文算法对极 化参数估计的可行性。

传统 PSO 算法的实验条件设置如下:可行解维



图 7 $T_k(y,\eta)$ 的等高线俯视图及粒子的初始化分布 (N=20)

Fig. 7 Top view of $T_k(\gamma, \eta)$ contour, and initial distribution of particles (*N*=20)

度 d=2,最大进化代数 $T_{max}=50$,学习因子 c_1 、 c_2 均为 2,最大惯性权重 $w_{max}=0.9$,最小惯性权重 $w_{min}=0.2$, 粒子的最大速度 $v_{max}=6$ 。本文算法的基本参数与上 述相同,步长缩放因子 $\alpha=0.01$ 。表 2 为不同粒子数 N 对传统 PSO 算法和本文算法收敛性能对比。其 中收敛性能指标包含算法的平均收敛代数为 $C_{interation}$,算法的平均收敛时间为 C_{time} ,算法的平均最 优适应度值为 C_{value} 。对于上述 3 个信号极化函数 (见式(32)~式(34)),每个信号的极化函数做 100 次 寻优实验。

$$C_{\text{interation}} = \frac{1}{300} \sum_{i=1}^{100} \sum_{k=1}^{3} (C_{\text{interation}})_{ik}$$
(32)

$$C_{\text{time}} = \frac{1}{300} \sum_{k=1}^{100} \sum_{k=1}^{3} (C_{\text{time}})_{ik}$$
(33)

$$C_{\text{vaule}} = \frac{1}{300} \sum_{i=1}^{100} \sum_{k=1}^{3} (C_{\text{vaule}})_{ik}$$
(34)

式中:(C_{interation})_{ik}为第 i 次实验中第 k 个信号极化函数的粒子收敛到最优值的迭代次数;(C_{time})_{ik}为第 i 次实验中第 k 个信号极化函数的粒子收敛到最优 值的迭代时间;(C_{vaule})_{ik}为第 i 次实验中第 k 个信号

表 2 不同粒子数的 PSO 算法和本文算法收敛性能 Table 2 Convergence performance between PSO algorithm and the proposed algorithm with different particle numbers

Ν	$C_{\text{interation}}$		$C_{\rm time}/{ m s}$		$C_{ m value}$	
	PSO算法	本文算法	PSO算法	本文算法	PSO算法	本文算法
10	25.6	13.4	0.005 9	0.008 2	0.007 6	0.007 1
15	17.86	10.66	0.006 7	0.009 1	0.006 6	0.006 2
20	16.8	8.66	0.008 5	0.009 9	0.005 4	0.005 3
25	15.73	8.46	0.010 3	0.011 6	0.005 4	0.005 3
30	13.8	8.53	0.010 8	0.012 9	0.005 4	0.005 3

极化函数的粒子收敛到的最优适应度值。

从实验数据可以看出,在 N 相同的条件下,本 文算法的平均迭代时间比传统 PSO 算法的平均迭 代时间略长,这是由于本文算法增加了粒子优化与 搜索策略,必然会导致时间的略微延长,但收敛 精度有所提高。当 N<20 时,2 种算法均未收敛到 最优值。当 N>20 时,2 种算法的收敛精度趋于 一致但收敛时间会随着 N 的增大而延长。故当 N=20 时,实验效果最佳 (表 2 中数据加粗),有利于 获得较好的收敛性能。 图 7 为每个信号极化函数的等高线俯视图及 N=20 时 2 种算法的初始化粒子分布图,由图可知, 本文算法的粒子比 PSO 算法的粒子更接近真实值, 故本文算法的收敛性能更好。

为增加对比,将文献 [26] 中的 WOA 引入极化 参数的估计并与 PSO、本文算法进行对比。实验 中, WOA 的螺旋形状 b取 1。图 8 为 3 种算法对 $T_1(\gamma,\eta)$ 、 $T_2(\gamma,\eta)$ 和 $T_3(\gamma,\eta)$ 的适应度进化曲线。由图 可知,本文算法的收敛速度最快,收敛代数最小,具 有明显的优势。适应度曲线最终收敛到 0 附近,该 适应度值所对应的 (γ_k,η_k)即为信号的极化参数估计 值($\hat{\gamma}_k, \hat{\eta}_k$)。





4.3 信噪比对估计精度的影响

本文利用均方根误差(root mean square error, RMSE)表征估计参数的精度,即

$$e_{\text{RMSE}}(y) = \sqrt{\frac{1}{kQ} \sum_{k=1}^{K} \sum_{i=1}^{Q} (y_{ik} - \hat{y}_k)^2}$$
(35)

式中:y为4个估计参数的任意一个;Q为蒙特卡罗 实验次数。令 SNR 以2dB为步长变化,且对于每个 信噪比点,Q=100,图 9~图 12分别为 DL-MUSIC 算法、Q-MUSIC 算法和本文算法对信号的方位角 θ 、俯仰角 φ 、极化幅角y和极化相位差 η 随着 SNR 从-4~10 dB的均方误差曲线。由图 9~图 12 可 知,本文算法整体具有更小的均方误差。在高 SNR 条件下,3种算法的均方误差相差较小,但在低 SNR条件下本文算法较其他2种算法有着明显的 优势,估计精度更高。

4.4 不同算法运算性能分析

在实际工程应用中,算法的运算时间是验证算





法能否成功应用的一项重要指标,因此,对 LV-MUSIC 算法、DL-MUSIC 算法、Q-MUSIC 算法和 本文算法的运算性能进行对比。在 MUSIC 类算法



图 10 俯仰角 φ 随 SNR 的 RMSE





图 11 极化幅角 y 随 SNR 的 RMSE Fig. 11 RMSE of polarization amplitude angle φ with respect to SNR



图 12 极化相位差 η 随 SNR 的 RMSE Fig. 12 RMSE of polarization phase angle η with respect to SNR

中,运算性能主要包括运算量和运算时间。算法的运算量主要与电磁矢量传感器维数 D、协方差矩阵 \hat{R}_x 的运算量、(θ , φ)搜索运算量、(γ , η)搜索运算量 有关。除 LV-MUSIC 算法外其余算法均是通过谱 搜索得到 (θ , φ),通过不同算法求取 (γ , η),因此, (γ , η)运算量没有统一的标准来衡量,故增加极化参 数搜索时间对比。表 3 为不同数据类型进行乘法 和加法产生的运算量。如 1 个复数和 1 个四元数 相乘定义为 8 次乘法和 4 次加法运算。此外,复数 取模定义为 2 次乘法和 1 次加法。

LV-MUSIC 算法中, 阵列接收数据矩阵 X为 6 $M \times L$ 维, 构造协方差矩阵 R_X 需要进行 $36M^2 \times L$ 次 复数乘法, 通过谱函数(见式(36))对四维参数每个 节点搜索需要进行 6M(6M-K) 次复数乘法和 (6M-K)次复数取模运算。

 $\left\|\boldsymbol{a}^{\mathrm{H}}(\boldsymbol{\theta},\boldsymbol{\varphi},\boldsymbol{\gamma},\boldsymbol{\eta})\boldsymbol{U}_{\mathrm{N}}\right\|^{2} \tag{36}$

DL-MUSIC 算法中, 阵列接收数据矩阵 X为 2 $M \times L$ 维,构造协方差矩阵 R_x 需要进行 4 $M \times L$ 次复 数乘法,通过谱函数(见式(37))估计 DOA。

$$\boldsymbol{a}_{s}^{H}(\theta,\varphi)\boldsymbol{U}_{N1}\boldsymbol{U}_{N1}^{H}\boldsymbol{a}_{s}(\theta,\varphi) - \frac{\left|\boldsymbol{a}_{s}^{H}(\theta,\varphi)\boldsymbol{U}_{N1}\boldsymbol{U}_{N2}^{H}\boldsymbol{a}_{s}(\theta,\varphi)\right|^{2}}{\boldsymbol{a}_{s}^{H}(\theta,\varphi)\boldsymbol{U}_{N2}\boldsymbol{U}_{N2}^{H}\boldsymbol{a}_{s}(\theta,\varphi)}$$
(37)

表 3 不同数据类型进行乘、加运算产生的运算量

 Table 3 Amount of computation generated by multiplication

and addition of different data types					
数据类型	运算类型 -	运算量/次			
		实数	复数	四元数	
实数	相乘	1(乘法)			
	相加	1(加法)			
复数	相乘	2(乘法)	4(乘法) 2(加法)	8(乘法) 4(加法)	
	相加	1(加法)	2(加法)	2(加法)	
四元数	相乘			16次(乘法) 12次(加法)	
	相加			4(加法)	

式中: U_{N1}、U_{N2}平分了噪声子空间,每个搜索节点需 要进行 (2*M*+1)(2*M*-K) 次复数乘法和 (2*M*-K) 次复 数取模运算。

Q-MUSIC 算法中, 阵列接收数据矩阵X为 $2M \times L$ 维,构造协方差四元数矩阵 R_x 需要进行 $4M^2 \times L$ 次复数乘法,通过谱函数(见式(38))估计 DOA。

$$\left\|\boldsymbol{q}^{\mathrm{H}}(\boldsymbol{\theta},\boldsymbol{\varphi})\boldsymbol{U}_{\mathrm{N}}\right\|^{2} \tag{38}$$

式中:q为复数向量;U_N为四元数,每个搜索节点需要进行2M(2M-K)次复数乘法和(2M-K)次复数乘法和(2M-K)次复数取模运算。

本文算法中, 阵列接收数据矩阵 X 为 3M×L 维, 构造协方差矩阵 **R**_x 需要进行 9M²×L 次复数乘 法, 通过式(18)估计 DOA, 每个节点搜索需要进行 3M(3M-K) 次复数乘法和 (3M-K) 次复数取模运算。

表 4 为在谱搜索间隔 n=1°, M=10, k=3(其余参数不变)情况下 4 种算法的运算性能,由表可知,由于本文算法仅在 DOA 的参数估计中进行谱峰搜索,较 LV-MUSIC 算法的运行时间缩短了约 2.1×10⁴倍。本文算法搜索运算量较 DL-MUSIC 和 Q-MUSIC 算法略大,这是由于传感器的维数不同。本文算法总运行时间较 DL-MUSIC 和 Q-MUSIC 略长,这是由于在极化参数的估计中引入了 PSO 算法和粒子优化策略,结合图 9~图 12 可知,本文算法以牺牲较短时间为代价的同时提升了一定的估计精度。

	表 4 MUSIC 类算法运算性能
Table 4	Computational performance among 4 MUSIC-type algorithms

)

算法	D	\widehat{R}_X 运算量/次	(θ,φ)运算量/次	(γ,η)搜索时间/s	总运行时间/s
LV-MUSIC	6	1.474 6×10 ⁷ (乘法) 7.372 8×10 ⁶ (加法)	7.240 2×10 ¹² (乘法) 3.620 0×10 ¹² (加法)		46 047.9
DL-MUSIC	2	1.638 4×10 ⁶ (乘法) 8.192 0×10 ⁵ (加法)	4.736 9×10 ⁷ (乘法) 2.698 4×10 ⁷ (加法)	0.0108	1.304 4
Q-MUSIC	2	1.638 4×10 ⁶ (乘法) 8.192 0×10 ⁵ (加法)	4.516 6×10 ⁷ (乘法) 2.225 8×10 ⁷ (加法)	0.1604	1.562 7
本文算法	3	3.686 4×10 ⁶ (乘法) 1.843 2×10 ⁶ (加法)	1.067 2×10 ⁸ (乘法) 5.336 3×10 ⁷ (加法)	0.081 3	2.168 2

5 结 论

本文提出了一种基于均匀圆阵矢量传感器的DOA和极化参数的联合估计方法。对四维谱函数降维,在极化参数未知的情况下优先进行DOA的估计,降低了搜索的难度。

2)提出了 DOA 谱峰搜索二级比较策略,能够 使雷达自动识别并存储谱峰坐标,将 DOA 的估计 值用于极化参数的估计过程从秒级减少至毫秒级, 显著提高了雷达系统实时信号处理的能力。

3) 采用基于精英反向学习和 Lévy 飞行策略的 改进 PSO 算法对信号的极化参数进行估计,本文 算法相比 PSO 算法和 WOA 有着更好的收敛性能。 通过对比仿真可知,本文算法虽然在运行时间方面 较其他降维算法略长,但估计精度有所提高,仍能 够满足高速运载平台中雷达信号实时性的要求。

参考文献(References)

- SCHMIDT R O. Multiple emitter location and signal parameter estimation[J]. IEEE Transactions on Antennas and Propagation, 1986, 34(3): 276-280.
- [2] CHENG Q, HUA Y B. Further study of the pencil-MUSIC algorithm[J]. IEEE Transactions on Aerospace and Electronic Systems, 1996, 32(1): 284-299.
- [3] LI J, COMPTON R T J. Two-dimensional angle and polarization estimation using the ESPRIT algorithm[J]. IEEE Transactions on Antennas and Propagation, 1992, 40(5): 550-555.
- [4] ROY R, PAULRAJ A, KAILATH T. ESPRIT: A subspace rotation approach to estimation of parameters of cisoids in noise[J]. IEEE Transactions on Acoustics, Speech and Signal Processing, 1986, 34(5): 1340-1342.
- [5] MIR H S, SAHR J D. Passive direction finding using airborne vector sensors in the presence of manifold perturbations[J]. IEEE Transactions on Signal Processing, 2007, 55(1): 156-164.
- [6] BLISS B, CHAN A, EAPEN A. Vector sensors array design[M]. Lexington: MIT Lincoin Laborary Presentation, 2004: 118-123.
- [7] FERRARA E, PARKS T. Direction finding with an array of antennas having diverse polarizations[J]. IEEE Transactions on Antennas and Propagation, 1983, 31(2): 231-236.
- [8] HUA Y. A pencil-MUSIC algorithm for finding two-dimensional angles and polarizations using crossed dipoles[J]. IEEE Transactions on Antennas and Propagation, 1993, 41(3): 370-376.
- [9] CHENG Q, HUA Y B. Performance analysis of the MUSIC and Pencil-MUSIC algorithms for diversely polarized array[J]. IEEE Transactions on Signal Processing, 1994, 42(11): 3150-3165.
- [10] GUO W N, YANG M L, CHEN B X, et al. Joint DOA and polarization estimation using MUSIC method in polarimetric MIMO radar[C]//Proceedings of the IET International Conference on Radar Systems. London: IET Press, 2012.
- [11] 王群. 基于极化敏感阵列的 DOA 估计算法研究[D]. 长春: 吉林 大学, 2011: 17-23.

WANG Q. Study of algorithm on DOA estimation based on polariz-

ation sensitive array[D]. Changchun: Jilin University, 2011: 17-23(in Chinese).

- [12] 李京书, 陶建武. 信号 DOA 和极化信息联合估计的降维四元数 MUSIC 方法[J]. 电子与信息学报, 2011, 33(1): 106-111.
 LI J S, TAO J W. The dimension reduction quaternion MUSIC algorithm for jointly estimating DOA and polarization[J]. Journal of Electronics & Information Technology, 2011, 33(1): 106-111(in Chinese).
- [13] 司伟建,朱瞳,张梦莹.平面极化天线阵列的 DOA 及极化参数降 维估计方法[J]. 通信学报, 2014, 35(12): 28-35.
 SI W J, ZHU T, ZHANG M Y. Dimension-reduction MUSIC for jointly estimating DOA and polarization using plane polarized arrays[J]. Journal on Communications, 2014, 35(12): 28-35(in Chinese).
- [14] HU X Y, LYU T T, ZHANG M, et al. MUSIC and improved MU-SIC algorithms for parameter estimation using a polarization sensitive array[C]//Proceedings of the 2021 IEEE 21st International Conference on Communication Technology. Piscataway: IEEE Press, 2021: 117-126.
- [15] 黄志强, 缪晨, 唐辉, 等. 基于联合 GA_MUSIC 算法的均匀圆阵 快速 DOA 估计[J]. 微波学报, 2021, 37(S1): 166-169.
 HUANG Z Q, MIAO C, TANG H, et al. Fast DOA estimation of uniform circular array based on joint GA_MUSIC algorithm[J].
 Journal of Microwaves, 2021, 37(S1): 166-169 (in Chinese).
- [16] JIN Y, HOU Y S, JIANG M. Frequency-DOA joint estimation by ant colony optimization[C]//Proceedings of the 2010 International Conference on Computer Application and Systems Modeling. Piscataway: IEEE Press, 2010: 314-319.
- [17] YANG Y P, HOU Y S, LIU X X. Two dimensional DOA estimation by ant colony optimization[C]//Proceedings of the 2010 International Conference on Intelligent System Design and Engineering Application. Piscataway: IEEE Press, 2010: 789-792.
- [18] 栾鹏程, 吴瑛. 改进遗传算法在 DOA 搜索中的应用[J]. 电光与控制, 2006, 13(3): 65-68.
 LUAN P C, WU Y. Application of improved genetic algorithm in DOA search algorithm[J]. Electronics Optics & Control, 2006, 13(3): 65-68(in Chinese).
- [19] 王炎.极化敏感阵列的二维 DOA 与极化参数估计算法研究[D]. 哈尔滨:哈尔滨工程大学, 2019: 13-20.
 WANG Y. Research on the estimation algorithms of two-dimensional DOA and polarization parameters[D]. Harbin: Harbin Engineering University, 2019: 13-20(in Chinese).
- [20] 张贤达. 信号处理中的线性代数[M]. 北京: 科学出版社, 1997: 27-31.

ZHANG X D. Linear algebra in signal processing[M]. Beijing: Science Press, 1997: 27-31(in Chinese).

- [21] SHI Y H, EBERHART R. A modified particle swarm optimizer[C]//Proceedings of the 1998 IEEE International Conference on Evolutionary Computation. Piscataway: IEEE Press, 1998: 69-73.
- [22] SHI Y H, EBERHART R C. Empirical study of particle swarm optimization[C]//Proceedings of the 1999 Congress on Evolutionary Computation. Piscataway: IEEE Press, 1999: 1945-1950.
- [23] 何庆,黄闽茗,王旭. 基于精英反向学习的逐维改进蜻蜓算法[J].
 南京师大学报 (自然科学版), 2019, 42(3): 65-72.
 HE Q, HUANG M M, WANG X. Elite opposition learning-based

1335

dimension by dimension improved dragonfly algorithm[J]. Journal of Nanjing Normal University(Natural Science Edition), 2019, 42(3): 65-72(in Chinese).

[24] 谢承旺, 王志杰, 夏学文. 应用档案精英学习和反向学习的多目标进化算法[J]. 计算机学报, 2017, 40(3): 757-772.
 XIE C W, WANG Z Z, XIA X W. Multi-objective evolutionary algorithm based on archive-elite learning and opposition-based learning[J]. Chinese Journal of Computers, 2017, 40(3): 757-772(in Chi-

nese).

- [25] 于建芳, 刘升, 王俊杰, 等. 融合莱维飞行与黄金正弦的蚁狮优化 算法[J]. 计算机应用研究, 2020, 37(8): 2349-2353.
 YU J F, LIU S, WANG J J, et al. Ant lion optimization algorithm integrating with Lévy flight and golden sine[J]. Application Research of Computers, 2020, 37(8): 2349-2353(in Chinese).
- [26] MIRJALILI S, LEWIS A. The whale optimization algorithm[J]. Advances in Engineering Software, 2016, 95(5): 51-67.

Joint estimation of DOA and polarization parameters based on uniform circle array with vector sensor

SHI Chunpeng¹, HE Huafeng^{1,*}, HE Yaomin¹, LI Zhen¹, YAN Shaoqiang², ZHANG Xiaoyu¹

(1. College of Missile Engineering, Rocket Force University of Engineer, Xi'an 710025, China;

2. College of Combat Support, Rocket Force University of Engineer, Xi'an 710025, China)

Abstract: To address the problem of multi-parameter joint estimation of incomplete electromagnetic vector sensors in the polarization sensitive array, a joint estimation method of Direction of Arrival (DOA) and polarization parameters is proposed based on a circle array model of the three-dimensional electromagnetic vector sensor plane. Using the characteristics of the array flow pattern matrix, a four-dimensional spectral function is decoupled, and the uncorrelation between DOA search and polarization parameter estimation is rigorously proved. A four-dimensional spectral search is optimized to a two-dimensional spectral search only related to DOA. A two-level comparison strategy of the DOA spectral peak search is proposed, which enables the radar to automatically identify the spectral peak coordinates and effectively improve the real-time performance of signal processing. The improved particle swarm optimization algorithm based on elite reverse learning and Lévy flight is proposed to estimate the polarization parameters, which improves the convergence performance of polarization parameters. The simulation results show that the proposed algorithm improves the estimation accuracy at the expense of a small amount of time.

Keywords: parameter estimation; direction of arrival search strategy; particle swarm optimization; elite opposition-based learning; Lévy flight; polarization sensitive array

Received: 2022-05-19; Accepted: 2022-08-19; Published Online: 2022-09-14 12: 54 URL: link.cnki.net/urlid/11.2625.V.20220913.1810.010

Foundation item: National Natural Science Foundation of China (61573366)

^{*} Corresponding author. E-mail: hhf0903@163.com
http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2022.0434

航空发动机润滑系统故障知识图谱构建及应用

吴闯1,2,张亮1,唐希浪1,*,崔利杰1,谢小月1

(1. 空军工程大学装备管理与无人机工程学院,西安710051; 2. 解放军 95478 部队,重庆 401329)

摘 要:由于航空发动机润滑系统结构功能复杂,基于现有的健康管理系统开展故障诊断存在可解释性不足及高度依赖专家经验的问题,提出一套面向航空发动机润滑系统的故障知识图谱构建方法。在结合专家知识设计润滑系统故障知识图谱本体概念的基础上,采用双向长短期记忆(BiLSTM)神经网络和条件随机场(CRF)等深度学习技术实现知识自主抽取,并基于余弦距离和Jaccard 相关系数法进行多源异构故障知识的融合。同时,基于构建的润滑系统故障知识图谱,实现润滑系统故障知识智能问答和故障归因分析应用。结果表明:知识图谱技术能够实现对润滑系统故障先验知识利用及故障原因解释,在智能故障诊断领域具有良好应用前景。

关键 词: 航空发动机; 润滑系统; 知识图谱; 深度学习; 知识问答

中图分类号: V239 文献标志码: A 文章编号: 1001-5965(2024)04-1336-11

航空发动机润滑系统(简称润滑系统)承担着 为航空发动机主轴轴承、机匣、传动齿轮等核心机 械装置提供润滑及冷却滑油的功能,是航空发动机 安全可靠工作的重要保障。当前,通过现有健康管 理系统对润滑系统开展状态参数监控,能够检测部 分故障问题并提供维修建议[1]。但由于润滑系统结 构功能复杂,其故障诊断结果的可解释性不足,仍 高度依赖专家知识和经验,需要维修人员对大量以 文本形式存在的维护操作规程、事故处置预案、排 故手册等资料进行反复学习、记忆和查询。一方 面,这种故障分析处理方式效率较低,不仅容易产 生"错、忘、漏",而且排故经验和能力难以实现共 享和传承,很难保证故障诊断的准确性和规范性; 另一方面,海量的故障先验知识却未得到有效管理 利用,造成数据资源浪费。因此,迫切需要一套可 解释的人工智能技术来提高润滑系统故障诊断效 率和故障知识管理水平。

知识图谱是近几年人工智能领域的研究热

点^[2],可以自主构建海量的"知识"网络,具备更强 大的语义表达、知识存储和推理能力,是新时代知 识工程发展的典型代表技术,是从感知智能到认知 智能的基石,能够为人们提供一种快速便捷的知识 检索和推理方式,已经广泛应用于智能问答、推 荐、决策等场景^[3]。

自 Google 公司在 2012 年提出知识图谱概念^[4] 至今,领域知识图谱的构建和应用越来越受到各行 各业的重视。在医疗诊断领域,侯梦薇等^[5]对医学 领域知识图谱构建技术展开了深入剖析,指出了当 前医学知识图谱研究存在的问题与挑战。在电网 故障处理领域,王骏东等^[6]提出了一种基于知识图 谱的配电网故障辅助决策方法,对知识图谱技术在 电网故障处理领域的应用进行了探索;Ou等^[7]针 对电力终端信息和故障信息等构建知识图谱,初步 实现了电力无线专网故障诊断和决策制定。在自 然灾害防治领域,杜志强等^[8]提出了自然灾害应急 知识图谱构建方法,实现了从多源数据到互联知识

收稿日期:2022-05-28; 录用日期:2022-08-26; 网络出版时间:2022-09-22 20:38 网络出版地址: link.cnki.net/urlid/11.2625.V.20220921.1519.004

基金项目:国家自然科学基金 (72201276);中国博士后科学基金 (2021M693941);西安市科协青年人才托举计划项目 (959202313098)

*通信作者. E-mail: tangxilang@sina.com

引用格式: 吴阔,张亮,唐希浪,等. 航空发动机润滑系统故障知识图谱构建及应用 [J]. 北京航空航天大学学报,2024,50(4):1336-1346. WUC, ZHANGL, TANGXL, et al. Construction and application of fault knowledge graph for aero-engine lubrication system [J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50(4): 1336-1346 (in Chinese).

的转化; 陶坤旺等¹⁹ 总结了面向一体化综合减灾的 知识图谱构建流程和关键技术, 可以实现对不同用 户的信息智能推送。

上述研究充分显示知识图谱技术在行业领域 具有显著的应用价值,目前将知识图谱技术引入 航空领域的研究仍处于起步阶段^[10]。韩涛等^[11]研 究了航空发动机故障领域文本中的实体自动抽取 方法,为知识图谱构建打下了一定基础;邱凌等^[12] 对比了航空制造知识图谱与通用知识图谱在构建 技术上的异同,提出了3个可行的应用方向及解 决方案;聂同攀等^[13]提出了面向飞机电源系统故 障诊断的知识图谱构建的关键技术,并开展智能 应用探索。上述研究为本文的知识图谱构建提供 了重要参考和依据。然而,目前的研究尚未系统 形成针对润滑系统故障诊断的知识图谱构建方法 和应用架构。因此,本文以航空维修保障需求为 牵引,采取知识驱动的故障诊断策略,提出一套面 向航空发动机润滑系统故障的知识图谱构建方 法,研究突破其中关键技术,并以某型国产航空发 动机近5年的润滑系统故障报告为原始语料进行 实例验证。

1 本文方法框架

知识图谱是一种用图模型来描述知识和客观 事物之间关联关系的技术手段,由节点和边组成^[14]。 根据知识覆盖范围可分为通用知识图谱和领域知 识图谱。面向润滑系统故障诊断的知识图谱属于 领域知识图谱,知识结构具有复杂性、多样性、异 构性的特点,对知识时效性和深层次的推理需求较 高,其构建需要充分考虑故障诊断的业务逻辑和语 义模型。本文方法框架如图1所示,包括基础数据 层、图谱构建层、知识推理层和图谱应用层,其中, BiLSTM 为 双 向 长 短 期 记 忆 (bi-directional long short-term memory), BERT(bidirection encoder representations from transformers)模型是一种基于 Transformer 架构的双向语言模型。



图 1 本文方法框架

Fig. 1 Framework of proposed method

其中,基础数据层主要包含润滑系统故障知识 的数据来源,以及各类结构化和非/半结构化的故障 知识,作为润滑系统故障实体和关系抽取的基础语 料。图谱构建层是通过人工和机器学习相结合的 方式从润滑系统故障文本中抽取故障知识,并进行 多源异构知识的融合,最终以三元组的结构化形式 储存在图数据库中。知识推理层是在构建的知识 图谱基础上,结合润滑系统故障诊断的业务场景、 维修经验和操作规则进行知识的逻辑关联挖掘,以 信息特征提取、相似度匹配、数据统计挖掘等方法 分析计算。图谱应用层是基于润滑系统故障诊断 应用需求与场景,实现功能模块到智能应用的转 变,在精细化场景中实现润滑系统故障知识的应用 探索。

2 关键技术

2.1 知识本体构建

知识本体是指与润滑系统故障知识内涵相关 的一组概念化的规范^[15]。知识图谱的模式层常通 过本体库进行管理,本体决定了知识图谱中的概念 节点和关联关系,是构建知识图谱的重要依据,相 当于知识库中的模具^[16]。

常用的本体构建方法有骨架法、七步法等,本 文采用成熟度较高的七步法构建知识本体模型^[17]。 其具体过程主要包括:①在故障知识本体构建之 前,与领域内专家和发动机保障人员共同探讨发动 机故障诊断的业务需求,明确构建本体的应用目 的;②查看其他领域(如医疗诊断、电网故障处理 等)是否有可复用的本体;③构建润滑系统故障领 域的概念体系,定义故障知识的核心概念;④借鉴 航空领域相关的国家标准、国家军用标准及行业标 准等对专业术语的定义,对故障相关概念进行规范 化处理;⑤定义相关概念的属性及取值范围;⑥收 集领域专家、基层保障人员对故障知识本体的评价 和改进意见;⑦回到第③步,对故障知识模型进行 改进和优化。通过多次迭代,最终构建一个具有实 用性、共通性的故障知识本体模型,如图2所示。



同时对关系设定了相应的定义域和值域,定义

域和值域的意义在于约束头尾实体,比如对于"失效原因"这个关系来说,其头实体只能是"故障模 式",尾实体则只能是"故障原因"。如果仅有头实 体而没有尾实体,则其关系不成立,反之亦然。

2.2 知识抽取

知识抽取是知识图谱构建过程中最核心的技术^[18],主要包括实体识别和关系抽取。

2.2.1 实体识别

润滑系统故障知识文本是大量实体的有机组 合,实体识别任务就是将文本中的具体实体按照 2.1节定义的本体类型进行识别和标注。润滑系统 故障实体识别是知识抽取模型的基础,实体识别的 具体效果将对知识图谱的构建质量起到至关重要 的作用。

对于结构化的润滑系统故障实体,如维修保障 支持系统数据表中单元构型,可以基于数据表将关 系型数据直接转换为"实体-关系-实体/属性"的三 元组。在对非/半结构化的润滑系统故障文本进行 实体识别时,本文采用一种基于 BiLSTM 神经网络 和条件随机场(conditional random field, CRF) 的识别方法^[19],同时通过加入 BERT 预训练模型^[20] 来实现实体识别任务,润滑系统故障实体识别模型 如图 3 所示,其中,B(begin)表示开始,C(check)表 示检查,E(end)表示结束,U(unit)表示组成单元, I(inside)表示内部,F(fault)表示故障,S(single)表示 单独实体,O(outside)表示实体外部。



Fig. 3 Model of lubrication system fault entity recognition

BERT 模型充分考虑字符之间、词语之间及句 子之间的关系特征,将原始的润滑系统故障文本转 换为向量表示,便于后续 BiLSTM 进行上下文特征 学习。BiLSTM 网络层用于自动提取句子的语义特 征,将文本中前后向的隐态结果进行结合,能够更

图 2 模式层本体设计 Fig. 2 Pattern layer ontology design

好地获取较长距离的双向语义依赖关系,可以有效 解决润滑系统故障文本中实体距离较远的依赖问 题。CRF 推理层是一种条件概率分布模型,用于解 决润滑系统故障文本序列化标注问题,将 BiLSTM 网络层的输出传递给 CRF 作为输入,选择出最大似 然概率的标注序列,为最后生成的预测标签添加 约束。

在训练模型前需要标注一定数量的文本对模型进行训练,标注方法采用的是 BIOES(begin-insideoutside-end-single)序列标注法^[21],按 8:1:1 的比例^[18] 划分标注语料为训练集、验证集和测试集,图 4 为 润滑系统故障语料实体标注的一个示例。期间不 断调整模型参数,确定模型最优性能,通过对训练 集上的数据进行参数训练,得到最优参数:词向量 维度为 300, Dropout 值为 0.5, 一次网络嵌入的数据 大小(batch-size)值为 16, 代数(epoch)为 50, 学习率 为 0.001, 优化算法使用 Adam。



图 4 实体标注示例 Fig. 4 Entity tagging example

例如,"裂纹"实体类型为故障模式 (failure mode, FM),相应的其第一个字符"裂"对应的标签为"B-FM",其最后一个字符"纹"对应的标签为"E-FM"。 2.2.2 关系抽取

润滑系统故障相关知识文本中的实体之间存 在着一定的关联关系,如何从非结构化润滑系统故 障知识文本中抽取其蕴含的关系是知识抽取的难 点。基于深度学习的关系抽取方法在提高关系分 类效率和关系抽取准确率方面具有很大优势^[22],因 此,本文采用基于自注意力机制的 BiLSTM 神经网 络模型来实现关系抽取,如图 5 所示,其中,*H*=(*h*₁, *h*₂,…,*h*_n)^T 为特征矩阵,*e*_n 为嵌入层的向量,*w*_n 为输 入的文字。

嵌入层用于对输入的润滑系统故障语句进行 向量化表示,得到语句的输入特征向量。BiLSTM 层利用双向神经网络模型,从输入特征向量中学习 故障文本句子中上下文信息和浅层语义特征,得到 更高级别的词向量特征。自注意力层则利用自注 意力机制,计算每一个单词的权重,从句子向量中 学习文本语句的深层次全局语义特征,得到句子的 全局特征向量结果,该结果能够反映2个词语在句 子中的关系。输出层则用于拼接全局特征向量和 局部特征向量,计算得到实体之间的关系向量。

基于自注意力机制的关系抽取模型中,自注意



图 5 润滑系统故障关系抽取模型

Fig. 5 Model of lubrication system fault relationship extraction

力权重向量 a 的计算和实体关系概率分布 p 的计 算是实现关系抽取的关键, 其计算式为

 $\boldsymbol{a} = \operatorname{softmax} \left(\boldsymbol{v}_{\operatorname{att}} \cdot \left[\tanh(\boldsymbol{H}) \right]^{\mathrm{T}} \right)$ (1)

$$\boldsymbol{p} = \operatorname{softmax} \left(\boldsymbol{W} \cdot [\tanh\left(\boldsymbol{a} \cdot \boldsymbol{H}\right)]^{\mathrm{T}} + \boldsymbol{b}^{\mathrm{T}} \right)$$
(2)

$$\boldsymbol{H} = (\boldsymbol{h}_1, \boldsymbol{h}_2, \cdots, \boldsymbol{h}_n)^{\mathrm{T}}$$
(3)

式中: v_{att}为维度为2u的权重向量, u为 BiLSTM 隐藏 层的大小; H为经过 BiLSTM 编码后得到的特征矩 阵, 维度为n×2u; W为表示词语重要性的权重矩 阵, 维度为c×2u, c为需要输出的关系类型数, 根据 2.1 节定义的关系类型c = 12; b为维度为n的偏置参 数向量。tanh(·)为双曲正切激活函数, softmax(·)为 归一化指数函数, 其表达式分别为

$$\tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}$$
(4)

softmax
$$(f_y) = \frac{\exp(f_y)}{\sum_{c=1}^{c} \exp(f_c)}$$
 (5)

2.3 知识融合

按照 2.2 节所述, 从非结构化的润滑系统故障 知识文本中抽取的知识可能存在大量模糊、重复的 数据, 知识融合的目的就是将其进行有效融合统 一, 提高知识图谱数据库的知识质量^[23]。知识融合 任务主要包括实体消歧和共指消解^[24]。

实体消歧技术用于处理同名实体指代不同客 观事物的问题,例如,"油滤"在有些文本中描述的 是"滑油系统的油滤",有的则描述的是"燃油系统 的油滤",因此,需要联系上下文的语义,明确同名 实体的正确含义。共指消解技术则用于处理多种 表达方式对应同一实体对象的问题,例如,"发 附"、"发动机附件机匣"、"附件机匣"均对应的 是"发动机附件机匣"这一个单元实体,尤其是人 工撰写的故障报告、排故经验等,用语不规范的现 象比较普遍,因此,需要统一规范的实体名称。本 文采用余弦距离和 Jaccard 相关系数相结合的方式 计算润滑系统故障实体之间的相似度,通过设定阀 值的方式判断待对齐实体是否匹配,实现知识融 合。其对应计算式为

$$D_{\cos}(s_1, s_2) = \frac{|A(s_1) \cap A(s_2)|}{\sqrt{|A(s_1)||A(s_2)|}}$$
(6)

$$D_{\text{Jaccard}}(s_1, s_2) = \frac{|A(s_1) \cap A(s_2)|}{|A(s_1) \cup A(s_2)|}$$
(7)

式中: s1和 s2分别为 2个润滑系统故障实体; A(s)为 实体 s的属性字符串。D值越大, 就代表两者的语 义相似度就越高, 表 1列出了一些实际的润滑系统 故障知识表述。从计算结果可知, 对于"油液分 析"和"油液化验" 2个不同表述的实体, 其 Jaccard 相似度为 0.33, 基于余弦距离的相似度达到 0.50, 实际上两者为同一概念, 应融合为同一实体。因 此, 可以通过设定相似度阀值, 结合 Jaccard 相关系 数和余弦距离来综合判断是否融合为同一实体。

表1 实体相似度计算实例

 Table 1
 Examples of entity similarity computation

实体1	实体2	$D_{\rm cosine}$	$D_{ m Jaccard}$
附件机匣	发附机匣	0.75	0.6
油液分析	油液化验	0.50	0.33
振动值超标	振动值异常	0.60	0.42
滑油泵组	滑油泵	0.87	0.75
滑油压力大	滑油压差大	0.80	0.66
金属屑信号器	金属末信号器	0.83	0.71

2.4 知识存储

通过知识抽取和知识融合等流程,已将多源异 构的润滑系统故障数据转换为结构化的知识,知识 存储则需要将各类知识以"实体-关系-实体/属性" 的三元组形式储存起来,用于支持大规模图数据的 有效管理和计算。

针对润滑系统故障知识结构明确、实体之间关 系丰富的特点,本文采用 ArangoDB 图数据库作为 存储系统,可以从多个维度展示润滑系统故障知识 图谱。ArangoDB 是一个开源的分布式原生多模型 图数据库,兼容图(graph)、文档(docdument)和键/ 值对(key/value)3种数据模型,可以采用统一的 ArangoDB 查询语言(ArangoDB query language, AQL) 查询多种数据模型,并允许在单词查询中混合使用 多种模型。与其他图数据库相比,在速度和性能方 面都具有极大优势,能够满足高性能领域需求,适 用于搭建高性能的应用程序。

2.5 基于知识图谱的智能应用

2.5.1 基于知识图谱的智能问答

基于知识图谱的问答系统(knowledge graph

based question answering, KBQA)是通过对用户输入 的问题进行语义解析和意图识别,将其转化为电脑 查询语言,在构建的知识图谱中寻求答案的一种问 答机制^[18]。KBQA架构包括问题分类、知识检索和 答案生成等技术,其中,问题分类是实现 KBQA 的 关键,问题分类的核心任务是识别用户所提问题的 意图,为构建答案生成策略提供有用的信息。

朴素贝叶斯分类器是一种基于概率统计的分 类模型,目前被广泛应用于电网、医疗、地铁等行 业智能问答技术的问题分类中^[25]。通过对用户的 问题分词提取特征词,从而实现问题分类,对应到 相应的问答模板中。本文采用这一分类方法通过 Jieba 分词后的问题特征词及自定义的问题类别作 为训练样本构建分类模型,实现对问题的意图识 别,以便将所得到的分类结果对应到相应的问答模 板中。具体地,令*x*=[*x*₁,*x*₂,…,*x_n*]表示问题分词后 的*n*个特征词向量,*C*={*C*₁,*C*₂,…,*C_m*}表示自定义的 *m*个类别。在每个类下特征词变量之间相互独立 的假设下,根据最大后验概率,朴素贝叶斯分类将 选择使后验概率最大的类作为对未知问题特征词 向量*x*的类标签,即

$$C(\mathbf{x}) = \underset{c_{j} \in C}{\operatorname{argmax}} P(c_{j}) \prod_{i=1}^{n} P(x_{i} | c_{j})$$
(8)

式中: *P*(*c_j*) 为类别 *c_j* 的先验概率, 可以通过样本中 *c_j* 类出现的频率来计算; *P*(*x_i*|*c_j*) 为在 *c_j* 类中 *x_i* 发生 的概率, *j*=1, 2, …, *m*。

关于 $P(x_i|c_j)$ 的计算, 一般采用词频来计算, 即 $P(x_i|c_j) = \frac{N_{x_i,c_j}}{N_c}$ (9)

式中: N_{x_i,c_j} 为 x_j 在 c_j 类的词频数; N_{c_j} 为 c_j 类所有词的词频数。

可知,式(9)的分子部分对整个计算结果影响 很大。同时,在进行分类时发现一些核心词,这些 核心词是一种特殊的关键词,其往往对某些分类具 有代表性,出现次数可能不多,但是对问题分类影 响很大。如果对于核心词 x,在计算 P(x,lc,)时仍采 用式(9)的话就会造成模型分类精度不高的结果。 为解决这一问题,需要修正这一计算式来提高核心 词的影响力。本文采用加权的方法对其进行了修 正,修正的计算式为

$$P^{*}(x_{i}|c_{j}) = \begin{cases} \frac{kN_{x_{i},c_{j}}}{N_{c_{j}}} & x_{i} \in c_{j} \\ \frac{N_{x_{i},c_{j}}}{N_{c_{j}}} & x_{i} \notin c_{j} \end{cases}$$
(10)

式中: k 为核心词影响问题分类的重要度, k>1。关 于问题类别核心词可以采用类别信息关键词和专 家经验的方式来选择。表 2 为本文针对不同问题 的类型所选择的一些特征核心词。

 Table 2
 Extraordinary key-word

特征核心词

表 2

问题分类	特征核心词
原因类	起因、导致、条件、故障树
事实类	组成、连接、是什么、系统、属性
统计类	有多少、集合、统计、总计
解决方案类	怎么办、处置、拆装、监控
预防措施类	提前、应急、保养、预先

2.5.2 基于知识图谱的故障归因分析

润滑系统部件数量繁多且关系密切,一个插头 未拧紧都有可能产生"蝴蝶效应",导致重大安全 事故发生,因此,查找故障根因对于润滑系统故障 诊断工作至关重要。基于知识图谱的故障归因分 析可以依据图谱中失效原因关系快速推理出可能 的故障原因,为维修人员提供排故辅助。

故障归因分析其实是知识问答的应用之一,属 于寻求原因类问题的答案,其更关注于对知识图谱 中故障原因关系的(一/多)跳链条查询,以便于深层 次地挖掘故障发生的根本原因。使用如图 6 所示 的故障现象 A 的故障根因挖掘为例,来介绍实现根 因分析的关键技术,其中, M_c(*i*,*j*)表示第 *i* 层故障原 因集合中的第 *j* 个故障原因。基于马尔可夫过程的 思想,假设某一故障原因的发生仅和与其有因果关 系的故障原因有关。因此,对于某一故障原因 M_{*j*}, 于是可以基 于 M_i 推算出 M_i 发生的概率,即

$$P\left(M_{j}|M_{i}\right) = \frac{n_{M_{i}M_{j}}}{n_{M_{i}}} \tag{11}$$

式中: n_M,为故障原因 M_i在历史故障数据中出现的 频数,可以通过图谱中关系节点的属性(次数)来统 计; n_{M,M},为故障原因 M_i导致 M_i发生的次数统计。

因此,可以通过上述这一原理对于图6中所显



示的各层故障原因发生的概率进行分析。例如,通 过第 *i* 层故障原因集合*M*_{ci}中的故障原因就可以推 测出第 *i*+1 层故障原因集合*M*_{ci}中的故障原因发生 的概率 *P*(*M*_{ci}|*M*_{ci})。于是,故障现象 *A* 经由知识图 谱中故障原因的单跳链条查询得到顶层故障原因 *M*_{ci},然后通过链式法则得到故障现象 *A* 下某故障 根因*M*_c(*i*,*i*)发生的概率为

$$P(M_{c}(i, j)|A) = P(M_{c}(i, j)|M_{c}(i-1, k)) \times P(M_{c}(i-1, k)|M_{c}(i-2, r)) \times \dots \times P(M_{c}(1, k)|A)$$
(12)

式中: *M_c(i-1,k)* 为故障原因*M_c(i,j)* 的上一层原因集 合中对应的故障原因,其他故障原因以此类推。基 于式 (12), 计算故障现象 *A* 发生时, 是由故障根因 *M_c*(3,1) 导致的概率为

$$P(M_{c}(3,1)|A) = P(M_{c}(3,1)|M_{c}(2,2)) \times P(M_{c}(2,2)|M_{c}(1,2)) \times P(M_{c}(1,2)|A)$$
(13)

通过采取式(12)计算方法,可以计算出导致某 一故障现象的每个根因的概率,以便于对这一故障 现象有一更清晰的认识。此外,还根据故障根因的 概率值大小制定排故的优先顺序。

3 润滑系统故障知识图谱构建实例 及应用

3.1 润滑系统故障数据介绍

本文以某型国产航空发动机润滑系统排故手 册、近5年故障报告和维修管理软件中的构型数据 为基础语料,利用本文方法构建润滑系统故障知识 图谱,所有数据均已做脱密处理。表3为其中典型 的非结构化语料示例。简要情况包含该故障报告 的故障时间、故障单位、故障模式和故障影响;排 查情况包含排查方法和故障原因;装备情况包含单 元的基本信息、属性;维修情况包含该故障模式相 应的维修措施。

表 3 润滑系统故障报告语料示例

Table 3 Corpus example of lubrication system failure report

情况	语料
简要情况	2022年X月X日,XX单位在执行飞行任务时,空中报发动 机"滑油金属含量超标"告警,飞行任务停止,飞行员降 低发动机转速后安全着陆。
排查情况	①着陆后,进行发动机滑油光谱分析,发现铁元素浓度值 超标;②进一步孔探该发附件机匣内齿轮齿面异常磨损; ③检查上次柔轴安装记录发现安装尺寸不够标准。
装备情况	发动机:型号XX;号码XXX;规定寿命XX;总使用时间 XX小时XX分;制造厂XXX。 附件机匣:型号XX;号码XXX;规定寿命XX;总使用时间 XX小时XX分;制造厂XXX。
维修措施	①更换附件机匣;②故障的附件机匣返厂检查修理;③更 换机匣后,清洗润滑系统,更换滑油3次。

对润滑系统故障报告中的数据进行预处理。 将2个实体单独整理成一个独立的语句,具体如下。

故障案例 "2022 年 X 月 X 日...安全着陆"关联的故障模式为滑油金属含量超标告警。

故障案例 "2022 年 X 月 X 日...安全着陆"的故障时间是 2022 年 X 月 X 日。

故障案例 "2022 年 X 月 X 日...安全着陆"的故 障单位是 XX 单位。

滑油金属含量超标告警的原因是铁元素浓度 值超标;铁元素浓度值超标是附件机匣齿面异常磨 损引起;附件机匣齿面异常磨损是柔轴安装尺寸不 标准引起。

附件机匣齿面异常磨损的维修措施是更换附 件机匣,清洗更换滑油。

装备情况可作为故障单元属性进行抽取。

3.2 知识抽取结果

在利用模型抽取知识前,需要标注一定的数据 训练模型,部分故障文本标注可视化效果如图7所示。



图 7 故障文本标注可视化

Fig. 7 Visualization of failure text annotation

本文共选取 600 篇文档进行人工标注,对于实体识别和关系抽取模型均按照 8:1:1 的比例^[18] 划分标注语料为训练集、验证集和测试集。抽取结果 准确性的评价,则采用 *F*₁分数进行评估,其计算 式为

$$F_1 = \frac{2PR}{P+R} \tag{14}$$

式中: P 为精确度, 表示识别正确的实体数和识别 实体数之比; R 为召回率, 表示识别正确实体数和 总实体数之比。图 8 为本文所建立的知识抽取模 型的整体评估结果, F₁值大于 80%, 基本满足领域 实体识别的一致性要求^[26], 可用于润滑系统故障文 本的知识抽取任务。

利用本文构建的知识抽取模型共抽取故障模 式、故障原因、部件、维修措施等润滑系统故障实 体 1712 个(见表 4),抽取其对应的失效原因、组 成、发现等关系 2654 条(见表 5)。



图 8 抽取模型整体评估

Fig. 8 General comment of extraction model

表 4 润滑系统故障知识图谱实体数量

 Table 4
 Entity statistic numbers of lubrication

system fault knowledge graph

故障模式	故障原因	排查方法	维修措施	预防措施	系统
295	331	240	149	93	3
故障案例	故障影响	故障时间	故障单位	故障单元	
262	7	169	20	143	

表 5	润滑系统故障知识图谱关系
Table 5	Relationship statistics of lubrication

system fault knowledge graph						
失效原因	发生时间	采取	引起	预防	造成	
459	225	212	253	15	214	
发生单位	组成	发现	关联	检查	发生	
222	198	257	188	116	395	

3.3 知识图谱构建结果

把故障文本中的实体作为知识图谱的节点,实体之间的关联关系作为连接图谱节点的边,存储在 ArangoDB 图数据库中,构成润滑系统故障知识图 谱。图 9 为部分润滑系统故障知识图谱。由图可 知,以润滑系统为中心点,展示了润滑系统组成结 构下的各子部件,以及与其相关联的故障模式、失 效原因等信息。通过"发生"、"失效原因"等关系 将故障模式和部件连接、将故障模式和故障原因连 接,构建了润滑系统不同实体间的相关关系,将整 个系统的故障知识连接起来,形成一个较为完整的 润滑系统知识体系,为知识图谱应用中知识查询和 调用提供可靠的数据支撑。

3.4 基于润滑系统故障知识图谱的智能应用

3.4.1 知识智能问答

润滑系统故障知识图谱通过构建润滑系统故障模式、故障原因、排查方法、维修措施等信息之间的全方位关联,借助知识图谱强大的语义网络,能够智能理解维修人员输入的自然语言问题,并返回正确答案,如图 10 所示。

在"知识问答"应用模块中查询滑油泵组的故 障模式,依据故障单元——滑油泵组和故障模式之

个

攵





图 9 润滑系统故障知识图谱可视化 Fig. 9 Visualization of lubrication system fault knowledge graph



图 10 基于润滑系统故障知识图谱的知识问答 Fig. 10 Knowledge Q&A based on lubrication system fault knowledge graph

间的关系,与知识图谱的图数据相匹配,以自然语言的形式返回滑油泵组的5种故障模式,并且可以深度溯源与答案有关的案例文件,为答案结果提供可解释性。系统还可以依据实体的相似度进行相关问题推荐,全方位追求答案的完整性。

同时,润滑系统故障知识图谱可以实现故障知 识更新功能。当上传新的故障文本后,可以深度挖 掘出新故障模式、新原因、新的处置方式和预防措 施,从而更新润滑系统故障知识图谱,为其他单位 故障诊断和处置提供借鉴,实现故障知识的可靠管 理和传承积淀。

3.4.2 故障归因分析

故障归因分析可以根据故障发生时间统计相 应故障案例中故障原因的频数,并结合故障影响级 别对可能的故障原因进行排序,向用户推荐,如图 11 所示,输入"滑油中金属含量超标"故障描述后,可 以通过故障原因解析,结合故障案例中故障原因发 生的频数统计,计算出 5 种第 1 层故障原因的概率 值,并进行排序。通过现场排查,补充故障失效信 息,进而提高"柔轴安装尺寸不标准"这个故障根 因的概率,并可以用图谱可视化形式展示,向用户 反馈最优的排故顺序。





4 结 论

基于 BERT+BiLSTM+CRF 的实体识别模型
 和自注意力机制的关系抽取模型,能够实现对润滑
 系统故障文本的自主化知识抽取。

2)利用本文提出的框架和技术,构建了一定数据规模的润滑系统故障知识图谱,可以为智能问答、归因分析等应用提供知识支撑,一定程度上为故障诊断提供了可解释性依据。

3)应用润滑系统故障知识图谱实现了非结构 化故障知识的结构化存储,可以有效整合领域知 识、汇聚专家经验,有利于提高历史故障知识利用 率和故障诊断效率,降低对技术骨干经验的依赖。

随着人工智能、大数据、云计算等新技术的继续发展,知识图谱在航空维修管理、质量控制、安全监察等环节具有广阔应用空间。后续将进一步完善知识抽取算法,扩大知识图谱的覆盖范围,开展面向图像、音视频、传感器等智能终端的知识工程探索。

参考文献(References)

[1] 程荣辉,张志舒,陈仲光.第四代战斗机动力技术特征和实现途径[J].航空学报,2019,40(3):022698.

CHENG R H, ZHANG Z S, CHEN Z G. Technical characteristics and implementation of the fourth-generation jet fighter engines[J]. Acta Aeronautica et Astronautica Sinica, 2019, 40(3): 022698 (in Chinese).

[2] 于皓, 张杰, 吴明辉, 等. 领域知识图谱快速构建和应用框架[J]. 智能系统学报, 2021, 16(5): 871-884.

YU H, ZHANG J, WU M H, et al. A framework for rapid construc-

tion and application of domain knowledge graphs[J]. CAAI Transactions on Intelligent Systems, 2021, 16(5): 871-884 (in Chinese).

- [3] 马忠贵, 倪润宇, 余开航. 知识图谱的最新进展、关键技术和挑战
 [J]. 工程科学学报, 2020, 42(10): 1254-1266.
 MA Z G, NI R Y, YU K H. Recent advances, key techniques and future challenges of knowledge graph[J]. Chinese Journal of Engineering, 2020, 42(10): 1254-1266 (in Chinese).
- [4] AMIT S. Introducing the knowledge graph: Things, not strings[EB/ OL]. (2012-05-16)[2022-05-27]. http://www.blog.goole/products/ search/introducing-knowledge-graph-things-not/.
- [5] 侯梦薇, 卫荣, 陆亮, 等. 知识图谱研究综述及其在医疗领域的应用[J]. 计算机研究与发展, 2018, 55(12): 2587-2599.
 HOU M W, WEI R, LU L, et al. Research review of knowledge graph and its application in medical domain[J]. Journal of Computer Research and Development, 2018, 55(12): 2587-2599 (in Chinese).
- [6] 王骏东,杨军,裴洋舟,等.基于知识图谱的配电网故障辅助决策 研究[J]. 电网技术, 2021, 45(6): 2101-2112.
 WANG J D, YANG J, PEI Y Z, et al. Distribution network fault assistant decision-making based on knowledge graph[J]. Power System Technology, 2021, 45(6): 2101-2112 (in Chinese).
- [7] OU Q H, ZHENG W J, QI W W, et al. Research on the construction method of knowledge graph for electric power wireless private network[C]//Proceedings of the 2020 IEEE 10th International Conference on Electronics Information and Emergency Communication. Piscataway: IEEE Press, 2020: 10-13.
- [8] 杜志强,李钰,张叶廷,等. 自然灾害应急知识图谱构建方法研究
 [J]. 武汉大学学报 (信息科学版), 2020, 45(9): 1344-1355.
 DU Z Q, LI Y, ZHANG Y T, et al. Knowledge graph construction method on natural disaster emergency[J]. Geomatics and Information Science of Wuhan University, 2020, 45(9): 1344-1355 (in Chinese).
- [9] 陶坤旺, 赵阳阳, 朱鹏, 等. 面向一体化综合减灾的知识图谱构建 方法[J]. 武汉大学学报 (信息科学版), 2020, 45(8): 1296-1302.

TAO K W, ZHAO Y Y, ZHU P, et al. Knowledge graph construction for integrated disaster reduction[J]. Geomatics and Information Science of Wuhan University, 2020, 45(8): 1296-1302 (in Chinese).

[10] 曹明, 王鹏, 左洪福, 等. 民用航空发动机故障诊断与健康管理: 现状、挑战与机遇-地面综合诊断、寿命管理和智能维护维修决 策[J]. 航空学报, 2022, 43(9): 625574.

CAO M, WANG P, ZUO H F, et al. Civil aero-engine diagnostics & health management: Current status, challenges and opportunitiescomprehensive off-board diagnosis, life management & intelligent condition based MRO[J]. Acta Aeronautica et Astronautica Sinica, 2022, 43(9): 625574 (in Chinese).

- [11] 韩涛,黄海松,姚立国.面向航空发动机故障知识图谱构建的实体抽取[J].组合机床与自动化加工技术,2021(10):69-73.
 HAN T, HUANG H S, YAO L G. Entity extraction for aero-engine fault knowledge graph[J]. Modular Machine Tool & Automatic Manufacturing Technique, 2021(10): 69-73 (in Chinese).
- [12] 邱凌,张安思,李少波,等. 航空制造知识图谱构建研究综述[J]. 计算机应用研究, 2022, 39(4): 968-977.
 QIU L, ZHANG A S, LI S B, et al. Advances in building knowledge graphs for aerospace manufacturing[J]. Application Research of Computers, 2022, 39(4): 968-977 (in Chinese).
- [13] 聂同攀,曾继炎,程玉杰,等.面向飞机电源系统故障诊断的知识 图谱构建技术及应用[J].航空学报,2022,43(8):625499.
 NIE T P, ZENG J Y, CHENG Y J, et al. Knowledge graph construction technology and its application in aircraft power system fault diagnosis[J]. Acta Aeronautica et Astronautica Sinica, 2022, 43(8):625499 (in Chinese).
- [14] 王昊奋,漆桂林,陈华钧.知识图谱方法、实践与应用[M].北京:
 电子工业出版社, 2019: 425-427.
 WANG H F, QI G L, CHEN H J. Knowledge graph method, prac-

tice and application[M]. Beijing: Publishing House of Electronics Industry Press, 2019: 425-427 (in Chinese).

- [15] 陈永当,杨海成,杜兵劳.基于本体的航空发动机设计知识组织 模型构建与分析[J].航空动力学报,2007,22(1):90-95.
 CHEN Y D, YANG H C, DU B L. Building and analysis of ontology-based knowledge organization model for aircraft engine design[J]. Journal of Aerospace Power, 2007, 22(1): 90-95 (in Chinese).
- [16] 朱庆,王所智,丁雨淋,等.铁路隧道钻爆法施工智能管理的安全 质量进度知识图谱构建方法[J]. 武汉大学学报 (信息科学版), 2022, 47(8): 1155-1164.

ZHU Q, WANG S Z, DING Y L, et al. Construction method of "safety-quality-schedule" knowledge graph for intelligent management of drilling and blasting construction of railway tunnels[J]. Geomatics and Information Science of Wuhan University, 2022, 47(8): 1155-1164 (in Chinese).

[17] 李震, 刘斌, 苗虹, 等. 基于本体的软件安全性需求建模和验证[J]. 北京航空航天大学学报, 2012, 38(11): 1445-1449.

LI Z, LIU B, MIAO H, et al. Modeling and verification of software safety requirement based on ontology[J]. Journal of Beijing Uni-

versity of Aeronautics and Astronautics, 2012, 38(11): 1445-1449 (in Chinese).

- [18] 刘峤, 李杨, 段宏, 等. 知识图谱构建技术综述[J]. 计算机研究与发展, 2016, 53(3): 582-600.
 LIU Q, LI Y, DUAN H, et al. Knowledge graph construction techniques[J]. Journal of Computer Research and Development, 2016, 53(3): 582-600 (in Chinese).
- [19] XU K, ZHOU Z F, HAO T Y, et al. A bidirectional LSTM and conditional random fields approach to medical named entity recognition [C]//Proceedings of the International Conference on Advanced Intelligent Systems and Informatics. Berlin: Springer, 2017: 355-365.
- [20] DEVLIN J, CHANG M W, LEE K, et al. Bert: pre-training of deep bidirectional transformers for language understanding[C]//Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Minneapolis: Association for Computational Linguistics, 2019: 4171-4186.
- [21] 吴赛赛, 周爱莲, 谢能付, 等. 基于深度学习的作物病虫害可视化知识图谱构建[J]. 农业工程学报, 2020, 36(24): 177-185.
 WU S S, ZHOU A L, XIE N F, et al. Construction of visualization domain-specific knowledge graph of crop diseases and pests based on deep learning[J]. Transactions of the Chinese Society of Agricultural Engineering, 2020, 36(24): 177-185 (in Chinese).
- [22] 罗欣, 陈艳阳, 耿昊天, 等. 基于深度强化学习的文本实体关系抽取方法[J]. 电子科技大学学报, 2022, 51(1): 91-99.
 LUO X, CHEN Y Y, GENG H T, et al. Entity relationship extraction from text data based on deep reinforcement learning[J]. Journal of University of Electronic Science and Technology of China, 2022, 51(1): 91-99 (in Chinese).
- [23] 赵晓娟, 贾焰, 李爱平, 等. 多源知识融合技术研究综述[J]. 云南 大学学报 (自然科学版), 2020, 42(3): 459-473.
 ZHAO X J, JIA Y, LI A P, et al. A survey of the research on multisource knowledge fusion technology[J]. Journal of Yunnan University (Natural Sciences Edition), 2020, 42(3): 459-473 (in Chinese).
- [24] 乔骥, 王新迎, 闵睿, 等. 面向电网调度故障处理的知识图谱框架 与关键技术初探[J]. 中国电机工程学报, 2020, 40(18): 5837-5849.
 QIAO J, WANG X Y, MIN R, et al. Framework and key technologies of knowledge-graph-based fault handling system in power grid[J]. Proceedings of the CSEE, 2020, 40(18): 5837-5849 (in Chinese).
- [25] 董丽叶. 基于知识图谱的 S 市地铁机电设备故障处理优化研究
 [D]. 石家庄: 河北科技大学, 2020: 43-48.
 DONG L Y. Fault treatment optimization of subway electromechanical equipment in S city based on knowledge graph[D]. Shijiazhuang: Hebei University of Science and Technology, 2020: 43-48 (in Chinese).
- [26] ARTSTEIN R, POESIO M. Inter-coder agreement for computational linguistics[J]. Computational Linguistics, 2008, 34(4): 555-596.

Construction and application of fault knowledge graph for aero-engine lubrication system

WU Chuang^{1, 2}, ZHANG Liang¹, TANG Xilang^{1, *}, CUI Lijie¹, XIE Xiaoyue¹

College of Equipment Management and Unmanned Aerial Vehicle Engineering, Air Force Engineering University, Xi'an 710051, China;
 2. 95478 of the PLA, Chongqing 401329, China)

Abstract: Due to the complex structure and function of aero-engine lubrication system, the existing health management system lacks sufficient interpretability and relies heavily on expert experience for fault diagnosis A method for constructing aero-engine lubrication system fault knowledge graph was proposed in this paper. By integrating expert knowledge, the concept of lubrication system fault knowledge graph ontology was designed. With the help of deep learning techniques such as bi-directional long short-term memory (BiLSTM) and conditional random field (CRF), we achieved the automatic extraction of unstructured knowledge. Next, based on the Cosine Distance and Jaccard coefficient, multi-source heterogeneous fault knowledge fusion was realized. In the end, incorporating the constructed areo-engine lubrication system fault knowledge graph, we achieved intelligent question and answer capabilities for lubrication system fault knowledge. The application results show that the knowledge graph technology can realize the utilization of prior knowledge of lubrication system faults and the explanation of fault causes, and has a good application prospect in the field of intelligent fault diagnosis.

Keywords: aero-engine; lubrication system; knowledge graph; deep learning; knowledge questions and answers

Received: 2022-05-28; Accepted: 2022-08-26; Published Online: 2022-09-22 20:38 URL: link.cnki.net/urlid/11.2625.V.20220921.1519.004

Foundation items: National Natural Science Foundation of China (72201276); China Postdoctoral Science Foundation (2021M693941); Xi'an Association for Science and Technology Youth Talent Support Program (959202313098)

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn **DOI**: 10.13700/j.bh.1001-5965.2022.0446

考虑推力意图的航空器连续爬升垂直剖面预测

杜卓铭1,张军峰1,*,苗洪连1,王斌2

(1. 南京航空航天大学 民航学院,南京 210016; 2. 中国民用航空中南地区空中交通管理局,广州 510403)

摘 要:航空器连续爬升剖面的准确预测可以提高冲突探测的可靠性和离场排序的精确 性。基于推力设置信息,提出航空器推力意图的建模方法。利用航空器的全能量方程,考虑风速向 量与温度信息,提出考虑推力意图的航空器连续爬升垂直剖面的预测方法。利用快速存取记录器 (QAR)数据,进行多案例的对比分析,以 QAR 每一数据帧为采样点,重点对预测剖面的真空速、 高度和燃油流量等进行误差分析。考察不同预测方法,对比到达爬升顶点 (TOC)的时间和距离的误 差。结果表明:采用所提预测方法可以将到达 TOC 时间平均绝对误差控制在 1 min 内;与不考虑推 力意图的预测方法相比,可以降低到达 TOC 时间平均绝对误差约 52%。

关键词:空中交通管理;航迹预测;连续爬升;性能模型;推力意图
 中图分类号:V352
 文献标志码:A
 文章编号:1001-5965(2024)04-1347-07

随着民航运输业高速发展,空中交通流量日益 增长,导致空域拥堵与航班延误问题亟待解决。航 迹的准确预测有利于提高管制员对空域态势的掌 控能力,助力管制决策支持、降低管制工作负荷、 提升安全运行水平、实现"双碳目标"^[1]。

航迹预测主要分为模型驱动与数据驱动2类 方法^[2]。前者主要通过航空器的运动学与动力学模 型,利用航空器性能参数,结合航空器意图与气象 信息,实现四维航迹的预测^[3]。后者通过构建输入 输出空间,利用回归方法实现航空器飞行时间的预 测^[4-5]。模型驱动方法的优势在于状态信息可具体 获知,数据驱动方法的优势在于影响因素可量化 考虑。

航迹预测对象往往根据运行场景可以分为:场 面滑行预测、离场航迹预测、航路飞行预测与进场 飞行预测。准确的离场航空器的垂直剖面预测,可 有效实现终端空域冲突探测与解脱^[6]、离场航班排 序与调度^[7-8]、基于时间的流量管理^[9]等。鉴于此, 本文聚焦离场航迹预测。

相较于进场运行,离场航空器往往沿着公布的标准仪表离场程序实施离场运行,因此,其航迹预测的重点与难点集中于垂直剖面预测。张军峰等^[10]提出了基于连续动态模型与离散动态模型的航迹预测方法,并通过航空器意图、初始状态、性能参数及环境信息等因素,降低离场四维航迹预测的不确定性。Sun等^[11-12]使用了广播式自动相关监视(automatic dependent surveillance-broadcast, ADS-B)数据估计航空器性能参数,基于贝叶斯推理方法对飞机初始质量进行了估计,提升离场航迹预测的准确性。Alligier等^[13-15]使用机器学习算法,基于实时轨迹数据估算航空器质量与速度意图,以10 min为预测时间,提升了航空器爬升高度预测的准确性。

航空器离场爬升的垂直剖面预测精确性主要 受航空器质量、速度及意图的影响。大量的历史运 行数据表明,为延长发动机的使用寿命,离场航空 器在起飞和爬升阶段往往使用减推力的模式。然

*通信作者. E-mail: zhangjunfeng@nuaa.edu.cn

收稿日期: 2022-05-31;录用日期: 2022-07-29;网络出版时间: 2022-09-14 13:56 网络出版地址: link.cnki.net/urlid/11.2625.V.20220913.1956.013 基金项目: 国家自然科学基金 (U1933117)

引用格式: 杜卓铭,张军峰, 苗洪连,等. 考虑推力意图的航空器连续爬升垂直剖面预测 [J]. 北京航空航天大学学报, 2024, 50 (4): 1347-1353. DU Z M, ZHANG J F, MIAO H L, et al. Aircraft vertical profile prediction for continuous climb based on thrust intention [J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50 (4): 1347-1353 (in Chinese).

而,现有的基于模型的预测方法往往直接使用最大 爬升推力,无疑对离场航空器垂直剖面预测的准确 性造成了影响。为此,本文立足于航空器离场爬升 的航迹预测,提出一种对航空器推力意图进行建模 的方法,根据推力意图,应用不同的推力设置,结合 航迹预测模型、风温信息、性能参数等,以期提升 离场航空器垂直剖面预测的精度。

1 航空器动力模型

1.1 全能量模型

全能量模型(total energy model, TEM)是航空器 垂直剖面预测的常用模型,对于起飞离场运行而 言,将航空器视作质点,通过分析质点受力及能量 转化,建立全能量方程,如式(1)所示。

$$(T-D)v = mg\frac{dh}{dt} + mv\frac{dv}{dt}$$
(1)

式中:*T*、*D*、*v*、*h*和*m*分别为航空器的推力、阻力、 真空速、高度和质量;g为重力加速度;t为时间。

将式(1)按航空器气动轴分解,考虑航空器燃 油消耗造成的质量减少,可得

$$\begin{cases} \frac{dv}{dt} = \dot{v} = \frac{T - D}{m} - g \sin \gamma \\ \frac{ds}{dt} = \dot{s} = v \cos \gamma \\ \frac{dh}{dt} = \dot{h} = v \sin \gamma \\ \frac{dm}{dt} = \dot{m} = -f \end{cases}$$
(2)

式中:s为水平距离;y为航径角;f为燃油流量。

1.2 推力阻力模型

全能量模型的推力可以通过航空器基础数据 (base of aircraft data, BADA)提供的推力模型进行 计算,其中 BADA 是由欧控实验中心提出,为实现 航空器轨迹仿真的航空器基础数据^[16]。BADA 的用户手册并没有对起飞爬升推力的挡位做明确 界定,而是提供了适用于起飞和爬升阶段的最大爬 升推力,如式(3)^[16]所示。

$$(T_{\text{maxclimb}})_{\text{ISA}} = C_{\text{Tc},1} \left(1 - \frac{h}{C_{\text{Tc},2}} + C_{\text{Tc},3} h^2 \right)$$
 (3)

式中: C_{Tc,1}、C_{Tc,2}、C_{Tc,3}分别代表了第1、第2、第3 最大爬升推力系数。由于航空器在爬升飞行中,随 高度变化,大气温度将偏离标准大气,需要对最大 爬升推力进行修正,如下:

$$T_{\text{maxclimb}} = (T_{\text{maxclimb}})_{\text{ISA}} \left(1 - C_{\text{Tc},5} \Delta T_{\text{eff}}\right)$$
(4)

式中: $\Delta T_{\text{eff}} = \Delta T - C_{\text{Tc},4}, C_{\text{Tc},4} \oplus C_{\text{Tc},5}$ 分别为第1、第2 推力温度系数。

推力计算也可以通过航空器噪声和性能

(aircraft noise and performance, ANP)数据库进行。 ANP数据库是由欧控实验中心与美国汽车工程师 协会合作开发和维护的标准数据库^[17]。ANP数据 库推力分为额定功率推力和非额定功率推力。指 定推力等级代表了在额定功率下发动机的推力情 况,提供了一个与航空器所在高度、速度、温度相 关的函数。对于起飞和爬升阶段,该模式分别提供 了最大起飞与最大爬升2种额定推力。推力计算 函数^[17]为

$$F_{\rm n}/\delta = E + FV_{\rm C} + G_{\rm A}h + G_{\rm B}h^2 + HT_{\rm temp}$$
(5)

式中: *F*_n为单发的静推力; δ为压力比,表示航空器 所在高度大气压力与标准大气压之比; *V*_c为校正空 速; *T*_{temp}为航空器所处位置的温度; *E*、*F*、*G*_A、 *G*_B、*H*分别为发动机推力常数或系数。

针对非额定功率下的发动机运行情况, ANP 数据库提供了参数控制模式实现推力的计算, 如式 (16)^[17]所示。

$$F_{n}/\delta = E + FV_{C} + G_{A}h + G_{B}h^{2} + HT + K_{1}r_{EPR} + K_{2}r_{EPR}^{2}$$
(6)

式中: K_1 和 K_2 分别为 ANP 数据库中与推力和发动 机压力比(engine pressure ratio, EPR)相关的参数, r_{EPR} 为 EPR 的值。

阻力建模如下:

$$D = \frac{C_D \rho v^2 S}{2} \tag{7}$$

式中: C_{D} 为阻力系数; ρ 为空气密度;S为机翼参考 面积。

1.3 燃油流量模型

离场爬升的垂直剖面预测中,航空器燃油流量 由推力和燃油消耗率计算得出,而燃油消耗率η在 BADA 中被表示为真空速的函数,如式(8)所示。

$$f = \eta T \tag{8}$$

$$\eta = C_{f1} \left(1 + \frac{v}{C_{f2}} \right) \tag{9}$$

式中: *C*_{f1}、*C*_{f2}为计算燃油消耗率的相关系数; η为燃油消耗率。

2 航空器意图模型

2.1 基础意图

航空器意图是指描述航空器当前飞行状态,并 指定航空器运动的一系列抽象化指令^[3]。图1为连 续爬升运行航空器的垂直剖面示意图,其中, TOC为爬升顶点(top of climb), CAS为校正空速 (calibrated air speed),本文研究范围主要涉及6个 阶段。



图 1 航空器连续爬升垂直剖面的基础意图示意图

Fig. 1 Aircraft basic intent for vertical profiles under continuous climb operation

阶段 | 航空器采用爬升推力,以固定的校正 空速爬升,爬升至 914.4 m。

阶段Ⅱ 航空器采用爬升推力,以等爬升率加速爬升直至128.6 m/s。

阶段Ⅲ 航空器采用爬升推力,以固定的校正 空速 128.6 m/s 爬升直至 3 048 m。

阶段Ⅳ 航空器采用爬升推力,以等爬升率加 速爬升,加速至机型的爬升速度。

阶段∨ 航空器采用爬升推力,以固定的校正 空速等速爬升,直至到达终端空域改平高度。

阶段 Ⅵ 航空器采用巡航推力,以固定的校正 空速等速巡航。

2.2 推力意图

基础意图主要控制了航空器到达指定高度或 指定速度前,应当保持的爬升方式(等速爬升、加速 爬升等),而推力意图则反映了在爬升过程中推力 挡位的设置。

随着民航地空数据链的发展,数据链带宽、稳定性逐步提高,未来航空器更多的实时信息能够得以下传,其中包括了航空器的自动油门模式和范围等信息。航空器油门范围反映了当前推力所处挡位,可以归纳为航空器的推力意图。若在垂直剖面预测时预先掌握了航空器的推力意图,预测模型根据推力意图输出合适的推力,可以大幅提高连续爬升剖面预测精度。通过整理快速存取记录器(quick access recorder, QAR)数据发现,目前 QAR 记录的油门范围(即推力意图)包括"MAN TOGA"、"THR CLB"等,与其在本文中对应的推力关系如表1所示。

图 2为 A330 机型在相同基础意图下,应用不同的推力意图产生的垂直剖面。由图可知,不同的

Table 1	Correspondence of thrust intention and thrust	
---------	---	--

推力意图与推力对应关系

表1

推力意图	描述	对应推力
MAN TOGA	最大起飞推力	ANP 最大起飞推力
THR CLB	爬升推力	BADA 最大爬升推力
SPEED	加速	保持原阶段推力
MAN FLX	灵活推力	110% BADA爬升推力
THR DCLB1	一挡减推力	ANP一挡减推力
THR DCLB2	二挡减推力	85% BADA推力
THR DCLB3	三挡减推力	ANP二挡减推力



图 2 不同推力意图对垂直剖面影响

Fig. 2 Impact on vertical profiles of different thrust intentions

推力意图对应不同的推力,而不同的推力对连续爬 升的垂直剖面会产生较为显著的影响:一方面,不 同阶段的爬升率是有差异的;另一方面,TOC的位 置及到达TOC的时间也是不同的。

3 离场连续爬升剖面预测

3.1 实验设置和评价指标

本文选择了一份 2017 年 1 月由北京首都国际 机场飞往成都双流国际机场的 QAR 数据作案例分 析(详见 3.2 节),包括了对 7 份 QAR 使用不同预测 模型的误差分析(详见 3.3 节)。

仿真验证中的航空器运行方式为连续爬升,机型为空客A330-200,发动机型号为RRTRENT772B。鉴于此,阶段I的校正空速设为82m/s;阶段IV的校正空速设为154.3m/s;阶段V的终端空域改平高度设为5100m。风温数据来自QAR中的记录信息,本文仅考虑顺逆风分量。由于起飞质量的估计不属于本文研究范畴,故使用实际的起飞质量。

图 3 为 QAR 记录的风分量与标准温差图。由 于 QAR 记录的风温信息时间有限,预测剖面的飞 行时间往往超过 QAR 飞行时间,故后续用风速 0 m/s、温差 0 ℃ 补全。而后续实验的对比部分,将 选取 QAR 飞行时间和预测飞行时间二者中较短时 间的作为对比范围。

几种方法分别为:传统方法是指剖面预测仅考



1350

Fig. 3 Wind component and standard temperature deviation

虑了基础意图,而没有使用风速、温度信息;考虑风 温方法指在传统方法的基础上,加入了 OAR 中风 的分量和大气温度对剖面的影响:本文预测方法则 是在考虑风温的基础上加入了本文提出的航空器 推力意图方法。

本文连续爬升剖面预测采样欧拉法数值积分 实现预测剖面生成。由于 OAR 数据帧记录间隔为 1s,为方便与OAR数据进行对比,故以1s为间隔 对预测剖面进行采样,使用绝对误差(absolute error. AE)作为评价指标:

(10) $\varepsilon_{\rm AE} = |\hat{x}_i - x_i|$

式中:x为状态向量,包括航空器高度h、真空速v、 燃油流量 f等状态信息。通过该指标可以直观的观 察预测剖面与实际剖面误差的绝对值。

爬升剖面预测中,到达 TOC 的时间和位置与 真实剖面的差异大小反映了剖面预测模型的性能, 使用指标平均绝对误差(mean absolute error, MAE) 可以真实的反映预测模型在多份数据下的平均预 测性能,定义如下:

$$\varepsilon_{\text{MAE}} = \frac{1}{n} \sum_{j=1}^{n} \varepsilon_{\text{AE},j} = \frac{1}{n} \sum_{j=1}^{n} \left| \hat{y}_j - y_j \right|$$
 (11)

式中:n为所使用QAR 样本数量; ŷ_i和y_i分别为实际 和预测剖面到达 TOC 的时间或距离。

3.2 预测结果对比

图 4 为连续爬升预测的高度、速度剖面对比, 包含了 QAR 记录剖面、传统方法预测剖面、考虑 风和温度影响的预测剖面,以及本文预测方法的预 测剖面。

在没有考虑推力意图的预测剖面中,由于不能 准确掌握推力意图,故航空器一直保持最大爬升推 力连续爬升,使得加速时机较早(见图 4(c)),总体爬 升率偏高,较早到达 TOC 点后,保持了长时间平飞 (见**冬**4(a))。

而本文预测方法中,根据推力意图使用了相应



Fig. 4 Prediction of altitude and speed profile

的减推力,使预测剖面更好地贴合实际剖面(见 图 4(b))。从总体看,预测剖面的前半段和真实剖 面较为一致,后阶段出现了一定的误差。这主要是 由于空管对于航空器在3048m以下校正空速最高 不得超过 128.6 m/s 的限制, 使得前半段速度、高度 剖面与 OAR 保持了较高的一致性。而后半段受到 管制等因素影响,真实剖面产生了少许平飞,使得 预测剖面产生了一定的误差。

图 5 为 3 种方法的推力预测结果。由图可知, 考虑推力意图的预测方法在起飞阶段中,使用了最 大灵活推力"MAN FLX",并在爬升阶段中切换为 三挡减推"THR DCLB3"。

由于推力无法测量,因此,QAR数据中没有直 接反映推力大小的数值。鉴于此,利用与推力息息 相关的燃油流量反映推力意图,如图6所示。由图



Fig. 6 Prediction of fuel flow

可知,考虑推力意图,其燃油流量与 QAR 记录的燃 油流量能够保持一致趋势。传统方法和考虑风温 方法在爬升的过程中始终采取了最大爬升推力,而 本文预测方法充分利用了油门范围的变化信息,对 推力进行了动态调整,与其他预测方法相比具有更 高的准确性。

值得注意的是,图 5 和图 6 中各有 2 处燃油流 量和推力大幅降低,这是由于油门范围切换为 "SPEED"所导致。由于该模式的流量与推力为非 额定功率推力,只有非额定功率的 ANP 数据库推 力模型(式(6))可以对此种推力进行预测,而在实 际运行过程中 EPR 难以获取(涉及商业机密),在后 续的研究中会加以考虑其他替代的方法。

3.3 误差分析

图 7 为航空器状态变量误差箱线图,包含了对 7 份在管制员干预下连续爬升剖面的 QAR 数据误 差分析。箱线图可以较好的反映数据的分布情况, 矩形外侧的加号代表了数据中的离群点,图中的数 值则代表了当前方法误差的中位数。

而本文预测方法较前二者,无论是其中位数还 是整体分布,有较为显著的提升。其真空速、燃油 流量和高度的中位数AE分别约为5.6 m/s、约 3.7 kg/min和0.13 km。最大误差分别降低了46.8%、 84.0%和53.6%

值得注意的是,本文预测方法较其他2种方法



Fig. 7 Absolute error box diagram of aircraft state variebles

产生了较多的离群点。经过分析发现,误差主要集中在"SPEED"阶段,这是由于该模式为非额定推力,无法直接根据该意图建模推力,故在本文中采取与前阶段一致的推力代替,因此,导致了离群点的产生,在后续的研究中需加以改进。

表 2 包含了上述 7 份 QAR 到达 TOC 时间和距离数据和各方法预测的误差。3 种方法的 TOC 时间的 MAE 分别为 95,90,40 s, TOC 位置的 MAE 分别为 12,13,11 km。各方法对 TOC 位置的预测的 MAE 约为 12 km,本文预测方法未见明显优势。在到达 TOC 时间的方面,本文预测方法误差降低显著, MAE 在 1 min 内, 仅为 40 s, 与其他方法相比, 最高降低误差约 52%。

值得注意的是,表2中第1组和第3组的TOC 点的位置和到达TOC的时间,没有体现本文预测 方法的优越性。但是考虑到,垂直剖面的预测是 过程量,包括了高度、速度及燃油流量预测。通 过对预测数据的进一步分析,第1组和第3组案 例的预测高度、预测速度、预测燃油流量的平均

	表 2 到达 TOC 时间位置误差
Table 2	Time and position error of arrival in TOC

组号 t _{TOC} /s s	a lam		$\varepsilon_{\rm AE}^t/{ m s}$			$\varepsilon_{\rm AE}^s/{ m km}$		
	31007811	传统方法	考虑气象方法	本文预测方法	传统方法	考虑气象方法	本文预测方法	
1	416	52	17	13	46	6	1	11
2	557	84	149	138	74	25	26	16
3	399	49	1	8	51	8	0	11
4	512	69	137	130	16	15	18	7
5	519	68	135	131	14	13	17	5
6	478	55	88	83	66	1	7	17
7	512	69	137	130	16	15	18	7

误差相较于传统方法降低约38%,30%、48%。因此,推力意图的引入,确实可以提升垂直剖面预测的精确性。

4 结 论

 本文考虑了推力意图,可以较为准确地实现 连续爬升垂直剖面预测。

2)本文预测方法与传统方法相比,真空速、燃油流量和高度误差分别降低了46.8%、84.0%和53.6%;到达爬升顶点的时间与实际的平均绝对误差为40s,比传统方法的误差降低约52%。

对于推力意图对应于非额定推力的情况,需要 对本文预测方法再优化。进一步,由于目前油门范 围信息暂未通过数据链下传,为使本文预测方法能 实际应用,应根据历史轨迹数据,建立推力意图的 预测模型,对推力意图加以预测。此外,如何将推 力意图引入数据驱动的垂直剖面预测,也是值得深 入研究的问题。

参考文献(References)

- [1] 曾雯, 胡荣, 宋文, 等. 中国民用航空器 CO₂ 减排潜力的区域划分
 [J]. 北京航空航天大学学报, 2023, 49(9): 2455-2462.
 ZENG W, HU R, SONG W, et al. Regional classification of CO₂ emission reduction potential of China's civil aircraft[J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49(9): 2455-2462 (in Chinese).
- [2] ZENG W L, CHU X, XU Z F, et al. Aircraft 4D trajectory prediction in civil aviation: A review[J]. Aerospace, 2022, 9(2): 91.
- [3] ZHANG J F, LIU J, HU R, et al. Online four dimensional trajectory prediction method based on aircraft intent updating[J]. Aerospace Science and Technology, 2018, 77: 774-787.
- [4] ZHANG J F, PENG Z H, YANG C W, et al. Data-driven flight time prediction for arrival aircraft within the terminal area[J]. IET Intelligent Transport Systems, 2022, 16(2): 263-275.
- [5] WANG Z Y, LIANG M, DELAHAYE D. A hybrid machine learning model for short-term estimated time of arrival prediction in terminal manoeuvring area[J]. Transportation Research Part C:Emerging Technologies, 2018, 95: 280-294.

- [6] LIANG M, DELAHAYE D, MARECHAL P. Conflict-free arrival and departure trajectory planning for parallel runway with advanced point-merge system[J]. Transportation Research Part C:Emerging Technologies, 2018, 95: 207-227.
- [7] FEUSER FERNANDES H, MÜLLER C. Optimization of the waiting time and makespan in aircraft departures: A real time non-iterative sequencing model[J]. Journal of Air Transport Management, 2019, 79: 101686.
- [8] HO-HUU V, HARTJES S, VISSER H G, et al. Integrated design and allocation of optimal aircraft departure routes[J]. Transportation Research Part D:Transport and Environment, 2018, 63: 689-705.
- [9] GUO Y, HU M H, ZOU B, et al. Air traffic flow management integrating separation management and ground holding: An efficiencyequity Bi-objective perspective[J]. Transportation Research Part B:Methodological, 2022, 155: 394-423.
- [10] 张军峰, 葛腾腾, 陈强, 等. 离场航空器四维航迹预测及不确定性 分析[J]. 西南交通大学学报, 2016, 51(4): 800-806. ZHANG J F, GE T T, CHEN Q, et al. 4D trajectory prediction and uncertainty analysis for departure aircraft[J]. Journal of Southwest Jiaotong University, 2016, 51(4): 800-806 (in Chinese).
- [11] SUN J Z, ELLERBROEK J, HOEKSTRA J M. WRAP: An opensource kinematic aircraft performance model[J]. Transportation Research Part C:Emerging Technologies, 2019, 98: 118-138.
- [12] SUN J Z, ELLERBROEK J, HOEKSTRA J M. Aircraft initial mass estimation using Bayesian inference method[J]. Transportation Research Part C:Emerging Technologies, 2018, 90: 59-73.
- [13] ALLIGIER R, GIANAZZA D, DURAND N. Learning the aircraft mass and thrust to improve the ground-based trajectory prediction of climbing flights[J]. Transportation Research Part C:Emerging Technologies, 2013, 36: 45-60.
- [14] ALLIGIER R, GIANAZZA D, DURAND N. Machine learning and mass estimation methods for ground-based aircraft climb prediction[J]. IEEE Transactions on Intelligent Transportation Systems, 2015, 16(6): 3138-3149.
- [15] ALLIGIER R, GIANAZZA D. Learning aircraft operational factors to improve aircraft climb prediction: A large scale multi-airport study[J]. Transportation Research Part C:Emerging Technologies, 2018, 96: 72-95.
- [16] Eurocontrol Experimental Center. User manual for the base of aircraft data (BADA) [Z]. Paris: Eurocontrol Experimental Center, 2014: 22.

[17] European Civil Aviation Conference. Report on standard method of computing noise contours around civil airports – Volume 2: Technical guide [S]. Paris: European Civil Aviation Conference, 2016: B5-B6

Aircraft vertical profile prediction for continuous climb based on thrust intention

DU Zhuoming¹, ZHANG Junfeng^{1,*}, MIAO Honglian¹, WANG Bin²

(1. College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China;

2. Central and Southern Regional Air Traffic Management Bureau of the Civil Aviation Administration of China, Guangzhou 510403, China)

Abstract: Accurate continuous climb profile prediction can improve conflict detection reliability and departure scheduling precision. A method of modeling aircraft thrust intention based on the thrust setting information is proposed. A vertical profile prediction method for the continuous ascent is suggested, taking into account temperature data, wind vector, thrust purpose, and the total energy model. The case studies and comparative analysis are based on quick access recorder (QAR) data. The analysis is focused on the error between the predicted and actual values of true airspeed, altitude, and fuel flow at each sampling data from QAR. In addition, the evaluation is done on the mean absolute error of the duration and distance to the top of climb (TOC) between the actual and anticipated values. The results indicate that the TOC arrival time mean absolute error at TOC can be reduced by approximately 52% compared to prediction methods without considering the thrust intent.

Keywords: air traffic management; trajectory prediction; continuous climb; performance model; thrust intention

Received: 2022-05-31; Accepted: 2022-07-29; Published Online: 2022-09-1413:56 URL: link.cnki.net/urlid/11.2625.V.20220913.1956.013

Foundation item: National Natural Science Foundation of China (U1933117)

^{*} Corresponding author. E-mail: zhangjunfeng@nuaa.edu.cn

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

DOI: 10.13700/j.bh.1001-5965.2022.0449

基于贝叶斯网络强化学习的复杂装备 维修排故策略生成

刘宝鼎1,于劲松1,*,韩丹阳2,唐荻音1,李鑫3

(1. 北京航空航天大学 自动化科学与电气工程学院,北京 100191; 2. 北京航空航天大学 仪器科学与光电工程学院,北京 100191;3. 中国运载火箭技术研究院,北京 100076)

摘 要:为解决传统启发式维修排故决策方法决策时间长、生成策略总成本高的问题,提 出一种基于贝叶斯网络(BN)结合强化学习(RL)进行复杂装备维修排故策略生成方法。为更好 地利用复杂装备模型知识,使用 BN进行维修排故知识表述,并且为更加贴近复杂装备实际情况, 依据故障模式、影响和危害性分析(FMECA)的故障概率,经合理转化后作为 BN 的先验概率;为 使用 RL 的决策过程生成维修排故策略,提出一种维修排故决策问题转化为 RL 问题的方法;为更 好地求解转化得到的强化学习问题,引入观测-修复动作对(O-A)以减小问题规模,并设置动作掩 码处理动态动作空间。仿真结果表明:在统一的性能指标下,所提 BN-RL 方法较传统方法获得更 高的指标值,证明该方法的有效性和优越性。

关 键 词: 强化学习; 贝叶斯网络; 维修排故策略生成; 复杂装备; 动态动作空间

中图分类号: TP206+.3

文献标志码: A 文章编号: 1001-5965(2024)04-1354-11

近年来,随着工业装备的复杂化、精密化和智能化,复杂装备不同故障源往往表现相近,导致常用的故障诊断方法只能定位至故障模糊组,无法准确判断故障位置和类型¹¹。常采取的对模糊组中部件全部替换或随机维修的方式存在针对性不强、效率低下、修复成本高的问题。为解决上述难题,需要通过维修排故决策,根据复杂装备当前状态和故障信息动态选择维修对象部件,实时确定观测、维修动作顺序,以生成针对性更强、效率更高、修复成本更低的维修排故策略。

然而,传统维修排故决策方法在面对复杂装备 时,由于问题规模大,导致维修排故策略生成耗时 长、维修排故效率低。因此,如何对复杂装备在更 短时间内高效生成成本更低的维修排故策略具有 重要研究意义。

维修排故决策通过观测和维修的交互不断收 集知识进行决策,因此,维修排故知识在决策过程 中起到重要作用,维修排故知识的高效表达也是维 修排故策略生成的关键环节。传统的决策树表达 排故知识单一、低效,其规模随节点呈指数型增 长。例如,包含2个维修节点和1个观测节点的决 策树有12个终极节点,在5个维修动作、3个观测 动作时迅速增长到近340000个终极节点^[2]。决策 树规模增长过快是维修排故知识管理复杂、策略生 成困难的根本原因,并且决策树无法表达复杂装备 结构知识,导致维修排故知识的缺失,因此,不适用 于复杂装备的维修排故决策。相比于决策树,贝叶 斯网络 (Bayesian networks, BN)^[3]规模随节点呈线

收稿日期:2022-05-31;录用日期:2022-07-22;网络出版时间:2022-08-0111:00 网络出版地址:link.cnki.net/urlid/11.2625.V.20220729.1212.001

基金项目:国家重点研发计划(2018YFB1403300);国家自然科学基金(51875018,71701008)

^{*}通信作者. E-mail: yujs@buaa.edu.cn

引用格式:刘宝鼎,于劲松,韩丹阳,等.基于贝叶斯网络强化学习的复杂装备维修排故策略生成[J].北京航空航天大学学报,2024, 50(4):1354-1364. LIU B D, YU J S, HAN D Y, et al. Complex equipment troubleshooting strategy generation based on Bayesian networks and reinforcement learning[J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50(4): 1354-1364 (in Chinese).

性增长,增长速度较慢,在表达结构知识方面具有 优势,适用于复杂装备。并且 BN 能高效处理维修 排故决策问题本身包含的不确定性和因果关系,因 此受到广泛应用^[4]。然而,传统基于 BN 的方法在 设置 BN 先验故障概率时通常采用各部件故障等概 率发生的方式,此时的 BN 模型相对脱离实际,导 致生成的策略成本升高。

在 BN 的基础上, 现有维修排故决策方法大多 以低预期修复成本为目标,采用启发式方法生成排 故策略。例如, Heckerman 等^[2]在单故障假设下提 出2种基于决策理论的排故策略生成方法,一种仅 考虑维修操作,另一种在维修前加入最多一步的间 接观测操作,与静态排故策略比较具有优势; Jensen 等^[5]在 Heckerman 等^[2]研究的基础上将维修 操作推广为非理想的情况,并提出向前两步间接观 测的方法;于劲松等⁶⁰为解决向前一步间接观测在 获取排故信息过程中的限制,基于 BN 和决策理论 提出向前多步排故策略牛成方法,并采用决策影响 图进行排故决策分析; Huang 等^[7]将 BN 与多准则 决策分析 (multi-criteria decision analysis, MCDA) 方 法结合进行维修排故决策,综合故障概率、风险、 时间、成本4个标准的影响计算动作的效用值。除 此之外,元启发式方法作为构建启发式方法的有效 范式,也被应用到维修排故决策问题中。例如, Vianna 等^[8] 使用结合多起点搜索的模拟退火算法 优化排故策略; de Oliveira 等^[9]将基于教学的优化 (teaching-learning-based optimization, TLBO) 算法应 用于维修排故中,获得了比遗传算法、模拟退火算 法更好的性能; Coelho 等^[10] 将混沌惯性权值应用 于 TLBO 的随机参数中,获得了比标准 TLBO 更好 的性能。然而,启发式方法多采用遍历思想,导致 问题求解耗时较长并且效果受方法参数设置影响 较大[11]。此外, 启发式方法获得的解不具有重用 性,仅对当前故障情况有效,因此,该方法在每次维 修排故决策中都需要重新运行一次以获得新的解, 浪费计算资源、降低维修排故效率。

启发式方法计算效率低,导致决策耗时长、维 修排故效率低,有进一步优化空间,因此,本文提出 一种维修排故决策问题转化为强化学习 (reinforcement learning, RL)问题的方法。RL方法 具有"离线学习、在线决策"的特性,不同于启发式 方法为不同故障情况分别生成策略,RL方法离线 阶段学习复杂装备整体特征,在线阶段进一步结合 具体故障情况生成维修排故策略,一次学习、多次 决策,因此,策略生成耗时更短、维修排故效率更 高。此外,RL的"动作选择"与组合优化问题在离 散决策空间内进行决策变量的最优选择具有天然

相似的特征^[12],并且 RL 方法采用动态规划的思想, 能够更好的解决维修排故决策这类动态序列优化 问题。目前,针对类似的组合优化问题的研究中, 使用 RL 的决策过程生成可行解的方法取得了较大 进展^[13]。例如, Zhang 等^[14]提出基于元学习的深度 RL方法用于多目标旅行商问题,较大多数基于学 习和基于迭代的方法有优势,并且在处理不同尺度 的问题时具有良好的泛化能力; Oren 等^[15]提出结 合 RL 和规划的方法用于机器调度问题, 在计算时 间和性能上优于传统方法; Almasan 等^[16] 基于深度 RL方法对无线通信网络的路由策略进行研究,实 验表明该方法具有很强的泛化能力,模型训练完成 后能够在线优化任意结构网络的路由策略。然而, 在 RL 应用于复杂装备维修排故决策问题的过程 中,模型节点数较多导致 RL 问题规模较大,并且维 修排故决策过程中可采取的动作动态变化导致 RL具有动态动作空间,进而难以收敛,给RL的应 用带来了困难。

综上所述,为解决传统维修排故方法在面对复杂装备时存在的决策耗时长、成本高、维修排故效率低下的问题,本文提出一种基于 BN 结合 RL 进行复杂装备维修排故策略生成的方法(简称本文BN-RL 方法)。首先,将故障模式、影响和危害性分析 (failure mode, effects and criticality analysis, FMECA)表中定义的故障概率经过转化后应用到BN 中,以更加贴近复杂装备实际情况;其次,提出了维修排故决策问题转化为 RL 问题的方法,以使用 RL 决策过程生成维修排故策略;最后,为解决对复杂装备进行维修排故决策时问题规模过大和动态动作空间导致 RL 难以收敛的问题,采用"观测修复"动作对(observation-action pair,O-A)大幅降低 RL 问题规模,并通过设置动作掩码处理动态动作空间。

1 维修排故决策问题

1.1 问题定义

维修排故流程是在设备发生故障时,对设备执 行序贯式测试和维修步骤,直至设备恢复正常状态 的过程⁶⁶。维修排故决策问题即以维修排故流程总 成本最低为目标,生成最优维修排故策略的问题。

1.2 问题要素

维修排故决策问题通常具备以下要素[17]。

1) 故障集合<u>F</u>

设备可能发生的所有故障集合。在设备发生 故障时,<u>F</u>中至少有一个故障模式发生故障,也可 能有多个故障模式同时发生故障;设备故障被排除 当且仅当<u>E</u>中所有故障模式状态均为正常。 2) 维修动作集合A

用于修复<u>F</u>中故障模式的维修动作集合。当设 备发生故障时,必须采取<u>A</u>中的动作才能达到修复 设备的目的。

3) 观测动作集合 0

对设备状态的观测动作集合。当设备故障时, Q中的观测动作不会修复设备的故障,但可以为维 修人员提供额外的设备状态信息。这些额外信息 通常可以缩小设备故障的待排查范围,起到缩短排 故流程、降低排故总代价的作用。

4) 设备状态集合<u>S</u>

设备在完整的排故流程中可能的状态集合。 通常情况下,设备是由多个可分离的部件c组成的 集合,记为Call。设备的故障可以解释为组成设备 的部件集合CCOMP中的一个或多个部件的故障,设 备的状态也可以表示为CCOMP中部件的状态集合,即

 $\underline{S} = \bigcup_{\underline{c_i} \in \mathcal{C}_{\text{COMP}}} \underline{s_{c_i}} \tag{1}$

式中: <u>c</u>_i为设备的组成部件; s_{ci}为第i个部件的状态。 5) 系统测试T

系统测试用于检验设备是否正常工作,是判断 维修排故流程是否继续的依据。系统测试只有 "YES"和"NO"2种结果。其中,"YES"代表系统 正常,"NO"代表系统故障。

6) 代价函数C(s)

代价函数描述设备当前状态s下执行各维修动 作或观测动作的代价。动作的代价可以由以下因 素中的一个或多个决定^[6]:动作所需时间t、动作复 杂程度d、动作所需费用m、动作的风险r、动作重 要性i,且与设备当前状态s相关,即

 $\underline{C(s)} = f(t, d, m, r, i|s) \tag{2}$

式中: f与t,d,m,r成正比,与i成反比,即

$$f(t,d,m,r,i|s) \propto t dm r \frac{1}{i}$$
(3)

与多数同类研究类似,本文只考虑最常用的动作所需费用因素*m*,此时有

$$\underline{C(s)} = f(m|s) \propto m \tag{4}$$

7) 决策策略<u>P</u>

排故决策者选择动作的策略,可以表示为由维 修动作集合<u>A</u>与观测动作集合<u>O</u>中的动作组成的静态序列,也可以表示为关于设备状态的动态函数。

1.3 问题假设

在问题要素的基础上,对维修排故问题加以合 理的假设和约束,能够避免过于泛化,更加具体地 描述问题。维修排故问题通常有以下假设^[17]。 初始故障假设。假设在维修排故流程开始 时,目标设备处于故障状态。否则无需进行维修排 故流程。

2) 完美动作假设。所有的维修动作和观测动 作在执行时都没有副作用,即不会引入新的设备 故障。

3)维修幂等性假设。多次执行重复的维修动 作不会使设备状态发生新的变化,即有效的维修动 作只需执行一次就能修复故障,而无效的维修动作 永远不会使设备恢复正常。

 4)观测幂等性假设。同一个观测动作的重复 执行始终会给出同样的观测结果,不会为维修人员 带来更多的信息。

5) 真实系统测试假设。系统测试T始终会准确地给出系统当前的真实状态:正常运行或依然故障。T通常在每一步维修动作后执行,以确定终止还是继续维修排故流程。通常情况下系统测试T的成本认为是0。

2 维修排故知识表示

为高效表示维修排故知识,需要结合维修排故 决策问题的特性构建 BN,并结合复杂装备真实情 况设定网络中的先验概率。

2.1 贝叶斯网络建模

基于 BN 的维修排故决策问题模型如图 1 所示。根据复杂装备部件可维修性和可观测性,可以将网络节点类型区分为如下类型。

1)问题定义节点。BN中问题定义节点用于表示装备的故障状态。在维修排故决策过程中,装备始终处于故障状态,因此,问题定义节点故障概率为1。

2)可观测部件节点、不可观测部件节点。维修的节点称为部件节点,对应维修排故决策问题中的维修动作,用于修复装备,如图1中节点C1~C6。



based on BN

其中节点 C_1 、节点 C_2 、节点 C_3 和节点 C_5 的故障状态 可被观测,对应观测和维修 2 种动作,称为可观测 部件节点,节点 C_4 和节点 C_6 不可观测其故障状态, 只对应维修动作,称为不可观测部件节点。

3)间接观测节点。可观测其状态但不可维修的节点称为间接观测节点,只对应观测动作,用于获取更多信息以辅助维修排故决策,如图1中节点 O₁、节点O₂和节点O₃。

4)隐含节点。BN中为更好表达装备结构知识 并使模型更易于解释而创建的节点称为隐含节点, 不对应维修动作或观测动作,如图1中节点C₇~C₁₀。

2.2 先验概率的确定

BN 需要网络中节点的先验概率以进行概率计 算和传播。传统方法为简化模型、便于计算通常等 概率设置各个部件的故障先验概率,但一定程度上 与复杂装备的实际情况不符。作为常用的设备可 靠性分析和评估方法,FMECA将记录各个故障模 式的发生概率^[18]。在维修排故决策问题的 BN 模型 中,各部件节点相应的故障模式发生概率即可作为 此节点的先验故障概率。

然而,故障概率在维修排故问题与 FMECA 中 的定义存在差异。在 FMECA 中,故障概率的定义 为在任务阶段内故障模式发生的先验概率,单位为 1/h^[19],其量级通常为 10⁻⁴、10⁻⁵;而在维修排故问题 中,设备始终处于故障状态,节点故障概率定义为 设备故障时此故障模式发生的条件概率。因此, FMECA 表中获取的故障概率 λ_p不能直接作为先验 概率,需要对其进行转化,以满足维修排故决策问 题中初始故障的要求。

根据初始故障假设和故障独立假设,各部件的 先验故障概率之和应为1,且根据单故障假设,当某 一部件*c*_i故障时,*C*_{all}中其他部件均为正常。因此, 本文定义 BN 中的先验概率为

$$p_{\rm f}^i = \frac{p_{\rm f}^i}{\sum p_{\rm f}^i} \tag{5}$$

 $\overline{p_{\rm f}^{i}} = \lambda_{\rm p}^{i} \prod_{j \neq i} \left(1 - \lambda_{\rm p}^{j} \right) \tag{6}$

式中: $p_{\rm f}$ 为第i个节点的先验故障概率; $\lambda_{\rm p}$ 为FMECA 表中定义的部件 c_i 故障的概率。

3 维修排故策略生成

使用 RL 方法解决维修排故决策问题之类的组合优化问题是当前的研究热点^[13]。并且维修排故决策问题中设备的后续状态只与当前状态和当前采取的措施有关,因此,满足 RL 的马尔可夫性基本

假设。基于以上原因,本文探索使用 RL 动作决策 过程进行维修排故策略生成。

3.1 维修排故决策问题向强化学习问题的转化

为使用 RL 方法解决维修排故决策问题,需要将维修排故决策问题转化为 RL 问题。RL 通过智能体与环境的不断交互进行学习,其基本要素^[20]包括环境的状态空间 S、智能体动作空间A、智能体策略 $\pi(a|s)$ 、状态转移概率 p(s'|s,a)和即时奖励 R(s,a,s'),其中, a 为智能体采取的动作, s'为动作后环境转移到的状态。一个 RL 问题可用五元组 (S,A,π,p,R) 表示。

RL 环境即维修排故决策问题中的对象设备, RL 智能体即排故决策问题中的决策者,上述五元 组(*S*,*A*,*π*,*p*,*R*)的转化方式如下(为便于区分,维修 排故决策问题中的符号均加下划线表示)。

1) 环境的状态空间S

RL环境的当前状态即维修排故决策问题中设 备当前状态,因此,RL状态空间可用维修排故决策 问题中设备状态集合表示,即

 $S \equiv \underline{S} \tag{7}$

在维修排故决策问题中,各部件状态用正常、 故障等离散状态表示,因此,设备状态集合转化成 的 RL 状态空间是离散的。

2) 智能体的动作空间A

RL智能体可采取的动作a可表示为维修排故 决策问题中决策者可以执行的动作a,因此,RL智 能体的动作空间可用决策者可执行的动作集合表 示,即

 $A \equiv \underline{A} \cup \underline{O}$

维修排故决策问题中动作分为观测和维修 2类,因此,转化成的RL动作空间是离散的。

3) 策略π

RL 中的策略π(a|s)即为排故决策问题中决策 者根据设备状态选择动作的策略<u>P</u>。

4) 状态转移概率p(s'|s,a)

由 1.3 节中的完美动作假设,设备的故障原因 和真实状态<u>sreal</u>在进入排故流程时即确定并且不会 发生变化,且由单故障假设,有且仅有一个设备部 件发生故障,记为c_f。

因此,当观测<u>c</u>_i时,环境状态将会向真实状态 <u>s_{real}</u>靠拢,下一时刻环境状态s_{i+1}相当于用<u>c</u>_i的真实 状态代替当前环境状态s_i中c_i的状态,即

$$s_{t+1} = s_t \setminus \left\{ \underline{c_i} | \underline{s_t} \right\} \cup \left\{ \underline{c_i} | \underline{s_{\text{real}}} \right\}$$
(9)

当维修 c_i 时,若 $c_i \neq c_f$,则环境状态不发生变

(8)

(10)

化,即

 $s_{t+1} = s_t$

当 $c_i = c_f$ 时环境状态发生变化, c_f 的状态由 "N_{NO}" 变为 "Y_{YES}", 即

$$s_{t+1} = s_t \setminus \left\{ \underline{c_f} = N_{\text{NO}} \right\} \cup \left\{ \underline{c_f} = Y_{\text{YES}} \right\}$$
(11)

进一步由14节中的完美修复假设,本节状态 转移概率均为1。

5) 即时奖励*R*(*s*,*a*,*s*')

RL 中即时奖励函数 R(s,a,s')的设置取决于 RL 的目标。这里的目标为最小化维修排故决策过 程产生的总代价, 而 RL 智能体会学习奖励最大的 策略,因此,需要

(12) $R(s, a, s') \propto -C$

根据本节对状态转移概率的讨论,环境下一个 状态s'由当前状态s和采取的动作a唯一确定,且智 能体的动作由其策略 $\pi(a|s)$ 和当前状态决定,则有

$$R(s,a,s') = R(s,a) = R(s)$$
 (13)

又根据固定成本假设,动作的成本与设备状态 无关,即C(s)=C,因此,即时奖励可表示为

R(s, a, s') = -C(14)

综上所述,对于一个维修排故决策问题,总可 以通过合适的要素转换将其转换成离散环境状态 空间、离散动作空间的 RL 问题来解决。

3.2 本文 BN-RL 方法

由 3.1 节论证,维修排故决策问题可通过适当 步骤转化为强化学习问题解决,因此,提出本文 BN-RL 方法,该方法原理框架如图2所示。



图 2 本文 BN-RL 方法原理框架

Fig. 2 Principal framework of proposed BN-RL method

本文 BN-RL 方法将复杂装备对应的 BN 模型 作为 RL 环境,智能体在离线学习阶段与环境进行 大量交互,根据反馈不断更新自身策略直至收敛; 在在线决策阶段根据学习到的策略结合实际故障 状态进行动作选择,进而生成维修排故策略。由于 转化成的 RL 问题具有离散状态空间和离散动作空 间,因此,本文以 Q-learning 算法为基础进行求解^[21]。

O-learning 算法^[22] 是一种基于价值的 RL 方法, 其中Q值表示Q(s,a),即在状态s采取动作a动作能 够获得收益的期望。Q-learning 算法基本思想^[23]是 将状态空间与动作空间构建成一张 Q 值表, 在离线 学习阶段采取以一定的探索率 ($\varepsilon \in (0,1)$)的 ε -greedy方式选择动作*a*,即以 ε 的概率贪婪选择 O值表中收益最大的动作, $1-\varepsilon$ 的概率随机选择动 作,接着根据环境反馈的r和s'更新 O 值表,其更新 方式^[24] 如式 (15)、式 (16) 所示:

 $Q(s,a) \leftarrow Q(s,a) + \alpha T_{\rm D}$ (15)

$$T_{\rm D} = r + \gamma \max Q(s', a') - Q(s, a) \tag{16}$$

式中:γ为折扣系数;T_n为时间差分;α为学习率。 最后得到的 O 值表即为智能体在环境中的策略;在 线决策阶段根据当前环境状态结合学习完成的 O 值表选取能够获得最大的收益的动作。Q-learning 算法伪代码^[25]如下。

1) 学习阶段

输入: RL 五元组(S,A,π, p,R)

输出:0 值表

1随机初始化 0 值表;

2设置学习率 α ,折扣系数 γ ,探索率 ε ;

3 for each episode do

- 4 重置s为初始状态 s_0 ;
- 5 for each step do
- 6 在 A 中 ε -greedy 选择动作 a;

7 执行动作a,观察反馈的r和s';

- 8 根据式 (15) 和式 (16) 更新 O 值表;
- 9 更新*s* ← *s*′:

10 until $s = s_{\text{terminal}}$;

11 until 收敛;

2) 决策阶段

输入: RL 五元组(S,A,π, p,R), 学习完成的 O 值表 输出:维修排故策略

1 while $s \neq s_{\text{terminal}}$ do

- 2 根据 O 值表在A 中贪婪选择动作a;
- 3 执行动作a,观察反馈的r和s';
- 更新*s* ← *s*′; 4
- 3.3 观测-修复动作对的引入

在 RL 模型中, 将智能体对设备模型中各节点

的状态认知分为3种:初始的模糊状态、正常状态 和故障状态,分别用-1、0、1来表示;智能体对节点 采取的动作有2种,分别为对可观测部件节点和间 接观测节点的观测动作和对所有部件节点的修复 动作。根据3.1节的论证,转化成的RL问题具有 离散环境状态空间和离散动作空间,因此,以RL动 作数与状态数的乘积表示其问题规模。此时的 RL问题规模大小可表示为

 $S_{c1} = (N_{obs} + N_{act}) \times 3^{N_{node}}$ (17)

式中: Nobs为观测动作数; Nact为维修动作数; Nnode为初始状态为模糊的节点数。

现定义对部件节点的 O-A,其定义为先对部件 节点进行观测,若状态为正常,则结束当前动作,若 为故障则直接进行修复。此时智能体可对节点采 取的动作变为对间接观测节点的观测动作和对所 有部件节点的 O-A 动作。并且在执行 O-A 动作 后,对应的部件节点必然处于正常状态,因此,此时 部件节点只有模糊 (-1) 和正常 (0) 这 2 种可能。引 入 O-A 动作后 RL 问题规模大小变为

$$S_{c2} = (N_{ind} + N_{comp}) \times 2^{N_{comp}} \times 3^{N_{node} - N_{comp}}$$
(18)

式中: N_{ind}为间接观测节点数; N_{comp}为部件节点数。

由于 $N_{ind} \leq N_{obs}, N_{comp} \leq N_{act}, 2^{N_{comp}} < 3^{N_{comp}}, 可得$ $S_{c2} < S_{c1}, 因此, RL 问题规模得到减小。$

3.4 动态动作空间的处理

由于 1.3 节中的动作幂等性和观测幂等性假 设,智能体重复执行相同的动作不会获得更多的信 息或修复设备,并将产生更大的成本。因此,智能体 不应重复执行相同的动作,即智能体的动作空间是 动态变化的。动态动作空间中的无效动作会给 RL 智能体的训练带来困难,导致算法收敛慢、收敛 效果不好的问题。

对于动态动作空间通常有2种解决方法:软限制,给不合法动作设置一个很大的负奖励以引导智能体不选择这类动作;硬限制,设置动作掩码来屏蔽不合法的动作,不允许智能体选择。软限制方法 在无效动作较多时将失效^[26],因此,本文选择硬限制方法,通过对不合法动作加以掩码屏蔽,避免智能体探索不合理的动作。

记动作掩码为 A_{mask} ,掩码为1表示动作可被选择,为0则不可被选择,其初始化值为 $A_{mask}(a_i) = 1, \forall a_i \in A, 并记掩码为1和为0的动作集合分别为 <math>A^1_{mask}, A^0_{mask}$ 。本文 BN-RL 方法使用加入动作掩码 后的 Q-learning 算法进行强化学习过程,加入动作 掩码后的 Q-learning 算法伪代码如下所示。

1) 学习阶段 **输入:**维修排故决策问题转化的五元组(S,A,π,p,R) 输出:0 值表 1随机初始化0值表; 2设置学习率 α ,折扣系数 γ ,探索率 ε : 3 for each episode do 4初始化 A_{mask} : $A_{\text{mask}}(a_i) = 1, \forall a_i \in A;$ 5 重置s为初始状态 s_0 ; 6 for each step do 7在 A_{mask}^1 中 ε -greedy选择动作a; 8执行动作a,观察反馈的r和s'; $9 A_{mask}(a) = 0;$ 10根据式(15)和式(16)更新 Q 值表; 11 更新*s* ← *s*′; 12 until $s = s_{\text{terminal}}$; 13 until 收敛; 2) 决策阶段 输入:维修排故决策问题转化的五元组

 (S,A,π,p,R) ,学习完成的 Q 值表

输出: 维修排故策略 1 初始化 A_{mask} : $A_{mask}(a_i) = 1, \forall a_i \in A$; 2 while $s \neq s_{terminal}$ do 3 根据 Q 值表在 A_{mask}^1 中贪婪选择动作a; 4 执行动作a, 观察反馈的 $r \pi s'$; 5 $A_{mask}(a) = 0$; 6 更新 $s \leftarrow s'$;

4 仿真分析

为验证本文 BN-RL 方法的有效性,选择"汽车 不能启动"问题^[2]作为实验背景,对比本文 BN-RL 方法与随机决策方法、静态决策方法、决策理论方法^[2] 和向前一步间接观测方法^[2]进行性能对比。

4.1 问题模型

"汽车不能启动"问题的 BN 模型如图 3 所示。 该网络共有 17 个节点,其中,包含 9 个部件节点(4 个 可观测、5 个不可观测),5 个间接观测节点,2 个隐含 节点和 1 个问题定义节点,网络中共有 17 条连接线。

4.2 对比决策方法简介

4.2.1 全知决策方法

全知决策方法假设维修时明确知道导致故障 的具体部位,并对其直接进行操作,代表问题的最 优情况。

4.2.2 随机决策方法

随机决策方法以随机的方式选择观测和维修 动作,直至设备故障被修复,代表问题的随机情况。 4.2.3 静态决策方法

静态决策方法按照代价从小到大的顺序选择



Fig. 3 BN model of "Car can't start" problem

动作,是解决维修排故决策问题的朴素方法。

4.2.4 决策理论方法

决策理论方法^[2] 是一种启发式方法,根据各部 件节点的效费比,即故障概率与观测费用之比来选 择部件进行维修。

4.2.5 向前一步间接观测方法

向前一步间接观测方法^[2] 是一种启发式方法, 在决策理论排故策略的基础上进行改善,在开始对部 件节点进行维修动作前执行观测操作以获取更多 的设备状态信息辅助决策,使决策更加准确。根据观 测操作执行后设备的预期修复成本选择观测操作。

4.3 结果分析

在仿真实验中,本文 BN-RL 方法与上述 5 种参与对比的决策方法所需要的装备模型部件故障概率参数和各动作成本参数均保持一致,以保证方法对比的合理性。

4.3.1 方法整体性能对比

对"汽车不能启动"问题场景随机生成1000 个故障样本用于仿真计算,各方法在样本整体上的 维修排故成本分布如图4所示,在仿真样本上的平 均决策耗时t和平均维修排故成本C对比如表1所 示,其中为便于比较,将排故时间t以全知决策方法 耗时为基准标准化。





由图 4 可知, 仅考虑维修排故成本C指标时, 本 文 BN-RL 方法与全知决策方法代表的最优情况类 似, 维修排故成本主要分布在低成本区域, 并且优 于其他非最优方法。但是由表 1 可看出, 同时考虑 t和C这 2 个衡量标准时, 除全知决策方法外的各方 法在 2 个指标下的表现不同步, 例如随机决策方法 和静态决策方法在决策时间方面表现优秀, 但维修 排故成本较高,决策理论方法^[2]、向前一步间接观 测方法^[2]等基于 BN 的方法在维修排故成本方面表 现更好但耗时更长。因此,需要构建一个新的性能 指标以衡量各方法在2指标下的综合表现。本文 设置对比指标*I*如式(19)~式(21)所示:

$$I = w_C \frac{C_{\max} - C}{C_{\max}} + w_t \frac{t_{\max} - t}{t_{\max}}$$
(19)

表1 各方法耗时和维修排故成本

of each method						
方法	С	t				
全知决策方法	25.46	1				
随机决策方法	92.83	1.67				
静态决策方法	134.31	2.11				
决策理论方法	45.61	37 990.20				
向前一步间接观测方法	38.23	383 244.21				
本文BN-RL方法	30.66	11.10				
$w_C = \frac{\Delta_C + \Delta_t}{\Delta_C}, w_t = \frac{\Delta_C + \Delta_t}{\Delta_t}$		(20)				

$$\Delta_C = \frac{C_{\max} - C_{\min}}{C_{\min}}, \Delta_t = \frac{t_{\max} - t_{\min}}{t_{\min}}$$
(21)

式中: C_{max}和C_{min}分别为参与对比的所有方法排故 成本最大值和最小值; t_{max}和t_{min}分别为排故时间最 大值和最小值。

各决策方法在上述指标s衡量下的性能表现对 比如图5所示,其中为便于比较,各方法性能指标 I是经过归一化的结果,最优为1,最差为0。



Fig. 5 Performance of each method

相比于随机排故方法和朴素的静态排故方法, 本文 BN-RL 方法排故时间稍长但策略的排故成本 大幅减少,策略大幅优化,因而性能指标大幅提高; 相比于同样基于 BN 的经典启发式决策方法,本文 BN-RL 方法在进一步减少排故成本的同时大幅缩 短排故时间,因而性能指标得到进一步提高。综上 所述,本文 BN-RL 方法优于其他非最优维修排故 决策方法。

4.3.2 引入 FMECA 效果验证

本文通过 2.2 节中所述方法转化 FMECA 分析 定义的部件故障概率 λ_p作为 BN 先验部件故障概 率,以改进传统方法中常用的等概率方式设置的先 验概率。

参与对比的非最优方法中静态决策方法和随 机决策方法不使用部件的故障概率信息进行维修 排故决策,因此,对 BN 的改进对其没有影响。对 于本文 BN-RL 方法和剩余 2 种基于 BN 进行维修 排故决策的方法, FMECA 的引入对其决策时间影 响较小,而在维修排故成本方面影响更大,因此,主 要对比各方法在引入 FMECA 前后的维修排故成 本,结果如图6 所示。

由图6可知,引入 FMECA 表中故障概率改进 等先验概率的 BN 后,本文 BN-RL 方法和其他 2 种 以 BN 为基础的方法都产生了具有更低维修排故成 本的策略。





4.3.3 加入 O-A 动作效果验证

"汽车不能启动"问题在 O-A 动作加入前后问题规模对比如图7 所示。

由图7可知,加入O-A动作前问题规模为 $18 \times 3^{16} \approx 7.7 \times 10^8$,加入O-A后问题规模变为14× $2^9 \times 3^7 \approx 1.6 \times 10^7$ 。因此,"汽车不能启动"问题在加入O-A动作后 RL问题规模大幅减小。



Fig. 7 Problem scale before and after introducing O-A action

4.3.4 设置动作掩码效果验证

为对比设置动作掩码前后 RL 训练的效果,每 训练1000次,测试一次智能体在4.3.1节中样本整 体上的平均决策时间和平均维修排故成本,共测 试 500次,最终将获得的性能数据通过式 (19)~ 式 (21)进行综合性能指标 I计算,并进行归一化以 方便对比。设置动作掩码前后性能指标对比如 图 8 所示。





由图 8 可知,设置动作掩码后, RL 训练过程更 快收敛,并且性能指标 s较未设置动作掩码时有较 大提升。因此,通过对不合法动作设置掩码屏蔽的 方式处理动态动作空间能够提升智能体的学习效 率,并提升算法整体性能。

5 结 论

1)根据维修排故决策问题特性对 FMECA 表中 定义的故障概率进行转化后应用到 BN 模型中,作 为先验故障概率。对于使用 BN 进行维修排故知识 表达的方法,相较于等故障概率方式设置先验概 率,本文 BN-RL 方法生成的维修排故策略在引入 FMECA 之后更加贴近实际,因而具有更低的维修 排故成本。

2)将针对部件节点的观测动作和维修动作相结合,形成 O-A 作为部件节点的新动作,大幅减小了 RL 问题规模。

3)根据维修排故决策问题特性,针对动态动作 空间,加入不合法动作掩码避免智能体探索无效动 作,使得 RL 智能体的学习效率得到提高,算法整体 性能得到提升。

为能够满足更一般、更加贴合实际的维修排故 决策问题场景,未来需要研究基于 RL 的排故方法 在更加松弛的假设条件下的有效性。

参考文献(References)

 [1] 郭文彬, 刘东, 王宇健. 功能交联条件下飞机混合增强故障诊断 方法[J]. 测控技术, 2022, 41(10): 107-113.

GUO W B, LIU D, WANG Y J. Hybrid enhancement fault diagnos-

is method of aircraft under functional crosslink condition[J]. Measurement & Control Technology, 2022, 41(10): 107-113(in Chinese).

- [2] HECKERMAN D, BREESE J S, ROMMELSE K. Decision-theoretic troubleshooting[J]. Communications of the ACM, 1995, 38: 49-57.
- [3] PEARL J. Bayesian networks: A model of self-activated memory for evidential reasoning[C]//Proceedings of the 7th Conference of the Cognitive Science Society. Irvine: Cognitive Science Society, 1985: 15-17.
- [4] SKAANNING C, JENSEN F V, KJÆRULFF U. Printer troubleshooting using Bayesian networks[M]. Berlin: Springer, 2003: 367-380.
- [5] JENSEN F V, KJÆRULFF U, KRISTIANSEN B, et al. The SAC-SO methodology for troubleshooting complex systems[J]. Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 2001, 15(4): 321-333.
- [6] 于劲松, 刘浩, 万九卿, 等. 贝叶斯网络结合决策理论的向前多步 排故策略[J]. 北京航空航天大学学报, 2014, 40(3): 298-303.
 YU J S, LIU H, WAN J Q, et al. Bayesian networks and decision theory-based forward multi-step troubleshooting strategy[J]. Journal of Beijing University of Aeronautics and Astronautics, 2014, 40(3): 298-303(in Chinese).
- [7] HUANG Y P, WANG Y S, ZHANG R J. Fault troubleshooting using Bayesian network and multicriteria decision analysis[J]. Advances in Mechanical Engineering, 2014, 6: 282013.
- [8] VIANNA W O L, RODRIGUES L R, YONEYAMA T, et al. Troubleshooting optimization using multi-start simulated annealing[C]//Proceedings of the 2016 Annual IEEE Systems Conference. Piscataway: IEEE Press, 2016: 1-6.
- [9] DE OLIVEIRA L S, RODRIGUES L R, YONEYAMA T. A comparative study of metaheuristics applied to troubleshooting optimization problems[C]//Proceedings of the XLIX Brazilian Symposium on Operational Research. Blumenau: SOBRAPO, 2017: 1783-1794.
- [10] COELHO D B P, RODRIGUES L R. A chaotic inertia weight

TLBO applied to troubleshooting optimization problems[C]//Proceedings of the 2020 IEEE Congress on Evolutionary Computation. Piscataway: IEEE Press, 2020: 1-8.

- [11] HUANG C W, LI Y X, YAO X. A survey of automatic parameter tuning methods for metaheuristics[J]. IEEE Transactions on Evolutionary Computation, 2020, 24(2): 201-216.
- [12] 李凯文,张涛,王锐,等.基于深度强化学习的组合优化研究进展[J].自动化学报,2021,47(11):2521-2537.

LI K W, ZHANG T, WANG R, et al. Research reviews of combinatorial optimization methods based on deep reinforcement learning[J]. Acta Automatica Sinica, 2021, 47(11): 2521-2537(in Chinese).

[13] 顾一凡. 基于强化学习的组合优化综述[J]. 软件导刊, 2021, 20(9): 74-77.

GU Y F. A survey on reinforcement learning for combinatorial optimization[J]. Software Guide, 2021, 20(9): 74-77(in Chinese).

- ZHANG Z Z, WU Z Y, ZHANG H, et al. Meta-learning-based deep reinforcement learning for multiobjective optimization problems[J].
 IEEE Transactions on Neural Networks and Learning Systems, 2022, 34(10): 7978-7991.
- [15] OREN J, ROSS C, LEFAROV M, et al. SOLO: Search online, learn offline for combinatorial optimization problems[C]//Proceedings of the International Symposium on Combinatorial Search. Palo Alto: AAAI Press, 2021, 12(1): 97-105.
- [16] ALMASAN P, SUÁREZ-VARELA J, RUSEK K, et al. Deep reinforcement learning meets graph neural networks: Exploring a routing optimization use case[J]. Computer Communications, 2022, 196: 184-194.
- [17] OTTOSEN T J. Solutions and heuristics for troubleshooting with dependent actions and conditional costs[D]. Aalborg: Aalborg University, 2012: 33-43.
- [18] MZOUGUI I, CARPITELLA S, CERTA A, et al. Assessing supply chain risks in the automotive industry through a modified MCDMbased FMECA[J]. Processes, 2020, 8(5): 579.

- [19] 李俊杰, 王尧, 张强, 等. 基于视情维修的涡轴发动机维修保障辅助决策体系研究[J]. 计算机测量与控制, 2021, 29(6): 205-211.
 LI J J, WANG Y, ZHANG Q, et al. Research on auxiliary decision-making system of turboshaft engine maintenance support based on condition-based maintenance[J]. Computer Measurement & Control, 2021, 29(6): 205-211(in Chinese).
- [20] 邱锡鹏. 神经网络与深度学习[M]. 北京: 机械工业出版社, 2020: 328-353.
 QIU X P. Neural networks and deep learning[M]. Beijing: China

Machine Press, 2020: 328-353(in Chinese).

- [21] 张秦浩, 敖百强, 张秦雪. Q-learning 强化学习制导律[J]. 系统工程与电子技术, 2020, 42(2): 414-419.
 ZHANG Q H, AO B Q, ZHANG Q X. Reinforcement learning guidance law of Q-learning[J]. Systems Engineering and Electron-
- [22] WATKINS C J, DAYAN P. Q-learning[J]. Machine Learning, 1992, 8(3): 279-292.

ics, 2020, 42(2): 414-419(in Chinese).

- [23] 龚铭凡, 徐海祥, 冯辉, 等. 基于改进 Q-Learning 的智能船舶局部 路径规划[J]. 船舶力学, 2022, 26(6): 824-833.
 GONG M F, XU H X, FENG H, et al. Ship local path planning based on improved Q-learning[J]. Journal of Ship Mechanics, 2022, 26(6): 824-833(in Chinese).
- [24] 黄鑫陈,陈光祖,郑敏,等. 基于 Q-learning 的飞行自组织网络 QoS 路由方法[J]. 中国科学院大学学报, 2022, 39(1): 134-143. HUANG X C, CHEN G Z, ZHENG M, et al. Q-learning based QoS routing for high dynamic flying Ad Hoc networks[J]. Journal of University of Chinese Academy of Sciences, 2022, 39(1): 134-143(in Chinese).
- [25] LOW E S, ONG P, CHEAH K C. Solving the optimal path planning of a mobile robot using improved Q-learning[J]. Robotics and Autonomous Systems, 2019, 115: 143-161.
- [26] HUANG S, ONTAÑÓN S. A closer look at invalid action masking in policy gradient algorithms[C]//Proceedings of the the International FLAIRS Conference. Gainesville: Library Press, 2022, 35: 1-6.

Complex equipment troubleshooting strategy generation based on Bayesian networks and reinforcement learning

LIU Baoding¹, YU Jinsong^{1, *}, HAN Danyang², TANG Diyin¹, LI Xin³

(1. School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China;

2. School of Instrumentation and Optoelectronic Engineering, Beihang University, Beijing 100191, China;

3. China Academy of Launch Vehicle Technology, Beijing 100076, China)

Abstract: To shorten the time spent and reduce the troubleshooting cost of traditional heuristic methods, a method of generating a troubleshooting strategy based on reinforcement learning (RL) and Bayesian networks (BN) is proposed for complex equipment. BN is used for the expression of knowledge to make better use of model knowledge of complex equipment. To get closer to the real scenario, the fault probability in the failure mode, effect, and critical analysis (FMECA) of complex equipment is converted and used as a prior probability in BN. A paradigm of converting troubleshooting problems into RL problems is proposed to generate a troubleshooting strategy by using the decision process of RL. The observation-action pair (O-A) is introduced to reduce the scale of the RL problem and the action masking is set to deal with dynamic action space. Simulation findings demonstrate the superiority of the proposed BN-RL method by demonstrating its remarkable performances compared to standard heuristic methods based on the proposed metrics.

Keywords: reinforcement learning; Bayesian networks; troubleshooting strategy generation; complex equipment; dynamic action space

Received: 2022-05-31; Accepted: 2022-07-22; Published Online: 2022-08-01 11:00 URL: link.cnki.net/urlid/11.2625.V.20220729.1212.001

Foundation items: National Key R&D Program of China (2018YFB1403300); National Natural Science Foundation of China (51875018, 71701008)

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn DOI: 10.13700/j.bh.1001-5965.2022.0450

基于 anchor-free 的光学遥感舰船关重部位检测算法

张冬冬, 王春平, 付强*

(陆军工程大学石家庄校区电子与光学工程系,石家庄 050003)

摘 要:针对基于深度学习的遥感舰船检测算法存在精细化程度不足、检测效率低的问题,提出一种基于 anchor-free 的光学遥感舰船关重部位检测算法。所提算法以全卷积的单阶段目标 检测 (FCOS)算法为基准,在主干网络中引入全局上下文模块,提高网络的特征表达能力;为更好 地描述目标的方向性,在预测阶段构建了具有方向表征能力的回归分支;对中心度函数进行优化, 使其具备方向感知和自适应能力。实验结果表明:在自建舰船关重部位数据集和 HRSC2016 上,所 提算法的平均精度 (AP) 比 FCOS 算法有显著提升;与其他算法相比,所提算法在检测速度和检测 精度上均表现优越,具有较高的检测效率。

关 键 词:深度学习;遥感图像; anchor-free; 舰船检测; 关重部位检测; 全卷积单阶段检测 中图分类号: TP753

文献标志码: A 文章编号: 1001-5965(2024)04-1365-10

目标检测是计算机视觉和模式识别领域的研究热点,被广泛地应用于军用和民用领域,诸如定 点打击、无人驾驶、智能监控等^[1-3]。近年来,随着 航天、遥感、传感器等技术的发展,遥感图像的质 量、分辨率、时效性等得到了大幅提升,利用遥感 图像进行目标检测与识别引起了学者的广泛关注。

舰船是海域中的重要目标,对其进行有效监测 对于维护国家安全、提高海上作战能力具有重要意 义^[4]。合成孔径雷达(synthetic aperture radar, SAR)^[5] 可以全天候、全时段、多角度的对目标实施监测, 通常被用于舰船检测任务。与 SAR 图像相比,光 学遥感图像包含丰富的细节信息,解译难度相对较 小。随着光学遥感成像技术的发展成熟,图像的分 辨率和质量得到了大幅提高,可以更加直观地呈现 目标的纹理、颜色和形状信息,在舰船检测上拥有 巨大优势。因此,本文以光学遥感图像为基础,对 舰船检测进行深入研究。

深度学习技术的飞速发展极大地推动了目标 检测领域的进步,基于深度学习的目标检测算法是

目前的主流算法。卷积神经网络(convolutional neural network, CNN)具有强大的特征提取能力,可以通过 学习提取高层次的抽象特征,为目标检测任务提供 所需的信息,使检测的精度得以提升。基于 CNN 的目标检测算法主要分为2类:①以 R-CNN(Region-CNN)^[6]、Faster R-CNN^[7]、Mask R-CNN^[8]等为代表的 双阶段网络检测算法; ②以 SSD(single shot multibox detector)^[9]、YOLO^[10]、RetinaNet^[11]等为代表的单阶 段网络检测算法。总体来说,双阶段检测网络检测 算法检测精度高,但速度慢;单阶段网络检测算法 速度快,但精度偏低。为提高光学遥感舰船的检测 精度,研究人员进行了深入研究并取得了突破性进 展。文献 [12] 通过增加特征融合模块和设计聚焦 分类损失函数来提高复杂场景下舰船的检测精 度。文献 [13] 针对遥感舰船的大长宽比和任意朝 向特性,在YOLOv3算法中引入方向参数以更好地 描述舰船目标,并且设计了一种旋转卷积集成模块 来提高算法对角度的敏感度。文献 [14] 通过由粗 到细逐步精化的方式实现快速、准确地检测舰船

收稿日期: 2022-05-31; 录用日期: 2022-08-08; 网络出版时间: 2022-08-18 11:07

网络出版地址: link.cnki.net/urlid/11.2625.V.20220817.1530.002

^{*}通信作者. E-mail: 1418748495@qq.com

引用格式: 张冬冬, 王春平, 付强. 基于 anchor-free 的光学遥感舰船关重部位检测算法 [J]. 北京航空航天大学学报, 2024, 50 (4): 1365-1374. ZHANG D D, WANG C P, FU Q. Ship's critical part detection algorithm based on anchor-free in optical remote sensing [J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50 (4): 1365-1374 (in Chinese).

目标。

上述算法针对舰船目标取得了较好的检测效 果,但仍存在以下不足:①将舰船视为整体目标,未 对其进行更加精细化的检测;②均是基于锚框 (anchor-based)的检测算法,锚框本身所存在的弊端 导致算法的训练和检测效率低下。下面将对上述 2点不足进行详细分析。

舰船是一类多层级的目标,一般由主船体、雷 达、控制室、油料舱和弹药舱等子部分构成。雷 达、控制室、油料舱、弹药舱和动力舱是舰船上较 为重要的组成部分,即为关重部位。为满足一些特 殊任务的需求,文献 [15-16] 对舰船要害部位的检 测算法进行了研究,但研究内容未涉及光学遥感图 像,在应用范围上仍存在局限性。不同类型的舰船 在构成上有一定的差异性,比如弹药舱仅存在于有 军用用途的舰船而不会存在于商用船只上。综合 考虑舰船各部分的普遍性、重要程度及在光学遥感 图像上的可视性,本文以控制室为研究对象,对舰 船的精细化检测进行探究。

Anchor-based 算法的实现过程可概括为:首先, 根据数据集中目标的特点预先设置大量锚框:然 后,对这些预设锚框进行分类和偏移修正:最后,通 过非极大值抑制(non-maximum suppression, NMS)^[17] 得到预测结果。通过使用预设锚框, anchorbased 算法增加了先验知识使训练更加稳定并能提 高目标召回能力,但也存在着一定的局限性:①普 适性低,预设锚框的数量、大小、长宽比等超参数 需要根据数据集中目标的特性进行人为设定,使其 难以适用于多种类型的数据集;②泛化能力弱, anchor-based 算法的性能对预设锚框的设置非常敏 感,并且由于锚框的预设类型有限使算法不能较好 的检测长宽比大的目标;③计算量大,为尽可能覆 盖所有目标,需要在图像的每个像素点都设置一组 相同的锚框,使整体的锚框数量非常庞大,但一副 图像所包含的目标数量有限,仅有一小部分锚框中 包含目标,导致正负样本严重不平衡,而且数量庞 大的锚框会占用更多的内存并增加计算量,严重影 响了算法的检测速度。针对以上问题,研究人员提 出了无锚框(anchor-free)的检测算法,一大批相关 算法应运而生。文献 [18] 将目标视为成对的关键 点,通过预测目标的左上角和右下角关键点来定位 目标。文献 [19] 在 ConerNet^[18] 的基础上增加了中 心关键点,使角点匹配更加准确,从而提高了检测 精度。文献 [20] 成功的将 anchor-free 部署在 YOLO 算法上,并取得了不错的检测效果。文献 [21] 提出 了一种全卷积单阶段目标检测 (fully convolutional one-stage object detection, FCOS)算法,该算法的处

理方式类似于语义分割,采用逐像素预测的方式实现目标的分类和回归,并通过构建中心度分支来抑制低质量预测框。与 anchor-based 算法相比, anchor-free 算法摆脱了锚框机制的制约, 对于不同的任务和目标拥有更强的泛化能力,更好的兼顾了检测的精度和速度。因为预测时缺少锚框作为基准, anchor-free 算法的回归范围较大,更易产生虚警,这一现象在复杂背景下尤为明显。

基于以上分析,本文以 FCOS 算法^[21] 为基准, 提出了一种针对光学遥感舰船关重部位的检测算 法。该算法先将全局上下文模块(global context block, GCB)^[22] 融入主干网络,提升网络的特征提取 能力;其次,改进回归分支,并使用"中心到角点" 的预测策略^[23],使网络具有描述目标方向的能力, 从而提高定位的精准度;另外,对中心度函数进行 优化,使其具有方向感知和自适应能力。

1 本文算法

1.1 FCOS 算法

FCOS 算法^[21] 是一种基于 anchor-free 的全卷积 算法,其结构与 RetinaNet^[11]类似,均由 ResNet+特 征金字塔网络 (feature pyramid networks, FPN)+Head 组成,不同的是 FCOS 算法^[21]对 FPN 的层数进行了 扩充并增加了中心度预测分支,如图 1 所示,其中, H和W分别为高度和宽度,s为上采样。与基于 anchorbased 的单阶段检测算法相比,FCOS 算法^[21]规避了 使用锚框所带来的问题,更加简单、灵活,仅凭借后 处理 NMS 就可以得到较高的检测精度。FCOS 算 法^[21]采用了语义分割的思想,在真实目标范围内进 行逐像素的分类和回归,通过这种密集预测的方式 来保证检测的召回率。

FCOS 算法^[21] 在训练的过程中会将真实框内的 每个像素都作为正样本,由于框内不可避免的会包 含一些背景,导致 FCOS 算法^[21] 的误检和虚检概率 较高,同时预测框数量增加会降低算法的处理速 度。针对上述问题,FCOS 算法^[21] 在检测头中增加 了中心度分支,该分支的预测值与分类置信度一起 决定预测框的得分,可以使低质量的预测框被抑制 掉,从而提高检测精度。中心度的真值*c**定义为

$$c^* = \sqrt{\frac{\min(l^*, r^*)}{\max(l^*, r^*)} \cdot \frac{\min(t^*, b^*)}{\max(t^*, b^*)}}$$
(1)

式中: l*, r*, t*, b*为样本点到真实框 4 条边的距离。

1.2 优化 FCOS 算法

1.2.1 优化主干网络

与自然图像相比,遥感图像中目标所处的场景 复杂多变,目标周围环境会对检测造成干扰,这一





问题在近岸舰船的检测任务中尤为突出。此外,相 对于舰船,关重部位特征较少,更易受到环境影 响。因此,需要在特征提取的过程中从全局出发将 更多的注意力放在目标上,为此在主干网络中加入 了 GCB。本文使用 Resnet-50 作为主干网络。

GCB主要由上下文模块和转换模块组成,如 图 2 所示,其中,C为通道数。在上下文模块中,输 入特征经过 1×1 和 softmax 函数得到注意力权重, 然后作用于原始输入获取全局上下文特征。在转 换模块中,通过先降维再升维的方式获取通道间的 依赖关系。特征模块的输出与原始输入进行逐像 素相加,以这种方式将全局上下文特征整合到每个 位置的特征上。GCB 的整个作用过程可表示为

$$y_{i} = x_{i} + W_{v2} \text{ReLU} \left[\text{LN} \left(W_{v1} \sum_{j=1}^{N_{p}} \frac{e^{W_{k} x_{j}}}{\sum_{m=1}^{N_{p}} e^{W_{k} x_{m}}} x_{j} \right) \right]$$
(2)

式中: *x*和*y*分别为输入和输出; *W*_k, *W*_{v1}, *W*_{v2}分别为 图 2 中 3 个卷积的权重; ReLU(·)为激活函数; LN(·) 为层归一化。



Fig. 2 Structure of GCB

经过 GCB 处理后,特征的表达能力得到增强, 可以引导网络更多的关注目标特征,忽略图像中的 背景信息,但是考虑到关重部位的特殊性,在检测 的过程中希望借助舰船的信息来辅助定位目标,为 避免过滤掉这部分信息,将 GCB 嵌入至 Resnet-50 的残差结构中,如图 3 所示。该结构实现了特征的 重用,提高了网络对目标特征关注度,同时避免了 部分信息的丢失影响目标的检测效果。



图 3 嵌入 GCB 的残差结构 Fig. 3 Residual structure of GCB embedded

为直观的展示出将 GCB 嵌入至 Resnet-50 的作 用效果,在有无 GCB 的情况下对比 P3、P4 和 P5 层 的特征图,如图 4 所示,其中,蓝色到红色表示背景 特征逐渐减少,即目标特征逐渐显著。由图可知, 增加 GCB 后,背景特征减少,目标特征能够更好地 得到表达,同时辅助信息也得到了保留。图 4 从定 性的角度证明了增加 GCB 的有效性,在消融实验 中将会从定量的角度加以验证。

1.2.2 回归分支

关重部位作为舰船的一部分同样具有方向性, 为更加准确的描述该目标,对 FCOS 算法^[21] 检测头 的回归分支进行改进,构建了中心点和角点 2 个回 归分支,如图 5 所示。中心点和角点回归分支分别 负责预测目标中心和中心到角点的偏移量,然后利 用"中心到角点"的预测策略来定位目标。"中心 到角点"的策略将预测分为 2 个阶段,第 1 阶段预 测目标的中心,第 2 阶段从预测的中心回归角点 (见图 6),该策略可表示为

$$\begin{cases} \tilde{c} = \operatorname{conv}_{256 \to 2}(X) \\ c_i = \operatorname{conv}_{256 \to 8}(X) + \tilde{c} \qquad i = 0, 1, \cdots, 3 \end{cases}$$
(3)

式中: *c*和 c₀、 c₁、 c₂、 c₃分别为目标中心点和 4 个



(a) 原图 (b) P3 (有GCB) (c) P3 (无GCB) (d) P4 (有GCB) (e) P4 (无GCB) (f) P5 (有GCB) (g) P5 (无GCB) 图 4 有无 GCB 情况下的特征可视化

Fig. 4 Visualization of features with and without GCB



图 5 改进后的检测头结构

Fig. 5 Improved detection heads structure



图 6 "中心到角点"的预测策略示意图 Fig. 6 Diagram of "center to corner" prediction strategy

角点的坐标; "+" 表示广播操作, 其将 č添加到每个 角分支输出的元组中。在实际预测中, 中心点的预 测难度要低于角点, 因为其维度较低, 并且与目标 的尺寸和方向无关。因此, "中心到角点"的策略 可以先通过优化中心损失来获得目标的大致位置, 然后从中心逐步完善角点。

1.2.3 定向自适应中心度

FCOS 算法^[21] 通过引入中心度来降低低质量预 测框的权重,可以使低质量预测框更容易被抑制 掉,从而提高检测的精度,但是该中心度的计算方 式是根据水平框制定的,不适用于旋转框的情况。

因此,将中心度的计算式进行扩展,将其推广到任 意四边形,对于给定旋转框(x₀,y₀,x₁,y₁,x₂,y₂,x₃,y₃) 及中心点(x,y),定向中心度可由下式计算得到:

$$\begin{aligned}
 D_{\text{center-ness}}(a, b, c, d) &= \left(\frac{\min(a, c)}{\max(a, c)} \cdot \frac{\min(b, d)}{\max(b, d)}\right)^{\alpha} \\
 a &= d(x_0, y_0, x_1, y_1, x, y) \\
 b &= d(x_1, y_1, x_2, y_2, x, y) \\
 c &= d(x_2, y_2, x_3, y_3, x, y) \\
 d &= d(x_3, y_3, x_0, y_0, x, y)
 \end{aligned}$$
(4)

式中: *α*为自适应参数; *d*(·)表示中心点与 2 个相邻 角点间的归一化垂直距离, 可根据下式进行 计算:

$$d(x_{i}, y_{i}, x_{j}, y_{j}, x, y) = \frac{\left| \left(x_{j} - x_{i} \right) (y_{i} - y) - (x_{i} - x) \left(y_{j} - y_{i} \right) \right|}{\sqrt{\left(x_{j} - x_{i} \right)^{2} + \left(y_{j} - y_{i} \right)^{2}}}$$
(5)

在式(1)中,中心度被定义成了几何平均值,缺 乏对目标形状的自适应能力。通过(4)式,将中心 度推广到了任意四边形,为使其具备自适应能力, 将α定义为

$$\alpha = \frac{\min(\sqrt{(x_0 - x_1)^2 + (y_0 - y_1)^2}, \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2})}{\max(\sqrt{(x_0 - x_1)^2 + (y_0 - y_1)^2}, \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2})}$$
(6)

测试中,位置(x,y)上用于 NMS 排序的最终分数 s_{xy}由分类置信度 p_{xy}和定向自适应中心度 c_{xy}共同决定,其表达式为

$$s_{x,y} = \sqrt{p_{x,y}c_{x,y}} \tag{7}$$

为便于理解定向自适应的作用,图7为在给定 旋转框中每个位置的中心度、定向中心度和定向自 适应中心度的权重热图。对比图7(a)和图7(b)可 知,定向中心度可以更好的适应旋转框,能够与旋 转框的边缘完全对齐,并适应其长宽比。对比 图 7(b)和图 7(c)可知,在α的调解下,中心度的衰 减率变大,靠近旋转框边缘位置的权重变得更低, 中心点周围高权重的区域变大,这样更利于抑制低 质量预测框和保留高质量预测框。



Fig. 7 Heat map for weights of different center-ness

2 实 验

实验在配置为 E5-2630v2 CPU、NVIDIA GTX-1080Ti GPU(11G video memory)、64G RAM 的图像 工作站上进行,以 Ubuntu18.04 为平台,在深度学 习框架 Pytorch1.6 下编译完成。训练过程中,网络 在随机梯度下降 (stochastic gradient descent, SGD) 法 下进行 36 轮迭代训练,初始学习率为 0.002 5,批 次大小为 4。在迭代 24 和 33 轮时,学习率分别降 低 10 倍。权重衰减和动量分别设置为 0.000 1 和 0.9。为验证本文算法的有效性,在自建数据集和 HRSC2016^[24]数据集上进行了多组实验,使用每秒处 理帧数(frame per second, FPS)^[25]、平均精度(average precision, AP)^[26]及网络参数量(parameters)^[16]来评 价网络的性能。

2.1 数据集

当前用于舰船检测的光学遥感数据集,比如 HRSC2016^[24]、DIOR^[27]、DOTA^[28]等,均以整体舰船 作为目标进行标注,多用于舰船整体的检测任务。 将舰船关重部位作为研究对象时,数据集中应包含 明显的部位信息,当前所存在的数据集无法满足这 一需求。针对数据集欠缺的问题,利用谷歌地图在 一些著名港口(珍珠港、横须贺港、朴次茅斯港等) 区域下载了1015张图像,构建了CP-Ship数据集, 数据集的详细信息如表1所示,图像样例如图8所 示。CP-Ship中图像尺度为800像素×600像素,使 用 LabelImg 和 roLabelImg 对图像中舰船的关重部 位进行标注。

表 1 CP-Ship 数据集和 HRSC2016^[24] 数据集详细信息 Table 1 Details of CP-Ship dataset and HRSC2016^[24] dataset

数据集	逐	像数量	关重部位数量	舰船数量 (CP-Ship)	
	CP-Ship	HRSC2016 ^[24]	(CP-Ship)		
训练集	812	849	2 295	2 342	
测试集	203	212	561	634	
总计	1 015	1 680	2 856	2 976	

HRSC2016^[24] 是唯一一个只用于舰船检测的公 开数据集,该数据集包含 20 多类不同外观、比例和 方向的舰船,图像样例如图 8 所示。数据集图像来 源于 6 个著名港口,图像分辨率在 0.4~2 m 之间, 大小从 300 像素×300 像素到 1 500 像素×900 像素 不等,大多数大于 1 000 像素×600 像素。HRSC2016^[24] 数据集包含 1 061 张标注图像,将其按照 4:1 的比例 划分为训练集和测试集,具体信息如表 1 所示。

2.2 消融实验

为评估嵌入 GCB、改进回归分支及定向自适 应中心度对提升算法检测能力的贡献,以 FCOS 算法^[21]为基准,在 CP-Ship数据集上进行消融实 验,实验结果如表2所示。表2中,"√"表示采用对 应的改进操作;算法1是原始的 FCOS 算法^[21];算 法2是在 FCOS 算法^[21]主干网络中嵌入 GCB 所形 成的算法;算法3是按照本文算法改进 FCOS 回归 分支后所形成的算法;算法4是算法2和算法3的 集合体;算法5是本文算法。

由表2可知,FCOS算法^[21]的AP值仅有



图 8 CP-Ship 数据集和 HRSC2016^[24] 数据集图像样例

Fig. 8 Image examples of CP-Ship dataset and HRSC2016^[24]

	Table 2	Experiment results of ablation						
算法	嵌入GCB	改进回归分支	定向自适应中心度	AP/%				
算法1				63.04				
算法2	\checkmark			64.96				
算法3		\checkmark		64.34				
算法4	\checkmark	\checkmark		66.51				
算法5	\checkmark	\checkmark	\checkmark	68.56				

63.04%,虽然其多层特征上进行检测,但是仍受背 景影响,导致检测效果不佳。在主干网络融入 GCB后,AP提高了1.92%,说明在增加特征提取阶 段增加上下文信息有利于规避背景的影响,提高网 络对目标特征的表达能力。通过改进回归分支,网 络对关重部位的描述能力增强,提高了网络的检测 能力。同时嵌入 GCB 和改进回归分支比增加单一

表 2 消融实验结果模块的 AP 增量Experiment results of ablation够较好的融合。改进回归分支 定向自适应中心度 AP/%京、大利乙乙(8,6)

模块的 AP 增量大,说明 2 个模块不会相互影响,能 够较好的融合。将所有改进融合后检测的 AP 值最 高,达到了 68.56%,大幅提高了 FCOS 算法^[21] 对关 重部位的检测能力。

2024 年

2.3 对比实验

为验证本文算法的有效性,将本文算法与 Faster R-CNN^[7](旋转框和水平框)、RetinaNet^[11](旋 转框和水平框)、R3Det^[14]、CornerNet^[18]、CenterNet^[19]、 YOLOX-L^[20]、VarifocalNet^[29]、BBAVectors^[30]、SASM^[31] 及FCOS算法^[21]进行对比分析。所用对比算法的 相关代码来源于 OpenMMLab平台。各算法在 CP-Ship 测试集上检测的定量结果如表 3 所示,其中, Rotated Faster R-CNN 和 Rotated RetinaNet 算法数据 来源于 https://github.com/open-mmlabl mmrotute。表 中 TP 表示正样本被正确识别的数量, FP 表示误检

表 3 不同算法在 CP-Ship 测试集上的定量结果

Table 3	Quantitative results of different algorithms on the CP-Ship test set	
Table 3	Quantitative results of different algorithms on the CP-Ship test set	

算法	主干网络	图像大小/像素	锚框类型	ТР	FP	AP/%	FPS	Paras/10 ⁶
Faster R-CNN ^[7]	ResNet-50	800×608	水平框	411	152	65.78	19.2	41.12
Rotated Faster R-CNN	ResNet-50	800×608	旋转框	425	138	68.60	11.9	41.12
RetinaNet ^[11]	ResNet-50	800×608	水平框	387	229	60.41	23.8	36.1
Rotated RetinaNet	ResNet-50	800×608	旋转框	388	162	60.61	13.9	36.13
R3Det ^[18]	ResNet-50	800×608	旋转框	393	146	64.28	11.3	41.58
CornerNet ^[18]	Hourglass-104	511×511	无锚框	419	1 274	59.73	3.0	200.95
CenterNet ^[19]	ResNet-18	512×512	无锚框	392	136	63.23	68.7	14.21
YOLOX-L ^[20]	CSPDarkNet	640×640	无锚框	441	189	72.71	27.8	54.15
VarifocalNet ^[29]	ResNet-50	800×608	无锚框	400	262	62.98	20.3	32.48
BBAVectors ^[30]	ResNet-50	800×608	无锚框	448	198	68.31	14.5	
SASM ^[31]	ResNet-50	800×608	无锚框	397	214	63.85	13.5	36.6
FCOS ^[21]	ResNet-50	800×608	无锚框	384	101	63.04	24.1	31.84
本文算法	ResNet-50+GCB	800×608	无锚框	417	89	68.56	21.3	32.49

的负样本数量, Paras 表示模型参数量。

表 3 中 Faster R-CNN^[7] 和 RetinaNet^[11] 分别是经 典的双阶段网络和单阶段网络检测算法,通过增加 角度参数,算法的误检数量均有一定量的降低,并 日算法的参数量几乎没有变化,但算法的推理速度 均急剧下降,究其原因是增加角度参数会使锚框数 量剧增,算法的计算量大幅增加。通过比较可以发 现,在Faster R-CNN^[7] 中增加角度参数对算法检测 能力提高的效果高于 RetinaNet^[11], 说明在单阶段网 络检测算法中仅依靠增加角度参数并不能大幅提 高算法的检测能力。相比而言,虽然本文算法相对 于 FCOS 算法^[21] 增加了少量的参数量, 但是由于未 使用锚框,很好的规避了锚框对检测速度的影响。 CornerNet^[18] 使用 Hourglass-104 作为主干网络, 严 重影响了网络的推理速度,并且该网络仅依赖嵌入 向量来判断角点是否属于同一目标,容易产生误检 现象。CenterNet^[19]在CornerNet^[18]的基础上增加了 中心点预测,大大降低了误检概率,但是由于使用 RestNet-18 作为主干网络, 网络的特征提取能力不 强,导致检测效果不佳。YOLOX-L^[20]使用较多的 参数得到了不错的检测效果,但算法受背景影响较 大,存在较多的误检目标。VarifocalNet^[29]是基于 FCOS 算法^[21]的优化算法,主要用于密集目标的检 测,其对关重部位的检测的 TP 值相对于 FCOS 算法^[21] 有了明显提升,但是检测精度并没未得到提升。 BBAVectors^[30]和 SASM^[31]均是基于 anchor-free 的 旋转框检测算法,与水平框检测算法相比,更适用 于具有任意方向的关重部位目标,在检测效果上较 一般的检测算法更为优越,但在检测速度上不占优 势。FCOS 算法^[21] 通过引入中心度大幅减少了误 检目标的数量,但受背景影响,目标特征无法被充

分提取,算法的检测能力较低。本文算法在 FCOS 算法^[21]的基础上进行优化,其检测精度达到 了与 Rotated Faster R-CNN 算法相当的水平,在检测 速度上超过了单阶段旋转框检测 R3Det^[18]算法,真 正兼顾了检测的速度和精度。

为进一步验证本文算法的鲁棒性及泛化能力, 将本文算法与对比算法在 HRSC2016^[24] 数据集上进 行实验,实验结果如表4所示。由表可知,本文算 法在 HRSC2016^[24] 数据集上仍取得了较好的检测效 果,其AP值达到了84.73%:本文算法的检测速度 优于其他具有旋转框检测结果的算法, FPS 为 22.5。Rotated Faster R-CNN 和 Rotated RetinaNet 算 法通过直接增加角度参数实现旋转目标的检测,在 不同的数据集上的检测表现不一致,相比之下,本 文算法具有较强的泛化能力,在不同数据集上具有 较好的鲁棒性。与一些先进的舰船检测算法^[32-33] 相比,本文算法对舰船目标的检测效果略有不足, 究其原因是没有很好的适应目标的多尺度特性,针 对这一问题,下一步将会对网络的多尺度特征提取 及表征能力进行增强,以提高对多尺度目标的检测 能力。

*P-R*曲线是根据准确率*P*和召回率*R*的关系建 立的,可以直观的反应出模型的好坏,一般认为曲 线与坐标轴所包围的面积越大,算法的性能越好。 不同算法在 CP-Ship 测试集和 HRSC2016^[24]数据集 上的 *P-R*曲线如图 9 所示。由图可知,本文算法的 *P-R*曲线较为平稳,并且能够包围绝大多数其他算 法的 *P-R*曲线,即与坐标轴所围面积较大。因此, 与其他算法相比,本文算法的整体性能表现优异。

为说明本文算法相对于 FCOS 算法^[21]的提升, 图 10 为 2 种算法在 CP-Ship 和 HRSC2016^[24]数据

	-		0				
算法	主干网络	图像大小	锚框类型	ТР	FP	AP/%	FPS
Faster R-CNN ^[7]	ResNet-50	800×608	水平框	557	93	84.33	20.6
Rotated Faster R-CNN	ResNet-50	800×608	旋转框	558	199	81.82	16.4
RetinaNet ^[11]	ResNet-50	800×608	水平框	543	151	81.11	22.7
Rotated RetinaNet	ResNet-50	800×608	旋转框	467	66	68.10	21.2
R3Det ^[18]	ResNet-50	800×608	旋转框	535	118	82.80	16.0
CornerNet ^[18]	Hourglass-104	511×511	无锚框	521	2 128	60.71	1.9
CenterNet ^[19]	ResNet-18	512×512	无锚框	541	328	75.25	46.6
YOLOX-L ^[20]	CSPDarkNet	640×640	无锚框	567	142	87.09	27.8
VarifocalNet ^[29]	ResNet-50	800×608	无锚框	558	144	85.04	20.8
BBAVectors ^[30]	ResNet-50	800×608	无锚框	561	219	86.19	15.8
SASM ^[31]	ResNet-50	800×608	无锚框	540	431	80.40	20.0
FCOS ^[21]	ResNet-50	800×608	无锚框	510	76	78.06	25.7
本文算法	ResNet-50+GCB	800×608	无锚框	558	52	84.73	22.5

表 4 不同算法在 HRSC2016^[24] 测试集上的定量结果 Table 4 Quantitative results of different algorithms on HRSC2016^[24] test set


聚1 (b) CP-Snip数据集结来2 (c) HRSC2016⁴⁴数据集结来1 (d) HRSC2016 图 10 不同数据集上检测的可视化结果



集上的检测可视化结果。由图可知,与FCOS算法^[21]相比,本文算法更适用于任意方向密集排列的目标,并且拥有较强的检测能力,漏检风险较小。

3 结 论

1) 在主干网络中嵌入 GCB 提升了特征表达能力,提高了算法的检测能力。

 2)改进回归分支增强了算法对关重部位目标 方向的描述能力。

3)"中心到角点"的预测策略降低了网络回归的难度,提升了算法的推理速度。

 4) 定向自适应中心度抑制了低质量边框, 大幅 提高了算法的检测精度。

5) 大量实验证明了本文算法的有效性。

参考文献(References)

[1] KANJIR U, GREIDANUS H, OŠTIR K. Vessel detection and classification from spaceborne optical images: A literature survey[J].
 Remote Sensing of Environment, 2018, 207: 1-26.

[2] 刘俊琦, 李智, 张学阳. 可见光遥感图像海面目标检测技术综述
[J]. 计算机科学, 2020, 47(3): 116-123.
LIU J Q, LI Z, ZHANG X Y. Review of maritime target detection in visible bands of optical remote sensing images[J]. Computer Science, 2020, 47(3): 116-123 (in Chinese).

2024年

- [3] 李晓斌, 江碧涛, 杨渊博, 等. 光学遥感图像目标检测技术综述[J]. 航天返回与遥感, 2019, 40(4): 95-104.
 LI X B, JIANG B T, YANG Y B, et al. A survey on object technology in optical remote sensing images[J]. Spacecraft recovery & Remote Sensing, 2019, 40(4): 95-104 (in Chinese).
- 【4】 张财广, 熊博莅, 匡纲要. 光学卫星遥感图像舰船目标检测综述
 [J]. 电波科学学报, 2020, 35(5): 637-647.
 ZHANG C G, XIONG B L, KUANG G Y. A survey of ship detection in optical satellite remote sensing images[J]. Chinese Journal of Radio Science, 2020, 35(5): 637-647 (in Chinese).
- [5] 张永胜, 计一飞, 董臻. 时-空变化的背景电离层对星载合成孔径 雷达方位向成像的影响分析[J]. 电子与信息学报, 2021, 43(10): 2781-2789.

ZHANG Y S, JI Y F, DONG Z. Research on background ionospheric impacts imposed by spatio-temporal variations on spaceborne synthetic aperture radar azimuth imaging[J]. Journal of Electronics & Information Technology, 2021, 43(10): 2781-2789 (in Chinese).

- [6] GIRSHICK R. Fast R-CNN[C]//Proceedings of the 2015 IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2015: 1440-1448.
- [7] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [8] HE K M, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN[C]//Proceedings of the 2017 IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2017: 2980-2988.
- [9] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot MultiBox detector[C]//Proceedings of the European Conference on Computer Vision. Berlin: Springer, 2016: 21-37.
- [10] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger[C]//Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 6517-6525.
- [11] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//Proceedings of the 2017 IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2017: 2999-3007.
- [12] 史文旭, 江金洪, 鲍胜利. 基于特征融合的遥感图像舰船目标检测算法[J]. 光子学报, 2020, 49(7): 57-67.
 SHI W X, JIANG J H, BAO S L. Ship detection method in remote sensing image based on feature fusion[J]. Acta Photonica Sinica, 2020, 49(7): 57-67 (in Chinese).
- [13] 吴止锾,李磊,高永明. 遥感图像舰船检测的旋转卷积集成 YOLOv3 模型[J]. 计算机工程与应用, 2019, 55(22): 146-151.
 WU Z H, LI L, GAO Y M. Rotation convolution ensemble YOLOv3 model for ship detection in remote sensing images[J]. Computer Engineering and Applications, 2019, 55(22): 146-151 (in Chinese).
- [14] YANG X, YAN J C, FENG Z M, et al. R3Det: Refined single-stage detector with feature refinement for rotating object[EB/OL]. (2020-12-08) [2020-12-10]. http://arxiv.org/abs/1908.05612.pdf.
- [15] 赵微, 惠斌, 张玉晓. 红外舰船目标的要害点检测算法[J]. 红外与 激光工程, 2014, 43(1): 48-52.
 ZHAO W, HUI B, ZHANG Y X. Detection algorithms of aim points of IR warship images[J]. Infrared and Laser Engineering, 2014, 43(1): 48-52 (in Chinese).
- [16] 李晨瑄,顾佼佼,王磊,等.多尺度特征融合的 Anchor-Free 轻量 化舰船要害部位检测算法[J].北京航空航天大学学报,2022, 48(10):2006-2019.

LI C X, GU J J, WANG L, et al. Warship's vital parts detection algorithm based on lightweight anchor-free network with multi-scale feature fusion[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(10): 2006-2019 (in Chinese).

- [17] 张长伦,张翠文,王恒友,等. 基于注意力机制的 NMS 在目标检测中的研究[J]. 电子测量技术, 2021, 44(19): 82-88.
 ZHANG C L, ZHANG C W, WANG H Y, et al. Research on non-maximum suppression based on attention mechanism in object detection[J]. Electronic Measurement Technology, 2021, 44(19): 82-88 (in Chinese).
- [18] LAW H, DENG J. CornerNet: Detecting objects as paired keypoints[C]//Proceedings of the European Conference on Computer Vis-

ion. Berlin: Springer, 2018: 765-781.

- [19] DUAN K W, BAI S, XIE L X, et al. CenterNet: Keypoint triplets for object detection[C]//Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2020: 6568-6577.
- [20] GE Z, LIU S T, WANG F, et al. YOLOX: Exceeding YOLO series in 2021[EB/OL]. (2021-08-06) [2021-08-10]. http://arxiv.org/abs/ 2107.08430.pdf.
- [21] TIAN Z, SHEN C H, CHEN H, et al. FCOS: A simple and strong anchor-free object detector[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(4): 1922-1933.
- [22] CAO Y, XU J R, LIN S, et al. GCNet: Non-local networks meet squeeze-excitation networks and beyond[C]//Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision Workshop. Piscataway: IEEE Press, 2020: 1971-1980.
- [23] LANG S, VENTOLA F, KERSTING K. DAFNe: A one-stage anchor-free approach for oriented object detection[EB/OL]. (2022-05-30) [2022-06-02]. http://arxiv.org/abs/2109.06148.pdf.
- [24] LIU Z K, YUAN L, WENG L B, et al. A high resolution optical satellite image dataset for ship recognition and some new baselines[C]//Proceedings of the 6th International Conference on Pattern Recognition Applications and Methods. Porto: Science and Technology Publications, 2017.
- [25] MA Y, CHAI L, JIN L Z, et al. AVS-YOLO: Object detection in aerial visual scene[J]. International Journal of Pattern Recognition and Artificial Intelligence, 2022, 36(1): 2250004.
- [26] GAO X, RAM S, PHILIP R C, et al. Selecting post-processing schemes for accurate detection of small objects in low-resolution wide-area aerial imagery[J]. Remote Sensing, 2022, 14(2): 255.
- [27] LI K, WAN G, CHENG G, et al. Object detection in optical remote sensing images: A survey and a new benchmark[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 159: 296-307.
- [28] XIA G S, BAI X, DING J, et al. DOTA: A large-scale dataset for object detection in aerial images[C]//Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 3974-3983.
- [29] ZHANG H Y, WANG Y, DAYOUB F, et al. VarifocalNet: An IoUaware dense object detector[C]//Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2021: 8510-8519.
- [30] YI J R, WU P X, LIU B, et al. Oriented object detection in aerial images with box boundary-aware vectors[C]//Proceedings of the 2021 IEEE Winter Conference on Applications of Computer Vision. Piscataway: IEEE Press, 2021: 2149-2158.
- [31] HOU L P, LU K, XUE J A, et al. Shape-adaptive selection and measurement for oriented object detection[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Washton D. C.: AAAI, 2022, 36(1): 923-932.
- [32] YANG Y, PAN Z X, HU Y X, et al. CPS-det: An anchor-free based rotation detector for ship detection[J]. Remote Sensing, 2021, 13(11): 2208.
- [33] ZHANG X R, WANG G C, ZHU P, et al. GRS-det: An anchor-free rotation ship detector based on gaussian-mask in remote sensing images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 59(4): 3518-3531.

Ship's critical part detection algorithm based on anchor-free in optical remote sensing

ZHANG Dongdong, WANG Chunping, FU Qiang*

(Department of Electronic and Optical Engineering, People Liberation Army Engineering University-Shijiazhuang, Shijiazhuang 050003, China)

Abstract: Low detection effectiveness and inadequate refinement plague the existing deep learning-based remote sensing ship detection technique. To address the above problems, an optical remote sensing ship critical part detection algorithm based on anchor-free is proposed. The proposed algorithm takes fully convolutional one-stage object detection (FCOS) as the benchmark algorithm and introduces a global context module in the backbone network to improve the feature representation capability of the network. In the prediction step, a regression branch with orientation representation capabilities is built to more accurately describe the orientation of targets. The centrality function is optimized to make it direction-aware and adaptive. The experimental results show that the average precision (AP) of the proposed algorithm is significant improved over FCOS algorithm on the self-built ship critical part dataset and HRSC2016, respectively. Compared with other algorithms, the proposed algorithm has superior performance in both detection speed and detection accuracy and has high detection efficiency.

Keywords: deep learning; remote sensing; anchor-free; ship detection; critical part detection; fully convolutional one-stage object detection

Received: 2022-05-31; Accepted: 2022-08-08; Published Online: 2022-08-18 11:07

URL: link.cnki.net/urlid/11.2625.V.20220817.1530.002

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn **DOI**: 10.13700/j.bh.1001-5965.2022.0462

面向无人驾驶车辆的行车安全场模型构建方法

谢楚安,任羿,杨德真*,冯强,孙博,王自力

(北京航空航天大学可靠性与系统工程学院,北京100191)

摘 要: 针对全面系统评价无人驾驶车辆行车安全方法欠缺的问题,提出一种改进的无人 驾驶车辆行车安全场模型。考虑无人驾驶车辆复杂道路因素,人工智能 (AI) 系统感知、决策、控 制 3 大模块的特性对无人驾驶车辆行车安全的影响,基于胡克定律,建立动态势能场及安全行为场 相结合的无人驾驶车辆行车安全场数学模型,以此表征道路上静止物体、运动物体与 AI 系统自身 等因素造成的行车风险。结合典型行驶场景的行车安全分析验证所提模型的正确性和可用性。

关键词:无人驾驶车辆;行车安全场;风险;行车安全评价;动态势能场

中图分类号: X981

文献标志码: A 文章编号: 1001-5965(2024)04-1375-09

随着科技的发展及自主化、智能化的不断进步,无人驾驶车辆的发展也逐渐成为了衡量一个国家科技创新和技术水平的重要标志。但近年来,特斯拉、谷歌、Uber等企业在进行无人驾驶车辆行车试验时屡发交通事故,导致乘客、司机受伤甚至丧命^[1]。无人驾驶车辆行车安全至关重要,已成为国内外关注热点^[2-3]。

在无人驾驶车辆的驾驶行为中,对车辆行驶过 程中的交通态势及周遭环境进行精确的风险评估, 使车辆有效地预防碰撞可实现安全行驶。因此,全 面系统地评价行车安全是其关键技术。经调研,国 内外学者对此提出了许多不同的方法,传统的行车 安全评价方法通常是基于车辆的运动参数及历史 数据分析对其进行安全评价,如根据车辆的行驶状 态及车辆的相对运动关系计算碰撞时间(time-tocollision, TTC)^[4]、车头时距(time headway, THW)^[5]、 行车安全距离等^[6],以此作为安全约束实现碰撞预 防。但此类评价方法指标单一,难以描述不同道路 条件、周边行人车辆动态交互性行为等复杂潜在行 车风险因素对行车安全的影响,局限性较大^[7]。随 着人工智能(arti-ficial intelligence, AI)、传感技术的 发展,一些学者^[8-9]开始借鉴人工势能场(artificial potential field, APF)^[10]应用于机器人避撞的思路, 提出将势能场理论融入到车辆行车安全评估中,如 重力场^[11-12]、弹性场^[13]等。但由于场量单一,其只 可应用于特定的驾驶场景。针对此问题,基于改进 APF 建立的行车安全场模型^[14]具有能动态反映交 通态势变化,同时考虑复杂行车风险因素对行车安 全影响的优点,可以相对准确地评价车辆在复杂交 通场景中的行车安全,描述人-车-路三者之间的耦 合机制。同时,行车安全场模型在车辆协同控 制^[15]、安全路径决策规划^[16]、车辆最优换道策略^[17-18] 等方面得到了重要的应用。

但现有的行车安全场模型在无人驾驶车辆的 应用上存在限制,例如,在其AI系统是否能检测并 适应复杂的交通环境^[19-21](如雨雪、坡度、弯曲道路 等),是否能对种类繁杂、具有高度动态不确定性及 开放性的周遭环境做出正确的交互^[22-24]等问题上 仍存在挑战。其中最重要的问题之一是,无人驾驶 车辆是一个集感知、决策、控制于一体的综合系

收稿日期:2022-06-08;录用日期:2022-09-11;网络出版时间:2022-11-0311:17 网络出版地址:link.cnki.net/urlid/11.2625.V.20221103.0856.001

*通信作者. E-mail: 09967@buaa.edu.cn

基金项目:可靠性与环境工程技术国家级重点实验室基金 (6142004210108)

引用格式: 谢楚安, 任羿, 杨德真, 等. 面向无人驾驶车辆的行车安全场模型构建方法 [J]. 北京航空航天大学学报, 2024, 50 (4): 1375-1383. XIE C A, REN Y, YANG D Z, et al. Construction method of driving safety field model for unmanned vehicles [J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50 (4): 1375-1383 (in Chinese).

统,在 AI 系统代替人类对车辆进行决策控制时,其 对无人驾驶车辆行车安全产生的影响同样不可忽 视。因无人驾驶车辆不存在可执行决策控制操作 的驾驶员,现有的行车安全场中与"人"有关的"行 为场"无法完全描述无人驾驶车辆 AI 系统的控制 行为。然而,现有对于无人驾驶车辆行车安全场的 研究^[25]大多直接排除了原有行车安全场中驾驶员 对行车安全施加的"行为场"的影响。因此,需综 合评估感知、决策、控制 3 方面对行车安全的风 险,对无人驾驶车辆的行车安全进行更为全面的 评估。

综上,为全面系统地评价无人驾驶车辆的行车 安全,本文拟针对复杂的驾驶环境及无人驾驶车辆 特性,以有人车行车安全场为基础,提出一种改进 的无人驾驶车辆行车安全场模型,并给出无人驾驶 车辆典型行驶场景的行车安全场建模案例,为无人 驾驶车辆安全性、可靠性设计提供依据,为无人驾 驶车辆的行车安全分析提供方法支持。

1 无人驾驶车辆行车安全场建模

结合无人驾驶车辆 AI 系统自主感知、控制、 决策的特征,基于胡克定律对现有行车安全场模型 做了改进,本文模型由道路上静止物体及运动物体 形成的动态势能场和无人驾驶车辆安全行为场组 成,其数学模型可表示为

 $\boldsymbol{E}_{\mathrm{S}} = \boldsymbol{E}_{\mathrm{V}} + \boldsymbol{E}_{\mathrm{D}} \tag{1}$

式中: Es为无人驾驶车辆行车安全场场强; Ev为动 态势能场场强; Eo为安全行为场场强。

1.1 动态势能场

动态势能场是用于表征道路上静止环境及运 动目标对行车安全影响的物理场。道路上的静止 物体主要包括靠边的车辆、施工区域、路障、开启 的井盖等。道路上的运动物体主要包括运动中的 无人驾驶车辆、行人、非机动车等。动态势能场场 强的大小和方向主要由物体之间的矢量距离、物理 类型、物体质量、运动状态和道路条件决定。

在无人驾驶车辆行驶时,会倾向于跟前方车辆 保持一个安全距离。当与前车的车间距离低于此 安全距离时,会控制车辆减速以避免与前车发生可 能的碰撞。此类现象与胡克定律现象相似,即弹簧 被压缩时,会产生一个反向的力,压缩的距离越大, 力也越大。如图1所示。

基于此,本文采用车辆间距与安全距离的差值 *l*-|*r_{ij}*|类比弹簧压缩量,并结合道路上物体的特性 及影响因素提出了物体*i*(*x_i*,*y_i*)在其周围环境条件下



2024 年

图 1 前后车车距与弹簧类比示意图

Fig. 1 Analogous diagram of front and rear car distance and spring compression

形成的动态势能场模型,如式(2)所示:

$$\boldsymbol{E}_{V_{ij}} = \boldsymbol{\Omega}_{i} \left(v_{i}^{2} a_{i} + 1 \right) \cos \theta_{i} \boldsymbol{Q}_{z} \left(l - \left| \boldsymbol{r}_{ij} \right| \right) \frac{\boldsymbol{r}_{ij}}{\left| \boldsymbol{r}_{ij} \right|}$$
(2)

式中: (x_i, y_i) 为物体 *i*的质心坐标; $r_{ij} = [x_{j} - x_i, y_j - y_i]$ 为物体 *i* 与 (x_j, y_j) 之间的矢量距离; $\Omega_i v_i^2$ 为物体 *i* 的等效动能; a_i 为物体 *i* 的加速度; θ_i 为无 人驾驶车辆相对于该物体的方向矢量与物体速度 方向的夹角; Q_z 为无人驾驶车辆所在第 *z* 段道路条 件的影响因子; Ω_i 为物体 *i* 的等效质量; *l*为无人驾 驶车辆与前方物体的安全距离, 且 $|r_{ij}| = l$ 始终应 满足 $|r_{ij}| \leq l$ 的关系式, $\frac{r_{ij}}{|r_{ij}|}$ 代表势能沿梯度下降的 方向; E_{V_ij} 表示物体 *i* 在一定道路条件下对整体交 通环境造成的风险程度, 数值越大, 场强越大, 表示 其造成的行车风险也越大。

等效质量*Q*_i及道路条件影响因子*Q*_z的具体表 达式将于 2.1 节及 2.2 节详细介绍。

1.2 安全行为场

近年的无人驾驶车辆事故基本归因于对周围 环境理解错误或对其他车辆行为判断错误。因此, 无人驾驶车辆不仅要在感知方面检测准确,还应在 正确决策的基础上做出精确的控制才能确保行车 安全。无人驾驶车辆的驾驶动作是由 AI 系统进行 控制的,分别分为感知、决策、控制 3 个方面。

感知是由车辆传感器来探测车辆周围的行驶 环境,如障碍物的位置、形状、类别及速度信息。 决策是指无人驾驶车辆通过得到的交通环境信息, 考虑周边交通态势及让行规则,与无人驾驶库中的 经验知识进行匹配,选择适合当下的驾驶行为。无 人驾驶车辆的控制主要是用于接收决策后的驾驶 指令,控制车辆稳定运行同时保证控制精度。

因此,基于无人驾驶车辆的特性,建立无人驾驶车辆的安全行为场,其是用于表征无人驾驶车辆 自身 AI 系统性能对行车安全影响的物理场。定义 无人驾驶车辆*i*(*x_i*,*y_i*)在其周围环境条件下形成的安 全行为场模型如式 (3) 所示:

$$\begin{aligned} \boldsymbol{E}_{\mathrm{D}_ij} &= \boldsymbol{E}_{\mathrm{V}_ij} \boldsymbol{D}_i = \boldsymbol{\Omega}_i \left(\boldsymbol{v}_i^2 \cdot \boldsymbol{a}_i + 1 \right) \cos \theta_i \boldsymbol{Q}_z \cdot \\ & \left(l - \left| \boldsymbol{r}_{ij} \right| \right) \frac{\boldsymbol{r}_{ij}}{\left| \boldsymbol{r}_{ij} \right|} \boldsymbol{D}_i \end{aligned} \tag{3}$$

式中: E_{D_ij} 为无人驾驶车辆 i 在一定道路条件下形成的安全行为场; D_i 为无人驾驶车辆 i 的 AI 驾驶风险因子, 详见 2.3 节。

2 行车风险因素构建

2.1 等效质量

由物体 *i* 自身的属性所产生的潜在风险损失由 等效质量 *Q*_i表示。其中,物体 *i* 的属性包括物体的 物理质量、物体类型及速度。根据文献 [26] 中行车 安全场的定义,物体 *i* 的等效质量 *Q*_i可表示为

$$Q_i = T_u m_i \left(\lambda_0 v_i^{\lambda_1} + \lambda_2 \right) \tag{4}$$

式中: T_u 为无人驾驶车辆与道路上类型 u 的物体相 撞产生的风险大小,其中,道路上的物体可分为行 人、非机动车、车辆、动物及施工区等 5 种类型,以 人员受伤或死亡的可能性大小作为风险损失的依 据对 T_u 赋值; m_i 为物体 i的实际质量; $\lambda_0 v_i^{\lambda_1}$ + λ_2 为包含速度的表达式,用于表征速度对风险的影 响,其中, λ_0 为速度的系数, λ_1 为速度的幂指数, λ_2 为常数, λ_0 、 λ_1 、 λ_2 的实际值可根据现实交通事 故数据拟合得到。理论上,道路上的物体速度越 大,造成的风险也会越大。因此,可采用速度的幂 函数项来表示行车风险的影响。

2.2 道路影响因子

第 z 个路段的道路条件对行车产生的潜在风险 主要与路面附着系数、道路曲率、坡度和能见度有 关^[27], 其影响因子 Q_z 可表示为

 $Q_{z} = Q_{z}(\mu_{z},\rho_{z},\tau_{z},\delta_{z}) = f_{\mu}(\mu_{z})f_{\rho}(\rho_{z})f_{\tau}(\tau_{z})f_{\delta}(\delta_{z})$ (5) 式中: μ_{z} 为路面附着系数; ρ_{z} 为道路曲率; τ_{z} 为道路 坡度; δ_{z} 为道路能见度。 $f(\cdot)$ 为各影响因素的风险 评价函数。

2.2.1 路面附着系数

路面附着系数是车轮和路面之间的静摩擦系数。路面附着系数越大,车辆就越不容易打滑。在 日常行驶中,雨天雪天的路面附着系数很低,车辆 制动力降低,车辆较易打滑,行车风险增加。

根据实验的精度需求,将实际行驶道路平均分为 m 个 路 段, 令 各 个 路 段 的 道 路 附 着 系 数 为 $\mu_1, \mu_2, \cdots, \mu_z, \cdots, \mu_m$ 。在实际交通环境中,最理想路 面条件为干燥水泥路面,其附着系数较大,车辆不 易打滑。因此,令标准路面附着系数 μ^* 为干燥水泥路面的附着系数,令实际驾驶道路中第 z 段道路的 附着系数为 $\mu_z(z = 1, 2, \cdots, m, \mu_z > 0)$ 。得到路面附着

系数风险评价函数为

$$f_{\mu}(\mu_z) = \frac{\mu}{\mu_z} \tag{6}$$

实际驾驶道路第 z 段道路的附着系数μ₂可通过 测量车轮制动抱死时的制动力,并除以车轮的载荷 而求得。

2.2.2 道路曲率因子

实际驾驶的道路并非完全平直,存在直道、弯 道、直角转弯、环路等情况。平直道路的行车风险 通常较低,而弯道的转弯半径越小,行车风险越 高。因此,用道路曲率衡量道路的平直程度,如 图2所示,以第z段道路的起始与道路中心线的交 点为原点,建立直角坐标系。将道路中心线过原点 的切线记为x轴,y轴与x轴垂直且相交于原点。



图 2 道路曲率直角坐标系示意图 Fig. 2 Schematic diagram of road curvature rectangular coordinate system

基于上述坐标系,设第z段道路中心线的函数 S_z表达式为

$$S_z = c_0 + c_1 x + c_2 x^2 + c_3 x^3 \tag{7}$$

式中: c_0, c_1, c_2, c_3 为待定系数。道路中心线 S_z 的函数 表达式可通过采样w个中心线上的点 s_1, s_2, \dots, s_w 的坐标值 $s_w(x_w, y_w)$ 进一步拟合确定。

由曲率公式 $\rho_z = \frac{|S''_z|}{(1+S'^2_z)^{\frac{3}{2}}}$ 可得点 $s_w(x_w, y_w)$ 处的曲率 ρ_z 如式 (8) 所示:

$$\rho_{z} = \frac{|2c_{2} + 6c_{3}x_{w}|}{\left[1 + \left(c_{1} + 2c_{2}x_{w} + 3c_{3}x_{w}^{2}\right)^{2}\right]^{\frac{3}{2}}}$$
(8)

平直道路中心线表达式的待定系数为 $c_0 = 0$, $c_1 = 0$, $c_2 = 0$, $c_3 = 0$,即平直道路中心线与x轴重合 且曲率始终为0。因此,本文以指数形式作为道路 曲率因子的风险评价函数 $f_o(\rho_2)$,即

$$f_{\rho}(\rho_z) = e^{\rho_z} = e^{\left[1 + \left(c_1 + 2c_2 x_w + 3c_3 x_w^2\right)^2\right]^{\frac{3}{2}}}$$
(9)

2.2.3 道路坡度

坡度是用于表示道路陡缓的程度,车辆在驶入 高架桥等道路时需要经过斜坡。通常来说,坡度越 大,行车风险越大。设坡面的垂直高度为 H,坡面 的水平宽度 B, 坡度示意图如图 3 所示。



Fig. 3 Schematic diagram of road slope

坡度通常用 2 点之间坡面垂直高度与坡面水 平宽度的比表示,同样将实际行驶道路分为 *m* 个路 段,以 $\tau_z(z=1,2,\cdots,m,\tau_z>0)$ 表示第 *z* 段道路的坡 度,令各个路段的道路坡度为 $\tau_1,\tau_2,\cdots,\tau_z,\cdots,\tau_m$ 。 其计算式为

$$\tau_z = \frac{H_z}{B_z} \tag{10}$$

水平道路的坡面垂直高度始终为0,即道路坡 度为0。因此,以指数形式作为道路坡度的风险评 价函数*f*_r(τ_z)为

$$f_{\tau}(\tau_z) = \mathrm{e}^{\tau_z} = \mathrm{e}^{\frac{H_z}{B_z}} \tag{11}$$

第 z 段道路的坡度τ_z可通过测量第 z 段及第 z+1 段道路之间坡面的垂直高度,并除以坡面水平 宽度求得。

2.2.4 道路能见度

能见度与行车安全密切相关,雾霾、暴雨、冰 雪等常规恶劣天气会导致道路能见距离降低,易对 前方物体距离估计错误,极可能发生多车连续追尾 的重大交通事故。道路能见度越低,行车风险 越高。

据标准 GB/T 33673-2017^[28] 知,当能见度大于 10 km 时视为能见度优。因此,令标准能见度 δ^* 为 晴朗干燥、无霾无尘粒天气的最远能见度距离,令 实际驾驶道路第 z 个路段的能见度为 $\delta_2(z = 1, 2, \cdots, m, \delta_z > 0)$,其表征了第 z 段道路上常规恶劣天气对 行车风险的影响。得到道路能见度的风险评价函 数为

$$f_{\delta}(\delta_z) = \frac{\delta^*}{\delta_z} \tag{12}$$

实际驾驶道路第 z 个路段的能见度δ.可通过大 气透射仪测量仪器测量,原理是让光束透过两固定 点之间的大气柱得到气柱的透射率,以此来推算能 见度的值。

综上,道路影响因子Q_z的表达式为

$$Q_{z} = f_{\mu}(\mu_{z}) f_{\rho}(\rho_{z}) f_{\tau}(\tau_{z}) f_{\delta}(\delta_{z}) = \frac{\mu^{*}}{\mu_{z}} e^{\rho_{z}} e^{\frac{H_{z}}{B_{z}}} \frac{\delta^{*}}{\delta_{z}}$$
(13)

2.3 人工智能驾驶风险因子

考虑到无人驾驶车辆的行驶过程主要分为感知、决策、控制3个过程。因此,AI驾驶风险相应 可表示为感知风险因子、决策风险因子及控制风险 因子。

由于感知、决策、控制过程在不同种类的无人 驾驶车辆 AI 系统上对无人驾驶车辆行车安全的影 响不完全一致,且不同种类的 AI 系统在感知、决 策、控制 3 方面的侧重有所不同。因此,应综合三 者影响,对感知、决策、控制 3 个风险因子分别设 置权重值,构成 AI 驾驶风险因子以评价无人驾驶 车辆自身 AI 系统对行车风险的影响。基于此,定 义无人驾驶车辆*i*(*x_i*,*y_i*)的 AI 驾驶风险因子*D_i*为

$$D_i = \omega_1 D_1 + \omega_2 D_2 + \omega_3 D_3 \tag{14}$$

式中: *D*₁为感知风险因子; *D*₂为决策风险因子; *D*₃为控制风险因子; ω₁,ω₂,ω₃为各影响因子的权重

系数,且
$$\sum_{i=1}^{e=3} \omega_e = 1_{\circ}$$

2.3.1 感知风险因子

无人驾驶车辆的感知即是对道路上的交通态 势进行识别,如行人、车辆、路障,乃至车道及交通 标识,并针对物体特征依据内在模型对物体分类, 而分类问题可使用机器学习评估方法中的查准率 评价。

查准率 P 是真正例样例数与真正例及假正例 样例数和之比。此处,将查准率定义为无人驾驶车 辆感知系统的神经网络模型分类正确的比例,即

$$P = \frac{T_P}{T_P + F_P} \tag{15}$$

式中: *T_p* 为真正例, 即无人驾驶车辆感知系统将物体正确识别的样例数; *F_p* 为假正例, 即无人车无人驾驶车辆感知系统将物体错误识别成其他物体的样例数。

由于查准率越高,无人驾驶车辆感知系统的神 经网络模型分类正确的比例越高,行车风险越低。 为使得感知风险因子与行车风险的变化趋势一致, 采用1与查准率的差表示感知风险因子,即

$$D_1 = 1 - P = 1 - \frac{T_P}{T_P + F_P} \tag{16}$$

2.3.2 决策风险因子

无人驾驶车辆对感知系统得到的交通环境信息进行决策,并评估当前轨迹的可行性,从而对行 车轨迹进行规划。对于规划的多条可行轨迹,无人 驾驶车辆决策系统会根据安全评估算法计算每一 条轨迹的碰撞概率,最终选择碰撞概率最小的轨迹 作为其行驶轨迹。

无人驾驶车辆决策系统所规划轨迹的可行性 体现了决策的准确性,而决策的准确性决定了无人 驾驶车辆在面对复杂交通态势的行车安全。要尽 可能地保证行车安全,需要在轨迹执行前对轨迹的 安全性进行评估,即无人驾驶车辆沿轨迹行驶可能 发生碰撞的概率。因此,本文采用模拟碰撞率评估 无人驾驶车辆决策系统对行车安全的影响,并定义 决策风险因子为

$$D_2 = \frac{\varsigma_0}{\varsigma^*} \tag{17}$$

式中: s*为无人驾驶车辆轨迹仿真次数; so为无人 驾驶车辆仿真中会发生碰撞的次数,碰撞次数越 多,表示决策系统能力越差,行车风险越高。

2.3.3 控制风险因子

无人驾驶车辆在完成当前轨迹安全性评估后, 安全的轨迹将输出到车辆控制系统。然而,无人驾 驶车辆的实际行驶路线与规划行驶路线并非是完 全重合的,会存在一定的误差,如图4所示,其中, *P*_{e,q}和*P*_{t,q}分别为第*q*个规划和实际轨迹的点,*d*_q为 2个点的距离差。



Fig. 4 Actual trajectory and planned trajectory of unmanned vehicle

在实际行驶轨迹及规划行驶轨迹上根据(η-1) 个固定时间差Δt,将轨迹分割成η段,可采用均方差 表示每一段实际轨迹与规划轨迹的相对位姿差。

基于此,本文采用响应精确率评估无人驾驶车 辆控制系统对行车安全的影响,并将其定义为控制 风险因子,即

$$D_{3} = \left(\frac{1}{\eta} \sum_{q=1}^{\eta} \left\| D_{\text{trans}} \left(d_{q} \right) \right\|^{2} \right)^{\frac{1}{2}}$$
(18)

式中: η为轨迹取样的粒度; D_{trans} (d_q)表示每一段实际轨迹与规划轨迹的相对位姿差。相对位姿差的 均方差越大, 实际行车轨迹与规划行车轨迹的偏差 越大, 行车风险越大。

综上, AI 驾驶风险因子 D, 的表达式为

$$D_{i} = \omega_{1} \left(1 - \frac{T_{P}}{T_{P} + F_{P}} \right) + \omega_{2} \frac{\varsigma_{0}}{\varsigma^{*}} + \omega_{3} \cdot \left(\frac{1}{\eta} \sum_{q=1}^{\eta} \left\| D_{\text{trans}} \left(d_{q} \right) \right\|^{2} \right)^{\frac{1}{2}}$$
(19)

2.4 行车风险因素待定系数求解

对等效质量*Ω*_i函数表达式式(4)中存在的待定 系数λ₀、λ₁、λ₂进行标定及求解。等效质量待定系 数的标定流程如图 5 所示。



Fig. 5 Flow chart for calibration of equivalent mass undetermined coefficient

由于等效质量 Q_i是道路上物体由于自身属性 产生的潜在风险损失,因此,以实际交通事故造成 的损失表示此潜在风险的严重程度。其中,交通事 故的损失通常是指事故中人员受伤及死亡的数 量。以一亿辆车行驶一公里为基准,根据真实交通 事故数据,可求解亿车公里的事故数及伤亡人数, 并定义事故平均伤亡人数 N 为风险损失指标,表达 式为

$$N = \gamma_{\rm d} \frac{N_{\rm d}}{N_{\rm s}} + \gamma_{\rm i} \frac{N_{\rm i}}{N_{\rm s}} \tag{20}$$

式中: N_d 为亿车公里事故中的死亡人数; N_i 为亿车 公里事故中的受伤人数; N_s 为亿车公里事故的发生 数量; γ_d 为事故致死的严重度权重, γ_i 为事故致伤的 严重度权重, 取值范围满足 $0 \le \gamma_i < \gamma_d \le 1, 0 \le \gamma_i < \gamma_d \le 1, 1, 0 \le \gamma_i < \gamma_d \le 1, 1, 1, 0 \le \gamma_i < \gamma_d \le 1, 1, 1, 0 \le \gamma_i < \gamma_d \le 1, 1, 1, 0 \le \gamma_i < \gamma_d \le 1, 1, 1, 0 \le \gamma_i < \gamma_d \le 1, 1, 1, 0 \le \gamma_i < \gamma_d \le 1, 1, 1, 0 \le \gamma_i < \gamma_d \le 1, 1, 1, 0 \le \gamma_i < \gamma_d \le 1, 1, 1, 0 \le \gamma_i < \gamma_d \le 1, 1, 1, 0 \le \gamma_i < \gamma_d \le 1, 1, 1, 0 \le \gamma_i < \gamma_d \le 1, 0 \le 1$

对于式 (4) 中等效质量 Ω_i 函数表达式中的 $\lambda_0 v_i^{\Lambda_1} + \lambda_2$,其是用来描述速度对行车风险影响的幂 函数式。以事故损失表示行车风险,建立事故平均 伤亡人数 N 与速度的关系,如式 (21) 所示:

$$N = \lambda_0 v_i^{\lambda_1} + \lambda_2 \tag{21}$$

采集真实交通场景的通行速度,并根据式 (20) 计算事故平均伤亡人数 N,最后对 N 及车辆速度 v 拟合回归,则可求得待定系数 λ₀、λ₁、λ₂的值。

3 典型场景案例分析

3.1 数据处理及参数求解

根据文献 [29] 及国家统计局的统计数据,可得 到中国 7 条高速公路一年内的平均车速、里程、交 通量、汽车交通事故发生数、事故死亡人数及事故 受伤人数。选取事故致死严重度权重γ_d = 0.7,事故 致伤严重度权重γ_i = 0.3,计算得到亿车公里死亡人 数、亿车公里受伤人数及亿车公里事故数。在此基 础上由式 (20)可得事故平均伤亡人数 N,结果如 表1所示。

Table 1 Calculation results of expressway data						
高速公路	平均车速/ (km·h ⁻¹)	亿车公里死亡人数/ (10 ⁸ 人·辆·km ⁻¹)	亿车公里受伤人数/ (10 ⁸ 人·辆·km ⁻¹)	亿车公里事故数/ (10 ⁸ 次·辆·km ⁻¹)	Ν	
成渝高速	87.61	7	23	23	0.513	
石太高速	71.00	9	29	29	0.517	
广佛高速	58.13	6	21	21	0.500	
京石高速	93.00	15	47	45	0.547	
沪宁高速	79.86	6	20	21	0.486	
沈大高速	79.50	5	18	19	0.468	
京津塘高速	88.70	10	32	31	0.535	

表1 高速公路数据计算结果 le 1 Calculation results of expressway data

对事故平均伤亡人数 N 及各公路的平均车速 v 依据 $\lambda_{0}v_{i}^{\lambda_{1}} + \lambda_{2}$ 的关系式进行幂拟合,得到如图 6 所示曲线。该曲线表达了车速与事故风险严重程 度的关系。





Fig. 6 Relationship between average highway speed and average number of casualties in accidents

根据拟合结果, λ₀、λ₁、λ₂的值分别为7.886×10⁻¹⁷、 7.534、0.4856, 代入式 (21) 中则可得到 N关于速度 的表达式为

$$N = 7.886 \times 10^{-17} v_i^{7.534} + 0.4856$$
 (22)

将式 (22) 代入等效质量 Ω_i的表达式,则可得到 等效质量 Ω_i的最终表达式为

 $\Omega_i = T_i m_i \left(7.886 \times 10^{-17} v_i^{7.534} + 0.4856 \right)$ (23)

3.2 无人驾驶车辆行车安全场分析

设定基础场景为单行双车道,前后车之间的安 全距离为5m,选取单条道路宽5m,长度50m,坡 度为0,能见度最佳且路面附着系数为0.8水泥平 直路段。设本车无人驾驶车辆起始位置为(0,0),道 路上存在无人驾驶车辆、静止车辆、非机动车、行 人等共8个目标,且各个目标的当前状态信息如

表2所示,其中,目标①~目标⑧如图7所示。

オ	₹2	典ゴ	凹功京	当刖	状心	信思	
Table 2	Τν	nical	scena	rio s	tate	infor	mati

目标	$(x_i, y_i)/m$	m _i / kg	$\frac{v_i}{(\mathbf{m}\cdot\mathbf{s}^{-1})}$	$a_i/(\mathbf{m}\cdot\mathbf{s}^{-2})$	$\theta_i/$
目标①	(7,1)	800	45	2	0
目标②	(30,4)	200	0	0	0
目标③	(25, 0)	1 000	50	1.5	-25
目标④	(40,3)	900	30	-1	2
目标⑤	(47,2.5)	90	6	1	0
目标⑥	(4,-3)	88	15	1	0
目标⑦	(18,-4)	900	0	0	0
目标⑧	(37,-2.5)	60	15	1	-10

基于该场景,运用本文模型可得到该场景下的 行车安全场场强图如图 7 所示。其中, x 轴表示道 路的宽度, y 轴表示道路的长度, z 轴表示行车安全 场场强的大小, 数值越大, 表示行车风险越大。

由图 7 可知,此时由自行车⑥及无人驾驶车辆① 叠加形成的行车安全场峰值较高,即无人驾驶车辆 要想从两车之间穿行存在较大风险,应等交通态势 缓和后行进,与假设的实际场景情况相符。在施工 区②附近,由行车安全场场强图可知该区域行车风 险较低,即无人驾驶车辆可在场强较低的该区域安 全通行。

为对比不同场景之间行车安全场场强的差异 以进一步说明本文模型的可行性及正确性,基于该 基础场景,对其做道路目标及道路条件的变换。

其中,设定场景1的道路上不存在静止目标施 工区②,场景2的道路上不存在运动目标机动车 ⑥及机动车⑧,设定场景3的道路条件为下雨过后



Fig. 7 Typical driving safety field case of unmanned vehicles

的路面,其附着系数为 0.5。在此基础上生成不同 的场景及对应的行车安全场场强变化图,如图 8 所示。根据场景 1 及场景 2 生成的无人驾驶车辆 行车安全场场强图可知,不同种类的静态及动态目 标对整个交通环境产生的风险大小存在影响,在基 础场景的道路上减少该目标的同时,对应位置的无 人驾驶车辆行车安全场场强也发生了变化。无人 驾驶车辆行车安全场可反映道路上各个物体所产 生的潜在行车风险。由场景3的无人驾驶车辆行 车安全场场强图可知,此时道路的总体场强值远远 大于基础场景的场强值,说明无人驾驶车辆在雨后 路面行驶的行车风险要高于干燥水泥路面上行驶 的行车风险。无人驾驶车辆行车安全场可反映不 同道路条件对行车风险的影响。



Fig. 8 Different driving case and field strength changes

4 结 论

 1)本文模型综合考虑无人驾驶车辆 AI 操作系 统感知、决策、控制 3 大模块的特性,可解决原有 行车安全场模型应用到无人驾驶车辆上的局限性 问题。 2)结合生成的无人驾驶车辆行车安全场场景可知,本文模型可直观观察到道路上不同种类的物体对整个交通环境产生的风险大小,从而为无人驾驶车辆行车安全风险评估提供依据,为无人驾驶车辆后续路径规划或避撞算法的应用奠定基础。

参考文献(References)

[1] 《中国公路学报》编辑部.中国汽车工程学术研究综述·2017[J]. 中国公路学报, 2017, 30(6): 1-197.

Editorial Department of China Journal of Highway and Transport. Review on China's automotive engineering research progress: 2017[J]. China Journal of Highway and Transport, 2017, 30(6): 1-197 (in Chinese).

- [2] XU Q, LI K Q, WANG J Q, et al. The status, challenges, and trends: An interpretation of technology roadmap of intelligent and connected vehicles in China[J]. Journal of Intelligent and Connected Vehicles, 2022, 5(1): 1-7.
- [3] YUAN Q, XU X C, WANG T, et al. Investigating safety and liability of autonomous vehicles: Bayesian random parameter ordered probit model analysis[J]. Journal of Intelligent and Connected Vehicles, 2022, 5(3): 199-205.
- [4] VAN D H. A time-based analysis of road user behaviour in normal and critical encounters[M/OL]. (1990-04-24) [2021-11-01]. https:// repository.tudelft.nl/islandora/object/uuid%3A8fb40be7-fae1-4481bc37-12a7411b85c7.
- [5] KIEFER R J, FLANNAGAN C A, JEROME C J. Time-to-collision judgments under realistic driving conditions[J]. Human Factors: The Journal of the Human Factors and Ergonomics Society, 2006, 48(2): 334-345.
- [6] FULLER R G C. Determinants of time headway adopted by truck drivers[J]. Ergonomics, 1981, 24(6): 463-474.
- [7] YUAN Q A, CHEN H Y. Factor comparison of passenger-vehicle to vulnerable road user crashes in Beijing, China[J]. International Journal of Crashworthiness, 2017, 22(3): 260-270.
- [8] WANG J Q, WU J, ZHENG X J, et al. Driving safety field theory modeling and its application in pre-collision warning system[J]. Transportation Research Part C:Emerging Technologies, 2016, 72: 306-324.
- [9] ROSSETTER E J, GERDES J C. Lyapunov based performance guarantees for the potential field lane-keeping assistance system[J]. Journal of Dynamic Systems, Measurement, and Control, 2006, 128(3): 510-522.
- [10] BYRNE S, NAEEM W, FERGUSON S. Improved APF strategies for dual-arm local motion planning[J]. Transactions of the Institute of Measurement and Control, 2015, 37(1): 73-90.
- [11] MILLER R, HUANG Q F. An adaptive peer-to-peer collision warning system[C]//Proceedings of the IEEE 55th Vehicular Technology Conference. Piscataway: IEEE Press, 2002: 317-321.
- [12] TSOURVELOUDIS N C, VALAVANIS K P, HEBERT T. Autonomous vehicle navigation utilizing electrostatic potential fields and fuzzy logic[J]. IEEE Transactions on Robotics and Automation, 2001, 17(4): 490-497.
- [13] SATTEL T, BRANDT T. From robotics to automotive: Lane-keeping and collision avoidance based on elastic bands[J]. Vehicle System Dynamics, 2008, 46(7): 597-619.
- [14] 王建强, 吴剑, 李洋. 基于人-车-路协同的行车风险场概念、原理 及建模[J]. 中国公路学报, 2016, 29(1): 105-114.

WANG J Q, WU J, LI Y. Concept, principle and modeling of driving risk field based on driver-vehicle-road interaction[J]. China Journal of Highway and Transport, 2016, 29(1): 105-114(in Chinese).

 [15] 贺启才,蔡少康,王康.基于场论的汽车避撞方法研究[J]. 汽车实 用技术, 2021, 46(9): 25-29.
 HE Q C, CAI S K, WANG K. Research on collision avoidance

method based on field theory[J]. Automobile Applied Technology, 2021, 46(9): 25-29(in Chinese).

- [16] 刘帅. 基于安全场的智能汽车个性化换道决策与规划算法研究
 [D]. 长春: 吉林大学, 2019: 69-88.
 LIU S. A personalized lane-changing decision-making and trajectory-planning method based on safety field for intelligent vehicles[D]. Changchun: Jilin University, 2019: 69-88 (in Chinese).
- [17] NI D H. A unified perspective on traffic flow theory. Part I: The field theory[J]. Applied Mathematical Sciences, 2013, 7: 1929-1946.
- [18] 杨杨, 任少杰, 杨正才. 基于改进型人工势场的无人车局部避障 [J]. 湖北汽车工业学院学报, 2020, 34(4): 5-10. YANG Y, REN S J, YANG Z C. Local obstacle avoidance based on improved artificial potential field[J]. Journal of Hubei University of Automotive Technology, 2020, 34(4): 5-10(in Chinese).
- [19] 张国辉, 王璇, 张雅楠, 等. 实际环境中多无人车协同路径规划模型研究[J]. 系统仿真学报, 2023, 35(2): 408-422. ZHANG G H, WANG X, ZHANG Y N, et al. Research on cooperative path planning model of multiple unmanned vehicles in real environment[J]. Journal of System Simulation, 2023, 35(2): 408-422 (in Chinese).
- [20] BERRY A J, HOWITT J, GU D W, et al. A continuous local motion planning framework for unmanned vehicles in complex environments[J]. Journal of Intelligent & Robotic Systems, 2012, 66: 477-494.
- [21] TRENTINI M, BECKMAN B, DIGNEY B. Control and learning for intelligent mobility of unmanned ground vehicles in complex terrains[C]//Proceedings of the Unmanned Ground Vehicle Technology VII. Orlando: SPIE, 2005, 5804: 203-216.
- [22] LI J X, DAI B, LI X H, et al. An Interaction-aware predictive motion planner for unmanned ground vehicles in dynamic street scenarios[J]. International Journal Of Robotics & Automation, 2019, 34(3): 203-215.
- [23] ZHANG C, BERGER C, DOZZA M. Social-IWSTCNN: A social interaction-weighted spatio- temporal convolutional neural network for pedestrian trajectory prediction in urban traffic scenarios[C]// Proceedings of the 2021 IEEE Intelligent Vehicles Symposium. Piscataway: IEEE Press, 2021: 1515-1522.
- [24] YU S Y, MALAWADE A V, MUTHIRAYAN D, et al. Scenegraph augmented data-driven risk assessment of autonomous vehicle decisions[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(7): 7941-7951.
- [25] 金立生, 郭柏苍, 谢宪毅, 等. 基于行车安全场模型的交叉口车辆 控制算法[J]. 西南交通大学学报, 2022, 57(4): 753-760. JIN L S, GUO B C, XIE X Y, et al. Cooperative control algorithm for vehicle at intersection based on driving safety field model[J]. Journal of Southwest Jiaotong University, 2022, 57(4): 753-760 (in Chinese).
- [26] WANG J Q, WU J, LI Y, et al. The concept and modeling of driving safety field based on driver-vehicle-road interactions[C]//Proceedings of the 17th International IEEE Conference on Intelligent

Transportation Systems. Piscataway: IEEE Press, 2014: 974-981.

[27] 吴剑. 考虑人-车-路因素的行车风险评价方法研究[D]. 北京:清 华大学, 2015: 17-18.

WU J. Research on driver-vehicle-road factors considered driving risk evaluation method[D]. Beijing: Tsinghua University, 2015: 17-18 (in Chinese).

[28] 国家质量监督检验检疫总局,中国国家标准化管理委员会.水平 能见度等级: GB/T 33673—2017[S].北京:中国标准出版社, 2017:1-8.

General Administration of Quality Supervision, Inspection and

Quarantine of the People's Republic of China, Standardization Administration of the People's Republic of China. Grade of horizontal visibility: GB/T 33673—2017[S]. Beijing: Standards Press of China, 2017: 1-8 (in Chinese).

[29] 裴玉龙, 程国柱. 高速公路车速离散性与交通事故的关系及车速 管理研究[J]. 中国公路学报, 2004, 17(1): 74-78. PEI Y L, CHENG G Z. Research on the relationship between discrete character of speed and traffic accident and speed management of freeway[J]. China Journal of Highway and Transport, 2004, 17(1): 74-78 (in Chinese).

Construction method of driving safety field model for unmanned vehicles

XIE Chu'an, REN Yi, YANG Dezhen*, FENG Qiang, SUN Bo, WANG Zili

(School of Reliability and Systems Engineering, Beihang University, Beijing 100191, China)

Abstract: In response to the shortage of comprehensive and systematic methods for evaluating the driving safety of unmanned vehicles, an enhanced driving safety field model is proposed, taking into account the impact of the complex road factors of unmanned vehicles as well as the characteristics of three artificial intelligence (AI) system modules: perception, decision-making, and control. A mathematical model of unmanned vehicle driving safety field combining dynamic potential field and safety behavior field is established based on Hooke's law to characterize the driving risks caused by static objects, moving objects and AI system itself on the road. The correctness and usability of the proposed model is verified by the driving safety analysis of typical driving scenarios.

Keywords: unmanned vehicles; driving safety field; risk; driving safety evaluation; dynamic potential field

Received: 2022-06-08; Accepted: 2022-09-11; Published Online: 2022-11-03 11:17 URL: link.cnki.net/urlid/11.2625.V.20221103.0856.001

Foundation item: National Key Laboratory of Reliability and Environmental Engineering Technology Foundation (6142004210108)

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

DOI: 10.13700/j.bh.1001-5965.2022.0474

碳化材料的烧蚀性能参数辨识方法

程莫瀚,李文光*,王志

(北京理工大学 宇航学院,北京 100081)

摘 要: 针对碳化材料烧蚀性能参数测量难度大的问题,对基于优化算法的烧蚀性能参数 辨识方法进行了研究。根据一维连续模型描述碳化烧蚀材料的内部热响应情况,与公开试验数据进 行对比验证,各测量点计算误差均小于 15%。采用基函数表征法对材料碳化前后的比热容和导热率 进行表征,基函数选取为切比雪夫多项式。根据 PICA 材料的温度测试试验数据,对材料的热解动 力学参数和基函数的待定系数采用遗传算法进行辨识,将辨识得到的材料烧蚀性能参数代入一维连 续模型中进行计算,与验证工况下的试验数据进行比对,计算得到的温度曲线平均相对误差小于 10%。

关键 词: 热防护材料; 烧蚀性能; 参数辨识; 优化算法; 数值计算

中图分类号: V250.3

文献标志码: A 文章编号: 1001-5965(2024)04-1384-08

随着各国航空航天工业的发展,飞行器高速飞 行带来的热防护问题不容忽视,具有良好隔热性能 的低密度碳化烧蚀材料被广泛应用于高速飞行器 的热防护系统。低密度碳化烧蚀材料的防热机理 非常复杂,包含传热、化学反应、气体流动和热物 性变化等过程,且不同防热材料的基体和强化材 料、热容吸热、化学反应不同,防热机理也会出现 差异。但经过多年研究,发现不同防热材料的烧蚀 过程存在相似之处,即随着烧蚀过程的进行,材料 会逐渐分为碳化层、热解层和原始层^[1]。

目前,有4种模型描述烧蚀材料的分层现象: 简化模型、热解面模型、热解层模型和连续模型^[2]。 简化模型忽略材料热解现象,将碳化层、热解层和 原始层简化为一层,根据该层的瞬时温度分布划分 为3个区域,并对每个区域求解能量方程,从而解 得区域内部的温度相应结果。这是最早的一类模 型,计算简单方便,但精度较差,且难以反映实际情 况。热解面模型将材料的热解区域假定为无厚度 的薄层,只考虑碳化层和原始层的热响应特性,热 解气体质量流率和温度响应由能量平衡关系式计 算,该模型相比简化模型更加精确。热解层模型将 材料划分为碳化层、热解层和原始层,分别考虑每 个特征区域的内部热响应特性,并根据傅里叶定律 建立导热微分方程,该模型更加接近真实情况,但 特征区域的划分存在一定困难。连续模型将3种 特征区域统一建模,无需人工分层,与热解层模型 相比更接近实际情况。

采用上述模型对材料热响应进行仿真时,需要 得知碳化烧蚀材料的烧蚀性能参数,包括反映热解 反应规律的热解动力学参数和材料本身的热物理 参数。在热解动力学参数辨识方面,目前多采用的 方法是通过缓慢升温实验得到材料剩余质量与温 度的关系,再采用 Coats-Redfern 积分法、Kissinger 法、Ozawa 法等方法得到所需的热解动力学参数^[3]。 黄娜等^[4] 采用热重质谱联用技术和热重分析法研 究了氨酚醛树脂的热解现象,并使用 Coats-Redfern 积分法对实验数据进行了动力学解析,得到了热解 动力学参数;张莹等^[5] 采用非等温法研究了钡酚醛 树脂的热解反应行为,并通过 Kissinger 法计算得到 了活化能;时天林等^[6] 采用热重分析法对酚醛树脂 的热解过程进行了分析,并通过 Ozawa 法和机理函 数法确定了热解动力学参数。但是有学者研究发

收稿日期: 2022-06-11; 录用日期: 2022-08-21; 网络出版时间: 2022-11-01 13:12

引用格式: 程莫瀚, 李文光, 王志. 碳化材料的烧蚀性能参数辨识方法 [J]. 北京航空航天大学学报, 2024, 50 (4): 1384-1391. CHENG M H, LI W G, WANG Z. Identification method of ablation performance parameters of carbonized materials [J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50 (4): 1384-1391 (in Chinese).

网络出版地址: link.cnki.net/urlid/11.2625.V.20221031.1437.002

^{*}通信作者. E-mail: liwenguang@bit.edu.cn

现,在实际的烧蚀过程中,材料升温速率极高,与实验方法建立的缓慢升温条件有较大差别,材料烧蚀性能参数的计算结果也会不同。文献[7-9]研究了酚醛树脂在温度为1000~2000 K之间的热解现象,发现材料在快速升温环境下的热解过程与缓慢升温环境下相比有显著不同。

针对材料本身的热物理参数,Telejko和 Malinowski^[10]采用Broyden-Fletcher-Goldfarb-Shanno 方法和有限元法辨识了材料导热率;Zueco等^[11]使 用网络模型反推了一维传热问题中均匀材料的导 热率;Char等^[12]提出了一种微分求积法来反推一维 传热问题中材料的导热率,该方法的特点是未假定 导热率的函数形式。近年来,智能优化算法得到发 展,被广泛应用于热物理参数辨识中。Ardakani和 Khodada^[13]在边界元法的基础上,结合粒子群算法 辨识了材料的导热率;曾建潮和崔志华^[14]基于粒子群 算法提出了 SPSO算法,能够保证搜索结果收敛于 全局最优解;Liu^[15]将遗传算法和Levenberg-Marquardt 方法相结合,对瞬态传热问题中材料的导热率和比 热容进行辨识,证明该方法具有可行性。

综上所述,针对材料热解动力学参数的辨识问题,现有方法大多建立在材料缓慢升温的测试环境 基础上,然而在实际的快速升温环境中,上述方法 得出的热解动力学参数不能满足实际情况,无法反 映升温速率较高的加热条件下烧蚀材料真实的防 热性能。针对材料本身的热物理参数辨识问题,现 有方法所依赖的模型大多为均匀材料的非稳态传 热模型,忽略高温条件下材料的热解效应,对烧蚀 材料的性能参数辨识具有局限性。因此,本文基于 一维连续模型,采用切比雪夫多项式构建导热率和 比热容的函数形式,依据碳化烧蚀材料快速升温试 验曲线,采用遗传算法辨识低密度碳化烧蚀材料的 烧蚀性能参数。

1 一维连续模型

1.1 热响应现象分析

碳化烧蚀材料的热响应现象可以用图 1 所示的分层模型来描述^[16]。

低密度碳化烧蚀材料在温度上升至一定温度 后,就会开始发生热解反应。热解反应正在发生的 区域称为热解层,尚未开始发生热解反应的区域称 为原始层,完全热解形成碳层的区域称为碳化层。 热解反应吸收一部分能量,产生的热解气体经由材 料内部流动,将流动过程中吸收的能量释放到边界 层中。在热解过程中,材料的组成和温度发生改 变,其烧蚀性能参数也随之改变。





1.2 控制方程

由于热解区域在仿真中难以界定,本文采用一 维连续模型^[17],以材料剩余密度为纽带,将热解过 程3种主要特征区域统一建模,从而有效改善热解 层模型人为分层引入的误差。

控制方程为

$$\rho C_{\rm p} \frac{\partial T}{\partial t} = \frac{\partial}{\partial x} \left(\lambda \frac{\partial T}{\partial x} \right) + \Delta H_{\rm c} \frac{\partial \rho}{\partial t} + \dot{m}_{\rm g} C_{\rm pg} \frac{\partial T}{\partial x} \qquad (1)$$

式中: ρ 为材料密度; C_p 为材料比热容; λ 为材料导 热率;T为温度; ΔH_c 为材料热解热; \dot{m}_g 为热解气体 质量流率; C_{pg} 为热解气体比热容。

$$\Delta H_{\rm c} = h_{\rm g} - \bar{h} \tag{2}$$

$$\bar{h} \triangleq \frac{\rho_{\rm v} h_{\rm v} - \rho_{\rm c} h_{\rm c}}{\rho_{\rm v} - \rho_{\rm c}} \tag{3}$$

式中: h_g 为热解气体焓; ρ_v 为原始材料密度; h_v 为原 始材料焓; ρ_c 为碳化材料密度; h_c 为碳化材料焓。

$$h_{\rm g} = \int_{T_{\rm ref}}^{T} C_{\rm pg} \mathrm{d}T \tag{4}$$

$$h_{\rm v} = \int_{T_{\rm ref}}^{T} C_{\rm pv} \mathrm{d}T \tag{5}$$

$$h_{\rm c} = \int_{T_{\rm ref}}^{T} C_{\rm pc} \mathrm{d}T \tag{6}$$

式中: *T*_{ref}为参考温度,取 300 K; *C*_{pv}为原始材料比 热容; *C*_{pc}为碳化材料比热容。

在材料受热烧蚀分解过程中,材料密度不断减小,其热解动力学规律可采用 Arrhenius 方程表征:

$$\frac{\partial \rho}{\partial t} = -A \exp\left(\frac{-E}{RT}\right) \rho_{\rm v} \left(\frac{\rho - \rho_{\rm c}}{\rho_{\rm v}}\right)^m \tag{7}$$

式中:A为指前因子;E为活化能;R为气体常数; m为反应级数。

为求解热解气体质量流率,引入质量守恒方程: $\frac{\partial \rho}{\partial t} = -\frac{\partial \dot{m}_s}{\partial x}$ (8)

烧蚀材料的导热率和比热容会随热解反应的 进行程度而变化,因此,引入碳化率φ来表征热解 反应的进行程度:

$$\varphi = \frac{\rho - \rho_{\rm v}}{\rho_{\rm v} - \rho_{\rm c}} \tag{9}$$

 $C_{\rm p} = \varphi C_{\rm pv} + (1 - \varphi) C_{\rm pc} \tag{10}$

 $\lambda = \varphi \lambda_{\rm v} + (1 - \varphi) \lambda_{\rm c} \tag{11}$

式中: λ_v为原始材料导热率; λ_c为碳化材料导热率。 **1.3 边界条件**

考虑气体产物质量引射的边界条件,可以建立 边界能量平衡关系^[18],如图2所示。





根据上述边界层能量平衡关系,建立边界层能 量平衡关系式:

$$q_{\rm in} = q_0 - \dot{m}_{\rm g} h_{\rm w} + \dot{m}_{\rm g} h_{\rm g} - \varepsilon \sigma \left(T_{\rm w}^4 - T_{\infty}^4 \right) \tag{12}$$

式中: q_{in} 为表面净吸收热流密度; q_0 为冷壁热流密度; h_w 为壁面焓; ε 为辐射系数; σ 为玻尔兹曼常数; T_w 为壁面温度; T_∞ 为环境温度。

式 (12) 中, 各项从左到右的含义为: 流入材料 内部的净吸收热流、冷壁热流、热解气体注入边界 层带走的热流、热解气体流动带来的热流及辐射热流。

1.4 数值求解

将上述控制方程(1)采用有限差分进行离散, 数值模拟过程如图3所示。首先,设定网格密度和 时间步长;然后,输入材料参数和工况条件,在每个 时间步长中使用有限差分法求解热响应控制方程 和边界控制方程,并将得到的温度和材料密度分布 用于下一时刻的热响应计算,直至工况计算时间结束。

对控制方程的空间导数项使用二阶中心差分, 时间导数项使用一阶向前差分,得到显式格式,具体为

$$\left(\rho C_{\rm p}\right)_{n}^{t} \frac{T_{n}^{t+1} - T_{n}^{t}}{\Delta t} = \frac{\lambda_{n+1/2}^{t} \frac{T_{n+1}^{t} - T_{n}^{t}}{\Delta x} - \lambda_{n-1/2}^{t} \frac{T_{n}^{t} - T_{n-1}^{t}}{\Delta x}}{\Delta x} + \left(\Delta H_{\rm c} \frac{\partial \rho}{\partial t}\right)_{-}^{t} + \left(\dot{m}_{\rm g} C_{\rm pg}\right)_{n}^{t} \frac{T_{n+1}^{t} - T_{n-1}^{t}}{2\Delta x} \quad (13)$$
$$\Leftrightarrow \frac{\partial \rho}{\partial t} \stackrel{\text{\tiny def}}{=} \omega_{z}$$



图 3 热响应模拟计算流程

Fig. 3 Flow chart of thermal response simulation calculation

$$\begin{cases}
A_n^t = \lambda_{n-1/2}^t \frac{\Delta t}{\left(\rho C_p\right)_n^t \left(\Delta x\right)^2} - \frac{\Delta t}{\left(\rho C_p\right)_n^t} \cdot \frac{\dot{m}_g C_{pg}}{2\Delta x} \\
D_n^t = \lambda_{n+1/2}^t \frac{\Delta t}{\left(\rho C_p\right)_n^t \left(\Delta x\right)^2} + \frac{\Delta t}{\left(\rho C_p\right)_n^t} \cdot \frac{\dot{m}_g C_{pg}}{2\Delta x} \\
B_n^t = \frac{1}{2} \left(A_n^t + D_n^t\right) \\
P_n^t = \frac{\Delta t}{\left(\rho C_p\right)_n^t} \left(\left(h_g - \bar{h}\right) \omega_z\right)_n^t
\end{cases}$$
(14)

得

$$T_n^{t+1} = A_n^t T_{n-1}^t + (1 - 2B_n^t) T_n^t + D_n^t T_{n+1}^t + P_n^t$$
(15)

式中: $\lambda'_{n+1/2}$ 、 $\lambda'_{n-1/2}$ 分别由相邻位置的 λ 值调和平均 求解:

$$\begin{cases} \frac{2}{\lambda_{n+1/2}^{t}} = \frac{1}{\lambda_{n}^{t}} + \frac{1}{\lambda_{n+1}^{t}} \\ \frac{2}{\lambda_{n-1/2}^{t}} = \frac{1}{\lambda_{n}^{t}} + \frac{1}{\lambda_{n-1}^{t}} \end{cases}$$
(16)

设防护层有N个节点,即n=1,2,…,N-1,N(其 中,1代表烧蚀表面,N代表内壁),由上述离散格 式可知,在求解n=1,n=N节点时,会出现n=0, n=N+1的情况。因此,构建虚拟节点如图4所示。

在烧蚀表面处给定表面净吸收热流密度q_{in}和 背壁净吸收热流密度q_{wall},由傅里叶公式:

$$-\lambda \frac{\partial T}{\partial t} = q_{\rm in} \tag{17}$$

$$\begin{cases} -\lambda_1 \frac{T_2 - T_0}{2\Delta x} = q_{\rm in} \\ T_0 = T_2 + \frac{2\Delta x q_{\rm in}}{\lambda_1} \end{cases}$$
(18)

$$\begin{cases} -\lambda_N \frac{T_{N+1} - T_{N-1}}{2\Delta x} = q_{\text{wall}} \\ T_{N+1} = T_{N-1} - \frac{2\Delta x q_{\text{wall}}}{\lambda_N} \end{cases}$$
(19)







故最终计算公式为

$$\begin{cases} T_{1}^{t+1} = A_{1}^{t} \frac{2\Delta x q_{\text{in}}}{\lambda_{1}^{t}} + (1 - 2B_{1}^{t}) T_{1}^{t} + (A_{1}^{t} + D_{1}^{t}) T_{2}^{t} + P_{1}^{t} \\ n = 1 \\ T_{n}^{t+1} = A_{n}^{t} T_{n-1}^{t} + (1 - 2B_{n}^{t}) T_{n}^{t} + D_{n}^{t} T_{n+1}^{t} + P_{n}^{t} \\ n = 2, \cdots, N-1 \\ T_{N}^{t+1} = -D_{N}^{t} \frac{2\Delta x q_{\text{wall}}}{\lambda_{N}^{t}} + (1 - 2B_{N}^{t}) T_{N}^{t} + (A_{1}^{t} + D_{1}^{t}) T_{N-1}^{t} + P_{1}^{t} \\ n = N \\ (20) \\ \frac{\partial \rho}{\partial t} = \frac{\rho_{n}^{t+1} - \rho_{n}^{t}}{\Delta t} \end{cases}$$

$$\rho_n^{t+1} = \omega_z \Delta t + \rho_n^t \tag{22}$$

 $P_n \qquad \Im_2 = P_n \qquad (=$

1.5 模型验证

为验证上述模型和求解方法的准确性,编写数 值求解程序,将计算结果与试验结果进行对比。文 献[19]给出了一种试验工况:持续时间为 800 s、热 流密度为 279.7 kW/m²的辐射热流作用于厚度为 30 mm 的碳纤维增强酚醛复合材料试件上。文献 [19]给出了试件不同位置处(距壁面 1 mm、5 mm、 10 mm 和 29 mm)的温升情况。

试件材料性能参数及初始条件列于表1中。

表1 材料性能及初始条件

材料性能	数值
$ ho_{ m v}/(m kg\cdot m^{-3})$	1 810
$ ho_{ m c}/(m kg{\cdot}m^{-3})$	1 440
$\lambda_v/(W\cdot m^{-1}\cdot^\circ\! C^{-1})$	$0.804 + 2.76 \times 10^{-4}T$
$\lambda_c/(W \cdot m^{-1} \cdot {}^\circ C^{-1})$	$0.955 + 8.42 \times 10^{-4}T - 4.07 \times 10^{-6}T^2 + 5.32 \times 10^{-9}T^3$
$C_v/(kJ\cdot kg^{-1}\cdot C^{-1})$	$1.089 + 1.09 \times 10^{-3} T$
$C_{\rm c}/({\rm kJ}\cdot{\rm kg}^{-1}^{\circ}{\rm C}^{-1})$	$0.87 + 1.02 \times 10^{-3} T$
$C_{\rm pg}/(\rm kJ\cdot \rm kg^{-1}\cdot ^{\circ}C^{-1})$	9.63
ε	0.9
$\sigma/\!(\mathbf{W}\!\cdot\!\mathbf{m}^{-2}\!\cdot\!\mathbf{K}^{-4})$	5.73×10^{-8}
T_0/K	300
$E/(kJ \cdot kmol^{-1})$	2.60×10^{5}
A/s^{-1}	1.98×10 ²⁹ (剩余率≥0.91), 8.16×10 ¹⁸ (剩余率<0.91)
m	17.33(剩余率≥0.91), 6.3(剩余率≥0.91)

碳化烧蚀材料不同位置处温度-时间历程的仿 真数值与试验结果对比如图 5 所示。图中,连续曲 线为程序仿真结果,离散点为试验结果。由此可 知,温度场的数值计算结果与试验结果趋势一致。

在热流持续加热 800 s 后, 取最终时刻的仿真 值与试验值进行对比, 其误差如表 2 所示。可以看 出, 试件不同厚度的仿真结果与试验值的相对误差 控制在 15% 以内, 说明本文建立的一维连续模型计 算精度较好。



图 5 不同位置处材料热响应数值仿真与试验结果对比

Fig. 5 Comparison between numerical simulation and experimental results of thermal response of materials at different positions

表 2 最终时刻不同位置温度对比

Table 2 Temperature comparison of different positions at the final moment

位置/mm —	温度	相对 泪 关 网	
	试验值	仿真值	相利 庆左/%
1	1 264.8	1 295.6	2.44
5	1 095.4	1 188.4	8.49
10	938.54	1 020.6	8.74
29	521.37	594.39	14.00

2 基于试验数据的参数辨识方法

2.1 烧蚀分析的参数化表征

根据第1节中建立的一维连续模型,可将影响 烧蚀特性的参数分为2部分:烧蚀材料的热物理特 性参数和热解动力学相关参数。其中,碳化烧蚀材 料的热物理参数表示为温度T相关的函数,热解动 力学相关参数表示为常数。

材料热物理参数中,比热容 $C_{pv}(T)$ 、 $C_{pc}(T)$ 和导 热率 $\lambda_v(T)$ 、 $\lambda_c(T)$ 是随温度变化的参数,对于这类 参数,选定一组基函数:

$$F_{kv} = \{G_{d_{kv(1)}}(T), \cdots, G_{d_{kv}(P_{kv})}(T)\}$$

$$F_{kc} = \{G_{d_{kc(1)}}(T), \cdots, G_{d_{kc}(P_{kc})}(T)\}$$

$$F_{cv} = \{G_{d_{cv(1)}}(T), \cdots, G_{d_{cv}(P_{cv})}(T)\}$$

$$F_{cc} = \{G_{d_{cc(1)}}(T), \cdots, G_{d_{cc}(P_{cc})}(T)\}$$
(23)

式中: F_{kv} 表示烧蚀材料热解前导热率随温度变化 函数子空间; F_{kc} 表示烧蚀材料碳化后导热率随温 度变化函数子空间; F_{cv} 表示烧蚀材料热解前比热 容随温度变化函数子空间; F_{cc} 表示烧蚀材料磷化 后比热容随温度变化函数子空间; $d_{kv(i)}(i = 1, \dots, p_{kv}),$ $d_{kc(i)}(i = 1, \dots, p_{kc}), d_{cv(i)}(i = 1, \dots, p_{cv}), d_{cc(i)}(i = 1, \dots, p_{cc})$ 包含了每个子空间中的特定基函数的指标。

将热物理参数进行泛函展开,进行表征:

$$\begin{cases} \lambda_{\rm v}(T) = \alpha_{\rm kv0} + \sum_{j=1}^{p_{\rm kv}} \alpha_{\rm kvj} G_{d_{\rm kv(j)}}(T) \\ \lambda_{\rm c}(T) = \alpha_{\rm kc0} + \sum_{j=1}^{p_{\rm kc}} \alpha_{\rm kcj} G_{d_{\rm kc(j)}}(T) \\ C_{\rm pv}(T) = \alpha_{\rm cv0} + \sum_{j=1}^{p_{\rm cv}} \alpha_{\rm cvj} G_{d_{\rm cv(j)}}(T) \\ C_{\rm pc}(T) = \alpha_{\rm cc0} + \sum_{j=1}^{p_{\rm cv}} \alpha_{\rm ccj} G_{d_{\rm cc(j)}}(T) \end{cases}$$
(24)

式中: α 为常数。

即可转换为

$$\begin{cases} \boldsymbol{\theta}_{1} = [\alpha_{kv0}, \cdots, \alpha_{kvj}, \alpha_{kc0}, \cdots, \alpha_{kcj}, \alpha_{cv0}, \cdots, \alpha_{cvj}, \\ \alpha_{cc0}, \cdots, \alpha_{ccj}] \\ \boldsymbol{\theta}_{2} = [\rho_{c}, A_{mg}, E_{mg}, m_{mg}] \\ \boldsymbol{\theta} = [\boldsymbol{\theta}_{1}, \boldsymbol{\theta}_{2}]^{\mathrm{T}} \end{cases}$$
(25)

2.2 切比雪夫多项式拟合法

使用多项式族进行拟合可以实现较好的精度,同时可对求解精度进行控制,而在不同的多项式中,切比雪夫多项式的拟合效果最佳^[20]。n阶切比 雪夫多项式为

$$G_{d(j)}(T) = \cos(j \cdot \arccos(T))$$
(26)

由式 (24) 选定阶次后,采用切比雪夫多项式即 可对碳化烧蚀材料的热物理参数进行表征。

2.3 辨识模型

为能够基于试验数据获得 2.1 节中烧蚀分析的 主要参数,并使参数辨识结果尽可能与实际值相接 近,本文构造以下费用函数作为目标函数:

$$S(\boldsymbol{\theta}) = \sqrt{\frac{1}{M} \sum_{i=1}^{M} \left(T_i^* - T_i(\boldsymbol{\theta}) \right)^2}$$
(27)

式中: M为试验数据记录温度的时刻数; T_i^* 为温度的试验值; T_i 为温度的计算值($i = 1 \sim M$)。

可将参数辨识问题转化为优化问题,优化模 型如式(28)所示:

$$\begin{array}{l} \text{Find } \boldsymbol{\theta} \\ \text{min } S\left(\boldsymbol{\theta}\right) \\ \\ \text{s.t.} \begin{cases} C_{\text{pvmin}} \leqslant C_{\text{pv}} \leqslant C_{\text{pvmax}} \\ C_{\text{pcmin}} \leqslant C_{\text{pc}} \leqslant C_{\text{pcmax}} \\ \lambda_{\text{vmin}} \leqslant \lambda_{\text{v}} \leqslant \lambda_{\text{vmax}} \\ \lambda_{\text{cmin}} \leqslant \lambda_{\text{v}} \leqslant \lambda_{\text{vmax}} \\ \lambda_{\text{cmin}} \leqslant \lambda_{\text{c}} \leqslant \lambda_{\text{cmax}} \\ (28) \\ \\ \text{s.t.} \begin{cases} C_{\text{pgmin}} \leqslant C_{\text{pg}} \leqslant C_{\text{pgmax}} \\ \rho_{\text{cmin}} \leqslant C_{\text{pg}} \leqslant C_{\text{pgmax}} \\ \rho_{\text{cmin}} \leqslant A_{i} \leqslant A_{i\text{max}} \\ E_{i\text{min}} \leqslant E_{i} \leqslant E_{i\text{max}} \\ m_{gimin} \leqslant m_{gi} \leqslant m_{gimax} \end{array}$$

3 算 例

为验证烧蚀性能参数估算方法的准确性和适用性,本文对 PICA 材料的烧蚀性能参数进行估算。文献 [21] 提供了 PICA 材料的温度测试试验,本文选取的试验条件列于表 3。

表 3 试验条件及编号 Table 3 Experimental conditions and numbers

编号	热流密度/ $\left(W \cdot \left(m^2 \cdot s\right)^{-1}\right)$	持续 时间/s	材料 厚度/cm	测量 位置/cm
1	5.68×10 ⁶	11	15.24	1.08
2	5.68×10 ⁶	11	15.24	1.93
3	1.90×10 ⁷	17	20.32	3.77
4	9.66×10 ⁶	22	10.16	2.56

选取编号1、编号2和编号3的试验条件作为 训练集,编号4的试验条件作为验证集。基函数选 定为切比雪夫多项式,其中,

$$\begin{cases} p_{kv} = p_{kc} = p_{cv} = 1 \\ p_{cc} = 3 \end{cases}$$
(29)

采用遗传算法进行参数辨识,种群数量设置为 20,交叉概率设置为0.8。参数辨识结果为

$$\begin{cases} \boldsymbol{\theta}_{1} = [\alpha_{kv0}, \alpha_{kv1}, \alpha_{kc0}, \alpha_{kc1}, \alpha_{kc2}, \alpha_{kc3}, \alpha_{cv0}, \alpha_{cv1}, \alpha_{cc0}, \alpha_{cc1}] = \\ [0.50, 2.84 \times 10^{-4}, 0.76, 3.09 \times 10^{-4}, 2.36 \times 10^{-6}, \\ 1.40 \times 10^{-9}, 544.09, 1.02, 3\ 054.3, 0.076] \\ \boldsymbol{\theta}_{2} = [\rho_{c}, A_{mg}, E_{mg}, m_{mg}] = \end{cases}$$

$$[195.77, 3.65 \times 10^{19}, 4.25 \times 10^{6}, 10.79]$$

(30)

收敛曲线、训练集和验证集温度对比曲线如

图 6~图 9 所示。

参数辨识结果在第5代时收敛,模型调用次数为2025次,目标函数最优值为25.8391。

由图 7、图 8 和图 9 可知,最终参数辨识结果在 训练集中仿真温度曲线与试验温度曲线基本贴合, 趋势相同且误差较小。由图 10 可知,在试验 4 中 仿真曲线升温时刻较试验有一定延迟,但参数结果 在验证集中总体来说表现良好,平均相对误差为 5.49%。











图 9 试验 3 温度对比曲线

Fig. 9 Experiment 3 temperature comparison curve



Fig. 10 Experiment 4 temperature comparison curve

4 结 论

本文对低密度碳化烧蚀材料内部热响应数值 计算和烧蚀性能参数辨识方法进行了研究,编写了 相应的计算程序,并对温度计算模型和辨识方法进 行了考核验证。分析得到如下结论:

 1)本文建立的一维连续模型在计算低密度碳 化烧蚀材料内部温度场时具有较高的精度,对于同 种工况不同测量位置的温度仿真结果与试验相比 误差不超过15%。

2)本文使用的烧蚀性能参数辨识方法能够较 好地得到参数结果,在验证工况下计算平均相对误 差小于10%。

参考文献(References)

- [1] 朱召贤,朱小飞,黄洪勇,等. 低密度树脂基防热材料研究进展[J]. 中国材料进展, 2019, 38(11): 1086-1092.
 ZHU Z X, ZHU Z F, HUANG H Y, et al. Research progress of lowdensity resin-based thermal protection materials[J]. Materials China, 2019, 38(11): 1086-1092(in Chinese).
- [2] 张志成,潘梅林,刘初平,等.高超声速气动热和热防护[M].北京: 国防工业出版社,2003.

ZHANG Z C, PAN M L, LIU C P, et al. Hypersonic aerodynamic thermal and thermal protection[M]. Beijing: National Defense Industry Press, 2003(in Chinese).

- [3] 张军,李伟,方国东,等. 树脂基防隔热复合材料高温响应分析方 法研究进展[J]. 宇航学报, 2020, 41(6): 739-748.
 ZHANG J, LI W, FANG G D, et al. Review of high temperature response analysis of resin matrix thermal protection and insulation composites[J]. Journal of Astronautics, 2020, 41(6): 739-748(in Chinese).
- [4] 黄娜, 刘亮, 王晓叶. 热重质谱联用技术对酚醛树脂热解行为及 动力学研究[J]. 宇航材料工艺, 2012, 42(2): 99-102.
 HUANG N, LIU L, WANG X Y. Pyrolysis and kinetics of phenolic resin by TG-MS analysis[J]. Aerospace Materials & Technology, 2012, 42(2): 99-102(in Chinese).
- [5] 张莹, 胡宏林, 蒋丽琴, 等. 典型结构钡酚醛树脂热解非等温动力 学及热稳定性[J]. 宇航材料工艺, 2021, 51(1): 17-24.
 ZHANG Y, HU H L, JIANG L Q, et al. Non-isothermal pyrolysis kinetic mechanism and thermal stabilization of phenolic resin with different characteristic structure[J]. Aerospace Materials & Technology, 2021, 51(1): 17-24(in Chinese).
- [6] 时天林,高丽娟,何也,等. 酚醛树脂保温板热解动力学研究[J]. 辽宁科技大学学报, 2021, 44(1): 38-42.
 SHI T L, GAO L J, HE Y, et al. Study on pyrolysis kinetics of phenolic resin insulation board[J]. Journal of Liaoning University of Science and Technology, 2021, 44(1): 38-42(in Chinese).
- [7] 王苏, 闫卫锋, 马伟, 等. 酚醛树脂热解的激波管实验[C]//第一届高超声速科技学术会议. 北京: 中国力学学会, 2008: 21-26.
 WANG S, YAN W F, MA W, et al. Shock tube study on pyrolysis of phenolic resin[C]//Proceedings of the 1st Hypersonic Technology Academic Conference. Beijing: The Chinese Society of Theoretical and Applies Mechanics, 2008: 21-26(in Chinese).
- [8] 闫卫锋,王苏,李帅辉,等. 酚醛树脂热解产物的激波管研究
 [C]//第十三届全国激波与激波管会议.北京:中国力学学会, 2008: 452-457.

YAN W F, WANG S, LI S H, et al. Shock tube study on pyrolysis products of phenolic resin[C]//Proceedings of the 13th National Conference on Shock Waves and Shock Tubes. Beijing: The Chinese Society of Theoretical and Applies Mechanics, 2008: 452-457(in Chinese).

- [9] 闫卫锋, 王苏, 马伟, 等. 酚醛树脂高温热解的激波管实验研究[J]. 力学学报, 2009, 41(4): 463-468.
 YAN W F, WANG S, MA W, et al. Shock tube study of phenolic resin pyrolysis[J]. Chinese Journal of Theoretical and Applied Mechanics, 2009, 41(4): 463-468(in Chinese).
- [10] TELEJKO T, MALINOWSKI Z. Application of an inverse solution to the thermal conductivity identification using the finite element

method[J]. Journal of Materials Processing Technology, 2004, 146: 145-155.

- [11] ZUECO J, ALHAMA F, FERNANDEZ C. Inverse determination of temperature dependent thermal conductivity using network simulation method[J]. Journal of Materials Processing Technology, 2006, 174(1-3): 137-144.
- [12] CHAR M I, CHANG F P, TAI B C. Inverse determination of thermal conductivity by differential quadrature method[J]. International Communications in Heat and Mass Transfer, 2008, 35: 113-119.
- [13] ARDAKANI D M, KHODADAD M. Identification of thermal conductivity and the shape of an inclusion using the boundary elements method and the particle swarm optimization algorithm[J]. Inverse Problems in Science & Engineering, 2009, 17(7): 855-870.
- [14] 曾建潮, 崔志华. 一种保证全局收敛的 PSO 算法[J]. 计算机研究 与发展, 2004, 41(8): 1333-1338.
 ZENG J C, CUI Z H. A guaranteed global convergence particle swarm optimizer[J]. Journal of Computer Research and Development, 2004, 41(8): 1333-1338(in Chinese).
- [15] LIU F B. A hybrid method for the inverse heat transfer of estimating fluid thermal conductivity and heat capacity[J]. International Journal of Thermal Sciences, 2011, 50(5): 718-724.
- [16] 郭瑾. 树脂基复合材料热解层模型及高温热物性测试研究[D]. 北京: 北京交通大学, 2021.
 GUO J. Study on measurement of high-temperature thermophysical properties and pyrolysis layer model of polymer matrix composites
 [D]. Beijing: Beijing Jiaotong University, 2021(in Chinese).
- [17] MOYER C B, RINDAL R A. An analysis of the coupled chemically reacting boundary layer and the charring ablator: NASA CR-1061[R]. Washington, D.C.: NASA, 1968.
- [18] 刘骁, 国义军, 刘伟, 等. 碳化材料三维烧蚀热响应有限元计算研究[J]. 宇航学报, 2016, 37(9): 1150-1156.
 LIU X, GUO Y J, LIU W, et al. Numerical simulation research on three-dimensional ablative thermal response of charring ablators[J]. Journal of Astronautics, 2016, 37(9): 1150-1156(in Chinese).
- [19] HENDERSON J B, WIEBELT J A, TANT M R. A model for the thermal response of polymer composite materials with experimental verification[J]. Journal of Composite Materials, 1985, 19(6): 579-595.
- [20] 孙斌, 于聪, 周王超, 等. NTC 热敏电阻特性曲线的拟合方法研究
 [J]. 中国计量学院学报, 2012, 23(1): 75-79.
 SUN B, YU C, ZHOU W C, et al. Study on curve fitting methods of NTC thermistor characteristics[J]. Journal of China University of Metrology, 2012, 23(1): 75-79(in Chinese).
- [21] TRAN H K, JOHNSON C, RASKY D, et al. Phenolic impregnated carbon ablators pica as thermal protection system for discovery missions: NASA TM-110440[R]. Washington, D.C.: NASA, 1997.

Identification method of ablation performance parameters of carbonized materials

CHENG Mohan, LI Wenguang^{*}, WANG Zhi

(School of Aerospace Engineering, Beijing Institute of Technology, Beijing 100081, China)

Abstract: In response to the difficulty in measuring the ablation performance parameters of thermal protection materials, this paper studies the identification method of ablation performance parameters based on optimization algorithms. The one-dimensional continuous model is used to describe the internal thermal response of carbonization ablation thermal protective materials. Compared with the public experimental data, the calculation error of each measurement point is less than 15%. The basis function characterisation approach uses the Chebyshev polynomial as the basis function to characterize the material's specific heat capacity and thermal conductivity both before and after carbonization. According to the temperature test experimental data of PICA materials in the public literature, the pyrolysis kinetic parameters of the materials and the undetermined coefficients of the basis function are identified by genetic algorithm, and the parameter identification results are substituted into the one-dimensional continuous model for calculation. Compared with the experimental data in the public literature, the average relative error of the temperature curves under the two verification conditions is less than 10%.

Keywords: thermal protective material; ablation performance; parameter identification; optimization algorithm; numerical calculation

Received: 2022-06-11; Accepted: 2022-08-21; Published Online: 2022-11-01 13:12

URL: link.cnki.net/urlid/11.2625.V.20221031.1437.002

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

DOI: 10.13700/j.bh.1001-5965.2022.0486

复杂多方向威胁下的导弹预警雷达优化部署方法

刘伟^{1,2},刘昌云³,郭相科^{3,*},樊良优¹,何晟¹,兰昊²

(1. 空军工程大学研究生院,西安710000; 2. 中国人民解放军 63768 部队,西安710000;

3. 空军工程大学 防空反导学院,西安 710000)

摘 要: 针对现有导弹预警雷达部署相对独立、协同困难,难以满足大规模对抗场景的现状,从远程预警雷达、跟踪识别雷达、机动式预警雷达不同的任务特点出发,建立应对复杂多方向 威胁的多型导弹预警雷达优化部署模型,在满足最优覆盖、协同交接、目标识别等任务约束下,解 决雷达协同部署问题。针对所提模型设计了一种基于云自适应的分区优化离散粒子群(CPBPSO)算法, 通过设计分区编码策略缩减算法求解空间、加入云自适应变异算子提高算法全局寻优和局部跳出能 力,使算法更适用于导弹预警雷达部署问题的处理。实例验证了所提模型在求解单方向、多方向威 胁场景部署问题的可行性,对比分析了CPBPSO 算法的有效性,基本满足导弹预警雷达最优化协同部署的需求。

关 键 词:导弹预警雷达;协同预警;优化部署模型;云自适应变异;粒子群算法

中图分类号: TN959.1

文献标志码: A 文章编号: 1001-5965(2024)04-1392-13

战略预警是当前及今后空天防御作战的研究 热点,弹道导弹预警又是战略预警的重中之重。导 弹预警雷达是导弹预警体系的骨干设备,其部署情 况决定着体系作战效能的发挥,在应对复杂导弹威 胁时,合理的雷达部署方案将能更好地完成"早预 警、准识别、稳交接"的作战目标^[1]。随着面临的导 弹威胁日益增多,如何科学设计导弹预警雷达部署 方案,满足多型雷达协同预警、多方向威胁协同探 测任务,实现预警资源的优化配置,是当前亟待解 决的问题。

现有导弹预警雷达部署方法研究相对较少,多 借鉴于防空雷达部署方法,在部署战略作用大、设 备结构杂、协同要求高的导弹预警雷达时,存在可 借鉴参考理论不足的问题。传统的雷达优化部署 方法以应对高机动、全空域的空气动力学目标的组 网部署问题为主^[24],主要是以威力衔接、重叠系数 等因素为约束,以雷达空域覆盖最优化为目标的部 署问题。例如,文献 [5] 基于空间覆盖率、覆盖重叠 率、频率干扰系数,将部署问题转化为多目标组合

优化问题: 文献 [6] 采用基于最优覆盖率的防空雷 达部署模型,对不同地形条件下的部署方案进行优 化评估; 文献 [7] 以联合探测概率、地形遮蔽等为约 束,提出基于最优覆盖的多传感器优化配置方法。 此外,在部署优化模型求解时,由于求解计算量大, 传统算法稳定性不足,大多数研究采用智能优化算 法进行问题求解,如混沌遗传算法 (chaos genetic algorithm, CGA)^[8]、改进粒子群算法^[9]、烟花算法^[10] 等,提高了搜索效率和性能。但是,上述研究还存 在以下问题:①在反导作战中,导弹预警雷达面对 的是弹道导弹这类攻击方向较为固定、弹道相对可 预测、地理跨度大的目标,因此,更加关注对某个威 胁方向的目标覆盖情况,纯粹追求探测空域全覆盖 则不适用于导弹预警雷达的部署与资源的利用; ②导弹预警体系中包含多型预警设备,强调基于预 报交接、接力跟踪的体系作战运用,因此,需建立针 对各型雷达任务特点的协同部署方法。鉴于此,本 文从导弹预警作战任务出发,基于弹道覆盖和多型 雷达协同探测需求,建立应对复杂导弹威胁的预警

收稿日期: 2022-06-12; 录用日期: 2022-08-10; 网络出版时间: 2022-09-08 13:32 网络出版地址: link.cnki.net/urlid/11.2625.V.20220907.1131.001

基金项目: 国家自然科学基金 (62106283)

*通信作者. E-mail: guosyanyu@163.com

引用格式:刘伟,刘昌云,郭相科,等.复杂多方向威胁下的导弹预警雷达优化部署方法 [J]. 北京航空航天大学学报,2024,50(4):1392-1404. LIU W, LIU CY, GUO X K, et al. Deployment optimization method for missile early warning radar under complex and multidirectional missile threats [J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50(4): 1392-1404 (in Chinese). 雷达优化部署方法。①提出一种基于协同探测、目标识别、关机点截获等约束的多型导弹预警雷达优化部署模型,解决复杂多方向威胁条件下的导弹预 警最优部署问题;②提出一种基于云自适应的分区 优化离散粒子群(cloud adaptive partition optimization binary particle swarm optimization, CPBPSO)算法的 模型求解方法,在编码策略、粒子变异等方面对算 法进行改进,该算法在处理大规模部署问题中具有 较好的求解性能。

1 问题分析

1.1 作战任务分析

导弹预警雷达属于探测系统中的骨干力量,承 担着弹道导弹发射后的早期预警、连续跟踪、精确 识别、拦截引导及拦截效果评估等任务。由于受雷 达设备性能、地理环境等因素限制,单一导弹预警 雷达只能完成导弹预警某个阶段的任务,需要针对 不同阶段的导弹目标特性,配置不同的雷达完成相 应的探测任务。典型地基预警雷达探测能力与弹 道导弹飞行距离之间的对比关系如图1所示。

从图 1 可以看出,对于飞行距离长达几千千米 的弹道导弹目标,单部雷达无法满足对导弹的探测 覆盖,特别是对于来自多个威胁方向的复杂对抗场 景,多型雷达协同组网探测显得尤为重要。雷达部 署问题的实质是:在保证探测资源最优的前提下, 能够满足复杂多方向威胁场景下的各型雷达的作 战需求,达到"尽早发现告警、威胁空域覆盖、全程 接续跟踪、精确识别目标"等预警作战任务。数学 模型表示为

$$\max Z(\mathbf{x}) = \max\{z_1(\mathbf{x}_1), z_2(\mathbf{x}_2), \cdots, z_n(\mathbf{x}_n)\}$$

s.t. $\mathbf{x}_i \in f(\mathbf{x}), i = 1, 2, \cdots, n$ (1)

式中:Z(x)为目标函数;f(x)为约束函数; z_n 为n维决策变量函数。







各型雷达之间通常难以同时取得最优,因此, 需根据各型雷达功能和预警任务特点,建立协同探 测约束,实现对多型预警雷达的最优化部署。

1.2 典型导弹预警雷达

导弹预警雷达以地基预警雷达为主,包括远程 预警雷达、跟踪识别雷达和机动式预警雷达等,其 作战任务各不相同[11-13],如图2所示。远程预警雷 达通常工作在 P、L 波段,具有作用距离远、预警范 围广的特点,主要用于对视距范围内的弹道导弹中 段飞行的预警探测,为导弹预警系统提供早期预警 信息,如"铺路爪"雷达。跟踪识别雷达通常工作 在 X 波段,具有较强的目标探测能力,可对目标进 行精确跟踪与识别,主要用于获取弹道导弹中段到 末段过渡区域的目标精确信息,对后续的导弹飞行 轨迹进行精确预测,也用于对拦截系统的引导,如 GBR-P 雷达等。机动式预警雷达通常工作在 P、S、 X 波段,具有机动性高、生存性强等特点,可根据防 御重点灵活部署,在前置预警、中段补盲、末段引 导等方面具有得天独厚的优势,如 AN/TPY-2 雷达 等。对于空基、海基导弹预警设备,本文不作重点讨 论。在导弹预警指控系统的紧密交链下,各型雷达



图 2 导弹预警系统协同预警示意图

Fig. 2 Schematic diagram of missile early warning system coordinated early warning

密切协同、交接引导,共同完成导弹预警作战任务。

2 量化指标分析

2.1 导弹威胁空域

在已知弹道导弹来袭方向和我方重点保卫要地的情况下,可通过典型弹道分析,基本推算出该方向所有来袭导弹的弹道集合,即威胁空域。如图3所示,已知导弹攻击方有*E*₁、*E*₃2个导弹发射井及在*E*₂、*E*₄附近活动的多部导弹发射车,则 *E*₁*E*₂*E*₃*E*₄为所有可能的导弹发射区域。当攻击方发射导弹打击防御方的保卫要地F时,来袭导弹所有可能的弹道(高弹道、最优能量弹道、低弹道)均在*E*₁*E*₂*E*₃*E*₄*F*威胁空域内。由于该空域难以用数学公式表达,本文采用无数条从发射区域射向保卫要地的典型弹道来逼近表示威胁空域。为便于计算,可选取威胁空域内*N*_{bm}条弹道的覆盖需求,则认为能够实现对该威胁空域的覆盖。



2.2 弹道覆盖模型

选取 2.1 节威胁空域模型内某条典型弹道*l*, 建立以下分析:以雷达所在位置 *O* 为原点,建立东-北-天坐标系*Oxyz*,其中, *x*轴指向正东方向, *y*轴指 向正北方向, *z*轴垂直于 *xOy*平面指向正上方向,如 图 4 所示。图中,阴影部分为地基预警雷达的威力 范围,可由包含等效探测距离 *r*、方位角 *A*、俯仰角 *E* 3 个变量的等效空间 *Q*_R表示,其中,不连续区域为





地形遮蔽影响,表达式为

$$\boldsymbol{\Omega}_{\mathrm{R}} = \left\{ (r, A, E) \middle| r \leqslant \sqrt{-\frac{\sigma p_{\mathrm{T}}}{2\sigma_{\mathrm{n}}^{2}\ln(1-p_{\mathrm{d}})}}, A \in [\alpha_{\min}, \alpha_{\max}], \\
E \in E_{\mathrm{g}} = \left\{ \begin{bmatrix} \max(\beta_{\min}, \beta_{\mathrm{gmax}, \alpha_{i}}), \beta_{\max} \end{bmatrix} & \beta_{\mathrm{gmax}, \alpha_{i}} < \beta_{\max} \\ \varnothing & \beta_{\mathrm{gmax}, \alpha_{i}} > \beta_{\max} \\ \end{bmatrix} \right. \\$$

$$(2)$$

式中: σ 为目标典型雷达散射截面 (radar cross section, RCS)值; σ_n 为噪声振幅的均方根值; p_T 为发 射机功率; p_d 为雷达检测概率; $[\alpha_{\min}, \alpha_{\max}]$ 为雷达俯 仰范围; E_g 为经过地形遮蔽修正的俯仰范围; β_{gmax,α_i} 为雷达在 α_i 方位上的最大遮蔽角^[14], 计算公 式为

$$\beta_{\text{gmax},\alpha_i} = \arctan\left(\frac{(r_{\text{e}} + h_{G\max,\alpha_i})\cos\gamma - r_{\text{e}} - h_{\text{R}}}{(r_{\text{e}} + h_{G\max,\alpha_i})\sin\gamma}\right) \quad (3)$$

其中: *r*_e为等效地球半径; *h*_{Gmax,α_i}为该方位的最高遮 蔽物 G的高程; *h*_R为雷达部署海拔; γ为遮蔽物与 雷达间的地心夹角。

地形遮蔽影响探测能力如图5所示。



图 5 地形遮蔽影响探测能力

Fig. 5 Terrain masking affects detection capability

设D为导弹飞行弹道上的任一点,该点相对雷达的位置状态为

$$\boldsymbol{R}_{D} = \begin{bmatrix} r \\ \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} |OD| \\ \arccos\left(\frac{|ON|^{2} + |OD'|^{2} - |ND'|^{2}}{2|OD'||ON|}\right) \\ \arctan\left(\frac{|DD'|}{|OD'|}\right) \end{bmatrix}$$
(4)

式中: D'点为D点在 xOy平面的投影点; ON为雷达 阵面法向在 xOy平面的投影。

所有满足 $R_D \subset \Omega_R$ 的弹道点,即为雷达对导弹的通视覆盖弧段,如图 4 中弹道弧段 $\widehat{D_sD_e}$ 。由于地形的遮蔽,弹道弧段可能被分为 n_g 个不连续的弧段,其总长度表示为

5)

$$I_{\text{detect}} = \sum_{j=1}^{n_{\text{g}}} \int_{t_{D_{\text{g},j}}}^{t_{D_{\text{g},j}}} |l'(t)| \,\mathrm{d}t \tag{1}$$

式中:曲线l'(t) = (x'(t), y'(t), z'(t))为经典弹道方程的 微分形式; D_{s_j}, D_{e_j} 分别为雷达对第j组观测弧段 的首、末点; $t_{D_{s_j}}, t_{D_{e_j}}$ 为其对应的导弹飞行时刻。

威胁弹道覆盖率可表示为

$$C_{\rm r} = \frac{l_{\rm detect}}{l_{\rm missile}} \tag{6}$$

雷达提供的预警时间比率为

$$E_{\rm r} = \frac{t_{\rm missile} - t_{D_{\rm s_{1}}}}{t_{\rm missile}} \tag{7}$$

雷达对目标的持续跟踪率为

$$T_{\rm r} = \frac{t_{D_{\rm c,n}} - t_{D_{\rm s,l}}}{t_{\rm missile}} \tag{8}$$

式中: *l*_{missile}为导弹飞行全程弹道弧长; *t*_{missile}为导弹飞行时间。

2.3 基于微多普勒频率的高识别弹道覆盖率

研究表明^[15-16],基于中段微多普勒特性的目标 识别能力与雷达观测角度有关,而雷达的部署位置 决定了目标相对雷达的观测角度和空间变化规律, 因此,可通过调整雷达部署位置,使目标处于有利 于识别的观测角度内。由于带宽的限制,本文仅讨 论跟踪识别雷达在中段目标识别的作用。文献 [16] 将弹头目标一个进动周期内微多普勒频率的极差 定义为微多普勒调制带宽 B_{nd} ,当 $B_{nd} \in [B_{ndmax}/2, B_{ndmax}]时,越容易观测到目标的微动特性。由于目$ 标在中段飞行过程中,几何参数和进动参数基本不 $变,<math>B_{nd}$ 由雷达平均视线角 ϕ 唯一确定,因此, ϕ 越接 近90°,微多普勒调制带宽越大,越有利于目标识 别。建立以下分析:设弹头为旋转对称体,建立以 弹头质心为原点的Oxyz进动坐标系和弹体坐标系,如 图 6 所示。

平均视线角ø由导弹相对于雷达的位置矢量 p和进动轴的方向矢量q确定,表达式为

$$\phi = \arccos\left(\frac{\boldsymbol{p} \cdot \boldsymbol{q}}{|\boldsymbol{p}| \cdot |\boldsymbol{q}|}\right) \tag{9}$$

当弹道上一点S满足 $R_S \subset \Omega_R$,且其微多普勒调制带宽 $B_{md_S} \in [B_{md max}/2, B_{md max}]$ 满足式(10)时,可认为雷达对处于S点目标的微多普勒识别能力较高。

$$B_{\text{md}_S} = \frac{4\omega \sin\theta \sin\left(\arccos\left(\frac{\boldsymbol{p} \cdot \boldsymbol{q}}{|\boldsymbol{p}| \cdot |\boldsymbol{q}|}\right)\right)}{\lambda} \left| d_{\text{h}} + d_{\text{v}} \frac{\cos\theta}{\sin\theta} \right|$$
(10)

式中: θ为弹头进动角; λ为波长; ω为进动角速度; d_v和 d_h分别为目标散射中心到质心的垂直距离和 水平距离。





定义满足上述要求的所有弹道弧段的集合为 雷达微多普勒高识别弧段,其弹道覆盖率为

$$I_{\rm r} = \frac{l_{\rm iden}}{l_{\rm detect}} \tag{11}$$

式中: liden为高识别弧段的长度。

2.4 雷达协同交接系数

雷达协同交接方式主要有握手交接和预测交 接^[17-18]。本文基于资源最优化考虑,雷达间均采用 预测交接。预测交接指前站雷达和后站雷达没有 足够的共视弹道,后站雷达主要依靠前站雷达探测 生成的预测弹道进行目标搜索,如图 7 所示。由于 测量误差的影响,导弹状态误差会在预报中不断放 大并形成误差管道,后站雷达需根据交接时刻导弹 状态协方差来确定搜索空域。因此,可通过后站搜 索能力和前站预测误差,推算最优目标交接时刻。

设前站预测的弹道在交接时刻 $t_{D_e}+\Delta t_e$ 的方位、 俯仰误差均方根为($\Delta \phi_e, \Delta \eta_e$)。为保证目标落入后 站截获屏的概率 $p_{cross} = 0.997$,后站截获屏单边厚度 应为交接误差均方根的3倍,考虑方位、俯仰两维 截获,截获屏厚度设置为7 $\Delta \phi_e \times 7\Delta \eta_e$ 。后站雷达波 束宽度为($\Delta \theta, \Delta \xi$),单个波位驻留时间为 T_{θ} 、检测概 率为 p_d ,则单屏搜索时间为

$$T_{\rm c} = \frac{7\Delta\phi_{\rm c} \times 7\Delta\eta_{\rm c}}{\Delta\theta \times \Delta\xi} T_{\theta} \tag{12}$$

为满足截获概率为p_{capture},需进行n屏搜索积累,即

$$n = \left[\frac{\ln\left(1 - \frac{p_{\text{capture}}}{p_{\text{cross}}}\right)}{\ln(1 - p_{\text{d}})}\right]$$
(13)





式中:[]为向上取整运算。而n屏搜索时间需小于 目标穿越搜索屏时间 T_p ,即 $nT_c \leq T_p$ 。因此,只有交 接时刻前站的预测误差($\Delta \phi_c, \Delta \eta_c$)满足式(14)时,方 可实现在交接时刻截获目标。

$$(\Delta\phi_{\rm c}, \Delta\eta_{\rm c}) \leq \arg\min\left(\left|n\frac{7\Delta\phi_{\rm c}\times7\Delta\eta_{\rm c}}{\Delta\theta\times\Delta\xi}T_{\theta} - \frac{r_{\rm m}\times\Delta\varphi}{v}\right|\right)$$
(14)

式中:v为导弹穿屏时的飞行速度;r_m为目标与雷达间的斜距。

定义 Δt_c 为最优预测交接时间间隔,当预测段飞行时间 $\Delta t_h \leq \Delta t_c$ 时,交接时刻前站的预测误差不大于($\Delta \phi_c, \Delta \eta_c$),后站可以 $p_{capture}$ 的概率截获目标。故雷达协同交接系数定义为

$$H_{\rm r} = \frac{\Delta t_{\rm c} - \Delta t_{\rm h}}{\Delta t_{\rm c}} \tag{15}$$

3 导弹预警雷达优化部署模型

根据上述问题分析,建立针对威胁空域的远程 预警雷达、跟踪识别雷达、机动式预警雷达优化部 署模型及协同预警部署模型。

3.1 单型导弹预警雷达部署模型

3.1.1 远程预警雷达部署模型

远程预警雷达主要执行对导弹目标的早期预 警任务,跟踪监视来袭导弹目标,为导弹预警体系 提供相应的早期预警信息。因此,在部署时应满足 尽早发现导弹目标、覆盖尽量多的弹道等任务需 求。针对某个威胁空域的远程预警雷达,部署位置 应满足如下目标函数:

$$\max Z_{\rm pr} = \sum_{i=1}^{N_{\rm bm}} (\lambda_{\rm p1} C_{\rm rpr,i} + \lambda_{\rm p2} E_{\rm rpr,i} + \lambda_{\rm p3} T_{\rm rpr,i})$$
(16)

式中: Z_{pr} 为远程预警雷达探测收益; $C_{rpr,i}$ 、 $E_{rpr,i}$ 、 $T_{rpr,i}$ 分别为远程预警雷达对威胁空域内第 *i* 条弹道 的弹道覆盖率、预警时间比率、持续跟踪率; λ_{p1} 、 λ_{p2} 、 λ_{p3} 为权重系数, 且 $\lambda_{p1} + \lambda_{p2} + \lambda_{p3} = 1$ 。 3.1.2 跟踪识别雷达部署模型

跟踪识别雷达主要执行对导弹目标的跟踪识 别任务,精确识别弹头目标,为部署在保卫要地周 边的反导拦截系统提供精确的弹道信息和目标指 示信息。跟踪识别雷达部署时,应覆盖尽可能多的 弹道高概率识别区,并满足对拦截系统的引导和拦 截效果的评估等任务需求。针对某个威胁空域的 跟踪识别雷达,部署位置应满足如下目标函数:

$$\max Z_{\rm xr} = \sum_{i=1}^{N_{\rm bm}} \left(\lambda_{\rm x1} C_{\rm rxr,i} + \lambda_{\rm x2} T_{\rm rxr,i} + \lambda_{\rm x3} I_{\rm rxr,i} \right)$$
(17)

式中: Z_{xr} 为跟踪识别雷达探测收益; $C_{rxr,i}$ 、 $T_{rxr,i}$ 、 $I_{rxr,i}$ 分别为跟踪识别雷达对威胁空域内第i条弹道 的弹道覆盖率、持续跟踪率、高识别弹道覆盖率; λ_{x1} 、 λ_{x2} 、 λ_{x3} 为权重系数,且有 $\lambda_{x1} + \lambda_{x2} + \lambda_{x3} = 1$ 。 3.1.3 机动式预警雷达部署模型

机动式预警雷达可执行前置预警或补盲预警 任务。前置预警是在能获取导弹关机点状态观测 信息的情况下,将机动式预警雷达前出部署获取导 弹关机点信息,从而准确快速地估计导弹后续的运 动轨迹,为导弹预警体系提供精确的弹道预测信 息。补盲预警是在应对新增导弹威胁且短期内无 法形成对该威胁空域的覆盖时,应急调用机动式预 警雷达对该空域的探测补盲,快速构建导弹预警体 系作战能力。由于补盲预警雷达部署需求与远程 预警雷达类似,在探测弹道时,侧重于观测预警体 系未覆盖的威胁空域,本文不作讨论。

在机动式预警雷达用作前置预警部署时,部署 位置应满足如下探测需求:

$$\max Z_{\rm fr} = \sum_{i=1}^{N_{\rm bm}} \left(\lambda_{\rm f1} C_{\rm rfr,i} + \lambda_{\rm f2} T_{\rm rfr,i} + \lambda_{\rm f3} E_{\rm rfr,i} + \lambda_{\rm f4} L_{\rm rfr,i} \right)$$
(18)

式中: $Z_{\rm fr}为前置预警雷达探测收益; C_{\rm ff,i}, T_{\rm ff,i}, E_{\rm ff,i}分别为前置预警雷达对威胁空域内第$ *i*条弹道的弹道覆盖率、持续跟踪率、预警时间比率; $<math>L_{\rm rf,i}$ 为雷达对导弹关机点 $D_{\rm off,i}$ 距离与雷达最大等效 探测能力的比值,距离越近越有利于探测到关机 点; $\lambda_{\rm f1}, \lambda_{\rm f2}, \lambda_{\rm f3}, \lambda_{\rm f4}$ 为权重系数,且 $\lambda_{\rm f1} + \lambda_{\rm f2} + \lambda_{\rm f3} + \lambda_{\rm f4} = 1$ 。

3.1.4 部署约束条件

各型预警雷达的部署位置*P*_r均需满足以下约 束:①只可部署在我方可部署区域*S*_{ed}内,以免遭受 打击;②不可部署在河流、湖泊等不具备部署条件 的地区*S*_{nd}。

同时,前置预警雷达的部署位置*P*_{fr}还应满足对 第*i*条弹道的关机点*D*_{off,i}探测覆盖的约束;跟踪识 别雷达部署位置*P*_{xr}还应满足对拦截系统射击范围 探测覆盖的约束,即能够同时探测到拦截系统远、 近界与第*i*条弹道的交点*D*_{cls,i}、*D*_{far,i},从而确保对反 导拦截系统的引导能力。

以上约束可表示为

$$\begin{cases} P_{\rm r} \in S_{\rm ed}, P_{\rm r} \notin S_{\rm nd} \\ P_{\rm fr} \in \{P_{\rm f} | \forall D_{\rm off,i} \in \boldsymbol{\Omega}_{P_{\rm f}} \} \\ P_{\rm xr} \in \{P_{\rm x} | (\forall D_{\rm cls,i} \in \boldsymbol{\Omega}_{P_{\rm c}}) \cap \forall (D_{\rm cls,i} \in \boldsymbol{\Omega}_{P_{\rm c}}) \} \end{cases}$$
(19)

式中: Ω_{P_t} 、 Ω_{P_x} 分别为部署在 P_t 的跟踪识别雷达的 探测威力空间、部署在 P_x 的跟踪识别雷达的探测威力空间。

3.2 协同预警优化部署

3.2.1 单方向协同预警优化部署模型

在协同预警作战场景下,远程预警雷达起着 "衔接"前置预警雷达和跟踪识别雷达的作用。因 此,可根据前置预警雷达的预测误差推算最优交接 时刻的弹道点D_{ps.i},即为远程预警雷达的探测首点; 根据跟踪识别雷达所需的目标交接需求反推前站 弹道末点D_{pe,i},即为远程预警雷达的探测末点。远 程预警雷达只有部署在能够同时覆盖D_{ps,i}、D_{pe,i}的 部署区域,才能有效衔接体系内的预警雷达。且雷 达协同交接系数H_r越高,越有利于目标的引导交接。 基于上述讨论,建立单方向来袭协同部署模型为

$$\max Z_{\rm sd} = \omega_1 Z_{\rm pr} + \omega_2 Z_{\rm xr} + \omega_3 Z_{\rm fr} + \omega_4 H_{\rm rf_p} + \omega_5 H_{\rm rp_x}$$
(20)

式中: Z_{sd} 为单方向协同预警雷达探测收益; H_{ff_p} 、 H_{rp_x} 分别为前置预警雷达对远程预警雷达交接的 协同交接系数、远程预警雷达对跟踪识别雷达交接 的协同交接系数; ω_1 、 ω_2 、 ω_3 、 ω_4 、 ω_5 为权重系数, 且有 $\omega_1 + \omega_2 + \omega_3 + \omega_4 + \omega_5 = 1$ 。

约束条件在满足式(19)的同时,还应满足:

 $P_{pr} \in \{P_{x} | (\forall D_{ps,i} \in \mathbf{\Omega}_{P_{p}}) \cap \forall (D_{pe,i} \in \mathbf{\Omega}_{P_{p}})\}$ (21) 式中: $\mathbf{\Omega}_{P_{p}}$ 为部署在 P_{p} 位置的远程预警雷达的探测 威力空间。

3.2.2 多方向协同预警优化部署模型

在面对多方向来袭的作战场景时,考虑到单部 雷达可兼顾多个威胁方向的预警任务,在配置预警 雷达时可进一步优化设备数量。优化雷达部署数 量的方法如下:对来自M个方向的威胁进行威胁空 域划分,逐空域计算各雷达的可部署区域。依次判 断第n部雷达对于第m个威胁方向的可部署区域 P_{n,m}和第m+1个方向可部署区域P_n,若有,则雷达部署区域P_n,有无重叠区 域P_n,若有,则雷达部署区域P_n内可兼顾 2 个威胁 方向;若无,则需在该方向上部署第n+1部雷达。 依次判断M个威胁,可得某型雷达部署的最小数量 N,以及每部雷达的可部署区域P_n(n = 1,2,…,N)。 由于目标容量的限制,任务分配方案最后还需根据 作战实际进行调整。

在确定雷达最优数量和部署约束后,在3.1节 讨论的单方向协同部署方法的基础上,按照每部雷 达最优部署和体系协同部署的优化需求,建立多方 向协同部署的模型如下:

$$\max Z_{md} = \sum_{m=1}^{M} \sum_{l=1}^{N_{bm}} (\omega_{m1} Z_{pr,l}^{m,n_{p}} + \omega_{m2} Z_{xr,l}^{m,n_{x}} + \omega_{m3} Z_{fr,l}^{m,n_{f}} + \omega_{m4} H_{rfp,l}^{m} + \omega_{m5} H_{rpx,l}^{m})$$
s.t.
$$\begin{cases}
P_{pr,n_{p}} \in P_{n_{p}}, P_{xr,n_{x}} \in P_{n_{x}}, P_{fr,n_{f}} \in P_{n_{f}}, \\
P_{pr} \notin S_{nd}, P_{xr} \notin S_{nd}, P_{fr} \notin S_{nd} \\
n_{p} = 1, 2, \cdots, n_{p}, n_{x} = 1, 2, \cdots, N_{x}, \\
n_{f} = 1, 2, \cdots, N_{f}
\end{cases}$$
(22)

式中: Z_{md} 为多方向协同预警雷达探测收益; $Z_{pr,l}^{m,n_p}$ 、 $Z_{xr,l}^{m,n_t}$ 、 $Z_{fr,l}^{m,n_t}$ 分别为第 n_p 、 n_x 、 n_f 部雷达面对第 m个威胁方向中第l条弹道的探测目标函数; $H_{fr,l}^m$ 、 $H_{rpx,l}^{m}$ 为第*m*个威胁方向第*l*条弹道的协同交接系数; N_{p} 、 N_{x} 、 N_{f} 分别为三型雷达最小部署数量; ω_{m1} 、 ω_{m2} 、 ω_{m3} 、 ω_{m4} 、 ω_{m5} 为权重系数,且有 $\omega_{m1} + \omega_{m2} + \omega_{m3} + \omega_{m4} + \omega_{m5} = 1$ 。

4 基于 CPBPSO 算法的模型求解方法

本文所研究的部署模型是一个多目标组合优 化问题,随着待部署雷达和威胁目标数量的增加, 问题的求解难度会急剧增加,传统的枚举法很难求 出最优解。群智能算法在求解组合优化问题中有较 好的表现。其中,粒子群(particle swarm optimization, PSO)算法^[19]相比于遗传算法、差分进化算法,具有 控制参数少、收敛速度快、易实现等优点^[20],更适 用于本文所研究问题。针对协同预警模型的特点 和求解要求,提出一种基于雷达部署参数的粒子编 码策略,并根据雷达威力空间特点进行编码分区, 实现了部署参数和粒子状态的有效映射。同时,为 解决传统 PSO 算法易陷入局部最优和早熟收敛的 问题,将云自适应变异算子和 PSO 算法结合,提出 CPBPSO 算法,具有较好的全局搜索能力和局部跳 出能力。

4.1 粒子群算法

假设在一个 a 维的目标搜索空间中, 由 N_s 个粒 子组成一个群落^[21-24], 其中, 第i个粒子的位置为 $X_i = (x_{i1}, x_{i2}, \cdots, x_{ia})$, 速度为 $V_i = (v_{i1}, v_{i2}, \cdots, v_{ia})$, 此 时, 每个粒子 X_i 表示一种雷达部署方案, 包含各型 雷达部署的经纬度、阵面法向信息。根据 3.2 节建 立的模型, 可以计算每个粒子的适应度值, 即每个 部署方案的协同探测收益。适应度值越大, 代表该 部署方案越好。采用多方向协同预警优化部署模 型时, 适应度函数表示为

$$Z_{\mathrm{md},i} = Z_{\mathrm{md}}(X_i) \tag{23}$$

粒子的状态更新公式为

$$\begin{cases} v_{id}^{(t+1)} = wv_{id}^{(t)} + c_1 r_1(p_{id}^{(t)} - x_{id}^{(t)}) + c_2 r_2(g_{id}^{(t)} - x_{id}^{(t)}) \\ x_{id}^{(t+1)} = \begin{cases} 1 & \text{randn}(0, 1) < \frac{1}{1 + e^{v_{id}^{(t)}}} \\ 0 & \notin \& \end{cases}$$
(24)

式中: d = 1,2,...,a; w 为惯性因子,采用线性递减权

当算法达到终止条件时,可以得到适应度值最 大的粒子**G**_T,即最优部署方案。

4.2 基于雷达部署参数的粒子分区编码策略

由于雷达的部署位置、阵面法向等输入参数 均为离散变量,本文采用一种等价表示待优化参数 的离散二进制编码方法。将离散的雷达部署参 数转码为一维行向量,各参数占据不同码位。编 码策略如下:每部雷达部署点位编码位数为 N_d = $[log_2(S_{able}/S_{radar})],阵面法向间隔编码位数为<math>N_a$ = $[log_2(2\pi/n_d)],其中,S_{able}$ 为雷达可部署域的总面积, S_{radar} 为雷达部署密度, n_d 为阵面法向密度。此时, 编码的粒子总维数为

$$a = N_{\rm p} \times (N_{\rm d_p} + N_{\rm a}) + N_{\rm x} \times (N_{\rm d_x} + N_{\rm a}) + N_{\rm f} \times (N_{\rm d_f} + N_{\rm a})$$
(25)

式中: N_p 、 N_x 、 N_f 为三型雷达的部署数量; N_{d_p} 、 N_{d_x} 、 N_{d_f} 为可部署点位的数量。

粒子的映射关系如图 8 所示。图中: P_{i_lat} 、 P_{i_lon} 、 P_{i_nor} 分别为第 i 部远程预警雷达的纬度、经度、法 向方向; X_{j_lat} 、 X_{j_lon} 、 X_{j_nor} 分别为第 j 部跟踪识别雷 达的纬度、经度、法向方向; F_{k_lat} 、 F_{k_lon} 、 F_{k_nor} 分别 为第 k 部机动式预警雷达的纬度、经度、法向方向。

但是,随着问题规模的扩大,算法的运算量、寻 优难度将进一步增加。为减少粒子寻优的空间,结 合本文模型,提出一种粒子分区编码策略。按照资 源最优化原则,雷达不应部署在同型雷达的威力空 间内,而是在各威胁空域分散开来,如图9所示。

如果2部同型雷达部署过于靠近,其探测空间 将有较大重叠,导致探测资源冗余。因此,对于第 N部同型雷达,粒子空间划分区域为 $\left\{P_{r}-\bigcup_{j=1}^{N}C_{j}\right\}$ 。 以远程预警雷达为例,在分区编码策略下,其粒子 空间映射关系如下:





Fig. 8 Particle mapping relationships









$$\begin{cases} [x_{p11}, x_{p12}, \cdots, x_{p1N_d}] \rightarrow \{P_{pr}\} \\ [x_{p21}, x_{p22}, \cdots, x_{p2N_d}] \rightarrow \{P_{pr} - C_1\} \\ \vdots \\ [x_{pi1}, x_{pi2}, \cdots, x_{piN_d}] \rightarrow \begin{cases} P_{pr} - \bigcup_{j=1}^{N_d} C_j \end{cases} \end{cases}$$

$$(26)$$

式中: P_{pr}为远程预警雷达可部署区域; C_j为第 j部 雷达威力空间在地面的投影范围。

采用本文分区编码策略,能够有效缩减搜索空间至2^{$\sum_{N_k \times (N_a + N_{a,k} - N_{c,k})}$}($N_k \times N_{d_k} \times N_{c_k}$ 分别为第k型雷达的数量、部署点位和探测覆盖区域编码位数),如 $N_p = 3 \times N_f = 5$, C_j 所包含点位面积约为0.25 S_{able} 、0.125 S_{able} ,则其总搜索空间将降至7.69%。采用分区粒子编码,避免了搜索空间过大的问题,提高寻优搜索的效率。

4.3 基于云自适应变异的粒子更新策略

为提高 PSO 算法的全局寻优和快速收敛能力, 在迭代过程中,必须使适应度小的粒子有较大的变 异,适应度大的粒子有较小的变异。为此,本文借 鉴云模型思想^[25-26],提出云自适应变异操作。在云 模型中,期望值 E_x 反映了云层的中心位置,熵 E_n 反 映了云层的离散程度,超熵 H_e 反映了云层的厚度, 如图 10 所示。当 $x > E_x$ 时,隶属度随x的增大而减 小;且云模型中云滴集中在 $[E_x - 3E_n, E_x + 3E_n]$,具有 较好的随机性和稳定倾向。将云模型这一特征运 用在粒子的变异中,适应度小的粒子进行较大的变 异,以求生成适应度大的粒子,反之,要求变异较 小,保留优良粒子。

云模型数学描述如式(27)所示:

$$\begin{cases} E_x = Z_{\text{md},t} \\ E_n = (Z_{\text{md},g} - \overline{Z}_{\text{md},t})/e \\ H_e = E_n/h \\ E'_n = \text{randn}(E_n, H_e) \end{cases}$$
(27)



式中: $\overline{Z}_{md,r}$ 为当前平均适应度; $Z_{md,g}$ 为当前最大适应 度; e、h为控制参数, e控制云的陡峭程度, 根据 " $3E_n$ "规则, 一般取 3, h控制云层的厚度, 一般为 10。

执行云自适应变异操作,即粒子 $X_i = (x_{i1}, x_{i2}, \cdots, x_{ia})$ 变异后得到 $X'_i = (x'_{i1}, x'_{i2}, \cdots, x'_{ia})$,具体如式(28)所示:

$$\begin{cases} x'_{id} = \begin{cases} x_{id} & \rho_i N_i(0,1)\varphi(Z_{\mathrm{md},i}) < 0.5\\ 1 - x_{id} & \rho_i N_i(0,1)\varphi(Z_{\mathrm{md},i}) \ge 0.5 \end{cases} \\ \varphi(Z_{\mathrm{md},i}) = \begin{cases} \mathrm{e}^{\frac{-(Z_{\mathrm{md},i} - E_x)^2}{2(E_a^i)^2}} & Z_{\mathrm{md},i} > E_x\\ p_c & Z_{\mathrm{md},i} \le E_x \end{cases} \end{cases}$$
(28)

式中: ρ_i 、 p_c 为控制参数, ρ_i 通常取 [0.625,5] 区间的 常数, p_c 通常取 [0.8,1] 区间的常数,可根据实验情况 具体调节; $\varphi(Z_{md,i})$ 的作用是保持相对优良粒子 $(Z_{md,i} > E_x)$,对于相对较差的粒子 $(Z_{md,i} \leq E_x)$ 进行 较大的变异。

同时,引入使用个体相似性 $S_{ij}^{(0)}$ 和种群多样性 $S_D^{(0)}$ 表征粒子间相似程度^[27],控制迭代过程中的变异 过程, $S_D^{(0)}$ 越小,说明第t代粒子中编码相似的粒子 越多,越容易陷入局部最优解。当迭代过程中粒子 的种群多样性低于阈值时,即 $S_D^{(0)} \leq S_{D0}$ 时,粒子执 行云自适应变异操作。具体如式(29)所示:

$$\begin{cases} S_{ij}^{(t)} = \frac{1}{a} \sum_{k=1}^{a} (x_{ik}^{(t)} \odot x_{jk}^{(t)}) \\ S_{D}^{(t)} = 1 - \frac{1}{a(a-1)/2} \sum_{i=1}^{a} \sum_{j=1}^{a} S_{ij}^{(t)} \end{cases}$$
(29)

式中: x⁽ⁱ⁾_{ik} ⊙ x⁽ⁱ⁾_{jk}为同或运算。即用第 t 代粒子 i 与粒子 j 之间相同的位数累加值来表示各粒子间的相似性, 用第 t 代粒子中粒子总相似度表征多样性。

4.4 基于 CPBPSO 算法的求解步骤

步骤1 初始化粒子群。根据来袭导弹威胁计 算各型雷达最小部署数量($N_{\rm R} = N_{\rm p} + N_{\rm x} + N_{\rm f}$),将每 部雷达部署涉及的参数进行编码处理,随机设置初 始值,得到初始群落,按照分区编码策略处理。其 中,第i个粒子二进制编码表示为

 $\begin{cases} X_{i} = (\cdots, x_{iP_{j,1}}, x_{iP_{j,2}}, \cdots, x_{iP_{j,n1}}, x_{iP_{j,n2}}, \cdots, x_{iX_{j,1}}, x_{iX_{j,2}}, \cdots, x_{iX_{j,n1}}, x_{iX_{j,n2}}, \cdots, x_{iF_{j,n1}}, x_{iF_{j,n2}}, \cdots, x_{iF_{j,n1}}, x_{iF_{j,n2}}, \cdots, x_{iF_{j,n1}}, x_{iF_{j,n2}}, \cdots) \\ V_{i} = (\cdots, v_{iP_{j,1}}, v_{iP_{j,2}}, \cdots, v_{iP_{j(N_{a}+N_{d,2})}}, \cdots, v_{iF_{j,1}}, v_{iF_{j,2}}, \cdots, v_{iF_{j(N_{a}+N_{d,2})}}, \cdots) \\ v_{iX_{j,2}}, \cdots, v_{iX_{j(N_{a}+N_{d,2})}}, \cdots, v_{iF_{j,1}}, v_{iF_{j,2}}, \cdots, v_{iF_{j(N_{a}+N_{d,2})}}, \cdots) \end{cases}$ (30)

式中: $P_j = 1, 2, \dots, N_p$, $X_j = 1, 2, \dots, N_x$, $F_j = 1, 2, \dots$, N_f , 每个点位的地理高程信息由该点所在的 DEM 数据确定。

步骤 2 适应度计算,更新全局极值。按照模型的目标函数和约束条件,计算每个粒子的适应度,更新个体极值 *p*_{best}。将每个粒子的适应度与粒子的最好位置比较,更新全局极值 *g*_{best}。

步骤 3 更新粒子状态。利用式(24)计算新的 粒子速度和位置,并更新粒子状态。

步骤 4 执行云自适应变异操作。判断本代粒 子的种群多样性是否低于阈值,即*S⁰_D* ≤ *S*_{D0}, 对多 样性低的本代粒子执行云自适应变异操作,并对超 出边界的数据进行边界吸收处理。

步骤 5 当算法达到停止条件时,停止搜索输 出结果,并对结果进行反编码处理,得到最优化雷 达部署方案;否则,返回到步骤 3 继续循环。

5 实例分析

针对上述分析和讨论,本文设计了导弹预警仿 真场景,用以验证部署模型的可行性。仿真场景搭 建和模型计算使用 STK11.0、MATLAB 2021 软件, 地理信息数据使用 ASTER GDEM V3 版本的高程 数据。

5.1 应对单方向威胁协同预警场景仿真

在单方向威胁协同预警场景中,模拟攻击方 从中国东南方向来袭。根据 3.2.1 节单方向协同预 警优化部署方法对上述场景进行求解。参数设置 如下:

1)作战场景想定为:攻击方在以 $E_1(27.4^{\circ}N, 133.1^{\circ}E)$ 、 $E_2(19.2^{\circ}N, 127.8^{\circ}E)$ 、 $E_3(27.2^{\circ}N, 129.0^{\circ}E)$ 、 $E_4(23.7^{\circ}N, 127.1^{\circ}E)$ 构成的威胁区域内部署导弹发 射平台,形成对我方保卫区域 $F(41.0^{\circ}N, 80.3^{\circ}E)$ 的威胁,如图 11 所示。随机选取威胁空域内的典型 高弹道 BM1、BM2 和低弹道 BM3、BM4 用以表示 威胁空域,其弹道数据由 STK 软件生成,如表 1 所 示。表中,弹道弧长为 $l_{missile}$,导弹飞行时间为 $t_{missile}$ 。 导弹在中段飞行速度不变,始终以关机点速度v飞 行。拦截系统部署在保卫要地F附近,假设拦截远 界 $L_{far}=300 \text{ km}、近界 L_{close}=50 \text{ km}$ 。

2) 雷达部署区域约束设置为:可部署区域选取 在中国境内,按照经纬度 0.1°的间隔,形成离散化



图 11 威胁空域示意

Fig. 11 Schematic diagram of threat airspace

表1 单方向威胁空域的典型弹道信息

 Table 1
 Typical ballistic information of single direction

 threat airspace
 1

弹道序号	发点 位置/(°)	落点 位置/(°)	弹道 弧长/km	关机点速度/ (km·s ⁻¹)	飞行 时间/s
1	(27.4,133.1)	(41.0,80.3)	7 036.7	6.534	1 934.7
2	(19.2,127.8)	(41.0,80.3)	7 072.5	6.498	1 928.3
3	(27.2,129.0)	(41.0,80.3)	5 140.5	5.865	1 186.9
4	(23.7,127.1)	(41.0,80.3)	5 258.1	5.904	1 203.4

网格, 阵面法向寻优间隔为 0.1°。为简化区域划分 复杂度, 可将不适合部署的点位 *S*_{nd} 对应的探测收 益设为 0。

3) 模型输入参数设置为: 三型雷达参照美军典 型预警雷达参数^[28]设置, 如表 2 所示。表中数据对 应第 2 节所述雷达威力范围, 包含最大探测距离r、 方位范围 A、俯仰范围 E等模型输入; 远程预警雷 达波束宽度 $\Delta \theta = \Delta \xi = 2^{\circ}$ (对应 3 dB 能量衰减), 跟 踪识别雷达波束宽度 $\Delta \theta = \Delta \xi = 0.3^{\circ}$, 单波束检测概 率 $p_d = 0.9$ 。模型中所有权重系数取均值, 如式(20) 中权重系数 $\omega_i = 0.2(i = 1, 2, \dots, 5)$, 在作战应用中可 根据实际进行调整。

表 2 雷达基本参数

 Table 2
 Basic parameter of radar

雷达类型	最大探测距离/km	方位范围/(°)	俯仰范围/(°)
远程预警雷达	3 000	±60	0~85
跟踪识别雷达	500	0~360	10~90
前置预警雷达	1 500	±53	10~85

4) CPBPSO 算法参数设置为:本文实验中,算法最大迭代次数T = 500,粒子个数 $N_s = 50$,粒子维数a = 81,其中,每型雷达经纬度编码 15 位、阵面法向编码 12 位。惯性因子采用线性递减策略,取 $w \in [0.4, 0.9]$,学习因子取 $c_1 = c_2 = 2$,种群多样性阈值取 $S_{D0} = 50\%$,云自适应控制参数 $\rho_i = 3$, $p_c = 0.9$ 。

经计算后,得到三型雷达最优部署和阵面法向 参数,如表3所示,部署方案如图12所示。

从仿真分析结果可见,在覆盖性方面,该部署

2024 年

方案基本实现对典型弹道的最优覆盖,弹道的协同 覆盖率分别为94.67%、83.05%、94.13%、92.07%(见 表4)。各雷达部署位置的地形遮蔽对探测性能影 响较小,较为明显的是跟踪识别雷达的等效探测距 离受到影响,形成了不规则的威力空间,但在来袭 方向上的遮蔽影响较小,如图13所示。

在协同性方面, 雷达间的交接均满足预测交接 需求。远程预警雷达部署位置更靠近发射点, 此时

	衣 5	田区儿化即省参奴
Table 3	Optim	al radar deployment parameter

雷达类型	部署点位	海拔/km	法向/(°)
前置预警雷达	(24.83°N 114.83°E)	0.029 3	353.5
远程预警雷达	(45.62°N 116.82°E)	0.897 9	148.6
跟踪识别雷达	(43.27°N 85.23°E)	3.7204	



图 12 单方向威胁优化部署方案

Fig. 12 Optimal radar deployment in single direction threat

表4 单方向威胁预警能力分析

 Table 4 Analysis of single direction threat early warning capability

弹道 名称	首点告警 时间/s	持续跟踪 时长/s	弹道覆盖 率/%	高识别 时长/s
BM1	T_0 +56.18	1 831.58	94.67	71.41
BM2	T_0 +39.21	1 601.45	83.05	56.10
BM3	T ₀ +50.36	1 117.22	94.13	113.88
BM4	T_0 +40.29	1 106.76	92.07	100.85

注:T₀为导弹发射时刻。



图 13 跟踪识别雷达地形遮蔽示意图 Fig. 13 Schematic of tracking and identifying radar terrain shadowing

更有利于提高远程预警雷达的预警时间。跟踪识 别雷达远离导弹落点的原因是部署位置距离落点 越近,探测收益越差。仿真结果和实际反导作战 中,对预警资源运用的认知基本一致,从而说明关 于三型预警雷达的协同部署模型和求解方法具有 较强的可信度。

5.2 应对多方向威胁协同预警场景仿真

在多方向威胁协同预警场景中,模拟攻击方从 中国各个方向来袭。仿真参数与5.1节参数相同, 增加威胁空域至6处,即假设攻击方在*E*₁~*E*₆等 6个区域部署导弹发射平台,构成对保卫区域*F*的 威胁,其典型弹道参数如表5所示。

在此场景下,经计算得到雷达最优部署、阵面 法向参数及所针对的作战方向,如表6所示。

表 5 多方向威胁空域的典型弹道信息

Table 5 Typical ballistic information of multi-direction

油道宮县	发点	落点	弹道	关机点速度/	飞行
産垣川・ラ	位置/(°)	位置/(°)	弧长/km	$(km \cdot s^{-1})$	时间/s
1	(56.3,123.5)	(41.0,80.3)	6 003.7	5.236	984.5
2	(37.0,127.0)	(41.0,80.3)	4 839.7	5.550	1 078.9
3	(23.7,130.0)	(41.0,80.3)	6 083.4	5.993	1 234.8
4	(13.7,105.0)	(41.0,80.3)	4 167.3	5.904	1 607.3
5	(7.5,91.5)	(41.0,80.3)	4 601.5	5.392	1 053.3
6	(1.6,60.0)	(41.0,80.3)	5 699.9	5.747	1 216.6

注:为便于表示,每个空域仅选填1条典型最优能量弹道。

表6 雷达部署及任务分配

Table 6 Radar deployment and task assignment

					_	_	_	_	_	_
雷达类型	名称	部署 点位/(°)	高程/ km	法向/ (°)	E_1	E_2	E_3	E_4	E_5	E_6
M. organitation	FBR1	(44.8,113.3)	1.278	7.2	\checkmark					
间置	FBR2	(25.7,119.1)	0.568	4.8		\checkmark	\checkmark			
田心	FBR3	(27.9,88.1)	4.758	88.4				\checkmark	\checkmark	
远程预警	PBR1	(29.6,109.4)	0.569	304.7	\checkmark	\checkmark	\checkmark			
雷达	PBR2	(38.6,90.2)	3.138	147.9				\checkmark	\checkmark	\checkmark
跟踪识别	XBR1	(42.4,81.4)	3.951		\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
雷达	XBR2	(41.7,77.6)	5.199		\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark

图 14 为面对多方向的导弹预警雷达协同部署 方案。其中,由于没有针对*E*₆区域的前置预警雷达 可部署区域,此方向未部署前置预警雷达;相对于 5.1 节单方向威胁场景部署方案,第 2 部前置预警 雷达部署位置更接近*E*₃区域,兼顾了*E*₂的探测;第 3 部前置预警雷达由于部署区域的约束限制,只能 前置部署至靠近威胁区域*E*₄、*E*₅的点位。2 部远程 预警雷达共同实现了对 6 个方向威胁的探测,其 中,第 1 部远程预警雷达的位置和方位相对于 5.1 节单方向威胁场景部署方案发生较大改变,以 满足对3个方向同时探测的最优部署需求。此外, 单部跟踪识别雷达即具备对6个方向探测覆盖,但 考虑到探测资源的饱和,故在次优解位置部署第 2部跟踪识别雷达。各雷达部署后,其地形遮蔽对 威胁方向的探测影响较小。通过仿真分析可得,该 部署方案满足对各威胁方向弹道的覆盖性和协同 性需求,实现对6个威胁方向的协同预警最优覆 盖,具体分析结果如表7所示,此处不再赘述。



图 14 多方向威胁优化部署方案

Fig. 14 Optimal radar deployment in multi-direction threat

表 7 多方向威胁预警能力分析

 Table 7 Analysis of mutli-direction threat early warning capability

威胁	首点告警	持续跟踪	弹道覆盖	高识别
区域	时间/s	时长/s	率/%	时长/s
E_1	$T_0 + 56.68$	905.28	92.12	76.11
E_2	T_0 +50.39	1 049.44	97.27	111.26
E_3	T_0 +42.07	1 022.90	82.84	93.425
E_4	T_0 +59.97	1 507.16	93.77	76.96
E_5	T_0 +51.88	980.93	93.13	77.16
E_6	T_0 +608.46	563.16	46.29	62.55

5.3 算法性能对比实验

设计对比实验对本文所提 CPBPSO 算法的寻 优性能进行分析。选取现有研究中同样针对雷达 部署问题而提出的改进算法,如 CGA^[2]、多目标粒 子 群 (multi-objective particle swarm optimization, MOPSO)算法^[29],形成对照实验。采用上述单方向、 多方向威胁 2 种典型场景进行计算,分别代表小规 模和大规模算例实验。算法对比结果如图 15 所 示。仿真结果表明,在小规模算例中,CGA 算法^[2] 跳出能力较差,易陷入局部最优解;MOPSO 算法^[29] 具有较快的收敛速度,但在收敛能力上不如 CPBPSO 算法。在大规模算例中,CPBPSO 算法在迭代中后 期也具有较好的局部跳出能力,相对于其他算法更 能得到最优结果。





6 结 论

 1)模型解决了导弹预警雷达协同部署问题,分 别建立远程预警雷达、跟踪识别雷达和机动式预警 雷达最优部署模型,实现单、多方向威胁条件下的 最优探测部署方案求解。

2) CPBPSO 算法所采用的编码策略设计、云自 适应变异等改进,实现了分区编码计算和粒子变异 更新,相比于现有算法,在解决复杂模型中收敛能 力较为优异。

3) 在本文背景下, 对于单方向威胁典型弹道覆 盖率分别为 94.67%、83.05%、94.13%、92.07%; 对于 来自 6 个方向的弹道覆盖率分别为 92.12%、97.27%、 82.84%、93.77%、93.13%、46.29%, 满足所提的覆盖 性和协同性需求。

为使本文提出的协同部署模型更加精细,仍需 进一步考虑雷达精度、目标容量等指标的模型表征; 同时,对于未来天基预警卫星、地基预警雷达、海 基预警雷达等预警设备协同运用为主的体系布局, 需开展多源异构预警资源的协同部署问题研究。

参考文献(References)

[1] CHRISTIAN A. US missile defence efforts and Chinese reserva-

tion in East Asia[J]. Asian Affairs, 2020(2): 605-620.

- [2] WEI P, ZHENG L, HUANG F, et al. Optimal deployment of radar network based on chaos genetic algorithm[C]//Proceedings of the Chinese Control and Decision Conference. Piscataway: IEEE Press, 2019: 6018- 6023.
- [3] LI J, DAI W, HU L. Research on intelligent station layout optimization of air defense radar network[C]//Proceedings of the 4th International Conference on Robot Systems and Applications. Piscataway: IEEE Press, 2021: 39-43.
- [4] YANG Y, ZHANG T, YI W. Deployment of multistatic radar system using multi objective particle swarm optimization[J]. IET Radar, Sonar & Navigation, 2018, 12(5): 485-493.
- [5] Ma Y, JIN H B, LI H, et al. Adaptive opposition-based particle swarm optimization algorithm and application research[C]//Proceedings of the IEEE 4th International Conference on Signal and Image Processing. Piscataway: IEEE Press, 2019: 518-523.
- [6] RAO D V, RAVISHANKAR M. A methodology for optimal deployment and effectiveness evaluation of air defence resources using game theory[J]. Sādhanā, 2020, 45(1): 1-15.
- [7] YU Z, SHAN G, XU G, et al. Method of multi-sensor optimal deployment for area coverage[C]//Proceedings of the International Conference on Electronics Technology. Piscataway: IEEE Press, 2018: 116-119.
- [8] 付鑫, 张峰, 冯占林. 基于并行计算的混沌遗传算法对反导预警 雷达部署优化研究[J]. 中国电子科学研究院学报, 2016, 11(3): 276-282.

FU X, ZHANG F, FENG Z L. Research of chaos genetic algorithm based on parallel computing for anti-missile warning radar disposition[J]. Journal of CAEIT, 2016, 11(3): 276-282(in Chinese).

- [9] TIAN M, CHEN S, ZHENG X, et al. Sensors deployment optimization in multi-dimensional space based on improved particle swarm optimization algorithm[J]. Journal of Systems Engineering and Electronics, 2018, 29(5): 969-982.
- [10] 田桂林,刘昌云,高嘉乐,等. 基于烟花算法的多传感器优化模型 部署[J]. 系统工程与电子技术, 2019, 41(8): 1742-1748.
 TIAN G L, LIU C Y, GAO J L, et al. Multi-sensor optimal disposition model based on fireworks algorithm[J]. Systems Engineering and Electronics, 2019, 41(8): 1742-1748(in Chinese).
- [11] 刘伟,刘昌云,陈晨,等. 反导预警装备部署方法综述[J]. 探测与 控制学报, 2022, 44(3): 27-33.
 LIU W, LIU C Y, CHEN C, et al. Review on deployment methods of anti-missile early waring equipment[J]. Journal of Detection & Control, 2022, 44(3): 27-33(in Chinese).
- [12] KORDA M, KRISTENSEN H M. US ballistic missile defenses[J]. Bulletin of the Atomic Scientists, 2019, 75(6): 295-306.
- [13] ZHENG Y, YAN S, QU S, et al. Research on design and evaluation method of anti-missile early warning and detection plan[C]//Proceedings of the IEEE 5th Information Technology, Networking, Electronic and Automation Control Conference. Piscataway: IEEE Press, 2021: 864-867.
- [14] FLORES R, ALMEIDA G F, TEJERINA G R, et al. Radar coverage over irregular terrain: A practical algorithm for multipath propagation[C]//Proceedings of the IEEE Radar Conference. Piscataway: IEEE Press, 2018: 1383-1388.

- [15] 徐志明, 艾小锋, 刘晓斌, 等. 基于散射中心滑动特性的双基地雷达 锥体目标微动特征提取方法[J]. 电子学报, 2021, 49(3): 461-469. XU Z M, AI X F, LIU X B, et al. Micro-motion feature extraction of bistatic radar cone-shaped targets based on characteristic analysis of sliding scattering centers[J]. Acta Electonica Sinica, 2021, 49(3): 461-469(in Chinese).
- [16] 马梁. 弹道中段目标微动特性及综合识别方法[D]. 长沙: 国防科 学技术大学, 2011: 49-68.

MA L. The micro-motion characteristic and combining classification of ballistic target[D]. Changsha: National University of Defense Technology, 2011: 49-68(in Chinese).

- [17] KIM J, CHO D H, LEE W C, et al. Optimal target assignment with seamless handovers for networked radars[J]. Sensors, 2019, 19(20): 4555.
- [18] 郑玉军, 田康生, 刘俊凯, 等. 早期预警雷达和多功能相控阵雷达的目标指示交接方法[J]. 兵工学报, 2017, 38(1): 106-113. ZHENG Y J, TIAN K S, LIU J K, et al. The target designation and handover method of early-warning radar and multifunction phased array radar[J]. Acta Armamentarii, 2017, 38(1): 106-113(in Chinese).
- [19] FREITAS D, LOPES L G, MORGADO F. Particle swarm optimization: A historical review up to the current developments[J]. Entropy, 2020, 22(3): 362.
- [20] WANG D S, TAN D P, LIU L. Particle swarm optimization algorithm: An overview[J]. Soft Computing, 2018, 22(2): 387-408.
- [21] GUO G Y, LV Y S, YANG X R, et al. Ballistic missile defense system deployment simulation based particle swarm optimization [C]//Proceedings of the IEEE CSAA Guidance, Navigation and Control Conference. Piscataway: IEEE Press, 2018: 1-4.
- [22] 魏法,杨明磊,何小静,等.基于改进粒子群算法的平面阵同时多 波束赋形方法[J].系统工程与电子技术,2022,44(6):1789-1797.
 WEI F, YANG M L, HE X J, et al. Simultaneous multi-beam forming method for planar array based on improved particle swarm algorithm[J]. Systems Engineering and Electronics, 2022, 44(6): 1789-1797(in Chinese).
- [23] LIU S, HUANG F, YAN B. Optimal design of multi-missile formation based on an adaptive SA-PSO algorithm[J]. Aerospace, 2022, 9(1): 21.
- [24] 练青坡, 王宏健, 袁建亚, 等. 基于粒子群优化算法的 USV 集群 协同避碰方法[J]. 系统工程与电子技术, 2019, 41(9): 2034-2040. LIAN Q P, WANG H J, YUAN J Y, et al. Collaborative collision avoidance method of USV cluster based on particle swarm optimization algorithm[J]. Systems Engineering and Electronics, 2019, 41(9): 2034-2040(in Chinese).
- [25] 黄飞腾, 翁国庆, 南佘荣, 等. 基于改进云自适应粒子群的多 DG 配电网 EV 充电站优化配置[J]. 中国电机工程学报, 2018, 38(2): 514-525.

HUANG F T, WENG G Q, NAN Y R, et al. Optimization of electric vehicle charging stations based on improved cloud adaptive particle swarm in distribution network with multiple DG[J]. Proceedings of the CSEE, 2018, 38(2): 514-525(in Chinese).

- [26] XIANG L I, CHEN J. A modified PSO algorithm based on cloud theory for optimizing the fuzzy PID controller[J]. Journal of Physics: Conference Series, 2022, 2183(1): 012014.
- [27] 陈国龙, 郭文忠, 陈羽中. 无线传感器网络任务分配动态联盟模

型与算法研究[J]. 通信学报, 2009, 30(11): 48-55. CHEN G L, GUO W Z, CHEN Y Z. Research on dynamic alliance of task allocation and its algorithm in wireless sensor network[J]. Journal on Communications, 2009, 30(11): 48-55(in Chinese).

[28] 韩晋山, 邢建平, 张浩, 等. 美国导弹预警系统的发展现状与趋势 分析[J]. 科技导报, 2019, 37(4): 91-95.

HAN J S, XING J P, ZHANG H, et al. Development of American

missile early warning system[J]. Science Technology Review, 2019, 37(4): 91-95(in Chinese).

[29] 唐嘉诚. 多功能雷达组网资源调度方法研究[D]. 成都: 电子科技 大学, 2021: 33-51.

TANG J C. Research on resource scheduling method of multifunctional radar network[D]. Chengdu: University of Electronic Science and Technology of China, 2021: 33-51(in Chinese).

Deployment optimization method for missile early warning radar under complex and multi-directional missile threats

LIU Wei^{1, 2}, LIU Changyun³, GUO Xiangke^{3, *}, FAN Liangyou¹, HE Sheng¹, LAN Hao²

(1. Graduate School, Air Force Engineering University, Xi' an 710000, China;

2. Unit 63768 of the PLA, Xi'an 710000, China;

3. Air and Missile Defense College, Air Force Engineering University, Xi'an 710000, China)

Abstract: An optimized deployment model of multiple missile early warning radars is established based on the distinct mission characteristics of early warning radar, tracking and identification radar, and transportable early warning radar in order to address the situation where the deployment of existing missile early warning radars is relatively independent, difficult to cooperate with, and difficult to meet the large-scale operation scenario. Under the constraints of optimal coverage, cooperative handover, and target identification, the cooperative deployment of early warning radars is solved. A cloud adaptive partition optimization binary particle swarm optimization (CPBPSO) algorithm is designed. In order to make the algorithm more suited for solving early warning radar deployment problems, the partition coding strategy is used to shrink the algorithm's solution space. The cloud adaptive mutation operator is then added to enhance the algorithm's global optimization and local jumping capability. The example verifies the feasibility of the model in solving the deployment problem of single-direction and multi-direction threat scenarios and analyzes the effectiveness of the CPBPSO algorithm, which basically meets the needs of the optimal cooperative deployment of missile early warning radar.

Keywords: missile early warning radar; collaborative early warning; deployment optimization model; cloud adaptive mutation; particle swarm optimization algorithm

Received: 2022-06-12; Accepted: 2022-08-10; Published Online: 2022-09-08 13:32 URL: link.cnki.net/urlid/11.2625.V.20220907.1131.001

Foundation item: National Natural Science Foundation of China (62106283)

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn **DOI**: 10.13700/j.bh.1001-5965.2022.0490

含不同直径微脉管层合板拉伸性能研究

冉广陵^{1,2},马埸浩^{1,3},安子乾¹,赵大方⁴,郭鑫¹,程小全^{1,2,*}

(1. 北京航空航天大学 航空科学与工程学院,北京 100191; 2. 北京航空航天大学 宁波创新研究院,宁波 315800;

3. 中国兵器工业导航与控制技术研究所,北京100089; 4. 中航复合材料有限责任公司,北京101300)

摘 要:针对微脉管对层合板力学性能的影响,对含不同直径微脉管层合板的拉伸性能进行研究。通过试验测量了含不同直径微脉管层合板的拉伸性能,得到拉伸强度与刚度。建立了含微脉管层合板的精细化有限元模型,综合考虑了微脉管引起的富树脂区、纤维局部弯曲和铺层局部纤维体积分数变化的影响,对含微脉管层合板的拉伸性能进行了分析,并得到了试验结果的验证。在此基础上,分别就微脉管方向和直径对层合板拉伸性能的影响进行了研究。结果表明:当微脉管垂直铺层方向时,使单向层合板纵向拉伸强度降低,但刚度基本不受影响;微脉管直径越大,拉伸强度降低幅度越大;与不含微脉管层合板相比,含直径为0.255 mm和0.4 mm 微脉管的层合板拉伸强度分别下降了21.9%和39.9%,但刚度变化不超过1.4%;当微脉管直径变化时,层合板拉伸损伤与破坏过程也发生变化。

关键 词:复合材料层合板;自修复;微脉管;拉伸性能;有限元

中图分类号: V215.2⁺2; TB332

文献标志码: A 文章编号: 1001-5965(2024)04-1405-11

复合材料由于比强度高、比模量大等优点在航空航天等领域得到广泛应用。然而,由于复合材料 层合板层间性能较低,受冲击时结构内部容易出现 基体开裂、层间分层,甚至纤维断裂等损伤,在结构 设计值确定时,通常须要对复合材料极限强度值进 行大幅折减,严重限制了复合材料优势的发挥。传 统的机械修理和胶接修理须要飞行器暂停工作,不 能实现损伤的实时修复,降低了结构的使用效率。 受生物体损伤自愈合机制的启发,人们提出了自修 复复合材料概念,并发展出多种自修复体系^[1-4],其 中,微脉管型自修复体系因能实现损伤的多次修复 受到研究人员的关注。

微脉管被埋入复合材料铺层之间时,由于微脉 管周围铺层的纤维连续且具有一定的弯曲刚度,在 微脉管边缘某些位置会形成富树脂区、纤维局部弯 曲及微脉管周围的铺层厚度变化。微脉管的引入 事实上是在层合板内引入了初始缺陷,会对结构造 成不利影响。如果这种影响过大,则影响结构的承 载能力,使其修理功能失去工程价值。因此,对于 微脉管型自修复体系,应充分研究微脉管对层合板 力学性能的影响,确保不会引起层合板刚度和强度 较大的下降。

在试验研究方面, Kousourakis 等^[5-7]研究了位 于铺层中面的不同形状(圆形和椭圆形)和截面尺 寸微脉管对 [0/90]₁₆铺层层合板的纵向(载荷方向 平行微脉管)和横向(载荷方向垂直微脉管)力学性 能、层合板层间力学性能及T形接头拉脱性能等的 影响,发现微脉管对层合板横向力学性能的影响较 纵向更为显著,随微脉管直径增大,层合板的弹性 模量和强度均逐渐降低,微脉管会钝化裂纹尖端或 改变扩展路径,使层合板层间 I 型断裂韧性提高, 此外,微脉管对T形接头的拉脱强度没有影响,但 会提高其失效位移和吸收的能量。Huang 等^[8-9]采 用试验和有限元方法研究了不同直径圆截面微脉

收稿日期: 2022-06-16; 录用日期: 2022-08-21; 网络出版时间: 2022-08-29 16:42

引用格式: 冉广陵, 马場浩, 安子乾, 等. 含不同直径徽脉管层合板拉伸性能研究 [J]. 北京航空航天大学学报, 2024, 50 (4): 1405-1415. RANGL, MAYH, ANZQ, et al. Study of tensile properties of laminates containing microvascular channels with different diameters [J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50 (4): 1405-1415 (in Chinese).

网络出版地址: link.cnki.net/urlid/11.2625.V.20220829.1421.002

^{*}通信作者. E-mail: xiaoquan_cheng@buaa.edu.cn

管对层合板压缩性能的影响及其损伤扩展过程,发现试验件先在微脉管左右两侧沿载荷方向出现裂纹,再在富树脂区内扩展,最终微脉管两侧铺层由于微屈曲而发生压缩破坏。Luterbacher等^[10]研究了微脉管对层合板静力和拉伸疲劳性能的影响,发现当微脉管面积占层合板截面积低于 0.5% 时,对层合板力学性能和损伤过程影响很小。Trask和Bond^[11]对含微脉管层合板的冲击性能进行了研究,发现其剩余压缩强度不超过完好层合板的 70%。

在数值分析方面, Nguyen 和 Orifici^[12] 通过有限 元方法系统研究了不同尺寸的圆形/椭圆形 2 种截 面形状微脉管对层合板横向和纵向刚度和强度的 影响。Al-Shawk 等^[13-14] 建立了含微脉管玻璃纤维 复合材料层合板拉伸载荷作用下的有限元模型, 研 究了微脉管形状、截面尺寸及层合板铺层顺序对富 树脂区应力分布和层合板失效行为的影响, 发现与 截面积相同的圆形微脉管相比, 椭圆形微脉管对层 合板力学性能的影响更小。

目前,国内外学者主要通过试验方法研究微脉 管对复合材料层合板力学性能的影响,仅有少数学 者通过数值方法对其进行研究。在数值研究中,一 般会考虑微脉管引起的富树脂区和纤维局部弯曲 的作用,但对微脉管周围铺层厚度和纤维体积分数 变化的影响未加以考虑,而其对复合材料的强度会 有直接的影响,因此不能忽视。

本文对含不同直径圆形微脉管层合板的拉伸 性能进行试验研究,在此基础上,提出一种综合考 虑富树脂区、纤维局部弯曲和铺层局部纤维体积分 数变化对含微脉管层合板精细化有限元模型,对含 垂直铺层方向微脉管的层合板的拉伸性能进行分 析,并总结了微脉管方向和直径的影响,为下一步 层合板中微脉管的设计提供基础。

1 试 验

1.1 试验件

试验件尺寸参照 ASTM D3039^[15] 进行设计,共4组,试验件参数及数量安排如表1所示。所有微脉管均铺设在层合板中部,平行于试验件宽度方向,沿长度方向间隔为10 mm,如图1所示。

表1 试验件参数

Table 1 Parameters of specimen

组别	铺层	尺寸(长×宽×厚)/	微脉管	微脉管	微脉管	数量
-11/44	110/24	(mm×mm×mm)	直径d/mm	间距/mm	层数	M±
1	[0 ₁₆]	250×15×2			0	6
2	$[0_8/(r)/0_8]$	250×15×2	0.255	10	1	5
3	$[0_8/(r)/0_8]$	250×15×2	0.4	10	1	5
4	[90 ₈ /(r)/90 ₈]	175×25×2	0.255	10	1	5



Fig. 1 Schematic diagram of microvascular channels' position in specimen

试验件材料为 T700/LT03A 单向带, 热压罐成型, 单层名义厚度为 0.125 mm。对于含有微脉管的试验件, 先按照设计的铺层和微脉管直径及位置, 在相应铺层之间预埋相应直径的钢丝, 加工成复合材料大板, 待复合材料大板固化完成后, 将钢丝抽出形成微脉管, 再根据设计要求将大板切割成试验件。

1.2 试验过程

拉伸试验参照 ASTM D3039^[15] 进行,在 Instron 880-100 kN 液压伺服试验机上进行。试验前,在试 验件两面各粘贴一个纵向应变片,其中,1#应变片 位于微脉管位置,2#应变片位于不含微脉管位置。 在第 2、3 组试验件中,一个面上沿纵向安装一个引 伸计(标距 25 mm),用来测量试验件的拉伸模量。 加载速度 2 mm/min。试验件试验状态如图 2 所示。



图 2 试验件试验状态 Fig. 2 Test condition of specimen

1.3 试验结果

试验测得的载荷-位移曲线如图 3 所示,根据载荷与不同部位应变测量数据计算得到的强度和局部刚度结果如表 2 所示。由第 2、3 组试验件结果可见,当微脉管布置垂直于受载方向时,微脉管的引入对单向层合板整体纵向刚度影响较小,不超过1.4%,但由不同部位测得的应变与应力数据计算得到的局部刚度却有较大差异。以第 2 组试验件为例,无微脉管位置应变片测得的试验件拉伸刚度和引伸计测得的刚度一致,约为 139 GPa;而由微脉管位置应变片测得的拉伸刚度偏高,约为 147 GPa。



图 3 试验测得的载荷-位移曲线

Fig. 3 Load-displacement curves measured in test

衣 4 拉仲氏铋纪未刈り	表 2	申试验结果对比
--------------	-----	---------

				1					
组别	试验件	微脉管 直径/	1#应变片所测 刚度(微脉	1#应变片所测 刚度平均值	1#应变片所测 刚度离散度	2#应变片所测 刚度(无微脉	2#应变片 刚度平:	片所测 2: 均值	#应变片所测 刚度离散度
	姍丂	mm	管处)/GPa	(微脉管处)/GPa	(微脉管处)/%	管处)/GPa	(无微脉管	处)/GPa (劧	毛微脉管处)/%
1	无管								
	1#	0.255	146.6			138.4			
2	2#	0.255	147.8	147.2	0.4	140.6	139.	7	0.8
	3#	0.255	147.1			140.1			
	1#	0.4	151.8			139.7			
3	2#	0.4	152.1	151.8	0.2	139.6	139.	6	0.1
	3#	0.4	151.6			139.5			
	1#	0.255	7.75			8.02			
	2#	0.255	7.68			7.90			
4	3#	0.255	7.75	7.73	0.37	8.12	8.	00	1.00
	4#	0.255	7.73			8.00			
	5#	0.255	7.73			7.97			
组别	试验件 编号	引伸计所测 刚度/GPa	引伸计所测刚度 平均值/GPa	引伸计所测刚度 平均下降比例/%	引伸计所测刚度 离散度/%	拉伸强度/ MPa	拉伸强度 平均值/MPa	拉伸强度平均 下降比例/%	勾 拉伸强度 。 离散度/%
1	无管		140.3		1.5		2 372.0		2.65
	1#	138.9				1 902.27			
2	2#	140.5	139.3	0.7	0.7	1 793.35	1 852.09	21.9	3.0
	3#	138.6				1 860.65			
	1#	135.0				1 454.54			
3	2#	140.5	138.3	1.4	2.1	1 391.91	1 424.59	39.9	2.2
	3#	139.5				1 427.32			
	1#					48.08			
	2#					47.99			
4	3#					45.52	48.05		3.53
	4#					48.40			
	5#					50.28			

Table 2 Comparison of tensile test results

第3组试验件也显示出类似的试验结果。因此,在 进行含微脉管试验件应变测量时,应注意测量方式 和测量点的选择。

表2显示,微脉管对单向层合板的纵向拉伸强 度影响较大,且微脉管直径越大,拉伸强度下降越 显著。与不含微脉管层合板的纵向拉伸强度相比, 含直径为 0.255 mm 和 0.4 mm 的微脉管分别使层合 板拉伸强度下降了 21.9%和 39.9%。2 组试验件均 在达到最大载荷后瞬间发生断裂,试验过程中未能 观测到损伤起始和扩展的过程。试验件最终的破 坏模式如图 4 所示,在多处断裂并伴有分层,主要 断裂面位于微脉管处。当微脉管布置垂直于受载


图 4 第 2、3 组试验件典型破坏模式



方向时,含微脉管单向层合板的横向拉伸强度为 48.05 MPa. 与 T700/LT03A 单向预浸带的横向拉伸 强度 55 MPa^[16]相比下降了 12.6%,但含微脉管层合 板的模量与不含微脉管层合板一致,均为8GPa;试 验件最终破坏模式均为微脉管所在截面断裂。

综上可知,微脉管对单向层合板的拉伸强度有 较大影响,对其刚度影响很小,但会改变层合板的 局部刚度。

有限元建模 2

2.1 单向预浸带力学性能和纤维体积分数的关系

对于单向预浸带等效模量的预测,选取 Chamis 模型[17],预测公式如下:

$$E_{11} = V^{\rm f} E_{11}^{\rm f} + V^{\rm m} E^{\rm m} \tag{1}$$

$$\mu_{12} = V^{f} \mu_{12}^{f} + V^{m} \mu^{m} \tag{2}$$

$$E_{22} = \frac{E_{22}^{\rm f} E^{\rm m}}{E_{22}^{\rm f} - \sqrt{V^{\rm f}} \left(E_{22}^{\rm f} - E^{\rm m} \right)} \tag{3}$$

$$G_{12} = \frac{G_{12}^{\rm f} G^{\rm m}}{G_{12}^{\rm f} - \sqrt{V^{\rm f}} \left(G_{12}^{\rm f} - G^{\rm m}\right)} \tag{4}$$

$$G_{23} = \frac{G_{23}^{\rm f}G^{\rm m}}{G_{23}^{\rm f} - \sqrt{V^{\rm f}} \left(G_{23}^{\rm f} - G^{\rm m}\right)} \tag{5}$$

式中:上标 f、m 分别表示纤维、树脂基体; V^f和V^m 分别为纤维和树脂的体积分数,且 $V^{f} + V^{m} = 1; E^{m}$ 、 G^{m} 和 μ^{m} 分别为树脂的弹性模量、剪切模量和泊松 比; Ef1和Ef2分别为纤维沿1方向和2方向的弹性 模量(1为沿纤维方向,2为垂直纤维方向); $\mu_{1,2}^{f}$ 和 G₁₂、G₂₃分别为纤维在相应平面内的泊松比和剪切 模量; E11、E22、G12和G23分别为单向预浸带对应方 向的模量;μ12为单向预浸带在对应方向的泊松比。

试验件所用 T700/LT03A 单向预浸带的纤维体 积分数约为0.6, 根据表3 所示参数^[16,18], 利用 Chamis 模型计算的单向预浸带性能参数和实际测量结果 对比如表4所示。可以看出, Chamis 模型计算值和 实际测量值非常接近。为更好地反映微脉管上下 铺层纤维体积分数变化对局部应力、应变分布的影 响,有限元中采用由 Chamis 模型计算得到的单向 预浸带弹性参数进行分析。

表 3 T700 纤维及 LT03A 树脂力学性能[16,18] Table 3 Mechanical properties of T700 fiber and LT03A resin^[16,18]

$E_{22}^{\mathrm{f}}E^{\mathrm{m}}$	(2)	$E_{11}^{\mathrm{f}}/\mathrm{GPa}$	$E_{22}^{\mathrm{f}}/\mathrm{GPa}$	G ^f ₁₂ /GPa	$G_{23}^{\mathrm{f}}/\mathrm{GPa}$	$\mu_{12}^{\rm f}$	μ^{f}_{23}	E ^m /GPa	G ^m /GPa	μ^{m}
$\overline{V^{\mathrm{f}}\left(E^{\mathrm{f}}_{22}-E^{\mathrm{m}} ight)}$	(3)	230.0	15.0	24.0	5.0	0.3	0.49	3.45	1.28	0.35

表 4 单向预浸带性能 Chamis 模型计算值和实际测量值对比

Table 4 Comparison of Chamis model calculation results and actual measured performance of unidirectional prepreg tape

测试方法	E_{11}/GPa	<i>E</i> ₁₁ 误差/%	E_{22}/GPa	E ₂₂ 误差/%	μ_{12}	μ_{12} 误差/%	G_{12} /GPa	G ₁₂ 误差/%	G_{23} /GPa	G ₂₃ 误差/%
Chamis模型计算值	139	0.02	8.55	(00	0.32	7 70	3.8	5.04	3.02	2.59
实际测量值	140.3	0.93	8.00	6.88	0.347	1.78	4.04	5.94	3.1	2.58

对于单向预浸带纵向拉伸强度 $X_{\rm T}$ 的预测,假设 直到单向预浸带破坏树脂仍表现为线弹性,则X_T为

$$X_{\rm T} = \beta \left(V^{\rm f} \sigma_{\rm f}^* + V^{\rm m} \frac{\sigma_{\rm f}^*}{E_{11}^{\rm f}} E^{\rm m} \right) \tag{6}$$

式中: σ_{f}^{*} 为纤维的拉伸强度; β 为强度折减系数。若 令β=1,由式(6)计算的单向预浸带纵向拉伸强度为 3211.8 MPa, 而实际测得的拉伸强度仅为2372 MPa, 因此, 取β=0.7385 以表征各种因素导致的单向预浸 带实际纵向拉伸强度和理论值的偏差。

根据式(1)~式(6),假设单向预浸带的纵向压 缩强度 $X_{\rm C}$ 、面内剪切强度 S_{12} 、 S_{23} 按纤维体积分数 等比例提高,保持纤维破坏时的压缩应变不变。而 对于单向预浸带的横向拉压强度Y_T和Y_c,由于其主 要由基体性能和纤维/基体界面性能决定,可认为其 不随纤维体积分数变化。

2.2 损伤模型

2.2.1 铺层损伤模型

表 5 为本文采用的二维 Hashin 复合材料铺层 失效准则^[19]。不考虑面外应力σ33、τ13和τ23对铺层

表 5 二维 Hashin 失效准则^[19]

失效模式	失效判据
纤维拉伸失效	$(\sigma_{11})^2 (\tau_{12})^2 > 1$
$(\sigma_{11} \ge 0)$	$\left(\frac{\overline{X_{T}}}{\overline{X_{T}}}\right) + \left(\frac{\overline{S_{12}}}{\overline{S_{12}}}\right) \ge 1$
纤维压缩失效	$\left(\sigma_{11}\right)^2 > 1$
$(\sigma_{11} < 0)$	$\left(\frac{1}{X_{\rm C}}\right) \neq 1$
基体拉伸失效	$\left(\sigma_{22}\right)^2$, $\left(\tau_{12}\right)^2$ 1
$(\sigma_{22} \ge 0)$	$\left(\frac{\overline{Y}_{T}}{\overline{Y}_{T}}\right) + \left(\frac{\overline{S}_{12}}{\overline{S}_{12}}\right) \ge 1$
基体压缩失效	$\left(\sigma_{22}\right)^2$, $\left[\left(Y_{\rm C}\right)^2$, $\left[\sigma_{22}, \left(\tau_{12}\right)^2\right]$, $\left[\left(Y_{\rm C}\right)^2\right]$
$(\sigma_{22} < 0)$	$\left(\frac{1}{2S_{13}}\right)^{-1} \left[\left(\frac{1}{2S_{23}}\right)^{-1}\right] \frac{1}{Y_{\rm C}} \left(\frac{1}{S_{12}}\right)^{-1} \right] \stackrel{\text{def}}{\rightarrow} 1$

1409

失效的影响, σ₁₁、σ₂₂和τ₁₂为对应方向的面内应 力。当材料满足上述失效起始判据后, 按表 6 所给 刚度退化系数进行衰减。

为考虑由面外应力导致的分层损伤,在层间加入 Cohesive 单元,选用二次应力准则作为界面损伤 起始判据,选取 B-K 准则模拟损伤扩展。

Table 6	Degradation	factor	of	material	stiffness	for	different
		failm	re	modes			

失效模式	退化系数
纤维拉伸失效	$E_{11} = 0.07 E_{11}, G_{12} = 0.07 G_{12}, \mu_{12} = 0.07 \mu_{12}$
纤维压缩失效	$E_{11}=0.14E_{11}, G_{12}=0.14G_{12}, \mu_{12}=0.14\mu_{12}$
其体拉伸生效	$E_{22} = 0.2E_{22}, G_{12} = 0.2G_{12}, G_{23} = 0.2G_{23},$
举冲 "	$\mu_{12} = 0.2\mu_{12}, \mu_{23} = 0.2\mu_{23}$
其休压缩生效	$E_{22} = 0.4E_{22}, G_{12} = 0.4G_{12}, G_{23} = 0.4G_{23},$
坐评压油八从	$\mu_{12} = 0.4 \mu_{12}, \mu_{23} = 0.4 \mu_{23}$

2.2.2 树脂损伤模型

对于富树脂区的树脂,采用最大切应力准则判 断树脂的损伤起始,如下:

 $\tau_{\max}/S \ge 1 \tag{7}$

式中: τ_{max} 为单元中的最大切应力; *S* 为树脂的剪切 强度。树脂满足损伤起始判据后, 按 $E^{m} = 0.4 E^{m}$ 、 $\mu^{m} = 0.4 \mu^{m}$ 进行衰减。

以上各判据均通过编写 ABAQUS 用户自定义 场变量 USDFLD 子程序实现。

2.3 有限元模型的建立

对于第2组含 0.255 mm 直径微脉管的层合板, 随机选取一个微脉管处的显微图像(见图 5),根据 试验件总厚度 2 mm 和各部分尺寸的相对比例关 系,测得各部分实际尺寸(见表 7),并由此建立含微 脉管层合板有限元模型,如图 6 所示。因层合板关 于微脉管中心对称,本文取 1/4 结构进行分析。在 x 方向建模长度为 2.5 mm,约为图 5 中富树脂区长 度的 2 倍,以包含受载过程中富树脂区尖端的不均



图 5 微脉管周围富树脂区长度和各铺层厚度测量 Fig. 5 Measurement of length of resin-rich zone around microvascular channel and thickness of each layer





匀应力场。对于 0°和 90°单向层合板沿 x 方向单向 受拉时, 层合板 y 方向的应力及影响很小, 因此, y 方向建模长度取为 0.2 mm, 以减少模型单元数 量, 提高计算效率。

编写 Python 脚本读取富树脂区上方铺层每个 单元的节点坐标, 计算单元的平均厚度, 并与标准 铺层厚度 0.125 mm 对比, 利用 Chamis 模型, 结合表 3 参数, 计算当前单元的纤维体积分数、弹性模量和 强度参数, 为单元赋予相应的材料属性。图 7 为图 6 中富树脂区上方铺层的纤维体积分数分布。图中 圆圈为微脉管位置示意。可以看出, 微脉管上方区 域的纤维体积分数普遍大于 0.6, 最靠近微脉管顶 点位置达到 0.816。赋予材料属性并细化网格后的 模型如图 8 所示。

本文所有有限元模型均基于商用有限元分析 软件 ABAQUS 建立,各铺层及富树脂区均选用三 维八节点线性缩减积分单元 (C3D8R)。在各铺层之 间、铺层1和富树脂区之间及铺层1和对称面之间



图 7 富树脂区上方铺层纤维体积分数分布 Fig. 7 Fiber volume fraction distribution in layers above resin-rich zone

加入 0 厚度的 Cohesive 单元(见图 6)。模型中的材料 参数如表 8 和表 9 所示。表中: $t_i^0(i表示n,s,t)$ 分别为 Cohesive 单元法向、纵向、横向的强度参数, $G_{C,i}(i表示$ n,s,t)为对应方向的断裂能释效率, $K_{ii}^0(i表示n,s,t)$ 为 对应方向的刚度参数, η 为材料参数。





Fig. 8 Refined mesh model after assigning material according to fiber volume fraction

$\chi_0 = 1/00/1100A + -1/2/2 + 1/4 Concave + /1/2$	JA 単回
---	-------

 Table 8
 T700/LT03A unidirectional prepreg tape and Cohesive unit parameters

	参数	数值
	E ₁₁ /GPa	139
T700/LT03A 单向预浸带	E_{22}/GPa	8.55
	G_{12}/GPa	3.8
	G_{23}/GPa	3.02
	μ_{12}	0.32
	μ_{23}	0.3
	$X_{\rm T}/{\rm MPa}$	2 3 7 2
	$X_{\rm C}/{\rm MPa}$	1 2 3 4
	$Y_{\rm T}/{\rm MPa}$	55.0
	$Y_{\rm C}/{\rm MPa}$	178
	<i>S</i> ₁₂ /MPa	125.0
	<i>S</i> ₂₃ /MPa	107
	$K_{\rm nn}^0/({\rm MPa}\cdot{\rm mm}^{-1})$	10 ⁵
	$K_{\rm ss}^0/({\rm MPa}\cdot{\rm mm}^{-1})$	10 ⁵
	$K_{\rm tt}^0/({\rm MPa}\cdot{\rm mm}^{-1})$	10 ⁵
	t_n^0/MPa	55
数值Cohesive	$t_{\rm s}^0/{ m MPa}$	80
单元[16]	t_t^0/MPa	80
	$G_{\mathrm{C,n}}/(\mathrm{N}\cdot\mathrm{mm}^{-1})$	0.276
	$G_{\mathrm{C},\mathrm{s}}/(\mathrm{N}\cdot\mathrm{mm}^{-1})$	0.807
	$G_{\mathrm{C,t}}/(\mathrm{N}\cdot\mathrm{mm}^{-1})$	0.807
	η	1.75

注: T700/LT03A单向预浸带参数根据Chamis 模型纤维体积分数为0.6 时计算。

表9 LT03A 树脂力学性能

Table 9 Mechanical properties of LT03A resin

E ^m /GPa	$\mu^{ m m}$	<i>S</i> /MPa
3.45	0.35	120

边界条件如图 9 所示。在 2 个对称面施加对称 约束,在模型+x 端面的节点和参考点 P_R之间建立 耦合关系,并对参考点 P_R 施加 x 方向的位移载荷。

第3组试验件所含微脉管直径为0.4 mm, 富树 脂区长度为3.35 mm, 建模方法与第2组相同, 此处 不再赘述。



Fig. 9 Finite element model boundary conditions

3 模型验证与失效机理分析

3.1 模型有效性验证

图 10 为结构破坏前(对应图 11 中 2 组曲线上的 D 点)的应力场分布。可见, 2.5 mm 长度外的应力场均分布均匀, 证明了为减少计算量而选用 2.5 mm 的建模长度是合理的。

2组有限元计算所得载荷-位移曲线如图 11 所示,试验和有限元拉伸强度结果对比如表 10 所示。第2组试验件峰值载荷为 360.74 N,相应的



图 10 含 0.255 mm 直径微脉管层合板破坏前的应力分布

Fig. 10 Stress distribution of laminate with 0.255 mm microvascular channel before damage





表 10 试验和有限元所得拉伸强度对比

 Table 10
 Comparison of tensile strength obtained from tests

and finite elements							
4日 모네	微脉签古公/mm	拉伸强	度/MPa	相对设美/0/			
组加	版 协 居且任/IIIII	试验	有限元	11月庆左/70			
2	0.255	1 852.09	1 803.7	-2.6			
3	0.4	1 424.59	1 430.0	0.38			

拉伸强度为 1 803.7 MPa, 对比试验测得的平均拉伸 强度 1 852.09 MPa, 相对误差为-2.6%。第 3 组试 验件计算结果峰值载荷为 286 N, 相应拉伸强度为 1 430.0 MPa, 对比试验平均拉伸强度 1 424.59 MPa, 相对误差为 0.38%。

根据第2组试验件的计算与试验结果,可对其 微脉管位置和不含微脉管位置的刚度进行对比分 析。图12所示有限元模型中*a*、*b*两处单元的*x*方 向应变-载荷曲线如图13所示,*a*点处应变始终小 于*b*点处应变。由图14可以发现,*a*点所在截面 处,越靠近微脉管的铺层应力(该铺层截面的拉伸 应力均值)越大,越靠近表面的铺层应力越小,导致 *a*点的拉伸应力小于*b*点,而2处拉伸模量基本相 同,最终使*a*点沿载荷方向的应变小于*b*点。

端部位移为 0.007 5 mm(层合板应变 3 000 ×10⁻⁶) 时, 层合板表层的纵向应变 ε_{11} 分布如图 15 所示。



图 13 有限元模型微脉管处和无微脉管处应变

Fig. 13 Finite element model strains at microvascular channel and non-microvascular channel

可以看出,微脉管上方的应变最小(约2599×10⁻⁶), 随着与微脉管的距离增大,层合板表层纵向应变逐 渐增大,到富树脂区尖端上方时达到最大值(约 3325×10⁻⁶),之后进入到正常铺层位置表层,纵向 应变为3000×10⁻⁶。

表 11 为 a、b 这 2 处局部刚度的计算结果与试验结果对比。可见,模型的计算结果与试验结果吻合,由此可以证明所建立的精细化有限元模型是有效的。



图 15 第 2 组试验件端部位移 0.007 5 mm 时层合板表层 纵向应变ε₁₁分布

2 780.5 2 720.0 2 659.5 2 599.0

Fig. 15 Laminate surface longitudinal strain ε_{11} distribution at 0.007 5 mm shift of the end of group 2 specimens

表 11 第 2 组试验件 2 处刚度有限元和试验结果对比

 Table 11
 Comparison of finite element and test results of stiffness at two locations for group 2 specimens

合要		相对招关网	
12.直.	有限元	试验	- 相利 庆左/%
a点微脉管处	156.0	147.25	6.0
b点无微脉管处	138.5	139.7	-0.9

为进一步证明在有限元建模中考虑纤维体积 分数变化的必要性,对不考虑微脉管引起铺层纤维 体积分数变化的情况进行计算,即模型中复合材料 全部采用表 8 中纤维体积分数为 0.6 时的 T700/LT03A 材料参数,其余部分的设置与原模型保持一致。第 2、3 组试验件有限元计算的载荷-位移曲线如图 16 所示,相应的拉伸强度和原模型及试验结果对比如 表 12 所示。可以看出,不考虑纤维体积分数变化



1412



时,计算的试验件拉伸强度低于试验结果,第2、 3组试验件的相对误差分别为-17.62%和-10.92%, 充分说明了在进行含微脉管层合板力学性能分析 时考虑微脉管引起纤维体积分数变化的必要性。

3.2 损伤机理分析

由于2组试验件均在达到最大载荷后瞬间发 生断裂,试验过程中未能观测到损伤起始和扩展的 具体过程,为此,本文结合有限元分析就微脉管对 层合板力学性能的影响机理进行分析。

对于含 0.255 mm 直径的微脉管层合板,图 11 中所示的 A~E 点为其出现损伤的关键点。

A 点时, 富树脂区尖端和铺层 1(具体位置见图 8) 之间的 Cohesive 单元开始出现损伤, 如图 17 (a) 所 示, 这是由于微脉管的引入使其上下纤维弯曲, 在 结构受拉过程中, 纤维趋于被拉直, 使此处沿铺层 厚度方向的拉应力σ₃₃较大所致, 如图 17 (b) 所示。

B点时,富树脂区在图 18(a)所示位置开始出现 损伤,这是由于随着拉伸载荷增加,富树脂区尖端 位置拉应力在该处产生了较大的沿层合板厚度方 向的压应力所致。

C点时, 铺层 1 和富树脂区之间及铺层 1 和对称面之间的 Cohesive 单元损伤扩展至图 19(a) 所示

表 12 不考虑纤维体积分数变化时有限元拉伸强度和原模型及试验结果对比

 Table 12
 Comparison of calculated tensile strength without considering change in fiber volume fraction with original model and test results

4日 兄山		拉伸强度/MP	a	相对误	差/%
坦劝	试验	有限元(不考虑纤维体积分数变化)	有限元(考虑纤维体积分数变化)	不考虑纤维体积分数变化	考虑纤维体积分数变化
2	1 852.1	1 525.8	1 603.7	-17.62	-13.41
3	1 424.6	1 269	1 430	-10.92	0.38







图 18 B 点富树脂区状态 Fig. 18 Resin-rich zone status at point B

Cohesive

单元损伤

(Avg: 75%) 0.900 0.825 0.750

0.675

 $\begin{array}{c} 0.600\\ 0.525\\ 0.450\\ 0.375\\ 0.300\\ 0.225\\ 0.150\\ 0.075 \end{array}$

纤维损伤 (Avg: 75%)

纤维损伤 (Avg: 75%)

1.0 0.5 0

1.0 0.5 - X

纤维拉伸 破坏

 \mathbf{x}

(b) D'点时纤维拉伸损伤位置, 铺层1和对称面分离

(c) F'点时纤维拉伸破坏区域和分层情况

图 20 第 3 组试验件有限元载荷-位移曲线各点对应损伤情况

Fig. 20 Damage corresponding to each point of finite element

load-displacement curve of specimen in group 3

对于含 0.4 mm 直径微脉管的层合板, 图 11 中

A'~F'为损伤关键点。A'和 B'处损伤与含直径为

0.255 mm 微脉管的层合板类似, 此处不再赘述, 至

C'点时该处 Cohesive 单元完全破坏, 铺层1和对称

面分离,并且分层继续向两侧扩展,如图 20(a) 所

示,至D'点时,铺层1已经与对称面完全分离,并且

与富树脂区之间也出现明显分层,这与含 0.255 mm

直径微脉管层合板的破坏模式明显不同,这是因为

随着微脉管直径的增加,靠近脉管的纤维弯折角度

讨大所致,而后位于微脉管上方的铺层开始出现纤

维拉伸破坏,如图 20 (b) 所示,此时载荷达到峰值,

位移继续增大时,纤维损伤在铺层1内向上扩展,

同时由于铺层1已经与对称面及富树脂区分离,整

个铺层向+z方向移动,载荷下降至E'点时,靠近表

面的铺层开始出现纤维拉伸破坏,这是由于铺层和

(a) C'点时富树脂区尖端铺层完全分离

完全分层

区域,并且铺层2和铺层3之间、铺层3和铺层4 之间的 Cohesive 单元也开始出现损伤。

D点时,载荷达到峰值,此时微脉管上方的铺层 出现纤维拉伸破坏,如图 19(b)所示,结构失去继续 承载的能力,载荷开始下降,纤维拉伸破坏向厚度 方向扩展。

载荷下降至 E 点时,纤维拉伸破坏贯穿层合板 整个截面,如图 19(c) 所示,并且各层之间均出现分 层,如图 19(d) 所示。



对称面分离后,所有铺层趋于与加载方向平行,即 铺层被拉直,而越靠近微脉管的铺层原始长度越 长,越靠近表面的铺层长度越短,在端部位移一致 的情况下,越靠近表面的铺层拉伸应变越大,因此, 在该处出现纤维拉伸破坏,载荷从 E'点下降至 F'点 过程中,纤维拉伸破坏从靠近表面铺层向下扩展, 同时铺层 2 在微脉管上方位置也出现纤维拉伸破 坏,铺层 1 和铺层 2 之间、铺层 2 和铺层 3 之间也 出现分层,如图 20 (c) 所示。

4 结 论

 1)当载荷方向垂直于微脉管布置方向时,直径 为 0.255 mm 微脉管使单向层合板纵向拉伸强度的 下降大于横向拉伸强度。

2)对于2mm厚的单向层合板,其内含横向布置的微脉管尺寸越大,拉伸强度越低,但刚度所受的影响很小。

3)建立含微脉管层合板有限元模型时,除考虑 由微脉管引起的富树脂区、局部纤维弯曲外,还应 考虑微脉管上下铺层的局部纤维体积分数变化,因 为其会使该区域铺层的模量和强度发生较大 变化。

4)微脉管直径对单向层合板的损伤过程会有 影响。直径较小(0.255 mm)时,微脉管上方的铺层 先出现纤维拉伸破坏,并逐渐沿厚度方向表面扩 展,层间分层出现在整个截面发生纤维拉伸破坏之 后;直径较大(0.4 mm)时,微脉管上方的铺层也是 先出现纤维拉伸破坏,但层间分层出现在整个截面 发生纤维拉伸破坏之前,同时由于铺层和对称面分 离,铺层被拉直,靠近表面的铺层出现纤维拉伸破 坏,并沿厚度方向逐渐向内扩展。

参考文献(References)

- [1] 马埸浩, 杜晓渊, 胡仁伟, 等. 微脉管型自修复复合材料研究进展
 [J]. 高分子材料科学与工程, 2018, 34(1): 166-172.
 MA Y H, DU X Y, HU R W, et al. Development of self-healing composite materials with microvascular networks[J]. Polymer Materials Science and Engineering, 2018, 34(1): 166-172(in Chinese).
- [2] 李元杰,律微波,孟宪铎. 微胶囊自修复聚合物材料的研究进展
 [J]. 工程塑料应用, 2005, 33(1): 68-70.
 LI Y J, LV W B, MENG X D. Research progress in self-repair polymer materials with microcapsule[J]. Engineering Plastics Applica-

tion, 2005, 33(1): 68-70(in Chinese).
[3] 梁大开,杨红.采用空心光纤自诊断、自修复智能结构的研究[J].

压电与声光, 2002, 24(4): 261-263. LIANG D K, YANG H. Research on self-diagnose and self-repair net in smart composite structure by hollow-center optic fiber[J].

[4] 印明勋,齐德胜,云庆文.航空复合材料自修复研究进展

[C]//2015 年第二届中国航空科学技术大会. 北京: 中国航空学会, 2015: 185-189.

YIN M X, QI D S, YUN Q W. Research development of self-healing aeronautical composite materials[C]//Proceedings of the 2nd China Aviation Science and Technology Conference in 2015. Beijing: Chinese Society of Aeronautics and Astronautics, 2015: 185-189(in Chinese).

- [5] KOUSOURAKIS A, BANNISTER M K, MOURITZ A P. Tensile and compressive properties of polymer laminates containing internal sensor cavities[J]. Composites Part A:Applied Science and Manufacturing, 2008, 39(9): 1394-1403.
- [6] KOUSOURAKIS A, MOURITZ A P, BANNISTER M K. Interlaminar properties of polymer laminates containing internal sensor cavities[J]. Composite Structures, 2006, 75(1-4): 610-618.
- [7] KOUSOURAKIS A, MOURITZ A P. The effect of self-healing hollow fibres on the mechanical properties of polymer composites
 [J]. Smart Materials and Structures, 2010, 19(8): 085021.
- [8] HUANG C Y, TRASK R S, BOND I P. Characterization and analysis of carbon fibre-reinforced polymer composite laminates with embedded circular vasculature[J]. Journal of the Royal Society Interface, 2010, 7(49): 1229-1241.
- [9] HUANG C Y, TRASK R S, BOND I P. Analytical study of vascular networks for self-healing composite laminates[C]//Proceedings of the 17th International Conference on Composite Materials. Edinburgh: ICCM, 2009: 27-31.
- [10] LUTERBACHER R, TRASK R S, BOND I P. Static and fatigue tensile properties of cross-ply laminates containing vascules for self-healing applications[J]. Smart Materials and Structures, 2016, 25(1): 015003.
- [11] TRASK R S, BOND I P. Bioinspired engineering study of plantae vascules for self-healing composite structures[J]. Journal of the Royal Society Interface, 2010, 7(47): 921-931.
- [12] NGUYEN A T T, ORIFICI A C. Structural assessment of microvascular self-healing laminates using progressive damage finite element analysis[J]. Composites Part A:Applied Science and Manufacturing, 2012, 43(11): 1886-1894.
- [13] AL-SHAWK A, TANABI H, SABUNCUOGLU B. Investigation of stress distributions in the resin rich region and failure behavior in glass fiber composites with microvascular channels under tensile loading[J]. Composite Structures, 2018, 192: 101-114.
- [14] TANABI H, AL-SHAWK A, SABUNCUOGLU B. Stress concentrations in composites with microvascular channels[J]. Procedia Structural Integrity, 2017, 6: 56-63.
- [15] ASTM. Standard test method for tensile properties of polymer matrix composite materials: ASTM D3039[S]. West Conshohocken: ASTM, 2000.
- [16] 马埸浩. 含自修复微脉管层合板力学性能研究[D]. 北京: 北京航空航天大学, 2021: 87.

MA Y H. Mechanical properties of composite laminates with selfhealing microvascular networks[D]. Beijing: Beihang University, 2021: 87(in Chinese).

[17] CHAMIS C C. Mechanics of composite materials: past, present, and future[J]. Journal of Composites, Technology and Research, 1989, 11(1): 3-14.

1415

[18] YOUNES R, HALLAL A, FARDOUN F, et al. Comparative review study on elastic properties modeling for unidirectional composite materials[M]//HU N. Composites and their properties. London: IntechOpen, 2012, 17: 391-408.

[19] HASHIN Z. Failure criteria for unidirectional fiber composite[J]. Journal of Applied Mechanics, 1980, 47(2): 329-334.

Study of tensile properties of laminates containing microvascular channels with different diameters

RAN Guangling^{1, 2}, MA Yihao^{1, 3}, AN Ziqian¹, ZHAO Dafang⁴, GUO Xin¹, CHENG Xiaoquan^{1, 2, *}

(1. School of Aeronautic Science and Engineering, Beihang University, Beijing 100191, China;

2. Ningbo Institute of Technology, Beihang University, Ningbo 315800, China;

3. Navigation and Control Technology Research Institute of China North Industries Group Corporation, Beijing 100089, China;

4. AVIC Composites Co., Ltd., Beijing 101300, China)

Abstract: In view of the influence of microvascular on the mechanical properties of laminates, the tensile properties of laminates with varying diameter microvascular channels were studied. Tensile strength and stiffness of laminates with varying diameter microvascular channels were determined by experimental measurements. A refined finite element model of laminates containing microvascular channels was established, which considered the effects of resin-rich area, fiber bending, and fiber volume fraction change, and the tensile properties of laminates containing microvascular channels were verified by the experimental results. On this basis, the effects of microvascular channel orientation and diameter were investigated. The results show that when the microvascular channel is laid in a vertical ply direction, it decreases the longitudinal tensile strength of the unidirectional laminate, but the stiffness is basically unaffected. The larger the diameter of the microvascular channels with diameters of 0.255 mm and 0.4 mm is reduced by 21.9% and 39.9%, respectively, when compared to the tensile properties of the laminate without the channels; however, its stiffness changes less than 1.4%. The damage process of laminate also changes with the microvascular channel diameter.

Keywords: composite laminate; self-healing; microvascular; tensile properties; finite element

Received: 2022-06-16; Accepted: 2022-08-21; Published Online: 2022-08-29 16:42 URL: link.cnki.net/urlid/11.2625.V.20220829.1421.002

^{*} Corresponding author. E-mail: xiaoquan_cheng@buaa.edu.cn

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

DOI: 10.13700/j.bh.1001-5965.2022.0492

多策略融合改进的自适应蜉蝣算法

蒋宇飞,许贤泽*,徐逢秋,高波

(武汉大学电子信息学院,武汉 430072)

摘 要:为改进蜉蝣算法全局搜索能力较差、种群多样性较小和自适应能力弱等问题,提 出一种多策略融合改进的自适应蜉蝣算法 (MIMA)。采用 Sin 混沌映射初始化蜉蝣种群,使种群能 够均匀分布在解空间中,提高初始种群质量,增强全局搜索能力;引入 Tent 混沌映射和高斯变异对 种群个体进行调节,增加种群多样性的同时调控种群密度,增强局部最优逃逸能力;引入不完全伽 马函数,重构自适应动态调节的重力系数,建立全局搜索和局部开发能力之间更好的平衡,进而提 升算法收敛精度,有利于提高全局搜索能力;采用随机反向学习 (ROBL)策略,增强全局搜索能 力,提高收敛速度并增强稳定性。利用经典测试函数集进行算法对比,并利用 Wilcoxon 秩和检验 分析算法的优化效果,证明改进的有效性和可靠性。实验结果表明:所提算法与其他算法相比,寻 优精度、收敛速度、稳定性都取得了较大提升。

关 键 词: 蜉蝣算法; 混沌映射; 高斯变异; 自适应动态调节; 随机反向学习 中图分类号: TP301.6

文献标志码: A 文章编号: 1001-5965(2024)04-1416-11

元启发式算法是对启发式算法的改进,是随机 算法与局部搜索算法相结合的产物,如粒子群优化 (particle swarm optimization, PSO)算法^[1]、遗传算法 (genetic algorithm, GA)、麻雀搜索算法 (sparrow search algorithm, SSA)^[2]、灰狼 (grey wolf optimization, GWO) 算法^[3]、萤火虫算法 (firefly algorithm, FA)^[4]等。

蜉蝣算法 (mayfly algorithm, MA) 受蜉蝣的飞行 行为和交配行为启发, 在结构上结合 PSO、GA、FA 算法进行改进, 是一种综合群体智能和进化算法主 要优点的新型元启发式算法。改进蜉蝣算法 (improved mayfly algorithm, IMA)由 Zervoudakis 和 Tsafarakis^[5]提出, 针对 MA 的收敛稳定性问题、开 发与探索之间不平衡的问题进行改进:①增加速度 限制;②为重力系数设置阻尼;③为婚礼舞蹈和随 机飞行设置阻尼;④对子代基因随机变异, IMA 在 包括 13 个 CEC2017 测试函数的 38 个数学基准函 数上均优于 MA。IMA 虽然继承了 MA 良好的性 能和优势,并取得了较好的效果,但前期全局搜索 能力不理想,导致后期即使具备较好的开发能力也 难以达到理想的求解精度和算法稳定性,变异过程 简单,种群多样性较小,易陷入局部最优解。

近年来,针对群体智能算法的缺点,研究人员 提出了不同策略优化改进方案,以改善其全局搜索 能力,增加种群多样性。文献[6]利用 Circle 混沌 映射初始化鲸鱼优化算法的种群,使种群初始化过 程中有更好的遍历性,降低了群体的盲目性,且引 入反向学习 (opposition-based learning, OBL)策略对 鲸鱼的位置求出反向解,提高了鲸鱼优化算法的搜 索能力和种群的多样性;文献[7]利用 Sin 混沌映 射初始化 MA 的种群,获得更高质量初始解,证明 了混沌映射方法对 MA 的有效性;文献[8]采用 Tent 混沌扰动和柯西变异优化种群,避免种群过于 "集中"或"分散",提高了种群的多样性,增强了探 索能力;文献[9]通过引入不完全的伽马函数和

收稿日期: 2022-06-16; 录用日期: 2022-09-23; 网络出版时间: 2022-10-10 10:36 网络出版地址: link.cnki.net/urlid/11.2625.V.20221009.1050.010 基金项目: 国家自然科学基金 (51975422)

*通信作者. E-mail: xxz@whu.edu.cn

引用格式: 蒋宇飞, 许贤泽, 徐逢秋, 等. 多策略融合改进的自适应蜉蝣算法 [J]. 北京航空航天大学学报, 2024, 50 (4): 1416-1426. JIANG Y F, XU X Z, XU F Q, et al. Multi-strategy fusion improved adaptive mayfly algorithm [J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50 (4): 1416-1426 (in Chinese). Beta 累加分布函数,构建自适应动态调节的重力系数,增强全局搜索和开发之间的平衡,提升了算法的自适应动态搜索能力。

为改善 IMA 的全局搜索能力, 增加种群多样 性,提高搜索的自适应性,本文提出多策略融合改 进的自适应蜉蝣算法 (multi-strategy fusion improved adaptive mayfly algorithm, MIMA)。①使用 Sin 混沌 映射初始化蜉蝣种群,使蜉蝣种群在解空间中均匀 分布,提高前期全局搜索能力;②引入 Tent 混沌映 射和高斯变异对种群个体进行调节,避免种群过于 "集中"或"分散",提高种群的多样性,增强局部最 优逃逸能力:③引入不完全伽马函数,重构自适应 动态调节的重力系数,对全局搜索能力和局部开发 能力建立更好的平衡;④引入随机反向学习 (random opposition-based learning, ROBL)策略, 增加种群的 多样性,补足全局搜索能力,提高收敛速度并增强 稳定性。通过经典测试函数集进行优化对比,验证 了 MIMA 的寻优精度、收敛速度、稳定性都取得较 大提升。

1 改进蜉蝣算法原理

IMA 通过将最优雄性个体与最优雌性个体交 配得到最优子代,寻优过程为:雄性蜉蝣的运动、雌 性蜉蝣的运动、蜉蝣交配与变异。

1.1 雄性蜉蝣的运动

设雄性蜉蝣总数为N,每只雄性蜉蝣根据自己的经验和邻居的经验,调整自身位置。假设x;为第 i个蜉蝣于时间步t在搜索空间U中的位置,通过在 当前位置上添加一个速度v;⁺¹来改变位置:

 $\mathbf{x}_{i}^{t+1} = \mathbf{x}_{i}^{t} + \mathbf{v}_{i}^{t+1} \tag{1}$

式中:雄性蜉蝣位置 $x_i \in U(x_{\min}, x_{\max})$,若添加速度 v_i^{t+1} 后超出U,则将其限制回最近边界值。

雄性蜉蝣的速度**v**^{t+1}表示为

$$v_{i,j}^{t+1} = \begin{cases} gv_{i,j}^{t} + a_1 e^{-\beta r_{g}^{t}} (p_{\text{best},i,j} - x_{i,j}^{t}) + \\ a_2 e^{-\beta r_{g}^{2}} (g_{\text{best},i} - x_{i,j}^{t}) & f(\boldsymbol{x}_{i}^{t}) > f_{\min} \\ gv_{i,j}^{t} + dr & f(\boldsymbol{x}_{i}^{t}) = f_{\min} \end{cases}$$
(2)

若改变后的速度超过范围[**v**_{min},**v**_{max}],则将其限制回最近边界值。式中: v^t_{i,j}为第*i*个雄性蜉蝣于时间步t在维度 j(j = 1,…,n)上的速度; x^t_{i,j}为第*i*个蜉蝣于时间步t在维度 j上的位置; g为重力系数,其迭代公式为

$$g^{t+1} = g^t g_{\text{damp}} \tag{3}$$

其中:g^t为时间步t时的重力系数;g_{damp}为重力系数 阻尼。

a1和a2为正吸引常数,分别用于缩放认知和社

会部分的贡献, β为固定的可见度系数, 用于限制蜉 蝣的可见度, r_p 、 r_g 分别为 x_i 与 $p_{\text{best},i}$ 、 x_i 与 g_{best} 之间 的笛卡儿距离, 计算公式为

$$\|\boldsymbol{x}_{i} - \boldsymbol{X}_{i}\| = \sqrt{\sum_{j=1}^{n} (x_{i,j} - X_{i,j})}$$
(4)

式中: $x_{i,j}$ 为蜉蝣 x_i 在维度j上的位置; X_i 对应 $p_{\text{best},i}$ 与 g_{best} 。

f表示适应度函数,越小则函数越适应, fmin即 雄性蜉蝣中最小适应度函数值。最优位置雄性蜉 蝣进行婚礼舞蹈,改变速度, d为婚礼舞蹈系数, r为 [-1,1] 之间的随机数。d的迭代公式为

$$d^{t+1} = d^t d_{damp} \tag{5}$$

式中: d'为时间步t时的婚礼舞蹈系数; d_{damp}为舞蹈 系数阻尼。

1.2 雌性蜉蝣的运动

雌性蜉蝣最大的行为特征在于飞向雄性蜉蝣 以繁殖。与雄性蜉蝣类似,设雌蜉蝣总数为N。

假设yⁱ为第i个雌性蜉蝣于时间步t在搜索空间 U中的位置,通过在当前位置上添加速度vⁱ⁺¹来改变 位置。雌雄蜉蝣位置范围为U(x_{min},x_{max}),若添加速 度vⁱ⁺¹后超出U,则将其限制回最近边界值。

吸引过程设定为最优雄性吸引最优雌性,第二 优雄性吸引第二优雌性。雌性蜉蝣速度为

$$v_{i,j}^{t+1} = \begin{cases} gv_{i,j}^{t} + a_3 e^{-\beta r_{ml}^{2}} (x_{i,j}^{t} - y_{i,j}^{t}) & f(\mathbf{y}_{i}^{t}) > f(\mathbf{x}_{i}^{t}) \\ gv_{i,j}^{t} + lr & f(\mathbf{y}_{i}^{t}) \le f(\mathbf{x}_{i}^{t}) \end{cases}$$
(6)

若改变后的速度超过范围[v_{min},v_{max}],则将其限 制回最近边界值。式中:v_{i,j}为第*i*个蜉蝣于时间步 *t*在维度*j*(*j* = 1,…,*n*)上的速度; y_{i,j}为第*i*个蜉蝣于 时间步*t*在维度*j*上的位置; *a*₃为正吸引常数; r_{mf}为 雌性蜉蝣与对应雄性蜉蝣之间的笛卡儿距离,通过 式(4)计算; *l*为随机飞行系数,当雌性蜉蝣没有被 雄性蜉蝣吸引时使用,此时雌性蜉蝣随机飞行,其 迭代公式为

$$l^{t+1} = l^t l_{damp} \tag{7}$$

式中: *l* 为时间步*t*时的随机飞行系数; *l*_{damp}为随机飞行系数阻尼。

1.3 蜉蝣交配及变异

交配过程以交叉算子表示,从雄性蜉蝣中选择 父本,雌性蜉蝣中选取母本,二者在本性别种群适 应度排名相同。采用优胜劣汰机制,将最优个体的 雄性和雌性蜉蝣进行繁殖得到最优个体,依次类 推,得到的2个子代表达式为

$$\begin{cases} c_1 = Lm + (1 - L)f + \sigma N(0, 1) \\ c_2 = Lf + (1 - L)m + \sigma N(0, 1) \end{cases}$$
(8)

式中: c₁为雄性蜉蝣子代; c₂为雌性蜉蝣子代; L为 [-1,1]范围内服从高斯分布的随机数; m为父本; f为母本; σN(0,1)表示服从高斯分布的、均值为 0、 方差为 1 的随机数组成的1×n矩阵。

2 本文算法

2.1 Sin 混沌映射初始化种群

群体初始化对当前大多数智能优化算法的效 率有很大影响,均匀分布的种群可以适度地扩大算 法的搜索范围,从而提高收敛速度和求解精度。混 沌映射具有随机性、遍历性和规律性特点,常用于 优化搜索问题,以保持种群多样性及跳出局部最优 解。本文以 Sin 混沌映射生成初始种群,定义如下:

 $y_{n+1} = \mu \sin(\pi y_n) \qquad 0 \le y_n \le 1 \tag{9}$

式中: y为迭代值; µ为 [0,1] 间的控制参数。

通过 Sin 混沌映射可以使生成的初始种群较为 均匀地充斥整个解空间,提高初始种群质量,利于 前期的全局搜索。

2.2 基于 Tent 混沌映射和高斯变异的种群调节

为增加种群多样性、提高求解进度,使算法不 易在迭代末期陷入局部最优区域而无法脱出,需要 在每个迭代环节末对种群进行调节。设*f*ⁱⁿ为蜉蝣 适应度函数值,*f*a为种群适应度函数值的平均值, 按如下方式判断:

1) 若 f_i^m < f_a,即出现"聚集"现象,进行高斯变 异,若新位置适应度函数值低于旧位置,则进行位 置替换。

2) 若 $f_i^m \ge f_a$,即出现"发散"现象,进行 Tent 混 沌映射,按同样原则进行位置替换。

2.2.1 Tent 混沌映射

参考文献 [8,10], 若利用混沌映射初始化种群 后再对种群进行调节, 使用多种混沌映射能避免多 次陷入重复点, 导致调节失效。Tent 混沌映射定义 如下:

$$y_{n+1} = \begin{cases} 2y_n & y_n \in [0, 0.5] \\ 0.5y_n & y_n \in (0.5, 1] \end{cases}$$
(10)

但 Tent 混沌映射容易在小循环周期和不动点 上出现问题,针对此,文献 [11] 在 Tent 混沌映射基 础上引入随机变量,改进后表达式为

$$y_{n+1} = \begin{cases} 2y_n + \operatorname{rand}(0, 1)/M & y_n \in [0, 0.5] \\ 0.5y_n + \operatorname{rand}(0, 1)/M & y_n \in (0.5, 1] \end{cases}$$
(11)

式中: *M*为混沌序列内粒子个数; rand(0,1)为 [0,1] 间的随机数。

2.2.2 高斯变异

高斯分布即正态分布,在自然界中,很多随机

因素和影响都可以近似地用高斯分布来描述,许多 概率分布可以也用高斯分布来近似或导出,其在统 计学等诸多领域有重大的影响力。高斯变异^[12]是 在原有个体的状态上加一个服从高斯分布的随机 向量,定义如下:

 $mutation(\mathbf{x}) = \mathbf{x}[1 + \sigma N(0, 1)]$ (12)

式中:x为蜉蝣原本的位置;mutation(x)为变异后的 蜉蝣位置。

由高斯分布特性可知,多数变异算子分布在原始位置周围,局部搜索能力强,能够提高优化算法的寻优精度,有利于跳出局部最优区域。同时,少数算子远离当前位置,提高了种群的多样性,有利于更好地搜索潜在区域,从而加快搜索速度,提高逃出局部最优解的概率。

2.3 不完全伽马函数的自适应重力系数

重力系数或称惯性权重,对算法的搜索能力和 开发能力具有一定的指导作用,体现出蜉蝣能够借 鉴一定先验行为的能力。文献 [13] 指出,较大的重 力系数具有良好的全局搜索能力,较小的重力系数 具有较好的局部开发能力。本文引入一种非线性 递减的自适应重力系数,从而更好地平衡全局搜索 和局部开发能力。为与 IMA 的重力系数g区别,设 为g(t),表达式为

$$g(t) = (1 - t/T)^{\sqrt[q]{t/T}}$$
(13)

式中: *t*为当前时间步; *T*为最大迭代次数; α为重力 系数的控制系数。

考虑到具有确定性数学表达式的重力系数按照固定规律搜索,很容易陷入局部最优,达不到真正的自适应机制。因此,引入不完全伽马函数^[14],重构自适应动态调节的重力系数g'(t),其表达式如下:

 $g'(t) = (1 - t/T)^{\sqrt[q]{t/T}} \Gamma(\lambda, 1 - t/T)$ (14)

式中: **Γ**为不完全的伽马函数; **λ**为大于 **0** 的随机变量, 取 0.1。

雄性蜉蝣的速度表达式更新如下:

$$v_{i,j}^{t+1} = \begin{cases} g'(t)v_{i,j}^{t} + a_{1}e^{-\beta r_{p}^{2}}(p_{\text{best},i,j} - x_{i,j}^{t}) + \\ a_{2}e^{-\beta r_{g}^{2}}(g_{\text{best}} - x_{i,j}^{t}) & f(\mathbf{x}_{i}^{t}) > f_{\min} \\ g'(t)v_{i,j}^{t} + dr & f(\mathbf{x}_{i}^{t}) = f_{\min} \\ \end{cases}$$
(15)

若改变后的速度超过范围[*v*_{min},*v*_{max}],则限制为边界值。

雌性蜉蝣的速度表达式更新如下:

$$v_{i,j}^{t+1} = \begin{cases} g'(t)v_{i,j}^{t} + a_{3}e^{-\beta r_{ml}^{2}}(x_{i,j}^{t} - y_{i,j}^{t}) & f(\mathbf{y}_{i}^{t}) > f(\mathbf{x}_{i}^{t}) \\ g'(t)v_{i,j}^{t} + lr & f(\mathbf{y}_{i}^{t}) \leqslant f(\mathbf{x}_{i}^{t}) \end{cases}$$
(16)

若改变后的速度超过范围[v_{min},v_{max}],则限制为边界值。

为确定合适的α值,利用经典测试函数f_i对不同参数独立运行 50次,最大迭代1000次,设置维度为20,记录不同α值时的平均值、最优值和标准差。

由表 1 可知, 重力系数的控制系数 α 对算法的 寻优能力有一定影响。当 α = 1时, MIMA 可以表现 出较好的搜索能力。

表 1 不同 α 值的搜索结果 Table 1 Search results with different α values

α	平均值	最优值	标准差
1	1.4275×10^{-40}	5.165 1×10 ⁻⁴⁸	4.4247×10^{-40}
1.5	$2.923 9 \times 10^{-40}$	$3.7552{\times}10^{-47}$	1.317 3×10 ⁻³⁹
2	1.043 5×10 ⁻³⁹	5.166 7×10 ⁻⁴⁷	4.285 5×10 ⁻³⁹
2.5	$3.823.0 \times 10^{-39}$	8.3485×10^{-50}	$2.3260{ imes}10^{-38}$
3	1.739 6×10 ⁻³⁸	$2.5809{ imes}10^{-48}$	6.760 1×10 ⁻³⁸

2.4 随机反向学习策略

OBL 策略^[15] 是一种于 2005 年提出的群智能领

域中的改进策略,其思想为: 在种群寻优的过程中, 以上下界之和减去当前解从而产生反向解,比较当 前解与反向解的适应度函数值,择优进入下一次迭 代。OBL 策略生成的反向解与当前解距离为定值, 缺乏随机性,无法有效增强种群多样性。因此, Long 等^[16]提出了改进的 ROBL 策略,加强全局搜 索能力,减少算法陷入局部最优解无法逃逸的概 率。变异公式如下:

$$mutation(\mathbf{x}) = \mathbf{x}_{min} + \mathbf{x}_{max} - r\mathbf{x}$$
(17)

2.5 算法流程

本文提出的基于 Tent 混沌映射和高斯变异的 种群调节及 ROBL 策略都采用贪心算法。简而言 之,即当且仅当变异后蜉蝣个体的适应度函数值减 小,才进行此次变异,蜉蝣个体的位置才会因此发 生改变。

综合引入的 4 项改进策略,本文提出的 MIMA 流程如图 1 所示。



图 1 MIMA 算法流程 Fig. 1 MIMA Algorithm Flow Chart

3 算法性能测试

为验证 MIMA 的有效性和优异性,本文以 IMA2^[17]、TCMA^[18] 2 种基于 IMA 的改进算法作为 对比,并引入 GWO 算法、正余弦优化算法 (sine cosine algorithm, SCA)、PSO 等经典优化算法进行 复现比较。

3.1 参数设置

实验选取种群规模为 40(雄性蜉蝣 20, 雌性蜉 蝣 20), 种群交叉个数为 20, 突变率为 0.01。各算法 具体参数设置为: MIMA、IMA、IMA2 和 TCMA 中, α =1.0, l=1, l_{damp} =0.99, d=5, d_{damp} =0.8, a_1 =1.0, a_2 =1.5, a_3 =1.5, β =2, g_0 =0.8; PSO 算法中, c_1 =1.494 45, c_2 = 1.494 45。

3.2 测试函数

实验利用 14 个经典基准测试函数进行测试, 其中, $f_1 \sim f_6$ 为不定维单峰函数、 $f_7 \sim f_{11}$ 为不定维多 峰函数, $f_{12} \sim f_{14}$ 为固定维函数。表 2 为测试函数的 具体信息。

3.3 结果分析

3.3.1 寻优精度和稳定性分析

在 MATLAB R2018b 环境下对所有测试函数进行仿真对比实验,不定维函数设置维度为 20,实验均独立运行 50次,最大迭代次数为1000。记录各算法的平均值、最优值、标准差,测试结果如表 3 所示。

由表3可知,对于不定维单峰函数,除f4外, MIMA 相对其他所有算法具有显著性优势。针对 f4测试时, MIMA、IMA、TCMA 均能找到理论最优 解0,但三者小概率出现进入局部最优解而无法在 1000次迭代内跳出,从而导致平均值、标准差差别 不大;而IMA2则可以稳定跳出该局部最优解。 对于不定维多峰函数, MIMA 的各项测试指标 除 f_7 、 f_{11} 外均显著优于其他算法。针对 f_7 测试时, MIMA 逊于 IMA2; 针对 f_{11} 测试时, 二者平均值较为 接近, 但 MIMA 方差更小, 具有更强的稳定性和鲁 棒性。

对于固定维函数,由于 IMA 本身实验效果已非 常接近理论最优解,所有基于 IMA 的改进算法均 难以取得大幅度的提升(仅 IMA2 在针对 f₁₃进行测 试时出现不能稳定获取最优解的情况),但 MIMA 在针对 f₁₂、f₁₃测试时,稳定性相对更好。

综上所述, MIMA 的优势直观上较为明显。 3.3.2 收敛曲线分析

为更直观地比较 7 种算法的收敛速度,分别用 7 种算法对测试函数进行实验,平均收敛曲线如图 2 所示。

由图 2 可知, 对于图 2(a)~图 2(f), MIMA 寻优 精度最高, 除图 2(d) 外, MIMA 收敛速度均优于其 他算法。图 2(d) 中, 起步阶段, MIMA、IMA、IMA2、

衣	2	垫/ 准/测证	(四3	X.
Table 2	Ber	nchmark	test	function

函数公式	维度	搜索空间	f_{\min}
$f_1\left(x\right) = \sum_{i=1}^n x_i^2$	20	[-100,100]	0
$f_2(x) = \sum_{i=1}^n x_i + \prod_{i=1}^n x_i $	20	[-10,10]	0
$f_3(x) = \max\left\{ x_i \right\}$	20	[-100,100]	0
$f_4(x) = \sum_{i=1}^n (x_i + 0.5)^2$	20	[-100,100]	0
$f_5(x) = \sum_{i=1}^n x_i^4 + random[0, 1)$	20	[-1.28.1.28]	0
$f_6(x) = \sum_{i=1}^n x_i ^{i+1}$	20	[-1,1]	0
$f_{7}(x) = -\sum_{i=1}^{n} \left[x_{i} \sin\left(\sqrt{x_{i}}\right) \right]$	20	[-500,500]	-8 379.66
$f_{8}(x) = \sum_{i=1}^{n} \left(x_{i}^{2} - 10\cos(2\pi x_{i}) + 10 \right)$	20	[-5.12,5.12]	0
$f_9(x) = -20 \exp\left(-0.2 \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}\right) - \exp\left(\frac{1}{n} \sum_{i=1}^n \cos(2\pi x_i)\right) + 20 + e$	20	[-32,32]	0
$f_{10}(x) = \frac{1}{4\ 000} \sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n} \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$	20	[-600,600]	0
$f_{11}(x) = \frac{\pi}{n} \left[10\sin^2(\pi y_i) + \sum_{i=1}^{n-1} (y_i - 1)^2 \left(1 + 10\sin^2(\pi y_{i+1}) \right) + (y_n - 1)^2 \right] + \sum_{i=1}^n u(x_i, 10, 100, 4)$	20	[-50,50]	0
$f_{12}(x) = \left[0.002 + \sum_{j=1}^{25} \frac{1}{j + \sum_{i=1}^{2} (x_i - a_{ij})^6} \right]^{-1}$	2	[-65.53,65.53]	1
$f_{13}(x) = \sum_{i=1}^{11} \left[a_i - \frac{x_i \left(b_i^2 + b_i x_2 \right)}{b_i^2 + b_i x_3 + x_4} \right]^2$	4	[-5,5]	3.075×10^{-4}
$f_{14}(x) = \left(x_2 - \frac{5.1}{4\pi^2}x_1^2 + \frac{5}{\pi}x_1 - 11\right)^2 + 10\left(1 - \frac{1}{8\pi}\right)\cos x_1 + 10$	2	[-5,10]	0.398

表 3 基准测试函数结果对比 Table 3 Comparison of benchmark test function results

				平均值			
测试函数	MIMA	IMA	PSO	GWO	SCA	IMA2	TCMA
f_1	1.486×10^{-192}	6.023×10 ⁻³⁹	2.007	1.945×10 ⁻⁶⁵	2.717×10 ⁻⁸	1.229×10 ⁻⁴⁴	2.310×10 ⁻⁴⁴
f_2	7.613×10 ⁻⁹⁸	2.106×10^{-16}	8.188×10^{-1}	2.540×10^{-38}	1.844×10^{-9}	1.576×10^{-24}	6.570×10 ⁻²¹
f_3	9.818×10^{-64}	2.106×10 ⁻³	8.653×10^{-1}	3.919×10 ⁻¹⁶	5.631×10^{-1}	4.224×10 ⁻⁴	2.976×10 ⁻⁴
f_4	4.930×10^{-34}	1.972×10^{-33}	2.283	4.616×10 ⁻¹	2.048	0	2.465×10^{-33}
f_5	1.254×10^{-4}	5.163×10^{-3}	5.900×10^{-3}	6.322×10^{-4}	6.508×10^{-3}	3.237×10^{-3}	4.164×10 ⁻³
f_6	1.864×10^{-294}	7.675×10^{-74}	4.774×10^{-8}	2.650×10^{-221}	4.803×10^{-9}	2.972×10^{-72}	2.413×10 ⁻⁷³
f_7	-7.445×10^{3}	-6.881×10^{3}	-5.169×10^{2}	-6.130×10 ³	-3.234×10^{3}	-8.380×10^{3}	-6.557×10 ³
f_8	0	2.815	1.909×10 ¹	4.887×10^{-1}	2.379	3.317×10^{-1}	4.345
f_9	8.882×10^{-16}	1.182	1.840	1.489×10^{-14}	8.627	2.599×10^{-14}	1.614
f_{10}	0	3.780×10^{-2}	1.764×10^{-1}	1.776×10^{-3}	6.454×10^{-2}	2.254×10^{-2}	2.146×10 ⁻¹⁶
f_{11}	2.382×10^{-32}	3.421×10^{-2}	4.423×10^{-2}	3.328×10^{-2}	2.622×10^{-1}	2.408×10^{-32}	5.183×10 ⁻³
f_{12}	9.980×10^{-1}	9.980×10^{-1}	1.267×10^{1}	4.131	9.982×10^{-1}	9.980×10^{-1}	9.980×10 ⁻¹
f_{13}	3.075×10^{-4}	3.075×10^{-4}	6.628×10^{-4}	4.376×10^{-3}	9.713×10 ⁻⁴	1.645×10^{-3}	3.075×10 ⁻⁴
f_{14}	3.979×10^{-1}	3.979×10 ⁻¹	3.979×10^{-1}	3.979×10 ⁻¹	3.987×10^{-1}	3.979×10 ⁻¹	3.979×10 ⁻¹
				最优值			
测试函数	MIMA	IMA	PSO	GWO	SCA	IMA2	ТСМА
f_1	5.722×10 ⁻¹⁹³	8.786×10 ⁻⁴⁹	6.672×10^{-1}	8.155×10 ⁻⁶⁸	1.571×10^{-14}	1.468×10 ⁻⁵¹	2.017×10 ⁻⁵¹
f_2	1.450×10^{-99}	1.962×10^{-26}	5.817×10^{-1}	3.137×10^{-39}	6.787×10^{-13}	2.663×10^{-30}	5.440×10 ⁻²⁷
f_3	1.010×10^{-97}	5.682×10 ⁻⁵	4.711×10^{-1}	8.026×10^{-18}	6.846×10 ⁻³	4.710×10 ⁻⁵	1.529×10 ⁻⁵
f_4	0	0	1.175	1.410×10^{-5}	1.579	0	0
f_5	1.227×10^{-7}	1.762×10^{-3}	9.874×10^{-4}	1.539×10^{-4}	4.763×10 ⁻⁴	1.334×10 ⁻³	1.468×10 ⁻³
f_6	0	3.927×10^{-93}	4.576×10 ⁻⁹	9.344×10^{-231}	9.462×10^{-22}	1.576×10^{-93}	1.037×10^{-86}
f7	-7.926×10^{3}	-7.787×10^{3}	-1.506×10^{3}	-8.067×10^{3}	-3.824×10^{3}	-8.380×10^{3}	-7.432×10 ³
f_8	0	5.684×10^{-14}	8.689	0	5.684×10^{-14}	0	9.950×10 ⁻¹
f_9	8.882×10^{-16}	2.931×10^{-14}	1.015	7.994×10^{-15}	1.877×10^{-8}	7.994×10^{-15}	6.401×10^{-6}
f_{10}	0	0	4.317×10^{-2}	0	8.378×10^{-13}	0	0
f_{11}	2.356×10 ⁻³²	2.356×10^{-32}	7.751×10^{-3}	6.514×10^{-3}	1.546×10^{-1}	2.356×10^{-32}	2.356×10 ⁻³²
f_{12}	9.980×10^{-1}	9.980×10^{-1}	1.267×10^{1}	9.980×10^{-1}	9.980×10^{-1}	9.980×10^{-1}	9.980×10 ⁻¹
f_{13}	3.075×10^{-4}	3.075×10^{-4}	3.166×10 ⁻⁴	3.075×10^{-4}	3.436×10 ⁻⁴	3.075×10^{-4}	3.075×10 ⁻⁴
f_{14}	3.979×10^{-1}	3.979×10^{-1}	3.979×10^{-1}	3.979×10^{-1}	3.979×10^{-1}	3.979×10^{-1}	3.979×10 ⁻¹
				标准差			
测试函数	MIMA	IMA	PSO	GWO	SCA	IMA2	TCMA
f_1	0	1.887×10 ⁻³⁷	7.375×10^{-1}	3.560×10 ⁻⁶⁵	1.025×10 ⁻⁷	2.807×10^{-44}	7.722×10 ⁻⁴⁴
f_2	1.161×10^{-97}	1.043×10^{-14}	1.153×10^{-1}	2.434×10^{-38}	4.370×10^{-9}	6.994×10^{-24}	2.832×10^{-20}
f3	4.279×10^{-63}	6.188×10^{-3}	1.758×10^{-1}	6.989×10^{-16}	1.159	7.688×10^{-4}	1.555×10^{-4}
f_4	2.415×10 ⁻³³	7.542×10^{-33}	7.023×10^{-1}	2.647×10^{-1}	2.190×10^{-1}	0	1.121×10^{-32}
f_5	1.318×10^{-4}	2.830×10^{-3}	3.497×10^{-3}	3.159×10^{-4}	5.658×10^{-3}	1.292×10^{-3}	2.459×10 ⁻³
f_6	0	4.000×10^{-73}	3.824×10^{-8}	0	1.413×10^{-8}	1.590×10^{-71}	8.527×10^{-73}
f7	2.102×10^{2}	3.168×10 ²	2.053×10^{2}	6.476×10^{2}	2.169×10 ²	1.012×10^{-11}	3.688×10 ²
f_8	0	1.888	5.034	1.946	5.766	6.957×10 ⁻¹	1.441
f_9	0	7.940×10 ⁻¹	5.516×10 ⁻¹	2.602×10^{-15}	9.000	1.917×10^{-14}	7.649×10^{-1}
f_{10}	0	5.098×10 ⁻²	6.514×10 ⁻²	5.066×10 ⁻³	1.382×10 ⁻¹	2.578×10 ⁻²	3.280×10 ⁻¹⁶
f_{11}	5.825×10^{-34}	7.800×10^{-2}	3.926×10 ⁻²	1.549×10^{-2}	4.723×10 ⁻²	1.187×10^{-33}	2.791×10 ⁻²
f_{12}	7.022×10^{-17}	1.088×10^{-16}	2.139×10 ⁻¹⁰	4.107	4.398×10 ⁻⁴	9.065×10 ⁻¹⁷	1.147×10^{-16}
f_{13}	3.048×10 ⁻¹⁹	1.870×10 ⁻¹⁹	2.497×10^{-4}	7.997×10 ⁻³	4.014×10 ⁻⁴	5.003×10 ⁻³	3.278×10 ⁻¹⁹
f_{14}	3.331×10 ⁻¹⁶	3.331×10^{-16}	2.714×10 ⁻⁵	4.969×10 ⁻⁵	8.580×10 ⁻⁴	3.331×10^{-16}	3.331×10^{-16}



Fig. 2 Average convergence curves of test functions

TCMA 均能快速收敛, 大概率收敛至 0, 而 IMA2 更为稳定, 其次为 MIMA。图 2(e) 中, MIMA 起步即快速收敛, 但被 GWO 追上, 不过 MIMA 寻优精度优于 GWO。在其他测试函数中, MIMA 的收敛曲线处于所有曲线左下角, 寻优精度远超其他算法几十个数量级, 表现出优异的收敛速度和寻优精度。这与表 3 中数据相互印证。

对于不定维多峰函数(见图 2(g)~图 2(k)),多数情况下,MIMA 收敛曲线明显远离其他算法,处于左下角,即各项指标皆远超其他算法。图 2(g)中,由于 SCA效果太差而舍弃,IMA2表现优异达到理论值,而 MIMA 收敛曲线整体处于 IMA 下方,寻优精度优于 IMA,仅次于 IMA2。图 2(h)中,MIMA 收敛曲线因陷入局部最优解而部分处于 GWO上方,但最终能跳出局部最优解收敛至理论最优解0,而 GWO 不能在迭代中跳出局部最优。图 2(k)中,MIMA 在迭代初期收敛速度最快,中期被 IMA2 赶上,但 IMA2 迭代中后期陷入停滞,最终 MIMA、IMA2、TCMA 三者收敛精度相近。

对于固定维函数(见图 2(1)~图 2(n)),5种算 法收敛曲线较为接近,但仍能看出 MIMA 的起步收 敛速度较快,一般处于所有曲线的左下方。

整体而言,所有测试函数平均收敛曲线均能明确显示 MIMA 收敛曲线的断层现象,表现出极强的局部最优逃逸能力;而另外 2 种基于 IMA 的改进算法由于变异手段较为简单,未能如 MIMA 一样多种变异手法互相补足,导致其陷入局部最优解后往往最多只能跳出三四次;其他经典算法的收敛曲线则基本上只能较为平滑地下降。此外, MIMA 的收敛曲线起点一般较低,即初始解质量较高,在图 2(b)、

图 2(d)、图 2(h)、图 2(i) 中表现尤为明显,这是加入 Sin 混沌映射初始化的结果;在图 2(j)中,使用 Tent 混沌映射初始化的 TCMA 与 MIMA 收敛曲线起点 接近,并且明显远离其他算法的起点,说明混沌映 射初始化能够较为有效地提高初始解质量。

综上所述, MIMA 在 14 个基准测试函数上的 收敛速度、寻优精度、全局搜索能力、局部最优逃 逸能力、稳定性等 5 个指标上,综合强于常用经典 优化算法和其他基于 IMA 的改进算法。由此可 见,本文算法具有更强的寻优能力。

3.3.3 Wilcoxon 秩和检验分析

仅根据算法运行 50次后的平均值、标准差等 指标对算法性能进行评价往往不够准确,为体现算 法改进的鲁棒性和公平性,利用统计检验来评估本 文算法相比于其他算法的优越性。因此,在 5%的 显著水平下进行 Wilcoxon 秩和检验。表 4 给出了 14 个测试函数的 MIMA 运行结果与其他算法的 Wilcoxon 秩和检验中计算的 p值。当p < 0.05时,即 可被认为是拒绝零假设的有力验证^[19]。由表 4 数 据可知,绝大部分的秩和检验 p值小于 0.05,说明 MIMA 的性能较之其他算法而言具有显著优异的 收敛性能。

3.3.4 消融实验分析

由于引入 4 项改进机制, MIMA 的寻优性能较为优异。为明确各改进机制的作用, 对 MIMA 进行 消融实验, 对不同策略进行实验测试。本文引入 2 个实验测试。基础的改进蜉蝣算法记为 IMA, 在 IMA 基础上引入 Sin 混沌映射初始化蜉蝣种群、 Tent 混沌映射和高斯变异对种群个体进行调节策 略的算法记为 CMA, 在 IMA 基础上引入自适应

表 4 Wilcoxon 秩和检验*p*值 Table 4 *p*-value for Wilcoxon's rank-sum results

		-				
测试函数	IMA	PSO	GWO	SCA	IMA2	TCMA
f_1	3.019 9×10 ⁻¹¹	3.019 9×10 ⁻¹¹	3.019 9×10 ⁻¹¹	3.019 9×10 ⁻¹¹	3.019 9×10 ⁻¹¹	3.019 9×10 ⁻¹¹
f_2	3.0199×10^{-11}	3.0199×10^{-11}	3.0199×10^{-11}	3.019 9×10 ⁻¹¹	3.0199×10^{-11}	$3.0199{\times}10^{-11}$
f_3	3.0199×10^{-11}	3.0199×10^{-11}	3.0199×10^{-11}	3.019 9×10 ⁻¹¹	3.0199×10^{-11}	$3.0199{\times}10^{-11}$
f_4	7.457 7×10 ⁻⁹	2.363 8×10 ⁻¹²	$2.363 8 \times 10^{-12}$	2.363 8×10 ⁻¹²	$1.6074{\times}10^{-1}$	3.0199×10^{-11}
f_5	3.0199×10^{-11}	3.0199×10^{-11}	$7.389.1 \times 10^{-11}$	3.019 9×10 ⁻¹¹	3.0199×10^{-11}	3.0199×10^{-11}
f_6	$6.478 \ 9 \times 10^{-12}$	$6.478 9 \times 10^{-12}$	$6.478 \ 9 \times 10^{-12}$	$6.478 \ 9 \times 10^{-12}$	$6.4789{ imes}10^{-12}$	$6.478 \ 9 \times 10^{-12}$
f_7	$2.530.6 \times 10^{-4}$	3.0199×10^{-11}	9.063 2×10 ⁻⁸	3.019 9×10 ⁻¹¹	$2.4306{\times}10^{-11}$	$2.194.7 \times 10^{-8}$
f_8	$2.530.6 \times 10^{-4}$	3.0199×10^{-11}	9.063 2×10 ⁻⁸	3.019 9×10 ⁻¹¹	$4.5270{ imes}10^{-12}$	$1.211 8 \times 10^{-12}$
f_9	$1.210 8 \times 10^{-12}$	1.211 8×10 ⁻¹²	$2.034 \ 1 \times 10^{-13}$	1.211 8×10 ⁻¹²	$1.0609{\times}10^{-12}$	$1.211 8 \times 10^{-12}$
f_{10}	$5.772.0 \times 10^{-11}$	1.211 8×10 ⁻¹²	8.152 3×10 ⁻²	1.211 8×10 ⁻¹²	1.6560×10^{-11}	2.911 2×10 ⁻²
f_{11}	1.152 9×10 ⁻⁹	$2.972.9 \times 10^{-11}$	$2.972.9 \times 10^{-11}$	2.972 9×10 ⁻¹¹	3.1958×10^{-7}	1.4735×10^{-7}
f_{12}	$3.049.0 \times 10^{-1}$	$3.157 8 \times 10^{-12}$	$1.988 \ 1 \times 10^{-11}$	$4.390.9 \times 10^{-12}$	$7.805~6 \times 10^{-1}$	6.324 1×10 ⁻¹
f_{13}	NaN	1.211 8×10 ⁻¹²	$1.211 8 \times 10^{-12}$	1.211 8×10 ⁻¹²	NaN	8.1404×10^{-2}
f_{14}	NaN	1.211 8×10 ⁻¹²	1.211 8×10 ⁻¹²	1.211 8×10 ⁻¹²	NaN	NaN

注:NaN表示算法性能相当。

2024 年

动态调节的重力系数和 ROBL 策略的算法记为 GMA。分别以 IMA、CMA、GMA、MIMA 在测试

函数上作对比实验,结果如表5所示。

从表5中可以看出,在不定维单峰函数fi~f6

表 5 消融实验结果

Table 5	Experimental	results	of	ablation
---------	--------------	---------	----	----------

测计远数		平均	值	
侧瓜函数	IMA	СМА	GMA	MIMA
f_1	6.023×10 ⁻³⁹	4.459×10^{-160}	8.516×10 ⁻¹⁵⁶	1.486×10^{-192}
f_2	2.106×10^{-16}	1.159×10^{-94}	1.604×10^{-82}	7.613×10 ⁻⁹⁸
f_3	2.106×10^{-3}	2.058×10^{-6}	2.844×10 ⁻⁵¹	9.818×10^{-64}
f_4	1.972×10^{-33}	7.396×10^{-34}	1.761×10^{-31}	4.930×10 ⁻³⁴
f_5	5.163×10 ⁻³	2.098×10^{-3}	1.379×10^{-4}	1.254×10^{-4}
f_6	7.675×10 ⁻⁷⁴	6.998×10 ⁻²²²	7.204×10^{-217}	1.864×10^{-294}
f_7	-6.881×10^{3}	-7.233×10 ³	-7.105×10^{3}	-7.445×10^{3}
f_8	2.815	1.678	0	0
f_9	1.182	4.370×10 ⁻¹⁵	8.882×10^{-16}	8.882×10^{-16}
f_{10}	3.780×10^{-2}	0	0	0
f_{11}	3.421×10^{-2}	1.662×10^{-2}	1.960×10^{-31}	2.382×10^{-32}
f_{12}	9.980×10^{-1}	9.980×10 ⁻¹	9.980×10 ⁻¹	9.980×10 ⁻¹
<i>f</i> 13	3.075×10^{-4}	1.310×10 ⁻³	3.075×10^{-4}	3.075×10^{-4}
f_{14}	3.979×10^{-1}	3.979×10 ⁻¹	3.979×10^{-1}	3.979×10 ⁻¹
		最优	值	
侧瓜图数	IMA	СМА	GMA	MIMA
f_1	8.786×10 ⁻⁴⁹	4.162×10^{-173}	4.062×10^{-196}	5.722×10 ⁻¹⁹³
f_2	1.962×10^{-26}	3.206×10 ⁻⁹⁵	4.226×10 ⁻⁹⁸	1.450×10 ⁻⁹⁹
f_3	5.682×10 ⁻⁵	1.443×10^{-24}	9.776×10 ⁻⁹⁸	1.010×10^{-97}
f_4	0	0	0	0
f_5	1.762×10^{-3}	3.955×10 ⁻⁴	5.288×10^{-6}	1.227×10^{-7}
f_6	3.927×10 ⁻⁹³	1.580×10^{-232}	7.363×10^{-229}	0
f7	-7.787×10^{3}	-7.906×10^{3}	-7.906×10^{3}	-7.926×10 ³
f_8	5.684×10^{-14}	0	0	0
f_9	2.931×10^{-14}	8.882×10^{-16}	8.882×10^{-16}	8.882×10^{-16}
f_{10}	0	0	0	0
f_{11}	2.356×10^{-32}	2.356×10 ⁻³²	2.404×10^{-32}	2.356×10 ⁻³²
f_{12}	9.980×10^{-1}	9.980×10^{-1}	9.980×10^{-1}	9.980×10 ⁻¹
f_{13}	3.075×10^{-4}	3.075×10^{-4}	3.075×10^{-4}	3.075×10^{-4}
f_{14}	3.979×10^{-1}	3.979×10 ⁻¹	3.979×10^{-1}	3.979×10 ⁻¹
测试或类		标准	差	
仍此四致	IMA	СМА	GMA	MIMA
f_1	1.887×10^{-37}	2.425×10 ⁻¹⁵⁹	5.945×10 ⁻¹⁵⁵	0
f_2	1.043×10^{-14}	1.268×10^{-94}	1.123×10 ⁻⁸¹	1.161×10 ⁻⁹⁷
f_3	6.188×10^{-3}	9.700×10^{-6}	1.240×10^{-50}	4.279×10 ⁻⁶³
f_4	7.542×10^{-33}	2.927×10 ⁻³³	5.428×10 ⁻³¹	2.415×10 ⁻³³
f_5	2.830×10 ⁻³	1.350×10 ⁻³	1.045×10 ⁻⁴	1.318×10^{-4}
f_6	4.000×10^{-73}	0	0	0
f_7	3.168×10 ²	2.789×10^{2}	4.038×10^{2}	2.102×10^{2}
f_8	1.888	2.873	0	0
f_9	7.940×10^{-1}	4.974×10^{-16}	0	0
f_{10}	5.098×10^{-2}	0	0	0
f_{11}	7.800×10^{-2}	5.627×10 ⁻²	5.236×10 ⁻³¹	5.825×10^{-34}
f_{12}	1.088×10^{-16}	5.439×10 ⁻¹⁷	2.259×10^{-10}	7.022×10^{-17}
f_{13}	1.870×10^{-19}	4.371×10 ⁻³	2.341×10 ⁻¹⁹	3.048×10 ⁻¹⁹
f_{14}	3.331×10 ⁻¹⁶	3.331×10 ⁻¹⁶	0	3.331×10 ⁻¹⁶

中,除 f_4 、 f_5 外,CMA、GMA、MIMA 结果均显著优 于 IMA,CMA、GMA 均使寻优性能有一定提升;在 f_4 中,GMA 略逊于 IMA,但CMA、GMA 结合后,MIMA 的平均值最好,稳定性最高;在 f_5 中,CMA、GMA 平 均值、均方差略优于 IMA,因此,MIMA 相对于 IMA 的总体提升幅度不大。

在不定维多峰函数 f₇~f₁₁中, CMA、GMA 对 MIMA 的寻优能力提升显著, 仅 f₇因为理论最优解 约为-8 379.66 而使得显示结果不够明显; 在 f₈、f₁₁ 中, CMA 不能稳定找到理论最优解 0, 对于平均值 提升幅度不大, 但 GMA 可以稳定收敛至理论最优 解 0, 对 MIMA 影响更大。

在固定维函数 f₁₂~f₁₄中,由于每种算法都无限 接近理论最优解, MIMA 的性能提升幅度不明显, 但 CMA、GMA、MIMA 的结果在 16 位小数显示状 态下可以看到均不逊于 IMA。

综上所述,通过模型消融实验证明本文所提出 的每一种策略都对 MIMA 的寻优性能有所帮助。

4 结 论

本文综合4种改进策略,基于改进蜉蝣算法, 提出了一种多策略融合改进的自适应蜉蝣算法。

1)本文算法在14种经典基准测试函数上,相 较于 IMA2、TCMA 2种基于 IMA 的改进算法和 GWO、SCA、PSO 等经典优化算法,有更加出色的 寻优能力和稳定性。

2)通过 Wilcoxon 秩和检验分析,验证了本文 算法相比于其他算法的优越性,通过消融实验验证 了所采用改进策略的有效性。

综上,本文算法相比其他算法,在寻优精度、 收敛速度、稳定性上都取得了较大提升,但其使用 了诸多映射方法,若将算法应用于复杂约束问题, 会存在困难。因此,对于多目标复杂问题的求解, 并将其应用复杂工程设计问题中,是未来的研究 重点。

参考文献(References)

- [1] RICCI F, ROKACH L, SHAPIRA B. Recommender systems handbook[M]. Berlin: Springer, 2011.
- [2] XUE J K, SHEN B. A novel swarm intelligence optimization approach: Sparrow search algorithm[J]. Systems Science & Control Engineering, 2020, 8(1): 22-34.
- [3] PANDA M, DAS B, PATI B B. Grey wolf optimization for global path planning of autonomous underwater vehicle[C]//Proceedings of the 3rd International Conference on Advanced Informatics for Computing Research. New York: ACM, 2019: 1-6.
- [4] YANG X S. Firefly algorithms for multimodal optimization[C]// Proceedings of the International Symposium on Stochastic Algorithms. Berlin: Springer, 2009: 1-10.

- [5] ZERVOUDAKIS K, TSAFARAKIS S. A mayfly optimization algorithm[J]. Computers & Industrial Engineering, 2020, 145: 106559.
- [6] 张达敏, 徐航, 王依柔, 等. 嵌入 Circle 映射和逐维小孔成像反向 学习的鲸鱼优化算法[J]. 控制与决策, 2021, 36(5): 1173-1180.
 ZHANG D M, XU H, WANG Y R, et al. Whale optimization algorithm for embedded Circle mapping and one-dimensional oppositional learning based small hole imaging[J]. Control and Decision, 2021, 36(5): 1173-1180(in Chinese).
- [7] 胡健,刘祥敏,毛伊敏,等. 基于 KD 树和混沌蜉蝣优化的并行谱 聚类算法[J]. 计算机集成制造系统, 2023, 29(12): 4001-4020.
 HU J, LIU X M, MAO Y M, et al. Parallel spectral clustering algorithm using KD tree and chaotic mayfly optimization algorithm[J]. Computer Integrated Manufacturing Systems, 2023, 29(12): 4001-4020(in Chinese).
- [8] 唐延强,李成海,宋亚飞,等. 自适应变异麻雀搜索优化算法[J]. 北京航空航天大学学报, 2023, 49(3):681-692. TANG Y Q, LI C H, SONG Y F, et al. Adaptive mutation sparrow search optimization algorithm[J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49(3):681-692(in Chinese).
- [9] 王义,张达敏,邹诚诚. 增强全局搜索和自适应蜉蝣算法[J]. 哈尔 滨工业大学学报, 2022, 54(11): 137-150.
 WANG Y, ZHANG D M, ZOU C C. Enhance global search and adaptive mayfly algorithm[J]. Journal of Harbin Institute of Technology, 2022, 54(11): 137-150(in Chinese).
- [10] 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, 2021, 47(8): 1712-1720.
 LU 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).
- [11] 张娜, 赵泽丹, 包晓安, 等. 基于改进的 Tent 混沌万有引力搜索算法[J]. 控制与决策, 2020, 35(4): 893-900.
 ZHANG N, ZHAO Z D, BAO X A, et al. Gravitational search algorithm based on improved Tent chaos[J]. Control and Decision, 2020, 35(4): 893-900(in Chinese).
- [12] WU K L, YANG M S. Alternative c-means clustering algorithms[J]. Pattern Recognition, 2002, 35(10): 2267-2278.
- [13] SHI Y. A modified particle swarm optimizer[C]//Proceedings of the IEEE International Conference on Evolutionary Computation. Piscataway: IEEE Press, 1998: 69-73.
- [14] 王梦娜, 王秋萍, 王晓峰. 基于 Iterative 映射和单纯形法的改进灰 狼优化算法[J]. 计算机应用, 2018, 38(S2): 16-20.
 WANG M N, WANG Q P, WANG X F. Improved grey wolf optimization algorithm based on Iterative mapping and simplex method[J]. Journal of Computer Applications, 2018, 38(S2): 16-20 (in Chinese).
- [15] TIZHOOSH H R. Opposition-based learning: A new scheme for machine intelligence[C]//Proceedings of the International Conference on International Conference on Computational Intelligence for Modelling, Control & Automation. Piscataway: IEEE Press, 2005: 695-701.
- [16] LONG W, JIAO J, LIANG X, et al. A random opposition-based learning grey wolf optimizer[J]. IEEE Access, 2019, 7: 113810-113825.
- [17] 邵瑞凝,杨博,束洪春,等.基于改进蜉蝣算法的光伏阵列最优重 构方法[J].电力系统自动化,2022,14(11):142-150.

SHAO R N, YANG B, SHU H C, et al. Improved mayfly algorithm for optimal PV array reconfiguration[J]. Automation of Electric Power Systems, 2022, 14(11): 142-150(in Chinese).

[18] 徐焕增,徐文倩,孔政敏. 基于 Tent 混沌序列改进的蜉蝣算法及
 其应用[J]. 控制工程, 2022, 29(3): 435-440.
 XU H Z, XU W O, KONG Z M. Mayfly algorithm based on Tent

chaotic sequence and its application[J]. Control Engineering of China, 2022, 29(3): 435-440(in Chinese).

[19] DERRAC J, GARCIA S, MOLINA D, et al. A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms[J]. Swarm & Evolutionary Computation, 2011, 1(1): 3-18.

Multi-strategy fusion improved adaptive mayfly algorithm

JIANG Yufei, XU Xianze^{*}, XU Fengqiu, GAO Bo

(School of Electronic Information, Wuhan University, Wuhan 430072, China)

Abstract: This paper proposes the multi-strategy fusion improved adaptive mayfly algorithm (MIMA), which addresses the shortcomings of the improved mayfly algorithm, including its low adaptive ability, minimal population diversity, and poor global search performance. Firstly, Sin chaos mapping was used to initialize the mayfly population so that the population could be uniformly distributed in the solution space, which improved the initial population quality and enhanced the global search ability. Second, in order to improve the local optimal escape ability, control population density, and boost population diversity, individuals in the population were exposed to Gaussian variation and Tent chaos mapping. Then, the incomplete gamma function was introduced to reconstruct the adaptive dynamic adjustment of gravity coefficients to establish a better balance between global search and local exploitation ability, which in turn improved the convergence accuracy of the algorithm and facilitated the potential of global search to find the optimal solution. Finally, the random opposition-based learning (ROBL) strategy was adopted to enhance the global search ability improve the convergence speed and enhance the stability. To demonstrate the efficacy and dependability of the four improvement measures, the algorithms were compared using the classical test function set and their optimization effect was examined using the Wilcoxon rank sum test. The experimental results show that compared with other algorithms, the MIMA has better searching accuracy, convergence speed, and stability.

Keywords: mayfly algorithm; chaotic mapping; Gaussian mutation; adaptive dynamic adjustment; random opposition-based learning

Received: 2022-06-16; Accepted: 2022-09-23; Published Online: 2022-10-10 10:36 URL: link.cnki.net/urlid/11.2625.V.20221009.1050.010

Foundation item: National Natural Science Foundation of China (51975422)

^{*} Corresponding author. E-mail: xxz@whu.edu.cn

《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	E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4期 2024年4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 BELJING UNIVERSITY OF **AERONAUTICS AND ASTRONAUTICS** (JBUAA)

(Monthly, Started in 1956) Vol.50 No.4 April 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 (010) 82316698 82316981 Telephone 82315499 82317448 E-mail jbuaa@buaa.edu.cn http://bhxb.buaa.edu.cn

中国标准连续出版物号: ISSN 1001-5965 CN 11-2625/V

国内定价: 50.00元/期

